Problem Set #0

1.

a) Sea $A \in \mathbb{R}^{n\times n}$ una matriz simétrica y sea $\nabla f(x)$ el gradiente de una función $f: \mathbb{R}^n \to \mathbb{R}$. Sea $f(x) = \frac{1}{2} x^T A x + b^T x$, $b \in \mathbb{R}^n$. ¿Qué es $\nabla f(x)$?

Tenemos $\nabla_x f(x) = \nabla_x \left(\frac{1}{2} x^T A x + b^T x \right)$. Sabemos que $df = (\nabla_x f)^T dx$, con ello:

$$df = \frac{1}{2}\left((dx)^T A x + x^T A dx \right) + b^T dx$$

Como $x^T A x$ es un escalar entonces $(x^T A x)^T = x^T A x$ y por ello,

$$df = \frac{1}{2}\left(((dx)^T A x)^T + x^T A dx \right) + b^T dx = \frac{1}{2}\left(x^T A^T dx + \right.$$

$$\left. + x^T A dx \right) + b^T dx = \frac{1}{2}\left(x^T A^T + x^T A \right) dx + b^T dx =$$

$$= \left[\frac{1}{2}\left(x^T A^T + x^T A \right) + b^T \right] dx = \left[\frac{1}{2}\left(A^T x + A x \right)^T + b^T \right] dx$$

$$= \left[\frac{1}{2}\left(A^T x + A x \right) + b \right]^T dx = \left[\frac{1}{2} \cdot 2 A x + b \right]^T dx = (Ax + b)^T dx$$

$$\Longrightarrow \nabla_x f(x) = Ax + b$$

b) Sea $f(x) = g(h(x))$, donde $g: \mathbb{R} \to \mathbb{R}$ y $h: \mathbb{R}^n \to \mathbb{R}$, ¿Qué es $\nabla_x f(x)$?

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

$$\Longrightarrow \nabla_x f(x) = \nabla_x g(h(x)) = g'(h(x))\, \nabla_x h(x)$$

c) Sea $f(x) = \frac{1}{2} x^T A x + b^T x$, ¿Qué es $\nabla^2 f(x)$?

De a) sabemos $\nabla f(x) = Ax + b$. Por lo tanto, $\nabla^2 f = \nabla(\nabla f) =$

$= \nabla(Ax+b) = \nabla g(x)$, siendo $g(x) = Ax+b$

Como en a)

$$dg = A\,dx \longrightarrow \nabla g(x) = A^T = A$$

Por lo tanto, $\nabla^2 f(x) = \nabla^2\left(\frac{1}{2} x^T A x + b^T x\right) = A$

d) Sea $f(x) = g(a^T x)$, ¿Qué es $\nabla f(x)$ y $\nabla^2 f(x)$?

$$\frac{\partial f}{\partial x_i} = \frac{\partial g(a^T x)}{\partial x_i} = \frac{\partial g}{\partial (a^T x)} \frac{\partial a^T x}{\partial x_i} = g'(a^T x) \frac{\partial (a^T x)}{\partial x_i}$$

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

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

2.

a) Sea $z \in \mathbb{R}^n$. Demostrar que $A = zz^T$ es positiva semidefinida.

Primero vemos si $A^T = A$: $A^T = (zz^T)^T = zz^T = A$ ✓

Segundo veamos si $x^T A x \geq 0 \ \forall x \in \mathbb{R}^n$:

$$x^T A x = x^T z z^T x$$

Como $x^T z$ es un escalar entonces $z^T x$ es el mismo escalar. Si llamamos λ a este escalar entonces

$$x^T A x = \lambda \lambda^T = \lambda^2 \geq 0$$

b) Sea $z \in \mathbb{R}^n$ un vector distinto de $\vec{0}$. Sea $A = zz^T$. ¿Cual es el espacio nulo de A? ¿Cual es el rango de A?

El espacio nulo es el conjunto x de vectores que satisfacen:

$$\text{Ker}(A) = \{ x \in \mathbb{R}^n : Ax = 0 \}$$

$$\Rightarrow Ax = zz^T x = 0, \ z(z^T x) = 0$$

Como $z \neq 0$, la unica solución es: $z^T x = 0$

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

El rango de A es 1.

c) Sea $A \in \mathbb{R}^{n \times n}$ (PSD) y sea $B \in \mathbb{R}^{m \times n}$. ¿Es BAB^T PSD?

Veamos la simetría: $(BAB^T)^T = BA^T B^T = BAB^T$ ✓

Veamos si $x^T BAB^T x \geq 0,\ \forall x \in \mathbb{R}^m$:

$$x^T BAB^T x = \underbrace{(\overbrace{x^T B}^{= v^T})}_{1\times n} A \underbrace{(B^T x)}_{n \times 1} = v^T A v \geq 0 \qquad \checkmark$$

3.

a) Sea $A \in \mathbb{R}^{n\times n}$ diagonalizable $(A = T\Lambda T^{-1})$. Sean $t^{(i)}$ las columnas de T, tal que $T = [t^{(1)}, \dots, t^{(n)}]$, donde $t^{(i)} \in \mathbb{R}^n$. Demostrar $At^{(i)} = \lambda_i t^{(i)}$.

Sabemos $A = T\Lambda T^{-1}$, entonces

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

Que para la fila K de A es: $(a_{Kj})\, t_j^{(K)} = t_j^{(K)} \lambda_K$

b) Sea $A = A^T$. Demostrar que si $U = [u^{(1)}, \dots, u^{(n)}]$ es ortogonal, donde $u^{(i)} \in \mathbb{R}^n$ y $A = U\Lambda U^T$, entonces $u^{(i)}$ es un autovalor de A y $Au^{(i)} = \lambda_i u^{(i)}$, donde $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$

Partimos de $A = U\Lambda U^T$, como U es ortogonal

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

Si tomamos un $u^{(i)}$:

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

c) Demostrar que si A es PSD, entonces $\lambda_i(A) \geq 0$ para cada i.

Si A es PSD entonces:

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

$$(t^{(i)})^T A t^{(i)} = \lambda_i (t^{(i)})^T t^{(i)} = \lambda_i \|t^{(i)}\|_2^2 = \lambda_i \geq 0$$