人工智能顶级实战工程师就业课程

数学基础-矩阵和线性代数

北风网 IBEIFENG.COM

主讲人：米老师

线性代数

- 线性(linear)指量(变量)与量(变量)之间按比例、成直线关系，在数学上可以理解为一阶导数为常数的函数；而非线性(non-linear)是指不成比例、没有直线关系，一阶导数不是常数的函数。
- 线性代数中的基本量指的是向量，基本关系是严格的线性关系；也就是可以简单的将线性代数理解为向量与向量之间的线性关系的映射。

一、向量

- 向量：是指具有n个互相独立的性质(维度)的对象的表示，向量常使用字母+箭头的形式进行表示，也可以使用几何坐标来表示向量，比如 $\vec{a}=\overrightarrow{OP}=xi+yj+zk$ ，可以用坐标(i,j,k)表示向量a
- 向量的模：向量的大小，也就是向量的长度，向量坐标到原点的距离，常记作|a|
- 单位向量：长度为一个单位(即模为1)的向量就叫做单位向量

向量的运算

- 设两向量为：$\vec{a}=(x_1,y_1)$，$\vec{b}=(x_2,y_2)$
- 向量的加法/减法满足平行四边形法则和三角形法则

$$\vec{a}+\vec{b}=(x_1+x_2,y_1+y_2)$$

$$\vec{a}-\vec{b}=(x_1-x_2,y_1-y_2)$$

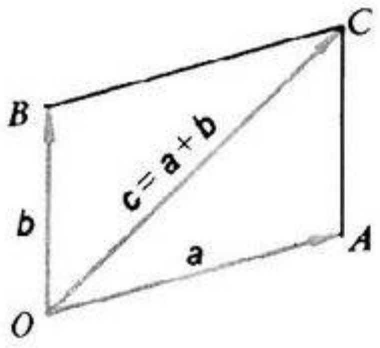


图 4 向量的加法

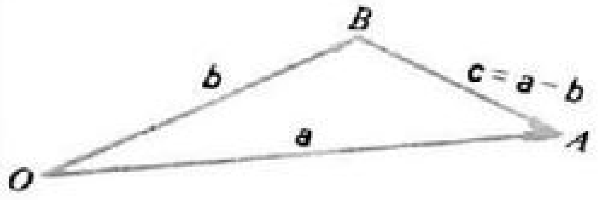


图 5 向量的减法

- 数乘：实数λ和向量a的乘积还是一个向量，记作λa，且|λa|= |λ| |a|；数乘的几何意义是将向量a进行伸长或者压缩操作

$$\lambda\vec{a}=(\lambda x_1,\lambda y_1)$$

向量的运算

- 设两向量为：$\vec{a}=(x_1,y_1)$，$\vec{b}=(x_2,y_2)$，并且a和b之间的夹角为:θ
- 数量积：两个向量的数量积(内积、点积)是一个数量/实数，记作 $\vec{a}\cdot\vec{b}$

$$\vec{a}\cdot\vec{b}=\left|\vec{a}\right|*\left|\vec{b}\right|*\cos\theta$$

- 向量积：两个向量的向量积(外积、叉积)是一个向量，记作 $\vec{a}\times\vec{b}$；

向量积即两个不共线非零向量所在平面的一组法向量。

$$\left|\vec{a}\times\vec{b}\right|=\left|\vec{a}\right|*\left|\vec{b}\right|*\sin\theta$$

正交向量

- 正交向量：如果两个向量的点积为零，那么称这两个向量互为正交向量；在几何意义上来讲，正交向量在二维/三维空间上其实就是两个向量垂直。
- 如果两个或者多个向量，它们的点积均为0，那么它们互相称为正交向量。

二、矩阵

- 矩阵：即描述线性代数中线性关系的参数，即矩阵是一个线性变换，可以将一些向量转换为另一些向量。
- 初等代数中，y=ax表示的是x到y的一种映射关系，其中a是描述这中关系的参数。
- 线性代数中，Y=AX表示的是向量X和Y的一种映射关系，其中A是描述这种关系的参数。

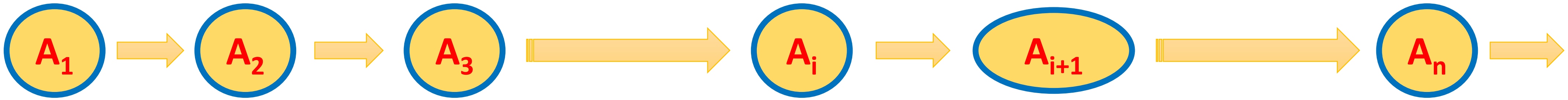
矩阵的直观表示

- 数域F中m*n个数排成m行n列，并括以圆括弧(或方括弧)的数表示成为数域F上的矩阵，通常用大写字母记作A或者A_{m*n}，有时也记作$A=(a_{ij})_{m*n}(i=1,2...,m;j=1,2,...n)$，其中$a_{ij}$表示矩阵A的第i行的第j列元素，当F为实数域R时，A叫做实矩阵，当F为复数域C时，A叫做复矩阵。

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

状态转移模型（讲完概率后再讲）

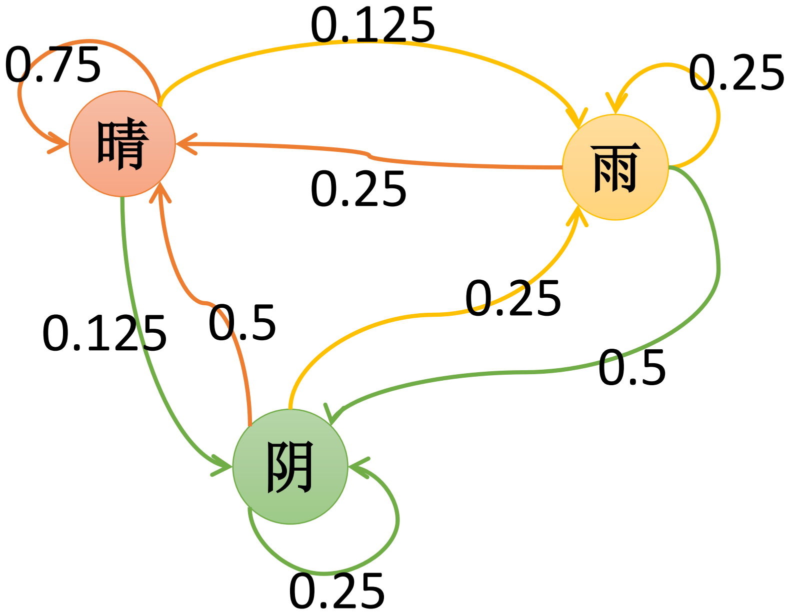
- 考虑某随机过程π，它的状态有n个，使用1~n表示。记在当前时刻t时位于i状态，它在t+1时刻位于j状态的概率为P(i,j)=P(j|i)；
    - 即状态转移的概率只依赖上一个状态值



状态转移模型

- 设将天气状态分为晴、阴、雨三种状态，假定某天的天气状态只和上一天的天气状态有关，状态使用1(晴)、2(阴)、3(雨)表示，转移概率矩阵如下：

今/明	晴	阴	雨
晴	0.75	0.125	0.125
阴	0.5	0.25	0.25
雨	0.25	0.5	0.25



概率转移矩阵

- 第n+1天天气状态为j的概率为：

$$\pi\left(X_{n+1}=j\right)=\sum_{i=1}^{K}\pi\left(X_{n}=i\right)\cdot P\left(X_{n+1}=j \mid X_{n}=i\right)$$

$$\Rightarrow \pi^{n+1}=\pi^{n}\cdot P$$

- 因此，矩阵P即为条件概率转移矩阵。
  - 矩阵P的第i行元素表示，在上一个状态为i的时候的分布概率，即每行元素的和必须为1

初始概率π[0.5,0.3,0.2]

$$P=\begin{bmatrix}0.75 & 0.125 & 0.125\\ 0.5 & 0.25 & 0.25\\ 0.25 & 0.5 & 0.25\end{bmatrix}$$

第n天	晴	阴	雨
0	0.5	0.3	0.2
1	0.575	0.2375	0.1875
2	0.5969	0.225	0.1781
3	0.6047	0.2199	0.1754
4	0.6073	0.2183	0.1744
5	0.6082	0.2177	0.1741
6	0.6085	0.2175	0.174
7	0.6086	0.2174	0.1739
8	0.6087	0.2174	0.1739
9	0.6087	0.2174	0.1739
10	0.6087	0.2174	0.1739
11	0.6087	0.2174	0.1739

初始概率π[0.1,0.6,0.3]

$$P=\begin{bmatrix}0.75 & 0.125 & 0.125\\ 0.5 & 0.25 & 0.25\\ 0.25 & 0.5 & 0.25\end{bmatrix}$$

第n天	晴	阴	雨
0	0.1	0.6	0.3
1	0.45	0.3125	0.2375
2	0.5531	0.2531	0.1937
3	0.5898	0.2293	0.1809
4	0.6022	0.2215	0.1763
5	0.6065	0.2188	0.1747
6	0.6079	0.2179	0.1742
7	0.6084	0.2176	0.174
8	0.6086	0.2174	0.1739
9	0.6087	0.2174	0.1739
10	0.6087	0.2174	0.1739
11	0.6087	0.2174	0.1739

平稳分布

- 初始概率不同，但经过若干次迭代后，π最终会收敛在某个分布上。
- 从而，这是转移概率矩阵P的性质，而非初始分布的性质，事实上，上述天气的转移矩阵P的n次幂，最终会收敛到(0.6087, 0.2174, 0.1739)，n>10

矩阵与向量

- 当m=1或者n=1的时候，称A为行向量或者列向量

$$\mathrm{A} = \begin{bmatrix} \mathbf{a}_{11} & a_{12} & \cdots & a_{1n} \end{bmatrix} \qquad A = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}$$

矩阵相等

- 对于两个矩阵A和B，当它们的行数相同，列数相同，并且对应位置上的元素都相等时，称矩阵A与B相等，记做A=B
- 即$a_{ij}=b_{ij}$，对所有i=1,2,......m;j=1,2......n都成立
- 若两个矩阵行数与列数分别相等，则为同型矩阵，如下：

$$\begin{bmatrix} 1 & 2 \\ 5 & 6 \\ 4 & 7 \end{bmatrix} \text{与} \begin{bmatrix} 14 & 3 \\ 7 & 4 \\ 5 & 9 \end{bmatrix}$$

