线性代数简介

demerzel

2019 年 3 月 22 日

矩阵是线性代数的主要研究对象。

定义：由 $n \times m$ 个数排成的 n 行 m 列的数表

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$

叫做一个 n 行 m 列的矩阵。a_{ij} 称为该矩阵的第 i 行第 j 列的元素。特别地，当 $n = m$ 时矩阵称为 n 阶方阵。

矩阵常用大写字母来表示。

行列式

对于 n 阶方阵 A 定义其行列式:

$$det(A) = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{vmatrix} = \sum_p (-1)^{sig(p)} a_{1,p_1} a_{2,p_2} \cdots a_{n,p_n}$$

其中 $sig(p)$ 表示排列 p 中的逆序对个数。显然根据对称性，$det(A)$ 也可以表示为

$$\sum_p (-1)^{sig(p)} a_{p_1,1} a_{p_2,2} \cdots a_{p_n,n}$$

行列式的性质

首先对矩阵定义转置 A^T 满足 $A^T_{ij} = A_{ji}$。

性质 1　$det(A) = det(A^T)$。

性质 1 说明式的行和列具有同等地位，因而凡是对行成立的性质，对列也一样成立。

性质 2　交换方阵的任意两行 (列)，行列式变号。

推论：若方阵的两行 (列) 完全一样，那么行列式等于零。

性质 3　方阵的某一行 (列) 中所有的元素都乘以同一个数 k, 行列式的值变为原来的 k 倍。

推论行列式中某一行 (列) 中所元素的公因子可以提到行列式符号的外面。

性质 4　行列式中如果有两行 (列) 元素成比例，则此行列式为零。

行列式的性质

性质 5

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1}+b_{i1} & a_{i2}+b_{i2} & \cdots & a_{im}+b_{im} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{vmatrix}$$

$$= \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{im} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ b_{i1} & b_{i2} & \cdots & b_{im} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{vmatrix}$$

性质 6 把行列式的某一行 (列) 的各元素都乘以同一数后加到另一行 (列) 对应的元素上，行列式的值不变。

行列式的展开

高阶行列式可以转化为低阶的行列式。为此，首先引入余子式和代数余子式的概念。

定义　在 n 阶行列式中，删去元素 a_{ij} 所在的第 i 行和第 j 列的元素，剩下的元素按原来的顺序构成一个 $n-1$ 阶行列式，称为元素 a_{ij} 的余子式，记作 M_{ij}。记 $A_{ij}=(-1)^{i+j}M_{ij}$ 称为其代数余子式。

那么对于任意的 i 都有

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

。

证明

考虑行列式

$$D1 = \begin{vmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \cdots & a_{nn} \end{vmatrix}$$

可以看出 $D1 = a_{11}A_{11}$。

再考虑行列式

$$D2 = \begin{vmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & a_{ij} & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{vmatrix}$$

通过交换行和列，可得 $D2 = a_{ij}A_{ij}$。

之后结合性质 5 即得证。

一个重要推论

对于 $i \neq j$，有

$$\sum_{k=1}^{n} a_{ik} A_{jk} = 0$$

证明：上式相当于将矩阵的第 j 行用第 i 行代替之后求行列式。

由于 $i \neq j$，所以矩阵中有两行完全相等，因此行列式等于零。

因此上面的定理可以写作

$$\sum_{k=1}^{n} a_{ik} A_{jk} = det(A)[i = j]$$

范德蒙 (Vandermonde) 行列式

$$D_n = \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ a_1^2 & a_2^2 & \cdots & a_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} = \prod_{i<j}(a_j - a_i)$$

证明使用数学归纳法。

$n = 1$ 时，$D_1 = \begin{vmatrix} 1 \end{vmatrix} = 1$，结论成立。

证明

$n > 1$ 时，考虑如何从行列式中消去 a_1。从最后一行开始，后一行减去前一行的 a_1 倍，得到

$$\begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & \cdots & a_n - a_1 \\ 0 & a_2(a_2 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_2^{n-2}(a_2 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix}$$

将其按照第一列展开得到。

$$a_{11}A_{11} = \begin{vmatrix} a_2 - a_1 & \cdots & a_n - a_1 \\ a_2(a_2 - a_1) & \cdots & a_n(a_n - a_1) \\ \vdots & \ddots & \vdots \\ a_2^{n-2}(a_2 - a_1) & \cdots & a_n^{n-2}(a_n - a_1) \end{vmatrix}$$

