

# Линейная алгебра - II

## 1 Евклидово пространство

**1.1.** Ранее мы определили линейное пространство как множество элементов (векторов), которые мы можем складывать между собой и умножать на числа. Это понятие оказалось чрезвычайно полезным в вопросах, связанных, например, с анализом решений систем линейных алгебраических уравнений. Однако на практике часто встречаются задачи, в которых такое пространство оказывается слишком бедным — в нем не ввести привычные нам со школьной геометрии понятия длины вектора, угла между векторами, расстояния между векторами и так далее. Для того, чтобы эти понятия ввести, нам понадобится важное понятие *скалярного произведения* векторов.

**1.1.1.** Рассмотрим линейное пространство  $V$  векторов.

**Определение 1.1.** Говорят, что на множестве векторов — элементов линейного пространства  $V$  введено скалярное произведение, если мы любым двум векторам  $x$  и  $y$  этого пространства можем сопоставить вещественное число  $(x, y)$ , причем это соответствие обладает следующими свойствами:

1. симметричность скалярного произведения:  $(x, y) = (y, x)$ ;
2. для любого  $\lambda \in \mathbb{R}$  скалярное произведение  $(\lambda x, y) = \lambda(x, y)$ ;
3. дистрибутивность скалярного произведения:  $(x_1 + x_2, y) = (x_1, y) + (x_2, y)$ ;
4. неотрицательность скалярного произведения:  $(x, x) \geq 0$ , причем  $(x, x) = 0$  тогда и только тогда, когда  $x = \mathbf{0}$ .

Линейное пространство, в котором определено скалярное произведение, называется линейным пространством со скалярным произведением или евклидовым пространством.

**1.1.2.** Покажем теперь, как нам в рассмотренных выше примерах линейных пространств корректно ввести скалярное произведение.

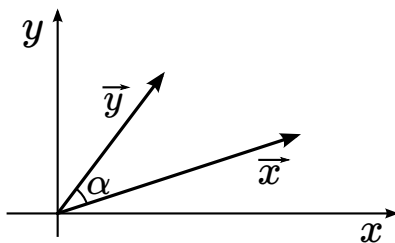


Рис. 1

**Пример 1.2.** Рассмотрим линейное пространство векторов (направленных отрезков) на плоскости  $\mathbb{R}^2$ . У любого такого вектора  $x$  мы можем измерить длину  $|x|$ , а также угол  $\alpha$  между двумя векторами  $x$  и  $y$  (рис 1). Тогда по определению мы можем в качестве скалярного произведения  $(x, y)$  пары векторов положить число

$$(x, y) = |x| \cdot |y| \cdot \cos \alpha.$$

**Пример 1.3.** Обобщением приведенного выше примера является линейное пространство  $\mathbb{R}^n$ , элементами которого являются наборы  $n$  вещественных чисел  $x := (\xi_1, \xi_2, \dots, \xi_n)$ . Сложение этих векторов и умножение на вещественные числа осуществляются в таком пространстве покомпонентно:

$$\lambda \cdot x = \lambda \cdot (\xi_1, \xi_2, \dots, \xi_n) := (\lambda \xi_1, \lambda \xi_2, \dots, \lambda \xi_n).$$

$$(x, y) := \xi_1 \eta_1 + \xi_2 \eta_2 + \dots + \xi_n \eta_n.$$

**Пример 1.4.** В том же самом линейном пространстве  $\mathbb{R}^n$   $n$ -мерных векторов скалярное произведение мы можем ввести и по другому. А именно, определим умножение двух векторов  $x$  и  $y$  по формуле

Набор чисел  $a_{ij}$ , задающий нам операцию умножения в пространстве  $\mathbb{R}^n$ , мы можем записать в виде матрицы

Давайте посмотрим, какие условия мы должны наложить на числа  $a_{ij}$  для того, чтобы они корректно определяли нам в линейном пространстве  $\mathbb{R}^n$  скалярное произведение.

$$(x, y) = (y, x) \quad \Longleftrightarrow \quad a_{ji} = a_{ij}.$$
$$(x, x) = \sum_{i,j=1}^n a_{ij} \xi_i \xi_j \geq 0 \quad \forall \quad \xi_i, \xi_j \in \mathbb{R},$$
$$(x, x) = 0$$

должно выполняться тогда и только тогда, когда все  $\xi_i = 0$ . Матрица  $\mathcal{A}$ , коэффициенты которой удовлетворяют описанным выше свойствам, называется *положительно определенной* матрицей.

Таким образом, определяемое матрицей  $\mathcal{A}$  произведение векторов задает в линейном пространстве  $\mathbb{R}^n$  скалярное произведение тогда и только тогда, когда матрица  $\mathcal{A}$  является симметричной положительно определенной матрицей.

Заметим сразу же, что если в качестве  $\mathcal{A}$  выбрать единичную матрицу

$$\mathcal{A} := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

то определяемое такой матрицей скалярное произведение примет вид

$$(x, y) := \xi_1 \eta_1 + \xi_2 \eta_2 + \dots + \xi_n \eta_n,$$

то есть мы получаем то самое простейшее скалярное произведение, которое мы ввели в примере 1.3. Введенное в данном примере скалярное произведение является, таким образом, его обобщением на случай произвольной положительно определенной симметричной матрицы  $\mathcal{A}$ .

**Пример 1.5.** Рассмотрим, наконец, линейное пространство  $C[a, b]$  функций, непрерывных на отрезке  $[a, b]$ . Введем на этом пространстве произведение по формуле

$$(f, g) := \int_a^b f(x) \cdot g(x) dx.$$

Можно показать, что такого рода операция удовлетворяет всем четырем свойствам скалярного произведения, то есть превращает данное линейное пространство в евклидово.

**1.2.** Покажем теперь, как с помощью понятия скалярного произведения ввести такие важные на практике понятия, как длина вектора или угол между векторами.

**1.2.1.** В качестве наводящих соображений вернемся к хорошо нам известному со школы случаю векторов на плоскости. Для этого частного случая мы скалярное произведение определили через длины направленных отрезков и угол между ними (рис. 1):

$$(\mathbf{x}, \mathbf{y}) = |\mathbf{x}| \cdot |\mathbf{y}| \cdot \cos \alpha. \quad (1)$$

Попробуем теперь получить отсюда обратные формулы, а именно, выразить длину  $|\mathbf{x}|$  вектора  $\mathbf{x}$ , а также угол  $\alpha$  между двумя векторами выразить через скалярное произведение  $(\mathbf{x}, \mathbf{y})$ . Легче всего, конечно же, выразить из (1) косинус угла между векторами:

$$\cos \alpha = \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| \cdot |\mathbf{y}|} \quad \implies \quad \alpha = \arccos \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| \cdot |\mathbf{y}|}.$$

Осталось теперь выразить через скалярное произведение длину вектора. Для этого нужно в определении (1) скалярного произведения положить  $\mathbf{x} = \mathbf{y}$ :

$$(\mathbf{x}, \mathbf{x}) = |\mathbf{x}| \cdot |\mathbf{x}| \cdot \cos 0 = |\mathbf{x}| \cdot |\mathbf{x}| \quad \implies \quad |\mathbf{x}| = \sqrt{(\mathbf{x}, \mathbf{x})}.$$

**1.2.2.** Итак, мы выразили в частном случае векторов на плоскости длину вектора и угол между векторами через скалярное произведение векторов. Теперь мы можем обобщить эти результаты и считать, что те же самые формулы определяют нам *по определению* и длину  $|\mathbf{x}|$  вектора, и угол  $\alpha$  между векторами в произвольном линейном пространстве со скалярным произведением. Именно,

**Определение 1.6.** Длиной  $|\mathbf{x}|$  вектора  $\mathbf{x}$  в евклидовом пространстве  $V$  называется вещественное число

$$|\mathbf{x}| := \sqrt{(\mathbf{x}, \mathbf{x})}.$$

**Определение 1.7.** Углом  $\alpha$  между двумя векторами  $\mathbf{x}$  и  $\mathbf{y}$  в евклидовом пространстве  $V$  называется вещественное число

$$\alpha := \arccos \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| \cdot |\mathbf{y}|}.$$

