К. Ю. Федоровский

Алгебра. Линейная алгебра

Текст лекций для факультета ФН

Москва — 2012

Оглавление

РАЗДЕЛ 1

Предварительные сведения

Мы начинаем изучение *линейной алгебры*. В самом широком смысле можно считать, что основная задача и основное содержание линейной алгебры как математической дисциплины состоит в разработке специального математического языка для выражения одной из наиболее общих и фундаментальных естественнонаучных идей – идеи *линейности*. Одним из наиболее важных частных случаев этой идеи является т.н. *принцип малых приращений*, согласно которому почти всякий естественный процесс почти всюду в малом линеен. На этом принципе основаны, например, многие разделы современного математического анализа и их приложения. Аппарат линейной алгебры оказался адекватным и удобным инструментом для формулировки многих фундаментальных физических законов.

С другой стороны, с позиции математика-пуриста, линейная алгебра – это раздел теории алгебраических структур, в котором рассматриваются структуры специального вида – линейные пространства и их отображения.

В предлагаемом курсе лекций учтено, что теория матриц и определителей, а также основные элементы теории систем линейных уравнений изучаются студентами в рамках курса *аналитической геометрии* в первом семестре. Соответственно, основные понятия этого курса считаются известными.

Многие простые утверждения о свойствах изучаемых объектов приводятся в курсе в качестве ***упражнений***. Читателю рекомендуется самостоятельно выполнять такие упражнения при чтении соответствующего материала. Упражнения, выполнение которых может вызвать затруднения, могут быть разобраны на семинарских занятиях по курсу.

1.1. Множества и отображения

Напомним ряд обозначений, которые будут использоваться в дальнейшем (в обоих частях курса на протяжении обоих семестров). Через $\mathfrak{S}(X)$ обозначим совокупность всех подмножеств некоторого множества X, при этом $\varnothing \in \mathfrak{S}(X)$ (где, как обычно, $\varnothing$ обозначает пустое множество) и $X \in \mathfrak{S}(X)$. Система подмножеств $\mathcal{S} \subset \mathfrak{S}(X)$ называется *нетривиальной*, если $\mathcal{S} \neq \{\varnothing, X\}$. Всюду в дальнейшем, если обратное не оговорено специально, запись $X \subset Y$ *не исключает* равенства $X = Y$. Таким образом, множества X и Y равны если и только если $X \subset Y$ и $Y \subset X$.

Пусть X и Y — некоторые множества.

Определение. *Если задан некоторый закон (правило), по которому для каждого элемента $x \in X$ ставится в соответствие некоторый элемент $f(x) \in Y$, то говорят, что задано отображение $f : X \to Y$. Множество X называется областью определения отображения f, а множество Y — его областью значений.*

Выражение $f(x)$ обозначается также символом fx. Сопоставление элементу $x \in X$ значения $f(x) \in Y$ обозначается также символом $x \mapsto f(x)$.

Из определения отображения непосредственно вытекает, что два отображения $f : X \to Y$ и $g : X \to Y$ *равны*, если $f(x) = g(x)$ для любого элемента $x \in X$.

Напомним несколько понятий, связанных с отображением $f : X \to Y$.

Определение. *Множество $\mathrm{Im}(f) = \{f(x) : \ x \in X\}$ называется образом f. Если $\mathrm{Im}(f) = Y$, то отображение f называется сюръективным (или "отображением на").*

Для любого элемента $y \in Y$ определяется множество $f^{-1}(y) = \{x \in X : \ f(x) = y\}$, которое называется *прообразом* элемента $y \in Y$. По определению прообраза, $f^{-1}(y) = \varnothing$ при $y \in Y \setminus \mathrm{Im}(f)$. Для любого подмножества $Y_0 \subset Y$ определяется его прообраз относительно отображения f: $f^{-1}(Y_0) = \{x \in X : \ f(x) \in Y_0\} = \bigcup_{y \in Y_0} f^{-1}(y)$.

Определение. *Отображение f называется инъективным, если из неравенства $x_1 \neq x_2$ вытекает, что $f(x_1) \neq f(x_2)$. Всякое инъективное и, одновременно, сюръективное отображение $f : X \to Y$ называется биективным.*

Легко привести пример отображения, не являющегося ни сюръективным ни инъективным. В самом деле, пусть $(\phi_n : \ n \in \mathbb{N})$ – последовательность чисел Фибоначчи, т.е. $\phi_1 = \phi_2 = 1$, а $\phi_n = \phi_{n-1} + \phi_{n-2}$ при $n \geqslant 3$. Нетрудно проверить, что отображения $f : \mathbb{N} \mapsto \mathbb{N}$ множества натуральных чисел в себя, определенное по правилу $f(n) = \phi_n$, обладает указанным свойством.

Рассмотрим теперь отображение, определенно правилом $x \mapsto x^2$. Это отображение, рассматриваемое как отображение $\mathbb{R} \to \mathbb{R}$, не является ни сюръективным, ни инъективным. Если отображение $x \mapsto x^2$, рассматривается как отображение $\mathbb{R} \to \mathbb{R}_+$ (где $\mathbb{R}_: = \{t \in \mathbb{R} : \ t \geqslant 0\}$), то оно является сюръективным, но не инъективным. И, наконец, отображение $x \mapsto x^2$, рассматривается как отображение $\mathbb{R}_+ \to \mathbb{R}_+$, является взаимно однозначным (биективным).

Определение. *Отображение* $\mathrm{id}_X : X \to X$*, определенное по правилу $x \mapsto x$, называется единичным (или тождественным) отображением.*

Другими словами, единичное отображение переводит каждый элемент своей области определения в себя. Часто тождественное отображение обозначается символом e_X.

Пусть задано отображение $f : X \to Y$. Тогда для любого подмножества $X_0 \subset X$ можно определить отображение $f_0 : X_0 \to Y$ по правилу $f_0(x) = f(x)$ при $x \in X_0$. Отображение f_0 называется *сужением* (или *ограничением*) отображения f на X_0 и обозначается символом $f_0 = f_{|X_0}$.

Если же заданы множества X_1 и Y_1 такие, что $X \subset X_1$ и $Y \subset Y_1$ и отображение $f_1 : X_1 \to Y_1$ такое, что $f_1(x) = f(x)$ для любого $x \in X$, то отображение f_1 называется *продолжением* отображения f с X на X_1. Понятия сужения и продолжения отображений широко используются во всех разделах математики.

Определение. *Если определены два отображения $g : X \to Y$ и $f : Y \to Z$, то отображение $f \circ g : X \to Z$, определенное соотношением*

$$(f \circ g)(x) = f(g(x)), \quad x \in X,$$

называется композицией отображений f и g.

Часто композицию отображений называют *произведением* отображений и, если это не вызывает путаницы, обозначают символом fg. Заметим, что операция композиции определена не для любых пар отображений. Для того, чтобы оно была определена, необходимо, чтобы область определения отображения f совпадала с областью значений отображения g.

Для любого отображения $f : X \to Y$ верны равенства $f \circ \mathrm{id}_X = f$ и $\mathrm{id}_Y \circ f = f$.

Предложение 1.1. *Операция композиции отображений ассоциативна, т.е. для любых отображений $h : X \to Y$, $g : Y \to Z$ и $f : Z \to W$ выполняется равенство*

$$f \circ (g \circ h) = (f \circ g) \circ h.$$

Доказательство. Рассмотрим два отображения $q_1 : X \to W$ и $q_2 : X \to W$ определенные, соответственно, равенствами $q_1 = f \circ (g \circ h)$ и $q_2 = (f \circ g) \circ h$. Как нетрудно проверить, оба отображения q_1 и q_2 определены при всех $x \in X$ и имеют свой областью значений множество W. Так как

$$q_1(x) = (f \circ (g \circ h))(x) = f((g \circ h)(x)) = f(g(h(x))) = (f \circ g)(h(x)) = ((f \circ g) \circ h)(x) = q_2(x)$$

для любого $x \in X$, то отображения $f \circ (g \circ h)$ и $(f \circ g) \circ h$ равны. $\square$

Нетрудно заметить, что операция композиции отображений не является коммутативной: для отображений $f, g : X \to X$, в общем случае, $f \circ g \neq g \circ f$. В самом деле, это равенство нарушается, например, для отображений $f, g : \{a, b\} \to \{a, b\}$ определенных равенствами $f(a) = b$, $f(b) = a$, $g(a) = g(b) = a$.

Определение. *Если $f : X \to Y$ и $g : Y \to X$ такие отображения, что*

$$f \circ g = \mathrm{id}_Y \quad \text{и} \quad g \circ f = \mathrm{id}_X, \tag{1.1}$$

то g называется обратным отображением для f и обозначается символом f^{-1}.

Непосредственно из определения обратного отображения вытекает, что если g является обратным отображением для f, то f является обратным отображением для g.

Предложение. *Если отображение $f : X \to Y$ имеет обратное отображение, то это обратное отображение единственно.*

Доказательство. Предположим, что существуют два отображения $g_{1,2} : Y \to X$ со свойствами $f \circ g_1 = \mathrm{id}_Y$, $g_1 \circ f = \mathrm{id}_X$ и $f \circ g_2 = \mathrm{id}_Y$, $g_2 \circ f = \mathrm{id}_X$. Тогда, в силу ассоциативности операции композиции отображений:

$$g_2 = \mathrm{id}_X \circ g_2 = (g_1 \circ f) \circ g_2 = g_1 \circ (f \circ g_2) = g_1 \circ \mathrm{id}_Y = g_1. \quad \square$$

Далее, имеет место следующее утверждение:

***Лемма* 1.2.** *Если $f : X \to Y$ и $g : Y \to X$ – такие отображения, что $g \circ f = \mathrm{id}_X$, то отображение f инъективно, а отображение g сюръективно.*

Доказательство. Начнем с проверки инъективности f. Пусть $x_1, x_2 \in X$ таковы, что $f(x_1) = f(x_2)$. Тогда

$$x_1 = \mathrm{id}_X(x_1) = (g \circ f)(x_1) = g(f(x_1)) = g(f(x_2)) = (g \circ f)(x_2) = \mathrm{id}_X(x_2) = x_2.$$

Далее, если x – произвольный элемент из X, то $x = \mathrm{id}_X(x) = g(f(x))$, т.е. для любого $x \in X$ существует $y \in Y$ такой, что $x = g(y)$ (это элемент $y = f(x)$). Т.е. f – инъективно, а g – сюръективно. $\square$

Пример. Только что доказанная лемма может быть проиллюстрирована следующим простым примером. Пусть $X = \mathbb{R}_+$, $Y = \mathbb{R}$, отображения $f : X \to Y$ и $g : Y \to X$ определены так, что $f : x \mapsto x^2$, а

$$g : y \mapsto \begin{cases} \sqrt{y}, & y \geqslant 0, \\ 0, & y < 0. \end{cases}$$

Видно, что $g \circ f$ — это отображение $\mathbb{R}_+ \to \mathbb{R}_+$, действующее по правилу $x \mapsto \sqrt{x^2} = x$, т.е. $g \circ f = \mathrm{id}_X$ при $X = \mathbb{R}_+$.

Теперь мы готовы сформулировать и доказать критерий обратимости отображений.

Теорема 1.3.

1. *Отображение $f : X \to Y$ обратимо если и только если оно биективно.*

2. *Если f биективно, то f^{-1} биективно и $(f^{-1})^{-1} = f$.*

3. *Если отображения $f : X \to Y$ и $h : Y \to Z$ биективны, то отображение $h \circ f$ биективно и $(h \circ f)^{-1} = f^{-1} \circ h^{-1}$.*

Доказательство. Докажем первое утверждение теоремы. Предположим, что отображение f обратимо. Пусть $g = f^{-1}$ – обратное отображение. Тогда, из Леммы 1.2 и соотношений (1.1) непосредственно вытекает, что f инъективно и сюръективно (т.е., биективно). Обратно, пусть отображение f биективно. Следовательно, для любого $x \in X$ существует единственное $y \in Y$ такое, что $y = f(x)$. Построим отображение $g : Y \to X$

так, что $g(y) = x$. Проверка того, что отображение g удовлетворяет соотношениями (1.1) очевидна и оставляется в качестве *упражнения*.

Докажем второе утверждение теоремы. Так как отображение f биективно, то существует обратное отображение f^{-1}, для которого $f^{-1} \circ f = \mathrm{id}_X$ и $f \circ f^{-1} = \mathrm{id}_Y$. Из этих соотношений и Леммы 1.2 вытекает биективность f^{-1} и равенство $(f^{-1})^{-1} = f$.

Третье утверждение теоремы доказывается аналогично. В самом деле, из биективности отображений f и h вытекает биективность их композиции $h \circ f$. Т.е. отображение $h \circ f$ обратимо. Из равенств

$$(h \circ f) \circ (f^{-1} \circ h^{-1}) = ((h \circ f) \circ f^{-1}) \circ h^{-1} = (h \circ (f \circ f^{-1}) \circ h^{-1} = h \circ h^{-1} = \mathrm{id}_Z,$$
$$(f^{-1} \circ h^{-1}) \circ (h \circ f) = f^{-1} \circ (h^{-1} \circ (h \circ f)) = f^{-1} \circ ((h^{-1} \circ h) \circ f) = f^{-1} \circ f = \mathrm{id}_X,$$

вытекает, что $(h \circ f)^{-1} = f^{-1} \circ h^{-1}$. □

Рассмотрим еще один интересный и поучительный пример. Пусть отображение $f : \mathbb{N} \to \mathbb{N}$ определено по правилу $x \mapsto x + 1$. Тогда ясно, что это отображение инъективно, но не сюръективно (элемент $x = 1 \in \mathbb{N}$ не имеет прообраза). Оказывается, что такой феномен возможен только в случае бесконечных множеств. А в случае конечных множеств справедливо следующее утверждение.

Предложение 1.4. *Если X – конечное множество, а отображение $f : X \to X$ инъективно, то оно биективно.*

Доказательство. Нам необходимо доказать сюръективность f. Для любого целого $k \geqslant 0$ определим отображение f^k по индукции: $f^0 = \mathrm{id}_X$, а $f^k = f^{k-1} \circ f$ при $k \geqslant 1$. Возьмем произвольное $x \in X$ и рассмотрим последовательность $f^0(x), f^1(x), \dots, f^n(x), \dots$. Так как X – конечное множество, то в этой последовательности должны быть повторения. Предположим, что $f^m(x) = f^n(x)$ при $n > m$. Если $m = 0$, то $f^n(x) = f^0(x) = x$. Если $m > 0$, то, в силу инъективности f имеет место равенство $f^{n-1}(x) = f^{m-1}(x)$. После последовательного выполнения “сокращений” мы получим элемент $x_1 = f^{n-m-1}(x)$, для которого $x = f(x_1)$. Т.е., f сюръективно. □

1.2. Отношения эквивалентности, факторизация множеств и отображений

Напомним, что *бинарным отношением* между двумя множествами X и Y называется любое подмножество $\Omega \subset X \times Y$, где

$$X \times Y = \{(x, y) : \ x \in X, \ y \in Y\}.$$

Бинарное отношение между X и X называется *бинарным отношением на X*.

Пусть на множестве X задано некоторое бинарное отношение Ω. Тот факт, что (упорядоченная) пара $(x, y) \in X \times X$ принадлежит отношению Ω будем обозначать $x \sim_\Omega y$. Для простоты обозначений, говоря о некотором произвольном бинарном отношении на X будет использовать обозначение $x \sim y$.

Определение. *Бинарное отношение $\sim$ на множестве X называется отношением эквивалентности, если для всех $x, x_1, x_2 \in X$ выполнены соотношения*

(1) *$x \sim x$ (рефлексивность);*
(2) *из $x_1 \sim x_2$ следует $x_2 \sim x_1$ (симметричность);*
(3) *из $x \sim x_1$ и $x_1 \sim x_2$ следует $x \sim x_2$ (транзитивность).*

Подмножество $\widetilde{x} = \{y \in X : \ y \sim x\} \subset X$ называется классом эквивалентности, содержащим элемент x.

Проверим корректность определения класса эквивалентности. Для этого необходимо установить, что, во-первых, $x \in \widetilde{x}$ и, во вторых, что класс $\widetilde{x}$ однозначно определяется любым своим представителем, т.е. $\widetilde{x}_1 = \widetilde{x}_2$ если и только если $x_1 \sim x_2$. В самом деле, так как $x \sim x$, то $x \in \widetilde{x}$. Далее, если $x_1 \sim x_2$ и $y \in \widetilde{x}_1$, то $y \sim x_1 \sim x_2$ и, следовательно, $y \in \widetilde{x}_2$. Таким образом, $\widetilde{x}_1 \subset \widetilde{x}_2$. Аналогично, $\widetilde{x}_2 \subset \widetilde{x}_1$ и, окончательно, $\widetilde{x}_1 = \widetilde{x}_2$. Обратно, если $x_1 \in \widetilde{x}_1 = \widetilde{x}_2$, то $x_1 \sim x_2$ по определению класса эквивалентности. Оказывается, что имеет место следующий общий факт.

Предложение 1.5. *Если $\sim$ — отношение эквивалентности на множестве X, то*

1. *Для любых $x_1 \in X$ и $x_2 \in X$ либо $\widetilde{x}_1 = \widetilde{x}_2$, либо $\widetilde{x}_1 \cap \widetilde{x}_2 = \varnothing$.*
2. *Найдется такое подмножество $X_\sim \subset X$, что $X = \bigsqcup_{x \in X_\sim} \widetilde{x}$.*

Доказательство. Для доказательства утверждения 1 предположим, что $\widetilde{x}_1 \cap \widetilde{x}_2 \neq \varnothing$. Тогда найдется $y \in \widetilde{x}_1 \cap \widetilde{x}_2$. Так как $y \sim x_1$ и $y \sim x_2$, то $x_1 \sim x_2$ и, следовательно, $\widetilde{x}_1 = \widetilde{x}_2$. Далее, из условия $x \in \widetilde{x}$ вытекает, что $X = \bigcup_{x \in X} \widetilde{x}$. Определим множество $X_\sim$ взяв по одному представителю из всех различных классов $\widetilde{x}$, $x \in X$. Ясно, что в этом случае $X = \bigcup_{x \in X} \widetilde{x} = \bigsqcup_{x \in X_\sim} \widetilde{x}$. □

Другими словами, множество классов эквивалентности по отношению $\sim$ является разбиением множества X в объединение непересекающихся подмножеств, каждое из которых является классом эквивалентности относительно $\sim$.

Приведем примеры отношений эквивалентности. Пусть $X = \mathbb{R}^2$ – плоскость с декартовой системой координат. Пусть P_1, P_2 — точки плоскости. Определим отношение $\sim_1$ следующим образом: $P_1 \sim_1 P_2$ если P_1 и P_2 лежат на одной вертикальной прямой. Ясно, что $\sim_1$ — это отношение эквивалентности. Плоскость $\mathbb{R}^2$ распадается на непересекающиеся классы эквивалентности — на множество вертикальных прямых.

Другой пример можно получить рассмотрев множество $X = \{(x, y) \in \mathbb{R}^2 : \ x > 0, \ y > 0\}$ и отношение $\sim_2$ определяемое таким образом, что $P_1 \sim_2 P_2$ если P_1 и P_2 лежат на одной гиперболе $\{(x, y) \in \mathbb{R}^2 : \ x > 0, \ y > 0, \ xy = \delta\}$, $\delta > 0$.

Предложение 1.6. *Если задано некоторое разбиение множества X на непересекающиеся подмножества $X = \bigsqcup_{\alpha \in A} C_\alpha$ (здесь A — некоторое множество индексов), то существует такое отношение эквивалентности на X, что множества C_α, $\alpha \in A$ будут классами эквивалентности по этому отношению.*

Доказательство. По условию, для любого $x \in X$ существует единственный индекс $\alpha(x) \in A$ такой, что $x \in C_{\alpha(x)}$. Определим отношение $\sim$ считая, что $y \sim x$ если и только если $y \in C_{\alpha(x)}$. При этом $\widetilde{x} \subset C_{\alpha(x)}$. Из утверждения 1 Предложения 1.5 и из того, что $C_\beta \cap C_\alpha = \varnothing$ при $\alpha \neq \beta$ вытекает, что $C_{\alpha(x)} \subset \widetilde{x}$. Следовательно, разбиение $X = \bigsqcup_{\alpha \in A} C_\alpha$ совпадает с разбиением X на классы эквивалентности по отношению $\sim$. □

Определение. *Разбиение множества X на классы эквивалентности, соответствующие отношению эквивалентности $\sim$ на X, обозначается $X/\sim$ и называется фактормножеством (множества X по отношению $\sim$).*

Сюръективное отображение

$$p : x \mapsto p(x) = \widetilde{x}$$

называется естественным отображением или канонической проекцией X на $X/\sim$.

Пусть X и Y – два множества, а $f : X \to Y$ – некоторое отображение. Определим бинарное отношение $\sim_f$ на X следующим образом:

$$x_1 \sim_f x_2 \quad \Longleftrightarrow \quad f(x_1) = f(x_2), \quad x_1, x_2 \in X.$$

В качестве ***упражнения*** предлагается проверить, что $\sim_f$ является отношением эквивалентности на X. Обозначим классы эквивалентности по этому отношению $\widetilde{x}_f$, т.е. $\widetilde{x}_f = \{z \in X : \ f(z) = f(x)\}$.

Определим отображение $\widetilde{f} : X/\sim_f \to Y$ следующим образом

$$\widetilde{f}(\widetilde{x}_f) = f(x). \tag{1.2}$$

Так как $\widetilde{x}_f = \widetilde{z}_f$ если и только если $f(x) = f(z)$, то соотношение (1.2) не зависит от выбора представителя x в классе $\widetilde{x}_f$. Таким образом, отображение $\widetilde{f}$ определено корректно. При этом отображение $\widetilde{f}$ является *инъективным.* В самом деле, $\widetilde{f}(\widetilde{x}_f) = \widetilde{f}(\widetilde{z}_f)$ если и только если $f(x) = f(z)$, что эквивалентно тому, что $\widetilde{x}_f = \widetilde{z}_f$.

Заметим, что равенство (1.2) можно представить в виде

$$(\widetilde{f} \circ p)(x) = f(x),$$

где p — каноническая проекция. Таким образом, нами получена *факторизация* (разложение) отображения $f : X \to Y$ в композицию $f = \widetilde{f} \circ p$ сюръективного отображения p и инъективного отображения $\widetilde{f}$.

В качестве ***упражнения*** предлагается проверить, что биективность отображения f равносильна сюръективности отображения $\widetilde{f}$.

Оказывается, что отображение $\widetilde{f}$ единственно в следующем смысле. Предположим, что существует другое отображение $f_1 : X/\sim_f \to X$ такое, что $f = f_1 \circ p$. Тогда из цепочки равенств $f_1(\widetilde{x}_f) = f_1(p(x)) = f(x) = \widetilde{f}(\widetilde{x}_f)$ вытекает, что $f_1 = \widetilde{f}$. Следовательно, *отображение $\widetilde{f}$, дающее факторизацию f вида $f = \widetilde{f} \circ p$ единственно.*

1.3. Отношения порядка, упорядоченные множества

Пусть на множестве X задано бинарное отношение $\preceq$. Если это отношение является рефлексивным (т.е. $x \preceq x$ для любого $x \in X$), антисимметричным (т.е. из условий $x_1 \preceq x_2$ и $x_2 \preceq x_1$ вытекает $x_1 = x_2$) и транзитивным т.е. из условий $x_1 \preceq x_2$ и $x_2 \preceq x_3$ вытекает $x_1 \preceq x_3$ для всех $x_1, x_2, x_3 \in X$), то отношение $\preceq$ называется *порядком* на X.

Вместе с отношением $\preceq$ обычно рассматривают отношение $\succeq$, определяемое так: $x_1 \succeq x_2$ если $x_2 \preceq x_1$. Кроме того, определяются отношения $\prec$ и $\succ$: $x_1 \prec x_2$ если $x_1 \preceq x_2$ и $x_1 \neq x_2$, а $x_1 \succ x_2$ если $x_1 \succeq x_2$ и $x_1 \neq x_2$.

В общем случае, пара элементов x_1, x_2 множества X может как находится в отношении $\preceq$, так и не находится в этом отношении. Если для каждых двух элементов x_1 и x_2 множества X верно одно из двух выражений $x_1 \preceq x_2$ или $x_2 \preceq x_1$, то множество X называется *вполне упорядоченным.*

Примерами отношений порядка могут служить, в частности, отношения $A \subset B$ на множестве $\mathfrak{S}(X)$ всех подмножеств некоторого множества X, стандартное отношение $\leqslant$ на множестве $\mathbb{R}$ всех вещественных чисел, отношение делимости на множестве $\mathbb{N}$ натуральных чисел.

1.4. Понятие числового поля

Пусть $\mathcal{F}$ – множество элементов (чисел) произвольной природы. Предположим, что на множестве $\mathcal{F} \times \mathcal{F} = \{(x, y) : \ x, y \in \mathcal{F}\}$ упорядоченных пар элементов множества $\mathcal{F}$ заданы две функции, скажем φ и ψ, принимающие значения в $\mathcal{F}$. Такие функции называются *алгебраическими операциями* на $\mathcal{F}$. Назовем φ *операцией сложения*, а ψ — *операцией умножения.* Выражение $\varphi(x, y)$ обозначим через $x + y$, а выражение $\psi(x, y)$ — через xy (используются также обозначения $x \times y$ и $x \cdot y$).

Определение. *Скажем, что множество $\mathcal{F}$, рассматриваемое совместно с операциям сложения $\varphi = +$ и умножения $\psi = \times$, образует поле, если*

(1) *Операция $+$ ассоциативна и коммутативна, т.е. для любых $x, y, z \in \mathcal{F}$ выполняются равенства $(x + y) + z = x + (y + z)$ и $x + y = y + x$.*

(2) *Существует элемент $0 \in \mathcal{F}$ такой что $0 + x = x + 0 = x$ для любого $x \in \mathcal{F}$. Этот элемент 0 называется нулем поля $\mathcal{F}$ (или нейтральным элементом относительно операции сложения).*

(3) *Для любого элемента $x \in \mathcal{F}$ существует элемент $y \in \mathcal{F}$ такой, что $x + y = y + x = 0$. Такой элемент y называется противоположным для x и обозначается символом $-x$.*

(4) *Операция $\times$ ассоциативна и коммутативна, т.е. для любых $x, y, z \in \mathcal{F}$ выполняются равенства $(xy)z = x(yz)$ и $xy = yx$.*

(5) *Существует элемент* $1 \in \mathcal{F}$ *такой что* $1x = x1 = x$ *для любого* $x \in \mathcal{F}$. *Этот элемент* 1 *называется единицей поля* $\mathcal{F}$ *(или нейтральным элементом относительно операции умножения).*

(6) *Для любого элемента* $x \in \mathcal{F}^* := \mathcal{F} \setminus \{0\}$ *существует элемент* z *такой, что* $xz = zx = 1$. *Такой элемент* z *называется обратным для* x *и обозначается символом* x^{-1}.

(7) *Операции* $+$ *и* $\times$ *связаны дистрибутивным законом, т.е. для любых элементов* $x, y, z \in \mathcal{F}$ *имеет место равенство* $x(y+z) = xy + xz$.

Говорят также, что $\mathcal{F}$ *образует поле (является полем) относительно операций* $+$ *и* $\times$. *Свойства (1)-(7) часто называют аксиомами поля.*

Из аксиом поля в частности вытекает, что для любого элемента $x \in \mathcal{F}$ выполняется равенство $x0 = 0x = 0$. Кроме того, если $\mathcal{F}$ нетривиальное поле, т.е если $\mathcal{F} \neq \{0\}$, то $1 \neq 0$. Проверка этих свойств поля оставляется в качестве *упражнения*.

Замечание. Противоположный элемент $-x$, существование которого для любого элемента $x \in \mathcal{F}$ утверждается в аксиоме (3) поля определен для каждого $x \in \mathcal{F}$ единственным образом. В самом деле, если $y_{1,2} \in \mathcal{F}$ такие, что $x + y_1 = y_1 + x = 0$ и $x + y_2 = y_2 + x = 0$, то $y_2 = y_2 + 0 = y_2 + x + y_1 = 0 + y_1 = y_1$. Аналогично доказывается единственность обратного элемента для любого элемента $x \in \mathcal{F}^*$.

Кроме того, если $\mathcal{F}$ — поле, $x, y \in \mathcal{F}$ — произвольные его элементы такие, что $xy = 0$, то хотя бы один из элементов x или y должен равняться нулю. В самом деле, если $x \neq 0$ то существует $x^{-1} \in \mathcal{F}$ такой, что $x^{-1}x = 1$ и из равенства $xy = 0$ вытекает, что $y = 1y = x^{-1}xy = x^{-1}0 = 0$. Аналогично разбирается случай $y \neq 0$. Установленное свойство называется свойством *отсутствия делителей нуля* в поле $\mathcal{F}$.

В поле $\mathcal{F}$ выражение $x + (-y)$ записывают в виде $x - y$. Операция $(x, y) \mapsto x - y$ называется операцией *вычитания*. Кроме того, выражение xy^{-1}, где $x \in \mathcal{F}$, а $y \in \mathcal{F}^*$ записывают в виде x/y, а операция $(x, y) \mapsto x/y$ называется операцией *деления*.

Важные *примеры* полей образуют множества *рациональных чисел* $\mathbb{Q}$ и *вещественных чисел* $\mathbb{R}$, рассматриваемые со стандартными операциями сложения и умножения.

Упражнение. Проверить, что множество $\mathbb{Q}(\sqrt{2}) := \{a + b\sqrt{2} : \ a, b \in \mathbb{Q}\}$ образует поле относительно следующий операций сложения $\oplus$ и умножения $\otimes$:

$$(a+b\sqrt{2}) \oplus (c+d\sqrt{2}) = (a+c)+(b+d)\sqrt{2}, \quad (a+b\sqrt{2}) \otimes (c+d\sqrt{2}) = (ac+2bd)+(ad+bc)\sqrt{2}.$$

Если два поля $\mathcal{F}$ и $\mathcal{F}'$ связаны соотношением $\mathcal{F}' \subset \mathcal{F}$, то говорят, что поле $\mathcal{F}'$ является *подполем* поля $\mathcal{F}$, а поле $\mathcal{F}$ — *расширением* поля $\mathcal{F}'$.

Пусть $\mathcal{F}_1$ – поле относительно операций $+$ и $\times$, а $\mathcal{F}_2$ –поле относительно операций $\oplus$ и $\otimes$. Поля $\mathcal{F}_1$ и $\mathcal{F}_2$ называются *изоморфными*, если существует биективное отображение $F : \mathcal{F}_1 \to \mathcal{F}_2$ такое, что для любых элементов $x, y \in \mathcal{F}_1$ выполнены равенства $F(x+y) = F(x) \oplus F(y)$ и $F(x \times y) = F(x) \otimes F(y)$.

1.5. Комплексные числа

Поле комплексных чисел. По определению, $\mathbb{C} = \{x + iy : \ x, y \in \mathbb{R}\}$, где i — символ и введены следующие *операции*:

$$z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2), \quad z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$$

при условии, что $z_{1,2} = x_{1,2} + iy_{1,2}$.

Упражнение. Проверить, что $\mathbb{C}$ является полем и показать, что подполе $\{x + i0 : x \in \mathbb{R}\}$ изоморфно $\mathbb{R}$ (далее они отождествляются).

Заметим, что $i^2 = (0 + i1)^2 = -1 + i0 = -1$.

Нулем и *единицей* в $\mathbb{C}$ являются $0 = 0 + i0$ и $1 = 1 + i0$ соответственно. При $z \neq 0$ *обратный элемент* числа $z \in \mathbb{C}$ находится по формуле:

$$\frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{x - iy}{x^2 + y^2} = \frac{x}{x^2 + y^2} + i\left(-\frac{y}{x^2 + y^2}\right),$$

где $\overline{z} = x - iy$ – число, *сопряженное* к $z = x + iy$.

Определение. *Выражение $z = x + iy$ называют алгебраической формой комплексного числа z, $x = \operatorname{Re} z$ – его действительной частью, а $y = \operatorname{Im} z$ – мнимой частью z.*

Тригонометрическая форма комплексного числа. При $z = x + iy$ положим $|z| = \sqrt{x^2 + y^2}$ – *модуль* числа z. При этом, если $r = |z|$, то $z\overline{z} = r^2$.

Пусть $z \in \mathbb{C}^*$. Существует (единственное) число $\theta_0 \in (-\pi, \pi]$ такое, что $x = r\cos\theta_0$, $y = r\sin\theta_0$. Это число называется *главным значением (полярного) аргумента* z и обозначается $\arg z$. Кроме того вводится $\operatorname{Arg} z = \{\theta_0 + 2k\pi \ : \ k \in \mathbb{Z}\}$ – *совокупный (полярный) аргумент* числа z. При любом $\theta \in \operatorname{Arg} z$ имеем $z = r(\cos\theta + i\sin\theta)$. Полезно заметить, что если $z = x + iy$ и $x > 0$ (z лежит в *правой* полуплоскости), то $\arg z = \operatorname{arctg}(y/x)$.

Определение. *Выражение $z = r(\cos\theta + i\sin\theta)$ называется тригонометрической формой комплексного числа z.*

Предложение (формула Муавра)**.** *Если $z = r(\cos\theta + i\sin\theta) \neq 0$, то при $n \in \mathbb{N}$*

$$z^n = r^n\big(\cos(n\theta) + i\sin(n\theta)\big).$$

Это утверждение вытекает из того, что если $\theta_{1,2} \in \operatorname{Arg}(z_{1,2})$, $r_{1,2} = |z_{1,2}|$, то

$$z_1z_2 = r_1r_2(\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)).$$

Проверка последнего соотношения не представляет сложностей, делается непосредственно на основании определения операций над комплексными числами и оставляется в качестве *упражнения*.

Корни степени n ($\sqrt[n]{z}$). Пусть $n \in \mathbb{N}$, $n \geqslant 2$. По определению, $w \in \sqrt[n]{z} \Longleftrightarrow w^n = z$. Из формулы Муавра следует, что при $z \neq 0$ совокупность $\sqrt[n]{z}$ состоит из n элементов $\{w_0, w_1, \dots, w_{n-1}\}$, находящихся по формуле

$$w_k = \sqrt[n]{z}_{(k)} = \sqrt[n]{r}\left(\cos\left(\frac{\varphi_0 + 2k\pi}{n}\right) + i\sin\left(\frac{\varphi_0 + 2k\pi}{n}\right)\right),$$

$k = 0, \dots, n-1$. Ясно, что $\sqrt[n]{0} = \{0\}$.

Экспоненциальная (показательная) форма записи комплексных чисел. По определению, при любом $z \in \mathbb{C}$,

$$e^z = \exp(z) = \lim_{n\to\infty}\left(1 + \frac{z}{n}\right)^n.$$

Пусть $z = x + iy$. *Проверим*, что $e^z = e^x(\cos y + i\sin y)$, т.е. $|e^z| = e^x$, $\operatorname{Arg}(e^z) = \{y + 2\pi k, \ k \in \mathbb{Z})$.

Действительно, пусть n – достаточно велико, тогда $1 + \frac{z}{n} = 1 + \frac{x}{n} + i\frac{y}{n}$ лежит в правой полуплоскости. По формуле Муавра находим:

$$\left|1 + \frac{z}{n}\right|^n = \left(\left(1 + \frac{x}{n}\right)^2 + \left(\frac{y}{n}\right)^2\right)^{n/2} = \exp\left(\frac{n}{2}\ln\left(1 + \frac{2x}{n} + o(n^{-1})\right)\right) \to e^x,$$

$$\arg\left(\left(1 + \frac{z}{n}\right)^n\right) = (\mod 2\pi) = n\operatorname{arctg}\left(\frac{y}{x+n}\right) \to y$$

при $n \to +\infty$. Что и требовалось.

Пусть теперь $z = x + iy \in \mathbb{C}$ и $w = u + iv \in \mathbb{C}$. Тогда
$e^z e^w = e^x e^u(\cos y + i\sin y)(\cos v + i\sin v) = e^{x+u}(\cos(y+v)+i\sin(y+v)) = e^{x+u+i(y+v)} = e^{z+w}$.

Из полученной формулы вытекает следующее важное свойство комплексных чисел. Пусть $z \neq 0$, $r = |z|$, $\varphi = \arg z$, тогда $z = re^{i\varphi}$; в частности, $\cos\varphi = (e^{i\varphi} + e^{-i\varphi})/2$ и $\sin\varphi = (e^{i\varphi} - e^{-i\varphi})/(2i)$.

Определение. *Выражение $z = re^{i\varphi}$ называется экспоненциальной (или показательной) формой комплексного числа z.*

Корни степени n из единицы. Рассмотрим уравнение $\omega^n = 1$ (где ω – комплексное число). Это уравнение имеет n различных корней $\omega_n^k = e^{2\pi ik/n} = \cos\left(2\pi\frac{k}{n}\right) + i\sin\left(2\pi\frac{k}{n}\right)$ при $k = 0, 1, \ldots, n-1$. Приведем ряд полезных свойств чисел ω_n^k, доказательство которых оставляется в качестве *упражнений*:

1. Числа ω_n^k геометрически расположены в вершинах правильного n-угольника, вписанного в окружность радиуса 1 с центром в точке $0 = 0 + 0i = (0,0)$ так, что точка $1 = 1 + 0i = (1,0)$ является одной из вершин этого многоугольника.

2. Если $\omega_n := e^{\frac{2\pi i}{n}}$, то $\omega_n^k = (\omega_n)^k$. Отсюда $\omega_n^{j+k} = \omega_n^{j+k \pmod n}$ и $\omega_n^{-1} = \omega_n^{n-1}$.

3. Для любых целых $n \geqslant 0$, $k \geqslant 0$ и $d > 0$ имеет место $\omega_{nd}^{kd} = \omega_n^k$.

4. Пусть n четно. Тогда $\omega_n^{n/2} = \omega_2 = -1$. Кроме того, если $\{\omega_n^k\}_{k=0}^{n-1}$ — совокупность всех корни степени n из единицы, то $\left\{(\omega_n^k)^2\right\}$ — совокупность всех корней степени $n/2$ из единицы, причем каждый корень степени $n/2$ входит в эту совокупность дважды.

Интересно отметить, что для любого $n \geqslant 1$ и для любого $k > 0$, не кратного n,

$$\sum_{j=0}^{n-1}(\omega_n^k)^j = \frac{(\omega_n^k)^n - 1}{\omega_n^k - 1} = \frac{(\omega_n^n)^k - 1}{\omega_n^k - 1} = \frac{1-1}{\omega_n^k - 1} = 0,$$

где $\omega_n^k \neq 1$ так как k не кратно n. Число $\omega_n = \omega_n^1$ называется *главным (первообразным) корнем степени n из единицы.*

Матричная модель комплексных чисел. В качестве *упражнения* предлагается проверить, что множество матриц

$$\left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} :\ a, b \in \mathbb{R}\right\} \subset \mathrm{M}_2(\mathbb{R}) \tag{1.3}$$

рассматриваемое со стандартными операциями матричного сложения и умножения образует поле. Единицей в этом поле является единичная матрица, а нулем – нулевая матрица.

Заметим, что любая матрица A вида (1.3) может быть записана в виде

$$A = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} =: aE + bJ,$$

а также, что

$$J^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E.$$

Таким образом, множество матриц вида (1.3) может быть отождествлено с множеством $\mathbb{C}$ комплексных чисел следующим образом: комплексному числу $a + bi$ ставиться в соответствие матрица $aE + bJ$ и наоборот.

Упражнение. Проверить, что при этом сумме и произведению комплексных чисел соответствуют сумма и произведение соответствующих матриц из (1.3) и наоборот.

Таким образом, поле C и поле матриц вида (1.3) изоморфны. Кроме того, множество матриц вида $\{aE :\ a \in \mathbb{R}\}$ образует подполе поля (1.3), изоморфное полю $\mathbb{R}$, а матричное соотношение $J^2 + E = 0$ является аналогом соотношения $i^2 = -1$ в $\mathbb{C}$.

РАЗДЕЛ 2

Линейные пространства

2.1. Понятие линейного пространства

Пусть $\mathcal{F}$ – некоторое поле. Элементы поля $\mathcal{F}$ будем называть *числами.*

Определение. *Множество $\mathcal{X}$ объектов произвольной природы называется линейным пространством над полем $\mathcal{F}$, если*

(1) *В $\mathcal{X}$ задана операция сложения, т.е. для любых элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ определен элемент $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y} \in \mathcal{X}$.*

(2) *Для любого элемента $\boldsymbol{x} \in \mathcal{X}$ и для любого числа $\lambda \in \mathcal{F}$ определен элемент $\lambda\boldsymbol{x} \in \mathcal{X}$. Этот элемент называется произведением элемента $\boldsymbol{x}$ на число λ.*

и если выполнены следующие условия

(3) *Операция сложения коммутативна, т.е. $\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$ для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.*

(4) *Операция сложения ассоциативна, т.е. $(\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z} = \boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z})$ для любых элементов $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{X}$.*

(5) *Существует элемент $\mathbf{0} \in \mathcal{X}$ такой, что $\mathbf{0} + \boldsymbol{x} = \boldsymbol{x}$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$. Этот элемент называется нулевым элементом (или нулем) пространства $\mathcal{X}$.*

(6) *Для любого элемента $\boldsymbol{x} \in \mathcal{X}$ существует элемент $\boldsymbol{x}' \in \mathcal{X}$ такой, что $\boldsymbol{x}+\boldsymbol{x}' = \mathbf{0}$. Элемент $\boldsymbol{x}'$ называется противоположным элементом для элемента $\boldsymbol{x}$.*

(7) *Имеет место равенство $1\boldsymbol{x} = \boldsymbol{x}$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$, где $1 = 1_{\mathcal{F}}$ – единица поля $\mathcal{F}$.*

(8) *Для любого элемента $\boldsymbol{x} \in \mathcal{X}$ и для любых чисел $\lambda, \mu \in \mathcal{F}$ верны равенства $\lambda(\mu\boldsymbol{x}) = (\lambda\mu)\boldsymbol{x}$ и $(\lambda + \mu)\boldsymbol{x} = \lambda\boldsymbol{x} + \mu\boldsymbol{x}$.*

(9) *Для любых элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ и для любого числа $\lambda \in \mathcal{F}$ верно равенство $\lambda(\boldsymbol{x} + \boldsymbol{y}) = \lambda\boldsymbol{x} + \lambda\boldsymbol{y}$.*

Линейное пространство над полем $\mathbb{R}$ вещественных чисел называется вещественным линейным пространством. Линейной пространство над полем $\mathbb{C}$ комплексных чисел называется комплексным линейным пространством.

Приведем несколько простых примеров линейных пространств. Ясно, что множества $\mathbb{R}$ и $\mathbb{C}$ со стандартными операциями сложения чисел и умножения вещественных и комплексных чисел на вещественные числа являются вещественными линейными пространствами. Более того, множество $\mathbb{C}$, также рассматриваемое со стандартными операциями сложения и умножения комплексных чисел является комплексным линейным пространством.

Рассмотрим теперь более экзотический пример. Пусть $\mathbb{R}_+$ – это множество всех положительных вещественных чисел. Определим сумму элементов $x, y \in \mathbb{R}_+$ этого множества как произведение xy, а умножение элемента $x \in \mathbb{R}_+$ этого множества на вещественное число α как x^α. Проверка всех аксиом вещественного линейного пространства в этом примере оставляется в качестве *упражнения*.

Пусть теперь $\mathbb{R}^n$, $n \in \mathbb{N}$ – множество упорядоченных наборов, состоящих из n произвольных вещественных чисел $\boldsymbol{x} = (x_1, \ldots, x_n)$. Операции сложения элементов множества $\mathbb{R}^n$ и их умножения на (вещественные) числа определяются следующим образом

$$(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n),$$
$$\lambda(x_1, \ldots, x_n) = (\lambda x_1, \ldots, \lambda x_n),$$

где $\lambda \in \mathbb{R}$. Проверка того, что $\mathbb{R}^n$ является линейным пространством оставляется в качестве ***упражнения***. Заметим, что нулевым элементом в пространстве $\mathbb{R}^n$ является $\mathbf{0} = (0, \ldots, 0)$, а противоположным элементом для элемента $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ – элемент $(-x_1, \ldots, -x_n)$. Пространство $\mathbb{R}^n$ часто называют *вещественным n-мерным координатным пространством.*

Вещественными линейными пространствами являются также совокупность $C([a, b])$ всех непрерывных вещественнозначных функций на некотором отрезке $[a, b]$ и множество $\mathbb{R}[t]_n$ всех многочленов с вещественными коэффициентами степени не выше n.

Однако совокупность всех многочленов от вещественной переменной с вещественными коэффициентами степени равной n не является линейным пространством (так как сумма двух таких многочленов может иметь меньшую степень). Не будет вещественным линейным пространством и совокупность всех многочленов степени не выше n с положительными вещественными коэффициентами (такие многочлены нельзя умножать на отрицательные числа).

Замечание. Здесь и далее используется стандартная общеупотребительная терминология, согласно которой элементы линейных пространств называются *векторами.*

Тот факт, что термин вектор обычно употребляется в более узком смысле, не только не приводит к недоразумениям, но и позволяет лучше понимать и, во многих случаях, предвидеть (на основе геометрической интуиции) результаты, справедливые для линейных пространств произвольной природы.

Предложение 2.1. *Пусть $\mathcal{X}$ – линейное пространство над полем $\mathcal{F}$.*

1. *Нулевой элемент в $\mathcal{X}$ является единственным. Для любого элемента $\boldsymbol{x} \in \mathcal{X}$ противоположный элемент $\boldsymbol{x}'$ является единственным.*

2. *Для любого $\boldsymbol{x} \in \mathcal{X}$ справедливо $0\boldsymbol{x} = \mathbf{0}$, а противоположный к $\boldsymbol{x}$ элемент в $\mathcal{X}$ равен $(-1)\boldsymbol{x}$ (здесь $0 = 0_{\mathcal{F}}$ и $1 = 1_{\mathcal{F}}$ – ноль и единица в $\mathcal{F}$).*

Доказательство. Предположим, что в некотором линейном пространстве $\mathcal{X}$ существуют два нулевых элемента: $\mathbf{0}_1$ и $\mathbf{0}_2$. Тогда из свойств нулевого элемента и коммутативности суммы элементов линейного пространства вытекает, что $\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2$ и $\mathbf{0}_1 + \mathbf{0}_2 = \mathbf{0}_2 + \mathbf{0}_1 = \mathbf{0}_1$, т.е. $\mathbf{0}_1 = \mathbf{0}_2$. Пусть теперь для некоторого элемента $\boldsymbol{x} \in \mathcal{X}$ существуют два противоположных элемента $\boldsymbol{y}_1$ и $\boldsymbol{y}_2$. Тогда

$$\boldsymbol{y}_1 = \boldsymbol{y}_1 + \mathbf{0} = \boldsymbol{y}_1 + (\boldsymbol{x} + \boldsymbol{y}_2) = (\boldsymbol{y}_1 + \boldsymbol{x}) + \boldsymbol{y}_2 = \mathbf{0} + \boldsymbol{y}_2 = \boldsymbol{y}_2.$$

Далее, если $x \in \mathcal{X}$, а $\boldsymbol{y}$ – противоположный элемент для $\boldsymbol{x}$, то

$$0\boldsymbol{x} = 0\boldsymbol{x} + \mathbf{0} = 0\boldsymbol{x} + (\boldsymbol{x} + \boldsymbol{y}) = (0\boldsymbol{x} + 1\boldsymbol{x}) + \boldsymbol{y} = (0 + 1)\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{y} = \mathbf{0}.$$

И, наконец,

$$\boldsymbol{x} + (-1)\boldsymbol{x} = (1 + (-1))\boldsymbol{x} = 0\boldsymbol{x} = \mathbf{0}. \quad \square$$

Замечание. Из аксиом линейного пространства и только что доказанного предложения вытекает, что для любых двух элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ линейного пространства $\mathcal{X}$ существует единственный элемент $\boldsymbol{z} \in \mathcal{X}$ такой, что $\boldsymbol{z} + \boldsymbol{y} = \boldsymbol{x}$. В самом деле, в качестве $\boldsymbol{z}$ нужно взять $\boldsymbol{x} + (-1)\boldsymbol{y}$. Этот элемент $\boldsymbol{z}$ называется *разностью* элементов $\boldsymbol{x}$ и $\boldsymbol{y}$ и обозначается $\boldsymbol{x} - \boldsymbol{y}$.

2.2. Линейная зависимость векторов в линейном пространстве

Рассмотрим линейное пространство $\mathcal{X}$ над полем $\mathcal{F}$ и некоторую совокупность векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ пространства $\mathcal{X}$.

Определение. *Выражение*

$$\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_n\boldsymbol{x}_n,$$

где $\alpha_1 \in \mathcal{F}, \ldots, \alpha_n \in \mathcal{F}$, называется линейной комбинацией элементов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$.

Линейная комбинация элементов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ *называется нетривиальной, если хотя бы один из ее коэффициентов* $\alpha_1, \ldots, \alpha_n$ *отличен от нуля. Линейная комбинация элементов* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, *у которой все коэффициенты* $\alpha_1, \ldots, \alpha_n$ *равны нулю называется тривиальной.*

Скажем далее, что элемент $\boldsymbol{y}$ линейного пространства $\mathcal{X}$ является линейной комбинацией элементов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathcal{X}$, если существует некоторая линейная комбинация элементов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$, равная $\boldsymbol{y}$.

Определение. *Элементы* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ *линейного пространства* $\mathcal{X}$ *называются линейно зависимыми, если существует их нетривиальная линейная комбинация, равная* $\mathbf{0}$.

Другими словами, векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ линейного пространства $\mathcal{X}$ линейно зависимы, если существуют n чисел $\alpha_1 \in \mathcal{F}, \ldots, \alpha_n \in \mathcal{F}$, хотя бы одно из которых не равно нулю, такие, что $\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_n \boldsymbol{x}_n = \mathbf{0}$

Определение. *Элементы* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ *линейного пространства* $\mathcal{X}$ *называются линейно независимыми, если их линейная комбинация равна* $\mathbf{0}$ *тогда и только тогда, когда она тривиальна.*

Предложение 2.2. *Элементы* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ *линейного пространства* $\mathcal{X}$ *линейно зависимы если и только если один из них является линейной комбинацией остальных.*

Доказательство. Пусть векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathcal{X}$ линейно зависимы. Тогда существуют числа $\alpha_1 \in \mathcal{F}, \ldots, \alpha_n \in \mathcal{F}$ такие, что $\mathbf{0} = \alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_n \boldsymbol{x}_n$ и хотя бы одно из чисел $\alpha_1, \ldots, \alpha_n$ отлично от нуля. Без ограничения общности предположим, что $\alpha_n \neq 0$. Тогда

$$\boldsymbol{x}_n = \beta_1 \boldsymbol{x}_1 + \cdots + \beta_{n-1} \mathsf{x}_{n-1},$$

где $\beta_k = -\alpha_k / \alpha_n$ при $k = 1, \ldots, n-1$. Т.е., элемент $\boldsymbol{x}_n$ является линейной комбинацией остальных. Обратно, пусть элемент $\boldsymbol{x}_1$ является линейной комбинацией элементов $\boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$, т.е. $\boldsymbol{x}_1 = \gamma_2 \boldsymbol{x}_2 + \cdots + \gamma_n \boldsymbol{x}_n$ для некоторых $\gamma_2 \in \mathcal{F}, \ldots, \gamma_n \in \mathcal{F}$. Тогда

$$(-1)\boldsymbol{x}_1 + \gamma_2 \boldsymbol{x}_2 + \cdots + \gamma_n \boldsymbol{x}_n = \mathbf{0},$$

а стоящая в левой части этого равенства линейная комбинация элементов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ является нетривиальной. □

Замечание. Пусть $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ и $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_m$ – элементы линейного пространства $\mathcal{X}$. Тогда любой набор вида $\{\mathbf{0}, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ будет линейно зависимым. Кроме того, если элементы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ линейно зависимы, то и любой набор вида $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \boldsymbol{y}_1, \ldots, \boldsymbol{y}_m\}$ будет линейно зависимым.

Пример 2.3. Рассмотрим n-мерное вещественное координатное пространство $\mathbb{R}^n$. Определим элементы $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ следующим образом:

$$\boldsymbol{a}_1 = (1, 0, 0, \ldots, 0, 0), \quad \boldsymbol{a}_2 = (0, 1, 0, \ldots, 0, 0), \quad \ldots, \quad \boldsymbol{a}_n = (0, 0, 0, \ldots, 0, 1),$$

т.е. у элемента $\boldsymbol{a}_k$, $k = 1 \ldots n$, компонента с номером k равна 1, а все остальные компоненты равны нулю. Проверим, что векторы $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ – линейно независимы и, что для любого вектора $\boldsymbol{x} \in \mathbb{R}^n$ набор векторов $\boldsymbol{x}, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ будет линейно зависимым.

Рассмотрим произвольную линейную комбинацию векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$ с коэффициентами $\alpha_1, \ldots, \alpha_n$. Из общих свойств линейного пространства и из определения пространства $\mathbb{R}^n$ вытекает, что

$$\alpha_1 \boldsymbol{a}_1 + \cdots + \alpha_n \boldsymbol{a}_n = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n.$$

Этот вектор является нулевым вектором пространства $\mathbb{R}^n$ если и только если $\alpha_1 = \cdots = \alpha_n = 0$. А последнее утверждение в точности означает линейную независимость векторов $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n$. Пусть теперь $\boldsymbol{x} = (x_1, \ldots, x_n)$ – произвольный вектор из пространства $\mathbb{R}^n$. Как и раньше

$$\boldsymbol{x} = (x_1, \ldots, x_n) = x_1 \boldsymbol{a}_1 + \cdots + x_n \boldsymbol{a}_n,$$

что означает, что вектор $\boldsymbol{x}$ есть линейная комбинация векторов $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ и, следовательно, векторы $\boldsymbol{x}, \boldsymbol{a}_1, \dots, \boldsymbol{a}_n$ линейно зависимым.

Замечание. Пусть $A = \|\alpha_{j\ell}\|$ – некоторая скалярная $k \times m$-матрица и пусть $\boldsymbol{X} = \|\boldsymbol{x}_{j\ell}\|$ – $m \times r$-матрица, составленная из векторов $\boldsymbol{x}_{j\ell}$ некоторого линейного пространства $\mathcal{X}$. Тогда определена $k \times r$ матрица $A\boldsymbol{X} = \boldsymbol{Y} = \|\boldsymbol{y}_{j\ell}\|$, где $\boldsymbol{y}_{j\ell} = \sum_{s=1}^{m} \alpha_{js}\boldsymbol{x}_{s\ell}$ при $j = 1 \dots k$ и $\ell = 1 \dots r$ (элементами которой снова являются векторы из $\mathcal{X}$). Нетрудно проверить (это оставляется в качестве ***упражнения***), что $(A_1 A_2)\boldsymbol{X} = A_1(A_2\boldsymbol{X})$ и $A(\boldsymbol{X}_1 + \boldsymbol{X}_2) = A\boldsymbol{X}_1 + A\boldsymbol{X}_2$ для любых скалярных матриц A, A_1 и A_2 и любых векторных матриц $\boldsymbol{X}$, $\boldsymbol{X}_1$ и $\boldsymbol{X}_2$ соответствующего размера. Аналогичным образом определяется произведение $\boldsymbol{X}A$ векторной $k \times m$-матрицы $\boldsymbol{X}$ на скалярную $m \times r$-матрицу A и проверяется выполнение свойств $\boldsymbol{X}(A_1A_2) = (\boldsymbol{X}A_1)A_2$ и $(\boldsymbol{X}_1 + \boldsymbol{X}_2)A = \boldsymbol{X}_1A + \boldsymbol{X}_2A$ для любых скалярных матриц A, A_1 и A_2 и любых векторных матриц $\boldsymbol{X}$, $\boldsymbol{X}_1$ и $\boldsymbol{X}_2$ соответствующего размера.

2.3. Базис и размерность линейного пространства

Определение. *Пусть $\mathcal{X}$ – линейное пространство над полем $\mathcal{F}$. Упорядоченная совокупность линейно независимых векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ в линейном пространстве $\mathcal{X}$ называется базисом пространства $\mathcal{X}$, если для каждого элемента $\boldsymbol{x} \in \mathcal{X}$ найдутся числа $x_1 \in \mathcal{F}, \dots, x_n \in \mathcal{F}$ такие, что имеет место равенство*

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \dots + x_n\boldsymbol{e}_n. \tag{2.1}$$

Другими словами, система линейно независимых векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ пространства $\mathcal{X}$ является базисом в $\mathcal{X}$, если любой вектор $\boldsymbol{x} \in \mathcal{X}$ представим в виде некоторой линейной комбинации векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$.

Предложение 2.4. *Пусть $\mathcal{X}$ – линейное пространство, а $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ – базис в $\mathcal{X}$. Тогда для любого вектора $\boldsymbol{x} \in \mathcal{X}$ числа $x_1, \dots, x_n$ такие, что $\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \dots + x_n\boldsymbol{e}_n$, определяются единственным образом.*

Доказательство. Предположим, что для некоторого элемент $\boldsymbol{x} \in \mathcal{X}$ имеют место два различных представления: $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n = x'_1\boldsymbol{e}_1 + \dots + x'_n\boldsymbol{e}_n$. Тогда $\boldsymbol{0} = \boldsymbol{x} - \boldsymbol{x} = (x_1 - x'_1)\boldsymbol{e}_1 + \dots + (x_n - x'_n)\boldsymbol{e}_n$. Так как векторы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ линейно независимы, то из последнего соотношения вытекает, что $x_1 = x'_1, x_2 = x'_2, \dots, x_n = x'_n$. $\square$

Определение. *Равенство (2.1) называется разложением вектора $\boldsymbol{x} \in \mathcal{X}$ по базису $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$. При этом числа $x_1, \dots, x_n$ называются координатами вектора $\boldsymbol{x}$ (относительно базиса $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$).*

В силу Предложения 2.4 понятия разложения вектора по базису и координат вектора определены корректно. Далее, из свойств линейного пространства и свойства единственности разложения по базису непосредственно вытекает следующее утверждение.

Предложение 2.5. *Пусть $\mathcal{X}$ – линейное пространство с базисом $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ и пусть*

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n \in \mathcal{X}, \quad \boldsymbol{y} = y_1\boldsymbol{e}_1 + \dots + y_n\boldsymbol{e}_n \in \mathcal{X}.$$

Пусть $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$, а $\boldsymbol{v} = \alpha\boldsymbol{x}$ при $\alpha \in \mathcal{F}$. Тогда

$$\boldsymbol{z} = (x_1 + y_1)\boldsymbol{e}_1 + \dots + (x_n + y_n)\boldsymbol{e}_n, \quad \boldsymbol{v} = (\alpha x_1)\boldsymbol{e}_1 + \dots + (\alpha x_n)\boldsymbol{e}_n.$$

Нам потребуется также следующее утверждение.

Предложение 2.6. *Пусть $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – базис линейного пространства $\mathcal{X}$. Векторы $\boldsymbol{x}_1 \in \mathcal{X}, \dots, \boldsymbol{x}_m \in \mathcal{X}$ линейно зависимы если и только если линейно зависимы строки матрицы*

$$\begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \dots & \dots & \dots \\ x_{m,1} & \cdots & x_{m,n} \end{pmatrix} \tag{2.2}$$

где $\boldsymbol{x}_j = x_{j,1}\boldsymbol{e}_1 + \dots + x_{j,n}\boldsymbol{e}_n$, $j = 1, \dots, m$.

Доказательство. Пусть векторы $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ линейно зависимы, а $\alpha_1 \in \mathcal{F}, \dots, \alpha_m \in \mathcal{F}$ – такие числа, не все равные нулю, что $\alpha_1\boldsymbol{x}_1 + \dots + \alpha_m\boldsymbol{x}_m = \mathbf{0}$. Тогда, используя разложения $\boldsymbol{x}_j = x_{j,1}\boldsymbol{e}_1 + \dots + x_{j,n}\boldsymbol{e}_n$, $j = 1, \dots, m$, и приводя подобные члены, получаем

$$(\alpha_1 x_{1,1} + \dots + \alpha_m x_{m,1})\mathbf{e}_1 + \dots + (\alpha_1 x_{1,n} + \dots + \alpha_m x_{m,n})\mathbf{e}_n = \mathbf{0},$$

откуда, в силу линейной независимости векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$, получаем, что

$$\alpha_1 x_{1,j} + \dots + \alpha_m x_{m,j} = 0, \quad j = 1, \dots, n,$$

что в точности означает линейную зависимость строк матрицы (2.2). Обратное утверждение доказывается аналогично. □

Рассмотрим приведенные выше примеры линейных пространств и найдем в них базисы. Первым делом отметим, что в линейном пространстве $\mathbb{R}^n$ система векторов $\boldsymbol{a}_1, \dots, \boldsymbol{a}_n$, введенная выше, образует базис.

Базис в пространстве $\mathbb{R}_+$ состоит из одного элемента, в качестве которого можно взять любое вещественное положительное ненулевое число. В самом деле, для любого $x \in \mathbb{R}_+ \setminus \{0\}$ и для любого $w \in \mathbb{R}_+ \setminus \{0\}$ верно равенство $x = w^\alpha$, где $\alpha = \log_w x$.

В линейном пространстве Π_n многочленов от одной переменной с вещественными коэффициентами степени не выше n базисом будет, например, $\{1, t, t^2, \dots, t^n\}$.

Пусть $\mathcal{X}$ – линейное пространство с базисом $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. Удобно связать с рассматриваемым базисом вектор-строку

$$\boldsymbol{e} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n).$$

Для произвольного элемента $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n \in \mathcal{X}$ рассмотрим вектор-столбец $[\boldsymbol{x}]_{\boldsymbol{e}} = (x_1, \dots, x_n)^\top$, составленный из координат элемента $\boldsymbol{x}$ относительно базиса $\boldsymbol{e}$. При этом

$$\boldsymbol{x} = \boldsymbol{e}\,[\boldsymbol{x}]_{\boldsymbol{e}}.$$

Заметим, что пространство всех столбцов $(x_1, \dots, x_n)^\top$, составленных из вещественных чисел, естественно отождествить с n-мерным координатным вещественным пространством $\mathbb{R}^n$.

Определение. *Линейное пространство $\mathcal{X}$ называется n-мерным (n – натуральное число), если в нем существует n линейно независимых векторов, а любая система, состоящая из $n+1$ вектора является линейно зависимой. Это число n называется размерностью пространства $\mathcal{X}$ и обозначается $\dim \mathcal{X}$. Линейное пространство $\mathcal{X}$ называется бесконечномерным, если в нем существует любое число линейно независимых векторов. В этом случае пишут $\dim \mathcal{X} = \infty$.*

Предложение 2.7. *Если $\mathcal{X}$ – линейное пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$, то любые n линейно независимых векторов из этого пространства образуют базис в $\mathcal{X}$.*

Доказательство. Пусть $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – некоторая система из n линейно независимых векторов в $\mathcal{X}$ (существование такой системы следует из того, что $\dim \mathcal{X} = n$). Пусть $\boldsymbol{x} \in \mathcal{X}$ – произвольный вектор. Тогда, из определения размерности вытекает, что система векторов $\{\boldsymbol{x}, \boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ линейно зависима и, следовательно, $\alpha\boldsymbol{x} + \alpha_1\boldsymbol{e}_1 + \dots + \alpha_n\boldsymbol{e}_n = 0$. В последнем равенстве коэффициент α заведомо не равен нулю, так как в противном случае нетривиальная линейная комбинация векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ равна нулю, что противоречит линейной независимости векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$. Итак (разделив последнее соотношение на α) получаем, что произвольный вектор $\boldsymbol{x} \in \mathcal{X}$ допускает разложение вида $\boldsymbol{x} = \beta_1\boldsymbol{e}_1 + \dots + \beta_n\boldsymbol{e}_n = 0$ (где $\beta_k = -\alpha_k/\alpha$ при $k = 1, \dots, n$), что означает в свою очередь, что $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ образует базис в пространстве $\mathcal{X}$. □

Теорема 2.8. *Если линейное пространство $\mathcal{X}$ имеет базис, состоящий из n, $n \in \mathbb{N}$, элементов, то $\dim \mathcal{X} = n$.*

Доказательство. Пусть система векторов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ является базисом в пространстве $\mathcal{X}$. По определению базиса, эта система состоит из линейно независимых векторов. Для доказательства утверждения теоремы нам необходимо проверить, что любая система $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n+1}\}$, состоящая из $n+1$ вектора является линейно зависимой. Разложим каждый вектор из этой системы по базису $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$:

$$\boldsymbol{x}_k = x_{k,1}\boldsymbol{e}_1 + \cdots + x_{k,n}\boldsymbol{e}_n,$$

где $x_{k,j}$ при $k = 1, \dots, n+1$ и $j = 1, \dots, n$ – некоторые числа. Согласно Предложению 2.6, линейная зависимость векторов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_{n+1}$ эквивалентна линейной зависимости строк матрицы

$$X = \begin{pmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \dots & \dots & \dots & \dots \\ x_{n+1,1} & x_{n+1,2} & \cdots & x_{n+1,n} \end{pmatrix}.$$

Строки этой матрицы заведомо являются линейно зависимыми в силу теоремы о базисном миноре, доказываемой в курсе аналитической геометрии. □

Возвращаясь к ранее рассмотренным ***примерам*** линейных пространств, заметим, что $\dim \mathbb{R}_+ = 1$, $\dim \mathbb{R}^n = n$. Далее, $\dim C([a,b]) = \infty$. В самом деле, для любого натурального числа n система функций из $C([a,b])$ вида $\{1, t, t^2, \dots, t^n\}$ является линейно независимой (проверка этого факта оставляется в качестве ***упражнения***).

2.4. Изоморфизм линейных пространств

В этом разделе будет установлено, что различные линейные пространства одной и той же размерности в смысле свойств, связанных с введенными в этих пространствах операциями, по существу не отличаются друг от друга.

Напомним, что задание на множестве $\mathcal{X}$ структуры линейного пространства над полем $\mathcal{F}$ предполагает задание двух операций: сложения элементов из $\mathcal{X}$ друг с другом и их умножения на числа из $\mathcal{F}$ (причем эти операции должны удовлетворять аксиомам линейного пространства).

Предположим, что поле $\mathcal{F}$ фиксировано, а все линейные пространства в этом разделе рассматриваются над полем $\mathcal{F}$.

Определение. *Два произвольных вещественных линейных пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ называются изоморфными, если существует биективное отображение $\varphi : \mathcal{X}_1 \to \mathcal{X}_2$ такое, что*

(1) $\varphi(\boldsymbol{x} + \boldsymbol{y}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$ *для любых* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_1$.
(2) $\varphi(\alpha\boldsymbol{x}) = \alpha\varphi(\boldsymbol{x})$ *для любого* $\boldsymbol{x} \in \mathcal{X}_1$ *и* $\alpha \in \mathcal{F}$.

Такое отображение φ называется изоморфизмом линейных пространств $\mathcal{X}_1$ и $\mathcal{X}_2$. Факт изоморфизма линейных пространств $\mathcal{X}_1$ и $\mathcal{X}_2$ будем обозначать $\mathcal{X}_1 \cong \mathcal{X}_2$.

Из этого определения вытекает, что изоморфные линейные пространства неразличимы по своей внутренней структуре, т.е. любые понятия, свойства и т.п., выраженные с помощью операций и аксиом линейного пространства, у изоморфных пространств совпадают.

Пусть пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ изоморфны и пусть $\varphi : \mathcal{X}_1 \to \mathcal{X}_2$ – соответствующий изоморфизм.

Тогда, во-первых, нулевому элементу $\mathbf{0}_1 \in \mathcal{X}_1$ соответствует нулевой элемент $\mathbf{0}_2 \in \mathcal{X}_2$. В самом деле, $\varphi(\mathbf{0}_1) = \varphi(0\boldsymbol{x}) = 0\varphi(\boldsymbol{x}) = \mathbf{0}_2$ для произвольного элемента $\boldsymbol{x} \in \mathcal{X}_1$.

Двигаясь далее, рассмотрим семейство векторов $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\} \subset \mathcal{X}_1$ и какую-нибудь их нетривиальную линейную комбинацию $\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_m\boldsymbol{x}_m$. Тогда $\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_m\boldsymbol{x}_m = \mathbf{0}_1$ если и только если $\varphi(\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_m\boldsymbol{x}_m) = \alpha_1\varphi(\boldsymbol{x}_1) + \cdots + \alpha_m\varphi(\boldsymbol{x}_m) = \mathbf{0}_2$. Из этого непосредственно вытекает следующее свойство изоморфных линейных пространств.

Предложение 2.9. *Максимальное количество линейно независимых элементов в каждом из изоморфных линейных пространств $\mathcal{X}_1$ и $\mathcal{X}_2$ одинаково.*

Следующая теорема выражает один из фундаментальных фактов теории линейных пространств.

Теорема 2.10. *Если два линейных пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ изоморфны, то они имеют одинаковую размерность. Любые два линейных пространства $\mathcal{X}_1$ и $\mathcal{X}_2$, имеющие одинаковую размерность $n \in \mathbb{N}$ изоморфны.*

Из Теоремы 2.10 вытекает, в частности, что линейные пространства разной размерности не могут быть изоморфны.

Доказательство. Первое утверждение теоремы является непосредственным следствием Предложения 2.9. Докажем второе утверждение теоремы. Рассмотрим два вещественных линейных пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ конечной размерности n. Выберем в пространстве $\mathcal{X}_1$ базис $\{\boldsymbol{e}_1^1, \ldots, \boldsymbol{e}_n^1\}$, а в пространстве $\mathcal{X}_2$ базис $\{\boldsymbol{e}_1^2, \ldots, \boldsymbol{e}_n^2\}$. Определим отображение $\varphi : \mathcal{X}_1 \to \mathcal{X}_2$ следующим образом: для элемента $\boldsymbol{x} \in \mathcal{X}_1$ вида $\boldsymbol{x} = x_1\boldsymbol{e}_1^1 + \cdots + x_n\boldsymbol{e}_n^1$ положим

$$\varphi(\boldsymbol{x}) = \varphi(x_1\boldsymbol{e}_1^1 + \cdots + x_n\boldsymbol{e}_n^1) = x_1\boldsymbol{e}_1^2 + \cdots + x_n\boldsymbol{e}_n^2.$$

Т.е., мы ставим в соответствие элементу $\boldsymbol{x} \in \mathcal{X}_1$ с координатами $x_1, \ldots, x_n$ относительно базиса $\{\boldsymbol{e}_1^1, \ldots, \boldsymbol{e}_n^1\}$ элемент $x_1\boldsymbol{e}_1^2 + \cdots + x_n\boldsymbol{e}_n^2$ в $\mathcal{X}_2$, имеющий такие же координаты относительно базиса $\{\boldsymbol{e}_1^2, \ldots, \boldsymbol{e}_n^2\}$, что и исходный элемент.

В силу единственности разложения элемента векторного пространства по базису (см. Предложение 2.4) построенное отображение φ является взаимно-однозначным отображением $\mathcal{X}_1$ на $\mathcal{X}_2$, а в силу Предложения 2.5 это взаимно однозначное отображение является изоморфизмом линейных пространств $\mathcal{X}_1$ и $\mathcal{X}_2$. □

2.5. Подпространства линейных пространств

Определение. *Подмножество $\mathcal{X}_0$ линейного пространства $\mathcal{X}$ над полем $\mathcal{F}$ называется подпространством пространства $\mathcal{X}$, если для любых элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_0$ имеет место $\boldsymbol{x} + \boldsymbol{y} \in \mathcal{X}_0$, а для любого числа $\alpha \in \mathcal{F}$ имеет место $\alpha\boldsymbol{x} \in \mathcal{X}_0$.*

Нетрудно проверить (это оставляется в качестве ***упражнения***), что $\mathcal{X}_0$ является линейным пространством относительно операций сложения векторов и умножения их на числа из $\mathcal{F}$, определенных в $\mathcal{X}$. Это свойство и оправдывает использование термина “подпространство” в отличии от более общего термина “подмножество”.

Пример 2.11. Простейшими примерами подпространств пространства $\mathcal{X}$ являются *нулевое подпространство* $\{\boldsymbol{0}\}$ и само пространство $\mathcal{X}$.

Упражнение. Проверить, что совокупность всех многочленов степени не выше некоторого натурального числа n с вещественными коэффициентами образует собственное подпространство пространства $C([a, b])$.

Упражнение. В линейной пространстве $\mathrm{M}_n(\mathbb{R})$ вещественных $n \times n$ матриц, $n \in \mathbb{N}$, привести примеры подмножеств являющихся и не являющихся подпространствами пространства $\mathrm{M}_n(\mathbb{R})$.

Определение. *Пусть $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ – произвольная совокупность векторов некоторого линейного пространства $\mathcal{X}$. Множество*

$$\mathrm{Span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\} := \{\alpha_1\boldsymbol{x}_1 + \cdots + \alpha_m\boldsymbol{x}_m : \ \alpha_1, \ldots, \alpha_m \in \mathcal{F}\}$$

называется линейной оболочкой векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$.

Другими словами, линейная оболочка совокупности векторов некоторого линейного пространства – это совокупность всех линейных комбинаций векторов рассматриваемой совокупности (включая, естественно, и тривиальную линейную комбинацию). Например, $\operatorname{Span}\{1, t, t^2, \dots, t^n\}$, $n \in \mathbb{N}$ – это совокупность всех многочленов с вещественными коэффициентами степени не выше n.

Предложение. *Линейная оболочка совокупности векторов некоторого линейного пространства является подпространством этого линейного пространства.*

Проверка справедливости этого утверждения оставляется в качестве *упражнения*.

Рассмотрим вопрос о размерности подпространств линейного пространства и, в частности, вопрос о размерности линейной оболочки. Можно утверждать, что размерность любого подпространства линейного пространства не превосходит размерности самого пространства. В самом деле, это вытекает из того, что любая линейно независимая система элементов подпространства $\mathcal{X}_0$ пространства $\mathcal{X}$ является и линейно независимой системой элементов в $\mathcal{X}$.

Далее, пусть $\mathcal{X}$ – линейное пространство и $\dim \mathcal{X} = n$, $n \in \mathbb{N}$. Если подпространство $\mathcal{X}_0$ пространства $\mathcal{X}$ не совпадает с $\mathcal{X}$, то $\dim \mathcal{X}_0 < \dim \mathcal{X}$ (*строго меньше*). Это вытекает из того, что если $\dim \mathcal{X}_0 = n$ и $\dim \mathcal{X} = n$, то всякий базис подпространства $\dim \mathcal{X}_0$ состоит из n элементов и является (в силу Предложения 2.7) базисом и всего пространства $\mathcal{X}$.

Пусть теперь $\mathcal{X}_0$ – подпространство линейного пространства $\mathcal{X}$ с базисом $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$. Заметим, что базисные элементы подпространства $\mathcal{X}_0$ нельзя, в общем случае, выбрать среди элементов базиса объемлющего пространства. В то же время, если элементы $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ образуют базис в $\mathcal{X}_0$ то систему элементов $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k$ можно дополнить некоторыми элементами $\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n$ пространства $\mathcal{X}$ до базиса в пространства $\mathcal{X}$. Проверим это. Пусть $k < n$. Тогда найдется такой элемент $\boldsymbol{v}_{k+1} \in \mathcal{X}$ такой, что элементы $\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}$ линейно независимы (если это не так, то $\dim \mathcal{X} = k < n$, что противоречит первоначальному предположению). Далее, если $k + 1 < n$ то найдется элемент $\boldsymbol{v}_{k+2} \in \mathcal{X}$ такой, что линейно независимой будет система элементов $\boldsymbol{v}_1 \dots, \boldsymbol{v}_{k+2}$. Продолжая эту конструкцию мы найдем n линейно независимых элементов в $\mathcal{X}$. Докажем теперь следующее важное утверждение о размерности линейной оболочки.

