# Матрицы и векторы

Алексей Померанцев

# Содержание

# 1 Введение

В этом документе собраны основные сведения из алгебры матриц и векторов, которые используются в хемометрике. Приведенный текст не может служить учебником по матричной алгебре — он скорее является конспектом, справочником в этой области. Более глубокое и систематическое изложение может быть найдено в литературе.

Текст разбит на две части названные — “Базовые сведения” и “Дополнительная информация”. В первой части изложены положения, минимально необходимые для понимания хемометрики, а во второй части — факты,которые необходимо знать для более глубокого постижения методов многомерного анализа. Изложение иллюстрируется примерами, выполненными в рабочей книге Excel, Matrix.xls, которая сопровождает этот документ.

Ссылки на примеры помещены в текст как объекты Excel. Эти примеры имеют абстрактный характер, они никак не привязаны к задачам аналитической химии. Реальные примеры использования матричной алгебры в хемометрике рассмотрены в других текстах, посвященных разнообразным хемометрическим приложениям.

Большинство измерений, проводимых в аналитической химии, являются не прямыми, а косвенными. Это означает, что в эксперименте вместо значения искомого аналита $C$ (концентрации) получается другая величина $x$ (сигнал), связанная, но не равная $C$, т.е. $x(C) \neq$ . Как правило, вид зависимости $x(C)$ не известен, однако, к счастью, в аналитической химии большинство измерений пропорциональны. Это означает, что при увеличении концентрации  в а раз, сигнал $X$ увеличится на столько же., т.е. $x(aC) = ax(C)$. Кроме того, сигналы еще и аддитивны, так что сигнал от пробы, в которой присутствуют два вещества с концентрациями $C_1$ и $C_2$, будет равен сумме сигналов от каждого компонента, т.е. $x(C_1+C_2) = x(C_1)+x(C_2)$.

Пропорциональность и аддитивность вместе дают линейность. Можно привести много примеров, иллюстрирующих принцип линейности, но достаточно упомянуть два самых ярких примера — хроматографию и спектроскопию. Вторая особенность, присущая эксперименту в аналитической химии — это многоканальность. Современное аналитическое оборудование одновременно измеряет сигналы для многих каналов. Например, измеряется интенсивность пропускания света сразу для нескольких длин волн, т.е. спектр. Поэтому в эксперименте мы имеем дело со множеством сигналов $x_1, x_2, ...., x_n$, характеризующих набор концентраций $C_1, C_2, ..., C_m$ веществ, присутствующих в изучаемой системе.

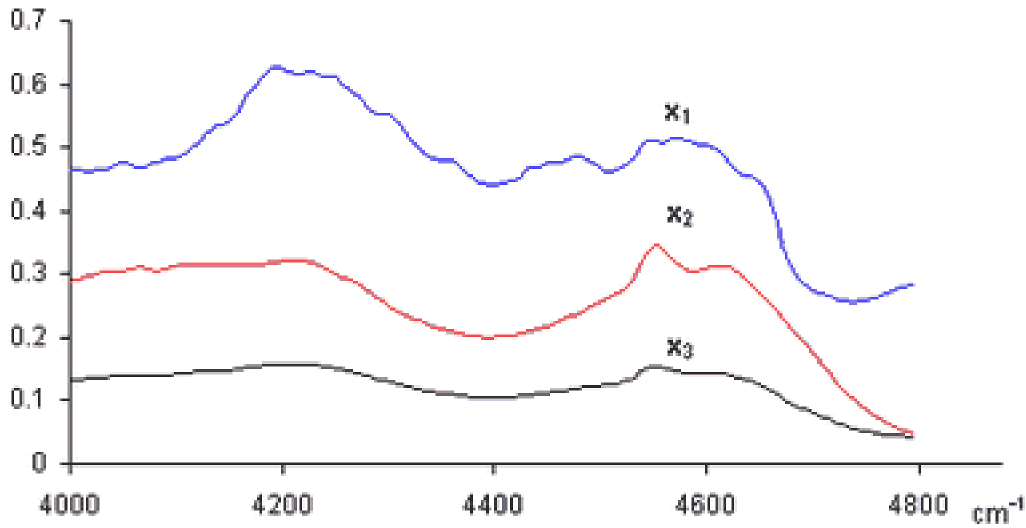


**Рис. 1.1.** Спектры

Итак, аналитический эксперимент характеризуется линейностью и многомерностью. Поэтому удобно рассматривать экспериментальные данные как векторы и матрицы и манипулировать с ними, используя аппарат матричной алгебры. Плодотворность такого подхода иллюстрирует пример, показанный на Рис. 1, где представлены три спектра, снятые для 200 длин волн от 4000 до 4796 cm−1. Первый ($x_1$) и второй ($x_2$) спектры получены для стандартных образцов, в которых концентрация двух веществ $A$ и $B$, известны: в первом образце $[A] = 0.5, [B] = 0.1$, а во втором образце $[A] = 0.2, [B] = 0.6$. Что можно сказать о новом, неизвестном образце, спектр которого обозначен $x_3$?

Рассмотрим три экспериментальных спектра $x_1$, $x_2$ и $x_3$ как три вектора размерности 200. Средствами линейной алгебры можно легко показать, что $x_3 = 0.1x_1 + 0.3x_2$, поэтому в третьем образце очевидно присутствуют только вещества $A$ и $B$ в концентрациях $[A] = 0.5{\times}0.1{+}0.2{\times}0.3 = 0.11$ и $[B] = 0.1{\times}0.1{+}0.6{\times}0.3 = 0.19$.

# 2 Базовые сведения

## 2.1 Матрицы

Матрицей называется прямоугольная таблица чисел, например

$$\mathbf{A}=\begin{vmatrix} 1.2 & -5.3 & 0.25 \\ 10.2 & 1.5 & -7.5 \\ 2.3 & -1.2 & 5.6 \\ 4.5 & -0.8 & 9.5 \end{vmatrix}$$

**Рис. 2.1.** Матрица

Матрицы обозначаются заглавными полужирными буквами ($\mathbf{A}$), а их элементы — соответствующими строчными буквами с индексами, т.е. $a_{ij}$. Первый индекс нумерует строки, а второй — столбцы. В хемометрике принято обозначать максимальное значение индекса той же буквой, что и сам индекс, но заглавной. Поэтому матрицу **A** можно также записать как
$a_{ij}, i = 1, ..., I; j = 1, ..., J$. Для приведенной в примере матрицы $I = 4$, $J = 3$ и $a_{23} = -7.5$.

Пара чисел $I$ и $J$ называется размерностью матрицы и обознается как $I \times J$. Примером матрицы в хемометрике может служить набор спектров, полученный для $I$ образцов на $J$ длинах волн.

## 2.2 Простейшие операции с матрицами

Матрицы можно *умножать на числа*. При этом каждый элемент умножается на это число. Например:

$$3 \ * \begin{vmatrix} 1.2 & -5.3 & 0.25 \\ 10.2 & 1.5 & -7.5 \\ 2.3 & -1.2 & 5.6 \\ 4.5 & -0.8 & 9.5 \end{vmatrix} = \begin{vmatrix} 3.6 & -16 & 0.75 \\ 30.6 & 4.5 & -23 \\ 6.9 & -3.6 & 16.8 \\ 13.5 & -2.4 & 28.5 \end{vmatrix}$$

**Рис. 2.2.** Умножение матрицы на число

Две матрицы одинаковой размерности можно поэлементно *складывать и вычитать*. Например:

$$\begin{vmatrix} 1.2 & -5.3 \\ 10.2 & 1.5 \\ 2.3 & -1.2 \\ 4.5 & -0.8 \end{vmatrix} + \begin{vmatrix} -2.5 & 13.8 \\ -4.2 & 11.2 \\ -3.5 & 4.8 \\ 1.4 & 6.5 \end{vmatrix} = \begin{vmatrix} -1.3 & 8.5 \\ 6 & 12.7 \\ -1.2 & 3.6 \\ 5.9 & 5.7 \end{vmatrix}$$

**Рис. 2.3.** Сложение матриц

В результате умножения на число и сложения получается матрица той же размерности.

