矩阵的秩与线性方程组

矩阵的秩

子矩阵

在`m*n`矩阵A中，任意取k行l列，选取交叉处元素组成的矩**行列式**称为A的`k*l`子矩阵。当k=l，称为矩阵A的一个**k阶子式**

如

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix}$$

选取1 2 3行和1 2 4列组成的一个3阶子式为

$$A = \begin{vmatrix} 1 & 2 & 4 \\ 5 & 6 & 8 \\ 9 & 10 & 12 \end{vmatrix}$$

秩

定义

若A中有一个**不为0的**r阶子式D，且所有的r+1阶子式全为0，那么D称为A的一个**最高阶非零子式**，r为矩阵的秩，记作R(A)。**零矩阵的秩为0**

- 若矩阵A有一个r阶子式不为0，则R(A)>=r。若所有r+1阶子式全为0，则R(A)<=r
- 若B是A的子矩阵，则$R(B) \leq R(A)$

满秩矩阵与降秩矩阵

若A是n阶方阵，则n阶子式只有一个|A|，则

- 若$|A| \neq 0$，R(A) = n，称为满秩矩阵
- |A|=0，R(A)<n，称为降秩矩阵

初等变换

初等变换不改变矩阵的秩

因此可以将矩阵化为行阶梯矩阵，从而直接求得矩阵的秩

性质

- $0 \leq R(A_{m \times n}) \leq min\{m, n\}$
- $R(A) = R(A^T)$
- 若P Q可逆，$R(PA) = R(AQ) = R(PAQ)$
- $max\{R(A), R(B)\} \leq R(A, B) \leq R(A) + R(B)$
- $R(A + B) \leq R(A, B) \leq R(A) + R(B)$
- A B是`m*n`和`n*s`矩阵，$R(AB) \leq min\{R(A), R(B)\}$
- A B是`m*n`和`n*s`矩阵，$R(AB) \geq R(A) + R(B) - n$ **西尔维斯特（Sylvester）不等式**
- A B是`m*n`和`n*s`矩阵，$R(A) + R(B) \leq n$

线性方程组解的判定

几个结论

对于n元线性方程组 Ax=b，$\tilde{A}$为增广矩阵(A b)

- 有解的充要条件为 $R(\tilde{A}) = R(A)$
- 有唯一解的充要条件为 $R(A) = R(\tilde{A}) = n$
- 有无穷多解的充要条件为 $R(A) = R(\tilde{A}) < n$

因此对于n元齐次线性方程组 Ax=0，有非零解的充要条件是 $R(A) < n$

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

分块矩阵的初等变换

分块初等变换

与矩阵初等变换相似，分别为

- 互换分块矩阵的行/列
- 用一个可逆矩阵左乘/右乘分块矩阵的某一行/列
- 把分块矩阵某一行/列左乘/右乘一个矩阵后加到另一行/列

分块单位矩阵

形如下式的矩阵

$$\begin{pmatrix} E_m & O \\ O & E_n \end{pmatrix}$$

分块初等矩阵

对分块单位矩阵做**一次**分块初等变换后得到的矩阵

对分块矩阵做一次初等变换相当于用一个分块初等矩阵左乘/右乘该矩阵

性质

- 分块初等矩阵可逆
- 若分块矩阵A经过有限次初等变换变为B，则两矩阵等价
- 分块初等变换不改变矩阵的秩

分块初等矩阵的秩

设A为`m*n`矩阵，B为`s*t`矩阵，C为`m*t`矩阵

- 定理1

$$R(\begin{pmatrix} A & C \\ O & B \end{pmatrix}) \geq R(A) + R(B)$$

- 定理2

$$R(\begin{pmatrix} A & O \\ O & B \end{pmatrix}) = R(A) + R(B)$$

- 弗罗贝尼乌斯不等式

设A为`m*n`矩阵 B为`n*s`矩阵 C为`s*t`矩阵

$$R(AB)+R(BC)\le R(B)+R(ABC)$$

将A B C换为X(`m*n`) E(`n*n`) Y(`n*p`)得西尔维斯特不等式

$$R(X)+R(Y)\le n+R(XY)$$

将A B C换为X Y O和X Y O，分别得$R(XY)\le R(Y)$和$R(XY)\le R(X)$，即上面性质中的不等式

$$R(AB)\le min\{R(A),R(B)\}$$

向量空间

向量

矩阵的向量分解

$\boldsymbol{\alpha}_i$表示**列向量**，$\boldsymbol{\beta}_i$表示**行向量**，

$$\boldsymbol{A}=\begin{pmatrix}\boldsymbol{\alpha}_1 & \boldsymbol{\alpha}_2 & \cdots & \boldsymbol{\alpha}_n\end{pmatrix}=\begin{pmatrix}\boldsymbol{\beta}_1\\ \boldsymbol{\beta}_2\\ \vdots\\ \boldsymbol{\beta}_n\end{pmatrix}$$

方程组的向量表示

AX=B可以表示为

$$x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_n\boldsymbol{\alpha}_n=\boldsymbol{\beta}$$