Теорема 2.12. *Пусть $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ – некоторая система векторов линейного пространства $\mathcal{X}$. Размерность их линейной оболочки $\operatorname{Span}\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\}$ равна максимальному количеству линейно независимых векторов в системе $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$. В частности, если $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ линейно независимы, то $\dim \operatorname{Span}\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\} = m$, а сами элементы $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ образуют базис в $\operatorname{Span}\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\}$.*

Доказательство. Предположим, что среди векторов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ имеется r линейно независимых векторов (без ограничения общности считаем, что это векторы $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r$), а любые $r+1$ векторов из $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ уже линейно зависимы. Пусть $k \in \{r+1, \dots, m\}$. Тогда найдутся числа $\alpha_1, \dots, \alpha_r$ и α_k из поля $\mathcal{F}$ такие, что

$$\alpha_1 \boldsymbol{x}_1 + \dots + \alpha_r \boldsymbol{x}_r + \alpha_k \boldsymbol{x}_k = \mathbf{0}.$$

Если $\alpha_k = 0$, то, в силу линейной независимости векторов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r$, получаем, что $\alpha_1 = \dots = \alpha_r = 0$, что невозможно, так как среди чисел $\alpha_1, \dots, \alpha_r, \alpha_k$ должны быть ненулевые. А если $\alpha_k \neq 0$, то вектор $\boldsymbol{x}_k$ представляется в виде линейной комбинации векторов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r$. Из этого вытекает, что каждая линейная комбинация элементов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_m$ может быть записана как линейная комбинация элементов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_r$. Таким образом, всякий элемент из $\operatorname{Span}\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\}$ является некоторой линейной комбинацией элементов $\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_r\}$, которые, по определению, образуют базис в $\operatorname{Span}\{\boldsymbol{x}_1, \dots, \boldsymbol{x}_m\}$. □

Замечание (о ранге матрицы)**.** Напомним, что ранг матрицы A определялся как порядок ее *базисного минора*, т.е. как число $r = \operatorname{rg} A$ такое, что у матрицы A существует отличный от нуля минор порядка r и не существует ненулевых миноров порядка, большего r. В предшествующих курсах было доказано, что ранг $\operatorname{rg} A$ матрицы A равен максимальному числу линейно независимых строк (или столбцов) матрицы A.

Сумма и пересечение подпространств. Пусть $\mathcal{X}_1$ и $\mathcal{X}_2$ – два подпространства линейного пространства $\mathcal{X}$. Обозначим через $\mathcal{X}_1 \cap \mathcal{X}_2$ пересечение множеств $\mathcal{X}_1$ и $\mathcal{X}_2$, а через $\mathcal{X}_1 + \mathcal{X}_2$ множество $\{\boldsymbol{x} + \boldsymbol{y} : \ \boldsymbol{x} \in \mathcal{X}_1, \boldsymbol{y} \in \mathcal{X}_2\}$.

Предложение 2.13. *Множества $\mathcal{X}_1 \cap \mathcal{X}_2$ и $\mathcal{X}_1 + \mathcal{X}_2$ являются подпространствами пространства $\mathcal{X}$.*

Упражнение. Доказать Предложение 2.13

Замечание. Подпространства $\mathcal{X}_1 \cap \mathcal{X}_2$ и $\mathcal{X}_1 + \mathcal{X}_2$ называются *пересечением* и *суммой* подпространств $\mathcal{X}_1$ и $\mathcal{X}_2$ соответственно.

Теорема 2.14. *Пусть $\mathcal{X}_1$ и $\mathcal{X}_2$ – два подпространства конечномерного линейного пространства $\mathcal{X}$. Тогда*

$$\dim \mathcal{X}_1 + \dim \mathcal{X}_2 = \dim(\mathcal{X}_1 \cap \mathcal{X}_2) + \dim(\mathcal{X}_1 + \mathcal{X}_2).$$

Доказательство. Пусть $k = \dim(\mathcal{X}_1 \cap \mathcal{X}_2)$ и пусть $\boldsymbol{e}_1, \dots, \boldsymbol{e}_k$ – базис в $\mathcal{X}_1 \cap \mathcal{X}_2$. Дополним этот базис до базиса $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell\}$ в пространстве $\mathcal{X}_1$ и до базиса $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ в пространстве $\mathcal{X}_2$.

Проверим, что элементы $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ являются базисом в $\mathcal{X}_1 + \mathcal{X}_2$. Первым делом проверим, что все эти элементы линейно независимы. Возьмем числа $\alpha_1, \dots, \alpha_\ell \in \mathbb{R}$, $\beta_1, \dots, \beta_k \in \mathbb{R}$ и $\gamma_1, \dots, \gamma_m \in \mathbb{R}$ и предположим, что

$$\alpha_1 \boldsymbol{v}_1 + \dots + \alpha_\ell \boldsymbol{v}_\ell + \beta_1 \boldsymbol{e}_1 + \dots + \beta_k \boldsymbol{e}_k + \gamma_1 \boldsymbol{w}_1 + \dots + \gamma_m \boldsymbol{w}_m = 0 \tag{2.3}$$

Из этого вытекает, что

$$\alpha_1 \boldsymbol{v}_1 + \dots + \alpha_\ell \boldsymbol{v}_\ell + \beta_1 \boldsymbol{e}_1 + \dots + \beta_k \boldsymbol{e}_k = -\gamma_1 \boldsymbol{w}_1 - \dots - \gamma_m \boldsymbol{w}_m.$$

Так как левая часть последнего равенства – это элемент $\mathcal{X}_1$, а правая – элемент $\mathcal{X}_2$, то и левая и правая части этого равенства принадлежат $\mathcal{X}_1 \cap \mathcal{X}_2$. Из этого, в частности, вытекает, что найдутся такие числа $\delta_1, \dots, \delta_k \in \mathbb{R}$, что

$$-\gamma_1 \boldsymbol{w}_1 - \dots - \gamma_m \boldsymbol{w}_m = \delta_1 \boldsymbol{e}_1 + \dots + \delta_k \boldsymbol{e}_k,$$

или, соответственно,

$$\delta_1 \boldsymbol{e}_1 + \dots + \delta_k \boldsymbol{e}_k + \gamma_1 \boldsymbol{w}_1 + \dots + \gamma_m \boldsymbol{w}_m = 0.$$

Так как элементы $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ линейно независимы, то $\gamma_1 = \dots = \gamma_m = \delta_1 = \dots = \delta_k = 0$. При этом, из равенства (2.3) вытекает, что

$$\alpha_1 \boldsymbol{v}_1 + \dots + \alpha_\ell \boldsymbol{v}_\ell + \beta_1 \boldsymbol{e}_1 + \dots + \beta_k \boldsymbol{e}_k = 0$$

откуда из линейной независимости элементов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell\}$ следует, что $\alpha_1 = \dots = \alpha_\ell = \beta_1 = \dots = \beta_k = 0$.

Итак, система $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ линейно независима. Так как любой элемент $\boldsymbol{x} \in \mathcal{X}_1 + \mathcal{X}_2$ имеет вид $\boldsymbol{x}_1 + \boldsymbol{x}_2$, где $\boldsymbol{x}_1 \in \mathcal{X}_1$ и $\boldsymbol{x}_2 \in \mathcal{X}_2$ и так как $\boldsymbol{x}_1$ представляется в виде линейной комбинации элементов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell\}$, а элемент $\boldsymbol{x}_2$ – в виде линейной комбинации элементов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$, то элемент $\boldsymbol{x}$ представляется в виде линейной комбинации элементов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$.

Таким образом, нами установлено, что система $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k, \boldsymbol{v}_1, \dots, \boldsymbol{v}_\ell, \boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$ образует базис в $\mathcal{X}_1 + \mathcal{X}_2$ и из этого вытекает справедливость утверждения теоремы. □

Замечание. Пусть $E \subset \mathcal{X}$ – некоторое подмножество линейного пространства $\mathcal{X}$. *Линейную оболочку* $\operatorname{Span}(E)$ множества E определим как наименьшее подпространство пространства $\mathcal{X}$, содержащее E. В качестве ***упражнения*** предлагается показать, что $\operatorname{Span}(E) = \bigcap \widetilde{\mathcal{X}}$, где пересечение берется по всем подпространствам $\widetilde{\mathcal{X}}$ пространства $\mathcal{X}$ таким, что $E \subset \widetilde{\mathcal{X}}$. Используя понятие линейной оболочки множества можно заметить, что $\mathcal{X}_1 + \mathcal{X}_2 = \operatorname{Span}(\mathcal{X}_1 \cup \mathcal{X}_2)$ (пересечение рассматривается в смысле теории множеств). Проверка этого утверждения также оставляется в качестве *упражнения*.

Разложение пространства в прямую сумму подпространств. Пусть $\mathcal{X}_1$ и $\mathcal{X}_2$ – два подпространства конечномерного линейного пространства $\mathcal{X}$, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$.

Определение. *Скажем, что пространство $\mathcal{X}$ представляет собой прямую сумму подпространств $\mathcal{X}_1$ и $\mathcal{X}_2$, если каждый элемент $\boldsymbol{x} \in \mathcal{X}$ может быть единственным образом представлен в виде суммы $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, где $\boldsymbol{x}_1 \in \mathcal{X}_1$, а $\boldsymbol{x}_2 \in \mathcal{X}_2$.*

Тот факт, что $\mathcal{X}$ представляет собой прямую сумму подпространств $\mathcal{X}_1$ и $\mathcal{X}_2$ записывается $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$. Равенство $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$ называют разложением пространства $\mathcal{X}$ в прямую сумму подпространств $\mathcal{X}_1$ и $\mathcal{X}_2$.

Теорема 2.15. *Пусть $\mathcal{X}_1$ и $\mathcal{X}_2$ – подпространства n-мерного пространства $\mathcal{X}$. Предположим, что (1) $\mathcal{X}_1 \cap \mathcal{X}_2 = \{\boldsymbol{0}\}$ и (2) $\dim \mathcal{X}_1 + \dim \mathcal{X}_2 = \dim \mathcal{X}$. Тогда $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$.*

Доказательство. Пусть векторы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ образуют базис в пространстве $\mathcal{X}_1$, а векторы $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m$ – базис в $\mathcal{X}_2$. Проверим, что система $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ образует базис во все пространстве $\mathcal{X}$. Так как $k + m = \dim \mathcal{X}$, то, в силу Предложения 2.7, нам достаточно установить линейную независимость системы $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$.

Возьмем произвольные числа $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$ и $\beta_1, \ldots, \beta_m \in \mathbb{R}$ и предположим, что

$$\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_k \boldsymbol{e}_k + \beta_1 \boldsymbol{w}_1 + \cdots + \beta_m \boldsymbol{w}_m = \boldsymbol{0}.$$

Тогда

$$\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_k \boldsymbol{e}_k = -\beta_1 \boldsymbol{w}_1 - \cdots - \beta_m \boldsymbol{w}_m,$$

а так как левая и правая части этого равенства представляют собой элементы из $\mathcal{X}_1$ и $\mathcal{X}_2$ соответственно и так как пересечение $\mathcal{X}_1$ и $\mathcal{X}_2$ состоит только из нулевого элемента, то $\alpha_1 \boldsymbol{e}_1 + \cdots + \alpha_k \boldsymbol{e}_k = \boldsymbol{0}$ и $\beta_1 \boldsymbol{w}_1 + \cdots + \beta_m \boldsymbol{w}_m = \boldsymbol{0}$. Так как системы $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k\}$ и $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ линейно независимы, то $\alpha_1 = \cdots = \alpha_k = \beta_1 = \cdots = \beta_m = 0$.

Мы доказали, что система $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k, \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m\}$ является базисом в $\mathcal{X}$. Пусть теперь $\boldsymbol{x} \in \mathcal{X}$ – произвольный элемент. Разложим его по найденному базису и получим

$$\boldsymbol{x} = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_k \boldsymbol{e}_k + \mu_1 \boldsymbol{w}_1 + \cdots + \mu_m \boldsymbol{w}_m,$$

т.е. $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$, где $\boldsymbol{x}_1 = \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_k \boldsymbol{e}_k \in \mathcal{X}_1$, а $\boldsymbol{x}_2 = \mu_1 \boldsymbol{w}_1 + \cdots + \mu_m \boldsymbol{w}_m \in \mathcal{X}_2$. Остается показать, что разложение элемента $\boldsymbol{x}$ в сумму $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ единственно.

В самом деле, пусть имеются два таких разложения: $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2 = \boldsymbol{x}_1' + \boldsymbol{x}_2'$. Отсюда получаем, что $\boldsymbol{x}_1' - \boldsymbol{x}_1 = \boldsymbol{x}_2' - \boldsymbol{x}_2$ и, так как $\boldsymbol{x}_1' - \boldsymbol{x}_1 \in \mathcal{X}_1$, а $\boldsymbol{x}_2' - \boldsymbol{x}_2 \in \mathcal{X}_2$ и, так как пересечение $\mathcal{X}_1$ и $\mathcal{X}_2$ содержит только нулевой элемент, то $\boldsymbol{x}_1' - \boldsymbol{x}_1 = \boldsymbol{0}$ и $\boldsymbol{x}_2' - \boldsymbol{x}_2 = \boldsymbol{0}$, т.е. $\boldsymbol{x}_1 = \boldsymbol{x}_1'$ и $\boldsymbol{x}_2 = \boldsymbol{x}_2'$. □

Замечание. Если пространство $\mathcal{X}$ представляет собой не прямую, а обычную сумму подпространств $\mathcal{X}_1$ и $\mathcal{X}_2$, то представление элемента $\boldsymbol{x} \in \mathcal{X}$ в виде суммы $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ элементов $\boldsymbol{x}_1 \in \mathcal{X}_1$ и $\boldsymbol{x}_2 \in \mathcal{X}_2$ не является единственным. В самом деле, если $\mathcal{X}$ – это не прямая сумма подпространств $\mathcal{X}_1$ и $\mathcal{X}_2$, то найдется такой элемент $\boldsymbol{x}_0 \in \mathcal{X}$, что $\boldsymbol{x}_0 = \boldsymbol{x}_{0,1} + \boldsymbol{x}_{0,2}$ и $\boldsymbol{x}_0 = \widetilde{\boldsymbol{x}}_{0,1} + \widetilde{\boldsymbol{x}}_{0,2}$, где $\boldsymbol{x}_{0,1} \neq \widetilde{\boldsymbol{x}}_{0,1} \in \mathcal{X}_1$ и $\boldsymbol{x}_{0,2} \neq \widetilde{\boldsymbol{x}}_{0,2} \in \mathcal{X}_2$. Тогда $\boldsymbol{0} = (\boldsymbol{x}_{0,1} - \widetilde{\boldsymbol{x}}_{0,1}) + (\boldsymbol{x}_{0,2} - \widetilde{\boldsymbol{x}}_{0,2})$. А так как нулевой вектор имеет неоднозначное представление в виде суммы векторов из $\mathcal{X}_1$ и $\mathcal{X}_2$, то и любой другой вектор из $\mathcal{X}$ также будет иметь неоднозначное представление в таком виде.

Пример. Пусть линейные подпространства $\mathcal{X}$ и $\mathcal{Y}$ заданы как линейные оболочки систем векторов $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k\}$ и $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell\}$. Предположим также, что векторы $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_k$ и $\boldsymbol{b}_1, \ldots, \boldsymbol{b}_\ell$ заданы их координатными столбцами, которые образуют матрицы A и B соответственно. Требуется найти размерность и базис суммы $\mathcal{X} + \mathcal{Y}$ и пересечения $\mathcal{X} \cap \mathcal{Y}$ эти подпространств.

Напомним, что матрица M размера $m \times n$ имеет *упрощенный вид*, если, во-первых, некоторые r ее столбцов ($r \geqslant 0$) являются первыми r столбцами единичной $m \times m$

матрицы и, во-вторых, при $r < m$ последние $m - r$ ее строк нулевые. Ранг матрицы, имеющей упрощенный вид, равен r.

Объединим матрицы A и B в матрицу $\|A|B\|$, первые k столбцов которой совпадают со столбцами матрицы A, а оставшиеся ℓ столбцов совпадают со столбцами матрицы B. Приведем матрицу $\|A|B\|$ к упрощенному виду $\|A'|B'\|$ при помощи элементарных преобразований *строк*. Тогда векторы, соответствующие базисным столбцам матрицы $\|A'|B'\|$, составляют базис $\mathcal{X} + \mathcal{Y}$, а соответствующие базисным столбцам, расположенным в A', составят базис в $\mathcal{X}$. Для того, чтобы найти размерности $\mathcal{Y}$ и $\mathcal{X} \cap \mathcal{Y}$, упростим теперь матрицу B' с помощью элементарных преобразований *столбцов* так, чтобы не менять ранее полученных базисных столбцов (назовем эти столбцы *основными*). В результате этого преобразования матрицы B' получим матрицу B''. Полный набор базисных столбцов матрицы B'' соответствует базису в $\mathcal{Y}$, а базисные столбцы B'', не являющиеся основными, соответствуют базису в $\mathcal{X} \cap \mathcal{Y}$. Обоснование этого метода и его проверка на конкретных числовых примерах оставляется в качестве *упражнения*.

Задача 2.1. Доказать, что сумма подпространств $\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2 + \cdots + \mathcal{X}_k$ является прямой если и только если для всех $j = 1, \ldots, k$ выполнено соотношение

$$\mathcal{X}_j \cap \big(\mathcal{X}_1 + \cdots + \mathcal{X}_{j-1} + \mathcal{X}_{j+1} + \cdots + \mathcal{X}_k\big) = \{\mathbf{0}\}.$$

Задача 2.2. Доказать, что сумма подпространств $\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2 + \cdots + \mathcal{X}_k$ является прямой если и только если $\dim \mathcal{X} = \sum_{j=1}^{k} \dim \mathcal{X}_j$.

Внешние и внутренние прямые суммы. Понятие прямой суммы было введено выше для подпространств некоторого объемлющего линейного пространства $\mathcal{X}$. Такие прямые суммы часто называют *внутренними* прямыми суммами.

Введем теперь понятие прямой суммы для произвольных линейных пространств $\mathcal{Y}_1$ и $\mathcal{Y}_2$ (рассматриваемых, разумеется, над одним и тем же полем $\mathcal{F}$). Определим *внешнюю прямую сумму* $\mathcal{Y} = \mathcal{Y}_1 \oplus \mathcal{Y}_2$ как множество $\mathcal{Y}_1 \times \mathcal{Y}_2 = \{(\boldsymbol{y}_1, \boldsymbol{y}_2) : \ \boldsymbol{y}_1 \in \mathcal{Y}_1, \ \boldsymbol{y}_2 \in \mathcal{Y}_2\}$, для элементов которого операции сложения элементов и их умножения на числа из $\mathcal{F}$ определены при помощи соотношения

$$\alpha(\boldsymbol{y}_1', \boldsymbol{y}_2') + \beta(\boldsymbol{y}_1'', \boldsymbol{y}_2'') = (\alpha \boldsymbol{y}_1' + \beta \boldsymbol{y}_1'', \alpha \boldsymbol{y}_2' + \beta \boldsymbol{y}_2'').$$

Множество $\widetilde{\mathcal{Y}}_1 := \{(\boldsymbol{y}_1, 0) : \ \boldsymbol{y}_1 \in \mathcal{Y}_1\} \subset \mathcal{Y}$ – это линейной подпространство пространства $\mathcal{Y}$, изоморфное пространству $\mathcal{Y}_1$ (соответствующий изоморфизм имеет вид $(\boldsymbol{y}_1, 0) \mapsto \boldsymbol{y}_1$). Аналогично, множество $\widetilde{\mathcal{Y}}_2 := \{(0, \boldsymbol{y}_2) : \ \boldsymbol{y}_2 \in \mathcal{Y}_2\} \subset \mathcal{Y}$ – это линейной подпространство пространства $\mathcal{Y}$, изоморфное пространству $\mathcal{Y}_2$.

Заметим, что имеет место соотношение

$$\mathcal{Y}_1 \oplus \mathcal{Y}_2 = \mathcal{Y} = \widetilde{\mathcal{Y}}_1 \oplus \widetilde{\mathcal{Y}}_2,$$

где первая прямая сумма внешняя, а вторая – внутренняя. Практически всюду в дальнейшем рассматриваемые прямые суммы будут внутренними и термин прямая сумма будет использоваться без дополнительных комментариев.

2.6. Факторпространства линейных пространств

Пусть дано некоторое линейное пространство $\mathcal{X}$ над полем $\mathcal{F}$ и его подпространство $\mathcal{U}$. Введем на $\mathcal{X}$ отношение $\sim_{\mathcal{U}}$ так, что $\boldsymbol{x} \sim_{\mathcal{U}} \boldsymbol{y}$ если и только если $\boldsymbol{x} - \boldsymbol{y} \in \mathcal{U}$. Нетрудно показать, что $\sim_{\mathcal{U}}$ является отношением эквивалентности на $\mathcal{X}$.

Рассмотрим фактормножество $\mathcal{X}/\sim_{\mathcal{U}}$. Обозначим классы эквивалентности по этому отношению через $\widetilde{\boldsymbol{x}}_{\mathcal{U}}$. При этом

$$\widetilde{\boldsymbol{x}}_{\mathcal{U}} = \{\boldsymbol{y} \in \mathcal{X} : \ \boldsymbol{y} \sim_{\mathcal{U}} \boldsymbol{x}\} = \{\boldsymbol{y} \in \mathcal{X} : \ y - x \in \mathcal{U}\} = \{\boldsymbol{y} \in \mathcal{X} : \ \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{u}, \ \boldsymbol{u} \in \mathcal{U}\}.$$

Множество $\boldsymbol{x} + \mathcal{U} = \{\boldsymbol{x} + \boldsymbol{u} : \ \boldsymbol{u} \in \mathcal{U}\}$ называется *смежным классом* пространства $\mathcal{X}$ по подпространству $\mathcal{U}$, вектор $\boldsymbol{x}$ называется *представителем* соответствующего смежного класса. Из Предложения 1.5 вытекает, что любые два смежных класса пространства $\mathcal{X}$ по подпространству $\mathcal{U}$ либо совпадают, либо не пересекаются.

Множество $\mathcal{X}/\sim_{\mathcal{U}}$ традиционно обозначают символом $\mathcal{X}/\mathcal{U}$. Оказывается, что на множестве $\mathcal{X}/\mathcal{U}$ можно ввести структуру линейного пространства над полем $\mathcal{F}$. В самом деле, если определить операции сложения элементов множества $\mathcal{X}/\mathcal{U}$ и умножения элементов из $\mathcal{X}/\mathcal{U}$ на числа $\lambda \in \mathcal{F}$ по правилам

$$\widetilde{\boldsymbol{x}}_{\mathcal{U}} + \widetilde{\boldsymbol{y}}_{\mathcal{U}} = \widetilde{\boldsymbol{z}}_{\mathcal{U}}, \quad \boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}, \qquad \lambda \widetilde{\boldsymbol{x}}_{\mathcal{U}} = \widetilde{(\lambda \boldsymbol{x})}_{\mathcal{U}},$$

то свойства (3)-(9) из определения линейного пространства проверяются непосредственно (это оставляется в качестве *упражнения*).

Определение. *Пространство $\mathcal{X}/\mathcal{U}$ называется факторпространством пространства $\mathcal{X}$ по подпространству $\mathcal{U}$.*

Рассмотрим следующий простой пример. Пусть $\mathcal{X} = \mathbb{R}^2$, а $\mathcal{U} = \{(x, 0) : \ x \in \mathbb{R}\} \subset \mathbb{R}^2$. Определим также подпространство $\mathcal{W}_a = \{(x, ax) : \ x \in \mathbb{R}\} \subset \mathbb{R}^2$ при вещественном $a \neq 0$. Тогда для любого $a \neq 0$ верно равенство $\mathcal{U} \oplus \mathcal{W}_a = \mathcal{X}$. Заметим, что множество $\mathcal{X}/\mathcal{U}$ в этом случае – это множество всех прямых, параллельных горизонтальной оси.

Установим, что справедливо следующее утверждение.

Предложение 2.16. *Пусть $\mathcal{X}$ – линейное пространство, $\mathcal{U}$ и $\mathcal{W}$ – такие его подпространства, что $\mathcal{X} = \mathcal{U} \oplus \mathcal{W}$. Тогда отображение $f : \boldsymbol{w} \mapsto \widetilde{\boldsymbol{w}}_{\mathcal{U}}$, $\boldsymbol{w} \in \mathcal{W}$, является изоморфизмом между $\mathcal{W}$ и $\mathcal{X}/\mathcal{U}$.*

Доказательство. Проверим, что f – биективное отображение. Так как произвольный элемент $\boldsymbol{x} \in \mathcal{X}$ можно (единственным образом) записать в виде $\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{w}$, где $\boldsymbol{u} \in \mathcal{U}$, а $\boldsymbol{w} \in \mathcal{W}$, то

$$\widetilde{\boldsymbol{x}}_{\mathcal{U}} = \widetilde{\boldsymbol{u}}_{\mathcal{U}} + \widetilde{\boldsymbol{w}}_{\mathcal{U}} = \mathcal{U} + \widetilde{\boldsymbol{w}}_{\mathcal{U}} = \widetilde{\boldsymbol{w}}_{\mathcal{U}} = f(\boldsymbol{w}),$$

т.е. отображение f сюръективно. Инъективность f (и, следовательно, его биективность) вытекает из следующих соображений. Пусть $\boldsymbol{w}_1 \in \mathcal{W}$ и $\boldsymbol{w}_2 \in \mathcal{W}$ таковы, что $f(\boldsymbol{w}_1) = f(\boldsymbol{w}_2)$. Тогда $\widetilde{\boldsymbol{w}}_{1\mathcal{U}} = \widetilde{\boldsymbol{w}}_{2\mathcal{U}}$. Тогда $\boldsymbol{w}_1 - \boldsymbol{w}_2 \in \mathcal{U}$ и, так как $\mathcal{U} \cap \mathcal{W} = \{\boldsymbol{0}\}$, то $\boldsymbol{w}_1 - \boldsymbol{w}_2 = \boldsymbol{0}$.

Остается заметить, что для любых $\boldsymbol{w}_1, \boldsymbol{w}_2 \in \mathcal{W}$ и $\alpha_1, \alpha_2 \in \mathcal{F}$ верны равенства

$$f(\alpha_1 \boldsymbol{w}_1 + \alpha_2 \boldsymbol{w}_2) = \alpha_1 \boldsymbol{w}_1 + \alpha_2 \boldsymbol{w}_2 + \mathcal{U} = (\alpha_1 \boldsymbol{w}_1 + \mathcal{U}) + (\alpha_2 \boldsymbol{w}_2 + \mathcal{U}) = \alpha_1 f(\boldsymbol{w}_1) + \alpha_2 f(\boldsymbol{w}_2).$$

из которых и вытекает, что f – изоморфизм $\mathcal{W}$ и $\mathcal{X}/\mathcal{U}$. □

В качестве следствия только что доказанного утверждения можно отметить, что для любого подпространства $\mathcal{U} \subset \mathcal{X}$ верно равенство $\dim \mathcal{X}/\mathcal{U} = \dim \mathcal{X} - \dim \mathcal{U}$.

2.7. Преобразование координат при преобразовании базиса конечномерного линейного пространства

Преобразование базисов. Пусть $\mathcal{X}$ – линейное пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$, и пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ – два базиса в $\mathcal{X}$. При $k = 1, \ldots, n$ разложим элемент $\boldsymbol{e}'_k \in \mathcal{X}$ по базису $\boldsymbol{e}$. При этом получим следующие соотношения ($k = 1, \ldots, n$):

$$\begin{cases} \boldsymbol{e}'_1 = \sigma_{11}\boldsymbol{e}_1 + \sigma_{21}\boldsymbol{e}_2 + \cdots + \sigma_{n1}\boldsymbol{e}_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \boldsymbol{e}'_k = \sigma_{1k}\boldsymbol{e}_1 + \sigma_{2k}\boldsymbol{e}_2 + \cdots + \sigma_{nk}\boldsymbol{e}_n \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \boldsymbol{e}'_n = \sigma_{1n}\boldsymbol{e}_1 + \sigma_{2n}\boldsymbol{e}_2 + \cdots + \sigma_{nn}\boldsymbol{e}_n \end{cases} \tag{2.4}$$

В матричном виде эти равенства записываются следующим образом

$$\boldsymbol{e}' = \boldsymbol{e}S,$$

где S – это $n \times n$-матрица, столбцами которой являются координатные столбцы векторов $\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n$ базиса $\boldsymbol{e}'$ в базисе $\boldsymbol{e}$:

$$S = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \multicolumn{3}{c}{\dots\dots\dots\dots} \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix}.$$

Определение. *Матрица S называется матрицей перехода от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$. Говорят также, что переход от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$ задается матрицей S.*

Заметим, что определитель $\det S$ матрицы S отличен от нуля. В самом деле, если $\det S = 0$, то это означает линейную зависимость столбцов матрицы S и, следовательно, линейную зависимость элементов $\{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$ (что невозможно, так как эти элементы образуют базис в $\mathcal{X}$).

Проверим, что обратный переход от базиса $\boldsymbol{e}'$ к базису $\boldsymbol{e}$ задается при помощи матрицы S^{-1}, обратной к матрице S. Напомним, что $S^{-1} = \|\tau_{jk}\|_{j,k=1\dots n}$, где $\tau_{jk} = S_{kj}/\det S$, а через S_{jk} обозначено алгебраическое дополнение элемента σ_{jk} матрицы S. Из уравнений (2.4) вытекает, что при $m = 1, \dots, n$ имеют место соотношения

$$\sum_{k=1}^{n} \tau_{km} \boldsymbol{e}'_k = \sum_{k=1}^{n} \frac{S_{mk}}{\det S} \sum_{j=1}^{n} \sigma_{jk} \boldsymbol{e}_j = \frac{1}{\det S} \sum_{j=1}^{n} \boldsymbol{e}_j \sum_{k=1}^{n} \sigma_{jk} S_{mk},$$

В силу известных свойств алгебраических дополнений, выражение $\sum_{k=1}^{n} \sigma_{jk} S_{mk}$ равно $\det S$ при $j = m$ и обращается в нуль при $j \neq m$, то $\boldsymbol{e}_m = \sum_{k=1}^{n} \tau_{km} \boldsymbol{e}'_k$, откуда и следует требуемое утверждение.

Замечание. Формулу перехода от базиса $\boldsymbol{e}'$ к базису $\boldsymbol{e}$ можно получить намного проще. В самом деле, так как $\boldsymbol{e}' = \boldsymbol{e}S$, то $\boldsymbol{e}'S^{-1} = \boldsymbol{e}SS^{-1} = \boldsymbol{e}E = \boldsymbol{e}$, где E – это единичная $n \times n$-матрица, т.е. $\boldsymbol{e} = \boldsymbol{e}'S^{-1}$.

Преобразование координат. Выясним теперь, как меняются координаты элемента $\boldsymbol{x} \in \mathcal{X}$ при переходе от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$. Предположим, что такой переход осуществляется при помощи матрицы перехода S. Первым делом установим, что

$$[\boldsymbol{x}]_{\boldsymbol{e}} = S[\boldsymbol{x}]_{\boldsymbol{e}'}.$$

В самом деле, пусть $[\boldsymbol{x}]_{\boldsymbol{e}} = (x_1, \dots, x_n)^\top$, а $[\boldsymbol{x}]_{\boldsymbol{e}'} = (x'_1, \dots, x'_n)^\top$. Тогда

$$\boldsymbol{x} = \sum_{k=1}^{n} x'_k \boldsymbol{e}'_k = \sum_{k=1}^{n} x'_k \sum_{j=1}^{n} \sigma_{jk} \boldsymbol{e}_k = \sum_{j=1}^{n} \boldsymbol{e}_j \sum_{k=1}^{n} \sigma_{jk} x'_k.$$

Так как матрица перехода S является невырожденной, то существует обратная матрица и из соотношения $[\boldsymbol{x}]_{\boldsymbol{e}} = S[\boldsymbol{x}]_{\boldsymbol{e}'}$ немедленно вытекает, что $[\boldsymbol{x}]_{\boldsymbol{e}'} = S^{-1}[\boldsymbol{x}]_{\boldsymbol{e}}$.

Собирая вместе полученные в этом разделе утверждения мы может сформулировать следующий результат

Теорема 2.17. *Пусть $\mathbf{e}$ и $\mathbf{e}'$ – два базиса в n-мерном линейном пространстве $\mathcal{X}$ и пусть S – соответствующая матрица перехода, т.е. $\mathbf{e}' = \mathbf{e}S$. Тогда матрица S невырождена, $\mathbf{e} = \mathbf{e}'S^{-1}$, а для любого вектора $\boldsymbol{x} \in \mathcal{X}$ верны равенства $[\boldsymbol{x}]_{\boldsymbol{e}} = S[\boldsymbol{x}]_{\boldsymbol{e}'}$ и $[\boldsymbol{x}]_{\boldsymbol{e}'} = S^{-1}[\boldsymbol{x}]_{\boldsymbol{e}}$.*

РАЗДЕЛ 3

Линейные операторы, линейные функционалы, полуторалинейные и билинейные функции

Все линейные пространства в этом разделе, если иное не оговорено специально, считаются *комплексными.*

3.1. Определение и основные свойства линейных операторов

Пусть $\mathcal{X}$ и $\mathcal{Y}$ – два линейных пространства и пусть $\dim \mathcal{X} = n$ и $\dim \mathcal{Y} = m$. Отображение $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$, сопоставляющее каждому элементу $\boldsymbol{x} \in \mathcal{X}$ некоторый элемент $\boldsymbol{y} \in \mathcal{Y}$, называется *оператором, действующим из* $\mathcal{X}$ *в* $\mathcal{Y}$. В дальнейшем мы будем использовать обозначения $\boldsymbol{y} = \mathcal{A}(\boldsymbol{x})$.

Определение. *Оператор* $\mathcal{A}$*, действующий из линейного пространства* $\mathcal{X}$ *в линейное пространство* $\mathcal{Y}$ *называется линейным, если для любых элементов* $\boldsymbol{x}_1 \in \mathcal{X}$ *и* $\boldsymbol{x}_2 \in \mathcal{X}$ *и для любого числа* $\alpha \in \mathbb{C}$ *выполнены следующие соотношения*

(1) $\mathcal{A}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \mathcal{A}(\boldsymbol{x}_1) + \mathcal{A}(\boldsymbol{x}_2)$;
(2) $\mathcal{A}(\alpha \boldsymbol{x}_1) = \alpha\, \mathcal{A}(\boldsymbol{x}_1)$.

Для линейного оператора $\mathcal{A}$ *кроме обозначения* $\mathcal{A}(\boldsymbol{x})$ *применяется также обозначение* $\mathcal{A}\,\boldsymbol{x}$*. В случае, когда* $\mathcal{Y} = \mathbb{C}$ *или* $\mathcal{Y} = \mathbb{R}$*, оператор* $\mathcal{A}$ *называют линейной формой или линейным функционалом на* $\mathcal{X}$*. В случае, когда* $\mathcal{Y} = \mathcal{X}$*, оператор* $\mathcal{A}$ *называют линейным преобразованием пространства* $\mathcal{X}$*.*

Из определения линейного оператора непосредственно вытекает, что $\mathcal{A}\,\mathbf{0} = \mathcal{A}(\mathbf{0} + \mathbf{0}) = \mathcal{A}\,\mathbf{0} + \mathcal{A}\,\mathbf{0}$, откуда $\mathcal{A}\,\mathbf{0} = \mathbf{0}$, т.е. образом нулевого вектора является нулевой вектор.

Нам потребуется определить еще одно понятие.

Определение. *Два линейных оператора* $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ *и* $\mathcal{B} : \mathcal{X} \to \mathcal{Y}$ *равны, если для любого* $\boldsymbol{x} \in \mathcal{X}$ *верно равенство* $\mathcal{A}\,\boldsymbol{x} = \mathcal{B}\,\boldsymbol{x}$ *(т.е., если образы любого вектора относительно этих операторов совпадают).*

Далее, *суммой* двух линейных операторов $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ и $\mathcal{B} : \mathcal{X} \to \mathcal{Y}$ назовем линейный оператор $\mathcal{A} + \mathcal{B}$, определенный следующим образом

$$(\mathcal{A} + \mathcal{B})(\boldsymbol{x}) = \mathcal{A}(\boldsymbol{x}) + \mathcal{B}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Кроме того, *произведением* линейного оператора $\mathcal{A} : \mathcal{X} \to \mathcal{Y}$ на число $\alpha \in \mathbb{C}$ назовем линейный оператор $\alpha\,\mathcal{A}$, определенный соотношением

$$(\alpha\,\mathcal{A})(\boldsymbol{x}) = \alpha\,\mathcal{A}(\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Нулевым оператором назовем такой оператор $\mathcal{O} : \mathcal{X} \to \mathcal{Y}$, что для любого $\boldsymbol{x} \in \mathcal{X}$ имеет место равенство $\mathcal{O}(\boldsymbol{x}) = \mathbf{0} \in \mathcal{Y}$.

Предложение 3.1. *Совокупность* $\mathfrak{L}(\mathcal{X}, \mathcal{Y})$ *всех линейных операторов, действующих из* $\mathcal{X}$ *в* $\mathcal{Y}$ *с введенными выше операциями сложения и умножения на комплексные числа является (комплексным) линейным пространством.*

Упражнение. Доказать Предложение 3.1.

Рассмотрим более подробно *свойства множества* $\mathfrak{L}(\mathcal{X}, \mathcal{X})$ линейных преобразований некоторого линейного пространства $\mathcal{X}$.

Произведением двух операторов $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ и $\mathcal{B} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ называется оператор $\mathcal{A}\,\mathcal{B} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$, определенный следующим соотношением

$$(\mathcal{A}\,\mathcal{B})(\boldsymbol{x}) = \mathcal{A}(\mathcal{B}(\boldsymbol{x})), \quad \forall \boldsymbol{x} \in \mathcal{X}.$$

Другими словами, оператор $\mathcal{A}\,\mathcal{B}$ как отображение является композицией отображений, осуществляемых операторами $\mathcal{A}$ и $\mathcal{B}$.

Замечание. Пусть даны три линейных пространства $\mathcal{X}$, $\mathcal{Y}$ и $\mathcal{Z}$. При помощи соотношения $(\mathcal{A}\,\mathcal{B})(\boldsymbol{x}) = \mathcal{A}(\mathcal{B}(\boldsymbol{x}))$ произведение $\mathcal{A}\,\mathcal{B}$ операторов $\mathcal{A}$ и $\mathcal{B}$ можно определить для операторов $\mathcal{A} \in \mathfrak{L}(\mathcal{Y},\mathcal{Z})$ и $\mathcal{B} \in \mathfrak{L}(\mathcal{X},\mathcal{Y})$.

Пример 3.2. Напомним, что $\mathbb{R}[t]_n$ – пространство многочленов с вещественными коэффициентами степени не выше n. Определим оператор $\mathcal{A}_1 : \mathbb{R}[t]_n \to \mathbb{R}^{n-1}$ следующим образом:

$$\mathcal{A}_1 : p_n t^n + \cdots + p_1 t + p_0 \mapsto (p_1, p_2, \ldots, p_n).$$

Далее, определим оператор $\mathcal{A}_2 : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ так, что

$$\mathcal{A}_2 : (x_1, x_2, \ldots, x_n) \mapsto (x_1, 2x_2, \ldots, nx_n),$$

а оператор $\mathcal{A}_3 : \mathbb{R}^{n-1} \to \mathbb{R}[t]_{n-1}$ так, что

$$\mathcal{A}_3 : (a_1, \ldots, a_n) \mapsto a_n t^{n-1} + \cdots + a_2 t + a_1.$$

В качестве ***упражнения*** предлагается проверить, что операторы $\mathcal{A}_1$, $\mathcal{A}_2$ и $\mathcal{A}_3$ являются линейными операторами. Заметим, что их произведение $\mathcal{A} = \mathcal{A}_3\,\mathcal{A}_2\,\mathcal{A}_1$ – это оператор, действующий из пространства $\mathbb{R}[t]_n$ в пространство $\mathbb{R}[t]_{n-1}$ следующим образом

$$\mathcal{A} : p_n t^n + \cdots + p_1 t + p_0 \mapsto n p_n t^{n-1} + \cdots + 2p_2 t + p_1,$$

т.е. $\mathcal{A}$ – это оператор дифференцирования d/dt.

Пусть снова $\mathcal{A}, \mathcal{B} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$.

Замечание. В общем случае $\mathcal{A}\,\mathcal{B} \neq \mathcal{B}\,\mathcal{A}$. В качестве ***упражнения*** предлагается привести пример такой ситуации.

Оператор $\mathcal{E} : \mathcal{X} \to \mathcal{X}$, действующий по правилу $\mathcal{E}(\boldsymbol{x}) = \boldsymbol{x}$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$ будем называть *тождественным* или *единичным*.

Непосредственно из определения произведения операторов вытекает, что для любого $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ и для любого $\boldsymbol{x} \in \mathcal{X}$ верны равенства

$$(\mathcal{A}\,\mathcal{E})(\boldsymbol{x}) = \mathcal{A}(\mathcal{E}(\boldsymbol{x})) = \mathcal{A}(\boldsymbol{x}), \qquad (\mathcal{E}\,\mathcal{A})(\boldsymbol{x}) = \mathcal{E}(\mathcal{A}(\boldsymbol{x})) = \mathcal{A}(\boldsymbol{x}),$$

откуда вытекает, что $\mathcal{A}\,\mathcal{E} = \mathcal{E}\,\mathcal{A} = \mathcal{A}$. Другими словами, тождественный оператор является двусторонним нейтральным элементов по умножению в $\mathfrak{L}(\mathcal{X},\mathcal{X})$.

Предложение 3.3. *Операция умножения линейных операторов удовлетворяет следующим свойствам:*

$$\alpha(\mathcal{A}\,\mathcal{B}) = (\alpha\,\mathcal{A})\,\mathcal{B};$$
$$(\mathcal{A}_1 + \mathcal{A}_2)\,\mathcal{B} = \mathcal{A}_1\,\mathcal{B} + \mathcal{A}_2\,\mathcal{B};$$
$$\mathcal{A}(\mathcal{B}_1 + \mathcal{B}_2) = \mathcal{A}\,\mathcal{B}_1 + \mathcal{A}\,\mathcal{B}_2;$$
$$(\mathcal{A}_1\,\mathcal{A}_2)\,\mathcal{A}_3 = \mathcal{A}_1(\mathcal{A}_2\,\mathcal{A}_3)$$

которые выполняются для любых операторов $\mathcal{A}, \mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3, \mathcal{B}, \mathcal{B}_1, \mathcal{B}_2 \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ *и для любого числа* $\alpha \in \mathbb{C}$. *При этом последнее свойство выражает ассоциативность операции умножения операторов, а два предпоследних свойства – дистрибутивность операции умножения операторов относительно сложения.*

Доказательство. Для любого $\boldsymbol{x} \in \mathcal{X}$ выполняются равенства

$$(\alpha(\mathcal{A}\,\mathcal{B}))(\boldsymbol{x}) = \alpha((\mathcal{A}\,\mathcal{B})(\boldsymbol{x})) = \alpha\,\mathcal{A}(\mathcal{B}(x)) = (\alpha\,\mathcal{A})(\mathcal{B}(\boldsymbol{x}));$$

$$((\mathcal{A}_1+\mathcal{A}_2)\,\mathcal{B})(\boldsymbol{x}) = (\mathcal{A}_1+\mathcal{A}_2)(\mathcal{B}(\boldsymbol{x})) = \mathcal{A}_1(\mathcal{B}(\boldsymbol{x})) + \mathcal{A}_2(\mathcal{B}(\boldsymbol{x})) = (\mathcal{A}_1\,\mathcal{B})(\boldsymbol{x}) + (\mathcal{A}_2\,\mathcal{B})(\boldsymbol{x});$$

чем доказываются первое и второе свойство. Третье свойство проверяется аналогично второму. Четвертое свойство является частным случаем ассоциативности операции композиции отображений. □

Для любого оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ и для любого натурального числа k определим *степень* $\mathcal{A}^k := \mathcal{A}\,\mathcal{A}\cdots\mathcal{A}$ (всего k сомножителей) оператора $\mathcal{A}$. Из свойства ассоциативности умножения операторов вытекает, что степень оператора обладает тем свойством, что для любых двух натуральный чисел k и s верно равенство $\mathcal{A}^{k+s} = \mathcal{A}^k\,\mathcal{A}^s$.

Введем теперь понятие обратного оператора.

Определение. *Оператор $\mathcal{B} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ является обратным оператором для оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$, если выполняются следующие равенства*

$$\mathcal{A}\,\mathcal{B} = \mathcal{B}\,\mathcal{A} = \mathcal{E}\,.$$

Заметим, что обратный оператор для оператора $\mathcal{A}$ может и не существовать. Справедлив, однако, следующий факт

Предложение. *Если обратный оператор существует, то он единственный.*

Доказательство. В самом деле, пусть некоторый оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ имеет два обратных оператора $\mathcal{B}_1$ и $\mathcal{B}_2$. Тогда

$$\mathcal{B}_1 = \mathcal{B}_1\,\mathcal{E} = \mathcal{B}_1\,\mathcal{A}\,\mathcal{B}_2 = \mathcal{E}\,\mathcal{B}_2 = \mathcal{B}_2\,. \quad \square$$

В рассуждении, доказывающем единственность обратного оператора по существу используется тот факт, что обратный оператор определен как *двусторонний* обратный.

Для оператора, обратного к оператору $\mathcal{A}$ используется обозначение $\mathcal{A}^{-1}$ (это обозначение оправдано в силу единственности обратного оператора). Непосредственно из определения обратного оператора вытекает, что $\mathcal{A}^{-1}\,\mathcal{A}\,\boldsymbol{x} = \boldsymbol{x}$.

Выясним теперь, при каких условиях на оператор $\mathcal{A}$ обратный к нему оператор *существует*. Определим следующие понятия.

Определение. *Скажем, что оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{Y})$ действует взаимно однозначно, если $\mathcal{A}\,\boldsymbol{x}_1 \neq \mathcal{A}\,\boldsymbol{x}_2$ для любых двух элементов $\boldsymbol{x}_1 \in \mathcal{X}$ и $\boldsymbol{x}_2 \in \mathcal{X}$ таких, что $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$. В этом случае оператор $\mathcal{A}$ называется также мономорфизмом.*

Кроме того, оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{Y})$ называется эпиморфизмом, если для любого $\boldsymbol{y} \in \mathcal{Y}$ найдется $\boldsymbol{x} \in \mathcal{X}$ такой, что $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$.

Предложение 3.4. *Оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{Y})$ является мономорфизмом если и только если из того, что $\mathcal{A}\,\boldsymbol{x} = \mathbf{0}$ вытекает $\boldsymbol{x} = \mathbf{0}$.*

Доказательство. В самом деле, из существования элемента $\boldsymbol{x}_0 \in \mathcal{X}$ такого, что $\boldsymbol{x}_0 \neq \mathbf{0}$, но $\mathcal{A}\,\boldsymbol{x}_0 = \mathbf{0}$ и из того, что $\mathcal{A}\,\mathbf{0} = \mathbf{0}$ вытекает, что оператор $\mathcal{A}$ не может быть мономорфизмом. Обратно, из существования двух элементов $\boldsymbol{x}_1 \in \mathcal{X}$ и $\boldsymbol{x}_2 \in \mathcal{X}$ таких, что $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, но $\mathcal{A}\,\boldsymbol{x}_1 = \mathcal{A}\,\boldsymbol{x}_2$ вытекает, что для элемента $\boldsymbol{x}_0 = \boldsymbol{x}_1 - \boldsymbol{x}_2 \neq \mathbf{0}$ выполнено равенство $\mathcal{A}\,\boldsymbol{x}_0 = \mathbf{0}$, которое противоречит условию. □

Предложение 3.5. *Пусть линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ – мономорфизм. Тогда $\mathcal{A}$ – эпиморфизм.*

Доказательство. Пусть $\dim\mathcal{X} = n$. Для доказательства этого предложения нам необходимо проверить, что каждый элемент $\boldsymbol{y} \in \mathcal{X}$ представим в виде $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$ для некоторого элемента $\boldsymbol{x} \in \mathcal{X}$. Для этого достаточно проверить, что n линейно независимых элементов $\boldsymbol{x}_1,\ldots,\boldsymbol{x}_n \in \mathcal{X}$ переходят при отображении $\boldsymbol{x} \mapsto \mathcal{A}\,\boldsymbol{x}$ в n линейно независимых элементов $\mathcal{A}\,\boldsymbol{x}_1,\ldots,\mathcal{A}\,\boldsymbol{x}_n \in \mathcal{X}$.

Пусть числа $\alpha_1, \dots, \alpha_n \in \mathbb{C}$ таковы, что

$$\alpha_1 \mathcal{A}\,\boldsymbol{x}_1 + \dots + \alpha_n \mathcal{A}\,\boldsymbol{x}_n = \boldsymbol{0}.$$

Тогда, из линейности $\mathcal{A}$ вытекает, что

$$\mathcal{A}(\alpha_1\boldsymbol{x}_1 + \dots + \alpha_n\boldsymbol{x}_n) = \boldsymbol{0}.$$

Так как $\mathcal{A}$ действует взаимно однозначно, то $\alpha_1\boldsymbol{x}_1 + \dots + \alpha_n\boldsymbol{x}_n = \boldsymbol{0}$. В силу линейной независимости элементов $\boldsymbol{x}_1, \dots, \boldsymbol{x}_n$ из последнего равенства вытекает, что $\alpha_1 = \dots = \alpha_n = 0$. Следовательно, элементы $\mathcal{A}\,\boldsymbol{x}_1, \dots, \mathcal{A}\,\boldsymbol{x}_n$ линейно независимы. □

Теорема 3.6. *Линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет обратный если и только если $\mathcal{A}$ действует взаимно однозначно.*

Доказательство. Предположим, что оператор $\mathcal{A}$ имеет обратный. Возьмем два произвольных элемента $\boldsymbol{x}_1 \in \mathcal{X}$ и $\boldsymbol{x}_2 \in \mathcal{X}$ и предположим, что $\mathcal{A}\,\boldsymbol{x}_1 = \mathcal{A}\,\boldsymbol{x}_2$. Тогда $\mathcal{A}(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \boldsymbol{0}$ и, так как существует $\mathcal{A}^{-1}$, то

$$\boldsymbol{x}_1 - \boldsymbol{x}_2 = \mathcal{A}^{-1}\,\mathcal{A}(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \mathcal{A}^{-1}\,\boldsymbol{0} = \boldsymbol{0}.$$

Итак, из равенства $\mathcal{A}\,\boldsymbol{x}_1 = \mathcal{A}\,\boldsymbol{x}_2$ вытекает, что $\boldsymbol{x}_1 = \boldsymbol{x}_2$ и, следовательно, оператор $\mathcal{A}$ действует взаимно однозначно.

Обратно, пусть оператор $\mathcal{A}$ действует взаимно однозначно. Тогда каждому элементу $\boldsymbol{y} \in \mathcal{X}$ соответствует элемент $\boldsymbol{x} \in \mathcal{X}$ такой, что $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$. Определим оператор $\mathcal{A}^{-1} : \mathcal{X} \to \mathcal{X}$ таким образом, что $\mathcal{A}^{-1}\,\boldsymbol{y} = \mathcal{A}^{-1}(\mathcal{A}\,\boldsymbol{x}) = \boldsymbol{x}$. Проверим линейность оператора $\mathcal{A}^{-1}$. Возьмем произвольные $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$ и $\alpha \in \mathbb{C}$ и найдем такие $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathcal{X}$, что $\mathcal{A}\,\boldsymbol{v}_j = \boldsymbol{x}_j$ при $j = 1, 2$. Тогда

$$\mathcal{A}^{-1}(\boldsymbol{x}_1 + \boldsymbol{x}_2) = \mathcal{A}^{-1}(\mathcal{A}\,\boldsymbol{v}_1 + \mathcal{A}\,\boldsymbol{v}_2) = \mathcal{A}^{-1}\,\mathcal{A}(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2 = \mathcal{A}^{-1}\,\boldsymbol{x}_1 + \mathcal{A}^{-1}\,\boldsymbol{x}_2,$$
$$\mathcal{A}^{-1}(\alpha\boldsymbol{x}_1) = \mathcal{A}^{-1}(\alpha\,\mathcal{A}\,\boldsymbol{v}_1) = \mathcal{A}^{-1}\,\mathcal{A}(\alpha\boldsymbol{v}_1) = \alpha\boldsymbol{v}_1 = \alpha\,\mathcal{A}^{-1}(\boldsymbol{x}_1).$$

Из определения оператора $\mathcal{A}^{-1}$ немедленно заключаем, что $\mathcal{A}^{-1}$ действительно является обратным оператором для оператора $\mathcal{A}$. □

Отметим, что доказательства утверждений Предложения 3.5 и Теоремы 3.6 существенно опираются на то, что на множестве $\mathcal{X}$ задана структура линейного пространства и на линейность оператора $\mathcal{A}$.

3.2. Ядро и образ линейного оператора

Определение. *Ядром линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется множество*

$$\operatorname{Ker} \mathcal{A} := \{\boldsymbol{x} \in \mathcal{X} :\ \mathcal{A}\,\boldsymbol{x} = \boldsymbol{0}\}.$$

Упражнение. Проверить, что множество $\operatorname{Ker} \mathcal{A}$ является линейным подпространством пространства $\mathcal{X}$.

Предложение. *Если $\operatorname{Ker} \mathcal{A} = \{\boldsymbol{0}\}$, то оператор $\mathcal{A}$ действует взаимно однозначно.*

Доказательство. В самом деле, из того, что $\operatorname{Ker} \mathcal{A}$ содержит только элемент $\boldsymbol{0}$ (в дальнейшем этот факт мы будем коротко обозначать $\operatorname{Ker} \mathcal{A} = \boldsymbol{0}$), то из условия $\mathcal{A}\,\boldsymbol{x} = \boldsymbol{0}$ вытекает, что $\boldsymbol{x} = \boldsymbol{0}$. А это в свою очередь означает, что если для некоторых элементов $\boldsymbol{x}_1 \in \mathcal{X}$ и $\boldsymbol{x}_2 \in \mathcal{X}$ выполняется равенство $\mathcal{A}\,\boldsymbol{x}_1 = \mathcal{A}\,\boldsymbol{x}_2$, то $\mathcal{A}(\boldsymbol{x}_1 - \boldsymbol{x}_2) = \boldsymbol{0}$ и, следовательно, $\boldsymbol{x}_1 = \boldsymbol{x}_2$. □

Из этого предложения непосредственно вытекает следующее утверждение.

Предложение. *Оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет обратный в том и только том случае, когда $\operatorname{Ker} \mathcal{A} = \boldsymbol{0}$.*

Определение. *Образом оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется множество*

$$\operatorname{Im} \mathcal{A} := \{\boldsymbol{y} \in \mathcal{X} :\ \exists \boldsymbol{x} \in \mathcal{X} \text{ такой, что } \mathcal{A}\,\boldsymbol{x} = \boldsymbol{y}\}.$$

Упражнение. Проверить, что множество $\operatorname{Im}\mathcal{A}$ является линейным подпространством пространства $\mathcal{X}$.

Замечание. Понятия ядра и образа можно распространить (очевидным образом) на общий случай линейных операторов, действующих из одного линейного пространства $\mathcal{X}$ в другое линейное пространство $\mathcal{Y}$.

Справедлив еще один критерий обратимости линейного оператора.

Предложение. *Оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет обратный в том и только том случае, когда $\operatorname{Im}\mathcal{A} = \mathcal{X}$.*

Теорема 3.7. *Пусть $\mathcal{X}$ – конечномерное линейное пространство, а $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. Тогда*

$$\dim(\operatorname{Im}\mathcal{A}) + \dim(\operatorname{Ker}\mathcal{A}) = \dim\mathcal{X}. \tag{3.1}$$

Доказательство. Пусть $\dim\mathcal{X} = n$, $n \in \mathbb{N}$. Так как $\operatorname{Ker}\mathcal{A}$ – это подпространство пространства $\mathcal{X}$, то найдется такое подпространство $\mathcal{V}$ пространства $\mathcal{X}$, что $\mathcal{X} = \mathcal{V} \oplus \operatorname{Ker}\mathcal{A}$. В самом деле, выберем в пространстве $\mathcal{X}$ базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ такой, что элементы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$ при некотором натуральном $k < n$ образуют базис в $\operatorname{Ker}\mathcal{A}$. При этом в качестве подпространства $\mathcal{V}$ можно взять линейную оболочку векторов $\boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n$.

Так как $\mathcal{X} = \mathcal{V} \oplus \operatorname{Ker}\mathcal{A}$, то

$$\dim(\operatorname{Ker}\mathcal{A}) + \dim\mathcal{V} = n$$

и нам остается проверить, что $\dim(\operatorname{Im}\mathcal{A}) = \dim\mathcal{V}$.

Заметим, что $\mathcal{V} \cap \operatorname{Ker}\mathcal{A} = \{\boldsymbol{0}\}$. Следовательно, оператор $\mathcal{A}_{|\mathcal{V}}$ – сужение оператора $\mathcal{A}$ с $\mathcal{X}$ на подпространство $\mathcal{V} \subset \mathcal{X}$ является мономорфизмом. На основании Предложения 3.5 оператор $\mathcal{A}_{|\mathcal{V}}$ является эпиморфизмом и, следовательно, изоморфизмом. Проверим теперь, что $\operatorname{Im}\mathcal{A} = \operatorname{Im}\mathcal{A}_{|\mathcal{V}}$. Пусть для $\boldsymbol{y} \in \mathcal{X}$ существует $\boldsymbol{x} \in \mathcal{X}$ такой, что $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$. Так как $\boldsymbol{x}$ единственным образом может быть представлен в виде $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{v}$, где $\boldsymbol{x}_0 \in \operatorname{Ker}\mathcal{A}$, а $\boldsymbol{v} \in \mathcal{V}$, то $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}_0 + \mathcal{A}\,\boldsymbol{v} = \mathcal{A}\,\boldsymbol{v}$. Т.е. для любого $\boldsymbol{y} \in \operatorname{Im}\mathcal{A}$ существует $\boldsymbol{v} \in \mathcal{V}$ такой, что $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{v}$. Окончательно получаем, что так как пространства $\mathcal{V}$ и $\operatorname{Im}\mathcal{A}$ изоморфны, то имеют одинаковую размерность. □

Можно доказать утверждение, в некотором смысле являющееся обратным к только что доказанной теореме. Имеет место

Предложение 3.8. *Пусть $\mathcal{X}_1$ и $\mathcal{X}_2$ – два подпространства n-мерного линейного пространства $\mathcal{X}$ такие, что $\mathcal{X}_1 \cap \mathcal{X}_2 = \{\boldsymbol{0}\}$ и $\dim\mathcal{X}_1 + \dim\mathcal{X}_2 = n$. Тогда найдется линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ такой, что $\mathcal{X}_1 = \operatorname{Im}\mathcal{A}$ и $\mathcal{X}_2 = \operatorname{Ker}\mathcal{A}$.*

Доказательство. Пусть $k = \dim\mathcal{X}_1$. Выберем в пространстве $\mathcal{X}$ базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ так, чтобы элементы $\boldsymbol{e}_{k+1}, \ldots, \boldsymbol{e}_n$ принадлежали пространству $\mathcal{X}_2$. Теперь в пространстве $\mathcal{X}_1$ выберем базис $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_k$.

Определим теперь линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ следующим образом: $\mathcal{A}\,\boldsymbol{e}_j = \boldsymbol{w}_j$ при $j = 1, \ldots, k$ и $\mathcal{A}\,\boldsymbol{e}_j = \boldsymbol{0}$ при $j = k+1, \ldots, n$. Соответственно, для произвольного элемента $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_k\boldsymbol{e}_k + x_{k+1}\boldsymbol{e}_{k+1} + \cdots + x_n\boldsymbol{e}_n$ положим $\mathcal{A}\,\boldsymbol{x} = x_1\boldsymbol{w}_1 + \cdots + x_k\boldsymbol{w}_k$. Проверка того факта, что определенный указанным образом оператор $\mathcal{A}$ является линейным и удовлетворяет всем требуемым условиям оставляется читателю в качестве простого *упражнения*. □

Определение. *Число $\operatorname{rg}\mathcal{A} := \dim(\operatorname{Im}\mathcal{A})$ называется рангом оператора $\mathcal{A}$. Число $\operatorname{df}\mathcal{A} := \dim(\operatorname{Ker}\mathcal{A})$ называется дефектом оператора $\mathcal{A}$.*

Замечание. В терминах введенных характеристик можно сказать, что оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ (где $\mathcal{X}$ – конечномерное линейное пространство) имеет обратный тогда и только тогда, когда $\operatorname{rg}\mathcal{A} = \dim\mathcal{X}$ или, что эквивалентно, тогда и только тогда, когда $\operatorname{df}\mathcal{A} = 0$ равен нулю. Кроме того, всегда выполняется равенство $\operatorname{rk}\mathcal{A} + \operatorname{df}\mathcal{A} = \dim\mathcal{X}$.

Предложение 3.9. *Пусть* $\dim \mathcal{X} = n$ *и пусть* $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ *и* $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. *Тогда*

$$\operatorname{rg}(\mathcal{A}\,\mathcal{B}) \leqslant \operatorname{rg}\mathcal{A}, \qquad \operatorname{rg}(\mathcal{A}\,\mathcal{B}) \leqslant \operatorname{rg}\mathcal{B}\,.$$

Доказательство. Так как $\operatorname{Im}(\mathcal{A}\,\mathcal{B}) \subset \operatorname{Im}\mathcal{A}$, то $\dim(\operatorname{Im}(\mathcal{A}\,\mathcal{B})) \leqslant \dim(\operatorname{Im}\mathcal{A})$ и, следовательно, первое неравенство установлено.

Так как $\mathcal{A}\,\mathcal{B} \neq \mathcal{B}\,\mathcal{A}$, то включение $\operatorname{Im}(\mathcal{A}\,\mathcal{B}) \subset \operatorname{Im}\mathcal{B}$ может, в общем случае, не иметь места. Поэтому для доказательства второго неравенства мы используем другие аргументы. Заметим, что $\operatorname{Ker}\mathcal{B} \subset \operatorname{Ker}(\mathcal{A}\,\mathcal{B})$. Следовательно, $\dim(\operatorname{Ker}\mathcal{B}) \leqslant \dim(\operatorname{Ker}(\mathcal{A}\,\mathcal{B}))$ и, в свою очередь, $n - \dim(\operatorname{Ker}(\mathcal{A}\,\mathcal{B})) \leqslant n - \dim(\operatorname{Ker}\mathcal{B})$, откуда, с учетом равенства (3.1) вытекает $\dim(\operatorname{Im}(\mathcal{A}\,\mathcal{B})) \leqslant \dim(\operatorname{Im}\mathcal{B})$. А это и есть в точности второе неравенство из условия теоремы. □

Предложение 3.10. *Пусть* $\mathcal{X}$ *–* n*-мерное линейное пространство, а* $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ *и* $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. *Тогда*

$$\operatorname{rg}(\mathcal{A}\,\mathcal{B}) \geqslant \operatorname{rg}\mathcal{A} + \operatorname{rg}\mathcal{B} - n \tag{3.2}$$

$$\operatorname{df}(\mathcal{A}\,\mathcal{B}) \leqslant \operatorname{df}\mathcal{A} + \operatorname{df}\mathcal{B}\,. \tag{3.3}$$

Доказательство. Сразу заметим, что неравенство (3.3) непосредственно вытекает из неравенства (3.2). В самом деле,

$$\operatorname{df}(\mathcal{A}\,\mathcal{B}) = n - \operatorname{rg}(\mathcal{A}\,\mathcal{B}) \leqslant n - (\operatorname{rg}\mathcal{A} + \operatorname{rg}\mathcal{B} - n) = (n - \operatorname{rg}\mathcal{A}) + (n - \operatorname{rg}\mathcal{B}) = \operatorname{df}\mathcal{A} + \operatorname{df}\mathcal{B}\,.$$

Докажем неравенство (3.2). Пусть $\mathcal{A}_0 := \mathcal{A}_{|\operatorname{Im}\mathcal{B}}$ – сужение оператора $\mathcal{A}$ на подпространство $\operatorname{Im}\mathcal{B} \subset \mathcal{X}$. Так как $\operatorname{Im}\mathcal{A}_0 = \operatorname{Im}(\mathcal{A}\,\mathcal{B})$ то $\operatorname{rg}\mathcal{A}_0 = \operatorname{rg}(\mathcal{A}\,\mathcal{B})$. Далее, так как $\operatorname{Ker}\mathcal{A}_0 \subset \operatorname{Ker}\mathcal{A}$, то $\operatorname{df}\mathcal{A}_0 \leqslant \operatorname{df}\mathcal{A}$. Используя эти неравенства получаем

$$\operatorname{rg}\mathcal{B} = \dim(\operatorname{Im}\mathcal{B}) = \operatorname{rg}\mathcal{A}_0 + \operatorname{df}\mathcal{A}_0 \leqslant \operatorname{rg}\mathcal{A}_0 + \operatorname{df}\mathcal{A} = \operatorname{rg}(\mathcal{A}\,\mathcal{B}) + \operatorname{df}\mathcal{A},$$

откуда

$$\operatorname{rg}(\mathcal{A}\,\mathcal{B}) \geqslant \operatorname{rg}\mathcal{B} - \operatorname{df}\mathcal{A} = \operatorname{rg}\mathcal{A} + \operatorname{rg}\mathcal{B} - n. \quad \square$$

Заметим, что из полученных утверждений и ранге произведения операторов вытекает, что если $\dim \mathcal{X} = n$ и $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ таков, что $\operatorname{rg}\mathcal{A} = n$, до $\operatorname{rg}(\mathcal{A}\,\mathcal{B}) = \operatorname{rg}(\mathcal{B}\,\mathcal{A}) = \operatorname{rg}\mathcal{B}$, где $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. Проверка этого соотношения оставляется в качестве ***упражнения***.

3.3. Матричная запись линейных операторов

Фиксируем некоторое число $n \in \mathbb{N}$.

Матрица линейного оператора в заданном базисе линейного пространства. Пусть $\mathcal{X}$ – n-мерное линейное пространство. Выберем и зафиксируем в пространстве $\mathcal{X}$ некоторый базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$.

Пусть $\mathcal{A}$ – произвольный линейный оператор, действующий из $\mathcal{X}$ в $\mathcal{X}$. Рассмотрим разложения элементов $\mathcal{A}\,\boldsymbol{e}_1, \ldots, \mathcal{A}\,\boldsymbol{e}_n$ по базису $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$:

$$\mathcal{A}\,\boldsymbol{e}_k = \sum_{j=1}^{n} a_{jk}\boldsymbol{e}_j$$

и определим матрицу

$$[\mathcal{A}]_{\boldsymbol{e}} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \hdotsfor{4} \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix},$$

столбцами которой являются координатные столбцы элементов $\mathcal{A}\,\boldsymbol{e}_1, \ldots, \mathcal{A}\,\boldsymbol{e}_n$ относительно базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$.

Пусть $\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{e}_k$ – произвольный элемент пространства $\mathcal{X}$. Кроме того, пусть $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$ и $\boldsymbol{y} = \sum_{j=1}^{n} y_j \boldsymbol{e}_j$. Тогда

$$\mathcal{A}\,\boldsymbol{x} = \sum_{k=1}^{n} x_k \sum_{j=1}^{n} a_{jk} \boldsymbol{e}_j = \sum_{j=1}^{n} \boldsymbol{e}_j \sum_{k=1}^{n} a_{jk} x_k$$

и, координаты элемента $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$ относительно базиса $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ выражаются следующим образом

$$y_j = \sum_{k=1}^{n} a_{jk} x_k.$$

Из этой формулы вытекает, что координатный столбец $[\boldsymbol{y}]_{\boldsymbol{e}} = (y_1, \dots, y_n)^\top$ выражается через координатный столбец $[\boldsymbol{x}]_{\boldsymbol{e}} = (x_1, \dots, x_n)^\top$ по формуле

$$[\boldsymbol{y}]_{\boldsymbol{e}} = [\mathcal{A}]_{\boldsymbol{e}} [\boldsymbol{x}]_{\boldsymbol{e}}.$$

Матрица $[\mathcal{A}]_{\boldsymbol{e}}$ называется матрицей линейного оператора $\mathcal{A}$ в базисе $\boldsymbol{e}$.

Приведем несколько простых свойств матрицы линейного оператора. Во-первых заметим, что *тождественному* оператору $\mathcal{E}$ отвечает *единичная матрица* E. Во вторых, имеет место следующее утверждение

Предложение 3.11. *Пусть $\mathcal{X}$ – линейное пространство, $\dim \mathcal{X} = n$, а $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – некоторый базис в $\mathcal{X}$. Тогда для любой $n \times n$-матрицы A существует единственный линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$, матрица которого в базисе $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ равна A.*

Доказательство. Определим требуемый оператор $\mathcal{A}$ следующим образом. Для базисных векторов $\boldsymbol{e}_k$, $k = 1, \dots, n$ (напомним, что базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ выбран и зафиксирован выше) положим $\mathcal{A}\,\boldsymbol{e}_k := \sum_{j=1}^{n} a_{jk} \boldsymbol{e}_j$, а для произвольного вектора $\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{e}_k \in \mathcal{X}$ определим $\mathcal{A}\,\boldsymbol{x}$ по линейности: $\mathcal{A}\,\boldsymbol{x} := \sum_{k=1}^{n} x_k\, \mathcal{A}\,\boldsymbol{e}_k$. Проверка того факта, что определенный таким образом оператор $\mathcal{A}$ является линейный оператором из $\mathfrak{L}(\mathcal{X}, \mathcal{X})$ и имеет матрицу A является несложным **упражнением** и оставляется читателю. Таким образом, существование требуемого линейного оператора доказано.

Для проверки единственности оператора $\mathcal{A}$ достаточно заметить, что любой оператор из $\mathfrak{L}(\mathcal{X}, \mathcal{X})$ однозначно определяется своими значениями на элементах базиса. В самом деле, если линейный оператор $\mathcal{B}$ обладает тем свойством, что $\mathcal{B}\,\boldsymbol{e}_j = \boldsymbol{0}$ для всех $j = 1, \dots, n$, то для любого вектора $\boldsymbol{x} = \alpha_1 \mathbf{e}_1 + \dots + \alpha_n \mathbf{e}_n$, $\alpha_1, \dots, \alpha_n \in \mathbb{C}$, имеет место равенство $\mathcal{B}\,\boldsymbol{x} = \alpha_1\, \mathcal{B}\,\boldsymbol{e}_1 + \dots + \alpha_n\, \mathcal{B}\,\boldsymbol{e}_n = \boldsymbol{0}$, т.е. $\mathcal{B} = \mathcal{O}$ – нулевой оператор. □

Из только что доказанного утверждения непосредственно вытекает следующее.

Предложение. *Если оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет в базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$ матрицу A, а оператор $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – матрицу B, то оператор $\mathcal{A}\,\mathcal{B}$ имеет в базисе $\boldsymbol{e}$ матрицу AB и для любых чисел $\alpha, \beta \in \mathbb{C}$ оператор $\alpha\,\mathcal{A} + \beta\,\mathcal{B}$ имеет в базисе $\boldsymbol{e}$ матрицу $\alpha A + \beta B$.*

В самом деле, если $\boldsymbol{y} = (\alpha\,\mathcal{A} + \beta\,\mathcal{B})\boldsymbol{x}$ при $\boldsymbol{x} \in \mathcal{X}$, то $\boldsymbol{y} = \alpha\,\mathcal{A}\,\boldsymbol{x} + \beta\,\mathcal{B}\,\boldsymbol{x}$ и имеют место равенства

$$[\boldsymbol{y}]_{\boldsymbol{e}} = \alpha A [\boldsymbol{x}]_{\boldsymbol{e}} + \beta B [\boldsymbol{x}]_{\boldsymbol{e}} = (\alpha A + \beta B)[\boldsymbol{x}]_{\boldsymbol{e}}.$$

Далее, если $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{z} = \mathcal{B}\,\boldsymbol{x}$, а $\boldsymbol{w} = \mathcal{A}\,\boldsymbol{z}$, то $\boldsymbol{w} = \mathcal{A}\,\mathcal{B}\,\boldsymbol{x}$ и справедливы равенства

$$[\boldsymbol{z}]_{\boldsymbol{e}} = B[\boldsymbol{x}]_{\boldsymbol{e}}, \qquad [\boldsymbol{w}]_{\boldsymbol{e}} = A[\boldsymbol{z}]_{\boldsymbol{e}} = AB[\boldsymbol{x}]_{\boldsymbol{e}} = (AB)[\boldsymbol{x}]_{\boldsymbol{e}}.$$

Предложение 3.12. *Ранг линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ равен рангу матрицы A этого оператора, т.е.* $\operatorname{rg} \mathcal{A} = \operatorname{rg} A$.

Доказательство. По определению, $\operatorname{rg} \mathcal{A} = \dim(\operatorname{Im} \mathcal{A})$, а $\operatorname{Im} \mathcal{A}$ можно представить как линейную оболочку элементов $\boldsymbol{w}_k = \mathcal{A}\,\boldsymbol{e}_k = \sum_{j=1}^{n} a_{jk} \boldsymbol{e}_j$, $k = 1, \dots, n$. Следовательно, ранг оператора $\mathcal{A}$ равен максимальному числу линейно независимых элементов в системе $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$. Так как все элементы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ линейно независимы, то число линейно независимых элементов $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ совпадает с числом линейно независимых столбцов матрицы A, т.е. равно рангу матрицы A. $\square$

Используя понятие ранга матрицы линейного оператора можно привести еще одно необходимое и достаточное условие существования обратного оператора $\mathcal{A}^{-1}$ для оператора $\mathcal{A}$: *для оператора $\mathcal{A}$ существует обратный если и только если ранг матрицы A этого оператора равен размерности n пространства $\mathcal{X}$.*

Замечание. Если для оператора $\mathcal{A}$ существует обратный оператор $\mathcal{A}^{-1}$, то существует и обратная матрица A^{-1} для матрицы A. При этом матрица A^{-1} является матрицей оператора $\mathcal{A}^{-1}$ в базисе $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$.

Преобразование матрицы линейного оператора при переходе от базиса к базису. Рассмотрим линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ и предположим, что в пространстве $\mathcal{X}$ выбраны два базиса $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$, причем переход от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$ осуществляется при помощи матрицы S, т.е. $\boldsymbol{e}'_k = \sum_{r=1}^n \sigma_{rk}\boldsymbol{e}_k$, $k = 1, \dots, n$. Пусть $A = (a_{jk})_{j,k=1}^n$ – матрица оператора $\mathcal{A}$ в базисе $\boldsymbol{e}$, а $A' = (a'_{jk})_{j,k=1}^n$ – матрица оператора $\mathcal{A}$ в базисе $\boldsymbol{e}'$. Тогда, при $k = 1, \dots, n$

$$\mathcal{A}\,\boldsymbol{e}'_k = \sum_{j=1}^n a'_{jk}\boldsymbol{e}'_j = \sum_{j=1}^n a'_{jk} \sum_{r=1}^n \sigma_{rj}\boldsymbol{e}_r = \sum_{r=1}^n \boldsymbol{e}_r \sum_{j=1}^n \sigma_{rj}a'_{jk};$$

$$\mathcal{A}\,\boldsymbol{e}'_k = \mathcal{A}\left(\sum_{j=1}^n \sigma_{jk}\boldsymbol{e}_j\right) = \sum_{j=1}^n \sigma_{jk}\,\mathcal{A}\,\boldsymbol{e}_j = \sum_{j=1}^n \sigma_{jk} \sum_{r=1}^n a_{rj}\boldsymbol{e}_r = \sum_{r=1}^n \boldsymbol{e}_r \sum_{j=1}^n a_{rj}\sigma_{jk}.$$

Так как $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – базис, то сравнивая последние выражения в полученных цепочках равенств получаем, что для любых $k = 1, \dots, n$ и $r = 1, \dots, n$ выполняются равенства

$$\sum_{j=1}^n \sigma_{rj}a'_{jk} = \sum_{j=1}^n a_{rj}\sigma_{jk}.$$

Эти равенства, в свою очередь означают, что выполняется следующее матричное равенство

$$SA' = AS.$$

Заметим, что матрица S, как матрица перехода от базиса к базису, имеет ранг n и, следовательно, существует обратная матрица S^{-1}. Поэтому последнее равенство можно переписать в следующем (окончательном) виде

$$A' = S^{-1}AS.$$

Нами доказана

Теорема 3.13. *Матрицы A и A' линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ в базисах $\boldsymbol{e}$ и $\boldsymbol{e}'$ пространства $\mathcal{X}$ связаны соотношением $A' = S^{-1}AS$, где S – матрица перехода от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$.*

Замечание. Утверждение Теоремы 3.13 можно было получить заметно проще. В самом деле, если $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} = \mathcal{A}\,\boldsymbol{x}$, то справедливы соотношения

$$\boldsymbol{e}A[\boldsymbol{x}]_{\boldsymbol{e}} = \boldsymbol{e}[\boldsymbol{y}]_{\boldsymbol{e}} = \boldsymbol{y} = \boldsymbol{e}'[\boldsymbol{y}]_{\boldsymbol{e}'} = \boldsymbol{e}'A'[\boldsymbol{x}]_{\boldsymbol{e}'} = \boldsymbol{e}SA'S^{-1}[\boldsymbol{x}]_{\boldsymbol{e}},$$

из которых немедленно вытекает соотношение $A' = S^{-1}AS$.

