Assignment 0

Problem 1 : Definitions

a) $f : R^n \rightarrow R$

$x \in R^n$

$$\nabla f(x) = \left[\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right]^T$$

b) $f : R^n \rightarrow R^m$

$x \in R^n$

$$f(x) = \left[f_1(x), f_2(x), \ldots, f_m(x)\right]^T$$

$$J = \frac{\partial f}{\partial x} \equiv \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

c) $\left.\begin{matrix} f : R^n \rightarrow R^m \\ x \in R^n \end{matrix}\right\} \Rightarrow$ Size of ∇f will be

$1 \times n = \underline{n}$

d) $\left.\begin{matrix} f : R^n \rightarrow R^m \\ x \in R^n \end{matrix}\right\} \rightarrow$ Size of $\frac{\partial f}{\partial x}$ will be

$\underline{m \times n}$

Problem 2: Linear

$\underline{f}(\underline{x}) = A\underline{x}$

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

$$\Longrightarrow \underline{f}(\underline{x}) = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$$

a) $$\frac{\partial \underline{f}}{\partial \underline{x}} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \underline{\underline{A}}$$

Since $\underline{f}: \mathbb{R}^2 \to \mathbb{R}^2$ this is the Jacobian of $\underline{f}(\underline{x})$.

b) $x \in \mathbb{R}^n$

A $m \times n$

$$\frac{\partial A\underline{x}}{\partial \underline{x}} = \underline{\underline{A}}$$

Problem 3: Nonlinear

$$f(x, y) = x^T G y$$

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad G = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

a) The dimension of $f(x,y)$ is $1 \times 1 = \underline{\underline{1}}$

$$\underline{\underline{\nabla_x f(x,y) = \left[\frac{\partial f}{\partial x}\right]^T}}$$

b) $$\nabla_x f(x,y) = \left[\frac{\partial f}{\partial x}\right]^T$$

$$= \left[\frac{\partial}{\partial x}\left(\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}\right)\right]^T$$

$$= \left[\frac{\partial}{\partial x}\left(\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} g_{11}y_1 + g_{12}y_2 + g_{13}y_3 \\ g_{21}y_1 + g_{22}y_2 + g_{23}y_3 \end{bmatrix}\right)\right]^T$$

$$= \left[\frac{\partial}{\partial x}\left((g_{11}y_1 + g_{12}y_2 + g_{13}y_3)x_1 + (g_{21}y_1 + g_{22}y_2 + g_{23}y_3)x_2\right)\right]^T$$

$$= \begin{bmatrix} g_{11}y_1 + g_{12}y_2 + g_{13}y_3 \\ g_{21}y_1 + g_{22}y_2 + g_{23}y_3 \end{bmatrix} = \underline{\underline{Gy}}$$

c) $\nabla_y f(x,y) = \left[\frac{\partial f}{\partial y}\right]^T$

$f(x,y)$ is a scalar function

$\longrightarrow \nabla_y (c^T y) = c$

$c^T = x^T G$

$\Longrightarrow c = (x^T G)^T = G^T x$

$\nabla_y f(x,y) = \underline{\underline{G^T x}}$

d) $f(x) = x^T H x$

$x \in R^n$

$H \in R^{n \times n}$

Introducing the vector $y = x$, the function f can be rewritten as:

$f(x) = f(x,y) = x^T H y$

$\nabla f(x) = \nabla_x f(x,y) + \nabla_y f(x,y)$

$= Hy + H^T x$

$y = x$

$\Longrightarrow \underline{\underline{\nabla f(x) = Hx + H^T x}}$

$H = H^T$

$$\longrightarrow \nabla f(x) = Hx + Hx = \underline{\underline{2Hx}}$$

Problem 4: Common case

$$\mathcal{L}(x, \lambda, \mu) = x^T G x + \lambda^T (Cx - d) + \mu^T (Ex - h)$$

a) $$\nabla_x \mathcal{L}(x, \lambda, \mu) = \nabla_x (x^T G x) + \nabla_x (\lambda^T (Cx - d)) + \nabla_x (\mu^T (Ex - h))$$

$$\nabla_x (x^T G x) = Gx + G^T x$$

$$\nabla_x (\lambda^T (Cx - d)) = \nabla_x (\lambda^T C x - \lambda^T d) = \nabla_x (\lambda^T C x) = (\lambda^T C)^T = C^T \lambda$$

$$\nabla_x (\mu^T (Ex - h)) = \nabla_x (\mu^T E x - \mu^T h) = \nabla_x (\mu^T E x) = (\mu^T E)^T = E^T \mu$$

$$\Longrightarrow \nabla_x \mathcal{L}(x, \lambda, \mu) = Gx + G^T x + C^T \lambda + E^T \mu$$

$G = G^T$

$$\longrightarrow \underline{\underline{\nabla_x \mathcal{L}(x, \lambda, \mu) = 2Gx + C^T \lambda + E^T \mu}}$$

b) $\nabla_\mu \mathcal{L}(x, \lambda, \mu) = \nabla_\mu (x^T G x) + \nabla_\mu (\lambda^T (Cx - d))$
$+ \nabla_\mu (\mu^T (Ex - h))$

$\nabla_\mu (x^T G x) = 0$

$\nabla_\mu (\lambda^T (Cx - d)) = 0$

$\nabla_\mu (\mu^T (Ex - h)) = \nabla_\mu (\mu^T E x) - \nabla_\mu (\mu^T h)$
$= Ex - h$

$\longrightarrow \underline{\underline{\nabla_\mu \mathcal{L}(x, \lambda, \mu) = Ex - h}}$

c) $\nabla_\lambda \mathcal{L}(x, \lambda, \mu) = \nabla_\lambda (x^T G x) + \nabla_\lambda (\lambda^T (Cx - d))$
$+ \nabla_\lambda (\mu^T (Ex - h))$

$\nabla_\lambda (x^T G x) = 0$

$\nabla_\lambda (\lambda^T (Cx - d)) = \nabla_\lambda (\lambda^T C x) - \nabla_\lambda (\lambda^T d)$
$= Cx - d$

$\nabla_\lambda (\mu^T (Ex - h)) = 0$

$\Longrightarrow \underline{\underline{\nabla_\lambda \mathcal{L}(x, \lambda, \mu) = Cx - d}}$