

# Linear Algebra

2024-07-01

## 矩阵定义以及运算

### 定义以及概念

矩阵为矩形数表；

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

上三角矩阵  $U$ .

$$\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

下三角矩阵  $L$ .

上下三角矩阵统称为三角矩阵。

$$\begin{bmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$$

对角矩阵  $\Lambda$ .

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

单位矩阵  $E, I$ .

矩阵线性运算

$$c_{ij} = a_{ij} + b_{ij}$$

$$C_{ij} = ac_{ij}$$

矩阵加法和数乘统称为矩阵的线性运算。

矩阵乘法

The element  $C_{ij}$  of the resulting matrix  $C$  is found by taking the dot product of the  $i$ th row of matrix  $A$  and the  $j$ th column of matrix  $B$ . This can be expressed as:

$$C_{ij} = \sum_{k=1}^n A_{ik} \times B_{kj}$$

Where  $A_{ik}$  is the element in the  $i$ th row and  $k$ th column of matrix  $A$ , and  $B_{kj}$  is the element in the  $k$ th row and  $j$ th column of matrix  $B$ .

- $AB = 0 \nRightarrow A = 0$  或  $B = 0$
- $A \cdot B \neq B \cdot A$
- $ABC = (AB)C$
- $A(B + C) = AB + AC$
- $k(AB) = A(kB) = ABk$

空间位置不能变，时间次序可以变。

$$2x_1 + 3x_2 = 8$$

$$x_1 - x_2 = 1$$

Can be represented in matrix form as:

$$\begin{bmatrix} 2 & 3 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \end{bmatrix}$$

**线性方程组的矩阵表示**系数矩阵，变量矩阵，常数矩阵。

对于矩阵乘法有如下理解：

左边乘一个矩阵为对该矩阵列向量的线性组合，线性组合的系数为右侧矩阵的列，该视角最为常用，可用于判断线性方程组解的情况；右乘一个矩阵为对该矩阵的行向量的线性组合，线性组合的系数为左侧矩阵的行，该视角可用于消元。

矩阵乘法也可理解为每一列乘每一行得到多个矩阵，然后相加。

## 矩阵转置

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

The transpose of  $A$ , denoted as  $A^T$  is obtained by flipping  $A$  over its main diagonal, and is defined as:

$$A^T = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}$$

- $(A^T)^T = A$

- $(A+B)^T = A^T + B^T$
- $(kA)^T = kA^T$
- $(AB)^T = B^T A^T$

对称矩阵  $A^T = A$ , 一定为方阵。

反对称矩阵  $A^T = -A$ , 一定为方阵, 对角线一定为 0.

## 分块矩阵

非分块矩阵可以认为每个元素都是一个只有一个数的块。

$$\alpha_1 x_1 + \alpha_2 x_2 = b$$

分块矩阵可用于矩阵的抽象运算, 利用分块矩阵表示线性方程组如上。

## 矩阵的逆与矩阵的初等变化

$$A_n B_n = B_n A_n = E_n$$

逆只能存在方阵之中, 且不一定可逆。逆具有唯一性。

$$\begin{aligned}(A^{-1})^{-1} &= A \\ (A+B)^{-1} &=? \\ (kA)^{-1} &= k^{-1}A^{-1} \\ (AB)^{-1} &= B^{-1}A^{-1} \\ (A^{-1})^T &= (A^T)^{-1}\end{aligned}$$

矩阵逆运算规律如上。矩阵的上标运算可以任意交换位置。

**矩阵初等变换**实际上就是方程组系数消元的抽象过程, 初等变换的一系列矩阵称为等价矩阵。

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xrightarrow{\text{行变换}} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

**初等方阵**单位矩阵  $E$  经过一次初等变换所形成的方阵  $P$ 。若  $P$  为初等方阵,  $PA = B$ , 则  $A \rightarrow B$  的变化等价于  $E \rightarrow P$  的初等行变换。若  $AP = B$ , 则  $A \rightarrow B$  等价于  $E \rightarrow P$  的初等列变换。实际上这里描述的是一个初等变换可由一个初等方阵来定义; 从矩阵乘法的行视角来看, 就是对右侧矩阵行向量的线性组合。

那么有初等方阵的逆矩阵:

1. 通过交换单位矩阵的第  $i$  行和第  $j$  行得到的初等矩阵  $E_{ij}$ , 其逆矩阵也是通过交换同样的两行得到的, 即  $E_{ij}^{-1} = E_{ij}$ 。
2. 将单位矩阵的第  $i$  行乘以一个非零常数  $k$  得到的初等矩阵  $E_i(k)$ , 其逆矩阵是将第  $i$  行乘以  $1/k$ , 即  $E_i(k)^{-1} = E_i(1/k)$ 。
3. 将单位矩阵的第  $i$  行加上第  $j$  行的  $k$  倍得到的初等矩阵  $E_{ij}(k)$ , 其逆矩阵是将第  $i$  行减去第  $j$  行的  $k$  倍, 即  $E_{ij}(k)^{-1} = E_{ij}(-k)$ 。

$$\begin{bmatrix} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \end{bmatrix} \xrightarrow{\text{初等行变换}} \begin{bmatrix} 1 & 0 & a_{11}^{-1} & -a_{12}a_{11}^{-1} \\ 0 & 1 & -a_{21}a_{11}^{-1} & a_{11}^{-1}(a_{11}a_{22} - a_{12}a_{21}) \end{bmatrix}$$

