计算方法B

Little Wolf

2024年10月13日

线性方程组的直接解法 1

题目. 1. 求下三角矩阵的逆矩阵的详细算法

解答.下三角矩阵L可逆,因为行列式是对角元乘积,因此所有对角元均不为零 考虑LX = I,已知下三角矩阵的逆矩阵一定是下三角矩阵 逐列进行计算:

对(j,j), 有 $l_{jj}x_{jj}=1 \iff x_{jj}=\frac{1}{l_{jj}}$ 读(i,j), 是L的第i行的非零元是(1:i), X的第j列的非零元是(j:n), 不妨取i>j, 那么有:

$$\sum_{k=j}^{i} l_{ik} x_{kj} = 0 \iff x_{ij} = -\frac{\sum_{k=j}^{i-1} l_{ik} x_{kj}}{l_{ii}}$$

我们已知 x_{jj} , 取i=j+1, 得到 $x_{j+1,j}=-\frac{l_{ij}x_{jj}}{l_{ii}}$ 下面取i=j+2, 得到 $x_{j+2,j}=-\frac{\sum_{k=j}^{j+1}l_{ik}x_{kj}}{l_{ii}}$, 依此下去, 得到X矩阵在第j列的(j,n)的元素, 注意X矩 阵的X(1:j-1,j)都是0.

因此, 完整的算法如下:

Algorithm 1 下三角矩阵求逆

Input: 满秩的下三角矩阵 L

Output: 逆矩阵 L^{-1} 初始化 L^{-1} , 全零矩阵

for $j = 1 : n \ do$

$$x_{jj} = \frac{1}{l_{jj}}$$

 $x_{jj} = \frac{1}{l_{jj}}$ for i = j + 1 : n do $x_{i,j} = -\frac{\sum_{k=j}^{i-1} l_{ik} x_{kj}}{l_{ii}}$

end for

end for

Return L^{-1}

题目. 4. 确定一个 3×3 的高斯变换L, 使得

$$L \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 7 \\ 8 \end{bmatrix}$$

解答.第二行加上了第一行成二,第三行加上了第一行成二,因此

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = I - l_1 e_1^T, \quad l_1 = \begin{pmatrix} 0 \\ -2 \\ -2 \end{pmatrix}$$

题目. 5. 如果 $A \in \mathbb{R}^{n \times n}$ 有三角分解, 并且都是非奇异的, 那么A = LU分解得到的L和U是唯一的.

解答. 不妨假设分解是不唯一的,有非奇异单位下三角矩阵 L_1, L_2 ,非奇异上三角矩阵 U_1, U_2 使得 $L_1U_1 = L_2U_2 \iff L_2^{-1}L_1 = U_2U_1^{-1}$, $L_2^{-1}L_1$ 是单位下三角, $U_2U_1^{-1}$ 是上三角 因此 $L_2^{-1}L_1 = U_2U_1^{-1} = I \iff L_1 = L_2, U_1 = U_2$

题目. 8.设 $A = [a_{ij}] \in \mathbb{R}^{n \times n}$ 是严格对角占优矩阵, 即

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

又假设经过一步Gauss消去之后, A有如下形状:

$$\begin{bmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{bmatrix}$$

证明: 矩阵 A_2 仍然是严格对角占优矩阵. 由此推断, 对于严格对角占优矩阵来说, 用Gauss消去 法和列主元Gauss消去法可以得到同样的结果.

解答. 设矩阵

$$A = \begin{pmatrix} a_{11} & \alpha^T \\ \beta & A_{22} \end{pmatrix}$$

那么做一步Gauss消元就是左乘 $L = I - l_1 e_1^T$

$$(I - l_1 e_1^T) A = \begin{pmatrix} 1 & 0 \\ -l_1 & I_{n-1} \end{pmatrix} \begin{pmatrix} a_{11} & \alpha^T \\ \beta & A_{22} \end{pmatrix} = \begin{pmatrix} a_{11} & \alpha^T \\ \beta - a_{11} l_1 & A_{22} - l_1 \alpha^T \end{pmatrix}$$

通过 $\beta - a_{11}l_1 = 0$, 得到 $\beta = a_{11}l_1$, 因此

$$\begin{pmatrix} a_{11} & \alpha^T \\ \beta - a_{11}l_1 & A_{22} - l_1\alpha^T \end{pmatrix} = \begin{pmatrix} a_{11} & \alpha^T \\ 0 & A_{22} - \frac{1}{a_{11}}\beta\alpha^T \end{pmatrix} := A^{(1)}$$

