第二章 线性代数方程组的数值解法

1 引言和线性代数基础知识

例 1.1. $Ax = b \Rightarrow x = A^{-1}b$
高斯消元法复杂度太高 $O(n^3)$，在 A 规模很大时，时间消耗过于巨大。

1.1 线性空间

定义 1.1. 定义了加减和数乘的非空集合，称为线性空间。

如 $R^n, C^n, R^{n\times m}(C^{n\times m}), C[a,b]$(定义在 [a,b] 上的连续函数)。

1.2 内积

定义 1.2. 对于定义于数域 $K(R\ or\ C)$ 上的线性空间 X，如果对于 $\forall u, v, w \in X$ 及 $\alpha \in K$，满足：

- $(u+v, w) = (u, w) + (v, w)$
- $(\alpha u, v) = \alpha(u, v)$
- $(u, v) = \overline{(v, u)}$
- $(u, u) \geq 0\ (u, u) = 0 \Leftrightarrow u = 0$

则称数 (u, v) 为 u 与 v 的内积，定义了内积的线性空间称为内积空间。

几种常见内积的定义：

- R^n 和 C^n 上的内积：
  设 $x, y \in R^n$，则 R^n 的内积可定义为：

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

若给定权系数 $w_i(i=1,2,\dots,n)$, 则 R^n 上带权 $\{w_i\}$ 的内积可定义为:

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

类似的，若 $x,y\in C^n$，则带权的内积:

$$(x,y)=\sum_{i=1}^{n} w_i x_i \overline{y_i}$$

- C[a,b] 上的内积:

  定义 1.3. *若定义在 $[a,b]$ 上的函数 $\rho(x)$ 满足:*

  - $\rho(x)\geq 0, \forall x\in(a,b)$
  - $\int_a^b x^k\rho(x)dx, \exists$ 且有限 $(k=0,1,2,\dots)$
  - 若对 $[a,b]$ 上的非负连续函数 $g(x)$ 有:
    $\int_a^b \rho(x)g(x)dx=0$，则 $g(x)\equiv 0$

  就称 $\rho(x)$ 为 $[a,b]$ 上的一个权函数。

  设 $f,g\in C[a,b]$，ρ 是 $[a,b]$ 上给定的权函数，则称:
  $(f,g)=\int_a^b \rho(x)f(x)g(x)dx$ 为 $C[a,b]$ 上函数 f,g 的内积。

1.3 内积空间的几个性质

- (Cauchy-Schwarz 不等式): 设 X 为一个内积空间, 则对 $\forall u,v\in X$ 有:

$$|(u,v)|^2\leq (u,u)\cdot(v,v) \tag{1.1}$$

- 设 X 为一个内积空间，$u_1,u_2,\dots,u_n\in X$，矩阵:

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

  则 G 非奇异的充要条件是 $u_1,u_2,\dots,u_n$ 线性无关

- (Gram-Schmidz 正交化方法)：若 $\{u_1, u_2, \ldots, u_n\}$ 是内积空间 X 中一个线性无关的序列，则可按下列公式：

$$\begin{cases} v_1 = u_1, \\ v_i = u_i - \sum_{k=1}^{i-1} \frac{(u_i, v_k)}{(v_k, v_k)} v_k, \quad i = 2, 3, \ldots, n \end{cases} \tag{1.2}$$

  构造一个正交序列 $\{v_1, v_2, \ldots, v_n\}$ 且满足：$(v_i, v_j) = 0, (i \neq j)$。

1.4 向量范数

例 1.2. 欧式范数或 *2*-范数：

$$||\mathbf{x}||_2 = \sqrt{(\mathbf{x}, \mathbf{x})} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

定义 1.4. 设对 $\forall \mathbf{x} \in R^n (or\ C^n)$，按一定的规则有一实值函数与之对应，记为 $||\mathbf{x}||$，若满足：

- 正定性：$||\mathbf{x}|| \geq 0$ 且 $||\mathbf{x}|| = 0 \Leftrightarrow x = 0$
- 齐次性：$||\alpha \mathbf{x}|| = |\alpha| \cdot ||\mathbf{x}||, \forall \alpha \in R(or\ C)$
- 三角不等式

$$||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||, \forall \mathbf{x}, \mathbf{y} \in R^n (or\ C^n) \tag{1.3}$$

则称 $||\mathbf{x}||$ 为向量 $\mathbf{x}$ 的范数，由（*1.3*）易证 $|||\mathbf{x}|| - ||\mathbf{y}||| \leq ||\mathbf{x} - \mathbf{y}||$ *(*自证*)*

证明. 范数三角不等式减法形式：
不妨设：$||\mathbf{x}|| \geq ||\mathbf{y}||$，则：

$$\begin{aligned} & |||\mathbf{x}|| - ||\mathbf{y}||| \\ =& ||\mathbf{x}|| - ||\mathbf{y}|| \\ =& ||\mathbf{x} - \mathbf{y} + \mathbf{y}|| - ||\mathbf{y}|| \\ \leq& ||\mathbf{x} - \mathbf{y}|| + ||\mathbf{y}|| - ||\mathbf{y}|| \\ =& ||\mathbf{x} - \mathbf{y}|| \end{aligned}$$

□

几种常见范数：

- $||\mathbf{x}||_\infty = \max \lim_{1 \leq i \leq n} |\mathbf{x}_i|$, ∞-范数
- $||\mathbf{x}||_1 = \sum_{i=1}^{n} |x_i|$，1-范数

- $||\mathbf{x}||_2 = (\sum_{i=1}^n x_i^2)^{\frac{1}{2}} = \sqrt{(x,x)}$，2-范数

证明. 二范数三角不等式：

$$
\begin{aligned}
&||\mathbf{x}+\mathbf{y}||_2 \\
=&(\mathbf{x}+\mathbf{y},\mathbf{x}+\mathbf{y})^{\frac{1}{2}} \\
=&[(\mathbf{x},\mathbf{x})+2(\mathbf{x},\mathbf{y})+(\mathbf{y},\mathbf{y})]^{\frac{1}{2}} \\
\le&[||\mathbf{x}||_2^2+2||\mathbf{x}||_2\cdot||\mathbf{y}||_2+||\mathbf{y}||_2^2]^{\frac{1}{2}} \\
=&||\mathbf{x}||_2+||\mathbf{y}||_2
\end{aligned}
$$

□

例 1.3. 思考题：对于任一种范数 $||\cdot||_\lambda$ 定义集合：

$$A = \{x|x \in R^3, ||x|| \le 1\}$$

试问当 $\lambda = 1, 2, \infty$ 时，A 代表什么样的图形？并画图

$$||x||_p = (\sum_{i=1} |x_i|^p)^{\frac{1}{p}}, p \in [1,\infty) \tag{1.4}$$

1.5 向量范数的性质

- 连续性：

  定理 1.1. 设给定 $A \in R^{n\times n}$，则对 R^n 上的每一种向量范数 $||\cdot||, x \in R^n$，$||Ax||$ 都是 x 的分量 $(x_1, x_2, \ldots, x_n)$ 的 n 元连续函数 $f(x_1, x_2, \ldots, x_n)$

- 等价性：

  定义 1.5. 线性空间 X 上定义了两种范数 $||\cdot||_\alpha$ 和 $||\cdot||_\beta$，若 $\exists$ 常数 $C_1, C_2 > 0$，使得

  $$C_1 \cdot ||\mathbf{u}||_\alpha \le ||\mathbf{u}||_\beta \le C_2 \cdot ||\mathbf{u}||_\alpha, \forall \mathbf{u} \in X$$

  则称 $||\cdot||_\alpha$ 和 $||\cdot||_\beta$ 是 X 上等价的范数。

- 传递性：
  若 $||\cdot||_\alpha$ 与 $||\cdot||_\beta$ 等价，$||\cdot||_\beta$ 与 $||\cdot||_\lambda$ 等价，则 $||\cdot||_\alpha$ 与 $||\cdot||_\lambda$ 等价。

  定理 1.2. R^n 上所有范数是彼此等价的。

1.6 矩阵范数

定义 1.6. 如果对 $R^{n\times n}$ 上的任一矩阵 A，对应一个实数 $||A||$，满足条件：$\forall A,B\in R^{n\times n}$ 和 $\alpha\in R$，

- $||A||>0$，且 $||A||=0\Leftrightarrow A=0$ （正定性）
- $||\alpha A||=|\alpha|\cdot||A||$ （齐次性）
- $||A+B||\leq||A||+||B||$ （三角不等式）
- $||AB||\leq||A||\cdot||B||$ （相容性）

则称 $||A||$ 为矩阵 A 的范数

易证：$|||A||-||B|||\leq||A-B||$（自证）

F-范数

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

性质 4 利用柯西不等式证明（和的平方 $\leq$ 平方的和）

定义 1.7. 对于 $R^{n\times n}$ 上给定的一种矩阵范数：
$||\cdot||_\beta$,R^n 上规定的一种向量范数 $||\cdot||_\alpha$，若有：

$$||A\mathbf{x}||_\alpha\leq||A||_\beta||\mathbf{x}||_\alpha,\forall A\in R^{n\times n},\mathbf{x}\in R^n \tag{1.5}$$

成立，则称上述矩阵范数和向量范数是相容的。或：

$$||A||_\beta \overset{\text{sup}}{=\!=} \frac{||A\mathbf{x}||_\alpha}{||\mathbf{x}||_\alpha} \tag{1.6}$$

定理 1.3. 设 $A\in R^{n\times n}$，$||\cdot||$ 是 R^n 上的向量范数，则：

$$||A||=\max_{||x||=1}||Ax||=\max_{x\neq 0}\frac{||Ax||}{||x||} \tag{1.7}$$

是一种矩阵范数，称其为由向量范数诱导出的矩阵范数（简称为诱导范数，或从属范数，自然范数），且满足相容性条件。

证明.
- 由定理1.1可知,$||Ax||$ 是 R^n 中有界闭集 $D=\{\mathbf{x}=(x_1,x_2,\ldots,x_n)^T,||x||=1\}$ 上的连续函数,$\therefore$ $||Ax||$ 在 D 上最大值存在,$\therefore$(1.7) 中 $\max_{||x||=1}||Ax||$ 的写法正确。

