课程准备知识

殷东生
dyin@math.tsinghua.edu.cn

清华大学数学科学系

2018年秋季学期

定义 (线性空间:linear space)

设$\mathbb{P}$是一个数域，一般取$\mathbb{P}=\mathbb{R},\mathbb{C}$，$V$是一个非空集合，在$V$上定义两种运算

- **加法**：$\forall u,v\in V$，有V中唯一的元素$(u+v)$与之对应，且

$$\begin{aligned} u+v &= v+u, \quad \forall u,v\in V, \\ u+(v+w) &= (u+v)+w, \quad \forall u,v,w\in V, \end{aligned}$$

  且V中有唯一的零元素θ使得

$$u+\theta=u,\forall u\in V.$$

  对每一个$u\in V$有唯一的负元素$-u$与之对应，使得

$$u+(-u)=\theta.$$

定义 (线性空间)

- **数乘**：$\forall \alpha \in \mathbb{P}, u \in V$，有$V$中的唯一的元素$\alpha u$与之对应，满足

$$
\begin{aligned}
1u &= u,\ \forall u \in V,\\
\alpha(\beta u) &= (\alpha\beta)u,\ \forall \alpha, \beta \in \mathbb{P}, u \in V,\\
\alpha(u+v) &= \alpha u + \alpha v,\\
(\alpha+\beta)u &= \alpha u + \beta u,\ \forall \alpha, \beta \in \mathbb{P}, u \in V,
\end{aligned}
$$

称V为数域$\mathbb{P}$上的线性空间(数域$\mathbb{P}$上的向量空间)。

注记 (子空间 subspace)

若线性空间 V 的一个子集 W 按照 V 的加法和数乘也是一个线性空间，则称 W 是 V 的线性子空间。

例 ($\mathbb{R}^n$和$\mathbb{C}^n$)

$\mathbb{R}^n$是n维实向量的全体，按照向量的加法以及实数与向量的乘法构成了实数域$\mathbb{R}$上的线性空间。通常，向量是指列向量，亦即 $\boldsymbol{x} \in \mathbb{R}^n$，则$\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T$，其中$x_1, x_2, \cdots, x_n \in \mathbb{R}$。$\mathbb{R}^n$中的零元素 就是零向量$\mathbf{0} = (0, 0, \cdots, 0)^T$。
类似的可定义$\mathbb{C}^n$。

例 (连续函数空间)

设 $[a, b]$ 为闭区间， 定义

$$C[a, b] := \{f | f \text{ 在 } [a, b] \text{ 上连续}\}$$

则 $C[a, b]$ 是一个线性空间。

例 (矩阵空间$\mathbb{R}^{m\times n}$与$\mathbb{C}^{m\times n}$)

$\mathbb{R}^{m\times n}$是所有m行n列实元素矩阵的全体

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix},$$

按矩阵的加法以及实数与矩阵的乘法构成了$\mathbb{R}$上的线性空间。它的零元素 是m行n列的零矩阵，其所有元素均为零。同理有$\mathbb{C}^{m\times n}$。

例

多项式空间

$$P_N := \{p(x)|p(x) = a_0 + a_1x + \cdots + a_Nx^N, a_i \in R(C)\}$$

线性独立

定义 (线性无关 linearly independent)

设 $(V,\mathbb{P})$是一个线性空间，$x_i \in V$, $i = 1, 2, \cdots, n$，若存在 $a_i \in \mathbb{P}$, $i = 1, 2, \cdots, n$，$\sum_{i=1}^{n} |a_i| > 0$，使得

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = 0,$$

则称 $\{x_i\}_{i=1}^{n}$ 是线性相关的，反之，则称线性无关的。

1. 在 P_N中，$\{1, x, \cdots, x^N\}$ 是线性无关的。
2. 在 $C[-\pi, \pi]$中，$\{1, \cos x, \sin x, \cdots, \cos nx, \sin nx\}$ 是线性无关的。

基和维数

若 $\{x_i\}_{i=1}^n$ 线性无关，且 $\forall x \in V$，有 $\{a_i\}_{i=1}^n$使得

$$x = a_1x_1 + a_2x_2 \cdots + a_nx_n,$$

则 $\{x_i\}_{i=1}^n$ 构成 S的一组基(basis)，空间的维数为 n。

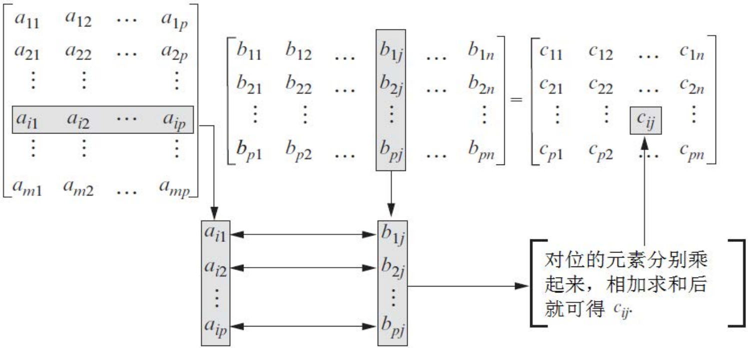
注记

线性空间的维数与基的选取无关，它是空间的内在性质。

1. P_N中 $\{1, x, \cdots, x^N\}$ 是一组基。$\dim P_N = N + 1$
2. $C[a,b]$中 $\forall N$, $\{1, x, \cdots, x^N\}$ 是线性无关的，$\dim C[a,b] = +\infty$。

存在$\mathbb{P}^{m\times p}\times\mathbb{P}^{p\times n}\to\mathbb{P}^{m\times n}$的一个运算，称为 矩阵乘法

$$AB=C,\ c_{ij}=\sum_{1\le k\le p}a_{ik}b_{kj}$$



Q：矩阵乘法的计算量？

定义 (迹)

设 $A=(a_{ij})_{i,j=1}^{n}\in\mathbb{C}^{n\times n}$，矩阵 A 的迹 *(trace)* 是其所有对角元元素之和。

$$trace(A)=\sum_{i=1}^{n}a_{ii}$$

定理

若 $A\in\mathbb{C}^{n\times n}$， $B\in\mathbb{C}^{n\times n}$，则

$$trace(AB)=trace(BA).$$

Matlab 求迹的命令为 trace

```
>> A = [1 5 1;2 -1 6;1 0 3]

A =
     1     5     1
     2    -1     6
     1     0     3

>> B = [2 3 0;3 -1 7;4 8 9]

B =
     2     3     0
     3    -1     7
     4     8     9

>> trace(A + B)

ans =
    13

>> 7*trace(A + B)

ans =
    91
```

定义 (非奇异矩阵 nonsingular matrix)

矩阵 $A \in \mathbb{R}^{n\times n}$ 称为非奇异的或可逆的*(invertible)*，若存在一个矩阵 $B \in \mathbb{R}^{n\times n}$使得

$$AB = BA = I.$$

其中 B 称为A 的逆矩阵*(inverse)*。若 A 的逆矩阵不存在，则称其是奇异的*(singular)*。

注记

若 $A_1, A_2, \cdots, A_m \in \mathbb{R}^{n\times n}$ 都非奇异，则$A_1A_2\cdots A_m$亦是非奇异的，且

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

定理 (线性方程组的可解性)