**逆矩阵的求法**将矩阵经过初等变换为单位矩阵, 同样的变化作用于单位矩阵的结果就是其逆矩阵。

## 行列式

**逆序数**在一个排列  $a_1, a_2, \dots, a_n$  中, 如果存在一对  $i, j$ , 满足  $i < j$  且  $a_i > a_j$ , 那么这对  $(a_i, a_j)$  称为一个逆序对。排列中所有逆序对的个数称为该排列的逆序数。

例如, 对于排列  $3, 1, 2$ , 它的逆序对有  $(3, 1)$  和  $(3, 2)$ , 所以它的逆序数是 2。

形式化地, 设排列为  $\sigma = (\sigma_1, \sigma_2, \dots, \sigma_n)$ , 则逆序数定义为:

$$\text{Inv}(\sigma) = \sum_{1 \leq i < j \leq n} \mathbf{1}_{\sigma_i > \sigma_j}$$

其中  $\mathbf{1}_{\sigma_i > \sigma_j}$  是指示函数，当  $\sigma_i > \sigma_j$  时取值为 1，否则取值为 0。

**余子式** 设  $A$  是一个  $n \times n$  的矩阵， $A_{ij}$  是  $A$  中去掉第  $i$  行和第  $j$  列后剩下的  $(n-1) \times (n-1)$  子矩阵。那么， $A$  的  $(i, j)$  元素对应的余子式定义为：

$$M_{ij} = \det(A_{ij})$$

其中  $\det(A_{ij})$  表示  $A_{ij}$  的行列式。

**代数余子式** 设  $A$  是一个  $n \times n$  的矩阵， $A_{ij}$  是  $A$  中去掉第  $i$  行和第  $j$  列后剩下的  $(n-1) \times (n-1)$  子矩阵。那么， $A$  的  $(i, j)$  元素对应的代数余子式定义为：

$$C_{ij} = (-1)^{i+j} \det(A_{ij})$$

**行列式** 行列式 (Determinant) 是一个函数，将方阵映射到标量，其共有  $n!$  项，其中每一项为不同行不同列元素的积；每一项的正负由排列的逆序数决定。记作  $\det(A)$  或  $|A|$ 。设  $A$  是一个  $n \times n$  的方阵，那么  $A$  的行列式  $\det(A)$  定义为：

$$\det(A) = \sum_{\sigma \in S_n} (\text{sgn}(\sigma) \prod_{i=1}^n a_{i, \sigma(i)})$$

其中， $S_n$  表示所有  $n$  个元素的排列， $\sigma$  是一个排列， $\text{sgn}(\sigma)$  表示排列  $\sigma$  的符号， $a_{i, \sigma(i)}$  表示矩阵  $A$  在第  $i$  行第  $\sigma(i)$  列的元素。对于  $n$  阶方阵  $A$ ，其行列式可通过以下方法计算：

- **三角矩阵**：对于三角矩阵，行列式是对角线上元素的代数余子式。
- **展开法**：通过任意一行（或列）的元素进行展开。例如，第  $i$  行的展开公式为：

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij})$$

其中， $a_{ij}$  是矩阵  $A$  中第  $i$  行第  $j$  列的元素， $A_{ij}$  是删除了第  $i$  行和第  $j$  列的子矩阵。

如果按第  $i$  行展开得到代数余子式，分别乘以第  $j(j \neq i)$  行对应的元素，其等于 0

**行列式的基本性质包括**

- 转置相等  $\det(A) = \det(A^T)$

- 换行变号交换行列式的两行（列），行列式变号。
- 乘数乘行行列式的某一行（列）乘以常数，行列式也乘以该常数。
- 倍加相等行列式中某一行（列）的常数倍加到另一行（列），行列式不变。
- 拆分分行行列式可以按行或列拆分。
- 零性质行列式中某一行（列）全为零，或有行或者列成比例，行列式为零。

### 行列式运算规律如下

- $|AB| = |A||B|$
- $|kA| = k^n |A|$
- 分块的对角矩阵的行列式等于对角线元素上的分块的行列式的乘积。

### 行列式计算

- 对角，三角行列式，次（副）对角元素乘积  $\rightarrow$  对角线乘积（正负由逆序数决定）
- 分块矩阵的行列式。
- 一杠一星，两杠一星行列式  $\rightarrow$  只有对角元素项。
- 箭头行列式，弓形行列式  $\rightarrow$  列变换为第一列为 0，列变换为对角行列式。
- 同行（列）行列式  $\rightarrow$  列，行变换为箭头行列式。
- $X$  形行列式  $\rightarrow$
- $ab$  矩阵。
- 范德蒙行列式  $\rightarrow$  确定不变行。

$$V_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ x_1 & x_2 & x_3 & \cdots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \cdots & x_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & x_3^{n-1} & \cdots & x_n^{n-1} \end{vmatrix}$$

### 伴随矩阵

**伴随矩阵**伴随矩阵是与原矩阵密切相关的一个矩阵，它由原矩阵的代数余子式构成。对于一个  $n$  阶方阵  $A$ ，其伴随矩阵记作  $adj(A)$ ，它的每个元素  $a_{ij}$  都是由原矩阵  $A$  的代数余子式  $C_{ij}$  替换得到的。具体来说，伴随矩阵的第  $i$  行第  $j$  列的元素是原矩阵第  $j$  行第  $i$  列的代数余子式，即  $adj(A)_{ij} = C_{ji}$ 。

