Линейная алгебра. Матрицы

1. Матрицы

Пусть имеется некоторое количество бутылок вина, для каждой из которых необходимо определить, из какого сорта винограда сделано вино. Числовые данные про эти бутылки имеют двумерную структуру, причем по строкам расположены *объекты*, а по столбцам — *признаки*.

Такого рода структуры встречаются очень часто и называются **матрицами**:

$$A = \begin{pmatrix} 12 & 7 & 21 & 31 & 11 \\ 45 & -2 & 14 & 27 & 19 \\ -3 & 15 & 36 & 71 & 26 \\ 4 & -13 & 55 & 34 & 15 \end{pmatrix}.$$

Элементы матрицы обозначаются следующим образом (первый индекс — номер строки, второй — номер столбца):

$$a_{12} = 7, \quad a_{31} = -3.$$

В данном примере матрица A имеет размер 4×5, то есть рассматриваются четыре бутылки вина, каждая из которых характеризуется пятью признаками. Пространство всех таких матриц обозначается $\mathbb{R}^{4\times 5}$.

Необходимо отметить, что вектор размера n тоже является матрицей (*вектор-столбец* или *вектор-строка*):

$$\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} w_1 & w_2 & w_3 & w_4 & w_5 \end{pmatrix}.$$

Операция *умножения* матрицы размера $m \times n$ на вектор-столбец размера $n \times 1$ задается следующим образом:

$$A\vec{w} = \begin{pmatrix} \sum_{i=1}^n a_{1i}w_i \\ \sum_{i=1}^n a_{2i}w_i \\ \dots \\ \sum_{i=1}^n a_{mi}w_i \end{pmatrix}.$$

Матрицы используются для работы с системами линейных уравнений. Пусть про каждую из четырех бутылок вина известно, подлинная она или нет (сделана ли она из правильного сорта винограда). Эту информацию можно «закодировать» с помощью векторов. Если первая и четвертая бутылки вина подлинные, а две другие — поддельные, то это можно записать в виде вектора:

$$\vec{y} = (1,\ 0,\ 0,\ 1).$$

Тогда можно произвести линейную классификацию и потребовать, чтобы сумма значений всех признаков с некоторыми весами $(w_1, \ldots, w_5)$ равнялась номеру класса. Получится *система уравнений*:

$$\begin{cases} 12w_1 + 7w_2 + 21w_3 + 31w_4 + 11w_5 = 1, \\ 45w_1 - 2w_2 + 14w_3 + 27w_4 + 19w_5 = 0, \\ -3w_1 + 15w_2 + 36w_3 + 71w_4 + 26w_5 = 0, \\ 4w_1 - 13w_2 + 55w_3 + 34w_4 + 15w_5 = 1. \end{cases}$$

Такую систему можно записать в матричном виде:

$$A\vec{w} = \vec{y}.$$

Пример умножения матрицы на вектор:

$$A = \begin{pmatrix} 12 & 7 & 21 & 31 & 11 \\ 45 & -2 & 14 & 27 & 19 \\ -3 & 15 & 36 & 71 & 26 \\ 4 & -13 & 55 & 34 & 15 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 46 \\ 55 \\ 63 \\ 33 \end{pmatrix}.$$

Нужно заметить, что при умножении матрицы $m \times n$ на вектор длины n получается вектор длины m. Таким образом, матрица задает некоторое *линейное преобразование* — линейную функцию из одного векторного пространства в другое.

2. Матричные операции

Помимо умножения матрицы на вектор определенного размера существует операция *умножения матрицы на матрицу*. Данная операция определена для матриц, одна из которых размера $m \times n$, а вторая — $n \times k$. Результатом такой операции будет матрица C размера $m \times k$:

$$C = AB, \quad c_{ij} = \sum_{p=1}^{n} a_{ip} b_{pj}.$$

Пример умножения матрицы на матрицу:

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 10 & 0 \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 4 \\ 0 & 0 & 4 \\ 10 & 0 & 0 \end{pmatrix}.$$

Чтобы понять, почему операция умножения матриц определяется именно так, нужно заметить, что матрица, по сути, задает линейное преобразование. Если, например, матрица A задает преобразование из векторного пространства V в пространство W, а матрица B — из W в T, то их композиция BA задает преобразование из V в T. Таким образом, BA задает матрицу преобразования-композиции A и B.

