Vector Calculus- 2

23 March 2024 14:03

Matrix Differentiation

To keep in mind:

$(A)_{n\times n} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}_{n\times n}$

$(B)_{n\times n} = \begin{bmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \\ b_{n1} & \cdots & b_{nn} \end{bmatrix}_{n\times n}$

$\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}_{n\times 1}$ $\quad \bar{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1}$

Formula 1: If $\varepsilon = \bar{x}^T\bar{x}$, $\dfrac{\partial \varepsilon}{\partial \bar{x}} = 2\bar{x}$

Proof:

$\bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ $\quad \bar{x}^T = [x_1 \ x_2 \ldots x_n]$

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

$$\varepsilon = x_1^2 + x_2^2 + \ldots x_n^2$$

$$\rightarrow \frac{\partial \varepsilon}{\partial x_i} = 2x_i \Rightarrow \frac{\partial \varepsilon}{\partial \bar{x}} = 2\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = 2\bar{x} \qquad \text{HP.}$$

Formula 2: If $\varepsilon = \bar{x}^T\bar{y}$, $\dfrac{\partial \varepsilon}{\partial \bar{x}} = \bar{y}$

Proof:

$\bar{x}^T = [x_1 \ x_2 \ldots x_n]_{1\times n}$ $\quad \bar{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1}$

$$\varepsilon = \bar{x}^T\bar{y} = x_1y_1 + x_2y_2 + \ldots x_ny_n$$

$$= \sum_{i=1}^{n} x_iy_i$$

$$\Rightarrow \frac{\partial \varepsilon}{\partial x_i} = y_i \Rightarrow \frac{\partial \varepsilon}{\partial \bar{x}} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \Rightarrow \frac{\partial \varepsilon}{\partial \bar{x}} = \bar{y} \qquad \text{hp.}$$

Formula 3: If $\varepsilon = (\bar{x}^T\bar{y})^2$, $\dfrac{\partial \varepsilon}{\partial \bar{x}} = 2(\bar{x}^T\bar{y})\bar{y}$

Proof:

$\bar{x}^T = [x_1 \ldots x_n]$ $\quad \bar{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$

$$\varepsilon = (\bar{x}^T \cdot \bar{y})^2 = (x_1y_1 + x_2y_2 + \ldots + x_ny_n)^2$$

$$\rightarrow \frac{\partial \varepsilon}{\partial x_i} = 2(x_iy_i)y_i \rightarrow \frac{\partial \varepsilon}{\partial \bar{x}} = 2(x^T\bar{y})\bar{y} \qquad \text{Hp.}$$

Formula 4: $\dfrac{\partial (A\bar{x})}{\partial \bar{x}^T} = A$

Proof:

$$A\bar{x} = \overset{n\times n}{\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}} \overset{n\times 1}{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots a_{nn}x_n \end{bmatrix}$$

$a_{ij}x_j \Rightarrow j$ is constant for each row.
i is varying

$$A\bar{x} = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_n \end{bmatrix}$$ where S_i is a row

$$S_i = a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \ldots a_{in}x_n$$

Namita Achyuthan

$$S_i = a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \dots a_{in}x_n$$
$$= \sum_{i=0}^{n} a_{ij} x_j$$

$$\Rightarrow \frac{\partial S_i}{\partial x_j} = \sum a_{ij}$$ [other terms vanish]

$$\Rightarrow \frac{\partial A\bar{x}}{\partial \bar{x}} = \begin{bmatrix} \partial s_1/\partial x_1 & \partial s_1/\partial x_2 & \dots & \partial s_1/\partial x_n \\ \vdots & & & \\ \partial s_n/\partial x_1 & \partial s_n/\partial x_2 & \dots & \partial s_n/\partial x_n \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & & & \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} = A$$

Matrix Diff^n Utilities.

Proposition A: If $\bar{y}_{m\times1}$, $\bar{x}_{n\times1}$, $A_{m\times n}$, A doesn't depend on $\bar{x}$:

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

→ diff of m×1 wrt n×1 yields independent m×n matrix

Proof: i^{th} element of $\bar{y}$: y_i

$$y_i = \sum_{j=1}^{n} a_{ij} x_j$$ [Same as prev. proof but y_i in place of S_i]

$$\frac{\partial y_i}{\partial x_i} = a_{ij}$$

$$\Rightarrow \frac{\partial \bar{y}}{\partial \bar{x}} = A$$ hp.