方阵

- 如果矩阵A中m等于n，那么称矩阵A为n阶矩阵(或n阶方阵)。

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

从左上到右下的对角线称为主对角线

从右上到左下的对角线称为次对角线

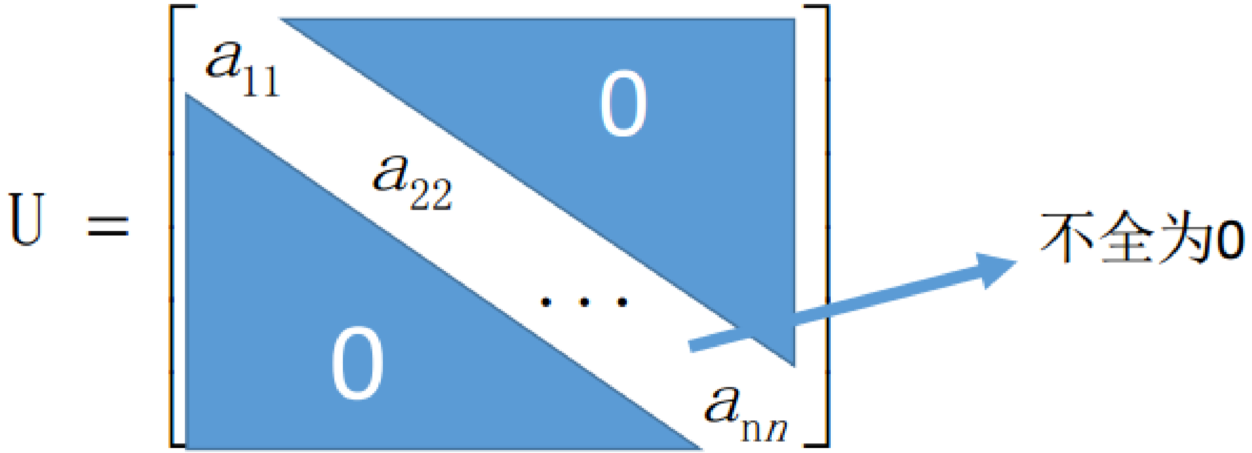
负矩阵、上三角矩阵、下三角矩阵

- 对于矩阵A（所有元素m*n），各个元素取相反数得到的矩阵称为为A的负矩阵，记为-A
- 对于矩阵 $A=\begin{bmatrix}1 & 2\\ 5 & 6\end{bmatrix}$ 的负矩阵为 $-A=\begin{bmatrix}-1 & -2\\ -5 & -6\end{bmatrix}$
- 上三角阵

$$U=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ & a_{22} & \cdots & a_{2n}\\ 0 & & \cdots & \cdots\\ & & & a_{nn}\end{bmatrix}\qquad L=\begin{bmatrix}a_{11} & & 0 & \\ a_{21} & a_{22} & & \\ \cdots & \cdots & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{nn}\end{bmatrix}$$

对角矩阵

- 既是上三角阵，又是下三角阵的矩阵称为对角方阵或对角矩阵
- 记为 $D = diag[a_{11} \quad a_{22} \quad \cdots \quad a_{nn}]$



- 对角元的和 $\sum_{i=1}^{n} a_{ii}$ 称为方阵A的迹，记为trA

单位矩阵

- 单位矩阵：n阶方阵中除了主对角线上的元素外，其它元素均为0，主对角线元素均为1，那么此时的n阶方阵叫做n阶单位矩阵。单位矩阵常用E或者I表示。

$$E = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

零矩阵

- 如果矩阵A中的所有元素(m*n个)均为0，那么此时矩阵A叫做零矩阵，可以记作0。

$$B=\begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

矩阵的加减法

- 矩阵的加法与减法要求进行操作的两个矩阵A和B具有相同的阶，假设A为m*n阶矩阵，B为m*n阶矩阵，那么C=A ± B也是m*n阶的矩阵，并且矩阵C的元素满足：$c_{ij} = a_{ij} \pm b_{ij}$

$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix} \quad B = \begin{Bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{Bmatrix} \quad C = A \pm B = \begin{Bmatrix} a_{11} \pm b_{11} & a_{12} \pm b_{12} & \dots & a_{1n} \pm b_{1n} \\ a_{21} \pm b_{21} & a_{22} \pm b_{22} & \dots & a_{2n} \pm b_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} \pm b_{m1} & a_{m2} \pm b_{m2} & \dots & a_{mn} \pm b_{mn} \end{Bmatrix}$$

矩阵的加减法

- 记 $A_{m\times n}=(a_{ij})$ $B_{m\times n}=(b_{ij})$

加法： $A_{m\times n}+B_{m\times n}=(a_{ij}+b_{ij})$

$$\begin{bmatrix}1 & 2\\5 & 6\end{bmatrix}+\begin{bmatrix}3 & 4\\7 & 8\end{bmatrix}=\begin{bmatrix}4 & 6\\12 & 14\end{bmatrix}$$

减法：

$A_{m\times n}-B_{m\times n}=(a_{ij}-b_{ij})$

$$\begin{bmatrix}6 & 8\\11 & 14\end{bmatrix}-\begin{bmatrix}3 & 4\\7 & 8\end{bmatrix}=\begin{bmatrix}3 & 4\\4 & 6\end{bmatrix}$$

运算律：

交换律 $A+B=B+A$

结合律 $(A+B)+C=A+(B+C)$

矩阵与数的乘法

- 数乘：将数λ与矩阵A相乘，就是将数λ与矩阵A中的每一个元素相乘，记作λA；结果C=λA，并且C中的元素满足： $c_{ij} = \lambda a_{ij}$

$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix} \qquad \lambda A = \begin{Bmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \dots & \dots & \dots & \dots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{Bmatrix}$$

矩阵与数的乘法

- 数乘： $\mathrm{kA} = (ka_{ij})$

$$3 \times \begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 3 \times 1 & 3 \times 2 \\ 3 \times 5 & 3 \times 6 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ 15 & 18 \end{bmatrix}$$

- 数乘运算律

结合律 $(\lambda\mu)A = \lambda(\mu A)$

分配律

$$(\lambda + \mu)A = \lambda A + \mu A$$

$$\lambda(A + B) = \lambda A + \lambda B$$

例1

$$A=\begin{bmatrix}1 & 0 & 2\\ 2 & -1 & 3\end{bmatrix},\ B=\begin{bmatrix}1 & 3 & 4\\ 1 & 0 & 5\end{bmatrix},$$

求 $-3A+2B$

$$-3A+2B=\begin{bmatrix}-3 & 0 & -6\\ -6 & 3 & -9\end{bmatrix}+\begin{bmatrix}2 & 6 & 8\\ 2 & 0 & 10\end{bmatrix}$$

$$=\begin{bmatrix}-1 & 6 & 2\\ -4 & 3 & 1\end{bmatrix}$$

矩阵与向量的乘法

- 假设A为m*n阶矩阵，x为n*1的列向量，则Ax为m*1的列向量，记作：$\vec{y} = A\vec{x}$

$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix} \qquad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \qquad \vec{y} = A\vec{x} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_m \end{pmatrix}$$

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$

矩阵与矩阵的乘法

- 矩阵的乘法仅当第一个矩阵A的列数和第二个矩阵B的行数相等时才能够定义，假设A为m*s阶矩阵，B为s*n阶矩阵，那么C=A*B是m*n阶矩阵，并且矩阵C中的元素满足: $c_{ij}=\sum_{k=1}^{s} a_{ik}b_{kj}$

$$A=\left\{\begin{matrix} a_{11} & a_{12} & \dots & a_{1s} \\ a_{21} & a_{22} & \dots & a_{2s} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{ms} \end{matrix}\right\} \qquad B=\left\{\begin{matrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \dots & \dots & \dots & \dots \\ b_{s1} & b_{s2} & \dots & b_{sn} \end{matrix}\right\} \qquad C=A*B=\left\{\begin{matrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{m1} & c_{m2} & \dots & c_{mn} \end{matrix}\right\}$$

矩阵与矩阵的乘法

- 例：

$$A=\begin{bmatrix}1&0&-1&2\\-1&1&3&0\\0&5&-1&4\end{bmatrix}\qquad B=\begin{bmatrix}0&3&4\\1&2&1\\3&1&-1\\-1&2&1\end{bmatrix}$$ 求C=AB

$$C=AB=\begin{bmatrix}1&0&-1&2\\-1&1&3&0\\0&5&-1&4\end{bmatrix}\begin{bmatrix}0&3&4\\1&2&1\\3&1&-1\\-1&2&1\end{bmatrix}=\begin{bmatrix}-5&6&7\\10&2&-6\\-2&17&10\end{bmatrix}$$

$$(AB)C=A(BC)\quad (A+B)C=AC+BC\quad C(A+B)=CA+CB$$

例3

$$A=\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}, B=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}, C=\begin{bmatrix}-1 & 2\\ 3 & 0\end{bmatrix}$$

$$AB=\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}=\begin{bmatrix}2 & 2\\ -2 & -2\end{bmatrix}$$

$$BA=\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

$$AC=\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}\begin{bmatrix}-1 & 2\\ 3 & 0\end{bmatrix}=\begin{bmatrix}2 & 2\\ -2 & -2\end{bmatrix}$$

$$A^2 \overset{\text{def}}{=\!=} AA=\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}\begin{bmatrix}1 & 1\\ -1 & -1\end{bmatrix}=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix}$$

注意

（1）只有当两个矩阵是同型矩阵时，才能进行加法运算.

（2）只有当第一个矩阵的列数等于第二个矩阵的行数时，两个矩阵才能相乘,且矩阵相乘不满足交换律.

（3）矩阵的数乘运算与行列式的数乘运算不同.

（4）矩阵乘法中存在化零因子，而实数乘法中不存在化零因子。

（5）在实数运算系统中，方程 $ax = 0(a \neq 0)$ 有唯一解，等价的有消去律，矩阵乘法中没有，消去律不成立.

矩阵的转置

- 矩阵的转置：把矩阵A的行和列互相交换所产生的矩阵称为A的转置矩阵，这一过程叫做矩阵的转置。 使用A^T表示A的转置

$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix} \qquad A^T = \begin{Bmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{Bmatrix}$$

矩阵的转置

- 例： $A = \begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}$ ，A的转置为 $A^T = \begin{bmatrix} 1 & 5 \\ 2 & 6 \end{bmatrix}$
- 转置的运算性质：

$$\left(A^T\right)^T = A$$

$$(\lambda A)^T = \lambda A^T$$

$$(AB)^T = B^T A^T$$

$$(A + B)^T = A^T + B^T$$

例4 $A=\begin{pmatrix}2&0&-1\\1&3&2\end{pmatrix}$, $B=\begin{pmatrix}1&7&-1\\4&2&3\\2&0&1\end{pmatrix}$, 求 $(AB)^T$.

