

矩阵分析 & 最优化方法 考题整理

目录

第一章 矩阵分析	2
I.1 今年题	3
I.1.1 填空题	3
I.1.2 大题	3
I.2 往年题	6
第二章 最优化方法	12
II.1 今年题	13
II.1.1 陈述题	13
II.1.2 大题	13
II.2 往年题 (2011)	17
II.2.1 简答题	17
II.2.2 大题	19

CHAPTER I

矩阵分析



SECTION I.1

今年题

考试分为填空和大题.

I.1.1 填空题

填空主要是矩阵的极限、级数、微分相关的题型. 不管是矩阵的级数还是极限, 一个明显的套路就是将矩阵对角化或者化成 Jordan 标准型, 再进行求解.

题 1

设矩阵 $A = \begin{bmatrix} 0 & a & a \\ a & 0 & a \\ a & a & 0 \end{bmatrix}$, 试求 $\lim_{m \rightarrow \infty} A^m = O$ 时 a 的取值范围为_____.

解. A 有特征值 $\lambda_1 = 2a$ (1 重), $\lambda_2 = -a$ (2 重), 且 A 可对角化, 存在可逆矩阵 P , 使得 $P^{-1}AP = D = \text{diag}\{2a, -a, -a\}$, 如果 $\lim_{m \rightarrow \infty} A^m = O$, 那么 $|2a| < 1$ 且 $|a| < 1$, a 取值范围 $(-\frac{1}{2}, \frac{1}{2})$. 解毕.

题 2

矩阵 A 满足 $\|A\| < 1$, 则 $\lim_{m \rightarrow \infty} A^m =$ _____.

解. 因为 $\|A^m\| \leq \|A\|^m \rightarrow 0$ 当 $m \rightarrow \infty$, 则所求极限为 O . 解毕.

I.1.2 大题

大题的第一题是关于盖氏圆盘定理的应用和推论. 这个题大致是先用 Geršgorin 圆盘定理判断特征值分布在哪些圆盘中, 这就需要用到定理的推论.

题 2

矩阵 $A \in \mathbb{C}^{m \times n}$, 证明:

$$\max_{i,j} |a_{ij}| \leq \|A\|_2 \leq \sqrt{mn} \max_{i,j} |a_{ij}|.$$



证明. 首先证明第一个不等式. 因为矩阵的 2-范数与向量的 2-范数是相容的, 取 $\varepsilon_j = (0, \dots, 1, \dots, 0)^T$ 为单位阵 I 的第 j 列. 又

$$\|A\|_2 = \|A\|_2 \|\varepsilon_j\|_2 \geq \|A\varepsilon_j\|_2,$$

$A\varepsilon_j = (a_{1j}, a_{2j}, \dots, a_{mj})^T$. 则

$$\|A\|_2 \geq \sqrt{\sum_{i=1}^m a_{ij}^2} \geq \max_i |a_{ij}|.$$

由 j 的任意性,

$$\|A\|_2 \geq \max_{i,j} |a_{ij}|.$$

第二个不等式. 记 $I_{ij}^{(k)}$ 为除了 (i, j) 元为 1 之外, 其余元素均为 0 的 k 阶方阵, 则

$$A = \sum_{i,j} I_{ii}^{(m)} A I_{jj}^{(n)},$$

其中 $I_{ii}^{(m)} A I_{jj}^{(n)}$ 是除了 (i, j) 元为 a_{ij} 之外其余元素为 0 的 $m \times n$ 矩阵, 则

$$\begin{aligned} \|A\|_2 &\leq \sum_{i,j} \left\| I_{ii}^{(m)} A I_{jj}^{(n)} \right\|_2 \\ &\leq \sum_{k=1}^{\min(n,m)} |a_{ij}| \\ &\leq \min(m, n) \max_{i,j} |a_{ij}| \\ &\leq \sqrt{mn} \max_{i,j} |a_{ij}|. \end{aligned}$$

证毕. ■

大题的第三个题是关于满秩分解的, 纯粹是硬算.

