数值线代の题库二

by 23 大数据 miracle

1. 证明:$x \in \mathcal{X}_{\text{LS}}$ 当且仅当

$$A^{\text{T}}Ax = A^{\text{T}}b$$

Solution: 设 $x \in \mathcal{X}_{\text{LS}}$. 即

$$x = \underset{x\in\mathbb{R}^n}{\operatorname{argmin}}\{\|b - Ax\|_2\}$$

令 $b = b_1 + b_2$, 其中 $b_1 \in \mathcal{R}(A)$ 且 $b_2 \in \mathcal{R}(A)^{\perp}$. 则有 $Ax = b_1$, 令

$$r(x) = b - Ax = b - b_1 = b_2 \in \mathcal{R}(A)^{\perp}$$

因此 $A^{\text{T}}r(x) = A^{\text{T}}b_2 = 0$. 将 $r(x) = b - Ax$ 代入 $A^{\text{T}}r(x) = 0$ 即得 $A^{\text{T}}Ax = A^{\text{T}}b$.

反之, 设 $x \in \mathbb{R}^n$ 满足 $A^{\text{T}}Ax = A^{\text{T}}b$, 则对任意的 $y \in \mathbb{R}^n$, 有

$$\begin{aligned}\|b - A(x+y)\|_2^2 &= \|b - Ax\|_2^2 - 2y^{\text{T}}A^{\text{T}}(b - Ax) + \|Ay\|_2^2 \\ &= \|b - Ax\|_2^2 + \|Ay\|_2^2 \\ &\geq \|b - Ax\|_2^2\end{aligned}$$

由此即得 $x \in \mathcal{X}_{\text{LS}}$. 综上, 命题得证.

2. 证明 QR 分解定理, 即设 $A \in \mathbb{R}^{m\times n}$ $(m \geq n)$, 则 A 有 QR 分解:

$$A = Q\begin{bmatrix} R \\ 0 \end{bmatrix}$$

其中 $Q \in \mathbb{R}^{m\times m}$ 是正交矩阵, $R \in \mathbb{R}^{n\times n}$ 是具有非负对角元的上三角阵；而且当 $m = n$ 且 A 非奇异时, 上述分解是唯一的.

Solution: 先证明 QR 分解的存在性. 对 n 用数学归纳法. 当 $n = 1$ 时, 此定理自然成立. 现假设已经证明定理对所有的 $p \times (n-1)$ 矩阵成立, 这里假定 $p \geq n-1$. 设 $A \in \mathbb{R}^{m\times n}$ 的第一列为 a_1, 则由 Householder 变换相关知识知存在正交矩阵 $Q_1 \in \mathbb{R}^{m\times m}$, 使得

$$Q_1^{\text{T}}a_1 = \|a_1\|_2 e_1$$

于是, 有

$$Q_1^{\mathrm{T}} A = \begin{bmatrix} \|a_1\|_2 & v^{\mathrm{T}} \\ 0 & A_1 \end{bmatrix}$$

对 $(m-1)\times(n-1)$ 矩阵 A_1 应用归纳法假设, 得

$$A_1 = Q_2 \begin{bmatrix} R_2 \\ 0 \end{bmatrix}$$

其中 Q_2 是 $(m-1)\times(m-1)$ 正交矩阵, 而 R_2 是具有非负对角元的 $(n-1)\times(n-1)$ 上三角阵. 这样, 令

$$Q = Q_1 \begin{bmatrix} 1 & 0 \\ 0 & Q_2 \end{bmatrix}, \quad R = \begin{bmatrix} \|a_1\|_2 & v^{\mathrm{T}} \\ 0 & R_2 \end{bmatrix}$$

则有

$$Q \begin{bmatrix} R \\ 0 \end{bmatrix} = Q_1 \begin{bmatrix} 1 & 0 \\ 0 & Q_2 \end{bmatrix} \begin{bmatrix} \|a_1\|_2 & v^{\mathrm{T}} \\ 0 & R_2 \\ 0 & 0 \end{bmatrix} = Q_1 Q_1^{\mathrm{T}} A = A$$

即 Q 与 R 满足定理的要求. 于是, 由归纳法原理知存在性得证.

再证唯一性. 设 $m=n$ 且 A 非奇异, 并假定 $A = QR = \tilde{Q}\tilde{R}$, 其中 $Q, \tilde{Q} \in \mathbb{R}^{m\times m}$ 是正交矩阵, $R, \tilde{R} \in \mathbb{R}^{n\times n}$ 是具有非负对角元的上三角阵. A 非奇异蕴涵着 R 和 $\hat{R}$ 的对角元均为正数. 因此, 我们有

$$\tilde{Q}^{\mathrm{T}} Q = \tilde{R} R^{-1}$$

此矩阵既是正交矩阵又是对角元均为正数的上三角阵, 故只能是单位矩阵, 从而必有 $\tilde{Q} = Q$, $\tilde{R} = R$, 即分解是唯一的.

3. 假设 x 和 y 是 $\mathbb{R}^n$ 中的两个单位向量. 给出一种使用 Givens 变换的算法, 计算一个正交矩阵 Q, 使得 $Qx = y$.

Solution: 核心思路: 使用 Givens 变换分别将 x 和 y 旋转到 e_1, 得到正交矩阵 Q_1 和 Q_2 满足 $Q_1 x = e_1$ 与 $Q_2 y = e_1$, 于是便有 $Q = Q_2^{\mathrm{T}} Q_1$ 满足 $Qx = y$.

以下是后面各函数的简要说明:

1. `givens`(a, b): 计算单次 Givens 变换的旋转参数 c, s.
2. `rotVec`(v, i, j, c, s): 对向量 v 的第 i, j 分量应用 Givens 变换 (参数为 c, s).
3. `rotMat`(Q, i, j, c, s): 对矩阵 Q 的第 i, j 行应用 Givens 变换 (参数为 c, s), 来记录总变换.
4. `rotFstAx`(v): 用 Givens 旋转将向量 v 逐步对齐到 e_1(主坐标轴), 返回累积正交矩阵 U.
5. `main`(x, y): 主函数, 即计算最后的正交矩阵 Q.

