冯果忱 于庚蒲 郃继福

矩阵迭代分析导论

吉林大[illegible]版社

矩阵迭代分析导论

冯果忱　于庚蒲　郤继福

吉 林 大 学 出 版 社

矩阵迭代分析导论

冯果忱　于庚蒲　邻继福

责任编辑：崔晓光　　封面设计：张沭沉

吉林大学出版社出版　　吉林省新华书店发行
（长春市东中华路29号）　　吉林大学印刷厂印刷

开本：850×1168毫米 1/32　　1991年12月第1版
印张：5.75　　1991年12月第1次印刷
字数：140千字　　印数：1—900册

ISBN 7-5601-1064-9/O·123　　定价：2.00元

前　言

1982年以来，我们为吉林大学计算数学专业的大学生和研究生开设了《矩阵迭代分析》课，讲授解线代数方程组的迭代法的一般理论，本书是为这门课程编写的教材。

本书试图用不太大的篇幅阐述解大规模线性代数方程组的迭代法一般理论及迭代法的构造原理。书中采用了《迭代法与二次泛函》[2]中格·伊·马尔丘克与尤·阿·库兹涅佐夫的观点，以二次泛函为工具，统一处理了迭代法的理论问题，特别地，还研究了在系数矩阵奇异时一些迭代法的收敛性问题。

为便于自学，本书自成系统，阅读本书只需一般的线性代数知识，在叙述上力求严谨，基本结果均有严格的证明。

自1983年以来，本书曾多次油印，黑龙江大学数学系及吉林工业大学应用数学系曾多次作为教材使用，他们通过教学实践曾对本书提出过宝贵意见，吉林大学计算数学教研室及出版社的同志们一直关心本书的出版，借此机会对他们表示衷心的谢意。

作者

1991年10月

目　　录

第一章　矩阵论与有限维空间的若干结论…………（1）
§1 矩阵谱的性质…………（1）
§2 不可约矩阵与对角占优矩阵…………（11）
§3 有限维空间的范数…………（16）
§4 线性方程组…………（24）
第二章　迭代法的一般理论…………（30）
§1 基本概念…………（30）
§2 定常迭代法…………（33）
§3 定常迭代法的若干收敛性条件…………（42）
§4 迭代法的收敛性与二次泛函极小化…………（48）
第三章　主要定常迭代法…………（55）
§1 Richardson 方法（RF 方法）与Jacobi 方法…………（55）
§2 逐步超松弛法（SOR方法）…………（62）
§3 分裂算法(Splitting-up method)…………（79）
§4 切比雪夫(Чебышсв) 加速与三项迭代法…………（97）
第四章　非负矩阵与迭代法的单调收敛性…………（111）
§1 非负矩阵…………（111）
§2 L-矩阵及其相关矩阵…………（119）
§3 主要迭代法的单调收敛性…………（122）
§4 正则分裂方法…………（126）
第五章　下降法…………（132）
§1 最速下降法…………（132）

§ 2 多步下降法…………………………………（140）
§ 3 极小迭代法…………………………………（150）
§ 4 共轭梯度法…………………………………（158）
§ 5 条件预优算法………………………………（166）
参考书目……………………………………………（177）

第一章　矩阵论及有限维空间的若干结论

§1　矩阵谱的性质

1.1 自共轭矩阵

考虑 $n\times n$ 复矩阵 $A=(a_{ij})$，a_{ij} 为复数，$i,j=1,2,\cdots,n$。A^T，$\overline{A}$分别表示A的转置矩阵与复共轭矩阵：$A^T=(a_{ji})$，$\overline{A}=(\bar{a}_{ij})$，$\bar{a}_{ij}$为$a_{ij}$的共轭复数。矩阵 $A^*\equiv\overline{A}^T$称为A的**共轭矩阵**。当 $A^*=A$ 时，称A为**自共轭矩阵**。显然，实对称矩阵为自共轭矩阵。