Нулевой матрицей называется матрица, состоящая из нулей. Она обозначается $\mathbf{O}$. Очевидно, что $\mathbf{A}+\mathbf{O} = \mathbf{A}$, $\mathbf{A} - \mathbf{A} = \mathbf{O}$ и $\mathbf{OA} = \mathbf{O}$.

Матрицу можно *транспонировать*. При этой операции матрица переворачивается, т.е. строки и столбцы меняются местами. Транспонирование обозначается штрихом, $\mathbf{A}'$ или индексом $\mathbf{A}^{\mathrm{t}}$. Таким образом, если $\mathbf{A} = \{a_{ij}, i = 1, ..., I; j = 1, ..., J\}$, то $\mathbf{A}^{\mathrm{t}} = \{a_{ji}, j = 1, ..., J; i = 1, ..., I\}$. Например:

$$\mathbf{A}=\begin{vmatrix} 1.2 & -5.3 & 0.25 \\ 10.2 & 1.5 & -7.5 \\ 2.3 & -1.2 & 5.6 \\ 4.5 & -0.8 & 9.5 \end{vmatrix} \qquad \mathbf{A}^{\mathrm{t}}=\begin{vmatrix} 1.2 & 10.2 & 2.3 & 4.5 \\ -5.3 & 1.5 & -1.2 & -0.8 \\ 0.25 & -7.5 & 5.6 & 9.5 \end{vmatrix}$$

**Рис. 2.4.** Транспонирование матрицы

Очевидно, что $(\mathbf{A}^{\mathrm{t}})^{\mathrm{t}} = \mathbf{A}, (\mathbf{A} + \mathbf{B})^{\mathrm{t}} = \mathbf{A}^{\mathrm{t}} + \mathbf{B}^{\mathrm{t}}$.

## 2.3 Умножение матриц

Матрицы можно перемножать, но только в том случае, когда они имеют соответствующие размерности. Почему это так, будет ясно из определения. Произведением матрицы $\mathbf{A}$, размерностью $I \times K$, и матрицы $\mathbf{B}$, размерностью $K \times J$, называется матрица $\mathbf{C}$, размерностью $I \times J$, элементами которой являются числа

$$c_{ij} = \sum_{k=1}^{K} a_{ik} b_{kj} = 1$$

Таким образом для произведения $\mathbf{AB}$ необходимо, чтобы число столбцов в левой матрице $\mathbf{A}$ было равно числу строк в правой матрице $\mathbf{B}$. Пример произведения матриц:

$$
\left|\begin{array}{rr} 2.2 & 5.8 \\ 1.5 & 6.2 \\ -2.5 & 7.4 \\ 4.2 & -2.3 \end{array}\right| * \left|\begin{array}{rrr} 5.2 & 3.2 & 3.8 \\ 6.8 & -7.5 & -1.9 \end{array}\right| = \left|\begin{array}{rrr} 50.88 & -36.5 & -2.66 \\ 49.96 & -41.7 & -6.08 \\ 37.32 & -63.5 & -23.6 \\ 6.2 & 30.69 & 20.33 \end{array}\right|
$$

**Рис. 2.5.** Произведение матриц

Правило перемножения матриц можно сформулировать так. Для того, чтобы найти элемент матрицы **C**, стоящий на пересечении $i$-ой строки и $j$-ого столбца ($c_{ij}$) надо поэлементно перемножить $i$-ую строку первой матрицы **A** на $j$-ый столбец второй матрицы **B** и сложить все результаты. Так в показанном примере, элемент из третьей строки и второго столбца, получается как сумма поэлементных произведений третьей строки **A** и второго столбца **B**.

$$
\left|\begin{array}{rr} \# & \# \\ \# & \# \\ -2.5 & 7.4 \\ \# & \# \end{array}\right| * \left|\begin{array}{rrr} \# & 3.2 & \# \\ \# & -7.5 & \# \end{array}\right| = \left|\begin{array}{rrr} \# & \# & \# \\ \# & \# & \# \\ \# & -63.5 & \# \\ \# & \# & \# \end{array}\right|
$$

**Рис. 2.6.** Элемент произведения матриц

Произведение матриц зависит от порядка, т.е. $\mathbf{AB} \neq \mathbf{BA}$, хотя бы по соображениям размерности. Говорят, что оно некоммутативно. Однако произведение матриц ассоциативно. Это означает, что $\mathbf{ABC} = (\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$. Кроме того, оно еще и дистрибутивно, т.е. $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$. Очевидно, что $\mathbf{AO} = \mathbf{O}$.

## 2.4 Квадратные матрицы

Если число столбцов матрицы равно числу ее строк ($I = J = N$), то такая матрица называется *квадратной*. В этом разделе мы будем рассматривать только такие матрицы. Среди этих матриц можно выделить матрицы, обладающие особыми свойствами.

Единичной матрицей (обозначается **I**, а иногда **E**) называется матрица, у которой все элементы равны нулю, за исключением диагональных, которые равны 1, т.е.

$$
\mathbf{I} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix}
$$

Очевидно $\mathbf{AI} = \mathbf{IA} = \mathbf{A}$.

Матрица называется *диагональной*, если все ее элементы, кроме диагональных ($a_{ii}$) равны нулю. Например

$$\begin{vmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -4 \end{vmatrix}$$

**Рис. 2.7.** Диагональная матрица

Матрица $\mathbf{A}$ называется *верхней треугольной*, если все ее элементы, лежащие ниже диагонали, равны нулю, т.е. $a_{ij} = 0$, при $i > j$. Например

$$\begin{vmatrix} 1 & 5 & 3 \\ 0 & 2 & -1 \\ 0 & 0 & -4 \end{vmatrix}$$

**Рис. 2.8.** Верхняя треугольная матрица

Аналогично определяется и нижняя треугольная матрица.

Матрица $\mathbf{A}$ называется *симметричной*, если $\mathbf{A}^{\mathrm{t}} = \mathbf{A}$. Иными словами $a_{ij} = a_{ji}$. Например

$$\begin{vmatrix} 1 & 5 & 3 \\ 5 & 2 & -1 \\ 3 & -1 & -4 \end{vmatrix}$$

**Рис. 2.9.** Симметричная матрица

Матрица $\mathbf{A}$ называется *ортогональной*, если $\mathbf{A}^{\mathrm{t}}\mathbf{A} = \mathbf{A}\mathbf{A}^{\mathrm{t}} = \mathbf{I}$.

Матрица называется *нормальной* если $\mathbf{A}^{\mathrm{t}}\mathbf{A} = \mathbf{A}\mathbf{A}^{\mathrm{t}}$.

## 2.5 След и определитель

Следом квадратной матрицы $\mathbf{A}$ (обозначается $\mathrm{Tr}(\mathbf{A})$ или $\mathrm{Sp}(\mathbf{A})$) называется сумма ее диагональных элементов.