题 4

矩阵 $X, A \in \mathbb{C}^{n \times n}$, 证明

$$\frac{d}{dX} \text{tr}(X^T A X) = (A + A^T)X.$$



证明. 设 $X = [x_1, x_2, \dots, x_n] = (x_{ij})_{n \times n}$, $A = (a_{ij})_{n \times n}$, 那么 $\text{tr}(X^T A X) = \sum_{k=1}^n x_k^T A x_k$. 又 $\frac{d}{dX} x_k^T A x_k = \left(\frac{d}{dx_{ij}} x_k^T A x_k \right)_{n \times n}$, 当 $j \neq k$, $\frac{d}{dx_{ij}} x_k^T A x_k = 0$, 当 $j = k$ 时, $\frac{d}{dx_{ik}} x_k^T A x_k = \varepsilon_i^T (A + A^T) x_k$, 则

$$\frac{d}{dX} x_k^T A x_k = \begin{bmatrix} 0 & \cdots & \varepsilon_1^T (A + A^T) x_k & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & \varepsilon_n^T (A + A^T) x_k & \cdots & 0 \end{bmatrix}.$$

得到

$$\begin{aligned} \frac{d}{dX} X^T A X &= (\varepsilon_i^T (A + A^T) x_j)_{n \times n} \\ &= I(A + A^T)X = (A + A^T)X. \end{aligned}$$

也就是要证的等式. ■

题 5

矩阵 $A, B \in \mathbb{C}^{m \times n}$ 且 $AB^H = B^H A = O$, 证明

$$(A + B)^+ = A^+ + B^+.$$

证明. 首先证明 $(A + B)(A^+ + B^+)(A + B) = A + B$.

$$\begin{aligned} (A + B)(A^+ + B^+)(A + B) &= AA^+A + AA^+B + AB^+A + AB^+B \\ &\quad + BA^+A + BA^+B + BB^+A + BB^+B \\ &= A + B + (AA^+B + AB^+A + AB^+B \\ &\quad + BA^+A + BA^+B + BB^+A). \end{aligned}$$

又因为

$$\begin{aligned} AA^+B &= (AA^+)^H B = (A^+)^H A^H B = O, \\ AB^+A &= AB^+BB^+A = AB^H(B^+)^H B^+A = O, \end{aligned}$$

同理得到 $AB^+B = BA^+A = BA^+B = BB^+A = O$, 那么

$$(A + B)(A^+ + B^+)(A + B) = A + B.$$



同理, 可得 $(A^+ + B^+)(A + B)(A^+ + B^+) = A^+ + B^+$.

下面证明 $[(A + B)(A^+ + B^+)]^H = (A + B)(A^+ + B^+)$. 因为

$$\begin{aligned} [(A + B)(A^+ + B^+)]^H &= (AA^+ + AB^+ + BA^+ + BB^+)^H \\ &= AA^+ + BB^+ + (AB^+ + BA^+)^H. \end{aligned}$$

又 $AB^+ = AB^+BB^+ = AB^H(B^+)^HB^H = O$, 同理 $BA^+ = O$, 那么 $(A + B)(A^+ + B^+) = AA^+ + BB^+$, 则

$$[(A + B)(A^+ + B^+)]^H = AA^+ + BB^+ = (A + B)(A^+ + B^+).$$

同理, $[(A^+ + B^+)(A + B)]^H = (A^+ + B^+)(A + B)$.

则 $(A + B)^+ = A^+ + B^+$. ■

大题第 6 题大致是: 对于线性方程组 $Ax = b$, 需要

1. 求解 A^+ .
2. 计算 $\beta = Ay_0$, 其中 y_0 是最小二乘解 (其实 y_0 就是 A^+b).

也是硬算题.