因此 $A^{(1)}(i,j) = a'_{ij} = a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}, \ \forall 1 \leq i,j \leq n.$ 因此对于矩阵 A_2 ,考虑 $2 \leq k \leq n$,有

$$\sum_{j=2,j\neq k}^{n} |a'_{kj}| = \sum_{j=2,j\neq k}^{n} |a_{kj} - \frac{a_{k1}a_{1j}}{a_{11}}| \le \sum_{j=2,j\neq k}^{n} |a_{kj}| + |\frac{a_{k1}}{a_{11}}| \sum_{j=2,j\neq k}^{n} |a_{1j}|$$

$$< |a_{kk}| - a_{k1} + |\frac{a_{k1}}{a_{11}}| (|a_{11}| - |a_{1k}|)$$

$$= |a_{kk}| - |\frac{a_{k1}a_{1k}}{a_{11}}| \le |a_{kk} - \frac{a_{k1}a_{1k}}{a_{11}}| = |a'_{kk}|$$

因此 A_2 是严格对角占优的.

因此,在高斯消去法的第k-1步后,因为右下角的矩阵是严格对角占优的,所以 $A^{(k-1)}$ 第k列的k之后绝对值最大元素就是 $|a_{kk}^{(k-1)}|$,因此列主元得到的结果是不交换,即与正常Gauss消去法得到一样的结果.

题目. 10. A是正定矩阵, 对A执行一步Gauss消去得到:

$$\begin{pmatrix} a_{11} & a_1^T \\ 0 & A_2 \end{pmatrix}$$

证明: A2是正定矩阵.

解答. A是正定矩阵. 但是A不一定是对称矩阵

但由于 $x^T A x = x^T A^T x = 0$ (since 这是一个标量),所以 $x^T (\frac{A - A^T}{2}) x = 0$,因此 $x^T (\frac{A + A^T}{2}) x = 0$ 反过来, $x^T (\frac{A + A^T}{2}) x = 0$,那么 $x^T (\frac{A + A^T}{2}) x = 0$,有 $x^T (\frac{A + A^T}{2}) x = x^T A x = 0$ 因此我们不妨假设正定矩阵A是对称的.

对于正定矩阵A, 经过一步Gauss消元之后得到:

$$A = \begin{pmatrix} a_{11} & \alpha^T \\ \alpha & A_{22} \end{pmatrix} \Rightarrow LA = \begin{pmatrix} a_{11} & \alpha^T \\ 0 & A_{22} - \frac{1}{a_{11}} \alpha \alpha^T \end{pmatrix}$$

要证明A2正定, 即:

$$x^T (A_{22} - \frac{1}{a_{11}} \alpha \alpha^T) x = 0, \quad \forall x \in \mathbb{R}^{n-1}$$

根据条件: (注意, 这里可以取 $\sqrt{a_{11}}$ 是因为A是正定矩阵, 所有对角元都是正的)

取 $y = -\frac{x^T\alpha}{a_{11}}$, 得到:

$$(\sqrt{a_{11}}y + \frac{x^T\alpha}{\sqrt{a_{11}}})^2 + x^T(A_{22} - \frac{1}{a_{11}}\alpha\alpha^T)x$$
$$= x^T(A_{22} - \frac{1}{a_{11}}\alpha\alpha^T)x \ge 0$$

且取0当且仅当x=0,得证.

解答. 订正的解法: 因为正定矩阵不一定是对称矩阵 考虑:

$$\begin{pmatrix} x & y^T \end{pmatrix} \begin{pmatrix} a_{11} & \alpha^T \\ \beta & A_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = a_{11}y^2 + x\alpha^T y + xy^T \beta + x^T A_{22}x$$

$$= (\sqrt{a_{11}}x + \frac{1}{\sqrt{a_{11}}}y^T \beta)(\sqrt{a_{11}}x + \frac{1}{\sqrt{a_{11}}}\alpha^T y) + y^T (A_{22} - \frac{1}{a_{11}}\beta\alpha^T)y$$

取 $x = -\frac{y^T\alpha}{a_{11}}$, 有:

$$\begin{pmatrix} x & y^T \end{pmatrix} \begin{pmatrix} a_{11} & \alpha^T \\ \beta & A_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = y^T \underbrace{(A_{22} - \frac{1}{a_{11}} \beta \alpha^T)}_{A_2} y$$

因此取任意的向量 $y \in \mathbb{R}^{n-1}$, 有 $y^T A_2 y \ge 0$

因此A正定可以推出 A_2 正定.

题目. 14. 假定已知 $A \in \mathbb{R}^{n \times n}$ 的三角分解A = LU, 设计算法来计算 A^{-1} .