Algorithm 1: 用 Givens 算法计算正交阵 Q, 满足 $Qx = y$, 且 $\|x\|_2 = \|y\|_2 = 1$

---

function: `[c, s]=givens(`a, b`)`

  if $|b| < \epsilon$ **then**

    $c = \text{sign}(a)$; $s = 0$

  else

    if $|a| < \epsilon$ **then**

      $c = 0$; $s = \text{sign}(b)$

    else

      $r = \sqrt{a^2 + b^2}$; $c = a/r$; $s = b/r$

    end

  end

end

---

function: `[v]=rotVec(`v, i, j, c, s`)`

  $\sigma = c \cdot v_i + s \cdot v_j$

  $v_j = -s \cdot v_i + c \cdot v_j$

  $v_i = \sigma$

end

---

function: `[Q]=rotMat(`Q, i, j, c, s`)`

  for $k = 1 : n$ **do**

    $\tau = c \cdot Q_{i,k} + s \cdot Q_{j,k}$

    $Q_{j,k} = -s \cdot Q_{i,k} + c \cdot Q_{j,k}$

    $Q_{i,k} = \tau$

  end

end

---

function: `[U]=rotFstAx(`v`)`

  $U = I_n$

  for $k = n : 2$ **do**

    $i = k - 1$; $j = k$

    $(c, s) = \texttt{givens}(v_i, v_j)$

    $v = \texttt{rotVec}(v, i, j, c, s)$

    $U = \texttt{rotMat}(U, i, j, c, s)$

  end

end

---

function: `[Q]=main(`x, y`)`

  $Q_1 = \texttt{rotFstAx}(x)$; $Q_2 = \texttt{rotFstAx}(y)$

  $Q = Q_2^{\mathrm{T}} Q_1$

end

其中关于该算法正确性的具体证明:

$$
\begin{aligned}
Qx &= (Q_2^{\mathrm{T}} Q_1)x \\
&= Q_2^{\mathrm{T}}(Q_1 x) \\
&= Q_2^{\mathrm{T}} e_1 \\
&= Q_2^{\mathrm{T}}(Q_2 y) \\
&= y
\end{aligned}
$$

证毕.

4. 设 x 和 y 是 $\mathbb{R}^n$ 中的两个非零向量. 给出一种算法来确定一个 Householder 变换 H, 使得 $Hx = \alpha y$, 其中 $\alpha \in \mathbb{R}$.

Solution: 当 x 和 y 线性相关时, 直接取 $H = I$. 否则按照以下思路计算出 H 即可:

$$
H = I - 2ww^{\mathrm{T}} \Rightarrow Hx = x - 2w(w^{\mathrm{T}}x) = \alpha y
$$

则有 w 平行于 $(x - \alpha y)$, 又 $\|w\|_2 = 1$, 故取

$$
w = \frac{x - \alpha y}{\|x - \alpha y\|_2}
$$

具体算法如下:

Algorithm 2: 用 Householder 变换求得对应矩阵 H, 满足 $Hx = \alpha y$, $\alpha \in \mathbb{R}$

function: `[H]=house`(x, y)
  $\alpha = \|x\|_2 / \|y\|_2$
  $z = x - \alpha y$
  $w = z / \|z\|_2$
  $H = I - 2ww^{\mathrm{T}}$
end

其中关于该算法正确性的具体证明 (此时 $z \neq 0$):
展开有

$$
\begin{aligned}
\|z\|_2^2 &= (x^{\mathrm{T}} - \alpha y^{\mathrm{T}})(x - \alpha y) \\
&= (x^{\mathrm{T}} - \alpha y^{\mathrm{T}})x + \alpha(\alpha y^{\mathrm{T}} - x^{\mathrm{T}})y \\
&= z^{\mathrm{T}}x + x^{\mathrm{T}}x - \alpha x^{\mathrm{T}}y \\
&= 2(z^{\mathrm{T}}x)
\end{aligned}
$$

于是

$$
\begin{aligned}
Hx &= x - 2ww^{\mathrm{T}}x \\
&= x - 2\left(\frac{z}{\|z\|_2}\right)\left(\frac{z^{\mathrm{T}}x}{\|z\|_2}\right) \\
&= x - \frac{2(z^{\mathrm{T}}x)}{\|z\|_2^2}z \\
&= x - z \\
&= x - (x - \alpha y) \\
&= \alpha y
\end{aligned}
$$

证毕.

5. 设 $A \in \mathbb{R}^{m\times n}$, 且存在 $X \in \mathbb{R}^{n\times m}$ 使得对于每一个 $b \in \mathbb{R}^m$, $x = Xb$ 均极小化 $\|Ax - b\|_2$. 证明:$AXA = A$ 和 $(AX)^{\mathrm{T}} = AX$.

Solution: 设 $R(A)$ 为线性映射 $f(x) = Ax$ 的值域, 则 $R(A)$ 为 $\mathbb{R}^m$ 的线性子空间.

1. 考虑 $b \in R(A)$, 这时存在 $y \in \mathbb{R}^n$ 使得 $Ay = b$. 因此 $\|Ax - b\|_2$ 最小值为 0, $Ax - b = (AX - E)b = 0$. 依次取 $b_1, b_2, \cdots, b_n$ 为 A 的列向量, 则有 $(AX - E)b_i = 0$, $i = 1, 2, \cdots, n$. 即
$$(AX - E)A = 0$$
由此得
$$AXA = A$$
2. 对于所有 $b \in \mathbb{R}^m$, 由于 $\|Ax - b\|_2$ 取到极小值, 由课本定理可知这时 $b - Ax$ 与 Ax 互相垂直, 即
$$(Ax)^{\mathrm{T}}(Ax - b) = 0$$
再将 $x = Xb$ 代入, 有
$$(AXb)^{\mathrm{T}}(AXb - b) = b^{\mathrm{T}}(AX)^{\mathrm{T}}(AX - E)b = 0$$
由于 b 的任意性, 可知
$$(AX)^{\mathrm{T}}(AX - E) = 0$$
因此,
$$(AX)^{\mathrm{T}}AX = (AX)^{\mathrm{T}}$$
两边取转置得:
$$(AX)^{\mathrm{T}}AX = AX$$
比较上两式, 即有
$$(AX)^{\mathrm{T}} = AX$$

另解: 因为 x 极小化 $\|Ax-b\|_2$, 所以 $\forall b\in\mathbb{R}^m$, 有 $A^{\mathrm{T}}b=A^{\mathrm{T}}Ax=A^{\mathrm{T}}AXb$, 于是

$$(A^{\mathrm{T}}AX-A^{\mathrm{T}})b=0$$

依次取 $b=e_1,e_2,\ldots,e_n$, 有

$$(A^{\mathrm{T}}AX-A^{\mathrm{T}})(e_1,e_2,\ldots,e_n)=(A^{\mathrm{T}}AX-A^{\mathrm{T}})I=0 \implies A^{\mathrm{T}}AX=A^{\mathrm{T}}$$

于是

$$A=X^{\mathrm{T}}A^{\mathrm{T}}A \implies AX=X^{\mathrm{T}}A^{\mathrm{T}}AX=(AX)^{\mathrm{T}}AX \text{ 对称}$$

则有 AX 对称, 即

$$(AX)^{\mathrm{T}}=AX$$

进而

$$AXA=(AX)^{\mathrm{T}}A=X^{\mathrm{T}}A^{\mathrm{T}}A=A$$

6. 利用等式

$$\|A(x+\alpha w)-b\|_2^2=\|Ax-b\|_2^2+2\alpha w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)+\alpha^2\|Aw\|_2^2$$

证明: 如果 $x\in\mathcal{X}_{\mathrm{LS}}$, 那么 $A^{\mathrm{T}}Ax=A^{\mathrm{T}}b$.