Пусть оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет матрицу A в базисе $\boldsymbol{e}$ и матрицу A' в базисе $\boldsymbol{e}'$, а оператор $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет матрицу B в базисе $\boldsymbol{e}$ и матрицу B' в базисе $\boldsymbol{e}'$. Тогда для любых чисел $\alpha, \beta \in \mathbb{C}$ оператор $\alpha\,\mathcal{A} + \beta\,\mathcal{B}$ имеет матрицу $\alpha A + \beta B$ в базисе $\boldsymbol{e}$. Далее,

$$S^{-1}(\alpha A + \beta B)S = \alpha S^{-1}AS + \beta S^{-1}BS = \alpha A' + \beta B',$$

т.е., оператор $\alpha\mathcal{A}+\beta\mathcal{B}$ имеет матрицу $\alpha A'+\beta B'$ в базисе $\boldsymbol{e}'$. А так как оператор $\mathcal{A}\mathcal{B}$ имеет в базисе $\boldsymbol{e}$ матрицу AB и так как

$$S^{-1}(AB)S = S^{-1}ASS^{-1}BS = A'B',$$

то оператор $\mathcal{A}\mathcal{B}$ имеет матрицу $A'B'$ в базисе $\boldsymbol{e}'$.

Кроме того,

$$\det A' = \det(S^{-1}AS) = \det S^{-1}\det A\det S = \det A$$

и, следовательно, *определитель матрицы линейного оператора не зависит от базиса, в котором взята матрица этого линейного оператора, а зависит только от этого оператора.*

Таким образом можно ввести понятие *определителя* $\det\mathcal{A}$ линейного оператора: $\det\mathcal{A} := \det A$, где A – матрица оператора $\mathcal{A}$ в некотором базисе.

Заметим также, что $S^{-1}ES = E$ и, следовательно, единичный оператор имеет единичную матрицу в любом базисе.

3.4. Линейные функционалы и их координаты

Пусть $\mathcal{X}$ – линейное пространство (вещественное или комплексное) конечной размерности n, $n \in \mathbb{N}$. Рассмотрим произвольные линейный функционал $\mathfrak{f} \in \mathfrak{L}(\mathcal{X}, \mathbb{C})$.

Предположим, что в пространстве $\mathcal{X}$ выбран некоторый базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. Рассмотрим вектор $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{x} = x_1\boldsymbol{e}_+ \dots + x_n\boldsymbol{e}_n$. Значение функционала f на векторе $\boldsymbol{x}$ равно

$$\mathfrak{f}(\boldsymbol{x}) = \mathfrak{f}(x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n) = x_1\mathfrak{f}(\boldsymbol{e}_1) + \dots + x_n\mathfrak{f}(\boldsymbol{e}_n). \tag{3.4}$$

Заметим, что набор чисел $\{f_1 := \mathfrak{f}(\boldsymbol{e}_1), \dots, f_n := \mathfrak{f}(\boldsymbol{e}_n)\}$ однозначно определяет данный линейный функционал f.

Определение. *Числа $(f_1, \dots, f_n)$ называются коэффициентами (или координатами) линейного функционала $\mathfrak{f}$ в базисе $\boldsymbol{e}$.*

Таким образом, с каждым линейным функционалом $\mathfrak{f}$ связывается вектор-строка

$$(\mathfrak{f})_{\boldsymbol{e}} = (f_1, \dots, f_n)$$

— строка коэффициентов линейного функционала $\mathfrak{f}$ в базисе $\boldsymbol{e}$. При этом имеет место следующее равенство, являющееся записью равенства (3.4) в матричной форме:

$$\mathfrak{f}(\boldsymbol{x}) = (\mathfrak{f})_{\boldsymbol{e}}[x]_{\boldsymbol{e}}.$$

Выясним, как меняются координаты линейного функционала при смене базиса. Имеет место следующее утверждение.

Предложение 3.14. *Пусть $\mathfrak{f} \in \mathfrak{L}(\mathcal{X}, \mathbb{C})$ — линейный функционал в n-мерном линейном пространстве $\mathcal{X}$ и пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$ – два базиса в $\mathcal{X}$. Пусть, кроме того, S – матрица перехода от $\boldsymbol{e}$ к $\boldsymbol{e}'$. Тогда*

$$(\mathfrak{f})_{\boldsymbol{e}'} = (\mathfrak{f})_{\boldsymbol{e}}S. \tag{3.5}$$

В самом деле, для любого вектора $\boldsymbol{x} \in \mathcal{X}$ выполнено $[\boldsymbol{x}]_{\boldsymbol{e}} = S[\boldsymbol{x}]_{\boldsymbol{e}'}$, откуда

$$\mathfrak{f}(\boldsymbol{x}) = (\mathfrak{f})_{\boldsymbol{e}'}[\boldsymbol{x}]_{\boldsymbol{e}'} = (\mathfrak{f})_{\boldsymbol{e}}[\boldsymbol{x}]_{\boldsymbol{e}} = (\mathfrak{f})_{\boldsymbol{e}}S[\boldsymbol{x}]_{\boldsymbol{e}'},$$

откуда и следует требуемое утверждение. Справедливость равенства (3.5) можно проверить и в координатной форме. Пусть $S = (\sigma_{jk})$, а $(\mathfrak{f})_{\boldsymbol{e}'} = (f'_1, \dots, f'_n)$. Тогда

$$f'_j := \mathfrak{f}(\boldsymbol{e}'_j) = \mathfrak{f}\left(\sum_{r=1}^{n}\sigma_{rj}\boldsymbol{e}_r\right) = \sum_{r=1}^{n}\sigma_{rj}\mathfrak{f}(\boldsymbol{e}_j) = \sum_{r=1}^{n}\sigma_{rj}f_r, \qquad j = 1, \dots, n.$$

3.5. Полуторалинейные функции в комплексных пространствах

Пусть $\mathcal{X}$ – комплексное линейное пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$.

Определение. *Числовая функция $\mathcal{F} : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$ называется полуторалинейной функцией (или полуторалинейной формой), если для любых элементов $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{X}$ и для любых чисел $\alpha, \beta \in \mathbb{C}$ выполняются соотношения*

(1) $\mathcal{F}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = \mathcal{F}(\boldsymbol{x}, \boldsymbol{z}) + \mathcal{F}(\boldsymbol{y}, \boldsymbol{z})$;
(2) $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = \mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) + \mathcal{F}(\boldsymbol{x}, \boldsymbol{z})$;
(3) $\mathcal{F}(\alpha\boldsymbol{x}, \beta\boldsymbol{y}) = \alpha\overline{\beta}\,\mathcal{F}(\boldsymbol{x}, \boldsymbol{y})$.

Другими словами, полуторалинейная форма – это числовая функция двух векторных аргументов, определенная на векторах линейного пространства $\mathcal{X}$, линейная по первому аргументу и антилинейная по второму.

Здесь а далее под *антилинейностью* функции $\Phi(\boldsymbol{x})$ векторного аргумента $\boldsymbol{x}$ понимается, что функция Φ обладает следующим свойством $\Phi(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \overline{\alpha}\Phi(\boldsymbol{x}) + \overline{\beta}\Phi(\boldsymbol{y})$ для любых векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ и любых чисел $\alpha, \beta \in \mathbb{C}$ (в отличии от свойства линейности, числовой множитель выносится со знаком комплексного сопряжения).

Матричная запись полуторалинейных форм. Пусть теперь $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – некоторый базис в (комплексном) линейном пространстве $\mathcal{X}$, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$, и пусть $\boldsymbol{x} = \sum_{j=1}^n x_j \boldsymbol{e}_j \in \mathcal{X}$ и $\boldsymbol{y} = \sum_{k=1}^n y_k \boldsymbol{e}_k \in \mathcal{X}$.

Пусть $\mathcal{F}$ – некоторая полуторалинейная форма на $\mathcal{X}$. Тогда

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{F}\left(\sum_{j=1}^n x_j \boldsymbol{e}_j, \sum_{k=1}^n y_k \boldsymbol{e}_k\right) = \sum_{j=1}^n \sum_{k=1}^n x_j \overline{y}_k\, \mathcal{F}(\boldsymbol{e}_j, \boldsymbol{e}_k).$$

Определим матрицу

$$[\mathcal{F}]_{\boldsymbol{e}} = \begin{pmatrix} f_{11} & f_{12} & \cdots & f_{1n} \\ f_{21} & f_{22} & \cdots & f_{2n} \\ \hdotsfor{4} \\ f_{n1} & f_{n2} & \cdots & f_{nn} \end{pmatrix}$$

где $f_{jk} = \mathcal{F}(\boldsymbol{e}_j, \boldsymbol{e}_k)$ при $j, k = 1, \ldots, n$.

Матрица $[\mathcal{F}]_{\boldsymbol{e}}$ называется матрицей полуторалинейной формы $\mathcal{F}$ в базисе $\boldsymbol{e}$.

Имеет место равенство

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top} F\, \overline{[\boldsymbol{y}]}_{\boldsymbol{e}}, \tag{3.6}$$

где $F = [\mathcal{F}]_{\boldsymbol{e}}$, которое выражает значение полуторалинейной формы на векторах $\boldsymbol{x}$ и $\boldsymbol{y}$ в матричной форме.

Из этого равенства вытекает, что если в пространстве $\mathcal{X}$ заданы два базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$, связанные между собой матрицей перехода S, то

$$[\mathcal{F}]_{\boldsymbol{e}'} = S^{\top}[\mathcal{F}]_{\boldsymbol{e}}\overline{S},$$

где $\overline{S}$ – матрица, состоящая из элементов, комплексно сопряженных к соответствующим элементам матрицы S.

В самом деле, для любых векторов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ имеют место равенства

$$[\boldsymbol{x}]_{\boldsymbol{e}'}^{\top}[\mathcal{F}]_{\boldsymbol{e}'}\overline{[\boldsymbol{y}]}_{\boldsymbol{e}'} = \mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top}[\mathcal{F}]_{\boldsymbol{e}}\overline{[\boldsymbol{y}]}_{\boldsymbol{e}} = [\boldsymbol{x}]_{\boldsymbol{e}'}^{\top}\big(S^{\top}[\mathcal{F}]_{\boldsymbol{e}}\overline{S}\big)\overline{[\boldsymbol{y}]}_{\boldsymbol{e}'},$$

откуда и вытекает требуемое равенство.

Следующее утверждение непосредственно вытекает из определения матрицы полуторалинейной формы и, по сути дела, не нуждается в доказательстве:

Предложение 3.15. *Имеет место взаимно однозначное соответствие между полуторалинейными формами, определенными на n-мерном (комплексном) пространстве $\mathcal{X}$ и матрицами $F \in \mathrm{M}_n(\mathbb{C})$, выражаемое формулой (3.6).*

Эрмитовы формы. Выделить один специальный класс полуторалинейных форм.

Определение. *Полуторалинейная форма $\mathcal{F}$ называется эрмитовой, если для любых $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ выполняется соотношение*

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \overline{\mathcal{F}(\boldsymbol{y}, \boldsymbol{x})}.$$

Эрмитовы формы играют весьма важную роль и будут подробно изучены позже. Термин "эрмитова форма" связан с именем французского математика Ш. Эрмита (Charles Hermite, 24.12.1822–14.01.1901).

Пусть $\mathcal{F}$ – эрмитова форма, а $F = (f_{jk})_{j,k=1}^{n} = [\mathcal{F}]_{\boldsymbol{e}}$ – матрица этой формы в некотором базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$. Тогда, для при $j, k \in \{1, \ldots, n\}$

$$f_{jk} = \mathcal{F}(\boldsymbol{e}_j, \boldsymbol{e}_k) = \overline{\mathcal{F}(\boldsymbol{e}_k, \boldsymbol{e}_j)} = \overline{f}_{kj}.$$

В матричной форме это свойство записывается так: $F = \overline{F}^{\top}$, где символом $\overline{F}$ обозначена матрица, состоящая из элементов, комплексно сопряженных к соответствующим элементам матрицы F.

Определение. *Матрица $A \in \mathrm{M}_n(\mathbb{C})$, обладающая тем свойством, что $A = A^*$, где $A^* = \overline{A}^{\top}$, называется эрмитовой.*

3.6. Билинейные формы в вещественных пространствах

Понятие полуторалинейной формы, введенное в случае комплексного линейного пространства в вещественном случае видоизменяется в понятие *билинейной формы*. Билинейные формы определяются практически аналогично полуторалинейным, только свойство антилинейности по второму аргументу сводится (в силу вещественности коэффициентов) к соответствующему свойству линейности. Итак, пусть $\mathcal{X}$ — *вещественное* линейное пространство, т.е. линейное пространство над полем $\mathbb{R}$ вещественных чисел.

Определение. *Функция $\mathcal{B} : \mathcal{X} \times \mathcal{X} \to \mathcal{X}$ называется билинейной функцией (или билинейной формой) на $\mathcal{X}$, если для любых элементов $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathcal{X}$ и для любого $\alpha \in \mathbb{R}$ верны равенства*

(1) $\mathcal{B}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{z}) = \mathcal{B}(\boldsymbol{x}, \boldsymbol{z}) + \mathcal{B}(\boldsymbol{y}, \boldsymbol{z})$;
(2) $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y} + \boldsymbol{z}) = \mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) + \mathcal{B}(\boldsymbol{x}, \boldsymbol{z})$;
(3) $\mathcal{B}(\alpha\boldsymbol{x}, \boldsymbol{y}) = \alpha\, \mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{B}(\boldsymbol{x}, \alpha\boldsymbol{y})$.

Простым примером билинейной формы является выражение $\mathfrak{f}(\boldsymbol{x})\mathfrak{g}(\boldsymbol{y})$, где $\mathfrak{f}$ и $\mathfrak{g}$ – линейные функционалы, определенные на пространстве $\mathcal{X}$.

Формально говоря, билинейные функции можно рассматривать и в комплексных пространствах. В этом случае билинейные функции определяются аналогично с той только разницей, что в определении надо рассматривать числа $\alpha \in \mathbb{C}$. Однако в дальнейшем мы будем иметь дело с билинейными формами в основном в вещественных пространствах.

Несмотря на то, что следующие конструкции по сути дела аналогичны рассмотренным ранее соответствующим конструкциям для полуторалинейных форм, мы приведем их для полноты изложения.

Матричная запись билинейных форм. Пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – некоторый базис в n-мерном пространстве $\mathcal{X}$ и пусть $\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{e}_j \in \mathcal{X}$ и $\boldsymbol{y} = \sum_{k=1}^{n} y_k \boldsymbol{e}_k \in \mathcal{X}$.

Пусть $\mathcal{B}$ – некоторая билинейная форма на $\mathcal{X}$. Тогда

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{B}\left(\sum_{j=1}^{n} x_j \boldsymbol{e}_j, \sum_{k=1}^{n} y_k \boldsymbol{e}_k\right) = \sum_{j=1}^{n}\sum_{k=1}^{n} x_j \overline{y}_k\, \mathcal{B}(\boldsymbol{e}_j, \boldsymbol{e}_k).$$

Матрица
$$[\mathcal{B}]_{\boldsymbol{e}} = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{pmatrix}$$
где $b_{jk} = \mathcal{B}(\boldsymbol{e}_j, \boldsymbol{e}_k)$ при $j, k = 1, \ldots, n$, называется *матрицей билинейной формы* $\mathcal{B}$ в базисе $\boldsymbol{e}$.

Имеет место равенство
$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top} B\, \overline{[\boldsymbol{y}]_{\boldsymbol{e}}}, \tag{3.7}$$
где $B = [\mathcal{B}]_{\boldsymbol{e}}$, которое выражает значение полуторалинейной формы на векторах $\boldsymbol{x}$ и $\boldsymbol{y}$ в матричной форме.

Из этого равенства вытекает, что если в пространстве $\mathcal{X}$ заданы два базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}_1', \ldots, \boldsymbol{e}_n'\}$, связанные между собой матрицей перехода S, то
$$[\mathcal{B}]_{\boldsymbol{e}'} = S^{\top}[\mathcal{B}]_{\boldsymbol{e}} S. \tag{3.8}$$
Эта формула доказывается в точности также, как и формула преобразования матрицы полуторалинейной формы в комплексном случае. Проверка необходимых деталей оставляется в качестве *упражнения.*

Аналогично случаю полуторалинейных форм имеет место следующее утверждение:

Предложение 3.16. *Имеет место взаимно однозначное соответствие между билинейными формами, определенными на n-мерном пространстве $\mathcal{X}$ и матрицами $B \in \mathrm{M}_n(\mathbb{R})$, выражаемое формулой (3.7).*

Из формулы (3.8) непосредственно вытекает, что для матриц B и B' билинейной формы $\mathcal{B}$ в базисах $\boldsymbol{e}$ и $\boldsymbol{e}'$ выполняется соотношение $\operatorname{rg} B' = \operatorname{rg} B$ и, следовательно, корректно определено следующее понятие ранга билинейной формы.

Определение. *Рангом билинейной формы $\mathcal{B}$, определенной на конечномерном вещественном линейном пространстве $\mathcal{X}$, называют ранг ее матрицы в произвольном базисе пространства $\mathcal{X}$. Ранг билинейной формы $\mathcal{B}$ обозначается $\operatorname{rg} \mathcal{B}$.*

Если $\operatorname{rg} \mathcal{B} = n = \dim \mathcal{X}$, то форма $\mathcal{B}$ называется невырожденной, а если $\operatorname{rg} \mathcal{B} < n = \dim \mathcal{X}$, то форма $\mathcal{B}$ называется вырожденной.

Замечание. Из записи билинейной формы $\mathcal{B}$ в виде
$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = [\boldsymbol{x}]_{\mathbf{e}}^{\top}[\mathcal{B}]_{\mathbf{e}}[\boldsymbol{y}]_{\mathbf{e}} = \sum_{j,k=1}^{n} b_{jk} x_j y_k,$$
где $\mathbf{e}$ – некоторый базис пространства $\mathcal{X}$, $(x_1, \ldots, x_n)^{\top} = [\boldsymbol{x}]_{\mathbf{e}}$ и $(y_1, \ldots, y_n)^{\top} = [\boldsymbol{x}]_{\mathbf{e}}$ – координатные столбцы векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ относительно базиса $\boldsymbol{e}$, а $(b_{jk})_{j,k=1\ldots n} = [\mathcal{B}]_{\boldsymbol{e}}$ – матрица формы $\mathcal{B}$ в базисе $\boldsymbol{e}$, непосредственно вытекает, что билинейная форма $\mathcal{B}$ – это однородный многочлен степени 2 относительно координат векторов $\boldsymbol{x}$ и $\boldsymbol{y}$. Это наблюдение полезно иметь в виду при изучении билинейных форм и работе с ними.

Симметричные и кососимметричные билинейные формы. Полезно выделить некоторые специальные классы билинейных форм.

Определение. *Билинейная форма $\mathcal{B}$ называется симметричной, если для всех $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ выполнено равенство $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{B}(\boldsymbol{y}, \boldsymbol{x})$. Билинейная форма $\mathcal{B}$ на $\mathcal{X}$ называется кососимметричной, если $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = -\mathcal{B}(\boldsymbol{y}, \boldsymbol{x})$ для всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.*

Заметим, что симметричные билинейные формы являются аналогом эрмитовых форм в случае вещественного пространства. Кроме того, каждая билинейная форма $\mathcal{B}$ может быть представлена в виде суммы
$$\mathcal{B} = \mathcal{B}_1 + \mathcal{B}_2,$$

где
$$\mathcal{B}_1(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\Big(\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) + \mathcal{B}(\boldsymbol{y},\boldsymbol{x})\Big)$$
– симметричная билинейная форма, а
$$\mathcal{B}_2(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\Big(\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) - \mathcal{B}(\boldsymbol{y},\boldsymbol{x})\Big)$$
– кососимметричная билинейная форма.

Пусть $\mathcal{B}$ – симметричная билинейная форма, а $B = (b_{jk})_{j,k=1}^n = [\mathcal{B}]_{\boldsymbol{e}}$ – матрица этой формы в некотором базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$. Тогда, для при $j,k \in \{1,\ldots,n\}$
$$b_{jk} = \mathcal{B}(\boldsymbol{e}_j,\boldsymbol{e}_k) = \mathcal{B}(\boldsymbol{e}_k,\boldsymbol{e}_j) = b_{kj}.$$
В матричной форме это свойство записывается так: $B = B^\top$.

Если же $\mathcal{B}$ – кососимметричная билинейная форма, то
$$b_{jk} = \mathcal{B}(\boldsymbol{e}_j,\boldsymbol{e}_k) = -\,\mathcal{B}(\boldsymbol{e}_k,\boldsymbol{e}_j) = -b_{kj},$$
а последнее равенство означает, что $B = -B^\top$.

Определение. *Матрица $A \in \mathrm{M}_n(\mathbb{R})$, обладающая тем свойством, что $A = A^\top$, называется симметричной (или симметрической). Матрица же $A \in \mathrm{M}_n(\mathbb{R})$, обладающая свойством $A = -A^\top$, называется кососимметричной (или кососимметрической).*

Легко проверяется и утверждение о том, что если матрица билинейной формы является симметричной (соответственно, кососимметричной), то такая билинейная форма является симметричной (соответственно, кососимметричной).

В самом деле, если $\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) = \sum_{j,k=1}^n b_{jk}x_jy_k$, где $B = (b_{jk})_{j,k=1}^n$ – заданная матрица, а числа x_j и y_k – координаты векторов $\boldsymbol{x},\boldsymbol{y} \in \mathcal{X}$ относительно базиса $\boldsymbol{e}$ в $\mathcal{X}$, то $\mathcal{B}(\boldsymbol{y},\boldsymbol{x}) = \sum_{j,k=1}^n b_{kj}x_jy_k$. Отсюда, с учетом симметричности (соответственно, кососимметричности) матрицы B непосредственно вытекает требуемое утверждение.

РАЗДЕЛ 4

Двойственное пространство. Сопряженный оператор

4.1. Двойственное пространство и двойственный базис

Рассмотрим линейное пространство $\mathcal{X}^* = \mathfrak{L}(\mathcal{X}, \mathbb{C})$ всех линейных функционалов на линейном пространстве $\mathcal{X}$. Напомним, что если $\mathfrak{f}, \mathfrak{g} \in \mathcal{X}^*$, а $\alpha, \beta \in \mathbb{C}$, то отображение $(\alpha\mathfrak{f} + \beta\mathfrak{g}) : \mathcal{X} \to \mathbb{C}$ определяется соотношением $(\alpha\mathfrak{f} + \beta\mathfrak{g})(\boldsymbol{x}) = \alpha\mathfrak{f}(\boldsymbol{x}) + \beta\mathfrak{g}(\boldsymbol{x})$, $\boldsymbol{x} \in \mathcal{X}$.

Определение. *Линейное пространство $\mathcal{X}^*$ называется двойственным пространством к пространству $\mathcal{X}$ (используются также термины сопряженное или дуальное пространство).*

Как было показано выше, при заданном базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$ имеет место взаимно однозначное соответствие $\Phi : \mathfrak{f} \mapsto (f_1, \dots, f_n)$ между линейными функционалами на $\mathcal{X}$ и вектор-строками из n элементов. Из правил сложения функционалов и умножения функционалов на числа вытекают следующие свойства отображения Φ. Пусть $\mathfrak{f}, \mathfrak{g} \in \mathcal{X}^*$, причем $(\mathfrak{f})_{\boldsymbol{e}} = (f_1, \dots, f_n)$, а $(\mathfrak{g})_{\boldsymbol{e}} = (g_1, \dots, g_n)$ и пусть $\lambda \in \mathbb{C}$. Тогда $(\mathfrak{f} + \mathfrak{g})_{\boldsymbol{e}} = (f_1 + g_1, \dots, f_n + g_n)$ и $(\lambda\mathfrak{f})_{\boldsymbol{e}} = (\lambda f_1, \dots, \lambda f_n)$. Таким образом, Φ – это *изоморфизм* линейных пространств $\mathcal{X}^*$ и $\mathbb{C}^n$ (рассматриваемого как пространство вектор-строк) и, следовательно, $\dim \mathcal{X}^* = \dim \mathcal{X} = n$.

Рассмотрим набор вектор-строк $\beta_j = (\beta_{j,1}, \dots, \beta_{j,n})$ при $j = 1, \dots, n$ такой, что

$$\beta_{j,k} = \begin{cases} 1, & j = k; \\ 0, & j \neq k \end{cases}$$

и введем набор линейных функционалов $\mathfrak{e}^1, \dots, \mathfrak{e}^n$ определенных в базисе $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ этими координатными строками. Непосредственное вычисление показывает, что для этих функционалов верны соотношения

$$\mathfrak{e}^j(\boldsymbol{e}_k) = \delta_{jk},$$

где δ_{jk} – символ Кронекера. При этом для любого вектора $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n$ имеем

$$\mathfrak{e}^j(\boldsymbol{x}) = \sum_{k=1}^{n} x_k \mathfrak{e}^j(\boldsymbol{e}_k) = \sum_{k=1}^{n} x_k \beta_{j,k} = x_j.$$

Так как вектор-строки β_j, $j = 1, \dots, n$ линейно независимы, то и функционалы $\mathfrak{e}^j$, $j = 1, \dots, n$, рассматриваемые как элементы пространства $\mathcal{X}^*$, будут линейно независимыми. Таким образом нами установлено следующее утверждение.

Теорема 4.1. *Пусть $\mathcal{X}$ – n-мерное линейное пространство. Тогда пространство $\mathcal{X}^*$ также имеет размерность n. Если $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – базис в $\mathcal{X}$, а $\mathfrak{e}^1, \dots, \mathfrak{e}^n$ – такие линейные функционалы, что $\mathfrak{e}^j(\boldsymbol{e}_k) = \delta_{jk}$, то $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$ – базис в $\mathcal{X}^*$.*

Определение. *Базис $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$ пространства $\mathcal{X}^*$ (см. формулировку теоремы 4.1) называется двойственным (дуальным или сопряженным) базисом для данного базиса $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$.*

Выясним смысл “двойственности”, содержащейся в понятии двойственного базиса. Во-первых заметим, что между базисами в $\mathcal{X}$ и $\mathcal{X}^*$ установлено взаимно-однозначное соответствие (которое каждому базису ставит с соответствие соответствующий сопряженный базис). Во вторых заметим, что выражение $\mathfrak{f}(\boldsymbol{x})$, где $\mathfrak{f} \in \mathcal{X}^*$, а $\boldsymbol{x} \in \mathcal{X}$, допускает двоякую трактовку. При фиксированном $\mathfrak{f}$ – это линейный функционал на $\mathcal{X}$, а при фиксированном $\boldsymbol{x}$ – линейный функционал на $\mathcal{X}^*$. Отталкиваясь от этого, запишем выражение $\mathfrak{f}(\boldsymbol{x})$ – значение функционала $\mathfrak{f}$ на векторе $\boldsymbol{x}$ в виде $\langle \mathfrak{f}|\boldsymbol{x}\rangle$, смысл этой записи будет ясен чуть позднее.

Заметим, что выражение $\langle\cdot|\cdot\rangle$ определяет отображение $\mathcal{X}^* \times \mathcal{X} \to \mathbb{C}$, линейное по каждому аргументу, т.е.

$$\langle \alpha\mathfrak{f} + \beta\mathfrak{g}|\boldsymbol{x}\rangle = \alpha\langle \mathfrak{f}|\boldsymbol{x}\rangle + \beta\langle \mathfrak{g}|\boldsymbol{x}\rangle, \qquad \langle \mathfrak{f}|\alpha\boldsymbol{x} + \beta\boldsymbol{y}\rangle = \lambda\langle \mathfrak{f}|\boldsymbol{x}\rangle + \beta\langle \mathfrak{f}|\boldsymbol{y}\rangle \tag{4.1}$$

при всех $\mathfrak{f}, \mathfrak{g} \in \mathcal{X}^*$, $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, $\alpha, \beta \in \mathbb{C}$.

Определение. *Пусть $\mathcal{V}$ и $\mathcal{W}$ – линейные пространства. Отображение декартова произведения $\mathcal{V} \times \mathcal{W} = \{(v, w) : \ v \in \mathcal{V}, w \in \mathcal{W}\}$ в $\mathbb{C}$ (или в $\mathbb{R}$), линейное по каждому из двух своих аргументов называют спариванием (происходит от английского слова paring) между пространствами $\mathcal{V}$ и $\mathcal{W}$.*

Спаривание между $\mathcal{X}^*$ и $\mathcal{X}$ определенное соотношением $\langle \mathfrak{f}|\boldsymbol{x}\rangle := \mathfrak{f}(\boldsymbol{x})$, $\mathfrak{f} \in \mathcal{X}^*$, $\boldsymbol{x} \in \mathcal{X}$ называется каноническим.

Разложим теперь произвольные элементы $\mathfrak{f} \in \mathcal{X}^*$ и $\boldsymbol{x} \in \mathcal{X}$ по двойственным базисам $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$ и $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ соответственно: $\mathfrak{f} = \sum_{k=1}^n f_k\mathfrak{e}^k$, и $\boldsymbol{x} = \sum_{j=1}^n x_j\boldsymbol{e}_j$. Тогда

$$\mathfrak{f}(\boldsymbol{x}) = \sum_{j=1}^n f_j\mathfrak{e}^j(\boldsymbol{x}) = \sum_{j=1}^n f_j \sum_{k=1}^n x_k\mathfrak{e}^j(\boldsymbol{e}_k) = \sum_{j=1}^n\sum_{k=1}^n x_kf_j\mathfrak{e}^j(\boldsymbol{e}_k) = \sum_{j=1}^n f_jx_j.$$

Заметим также, что координаты $(f_1, \dots, f_n)$ линейного функционала $\mathfrak{f} \in \mathcal{X}^*$ в базисе $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$ и координаты $(x_1, \dots, x_n)$ вектора $\boldsymbol{x} \in \mathcal{X}$ в базисе $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$, сопряженном к $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$, могут быть вычислены по формулам

$$x_k := \langle \mathfrak{e}^k|\boldsymbol{x}\rangle \qquad \text{и} \qquad f_k := \langle \mathfrak{f}|\boldsymbol{e}_k\rangle \tag{4.2}$$

соответственно. В самом деле, $\langle \mathfrak{e}^k|\boldsymbol{x}\rangle = \left\langle \mathfrak{e}^k\middle|x_1\boldsymbol{e}_1 + \dots + x_n\boldsymbol{e}_n\right\rangle = \sum_{j=1}^n x_j\langle \mathfrak{e}^k|\boldsymbol{e}_j\rangle = x_k$, а $\langle \mathfrak{f}|\boldsymbol{e}_k\rangle = \left\langle f_1\mathfrak{e}^1 + \dots + f_n\mathfrak{e}^n\middle|\boldsymbol{e}_k\right\rangle = \sum_{j=1}^n f_j\langle \mathfrak{e}^j|\boldsymbol{e}_k\rangle = f_k$. Формулы (4.2) также можно рассматривать как проявление свойства двойственности базисов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$.

Пример 4.2. Рассмотрим пространство $\mathbb{R}[t]_n$. Отображение $\mathfrak{f}_\xi : P \mapsto P(\xi)$, ставящее многочлену $P \in \Pi_n$ его значение в точке $\xi \in \mathbb{R}$ является, очевидно, линейным функционалом на пространстве $\mathbb{R}[t]_n$. Можно показать, что совокупность функционалов $\mathfrak{f}_\xi$ при $\xi \in \{\xi_0, \dots, \xi_n\}$, где $\xi_0, \dots, \xi_n$ – различные точки из $\mathbb{R}$, является линейно независимой. Следовательно, она является базисом в $\mathbb{R}[t]_n^*$.

Другой пример базиса в $\mathbb{R}[t]_n^*$ можно получить рассмотрев семейство линейных функционалов $\mathfrak{f}_k : P \mapsto aP^{(k)}(\xi)$, где $a \in \mathbb{R}$ и $\xi \in \mathbb{R}$ – некоторые фиксированные параметры, а $k = 0, 1, \dots, n$. Проверка линейной независимости функционалов $\mathfrak{f}_k$ оставляется в качестве несложного *упражнения*.

Можно проверить, что набор функционалов $\mathfrak{f}_k^0 : P \mapsto P^{(k)}(0)/k!$ образует базис в $\mathbb{R}[t]_n^*$, двойственный к базису $\{1, t, \dots, t^n\}$ в $\mathbb{R}[t]_n$, а набор функционалов $\mathfrak{f}_k^\xi : P \mapsto P^{(k)}(\xi)/k!$ при $k = 0, 1, \dots, n$ образует базис в $\mathbb{R}[t]_n^*$, двойственный к базису $\{1, (t - \xi), \dots, (t - \xi)^n\}$ в пространстве $\mathbb{R}[t]_n$.

4.2. Рефлексивность

Напомним, что если $\dim \mathcal{X} = n$, то $\dim \mathcal{X}^* = n$. Следовательно, на основании теоремы об изоморфизме конечномерных линейных пространств одинаковой размерности, пространства $\mathcal{X}$ и $\mathcal{X}^*$ изоморфны. В силу аналогичных соображений изоморфными будут также пространства $\mathcal{X}^*$ и $\mathcal{X}^{**} = (\mathcal{X}^*)^*$. На первый взгляд кажется, что пространство $\mathcal{X}^{**}$ – это пространство весьма экзотической природы. В самом деле, оно состоит из линейных функционалов на пространстве линейных функционалов на $\mathcal{X}$. Но, как оказывается, между пространствами $\mathcal{X}$ и $\mathcal{X}^{**}$ существует весьма простая и естественная связь. Эта связь устанавливается следующей теоремой.

Теорема 4.3. *Отображение $\varepsilon : \mathcal{X} \to \mathcal{X}^{**}$, определяемое следующим образом*

$$\varepsilon(\boldsymbol{x}) = \varepsilon_{\boldsymbol{x}}, \quad \varepsilon_{\boldsymbol{x}}(\mathfrak{f}) = \mathfrak{f}(\boldsymbol{x}), \qquad \boldsymbol{x} \in \mathcal{X}, \quad \mathfrak{f} \in \mathcal{X}^*, \quad \varepsilon_{\boldsymbol{x}} \in \mathcal{X}^{**},$$

является изоморфизмом. Оно называется каноническим (или естественным) изоморфизмом между пространствами $\mathcal{X}$ и $\mathcal{X}^{**}$.

Доказательство. Линейность ε проверяется непосредственно исходя из его определения. В самом деле, если $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, а $\alpha, \beta \in \mathbb{C}$, то для любого линейного функционала $\mathfrak{f} \in \mathcal{X}^*$ имеет место цепочка равенств

$$\varepsilon_{\alpha\boldsymbol{x}+\beta\boldsymbol{y}}(\mathfrak{f}) = \mathfrak{f}(\alpha\boldsymbol{x} + \beta\boldsymbol{y}) = \alpha\mathfrak{f}(\boldsymbol{x}) + \beta\mathfrak{f}(\boldsymbol{y}) = \alpha\varepsilon_{\boldsymbol{x}}(\mathfrak{f}) + \beta\varepsilon_{\boldsymbol{y}}(\mathfrak{f}) = (\alpha\varepsilon_{\boldsymbol{x}} + \beta\varepsilon_{\boldsymbol{y}})(\mathfrak{f}).$$

Проверим теперь биективность отображения ε. Выберем в пространствах $\mathcal{X}$ и $\mathcal{X}^*$ двойственные базисы $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ и $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$ соответственно. Тогда

$$\varepsilon_{\boldsymbol{e}_j}(\mathfrak{e}^k) = \mathfrak{e}^k(\boldsymbol{e}_j) = \delta_{k,j}.$$

Используя рассуждения, аналогичные приведенным при доказательстве теоремы 4.1 получим, что $\mathcal{X}^{**} = \mathrm{Span}\{\varepsilon_{\boldsymbol{e}_1}, \dots, \varepsilon_{\boldsymbol{e}_n}\}$, т.е. набор элементов $\{\varepsilon_{\boldsymbol{e}_1}, \dots, \varepsilon_{\boldsymbol{e}_n}\}$ образует базис в $\mathcal{X}^{**}$, двойственный к базису $\{\mathfrak{e}^1, \dots, \mathfrak{e}^n\}$ в пространстве $\mathcal{X}^*$. Из этого вытекает как сюръективность, так и инъективность отображения ε. □

Замечание. Построенный в Теореме 4.3 изоморфизм ε назван естественным (каноническим) в силу того, что он не зависит от выбора конкретного базиса.

Определение. *Свойство линейный пространств, состоящее в том, что между пространствами $\mathcal{X}$ и $\mathcal{X}^{**}$ существует естественный изоморфизм носит название рефлексивность.*

Замечание. Конечномерные линейные пространства рефлексивны. Однако это, в общем случае, не так для бесконечномерных пространств. Соответствующие примеры изучаются в курсе функционального анализа.

Рефлексивность позволяет отождествить пространства $\mathcal{X}$ и $\mathcal{X}^{**}$. При этом уже пространство $\mathcal{X}$ можно понимать как пространство линейных функционалов на $\mathcal{X}^*$. При таком понимании формулы спаривания приобретают симметричный вид $\boldsymbol{x}(\mathfrak{f}) = \langle \mathfrak{f} | \boldsymbol{x} \rangle = \mathfrak{f}(\boldsymbol{x})$, где в первом равенстве элемент $\boldsymbol{x}$ рассматривается как элемент $\mathcal{X}^{**}$, а в последнем – как элемент $\mathcal{X}$.

Из теоремы 4.1 вытекает следующее утверждение.

***Следствие* 4.4.** *Если пространство $\mathcal{X}$ рефлексивно, то для любого базиса в $\mathcal{X}^*$ существует однозначно определенный двойственный ему базис в $\mathcal{X}$.*

Пример 4.5. Проверим, что система линейных функционалов $\mathfrak{f}_k \in \mathbb{R}[t]_n^*$, определенных при $k = 0, \dots, n$ равенствами

$$\mathfrak{f}_k(P) = P(k), \quad P \in \mathbb{R}[t]_n,$$

образует базис в пространстве $\mathbb{R}[t]_n^*$. В самом деле, $\dim \mathbb{R}[t]_n^* = \dim \mathbb{R}[t]_n = n + 1$, а система $\{\mathfrak{f}_0, \dots, \mathfrak{f}_n\}$ состоит из $n+1$ элемента. Если мы докажем линейную независимость элементов $\mathfrak{f}_0, \dots, \mathfrak{f}_n$, то, на основании Предложения 2.7, функционалы $\mathfrak{f}_0, \dots, \mathfrak{f}_n$ будут образовывать базис в $\mathbb{R}[t]_n^*$. Рассмотрим произвольную линейную комбинацию

$$\mathfrak{f} = \alpha_0\mathfrak{f}_0 + \alpha_1 f_1 + \dots + \alpha_n\mathfrak{f}_n$$

и предположим, что $\mathfrak{f} \equiv 0$. Это означает, что $\mathfrak{f}(P) = 0$ для любого $P \in \mathbb{R}[t]_n$. Рассмотрим многочлен $Q(t) = \prod_{k=0}^n (t - k)$, обращающийся в нуль в точках $0, 1, \dots, n$ и многочлены $P_k(t) := Q(t)/(t-k)$ при $k = 1, \dots, n$. Так как многочлен Q содержит множитель $(t-k)$ для любого $k = 1, \dots, n$, то все P_k в самом деле являются многочленами. При этом $P_k(m) = 0$ при $m = 0, 1, \dots, n$ и $m \neq k$ и $P_k(k) \neq 0$. Далее,

$$0 = \mathfrak{f}(P_k) = \sum_{j=0}^n \alpha_j\mathfrak{f}_j(P_k) = \sum_{j=0}^n \alpha_j P_k(j) = \alpha_k P_k(k)$$

для любого $k = 0, 1, \ldots, n$ и, следовательно, $\alpha_k = 0$ для всех $k = 0, 1, \ldots, n$. А последнее в точности означает, что система $\{\mathfrak{f}_0, \ldots, \mathfrak{f}_n\}$ линейно независима.

Итак, $\{\mathfrak{f}_0, \ldots, \mathfrak{f}_n\}$ – базис в $\mathbb{R}[t]_n^*$. Для того, чтобы найти базис в $\mathbb{R}[t]_n$, сопряженный базису $\{\mathfrak{f}_0, \ldots, \mathfrak{f}_n\}$ в $\mathbb{R}[t]_n^*$ необходимо найти такие многочлены $\widetilde{P}_k \in \mathbb{R}[t]_n$, что $\mathfrak{f}_k(\widetilde{P}_k) = 1$ при $k = 0, 1, \ldots, n$, а $\mathfrak{f}_j(\widetilde{P}_k) = 0$ при $j = 0, 1, \ldots, n$ и $j \neq k$. Из определения многочленов P_k при $k = 0, 1, \ldots, n$ непосредственно вытекает, что можно определить многочлены $\widetilde{P}_k$ соотношением $\widetilde{P}_k(t) := P_k(t)/P_k(k)$. Проверка того, что система многочленов $\{\widetilde{P}_0, \ldots, \widetilde{P}_n\}$ линейно независима очевидна и оставляется в качестве *упражнения*.

4.3. Условия линейной независимости

В этом параграфе мы еще раз обсудим понятие линейной зависимости элементов некоторого линейного пространства $\mathcal{X}$ и установим один интересный рабочий критерий линейной зависимости, формулируемый в терминах двойственного пространства $\mathcal{X}^*$. Мы начнем с доказательства следующего утверждения

Предложение 4.6. *Если элементы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m \in \mathcal{X}$ линейно зависимы, то для любых линейных функционалов $\mathfrak{f}_1, \ldots, \mathfrak{f}_m \in \mathcal{X}^*$ имеет место равенство $\det F = 0$, где*

$$F = \begin{pmatrix} \mathfrak{f}_1(\boldsymbol{x}_1) & \cdots & \mathfrak{f}_1(\boldsymbol{x}_m) \\ \multicolumn{3}{c}{\dotfill} \\ \mathfrak{f}_m(\boldsymbol{x}_1) & \cdots & \mathfrak{f}_m(\boldsymbol{x}_m) \end{pmatrix}.$$

Доказательство. Так как элементы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ линейно зависимы, то один из этих элементов является линейной комбинацией двух других. Пусть, например,

$$\boldsymbol{x}_m = \alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m-1} \boldsymbol{x}_{m-1}.$$

Преобразуем матрицу F следующим образом: из последнего столбца вычтем первый, умноженный на α_1, затем второй, умноженный на α_2 и, наконец, столбец с номером $(m-1)$, умноженный на α_{m-1}. При таком преобразовании значение $\det F$ не меняется. При этом, последний столбец преобразованной матрицы будет состоять из элементов

$$\mathfrak{f}_j(\boldsymbol{x}_m) - \sum_{r=1}^{m-1} \alpha_r \mathfrak{f}_j(\boldsymbol{x}_r) = \mathfrak{f}_j(\boldsymbol{x}_m - \alpha_1 \boldsymbol{x}_1 - \cdots - \alpha_{m-1} \boldsymbol{x}_{m-1}) = \mathfrak{f}_j(0) = 0, \quad j = 1, \ldots, m.$$

Следовательно, $\det F = 0$. □

Упражнение. Пусть матрица F в Предложении 4.6 невырождена. Доказать, что и система $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m\}$ и система $\{\mathfrak{f}_1, \ldots, \mathfrak{f}_m\}$ будут в этом случае линейно независимы.

Предложение 4.7. *Если система линейных функционалов $\{\mathfrak{f}_1, \ldots, \mathfrak{f}_n\}$ образует базис в пространстве $\mathcal{X}^*$, сопряженном к линейному пространству $\mathcal{X}$ размерности n, то векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathcal{X}$ будут линейной независимы если и только если*

$$\det F \neq 0, \qquad \text{где} \quad F = \begin{pmatrix} \mathfrak{f}_1(\boldsymbol{x}_1) & \cdots & \mathfrak{f}_1(\boldsymbol{x}_n) \\ \multicolumn{3}{c}{\dotfill} \\ \mathfrak{f}_n(\boldsymbol{x}_1) & \cdots & \mathfrak{f}_n(\boldsymbol{x}_n) \end{pmatrix}.$$

Доказательство. Из Предложения 4.6 следует, что из линейной зависимости векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ вытекает равенство $\det F = 0$. Если векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ линейно независимы, то они образуют базис в $\mathcal{X}$ (см. Предложение 2.7). Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – базис в $\mathcal{X}$, двойственный к базису $\{\mathfrak{f}_1, \ldots, \mathfrak{f}_n\}$. Пусть

$$\boldsymbol{x}_j = x_{1j} \boldsymbol{e}_1 + \cdots + x_{nj} \boldsymbol{e}_n.$$

Тогда матрица

$$X := \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \multicolumn{3}{c}{\dotfill} \\ x_{n1} & \cdots & x_{nn} \end{pmatrix}$$

является матрицей перехода от базиса $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ к базису $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. По определению матрицы перехода, $\det X \neq 0$. Далее, в силу определения двойственного базиса,

$$\mathfrak{f}_j(\boldsymbol{x}_k) = \mathfrak{f}_j(x_{1k}\boldsymbol{e}_1 + \cdots + x_{nk}\boldsymbol{e}_n) = \sum_{r=1}^{n} x_{rk}\mathfrak{f}_j(\boldsymbol{e}_r) = x_{jk}.$$

Следовательно, $F = X$ – матрица координат векторов $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n\}$ в базисе, двойственном к базису $\{\mathfrak{f}_1, \ldots, \mathfrak{f}_n\}$. Окончательно заключаем, что $\det F \neq 0$. □

Опираясь на только что доказанные Предложения 4.6 и 4.7 установим следующий критерий линейной зависимости векторов линейного пространства $\mathcal{X}$.

Теорема 4.8. *Пусть система линейных функционалов $\{\mathfrak{f}_1, \ldots, \mathfrak{f}_n\}$ образует базис в пространстве $\mathcal{X}^*$, сопряженном к линейному пространству $\mathcal{X}$ размерности n. Тогда число линейно независимых векторов среди векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathcal{X}$ равно порядку наибольшего отличного от нуля определителя матрицы вида $\big(\mathfrak{f}_t(\boldsymbol{x}_j)\big)$, где $1 \leqslant t = t_1, \ldots, t_m \leqslant n$, $1 \leqslant j = j_1, \ldots, j_m \leqslant k$.*

Доказательство. Пусть r – число линейно независимых векторов среди $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$. Так как любые $m > r$ векторов из этого набора линейно зависимы, то определитель соответствующей матрицы равен нулю в силу Предложения 4.6.

Найдем среди матриц вида $\big(\mathfrak{f}_t(\boldsymbol{x}_j)\big)$, где $1 \leqslant t = t_1, \ldots, t_r \leqslant n$, $1 \leqslant j = j_1, \ldots, j_r \leqslant k$ матрицу с отличным от нуля определителем.

Рассмотрим систему линейных функционалов $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$, где $\mathfrak{g}_\ell$ при $\ell = 1, \ldots, n$ – ограничение функционала $\mathfrak{f}_\ell$ на подпространство $\mathcal{W} = \operatorname{Span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$.

Заметим, что $\operatorname{Span}\{\mathfrak{g}_1, \ldots, \mathfrak{g}_n\} = \mathcal{W}^*$. В самом деле, включение $\operatorname{Span}\{\mathfrak{g}_1, \ldots, \mathfrak{g}_n\} \subset \mathcal{W}^*$ непосредственно следует из определения двойственного пространства. Рассмотрим теперь произвольный элемент $\mathfrak{g} \in \mathcal{W}^*$. Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r\}$ – базис в $\mathcal{W}$. Дополним его до базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_r, \boldsymbol{e}_{r+1}, \ldots, \boldsymbol{e}_n\}$ в пространстве $\mathcal{X}$. Пусть $\mathfrak{f} \in \mathcal{X}^*$ такой линейный функционал, что $\mathfrak{f}(\boldsymbol{e}_\ell) = \mathfrak{g}(\boldsymbol{e}_\ell)$ при $\ell = 1, \ldots, r$, а $\mathfrak{f}(\boldsymbol{e}_\ell) = 0$ при $\ell = r+1, \ldots, n$. Такой функционал существует, так как в $\mathcal{X}^*$ существуют функционалы, принимающие любые вещественные значения на элементах базиса пространства $\mathcal{X}$. Так как функционалы $\mathfrak{f}_1, \ldots, \mathfrak{f}_n$ образуют базис в $\mathcal{X}^*$, то $\mathfrak{f} = \sum_{s=1}^{n} \beta_s \mathfrak{f}_s$. Рассмотрим ограничения правой и левой частей этого равенства на $\mathcal{W}$. По определению функционала $\mathfrak{f}$ получаем, что $\mathfrak{g} = \sum_{s=1}^{n} \beta_s \mathfrak{g}_s$, откуда $\mathfrak{g} \in \operatorname{Span}\{\mathfrak{g}_1, \ldots, \mathfrak{g}_n\}$. Требуемое равенство доказано.

Напомним, что $\dim \mathcal{W} = \dim \mathcal{W}^* = r$. Выберем теперь из набора $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ систему из r линейно независимых векторов $\boldsymbol{x}_{j_1}, \ldots, \boldsymbol{x}_{j_r}$, а из набора $\mathfrak{g}_1, \ldots, \mathfrak{g}_n$ выберем систему из r линейно независимых функционалов $\mathfrak{g}_{t_1}, \ldots, \mathfrak{g}_{t_r}$. Из Предложения 4.7

$$\det \begin{pmatrix} \mathfrak{g}_{t_1}(\boldsymbol{x}_{j_1}) & \cdots & \mathfrak{g}_{t_1}(\boldsymbol{x}_{j_r}) \\ \multicolumn{3}{c}{\dotfill} \\ \mathfrak{g}_{t_r}(\boldsymbol{x}_{j_1}) & \cdots & \mathfrak{g}_{t_r}(\boldsymbol{x}_{j_r}) \end{pmatrix} \neq 0.$$

Остается заметить, что $\mathfrak{g}_\ell(\boldsymbol{x}_j) = \mathfrak{f}_\ell(\boldsymbol{x}_j)$ при $\ell = 1, \ldots, n$, $j = 1, \ldots, k$ по определению функционалов $\mathfrak{g}_\ell$. □

Рассмотрим систему линейных однородных уравнений

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0 \\ \dotfill \\ a_{n1}x_1 + \cdots + a_{nn}x_n = 0 \end{cases} \tag{4.3}$$

Если считать числа $x_1, \ldots, x_n$ координатами некоторого вектора $\boldsymbol{x}$ из n-мерного линейного пространства $\mathcal{X}$ в некотором определенном базисе в $\mathcal{X}$, то каждое выражение виде $a_{j1}x_1 + \cdots + a_{jn}x_n$, $j = 1, \ldots, n$, можно рассматривать как запись некоторого линейного функционала $\mathfrak{f}_j$. Таким образом, система (4.3) может быть записана в виде

$$\mathfrak{f}_1(\boldsymbol{x}) = 0, \ldots, \mathfrak{f}_m(\boldsymbol{x}) = 0, \tag{4.4}$$

где $\mathfrak{f}_1, \dots, \mathfrak{f}_m$ – некоторые линейные функционалы из $\mathcal{X}^*$. Верно и обратное наблюдение – каждая система уравнений вида (4.4) может быть записана в виде обычной системы линейных однородных уравнений вида (4.3). Для этого необходимо зафиксировать в пространстве $\mathcal{X}$ некоторый базис и расписать каждое из уравнений $\mathfrak{f}_j(\boldsymbol{x}) = 0$ в координатном виде. Опираясь на эти наблюдения можно сформулировать и доказать следующее утверждение.

Предложение 4.9. *Пусть $\mathcal{X}$ – n-мерное линейное пространство. Если система линейных функционалов $\mathfrak{f}_1, \dots, \mathfrak{f}_m$ содержит ровно r линейно независимых элементов, то размерность пространства решений системы (4.4) равна $n - r$.*

Любое подпространство $\mathcal{W} \subset \mathcal{X}$ является подпространством решений некоторой системы вида (4.4).

Доказательство. Пусть (без ограничения общности) элементы $\mathfrak{f}_1, \dots, \mathfrak{f}_r$ линейно независимы. Тогда остальные функционалы $\mathfrak{f}_j$ при $j \geqslant r$ будут линейными комбинациями элементов $\mathfrak{f}_1, \dots, \mathfrak{f}_r$. Следовательно, система (4.4) равносильна системе

$$\mathfrak{f}_1(\boldsymbol{x}) = 0, \dots, \mathfrak{f}_r(\boldsymbol{x}) = 0. \tag{4.5}$$

Пусть $\mathfrak{e} = \{\mathfrak{f}_1, \dots, \mathfrak{f}_r, \mathfrak{h}_{r+1}, \dots, \mathfrak{h}_n\}$ – базис в $\mathcal{X}^*$ и пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – базис в $\mathcal{X}$, двойственный к $\mathfrak{e}$. Тогда для любого $\boldsymbol{x} = \sum_{s=1}^n x_s \boldsymbol{e}_s$ система (4.5) будет иметь вид

$$x_1 = \dots = x_r = 0.$$

Т.е. пространство решений системы (4.5) состоит из векторов вида $\boldsymbol{x} = \sum_{s=r+1}^n x_s \boldsymbol{e}_s$ и имеет размерность $n - r$ так как вектора $\boldsymbol{e}_{r+1}, \dots, \boldsymbol{e}_n$ линейно независимы.

Для доказательства второго утверждения возьмем базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ в $\mathcal{W}$ такой, что $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \dots, \boldsymbol{e}_n\}$ – некоторый базис в пространстве $\mathcal{X}$. Вектор $\boldsymbol{x} \in \mathcal{X}$ принадлежит $\mathcal{W}$ если и только если $x_{m+1} = \dots = x_n = 0$, где $(x_1, \dots, x_n)^\top$ координаты вектора $\boldsymbol{x}$ в рассматриваемом базисе. Если $\{\mathfrak{f}_1, \dots, \mathfrak{f}_n\}$ – базис в $\mathcal{X}^*$, двойственный к рассматриваемому базису в $\mathcal{X}$, то уравнение $x_s = 0$ – это в точности уравнение $\mathfrak{f}_s(\boldsymbol{x}) = 0$. Итак,

$$\mathcal{W} = \{\boldsymbol{x} : \ \mathfrak{f}_{m+1}(\boldsymbol{x}) = \dots = \mathfrak{f}_n(\boldsymbol{x}) = 0\}. \quad \square$$

4.4. Общее понятие сопряженного оператора

Пусть $\mathcal{X}$ и $\mathcal{Y}$ – два конечномерных линейных пространства, а $\mathcal{X}^*$ и $\mathcal{Y}^*$ – соответствующие сопряженные пространства. Пусть, как и раньше, выражение $\langle \mathfrak{f} | \boldsymbol{x} \rangle$, $\mathfrak{f} \in \mathcal{X}^*$, $\boldsymbol{x} \in \mathcal{X}$ обозначает спаривание между $\mathcal{X}^*$ и $\mathcal{X}$, а выражение $\langle \mathfrak{g} | \boldsymbol{y} \rangle$ – спаривание между $\mathcal{Y}^*$ и $\mathcal{Y}$.

Пусть теперь определен линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{Y})$. Для любого линейного функционала $\mathfrak{f} \in \mathcal{Y}^*$ выражение $\langle \mathfrak{f} | \mathcal{A}\, \boldsymbol{x} \rangle$, $\boldsymbol{x} \in \mathcal{X}$ определяет некоторый линейный функционал на $\mathcal{X}$. В самом деле,

$$\langle \mathfrak{f} | \mathcal{A}(\alpha_1 \boldsymbol{x} + \alpha_2 \boldsymbol{x}_2) \rangle = \langle \mathfrak{f} | \alpha_1 \mathcal{A}\, \boldsymbol{x}_1 + \alpha_2 \mathcal{A}\, \boldsymbol{x}_2 \rangle = \alpha_1 \langle \mathfrak{f} | \mathcal{A}\, \boldsymbol{x}_1 \rangle + \alpha_2 \langle \mathfrak{f} | \mathcal{A}\, \boldsymbol{x}_2 \rangle$$

для любых $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$ и для любых чисел α_1, α_2, так как спаривание является линейным по обоим своим аргументам. Следовательно, определено отображение $\mathfrak{f} \mapsto \langle \mathfrak{f} | \mathcal{A}(\cdot) \rangle$ пространства $\mathcal{Y}^*$ линейных функционалов на $\mathcal{Y}$ в пространство $\mathcal{X}^*$ линейных функционалов на $\mathcal{X}$, связанное с рассматриваемым линейным оператором $\mathcal{A}$. Это отображение называют *сопряженным оператором* для оператора $\mathcal{A}$ и *обозначают* $\mathcal{A}^*$:

$$(\mathcal{A}^* \mathfrak{f})(\boldsymbol{x}) = \langle \mathfrak{f} | \mathcal{A}\, \boldsymbol{x} \rangle, \ \boldsymbol{x} \in \mathcal{X}.$$

Проверим, что $\mathcal{A}^*$ является линейным оператором. В самом деле,

$$\mathcal{A}^*(\alpha \mathfrak{f} + \beta \mathfrak{g})(\boldsymbol{x}) = \langle \alpha \mathfrak{f} + \beta \mathfrak{g} | \mathcal{A}\, \boldsymbol{x} \rangle = \alpha \langle \mathfrak{f} | \mathcal{A}\, \boldsymbol{x} \rangle + \beta \langle \mathfrak{g} | \mathcal{A}\, \boldsymbol{x} \rangle = \alpha\, \mathcal{A}^* \mathfrak{f}(\boldsymbol{x}) + \beta\, \mathcal{A}^* \mathfrak{g}(\boldsymbol{x})$$

для любых $\mathfrak{f}, \mathfrak{g} \in \mathcal{Y}^*$, для любых чисел α, β и для любого вектора $\boldsymbol{x} \in \mathcal{X}$.

Итак, для линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{Y})$ мы определили сопряженный линейный оператор $\mathcal{A}^* \in \mathfrak{L}(\mathcal{Y}^*, \mathcal{X}^*)$ (обратите внимание на "*направление действия*" оператора $\mathcal{A}^*$).

Пусть теперь $\mathcal{Y} = \mathcal{X}$. Непосредственно из определения сопряженного оператора и свойства линейности спаривания по обоим аргументам вытекают следующие свойства сопряженного оператора:

(1) Оператор $\mathcal{E}^*$ (где $\mathcal{E}$ – тождественный оператор на $\mathcal{X}$), совпадает с тождественным оператором на линейном пространстве $\mathcal{X}^*$.

(2) $(\alpha\,\mathcal{A})^* = \alpha\,\mathcal{A}^*$ для любого комплексного числа α. Это вытекает из того, что для любого $\mathfrak{f} \in \mathcal{X}^*$ и для любого $\boldsymbol{x} \in \mathcal{X}$ верны равенства $\langle(\alpha\,\mathcal{A})^*\mathfrak{f}|\boldsymbol{x}\rangle = \langle\mathfrak{f}|(\alpha\,\mathcal{A})\boldsymbol{x}\rangle = \langle\mathfrak{f}|\alpha(\mathcal{A}\,\boldsymbol{x})\rangle = \alpha\langle\mathfrak{f}|\,\mathcal{A}\,\boldsymbol{x}\rangle = \alpha\langle\mathcal{A}^*\,\mathfrak{f}|\boldsymbol{x}\rangle$.

(3) Для любых линейных операторов $\mathcal{A}, \mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеют место равенства

$$(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*, \qquad (\mathcal{A}\,\mathcal{B})^* = \mathcal{B}^*\,\mathcal{A}^*.$$

Проверим, например, последнее равенство. Для любого $\mathfrak{f} \in \mathcal{X}^*$ и для любого $\boldsymbol{x} \in \mathcal{X}$ справедливы следующие равенства, вытекающие из определения сопряженного оператора: $\langle(\mathcal{A}\,\mathcal{B})^*\mathfrak{f}|\boldsymbol{x}\rangle = \langle\mathfrak{f}|(\mathcal{A}\,\mathcal{B})\boldsymbol{x}\rangle = \langle\mathfrak{f}|\,\mathcal{A}(\mathcal{B}\,\boldsymbol{x})\rangle = \langle\mathcal{A}^*\,\mathfrak{f}|\,\mathcal{B}\,\boldsymbol{x}\rangle = \langle\mathcal{B}^*\,\mathcal{A}^*\,\mathfrak{f}|\boldsymbol{x}\rangle$.

Теорема 4.10. *Если в некотором базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$ оператор $\mathcal{A}$ имеет матрицу A, то сопряженный оператор в базисе $\{\mathfrak{e}^1, \ldots, \mathfrak{e}^n\}$ пространства $\mathcal{X}^*$, двойственном базису $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ имеет матрицу $A^\top$.*

Доказательство. Пусть $n = \dim\mathcal{X}$, а $A = (a_{jk})_{j,k=1}^n$. Тогда $\mathcal{A}\,\boldsymbol{e}_k = \sum_{j=1}^n a_{jk}\boldsymbol{e}_j$ и, учитывая двойственность базисов $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\{\mathfrak{e}^1, \ldots, \mathfrak{e}^n\}$, получаем, что

$$\langle\mathfrak{e}^m|\,\mathcal{A}\,\boldsymbol{e}_k\rangle = \sum_{j=1}^n a_{jk}\langle\mathfrak{e}^m|\boldsymbol{e}_j\rangle = a_{mk}.$$

Пусть теперь A^* – матрица оператор $\mathcal{A}^*$ в базисе $\{\mathfrak{e}^1, \ldots, \mathfrak{e}^n\}$ и $A^* = (a^*_{rm})_{r,m=1}^n$, т.е. $\mathcal{A}^*\,\mathfrak{e}^m = \sum_{r=1}^n a^*_{rm}\mathfrak{e}^r$ и, следовательно,

$$\langle\mathcal{A}^*\,\mathfrak{e}^m|\boldsymbol{e}_k\rangle = \sum_{r=1}^n a^*_{rm}\langle\mathfrak{e}^r|\boldsymbol{e}_k\rangle = a^*_{km}.$$

Так как, по определению сопряженного оператора $\langle\mathcal{A}^*\,\mathfrak{e}^m|\boldsymbol{e}_k\rangle = \langle\mathfrak{e}^m|\,\mathcal{A}\,\boldsymbol{e}_k\rangle$, то $a^*_{km} = a_{mk}$ при всех $k, m = 1, \ldots, n$. □

Выше была установлено, что если $\dim\mathcal{X} = n < \infty$, то существует канонический изоморфизм между пространствами $\mathcal{X}$ и $\mathcal{X}^{**}$. В терминах сопряженных операторов этот факт можно выразить следующим образом: $\mathcal{A}^{**} = \mathcal{A}$. В самом деле, так как пространство $\mathcal{X}$ рефлексивно, то всякий линейный функционал на $\mathcal{X}^*$ можно представить в виде $\mathfrak{f} \mapsto \langle\mathfrak{f}|\boldsymbol{x}\rangle$, где $\mathfrak{f}$ – произвольный элемент $\mathcal{X}^*$, а $\boldsymbol{x} \in \mathcal{X}$ – фиксированный вектор. В частности, $\langle\mathcal{A}^*\,\mathfrak{f}|\boldsymbol{x}\rangle = \langle\mathfrak{f}|\boldsymbol{y}\rangle$. По определению сопряженного оператора $\boldsymbol{y} = \mathcal{A}^{**}\,\boldsymbol{x}$, т.е.

$$\langle\mathfrak{f}|\,\mathcal{A}\,\boldsymbol{x}\rangle = \langle\mathcal{A}^*\,\mathfrak{f}|\boldsymbol{x}\rangle = \langle\mathfrak{f}|\,\mathcal{A}^{**}\,\boldsymbol{x}\rangle,$$

а это равенство, верное для любых $\mathfrak{f} \in \mathcal{X}^*$ и $\boldsymbol{x} \in \mathcal{X}$, показывает, что $\mathcal{A} = \mathcal{A}^{**}$.

РАЗДЕЛ 5

Линейные операторы, продолжение теории

5.1. Инвариантные подпространства линейных операторов

Пусть $\mathcal{X}$ – некоторое n-мерное линейное пространство, $\mathcal{A}$ – линейный оператор, действующий в пространстве $\mathcal{X}$.

Определение. *Подпространство $\mathcal{X}_1$ пространства $\mathcal{X}$ называют инвариантным подпространством относительно $\mathcal{A}$, или $\mathcal{A}$-инвариантным подпространством, если для любого элемента $\boldsymbol{x} \in \mathcal{X}_1$ выполнено $\mathcal{A}\,\boldsymbol{x} \in \mathcal{X}_1$. Короче говоря, если $\mathcal{A}\,\mathcal{X}_1 \subset \mathcal{X}_1$.*

Пример. $\operatorname{Ker}\mathcal{A}$ и $\operatorname{Im}\mathcal{A}$ являются $\mathcal{A}$-инвариантными подпространствами.

Предположим теперь, что оператор $\mathcal{A}$ имеет инвариантное подпространство $\mathcal{Y} \subset \mathcal{X}$. Выберем некоторый базис $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m$ в пространстве $\mathcal{Y}$ и дополним его векторами $\boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n$ до базиса $\boldsymbol{e}$ в $\mathcal{X}$. Так как $\mathcal{A}\,\mathcal{Y} \subset \mathcal{Y}$, то, по определению матрицы линейного оператора, матрица A оператора $\mathcal{A}$ в базисе $\boldsymbol{e}$ будет иметь вид

$$A = \begin{pmatrix} A_1 & B \\ 0 & A_2 \end{pmatrix},$$

где A_1 – это $m \times m$ матрица. Можно сказать, что A_1 – это матрица оператора $\mathcal{A}_{|\mathcal{Y}}$ – ограничения оператора $\mathcal{A}$ на подпространство $\mathcal{Y}$ – в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$. Нижний $(n-m) \times m$ блок матрицы A состоит из нулей, так как оператор $\mathcal{A}$ переводит подпространство $\mathcal{Y}$ в себя. Заметим теперь, что $m \times (n-m)$ матрица B обращается в нуль если и только если подпространство $\mathcal{W} := \operatorname{Span}\{\boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n\}$ также будет инвариантно относительно оператора $\mathcal{A}$.

Напомним, что $n \times n$-матрица называется *блочно-диагональной*, если она имеет вид

$$\begin{pmatrix} D_1 & 0 & \cdots & 0 \\ 0 & D_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & D_m \end{pmatrix}$$

$m \leqslant n$, где $D_1, \ldots, D_m$ – квадратные матрицы размера $n_k \times n_k$, $k = 1, \ldots, m$, $\sum_{k=1}^{m} n_k = n$, а символом 0 обозначены нулевые матрицы соответствующего размера. Для такой блочно-диагональной матрицы будем использовать обозначение $\operatorname{Diag}\{D_1, \ldots, D_m\}$. Если же все элементы некоторой матрицы за исключением элементов, стоящих на главной диагонали, равны нулю, то такая матрица будет называться *диагональной*. Диагональной матрица

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & d_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & d_n \end{pmatrix}$$

будем обозначаться символом $\operatorname{Diag}(d_1, d_2, \ldots, d_n)$.

Таким образом, если пространство $\mathcal{X}$ разложено в прямую сумму $\mathcal{X} = \mathcal{Y} \oplus \mathcal{W}$, оба слагаемых в которой – $\mathcal{A}$-инвариантные подпространства, то в $\mathcal{X}$ существует базис, в котором матрица A оператора $\mathcal{A}$ имеет блочный вид $A = \operatorname{Diag}\{A_1, A_2\}$, где A_1 и A_2 некоторые квадратные матрицы размера $m \times m$ и $(n-m) \times (n-m)$ соответственно. В таком случае также говорят, что оператор $\mathcal{A}$ есть *прямая сумма* операторов $\mathcal{A} = \mathcal{A}_{|\mathcal{Y}} \oplus \mathcal{A}_{|\mathcal{W}}$.

Совершенно аналогично понятию прямой сумму двух подпространств можно ввести и понятие прямой суммы любого конечного числа подпространств. Детали этого построения предлагается провести в качестве *упражнения*.

Абсолютно аналогично только что разобранному случаю разложения пространства $\mathcal{X}$ в прямую сумму двух $\mathcal{A}$-инвариантных подпространств, проверяется, что если пространство $\mathcal{X}$ разложено в прямую сумму p штук, $p \in \mathbb{N}$, $\mathcal{A}$-инвариантных подпространств $\mathcal{X} = \mathcal{Y}_1 \oplus \cdots \oplus \mathcal{Y}_p$, то существует базис, в котором матрица A оператора $\mathcal{A}$ имеет блочный вид $A = \mathrm{Diag}\{A_1, \ldots, A_p\}$, где A_j (при $j = 1, \ldots, p$) – некоторые квадратные матрицы.

Отметим, что при разложении пространства $\mathcal{X}$ в прямую сумму подпространств $\mathcal{X} = \mathcal{Y} \oplus \mathcal{W}$, где первое слагаемое является $\mathcal{A}$-инвариантным подпространством, второе слагаемое совершенно не обязано быть $\mathcal{A}$-инвариантным подпространством (в качестве *упражнения* предлагается привести пример такой ситуации в случае оператора дифференцирования).

Понятие о фактороператоре. Предположим теперь, что линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет инвариантное подпространство $\mathcal{U}$, т.е. $\mathcal{A}\mathcal{U} \subset \mathcal{U}$. Рассмотрим факторпространство $\mathcal{X}/\mathcal{U}$ и определим оператор $\widetilde{\mathcal{A}} : \mathcal{X}/\mathcal{U} \to \mathcal{X}/\mathcal{U}$ по правилу

$$\widetilde{\mathcal{A}}\widetilde{\boldsymbol{x}} = \widetilde{(\mathcal{A}\,\boldsymbol{x})},$$

где $\widetilde{\boldsymbol{x}} = \boldsymbol{x}_{\mathcal{U}} = \boldsymbol{x} + \mathcal{U}$ – класс эквивалентности из $\mathcal{X}/\mathcal{U}$, а $\boldsymbol{x}$ – некоторый его представитель. Проверим, что такое определение оператора $\widetilde{\mathcal{A}}$ не зависит от выбора представителя $\boldsymbol{x}$ в классе $\widetilde{\boldsymbol{x}}$. В самом деле, если $\boldsymbol{y}_1 \in \widetilde{\boldsymbol{x}}$ и $\boldsymbol{y}_2 \in \widetilde{\boldsymbol{x}}$, то $\boldsymbol{y}_1 - \boldsymbol{y}_2 \in \mathcal{U}$ и, так как $\mathcal{A}\mathcal{U} \subset \mathcal{U}$, то $\mathcal{A}\,\boldsymbol{y}_1 - \mathcal{A}\,\boldsymbol{y}_2 = \mathcal{A}(\boldsymbol{y}_1 - \boldsymbol{y}_2) \in \mathcal{U}$. Следовательно, $\mathcal{A}\,\boldsymbol{y}_1 + \mathcal{U} = \mathcal{A}\,\boldsymbol{y}_2 + \mathcal{U}$.

Легко проверяется (это оставляется в качестве *упражнения*, что $\widetilde{\mathcal{A}} \in \mathfrak{L}(\mathcal{X}/\mathcal{U}, \mathcal{X}/\mathcal{U})$ (т.е. $\widetilde{\mathcal{A}}$ является линейным оператором).

Определение. *Оператор* $\widetilde{\mathcal{A}} = \widetilde{\mathcal{A}}_{\mathcal{U}}$ *называется фактороператором.*

Заметим, что понятие фактороператора имеет смысл только в случае *инвариантного* подпространства $\mathcal{U}$.

Предположим теперь, что пространство $\mathcal{X}$ разложено в прямую сумму двух $\mathcal{A}$-инвариантных подпространств $\mathcal{X} = \mathcal{U} \oplus \mathcal{W}$, $\mathcal{A}\,\mathcal{U} \subset \mathcal{U}$, $\mathcal{A}\,\mathcal{W} \subset \mathcal{W}$. Тогда

$$\mathcal{A} = \mathcal{A}_{|\mathcal{U}} \oplus \mathcal{A}_{|\mathcal{W}}\,.$$

Напомним, что согласно Предложению 2.16 отображение $f : \boldsymbol{w} \mapsto \widetilde{\boldsymbol{w}}_{\mathcal{U}}$, $\boldsymbol{w} \in \mathcal{W}$, является изоморфизмом между $\mathcal{W}$ и $\mathcal{X}/\mathcal{U}$. При этом для любого $\boldsymbol{w} \in \mathcal{W}$ получаем, что

$$f(\mathcal{A}_{|\mathcal{W}}\,\boldsymbol{w}) = \mathcal{A}_{|\mathcal{W}}\,\boldsymbol{w} + \mathcal{U} = \widetilde{\mathcal{A}}_{\mathcal{U}}(\boldsymbol{w} + \mathcal{U}) = \widetilde{\mathcal{A}}_{\mathcal{U}}(f(\boldsymbol{w})),$$

т.е. $f \circ \mathcal{A}_{|\mathcal{W}} = \widetilde{\mathcal{A}}_{\mathcal{U}} \circ f$. Последнее равенство можно понимать так, что оператор $\widetilde{\mathcal{A}}_{\mathcal{U}}$ в определенном смысле “подобен” оператору $\mathcal{A}_{|\mathcal{W}}$.

5.2. Минимальный многочлен линейного оператора

Пусть $\mathcal{X}$ – конечномерное комплексное линейное пространство, $\dim \mathcal{X} = n$ и пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – линейный оператор на $\mathcal{X}$.

Пусть P – некоторый многочлен от переменной t с вещественными или комплексными коэффициентами

$$P(t) = p_0 t^m + p_1 t^{m-1} + \cdots + p_{m-1} t + p_m,$$

Тогда формула

$$P(\mathcal{A}) = p_0\,\mathcal{A}^m + p_1\,\mathcal{A}^{m-1} + \cdots + p_{m-1}\,\mathcal{A} + p_m\,\mathcal{E}\,.$$

корректно определяет линейный оператор $P(\mathcal{A}) \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. Напомним также, что многочлен P называется *нормализованным*, если $p_0 = 1$.

Определение. *Пусть P – многочлен. Если $P(\mathcal{A}) = \mathcal{O}$, то говорят, что P аннулирует оператор $\mathcal{A}$. Нормализованный многочлен минимальной степени, аннулирующий оператор $\mathcal{A}$ называется минимальным многочленом для $\mathcal{A}$. Минимальный многочлен оператора $\mathcal{A}$ обозначается символом $M_{\mathcal{A}}$,*

$$M_{\mathcal{A}}(t) = t^m + \mu_1 t^{m-1} + \cdots + \mu_m.$$

Предложение. *Для любого оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ существует минимальный многочлен.*

Доказательство. Рассмотрим пространство $\mathfrak{L}(\mathcal{X}, \mathcal{X})$ и заметим, что $\dim \mathfrak{L}(\mathcal{X}, \mathcal{X}) = n^2$. Выберем наименьшее число $m \in \mathbb{Z}_+$ так, что линейные операторы $\mathcal{E}, \mathcal{A}, \mathcal{A}^2, \ldots, \mathcal{A}^{m-1}$ линейно независимы в пространстве $\mathfrak{L}(\mathcal{X}, \mathcal{X})$, а линейные операторы $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^m$ уже линейно зависимы. Это возможно так как пространство $\mathfrak{L}(\mathcal{X}, \mathcal{X})$ конечномерно. Из линейной зависимости операторов $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^m$ вытекает, что существуют такие числа $\alpha_0, \ldots, \alpha_m$, не все из которых равны нулю, что

$$\alpha_0 \mathcal{A}^m + \alpha_1 \mathcal{A}^{m-1} + \cdots + \alpha_{m-1} \mathcal{A} + \alpha_m \mathcal{E} = \mathcal{O},$$

а из линейной независимости операторов $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^{m-1}$ следует, что $\alpha_0 \neq 0$. Тогда,

$$M_{\mathcal{A}}(t) = t^m + \mu_1 t^{m-1} + \cdots + \mu_m,$$

где $\mu_k = \alpha_k / \alpha_0$ при $k = 1, \ldots, m$. □

Предложение. *Пусть $M_{\mathcal{A}}(t) = t^m + \mu_1 t^{m-1} + \cdots + \mu_m$ – минимальный многочлен оператора $\mathcal{A}$. Если $\mu_m \neq 0$, то оператор $\mathcal{A}$ обратим.*

Доказательство. По определению минимального многочлена,

$$\mathcal{A}^m + \mu_1 \mathcal{A}^{m-1} + \cdots + \mu_{m-1} \mathcal{A} + \mu_m \mathcal{E} = \mathcal{O},$$

откуда (с учетом того, что $\mu_m \neq 0$)

$$\mathcal{E} = \mathcal{A} \left(-\frac{1}{\mu_m} \mathcal{A}^{m-1} - \frac{\mu_1}{\mu_n} \mathcal{A}^{m-2} - \cdots - \frac{\mu_{m-1}}{\mu_m} \mathcal{E} \right).$$

Выражение, стоящее в скобках, определяет оператор, обратный для $\mathcal{A}$. □

Заметим также, что справедливо следующее свойство, выражающее "минимальность" минимального многочлена в несколько иных терминах:

Предложение. *Любой многочлен P, аннулирующий оператор $\mathcal{A}$, имеет вид $P = QM_{\mathcal{A}}$ для некоторого многочлена Q.*

Доказательство. Пусть P – многочлен такой, что $P(\mathcal{A}) = \mathcal{O}$. Существуют многочлены Q и R такие, что $P = QM_{\mathcal{A}} + R$ и $\deg R < \deg M_{\mathcal{A}}$. Так как $M_{\mathcal{A}}$ – минимальный многочлен для $\mathcal{A}$, то

$$\mathcal{O} = P(\mathcal{A}) = Q(\mathcal{A}) M_{\mathcal{A}}(\mathcal{A}) + R(\mathcal{A}) = R(\mathcal{A}).$$

Полученное равенство противоречит тому, что $M_{\mathcal{A}}$ – минимальный многочлен для $\mathcal{A}$. Следовательно, $R \equiv 0$ и многочлен P, аннулирующий оператор $\mathcal{A}$ делится на $M_{\mathcal{A}}$ без остатка. □

Установим существование инвариантных подпространств у линейного оператора, действующего в комплексном или вещественном линейном пространстве.

