第一章 数值方法引论

§1.1 线性空间

- **数域：是复数集合C的一个子集，包含数0和1，该子集中任何两个数的和、差、积、商(除数不为零)仍属于该子集。**

例如：实数域、复数域、有理数域等。

线性空间：设P是一个数域，V是一个非空集合，在V上定义两种运算：

（1）加法

$$u + \mathbf{v} = \mathbf{v} + \mathbf{u}, \qquad \forall\, \mathbf{u}, \mathbf{v} \in \mathbf{V}$$

$$u + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}, \quad \forall\, \mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbf{V}$$

且V中存在唯一的零元素(记为0)，使得

$$u + \mathbf{0} = \mathbf{u}, \qquad \forall\, \mathbf{u} \in \mathbf{V}$$

对每个$\mathbf{u} \in \mathbf{V}$，存在唯一的负元素(记为$-\mathbf{u}$)，使得

$$u + (-\mathbf{u}) = \mathbf{0}, \qquad \forall\, \mathbf{u} \in \mathbf{V}$$

（2）数乘

对任意的$\alpha \in P$和$u \in V$，在V中有唯一的元素（记为αu）与之对应，满足

$$1\boldsymbol{u} = \mathbf{u}, \quad \forall\, \mathbf{u} \in \mathbf{V},$$

$$\boldsymbol{\alpha}(\boldsymbol{\beta u}) = (\boldsymbol{\alpha\beta})\mathbf{u}, \qquad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbf{P}, \forall\, \mathbf{u} \in \mathbf{V}$$

$$\boldsymbol{\alpha}(\boldsymbol{u} + \mathbf{v}) = \boldsymbol{\alpha}\mathbf{u} + \boldsymbol{\alpha}\mathbf{v}, \qquad \forall \boldsymbol{\alpha} \in \mathbf{P}, \forall\, \mathbf{u}, \mathbf{v} \in \mathbf{V}$$

$$(\boldsymbol{\alpha} + \boldsymbol{\beta})\boldsymbol{u} = \boldsymbol{\alpha}\mathbf{u} + \boldsymbol{\beta}\boldsymbol{u}, \qquad \forall \boldsymbol{\alpha}, \boldsymbol{\beta} \in \mathbf{P}, \qquad \forall\, \mathbf{u} \in \mathbf{V}$$

称**V**为一个数域**P**上的线性空间**(**或数域**P**上的向量空间**).**

定义1 设集合S是数域P上的线性空间，元素$x_1,\cdots,x_n\in S$，如果存在不全为零的数$\alpha_1,\cdots,\alpha_n\in P$，使得

$$\alpha_1 x_1+\cdots+\alpha_n x_n=0, \tag{1.1}$$

则称$x_1,\cdots,x_n$**线性相关**. 否则，若(1.1)只对$\alpha_1=\cdots=\alpha_n=0$成立，则称$x_1,\cdots,x_n$**线性无关**.

$$S=\mathrm{span}\{x_1,\cdots,x_n\}.\cdots$$

$$H_n=\mathrm{span}\{1,x,\cdots,x^n\}.\cdots$$

有限维空间 vs 无限维空间.

$\mathbf{R}^n$, $C[a,b]$, $\cdots$

§ 1.2 内积与内积空间

定义2 R^n中向量x及y定义内积：$(x,y)=x_1y_1+\cdots,x_ny_n$.

定义3 设X是数域K(R或C)上的线性空间，对$\forall u,v\in X$，有K中一个数与之对应，记为(u,v)，并满足条件：

(1) $(u,v)=\overline{(v,u)},\ \forall u,v\in X$;

(2) $(\alpha u,v)=\alpha(u,v),\quad \alpha\in \mathrm{R}$;

(3) $(u+v,w)=(u,w)+(v,w),\ \forall u,v,w\in X$;

(4) $(u,u)\geq 0$，当且仅当$u=0$时，$(u,u)=0$.