伴随矩阵与原矩阵的行列式之间有如下关系：

$$A \cdot adj(A) = adj(A) \cdot A = \det(A) \cdot I$$

其中， $I$  是单位矩阵， $\det(A)$  是矩阵  $A$  的行列式。这个关系表明，伴随矩阵可以用于计算矩阵的逆，当  $\det(A) \neq 0$  时，矩阵  $A$  可逆，且其逆矩阵  $A^{-1}$  可以通过伴随矩阵来计算：

$$A^{-1} = \frac{1}{\det(A)} \cdot adj(A)$$

### 对角矩阵（副对角）相关公式

对角矩阵  $A, B$

- $AB = BA \rightarrow$  对角线元素的乘积。
- 对角矩阵的幂  $\rightarrow$  对角线元素的幂。
- 逆  $\rightarrow$  对角线元素的倒数。
- 行列式  $\rightarrow$  对角线元素的乘积。

副对角矩阵

- 逆  $\rightarrow$  副对角线元素的倒数，转置。
- 行列式  $\rightarrow$  副对角线元素乘积，正负由逆序数确定。

**分块对角矩阵相关公式**同上。副对角分块矩阵的行列式为  $-1^{m \cdot n} |A| |B|$

### 矩阵运算规律总结

- $(A + B)$  的上标运算，只有转置有公式。
- 伴随矩阵的运算全都来源于其定义。
- 若  $AB = BA = E$ ，则  $(A + B)$  的幂次运算可按二项式展开。



## 矩阵的秩

**矩阵的秩 (Rank)** 通过其最高阶非零子式来定义。设  $A$  是一个  $m \times n$  的矩阵。矩阵的秩是  $A$  中阶数最大的非零子式的阶数。具体定义如下：

设  $A$  是一个  $m \times n$  的矩阵，若  $A$  中存在一个  $k \times k$  的子矩阵（即从  $A$  中取出  $k$  行  $k$  列构成的子矩阵）的行列式不为零，但所有  $k+1 \times k+1$  的子矩阵的行列式均为零，则称矩阵  $A$  的秩为  $k$ 。

换句话说：

$$\text{rank}(A) = k$$

其中  $k$  是满足矩阵  $A$  中存在  $k \times k$  非零行列式的最大整数。

**行阶梯矩阵 (Row Echelon Matrix)** 行阶梯型矩阵是一种特殊形式的矩阵，它满足以下条件：

- 所有零行（全为零的行）都排在非零行的下面。
- 每一非零行的首个非零元素（称为主元）位于其前一行主元的右边。

$$\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 6 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**行最简形矩阵 (Reduced Row Echelon Form, RREF)** 行最简形矩阵 (Reduced Row Echelon Form, RREF) 是行阶梯矩阵的进一步简化形式，满足以下条件：

- 它是一个行阶梯矩阵。
- 每个主元所在列的其他元素均为零。
- 每个主元为 1。

$$\begin{bmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**矩阵秩的性质如下**

- $R(AB) \leq R(A)$  以及  $R(B)$
- $R(AB) \geq R(A) + R(B) - n$  (若  $AB = 0, R(A) + R(B) \leq n$ )
- $R(A^*) = n, 1, 0 (R(A) = n, R(A) = n - 1, R(A) < n - 1)$

初等变化不改变矩阵的秩，将矩阵经过初等变换为最简形矩阵可用于求矩阵的秩

矩阵视角下的线性方程组

**克莱姆法则 (Cramer's Rule)** 克莱姆法则是一种用行列式求解线性方程组的方法。假设我们有一个线性方程组：

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

我们可以将其表示为矩阵形式：

$$A\mathbf{x} = \mathbf{b}$$

其中， $A$  是系数矩阵， $\mathbf{x}$  是未知数向量， $\mathbf{b}$  是常数向量：

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

克莱姆法则指出，如果行列式  $\det(A) \neq 0$ ，那么线性方程组有唯一解，每个未知数  $x_i$  的解可以通过以下公式求得：

$$x_i = \frac{\det(A_i)}{\det(A)}$$

其中，矩阵  $A_i$  是将矩阵  $A$  的第  $i$  列替换为向量  $\mathbf{b}$  后得到的矩阵。

对于非齐次线性方程组  $A\mathbf{x} = \mathbf{b}$ ，其解的判定如下：

- 如果  $\det(A) \neq 0$ , 则方程组有唯一解。
- 如果  $\det(A) = 0$ , 则方程组有无穷多解, 或者无解。

对于齐次线性方程组  $A\mathbf{x} = \mathbf{0}$ , 其解的判定如下:

- 如果  $\det(A) \neq 0$  则方程组只有零解。
- 如果  $\det(A) = 0$ , 则方程组有无穷多非零解。

**高斯消元法 (初等变换) 解方程组**高斯消元法的抽象就是对矩阵的初等变换。

对于非齐次方程组, 其解的判定如下:

- 若  $\text{rank}([A, b]) > \text{rank}(A)$ , 方程组无解 - 方程角度出现了矛盾方程。
- 若  $\text{rank}([A, b]) = \text{rank}(A)$ , 若  $A_{mn}$ ,  $\text{rank}(A) < n$  则有无穷多解,  $\text{rank}(A) = n$ , 则只有唯一解。 - 方程角度  $\text{rank}(A)$  就是方程组的约束条件,  $n$  就是未知数的个数