SECTION 1.2

往年题

题 1

求矩阵 $A = \begin{bmatrix} 2 & -1 & -1 \\ 2 & -1 & -2 \\ -1 & -1 & 2 \end{bmatrix}$ 的行列式因子、不变因子和初等因子.

解. 矩阵 $\lambda I - A$ 的 Smith 标准型为

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & (\lambda - 1)(\lambda + 1)(\lambda - 3) \end{bmatrix}$$



行列式因子: $D_1 = 1, D_2 = 1, D_3 = (\lambda - 1)(\lambda + 1)(\lambda - 3)$.

不变因子: $d_1 = 1, d_2 = 1, d_3 = (\lambda - 1)(\lambda + 1)(\lambda - 3)$.

初等因子: $\lambda - 1, \lambda + 1, \lambda + 3$.

解毕.

题 2

设 A 是 n 阶常数对称矩阵. 向量函数 $y = (y_1(t), y_2(t), \dots, y_n(t))^T$, 二次型 $f(y) = y^T A y$, 证明: $\frac{df}{dt} = 2y^T A \frac{dy}{dt}$.

证明. 注意到 A 对称,

$$\begin{aligned} \frac{df}{dt} &= \frac{dy^T}{dt} A y + y^T A \frac{dy}{dt} \\ &= 2y^T A \frac{dy}{dt}. \end{aligned}$$

即证. ■

题 3

求 $A = \begin{bmatrix} -1 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$ 的奇异值分解.

解. 求 $A^T A$. $A^T A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ -2 & 0 & 4 \end{bmatrix}$.

求 $A^T A$ 的特征值和特征向量. $A^T A$ 的特征值为 $\lambda_1 = 5, \lambda_2 = 1, \lambda_3 = 0$, 对应的特

征向量按照矩阵写出是 $V = \begin{bmatrix} -\frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ 0 & 1 & 0 \\ \frac{2}{\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{bmatrix}$.

利用 $A = U \begin{bmatrix} S & O \\ O & O \end{bmatrix} V^H$ 求解 U (或 V). 已知 $S = \text{diag}\{\sqrt{5}, 1\}$, 求解得 $U = I$.
解毕.



题 4

设 $m \times n$ 阶矩阵 A , $\text{rank}(A) = r$, 证明: $A = CD$ 存在, 其中 C 是 $m \times r$ 阶列满秩阵, D 是 $r \times n$ 阶行满秩阵.

证明. 存在可逆矩阵 P, Q , 使得 $PAQ = B = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$, 则

$$\begin{aligned} A &= P^{-1}BQ^{-1} \\ &= P^{-1} \begin{bmatrix} I_r \\ O \end{bmatrix} \begin{bmatrix} I_r & O \end{bmatrix} Q^{-1}, \end{aligned}$$

令 $C = P^{-1} \begin{bmatrix} I_r \\ O \end{bmatrix}$, $D = \begin{bmatrix} I_r & O \end{bmatrix} Q^{-1}$, 它们分别是列满秩和行满秩矩阵. ■

题 5

设矩阵 $A = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, 求 e^{tA} .

解. 令 $P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}$, $J = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$, 则

$$P^{-1}AP = J.$$

计算 e^{tA} :

$$\begin{aligned} e^{tA} &= e^{tPJP^{-1}} = Pe^{tJ}P^{-1} \\ &= P \exp(tI + tB)P^{-1} \\ &= P \exp(tI) \exp(tB)P^{-1} \\ &= e^t P \exp(tB)P^{-1} \\ &= e^t P \left(I + tB + \frac{t^2 B^2}{2} \right) P^{-1} \end{aligned}$$



$$= e^t \begin{bmatrix} 1 & t & t^2/2 + t \\ 0 & 1 & t \\ 0 & 0 & 1 \end{bmatrix}.$$

解毕.

题 6

设矩阵 $H \in \mathbb{C}^{n \times n}$, 证明其广义逆 $H^+ = H$ 当且仅当 H^2 是幂等 Hermite 矩阵且 $\text{rank}(H^2) = \text{rank}(H)$.