若矩阵$A \in \mathbb{R}^{n\times n}$ 可逆，则方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 有唯一解$\boldsymbol{x} = A^{-1}\boldsymbol{b}$。

注记

$A \in \mathbb{R}^{n\times n}$ 是非奇异的$\Longleftrightarrow$ 齐次方程 $A\boldsymbol{x} = \boldsymbol{0}$只有 $\boldsymbol{0}$解，亦称为平凡解。

Matlab 求逆矩阵的命令为 inv(A)。

```
>> format rational;
>> A = [1 3 -1; 4 1 6; 0 2 3]
A =
1 3 -1
4 1 6
0 2 3

>> A_inv = inv(A)
A_inv =
9/53   11/53 -19/53
12/53  -3/53 10/53
-8/53 2/53   11/53

>> B = [1 4 0; 3 5 1; 2 -7 8]
B =
1   4 0
3   5 1
2  -7 8

>> B_inv = inv(B)
B_inv =
-47/41 32/41 -4/41
22/41 -8/41   1/41
31/41 -15/41 7/41

>> inv(A*B)
ans =
-7/2173 -621/2173 1169/2173
94/2173 268/2173 -487/2173
43/2173 400/2173 -662/2173

>> B_inv * A_inv
ans =
-7/2173 -621/2173 1169/2173
94/2173 268/2173 -487/2173
43/2173 400/2173 -662/2173
```

定理

1. A非奇异当且仅当 $\det A \neq 0$;
2. A奇异当且仅当 $\det A = 0$;
3. 齐次方程 $A\boldsymbol{x} = \boldsymbol{0}$ 有非平凡解当且仅当 $\det A = 0$。

Matlab 求行列式的命令为 det(A)，而求方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 的命令为 “\”，用 $A\backslash\boldsymbol{b}$来得到 $\boldsymbol{x}$。

```
>> A = [1 6 25;16 32 19;56 53 5];
>> det(A)

ans =
     -18543

>> A\[1 2 5]'

ans =
    0.0910
   -0.0053
    0.0376

>> B = [1 3 2;5 14 7;2 5 1];
>> det(B)

ans =
 -3.3307e-015

>> size(null(B),2) % compute the nullity of B

ans =
     1

>> B\[1 2 5]'
Warning: Matrix is close to singular or badly scaled.
         Results may be inaccurate. RCOND = 4.587698e-018.

ans =
  1.0e+016 *

    1.2610
   -0.5404
    0.1801
```

定义 (矩阵的转置)

设 $A \in \mathbb{R}^{m\times n}$，它的转置 *(transpose)* A^T 可通过交换 A 的行和列得到，亦即

$$A = [a_{ij}] \Longrightarrow (A^T)_{ij} = a_{ji}, \quad A^T \in \mathbb{R}^{n\times m}$$

例

$$A = \begin{pmatrix} 1 & 9 & 0 \\ 3 & 7 & 15 \\ 4 & 8 & 1 \\ -7 & 12 & 3 \end{pmatrix} \qquad A^T = \begin{pmatrix} 1 & 3 & 4 & -7 \\ 9 & 7 & 8 & 12 \\ 0 & 15 & 1 & 3 \end{pmatrix}$$

$$(A^T)^T = A, \quad (AB)^T = B^T A^T, \quad (A^T)^{-1} = (A^{-1})^T$$

定义 (对称矩阵)

矩阵 $A \in \mathbb{R}^{n\times n}$ 称为**对称的**(*symmetric*)若 $A^T = A$。即

$$a_{ij} = a_{ji}, \forall 1 \leq i,j \leq n.$$

亦即交换 A 的行列，A不变。

推论

若$A \in \mathbb{R}^{m\times n}$，则 $A^TA \in \mathbb{R}^{n\times n}$是对称矩阵。

Matlab 求矩阵的转置命令为 ′

```
>> A = [1 8 -1; 3 -9 15; -1 5 3]

A =
     1     8    -1
     3    -9    15
    -1     5     3

>> A_TA = A'*A

A_TA =
    11   -24    41
   -24   170  -128
    41  -128   235

>> A_TA - A_TA'

ans =
     0     0     0
     0     0     0
     0     0     0
```

对于方阵$A \in \mathbb{C}^{n\times n}$，若有$\lambda \in \mathbb{C}$和非零向量$\boldsymbol{x} \in \mathbb{C}^n$，使得

$$A\boldsymbol{x} = \lambda \boldsymbol{x}$$

则称$\boldsymbol{x}$为特征值(eigenvalue)λ对应的特征向量(eigenvector)。

定义 (特征值)

$A \in \mathbb{R}^{n\times n}$，多项式 $p(\lambda) = \det(\lambda I - A)$ 称为 A 的特征多项式*(characteristic polynomial)*，方程 $p(\lambda) = 0$ 称为特征方程*(characteristic equation)*，若λ 是 p 的根，则 λ是 A 的特征值，若 $\boldsymbol{v}$ 是满足 $A\boldsymbol{v} = \lambda\boldsymbol{v}$ 的非零列向量，则 $\boldsymbol{v}$ 是 A 的一个特征向量。

问题

上述特征值的定义是否可以用来数值求解特征值？

定义 (谱)

对任一$A \in \mathbb{C}^{n\times n}$，其有$n$个特征值（可以相同）。全体特征值的集合称为$A$的 谱(*spectrum*)，记作 $\sigma(A)$，即

$$\sigma(A) \overset{def}{=} \{\lambda_1, \cdots, \lambda_n\}$$

$$\rho(A) \overset{def}{=} \max_{\lambda\in\sigma(A)} |\lambda|$$

称为A的谱半径(*spectrum radius*)。

$$\begin{aligned} \text{trace}(A) &= \sum_{1\le i\le n} \lambda_i, \\ \det A &= \lambda_1\lambda_2\cdots\lambda_n \end{aligned}$$

定义 (可对角化)

设$A \in \mathbb{P}^{n\times n}$，若存在非奇异阵$P \in \mathbb{P}^{n\times n}$，使得

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

为对角阵，则称A可对角化(*diagonalizable*)。A可对角化的充要条件是它有n个线性无关的特征向量。

若A是实对称矩阵，则A的特征值都是实数，可以对角化。

定理

矩阵$A \in \mathbb{R}^{n\times n}$的不同特征值对应的特征向量是线性无关的。若$A$有$n$个不同的特征值，则$A$可对角化。

若A具有重特征值，即特征方程有重根，则

$$\det(\lambda I - A) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2} \cdots (\lambda - \lambda_s)^{n_s}, \lambda_i \neq \lambda_j, \text{if } i \neq j,$$

亦即λ_i 是特征方程的n_i重根。n_i称为特征值λ_i的代数重数。

$$n_1 + n_2 + \cdots + n_s = n, \;\; n_i \geq 1, 1 \leq i \leq s.$$

设λ_i对应的最大线性无关特征向量的个数为m_i，则m_i就是齐次方程组$(\lambda_i I - A)\boldsymbol{x} = 0$的基础解系所包含最大线性无关解的个数，$m_i$称为特征值$\lambda_i$ 的几何重数。

一般地

$$m_i \leq n_i, \;\; i = 1, 2, \cdots, s.$$

定理

设A具有重特征值，则A可对角化的充要条件为每个特征值的几何重数和代数重数相等。

Matlab 中求特征值和特征向量的命令为 eig(A)，若要得到特征值和其对应的特征向量可用

»[V D]=eig(A)