- 对于 $\forall x \neq 0$,

$$\because \frac{||Ax||}{||x||} = ||A \cdot \frac{x}{||x||}||, \text{且}||\frac{x}{||x||}|| = 1$$

$$\therefore \max_{x \neq 0} \frac{||Ax||}{||x||} = \max_{x \neq 0} ||A \cdot \frac{x}{||x||}|| \xlongequal{let y=\frac{x}{||x||}} \max_{||y||=1} ||Ay|| \xlongequal{\text{记 y 为 x}} \max_{||x||=1} ||Ax||$$

- – 正定性：显然 $||A|| \geq 0$，另外：
  若 $A = 0$，则：$||A|| = \max_{||x||=1} ||Ax|| = 0$，反之，若 $||A|| = 0$，i.e. 对 $\forall$ 非零向量 x 有：

$$||Ax|| = 0 \Rightarrow Ax = 0 \Rightarrow A = 0$$

  – 齐次性：对 $\forall \alpha \in R$，有：

$$\begin{aligned} & ||\alpha A|| \\ =& \max_{||x||=1} ||\alpha Ax|| \\ =& \max_{||x||=1} |\alpha| ||Ax|| \\ =& |\alpha| \max_{||x||=1} ||Ax|| \\ =& |\alpha| \cdot ||A|| \end{aligned}$$

  – 三角不等式：$\forall A, B \in R^{n \times n}$，有：

$$\begin{aligned} & ||A+B|| \\ =& \max_{||x||=1} ||(A+B)x|| \\ =& \max_{||x||=1} ||Ax + Bx|| \\ \leq& \max_{||x||=1} (||Ax|| + ||Bx||) \\ \leq& ||A|| + ||B|| \end{aligned}$$

  – 相容性：由1.7显然有 $||Ax|| \leq ||A|| ||x||, \forall x \in R^n$,
  $\therefore A, B \in R^{n \times n}$, 有

$$
\begin{aligned}
&||AB|| \\
&= \max_{||x||=1} ||ABx|| \\
&= \max_{||x||=1} ||A(Bx)|| \\
&\le \max_{||x||=1} ||A|| \cdot ||Bx|| \\
&\le ||A|| \max_{||x||=1} ||Bx|| \\
&= ||A|| \cdot ||B||
\end{aligned}
$$

综上所述，$||A||$ 是 $R^{n\times n}$ 上的一种矩阵范数。

由定义及 $||A||$ 是 $R^{n\times n}$ 是一种范数，即可知其相容性条件成立。 □

注释：

- 对 $\forall$ 的从属范数，$||I|| = 1$
- 矩阵的任一从属范数一定与所给定的向量范数相容，但相容未必具有从属关系。如：$||Ax||_2 \le ||A||_F \cdot |||x||_2$，但不具有从属关系。
事实上，$||I||_F = \sqrt{n}$，另外，$||I||_F = \max_{x\neq 0} \frac{||Ix||}{||x||} = 1$
$\therefore ||A||_F$ 与 $||x||_2$ 没有关系。

定理 1.4. 设 $x \in R^n, A \in R^{n\times n}$，则：

- $||A||_\infty = \max_{||x||_\infty=1} ||Ax||_\infty = \max_{1\le i\le n} \sum_{j=1}^n |a_{ij}|$ *(∞-范数 or 行范数)*
- $||A||_1 = \max_{||x||_1=1} ||Ax||_1 = \max_{1\le j\le n} \sum_{i=1}^n |a_{ij}|$ （*1-范数 or 列范数*）
- $||A||_2 = \max_{||x||_2=1} ||Ax||_2 = \sqrt{\lambda_{\max(A^TA)}}$ （*2-范数 or 谱范数*）
其中，$\lambda_{\max(A^TA)} = \rho(A^TA)$*(*$\rho$ *表示谱半径)* 表示 A^TA 的最大特征值。
一般的，$\rho(A) = \max_{1\le i\le n} |\lambda_i|$，$\lambda_i$ 为 A 的特征值。

证明.
- 若 $A = 0$ 时，显然成立，不妨设 $A \neq 0$
$\forall x \in R^n$，且 $||x||_\infty = 1$，则：

$$
||Ax||_\infty = \max_{1\le i\le n} |\sum_{j=1}^n a_{ij}x_j| \le \max_i \sum_{j=1}^n |a_{ij}||x_j| \le \max_i \sum_{j=1}^n |a_{ij}| \xlongequal{\text{记}} \mu
$$

$\therefore \max_{||x||_\infty=1} ||Ax||_\infty \le \mu$
事实上，假设第 i_0 行使 $\mu = \sum_{j=1} |a_{i_0j}|$ 成立，
取 $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, \dots, x_n^{(0)})^T$，其中：

$$
x_j^{(0)} = \begin{cases} 1, a_{i_0j} \ge 0 \\ -1, a_{i_0j} < 0 \end{cases} \quad j = 1, 2, \dots, n
$$

显然 $||x^{(0)}||_\infty = 1$ 且

$$\max_{||x||_\infty=1} ||Ax||_\infty \geq ||Ax^{(0)}||_\infty = \max_i |\sum_{j=1}^n a_{ij}x_j^{(0)}| \geq |\sum_{j=1}^n a_{i_0j}x_j^{(0)}| = \sum_{j=1}^n |a_{i_0j}| = \mu$$

$\therefore \max_{||x||_\infty=1} ||Ax||_\infty = \mu = \max_{1\leq i\leq n} \sum_{j=1}^n |a_{ij}|$

- $\because ||Ax||_2^2 = (Ax, Ax) = x^TA^TAx \geq 0, \forall x \in R^n$
  $\therefore A^TA$ 为非负的对称实矩阵，其特征值均为非负的实数，不妨假设依次排序为：$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$
  对应于一组规范的正交特征向量组 $\{u_1, \ldots, u_n\}$
  对于 $\forall x \in R^n$，可表示为：$x = \sum_{i=1}^n \alpha_i u_i$
  对于满足 $||x||_2 = 1$ 的任意 x 有：

$$||x||_2^2 = (x, x) = \sum_{i=1}^n \alpha_i^2 = 1$$

$$\begin{aligned}
&\therefore ||Ax||_2^2 = (Ax, Ax) = x^TA^TAx = (A^TAx, x) \\
&= (A^TA\textstyle\sum_{i=1}^n \alpha_i u_i, \sum_{i=1}^n \alpha_i u_i) \\
&= (\textstyle\sum_{i=1}^n \alpha_i\lambda_i u_i, \sum_{i=1}^n \alpha_i u_i) \\
&= \textstyle\sum_{i=1}^n \lambda_i\alpha_i^2 \\
&\leq \lambda_1 \textstyle\sum_{i=1}^n \alpha_i^2 \\
&= \lambda_1 \\
&\therefore \max_{||x||_2=1} ||Ax||_2 \leq \sqrt{\lambda_1}
\end{aligned}$$

  特别地，若取 $x = u_1$，则：
  $||Au_1||_2^2 = (A^TAu_1, u_1) = \lambda_1$，$\therefore ||Au_1||_2 = \sqrt{\lambda_1}$
  $\max_{||x||_2=1} ||Ax||_2 \geq ||Au_1||_2 = \sqrt{\lambda_1}$
  $\therefore \max_{||x||_2=1} ||Ax||_2 = \sqrt{\lambda_1} = \sqrt{\lambda_{\max(A^TA)}}$

□

推论（自证）若 $A \in R^{n\times n}$ 为对称阵，则 $||A||_2 = \rho(A)$
$A^T = A \Rightarrow \lambda_{A^TA} = \lambda_A^2$

定理 1.5. 设 $||x||_\alpha$ 为 R^n 上任一种向量范数，从属于它的矩阵范数记为：$||A||_\alpha$，并设 $P \in R^{n\times n}, det(P) \neq 0$，则：

- R^n 到 R 的映射 $P : x \Rightarrow ||Px||_\alpha$ 定义了 R^n 上另一种向量范数，记为：$||x||_{P,\alpha} = ||Px||_\alpha$

- 从属向量范数 $||x||_{P,\alpha}$ 的矩阵范数为：

$$||A||_{P,\alpha} = ||PAP_\alpha^{-1}||$$

1.7 矩阵范数的性质

定理 1.6. • 设 $||\cdot||$ 为 $R^{n\times n}$ 上任一种（从属或非从属）的矩阵范数，则对 $\forall A\in R^{n\times n}$，有：

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

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

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

证明. • 设 $x\neq 0, Ax=\lambda x$，且 $|\lambda|=\rho(A)$，则必存在向量 $y\in R^n$，使得 $xy^T\neq 0$，于是有：

$$\rho(A)||xy^T||=||\lambda xy^T||=||Axy^T||\le ||A||\cdot||xy^T||\Rightarrow \rho(A)\le||A||$$

• 对 $\forall A\in R^{n\times n},\exists T\ s.t.\ TAT^{-1}=J$ 为 Jordan 标准形，其中 $J=diag(J_1,J_2,\dots,J_s)$，$J_i$ 为 Jordan 块：

$$J_i=\begin{bmatrix}\lambda_i & 1 & \ddots\\ & \lambda_i & 1\\ \ddots & \ddots & \lambda_i\end{bmatrix}, i=1,2,\dots,s$$

任取 $\varepsilon>0$，并定义 $D_\varepsilon\in R^{n\times n}$ 为：

$$D_\varepsilon=diag(1,\varepsilon,\dots,\varepsilon^{n-1})$$

易证，$D_\varepsilon^{-1}JD_\varepsilon$ 仍为块对角阵，且分块与 J 相同，即：$\widetilde{J}=D_\varepsilon^{-1}JD_\varepsilon=diag(\widetilde{J}_1,\widetilde{J}_2,\dots,\widetilde{J}_s)$ 其中：

$$\widetilde{J}_i=\begin{bmatrix}\lambda_i & \varepsilon & \ddots\\ & \lambda_i & \varepsilon\\ \ddots & \ddots & \lambda_i\end{bmatrix}, i=1,2,\dots,s$$

$$\therefore\ ||\widetilde{J}||_\infty=||D_\varepsilon^{-1}JD_\varepsilon||_\infty\le\max_i|\lambda_i|+\varepsilon=\rho(A)+\varepsilon$$