提取每一列的公因子。

$$\prod_{2\le i\le n}(a_i-a_1)\begin{vmatrix} 1 & \cdots & 1 \\ a_2 & \cdots & a_n \\ \vdots & \ddots & \vdots \\ a_2^{n-2} & \cdots & a_n^{n-2} \end{vmatrix} = \prod_{i<j}(a_j-a_i)$$

克拉默法则

设有 n 个未知数的 n 个线性方程组为

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

其系数构成一个方阵，对于第 k 列，求出其所以代数余子式 $A_{1k}, A_{2k}, \cdots, A_{nk}$。之后把第 i 个等式左右两边乘上 A_{ik} 得到

$$\begin{cases} a_{11}A_{1k}x_1 + a_{12}A_{1k}x_2 + \cdots + a_{1n}A_{1k}x_n = b_1A_{1k} \\ a_{21}A_{2k}x_1 + a_{22}A_{2k}x_2 + \cdots + a_{2n}A_{2k}x_n = b_2A_{2k} \\ \qquad\qquad\vdots \\ a_{n1}A_{nk}x_1 + a_{n2}A_{nk}x_2 + \cdots + a_{nn}A_{nk}x_n = b_nA_{nk} \end{cases}$$

克拉默法则

将这 n 个等式相加得到

$$(\sum_{i=1}^{n} a_{i1}A_{ik})x_1 + (\sum_{i=1}^{n} a_{i2}A_{ik})x_2 + \cdots + (\sum_{i=1}^{n} a_{in}A_{ik})x_n = (\sum_{i=1}^{n} b_iA_{ik})$$

根据代数余子式的性质，等式的左边等于 $det(A)x_k$，右边等于

$$\begin{vmatrix} a_{11} & \cdots & b_{1k} & \cdots & a_{1n} \\ a_{21} & \cdots & b_{2k} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & b_{nk} & \cdots & a_{nn} \end{vmatrix}$$

记作 $det(A_k)$。因此，若 $det(A) \neq 0$，则线性方程组有唯一解 $x_k = \frac{det(A_k)}{det(A)}$。

向量

只有一行的矩阵称为行向量，只有一列的矩阵称为列向量。向量通常用小写字母表示。

为避免混淆，行向量通常记作 $(a_1, a_2, \cdots, a_n)$，列向量记作 $(a_1, a_2, \cdots, a_n)^T$。

一个 n 维向量可以对应到 R^n 空间中的一个点。

线性变换

现在有一组变量 $x_1, x_2, \cdots, x_m$，求另一组变量 $y_1, y_2, \cdots, y_n$ 满足

$$\begin{cases} y_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1m}x_m \\ y_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2m}x_m \\ \quad\vdots \\ y_n = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nm}x_m \end{cases}$$

那么从这种一组变量到另一组变量的变换叫做线性变换。

将线性变换的系数构成一个矩阵，就可以用一个矩阵来描述该线性变换。再把这两组变量构成两个向量，那么可以定义矩阵和向量的乘法

$$Ax = y$$

从线性变换的角度看矩阵

在 R^n 空间中取一组向量 $e_i = (0, \cdots, 1, \cdots, 0)^T$，其中只有第 i 位为 1，这叫做一组基。那么向量 x 可以表示为

$$\sum_{i=1}^{n} x_i e_i$$

现有一个 $m \times n$ 的矩阵 $A = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix}$，那么 $y = Ax$ 就可以表示为

$$\sum_{i=1}^{n} x_i a_i$$

即一个线性变换可以看成基向量的变换。矩阵的 n 个列向量表示 n 个新的基。如果从图像上看，线性变换可以看成对空间进行的拉伸操作。

从向量的角度看行列式

行列式可以看作 n 个 R^n 空间中的向量所形成的有向体积。

首先对于对角矩阵，其行列式等于主对角线元素之积，由于对角矩阵的所有向量显然都是正交的，所以也等于这些向量形成的有向体积。

对于其他矩阵，求行列式时会运用行列式的性质来化为对角矩阵，即不断将某一行乘上一个数加到另一行上。

假设被加的行向量是 x，那么其他的 $n-1$ 个向量肯定同时位于某个 $n-1$ 维超平面内，那么所求的 n 维超体积就等于其他 $n-1$ 个向量形成的 $n-1$ 维超体积乘上 x 到该超平面的“距离”。

