CS 229 Fall 2018 Problem Set #0

1) a) * $f(x) = \frac{1}{2} x^T A x + b^T x$

* $A^T = A$; $A \in \mathbb{R}^{n \times n}$
* $b \in \mathbb{R}^n$

Problem: $\nabla_x f(x) = ?$

$$\nabla_x f(x) = \frac{1}{2} \nabla_x x^T A x + \nabla_x b^T x$$

$$= \frac{1}{2} \nabla_x \left(\sum_{i=0}^{n} \sum_{j=0}^{n} A_{ij} x_i x_j \right) + b$$

$$= \frac{1}{2} \nabla_x \left[\sum_{i \neq k} \sum_{j \neq k} A_{ij} x_i x_j + \sum_{i \neq k} A_{ik} x_i x_k + \sum_{j \neq k} A_{kj} x_j x_k + A_{kk} x_k^2 \right] + b$$

$$= \frac{1}{2} \nabla_x \left[\sum_{i \neq k} A_{ik} x_i x_k + \sum_{j \neq k} A_{kj} x_j x_k + A_{kk} x_k^2 \right] + b$$

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

$$= \frac{1}{2} \sum_{i=0}^{n} A_{ik} x_i + \sum_{j=0}^{n} A_{kj} x_j + b$$

$$= Ax + b$$

b) * $g : \mathbb{R} \Rightarrow \mathbb{R}$ and g differentiable
* $h : \mathbb{R}^n \Rightarrow \mathbb{R}$ and h differentiable

* $f(x) = g(h(x))$

$\nabla f(x) = ?$

$$\nabla f(x) = \frac{\partial g}{\partial h} \times \nabla_x h(x)$$

c) * $f(x) = \frac{1}{2} x^T A x + b^T x$

* $A = A^T$

* $b \in \mathbb{R}^n$

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

$\Rightarrow \nabla_x (Ax + b)$ (question (a))

$\Rightarrow A^T = A$ (A is symmetric)

d) * $f(x) = g(a^T x)$; g continuously differentiable

* $a \in \mathbb{R}^n$

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

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

$$= g'(a^T x) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = g'(a^T x)\, a$$

• $\nabla^2 f(x) = \nabla \left(g'(a^T x)\, a \right)$

$$= \begin{pmatrix} \frac{g'(a^T x) a_1}{\partial x_1} & \cdots & \frac{g'(a^T x) a_n}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{g'(a^T x) a_1}{\partial x_n} & \cdots & \frac{g'(a^T x) a_n}{\partial x_n} \end{pmatrix}$$

$$= \begin{pmatrix} g''(a^T x) a_1^2 & \cdots & g''(a^T x) a_1 a_n \\ \vdots & \ddots & \vdots \\ g''(a^T x) a_1 a_n & \cdots & g''(a^T x) a_n^2 \end{pmatrix}$$

$$= g''(a^T x)\, a a^T$$

2) a) * $z \in \mathbb{R}^n$

* $A = zz^T$

Problem: Prove that A is positive semi-definite

Let $x \in \mathbb{R}^n$. A is positive semi-definite, if and only if $x^T A x \geq 0$

$\Rightarrow x^T A x = x^T z z^T x = \sum_{i=0}^{n} x_i z_i \wedge \sum_{j=0}^{n} x_j z_j$

$= \left(\sum_{i=0}^{n} x_i z_i \right)^2 \geq 0$

b) * $A = zz^T$
* $z \in \mathbb{R}^n$

Problem: What is the null-space and rank of A?

$Nul(A) = \{ x \in \mathbb{R}^n \}$ such that $Ax = 0$

$Ax = 0$

$\Rightarrow z z^T x = 0$

$\Rightarrow Nul(A) = \{ x \in \mathbb{R}^n \mid z^T x = 0 \}$

c) * $A \in \mathbb{R}^{n \times n}$ and A, PSD

* $B \in \mathbb{R}^{m \times n}$ arbitrary

Problem: Is BAB^T PSD?

$\forall y \in \mathbb{R}^m$, $B^T y \left(= (y^T B)^T\right) \in \mathbb{R}^n$

$\Rightarrow y^T B A B^T y$ *

$\Rightarrow (y^T B) A (B^T y)$

$\Rightarrow (B^T y)^T A (B^T y)$

Let $B^T y$ be $x \in \mathbb{R}^n$

$\Rightarrow x^T A x \geq 0$

Hence BAB^T is PSD

3) * $A \in \mathbb{R}^{n \times n}$ and diagonalisable

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

* $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$

* $T = [t_{(1)}, \dots, t_{(n)}]$

Promblem: Show $A t^{(i)} = \lambda_i t^{(i)}$

$\forall i \in \{1, \dots, n\}, u \mid u = T^{-1} t_{(i)},$

$u \in \{0,1\}^n : u_i = 1$ and $u_{k \neq i} = 0$

$\Rightarrow A t^{(i)} = T \Lambda T^{-1} t^{(i)}$

$= T \Lambda u^{(i)}$ where $u^{(i)}$ is a onehot vector of $\mathbb{R}^n$, taking the value 1 in the i^{th} position

$= T \lambda_i$

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

b) * $U \in \mathbb{R}^{n \times n}$ orthogonal

~~*COTAUTA~~

* $A = U \Lambda U^T$

* Problem: Prove $A u^{(i)} = \lambda_i u^{(i)}$

Like the last question, $\forall i \in \{0, \dots, n\}$

$U^T u_i$ gives a one-hot vector where the i^{th} coefficient takes the value 1 and all others take the value 0 (cause U orthogonal $\Leftrightarrow U^T U = I$)

$$\Rightarrow A u^{(i)} = U \Lambda U^T u^{(i)}$$

$$= U \lambda_i$$

$$= \cancel{\lambda_i u_i} \; \lambda_i u^{(i)}$$

c) Problem: If A is PSD, show that $\lambda_i(A) \geq 0$ for each i.

We saw in the previous question that

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

If A PSD $\Rightarrow u^{(i)T} A u^{(i)} \geq 0$

$$\Rightarrow u^{(i)T} \lambda_i u^{(i)} \geq 0$$

$$\Rightarrow \lambda_i \underbrace{\sum_{j=0}^{n} \left(u_j^{(i)}\right)^2}_{\geq 0} \geq 0$$

$\Rightarrow \lambda_i \geq 0$ for each i