证明. “ \Rightarrow ”. 这时 $H^+ = H, H^2 = H^+H = (H^+H)^H = (H^2)^H$, 则 H^2 是 Hermite 阵; $(H^2)^2 = (H^+H)^2 = H^H H^+ H = H^+ H = H^2$, 则 H^2 幂等; $H = HH^+H = H^3, \text{rank}(H) \leq \text{rank}(H^2)$, 则 $\text{rank}(H) = \text{rank}(H^2)$.

“ \Leftarrow ”. 因为 H^2 是 Hermite 阵, 那么 $(HH)^H = HH$. 又 $\text{rank}(H^2) = \text{rank}(H)$, 则 $R(H^2) = R(H)$, 即存在 A , 使得 $H^2A = H$,

$$HHH = HHA = H.$$

则 $H = H^+$. ■

题 7

矩阵 $A \in \mathbb{C}^{m \times n}$, 若存在 $G \in \mathbb{C}^{n \times m}, S \in \mathbb{C}^{m \times m}$ 和 $T \in \mathbb{C}^{n \times n}$ 满足

$$AGA = A, \tag{I.1}$$

$$G = A^H S, \tag{I.2}$$

$$G = TA^H, \tag{I.3}$$

则 G 唯一确定且 $G = A^+$.

证明. 根据式 (I.2) 和 (I.3),

$$A^H G^H G = A^H A T^H A^H S = (A T^H A)^H S = (AGA)^H S = A^H S = G.$$

类似, $GG^H A^H = G$, 也就是如果有 (I.1)、(I.2) 和 (I.3) 成立, 一定就有 $A^H G^H G =$



$G, GG^H A^H = G$ 成立. 不妨令 $S = G^H G, T = GG^H$, 则 $S^H = S, T^H = T$,

$$(GA)^H = (A^H SA)^H = A^H S^H A = A^H SA = GA,$$

$$(AG)^H = (ATA^H)^H = AT^H A^H = ATA^H = AG,$$

$$GAG = (GA)^H G = A^H G^H G = A^H S = G.$$

则 $G = A^+$. 从上面求解 G 的过程来看, G 唯一确定. ■

题 8

设 $A = \begin{bmatrix} 1 & 2 & 2 & 1 \\ 1 & 1 & 1 & 1 \\ 2 & 1 & 1 & 2 \end{bmatrix}, b = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$

求:

1. A^+ ;
2. 用广义逆矩阵性质判别 $Ax = b$ 是否相容;
3. 若相容, 求出最小范数解; 若不相容, 求出最小范数二乘解.

解. 1. 将 A 奇异值分解, 设 $A = U \begin{bmatrix} S & O_1 \\ O_2 & O_3 \end{bmatrix} V^H$. 在形式上,

$$\begin{aligned} A^+ &= V \begin{bmatrix} S^{-1} & O_2^T \\ O_1^T & O_3^T \end{bmatrix} U^H \\ &= \begin{bmatrix} -0.1818 & 0.0455 & 0.3182 \\ 0.3182 & 0.0455 & -0.1818 \\ 0.3182 & 0.0455 & -0.1818 \\ -0.1818 & 0.0455 & 0.3182 \end{bmatrix}. \end{aligned}$$

(或者用满秩分解 $A = FG, A^+ = G^H(F^H FGG^H)^{-1}F^H$.)



2. A^+ 当然是 $A^{(1)}$, 计算 AA^+b , 得到

$$AA^+b = \begin{bmatrix} 1.0909 \\ 0.7273 \\ 1.0909 \end{bmatrix} \neq b.$$

方程组不相容.(或者说明 $\text{rank}([A \ b]) \neq \text{rank}(A)$.)

3. 最小范数二乘解为

$$x_0 = A^+b = \begin{bmatrix} 0.1818 \\ 0.1818 \\ 0.1818 \\ 0.1818 \end{bmatrix}.$$

解毕.