现在将 x 加上 ky，由于 y 属于该超平面，所以 ky 平行于该超平面，所以 $x+ky$ 的“距离”并不会改变，因此有向体积也不会改变。

因此行列式也可以表示“线性变换的压缩率”这种东西。

矩阵的运算

之前已经定义了矩阵和向量的乘法。

加法为将矩阵个位置对应相加，显然加法满足交换律和结合律。矩阵的数乘为将矩阵各个元素都乘上该数。

设 A 是 $n \times s$ 矩阵，B 是 $s \times m$ 矩阵，那么定义 $C = AB$ 满足

$$C = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_m \end{bmatrix}$$

即将 B 的每个列向量与 A 分别相乘，再合起来。矩阵乘法可以看作是线性变换的复合。

容易验证矩阵乘法满足结合律，分配律，不满足交换律。

由于 n 阶方阵的乘积还是 n 阶方阵，所以对于方阵还可以定义幂运算。

关于矩阵乘法和行列式有性质

$$|AB| = |A||B|$$

可以理解为两次线性变换复合后的压缩率等于各自的压缩率之积。

证　构造一个 $2n$ 阶的方阵

$$\begin{bmatrix} A & 0 \\ -E & B \end{bmatrix}$$

于是

$$\begin{aligned} &\begin{vmatrix} A & 0 \\ -E & B \end{vmatrix} \\ =&\begin{vmatrix} A & AB \\ -E & 0 \end{vmatrix} \\ =&(-1)^n \begin{vmatrix} AB & A \\ 0 & -E \end{vmatrix} \end{aligned}$$

因此 $|A||B| = (-1)^n|AB|(-1)^n = |AB|$

矩阵的逆

设 A 为 n 阶方阵，若存在 n 阶方阵 B 满足

$$AB = BA = E$$

则称 B 为 A 的逆矩阵。可以证明，这样的 B 是唯一的，因此记 A 的逆为 A^{-1}，即 $A^{-1} = B$

定理 $|A| \neq 0$ 是矩阵 A 可逆的充分必要条件。

证 首先若 A 可逆，则存在 A^{-1} 满足 $AA^{-1} = E$。于是 $|A||A^{-1}| = 1$，所以 $|A| \neq 0$。

若 $|A| \neq 0$，则构造

$$A^{-1} = \frac{1}{|A|}\begin{bmatrix} A_{11} & A_{21} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{nn} \end{bmatrix}$$

即将 A 中每个元素用其代数余子式代替，转置后再数乘上 $\frac{1}{|A|}$ 得到。容易验证该矩阵与 A 相乘的结果是单位矩阵。

这个矩阵被称为伴随矩阵。

矩阵的逆有性质

$$(AB)^{-1} = B^{-1}A^{-1}$$

证

$$\begin{aligned}(AB)^{-1} &= B^{-1}B(AB)^{-1} \\ &= B^{-1}A^{-1}AB(AB)^{-1} \\ &= B^{-1}A^{-1}\end{aligned}$$

该性质可以推广到任意多个矩阵

$$(A_1A_2\cdots A_k)^{-1} = A_k^{-1}\cdots A_2^{-1}A_1^{-1}$$

利用向量可以把线性方程组表示为以下形式

$$Ax = b$$

于是若 A 可逆，那么可以得到 $x = A^{-1}b$，这和克拉默法则得到的解是相同的。

矩阵的秩

设 A 是一个 $n \times m$ 的矩阵，在 A 中选取 k 行 k 列，形成的 k^2 个交叉点处的元素构成了一个 k 阶方阵，该方阵的行列式的值叫做矩阵 A 的一个 k 阶子式。

若矩阵 A 中最高阶的非零子式阶数为 r，那么称矩阵 A 的秩是 r。

矩阵 A 的秩用 $r(A)$ 来表示。

矩阵的初等变换

矩阵的初等变换是线性代数中的基本运算，它在求矩阵的秩、求矩阵的逆和解线性方程组等方面起着重要的作用。前面求行列式的时候实际上已经用过初等变换了。

矩阵的初等变换有三种

1. 对换变换：对调两行 (列)
2. 数乘变换：以数 $k(k \neq 0)$ 乘某一行 (列) 中的所有元素
3. 倍加变换：把某一行 (列) 所有元素的 k 倍加到另一行 (列) 对应的元素上去