**1.2.3.** Корректность определения длины вектора  $\mathbf{x}$  сомнений не вызывает: согласно четвертой аксиоме евклидова пространства, скалярное произведение всегда неотрицательно, и поэтому арифметический корень из такого числа всегда корректно определен. С определением же угла между векторами все немного сложнее: для того, чтобы доказать корректность введенного выше определения, нам нужно убедиться в том, что

$$-1 \leq \cos \alpha \leq 1 \quad \Longleftrightarrow \quad -1 \leq \frac{(\mathbf{x}, \mathbf{y})}{|\mathbf{x}| \cdot |\mathbf{y}|} \leq 1.$$

Последние неравенства равносильны следующему единственному неравенству:

$$\frac{(\mathbf{x}, \mathbf{y})^2}{|\mathbf{x}|^2 \cdot |\mathbf{y}|^2} \leq 1 \quad \Longleftrightarrow \quad (\mathbf{x}, \mathbf{y})^2 \leq |\mathbf{x}|^2 \cdot |\mathbf{y}|^2.$$

Последнее неравенство называется *неравенством Коши-Буняковского*. Итак, для доказательства корректности определения угла между векторами в евклидовом пространстве нам осталось доказать неравенство Коши-Буняковского.

**1.2.4.** Для доказательства неравенства Коши-Буняковского рассмотрим вектор  $\mathbf{x} - t\mathbf{y}$ , где  $t$  — произвольный вещественный параметр. Согласно четвертой аксиоме скалярного произведения, скалярное произведение этого вектора на самого себя больше или равно нулю:

$$(\mathbf{x} - t\mathbf{y}, \mathbf{x} - t\mathbf{y}) \geq 0 \quad \Longleftrightarrow \quad (\mathbf{x}, \mathbf{x}) - 2t \cdot (\mathbf{x}, \mathbf{y}) + t^2 \cdot (\mathbf{y}, \mathbf{y}) \geq 0.$$

Стоящее в левой части последнего неравенства выражение мы можем рассматривать как квадратный относительно параметра  $t$  трехчлен. Геометрически любой такой трехчлен представляет собой параболу на плоскости  $(t, f(t))$ . Тот факт, что такой трехчлен больше или равен нулю при любых  $t$ , означает, что график этой параболы находится либо целиком над осью абсцисс, либо касается этой параболы в единственной точке (рис. 2). С алгебраической же точки зрения это



Рис. 2

означает, что квадратное уравнение

$$t^2 \cdot (\mathbf{y}, \mathbf{y}) - 2t \cdot (\mathbf{x}, \mathbf{y}) + (\mathbf{x}, \mathbf{x}) = 0$$

либо вовсе не имеет корней, либо имеет единственный корень. А это, в свою очередь, возможно тогда и только тогда, когда дискриминант данного квадратного уравнения меньше или равен нулю:

$$D = 4(\mathbf{x}, \mathbf{y})^2 - 4(\mathbf{x}, \mathbf{x})(\mathbf{y}, \mathbf{y}) \leq 0 \quad \Longleftrightarrow \quad (\mathbf{x}, \mathbf{y})^2 \leq |\mathbf{x}|^2 \cdot |\mathbf{y}|^2.$$

Тем самым неравенство Коши-Буняковского доказано.

## 2 Базис в евклидовом пространстве

**2.1.** Как мы уже говорили ранее, в качестве базиса в обычном  $n$ -мерном линейном пространстве мы можем выбирать абсолютно произвольный набор из  $n$  линейно независимых векторов. При этом у нас нет никаких особенных предпочтений в том, как именно нам эти векторы выбирать. В линейном же пространстве со скалярным произведением ситуация несколько иная — в этом пространстве в качестве базиса удобно выбирать так называемый набор из  $n$  взаимно ортогональных векторов.

**2.1.1.** Начнем с определения ортогональности пары векторов.

**Определение 2.1.** Два ненулевых вектора  $\mathbf{x}$ ,  $\mathbf{y}$  в евклидовом пространстве называются ортогональными, если их скалярное произведение равно нулю:

$$(\mathbf{x}, \mathbf{y}) = 0.$$

Рассмотрим в качестве примера два вектора на плоскости  $\mathbb{R}^2$ . Скалярное произведение двух ненулевых векторов на этой плоскости

$$(\mathbf{x}, \mathbf{y}) = |\mathbf{x}| \cdot |\mathbf{y}| \cdot \cos \alpha$$

равняется нулю тогда и только тогда, когда угол  $\alpha = 90^\circ$ . Аналогичный факт, естественно, имеет место и для произвольного евклидова пространства  $V$ .

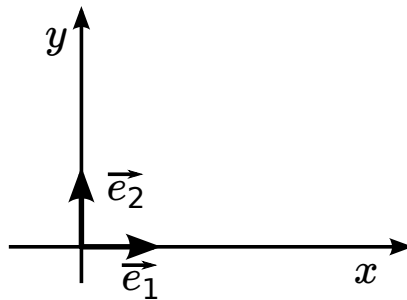


Рис. 3

**2.1.2.** Вернемся теперь в вопросу выбора базиса в евклидовом пространстве. В качестве примера вновь рассмотрим евклидово пространство векторов на плоскости (рис. 3). В этом простейшем случае мы всегда можем выбрать два ортогональных друг другу вектора  $\mathbf{e}_1$  и  $\mathbf{e}_2$ , имеющих единичную длину. Еще со школьной геометрии мы знаем, что задание таких векторов, по сути, равносильно введению на плоскости прямоугольной системы координат. На языке же линейной алгебры это означает, что эти векторы образуют на плоскости так называемый *ортонормированный базис*.

**Определение 2.2.** Говорят, что набор из  $n$  ненулевых векторов  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  образует ортогональный базис, если все эти векторы попарно ортогональны друг другу. В случае, если длина каждого вектора равна единице, базис называют ортонормированным. Иными словами, набор векторов  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  образует ортонормированный базис, если

$$(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad \forall i \neq j \quad \text{и} \quad (\mathbf{e}_i, \mathbf{e}_i) = 1 \quad \forall i = 1, \dots, n.$$

Нам, однако, нужно убедиться в корректности данного определения, а именно, показать, что векторы, входящие в набор  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , линейно независимы. Так как их ровно  $n$  штук, то тем самым мы удостоверимся в том, что они действительно образуют базис нашего пространства.

Для проверки линейной независимости нам достаточно доказать, что равенство

$$\lambda_1 \mathbf{e}_1 + \lambda_2 \mathbf{e}_2 + \dots + \lambda_n \mathbf{e}_n = \mathbf{0}$$

возможно тогда и только тогда, когда все  $\lambda_i = 0$ . Для этого домножим это равенство скалярно на вектор  $\mathbf{e}_1$ :

$$\lambda_1(\mathbf{e}_1, \mathbf{e}_1) + \lambda_2(\mathbf{e}_2, \mathbf{e}_1) + \dots + \lambda_n(\mathbf{e}_n, \mathbf{e}_1) = (\mathbf{0}, \mathbf{e}_1) = 0.$$

По условию, все скалярные произведения вида  $(\mathbf{e}_i, \mathbf{e}_1) = 0$  для всех  $i \neq 1$ , а скалярное произведение  $(\mathbf{e}_1, \mathbf{e}_1) = 1$ . Следовательно, мы из последнего соотношения получаем, что

$$\lambda_1 \cdot 1 = 0 \quad \implies \quad \lambda_1 = 0.$$

Действуя аналогично с оставшимися векторами, то есть домножая исходное равенство скалярно на векторы  $\mathbf{e}_2, \dots, \mathbf{e}_n$ , мы убедимся, что все  $\lambda_i$  в этом равенстве обязаны равняться нулю. Линейная независимость набора векторов  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  тем самым нами полностью доказана.

**2.1.3.** Предположим теперь, что у нас в  $n$ -мерном пространстве  $V$  имеется какой-то набор  $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$  линейно независимых векторов, не обязательно ортонормированный. Очень часто на практике возникает задача превратить его в набор из  $n$  попарно-ортогональных векторов  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ . Эта процедура носит специальное название — процесс ортогонализации Грамма-Шмидта. Опишем эту процедуру.