解法一

$$AB=\begin{pmatrix}2&0&-1\\1&3&2\end{pmatrix}\begin{pmatrix}1&7&-1\\4&2&3\\2&0&1\end{pmatrix}=\begin{pmatrix}0&14&-3\\17&13&10\end{pmatrix},$$

$$(AB)^T=\begin{pmatrix}0&17\\14&13\\-3&10\end{pmatrix}.$$

解法二

$$(AB)^T=B^TA^T=\begin{pmatrix}1&4&2\\7&2&0\\-1&3&1\end{pmatrix}\begin{pmatrix}2&1\\0&3\\-1&2\end{pmatrix}=\begin{pmatrix}0&17\\14&13\\-3&10\end{pmatrix}.$$

方阵行列式

- 行列式是数学的一个函数，可以看作在几何空间中，一个线性变换对“面积”或“体积”的影响。
- 方阵行列式，n阶方阵A的行列式表示为|A|或者det(A)
  - 1X1的方阵，其行列式等于该元素本身。 $A=\left(a_{11}\right) \quad |A|=a_{11}$
  - 2X2的方阵，其行列式用主对角线元素乘积减去次对角线元素的乘积。

$$A=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \qquad |A|=a_{11}*a_{22}-a_{12}*a_{21}$$

方阵行列式

- 方阵行列式，n阶方阵A的方阵行列式表示为|A|或者det(A)；n阶方阵A的行列式计算规则为：主对角线元素乘积和减去次对角线元素乘积和。设r_i为第i个主对角线的积，l_i为第i个次对角线的积。0≤i≤n

$$A=\begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix}$$

$$r_i=\prod_{k=1}^{i} a_{k(n+k-i)} * \prod_{k=i+1}^{n} a_{k(k-i)}$$

了解

$$l_i=\prod_{k=1}^{i} a_{k(i-k+1)} * \prod_{k=i+1}^{n} a_{k(n-k+i+1)}$$

$$|A|=\sum_{i=1}^{n} r_i-\sum_{i=1}^{n} l_i$$

方阵行列式

- 根据方阵行列式的计算规则可以得到三阶方阵A的行列式为：

$$A = \left\{ \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right\}$$

$$r_i = \prod_{k=1}^{i} a_{k(3-i+k)} * \prod_{k=i+1}^{3} a_{k(k-i)}$$

$$l_i = \prod_{k=1}^{i} a_{k(i-k+1)} * \prod_{k=i+1}^{3} a_{k(4-k+i)}$$

了解

$$|A| = \sum_{i=1}^{3} r_i - \sum_{i=1}^{3} l_i$$

$$= a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} + a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

代数余子式

- 在一个n阶的行列式A中，把元素a_{ij}(i,j=1,2,3,....,n)所在的行和列划去后，剩下的$(n-1)^2$个元素按照原来的次序组成的一个n-1阶行列式M_{ij}，称为元素a_{ij}的余子式。M_{ij}带上符号$(-1)^{(i+j)}$称为a_{ij}的代数余子式，记作：$A_{ij}=(-1)^{(i+j)}M_{ij}$

$$\forall 1 \le j \le n, |A| = \sum_{i=1}^{n} a_{ij} \cdot (-1)^{i+j} M_{ij}$$

$$\forall 1 \le i \le n, |A| = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} M_{ij}$$

$$A = \begin{Bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{Bmatrix}$$

$$|A| = a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} + a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

代数余子式

- 行列式计算例子

$$A = \left\{ \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{matrix} \right\}$$

$$|A| = a_{13}a_{21}a_{32} + a_{12}a_{23}a_{31} + a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$$

例1 计算

$$\begin{vmatrix} 3 & 1 & 4 \\ -1 & 2 & -5 \\ 1 & 3 & 0 \end{vmatrix} \xlongequal{\text{按第3行展开}} 1\times\begin{vmatrix} 1 & 4 \\ 2 & -5 \end{vmatrix} + 3\times\left(-\begin{vmatrix} 3 & 4 \\ -1 & -5 \end{vmatrix}\right)$$

$$= -13 + 3\times 11 = 20$$

伴随矩阵

- 对于n阶方阵的任意元素a_{ij}都有各自的代数余子式$A_{ij}=(-1)^{(i+j)}M_{ij}$，将所有的代数余子式按照次序进行排列，可以得到一个n阶的方阵A^*；那么A^*称为矩阵A的伴随矩阵。
  - 注意：A_{ij}位于A^*的第j行第i列

$$A \cdot A^* = |A| \cdot E$$

$$A^* = \begin{Bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \cdots & \cdots & \cdots & \cdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{Bmatrix}$$

方阵的逆

- 设A是数域上的一个n阶方阵，若在相同的数域上存在另一个n阶方阵B，使得AB=BA=E，那么称B为A的逆矩阵，而A被称为可逆矩阵或非奇异矩阵。如果A不存在逆矩阵，那么A称为奇异矩阵。A的逆矩阵记作：A^{-1}
- 具有以下性质：
  - 如果矩阵A是可逆的，那么矩阵A的逆矩阵是唯一的。
  - A的逆矩阵的逆矩阵还是A，记作$(A^{-1})^{-1}=A$
  - 可逆矩阵A的转置矩阵A^T也可逆，并且$(A^T)^{-1}=(A^{-1})^T$
  - 若矩阵A可逆，则矩阵A满足消去律，即AB=AC => B=C
  - 矩阵A可逆的充要条件是行列式|A|不等于0

运算规律

$$A\cdot A^{*}=|A|\cdot E\Rightarrow A\cdot\frac{A^{*}}{|A|}=\cdot E\Rightarrow A^{-1}=\frac{A^{*}}{|A|}$$

(1) A 可逆 $\Rightarrow A^{-1}$ 可逆，且 $(A^{-1})^{-1}=A$

(2) A 可逆，$k\neq 0\Rightarrow kA$ 可逆，且 $(kA)^{-1}=\frac{1}{k}A^{-1}$

(3) A，B 同阶可逆 $\Rightarrow AB$ 可逆，且 $(AB)^{-1}=B^{-1}A^{-1}$

(4) A 可逆 $\Rightarrow A^{T}$ 可逆，且 $(A^{T})^{-1}=(A^{-1})^{T}$

(5) A 可逆 $\Rightarrow\left|A^{-1}\right|=\frac{1}{|A|}$

方阵的逆

$$\forall 1 \le i \le n, |A| = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} M_{ij}$$

$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Bmatrix} \qquad A^* = \begin{Bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \dots & \dots & \dots & \dots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{Bmatrix}$$

$$A \cdot A^* = \begin{Bmatrix} |A| & 0 & \dots & 0 \\ 0 & |A| & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & |A| \end{Bmatrix} = |A| E \Rightarrow A^{-1} = \frac{1}{|A|} A^*$$

例2

求方阵 $A=\begin{pmatrix}1&2&3\\2&2&1\\3&4&3\end{pmatrix}$ 的逆矩阵.

解 $\because\ |A|=\begin{vmatrix}1&2&3\\2&2&1\\3&4&3\end{vmatrix}=2\neq 0,\quad \therefore\ A^{-1}$存在.

$$A_{11}=\begin{vmatrix}2&1\\4&3\end{vmatrix}=2,\qquad A_{12}=-\begin{vmatrix}2&1\\3&3\end{vmatrix}=-3,$$

同理可得 $A_{13}=2,\ A_{21}=6,\ A_{22}=-6,\ A_{23}=2,$

$A_{31}=-4,\ A_{32}=5,\ A_{33}=-2,$

得 $$A^*=\begin{pmatrix} 2 & 6 & -4 \\ -3 & -6 & 5 \\ 2 & 2 & -2 \end{pmatrix},$$

故

$$A^{-1}=\frac{1}{|A|}A^*=\frac{1}{2}\begin{pmatrix} 2 & 6 & -4 \\ -3 & -6 & 5 \\ 2 & 2 & -2 \end{pmatrix}=\begin{pmatrix} 1 & 3 & -2 \\ -3/2 & -3 & 5/2 \\ 1 & 1 & -1 \end{pmatrix}.$$

矩阵的初等变换

引例

一、消元法解线性方程组

分析：用消元法解下列方程组的过程.

求解线性方程组

$$\begin{cases} 2x_1 - x_2 - x_3 + x_4 = 2, & ① \\ x_1 + x_2 - 2x_3 + x_4 = 4, & ② \\ 4x_1 - 6x_2 + 2x_3 - 2x_4 = 4, & ③ \div 2 \\ 3x_1 + 6x_2 - 9x_3 + 7x_4 = 9, & ④ \end{cases} \qquad (1)$$

解

$$(1)\ \xrightarrow[③\div 2]{①\leftrightarrow ②}\ \begin{cases} \cancel{x_1 + x_2 - 2x_3 + x_4 = 4,} & ① \\ 2x_1 - x_2 - x_3 + x_4 = 2, & ② \\ 2x_1 - 3x_2 + x_3 - x_4 = 2, & ③ \\ 3x_1 + 6x_2 - 9x_3 + 7x_4 = 9, & ④ \end{cases} \quad (B_1)$$

$$\xrightarrow[\substack{③-2①\\ ④-3①}]{②-③}\ \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4, & ① \\ \cancel{2x_2 - 2x_3 + 2x_4 = 0,} & ② \\ -5x_2 + 5x_3 - 3x_4 = -6, & ③ \\ 3x_2 - 3x_3 + 4x_4 = -3, & ④ \end{cases} \quad (B_2)$$

$$\xrightarrow[\substack{③+5② \\ ④-3②}]{② \times \frac{1}{2}} \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4, & ① \\ x_2 - x_3 + x_4 = 0, & ② \\ 2x_4 = -6, & ③ \\ x_4 = -3, & ④ \end{cases} \quad (B_3)$$

$$\xrightarrow[④-2③]{③ \leftrightarrow ④} \begin{cases} x_1 + x_2 - 2x_3 + x_4 = 4, & ① \\ x_2 - x_3 + x_4 = 0, & ② \\ x_4 = -3, & ③ \\ 0 = 0, & ④ \end{cases} \quad (B_4)$$

用“回代”的方法求出解：

于是解得 $\begin{cases} x_1 = x_3 + 4 \\ x_2 = x_3 + 3 \\ x_4 = -3 \end{cases}$ 其中x_3为任意取值.