齐次方程组, 其解的判定如下:

- 若  $\text{rank}(A_{mn}) = n$ , 则只有零解。
- 若  $\text{rank}(A_{mn}) < n$ , 则有非零解。

## 几何直观

**向量**一个向量完全由其基向量以及各基向量方向上的标量决定, 例如向量  $\alpha = [a, b]$ , 表示其在两个基  $i, j$  的倍数。

**矩阵**一个矩阵代表了空间的变化, 为了确定一个向量在变换后的位置, 只需要追踪基向量的位置。实际上矩阵  $A = [e_1|e_2]$  的每个列向量  $e_1, e_2$  就是变换后的基向量。将每个基向量  $e_1, e_2$  乘以原向量  $\alpha$  每个方向上的标量  $a, b$ , 再加起来就得到变换后的向量  $\beta$ 。

**初等变换**向量的线性组合, 重新组合后的向量仍然在同一向量空间。

**行列式倍加相等**向量的重新线性组合。

$$\beta = ae_1 + be_2$$

**行列式**行列式的几何含义空间变换以后由基向量所张成空间的变化比例。 $|A| = 0$ , 则说明  $A$  并不是满秩的, 会造成空间的降维, 其行列式为 0, 且由于发生了空间的坍塌, 由低维空间返回到

原高维空间是不可能得，即其逆变换不存在，即  $A^{-1}$  不存在。所以实际上行列式，秩，是对矩阵所张成空间的维度的不同角度的描述。

## 向量以及向量空间

### 定义以及定理

**线性组合**线性组合是指通过对向量进行加权求和得到一个新的向量。假设有向量  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  和系数  $c_1, c_2, \dots, c_n$ ，它们的线性组合表示为：

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n$$

**线性表示**线性表示是指一个向量可以表示为其他向量的线性组合。假设向量  $\mathbf{v}$  可以表示为向量  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$  的线性组合，则有：

$$\mathbf{v} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n$$

**线性相关**如果存在一组不全为零的标量  $c_1, c_2, \dots, c_n$ ，使得这些向量的线性组合为零向量：

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \mathbf{0}$$

换句话说，如果至少有一个向量可以表示为其他向量的线性组合，则这些向量是线性相关的。对应于齐次方程组有非零解。

**线性无关**如果只有当所有标量  $c_1, c_2, \dots, c_n$  均为零时，这些向量的线性组合才为零向量：

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n = \mathbf{0} \implies c_1 = c_2 = \dots = c_n = 0$$

换句话说，任何一个向量都不能表示为其他向量的线性组合，则这些向量是线性无关的。对应于齐次方程组只有零解。

假设我们有一组线性无关的向量  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ 。这些向量的线性组合只有在所有系数都为零时，才会等于零向量：

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n = \mathbf{0} \implies c_1 = c_2 = \cdots = c_n = 0$$

现在, 假设我们引入一个新向量  $\mathbf{u}$ 。如果将这个新向量加到上述向量组中后变得线性相关, 则有:

$$c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_n \mathbf{v}_n + c_{n+1} \mathbf{u} = \mathbf{0}$$

并且存在不全为零的系数  $c_1, c_2, \dots, c_n, c_{n+1}$  使得上式成立。

由于  $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$  是线性无关的, 因此  $\mathbf{u}$  可以唯一地表示为这些向量的线性组合。

**向量组的部分与整体定理** 整体线性无关, 那么部分线性无关; 部分线性相关, 那么整体线性相关。

**向量组的伸长与缩短定理** 短线性无关, 那么长线性无关; 长线性相关, 那么短线性相关。

**极大无关组与向量组的秩** 向量组中无关向量的数量, 数量上等于矩阵的秩。

**向量个数与向量维数**

**等价向量组** 向量组与它的任意极大无关组等价。

**向量组的秩与向量组的线性表示定理** 如果向量组  $\alpha_1, \alpha_2, \dots$ , 可由向量组  $\beta_1, \beta_2, \dots$  表示, 那么  $\text{rank}(B) \geq \text{rank}(A)$ 。

**向量组的臃肿与紧凑定理**

**向量空间** 是一个由向量构成的集合, 这个集合对于向量的加法和标量乘法是封闭的。更正式地说, 一个集合  $V$  配备了一对操作: 加法 (记作  $+$ ) 和标量乘法 (记作  $\cdot$ ), 如果满足以下性质, 则称  $V$  为一个向量空间:

- 封闭性, 对于所有  $\mathbf{u}, \mathbf{v} \in V$ , 有  $\mathbf{u} + \mathbf{v} \in V$  和  $c \cdot \mathbf{u} \in V$ , 其中  $c$  是任意标量。
- 结合律, 对于所有  $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ , 有  $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ 。
- 交换律, 对于所有  $\mathbf{u}, \mathbf{v} \in V$ , 有  $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ 。
- 存在零向量, 存在一个向量  $\mathbf{0} \in V$ , 使得对于所有  $\mathbf{v} \in V$ , 有  $\mathbf{v} + \mathbf{0} = \mathbf{v}$ 。
- 存在加法的逆元, 对于每个  $\mathbf{v} \in V$ , 存在一个向量  $-\mathbf{v} \in V$ , 使得  $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ 。
- 标量乘法的分配律, 对于所有  $c, d$  是标量, 和所有  $\mathbf{u}, \mathbf{v} \in V$ , 有  $c \cdot (\mathbf{u} + \mathbf{v}) = c \cdot \mathbf{u} + c \cdot \mathbf{v}$  和  $(c + d) \cdot \mathbf{u} = c \cdot \mathbf{u} + d \cdot \mathbf{u}$ 。
- 标量乘法的单位元, 对于所有  $\mathbf{v} \in V$ , 有  $1 \cdot \mathbf{v} = \mathbf{v}$ , 其中  $1$  是标量乘法的单位元。

