矩阵分析

作者：Kamden Wang

时间：December 25, 2022

主要参考资料：

《矩阵数值分析》刑志栋，

《Numerical Analysis》Timothy Sauer,

《Matrix Analysis》Horn

目录

第 1 章 向量范数与矩阵范数 **1**
1.1 向量与矩阵序列的极限 1
1.2 向量范数 1
1.3 矩阵范数 2
1.4 谱半径与收敛性定理 5

第 2 章 线性代数方程组的直接解法 **6**
2.1 简单情形 6
2.2 Gauss 消去法 6
2.3 三角分解 6
2.4 正交三角分解 7
2.5 线性矛盾方程组的最小二乘法 8
2.6 方程组的条件问题和算法的确误差分析 9

第 3 章 线性代数方程组的迭代解法 **10**
3.1 引言 10
3.2 常用的迭代方法 11
3.2.1 Jacobi 方法 11
3.2.2 Gauss-Seidel 方法与 SOR 12
3.3 共轭方向法 13

第 4 章 矩阵特征问题的形态 **14**
4.1 特征值的估计及极值性质 14

第 1 章 向量范数与矩阵范数

1.1 向量与矩阵序列的极限

定义 1.1 (向量序列与矩阵序列的极限)

向量序列与矩阵序列的极限即是每个元素趋于极限后的整体, 若其中有一个元素极限不存在, 那么就说该向量或矩阵序列不存在极限

定理 1.1

向量序列和矩阵序列的极限满足极限的运算法则

定义 1.2

设 $\boldsymbol{C}^n$ 上的向量序列 $\{x_k\}$ 以及向量 x_0 满足: 对给定的向量范数 $||\cdot||_v$, $\lim_{k\to\infty}||x_k - x_0||_v = 0$ 则称序列依该范数收敛于 x_0, 并记 $\lim_{k\to\infty} x_k = x_0$

笔记 矩阵序列同理, 详细参考课件

例题 1.1 课件 p20,21

1.2 向量范数

定义 1.3 (向量范数)

对任意向量 $\boldsymbol{x}, \boldsymbol{y}$ 以及复数 α, 函数 $f(x) = ||x||$ 满足以下三个条件:

1. 非负性:$||\boldsymbol{x}|| \leq 0, ||\boldsymbol{x}|| = 0 \iff \boldsymbol{x} = \boldsymbol{0}$
2. 齐次性:$||\alpha\boldsymbol{x}|| = |\alpha|||\boldsymbol{x}||$
3. 三角不等式:$||\boldsymbol{x}+\boldsymbol{y}|| \leq ||\boldsymbol{x}|| + ||\boldsymbol{y}||$

称函数 $||\cdot||$ 为复数域上一个**向量范数**

推论 1.1 (三角不等式的推广)

$$\Big| ||\boldsymbol{x}|| - ||\boldsymbol{y}|| \Big| \leq ||\boldsymbol{x} - \boldsymbol{y}||$$

证明 令 $\boldsymbol{y} = \boldsymbol{x} + (\boldsymbol{y} - \boldsymbol{x})$ 代入运算即证

例题 1.2 对任给 $\boldsymbol{x} = (x_1, x_2, x_3)^T$, 试问如下实值是否构成向量范数?

1. $|x_1| + |2x_2 + x_3|$
2. $|x_1| + |2x_2| - 5|x_3|$
3. $|x_1|^4 + |x_2|^4 + |x_3|^4$
4. $|x_1| + 3|x_2| + 2|x_3|$

解

1. 不满足非负性, 取 $\boldsymbol{x} = (0, 2, -4)^T$
2. 不满足非负性, 取 $\boldsymbol{x} = (0, 5, 2)^T$
3. 不满足齐次性
4. 构成向量范数

定义 1.4 (L_p 范数 (Holder 范数))

$$||\boldsymbol{x}||_p = (\sum_{i=1}^{\infty} |x_i|_p)^{\frac{1}{p}}$$

♣

定义 1.5 (加权范数 (椭圆范数))

$$||\boldsymbol{x}||_A = ||A\boldsymbol{x}||$$

♣

例题 1.3 课件 p30

定理 1.2 (向量范数等价性定理)

对 n 维向量空间上的任意两个不同的向量范数 $||\cdot||$ 和 $||\cdot||'$, 存在常数 $c_2 \geq c_1 > 0$ 使得对任意 x, 有

$$c_1||\boldsymbol{x}|| \leq ||\boldsymbol{x}||' \leq c_2||\boldsymbol{x}||$$

♡

推论 1.2 (P-范数等价性)

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

或者，将它们结合，形成一个不等式链

$$||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_2 \leq ||\boldsymbol{x}||_1 \leq \sqrt{n}||\boldsymbol{x}||_2 \leq n||\boldsymbol{x}||_\infty$$