CHAPTER II

最优化方法



SECTION II.1

今年题

分为两部分: 陈述题和解答题 (解答题最后两道二选一).

II.1.1 陈述题

陈述题既包括定义比如凸集和可行域的概念, 也包括其他的一些基本的东西, 比如说是信赖域方法简述和信赖域半径确定, 最速下降法算法步骤, 带不等式约束的外罚函数等.

II.1.2 大题

大题差不多就是老师强调的重点.

题 1

定义在 \mathbb{R}^3 上的函数 $f(x) = \frac{1}{3}x_1^3 + \frac{1}{3}x_2^3 - 2x_2 - x_1$, 求 f 的平稳点, 在平稳点中哪些是最优解?

思路

若 x 是 f 的平稳点, 那么有 $\nabla f(x) = 0$, 这样求解可以得到平稳点.

要想找到最优解, 首先需要利用最优性条件的二阶必要条件判断那些是局部最小点, 然后在这些局部最小点中寻找.

对于本题来说, 因为是无约束优化问题, 而且 f 是三次的, 所以没有最小点.

题 2

考虑以下优化问题:

$$\begin{aligned} \min \quad & (x_1 - 3)^2 + (x_2 - 1)^2 \\ \text{s.t.} \quad & \begin{cases} -x_1^2 + x_2 \geq 0, \\ 2x_1 + x_2 - 3 \geq 0. \end{cases} \end{aligned}$$

求出 KKT 点.



思路

首先写出这个问题的 KKT 条件, 然后分类讨论, 排除到最后只有一种情况, 是一个不好求解的三次方程.

题 3

考虑无约束优化问题:

$$\min f(x) = (-x_1 + x_2 + x_3)^2 + (x - 1 - x_2 + x_3)^2 + (x_1 + x_2 - x_3)^2.$$

初始点 $x_0 = (\frac{1}{2}, 1, \frac{1}{2})^T$, 用共轭梯度法 (HS/CW/FR/PRP/Dixon/DY) 进行两步迭代.

思路 – 以 FR 为例

算法不唯一, 其中一个:

算法. 令 $f(x) = \frac{1}{2}x^T Gx - bx$, 其中 $G = \nabla^2 f(x)$.

步 1. 初始步: 给出 x_0 , 计算 $r_0 = g_0 = \nabla f(x_0)$, 令 $d_0 = -r_0, k := 0$.

步 2. 如果 $k > 2$, 停止.

步 3. 依次计算

$$\alpha_k = \frac{r_k^T r_k}{d_k^T G d_k},$$

$$x_{k+1} = x_k + \alpha_k d_k,$$

$$r_{k+1} = r_k + \alpha_k G d_k,$$

$$\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k},$$

$$d_{k+1} = -r_{k+1} + \beta_k d_k.$$

步 4. 令 $k := k + 1$, 转步 2.

没记错的话答案是 $(0, 0, 0)^T$.



题 4

有两个优化问题:

$$\begin{aligned} \text{(I)} \quad & \min f(x) \\ & \text{s.t. } a_i^T x - b_i \geq 0 \quad i \in \mathcal{I}, \\ & \quad \quad a_i^T x - b_i = 0 \quad i \in \mathcal{E}. \\ \text{(II)} \quad & \min \nabla f(x)^T d \\ & \text{s.t. } a_i^T d \geq 0 \quad i \in \mathcal{I}, \\ & \quad \quad a_i^T d = 0 \quad i \in \mathcal{E}. \end{aligned}$$

设 d 是 (II) 的解, 证明对 (I) 的任意可行点 x , d 是 x 的可行方向.