容易验证，矩阵A与其共轭矩阵满足下列关系式：

$$(Ax,y)=(x,A^*y),\qquad \forall (x,y)\in C^n$$

这里C^n表示n维复向量空间，$(\cdot,\cdot)$表示向量内积：

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

其中$x_i,y_i(i=1,2,\cdots,n)$分别为向量x,y的分量。$x=(x_1,\cdots,x_n)^T$，$y=(y_1,\cdots,y_n)^T$。

从上述事实容易看出，自共轭矩阵也可以按下列方式定义。

定义1.1 $n\times n$阶复矩阵 $A=(a_{ij})$ 称为**自共轭矩阵**，如果对任意n维复向量$x,y\in C^n$有

$$(Ax,y)=(x,Ay).$$

设$Q\subseteq C^n$ 为r维子空间，$r\leqslant n$，如果

$$(Ax,y)=(x,Ay), \qquad \forall x,y\in Q$$

则称A于Q上自共轭。

设$u^{(1)},\cdots,u^{(r)}$为Q中正规直交基：

$$(u^{(i)},u^{(j)})=\delta_{ij}, \quad i,j=1,2,\cdots r.$$

$$\delta_{ij}=\begin{cases}0, & i\neq j\\ 1, & i=j\end{cases}$$

在C^n中补充选取$u^{(r+1)},\cdots,u^{(n)}$，使$u^{(1)},\cdots,u^{(n)}$构成C^n中的正规直交基，以$u^{(1)},\cdots,u^{(n)}$为列向量构成的矩阵

$$U=(u^{(1)},\cdots,u^{(n)})$$

为酉矩阵：$UU^*=U^*U=I$.若记

$$U_1=(u^{(1)},\cdots,u^{(r)}), \quad U_2=(u^{(r+1)},\cdots,u^{(n)})$$
$$U=(U_1,U_2)$$

则

$$\tilde{A}\equiv U^*AU=\begin{pmatrix}U_1^*\\ U_2^*\end{pmatrix}A(U_1,U_2)$$
$$=\begin{pmatrix}U_1^*AU_1 & U_1^*AU_2\\ U_2^*AU_1 & U_2^*AU_2\end{pmatrix}$$

$U_1^*AU_1$与$U_2^*AU_2$分别为$r\times r$阶及$(n-r)\times(n-r)$阶矩阵。容易证明，A于Q上自共轭的充分必要条件为$U_1^*AU_1$是自共轭矩阵。此外，如果Q为A的不变子空间，则$U_2^*AU_1=O$。

矩阵特征值的集合称为**矩阵的谱**。自共轭矩阵的谱有许多特别的性质。

首先，自共轭矩阵的所有特征值均为实数，其特征向量组是完全的，而且不同特征值对应的特征向量是相互直交的。因而自共轭矩阵具有完全直交特征向量组。显然可以使特征向量的长度均为1。于是，它们构成正规直交向量组。

其次，任给$n\times n$阶自共轭矩阵A，设其特征值为$\lambda_1,\cdots,\lambda_n$,

相应的特征向量为$u^{(1)},\cdots,u^{(n)}$，假设它们是正规直交的。若记$U=(u^{(1)},\cdots,u^{(n)})$，则$U$为酉矩阵，并且，由于$Au^{(i)}=\lambda_i u^{(i)}$，$i=1,2,\cdots,n$，因此有

$$U^*AU=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}$$

即对任何自共轭矩阵A，都存在酉矩阵U，使U^*AU为对角矩阵，其对角元素为A的特征值，而矩阵U的每一列为A的特征向量。

此外，自共轭矩阵特征值具有重要的极值性质：

设A为$n\times n$阶自共轭矩阵。对任何$x\in C^n, x\neq \mathbf{0}$，

$$\lambda(A,x)\equiv\frac{(Ax,x)}{(x,x)}$$

称为**Rayleigh商**。

对任何$x\neq\mathbf{0}$，令$y=x/(x,x)^{1/2}$，则$(y,y)=1$，且

$$\lambda(A,x)=\frac{(Ax,x)}{(x,x)}=(Ay,y)$$

从而

$$\max_{x\neq 0}\lambda(A,x)=\max_{(y,y)=1}(Ay,y)$$

$$\min_{x\neq 0}\lambda(A,x)=\min_{(y,y)=1}(Ay,y)$$

根据Weierstrass定理，$(Ay,y)=\sum\limits_{i,j=1}^{n}a_{ij}y_i\bar{y}_j$于集合$(y,y)=\sum\limits_{i=1}^{n}|y_i|^2=1$上达到其极大值，记此极大值为$\lambda_M$，即存在$y^M$，$(y^M,y^M)=1$，使