♡

证明

$$||\boldsymbol{x}||_\infty = \max_{1\leq i\leq n} |x_i| \leq \sum_{i=1}^{n} |x_i| = ||\boldsymbol{x}||_1 \leq \sum_{i=1}^{n} (\max_{1\leq i\leq n} |x_i|) = n||\boldsymbol{x}||_\infty$$

$$\frac{1}{\sqrt{n}}||\boldsymbol{x}||_1 = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}|x_i| = \frac{1}{\sqrt{n}}\sqrt{(\sum_{i=1}^{n}|x_i|)^2} \leq \frac{1}{\sqrt{n}} \times \sqrt{1^2+1^2+\cdots+1^2}\sqrt{\sum_{i=1}^{n}{x_i}^2} = ||\boldsymbol{x}||_2$$

$$\frac{1}{\sqrt{n}}||\boldsymbol{x}||_2 = \frac{1}{\sqrt{n}}\sqrt{\sum_{i=1}^{n}{x_i}^2} \leq ||\boldsymbol{x}||_\infty \leq ||\boldsymbol{x}||_2$$

1.3 矩阵范数

定义 1.6 (矩阵范数)

对任意矩阵 A, B 和复数 α, 函数 $f(A) = ||A||$ 满足以下三条件:

1. 非负性:$||A|| \leq 0, ||A|| = 0 \iff A = 0$
2. 齐次性:$||\alpha A|| = |\alpha|||A||$
3. 三角不等式:$||A+B|| \leq ||A|| + ||B||$

加上满足下述的相容性条件, 称函数 $||\cdot||$ 为复数域上的一个矩阵范数

♣

相比于向量范数矩阵范数应考虑到矩阵乘法

定义 1.7 (矩阵范数相容性)

对任意矩阵 $A \in \boldsymbol{C}^{m\times l}, B \in \boldsymbol{C}^{l\times n}$, 满足

$$||AB|| \le ||A|| \cdot ||B||$$

命题 1.1 (由向量的 1,2 ∞-范数直接推广的函数)

$$||A||_{m_1} = \sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|$$

$$||A||_F = (\sum_{i=1}^{n}\sum_{j=1}^{n}{a_{ij}}^2)^{\frac{1}{2}}$$

$$||A||_{m_\infty} = \max_{ij}|a_{ij}|$$

根据相容性的定义验证发现前两个属于矩阵范数, 最后一个不属于矩阵范数, 但定义其为

$$||A||_{m_\infty} = \sqrt{m\cdot n}\max_{ij}|a_{ij}|$$

即可构成 A 的一种范数

定理 1.3 (矩阵范数等价性定理)

对 $n\times n$ 维向量空间上的任意两个不同的矩阵范数 $||\cdot||$ 和 $||\cdot||'$, 存在常数 $c_2 \ge c_1 > 0$ 使得对任意 A, 有

$$c_1||A|| \le ||A||' \le c_2||A||$$

矩阵与向量的乘积在矩阵计算中经常出现, 所以我们希望矩阵范数和向量范数最好有某种协调性

命题 1.2 (矩阵范数与向量范数相容性)

对于矩阵范数 $||\cdot||_M$ 与向量范数 $||\cdot||_V$, 有

$$||A\boldsymbol{x}||_v \le ||A||_M||\boldsymbol{x}||_V$$

定义 1.8 (矩阵算子范数)

定义

$$||A||_M = \max_{\boldsymbol{x}\neq 0}\frac{||A\boldsymbol{x}||_V}{||\boldsymbol{x}||_V} = \max_{||\boldsymbol{x}||_V=1}||A\boldsymbol{x}||_V$$

为从属向量范数的矩阵范数, 简称从属范数或算子范数

笔记 两个等号成立的证明, 课件 p45

首先来看存在性, 由于 $||A\boldsymbol{x}||_V$ 是 $\boldsymbol{C}^n$ 上的连续函数,$D = \left\{\boldsymbol{x}\middle|||\boldsymbol{x}||_V = 1\right\}$ 为有界闭集, 因此存在 $\boldsymbol{x_0}$ 使得 $||A||_M = \max_{||\boldsymbol{x}||_V=1}||A\boldsymbol{x}||_V$, 再来看算子范数是否是矩阵范数, 根据定义可证明算子范数是矩阵范数, 课件 p44

推论 1.3 (从属于向量 p-范数的算子范数)

A 为 $m \times n$ 矩阵

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

$$||A||_2 = \sqrt{\lambda_{max}(A^T A)} \qquad \text{(谱范数)}$$

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

同时它们有如下关系

$$||A||_2{}^2 \le ||A||_1 ||A||_\infty$$

♡

证明 课件 p46-p50

定义 1.9 (Frobenius 矩阵范数)