Например,

$$\mathrm{Sp}\begin{vmatrix} 1 & 5 & 3 \\ 5 & 2 & -1 \\ 3 & -1 & -4 \end{vmatrix} = -1$$

**Рис. 2.10.** След матрицы

Очевидно, что

$$\mathrm{Sp}(\alpha\mathbf{A}) = \alpha\mathrm{Sp}(\mathbf{A})$$

$$\mathrm{Sp}(\mathbf{A} + \mathbf{B}) = \mathrm{Sp}(\mathbf{A}) + \mathrm{Sp}(\mathbf{B})$$

Можно показать, что

$$\mathrm{Sp}(\mathbf{A}) = \mathrm{Sp}(\mathbf{A}^{\mathrm{t}}), \mathrm{Sp}(\mathbf{I}) = N$$

а также, что

$$\mathrm{Sp}(\mathbf{AB}) = \mathrm{Sp}(\mathbf{BA})$$

Другой важной характеристикой квадратной матрицы является ее определитель (обозначается $\det(\mathbf{A})$). Определение определителя в общем случае довольно сложно, поэтому мы начнем с простейшего варианта — матрицы $\mathbf{A}$ размерностью $(2 \times 2)$. Тогда

$$\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}$$

Для матрицы (3×3) определитель будет равен

$$\det(\mathbf{A}) = a_{11}a_{22}a_{33} - a_{13}a_{21}a_{31} + a_{12}a_{23}a_{31} - a_{12}a_{21}a_{33} + a_{13}a_{21}a_{31} - a_{11}a_{23}a_{32}$$

В случае матрицы $(N \times N)$ определитель вычисляется как сумма $1 \cdot 2 \cdot 3 \cdot \ldots \cdot N = N!$ слагаемых, каждый из которых равен

$$(-1)^r a_{1k_1} a_{2k_2} \dots a_{Nk_N}$$

Индексы $k_1, k_2, \dots, k_N$ определяются как всевозможные упорядоченные перестановки $r$ чисел в наборе $(1, 2, \dots, N)$. Вычисление определителя матрицы — это сложная процедура, которую на практике осуществляется с помощью специальных программ. Например,

$$\det \begin{vmatrix} 1.2 & -5.3 & 0.25 & 2.2 & 13.8 \\ 10.2 & 1.5 & -7.5 & 1.5 & 11.2 \\ 2.3 & -1.2 & 5.6 & -2.5 & 4.8 \\ 4.5 & -0.8 & 9.5 & 4.2 & 6.5 \\ 2.5 & 8.2 & -3.5 & 5.9 & 1.1 \end{vmatrix} = 56044.308595$$

**Рис. 2.11.** Определитель матрицы

Отметим только очевидные свойства:

$$\det(\mathbf{I}) = 1, \det(\mathbf{A}) = \det(\mathbf{A}^{\mathrm{t}})$$

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B})$$

## 2.6 Векторы

Если матрица состоит только из одного столбца ($J = 1$), то такой объект называется *вектором.* Точнее говоря, вектором-столбцом. Например:

$$\mathbf{a} = \begin{bmatrix} 1.2 \\ 10.2 \\ 2.3 \\ 4.5 \end{bmatrix}$$

Можно рассматривать и матрицы, состоящие из одной строки, например:

$$\mathbf{b} = \begin{bmatrix} 1.2 & -5.3 & 0.25 \end{bmatrix}$$

Этот объект также является вектором, но *вектором-строкой.* При анализе данных важно понимать, с какими векторами мы имеем дело — со столбцами или строками. Так спектр, снятый для одного образца можно рассматривать как вектор-строку. Тогда набор спектральных интенсивностей на какой-то длине волны для всех образцов нужно трактовать как вектор-столбец.

Размерностью вектора называется число его элементов.

Ясно, что всякий вектор-столбец можно превратить в вектор-строку транспонированием, т.е.

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{bmatrix}^{\mathrm{t}} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \end{bmatrix}$$

В тех случаях, когда форма вектора специально не оговаривается, а просто говорится вектор, то имеют в виду вектор-столбец. Мы тоже будем придерживаться этого правила. Вектор обозначается строчной прямой полужирной буквой. Нулевым вектором называется вектор, все элементы которого раны нулю. Он обозначается $\mathbf{0}$.

## 2.7 Сложение и умножение векторов

Векторы можно складывать и умножать на числа так же, как это делается с матрицами. Например,

$$\begin{vmatrix} 1.2 \\ 10.2 \\ 2.3 \end{vmatrix} + \begin{vmatrix} 2.3 \\ -1.2 \\ 5.6 \end{vmatrix} = \begin{vmatrix} 3.5 \\ 9 \\ 7.9 \end{vmatrix} \qquad 3 * \begin{vmatrix} 2.3 \\ -1.2 \\ 5.6 \end{vmatrix} = \begin{vmatrix} 6.9 \\ -3.6 \\ 16.8 \end{vmatrix}$$

**Рис. 2.12.** Операции с векторами

Два вектора $\mathbf{x}$ и $\mathbf{y}$ называются *колинеарными*, если существует такое число $\alpha$, что $\alpha\mathbf{x} = \mathbf{y}$.

Два вектора одинаковой размерности $N$ можно перемножить. Пусть имеются два вектора $\mathbf{x} = (x_1, x_2, ..., x_N)^{\mathrm{t}}$ и $\mathbf{y} = (y_1, y_2, ..., y_N)$t. Руководствуясь правилом перемножения “строка на столбец”, мы можем составить из них два произведения: $\mathbf{x}^{\mathrm{t}}\mathbf{y}$ и $\mathbf{x}\mathbf{y}^{\mathrm{t}}$.

Первое произведение:

$$
\mathbf{x}^{\mathrm{t}}\mathbf{y} = \begin{bmatrix} x_1 & x_2 & \dots & x_N \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_N \end{bmatrix} = x_1 y_1 + x_2 y_2 + \dots + x_N y_n
$$

называется *скалярным* или *внутренним*. Его результат — это число. Для него также используется обозначение $(\mathbf{x}, \mathbf{y}) = \mathbf{x}^{\mathrm{t}}\mathbf{y}$. Например,

$$
\begin{vmatrix} 1 & 5 & 2 \end{vmatrix} * \begin{vmatrix} 3 \\ -2 \\ 5 \end{vmatrix} = 3
$$

**Рис. 2.13.** Внутреннее (скалярное) произведение

Второе произведение:

$$
\mathbf{x}\mathbf{y}^{\mathrm{t}} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_N \end{bmatrix} \begin{bmatrix} y_1 & y_2 & \dots & y_N \end{bmatrix} = \begin{bmatrix} x_1 y_1 & x_1 y_2 & \dots & x_1 y_N \\ x_2 y_1 & x_2 y_2 & \dots & x_2 y_N \\ \dots & \dots & \dots & \dots \\ x_N y_1 & x_N y_2 & \dots & x_N y_N \end{bmatrix}
$$

называется *внешним*. Его результат — это матрица размерности $(N \times N)$. Например,

$$
\begin{vmatrix} 3 \\ -2 \\ 5 \end{vmatrix} * \begin{vmatrix} 1 & 5 & 2 \end{vmatrix} = \begin{vmatrix} 3 & 15 & 6 \\ -2 & -10 & -4 \\ 5 & 25 & 10 \end{vmatrix}
$$

**Рис. 2.14.** Внешнее произведение

Векторы, скалярное произведение которых равно нулю, называются *ортогональными*.

## 2.8 Норма вектора и угол между векторами

Скалярное произведение вектора самого на себя называется скалярным квадратом. Эта величина

$$(\mathbf{x}, \mathbf{x}) = \mathbf{x}^{\mathrm{t}}\mathbf{x} = \sum_{n=1}^{N} x_n^2$$

определяет квадрат *длины* вектора x. Для обозначения длины (называемой также *нормой* вектора) используется обозначение

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

Например,

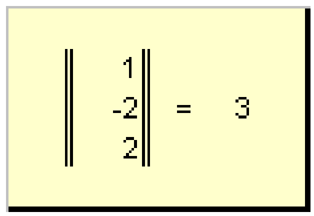


**Рис. 2.15.** Норма вектора

Вектор единичной длины ($\|\mathbf{x}\| = 1$) называется *нормированным*. Ненулевой вектор ($\mathbf{x} \neq 0$) можно нормировать, разделив его на длину, т.е. $\mathbf{x} = \|\mathbf{x}\|(\mathbf{x}\|\mathbf{x}\|) = \|\mathbf{x}\|\mathbf{e}$. Здесь $\mathbf{e} = \mathbf{x}/\|\mathbf{x}\|$ — нормированный вектор.

