II Gradientes:

1-) Sea la matriz $A \in \mathbb{R}^{n \times n}$ simétrica y sea el vector $x \in \mathbb{R}^n$

- Encuentre $\nabla_{\underline{x}} f(\underline{x})$ con $f(\underline{x}) = \frac{1}{2} \cdot \underline{x}^T \cdot A \cdot \underline{x} + \underline{b}^T \underline{x}$, Utilizando las propiedades de gradientes y trazas vistas en clase (10 pts).

$$\cdot \; f(\underline{x}) = \frac{1}{2} \cdot \underbrace{\underbrace{\underline{x}^T \cdot A}_{1 \times n} \cdot \underline{x}}_{1\times 1 = \text{escalar}} + \underbrace{\underline{b}^T \underline{x}}_{1 \times 1 = \text{esalar}}$$

- $A^{n \times n}$
$x^{n \times 1}$
$b^{n \times 1}$

$$\therefore \nabla_{\underline{x}} f(\underline{x}) = \frac{1}{2} \cdot \nabla_{\underline{x}} tr(\underline{x}^T A \underline{x}) + tr(\underline{b}^T \underline{x})$$

} "Matrix Codebook"

$\downarrow$ ec. 108 $\downarrow$ $\downarrow$ ec. 101 $\downarrow$

$$= \frac{1}{2} \cdot (A\underline{x} + A'\underline{x}) + \underline{b} \; ; \; \#\# \; A \text{ es simétrico} \Rightarrow A = A'$$

$$\boxed{\therefore \nabla_{\underline{x}} f(x) = A\underline{x} + \underline{b}}$$

5-) Demuestre que para $f(\underline{\theta}) = \frac{1}{2} \| X\underline{\theta} - \underline{y} \|^2$ | $\underline{\theta} \in \mathbb{R}^n$, $X \in \mathbb{R}^{m\times n}$

$\rightarrow \nabla_{\underline{\theta}} f(\underline{\theta}) = X^T (X\underline{\theta} - \underline{y})$

* nota: $\|v\|^2 = v^T \cdot v$

$\therefore f(\underline{\theta}) = \frac{1}{2} (X\underline{\theta} - \underline{y})^T (X\underline{\theta} - \underline{y})$

$= \frac{1}{2} [(X\underline{\theta})^T - \underline{y}^T][X\underline{\theta} - \underline{y}]$; $(A+B)^T = A^T + B^T$

$= \frac{1}{2} [\underline{\theta}^T X^T - \underline{y}^T][X\underline{\theta} - \underline{y}]$; $(AB)^T = B^T A^T$

$= \frac{1}{2} [\underline{\theta}^T X^T X \underline{\theta} - \underbrace{\underline{\theta}^T X^T \underline{y}} - \underline{y}^T X \underline{\theta} - \cancel{\underline{y}^T \underline{y}}]$ $\rightarrow 0$; a la hora de derivar respecto a $\underline{\theta}$ para encontrar el gradiente esa expresión será igual a 0 ya que no hay terminos dependientes de $\underline{\theta}$

$C^T B^T A^T = (ABC)^T$

$\Rightarrow \underline{\theta}^T X^T \underline{y} = \underbrace{(\underline{y}^T X \underline{\theta})^T}$

y además esta expresión se resume a un escalar y $s = s^T \Rightarrow (\underline{y}^T X \underline{\theta})^T = (\underline{y}^T X \underline{\theta})$

$\therefore f(\underline{\theta}) = \frac{1}{2} [\underline{\theta}^T X^T X \underline{\theta} - 2(\underline{y}^T X \underline{\theta})]$

$\therefore \nabla_{\underline{\theta}} f(\underline{\theta}) = \frac{1}{2} \underbrace{tr(\underline{\theta}^T X^T X \underline{\theta})} - \underbrace{tr(\underline{y}^T X \underline{\theta})}$

usando: $\nabla_{A^T} tr\{ABA^T C\} = B^T A^T C^T + B A^T C$

con: $A = \theta^T$
$B = X^T X$
$A^T = \theta$
$C = I$

$\nabla_X tr(AXB) = A^T B^T$

con $A = y^T X$
$X = \underline{\theta}$
$B = I$

$$\boxed{\therefore \nabla_{\underline{\theta}} f(\underline{\theta}) = X^T X \underline{\theta} - \underline{X}^T y = X^T (X\underline{\theta} - y)}$$