D 是由特征值组成的对角阵，V 的列向量是特征值对应的特征向量。如求 B 的特征对。

$$B = \begin{pmatrix} 1 & 6 & 3 \\ -1 & 4 & 9 \\ 12 & 35 & 1 \end{pmatrix}$$

```
>> [V E] = eig(B)

V =
    0.0118    0.9119    0.2500
    0.4211   -0.3220    0.4278
   -0.9069    0.2545    0.8686

E =
  -15.4092         0         0
         0   -0.2812         0
         0         0   21.6905
>> eig(B)

ans =
  -15.4092
   -0.2812
   21.6905
```

定理 (Jordan标准型)

任一复方阵可以通过相似变换化为Jordan标准型J，

$$P^{-1}AP = J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_p \end{pmatrix}$$

每个J_i对应一个特征值λ_i，它是r_i个小块组成的块对角阵

$$J_i = \begin{pmatrix} J_{i1} & & & \\ & J_{i2} & & \\ & & \ddots & \\ & & & J_{ir_i} \end{pmatrix}, J_{ik} = \begin{pmatrix} \lambda_i & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_i & 1 \\ & & & \lambda_i \end{pmatrix}$$

内积空间

定义 (内积(inner product))

设 V 是数域 $\mathbb{P}$ 上的线性空间，内积 $(\cdot,\cdot): V\times V\to\mathbb{P}$，对于$\forall u,v\in V,\exists!(u,v)\in\mathbb{P}$与之对应，且有：

1. $(u+v,w)=(u,w)+(v,w)$，$\forall u,v,w\in V$;
2. $(\alpha u,v)=\alpha(u,v)$，$\forall u,v\in V,\alpha\in\mathbb{P}$;
3. $(u,v)=\overline{(v,u)}$，$\forall u,v\in V$;
4. $(u,u)\geq 0$，$\forall u\in V$，且 $(u,u)=0\Leftrightarrow u=\theta$.

则称 (u,v) 是 u 和 v 的内积，定义了内积的空间称为内积空间。

例 ($\mathbb{R}^n$和 $\mathbb{C}^n$的内积)

$\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$, 则内积 $(\boldsymbol{x},\boldsymbol{y})$定义为

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_iy_i=\boldsymbol{y}^T\boldsymbol{x}.$$

若 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{C}^n$, 则内积定义为

$$(\boldsymbol{x},\boldsymbol{y})=\sum_{i=1}^{n}x_i\bar{y}_i=\bar{\boldsymbol{y}}^T\boldsymbol{x}.$$

如果给定$\omega_i\in\mathbb{R},\omega_i>0,1\le i\le n$(权系数), 可定义一种带权内积:

$$(\boldsymbol{x},\boldsymbol{y})_\omega=\sum_{i=1}^{n}\omega_ix_iy_i$$

正交向量

定义 (正交(orthogonal))

若向量 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$ 满足$(\boldsymbol{x},\boldsymbol{y})=0$，则称它们是正交的。两个向量的集合 X 和 Y，若每个 $\boldsymbol{x}\in X$ 都和每个 $\boldsymbol{y}\in Y$ 正交，则称 X 与 Y 正交。

S 是非零向量的集合，若 $\forall\boldsymbol{x},\boldsymbol{y}\in S,\boldsymbol{x}\neq\boldsymbol{y}\Rightarrow(\boldsymbol{x},\boldsymbol{y})=0$，则其为正交集合。

定理

正交集合 S 中的向量是线性无关的。

推论

若正交集合 $S\subseteq\mathbb{R}^n$ 有 n 个向量，则它是$\mathbb{R}^n$ 的一组基。

正交矩阵

定义 (正交矩阵(orthogonal matrix))

设$A \in \mathbb{R}^{n\times n}$，若满足

$$A^TA = I,$$

则称A为正交矩阵。

正交矩阵有如下的性质：

1. A不同的列向量相互正交，且各列向量的2-范数为1。
2. $A^{-1} = A^T$，且A^T也是正交矩阵。
3. $|\det A| = 1$。
4. 若A和B是同阶的正交矩阵，则AB和BA都是正交矩阵。

定义 (L^2 内积)

设$f(x), g(x) \in C[a,b]$，则它们的L^2 内积为

$$(f,g) = \int_a^b f(x)g(x)dx.$$

定义 (权函数(weight function))

若定在 $[a,b]$ 上的可积函数 $\rho(x)$ 满足

1. $\rho(x) \geq 0, \forall x \in [a,b]$;
2. 在 $[a,b]$ 的任一子区间上$\rho(x)$ 不恒为零。

称 $\rho(x)$ 为 $[a,b]$ 上的一个权函数。

利用权函数可定义带权L^2 内积：

$$(f,g)_\rho = \int_a^b \rho(x)f(x)g(x)dx$$

定理 (Cauchy-Schwarz不等式)

设V 是一个内积空间，则对任一的 u, v有

$$|(u, v)^2| \leq (u, u)(v, v),$$

等号当且仅当 u, v 线性相关时成立。

数学家简介 (柯西(Cauchy))

柯西*(Cauchy, 1789—1857)*是法国数学家、物理学家、天文学家。柯西在数学上的最大贡献是在微积分中引进了极限概念，并以极限为基础建立了逻辑清晰的分析体系。复变函数的微积分理论就是由他创立的。*1821*年柯西提出极限定义的方法，把极限过程用不等式来刻画，后经魏尔斯特拉斯改进，成为现在所说的柯西极限定义或叫$\epsilon-\delta$定义。当今所有微积分的教科书都还沿用着柯西等人关于极限、连续、导数、收敛等概念的定义。柯西对定积分作了最系统的开创性工作，他把定积分定义为和的“极限”。在定积分运算之前，强调必须确立积分的存在性。他利用中值定理首先严格证明了微积分基本定理。使数学分析的基本概念得到严格的论述。把微积分及其推广从对几何概念、运动和直观了解的完全依赖中解放出来，并使微积分发展成现代数学最基础最庞大的数学学科。

Faá di Bruno $\mapsto$ Peano $\mapsto$ Russell

Gram-Schmidt正交化方法

若$\{u_1, u_2, \cdots, u_n\}$ 是内积空间 V 中的一个线性无关元素系列，则

$$\begin{cases} v_1 = u_1, \\ v_i = u_i - \displaystyle\sum_{k=1}^{i-1} \frac{(u_i, v_k)}{(v_k, v_k)} v_k, \ i = 2, 3 \cdots \end{cases}$$

生成V 中的一个正交序列 $\{v_1, v_2, \cdots, v_n\}$ 是 $\mathrm{span}\{u_1, u_2, \cdots, u_n\}$ 的一组基。

注记

Gram-Schmidt 正交化以丹麦数学家 *Gram* 和德国数学家 *Schmidt*的名字命名，但是最先由*Laplace* 和 *Cauchy* 给出。

数学家简介 (Pierre-Simon Laplace)

拉普拉斯(1749—1827)是法国分析学家、概率论学家和物理学家，法国科学院院士。是天体力学的主要奠基人、天体演化学的创立者之一， 还是分析概率论的创始人，是应用数学的先驱。他发表的天文学、数学和物理学的论文有270多篇，专著合计有4006多页。其中最有代表性的专著有《天体力学》、 《宇宙体系论》（中译本1978年版）和《概率分析理论》（1812）。