而 $D_\varepsilon^{-1}T=P$ 为非异阵，由定理1.5可知：$||D_\varepsilon^{-1}Tx||_\infty$ 定义了 R^n 上的一种向量范数 $||x||_{P,\infty}$ 且从属于该向量范数的矩阵范数为：

$$||A||=||D_\varepsilon^{-1}TAT^{-1}D_\varepsilon||_\infty=||D_\varepsilon^{-1}JD_\varepsilon||_\infty\le\rho(A)+\varepsilon$$

□

定理 1.7. 对于 $R^{n\times n}$ 的任意两种范数 $||\cdot||_\alpha$ 和 $||\cdot||_\beta$，存在常数 M 和 $m(M \geq m > 0)$ 使得：

$$m||A||_\alpha \leq ||A||_\beta \leq M||A||_\alpha, \forall A \in R^{n\times n}$$

即 $R^{n\times n}$ 上所有矩阵范数是等价的。

定理 1.8. 设 $||\cdot||$ 是 $R^{n\times n}$ 上的算子范数，若 $||B|| \leq 1$，则 $I \pm B$ 为非异阵，且 $||(I+B)^{-1}|| \leq \frac{1}{1-||B||}$

证明. (反证法)
若 $det(A) = 0$ 则 $(I-B)x = 0$ 有非零解，$i.e.$ $\exists x_0 \neq 0$使$Bx_0 = x_0 \Rightarrow \frac{||Bx_0||}{||x_0||} = 1$，而 $||B|| = \max_{x\neq 0} \frac{||Bx||}{||x||} \geq \frac{||Bx_0||}{||x_o||} = 1$，与已知矛盾。
$\therefore$ $(I-B)$ 是非异阵，同理可证 $I+B$ 也为非异阵。
记 $D = (I-B)^{-1}$，则：

$$1 = ||I|| = ||D(I-B)|| = ||D - DB|| \geq ||D|| - ||DB|| \geq$$
$$||D|| - ||D|| \cdot ||B|| = ||D|| \cdot (1 - ||B||) \Rightarrow ||(I-B)^{-1}|| \leq \frac{1}{1-||B||}$$

□

2 Gauss 消去法和矩阵的 LU 分解

设

$$Ax = b \tag{2.1}$$

其中：

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}, x = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

且 $det(A) \neq 0$，记 $[A|b] = [A^{(1)}|b^{(1)}]$ 其中 $A^{(1)} = (a_{ij}^{(1)})$，$b^{(1)} = (b_i^{(1)})$

2.1 Gauss 消去法

2.1.1 消元过程

- 设 $a_{11}^{(1)} \neq 0$，并且选取 $a_{11}^{(1)}$ 作为主元，计算乘数 $l_{i1} = a_{i1}^{(1)}/a_{11}^{(1)}(i = 2, \dots, n)$；用 $-l_{i1}$ 乘以第 1 行并加到第 i 行 $(i = 2, 3, \dots, n)$，则

$[A^{(1)}|b^{(1)}] \Rightarrow [A^{(2)}|b^{(2)}]$

$$[A^{(2)}|b^{(2)}] = \left[\begin{array}{cccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} & b_n^{(2)} \end{array}\right]$$

其中，

$$\begin{cases} a_{ij}^{(2)} = a_{ij}^{(1)} - l_{i1}a_{1j}^{(1)}, & i,j = 2,3,\dots,n \\ b_i^{(2)} = b_i^{(1)} - l_{i1}b_i^{(1)}, & i = 2,3,\dots,n \end{cases}$$

对应的方程 $A^{(2)}x = b^{(2)} \Leftrightarrow Ax = b$

- 假设消元过程已进行了 $k-1$ 步，得到方程组 $A^{(k)}x = b^{(k)}$，其对应的增广矩阵为：

$$[A^{(k)}|b^{(k)}] = \left[\begin{array}{cccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \ddots & & & & \vdots \\ & & & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} & b_k^{(k)} \\ & & & \vdots & \dots & \dots & \vdots \\ & & & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} & b_n^{(k)} \end{array}\right] \tag{2.2}$$

设 $a_{kk}^{(k)} \neq 0$，并选 $a_{kk}^{(k)}$ 为主元素，计算 $l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}}, i = k+1,\dots,n$，分别用 $-l_{ik}$ 乘以第 k 行并加到第 k+1 至第 n 行，则2.2式可变为：

$$[A^{(k+1)}|b^{(k+1)}] = \left[\begin{array}{cccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ & \ddots & & & & & \vdots \\ & & a_{kk}^{(k)} & a_{k,k+1}^{(k)} & \dots & a_{kn}^{(k)} & b_k^{(k)} \\ & & & a_{k+1,k+1}^{(k+1)} & \dots & a_{k+1,n}^{(k+1)} & b_{k+1}^{(k+1)} \\ & & & \vdots & & & \vdots \\ & & & a_{n,k+1}^{(k+1)} & \dots & a_{nn}^{(k+1)} & b_n^{(k+1)} \end{array}\right]$$

其中，

$$\begin{cases} a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik}a_{kj}^{(k)}, & i,j = k+1,\dots,n \\ b_i^{(k+1)} = b_i^{(k)} - l_{ik}b_k^{(k)}, & i = k+1,\dots,n \end{cases}$$

对应的方程组为 $A^{(k+1)}x = b^{(k+1)}$

- 继续这一过程且 $a_{ii}^{(i)} \neq 0, i = 1,2,\dots,n-1$（称为主元），直至所有主

对角线以下的元素消为 0，即可得上三角方程组 $A^{(n)}x = b^{(n)}$

$$[A^{(n)}|b^{(n)}] = \left[\begin{array}{ccccc|c} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & a_{1n}^{(1)} & b_1^{(1)} \\ & a_{22}^{(2)} & \dots & \dots & a_{2n}^{(2)} & b_2^{(2)} \\ & & \ddots & & \vdots & \vdots \\ & & & \ddots & \vdots & \vdots \\ & & & & a_{nn}^{(n)} & b_n^{(n)} \end{array}\right]$$

2.1.2 回代过程

$A^{(n)}x = b^{(n)}$ 的解即为 $Ax = b$ 的解为：

$$\begin{cases} x_n = b_n^{(n)}/a_{nn}^{(n)} \\ x_k = (b_k^{(k)} - \sum_{j=k+1}^{n} a_{kj}^{(k)} x_j)/a_{kk}^{(k)}, \quad k = n-1, n-2, \dots, 1 \end{cases}$$

定理 2.1. $a_{ii}^{(i)} \neq 0 (i = 1, 2, \dots, k)$ 的充要条件是 A 的顺序主子式 $D_i \neq 0, i = 1, \dots, k, k \leq n$

归纳法.：

- 充分性：当 $k = 1$ 时，显然成立，
  假设定理对 $k-1$ 成立. i.e. 由 $D_i \neq 0, (i = 1, 2, \dots, k-1) \Rightarrow a_{ii}^{(i)} \neq 0, i = 1, 2, \dots, k-1$
  证对 k 亦成立，i.e. 证 $D_i \neq 0, i = 1, 2, \dots, k \Rightarrow a_{ii}^{(i)} \neq 0, i = 1, 2, \dots, k$
  由归纳假设 $a_{ii}^{(i)} \neq 0, i = 1, 2, \dots, k-1$ 于是 Gauss 消去法将 $A^{(1)} = A$ 化为 $A^{(k)}$，即

$$A^{(1)} \to A^{(k)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & \dots & \dots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & \ddots & & & \\ & & & a_{kk}^{(k)} & \dots & a_{kn}^{(k)} \\ & & & \vdots & \dots & \dots \\ & & & a_{nk}^{(k)} & \dots & a_{nn}^{(k)} \end{bmatrix}$$

  由线性代数知识，消去过程中的行变换不影响 A 的顺序主子式的值。

$$D_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} = a_{11}^{(1)} a_{22}^{(2)} \cdots a_{kk}^{(k)} \neq 0 \tag{2.3}$$

  由假设 $D_i \neq 0, i = 1, 2, \dots, k, \therefore$ 由上式得，

$$a_{ii}^{(i)} \neq 0, i = 1, 2, \dots, k$$

- 必要条件：由2.3式可证。

□

推论：若 A 的顺序主子式 $D_i \neq 0, i = 1, 2, \ldots, n-1$, 则 $a_{11}^{(1)} = D_1, a_{ii}^{(i)} = D_i/D_{i-1}, i = 2, 3, \ldots, n$（由2.3易知）

定理 2.2. 对于 $Ax = b$，其中 A 非异阵，若 A 的顺序主子式 $D_i \neq 0, i = 1, 2, \ldots, n-1$，则可用 *Gauss* 消去法求出方程组的解。

定理 2.3. A 对称正定 $\Rightarrow a_{kk}^{(k)} > 0, k = 1, 2, \ldots, n$

定理 2.4. A 为严格对角占优阵 $\Rightarrow a_{kk}^{(k)} \neq 0, k = 1, 2, \ldots, n$

定义 2.1 (严格对角占优阵)**.**

$$|a_{kk}| > \sum_{j=1, j\neq k}^{n} |a_{kj}|, k = 1, 2, \ldots, n$$

2.2 矩阵的 LU 分解

定义 2.2 (LU 分解)**.** $A = LU$，L 为单位下三角矩阵，U 为上三角矩阵

$$Ax = b \Leftrightarrow LUx = b \Leftrightarrow \begin{cases} Ly = b \\ Ux = y \end{cases}$$

第 k 步消元等价于用下列矩阵：

$$L_k = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & -l_{k+1,k} & 1 & & \\ & \vdots & & \ddots & \\ & -l_{nk} & & & 1 \end{bmatrix}$$

左乘 $[A^{(k)}|b^{(k)}]$, i.e.

$$L_k[A^{(k)}|b^{(k)}] = [A^{(k+1)}|b^{(k+1)}]$$

第 $n-1$ 步消元等价于用 L_{n-1} 左乘 $[A^{(n-1)}|b^{(n-1)}]$, i.e.