Solution: 定义泛函:

$$f(x)=\|Ax-b\|_2^2$$

对于任意 $x\in\mathcal{X}_{LS}$, x 是 $f(x)$ 的一个极小值点. 由于 f 是二次函数, 其在 $\mathbb{R}^n$ 上处处可微（特别地, 任意方向导数均存在）, 且定义域为开集, 故 x 作为内点极小值点, 其方向导数在任意方向为零.

考虑任意方向向量 $w\in\mathbb{R}^n$. 方向导数在方向 w 上定义为:

$$\lim_{\alpha\to 0}\frac{f(x+\alpha w)-f(x)}{\alpha}$$

利用题中等式有:

$$\frac{f(x+\alpha w)-f(x)}{\alpha}=2w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)+\alpha\|Aw\|_2^2$$

取极限 $\alpha\to 0$:

$$\lim_{\alpha\to 0}\frac{f(x+\alpha w)-f(x)}{\alpha}=2w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)$$

由方向导数为零:

$$w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)=0$$

此式对所有 $w\in\mathbb{R}^n$ 成立. 若固定 $v=A^{\mathrm{T}}(Ax-b)$, 则 $w^{\mathrm{T}}v=0$ 对所有 w 成立. 取 $w=v$, 得:

$$v^{\mathrm{T}}v=\|v\|_2^2=0$$

故 $v=0$, 即:

$$A^{\mathrm{T}}Ax = A^{\mathrm{T}}b$$

另解: 注意到, $\forall \alpha \in \mathbb{R}$ 有

$$\begin{aligned} 0 &\leq \|A(x+\alpha w)-b\|_2^2 - \|Ax-b\|_2^2 \\ &= 2\alpha w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b) + \alpha^2\|Aw\|_2^2 \\ &= \|Aw\|_2^2\alpha^2 + 2w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)\alpha \end{aligned}$$

于是, 考虑其对应一元二次方程的判别式

$$\Delta = 4[w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)]^2 - 4\cdot\|Aw\|_2^2\cdot 0 \leq 0$$

即

$$w^{\mathrm{T}}A^{\mathrm{T}}(Ax-b)=0 \xRightarrow{w\text{的任意性}} A^{\mathrm{T}}Ax = A^{\mathrm{T}}b$$

7. 若迭代矩阵 M 的范数 $\|M\| = q < 1$, 则迭代法 $x_k = Mx_{k-1}+g$ 所产生的近似解 x_k 与准确解 x_* 的误差有如下估计式:

$$\|x_k - x_*\| \leq \frac{q^k}{1-q}\|x_1-x_0\|, \quad \|x_k-x_*\| \leq \frac{q}{1-q}\|x_{k-1}-x_k\|$$

Solution: 以下分别对两个估计式进行证明:

1. 令 $y_k = x_k - x_*$ 并结合 $x_* = Mx_*+g$, 有

$$y_k = My_{k-1} \xRightarrow{\text{迭代}} y_k = M^k y_0$$

两边取范数, 得

$$\|y_k\| = \|M^k y_0\| \leq \|M\|^k\|y_0\| = q^k\|y_0\|$$

现在估计 y_0. 根据定义, 我们有

$$\begin{aligned} \|y_0\| &= \|x_0 - x_*\| \\ &\leq \|x_0-x_1\| + \|x_1-x_*\| \\ &= \|x_0-x_1\| + \|My_0\| \\ &\leq \|x_0-x_1\| + q\|y_0\| \end{aligned}$$

从而有

$$\|y_0\| \leq \frac{1}{1-q}\|x_0-x_1\|$$

将此不等式代入 $\|y_k\| \leq q^k\|y_0\|$, 即有:

$$\|x_k-x_*\| \leq \frac{q^k}{1-q}\|x_1-x_0\|$$

2. 因为

$$\begin{aligned}\|x_k - x_*\| &= \|M(x_{k-1} - x_*)\| \le q\|x_{k-1} - x_*\| \\ &\le q\|x_{k-1} - x_k\| + q\|x_k - x_*\|\end{aligned}$$

所以有

$$\|x_k - x_*\| \le \frac{q}{1-q}\|x_{k-1} - x_k\|$$

另解: 第二个估计式同上, 下证明第一个估计式:

使用数学归纳法, 首先当 $k = 1$ 时, 需证

$$\|x_1 - x_*\| \le \frac{q}{1-q}\|x_1 - x_0\|$$

这就是第二个估计式在 $k = 1$ 时的情形.

然后假设在 $k - 1$ 下估计式成立, 即

$$\|x_{k-1} - x_*\| \le \frac{q^{k-1}}{1-q}\|x_1 - x_0\|$$

于是

$$\|x_k - x_*\| = \|M(x_{k-1} - x_*)\| \le q \cdot \frac{q^{k-1}}{1-q}\|x_1 - x_0\| = \frac{q^k}{1-q}\|x_1 - x_0\|$$

由归纳假设知第一个估计式得证, 证毕.

8. 若线性方程组 $Ax = b$ 的系数矩阵 A 对称, 而且其对角元 $a_{ii} > 0$ $(i = 1, \cdots, n)$, 则 Jacobi 迭代法收敛的充分必要条件是 A 和 $2D - A$ 都正定.

Solution: 记 $B = D^{-1}(L + U) = D^{-1}(D - A) = I - D^{-1}A$, 而 $D = \mathrm{diag}(a_{ii})$ 的对角元均大于零, 故

$$B = I - D^{-1}A = D^{-\frac{1}{2}}(I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}.$$

由 A 的对称性推出 $I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 也是实对称的, 特征值均为实数, 而且它与 B 相似, 有相同的特征值, 从而 B 的特征值均为实数. 此外, 由上式立即可得

$$I - B = D^{-\frac{1}{2}}(D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}},$$

$$I + B = D^{-\frac{1}{2}}(2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}.$$

又

$$\begin{aligned}\text{Jacobi 迭代法收敛} &\Leftrightarrow \rho(B) < 1, \text{即矩阵}B\text{的所有特征值}\lambda\text{满足}|\lambda| < 1 \\ &\Leftrightarrow I - B\text{和}I + B\text{的特征值均为正实数}\ (B\text{的特征值均为实数}) \\ &\Leftrightarrow D^{-\frac{1}{2}}AD^{-\frac{1}{2}}\text{和}2I - D^{-\frac{1}{2}}AD^{-\frac{1}{2}}\text{的特征值均为正实数} \\ &\Leftrightarrow A\text{和}2D - A\text{均正定}\end{aligned}$$

从而定理得证.

9. 若矩阵 A 是严格对角占优的或不可约对角占优的, 则 A 非奇异且则 Jacobi 迭代法和 G-S 迭代法都收敛.