设 A 为 $m \times n$ 阶矩阵

$$||A||_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}{}^2}$$

♣

例题 1.4 课件 p53

定理 1.4 (酉矩阵的范数不变性)

对于酉矩阵 $U^H U = UU^H = I$, 有

$$||U||_2 = 1, ||AU||_2 = ||UA||_2 = ||A||_2$$

♡

证明 讲义 p54

推论 1.4 (F-范数的酉不变性)

设 A 为 n 阶方阵, 则存在 n 阶酉矩阵 U, V, 使得

$$||UA||_F = ||AV||_F = ||UAV||_F = ||A||_F$$

♡

证明 讲义 p55

总结

1. 任意给定的矩阵范数必然存在与之相容的向量范数；任意给定的向量范数必然存在与之相容的矩阵范数 (如算子范数)
2. 一个矩阵范数可以与多种向量范数相容 (矩阵的 m1-范数与矩阵的 p-范数相容)；多种矩阵范数可以与一个向量范数相容 (矩阵的 F-范数、2-范数与向量的 2-范数相容)。
3. 算子范数一定与所定义的向量范数相容，但是矩阵范数与向量范数相容却未必有从属关系。(矩阵的 F-范数与向量 2-相容，但无从属关系)。
4. 并非任意的矩阵范数与任意的向量范数相容 (课件 p57)

1.4 谱半径与收敛性定理

定义 1.10 (谱和谱半径)

把矩阵 A 所有特征值组成的集合称为谱, 并记为 $\sigma(A)$; 把矩阵 A 的特征值模的最大值定义为其谱半径, 并记其为 $\rho(A)$, 即

$$\sigma(A) = \{\lambda | det(\lambda I - A) = 0\} \qquad \rho(A) = \max_{1 \le i \le n} |\lambda_i|$$

♣

笔记 对称矩阵的 2-范数与其谱半径相等

定理 1.5

1. 对任意一种矩阵 A 的从属范数, 有 $\rho(A) \le ||A||$
2. 对任意 $\varepsilon > 0$, 存在 A 的一种矩阵范数 $||\cdot||$ 使得

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

♡

证明 课件 p4

定理 1.6

设矩阵 A 是 n 阶方阵, 则 $A^m \to 0$ $(m \to \infty)$ 的充要条件是 $\rho(A) < 1$

♡

例题 1.5 课件 p11

证明 课件 p8

定义 1.11 (矩阵级数)

设 A 为 $\boldsymbol{C}^{m\times n}$ 中的矩阵序列, 称

$$A_1 + A_2 + \cdots + A_k + \cdots$$

为矩阵序列构成的矩阵级数, 记为 $\sum_{k=1}^{\infty} A_k$

♣

定理 1.7

级数 $I + A + A^2 + \cdots + A^m + \cdots$ 收敛的充要条件是 $A^m \to 0$ $(m \to \infty)$ 若该矩阵级数收敛, 则收敛至 $(I - A)^{-1}$

♡

例题 1.6 课件 p15

证明 课件 p14

定理 1.8

如果矩阵 A 的一种范数小于 1, 则对任何非负整数 k 下面关系式成立:

$$||(I - A)^{-1} - (I + A + A^2 + \cdots + A^k)|| \le \frac{||A||^{k+1}}{1 - ||A||}$$

♡

定理 1.9

设 A, C 是 n 阶矩阵, 设 A 可逆, 并且 $||A^{-1}|| \le \alpha$ 若 $||A - C|| \le \beta$ $\quad \alpha\beta < 1$ 那么 C 也是可逆的, 且

$$||C^{-1}|| \le \frac{\alpha}{1 - \alpha\beta}$$

♡

证明 两个定理证明见图片补充 1(1),1(2),1(3)

第 2 章 线性代数方程组的直接解法

2.1 简单情形

定理 2.1
上三角方程组有唯一解的充要条件是主元全不为零 ♡

见图片补充 2(1)

2.2 Gauss 消去法

命题 2.1 (Gauss 消去法两大过程)
1. 消元过程: 经过 n-1 步消元将原方程组化成等价的上三角形方程组
2. 回代过程: 利用逐步回代, 求解三角形方程组 ♠

结论 Gauss 消元法消元过程乘除法的总运算次数为:$\frac{2n^3+3n^2-5n}{6}$, 回代过程的运算次数为:$\frac{n^2+n}{2}$

命题 2.2 (Gauss 消去法主要缺点)
1. 遇上某一步的主元素为零时, 消元就不能进行了
2. 有时尽管主元不为零, 但它可能是一个绝对值很小的数, 在数值计算角度上会引起很大的舍入误差, 解的可靠性差, 甚至无准确度可言. ♠

见图片补充 2(2) 2(6)