Proposition B: If $\bar{y}_{m\times1}$, $\bar{x}_{n\times1} = f(\bar{z})$, $A_{m\times n}$, doesn't depend on $\bar{x}$ or $\bar{z}$:

$$\frac{\partial \bar{y}}{\partial \bar{z}} = A \frac{\partial \bar{x}}{\partial \bar{z}}$$

Proof: $\bar{y} = A\bar{x}$

wkt $y_i = \sum_{j=1}^{n} a_{ij} x_i$ [one row]

$\bar{x} = f(\bar{z})$

$$\Rightarrow \frac{\partial y_i}{\partial z_j} = \sum_{j=1}^{n} a_{ij} \frac{\partial x_j}{\partial z_j}$$

$$\Rightarrow \frac{\partial \bar{y}}{\partial \bar{z}} = \frac{\partial \bar{y}}{\partial \bar{x}} \cdot \frac{\partial \bar{x}}{\partial \bar{z}} = A \frac{\partial \bar{x}}{\partial \bar{z}}$$

Namita Achyuthan

Proposition C: $\bar{y}_{m\times1}$, $\bar{x}_{n\times1}$, A doesn't depend on $\bar{x}$ or $\bar{y}$. α = scalar = $\bar{y}^TA\bar{x}$

$$\frac{\partial\alpha}{\partial\bar{x}} = \bar{y}^TA \qquad \frac{\partial\alpha}{\partial\bar{y}} = \bar{x}^TA^T$$

Proof: ① $\alpha = \bar{y}^TA\bar{x}$

Let $\bar{\omega}^T = \bar{y}^TA$

$\Rightarrow \alpha = \bar{\omega}^T\bar{x}$

$\frac{\partial\alpha}{\partial\bar{x}} = \bar{\omega}^T = \bar{y}^TA$ ✓

② $\alpha^T = \alpha$

$\Rightarrow \alpha^T = (\bar{y}^TA\bar{x})^T$

$= \bar{x}^TA^T\bar{y}$

Let $\bar{\omega}^T = \bar{x}^TA^T$

$\Rightarrow \alpha^T = \alpha = \bar{\omega}^T\bar{y}$ ✓

$\frac{\partial\alpha}{\partial\bar{y}} = \bar{\omega}^T = \bar{x}^TA^T$ Hp

Proposition D: If $\bar{x}_{n\times1}$, $A_{m\times n}$ doesn't depend on $\bar{x}$, α = scalar = $\bar{x}^TA\bar{x}$

$$\frac{\partial\alpha}{\partial\bar{x}} = \bar{x}^T(A+A^T)$$

Proof:

$\alpha = \bar{x}^TA\bar{x}$

$= \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_ix_j$ [two for loops]

go row by row. Let k represent the row.

$\frac{\partial\alpha}{\partial x_k} = \sum_{j=1}^{n} a_{kj}x_j + \sum_{i=1}^{n} a_{ik}x_i$

$\frac{\partial\alpha}{\partial\bar{x}} = \bar{x}^TA^T + \bar{x}^TA = \bar{x}^T(A^T+A)$ Hp.

Proposition E: Same as D, but $A_{m\times n}$ is symmetric $\Rightarrow A^T = A$.

$$\frac{\partial\alpha}{\partial\bar{x}} = 2\bar{x}^TA$$

Proof: $\frac{\partial\alpha}{\partial\bar{x}} = \bar{x}^T(A^T+A)$

$= 2\bar{x}^TA$ Hp.