Теорема 5.1. *Пусть $\mathcal{A}$ – линейный оператор, действующий в комплексном или вещественном линейном пространстве $\mathcal{X}$.*

1. *Если $\mathcal{X}$ – комплексное линейное пространство, то $\mathcal{A}$ всегда имеет одномерное инвариантное подпространство.*

2. *Если $\mathcal{A}$ – вещественное линейное пространство, то $\mathcal{A}$ всегда имеет одномерное или двумерное инвариантное подпространство.*

Доказательство. Пусть $\mathcal{X}$ – комплексное линейное пространство. Многочлен $M_{\mathcal{A}}$ (минимальный многочлен оператора $\mathcal{A}$), рассматриваемый как многочлен из $\mathbb{C}[t]$ всегда имеет корни (возможно, комплексные). Пусть λ – один из корней многочлена $M_{\mathcal{A}}$. Тогда $M_{\mathcal{A}}(t) = (t-\lambda)Q(t)$, где Q – некоторый многочлен, причем $Q(\mathcal{A}) \neq \mathcal{O}$. Следовательно, найдется вектор $\boldsymbol{y} \in \mathcal{X}$ такой, что $\boldsymbol{x} := Q(\mathcal{A})\boldsymbol{y} \neq \mathbf{0}$. Заметим далее, что

$$(\mathcal{A} - \lambda\,\mathcal{E})\boldsymbol{x} = (\mathcal{A} - \lambda\,\mathcal{E})Q(\mathcal{A})\boldsymbol{y} = M_{\mathcal{A}}(\mathcal{A})\boldsymbol{y} = \mathbf{0}.$$

Таким образом, $\mathcal{A}\,\boldsymbol{x} = \lambda\boldsymbol{x}$. Рассмотрим совокупность векторов $\mathcal{X}(\lambda) := \{\alpha\boldsymbol{x} :\ \alpha \in \mathbb{C}\}$. Ясно, что $\mathcal{X}_(\lambda)$ является подпространством пространства $\mathcal{X}$. Пусть теперь $\alpha \in \mathbb{C}$ – произвольно. Тогда

$$\mathcal{A}(\alpha\boldsymbol{x}) = \alpha\,\mathcal{A}(\boldsymbol{x}) = \alpha\lambda\boldsymbol{x} \in \mathcal{X}(\lambda).$$

Следовательно, подпространство $\mathcal{X}(\lambda)$ является $\mathcal{A}$-инвариантным.

Пусть теперь $\mathcal{X}$ – вещественное пространство. Снова рассмотрим минимальный многочлен $M_{\mathcal{A}}$ оператора $\mathcal{A}$. Теперь $M_{\mathcal{A}} \in \mathbb{R}[t]$ и возможны два случая. Во-первых, многочлен $M_{\mathcal{A}}$ может иметь вещественный корень λ. Тогда также, как и в случае комплексного пространства, устанавливается, что существует вектор $\boldsymbol{x} \neq 0$ такой, что $\mathcal{A}\,\boldsymbol{x} = \lambda\boldsymbol{x}$, а совокупность векторов $\mathcal{X}(\lambda) = \{\alpha\boldsymbol{x} :\ \alpha \in \mathbb{R}\}$ является $\mathcal{A}$-инвариантным подпространством.

Пусть теперь минимальный многочлен $M_{\mathcal{A}}$ не имеет вещественных корней. Тогда многочлен $M_{\mathcal{A}}$ может быть представлен в виде

$$M_{\mathcal{A}}(t) = (t^2 - at - b)Q(t).$$

Как и в предыдущем случае $Q(\mathcal{A}) \neq \mathcal{O}$ и, следовательно, найдется вектор $\boldsymbol{y}$ такой, что $\boldsymbol{x} := Q(\mathcal{A})\boldsymbol{y} \neq \mathbf{0}$. Далее

$$(\mathcal{A}^2 - a\,\mathcal{A} - b\,\mathcal{E})\boldsymbol{x} = (\mathcal{A}^2 - a\,\mathcal{A} - b\,\mathcal{E})Q(\mathcal{A})\boldsymbol{y} = M_{\mathcal{A}}(\mathcal{A})\boldsymbol{y} = \mathbf{0},$$

откуда $\mathcal{A}^2\,\boldsymbol{x} = a\,\mathcal{A}\,\boldsymbol{x} + b\boldsymbol{x}$.

Проверим, что вектор $\boldsymbol{y}$ можно выбрать так, что соответствующие векторы $\boldsymbol{x}$ и $\mathcal{A}\,\boldsymbol{x}$ будут линейно независимы. В самом деле, предположим, что это не так. Тогда для любого вектора $\boldsymbol{v}$, для которого $Q(\mathcal{A})\boldsymbol{v} \neq \mathbf{0}$ найдется такое $\beta \in \mathbb{R}$, что

$$Q_1(\mathcal{A})\boldsymbol{v} = \mathcal{A}(Q(\mathcal{A})\boldsymbol{v}) + \beta Q(\mathcal{A})\boldsymbol{v} = \mathbf{0},$$

где $Q_1(t) = (t+\beta)Q(t)$. Если же $Q(\mathcal{A})\boldsymbol{u} = \mathbf{0}$, то и $Q_1(\mathcal{A})\boldsymbol{u} = \mathbf{0}$. Таким образом, $Q_1(\mathcal{A})\boldsymbol{z} = \mathbf{0}$ для любого вектора $\boldsymbol{z} \in \mathcal{X}$. Но степень многочлена Q_1 на единицу меньше, чем степень *минимального* многочлена $M_{\mathcal{A}}$ оператора $\mathcal{A}$. Полученное противоречие и доказывает существование вектора $\boldsymbol{y}$ такого, что $\boldsymbol{x} = Q(\mathcal{A})\boldsymbol{y} \neq \mathbf{0}$, а векторы $\boldsymbol{x}$ и $\mathcal{A}\,\boldsymbol{x}$ линейно независимы.

Рассмотрим подпространство $\mathrm{Span}\{\boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}\}$. Для любого вектора $\boldsymbol{w} = \alpha_1\boldsymbol{x} + \alpha_2\,\mathcal{A}\,\boldsymbol{x}$ из этого подпространства получаем

$$\mathcal{A}\,\boldsymbol{w} = (\alpha_1 + a\alpha_2)\,\mathcal{A}\,\boldsymbol{x} + b\alpha_2\boldsymbol{x} \in \mathrm{Span}\{\boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}\},$$

т.е. $\mathrm{Span}\{\boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}\}$ является двумерным инвариантным подпространством для $\mathcal{A}$. □

5.3. Характеристический многочлен, собственные числа и собственные векторы линейных операторов

Пусть $\mathcal{X}$ – некоторое n-мерное линейное пространство, $\mathcal{A}$ – линейный оператор, действующий в пространстве $\mathcal{X}$.

Определение. *Выражение*

$$P_{\mathcal{A}}(\lambda) := \det(\mathcal{A} - \lambda\,\mathcal{E}),$$

где символом $\mathcal{E}$ обозначен единичный оператор, называется характеристическим многочленом оператора $\mathcal{A}$.

Проверим первым делом, что выражение $P_{\mathcal{A}}$ в самом деле является многочленом от переменной λ. Пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ – некоторый базис в $\mathcal{X}$ и пусть A – матрица оператора $\mathcal{A}$ в этом базисе. Тогда, по определению определителя линейного оператора,

$$P_{\mathcal{A}}(\lambda) = \det[A - \lambda E],$$

где E – единичная матрица. Т.е., $P_{\mathcal{A}}$ – это характеристический многочлен матрицы A и, следовательно, $P_{\mathcal{A}}$ является многочленом от переменной λ и $\deg P_{\mathcal{A}} = n$.

Пусть

$$P_{\mathcal{A}}(\lambda) = P_n\lambda^n + P_{n-1}\lambda^{n-1} + \cdots + P_1\lambda + P_0.$$

Так как значение определителя $\det(\mathcal{A} - \lambda\,\mathcal{E})$ не зависит от выбора базиса, то и все коэффициенты $P_n, \dots, P_1, P_0$ характеристического многочлена оператора $\mathcal{A}$ на зависят от выбора базиса, а зависят только от оператора $\mathcal{A}$. Например, коэффициент P_{n-1} характеристического многочлена, который может быть записан в виде $P_{n-1} = a_{11} + \cdots + a_{nn}$, где $A = (a_{jk})_{j,k=1}^{n}$ – матрица оператора $\mathcal{A}$ в некотором базисе (проверка приведенного равенства для коэффициента P_{n-1} оставляется в качестве *упражнения*), называется *следом* линейного оператора $\mathcal{A}$ и обозначается $\operatorname{Tr}\mathcal{A}$.

Определение. *Уравнение $P_{\mathcal{A}}(\lambda) = 0$ называют характеристическим уравнением для оператора $\mathcal{A}$.*

Определение. *Число λ называется собственным числом или собственным значением оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$, если найдется ненулевой элемент $\boldsymbol{x}_\lambda \in \mathcal{X}$ такой, что $\mathcal{A}\,\boldsymbol{x}_\lambda = \lambda\boldsymbol{x}_\lambda$. При этом соответствующий элемент $\boldsymbol{x}_\lambda$ называется собственным вектором оператора $\mathcal{A}$, соответствующим собственному числу λ.*

Теорема 5.2. *Число λ является собственным значением оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ тогда и только тогда, когда λ является корнем характеристического многочлена $P_{\mathcal{A}}$ оператора $\mathcal{A}$.*

Доказательство. Пусть λ – собственное значение оператора $\mathcal{A}$ и пусть $\boldsymbol{x}$ – соответствующий этому собственному значению собственный вектор. Тогда, из равенства $\mathcal{A}\,\boldsymbol{x} = \lambda\boldsymbol{x}$ вытекает, что $(\mathcal{A} - \lambda\,\mathcal{E})\boldsymbol{x} = \mathbf{0}$. Так как $\boldsymbol{x}$ – ненулевой элемент, то из последнего равенства вытекает, что $\operatorname{Ker}(\mathcal{A} - \lambda\,\mathcal{E}) \neq \mathbf{0}$, т.е. $\dim(\operatorname{Ker}(\mathcal{A} - \lambda\,\mathcal{E})) \geqslant 1$. Из этого вытекает, что $\dim(\operatorname{Im}(\mathcal{A} - \lambda\,\mathcal{E})) \leqslant (n-1)$. По определению, размерность образа линейного оператора равняется его рангу, и, следовательно, ранг матрицы оператора $\mathcal{A} - \lambda\,\mathcal{E}$ (в любом базисе пространства $\mathcal{X}$) меньше n. А это, в свою очередь, означает, что определитель этой матрицы равен нулю и, следовательно, $\det(\mathcal{A} - \lambda\,\mathcal{E}) = 0$. Т.е., собственное число λ является корнем уравнения $P_{\mathcal{A}}(\lambda) = 0$.

Пусть теперь λ – корень уравнения $P_{\mathcal{A}}(\lambda) = 0$. Тогда ранг оператора $\mathcal{A} - \lambda\,\mathcal{E}$ меньше n и, следовательно, $\dim(\operatorname{Im}(\mathcal{A} - \lambda\,\mathcal{E})) \leqslant (n-1)$. Отсюда вытекает, что $\operatorname{Ker}(\mathcal{A} - \lambda\,\mathcal{E}) \neq \mathbf{0}$, т.е., что существует ненулевой элемент $\boldsymbol{x} \in \operatorname{Ker}(\mathcal{A} - \lambda\,\mathcal{E})$. А это, по определению, означает, что λ – собственное значение оператора $\mathcal{A}$ с собственным вектором $\boldsymbol{x}$. ☐

Замечание. Так как всякий многочлен степени n имеет n комплексных корней (с учетом кратности), то любой линейный оператор, действующий в n-мерном (комплексном) линейном пространстве имеет n собственных значений (с учетом кратности).

Важно понимать, что последнее замечание относится только к комплексным пространствам и непосредственно связано с фактом алгебраической замкнутости поля комплексных чисел.

Поясним, что будет происходить в случае вещественного пространства. В этом случае собственными значениями будут только вещественные корни характеристического уравнения. Приведем пример оператора, не имеющего собственных векторов. Рассмотрим в пространстве $\mathbb{R}^2$ оператор с матрицей

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

где $\theta \in (0, \pi)$ (в качестве *упражнения* предлагается выяснить геометрический смысл этого оператора) и заметим, что его характеристическое уравнение $\lambda^2 - 2\lambda\cos\theta + 1 = 0$ не имеет вещественных корней.

Предложение 5.3. *Для того, чтобы матрица A линейного оператора $\mathcal{A}$ в некотором базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ была диагональной необходимо и достаточно, чтобы все базисные векторы были собственными векторами для оператора $\mathcal{A}$.*

Доказательство. Если все базисные векторы $\boldsymbol{e}_k$, $k = 1, \ldots, n$ являются собственными векторами оператора $\mathcal{A}$, то $\mathcal{A}\,\boldsymbol{e}_k = \lambda_k \boldsymbol{e}_k$ при $k = 1, \ldots, n$ и, следовательно, матрица оператора $\mathcal{A}$ имеет в базисе $\boldsymbol{e}$ вид $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$. Обратно, если матрица A оператора $\mathcal{A}$ имеет в некотором базисе $\boldsymbol{e}$ вид $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$, то $\mathcal{A}\,\boldsymbol{e}_k = \lambda_k \boldsymbol{e}_k$, $k = 1, \ldots, n$ и, так как все базисные векторы ненулевые, то они являются собственными векторами для $\mathcal{A}$. □

Предложение 5.4. *Пусть собственные значения $\lambda_1, \ldots, \lambda_k$ линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ различны. Тогда соответствующие этим собственным значениям собственные векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ линейно независимы.*

Доказательство. Проверим вначале, что два собственных вектора $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$, отвечающие двум различным собственным значениям λ_1 и λ_2 линейно независимы. Пусть $\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 = \mathbf{0}$. Тогда $\alpha_1 \lambda_1 \boldsymbol{x}_1 + \alpha_2 \lambda_2 \boldsymbol{x}_2 = \mathcal{A}(\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2) = \mathcal{A}\,\mathbf{0} = \mathbf{0}$ и

$$\mathbf{0} = (\alpha_1 \lambda_1 \boldsymbol{x}_1 + \alpha_2 \lambda_2 \boldsymbol{x}_2) - \lambda_2(\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2) = \alpha_1(\lambda_1 - \lambda_2)\boldsymbol{x}_1.$$

Так как $\boldsymbol{x}_1 \neq \mathbf{0}$ (как собственный вектор) и $\lambda_1 \neq \lambda_2$ по условию, то $\alpha_1 = 0$. Отсюда $\alpha_2 \boldsymbol{x}_2 = \mathbf{0}$ и, так как $\boldsymbol{x}_2 \neq \mathbf{0}$, то $\alpha_2 = 0$. Таким образом, из $\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 = \mathbf{0}$ вытекает, что $\alpha_1 = \alpha_2 = 0$, что, в свою очередь, означает линейную независимость векторов $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$.

Предположим теперь, что при некотором натуральном $m > 1$ система собственных векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$, соответствующих различным собственным значениям $\lambda_1, \ldots, \lambda_m$ линейно независима.

Добавим к этой системе собственный вектор $\boldsymbol{x}_{m+1}$, соответствующий собственному значению λ_{m+1}, отличному от $\lambda_1, \ldots, \lambda_m$ и докажем, что получившаяся система из $m+1$ собственного вектора $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m, \boldsymbol{x}_{m+1}$ будет линейно независимой. Рассмотрим линейную комбинацию $\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \boldsymbol{x}_{m+1}$ и предположим, что $\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \boldsymbol{x}_{m+1} = \mathbf{0}$. Тогда

$$\alpha_1 \lambda_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \lambda_{m+1} \boldsymbol{x}_{m+1} = \mathcal{A}(\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \boldsymbol{x}_{m+1}) = \mathcal{A}\,\mathbf{0} = \mathbf{0}.$$

Кроме того, $\lambda_{m+1}(\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \boldsymbol{x}_{m+1}) = \mathbf{0}$ и, следовательно,

$$\sum_{j=1}^{m} (\lambda_j - \lambda_{m+1}) \alpha_j \boldsymbol{x}_j = \mathbf{0}.$$

Так как $\lambda_j \neq \lambda_{m+1}$ при $j = 1, \ldots, m$ по условию теоремы и так как векторы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m$ линейно независимы в силу сделанного выше предположения, то из последнего равенства вытекает, что $\alpha_1 = \cdots = \alpha_m = 0$. Таким образом, $\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \boldsymbol{x}_{m+1} = \alpha_{m+1} \boldsymbol{x}_{m+1}$ и, так как $\boldsymbol{x}_{m+1} \neq \mathbf{0}$ как собственный вектор, то $\alpha_{m+1} = 0$. Таким образом, из равенства $\alpha_1 \boldsymbol{x}_1 + \cdots + \alpha_{m+1} \boldsymbol{x}_{m+1} = \mathbf{0}$ вытекает, что $\alpha_1 = \cdots = \alpha_{m+1} = 0$ и система векторов $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_m, \boldsymbol{x}_{m+1}$ линейно независима.

Повторив последовательно приведенное рассуждение при $m = 2, 3, \ldots, k-1$ мы получим требуемое утверждение. □

Из Предложения 5.4 вытекает, что если оператор $\mathcal{A}$ имеет n различных собственных значений, то существует базис пространства $\mathcal{X}$, в котором матрица этого оператора имеет диагональный вид.

Упражнение. Проверить, что множество $\mathcal{X}_\lambda$ собственных векторов оператора $\mathcal{A}$, соответствующих собственному значению λ, дополненное нулевым вектором, является $\mathcal{A}$-инвариантным подпространством.

Пример 5.5. Пусть зафиксирован некоторый базис 3-х мерного линейного пространства и пусть линейный оператор $\mathcal{A}$ имеет в этом базисе матрицу

$$\begin{pmatrix} 1 & 1-i & 0 \\ 1+i & 3 & i \\ 0 & -i & 1 \end{pmatrix}.$$

Вычислим характеристический многочлен этого оператора:

$$P_{\mathcal{A}}(\lambda) = \det(\mathcal{A} - \lambda\,\mathcal{E}) = \lambda(1-\lambda)(\lambda-4).$$

Таким образом, оператор $\mathcal{A}$ имеет три различных собственный значения $\lambda_1 = 0$, $\lambda_2 = 1$ и $\lambda_3 = 4$. Найдем собственный вектор $\boldsymbol{x}_1$, соответствующий собственному значению $\lambda_1 = 0$. Так как любой собственный вектор $\boldsymbol{x}_\lambda$, соответствующий собственному значению λ удовлетворяет уравнению $(\mathcal{A} - \lambda\,\mathcal{E})\boldsymbol{x}_\lambda = \mathbf{0}$, то для компонент ξ_1, η_1, ζ_1 вектора $\boldsymbol{x}_1 = (\xi_1, \eta_1, \zeta_1)^\top$ получаем систему линейных уравнений

$$\begin{pmatrix} 1 & 1-i & 0 \\ 1+i & 3 & i \\ 0 & -i & 1 \end{pmatrix} \begin{pmatrix} \xi_1 \\ \eta_1 \\ \zeta_1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Для компонент $\xi_{2,3}, \eta_{2,3}, \zeta_{2,3}$ собственных векторов $\boldsymbol{x}_2$ и $\boldsymbol{x}_3$ получаем, соответственно, системы линейных уравнений

$$\begin{pmatrix} 0 & 1-i & 0 \\ 1+i & 2 & i \\ 0 & -i & 0 \end{pmatrix} \begin{pmatrix} \xi_2 \\ \eta_2 \\ \zeta_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{и} \quad \begin{pmatrix} -3 & 1-i & 0 \\ 1+i & -1 & i \\ 0 & -i & -3 \end{pmatrix} \begin{pmatrix} \xi_3 \\ \eta_3 \\ \zeta_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Решая эти системы получаем

$$\boldsymbol{x}_1 = (1+i, -i, 1)^\top, \quad \boldsymbol{x}_2 = (1, 0, i-1)^\top, \quad \boldsymbol{x}_3 = (1+i, 3i, 1)^\top.$$

5.4. Теорема Гамильтона-Кэли

Имеет место следующее важное утверждение.

Теорема 5.6 (теорема Гамильтона-Кэли)**.** *Для любого линейного оператора $\mathcal{A}$, действующего в комплексном линейном пространстве $\mathcal{X}$ верно равенство $P_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$.*

Доказательство. Для доказательство этого утверждения мы заметим, первым делом, что для любого линейного оператора $\mathcal{A}$, действующего в комплексном линейном пространстве $\mathcal{X}$, существует базис, в котором матрица этого оператора имеет верхнетреугольный вид.

Для доказательства этого факта мы покажем, что $\mathcal{A}$ обладает инвариантным подпространством размерности $n-1$. В самом деле, рассмотрим сопряженный оператор $\mathcal{A}^*$ (действующий в пространстве $\mathcal{X}^*$). Пусть λ – собственное значение $\mathcal{A}^*$, а $\mathfrak{f} \in \mathcal{X}^*$ – соответствующий λ собственный элемент оператора $\mathcal{A}$. Так как $\mathfrak{f} \not\equiv 0$, то $\dim \operatorname{Ker} \mathfrak{f} = n-1$. Далее, если $\boldsymbol{x} \in \operatorname{Ker} \mathfrak{f}$, то

$$\mathfrak{f}(\mathcal{A}\,\boldsymbol{x}) = \langle \mathfrak{f} |\, \mathcal{A}\,\boldsymbol{x} \rangle = \langle \mathcal{A}^*\, \mathfrak{f} | \boldsymbol{x} \rangle = \lambda \langle \mathfrak{f} | \boldsymbol{x} \rangle = \lambda \mathfrak{f}(\boldsymbol{x}) = 0,$$

т.е. $\mathcal{A}\,\boldsymbol{x} \in \operatorname{Ker} \mathfrak{f}$. Таким образом, $\operatorname{Ker} \mathfrak{f}$ – искомое $(n-1)$-мерное инвариантное подпространство для $\mathcal{A}$.

Теперь мы можем доказать существование базиса, в котором матрица оператора $\mathcal{A}$ имеет верхнетреугольный вид по индукции. В самом деле, существует подпространство $\mathcal{W}$, $\dim \mathcal{W} = n-1$, такое, что $\mathcal{A}\,\mathcal{W} \subset \mathcal{W}$. По предположению индукции, в подпространстве $\mathcal{W}$ существует такой базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{n-1}\}$, что

$$\mathcal{A}\,\boldsymbol{e}_j = \lambda_j \boldsymbol{e}_j + \boldsymbol{w}_j, \qquad \text{где} \quad \boldsymbol{w}_j \in \operatorname{Span}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{j-1}\}.$$

Пусть $\boldsymbol{e}_n \in \mathcal{X} \setminus \mathcal{W}$ – произвольный вектор. Тогда $\mathcal{X} = \mathrm{Span}\{\mathcal{W}, \boldsymbol{e}_n\}$. Остается заметить, что $\mathcal{A}\,\boldsymbol{e}_n = \lambda_n \boldsymbol{e}_n + \boldsymbol{w}'$, где $\boldsymbol{w}' \in \mathcal{W}$. Таким образом, в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ матрица оператора $\mathcal{A}$ имеет верхнетреугольный вид.

Дальнейшее рассуждение будем вести в этом базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Определим подпространства

$$\mathcal{W}_k := \mathrm{Span}\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_{n-k-1}, \boldsymbol{e}_{n-k}\}$$

так, что

$$\mathcal{X} = \mathcal{W}_0 \supset \mathcal{W}_1 \supset \cdots \supset \mathcal{W}_{n-1} \supset \mathcal{W}_n = \{\mathbf{0}\}.$$

Заметим, что $(\mathcal{A} - \lambda_{n-k}\,\mathcal{E})\boldsymbol{e}_{n-k} \in \mathcal{W}_{k+1}$, откуда $(\mathcal{A} - \lambda_{n-k}\,\mathcal{E})\mathcal{W}_k \subset \mathcal{W}_{k+1}$. Далее,

$$\begin{aligned} P_{\mathcal{A}}(\mathcal{A})\mathcal{X} &= (\mathcal{A} - \lambda_1\,\mathcal{E})(\mathcal{A} - \lambda_2\,\mathcal{E})\cdots(\mathcal{A} - \lambda_{n-1}\,\mathcal{E})(\mathcal{A} - \lambda_n\,\mathcal{E})\mathcal{W}_0 \subset \\ &\subset (\mathcal{A} - \lambda_1\,\mathcal{E})(\mathcal{A} - \lambda_2\,\mathcal{E})\cdots(\mathcal{A} - \lambda_{n-1}\,\mathcal{E})\mathcal{W}_1 \subset \cdots \subset (\mathcal{A} - \lambda_1\,\mathcal{E})\mathcal{W}_{n-1} = \{\mathbf{0}\}, \end{aligned}$$

а равенство $P_{\mathcal{A}}(\mathcal{A})\mathcal{X} = \mathbf{0}$ эквивалентно тому, что $P_{\mathcal{A}}(\mathcal{A}) = \mathcal{O}$. □

5.5. Циклические векторы линейных операторов

Нильпотентные операторы. Напомним, что линейный оператор $\mathcal{A}$ называется *нильпотентным*, если $\mathcal{A}^m = \mathcal{O}$ для некоторого $m \in \mathbb{Z}_+$. Минимальное число $m \in \mathbb{Z}_+$ с таким свойством называется *индексом нильпотентности*.

Предложение 5.7. *Пусть $\mathcal{A}$ – нильпотентный оператор с индексом нильпотентности m. Тогда $m \leqslant n$.*

Доказательство. Так как оператор $\mathcal{A}^{m-1} \neq \mathcal{O}$ (по определению индекса нильпотентности), то найдется вектор $\boldsymbol{x} \in \mathcal{X}$ такой, что $\mathcal{A}^{m-1}\boldsymbol{x} \neq 0$. Проверим, что векторы $\boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}, \ldots, \mathcal{A}^{m-1}\boldsymbol{x}$ линейно независимы. Для этого рассмотрим линейную комбинацию

$$\boldsymbol{w} := \alpha_0\boldsymbol{x} + \alpha_1\,\mathcal{A}\,\boldsymbol{x} + \cdots + \alpha_{m-1}\,\mathcal{A}^{m-1}\boldsymbol{x}$$

и предположим, что $\boldsymbol{w} = \mathbf{0}$. Так как $\mathcal{A}^{m-1}\boldsymbol{w} = \mathbf{0}$ и так как $\mathcal{A}^k = \mathcal{O}$ при $k \geqslant m$, то $\alpha_0\,\mathcal{A}^{m-1}\boldsymbol{x} = \mathbf{0}$, откуда $\alpha_0 = 0$. Далее, $\mathcal{A}^{m-2}\boldsymbol{w} = \mathbf{0}$, но

$$\mathcal{A}^{m-2}\boldsymbol{w} = \alpha_0\,\mathcal{A}^{m-2}\boldsymbol{x} + \alpha_1\,\mathcal{A}^{m-1}\boldsymbol{x} = \alpha_1\,\mathcal{A}^{m-1}\boldsymbol{x}.$$

Следовательно $\alpha_1 = 0$. Рассматривая последовательно равенства $\mathcal{A}^{m-3}\boldsymbol{w} = \mathbf{0}$, ..., $\mathcal{A}\,\boldsymbol{w} = \mathbf{0}$ и $\boldsymbol{w} = \mathbf{0}$ получим, что $\alpha_2 = \alpha_3 = \cdots = \alpha_{m-1} = 0$. Таким образом, векторы $\boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}, \ldots, \mathcal{A}^{m-1}\boldsymbol{x}$ линейно независимы. Отсюда немедленно вытекает, что $m \leqslant n$. □

Пусть $\mathcal{X}$ – линейное пространство размерности $\dim\mathcal{X} = n$ и пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – нильпотентный оператор с индексом нильпотентности n. Тогда, на основании доказанного выше Предложение 5.7, найдется такой вектор $\boldsymbol{v}$, что векторы $\boldsymbol{v}, \mathcal{A}\,\boldsymbol{v}, \ldots, \mathcal{A}^{n-1}\boldsymbol{v}$ являются линейно независимыми.

В этом случае система векторов $\{\mathcal{A}^{n-1}\boldsymbol{v}, \mathcal{A}^{n-2}\boldsymbol{v}, \ldots, \mathcal{A}\,\boldsymbol{v}, \boldsymbol{v}\}$ будет базисом в $\mathcal{X}$ (так как состоит из n линейно независимых векторов), а $\mathcal{X} = \mathrm{Span}\{\boldsymbol{v}, \mathcal{A}\,\boldsymbol{v}, \ldots, \mathcal{A}^{n-1}\boldsymbol{v}\}$.

Циклические векторы. Ситуация, когда существует вектор $\boldsymbol{v} \in \mathcal{X}$ такой, что $\mathcal{X} = \mathrm{Span}\{\boldsymbol{v}, \mathcal{A}\,\boldsymbol{v}, \ldots, \mathcal{A}^{n-1}\boldsymbol{v}\}$ возможна не только в случае, когда оператор $\mathcal{A}$ является нильпотентным.

Определение. *Вектор $\boldsymbol{v}$ называется циклическим для оператора $\mathcal{A}$, действующего в n-мерном линейном пространстве $\mathcal{X}$, если $\mathcal{X} = \mathrm{Span}\{\boldsymbol{v}, \mathcal{A}\,\boldsymbol{v}, \ldots, \mathcal{A}^{n-1}\boldsymbol{v}\}$. Само пространство $\mathcal{X}$ в этом случае называется циклическим относительно $\mathcal{A}$.*

Ясно, что если $\boldsymbol{v}$ – циклический вектор для $\mathcal{A}$, то векторы $\{\boldsymbol{v}, \mathcal{A}\,\boldsymbol{v}, \ldots, \mathcal{A}^{n-1}\boldsymbol{v}\}$ являются линейно независимыми.

Пусть $\boldsymbol{v}$ – циклический вектор для оператора $\mathcal{A}$, действующего в n-мерном пространстве $\mathcal{X}$. Пусть также

$$\boldsymbol{e} = \{\boldsymbol{e}_1 = \mathcal{A}^{n-1}\boldsymbol{v}, \boldsymbol{e}_2 = \mathcal{A}^{n-2}\boldsymbol{v}, \ldots, \boldsymbol{e}_{n-1} = \mathcal{A}\,\boldsymbol{v}, \boldsymbol{e}_n = \boldsymbol{v}\}$$

– соответствующий циклический базис. В этом случае

$$[\mathcal{A}]_{\boldsymbol{e}} = \begin{pmatrix} a_1 & 1 & 0 & \cdots & 0 \\ a_2 & 0 & 1 & \cdots & 0 \\ \multicolumn{5}{c}{\dots\dots\dots\dots\dots\dots} \\ a_{n-1} & 0 & 0 & \cdots & 1 \\ a_n & 0 & 0 & \cdots & 0 \end{pmatrix}, \tag{5.1}$$

где $a_1, \dots, a_n$ – коэффициенты разложения вектора $\mathcal{A}^n \boldsymbol{v}$ по базису $\boldsymbol{e}$, т.е. $\mathcal{A}^n \boldsymbol{v} = \sum_{k=1}^n a_k \mathcal{A}^{n-k} \boldsymbol{v}$. Заметим, что в общем случае вектор $\mathcal{A}^n \boldsymbol{v}$ *не обязан* равняться нулю (это так если оператор $\mathcal{A}$ является нильпотентным с индексом n).

Имеет место и обратное наблюдение. Пусть оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет в некотором базисе $\boldsymbol{e} = (\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ пространства $\mathcal{X}$ матрицу вида (5.1). Тогда $\mathcal{A}^{n-k} \boldsymbol{e}_n = \boldsymbol{e}_k$ при $k = 1, \dots, n$ и, согласно определению, вектор $\boldsymbol{e}_n$ будет циклическим для $\mathcal{A}$.

Пример 5.8. Пусть оператор $\mathcal{A}$ действует в пространстве $\mathbb{R}[t]_n$ многочленов с вещественными коэффициентами степени не выше n по правилу

$$\mathcal{A} : p_n t^n + p_{n-1} t^{n-1} + \dots + p_1 t + p_0 \mapsto p_0 t^n + (p_n + p_0) t^{n-1} + \dots + (p_2 + p_0) t + (p_1 + p_0).$$

Ясно, что $\mathcal{A}$ – линейный оператор (проверка оставляется в качестве ***упражнения***) и, что в базисе $\{1, t, \dots, t^n\}$ оператор $\mathcal{A}$ имеет матрицу вида (5.1) в которой первый столбец состоит из единиц. Легко проверить, что элемент t^n будет циклическим для $\mathcal{A}$.

Предложение 5.9. *Пусть n-мерное пространство $\mathcal{X}$ является циклическим относительно линейного оператора $\mathcal{A}$. Тогда элементы $a_1, \dots, a_n$ матрицы (5.1) оператора $\mathcal{A}$ не зависят от выбора циклического базиса.*

Доказательство. Пусть $\boldsymbol{v}$ – некоторый циклический вектор для $\mathcal{A}$, а

$$\boldsymbol{e} = \{\boldsymbol{e}_1 = \mathcal{A}^{n-1} \boldsymbol{v}, \boldsymbol{e}_2 = \mathcal{A}^{n-2} \boldsymbol{v}, \dots, \boldsymbol{e}_{n-1} = \mathcal{A} \boldsymbol{v}, \boldsymbol{e}_n = \boldsymbol{v}\}$$

– соответствующий циклический базис. Пусть также $\mathcal{A}^n \boldsymbol{v} = \sum_{k=1}^n a_k \mathcal{A}^{n-k} \boldsymbol{v}$. Рассмотрим нормализованный многочлен

$$P(t) = t^n - \sum_{k=1}^n a_k t^{n-k}$$

и заметим, что

$$P(\mathcal{A})(\boldsymbol{e}_k) = P(\mathcal{A})(\mathcal{A}^{n-k} \boldsymbol{v}) = \mathcal{A}^{n-k} \left(\mathcal{A}^n \boldsymbol{v} - \sum_{k=1}^n a_k \mathcal{A}^{n-k} \boldsymbol{v} \right) = \mathcal{A}^{n-k} \mathbf{0} = 0$$

Так как векторы $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ образуют базис в $\mathcal{X}$, то из последнего равенства вытекает, что $P(\mathcal{A}) \boldsymbol{x} = \mathbf{0}$ для любого вектора $\boldsymbol{x} \in \mathcal{X}$ и, следовательно, $P(\mathcal{A}) = \mathcal{O}$.

Предположим теперь, что существует какой-то нормализованный многочлен $Q(t) = t^m + q_1 t^{m-1} + \dots + q_m$ степени меньше $m < n$ такой, что $Q(\mathcal{A}) = \mathcal{O}$. Тогда $Q(\mathcal{A}) \boldsymbol{v} = \mathbf{0}$ и

$$\mathcal{A}^m \boldsymbol{v} + q_1 \mathcal{A}^{m-1} \boldsymbol{v} + \dots + q_{m-1} \mathcal{A} \boldsymbol{v} + q_m \boldsymbol{v} = \mathbf{0},$$

что противоречит линейной независимости векторов $\{\boldsymbol{v}, \mathcal{A} \boldsymbol{v}, \dots, \mathcal{A}^{n-1} \boldsymbol{v}\}$.

Следовательно, многочлен P – это минимальный многочлен для оператора $\mathcal{A}$, а числа $a_1, \dots, a_n$ – это коэффициенты этого многочлена. Таким образом, эти числа зависят только от оператора $\mathcal{A}$, но не от выбора циклического базиса. □

***Следствие* 5.10.** *Конечномерное пространство $\mathcal{X}$ является циклическим относительно линейного оператора $\mathcal{A}$ если и только если размерность $\mathcal{X}$ совпадает со степенью минимального многочлена оператора $\mathcal{A}$.*

Доказательство. При доказательстве Предложения 5.9 было показано, что если пространство $\mathcal{X}$ является циклическим относительно $\mathcal{A}$, то степень минимального многочлена совпадает с размерностью пространства $\mathcal{X}$. Обратное утверждение проверяется следующим образом. Пусть степень минимального многочлена оператора $\mathcal{A}$ равняется $n = \dim \mathcal{X}$. Тогда операторы $\mathcal{E}, \mathcal{A}, \ldots, \mathcal{A}^n$ линейно зависимы, а операторы $\mathcal{A}, \mathcal{A}, \ldots, \mathcal{A}^{n-1}$ линейно независимы. Из линейной независимости операторов $\mathcal{A}, \mathcal{A}, \ldots, \mathcal{A}^{n-1}$ вытекает существование такого вектора $\boldsymbol{v} \neq \mathbf{0}$, что векторы $\boldsymbol{v}, \mathcal{A}\,\boldsymbol{v}, \ldots, \mathcal{A}^{n-1}\,\boldsymbol{v}$ линейно независимы. Так как этих векторов n штук, то они образуют базис в пространстве $\mathcal{X}$. При этом ясно, что вектор $\boldsymbol{v}$ и будет циклическим для $\mathcal{A}$. $\square$

Таким образом, из Теоремы Гамильтона-Кэли и из Следствия 5.10 вытекает, что конечномерное комплексное линейное пространство $\mathcal{X}$ является циклическим относительно линейного оператора $\mathcal{A}$ если и только если минимальный многочлен этого оператора совпадает с его нормализованным характеристическим многочленом.

РАЗДЕЛ 6

Канонический вид линейный операторов

Пусть $\mathcal{X}$ – n-мерное комплексное линейное пространства и пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – действующий в пространстве $\mathcal{X}$ линейный оператор.

Естественно возникает вопрос об отыскании такого базиса в $\mathcal{X}$, в котором матрица A оператора $\mathcal{A}$ имеет как можно более простой вид. В силу ряда причин, которые станут ясны несколько позже, задачу об отыскании такого базиса называют также задачей о приведении оператора $\mathcal{A}$ к *каноническому виду*.

Как уже отмечалось ранее, если оператор $\mathcal{A}$ имеет n штук различных собственных значений $\lambda_1 \neq \cdots \neq \lambda_n$, то соответствующие этим собственным значениям собственные векторы линейно независимы. Они образуют базис, в котором матрица оператора $\mathcal{A}$ приобретает диагональный вид $\mathrm{Diag}\{\lambda_1, \dots, \lambda_n\}$ (см. Предложение 5.3).

Кроме того, в разделе 5.1 было показано, что если пространство $\mathcal{X}$ разложено в прямую сумму p штук, $p \in \mathbb{N}$, $\mathcal{A}$-инвариантных подпространств $\mathcal{X} = \mathcal{Y}_1 \oplus \cdots \oplus \mathcal{Y}_p$, то существует базис, в котором матрица оператора $\mathcal{A}$ имеет блочный вид $\mathrm{Diag}\{A_1, \dots, A_p\}$, где A_j (при $j = 1, \dots, p$) – некоторые квадратные матрицы.

Пусть теперь среди собственных значений $\lambda_1, \dots, \lambda_n$ оператора $\mathcal{A}$ есть кратные. Нашей ближайшей целью является доказательство существования базиса, в котором матрица оператора $\mathcal{A}$ имеет блочную структуру вида

$$\mathrm{Diag}\{\Lambda_1, \dots, \Lambda_\ell\}, \tag{6.1}$$

где $n_k \times n_k$-матрицы Λ_k при $k = 1, \dots, \ell$ – это матрицы вида

$$\Lambda_k = \begin{pmatrix} \lambda_k & 1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_k & 1 & \cdots & 0 & 0 \\ \hdotsfor{6} \\ 0 & 0 & 0 & \cdots & \lambda_k & 1 \\ 0 & 0 & 0 & \cdots & 0 & \lambda_k \end{pmatrix}$$

Матрицы вида Λ_k называют *жордановыми клетками* матрицы A, а про матрицу (6.1), составленную из жордановых клеток, говорят, что она записана в *жордановой форме*. Для удобства будем обозначать матрицы Λ_k через $J(\lambda_k, n_k)$ указывая размер соответствующей жордановой клетки и определяющее ее число λ_k.

Ясно, что числа $\lambda_1, \dots, \lambda_\ell$, определяющие жордановы клетки матрицы (6.1) – это собственные числа соответствующей матрицы.

6.1. Собственные и корневые инвариантные подпространства

Пусть λ – собственное значение оператора $\mathcal{A}$, т.е. $\det(\mathcal{A} - \lambda\,\mathcal{E}) = 0$. Вспомним данное выше понятие *собственного вектора* линейного оператора: элемент $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{x} \neq \boldsymbol{0}$, называется собственным вектором оператора $\mathcal{A}$, соответствующим собственному значению λ, если $\mathcal{A}\,\boldsymbol{x} = \lambda \boldsymbol{x}$.

Ведем подпространство

$$\mathcal{X}_{\mathcal{A}}(\lambda) := \big\{\boldsymbol{x} \in \mathcal{X} : \ (\mathcal{A} - \lambda\,\mathcal{E})\boldsymbol{x} = \boldsymbol{0}\big\},$$

содержащее все собственные векторы, соответствующих данному собственному значению λ оператора $\mathcal{A}$. Как уже отмечалось выше, $\mathcal{A}\,\mathcal{X}_{\mathcal{A}}(\lambda) \subset \mathcal{X}_{\mathcal{A}}(\lambda)$, т.е. $\mathcal{X}_{\mathcal{A}}(\lambda)$ – это $\mathcal{A}$-инвариантное подпространство.

Подпространство $\mathcal{X}_{\mathcal{A}}(\lambda)$ называется собственным подпространством, соответствующим собственному значению λ оператора $\mathcal{A}$.

Введем еще одно понятие, связанное с понятием собственного числа линейного оператора $\mathcal{A}$.

Определение. *Элемент $\boldsymbol{x} \in \mathcal{X}$ называется присоединенным элементом (или присоединенный вектором) оператора $\mathcal{A}$, отвечающим собственному значению λ, если для некоторого целого числа $m \geqslant 1$ выполняются соотношения*

$$(\mathcal{A} - \lambda\, \mathcal{E})^m \boldsymbol{x} \neq \mathbf{0}, \qquad (\mathcal{A} - \lambda\, \mathcal{E})^{m+1} \boldsymbol{x} = \mathbf{0}.$$

Число m называется порядком присоединенного элемента $\boldsymbol{x}$.

Заметим, что если $\boldsymbol{x}$ – присоединенный вектор порядка m для оператора $\mathcal{A}$, то вектор $(\mathcal{A} - \lambda\, \mathcal{E})^m \boldsymbol{x}$ является собственным вектором для оператора $\mathcal{A}$.

Рассмотрим совокупность векторов

$$\widehat{\mathcal{X}}_{\mathcal{A}}(\lambda) := \big\{\boldsymbol{x} \in \mathcal{X} :\ (\mathcal{A} - \lambda\, \mathcal{E})^k \boldsymbol{x} = \mathbf{0} \text{ для некоторого } k \in \mathbb{Z}_+\big\}$$

и проверим, что она является подпространством пространства $\mathcal{X}$. В самом деле, если $\boldsymbol{x}_1 \in \widehat{\mathcal{X}}_{\mathcal{A}}(\lambda)$ и $\boldsymbol{x}_2 \in \widehat{\mathcal{X}}_{\mathcal{A}}(\lambda)$, то существуют такие $k_1 \in \mathbb{Z}_+$ и $k_2 \in \mathbb{Z}_+$, что $\mathcal{A} - \lambda\, \mathcal{E})^{k_1} \boldsymbol{x}_1 = \mathbf{0}$ и $\mathcal{A} - \lambda\, \mathcal{E})^{k_2} \boldsymbol{x}_2 = \mathbf{0}$. Тогда, если $k = \max\{k_1, k_2\}$, а $\alpha_1, \alpha_2 \in \mathbb{C}$, то

$$(\mathcal{A} - \lambda\, \mathcal{E})^k (\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2) = \alpha_1 (\mathcal{A} - \lambda\, \mathcal{E})^k \boldsymbol{x}_1 + \alpha_2 (\mathcal{A} - \lambda\, \mathcal{E})^k \boldsymbol{x}_2 = \mathbf{0},$$

откуда вытекает, что $\alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 \in \widehat{\mathcal{X}}_{\mathcal{A}}(\lambda)$.

Подпространство $\widehat{\mathcal{X}}_{\mathcal{A}}(\lambda)$ называется корневым подпространством, соответствующим собственному значению λ оператора $\mathcal{A}$.

Ясно, что $\mathcal{X}_{\mathcal{A}}(\lambda) \subset \widehat{\mathcal{X}}_{\mathcal{A}}(\lambda)$, но это включение, в общем случае, является строгим. Примером такой ситуации является случай нильпотентного оператора с индексом нильпотентности n (проверка деталей оставляется в качестве *упражнения*).

Если из контекста ясно, о каком операторе $\mathcal{A}$ идет речь, то вместо обозначений $\mathcal{X}_{\mathcal{A}}(\lambda)$ и $\widehat{\mathcal{X}}_{\mathcal{A}}(\lambda)$ для собственных и корневых инвариантных подпространств оператора $\mathcal{A}$ использовать, соответственно, обозначения $\mathcal{X}(\lambda)$ и $\widehat{\mathcal{X}}(\lambda)$.

Пусть $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p\}$ – базис в $\widehat{\mathcal{X}}(\lambda)$. Тогда существуют такие числа $r_1, \ldots, r_p \in \mathbb{Z}_+$, что $(\mathcal{A} - \lambda\, \mathcal{E})^{r_j} \boldsymbol{x}_j = \mathbf{0}$ при $j = 1, \ldots, p$. Из этого следует, что $(\mathcal{A} - \lambda\, \mathcal{E})^r \boldsymbol{x} = \mathbf{0}$ при $r = \max\{r_1, \ldots, r_p\}$ для любого $\boldsymbol{x} \in \widehat{\mathcal{X}}(\lambda)$. Следовательно, ограничение оператора $\mathcal{A} - \lambda\, \mathcal{E}$ на подпространство $\widehat{\mathcal{X}}(\lambda)$ является *нильпотентным* оператором.

Так как $\dim \widehat{\mathcal{X}}(\lambda) \leqslant n$, то

$$\widehat{\mathcal{X}}(\lambda) = \{\boldsymbol{x} \in \mathcal{X} :\ (\mathcal{A} - \lambda\, \mathcal{E})^n \boldsymbol{x} = \mathbf{0}\}.$$

6.2. Канонический вид линейного оператора

Предположим, что характеристический многочлен $P = P_{\mathcal{A}}$ оператора $\mathcal{A}$ имеет вид

$$P(t) = \prod_{j=1}^{m} (t - \lambda_j)^{k_j},$$

где $\lambda_1, \ldots, \lambda_m$, $m \leqslant n$, – все характеристические числа оператора $\mathcal{A}$ кратности $k_1, \ldots, k_m$ соответственно; при этом $\sum_{j=1}^m k_j = n$.

Имеет место следующее утверждение

Теорема 6.1. *В приведенных выше обозначениях справедливы следующие утверждения.*

1. *Пространство $\widehat{\mathcal{X}}(\lambda_j)$ является $\mathcal{A}$-инвариантным,* $\dim \widehat{\mathcal{X}}(\lambda_j) = k_j$ *и*

$$\mathcal{X} = \widehat{\mathcal{X}}(\lambda_1) \oplus \cdots \oplus \widehat{\mathcal{X}}(\lambda_m).$$

2. *Оператор* $\mathcal{A} - \lambda_j \mathcal{E}$ *нильпотентен на* $\widehat{\mathcal{X}}(\lambda_j)$ *и невырожден на* $\mathcal{X}_j := \mathcal{X} \ominus \widehat{\mathcal{X}}(\lambda_j)$. *Число* λ_j *– единственное собственное значение оператора* $\mathcal{A}_{|\mathcal{X}_j}$.

Доказательство. Рассмотрим многочлены

$$P_j(t) = \frac{P_{\mathcal{A}}(t)}{(t-\lambda_j)^{k_j}}.$$

Ясно, что НОД$(P_1, \ldots, P_m) = 1$ (мы предполагаем, что читатель знаком с понятием наибольшего общего делителя многочленов). Тогда существуют такие многочлены $Q_j \in \mathbb{C}[t]$, $j = 1, \ldots, m$, что

$$\sum_{j=1}^{m} P_j(t) Q_j(t) = 1.$$

Читатель, не знакомый с основными фактами теории многочленов, может воспринимать последнюю формулу как известный факт, он будет доказан в курсе алгебры в 3-м семестре.

Определим подпространства

$$\mathcal{Y}_j := P_j(\mathcal{A}) Q_j(\mathcal{A}) \mathcal{X}, \qquad j = 1, \ldots, m$$

и заметим, что эти подпространства являются $\mathcal{A}$-инвариантными. В самом деле,

$$\mathcal{A}\, \mathcal{Y}_j = P_j(\mathcal{A}) Q_j(\mathcal{A})\, \mathcal{A}\, \mathcal{X} \subset P_j(\mathcal{A}) Q_j(\mathcal{A}) \mathcal{X}$$

так как оператор $\mathcal{A}$ коммутирует с $Q(\mathcal{A})$ для любого $Q \in \mathbb{C}[t]$.

Из теоремы Гамильтона-Кэли следует, что

$$(\mathcal{A} - \lambda_j \mathcal{E})^{k_j} \mathcal{Y}_j = P(\mathcal{A}) Q_j(\mathcal{A}) \mathcal{X} = \{\mathbf{0}\},$$

откуда $\mathcal{Y}_j \subset \widehat{\mathcal{X}}(\lambda_j)$. Далее, так как $\sum_{j=1}^{m} P_j(t) Q_j(t) = 1$, то

$$\mathcal{E} = \sum_{j=1}^{m} P_j(\mathcal{A}) Q_j(\mathcal{A}).$$

По определению подпространств $\mathcal{Y}_j$ из последнего равенства вытекает, что $\mathcal{X} = \mathcal{Y}_1 + \cdots + \mathcal{Y}_m$ и, тем более, что

$$\mathcal{X} = \widehat{\mathcal{X}}_1 + \cdots + \widehat{\mathcal{X}}_m.$$

Пусть $\boldsymbol{x} \in \widehat{\mathcal{X}}(\lambda_j) \cap \mathcal{X}_j$. Тогда $(\mathcal{A} - \lambda_j \mathcal{E})^n \boldsymbol{x} = \mathbf{0}$. Далее, представим вектор $\boldsymbol{x} \in \mathcal{X}_j$ как сумму $\boldsymbol{x} = \sum_{s \neq j} \boldsymbol{x}_s$, где векторы $\boldsymbol{x}_s \in \widehat{\mathcal{X}}(\lambda_s)$ при $s = 1, \ldots, m$, $s \neq j$. Далее $(\mathcal{A} - \lambda_s \mathcal{E})^n \boldsymbol{x}_s = \mathbf{0}$ для любого s и, следовательно,

$$\left(\prod_{s \neq j} (\mathcal{A} - \lambda_s \mathcal{E})^n \right) \boldsymbol{x} = \mathbf{0}.$$

Так как многочлены $(t - \lambda_j)^n$ и $\prod_{s \neq j} (t - \lambda_s)^n$ взаимно просты, то существуют такие многочлены $F_1, F_2 \in \mathbb{C}[t]$, что

$$(t - \lambda_j)^n F_1(t) + \left(\prod_{s \neq j} (t - \lambda_s)^n \right) F_2(t) = 1.$$

Из последнего равенства получаем, что

$$\boldsymbol{x} = F_1(\mathcal{A})(\mathcal{A} - \lambda_j \mathcal{E})^n \boldsymbol{x} + F_2(\mathcal{A}) \left(\prod_{s \neq j} (\mathcal{A} - \lambda_s \mathcal{E})^n \right) \boldsymbol{x} = \mathbf{0}.$$

Таким образом пространства $\widehat{\mathcal{X}}(\lambda_j)$ и $\mathcal{X}_j$ не пересекаются. Следовательно,

$$\mathcal{X} = \widehat{\mathcal{X}}(\lambda_1) \oplus \cdots \oplus \widehat{\mathcal{X}}(\lambda_m).$$

Одновременно доказано, что

$$\mathcal{Y}_j = \widehat{\mathcal{X}}(\lambda_j) = P_j(\mathcal{A})Q_j(\mathcal{A})\mathcal{X}$$

и, следовательно, $\widehat{\mathcal{X}}(\lambda_j)$ является $\mathcal{A}$-инвариантным подпространством. В частности,

$$(\mathcal{A} - \lambda_j\,\mathcal{E})^n \widehat{\mathcal{X}}(\lambda_j) = \mathbf{0},$$

т.е. оператор $\mathcal{A} - \lambda_j\,\mathcal{E}$ нильпотентен на $\widehat{\mathcal{X}}(\lambda_j)$.

Минимальным многочленом для оператора $\mathcal{A}$, рассматриваемого как оператор на подпространстве $\widehat{\mathcal{X}}(\lambda_j)$ будет некоторый делитель многочлена $(t-\lambda_j)^{k_j}$. Из этого, в частности, вытекает, что число λ_j – единственное собственное значение оператора $A_{|\widehat{\mathcal{X}}(\lambda_j)}$. Матрицей оператора $\mathcal{A}$ в базисе, составленном из базисов подпространств $\widehat{\mathcal{X}}(\lambda_j)$, $j = 1, \ldots, m$, имеет вид $\mathrm{Diag}\{A_1, \ldots, A_m\}$, где A_j, $j = 1, \ldots, m$, – матрица порядка $k'_j = \dim \widehat{\mathcal{X}} * (\lambda_j)$ с единственным собственным значением λ_j. Так как $P_{A_j}(t) = (t-\lambda_j)^{k'_j}$ и так как $P = \prod_{j=1}^m P_{A_j}$, то $k'_j = k_j$ при всех $j = 1, \ldots, m$.

Проверим невырожденность оператора $\mathcal{A} - \lambda_j\,\mathcal{E}$ на $\mathcal{X}_j = \mathcal{X} \ominus \widehat{\mathcal{X}}(\lambda_j)$. Если предположить, что $\mathrm{Ker}(\mathcal{A} - \lambda_j\,\mathcal{E}) \cap \mathcal{X}_j \neq \{\mathbf{0}\}$, то существует такой вектор $\boldsymbol{x} \in \mathcal{X}_j$, что $\boldsymbol{x} \neq \mathbf{0}$ и $\mathcal{A}\,\boldsymbol{x} = \lambda_j \boldsymbol{x}$. Таким образом, в пространстве $\mathcal{X}_j$ существует собственный вектор для $\mathcal{A}$, соответствующий собственному числу λ_j. Но это противоречит тому, что характеристический многочлен оператора $\mathcal{A}$ на $\mathcal{X}_j$ – это многочлен $\prod_{s\neq j}(t-\lambda_s)^{k_s}$, среди корней которого нет λ_j. □

С учетом только что доказанной теоремы нам остается изучить вопрос о приведении к жордановой форме оператора $\mathcal{A}$ в случае, когда $\mathcal{A}$ имеет единственное собственное значение λ, а $(\mathcal{A} - \lambda\,\mathcal{E})^m = \mathcal{O}$ для некоторого $m \leqslant n$. Таким образом, заменой $\mathcal{B} = \mathcal{A} - \lambda\,\mathcal{E}$ задача сведена к задача о нахождении канонического вида для нильпотентного оператора $\mathcal{B}$ индекса нильпотентности m.

Для произвольного вектора $\boldsymbol{x} \in \mathcal{X}$ рассмотрим подпространство

$$\mathcal{B}\{\boldsymbol{x}\} := \mathrm{Span}\{\boldsymbol{x}, \mathcal{B}\,\boldsymbol{x}, \ldots, \mathcal{B}^{m(\boldsymbol{x})-1}\},$$

где $m(\boldsymbol{x}) \leqslant m$ – наименьшее натуральное число такое, что $\mathcal{B}^{m(\boldsymbol{x})}\,\boldsymbol{x} = \mathbf{0}$. Подпространство $\mathcal{B}\{\boldsymbol{x}\}$ называется *циклическим подпространством*, порожденным оператором $\mathcal{B}$ и вектором $\boldsymbol{x}$.

Предложение 6.2. *Пусть $\mathcal{B}$ – нильпотентный оператор, действующий в линейном пространстве $\mathcal{X}$, $\dim \mathcal{X} = n$. Тогда $\mathcal{X}$ представимо в виде прямой суммы циклических подпространств, порожденных оператором $\mathcal{B}$.*

Доказательство. Воспользуемся индукцией по размерности пространства $\mathcal{X}$. Так как оператор $\mathcal{B}$ является нильпотентным, то существует такой базис, в котором его матрица имеет верхнетреугольный вид с нулями на главной диагонали. Пусть $\mathcal{U} \subset \mathcal{X}$ – подпространство пространства $\mathcal{X}$, являющееся линейной оболочкой первых $n-1$ векторов указанного базиса. Тогда $\mathcal{B}\,\mathcal{X} \subset \mathcal{U}$ (в частности. $\mathcal{U}$ является $\mathcal{B}$-инвариантным подпространством).

По предположению индукции существуют такие векторы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_k$, что

$$\mathcal{U} = \mathcal{B}\{\boldsymbol{e}_1\} \oplus \cdots \oplus \mathcal{B}\{\boldsymbol{e}_k\}.$$

обозначим $m_j = m(\boldsymbol{e}_j)$ при $j = 1, \ldots, k$ так, что $m_j \leqslant m$, где m – индекс нильпотентности $\mathcal{B}$. Без ограничения общности можно считать, что $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_k$.

Выберем далее вектор $\boldsymbol{x}$ так, чтобы $\mathcal{X} = \mathrm{Span}\{\boldsymbol{x}, \mathcal{U}\}$, а $\mathcal{B}\,\boldsymbol{x} = \sum_{j=1}^k \alpha_j \boldsymbol{e}_j$. Для этого возьмем произвольный вектор $\boldsymbol{y}$ такой, что $\mathcal{X} = \mathrm{Span}\{\boldsymbol{y}, \mathcal{U}\}$. Далее, так как $\mathcal{B}\,\boldsymbol{y} \in \mathcal{U}$, то найдется такой вектор $\boldsymbol{u} \in \mathcal{U}$, что $\mathcal{B}\,\boldsymbol{y} = \sum_{j=1}^k \alpha_j \boldsymbol{e}_j + \mathcal{B}\,\boldsymbol{u}$. Осталось положить $\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{u}$.

Итак, пусть $\mathcal{X} = \mathrm{Span}\{\boldsymbol{x}, \mathcal{U}\}$, а $\mathcal{B}\,\boldsymbol{x} = \sum_{j=1}^k \alpha_j \boldsymbol{e}_j$. Возможны два случая.

Во-первых, может оказаться, что $\alpha_1 = \cdots = \alpha_k = 0$. Тогда очевидно, что $\mathcal{B}\{\boldsymbol{x}\} = \operatorname{Span}\{\boldsymbol{x}\}$, а

$$\mathcal{X} = \mathcal{U} \oplus \mathcal{B}\{\boldsymbol{x}\} = \mathcal{B}\{\boldsymbol{e}_1\} \oplus \cdots \oplus \mathcal{B}\{\boldsymbol{e}_k\} \oplus \mathcal{B}\{\boldsymbol{x}\}.$$

Второй случай состоит в том, что существует такое $\ell < k$, что $\alpha_1 = \cdots = \alpha_{\ell-1} = 0$, но $\alpha_\ell \neq 0$. Тогда $\mathcal{B}\,\boldsymbol{x} = \sum_{j=\ell}^{k} \alpha_j \boldsymbol{e}_j$. В этом случае рассмотрим набор векторов

$$\widetilde{\boldsymbol{e}}_j = \boldsymbol{e}_j \quad \text{при } j \neq \ell, \qquad \widetilde{\boldsymbol{e}}_\ell = \frac{1}{\alpha_\ell}\boldsymbol{x}$$

и определим числа $\beta_j = \alpha_j/\alpha_\ell$ при $j = 1, \ldots, k$. Заметим при этом, что вектор

$$\boldsymbol{w} := \mathcal{B}\,\widetilde{\boldsymbol{e}}_\ell = \boldsymbol{e}_\ell + \sum_{j=\ell+1}^{k} \beta_j \boldsymbol{e}_j.$$

Так как векторы $\boldsymbol{e}_j$ были упорядочены таким образом, что $m_1 \geqslant m_2 \geqslant \cdots \geqslant m_k$, то $\mathcal{B}^{m_\ell}\,\boldsymbol{w} = \boldsymbol{0}$. Так как $\mathcal{U}$ представляет собой прямую сумму циклических подпространств $\mathcal{B}\{\boldsymbol{e}_j\}$, то $\mathcal{B}^{m_\ell - 1}\,\boldsymbol{w} \neq 0$.

Для завершения доказательства остается заметить, что сумма $\mathcal{B}\{\boldsymbol{w}\} + \sum_{j\neq\ell} \mathcal{B}\{\boldsymbol{e}_j\}$ также является прямой суммой и совпадает с $\mathcal{U}$. Проверка этого факта оставляется в качестве несложного *упражнения*.

Так как $\mathcal{B}\{\boldsymbol{w}\} \subset \mathcal{B}\{\widetilde{\boldsymbol{e}}_\ell\}$, то

$$\mathcal{X} = \mathcal{B}\{\widetilde{\boldsymbol{e}}_1\} \oplus \cdots \oplus \mathcal{B}\{\widetilde{\boldsymbol{e}}_k\}. \quad \square$$

Остается заметить, что матрица ограничения нильпотентного оператора $\mathcal{B}$ на любое его циклическое подпространство имеет вид жордановой клетки (этот факт явно вытекает из вида матрицы оператора в циклическом базисе, установленного ранее). Из этого наблюдения и Предложения 6.2 вытекает, что матрица нильпотентного оператора может быть приведена к жордановой нормальной форме.

Определение числа жордановых клеток заданного размера при записи линейного оператора в каноническом виде. Пусть теперь $\mathcal{A}$ – произвольный линейный оператор, рассмотренный выше. Получим формулы для вычисления числа $N(\lambda, m)$ жордановых клеток порядка m, соответствующих собственному значению λ, в матрице жордановой нормальной формы оператора $\mathcal{A}$.

Пусть $\mathcal{X} = \widehat{\mathcal{X}}(\lambda) \oplus \mathcal{X}'$ – прямая сумма корневого подпространства $\widehat{\mathcal{X}}(\lambda)$, соответствующего собственному значению λ, и его дополнения $\mathcal{X}'$ и пусть

$$\widehat{\mathcal{X}}(\lambda) = (\mathcal{A} - \lambda\,\mathcal{E})\{\boldsymbol{e}_1\} \oplus \cdots \oplus (\mathcal{A} - \lambda\,\mathcal{E})\{\boldsymbol{e}_k\}.$$

Последнее разложение имеет место так как оператор $\mathcal{A} - \lambda\,\mathcal{E}$ нильпотентен на $\widehat{\mathcal{X}}(\lambda)$ в силу утверждения Теоремы 6.1. Кроме того $\mathcal{X}' = \sum_{\nu\neq\lambda} \widehat{\mathcal{X}}(\nu)$, где (прямая) сумма берется по всем собственным числам ν оператора $\mathcal{A}$, отличным от λ.

Вычислим $r_t := \operatorname{rg}(\mathcal{A} - \lambda\,\mathcal{E})^t = \dim(\mathcal{A} - \lambda\,\mathcal{E})^t \mathcal{X}$:

$$r_t = \sum_{j=1}^{k} \dim(\mathcal{A} - \lambda\,\mathcal{E})^t (\mathcal{A} - \lambda\,\mathcal{E})\{\boldsymbol{e}_j\} + \dim(\mathcal{A} - \lambda\,\mathcal{E})^t \mathcal{X}'.$$

Сейчас нам удобно считать, что векторы $\boldsymbol{e}_j$, $j = 1, \ldots, k$ пронумерованы так, что числа $m_j = m(\boldsymbol{e}_j)$ упорядочены следующим образом: $m_1 \leqslant \cdots \leqslant m_k$.

Если $m_j \leqslant t$, то $(\mathcal{A} - \lambda\,\mathcal{E})^t(\mathcal{A} - \lambda\,\mathcal{E})\{\boldsymbol{e}_j\} = \boldsymbol{0}$. При $m_j > t$ верно равенство

$$(\mathcal{A} - \lambda\,\mathcal{E})^t(\mathcal{A} - \lambda\,\mathcal{E})\{\boldsymbol{e}_j\} = \operatorname{Span}\{(\mathcal{A} - \lambda\,\mathcal{E})^t \boldsymbol{e}_j, \ldots, (\mathcal{A} - \lambda\,\mathcal{E})^{m_j - 1} \boldsymbol{e}_j\},$$

откуда $\dim(\mathcal{A} - \lambda\,\mathcal{E})^t(\mathcal{A} - \lambda\,\mathcal{E})\{\boldsymbol{e}_j\} = m_j - t$.

Так как оператор $\mathcal{A} - \lambda\,\mathcal{E}$ невырожден на $\mathcal{X}'$, то $\dim(\mathcal{A} - \lambda\,\mathcal{E})^t \mathcal{X}' = \dim \mathcal{X}'$. Таким образом,

$$r_t = \sum_{j:\ m_j > t} (m_j - t) + \dim \mathcal{X}'.$$

Из последнего равенства вытекает, что

$$r_t - r_{t+1} = \sum_{j:\ m_j > t} (m_j - t) - \sum_{j:\ m_j > t+1} (m_j - t - 1) = \sum_{j:\ m_j \geqslant t+1} 1 = \sum_{s=1}^{n} N(\lambda, t+s).$$

В последней сумме некоторое число последних слагаемых может быть равно нулю и учтено, что $N(\lambda, p) = 0$ при $p > n$. Таким образом, вычисляя разность $(r_{m-1} - r_m) - (r_m - r_{m+1})$ мы приходим к следующей формуле

$$N(\lambda, m) = r_{m-1} - 2r_m + r_{m+1}, \tag{6.2}$$

где, для определенности положено $r_0 = n$.

Заметим также, что все приведенные рассуждения будут справедливы и в вещественном пространстве $\mathcal{X}$ при дополнительно предположении, что характеристический многочлен оператора $\mathcal{A}$ имеет n (вещественных) корней (с учетом их кратности). Суммируя все вышесказанное можно сформулировать следующее утверждение.

Теорема 6.3. *Пусть $\mathcal{A}$ – это или произвольный линейный оператор, действующий в комплексном n-мерном линейном пространстве или такой линейный оператор, действующий в вещественном n-мерном линейном пространстве, что характеристический многочлен $P_{\mathcal{A}}$ имеет n (вещественных) корней (с учетом их кратности). Тогда для оператора $\mathcal{A}$ существует базис, состоящий из собственных и присоединенных векторов, в котором матрица оператора A имеет жорданову форму. При этом такое представление оператора $\mathcal{A}$ единственно с точностью до порядка следования жордановых клеток.*

Утверждение Теоремы 6.3 оправдывает использования термина *канонический вид* для представления оператора $\mathcal{A}$ в жордановой форме. В самом деле, собственные числа $\lambda_1, \ldots, \lambda_m$ и числа $N(\lambda_j, k)$ зависят только от оператора $\mathcal{A}$, но не от способа построения базиса из собственных и присоединенных векторов. Кроме того, из полученных формул для чисел $N(\lambda_j, k)$ следует, в частности, единственность (с точностью до порядка следования клеток) жордановой формы (канонического вида) оператора $\mathcal{A}$.

Пример отыскания канонического вида линейного оператора. Пусть линейный оператор $\mathcal{A}$, действующий в (четырехмерном) пространстве $\mathbb{R}^4$ задан в некотором базисе этого пространства матрицей

$$A = \begin{pmatrix} 3 & 0 & 1 & 1 \\ 0 & 2 & 0 & -1 \\ 5 & 0 & 7 & 8 \\ -4 & 0 & -4 & -4 \end{pmatrix}$$

Характеристический многочлен этой матрицы имеет вид

$$\det(A - \lambda E) = \lambda^4 - 8\lambda^3 + 24\lambda^2 - 32\lambda + 16.$$

Этот многочлен имеет корень $\lambda = 2$ кратности $k = 4$. Таким образом, все пространство $\mathbb{R}^4$ совпадает с корневым подпространством $\widehat{\mathcal{X}}_{\mathcal{A}}(2)$.

Для того, чтобы найти канонический вид оператора $\mathcal{A}$ вычислим числа $N(2,1)$, $N(2,2)$, $N(2,3)$ и $N(2,4)$ жордановых клеток размера 1×1, 2×2, 3×3 и 4×4, соответствующий собственному значению $\lambda = 2$ в канонической матрице оператора $\mathcal{A}$. Имеем $r_0 = 4$,

$$r_1 = \operatorname{rg}(A - 2E) = \operatorname{rg} \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 5 & 0 & 5 & 8 \\ -4 & 0 & -4 & -6 \end{pmatrix} = 2;$$

$r_2 = \mathrm{rg}(A-2E)^2 = 1$, а $(A-2E)^3 = 0$. Таким образом,

$$N(2,1) = r_0 - 2r_1 + r_2 = 4 - 4 + 1 = 1;$$
$$N(2,2) = r_1 - 2r_2 + r_3 = 2 - 2 + 0 = 0;$$
$$N(2,3) = r_2 - 2r_3 + r_4 = 1 - 0 + 0 = 1;$$

а $N(2,4) = 0$. Из этого вытекает, что канонический вид оператора $\mathcal{A}$ – это одна из следующих матриц

$$\begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \qquad \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Найдем теперь базис из собственных и присоединенных векторов, в котором матрица оператора $\mathcal{A}$ имеет канонический вид.

Для этого, во-первых, найдем собственное подпространство $\mathcal{X}_{\mathcal{A}}(2)$ оператора $\mathcal{A}$, соответствующее собственному значению $\lambda = 2$. Для этого решим следующую систему уравнений, неизвестными в которой являются координаты x_1, x_2 ,x_3 и x_4 собственных векторов оператора $\mathcal{A}$ в исходном базисе:

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 5 & 0 & 5 & 8 \\ -4 & 0 & -4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Решение этой системы имеет вид

$$\big(a, b, -a, 0\big)^\top,$$

где $a, b \in \mathbb{R}$ – произвольные числа. Таким образом, $\dim \mathcal{X}_{\mathcal{A}}(2) = 2$. Так как размерность собственного подпространства меньше, чем размерность коренного, нам необходимо продолжить отыскание присоединенных векторов для собственного значения $\lambda = 2$. Запишем систему уравнений на координаты присоединенного вектора порядка 1:

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 5 & 0 & 5 & 8 \\ -4 & 0 & -4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} a \\ b \\ -a \\ 0 \end{pmatrix}.$$

Нетрудно проверить, что эта система будет совместна при условии $b = 2a$. Выберем собственный вектор $\boldsymbol{x}_1 = (1, 2, -1, 0)^\top$ удовлетворяющий этому условия и решим предыдущую систему при $a = 1$ и $b = 2$. Решение этой системы имеет вид

$$\big(3 - t, s, t, -2\big)^\top$$

где $t, s \in \mathbb{R}$ – произвольные числа.

Для отыскания присоединенного вектора порядка 2 мы рассмотрим систему уравнений

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 5 & 0 & 5 & 8 \\ -4 & 0 & -4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 3 - t \\ s \\ t \\ -2 \end{pmatrix}.$$

Условием совместности этой системы будет $s = 5 - 2t$. Выберем присоединенный вектор порядка 1 так, что $t = 0$ и $s = 5$. Обозначим этот вектор через $\boldsymbol{x}_2$, т.е. $\boldsymbol{x}_2 = (3, 5, 0, -2)^\top$ и решим последнюю систему в этом случае. Получим общий вид присоединенного вектора порядка 2

$$\big(8 - p, q, p, -5\big)^\top.$$

Заметим, что система

$$\begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \\ 5 & 0 & 5 & 8 \\ -4 & 0 & -4 & -6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 8-p \\ q \\ p \\ -5 \end{pmatrix},$$

определяющая координаты присоединенного вектора порядка 3 уже не будет совместной. Поэтому присоединенный вектор $\boldsymbol{e}_3$ порядка 2 мы выбираем произвольно, не требуя соблюдения никаких дополнительных условий. Пусть $\boldsymbol{x}_3 = (8, 0, 0, -5)^\top$.

Нам осталось выбрать вектор $\boldsymbol{x}_4$ так, чтобы он, совместно с уже выбранными векторами $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ и $\boldsymbol{x}_3$, образовывал канонический базис для оператора $\mathcal{A}$. Вспомним, что из двумерного собственного подпространства $\mathcal{X}_{\mathcal{A}}(2)$ мы выбрали пока только один вектор (вектор $\boldsymbol{x}_1$). Выберем $\boldsymbol{x}_4 \in \mathcal{X}_{\mathcal{A}}(2)$ так, чтобы он был линейно независим с $\boldsymbol{x}_1$. Например, $\boldsymbol{x}_4 = (1, 0, -1, 0)^\top$. Составим из координатных столбцов векторов $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ и $\boldsymbol{x}_4$ матрицу

$$S = \begin{pmatrix} 1 & 3 & 8 & 1 \\ 2 & 5 & 0 & 0 \\ -1 & 0 & 0 & -1 \\ 0 & -2 & -5 & 0 \end{pmatrix}$$

и заметим, что

$$S^{-1}AS = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Упражнение. Проверить, что оператор $\mathcal{E} - t\frac{d^2}{dt^2}$, действующий в пространстве $\mathbb{R}[t]_2$ имеет канонический вид

$$\begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

причем соответствующий канонический базис состоит из многочленов $(t, -t^2/2, 1)$.

6.3. Альтернативное доказательство теоремы о каноническом виде линейного оператора

Утверждение о каноническом виде линейного оператора можно сформулировать и в следующем виде:

Теорема 6.4. *Пусть $\mathcal{A}$ – линейный оператор, действующий в n-мерном эрмитовом пространстве $\mathcal{X}$. Существует базис*

$$\boldsymbol{e}_J = \left\{ \boldsymbol{e}_{k,m} : \; k = 1, \ldots, \ell, \; m = 1, \ldots, n_k, \; \sum_{k=1}^{\ell} n_k = n \right\}$$

состоящий из собственных и присоединенных векторов оператора $\mathcal{A}$ такой, что действие оператора $\mathcal{A}$ в базисе $\boldsymbol{e}_J$ описывается следующим образом

$$\begin{aligned} &\mathcal{A}\,\boldsymbol{e}_{k,1} = \lambda_k \boldsymbol{e}_{k,1}, \quad k = 1, \ldots, \ell; \\ &\mathcal{A}\,\boldsymbol{e}_{k,m} = \lambda_k \boldsymbol{e}_{k,m} + \boldsymbol{e}_{k,m-1}, \quad k = 1, \ldots, \ell, \; m = 2, \ldots, n_k. \end{aligned}$$

Из условий сформулированной теоремы вытекает, что векторы $\boldsymbol{e}_{k,1}$ при $k = 1, \ldots, \ell$ – это собственные векторы оператора $\mathcal{A}$, соответствующие собственным значениям λ_k. Соответственно, векторы $\boldsymbol{e}_{k,m}$ при $k = 1, \ldots, \ell$ и $m = 2, \ldots, n_k$, являются присоединенными векторами порядка m, также отвечающими собственным значениям λ_k.

Из доказательства Теоремы 6.4 будет следовать, что жорданова форма матрицы определена единственным образом с точностью до порядка следования жордановых клеток. Этот порядок совпадает с выбранным способом нумерации собственных значений.