$$L_{n-1}[A^{(n-1)}|b^{(n-1)}] = [A^{(n)}|b^{(n)}]$$

因此经过 $n-1$ 次消元后，即矩阵 $[A^{(n)}|b^{(n)}]$ 左乘 $L_1, L_2, \ldots, L_{n-1}$ 有：

$$[A^{(n)}|b^{(n)}] = L_{n-1}[A^{(n-1)}|b^{(n-1)}] = \cdots = L_{n-1}L_{n-2}\ldots L_1[A^{(1)}|b^{(1)}]$$

$$\therefore U \triangleq A^{(n)} = L_{n-1}L_{n-2}\dots L_1 A^{(1)}, b^{(n)} = L_{n-1}L_{n-2}\dots L_1 b^{(1)}$$

$$\Rightarrow L_1^{-1}L_2^{-1}\dots L_{n-1}^{-1}U \triangleq LU(\text{Doolittle 分解})$$

易证:

$$L_k^{-1} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & l_{k+1,k} & 1 & & \\ & \vdots & & \ddots & \\ & l_{nk} & 1 & & 1 \end{bmatrix}$$

$$\therefore L = L_1^{-1}L_2^{-1}\dots L_{n-1}^{-1} = \begin{bmatrix} 1 & & & & \\ l_{21} & \ddots & & & \\ l_{31} & l_{3,2} & 1 & & \\ \vdots & \vdots & & \ddots & \\ l_{n1} & l_{n2} & 1 & & 1 \end{bmatrix}, |A| \neq 0$$

定理 2.5 (LU 分解定理)**.** 非奇异阵 $A \in R^{n\times n}$, 若其顺序主子式 $D_i \neq 0(i = 1, 2, \dots, n-1)$, 则存在唯一的单位下三角阵 L 和上三角阵 U, 使得 $A = LU$

反证法. 假设 A 的 LU 分解不唯一, i.e. $\exists A = L_1U_1 = L_2U_2 \Rightarrow U_1 = L_1^{-1}L_2U_2 \Rightarrow U_1U_2^{-1} = L_1^{-1}L_2 = I \Rightarrow L_1 = L_2, U_1 = U_2$ □

注释:

- 当 A 为奇异阵，且 $D_i \neq 0(i = 1, 2, \dots, n-1)$ 时，定理2.5仍成立。
- A 的 k 阶 $(k = 1, \dots, n)$ 顺序主子式

$$D_k = \begin{vmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{k1} & \dots & a_{kk} \end{vmatrix} = a_{11}^{(1)}a_{22}^{(2)}\cdots a_{kk}^{(k)} = u_{11}u_{22}\cdots u_{kk}$$

- 若 A 的顺序主子式 $D_k \neq 0(k = 1, 2, \dots, n)$，则存在唯一的下三角阵 $\widetilde{L}$ 和单位上三角阵 $\widetilde{U}$ 使得:

$$A = \widetilde{L}\widetilde{U}$$

称为 Crout 分解。

-

定理 2.6 (LDU 分解定理). 设 $A \in R^{n\times n}$，则存在唯一的单位下、上三角阵 L、U 及对角阵 D 使得 $A = LDU$ 的充要条件是 A 的顺序主子式 $D_k \neq 0, k = 1, 2, \ldots, n-1$，$A = L\widetilde{U}$，而 $\widetilde{U} = DU$

3 主元素消去法

3.1 列主元 Gauss 消去法

Gauss 消去过程：

$$A^{(k)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \ldots & \ldots & \ldots & a_{1n}^{(1)} \\ & a_{22}^{(2)} & & & & \vdots \\ & & \ddots & & & \vdots \\ & & & a_{kk}^{(k)} & \ldots & a_{kn}^{(k)} \\ & & & \vdots & \ldots & \ldots \\ & & & a_{nk}^{(k)} & \ldots & a_{nn}^{(k)} \end{bmatrix}$$

需要列主元的两种情况：

- $a_{kk}^{(kk)} = 0$
- $|a_{kk}^{(kk)}|$ 很小，$l_{ik} = \frac{a_{ik}^{(k)}}{a_{kk}^{(k)}} > 1$ 会放大误差

3.1.1 消去过程

对 $k = 1, 2, \ldots, n-1$ 做：

- 选列主元：
$$|a_{i_k k}^{(k)}| = \max_{k \leq i \leq n} |a_{ik}^{(k)}|$$
- 若 $a_{i_k k}^{(k)} = 0$，则停止计算 ($\because det(A) = 0$), 否则：
- 若 $i_k \neq k$，则换行：$a_{kj}^{(k)} \leftrightarrow a_{i_k k}^{(k)}$，$b_k^{(k)} \leftrightarrow b_{i_k}^{(k)}$
- 消元：对 $i = k+1, \ldots, n$ 做 (a) - (c)
$$(a) l_{ik} = a_{ik}^{(k)} / a_{kk}^{(k)}$$
  对 $j = k+1, \ldots, n$ 做 (b)
$$(b) a_{ij}^{(k+1)} = a_{ij}^{(k)} - l_{ik} a_{kj}^{(k)}, b_i^{(k+1)} = b_i^{(k)} - l_{ik} b_k^{(k)}$$

3.1.2 回代过程

- 若 $a_{nn}^{(n)} = 0$，则停止计算 ($\because det(A) = 0$)，否则
- $x_n = b_n^{(n)}/a_{nn}^{(n)}$
- 对 $i = n-1, n-2, \ldots, 1$ 做:

$$x_i = (b_i^{(i)} - \sum_{j=j+1}^{n} a_{ij}^{(i)} x_j)/a_{ii}^{(i)}$$

3.2 完全选主元消去法

$$|a_{i_k j_k}^{(k)}| = \max_{\substack{k \le i \le n \\ k \le j \le n}} |a_{ij}^{(k)}|$$

4 直接分解法

定理 4.1. *若 A 为非奇异阵，则 ∃ 排列阵 P(初等置换阵乘积)，使得 $PA = LU$ 其中 L 为单位下三角阵，U 为上三角阵 $A = LU$*

$$Ax = b \Leftrightarrow LUx = b \Leftrightarrow \begin{cases} Ly = b \\ Ux = y \end{cases}$$

4.1 Doolittle 分解法

假设 A 非异阵，且 $A = LU$，i.e.

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & & & \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \\ \ldots & \ldots & \ddots & \\ l_{n1} & l_{n2} & \ldots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \ldots & u_{1n} \\ & u_{22} & \ldots & u_{2n} \\ & & \ddots & \vdots \\ & & & u_{nn} \end{bmatrix}$$

对于 L 和 U，由于 $i < j$ 时，$l_{ij} = 0$，当 $i > j, u_{ij} = 0, \therefore$ 由 $A = LU$ 可得:

$$a_{kj} = \sum_{r}^{n} = 1 l_{kr} u_{rj} = \sum_{r=1}^{\min(k,j)} l_{kr} u_{kj}, k, j = 1, 2, \ldots, n \tag{4.1}$$

显然:

$$u_{1j} = a_{1j}, j = 1, 2, \ldots, n \tag{4.2}$$

当 $j = 1$ 时，由4.1可得，

$$a_{k1} = l_{k1} u_{11} \Rightarrow l_{k1} = a_{k1}/u_{11}, k = 2, 3, \ldots n \tag{4.3}$$

假设 U 的第 1 行至第 $k-1$ 行和 L 的第 1 列至第 $k-1$ 列都已求出，则可计算 U 的第 k 行元素

$$\begin{aligned} &\because a_{kj} = l_{kk}u_{kj} + \textstyle\sum_{r=1}^{k-1} l_{kr}u_{rj} \\ &\therefore u_{kj} = a_{kj} - \textstyle\sum_{r=1}^{k-1} l_{kr}u_{rj}, j = k, k+1, \ldots, n \end{aligned} \tag{4.4}$$

$$\begin{aligned} &\because a_{ik} = l_{ik}u_{kk} + \textstyle\sum_{r=1}^{k-1} l_{ir}u_{rk} \\ &\therefore l_{ik} = (a_{ik} - \textstyle\sum_{r=1}^{k-1} l_{ik}u_{rk})/u_{kk}, i = k+1, \ldots, n \end{aligned} \tag{4.5}$$

利用（4.2）（4.3）（4.4）（4.5）将 A 分解为 LU，这种分解方法称为 Doolittle 三角分解。

$$Ax = b \Leftrightarrow LUx = b \Leftrightarrow \begin{cases} Ly = b \\ Ux = y \end{cases} \Rightarrow y_i = b_i - \sum_{r=1}^{i-1} l_{ir}y_r, i = 1, 2, \ldots, n \tag{4.6}$$

$i = 1$ 时，$\sum_{r=1}^{0} l_{ir}y_r = 0$

$$x_i = (y_i - \sum_{r=i+1}^{n} u_{ir}x_r)/u_{ii}, i = n, n-1, \ldots, 1 \tag{4.7}$$

首先用（4.2）（4.3）（4.4）（4.5）将 A 分解为 LU，再用（4.6）(4.7) 求出方程组的解，称为 Doolitle 三角分析方法。

4.2 三对角方程组的追赶法

设有方程 $Ax = d$，且 $D_i \neq 0, i = 1, 2, \ldots, n$ 其中：

$$A = \begin{bmatrix} b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & c_{n-1} \\ & & & a_n & b_n \end{bmatrix}$$

设 $A = LU$，易证：

$$L = \begin{bmatrix} 1 & & & & \\ l_2 & 1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & l_n & 1 \end{bmatrix}, U = \begin{bmatrix} u_1 & c_1 & & & \\ & u_2 & c_2 & & \\ & & \ddots & \ddots & \\ & & & \ddots & c_{n-1} \\ & & & & u_n \end{bmatrix}$$

$\therefore$ 由 $A=LU$ 即可确定:

$$\begin{cases} u_1 = b_1 \\ l_i = a_i/u_{i-1}, i=2,3,\ldots,n \\ u_i = b_i - l_i c_{i-1}, i=2,3,\ldots,n \end{cases} \tag{4.8}$$

$$LUx=b \Leftrightarrow \begin{cases} Ly=b \\ Ux=y \end{cases} \Rightarrow \begin{cases} y_1 = d_1 \\ y_i = d_i - l_i y_{i-1}, i=2,3,\ldots,n \end{cases} \tag{4.9}$$

$$\begin{cases} x_n = y_n/u_n \\ x_i = (y_i - c_i x_{i+1})/u_i, i=n-1,n-2,\ldots,1 \end{cases} \tag{4.10}$$

