

运筹学基础：课后练习 1 答案

1. (a) 因为 $\lim_{k \rightarrow \infty} \frac{1}{k!} = 0$ 设 $x^* = 0$ 。

$\lim_{k \rightarrow \infty} \frac{|x^{k+1}|}{|x^k|} = \lim_{k \rightarrow \infty} \frac{\frac{1}{(k+1)!}}{\frac{1}{k!}} = \lim_{k \rightarrow \infty} \frac{1}{k+1} = 0$ 。因为 $\lim_{k \rightarrow \infty} \frac{|x^{k+1}|}{|x^k|} = 0$ ，所以点列 $\{x^k\}$ 超线性收敛。

- (b) 当 k 为奇数时： $k = 2m$ ， $x^{2m} = (\frac{1}{4})^{2^m}$ ；

当 k 为偶数时： $k = 2m + 1$ ， $x^{2m+1} = \frac{x^{2m}}{2m+1}$ 。

因此， $\lim_{k \rightarrow \infty} x^k = 0$ 设 $x^* = 0$ 。

对于偶数 $k = 2m$ ， $\frac{|x^{2m+1}|}{|x^{2m}|} = \frac{1}{2m+1}$ ，

对于奇数 $k = 2m + 1$ ， $\frac{|x^{2m+2}|}{|x^{2m+1}|} = \frac{(\frac{1}{4})^{2^{m+1}}}{\frac{(\frac{1}{4})^{2^m}}{2m+1}} = (2m+1)(\frac{1}{4})^{2^{m+1}-2^m}$ 。

$\lim_{k \rightarrow \infty} \frac{|x^{k+1}|}{|x^k|} = 0$ ，所以 $\{x^k\}$ 超线性收敛。

2. (a) 已知 $f(x) = x_1^2 + x_2^2$ ， $x^k = (1 + \frac{1}{2^k})(\cos k, \sin k)^T$ ，

则 $f(x^k) = (1 + \frac{1}{2^k})^2(\cos^2 k + \sin^2 k) = (1 + \frac{1}{2^k})^2$ 。 $\lim_{k \rightarrow \infty} f(x^k) = 1$ 。

$\frac{|f(x^{k+1})-1|}{|f(x^k)-1|} = \frac{(1+\frac{1}{2^{k+1}})^2-1}{(1+\frac{1}{2^k})^2-1} = \frac{1+\frac{1}{2^k}+\frac{1}{2^{2k+2}}-1}{1+\frac{1}{2^{k-1}}+\frac{1}{2^{2k}}-1} = \frac{\frac{1}{2^k}+\frac{1}{2^{2k+2}}}{\frac{1}{2^{k-1}}+\frac{1}{2^{2k}}} \lim_{k \rightarrow \infty} \frac{|f(x^{k+1})-1|}{|f(x^k)-1|} = \frac{1}{2}$ ，

所以 $\{f(x^{k+1})\}$ 线性收敛 (Q -线性收敛)。

- (b) 由于 $\lim_{k \rightarrow \infty} x^k = \lim_{k \rightarrow \infty} (1 + \frac{1}{2^k})(\cos k, \sin k)^T$ ，由于 $\cos k$ 和 $\sin k$ 是周期函数，所以 $\{x^{k+1}\}$ 不收敛。

3. 矩阵 A 的奇异值分解为 $A = U\Sigma V^T$ ，其中 U 和 V 是正交矩阵， $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r, 0, \dots, 0)$ ， $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ 。

对于任意 $\|x\|_2 = 1$ ，有 $\|Ax\|_2 = \|U\Sigma V^T x\|_2 = \|\Sigma V^T x\|_2$ 。

令 $y = V^T x$ ，则 $\|y\|_2 = 1$ ， $\|\Sigma y\|_2 = \sqrt{\sum_{i=1}^n \sigma_i^2 y_i^2} \leq \sigma_1$ 。

当 $y = (1, 0, \dots, 0)^T$ 时， $\|\Sigma y\|_2 = \sigma_1$ 。

所以 $\sigma_1(A) = \max_{\|x\|_2=1} \|Ax\|_2$ 。

4. 证明:

(a)

$$\|AB\|_F^2 = \sum_{i=1}^n \|Ab_i\|_2^2$$

因为 $\|Ab_i\|_2 \leq \|A\|_2 \|b_i\|_2$, 所以:

$$\|AB\|_F^2 \leq \sum_{i=1}^n \|A\|_2^2 \|b_i\|_2^2 = \|A\|_2^2 \sum_{i=1}^n \|b_i\|_2^2 = \|A\|_2^2 \|B\|_F^2$$

两边开平方, 得到:

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

(b) 对矩阵 A 和 B 进行奇异值分解:

$$A = U\Sigma V^T, \quad B = X\Gamma Y^T$$

代入内积:

$$\langle A, B \rangle = \text{tr}(A^T B) = \text{tr}(V\Sigma U^T X\Gamma Y^T) = \text{tr}(\Gamma Y^T V\Sigma U^T X)$$

令 $W = Y^T V$, $Z = U^T X$, 则:

$$\langle A, B \rangle = \text{tr}(\Gamma W\Sigma Z)$$

由于 W 和 Z 是正交矩阵的乘积, 它们的元素满足 $|W_{ij}| \leq 1$ 和 $|Z_{ji}| \leq 1$ 。因此:

$$\langle A, B \rangle = \sum_{i,j} \sigma_i(B) \sigma_j(A) W_{ij} Z_{ji} \leq \sum_{i,j} \sigma_i(B) \sigma_j(A)$$

谱范数 $\|A\|_2$ 是矩阵 A 的最大奇异值, 即 $\|A\|_2 = \sigma_1(A)$ 。核范数 $\|B\|_*$ 是矩阵 B 的奇异值之和, 即 $\|B\|_* = \sum_{i=1}^n \sigma_i(B)$ 。

因此:

$$\langle A, B \rangle \leq \sum_{i,j} \sigma_i(B) \sigma_j(A) \leq \sigma_1(A) \sum_{i=1}^n \sigma_i(B) = \|A\|_2 \|B\|_*$$

5. 先证明 A 可逆。分块矩阵的行列式公式: $\det(A) = \det(I - BB^T)$, 因为 $\|B\|_2 < 1$, BB^T 的所有特征值 λ_i 满足 $|\lambda_i| \leq \|B\|_2^2 < 1$, 所以 $\det(I - BB^T) \neq 0$, A 可逆。

然后求矩阵 A 的特征值。已知 $A = \begin{bmatrix} I & B \\ B^T & I \end{bmatrix}$, 设 λ 是 A 的特征值, $z = \begin{bmatrix} x \\ y \end{bmatrix}$

是对应的特征向量, 其中 $x, y \in \mathbb{R}^n$ 。则 $A \begin{bmatrix} x \\ y \end{bmatrix} = \lambda \begin{bmatrix} x \\ y \end{bmatrix}$, 即

$$\begin{cases} x + By = \lambda x \\ B^T x + y = \lambda y \end{cases}.$$

由第一个方程可得 $By = (\lambda - 1)x$, 两边左乘 B^T , 得到 $B^T By = (\lambda - 1)B^T x$ 。

由第二个方程可得 $B^T x = (\lambda - 1)y$, 则有 $B^T By = (\lambda - 1)^2 y$ 。

同理, 可得 $BB^T x = (\lambda - 1)^2 x$ 。

这表明 $(\lambda - 1)^2$ 是 $B^T B$ 和 BB^T 的特征值。

设 σ 是 B 的奇异值, 则 σ^2 是 $B^T B$ 的特征值, 我们有 $(\lambda - 1)^2 = \sigma^2$, 所以 $\lambda = 1 \pm \sigma$ 。

根据矩阵 ℓ_2 范数的定义 $\|A\|_2 = \max_{\|z\|_2=1} \|Az\|_2$, 且矩阵的 ℓ_2 范数等于其最大奇异值, 而矩阵的奇异值是其特征值的绝对值。因为 $\|B\|_2$ 是 B 的最大奇异值, A 的特征值为 $\lambda = 1 \pm \sigma$ (σ 是 B 的奇异值), 要使 $|\lambda|$ 最大, 当取 $\lambda = 1 + \|B\|_2$ 时 (因为 $\sigma \geq 0$, 且 $\|B\|_2$ 是 B 的最大奇异值), 所以 $\|A\|_2 = 1 + \|B\|_2$ 。

接着求 A^{-1} 。先求 A 的逆矩阵, 已知 $A = \begin{bmatrix} I & B \\ B^T & I \end{bmatrix}$, 利用分块矩阵求逆公式

$$\begin{bmatrix} M & N \\ P & Q \end{bmatrix}^{-1} = \begin{bmatrix} M^{-1} + M^{-1}N(Q - PM^{-1}N)^{-1}PM^{-1} & -M^{-1}N(Q - PM^{-1}N)^{-1} \\ -(Q - PM^{-1}N)^{-1}PM^{-1} & (Q - PM^{-1}N)^{-1} \end{bmatrix}$$

对于 A , $M = I$, $N = B$, $P = B^T$, $Q = I$, 则

$$A^{-1} = \begin{bmatrix} (I - BB^T)^{-1} & -(I - BB^T)^{-1}B \\ -B^T(I - BB^T)^{-1} & (I - B^T B)^{-1} \end{bmatrix}.$$

