Homework 1 (MATH7243)

1) Assume $\vec{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $\vec{b} \in \mathbb{R}^m$. Let $f(\vec{x}) = \vec{b}^T A \vec{x}$. Find ∇f.

$$= \vec{b}^T (A\vec{x})$$

$$A\vec{x} = \begin{bmatrix} \sum_{i=1}^{n} \vec{x}_{1i} A_i \\ \vdots \\ \sum_{i=1}^{n} \vec{x}_{ni} A_i \end{bmatrix} \Rightarrow f(\vec{x}) = \vec{b}^T \begin{bmatrix} \sum_{i=1}^{n} \vec{x}_{1i} A_i \\ \vdots \\ \sum_{i=1}^{n} \vec{x}_{ni} A_i \end{bmatrix}$$

$$\Rightarrow f(\vec{x}) = \sum_j \sum_i \vec{b}^T \vec{x}_{ij} A_i$$

$$\frac{d\, f(\vec{x})}{d(\vec{x})} = \begin{bmatrix} \frac{df}{dx_{11}} & \frac{df}{dx_{12}} & \cdots & \frac{df}{dx_{1n}} \\ \vdots & & & \vdots \\ \frac{df}{dx_{m1}} & \frac{df}{dx_{m2}} & \cdots & \frac{df}{dx_{mn}} \end{bmatrix} = \begin{bmatrix} b_1^T A_1 & b_1^T A_2 & \cdots \\ \vdots & & \\ b_m^T A_1 & & b_m^T A_n \end{bmatrix}$$

2) Assume $\vec{x} \in \mathbb{R}^n$. Find $\frac{d\, \vec{x}^T x}{dx}$

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad x^T = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}$$

$$\vec{x}^T \cdot \vec{x} = \sum_{i=1}^{n} x_i x_i = \sum_{i=1}^{n} x_i^2 =$$

$$\frac{d(\vec{x}^T \cdot \vec{x})}{dx} = 2x_i \Rightarrow \frac{d x^T x}{dx} = (2x_1, \ldots, 2x_n) = 2x^T$$

3) Assume $\vec{x}$ and $\vec{a} \in \mathbb{R}^n$. Find $\frac{d}{dx}(\vec{x}^T.\vec{a})^2$

$$\vec{x}^T = [x_1^T \dots x_n^T] \quad , \quad \vec{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$

$$(\vec{x}^T.\vec{a})^2 = \left(\sum_{i=1}^{n} x_i a_i\right)^2 \Rightarrow \frac{d(\vec{x}^T.\vec{a})^2}{d\vec{x}}$$

Chain Rule : Putting a bar to show constant

$$\Rightarrow \frac{d}{d\vec{x}}\left(\vec{x}^{T^2}.\vec{a}^2\right) = \frac{d(\vec{x}^{T^2}.\underline{\vec{a}})}{d\vec{x}} + \cancel{\frac{d(\underline{\vec{x}}^T.a^2)}{d\vec{x}}}^{=0}$$

$$= 2x^T a^2 + 0 = \underline{2x^T a^2}$$

4) $\vec{x} : \mathbb{R}^n \to \mathbb{R}^m$ is a map sending $z \in \mathbb{R}^n$ to $\vec{x}(\vec{z}) \in \mathbb{R}^m$
$\vec{y} : \mathbb{R}^n \to \mathbb{R}^m$ and A is $m \times m$

$$\frac{d(\vec{y}^T A\vec{x})}{d\vec{z}} = \frac{d\vec{y}}{d\vec{z}} A\vec{x} + \frac{d\vec{x}}{d\vec{z}} A^T \vec{y}$$

$$\vec{u} = \vec{y}^T A\vec{x} \quad , \quad \vec{v}^T = \vec{y}^T.A$$

$$\Rightarrow \vec{u} = \vec{v}^T.\vec{x}$$

Partial derivative w.r.t z : $\frac{d\vec{u}}{d\vec{z}} = \frac{d(\vec{v}^T.\vec{x})}{dz}$

Putting bars to show constant

$$\Rightarrow \frac{d\vec{u}}{d\vec{z}} = \frac{d(\vec{v}^T.\underline{\vec{x}})}{d\vec{z}} + \frac{d(\underline{\vec{v}}^T.\vec{x})}{d\vec{z}}$$

$$\frac{d\vec{u}}{d\vec{z}} = \frac{d\vec{u}}{d\vec{y}}.\frac{d\vec{y}}{d\vec{z}} + \frac{d\vec{u}}{d\vec{x}}.\frac{d\vec{x}}{d\vec{z}}$$

$$\frac{d(\vec{y}^T A\vec{x})}{d\vec{z}} = \frac{d\vec{y}}{d\vec{z}} A\vec{x} + \frac{d\vec{x}}{d\vec{z}} A^T.\vec{y}$$