$$\lambda_M=\max_{(y,y)=1}(Ay,y)=(Ay^M,y^M)$$

类似地，存在λ_m, y^m，满足$(y^m, y^m)=1$，且

$$\lambda_m = \min_{(y,y)=1}(Ay, y) = (Ay^m, y^m)$$

设A的特征值为

$$\lambda_1(A) \geqslant \lambda_2(A) \geqslant \cdots \geqslant \lambda_n(A)$$

相应特征向量为

$$x^{(1)}, x^{(2)}, \cdots, x^{(n)}$$

则由

$$Ax^{(i)} = \lambda_i(A)x^{(i)}, \qquad i=1,2,\cdots,n$$

得

$$\lambda(A, x^i) = \lambda_i(A), \qquad i=1,2,\cdots,n$$

从而

$$\lambda_M \geqslant \lambda_i(A) \geqslant \lambda_m, \qquad i=1,2,\cdots,n$$

利用 Lagrange 乘数法容易求出(Ay,y)满足$(y,y)=1$的任何极值点y^*均满足关系式

$$Ay^* = \mu y^*$$

即极值点为A的特征向量。这样，我们得到y^M及y^m分别为A的特征向量，从而λ_M及λ_m为相应的特征值，于是应有

$$\lambda_M = \lambda_1(A), \lambda_m = \lambda_n(A)$$

综合上述讨论，我们得到

定理1.1 设A为$n \times n$阶自共轭矩阵，λ_1, λ_n分别为A的最大与最小特征值，则

$$\lambda_1 = \max_{x \neq 0} \lambda(A, x) \equiv \max_{x \neq 0} \frac{(Ax, x)}{(x, x)}$$

$$\lambda_n = \min_{x \neq 0} \lambda(A, x) \equiv \min_{x \neq 0} \frac{(Ax, x)}{(x, x)}$$

并且使$\lambda(A,x)$取最大与最小的$x^{(1)}, x^{(n)}$，分别为相应于λ_1及λ_n的特征向量。

1.2 正定矩阵

定义1.2 设 $Q\subseteq R^n$ 为n维实向量空间中给定集合，如果对任何 $x\in Q\setminus\{0\}$有$(Ax,x)>0(\geqslant 0)$，则称A于 Q 上**正定**（**正半定**），当$Q=R^n$ 时，称A**正定**（**正半定**）。

注意，在上述正定性的定义中，并未假定A的对称性，这与一般线性代数书中的定义不同。

对于任意实矩阵A，由于

$$(Ax,x)=(x,A^Tx)=(A^Tx,x)$$

故有

$$(Ax,x)=(A_sx,x)$$

其中$A_s=(A+A^T)/2$称为A的**对称部分**。由定义直接推出

命题1.1 矩阵A正定，当且仅当其对称部分$A_s=(A+A^T)/2$正定。

命题1.2 矩阵A正定，当且仅当其一切主子矩阵正定。

证明 充分性显然。为证明必要性，只需如此选取x，当该主子矩阵中不包含某一行（列），例如，不包含第m行（列）时，x的第m个分量取成0，然后按定义即可得证。

命题1.3 对称矩阵A正定，当且仅当其全部特征值为正。

证明 由于A为对称矩阵，据定理1.1，A的最小特征值λ_n具有下述极值性质：

$$\begin{aligned}\lambda_n&=\min_{x\neq 0}\frac{(Ax,x)}{(x,x)}\\&=(Ay^m,y^m)\end{aligned}$$