则称(u,v)为X上的u与v的**内积**. 定义了内积的线性空间称为**内积空间**.

$\overline{(v,u)}$为(u,v)的共轭，当$\mathrm{K}=\mathrm{R}$时$(v,u)=(u,v)$.

定理2 设X为一个内积空间，对$\forall u,v\in X$,有

$$|(u,v)|^2\le (u,u)(v,v).$$

称为$Cauchy-Schwarz$不等式.

定理3 设X为一个内积空间，$\boldsymbol{u_1,u_2,\cdots,u_n\in X}$,矩阵

$$\boldsymbol{G=\begin{bmatrix}(u_1,u_1) & (u_2,u_1) & \cdots & (u_n,u_1)\\ (u_1,u_2) & (u_2,u_2) & \cdots & (u_n,u_2)\\ \vdots & \vdots & \cdots & \vdots\\ (u_1,u_n) & (u_2,u_n) & \cdots & (u_n,u_n)\end{bmatrix}}$$

称为$\boldsymbol{Gram}$矩阵，则$\boldsymbol{G}$非奇异的充要条件是$\boldsymbol{u_1,u_2,\cdots,u_n}$线性无关.

证明 1) G非奇异 $\Leftrightarrow$ 以G为系数矩阵的齐次线性 方程组

$$(\sum_{j=1}^{n} \alpha_j u_j, u_k) = \sum_{j=1}^{n} (u_j, u_k)\alpha_j = 0, \quad k = 1,\cdots,n.$$

只有零解。

2) $$\sum_{j=1}^{n} \alpha_j u_j = 0 \Leftrightarrow (\sum_{j=1}^{n} \alpha_j u_j, \sum_{j=1}^{n} \alpha_j u_j) = 0$$

$$\Leftrightarrow (\sum_{j=1}^{n} \alpha_j u_j, u_k) = 0, \ k = 1,\cdots,n.$$

$\therefore G$非奇异 $\Rightarrow u_1, u_2, \cdots, u_n$线性无关(反证法)　反之亦然.

在内积空间X上可以由内积导出一种 范数,即对$u \in X$,记

$$\| u \| = \sqrt{(u,u)}, \qquad (1.10)$$

易证它满足范数定义的 正定性和齐次性,而三角不等式由$Cauchy-Schwarz$不等式立得.

例1 考察$\mathbf{R}^n$与$\mathbf{C}^n$的内积和范数.

设$x=(x_1,\cdots,x_n)^T, y=(y_1,\cdots,y_n)^T\in\mathbf{R}^n$, 则定义

内积 $(x,y)=\sum_{i=1}^{n}x_iy_i$; 范数 $\|x\|_2=\left(\sum_{i=1}^{n}x_i^2\right)^{1/2}$.

若给定$\omega_i>0(i=1,\cdots,n)$为权系数,则定义

内积 $(x,y)=\sum_{i=1}^{n}\omega_ix_iy_i$; 范数 $\|x\|_2=\left(\sum_{i=1}^{n}\omega_ix_i^2\right)^{1/2}$.

若$x,y\in\mathbf{C}^n$, 则定义加权内积 $(x,y)=\sum_{i=1}^{n}\omega_ix_i\overline{y}_i$.

定义4 设$\rho(x)$是区间$[a,b]$上的非负函数，如果满足条件

(1) $\int_a^b x^k \rho(x)\mathrm{d}x$存在，$k = 0,1,2,\cdots$;

可以有限或无限区间

(2) 对于$[a,b]$上的非负连续函数$g(x)$，若$\int_a^b g(x)\rho(x)\mathrm{d}x = 0$，

则在$[a,b]$上$g(x) \equiv 0$;

就称$\rho(x)$为$[a,b]$上的**权函数**.