Рассматривается некоторая система линейных уравнений $A\vec{x} = \vec{b}$. Вектор $\vec{x}$ может выражается через другой вектор, например $\vec{x} = B\vec{z}$. Тогда:

$$A(B\vec{z}) = \vec{b} \quad \Rightarrow \quad AB\vec{z} = \vec{b},$$

следовательно, можно решать систему уравнений с матрицей AB.

Для матриц одинакового размера определена также операция *сложения*:

$$C = A + B, \quad c_{ij} = a_{ij} + b_{ij}.$$

Похожим образом вводится операция *умножения на число*:

$$C = \alpha B, \quad c_{ij} = \alpha b_{ij}.$$

Пусть матрица A имеет вид:

$$A = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 3 \\ 2 & 1 & 0 \\ 5 & 1 & 1 \end{pmatrix}.$$

Главной диагональю матрицы называются все ее элементы, у которых первый индекс равен второму (a_{ii}). У матрицы A такими элементами являются 1, 3, 0.

Операция *транспонирования матрицы* — это поворот относительно главной диагонали. Матрица A размера $m \times n$ становится в результате матрицей A^T размера $n \times m$:

$$a_{ij}^T = a_{ji}.$$

Строки исходной матрицы становятся столбцами транспонированной матрицы. В примере с матрицей A:

$$A = \begin{pmatrix} 1 & 4 & 1 \\ 2 & 3 & 3 \\ 2 & 1 & 0 \\ 5 & 1 & 1 \end{pmatrix} \quad \Rightarrow \quad A^T = \begin{pmatrix} 1 & 2 & 2 & 5 \\ 4 & 3 & 1 & 1 \\ 1 & 3 & 0 & 1 \end{pmatrix}.$$

3. Ранг и определитель

Пусть имеются два вектора с координатами $(1, 2)$ и $(3, 1)$. Их можно дополнить до параллелепипеда (см. рис. 1).

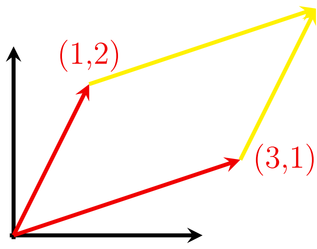

Рис. 1.

Для вычисления площади получившейся фигуры нужно составить матрицу из данных векторов:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}.$$

Площадь параллелепипеда будет равна **определителю** матрицы:

$$\det A = |A| = a_{11}a_{22} - a_{12}a_{21}.$$

В данном примере определитель равен -5. Знак здесь зависит от ориентации векторов относительно друг друга.

Определитель определен только для квадратной матрицы. Существует много способов задать определитель, например, это объем параллелограмма из n векторов. Можно также вводить данное понятие рекурсивно (*теорема Гаусса*).

В случае двух линейно зависимых векторов площадь фигуры, образованной векторами, равна нулю (см. рис. 2).

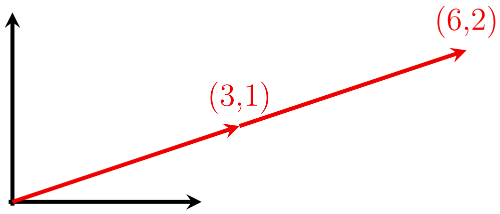

Рис. 2.

Свойства определителя:

1) Определитель равен нулю тогда и только тогда, когда строки или столбцы матрицы линейно зависимы.
2) $|A| = \left|A^T\right|$.
3) $|AB| = |A| \times |B|$.