证明. 对任意的 $t > 0$, x 是 (I) 可行域的点, d 是 (II) 的解, 有

$$\begin{aligned} a_i^T(x + td) - b_i &\geq a_i^T x - b_i + ta_i^T d \geq 0, \quad i \in \mathcal{I}, \\ a_i^T(x + td) - b_i &= a_i^T x - b_i + ta_i^T d = 0, \quad i \in \mathcal{E}. \end{aligned}$$

则 $x + td$ 是可行域中的点.



其实要是说明 d 是可行方向, 还需要说明 $d \neq 0$, 但我不知道是题目不严谨还是确实有解 $d \neq 0$ 恒成立, 可以再验证一下.



题 5

d 是 $f(x)$ 的一个下降方向, 令

$$H = I - \frac{dd^T}{d^T \nabla f(x)} - \frac{\nabla f(x) \nabla f(x)^T}{\nabla f(x)^T \nabla f(x)}.$$

再令 $p = -H \nabla f(x)$, 证明 p 是 f 在 x 处的一个下降方向.

证明. 定义验证 d 是 f 的一个下降方向, 则 $\nabla f(x)^T d < 0$.

$$\begin{aligned} \nabla f(x)^T p &= -\nabla f(x)^T H \nabla f(x) \\ &= -\left(\nabla f(x)^T \nabla f(x) - \frac{\nabla f(x)^T d d^T \nabla f(x)}{d^T \nabla f(x)} \right. \\ &\quad \left. - \frac{\nabla f(x)^T \nabla f(x) \nabla f(x)^T \nabla f(x)}{\nabla f(x)^T \nabla f(x)} \right) \\ &= -(\nabla f(x)^T \nabla f(x) - d^T \nabla f(x) - \nabla f(x)^T \nabla f(x)) \\ &= d^T \nabla f(x) < 0. \end{aligned}$$



即 p 是一个下降方向. ■

题 6

二选一:

(1) G 是 n 阶正定矩阵, \mathbb{R}^n 上的函数 $f(x) = \frac{1}{2}x^T Gx + b^T x$, 用共轭方向法求最小值, 证明精确线性搜索在第 n 步得到准确值.(题目大致是这样)

(2)

$$\begin{aligned} \min \quad & \frac{1}{2} [(x_1 - 1)^2 + x_2^2] \\ \text{s.t.} \quad & x_1 = \beta x_2^2. \end{aligned}$$

β 取何值时, $(0, 0)^T$ 是局部极小点.

证明. (1) 问. 参考课本 P148 定理 4.3.3 的证明. 因为 G 正定, 且共轭方向 d_0, d_1, \dots 线性无关, 故只需要证明对所有的 $i \leq n-1$, 有

$$g_{i+1}^T d_j = 0, \quad j = 0, \dots, i, \quad (\text{II.1})$$

就可得出定理的结论. 因为当 $i = n-1$ 时, $g_n^T d_j = 0 (j = 0, \dots, n-1)$, 于是 $g_n = 0$, 从而 x_n 是极小点.

现在证明(II.1). 由 $g_k = Gx_k + b$,

$$g_{k+1} - g_k = G(x_{k+1} - x_k) = \alpha G d_k,$$

又, 在精确线性搜索下, $g_{k+1}^T d_k = 0$, 故当 $j < i$ 时,

$$\begin{aligned} g_{i+1}^T d_j &= g_{j+1}^T d_j + \sum_{k=j+1}^i (g_{k+1} - g_k)^T d_j \\ &= g_{j+1}^T d_j + \sum_{k=j+1}^i \alpha_k d_k^T G d_j \\ &= 0. \end{aligned}$$

当 $j = i$ 时, 直接由精确线性搜索可知

$$g_{i+1}^T d_i = 0.$$

从而(II.1)成立. ■



思路 – (2) 问

将 $x_1 = \beta x_2^2$ 代入目标函数, 化成单变量的无约束优化, 且 $(0, 0)^T$ 是局部极小当且仅当 $x_2 = 0$ 是局部极小, 此时 $f'(0) = 0, f''(0) \geq 0$, 然后解出 β 的范围.