定理 4.2. 设三对角阵 A 满足对角占优条件，$|b_1|>|c_1|>0, |b_n|>|a_n|>0, |b_i| \geq |a_i|+|c_i|$ 则 A 非奇异，且 $y_i \neq 0, i=1,2,\ldots,n$ $0<|\frac{c_i}{u_i}|<1, i=1,2,\ldots,n-1, |b_i|-|a_i|<|u_i|<|b_i|+|a_i|, i=2,3,\ldots,n$

4.3 Cholesky 分解（平方根）法

对称：$A^T=A$；**正定**：$\lambda_i>0$，$D_i>0$

假设 $det(A) \neq 0$，且 $D_i \neq 0, i=1,2,\ldots,n-1$ 则 $A=LU$ 进一步：

$$U = \begin{bmatrix} u_{11} & & & \\ & u_{22} & & \\ & & \ddots & \\ & & & u_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12}/u_{11} & \ldots & u_{1n}/u_{11} \\ & 1 & \ldots & u_{2n}/u_{22} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} = D\widetilde{U}$$

i.e. $A=LD\widetilde{U}$ 这就是 A 的 LDU 分解。

定理 4.3. （对称阵的三角分解定理）设 $A \in R^{n\times n}$，且 A 的顺序主子式 $D_i \neq 0, i=1,2,\ldots,n$ 则 $\exists$ 唯一单位下三角阵 L 和对角阵 D 使得 $A=LDL^T$

证明. $\because A=A^T=LD\widetilde{U}=\widetilde{U}^T DL^T \overset{\text{分解唯一性}}{\Rightarrow} L=\widetilde{U}^T \Rightarrow \widetilde{U}=L^T$

$\therefore A=LD\widetilde{U}=LDL^T$ □

定理 4.4. 设 $A \in R^{n\times n}$，且 A 对称正定，则 $\exists$ 唯一的对角元素为正的下三角阵 L，使得 $A=LL^T$，由

$$A=LL^T, L=\begin{bmatrix} l_{11} & & & & \\ l_{21} & l_{22} & & & \\ \vdots & \vdots & l_{33} & & \\ \vdots & \vdots & & \ddots & \\ l_{n1} & l_{n2} & \ldots & \ldots & l_{nn} \end{bmatrix}$$

$$a_{ij} = (\sum_{k=1}^{j} l_{ik}l_{jk}) = \sum_{k=1}^{j-1} l_{ik}l_{jk} + l_{ij}l_{jj}, i = j, j+1, \dots, n$$

$\Rightarrow$

$$l_{jj} = \sqrt{a_{jj} - \sum_{k=1}^{j-1} l_{jk}^2}, j = 1, 2, \dots, n \tag{4.11}$$

$$l_{ij} = (a_{ij} - \sum_{k=1}^{j-1} l_{ik}l_{jk})/l_{jj}, i = j+1, \dots, n, j = 1, 2, \dots, n-1 \tag{4.12}$$

方程组的解：

$$y_i = (b_i - \sum_{k=1}^{i-1} l_{ik}y_k)/l_{ii}, i = 1, 2, \dots, n \tag{4.13}$$

$$x_i = (y_i - \sum_{k=i+1}^{n} l_{ki}x_k)/l_{ii}, i = n, n-1, \dots, 1 \tag{4.14}$$

证明. A 对称正定 $\therefore D_i = u_{11}u_{22}\dots u_{ii} > 0 \therefore u_{ii} > 0, i = 1, 2, \dots, n$
$\therefore D = D^{\frac{1}{2}}D^{\frac{1}{2}}$，其中 $D^{\frac{1}{2}} = diag(\sqrt{u_{11}}, \sqrt{u_{22}}, \dots, \sqrt{u_{nn}})$
$\therefore A = LD^{\frac{1}{2}}D^{\frac{1}{2}}L^T = (LD^{\frac{1}{2}})(LD^{\frac{1}{2}})^T = \widetilde{L}\widetilde{L}^T$
由 $A = LDL^T$ 的分解唯一性 $\Rightarrow A = \widetilde{L}\widetilde{L}^T$ 也是唯一的。 □

5 条件数和摄动理论初步

例 5.1.

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.00001x_2 = 2 \end{cases} \Rightarrow x_1 = 2, x_2 = 0$$

若上式变为：

$$\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.00001x_2 = 2.00001 \end{cases} \Rightarrow x_1 = 1 = x_2$$

这就是病态方程组。

定义 5.1. 设 $Ax = b$，当 A 和 b 有微小变化 ΔA 和 Δb 就引起解向量 x 的很大变化，称 A 为关于解方程组和矩阵求逆的病态矩阵。称相应的方程组为病态方程组，反之，若 ΔA 和 Δb 很小，Δx 也很小，就称 A 为良态矩阵和称 $Ax = b$ 为良态方程组。

5.1 右端项 b 的摄动对解的影响

假设 $det(A) \neq 0$，b 有扰动 Δb，则 $Ax = b$ 的解也要产生摄动 Δx，i.e. 方程组 $Ax = b$ 变成了 $A(x + \Delta x) = b + \Delta b \Rightarrow \Delta x = A^{-1}\Delta b \Rightarrow ||\Delta x|| \leq ||A^{-1}|| \cdot ||\Delta b||$

$\because ||x|| \geq \frac{||Ax||}{||A||} = \frac{||b||}{||A||} \Rightarrow \frac{1}{||x||} \leq \frac{||A||}{||b||}$

$\therefore \frac{||\Delta x||}{||x||} \leq ||A|| \cdot ||A^{-1}|| \frac{||\Delta b||}{||b||}$

5.2 A 的摄动对解的影响

设 $det(A) \neq 0$, A 有摄动 ΔA, 相应地由 $x \to x + \Delta x$, i.e. $Ax = b, (A + \Delta A)(x + \Delta x) = b \Rightarrow (A + \Delta A)\Delta x = -\Delta A x \Rightarrow A(I + A^{-1}\Delta A)\Delta x = -\Delta A x$

当 $||A^{-1}|| \cdot ||\Delta A|| < 1$ 时,由矩阵范数定理1.8($||(I \pm B)^{-1}|| \leq \frac{1}{1-||B||}(||B|| < 1)$)，可知，$(I + A^{-1}\Delta A)^{-1}$ 存在

$$\therefore \Delta x = -(I + A^{-1}\Delta A)^{-1}A^{-1}\Delta A x$$

$$\therefore ||\Delta x|| \leq ||(I + A^{-1}\Delta A)^{-1}||||A^{-1}||||\Delta A||||x|| \leq \frac{||A^{-1}||||\Delta A||}{1 - ||A^{-1}\Delta A||}||x||$$

$$\Rightarrow \frac{||\Delta x||}{||x||} \leq \frac{||A||||A^{-1}|| \cdot \frac{||\Delta A||}{||A||}}{1 - ||A||||A^{-1}||\frac{||\Delta A||}{||A||}} = \frac{cond(A)\frac{||\Delta A||}{||A||}}{1 - cond(A)\frac{||\Delta A||}{||A||}}$$

定义 5.2. *对于非异阵 A，$||\cdot||$ 为一种矩阵的从属范数，称数 $||A|| \cdot ||A^{-1}||$ 为 A 的条件数，记为 $cond(A) = ||A|| \cdot ||A^{-1}||$*

常用条件数:

$$cond(A)_\infty = ||A||_\infty \cdot ||A^{-1}||_\infty$$

$$cond(A)_2 = ||A||_2||A^{-1}||_2 = \sqrt{\frac{\lambda_{\max}(A^TA)}{\lambda_{\min}(A^TA)}}$$

特别地当：$A^T = A$ 时，$cond(A)_2 = \frac{|\lambda_1|}{|\lambda_n|}$，其中 λ_1 和 λ_n 分别为 A 的绝对值最大和最小的特征值。

条件数性质：（运用线性代数知识证明）

1. $cond(A) \geq 1, cond(A) = cond(A^{-1})$
2. $cond(cA) = cond(A), \forall c \in R, c \neq 0$
3. 若矩阵 A 正交阵，则 $cond(A)_2 = 1$
4. 若 U 为正交阵，则 $cond(A)_2 = cond(AU)_2 = cond(UA)_2$

5. $cond(A) \geq \frac{|\lambda_1|}{|\lambda_n|}$，其中 λ_1, λ_n 分别为绝对值最大，最小的特征值。(证明：定理1.6)

定理 5.1. 设 $Ax = b, det(A) \neq 0$，b 为非零向量且 Δb 和 ΔA 分别为 b 和 A 的扰动量，若 $||A^{-1}|| \cdot ||\Delta A|| < 1$ 则有事前误差估计式：

$$\frac{||\Delta x||}{||x||} \leq \frac{cond(A)}{1 - cond(A)\frac{||\Delta A||}{||A||}}[\frac{||\Delta b||}{||b||} + \frac{||\Delta A||}{||A||}] \tag{5.1}$$

例 5.2. 已知方程组 $\begin{cases} x_1 + x_2 = 2 \\ x_1 + 1.000001x_2 = 2 \end{cases}$，右端项 b 有扰动 $\Delta b = (0, 10^{-5})^T$，求其系数矩阵 A 的条件数 $cond(A)_\infty$，说明 Δb 对解的影响，并分析其性态。

解：$\because A = \begin{bmatrix} 1 & 1 \\ 1 & 1 + 10^{-6} \end{bmatrix}, A^{-1} = \begin{bmatrix} 1 + 10^6 & -10^6 \\ -10^6 & 10^6 \end{bmatrix}$

$\therefore cond(A)_\infty = ||A||_\infty ||A^{-1}||_\infty = (2 + 10^{-6})(1 + 2 \times 10^6) > 4 \times 10^6$ 由于条件数很大，方程组病态，A 为病态矩阵。

右端项对解的影响：$\frac{||\Delta x||}{||x||} \leq cond(A)\frac{||\Delta b||}{||b||} = 2000\%$

定理 5.2. 设 $det(A) \neq 0$，x 和 $\bar{x}$ 分别为 $Ax = b$ 的准确解和近似解，$r = b - A\bar{x}$ 为残差。则：

$$\frac{1}{cond(A)}\frac{||r||}{||b||} \leq \frac{||x - \bar{x}||}{||x||} \leq cond(A)\frac{||r||}{||b||} \tag{5.2}$$