Скалярное произведение определяет и угол $\phi$ между двумя векторами $\mathbf{x}$ и $\mathbf{y}$:

$$\cos\phi = \frac{(\mathbf{x}, \mathbf{y})}{\|\mathbf{x}\|\|\mathbf{y}\|}$$

Если вектора ортогональны, то $\cos\phi = 0$ и $\phi = \pi/2$, а если они колинеарны, то $\cos\phi = 1$ и $\phi = 0$.

Векторы называются *ортонормированными*, если все они нормированы и попарно ортогональны.

## 2.9 Векторное представление матрицы

Каждую матрицу $\mathbf{A}$ размера $I \times J$ можно представить как набор векторов:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_J \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \dots \\ \mathbf{b}_I \end{bmatrix}$$

Здесь каждый вектор $a_j$ является $j$-ым столбцом, а вектор-строка $b_i$ является $i$-ой строкой матрицы $\mathbf{A}$:

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \dots \\ a_{Ij} \end{bmatrix}$$

$$\mathbf{b}_i = \begin{bmatrix} a_{i1} & a_{i2} & \dots & a_{iJ} \end{bmatrix}$$

## 2.10 Линейно зависимые векторы

Векторы одинаковой размерности ($N$) можно складывать и умножать на число, также как матрицы. В результате получится вектор той же размерности. Пусть имеется несколько векторов одной размерности: $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K$ и столько же чисел $\alpha$: $\alpha_1, \alpha_2, \dots, \alpha_K$. Вектор:

$$\mathbf{y} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_K \mathbf{x}_K$$

называется *линейной комбинацией* векторов $\mathbf{x}_k$.

Если существуют такие ненулевые числа $\alpha_k \neq 0, k = 1, ..., K$, что $\mathbf{y} = 0$, то такой набор векторов $\mathbf{x}_k$ называется *линейно зависимым*. В противном случае векторы называются *линейно независимыми*. Например, векторы $\mathbf{x}_1 = (2, 2)^{\mathrm{t}}$ и $\mathbf{x}_2 = (-1, -1)^{\mathrm{t}}$ линейно зависимы, т.к. $\mathbf{x}_1 + 2\mathbf{x}_2 = 0$.

## 2.11 Ранг матрицы

Рассмотрим набор из $K$ векторов $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_K$ размерности $N$. *Рангом* этой системы векторов называется максимальное число линейно-независимых векторов. Например в наборе

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \mathbf{x}_4 = \begin{bmatrix} 3 \\ 2 \\ 0 \end{bmatrix}$$

имеются только два линейно независимых вектора, например $\mathbf{x}_1$ и $\mathbf{x}_2$, поэтому ее ранг равен 2.

Очевидно, что если векторов в наборе больше, чем их размерность ($K > N$), то они обязательно линейно зависимы.

*Рангом матрицы* (обозначается $\mathrm{rank}(\mathbf{A})$) называется ранг системы векторов, из которых она состоит. Хотя любую матрицу можно представить двумя способами (векторы столбцы или строки), это не влияет на величину ранга, т.к.

$$\mathrm{rank}(\mathbf{A}) = \mathrm{rank}(\mathbf{A}^{\mathrm{t}}).$$

## 2.12 Обратная и псевдообратная матрицы

Квадратная матрица $\mathbf{A}$ называется *невырожденной*, если она имеет единственную *обратную* матрицу $\mathbf{A}^{-1}$, определяемую условиями

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

Обратная матрица существует не для всех матриц. Необходимым и достаточным условием невырожденности является $\det(\mathbf{A}) \neq 0$ или $\operatorname{rank}(\mathbf{A}) = N$.

Обращение матрицы — это сложная процедура, для выполнения которой существуют специальные программы. Например:

$$\begin{vmatrix} 1.2 & -5.3 & 0.25 & 2.2 & 13.8 \\ 10 & 1.5 & -7.5 & 1.5 & 11.2 \\ 2.3 & -1.2 & 5.6 & -2.5 & 4.8 \\ 4.5 & -0.8 & 9.5 & 4.2 & 6.5 \\ 2.5 & 8.2 & -3.5 & 5.9 & 1.1 \end{vmatrix}^{-1} = \begin{vmatrix} -0.099 & 0.104 & -0.065 & 0.0926 & -0.083 \\ -0.003 & -0.039 & 0.1782 & -0.082 & 0.1452 \\ -0.016 & -0.035 & 0.0633 & 0.0447 & 0.0097 \\ 0.0223 & -0.006 & -0.202 & 0.1126 & -0.003 \\ 0.0766 & -0.022 & 0.1052 & -0.058 & 0.0633 \end{vmatrix}$$

**Рис. 2.16.** Обращение матрицы

Приведем формулы для простейшего случая — матрицы $2 \times 2$.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} a_{11} & -a_{12} \\ -a_{21} & a_{22} \end{bmatrix}.$$

Если матрицы $\mathbf{A}$ и $\mathbf{B}$ невырождены, то

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$$

Если матрица $\mathbf{A}$ вырождена и обратная матрица не существует, то в некоторых случаях можно использовать *псевдообратную* матрицу, которая определяется как такая матрица $\mathbf{A}^{+}$, что

$$\mathbf{A}\mathbf{A}^{+}\mathbf{A} = \mathbf{A}$$

Псевдобратная матрица — не единственная и ее вид зависит от способа построения. Например для прямоугольной матрицы можно использовать метод Мура-Пенроуза.

Если число столбцов меньше числа строк, то

$$\mathbf{A}^{+} = (\mathbf{A}^{t}\mathbf{A})^{-1}\mathbf{A}^{t}$$

Например,

$$\mathbf{A} = \begin{vmatrix} 1.2 & -1.1 & 0.3 \\ 1.1 & 1.5 & 0.8 \\ 2.3 & -1.2 & -0.5 \\ -1.2 & -0.8 & 2.1 \end{vmatrix} \quad \mathbf{A}^{t} = \begin{vmatrix} 1.2 & 1.1 & 2.3 & -1.2 \\ -1.1 & 1.5 & -1.2 & -0.8 \\ 0.3 & 0.8 & -0.5 & 2.1 \end{vmatrix} \quad \mathbf{A}^{t}\mathbf{A} = \begin{vmatrix} 9.38 & -1.47 & -2.49 \\ -1.47 & 5.54 & -0.16 \\ -2.49 & -0.16 & 5.36 \end{vmatrix}$$

$$(\mathbf{A}^{t}\mathbf{A})^{-1} = \begin{vmatrix} 0.13 & 0.04 & 0.06 \\ 0.04 & 0.19 & 0.02 \\ 0.06 & 0.02 & 0.22 \end{vmatrix} \quad \mathbf{A}^{+} = (\mathbf{A}^{t}\mathbf{A})^{-1}\mathbf{A}^{t} = \begin{vmatrix} 0.13 & 0.24 & 0.22 & -0.06 \\ -0.16 & 0.34 & -0.16 & -0.15 \\ 0.10 & 0.27 & 0.01 & 0.36 \end{vmatrix} \quad \mathbf{A}\mathbf{A}^{+}\mathbf{A} = \begin{vmatrix} 1.2 & -1.1 & 0.25 \\ 1.1 & 1.5 & 0.8 \\ 2.3 & -1.2 & -0.5 \\ -1.2 & -0.8 & 2.1 \end{vmatrix}$$