解答. 形式上: $A^{-1} = U^{-1}L^{-1}$

方法一: 根据第1题的算法, 我们有了下三角矩阵求逆的算法, 那么可以得到 L^{-1} , 以及 $(U^T)^{-1} = (U^{-1})^T$

自然就得到了 $A^{-1} = ((U^T)^{-1})^T L^{-1}$.

这样的计算复杂度是 $O(n^3) + O(n^3) + O(n^3) = O(n^3)$, 注意加法的最后一个是矩阵乘法

方法二: 因为 $AA^{-1} = I$, 那么 A^{-1} 的第j列是 $Ax_j = e_j$ 的解, 即 $LUx_j = e_j$ 的解, 那么先解 $Ly = e_j$, 再解 $Ux_j = y$ 就可以解出 x_j .

这样的计算复杂度是 $n \cdot O(n^2) = O(n^3)$,因为前代法和回代法的复杂度都是 $O(n^2)$

绷: 题目是求 $A^{-1}(i:j)$, 那么直接用方法二即可:

 $Ax_i = e_i$ 的解向量的地i个元素, $O(n^2)$

题目. 19. 若 $A = LL^T$ 是A的Cholesky分解, 试证: L的i阶顺序主子阵 L_i 正好是A的i阶顺序主子阵 A_i 的Cholesky因子.

解答. 根据Cholesky分解,有:

$$A = LL^{T} = \begin{pmatrix} L_{i} & 0 \\ L_{21} & L_{22} \end{pmatrix} \begin{pmatrix} L_{i}^{T} & L_{21}^{T} \\ 0 & L_{22}^{T} \end{pmatrix} = \begin{pmatrix} L_{i}L_{i}^{T} & L_{i}L_{21}^{T} \\ L_{21}L_{i}^{T} & L_{22}L_{22}^{T} \end{pmatrix}$$

因此自然有: $A_i = L_i L_i^T$

2 线性方程组的敏度分析

题目. 2. 证明: 当且仅当x和y线性相关且 $x^Ty \ge 0$ 时, 才有:

$$||x + y||_2 = ||x||_2 + ||y||_2$$

解答. (\Rightarrow) : 已知 $y = kx, x^Ty \ge 0$, 那么 $y = kx, k \ge 0$, 那么等式成立:

$$||x + y||_2 = (1 + k)||x||_2 = ||x||_2 + ||y||_2$$

(\Leftarrow): 已知等式成立, 考虑y关于x的正交分解 $y=kx+(y-kx), k=\frac{x^Ty}{x^Tx}$, 又因为等式成立, 两边平方后代入:

$$\begin{aligned} \|x+y\|_2^2 &= (1+k)^2 \|x\|_2^2 + \|y-kx\|_2^2 \\ \|x\|_2^2 + \|y\|_2^2 + 2\|x\|_2 \|y\|_2 &= \|x\|_2^2 + k^2 \|x\|_2^2 + \|y-kx\|_2^2 + 2\|x\|_2 \|y\|_2 \\ \|x+y\|_2 &= \|x\|_2 + \|y\|_2 \Rightarrow k \|x\|_2 = \|kx+y-kx\|_2 \\ &\Rightarrow \|y-kx\|_2^2 = 0 \iff y = kx \end{aligned}$$

因此x和y线性相关, 又因为:

$$||x + y||_2 = |1 + k|||x||_2 = (1 + |k|)||x||_2$$

故
$$k > 0$$
, 即 $x^T y > 0$.

题目. 3. 证明: 如果 $A = [a_1, \dots, a_n]$ 是按列分块的, 那么

$$||A||_F^2 = ||a_1||_2^2 + \dots + ||a_n||_2^2$$

解答.

$$||A||_2 = \sum_{j=1}^n (\sum_{i=1}^n a_{ij}^2) = \sum_{j=1}^n ||a_j||_2^2$$

得证

题目. 4. 证明:

$$||AB||_F \le ||A||_2 ||B||_F$$

 $||AB||_F \le ||A||_F ||B||_2$

解答. 方法一: 考虑 $||AB||_F^2 = \operatorname{tr}(B^*A^*AB^*)$

对于Hermite矩阵 A^*A , 首先, 它的对角元都是非负实数(因为第i个对角元是A的第i列的模平方) 其次, 根据谱定理, 它可以酉对角化, 即存在酉矩阵U, 使得 $A^*A = U^*DU$, D是对角矩阵, 且对角元都是非负实数.