若矩阵 A 通过有限次初等变换变成矩阵 B，则称 A 与 B 等价。等价关系具有自反性，对称性和传递性。

矩阵的初等变换不会改变矩阵的秩，前两种变换比较显然，对于倍加变换，设矩阵 B 是由矩阵 A 将第 i 行的 k 倍加到第 j 行上得到。

设矩阵 A 的秩为 r，那么对于 B 中的任一 $r+1$ 阶子式 D_{r+1}

1. 不包含 j 行元素，那么 D_{r+1} 也是 A 的子式，故 $D_{r+1}=0$。
2. 包含 i,j 两行元素，由行列式的性质可得，D_{r+1} 与 A 中对应子式的值相同，故 $D_{r+1}=0$。
3. 包含 j 行元素而不包含 i 行元素，即

$$D_{r+1}=\begin{vmatrix} \cdots & \cdots & \cdots \\ a_{j,t_1}+ka_{i,t_1} & \cdots & a_{j,t_{r+1}}+ka_{i,t_{r+1}} \\ \cdots & \cdots & \cdots \end{vmatrix}$$

$$=\begin{vmatrix} \cdots & \cdots & \cdots \\ a_{j,t_1} & \cdots & a_{j,t_{r+1}} \\ \cdots & \cdots & \cdots \end{vmatrix}+k\begin{vmatrix} \cdots & \cdots & \cdots \\ a_{i,t_1} & \cdots & a_{i,t_{r+1}} \\ \cdots & \cdots & \cdots \end{vmatrix}$$

两个子式都属于 A，故 $D_{r+1}=0$。

由此有 $r(B)\le r(A)$。反过来 B 也是由 A 经过一次倍加变换得到的，即 $r(A)\le r(B)$，所以 $r(A)=r(B)$。

标准形和行阶梯形

由于初等变换不改变矩阵的秩，所以可以用初等变换来化简矩阵。

运用初等变换，所有的矩阵都可以被化成下面的形式，其中 E_r 是 r 阶单位矩阵。

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

那么矩阵的秩就是 r 即左上角单位矩阵的阶数。这种形式被称为矩阵的标准形。

实际上求秩的时候并不需要化到标准形。只需化为行阶梯形矩阵即可，其特点是

1. 若有零行，则零行全部在矩阵的下方。
2. 从第一行起，每行第一个非零元素前面零的个数逐行增加。

对于这样的矩阵，可画出一条阶梯线，线的下方全为 0。则矩阵的秩就是非零行数。

初等方阵

矩阵的初等变换可以通过乘一些特殊的矩阵来表示，这些特殊的矩阵被称为初等方阵。

三种初等变换所对应的初等方阵分别是

$$
\begin{bmatrix}
1 & & & & & & \\
& \ddots & & & & & \\
& & 0 & \cdots & 1 & & \\
& & \vdots & \ddots & \vdots & & \\
& & 1 & \cdots & 0 & & \\
& & & & & \ddots & \\
& & & & & & 1
\end{bmatrix}
\begin{bmatrix}
1 & & & & \\
& \ddots & & & \\
& & k & & \\
& & & \ddots & \\
& & & & 1
\end{bmatrix}
\begin{bmatrix}
1 & & & & \\
& \ddots & & & \\
& \vdots & \ddots & & \\
& k & \cdots & \ddots & \\
& & & & 1
\end{bmatrix}
$$

于是，设 A 是一个 $n \times m$ 的矩阵，那么对 A 进行一次初等行变换等价于左乘上一个 n 阶的对应初等矩阵，而列变换等价与右乘上一个 m 阶的对应初等矩阵。

运用初等变换求逆

由前面的讨论可知，任何矩阵在经过有限次初等变换后都可以化成 $\begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}$ 的形式。用初等方阵可以表示为

$$P_l \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_r = \begin{bmatrix} E_r & 0 \\ 0 & 0 \end{bmatrix}$$

对于可逆矩阵，其标准形显然是 E，于是上式变为

$$P_l \cdots P_2 P_1 A Q_1 Q_2 \cdots Q_r = E$$
$$P_l \cdots P_2 P_1 A = Q_r^{-1} Q_{r-1}^{-1} \cdots Q_1^{-1}$$
$$Q_1 \cdots Q_{r-1} Q_r P_l \cdots P_2 P_1 A = E$$

重编号得 $P_l \cdots P_2 P_1 A = E$，所以 $A^{-1} = P_l \cdots P_2 P_1$。于是我们只需将 A 通过初等行变换化为单位矩阵，并将这些行变换在 E 上也做一遍，就可以得到 A^{-1} 了。