**向量在基下的坐标**在向量空间中，一个向量  $\mathbf{v}$  可以通过基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  下的坐标来唯一表示。如果  $\mathbf{v}$  可以表示为基向量的线性组合，那么存在一组系数  $x_1, x_2, \dots, x_n$  使得：

$$\mathbf{v} = x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \dots + x_n\mathbf{e}_n$$

系数  $x_1, x_2, \dots, x_n$  就是向量  $\mathbf{v}$  在基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  下的坐标。

**过渡矩阵**过渡矩阵是一个与基变换相关的矩阵。如果有两组基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  和  $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ ，过渡矩阵  $P$  是一个矩阵，它的列是第二个基向量在第一个基下的坐标。如果  $\mathbf{f}_i$  在基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$  下的坐标是列向量  $\mathbf{p}_i$ ，那么过渡矩阵  $P$  可以表示为：

$$P = [\mathbf{p}_1 | \mathbf{p}_2 | \dots | \mathbf{p}_n]$$

过渡矩阵可以用来转换一个向量在两个基下的坐标。

**向量的内积**两个向量  $\mathbf{u}$  和  $\mathbf{v}$  的内积（点积）是一个标量，定义为：

$$\mathbf{u} \cdot \mathbf{v} = u_1v_1 + u_2v_2 + \dots + u_nv_n$$

内积也可以表示为：

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \theta$$

其中  $\theta$  是向量  $\mathbf{u}$  和  $\mathbf{v}$  之间的夹角。

**向量的长度（范数）**向量  $\mathbf{v}$  的长度或范数是向量的内积的平方根，通常指欧几里得范数：

$$\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

**向量的夹角**两个非零向量  $\mathbf{u}$  和  $\mathbf{v}$  之间的夹角  $\theta$  可以通过它们的内积和范数来计算：

$$\cos \theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \|\mathbf{v}\|}$$

**正交基**如果一组基向量中的任意两个都是正交的，即它们的内积为零，那么这组基称为正交基。对于正交基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ，满足：

$$\mathbf{e}_i \cdot \mathbf{e}_j = 0 \quad \text{for } i \neq j$$

**标准正交基**如果正交基中的每个基向量的长度都是 1，那么这组基称为标准正交基。对于标准正交基  $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ ，满足：

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$$

其中  $\delta_{ij}$  是克罗内克 delta 函数。

**正交矩阵**正交矩阵是一个方阵，其列向量构成一个标准正交基。对于一个  $n \times n$  的正交矩阵  $Q$ ， $Q^T Q = E$ ；它满足以下条件：

- 列向量是线性无关的。
- 列向量的长度都是 1。
- 列向量与自身正交，即对于任意两个不同的列向量  $\mathbf{q}_i$  和  $\mathbf{q}_j$  ( $i \neq j$ )，满足：

$$\mathbf{q}_i \cdot \mathbf{q}_j = \delta_{ij}$$

其中  $\delta_{ij}$  是克罗内克 delta 函数。

### 正交矩阵的性质

- 正交矩阵上标运算也是正交矩阵。
- 正交矩阵的转置等于其逆矩阵，即  $Q^T = Q^{-1}$ 。
- 若  $A, B$  都为正交矩阵，则  $AB$  及  $BA$  都是正交矩阵。
- $(A\alpha, A\beta) = (\alpha, \beta)$ ,  $\|A\alpha\| = \|\alpha\|$ ,  $\langle A\alpha, A\beta \rangle = \langle \alpha, \beta \rangle$
- 正交矩阵的行列式等于  $\pm 1$ ，即  $\det(Q) = \pm 1$ 。
- 正交矩阵的奇异值都是 1 或 -1。
- 正交矩阵的特征值是  $\pm 1$ 。

## 向量空间视角下的线性方程组

**方程组表示 (System of Equations)** 线性方程组可以直接以方程的形式表示。假设有  $n$  个方程和  $m$  个未知数  $x_1, x_2, \dots, x_m$ , 方程组形式如下:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2m}x_m = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nm}x_m = b_n \end{cases}$$

**矩阵表示 (Matrix Form)** 使用矩阵表示, 可以将上述线性方程组表示为:

$$A\mathbf{x} = \mathbf{b}$$

其中, 矩阵  $A$  和向量  $\mathbf{x}$ 、 $\mathbf{b}$  分别为:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

**向量表示 (Vector Form)** 向量表示法强调线性组合的观点, 可以写成:

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \dots + a_m\mathbf{x}_m = \mathbf{b}$$