因此有:

$$||AB||_F^2 = \operatorname{tr}(B^*A^*AB) = \operatorname{tr}(B^*U^*DUB) = \operatorname{tr}(DB^*U^*UB)$$

$$\leq \max_i(d_{ii})\operatorname{tr}(B^*U^*UB) = \max_i(d_{ii})\operatorname{tr}(U^*B^*BU)$$

$$= \max_i(d_{ii})\operatorname{tr}(B^*B) = ||A||_2^2||B||_F^2$$

同理有:

$$||AB||_F^2 = \operatorname{tr}(B^*A^*AB) = \operatorname{tr}(A^*B^*BA) = \operatorname{tr}(A^*\tilde{U}^*\tilde{D}\tilde{U}A) = \operatorname{tr}(\tilde{D}A^*\tilde{U}^*\tilde{U}A)$$

$$\leq \max_i(\tilde{d}_{ii})\operatorname{tr}(A^*\tilde{U}^*\tilde{U}A) = \max_i(\tilde{d}_{ii})\operatorname{tr}(\tilde{U}^*A^*A\tilde{U})$$

$$= \max_i(\tilde{d}_{ii})\operatorname{tr}(A^*A) = ||B||_2^2||A||_F^2 = ||A||_F^2||B||_2^2$$

上面的方法用到的性质有:

- 矩阵成绩的trace中, 乘积可以互换
- 矩阵的trace在相似变化下保持不变, 因为 $tr(U^*CU) = tr(U^*UC) = tr(C)$.

方法二:

考虑 $||AB||_F^2 = \sum_{i=1}^n ||AB[i,:]||_2^2$,而AB的第i行就是A的第i行乘B: $AB[i,:] = \alpha_i^T B$,因此

$$||AB||_F^2 = \sum_{i=1}^n ||AB[i,:]||_2^2 = \sum_{i=1}^n ||\alpha_i^T B||_2^2 \le \sum_{i=1}^n ||\alpha_i^T ||_2^2 ||B||_2^2 = ||A||_F^2 ||B||_2^2$$

换一个角度, AB的第j行就是A乘上B的第j列: $AB[:,j] = A\beta_j$, 因此:

$$||AB||_F^2 = \sum_{j=1}^n ||AB[:,j]||_2^2 = \sum_{j=1}^n ||A\beta_j||_2^2 \le \sum_{j=1}^n ||A||_2^2 ||\beta_j||_2^2 = ||A||_2^2 ||B||_F^2$$

这里用到了:

• 矩阵范数和向量范数的相容性: $||Ax||_2 \le ||A||_2 ||x||_2$

题目的注记. (1) F范数(平方)的两种表示: $\|A\|_F^2 = \sum_{i,j} |a_{ij}|^2 = \operatorname{tr}(A^T A)$, 注意在复矩阵的情况下 为tr(A*A), 共轭转置.

- (2) 共轭转置不变的复矩阵,即Hermite矩阵,是可以酉对角化的(即,实矩阵意义下的正交对角化)
- (3) $A^T A$ 的特征值都是非负实数

题目. 8. 若||A|| < 1, 且||I|| = 1, 证明:

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

解答.因为||A|| < 1,且任意的矩阵范数都大于等于谱半径,因此 $\rho(A) \le ||A|| < 1$ 因为 $\rho(A) < 1 \iff \sum_{k=0}^{\infty} A^k$ 收敛, 因此 $\sum_{k=0}^{\infty} A^k$ 存在 因为 $(I-A)(I+A+\cdots+A^k) = I-A^{k+1}$,且 $||A^{k+1}|| \le ||A||^{k+1} \to 0 \Rightarrow A^{k+1} \to \mathbf{0}$ 因此可以证明: $(I-A)\sum_{k=0}^{\infty} A^k = I$, 即I-A可逆, $(I-A)^{-1} = \sum_{k=0}^{\infty} A^k$ 注意题目条件||I|| = 1, 我们有

$$\|(I-A)^{-1}\| = \|\sum_{k=0}^{\infty} A^k\| \le \sum_{k=0}^{\infty} \|A^k\| \le \sum_{k=0}^{\infty} \|A\|^k = \frac{\|I\| \cdot (1 - \lim_{k \to \infty} \|A\|^k)}{1 - \|A\|} = \frac{1}{1 - \|A\|}$$

题目的注记. 需要记住的结论:

- A的任意矩阵范数大于等于谱半径: $\rho(A) \leq ||A||$, 即, 这一条对任意矩阵成立
- A给定, 谱半径是矩阵(算子)范数的下确界, $\forall \epsilon > 0$, 存在算子范数, 使得 $||A|| \leq \rho(A) + \epsilon$
- $\lim_{k\to\infty} A^k = 0 \iff \sum_{k=0}^{\infty} A^k$ 收敛 $\iff \rho(A) < 1$, 若收敛, $(I-A)^{-1} = \sum_{k=0}^{\infty} A^k$

题目. 11. 设

$$A = \begin{bmatrix} 375 & 374 \\ 752 & 750 \end{bmatrix}$$

- (1) 计算 A^{-1} 和 $\kappa_{\infty}(A)$

$$Ax = b$$
, $A(x + \delta x) = b + \delta b$

而且 $\frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}}$ 很小,但 $\frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}}$ 很大 (2) 选择 $b,\delta b,x,\delta x$,使得

$$Ax = b$$
, $A(x + \delta x) = b + \delta b$

而且 $\frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}}$ 很小,但 $\frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}}$ 很大

解答. 观察矩阵, 设a = 374, 那么:

$$A = \begin{bmatrix} a & a-1 \\ 2(a+1) & 2a \end{bmatrix}$$

记住了二阶矩阵求逆的标准公式:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{|ad - bc|} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

因此有:

$$A^{-1} = \frac{1}{2} \begin{pmatrix} 2a & 1-a \\ -2a-2 & a \end{pmatrix} = \begin{pmatrix} 375 & -187 \\ -376 & 187.5 \end{pmatrix}$$

计算 $\kappa_{\infty}(A) = ||A||_{\infty} ||A^{-1}||_{\infty},$ 有(行范数计算)

$$\kappa_{\infty}(A) = ||A||_{\infty} ||A^{-1}||_{\infty} = (752 + 750) \cdot (376 + 187.5) = 846377$$

(2) 考虑先进行敏度分析, $A(x + \delta x) = b + \delta b$, 因为Ax = b, 所以 $\delta x = A^{-1}\delta b$, 因此有 $\|\delta x\|_{\infty} \le \|A^{-1}\|_{\infty} \|\delta b\|_{\infty}$, 两边同时除以 $\|b\|_{\infty} \le \|A\|_{\infty} \|x\|_{\infty}$, 有:

$$\frac{\|\delta x\|_{\infty}}{\|A\|_{\infty}\|x\|_{\infty}} \le \frac{\|\delta x\|_{\infty}}{\|b\|_{\infty}} \le \|A^{-1}\|_{\infty} \frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}}$$

$$\Rightarrow \frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}} \le \|A\|_{\infty} \|A^{-1}\|_{\infty} \frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}} = \kappa(A) \frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}}$$

虽然我们已知 $\kappa(A)$ 很大, 但这似乎对我们的构造没有什么帮助. 因为这个不等式是一个关于上下界的估计, 但实际值可以不在等号附近.

好的思路: 因为矩阵A的条件数很大(实际上是 $||A||_{\infty}$ 和 $||A^{-1}||_{\infty}$ 都不小), 所以对于 $A\delta x = \delta b$ (已知 δx , $\hat{p}\delta b$)和 $\delta x = A^{-1}\delta b$ (已知 δb , $\hat{p}\delta x$), 即使已知的对象比较小, 被计算的部分也会被放得很大.

取
$$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
, 那么 $b = Ax = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, 先取 $\delta b = \begin{pmatrix} \epsilon \\ 0 \end{pmatrix}$, 计算得到 $\delta x = A^{-1}\delta b = \begin{pmatrix} 375\epsilon \\ -376\epsilon \end{pmatrix}$, 那么 $(\epsilon > 0)$

$$\begin{split} \frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}} &= \frac{\epsilon}{2}, \frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}} = \frac{376\epsilon}{1} = 376\epsilon\\ \epsilon &= 1 \Rightarrow \frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}} = \frac{1}{2}, \frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}} = 376 \end{split}$$

(3) 类似于(2), 取
$$x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$
, 那么 $b = Ax = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, 先取 $\delta x = \begin{pmatrix} \epsilon \\ 0 \end{pmatrix}$, 计算 $\delta b = A\delta x = \begin{pmatrix} 375\epsilon \\ 752\epsilon \end{pmatrix}$, 那么 $(\epsilon > 0)$

$$\frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}} = \frac{752\epsilon}{2}, \frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}} = \frac{\epsilon}{1} = \epsilon$$

$$\epsilon = 1 \Rightarrow \frac{\|\delta b\|_{\infty}}{\|b\|_{\infty}} = 376, \frac{\|\delta x\|_{\infty}}{\|x\|_{\infty}} = 1$$