然后求 A^{-1} 的特征值, 设 μ 是 A^{-1} 的特征值, 由 A 的特征值 λ 与 A^{-1} 的特征值 μ 的关系得 $\mu = \frac{1}{\lambda}$ 。

已知 A 的特征值为 $\lambda = 1 \pm \sigma$, 要使 $|\mu|$ 最大, 因为 $\|B\|_2 < 1$, $\sigma \in [0, \|B\|_2]$, 则 $|\mu|$ 最大时, λ 取最小的非零值 $1 - \|B\|_2$ (因为 $\lambda = 1 \pm \sigma$, 当 $\sigma = \|B\|_2$ 且取 $\lambda = 1 - \|B\|_2$ 时, $\frac{1}{|\lambda|}$ 最大), 所以 $\|A^{-1}\|_2 = \frac{1}{1 - \|B\|_2}$ 。

综上,

$$\|A\|_2 \|A^{-1}\|_2 = \frac{1 + \|B\|_2}{1 - \|B\|_2}$$

6. 证明 $\langle A, B \rangle \geq 0$ (A, B 为半正定矩阵) 因为 A 是半正定矩阵, 存在正交矩阵 U 使得 $A = U\Lambda U^T$, 其中 $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$, $\lambda_i \geq 0, i = 1, \dots, n$ 。
 $\langle A, B \rangle = \text{Tr}(A^T B) = \text{Tr}(U\Lambda U^T B) = \text{Tr}(\Lambda U^T B U)$ 令 $C = U^T B U$, C 也是半正定矩阵, 其对角线元素非负。 $\langle A, B \rangle = \text{Tr}(\Lambda C) = \sum_{i=1}^n \lambda_i c_{ii} \geq 0$

7. 具体推导见课件

(a) $\nabla f(X) = ab^T$

(b) $\nabla f(X) = (A + A^T)X$

(c) $\nabla f(X) = X^{-T}$

8. (a) 1.-4. 套用定义即可得证。

- (b) 5. 设 $(x_1, t_1), (x_2, t_2) \in \mathcal{C}$, 即 $t_1 \geq x_1^T x_1, t_1 \geq 0$ 且 $t_2 \geq x_2^T x_2, t_2 \geq 0$ 。
 对于任意 $\lambda \in [0, 1]$, 令 $(x, t) = \lambda(x_1, t_1) + (1 - \lambda)(x_2, t_2) = (\lambda x_1 + (1 - \lambda)x_2, \lambda t_1 + (1 - \lambda)t_2)$ 。

$$\begin{aligned} x^T x &= (\lambda x_1 + (1 - \lambda)x_2)^T (\lambda x_1 + (1 - \lambda)x_2) \\ &= \lambda^2 x_1^T x_1 + 2\lambda(1 - \lambda)x_1^T x_2 + (1 - \lambda)^2 x_2^T x_2 \\ &\leq \lambda^2 x_1^T x_1 + (1 - \lambda)^2 x_2^T x_2 \\ &\leq \lambda^2 t_1 + (1 - \lambda)^2 t_2 = t, \end{aligned}$$

第一个不等号根据柯西 - 施瓦茨不等式 $2x_1^T x_2 \leq x_1^T x_1 + x_2^T x_2$ 。

9. (a) 套用定义即可得证。

- (b) 设 $f(x) = \frac{1}{2}x^T A x - b^T x + c$. 利用二次罚函数, 令

$$\tilde{f}(x) = f(x) + \frac{\lambda}{2}(g^T x + h)^2,$$

其中 $\lambda \in \mathbb{R}$ 是给定的常数, 满足 $A + \lambda g g^T \succeq 0$ 。设

$$C'' := \{x \in \mathbb{R}, \tilde{f}(x) \leq 0\},$$

根据 (a) 中的结论, 若 $A + \lambda g g^T$ 正定, 则 C'' 是凸集。在 $g^T x + h = 0$ 上, $C' = \{x : \tilde{f}(x) \leq 0, g^T x + h = 0\}$ 与集合 C'' 等价, 得证。

10. 所有函数均为凸函数。

$$(a) \ f^*(y) = \sum_{i=1}^n e^{y_i - 1}.$$

$$(b) \ f^*(Y) = \begin{cases} -n - \ln \det(-Y), & Y \preceq 0 \\ +\infty, & \text{otherwise} \end{cases}.$$

$$(c) \ f^*(y) = \begin{cases} 0, & y_i \leq 0, \sum_{i=1}^n y_i = 1 \\ +\infty, & \text{otherwise} \end{cases}.$$

$$(d) \ f^*(y, s) = \begin{cases} -\ln(1 - y^T y - s^2) - 1, & y^T y + s^2 < 1 \\ +\infty, & \text{otherwise} \end{cases}.$$