例2 设$f(x),g(x)\in C[a,b]$, $\rho(x)$为$[a,b]$上的权函数,则可定义内积

$$(f,g)=\int_a^b \rho(x)f(x)g(x)\mathrm{d}x.$$

$\rho\equiv 1,(f,g)=\int_a^b f(x)g(x)\mathrm{d}x.$

容易验证内积定义中的四个性质，并导出范数

$$\|f(x)\|_2=\left(\int_a^b \rho(x)f^2(x)\mathrm{d}x\right)^{1/2}.$$

$\|f(x)\|_2=\left(\int_a^b f^2(x)\mathrm{d}x\right)^{1/2}.$

设$\varphi_0,\cdots,\varphi_n \in C[a,b]$,则$Gram$矩阵为

$$G = G(\varphi_0,\cdots,\varphi_n)$$

$$= \begin{bmatrix} (\varphi_0,\varphi_0) & (\varphi_0,\varphi_1) & \cdots & (\varphi_0,\varphi_n) \\ (\varphi_1,\varphi_0) & (\varphi_1,\varphi_1) & \cdots & (\varphi_1,\varphi_n) \\ \vdots & \vdots & \cdots & \vdots \\ (\varphi_n,\varphi_0) & (\varphi_n,\varphi_1) & \cdots & (\varphi_n,\varphi_n) \end{bmatrix}$$

根据定理3, $\varphi_0,\cdots,\varphi_n$线性无关 $\Leftrightarrow \det(G) \neq 0.$

§ 1.3 范数与赋范线性空间

定义2 设S是实数域上的线性空间， $x \in S$, 如果存在唯一实数 $\|\cdot\|$, 满足条件

(1) $\|x\| \ge 0$, 当且仅当$x = 0$时, $\|x\| = 0$; (正定性)

(2) $\|\alpha x\| = |\alpha| \|x\|$, $\alpha \in \mathbf{R}$; (齐次性)

(3) $\|x + y\| \le \|x\| + \|y\|$, $x, y \in S$. (三角不等式)

则称 $\|\cdot\|$ 为线性空间S上的**范数**, S与 $\|\cdot\|$ 一起称为**赋范线性空间**, 记为X.

例如，对$\mathbf{R}^n$上的向量$x=(x_1,\cdots,x_n)^T$，有 三种常用范数：

$$\|x\|_\infty=\max_{1\le i\le n}|x_i|,\qquad \text{称为}\infty-\text{范数或最大范数},$$

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

$$\|x\|_2=\left(\sum_{i=1}^{n}x_i^2\right)^{\frac{1}{2}},\qquad \text{称为}2-\text{范数}.$$

类似地，对$C[a,b]$上的$f(x)$，可定义三种常用范数

$$\|f\|_\infty=\max_{a\le x\le b}|f(x)|,\qquad \text{称为}\infty-\text{范数},$$

$$\|f\|_1=\int_a^b|f(x)|\,\mathrm{d}x,\qquad \text{称为}1-\text{范数},$$

$$\|f\|_2=\left(\int_a^b f^2(x)\,\mathrm{d}x\right)^{\frac{1}{2}},\qquad \text{称为}2-\text{范数}.$$

§ 1.4 向量和矩阵范数 /* Norms of Vectors and Matrices */

—— 为了误差的度量

➢ 向量范数 /* vector norms */

定义 R^n空间的**向量范数** $\|\cdot\|$ 对任意 $\vec{x}, \vec{y} \in$ 满足下列条件：

(1) $\|\vec{x}\| \geq 0$; $\|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$ （**正定性** /* positive definite */)

(2) $\|\alpha\vec{x}\| = |\alpha| \cdot \|\vec{x}\|$ 对任意 $\alpha \in C$ （**齐次性** /* homogeneous */)

(3) $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$ （**三角不等式** /* triangle inequality */)