На первом шаге выберем в качестве вектора  $\mathbf{e}_1$  вектор  $\mathbf{f}_1$ , то есть положим

$$\mathbf{e}_1 = \mathbf{f}_1.$$

На втором шаге нашей процедуры положим вектор  $\mathbf{e}_2$  равным линейной комбинации векторов  $\mathbf{f}_2$  и  $\mathbf{e}_1$ :

$$\mathbf{e}_2 = \mathbf{f}_2 + \alpha_1 \cdot \mathbf{e}_1.$$

Коэффициент  $\alpha_1$  при этом найдем из условия ортогональности векторов  $\mathbf{e}_1$  и  $\mathbf{e}_2$ :

$$(\mathbf{e}_1, \mathbf{e}_2) = 0 \quad \iff \quad (\mathbf{e}_1, \mathbf{f}_2 + \alpha_1 \cdot \mathbf{e}_1) = 0 \quad \iff \quad (\mathbf{e}_1, \mathbf{f}_2) + \alpha_1(\mathbf{e}_1, \mathbf{e}_1) = 0.$$

Так как вектор  $\mathbf{e}_1 = \mathbf{f}_1$  принадлежит набору линейно независимых векторов, то он отличен от  $\mathbf{0}$ , а потому скалярное произведение  $(\mathbf{e}_1, \mathbf{e}_1) \neq 0$ . Следовательно, мы из полученного нами выше равенства всегда можем определить коэффициент  $\alpha_1$ , обеспечивающий нам условие ортогональности векторов  $\mathbf{e}_1$  и  $\mathbf{e}_2$ :

$$\alpha_1 = -\frac{(\mathbf{e}_1, \mathbf{f}_2)}{(\mathbf{e}_1, \mathbf{e}_1)}.$$

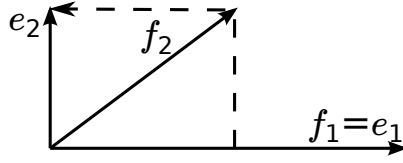


Рис. 4

Описанная выше процедура проиллюстрирована на рис 4. Мы вначале имели какие-то два вектора  $\mathbf{f}_1$  и  $\mathbf{f}_2$ , не лежащие на одной и той же прямой. В качестве вектора  $\mathbf{e}_1$  мы выбрали вектор  $\mathbf{f}_1$ . Нам осталось каким-то образом подправить вектор  $\mathbf{f}_2$  так, чтобы подправленный вектор — вектор  $\mathbf{e}_2$  — оказался бы ортогональным вектору  $\mathbf{e}_1$ . Для этого мы берем вектор  $\mathbf{f}_2$  и вычитаем из него вектор  $((\mathbf{e}_1, \mathbf{f}_2)/(\mathbf{e}_1, \mathbf{e}_1))\mathbf{e}_1$  — по сути, вектор, параллельный  $\mathbf{e}_1$  и равный по величине проекции вектора  $\mathbf{f}_2$  на вектор  $\mathbf{e}_1$ . В результате мы получаем вектор  $\mathbf{e}_2$ , ортогональный выбранному ранее вектору  $\mathbf{e}_1$ .

Опишем теперь еще один шаг процедуры ортогонализации. Для этого в качестве вектора  $\mathbf{e}_3$  выберем линейную комбинацию векторов вида

$$\mathbf{e}_3 = \mathbf{f}_3 + \beta_1 \cdot \mathbf{e}_1 + \beta_2 \cdot \mathbf{e}_2.$$

Коэффициенты же  $\beta_1$  и  $\beta_2$  мы постараемся подобрать так, чтобы получившийся вектор  $\mathbf{e}_3$  оказался ортогональным как вектору  $\mathbf{e}_1$ , так и вектору  $\mathbf{e}_2$ . Из условия  $(\mathbf{e}_1, \mathbf{e}_3) = 0$  получаем, что

$$(\mathbf{e}_1, \mathbf{f}_3) + \beta_1(\mathbf{e}_1, \mathbf{e}_1) + \beta_2(\mathbf{e}_1, \mathbf{e}_2) = 0.$$

Учитывая ортогональность  $(\mathbf{e}_1, \mathbf{e}_2) = 0$  векторов  $\mathbf{e}_1$  и  $\mathbf{e}_2$ , а также тот факт, что скалярное произведение  $(\mathbf{e}_1, \mathbf{e}_1) \neq 0$ , мы из последнего уравнения находим  $\beta_1$  как

$$\beta_1 = -\frac{(\mathbf{e}_1, \mathbf{f}_3)}{(\mathbf{e}_1, \mathbf{e}_1)}.$$

Действуя аналогично, мы из условия  $(\mathbf{e}_2, \mathbf{e}_3) = 0$  определяем коэффициент  $\beta_2$  по формуле

$$\beta_2 = -\frac{(\mathbf{e}_2, \mathbf{f}_3)}{(\mathbf{e}_2, \mathbf{e}_2)}.$$

Теперь уже достаточно очевидно, как описать общий,  $k$ -й шаг процедуры ортогонализации Грамма-Шмидта — нам нужно в качестве вектора  $\mathbf{e}_k$  выбрать линейную комбинацию вида

$$\mathbf{e}_k = \mathbf{f}_k + \lambda_1 \cdot \mathbf{e}_1 + \lambda_2 \cdot \mathbf{e}_2 + \dots + \lambda_{k-1} \mathbf{e}_{k-1},$$

а коэффициенты  $\lambda_i$  найти из условия ортогональности векторов  $\mathbf{e}_k$  и  $\mathbf{e}_i$ .

**Замечание 2.3.** В результате процесса ортогонализации мы получаем набор  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  попарно ортогональных векторов. Поделив каждый вектор на его длину, мы из него всегда можем построить ортонормированный базис.

**Замечание 2.4.** Вовсе не обязательно в качестве  $n$  выбирать размерность исходного евклидова пространства — мы всегда можем взять произвольный набор из  $m \leq n$  линейно независимых векторов и превратить его в ортогональный.

**2.2.** Давайте теперь поймем, чем же ортогональный базис лучше неортогонального.

**2.2.1.** Как мы знаем, если у нас в пространстве имеется базис, то мы по этому базису любой вектор можем разложить. Возьмем два вектора  $\mathbf{x}$  и  $\mathbf{y}$ , выберем в евклидовом пространстве  $V$  ортонормированный базис  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  и разложим эти два вектора по выбранному базису:

$$\mathbf{x} = \xi^1 \mathbf{e}_1 + \xi^2 \mathbf{e}_2 + \dots + \xi^n \mathbf{e}_n, \quad \mathbf{y} = \eta^1 \mathbf{e}_1 + \eta^2 \mathbf{e}_2 + \dots + \eta^n \mathbf{e}_n.$$

Здесь  $\xi^1, \dots, \xi^n$  есть координаты вектора  $\mathbf{x}$  в базисе  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , а  $\eta^1, \dots, \eta^n$  — координаты вектора  $\mathbf{y}$  в том же базисе.

Вычислим скалярное произведение векторов  $\mathbf{x}$  и  $\mathbf{y}$ :

$$(\mathbf{x}, \mathbf{y}) = (\xi^1 \mathbf{e}_1 + \xi^2 \mathbf{e}_2 + \dots + \xi^n \mathbf{e}_n, \eta^1 \mathbf{e}_1 + \eta^2 \mathbf{e}_2 + \dots + \eta^n \mathbf{e}_n).$$

Учитывая, что

$$(\mathbf{e}_i, \mathbf{e}_j) = 0 \quad \forall i \neq j \quad \text{и} \quad (\mathbf{e}_i, \mathbf{e}_i) = 1 \quad \forall i = 1, \dots, n,$$

мы получаем, что

$$(\mathbf{x}, \mathbf{y}) = \xi^1 \eta^1 + \xi^2 \eta^2 + \dots + \xi^n \eta^n. \quad (2)$$

Иными словами, в ортонормированном базисе скалярное произведение любых двух векторов равняется сумме произведений координат этих векторов в базисе  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ .

**2.2.2.** Давайте теперь возьмем произвольный базис  $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$  в исходном пространстве. Раскладывая два произвольных вектора  $\mathbf{x}$  и  $\mathbf{y}$  по этому базису

$$\mathbf{x} = \xi^1 \mathbf{f}_1 + \xi^2 \mathbf{f}_2 + \dots + \xi^n \mathbf{f}_n, \quad \mathbf{y} = \eta^1 \mathbf{f}_1 + \eta^2 \mathbf{f}_2 + \dots + \eta^n \mathbf{f}_n,$$

а затем вычисляя скалярное произведение этих векторов, мы получим, что

$$(\mathbf{x}, \mathbf{y}) = \sum_{i,j=1}^n a_{ij} \xi^i \eta^j, \quad \text{где} \quad a_{ij} = (\mathbf{f}_i, \mathbf{f}_j).$$

Как видно, в произвольном, не ортонормированном базисе скалярное произведение векторов выражается через свои координаты в этом базисе по формуле, значительно более сложной по сравнению с формулой (2).

Иными словами, в произвольном базисе скалярное произведение пары векторов задается некоторой матрицей  $\mathcal{A} = \{a_{ij}\}$ ,  $a_{ij} = (\mathbf{f}_i, \mathbf{f}_j)$ . В случае же ортонормированного базиса эта матрица является единичной.

**2.2.3.** Теперь рассмотрим разложение произвольного вектора  $\mathbf{x}$  по базису

$$\mathbf{x} = \xi^1 \mathbf{e}_1 + \xi^2 \mathbf{e}_2 + \dots + \xi^n \mathbf{e}_n$$

и зададимся следующим вопросом — а как же нам найти коэффициенты  $\xi^i$  в таком разложении? Оказывается, в случае, когда базис — ортонормированный, сделать это очень просто. Для этого домножим скалярно левую и правую часть выписанного выше равенства на вектор  $\mathbf{e}_i$ :

$$(\mathbf{x}, \mathbf{e}_i) = \xi^1 (\mathbf{e}_1, \mathbf{e}_i) + \xi^2 (\mathbf{e}_2, \mathbf{e}_i) + \dots + \xi^n (\mathbf{e}_n, \mathbf{e}_i).$$

В силу ортонормированности базиса все скалярные произведения, кроме  $(\mathbf{e}_i, \mathbf{e}_i)$ , обратятся в ноль. С учетом же равенства  $(\mathbf{e}_i, \mathbf{e}_i) = 1$  мы получим следующую удобную формулу для вычисления коэффициентов  $\xi^i$  в разложении вектора по базису:

$$\xi^i = (\mathbf{x}, \mathbf{e}_i).$$



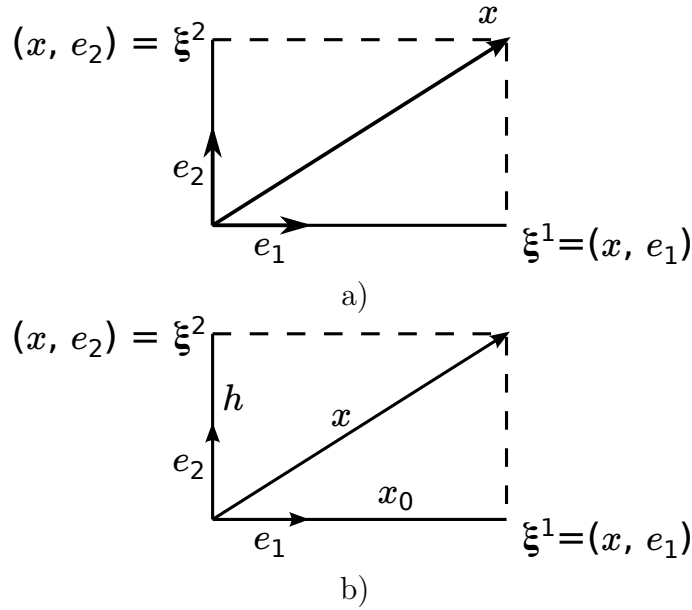


Рис. 5

В качестве иллюстрации к вышесказанному вновь рассмотрим векторы на плоскости (рис.5,а). Выберем на плоскости пару ортонормированных векторов  $\mathbf{e}_1$  и  $\mathbf{e}_2$ . Тогда любой вектор  $\mathbf{x}$  на плоскости может быть представлен в виде линейной комбинации базисных векторов с коэффициентами  $\xi^i$ , равными скалярным произведениям  $(\mathbf{x}, \mathbf{e}_i)$ .

**2.3.** Следующее очень важное понятие, которое нам нужно изучить, связано с идеей проектирования вектора на подпространство.

**2.3.1.** Вернемся к только что рассмотренному примеру вектора на плоскости (рис. 5,б). С геометрической точки зрения координаты  $\xi^i$  вектора  $\mathbf{x}$  в ортонормированном базисе  $\{\mathbf{e}_1, \mathbf{e}_2\}$  есть не что иное, как проекции этого вектора  $\mathbf{x}$  на взаимно-перпендикулярные прямые, проходящие через векторы  $\mathbf{e}_1$  и  $\mathbf{e}_2$  базиса. Иными словами, скалярное произведение  $(\mathbf{x}, \mathbf{e}_i)$  определяет нам в данном случае проекцию  $\mathbf{x}_0$  вектора  $\mathbf{x}$  на одномерное подпространство, натянутое на вектор  $\mathbf{e}_i$ , а координата  $\xi^1 = (\mathbf{x}, \mathbf{e}_1)$  задает длину этой проекции. При этом длина  $\xi^2 = (\mathbf{x}, \mathbf{e}_2)$  вектора  $\mathbf{h}$  — проекции вектора  $\mathbf{x}$  на  $\mathbf{e}_2$  — задает нам кратчайшее расстояние от вектора  $\mathbf{x}$  до одномерного подпространства, натянутого на вектор  $\mathbf{e}_1$ .

**2.3.2.** Давайте теперь обобщим те понятия, с которыми мы встретились в предыдущем примере, на общий случай линейного пространства  $V$  со скалярным произведением. Пусть  $V_0$  есть некоторое подпространство исходного пространства  $V$ .

**Определение 2.5.** Говорят, что вектор  $\mathbf{h} \in V$  ортогонален подпространству  $V_0$ , если он ортогонален любому вектору  $\mathbf{x}_0$  этого подпространства:

$$(\mathbf{h}, \mathbf{x}_0) = 0 \quad \forall \mathbf{x}_0 \in V_0.$$

Пусть теперь  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  есть ортонормированный набор векторов, образующий базис подпространства  $V_0$ . Ясно, что ортогональность  $\mathbf{h}$  произвольному вектору  $\mathbf{x}_0 \in V_0$  равносильна ортогональности  $\mathbf{h}$  каждому вектору нашего базиса:

$$(\mathbf{h}, \mathbf{x}_0) = 0 \quad \forall \mathbf{x}_0 \in V_0 \quad \Longleftrightarrow \quad (\mathbf{h}, \mathbf{e}_i) = 0 \quad \forall i = 1, \dots, m.$$

Рассмотрим, наконец, произвольный вектор  $\mathbf{x}$  исходного пространства и поставим следующую задачу: найти вектор  $\mathbf{x}_0 \in V_0$ , такой, чтобы вектор  $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$  был бы ортогонален подпространству  $V_0$ . Вектор  $\mathbf{x}_0$  называется при этом *ортогональной проекцией* вектора  $\mathbf{x}$  на подпространство  $V_0$ . Оказывается, что с геометрической точки зрения данная задача равносильна задаче поиска кратчайшего расстояния от точки, задаваемой концом вектора  $\mathbf{x}$ , до подпространства  $V_0$ .