Lapalce 方程 $-\Delta u = 0;$

Laplace 变换 $F[s] = \mathcal{L}[f(t)] = \int_0^{+\infty} f(t)e^{-st}dt.$

若要求归一化即 $\|q_i\|_2 = 1, 1 \le i \le n$，则由Gram-Schmidt 正交化可得 QR 分解。设

$$u_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix}$$

1. 归一化：$q_1 = \frac{1}{2}(1, -1, 1, 1)^T$。
2. 计算投影:$(u_2, q_1) = 2$，接着

$$\hat{q} = u_2 - (u_2, q_1)q_1 = \begin{pmatrix} 2 \\ 0 \\ 0 \\ 2 \end{pmatrix} - 2 \times \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

3. 归一化：$q_2 = \frac{1}{2}(1, 1, -1, 1)^T$。

经典的Gram-Schmidt正交化

算法 (Classical Gram-Schmidt)

1. *compute* $r_{11} := \|\boldsymbol{u}_1\|_2$. *If* $r_{11} = 0$ *stop, else compute* $\boldsymbol{q}_1 := \boldsymbol{u}_1 / r_{11}$
2. *For* $j = 2, \cdots, n$. *Do*
3. *Compute* $r_{ij} := (\boldsymbol{u}_j, \boldsymbol{q}_i)$ *for* $i = 1, 2, \cdots, j-1$
4. $\hat{\boldsymbol{q}} := \boldsymbol{u}_j - \sum_{i=1}^{j-1} r_{ij} \boldsymbol{q}_i$
5. $r_{jj} := \|\hat{q}\|_2$
6. *If* $r_{jj} = 0$ *then stop, else* $\boldsymbol{q}_j := \hat{\boldsymbol{q}} / r_{jj}$
7. *EndDo*

若$\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ 线性无关，则算法 n 步完成。

定义矩阵R的元素为

$$
\begin{aligned}
r_{ij} &= \begin{cases} (\boldsymbol{u}_j, \boldsymbol{q}_i), & i < j \\ \|\hat{\boldsymbol{q}}\|_2, & i = j \\ 0, & i > j \end{cases} \\
\Rightarrow \boldsymbol{u}_j &= r_{1j}\boldsymbol{q}_1 + r_{2j}\boldsymbol{q}_2 + \cdots + r_{jj}\boldsymbol{q}_j, \quad 1 \le j \le n.
\end{aligned}
$$

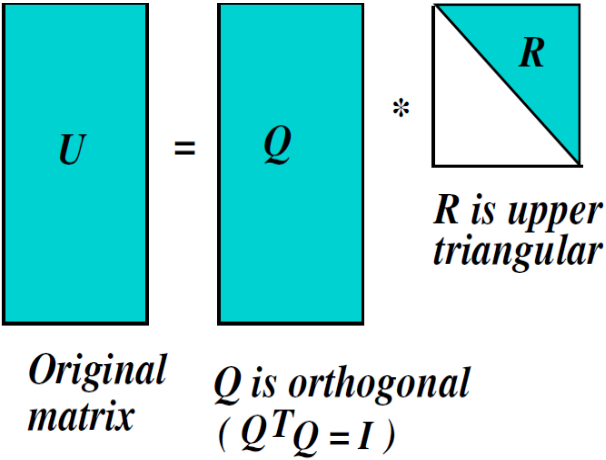
令 $U = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n]$，$Q = [\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n]$，且令 $R \in \mathbb{R}^{n\times n}$ 是上面定义的矩阵的上三角部分

$$R = \{r_{ij}\}_{i,j=1}^{n}.$$

则有 U 的QR分解(QR factorization)

$$U = QR$$

若 $U \in \mathbb{R}^{m\times n}$ 列满秩，则有



Matlab 的 Classical Gram-Schmidt 算法的程序和计算结果：

```
function [Q R]=cgs(A)
[m,n]=size(A);Q=zeros(m,n);R=zeros(n,m);%以使矩阵具有正确的大小。
for j=1:n %Gram-Schmidt正交化
v=A(:,j); %v初始化为A的第j列
for i=1:j-1
R(i,j)=Q(:,i)'*A(:,j); %为提高精确度用A(:,j)代替v
v=v-R(i,j)*Q(:,i); %减去投影
end %v现在和q1,…,qj 正交
R(j,j)=norm(v);
Q(:,j)=v/R(j,j); %将v规范化使之成为单位向量qj
end
```

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}, \quad \epsilon = 10^{-8}$$

```
>> epsilon = 1.0e-8;
A = [1 1 1;epsilon 0 0;0 epsilon 0;0 0 epsilon];
>> [Q R]=cgs(A)

Q =

    1.0000         0         0
    0.0000   -0.7071   -0.7071
         0    0.7071         0
         0         0    0.7071


R =

    1.0000    1.0000    1.0000         0
         0    0.0000         0         0
         0         0    0.0000         0

>> Q'*Q

ans =

    1.0000   -0.0000   -0.0000
   -0.0000    1.0000    0.5000
   -0.0000    0.5000    1.0000
```

改进的方法(modified Gram-Schmidt MGS)算法

算法 (MGS)

1. *Define* $r_{11} := \|\boldsymbol{u}_1\|_2$. *If* $r_{11} = 0$ *Stop, else* $\boldsymbol{q}_1 := \boldsymbol{u}_1 / r_{11}$
2. *For* $j = 2, \cdots, n$, *Do*
3. *Define* $\hat{\boldsymbol{q}} := \boldsymbol{u}_j$
4. *For* $i = 1, \cdots, j-1$, *Do*
5. $r_{ij} := (\hat{\boldsymbol{q}}, \boldsymbol{q}_i)$
6. $\hat{\boldsymbol{q}} := \hat{\boldsymbol{q}} - r_{ij}\boldsymbol{q}_i$
7. *EndDo*
8. *Compute* $r_{jj} := \|\hat{\boldsymbol{q}}\|_2$
9. *If* $r_{jj} = 0$ *then Stop, else* $\boldsymbol{q}_j := \hat{\boldsymbol{q}} / r_{jj}$
10. *EndDo*

Matlab 的修正的Gram-Schmidt 算法的程序和计算结果：

```
% Modified Gram-Schmidt method
function [Q,R] = mgs(A)
[m,n]=size(A);
Q=zeros(m,n); %矩阵的行和列
Q(1:m,1) = A(1:m,1);
R=zeros(n); R(1,1)=1;
for k = 1:n
R(k,k) = norm (A(1:m,k));
Q(1:m,k) = A(1:m,k)/R(k,k);%对本次得到的正交向量进行归一化
for j=k+1:n
R(k,j) = Q(1:m,k)'* A(1:m,j);
A(1:m,j) = A(1:m,j) - Q(1:m,k)*R(k,j);%对剩余向量进行修正
end
end
```