或令$x_3 = c$,方程组的解可记作

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} c+4 \\ c+3 \\ c \\ -3 \end{pmatrix}, \quad \text{即} x = c\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ 0 \\ -3 \end{pmatrix} \qquad (2)$$

其中c为任意常数.

小结：

1．上述解方程组的方法称为消元法.

2．始终把方程组看作一个整体变形，用到如下三种变换

（1）交换两个方程次序；

（ⓘ 与ⓙ 相互替换）

（2）以不等于 0 的数乘某个方程；

（以 ⓘ$\times k$ 替换 ⓘ）

（3）一个方程的k倍加到另一个方程上.

（以 ⓘ$+k$ⓙ 替换 ⓘ）

3．上述三种变换都是可逆的．

$$若(A)\xrightarrow{ⓘ\leftrightarrow ⓙ}(B),\ 则(B)\xrightarrow{ⓘ\leftrightarrow ⓙ}(A);$$

$$若(A)\xrightarrow{ⓘ\times k}(B),\ 则(B)\xrightarrow{ⓘ\div k}(A);$$

$$若(A)\xrightarrow{ⓘ+kⓙ}(B),\ 则(B)\xrightarrow{ⓘ-kⓙ}(A).$$

由于三种变换都是可逆的，所以变换前的方程组与变换后的方程组是同解的．故这三种变换是同解变换．

因为在上述变换过程中，仅仅只对方程组的系数和常数进行运算，未知量并未参与运算.

若记

$$B=(A,b)=\begin{pmatrix} 2 & -1 & -1 & 1 & 2 \\ 1 & 1 & -2 & 1 & -4 \\ 4 & -6 & 2 & -2 & 4 \\ 3 & 6 & -9 & 7 & 9 \end{pmatrix}$$

则对方程组的变换完全可以转换为对矩阵***B***(方程组（1）的增广矩阵）的变换.

二、矩阵的初等变换

定义 下面三种变换称为矩阵的初等行变换:

(1) 对调两行（对调 i,j 两行,记作 $r_i \leftrightarrow r_j$）；

(2) 以数 $k \neq 0$ 乘以某一行的所有元素;

（第 i 行乘 k,记作 $r_i \times k$）

(3) 把某一行所有元素的 k 倍加到另一行对应的元素上去（第 j 行的 k 倍加到第 i 行上记作 $r_i + kr_j$）.

同理可定义矩阵的初等列变换(所用记号是把“r”换成“c”).

定义 矩阵的**初等列变换**与**初等行变换**统称为**初等变换.**

初等变换的逆变换仍为初等变换, 且变换类型相同.

$r_i \leftrightarrow r_j$ 逆变换 $r_i \leftrightarrow r_j$;

$r_i \times k$ 逆变换 $r_i \times (\frac{1}{k})$ 或 $r_i \div k$;

$r_i + kr_j$ 逆变换 $r_i + (-k)r_j$ 或 $r_i - kr_j$.

如果矩阵 A 经有限次初等变换变成矩阵 B，就称矩阵 A 与 B 等价，记作 $A \sim B$.

等价关系的性质：

（1）自反性：$A \sim A$

（2）对称性：若$A \sim B$，则$B \sim A$

（3）传递性：若$A \sim B$,且$B \sim C$,则$A \sim C$

具有上述三条性质的关系称为等价.

例如，两个线性方程组同解，

就称这两个线性方程组等价

用矩阵的初等行变换 解方程组（1）：

$$B=\begin{pmatrix} 2 & -1 & -1 & 1 & 2 \\ 1 & 1 & -2 & 1 & -4 \\ 4 & -6 & 2 & -2 & 4 \\ 3 & 6 & -9 & 7 & 9 \end{pmatrix}$$

$$\xrightarrow[r_3 \div 2]{r_1 \leftrightarrow r_2} \begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 2 & -1 & -1 & 1 & 2 \\ 2 & -3 & 1 & -1 & 2 \\ 3 & 6 & -9 & 7 & 9 \end{pmatrix} = B_1$$

$$\underset{r_4-3r_1}{\overset{\substack{r_2-r_3\\ r_3-2r_1}}{\sim}}\begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 0 & 2 & -2 & 2 & 0 \\ 0 & -5 & 5 & -3 & -6 \\ 0 & 3 & -3 & 4 & -3 \end{pmatrix} = B_2$$

$$\underset{r_4-3r_2}{\overset{\substack{r_2 \div 2\\ r_3+5r_2}}{\sim}}\begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 1 & -3 \end{pmatrix} = B_3$$

$$\underset{r_4-2r_3}{\overset{r_3 \leftrightarrow r_4}{\sim}} \begin{pmatrix} 1 & 1 & -2 & 1 & 4 \\ 0 & 1 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B_4$$

$$\underset{r_2-r_3}{\overset{r_1-r_2}{\sim}} \begin{pmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B_5$$

B_5 对应的方程组为 $\begin{cases} x_1 = x_3 + 4 \\ x_2 = x_3 + 3 \\ x_4 = -3 \end{cases}$

或令$x_3 = c$,方程组的解可记作

$$x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} c+4 \\ c+3 \\ c \\ -3 \end{pmatrix} = c\begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 4 \\ 3 \\ 0 \\ -3 \end{pmatrix}$$

其中c为任意常数.

矩阵 B_4 和 B_5 都称为行阶梯形矩阵.

特点:

（1）、可划出一条阶梯线，线的下方全为零；

$$\begin{pmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = B_5$$

（2）、每个台阶只有一行；

台阶数即是非零行的行数，阶梯线的竖线后面的第一个元素为非零元，即非零行的第一个非零元.

行阶梯形矩阵B_5还称为行最简形矩阵，即非零行的第一个非零元为1，且这些非零元所在的列的其他元素都为零.

对于任何矩阵$A_{m\times n}$，总可经过有限次初等行变换把它变为行阶梯形和行最简形.

注意：行最简形矩阵是由方程组唯一确定的，行阶梯形矩阵的行数也是由方程组唯一确定的.

行最简形矩阵再经过初等列变换，可化成标准形.

例如

$$B_5 = \begin{pmatrix} 1 & 0 & -1 & 0 & 4 \\ 0 & 1 & -1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\underset{c_5 - 4c_1 - 3c_2 + 3c_3}{\overset{\substack{c_3 \leftrightarrow c_4 \\ c_4 + c_1 + c_2}}{\sim}} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = F$$

矩阵 F 称为矩阵 B 的标准形.

特点：F的左上角是一个单位矩阵，其余元素全为零.

$m\times n$ 矩阵 A 总可经过初等变换化为 标准形

$$F=\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}_{m\times n}$$

此标准形由 m,n,r 三个数唯一确定，其中 r 就是行阶梯形矩阵中非零行 的行数.

所有与矩阵 A 等价的矩阵组成的一个集合，称为一个等价类，标准形 F是这个等价类中最简单的矩阵.

定理 (等价标准形定理)

用初等变换必能将任何一个矩阵化为如下等价标准形：

$$\begin{bmatrix} E_r & O \\ O & O \end{bmatrix}$$ 等价标准形是唯一的。

定理 设A与B为 $m \times n$ 矩阵，那么：

(1) $A \overset{r}{\sim} B$ 的充分必要条件是存在 m阶可逆矩阵P，使

$$PA = B;$$

(2) $A \overset{c}{\sim} B$ 的充分必要条件是存在 n阶可逆矩阵Q，使

$$AQ = B;$$

(3) $A \sim B$ 的充分必要条件是存在 m阶可逆矩阵P及 n阶可逆矩阵Q，使 $PAQ = B;$

推论 方阵 A 可逆的充分必要件是 $A \overset{r}{\sim} E$

定理（1）表明 $A \overset{r}{\sim} B$ 即A 经一系列初等行变换变为B,则 有可逆矩阵P,使 $PA = B$; 如何求P？

由于 $PA = B \Leftrightarrow \begin{cases} PA = B \\ PE = P \end{cases}$,

$$\Leftrightarrow P(A, E) = (B, P) \Leftrightarrow (A, E) \overset{r}{\sim} (B, P)$$

例1 设 $A = \begin{pmatrix} 2 & -1 & -1 \\ 1 & 1 & -2 \\ 4 & -6 & 2 \end{pmatrix}$ 的行最简形矩阵为F

求可逆矩阵P使 $PA = F$.

解：把A用初等行变换化为行最简形F,同时求P,对 (A, E) 作初等行变换A化成行最简形F,同时E就化为成了P;

$$(A,E)=\begin{pmatrix} 2 & -1 & -1 & 1 & 0 & 0 \\ 1 & 1 & -2 & 0 & 1 & 0 \\ 4 & -6 & 2 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[r_2-2r_1]{\substack{r_1\leftrightarrow r_2 \\ r_3-2r_2}} \begin{pmatrix} 1 & 1 & -2 & 0 & 1 & 0 \\ 0 & -3 & 3 & 1 & -2 & 0 \\ 0 & -4 & 4 & -2 & 0 & 1 \end{pmatrix}$$

$$\xrightarrow[r_3+4r_2]{\substack{r_2-r_3 \\ r_1-r_2}} \begin{pmatrix} 1 & 0 & -1 & -3 & 3 & 1 \\ 0 & 1 & -1 & 3 & -2 & -1 \\ 0 & 0 & 0 & 10 & -8 & -3 \end{pmatrix}$$

$$F=\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}, P=\begin{pmatrix} -3 & 3 & 1 \\ 3 & -2 & -1 \\ 10 & -8 & -3 \end{pmatrix}, \text{使} \quad PA=F.$$

例2 设 $A=\begin{pmatrix} 0 & -2 & 1 \\ 3 & 0 & -2 \\ -2 & 3 & 0 \end{pmatrix}$，证明 A 可逆，求 A^{-1}