矩阵分块

在矩阵中画很多条横线和竖线，会将矩阵分成若干块。这样就形成了一个新的矩阵，这个矩阵中的元素也是矩阵，这个矩阵叫做分块矩阵。

分块矩阵的加法和数乘和普通矩阵一样。分块矩阵 A 和 B 的乘法要求 A 的行和 B 的列分法一致。之后按照普通矩阵的乘法规则运算即可。

一道题目

现在有 125 台计算机和两个 125 阶方阵 A 和 B，第 i 个计算机中存储了 $A_{i,1}, A_{i,2}, \cdots, A_{i,125}$ 和 $B_{i,1}, B_{i,2}, \cdots, B_{i,125}$。每台计算机都有很多存储空间，任意两个计算机之间都存在通信管道，计算机之间只能通过通信管道来传递信息。

每回合每台计算机将依次执行下列操作

1. 在本地进行运算，比如读取本地存储的信息，其他计算机发来的信息，进行计算等等。
2. 对于每一台其他的计算机，可以选择发送一个数，注意这个数需要在下一回合才能被目标计算机读取到。

请设计一种方法，在尽量少的回合内计算出 $C = AB$，并将 $C_{i,1}, C_{i,2}, \cdots, C_{i,125}$ 存储到第 i 台计算机内。

所有的变量都是 unsigned int 型，乘法自然溢出即可。

暴力

用 125 个回合将 B 矩阵传到每一个计算机内，第 126 回合计算即可。

更快一些的做法

首先将矩阵的行和列都分成 5 个大小 25 的块，这样 A 和 B 都变成了 5×5 的分块矩阵，其中的每个元素都是 25×25 的矩阵。

这样在做分块矩阵乘法时，一共需要计算 $5^3 = 125$ 次 25×25 的矩阵乘法。

我们将着 125 次矩阵乘法的计算任务分配给 125 台计算机。然后按照下面的方法做。

- 1 到 50 回合，将所需要的数据传输到对应的计算机上。
- 51 回合各自完成计算。
- 51 回合完成计算后开始往回传输，由于传输量小了一半，所以到 75 回合就传完了。
- 76 回合各自把结果加起来，完成计算。

其实还可以更快，可是我不会。

向量的线性相关

直接给出定义：设有 n 维向量 $\beta, \alpha_1, \alpha_2, \cdots, \alpha_m$，如果存在一组数 $k_1, k_2, \cdots, k_m$ 满足

$$\beta = k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m$$

则称 β 是 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 的线性组合，也称可以被线性表示。

线性相关的定义：设有 n 维向量 $\alpha_1, \alpha_2, \cdots, \alpha_m$，如果存在一组不全为零的数 $k_1, k_2, \cdots, k_m$ 满足

$$k_1\alpha_1 + k_2\alpha_2 + \cdots + k_m\alpha_m = 0$$

则称向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 线性相关，否则是线性无关的。也就是说，若向量组是线性无关的，上式只有 $k_1 = k_2 = \cdots = k_m = 0$ 时才成立。

可以看出，向量组中存在一个向量可以被其他的向量线性表示是向量组线性相关的充分必要条件。

一点性质

性质 1　若 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性相关，则 $\alpha_1, \alpha_2, \cdots, \alpha_r, \alpha_{r+1}$ 也线性相关。

这说明一个向量组若部分组线性相关，则整体线性相关。反之，若一个向量组线性无关，则它的任一部分组也线性无关。

性质 2　设有两个向量组

- $A : \alpha_i = (a_{1,i}, a_{2,i}, \cdots, a_{r,i})$
- $B : \beta_i = (a_{1,i}, a_{2,i}, \cdots, a_{r,i}, a_{r+1,i})$

即 β_i 是由 α_i 加上一个分量得到的。若向量组 A 线性无关，则向量组 B 也线性无关。

向量组 $\alpha_1, \alpha_2, \cdots, \alpha_m$ 构成矩阵 $A = (\alpha_1, \alpha_2, \cdots, \alpha_m)$，向量组线性相关相当于 $Ax = 0$ 有非零解。有如下定理

n 元齐次线性方程组 $Ax = 0$ 有非零解的充分必要条件是系数矩阵 A 的秩 $r(A) < n$

证　必要性：用反证法，若 $r(A) = n$，则存在一个 n 阶非零子式，那么根据克拉默法则，该子式对应的 n 个方程只有零解，矛盾。

充分性，因为 $r(A) < n$，那么将系数矩阵 A 化为行阶梯形式后会有 $n - r$ 个自由变量，令它们取非零值即可。

由该定理可得

向量组 $A = (\alpha_1, \alpha_2, \cdots, \alpha_m)$ 线性无关的充要条件是 $r(A) = m$。

向量组的等价

现有两个向量组 A, B，若 A 中的每一个向量都可由向量组 B 线性表示，则称向量组 A 可由向量组 B 线性表示。若向量组 B 也可由向量组 A 线性表示，则称两个向量组等价。