Proposition F: $\bar{y}_{m\times1} = f(\bar{z})$, $\bar{x}_{n\times1} = f(\bar{z})$, α = scalar = $\bar{y}^T\bar{x}$

$$\frac{\partial\alpha}{\partial\bar{z}} = \bar{x}^T\frac{\partial\bar{y}}{\partial\bar{z}} + \bar{y}^T\frac{\partial\bar{x}}{\partial\bar{z}}$$

Proof: $\bar{y}^T = [y_1 \cdots y_n]$ $\quad \bar{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$

$\alpha = \bar{y}^T\bar{x} = \sum_{i=1}^{n} x_iy_i$

$\frac{\partial\alpha}{\partial z_k} = \sum_{i=1}^{n} x_i\frac{\partial y_i}{\partial z_k} + y_i\frac{\partial x_i}{\partial z_k} \Rightarrow \frac{\partial\alpha}{\partial\bar{z}} = \frac{\partial\alpha}{\partial\bar{y}}\cdot\frac{\partial\bar{y}}{\partial\bar{z}} + \frac{\partial\alpha}{\partial\bar{x}}\cdot\frac{\partial\bar{x}}{\partial\bar{z}}$

$= \bar{x}^T\frac{d\bar{y}}{d\bar{z}} + \bar{y}^T\frac{\partial\bar{x}}{\partial\bar{z}}$ hp.

Namita Achyuthan

$$= \bar{x}^T \frac{d\bar{y}}{d\bar{z}} + \bar{y}^T \frac{\partial \bar{x}}{\partial \bar{z}} \quad hp.$$

Proposition G: $\bar{x}_{n\times 1} = f(\bar{z})$, $\alpha = \text{scalar} = \bar{x}^T\bar{x}$

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

Proof:

$$\frac{\partial \alpha}{\partial \bar{z}} = \frac{\partial \alpha}{\partial \bar{x}} \cdot \frac{\partial \bar{x}}{\partial \bar{z}} + \frac{\partial \alpha}{\partial \bar{x}} \cdot \frac{\partial \bar{x}}{\partial \bar{z}}$$

$$= 2\left[\frac{\partial \alpha}{\partial \bar{x}} \cdot \frac{\partial \bar{x}}{\partial \bar{z}}\right]$$

$$= 2\bar{x}^T \frac{\partial \bar{x}}{\partial \bar{z}} \quad hp.$$

Proposition H: $\bar{y}_{m\times 1} = f(\bar{z})$, $\bar{x}_{n\times 1} = f(\bar{z})$, $A_{m\times n}$ doesn't depend on $\bar{z}$, $\alpha = \text{scalar} = \bar{y}^T A \bar{x}$

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

Proof: wkt

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

let $w^T = y^T A$

$\Rightarrow \alpha = w^T \bar{x}$

$$\frac{\partial \alpha}{\partial \bar{z}} = \bar{x}^T \frac{\partial \bar{w}}{\partial \bar{z}} + \bar{w}^T \frac{\partial \bar{x}}{\partial \bar{z}}$$

$$= x^T A^T \frac{\partial \bar{y}}{\partial \bar{z}} + y^T A \frac{\partial \bar{x}}{\partial \bar{z}} \quad hp.$$

Proposition I: $y_{m\times 1} = f(\bar{z})$, $x_{n\times 1} = f(z)$, $A_{m\times n} \neq f(\bar{z})$, $\alpha = \text{scalar} = \bar{x}^T A \bar{x}$

$$\frac{\partial \alpha}{\partial \bar{z}} = \bar{x}^T (A + A^T) \frac{\partial \bar{x}}{\partial \bar{z}}$$

Proof: same as before.

Proposition J: same as I, but $A_{n\times n}$ is symmetric

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

Proof: same as before

Proposition K: A is a non-singular matrix, $\{a_{ij}\} \in A$, $a_{ij} = f(\alpha)$

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

$$\frac{\partial A}{\partial \alpha} = \begin{bmatrix} \frac{\partial a_{11}}{\partial \alpha} & \frac{\partial a_{12}}{\partial \alpha} & \cdots & \frac{\partial a_{1n}}{\partial \alpha} \\ \vdots & & & \\ \frac{\partial a_{n1}}{\partial \alpha} & \cdots & & \frac{\partial a_{nn}}{\partial \alpha} \end{bmatrix}$$

Proof: A is non-singular $\Rightarrow AA^{-1} = I$

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

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

$$\Rightarrow \frac{\partial A^{-1}}{\partial \alpha} = -A^{-1} \frac{\partial A}{\partial \alpha} A^{-1}$$

Namita Achyuthan

$\partial\alpha \qquad \partial\alpha$

$$\Rightarrow \frac{\partial A^{-1}}{\partial \alpha} = -A^{-1}\frac{\partial A}{\partial \alpha}A^{-1}$$