常用向量范数：

$$\|\vec{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|\vec{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}$$

$$\|\vec{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$$

$$\|\vec{x}\|_\infty = \max_{1\leq i\leq n} |x_i|$$

注： $\lim_{p\to\infty} \|\vec{x}\|_p = \|\vec{x}\|_\infty$

定义 向量序列 $\{\vec{x}^{(k)}\}$**收敛**于向量 $\vec{x}^*$是指对每一个 $1 \le i \le n$ 都有 $\lim\limits_{k\to\infty} x_i^{(k)} = x_i^*$。 可以理解为 $\|\vec{x}^{(k)} - \vec{x}^*\|_\infty \to 0$

事实上, $\|\vec{x}^{(k)} - \vec{x}^*\|_\infty = \max(|\vec{x}_1^{(k)} - \vec{x}_1^*|, \cdots, |\vec{x}_n^{(k)} - \vec{x}_n^*|)$

从而当 $k \to \infty$ 时 $\vec{x}^{(k)} \to \vec{x}^*$ 与 $\|\vec{x}^{(k)} - \vec{x}^*\|_\infty \to 0$ 等价。

定义 若存在常数 $C > 0$ 使得对任意 $\vec{x}$ … ^{n} 有 $\|\vec{x}\|_A \le C \|\vec{x}\|_B$, 则称**范数 $\|\cdot\|_A$ 比范数 $\|\cdot\|_B$ 强**。

定义 若范数 $\|\cdot\|_A$ 比 $\|\cdot\|_B$ 强，同时 … 强，即存在常数 C_1、$C_2 > 0$ 使得 $C_1 \|\vec{x}$ … $\|\cdot\|_A$ 和 $\|\cdot\|_B$ **等价**。

可以理解为对任何向量范数都成立。

引理1 设$\|x\|$是在R^n上一个范数，则$\|x\|$是x的分量$x_1, x_2, \cdots x_n$的连续函数。

证明:对任何 $x=(x_1,\cdots,x_n)^T=\sum_{j=1}^{n} x_j e_j$,

$$y=(y_1,\cdots,y_n)^T=\sum_{j=1}^{n} y_j e_j$$

从而

$$\big|\|x\|-\|y\|\big| \le \|x-y\| = \|\sum_{j=1}^{n}(x_j-y_j)e_j\|$$

$$\le \sum_{j=1}^{n} |(x_j-y_j)| \cdot \|e_j\| \quad \le \max_j |(x_j-y_j)| \cdot \sum_{j=1}^{n} \|e_j\|$$

当$y \to x$时，$\|y\| \to \|x\|$，即$\|x\|$为x的连续函数。

定理 R^n上一切范数都等价。

证明：只需证任意范数$\|x\|_p$和$\|x\|_2$等价。

考虑单位球面 $S=\{y \in R^n, \|y\|_2=1\}$ 是 R^n 中的有界闭集。

由引理1，可知$\|x\|_p$连续并在S上达到最大值M和最小值m，

对任何$x \in R^n, x \neq 0$, 有 $x/\|x\|_2 \in S$, 从而

$$m \leq \left\| \frac{x}{\|x\|_2} \right\|_p \leq M$$

因此当 $x \neq 0$,

$$m\|x\|_2 \leq \|x\|_p \leq M\|x\|_2$$

$x = 0, \|x\|_2 = \|x\|_p = 0$ ，上式也成立。

常用向量范数等价形式:

$$\frac{1}{\sqrt{n}}\|x\|_1 \leq \|x\|_2 \leq \|x\|_1$$

$$\|x\|_\infty \leq \|x\|_1 \leq n\|x\|_\infty$$

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n}\|x\|_\infty$$

➢ 矩阵范数 /* matrix norms */

若$\|\cdot\|$是 R^n上的任何向量范数，则对任一n阶实方阵A，定义矩阵范数：

$$\|A\| = \max_{x\neq 0}\frac{\|Ax\|}{\|x\|} = \max_{\|x\|=1}\|Ax\|$$

由向量范数导出的矩阵范数也称为**从属范数**。

性质 $R^{n\times n}$空间的矩阵范数 $\|\cdot\|$ 对任意 $A, B \in R$ 满足：

(1) $\|A\|\geq 0$；$\|A\|=0 \Leftrightarrow A=0$ (正定性 /* positive definite */)