Рангом системы строк матрицы A называется максимальное число линейно независимых строк среди них. **Рангом системы столбцов** матрицы A называется максимальное число линейно независимых столбцов среди них. Ранги матрицы по столбцам и по строкам всегда совпадают и обозначаются $\operatorname{rg} A$. Ранг характеризует количество информации в матрице.

Пусть матрица A имеет вид:

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 0{,}5 & 1 & 0{,}5 \end{pmatrix}.$$

В такой матрице все строки являются линейно зависимыми, при этом столбцы также линейно зависимы. Ранг матрицы A равен 1, следовательно, всю матрицу можно заменить на одну строку:

$$A = \begin{pmatrix} 1 & 2 & 1 \end{pmatrix}.$$

4. Системы линейных уравнений

Системой линейных уравнений называется набор из n уравнений, каждое из которых — линейное:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2, \\ \dots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_n. \end{cases}$$

В матричном виде система записывается следующим образом:

$$A\vec{x} = \vec{b}.$$

Задача заключается в нахождении вектора $\vec{x}$, удовлетворяющего данному уравнению. Возможны три случая: система имеет бесконечно много, одно или ни одного решения.

Пусть имеются два неизвестных — x и y. Линейная комбинация этих неизвестных (уравнение вида $a_1x + a_2y = b$) задает прямую на плоскости. Если уравнение одно, то любая точка, лежащая на этой прямой, будет решением, значит, решений бесконечно много (см. рис. 3).

Если уравнений два, то необходимо найти пересечение двух прямых, задаваемых этими уравнениями. Эта точка будет единственным решением (см. рис. 4).

Если уравнений три, то прямых тоже три, причем любые две из них пересекаются, но все три не сходятся в одной точке. Решений не существует (см. рис. 5).

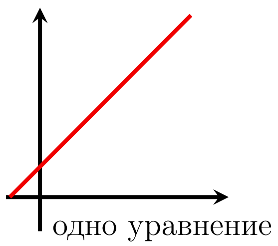

Рис. 3.

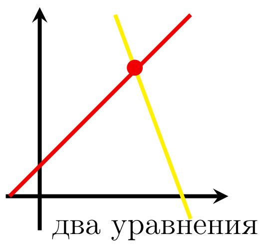

Рис. 4.

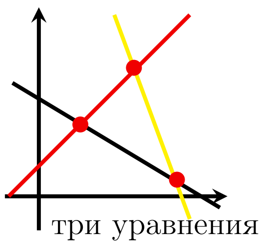

Рис. 5.

Для определения, какой из описанных случаев имеет место, понадобится ранг матрицы. Необходимо сравнить ранг матрицы A и ранг расширенной матрицы $\left(A|\vec{b}\right)$. Если ранги равны, то существует хотя бы одно решение. Если же второй ранг больше, то решений нет.

В случае равенства рангов если ранг A равен числу неизвестных, то решение одно. Если этот ранг меньше числа неизвестных, то решений бесконечно много.

Существует множество численных методов решения систем линейных уравнений. Однако, если решений бесконечно много, то необходимо ввести некоторые дополнительные требования (например, найти решение с наименьшей нормой).

Матрица A^{-1} называется **обратной** к матрице A, если:

$$AA^{-1} = A^{-1}A = I,$$

где I — единичная матрица.

Операция обращения определена только для квадратных матриц.

Теорема 1. *Обратная матрица к матрице A существует тогда и только тогда, когда:*

$$\det A \neq 0.$$

Для системы линейных уравнений с квадратной матрицей A можно записать:

$$A\vec{x} = \vec{b} \quad \Rightarrow \quad \vec{x} = A^{-1}\vec{b}.$$

5. Типы матриц

Матрица задает линейное преобразование одного векторного пространства в другое. Если матрица квадратная, то она отображает векторное пространство в это же пространство. Например, следующее преобразование поворачивает, отражает и растягивает единичный квадрат (см. рис. 6):

$$T(x, y) = \begin{bmatrix} -1 & -1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