Solution: 以下分两部分进行证明:

1. 对两种矩阵非奇异性的证明:
   当 A 为严格对角占优时: 用反证法. 假设 A 奇异, 则齐次方程组 $Ax=0$ 有非零解 x. 不妨取 $|x_i| = \|x\|_\infty = 1$, 则有

$$|a_{ii}| = |a_{ii}x_i| = \left|\sum_{j=1}^n a_{ij}x_j\right| \leq \sum_{\substack{j=1\\j\neq i}}^n |a_{ij}|.$$

   这与 A 严格对角占优矛盾, 因此 A 非奇异.

   当 A 为不可约对角占优时: 仍用反证法. 设 x 满足 $\|x\|_\infty = 1$, 使得 $Ax = 0$, 并定义

$$\mathcal{S} = \{i : |x_i| = 1\},$$

$$\mathcal{T} = \{k : |x_k| < 1\}.$$

   显然 $\mathcal{S} \cup \mathcal{T} = \mathcal{W} = \{1, 2, \cdots, n\}$, $\mathcal{S} \cap \mathcal{T} = \emptyset$, 而且 $\mathcal{T}$ 非空. 这是因为: 假设 $\mathcal{T}$ 为空集, 则 x 的各个分量的绝对值均为 1, 那么不论 i 为 $\mathcal{S}$ 中的何值均有

$$|a_{ii}| \leq \sum_{\substack{j=1\\j\neq i}}^n |a_{ij}|.$$

   这与 A 弱严格对角占优矛盾. 因此, 在 $\mathcal{T}$ 非空的情况下, 因为 A 不可约, 必定存在 i, k, 使得

$$a_{ik} \neq 0, \quad i \in \mathcal{S}, k \in \mathcal{T}.$$

   于是, $|a_{ik}x_k| < |a_{ik}|$, 并且

$$\begin{aligned}
|a_{ii}| &\leq \sum_{j\in\mathcal{S}, j\neq i} |a_{ij}||x_j| + \sum_{j\in\mathcal{T}} |a_{ij}||x_j| \\
&< \sum_{j\in\mathcal{S}, j\neq i} |a_{ij}| + \sum_{j\in\mathcal{T}} |a_{ij}| \\
&= \sum_{j\neq i} |a_{ij}|.
\end{aligned}$$

   这又与 A 弱严格对角占优矛盾, 则此时 A 同样非奇异.

2. 对两种迭代法收敛性的证明:

若 A 是严格对角占优的或不可约对角占优的, 则对每个 i, 必有 $|a_{ii}| > 0$. 因此 D 可逆.

Jacobi 迭代法收敛性证明: 假设 Jacobi 迭代矩阵 $B_1 = D^{-1}(L+U)$ 的某个特征值 $|\lambda| \geq 1$. 考察矩阵 $\lambda D - L - U$, 它满足与原矩阵 A 相同的占优条件（严格对角占优或不可约对角占优）, 故 $\lambda D - L - U$ 非奇异. 由

$$\lambda I - B_1 = \lambda I - D^{-1}(L+U) = D^{-1}(\lambda D - L - U)$$

可得

$$\det(\lambda I - B_1) = \det(D^{-1})\det(\lambda D - L - U) \neq 0$$

这与 λ 是 B_1 的特征值矛盾. 因此 B_1 的所有特征值模均小于 1, Jacobi 迭代法收敛.

G-S 迭代法收敛性证明: 设 G-S 迭代矩阵为 $B_2 = (D-L)^{-1}U$. 假设存在特征值 λ 满足 $|\lambda| \geq 1$. 考察矩阵:

$$\lambda D - \lambda L - U = \lambda(D-L) + [A - (D-L)]$$

同样的, 与矩阵 A 有相同的占优条件, 因此 $\lambda D - \lambda L - U$ 非奇异, 于是,

$$\begin{aligned}\det(\lambda I - B_2) &= \det\left(\lambda I - (D-L)^{-1}U\right) \\ &= \det((D-L)^{-1})\det(\lambda D - \lambda L - U) \\ &\neq 0\end{aligned}$$

与 λ 是 B_2 的特征值矛盾. 因此 B_2 的所有特征值模均小于 1, G-S 迭代法收敛.

综上, 所有命题皆得证.

10. 设 $B \in \mathbb{R}^{n\times n}$ 满足 $\rho(B) = 0$. 证明对于任意的 $g, x_0 \in \mathbb{R}^n$, 迭代格式 $x^{(k+1)} = Bx^{(k)} + g$, $k = 0, 1, 2, \ldots$ 最多迭代 n 次就可以得到方程组 $x = Bx + g$ 的精确解.

Solution: 由于 B 谱半径为零, 故特征值都为零. 于是存在可逆矩阵 Q 和矩阵

$$J = \begin{bmatrix} 0 & \alpha_1 & * & \cdots & * \\ & 0 & \alpha_2 & \ddots & \vdots \\ & & \ddots & \ddots & * \\ & & & 0 & \alpha_{n-1} \\ & & & & 0 \end{bmatrix}$$

使得 $B = Q^{-1}JQ$. 注意到 $J^n = 0$, 因此 $B^n = Q^{-1}J^nQ = 0$. 这说明 $y_n = B^n y_0 = 0$.
即 $x_n = x_* \Leftrightarrow$ 收敛到精确解.

另解: 方程 $x = Bx + g$ 等价于 $(I - B)x = g$. 因 B 是幂零矩阵, 则 $I - B$ 可逆, 于是该方程的解存在唯一, 记为 x_*.
由迭代格式有

$$y_{k+1} = By_k \Rightarrow y_k = B^k y_0$$

考虑核空间升链, 于是

$$\{0\} = \ker(B) \subseteq \ker(B^2) \subseteq \cdots \subseteq \ker(B^m) = \mathbb{R}^n$$

其中 m 为幂零指数（即 $B^m = 0$）. 令 $d_k = \dim\ker(B^k)$, 则

$$0 = d_0 \leq d_1 \leq \cdots \leq d_m = n$$

事实上, 对于任意矩阵 B, 若存在 $p \in \mathbb{N}$ 使得 $\ker(B^p) = \ker(B^{p+1})$, 则有

$$\ker(B^p) = \ker(B^{p+1}) = \cdots = \ker(B^s),\ \forall s \geq p$$

于是存在 $0 < p \leq m$, 使得

$$0 = d_0 < \cdots < d_p = \cdots = d_m = n$$

而前面递增的部分每步维数至少增加 1, 故

$$n = d_m \geq d_0 + p = p \Rightarrow n \geq p$$

因而 $\ker(B^n) = \ker(B^p) = \ker(B^m) = \mathbb{R}^n$, 于是

$$y_n = B^n y_0 = 0 \Rightarrow x_n = x_*$$

补充对上述所用定理的证明:
已知存在 p 满足 $\ker(B^p) = \ker(B^{p+1})$, 证明:

$$\ker(B^k) = \ker(B^p),\ \forall k > p$$

注意到事实上只需证明 $\ker(B^{p+1}) \supseteq \ker(B^{p+2})$ 即可:
考虑 $\forall x \in \ker(B^{p+2})$, 即 $B^{p+2}x = 0$, 则 $B^{p+1}(Bx) = 0$, 故 $Bx \in \ker(B^{p+1}) = \ker(B^p)$. 于是 $B^p(Bx) = B^{p+1}x = 0$, 即 $x \in \ker(B^{p+1})$, 自然便有 $\ker(B^{p+1}) \supseteq \ker(B^{p+2})$, 证毕.