其中y^m为长度等于1的非零向量，因而$\lambda_n>0$。

命题1.4 矩阵A的特征值的实部$\mathrm{Re}\lambda(A)$满足不等式：

$$\mu_m\leqslant\mathrm{Re}\lambda(A)\leqslant\mu_M$$

其中

$$\mu_m = \min_{x \neq 0} \frac{(A_s x, x)}{(x, x)}$$

$$\mu_M = \max_{x \neq 0} \frac{(A_s x, x)}{(x, x)}$$

分别为A_s的最小及最大特征值。

证明 设y为相应于特征值$\lambda(A)$的特征向量，我们有

$$(Ay, y) = \lambda(A)(y, y)$$
$$(y, Ay) = \lambda(A)(y, y)$$

注意A为实的，因而$(y, Ay) = (A^T y, y)$，由上式得到

$$(A_s y, y) = \left(\frac{A + A^T}{2} y, y\right) = \mathrm{Re}\lambda(A)(y, y)$$

由此显然有

$$\mu_m \leqslant \frac{(A_s y, y)}{(y, y)} \leqslant \mu_M$$

从而

$$\mu_m \leqslant \mathrm{Re}\lambda(A) \leqslant \mu_M$$

推论 当A正定时，A的特征值实部$\mathrm{Re}\lambda(A)$为正数。

对于正半定矩阵，也有与命题 1.1—1.4 平行的结论。此外，还可建立

命题1.5 设矩阵A正半定，$Q \subseteq R^n$，且集合Q与A_s的**核空间**$\mathrm{Ker}A_s \equiv \mathrm{Ker}(A + A^T)$正交，则对任何$x \in Q$，有

$$(Ax, x) \geqslant \alpha(x, x)$$

其中α为A_s的最小非零特征值，从而A于Q上正定。

证明 由于A_s为对称矩阵，因而$\mathrm{Ker}A_s$为由A_s的零特征值对应的特征向量张成的子空间，从而Q属于由A_s的非零特征值对应的特征向量张成的子空间。故对任何$x \in Q$有

$$\frac{(Ax, x)}{(x, x)} = \frac{(A_s x, x)}{(x, x)} \geqslant \inf_{y \in Q} \frac{(A_s y, y)}{(y, y)}$$

$$\geqslant \inf_{y \perp \mathrm{Ker} A_s} \frac{(A_s y, y)}{(y, y)} = \alpha$$

命题1.6 设$Q \subseteq R^n$为闭集，并假定集合

$$Q_1 = \{y \mid y = x/(x,x)^{1/2}, \mathbf{0} \neq x \in Q\} \subseteq Q,$$

而且A于Q上正定，则存在正的常数α，使得

$$(Ax, x) \geqslant \alpha(x, x), \forall x \in Q$$

证明 由于Q为闭集，因而Q_1也是闭集。显然Q_1为有界集。对任何$x \in Q \backslash \{\mathbf{0}\}$，

$$\frac{(Ax, x)}{(x, x)} = \left(A \frac{x}{(x,x)^{1/2}}, \frac{x}{(x,x)^{1/2}}\right)$$

$$= (Ay, y), y \in Q_1$$

由于A于Q上正定，特别，对任何$y \in Q_1$，有$(Ay, y) > 0$，而Q_1为有界闭集，故有α，使

$$0 < \alpha = \inf_{y \in Q_1} (Ay, y)$$

从而

$$(Ax, x) \geqslant \alpha(x, x)$$

对于$x = \mathbf{0}$，上式显然成立。

1.3 可交换矩阵及乘积矩阵AB的特征值

一般说来，矩阵乘法是不可交换的。给定矩阵A、B，若有$AB = BA$，就说A, B是**可交换**的。可交换矩阵的特征值及特征向量有密切关系。

命题1.7 任意可交换矩阵A与B总有相同的特征向量。

证明 设x是矩阵A相应于特征值λ的特征向量：$Ax = \lambda x$，$x \neq \mathbf{0}$。则由A与B的可交换性得到

(1.1) $\quad AB^k x = \lambda^k B x, k = 0, 1, 2, \cdots$.