Действительно, возьмем произвольный вектор  $\mathbf{y}_0 \in V_0$ , отличный от  $\mathbf{x}_0$  и покажем, что расстояние  $|\mathbf{x} - \mathbf{y}_0|$  от  $\mathbf{x}$  до этого вектора  $\mathbf{y}_0$  всегда больше расстояния  $|\mathbf{x} - \mathbf{x}_0|$ . Для этого воспользуемся теоремой Пифагора (упражнение ??): так как вектор  $\mathbf{y}_0 - \mathbf{x}_0$  принадлежит подпространству  $V_0$ , то он ортогонален вектору  $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$ , и, по теореме Пифагора,

$$|\mathbf{x} - \mathbf{y}_0| = |\mathbf{x} - \mathbf{x}_0 + \mathbf{x}_0 - \mathbf{y}_0| = |\mathbf{x} - \mathbf{x}_0| + |\mathbf{x}_0 - \mathbf{y}_0| \implies |\mathbf{x} - \mathbf{y}_0| > |\mathbf{x} - \mathbf{x}_0|.$$

**2.3.3.** Покажем теперь, как же нам практически решить задачу поиска ортогональной проекции  $\mathbf{x}_0$  вектора  $\mathbf{x}$  на заданное подпространство  $V_0$ . Предположим, что набор  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$  образует ортонормированный базис подпространства  $V_0$ . Как следствие, мы любой вектор из этого подпространства, и, в частности, вектор  $\mathbf{x}_0$ , можем разложить по этому базису:

$$\mathbf{x}_0 = \xi^1 \mathbf{e}_1 + \dots + \xi^m \mathbf{e}_m.$$

Наша задача, таким образом, свелась к задаче поиска координат  $\xi^i$  вектора  $\mathbf{x}_0$  в этом базисе. А эти координаты мы можем найти из условия ортогональности вектора  $\mathbf{h} = \mathbf{x} - \mathbf{x}_0$  каждому из базисных векторов  $\mathbf{e}_i$ :

$$(\mathbf{h}, \mathbf{e}_i) = 0 \iff (\mathbf{x} - \mathbf{x}_0, \mathbf{e}_i) = 0 \iff (\mathbf{x}, \mathbf{e}_i) = (\mathbf{x}_0, \mathbf{e}_i).$$

Подставляя теперь в это равенство разложение вектора  $\mathbf{x}_0$  по базису подпространства  $V_0$ , мы получаем следующую систему линейных алгебраических уравнений для определения неизвестных коэффициентов  $\xi^i$ :

$$\begin{cases} \xi^1(\mathbf{e}_1, \mathbf{e}_1) + \xi^2(\mathbf{e}_2, \mathbf{e}_1) + \dots + \xi^n(\mathbf{e}_n, \mathbf{e}_1) = (\mathbf{x}, \mathbf{e}_1) \\ \xi^1(\mathbf{e}_1, \mathbf{e}_2) + \xi^2(\mathbf{e}_2, \mathbf{e}_2) + \dots + \xi^n(\mathbf{e}_n, \mathbf{e}_2) = (\mathbf{x}, \mathbf{e}_2) \\ \dots \\ \xi^1(\mathbf{e}_1, \mathbf{e}_n) + \xi^2(\mathbf{e}_2, \mathbf{e}_n) + \dots + \xi^n(\mathbf{e}_n, \mathbf{e}_n) = (\mathbf{x}, \mathbf{e}_n) \end{cases} \quad (3)$$

В силу ортонормированности векторов  $\mathbf{e}_i$  мы отсюда сразу же получаем, что

$$\xi^i = (\mathbf{x}, \mathbf{e}_i). \quad (4)$$

**Замечание 2.6.** Все проведенные выше рассуждения остаются справедливыми и в случае, когда мы вместо ортонормированного набора векторов берем в подпространстве  $V_0$  произвольный набор из  $m$  линейно независимых векторов  $\{\mathbf{e}_1, \dots, \mathbf{e}_m\}$ . Единственное отличие состоит в том, что коэффициенты  $\xi^i$  вектора  $\mathbf{x}_0$  проекции  $\mathbf{x}$  на подпространство  $V_0$  вычисляются не по явным формулам (4), а из решения линейной системы алгебраических уравнений (3).

**2.4.** Вернемся к замечанию 2.6 и зададимся следующим вопросом — а зачем же нам усложнять задачу и выбирать в подпространстве  $V_0$  неортогональный базис? Оказывается, что к этой ситуации мы естественным образом приходим при решении задач *методом наименьших квадратов*.

**2.4.1.** Метод наименьших квадратов довольно часто возникает в следующих задачах. Предположим, что мы изучаем некоторый сложный процесс, зависящий от каких-либо  $m$  управляющих параметров  $x_1, \dots, x_m$ . Меняя значения этих параметров, мы получаем различные значения функции  $y = f(x_1, \dots, x_m)$ , описывающей каким-то образом изучаемый нами процесс. В простейшем случае мы можем считать, что зависимость  $y$  от  $x_1, \dots, x_m$  линейна, то есть что функция  $y = f(x_1, \dots, x_m)$  имеет следующий вид:

$$y = f(x_1, \dots, x_m) = c_1 x_1 + \dots + c_m x_m.$$

Если бы нам удалось построить какую-то теоретическую модель изучаемого процесса, то мы из этой модели определили бы коэффициенты  $c_1, \dots, c_m$ . Однако часто изучаемые процессы оказываются настолько сложными, что теоретическую модель построить не удастся. Как следствие, коэффициенты  $c_i$ , описывающие, по сути дела, математическую модель изучаемого явления, приходится находить из эксперимента. В идеале нам достаточно провести ровно  $m$  экспериментов, подбирая на каждом из этих экспериментов свои значения  $x_{ij}$  управляющих параметров и снимая реакцию  $y_j$  управляющей системы на эти воздействия. В результате мы бы получили систему из  $m$  уравнений с  $m$  неизвестными  $c_m$ , из которой мы бы и определили искомые параметры  $c_i$ , описывающие реакцию системы на изменение управляющих параметров  $x_i$ .

Проблема, однако, состоит в том, что любые измерения в реальном эксперименте проводятся с той или иной погрешностью. Как следствие, как управляющие параметры  $x_{ij}$ , так и реакция системы  $y_j$  в  $j$ -м эксперименте определяются неточно. Поэтому на практике вместо  $m$  проводят  $n \gg m$  экспериментов, на которых получают избыточный набор данных  $x_{ij}, y_j$ . В результате для определения  $c_i$  получается заведомо переопределенная система линейных алгебраических уравнений вида

$$\begin{cases} x_{11}c_1 + x_{12}c_2 + \dots + x_{1m}c_m = y_1 \\ x_{21}c_1 + x_{22}c_2 + \dots + x_{2m}c_m = y_2 \\ \dots \\ x_{n1}c_1 + x_{n2}c_2 + \dots + x_{nm}c_m = y_n \end{cases}$$

Точные решения этой системы искать бессмысленно — условие  $n \gg m$ , как правило, гарантирует нам то, что таких решений полученная система не имеет. Вместо этого числа  $c_i$  находят приближенно — из условия минимизации так называемой невязки, под которой понимается сумма квадратов отклонений левых частей системы от свободных членов:

$$\sum_{k=1}^n (y_k - c_1 x_{k1} - \dots - c_m x_{km})^2 \rightarrow \min. \quad (5)$$

Про коэффициенты  $c_i$ , найденные из условия минимизации невязки, как раз и говорят, что они найдены с помощью метода наименьших квадратов.

**2.4.2.** Казалось бы, для определения коэффициентов  $c_i$  нам нужно владеть навыками решения экстремальных задач вида (5). Однако оказывается, что задача (5) сводится к известной нам задаче о проектировании вектора на линейное подпространство.

Действительно, введем в некотором  $n$ -мерном линейном пространстве  $V$  векторы

$$e_1 := \begin{bmatrix} x_{11} \\ x_{21} \\ \dots \\ x_{n1} \end{bmatrix}, \quad e_2 := \begin{bmatrix} x_{12} \\ x_{22} \\ \dots \\ x_{n2} \end{bmatrix}, \quad \dots \quad e_m := \begin{bmatrix} x_{1m} \\ x_{2m} \\ \dots \\ x_{nm} \end{bmatrix}, \quad f := \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}.$$

Несложно убедиться в том, что сумма

$$\sum_{k=1}^n (y_k - c_1 x_{k1} - \dots - c_m x_{km})^2 = (\mathbf{f} - c_1 \mathbf{e}_1 - \dots - c_m \mathbf{e}_m, \mathbf{f} - c_1 \mathbf{e}_1 - \dots - c_m \mathbf{e}_m),$$

то есть выражается через скалярное произведение вектора  $\mathbf{f} - \mathbf{f}_0$  на самого себя, где

$$\mathbf{f}_0 := c_1 \mathbf{e}_1 + \dots + c_m \mathbf{e}_m.$$

Скалярное же произведение  $(\mathbf{f} - \mathbf{f}_0, \mathbf{f} - \mathbf{f}_0)$  есть не что иное, как квадрат расстояния вектора  $\mathbf{f}$  до вектора  $\mathbf{f}_0$ .