11. 设 A 是具有正对角元素的非奇异对称矩阵. 证明: 若求解方程组 $Ax = b$ 的 G-S 迭代法对任意初始近似皆收敛, 则 A 必定是正定的.

Solution: 设 A 为对称非奇异矩阵, 且对角元全为正. 令

$$A = D - L - L^{\mathrm{T}}$$

其中 D 为全为正对角元的对角矩阵, L 为严格下三角矩阵.

G-S 迭代格式为

$$x_{k+1} = D^{-1}Lx_{k+1} + D^{-1}L^{\mathrm{T}}x_k + D^{-1}b$$

设精确解为 x_*, 于是有

$$x_* = D^{-1}Lx_* + D^{-1}L^{\mathrm{T}}x_* + D^{-1}b$$

定义误差向量

$$y_k = x_k - x_*$$

则误差递推关系为

$$y_{k+1} = D^{-1}Ly_{k+1} + D^{-1}L^{\mathrm{T}}y_k$$

即

$$(D-L)y_{k+1} = L^{\mathrm{T}}y_k$$

具体证明:

1. **引理推导**: 定义

   $$\epsilon_k = y_k - y_{k+1}$$

   可推得

   $$Ay_{k+1} = Dy_{k+1} - Ly_{k+1} - L^{\mathrm{T}}y_{k+1} = L^{\mathrm{T}}\epsilon_k$$

   进一步有（优先消去 y_k, 则比较容易推出）

   $$\begin{aligned} y_k^{\mathrm{T}}Ay_k - y_{k+1}^{\mathrm{T}}Ay_{k+1} &= (y_{k+1}+\epsilon_k)^{\mathrm{T}}A(y_{k+1}+\epsilon_k) - y_{k+1}^{\mathrm{T}}Ay_{k+1} \\ &= \epsilon_k^{\mathrm{T}}Ay_{k+1} + y_{k+1}^{\mathrm{T}}A\epsilon_k + \epsilon_k^{\mathrm{T}}A\epsilon_k \\ &= \epsilon_k^{\mathrm{T}}L^{\mathrm{T}}\epsilon_k + \epsilon_k^{\mathrm{T}}L\epsilon_k + \epsilon_k^{\mathrm{T}}A\epsilon_k \\ &= \epsilon_k^{\mathrm{T}}D\epsilon_k \end{aligned}$$

   由于 D 正定, 故当 $\epsilon_k \neq 0$ 时,

   $$\epsilon_k^{\mathrm{T}}D\epsilon_k > 0$$

   即

   $$y_k^{\mathrm{T}}Ay_k > y_{k+1}^{\mathrm{T}}Ay_{k+1}$$

   说明未收敛时 $y_k^{\mathrm{T}}Ay_k$ 严格递减.

2. **反证得结论**: 假设 A 不是正定的, 则存在非零向量 x_0 使得

   $$x_0^{\mathrm{T}}Ax_0 \leq 0$$

   令 $b=0$（当 $b \neq 0$ 时, 由于 A 非奇异, 线性方程组 $Ax=b$ 也等价于 $A(x-A^{-1}b)=0$, 则作变量代换 $z = x - A^{-1}b$ 得 $Az=0$）, 则精确解 $x_*=0$ 且 $y_k = x_k$.

   但由前面的引理, 同时结合 $y_0 = x_0 \neq 0$（即初始向量非最终精确解）, 有

   $$y_k^{\mathrm{T}}Ay_k \leq y_1^{\mathrm{T}}Ay_1 < y_0^{\mathrm{T}}Ay_0 \leq 0,\ \forall k \in \mathbb{N}$$

故存在 $\epsilon' = y_1^{\mathrm{T}} A y_1$ 使得

$$y_k^{\mathrm{T}} A y_k \le \epsilon' < 0,\ \forall k \in \mathbb{N}$$

另一方面, 由 G-S 迭代收敛性假设应有（λ 指矩阵 A 模长最大的特征值）:

$$\|y_k^{\mathrm{T}} A y_k\|_2 \le |\lambda| \|y_k\|_2^2 = |\lambda| \|x_k\|_2^2 \to 0,\ k \to \infty$$

即

$$\lim_{k\to\infty} y_k^{\mathrm{T}} A y_k = 0$$

矛盾, 故假设不成立, A 必为正定矩阵.

12. 若存在对称正定矩阵 P, 使得

$$B = P - H^{\mathrm{T}} P H$$

为对称正定矩阵, 求证: 迭代法

$$x_{k+1} = H x_k + b, \quad k = 0, 1, 2, \cdots$$

收敛.

Solution: 设 λ 为 H 的特征值, x 为 H 属于 λ 的一个特征向量. 则（结合 H 一定是实的）

$$x^{\mathrm{H}} B x = x^{\mathrm{H}} P x - (Hx)^{\mathrm{H}} P (Hx) = (1 - |\lambda|^2) x^{\mathrm{H}} P x.$$

不妨设 $x = u + iv$, 则一定有 $u \ne 0$ 或者 $v \ne 0$, 于是（结合 B 的对称正定性）

$$x^{\mathrm{H}} B x = (u^{\mathrm{T}} - i v^{\mathrm{T}}) B (u + iv) = u^{\mathrm{T}} B u + v^{\mathrm{T}} B v > 0$$

同理 $x^{\mathrm{H}} P x > 0$, 从而 $1 - |\lambda|^2 > 0$, 即 $|\lambda| < 1 \Leftrightarrow \rho(H) < 1$（$\lambda$ 是任意的）. 因此迭代法 $x_{k+1} = H x_k + b$ 收敛.

13. 证明: 若系数矩阵 A 是严格对角占优的或不可约对角占优的, 且松驰因子 $\omega \in (0, 1)$, 则 SOR 迭代法收敛.