(2) $\|\alpha A\|=|\alpha|\cdot\|A\|$ 对任意$\alpha\in C$ (齐次性 /* homogeneous */)

(3) $\|A+B\|\leq\|A\|+\|B\|$ (三角不等式 /* triangle inequality */)

(4)* $\|Ax\|\leq\|A\|\cdot\|x\|$ (相容 /* consistent */ 任意 $x\in R^n$)

(5)* $\|AB\|\leq\|A\|\cdot\|B\|$ (相容 /* consistent */)

常用矩阵范数：

Frobenius 范数 $\|A\|_F=\sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2}$ —— 向量$\|\cdot\|_2$的直接推广

对方阵$A\in R^{n\times n}$以及 $\vec{x}\in R^n$有 $\|A\vec{x}\|_2\le\|A\|_F\cdot\|\vec{x}\|_2$

$$\|A\cdot B\|_F^2=\sum_{j=1}^{n}\|Ab_j\|_2^2\le\sum_{j=1}^{n}\|A\|_F^2\cdot\|b_j\|_2^2=\|A\|_F^2\cdot\|B\|_F^2$$

利用Cauchy不等式 $|\vec{x}\cdot\vec{y}|\le\|\vec{x}\|_2\cdot\|\vec{y}\|_2$ 可证。

从属范数 /* operator norm */

由向量范数 $\|\cdot\|_p$ 导出关于矩阵 $A\in R^{n\times n}$ 的 p 范数：

$$\|A\|_p=\max_{\vec{x}\ne\vec{0}}\frac{\|A\vec{x}\|_p}{\|\vec{x}\|_p}=\max_{\|\vec{x}\|_p=1}\|A\vec{x}\|_p$$

则 $\|AB\|_p\le\|A\|_p\,\|B\|_p$

$\|A\vec{x}\|_p\le\|A\|_p\,\|\vec{x}\|_p$

矩阵 A^TA 的最大特征根 /*谱半径$\rho(A^TA)$ */

特别有： $\|A\|_\infty=\max_{1\le i\le n}\sum_{j=1}^{n}|a_{ij}|$ （行和范数）

$\|A\|_1=\max_{1\le j\le n}\sum_{i=1}^{n}|a_{ij}|$ （列和范数）

$\|A\|_2=\sqrt{\lambda_{\max}(A^TA)}$ （谱范数 /* spectral norm */ ）

以 $\|A\|_1$ 范数为例证明其从属范数的形式：

$$\|Ax\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} x_j \right| \le \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}| \cdot |x_j|$$

$$= \sum_{j=1}^{n} |x_j| \cdot \sum_{i=1}^{n} |a_{ij}| \le (\max_j \sum_{i=1}^{n} |a_{ij}|) \cdot \|x\|_1$$

故对任何 $\|x\|_1 = 1$ ，$\|Ax\|_1 \le \max_j \sum_{i=1}^{n} |a_{ij}|$

假设 $\max_j \sum_{i=1}^{n} |a_{ij}| = \sum_{i=1}^{n} |a_{ik}|$ ，**令** $y_j = \begin{cases} 1 & j = k, \\ 0 & j \ne k, \end{cases}$ **则有**

$$\|y\|_1 = 1,$$

并且得 $\max_{\|x\|_1 = 1} \|Ax\|_1 \ge \|Ay\|_1 = \sum_{i=1}^{n} \left| \sum_{j=1}^{n} a_{ij} y_j \right| = \sum_{i=1}^{n} |a_{ik}| = \max_j \sum_{i=1}^{n} |a_{ij}|$