称为事后误差估计式。

证明. $\because x = A^{-1}b, \bar{x} = A^{-1}(b - r) \therefore x - \bar{x} = A^{-1}r$

$$\begin{aligned}
\frac{1}{cond(A)}\frac{||r||}{||b||} &= \frac{||AA^{-1}r||}{||A|| \cdot ||A^{-1}|| \cdot ||b||} \\
&\leq \frac{||A||||A^{-1}r||}{||A|| \cdot ||A^{-1}|| \cdot ||b||} = \frac{||A^{-1}r||}{||A^{-1}|| \cdot ||b||} \\
&\leq \frac{||x - \bar{x}||}{||x||} \\
&\leq \frac{||A^{-1}||||r||||A||}{||A^{-1}b||||A||} \leq cond(A)\frac{||r||}{||AA^{-1}b||} \\
&= cond(A)\frac{||r||}{||b||}
\end{aligned}$$

□

6 迭代法的基本概念

6.1 迭代法的一般形式

对于方程：

$$Ax = b, det(A) \neq 0, b \in R^n \tag{6.1}$$

将6.1转化为等价的方程组：

$$x = Bx + f, B \in R^{n\times n}, f \in R^n \tag{6.2}$$

由6.2可构造如下迭代公式：

$$x^{(k+1)} = Bx^{(k)} + f, k = 0, 1, 2, \dots \tag{6.3}$$

其中：B 称为迭代矩阵，f 为右端项，给定 $x^{(0)}$ 则可利用6.3得向量序列：$\{x^{(k)}\}, x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$，若 $x^{(k)}$ 收敛，i.e.$x^{(k)} \to x^*, k \to \infty$，则有：

$$x^* = Bx^* + f$$

6.2 向量序列和矩阵序列的收敛性

定义 6.1. 设 $\{x^{(k)}\}$ 为 R^n 中的向量序列，$x^* \in R^n$，若 $\lim_{k\to\infty} ||x^{(k)} - x^*|| = 0$，其中 $||\cdot||$ 为向量范数，则称 $\{x^{(k)}\}$ 收敛于 x^*，记为 $\lim_{k\to\infty} x^{(k)} = x^*$。

定理 6.1. R^n 中的 $\{x^{(k)}\}$ 收敛于 $x^* \in R^n$，当且仅当 $\lim_{k\to\infty} x_i^{(k)} = x_i^*, i = 1, 2, \dots, n$。其中 $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^T, x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$

证明. 若 $\{x^{(k)}\}$ 收敛，则 $\lim_{k\to\infty} ||x^{(k)} - x^*|| = 0$，又因为：$0 \leq |x_i^{(k)} - x_k^*| \leq \max_{n\geq i\geq 1} |x_i^{(k)} - x_k^*| = ||x^{(k)} - x^*||_\infty \to 0$
$\therefore \lim_{k\to\infty} x^{(k)} = x^*$

反之，若 $\lim_{k\to\infty} x_i^{(k)} = x_i^*, i = 1, 2, \dots, n$，则可证：$\lim_{k\to\infty} ||x^{(k)} - x^*||_\infty = 0$ 又因为向量范数的等价性1.5，有 $||x^{(k)} - x^*|| \leq C||x^{(k)} - x^*||_\infty \to 0$ □

定义 6.2. 定义了范数 $||\cdot||$ 的空间 $R^{n\times n}$ 中，若 $\exists A \in R^{n\times n}$，使 $\lim_{k\to\infty} ||A^{(k)-A}|| = 0$，则称 $\{A^{(k)}\}$ 收敛于 A，记为：$\lim_{k\to\infty} A^{(k)} = A$

定理 6.2. 【自证】设 $A^{(k)} = (a_{ij}^{(k)})_n, (k = 1, 2, \dots, \infty), A = (a_{ij})_n$ 为 $R^{n\times n}$ 中的矩阵，则 $\{A^{(k)}\}$ 收敛于 A 的充要条件为：

$$\lim_{k\to\infty} a_{ij}^{(k)} = a_{ij}$$

定理 6.3.

$$\lim_{k\to\infty} A^{(k)} = 0 \Leftrightarrow \lim_{k\to\infty} A^{(k)}x = 0, \forall x \in R^n$$

证明.
- 必要性：对任一从属范数：$||A^{(k)}x|| \le ||A^{(k)}||||x|| \to 0$
- 充分性：取 $x = e_j$，则由 $\lim_{k\to\infty} A^{(k)}e_j = 0$，可知 $A^{(k)}$ 的第 j 列为 0，当 $j = 1, 2, \ldots, n$ 时，即可证明 $\lim_{k\to\infty} A^{(k)} = 0$

□

定理 6.4. 设 $B \in R^{n\times n}$ 则有下列三个命题等价：

- $\lim_{k\to\infty} B^k = 0$
- $\rho(B) < 1$
- 至少存在一种从属矩阵范数 $||\cdot||$，使得 $||B|| < 1$

证明. 逐个证明：

(1)->(2) 反证法：假设 B 有一个特征值 λ 满足 $|\lambda| \ge 1$，则有 $\lambda \ne 0$ 使得 $Bx = \lambda x \Rightarrow B^k x = \lambda B^{k-1}x = \lambda^k x, \therefore k- > \infty, B^k x \ne 0$ 根据 $(\lim_{k\to\infty} A^{(k)} = 0 \Leftrightarrow \lim_{k\to\infty} A^{(k)}x = 0, \forall x \in R^n)$ 得，$\lim_{k\to\infty} B^k \ne 0$ 与已知矛盾，$\therefore \rho(B) < 1$

(2)->(3) 由定理1.6知，对 $\forall\varepsilon > 0, \exists||\cdot|| s.t. ||B|| \le \rho(B) + \varepsilon \Rightarrow ||B|| < \rho(B) + 2\varepsilon, \because \rho(B) < 1, \text{choose } \varepsilon = \frac{1-\rho(B)}{2} > 0 \therefore ||B|| < \rho(B) + 2\varepsilon = \rho(B) + 2\frac{1-\rho(B)}{2} = 1$

(3)->(1) 由相容性条件可得:$||B^k - 0|| \le ||B||^k \because ||B|| < 1 \Rightarrow \lim_{k\to\infty} ||B||^k = 0 \Rightarrow \lim_{k\to\infty} B^k = 0$ □

6.3 迭代方法的收敛性

$x^{(k+1)} = Bx^{(k)} + f$，若收敛，则有 $x^* = Bx^* + f$，则 $e^{(k)} = x^{(k)} - x^* = Bx^{(k-1)} - Bx^* = \cdots = B^k e^{(0)}, e^{(0)} = x^{(0)} - x^*$

定理 6.5. (重要) 对于任意初值 $x^{(0)}$ 和右端项 f, 迭代方法 $x^{(k+1)} = Bx^{(k)} + f$ 收敛的**充要条件**是：$\rho(B) < 1$

证明. 由以上分析有：$e^{(k)} = x^{(k)} - x^* = B^k e^{(0)}$，其中 $e^{(0)}$ 为与 k 无关的任意初始误差

$\therefore$ 迭代法的收敛性等价于 $\{e^{(k)}\}$ 的收敛性。

而由 $e^{(0)}$ 的任意性，及定理6.3知：$\{e^{(k)}\}$ 的收敛性等价于 $\{B^k\}$ 的收敛性，

而 $\{B^k\}$ 的收敛性与 $\rho(B) < 1$ 等价
$\therefore$ 迭代法收敛 $\Leftrightarrow \rho(B) < 1$

$$x^{(k)} \to x^* \Leftrightarrow B^k e^{(0)} \to 0 \Leftrightarrow B^k \to 0 \Leftrightarrow \rho(B) < 1$$

□

定理 6.6. ***(充分条件)*** *对于 $x^{(k+1)} = Bx^{(k)} + f$，若 $\exists || \cdot ||_\lambda, s.t. \ ||B||_\lambda < 1$ 则:*

1. 迭代法收敛；i.e. $x^{(k)} \to x^*, k \to \infty$

2.
$$||x^{(k)} - x^*||_\lambda \leq \frac{||B||_\lambda}{1 - ||B||_\lambda} ||x^{(k)} - x^{(k-1)}||_\lambda$$

3.
$$||x^{(k)} - x^*||_\lambda \leq \frac{||B||_\lambda^k}{1 - ||B||_\lambda} ||x^{(0)} - x^*||_\lambda$$

证明. 1. $\rho(B) < ||B|| < 1$

2. $\because x^{(k)} - x^* = B(x^{(k-1)} - x^*) = B(x^{(k)} - x^*) + B(x^{(k-1)} - x^{(k)}) \Rightarrow (I - B)(x^{(k)} - x^*) = B(x^{(k-1)} - x^{(k)})$
$\because ||B|| < 1, \therefore \exists (I - B)^{-1} \therefore x^{(k)} - x^* = -(I - B)^{-1} B(x^{(k)} - x^{(k-1)})$
$\therefore ||x^{(k)} - x^*||_\lambda \leq ||(I - B)^{-1}||_\lambda ||B||_\lambda ||x^{(k)} - x^{(k-1)}||_\lambda \leq \frac{||B||_\lambda}{1 - ||B||_\lambda} ||x^{(k)} - x^{(k-1)}||_\lambda$

3. $\because ||x^{(k)} - x^{(k-1)}||_\lambda = ||B(x^{(k-1)} - x^{(k-2)})||_\lambda \leq ||B||_\lambda ||x^{(k-1)} - x^{(k-2)}||_\lambda \leq \cdots \leq ||B||_\lambda^{k-1} ||x^{(1)} - x^{(0)}||_\lambda$
将上式代入 2 中，即可得到 3.