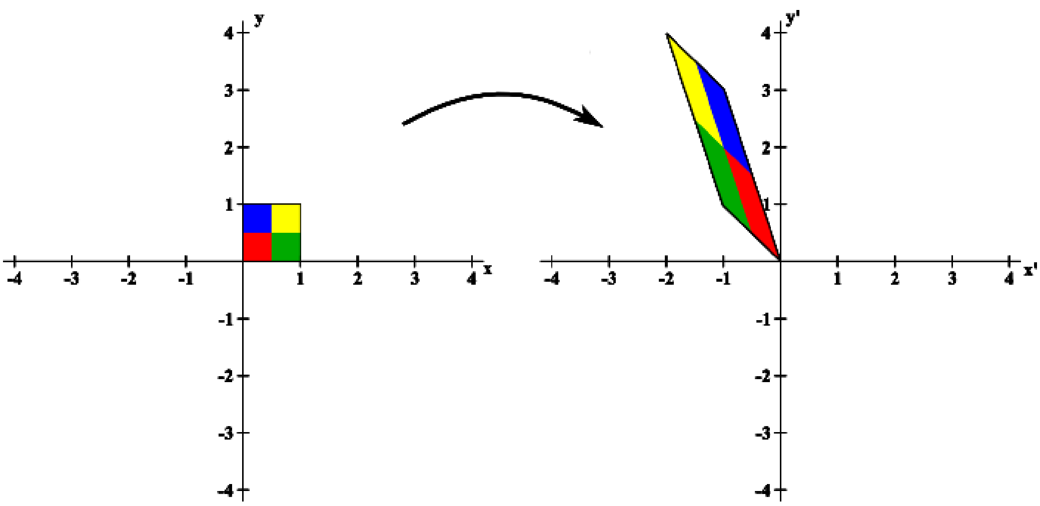

Рис. 6.

Важным типом матриц являются **диагональные матрицы**. У таких матриц вне главной диагонали находятся только нули. Частным случаем диагональных матриц является единичная матрица I.

Диагональные матрицы растягивают i-ю координату в a_{ii} раз. Например, следующая матрица растягивает первую координату в два раза (см. рис. 7):

$$D = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}.$$

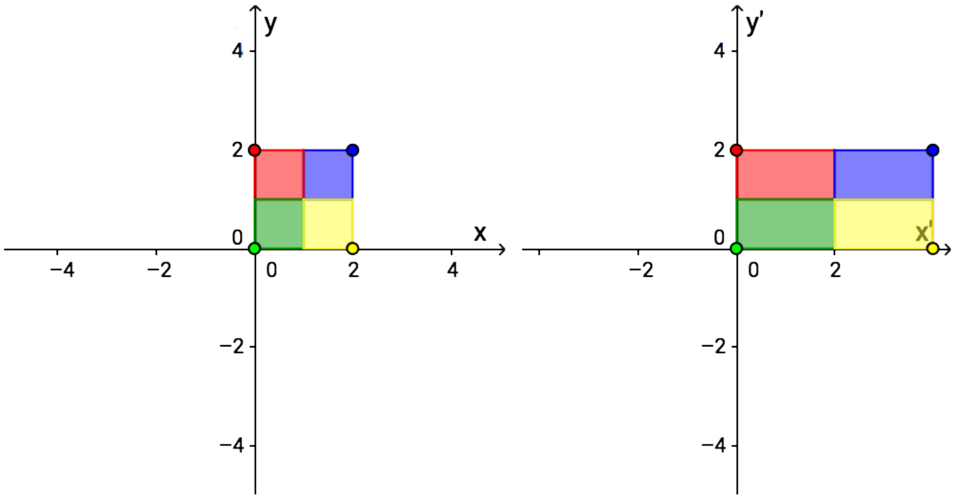

Рис. 7.