向量空间的公理

公理

令V为一定义了加法（即V中每一对元素x和y，唯一对应于V中一个元素x+y）和标量乘法（V中每一个元素x和一个标量a，唯一对应于V中一个元素ax）运算的集合

- 交换律 对V中任何x和y，$x+y=y+x$
- 加法结合律 对V中任何x y z，$(x+y)+z=x+(y+z)$
- 零元 V中存在一个元素0，满足对任意的$x\in V$，有$x+0=x$
- 逆元 对每一个$x\in V$，存在V中一个元素-x，满足$x+(-x)=0$
- 标量乘法分配率 对任意标量a和V中的元素x和y，有$a(\boldsymbol{x}+\boldsymbol{y})=a\boldsymbol{x}+a\boldsymbol{y}$
- 标量乘法结合律 对任意标量a和b及V中元素x，有$(ab)\boldsymbol{x}=a(b\boldsymbol{x})$
- 单位元 对V中所有x，有$1\cdot\boldsymbol{x}=\boldsymbol{x}$
- 对任意标量a和b及V中元素x，有$(a+b)\boldsymbol{x}=a\boldsymbol{x}+b\boldsymbol{x}$

封闭性

标量乘法的封闭性：$\text{若}x\in V\text{，且}a\text{为标量，则}ax\in V$

向量加法的封闭性：$\text{若}x,y\in V\text{，则}x+y\in V$

其他性质

若V为向量空间，x为V中任意元素，则

- $0\boldsymbol{x}=\boldsymbol{0}$
- $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{0}$蕴含$\boldsymbol{y}=-\boldsymbol{x}$，即加法逆元是唯一的
- $(-1)\boldsymbol{x}=-\boldsymbol{x}$

几个典型向量空间

C[a, b]

C[a,b]表示所有定义在闭区间[a,b]上的实值连续函数，全集为一个函数集合。若向量为C[a,b]中的函数，对所有[a, b]中的x：

- 两个函数和f+g定义为

$$(f+g)(x)=f(x)+g(x)$$

- 标量乘法运算定义为

$$(af)(x)=af(x)$$

要证明其为向量空间，需要证明上述定义符合公理：

- $$(f+g)(x)=f(x)+g(x)=g(x)+f(x)=(g+f)(x)$$
- $$(f+g)(x)+h(x)=f(x)+g(x)+h(x)=f(x)+(g+h)(x)$$
- 存在函数z(x)=0，则 f(x) + z(x) = f(x) + 0 = f(x)

如上一条一条验证（不写了）

P_n

P_n表示次数小于n的所有多项式的集合，定义p+q和ap为对所有的实数x，有

$$(p+q)(x)=p(x)+q(x)$$
$$(ap)(x)=ap(x)$$

也可一条一条验证。其中存在零向量z(x)

$$z(x)=0x^{n-1}+0x^{n-2}+\cdots+0x+0$$

子空间

若S为向量空间V的非空子集，且满足

- 对任意标量a，若$x\in S$，有$ax\in S$
- 若$x\in S$且$y\in S$，则$x+y\in S$

即S中的标量乘和加法运算满足封闭性，其中所有向量空间S都包含零子空间$\{\boldsymbol{0}\}$

证明S非空的最简单的方法即证明零向量在S中

$C^n[a,b]$

令C[a,b]表示定义域[a,b]上的所有函数，$C^n[a,b]$为定义域[a,b]上的所有n阶连续可导的函数f的集合，而由于n阶可导函数f与标量a的乘积得到的函数af也是n阶可导的，且对于n阶可导函数g，f+g也是n阶可导的，因此$C^n[a,b]$是C[a,b]的子空间

矩阵的零空间

令A为一`m*n`的矩阵，令N(A)为所有齐次方程组Ax=0的解的集合，即

$$N(A)=\{\boldsymbol{x}\in R^n|A\boldsymbol{x}=\boldsymbol{0}\}$$

则由于

$$A(a\boldsymbol{x})=aA\boldsymbol{x}=a\boldsymbol{0}=\boldsymbol{0}$$

即标量乘法具有封闭性

此外

$$A(\boldsymbol{x}+\boldsymbol{y})=A\boldsymbol{x}+A\boldsymbol{y}=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$$

即加法具有封闭性

则$N(A)$为R^n的一个子空间，所有齐次方程组$A\boldsymbol{x}=\boldsymbol{0}$的解的集合构成了$R^n$的一个子空间。这里$N(A)$称为$A$的零空间

向量集合的张成

令v1, v2, ... , vn为向量空间V中的向量，a1v1 + a2v2 + ... + an vn（a1, a2, ... , an为标量）称为向量v1, v2, ... , vn的线性组合。向量v1, v2, ... , vn的所有线性组合构成的集合称为v1, ... , vn的**张成**，记为$Span(v_1,...,v_n)$

若v1, v2, ... , vn为向量空间V中的元素，则$Span(v_1,v_2,...,v_n)$为V的一个子空间

向量空间的张集

若向量空间中的每个向量都可写为$v_1,v_2,\cdots,v_n$的一个线性组合，则称$\{v_1,v_2,\cdots,v_n\}$是V的一个**张集**

向量组的线性相关性

线性组合

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta}$是n维向量，若存在一组数$k_1,k_2,\cdots,k_m$，使得

$$\boldsymbol{\beta}=k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m$$

则称向量$\boldsymbol{\beta}$是向量组$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$的**线性组合**，$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$**线性表示**

充要条件

$\boldsymbol{\beta}$可由$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$**线性表示** 等价于