Доказательство теоремы 6.4. Доказательство проведем используя индукцию по размерности пространства. Для пространства размерности 1 утверждение теоремы очевидно. Пусть теперь $n > 1$ и утверждение теоремы справедливо для пространств размерности, меньшей n. Установим, что утверждение теоремы справедливо и для операторов в пространствах размерности n.

Пусть λ – собственное число оператора $\mathcal{A}$. Это число является (см. Теорему 5.2) корнем характеристического уравнения $\det(\mathcal{A} - \lambda\,\mathcal{E}) = 0$. Следовательно, для оператора $\mathcal{B} = \mathcal{A} - \lambda\,\mathcal{E}$

$$r := \operatorname{rg} \mathcal{B} < n.$$

Оператор $\mathcal{B}$ отображает пространство $\mathcal{X}$ на некоторое подпространство $\operatorname{Im}\mathcal{B}$. Поэтому, оператор $\mathcal{B}$ отображает подпространство $\operatorname{Im}\mathcal{B}$ в это же подпространство.

По предположению индукции, в $\operatorname{Im}\mathcal{B}$ существует базис

$$\boldsymbol{h} = \left\{ \boldsymbol{h}_{k,m} : \ k = 1, \ldots, p, \ m = 1, \ldots, r_k, \ \sum_{k=1}^{p} r_k = r \right\}$$

такой, что оператор $\mathcal{B}$ действует в $\boldsymbol{h}$ по правилам

$$\left.\begin{array}{l} \mathcal{B}\,\boldsymbol{h}_{k,1} = \mu_k \boldsymbol{h}_{k,1}, \quad k = 1, \ldots, p; \\ \mathcal{B}\,\boldsymbol{h}_{k,m} = \mu_k \boldsymbol{h}_{k,m} + \boldsymbol{h}_{k,m-1}, \quad k = 1, \ldots, p, \ m = 2, \ldots, r_k. \end{array}\right\} \tag{6.3}$$

Итак, оператор $\widetilde{\mathcal{B}} = \mathcal{B}_{|\operatorname{Im}\mathcal{B}}$ имеет в базисе $\boldsymbol{h}$ матрицу $\widetilde{B}$, которая имеет вид

$$\widetilde{B} = \operatorname{Diag}\{J(\mu_1, r_1), \ldots, J(\mu_p, r_p)\}.$$

Предположим, что только первые m_1 собственных значений оператора $\widetilde{\mathcal{B}}$ равны нулю, а остальные – отличны от нуля.

Заметим, что $\operatorname{rg} J(0, t) = t - 1$, а $\operatorname{rg} J(\alpha, t) = t$ при $\alpha \neq 0$. Отсюда следует, что

$$\operatorname{rg} \widetilde{B} = \sum_{k=1}^{p} r_k - m_1 = r - m_1.$$

Следовательно, $\dim \operatorname{Ker} \widetilde{\mathcal{B}} = m_1$. В самом деле, $\operatorname{rg} \widetilde{B} = \dim \operatorname{Im} \mathcal{B}$, а $\dim \operatorname{Ker} \widetilde{\mathcal{B}} = r - \dim \operatorname{Im} \widetilde{\mathcal{B}}$. Более того,

$$\operatorname{Ker} \widetilde{\mathcal{B}} = \operatorname{Span}\{\boldsymbol{h}_{1,1}, \ldots, \boldsymbol{h}_{m_1,1}\}$$

и вектора $\boldsymbol{h}_{1,1}, \ldots, \boldsymbol{h}_{m_1,1}$ в силу своей линейной независимости образуют базис в $\operatorname{Ker}\widetilde{\mathcal{B}}$. Напомним, что $\operatorname{Ker}\widetilde{\mathcal{B}} \subset \operatorname{Ker}\mathcal{B}$. Дополним базис $\boldsymbol{h}_{1,1}, \ldots, \boldsymbol{h}_{m_1,1}$ в $\operatorname{Ker}\widetilde{\mathcal{B}}$ до базиса в $\operatorname{Ker}\mathcal{B}$, добавив в него вектора $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_{m_0}$, где $m_0 = n - r - m_1$. Так как $\boldsymbol{g}_k \in \operatorname{Ker}\mathcal{B}$ при $k = 1, \ldots, m_0$, то $\mathcal{B}\,\boldsymbol{g}_k = 0$.

Так как векторы $\boldsymbol{h}_{k,r_k}$ при $k = 1, \ldots, m_1$ принадлежат $\operatorname{Im}\mathcal{B}$, то существуют такие векторы $\boldsymbol{w}_k \in \mathcal{X}$, что $\mathcal{B}\,\boldsymbol{w}_k = \boldsymbol{h}_{k,r_k}$.

Проверим, что векторы $\boldsymbol{h}_{k,m}$ при $k = 1, \ldots, p$ и $m = 1, \ldots, r_k$, векторы $\boldsymbol{g}_k$ при $k = 1, \ldots, m_0$ и векторы $\boldsymbol{w}_k$ при $k = 1, \ldots, m_1$ линейно независимы. Для этого рассмотрим произвольную линейную комбинацию этих векторов вида

$$\boldsymbol{z} := \sum_{k=1}^{p} \sum_{m=1}^{r_k} \alpha_{km} \boldsymbol{h}_{k,m} + \sum_{k=1}^{m_0} \beta_k \boldsymbol{g}_k + \sum_{k=1}^{m_1} \gamma_k \boldsymbol{w}_k$$

и предположим, что $\boldsymbol{z} = \boldsymbol{0}$. Тогда

$$\mathcal{B}\,\boldsymbol{z} = \sum_{k=1}^{p} \alpha_{k1} \mu_k \boldsymbol{h}_{k,1} + \sum_{k=1}^{p} \sum_{m=2}^{r_k} \alpha_{km} (\mu_k \boldsymbol{h}_{k,m} + \boldsymbol{h}_{k,m-1}) + \sum_{k=1}^{m_1} \gamma_k \boldsymbol{h}_{k,r_k} = \boldsymbol{0}.$$

Итак, выражение для $\mathcal{B}\,\boldsymbol{z}$ представляет собой линейную комбинацию базисных векторов $\boldsymbol{h}_{k,m}$. Следовательно, из равенства $\mathcal{B}\,\boldsymbol{z} = \boldsymbol{0}$ получаем, что коэффициенты при векторах $\boldsymbol{h}_{k,m}$ в этой линейной комбинации равны нулю. Заметим, что число m_1 было выбрано так, что при $k \leqslant m_1$ имеют место равенства $\mu_k = 0$. Следовательно (детали проверки оставляются в качестве *упражнения*), $\gamma_k = 0$ при $k = 1, \ldots, m_1$. Из этого вытекает, что

$$\sum_{k=1}^{m_0} \beta_k \boldsymbol{g}_k = -\sum_{k=1}^{p} \sum_{m=1}^{r_k} \alpha_{km} \boldsymbol{h}_{k,m}.$$

Вектор $\boldsymbol{g} := \sum_{k=1}^{m_0} \beta_k \boldsymbol{g}_k$ принадлежит $\operatorname{Ker} \mathcal{B}$ (так как векторы $\boldsymbol{g}_k$ при $k = 1, \ldots, m_0$ – это часть базиса в подпространстве $\operatorname{Ker} \mathcal{B}$). Из последнего равенства вытекает, что вектор $\boldsymbol{g} \in \operatorname{Im} \widetilde{\mathcal{B}}$ (так как векторы $\boldsymbol{h}_{k,m}$ образуют базис в $\operatorname{Im} \widetilde{\mathcal{B}}$). Отсюда вытекает, что так как $\operatorname{Ker} \widetilde{\mathcal{B}} = \operatorname{Im} \widetilde{\mathcal{B}} \cap \operatorname{Ker} \mathcal{B}$, то $\boldsymbol{g} \in \operatorname{Ker} \widetilde{\mathcal{B}}$. Из этого вытекает, что $\boldsymbol{g} = \sum_{k=1}^{m_1} \gamma'_k \boldsymbol{h}_{k,1}$. Так как пересечение линейных оболочек векторов $\boldsymbol{g}_k$ при $k = 1, \ldots, m_1$ и $\boldsymbol{h}_{k,1}$ при $k = 1, \ldots, m_1$ содержит только нулевой элемент, то вектор $\boldsymbol{g} = \boldsymbol{0}$, откуда последовательно получаем, что $\beta_k = 0$ при $k = 1, \ldots, m_1$ и $\alpha_{km} = 0$ при $k = 1, \ldots, p$ и $m = 1, \ldots, r_k$. Следовательно, указанный набор векторов линейно независим. Заметим, что он состоит из $m_0 + m_1 + \sum_{k=1}^{p} r_k = r + m_0 + m_1$ векторов. Так как $m_0 = n - r - m_1$, то в выбранной системе n векторов и, в силу линейной независимости, они образуют базис в $\mathcal{X}$.

Введем дополнительное обозначение $\boldsymbol{h}_{k,r_k+1} := \boldsymbol{w}_k$. Рассмотрим теперь элементы найденного базиса в $\mathcal{X}$ в следующем порядке

$$\left\{ \begin{array}{l} \boldsymbol{g}_1; \ldots; \boldsymbol{g}_{m_0}; \\ \{\boldsymbol{h}_{k,1}, \ldots, \boldsymbol{h}_{k,r_k}, \boldsymbol{h}_{k,r_k+1}\}, \quad k = 1, \ldots, m_1; \\ \{\boldsymbol{h}_{k,1}, \ldots, \boldsymbol{h}_{k,r_k}\}, \quad k = m_1 + 1, \ldots, p \end{array} \right\}$$

и заметим, что действие оператора $\mathcal{B}$ в этом базисе описывается соотношениями (6.3) и соотношениями $\mathcal{B}\, \boldsymbol{g}_k = \boldsymbol{0}$ при $k = 1, \ldots, m_0$ и $\mathcal{B}\, \boldsymbol{h}_{k,r_k+1} = h_{k,r_k}$ при $k = 1, \ldots, m_1$. следовательно, в построенном базисе оператор $\mathcal{B} = \mathcal{A} - \lambda\, \mathcal{E}$ действует по правилу, требуемому условием теоремы. Отсюда окончательно получаем, что оператор $\mathcal{A} = \mathcal{B} + \lambda\, \mathcal{E}$ действует в этом базисе так, как предписано условиями теоремы. $\square$

РАЗДЕЛ 7

Евклидовы и эрмитовы пространства

7.1. Определение и основные свойства евклидовых пространств

Из курса аналитической геометрии (и, отчасти, из обычного курса геометрии, входящего в школьную программу по математике) известны понятия *скалярного произведения* векторов на плоскости и в трехмерном пространстве и связанные с ним понятия длины вектора и угла между векторами. Оказывается, все эти понятия можно естественным образом обобщить на случай пространства произвольной размерности и природы. Более того, это можно сделать так, что при решении различных задач можно будет использовать наработанную в двумерном и трехмерном случаях геометрическую интуицию в полном объеме.

Определение. *Вещественное линейное пространство $\mathcal{X}$ называется евклидовым пространством, если для любых двух элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ определено вещественное число $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, которое называется скалярным произведением элементов $\boldsymbol{x}$ и $\boldsymbol{y}$ и которое удовлетворяет следующим условиям*

(1) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$ *для любых* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$*;*
(2) $\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y} \rangle = \langle \boldsymbol{x}_1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}_2, \boldsymbol{y} \rangle$ *для любых* $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in \mathcal{X}$*;*
(3) $\langle \alpha\boldsymbol{x}, \boldsymbol{y} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ *для любых* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ *и для любого* $\alpha \in \mathbb{R}$*;*
(4) $\langle \boldsymbol{x}, \boldsymbol{x} \rangle \geqslant 0$ *для любого* $\boldsymbol{x} \in \mathcal{X}$*, а* $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = 0$ *только если* $\boldsymbol{x} = \boldsymbol{0}$.

Замечание. Так как евклидово пространство является линейным пространством (над полем $\mathbb{R}$), в нем определены и имеют тот же смысл все понятия, введенные в предыдущей лекции для общих линейных пространств (линейная зависимость и независимость систем векторов, базис, координаты векторов, размерность пространства и т.п.). Говорят также, что задание скалярного произведения $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ задает в (вещественном) линейном пространстве $\mathcal{X}$ *структуру евклидова пространства.*

Часто, для того, чтобы подчеркнуть, что скалярное произведение удовлетворяет условиям (1)-(4), используют термин *евклидово скалярное произведение.*

Замечание. Обозначение $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ (скалярное произведение векторов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$) следует четко отличать от обозначения $\langle \mathfrak{f} | \boldsymbol{x} \rangle$ (от действия линейного функционала $\mathfrak{f} \in \mathcal{X}^*$ на вектор $\boldsymbol{x} \in \mathcal{X}$). Символ "вертикальной черты" применяется во втором обозначении для того, чтобы подчеркнуть, что первым аргументом в выражении $\langle \cdot | \cdot \rangle$ должен быть линейный функционал (т.е. элемент множества $\mathcal{X}^*$), а вторым аргументом – элемент $\boldsymbol{x} \in \mathcal{X}$.

Пример 7.1. В качестве первого примера евклидовых пространств рассмотрим линейное пространство $\mathbb{R}^n$, $n \in \mathbb{N}$, которое было введено в предыдущей лекции. Определим скалярное произведение $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ двух элементов $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ и $\boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$ следующим образом

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n.$$

Проверка выполнения свойств (1)–(4) скалярного произведения для определенной таким образом величины $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ оставляется в качестве легкого *упражнения*.

Пространство $\mathbb{R}^n$ с введенным указанным способом скалярным произведением также обозначается символом $\mathbb{R}^n$.

Заметим, что структура евклидова пространства в заданном линейном пространстве $\mathcal{X}$ может быть, в общем случае, введена многими различными способами. В самом деле, имеет место следующая конструкция.

Пример 7.2. Рассмотрим снова линейное пространство $\mathbb{R}^n$ и определим в нем скалярное произведение, отличное от скалярного произведения $\langle\cdot,\cdot\rangle$ из Примера 7.1. Пусть $A = (a_{jk})$ – некоторая положительно определенная симметричная $n \times n$-матрица, т.е. матрица A такова, что $A = A^\top$ и

$$\sum_{j,k=1}^{n} a_{jk}\xi_j\xi_k > 0$$

при всех вещественных $\xi_1, \dots, \xi_n$, одновременно не равных нулю.

В качестве ***упражнения*** предлагается проверить, что выражение

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle_A := \sum_{j,k=1}^{n} a_{jk}x_j y_k,$$

где $\boldsymbol{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ и $\boldsymbol{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$ удовлетворяет всем свойствам скалярного произведения.

Еще один интересный и важный **пример** (бесконечномерного) евклидова пространства – это линейное пространство $C([a,b])$, состоящее из всех непрерывных на отрезке $[a,b]$ вещественнозначных функций одной вещественной переменной, в котором скалярное произведение определено по формуле

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt, \qquad f, g \in C([a,b]).$$

Основные свойства евклидовых пространств. Пусть $\mathcal{X}$ – евклидово пространство со скалярным произведением $\langle \boldsymbol{x}, \boldsymbol{y}\rangle$ (конечномерное или бесконечномерное).

Предложение 7.3 (неравенство Коши-Буняковского)**.** *Для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ верно неравенство*

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle^2 \leqslant \langle \boldsymbol{x}, \boldsymbol{x}\rangle\langle \boldsymbol{y}, \boldsymbol{y}\rangle.$$

Доказательство. Из свойства (4) скалярного произведения вытекает, что для любого числа $\alpha \in \mathbb{R}$ и для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ имеет место неравенство

$$\langle \alpha\boldsymbol{x} - \boldsymbol{y}, \alpha\boldsymbol{x} - \boldsymbol{y}\rangle \geqslant 0.$$

Из свойств (1)–(3) скалярного произведения из этого неравенства вытекает, что

$$\alpha^2\langle \boldsymbol{x}, \boldsymbol{x}\rangle - 2\alpha\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{y}, \boldsymbol{y}\rangle \geqslant 0.$$

Так как необходимым и достаточным условием неотрицательности многочлена второй степени с положительным старшим коэффициентом является неположительность его дискриминанта, то

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle^2 - \langle \boldsymbol{x}, \boldsymbol{x}\rangle\langle \boldsymbol{y}, \boldsymbol{y}\rangle \leqslant 0. \quad \square$$

Напомним, что для векторов в двумерном и трехмерном пространстве естественным образом вводится понятие длины вектора. Введем аналогичное понятие в случае произвольного вещественного линейного пространства.

Определение. *Линейное пространство $\mathcal{X}$ называется нормированным, если для любого элемента $\boldsymbol{x} \in \mathcal{X}$ определено вещественное число $\|\boldsymbol{x}\|$, которое называется нормой элемента $\boldsymbol{x}$ и удовлетворяет следующим условиям*

(1) $\|\boldsymbol{x}\| > 0$ *для любого ненулевого элемента* $\boldsymbol{x} \in \mathcal{X}$, *а* $\|\boldsymbol{x}\| = 0$ *только если* $\boldsymbol{x} = \boldsymbol{0}$;
(2) $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$ *для любых* $\boldsymbol{x} \in \mathcal{X}$ *и* $\alpha \in \mathbb{R}$;
(3) *для любых* $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ *справедливо неравенство* $\|\boldsymbol{x} + \boldsymbol{y}\| \leqslant \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$, *которое называется неравенством треугольника или неравенством Минковского.*

Оказывается, что в произвольном евклидовом пространстве можно определить *норму* (или *длину*) вектора таким образом, что эта величина выражается в терминах скалярного произведения естественным образом.

Предложение 7.4. *Пусть $\mathcal{X}$ – евклидово пространство со скалярным произведением $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$. Тогда величина*

$$\|\boldsymbol{x}\| := \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$$

является нормой в $\mathcal{X}$. Таким образом, в евклидовом пространстве $\mathcal{X}$ естественно вводится структура нормированного пространства, а норма $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$, определенная при помощи скалярного произведения в $\mathcal{X}$, называется евклидовой нормой.

Доказательство. Пусть величина $\|\boldsymbol{x}\|$ определена соотношением $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ при $\boldsymbol{x} \in \mathcal{X}$. Из свойства (4) скалярного произведения вытекает, что эта величина корректно определена для всех $\boldsymbol{x} \in \mathcal{X}$. Проверим, что она удовлетворяет свойствам нормы. Свойство (1) нормы очевидно: $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} \geqslant 0$ для любого $\boldsymbol{x} \in \mathcal{X}$. Далее, пусть $\boldsymbol{x} \in \mathcal{X}$ – произвольный элемент рассматриваемого евклидова пространства, а $\alpha \in \mathbb{R}$ – произвольное число. Тогда

$$\|\alpha\boldsymbol{x}\| = \sqrt{\langle \alpha\boldsymbol{x}, \alpha\boldsymbol{x} \rangle} = \sqrt{\alpha^2 \langle \boldsymbol{x}, \boldsymbol{x} \rangle} = |\alpha| \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = |\alpha| \|\boldsymbol{x}\|$$

в силу свойств (1) и (3) скалярного произведения. Остается проверить, что величина $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ удовлетворяет неравенству треугольника. Из неравенства Коши-Буняковского вытекает, что

$$\begin{aligned}\|\boldsymbol{x}+\boldsymbol{y}\| = \sqrt{\langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y} \rangle} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle} \leqslant \\ \leqslant \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle + 2\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}\sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle} + \langle \boldsymbol{y}, \boldsymbol{y} \rangle} = \\ = \sqrt{\left(\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} + \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle}\right)^2} = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} + \sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle} = \|\boldsymbol{x}\| + \|\boldsymbol{y}\|\end{aligned}$$

для произвольных векторов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. Итак, величина $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$ удовлетворяет всем условиям нормы. □

Далее, в произвольном евклидовом пространстве $\mathcal{X}$ можно ввести понятия *угла* между векторами.

Определение. *Назовем углом между векторам $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ такой угол $\theta \in [0, \pi]$, для которого*

$$\cos\theta = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}\sqrt{\langle \boldsymbol{y}, \boldsymbol{y} \rangle}}.$$

Это определение корректно, так как в силу неравенства Коши-Буняковского, выражение, стоящее в правой части приведенной цепочки равенств, по модулю не превосходит единицы.

Следующее понятие также обобщает на случай произвольного евклидова пространства хорошо известное из школьной программы и из курса аналитической геометрии понятие ортогональных векторов.

Определение. *Элементы $\boldsymbol{x}$ и $\boldsymbol{y}$ евклидова пространства $\mathcal{X}$ называются ортогональными, если $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$.*

Заметим, что если элементы $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ ортогональны, т.е, если $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, то

$$\|\boldsymbol{x}+\boldsymbol{y}\|^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle + 2\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{x} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle = \|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2.$$

Это соотношение можно трактовать как естественный аналог в случае произвольного вещественного евклидова пространства классической теоремы Пифагора о том, что в прямоугольном треугольнике квадрат гипотенузы равен сумме квадратов катетов. Как нетрудно заметить, этот результат тривиально обобщается а случай n попарно ортогональных элементов в $\mathcal{X}$: если элементы $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathcal{X}$, $n \in \mathbb{N}$, таковы, что $\langle \boldsymbol{x}_j, \boldsymbol{x}_k \rangle = 0$ при $j \neq k$, $j, k \in \{1, \ldots, n\}$, то

$$\|\boldsymbol{x}_1 + \boldsymbol{x}_2 + \cdots + \boldsymbol{x}_n\|^2 = \|\boldsymbol{x}_1\|^2 + \|\boldsymbol{x}_2\|^2 + \cdots + \|\boldsymbol{x}_n\|^2.$$

В рассмотренном выше примере евклидова пространства $\mathbb{R}^n$, $n \in \mathbb{N}$, норма любого элемента $\boldsymbol{x} = (x_1, \ldots, x_n)$ определяется равенством $\|\boldsymbol{x}\| = \sqrt{x_1^2 + \cdots + x_n^2}$, а в случае евклидова пространства $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_A)$, где A – симметричная неотрицательно определенная матрица, норма элемента $\boldsymbol{x} = (x_1, \ldots, x_n)$ равна

$$\|\boldsymbol{x}\| = \sqrt{\sum_{j,k=1}^{n} a_{jk} x_j x_k}.$$

В качестве ***упражнения*** предлагается выписать неравенство Коши-Буняковского и неравенство треугольника в этих примерах, а также определение нормы элемента и соответствующие неравенства в пространстве $C([a, b])$.

7.2. Ортонормированный базис конечномерного евклидова пространства

Определим *символ Кронекера* $\delta_{j,k}$, где $j, k \in \mathbb{N}$ соотношением

$$\delta_{j,k} = \begin{cases} 1 & \text{при} \quad j = k; \\ 0 & \text{при} \quad j \neq k \end{cases}$$

Всюду в дальнейшем в этой лекции $n \in \mathbb{N}$ – некоторое фиксированное натуральное число.

Определение. *Пусть $\mathcal{X}$ – евклидово пространство, $\dim \mathcal{X} = n$ и пусть $m \leqslant n$ – натуральное число. Скажем, что элементы $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ евклидова пространства $\mathcal{X}$, образуют ортонормированную систему ранга m в $\mathcal{X}$, если*

$$\langle \boldsymbol{e}_j, \boldsymbol{e}_k \rangle = \delta_{j,k} \qquad \text{для любых} \quad j, k = 1, \ldots, m. \tag{7.1}$$

Предложение 7.5. *Любая ортонормированная система $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ ранга n в n-мерном евклидовом пространстве $\mathcal{X}$ образует базис в $\mathcal{X}$.*

Доказательство. Так как ортонормированная система ранга n состоит из n векторов, то, учитывая Предложение 2.7, для доказательства нам необходимо и достаточно проверить линейную независимость векторов ортонормированной системы $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. Рассмотрим произвольную линейную комбинацию

$$\boldsymbol{w} = \alpha_1 \boldsymbol{e}_1 + \alpha_2 \boldsymbol{e}_2 + \cdots + \alpha_n \boldsymbol{e}_n$$

элементов ортонормированной системы. Из свойств скалярного произведения и соотношений ортогональности (7.1), при $m = n$, получаем, что $\langle \boldsymbol{w}, \boldsymbol{e}_k \rangle = \alpha_k$ при $k = 1, \ldots, n$.

Если мы теперь предположим, что $\boldsymbol{w} = \boldsymbol{0}$, то из равенств $\alpha_k = \langle \boldsymbol{w}, \boldsymbol{e}_k \rangle$, $k = 1, \ldots, n$ будет следовать, что

$$\alpha_k = \langle \boldsymbol{0}, \boldsymbol{e}_k \rangle = 0$$

для любого $k = 1, \ldots, n$. Итак, только тривиальная линейная комбинация векторов ортонормированной системы ранга n равна нулю и, следовательно, такая система будет линейно независима в $\mathcal{X}$. А из этого, напомним, уже вытекает, что векторы ортонормированной системы ранга n образуют базис в n-мерного пространстве $\mathcal{X}$. $\square$

Определение. *Ортонормированным базисом в n-мерном евклидовом пространстве $\mathcal{X}$ называют любую ортонормированную систему $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ ранга n, или, что эквивалентно, любой базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$, элементы которого удовлетворяют соотношениям ортогональности (7.1) при $m = n$.*

Отметим, что понятие ортонормированного базиса – это естественное обобщение на случай произвольного вещественного евклидова пространства понятия декартова прямоугольного базиса в аналитической геометрии.

Оказывается, в произвольном конечномерном евклидовом пространстве можно найти ортонормированный базис. Имеет место следующее утверждение.

Теорема 7.6. *Во всяком n-мерном евклидовом пространстве $\mathcal{X}$ существует ортонормированный базис.*

Доказательство. Доказательство этой теоремы основано на специальном алгоритме построения ортонормированного базиса на основе произвольно выбранного базиса линейного пространства $\mathcal{X}$.

Пусть $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$ – некоторый базис в $\mathcal{X}$. Отталкиваясь от него мы построим ортонормированный базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в $\mathcal{X}$.

Вектор $\boldsymbol{e}_1$ мы определим соотношением $\boldsymbol{e}_1 := \boldsymbol{z}_1/\|\boldsymbol{z}_1\|$, так что $\|\boldsymbol{e}_1\| = 1$. Двигаясь далее мы определим вектор

$$\boldsymbol{w}_2 := \boldsymbol{z}_2 - \langle \boldsymbol{z}_2, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 = \boldsymbol{z}_2 - \frac{\langle \boldsymbol{z}_2, \boldsymbol{z}_1 \rangle}{\|\boldsymbol{z}_1\|^2} \boldsymbol{z}_1.$$

Из того, что векторы $\boldsymbol{z}_1$ и $\boldsymbol{z}_2$ являются элементами базиса в $\mathcal{X}$ вытекает, что $\boldsymbol{w}_2 \neq \boldsymbol{0}$ и, следовательно, $\|\boldsymbol{w}_2\| \neq 0$. Определим теперь вектор $\boldsymbol{e}_2 := \boldsymbol{w}_2/\|\boldsymbol{w}_2\|$, так что $\|\boldsymbol{e}_2\| = 1$. Легко проверяется, что $\langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle = 0$. В самом деле

$$\langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle = \frac{1}{\|\boldsymbol{w}_2\|} \langle \boldsymbol{e}_1, \boldsymbol{w}_2 \rangle = \frac{1}{\|\boldsymbol{w}_2\|} \big(\langle \boldsymbol{e}_1, \boldsymbol{z}_2 \rangle - \langle \boldsymbol{z}_2, \boldsymbol{e}_1 \rangle \big) = 0.$$

Итак, векторы $\boldsymbol{e}_1$ и $\boldsymbol{e}_2$ удовлетворяют соотношениям ортогональности (7.1) при $m = 2$. Сделаем еще один шаг рассматриваемого алгоритма. Для этого определим вектор

$$\boldsymbol{w}_3 := \boldsymbol{z}_3 - \langle \boldsymbol{z}_3, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2 - \langle \boldsymbol{z}_3, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1.$$

Как и в случае вектора $\boldsymbol{w}_2$ проверяется, что вектор $\boldsymbol{w}_3$ является нетривиальной линейной комбинацией базисных векторов $\boldsymbol{z}_1$, $\boldsymbol{z}_2$ и $\boldsymbol{z}_3$ и, следовательно, $\boldsymbol{w}_3 \neq \boldsymbol{0}$ и $\|\boldsymbol{w}_3\| \neq 0$. Далее,

$$\langle \boldsymbol{w}_3, \boldsymbol{e}_j \rangle = \langle \boldsymbol{z}_3, \boldsymbol{e}_j \rangle - \langle \boldsymbol{z}_3, \boldsymbol{e}_j \rangle = 0, \quad j = 1, 2.$$

Таким образом, при $\boldsymbol{e}_3 := \boldsymbol{w}_3/\|\boldsymbol{w}_3\|$ система векторов $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ удовлетворяет соотношениям ортогональности (7.1) при $m = 3$.

Заметим, что вектор $\boldsymbol{e}_1$ выражается только через вектор $\boldsymbol{z}_1$, вектор $\boldsymbol{e}_2$ только через $\boldsymbol{z}_1$ и $\boldsymbol{z}_2$, а вектор $\boldsymbol{e}_3$ только через $\boldsymbol{z}_1$, $\boldsymbol{z}_2$ и $\boldsymbol{z}_3$.

Предположим теперь, что мы сделали m шагов алгоритма при некотором $m < n$ и построили ортонормированную систему $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ ранга m, причем каждый вектор $\boldsymbol{e}_k$ из этой системы представим в виде линейной комбинации векторов $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_k$. Рассмотрим теперь вектор

$$\boldsymbol{w}_{m+1} = \boldsymbol{z}_{m+1} - \sum_{\ell=1}^{m} \langle \boldsymbol{z}_{m+1}, \boldsymbol{e}_\ell \rangle \, \boldsymbol{e}_\ell.$$

Так как $\boldsymbol{w}_{m+1}$ выражается в виде нетривиальной линейной комбинации базисных векторов $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{m+1}$, то $\boldsymbol{w}_{m+1} \neq \boldsymbol{0}$ и $\|\boldsymbol{w}_{m+1}\| \neq 0$. Так как система векторов $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ является ортонормированной, то из соотношений ортогональности (7.1) вытекает, что

$$\langle \boldsymbol{w}_{m+1}, \boldsymbol{e}_k \rangle = \langle \boldsymbol{z}_{m+1}, \boldsymbol{e}_k \rangle - \langle \boldsymbol{z}_{m+1}, \boldsymbol{e}_k \rangle = 0$$

при $k = 1, \ldots, m$.

Если определить вектор $\boldsymbol{e}_{m+1} := \boldsymbol{w}_{m+1}/\|\boldsymbol{w}_{m+1}\|$, то система $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}\}$ будет ортонормированной системой ранга $m+1$, причем все ее вектора являются линейными комбинациями векторов $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_{m+1}$.

Выполнив n шагов описанного алгоритма мы получим в $\mathcal{X}$ ортонормированную систему $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ ранга n. А она, согласно определению, и является ортонормированным базисом в $\mathcal{X}$. $\square$

Определение. *Алгоритм построения ортонормированного базиса, описанный явно при доказательстве теоремы 7.6 называется процессом ортогонализации системы $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n\}$ линейно независимых векторов. Этот процесс еще называют процессом ортогонализации Грама-Шмидта.*

Пример 7.7. Система векторов

$$\boldsymbol{e}_1 = (1, 0, 0, \ldots, 0, 0, 0), \quad \boldsymbol{e}_2 = (0, 1, 0, \ldots, 0, 0, 0), \quad \ldots, \quad \boldsymbol{e}_n = (0, 0, 0, \ldots, 0, 0, 1)$$

является ортонормированным базисом в пространстве $\mathbb{R}^n$, $n \in \mathbb{N}$.

Свойства ортонормированных базисов. Рассмотрим некоторый произвольный (не обязательно ортонормированный) базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в n-мерном евклидовом пространстве $\mathcal{X}$. Для элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{Y}$ имеет место соотношение

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \left\langle \sum_{j=1}^{n} x_j \boldsymbol{e}_j, \sum_{k=1}^{n} y_k \boldsymbol{e}_k \right\rangle = \sum_{j,k=1}^{n} x_j y_k \langle \boldsymbol{e}_j, \boldsymbol{e}_k \rangle = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top} G(\boldsymbol{e}) [\boldsymbol{y}]_{\boldsymbol{e}}, \tag{7.2}$$

где $[\boldsymbol{x}]_{\boldsymbol{e}} = (x_1, \ldots, x_n)^{\top}$ и $[\boldsymbol{y}]_{\boldsymbol{e}} = (y_1, \ldots, y_n)^{\top}$ – координатные столбцы векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в базисе $\boldsymbol{e}$, а матрица $G(\boldsymbol{e}) = G(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ определена следующим образом

$$G(\boldsymbol{e}) = G(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n) = \begin{pmatrix} \langle \boldsymbol{e}_1, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_1, \boldsymbol{e}_2 \rangle & \cdots & \langle \boldsymbol{e}_1, \boldsymbol{e}_n \rangle \\ \langle \boldsymbol{e}_2, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_2, \boldsymbol{e}_2 \rangle & \cdots & \langle \boldsymbol{e}_2, \boldsymbol{e}_n \rangle \\ \cdots & \cdots & \cdots & \cdots \\ \langle \boldsymbol{e}_n, \boldsymbol{e}_1 \rangle & \langle \boldsymbol{e}_n, \boldsymbol{e}_2 \rangle & \cdots & \langle \boldsymbol{e}_n, \boldsymbol{e}_n \rangle \end{pmatrix}.$$

Определение. *Матрица $G(\boldsymbol{z}) = G(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$ называется матрицей Грама системы векторов $\boldsymbol{z} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$.*

Имеет место следующее важное свойство матрицы Грама.

Теорема 7.8. *Система векторов $\boldsymbol{z} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$ в n-мерном евклидовом пространстве линейно независима тогда и только тогда, когда матрица Грама $G(\boldsymbol{z})$ невырождена.*

Доказательство. Предположим, что векторы $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$ линейно зависимы. Тогда один из них, скажем, $\boldsymbol{z}_m$, является невырожденной линейной комбинацией остальных: $\boldsymbol{z}_m = \alpha_1 \boldsymbol{z}_1 + \cdots + \alpha_{m-1} \boldsymbol{z}_{m-1}$. Но из определения матрицы Грама и свойств скалярного произведения немедленно вытекает, что в этом случае последняя строка матрицы Грама есть линейная комбинация ее первых $m-1$ строк и, следовательно, матрица Грама линейно зависимой системы будет вырождена.

Пусть теперь матрица Грама $G(\boldsymbol{z})$ системы векторов $\boldsymbol{z} = \{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m\}$ вырождена. Тогда найдутся такие числа $\beta_1, \ldots, \beta_m$, не все из которых равны нулю, такие, что

$$\beta_1 \langle \boldsymbol{z}_j, \boldsymbol{z}_1 \rangle + \beta_2 \langle \boldsymbol{z}_j, \boldsymbol{z}_2 \rangle + \cdots + \beta_m \langle \boldsymbol{z}_j, \boldsymbol{z}_m \rangle = 0, \quad j = 1, \ldots, m.$$

Из этого следует, что $\langle \boldsymbol{z}_j, \beta_1 \boldsymbol{z}_1 + \cdots + \beta_m \boldsymbol{z}_m \rangle = 0$ для всех $j = 1, \ldots, m$ и, следовательно, $\beta_1 \boldsymbol{z}_1 + \cdots + \beta_m \boldsymbol{z}_m \perp \operatorname{Span}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$. Но так как вектор $\beta_1 \boldsymbol{z}_1 + \cdots + \beta_m \boldsymbol{z}_m$ принадлежит подпространству $\operatorname{Span}(\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m)$, то $\beta_1 \boldsymbol{z}_1 + \cdots + \beta_m \boldsymbol{z}_m = \boldsymbol{0}$. Последнее соотношение в точности означает линейную зависимость векторов $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_m$. $\square$

Следствие. *Набор из n векторов $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в n-мерном евклидовом пространстве $\mathcal{X}$ образует базис если и только если матрица Грама $G(\boldsymbol{e})$ невырождена.*

Установим следующее характеристическое свойство ортонормированных базисов в евклидовом пространстве.

Теорема 7.9. *Пусть $\mathcal{X}$ – n-мерное евклидово пространство со скалярным произведением $\langle \cdot, \cdot \rangle$. Базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ в $\mathcal{X}$ является ортонормированным если и только если $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{n} x_j y_j$ для любых векторов $\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{e}_j \in \mathcal{X}$ и $\boldsymbol{y} = \sum_{k=1}^{n} y_k \boldsymbol{e}_k \in \mathcal{X}$.*

Доказательство. Пусть базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ в $\mathcal{X}$ является ортонормированным. Тогда, в силу соотношений ортогональности (7.1) при $m = n$ получаем, что $G(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) = E$, где E – единичная $n \times n$-матрица, а формула (7.2) примет, соответственно вид

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top} G(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)[\boldsymbol{y}]_{\boldsymbol{e}} = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top} E[\boldsymbol{y}]_{\boldsymbol{e}} = \sum_{j=1}^{n} x_j y_j.$$

Докажем обратную импликацию. Пусть базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ в $\mathcal{X}$ таков, что для любых двух векторов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{Y}$ имеет место равенство

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{j=1}^{n} x_j y_j.$$

Пусть $\ell \in \{1, \dots, n\}$. Так как в координатном столбце $[\boldsymbol{e}_\ell]_{\boldsymbol{e}} = (e_{\ell,1}, \dots, e_{\ell,n})^{\top}$ координата $e_{\ell,\ell} = 1$, а $e_{\ell,j} = 0$ при $j \in \{1, \dots, n\} \setminus \{\ell\}$, то при $k, r \in \{1, \dots, n\}$

$$\langle \boldsymbol{e}_k, \boldsymbol{e}_r \rangle = \sum_{j=1}^{n} e_{k,j} e_{r,j} = \begin{cases} 1, & \text{если} \quad k = r; \\ 0, & \text{если} \quad k \neq r. \end{cases}$$

Итак, для системы векторов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ выполнены соотношения ортогональности (7.1) при $m = n$ и, следовательно, базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ является ортонормированным. □

Пусть теперь

$$\boldsymbol{x} = x_1 \boldsymbol{e}_1 + \dots + x_n \boldsymbol{e}_n$$

– разложение некоторого элемента $\boldsymbol{x} \in \mathcal{X}$ по ортонормированному базису $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ в $\mathcal{X}$. Выясним, каким образом можно вычислять координаты $x_1, \dots, x_n$ элемента $\boldsymbol{x}$ относительно этого базиса. Из свойств ортонормированного базиса и из свойств скалярного произведения вытекает (проверка необходимых деталей оставляется в качестве *упражнения*), что

$$\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle = \left\langle \sum_{j=1}^{n} x_j \boldsymbol{e}_j, \boldsymbol{e}_k \right\rangle = \sum_{j=1}^{n} x_j \langle \boldsymbol{e}_j, \boldsymbol{e}_k \rangle = x_k,$$

при любом натуральном $k \leqslant n$. Другими словами, справедливо следующее утверждение.

Предложение. *Координаты произвольного вектора евклидова пространства относительно заданного ортонормированного базиса равны скалярным произведениям этого вектора на соответствующие элементы базиса.*

Определение. *Пусть $\boldsymbol{x} \in \mathcal{X}$ – произвольный элемент рассматриваемого евклидова пространства, а $\boldsymbol{c} \in \mathcal{X}$ такой элемент, что $\|\boldsymbol{c}\| = 1$. По аналогии с обычной геометрией, величина $\langle \boldsymbol{x}, \boldsymbol{c} \rangle$ называется проекцией элемента $\boldsymbol{x}$ на вектор $\boldsymbol{c}$.*

С использованием этой терминологии, координаты произвольного элемента евклидова пространства относительно некоторого ортонормированного базиса равны проекциям этого элемента на соответствующие векторы базиса.

Разложение конечномерного евклидова пространства в прямую сумму подпространства и его ортогонального дополнения. Пусть $n \in \mathbb{N}$ и пусть $\mathcal{X}$ – произвольное n-мерное евклидово пространство. Пусть $\mathcal{Y}$ – некоторое подпространство линейного пространства $\mathcal{X}$.

Определение. *Совокупность $\mathcal{Y}^{\perp} := \{\boldsymbol{x} \in \mathcal{X} : \ \langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ для любого $\boldsymbol{y} \in \mathcal{Y}\}$ называется ортогональным дополнением подпространства $\mathcal{Y}$ в $\mathcal{X}$.*

Упражнение. Проверить, что $\mathcal{Y}^{\perp}$ является подпространством в $\mathcal{X}$.

Предложение 7.10. *Пусть $\mathcal{X}$ – некоторое n-мерное евклидово пространство, $n \in \mathbb{N}$, а $\mathcal{Y}$ – подпространство пространства $\mathcal{X}$. Тогда $\mathcal{X} = \mathcal{Y} \oplus \mathcal{Y}^{\perp}$.*

Доказательство. Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m\}$ – произвольный ортонормированный базис в $\mathcal{Y}$. Дополним этот базис векторам $\boldsymbol{z}_{m+1}, \ldots, \boldsymbol{z}_n$ до базиса во всем пространстве $\mathcal{X}$. Применим к семейству линейно независимых векторов $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{z}_{m+1}, \ldots, \boldsymbol{z}_n\}$ процесс ортогонализации, описанный в предыдущем разделе. Из соотношений ортогональности непосредственно вытекает, что первые m векторов в результате этого процесса не изменятся. В результате ортогонализации получаем ортонормированный базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n\}$ во всем пространстве $\mathcal{X}$. Представим любой элемент $\boldsymbol{x} \in \mathcal{X}$ в виде

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_m\boldsymbol{e}_m + x_{m+1}\boldsymbol{e}_{m+1} + \cdots + x_n\boldsymbol{e}_n$$

(напомним, что такое представление единственно) и определим векторы

$$\boldsymbol{x}' = x_1\boldsymbol{e}_1 + \cdots + x_m\boldsymbol{e}_m \in \mathcal{Y} \qquad \text{и} \qquad \boldsymbol{x}'' = x_{m+1}\boldsymbol{e}_{m+1} + \cdots + x_n\boldsymbol{e}_n.$$

Так как $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_m, \boldsymbol{e}_{m+1}, \ldots, \boldsymbol{e}_n\}$ – это ортонормированный базис в $\mathcal{X}$, то $\langle \boldsymbol{x}'', \boldsymbol{e}_k \rangle = 0$ при $k = 1, \ldots, m$ и, следовательно, $\langle \boldsymbol{x}'', \boldsymbol{w} \rangle = 0$ для любого вектора $\boldsymbol{w} \in \mathcal{Y}$, т.е. $\boldsymbol{x}'' \in \mathcal{Y}^{\perp}$. Таким образом установлено, что произвольный элемент $\boldsymbol{x} \in \mathcal{X}$ единственным образом представляется в виде $\boldsymbol{x} = \boldsymbol{x}' + \boldsymbol{x}''$, $\boldsymbol{x}' \in \mathcal{Y}$, $\boldsymbol{x}'' \in \mathcal{Y}^{\perp}$. Следовательно, $\mathcal{X} = \mathcal{Y} \oplus \mathcal{Y}^{\perp}$. $\square$

Изоморфизм конечномерных евклидовых пространств. Пусть евклидовы пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ изоморфны как линейные пространства. При этом изоморфизм $\varphi : \mathcal{X}_1 \to \mathcal{X}_2$ линейных пространств не "видит" евклидовой структуры в $\mathcal{X}_1$ и $\mathcal{X}_2$ и может не сохранять скалярное произведение. Следовательно, при изучении евклидовых пространств естественно выделить такие изоморфизмы φ линейных пространств $\mathcal{X}_1$ и $\mathcal{X}_2$, которые сохраняют скалярное произведение. Это замечание оправдывает введение следующего понятия.

Определение. *Два произвольных евклидовых пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ называются изоморфными, если они изоморфны как линейные пространства, причем изоморфизм линейных пространств $\varphi : \mathcal{X}_1 \to \mathcal{X}_2$ дополнительно обладает тем свойством, что $\langle \varphi(\boldsymbol{x}), \varphi(\boldsymbol{y}) \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_1$. Факт изоморфизма евклидовых пространств $\mathcal{X}_1$ и $\mathcal{X}_2$ будем записывать символом $\mathcal{X}_1 \cong \mathcal{X}_2$.*

Из определения изоморфизма линейных пространств вытекает, что евклидовы пространства $\mathcal{X}_1$ и $\mathcal{X}_2$ изоморфны, если существует такое биективное отображение $\varphi : \mathcal{X}_1 \to \mathcal{X}_2$, что $\varphi(\boldsymbol{x} + \boldsymbol{y}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$ для любых элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_1$, $\varphi(\alpha\boldsymbol{x}) = \alpha\varphi(\boldsymbol{x})$ для любого $\boldsymbol{x} \in \mathcal{X}_1$ и $\alpha \in \mathbb{R}$ и $\langle \varphi(\boldsymbol{x}), \varphi(\boldsymbol{y}) \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle$ для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}_1$.

Предложение 7.11. *Любые два конечномерных евклидовых пространства, имеющие одинаковую размерность n, изоморфны.*

Доказательство. При доказательстве Теоремы 2.10 мы непосредственно устанавливали изоморфизм между двумя конечномерными линейными пространствами размерности n. Аналогичное рассуждение можно было применить и при доказательстве этого утверждения. Однако мы будем рассуждать несколько иначе. В самом деле, нам достаточно проверить, что любое n-мерное евклидово пространство $\mathcal{X}$ изоморфно пространству $\mathbb{R}^n$ (рассматриваемому со стандартным скалярным произведением). В самом деле, пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – некоторый ортонормированный базис в пространстве $\mathcal{X}$, т.е. для любого элемента $\boldsymbol{x} \in \mathcal{X}$ существует единственное представление $\boldsymbol{x} = x_1\boldsymbol{e}_1 + \cdots + x_n\boldsymbol{e}_n$. Определим отображение $\varphi : \mathcal{X} \to \mathbb{R}^n$ по правилу

$$\varphi(\boldsymbol{x}) = (x_1, \ldots, x_n) \in \mathbb{R}^n.$$

Из единственности разложения элемента по базису следует биективность этого отображения. Из Предложения 2.5 вытекает, что отображение φ – это изоморфизм линейных пространств $\mathcal{X}$ и $\mathbb{R}^n$. Далее, из Теоремы 7.9 вытекает, что скалярное произведение элементов $\boldsymbol{x} = \sum_{j=1}^n x_j \boldsymbol{e}_j \in \mathcal{X}$ и $\boldsymbol{y} = \sum_{j=1}^n y_j \boldsymbol{e}_j \in \mathcal{X}$ равно $\sum_{j=1}^n x_j y_j$ (так как базис $\boldsymbol{e}$ ортонормированный). Но это в точности равно скалярному произведению строк $(x_1, \ldots, x_n)$ и $(y_1, \ldots, y_n)$ в $\mathbb{R}^n$. Таким образом, построенное отображение φ – изоморфизм евклидовых пространств $\mathcal{X}$ и $\mathbb{R}^n$. □

Это утверждение позволяет использовать геометрическую интуицию, развитую при решении геометрических задач на плоскости и в трехмерном пространстве, при изучении различных свойств объектов в конечномерных евклидовых пространствах. Разумеется, "переносимыми" являются только те результаты, которые формулируются в терминах линейной и евклидовой структур, т.е. в терминах линейных операций, скалярных произведений, норм векторов и углов между векторами и т.п.

7.3. Ортонормированные базисы и ортогональные матрицы

Выясним, какими свойствами обладает матрица перехода от одного ортонормированного базиса в евклидовом пространстве к другому. Пусть $\mathcal{X}$ – евклидово пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$, и пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ и $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ – два *ортонормированных* базиса в $\mathcal{X}$. Пусть $S = (\sigma_{jk})$ – матрица перехода от $\boldsymbol{e}$ к $\boldsymbol{e}'$, т.е. $\boldsymbol{e}' = \boldsymbol{e}S$.

Так как $\boldsymbol{e}'_k = \sum_{j=1}^n \sigma_{jk}\boldsymbol{e}_j$ и так как базисы $\boldsymbol{e}$ и $\boldsymbol{e}'$ ортогональны, то

$$\delta_{k,\ell} = \langle \boldsymbol{e}'_k, \boldsymbol{e}'_\ell \rangle = \sum_{j=1}^n \sum_{t=1}^n \sigma_{jk}\sigma_{t\ell}\langle \boldsymbol{e}_j, \boldsymbol{e}_t \rangle = \sum_{j=1}^n \sigma_{jk}\sigma_{j\ell}. \tag{7.3}$$

Исходя из правил умножения матриц, последнее соотношение в точности означает, что $S^\top S = E$, где E – единичная $n \times n$ матрица, а $A^\top$ обозначает матрицу, транспонированную к матрице A. Проверка этого факта оставляется в качестве *упражнения*.

Таким образом, матрица S перехода от одного ортонормированного базиса к другому обладает тем свойством, что $S^\top S = E$. Из этого вытекает, что $S^\top = S^{-1}$ и, так как, $SS^{-1} = S^{-1}S = E$, то $SS^\top = E$ в рассматриваемом случае.

Определение. *Вещественная $n \times n$-матрица A называется ортогональной, если она удовлетворяет соотношению $A^\top A = E$.*

Предложение. *Ортогональная матрица A невырождена,* $\det A = \pm 1$ *и* $AA^\top = E$.

Доказательство. В самом деле, пусть матрица A ортогональна, т.е. $A^\top A = E$. Следовательно, так как $\det A^\top = \det A$, то $(\det A)^2 = 1$ и следовательно, матрица A невырождена и $\det A = \pm 1$. Отсюда вытекает, что существует обратная матрица, а из свойства ортогональности вытекает, что $A^\top = A^{-1}$. А так как $AA^{-1} = A^{-1}A = E$, то $AA^\top = E$. □

Итак, если базисы $\boldsymbol{e}$ и $\boldsymbol{e}'$ ортонормированы, то матрица перехода S от $\boldsymbol{e}$ к $\boldsymbol{e}'$ является ортогональной. Верно и обратное утверждение: если базис $\boldsymbol{e}$ ортонормированный, а матрица S перехода от $\boldsymbol{e}$ к некоторому базису $\boldsymbol{e}'$ ортогональна, то базис $\boldsymbol{e}'$ также будет ортогональным. В самом деле, достаточно заметить, что имеют место два следующих простых факта. Во-первых, если $\boldsymbol{e}$ – произвольный (не обязательно ортонормированный) базис в евклидовом пространстве $\mathcal{X}$, то

$$G(\boldsymbol{e}) = \boldsymbol{e}^\top \boldsymbol{e}. \tag{7.4}$$

Во-вторых, если S – матрица перехода от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$, то

$$G(\boldsymbol{e}') = \boldsymbol{e}'^\top \boldsymbol{e}' = (\boldsymbol{e}S)^\top(\boldsymbol{e}S) = S^\top \boldsymbol{e}^\top \boldsymbol{e} S = S^\top G(\boldsymbol{e}) S. \tag{7.5}$$

Из формул (7.4) и (7.5) немедленно вытекает, что если базис $\boldsymbol{e}$ – ортонормированный, а матрица S перехода от $\boldsymbol{e}$ к $\boldsymbol{e}'$ ортогональна, то

$$G(\boldsymbol{e}') = S^\top G(\boldsymbol{e}) S = S^\top E S = S^\top S = E,$$

т.е. базис $\boldsymbol{e}'$ – ортонормированный. Таким образом, доказано следующее утверждение.

Теорема 7.12. *Матрица перехода от одного ортонормированного базиса к другому является ортогональной матрицей. Любая ортогональная матрица является матрицей перехода от одного ортонормированного базиса к некоторому другому.*

В завершении этого раздела отметим еще одно простое, но полезное свойство ортогональных матриц: если ортогональная матрица $S = (\sigma_{jk})$ является матрицей перехода от ортонормированного базиса $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ к ортонормированному базису $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n\}$, то $\sigma_{jk} = \langle \boldsymbol{e}_j, \boldsymbol{e}'_k \rangle = \cos\theta_{jk}$, где θ_{jk} – угол между векторами $\boldsymbol{e}_j$ и $\boldsymbol{e}'_k$.

7.4. Эрмитовы пространства

Во многих теоретических и прикладных задачах, требующих использования методов линейной алгебры естественно возникают не вещественные, а комплексные линейные пространства. Соответственно, возникает необходимость рассматривать в таких пространствах и скалярные произведения. Однако заметим, что непосредственный перенос скалярного произведения со случая вещественных пространств на случай пространств комплексных оказывается невозможным.

В самом деле, пусть $\mathcal{X}$ – комплексное линейное пространство (т.е. линейное пространство над полем $\mathcal{F} = \mathbb{C}$ комплексных чисел). Предположим, что в $\mathcal{X}$ задано евклидово скалярное произведение $\langle \cdot, \cdot \rangle_{\mathcal{E}}$. Тогда если $\boldsymbol{x} \in \mathcal{X}$ такой вектор, что $\langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\mathcal{E}} > 0$, то, в силу свойств евклидова скалярного произведения, $\langle i\boldsymbol{x}, i\boldsymbol{x} \rangle_{\mathcal{E}} = -\langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\mathcal{E}} < 0$, что противоречит неотрицательности евклидова скалярного произведения.

Таким образом, для "обслуживания" комплексных пространств необходимо придумать какое-то другое скалярное произведение, более адекватное "комплексной природе" пространства $\mathcal{X}$. Ясно, что какие-то из свойств евклидова скалярного произведения придется для этого ослабить. Оказывается, что достаточно несколько ослабить свойство коммутативности, согласно которому $\langle \boldsymbol{x}, \boldsymbol{y} \rangle_{\mathcal{E}} = \langle \boldsymbol{y}, \boldsymbol{x} \rangle_{\mathcal{E}}$ для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.

Определение. *Комплексное линейное пространство $\mathcal{X}$ является эрмитовым пространством, если для любых двух элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ определено комплексное число $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$, которое называется эрмитовым скалярным произведением $\boldsymbol{x}$ и $\boldsymbol{y}$ такое, что*

(1) $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \overline{\langle \boldsymbol{y}, \boldsymbol{x} \rangle}$ *для любых элементов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, черта обозначает комплексное сопряжение;*
(2) $\langle \boldsymbol{x}_1 + \boldsymbol{x}_2, \boldsymbol{y} \rangle = \langle \boldsymbol{x}_1, \boldsymbol{y} \rangle + \langle \boldsymbol{x}_2, \boldsymbol{y} \rangle$ *для любых элементов $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{y} \in \mathcal{X}$;*
(3) $\langle \lambda\boldsymbol{x}, \boldsymbol{y} \rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ *для любого элемента $\boldsymbol{x} \in \mathcal{X}$ и $\lambda \in \mathbb{C}$;*
(4) *для любого $\boldsymbol{x} \in \mathcal{X}$ величина $\langle \boldsymbol{x}, \boldsymbol{x} \rangle$ – это некоторое неотрицательное вещественное число, которое обращается в нуль если и только если $\boldsymbol{x} = \boldsymbol{0}$.*

Замечание. Говорят, что задание эрмитова скалярного произведения $\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ задает в (комплексном) линейном пространстве $\mathcal{X}$ *структуру эрмитова пространства.*

Эрмитово скалярное произведение обладает следующими свойствами:

(5) $\langle \boldsymbol{x}, \lambda\boldsymbol{y} \rangle = \overline{\lambda}\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$ и $\lambda \in \mathbb{C}$;
(6) $\langle \boldsymbol{x}, \boldsymbol{y}_1 + \boldsymbol{y}_2 \rangle = \langle \boldsymbol{x}, \boldsymbol{y}_1 \rangle + \langle \boldsymbol{x}, \boldsymbol{y}_2 \rangle$ для любых элементов $\boldsymbol{x}, \boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathcal{X}$;

Проверка этих свойств оставляется в качестве простого *упражнения*.

Пример 7.13. Рассмотрим пространство $\mathbb{C}^n$, элементами которого служат всевозможные упорядоченные наборы, состоящие из n комплексных чисел, $n \in \mathbb{N}$. Двум векторам $\boldsymbol{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ и $\boldsymbol{w} = (w_1, \dots, w_n) \in \mathbb{C}^n$ поставим в соответствие число

$$\langle z, w \rangle = z_1\overline{w}_1 + \dots + z_n\overline{w}_n.$$

Легко проверить (проверка оставляется в качестве *упражнения*), что это выражение задает в $\mathbb{C}^n$ эрмитову структуру.

Аналогично, совокупность всех непрерывных комплекснозначных функций на некотором отрезке $[a,b]$ числовой прямой со скалярным произведением, определенным по формуле

$$\langle f,g\rangle = \int_a^b f(t)\overline{g(t)}\,dt, \qquad f,g\in C([a,b])$$

будет эрмитовым пространством.

Упражнение. Выяснить, какими свойствами должна обладать комплексная $n\times n$ матрица $A=(a_{jk})$, $n\in\mathbb{N}$, чтобы выражение

$$\langle z,w\rangle = \sum_{j,k=1}^{n} a_{jk} z_j \overline{w}_k$$

определяло скалярное произведение и чтобы множество $\mathbb{C}^n$ с этим скалярным произведением было эрмитовым пространством.

Неравенство Коши-Буняковского и норма в эрмитовом пространстве. Пусть $\boldsymbol{x}$ и $\boldsymbol{y}$ – произвольные элементы эрмитова пространства $\mathcal{X}$. Тогда

$$\langle \lambda\boldsymbol{x}-\boldsymbol{y}, \lambda\boldsymbol{x}-\boldsymbol{y}\rangle \geqslant 0$$

для любого $\lambda\in\mathbb{C}$. В силу свойств скалярного произведения получаем, что

$$\langle \lambda\boldsymbol{x}-\boldsymbol{y}, \lambda\boldsymbol{x}-\boldsymbol{y}\rangle = \lambda\overline{\lambda}\langle\boldsymbol{x},\boldsymbol{x}\rangle - \lambda\langle\boldsymbol{x},\boldsymbol{y}\rangle - \overline{\lambda}\langle\boldsymbol{y},\boldsymbol{x}\rangle + \langle\boldsymbol{y},\boldsymbol{y}\rangle$$

и, следовательно, первое неравенство этого подраздела имеет вид

$$|\lambda|^2\langle\boldsymbol{x},\boldsymbol{x}\rangle - \lambda(\boldsymbol{x},\boldsymbol{y}) - \overline{\lambda\langle\boldsymbol{x},\boldsymbol{y}\rangle} + \langle\boldsymbol{y},\boldsymbol{y}\rangle \geqslant 0$$

Пусть теперь $\langle\boldsymbol{x},\boldsymbol{y}\rangle = |\langle\boldsymbol{x},\boldsymbol{y}\rangle|e^{i\theta}$ и возьмем число $\lambda = re^{-i\theta}$, где $r\in\mathbb{R}$. Тогда $\lambda\langle\boldsymbol{x},\boldsymbol{y}\rangle = |\langle\boldsymbol{x},\boldsymbol{y}\rangle|$ и для любого $r\in\mathbb{R}$ будет выполнено неравенство

$$r^2\langle\boldsymbol{x},\boldsymbol{x}\rangle - 2r|\langle\boldsymbol{x},\boldsymbol{y}\rangle| + \langle\boldsymbol{y},\boldsymbol{y}\rangle \geqslant 0,$$

из которого, как и в случае обычного (вещественного) неравенства Коши-Буняковского вытекает, что

$$|\langle\boldsymbol{x},\boldsymbol{y}\rangle|^2 \leqslant \langle\boldsymbol{x},\boldsymbol{x}\rangle\,\langle\boldsymbol{y},\boldsymbol{y}\rangle.$$

В случае эрмитовых пространств, как и в случае евклидовых пространств всякое эрмитово пространство будет нормированными, если норму ввести используя (эрмитово) скалярное произведение: $\|\boldsymbol{x}\| = \sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}$. При этом, в частности, будет справедливо неравенство треугольника.

Ортонормированный базис и его свойства. Заметим, что в случае комплексного евклидова пространства понятие угла между элементами теряет смысл (скалярное произведение – комплексное число). Однако понятие ортогональности векторов сохраняется. Соответственно сохраняется понятие ортонормированного базиса, все свойства ортонормированных базисов, изученные выше, аналогичным образом работает процесс ортогонализации.

Небольшое изменение претерпит выражение скалярного произведения через координаты элементов в некотором ортонормированном базисе $\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n\}$ комплексного евклидова пространства $\mathcal{X}$:

$$\langle\boldsymbol{x},\boldsymbol{y}\rangle = \left\langle \sum_{j=1}^{n} x_j\boldsymbol{e}_j, \sum_{k=1}^{n} y_k\boldsymbol{e}_k \right\rangle = \sum_{j,k=1}^{n} x_j\overline{y}_k\langle\boldsymbol{e}_j,\boldsymbol{e}_k\rangle = \sum_{j=1}^{n} x_j\overline{y}_j.$$

Как и в случае евклидовых пространств такое выражение скалярного произведения через координаты вектора является характеристическим свойством ортонормированного базиса.

Заметим, что если $\{\boldsymbol{e}_1,\ldots,\boldsymbol{e}_n\}$ – какой-нибудь (не обязательно ортонормированный) базис в n-мерном эрмитовом пространстве $\mathcal{X}$, то эрмитово скалярное произведение

$\langle \boldsymbol{x}, \boldsymbol{y} \rangle$ векторов $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ выражается через через матрицу Грама $G(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n)$ системы векторов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ (которая вводится абсолютно аналогично вещественному случаю) следующим образом

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = [\boldsymbol{x}]_{\boldsymbol{e}}^{\top} G(\boldsymbol{e}_1, \dots, \boldsymbol{e}_n) \overline{[\boldsymbol{y}]_{\boldsymbol{e}}} = \sum_{j,k=1}^{n} x_j \overline{y}_k \langle \boldsymbol{e}_j, \boldsymbol{e}_k \rangle,$$

где $\overline{[\boldsymbol{y}]_{\boldsymbol{e}}} = (\overline{y}_1, \dots, \overline{y}_n)^{\top}$ при $[\boldsymbol{y}]_{\boldsymbol{e}} = (y_1, \dots, y_n)^{\top}$. Соответственно, базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ эрмитова пространства $\mathcal{X}$ является ортонормированным если и только если его матрица Грама является единичной матрицей.

Доказательства обоих приведенных выше необходимых и достаточных условий, при которых базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ эрмитова пространства $\mathcal{X}$ является ортонормированным проходят аналогично доказательству Теоремы 7.9.

Определение. *Комплексная $n \times n$-матрица U называется унитарной, если она удовлетворяет соотношению $U^*U = E$, где $U^* = \overline{U}^{\top}$, а $\overline{U}$ – матрица, комплексно сопряженная матрице U. Матрица U^* называется матрицей, эрмитово сопряженной матрице U.*

Предложение. *Унитарная матрица U невырождена, $|\det U| = 1$ и $UU^* = E$.*

Это утверждение доказывается абсолютно аналогично соответствующему свойству ортогональных матриц. Доказательство предлагается провести в качестве *упражнения*. Далее, абсолютно аналогично случаю евклидовых пространства устанавливается следующее утверждение.

Теорема 7.14. *Матрица перехода от одного ортонормированного базиса к другому в эрмитовом пространстве является унитарной. Обратно, всякая унитарная матрица является матрицей перехода от одного ортонормированного базиса к некоторому другому ортонормированному базису.*

Точный смысл понятия эрмитовой сопряженности матриц будет выяснен позже. Ясно, что имеет место следующее утверждение.

Предложение. *Вещественная ортогональная матрица является унитарной.*

Замечание. В связи с утверждением последнего Предложения необходимо заметить, что понятия эрмитова скалярного произведения, эрмитовой сопряженности и унитарности матриц и т.п. – это *естественное обобщение* соответствующих понятий (евклидова скалярного произведения, ортогональности матриц и т.п.), возникших в теории вещественных линейных и евклидовых пространств на случай комплексных пространств.

7.5. Линейные функционалы и полуторалинейные формы в эрмитовом пространстве

Пусть $\mathcal{X}$ – эрмитово пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$.

Общий вид линейного функционала в эрмитовом пространстве. Установим следующий, важный для дальнейших приложений, результат:

Теорема 7.15. *Пусть $\mathcal{X}$ – эрмитово пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$. Для любого линейного функционала $\mathfrak{f} \in \mathfrak{L}(\mathcal{X}, \mathbb{C})$ существует единственный вектор $\boldsymbol{w} \in \mathcal{X}$ такой, что*

$$\mathfrak{f}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{w} \rangle \tag{7.6}$$

для всех $\boldsymbol{x} \in \mathcal{X}$, где $\langle \cdot, \cdot \rangle$ – (эрмитово) скалярное произведение в $\mathcal{X}$.

Доказательство. Начнем с доказательства существования требуемого вектора $\boldsymbol{w}$. Выберем ортонормированный базис $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в $\mathcal{X}$ и рассмотрим элемент

$$\boldsymbol{w} = \sum_{k=1}^{n} w_k \boldsymbol{e}_k := \sum_{k=1}^{n} \overline{\mathfrak{f}(\boldsymbol{e}_k)}\, \boldsymbol{e}_k$$

(напомним, что черта означает комплексное сопряжение). Для произвольного элемента $\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{e}_k \in \mathcal{X}$ получаем

$$\mathfrak{f}(\boldsymbol{x}) = \sum_{k=1}^{n} x_k \mathfrak{f}(\boldsymbol{e}_k) = \sum_{k=1}^{n} x_k \overline{w}_k = \langle \boldsymbol{x}, \boldsymbol{w} \rangle,$$

т.е. элемент $\boldsymbol{w}$ удовлетворяет условиям доказываемого утверждения.

Проверим теперь единственность такого элемента $\boldsymbol{w}$. Пусть, от противного, найдутся два элемента $\boldsymbol{w}_1 \in \mathcal{X}$ и $\boldsymbol{w}_2 \in \mathcal{X}$ такие, что $\mathfrak{f}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{w}_1 \rangle = \langle \boldsymbol{x}, \boldsymbol{w}_2 \rangle$. Тогда для любого $\boldsymbol{x} \in \mathcal{X}$ имеет место равенство $\langle \boldsymbol{x}, \boldsymbol{w}_1 - \boldsymbol{w}_2 \rangle = 0$. Подставляя в это равенство вместо $\boldsymbol{x}$ элемент $\boldsymbol{w}_1 - \boldsymbol{w}_2 \in \mathcal{X}$, получаем, что $\|\boldsymbol{w}_1 - \boldsymbol{w}_2\|^2 = \langle \boldsymbol{w}_1 - \boldsymbol{w}_2, \boldsymbol{w}_1 - \boldsymbol{w}_2 \rangle = 0$ и, следовательно, $\boldsymbol{w}_1 = \boldsymbol{w}_2$. $\square$

В случае вещественных линейных пространств совершенно аналогично можно доказать (это предлагается сделать в качестве *упражнения*) следующее утверждение.

Теорема 7.16. *Пусть $\mathcal{X}$ – n-мерное евклидово пространство, $n \in \mathbb{N}$. Для любой линейной формы $\mathfrak{f} \in \mathfrak{L}(\mathcal{X}, \mathbb{R})$ существует единственный вектор $\boldsymbol{w} \in \mathcal{X}$ такой, что $\mathfrak{f}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{w} \rangle$ для всех $\boldsymbol{x} \in \mathcal{X}$, где $\langle \cdot, \cdot \rangle$ – (евклидово) скалярное произведение в $\mathcal{X}$.*

Каноническое спаривание $\mathcal{X}$ и $\mathcal{X}^*$ в случае эрмитова пространства. Пусть $\mathcal{X}$ – n-мерное эрмитово пространство с (эрмитовым) скалярным произведением $\langle \cdot, \cdot \rangle$. Рассмотрим сопряженное пространство $\mathcal{X}^*$. Так как $\dim \mathcal{X} = n$ и $\dim \mathcal{X}^* = n$, то $\mathcal{X}$ и $\mathcal{X}^*$ изоморфны как линейные пространства.

Построим специальный изоморфизм между $\mathcal{X}$ и $\mathcal{X}^*$ следующим образом: элементу $\boldsymbol{x} \in \mathcal{X}$ поставим в соответствие линейный функционал $\mathfrak{f} \in \mathcal{X}^*$ такой, что $\mathfrak{f}(\boldsymbol{y}) = \langle \boldsymbol{y}, \boldsymbol{x} \rangle$ (напомним, что в силу Теоремы 7.15 функционал f определяется таким вектором $\boldsymbol{x}$ единственным образом). Этот изоморфизм естественно назвать каноническим (или естественным) изоморфизмом между $\mathcal{X}$ и $\mathcal{X}^*$.

Пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – ортонормированный базис в $\mathcal{X}$, т.е. скалярные произведения базисных векторов равны $\langle \boldsymbol{e}_j, \boldsymbol{e}_k \rangle = \delta_{j,k}$. В этом случае линейные функционалы $\mathfrak{e}^j(\cdot) = \langle \cdot, \boldsymbol{e}_j \rangle$, $j = 1, \ldots, n$, определяемые векторами базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, будут образовывать базис, двойственный к базису $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. В самом деле, $\mathfrak{e}^j(\boldsymbol{e}_k) = \langle \boldsymbol{e}_k, \boldsymbol{e}_j \rangle = \delta_{k,j}$. Для построенного базиса $\{\mathfrak{e}^1, \ldots, \mathfrak{e}^n\}$ будет использоваться обозначение $\widehat{\boldsymbol{e}}$, где $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – исходный ортонормированный базис.

Далее, для произвольного вектора $\boldsymbol{x} \in \mathcal{X}$ определим вектор $\overline{\boldsymbol{x}}$ следующим образом: если $(x_1, \ldots, x_n)^\top$ – координатный столбец вектора $\boldsymbol{x}$ в рассматриваемом (ортонормированном базисе), то вектор $\overline{\boldsymbol{x}} = \overline{x}_1 \boldsymbol{e}_1 + \cdots \overline{x}_n \boldsymbol{e}_n$ (черта, как обычно, означает комплексное сопряжение). Таким образом, если $\mathcal{X}$ – вещественное пространство, то векторы $\boldsymbol{x}$ и $\overline{\boldsymbol{x}}$ совпадают.

Напомним, что каноническим спариванием между пространствами $\mathcal{X}^*$ и $\mathcal{X}$ является $\langle \mathfrak{f} | \boldsymbol{x} \rangle = \mathfrak{f}(\boldsymbol{x})$. Из доказательства Теоремы 7.15 вытекает, что оно задается формулой

$$\langle \mathfrak{f} | \boldsymbol{x} \rangle = \langle \boldsymbol{x}, \overline{\boldsymbol{f}} \rangle,$$

где $\mathfrak{f} \in \mathcal{X}^*$, $\boldsymbol{x} \in \mathcal{X}$, а $\boldsymbol{f} \in \mathcal{X}$ – такой вектор, что его координаты относительно базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ совпадают с координатами элемента $\mathfrak{f}$ относительно базиса $\{\mathfrak{e}^1, \ldots, \mathfrak{e}^n\}$.

Общий вид полуторалинейных форм в эрмитовом пространстве. Пусть $\mathcal{X}$ – эрмитово пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$ и пусть $\mathcal{F}$ – полуторалинейная форма в $\mathcal{X}$. Установим результат об общем виде полуторалинейной формы в эрмитовом пространстве, аналогичный результату Теоремы 7.15 об общем виде линейного функционала.

Теорема 7.17. *Пусть $\mathcal{F}$ – полуторалинейная форма в эрмитовом пространстве $\mathcal{X}$. Тогда существует единственный линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ такой, что*

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \mathcal{A}\, \boldsymbol{y} \rangle \tag{7.7}$$

для всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$, где $\langle \cdot, \cdot \rangle$ – (эрмитово) скалярное произведение в $\mathcal{X}$.

Доказательство. Пусть $\boldsymbol{y}$ – произвольный фиксированный элемент пространства $\mathcal{X}$. Тогда $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y})$, как функция одного аргумента $\boldsymbol{x}$, является линейно формой на пространстве $\mathcal{X}$. В силу Теоремы 7.15 существует единственный $\boldsymbol{w} = \boldsymbol{w}_{\boldsymbol{y}} \in \mathcal{X}$ такой, что

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \boldsymbol{w}_{\boldsymbol{y}} \rangle.$$

Таким образом, каждому элементу $\boldsymbol{y} \in \mathcal{X}$ однозначно поставлен в соответствие элемент $\boldsymbol{w}_{\boldsymbol{y}} \in \mathcal{X}$. По определению это означает, что определен оператор $\mathcal{A}$ такой, что $\boldsymbol{w}_{\boldsymbol{y}} = \mathcal{A}\, \boldsymbol{y}$. Линейность этого оператора вытекает из свойств (1)-(3) полуторалинейных форм, детали проверки этой линейности оставляются в качестве *упражнения*.

Проверим теперь единственность такого оператора $\mathcal{A}$. Пусть $\mathcal{A}_1$ и $\mathcal{A}_2$ – два оператора из $\mathfrak{L}(\mathcal{X}, \mathcal{X})$ таких, что $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \mathcal{A}_1 \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathcal{A}_2 \boldsymbol{y} \rangle$. Из этого следует, что при всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ верно равенство $\langle \boldsymbol{x}, \mathcal{A}_1 \boldsymbol{y} - \mathcal{A}_2 \boldsymbol{y} \rangle = 0$. Подставляя в это равенство $\mathcal{A}_1 \boldsymbol{y} - \mathcal{A}_2 \boldsymbol{y} \in \mathcal{X}$ вместо $\boldsymbol{x}$ получим $\| \mathcal{A}_1 \boldsymbol{y} - \mathcal{A}_2 \boldsymbol{y} \| = \langle \mathcal{A}_1 \boldsymbol{y} - \mathcal{A}_2 \boldsymbol{y}, \mathcal{A}_1 \boldsymbol{y} - \mathcal{A}_2 \boldsymbol{y} \rangle = 0$ и, следовательно, для любого $\boldsymbol{y} \in \mathcal{X}$ верно равенство $\mathcal{A}_1 \boldsymbol{y} = \mathcal{A}_2 \boldsymbol{y}$. Следовательно, $\mathcal{A}_1 = \mathcal{A}_2$ и теорема доказана. $\square$

Замечание. Утверждение Теоремы 7.17 остается верным и в случае, когда $\mathcal{X}$ – евклидово пространство, а $\mathcal{F}$ – *билинейная форма* на $\mathcal{X}$. В этом случае также существует единственный линейный оператор $\mathcal{A}$ на $\mathcal{X}$, такой, что $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \mathcal{A}\, \boldsymbol{y} \rangle$, где $\langle \cdot, \cdot \rangle$ – (евклидово) скалярное произведение в $\mathcal{X}$.

Замечание. Наряду с представлением (7.7) полуторалинейная форма $\mathcal{F}$ на комплексном евклидовом пространстве $\mathcal{X}$ может быть представлена в виде

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}'\, \boldsymbol{x}, \boldsymbol{y} \rangle,$$

причем, как и в случае представления (7.7), оператор $\mathcal{A}'$ определяется однозначно.

Проверим этот факт. В самом деле, функция $\mathcal{F}_1(\boldsymbol{y}, \boldsymbol{x}) := \overline{\mathcal{F}(\boldsymbol{x}, \boldsymbol{y})}$ является полуторалинейной формой (это непосредственно вытекает из свойств (1)-(3) полуторалинейных форм, проверка оставляется в качестве *упражнения*). Далее, для формы $\mathcal{F}_1$ существует единственный линейный оператор $\mathcal{A}_1$ такой, что $\mathcal{F}_1(\boldsymbol{y}, \boldsymbol{x}) = \langle \boldsymbol{y}, \mathcal{A}_1 \boldsymbol{x} \rangle$. Тогда $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \overline{\mathcal{F}_1(\boldsymbol{y}, \boldsymbol{x})} = \overline{\langle \boldsymbol{y}, \mathcal{A}_1 \boldsymbol{x} \rangle} = \langle \mathcal{A}_1 \boldsymbol{x}, \boldsymbol{y} \rangle$.

Найдем связь матрицы полуторалинейной формы и матрицы представляющего ее линейного оператора. Пусть $\boldsymbol{e}$ – ортонормированный базис в (эрмитовом) пространстве $\mathcal{X}$ – эрмитово, а базис $\boldsymbol{e}$ является ортонормированным. Нам удобно использовать представление полуторалинейной формы $\mathcal{F}$ в виде

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \mathcal{A}\, \boldsymbol{y} \rangle,$$

где $\mathcal{A}$ – представляющий форму $\mathcal{F}$ линейный оператор. Пусть A – матрица оператора $\mathcal{A}$, а F – матрица формы $\mathcal{F}$ в базисе $\boldsymbol{e}$. Тогда

$$f_{jk} = \mathcal{F}(\boldsymbol{e}_j, \boldsymbol{e}_k) = \langle \boldsymbol{e}_j, \mathcal{A}\, \boldsymbol{e}_k \rangle = \left\langle \boldsymbol{e}_j, \sum_{r=1}^{n} a_{rk} \boldsymbol{e}_r \right\rangle = \sum_{r=1}^{n} \langle \boldsymbol{e}_j, a_{rk} \boldsymbol{e}_r \rangle = \sum_{r=1}^{n} \bar{a}_{rk} \langle \boldsymbol{e}_j, \boldsymbol{e}_r \rangle = \bar{a}_{jk},$$

причем последнее равенство имеет место в силу того, что базис $\boldsymbol{e}$ – ортонормированный. Полученное равенство в точности означает, что $F = A^*$.