**Рис. 2.17.** Псевдообращение матрицы

Если же число столбцов больше числа строк, то

$$\mathbf{A}^{+} = \mathbf{A}^{t}(\mathbf{A}\mathbf{A}^{t})^{-1}$$

## 2.13 Умножение вектора на матрицу

Вектор $\mathbf{x}$ можно умножать на матрицу $\mathbf{A}$ подходящей размерности. При этом вектор-столбец умножается справа $\mathbf{Ax}$, а вектор строка — слева $\mathbf{x}^{t}\mathbf{A}$. Если размерность вектора $J$, а размерность матрицы $I \times J$ то в результате получится вектор размерности $I$. Например,

$$\left|\begin{matrix} 1.2 & -3.5 & 1.8 & -2.2 & -5.3 \\ 10 & -2.8 & 3.2 & -2.8 & 1.5 \\ 2.3 & -6.5 & -9.2 & 6.5 & -1.2 \end{matrix}\right| * \left|\begin{matrix} 1.2 \\ 0.5 \\ 0.8 \\ 2.1 \\ -1.4 \end{matrix}\right| = \left|\begin{matrix} 3.93 \\ 5.42 \\ 7.48 \end{matrix}\right|$$

**Рис. 2.18.** Умножение вектора на матрицу

Если матрица $\mathbf{A}$ — квадратная $(I \times I)$, то вектор $\mathbf{y} = \mathbf{Ax}$ имеет ту же размерность, что и $\mathbf{x}$. Очевидно, что

$$\mathbf{A}(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1\mathbf{Ax}_1 + \alpha_2\mathbf{Ax}_2$$

Поэтому матрицы можно рассматривать как линейные преобразования векторов. В частности $\mathbf{Ix} = \mathbf{x}$, $\mathbf{Ox} = 0$.

# 3 Дополнительная информация

## 3.1 Системы линейных уравнений

Пусть $\mathbf{A}$ — матрица размером $I \times J$, а $\mathbf{b}$ — вектор размерности $J$. Рассмотрим уравнение

$$\mathbf{Ax} = \mathbf{b}$$

относительно вектора $\mathbf{x}$, размерности $I$. По сути — это система из $I$ линейных уравнений с $J$ неизвестными $\mathbf{x}_1, \dots, \mathbf{x}_J$. Решение существует в том, и только в том случае, когда

$$\operatorname{rank}(\mathbf{A}) = \operatorname{rank}(\mathbf{B}) = R,$$

где $\mathbf{B}$ — это расширенная матрица размерности $I \times (J+1)$, состоящая из матрицы $\mathbf{A}$, дополненной столбцом $\mathbf{b}$, $\mathbf{B} = (\mathbf{Ab})$. В противном случае уравнения несовместны.

Если $R = I = J$, то решение единственно

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

Если $R < I$, то существует множество различных решений, которые можно выразить через линейную комбинацию $J - R$ векторов. Система однородных уравнений $\mathbf{Ax} = \mathbf{0}$ с квадратной матрицей $\mathbf{A}$ $(N \times N)$ имеет нетривиальное решение $(\mathbf{x} \neq \mathbf{0})$ тогда и только тогда, когда $\det(\mathbf{A}) = 0$. Если $R = \operatorname{rank}(\mathbf{A}) < N$, то существуют $N - R$ линейно независимых решений.

## 3.2 Билинейные и квадратичные формы

Если $\mathbf{A}$ — это квадратная матрица , а $\mathbf{x}$ и $\mathbf{y}$ — вектора соответствующей размерности, то скалярное произведение вида $\mathbf{x}^{\mathrm{t}}\mathbf{Ay}$ называется *билинейной* формой, определяемой матрицей $\mathbf{A}$. При $\mathbf{x} = \mathbf{y}$ выражение $\mathbf{x}^{\mathrm{t}}\mathbf{Ax}$ называется *квадратичной* формой.

## 3.3 Положительно определенные матрицы

Квадратная матрица $\mathbf{A}$ называется *положительно определенной*, если для любого ненулевого вектора $\mathbf{x} \neq \mathbf{0}$,

$\mathbf{x}^t\mathbf{A}\mathbf{x} > 0$.

Аналогично определяются *отрицательно* ($\mathbf{x}^t\mathbf{A}\mathbf{x} < 0$), *неотрицательно* ($\mathbf{x}^t\mathbf{A}\mathbf{x} \geq 0$) и *неположительно* ($\mathbf{x}^t\mathbf{A}\mathbf{x} \leq 0$) определенные матрицы.

## 3.4 Разложение Холецкого

Если симметричная матрица $\mathbf{A}$ положительно определена, то существует единственная треугольная матрица $\mathbf{U}$ с положительными элементами, для которой

$$\mathbf{A} = \mathbf{U}^t\mathbf{U}$$

Например,

$$\begin{vmatrix} 4 & 2 & 6 \\ 2 & 10 & 9 \\ 6 & 9 & 14 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 0 \\ 1 & 3 & 0 \\ 3 & 2 & 1 \end{vmatrix} * \begin{vmatrix} 2 & 1 & 3 \\ 0 & 3 & 2 \\ 0 & 0 & 1 \end{vmatrix}$$

**Рис. 3.1.** Разложение Холецкого

## 3.5 Полярное разложение

Пусть $\mathbf{A}$ — это невырожденная квадратная матрица размерности $N \times N$. Тогда существует однозначное *полярное* представление

$$\mathbf{A} = \mathbf{S}\mathbf{R},$$

где $\mathbf{S}$ — это неотрицательная симметричная матрица, а $\mathbf{R}$ — это ортогональная матрица. Матрицы $\mathbf{S}$ и $\mathbf{R}$ могут быть определены явно:

$$\mathbf{S} = (\mathbf{A}\mathbf{A}^t)^{1/2}$$

$$\mathbf{R} = \mathbf{S}^{-1}\mathbf{A} = (\mathbf{A}\mathbf{A}^t)^{-1/2}\mathbf{A}.$$

Например,

$$\mathbf{A}=\begin{vmatrix} 4 & 2 & 6 \\ 2 & 10 & 9 \\ 3 & 8 & 5 \end{vmatrix} \quad \mathbf{AA}^t=\begin{vmatrix} 56 & 82 & 58 \\ 82 & 185 & 131 \\ 58 & 131 & 98 \end{vmatrix} \quad \mathbf{S}=\begin{vmatrix} -1.4 & 5.89 & 4.41 \\ 5.89 & 9.31 & 7.98 \\ 4.41 & 7.98 & 3.85 \end{vmatrix} \quad \mathbf{S}^2=\begin{vmatrix} 56 & 82 & 58 \\ 82 & 185 & 131 \\ 58 & 131 & 98 \end{vmatrix}$$

$$\mathbf{R}=\begin{vmatrix} -0.5 & 0.85 & -0.2 \\ 0.81 & 0.53 & 0.27 \\ -0.3 & 0.01 & 0.95 \end{vmatrix} \quad \mathbf{RR}^t=\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \quad \mathbf{SR}=\begin{vmatrix} 4 & 2 & 6 \\ 2 & 10 & 9 \\ 3 & 8 & 5 \end{vmatrix}$$

**Рис. 3.2.** Полярное разложение

Если матрица $\mathbf{A}$ вырождена, то разложение не единственно — а именно: $\mathbf{S}$ по-прежнему одна, а вот $\mathbf{R}$ может быть много. Полярное разложение представляет матрицу $\mathbf{A}$ как комбинацию сжатия/растяжения $\mathbf{S}$ и поворота $\mathbf{R}$.