A 可由 B 线性表示，用矩阵乘法的方式可以表达为

$$A = BK$$

其中 K 是系数矩阵。

向量组之间的线性表示具有传递性，即若向量组 A 可由向量组 B 线性表示，向量组 B 可由向量组 C 线性表示，则向量组 A 可由向量组 C 线性表示。

极大线性无关组

设向量组 A 的一个子组 $a_1, a_2, \cdots, a_r$ 满足

1. $a_1, a_2, \cdots, a_r$ 线性无关。
2. 向量组 A 中每一个向量均可由此子组线性表示。

则称 $a_1, a_2, \cdots, a_r$ 是向量组 A 的一个极大线性无关组。把 r 称作该向量组的秩。由定义可知，向量组和其极大线性无关组是等价的。

极大线性无关组可能不止一个，由于它们都和原向量组等价，所以它们互相也是等价的。

定理　设向量组 A 的秩为 r，向量组 B 的秩为 s，若向量组 A 能由向量组 B 线性表示，则 $r \le s$。

证　设向量组 A 的极大线性无关组为 A_0，向量组 B 的极大线性无关组为 B_0。因向量组 A_0 能由向量组 A 线性表示，向量组 A 能由向量组 B 线性表示，向量组 B 能由向量组 B_0 线性表示，故向量组 A_0 能由向量组 B_0 线性表示，即存在 $s \times r$ 的矩阵 K 使得

$$A_0 = B_0 K$$

假设 $r > s$，那么 $r(K) \le s < r$，那么方程组 $Kx = 0$ 存在一组非零解 $x \neq 0$。将上式两边右乘上 x 得

$$A_0 x = 0$$

即 A_0 是线性相关的，矛盾，所以 $r \le s$。

向量组的秩与矩阵的秩

设 A 是一个 $n\times m$ 的矩阵，可把 A 看作 m 个列向量构成的向量组，也可把 A 看作 n 个行向量构成的矩阵。这两个向量组的秩分别叫做列秩和行秩。有定理：矩阵的秩等于列秩，也等于行秩。

证　若 $r(A)=0$ 则显然成立，现设 $r(A)=r>0$，那么 A 中必存在一个非零的 r 阶子式，不失一般性地设该子式选取的列集合就是前 r 列，由向量组线性无关的充要条件可知这 r 个列向量是线性无关的。

又因为 A 中所有的 $r+1$ 阶子式都等于零，所以任选一个大小为 $r+1$ 的列向量组都是线性相关的，由此可得所有的列向量都可以由前 r 个列向量线性表示，于是前 r 个列向量就是列向量组的一个极大线性无关组，于是列秩等于矩阵的秩。

同理可得行秩也等于矩阵的秩。

小练习

证明 $r(AB) \leq min(r(A), r(B))$。

证明 $r(A+B) \leq r(A) + r(B)$。

向量空间

设 V 是 n 维向量构成的非空集合，且满足

1. 若 $\alpha, \beta \in V$，则 $\alpha + \beta \in V$。
2. 若 $\alpha \in V, k \in R$，则 $k\alpha \in V$。

那么称集合 V 是向量空间。

在 V 中选取向量组 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 满足

1. $\alpha_1, \alpha_2, \cdots, \alpha_r$ 线性无关。
2. V 中的任何向量都是 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 的线性组合。

则称 $\alpha_1, \alpha_2, \cdots, \alpha_r$ 为向量空间的一组基，r 称为向量空间的维数。不难看出，基就是向量空间中的极大线性无关组，维数等同于秩。

齐次线性方程组

n 元齐次线性方程组 $Ax=0$ 总是有解的，即 $x=0$，由之前的讨论可知，非零解存在的条件是 $r(A)<n$，接下来研究这些解的性质。

可以发现两个性质

1. 若 u_0, u_1 是方程组的解，那么 u_0+u_1 也是方程组的解。
2. 若 u 是方程组的解，那么 ku 也是方程组的解。

这表明齐次线性方程组的解集 S 是一个向量空间。

解空间的基

设 $r(A)=r$，不失一般性，设 A 的前 r 个列向量是线性无关的，那么对 A 进行初等行变换可以得到

$$\begin{bmatrix} \mathrm{E} & \begin{matrix} b_{1,r+1} & b_{1,r+2} & \cdots & b_{1,n} \\ b_{2,r+1} & b_{1,r+2} & \cdots & b_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{r,r+1} & b_{1,r+2} & \cdots & b_{r,n} \end{matrix} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ x_n \end{bmatrix} = 0$$