2.3 三角分解

定理 2.2
设 A 是一个 $n \times n$ 的方阵, 它的所有顺序子式 A_i 非奇异, 则有

$$A = L\,R$$

其中 L 为下三角矩阵 (**L**ower triangular matrix), 当 L 或 R 的对角元素给定之后, 这种分解是唯一的. ♡

定理 2.3 (Doolittle Crout 分解)
Doolittle 分解即 L 主对角线都为 1 的 $L\,R$ 分解
Crout 分解即 R 主对角线都为 1 的 $L\,R$ 分解 ♡

例题 2.1 课件 p10,11

定理 2.4 (Cholesky 分解 (正定矩阵))
若 A 为 n 阶正定矩阵, 那么存在一个 n 阶下三角矩阵 L 使得

$$A = L\,L^T$$ ♡

结论 Cholesky 算法的计算量为 $\frac{n^3+6n^2+8n}{6}$

定理 2.5

设 n 阶矩阵 A 正定, 那么存在 n 阶正定矩阵 B 使得

$$A = B^2$$

定义 2.1 (改进平方根法)

在原来的 Cholesky 分解的条件下, 可将分解改为

$$A = L\,D^{\frac{1}{2}}\,D^{\frac{1}{2}}\,L^T = L\,D\,L^T$$

见图片补充 2(7) 2(9)

定义 2.2

设 A 为 n 阶方阵, 若对 $|i-j| > m$ 时, 有 $a_{ij} = 0$, 就称 A 是带宽为 $2m+1$ 的带状矩阵

定理 2.6

若 n 阶方阵 A 带宽为 $2m+1$, 假设 A 可以分解为下三角阵和上三角阵的乘积, 即 $A = L\,R$, 那么 L 和 R 同时具有带宽 $2m+1$

2.4 正交三角分解

定义 2.3 (初等矩阵)

初等矩阵表示为

$$E(\boldsymbol{u}, \boldsymbol{v}; \alpha) = I - \alpha \boldsymbol{u}\boldsymbol{v}^T, \qquad \boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{C}^n, \alpha \in \boldsymbol{C}$$

此时 $\boldsymbol{u}\boldsymbol{v}^T$ 为一个秩一矩阵, 且唯一一个非零特征值为 $\boldsymbol{v}^T\boldsymbol{u}$

事实上,

$$A\boldsymbol{u} = (\boldsymbol{u}\boldsymbol{v}^T)\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{v}^T\boldsymbol{u}) = (\boldsymbol{u}\boldsymbol{v}^T)\boldsymbol{u} = \lambda\boldsymbol{u}$$

定义 2.4 (Householder 矩阵)

Householder 矩阵, 记为 $H(\boldsymbol{\omega})$ 定义为

$$H(\boldsymbol{\omega}) = E(\boldsymbol{\omega}, \boldsymbol{\omega}; 2) = I - 2\boldsymbol{\omega}\boldsymbol{\omega}^H, \quad \boldsymbol{\omega}^H\boldsymbol{\omega} = 1, \boldsymbol{\omega} \in \boldsymbol{C}^n$$

性质 Householder 矩阵的性质

1. 对称性: $H(\boldsymbol{\omega})^H = H$
2. 正交性: $H(\boldsymbol{\omega})H(\boldsymbol{\omega})^H = I$
3. 保范数: $||H(\boldsymbol{\omega})\boldsymbol{x}||_2 = ||\boldsymbol{x}||_2$

证明 见课件 p5

定理 2.7 (镜面反射性质)

设 $\boldsymbol{x}$ 是 $\boldsymbol{R}^n$ 中的一个非零向量, $\boldsymbol{g}$ 是 $\boldsymbol{R}^n$ 中任意一个单位向量, 则存在 Householder 矩阵, 使得

$$H(\boldsymbol{\omega})\boldsymbol{x} = ||\boldsymbol{x}||_2\boldsymbol{g}$$

证明 课件 p7

定理 2.8

设 $\boldsymbol{x}, \boldsymbol{y}$ 是 $\boldsymbol{C}^n$ 中的两个向量, 则存在 Householder 矩阵 $H(\boldsymbol{\omega})$ 满足 $H(\boldsymbol{\omega})\boldsymbol{x} = \boldsymbol{y}$ 的充要条件是

1. $||\boldsymbol{x}||_2 = ||\boldsymbol{y}||_2 = (\boldsymbol{x}^H \boldsymbol{x})^{\frac{1}{2}}$
2. $\boldsymbol{x}^H \boldsymbol{y} = \boldsymbol{y}^H \boldsymbol{x}$

♡

证明 课件 p8

例题 2.2 课件 p11

例题 2.3 用正交三角分解解线性方程组, 课件 p19,p24

定理 2.9 (Shur 分解)