其中  $a_i$  表示向量  $\mathbf{x}_i$  的系数。

### 齐次方程组的解

齐次方程组是由线性方程组成的方程组, 其形式如下:

$$Ax = 0$$

假设  $\xi_1, \xi_2$  是其解向量, 那么  $k\xi_1, \xi_1 + \xi_2, k\xi_1 + k\xi_2$  也是其解向量, 这些所有的向量构成解空间, 其中任意一组解称为其基础解系, 也就是解空间的一组基。



其中  $A$  是一个  $m \times n$  的矩阵,  $x$  是一个  $n \times 1$  的列向量。齐次方程组的通解是指满足方程组的解的集合。对于齐次方程组, 如果  $A$  的秩  $r(A) < n$ , 则齐次方程组有非零解。这些解可以表示为  $A$  的零空间的基向量的线性组合, 其解的向量个数为  $n - r(A)$  即:

$$x = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_k \mathbf{v}_k$$

其中  $c_1, c_2, \dots, c_k$  是任意常数,  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  是  $A$  的零空间的基向量。

### 非齐次方程组的通解

导出组, 非齐次方程组对应的齐次方程组。

- 若  $\xi_1, \xi_2$  是非齐次方程组的两个解, 则  $\xi_1 - \xi_2$  是其导出组  $Ax = 0$  的解。
- 若  $n$  是非齐次线性方程组  $Ax = b$  的解,  $\xi$  是其导出组的解, 那么  $\xi + n$  是  $Ax = b$  的解。

非齐次方程组是由线性方程组成的方程组, 其形式如下:

$$Ax = b$$

其中  $A$  是一个  $m \times n$  的矩阵,  $x$  是一个  $n \times 1$  的列向量,  $b$  是一个  $m \times 1$  的列向量。非齐次方程组的通解是指满足方程组的解的集合。对于非齐次方程组, 如果  $A$  的秩  $r(A) = n$ , 则方程组有唯一解。这个解可以表示为  $A$  的零空间的基向量的线性组合加上  $A$  的零空间的补空间的基向量的线性组合, 即:

$$x = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_k \mathbf{v}_k + d_1 \mathbf{w}_1 + d_2 \mathbf{w}_2 + \cdots + d_m \mathbf{w}_m$$

其中  $c_1, c_2, \dots, c_k, d_1, d_2, \dots, d_m$  是任意常数,  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$  是  $A$  的零空间的基向量,  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m$  是  $A$  的零空间的补空间的基向量。

非齐次线性方程组的通解就是: 导出组的通解 + 其一个特解。

## 几何直观

**向量空间中有方向的箭头。**

**点积**向量  $\alpha$  投影到  $\beta$  上的长度与向量  $\beta$  长度的乘积。对于点积运算，向量总是对应一个对偶向量，其为  $\alpha^T$ ，也就是说点积也可以理解为向量  $\beta$  经过线性变换  $\alpha^T$  后，压缩为一个一维的向量。如果  $\|\alpha\| = 1$ ，则  $\alpha, \beta$  的点积为  $\beta$  在单位基向量  $\alpha$  的表示。这个应用在 PCA 等线性降维方法中。

**齐次方程组的自由变量**列向量的个数 - 列向量的极大无关组。多余的向量的系数可以任意取，最后由极大无关组，即这个空间的基向量，决定最终的系数。

**线性方程组的向量张成的空间**无解，则说明  $\alpha_1, \alpha_2, \dots, \alpha_n$  与  $b$  不在同一个向量空间，若有解则说明  $\alpha_1, \alpha_2, \dots, \alpha_n$  与  $b$  在同一个向量空间。进一步地，若  $\alpha_1, \alpha_2, \dots, \alpha_n$  的数量大于  $A$  空间维数，则有多余的向量，那么构成  $b$  的方式则不止一个；若  $\alpha_1, \alpha_2, \dots, \alpha_n$  的数量等于  $A$  空间维数，则只有一种表示方法。

## 矩阵变换

### 特征值 (Eigenvalues) 特征向量 (Eigenvectors)

**特征值**是线性代数中矩阵理论的一个基本概念。对于一个给定的方阵  $A$ ，特征值是满足方程  $Av = \lambda v$  的数  $\lambda$ ，其中  $v$  是非零向量，称为特征向量。特征值揭示了矩阵的某些代数属性，比如矩阵的行列式和迹（对角线元素之和）与特征值有关。

**特征向量**是与特征值相关联的非零向量，它表示了矩阵  $A$  在某个方向上的拉伸或压缩的比例。当一个矩阵作用在它的特征向量上时，结果是该特征向量的标量倍，这个标量就是相应的特征值。

由定义可得：

$$(A - \lambda E)v = 0$$

由特征值的定义， $v \neq 0$ ，则该齐次方程组有非零解，即  $|A - \lambda E| = 0$ ，可求得特征值，将特征值代入  $(A - \lambda E)v = 0$  即可得到该特征值对应的特征向量。

### 关于特征值的定理

- 特征值的和等于矩阵的迹  $tr(A)$ 。
- 特征值的积等于矩阵行列式。
- 考虑一个变换  $f$  其作用于矩阵  $A$ , 那么其对应的特征值有同样的变换, 但是  $f(\lambda)$  对应的特征向量不变。
- 矩阵互不相等的特征值对应的特征向量是线性无关的。
- 几何重数不大于代数重数。
- 转置矩阵的特征值相等。
- 若  $n$  阶矩阵  $g(A) = 0$ , 那么其所有特征值  $g(\lambda) = 0$