5) $A(x) : \mathbb{R} \to \mathbb{R}^{n \times n}$, Show $A(x)$ is invertible

$$\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}A^{-1} \quad ; \quad A \cdot A^{-1} = I$$

$$\frac{d}{dx}(A \cdot A^{-1}) = \frac{d}{dx}(I)$$

Putting bars to show constant

$$= \frac{d(\underline{A} \cdot A^{-1})}{dx} + \frac{d(A \cdot \underline{A^{-1}})}{dx} = 0$$

$$= A^{-1}\frac{dA}{dx} + A\frac{dA^{-1}}{dx} = 0$$

$$A\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}$$

$$\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx} \cdot A^{-1}$$

6.) $\vec{x}$ and $\beta \in \mathbb{R}^p$. Prove $\frac{d\vec{x}^T\beta}{d\vec{x}} = \beta$

Putting bars to show constant

$$\frac{d\vec{x}^T\beta}{dx} = \frac{d\vec{x}^T\underline{\beta}}{dx} + \cancel{\frac{d\underline{\vec{x}}^T\beta}{dx}}^{0} = \beta + 0 = \beta$$

7) Y is an n vector, Y depend on X & X depends on $Z \in \mathbb{R}^q$

$X: \mathbb{R}^q \to \mathbb{R}^p$ & $Y: \mathbb{R}^p \to \mathbb{R}^n$

$$\frac{dY}{dZ} = \frac{dX}{dZ}\frac{dY}{dX}$$

$$X = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \quad Z = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}$$

$$\left(\frac{dY}{dZ}\right)^T = \begin{bmatrix} \frac{dy_1}{dz_1} & \cdots & \frac{dy_1}{dz_n} \\ \vdots & \ddots & \vdots \\ \frac{dy_n}{dz_1} & \cdots & \frac{dy_n}{dz_n} \end{bmatrix}$$

$$\frac{dy_i}{dz_j} = \sum_{q=1} \frac{dy_i}{dx_q}\frac{dx_q}{dz_j} \qquad \begin{matrix} i = 1, 2, \dots, n \\ j = 1, 2, \dots, n \end{matrix}$$

$$\left(\frac{dY}{dZ}\right)^T = \begin{bmatrix} \sum_q \frac{dy_1}{dx_q}\frac{dx_q}{dz_1} & \cdots & \sum_q \frac{dy_1}{dx_q}\frac{dx_q}{dz_n} \\ \vdots & & \vdots \\ \sum_q \frac{dy_m}{dx_q}\frac{dx_q}{dz_1} & \cdots & \sum_q \frac{dy_m}{dx_q}\frac{dx_q}{dz_n} \end{bmatrix} = \left(\frac{dY}{dX}\right)^T \left(\frac{dX}{dZ}\right)^T$$

$$\left(\frac{dY}{dZ}\right)^T = \left(\frac{dY}{dX}\frac{dX}{dZ}\right)^T \Rightarrow \frac{dY}{dZ} = \frac{dY}{dX}\frac{dX}{dZ}$$

Order <u>Maters</u> due to Matrix multiplication

8) $Z : \mathbb{R}^p \to \mathbb{R}$ depends on $\vec{x} \in \mathbb{R}^p$.
Y n-vector depends on $\vec{x} \in \mathbb{R}^p$

Prove $$\frac{d}{d\vec{x}}(zY) = z\frac{dY}{d\vec{x}} + \frac{dz}{dx}Y^T$$

$$\frac{d}{d\vec{x}}(zY) = \frac{d}{dx_j}z(Y_i) = \frac{d}{dx_j}\left(z\left(\frac{dy_i}{dx_j}\right)\right)$$

$$= z\frac{dy_i}{dx_i} + \frac{dz}{dx_i}(y_i) = z\left(\frac{dy_i}{dx_i}\right)_{p\times n} + \left(\frac{dz}{dx_i}\right)_{p\times 1}(y_i)_{1\times n}$$

$$\Rightarrow z\frac{dY}{dx} + \frac{dz}{dx}Y^T$$