解：对 (A,E) 作初等行变换化成 (F,P)，F 为 A 的行最简形，如果 $F=E$，则 A 可逆，且 $P=A^{-1}$

$$(A,E)=\begin{pmatrix} 0 & -2 & 1 & 1 & 0 & 0 \\ 3 & 0 & -2 & 0 & 1 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix} \underset{r_2\leftrightarrow r_1}{\overset{r_2+r_3}{\sim}} \begin{pmatrix} 1 & 3 & -2 & 0 & 1 & 1 \\ 0 & -2 & 1 & 1 & 0 & 0 \\ -2 & 3 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\overset{r_3+2r_1}{\sim} \begin{pmatrix} 1 & 3 & -2 & 0 & 1 & 1 \\ 0 & -2 & 1 & 1 & 0 & 0 \\ 0 & 9 & -4 & 0 & 2 & 3 \end{pmatrix} \underset{r_2\leftrightarrow r_3}{\overset{r_3+4r_2}{\sim}} \begin{pmatrix} 1 & 3 & -2 & 0 & 1 & 1 \\ 0 & 1 & 0 & 4 & 2 & 3 \\ 0 & -2 & 1 & 1 & 0 & 0 \end{pmatrix}$$

$$\overset{r_3+2r_2}{\sim}\begin{pmatrix} 1 & 3 & -2 & 0 & 1 & 1 \\ 0 & 1 & 0 & 4 & 2 & 3 \\ 0 & 0 & 1 & 9 & 4 & 6 \end{pmatrix} \overset{r_1+2r_3}{\sim}\begin{pmatrix} 1 & 3 & 0 & 18 & 9 & 13 \\ 0 & 1 & 0 & 4 & 2 & 3 \\ 0 & 0 & 1 & 9 & 4 & 6 \end{pmatrix}$$

$$\overset{r_1-3r_2}{\sim}\begin{pmatrix} 1 & 0 & 0 & 6 & 3 & 4 \\ 0 & 1 & 0 & 4 & 2 & 3 \\ 0 & 0 & 1 & 9 & 4 & 6 \end{pmatrix}$$

故A可逆，且 $$A^{-1}=\begin{pmatrix} 6 & 3 & 4 \\ 4 & 2 & 3 \\ 9 & 4 & 6 \end{pmatrix}$$

本例给出了一种利用行初等变换求逆矩阵的方法，对高阶矩阵更实用。

例3 求解矩阵方程 $AX = B$, 其中

$$A=\begin{pmatrix}2&1&-3\\1&2&-2\\-1&3&2\end{pmatrix}, B=\begin{pmatrix}1&-1\\2&0\\-2&5\end{pmatrix}$$

解：设可逆矩阵P使$PA=F$,则 $P(A,B)=(F,PB)$

故对矩阵(A,B)作初等行变换把A变为F,同时B变为PB,如果$F=E$,
则A可逆， $P=A^{-1}, PAX=PB, PAX=A^{-1}AX=X=PB$

$$(A,B)=\begin{pmatrix}2&1&-3&1&-1\\1&2&-2&2&0\\-1&3&2&-2&5\end{pmatrix}\sim\begin{pmatrix}1&2&-2&2&0\\0&-3&1&-3&-1\\0&5&0&0&5\end{pmatrix}$$

$$\sim\begin{pmatrix}1&2&-2&2&0\\0&1&0&0&1\\0&0&1&-3&2\end{pmatrix}\sim\begin{pmatrix}1&0&0&-4&2\\0&1&0&0&1\\0&0&1&-3&2\end{pmatrix} \qquad X=A^{-1}B=\begin{pmatrix}-4&2\\0&1\\-3&2\end{pmatrix}$$

矩阵的秩

- 在m*n矩阵A中，任取k行k列，不改变这k^2个元素的在A中的次序，得到k阶方阵，称为矩阵A的k阶子式。
  - m*n阶矩阵A的k阶子式有 $C_m^k C_n^k$ 个
- $$A = \begin{bmatrix} 1 & 0 & 2 & 5 \\ 2 & 3 & 4 & 0 \\ 5 & 2 & 2 & 1 \\ 0 & 2 & 1 & 3 \end{bmatrix}$$ 取第1、2、4行与1、3、4列得到3阶子式 $$\begin{vmatrix} 1 & 2 & 5 \\ 2 & 4 & 0 \\ 0 & 1 & 3 \end{vmatrix}$$

矩阵的秩

- 在m*n矩阵A中，任取k行k列，不改变这k^2个元素的在A中的次序，得到k阶方阵，称为矩阵A的k阶子式。
  - m*n阶矩阵A的k阶子式有 $C_m^k C_n^k$
- 设在矩阵A中有一个不等于0的r阶子式D，且所有r+1阶子式(如果存在)全等于0，那么D称为矩阵A的最高阶非零子式，r称为矩阵A的秩，记作R(A)=r
  - n*n的可逆矩阵，秩为n
  - 可逆矩阵又称为满秩矩阵
  - 矩阵的秩等于它行(列)向量组的秩

例2

求矩阵 $B=\begin{pmatrix} 2 & -1 & 0 & 3 & -2 \\ 0 & 3 & 1 & -2 & 5 \\ 0 & 0 & 0 & 4 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ 的秩.

解 $\because$ B是一个行阶梯形矩阵，其非零行有3行，

$\therefore$ B 的所有 4 阶子式全为零.

而 $\begin{vmatrix} 2 & -1 & 3 \\ 0 & 3 & -2 \\ 0 & 0 & 4 \end{vmatrix} \neq 0,$ $\therefore$ $R(B)=3.$

显然，非零行的行数为所求的秩

此方法简单!

阶梯形矩阵的秩就是其非零行数!

对于一般的矩阵，当行数和列数较高时，按定义计算矩阵的秩是很麻烦的，容易遗漏子式。

由上例可知行阶梯形矩阵的秩就等于其非零行的行数，一看便知不须计算.

因为对于任何矩阵$A_{m\times n}$,总可经过有限次初等行变换把他变为行阶梯形.

自然想到用初等行变换把矩阵化为行阶梯形矩阵，

但两个等价矩阵的秩是否相等呢？

问题：经过初等变换矩阵的秩变吗？

二、秩的求法

秩的基本定理 若 $A\sim B$，则：$R(A)=R(B)$

初等变换不改变矩阵的秩。

推论： 设可逆矩阵P，Q 使 $PAQ=B$，则

$R(A)=R(B)$

注： 该定理回答了矩阵标准形 $A\to\begin{bmatrix}E_r & 0\\ 0 & 0\end{bmatrix}$

中 r 是唯一的。它就是矩阵 A 的秩。

于是得到求秩的方法：$A\xrightarrow{\text{行变换}}F$(行阶梯形矩阵)

则：R(A)＝F的台阶数

例3

设 $A=\begin{pmatrix} 3 & 2 & 0 & 5 & 0 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 1 & 6 & -4 & -1 & 4 \end{pmatrix}$，求矩阵 A 的秩，并求 A 的一个最高阶非零子式．

解　对A作初等行变换，变成行阶梯形矩阵：

$$A \underset{\sim}{r_1 \leftrightarrow r_4} \begin{pmatrix} 1 & 6 & -4 & -1 & 4 \\ 3 & -2 & 3 & 6 & -1 \\ 2 & 0 & 1 & 5 & -3 \\ 3 & 2 & 0 & 5 & 0 \end{pmatrix}$$

$$\xrightarrow[\substack{r_3-2r_1\\ r_4-3r_1}]{r_2-r_4}\begin{pmatrix}1 & 6 & -4 & -1 & 4\\ 0 & -4 & 3 & 1 & -1\\ 0 & -12 & 9 & 7 & -11\\ 0 & -16 & 12 & 8 & -12\end{pmatrix}$$

$$\xrightarrow[r_4-4r_2]{r_3-3r_2}\begin{pmatrix}1 & 6 & -4 & -1 & 4\\ 0 & -4 & 3 & 1 & -1\\ 0 & 0 & 0 & 4 & -8\\ 0 & 0 & 0 & 4 & -8\end{pmatrix}$$

$$\underset{\sim}{r_4 - r_3} \begin{pmatrix} 1 & 6 & -4 & -1 & 4 \\ 0 & -4 & 3 & 1 & -1 \\ 0 & 0 & 0 & 4 & -8 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

由阶梯形矩阵有三个非零行可知 $R(A) = 3.$

求 A 的一个最高阶子式.

$\because \quad R(A)=3,$ 知A的最高阶非零子式为3阶.

A 的 3 阶子式共有 $C_4^3 \bullet C_5^3 = 40$ 个.

考察A的行阶梯形矩阵,

记$A=(a_1,a_2,a_3,a_4,a_5)$,则矩阵$B=(a_1,a_2,a_4)$的行阶梯形矩阵为

$$\begin{pmatrix} 1 & 6 & -1 \\ 0 & -4 & 1 \\ 0 & 0 & 4 \\ 0 & 0 & 0 \end{pmatrix} \qquad \because \quad R(B)=3,$$

故 B 中必有 3 阶非零子式．且共有 4 个.

计算B的前三行构成的子式

$$\begin{vmatrix} 3 & 2 & 5 \\ 2 & 0 & 5 \\ 3 & -2 & 6 \end{vmatrix} = \begin{vmatrix} 3 & 2 & 5 \\ 2 & 0 & 5 \\ 6 & 0 & 11 \end{vmatrix}$$

$$= -2\begin{vmatrix} 2 & 5 \\ 6 & 11 \end{vmatrix} = 16 \neq 0.$$

则这个子式便是 A 的一个最高阶非零子式.

向量组

- 向量组：有限个相同维数的行向量或列向量组合成的一个集合就叫做向量组

$$A=\left(\vec{a}_1,\vec{a}_2,\vec{a}_2.....\vec{a}_n....\right)$$

- 向量组是有多个向量构成，可以表示为矩阵

向量组等价

- 向量b能由向量组A:$a_1,a_2,\ldots a_m$线性表示的充要条件为矩阵$A=(a_1,a_2,\ldots a_m)$的秩等于矩阵$B=(a_1,a_2,\ldots a_m,b)$的秩。
- 设有两个向量组A:$a_1,a_2,\ldots a_m$及B:$b_1,b_2,\ldots b_m$，若B组的向量都能由向量组A线性表示，则称向量组B能由向量组A线性表示。若向量组A和向量组B能相互线性表示，则称两个向量组等价。

系数矩阵

- 将向量组A和B所构成的矩阵依次记作$A=(a_1,a_2,\ldots a_m)$和$B=(b_1,b_2,\ldots b_n)$，且组B能由向量组A线性表示。即对每个向量b_j，存在$k_{1j},k_{2j},\ldots,k_{mj}$；使得

$$b_j = k_{1j}a_1 + k_{2j}a_2 + \ldots + k_{mj}a_m = \begin{pmatrix} a_1 & a_2 & \ldots & a_m \end{pmatrix} \begin{pmatrix} k_{1j} \\ k_{2j} \\ \ldots \\ k_{mj} \end{pmatrix}$$

- 从而可以得到系数矩阵K

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_m \end{pmatrix} \begin{bmatrix} k_{11} & k_{12} & \ldots & k_{1n} \\ k_{21} & k_{22} & \ldots & k_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ k_{m1} & k_{m2} & \ldots & k_{mn} \end{bmatrix} = \begin{pmatrix} b_1 & b_2 & \ldots & b_n \end{pmatrix}$$

正交矩阵

- 若n阶方阵A满足$A^TA=E$，则称A为正交矩阵，简称正交阵。
  - A是正交阵的充要条件：A的列(行)向量都是单位向量，且两两正交。
- 若A为正交矩阵，x为向量，则Ax称为正交变换。
  - 正交变换不改变向量的长度。
- 正交矩阵的性质：
  - 若A为正交矩阵，则逆矩阵A^{-1}也为正交矩阵
  - 若P、Q为正交矩阵，那么P*Q也为正交矩阵

对称矩阵

- 元素以对角线为对称轴对应相等的矩阵就叫做对称矩阵
- 对称矩阵具有的特性：
  - 对称矩阵中a_{ij}等于a_{ji}
  - 对称矩阵一定是方阵；并且对于任何的方阵A，$A+A^T$是对称矩阵
  - 除对角线外的其它元素均为0的矩阵叫做对角矩阵
  - 矩阵中的每个元素都是实数的对称矩阵叫做实对称矩阵

$$A=\begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{Bmatrix}$$

$$a_{12}=a_{21}$$
$$a_{ij}=a_{ji}$$
$$a_{1n}=a_{n1}$$

线性方程组

设有 n 个未知数 m 个方程的线性方程组

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases} \quad (1)$$

（1）式可以写成以向量 x 为未知元的向量方程

$Ax = b$，（2）

以后线性方程组（1）与向量方程（2）将混同使用而

不加区分，解与解向量的名称亦不加区别.