#### 关于实对称矩阵特征值的定理

- 特征值都是实数, 对应的特征向量也是实向量。
- 不同特征值对应特征向量两两正交。
- 几何重数等于其代数重数。

#### 相似矩阵矩阵对角相似化

矩阵  $A$  与  $B$  相似定义为:

$$B = P^{-1}AP$$

相似矩阵的性质:

- $A$  与  $B$  等价。
- $R(A) = R(B)$
- $\lambda_A = \lambda_B$
- $\det(A) = \det(B)$
- $|A| = |B|$
- $tr(A) = tr(B)$
- $f(A) = f(B)$

**相似对角化**将矩阵  $A$  与一个对角矩阵  $\Lambda$  进行相似。实际上, 对于形如  $P^{-1}AP$  的矩阵变化或者相似的过程, 其几何含义为将  $A$  所表示的线性变化转换为以矩阵  $P$  的列向量为基向量的坐标系下的线性变化。相似矩阵为同一线性变化在不同基下的描述。特别地, 为了得到与  $\Lambda$  的相似, 那么选择  $A$  的特征向量构成的基, 则该线性变化在特征基的描述下即为对角矩阵。

$$P^{-1}AP = \Lambda$$

**相似对角化充要条件：**

- $n$  个线性无关的特征向量
- 代数重数等于几何重数。

**求解**矩阵  $P$  就是  $A$  的特征向量按列构成的矩阵。

**实对称矩阵的相似对角化**对于实对称矩阵  $A$ ，总可以找到其正交的特征向量，并将其单位化为正交矩阵  $Q$ ，使得其相似对角化：

$$Q^{-1}AQ = \Lambda \rightarrow Q^T AQ = \Lambda$$

## 二次型

**二次型**定义为二次多项式函数的矩阵形式，其中  $A$  为实对称矩阵。

$$f = x^T Ax$$

**合同矩阵**对于合同矩阵，矩阵  $A$  实际上为一个  $(0, 2)$  的张量，其是空间或流形几何一个量，称之为度规，形如  $C^T AC$  的形式实际上是将某个度规，转化为以矩阵  $C$  的列向量为基的坐标下的描述。而对于二次型而言，该矩阵为实对称矩阵，形似上其相似与合同一致。

$$B = C^T AC$$

- $r(A) = r(B)$
- 若  $A$  为对称矩阵，那么  $B$  也为对称矩阵。

**二次型化为标准形：**

- 令  $x = Qy$ ，则  $x^T Ax = y^T Q^T A Q y$ ，其中矩阵  $Q$  为矩阵  $A$  特征向量对应的单位正交特征向量。(实际上为实对称矩阵相似对角化的过程，实对称矩阵的相似和合同一致)
- 配方法。

**惯性定理：**无论怎样的可逆的线性变换使得其为标准二次型，其正，负平方项的个数一样。

**正定（负定）矩阵：**假设有二次型  $f$ , 若  $f > 0$ , 则为正定矩阵  $f \geq 0$  则为负正定矩阵。

**正定矩阵的性质如下**

**负定矩阵的性质如下**

**等价，相似，合同的判定**

- 秩相等，则等价；反之，亦成立。
- 特征值相等，则可与同一个对角矩阵相似，则相似。
- 正负惯性指数，或者正负特征值一致，则合同。

## 几何直观

**等价矩阵**对于形如  $PAQ = B$  形式的变换，其中  $P, Q$  为可逆矩阵。左乘一个可逆矩阵等价于进行初等行变换，右乘一个可逆矩阵等价于进行初等列变换。几何上， $B$  是  $A$  经过初等变换得到的矩阵，两者等价意味着个矩阵所张成的行空间与列空间是等价的。

**相似矩阵**对于形如  $P^{-1}AP$  的矩阵变化或者相似的过程，其几何含义为将  $A$  所表示的线性变化转换为以矩阵  $P$  的列向量的为基向量的坐标系下的线性变化。相似矩阵为同一线性变化在不同基下的描述。特别地，为了得到与  $A$  的相似，那么选择  $A$  的特征向量构成的基，则该线性变化在特征基的描述下即为对角矩阵。

**合同矩阵**对于形如  $P^TAP$  的形式，矩阵  $A$  实际上为一个  $(0, 2)$  的张量，其是空间或流形几何一个量，称之为度规，形如  $C^TAC$  的形式实际上是将某个度规，转化为以矩阵  $C$  的列向量为基的坐标下的描述。特别地，对于二次型而言，该矩阵为实对称矩阵，形似上其相似与合同一致。

## 矩阵分解

$$A = LU$$

$LU$  分解将矩阵分解为一个下三角以及上三角矩阵乘积；实际上，对于线性方程组进行高斯消元时，在不设计行交换时，每一次的行变换都可以用一个下三角矩阵表示，那么其乘积以及逆都是下三角矩阵  $L$ ，该下三角矩阵，的每一个系数记录了每次操作的消元系数。

$$\begin{bmatrix} 1 & 0 & 0 \\ a_{21} & 1 & 0 \\ a_{31} & a_{32} & 1 \end{bmatrix}$$