Solution: 用反证法, 现假定某个复数 $|\lambda| \ge 1$ 是系数矩阵 A 的 SOR 迭代矩阵 $L_\omega = (D - \omega L)^{-1}[(1-\omega)D + \omega U]$ 的特征值 (由 A 的对角占优性可知 L_ω 一定存在), 不妨令 $\lambda = \alpha + \beta i$, 则有 $\alpha^2 + \beta^2 \ge 1$, 考虑矩阵 $(\lambda + \omega - 1)D - \lambda\omega L - \omega U$, 于是有:

$$\begin{aligned}
|\lambda + \omega - 1|^2 - |\lambda|^2 \omega^2 &= (\alpha + \omega - 1)^2 + \beta^2 - \omega^2(\alpha^2 + \beta^2) \\
&= \alpha^2 - 2\alpha(1-\omega) + (1-\omega)^2 + \beta^2 - \omega^2(\alpha^2 + \beta^2) \\
&= (1 - \omega^2)(\alpha^2 + \beta^2) - 2\alpha(1-\omega) + (1-\omega)^2 \\
&= (1-\omega)\left[(\alpha - 1)^2 + \beta^2 + \omega(\alpha^2 + \beta^2 - 1)\right] \\
&\ge 0
\end{aligned}$$

从而
$$|\lambda+\omega-1|^2 \geq |\lambda|^2\omega^2 \geq \omega^2$$
即
$$|\lambda+\omega-1| \geq |\lambda|\omega \geq \omega$$
于是由 A 的严格对角占优或不可约对角占优性质可知 $(\lambda+\omega-1)D-\lambda\omega L-\omega U$ 也是严格对角占优或不可约对角占优的. 则 $(\lambda+\omega-1)D-\lambda\omega L-\omega U$ 也是非奇异的. 而
$$\begin{aligned}\det(\lambda I-L_\omega) &= \det\left(\lambda I-(D-\omega L)^{-1}[(1-\omega)D+\omega U]\right)\\ &= \det\left((D-\omega L)^{-1}\right)\det\left((\lambda+\omega-1)D-\lambda\omega L-\omega U\right)\\ &\neq 0\end{aligned}$$
矛盾, 因此 λ 不是 SOR 迭代矩阵 L_ω 的特征值. 由 λ 的任意性可知, L_ω 的特征值都满足 $|\lambda|<1$, 从而 SOR 迭代收敛.

另解: 下仅证
$$|\lambda+\omega-1| \geq |\lambda|\omega \geq \omega$$
令 $z=\frac{1}{\lambda}$, 则 $|z|\leq 1$. 考虑函数
$$f(z)=\frac{\omega}{1-z(1-\omega)}$$
由于 $0<1-\omega<1$, 当 $|z|\leq 1$ 时, $|z(1-\omega)|<1$, 故 $f(z)$ 在闭单位圆盘 $\overline{D(0,1)}$ 上解析.
当 $|z|=1$ 时, 有
$$|1-z(1-\omega)| \geq |1-(1-\omega)| = \omega$$
因此
$$|f(z)| = \left|\frac{\omega}{1-z(1-\omega)}\right| \leq 1$$
由最大模原理, 在 $|z|\leq 1$ 上恒有 $|f(z)|\leq 1$. 特别地, 当 $0<|z|\leq 1$ 时,
$$\left|\frac{\omega}{1-z(1-\omega)}\right| \leq 1 \quad \Rightarrow \quad |1-z(1-\omega)| \geq \omega$$
代入 $z=\frac{1}{\lambda}$, 得
$$\left|1-\frac{1}{\lambda}(1-\omega)\right| \geq \omega \quad \Rightarrow \quad |\lambda-(1-\omega)| \geq |\lambda|\omega$$
证毕.

14. 证明: 若 A 为具有正对角元的实对称矩阵, 则 JOR 方法收敛的充分必要条件是 A 与 $2\omega^{-1}D-A$ 均为正定对称矩阵. 其中, JOR 迭代为
$$x_{k+1}=x_k-\omega D^{-1}(Ax_k-b)$$
或者
$$x_{k+1}=(I-\omega D^{-1}A)x_k+\omega D^{-1}b$$

Solution: $D = \mathrm{diag}(a_{11}, \ldots, a_{nn})$ 的对角元均为正数, 所以记

$$B_\omega = I - \omega D^{-1}A = D^{-\frac{1}{2}}(I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}})D^{\frac{1}{2}}.$$

即 B_ω 和 $I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 相似, 因此它们有相同的特征值. 因为 A 是实对称矩阵, 所以 $I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 也是实对称矩阵, 则有 B_ω 和 $I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的特征值相同且都为实数.

1. **必要性:** 设 JOR 方法收敛, 即 $\rho(B_\omega) < 1$, 所以 $I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 的任一特征值 α 都满足 $-1 < \alpha < 1$. 于是考虑矩阵

$$I - (I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}) = \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$$

的任一特征值满足 $1 - \alpha > 0$, 所以 $\omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 正定, 又 $\omega > 0$ (一般如此规定), 故矩阵 A 正定. 再考虑

$$I + (I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}) = 2I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}} = \omega D^{-\frac{1}{2}}(2\omega^{-1}D - A)D^{-\frac{1}{2}}$$

的任一特征值满足 $1 + \alpha > 0$, 所以 $2I - \omega D^{-\frac{1}{2}}AD^{-\frac{1}{2}}$ 正定, 同样有矩阵 $2\omega^{-1}D - A$ 也正定.

2. **充分性:** 因为 A 正定, 且

$$A = \omega^{-1}D(I - B_\omega) = \omega^{-1}D^{\frac{1}{2}}D^{\frac{1}{2}}(I - B_\omega)D^{-\frac{1}{2}}D^{\frac{1}{2}}$$

所以 $I - B_\omega$ 和 A 一样是正定矩阵, 任一特征值都大于 0, 故 B_ω 的特征值均小于 1. 又因为 $2\omega^{-1}D - A$ 正定, 且

$$2\omega^{-1}D - A = \omega^{-1}D(I + B_\omega) = \omega^{-1}D^{\frac{1}{2}}D^{\frac{1}{2}}(I + B_\omega)D^{-\frac{1}{2}}D^{\frac{1}{2}}$$

所以 $I + B_\omega$ 正定, 任一特征值都大于 0, 故 B_ω 的特征值均大于 -1. 因此 $\rho(B_\omega) < 1$, 即有 JOR 方法收敛.

另解: 可以参考题 8 直接证明等价性.

15. 设 x_k 由最速下降法产生, 证明:

$$\phi(x_k) \le \left[1 - \frac{1}{\kappa_2(A)}\right]\phi(x_{k-1})$$

其中 $\kappa_2(A) = \|A\|_2\|A^{-1}\|_2$.

Solution: 有 A 对称正定且 $\kappa_2(A) = \|A\|_2\|A^{-1}\|_2 > 0$. 又最速下降法的递推公式:

$$\phi(x_k) = \phi(x_{k-1}) - \frac{(r_{k-1}^{\mathrm{T}}p_{k-1})^2}{p_{k-1}^{\mathrm{T}}Ap_{k-1}} = \phi(x_{k-1}) - \frac{(r_{k-1}^{\mathrm{T}}r_{k-1})^2}{r_{k-1}^{\mathrm{T}}Ar_{k-1}}$$

1. 当 $\phi(x_{k-1}) \le 0$ 时, 由最速下降法的性质知 $\phi(x_k) < \phi(x_{k-1}) \le 0$, 于是

   $$\phi(x_k) < \phi(x_{k-1}) \le \left[1 - \frac{1}{\kappa_2(A)}\right]\phi(x_{k-1})$$

   不等式成立.

