Informe: Optimizacion Convexa Taller No. 2

Estudiante: Ruben Rodriguez
Código: 21819696
Grupo: O1
Escuela de Ingeniería de Sistemas e Informática
Universidad Industrial de Santander

14 de mayo de 2022

1. Introducción

Development of exercises on notation, Gradient and demonstrations.

2. Desarrollo

1. Basic notation. In front of each symbol write its meaning
   - R : denotes the set of real numbers.
   - R_+ : denotes the set of positive real numbers.
   - R_{++} : denotes the set of positive real numbers except zero.
   - R_+^{mxn}: denotes the set of matrices of positive real numbers of size $m * n$.
   - N_+^k: denotes the set of vectors of positive natural numbers of size k.
   - S_+^k : denotes the set of symmetric positive semi-defined matrices.
   - $f : R^p \rightarrow R^q$: denotes a function with domain R^p and range in R^q.
2. Calculate the $\nabla f(x)$ respect to x of the following function
   - $||y - Hx||_2^2$: (Demonstration in annex; 1)
   - $||y - Hx||_2$: (Demonstration in annex; 1)
   - $||y - Hx||_1$: (Demonstration in annex; 2)

3. For the special case where A is a symmetric matrix and

$$\alpha = x^T A x$$

where $x \in R^n$ and $A \in R^{nxn}$. Then, calculate $\frac{\partial \alpha}{\partial x}$. Differs the result if A is not symmetric?.

Answ: If A is not symmetric the answer change. (Demonstration in annex;3)

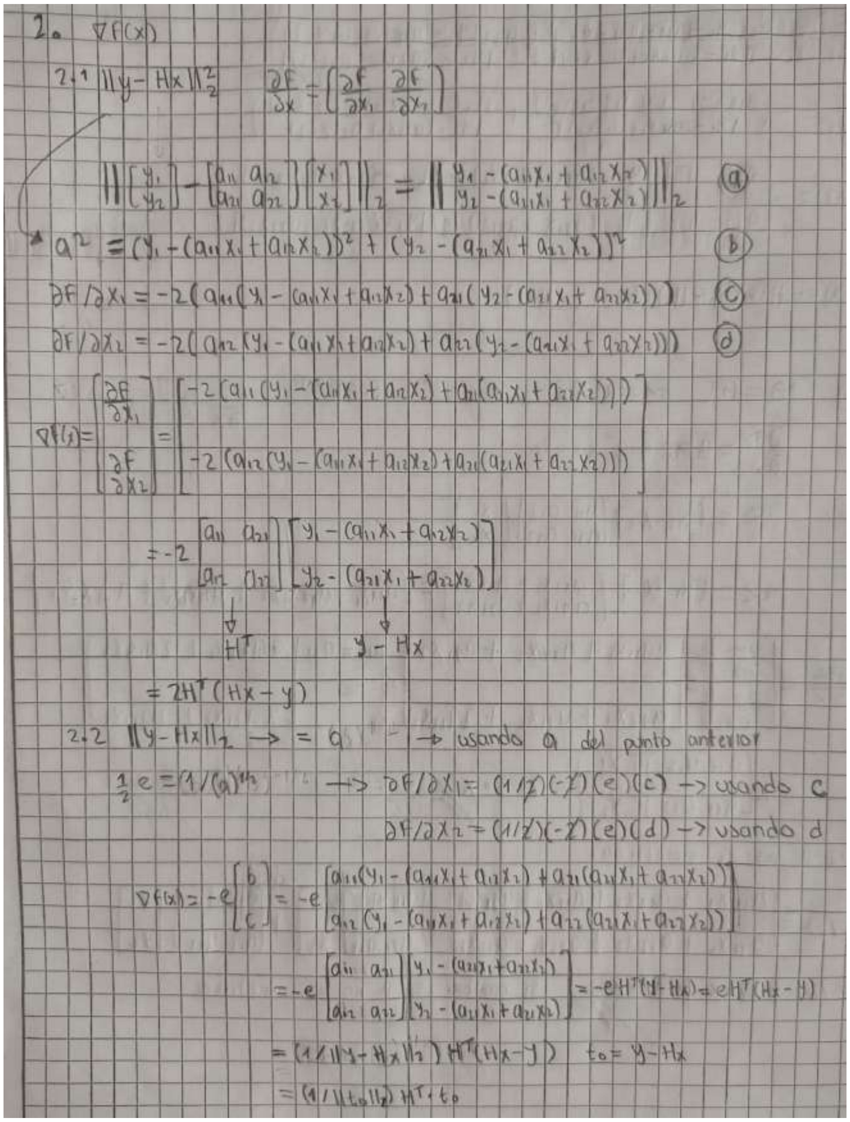
4. Let $A(\alpha)$ be a $m * n$ matrix whose elements are functions of the scalar parameter α. Then the derivative of the matrix $A(\alpha)$ with respect to the scalar parameter α is the $m * n$ matrix of element-by-element derivatives:

$$\frac{\partial \mathbf{A}(\alpha)}{\partial \alpha} = \begin{bmatrix} \frac{\partial a_{11}}{\partial \alpha} & \frac{\partial a_{12}}{\partial \alpha} & \dots & \frac{\partial a_{1n}}{\partial \alpha} \\ \frac{\partial a_{21}}{\partial \alpha} & \frac{\partial a_{22}}{\partial \alpha} & \dots & \frac{\partial a_{2n}}{\partial \alpha} \\ \vdots & \vdots & & \vdots \\ \frac{\partial a_{ma}}{\partial \alpha} & \frac{\partial a_{m2}}{\partial \alpha} & \dots & \frac{\partial a_{man}}{\partial \alpha} \end{bmatrix}$$

Then let $A(\alpha)$ be a nonsingular, $m * m$ matrix whose elements are functions of the scalar parameter α. Then show what:

$\frac{\partial A(\alpha)^{-1}}{\partial \alpha} = -A(\alpha)^{-1} \frac{\partial A(\alpha)}{\partial \alpha} A(\alpha)^{-1}$ (Demonstration in annex; 4)

3. Anexos



2. $\nabla f(x)$

2.1 $\|y - Hx\|_2^2$ $\quad \frac{\partial f}{\partial x} = \left[\frac{\partial f}{\partial x_1} \; \frac{\partial f}{\partial x_2}\right]$

$$\left\| \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} - \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \right\|_2 = \left\| \begin{matrix} y_1 - (a_{11}x_1 + a_{12}x_2) \\ y_2 - (a_{21}x_1 + a_{22}x_2) \end{matrix} \right\|_2 \quad (a)$$

$$a^2 = (y_1 - (a_{11}x_1 + a_{12}x_2))^2 + (y_2 - (a_{21}x_1 + a_{22}x_2))^2 \quad (b)$$

$$\partial f / \partial x_1 = -2(a_{11}(y_1 - (a_{11}x_1 + a_{12}x_2) + a_{21}(y_2 - (a_{21}x_1 + a_{22}x_2))) \quad (c)$$

$$\partial f / \partial x_2 = -2(a_{12}(y_1 - (a_{11}x_1 + a_{12}x_2) + a_{22}(y_2 - (a_{21}x_1 + a_{22}x_2))) \quad (d)$$

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} -2(a_{11}(y_1 - (a_{11}x_1 + a_{12}x_2) + a_{21}(a_{21}x_1 + a_{22}x_2))) \\ -2(a_{12}(y_1 - (a_{11}x_1 + a_{12}x_2) + a_{21}(a_{21}x_1 + a_{22}x_2))) \end{bmatrix}$$

$$= -2 \underbrace{\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}}_{H^T} \underbrace{\begin{bmatrix} y_1 - (a_{11}x_1 + a_{12}x_2) \\ y_2 - (a_{21}x_1 + a_{22}x_2) \end{bmatrix}}_{y - Hx}$$

$$= 2H^T(Hx - y)$$

2.2 $\|y - Hx\|_2 \rightarrow = a$ $\rightarrow$ usando a del punto anterior

$\frac{1}{2} e = 1/(a)^{1/2}$ $\longrightarrow \partial f/\partial x_1 = (1/\cancel{2})(-\cancel{2})(e)(c) \rightarrow$ usando c