从而得

$$\|A\|_1 = \max_{\|x\|_1 = 1} \|Ax\|_1 = \max_j \sum_{i=1}^{n} |a_{ij}|$$

注：☞ Frobenius 范数**不是**从属范数。

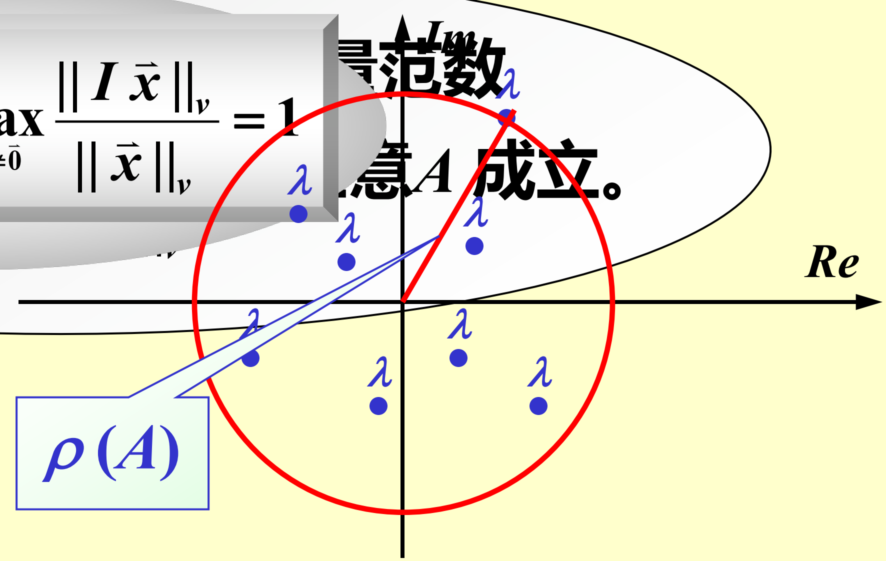
☞ 我们只关心有相容性的范数，**从属范数**总是相容的。

☞ 即使 A 中元素全为实数，其特征根和相应特征向量 /* eigenvector */ 仍可能为复数。将上述定义中绝对值换成**复数模**均成立。

$$\sqrt{n} = \| I \|_F \neq \max_{\vec{x}\neq\vec{0}} \frac{\| I\vec{x} \|_v}{\| \vec{x} \|_v} = 1$$

➢ 谱半径 /* spectral radius */ …量范数 $\| \cdot \|$ …意 A 成立。

定义 矩阵 A 的**谱半径**记为 $\rho(A) = \max_{1\le i\le n} |\lambda_i|$，其中 λ_i 为 A 的特征根。



定理 （1）对任意从属范数 $\|\cdot\|$，有 $\rho(A) \le \|A\|$

（2）存在一种ε从属范数（ ε >0 ）， $\|A\|_{\varepsilon} \le \rho(A) + \varepsilon$

证明： 由从属范数的相容性，得到 $\|A\vec{x}\| \le \|A\|\cdot\|\vec{x}\|$

将任意一个特征根 λ 所对应的特征向量 $\vec{u}$ 代入

$$|\lambda|\cdot\|\vec{u}\| = \|\lambda\vec{u}\| = \|A\vec{u}\| \le \|A\|\cdot\|\vec{u}\|$$

∎

(2)的证明略，可见《矩阵计算的理论和方法》徐树方编。

定理 若A对称，则有 $\|A\|_2 = \rho(A)$

证明： $\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)} = \sqrt{\lambda_{\max}(A^2)}$