Еще одним типом матриц являются **ортогональные матрицы**. Для таких матриц транспонированная версия является обратной:

$$A^T A = AA^T = I.$$

Такие матрицы сохраняют длины векторов и скалярные произведения:

$$||A\vec{x}|| = ||\vec{x}||, \quad \langle A\vec{x}, A\vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle.$$

Отсюда следует, что ортогональные матрицы сохраняют углы. Значит, они задают преобразования типа поворотов и вращений, например, следующая матрица поворачивает единичный квадрат примерно на 20° (см. рис. 8):

$$A = \begin{pmatrix} 0{,}96 & 0{,}28 \\ -0{,}28 & 0{,}96 \end{pmatrix}.$$

С помощью ортогональных матриц можно также задать определитель. Определитель — это характеристика того, во сколько раз увеличится площадь единичного квадрата после применения линейного преобразования. Следовательно, у ортогональных матриц:

$$\det A = \pm 1.$$

У **симметричных матриц** транспонированная версия совпадает с исходной:

$$A = A^T.$$

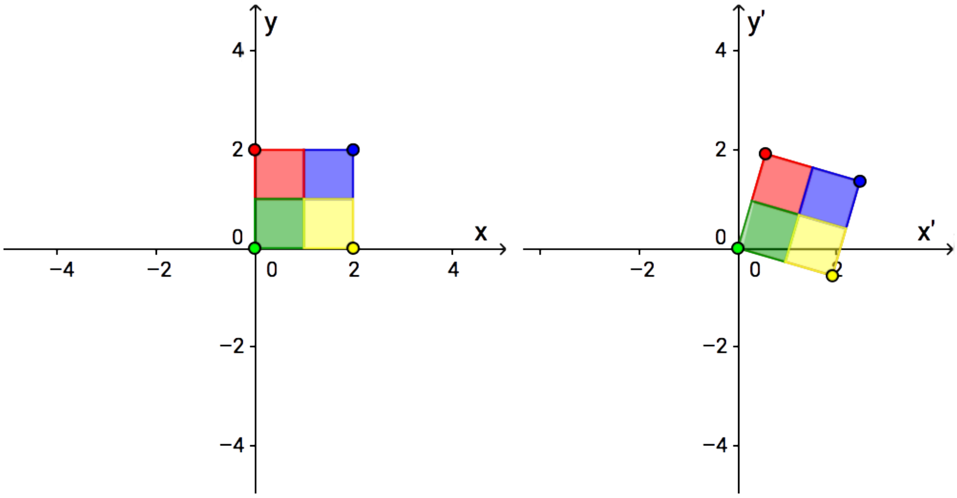

Рис. 8.

Любую симметричную матрицу A можно представить в виде произведения:

$$A = QDQ^T,$$

где D — диагональная матрица, а Q — ортогональная. Например, следующая матрица поворачивает и растягивает единичный квадрат, как показано на рис. 9:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

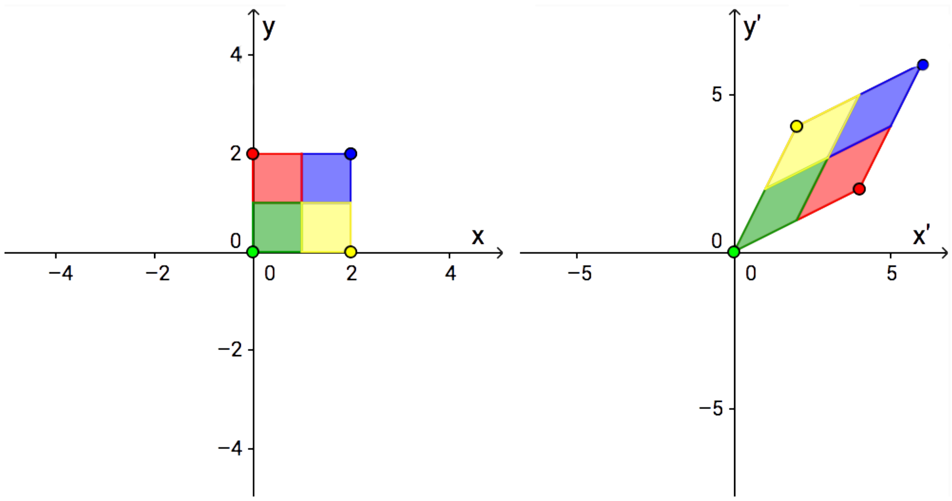

Рис. 9.