如上矩阵表示在对一列进行消元时，进行的操作  $r_2 = r_2 + a_{21}r_1$ ,  $r_3 = a_{31}r_1$ ；在第二列进行消元时， $r_3 = a_{32}r_2$

$$A = P\Lambda P^{-1}$$

考虑方阵  $A$ ， $A$  代表了一个空间变换，空间的变换只对该矩阵的特征向量起到缩放作用。那么  $A$  所代表的空间变换，我们可以考虑先把向量  $x$  转化为用矩阵  $A$  的特征向量构成的基底表示，即  $Q^{-1}x$ ，其中  $Q$  为方阵  $A$  的特征向量构成的矩阵（为什么是  $Q^{-1}$  而不是  $Q$ ？因为特征向量是从原矩阵求解而来，它们都是在同一视角下的表示，要变化到特征向量的视角则是其逆  $Q^{-1}$ ）。那么矩阵  $A$  对这个新的  $x$ ，即  $Q^{-1}x$  的作用就仅仅是缩放（在每个维度上进行缩放），这个缩放用一个新的矩阵  $\Sigma$  表示，其为对角矩阵，对角线上即是对应的特征值，那么现在整个变化记为  $\Sigma Q^{-1}x$ 。完成变化后，我们回到原来的基的视角，即左乘  $Q$ ，整个变化即如下：

$$Ax = Q\Sigma Q^{-1}x \rightarrow A = Q\Sigma Q^{-1}$$

$$A = U\Sigma V^*$$

考虑向量  $[V_1 \ V_2]$ ，其为一组正交的向量，如果其在经历空间变换  $M$  以后仍然映射为一组正交的向量  $[U_1 \ U_2]$ 。那么，我们直接在向量  $[U_1 \ U_2]$  的方向上选择一组基  $[u_1 \ u_2]$ ，那么向量  $[v_1 \ v_2]$ ，在经历空间变化以后，在基  $[u_1 \ u_2]$  上的表示为， $[u_1 \ u_2] * \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$ 。即：

$$M[v_1, v_2] = [u_1, u_2] * \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} \rightarrow MV = U\Sigma \rightarrow M = U\Sigma V^T$$

如此，对于奇异值分解，我们有如下的直观理解：对于矩阵  $M$ ，其对一组向量  $V$  在变换以后仍然为正交的向量。在进行变换的时候考虑直接转换到向量  $V$  的视角之下，然后进行缩放  $\Sigma$ ，以及其它变换  $E$ （旋转，投影），最后变换回原来的视角。即：

$$M = VE\Sigma V^T \rightarrow M = U\Sigma V^T$$

可以看到奇异值与特征值分解的区别在于，选择的视角不同。当选择特征向量视角时，变换只会有缩放即  $\Sigma$  变换，当选择任意的正交向量视角时，变换不仅包含缩放还含有旋转以及投影等即  $E\Sigma$ 。

需要注意的是，虽然对于特征分解以及奇异值分解从视角转换的角度去解释了。但是，矩阵分解同样可以从空间变换的角度理解。例如对于奇异值分解有如下解释：

第一个变换  $V^T$  将单位正交向量  $v_1, v_2$  转到水平和垂直方向、 $\Sigma$  相当于对  $v_1, v_2$  进行放缩、 $U$  将放缩后的向量旋转到最后的位置。

## 线性代数的应用

### 线性降维

线性降维方法主要就是 *PCA*(*principal componet analysis*) 以及 *SVD*(*sigular value decomposition*)，以及 *PCA* 的扩展 *MDS*(*Multtidimensional Scaling*)。*PCA* 与 *MDS* 的差别在于 *PCA* 考虑的是样本的特征，寻求在低维空间下保留方差较大的特征的信息，所以通过对特征之间的协方差矩阵进行特征值分解矩阵，得到在低维空间下能够保留方差较大特征的正交基，是对特征的线性组合，而 *MDS* 考虑的是样本之间的相似度矩阵，通过对相似度矩阵矩阵进行特征分解找到在低维空间下能够保留样本最大距离的正交基。

#### PCA

样本矩阵  $m * n$ ， $n$  个样本， $m$  个特征。

1. 计算样本矩阵协方差矩阵。
2. 协方差矩阵特征分解，找到使得协方差最大的方向。
3. 特征向量单位化 (使得其长度为 1)，特征向量与样本做点积，因为特征向量长度为 1，其点积就是样本在特征向量方向上的投影长度，代表样本在这个方向上的坐标，选择多少个特征向量就为降到多少维。

#### MDS

*MDS* 是一类基于样本距离进行映射的方法。*MDS* 针对不同类型的数据有不同的方法。*PCoA* 是 *MDS* 针对于数值数据分析的方法。

## PCoA

考虑一个样本的 *dissimilarity matrix*, 也就是一个含有样本之间距离或不相似度量的矩阵, 记为  $D$ 。

- *The Torgerson method*

首先对  $D$  进行 *double centering* 得到 *double-centered matrix*, 记为  $B$ 。然后对矩阵  $B$  进行奇异值分解。

*double centering1*):

$$D^2 = [d_{ij}^2]$$

*double centering2*):

$$B = -\frac{1}{2}CD^2C$$

其中  $C = I - \frac{1}{n}J_n$

- *The iterative method*

该方法更加实用, 可以用于非欧式距离矩阵。

$$Stress_D(x_1, x_2, \dots, x_N) = \left( \sum_{i \neq j=1, \dots, N} (d_{ij} - \|x_i - x_j\|)^2 \right)^{\frac{1}{2}}$$