解空间的基

那么 $x_{r+1},\cdots,x_n$ 这些变量都是自由元，令 $x_{r+i}=k_i$，解得

$$
\begin{array}{llllll}
x_1 & = -k_1 b_{1,r+1} & -k_2 b_{1,r+2} & \cdots & -k_{n-r} b_{1,n} \\
x_2 & = -k_1 b_{2,r+1} & -k_2 b_{2,r+2} & \cdots & -k_{n-r} b_{2,n} \\
\vdots & & & & \\
x_r & = -k_1 b_{r,r+1} & -k_2 b_{r,r+2} & \cdots & -k_{n-r} b_{r,n} \\
x_{r+1} & = k_1 & & & \\
x_{r+2} & = & k_2 & & \\
\vdots & & & \ddots & \\
x_n & = & & & k_{n-r}
\end{array}
$$

解空间的基

将上式写成向量形式，并把系数提出来可得

$$x = k_1 \begin{bmatrix} -b_{1,r+1} \\ -b_{2,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + k_2 \begin{bmatrix} -b_{1,r+2} \\ -b_{2,r+2} \\ \vdots \\ -b_{r,r+2} \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \cdots + k_{n-r} \begin{bmatrix} -b_{1,n} \\ -b_{2,n} \\ \vdots \\ -b_{r,n} \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

于是求得一组基。并且解空间的维度等于未知元个数减去系数矩阵的秩。

非齐次线性方程组

设有非齐次线性方程组 $Ax=b$，称 A 为方程组的系数矩阵，$A'=(A|b)$ 为其增广矩阵。

首先要判断该方程组是否有解，有解的充要条件是：$r(A)=r(A')$

证　必要性，若 $Ax=b$ 有解，即 b 是 A 列向量组的线性组合，那么 A 的极大线性无关组也是 A' 的极大线性无关组，所以 $r(A)=r(A')$。

充分性类似。

对于非齐次线性方程组的解有两个显然的性质

1. 若 u_0, u_1 是 $Ax=b$ 的解，那么 u_0-u_1 是 $Ax=0$ 的解。
2. 若 u 是 $Ax=b$ 的解，v 是 $Ax=0$ 的解，那么 $u+v$ 是 $Ax=b$ 的解。

由此可得 $Ax=b$ 的通解形式为

$$u^*+v$$

其中 u^* 是 $Ax=b$ 的一组特解，v 是 $Ax=0$ 的解。

向量的内积

直观起见，先引入向量模长的概念，n 维向量 a 的模长 $|a|$ 等于

$$\sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$$

然后再定义向量的内积

$$a \cdot b = |a||b|cos\theta$$

其中 θ 表示两个向量之间的夹角。可以证明，内积满足交换律和分配率。

关于 R^n 空间中的夹角，我们只知道不同坐标轴上的向量之间的夹角是 $90°$。那么根据内积的分配律，有如下结论

$$\begin{aligned} a \cdot b &= (a_1 e_1 + a_2 e_2 + \cdots + a_n e_n) \cdot (b_1 e_1 + b_2 e_2 + \cdots + b_n e_n) \\ &= a_1 e_1 \cdot b_1 e_1 + a_2 e_2 \cdot b_2 e_2 + \cdots + a_n e_n \cdot b_n e_n \\ &= a_1 b_1 + a_2 b_2 + \cdots + a_n b_n \end{aligned}$$

由此 $a \cdot b$ 也可记作 $a^T b$

正交向量组与正交矩阵

若 $a^Tb = 0$，则称向量 a 与 b 正交，若一个向量组中的向量两两正交，则称这是一个正交向量组。正交向量组必定是线性无关的。

若 n 阶方阵 A 满足 $AA^T = E$，则称 A 是正交矩阵。

可以发现，正交矩阵的列向量组是一个正交向量组，且所有列向量的模长都是 1，行向量组同理。

相似矩阵

设有 n 阶方阵 A, B，若存在可逆矩阵 P 满足

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

则称 B 是 A 的相似矩阵。可以发现相似关系具有自反性，对称性和传递性。

相似矩阵有一个很好的性质

$$B^k = P^{-1}A^kP$$

因此如果能找到矩阵 P 使得 B 尽可能的简单，那么矩阵的运算就可以得到简化。而对角矩阵就是最简单的矩阵。方便起见，把对角矩阵记作 $diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$。