A对称

若λ是 A 的一个特征根 则λ^2 必是 A^2 的特征根。

$\Rightarrow \quad \lambda_{\max}(A^2) = \lambda^2(A)$ 某个 A 的特征根λ成立

又：对称矩阵的特 实数，即 $\lambda^2(A)$ 为非负实数，

故得证。 ∎

所以2-范数亦称为**谱范数**。

定理 若矩阵 B 对某个从属范数满足 $\|B\| < 1$，则必有

① $I \pm B$ 可逆；　　② $\left\|(I \pm B)^{-1}\right\| \le \dfrac{1}{1-\|B\|}$

证明：① 若不然，则 $(I \pm B)\vec{x} = \vec{0}$ 有非零解，即存在非零向量 $\vec{x}_0$ 使得 $\pm B\vec{x}_0 = -\vec{x}_0 \;\Rightarrow\; \dfrac{\|B\vec{x}_0\|}{\|\vec{x}_0\|} = 1 \;\Rightarrow\; \|B\| \ge 1$ ✓

② $$(I \pm B)^{-1} \pm B(I \pm B)^{-1} = (I \pm B)(I \pm B)^{-1} = I$$

$$\Rightarrow\; (I \pm B)^{-1} = I \mp B(I \pm B)^{-1}$$

$$\Rightarrow\; \|(I \pm B)^{-1}\| \le 1 + \|B\| \cdot \|(I \pm B)^{-1}\|$$

应用Neumann引理

$$\|(I \pm B)^{-1}\| \le \sum_{k=0}^{\infty} \|(\pm B)^k\| \le \sum_{k=0}^{\infty} \|B\|^k = 1/(1-\|B\|)$$

∎

§1.5 矩阵特征值、相似变换

设$A \in R^{n\times n}$,

一、特征值与特征向量

$$\boxed{Ax = \lambda x} \qquad (\lambda \in C,\ x \neq 0)$$

性质：

(1) $\mathrm{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn} = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

(2) $\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n$

(3) 若 $B = P^{-1}AP$ (称A与B相似)，则 A 与 B 具有相同的特征值和相同的谱（全体特征值的集合）；

x 是 B 的特征向量 $\Leftrightarrow$ Px 是 A 的特征向量

定理 设A∈$R^{n\times n}$，A非奇异⇔⋯?

A可对角化的条件：

1）A与对角阵相似；

2）A有n个线性无关的特征向量；

特别，实对称矩阵的特征值都是实数，且可对角化。

3）A有n个不同的特征值，即A的特征方程的根都是单根， 即$\boldsymbol{det}(\lambda\mathbf{I}-\mathbf{A})=(\lambda-\lambda_{\mathbf{1}})(\lambda-\lambda_{\mathbf{2}})\cdots(\lambda-\lambda_{\boldsymbol{n}})$.

二、代数重数、几何重数

定义：若A的特征方程有重根，

$$det(\lambda I - A) = (\lambda - \lambda_1)^{n_1}(\lambda - \lambda_2)^{n_2}\cdots(\lambda - \lambda_s)^{n_s}$$

其中$\lambda_i \neq \lambda_j$，$i \neq j$，λ_i是特征方程的n_i重根，称为λ_i的代数重数，$n_i \geq 1, i = 1, 2, \cdots, s$，

$$n_1 + n_2 + \cdots + n_s = n$$

定义：设λ_i对应的最大线性无关特征向量的个数为m_i，m_i是齐次线性方程组

$$det(\lambda_i I - A)x = 0$$

的基础解系所含线性无关解的个数，m_i称为λ_i的几何重数，即特征子空间V_{λ_i}的维数。

一般，$m_i \leq n_i, i = 1, 2, \cdots, s.$

定理：若A有重特征值，则A相似于对角阵(可对角化)
$\Leftrightarrow$A有n个线性无关的特征向量

$$\Leftrightarrow m_1 + m_2 + \cdots + m_s = n$$

$$\Leftrightarrow m_i = n_i, \qquad 1 = 1, 2, 3, \cdots, n$$

$$\Leftrightarrow r(\lambda_i \mathrm{I} - \mathrm{A}) = n - n_i$$

当特征值的几何重数≠代数重数时，可以把A化成Jordan标准形。

定理(若当标准形)设方阵A有s个不同的特征值λ_i，则A可通过相似变换化为Joran标准形J，存在可逆阵P，使得

$$P^{-1}AP = J = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_s \end{bmatrix},$$

$$J_i = \begin{bmatrix} J_{i1} & & \\ & \ddots & \\ & & J_{im_1} \end{bmatrix} \qquad J_{ik} = \begin{bmatrix} \lambda i & 0 & \cdots & 0 & 0 \\ 1 & \lambda i & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & & \cdots & & \cdots \\ 0 & 0 & \cdots & 1 & \lambda i \end{bmatrix}.$$

