1.(a) $\frac{1}{2}x^TAx = \frac{1}{2}\underbrace{(x_1, x_2 \cdots x_n)}_{1\times n}\underbrace{\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}}_{n\times n}\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{n\times 1}$

$= \frac{1}{2}\underbrace{(A_{11}x_1 + A_{21}x_2 + \cdots A_{n1}x_n, \cdots, A_{1n}x_1 + A_{2n}x_2 + \cdots A_{nn}x_n)}_{1\times n}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

$= \frac{1}{2}[(A_{11}x_1 + A_{21}x_2 + \cdots A_{n1}x_n)x_1 + \cdots (A_{1n}x_1 + A_{2n}x_2 + \cdots A_{nn})x_n]$

$= \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij}x_ix_j$

对于任意 $k=1, \cdots n$. 有:

$$\frac{\partial}{\partial x_k}\frac{1}{2}x^TAx = \frac{\partial}{\partial x_k}\frac{1}{2}\Big(\sum_{i=1, i\neq k}^{n} A_{ik}x_ix_k + \sum_{j=1, j\neq k}^{n} A_{kj}x_kx_j + A_{kk}x_k^2 + \sum_{i=1, i\neq k}^{n}\sum_{j=1, j\neq k}^{n} A_{ij}x_ix_j\Big)$$

$$= \frac{1}{2}\sum_{i=1, i\neq k}^{n} A_{ik}x_i + \frac{1}{2}\sum_{j=1, j\neq k}^{n} A_{kj}x_j + A_{kk}x_k$$

$$= \sum_{i=1}^{n} A_{ki}x_i \quad (A_{ij} = A_{ji})$$

当 $k=1, \cdots, n$

$$\frac{\partial}{\partial x_k}\frac{1}{2}x^TAx = \underbrace{\begin{pmatrix} \sum_{i=1}^{n} A_{1i}x_i \\ \sum_{i=1}^{n} A_{2i}x_i \\ \vdots \\ \sum_{i=1}^{n} A_{ni}x_i \end{pmatrix}}_{n\times 1} = Ax$$

$$\frac{\partial}{\partial x_k}b^Tx = \frac{\partial}{\partial x_k}(b_1, b_2 \cdots b_n)\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \frac{\partial}{\partial x_k}\sum_{i=1}^{n} b_ix_i = b_k$$

当 $k=1. \cdots n$

$$\frac{\partial}{\partial x_k}b^Tx = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = b.$$

总结:对于对称矩阵A.

$\nabla(\frac{1}{2}x^TAx) = Ax$

$\nabla(b^Tx) =$

$\therefore \nabla f(x) = \nabla(\frac{1}{2}x^TAx + b^Tx) = Ax + b$

CS 229, Fall 2018

Problem Set #0 Solutions: Linear Algebra and Multivariable Calculus

1.

(a)

$$\nabla f(x) = \nabla(\frac{1}{2}x^T Ax + b^T x) = Ax + b \rightarrow \text{详细看手写}$$

(b)

$$\frac{\partial g(h(x))}{\partial x_i} = \frac{\partial g(h(x))}{\partial h(x)} \frac{\partial h(x)}{\partial x_i} = g'(h(x))\frac{\partial h(x)}{\partial x_i}$$

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

(c)

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

$$Ax+b = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ \vdots & \vdots & & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$
$$= \begin{pmatrix} A_{11}x_1 + A_{12}x_2 + \cdots A_{1n}x_n \\ A_{21}x_1 + \cdots \quad A_{2n}x_n \\ \vdots \\ A_{n1}x_1 + \cdots \quad A_{nn}x_n \end{pmatrix} + \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

$$\frac{\partial(Ax+b)}{\partial x_1} = \begin{pmatrix} A_{11} \\ A_{21} \\ \vdots \\ A_{n1} \end{pmatrix}$$

(d)

$$\nabla f(x) = \nabla g(a^T x) = g'(a^T x)\nabla(a^T x) = g'(a^T x)a$$

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

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

$$\nabla^2 f(x) = \nabla^2 g(a^T x) = g''(a^T x)\begin{bmatrix} a_1a_1 & a_1a_2 & \dots & a_1a_n \\ a_2a_1 & a_2a_2 & \dots & a_2a_n \\ \vdots & \vdots & \ddots & \vdots \\ a_na_1 & a_na_2 & \dots & a_na_n \end{bmatrix} = g''(a^T x)aa^T$$

2.

(a)

$$A^T = (zz^T)^T = zz^T = A$$

$$x^T Ax = x^T zz^T x = x^T z(x^T z)^T = (x^T z)^2 \geq 0$$

(b)

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

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

(c)

$$(BAB^T)^T = BA^T B^T = BAB^T$$

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

3.

(a)

$$A = T\Lambda T^{-1}$$

$$AT = T\Lambda$$

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

$$[\,At^{(1)} \quad At^{(2)} \quad \dots \quad At^{(n)}\,] = [\,\lambda_1 t^{(1)} \quad \lambda_2 t^{(2)} \quad \dots \quad \lambda_n t^{(n)}\,]$$

$$At^{(i)} = \lambda_i t^{(i)}$$

(b)

$$A = U\Lambda U^T$$

$$AU = U\Lambda U^T U = U\Lambda$$

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

$$[\,Au^{(1)} \quad Au^{(2)} \quad \dots \quad Au^{(n)}\,] = [\,\lambda_1 u^{(1)} \quad \lambda_2 u^{(2)} \quad \dots \quad \lambda_n u^{(n)}\,]$$

$$Au^{(i)} = \lambda_i u^{(i)}$$

(c)

$$At^{(i)} = \lambda_i t^{(i)}$$

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