线性方程组$\boldsymbol{\beta}=x_1\boldsymbol{\alpha}_1+x_2\boldsymbol{\alpha}_2+\cdots+x_m\boldsymbol{\alpha}_m$有解 等价于

矩阵$\boldsymbol{A}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m)$与增广矩阵$\tilde{\boldsymbol{A}}=(\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m,\boldsymbol{\beta})$的秩相等

唯一线性表示

由上面可知，要使线性表示唯一，必有

$$R(\boldsymbol{A})=R(\tilde{\boldsymbol{A}})=m$$

线性相关性

设$\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_m$是n维向量，若存在一组**不全为0**的数$k_1,k_2,\cdots,k_m$，使得

$$k_1\boldsymbol{\alpha}_1+k_2\boldsymbol{\alpha}_2+\cdots+k_m\boldsymbol{\alpha}_m=\mathbf{0}$$

则称向量组**线性相关**，否则为**线性无关**

充要条件

向量组$a_1,a_2,\cdots,a_n$线性相关的充要条件是它构成的矩阵$A=(a_1,a_2,\cdots,a_n)$的行列式|A| = 0。反之即为线性无关的充要条件。

函数的线性相关性

若函数$f_1,f_2,\cdots,f_n$在向量空间$C^{n-1}[a,b]$内（即这些函数都是n-1阶连续可导函数），则可以由此判断该函数向量的线性相关性：

若函数向量线性相关，则有

$$c_1f_1(x)+c_2f_2(x)+\cdots+c_nf_n(x)=0$$

两端同时依次取n阶导数，可以得到下列`n*n`方阵

$$\begin{bmatrix} f_1(x) & f_2(x) & \cdots & f_n(x) \\ f_1'(x) & f_2'(x) & \cdots & f_n'(x) \\ \vdots & \vdots & & \vdots \\ f_1^{(n-1)}(x) & f_2^{(n-1)}(x) & \cdots & f_n^{(n-1)}(x) \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

其中右边的方阵称为**朗斯基行列式**

若存在一个点x0，使得上式不成立，则函数**线性无关。**但反之，若上式恒成立不能直接说明函数线性相关。

一个例子

判断函数e^x,e^{-x}的线性相关性

$$W[e^x,e^{-x}]=\begin{bmatrix} e^x & e^{-x} \\ e^x & -e^{-x} \end{bmatrix}=-2$$

因此线性无关

基和维数

基

若向量空间张集中的元素是线性无关的，则它就是最小的。**最小张集**可以构成向量的**基。**

当且仅当向量空间V中的向量$v_1, v_2, \cdots, v_n$，满足

- $v_1, v_2, \cdots, v_n$线性无关
- $v_1, v_2, \cdots, v_n$张成V

则称为向量空间V的基

坐标变换

若向量x由一对基[e1, e2]的坐标变换到[u1, u2]的坐标，则应先求**转移矩阵**U

$$U = \begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix}$$

其中u1，u2为新的基在原本的基下的坐标的**列向量**

转移矩阵U代表了从[u1, u2]坐标变换到[e1, e2]的矩阵，因此反过来从[e1, e2]变换到[u1, u2]的矩阵为U^{-1}，即

$$\boldsymbol{c} = U^{-1}\boldsymbol{x}$$

其中c为新坐标的列向量，x为旧坐标的列向量

若问题为两个非标准基[v1, v2]到[u1, u2]的变换，实际上即为

$$V\boldsymbol{c} = U\boldsymbol{d}$$

则可以看成两步，第一步求v到标准基e的变换V，第二步求标准基e到u的变换U^{-1}，因此坐标变换公式为

$$\boldsymbol{d} = U^{-1}V\boldsymbol{c}$$

行空间和列空间

若A为一个`m*n`矩阵，由A的行向量张成的$R^{1\times n}$的子空间称为A的**行空间。**同理可定义A的**列空间**

- 行等价矩阵的行空间相同
- 行空间的维数即为矩阵的秩

线性方程组的相容性

相容即线性方程组的各个方程可以同时成立

- 一个线性方程组Ax = b相容的**充要条件**是b在A的列空间
- 若A为一`m*n`矩阵，当且仅当A的列向量张成R^m时，对每一个$b \in R^m$，线性方程组Ax=b是相容的。当且仅当A的列向量线性无关时，对每一$b \in R^m$，方程组Ax=b至多有一个解

证明见最后

注意这里张成R^m显然蕴含了$n \geq m$的条件

列空间

- 若A为一`m*n`矩阵，则其行空间的维数等于列空间的维数

线性变换

定义

一个将向量空间V映射到向量空间W的映射L，若对于所有的$v_1, v_2 \in V$，及所有标量a和b，有

$$L(av_1 + bv_2) = aL(v_1) + bL(v_2)$$

称其为线性变换，记为$L : V \rightarrow W$

若W就是V，则称L为V上的线性算子

证明

要证明L为线性变换，只需证明

- $L(v_1 + v_2) = L(v_1) + L(v_2)$
- $L(av_1) = aL(v_1)$

性质

若L为一从向量空间V到W的线性变换，则

- $L(\mathbf{0}_V) = \mathbf{0}_W$ 即线性变换后0向量仍为0向量
- $L(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1L(v_1) + a_2L(v_2) + \cdots + a_nL(v_n)$
- 对于$v \in V, L(-v) = -L(v)$