Замечание. Совершенно аналогично проверяется, что если пространство $\mathcal{X}$ евклидово, а $\mathcal{F}$ – билинейная форма на $\mathcal{X}$, то ее матрица в любом ортонормированном базисе совпадает с матрицей линейного оператора, представляющего эту билинейную форму в виде (7.7).

Упражнение. Выяснить как связаны матрица F полуторалинейной формы $\mathcal{F}$ с матрицей A_1 линейного оператора $\mathcal{A}_1$, представляющего $\mathcal{F}$ в виде $\mathcal{F}(\boldsymbol{x},\boldsymbol{y}) = \langle \mathcal{A}_1\, \boldsymbol{x}, \boldsymbol{y}\rangle$. Во что превратиться эта связь в случае билинейных форм?

7.6. Норма линейного оператора

В этом разделе $\mathcal{X}$ – вещественное или комплексное *нормированное* пространство с нормой $\|\cdot\|$. Рассмотрим пространство $\mathfrak{L}(\mathcal{X},\mathcal{X})$ и выясним вопрос, как можно ввести на нем структуру нормированного пространства. Оказывается, что существует по меньшей мере два способа определить норму в пространстве $\mathfrak{L}(\mathcal{X},\mathcal{X})$ так, чтобы эта норма была естественно согласованной с нормой $\|\cdot\|$ в исходном пространстве $\mathcal{X}$.

Рассмотрим линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ и *определим* величину

$$\|\,\mathcal{A}\,\|_s := \sup\{\|\,\mathcal{A}\,\boldsymbol{x}\| :\ \boldsymbol{x} \in \mathcal{X},\ \|\boldsymbol{x}\| = 1\}.$$

Заметим, однако, что в общем случае величина $\|\,\mathcal{A}\,\|_s$ может и не существовать.

Пример 7.18. Рассмотрим пространство $\mathbb{R}[t]$ многочленов с вещественными коэффициентами от одной вещественной переменной с нормой, определенной соотношением

$$\|P\|^2 = \int_0^1 P(t)^2\,dt, \qquad P \in \mathbb{R}[t]$$

Пусть $\mathcal{D} : P \mapsto P'$ – оператор дифференцирования. Так как $\|\sqrt{2n+1}\; t^n\| = 1$, то

$$\|\,\mathcal{D}(\sqrt{2n+1}\; t^n)\| = n\sqrt{2n-1}\;\|t^{n-1}\| = n\sqrt{(2n+1)/(2n-1)},$$

а последняя величина с ростом n неограничено возрастает.

Предложение 7.19. *Если пространство $\mathcal{X}$ конечномерно, то для любого линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ величина $\|\,\mathcal{A}\,\|_s$ существует и обладает всеми свойствами нормы (см. раздел 7.1).*

Доказательство. Пусть $\mathcal{X}$ – комплексное пространство размерности $\dim\mathcal{X} = n$. Рассмотрим в пространстве $\mathbb{C}^n$ множество $S = \{z = (z_1,\dots,z_n) :\ |z_1|^2 + \dots + |z_n|^2 = 1\}$. Это множество – компакт (т.е. замкнутое и ограниченное множество) в $\mathbb{C}^n$. Рассмотрим далее функцию $f(z) = \|\,\mathcal{A}\,\boldsymbol{v}_z\|$ при $z \in S$, где $\boldsymbol{v}_z \in \mathcal{X}$ – это такой вектор, что $[\boldsymbol{v}_z] = (z_1,\dots,z_n)^\top$ (относительно некоторого базиса в $\mathcal{X}$). Заметим, что функция f непрерывна на S (это непосредственно вытекает из определения нормы и определения линейного оператора; проверка оставляется в качестве *упражнения*). По теореме Больцано-Вейерштрасса, непрерывная на компакте S функция f ограничена и достигает на нем своей точной верхней грани. Следовательно, найдется вектор $\boldsymbol{v}_0 \in \mathcal{X}$ такой, что $\|\,\mathcal{A}\,\|_s = \|\,\mathcal{A}\,\boldsymbol{v}_0\|$.

Проверим, что величина $\|\,\mathcal{A}\,\|_S$ обладает свойствами нормы. Обозначим через $\mathcal{S}(\mathcal{X})$ множество $\mathcal{S}(\mathcal{X}) := \{\boldsymbol{x} \in \mathcal{X} :\ \|\boldsymbol{x}\| = 1\}$. Заметим, что

$$\begin{aligned}\|\,\mathcal{A}\,\|_s = 0 \quad &\Longleftrightarrow \quad \forall \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\ \|\,\mathcal{A}\,\boldsymbol{x}\| = 0 \quad \Longleftrightarrow \\ &\Longleftrightarrow \quad \forall \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\ \ \mathcal{A}\,\boldsymbol{x} = 0 \quad \Longleftrightarrow \quad \forall \boldsymbol{x} \in \mathcal{X}\ \ \mathcal{A}\,\boldsymbol{x} = 0 \quad \Longleftrightarrow \quad \mathcal{A} = \mathcal{O}\,.\end{aligned}$$

Таким образом, $\|\,\mathcal{A}\,\|_S = 0$, если и только если $\mathcal{A}$ – нулевой оператор. Далее, $\forall \alpha \in \mathbb{C}$

$$\|\alpha\,\mathcal{A}\,\|_s = \sup\{\|\alpha\,\mathcal{A}\,\boldsymbol{x}\| :\ \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\} = |\alpha|\sup\{\|\,\mathcal{A}\,\boldsymbol{x}\| :\ \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\} = |\alpha|\|\,\mathcal{A}\,\|_s$$

и, наконец, для любых двух операторов $\mathcal{A},\mathcal{B} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$

$$\begin{aligned}\|\,\mathcal{A} + \mathcal{B}\,\|_s &= \sup\{\|\,\mathcal{A}\,\boldsymbol{x} + \mathcal{B}\,\boldsymbol{x}\| :\ \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\} \leqslant \\ &\leqslant \sup\{\|\,\mathcal{A}\,\boldsymbol{x}\| :\ \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\} + \sup\{\|\,\mathcal{B}\,\boldsymbol{x}\| :\ \boldsymbol{x} \in \mathcal{S}(\mathcal{X})\} = \|\,\mathcal{A}\,\|_s + \|\,\mathcal{B}\,\|_s. \quad \square\end{aligned}$$

Замечание. Пусть пространство $\mathcal{X}$ конечномерно, так что норма $\|\,\mathcal{A}\,\|_s$ существует для любого оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$. При этом верно неравенство

$$\|\,\mathcal{A}\,\boldsymbol{x}\| \leqslant \|\,\mathcal{A}\,\|_s\|\boldsymbol{x}\| \qquad \text{для любого} \quad \boldsymbol{x} \in \mathcal{X}. \tag{7.8}$$

Упражнение. Доказать неравенство (7.8).

Определение. *Пусть $\mathcal{X}$ – нормированное линейное пространство. Линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется ограниченным если существует неотрицательное вещественное число N такое, что для любого вектора $\boldsymbol{x} \in \mathcal{X}$ верно неравенство*

$$\|\mathcal{A}\,\boldsymbol{x}\| \leqslant N\,\|\boldsymbol{x}\|. \tag{7.9}$$

Нижняя грань $\inf N$ *множества всех констант N, для которых справедливо неравенство (7.9), обозначается символом $\|\mathcal{A}\|_i$.*

Предложение 7.20. *Если пространство $\mathcal{X}$ – конечномерно, то каждый оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ ограничен. В этом случае $\|\mathcal{A}\|_i = \|\mathcal{A}\|_s$.*

Доказательство. Из неравенства (7.8) следует, что неравенство (7.9) выполняется с некоторым $N \leqslant \|\mathcal{A}\|_s$. Так что $\|\mathcal{A}\|_i \leqslant \|\mathcal{A}\|_s$. С другой стороны, для любого $\boldsymbol{x} \in \mathcal{S}(\mathcal{X})$ имеет место неравенство $\|\mathcal{A}\,\boldsymbol{x}\| \leqslant \|\mathcal{A}\|_i\|\boldsymbol{x}\| = \|\mathcal{A}\|_i$, откуда $\|\mathcal{A}\|_s \leqslant \|\mathcal{A}\|_i$. Итак, $\mathcal{A}$ ограничен и $\|\mathcal{A}\|_i = \|\mathcal{A}\|_s$. Что и требовалось. □

Определение. *В случае конечномерного линейного нормированного пространства $\mathcal{X}$ величину $\|\mathcal{A}\| = \|\mathcal{A}\|_s = \|\mathcal{A}\|_i$ называют нормой оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$.*

Оператор дифференцирования в примере 7.18 неограничен.

Упражнение. Проверить, что оператор $M : P(t) \mapsto tP(t)$ в пространстве $\mathbb{R}[t]$ ограничен и найти его норму.

Норма единичного оператора $\mathcal{E} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ равна 1, так как $\|\mathcal{E}\,\boldsymbol{x}\| = \|\boldsymbol{x}\|$ для любого $\boldsymbol{x} \in \mathcal{X}$. Рассмотрим еще один важный для дальнейшего пример.

Пример 7.21. Пусть $\mathcal{X}$ – евклидово пространство со скалярным произведением $\langle\cdot,\cdot\rangle$, и пусть оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ имеет в некотором ортонормированном базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$ диагональный вид, т.е. $\mathcal{A}\,\boldsymbol{e}_j = \lambda_j \boldsymbol{e}_j$ при $j = 1, \ldots, n$, где $\lambda_j \in \mathbb{R}$ – некоторые вещественные числа (ясно, что все они будут собственными числами оператора $\mathcal{A}$). Без ограничения общности предположим, что $\lambda_1 \geqslant \cdots \geqslant \lambda_n$. Возьмем произвольный вектор $\boldsymbol{x} \in \mathcal{X}$. Если $\boldsymbol{x} = \sum_{j=1}^n x_j \boldsymbol{e}_j$, то $\mathcal{A}\,\boldsymbol{x} = \sum_{j=1}^n \lambda_j x_j \boldsymbol{e}_j$ и

$$\|\mathcal{A}\,\boldsymbol{x}\|^2 = \langle \mathcal{A}\,\boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}\rangle = \sum_{j=1}^n \lambda_j^2 x_j^2 \leqslant \lambda_1^2 \sum_{j=1}^n x_j^2 = \lambda_1^2 \|\boldsymbol{x}\|^2.$$

Из этого следует, что $\|\mathcal{A}\,\boldsymbol{x}\| \leqslant |\lambda_1|\|\boldsymbol{x}\|$ и, окончательно, $\|\mathcal{A}\| \leqslant |\lambda_1|$. А так как $\|\mathcal{A}\,\boldsymbol{e}_1\| = |\lambda_1|$, то $\|\mathcal{A}\| = |\lambda_1|$.

Выясним в завершении этого раздела, как ведет себя норма относительно операции композиции линейный операторов. Имеет место следующий результат.

Предложение 7.22. *Если $\mathcal{A}, \mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – два ограниченных оператора на нормированном линейном пространстве $\mathcal{X}$, то оператор $\mathcal{A}\,\mathcal{B}$ является ограниченным и*

$$\|\mathcal{A}\,\mathcal{B}\| \leqslant \|\mathcal{A}\|\|\mathcal{B}\|. \tag{7.10}$$

Доказательство. В самом деле, для любого $\boldsymbol{x} \in \mathcal{X}$ имеют место неравенства $\|(\mathcal{A}\,\mathcal{B})(\boldsymbol{x})\| = \|\mathcal{A}(\mathcal{B}\,\boldsymbol{x})\| \leqslant \|\mathcal{A}\|\,\|\mathcal{B}\,\boldsymbol{x}\| \leqslant \|\mathcal{A}\|\,\|\mathcal{B}\|\,\|\boldsymbol{x}\|$, из которых все и вытекает. □

Следствие 7.23. *Если $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – ограниченный оператор на нормированном линейном пространстве $\mathcal{X}$, то $\|\mathcal{A}^m\| \leqslant \|\mathcal{A}\|^m$ для любого $m \in \mathbb{N}$.*

Пример 7.24. Рассмотрим в пространстве $\mathbb{R}^2$ операторы $\mathcal{P}_x$ и $\mathcal{P}_y$ проектирования на координатные оси OX и OY соответственно. Как несложно подсчитать, $\|\mathcal{P}_x\| = \|\mathcal{P}_y\| = 1$, а $\mathcal{P}_x\,\mathcal{P}_y = \mathcal{P}_y\,\mathcal{P}_x = \mathcal{O}$. Так что неравенство (7.10) не может, в общем случае, превратится в равенство.

РАЗДЕЛ 8

Линейные операторы в эрмитовых и евклидовых пространствах

В этом разделе, если не оговорено противное, мы будем считать, что $\mathcal{X}$ – это эрмитово пространство, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$. При этом через $\langle \cdot, \cdot \rangle$ мы будем, как обычно, обозначать (эрмитово) скалярное произведение в $\mathcal{X}$.

8.1. Самосопряженные операторы в эрмитовом пространстве

Определение. *Оператор $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется эрмитово сопряженным к оператору $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$, если для любых элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ выполняется равенство*

$$\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathcal{B}\, \boldsymbol{y} \rangle. \tag{8.1}$$

Замечание. Еще раз обратим внимание, что в этом определении рассматриваются только операторы, действующие в пространстве $\mathcal{X}$, а свойство сопряженности формулируется в терминах скалярного произведения в $\mathcal{X}$. Таким образом, введенное здесь понятие эрмитово сопряженного оператора работает только в эрмитовом пространстве и *отличается* от общего определения сопряженного оператора.

Однако в дальнейшем мы будем использовать называть эрмитово сопряженные операторы просто сопряженными, а корректность использования такой терминологии будет показана в разделе 8.2 ниже).

Линейность сопряженного оператора легко проверить непосредственно. В самом деле, пусть $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{y}_1 \in \mathcal{X}$, $\boldsymbol{y}_2 \in \mathcal{X}$ и пусть $\alpha \in \mathbb{C}$, $\beta \in \mathbb{C}$. Тогда

$$\langle \mathcal{A}\, \boldsymbol{x}, \alpha \boldsymbol{y}_1 + \beta \boldsymbol{y}_2 \rangle = \overline{\alpha} \langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y}_1 \rangle + \overline{\beta} \langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y}_2 \rangle = \overline{\alpha} \langle \boldsymbol{x}, \mathcal{B}\, \boldsymbol{y}_1 \rangle + \overline{\beta} \langle \boldsymbol{x}, \mathcal{B}\, \boldsymbol{y}_2 \rangle = \langle \boldsymbol{x}, \mathcal{B}(\alpha \boldsymbol{y}_1 + \beta \boldsymbol{y}_2) \rangle.$$

Предложение 8.1. *Для любого линейного оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ существует единственный сопряженный оператор.*

Доказательство. Заметим, что выражение $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) := \langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y} \rangle$ представляет собой полуторалинейную форму на пространстве $\mathcal{X}$. Согласно утверждению Теоремы 7.17 существует единственный линейный оператор $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ такой, что форма $\mathcal{F}$ представляется в виде $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \mathcal{B}\, \boldsymbol{y} \rangle$. Следовательно, $\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathcal{B}\, \boldsymbol{y} \rangle$. По определению, оператор $\mathcal{B}$ является сопряженным к $\mathcal{A}$ оператором. Единственность сопряженного оператора вытекает из единственности представления полуторалинейной формы $\mathcal{F}$ в виде $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{x}, \mathcal{B}\, \boldsymbol{y} \rangle$. □

Определение. *Оператор, сопряженный к оператору $\mathcal{A}$ в смысле определения, данного в этом разделе, обозначается $\mathcal{A}^*$.*

Замечание. Корректность использования символа $\mathcal{A}^*$ для обозначения оператора, сопряженного к оператору $\mathcal{A}$ в смысле только что данного определения будет обоснована в разделе 8.2 ниже.

Предложение 8.2. *Пусть $\mathcal{X}$ – n-мерное эрмитово пространство и пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ и $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – линейные операторы. Тогда*

(1) $\mathcal{E}^* = \mathcal{E}$;
(2) $(\mathcal{A} + \mathcal{B})^* = \mathcal{A}^* + \mathcal{B}^*$;
(3) $(\alpha\, \mathcal{A})^* = \overline{\alpha}\, \mathcal{A}^*$ *для любого* $\alpha \in \mathbb{C}$;
(4) $(\mathcal{A}^*)^* = \mathcal{A}$;
(5) $(\mathcal{A}\, \mathcal{B})^* = \mathcal{B}^*\, \mathcal{A}^*$.

Доказательство. Проверка свойств (1) и (2) элементарна и оставляется в качестве *упражнения*. Проверим свойства (3)-(5). Пусть $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ – произвольные элементы. Тогда

$$\langle(\alpha\,\mathcal{A})\boldsymbol{x}, \boldsymbol{y}\rangle = \alpha\langle\mathcal{A}\,\boldsymbol{x}, \boldsymbol{y}\rangle = \alpha\langle\boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{y}\rangle = \langle\boldsymbol{x}, (\overline{\alpha}\,\mathcal{A}^*)\boldsymbol{y}\rangle$$

с учетом свойств (эрмитова) скалярного произведения и, следовательно, свойство (3) выполняется. Свойство (4) устанавливается следующим образом:

$$\langle\mathcal{A}\,\boldsymbol{x}, \boldsymbol{y}\rangle = \langle\boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{y}\rangle = \overline{\langle\mathcal{A}^*\,\boldsymbol{y}, \boldsymbol{x}\rangle} = \overline{\langle\boldsymbol{y}, (\mathcal{A}^*)^*\boldsymbol{x}\rangle} = \langle(\mathcal{A}^*)^*\boldsymbol{x}, \boldsymbol{y}\rangle.$$

И, наконец, свойство (5) вытекает из следующей цепочки равенств:

$$\langle(\mathcal{A}\,\mathcal{B})\boldsymbol{x}, \boldsymbol{y}\rangle = \langle\mathcal{A}(\mathcal{B}\,\boldsymbol{x}), \boldsymbol{y}\rangle = \langle\mathcal{B}\,\boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{y}\rangle = \langle\boldsymbol{x}, \mathcal{B}^*(\mathcal{A}^*\,\boldsymbol{y})\rangle = \langle\boldsymbol{x}, (\mathcal{B}^*\,\mathcal{A}^*)\boldsymbol{y}\rangle. \quad \square$$

Замечание. В случае, когда $\mathcal{X}$ – евклидово (т.е. вещественное) пространство, понятие сопряженного оператора вводится аналогично и свойства (1), (2), (4) и (5) сопряженного оператора остаются в силе. Их доказательства проходят аналогично тому, как доказывались соответствующие свойства в Предложении 8.2 за тем исключением, что в проверке свойства (4) не надо использовать операцию комплексного сопряжения. Свойство (3) в этом случае измениться на следующее $(\alpha\,\mathcal{A})^* = \alpha\,\mathcal{A}^*$, $\alpha \in \mathbb{R}$, но его доказательство аналогично доказательству свойства (3) для комплексных пространств.

Определение. *Линейный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется самосопряженным, если $\mathcal{A} = \mathcal{A}^*$ (это определение работает как в случае комплексного, так и в случае вещественного пространства $\mathcal{X}$).*

Простейшим примером самосопряженного оператора является единичный оператор $\mathcal{E}$ (см. свойство (1) в Предложении 8.2). Имеет место следующий результат о разложении произвольного оператора, действующего в эрмитовом пространстве $\mathcal{X}$ на "действительную" и "мнимую" части.

Предложение 8.3. *Пусть $\mathcal{X}$ – эрмитово пространство. Тогда для любого оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ справедливо разложение*

$$\mathcal{A} = \mathcal{A}_R + i\,\mathcal{A}_I, \tag{8.2}$$

где $\mathcal{A}_R$ и $\mathcal{A}_I$ – самосопряженные операторы.

В самом деле, из Предложения 8.2 о свойствах сопряженных операторов вытекает, что операторы $\mathcal{A}_R$ и $\mathcal{A}_I$, определенные следующим образом:

$$\mathcal{A}_R := \frac{1}{2}\Big(\mathcal{A} + \mathcal{A}^*\Big), \qquad \mathcal{A}_I := \frac{i}{2}\Big(\mathcal{A}^* - \mathcal{A}\Big),$$

являются самосопряженными и разложение (8.2) имеет место.

Для того, чтобы сформулировать следующее свойство самосопряженных операторов напомним, что два линейных оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ и $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ *коммутируют*, если $\mathcal{A}\,\mathcal{B} = \mathcal{B}\,\mathcal{A}$. Имеет место следующее утверждение.

Предложение 8.4. *Пусть операторы $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ и $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ являются самосопряженными. Тогда оператор $\mathcal{A}\,\mathcal{B}$ является самосопряженным если и только если операторы $\mathcal{A}$ и $\mathcal{B}$ коммутируют.*

Доказательство. В самом деле, в силу свойства (5) сопряженного оператора (см. Предложение 8.2), самосопряженности $\mathcal{A}$ и $\mathcal{B}$ и в силу того, что $\mathcal{A}\,\mathcal{B} = \mathcal{B}\,\mathcal{A}$ получаем

$$(\mathcal{A}\,\mathcal{B})^* = \mathcal{B}^*\,\mathcal{A}^* = \mathcal{B}\,\mathcal{A} = \mathcal{A}\,\mathcal{B}$$

и следовательно, оператор $\mathcal{A}\,\mathcal{B}$ является самосопряженным. Обратно, если оператор $\mathcal{A}\,\mathcal{B}$ является самосопряженным, то

$$\mathcal{A}\,\mathcal{B} = (\mathcal{A}\,\mathcal{B})^* = \mathcal{B}^*\,\mathcal{A}^* = \mathcal{B}\,\mathcal{A}$$

(последнее равенство верно в силу самосопряженности $\mathcal{A}$ и $\mathcal{B}$) и, следовательно, операторы $\mathcal{A}$ и $\mathcal{B}$ коммутируют. $\square$

Отметим еще три полезных свойства самосопряженных операторов.

Предложение 8.5. *Пусть оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – самосопряженный. Тогда*

(1) $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$ *для любого* $\boldsymbol{x} \in \mathcal{X}$*;*
(2) *все собственные значения оператора* $\mathcal{A}$ *вещественны;*
(3) *собственные векторы оператора* $\mathcal{A}$*, отвечающие различным собственным значения, ортогональны.*

Доказательство. Первое утверждение вытекает из того, что

$$\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}, \mathcal{A}\,\boldsymbol{x}\rangle = \overline{\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle},$$

где первое равенство следует из самосопряженности $\mathcal{A}$, а второе – из свойств скалярного произведения в эрмитовом пространстве.

Пусть теперь λ – собственное число оператора $\mathcal{A}$ и пусть $\boldsymbol{x}$ – соответствующий этому собственному числу собственный вектор. Тогда $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \lambda\|\boldsymbol{x}\|^2$ и, так как $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$ (в силу предыдущего утверждения), а $\|\boldsymbol{x}\| \in \mathbb{R}$ (по определению нормы), $\lambda \in \mathbb{R}$.

Остается проверить ортогональность собственных векторов самосопряженного оператора $\mathcal{A}$, отвечающих разным собственным значениям. Пусть $\lambda_1 \neq \lambda_2$ – два собственных числа оператора $\mathcal{A}$ и пусть $\boldsymbol{x}_1$ и $\boldsymbol{x}_2$ – соответствующие собственные векторы. В этом случае, $\langle \mathcal{A}\,\boldsymbol{x}_1, \boldsymbol{x}_2\rangle = \lambda_1\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle$ и $\langle \boldsymbol{x}_1, \mathcal{A}\,\boldsymbol{x}_2\rangle = \lambda_2\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle$ (в последнем равенстве учтено, что $\lambda_2 \in \mathbb{R}$). Из самосопряженности $\mathcal{A}$ следует, что $\langle \mathcal{A}\,\boldsymbol{x}_1, \boldsymbol{x}_2\rangle = \langle \boldsymbol{x}_1, \mathcal{A}\,\boldsymbol{x}_2\rangle$ и, следовательно, $\lambda_1\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle = \lambda_2\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle$, или $(\lambda_1 - \lambda_2)\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle = 0$. Из последнего равенства (так как $\lambda_1 \neq \lambda_2$) вытекает, что $\langle \boldsymbol{x}_1, \boldsymbol{x}_2\rangle = 0$. □

Матрица самосопряженного оператора. Пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – *самосопряженный* оператор, действующий в эрмитовом пространстве $\mathcal{X}$.

Предложение 8.6. *Самосопряженный оператор в эрмитовом пространстве в произвольном ортонормированном базисе имеет эрмитову матрицу.*

Если оператор $\mathcal{A}$*, действующий в эрмитовом пространстве* $\mathcal{X}$ *имеет в некотором ортонормированном базисе* $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ *пространства* $\mathcal{X}$ *эрмитову матрицу* A*, то он является самосопряженным.*

Доказательство. Напомним, что матрица $M \in \mathrm{M}_n(\mathbb{C})$ называется *эрмитовой*, если $M = M^* = \overline{M}^\top$. Проверим первое утверждение.

Выберем в $\mathcal{X}$ некоторый ортонормированный базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, где $n = \dim \mathcal{X}$ – размерность $\mathcal{X}$. Обозначим через $A = (a_{jk})_{j,k=1}^n$ матрицу оператора $\mathcal{A}$ в базисе $\boldsymbol{e}$. Тогда:

$$\langle \mathcal{A}\,\boldsymbol{e}_j, \boldsymbol{e}_k\rangle = \sum_{r=1}^n a_{rj}\langle \boldsymbol{e}_r, \boldsymbol{e}_k\rangle = a_{kj}, \qquad \langle \boldsymbol{e}_j, \mathcal{A}\,\boldsymbol{e}_k\rangle = \sum_{r=1}^n \overline{a}_{rk}\langle \boldsymbol{e}_j, \boldsymbol{e}_r\rangle = \overline{a}_{jk}$$

для всех $j, k = 1, \ldots, n$. Итак, для элементов матрицы самосопряженного оператора в ортонормированном базисе выполняется соотношение $a_{jk} = \overline{a}_{kj}$ при всех $j, k = 1, \ldots, n$. В матричной форме это соотношение имеет вид $A = A^*$, откуда A – эрмитова матрица.

Докажем второе утверждение. Пусть, как обычно, $\mathcal{A}^*$ – оператор, (эрмитово) сопряженный с $\mathcal{A}$. Предположим, что оператор $\mathcal{A}^*$ имеем матрицу $A' = (a'_{jk})_{j,k=1}^n$. Тогда:

$$\langle \mathcal{A}\,\boldsymbol{e}_j, \boldsymbol{e}_k\rangle = \sum_{r=1}^n a_{rj}\langle \boldsymbol{e}_r, \boldsymbol{e}_k\rangle = a_{kj}, \qquad \langle \boldsymbol{e}_j, \mathcal{A}^*\,\boldsymbol{e}_k\rangle = \sum_{r=1}^n \overline{a}'_{rk}\langle \boldsymbol{e}_j, \boldsymbol{e}_r\rangle = \overline{a}'_{jk}$$

для всех $j, k = 1, \ldots, n$. Учитывая определения сопряженного оператора и эрмитовой матрицы получаем, что $a'_{jk} = \overline{a}_{kj} = a_{jk}$ при всех $j, k = 1, \ldots, n$. Следовательно, $A = A*$. Итак, матрицы операторов $\mathcal{A}$ и $\mathcal{A}^*$ в базисе $\boldsymbol{e}$ совпадают, откуда $\mathcal{A} = \mathcal{A}^*$. □

Критерии эрмитовости полуторалинейных форм. В лекции 3 (см. раздел 3.5) было введено понятие полуторалинейной формы в эрмитовом пространстве и дано определение *эрмитовой формы*. Кроме того, в Теореме 7.17 было установлено, что всякая полуторалинейная форма $\mathcal{F}$, определенная в эрмитовом пространстве, может быть единственным образом представлена в виде

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle,$$

где $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – некоторый линейный оператор. Выясним два свойства эрмитовых форм, которые будут использоваться в дальнейшем.

Предложение 8.7. *Пусть в эрмитовом пространстве $\mathcal{X}$ задана полуторалинейная форма $\mathcal{F}$ и пусть $\mathcal{A}$ такой линейный оператор, что $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle$ при всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. Тогда форма $\mathcal{F}$ является эрмитовой если и только если выполняется любое из следующих двух условий:*

(1) *Линейный оператор $\mathcal{A}$ является самосопряженным.*
(2) *$\mathcal{F}(\boldsymbol{x}, \boldsymbol{x}) \in \mathbb{R}$ для любого $\boldsymbol{x} \in \mathcal{X}$.*

Доказательство. Начнем с проверки того факта, что условие (1) эквивалентно эрмитовости формы $\mathcal{F}$. Если оператор $\mathcal{A}$ самосопряженный, т.е., если $\mathcal{A} = \mathcal{A}^*$, то при всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ получаем

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathcal{A}\,\boldsymbol{y} \rangle = \overline{\langle \mathcal{A}\,\boldsymbol{y}, \boldsymbol{x} \rangle} = \overline{\mathcal{F}(\boldsymbol{y}, \boldsymbol{x})}.$$

Обратно, если форма $\mathcal{F}$ является эрмитовой, то при всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ имеют место равенства

$$\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle = \mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \overline{\mathcal{F}(\boldsymbol{y}, \boldsymbol{x})} = \overline{\langle \mathcal{A}\,\boldsymbol{y}, \boldsymbol{x} \rangle} = \langle \boldsymbol{x}, \mathcal{A}\,\boldsymbol{y} \rangle,$$

т.е., оператор $\mathcal{A}$ является самосопряженным.

Утверждение (2) также эквивалентно эрмитовости формы $\mathcal{F}$. В самом деле, эрмитовость формы $\mathcal{F}$ эквивалентна (по доказанному выше) самосопряженности оператора $\mathcal{A}$, которая, в свою очередь, эквивалентна тому, что $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x} \rangle \in \mathbb{R}$ для любого $\boldsymbol{x} \in \mathcal{X}$ (см. Предложение 8.9). Остается напомнить, что $\mathcal{F}(\boldsymbol{x}, \boldsymbol{x}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x} \rangle$. □

8.2. Самосопряженные операторы в эрмитовых пространствах с точки зрения общего понятия сопряженного оператора

Пусть теперь $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – некоторый линейный оператор, действующий в линейном пространстве $\mathcal{X}$.

Ранее было введено понятие сопряженного оператора $\mathcal{A}^*$. При этом, сопряженный оператор – это оператор, который определен не на исходном пространстве $\mathcal{X}$, а на двойственном ему пространстве $\mathcal{X}^*$.

В случае, когда пространство $\mathcal{X}$ наделено эрмитовой (или евклидовой) структурой, в разделе 8.1 был определен оператор $\mathcal{A}^*$ – эрмитово сопряженный к $\mathcal{A}$.

Пусть теперь $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – некоторый ортонормированный базис в $\mathcal{X}$, а $A = (a_{jk})$ – матрица оператора $\mathcal{A}$ в этом базисе. Пусть также $\mathcal{A}^* \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – оператор, эрмитово сопряженный с $\mathcal{A}$. При этом, если $A^* = (a^*_{jk})$ – матрица оператора $\mathcal{A}^*$ в базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, то (см. Предложение 8.6) $a^*_{jk} = \overline{a}_{kj}$ при всех $j, k = 1, \ldots, n$. В этом случае, для канонического спаривания $\langle \mathfrak{f} | \boldsymbol{x} \rangle$ между $\mathcal{X}$ и $\mathcal{X}^*$, определенного выше, получаем

$$\langle \mathfrak{f} |\, \mathcal{A}\,\boldsymbol{x} \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_k \overline{f}_j,$$

где $f_1, \ldots, f_n$ – координаты элемента $\mathfrak{f} \in \mathcal{X}^*$ относительно базиса $\widehat{\boldsymbol{e}} = \{\mathfrak{e}^1, \ldots, \mathfrak{e}^n\}$ (см. раздел 7.5). Аналогично, если $f^*_1, \ldots, f^*_n$ – координаты относительно этого же базиса элемента $\mathcal{A}^*\,\mathfrak{f}$, то $f^*_j = \sum_{k=1}^{n} a^*_{jk} f_k$ и

$$\langle \mathcal{A}^*\,\mathfrak{f} | \boldsymbol{x} \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} \overline{a^*_{jk}} x_j \overline{f}_k = \sum_{j=1}^{n} \sum_{k=1}^{n} \overline{a^*_{kj}} x_k \overline{f}_j = \sum_{j=1}^{n} \sum_{k=1}^{n} a_{jk} x_k \overline{f}_j = \langle \mathfrak{f} |\, \mathcal{A}\,\boldsymbol{x} \rangle.$$

Таким образом, исходя из определения эрмитово сопряженного оператора, мы получаем следующее равенство

$$\langle \mathcal{A}^* \mathfrak{f} | \boldsymbol{x} \rangle = \langle \mathfrak{f} | \mathcal{A}\, \boldsymbol{x} \rangle$$

для канонического спаривания между $\mathcal{X}^*$ и $\mathcal{X}$. Таким образом, эрмитово сопряженный оператор $\mathcal{A}^*$ является сопряженным к $\mathcal{A}$ и с точки зрения общего понятия сопряженного оператора.

8.3. Свойства самосопряженных операторов

Напомним, что в разделе 7.6 понятие нормы оператора было введено и изучено для линейных операторов, действующих в линейном нормированном пространстве.

Оказывается, что в случае самосопряженного оператора, действующего в эрмитовом пространстве $\mathcal{X}$, существует следующая удобная в формула для вычисления нормы.

Предложение 8.8. *Пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – самосопряженный оператор. Тогда*

$$\| \mathcal{A} \| = \sup\{ |\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle| : \ \boldsymbol{x} \in \mathcal{X}, \|\boldsymbol{x}\| = 1 \}.$$

Доказательство. Для произвольного элемента $\boldsymbol{x} \in \mathcal{X}$ справедливо следующее неравенство, которое непосредственно вытекает из неравенства Коши-Буняковского:

$$|\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle| \leqslant \| \mathcal{A}\, \boldsymbol{x} \| \|\boldsymbol{x}\|.$$

Далее, с учетом неравенства $\| \mathcal{A}\, \boldsymbol{x} \| \leqslant \| \mathcal{A} \| \|\boldsymbol{x}\|$ получаем, что

$$|\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle| \leqslant \| \mathcal{A} \| \|\boldsymbol{x}\|^2.$$

Из этого неравенства вытекает, что число $\mu = \sup\{ |\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle| : \ \boldsymbol{x} \in \mathcal{X}, \|\boldsymbol{x}\| = 1 \}$ удовлетворяет неравенству $\mu \leqslant \| \mathcal{A} \|$. Так как при $\boldsymbol{y} \in \mathcal{X}$ и $\boldsymbol{y} \neq \boldsymbol{0}$ имеет место равенство

$$\langle \mathcal{A}\, \boldsymbol{y}, \boldsymbol{y} \rangle = \|\boldsymbol{y}\|^2 \left\langle \mathcal{A}\left(\frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \right), \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \right\rangle,$$

то, учитывая определение числа μ получаем, что

$$|\langle \mathcal{A}\, \boldsymbol{y}, \boldsymbol{y} \rangle| \leqslant \mu \|\boldsymbol{y}\|^2.$$

Из свойств скалярного произведения легко выводится следующее соотношение (проверка которого оставляется в качестве *упражнения*):

$$4 \operatorname{Re}\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y} \rangle = \langle \mathcal{A}(\boldsymbol{x} + \boldsymbol{y}), \boldsymbol{x} + \boldsymbol{y} \rangle - \langle \mathcal{A}(\boldsymbol{x} - \boldsymbol{y}), \boldsymbol{x} - \boldsymbol{y} \rangle, \quad \boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}.$$

Из этого неравенства вытекает, что

$$4 |\operatorname{Re}\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y} \rangle| \leqslant \mu \|\boldsymbol{x} + \boldsymbol{y}\|^2 + \mu \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2\mu (\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2).$$

Применяя это неравенство при $\|\boldsymbol{x}\| = \|\boldsymbol{y}\| = 1$ получаем

$$|\operatorname{Re}\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{y} \rangle| \leqslant \mu.$$

Остается в этом неравенстве положить $\boldsymbol{y} = \mathcal{A}\, \boldsymbol{x} / \| \mathcal{A}\, \boldsymbol{x} \|$ (ясно, что $\|\boldsymbol{y}\| = 1$) и учитывая, что число $\langle \mathcal{A}\, \boldsymbol{x}, \mathcal{A}\, \boldsymbol{x} \rangle = \| \mathcal{A}\, \boldsymbol{x} \|^2$ является вещественным, получить неравенство $\| \mathcal{A}\, \boldsymbol{x} \| \leqslant \mu$ при $\|\boldsymbol{x}\| = 1$. По определению нормы линейного оператора из последнего неравенства вытекает, что $\| \mathcal{A} \| \leqslant \mu$. Так как ранее было показано, что $\mu \leqslant \| \mathcal{A} \|$, то окончательно получаем, что

$$\| \mathcal{A} \| = \mu = \sup\{ |\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle| : \ \boldsymbol{x} \in \mathcal{X}, \|\boldsymbol{x}\| = 1 \}. \quad \square$$

Критерий самосопряженности линейного оператора.

Предложение 8.9. *Оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ является самосопряженным тогда и только тогда, когда $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$ для любого $\boldsymbol{x} \in \mathcal{X}$.*

Доказательство. Как было показано при доказательстве Предложения 8.5(1), если оператор $\mathcal{A}$ самосопряженный, то $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$ для любого $\boldsymbol{x} \in \mathcal{X}$.

Докажем теперь обратное утверждение. На основании Предложения 8.3 оператор $\mathcal{A}$ может быть записан в виде $\mathcal{A} = \mathcal{A}_R + i\,\mathcal{A}_I$, где $\mathcal{A}_R$ и $\mathcal{A}_I$ – самосопряженные операторы. Отсюда, для произвольного $\boldsymbol{x} \in \mathcal{X}$, получаем, что $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \mathcal{A}_R\,\boldsymbol{x}, \boldsymbol{x}\rangle + i\langle \mathcal{A}_I\,\boldsymbol{x}, \boldsymbol{x}\rangle$ и, так как $\mathcal{A}_R$ и $\mathcal{A}_I$ – самосопряженные операторы, то $\langle \mathcal{A}_R\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$ и $\langle \mathcal{A}_I\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$ и, следовательно,

$$\mathrm{Re}\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \mathcal{A}_R\,\boldsymbol{x}, \boldsymbol{x}\rangle, \qquad \mathrm{Im}\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \mathcal{A}_I\,\boldsymbol{x}, \boldsymbol{x}\rangle.$$

Теперь, если $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \in \mathbb{R}$, то для любого $\boldsymbol{x} \in \mathcal{X}$ имеет место равенство $\langle \mathcal{A}_I\,\boldsymbol{x}, \boldsymbol{x}\rangle = 0$. В силу Предложения 8.8 из этого вытекает, что $\|\,\mathcal{A}_I\,\| = 0$ и, следовательно, $\mathcal{A}_I = \mathcal{O}$. Следовательно, $\mathcal{A} = \mathcal{A}_R$ и таким образом, $\mathcal{A}$ является самосопряженным. $\square$

Свойства собственных значений самосопряженных операторов. Начнем со следующего наблюдения: *Любое собственное значение λ самосопряженного оператора $\mathcal{A}$, действующего в эрмитовом пространстве $\mathcal{X}$ может быть представлено в виде $\lambda = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$ для некоторого элемента $\boldsymbol{x} \in \mathcal{X}$ с условием $\|\boldsymbol{x}\| = 1$.* В самом деле, если λ – собственное значение для $\mathcal{A}$, то существует собственный вектор $\boldsymbol{w}$, соответствующий этому собственному значению, причем, по определению, $\boldsymbol{w} \neq \boldsymbol{0}$. Итак, $\mathcal{A}\,\boldsymbol{w} = \lambda\boldsymbol{w}$. Пусть $\boldsymbol{x} = \boldsymbol{w}/\|\boldsymbol{w}\|$. При этом легко проверить, что $\mathcal{A}\,\boldsymbol{x} = \lambda\boldsymbol{x}$ и, следовательно, $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \lambda\|\boldsymbol{x}\|^2 = \lambda$.

Пусть теперь $\mathcal{A}$ – произвольный самосопряженный оператор и пусть

$$m := \inf\{\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle :\ \boldsymbol{x} \in \mathcal{X}, \|\boldsymbol{x}\| = 1\}, \qquad M := \sup\{\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle :\ \boldsymbol{x} \in \mathcal{X}, \|\boldsymbol{x}\| = 1\}.$$

Тогда для любого собственного значения λ оператора $\mathcal{A}$ справедливо неравенство $m \leqslant \lambda \leqslant M$ (эти неравенства имеют смысл так как собственные значения самосопряженного оператора вещественны).

Так как скалярное произведение $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$ представляет собой непрерывную функцию от $\boldsymbol{x}$, то на замкнутом множестве $\{\boldsymbol{x} \in \mathcal{X} :\ \|\boldsymbol{x}\| = 1\}$ эта функция ограничена и достигает своих точной верхней грани M и своей точной нижней грани m.

Неравенство $m \leqslant \lambda \leqslant M$ вытекает из того, что каждое собственное значение λ самосопряженного оператора $\mathcal{A}$ представляется в виде $\lambda = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$ для некоторого элемента $\boldsymbol{x} \in \mathcal{X}$ с условием $\|\boldsymbol{x}\| = 1$.

Теорема 8.10. *Пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – самосопряженный оператор. Тогда*

(1) *Если дополнительно $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \geqslant 0$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$, то число $\|\,\mathcal{A}\,\|$ является наибольшим собственным числом оператора $\mathcal{A}$.*

(2) *Числа m и M, определенные выше, являются наименьшим и наибольшим собственными значениями оператора $\mathcal{A}$.*

Утверждение (1) Теоремы 8.10 целесообразно сравнить с утверждением, полученным в Примере 7.24. Доказательство Теоремы 8.10 удобно объединить с доказательством следующего утверждения.

Теорема 8.11. *Пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – самосопряженный оператор в n-мерном эрмитовом пространстве $\mathcal{X}$. Тогда у оператора $\mathcal{A}$ существует n линейно независимых попарно ортогональных и единичных собственных векторов.*

Доказательство Теорем 8.10 и 8.11. Начнем с доказательства первого утверждения Теоремы 8.10. Из Предложения 8.8 следует, что $\|\,\mathcal{A}\,\| = \sup_{\|\boldsymbol{x}\|=1}|\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle|$. Так как $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \geqslant 0$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$, то $\|\,\mathcal{A}\,\| = \sup_{\|\boldsymbol{x}\|=1}\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$. Так как скалярное произведение $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$ представляет собой непрерывную функцию от

$\boldsymbol{x}$, то найдется элемент $\boldsymbol{x}_0 \in \mathcal{X}$ такой, что $\|\boldsymbol{x}_0\| = 1$ и $\langle \mathcal{A}\,\boldsymbol{x}_0, \boldsymbol{x}_0\rangle = \|\,\mathcal{A}\,\|$. Пусть теперь $\lambda_0 := \|\,\mathcal{A}\,\|$. Тогда $\|\,\mathcal{A}\,\| = \langle \mathcal{A}\,\boldsymbol{x}_0, \boldsymbol{x}_0\rangle = \|\,\mathcal{A}\,\boldsymbol{x}_0\|$. В самом деле, в силу выбора $\boldsymbol{x}_0$, $\|\,\mathcal{A}\,\| = \langle \mathcal{A}\,\boldsymbol{x}_0, \boldsymbol{x}_0\rangle \leqslant |\langle \mathcal{A}\,\boldsymbol{x}_0, \boldsymbol{x}_0\rangle| \leqslant \|\,\mathcal{A}\,\boldsymbol{x}_0\|\,\|\boldsymbol{x}_0\| = \|\,\mathcal{A}\,\boldsymbol{x}_0\|$ и $\|\,\mathcal{A}\,\boldsymbol{x}_0\| \leqslant \|\,\mathcal{A}\,\|\,\|\boldsymbol{x}_0\| = \|\,\mathcal{A}\,\|$. Отсюда вытекает, что

$$\|(\mathcal{A} - \lambda_0\,\mathcal{E})\boldsymbol{x}_0\|^2 = \|\,\mathcal{A}\,\boldsymbol{x}_0\|^2 - 2\lambda_0\langle \mathcal{A}\,\boldsymbol{x}_0, \boldsymbol{x}_0\rangle + \lambda_0^2\|\boldsymbol{x}_0\|^2 = \|\,\mathcal{A}\,\|^2 - 2\|\,\mathcal{A}\,\|\|\,\mathcal{A}\,\| + \|\,\mathcal{A}\,\|^2 = 0.$$

Таким образом, $(\mathcal{A} - \lambda_0\,\mathcal{E})\boldsymbol{x}_0 = \mathbf{0}$ или $\mathcal{A}\,\boldsymbol{x}_0 = \lambda_0\boldsymbol{x}_0$ и, окончательно, $\lambda_0 = \|\,\mathcal{A}\,\|$ является собственным значением оператора $\mathcal{A}$. То, что λ_0 – наибольшее собственное значение вытекает из утверждений, приведенных в начале раздела.

Докажем теперь второе утверждение Теоремы 8.10. Для этого нам достаточно проверить, что числа m и M являются собственными значениями оператора $\mathcal{A}$. Первым делом проверим, что собственным значением является число M. Рассмотрим самосопряженный оператор $\mathcal{B} = \mathcal{A} - m\,\mathcal{E}$. Заметим, что

$$\langle \mathcal{B}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle - m\langle \boldsymbol{x}, \boldsymbol{x}\rangle \geqslant 0$$

для любого $\boldsymbol{x} \in \mathcal{X}$. Таким образом, оператор $\mathcal{B}$ удовлетворяет условиям первого утверждения теоремы. Следовательно, число $\|\,\mathcal{B}\,\|$ равно максимальному собственному значению оператора $\mathcal{B}$. Отсюда

$$\|\,\mathcal{B}\,\| = \sup_{\|\boldsymbol{x}\|=1} \langle \mathcal{B}\,\boldsymbol{x}, \boldsymbol{x}\rangle = \sup_{\|\boldsymbol{x}\|=1} \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle - m = M - m.$$

Итак, число $M - m$ – собственное значение оператора $\mathcal{B}$. Это означает, что существует ненулевой вектор $\boldsymbol{x}_0$ такой, что $\mathcal{B}\,\boldsymbol{x}_0 = (M - m)\boldsymbol{x}_0$. Так как $\mathcal{B} = \mathcal{A} - m\,\mathcal{E}$, то $\mathcal{B}\,\boldsymbol{x}_0 = \mathcal{A}\,\boldsymbol{x}_0 - m\boldsymbol{x}_0$. Следовательно $\mathcal{A}\,\boldsymbol{x}_0 - m\boldsymbol{x}_0 = M\boldsymbol{x}_0 - m\boldsymbol{x}_0$ и, окончательно, $\mathcal{A}\,\boldsymbol{x}_0 = M\boldsymbol{x}_0$ и M является собственным значением оператора $\mathcal{A}$.

Для того, чтобы проверить, что число m является собственным значением оператора $\mathcal{A}$ рассмотрим оператор $\mathcal{A}_1 = -\,\mathcal{A}$. Число $-m$ – это верхняя грань величины $\langle \mathcal{A}_1\,\boldsymbol{x}, \boldsymbol{x}\rangle$ при $\|\boldsymbol{x}\| = 1$. По доказанному выше, число $-m$ – собственное значение оператора $\mathcal{A}_1$ и, следовательно, число m – собственное значение оператора $\mathcal{A}$.

Нам остается доказать Теорему 8.11. Пусть λ_1 – максимальное собственное значение оператора $\mathcal{A}$ и пусть $\boldsymbol{e}_1$ – соответствующий этому собственному значению собственный вектор, нормированный условием $\|\boldsymbol{e}_1\| = 1$ (в начале данного раздела были приведены аргументы, подтверждающие возможность такого выбора вектора $\boldsymbol{e}_1$.

Определим подпространство $\mathcal{X}_1 \subset \mathcal{X}$ состоящее из всех элементов $\boldsymbol{x} \in \mathcal{X}$, ортогональных $\boldsymbol{e}_1$. Так как $\boldsymbol{e}_1$ – собственный вектор, то $\mathcal{X}_1$ – $\mathcal{A}$-инвариантное подпространство (проверка этого простого факта оставляется в качестве *упражнения*). Ограничение оператора $\mathcal{A}$ на подпространство $\mathcal{X}_1$ – это самосопряженный оператор, действующий на $\mathcal{X}_1$. У этого оператора также имеется максимальное собственное значение λ_2 и соответствующий собственный вектор $\boldsymbol{e}_2 \in \mathcal{X}_1$ такой, что $\|\boldsymbol{e}_2\| = 1$. Ясно, что $\boldsymbol{e}_2 \perp \boldsymbol{e}_1$ (значок $\perp$ обозначает ортогональность). Пусть теперь подпространство $\mathcal{X}_2$ состоит из всех элементов $\boldsymbol{x} \in \mathcal{X}$ таких, что $\boldsymbol{x} \perp \boldsymbol{e}_1$ и $\boldsymbol{x} \perp \boldsymbol{e}_2$. В этом подпространстве мы найдем следующий собственный вектор $\boldsymbol{e}_3$. Повторив описанную процедуру еще $n-3$ раза мы построим требуемую системы взаимно ортогональный собственных векторов. □

Замечание. В дальнейшем собственные значения $\lambda_1, \ldots, \lambda_n$ (рассматриваемые с учетом их кратности) самосопряженного оператора, действующего в n-мерном евклидовом пространстве, будем нумеровать в порядке убывания их значений: $\lambda_1 \geqslant \lambda_2 \geqslant \cdots \geqslant \lambda_n$. В этом случае, соответствующие собственные векторы $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ можно выбрать так, что будут иметь место соотношения $\langle \boldsymbol{e}_j, \boldsymbol{e}_k\rangle = \delta_{jk}$, $j, k = 1, \ldots, n$.

При доказательстве Теоремы 8.11 было неявным образом установлено следующее утверждение: если $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ – собственные значения самосопряженного оператора $\mathcal{A}$, то при $k = 2, \ldots, n$ справедливо соотношение

$$\lambda_k = \max\left\{\frac{\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle}{\langle \boldsymbol{x}, \boldsymbol{x}\rangle} :\ \boldsymbol{x} \perp \operatorname{Span}\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{k-1}\}\right\}.$$

На самом деле можно утверждать больше. А именно имеет место следующее важное свойство собственных значений самосопряженных операторов. При натуральном $k \leqslant n$ обозначим через $\mathfrak{S}_k(\mathcal{X})$ совокупность всех k-мерных подпространств пространства $\mathcal{X}$.

Теорема 8.12. *Пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – самосопряженный оператор и пусть $\lambda_1 \geqslant \dots \geqslant \lambda_n$ – его собственные значения. Тогда, при $k = 2, \dots, n$ имеет место равенство*

$$\lambda_k = \min_{\mathcal{V} \in \mathfrak{S}_{k-1}(\mathcal{X})} \max_{\boldsymbol{x} \perp \mathcal{V}} \frac{\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}.$$

Доказательство. Как было установлено в Замечании, приведенном после доказательства Теоремы 8.11,

$$\lambda_k = \max_{\boldsymbol{x} \perp \mathcal{V}_{k-1}} \frac{\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle},$$

где $\mathcal{V}_{k-1} = \operatorname{Span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_{k-1}\}$. Соответственно, для доказательства теоремы нам необходимо проверить, что для любого $\mathcal{V} \in \mathfrak{S}_{k-1}(\mathcal{X})$ имеет место неравенство

$$\max_{\boldsymbol{x} \perp \mathcal{V}} \frac{\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} \geqslant \max_{\boldsymbol{x} \perp \mathcal{V}_{k-1}} \frac{\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \lambda_k \tag{8.3}$$

(другим словами, надо показать, что требуемый в условии теоремы минимум достигается именно на подпространстве $\mathcal{V}_{k-1}$). Будем доказывать соотношение (8.3). Пусть $\mathcal{V}^\perp$ – ортогональное дополнение подпространства $\mathcal{V}$. При этом $\dim \mathcal{V}^\perp = n - k + 1$. Далее, $\dim \mathcal{V}^\perp + \dim \mathcal{V}_k = n - k + 1 + k = n + 1 > n$ и, следовательно, в подпространстве $\mathcal{V}^\perp \cap \mathcal{V}_k$ содержится ненулевой элемент. Следовательно, существует элемент $\widetilde{\boldsymbol{x}} \in \mathcal{V}_k$ такой, что $\widetilde{\boldsymbol{x}} \perp \mathcal{V}$ и $\|\widetilde{\boldsymbol{x}}\| = 1$. Так как $\widetilde{\boldsymbol{x}} \in \mathcal{V}_k$, то $\widetilde{\boldsymbol{x}} = \sum_{j=1}^k c_j \boldsymbol{e}_j$. Так как $\|\widetilde{\boldsymbol{x}}\| = 1$, а базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$ – ортонормированный, то

$$\sum_{j=1}^k |c_j|^2 = \|\widetilde{\boldsymbol{x}}\|^2 = 1 \tag{8.4}$$

(проверка первого равенства оставляется в качестве *упражнения*). Далее, так как $\boldsymbol{e}_k$, $k = 1, \dots, n$ – собственные векторы для $\mathcal{A}$, соответствующие собственным значениям $\lambda_1, \dots, \lambda_n$, то

$$\mathcal{A}\, \widetilde{\boldsymbol{x}} = \mathcal{A}\left(\sum_{j=1}^k c_j \boldsymbol{e}_j \right) = \sum_{j=1}^k c_j\, \mathcal{A}\, \boldsymbol{e}_j = \sum_{j=1}^k c_j \lambda_j \boldsymbol{e}_j.$$

Отсюда, учитывая взаимную ортогональность собственных векторов $\boldsymbol{e}_k$, $k = 1, \dots, n$, неравенства $\lambda_1 \geqslant \dots \geqslant \lambda_{k-1} \geqslant \lambda_k$ и соотношения (8.4) получаем

$$\langle \mathcal{A}\, \widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{x}} \rangle = \left\langle \sum_{j=1}^k c_j \lambda_j \boldsymbol{e}_j, \sum_{r=1}^k c_r \boldsymbol{e}_r \right\rangle = \sum_{j=1}^k \lambda_j |c_j|^2 \geqslant \lambda_k \sum_{j=1}^k |c_j|^2 = \lambda_k.$$

Теперь, для любого ненулевого $\boldsymbol{x} \in \mathcal{V}$ имеем

$$\max_{\boldsymbol{x} \perp \mathcal{V}} \frac{\langle \mathcal{A}\, \boldsymbol{x}, \boldsymbol{x} \rangle}{\langle \boldsymbol{x}, \boldsymbol{x} \rangle} = \max_{\boldsymbol{x} \perp \mathcal{V}} \left\langle \mathcal{A}\left(\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \right), \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \right\rangle \geqslant \langle \mathcal{A}\, \widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{x}} \rangle \geqslant \lambda_k,$$

откуда вытекает соотношение (8.3). □

Спектральное разложение самосопряженных операторов. Пусть $\mathcal{A}$ – самосопряженный оператор, действующий в евклидовом пространстве $\mathcal{X}$ и имеющий собственные значения $\lambda_1 \geqslant \dots \geqslant \lambda_n$. Пусть $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ – ортонормированный базис в $\mathcal{X}$, состоящий из собственных векторов оператора $\mathcal{A}$, соответствующих собственным значениям $\lambda_1, \dots, \lambda_n$. Пусть $\boldsymbol{x} \in \mathcal{X}$. Тогда, как было показано выше,

$$\boldsymbol{x} = \sum_{k=1}^n \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \boldsymbol{e}_k$$

и, так как $\mathcal{A}\,\boldsymbol{e}_k = \lambda_k \boldsymbol{e}_k$ при $k = 1, \ldots, n$, то

$$\mathcal{A}\,\boldsymbol{x} = \sum_{k=1}^{n} \lambda_k \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \boldsymbol{e}_k.$$

Определение. *Оператор $\mathcal{P}_k$, $k = 1, \ldots, n$, определенный соотношением*

$$\mathcal{P}_k\,\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \boldsymbol{e}_k,$$

назовем проектором на одномерное подпространство, порожденное вектором $\boldsymbol{e}_k$.

Упражнение. Проверить, что оператор $\mathcal{P}_k$, $k = 1, \ldots, n$, является *самосопряженным* линейным оператором.

Пусть $k, j = 1, \ldots, n$. Тогда

$$(\mathcal{P}_j\,\mathcal{P}_k)\boldsymbol{x} = \mathcal{P}_j(\mathcal{P}_k\,\boldsymbol{x}) = \mathcal{P}_j(\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \boldsymbol{e}_k) = \langle \boldsymbol{x}, \boldsymbol{e}_k \rangle \langle \boldsymbol{e}_k, \boldsymbol{e}_j \rangle \boldsymbol{e}_j = \begin{cases} \langle \boldsymbol{x}, \boldsymbol{e}_j \rangle \boldsymbol{e}_j, & \text{при } j = k \\ \boldsymbol{0}, & \text{при } j \neq k \end{cases}$$

Таким образом, проекторы обладают следующими свойствами

(1) $\mathcal{P}_k^2 = \mathcal{P}_k$ и, следовательно, $\mathcal{P}_k^\ell = \mathcal{P}$ для любого $\ell \in \mathbb{N}$;
(2) $\mathcal{P}_j\,\mathcal{P}_k = \mathcal{O}$ при $k \neq j$.
(3) Каждый проектор $\mathcal{P}_k$, $k = 1, \ldots, n$, коммутирует с любым оператором, с которым коммутирует $\mathcal{A}$.

Далее, для элементов $\boldsymbol{x} \in \mathcal{X}$ и $\mathcal{A}\,\boldsymbol{x}$ по определению проектора получаем

$$\boldsymbol{x} = \sum_{k=1}^{n} \mathcal{P}_k\,\boldsymbol{x} \qquad \text{и} \qquad \mathcal{A}\,\boldsymbol{x} = \sum_{k=1}^{n} \lambda_k\,\mathcal{P}_k\,\boldsymbol{x}. \tag{8.5}$$

Из этих равенств, во-первых, следует, что

$$\sum_{k=1}^{n} \mathcal{P}_k = \mathcal{E}\,.$$

Во-вторых, из равенств (8.5) мы получаем так называемое *спектрально разложение самосопряженного оператора*

$$\mathcal{A} = \sum_{k=1}^{n} \lambda_k\,\mathcal{P}_k\,.$$

Заметим, что из свойств (1) и (2) проекторов вытекает, что для любого натурального ℓ имеет место

$$\mathcal{A}^\ell = \sum_{k=1}^{n} \lambda_k^\ell\,\mathcal{P}_k\,.$$

Рассмотрим теперь произвольный полином $P(x) = a_r x^r + \cdots + a_1 x + a_0$ и, по определению положим $P(\mathcal{A}) = a_r\,\mathcal{A}^r + \cdots + a_1\,\mathcal{A} + a_0\,\mathcal{E}$. Используя спектральное разложение для оператора $\mathcal{A}$ и приводя подобные слагаемые получаем, что $P(\mathcal{A}) = \sum_{k=1}^{n} P(\lambda_k)\,\mathcal{P}_k$.

Используя спектральное разложение самосопряженного оператора и утверждение Теоремы 5.2 о том, что собственные числа линейного оператора являются корнями его характеристического многочлена, можно получить тривиальное доказательство теоремы Гамильтона-Кэли (см. Теорему 5.6). В самом деле, пусть $P_{\mathcal{A}}$ – характеристический многочлен самосопряженного оператора $\mathcal{A}$. Тогда (так как $P_{\mathcal{A}}(\lambda_k) = 0$ для любого собственного значения λ_k оператора $\mathcal{A}$)

$$P_{\mathcal{A}}(\mathcal{A}) = \sum_{k=1}^{n} P_{\mathcal{A}}(\lambda_k)\,\mathcal{P}_k = \mathcal{O}\,.$$

Положительные операторы. Корни степени ℓ из оператора.

Определение. *Самосопряженный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется положительным, если для любого $\boldsymbol{x} \in \mathcal{X}$ выполняется неравенство $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle \geqslant 0$. Если оператор $\mathcal{A}$ положителен и из условия $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle = 0$ вытекает, что $\boldsymbol{x} = 0$, то оператор $\mathcal{A}$ называется положительно определенным. Положительность и положительная определенность оператора $\mathcal{A}$ обозначаются $\mathcal{A} \geqslant 0$ и $\mathcal{A} > 0$ соответственно.*

Так как из самосопряженности оператора $\mathcal{A}$ вытекает, что любое собственное значение λ оператора $\mathcal{A}$ можно представить в виде $\lambda = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$ для некоторого элемента $\boldsymbol{x} \in \mathcal{X}$, $\|\boldsymbol{x}\| = 1$, то *каждое собственное значение положительного (соответственно, положительно определенного) оператора неотрицательно (соответственно, положительно).*

Определение. *Пусть ℓ – натуральное число. Корнем степени ℓ из оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется оператор $\mathcal{B} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ такой, что $\mathcal{B}^\ell = \mathcal{A}$. При этом корень степени ℓ из оператора $\mathcal{A}$ обозначается $\mathcal{A}^{1/\ell}$ или $\sqrt[\ell]{\mathcal{A}}$. При $\ell = 2$ традиционно пишут $\sqrt{\mathcal{A}}$ вместо $\sqrt[2]{\mathcal{A}}$.*

Теорема 8.13. *Для любого положительного самосопряженного оператора $\mathcal{A}$ и для любого натурального числа ℓ существует положительный корень $\mathcal{A}^{1/\ell}$ степени ℓ из оператора $\mathcal{A}$.*

Доказательство. Пусть $\lambda_1 \geqslant \cdots \geqslant \lambda_n$ – собственные значения оператора $\mathcal{A}$ и пусть $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – ортонормированный базис в $\mathcal{X}$, состоящий из соответствующих этим собственным значениям собственных векторов оператора $\mathcal{A}$. Пусть, далее, $\mathcal{P}_k$ – проекторы на одномерные подпространства, порожденные векторами $\boldsymbol{e}_k$, $k = 1, \ldots, n$. Таким образом имеет место спектральное разложение оператора $\mathcal{A}$:

$$\mathcal{A} = \sum_{k=1}^{n} \lambda_k\, \mathcal{P}_k\,.$$

Так как $\lambda_k \geqslant 0$ при $k = 1, \ldots, n$ (в силу положительности оператора $\mathcal{A}$), то можно определить линейный оператор

$$\mathcal{B} := \sum_{k=1}^{n} \sqrt[\ell]{\lambda_k}\, \mathcal{P}_k,$$

положительность которого немедленно вытекает из положительности проекторов $\mathcal{P}_k$, $k = 1, \ldots, n$ (проверка оставляется в качестве *упражнения*). Из свойств (1) и (2) проекторов (см. выше) вытекает, что $\mathcal{B}^\ell = \mathcal{A}$ и, следовательно, (положительный) оператор $\mathcal{B}$ и является корнем степени ℓ из оператора $\mathcal{A}$. $\square$

Упражнение. Доказать, что для положительного самосопряженного оператора $\mathcal{A}$ существует единственный положительный оператор $\mathcal{A}^{1/\ell}$. Проверить, что в ортонормированном базисе $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ составленном из собственных векторов оператора $\mathcal{A}$ матрица оператора $\mathcal{A}^{1/\ell}$ имеет вид $\mathrm{Diag}\{\sqrt[\ell]{\lambda_1}, \ldots, \sqrt[\ell]{\lambda_n}\}$.

8.4. Самосопряженные операторы и билинейные формы – случай евклидова пространства

В этом разделе мы кратко рассмотрим самосопряженные операторы и билинейные формы, определенные в (вещественном) евклидовом пространстве.

Пусть $\mathcal{X}$ – вещественное евклидово пространство размерности $n \in \mathbb{N}$ со скалярным произведением $\langle \cdot, \cdot \rangle$. Аналогично случаю (комплексного) эрмитова пространства для оператора $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ вводится *сопряженный оператор* $\mathcal{A}^*$: это такой оператор, что

$$\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{y}\rangle$$

при любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. Как и в эрмитовом случае доказывается, что для любого линейного оператора существует единственный сопряженный оператор (что, как и в эрмитовом случае, оправдывает обозначение $\mathcal{A}^*$).

Следующее утверждение доказывается абсолютно аналогично Предложению 8.6 (проверка необходимых деталей оставляется в качестве *упражнения*):

Предложение 8.14. *Самосопряженный оператор в евклидовом пространстве в произвольном ортонормированном базисе имеет симметричную матрицу.*

Если оператор $\mathcal{A}$, действующий в евклидовом пространстве $\mathcal{X}$ имеет в некотором ортонормированном базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$ симметричную матрицу A, то он является самосопряженным.

Далее, имеет место важное свойство характеристических корней самосопряженного оператора, действующего в (вещественном) евклидовом пространстве:

Предложение 8.15. *Все корни характеристического многочлена самосопряженного оператора $\mathcal{A}$, действующего в вещественном евклидовом пространстве $\mathcal{X}$ являются вещественными.*

Доказательство. Выберем базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в $\mathcal{X}$ и пусть $A = (a_{jk})$ – матрица оператора $\mathcal{A}$ в этом базисе ($a_{jk} \in \mathbb{R}$ при $j, k = 1, \ldots, n$). Пусть $\lambda = \alpha + i\beta$ – корень характеристического уравнения $\det(\mathcal{A} - \lambda\,\mathcal{E}) = 0$. Рассмотрим систему линейных однородных уравнений

$$\sum_{k=1}^{n} a_{jk} z_k = \lambda z_k, \quad j = 1, \ldots, n,$$

относительно неизвестных $z_1, \ldots, z_n$, где $\lambda = \alpha + i\beta$. Так как определитель этой системы равен нулю, то существует ненулевое решение $z_k = x_k + iy_k$, $k = 1, \ldots, n$. Отделяя вещественные и мнимые части соответствующих выражений получаем

$$\sum_{k=1}^{n} a_{jk} x_k = \alpha x_k - \beta y_k, \qquad \sum_{k=1}^{n} a_{jk} y_k = \alpha y_k + \beta x_k, \qquad j = 1, \ldots, n.$$

Рассмотрим векторы $\boldsymbol{x} = \sum_{k=1}^{n} x_k \boldsymbol{e}_k$ и $\boldsymbol{y} = \sum_{k=1}^{n} y_k \boldsymbol{e}_k$. Для них верны соотношения

$$\mathcal{A}\,\boldsymbol{x} = \alpha\boldsymbol{x} - \beta\boldsymbol{y}, \qquad \mathcal{A}\,\boldsymbol{y} = \alpha\boldsymbol{y} + \beta\boldsymbol{x}.$$

Отсюда следует, что

$$\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y} \rangle - \beta\langle \boldsymbol{y}, \boldsymbol{y} \rangle, \qquad \langle \boldsymbol{x}, \mathcal{A}\,\boldsymbol{y} \rangle = \alpha\langle \boldsymbol{x}, \boldsymbol{y} \rangle + \beta\langle \boldsymbol{x}, \boldsymbol{x} \rangle. \tag{8.6}$$

Если $\mathcal{A}$ – самосопряженный, то $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle = \langle \boldsymbol{x}, \mathcal{A}\,\boldsymbol{y} \rangle$ и, следовательно, вычитая первое равенство в (8.6) из второго получаем, что $\beta(\langle \boldsymbol{x}, \boldsymbol{x} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle) = 0$. Так как мы выбрали ненулевое решение $z_k = x_k + iy + k$ исходной линейной системы, то величина в скобках не равна нулю. Следовательно, $\beta = 0$, но $\beta = \operatorname{Im}\lambda$. □

Доказательство следующего утверждения ничем не отличается от доказательства соответствующего утверждения в эрмитовом случае и оставляется в качестве *упражнения*:

Предложение 8.16. *Пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – самосопряженный оператор в (вещественном) евклидовом пространстве $\mathcal{X}$. Тогда в $\mathcal{X}$ существует ортонормированный базис, состоящий из собственных векторов оператора $\mathcal{A}$.*

Аналогично Теореме 7.17 доказывается утверждение о том, что всякая билинейная форма $\mathcal{F}$ на евклидовом пространстве $\mathcal{X}$ может быть представлена в виде $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle$, где $\mathcal{A}$ – некоторый, действующий в $\mathcal{X}$, линейный оператор. Это утверждение предлагается доказать в качестве *упражнения*. Следующее свойство билинейных форм является аналогом Предложения 8.7

Предложение 8.17. *Пусть $\mathcal{F}$ – билинейная форма, определенная на вещественном евклидовом пространстве $\mathcal{X}$ и пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ такой линейный оператор, что $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y}\rangle$ при всех $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. Тогда форма $\mathcal{F}$ является симметричной если и только если оператор $\mathcal{A}$ самосопряженный.*

В самом деле, из самосопряженности оператора $\mathcal{A}$ вытекает симметричность формы $\mathcal{F}$ так как $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \mathcal{A}\,\boldsymbol{y}\rangle = \langle \mathcal{A}\,\boldsymbol{y}, \boldsymbol{x}\rangle = \mathcal{F}(\boldsymbol{y}, \boldsymbol{x})$ для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$. Проверка обратного утверждения оставляется в качестве *упражнения*.

8.5. Нормальные операторы

Пусть $\mathcal{X}$ – конечномерное эрмитово пространство, $\dim \mathcal{X} = n$, с эрмитовым скалярным произведением $\langle \cdot, \cdot \rangle$. Пусть, как обычно, $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$ при $\boldsymbol{x} \in \mathcal{X}$.

Определение. *Линейный оператор $\mathcal{N} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ называется нормальным, если справедливо соотношение $\mathcal{N}^*\,\mathcal{N} = \mathcal{N}\,\mathcal{N}^*$.*

Из определения сопряженного оператора непосредственно вытекает, что $(\lambda\,\mathcal{E})^* = \overline{\lambda}\,\mathcal{E}$ для любого $\lambda \in \mathbb{C}$. Следовательно, для любого оператора $\mathcal{A}$ и для любого $\lambda \in \mathbb{C}$ имеет место равенство $(\mathcal{A} - \lambda\,\mathcal{E})^* = \mathcal{A}^* - \overline{\lambda}\,\mathcal{E}$ и операторы $\mathcal{A}$ и $\mathcal{A} - \lambda\,\mathcal{E}$ являются и не являются нормальными одновременно.

Далее, если $\mathcal{N} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ – *нормальный* оператор, а $\boldsymbol{x} \in \mathcal{X}$ – произвольный вектор, то

$$\|\,\mathcal{N}\boldsymbol{x}\|^2 = \langle \mathcal{N}\boldsymbol{x}, \mathcal{N}\boldsymbol{x}\rangle = \langle \boldsymbol{x}, \mathcal{N}^*\,\mathcal{N}\boldsymbol{x}\rangle = \langle \boldsymbol{x}, \mathcal{N}\mathcal{N}^*\,\boldsymbol{x}\rangle = \langle \mathcal{N}^*\,\boldsymbol{x}, \mathcal{N}^*\,\boldsymbol{x}\rangle = \|\,\mathcal{N}^*\,\boldsymbol{x}\|^2. \tag{8.7}$$

Пусть λ – произвольное комплексное число. Из равенства (8.7), примененного к нормальному оператору $\mathcal{N} - \lambda\,\mathcal{E}$ вытекает, что $\|\,\mathcal{N}\boldsymbol{x} - \lambda\boldsymbol{x}\| = \|\,\mathcal{N}^*\,\boldsymbol{x} - \overline{\lambda}\boldsymbol{x}\|$ для любого вектора $\boldsymbol{x} \in \mathcal{X}$ а из этого равенства уже вытекает, что $\mathcal{N}\boldsymbol{x} = \lambda\boldsymbol{x}$ если и только если $\mathcal{N}^*\,\boldsymbol{x} = \overline{\lambda}\boldsymbol{x}$. Из этого наблюдения вытекает следующее утверждение.

Предложение 8.18. *Для любого собственного значения λ нормального оператора $\mathcal{N} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ существует единичный собственный вектор, отвечающий этому собственному значению, который в то же время является собственным для оператора $\mathcal{N}^*$.*

Замечание. Ранее нами были изучены самосопряженные операторы, которые определялись соотношением $\mathcal{A} = \mathcal{A}^*$. Следовательно, всякий самосопряженный оператор является нормальным. Нормальным также будет оператор, определяемый условием $\mathcal{A}^* = \mathcal{A}^{-1}$. Ясно также, что существуют нормальные операторов, которые не являются ни самосопряженными ни унитарными. Например, оператор с матрицей $\mathrm{Diag}\{1, 1, 2, 3i\}$.

Интерес к изучению нормальных операторов обусловлен, в частности, следующими двумя причинами. Во-первых, понятие нормального оператора естественным образом вводится в в случае бесконечномерных пространства, где такие операторы возникают во многих задачах. Во-вторых, нормальные операторы естественно возникают в связи с так называемой *проблемой диагонализируемости* линейных операторов.

Обсудим эту проблему более подробно.

Определение. *Скажем, что оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$, действующий в эрмитовом пространстве $\mathcal{X}$ является диагонализируемым, если в $\mathcal{X}$ существует такой ортонормированный базис, матрица оператора $\mathcal{A}$ в котором имеет диагональный вид.*

Из Теоремы 8.11 вытекает, что если оператор $\mathcal{A}$ самосопряженный, то существует ортонормированный базис в пространстве $\mathcal{X}$, состоящий из собственных векторов оператора $\mathcal{A}$. Отсюда следует, что всякий самосопряженный оператор является диагонализируемым.

Оказывается, что введенное свойство нормальности оказывается необходимым и достаточным условием диагонализируемости линейного оператора. Имеет место следующий результат.

Теорема 8.19. *Пусть $\mathcal{X}$ – n-мерное эрмитово пространство и пусть $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. Тогда $\mathcal{A}$ является диагонализируемым если и только если $\mathcal{A}$ является нормальным.*

Доказательство. Пусть $\mathcal{A}$ является диагонализируемым. Тогда существует базис ортонормированный базис $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ такой, что $\mathcal{A}\,\boldsymbol{e}_j = \lambda_j \boldsymbol{e}_j$, где $j = 1, \dots, n$, а λ_j – некоторые комплексные числа. Тогда $\mathcal{A}^*\,\boldsymbol{e}_j = \overline{\lambda}_j \boldsymbol{e}_j$. Следовательно, $\mathcal{A}\,\mathcal{A}^*\,\boldsymbol{e}_j - \mathcal{A}^*\,\mathcal{A}\,\boldsymbol{e}_j = \lambda_j \overline{\lambda}_j \boldsymbol{e}_j - \overline{\lambda}_j \lambda_j \boldsymbol{e}_j = \boldsymbol{0}$ и, следовательно, $\mathcal{A}\,\mathcal{A}^* = \mathcal{A}^*\,\mathcal{A}$.

Обратно, пусть $\mathcal{A}$ – нормальный оператор. Установим, что в пространстве $\mathcal{X}$ существует ортонормированный базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$, состоящий из собственных векторов операторов $\mathcal{N}$ и $\mathcal{N}^*$.

Из Предложения 8.18 вытекает, что операторы $\mathcal{N}$ и $\mathcal{N}^*$ имеют в $\mathcal{X}$ общий собственный вектор $\boldsymbol{e}_1$, имеющий единичную норму и соответствующий собственным значениям λ и $\overline{\lambda}$ операторов $\mathcal{N}$ и $\mathcal{N}^*$ соответственно.