## 3.6 Собственные векторы и значения

Пусть $\mathbf{A}$ — это квадратная матрица. Вектор $\mathbf{v}$ называется собственным вектором матрицы $\mathbf{A}$, если

$$\mathbf{Av} = \lambda\mathbf{v},$$

где число $\lambda$ называется *собственным значением* матрицы $\mathbf{A}$. Таким образом преобразование, которое выполняет матрица $\mathbf{A}$ над вектором $\mathbf{v}$, сводится к простому растяжению или сжатию с коэффициентом $\lambda$. Собственный вектор определяется с точностью до умножения на константу $\alpha \neq 0$, т.е. если $\mathbf{v}$ — собственный вектор, то и $\alpha\mathbf{v}$ — тоже собственный вектор.

## 3.7 Собственные значения

У матрицы $\mathbf{A}$ , размерностью $(N \times N)$ не может быть больше чем $N$ собственных значений. Они удовлетворяют *характеристическому уравнению*

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0,$$

являющемуся алгебраическим уравнением $N$-го порядка. В частности, для матрицы 2×2 характеристическое уравнение имеет вид

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \det\left(\begin{bmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{bmatrix}\right) = (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} = 0$$

Например,

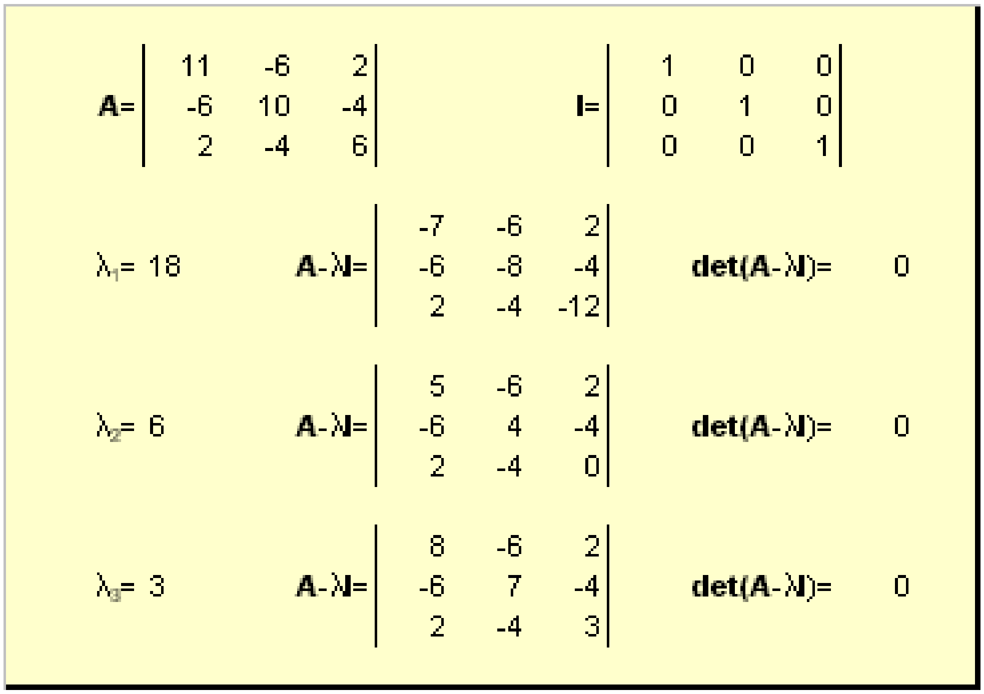


**Рис. 3.3.** Собственные значения

Набор собственных значений $\lambda_1, \dots, \lambda_N$ матрицы $\mathbf{A}$ называется *спектром* $\mathbf{A}$.

Спектр обладает разнообразными свойствами. В частности

$$\det(\mathbf{A}) = \lambda_1 \times \dots \times \lambda_N$$

$$\mathrm{Sp}(\mathbf{A}) = \lambda_1 + \dots + \lambda_N$$

Собственные значения произвольной матрицы могут быть комплексными числами, однако если матрица симметричная ($\mathbf{A}^{\mathrm{t}} = \mathbf{A}$), то ее собственные значения вещественны.

## 3.8 Собственные векторы

У матрицы $\mathbf{A}$, размерностью $(N \times N)$ не может быть больше чем $N$ собственных векторов, каждый из которых соответствует своему собственному значению. Для определения собственного вектора $\mathbf{v}_n$ нужно решить систему однородных уравнений

$$(\mathbf{A} - \lambda_n \mathbf{I})\mathbf{v}_n = \mathbf{0}$$

Она имеет нетривиальное решение, поскольку $\det(\mathbf{A} - \lambda_n \mathbf{I}) = 0$.

Например,

$$\mathbf{A}=\begin{vmatrix} 11 & -6 & 2 \\ -6 & 10 & -4 \\ 2 & -4 & 6 \end{vmatrix} \qquad \mathbf{I}=\begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix}$$

$$\lambda_1 = 18 \qquad \mathbf{v}_1=\begin{vmatrix} 2 \\ -2 \\ 1 \end{vmatrix} \qquad \begin{vmatrix} -7 & -6 & 2 \\ -6 & -8 & -4 \\ 2 & -4 & -12 \end{vmatrix} * \begin{vmatrix} 2 \\ -2 \\ 1 \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ 0 \end{vmatrix}$$

$$\lambda_2 = 6 \qquad \mathbf{v}_2=\begin{vmatrix} 2 \\ 1 \\ -2 \end{vmatrix} \qquad \begin{vmatrix} 5 & -6 & 2 \\ -6 & 4 & -4 \\ 2 & -4 & 0 \end{vmatrix} * \begin{vmatrix} 2 \\ 1 \\ -2 \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ 0 \end{vmatrix}$$

$$\lambda_3 = 3 \qquad \mathbf{v}_3=\begin{vmatrix} 1 \\ 2 \\ 2 \end{vmatrix} \qquad \begin{vmatrix} 8 & -6 & 2 \\ -6 & 7 & -4 \\ 2 & -4 & 3 \end{vmatrix} * \begin{vmatrix} 1 \\ 2 \\ 2 \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ 0 \end{vmatrix}$$

**Рис. 3.4.** Собственные векторы

Собственные вектора симметричной матрицы ортогональны.

## 3.9 Эквивалентные и подобные матрицы

Две прямоугольные матрицы **A** и **B** одной размерности $I \times J$ *эквивалентны*, если существуют такие квадратные матрицы **S**, размерности $I \times I$, и **T**, размерности $J \times J$, что:

$$\mathbf{B} = \mathbf{SAT}$$

Эквивалентные матрицы имею один и тот же ранг.

Две прямоугольные матрицы **A** и **B** одной размерности $N \times N$ *подобны*, если существует такая невырожденная матрица **T**, что:

$$\mathbf{B} = \mathbf{T}^{-1}\mathbf{AT}$$

Матрица **T** называется *преобразованием подобия.*

Подобные матрицы имеют один и тот же ранг, след, определитель и спектр.

## 3.10 Приведение матрицы к диагональному виду

Нормальную (в частности симметричную) матрицу $\mathbf{A}$ можно привести к диагональному виду преобразованием подобия:

$$\mathbf{A} = \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{-1}$$

Здесь $\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_N)$ — это диагональная матрица, элементами которой являются собственные значения матрицы $\mathbf{A}$, а $\mathbf{T}$ — это матрица, составленная из соответствующих собственных векторов матрицы $\mathbf{A}$, т.е. $\mathbf{T} = (\mathbf{v}_1, \dots, \mathbf{v}_N)$.