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \epsilon & 0 & 0 \\ 0 & \epsilon & 0 \\ 0 & 0 & \epsilon \end{bmatrix}, \quad \epsilon = 10^{-8}$$

```
>> epsilon = 1.0e-8;
A = [1 1 1;epsilon 0 0;0 epsilon 0;0 0 epsilon];
>> [Q R]=mgs(A)

Q =

    1.0000         0         0
    0.0000   -0.7071   -0.4082
         0    0.7071   -0.4082
         0         0    0.8165


R =

    1.0000    1.0000    1.0000
         0    0.0000    0.0000
         0         0    0.0000

>> Q'*Q

ans =

    1.0000   -0.0000   -0.0000
   -0.0000    1.0000    0.0000
   -0.0000    0.0000    1.0000
```

定义 (范数与线性赋范空间)

设V是数域$\mathbb{P}$上的线性空间，定义$\|\cdot\|: u \to \|u\|$是V到$\mathbb{R}$的一个线性映射，满足

1. 正定性(*positivity*)：$\|u\| \geq 0, \forall u \in V$；且$\|u\| = 0 \Leftrightarrow u = \theta$。
2. 齐次性(*scaling*)：$\|\alpha u\| = |\alpha|\|u\|, \forall u \in V, \alpha \in \mathbb{P}$。
3. 三角不等式(*triangle inequality*)：$\|u + v\| \leq \|u\| + \|v\|, \forall u, v \in V$。

称$\|\cdot\|$为V的范数(模)(*norm*)，定义了范数的线性空间称为赋范线性空间。

定义 (距离)

设X是任一非空集合，X中的任意两点x,y有$d(x,y)\in\mathbb{R}$与之对应且满足：

1. 非负性：$d(x,y)\geq 0$，且$d(x,y)=0$ 当且仅当 $x=y$;
2. 对称性： $d(x,y)=d(y,x)$;
3. 三角不等式： $d(x,y)\leq d(x,z)+d(z,y)$。

称$d(x,y)$为X上的一个距离，定义了距离d的集合 X 称为一个距离空间(X,d)，简记为 X。

有了范数就可以定义距离，设V是赋范空间，则$u,v\in V$的距离为

$$d(u,v)=\|u-v\|.$$

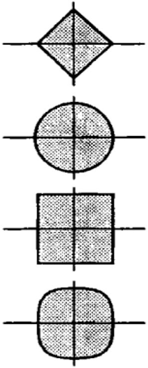
$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^T$可定义范数结构

$$\|\boldsymbol{x}\|_\infty=\max_{1\le i\le n}|x_i|,\quad \infty-\text{范数},$$

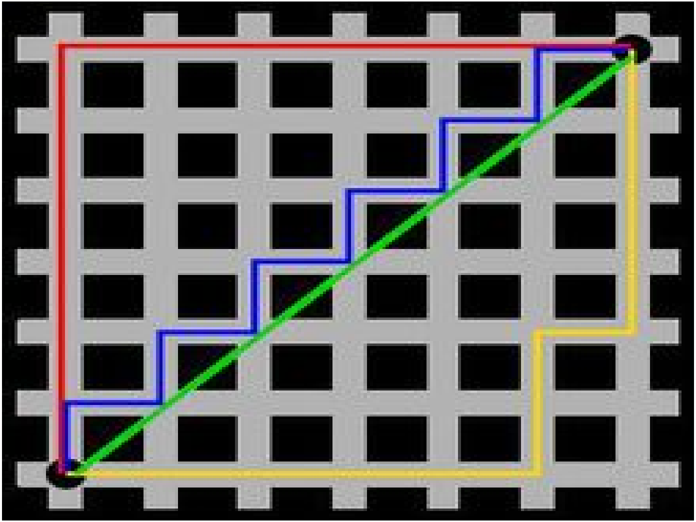
$$\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_i|,\quad 1-\text{范数},$$

$$\|\boldsymbol{x}\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{1/2},\quad 2-\text{范数},$$

$$\|\boldsymbol{x}\|_p=(\sum_{i=1}^{n}|x_i|^p)^{1/p},\quad p-\text{范数}.$$



曼哈顿距离

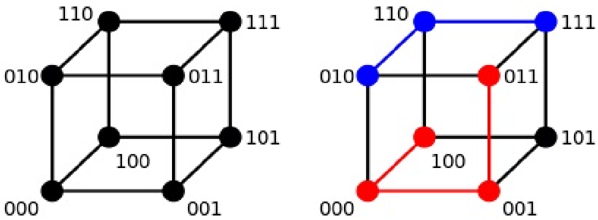


Hamming 距离

定义 (Hamming 距离)

*Hamming 距离*表示两个等长字符串在对应位置上不同字符的数目，汉明距离度量了通过替换字符的方式将字符串x变成y所需要的最小的替换次数。

Hamming 距离主要应用在通信编码领域上，用于制定可纠错的编码体系。在机器学习中，汉明距离也常常被用于作为一种距离的度量方式。



例

设 $\boldsymbol{u}=[-1,-9,2]^T$，则

$$
\begin{aligned}
\|\boldsymbol{u}\|_1 &= |-1|+|-9|+|2| = 12,\\
\|\boldsymbol{u}\|_2 &= \sqrt{|-1|^2+|-9|^2+|2|^2} = \sqrt{86},\\
\|\boldsymbol{u}\|_\infty &= \max\{|-1|,|-9|,|2|\} = 9,\\
\|\boldsymbol{u}\|_5 &= \left(|-1|^5+|-9|^5+|2|^5\right)^{1/5} = 9.0010.
\end{aligned}
$$

Matlab 中用 norm 命令来求向量的范数。

$$\boldsymbol{v}=[1,-7,2]^T$$

```
>> norm(v, 1)
ans =
10.0000
>> norm(v, 'inf')
ans =
7
>> norm(v, 2)
ans =
7.3485
```

通过内积可定义 2-范数

$$\|x\|_2 = \sqrt{(x,x)}.$$

设 $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$，其夹角为 α，则

$$\cos\alpha = \frac{(\boldsymbol{x},\boldsymbol{y})}{\|\boldsymbol{x}\|_2\|\boldsymbol{y}\|_2}.$$

由上面的公式亦可得出Cauchy-Schwarz 不等式，且

$$\begin{aligned}(\boldsymbol{u},\boldsymbol{v}) &= \|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2, \quad \alpha = 0,\\ (\boldsymbol{u},\boldsymbol{v}) &= -\|\boldsymbol{u}\|_2\|\boldsymbol{v}\|_2, \quad \alpha = \pi.\end{aligned}$$

说明二者平行时，等号成立。

注记

内积可用于自然语言处理中的新闻分类。

稀疏性

设$\boldsymbol{x} \in \mathbb{R}^n$，其0-范数为

$$\|\boldsymbol{x}\|_0 = \#(i|x_i \neq 0)$$

0-范数的大小反映了向量$\boldsymbol{x}$的稀疏程度，极小化0-范数可实现“压缩感知(compressed sensing)”和“稀疏编码(sparse coding)”。但是0-范数很难优化求解（NP难问题），而1-范数是-0范数的最优凸近似， 且比0-范数要容易优化求解。所以1-范数也称为“稀疏规则算子”（Lasso regularization）。

$$\begin{array}{l}\min \|x\|_0 \\ \text{s.t. } Ax = b\end{array} \quad \xLeftrightarrow{\text{在一定条件下，以概率1意义下等价}} \quad \begin{array}{l}\min \|x\|_1 \\ \text{s.t. } Ax = b\end{array}$$

