MATRIX CALCULUS

1. DERIVATIVES WITH SCALARS : (= DERIVATIVES WITH RESPECT TO SALARS)

x IS A SCALAR

(a) y IS A SCALAR : $\frac{\partial y}{\partial x}$

(b) y IS A VECTOR

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} \quad ; \quad \frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x} \\ \frac{\partial y_2}{\partial x} \\ \vdots \\ \frac{\partial y_m}{\partial x} \end{bmatrix}$$

(c) y IS A MATRIX $m \times n$

$$y = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix} \quad ; \quad \frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_{11}}{\partial x} & \frac{\partial y_{12}}{\partial x} & \cdots & \frac{\partial y_{1n}}{\partial x} \\ \frac{\partial y_{21}}{\partial x} & \frac{\partial y_{22}}{\partial x} & \cdots & \frac{\partial y_{2n}}{\partial x} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_{m1}}{\partial x} & \frac{\partial y_{m2}}{\partial x} & \cdots & \frac{\partial y_{mn}}{\partial x} \end{bmatrix}$$

OR:

$$y = [y_{ij}]_{m \times n} \quad ; \quad \frac{\partial y}{\partial x} = \left[\frac{\partial y_{ij}}{\partial x} \right]_{m \times n}$$

2. DERIVATIVE OF A SCALAR WITH RESPECT TO A MATRIX.

y IS A SCALAR

$x = [x_{ij}]$ IS AN $m \times n$ MATRIX

$\frac{\partial y}{\partial x} = \left[\frac{\partial y}{\partial x_{ij}}\right]$ IS AN $m \times n$ MATRIX

3. DERIVATIVES WITH RESPECT TO VECTORS

x IS A VECTOR OF SIZE n :

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

(a) y IS A SCALAR (SCALAR)

$$\frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y}{\partial x_1} \\ \frac{\partial y}{\partial x_2} \\ \vdots \\ \frac{\partial y}{\partial x_n} \end{bmatrix} = \text{GRADIENT OF } y = \nabla y$$

(b) y IS A VECTOR

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} ; \quad \frac{\partial y}{\partial x} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_2}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \frac{\partial y_1}{\partial x_2} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_n} & \frac{\partial y_2}{\partial x_n} & \cdots & \frac{\partial y_m}{\partial x_n} \end{bmatrix}$$

$$= \begin{bmatrix} \nabla y_1 & \nabla y_2 & \cdots & \nabla y_m \end{bmatrix}$$

RESULTS

y IS $m \times 1$

x IS $n \times 1$

A IS $m \times n$

1. $\boxed{y = Ax} \Rightarrow \boxed{\dfrac{\partial y}{\partial x} = A^T}$; A IS INDEPENDENT OF x

PROOF:

$y_1 = \sum_{k=1}^{n} a_{1k} x_k$	$y_2 = \sum_{k=1}^{n} a_{2k} x_k$	$\ldots \quad y_m = \sum_{k=1}^{n} a_{mk} x_k$
$\dfrac{\partial y_1}{\partial x_1} = a_{11}$	$\dfrac{\partial y_2}{\partial x_1} = a_{21}$	$\ldots \quad \dfrac{\partial y_m}{\partial x_1} = a_{m1}$
$\dfrac{\partial y_1}{\partial x_2} = a_{12}$	$\dfrac{\partial y_2}{\partial x_2} = a_{22}$	$\ldots \quad \dfrac{\partial y_m}{\partial x_2} = a_{m2}$
$\vdots$	$\vdots$	
$\dfrac{\partial y_1}{\partial x_n} = a_{1n}$	$\dfrac{\partial y_2}{\partial x_n} = a_{2n}$	$\ldots \quad \dfrac{\partial y_m}{\partial x_n} = a_{mn}$
∇y_1	∇y_2	$\ldots \quad \nabla y_m$

$$\frac{\partial y}{\partial x} = \left[\nabla y_1 \quad \nabla y_2 \cdots \nabla y_m \right] = A^T \qquad \square$$