Например,

$$\mathbf{A} = \begin{vmatrix} 11 & -6 & 2 \\ -6 & 10 & -4 \\ 2 & -4 & 6 \end{vmatrix} \qquad \mathbf{T} = \begin{vmatrix} 2 & 2 & 1 \\ -2 & 1 & 2 \\ 1 & -2 & 2 \end{vmatrix} \qquad \mathbf{T}^{-1} = \begin{vmatrix} 0.22 & -0.2 & 0.11 \\ 0.22 & 0.11 & -0.2 \\ 0.11 & 0.22 & 0.22 \end{vmatrix}$$

$$\boldsymbol{\Lambda} = \begin{vmatrix} 18 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 3 \end{vmatrix} \qquad \mathbf{T}\boldsymbol{\Lambda}\mathbf{T}^{-1} = \begin{vmatrix} 11 & -6 & 2 \\ -6 & 10 & -4 \\ 2 & -4 & 6 \end{vmatrix}$$

**Рис. 3.5.** Приведение к диагональному виду

## 3.11 Разложение по сингулярным значениям

Пусть имеется прямоугольная матрица $\mathbf{A}$ размерностью $I \times J$ ранга $R$ ($I \le J \le R$). Ее можно разложить в произведение трех матриц $\mathbf{P}_R$ ($I \times R$), $\mathbf{D}_R$ ($R \times R$) и $\mathbf{Q}_R$ ($J \times R$) —

$$\mathbf{A} = \mathbf{P}_R \mathbf{D}_R \mathbf{Q}_R^{\mathrm{t}}$$

так, чтобы —

$$\mathbf{P}_R^{\mathrm{t}} \mathbf{P}_R = \mathbf{Q}_R^{\mathrm{t}} \mathbf{Q}_R = \mathbf{I}_R$$

Здесь $\mathbf{P}_R$ — матрица, образованная $R$ ортонормированными собственными векторами $\mathbf{p}_r$ матрицы $\mathbf{A}\mathbf{A}^{\mathrm{t}}$, соответствующим $R$ наибольшим собственным значениям $\lambda_r$:

$$\mathbf{A}\mathbf{A}^{\mathrm{t}}\mathbf{p}_r = \lambda_r \mathbf{p}_r$$

$\mathbf{Q}_R$ — матрица, образованная $R$ ортонормированными собственными векторами $\mathbf{q}_r$ матрицы $\mathbf{A}^{\mathrm{t}}\mathbf{A}$:

$$\mathbf{A}^{\mathrm{t}}\mathbf{A}\mathbf{q}_r = \lambda_r \mathbf{q}_r$$

$\mathbf{D}_R = \mathrm{diag}(\sigma_1, \ldots, \sigma_R)$ — положительно определенная диагональная матрица, элементами которой являются $\sigma_1 \geq \cdots \geq \sigma_R \geq 0$ — *сингулярные значения* матрицы $\mathbf{A}$, равные квадратным корням из собственных значений матрицы $\mathbf{A}^{\mathrm{t}}\mathbf{A}$:

$$\sigma_r = \sqrt{\lambda_r}$$

Пример,

$$
\mathbf{A}=\begin{vmatrix}
-0.588 & 0.8317 & 0.7892 & 0.6727\\
0.1017 & 0.0217 & 0.1092 & 0.0647\\
-0.118 & 0.2817 & 0.2192 & 0.2167\\
0.1817 & -0.138 & -0.041 & -0.067\\
-0.738 & -0.868 & -0.671 & -0.776\\
0.0617 & 0.0917 & -0.041 & 0.0357\\
-0.108 & 0.3417 & 0.2492 & 0.2597\\
0.1017 & 0.1217 & 0.1592 & 0.1347\\
0.4517 & -0.328 & -0.601 & -0.359\\
0.2717 & -0.158 & -0.071 & -0.08\\
0.2017 & -0.148 & -0.211 & -0.138\\
0.1817 & -0.048 & 0.1092 & 0.0377
\end{vmatrix}
\qquad
\mathbf{P}_R=\begin{vmatrix}
0.631 & -0.403 & -0.155\\
0.0464 & 0.0908 & 0.2616\\
0.1948 & -0.066 & -0.145\\
-0.074 & 0.1395 & 0.2841\\
-0.579 & -0.756 & 0.0623\\
0.021 & 0.0661 & -0.294\\
0.2297 & -0.049 & -0.205\\
0.1048 & 0.104 & 0.1551\\
-0.358 & 0.3512 & -0.59\\
-0.094 & 0.2152 & 0.2872\\
-0.14 & 0.1563 & -0.115\\
0.0168 & 0.1509 & 0.4532
\end{vmatrix}
$$

$$I=12,\quad J=4,\quad R=3$$

$$
\mathbf{Q}_R=\begin{vmatrix}
-0.092 & 0.9865 & 0.1065\\
0.6136 & 0.093 & -0.662\\
0.5788 & -0.055 & 0.7412\\
0.5291 & 0.1233 & -0.024
\end{vmatrix}
\qquad
\mathbf{D}_R=\begin{vmatrix}
2.1825 & 0 & 0\\
0 & 1.149 & 0\\
0 & 0 & 0.2898
\end{vmatrix}
$$

$$
\mathbf{P}_R\mathbf{D}_R\mathbf{Q}_R^{t}=\begin{vmatrix}
-0.588 & 0.8317 & 0.7892 & 0.6727\\
0.1017 & 0.0217 & 0.1092 & 0.0647\\
-0.118 & 0.2817 & 0.2192 & 0.2167\\
0.1817 & -0.138 & -0.041 & -0.067\\
-0.738 & -0.868 & -0.671 & -0.776\\
0.0617 & 0.0917 & -0.041 & 0.0357\\
-0.108 & 0.3417 & 0.2492 & 0.2597\\
0.1017 & 0.1217 & 0.1592 & 0.1347\\
0.4517 & -0.328 & -0.601 & -0.359\\
0.2717 & -0.158 & -0.071 & -0.08\\
0.2017 & -0.148 & -0.211 & -0.138\\
0.1817 & -0.048 & 0.1092 & 0.0377
\end{vmatrix}
$$

**Рис. 3.6.** SVD разложение

Дополняя матрицы $\mathbf{P}_R$ и $\mathbf{Q}_R$ ортонормированными столбцами, а матрицу $\mathbf{D}_R$ нулевыми значениями, можно сконструировать матрицы $\mathbf{P}$ $(I \times J)$, $\mathbf{D}$ $(J \times J)$ и $\mathbf{Q}$ $(J \times J)$ такие, что

$$\mathbf{A} = \mathbf{P}_R\mathbf{D}_R\mathbf{Q}_R^{\mathrm{t}} = \mathbf{PDQ}^{\mathrm{t}}$$

Об использовании SVD рассказано в других пособиях MatLab. Руководство для начинающих и Метод главных компонент (PCA)

## 3.12 Линейное пространство и базис

Рассмотрим все возможные векторы размерности $N$. Это множество называется *линейным пространством* размерности $N$ и обозначается $\mathbb{R}^N$ . Так как в $\mathbb{R}^N$ включены все возможные векторы, то любая линейная комбинация векторов из $\mathbb{R}^N$ будет также принадлежать этому пространству.

Любой набор из $N$ линейно независимых векторов называется *базисом* в пространстве $\mathbb{R}^N$. Простейший пример базиса — это набор векторов

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \dots \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \dots \\ 0 \end{bmatrix}, \dots, \mathbf{e}_N = \begin{bmatrix} 0 \\ 0 \\ \dots \\ 1 \end{bmatrix}$$

в каждом из которых только один элемент равен 1, а остальные равны нулю. Тогда любой вектор $\mathbf{x} = (x_1, x_2, \dots, x_N)^{\mathrm{t}}$ может быть представлен как линейная комбинация $\mathbf{x} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_N\mathbf{e}_N$ базисных векторов.

Базис, составленный из попарно ортогональных векторов, называется *ортогональным*, а если базисные вектора еще и нормированы, то этот базис называется *ортонормированным*.

## 3.13 Геометрическая интерпретация

Линейному пространству можно дать удобную геометрическую интерпретацию. Представим себе $N$-мерное пространство, в котором базисные вектора задают направления осей координат. Тогда произвольный вектор $\mathbf{x} = (x_1, x_2, \dots, x_N)^{\mathrm{t}}$ можно изобразить точкой в этом пространстве с координатами $(x_1, x_2, \dots, x_N)$.