Scanned with CamScanner

6-) Encuentre ahora el gradiente $\nabla_X f(X)$ con:

$$\underbrace{f(X) = \tfrac{1}{2}\|X\underline{\theta} - \underline{y}\|^2}_{}$$

$X \in \mathbb{R}^{m \times n}$

$\underline{\theta} \in \mathbb{R}^n$

partiendo de la forma desarrollada de esta expresión.

$$\rightarrow f(X) = \tfrac{1}{2}[\underline{\theta}^T X^T X \underline{\theta} - \underbrace{\underline{\theta}^T X^T \underline{y}}_{} - \underline{y}^T X \underline{\theta} + \cancel{\underline{y}^T \underline{y}}]$$

0, por no haber elementos dependientes de X

$$C^T B^T A^T = (ABC)^T$$

$$\therefore \underline{\theta}^T X^T \underline{y} = \underbrace{(\underline{y}^T X \underline{\theta})^T = \underline{y}^T X \underline{\theta}}_{\text{por ser un escalar}}$$

$$\therefore f(X) = \tfrac{1}{2}(\underline{\theta}^T X^T X \underline{\theta}) - \underline{y}^T X \underline{\theta}$$

$$\nabla_X f(X) = \tfrac{1}{2}\underbrace{\nabla_X tr(X\underline{\theta}\underline{\theta}^T X^T)}_{} - \underbrace{\nabla_X tr(\underline{y}^T X \underline{\theta})}_{}$$

$$\nabla_X tr(X B X^T) = X B^T + X B$$

con $X = X$

$B = \underline{\theta}\underline{\theta}^T$

$$\nabla_X tr(A X B) = A^T B^T$$

con $A = \underline{y}^T$

$X = X$

$B = \underline{\theta}$

$$= \tfrac{1}{2}(X[\underline{\theta}\underline{\theta}^T]^T + X[\underline{\theta}\underline{\theta}^T]) - \underline{y}\underline{\theta}^T$$

$$\boxed{\therefore \nabla_X f(X) = X\underline{\theta}\underline{\theta}^T - \underline{y}\underline{\theta}^T = (X\underline{\theta} - \underline{y})\underline{\theta}^T}$$

Scanned with CamScanner

7) Encuentre $\nabla_x f(x)$ con $\underbrace{f(x) = g(h(x))}$, $g: R \to R \wedge h: R^n \to R$

$f(x)$ es una función compuesta por lo que se debe usar "regla de la cadena" para obtener el máximo cambio en una dirección.

$$\boxed{\therefore \nabla_x f(x) = g'(h(x)) \cdot h'(x)}$$

8-) Ahora $\underbrace{f(x) = g(a^T x)}$ con $g: R \to R \wedge a \in R^n$

función compuesta igual que en el caso anterior; $f(x) = g(h(x))$

$* \; a^T x =$ escalar $\Rightarrow tr(a^T x) = a$

$* \; tr(AXB) = A^T$, con $A = a^T$, $X = x \wedge B = I$

$$\boxed{\therefore \nabla_x f(x) = g'(a^T x) \cdot (a^T x)' = g'(a^T x) \cdot a}$$

III Matrices positivas definidas:

1·) Demuestre que $A = \underline{z}\underline{z}^T$ es positiva semidefinida si $\underline{z} \in R^n$

Si $\forall \underline{x}$ $\underline{x}^T A \underline{x} \geq 0 \Rightarrow A$ positiva semidefinida

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

$\boxed{\therefore A \text{ es positiva definida } \forall \underline{x}}$

2-) Sea $\underline{z} \in R^n$ no nulo $\wedge$ $A = zz^T$ Encuentre el espacio nulo de A y su rango.

$$\left.\begin{array}{l} \cdot\ z = n \times 1 \\ \cdot\ A = n \times n \end{array}\right\} rg(A) \leq \min(n, n) \Rightarrow rg(A) = n$$

$\rightarrow$ Los espacios nulos surgen cuando $z^T x$ son ortogonales.