Определим множество (подпространство) $\mathcal{X}_1 = \{\boldsymbol{x} \in \mathcal{X} : \ \langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle = 0\}$. Проверим, что $\mathcal{X}_1$ является одновременно $\mathcal{N}$- и $\mathcal{N}^*$-инвариантным подпространством. В самом деле, если элемент $\boldsymbol{x} \in \mathcal{X}$ ортогонален элементу $\boldsymbol{e}_1$, то

$$\langle \mathcal{N}\,\boldsymbol{x}, \boldsymbol{e}_1 \rangle = \langle \boldsymbol{x}, \mathcal{N}^*\,\boldsymbol{e}_1 \rangle = \langle \boldsymbol{x}, \overline{\lambda} \boldsymbol{e}_1 \rangle = \lambda \langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle = 0$$

и, аналогично,

$$\langle \mathcal{N}^*\,\boldsymbol{x}, \boldsymbol{e}_1 \rangle = \langle \boldsymbol{x}, \mathcal{N}\,\boldsymbol{e}_1 \rangle = \langle \boldsymbol{x}, \lambda \boldsymbol{e}_1 \rangle = \overline{\lambda} \langle \boldsymbol{x}, \boldsymbol{e}_1 \rangle = 0.$$

Применим утверждение Предложения 8.18 к подпространству $\mathcal{X}_1$ и найдем в нем общий единичный собственный вектор $\boldsymbol{e}_2$ для операторов $\mathcal{N}_{|\mathcal{X}_1}$ и $\mathcal{N}^*_{|\mathcal{X}_1}$. Затем мы определим подпространство $\mathcal{X}_2 = \{\boldsymbol{x} \in \mathcal{X}_1 : \ \langle \boldsymbol{x}, \boldsymbol{e}_2 \rangle = 0\}$ и повторим для этого подпространства описанную выше процедуру. В результате придем к вектору $\boldsymbol{e}_3$ и подпространству $\mathcal{X}_3$. Повторяя этот процесс необходимое (конечное, так как пространство $\mathcal{X}$ конечномерно) число раз мы получим ортонормированный базис в $\mathcal{X}$ с требуемыми свойствами. $\square$

Заметим, что самосопряженный оператор – это нормальный оператор, у которого все собственные значения действительны. Соответственно, в случае операторов, действующих в вещественных пространствах, нормальный оператор является самосопряженным.

Продолжим изучение свойств нормальных операторов. Пусть линейное пространство $\mathcal{X}$ разложено в прямую сумму $\mathcal{X} = \mathcal{Y} \oplus \mathcal{Z}$ подпространств и пусть $\boldsymbol{x} \in \mathcal{X}$. Тогда существует единственное разложение $\boldsymbol{x} = \boldsymbol{y} + \boldsymbol{z}$, где $\boldsymbol{y} \in \mathcal{Y}$, $\boldsymbol{z} \in \mathcal{Z}$.

Определение. *Оператор $\mathcal{P} = \mathcal{P}_{\mathcal{Y},\mathcal{Z}}$, определенный соотношением $\mathcal{P}\,\boldsymbol{x} = \boldsymbol{y}$ называют проектором на подпространство $\mathcal{Y}$ вдоль подпространства $\mathcal{Z}$.*

Замечание. Ясно, что $\mathcal{P}^2 = \mathcal{P}$. В связи с этим заметим, что любой линейный оператор $\mathcal{B}$, удовлетворяющий условию $\mathcal{B}^2 = \mathcal{B}$, называется *проектором.*

Рассмотрим теперь произвольный нормальный оператор $\mathcal{N} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$, действующий в n-мерном эрмитовом пространстве $\mathcal{X}$.

Пусть $\lambda_1, \dots, \lambda_m$, $m \leqslant n$, – все различные собственные значения оператора $\mathcal{N}$, а $\mathcal{V}_1, \dots, \mathcal{V}_m$ – собственные подпространства оператора $\mathcal{N}$, соответствующее собственным значениям $\lambda_1, \dots, \lambda_m$. При доказательстве Теоремы 8.19 был, по-существу, установлен тот факт, что $\mathcal{X} = \mathcal{V}_1 \oplus \dots \oplus \mathcal{V}_m$ (см. также Предложение 7.10).

Положим $\mathcal{W}_j := \bigoplus_{k \neq j} \mathcal{V}_k$ и определим $\mathcal{P}_j := \mathcal{P}_{\mathcal{V}_j, \mathcal{W}_j}$ – проектор на $\mathcal{V}_j$ вдоль $\mathcal{W}_j$. Тогда из определения проекторов $\mathcal{P}_j$ вытекает, что $\mathcal{P}_1 + \dots + \mathcal{P}_m = \mathcal{E}$, а также, что $\mathcal{P}_j\,\mathcal{P}_k = \mathcal{P}_k\,\mathcal{P}_j = \delta_{jk}\,\mathcal{P}_j$ при всех $k, j = 1, \dots, m$. Далее, для любого $\boldsymbol{x} \in \mathcal{X}$,

$$\begin{aligned} \mathcal{N}\,\boldsymbol{x} = \mathcal{N}(\mathcal{E}\,\boldsymbol{x}) = \mathcal{N}(\mathcal{P}_1\,\boldsymbol{x} + \dots + \mathcal{P}_m\,\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}_1 + \dots + \boldsymbol{x}_m) = \\ = \lambda_1 \boldsymbol{x}_1 + \dots + \lambda_m \boldsymbol{x}_m = \lambda_1\,\mathcal{P}_1\,\boldsymbol{x} + \dots + \lambda_m\,\mathcal{P}_m\,\boldsymbol{x}, \end{aligned}$$

где $\boldsymbol{x}_j = \mathcal{P}_j\,\boldsymbol{x}$, $j = 1,\dots,m$, откуда вытекает, что верно разложение

$$\mathcal{N} = \lambda_1\,\mathcal{P}_1 + \dots + \lambda_m\,\mathcal{P}_m, \tag{8.8}$$

которое называется *спектральным разложением оператора* $\mathcal{N}$.

Упражнение. Определим многочлены

$$P_j(z) = \prod_{k=1,\ k\neq j}^{m} \frac{z-\lambda_k}{\lambda_j-\lambda_k}.$$

Проверить, что $P_j(\lambda_k) = \delta_{jk}$ при $j,k = 1,\dots,m$ и, что $P_j(\mathcal{N}) = \mathcal{P}_j$ при $j = 1,\dots,m$. Заметим, что если оператор $\mathcal{N}$ является самосопряженным, то все числа λ_j, $j = 1,\dots,m$ вещественные и, соответственно, многочлен $\mathcal{P}_j$ также будет вещественным.

Из Предложения 8.3 следует, что любой оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ может быть представлен в виде $\mathcal{A} = \mathcal{A}_R + i\,\mathcal{A}_I$, где $\mathcal{A}_R$ и $\mathcal{A}_I$ – самосопряженные операторы.

Рассмотрим произвольный нормальный оператор $\mathcal{N}$ и разложим его указанным образом $\mathcal{N} = \mathcal{A}_R + i\,\mathcal{A}_I$. Тогда $\mathcal{N}^* = \mathcal{A}_R - i\,\mathcal{A}_I$, откуда $\mathcal{A}_R = \frac{1}{2}(\mathcal{N}+\mathcal{N}^*)$, а $\mathcal{A}_I = \frac{1}{2i}(\mathcal{N}-\mathcal{N}^*)$. При этом из равенства $\mathcal{N}\,\mathcal{N}^* = \mathcal{N}^*\,\mathcal{N}$ вытекает, что $\mathcal{A}_R\,\mathcal{A}_I = \mathcal{A}_I\,\mathcal{A}_R$.

Пусть теперь оператор $\mathcal{A}$ представлен в виде $\mathcal{A} = \mathcal{A}_R + i\,\mathcal{A}_I$, где самосопряженные операторы $\mathcal{A}_R$ и $\mathcal{A}_I$ коммутируют, т.е. $\mathcal{A}_R\,\mathcal{A}_I = \mathcal{A}_I\,\mathcal{A}_R$. Тогда $\mathcal{A}\,\mathcal{A}^* = \mathcal{A}_R^2 + \mathcal{A}_I^2 = \mathcal{A}^*\,\mathcal{A}$ и, следовательно, оператор $\mathcal{A}$ является нормальным.

8.6. Унитарные и ортогональные операторы

Пусть, как и выше, $\mathcal{X}$ – конечномерное эрмитово пространство, $\dim\mathcal{X} = n$, с эрмитовым скалярным произведением $\langle\cdot,\cdot\rangle$. Пусть, как обычно, $\|\boldsymbol{x}\| = \sqrt{\langle\boldsymbol{x},\boldsymbol{x}\rangle}$ при $\boldsymbol{x} \in \mathcal{X}$.

Определение. *Оператор $\mathcal{U} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ называется унитарным, если для любых элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ справедливо соотношение*

$$\langle\mathcal{U}\,\boldsymbol{x},\mathcal{U}\,\boldsymbol{y}\rangle = \langle\boldsymbol{x},\boldsymbol{y}\rangle.$$

Сразу заметим, что если оператор $\mathcal{U}$ является унитарным, то $\|\,\mathcal{U}\,\boldsymbol{x}\| = \|\boldsymbol{x}\|$ для любого элемента $\boldsymbol{x} \in \mathcal{X}$.

По определению, унитарный оператор сохраняет скалярное произведение. Следовательно, если $\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_n\}$ – ортонормированный базис в $\mathcal{X}$, а $\mathcal{U}$ – унитарный оператор, то набор векторов $\{\mathcal{U}\,\boldsymbol{e}_1,\dots,\mathcal{U}\,\boldsymbol{e}_n\}$ также образует ортонормированный базис в $\mathcal{X}$.

Установим следующий критерий унитарности линейного оператора.

Теорема 8.20. *Оператор $\mathcal{U} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ является унитарным если и только если $\mathcal{U}^*\,\mathcal{U} = \mathcal{U}\,\mathcal{U}^* = \mathcal{E}$.*

Доказательство. Пусть линейный оператор $\mathcal{U} \in \mathfrak{L}(\mathcal{X},\mathcal{X})$ является унитарным, т.е. $\langle\mathcal{U}\,\boldsymbol{x},\mathcal{U}\,\boldsymbol{y}\rangle = \langle\boldsymbol{x},\boldsymbol{y}\rangle$ для любых элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$. По определению сопряженного оператора, условие унитарности оператора $\mathcal{U}$ можно записать в виде $\langle\mathcal{U}^*\,\mathcal{U}\,\boldsymbol{x},\boldsymbol{y}\rangle = \langle\boldsymbol{x},\boldsymbol{y}\rangle$. Отсюда следует, что для любых элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ выполняется равенство

$$\langle(\mathcal{U}^*\,\mathcal{U} - \mathcal{E})\boldsymbol{x},\boldsymbol{y}\rangle = 0.$$

Подставив в это равенство $\boldsymbol{y} = \mathcal{U}^*\,\mathcal{U}\,\boldsymbol{x} - \boldsymbol{x}$ получим, что для любого $\boldsymbol{x} \in \mathcal{X}$ имеет место равенство $\|(\mathcal{U}^*\,\mathcal{U} - \mathcal{E})\boldsymbol{x}\| = 0$ и, следовательно, $\mathcal{U}^*\,\mathcal{U} - \mathcal{E} = \mathcal{O}$, т.е. $\mathcal{U}^*\,\mathcal{U} = \mathcal{E}$. Из этого равенства и из того, что пространство $\mathcal{X}$ конечномерно, вытекает, что $\mathcal{U}\,\mathcal{U}^* = \mathcal{E}$.

Проверим теперь справедливость обратного утверждения. Пусть $\mathcal{U}^*\,\mathcal{U} = \mathcal{U}\,\mathcal{U}^* = \mathcal{E}$ и исходя из определения сопряженного оператора получаем, что для любых элементов $\boldsymbol{x} \in \mathcal{X}$ и $\boldsymbol{y} \in \mathcal{X}$ справедливы равенства

$$\langle\mathcal{U}\,\boldsymbol{x},\mathcal{U}\,\boldsymbol{y}\rangle = \langle\boldsymbol{x},\mathcal{U}^*\,\mathcal{U}\,\boldsymbol{y}\rangle = \langle\boldsymbol{x},\mathcal{E}\,\boldsymbol{y}\rangle = \langle\boldsymbol{x},\boldsymbol{y}\rangle,$$

что в точности означает унитарность оператора $\mathcal{U}$. □

Следствие. *Оператор* $\mathcal{U} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ *– унитарный если и только если* $\mathcal{U}^* = \mathcal{U}^{-1}$.

Легко проверить, что любое собственное значение λ унитарного оператора $\mathcal{U}$ удовлетворяет условию $|\lambda| = 1$. В самом деле, пусть λ – собственное значение унитарного оператора $\mathcal{U}$, соответствующее единичному собственному вектору $\boldsymbol{e}_\lambda$, т.е. $\mathcal{U}\,\boldsymbol{e}_\lambda = \lambda \boldsymbol{e}_\lambda$ и $\|\boldsymbol{e}_\lambda\| = 1$. Отсюда немедленно заключаем, что

$$|\lambda| = \|\lambda \boldsymbol{e}_\lambda\| = \|\,\mathcal{U}\,\boldsymbol{e}_\lambda\| = \|\boldsymbol{e}_\lambda\| = 1.$$

Из Теоремы 8.20 вытекает, что всякий унитарный оператор является нормальным. Следовательно, для унитарного оператора существует базис, в котором его матрица является диагональной. Более того, так как все собственные значения унитарного оператора по модулю равны единице, то для любого унитарного оператора найдется базис, в котором его матрица может быть записана в виде $\mathrm{Diag}\{e^{i\theta_1}, \ldots, e^{i\theta_n}\}$, $\theta_j \in \mathbb{R}$, $j = 1, \ldots, n$.

Напомним, что если A – комплексная $n \times n$ матрица, то через A^* обозначается эрмитово сопряженная матрица для матрицы A, т.е. $A^* = \overline{A}^\top$.

Определение. *Комплексная* $n \times n$ *матрица* A *называется унитарной, если* $A^*A = AA^* = E$

Упражнение. Проверить, что если $\mathcal{X}$ – эрмитово пространство, $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – ортонормированный базис в $\mathcal{X}$, то оператор $\mathcal{U}$, действующий в $\mathcal{X}$ является унитарным если и только если его матрица в базисе $\boldsymbol{e}$ является унитарной.

Ортогональные операторы в евклидовых пространствах. Ортогональные операторы являются аналогом унитарных операторов в вещественном случае. Пусть $\mathcal{X}$ – (вещественное) евклидово пространство со скалярным произведением $\langle \cdot, \cdot \rangle$.

Определение. *Оператор* $\mathcal{P} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ *называется ортогональным, если для любых элементов* $\boldsymbol{x} \in \mathcal{X}$ *и* $\boldsymbol{y} \in \mathcal{X}$ *имеет место равенство*

$$\langle \mathcal{P}\,\boldsymbol{x}, \mathcal{P}\,\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{y} \rangle.$$

По определению, ортогональный оператор сохраняет скалярное произведение. Следовательно, если $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – ортонормированный базис в $\mathcal{X}$, а $\mathcal{P}$ – ортогональный оператор, то набор векторов $\{\mathcal{P}\,\boldsymbol{e}_1, \ldots, \mathcal{P}\,\boldsymbol{e}_n\}$ также образует ортонормированный базис в $\mathcal{X}$.

Теорема 8.21. *Оператор* $\mathcal{P} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ *в вещественном евклидовом пространстве* $\mathcal{X}$ *является ортогональным, если и только если* $\mathcal{P}^* = \mathcal{P}^{-1}$ *(т.е., если сопряженный оператор к* $\mathcal{P}$ *совпадает с обратным).*

Доказательство этого утверждения является дословным повторением доказательства Теоремы 8.20.

Напомним, что ранее было введено понятие *ортогональной матрицы*: вещественная $n \times n$ матрица A называется ортогональной, если $A^\top A = AA^\top = E$. Как и в случае унитарных операторов, в рассматриваемой ситуации имеет место следующее утверждение. Пусть $\mathcal{X}$ – евклидово пространство, а $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ – ортонормированный базис в $\mathcal{X}$. Тогда оператор $\mathcal{P}$, действующий в $\mathcal{X}$ является ортогональным если и только если его матрица в базисе $\boldsymbol{e}$ ортогональна. Доказательство этого утверждения, равно как и соответствующего утверждения про унитарные операторы оставляется в качестве *упражнения*.

Канонический вид ортогонального оператора. Найдем канонический вид ортогонального оператора $\mathcal{P}$, действующего в (вещественном) евклидовом пространстве $\mathcal{X}$. Если $\dim \mathcal{X} = 1$, то любой вектор $\boldsymbol{x} \in \mathcal{X}$ имеет вид $\boldsymbol{x} = \alpha \boldsymbol{e}$, где $\alpha \in \mathbb{R}$, а $\boldsymbol{e}$ – вектор, порождающий пространство $\boldsymbol{x}$. Так как $\mathcal{P}\,\boldsymbol{e} = \lambda \boldsymbol{e}$ и $\langle \mathcal{P}\,\boldsymbol{e}, \mathcal{P}\,\boldsymbol{e} \rangle = \lambda^2 \langle \boldsymbol{e}, \boldsymbol{e} \rangle$, то из ортогональности $\mathcal{P}$ получаем, что $\lambda = \pm 1$. Следовательно, в одномерном пространстве существует только два ортогональных преобразования: $\boldsymbol{x} \mapsto \boldsymbol{x}$ и $\boldsymbol{x} \mapsto -\boldsymbol{x}$.

Пусть теперь $\dim \mathcal{X} = 2$ и пусть оператор $\mathcal{P}$ имеет матрицу

$$P = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

где $a, b, c, d \in \mathbb{R}$. Так как матрица P должна быть ортогональной, то $PP^\top = P^\top P = E$. Из этих равенств вытекают следующие соотношения на числа a, b, c, d:

$$a^2 + b^2 = 1, \quad a^2 = d^2, \quad b^2 = c^2, \quad ac + bd = 0, \quad ab + cd = 0.$$

Из этих равенств вытекает, что матрица P может иметь только следующий вид:

$$P_{\theta+} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad P_{\theta-} = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}.$$

Заметим, что $\det P_{\theta+} = 1$, а $\det P_{\theta-} = -1$. Матрица $P_{\theta+}$ задает очевидное преобразование – поворот в плоскости, образованной векторами $\boldsymbol{e}_1$ и $\boldsymbol{e}_2$ на угол θ. Так как

$$P_{\theta-} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} P_{\theta+},$$

то матрица $P_{\theta-}$ задает преобразование, сводящееся к последовательному выполнению поворота на угол θ и отражения относительно оси $\boldsymbol{e}_1$.

Определим операторы $\mathcal{P}_{\theta+}$ и $\mathcal{P}_{\theta-}$ так, что в стандартном базисе $\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ их матрицы равны соответственно $P_{\theta+}$ и $P_{\theta-}$. Из ортогональности оператора $\mathcal{P}_{\theta+}$ вытекает, что векторы $\mathcal{P}_{\theta+}\,\boldsymbol{e}_1$ и $\mathcal{P}_{\theta+}\,\boldsymbol{e}_2$ образуют ортонормированный базис, в котором матрица оператора $\mathcal{P}_{\theta-}$ имеем матрицу

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

т.е. является диагональной.

Для того, чтобы выяснить общий вид ортогонального оператора $\mathcal{P}$ в вещественном евклидовом пространстве $\mathcal{X}$ используем следующий технический прием, называемый *комплексификацией* пространства $\mathcal{X}$. Рассмотрим комплексное линейное пространство $\mathcal{X}_\mathbb{C}$, определенное следующим образом

$$\mathcal{X}_\mathbb{C} := \{\boldsymbol{x} + i\boldsymbol{y} : \ \boldsymbol{x} \in \mathcal{X}, \boldsymbol{y} \in \mathcal{X}\},$$

причем операции сложения векторов в $\mathcal{X}_\mathbb{C}$ и умножения векторов из $\mathcal{X}_\mathbb{C}$ на комплексные числа определены следующим образом

$$(\boldsymbol{x}_1 + i\boldsymbol{y}_1) + (\boldsymbol{x}_2 + i\boldsymbol{y}_2) = (\boldsymbol{x}_1 + \boldsymbol{x}_2) + i(\boldsymbol{y}_1 + \boldsymbol{y}_2) \in \mathcal{X}_\mathbb{C},$$
$$(\alpha + i\beta)(\boldsymbol{x} + i\boldsymbol{y}) = (\alpha\boldsymbol{x} - \beta\boldsymbol{y}) + i(\alpha\boldsymbol{y} + \beta\boldsymbol{x}) \in \mathcal{X}_\mathbb{C}.$$

Непосредственная проверка того факта, что $\mathcal{X}_\mathbb{C}$ – комплексное линейной пространство оставляется в качестве *упражнения*. Кроме того, в качестве *упражнения* предлагается проверить, что если система векторов $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ образует базис в $\mathcal{X}$, то эти же векторы образуют базис в $\mathcal{X}_\mathbb{C}$.

Пусть $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y} \in \mathcal{X}_\mathbb{C}$. Тогда вектор $\overline{\boldsymbol{z}} = \boldsymbol{x} - i\boldsymbol{y} \in \mathcal{X}_\mathbb{C}$ называется *комплексно сопряженным* вектором к вектору $\boldsymbol{z}$ (используется также термин “просто” сопряженный, если это не приводит к терминологической путанице). В качестве *упражнения* предлагается проверить, что если $\boldsymbol{z} \in \mathcal{X}_\mathbb{C}$, а $a \in \mathbb{C}$, то $\overline{a\boldsymbol{z}} = \overline{a} \cdot \overline{\boldsymbol{z}}$.

Если в вещественном линейном пространстве $\mathcal{X}$ задано скалярное произведение $\langle \cdot, \cdot \rangle$, то скалярное произведение в $\mathcal{X}_\mathbb{C}$ можно ввести следующим образом: для $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y} \in \mathcal{X}_\mathbb{C}$ и $\boldsymbol{w} = \boldsymbol{u} + i\boldsymbol{v} \in \mathcal{X}_\mathbb{C}$ положим

$$\langle \boldsymbol{z}, \boldsymbol{w} \rangle_\mathbb{C} := \langle \boldsymbol{x}, \boldsymbol{u} \rangle + \langle \boldsymbol{y}, \boldsymbol{v} \rangle + i(\langle \boldsymbol{y}, \boldsymbol{u} \rangle - \langle \boldsymbol{x}, \boldsymbol{v} \rangle).$$

Проверка того, что выражение $\langle \boldsymbol{z}, \boldsymbol{w} \rangle_\mathbb{C}$ задает эрмитово скалярное произведение в $\mathcal{X}_\mathbb{C}$ оставляется в качестве *упражнения*.

Заметим также, что введенное скалярное произведение обладает таким свойством, что $\langle \overline{\boldsymbol{z}}, \overline{\boldsymbol{w}} \rangle_\mathbb{C} = \overline{\langle \boldsymbol{z}, \boldsymbol{w} \rangle}_\mathbb{C}$. Если в дальнейшем из контекста ясно, о каком пространстве ($\mathcal{X}$

или $\mathcal{X}_{\mathbb{C}}$) идет речь, то мы будем использовать одинаковое обозначение для скалярного произведения в $\mathcal{X}$ и для его комплексификации.

Если в пространстве $\mathcal{X}$ задан линейный оператор $\mathcal{A}$, то он может быть продолжен на пространство $\mathcal{X}_{\mathbb{C}}$ следующим естественным образом: если $\boldsymbol{z} = \boldsymbol{x} + i\boldsymbol{y} \in \mathcal{X}_{\mathbb{C}}$, то $\mathcal{A}\,\boldsymbol{z} = \mathcal{A}\,\boldsymbol{x} + i\,\mathcal{A}\,\boldsymbol{y}$. Ясно, что при таком определении $\mathcal{A}$ становится линейным оператором на $\mathcal{X}_{\mathbb{C}}$ (проверка оставляется в качестве *упражнения*).

Рассмотрим теперь оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ и сопряженный оператор $\mathcal{A}^*$. Продолжим оба эти оператора на пространство $\mathcal{X}_{\mathbb{C}}$ так как было описано выше. Тогда продолжения этих операторов останутся взаимно сопряженными. В самом деле, при всех $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{v} \in \mathcal{X}$ получаем

$$\langle \mathcal{A}(\boldsymbol{x} + i\boldsymbol{y}), \boldsymbol{u} + i\boldsymbol{v} \rangle = \langle \mathcal{A}\,\boldsymbol{x} + i\,\mathcal{A}\,\boldsymbol{y}, \boldsymbol{u} + i\boldsymbol{v} \rangle = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{u} \rangle + \langle \mathcal{A}\,\boldsymbol{y}, \boldsymbol{v} \rangle - i\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{v} \rangle + i\langle \mathcal{A}\,\boldsymbol{y}, \boldsymbol{u} \rangle =$$
$$= \langle \boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{u} \rangle + \langle \boldsymbol{y}, \mathcal{A}^*\,\boldsymbol{v} \rangle - i\langle \boldsymbol{x}, \mathcal{A}^*\,\boldsymbol{v} \rangle + i\langle \boldsymbol{y}, \mathcal{A}^*\,\boldsymbol{u} \rangle = \langle \boldsymbol{x} + i\boldsymbol{y}, \mathcal{A}^*(\boldsymbol{u} + i\boldsymbol{v}) \rangle.$$

Из определения продолжения оператора $\mathcal{A}$ с пространства $\mathcal{X}$ на пространство $\mathcal{X}_{\mathbb{C}}$ немедленно вытекает, что продолжение нормального оператора будет нормальным оператором, а продолжение ортогонального оператора будет унитарным оператором. Проверка этих фактов оставляется в качестве *упражнения*.

Итак, рассмотрим нормальный оператор $\mathcal{A}$ в $\mathcal{X}$ и его продолжение на $\mathcal{X}_{\mathbb{C}}$. Вспомним, что у нормального оператора, действующего в эрмитовом пространстве существует ортонормированный базис из собственный векторов, в котором матрица этого оператора имеет диагональный вид.

Для вещественных собственных значений (продолженного) оператора $\mathcal{A}$ можно выбрать вещественные собственные векторы (т.е. собственные векторы, лежащие в $\mathcal{X}$). В самом деле, система уравнений, из которой находятся координаты собственных векторов относительно базиса $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ в $\mathcal{X}$ в случае вещественного собственного значения – это система линейных однородных уравнений с вещественной матрицей.

Если λ – комплексное собственное значение оператора $\mathcal{A}$, то $\overline{\lambda}$ – также является собственным значением $\mathcal{A}$. В самом деле, если $\boldsymbol{z}$ – собственный вектор, соответствующий собственному значению λ, т.е. $\mathcal{A}\,\boldsymbol{z} = \lambda\boldsymbol{z}$, то $\mathcal{A}\,\overline{\boldsymbol{z}} = \overline{\mathcal{A}\,\boldsymbol{z}} = \overline{\lambda\boldsymbol{z}} = \overline{\lambda} \cdot \overline{\boldsymbol{z}}$.

Более того, комплексные собственные значений появляются парами с одинаковой кратностью. Выберем ортонормированный базис из собственных векторов, принадлежащих некоторому собственному значению $\lambda = \alpha + i\beta$ при $\beta \neq 0$. Как показано выше, базис из собственных векторов, принадлежащих собственному значению $\overline{\lambda}$ можно взять из векторов, комплексно сопряженных с соответствующими базисными собственными векторами для λ. Эти базисные векторы также будут попарно ортогональны и ортонормированы. Для каждой пары $\boldsymbol{x} + i\boldsymbol{y}$ и $\boldsymbol{x} - i\boldsymbol{y}$ рассмотрим подпространство $\mathrm{Span}_{\mathbb{C}}\{\boldsymbol{x} + i\boldsymbol{y}, \boldsymbol{x} - i\boldsymbol{y}\}$. Все эти подпространства $\mathcal{A}$-инвариантны, попарно ортогональны и ортогональны всем вещественным собственным векторам, соответствующим вещественным собственным значениям.

Из соотношения $\mathrm{Span}_{\mathbb{C}}\{\boldsymbol{x} + i\boldsymbol{y}, \boldsymbol{x} - i\boldsymbol{y}\} = \mathrm{Span}_{\mathbb{C}}\{\boldsymbol{x}, \boldsymbol{y}\}$ вытекает, что $\mathrm{Span}_{\mathbb{C}}\{\boldsymbol{x}, \boldsymbol{y}\} = (\mathrm{Span}_{\mathbb{R}}\{\boldsymbol{x}, \boldsymbol{y}\})_{\mathbb{C}}$. Из ортогональности собственных векторов $\boldsymbol{x} + i\boldsymbol{y}$ и $\boldsymbol{x} - i\boldsymbol{y}$ (принадлежащих различным собственным значениям) следует

$$0 = \langle \boldsymbol{x} + i\boldsymbol{y}, \boldsymbol{x} - i\boldsymbol{y} \rangle = \langle \boldsymbol{x}, \boldsymbol{x} \rangle - \langle \boldsymbol{y}, \boldsymbol{y} \rangle + 2i\langle \boldsymbol{x}, \boldsymbol{y} \rangle,$$

причем последнее равенство вытекает из того, что в вещественном евклидовом пространстве $\mathcal{X}$ скалярное произведение симметрично. Отделяя в последнем равенстве вещественные и мнимые части получаем, что $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$, т.е. векторы $\boldsymbol{x}$ и $\boldsymbol{y}$ ортогональны. Кроме того, получаем, что $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \langle \boldsymbol{y}, \boldsymbol{y} \rangle$. Так как вектор $\boldsymbol{x} + i\boldsymbol{y}$ имеет единичную норму, то $\langle \boldsymbol{x}, \boldsymbol{x} \rangle + \langle \boldsymbol{y}, \boldsymbol{y} \rangle = 1$. Отсюда окончательно получаем, что $\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \langle \boldsymbol{y}, \boldsymbol{y} \rangle = 1/2$ и, следовательно, векторы $\sqrt{2}\boldsymbol{x}$ и $\sqrt{2}\boldsymbol{y}$ имеют единичную норму.

Далее, для собственного вектора $\boldsymbol{x} + i\boldsymbol{y}$, принадлежащего собственному значению $\lambda = \alpha + i\beta$, получаем

$$\mathcal{A}(\boldsymbol{x} + i\boldsymbol{y}) = (\alpha + i\beta)(\boldsymbol{x} + i\boldsymbol{y}),$$

откуда, отделяя вещественные и мнимые части, получаем

$$\mathcal{A}\,\boldsymbol{x} = \alpha\boldsymbol{x} - \beta\boldsymbol{y}, \qquad \mathcal{A}\,\boldsymbol{y} = \beta\boldsymbol{x} + \alpha\boldsymbol{y}.$$

Итак, для нормального оператора, действующего в вещественном евклидовом пространстве существует ортонормированный базис, состоящий из собственных векторов, соответствующих вещественным собственным значениям и из умноженных на $\sqrt{2}$ вещественных и мнимых частей собственных векторов, соответствующих комплексным собственным значениям. Одномерные подпространства, натянутые на соответствующие собственные векторы и двумерные подпространства, натянутые на компоненты комплексных собственных векторов $\mathcal{A}$-инвариантны и, следовательно, матрица оператора $\mathcal{A}$ в таком базисе имеет блочно-диагональный вид и состоит из 1×1 и 2×2-блоков. Более того, каждый 2×2-блок такой матрицы имеет вид

$$\begin{pmatrix} \alpha & \beta \\ -\beta & \alpha \end{pmatrix}.$$

Если теперь оператор $\mathcal{P}$ – ортогонален, то структура его одномерных и двумерных собственных инвариантных подпространств была изучена выше и, как было установлено, каждый 1×1-блок матрицы ортогонального оператора имеет вид (± 1), а каждый 2×2-блок – вид

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Окончательно получаем, что в вещественном евклидовом пространстве существует ортонормированный базис, в котором матрица ортогонального оператора $\mathcal{P}$ имеет вид

$$\mathrm{Diag}\left\{1,\ldots,1,-1,\ldots,-1, \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix}, \ldots, \begin{pmatrix} \cos\theta_\ell & -\sin\theta_\ell \\ \sin\theta_\ell & \cos\theta_\ell \end{pmatrix}\right\}.$$

РАЗДЕЛ 9

Квадратичные формы и уравнения гиперповерхностей второго порядка

В этой лекции мы будем всюду считать, что $\mathcal{X}$ – вещественное линейное пространство, $\dim\mathcal{X} = n$, $n \in \mathbb{N}$. В ряде случаев мы будем также предполагать, что $\mathcal{X}$ является евклидовым пространством со скалярным произведением $\langle\cdot,\cdot\rangle$.

Основным объектом изучения в этом разделе будет понятие *квадратичной формы*. Это понятие тесно связано с понятием *билинейной формы*, которое уже возникало в нашем курсе (соответствующее определение было дано в разделе 3.6; см. также 8.4).

9.1. Квадратичные формы

Пусть $\mathcal{B}$ – *симметричная* билинейная форма на $\mathcal{X}$ (т.е. $\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) = \mathcal{B}(\boldsymbol{y},\boldsymbol{x})$ для всех $\boldsymbol{x},\boldsymbol{y} \in \mathcal{X}$).

Определение. *Квадратичной формой называется функция* $\mathcal{Q} : \mathcal{X} \to \mathbb{R}$*, определенная равенством* $\mathcal{Q}(\boldsymbol{x}) = \mathcal{B}(\boldsymbol{x},\boldsymbol{x})$*. Симметричная билинейная форма* $\mathcal{B}$ *называется полярной к квадратичной форме* $\mathcal{Q}$*.*

Иногда для квадратичной формы не вводится дополнительного обозначения, а используется обозначение $\mathcal{B}(\boldsymbol{x},\boldsymbol{x})$, где $\mathcal{B}$ – соответствующая симметричная билинейная форма.

Нетрудно проверить, что квадратичная форма $\mathcal{Q}$ и полярная билинейная форма $\mathcal{B}$ связаны соотношением

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) = \frac{1}{2}\Big(\mathcal{Q}(\boldsymbol{x}+\boldsymbol{y}) - \mathcal{Q}(\boldsymbol{x}) - \mathcal{Q}(\boldsymbol{y})\Big).$$

В самом деле,

$$\mathcal{B}(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{x}+\boldsymbol{y}) = \mathcal{B}(\boldsymbol{x},\boldsymbol{x}) + \mathcal{B}(\boldsymbol{x},\boldsymbol{y}) + \mathcal{B}(\boldsymbol{y},\boldsymbol{x}) + \mathcal{B}(\boldsymbol{y},\boldsymbol{y})$$

и, для проверки требуемого свойства, нам остается воспользоваться свойством симметрии билинейной формы $\mathcal{B}$.

Пусть теперь в пространстве $\mathcal{X}$ выбран некоторый базис $\boldsymbol{e} = \{\boldsymbol{e}_1,\dots,\boldsymbol{e}_n\}$ и пусть $B = (b_{jk})$ – матрица билинейной формы $\mathcal{B}$ в базисе $\boldsymbol{e}$. В силу симметричности билинейной формы $\mathcal{B}$, матрица B является симметричной. Как было показано выше, билинейная форма $\mathcal{B}$ может быть представлена в виде

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{y}) = [\boldsymbol{x}]^\top B[\boldsymbol{y}] = \sum_{j,k=1}^{n} b_{jk}x_jy_k,$$

где $[x]^\top = (x_1,\dots,x_n)$ и $[y]^\top = (y_1,\dots,y_n)$ – координатные столбцы векторов $\boldsymbol{x}$ и $\boldsymbol{y}$ в базисе $\boldsymbol{e}$. Отсюда вытекает, что квадратичная форма $\mathcal{B}(\boldsymbol{x},\boldsymbol{x})$ представляется в виде

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = [x]^\top B[x] = \sum_{j,k=1}^{n} b_{jk}x_jx_k.$$

Матрица B называется также *матрицей квадратичной формы* $\mathcal{Q}(\boldsymbol{x}) = \mathcal{B}(\boldsymbol{x},\boldsymbol{x})$. Верно и обратное, по любой симметричной матрице B можно определить квадратичную форму $\mathcal{B}$ соотношением $\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = [\boldsymbol{x}]^\top B[\boldsymbol{x}]$. Ясно, что при переходе от базиса к базису матрица квадратичной формы преобразуется также, как и матрица билинейной формы. Как и в случае билинейных форм, *рангом квадратичной формы* называется ранг матрицы этой формы в любом базисе пространства $\mathcal{X}$. Если ранг квадратичной формы равен размерности пространства, то такая форма называется *невырожденной*, а если ранг квадратичной формы строго меньше размерности пространства, то такая форма называется *вырожденной*.

Определение. *Квадратичная форма* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x})$ *называется*

(1) *положительно определенной, если для любого* $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{x} \neq \mathbf{0}$ *выполнено неравенство* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) > 0$*;*
(2) *отрицательно определенной, если для любого* $\boldsymbol{x} \in \mathcal{X}$, $\boldsymbol{x} \neq \mathbf{0}$ *выполнено неравенство* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) < 0$*;*
(3) *знакопеременной, если существуют такие элементы* $\boldsymbol{x} \in \mathcal{X}$ *и* $\boldsymbol{y} \in \mathcal{X}$*, что выполнены неравенства* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) > 0$ *и* $\mathcal{B}(\boldsymbol{y}, \boldsymbol{y}) < 0$*;*
(4) *неотрицательно определенной, если для всех элементов* $\boldsymbol{x} \in \mathcal{X}$ *выполнено неравенство* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$*, но существует* $\boldsymbol{x}_0 \neq \mathbf{0}$ *такой, что* $\mathcal{B}(\boldsymbol{x}_0, \boldsymbol{x}_0) = 0$.
(5) *неположительно определенной, если для всех элементов* $\boldsymbol{x} \in \mathcal{X}$ *выполнено неравенство* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \leqslant 0$*, но существует* $\boldsymbol{x}_0 \neq \mathbf{0}$ *такой, что* $\mathcal{B}(\boldsymbol{x}_0, \boldsymbol{x}_0) = 0$.

Если квадратичная форма положительно или отрицательно определена, то ее называют знакоопределенной. Если квадратичная форма неотрицательно или неположительно определена, то ее называют квазизнакоопределенной.

Ниже будут получен ряд признаков, позволяющих судить о принадлежности квадратичной формы к одному из четырех типов, введенных в сформулированном выше определении.

Проверим, что если $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$ – это билинейная форма, полярная к положительно определенной квадратичной форме $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x})$, то $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$ удовлетворяет всем свойствам скалярного произведения в вещественном евклидовом пространстве. Итак, свойство $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{B}(\boldsymbol{y}, \boldsymbol{x})$ выполнено, так как билинейная форма $\mathcal{B}$ является симметричной (это вытекает из определения полярной формы). Необходимые свойства линейности вытекают из соответствующих свойств линейности билинейной формы (проверка оставляется в качестве *упражнения*). Четвертая аксиома скалярного произведения, т.е. свойство $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$ при всех $\boldsymbol{x} \in \mathcal{X}$ и $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = 0$ при $\boldsymbol{x} = \mathbf{0}$, вытекает из положительной определенности квадратичной формы $\mathcal{B}$.

Замечание. Набор аксиом скалярного произведения можно рассматривать как совокупность условий, определяющих билинейную форму, полярную некоторой положительно определенной квадратичной форме. Из этого вытекает, что скалярное произведение в линейных пространствах может быть задано при помощи таких билинейных форм.

9.2. Канонический вид квадратичных форм

Рассмотрим ряд методов, позволяющих найти в пространстве $\mathcal{X}$ такой базис, по отношению к которому некоторая заданная квадратичная форма $\mathcal{B}$ представляется в *каноническом виде*, т.е. в виде

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \lambda_1 x_1^2 + \cdots + \lambda_n x_n^2,$$

где $(x_1, \ldots, x_n)^\top$ – координатный столбец вектора $\boldsymbol{x}$ в соответствующем базисе.

Так как каждому преобразованию базиса соответствует невырожденное линейное преобразование координат и наоборот, каждому невырожденному линейному преобразования координат соответствует преобразование базиса, то вопрос о приведении формы к каноническому виду можно решать путем выбора соответствующего невырожденного линейного преобразования координат.

Метод Лагранжа. Изложим первый метод приведения квадратичной формы к каноническому виду. Этот метод основан на последовательном дополнении квадратного трехчлена до полного квадрата. Соответствующий метод носит название "метода Лагранжа". Одновременно с описанием метода Лагранжа мы докажем следующее утверждение

Теорема 9.1. *Любая квадратичная форма $\mathcal{B}(\boldsymbol{x},\boldsymbol{x})$, заданная в n-мерном вещественном евклидовом пространстве $\mathcal{X}$ некоторым невырожденным линейным преобразованием может быть приведена к каноническому виду.*

Пусть $\mathcal{B} \not\equiv 0$ и пусть в некотором базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ форма $\mathcal{B}$ имеет вид

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = \sum_{j,k=1}^{n} b_{jk} x_j x_k,$$

где $(x_1, \dots, x_n)^\top$ – координаты вектора $\boldsymbol{x}$ относительно базиса $\boldsymbol{e}$.

Первым делом проверим, что при помощи линейного невырожденного преобразования форму $\mathcal{B}$ можно преобразовать таким образом, что коэффициент при квадрате первой координаты вектора $\boldsymbol{x}$ (в некотором новом базисе) будет отличен от нуля.

Если это уже так в базисе $\boldsymbol{e}$, то соответствующее преобразование берется тождественным. Если $b_{11} = 0$, но если отличен от нуля коэффициент при квадрате какой-то другой координаты, то перестановкой векторов базиса мы получаем требуемую ситуацию. И, наконец, если все коэффициенты при квадратах координат равны нулю, то требуемое преобразование построим следующим образом. Предположим, без ограничения общности, что $b_{12} \neq 0$ и рассмотрим преобразование координат

$$x_1' = (x_1 + x_2)/2, \qquad x_2' = (x_1 - x_2)/2, \qquad x_j' = x_j, \quad j = 3, \dots, n.$$

В новых координатах

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = 2b_{12}(x_1'^2 - x_2'^2) + \cdots$$

Итак, считаем, что $b_{11} \neq 0$. Представим форму $\mathcal{B}$ в виде

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = b_{11}x_1^2 + 2b_{12}x_1x_3 + \cdots + 2b_{1n}x_1x_n + \sum_{j,k=2}^{n} b_{jk}x_jx_k$$

и заметим, что

$$b_{11}x_1^2 + 2b_{12}x_1x_3 + \cdots + 2b_{1n}x_1x_n = b_{11}\left(x_1 + \frac{b_{12}}{b_{11}}x_2 + \cdots + \frac{b_{1n}}{b_{11}}x_n\right)^2 -$$
$$- \frac{b_{12}^2}{b_{11}}x_2^2 - \cdots - \frac{b_{1n}^2}{b_{11}}x_n^2 - 2\frac{b_{12}b_{13}}{b_{11}}x_2x_3 - \cdots - 2\frac{b_{1(n-1)}b_{1n}}{b_{11}}x_{n-1}x_n.$$

Ясно, что теперь выражение для $\mathcal{B}(\boldsymbol{x},\boldsymbol{x})$ можно записать в виде

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = b_{11}\left(x_1 + \frac{b_{12}}{b_{11}}x_2 + \cdots + \frac{b_{1n}}{b_{11}}x_n\right)^2 + \sum_{j,k=2}^{n} \widetilde{b}_{jk}x_jx_k,$$

где $\widetilde{b}_{jk}$ – новые коэффициенты.

Рассмотрим теперь преобразование координат

$$x_1' = x_1 + \frac{b_{12}}{b_{11}}x_2 + \cdots + \frac{b_{1n}}{b_{11}}x_n, \qquad x_j' = x_j, \quad j = 2, \dots, n.$$

В новых координатах форма $\mathcal{B}$ имеет вид

$$\mathcal{B}(\boldsymbol{x},\boldsymbol{x}) = b_{11}x_1'^2 + \sum_{j,k=2}^{n} \widetilde{b}_{jk}x_j'x_k'.$$

Аналогичный процесс применим теперь к квадратичной форме $\sum_{j,k=2}^{n} \widetilde{b}_{jk}x_j'x_k'$. Эта форма или тождественно равна нулю (в этом случае вопрос о приведении исходной формы к каноническому виду решен) или не равна тождественно нулю. Во втором случае к форме $\sum_{j,k=2}^{n} \widetilde{b}_{jk} x'_j x'_k$ применим описанный выше процесс, причем при преобразовании координат нам не надо изменять координату x'_1, а надо работать только с оставшимися координатами $x'_2, \dots, x'_n$. Ясно, что за конечное число шагов этот процесс приведет нас к каноническому виду квадратичной формы $\mathcal{B}$.

Одновременно с завершением описания процесса приведения формы к каноническому виду методом Лагранжа мы доказали теорему 9.1.

Базис, в котором квадратичная форма имеет канонический вид часто называют *каноническим*, а соответствующие коэффициенты $\lambda_1, \dots, \lambda_n$ квадратичной формы – *каноническими* коэффициентами. Ясно, что канонический базис для квадратичной формы не является единственным. Соответственно, термины "канонический базис" и "канонические коэффициенты" являются не вполне адекватными. Однако они применяются в силу существующей традиции.

Заметим также, что в результате приведения квадратичной формы к каноническому виду не все канонические коэффициенты λ_j отличны от нуля. Оставив в каноническом разложении только ненулевые коэффициенты и, переставив, если надо, элементы канонического базиса, можно привести квадратичную форму к виду

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \lambda_1 x_1^2 + \dots + \lambda_k x_k^2.$$

В этом представлении, очевидно, $k \leqslant n$. Так как ранг квадратичной формы равен рангу ее матрицы в любом базисе, то отсюда следует, что *число отличных от нуля канонических коэффициентов квадратичной формы равно рангу этой формы.*

Метод Якоби. Пусть в пространстве $\mathcal{X}$ задан базис, $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$. Скажем, по определению, что преобразование базисных векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ является *треугольным*, оно имеет следующий вид

$$\begin{aligned}
&\boldsymbol{w}_1 = \boldsymbol{e}_1, \\
&\boldsymbol{w}_2 = \alpha_{21}\boldsymbol{e}_1 + \boldsymbol{e}_2, \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\
&\boldsymbol{w}_n = \alpha_{n1}\boldsymbol{e}_1 + \alpha_{n2}\boldsymbol{e}_2 + \dots + \boldsymbol{e}_n.
\end{aligned}$$

Заметим, что определитель матрицы треугольного преобразования равен 1, следовательно такая матрица невырождена и, следовательно, векторы $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ образуют базис.

Пусть $\mathcal{B}$ – квадратичная форма, заданная на пространстве $\mathcal{X}$ и пусть B – матрица этой квадратичной формы в базисе $\boldsymbol{e}$. Рассмотрим набор определителей

$$\Delta_1 = b_{11}, \quad \Delta_2 = \det\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}, \quad \cdots, \quad \Delta_{n-1} = \det\begin{pmatrix} b_{11} & \cdots & b_{1(n-1)} \\ \dots & \dots & \dots \\ b_{(n-1)1} & \cdots & b_{(n-1)(n-1)} \end{pmatrix}.$$

Теорема 9.2. *Пусть все миноры $\Delta_1, \dots, \Delta_{n-1}$ матрицы B квадратичной формы $\mathcal{B}$ в базисе $\boldsymbol{e}$ отличны от нуля. Тогда существует единственное треугольное преобразование базисных векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$, приводящее квадратичную форму $\mathcal{B}$ к каноническому виду.*

Доказательство. Коэффициенты c_{jk}, $j, k = 1, \dots, n$, квадратичной формы $\mathcal{B}$ в базисе $\boldsymbol{w} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ вычисляются по формулам $c_{jk} = \mathcal{B}(\boldsymbol{w}_j, \boldsymbol{w}_k)$. Если в базисе $\boldsymbol{w}$ форма $\mathcal{B}$ имеет канонический вид, то $c_{jk} = 0$ при $j \neq k$. Для доказательства утверждения необходимо найти такой базис $\boldsymbol{w}$, в котором будут выполняться равенства $\mathcal{B}(\boldsymbol{w}_j, \boldsymbol{w}_k) = 0$ при $j \neq k$. Если предположить, что переход от базиса $\boldsymbol{e}$ к базису $\boldsymbol{w}$ осуществляется треугольным преобразованием, то для выполнения условий $\mathcal{B}(\boldsymbol{w}_j, \boldsymbol{w}_k) = 0$ при $j \neq k$ нам достаточно чтобы выполнялись следующие равенства (указанная достаточность вытекает из линейности формы $\mathcal{B}$, проверка необходимых деталей оставляется в качестве *упражнения*)

$$\mathcal{B}(\boldsymbol{e}_1, \boldsymbol{w}_j) = 0, \quad \mathcal{B}(\boldsymbol{e}_2, \boldsymbol{w}_j) = 0, \quad \dots, \quad \mathcal{B}(\boldsymbol{e}_{j-1}, \boldsymbol{w}_j) = 0, \qquad j = 2, \dots, n.$$

Так как $\boldsymbol{w}_j = \alpha_{j1}\boldsymbol{e}_1 + \alpha_{j2}\boldsymbol{e}_2 + \dots + \alpha_{j,j-1}\boldsymbol{e}_{j-1} + \boldsymbol{e}_j$, то для коэффициентов α_{jk} получается следующая система линейных уравнений (напомним, что $B = (b_{jk})$ – матрица квадратичной формы $\mathcal{B}$ в базисе $\boldsymbol{e}$)

$$\sum_{r=1}^{j-1} \alpha_{jr} b_{pr} + b_{pj}, \qquad p = 1, \dots, j-1.$$

Определитель этой системы равен Δ_{j-1}. Так как $\Delta_{j-1} \neq 0$, то рассматриваемая система имеет единственное решение. Следовательно, существование и единственность треугольного преобразования базиса, приводящего квадратичную форму к каноническому виду доказано. □

Выведем формулы для вычисления коэффициентов α_{kj} соответствующего треугольного преобразования и канонических коэффициентов формы $\mathcal{B}$. Пусть $\Delta_{k-1,j}$ – это минор матрицы B, полученный из строк с номерами $1, 2, \dots, k-1$ и столбцов $1, 2, \dots, j-1, j+1, \dots, k$. Используя формулы Крамера решаем выписанную при доказательстве теоремы 9.2 линейную систему на α_{jk} и получаем, что

$$\alpha_{kj} = (-1)^{k+1} \frac{\Delta_{k-1,j}}{\Delta_{k-1}}.$$

Так как $\lambda_j = \mathcal{B}(\boldsymbol{w}_j, \boldsymbol{w}_j)$, то

$$\lambda_j = \mathcal{B}(\boldsymbol{w}_j, \boldsymbol{w}_j) = \mathcal{B}(\boldsymbol{e}_j, \boldsymbol{w}_j) = \sum_{r=1}^{j-1} \alpha_{jr} b_{rj} + b_{jj}.$$

Из этой формулы, подставляя выражения для α_{jk} и учитывая формулу разложения определителя по строке, получаем (проверка технических деталей, традиционно, оставляется в качестве *упражнения*), что

$$\lambda_j = \frac{\sum_{r=1}^{j-1} (-1)^{j+r} b_{rj} \Delta_{j-1,r} + b_{jj} \Delta_{j-1}}{\Delta_{j-1}}.$$

числитель последнего выражения представляет собой формулу разложения определителя Δ_j по строке с номером j. Следовательно,

$$\lambda_1 = \Delta_1, \quad \lambda_2 = \Delta_2/\Delta_1, \quad \dots, \quad \lambda_n = \Delta_n/\Delta_{n-1}, \tag{9.1}$$

где $\Delta_n = \det B$.

В дальнейшем, через $\Delta_k(A)$, $k = 1, \dots, n$, будем обозначать угловой минор порядка k матрицы A; при этом, через $\Delta_n(A)$, естественно, обозначается величина $\det A$.

Закон инерции квадратичных форм. Как было отмечено выше, ранг квадратичной формы равен числу отличных от нуля канонических коэффициентов этой формы. Следовательно, при любом способе привидения квадратичной формы к каноническому виду число ненулевых канонических коэффициентов остается неизменным. Оказывается, что от способа приведения квадратичной формы к каноническому виду не зависит не только это число, но и числа положительных и отрицательных канонических коэффициентов.

Пусть квадратичная форма $\mathcal{B}$ приведена к каноническому виду

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \lambda_1 x_1^2 + \dots + \lambda_k x_k^2,$$

причем $\lambda_1, \dots, \lambda_k$ – все отличные от нуля канонические коэффициенты. Пусть

$$\lambda_1 > 0, \dots, \lambda_q > 0, \lambda_{q+1} < 0, \dots, \lambda_k < 0.$$

В координатах

$$y_j = x_j/\sqrt{\lambda_j}, \quad j = 1, \ldots, q, \qquad y_j = x_j/\sqrt{-\lambda_j}, \quad j = q+1, \ldots, k$$

форма $\mathcal{B}$ имеет вид

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = y_1^2 + \cdots + y_q^2 - y_{q+1}^2 - \cdots - y_k^2,$$

причем к этому виду квадратичная форма приводится некоторым невырожденным линейным преобразованием координат.

Определение. *Число положительных канонических коэффициентов $I_+(\mathcal{B})$ квадратичной формы $\mathcal{B}$ называется ее положительным индексом инерции этой формы, а число $I_-(\mathcal{B})$ отрицательных канонических коэффициентов – отрицательным индексом инерции. Разность между положительным и отрицательным индексами инерции квадратичной формы называется ее сигнатурой и обозначается $\operatorname{sg}\mathcal{B}$.*

Имеют место следующие соотношения

$$I_+(\mathcal{B}) + I_-(\mathcal{B}) = \operatorname{rg}\mathcal{B}, \qquad I_+(\mathcal{B}) = \frac{1}{2}\big(\operatorname{rg}\mathcal{B} + \operatorname{sg}\mathcal{B}\big), \quad I_-(\mathcal{B}) = \frac{1}{2}\big(\operatorname{rg}\mathcal{B} - \operatorname{sg}\mathcal{B}\big).$$

Теорема 9.3. *Числа $I_+(\mathcal{B})$, $I_-(\mathcal{B})$ и $\operatorname{sg}\mathcal{B}$ не зависят от способа приведения квадратичной формы $\mathcal{B}$ к каноническому виду.*

Доказательство. Предположим, что при помощи некоторого линейного невырожденного преобразования координат рассматриваемая квадратичная форма $\mathcal{B}$ приведена к нормальному виду

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \xi_1^2 + \cdots + \xi_q^2 - \xi_{q+1}^2 - \cdots - \xi_k^2,$$

а при помощи какого-то другого линейного невырожденного преобразования координат эта форма приведена к виду

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \zeta_1^2 + \cdots + \zeta_p^2 - \zeta_{p+1}^2 - \cdots - \zeta_k^2.$$

Нам необходимо доказать, что $p = q$. Предположим, что $p > q$. Заметим, что координаты ξ_j, $j = 1, \ldots, n$ получены из некоторых исходных координат $x_1, \ldots, x_n$ линейным невырожденным преобразованием. Аналогично, координаты ζ_k, $k = 1, \ldots, n$ получены из тех же координат $x_1, \ldots, x_n$ каким-то другим невырожденными линейным преобразованием. При этом набор соотношений

$$\xi_1 = 0, \quad \ldots, \quad \xi_q = 0, \quad \zeta_{p+1} = 0, \quad \ldots, \quad \zeta_n = 0$$

можно рассматривать как систему линейных однородных уравнений относительно координат $x_1, \ldots, x_n$ некоторого вектора в исходном базисе. Так как $p > q$, то число уравнений в этой системе меньше n и, следовательно, эта система имеет ненулевое решение. Обозначим это решение через $\boldsymbol{y}$. При этом получаем, что

$$\mathcal{B}(\boldsymbol{y}, \boldsymbol{y}) = -\xi_{q+1}^2 - \cdots - \xi_k^2 = \zeta_1^2 + \cdots + \zeta_p^2.$$

Но это возможно только в случае, когда $\xi_{q+1} = \cdots = \xi_k = 0$ и $\zeta_1 = \cdots = \zeta_p = 0$. А последняя цепочка равенств означает, что в некотором базисе ненулевой вектор $\boldsymbol{y}$ имеет все нулевые координаты, что, очевидно, невозможно. Следовательно, предположение о том, что $p > q$ приводит к противоречию. Аналогично рассматривается случай, когда $q > p$. Следовательно, $p = q$ и требуемое утверждение доказано. $\square$

Классификация квадратичных форм. Имеют место следующие критерии знакоопределенности квадратичной формы, выраженные в терминах ее индексов инерции.

Предложение 9.4. *Квадратичная форма $\mathcal{B}$, определенная в n-мерном вещественном линейном пространстве является*

(1) *положительно определенной если и только если $I_+(\mathcal{B}) = n$.*
(2) *отрицательно определенной если и только если $I_-(\mathcal{B}) = n$.*

(3) *знакопеременной если и только если* $I_+(\mathcal{B}) > 0$ *и* $I_-(\mathcal{B}) > 0$.
(4) *неотрицательно определенной если и только если* $I_+(\mathcal{B}) < n$ *и* $I_-(\mathcal{B}) = 0$.
(5) *неположительно определенной если и только если* $I_+(\mathcal{B}) = 0$ *и* $I_-(\mathcal{B}) < n$.

Во всех случаях доказательство утверждений предложения 9.4 непосредственно вытекает из возможности представления квадратичной формы в нормальном виде

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \xi_1^2 + \dots + \xi_p^2 - \xi_{p+1}^2 - \dots - \xi_k^2.$$

Проверим, например, приведенный выше критерий положительной определенности квадратичной формы. Пусть форма $\mathcal{B}$ положительно определена, т.е. $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) \geqslant 0$ для любого вектора $\boldsymbol{x}$ и $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = 0$ только при $\boldsymbol{x} = \boldsymbol{0}$. Из этого непосредственно вытекает, что отрицательных слагаемых в нормальном виде формы $\mathcal{B}$ быть не может. Т.е. $p = k$. Далее, если $k < n$, то можно предъявить вектор $\boldsymbol{x}_1 \neq \boldsymbol{0}$ такой, что $\mathcal{B}(\boldsymbol{x}_1, \boldsymbol{x}_1) = 0$ (достаточно взять ненулевой вектор с нулевыми первыми k координатами). Следовательно, $k = n$ и, окончательно, $p = n$. Обратное утверждение очевидно: форма $\sum_{j=1}^n \xi_j^2$ положительно определена. Проверка остальных четырех утверждений осуществляется аналогично и оставляется в качестве *упражнения*.

Критерий Сильвестра знакоопределенности квадратичной формы. Пусть квадратичная форма $\mathcal{B}$ в некотором базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ задана матрицей $B = (b_{jk})$ и пусть $\Delta_k = \Delta_k(B)$ – это определитель матрицы $(b_{jr} : \ j, r = 1 \dots k)$ при $k = 1, \dots, n$.

Теорема 9.5 (критерий Сильвестра)**.** *Для того чтобы квадратичная форма $\mathcal{B}$ с матрицей B была положительно определенной, необходимо и достаточно, чтобы выполнялись следующие неравенства*

$$\Delta_1(B) > 0, \quad \Delta_2(B) > 0, \quad \dots, \quad \Delta_n(B) > 0. \tag{9.2}$$

Квадратичная форма $\mathcal{B}$ является отрицательно определенной если и только если

$$\Delta_1(B) < 0, \quad \Delta_2(B) > 0, \quad \Delta_3(B) < 0, \quad \Delta_4(B) > 0, \quad \dots, \tag{9.3}$$

т.е. знаки определителей $\Delta_k(B)$ чередуются, причем $\Delta_1(B) < 0$.

Условия (9.2) и (9.3) назовем *условиями Сильвестра.*

Доказательство. Начнем с проверки необходимости условий Сильвестра. Из положительной (или отрицательной) определенности формы $\mathcal{B}$ вытекает, что $\Delta_k(B) \neq 0$ при всех $k = 1, \dots, n$. В самом деле, пусть $\Delta_k(B) = 0$ при каком-то k. Тогда система однородных линейных уравнений

$$\sum_{j=1}^{k} b_{rj} x_j = 0, \qquad r = 1, \dots, k$$

имеет ненулевое решение $(\widetilde{x}_1, \dots, \widetilde{x}_k)$. Так как $\sum_{j,r=1}^k b_{rj} \widetilde{x}_r \widetilde{x}_j = 0$, то $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = 0$ для вектора $\boldsymbol{x}$ с координатами $(\widetilde{x}_1, \dots, \widetilde{x}_k, 0, \dots, 0)$. Так как $\boldsymbol{x} \neq \boldsymbol{0}$, то возникает противоречие с положительной (или отрицательной) определенностью формы $\mathcal{B}$.

Итак, $\Delta_k(B) \neq 0$ при всех $k = 1, \dots, n$. Применяя метод Якоби, находим канонические коэффициенты $\lambda_1, \dots, \lambda_n$ формы $\mathcal{B}$. В силу (9.1) они равны

$$\lambda_1 = \Delta_1(B), \quad \lambda_2 = \Delta_2(B)/\Delta_1(B), \quad \dots, \quad \lambda_n = \Delta_n(B)/\Delta_{n-1}(B).$$

Если квадратичная форма $\mathcal{B}$ – положительно определена, то все канонические коэффициенты положительные. Следовательно,

$$\Delta_1(B) = \lambda_1 > 0, \qquad \Delta_2(B) = \lambda_2 \Delta_1(B) > 0, \qquad \dots, \qquad \Delta_n(B) = \lambda_n \Delta_{n-1}(B) > 0.$$

Если квадратичная форма $\mathcal{B}$ – отрицательно определена, то все канонические коэффициенты отрицательны. Аналогично случаю положительной определенности из формул для канонических коэффициентов вытекает, что $\Delta_1(B) = \lambda_1 < 0$. Далее, $\Delta_2(B) = \lambda_2 \Delta_1(B) > 0$ (как произведение двух отрицательных чисел). И так далее.

Продолжая эту цепочку неравенств получим, что знаки определителей $\Delta_k(B)$ при $k = 1, \ldots, n$ чередуются.

Достаточность условий Сильвестра вытекает из того, что если рассматриваемую квадратичную форму привести в каноническому виду методом Якоби, то в первом случае все канонические коэффициенты будут положительны, а во втором случае – все они будут отрицательны (но в обоих случаях все канонические коэффициенты будут отличны от нуля). □

9.3. Билинейные и квадратичные формы в евклидовом пространстве

В этом разделе мы напомним ряд фактов, установленных для билинейных форм в разделе 3.5. На самом деле, в этом разделе речь шла о свойствах полуторалинейных форм в эрмитовом пространстве, но были сделаны необходимые комментарии и том, как соответствующие свойства будут выглядеть для билинейных форм в евклидовом пространстве.

Первым делом напомним, что для любой билинейной формы $\mathcal{B}$ на евклидовом пространстве $\mathcal{X}$ существует единственный оператор $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$ такой, что $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle$ для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$.

Из Предложения 8.17 вытекает, что билинейная форма $\mathcal{B}$ является симметричной тогда и только тогда, когда оператор $\mathcal{A}$ самосопряженный. В этом случае, из предложений 8.15 и 8.16 вытекает, что все собственные значения $\lambda_1, \ldots, \lambda_n$ оператора $\mathcal{A}$ вещественны и, что существует ортонормированный базис в $\mathcal{X}$, состоящий из собственных векторов оператора $\mathcal{A}$ и в этом базисе матрица оператора $\mathcal{A}$ имеет диагональный вид.

Предложение 9.6. *Пусть $\mathcal{B}$ – симметричная билинейная форма, определенная в вещественном евклидовом пространстве $\mathcal{X}$. Тогда существует ортонормированный базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в $\mathcal{X}$ и набор чисел $\lambda_k \in \mathbb{R}$, $k = 1, \ldots, n$ такие, что*

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \sum_{k=1}^{n} \lambda_k x_k^2,$$

где x_k, $k = 1, \ldots, n$ – координаты вектора $\boldsymbol{x}$ относительно базиса $\boldsymbol{e}$.

Доказательство. Пусть $\mathcal{A}$ – самосопряженный оператор в $\mathcal{X}$ такой, что $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y} \rangle$. Пусть $\lambda_1, \ldots, \lambda_n$ – (вещественные) собственные числа $\mathcal{A}$ и пусть $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ – соответствующий набор ортонормированных собственных векторов (существование $\mathcal{A}$, λ_k и $\boldsymbol{e}_k$ было выяснено выше). Если $\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{e}_j$, то $\mathcal{A}\,\boldsymbol{x} = \sum_{j=1}^{n} \lambda_j x_j \boldsymbol{e}_j$ и, с учетом ортогональности векторов $\boldsymbol{e}_k$ при $k = 1, \ldots, n$ получаем, что $\langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x} \rangle = \sum_{j=1}^{n} \lambda_j x_j^2$. □

Аналогично рассмотренному выше случаю квадратичных форм в эрмитовых пространствах в вещественных евклидовых пространствах имеет место утверждение о возможности одновременного приведения в сумме квадратов двух квадратичных форм.

Предложение 9.7. *Пусть $\mathcal{B}$ и $\mathcal{D}$ – симметричные билинейные формы, определенные в вещественном линейной пространстве $\mathcal{X}$. Предположим, что форма $\mathcal{D}$ является положительно определенной. Тогда в пространстве $\mathcal{X}$ можно выбрать базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, в котором обе соответствующие квадратичные формы имеют канонический вид, т.е.*

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \sum_{k=1}^{n} \lambda_k x_k^2, \qquad \mathcal{D}(\boldsymbol{x}, \boldsymbol{x}) = \sum_{k=1}^{n} x_k^2,$$

где $x_1, \ldots, x_n$ – координаты вектора $\boldsymbol{x}$ относительно базиса $\boldsymbol{e}$, а $\lambda_1, \ldots, \lambda_n$ – некоторые вещественные числа.

Для доказательства этого предложения необходимо заметить, что при сделанных предположениях относительно билинейной формы $\mathcal{D}$ выражение $\mathcal{D}(\boldsymbol{x}, \boldsymbol{y})$, $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ задает скалярное произведение в $\mathcal{X}$ (проверка аксиом скалярного произведения в этом случае оставляется в качестве ***упражнения***). После этого остается применить результат предложения 9.6 для евклидова пространства $(\mathcal{X}, \mathcal{D}(\cdot, \cdot))$. Проверка оставшихся деталей оставляется в качестве ***упражнения***.

Рассмотрим последний результат более подробно. Пусть, как и в Предложении 9.7, $\mathcal{B}$ и $\mathcal{D}$ – две симметричные билинейные формы. Выражение

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) - \lambda\, \mathcal{D}(\boldsymbol{x}, \boldsymbol{x}),$$

где λ – параметр, называется *пучком квадратичных форм*. Если форма $\mathcal{D}$ является положительно определенной, то соответствующий пучок называется *регулярным*.

Предположим, что в некотором ортонормированном пространства $\mathcal{X}$ формы $\mathcal{B}$ и $\mathcal{D}$ имеют матрицы $B = (b_{jk})_{j,k=1}^n$ и $D = (d_{jk})_{j,k=1}^n$ соответственно. В дальнейшем все рассуждения будем вести в этом базисе.

Уравнение

$$\det(B - \lambda D) = 0,$$

называется *характеристическим уравнением пучка* $\mathcal{B} - \lambda\, \mathcal{D}$, а его решения – *характеристическими числами* этого пучка.

Пусть λ_0 – корень уравнения $\det(B - \lambda D) = 0$. Так как матрица $B - \lambda_0 D$ является вырожденной, то найдется такой *ненулевой* столбец $\zeta = (\zeta_1, \ldots, \zeta_n)^\top$, что $B\zeta = \lambda_0 D\zeta$. Вектор $\boldsymbol{z}$, координатный столбец которого в рассматриваемом базисе совпадает со столбцом ζ будем называть *главным вектором пучка* $\mathcal{B} - \lambda\, \mathcal{D}$. Имеет место следующее утверждение.

Предложение 9.8. *Характеристическое уравнение* $\det(B - \lambda D) = 0$ *регулярного пучка* $\mathcal{B} - \lambda\, \mathcal{D}$ *квадратичных форм всегда имеет* n *штук вещественных корней* $\lambda_1, \ldots, \lambda_n$. *Соответствующие этим корням главные векторы* $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ *этого пучка могут быть выбраны так, чтобы выполнялись соотношение*

$$\mathcal{D}(\boldsymbol{z}_j, \boldsymbol{z}_k) = \delta_{j,k}, \quad j, k = 1, \ldots, n.$$

Доказательство. Пусть λ – характеристическое число пучка $\mathcal{B} - \lambda\, \mathcal{D}$, а $\boldsymbol{z}$ – соответствующий главный вектор. Тогда $B[\boldsymbol{z}] = \lambda D[\boldsymbol{z}]$, где $[\boldsymbol{z}]$ – координатный столбец вектора $\boldsymbol{z}$ в рассматриваемом ортонормированном базисе. Из этого вытекает, что $\boldsymbol{z}$ – это *собственный вектор* для матрицы $M = D^{-1}B$, соответствующий собственному значению λ этой матрицы.

Заметим, что матрица M может быть записана в виде

$$M = F^{-1}GF,$$

где $F = \sqrt{D}$ (корень из матрицы D можно извлекать в силу того, что она является симметричной и положительно определенной), а $G = F^{-1}BF^{-1}$. При этом ясно, что матрица G является симметричной.

Из предложений 8.15 и 8.16 вытекает, что матрица G имеет n штук вещественных собственных значений $\lambda_1, \ldots, \lambda_n$ и n штук попарно ортогональных собственных векторов $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ (эти векторы образуют ортонормированный базис в $\mathcal{X}$). Определим векторы $\boldsymbol{z}_k := F^{-1}\boldsymbol{v}_k$, $k = 1, \ldots, n$. Так как матрица F невырождена (проверка этого факта оставляется в качестве ***упражнения***), то система векторов $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$ является линейно независимой. Кроме того, при $k = 1, \ldots, n$ верны равенства

$$M[\boldsymbol{z}_k] = F^{-1}GFF^{-1}[\boldsymbol{v}_k] = \lambda_k[\boldsymbol{z}_k],$$

а, при $j, k = 1, \ldots, n$, равенства

$$\mathcal{D}(\boldsymbol{z}_j, \boldsymbol{z}_k) = [\boldsymbol{z}_j]^\top D[\boldsymbol{z}_k] = [\boldsymbol{z}_j]^\top FF[\boldsymbol{z}_k] = [\boldsymbol{v}_j]^\top[\boldsymbol{v}_k] = \delta j, k.$$

Последнее равенство вытекает из того, что базис $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ – ортонормированный. $\square$

Замечание. В условиях Предложения 9.8 при всех $j, k = 1, \ldots, n$ имеют место равенства:

$$\mathcal{B}(\boldsymbol{z}_j, \boldsymbol{z}_k) = [\boldsymbol{z}_j]^\top B[\boldsymbol{z}_k] = [\boldsymbol{z}_j]^\top \lambda_k B[\boldsymbol{z}_k] = \lambda_k\, \mathcal{D}(\boldsymbol{z}_j, \boldsymbol{z}_k) = \lambda_k \delta_{j,k}.$$

Таким образом, в базисе $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$, составленном из главных векторов пучка квадратичных форм $\mathcal{B} - \lambda\, \mathcal{D}$ матрицы форм $\mathcal{B}$ и $\mathcal{D}$ имеют диагональный вид. Предложение 9.8 дает простой конструктивный метод приведения пары форм к каноническому виду и к отысканию соответствующего преобразования координат. В самом деле, все сводится к отысканию матрица перехода S от исходного ортонормированного базиса (в котором велись все рассуждения) к базису $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_n$.

В качестве ***упражнения*** предлагается проверить, что любой базис, в котором пара квадратичных форм одновременно принимает диагональный вид, состоит из главных векторов соответствующего пучка квадратичных форм.

9.4. Уравнения гиперповерхностей второго порядка, их инварианты и упрощение аффинными преобразованиями

Далее мы будем считать, что $\mathcal{X}$ – евклидово пространство со скалярным произведением $\langle\cdot,\cdot\rangle$, $\dim \mathcal{X} = n$, $n \in \mathbb{N}$.

Общее уравнение гиперповерхности второго порядка. Пусть на пространстве $\mathcal{X}$ заданы билинейная форма $\mathcal{B}$ и линейный функционал $\mathfrak{f}$.

Определение. *Общее уравнение гиперповерхности второго порядка – это уравнение вида*

$$\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) + 2\mathfrak{f}(\boldsymbol{x}) + c = 0, \tag{9.4}$$

где $\boldsymbol{x} \in \mathcal{X}$, а $c \in \mathbb{R}$ – некоторое заданное число.

В рамках данной лекции мы будем изучать, как изменяется уравнение (9.4) при специальных преобразованиях пространства $\mathcal{X}$, которые будем называть *аффинными* (определение см. ниже). Нашей первой целью будет получение ответов на следующие вопросы: какой наиболее простой вид можно придать уравнению (9.4) аффинными преобразованиями и как найти аффинное преобразование, приводящее уравнение (9.4) к такому виду? Вторая цель состоит в отыскание разумной системы инвариантов уравнения (9.4). Инвариант уравнения (9.4) – это такое выражение, связывающее входящие в уравнение (9.4) объекты $\mathcal{B}$, $\mathfrak{f}$ и c, значение которого не меняется при аффинных преобразованиях пространства $\mathcal{X}$. И, наконец, третья цель состоит в построении классификации уравнений вида (9.4). Здесь предполагается разбить все уравнения вида (9.4) на некоторое число классов так, что любое аффинное преобразование преобразовывает уравнение каждого класса в уравнение, принадлежащее этому же классу.

Определим теперь, что такое аффинное преобразование пространства $\mathcal{X}$.

Определение. *Параллельным переносом назовем преобразование пространства $\mathcal{X}$ вида $\boldsymbol{x} \mapsto \boldsymbol{x} + \boldsymbol{v}$, где $\boldsymbol{v}$ – фиксированный вектор из $\mathcal{X}$.*

Линейным преобразованием пространства $\mathcal{X}$ назовем любое преобразование вида $\boldsymbol{x} \mapsto \mathcal{A}\,\boldsymbol{x}$, где $\mathcal{A} \in \mathfrak{L}(\mathcal{X}, \mathcal{X})$. Если $\mathcal{A}$ – обратимый оператор, то соответствующее преобразование называется невырожденным. Если $\mathcal{A}$ – ортогональный оператор, то соответствующее преобразование называется ортогональным.

Преобразование $\Phi : \mathcal{X} \to \mathcal{X}$ назовем аффинным, если Φ есть композиция невырожденных линейных преобразований и параллельных переносов.

Рассмотрим некоторый ортонормированный базис $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ в пространстве $\mathcal{X}$. Пусть $(x_1, \ldots, x_n)^\top = [x]_{\boldsymbol{e}}$ – координаты элемента $\boldsymbol{x}$ в базисе $\boldsymbol{e}$. Пусть матрица $B = (b_{jk})$ – матрица билинейной формы $\mathcal{B}$, т.е. $b_{jk} = \mathcal{B}(\boldsymbol{e}_j, \boldsymbol{e}_k)$ при $j, k = 1, \ldots, n$.

Пусть также $\boldsymbol{f}$ – это такой элемент $\mathcal{X}$, что $\mathfrak{f}(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{f} \rangle$, а $(f_1, \dots, f_n)^\top = [\boldsymbol{f}]_{\boldsymbol{e}}$. Тогда уравнение (9.4) может быть записано в виде

$$\sum_{j,k=1}^{n} b_{jk} x_j x_k + 2 \sum_{k=1}^{n} f_k x_k + c = 0.$$

Определение. *Слагаемое* $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \sum_{j,k=1}^{n} b_{jk} x_j x_k$ *называется группой старших членов уравнения гиперповерхности второго порядка. Слагаемое* $\mathfrak{f}(\boldsymbol{x}) + c = \sum_{k=1}^{n} f_k x_k + c$ *называется линейной частью уравнения гиперповерхности второго порядка.*

Кроме матрицы B будем также рассматривать матрицу

$$D = \begin{pmatrix} b_{11} & \cdots & b_{1n} & f_1 \\ \multicolumn{4}{c}{\dotfill} \\ b_{n1} & \cdots & b_{nn} & f_n \\ f_1 & \cdots & f_n & c \end{pmatrix}.$$

Преобразование уравнения (9.4) при параллельном переносе. Рассмотрим параллельный перенос $\boldsymbol{x} = \boldsymbol{x}' + \widetilde{\boldsymbol{x}}$, где $\widetilde{\boldsymbol{x}}$ – некоторый фиксированный элемент из $\mathcal{X}$. Тогда из $(x_1, \dots, x_n)^\top = [\boldsymbol{x}]$, $(x_1', \dots, x_n') = [\boldsymbol{x}']$ и $(\widetilde{x}_1, \dots, \widetilde{x}_n)^\top = [\widetilde{\boldsymbol{x}}]$ вытекает, что $x_k = x_k' + \widetilde{x}_k$ при всех $k = 1, \dots, n$.