2. 当 $\phi(x_{k-1}) > 0$ 时, 由于 A^{-1} 正定, 注意到有

   $$r_{k-1}^{\mathrm{T}}A^{-1}r_{k-1} = (b - Ax_{k-1})^{\mathrm{T}}A^{-1}(b - Ax_{k-1}) = \phi(x_{k-1}) + b^{\mathrm{T}}A^{-1}b \ge \phi(x_{k-1}) > 0$$

   代入递推公式, 有

   $$\begin{aligned}\phi(x_k) &= \phi(x_{k-1})\left[1 - \frac{(r_{k-1}^{\mathrm{T}}r_{k-1})^2}{\phi(x_{k-1})\cdot(r_{k-1}^{\mathrm{T}}Ar_{k-1})}\right] \\ &\le \phi(x_{k-1})\left[1 - \frac{(r_{k-1}^{\mathrm{T}}r_{k-1})\cdot(r_{k-1}^{\mathrm{T}}r_{k-1})}{(r_{k-1}^{\mathrm{T}}A^{-1}r_{k-1})\cdot(r_{k-1}^{\mathrm{T}}Ar_{k-1})}\right]\end{aligned}$$

   又由 Cauchy-Schwarz 不等式, 有

   $$r_{k-1}^{\mathrm{T}}Ar_{k-1} \le \|r_{k-1}\|_2\|Ar_{k-1}\|_2 \le \|A\|_2\|r_{k-1}\|_2^2 = \|A\|_2 \cdot r_{k-1}^{\mathrm{T}}r_{k-1}$$

   且

   $$r_{k-1}^{\mathrm{T}}A^{-1}r_{k-1} \le \|r_{k-1}\|_2\|A^{-1}r_{k-1}\|_2 \le \|A^{-1}\|_2\|r_{k-1}\|_2^2 = \|A^{-1}\|_2 \cdot r_{k-1}^{\mathrm{T}}r_{k-1}$$

   代入得

   $$\phi(x_k) \le \phi(x_{k-1})\left[1 - \frac{1}{\|A\|_2\|A^{-1}\|_2}\right] = \phi(x_{k-1})\left[1 - \frac{1}{\kappa_2(A)}\right]$$

综上, 不等式 $\phi(x_k) \le \left[1 - \frac{1}{\kappa_2(A)}\right]\phi(x_{k-1})$ 得证.

16. 证明: 当最速下降法在有限步求得极小值时, 最后一步迭代的下降方向必是 A 的一个特征向量.

Solution: 假定在 $k+1$ 步迭代后, 得到了精确解 $x_{k+1} = x_*$, 即

$$x_* = x_k + \frac{r_k^{\mathrm{T}}r_k}{r_k^{\mathrm{T}}Ar_k}r_k$$

从而有

$$b = Ax_* = Ax_k + \frac{r_k^{\mathrm{T}}r_k}{r_k^{\mathrm{T}}Ar_k}Ar_k$$

记:

$$\lambda = \frac{r_k^{\mathrm{T}}Ar_k}{r_k^{\mathrm{T}}r_k}$$

又 $r_k = b - Ax_k$, 整理可得

$$Ar_k = \lambda r_k$$

17. $A \in \mathbb{R}^{n\times n}$ 对称正定, $p_1, \ldots, p_k \in \mathbb{R}^n$ 满足 $p_i^{\mathrm{T}} A p_j = 0$, $i \neq j$. 证明 $p_1, \ldots, p_k$ 线性无关.

Solution: 设有 $\alpha_1, \ldots, \alpha_k$ 满足

$$\alpha_1 p_1 + \cdots + \alpha_k p_k = 0$$

则对一切 $i = 1, \ldots, k$, 上式左乘 $p_i^{\mathrm{T}} A$ 有

$$0 = p_i^{\mathrm{T}} A(\alpha_1 p_1 + \cdots + \alpha_k p_k) = \alpha_i p_i^{\mathrm{T}} A p_i$$

由于 $p_i \neq 0$, 有 $p_i^{\mathrm{T}} A p_i \neq 0$, 由此得出 $\alpha_i = 0, \forall i = 1, \ldots, k$. 因此, $p_1, p_2, \ldots, p_k$ 线性无关.

18. 设 A 是一个只有 k 个互不相同特征值的 $n \times n$ 实对称矩阵, r 是任一 n 维实向量. 证明: 子空间 $\mathrm{span}\{r, Ar, \ldots, A^{n-1}r\}$ 的维数至多是 k.

Solution: 设 A 的 n 个线性无关的特征向量为 $x_1, x_2, \ldots, x_n$, 对应的特征值为 $\lambda_1, \lambda_2, \ldots, \lambda_n$. 又设 $r = k_1 x_1 + k_2 x_2 + \cdots + k_n x_n$. 则

$$A^l r = k_1 \lambda_1^l x_1 + k_2 \lambda_2^l x_2 + \cdots + k_n \lambda_n^l x_n \quad (l = 0, 1, 2, \ldots)$$

考察矩阵 $[r, Ar, \ldots, A^{n-1}r] = [k_1 x_1, k_2 x_2, \ldots, k_n x_n]B$, 其中 B 恰是范德蒙德矩阵:

$$B = \begin{bmatrix} 1 & \lambda_1 & \cdots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \cdots & \lambda_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \cdots & \lambda_n^{n-1} \end{bmatrix}$$

由于 $\lambda_1, \lambda_2, \ldots, \lambda_n$ 中只有 k 个不同的值, 于是 $\mathrm{rank}(B) = k$, 从而 $\mathrm{rank}[r, Ar, \ldots, A^{n-1}r] \leq \mathrm{rank}(B) = k$.

另解: 由于 A 是实对称矩阵, 因此 A 可对角化, 其对应最小多项式无重根, 即有

$$m(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_k)$$

其中 $\lambda_1, \lambda_2, \ldots, \lambda_k$ 是 A 的特征值. 因此有

$$A^k + c_{k-1}A^{k-1} + \cdots + c_1 A + c_0 I = 0$$

其中 $c_{k-1}, \ldots, c_0$ 为常数.
于是 A^k 可以被 $r, Ar, \ldots, A^{k-1}r$ 线性表出, 则子空间

$$\mathrm{span}\{r, Ar, \ldots, A^{k-1}r, A^k r\} = \mathrm{span}\{r, Ar, \ldots, A^{k-1}r\} \triangleq S$$

类似地, 我们可以递推（在 A 满足的等式两边同时左乘或者右乘矩阵 A）得到

$$\mathrm{span}\{r, Ar, \ldots, A^{n-1}r\} = \mathrm{span}\{r, Ar, \ldots, A^{k-1}r\},\ \forall n \geq k$$

又由于 S 的维数至多为 k, 故 $\mathrm{span}\{r, Ar, \ldots, A^{n-1}r\}$ 的维数至多为 k.