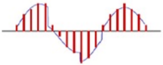
稀疏性可以轻松实现特征选择和可解释性。

传统的数字信号采样定律就是有名的香农采样定理，又称那奎斯特采样定律：

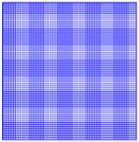
定理 (Shannon采样定理)

为了不失真地恢复模拟信号，采样频率应该不小于模拟信号频谱中最高频率的2倍。

下图分别为在时域和空域上的数字化采集



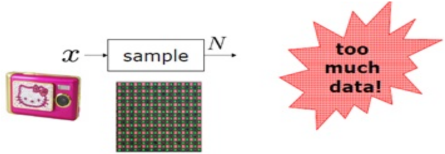
time



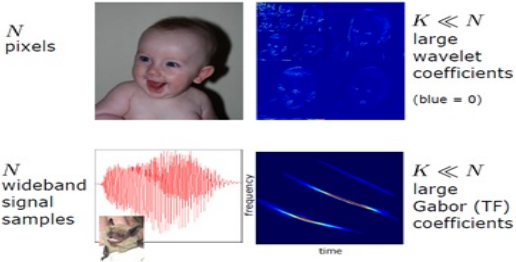
space

基于香农采样定理，目前传统图像信号采集设备的采样过程：

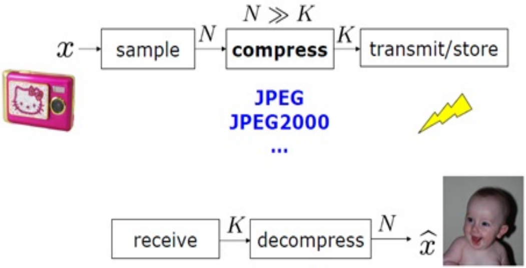
1、按照Nyquist采样率进行均匀采样，得到可以无失真恢复模拟信号的数字信号



信号存在冗余，即信号具有稀疏性：

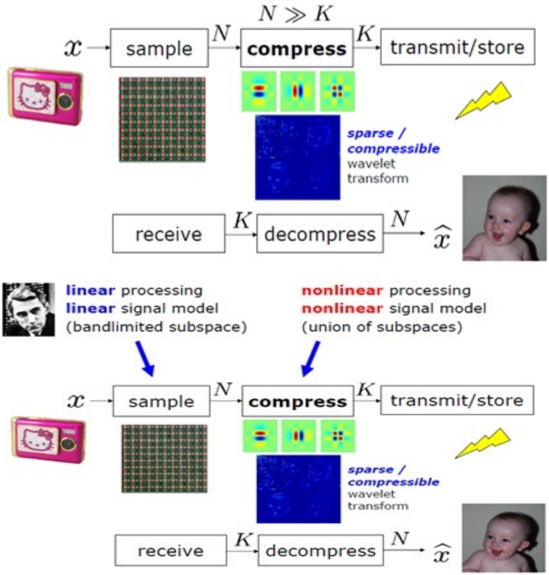


2、上述步骤得到的数字信号的数据量比较大，一方面不利于存储和传输，另一方面该数字信号本来存在很多冗余，可以对其进一步的压缩，于是就通过各种编码方法对数据进行有效的压缩：



相机的传感器通过将模拟信号（光）转换为数字信号，如N pixel的图像信号，之后又通过压缩编码算法将N pixel的图像信号转化为K个系数表示的数据， 而$K \ll N$，那么问题来了，为什么费了一番心思获得了N个采样值，却最后又通过复杂的编码算法将之压缩成K个数值？

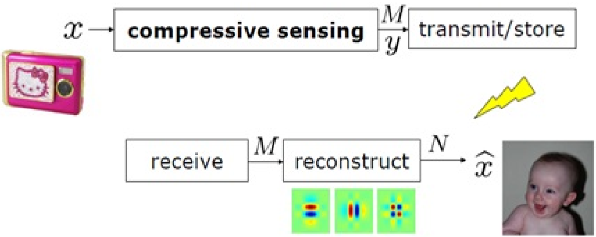
基于以上的疑问，引出了压缩感知的概念：



压缩感知

顾名思义，就是感知压缩，直接获取压缩后的数据。即在采集的时候，直接采集有效的M个测量值，而非满足Nyquist采样定理的N个采样值（$M \ll N$）。

$$K \approx \underline{M} \ll N$$



定义 (范数等价)

设$\|\cdot\|_\alpha$和$\|\cdot\|_\beta$是线性空间V的两个范数。若存在正的常数C_1和C_2，使得

$$C_1\|u\|_\alpha \le \|u\|_\beta \le C_2\|u\|_\alpha, \ \forall u \in S$$

则称范数$\|\cdot\|_\alpha$和范数$\|\cdot\|_\beta$等价。

定理

有限维空间中的任意两个范数等价。

$$\boldsymbol{x} \in \mathbb{R}^n \Longrightarrow \begin{array}{rcl} \|\boldsymbol{x}\|_\infty & \le & \|\boldsymbol{x}\|_2 \le \sqrt{n}\|\boldsymbol{x}\|_\infty, \\ \|\boldsymbol{x}\|_\infty & \le & \|\boldsymbol{x}\|_1 \le n\|\boldsymbol{x}\|_\infty, \\ \|\boldsymbol{x}\|_2 & \le & \|\boldsymbol{x}\|_1 \le \sqrt{n}\|\boldsymbol{x}\|_2. \end{array}$$

对称阵的性质

如果 $A \in \mathbb{R}^{n\times n}$，$A^T = A$，则称 A为 对称阵。它有下列性质：

1. A的特征值均为实数，且有n个线性无关的特征向量。
2. A对应于不同特征值的特征向量必正交。
3. 存在正交阵Q使得$Q^{-1}AQ$为对角阵。

若还满足

$$(A\boldsymbol{x}, \boldsymbol{x}) = \boldsymbol{x}^T A\boldsymbol{x} \geq 0, \ \forall \boldsymbol{x} \in \mathbb{R}^n,$$

则称A为半正定阵(semidefinite matrix)。若上式等号只当$\boldsymbol{x} = \mathbf{0}$时成立，则称$A$为正定阵(definite matrix)。

实对称矩阵为正定阵的充要条件是它的所有特征值都是正数，或者所有顺序主子式都是正定的。

定义 (矩阵范数)

$\mathbb{P}^{n\times n}$中的范数$\|\cdot\|$：是$\mathbb{P}^{n\times n}\to\mathbb{R}:A\to\|A\|$的一个映射满足

- 正定性 $\|A\|\geq 0$, $\|A\|=0\Leftrightarrow A=0$;
- 齐次性 $\|\alpha A\|=|\alpha|\|A\|$;
- 三角不等式 $\|A+B\|\leq\|A\|+\|B\|$;
- 次可乘性(submultiplicative) $\|AB\|\leq\|A\|\|B\|$。

例 (Frobenius范数)

$$A \in \mathbb{R}^{n\times n}, \quad \|A\|_F \stackrel{def}{=} \left(\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2}$$

$$\begin{aligned}
\|AB\|_F^2 &= \sum_{i=1}^{n}\sum_{j=1}^{n}|\sum_{k=1}^{n}a_{ik}b_{kj}|^2 \\
&\le \sum_{i=1}^{n}\sum_{j=1}^{n}\left(\sum_{k=1}^{n}|a_{ik}|^2\sum_{k=1}^{n}|b_{kj}|^2\right) \\
&= \left(\sum_{i=1}^{n}\sum_{k=1}^{n}|a_{ik}|^2\right)\left(\sum_{j=1}^{n}\sum_{k=1}^{n}|b_{kj}|^2\right) \\
&= \|A\|_F^2\|B\|_F^2
\end{aligned}$$

