Ps0问题回答

2022年7月6日 星期三 12:30

复习

1. 称 $A \in \mathbb{R}^{n\times n}$ 在 $A = A^T$ 时为对称矩阵，也就是对全体 i,j 均有 $A_{ij} = A_{ji}$

又称梯度 $\nabla f(x)$ 是 $\mathbb{R}^n \to \mathbb{R}$ 的映射，也就是 n 维向量的偏导

$$\nabla f(x) = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) \\ \frac{\partial}{\partial x_2} f(x) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x) \end{bmatrix} \quad 当 \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

海森 $\nabla^2 f(x)$ 为 $\mathbb{R}^n \to \mathbb{R}$ 的映射，同时形成一个实对称矩阵的二阶偏导

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2}{\partial x_1^2} f(x) & \frac{\partial^2}{\partial x_1 \partial x_2} f(x) & \cdots & \frac{\partial^2}{\partial x_1 \partial x_n} f(x) \\ \frac{\partial^2}{\partial x_2 \partial x_1} f(x) & \frac{\partial^2}{\partial x_2^2} f(x) & \cdots & \frac{\partial^2}{\partial x_2 \partial x_n} f(x) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \frac{\partial^2}{\partial x_n \partial x_2} f(x) & \cdots & \frac{\partial^2}{\partial x_n \partial x_n} f(x) \end{bmatrix}$$

a) $f(x) = \frac{1}{2} x^T A x + b^T x$，$A$是对称阵，$b \in \mathbb{R}^n$ 为向量

$\nabla f(x) = ?$

「解答」：
$$\nabla f(x) = \frac{1}{2}\Big[\sum_{i\neq k} \frac{\partial}{\partial x_k}\Big(\sum_{j=1}^{n} a_{ij} x_i x_j\Big) + \frac{\partial}{\partial x_k}\Big(\sum_{j=1}^{n} a_{kj} x_k x_j\Big)\Big] + b$$
$$= \frac{1}{2}\Big[\sum (0 + a_{ik} x_i) + \sum_{j\neq k} a_{kj} x_j + 2a_{kk} x_k\Big] + b$$
$$= \frac{1}{2}\Big[\sum_{i=1}^{n} a_{ik} x_i + \sum_{j=1}^{n} a_{kj} x_j\Big] + b$$
$$= \frac{1}{2}\big[x^T A + A x\big] + b$$
$$= Ax + b$$

b) $f(x) = g(h(x))$，$\nabla f(x) = ?$

「解答」：$\frac{\partial g(h(x))}{\partial x} = g'(h(x)) \cdot \frac{\partial h(x)}{\partial x}$

则 $\nabla f(x) = \nabla g(h(x)) = g'(h(x)) \cdot \nabla h(x)$

c) $f(x) = \frac{1}{2} x^T A x + b^T x$，$A$实对称阵，$b$是向量

$\nabla^2 f(x) = ?$

「解答」：由 (a)，$\nabla f(x) = Ax + b$

$$则 \quad \nabla^2 f(x) = \begin{bmatrix} \frac{\partial Ax+b}{\partial x_1} & \frac{\partial \nabla Ax+b}{\partial x_2} & \cdots\cdots & \frac{\partial \nabla Ax+b}{\partial x_n} \end{bmatrix}$$
$$= \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix} = A$$

d) $f(x) = g(a^T x)$，$\nabla f(x) = ?$ $\nabla^2 f(x) = ?$

「解答」：$\nabla f(x) = \nabla g(a^T x) = g'(a^T x) \nabla (a^T x) \cdot a$

$$\frac{\partial^2 g(h(x))}{\partial x_i \partial x_j} = \frac{\partial^2 g(h(x))}{\partial (h(x))^2} \cdot \frac{\partial h(x)}{\partial x_i} \cdot \frac{\partial h(x)}{\partial x_j} = g''(h(x)) \cdot \frac{\partial h(x)}{\partial x_i} \cdot \frac{\partial h(x)}{\partial x_j}.$$

$$\frac{\partial^2 g(a^T x)}{\partial x_i \partial x_j} = g''(a^T x) \frac{\partial (a^T x)}{\partial x_i} \cdot \frac{\partial (a^T x)}{\partial x_j} = g''(a^T x) \cdot a_i a_j$$

$$\nabla^2 f(x) = \nabla^2 g(a^T x) = g''(a^T x) = \begin{bmatrix} a_1 a_1 & a_1 a_2 & \cdots & a_1 a_n \\ a_2 a_1 & a_2 a_2 & \cdots & a_2 a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_n a_1 & a_n a_2 & \cdots & a_n a_n \end{bmatrix} = g''(a^T x) a a^T.$$

2. a) $z \in \mathbb{R}^n$，证明 $A = z z^T$ 为正定矩阵.

「解答」：$A^T = (z z^T)^T = z \cdot z^T = A$

则 $x^T A x = x^T z \cdot z^T x = (z^T x)^T \cdot z^T x$
$= (x^T z)^2 \geq 0$

b) $z \in \mathbb{R}^n$ 为非0 n 一向量，$A = z z^T$，求 A 的零空间以及 A 的秩.

「正解」：$N(A) = \{ x \in \mathbb{R}^n : x^T z = 0 \}$

$R(A) = R(z z^T) = 1$

c) $A \in \mathbb{R}^{n\times n}$ 为正定阵，$B \in \mathbb{R}^{m\times n}$ 为任意矩阵，BAB^T 为 PSD？是，证；否，举例.

「正解」：$(BAB^T)^T = BA^T B^T = BAB^T$

$x^T B A B^T x = (x^T B) A (x^T B)^T \geq 0$

3. a) $A = T \Lambda T^{-1}$ $AT = T\Lambda$

$$A[t^{(1)}, t^{(2)} \cdots t^{(n)}] = [t^{(1)}\ t^{(2)} \cdots t^{(n)}] \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ & & \lambda_3 & \\ 0 & & & \ddots & \lambda_n \end{bmatrix}$$

则 $A t^{(i)} = \lambda_i t^{(i)}$

b) $A = U \Lambda U^T$，$AU = U\Lambda$

$$A[u^{(1)}, u^{(2)} \cdots u^{(n)}] = \begin{bmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \ddots & \lambda_n \end{bmatrix}$$

则 $A u^{(i)} = \lambda_i u^{(i)}$

c) $A t^{(i)} = \lambda_i t^{(i)}$

$(t^{(i)})^T A t^{(i)} = \lambda_i \| t^{(i)} \|_2 = \lambda_i \geq 0$