Theoretical Machine Learning

Aman Sanwal

May 2024

1 Answer 1 Part A

$A_{m*n} =$

$$\begin{bmatrix} A_{11} & A_{12}\dots & A_{1n} \\ A_{21} & A_{22}\dots & A_{2n} \\ ..... & ....\dots & .... \\ A_{n1} & A_{n2}\dots & A_{nn} \end{bmatrix}$$

$x = [x1x2.....xn]^T$

$Ax = [\sum_{i=1}^{n} A_{1i}x_i \sum_{i=1}^{n} A_{2i}x_i \cdots \sum_{i=1}^{n} A_{mi}x_i]^T$

$\frac{d(Ax)}{dx} = \frac{d}{dx} [\sum_{i=1}^{n} A_{1i}x_i \sum_{i=1}^{n} A_{2i}x_i \cdots \sum_{i=1}^{n} A_{mi}x_i]^T$
$= \begin{bmatrix} \frac{d[\sum_{i=1}^{n} A_{1i}x_i}{dx1} & \frac{d[\sum_{i=1}^{n} A_{1i}x_i}{dx2} & \dots & \frac{d[\sum_{i=1}^{n} A_{1i}x_i}{dxn} \\ \frac{d[\sum_{i=1}^{n} A_{2i}x_i}{dx1} & \frac{d[\sum_{i=1}^{n} A_{2i}x_i}{dx2} & \dots & \frac{d[\sum_{i=1}^{n} A_{2i}x_i}{dxn} \\ ..... & ....\dots & & .... \\ \frac{d[\sum_{i=1}^{n} A_{mi}x_i}{dx1} & \frac{d[\sum_{i=1}^{n} A_{mi}x_i}{dx2} & \dots & \frac{d[\sum_{i=1}^{n} A_{mi}x_i}{dxn} \end{bmatrix}$

now since we have that $\frac{d[\sum_{i=1}^{n}(A_{ji}x_i)}{dxk} = A_{jk}$ hence we will get that
$\frac{d(Ax)}{dx} = \begin{bmatrix} A_{11} & A_{12}\dots & A_{1n} \\ A_{21} & A_{22}\dots & A_{2n} \\ ..... & ....\dots & .... \\ A_{n1} & A_{n2}\dots & A_{nn} \end{bmatrix}$
$= A$

2 Answer 1 Part B

$x = [x1x2.....xn]^T$

$$A = \begin{bmatrix} A_{11} & A_{12}\ldots & A_{1n} \\ A_{21} & A_{22}\ldots & A_{2n} \\ ..... & ....\ldots & .... \\ A_{n1} & A_{n2}\ldots & A_{nn} \end{bmatrix}$$

on matrix multiplication calculation we get that

$x^T Ax = \sum_{i=1}^{n}\sum_{j=1}^{n} x_j A_{ji} x_i = R(say)$

$\frac{d(R)}{dx} = \begin{bmatrix} \frac{d(R)}{dx1} & \frac{d(R)}{dx2} & \cdots & \frac{d(R)}{dxn} \end{bmatrix}$
$= \left[\sum_{i=1}^{n}(A_{1i} + A_{i1})x_i \quad \sum_{i=1}^{n}(A_{2i} + A_{i2})x_i \ldots \quad \sum_{i=1}^{n}(A_{ni} + A_{in})x_i\right]$

Also we have that
$x^T(A + A^T) = x^T A + x^T A^T$

$= (\sum_{i=1}^{n} A_{1i}x_i \sum_{i=1}^{n} A_{2i}x_i \ldots \sum_{i=1}^{n} A_{ni}x_i) + (\sum_{i=1}^{n} A_{i1}x_i \sum_{i=1}^{n} A_{i2}x_i \ldots \sum_{i=1}^{n} A_{in}x_i)$
$=$
$= \left[\sum_{i=1}^{n}(A_{1i} + A_{i1})x_i \quad \sum_{i=1}^{n}(A_{2i} + A_{i2})x_i \ldots \quad \sum_{i=1}^{n}(A_{ni} + A_{in})x_i\right]$

$= \frac{d(R)}{dx}$
$= \frac{d(x^T Ax)}{dx}$

HENCE WE PROVED THAT

$\frac{d(x^T Ax)}{dx} = x^T(A + A^T)$

3 Answer 2

The dimensions would be m*n*k

4 Answer 3 Part A

$$\frac{d(\begin{bmatrix} 2sin^2xcosy \\ x^2 + 3e^y \end{bmatrix})}{d(\begin{bmatrix} x \\ y \end{bmatrix})} = \begin{bmatrix} 4sinxcosxcosy & -2sin^2xsiny \\ 2x & 3e^y \end{bmatrix}$$

5 Answer 3 Part B

$$\frac{d(\begin{bmatrix} 3x^2y + xyzw \\ sin(x^2 + yw - z) \end{bmatrix})}{d(\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix})}$$

$$= \begin{bmatrix} 6xy + yzw & 3x^2 + xzw & xyw & xyz \\ (cos(x^2 + yw - z))2x & (cos(x^2 + yw - z))w & -cos(x^2 + yw - z) & ycos(x^2 + yw - z) \end{bmatrix}$$

6 Answer 4

$x = [x1x2.....xn]^T$

$B = [B1B2.....Bn]^T$

$B^Tx = \sum_{i=1}^{n} B_i x_i$ which is clearly of dim 1*1

$$\frac{d(e^{B^Tx})}{dx} = \begin{bmatrix} \frac{d(e^{B^Tx})}{dx1} & \frac{d(e^{B^Tx})}{dx2}.......... & \frac{d(e^{B^Tx})}{dxn} \end{bmatrix}$$

$$= \begin{bmatrix} e^{B^Tx}\frac{d(B^Tx)}{dx1} & e^{B^Tx}\frac{d(B^Tx)}{dx2}.......... & e^{B^Tx}\frac{d(B^Tx)}{dxn} \end{bmatrix}$$

$$= e^{B^Tx} \begin{bmatrix} \frac{d(B^Tx)}{dx1} & \frac{d(B^Tx)}{dx2}.......... & \frac{d(B^Tx)}{dxn} \end{bmatrix}$$

$$= e^{B^Tx}\frac{d(B^Tx)}{dx}$$

$$= B^Te^{B^Tx}$$