SECTION II.2

往年题 (2011)

只写出我们学过的知识点的题.

II.2.1 简答题

题 1

请给出凸函数的定义及 f 是凸集 D 上的凸函数的充要条件.

解. 定义. 对任意的 $x, y \in \mathbb{R}^n$ 及任意的 $0 \leq \lambda \leq 1$, 有

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

充要条件.(定义不就是吗, 不知道为什么这么问) 对任意的 $x, y \in D$ 及任意的 $0 \leq \lambda \leq 1$, 有

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

题 2

请给出求解无约束优化问题 $\min f(x)$ 的一般算法步骤.

解. 参考算法:

算法. step 1. 给定初始点 x_0 , 令 $k := 0$.

step 2. 确定搜索方向 d_k .

step 3. 确定搜索步长 α_k , 满足

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k).$$



step 4. 判断终止条件, 满足则终止; 否则, $k := k + 1$, 转 step 2.

题 3

对于无约束优化问题, 构造搜索方向 $d^{(k)}$ 有那些方法? 请给出 $d^{(k)}$ 的四种表达式.

解. 最速下降法. $d_k = -g_k$.

牛顿法. $d_k = -G_k^{-1}g_k$.

共轭梯度法. 以 FR 为例, $r_{k+1} = r_k + \alpha_k G d_k$, $\beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}$, $d_{k+1} = -r_{k+1} + \beta_k d_k$.

拟牛顿法. $d_k = -B_k^{-1}g_k$.

解毕.

题 4

请给出在 x 处下降方向 d 所满足的条件.

解.

$$\nabla f(x)^T d < 0.$$

题 5

请给出内点惩罚函数法中的惩罚函数/障碍函数的表达式.

解. 对数障碍函数

$$P(x; \mu) = f(x) - \mu \sum_{i \in \mathcal{I}} \log c_i(x).$$

分数障碍函数

$$P(x; \mu) = f(x) - \mu \sum_{i \in \mathcal{I}} \frac{1}{c_i(x)}.$$



题 6

请给出关于对称正定矩阵 A 共轭的含义, 并给出矩阵 $A = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$ 的一组共轭方向.

思路

向量 d_1, d_2 满足 $d_1^T A d_2 = 0, d_1, d_2$ 是共轭的.

矩阵 A 只有两个共轭方向, 令 $d_1 = (1, 0)^T$, 代入上式解 d_2 .

II.2.2 大题

题 1

对于无约束优化问题 $\min f(x)$:

1. 请给出求解无约束问题一阶必要性条件;
2. 请给出无约束问题最优解的二阶必要条件;
3. 请给出无约束问题最优解的二阶充分条件;

解. 一阶必要条件. 设 $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ 在开集上连续可微, 若 $x^* \in D$ 是局部极小点, 则

$$g(x^*) = 0.$$

二阶必要条件. 设 $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ 在开集 D 上二阶连续可微, 若 $x^* \in D$ 是局部极小点, 则

$$g(x^*) = 0, \quad G(x^*) \geq 0.$$

二阶充分条件. 设 $f: D \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ 在开集 D 上二阶连续可微, 则 $x^* \in D$ 是 f 的严格局部极小点的充分条件是

$$g(x^*) = 0 \text{ 和 } G(x^*) \text{ 是正定矩阵.}$$



题 2

设有非线性方程组:

$$\begin{cases} f_1(x) = x_1^3 - 2x_2^2 - 1 = 0, \\ f_2(x) = 2x_1 + x_2 - 2 = 0. \end{cases}$$

1. 列出求解这个方程组的非线性最小二乘问题的数学模型;
2. 写出求解该问题的 Gauss-Newton 法的具体形式;
3. 取初始可行点 $x_0 = (2, 2)^T$, 迭代一次.

解. 1. $\min \frac{1}{2}(f_1^2(x) + f_2^2(x))$.