□

注释:

- 若仅仅是为了判断迭代算法的收敛性，则定理中的条件还可放宽为: $\exists$ 某一种范数使得 $||B||_\lambda < 1$
- 方法的收敛性与右端项 f 无关
- 从定理可看出，$||B||_\lambda$ 不是很靠近 1，如果要求 $||x^{(k)} - x^*||_\lambda < \varepsilon$，只需要使相邻两次的 $||x^{(k)} - x^{(k-1)}||_\lambda < \varepsilon$ 即可
- 判断一种方法的迭代次数，用定理6.3.3 可以解析求出。

定义 6.3. **（重要概念）** 称 $R(B) = -ln(\rho(B))$ 为迭代法 $x^{(k+1)} = Bx^{(k)} + f$ 的渐进收敛率或渐进收敛速度。

7 Jacobi 方法和 GS 迭代法

7.1 J 法

假设 $det(A) \neq 0, A = D - L - U$，其中：$D = diag(a_{11}, a_{22}, \dots, a_{nn})$

$$L = \begin{bmatrix} 0 & & & \\ -a_{21} & 0 & & \\ \vdots & \vdots & \ddots & \\ -a_{n1} & -a_{n2} & \dots & 0 \end{bmatrix}, U = \begin{bmatrix} 0 & -a_{12} & \dots & -a_{1n} \\ & \ddots & \dots & -a_{2n} \\ & & \ddots & \vdots \\ & & & 0 \end{bmatrix}$$

将 $A = D - L - U$ 代入 $Ax = b$ 则 $Dx = (L+U)x + b \Rightarrow x = D^{-1}[(L+U)x + b] = D^{-1}(L+U)x + D^{-1}b \Rightarrow B = D^{-1}(L+U), f = D^{-1}b$

由此可得 J 法：

$$\begin{aligned} & x^{(k+1)} = Bx + f \\ & B = D^{-1}(L+U) = I - D^{-1}A \\ & f = D^{-1}b \end{aligned} \tag{7.1}$$

其分量形式：

$$x_i^{(k+1)} = \frac{1}{a_{ii}}[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}], j = 1, 2, \dots, n \tag{7.2}$$

7.2 GS 方法

分量形式：

$$x_i^{(k+1)} = \frac{1}{a_{ii}}[b_i - \sum_{j=1}^{i-1} a_{ij}\mathbf{x_j^{(k+1)}} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}], j = 1, 2, \dots, n \tag{7.3}$$

写成向量形式有：

$$\begin{aligned} & x^{(k+1)} = D^{-1}(b + Lx^{(k+1)} + Ux^{(k)}) \\ & \Leftrightarrow (D-L)x^{(k+1)} = b + Ux^{(k)} \\ & \Leftrightarrow x^{(k+1)} = (D-L)^{-1}Ux^{(k)} + (D-L)^{-1}b \\ & = (I - (D-L)^{-1}A)x^{(k)} + (D-L)^{-1}b \end{aligned} \tag{7.4}$$

例 7.1.

$$\begin{cases} 10x_1 - x_2 - 2x_3 = 72 \\ -x_1 + 10x_2 - 2x_3 = 83 \\ -x_1 - x_2 + 5x_3 = 42 \end{cases}$$

易知解 $x^* = (11, 12, 13)$

解：J 法：

$$
\begin{aligned}
x_1^{(k+1)} &= \tfrac{1}{10}(72 + x_2^{(k)} + 2x_3^k) \\
x_2^{(k+1)} &= \tfrac{1}{10}(83 + x_1^{(k)} + 2x_3^{(k)}) \\
x_3^{(k+1)} &= \tfrac{1}{5}(42 + x_1^{(k)} + x_2^{(k)})
\end{aligned}
$$

取 $x^{(0)} = (0,0,0)^T$ 则：$x^{(9)} = (10.9994, 11.9994, 12.9992)^T$，误差 $||x^{(9)} - x^*||_\infty = 0.0008$

GS 法：

$$
\begin{aligned}
x_1^{(k+1)} &= \tfrac{1}{10}(72 + x_2^{(k)} + 2x_3^k) \\
x_2^{(k+1)} &= \tfrac{1}{10}(83 + x_1^{(k+1)} + 2x_3^{(k)}) \\
x_3^{(k+1)} &= \tfrac{1}{5}(42 + x_1^{(k+1)} + x_2^{(k+1)})
\end{aligned}
$$

取 $x^{(0)} = (0,0,0)^T$ 则：$||x^{(6)} - x^*||_\infty = 0.0001$，显然 GS 法收敛好于 J 法。

7.3 J 法 GS 法的收敛性

收敛的充要条件：$\rho(B) < 1$，充分条件：$||B|| < 1$

定义 7.1.

$$|a_{ii}| > \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|, i = 1, 2, \ldots, n$$

称 A 为严格对角占优阵。

$$|a_{ii}| \geq \sum_{\substack{j=1\\ j\neq i}}^{n} |a_{ij}|, i = 1, 2, \ldots, n$$

且其中至少有一个不等式严格成立，则称 A 为弱对角占优阵。

定义 7.2. 设 $A \in R^{n\times n}$，若不能找到排列阵 P 使得 $P^TAP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$（其中 A_{11}, A_{22} 均为方阵）成立，则称 A 为不可约的。

例 7.2.

$$B = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 2 \end{bmatrix} \rightarrow \text{1-3 行交换} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \rightarrow \text{1,3 列交换} \begin{bmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

定理 7.1. 若 $A = (a_{ij})_n \in R^{n\times n}$ 为严格对角占优阵或不可约弱对角占优阵，则 $a_{ii} \neq 0, i = 1, 2, \ldots, n$，且 A 为非奇异阵。

定理 7.2. 若 A 为严格对角占优阵或不可约弱对角占优阵，则 $\forall x^{(0)}$，方程 $Ax = b$ 的 J 法和 GS 法均收敛。

反证法. (目标 $\rho(G)<1$)

设 G 有一个特征值 λ 满足 $|\lambda|\geq 1$，则 $|\lambda I-(D-L)^{-1}U|=0\Rightarrow |I-\lambda^{-1}(D-L)^{-1}U|=0\Rightarrow |(D-L)^{-1}||D-L-\lambda^{-1}U|=0$

$\because a_{ii}\neq 0\therefore |(D-L)^{-1}|\neq 0$，而 $A=D-L-U$ 与 $D-L-\lambda^{-1}U$ 的零元素与非零元素位置完全一样，所以 $D-L-\lambda^{-1}U$ 也是不可约的。

又 $\because |\lambda|\geq 1, D-L-\lambda^{-1}U$ 也是弱对角占优矩阵。根据定理7.1有 $|D-L-\lambda^{-1}U|\neq 0$，矛盾，证明 $\rho(B)<1$。 □

定理 7.3. 设 A 对称，且 $a_{ii}>0, i=1,2,\ldots,n$ 则 J 法收敛 $\Leftrightarrow A$ 及 $2D-A$ 均正定。

证明. $\because a_{ii}>0, i=1,2,\ldots,n$
$\therefore D=D^{\frac{1}{2}}D^{\frac{1}{2}}$
而 $B=I-D^{-1}A=D^{-\frac{1}{2}}(I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}$，说明 B 与 $I-D^{-\frac{1}{2}}AD^{\frac{1}{2}}$ 相似。

- 必要性:
  若 J 法收敛，则 $\rho(B)<1$。设 $D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值 μ（实数）则: $\lambda(B)=1-\mu, \therefore |1-\mu|<1\Rightarrow 0<\mu<2$
  $\because D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$正定 $\therefore \forall x\in R^n, (D^{-\frac{1}{2}}x)^T AD^{-\frac{1}{2}}x=x^T D^{-\frac{1}{2}}AD^{-\frac{1}{2}}x>$ 0A 正定。
  又 $\because \lambda(2I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})=2-\mu\in(0,2)$
  $\therefore 2D-A$ 正定。

- 充分性:
  $\because A$正定
  $\therefore D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$正定
  $\therefore \lambda(D^{-\frac{1}{2}}AD^{-\frac{1}{2}})>0$
  $\therefore \lambda(B)<1\ (B=I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})$
  又 $\because -B=-D^{-\frac{1}{2}}(I-D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{-\frac{1}{2}}=D^{-\frac{1}{2}}(I-D^{-\frac{1}{2}}(2D-A)D^{-\frac{1}{2}})D^{-\frac{1}{2}}\ \therefore -\lambda(B)<1\Rightarrow \lambda(B)>-1$

□

定理 7.4. 设 A 对称正定，则方程 $Ax=b$ 的 GS 法收敛。

例 7.3.

- $$\begin{cases} x_1+2x_2-2x_3=1 \\ x_1+x_2+x_3=1 \\ 2x_1+2x_2+x_3=1 \end{cases}$$

  J 法收敛，GS 法发散

- $$\begin{cases} x_1 + 0.8x_2 + 0.8x_3 = 2.6 \\ 0.8x_1 + x_2 + 0.8x_3 = 2.6 \\ 0.8x_1 + 0.8x_2 + x_3 = 2.6 \end{cases}$$

  GS 法收敛，*J* 法发散

例 7.4. 当 $a_{ii} > 0$，且 $a_{ij} \le 0, i \ne j$ 时，可证下列四种情况只有一种成立。

1. $\rho(G) = \rho(B) = 0$
2. $0 < \rho(G) < \rho(B) < 1$
3. $\rho(G) = \rho(B) = 1$
4. $1 < \rho(B) < \rho(G)$

通常情况下，*GS* 方法好于 *J* 法，但不是所有情况。

8 超松弛迭代法

8.1 SOR 法构造

记 $\Delta x = (\Delta x_1, \Delta x_2, \ldots, \Delta x_n)^T = x^{(k+1)} - x(k)$
则 GS 法可写成：$x^{(k+1)} = x^{(k)} + \Delta x$ 其中 $\Delta x_i = \frac{1}{a_{ii}}[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}] - x_i^{(k)}$
引入参数 w，即可得 SOR 法：

$$x^{(k+1)} = x^{(k)} + w\Delta x$$

i.e.

$$x_i^{(k+1)} = (1-w)x_i^{(k)} + \frac{w}{a_{ii}}[b_i - \sum_{j=1}^{i-1} a_{ij}x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij}x_j^{(k)}] \tag{8.1}$$

将 a_{ii} 乘到等号左边，写成向量形式：

$$\begin{aligned} &Dx^{(k+1)} = (1-w)Dx^{(k)} + w[b + Lx^{(k+1)} + Ux^{(k)}] \\ &\Rightarrow x^{(k+1)} = (D-wL)^{-1}[(1-w)D + wU]x^{(k)} + (D-wL)^{-1}wb \end{aligned}$$

令 $L_w = (D-wL)^{-1}[(1-w)D + wU], f = (D-wL)wb$，则有

$$x^{(k+1)} = L_w x^{(k)} + f \tag{8.2}$$

例 8.1.

$$\begin{bmatrix} 4 & 3 & 0 \\ 3 & 4 & -1 \\ 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 24 \\ 30 \\ -24 \end{bmatrix}$$