象与核

$L : V \rightarrow W$为一线性变换

核

$$ker(L) = \{v \in V | L(v) = 0_w\}$$

即变换后为0向量的所有向量组成的集合

象

$$L(S) = \{w \in W | w = L(v), \text{对某个} v \in S\}$$

即子空间S的所有向量经过线性变换后的向量的集合

整个向量空间的象L(V)称为L的**值域**

几个性质

令S为V的任意子空间，有

- ker(L)为V的一个子空间
- L(S)为W的一个子空间

证明两者对标量乘法和加法的封闭性即可

线性变换的矩阵表示

- 若L为一个从R^n到R^m的线性变换L，存在一个`m*n`矩阵A，使得
$$L(x) = Ax$$
  且A的第j个列向量为
$$a_j = L(e_j)$$

对于向量空间V和W基为任意基的情况下，也有类似定理

- 若E=[v1, ... , vn]和F=[w1, ... , wn]分别为向量空间V和W的有序基，则对每一线性变换$L : V \to W$，存在一个`m*n`矩阵A，使得对每一个向量v，有
$$[L(v)]_F = A[v]_E$$
  其中$[v]_E$表示在以E为基底情况下的坐标，则
$$a_j = [L(v_j)]_F$$

下面给出一个定理和其应用

- 令E=[u1, ... , un]和F=[b1, ... , bm]分别为R^n和R^m的有序基，若$L : R^n \to R^m$为一线性变换，且A为L相应于E和F的表示矩阵，则
$$a_j = B^{-1}L(u_j)$$
  其中$B = (b_1, \cdots, b_n)$

因此可以使用增广矩阵来计算A，即

$$\left[\, b_1 \quad \cdots \quad b_n \mid L(u_1) \quad \cdots \quad L(u_n) \,\right] = [\, I \mid A \,]$$

相似性

若B为L相应于[u1, ... , un]的表示矩阵，A为L相应于[e1, ... , en]的表示矩阵，U为从[u1, ... , un]到[e1, ... , en]的转移矩阵，则

$$B = U^{-1}AU$$

几个常见线性变换

R^2上的几个常见变换

拉伸

$$L(\boldsymbol{x}) = a\boldsymbol{x}$$

投影

如投影到x1轴

$$L(\boldsymbol{x}) = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} \boldsymbol{x}$$

轴对称

如对x1轴做轴对称

$$L(\boldsymbol{x}) = \begin{bmatrix} x_1 \\ -x_2 \end{bmatrix} \boldsymbol{x}$$

正交性

标量积

标量积

定义x和y的标量积为

$$x^T y = x_1 y_1 + \cdots + x_n y_n$$

欧几里得长度

$$||x|| = \sqrt{x^T x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

向量夹角

$$cos\theta = \frac{x^T y}{||x||\ ||y||}$$

柯西-施瓦茨不等式

若x和y为两个向量，则

$$|x^T y| \le ||x||\ ||y||$$

正交

若$x^T y = 0$，则称向量x和y正交，记为$x \perp y$

投影

标量投影

设y的单位向量为u，则x到y的投影为au，其中a为标量，称为x到y的标量投影

$$a = ||x||cos\theta = \frac{||x||\ ||y||cos\theta}{||y||} = \frac{x^T y}{||y||}$$

向量投影

由上述假设，则x在y上的投影为

$$p = au = a\frac{y}{||y||} = \frac{x^T y}{y^T y} y$$

正交子空间

定义

设X和Y为R^n的两个子空间，若对每一个$x \in X$及$y \in Y$都有$x^T y = 0$，则称X和Y为正交的。记为$X \perp Y$

正交补

$$
\text{令}Y\text{为}R^n\text{的子空间，}R^n\text{中所有与}Y\text{中的每一向量正交的向量集合记为}Y^\perp\text{，则} \\
Y^\perp = \{x \in R^n | x^T y = 0, y \in Y\}
$$

称为Y的正交补

基本子空间

对于`m*n`矩阵A，向量$b \in R^m$在A的列空间的充要条件是对于某$x \in R^n$，$b = Ax$。若将A看成是将R^n映射R^m的线性变换，则A的列空间和A的值域是相同的。记A的值域为R(A)，则

$$
\begin{aligned}
R(A) &= \{b \in R^m | b = Ax\text{，对某}x \in R^n\} \\
&= A\text{的列空间}
\end{aligned}
$$

同理，A的转置的列空间有

$$
\begin{aligned}
R(A^T) &= \{y \in R^n | y = A^T x\text{，对某}x \in R^m\} \\
&= A\text{的行空间}
\end{aligned}
$$

此外，由于对于N(A)中每个元素x，有Ax = 0，所以$R(A^T) \perp N(A)$

基本子空间定理

- 若A为一`m*n`矩阵，则$N(A) = R(A^T)^\perp, N(A^T) = R(A)^\perp$

其他几个定理

书P192

- $$\text{若}S\text{为}R^n\text{的一个子空间，则}dimS + dimS^\perp = n\text{。} \\ \text{此外若}\{x_1, \cdots, x_r\}\text{为}S\text{的一组基，且}\{x_{r+1}, \cdots, x_n\}\text{为}S^\perp\text{的一组基，则}\{x_1, \cdots, x_r, x_{r+1}, \cdots, x_n\}\text{为}R^n\text{的一组基}$$
- $$\text{若}S\text{为}R^n\text{的一个子空间，则} \\ R^n = S \oplus S^\perp$$
- $$\text{若}S\text{为}R^n\text{的一个子空间，则}(S^\perp)^\perp = S$$