6. Собственные числа и векторы

Необходимо понять, что происходит при линейном преобразовании с отдельными векторами. Например, красный вектор на рис. 10 в результате преобразования меняет свое направление, а синий вектор направлен в ту же сторону.

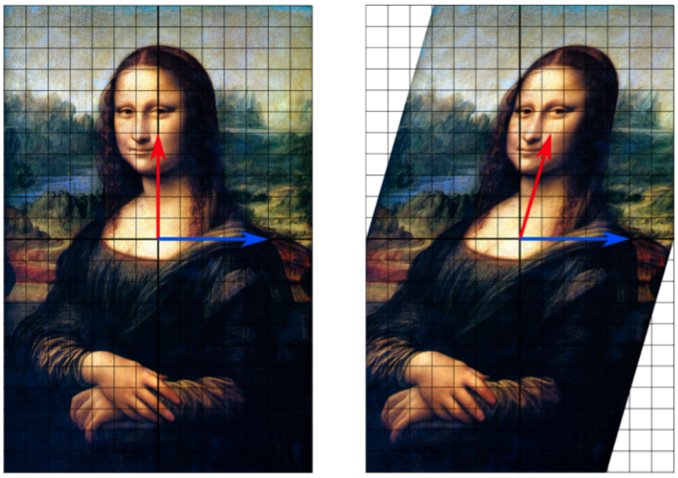

Рис. 10.

Векторы, которые не меняют направление после преобразования, называются **собственными**:

$$A\vec{x} = \lambda\vec{x}, \quad x \neq 0.$$

При этом λ называется **собственным значением**. У матрицы размера $n \times n$ может быть не более n собственных значений.

В качестве примера будет рассмотрена следующая матрица:

$$A = \begin{pmatrix} 2 & 1 \\ 1{,}5 & 2 \end{pmatrix}.$$

Такая матрица переводит единичный квадрат в некоторый ромб (см. рис. 11).

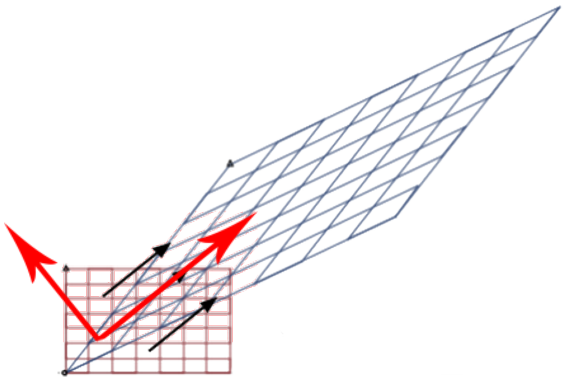

Рис. 11.

Первый собственный вектор и собственное значение матрицы A имеют вид:

$$\vec{\nu}_1 \approx \big(0{,}632456,\ 0{,}774597\big), \quad \lambda_1 \approx 3{,}22474.$$

Второй собственный вектор будет перпендикулярен первому:

$$\vec{\nu}_2 \approx \big(-0{,}632456,\ 0{,}774597\big), \quad \lambda_2 \approx 0{,}775255.$$

Собственные векторы необходимы при решении задач об уменьшении размерности матрицы с максимальным сохранением в ней информации, так как они дают наиболее характерные преобразования матрицы. В задаче о понижении числа признаков собственные векторы будут показывать, на какие оси нужно проецировать данные, чтобы максимально сохранить дисперсию в них (*метод главных компонент*).