Problem Set #0

① if not symmetric, $\nabla(x^T A x) = (A + A^T)x$

a) Let $f(x) = \frac{1}{2} x^T A x + b^T x$, A symmetric, $b \in \mathbb{R}^n$

$\nabla f(x) = \frac{1}{\cancel{2}} \cancel{2} A x + b = Ax + b$ ∎

b) $f(x) = g(h(x))$ $g: \mathbb{R} \to \mathbb{R}$ diff, $h: \mathbb{R}^n \to \mathbb{R}$ diff.

$$\nabla_x f(x) = \nabla_x g(h(x)) = \frac{d}{dh} g(h) \begin{bmatrix} \frac{\partial h(x)}{\partial x_1} \\ \vdots \\ \frac{\partial h(x)}{\partial x_n} \end{bmatrix} = \frac{d}{dh} g(h) \nabla_x h(x)$$ ∎

c) $\nabla_x f(x) = 2Ax + b$

$$J \nabla_x f(x) = J(2Ax) + J(\cancel{b})^{\to 0} = 2 \cdot J \begin{bmatrix} \sum_{i=1}^n A_{1i} x_i \\ \sum_{i=1}^n A_{2i} x_i \\ \vdots \\ \sum_{i=1}^n A_{ni} x_i \end{bmatrix} = 2 \cdot \begin{bmatrix} (\nabla_x \sum_{i=1}^n A_{1i} x_i) \\ \vdots \\ (\nabla_x \sum_{i=1}^n A_{ni} x_i) \end{bmatrix} = 2 \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}$$

$= 2A$ ∎

d) $f(x) = g(a^T x)$, $g: \mathbb{R} \to \mathbb{R}$ cont diff and $b \in \mathbb{R}^n$

$$\nabla f(x) = \frac{dg(y)}{dy} \nabla_x (a^T x) = \left[\frac{d}{dy} g(y)\right] a = \begin{bmatrix} \frac{d}{dy} g(y)\, a_1 \\ \vdots \\ \frac{d}{dy} g(y)\, a_2 \end{bmatrix}$$

$$\nabla^2 f(x) = g''(a^T x) \cdot (a a^T)$$

②

a) Let $A = zz^T$. Consider $x \in \mathbb{R}^n$.

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

$\Rightarrow$ A is positive semi-definite.

b) $z \in \mathbb{R}^n$ and non-zero. $A = zz^T$

Null space of $A = zz^T$ is $\{x : Ax = 0\} = \{x : zz^T x = 0\}$

Null $A \perp$ row A. row A = span$\{z_1 z, z_2 z, \ldots, z_n z\}$ = span$\{z\}$ ($\Rightarrow$ rank$(A) = 1$)

$\Rightarrow$ Null $A = \{x : z^T x = 0\}$

c) $A \in \mathbb{R}^{n \times n}$ be positive semidef, $B \in \mathbb{R}^{m \times n}$ be given.

Note: $(BAB^T)^T = BA^T B^T = BAB^T$ b/c A is PSD (by defn symmetric)

so BAB^T is symmetric. Let $x \in \mathbb{R}^n$

$x^T (BAB^T) x = (B^T x)^T A (B^T x) \geq 0$ since A is PSD and

$B^T x \in \mathbb{R}^n$ ∎

③

a) Let $A \in \mathbb{R}^{n \times n}$ and A is diagonalizable, i.e., $A = T \Lambda T^{-1}$.

Note, $A = T \Lambda T^{-1} \Rightarrow AT = T\Lambda$.

$$\Rightarrow \quad A \left[t^{(1)} \; t^{(2)} \cdots t^{(n)} \right] = \left[t^{(1)} \; t^{(2)} \cdots t^{(n)} \right] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

$$\Rightarrow \quad \left[At^{(1)} \; At^{(2)} \cdots At^{(n)} \right] = \left[\lambda_1 t^{(1)} \; \lambda_2 t^{(2)} \cdots \lambda_n t^{(n)} \right]$$

$$\Rightarrow \quad At^{(i)} = \lambda_i t^{(i)} \; \blacksquare$$

b) Follows from above and noting that $U^{-1} = U^T$ so U^{-1} always exists

c) Let $x \in \mathbb{R}^n$, $x^T A x = x^T U \Lambda U^T x = (U^T x)^T \Lambda (U^T x) \geq 0$

Then it must be the case Λ is PSD too $\Rightarrow \lambda_i \geq 0 \; \blacksquare$