19. 证明: 如果系数矩阵 A 至多有 l 个互不相同的特征值, 则共轭梯度法至多 l 步就可得到方程组 $Ax = b$ 的精确解.

Solution: 由题 18 可知 Krylov 子空间 $\mathrm{span}\{r, Ar, \ldots, A^{n-1}r\}$ 的维数最多是 l , 所以共轭梯度法至多 l 步便可得到方程组 $Ax = b$ 的精确解.

20. 证明: 用共轭梯度法求得的 x_k 有如下的误差估计:

$$\|x_k - x_*\|_2 \leq 2\sqrt{\kappa_2}\left(\frac{\sqrt{\kappa_2}-1}{\sqrt{\kappa_2}+1}\right)^k \|x_0 - x_*\|_2$$

其中 $\kappa_2 = \kappa_2(A) = \|A\|_2\|A^{-1}\|_2$.

Solution: 由课本定理可知

$$\|x_k - x_*\|_A \leq 2\left(\frac{\sqrt{\kappa_2}-1}{\sqrt{\kappa_2}+1}\right)^k \|x_0 - x_*\|_A$$

故只需证明

$$\frac{\|x_k - x_*\|_2}{\|x_0 - x_*\|_2} \leq \sqrt{\kappa_2}\frac{\|x_k - x_*\|_A}{\|x_0 - x_*\|_A}$$

记

$$\alpha = x_k - x_*, \quad \beta = x_0 - x_*$$

则只需证

$$(\alpha^{\mathrm{T}}\alpha)(\beta^{\mathrm{T}}A\beta) \leq \|A\|_2\|A^{-1}\|_2(\beta^{\mathrm{T}}\beta)(\alpha^{\mathrm{T}}A\alpha)$$

事实上, 对于任一正定矩阵 M 都有

$$\|M\|_2^2 = (\sqrt{\|M^{\mathrm{T}}M\|_2})^2 = \|M^2\|_2$$

由 A 的正定性可知 $A^{1/2}$ 和 $A^{-1/2}$ 也是正定的（正定矩阵的平方根分解）, 于是

$$\alpha^{\mathrm{T}}\alpha = \|\alpha\|_2^2 = \|A^{1/2}A^{-1/2}\alpha\|_2^2 \leq \|A^{-1/2}\|_2^2\|A^{1/2}\alpha\|_2^2 = \|A^{-1}\|_2(\alpha^{\mathrm{T}}A\alpha)$$

同样有

$$\beta^{\mathrm{T}}A\beta = \|A^{1/2}\beta\|_2^2 \leq \|A^{1/2}\|_2^2\|\beta\|_2^2 = \|A\|_2(\beta^{\mathrm{T}}\beta)$$

上两式相乘即为所需, 证毕.

21. 设 $A \in \mathbb{R}^{n\times n}$ 是对称正定的, $\mathcal{X}$ 是 $\mathbb{R}^n$ 的一个 k 维子空间. 证明: $x_k \in \mathcal{X}$ 满足

$$\|x_k - A^{-1}b\|_A = \min_{x\in\mathcal{X}} \|x - A^{-1}b\|_A$$

的充分必要条件是

$$r_k = b - Ax_k$$

垂直于子空间 $\mathcal{X}$, 其中 $b \in \mathbb{R}^n$ 是任意给定的.

Solution: 直接从充分性和必要性两方面分别证即可.

充分性: 当 $r_k \perp \mathcal{X}$, 即

$$(b - Ax_k)^{\mathrm{T}}x = 0 \Leftrightarrow b^{\mathrm{T}}x = x_k^{\mathrm{T}}Ax, \forall x \in \mathcal{X}$$

和

$$(b - Ax_k)^{\mathrm{T}}x_k = 0 \Leftrightarrow b^{\mathrm{T}}x_k = x_k^{\mathrm{T}}Ax_k$$

时, 要证明

$$\begin{aligned}\|x_k - A^{-1}b\|_A = \min_{x\in\mathcal{X}} \|x - A^{-1}b\|_A &\Leftrightarrow \|x_k - A^{-1}b\|_A \le \|x - A^{-1}b\|_A \\ &\Leftrightarrow x_k^{\mathrm{T}}Ax_k - 2b^{\mathrm{T}}x_k \le x^{\mathrm{T}}Ax - 2b^{\mathrm{T}}x\end{aligned}$$

只需证 $\forall x \in \mathcal{X}$, 恒成立

$$x^{\mathrm{T}}Ax - 2x_k^{\mathrm{T}}Ax + x_k^{\mathrm{T}}Ax_k \ge 0$$

然而显然有

$$x^{\mathrm{T}}Ax - 2x_k^{\mathrm{T}}Ax + x_k^{\mathrm{T}}Ax_k = \|x - x_k\|_A^2 \ge 0$$

则充分性得证.

必要性: 当

$$\|x_k - A^{-1}b\|_A = \min_{x\in\mathcal{X}} \|x - A^{-1}b\|_A$$

时, $\forall x \in \mathcal{X}$, $\forall \lambda \in \mathbb{R}$ 有

$$\|x_k - A^{-1}b\|_A \le \|(x_k + \lambda x) - A^{-1}b\|_A$$

即

$$x_k^{\mathrm{T}}Ax_k - 2b^{\mathrm{T}}x_k \le (x_k + \lambda x)^{\mathrm{T}}A(x_k + \lambda x) - 2b^{\mathrm{T}}(x_k + \lambda x)$$

于是

$$(x^{\mathrm{T}}Ax)\lambda^2 + 2(x^{\mathrm{T}}Ax_k - b^{\mathrm{T}}x)\lambda \ge 0$$

考虑其对应一元二次方程的判别式

$$\Delta = 4(x^{\mathrm{T}}Ax_k - b^{\mathrm{T}}x)^2 - 4\cdot(x^{\mathrm{T}}Ax)\cdot 0 \le 0$$

即

$$x^{\mathrm{T}}(Ax_k - b) = 0, \forall x \in \mathcal{X} \Leftrightarrow r_k \perp \mathcal{X}$$

必要性得证.

另解: 必要性同上, 下证充分性: 注意到 $\forall x \in \mathcal{X}$, 有

$$\begin{aligned}
\|x - A^{-1}b\|_A^2 &= \|x_k - A^{-1}b + x - x_k\|_A^2 \\
&= \|x_k - A^{-1}b\|_A^2 + \|x - x_k\|_A^2 + 2\langle x_k - A^{-1}b, x - x_k\rangle_A \\
&= \|x_k - A^{-1}b\|_A^2 + \|x - x_k\|_A^2 + 2\langle Ax_k - b, x - x_k\rangle \\
&\geq \|x_k - A^{-1}b\|_A^2
\end{aligned}$$

即原命题得证.

再另解: 可以考虑内积空间中的投影定理.