由最后一个定理可得，若集合A为B的正交补，则B也为A的正交补

且由后两个定理和基本子空间定理，有一个重要的结论：

$$
\text{对于}R^n\text{中的任意向量}x\text{，设}A\text{为}m \times n\text{矩阵，} \\
\text{因为}R(A^T) = N(A)^\perp\text{，所以}R^n\text{中的任意向量}x\text{必可由}R(A^T)\text{中的一组基和}N(A)\text{中的一组基共同表示，即：} \\
x = y + z,\ \ y \in R(A^T),\ \ z \in N(A)
$$

此外，从上面结论继续推导，可以得出，A可以确定一个$R(A)$和$R(A^T)$间的一一映射，因为

$$
x = y + z,\ \ y \in R(A^T),\ \ z \in N(A)\text{，则} \\
Ax = Ay + Az = Ay\text{，所以} \\
R(A) = \{Ax | x \in R^n\} = \{Ay | y \in R(A^T)\}
$$

最小二乘问题

定义

最小二乘问题可化为一个超定的线性方程组，即方程数多于变量数。对于`m*n`矩阵A组成的方程组Ax=b，其中m>n，不能期望找到一个向量x使得Ax=b，但可以寻找一个向量使其最接近b。可以具体描述如下

构造一个残差

$$r(x) = b - Ax$$

距离为

$$||b - Ax|| = ||r(x)||$$

希望找到一个向量x，使得残差r(x)最小，也即最小化||r(x)||

几个定理

- $$\text{令}S\text{为}R^m\text{的一个子空间，对每一}b \in R^m\text{，在}S\text{中存在一个唯一元素}p\text{和}b\text{最接近，即对任意}y \neq p\text{，有}$$ $$||b - y|| > ||b - p||$$ $$\text{且}p\text{和}b\text{最接近的充要条件是}b - p \in S^{\perp}$$
-

相似矩阵

方阵的特征值和特征向量

定义

设A是n阶方阵，如果存在**数**λ和n维非零列向量x，使得

$$Ax = \lambda x$$

则称数λ为矩阵A的**特征值**，非零列向量x称为矩阵A的属于特征值λ的**特征向量**

注意：零向量不能作为特征向量，但特征值可以为0。所以降秩矩阵必有特征值0，因为降秩矩阵R(A)<n，|A|=0，所以齐次线性方程组$Ax = 0$存在非零解a，所以$Aa = 0a$，

特征方程与特征多项式

定义

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

这里将方程化为了一个齐次线性方程，因此若有λ使得x有非零解，则数λ是A的一个特征值。由齐次线性方程组有非零解的充要条件可知

$$|\lambda E - A| = 0$$

展开得

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

$$= (\lambda - a_{11})B_{11} + -a_{12}B_{12} + \cdots + -a_{1n}B_{1n}$$

$$= \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

称$|\lambda E - A|$为A的**特征多项式**，方程$|\lambda E - A| = 0$为**特征方程**

由特征多项式可知，特征多项式可以有n个根，若λ为A的单根，则称其为A的**单特征值。**若λ为A的k重根，则称其为A的**k重特征值**，k为λ的**代数重数**

解法

- 计算特征多项式$\lambda E - A$
- 求$|\lambda E - A| = 0$的所有根$\lambda_1, \lambda_2, \cdots, \lambda_n$，即A的全部特征值
- 对每个特征值λ_i求齐次线性方程组$(\lambda_i E - A)x = 0$的一个基础解系

一些问题的解释和证明

矩阵的秩

初等变换

设A经过初等变换后变为B，R(A)=r，D是B的任意r+1阶子式，|D|=0

证明$R(A) \geq R(B)$

交换r

交换i1,i2行

- 若D本身不含i1和i2行，显然|D|不变
- 若D既含i1又含i2行，显然|D|还是不变（相当于行列式内的初等变换）
- 若D含i1不含i2，则因为对于A来说，必有r+1阶子式D2与D相同

  例如D由B的第1 2 3行和第4 5 6列构成，A到B的变换为交换了第4列和第9列，则D相当于A的由1 2 3行和5 6 9列构成的子式。由于该子式仍是A的r+1阶子式，因此|D|=0

kr

显然对于|D|=0，某行乘常数不改变结果

r+kr

$r_i + kr_j$

- D不含第i行，显然无影响，|D|=0
- D含i和j行，相当于行列式内初等变换，|D|=0
- D含i行不含j行，则如下，显然分解出的两个行列式都是A的r+1阶子式，所以|D|=0

$$
D=\begin{vmatrix} \vdots \\ r_i+kr_j \\ \vdots \end{vmatrix}=\begin{vmatrix} \vdots \\ r_i \\ \vdots \end{vmatrix}+k\begin{vmatrix} \vdots \\ r_j \\ \vdots \end{vmatrix}
$$