设向量列$x,Bx,B^2x,\cdots$的前m个向量线性无关，第$m+1$个向量$B^m x$可表为前m个向量的线性组合。于是，由$x,Bx,\cdots,B^{m-1}x$张成的子空间S是B的不变子空间：对$\forall y\in S$，有$By\in S$。因而存在B的特征向量$v\in S$：$Bv=\mu v, v\neq 0$。另一方面，由(1.1)，$x,Bx,\cdots,B^{m-1}x$为矩阵A相应于同一特征值λ的特征向量，因而其任意非零线性组合也是A相应于λ的特征向量。于是v是A相应于特征值λ的特征向量。

命题1.8 设矩阵A、B有完全的特征向量系，那么，A与B可交换的充分必要条件是A与B有公共的特征向量系。

证明 充分性。设$u^{(1)},\cdots,u^{(n)}$是A与B公共的特征向量系。记$U=(u^{(1)},\cdots,u^{(n)})$，则有

$$U^{-1}AU=\Lambda_1,\quad U^{-1}BU=\Lambda_2$$

其中$\Lambda_1=\mathrm{diag}(\lambda_i(A))$，$\Lambda_2=\mathrm{diag}(\lambda_i(B))$。显然$\Lambda_1\Lambda_2=\Lambda_2\Lambda_1$，从而

$$U^{-1}AUU^{-1}BU=U^{-1}BUU^{-1}AU$$

由此得$AB=BA$。

必要性。若$AB=BA$，设u为A相应于$\lambda(A)$的任一特征向量：$Au=\lambda(A)u$，我们有

$$BAu=\lambda(A)Bu$$

从而

$$A(Bu)=\lambda(A)(Bu)$$

这表明矩阵A的任何特征子空间均为B的不变子空间。因而如果U是以A的特征向量为列构成的矩阵，并且把同一特征值对应的特征向量排在一起，则$U^{-1}BU$具有分块对角形式：

$$U^{-1}BU=\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_m \end{pmatrix}$$

由于B有完全特征向量系，因而存在矩阵V，使

$$V^{-1}\begin{pmatrix} B_1 & & & \\ & B_2 & & \\ & & \ddots & \\ & & & B_m \end{pmatrix}V=\operatorname{diag}(\lambda_i(B))$$

即$V^{-1}U^{-1}BUV=\operatorname{diag}(\lambda_i(B))$。并且$V$可以取成$U^{-1}BU$ 相同结构的分块对角形式：

$$V=\begin{pmatrix} V_1 & & & \\ & V_2 & & \\ & & \ddots & \\ & & & V_m \end{pmatrix}$$

另一方面

$$U^{-1}AU=\begin{pmatrix} \lambda_1(A)I_1 & & & \\ & \lambda_2(A)I_2 & & \\ & & \ddots & \\ & & & \lambda_m(A)I_m \end{pmatrix}$$

容易算出

$$V^{-1}U^{-1}AUV=\begin{pmatrix} \lambda_1(A)I_1 & & & \\ & \lambda_2(A)I_2 & & \\ & & \ddots & \\ & & & \lambda_m(A)I_m \end{pmatrix}$$

以上分析表明矩阵UV的列是矩阵A与B的共同的特征向量系。

推论 若矩阵 A、B 可交换，并且它们有完全的特征向量系，则A、B与AB有相同的特征向量系。

当A、B不可交换时，矩阵 AB 与 BA 未必有相同的特征向量，但是我们有

命题1.9 对任何$n\times n$阶矩阵A、B，AB与 BA 有相同的谱，并且对于非零特征值λ，有

$$\dim[\ker(\lambda I - AB)] = \dim[\ker(\lambda I - BA)].$$

证明 对任意复数 μ，我们有

$$\begin{pmatrix} I & O \\ -B & \mu I \end{pmatrix}\begin{pmatrix} \mu I & A \\ B & \mu I \end{pmatrix} = \begin{pmatrix} \mu I & A \\ O & \mu^2 I - BA \end{pmatrix}$$

$$\begin{pmatrix} \mu I & -A \\ O & I \end{pmatrix}\begin{pmatrix} \mu I & A \\ B & \mu I \end{pmatrix} = \begin{pmatrix} \mu^2 I - AB & O \\ B & \mu I \end{pmatrix}$$