$\partial f/\partial x_2 = (1/\cancel{2})(-\cancel{2})(e)(d) \rightarrow$ usando d

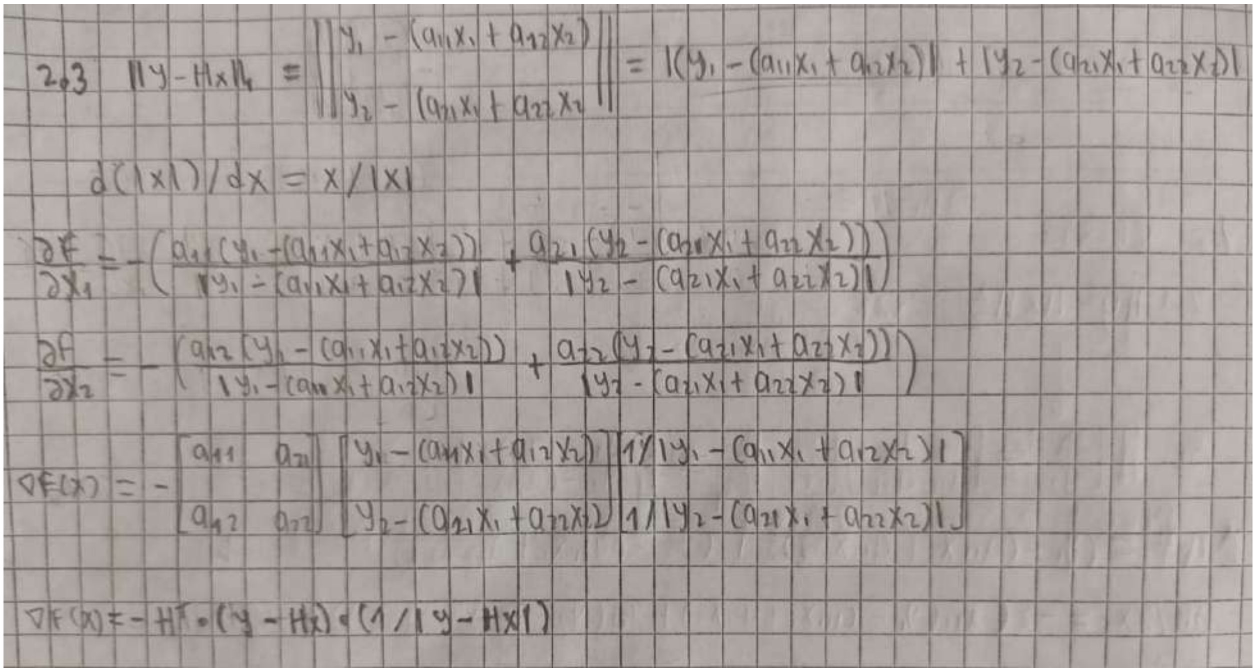
$$\nabla f(x) = -e \begin{bmatrix} b \\ c \end{bmatrix} = -e \begin{bmatrix} a_{11}(y_1 - (a_{11}x_1 + a_{12}x_2) + a_{21}(a_{21}x_1 + a_{22}x_2)) \\ a_{12}(y_1 - (a_{11}x_1 + a_{12}x_2) + a_{22}(a_{21}x_1 + a_{22}x_2)) \end{bmatrix}$$