设 A 是 $n \times n$ 复矩阵, 则存在酉阵 Ω, 使得

$$A = \Omega T \Omega^H$$

其中 T 为上下角矩阵

♡

证明 课件 p3

2.5 线性矛盾方程组的最小二乘法

定义 2.5 (最小二乘问题)

设 A 是一个 $m \times n$ 的矩阵, $\boldsymbol{b}$ 是一个 m 维向量, 考虑确定一个向量 $\boldsymbol{x}$, 使函数 $\rho^2(\boldsymbol{x}) = ||\boldsymbol{b} - A\boldsymbol{x}||^2$ 达到极小 ♣

定理 2.10

最小二乘问题恒有解, 且解唯一的充要条件是

$$Null(A) = 0$$

即零空间只包括零向量

♡

证明 课件 p12

定理 2.11

$\boldsymbol{x}^* \in \boldsymbol{R}^n$ 是最小二乘问题解的充要条件是

$$A^T r = A^T(\boldsymbol{b} - A\boldsymbol{x}^*) = 0$$

或

$$A^T A \boldsymbol{x}^* = A^T \boldsymbol{b}$$

♡

证明 课件 p13

命题 2.3 (最小二乘解法)

1. 正规方程法
2. Householder 变换法

♠

例题 2.4 见课件 p15,16,17

2.6 方程组的条件问题和算法的确误差分析

定义 2.6

若方程组 $A\boldsymbol{x}=\boldsymbol{b}$ 中 A 或 $\boldsymbol{b}$ 的微小变换就会引起方程组的巨大变化，则称该方程组为**病态方程组** (ill-conditioned equation system),A 为**病态矩阵** (ill-conditioned matrix), 否则就称为良态方程组和良态矩阵 ♣

定义 2.7 (条件数)

对于非奇异矩阵 A , 其条件数记为 $cond(A)$，定义为

$$cond(A)=||A||||A^{-1}||$$

♣

$cond(A)$ 越大, 解的相对误差界可能越大, A 对求解线性方程组来说就越可能呈现病态. 但 $cond(A)$ 多大 A 才算病态, 通常没有具体的定量标准; $cond(A)$ 越小, 解的相对误差界越小, 反之, 呈现病态.

性质

1. $Cond(A)\geq 1$
2. $Cond(A)=Cond(A^{-1})$
3. $Cond(\alpha A)=Cond(A)$
4. $Cond(AB)\leq Cond(A)Cond(B)$

证明 见课件 p22,23

推论 2.1 (条件数等价性)

由矩阵范数的等价性可以推出, $\boldsymbol{R}^{n\times n}$ 上任意两个范数的条件数 $cond_\alpha(A), cond_\beta(A)$ 都是等价的, 即存在常数 c_1,c_2, 使得

$$c_1cond_\alpha(A)\leq cond_\beta(A)\leq c_2cond_\alpha(A)$$

这样, 一个矩阵在 α 范数下是病态的, 那么在 β 范数下也是病态的 ♡

例题 2.5 课件 p26

第 3 章 线性代数方程组的迭代解法

3.1 引言

定义 3.1

迭代解法，顾名思义，是给定一个初始解通过迭代一步步逼近精确解的方法，对于一个给定的线性方程组

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

这里我们记 $\{\boldsymbol{x}^{(k)}\}$ 为迭代序列，同时记 $\boldsymbol{e}^{(k)} = \boldsymbol{x}^{(k)} - \boldsymbol{x}^*$ 为误差向量（error vector），$\boldsymbol{r}^{(k)} = A\boldsymbol{x}^{(k)} - \boldsymbol{b}$ 为残差向量（residual vector）

♣

定理 3.1

若方程组 $A\boldsymbol{x} = \boldsymbol{b}$ 的系数矩阵 A 非奇异，则向量序列 $\{\boldsymbol{x}^{(k)}\}$ 收敛到方程组的解 $\boldsymbol{x}^*$ 的充要条件是残差序列 $\boldsymbol{r}^{(k)} = A\,\boldsymbol{x}^{(k)} - \boldsymbol{b}$ 收敛到零

♡

证明 课件 p7,8

定理 3.2

将线性方程组化为

$$\boldsymbol{x}^{(k+1)} = B\,\boldsymbol{x}^{(k)} + \boldsymbol{f}$$

线性迭代法对任意初始向量 $\boldsymbol{x}^0$ 和右端向量 $\boldsymbol{f}$ 收敛的充要条件是迭代矩阵的谱半径 $\rho(B) < 1$

♡

定义 3.2 (严格对角占优)

若对每个 $1 \leq i \leq n, |a_{ii}| > \sum_{i \neq j} |a_{ij}|$ 则称矩阵 $A = (a_{ij})$ 是严格对角占优的，换句话说，严格对角占优的矩阵每个主对角元在绝对值上要比所在行的其他所有元素的绝对值和更大

♣

定理 3.3 (Hadamard)

如果矩阵是奇异的，则至少存在一个对角元的绝对值不大于同行其他元素绝对值之和.