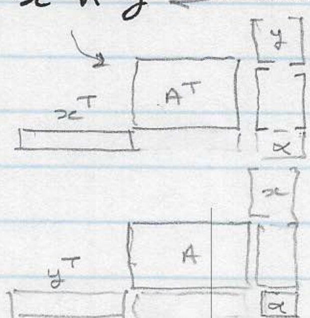
2. $\boxed{\alpha = y^T A x}$ SCALAR A IS INDEPENDENT OF x & y

$\boxed{\dfrac{\partial \alpha}{\partial x} = A^T y}$ GRADIENT COLUMN VECTOR (SIZE m)

$\boxed{\dfrac{\partial \alpha}{\partial y} = A x}$ GRADIENT COLUMN VECTOR (SIZE m)

PROOF: $\alpha = y^T A x = (y^T A x)^T = x^T A^T y$ ← SCALAR

$$\frac{\partial \alpha}{\partial x} = A^T y$$



$$\frac{\partial \alpha}{\partial y} = A x$$

3. $\boxed{\alpha = x^T A x}$ ~~A IS INDE~~

x IS $n \times 1$

A IS $n \times n$ A IS INDEPENDENT OF x

$\boxed{\dfrac{\partial \alpha}{\partial x} = (A + A^T) x}$ → COLUMN VECTOR (SIZE m) $\boxed{\dfrac{\partial^2 \alpha}{\partial x^2} = (A + A^T)}$

4. IN THE LAST RESULT, IF A IS SYMMETRIC, THEN

$\boxed{\dfrac{\partial \alpha}{\partial x} = 2 A x}$

$\boxed{\dfrac{\partial^2 \alpha}{\partial x^2} = 2 A}$

5. $\alpha = y^T x \longrightarrow$ SCALAR

y IS $m \times 1$ FUNCTION OF VECTOR z

x " " " " " "

z IS A VECTOR

$$\boxed{\frac{\partial \alpha}{\partial z} = x^T \frac{\partial y}{\partial z} + y^T \frac{\partial x}{\partial z}} \longleftarrow \text{GRADIENT}$$

PROOF: $\alpha = \sum_{j=1}^{m} x_j y_j = x^T y = y^T x$

$$\frac{\partial \alpha}{\partial z_k} = \sum_{j=1}^{m} x_j \frac{\partial y_j}{\partial z_k} + \sum_{j=1}^{m} y_j \frac{\partial x_j}{\partial z_k} \; ; \quad k = 1, 2, \cdots, m$$

$$\frac{\partial \alpha}{\partial z_k} = \sum_{j=1}^{m} \frac{\partial \alpha}{\partial y_j} \cdot \frac{\partial y_j}{\partial z_k} + \sum_{j=1}^{m} \frac{\partial \alpha}{\partial x_j} \cdot \frac{\partial x_j}{\partial z_k}$$

$$= \frac{\partial \alpha}{\partial y} \cdot \frac{\partial y}{\partial z_k} + \frac{\partial \alpha}{\partial x} \cdot \frac{\partial x}{\partial z_k}$$

DOT PRODUCT

$$\frac{\partial \alpha}{\partial z} = \frac{\partial \alpha}{\partial y} \cdot \frac{\partial y}{\partial z} + \frac{\partial \alpha}{\partial x} \cdot \frac{\partial x}{\partial z} = x^T \frac{\partial y}{\partial z} + y^T \frac{\partial x}{\partial z} \quad \square$$

6. $\alpha = x^T x \longleftarrow$ SCALAR

x IS $m \times 1$; FUNCTION OF z

$$\boxed{\frac{\partial \alpha}{\partial z} = 2 x^T \frac{\partial x}{\partial z}}$$

MATRIX DERIVATIVE CONVENTIONS

THERE ARE TWO TYPES OF DERIVATIVES WITH MATRICES (WITH RESPECT TO THE LAYOUT OF THE DERIVATIVE)

ORGANIZATION

y = $m \times 1$ VECTOR

x = $n \times 1$ VECTOR

A = $m \times n$ MATRIX

$y = Ax$

CONVENTION 1: $\frac{\partial y}{\partial x} = A$

CONVENTION 2: $\frac{\partial y}{\partial x} = A^T$

∂x

CAVEAT DIFFERENT AUTHORS USE DIFFERENT CONVENTIONS.

SOME TIMES THE CONVENTIONS ARE ALSO MIXED-UP!! □