2. Jacobian 矩阵为 $A, f_k = f(x_k)$, 格式

$$x_{k+1} = x_k + \delta_k = x_k - (A_k^T A_k)^{-1} A_k^T f_k. \quad (\text{II.2})$$

3.

算法. 步 1. 给定初始点 x_0 , 置 $k = 0$.

步 2. 将 x_k 代入(II.2)得 x_{k+1} .

题 3

证明: B_k 对称正定, B_{k+1} 由 BFGS 校正公式

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{y_k^T s_k}$$

确定, 那么 B_{k+1} 对称正定的充要条件是 $y_k^T s_k > 0$.

证明. 可以借鉴课本 P163 定理 4.4.2 的证明.

“ \Rightarrow ”. 考虑用归纳法证明

$$z^T B_{k+1} z > 0, \quad z \neq 0. \quad (\text{II.3})$$

根据初始的选择, B_0 正定. 假设对某个 k 结论成立, 记 $B_k = LL^T$ 为 B_k 的 Cholesky



分解. 设

$$a = L^T z, \quad b = L^T s_k. \quad (\text{II.4})$$

则

$$\begin{aligned} z^T B_{k+1} z &= z^T \left(B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} \right) z + z^T \frac{y_k y_k^T}{s_k^T y_k} z \\ &= a^T a - \frac{(a^T b)^2}{b^T b} + \frac{(z^T y_k)^2}{s_k^T y_k} \end{aligned} \quad (\text{II.5})$$

根据 Cauchy 不等式,

$$a^T a - \frac{(a^T b)^2}{b^T b} \geq 0, \quad (\text{II.6})$$

又因为 $y_k^T s_k > 0$, 则

$$z^T H_{k+1} z \geq 0.$$

下面证明, (II.5) 右边两项至少一项严格大于 0. 若第一项等于 0, 由 Cauchy 不等式, a 与 b 平行, 则 z 与 s_k 平行. 设 $z = \beta s_k, \beta \neq 0$, 这时,

$$\frac{(z^T y_k)^2}{s_k^T y_k} = \beta^2 y_k^T s_k > 0,$$

则第一项严格大于 0 时, 第二项也严格大于 0.

若第二项为 0, 则 $z^T y_k = 0$, z 与 s_k 不平行, 从而第一项大于 0, 否则, $z = \beta s_k, \beta \neq 0$, 又 $s_k^T y_k > 0$, 则 $z^T y_k = \beta s_k^T y_k \neq 0$, 矛盾. 充分性得证.

“ \Leftarrow ”. 必要性可类似证明. ■



对于上一页的题, 其实在课本 P163 定理 4.4.2 中已经说明 $s_k^T y_k > 0$ 当且仅当 H_{k+1} 正定, 又因为 $H_{k+1} = B_{k+1}^{-1}$, 则 H_{k+1} 正定当且仅当 B_{k+1} 正定, 这样也证明了结论.



题 4

对于下列约束优化最问题

$$\begin{aligned} \min \quad & f(x) = (x_1 - 3)^2 + (x_2 - 2)^2 \\ \text{s.t.} \quad & x_1 + 2x_2 = 4, \\ & x_1^2 + x_2^2 \leq 5, \\ & x_1, x_2 \geq 0. \end{aligned}$$

1. 请给出求解一般约束优化问题的 KKT 一阶必要条件;
2. 验证 $\bar{x} = (2, 1)^T$ 是否为该约束最优化问题的 KKT 点.

思路

直接用 KKT 条件.

题 5

问题:

$$\begin{aligned} \min \quad & f(x) = x_1^2 - x_1x_2 + 2x_2^2 - x_1 - 10x_2 \\ \text{s.t.} \quad & -3x_1 - 2x_2 \geq -6 \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

证明: 该问题是凸二次规划问题.

证明. 一方面, 可行域 D 是凸集.

另一方面, $\nabla^2 f(x) = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}$ 是正定矩阵, f 是凸函数. ■