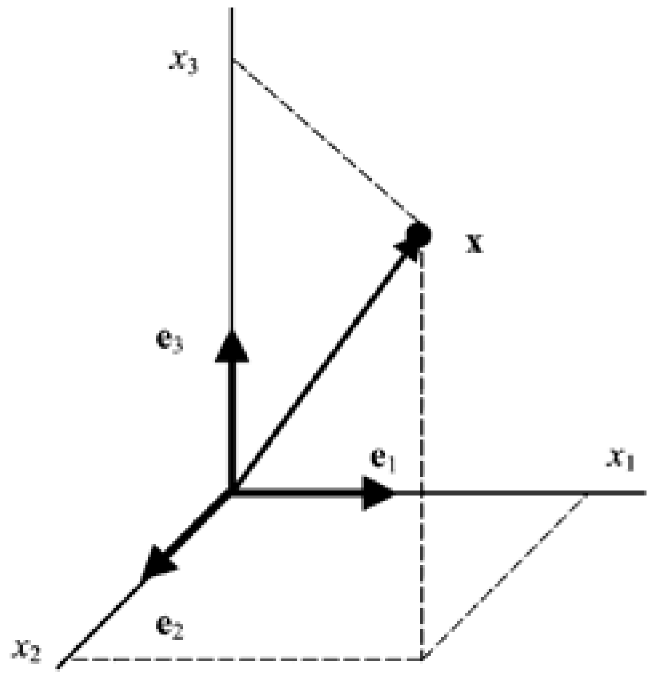


**Рис. 3.7.** Координатное пространство

## 3.14 Множественность базисов

В линейном пространстве могут быть неограниченное число базисов. Так, в пространстве $\mathbb{R}^3$ помимо обычного ортонормированного базиса

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \mathbf{e}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

можно установить и другой ортонормированный базис, например

$$\mathbf{b}_1 = \begin{bmatrix} \sqrt{0.5} \\ 0.5 \\ 0.5 \end{bmatrix}, \mathbf{b}_2 = \begin{bmatrix} -\sqrt{0.5} \\ 0.5 \\ 0.5 \end{bmatrix}, \mathbf{b}_3 = \begin{bmatrix} 0 \\ -\sqrt{0.5} \\ \sqrt{0.5} \end{bmatrix}$$

Каждый базис можно представить матрицей $\mathbf{B} = (\mathbf{b}_1, ..., \mathbf{b}_N)$, составленной из базисных векторов. Переход от одного базиса к другому осуществляется с помощью невырожденной квадратной матрицы $\mathbf{T}$, т.е. $\mathbf{B}_2 = \mathbf{TB}_1$.

## 3.15 Подпространство

Пусть имеется набор из $K$ линейно независимых векторов $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_K$ в пространстве $\mathbb{R}^N$. Рассмотрим все возможные линейные комбинации этих векторов

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \cdots + \alpha_K \mathbf{x}_K$$

О получившимся множестве $Q$ говорят, что оно является *линейной оболочкой* или что оно *натянуто* на векторы $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K$. По определению линейного пространства это множество $Q$ само является линейным пространством размерности $K$. При этом оно принадлежит пространству $\mathbb{R}^N$, поэтому $Q$ называется линейным подпространством $\mathbb{R}^K$ в пространстве $\mathbb{R}^N$.

## 3.16 Проекция на подпространство

Рассмотрим подпространство $\mathbb{R}^K$, натянутое на векторы $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_K)$ в пространстве $\mathbb{R}^N$. Матрица базиса $\mathbf{X}$ имеет размерность $(N \times K)$. Любой вектор $\mathbf{y}$ из $\mathbb{R}^N$ может быть спроецирован на подпространство $\mathbb{R}^K$, т.е. представлен в виде

$$\mathbf{y} = \mathbf{y}^{\parallel} + \mathbf{y}^{\perp}$$

где вектор $\mathbf{y}^{\parallel}$ принадлежит $\mathbb{R}^K$, а вектор $\mathbf{y}^{\perp}$ ортогонален $\mathbf{y}^{\parallel}$.

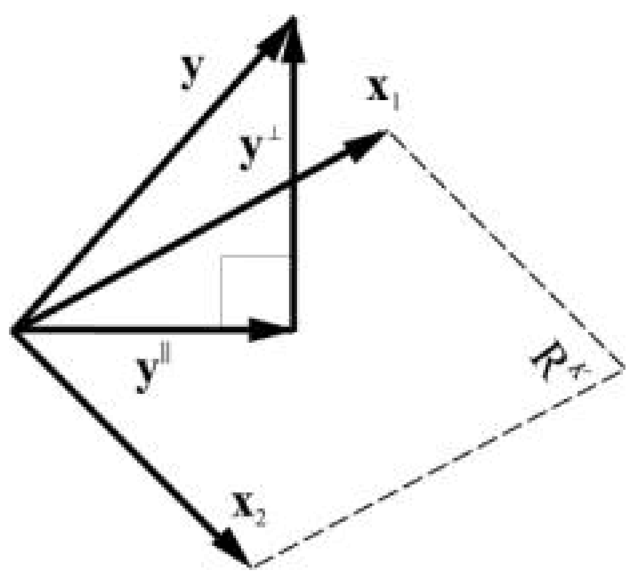


**Рис. 3.8.** Проекция на подпространство

Проекцию $\mathbf{y}^{\parallel}$ можно представить как результат действия проекционной матрицы $\mathbf{P}$

$$\mathbf{y}^{\parallel} = \mathbf{P}\mathbf{y}$$

Проекционная матрица определяется как

$$\mathbf{P} = \mathbf{X}(\mathbf{X}^{t}\mathbf{X})^{-1}\mathbf{X}^{t}$$

Пример:

$$\mathbf{y}=\begin{vmatrix}2\\0\\1\end{vmatrix}\quad \mathbf{x}_1=\begin{vmatrix}1\\1\\0\end{vmatrix}\quad \mathbf{x}_2=\begin{vmatrix}2\\2\\1\end{vmatrix}\quad \mathbf{X}=\begin{vmatrix}1&2\\1&2\\0&1\end{vmatrix}$$

$$\mathbf{X}^{t}\mathbf{X}=\begin{vmatrix}2&4\\4&9\end{vmatrix}\qquad (\mathbf{X}^{t}\mathbf{X})^{-1}=\begin{vmatrix}4.5&-2\\-2&1\end{vmatrix}$$

$$\mathbf{P}=\begin{vmatrix}0.5&0.5&0\\0.5&0.5&0\\0&0&1\end{vmatrix}\quad \mathbf{y}^{\parallel}=\mathbf{P}\mathbf{y}=\begin{vmatrix}1\\1\\1\end{vmatrix}\quad \mathbf{y}^{\perp}=(\mathbf{I}-\mathbf{P})\mathbf{y}=\begin{vmatrix}1\\-1\\0\end{vmatrix}$$

$$(\mathbf{y}^{\parallel},\mathbf{y}^{\perp})=0\qquad \mathbf{y}^{\parallel}=\mathbf{x}_2-\mathbf{x}_1=\begin{vmatrix}1\\1\\1\end{vmatrix}\qquad \mathbf{y}^{\parallel}+\mathbf{y}^{\perp}=\begin{vmatrix}2\\0\\1\end{vmatrix}$$

**Рис. 3.9.** Проекционное разложение

# 4 Заключение

Матричные методы активно используются при анализе данных, в том числе и хемометрическими методами.

Примеры приведены в пособиях

- Матричные операции в Excel
- Метод главных компонент (PCA)
- Калибровка
- Классификация
- Разрешение многомерных кривых

И в сопровождающих пособия книгах Excel.