$$= -e \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 - (a_{11}x_1 + a_{12}x_2) \\ y_2 - (a_{21}x_1 + a_{22}x_2) \end{bmatrix} = -eH^T(y - Hx) = eH^T(Hx - y)$$

$$= (1/\|y - Hx\|_2) H^T(Hx - y) \qquad t_0 = y - Hx$$

$$= (1/\|t_0\|_2) H^T \cdot t_0$$

Figura 1: Demostracion punto 2.1 y 2.2



2.3 $\|y - Hx\|_1 = \left\| \begin{matrix} y_1 - (a_{11}x_1 + a_{12}x_2) \\ y_2 - (a_{21}x_1 + a_{22}x_2) \end{matrix} \right\| = |(y_1 - (a_{11}x_1 + a_{12}x_2)| + |y_2 - (a_{21}x_1 + a_{22}x_2)|$

$d(|x|)/dx = x/|x|$

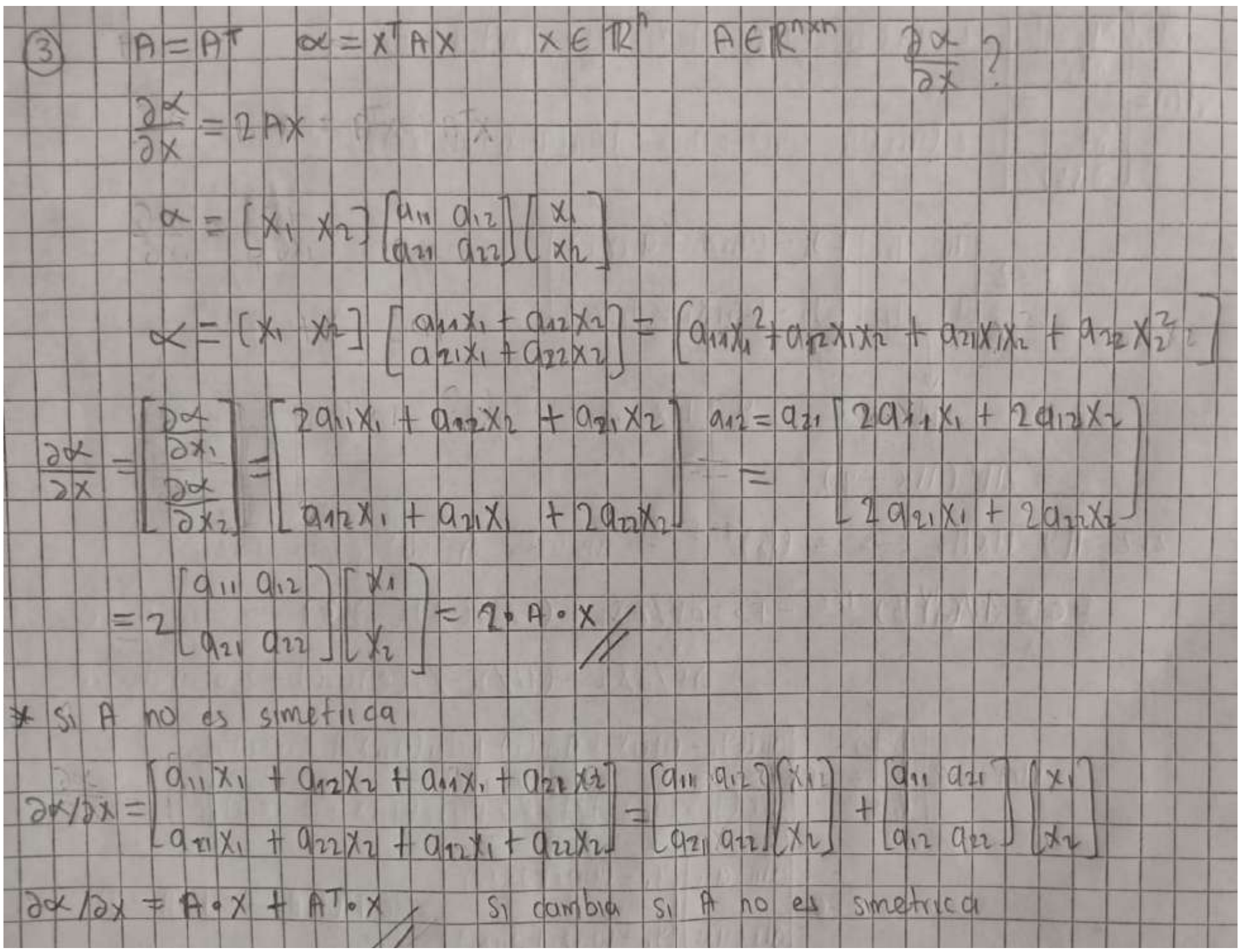
$$\frac{\partial f}{\partial x_1} = -\left(\frac{a_{11}(y_1 - (a_{11}x_1 + a_{12}x_2))}{|y_1 - (a_{11}x_1 + a_{12}x_2)|} + \frac{a_{21}(y_2 - (a_{21}x_1 + a_{22}x_2))}{|y_2 - (a_{21}x_1 + a_{22}x_2)|} \right)$$