上述过程可以证明$R(A)\geq R(B)$，因为初等变换可逆，因此也有$R(A)\leq R(B)$，综上，$R(A)=R(B)$

推论

因为A乘所有可逆矩阵都可视作初等变换，因此对于可逆矩阵P Q，有

$$
R(A)=R(PA)=R(AQ)=R(PAQ)
$$

性质

几个不等式

设A是`m*n`矩阵，B是`n*s`矩阵

- $max\{R(A),R(B)\}\leq R(A,B)\leq R(A)+R(B)$
  - $max\{R(A),R(B)\}\leq R(A,B)$ 因为A，B的所有非零子式均为(A,B)的非零子式，所以必有
  - $R(A,B)\leq R(A)+R(B)$ 因为对于A必存在一个P，使得其与行阶梯型矩阵U有关系 $A=P_1U_1$，对于B同样。而由矩阵的秩的定义可得，$R(A)=R(U_1)$，因此 $R(A,B)=R(\begin{pmatrix} P_1 & P_2 \end{pmatrix}\begin{pmatrix} U_1 & O \\ O & U_2 \end{pmatrix})=R(\begin{pmatrix} U_1 & O \\ O & U_2 \end{pmatrix})\leq R(U_1)+R(U_2)=R(A)+R(B)$
- $R(A+B)\leq R(A)+R(B)$

  将(A+B)组合为分块矩阵(A+B B)，假设A B均为n阶矩阵，则使1~n列减去n+1~2n列，有(A+B B)=(A B)，因此

  $$
  R(A+B)\leq R(\begin{pmatrix} A+B & B \end{pmatrix})=R(\begin{pmatrix} A & B \end{pmatrix})\leq R(A)+R(B)
  $$
- $R(AB)\leq min\{R(A),R(B)\}$

  因为$\begin{pmatrix} A & AB \end{pmatrix}=\begin{pmatrix} A & O \end{pmatrix}\begin{pmatrix} E_n & B \\ O & E_s \end{pmatrix}$，又因为$\begin{vmatrix} E_n & B \\ O & E_s \end{vmatrix}\neq 0$，所以为可逆矩阵，由性质3，若P可逆，则`R(AP) = R(A)`，即

  $$
  R(\begin{pmatrix} A & AB \end{pmatrix})=R(\begin{pmatrix} A & O \end{pmatrix}\begin{pmatrix} E_n & B \\ O & E_s \end{pmatrix})=R(\begin{pmatrix} A & O \end{pmatrix})=R(A)
  $$

  又因为AB是(A AB)的子矩阵，因此$R(AB)\leq R(\begin{pmatrix} A & AB \end{pmatrix})$，综上，$R(AB)\leq R(A)$

西尔维斯特不等式

$R(AB)\geq R(A)+R(B)-n$

将B分解为行最简式，有

$$
B=P\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q
$$

令

$$
AB=AP\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q=\begin{pmatrix} P_1 & P_2 \end{pmatrix}\begin{pmatrix} E_r & O \\ O & O \end{pmatrix}Q=\begin{pmatrix} P_1 & O \end{pmatrix}Q
$$

所以

$$R(AB)=R((\begin{matrix}P_1 & O\end{matrix})Q)=R((\begin{matrix}P_1 & O\end{matrix}))=R(P_1)$$

由上面的分析，因为B为`n*s`矩阵，因为这里的P Q都是可逆矩阵，所以P为`n*n`矩阵，Q为`s*s`矩阵

因为A为`m*n`矩阵，所以AP为`m*n`矩阵，则P1为`m*r`矩阵，P2为`m*(n-r)`矩阵

由性质1，$R(P_2)\leq min\{m,n-r\}$，所以必有$R(P_2)\leq n-r$

则

$$R(A)=R(AP)=R((\begin{matrix}P_1 & P_2\end{matrix}))\leq R(P_1)+R(P_2)=R(P_1)+n-r$$

所以

$$R(P_1)\geq R(A)+r-n$$

由B的分解式得，R(B) = r，所以

$$R(AB)=R(P_1)\geq R(A)+r-n=R(A)+R(B)-n$$

得证

线性方程组解的判定

n元线性方程组

对于增广矩阵$\tilde{A}=(\begin{matrix}A & b\end{matrix})$，必可化为行阶梯型矩阵。设其化为下列行阶梯型矩阵

$$\begin{pmatrix}
1 & 0 & \cdots & 0 & b_{1r+1} & \cdots & b_{1n} & d_1\\
0 & 1 & \cdots & 0 & b_{2r+1} & \cdots & b_{2n} & d_2\\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & 1 & b_{rr+1} & \cdots & b_{rn} & d_r\\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & d_{r+1}\\
\vdots & \vdots & & \vdots & \vdots & & \vdots & \vdots\\
0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0
\end{pmatrix}$$

则结果为

$$\begin{cases}
x_1+b_{1r+1}x_{r+1}+\cdots b_{1n}x_n=d_1\\
x_2+b_{2r+1}x_{r+1}+\cdots b_{2n}x_n=d_2\\
\cdots\\
x_r+b_{rr+1}x_{r+1}+\cdots b_{rn}x_n=d_r\\
0=d_{r+1}
\end{cases}$$