于上式两端取行列式得到

$$\mu^n \det(X) = \mu^n \det(\mu^2 I - BA)$$

$$\mu^n \det(X) = \mu^n \det(\mu^2 I - AB)$$

其中

$$X = \begin{pmatrix} \mu I & A \\ B & \mu I \end{pmatrix}$$

由此得到

$$\det(\mu^2 I - AB) = \det(\mu^2 I - BA)$$

即AB与BA有相同的特征多项式，因而它们的谱相同。

设$\lambda \neq 0$为AB的特征值，$\{\mu^{(i)}\}$为相应的最大线性无关特征向量系。由$ABu^{(i)} = \lambda u^{(i)}$，两端乘以$B$，得到

$$(1.2) \qquad BA(Bu^{(i)}) = \lambda(Bu^{(i)})$$

记$\mathbf{v}^{(i)} = B\mathbf{u}^{(i)}$，有

$$A\mathbf{v}^{(i)} = AB\mathbf{u}^{(i)} = \lambda \mathbf{u}^{(i)}$$

而 $\mathbf{u}^{(i)} \neq \mathbf{0}$，从而$\mathbf{u}^{(i)} \neq \mathbf{0}$。从(1.2)式看出，$\mathbf{v}^{(i)}$是 BA 相应于 λ 的特征向量。再由

$$\mathbf{u}^{(i)} = \frac{1}{\lambda} A\mathbf{v}^{(i)}$$

推出，当$\{\mathbf{u}^{(i)}\}$线性无关时，$\{\mathbf{v}^{(i)}\}$也线性无关。

对于AB的零特征值，AB与 BA 的最大线性无关特征向量数未必相同，例如

$$A=\begin{pmatrix}1 & -1\\ 1 & -1\end{pmatrix},\qquad B=\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}$$

这时$AB=O,BA=2A$，它们都只有零特征值，然而A的特征向量为$\alpha(1,1)^T,\alpha\neq 0$，而任何二维非零向量均为零矩阵的特征向量。

命题1.10 若A、B均为对称矩阵，则AB与BA之同一特征值相应的最大线性无关特征向量数相同。

证明 由于A,B对称，故$(AB)^T=B^TA^T=BA$。命题显然成立。

§2 不可约矩阵与对角占优矩阵

定义2.1 $n\times n$阶矩阵$A=(a_{ij})$称为是**可约的**，如果$n>1$，且存在$N=\{1,2,\cdots,n\}$的一个非空不交分划S、T，$S\cup T=N$，使对任何$i\in S$，$j\in T$有$a_{ij}=0$。否则称A为**不可约的**。规定一阶矩阵是不可约的。

设$(i_1,i_2,\cdots,i_n)$是$(1,2,\cdots,n)$的一个置换，矩阵$P=(e^{(i_1)},e^{(i_2)},\cdots,e^{(i_n)})$是将单位矩阵$I=(e^{(1)},e^{(2)},\cdots,e^{(n)})$的列按$i_1,i_2,\cdots,i_n$的次序重新排列得到的，称为**重排矩阵**。显然有$P^TP=PP^T=I$，而AP是将矩阵A的列按$i_1,i_2,\cdots,i_n$的次序重排得到的，P^TA则是将A的行按$i_1,i_2,\cdots,i_n$的次序重排的结果。

定理2.1 $n>1$阶矩阵$A=(a_{ij})$是可约的，当且仅当存在重排矩阵P，使

$$P^TAP=\begin{pmatrix}F & O\\ G & H\end{pmatrix}\tag{2.1}$$

其中F,H为方阵。

证明 若A可约，则有相应的分划$S\cup T=N$，使当$i\in S$，