线性方程组（1）如果有解，就称它是**相容的**，如果无解，就称它**不相容**.

定理 3　n 元线性方程组 $Ax = b$

(i) 无解的充要条件是 $R(A) < R(A,\ b)$；

(ii) 有唯一解的充要条件是

$$R(A) = R(A,\ b) = n\ ;$$

(iii) 有无穷多解的充要条件是

$$R(A) = R(A,\ b) < n\ .$$

求解线性方程组的步骤：

步骤 1 对于非齐次线性方程组，把它的增广矩阵 B 化成行阶梯形，从中可同时看出 $R(A)$ 和 $R(B)$. 若 $R(A) < R(B)$ ，则方程组无解.

步骤 2 若 $R(A) = R(B)$ ，则进一步把 B 化成行最简形. 而对于齐次线性方程组，则把系数矩阵 A 化成行最简形.

步骤 3 设 $R(A) = R(B) = r$，把行最简形中 r 个非零行的非零首元所对应的未知量取作非自由未知量，其余 $n - r$ 个未知量取作自由未知量，并令自由未知量分别等于 $c_1, c_2, \cdots, c_{n-r}$，由 B（或 A）的行最简形，即可写出含 $n - r$ 个参数的通解.

例1 求解齐次线性方程组

$$\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 = 0 \\ 2x_1 + x_2 - 2x_3 - 2x_4 = 0 \\ x_1 - x_2 - 4x_3 - 3x_4 = 0 \end{cases}$$

解 对系数矩阵A施行初等行变换化为最简阶梯形:

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 \\ 2 & 1 & -2 & -2 \\ 1 & -1 & -4 & -3 \end{pmatrix} \xrightarrow[r_3 - r_1]{r_2 - 2r_1} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & -3 & -6 & -4 \end{pmatrix}$$

$$\xrightarrow{r_3 - r_2} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & -3 & -6 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{-\frac{1}{3} r_2} \begin{pmatrix} 1 & 2 & 2 & 1 \\ 0 & 1 & 2 & 4/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow{r_1-2r_2}\begin{pmatrix}1 & 0 & -2 & -5/3\\ 0 & 1 & 2 & 4/3\\ 0 & 0 & 0 & 0\end{pmatrix}$$ 等价式: $$\begin{cases}x_1 = 2x_3+\frac{5}{3}x_4\\ x_2=-2x_3-\frac{4}{3}x_4\end{cases}$$

令 $x_3 = c_1, x_4 = c_2$

写出参数形式的通解,再改写为向量形式:

通解 $$\begin{cases}x_1 = 2c_1+\frac{5}{3}c_2\\ x_2=-2c_1-\frac{4}{3}c_2\\ x_3=c_1\\ x_4=c_2\end{cases}$$ 即 $$\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=c_1\begin{pmatrix}2\\-2\\1\\0\end{pmatrix}+c_2\begin{pmatrix}5/3\\-4/3\\0\\1\end{pmatrix}.$$

其中 c_1, c_2 为任意实数。

例2　求解非齐次线性方程组

$$\begin{cases} x_1 - 2x_2 + 3x_3 - x_4 = 1 \\ 3x_1 - x_2 + 5x_3 - 3x_4 = 2 \\ 2x_1 + x_2 + 2x_3 - 2x_4 = 3 \end{cases}$$

解 对增广矩阵只用行变换化阶梯形

$$B = \left[\begin{array}{cccc|c} 1 & -2 & 3 & -1 & 1 \\ 3 & -1 & 5 & -3 & 2 \\ 2 & 1 & 2 & -2 & 3 \end{array}\right] \xrightarrow{r} \left[\begin{array}{cccc|c} 1 & -2 & 3 & -1 & 1 \\ 0 & 5 & -4 & 0 & -1 \\ 0 & 0 & 0 & 0 & 2 \end{array}\right]$$

最后一行对应的方程是：$0 = 2$，所以无解。

例3 解方程组 $\begin{cases} x_1 + 2x_2 + x_4 = 3 \\ x_1 + 2x_2 + x_3 - 3x_4 = 8 \\ 2x_1 + 4x_2 + 2x_4 = 6 \\ x_1 + 2x_2 - x_3 + 5x_4 = -2 \end{cases}$

第一步:把增广矩阵用行变换化阶梯形,如果 $\mathrm{R}(A) \neq \mathrm{R}(B)$,则无解.

如果 $\mathrm{R}(A) = \mathrm{R}(B)$,则继续化为最简阶梯形。

$$B = \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 3 \\ 1 & 2 & 1 & -3 & 8 \\ 2 & 4 & 0 & 2 & 6 \\ 1 & 2 & -1 & 5 & -2 \end{array}\right] \xrightarrow{r} \left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 3 \\ 0 & 0 & 1 & -4 & 5 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

问:此时 $\mathrm{R}(A) = \mathrm{R}(B) = r = 2$ 其含义是 独立(或有效)方程的个数。

第二步：写出等价的(独立的)方程组，保留第一个未知数在左边其余的移到右边，移到右边的称为自由变量。

$$\left[\begin{array}{cccc|c} 1 & 2 & 0 & 1 & 3 \\ 0 & 0 & 1 & -4 & 5 \end{array}\right] \longrightarrow \begin{cases} x_1 = -2x_2 - x_4 + 3 \\ x_3 = 4x_4 + 5 \end{cases}$$

问：自由变量的个数 $= n - r = 4 - 2$

即未知数的个数减去独立方程的个数。

问：何时有唯一解？何时有无穷多解？

当出现自由变量时，令自由变量为任意数就可得到无穷多解，当没有自由变量时有唯一解。即当 $\mathrm{R}(A) = \mathrm{R}(B) < n$ 时，有无穷多解，当 $\mathrm{R}(A) = \mathrm{R}(B) = n$ 时有唯一解。

第三步令自由变量为任意实数，写出通解。再改写为向量形式。

$$\begin{cases} x_1 = -2x_2 - x_4 + 3 \\ x_3 = 4x_4 + 5 \end{cases}$$

令 $x_2 = k_1, x_4 = k_2$ 通解 $\begin{cases} x_1 = -2k_1 - k_2 + 3 \\ x_2 = k_1 \\ x_3 = 4k_2 + 5 \\ x_4 = k_2 \end{cases}$

即 $x = k_1 \begin{bmatrix} -2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} -1 \\ 0 \\ 4 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ 0 \\ 5 \\ 0 \end{bmatrix}$ k_1, k_2为任意常数

两个基本定理

由定理 3 容易得出线性方程组理论中两个最基本的定理.

定理 4 n 元齐次线性方程组 $Ax=0$ 有非零解的充要条件是 $R(A)<n$.

由定理 4 可得如下推论:

推论 当 $m<n$ 时，齐次线性方程组 $A_{m\times n}x=0$ 一定有非零解.

定理 5 线性方程组 $Ax = b$ 有解的充要条件是 $R(A) = R(A, b)$.

为了下一章论述的需要，下面把定理 5 推广到矩阵方程.

定理 6 矩阵方程 $AX = B$ 有解的充要条件是 $R(A) = R(A, B)$.

一 齐次线性方程组的解

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_4 = 0 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \end{cases} \quad (1)$$

记 $A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

$$\boldsymbol{Ax} = \mathbf{0} \qquad (2)$$

$$A(\xi_1 + \xi_2) = A\xi_1 + A\xi_2 = \mathbf{0} + \mathbf{0} = \mathbf{0}$$

$$A(k\xi_1) = kA(\xi_1) = k\mathbf{0} = \mathbf{0}$$

齐次线性方程组的解集的最大无关组称为该齐次线性方程组的基础解系. 由上面的讨论可知, 要求齐次线性方程组的通解, 只需求出它的基础解系. 为此, 我们可用初等变换的方法来求齐次线性方程组的基础解系.

齐次方程组解的结构定理

齐次方程组 $A_{m\times n}X=0$ 的基础解系所含向量个数为

$$n-r \quad (r=R(A))$$

设一个基础解系为：$\xi_1,\xi_2,\cdots,\xi_{n-r}$

则通解为：$x=k_1\xi_1+k_2\xi_2+\cdots+k_{n-r}\xi_{n-r}\ (k_i\in R)$

例1 求 $\begin{cases} x_1+x_2-x_3-x_4=0 \\ 2x_1-5x_2+3x_3+2x_4=0 \\ 7x_1-7x_2+3x_3+x_4=0 \end{cases}$ 基础解系和通解.