$$\frac{\partial f}{\partial x_2} = -\left(\frac{a_{12}(y_1 - (a_{11}x_1 + a_{12}x_2))}{|y_1 - (a_{11}x_1 + a_{12}x_2)|} + \frac{a_{22}(y_2 - (a_{21}x_1 + a_{22}x_2))}{|y_2 - (a_{21}x_1 + a_{22}x_2)|} \right)$$

$$\nabla f(x) = -\begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} y_1 - (a_{11}x_1 + a_{12}x_2) \\ y_2 - (a_{21}x_1 + a_{22}x_2) \end{bmatrix} \begin{bmatrix} 1/|y_1 - (a_{11}x_1 + a_{12}x_2)| \\ 1/|y_2 - (a_{21}x_1 + a_{22}x_2)| \end{bmatrix}$$

$$\nabla f(x) = -H^T \cdot (y - Hx) \cdot (1/|y - Hx|)$$

Figura 2: Demostracion punto 2.3



3) $A=A^T$ $\quad \alpha = x^T A x$ $\quad x \in \mathbb{R}^n$ $\quad A \in \mathbb{R}^{n\times n}$ $\quad \frac{\partial \alpha}{\partial x}$?

$$\frac{\partial \alpha}{\partial x} = 2Ax$$

$$\alpha = [x_1 \ x_2] \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\alpha = [x_1 \ x_2] \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix} = [a_{11}x_1^2 + a_{12}x_1x_2 + a_{21}x_1x_2 + a_{22}x_2^2]$$

$$\frac{\partial \alpha}{\partial x} = \begin{bmatrix} \frac{\partial \alpha}{\partial x_1} \\ \frac{\partial \alpha}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 2a_{11}x_1 + a_{12}x_2 + a_{21}x_2 \\ a_{12}x_1 + a_{21}x_1 + 2a_{22}x_2 \end{bmatrix} \quad a_{12} = a_{21} \quad = \begin{bmatrix} 2a_{11}x_1 + 2a_{12}x_2 \\ 2a_{21}x_1 + 2a_{22}x_2 \end{bmatrix}$$

