Математика для Data Science. Линейная алгебра. Шпаргалка

Содержание

Третья неделя. Определитель и скалярное произведение

Определитель

Матрица 2×2

Поставим в соответствие произвольной матрице из действительных чисел $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ параллелограмм натянутый на векторы столбцы этой матрицы. Под словом "натянутый" подразумевается, что две стороны параллелограмма образуются векорами $\begin{pmatrix} a \\ b \end{pmatrix}$ и $\begin{pmatrix} c \\ d \end{pmatrix}$, а остальные получаются из них параллельным переносом.

Площадь этого параллелограмма в общем случае равна $|ad - bc|$.

Базис на плоскости $\vec{e}_1, \vec{e}_2$ называется *правым*, если кратчайший поворот от вектора $\vec{e}_1$ к вектору $\vec{e}_2$ происходит против часовой стрелки. В противном случае базис называется *левым*.

В случае, если столбцы матрицы $\begin{pmatrix} a \\ b \end{pmatrix}$ и $\begin{pmatrix} c \\ d \end{pmatrix}$ задают левую пару векторов, то *ориентированная площадь* параллелограмма будет отрицательной. Если правую пару векторов — положительной.

Ориентированная площадь паралллограмма натянутого на столбцы матрицы $\begin{pmatrix} a & c \\ b & d \end{pmatrix}$ равняется $ad - bc$.

Матрица 3×3

Рассмотрим параллелепипед, натянутый на столбцы матрицы $\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}$. Его объём будет равен $|x_{11}x_{22}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31} - x_{23}x_{32}x_{11} - x_{33}x_{12}x_{21}|$.

Пусть $\vec{x}, \vec{y}$ и $\vec{z}$ — базис в трехмерном пространстве. Представим, что мы смотрим на плоскость, в которой лежат векторы $\vec{x}$ и $\vec{y}$, из конца вектора $\vec{z}$. Тогда ориентация трехмерного базиса совпадает с двумерной ориентацией пары векторов $\vec{x}$ и $\vec{y}$.

Ориентированный объем параллелепипеда натянутого на столбцы матрицы $\begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix}$ равняется $x_{11}x_{22}x_{33} + x_{12}x_{23}x_{31} + x_{13}x_{21}x_{32} - x_{13}x_{22}x_{31} - x_{23}x_{32}x_{11} - x_{33}x_{12}x_{21}$.

Матрица $\mathbf{n} \times \mathbf{n}$

Пусть M — квадратная матрица действительных чисел размера $n \times n$. Ее *определитель*, то есть n-мерный ориентированный объем параллелепипеда натянутого на столбцы этой матрицы имеет следующие свойства:

1. Ориентированный объем единичного куба равен 1, то есть определитель единичной матрицы равен 1. Или коротко $\det E = 1$.

   Единичным n-мерным кубом называют параллелепипед, натянутый на столбцы единичной матрицы $n \times n$.

2. Если один из столбцов матрицы домножить на $\lambda \in \mathbb{R}$, то n-мерный ориентированный объем (определитель) тоже домножится на λ.

3. Если поменять два столбца матрицы местами, то определитель поменяет знак. Это свойство называют *кососимметричностью* по столбцам.

4. Объем параллелепипеда натянутого на векторы $\vec{a}_1, (\vec{b} + \vec{c}), \vec{a}_3, \ldots, \vec{a}_n$ равен сумме объемов параллелограммов натянутых на векторы $\vec{a}_1, \vec{b}, \vec{a}_3, \ldots, \vec{a}_n$ и векторы $\vec{a}_1, \vec{c}, \vec{a}_3, \ldots, \vec{a}_n$.

Утверждения

1. В результате линейного преобразования A объем любой фигуры меняется в $\det A$ раз.

2. Для квадратных матриц одного размера выполнено $\det(BA) = \det B \cdot \det A$.

Матрица, у которой все элементы ниже главной диагонали равны нулю, называется *верхнетреугольной*:
$$\begin{pmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{pmatrix}$$

Как считать определитель

1. Применяем к матрице метод Гаусса и получаем верхнетреугольную матрицу. Определитель этой верхнетреугольной матрицы равен или противоположен определителю изначальной матрицы (в зависимости от того, сколько раз мы использовали элементарные преобразования первого типа).

2. Если на диагонали получается 0, то определитель равен нулю. Иначе применяем преобразования второго типа, чтобы превратить верхнетреугольную матрицу в диагональную:

   Вычтем из столбцов со 2-ого по n-ый первый столбец домноженный на $\frac{1}{a_{11}}$, потом вычтем из столбцов с 3-его по n-ый второй столбец домноженный на $\frac{1}{a_{22}}$ и так далее.

   Определитель при этом не меняется.

3. Считаем определитель диагональной матрицы.

Переход в другой базис и обратная матрица

Квадратная матрица A называется *матрицей перехода* от базиса $\boldsymbol{h}$ к базису $\boldsymbol{g}$ в $\mathbb{R}^n$, если для всех векторов $\vec{v} \in \mathbb{R}^n$ выполнено следующее. Если $(v_1, v_2, \dots v_n)$ — координаты вектора в базисе h, то координаты $\vec{v}$ в базисе $\boldsymbol{g}$ — это $A \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$.