精确解：$x^* = (3, 4, -5)^T$

用 $w = 1(GS), x^{(7)} = (3.013411, 3.988824, -5.002794)^T$

$w = 1.2(SOR), x^{(7)} = (3.00049, 4.000258, -5.000348)^T$

定理 8.1. （SOR 法收敛的必要条件）

$\because SOR$ 法收敛,$\therefore \rho(L_w) < 1$

设 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 为 L_w 的特征值,则 $|L_w| = |\prod_{i=1}^n \lambda_i| \leq |\lambda_1||\lambda_2|\ldots|\lambda_n| \leq [\rho(L_w)]^n < 1 \Rightarrow |L_w|^{\frac{1}{n}} \leq \rho(L_w) < 1$ 而 $det(L_w) = det(D - wL)^{-1} det[(1 - w)D + wU] = detD^{-1} det[(1 - w)D] = (1 - w)^n, \therefore |1 - w| < 1 \Rightarrow 0 < w < 2$

定理 8.2. 若 A 为对称正定阵，则 SOR 法收敛的充要条件为 $0 < w < 2$

定义 8.1. 若存在排列阵 P 使 $PAP^T = \begin{bmatrix} D_1 & M_1 \\ M_2 & D_2 \end{bmatrix}$，其中 D_1, D_2 为对角阵，称 A 是 *2*-循环的。若矩阵 $\alpha D^{-1}L + \alpha^{-1}D^{-1}U$ 的特征值都与 α 无关，则 A 是相容次序矩阵。

定理 8.3. 设 $A \in R^{n\times n}$ 是 *2*-循环和相容次序的，$B = I - D^{-1}A$ 的特征值全为实数，且 $\mu = \rho(B) < 1$，则：

$$\rho(L_w) = \begin{cases} \frac{[w\mu + \sqrt{w^2\mu^2 - 4(w-1)}]^2}{4}, 0 < w < w_{opt} \\ w - 1, w_{opt} \leq w < 2 \end{cases}$$

其中，$\rho(w_{opt}) = \min \rho(L_w) = w_{opt} - 1, w_{opt} = \frac{2}{1+\sqrt{1-\mu^2}}$，称为最佳松弛因子，且当 $0 < 2 < w_{opt}$ 时，$\rho(L_w)$ 是 w 的单减函数，当 $w_{opt} \leq w \leq 2$ 时，$\rho(L_w)$ 是 w 的单增函数。

1. 当 w=1 时，$\rho(L_1) = \mu^2 = \rho(B)^2$，$R(L_w) = -ln\mu^2 = 2R(B)$，GS 法收敛速度为 J 法的 2 倍。
2. 显然 $\rho(L_w) \geq \rho(L_{w_{opt}}) = w_{opt} - 1, w_{opt} \geq 1$
3.

   定理 8.4. 设 A 是对称正定的三对角阵，则：$\rho(G) = \rho(B)^2 < 1$，且 $w_{opt} = \frac{2}{1+\sqrt{1-\rho(B)^2}}$

例 8.2.

$$Ax=b, A=\begin{bmatrix}4 & 3 & 0\\ 3 & 4 & -1\\ 0 & -1 & 4\end{bmatrix}$$

显然 A 为对称正定三对角阵（利用顺序主子式均大于 0 可以判断正定）

$\because B=I-D^{-1}A=\begin{bmatrix} & -\frac{3}{4} & \\ -\frac{3}{4} & & \frac{1}{4}\\ & \frac{1}{4} & \end{bmatrix}, \therefore \rho(B)=\sqrt{5/8}\approx 0.790<1, \rho(G)=\rho(B)^2=0.625$

$$w_{opt}=\frac{2}{1+\sqrt{1-0.625}}\approx 1.24$$

8.2 块松弛迭代法

设 $A=\begin{bmatrix}A_{11} & A_{12} & \dots & A_{1N}\\ A_{21} & A_{22} & \dots & A_{2N}\\ \dots & & & \\ A_{N1} & A_{N2} & \dots & A_{NN}\end{bmatrix}$ 其中 A_{ii} 为 $n_i\times n_i$ 的非奇异方阵，且 $n_1+n_2+\cdots+n_N=n$，有：

$$A_{ii}x_i^{(k+1)}=(1-w)A_{ii}x_i^{(k)}+w[b_i-\sum_{j=1}^{i-1}A_{ij}x_j^{(k+1)}-\sum_{j=i+1}^{n}A_{ij}x_j^{(k)}] \tag{8.3}$$

其中 $x_i^{(k)}, b_i$ 均为 n_i 个分量的向量。

由（8.3）可得：

$$x^{(k+1)}=(D-wL)^{-1}[(1-w)D+wU]x^{(k)}+w(D-wL)^{-1}b$$

其中 $X=(x_1^T,x_2^T,\dots,x_N^T)^T, b=(b_1^T,b_2^T,\dots,b_N^T)^T$

9 共轭梯度法

系数矩阵对称正定

若 $Ax=b$ 其中 $A>0$，则求解可转化为求下列二次函数：

$$\phi(x)=\frac{1}{2}x^TAx-x^Tb=\frac{1}{2}(Ax,x)-(b,x) \tag{9.1}$$

的最小值点问题。

定理 9.1. 设 A 对称正定，则 x^* 为方程组的解的充要条件是 $\phi(x^*)=\min_{x\in R^n}\phi(x)$

证明. 定义如下函数:

$$F(x)=\frac{1}{2}(A^{-1}r,r)\geq 0$$

其中 $r=b-Ax$，将 r 代入上式有:

$$F(x)=\phi(x)+\frac{1}{2}(A^{-1}b,b)$$

其中 $\phi(x)=\frac{1}{2}x^TAx-x^Tb=\frac{1}{2}(Ax,x)-(b,x)$ $F(x)$ 最小值点 $\Leftrightarrow\phi(x)$ 最小值点 $\Rightarrow r=0, i.e. Ax^*=b$ □

9.1 最速下降法

选取初值 $x^{(0)}$，则有:$-\nabla\phi(x)=-(\frac{\partial\phi}{\partial x_1},\frac{\partial\phi}{\partial x_2},\ldots,\frac{\partial\phi}{\partial x_n})^T|_{x=x^{(0)}}=b-Ax^{(0)}=r^{(0)}$。可令 $\frac{d\phi(x^{(0)}+\alpha r^{(0)})}{d\alpha}=0$，则可得 $\alpha=\frac{(r^{(0)},r^{(0)})}{(Ar^{(0)},r^{(0)})}=\alpha_0$，则 $x^{(1)}=x^{(0)}+\alpha_0r^{(0)}$ 是使得 $\phi(x)$ 下降最快，并且使之达到最小值点的极值点，然后再从 $x^{(1)}$ 出发，寻找一个使得 $\phi(x)$ 下降最快的方向 $r^{(1)}$ 和步长 α_1，同理可知：$r^{(1)}=b-Ax^{(1)},\alpha_1=\frac{(r^{(1)},r^{(1)})}{(Ar^{(1)},r^{(1)})}$，则可得：$x^{(2)}=x^{(1)}+\alpha_1r^{(1)}$，依次类推可得：$x^{(3)},x^{(4)},\ldots,x^{(k)}$，综上可得如下最速下降算法:

$$\begin{array}{l} r^{(k)}=b-Ax^{(k)} \\ \alpha_k=\frac{(r^{(k)},r^{(k)})}{(Ar^{(k)},r^{(k)})} \\ x^{(k+1)}=x^{(k)}+\alpha_kr^{(k)} \end{array} \tag{9.2}$$

易证：$\phi(x^{(k+1)})<\phi(x^{(k)}),|x^{(k+1)}-x^{(k)}|<\varepsilon$

9.2 共轭梯度法

对于 $A>0$ 称满足 $(AP^{(i)},P^{(j)})=0,i\neq j$ 的向量组 $\{P^{(i)}\}$ 为 A 共轭向量组，如果按方向 $P^{(0)},P^{(1)},\ldots,P^{(k-1)}$ 已进行了 k 次一维搜索，求得了 $x^{(k)}$，下一步求 $x^{(k+1)}$，则需要进行一维搜索使 $\phi(x^{(k)}+\alpha P^{(k)})$ 极小，则可令 $\frac{d\phi(x^{(k)}+\alpha p^{(k)})}{d\alpha}=0\Rightarrow\alpha_k=\frac{(r^{(k)},p^{(k)})}{(Ap^{(k)},p^{(k)})}$

由此可得: $x^{(k+1)}=x^{(k)}+\alpha_kp^{(k)},r^{(k+1)}=b-Ax^{(k+1)}=r^{(k)}-\alpha_kAP^{(k)}$，其中 $(AP^{(i)},P^{(j)})=0,i\neq j,p^{(0)}=r^{(0)}=b-Ax^{(0)}$

按此方法有如下性质:

$$\begin{array}{l} \phi(x^{(k+1)})=\min_\alpha\phi(x^{(k)}+\alpha p^{(k)}) \\ \phi(x^{(k+1)})=\min_{x\in\mathrm{span}\{p^{(0)},p^{(1)},\ldots,p^{(k)}\}}\phi(x) \end{array}$$

开始时选取 $p^{(0)}=r^{(0)}$，然后选取 $p^{(k)}=r^{(k)}+\beta_{k-1}p^{(k-1)}$，其中 β_{k-1} 由 A-共轭性确定，i.e.，由 $(Ap^{(k)},p^{(k-1)})=0$ 确定为：$\beta_{k-1}=-\frac{(r^{(k)},Ap^{(k-1)})}{(p^{(k-1)},Ap^{(k-1)})}$

综上可得如下 CG 算法：给定：$x^{(0)}, p^{(0)} = r^{(0)}, r^{(0)} = b - Ax^{(0)}$

$$\alpha_k = \frac{(r^{(k)}, r^{(k)})}{(Ap^{(k)}, p^{(k)})} \tag{9.3}$$

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)} \tag{9.4}$$

$$r^{(k+1)} = r^{(k)} - \alpha_k A p^{(k)} \tag{9.5}$$

$$\beta_k = \frac{(r^{(k+1)}, r^{(k+1)})}{(r^{(k)}, r^{(k)})} \tag{9.6}$$

$$p^{(k+1)} = r^{(k+1)} + \beta_k p^{(k)} \tag{9.7}$$