♡

证明 课件 p14

笔记 Hadamard 定理告诉我们严格对角阵占优矩阵一定非奇异.

定理 3.4

设 $||B|| < 1$ 则线性迭代法有如下估计式

$$||\boldsymbol{x}^k - \boldsymbol{x}^*|| \leq \frac{||B||}{1 - ||B||}||\boldsymbol{x}^k - \boldsymbol{x}^{k-1}||$$

或

$$||\boldsymbol{x}^k - \boldsymbol{x}^*|| \leq \frac{||B||}{1 - ||B||}||(||\boldsymbol{f}^k|| + 2||\boldsymbol{x}^0||)$$

♡

证明 课件 p19,20,21

3.2 常用的迭代方法

设线性方程组

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

则方程可化为

$$(D + L + U)\boldsymbol{x} = \boldsymbol{b}$$

以下方法会在该基础下展开

3.2.1 Jacobi 方法

Jacobi 方法是求解方程组的不动点迭代的一种形式. 不动点迭代 (FPI) 的第一步是改写方程组, 求解未知量. Jacobi 方法的第一步按下列标准方式进行: 求解第 i 个方程来得到第 i 个未知量; 然后和不动点迭代一样, 从某一初始估计开始进行迭代

定理 3.5 (Jacobi)

矩阵形式的 Jacobi 方法

$$\boldsymbol{x}_0 = \text{初始向量}$$

$$\boldsymbol{x}_{k+1} = -D^{-1}(L + U)\boldsymbol{x}_k + D^{-1}\boldsymbol{b} \qquad k = 0, 1, \cdots$$

笔记 证明很简单，将系数矩阵拆分后进行运算即可，运用了对角矩阵逆的易算性。Jacobi 方法的向量形式根据矩阵形式展开即可得到，迭代停止条件示需求而定。

例题 3.1 课件 p28

定理 3.6 (Frobenius-Mises)

若 $n \times n$ 矩阵 A 是严格对角占优的, 则成立:

1. A 为非奇异矩阵
2. 对每个向量 $\boldsymbol{b}$ 以及每一个初始向量，Jacobi 方法收敛到唯一解

证明 课件 p29,30

定理 3.7

以下两个命题是等价的, 设 A 具有正对角元且为对称阵

1. Jacobi 方法收敛
2. A 和 $2D - A$ 同为正定对称矩阵

证明 我们先做一些准备工作,$A = D + L + U$

$$B = -D^{-1}(L + U) = -D^{-1}(A - D) = -D^{-1}A + I \tag{3.1}$$

$$= I - D^{-1}AI = D^{-1/2}D^{1/2} - D^{-1/2}D^{-1/2}AD^{-1/2}D^{1/2} = D^{-1/2}(I - D^{-1/2}AD^{-1/2})D^{1/2} \tag{3.2}$$

$$2D - A = 2D^{1/}D^{1/2} - D^{1/2}D^{-1/2}AD^{-1/2}D^{1/2} = (D^{1/2})^T(2I - D^{-1/2}AD^{-1/2})D^{1/2} \tag{3.3}$$

必要性: 假设 Jacobi 方法收敛，那么 B 的谱半径 $\rho(B) < 1$, 令 $Q = D^{-1/2}AD^{-1/2}$, 根据上面的等式，B 与 $I - D^{-1/2}AD^{-1/2}$ 相似，具有相同的特征值, 那么可得

$$|\lambda(I - Q)| < 1 \Rightarrow |1 - \lambda(Q) < 1| \Rightarrow 0 < \lambda(Q) < 2$$

又因为 $Q^T = Q$ 那么 Q 对称. 因此 Q 正定，又因为 Q, A 合同，所以 A 正定；同理，$2D - A, 2I - Q$ 合同，$\lambda(2I - Q) = 2 - \lambda(Q) > 0$ 因此 $2D - A$ 正定。

充分性: 有如下关系链，A 对称正定，Q, A 合同 $\Rightarrow Q$ 对称正定 $\Rightarrow \lambda(Q) > 0$;　$2D - A$ 对称正定, $2I - Q, 2D - A$ 合同 $\Rightarrow 2I - Q$ 对称正定 $\Rightarrow \lambda(2I - Q) > 0 \Rightarrow \lambda(Q) < 2$
综上可得 $|\lambda(Q) - 1| < 1 \Rightarrow |\lambda(I - Q)| = |1 - \lambda(Q)| < 1$，因为 $B, I - Q$ 相似，因此 $\lambda(B) < 1$，即 $\rho(B) < 1$，Jacobi 方法收敛。

3.2.2 Gauss-Seidel 方法与 SOR

与 Jacobi 方法密切相关的一种迭代叫做 Gauss-Seidel 方法. Gauss-Seidel 方法与 Jacobi 方法之间仅有的差别是, 前者在每一步用到最新校正过的未知量的值, 即便是校正发生在当前步. 若 Gauss-Seidel 方法是收敛的, 它常常比 Jacobi 方法收敛得更快. 但它们是相互独立的, 即可以同时收敛与不收敛, 也可以收敛一个另外一个不收敛.