Пусть $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$ – симметричная билинейная форма, полярная квадратичной форме $\mathcal{B}$. Тогда уравнение $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) + 2\mathfrak{f}(\boldsymbol{x}) + c = 0$ вида (9.4) запишется в виде

$$\mathcal{B}(\boldsymbol{x}' + \widetilde{\boldsymbol{x}}, \boldsymbol{x}' + \widetilde{\boldsymbol{x}}) + 2\mathfrak{f}(\boldsymbol{x}' + \widetilde{\boldsymbol{x}}) + c = \mathcal{B}(\boldsymbol{x}', \boldsymbol{x}') + 2(\mathcal{B}(\boldsymbol{x}', \widetilde{\boldsymbol{x}}) + \mathfrak{f}(\boldsymbol{x}')) + (\mathcal{B}(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) + 2\mathfrak{f}(\widetilde{\boldsymbol{x}}) + c).$$

Это уравнение, в свою очередь, может быть записано как уравнение

$$\mathcal{B}(\boldsymbol{x}', \boldsymbol{x}') + 2\mathfrak{f}'(\boldsymbol{x}') + c' = 0$$

вида (9.4), где $\mathfrak{f}'(\boldsymbol{x}') = \mathcal{B}(\boldsymbol{x}', \widetilde{\boldsymbol{x}}) + \mathfrak{f}(\boldsymbol{x}')$ и $c' = \mathcal{B}(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) + 2\mathfrak{f}(\widetilde{\boldsymbol{x}}) + c$.

Непосредственно проверяется, что $\mathcal{B}(\boldsymbol{x}', \boldsymbol{x}') = \sum_{j,k=1}^{n} b_{jk} x_j' x_k'$ (т.е. коэффициенты b_{jk} не меняются). Таким образом, при параллельном переносе группа старших членов уравнения гиперповерхности второго порядка сохраняет свой вид. Далее,

$$\mathcal{B}(\boldsymbol{x}', \widetilde{\boldsymbol{x}}) = \sum_{k=1}^{n} x_k' \sum_{j=1}^{n} b_{jk} \widetilde{x}_j, \quad \mathfrak{f}(\boldsymbol{x}') = \sum_{k=1}^{n} f_k x_k', \quad \mathcal{B}(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{x}}) = \sum_{j,k=1}^{n} b_{jk} \widetilde{x}_j \widetilde{x}_k, \quad \mathfrak{f}(\widetilde{\boldsymbol{x}}) = \sum_{k=1}^{n} f_k \widetilde{x}_k.$$

Отсюда вытекает, что

$$\mathfrak{f}'(\boldsymbol{x}') = \sum_{k=1}^{n} f_k' x_k' = \sum_{k=1}^{n} x_k' \left(\sum_{j=1}^{n} b_{jk} \widetilde{x}_j + f_k \right), \quad c' = \sum_{j,k=1}^{n} b_{jk} \widetilde{x}_j \widetilde{x}_k + 2 \sum_{k=1}^{n} f_k \widetilde{x}_k + c,$$

а используя определение величин $f_k' = \sum_{j=1}^{n} b_{jk} \widetilde{x}_j + f_k$ можно получить несколько более простое и удобное в дальнейшем выражение для коэффициента c':

$$c' = \sum_{k=1}^{n} (f_k' + f_k) \widetilde{x}_k + c.$$

Преобразование уравнения (9.4) при ортогональном преобразовании. Пусть задано некоторое ортогональное преобразование пространства $\mathcal{X}$ и пусть S – матрица этого преобразования. Пусть $\boldsymbol{e}' = \boldsymbol{e}S$. Тогда $[\boldsymbol{x}]_{\boldsymbol{e}} = S[\boldsymbol{x}]_{\boldsymbol{e}'}$ и уравнение (9.4) преобразуется следующим образом

$$\begin{aligned} &\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) + 2\mathfrak{f}(\boldsymbol{x}) + c = [\boldsymbol{x}]_{\boldsymbol{e}}^\top B [\boldsymbol{x}]_{\boldsymbol{e}} + 2[\boldsymbol{x}]_{\boldsymbol{e}}^\top [\boldsymbol{f}] + c = \\ &= (S[\boldsymbol{x}]_{\boldsymbol{e}'})^\top B (S[\boldsymbol{x}]_{\boldsymbol{e}}') + 2(S[\boldsymbol{x}]_{\boldsymbol{e}'})^\top [\boldsymbol{f}] + c = [\boldsymbol{x}]_{\boldsymbol{e}'}^\top (S^{-1} B S) [\boldsymbol{x}]_{\boldsymbol{e}'} + 2[\boldsymbol{x}]_{\boldsymbol{e}'}^\top (S^{-1}[\boldsymbol{f}]) + c, \end{aligned} \tag{9.5}$$

где учтено, что $S^{-1} = S^\top$, так как S – ортогональная матрица.

Заметим, что матрица квадратичной формы при ортогональном преобразовании изменяется как матрица некоторого линейного оператора. Проверим, что оператор, матрица которого в некотором ортонормированном базисе совпадает с матрицей квадратичной формы $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x})$ является самосопряженным. В самом деле, пусть $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$ – симметричная билинейная форма, полярная для квадратичной формы $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x})$. В силу Предложения 8.17 билинейная форма $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y})$ может быть представлена в виде $\mathcal{B}(\boldsymbol{x}, \boldsymbol{y}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{y}\rangle$, где $\mathcal{A}$ – самосопряженный оператор. Следовательно, рассматриваемая квадратичная форма может быть записана в виде $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \langle \mathcal{A}\,\boldsymbol{x}, \boldsymbol{x}\rangle$. Проверим, что в ортонормированном базисе $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ матрицы оператора $\mathcal{A}$ и квадратичной формы $\mathcal{B}$ совпадают. Пусть $(b_{jk})_{j,k=1}^n$ – матрица квадратичной формы $\mathcal{B}$ в базисе $\boldsymbol{e}$ и пусть $(a_{jk})_{j,k=1}^n$ – матрица оператора $\mathcal{A}$ в базисе $\boldsymbol{e}$. По определению, $b_{jk} = \mathcal{B}(\boldsymbol{e}_j, \boldsymbol{e}_k)$. Далее, $\mathcal{A}\,\boldsymbol{e}_j = \sum_{r=1}^n a_{rj}\boldsymbol{e}_r$ и, так как базис $\boldsymbol{e}$ является ортонормированным, то

$$b_{jk} = \mathcal{B}(\boldsymbol{e}_j, \boldsymbol{e}_k) = \langle \mathcal{A}\,\boldsymbol{e}_j, \boldsymbol{e}_k\rangle = \sum_{r=1}^n a_{jr}\langle \boldsymbol{e}_r, \boldsymbol{e}_k\rangle = a_{jk}$$

для всех $j, k = 1, \dots, n$. Итак матрицы квадратичной формы $\mathcal{B}$ и соответствующего самосопряженного оператора $\mathcal{A}$ в базисе $\boldsymbol{e}$ совпадают.

Инварианты уравнения гиперповерхности второго порядка.

Определение. *Назовем инвариантом уравнения (9.4) такую функцию от коэффициентов уравнения (9.4), которая не меняется при аффинных преобразованиях.*

Имеет место следующее утверждение

Теорема 9.9. *Коэффициенты характеристического многочлена матрицы B и, в частности, ее определитель $\det B$ и след $\operatorname{Tr} B$, а также определитель $\det D$ матрицы D являются инвариантами уравнения (9.4).*

Доказательство. Ясно, что нам достаточно проверить инвариантность указанных величин при параллельном переносе и при ортогональном преобразовании по отдельности.

Как уже было установлено, при параллельном переносе матрица B и, следовательно, все ее характеристики не меняются.

Проверим инвариантность величины $\det D$ при параллельном переносе. Пусть при параллельном переносе $\boldsymbol{x} = \boldsymbol{x}' + \widetilde{\boldsymbol{x}}$ матрица D преобразуется в матрицу D' и

$$\det D' = \det\begin{pmatrix} b_{11} & \cdots & b_{1n} & f_1' \\ \multicolumn{4}{c}{\dotfill} \\ b_{n1} & \cdots & b_{nn} & f_n' \\ f_1' & \cdots & f_n' & c' \end{pmatrix} = \det\begin{pmatrix} b_{11} & \cdots & b_{1n} & f_1' \\ \multicolumn{4}{c}{\dotfill} \\ b_{n1} & \cdots & b_{nn} & f_n' \\ f_1 & \cdots & f_n & c + \sum\limits_{k=1}^n f_k\widetilde{x}_k \end{pmatrix}$$

Последнее равенство получается следующим образом: из последней строчки матрицы D' вычитаем первую строку, умноженную на $\widetilde{x}_1$, вторую, умноженную на $\widetilde{x}_2$ и так далее, до предпоследней строки, умноженной на $\widetilde{x}_n$. Напомним, что такие преобразования матрицы не меняют значения ее определителя. Далее, вычитая из последнего столбца уже полученной матрицы первые n столбцов, умноженных, соответственно, на числа $\widetilde{x}_1, \dots, \widetilde{x}_n$ получаем, что $\det D' = \det D$ (проверка оставшихся деталей оставляется в качестве *упражнения*.

Рассмотрим теперь случай ортогонального преобразования. Выше было показано, что при таком преобразовании матрица B изменяется как матрица некоторого линейного оператора и, следовательно, характеристический многочлен матрицы B при ортогональном преобразовании не меняется.

Остается проверить инвариантность величины $\det D$ при ортогональном преобразовании. Предположим, что соответствующее ортогональное преобразование задано (ортогональной) матрицей S. Пусть матрица Q задана следующим образом

$$Q := \begin{pmatrix} S & 0 \\ 0 & 1 \end{pmatrix},$$

где 0 – $n \times 1$ и $1 \times n$ блоки, состоящие из нулей. Как нетрудно проверить, матрица Q является ортогональной (т.е. $Q^{-1} = Q^{\top}$). Следовательно, соответствующее преобразование является ортогональным.

Заметим, что матрица D преобразуется при ортогональном преобразовании с матрицей S так же, как матрица D, рассматриваемая как матрица квадратичной формы, преобразуется при преобразовании с матрицей Q. А так как определитель матрицы квадратичной формы не меняется при ортогональном преобразовании, то величина $\det D$ инвариантна. $\square$

***Следствие* 9.10.** *Величины* $\operatorname{rg} D$ *и* $\operatorname{rg} D$ *также являются инвариантами уравнения гиперповерхности второго порядка (9.4).*

Этот факт непосредственно вытекает из рассуждений, использовавшихся при доказательстве теоремы 9.9.

Упрощение уравнения гиперповерхности второго порядка аффинными преобразованиями.

1. Перейдем непосредственно к упрощению общего уравнения гиперповерхности второго порядка при помощи параллельного переноса и ортогональных преобразований. Первым делом рассмотрим следующий вопрос: найти такой параллельный перенос, при котором линейная часть уравнения будет содержать только свободный член.

Итак, нам необходимо найти параллельный перенос при котором все коэффициенты $f_k' = 0$, $k = 1 \ldots n$. Используя выражения этих коэффициентов через элементы исходной матрицы B и координаты элемента $\widetilde{\boldsymbol{x}}$, задающего параллельный перенос получаем, что

$$\sum_{j=1}^{n} b_{jk}\widetilde{x}_j + f_k = 0, \quad k = 1, \ldots, n.$$

Этот набор уравнений представляет систему линейных уравнений относительно неизвестных $\widetilde{x}_1, \ldots, \widetilde{x}_n$ с матрицей B.

Эта система уравнений называется *уравнениями центра гиперповерхности второго порядка* (или, для краткости, *уравнениями центра*), а элемент $\widetilde{\boldsymbol{x}} = (\widetilde{x}_1, \ldots, \widetilde{x}_n)^{\top}$ – решение уравнений центра (если он существует) – называется *центром гиперповерхности второго порядка*, заданной уравнением (9.4).

Здесь мы сознательно допускаем некоторую терминологическую неточность – мы рассматриваем уравнение (9.4), а не геометрический объект, но в данном случае это упрощает изложение и не приводит к недоразумениям.

Пусть для уравнения (9.4) соответствующая система уравнений центра имеет решение $\widetilde{\boldsymbol{x}}$. Тогда, в результате параллельного переноса $\boldsymbol{x} = \boldsymbol{x}' + \widetilde{\boldsymbol{x}}$ уравнение (9.4) примет вид

$$\sum_{j,k=1}^{n} b_{jk}x_j'x_k' + c' = 0. \tag{9.6}$$

Заметим также, что в случае существования центра $\widetilde{\boldsymbol{x}}$, коэффициент c' уравнения (9.6), где $\boldsymbol{x} = \boldsymbol{x}' + \widetilde{\boldsymbol{x}}$, может быть вычислен по формуле

$$c' = \det D / \det B.$$

Это немедленно вытекает из инвариантности $\det B$ и $\det D$ и их вида матрицы D'.

Определение. *Если уравнения центра имеют единственное решение, то соответствующее уравнение (9.4) называется уравнением центральной гиперповерхности второго порядка. В противном случае уравнение (9.4) называется уравнением нецентральной гиперповерхности второго порядка.*

Итак, уравнение центральной гиперповерхности второго порядка после параллельного переноса $\boldsymbol{x} = \boldsymbol{x}' + \widetilde{\boldsymbol{x}}$ (где $\widetilde{\boldsymbol{x}}$ – центр) запишется в виде (опускаем "штрихи" в уравнении (9.6) для упрощения обозначений)

$$\sum_{j,k=1}^{n} b_{jk}x_jx_k + \frac{\det D}{\det B} = 0. \tag{9.7}$$

Можно заметить, что если $\boldsymbol{x} = (x'_1, \ldots, x'_n)^\top$ удовлетворяет некоторому уравнению центральной гиперповерхности второго порядка, то и $-\boldsymbol{x} = (-x'_1, \ldots, -x'_n)^\top$ удовлетворяет этому же уравнению. Это свойство симметрии и объясняет название элемента $\widetilde{\boldsymbol{x}}$ центром.

2. Как было установлено выше (см. Предложения 8.17, 8.15 и 8.16) для квадратичной формы $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x})$ существует ортонормированный базис $\boldsymbol{e}' = \{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$ в котором $\mathcal{B}(\boldsymbol{x}, \boldsymbol{x}) = \sum_{j=1}^n \lambda_j x'_j$, где $(x'_1, \ldots, x'_n) = [\boldsymbol{x}]_{\boldsymbol{e}'}$ – координатный столбец вектора $\boldsymbol{x}$ относительно $\boldsymbol{e}'$, а $\lambda_1, \ldots, \lambda_n$ – собственные числа (самосопряженного) оператора $\mathcal{A}$, матрица которого в ортонормированном базисе совпадает с матрицей квадратичной формы $\mathcal{B}$.

При переходе от ортонормированного базиса $\boldsymbol{e} = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ к ортонормированному базису $\boldsymbol{e}'$ (напомним, что такой переход осуществляется при помощи ортогонального преобразования) квадратичные и линейные члены уравнения (9.4) гиперповерхности второго порядка преобразуются в соответствии с формулой (9.5) независимо друг от друга. Следовательно, в базисе $\boldsymbol{e}'$ уравнение (9.4) примет вид

$$\sum_{k=1}^{n} \lambda_k {x'_k}^2 + 2\sum_{k=1}^{n} f'_k x'_k + c = 0, \tag{9.8}$$

где $(f'_1, \ldots, f'_n)^\top = S^{-1}\boldsymbol{f}$, а S – соответствующая ортогональная матрица перехода от базиса $\boldsymbol{e}$ к базису $\boldsymbol{e}'$.

Определение. *Преобразование уравнения гиперповерхности второго порядка от вида (9.4) к виду (9.8) традиционно называется стандартным упрощением уравнения (9.4) (более подробно – стандартным упрощением при помощи ортогонального преобразования).*

9.5. Классификация уравнений гиперповерхностей второго порядка

Центральный случай. Если уравнение (9.4) является уравнением *центральной* гиперповерхности второго порядка, то для максимально возможного упрощения уравнения этой поверхности выполним последовательно два действия. Во-первых, найдем решение $\widetilde{\boldsymbol{x}}$ соответствующих уравнений центра и совершим параллельный перенос $\boldsymbol{x} = \boldsymbol{x}' + \widetilde{\boldsymbol{x}}$. При этом уравнение примет вид (9.7) (всюду в дальнейшем мы будем указывать применяемые преобразования и сразу опускать "штрихи" для упрощения обозначений). Далее, применяя стандартное упрощение при помощи ортогонального преобразования приводим уравнение вида (9.7) к виду

$$\sum_{k=1}^{n} \lambda_k x_k^2 + \frac{\det D}{\det B} = 0, \tag{9.9}$$

где все коэффициенты $\lambda_1, \ldots, \lambda_n$ отличны от нуля. В самом деле, матрица квадратичной части уравнения (9.4) в базисе, в котором уравнение гиперповерхности второго порядка имеет вид (9.9), имеет диагональный вид $\mathrm{Diag}\{\lambda_1, \ldots, \lambda_n\}$ и, так как величина $\det B$ – инвариант уравнения (9.4), то $\det B = \lambda_1 \times \cdots \times \lambda_n$. Остается заметить, что так как рассматриваемое уравнение является уравнением центральной гиперповерхности второго порядка, то $\det B \neq 0$.

Предположим, что $\lambda_j > 0$ при $j = 1, \dots, p$ и $\lambda_j < 0$ при $j = p+1, \dots, n$. При этом $1 \leqslant p \leqslant n$ – некоторое число, являющееся инвариантом квадратичной части уравнения (9.4). Выполнения этого предположения всегда можно добиться перестановкой векторов базиса, в котором уравнение (9.4) имеет вид (9.9).

Пусть теперь $\sigma = \operatorname{sgn}\bigl(\det D/ \det B\bigr)$, $\delta = \bigl|\det B/ \det D\bigr|$ и пусть при $k = 1, \dots, n$

$$\alpha_k := \begin{cases} \dfrac{1}{\sqrt{\delta|\lambda_k|}} & \text{при } \sigma \neq 0, \\ \dfrac{1}{\sqrt{|\lambda_k|}} & \text{при } \sigma = 0. \end{cases}$$

Используя эти обозначения можно переписать уравнение (9.9) в виде (мы снова сохраняем обозначение x_k, $k = 1, \dots, n$ для преобразованных координат)

$$\frac{x_1^2}{\alpha_1^2} + \dots + \frac{x_p^2}{\alpha_p^2} - \frac{x_{p+1}^2}{\alpha_{p+1}^2} - \dots - \frac{x_n^2}{\alpha_n^2} + \sigma = 0. \tag{9.10}$$

На основании уравнения (9.10) построим следующую *классификацию* уравнений центральных гиперповерхностей второго порядка. Пусть уравнение гиперповерхности второго порядка в некотором базисе имеет вид (9.10).

1: $p = n$**, а** $\sigma = -1$ **или** $p = 0$**, а** $\sigma = 1$**.** В этом случае уравнение (9.10) называется *уравнением* $(n-1)$*-мерного эллипсоида*. Если при этом $\alpha_1 = \dots = \alpha_n = R$, то соответствующее уравнением называется *уравнением* $(n-1)$*-мерной сферы*.

2: $p = 0$ **и** $\sigma = -1$ **или** $p = n$ **и** $\sigma = 1$**.** В этом случае уравнение (9.10) называется *уравнением мнимой поверхности* или *уравнением мнимого эллипсоида.*.

3: $0 < p < n$ **и** $\sigma \neq 0$**.** В этом случае соответствующее уравнение (9.10) называется *уравнением гиперболоида*.

4: $\sigma = 0$**.** В этом случае уравнение (9.10) называется *уравнением вырожденной поверхности*. В частности, при $p = 0$ или $p = n$ уравнение (9.10) называется *уравнением вырожденного эллипсоида*.

Нецентральный случай. Пусть теперь уравнение (9.4) – это уравнение *нецентральной* гиперповерхности второго порядка. В этом случае $\det B = 0$. Упрощение такого уравнения начнем со стандартного упрощения при помощи ортогонального преобразования. При этом уравнение (9.4) приведем к виду (9.8). Так как $\lambda_1 \times \dots \times \lambda_n = \det B = 0$ (здесь λ_k, $k = 1, \dots, n$ – коэффициенты при квадратах координат в уравнении (9.8)), то по крайней мере один коэффициент λ_k равен нулю. Предположим (переставляя если надо базисные векторы), что $\lambda_1 \neq 0, \dots, \lambda_p \neq 0$ и, что $\lambda_{p+1} = \dots = \lambda_n = 0$ при некотором p, $1 \leqslant p < n$. В этом случае перепишем уравнение (9.8) в виде

$$\sum_{k=1}^{p} \lambda_k {x'_k}^2 + 2\sum_{k=1}^{p} f'_k x'_k + 2\sum_{k=p+1}^{n} f'_k x'_k + c = 0.$$

Выделяя полные квадраты по переменным преобразуем последнее уравнение следующим образом

$$\sum_{k=1}^{p} \lambda_k \left(x'_k + \frac{f'_k}{\lambda_k} \right)^2 + 2\sum_{k=p+1}^{n} f'_k x'_k + \left(c - \sum_{k=1}^{p} \frac{{f'_k}^2}{\lambda_k} \right) = 0.$$

Пусть теперь $x_k = x'_k + f'_k/\lambda_k$ при $k = 1 \dots p$ и $x_k = x'_k$ при $k = p+1 \dots n$ (теперь символы x_k без “штрихов” обозначают новые координаты, отличные от первоначальных, это сделано для упрощения обозначений и, так как все описанные шаги упрощения исходного уравнений нецентральной гиперповерхности второго порядка осуществляются строго последовательно, то путаницы в обозначениях не возникает) последнее уравнение можно записать в виде

$$\sum_{k=1}^{p} \lambda_k x_p^2 + 2 \sum_{k=p+1}^{n} f'_k x_k + c', \tag{9.11}$$

где $c' = c - \sum_{k=1}^{p} f'^2_k / \lambda_k$.

Для дальнейшего упрощения уравнения (9.11) найдем ортогональное преобразование при котором первые p базисных векторов не меняются, а базисные векторы $\boldsymbol{e}_{p+1}, \ldots, \boldsymbol{e}_n$ меняются так, что член $\sum_{k=p+1}^{n} f'_k x_k$ в уравнении (9.11) будет иметь вид $\mu x'_n$ (для новых координат снова использовано обозначение $x'_1, \ldots, x'_n$, эти координаты не надо путать со "штрихованными" координатами, появлявшимися ранее). Если все числа $f'_{p+1}, \ldots, f'_n$ равны нулю, то ничего делать не нужно – достаточно просто положить $\mu = 0$. Пусть теперь хотя бы один из коэффициентов $f'_{p+1}, \ldots, f'_n$ отличен от нуля. В пространстве $\mathcal{Y} = \operatorname{Span}\{\boldsymbol{e}_{p+1}, \ldots, \boldsymbol{e}_n\}$ выражение $\sum_{k=p+1}^{n} f'_k x_k$ представляет собой линейный функционал $\mathfrak{h}(\boldsymbol{y})$. В силу Теоремы 7.15 линейный функционал f может быть единственным образом представлен в виде $\mathfrak{h}(\boldsymbol{y}) = \langle \boldsymbol{h}, \boldsymbol{y} \rangle$, где $\boldsymbol{y} \in \mathcal{Y}$ – произвольный, а $\boldsymbol{h} \in \mathcal{Y}$ – (однозначно определенный функционалом $\mathfrak{h}$) фиксированный вектор. Пусть базисный вектор $\boldsymbol{e}'_n := \boldsymbol{h}/\|\boldsymbol{h}\|$ и пусть $\boldsymbol{e}'_{p+1}, \ldots, \boldsymbol{e}'_{n-1}$ – векторы, дополняющие $\boldsymbol{e}'_n$ до ортонормированного базиса в пространстве $\mathcal{Y}$. Тогда $\boldsymbol{h} = \mu \boldsymbol{e}'_n$, где $\mu = \|\boldsymbol{h}\|$ и, для любого $\boldsymbol{y} \in \mathcal{Y}$

$$\mathfrak{h}(\boldsymbol{y}) = \langle \boldsymbol{h}, \boldsymbol{y} \rangle = \mu \langle \boldsymbol{e}'_n, \boldsymbol{y} \rangle = \mu y_n,$$

где $y_{p+1}, \ldots, y_n$ – координаты вектора $\boldsymbol{y}$ в базисе $\{\boldsymbol{e}'_{p+1}, \ldots, \boldsymbol{e}'_n\}$.

При переходе от ортонормированного базиса $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$, в котором уравнение (9.4) имеет вид (9.11), к базису $\{\boldsymbol{e}'_1, \ldots, \boldsymbol{e}'_n\}$, описанному выше, первые p квадратичных членов уравнения не изменятся (а других квадратичных членов в уравнении (9.11) нет), а (независимо меняющиеся) линейные члены преобразуются так, что для вектора $\boldsymbol{y} = x_{p+1}\boldsymbol{e}_{p+1} + \cdots + x_n \boldsymbol{e}_n$

$$\sum_{k=p+1}^{n} f'_k x_k = \mathfrak{h}(\boldsymbol{y}) = \mu y_n.$$

Таким образом, уравнение (9.11) может быть преобразовано к следующему виду (в котором мы опять меняем обозначения координат на $x_1, \ldots, x_n$ для простоты записи):

$$\sum_{k=1}^{p} \lambda_k x_k^2 + 2\mu x_n + \gamma = 0, \tag{9.12}$$

где $\gamma = c'$ (из формулы (9.11)). На основании уравнения (9.12) построим следующую *классификацию* уравнений нецентральных гиперповерхностей второго порядка.

1: $\mu \neq 0$ и $p = n-1$. Группу младших членов уравнения (9.12) можно преобразовать следующим образом

$$2\mu x_n + \gamma = 2\mu \left(x_n + \frac{\gamma}{2\mu} \right)$$

и, переходя к новой координате $x_n := x_n - \gamma/(2\mu)$ (этот переход осуществляется параллельным переносом), получаем окончательно упрощенный вид уравнения

$$\lambda_1 x_1^2 + \cdots + \lambda_{n-1} x_{n-1}^2 + 2\mu x_n = 0.$$

Уравнение такого вида называется *уравнением параболоида*.

2: $\mu = 0$, а $p < n$. В этом случае уравнение (9.4) запишется в виде

$$\lambda_1 x_1^2 + \dots + \lambda_p x_p^2 + \gamma = 0.$$

Это уравнение в подпространстве $\operatorname{Span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_p\}$ является каноническим уравнением центральной гиперповерхности второго порядка. Таким образом, исходное уравнение (9.4) можно назвать *уравнением центрального цилиндра.*

Используя геометрический язык, можно сказать, что в рассматриваемом случае уравнение (9.4) задает в подпространстве $\operatorname{Span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_p\}$ пространства $\mathcal{X}$ центральную гиперповерхность Σ и, более того, является каноническим уравнением этой гиперповерхности. Таким образом, исходная поверхность Γ является *центральным цилиндром*, *основанием* которого является поверхность Σ, а *образующими* служат гиперплоскости, параллельные $\operatorname{Span}\{\boldsymbol{e}_{p+1}, \dots, \boldsymbol{e}_n\}$.

3: $\mu \neq 0$ и $p < n - 1$. В этом случае уравнение (9.12) имеет вид

$$\lambda_1 x_1^2 + \dots + \lambda_p x_p^2 + 2\mu x_n = 0.$$

В подпространстве $\operatorname{Span}\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_p, \boldsymbol{e}_n\}$ пространства $\mathcal{X}$ это уравнение является уравнением параболоида. Соответственно, исходное уравнение можно назвать *уравнением параболоидального цилиндра.*

Пример 9.11. Рассмотрим уравнение

$$\tfrac{1}{2}x^2 + \tfrac{1}{2}y^2 + \tfrac{1}{2}z^2 + \tfrac{3}{2}u^2 + xz + \sqrt{3}yu - \sqrt{2}x - 2\sqrt{3}y + \sqrt{2}z + 2u - 2v + 1 = 0 \qquad (9.13)$$

в пространстве $\mathbb{R}^5 = \mathbb{R}^5_{x,y,z,u,v}$. Это уравнение является уравнением гиперповерхности второго порядка и записывается в виде (9.4), где

$$B = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{\sqrt{3}}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & 0 & \frac{3}{2} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \boldsymbol{f} = \begin{pmatrix} -\frac{\sqrt{2}}{2} \\ -\sqrt{3} \\ \frac{\sqrt{2}}{2} \\ 1 \\ -1 \end{pmatrix}.$$

Найдем собственные числа матрицы B. Решая уравнение $\det(B - \lambda E) = 0$ получаем:

$$\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = \lambda_4 = \lambda_5 = 0.$$

Находим соответствующие собственные подпространства

$$\begin{aligned} \mathcal{V}_1 &= \operatorname{Span}\{(1, 0, 1, 0, 0)^\top\}, \\ \mathcal{V}_2 &= \operatorname{Span}\{(0, 1, 0, \sqrt{3}, 0)^\top\}, \\ \mathcal{V}_3 &= \operatorname{Span}\{(-1, 0, 1, 0, 0)^\top, (0, -\sqrt{3}, 0, 1, 0)^\top, (0, 0, 0, 0, 1)^\top\}. \end{aligned}$$

Далее находим ортонормированный базис, состоящий из соответствующих собственных векторов

$$\boldsymbol{e}_1 = \tfrac{1}{2}(\sqrt{2}, 0, \sqrt{2}, 0, 0)^\top, \quad \boldsymbol{e}_2 = \tfrac{1}{2}(0, 1, 0, \sqrt{3}, 0)^\top,$$
$$\boldsymbol{e}_{3,1} = \tfrac{1}{2}(-\sqrt{2}, 0, \sqrt{2}, 0, 0)^\top, \quad \boldsymbol{e}_{3,2} = \tfrac{1}{2}(0, -\sqrt{3}, 0, 1, 0)^\top, \quad \boldsymbol{e}_{3,3} = (0, 0, 0, 0, 1)^\top.$$

и выписываем соответствующую матрицу перехода

$$S = \begin{pmatrix} \frac{\sqrt{2}}{2} & 0 & -\frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & -\frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{2}}{2} & 0 & \frac{\sqrt{2}}{2} & 0 & 0 \\ 0 & \frac{\sqrt{3}}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Теперь мы готовы преобразовать (упростить) уравнение (9.13) при помощи ортогонального преобразования с матрицей S. Имеем

$$S^{-1}BS = \operatorname{Diag}\{1, 2, 0, 0, 0\}, \qquad S^{-1}\boldsymbol{f} = (0, 0, 1, 2, -1)^\top,$$

т.е. уравнение (9.13) приведено к виду

$$x^2 + 2y^2 + 2z + 4u - 2v + 1 = 0. \tag{9.14}$$

Рассмотрим пространство $\mathbb{R}^3 = \mathbb{R}^3_{z,u,v}$ и линейный функционал $z + 2u - v$. Вычислим $\|(1, 2, -1)^\top\| = \sqrt{6}$, дополним вектор $(\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, -\frac{\sqrt{6}}{6})^\top$ до ортонормированной системы в $\mathbb{R}^3_{z,u,v}$ и выпишем матрицу Q, столбцы которой – это координатные столбцы соответствующих векторов:

$$Q = \begin{pmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{6} \\ 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{6}}{3} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{3}}{3} & -\frac{\sqrt{6}}{6} \end{pmatrix}.$$

Преобразуем теперь уравнение (9.14) при помощи ортогонального преобразования в $\mathbb{R}^5$ с матрицей $\mathrm{Diag}\{1, 1, Q\}$. Получим уравнение $x^2 + 2y^2 + 2\sqrt{6}v + 1 = 0$, которое при помощи параллельного переноса $v \mapsto v - \frac{1}{2\sqrt{6}}$ приводится к каноническому виду

$$x^2 + 2y^2 + 2\sqrt{6}v = 0.$$

РАЗДЕЛ 10

Дополнительный материал

10.1. Решение проблемы собственных значений методом вращений

Рассмотрим одну из важных вычислительный задач линейной алгебры – так называемую *полную проблему собственных значений* (так называют проблему отыскания всех собственных значений и отвечающих им собственных векторов заданной матрицы).

Все используемые на практике методы решения задач линейной алгебры можно естественно разделить на два класса – *точные методы* и *итерационные методы*. Под точными понимаются методы, которые позволяют получить точные значения требуемых величин в результате вычисления по некоторым известным формулам. Примером точного метода решения упомянутой проблемы собственных значений может быть метод, основанный на непосредственном решении характеристического уравнения (такой метод применим, например, для любой матрицы размера 4×4 и меньше). Итерационные методы позволяют получить значение требуемой величины только в виде предела последовательности величин, построение которых осуществляется в результате некоторого единообразного вычислительного процесса, который называется процессом итераций или последовательных приближений.

Пусть A – симметричная вещественная $n \times n$-матрица, $n \in \mathbb{N}$. В данном разделе мы рассмотрим вопрос о приближенном нахождении собственных чисел матрицы A при помощи достаточно простого приближенного алгоритма, основанного на свойствах ортогональных преобразований. Нашей задачей будет найти такую ортогональную матрицу U, что матрица $U^{-1}AU$ будет диагональной. Так как

$$\det(U^{-1}AU - \lambda E) = \det\big(U^{-1}(A - \lambda E)U\big) = \det(A - \lambda E),$$

то диагональные элементы матрицы $U^{-1}AU$ и будут собственными значениями матрицы A.

Для матрицы $A \in M_n(\mathbb{R})$, $A = (a_{jk})_{j,k=1}^n$, рассмотрим величину $\|A\|_*$, определенную следующим образом:

$$\|A\|_*^2 = \sum_{j,k=1}^n a_{jk}^2.$$

В качестве ***упражнения*** предлагается проверить, что величина $\|A\|_*$ обладает всеми свойствами нормы в пространстве $\mathrm{M}_n(\mathbb{R})$. Кроме того, непосредственно из определения величины $\|A\|_*$ вытекает, что

$$\|A\|_*^2 = \mathrm{Tr}(A^\top A). \tag{10.1}$$

Величину $\|A\|_*$ называют *сферической нормой* матрицы A. Во многих случаях символ $*$ в обозначении сферической нормы опускают и пишут просто $\|A\|$.

Из формулы (10.1) вытекает, в частности, что для любых ортогональных матриц U и V верно равенство

$$\|UAV\|_* = \|A\|_*. \tag{10.2}$$

В самом деле, так как $U^\top U = E$ и $V^\top V = E$, то

$$\begin{aligned}\|UAV\|_*^2 &= \mathrm{Tr}((UAV)^\top(UAV)) = \mathrm{Tr}((AV)^\top U^\top UAV) = \mathrm{Tr}((AV)^\top(AV)) = \\ &= \|AV\|_*^2 = \|VA\|_*^2 = \mathrm{Tr}((VA)^\top(VA)) = \mathrm{Tr}(A^\top V^\top VA) = \mathrm{Tr}(A^\top A) = \|A\|_*^2.\end{aligned}$$

Заметим, что для любой матрицы A справедливо неравенство

$$\sum_{j=1}^{n} a_{jj}^2 \leqslant \|A\|_*^2,$$

которое обращается в равенство если и только если матрица A диагональна. Будем рассматривать преобразования матрицы A вида $A \mapsto T^\top AT$, где T – ортогональная матрица. Такие преобразования будем называть *ортогональными преобразованиями* матрицы A. Предположим, что нам удалось найти такое ортогональное преобразование матрицы A, при котором сумма квадратов ее элементов, лежащих вне главной диагонали, уменьшается. Тогда, так как норма матрицы не меняется при ортогональном преобразовании, то сумма квадратов диагональных элементов будет, наоборот, увеличиваться. Опираясь на это соображение можно построить метод приведения матрицы A в диагональному виду при помощи ортогональных преобразований. Этот метод называется *метод вращений*.

В основе метода вращений лежит использование ортогональных преобразований вида $A \mapsto T^\top AT$, где $T = T_{pq}(\theta) = (\tau_{pq})_{p,q=1}^n$ такая матрица, у которой $\tau_{rr} = 1$ при $r \neq p, q$, $\tau_{pp} = \tau_{qq} = \cos\theta$, $\tau_{pq} = -\sin\theta$ и $\tau_{qp} = \sin\theta$, а остальные элементы равны нулю.

В результате прямых вычислений, проведение которых является простым упражнением на умножение матриц, получаем, что если матрица $A = (a_{jk})_{j,k=1}^n$ преобразуется при помощи ортогонального преобразования $A' := T_{pq}(\theta)^\top A T_{pq}(\theta)$ то для элементов a'_{jk}, $j, k = 1, \ldots, n$ матрицы A' будут верны следующие соотношения, записываемые для удобства и наглядности в матричном виде:

$$a'_{jk} = a_{jk}$$

при $j \neq p, q$ и $k \neq p, q$;

$$\begin{pmatrix} a'_{jp} & a'_{jq} \end{pmatrix} = \begin{pmatrix} a_{jp} & a_{jq} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

при $j \neq p, q$;

$$\begin{pmatrix} a'_{pk} \\ a'_{qk} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_{pk} \\ a_{qk} \end{pmatrix}$$

при $k \neq p, q$ и, наконец,

$$\begin{pmatrix} a'_{pp} & a'_{pq} \\ a'_{qp} & a'_{qq} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

Обозначим для произвольной матрицы $B = (b_{jk})_{j,k=1}^n$ через $\Delta(B)$ величину

$$\Delta(B) = \|B\|_*^2 - \sum_{j=1}^{n} b_{jj}^2,$$

т.е. сумму внедиагональных элементов матрицы B.

Из условия симметричности матрицы A и из приведенных соотношений на элементы матрицы A' вытекает, что

$$\Delta(A') = \Delta(A) - 2a_{pq}^2 + \frac{1}{2}\big((a_{qq} - a_{pp})\sin 2\theta + 2a_{pq}\cos 2\theta\big)^2.$$

Из этой формулы легко увидеть такой способ ортогонального преобразования матрицы A, при котором величина $\Delta(A)$ уменьшается максимально возможным способом. Для этого, во-первых, надо выбрать числа p и q так, чтобы $a_{pq}^2 = \max\{a_{jk}^2 : \ j, k = 1, \ldots, n, \ j \neq k\}$. Затем надо выбрать угол θ так, чтобы выполнялось соотношение

$$(a_{qq} - a_{pp})\sin 2\theta + 2a_{pq}\cos 2\theta = 0,$$

т.е. так, что

$$\operatorname{tg} 2\theta = \frac{2a_{pq}}{a_{pp} - a_{qq}}, \quad |\theta| \leqslant \frac{\pi}{4}.$$

При таком выборе p, q и θ получим, что

$$\Delta(A') = \Delta(A) - 2a_{pq}^2,$$

причем a_{pq} – наибольший по модулю элемент матрицы A.

Пусть теперь последовательно вычисляются матрицы A_m при $m = 0, 1, \ldots$ следующим образом: $A_0 = A$, $A_{m+1} = (A_m)'$ при $m \geqslant 1$, где символ "штрих" обозначает описанное выше ортогональное преобразование соответствующей матрицы. Тогда

$$\Delta(A_{m+1}) = \Delta(A_m) - 2\big(a_{p_m q_m}^{(m)}\big)^2.$$

Так как общее число внедиагональных элементов матрицы A_m равно $n(n-1)$, а $a_{p_m q_m}^{(m)}$ – максимальный по модулю внедиагональный элемент матрицы A_m, то имеет место неравенство

$$\big(a_{p_m q_m}^{(m)}\big)^2 \geqslant \frac{\Delta(A_m)}{n(n-1)}$$

Таким образом, имеет место оценка

$$\Delta(A_{m+1}) \leqslant \Delta(A_m)\left(1 - \frac{2}{n(n-1)}\right).$$

Следовательно, мы приходим к неравенству

$$\Delta(A_{m+1}) \leqslant \Delta(A)\left(1 - \frac{2}{n(n-1)}\right)^{m+1}$$

из которого вытекает, что

$$\lim_{m\to\infty} \Delta(A_{m+1}) = 0.$$

Таким образом, сумма квадратов внедиагональных элементов матриц A_m последовательно уменьшается при $m \to \infty$ и, следовательно, эта последовательность матриц сходится к диагональной матрице. Так как при ортогональных преобразованиях вида $A \mapsto A'$ собственные значения не меняются, то диагональные элементы (диагональной) матрицы, предельной для последовательности (A_m) являются собственными числами исходной матрицы A.

10.2. Псевдообратная матрица

Если A – невырожденная квадратная матрица, то для нее существует обратная матрица A^{-1} такая, что $AA^{-1} = A^{-1}A = E$. Однако понятие обратной матрицы не имеет смысла для не квадратных и для квадратных, но вырожденных матриц. В этом разделе мы покажем, что для произвольной прямоугольной матрицы A существует матрица A^+, которая обладает определенными свойствами обратной матрицы. Эта матрица A^+ будет названа *псевдообратной*. При этом для квадратной невырожденной матрицы псевдообратная матрица будет совпадать с обратной.

В этом разделе через $\mathrm{M}_{m,n}(\mathbb{R})$ мы обозначим совокупность всех вещественных $m\times n$-матриц, а через $\mathrm{M}_{m,n}(\mathbb{C})$ – совокупность всех комплексных $m \times n$-матриц, $m, n \in \mathbb{N}$.

Скелетное разложение матриц. Пусть $A \in \mathrm{M}_{m,n}(\mathbb{C})$ и пусть $\operatorname{rg} A = r$. Тогда найдутся такие матрицы $B \in \mathrm{M}_{m,r}(\mathbb{C})$ и $C \in \mathrm{M}_{r,n}(\mathbb{C})$, что $\operatorname{rg} B = \operatorname{rg} C = r$ и $A = BC$.

В самом деле, выберем r линейно независимых столбцов матрицы A и составим из этих столбцов матрицу B. Заметим, что любой столбец матрицы A выражается как линейная комбинация столбцов матрицы B. Составим матрицу C следующим образом. Возьмем k-ый столбец матрицы A и представим его в виде линейной комбинации столбцов матрицы B с коэффициентами $c_{1k}, \ldots, c_{rk}$; эти коэффициенты и составят k-ый столбец матрицы C. Например, для матрицы

$$A = \begin{pmatrix} 1 & -1 & 2 & 0 \\ -1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix}$$

получаем: $\operatorname{rg} A = 2$ и

$$B = \begin{pmatrix} 1 & -1 \\ -1 & 2 \\ 0 & 1 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & -1 & 1 \end{pmatrix}.$$

Заметим, что квадратные $r \times r$-матрицы B^*B и CC^* являются невырожденными. В самом деле, если $\boldsymbol{x} \in \mathbb{C}^r$ такой вектор, что $B^*B\boldsymbol{x} = \mathbf{0}$, то

$$(B\boldsymbol{x})^*(B\boldsymbol{x}) = \boldsymbol{x}^*B^*B\boldsymbol{x} = \boldsymbol{x}^*\mathbf{0} = \mathbf{0},$$

откуда $B\boldsymbol{x} = \mathbf{0}$. Так как выражение $B\boldsymbol{x}$ – это линейная комбинация столбцов матрицы B, то $\boldsymbol{x} = 0$. Итак, система уравнений $B^*B\boldsymbol{x} = \mathbf{0}$ имеет только нулевое решение. Следовательно, $\det B \neq 0$. Аналогично проверяется невырожденность матрицы CC^*.

Построенное таким образом разложение матрицы A в виде $A = BC$ называется *скелетным разложением* матрицы A.

Существование и основные свойства псевдообратной матрицы. Для матрицы $A \in \mathrm{M}_{m,n}(\mathbb{C})$ рассмотрим матричное уравнение

$$AXA = A.$$

Ясно, что если A – квадратная невырожденная матрица, то это уравнение имеет единственное решение $X = A^{-1}$. В случае произвольной прямоугольной матрицы A решение X этого уравнения – это матрица размера $n \times m$ и, в общем случае, уравнение $AXA = A$ имеет бесконечно много решений. Оказывается, что среди этих решений существует только одно такое, что его строки и столбцы являются линейными комбинациями строк и столбцов исходной матрицы A. Именно это решение и будет псевдообратной матрицей для матрицы A. Перейдем к точным определениям.

Определение. *Пусть $A \in \mathrm{M}_{m,n}(\mathbb{C})$. Матрица $A^+ \in \mathrm{M}_{n,m}(\mathbb{C})$ называется псевдообратной матрицей для A, если*

$$AA^+A = A, \tag{10.3}$$

и, кроме того,

$$\exists\, U, V \in \mathrm{M}_m(\mathbb{C}) \quad \textit{такие, что} \quad A^+ = UA^* = A^*V. \tag{10.4}$$

Условие (10.4) в точности означает, что строки и столбцы матрицы A^+ являются линейными комбинациями строк и столбцов матрицы A.

Теорема 10.1. *Пусть $A \in \mathrm{M}_{m,n}(\mathbb{C})$, $n, m \in \mathbb{N}$ и пусть $A = BC$ – скелетное разложение для A. Тогда*

1. *Псевдообратная матрица A^+ определяется соотношениями (10.3) и (10.4) единственным образом.*

2. $A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$.

Доказательство. Докажем вначале единственность псевдообратной матрицы.

Предположим, что существуют две матрицы A_1^+ и A_2^+, удовлетворяющие условиям (10.3) и (10.4). Обозначим через $U_{1,2}$ и $V_{1,2}$ матрицы U и V из условия (10.4) для матриц A_1^+ и A_2^+ соответственно. Пусть

$$D := A_1^+ - A_2^+, \qquad U := U_1 - U_2, \qquad V := V_1 - V_2.$$

Тогда $ADA = A(A_1^+ - A_2^+)A = 0$, а $D = UA^* = A^*V$. Следовательно,

$$(DA)^*(DA) = A^*D^*DA = A^*V^*ADA = 0,$$

откуда $DA = 0$. Так $DD^* = ADU^* = 0$, то $DD^* = 0$ и, окончательно, $D = 0$.

Перейдем к доказательству второго утверждения теоремы. Непосредственно проверяем, что если $X = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$, то

$$AXA = BC^*(CC^*)^{-1}(B^*B)^{-1}B^*C = BC = A.$$

Кроме того, если определить матрицу $Y = (CC^*)^{-1}(B^*B)^{-1}$, то

$$X = C^*YB^* = \big(C^*Y(CC^*)^{-1}C\big)A^* = A^*\big(B(B^*B)^{-1}YB^*\big).$$

Отсюда следует, что матрица X удовлетворяет условиям (10.3) и (10.4). Таким образом, с учетом первого утверждения теоремы, $X = A^+$. □

Интересно отметить, что матрица $(B^*B)^{-1}B^*$ является псевдообратной матрицей для матрицы B, и матрица $C^*(CC^*)^{-1}$ является псевдообратной для матрицы C. Проверка этого факта оставляется в качестве ***упражнения***. Заметим, что при доказательстве Теоремы 10.1 установлено, что для скелетного разложения $A = BC$ справедливо также равенство $A^+ = C^+B^+$.

Необходимо также заметить, что скелетное разложение $A = BC$ не определяется однозначно. Однако из Теоремы 10.1 вытекает, что любое скелетное разложение матрицы A приведет в одной и той же псевдообратной матрице A^+.

В качестве упражнения предлагается также проверить, что для квадратной невырожденной матрицы A верно равенство $A^+ = A^{-1}$.

Предложение 10.2. *Пусть* $A \in \mathrm{M}_{m,n}(\mathbb{C})$, $m, n \in \mathbb{N}$. *Тогда*

1. $(A^*)^+ = (A^+)^*$.
2. $(A^+)^+ = A$.
3. *Матрица* AA^+ *является эрмитовой и* $(AA^+)^2 = AA^+$.
4. *Матрица* A^+A *является эрмитовой и* $(A^+A)^2 = A^+A$.

Для доказательства первого утверждения достаточно заметить, что если $A = BC$ – скелетное разложение матрицы A, то $A^* = C^*B^*$ – скелетное разложение матрицы A^*. Все остальные утверждения Предложения 10.2 доказываются непосредственными матричными вычислениями. Это предлагается проделать в качестве ***упражнения***.

Необходимо заметить, что если представление матрицы $A = B_1C_1$ не является ее скелетным разложением, то равенство $A^+ = C_1^+B_1^+$ может не иметь места. Пусть, например,

$$A = \begin{pmatrix} 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = BC.$$

Тогда

$$A^+ = A^{-1} = \begin{pmatrix} 1 \end{pmatrix}, \quad B^+ = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad C^+ = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad C^+B^+ = \begin{pmatrix} \frac{1}{2} \end{pmatrix} \neq A^+.$$

Экстремальное свойство псевдообратной матрицы. Пусть $A \in \mathrm{M}_{m,n}(\mathbb{C})$, а $\boldsymbol{b} \in \mathbb{C}^m$. В этом разделе через $\|\cdot\|$ будет обозначена стандартная норма в $\mathbb{C}^k$, значение k будет ясно из контекста. Рассмотрим систему линейных уравнений

$$A\boldsymbol{x} = \boldsymbol{b}, \qquad \boldsymbol{x} \in \mathbb{C}^n \tag{10.5}$$

которая, в общем случае, может быть несовместной. Нам потребуется ввести следующее понятие.

Определение. *Вектор* $\boldsymbol{x}_0$ *называется наилучшим приближенным решением системы (10.5), если выполняются следующие два условия:*

(1) $\|\boldsymbol{b} - A\boldsymbol{x}_0\| = \inf_{\boldsymbol{x}\in\mathbb{C}^n} \|\boldsymbol{b} - A\boldsymbol{x}\|$ *и*
(2) $\|\boldsymbol{x}_0\| = \inf \|\boldsymbol{x}\|$, *где* $\inf$ *берется по всем векторам* $\boldsymbol{x} \in \mathbb{C}^n$ *таким, что*

$$\|\boldsymbol{b} - A\boldsymbol{x}\| = \inf_{\boldsymbol{y}\in\mathbb{C}^n} \|\boldsymbol{b} - A\boldsymbol{y}\|.$$

Другими словами, вектор $\boldsymbol{x}_0$ – наилучшее приближенное решение системы (10.5), если при $\boldsymbol{x} = \boldsymbol{x}_0$ величина *квадратичного отклонения* $\|\boldsymbol{b} - A\boldsymbol{x}\|$ достигает своего наименьшего значения и среди всех векторов $\boldsymbol{x}$, для которых это отклонение имеет минимальное значение, вектор $\boldsymbol{x}_0$ имеет наименьшую норму.

Теорема 10.3. *Для любой матрицы* $A \in \mathrm{M}_{m,n}(\mathbb{C})$ *и для любого вектора* $\boldsymbol{b} \in \mathbb{C}^m$ *наилучшее приближенное решение* $\boldsymbol{x}_0$ *системы (10.5) существует и равно* $\boldsymbol{x}_0 = A^+\boldsymbol{b}$.

Доказательство. Пусть $\boldsymbol{x}_0 = A^+\boldsymbol{b}$. Для любого $\boldsymbol{x} \in \mathbb{C}^n$ рассмотрим разность

$$\boldsymbol{b} - A\boldsymbol{x} = \boldsymbol{b} - A\boldsymbol{x}_0 + A\boldsymbol{x}_0 - A\boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v},$$

где $\boldsymbol{u} = \boldsymbol{b} - A\boldsymbol{x}_0 = \boldsymbol{b} - AA^+\boldsymbol{b}$, а $\boldsymbol{v} = A(\boldsymbol{x}_0 - \boldsymbol{x})$.

Вычислим матричное произведение $\boldsymbol{v}^*\boldsymbol{u}$:

$$\boldsymbol{v}^*\boldsymbol{u} = (\boldsymbol{x}_0 - \boldsymbol{x})^*A^*(\boldsymbol{b} - AA^+\boldsymbol{b}) = (\boldsymbol{x}_0 - \boldsymbol{x})^*(A^* - A^*AA^+)\boldsymbol{b}.$$

Пусть $A = BC$ – скелетное разложение матрицы A. Тогда

$$A^*AA^+ = C^*B^*BCC^*(CC^*)^{-1}(B^*B)^{-1}B^* = C^*B^* = A^*,$$

откуда $\boldsymbol{v}^*\boldsymbol{u} = 0$. Следовательно, $\boldsymbol{u}^*\boldsymbol{v} = (\boldsymbol{v}^*\boldsymbol{u})^* = 0$. Из этих равенств получаем, что

$$\|\boldsymbol{b} - A\boldsymbol{x}\|^2 = (\boldsymbol{u} + \boldsymbol{v})^*(\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{u}^*\boldsymbol{u} + \boldsymbol{u}^*\boldsymbol{v} + \boldsymbol{v}^*\boldsymbol{u} + \boldsymbol{v}^*\boldsymbol{v} = \\ = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 = \|\boldsymbol{b} - A\boldsymbol{x}_0\|^2 + \|\boldsymbol{v}\|^2 \geqslant \|\boldsymbol{b} - A\boldsymbol{x}_0\|^2.$$

Итак, вектор $\boldsymbol{x}_0 = A^+\boldsymbol{b}$ минимизирует квадратичное отклонение. Осталось проверить, что он имеет минимальную норму среди всех таких векторов. Пусть теперь $\boldsymbol{x} \in \mathbb{C}^n$ такой вектор, что

$$\|\boldsymbol{b} - A\boldsymbol{x}\| = \inf_{\boldsymbol{y}\in\mathbb{C}^n} \|\boldsymbol{b} - A\boldsymbol{y}\|$$

и пусть $\boldsymbol{w} = \boldsymbol{x} - \boldsymbol{x}_0$. Так как

$$\|\boldsymbol{b} - A\boldsymbol{x}\|^2 = \|\boldsymbol{b} - A\boldsymbol{x}_0\|^2 + \|A\boldsymbol{w}\|^2,$$

и так как $\|\boldsymbol{b} - A\boldsymbol{x}\| = \|\boldsymbol{b} - A\boldsymbol{x}_0\|$, то $A\boldsymbol{w} = 0$. Так как, по определению псевдообратной матрицы, $A^+ = A^*V$ для некоторой матрицы $V \in \mathrm{M}_m(\mathbb{C})$, то

$$\boldsymbol{x}_0^*\boldsymbol{w} = (A^*\boldsymbol{b})^*\boldsymbol{w} = (A^*V\boldsymbol{b})^*\boldsymbol{w} = \boldsymbol{b}^*V^*A\boldsymbol{w} = 0$$

и $\boldsymbol{w}^*\boldsymbol{x}_0 = (\boldsymbol{w}_0^*\boldsymbol{w})^* = 0$. Из этого, аналогично предыдущему рассуждению, получаем, что

$$\|\boldsymbol{x}\|^2 = \|\boldsymbol{x}_0 + \boldsymbol{w}\|^2 = \|\boldsymbol{x}_0\|^2 + \|\boldsymbol{w}\|^2 \geqslant \|\boldsymbol{x}_0\|^2,$$

причем равенство имеет место если и только если $\boldsymbol{w} = 0$, т.е. если и только если $\boldsymbol{x} = \boldsymbol{x}_0$. □

Замечание. Пусть $A \in \mathrm{M}_{m,n}(\mathbb{C})$ и $G \in \mathrm{M}_{m,k}(\mathbb{C})$ – произвольные матрицы. Для матричного уравнения

$$AX = G$$

понятие наилучшего приближенного решения $X_0 \in \mathrm{M}_{n,k}(\mathbb{C})$ вводится аналогично понятию наилучшего приближенного решения для системы (10.5) (только вместо нормы вектора используется уже знакомая нам сферическая норма матрицы). Аналогично тому, как это проделано при доказательстве Теоремы 10.3 устанавливается, что для матричного уравнения $AX = G$ всегда существует наилучшее приближенное решение X_0, выражающееся формулой

$$X_0 = A^+G.$$

РАЗДЕЛ 11

Программа и задачи к экзамену

11.1. Программа экзамена

1. Понятие отображения. Инъективные, сюръективные и биективные отображения. Композиция отображений и ее свойства. Ассоциативность композиции отображений. Обратное отображение и его свойства. Критерии обратимости отображений.

2. Бинарные отношения и отношения эквивалентности. Свойства отношений эквивалентности. Понятие фактормножества и канонической проекции на фактормножество. Фактоотображения и их свойства.

3. Вещественные и комплексные линейные пространства и их свойства. Линейная зависимость и линейная независимость систем векторов в линейных пространствах. Свойства линейно зависимых и линейно независимых систем. Критерий линейной независимости.

4. Базис и размерность линейного пространства. Свойства базиса. Связь размерности линейного пространства с количеством элементов базиса. Координаты элемента линейного пространства относительно базиса и их свойства. Изоморфизм линейных пространств и его свойства. Теорема о том, что любые два конечномерных линейных пространства, имеющие одинаковую размерность, изоморфны.

5. Подпространства линейного пространства, линейные оболочки систем векторов. Сумма и пересечение подпространств и их свойства. Нахождение базиса в сумме и пересечении подпространств, заданных как линейные оболочки систем векторов. Теоремы о размерности суммы подпространств и о разложении пространства в прямую сумму подпространств. Внутренние и внешние прямые суммы. Факторпространства линейных пространств и их свойства.

6. Матрица перехода от одного базиса к другому в линейном пространстве. Формулы преобразования координат вектора при замене базиса.

7. Линейные операторы и их основные свойства. Критерий обратимости линейного оператора. Ядро, образ, ранг и дефект линейного оператора и их свойства. Неравенства для ранга и дефекта произведения линейных операторов.

8. Матрица линейного оператора и ее свойства. Формула преобразования матрицы линейного оператора при замене базиса.

9. Линейные функционалы и их координаты. Формула преобразования координат линейного функционала при замене базиса. Полуторалинейные и билинейные функции (формы), их свойства. Матрицы полуторалинейных и билинейных форм и их преобразование при замене базиса.

10. Двойственное пространство. Сопряженный базис и его свойства. Существование канонического изоморфизма $\mathcal{X} \to \mathcal{X}^{**}$ в случае конечномерного пространства $\mathcal{X}$. Критерий линейной зависимости и интерпретация решений систем линейных уравнений в терминах сопряженного пространства. Сопряженный оператор и его свойства.

11. Инвариантные подпространства линейного оператора. Минимальный многочлен линейного оператора и его свойства. Характеристический многочлен линейного оператора, собственные числа и собственные векторы линейных операторов, их основные свойства. Теорема Гамильтона-Кэли. Циклические векторы линейных операторов.

12. Собственные и корневые подпространства линейных операторов. Их свойства. Теорема о приведении линейного оператора к каноническому виду. Определение структуры матрицы линейного оператора в канонической (жордановой форме).

13. Скалярное произведение, способы его задания и свойства. Евклидовы и эрмитовы пространства и их свойства. Неравенство Коши-Буняковского в вещественном и комплексном случаях.

14. Норма вектора и ее свойства, понятие нормированного пространства. Норма в евклидовом и эрмитовом пространствах и угол между векторами в евклидовом пространстве. Норма линейного оператора в евклидовом и эрмитовом пространствах.

15. Ортонормированные базисы в евклидовом пространстве и их свойства. Ортогональные матрицы. Матрица Грамма и ее свойства. Алгоритм ортогонализации Грама-Шмидта. Ортогональное дополнение подпространства в евклидовом пространстве. Разложении евклидова (эрмитова) пространства в прямую сумму подпространства и его ортогонального дополнения. Изоморфизм евклидовых (эрмитовых) пространств и его свойства.

16. Теоремы об общем виде линейных функционалов и полуторалинейных (билинейных) функций в эрмитовом (евклидовом) пространствах.

17. Эрмитово сопряженные и самосопряженные операторы в эрмитовом (евклидовом) пространстве. Вид матрицы самосопряженного оператора в евклидовом и эрмитовом пространствах. Свойства собственных чисел и собственных векторов самосопряженного оператора. Минимаксное свойстве собственных чисел самосопряженного оператора. Спектральное разложение самосопряженного оператора.

18. Нормальные операторы и их основные свойства. Критерий нормальности оператора в n-мерном евклидовом пространстве. Спектральное разложение нормальных операторов. Унитарные и ортогональные операторы и их свойства. Комплексификация вещественных евклидовых пространств и линейных операторов. Канонический вид унитарных и ортогональных операторов.

19. Квадратичные формы. Канонический вид квадратичной формы. Методы Лагранжа и Якоби приведения квадратичной формы к каноническому виду. Приведение пары квадратичных форм к сумме квадратов. Индекс инерции квадратичной формы, положительный и отрицательный индексы инерции квадратичной формы. Закон инерции квадратичных форм. Положительно определенные и отрицательно определенные квадратичные формы, знакоопределенные и знакопеременные квадратичных формы. Критерий Сильвестра положительной и отрицательной определенности квадратичной формы.

20. Общее уравнение гиперповерхности второго порядка и его преобразование при помощи аффинных преобразований. Инварианты общего уравнения гиперповерхности второго порядка. Классификация центральных и нецентральных уравнений гиперповерхностей второго порядка.

11.2. Задачи к экзамену

1. Пусть линейный оператор $\mathcal{A}$ действует в пространстве $M_2(\mathbb{R})$ одним из следующих способов:

$$\mathcal{A}(X) = X\begin{pmatrix} a & c \\ -c & a \end{pmatrix}, \quad \mathcal{A}(X) = X\begin{pmatrix} a & 1 \\ 0 & a \end{pmatrix} - \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}X, \quad \mathcal{A}(X) = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}X$$

где $X \in M_2(\mathbb{R})$. Найти матрицу оператора $\mathcal{A}$ в базисе из матричных единиц, собственные и корневые инвариантные подпространства оператора $\mathcal{A}$.

2. Проверить, что система Φ линейных функционалов

$$f_1(X) = x + y, \quad f_2(X) = x - y, \quad f_3(X) = z, \quad f_4(X) = w, \qquad X = \begin{pmatrix} x & y \\ z & w \end{pmatrix} \in M_2(\mathbb{R})$$

образует базис в пространстве $M_2(\mathbb{R})^*$ и найти в $M_2(\mathbb{R})$ базис $\boldsymbol{e}$, сопряженный базису Φ в $M_2(\mathbb{R})^*$. Выписать матрицу линейного оператора $\mathcal{B} : X \mapsto XA$, где $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R})$ заданная матрица, в базисе $\boldsymbol{e}$ и найти $\mathcal{B}^*\varphi$, где $\varphi(X) = \operatorname{Tr} X$.

3. Проверить, что система Φ линейных функционалов

$$f_1 = P(0), \quad f_2 = P'(0), \quad f_3(P) = P(-1), \quad f_4(P) = P(1), \qquad P \in \mathbb{R}[t]_3,$$

образует базис в пространстве $\mathbb{R}[t]_3^*$ и найти в пространстве $\mathbb{R}[t]_3$ базис $\boldsymbol{e}$, сопряженный базису Φ в $\mathbb{R}[t]_3^*$. Выписать матрицу линейного оператора $\mathcal{B} = 2\frac{d}{dt} - \mathcal{E}$ в базисе $\boldsymbol{e}$ и найти $\mathcal{B}^*\varphi$, где $\varphi(P) = P(-1) + 4P(0) + P(1)$.

4. В пространстве функций $\operatorname{Span}\{\cos t, \sin t, \ldots, \cos(nt), \sin(nt)\}$, $n \in \mathbb{N}$, найти все подпространства, инвариантные относительно оператора

$$\mathcal{A}(f) = \int_{-t}^{t} f(s)\, ds.$$

5. В пространстве функций $\operatorname{Span}\{\cos t, \sin t, \ldots, \cos(nt), \sin(nt)\}$, $n \in \mathbb{N}$, найти все подпространства, инвариантные относительно оператора

$$\mathcal{A}(f) = \frac{df}{dt}.$$

6. Для произвольного многочлена $P \in \mathbb{R}[t]$ определим многочлен $P_{a,b}$, $a, b \in \mathbb{R}$, $b \neq \pm 1$, при помощи соотношения $P_{a,b}(t) := P(a - bt)$. Найти канонический вид линейного оператора $\mathcal{A}$, который действует в пространстве $\mathbb{R}[t]_n$, $n \in \mathbb{N}$, следующим образом $\mathcal{A}(P) = P_{a,b}$.

7. В пространстве $\mathbb{R}[t, s]_2$ многочленов $P(t, s)$ от двух переменных степени не выше 2 действует оператор

$$\mathcal{A} : P \mapsto \widetilde{P}, \qquad \text{где } \widetilde{P}(t, s) = P(t - 1, s - 1).$$

Найти канонический вид оператора $\mathcal{A}$.

8. В пространстве $\mathbb{R}[t, s]_2$ многочленов от двух переменных t и s степени не выше 2 найти канонический вид оператора

$$\mathcal{A} = \frac{1}{2}\frac{\partial}{\partial t} - \frac{1}{3}\frac{\partial}{\partial s}$$

и преобразование, приводящее этот оператор к каноническому виду

9. Проверить, что выражение $Q(X) := \operatorname{Tr}(X\overline{X})$ задает на пространствах $M_n(\mathbb{R})$ и $M_n(\mathbb{C})$ квадратичные формы. Определить ранг этих форм и их положительные и отрицательные индексы инерции.

10. Пусть $\boldsymbol{e} = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_k\}$ — базис подпространства $\mathcal{W}$ евклидова пространства $\mathcal{X}$. Доказать, что ортогональная проекция произвольного вектора $\boldsymbol{x} \in \mathcal{X}$ на $\mathcal{W}$ равна сумме его проекций на подпространства $\{\alpha \boldsymbol{e}_j : \alpha \in \mathbb{R}\}$, $j = 1, \dots, k$, если и только если базис $\boldsymbol{e}$ ортогональный.

11. Пусть для любого вектора $\boldsymbol{x}$ евклидова пространства $\mathcal{X}$ сумма его ортогональных проекций на подпространства $\mathcal{X}_1 \subset \mathcal{X}$ и $\mathcal{X}_2 \subset \mathcal{X}$ равна его ортогональной проекции на $\mathcal{X}_1 + \mathcal{X}_2$. Доказать, что $\mathcal{X}_1 \perp \mathcal{X}_2$.

12. Пусть $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ – базис подпространства $\mathcal{U}$ вещественного евклидова пространства $\mathcal{X}$. Доказать, что ортогональная проекция произвольного вектора $\boldsymbol{x} \in \mathcal{X}$ на $\mathcal{U}$ равна сумме его проекций на одномерные подпространства, натянутые на $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$, тогда и только тогда, когда базис $\boldsymbol{u}_1, \dots, \boldsymbol{u}_k$ ортогональный

13. Пусть $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ — ортонормированная система векторов в n-мерном эрмитовом пространстве $\mathcal{X}$. Доказать, что для любого вектора $\boldsymbol{x} \in \mathcal{X}$ верно неравенство

$$\sum_{k=1}^{m} |\langle \boldsymbol{x}, \boldsymbol{e}_k \rangle|^2 \leqslant \|\boldsymbol{x}\|^2,$$

причем равенство достигается для любого $\boldsymbol{x}$ тогда и только тогда, когда $k = n$.

14. Доказать, что квадрат расстояния от вектора $\boldsymbol{x}$ эрмитова пространства до подпространства с базисом $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$ равен отношению определителей Грама систем векторов $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m, \boldsymbol{x}\}$ и $\{\boldsymbol{e}_1, \dots, \boldsymbol{e}_m\}$.

15. Пусть $\mathcal{W}$ – k-мерное подпространство n-мерного вещественного евклидова пространства $\mathcal{X}$ и пусть $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ – ортонормированный базис в $\mathcal{X}$. Рассмотрим векторы $\boldsymbol{e}'_1, \dots, \boldsymbol{e}'_n$ – ортогональные проекции векторов $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ на $\mathcal{W}$. Доказать, что $\sum_{j=1}^{n} \|\boldsymbol{e}'_j\|^2 = k$.

16. Пусть $\boldsymbol{e}_1, \dots, \boldsymbol{e}_n$ – ортонормированный базис в вещественном евклидовом пространстве $\mathcal{X}$, а система векторов $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ этого пространства такова, что $\sum_{j=1}^{n} \|\boldsymbol{e}_j - \boldsymbol{w}_j\| < 1$. Доказать, что система векторов $\boldsymbol{w}_1, \dots, \boldsymbol{w}_n$ линейно независима.

17. Доказать, что всякий линейный функционал f на пространстве $M_n(\mathbb{R})$ (или на пространстве $M_n(\mathbb{C})$) имеет вид $f(X) = \mathrm{Tr}(AX)$, где $A = A_f \in M_n(\mathbb{R})$ (или $A = A_f \in M_n(\mathbb{C})$) – некоторая однозначно определенная матрица.

18. Пусть $Q \in \mathbb{R}[t]$ – некоторый фиксированный многочлен. Какие из следующий выражений определяют линейные функционалы на пространстве $\mathbb{R}[t]_n$, $n \in \mathbb{N}$:

(1) $P \mapsto \int_{-2}^{1} P(t)Q(t)\,dt$,
(2) $P \mapsto \int_{0}^{1} P(t)Q(t^2)\,dt$,
(3) $P \mapsto \int_{0}^{1} P(t)Q(t)^2\,dt$,
(4) $P \mapsto P'''(-1)$,
(5) $P \mapsto \int_{0}^{1} (P+Q)^2\,dt$?

19. Пусть $\mathcal{X}$ – векторное пространство, и пусть $f, g \in \mathcal{X}^*$ таковы, что $\mathrm{Ker}\, f = \mathrm{Ker}\, g$. Доказать, что тогда $g = \lambda f$ для некоторого скаляра λ.

20. Пусть $\boldsymbol{x}$ – ненулевой вектор пространства $\mathcal{X}$. Однозначно ли определяется функция $f \in \mathcal{X}^*$ условием $f(\boldsymbol{x}) = 1$?

21. Пусть $\varPhi : \Pi_n \to \Pi_n$ – отображение, определенное по правилу $\varPhi(P(t)) = tP'(t) - P(t)$. Проверить линейность $\varPhi$, найти $\mathrm{Ker}\,\varPhi$ и вычислить $\mathrm{rg}\,\varPhi$.

22. Показать, что отображение $F_T : X \mapsto T^{-1}XT$, определенное невырожденной матрицей $T \in M_n(\mathbb{R})$, линейно на $M_n(\mathbb{R})$ и обладает свойством $F_T(XY) = F_T(X)F_T(Y)$.

23. Пусть $\mathcal{A}$ – линейный оператор на n-мерном линейном пространстве $\mathcal{X}$. Доказать, что $\mathcal{A} = \mathcal{A}^2$ если и только если $\operatorname{rg}\mathcal{A} + \operatorname{rg}(\mathcal{E} - \mathcal{A}) = n$.

24. Вычислить $J_k(\lambda)^m$ для любого натурального m, где $J_k(\lambda)$ – жорданова $k \times k$ клетка с собственным числом λ.

25. Пусть $\mathcal{F}$ – эрмитова форма в эрмитовом пространстве $\mathcal{X}$, а $\mathcal{Q}(\boldsymbol{x}) = \mathcal{F}(\boldsymbol{x}, \boldsymbol{x})$. Доказать, что для любых $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{X}$ верно равенство

$$\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \frac{1}{4}\big(\mathcal{Q}(\boldsymbol{x} + \boldsymbol{y}) - \mathcal{Q}(\boldsymbol{x} - \boldsymbol{y}) + i\,\mathcal{Q}(\boldsymbol{x} + i\boldsymbol{y}) - i\,\mathcal{Q}(\boldsymbol{y} - i\boldsymbol{y})\big).$$

26. Проверить, какие из следующий функций будут билинейными формами в соответствующих пространствах. Для каждой билинейной формы канонический вид квадратичной формы, определенной ее симметричной частью:

$$f(A, B) = \operatorname{Tr}(A^\top B),\ A, B \in \mathrm{M}_2(\mathbb{R}), \quad f(P, Q) = \int_0^1 P'(t)Q(t)\,dt,\ P, Q \in \mathbb{R}[t]_3.$$

27. Найти связь между матрицами A, B, G линейных операторов $\mathcal{A}$ и $\mathcal{B}$ и билинейной функции $\mathcal{G}(\cdot, \cdot)$ в некотором базисе пространства $\mathcal{X}$, $\dim \mathcal{X} = n$, и матрицей F (в том же базисе) билинейной функции $\mathcal{F}(\cdot, \cdot)$, определенной равенством $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \mathcal{G}(\mathcal{A}\,\boldsymbol{x}, \mathcal{B}\,\boldsymbol{y})$.

28. Доказать, что всякая билинейная (полуторалинейная) функция $\mathcal{F}$ ранга 1 может быть представлена в виде произведения двух линейных функций $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \mathfrak{f}(\boldsymbol{x})\mathfrak{g}(\boldsymbol{y})$ (соответственно $\mathcal{F}(\boldsymbol{x}, \boldsymbol{y}) = \mathfrak{f}(\boldsymbol{x})\mathfrak{g}(\overline{\boldsymbol{y}})$). К какому простейшему виду можно привести матрицу функции $\mathcal{F}$ с помощью замены базиса?

29. Доказать, что определитель целочисленной кососимметрической матрицы является квадратом целого числа.

30. Доказать, что для любой положительно определённой симметрической билинейной (эрмитовой) функции $\mathcal{F}$ выполнено неравенство

$$\sqrt{\mathcal{F}(\boldsymbol{x} + \boldsymbol{y}, \boldsymbol{x} + \boldsymbol{y})} \leqslant \sqrt{\mathcal{F}(\boldsymbol{x}, \boldsymbol{x})} + \sqrt{\mathcal{F}(\boldsymbol{y}, \boldsymbol{y})},$$

причём равенство имеет место тогда и только тогда, когда $\alpha\boldsymbol{x} = \beta\boldsymbol{y}$, где α, β – неотрицательные вещественные числа, не равные нулю одновременно.

31. Не применяя критерия Сильвестра, доказать, что для положительной определенности квадратичной функции $\sum_{j,k=1}^{n} a_{jk}x_j\overline{x}_k$ условие $a_{jj} > 0$ $(j = 1, \ldots, n)$ является необходимым, но не достаточным.

32. Доказать, что если два линейных оператора ранга 1 имеют равные ядра и равные образы, то они перестановочны.

33. Найти общий вид матрицы линейного оператора $\mathcal{A}$ в базисе, первые k векторов которого составляют: (а) базис ядра оператора $\mathcal{A}$; (б) базис образа оператора $\mathcal{A}$.

34. Найти характеристические числа матрицы $A^\top A$, где – матрица-строка $(a_1, \ldots, a_n)$.

35. Доказать, что матрица нильпотентна тогда и только тогда, когда все ее характеристические числа равны нулю.

36. Доказать, что для всякого линейного оператора $\mathcal{A}$ ранга 1 в комплексном линейном пространстве существует такое число α, что $\mathcal{A}^2 = \alpha\,\mathcal{A}$.

37. Доказать, что минимальный многочлен клеточно-диагональной матрицы равен наименьшему общему кратному минимальных многочленов её клеток.

38. Найти минимальный многочлен: (а) тождественного оператора; (б) нулевого оператора; (в) оператора проектирования n-мерного пространства $\mathcal{X}$ на k-мерное подпространство $\mathcal{Y}$ $(0 < k < n)$; (г) оператора отражения; (д) нильпотентного оператора индекса k.

39. Доказать, что минимальный многочлен матрицы порядка $\geqslant 2$ и ранга 1 имеет степень 2.

40. Что можно сказать о жордановой форме матрицы линейного оператора $\mathcal{A}$ в комплексном пространстве, если $\mathcal{A}^3 = \mathcal{A}^2$?

41. Пусть $\lambda_1, \ldots, \lambda_n$ – собственные значения матрицы $A \in \mathrm{M}_n(\mathbb{C})$. Доказать, что: (а) для любого натурального числа k верно равенство $\operatorname{Tr} A^k = \lambda_1^k + \cdots + \lambda_n^k$; (б) коэффициенты характеристического полинома матрицы A являются многочленами от $\operatorname{Tr} A, \ldots, \operatorname{Tr} A^n$; (в) если $\operatorname{Tr} A = \operatorname{Tr} A^2 = \cdots = \operatorname{Tr} A^n = 0$, то матрица A нильпотентна.

Литература

[1] Кострикин А. И. Введение в алгебру. Часть I. Основы алгебры. М.: Физматлит, 2004. 272 с.

[2] Кострикин А. И. Введение в алгебру. Часть 2. Линейная алгебра. М.: Физматлит, 2000. 368 с.

[3] Кострикин А. И. Введение в алгебру. Часть 3. Основные структуры. М.: Физматлит, 2004. 272 с.

[4] Сборник задач по алгебре / Под ред. А. И. Кострикина. М.: Физматлит, 2001. 464 с.

[5] Мальцев А. И. Основы линейной алгебры. М.: Наука, 1975. 400 с.

[6] Ильин В. А., Позняк Э. Г. Линейная алгебра. М.: Наука, 1978. 304 с.

[7] Проскуряков И. В. Сборник задач по линейной алгебре. М.: Наука, 1970. 384 с.

[8] Беклемишева Л. А., Петрович А. Ю., Чубаров И. А. Сборник задач по аналитической геометрии и линейной алгебре. М.: Физматлит, 2004. 496 с.

[9] Сборник задач по аналитической геометрии и линейной алгебре / Под ред. Ю. М. Смирнова. М.: Логос, 2005. 376 с.