定义 (相容性)

称矩阵范数与向量范数相容：若

$$\|A\boldsymbol{x}\| \leq \|A\| \cdot \|\boldsymbol{x}\|, \ \forall A \in \mathbb{P}^{n\times n}, \ \forall \boldsymbol{x} \in \mathbb{P}^{n}.$$

矩阵的F-范数与向量的2-范数相容：

$$\begin{aligned}
\|A\boldsymbol{x}\|_2^2 &= \sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right)^2 \\
&\leq \sum_{i=1}^{n}\left(\sum_{j=1}^{n}|a_{ij}|^2 \sum_{j=1}^{n} x_j^2\right) \\
&= \|A\|_F^2 \cdot \|\boldsymbol{x}\|_2^2
\end{aligned}$$

$\|A\boldsymbol{x}\|$ 是$\boldsymbol{x} \in \mathbb{R}^n$的连续函数，因为

$$\begin{aligned} \|A(\boldsymbol{x}-\boldsymbol{y})\| &\leq C\|A(\boldsymbol{x}-\boldsymbol{y})\|_2 \\ &\leq C\|A\|_F\|\boldsymbol{x}-\boldsymbol{y}\|_2 \\ &\leq C\|A\|_F C'\|\boldsymbol{x}-\boldsymbol{y}\| \end{aligned}$$

设$\|\cdot\|$ 是 $\mathbb{R}^n$ 中任意给定的一向量范数，集合

$$D = \{\boldsymbol{x} | \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^T \in \mathbb{R}^n, \|\boldsymbol{x}\| = 1\}$$

是一个有界闭集，又 $\|A\boldsymbol{x}\|$ 是 $\boldsymbol{x}$ 的连续函数，则在 D 上可达极大值，即存在 $\boldsymbol{x}_0 \in D$ 使得

$$\|A\boldsymbol{x}_0\| = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\|.$$

对于任意 $\boldsymbol{x} \in \mathbb{R}^n$,

$$\frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \|A\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}\|, \quad \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \in D$$

则 $\frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}$ 在$\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \neq \mathbf{0}$ 上有最大值，则可定义

定义 (从属范数)

设$\|\cdot\|$是一个向量范数，对$\forall A \in \mathbb{R}^{n\times n}$,

$$\|A\| \overset{def}{=} \max_{\boldsymbol{x}\neq 0} \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|} = \max_{\|\boldsymbol{x}\|=1} \|A\boldsymbol{x}\|$$

定义了$\mathbb{R}^{n\times n}$上的一种范数，称为从属于给定向量范数的矩阵范数， 简称从属范数(*subordinate matrix norm*)，亦称算子范数 (*operator norm*)。

- 若$\|A\| = 0$，则$A\boldsymbol{x} = 0, \forall \boldsymbol{x}$，于是$A = 0$；
- 显然$\|\alpha A\| = |\alpha|\|A\|$；
- 显然$\|A + B\| \leq \|A\| + \|B\|$；
- 因为$\|A\boldsymbol{x}\| \leq \|A\|\|\boldsymbol{x}\|$，从属范数是相容的。

$$\|AB\| = \max_{\|\boldsymbol{x}\|=1} \|AB\boldsymbol{x}\| \leq \max_{\|\boldsymbol{x}\|=1} \|A\|\|B\boldsymbol{x}\| = \|A\|\|B\|$$

- 单位矩阵I的任意从属范数为1。
- 相容的范数未必是从属范数。如$\|\cdot\|_F$与$\|\cdot\|_2$相容，但$\|I\|_F = \sqrt{n}$。

设$A \in \mathbb{P}^{n\times n}$，则

$$\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}|$$

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}|$$

$$\|A\|_2 = [\rho(A^T A)]^{1/2}$$

证明$\|A\|_2 = [\rho(A^T A)]^{1/2}$

若 $A = A^T$，则 $\rho(A) = \|A\|_2$，这是否说明 $\rho(A)$ 是一个矩阵范数？

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

则 $\rho(A) = 0$，但 $A \neq 0$。不满足正定性。
考虑 $B = A^T$，则

$$\begin{aligned} \rho(A+B) &= 1, \\ \rho(A) + \rho(B) &= 0. \end{aligned}$$

则

$$\rho(A) + \rho(B) < \rho(A+B),$$

三角不等式不成立。

Matlabe 求矩阵范数

Matlab 中求矩阵范数的命令为 norm，除了求 $1-,2-,\infty-$ 范数，还可求 Frobenius 范数。

$$
\begin{aligned}
A &= \begin{bmatrix} -2 & 1 & -8 & 1 \\ 0 & -4 & -21 & 18 \\ -33 & 16 & -6 & 20 \\ 14 & -20 & -18 & 5 \\ 8 & -1 & 12 & 16 \end{bmatrix} \\
\|A\|_\infty &= 33+16+6+20=75 \\
\|A\|_1 &= 8+21+6+18+12=65
\end{aligned}
$$

```
>> norm(A,'inf')
ans =
    75
>> norm(A,1)
ans =
    65
>> norm(A,'fro')
ans =
  63.5767
```

按照定义，可分别用 Matlab 中的 eig 和 norm 命令来求矩阵的 2-范数。

$$A = \begin{bmatrix} 1 & 13 & 5 & -9 \\ 12 & 55 & 5 & -6 \\ 18 & 90 & 1 & -1 \\ 3 & 0 & 2 & 3 \end{bmatrix}$$

```
>> E = eig(A'*A)

E =
  1.0e+004 *

    0.0000
    0.0021
    0.0131
    1.1802

>> sqrt(max(E))

ans =

  108.6373

>> norm(A,2)

ans =

  108.6373

>> norm(A)  % default without
   second argument is the 2-norm
ans =
  108.6373
```

定理 (谱半径和范数)

设$\|\cdot\|$为任一矩阵范数，则$\|A\| \geq \rho(A)$。

若取$\mathbb{R}^{n\times n}$的范数为$\|\cdot\|_\infty$，$\|\cdot\|_1$和$\|\cdot\|_2$之一，则有

$$\rho(A) \le \|A\|.$$

定理

$\forall A \in \mathbb{R}^{n\times n}$及实数$\epsilon > 0$，存在一种从属的矩阵范数$\|\cdot\|$，使得

$$\|A\| \le \rho(A) + \epsilon$$

定理

设$\|\cdot\|$为$\mathbb{R}^{n\times n}$一从属范数。若$\|B\|<1$，则$I+B$非奇异，且

$$\|(I+B)^{-1}\| \le \frac{1}{1-\|B\|}$$

推论 (摄动引理)

设$A, C \in \mathbb{R}^{n\times n}$，若$A^{-1}$存在，$\|A^{-1}\|\|A-C\| < 1$，则

$$\|C^{-1}\| \leq \frac{\|A^{-1}\|}{1-\|A^{-1}\|\|A-C\|}$$

定义 (实初等矩阵)