定理 3.8 (Gauss-Seidel)

矩阵形式的 Gauss-Seidel 方法

$$\boldsymbol{x}_0 = \text{初始向量}$$

$$\boldsymbol{x}_{k+1} = -(D+L)^{-1}U\boldsymbol{x}_k + (D+L)^{-1}\boldsymbol{b} \qquad k = 0, 1, \cdots$$

♡

定理 3.9 (Geiringor)

若 $n \times n$ 矩阵 A 是严格对角占优的，则成立：

1. A 是非奇异矩阵
2. 对每个向量 $\boldsymbol{b}$ 以及每一个初始向量，Gauss-Seidel 方法收敛到唯一解

♡

逐次超松弛 (Successive Over-Relaxation) 的方法采用 Gauss-Seidel 方法趋向与解的方向并试图加速收敛. 设 ω 是一实数，定义新估计量 x_{k+1} 的每个分量为 ω 乘上 Gauss-Seidel 公式与 $1 - \omega$ 乘当前估计量 x_k 的加权平均. ω 称为松弛参数 (relaxation parameter), $\omega > 1$ 时被称为超松弛 (over-relaxation), $\omega < 1$ 时被称为低松弛 (under-relaxation), $\omega = 1$ 时即为 Gauss-Seidel 方法

定理 3.10 (SOR)

矩阵形式的 SOR 方法

$$\boldsymbol{x}_0 = \text{初始向量}$$

$$\boldsymbol{x}_{k+1} = (D+\omega L)^{-1}[(1-\omega)D - \omega U]\boldsymbol{x}_k + (D+\omega L)^{-1}\omega\boldsymbol{b} \qquad k = 0, 1, \cdots$$

♡

笔记 ω 决定了迭代矩阵的谱半径, 使得其谱半径最小的参数 ω 称为最佳松弛因子

定理 3.11

若系数矩阵严格对角占优，则在 $0 < \omega \leq 1$ 时,SOR 方法收敛

♡

证明 采用反证法证明, 若系数矩阵 A 严格对角占优，设 SOR 方法不收敛. 那么迭代矩阵的谱半径 $\rho \geq 1$, 设迭代矩阵为 B .

$$B_\omega = (D+\omega L)^{-1}[(1-\omega)D - \omega U]$$

那么 $|\lambda(B)| \leq 1$, 下面求 B_ω 的谱半径

$$det(\lambda I - (D+\omega L)^{-1}[(1-\omega)D - \omega U]) = 0 \Rightarrow det(\lambda(D+\omega L) - (1-\omega)D + \omega U) = 0$$

进一步化简

$$det((\lambda+\omega-1)D + \lambda\omega L + \omega U) = 0 \Rightarrow det(D + \frac{\lambda\omega}{\lambda+\omega-1}L + \frac{\omega}{\lambda+\omega-1}U) = 0$$

令

$$Q = D + \frac{\lambda\omega}{\lambda+\omega-1}L + \frac{\omega}{\lambda+\omega-1}U$$

那么 Q 是奇异阵，根据 Hadamard 定理可得。Q 一定不是严格对角占优的，现在根据 Q 不是严格对角占优来证明 A 也不是严格对角占优的，从来导出矛盾。我们只需要证明

$$\left|\frac{\lambda\omega}{\lambda+\omega-1}\right| \le 1 \qquad \left|\frac{\omega}{\lambda+\omega-1}\right| \le 1$$

由于 $|\lambda| \ge 1$, 第二个不等式显然成立，现证第一个不等式成立，相当于证明

$$|\lambda\omega|^2 \le |\lambda+\omega-1|^2 \quad (|\lambda| \ge 1, 0 < \omega \le 1)$$

设 $\lambda = a + bi$, 那么

$$|\lambda\omega|^2 = \omega^2(a^2+b^2) \qquad |\lambda+\omega-1|^2 = |a+\omega-1|^2 + b^2$$

$$|\lambda+\omega-1|^2 - |\lambda\omega|^2 = (a^2+b^2)(1-\omega^2) + 2a(\omega-1) + (1-\omega)^2 = (1-\omega)[(a-1)^2 + b^2 + (a^2+b^2-1)\omega] \ge 0$$