Заметим теперь, что  $\mathbf{f}$  есть некоторый фиксированный вектор пространства  $V$ . При этом все, что мы знаем о пока неизвестном нам векторе  $\mathbf{f}_0$  — это лишь то, что он должен лежать в линейном подпространстве  $V_0$ , натянутом на  $m$  векторов  $\mathbf{e}_1, \dots, \mathbf{e}_m$ . Следовательно, задача минимизации расстояния от вектора  $\mathbf{f}$  до вектора  $\mathbf{f}_0$  равносильна хорошо уже нам известной задаче о поиске кратчайшего расстояния от вектора  $\mathbf{f}$  до линейного подпространства. Мы знаем, в частности, что это расстояние будет минимальным в том и только в том случае, когда вектор  $\mathbf{f}_0$  есть проекция вектора  $\mathbf{f}$  на линейное подпространство. При этом координаты вектора  $\mathbf{f}_0$  находятся из решения системы линейных алгебраических уравнений (3), которая для данных обозначений имеет следующий вид:

$$\begin{cases} c_1(\mathbf{e}_1, \mathbf{e}_1) + c_2(\mathbf{e}_2, \mathbf{e}_1) + \dots + c_m(\mathbf{e}_m, \mathbf{e}_1) &= (\mathbf{f}, \mathbf{e}_1) \\ c_1(\mathbf{e}_1, \mathbf{e}_2) + c_2(\mathbf{e}_2, \mathbf{e}_2) + \dots + c_m(\mathbf{e}_m, \mathbf{e}_2) &= (\mathbf{f}, \mathbf{e}_2) \\ &\dots \\ c_1(\mathbf{e}_1, \mathbf{e}_m) + c_2(\mathbf{e}_2, \mathbf{e}_m) + \dots + c_m(\mathbf{e}_m, \mathbf{e}_m) &= (\mathbf{f}, \mathbf{e}_m) \end{cases}$$

У этой системы уже  $m$  неизвестных и  $m$  уравнений, и из этой системы мы однозначно определим нужные нам коэффициенты  $c_i$ .

**2.4.3.** В качестве простейшего примера полученных выше результатов рассмотрим систему, состояние  $y$  которой зависит только от одного управляющего параметра  $x$ . Иными словами, поведение этой системы описывается единственным линейным уравнением вида

$$y = cx.$$

Нам нужно как-то определить неизвестный коэффициент  $c$ . Предположим, что для его опреде-

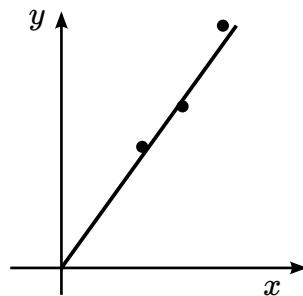


Рис. 6

ления мы провели три эксперимента, и определили из этих экспериментов для трех различных значений  $x_j$  управляющих параметров  $x_1 = 2$ ,  $x_2 = 3$ ,  $x_3 = 4$  три следующих значения функции  $y$ :  $y_1 = 3$ ,  $y_2 = 4$ ,  $y_3 = 6$ . Отложим теперь на плоскости  $(x, y)$  эти три точки (рис. 6).

С геометрической точки зрения мы хотим подобрать тангенс  $c$  угла наклона прямой  $y = cx$ , проходящей через начало координат, так, чтобы сумма квадратов уклонов этой прямой от трех измеренных нами точек был бы минимальным.

Мы свели данную задачу к задаче о проектировании вектора  $\mathbf{f}$  на одномерное линейное подпространство  $V_0$ , натянутое на вектор  $\mathbf{x}$ , где

$$\mathbf{f} = \begin{bmatrix} 3 \\ 4 \\ 6 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}.$$

Коэффициент  $c$  в этой задаче определяется в данном случае из единственного линейного алгебраического уравнения вида

$$(\mathbf{f}, \mathbf{e}_1) = c(\mathbf{e}_1, \mathbf{e}_1) \quad \Longleftrightarrow \quad 42 = c \cdot 29$$

и равен  $42/29$ .

## 3 Линейные операторы

**3.1.** Последний параграф этого небольшого введения в курс линейной алгебры мы посвятим изучению линейных операторов.

**3.1.1.** Как мы с вами знаем, основные операции над векторами в классическом линейном пространстве — это сложение векторов и умножение вектора на скаляр. Когда нам понадобилось ввести такие характеристики векторов, как длина вектора или угол между двумя векторами, мы дополнили линейное пространство до евклидова, добавив операцию скалярного умножения векторов. Однако на практике этих трех операций, чаще всего, не хватает — нам зачастую требуются операции, которые бы позволяли выполнять некоторые преобразования над векторами — например, поворачивать векторы, растягивать их и так далее. Такие операции в линейной алгебре как раз и описываются с помощью линейных операторов.

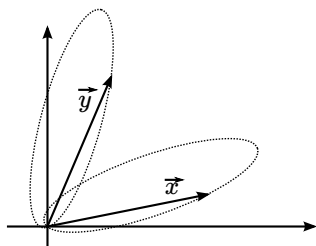


Рис. 7

Прежде чем давать формальное определение линейного оператора, рассмотрим достаточно простой пример. Рассмотрим двумерную плоскость  $\mathbb{R}^2$  и нарисуем на этой плоскости какую-то геометрическую фигуру, например, эллипс (рис. 7). Этот эллипс может служить, например, простейшей математической моделью некоторого механического объекта, например, механической руки. Каждую точку такого эллипса мы можем описать вектором  $\mathbf{x} \in \mathbb{R}^2$ . Предположим теперь, что мы хотим повернуть наш эллипс на некоторый угол  $\alpha$ . С формальной точки зрения такой поворот можно представлять себе так: если до поворота некоторая точка эллипса описывалась вектором  $\mathbf{x}$ , то после поворота та же точка должна описываться некоторым новым

вектором  $\mathbf{y}$ , повернутым на угол  $\alpha$  относительно исходного вектора  $\mathbf{x}$ . С формальной точки зрения такой поворот мы как раз и можем рассматривать как некоторое преобразование  $A$ , переводящее произвольный вектор  $\mathbf{x}$  исходного пространства в некоторый новый вектор  $\mathbf{y}$  этого же линейного пространства:

$$A : \mathbf{x} \mapsto \mathbf{y} =: A\mathbf{x}.$$

Теперь мы готовы дать формальное определение оператора.

**Определение 3.1.** Оператором  $A$  в линейном пространстве  $V$  называется отображение, сопоставляющее любому вектору  $\mathbf{v} \in V$  вектор  $\mathbf{y}$  того же самого пространства.

**3.1.2.** В общем случае преобразование  $A$ , сопоставляющее любому вектору  $\mathbf{x}$  некоторый новый вектор  $\mathbf{y}$ , может быть достаточно сложным. В линейной алгебре изучаются простейшие преобразования такого рода — так называемые линейные операторы.

**Определение 3.2.** Оператор

$$A : \mathbf{x} \mapsto \mathbf{y} =: A\mathbf{x}$$

называется линейным, если выполняются следующие два условия:

1.  $A(\mathbf{x}_1 + \mathbf{x}_2) = A(\mathbf{x}_1) + A(\mathbf{x}_2)$ ;
2.  $A(\lambda \cdot \mathbf{x}) = \lambda \cdot A(\mathbf{x})$ .