则若$d_{r+1}\neq 0$，则原方程组无解。所以当$R(A)\neq R(\tilde{A})$时，原方程无解

若r=n，即增广矩阵没有零行，则

$$\begin{cases}
x_1=d_1\\
x_2=d_2\\
\cdots\\
x_n=d_n
\end{cases}$$

若r<n，则

$$\begin{cases} x_1 + b_{1r+1}x_{r+1} + \cdots + b_{1n}x_n = d_1 \\ x_2 + b_{2r+1}x_{r+1} + \cdots + b_{2n}x_n = d_2 \\ \cdots \\ x_r + b_{rr+1}x_{r+1} + \cdots + b_{rn}x_n = d_r \end{cases}$$

即

$$\begin{cases} x_1 = d_1 - b_{1r+1}x_{r+1} - \cdots - b_{1n}x_n \\ x_2 = d_2 - b_{2r+1}x_{r+1} - \cdots - b_{2n}x_n \\ \cdots \\ x_r = d_r - b_{rr+1}x_{r+1} - \cdots - b_{rn}x_n \end{cases}$$

推论

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

n元齐次线性方程组

增广矩阵为(A, 0)

若r=n，则可以初等变换化为

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

即方程可化为

$$\begin{cases} x_1 \qquad = 0 \\ \quad x_2 \qquad = 0 \\ \cdots \\ \qquad x_r = 0 \end{cases}$$

因此对于n元齐次线性方程组，r<n时才有非零解

分块矩阵的秩

定理1 2

$$R(\begin{pmatrix} A & O \\ O & B \end{pmatrix}) = R(A) + R(B)$$

设R(A)=r，R(B)=s，则

$$R(\begin{pmatrix} A & O \\ O & B \end{pmatrix}) = R(\begin{pmatrix} E_r & O & O & O \\ O & O & O & O \\ O & O & E_s & O \\ O & O & O & O \end{pmatrix}) = R(\begin{pmatrix} E_r & O & O & O \\ O & E_s & O & O \\ O & O & O & O \\ O & O & O & O \end{pmatrix}) = R(A) + R(B)$$

由上面的证明，可以直观地证明第一个不等式

$$R(\begin{pmatrix} A & C \\ O & B \end{pmatrix}) = R(A) + R(B)$$

因为若C的秩大于A的秩，显然分块矩阵的秩大于$R(\begin{pmatrix} A & O \\ O & B \end{pmatrix})$

弗罗贝尼乌斯不等式

$$R(AB)+R(BC)\leq R(B)+R(ABC)$$

因为

$$\begin{pmatrix} B & O \\ O & ABC \end{pmatrix} \overset{r_2+Ar_1}{=} \begin{pmatrix} B & O \\ AB & ABC \end{pmatrix} \overset{c_2-Cc_1}{=} \begin{pmatrix} B & -BC \\ AB & O \end{pmatrix} \overset{-c_2}{=} \begin{pmatrix} B & BC \\ AB & O \end{pmatrix} = \begin{pmatrix} AB & O \\ B & BC \end{pmatrix}$$

所以

$$R(B)+R(ABC)=R(\begin{pmatrix} B & O \\ O & ABC \end{pmatrix})=R(\begin{pmatrix} AB & O \\ B & BC \end{pmatrix})\geq R(AB)+R(BC)$$

向量空间

向量空间的公理

其他性质

$$\begin{aligned} &1.\text{因为}\boldsymbol{x}=1\boldsymbol{x}=(1+0)\boldsymbol{x}=\boldsymbol{x}+0\boldsymbol{x}, \\ &\quad\text{所以}\boldsymbol{x}+(-\boldsymbol{x})=\boldsymbol{x}+(-\boldsymbol{x}-0\boldsymbol{x})=\boldsymbol{x}+(-\boldsymbol{x})-0\boldsymbol{x}=\mathbf{0}-0x=\mathbf{0} \\ &2.\boldsymbol{x}+\boldsymbol{y}=\mathbf{0}=\boldsymbol{x}+(-\boldsymbol{x}), \\ &\quad\text{所以}\boldsymbol{y}=-\boldsymbol{x} \\ &3.\mathbf{0}=0\boldsymbol{x}=[1+(-1)]\boldsymbol{x}=\boldsymbol{x}+(-1)\boldsymbol{x}, \\ &\quad\text{又因为}\mathbf{0}=\boldsymbol{x}+(-\boldsymbol{x})\text{，所以}(-1)\boldsymbol{x}=-\boldsymbol{x} \end{aligned}$$

子空间

向量集合的张成

- 若v1, v2, ... , vn为向量空间V中的元素，则Span(v1, v2, ..., vn)为V的一个子空间

先证明标量乘法，令b为一标量，v=a1 v1+a2 v2+...+an vn为Span(v1, v2, ... , vn)中的任意一个元素，由于

$$b\boldsymbol{v}=b(a_1\boldsymbol{v}_1+\cdots+a_n\boldsymbol{v}_n)=(ba_1)\boldsymbol{v}_1+\cdots+(ba_n)\boldsymbol{v}_n$$

因此bv也在Span(v1, v2, ... , vn)中

再证明加法，v = a1 v1 + ... + an vn， w = b1 v1 + ... + bn vn

$$\begin{aligned} \boldsymbol{v}+\boldsymbol{w}&=a_1\boldsymbol{v}_1+\cdots+a_n\boldsymbol{v}_n+b_1\boldsymbol{v}_1+\cdots+b_n\boldsymbol{v}_n \\ &=(a_1+b_1)\boldsymbol{v}_1+\cdots+(a_n+b_n)\boldsymbol{v}_n \end{aligned}$$