解 对系数矩阵A作初等行变换, 变为行最简形矩阵,

有 $$A=\begin{pmatrix} 1 & 1 & -1 & -1 \\ 2 & -5 & 3 & 2 \\ 7 & -7 & 3 & 1 \end{pmatrix}$$

$$\sim\begin{pmatrix} 1 & 0 & -\frac{2}{7} & -\frac{3}{7} \\ 0 & 1 & -\frac{5}{7} & -\frac{4}{7} \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\begin{cases} x_1 = \dfrac{2}{7}x_3 + \dfrac{3}{7}x_4 \\ x_2 = \dfrac{5}{7}x_3 + \dfrac{4}{7}x_4 \end{cases} \qquad (*)$$

令 $\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 则 $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{2}{7} \\ \frac{5}{7} \end{pmatrix}, \begin{pmatrix} \frac{3}{7} \\ \frac{4}{7} \end{pmatrix}$

合起来得基础解系：

基础解系为 $\xi_1=\begin{pmatrix}\frac{2}{7}\\ \frac{5}{7}\\ 1\\ 0\end{pmatrix}, \xi_2=\begin{pmatrix}\frac{3}{7}\\ \frac{4}{7}\\ 0\\ 1\end{pmatrix},$

通解为 $\begin{pmatrix}x_1\\ x_2\\ x_3\\ x_4\end{pmatrix}=c_1\begin{pmatrix}\frac{2}{7}\\ \frac{5}{7}\\ 1\\ 0\end{pmatrix}+c_2\begin{pmatrix}\frac{3}{7}\\ \frac{4}{7}\\ 0\\ 1\end{pmatrix}, (c_1,c_2\in R)$

$$\begin{cases} x_1 = \dfrac{2}{7}x_3 + \dfrac{3}{7}x_4 \\ x_2 = \dfrac{5}{7}x_3 + \dfrac{4}{7}x_4 \end{cases} \qquad (*)$$

令 $\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 7 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -7 \end{pmatrix}$ 则 $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \begin{pmatrix} -3 \\ -4 \end{pmatrix}$

合起来得基础解系：$\eta_1 = \begin{pmatrix} 2 \\ 5 \\ 7 \\ 0 \end{pmatrix}, \eta_2 \begin{pmatrix} -3 \\ -4 \\ 0 \\ -7 \end{pmatrix}$, 两基础解系等价，方程组通解的形式不唯一。

注:由最大无关组的性质可知方程组(1)的任何$n-r$个线性无关的解都可构成它的基础解系.从而齐次线性方程组的基础解系不唯一,

通解形式也不唯一.当r=n时，只有零解没有基础解系, s=$\{0\}$.

非齐次方程组解的结构定理

设η^*是 非齐次方程组$A_{m\times n}X=b$ 的一特解

则当非齐次线性方程组有无穷多解时其通解为:

$$x=k_1\xi_1+k_2\xi_2+\cdots+k_{n-r}\xi_{n-r}+\eta^*\ (k_i\in R)$$

其中$k_1\xi_1+\cdots+k_{n-r}\xi_{n-r}$为对应齐次线性方程组的通解

例5 求解方程组 $\begin{cases} x_1 - x_2 - x_3 + x_4 = 0, \\ x_1 - x_2 + x_3 - 3x_4 = 1, \\ x_1 - x_2 - 2x_3 + 3x_4 = -1/2. \end{cases}$

解

$$B = \left(\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 0 \\ 1 & -1 & 1 & -3 & 1 \\ 1 & -1 & -2 & 3 & -1/2 \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & -1 & 0 & -1 & 1/2 \\ 0 & 0 & 1 & -2 & 1/2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right),$$

可见 $R(A) = R(B) = 2 < 4$, 故方程组有无穷多解

$$\begin{cases} x_1 = x_2 + x_4 + 1/2 \\ x_3 = \quad 2x_4 + 1/2 \end{cases}$$

$$\Rightarrow \begin{cases} x_1 = x_2 + x_4 + 1/2 \\ x_2 = x_2 \\ x_3 = \quad 2x_4 + 1/2 \\ x_4 = \quad x_4 \end{cases} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + C_2 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix} \quad (C_1, C_2 \in R).$$

$$\begin{cases} x_1 = x_2 + x_4 + 1/2, \\ x_3 = \quad\ 2x_4 + 1/2. \end{cases}$$

取 $x_2 = x_4 = 0$, 则 $x_1 = x_3 = \dfrac{1}{2}$, 即得方程组的一个解

$$\eta^* = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix}.$$

在对应的齐次线性方程组$\begin{cases} x_1 = x_2 + x_4, \\ x_3 = 2x_4 \end{cases}$中,取

$$\begin{pmatrix} x_2 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{及} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \text{ 则} \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{及} \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

即得对应的齐次线性方程组的基础解系

$$\xi_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \xi_2 = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix},$$

于是所求通解为

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = c_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 0 \\ 2 \\ 1 \end{pmatrix} + \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \\ 0 \end{pmatrix}, (c_1, c_2 \in R).$$

特征值和特征向量

- A为n阶矩阵，若数λ和n维非0列向量x满足Ax=λx，那么数λ称为A的**特征值**，x称为A的对应于特征值λ的**特征向量**。并且|λE-A|叫做A的**特征多项式**。当特征多项式等于0的时候，称为A的特征方程，特征方程是一个齐次线性方程组，求解特征值的过程其实就是求解特征方程的解。

$$
\begin{aligned}
& Ax = \lambda x \\
\Rightarrow\ & Ax = \lambda E x \\
\Rightarrow\ & (\lambda E - A)x = 0
\end{aligned}
$$

$$
|\lambda E - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{vmatrix} = 0
$$

（1）设 α 是矩阵 A 的属于特征值 λ_0 的特征向量，则对任意常数 $k \neq 0$， $k\alpha$ 也是 A 的属于 λ_0 的特征向量；

（2）若 α,β 都是 A 的属于特征值 λ_0 的特征向量，则 $k\alpha+l\beta$ $(k,l$ 不全为零$)$ 也是 A 的属于 λ_0 的特征向量。

特征值和特征向量求解

例1 求$A=\begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}$的特征值和特征向量.

解 A的特征多项式为

$$|A-\lambda E|=\begin{vmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{vmatrix}=(3-\lambda)^2-1$$

$$=8-6\lambda+\lambda^2=(4-\lambda)(2-\lambda)$$

令$|A-\lambda E|=0$得A的特征值为$\lambda_1=2, \lambda_2=4.$

当$\lambda_1=2$时,对应的特征向量应满足

$$\begin{pmatrix} 3-2 & -1 \\ -1 & 3-2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}=\begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

即 $$\begin{cases} x_1 - x_2 = 0, \\ -x_1 + x_2 = 0. \end{cases}$$

解得 $x_1 = x_2$，所以对应的特征向量可取为 $p_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

当 $\lambda_2 = 4$时,由

$$\begin{pmatrix} 3-4 & -1 \\ -1 & 3-4 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \text{即} \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

解得 $x_1 = -x_2$,所以对应的特征向量可取为

$$p_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

得对应于特征值的全部特征向量为$c_1 p_1, c_2 p_2 (c_1 c_2 \neq 0)$

例2 求矩阵$A=\begin{pmatrix}-1&1&0\\-4&3&0\\1&0&2\end{pmatrix}$的特征值和特征向量.

解 A的特征多项式为

$$|A-\lambda E|=\begin{vmatrix}-1-\lambda&1&0\\-4&3-\lambda&0\\1&0&2-\lambda\end{vmatrix}=(2-\lambda)(1-\lambda)^2,$$

令$|A-\lambda E|=0$得A的特征值为$\lambda_1=2,\lambda_2=\lambda_3=1.$

当$\lambda_1=2$时,解方程$(A-2E)x=0.$由

$$A-2E=\begin{pmatrix}-3 & 1 & 0\\ -4 & 1 & 0\\ 1 & 0 & 0\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系 $p_1=\begin{pmatrix}0\\0\\1\end{pmatrix},$

所以$k_1 p_1 (k_1 \neq 0)$是对应于$\lambda_1 = 2$的全部特征值.

当$\lambda_2=\lambda_3=1$时,解方程$(A-E)x=0$.由

$$A-E=\begin{pmatrix}-2 & 1 & 0\\ -4 & 2 & 0\\ 1 & 0 & 1\end{pmatrix}\sim\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & 2\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系 $p_2=\begin{pmatrix}-1\\ -2\\ 1\end{pmatrix},$

所以$k_2p_2(k_2\neq 0)$是对应于$\lambda_2=\lambda_3=1$的全部特征值.

例3

设 $A=\begin{pmatrix}-2 & 1 & 1\\ 0 & 2 & 0\\ -4 & 1 & 3\end{pmatrix}$,求A的特征值与特征向量.

解

$$\left|A-\lambda E\right|=\begin{vmatrix}-2-\lambda & 1 & 1\\ 0 & 2-\lambda & 0\\ -4 & 1 & 3-\lambda\end{vmatrix}$$

$$=-(\lambda+1)(\lambda-2)^2,$$

令 $-(\lambda+1)(\lambda-2)^2=0$

得A的特征值为$\lambda_1=-1,\lambda_2=\lambda_3=2.$

当$\lambda_1=-1$时,解方程$(A+E)x=0$.由

$$A+E=\begin{pmatrix}-1 & 1 & 1\\ 0 & 3 & 0\\ -4 & 1 & 4\end{pmatrix}\sim\begin{pmatrix}1 & 0 & -1\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系 $p_1=\begin{pmatrix}1\\0\\1\end{pmatrix}$,

故对应于$\lambda_1=-1$的全体特征向量为

$$k_1 \boldsymbol{p}_1 \qquad (k_1\neq 0).$$

当$\lambda_2=\lambda_3=2$时，解方程$(A-2E)x=0$.由

$$A-2E=\begin{pmatrix}-4 & 1 & 1\\ 0 & 0 & 0\\ -4 & 1 & 1\end{pmatrix}\sim\begin{pmatrix}-4 & 1 & 1\\ 0 & 0 & 0\\ 0 & 0 & 0\end{pmatrix},$$

得基础解系为:

$$p_2=\begin{pmatrix}0\\ 1\\ -1\end{pmatrix},\qquad p_3=\begin{pmatrix}1\\ 0\\ 4\end{pmatrix},$$

所以对应于$\lambda_2=\lambda_3=2$的全部特征向量为：

$$k_2p_2+k_3p_3\quad (k_2,k_3\text{不同时为}0).$$

特征值和特征向量求解2

$$A=\begin{vmatrix}5 & 0 & 0\\0 & 2 & 0\\0 & 0 & 6\end{vmatrix}$$

$$|\lambda E-A|=\begin{vmatrix}\lambda-5 & 0 & 0\\0 & \lambda-2 & 0\\0 & 0 & \lambda-6\end{vmatrix}=(\lambda-5)(\lambda-2)(\lambda-6)=0$$

$$\Rightarrow \lambda_1=5,\ \ \lambda_2=2,\ \ \lambda_3=6$$

$$\lambda_1=1\Rightarrow (E-A)x=0$$

$$E-A=\begin{bmatrix}0 & 0 & 0\\0 & 3 & 0\\0 & 0 & -1\end{bmatrix} \Rightarrow \begin{cases}x_2=0\\x_3=0\end{cases} \overset{令x_1=1}{\Rightarrow} \xi_1=\begin{bmatrix}1\\0\\0\end{bmatrix} \quad \xi_2=\begin{bmatrix}0\\1\\0\end{bmatrix} \quad \xi_3=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