那么矩阵 P 满足

$$\begin{aligned} P^{-1}AP &= diag(\lambda_1, \lambda_2, \cdots, \lambda_n) \\ AP &= Pdiag(\lambda_1, \lambda_2, \cdots, \lambda_n) \\ [Ap_1|Ap_2|\cdots|Ap_n] &= [\lambda_1 p_1|\lambda_2 p_2|\cdots|\lambda_n p_n] \end{aligned}$$

即对于 P 的每一个列向量有 $Ap_i = \lambda_i p_i$，且 P 的所有列向量是线性无关的。

特征值与特征向量

设 A 是 n 阶方阵，如果数 λ 和 n 维非零向量 x 满足

$$Ax = \lambda x$$

那么称 λ 是矩阵 A 的特征值，x 是对应特征值 λ 的特征向量。
如何求特征值和特征向量呢，我们把上式写成

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

这是一个齐次线性方程组，它有非零解的条件是系数矩阵的行列式等于零。即

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

左边的行列式展开后是一个关于 λ 的多项式，称为 A 的特征多项式，记作 $f(\lambda)$。

所以矩阵的特征值就是方程 $f(\lambda)=0$ 的所有解。由代数基本定理可知 $f(\lambda)=0$ 在复数域内必有 n 个根（重根算两次）。

如果现在求出了一个特征值 λ，那么

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

求出解空间就可以得到对应的特征向量了。

至此可得方阵 A 与对角矩阵 B 相似的充要条件是 A 有 n 个线性无关的特征向量。

关于特征值有两个性质。

1. $\prod_{i=1}^{n} \lambda_i = det(A)$
2. $\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} A_{ii}$

也这两个性质不知道有什么用。

关于特征多项式有一个定理

$$f(A) = 0$$

这个定理可以用来优化一些阶数很高的线性递推。

一个定理

设 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 是 n 阶方阵 A 的 m 个特征值，$p_1, p_2, \cdots, p_m$ 是其对应的特征向量。

若 $\lambda_1, \lambda_2, \cdots, \lambda_m$ 互不相同，则 $p_1, p_2, \cdots, p_m$ 线性无关。

证　若存在一组不全为零的数使得

$$x_1 p_1 + x_2 p_2 + \cdots + x_m p_m = 0$$

上式不断左乘矩阵 A 之后有

$$(x_1 p_1, x_2 p_2, \cdots, x_m p_m) \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{m-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{m-1} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_m & \cdots & \lambda_m^{m-1} \end{bmatrix} = 0$$

上式中矩阵的行列式是范德蒙行列式，从而该矩阵可逆，可以推出所有的 x_i 都等于零。

另一个定理

相似矩阵的特征多项式和特征值相同。

证　由于 $B=P^{-1}AP$，则

$$|B-\lambda E|=|P^{-1}AP-P^{-1}(\lambda E)P| \quad =|P^{-1}(A-\lambda E)P|=|p^{-1}||A-\lambda E||P| \quad =|A-\lambda E|$$

线性空间

线性空间是向量空间 R^n 的推广。

设 V 是一个非空集合，R 是实数域，在 V 中定义两种运算，一种称为加法，另一种称为数乘。这两种运算都满足封闭性，并且满足下面 8 条运算法则：

1. 加法交换律 $\alpha+\beta=\beta+\alpha$
2. 加法结合律 $(\alpha+\beta)+\gamma=\alpha+(\beta+\gamma)$
3. V 中存在零元素 0，$\forall\alpha\in V, \alpha+0=\alpha$
4. V 中每个元素有负元素，即 $\forall\alpha\in V$，存在 $\beta\in V$，使 $\alpha+\beta=0$，记为 $\beta=-\alpha$
5. $1\cdot\alpha=\alpha$
6. 数乘结合律 $(kl)\alpha=k(l\alpha)$
7. 分配律 $(k+l)\alpha=k\alpha+l\alpha$
8. 分配律 $k(\alpha+\beta)=k\alpha+k\beta$

称 V 为线性空间，V 中的元素称为向量。

线性空间的同构

和向量空间类似，线性空间存在基，随意找一组基 $\alpha_1, \alpha_2, \cdots, \alpha_n$，那么线性空间中的向量都可以表示为这组基的线性组合，于是每个向量都可以用一组数来表示。

于是 V 中的向量就和 R_n 中的向量一一对应了，因此可以说 V 与 R_n 同构，显然所有维数相同的线性空间都是同构的。

因此线性空间的结构完全被它的维数所决定。