В частности, матрицей перехода от базиса $\boldsymbol{f}$ к стандартному является матрица, в которой по столбцам записаны координаты базиса $\boldsymbol{f}$.

Для квадратной матрицы A матрица B называется *обратной матрицей*, если $BA = E$. Обозначение A^{-1}.

Свойства обратной матрицы

1. Для любой матрицы полного ранга существует обратная матрица.

2. При домножении матрицы B на A^{-1} слева или справа и результат не обязательно будет одинаковый: $A^{-1}B$ не обязательно равно BA^{-1}.

3. Если $BA = E$ и $AC = E$ то $B = C$. То есть любая матрица, которая является обратной слева, будет и обратной справа. И наоборот.

4. Обратная матрица единственна. То есть, если $B_1A = E$ и $B_2A = E$, то $B_1 = B_2$.

5. Для каждой матрицы полного ранга единственным образом определена обратная матрица.

Пусть f — произвольный базис в $\mathbb{R}^n$, F — матрица перехода от f к стандартному базису.

Если $(v_1^f, v_2^f, \dots, v_n^f)$ — координаты вектора $\vec{v}$ в базисе f, то $\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = F \begin{pmatrix} v_1^f \\ v_2^f \\ \vdots \\ v_n^f \end{pmatrix}$ — координаты в стандартном базисе.

С другой стороны, F^{-1} — матрица перехода от стандартного базиса к f и при этом $\begin{pmatrix} v_1^f \\ v_2^f \\ \vdots \\ v_n^f \end{pmatrix} = F^{-1} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}$

Приведение матрицы к единичной

1. Пусть дана квадратная матрица A полного ранга. С помощью первых двух элементарных преобразований приведём её к диагональному виду.

2. Применим *третье элементарное преобразование*: умножение столбца (или строки) на ненулевое число. С его помощью мы каждый элемент на диагонали матрицы сможем привести к единице, домножив i-ый столбец на $\frac{1}{a_{ii}}$.

Метод Жордана-Гаусса поиска обратной матрицы

Для всей цепочки элементарных преобразований в описанном выше алгоритме существует последовательность матриц $M_1, M_2, \dots, M_k$ такая, что M_1 соответствует первому примененному элементарному преобразованию, M_2 — второму и т.д. Так как $M_k M_{k-1} \dots M_2 M_1 A = E$. Значит, $M_k M_{k-1} \dots M_2 M_1$ — как раз и есть A^{-1}.

Матрица перехода в общем случае

Если

- T_f — матрица перехода от $\boldsymbol{f}$ к стандартному базису, то есть записанные по столбцам векторы базиса $\boldsymbol{f}$
- T_g — матрица перехода от $\boldsymbol{g}$ к стандартному базису, то есть записанные по столбцам векторы базиса $\boldsymbol{g}$

то

- $T_g^{-1} T_f$ — матрица перехода от f к g
- $T_f^{-1} T_g$ — матрица перехода от $\boldsymbol{g}$ к $\boldsymbol{f}$

Линейное преобразование записанное в другом базисе Пусть в базисе $\boldsymbol{f}$ некоторое линейное преобразование t задано при помощи матрицы A и матрица перехода от $\boldsymbol{g}$ к $\boldsymbol{f}$ это C.. Тогда матрица преобразования t в базисе $\boldsymbol{g}$ записывается как $C^{-1}AC$.

Длина, углы и скалярное произведение

Длина вектора $\vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ определяется так: $||\vec{x}|| = \sqrt{x_1^2 + \dots + x_n^2}$.

Свойства длины

1. Длина любого вектора неотрицательна.

2. Длина вектора равна нулю тогда и только тогда, когда вектор равен $\vec{0}$.

3. Длина вектора $\alpha\vec{x}$ равна $|\alpha| \cdot ||\vec{x}||$.

4. Неравенство треугольника: $||\vec{x}|| + ||\vec{y}|| \geq ||\vec{x} + \vec{y}||$ для любых $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Теорема Пифагора. Дан треугольник со сторонами a, b, c. Угол между сторонами с длинами a и b прямой, если и только если $a^2 + b^2 = c^2$.

Векторы, угол между которыми прямой, называются *ортогональными* или *перпендикулярными.*
Нулевой вектор считается одновременно ортогональным и параллельным любому вектору.

Теорема косинусов. Дан треугольник со сторонами a, b, c. Угол между сторонами с длинами a и b равен α. Тогда выполнено соотношение $a^2 + b^2 - 2ab\cos\alpha = c^2$.
В частности, $\cos\alpha = \frac{||\vec{x}||^2 + ||\vec{y}||^2 - ||\vec{x} - \vec{y}||^2}{2||\vec{x}|| \cdot ||\vec{y}||}$ при $||\vec{x}|| > 0$ и $||\vec{y}|| > 0$.

Скалярным произведением векторов $\vec{x} \in \mathbb{R}^n$ и $\vec{y} \in \mathbb{R}^n$ называется число
$\langle \vec{x}, \vec{y} \rangle := x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum\limits_{i=1}^{n} x_i y_i$.