特征值的性质

- n阶方阵$A=(a_{ij})$的所有特征根λ_1、λ_2.....λ_n，则有

$$\lambda_1+\lambda_2+\ldots+\lambda_n=\sum_{i=1}^{n}a_{ii} \qquad \lambda_1\lambda_2\ldots\lambda_n=|A|$$

- 若λ是可逆矩阵A的一个特征根，x为对应的特征向量：
  - 则1/λ是矩阵A^{-1}的一个特征根，x仍为对应的特征向量。
  - 则λ^m是矩阵A^m的一个特征根，x仍为对应的特征向量。
- 设λ_1、λ_2.....λ_n是方阵A的互不相同的特征值，x_i是λ_i的特征向量，则$x_1,x_2\ldots x_n$线性无关，即不相同特征值的特征向量线性无关。

可对角化矩阵

- 如果一个n阶方阵A相似于对角矩阵，也就是，如果存在一个可逆矩阵P使得$P^{-1}AP$是对角矩阵，则称矩阵A为可对角化矩阵。并且最终对角矩阵的特征值就是矩阵A的特征值。

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

- 矩阵对角化在PCA、白化等机器学习领域应用的比较多。

于是对于矩阵是否可以对角化,判断如下:

1.由$|A-\lambda E|=0 \Rightarrow$ 求出所有的特征值λ_i

2.如果所有的特征值都是单根,则A一定能对角化

3.如果A的特征值有重根,则对每个λ_i,求齐次方程组

$(A-\lambda_i E)X=0$ 的基础解系, 如果基础解系所含向量

的个数等于λ_i的重根数或等于$n-R(A-\lambda_i E)$, 则A可以

对角化且这些 基础解系排成的矩阵为相似变换矩阵.

例1 设 $A=\begin{pmatrix} 4 & 6 & 0 \\ -3 & -5 & 0 \\ -3 & -6 & 1 \end{pmatrix}$

A能否对角化？若能对角化,则求出可逆矩阵P,使$P^{-1}AP$为对角阵.

解

$$|A-\lambda E|=\begin{vmatrix} 4-\lambda & 6 & 0 \\ -3 & -5-\lambda & 0 \\ -3 & -6 & 1-\lambda \end{vmatrix}=-(\lambda-1)^2(\lambda+2)$$

所以A的全部特征值为 $\lambda_1=\lambda_2=1, \lambda_3=-2.$

将$\lambda_1=\lambda_2=1$代入$(A-\lambda E)x=0$得方程组

$$\begin{cases} 3x_1+6x_2=0 \\ -3x_1-6x_2=0 \\ -3x_1-6x_2=0 \end{cases}$$

解之得基础解系

$$\xi_1=\begin{pmatrix}-2\\1\\0\end{pmatrix},\qquad \xi_2=\begin{pmatrix}0\\0\\1\end{pmatrix}.$$

将$\lambda_3=-2$代入$(A-\lambda E)x=0$,得方程组的基础解系

$$\xi_3=(-1,1,1)^T.$$

由于 ξ_1,ξ_2,ξ_3 线性无关. 所以 A 可对角化.

令 $$P=(\xi_1,\xi_2,\xi_3)=\begin{pmatrix}-2 & 0 & -1\\ 1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}$$

则有 $$P^{-1}AP=\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & -2\end{pmatrix}.$$

注意

令 $$P=(\xi_1,\xi_2,\xi_3)=\begin{pmatrix}-1 & 0 & -2\\ 1 & 0 & 1\\ 1 & 1 & 0\end{pmatrix}$$

则有 $$P^{-1}AP=\begin{pmatrix}-2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}.$$

即矩阵P的列向量和对角矩阵中特征值的位置要相互对应.

正定矩阵

- 对于n阶方阵A，若任意n阶向量x，都有$x^TAx>0$，则称矩阵A为正定矩阵。
- 若$x^TAx>=0$，则矩阵A为半正定矩阵。

奇异矩阵

- 若方阵A的行列式的值等于0，那么方阵A叫做奇异矩阵，否则叫做非奇异矩阵。
- 可逆矩阵就是非奇异矩阵，非奇异矩阵也是可逆矩阵。
- 若A为奇异矩阵，则Ax=0有无穷解，Ax=b有无穷解或者无解。
- 若A为非奇异矩阵，则Ax=0有且只有唯一零解，Ax=b有唯一解。

QR分解

- 对于m*n的列满秩矩阵A，必有：$A_{m*n} = Q_{m*m} \cdot R_{m*n}$
- 其中Q为正交矩阵，R为非奇异上三角矩阵，当要求R的对角线元素为正的时候，该分解唯一。
- 该分解叫做QR分解，常用于求解A的特征值、A的逆等问题。

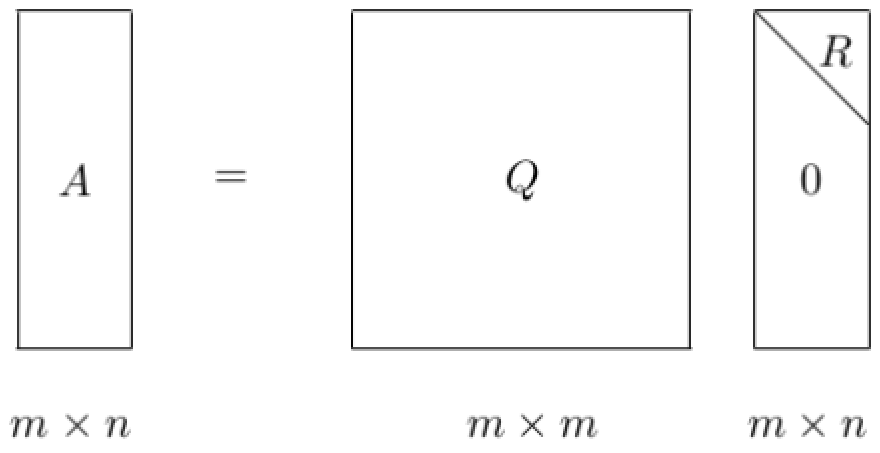
$$A = QR \Rightarrow A_1 = Q^T A Q = RQ$$

$$\cdots\cdots \qquad \Rightarrow A_k \to diag\{\lambda_1, \lambda_2 \ldots\lambda_n\}$$

$$A_k = Q_k R_k \Rightarrow A_{k+1} = R_k Q_k$$

QR分解

- QR分解是将矩阵分解为一个正交矩阵与上三角矩阵的乘积



- 这其中， Q为正交矩阵，$Q^TQ=I$，R为上三角矩阵。
- 实际中，QR分解经常被用来解线性最小二乘问题。

SVD

- 奇异值分解(Singular Value Decomposition)是一种重要的矩阵分解方法，可以看做是对称方阵在任意矩阵上的推广。
- 假设A为一个m*n阶实矩阵，则存在一个分解使得：

$$A_{m*n} = U_{m*m}\Sigma_{m*n}V_{n*n}^{T}$$

  - 通常将奇异值由大到小排列，这样Σ便能由A唯一确定了。
- 与特征值、特征向量的概念相对应，则：
  - Σ对角线上的元素称为矩阵A的奇异值
  - U和V称为A的左/右奇异向量矩阵

向量的导数

- A为m*n的矩阵，x为n*1的列向量，则Ax为m*1的列向量，记作 $\vec{y} = A \cdot \vec{x}$

$$A = \begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix} \qquad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} \qquad \vec{y} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \dots. \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n \end{pmatrix}$$

$$\frac{\partial \vec{y}}{\partial \vec{x}} = \frac{\partial A\vec{x}}{\partial \vec{x}} = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \dots & \dots & \dots & \dots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix} = A^T$$

向量的导数

- 向量偏导公式

$$\frac{\partial A\vec{x}}{\partial \vec{x}} = A^T$$

$$\frac{\partial A\vec{x}}{\partial \vec{x}^T} = A$$

$$\frac{\partial \left(\vec{x}^T A\right)}{\partial \vec{x}} = A$$

标量对向量的导数

- A为n*n的矩阵，x为n*1的列向量，记 $y = \vec{x}^T \cdot A \cdot \vec{x}$
- 同理可得：$\frac{\partial y}{\partial \vec{x}} = \frac{\partial\left(\vec{x}^T \cdot A \cdot \vec{x}\right)}{\partial \vec{x}} = \left(A^T + A\right) \cdot \vec{x}$
- 若A为对称矩阵，则有 $\frac{\partial\left(\vec{x}^T A \vec{x}\right)}{\partial \vec{x}} = 2A\vec{x}$

标量对向量的导数

$$A=\begin{Bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{Bmatrix} \qquad \vec{x}=\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}$$

$$\vec{x}^{T}\cdot A\cdot \vec{x}=(x_1,x_2\dots x_n)\left(\sum_{j=1}^{n}a_{1j}x_j,\sum_{j=1}^{n}a_{2j}x_j,\dots\sum_{j=1}^{n}a_{nj}x_j\right)^{T}=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j$$

$$\frac{\partial\left(\vec{x}^{T}\cdot A\cdot \vec{x}\right)}{\partial \vec{x}}=\left(\sum_{j=1}^{n}a_{ij}x_j\right)+\left(\sum_{j=1}^{n}a_{ji}x_j\right)=\sum_{j=1}^{n}\left(a_{ij}+a_{ji}\right)x_j$$

标量对方阵的导数

- A为n*n的矩阵，|A|为A的行列式，计算 $\frac{\partial |A|}{\partial A}$

$$\forall 1 \le i \le n, |A| = \sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} M_{ij}$$

$$\frac{\partial |A|}{\partial a_{ij}} = \frac{\partial \left(\sum_{j=1}^{n} a_{ij} \cdot (-1)^{i+j} M_{ij} \right)}{\partial a_{ij}} = (-1)^{i+j} M_{ij} = A_{ij}$$

$$\frac{\partial |A|}{\partial A} = \left(A^*\right)^T = |A| \cdot \left(A^{-1}\right)^T$$

参考文献

1.北京大学数学系 王萼芳 石生明，高等代数，高等教育出版社，2013；

2.郝志峰 谢国瑞 方文波 汪国强，线性代数，高等教育出版社，2010；

3.同济大学数学系，线性代数，高等教育出版社，2012。

THANK YOU

上海育创网络科技有限公司