因此得到 A 不是严格对角占优的，矛盾。

定理 3.12

若 A 实对称正定, 则 SOR 方法收敛的充要条件是 $0 < \omega < 2$ ♡

证明 课件 p7,8,9

3.3 共轭方向法

定义 3.3

Q 为 $n \times n$ 的实对称矩阵，对于方向 $d_0, d_1, \cdots, d_m$ 若对所有 $i \ne j$ 有 ${d_i}^T Q d_j = 0$, 则称它们是关于 Q 共轭的 ♣

引理 3.1

Q 为 $n \times n$ 的对称矩阵，若方向 $d_0, d_1, \cdots, d_m \in R^n, k \le n-1$ 非零且关于 Q 共轭，那么它们是线性无关的 ♡

证明 课件 p14

笔记 正交是共轭的特殊情况，在 n 维向量空间中，R^n 非零的共轭向量的个数不超过 n

定理 3.13 (线性方程组与极小值问题)

若 A 为 n 阶实对称正定矩阵, $b \in R^n$，则

$$Ax^* = b \iff f(x^*) = \min_{x \in R^n} f(x)$$

其中 $f(x) = \frac{1}{2}\langle x, Ax\rangle - \langle x, b\rangle$ ♡

证明 课件 p4,5

命题 3.1 (求二次函数极小点的一类方法)

1. 最速向量法
2. 共轭梯度法 (共轭斜量法) ♠

例题 3.2 最速向量法例题: 课件 p2

共轭方向法例题: 第四章课件 p3,p4

第 4 章 矩阵特征问题的形态

4.1 特征值的估计及极值性质

一般矩阵特征值的定位或估计

1. 对角矩阵或三角矩阵的特征值是对角线元素
2. 幂等矩阵的特征值为 0 或 1

   证明

$$
\begin{aligned}
A^2 &= A \\
A^2x &= Ax \\
\lambda^2 x &= \lambda x \\
(\lambda^2 - \lambda)x &= 0 \\
\lambda &= 0\ or\ 1
\end{aligned}
$$

3. Hermite 矩阵的特征值都是实数

   证明 关键在于证明 $\lambda = \overline{\lambda}$

   根据 $Ax = \lambda x$ 等式两端同取共轭转置再乘上 x，可以得到 $x^H Ax = \overline{\lambda}x$ 再取共轭转置后比较等式即可得证。

我们知道矩阵 A 的所有特征值都位于复平面上以原点为圆心，$||A||$ 为半径的的圆上，然而能否用更精确的方法来估计呢？答案是肯定的。

定理 4.1 (Gersgorin discs theorem)

设 A 为 $n \times n$ 的方阵，又设

$$
R_i{}'(A) = \sum_{j=1, j\neq i}^{n} |a_{ij}| \qquad 1 \leq i \leq n
$$

表示 A 的去心绝对行和，考虑到 n 个 Gersgorin 圆盘

$$
\{z \in \boldsymbol{C} : |z - a_{ii}| \leq R_i{}'(A)\} \qquad 1 \leq i \leq n
$$

矩阵 A 的所有特征值都位于 n 个 Gersgorin 圆盘的并上，即

$$
G(A) = \bigcup_{i=1}^{n} \{z \in \boldsymbol{C} : |z - a_{ii}| \leq R_i{}'(A)\}
$$

此外，若这 n 个圆盘中有 k 个圆盘的并形成一个连通区域，且它与余下的 $n-k$ 个圆盘都不相交, 则在这个区域中恰好有 k 个特征值

♡

证明 Horn p388, 课件 p11,p17,18

推论 4.1

对于去心绝对列和所形成的圆盘,Gersgorin discs theorem 依旧成立

♡

例题 4.1 课件 p12,13

推论 4.2

若 A 为 $n \times n$ 的方阵，则

$$
\rho(A) \leq min\left\{ \max_i \sum_{j=1}^{n} |a_{ij}|, \max_j \sum_{i=1}^{n} |a_{ij}| \right\}
$$

♡

证明 课件 p21

推论 4.3

Perron 若 A 为 $n \times n$ 的方阵, 且设 $p_1, \cdots, p_n$ 是正实数，则 A 的所有特征值位于区域

$$\bigcup_{i=1}^{n} \left\{ z \in \boldsymbol{C} : |z - a_{ii} \leq \frac{1}{p_i} \sum_{j \neq i} p_j |a_{ij}| \right\} = G(D^{-1}AD)$$

♡

证明 课件 p23

例题 4.2 课件 p25

定理 4.2 (Rayleigh)

设 A 为 Hermitian 矩阵，设其特征值按大小依次排列，设其 k 个标准正交特征向量为 x_i, S 为它们张成的空间, 那么

$$\lambda_1 = \min_{x \in S, x \neq 0} \frac{x^* A x}{x^* x} \leq \max_{x \in S, x \neq 0} \frac{x^* A x}{x^* x} = \lambda_k$$

♡

证明 课件 p27,28,29,30,31