**3.1.3.** Прежде чем приводить примеры линейных операторов, давайте поймем, как же нам эти операторы задавать. Основная проблема в задании такого оператора состоит в следующем. При определении оператора мы говорили, что оператор — это некоторое правило, которое *любому* вектору  $\mathbf{x}$  линейного пространства сопоставляет какой-то другой вектор  $\mathbf{y}$  того же пространства. Но векторов в любом пространстве бесконечно много, и на первый взгляд совершенно непонятно, как же нам описать для каждого из них правило, переводящее их в другие векторы.

Здесь нам на помощь вновь приходит понятие базиса. Мы знаем, что в любом  $n$ -мерном линейном пространстве  $V$  у нас существует базис, то есть набор  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  из  $n$  линейно независимых векторов. При этом мы любой вектор  $\mathbf{x} \in V$  можем по этому базису разложить:

$$\mathbf{x} = \xi^1 \cdot \mathbf{e}_1 + \xi^2 \cdot \mathbf{e}_2 + \dots + \xi^n \cdot \mathbf{e}_n.$$

Напомним, что коэффициенты  $\xi^i$  в разложении вектора  $\mathbf{x}$  по базису носят название координат вектора  $\mathbf{x}$  в базисе  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ .

Поддействуем теперь оператором  $A$  на вектор  $\mathbf{x}$ . В силу линейности оператора мы получаем, что

$$A\mathbf{x} = \xi^1 \cdot A\mathbf{e}_1 + \xi^2 \cdot A\mathbf{e}_2 + \dots + \xi^n \cdot A\mathbf{e}_n.$$

Рассмотрим теперь образы  $\mathbf{g}_i = A\mathbf{e}_i$  базисных векторов под действием линейного оператора  $A$ . Предположим, что мы знаем, как для заданного оператора  $A$  выглядят векторы  $\mathbf{g}_i$ , то есть мы знаем, в какие именно векторы перешли у нас базисные векторы  $\mathbf{e}_i$  под действием оператора  $A$ . Но тогда мы будем знать, куда перешел у нас *любой* вектор  $\mathbf{x}$  линейного пространства  $V$ . Действительно, если исходный вектор  $\mathbf{x}$  определялся своими координатами  $\xi^i$  в разложении по базису  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ , то его образ определяется теми же самыми координатами  $\xi^i$ , но уже в разложении этого образа по системе векторов  $\{\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n\}$ . Иными словами, зная, куда у нас

при отображении  $A$  перейдут  $n$  базисных векторов, мы сможем определить, куда у нас перейдет совершенно произвольный вектор  $\mathbf{x}$  нашего пространства.

Посмотрим теперь на образы  $\mathbf{g}_k := A\mathbf{e}_k$  базисных векторов  $\mathbf{e}_k$ . Так как эти образы вновь представляют собой некоторые векторы того же самого пространства  $V$ , а векторы  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  образуют в этом пространстве базис, то мы любой из векторов  $\mathbf{g}_k$  можем по этому базису разложить:

$$\mathbf{g}_k = a_{1k} \cdot \mathbf{e}_1 + a_{2k} \cdot \mathbf{e}_2 + \dots + a_{nk} \cdot \mathbf{e}_n.$$

У нас имеется  $n$  векторов  $\mathbf{g}_k$ , и каждый из них определяется набором из  $n$  чисел  $a_{ik}$ . Следовательно, всю совокупность векторов  $\mathbf{g}_k$  можно определить  $n^2$  числами  $a_{ik}$ , которые удобно записать в виде матрицы

$$\mathcal{A} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Задание этих  $n^2$  чисел, или, что то же самое, задание матрицы  $\mathcal{A}$  полностью определяет нам действие оператора  $A$  на любом из векторов линейного пространства  $V$ . Обратно, любой оператор  $A$  при фиксированном базисе  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  однозначно определяет нам некоторую матрицу  $\mathcal{A}$ . Иными словами, мы любой линейный оператор можем описывать с помощью матрицы — так называемой *матрицы линейного оператора  $A$  в базисе  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$* .

**3.1.4.** Приведем теперь наиболее популярные и полезные на практике примеры линейных операторов. Все эти примеры мы будем для наглядности иллюстрировать на примере векторов на плоскости. При этом мы будем считать, что у нас на множестве таких векторов задано также и скалярное произведение, а в качестве базиса выбирать ортонормированный базис  $\{\mathbf{e}_1, \mathbf{e}_2\}$  евклидова пространства  $\mathbb{R}^2$ .

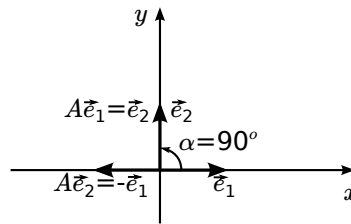


Рис. 8

**Пример 3.3.** Рассмотрим оператор  $A$  поворота векторов на угол в  $90^\circ$  (рис. 8). Для задания оператора  $A$  нам достаточно понять, куда у нас переходят два базисных вектора  $\mathbf{e}_1$  и  $\mathbf{e}_2$ . Как хорошо видно из рис. 8, вектор  $\mathbf{e}_1$  переходит при повороте на  $90^\circ$  в вектор  $\mathbf{e}_2$ , а вектор  $\mathbf{e}_2$  переходит в вектор  $-\mathbf{e}_1$ :

$$A\mathbf{e}_1 = \mathbf{e}_2 = 0 \cdot \mathbf{e}_1 + 1 \cdot \mathbf{e}_2, \quad A\mathbf{e}_2 = -\mathbf{e}_1 = -1 \cdot \mathbf{e}_1 + 0 \cdot \mathbf{e}_2.$$

Каждый вектор коэффициентов в разложении вектора  $\mathbf{g}_i$  по базису задает нам  $i$ -й столбец матрицы  $\mathcal{A}$  оператора. Следовательно, в рассматриваемом примере эта матрица имеет следующий вид:

$$\mathcal{A} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

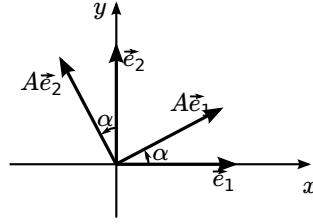


Рис. 9

**Пример 3.4.** Обобщением приведенного выше примера является поворот на произвольный угол  $\alpha$  (рис. 9). Вектор  $\mathbf{e}_1$  переходит при таком повороте в вектор

$$A\mathbf{e}_1 = \cos \alpha \cdot \mathbf{e}_1 + \sin \alpha \cdot \mathbf{e}_2,$$

а вектор  $\mathbf{e}_2$  переходит в вектор

$$A\mathbf{e}_2 = -\sin \alpha \cdot \mathbf{e}_1 + \cos \alpha \cdot \mathbf{e}_2.$$

Как следствие, матрица  $\mathcal{A}$  оператора  $A$  поворота на угол  $\alpha$  имеет вид

$$\mathcal{A} = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

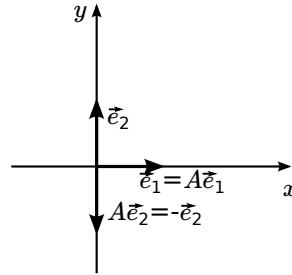


Рис. 10

**Пример 3.5.** Рассмотрим теперь оператор  $A$  отражения относительно одной из осей, например, относительно оси абсцисс (рис. 10). В этом случае вектор  $\mathbf{e}_1$  остается на месте, а вектор  $\mathbf{e}_2$  меняется на противоположный:

$$A\mathbf{e}_2 = -\mathbf{e}_2.$$

Поэтому матрица  $\mathcal{A}$  оператора  $A$  в этом случае записывается в виде

$$\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

**Пример 3.6.** Следующим достаточно часто встречающимся на практике оператором является оператор  $A$  растяжения. На рис 11. приведен пример оператора растяжения на одно и то же число, равное двум. Так как в этом случае

$$A\mathbf{e}_1 = 2\mathbf{e}_1, \quad A\mathbf{e}_2 = 2\mathbf{e}_2,$$

то матрица  $\mathcal{A}$  этого оператора  $A$  выглядит так:

$$\mathcal{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$



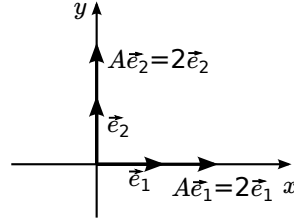


Рис. 11

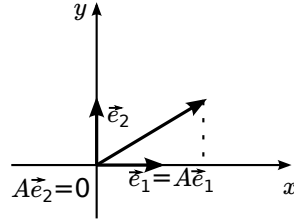


Рис. 12

**Пример 3.7.** В заключение рассмотрим так называемый оператор  $A$  проектирования вектора на подпространство. Рис. 12 иллюстрирует действие простейшего оператора такого рода — оператор проектирования на одномерное подпространство, натянутое на базисный вектор  $\mathbf{e}_1$ . При таком действии сам вектор  $\mathbf{e}_1$  не меняется, а ортогональный к нему вектор  $\mathbf{e}_2$  обращается в  $\mathbf{0}$ . Матрица  $\mathcal{A}$  оператора проектирования имеет, очевидно, следующий вид:

$$\mathcal{A} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}.$$

**3.2.** Поговорим теперь о действиях над линейными операторами.

**3.2.1.** Довольно естественным выглядит тот факт, что мы можем перемножать линейные операторы между собой. Действительно, предположим, что мы хотим совершить вначале некоторое действие, описываемое оператором  $B$  (например, поворот на угол  $\alpha$ ), а затем совершить второе действие, описываемое оператором  $A$  (например, растяжение на число  $c$ ). Тогда сложное действие, которое состоит вначале в повороте всех векторов на угол  $\alpha$ , а затем в растяжении этих векторов на одно и то же число  $c$ , описывается оператором  $C$ , равным произведению векторов  $A$  и  $B$ :

$$C\mathbf{x} = (A \cdot B)\mathbf{x} := A(B\mathbf{x})$$

(вначале применяем к вектору  $\mathbf{x}$  оператор  $B$ , а затем на полученный вектор  $B\mathbf{x}$  действуем оператором  $A$ ).

Несложно убедиться в том, что произведение пары линейных операторов  $A$  и  $B$  вновь будет линейным же оператором. Далее, если оператор  $A$  задается матрицей  $\mathcal{A}$ , а оператор  $B$  — матрицей  $\mathcal{B}$ , то оператор  $C$  их произведения задается матрицей  $\mathcal{C}$ , равной произведению матриц  $\mathcal{A}$  и  $\mathcal{B}$ . Например, в двумерном случае имеем

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \quad \Longrightarrow \quad \mathcal{C} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

**3.2.2.** Понятно, что наряду с умножением операторов мы можем определить и сложение двух операторов.

**Определение 3.8.** Суммой  $A + B$  операторов  $A$  и  $B$  называется оператор, сопоставляющий каждому вектору  $\mathbf{x}$  вектор  $A\mathbf{x} + B\mathbf{x}$ .

Понятно, что матрица  $\mathcal{C}$  суммы операторов есть сумма матриц  $\mathcal{A}$  и  $\mathcal{B}$  исходных операторов  $A$  и  $B$ .

**3.3.** Оказывается, что с любым оператором  $A$  естественным образом связаны два подпространства — ядро  $\text{Ker } A$  и образ  $\text{Im } A$  этого оператора.

**3.3.1.** Начнем с определений этих двух понятий.

**Определение 3.9.** Ядром  $\text{Ker } A$  оператора  $A$  называется множество всех векторов пространства  $V$ , образы которых совпадают с нулевым вектором  $\mathbf{0}$ :

$$\text{Ker } A := \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}.$$

**Определение 3.10.** Образом  $\text{Im } A$  оператора  $A$  называется множество всех векторов вида  $A\mathbf{x}$ , то есть множество векторов  $\mathbf{y}$ , которые получаются в результате действия оператора  $A$  на все векторы исходного пространства  $V$ :

$$\text{Im } A := \{\mathbf{y} : \mathbf{y} = A\mathbf{x}, \mathbf{x} \in V\}.$$

**3.3.2.** В качестве примера давайте рассмотрим оператор  $A$  проектирования векторов двумерного пространства  $\mathbb{R}^2$  на ось абсцисс (рис. 12). Для любого вектора

$$\mathbf{x} = \xi_1 \mathbf{e}_1 + \xi_2 \mathbf{e}_2$$

он сохраняет без изменения проекцию этого вектора на ось абсцисс и обнуляет компоненту этого вектора, связанную с проекцией на ось ординат. Как следствие, ядром такого оператора являются все векторы, параллельные вектору  $\mathbf{e}_2$ : все такие векторы будут оператором  $A$  переводиться в вектор  $\mathbf{0}$ . Иными словами, ядро

$$\text{Ker } A = \{\mathbf{e}_2\}$$

есть одномерное подпространство, натянутое на вектор  $\mathbf{e}_2$ . Образ же этого оператора — это все векторы, лежащие на оси абсцисс. Такие векторы образуют линейное подпространство

$$\text{Im } A = \{\mathbf{e}_1\},$$

натянутое на вектор  $\mathbf{e}_1$ .

**3.4.** В заключение нашего небольшого введения в линейную алгебру давайте очень кратко поговорим о таких важных понятиях, как собственный вектор и собственное значение оператора  $A$ .

**3.4.1.** Начнем с определения.

**Определение 3.11.** Вектор  $\mathbf{e} \neq \mathbf{0}$  линейного пространства  $V$  называется *собственным вектором* оператора  $A$ , если действие оператора  $A$  на этот вектор сводится к умножению вектора  $\mathbf{e}$  на некоторое вещественное число  $\lambda$ :

$$A\mathbf{e} = \lambda \cdot \mathbf{e}.$$

При этом само это число  $\lambda$  называется *собственным значением* оператора  $A$ .

Смысл этих понятий можно себе представлять так. В общем случае оператор  $A$  может быть достаточно сложным, и действие этого оператора на векторы линейного пространства часто достаточно тяжело описать. Однако даже для сколь угодно сложного оператора среди множества всех векторов есть выделенное подмножество векторов — собственных векторов этого оператора, на которые данный оператор действует особенно просто. Именно, действие оператора на собственные векторы сводится к растяжению или сжатию векторов из этого подмножества.

**Замечание 3.12.** Из определения собственного вектора следует, что любой такой вектор определен с точностью до константы. Действительно, пусть  $\mathbf{e}$  есть некоторый собственный вектор оператора  $A$ . Тогда любой вектор вида  $c \cdot \mathbf{e}$ , где  $c$  — произвольное вещественное число, также является собственным вектором оператора  $A$ :

$$A\mathbf{e} = \lambda \cdot \mathbf{e} \quad \implies \quad A(c \cdot \mathbf{e}) = c \cdot (A\mathbf{e}) = c \cdot (\lambda \cdot \mathbf{e}) = \lambda \cdot (c \cdot \mathbf{e}).$$

Поэтому для определенности часто считают, что любой собственный вектор имеет единичную длину.

**3.4.2.** В качестве примера понятия собственного вектора можно привести хорошо известную в механике теорему Эйлера. Предположим, что у нас имеется твердое тело, закрепленное в некоторой точке, и мы изучаем вращение такого тела вокруг этой неподвижной точки. В принципе, траектория вращения каждой из точек твердого тела может быть сколь угодно сложной. Зафиксируем, однако, два момента — начальный момент, в котором наше тело было еще неподвижным, а также конечный момент, в который мы тело вращать прекратили. Теорема Эйлера утверждает тогда, что любое такое перемещение тела мы можем рассматривать как результат вращения этого тела вокруг выделенной оси, проходящей через закрепленную точку нашего твердого тела, из начального положения тела в конечное.

На языке линейной алгебры теорема Эйлера означает следующее: любой оператор трехмерного вращения имеет собственный вектор  $\mathbf{e}$ , причем соответствующее этому собственному вектору  $\mathbf{e}$  собственное число  $\lambda = 1$ . Физический смысл этого собственного вектора следующий: вектор  $\mathbf{e}$  как раз и задает нам ту самую ось вращения, вокруг которой мы должны повернуть твердое тело так, чтобы перевести его из начального положения в конечное.