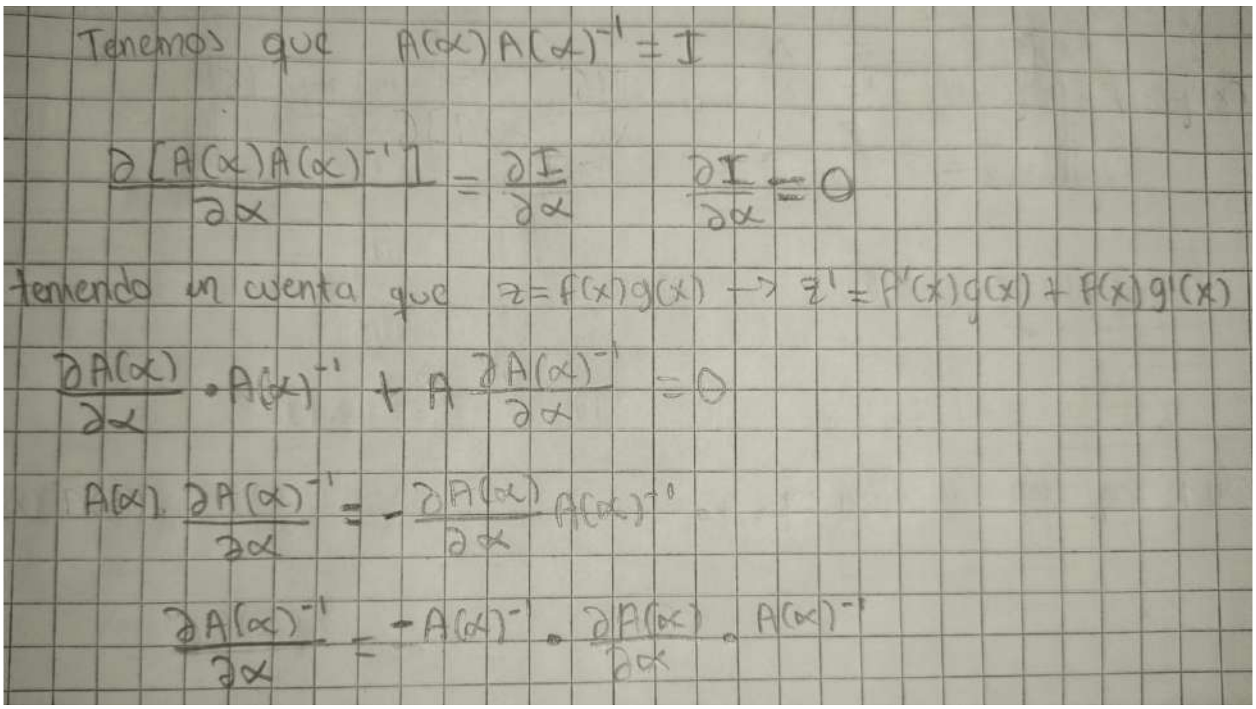
$$= 2 \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2 \cdot A \cdot x$$

* Si A no es simetrica

$$\partial\alpha/\partial x = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{11}x_1 + a_{22}x_2 \\ a_{21}x_1 + a_{22}x_2 + a_{12}x_1 + a_{22}x_2 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$\partial\alpha/\partial x = A \cdot x + A^T \cdot x$$ Si cambia si A no es simetrica

Figura 3: Demostracion punto 3



Tenemos que $A(\alpha)A(\alpha)^{-1} = I$

$$\frac{\partial[A(\alpha)A(\alpha)^{-1}]}{\partial\alpha} = \frac{\partial I}{\partial\alpha} \qquad \frac{\partial I}{\partial\alpha} = 0$$

teniendo en cuenta que $z = f(x)g(x) \rightarrow z' = f'(x)g(x) + f(x)g'(x)$

$$\frac{\partial A(\alpha)}{\partial\alpha}\cdot A(\alpha)^{-1} + A\frac{\partial A(\alpha)^{-1}}{\partial\alpha} = 0$$

$$A(\alpha)\cdot\frac{\partial A(\alpha)^{-1}}{\partial\alpha} = -\frac{\partial A(\alpha)}{\partial\alpha}A(\alpha)^{-1}$$

$$\frac{\partial A(\alpha)^{-1}}{\partial\alpha} = -A(\alpha)^{-1}\cdot\frac{\partial A(\alpha)}{\partial\alpha}\cdot A(\alpha)^{-1}$$

Figura 4: Demostracion punto 4