其中m是λ_i的代数重数.

三、特殊矩阵

1） 对角矩阵、三对角矩阵、上（下）三角矩阵
2） 上海森伯(Hessenberg)阵：方阵A，若 $\mathbf{i} > \mathbf{j} + \mathbf{1}$ 时，有 $\boldsymbol{a}(\boldsymbol{i},\boldsymbol{j}) = \mathbf{0}$.
3） 对称矩阵 $\boldsymbol{A}^{\boldsymbol{T}} = \mathbf{A}$
4） 埃尔米特矩阵 （$\boldsymbol{a}(\boldsymbol{i},\boldsymbol{j}) = \overline{\boldsymbol{a}(\boldsymbol{j},\boldsymbol{i})}$）
5） 对称正定矩阵 A实对称且 $(\mathbf{Ax},\mathbf{x}) > \mathbf{0}$，$\mathbf{x} \in \boldsymbol{R}^{\boldsymbol{n}}$
6） 正交矩阵 $\boldsymbol{A}^{\boldsymbol{T}}\mathbf{A} = \mathbf{I}$
7） 酉矩阵 复矩阵$\boldsymbol{U}$满足 $\boldsymbol{U}^{*}\boldsymbol{U} = \boldsymbol{I}$

定理 若$A\in R^{nxn}$对称正定矩阵，则⇒…?

实对称矩阵⇒

(1) A**的特征值都是实数，且有n个线性无关的特征向量；**

(2) A**特征值的特征向量正交；**

(3) **存在正交阵P，使得$P^{-}AP$为对角阵；**

(4) $\|\mathbf{A}\|_2 = \rho(A)$.

正定矩阵⇒

(1) A**的特征值都大于零；**

(2) **所有顺序主子式**$det(A_i) > \mathbf{0}$**；**$\mathbf{i} = \mathbf{1}, \mathbf{2}, \cdots, \mathbf{n}$

定理 若$A\in R^{nxn}$对称矩阵，则对称正定矩阵<=…?

四、初等矩阵

初等变换：对换、倍加、倍乘

定义： 设$u, v \in R^n, \sigma \in R, \sigma \neq 0$，矩阵

$$E(u, v; \sigma) = I - \sigma u v^T$$

称为**实初等矩阵**，其中I是n阶单位阵。

例. $n = 3$，$u = (u_1, u_2, u_3)^T, v = (v_1, v_2, v_3)^T$,

$$E(u, v; \sigma) = \begin{pmatrix} 1 - \sigma u_1 v_1 & -\sigma u_1 v_2 & -\sigma u_1 v_3 \\ -\sigma u_2 v_1 & 1 - \sigma u_2 v_2 & -\sigma u_2 v_3 \\ -\sigma u_3 v_1 & -\sigma u_3 v_2 & 1 - \sigma u_3 v_3 \end{pmatrix}$$

问题：如何使得初等矩阵$E(u, v; \sigma)$可逆？

设$u, v \in R^n, \sigma, \tau \in R$，则

$$E(u,v;\sigma)E(u,v;\tau) = (I - \sigma uv^T)(I - \tau uv^T)$$
$$= I - (\sigma + \tau - \sigma\tau v^T u)uv^T$$

取$\sigma + \tau - \sigma\tau v^T u=0$，即$\tau = \frac{\sigma}{\sigma v^T u-}$，

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