设$\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$，$\sigma \in \mathbb{R}$，称

$$E(\boldsymbol{u}, \boldsymbol{v}; \sigma) \overset{def}{=} \boldsymbol{I} - \sigma \boldsymbol{u}\boldsymbol{v}^T$$

为实初等矩阵。

设$\boldsymbol{u} = (u_1, u_2, \cdots, u_n)^T$，$\boldsymbol{v} = (v_1, v_2, \cdots, v_n)^T$， 则

$$
\begin{aligned}
\boldsymbol{u}\boldsymbol{v}^T &= \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix} \begin{pmatrix} v_1 & v_2 & \cdots & v_n \end{pmatrix} \\
&= \begin{pmatrix} u_1v_1 & u_1v_2 & \cdots & u_1v_n \\ u_2v_1 & u_2v_2 & \cdots & u_2v_n \\ \dots & \dots & \dots & \dots \\ u_nv_1 & u_nv_2 & \cdots & u_nv_n \end{pmatrix}
\end{aligned}
$$

引理

初等矩阵若可逆，其逆也是（同类型）初等矩阵。

证明： 因为

$$
\begin{aligned}
E(\boldsymbol{u},\boldsymbol{v};\sigma)E(\boldsymbol{u},\boldsymbol{v};\tau) &= (\boldsymbol{I}-\sigma\boldsymbol{u}\boldsymbol{v}^T)(\boldsymbol{I}-\tau\boldsymbol{u}\boldsymbol{v}^T) \\
&= \boldsymbol{I}-\left(\sigma+\tau-\sigma\tau(\boldsymbol{v}^T\boldsymbol{u})\right)\boldsymbol{u}\boldsymbol{v}^T,
\end{aligned}
$$

若令

$$\sigma+\tau-\sigma\tau(\boldsymbol{v}^T\boldsymbol{u})=0,$$

即

$$\tau=\frac{\sigma}{\sigma\boldsymbol{v}^T\boldsymbol{u}-1},$$

则

$$E(\boldsymbol{u},\boldsymbol{v};\sigma)^{-1}=E(\boldsymbol{u},\boldsymbol{v};\tau).$$

- 若$\boldsymbol{v}^T\boldsymbol{u}=0$，即$\boldsymbol{u}$与$\boldsymbol{v}$垂直，则有$\tau=-\sigma$，

$$E(\boldsymbol{u},\boldsymbol{v};\sigma)^{-1}=E(\boldsymbol{u},\boldsymbol{v};-\sigma).$$

- $\det E(\boldsymbol{u},\boldsymbol{v};\sigma)=1-\sigma\boldsymbol{v}^T\boldsymbol{u}$。原因在于$\boldsymbol{u}\boldsymbol{v}^T$的所有特征值为0和$\boldsymbol{v}^T\boldsymbol{u}$。

- 初等排列阵$\boldsymbol{I}_{ij} \overset{def}{=} E(\mathbf{e}_i - \mathbf{e}_j, \mathbf{e}_i - \mathbf{e}_j; 1)$。
  初等排列阵$\boldsymbol{I}_{ij}$左乘一个矩阵表示将该矩阵的第i行和第j行互换。 初等排列阵$\boldsymbol{I}_{ij}$右乘一个矩阵则表示将该矩阵的第i列和第j列互换。
  $\boldsymbol{I}_{ij}^{-1} = \boldsymbol{I}_{ij}$。
- $E(\mathbf{e}_j, \mathbf{e}_i; -\alpha) = \boldsymbol{I} + \alpha \mathbf{e}_j \mathbf{e}_i^T$
  左乘一个矩阵表示将该矩阵的第i行的α倍加到第j行上去。
  $E(\mathbf{e}_j, \mathbf{e}_i; -\alpha) = \boldsymbol{I} + \alpha \mathbf{e}_j \mathbf{e}_i^T$右乘一个矩阵表示将该矩阵 的第$j$列的$\alpha$倍加到第$i$列上去。

若记$\boldsymbol{l}_j = (0, \cdots, 0, l_{j+1,j}, \cdots, l_{n,j})^T$，其前$j$个分量为零。令

$$L_j(\boldsymbol{l}_j) = E(\boldsymbol{l}_j, \mathbf{e}_j; -1) = \boldsymbol{I} + \boldsymbol{l}_j \mathbf{e}_j^T = \begin{pmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & l_{j+1,j} & 1 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & l_{n,j} & 0 & \cdots & 1 \end{pmatrix}$$

注记

$L_j(\boldsymbol{l}_j)$左乘一个矩阵表示将j行的$l_{k,j}$倍加到k行上去，k从$j+1$行到n行。

由

$$\mathbf{e}_j^T \boldsymbol{l}_j = 0 \Rightarrow L_j(\boldsymbol{l}_j)^{-1} = L_j(-\boldsymbol{l}_j).$$

当$i \le j$时，

$$\begin{aligned} L_i(l_i)L_j(l_j) &= (I + l_i e_i^T)(I + l_j e_j^T) \\ &= I + l_i e_i^T + l_j e_j^T \end{aligned}$$

因而一个（单位）下三角阵可以写成

$$L = L_1(l_1)L_2(l_2)\cdots L_{n-1}(l_{n-1}).$$

定义

设$A \in \mathbb{R}^{n\times n}$，若存在排列矩阵$P \in \mathbb{R}^{n\times n}$使得

$$P^TAP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

如果A_{11}, A_{22}为方阵，则称A为可约矩阵*(reducible matrix)*；否则，称A不可约*(irreducible)*。

注记 (判别准则)

矩阵 $A \in \mathbb{R}^{n\times n}$ 是可约的，若它的指标1到n可被分成两个互不相交的非空集合 $S = \{i_1, i_2, \cdots, i_\alpha\}$和 $T = \{j_1, j_2, \cdots, j_\beta\}$，$\alpha + \beta = n$， 使得

$$a_{i_p, j_q} = 0, \ \ 1 \le p \le \alpha, \ 1 \le q \le \beta.$$

例 (可约矩阵)

$$
A=\begin{bmatrix}
0 & 0 & 1 & 0 & 0\\
1 & 0 & 0 & 2 & 3\\
8 & 0 & 0 & 0 & 9\\
5 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & 12 & 0
\end{bmatrix},
$$

$$
S=\{1,3,4,5\},\quad T=\{2\},
$$

且

$$
a_{12}=a_{32}=a_{42}=a_{52}=0
$$

因此 A 可约。

若A是可约阵，因为$P^T = P^{-1}$，方程组$A\boldsymbol{x} = \boldsymbol{b}$可写为

$$
\begin{aligned}
P^T A P\boldsymbol{y} &= P^T\boldsymbol{b}, \\
\boldsymbol{y} &= P^T\boldsymbol{x}.
\end{aligned}
$$

方程组有分块形式

$$
\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{bmatrix} \bar{\boldsymbol{b}}_1 \\ \bar{\boldsymbol{b}}_2 \end{bmatrix}
$$

注记

若矩阵可约，则原方程可以化为先解一个低阶方程组$A_{22}\boldsymbol{y}_2 = \bar{\boldsymbol{b}}_2$，再解$A_{11}\boldsymbol{y}_1 = \bar{\boldsymbol{b}}_1 - A_{12}\boldsymbol{y}_2$。