Свойства скалярного произведения

1. $\langle\vec{x},\vec{x}\rangle = ||\vec{x}||^2$
2. $\langle\vec{x},\vec{y}\rangle = \langle\vec{y},\vec{x}\rangle$, то есть скалярное произведение коммутативно.
3. $\langle c\vec{x},\vec{y}\rangle = c\langle\vec{x},\vec{y}\rangle$ для любого $c\in\mathbb{R}$
4. $\langle\vec{x}+\vec{y},\vec{z}\rangle = \langle\vec{x},\vec{z}\rangle + \langle\vec{y},\vec{z}\rangle$
5. $||\vec{x}+\vec{y}||^2 = \langle\vec{x},\vec{x}\rangle + 2\langle\vec{x},\vec{y}\rangle + \langle\vec{y},\vec{y}\rangle = ||x||^2 + 2\langle\vec{x},\vec{y}\rangle + ||y||^2$.
6. $\langle\vec{x},\vec{y}\rangle = \vec{x}^T\vec{y}$.
7. Для любого линейного отображения $f:\mathbb{R}^n\to\mathbb{R}$ можно найти вектор $\vec{a}$, такой что $f(\vec{x}) = \langle\vec{a},\vec{x}\rangle$
8. $\langle\vec{x},\vec{y}\rangle = 0$ если и только если вектора $\vec{x}$ и $\vec{y}$ перпендикулярны.
9. Все $\vec{e}_i$ ортогональны друг другу.
10. $\langle\vec{x},\vec{y}\rangle = \cos\alpha\cdot||\vec{x}||\cdot||\vec{y}||$, где α это угол между векторами $\vec{x}$ и $\vec{y}$

Следующие определения скалярного произведения эквивалентны:

1. $\langle\vec{x},\vec{y}\rangle = \sum\limits_{i=1}^{n} x_iy_i$ — алгебраическое определение
2. $\langle\vec{x},\vec{y}\rangle = \cos\alpha\cdot||\vec{x}||\cdot||\vec{y}||$ — геометрическое определение
3. $\langle\vec{x},\vec{y}\rangle = \vec{x}^T\vec{y}$ — определение через "строка на столбец"
4. $\langle\vec{x},\vec{y}\rangle = \frac{1}{2}(||\vec{x}+\vec{y}||^2 - ||\vec{x}||^2 - ||\vec{y}||^2)$ — определение через длины

Ортогональные матрицы

Линейное преобразование $Q:\mathbb{R}^n\to\mathbb{R}^n$, сохраняющее длины всех векторов, называется *ортогональным*. Другими словами, Q ортогонально, если $||Q\vec{x}|| = ||\vec{x}||$ для всех $\vec{x}\in\mathbb{R}^n$.

Теорема. Линейное преобразование Q сохраняет длины векторов тогда и только тогда, когда оно сохраняет скалярные произведения. То есть когда $\langle Q\vec{x},Q\vec{y}\rangle = \langle\vec{x},\vec{y}\rangle$ для любых $\vec{x},\vec{y}$.

Любой набор векторов в $\mathbb{R}^n$, удовлетворяющий свойствам

1. в наборе n элементов
2. длина каждого вектора набора равна 1
3. эти векторы попарно ортогональны (то есть любые два различных вектора из набора ортогональны друг другу)

называется *ортонормированным базисом*.

Следующие утверждения равносильны:

1. Q сохраняет длины векторов, то есть $||Q\vec{x}|| = ||\vec{x}||$
2. Q сохраняет скалярные произведения, то есть $\langle Q\vec{x},Q\vec{y}\rangle = \langle\vec{x},\vec{y}\rangle$
3. $Q^TQ = E$
4. $QQ^T = E$
5. $Q^T = Q^{-1}$
6. Q переводит базис $\vec{e}_1,\ldots,\vec{e}_n$ в ортонормированный базис

Теорема. Для любого ортогонального преобразования существует ортонормированный базис, в котором матрица преобразования выглядит так:

$$\begin{pmatrix}
1 & & & & & & & & & & & \\
& \ddots & & & & & & & & & & \\
& & 1 & & & & & & & & & \\
& & & -1 & & & & & & & & \\
& & & & \ddots & & & & & & & \\
& & & & & -1 & & & & & & \\
& & & & & & \cos\alpha_1 & -\sin\alpha_1 & & & & \\
& & & & & & \sin\alpha_1 & -\cos\alpha_1 & & & & \\
& & & & & & & & \ddots & & & \\
& & & & & & & & & & \cos\alpha_k & -\sin\alpha_k \\
& & & & & & & & & & \sin\alpha_k & -\cos\alpha_k
\end{pmatrix}$$

где все не написанные коэффициенты матрицы равны 0.

То есть есть в этом базисе ортогональное преобразование

- оставляет неподвижными первые несколько элементов базиса
- следующие несколько элементов базиса заменяет на противоположные
- оставшиеся элементы базиса разбивает на пары и делает поворот в каждой из плоскостей, заданной парой элементов базиса

То есть любое ортогональное преобразование состоит из простых отображений: отражений и поворотов.