Theoretical Machine Learning

Theoretical Assignment 1 Solutions

Problem 1. Prove the following :

(a) A is an m x n matrix. $\vec{x}$ is an n x 1 vector. Then,
$\frac{d}{d\vec{x}}[A\vec{x}] = A$

(b) A is an n x n matrix, $\vec{x}$ is an n x 1 vector. Then,
$\frac{d}{d\vec{x}}[\vec{x}^T A\vec{x}] = \vec{x}^T(A + A^T)$

(a) $$A = \begin{bmatrix} A_{1,1} & A_{1,2} & . & . & . & A_{1,n} \\ A_{2,1} & A_{2,2} & . & . & . & A_{2,n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ A_{m,1} & A_{m,2} & . & . & . & A_{m,n} \end{bmatrix}$$

$$A\vec{x} = \begin{bmatrix} A_{1,1}x_1 + A_{1,2}x_2 + \ldots + A_{1,n}x_n \\ A_{2,1}x_1 + A_{2,2}x_2 + \ldots + A_{2,n}x_n \\ . \\ . \\ A_{m,1}x_1 + A_{m,2}x_2 + \ldots + A_{m,n}x_n \end{bmatrix}$$

$$\frac{\partial}{\partial\vec{x}}[A\vec{x}] = \begin{bmatrix} \frac{\partial}{\partial\vec{x}}[A_{1,1}x_1 + A_{1,2}x_2 + \ldots + A_{1,n}x_n] \\ \frac{\partial}{\partial\vec{x}}[A_{2,1}x_1 + A_{2,2}x_2 + \ldots + A_{2,n}x_n] \\ . \\ . \\ \frac{\partial}{\partial\vec{x}}[A_{m,1}x_1 + A_{m,2}x_2 + \ldots + A_{m,n}x_n] \end{bmatrix} = \begin{bmatrix} A_{1,1} & A_{1,2} & . & . & . & A_{1,n} \\ A_{2,1} & A_{2,2} & . & . & . & A_{2,n} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ A_{m,1} & A_{m,2} & . & . & . & A_{m,n} \end{bmatrix} = A$$

(b) $$\vec{x}^T A\vec{x} = \begin{bmatrix} A_{1,1}x_1 + A_{2,1}x_2 + \ldots + A_{n,1}x_n & . & . & . & A_{1,n}x_1 + A_{2,n}x_2 + \ldots + A_{n,n}x_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ . \\ . \\ . \\ x_n \end{bmatrix}$$

The final output is
$\Sigma_{i\neq j}A_{i,j}x_ix_j + \Sigma_i A_{i,i}x_i^2$
So, $\frac{\partial}{\partial x_i}[\Sigma_{i\neq j}A_{i,j}x_ix_j + \Sigma_i A_{i,i}x_i^2] = \Sigma_{j\neq i}(A_{i,j} + A_{j,i})x_j + 2A_{i,i}x_i$ [Cross Check this step!]
Extending this, we get the final output,
$\vec{x}^T(A + A^T)$

Problem 2. Suppose you have a matrix of dimension m x n, and you differentiating wrt a k x 1 vector, what is the dimension of the final result?

For each element in the m x n matrix, differentiating wrt a k x 1 gives a 1 x k vector. So, we get a **m x nk** matrix.

Problem 3. Solve :

(a) $\dfrac{d\begin{bmatrix} 2sin^2(x)cos(y) \\ x^2 + 3e^y \end{bmatrix}}{d\begin{bmatrix} x \\ y \end{bmatrix}}$

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

(a) $\dfrac{\partial[2sin^2(x)cos(y)]}{\partial\begin{bmatrix} x \\ y \end{bmatrix}} = \begin{bmatrix} \frac{\partial[2sin^2(x)cos(y)]}{\partial x} & \frac{\partial[2sin^2(x)cos(y)]}{\partial y} \end{bmatrix} = \begin{bmatrix} 4sin(x)cos(y) & -2sin^2(x)sin(y) \end{bmatrix}$

$\dfrac{\partial[x^2+3e^y]}{\partial\begin{bmatrix} x \\ y \end{bmatrix}} = \begin{bmatrix} \frac{\partial[x^2+3e^y]}{\partial x} & \frac{\partial[x^2+3e^y]}{\partial y} \end{bmatrix} = \begin{bmatrix} 6x & 3e^y \end{bmatrix}$

$\dfrac{d\begin{bmatrix} 2sin^2(x)cos(y) \\ x^2 + 3e^y \end{bmatrix}}{d\begin{bmatrix} x \\ y \end{bmatrix}} = \begin{bmatrix} 4sin(x)cos(y) & -2sin^2(x)sin(y) \\ 6x & 3e^y \end{bmatrix}$

(b) Similarly,

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

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

Problem 4. BONUS : β is an n x 1 vector. $\vec{x}$ is an n x 1 vector. Solve for : $\frac{d}{d\vec{x}}[e^{\beta^T\vec{x}}]$

$\frac{d[e^{\beta^T\vec{x}}]}{d\vec{x}} = \frac{d[\beta^T\vec{x}]}{d\vec{x}}\frac{d[e^{\beta^T\vec{x}}]}{d[\beta^T\vec{x}]} = \beta^T e^{\beta^T\vec{x}}$