因此v+w也在Span(v1, v2, ... , vn)中

基和维数

坐标变换

变换过程如下，设

$$\begin{aligned} x&=x_1e_1+x_2e_2 \\ u_1&=a_{11}e_1+a_{21}e_2 \\ u_2&=a_{12}e_1+a_{22}e_2 \end{aligned}$$

（注意这里a的下标与推导后的结果构成的矩阵有关）

此外

$$x = c_1 u_1 + c_2 u_2$$

则

$$\begin{aligned} x &= c_1(a_{11}e_1 + a_{21}e_2) + c_2(a_{12}e_1 + a_{22}e_2) \\ &= (a_{11}c_1 + a_{12}c_2)e_1 + (a_{21}c_1 + a_{22}c_2)e_2 \\ &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} \\ &= \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

这里左边的矩阵即为[u1, u2]到[e1, e2]的**转移矩阵**。而我们要求[e1, e2]到[u1, u2]的转移矩阵，只需要求逆即可，即

$$\begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

行空间和列空间

线性方程的相容性

- 一个线性方程组Ax = b相容的**充要条件**是b在A的列空间
- 若A为一`m*n`矩阵，当且仅当A的列向量张成R^m时，对每一个$b \in R^m$，线性方程组Ax=b是相容的。当且仅当A的列向量线性无关时，对每一$b \in R^m$，方程组Ax=b至多有一个解

1. 显然，因为若b不在A的列空间中，则说明b不能由A的列向量线性组合而成，也意味着Ax=b的增广矩阵秩大于n
2. 前半部分可以由上一个定理得出。

   第二部分充分性：若A的列向量线性无关，假设此时有两个解x1和x2，有$Ax_1 = b$，$Ax_2 = b$，则$A(x_1 - x_2) = 0$。又因为列向量线性无关，因此方程Ax=0只有平凡解（x=0，注意这里都是向量）。因此$x_1 = x_2$

   第二部分必要性：若Ax=b至多只有一个解，则Ax=0只有平凡解，因此可以推出线性无关

列空间

- 若A为一`m*n`矩阵，则其行空间的维数等于列空间的维数

书P142

若A的行阶梯型为U，则U有r个首1元素，U中对应首1元素的列必定线性无关，但其不构成A的列空间的基，因为

线性变换

线性变换的矩阵表示

- 若L为一个从R^n到R^m的线性变换L，存在一个`m*n`矩阵A，使得

  $$L(x) = Ax$$

  且A的第j个列向量为

  $$a_j = L(e_j)$$

$$
\begin{gathered}
\text{设}x = x_1 e_1 + \cdots + x_n e_n\text{，则} \\
L(x) = x_1 L(e_1) + \cdots + x_n L(e_n) \\
\text{所以若令}a_j = L(e_j)\text{，}A = [a_1, \cdots, a_n] \\
\text{则就有}
\end{gathered}
$$

$$
\begin{aligned}
L(x) &= x_1 L(e_1) + \cdots + x_n L(e_n) \\
&= x_1 a_1 + \cdots + x_n a_n \\
&= \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= Ax
\end{aligned}
$$

$$\text{其中}a_j\text{为列向量}$$

正交性

正交子空间

基本子空间

- 基本子空间定理：若A为一`m*n`矩阵，则$N(A) = R(A^T)^\perp, N(A^T) = R(A)^\perp$

由于$N(A) \perp R(A^T)$，所以$N(A) \subset R(A^T)^\perp$。

又因为任取$x \in R(A^T)^\perp$，都有$A_i^T \perp x$，即$(A_i^T)^T x = 0$，即$A_i x = 0$。

所以$N(A) = R(A^T)^\perp$

最小二乘问题

几个定理

- $$\begin{gathered}\text{令}S\text{为}R^m\text{的一个子空间，对每一}b \in R^m\text{，在}S\text{中存在一个唯一元素}p\text{和}b\text{最接近，即对任意}y \neq p\text{，有} \\ ||b - y|| > ||b - p|| \\ \text{且}p\text{和}b\text{最接近的充要条件是}b - p \in S^\perp\end{gathered}$$

先证第一个不等式

$$
\begin{gathered}
R^m\text{中每个元素}b\text{都可表示为}b = p + z\text{，其中}p \in S, z \in S^\perp\text{，则若}y\text{为}S\text{中任意元素} \\
||b - y||^2 = ||(b - p) + (p - y)||^2 \\
\text{又因为}p - y \in S\text{，}b - p = z \in S^\perp\text{，则由两者正交，有} \\
||b - y||^2 = ||b - p||^2 + ||p - y||^2 \\
\text{所以由于}y \neq p \\
||b - y|| > ||b - p||
\end{gathered}
$$

再证充要条件

$$
\begin{gathered}
\text{由于上面不等式成立，假设对于}q\text{，有}b - q \notin S^\perp \\
\text{则由于}b = p + z\text{，且又因为}b - q \notin S^\perp\text{，则}b - q \in S\text{，即}q \in S
\end{gathered}
$$