# Deep Learning Homework1 Part3

Lidian Lin

January 19th, 2020

## a Proof: $\nabla_x(x^T a) = \nabla_x(a^T x) = a$

$$\because x^T a = \begin{bmatrix} x_1 & x_2 & ... & x_n \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{bmatrix} = \sum_{i=1}^{n} x_i a_i$$

$$\therefore \nabla_x(x^T a) = \begin{bmatrix} \frac{\partial x_1 a_1}{\partial x_1} \\ \frac{\partial x_2 a_2}{\partial x_2} \\ ... \\ \frac{\partial x_n a_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{bmatrix} = a$$

$$\text{similarly, } \nabla_x(a^T x) = \begin{bmatrix} \frac{\partial a_1 x_1}{\partial x_1} \\ \frac{\partial a_2 x_2}{\partial x_2} \\ ... \\ \frac{\partial a_n x_n}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ ... \\ a_n \end{bmatrix} = a$$

$$\therefore \nabla_x(x^T a) = \nabla_x(a^T x) = a$$

## b Proof: $\nabla_x(x^TAx) = (A^T + A)x$

$\because x^TAx = \begin{bmatrix} x_1 & x_2 & ... & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{n1} & a_{n2} & ... & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{bmatrix}$

$= \sum_{j=1}^n \sum_{i=1}^n x_j x_i a_{ij}$

$\therefore \nabla_x(x^TAx) = [\frac{\partial x^TAx}{\partial x_k}] = [\frac{\partial \sum_{j=1}^n \sum_{i=1}^n x_j x_i a_{ij}}{\partial x_k}] = [\frac{\partial}{\partial x_k}(x_1 \sum_{i=1}^n x_i a_{i1} + ... + x_k \sum_{i=1}^n x_i a_{ik} + ... + x_n \sum_{i=1}^n x_i a_{in})]$

$= [x_1 a_{k1} + ... + x_k a_{kk} + \sum_{i=1}^n x_i a_{ik} + ... + x_n a_{kn}]$

$= [\sum_{j=1}^n x_j a_{kj} + \sum_{i=1}^n x_i a_{ik}]$

$= [A_{(k)}x + A^{(k)}x]$

$= (A + A^T)x$

## c Proof: $\nabla_x(x^TAx) = 2Ax$

$\because \nabla_x(x^TAx) = (A^T + A)x \quad and \quad A = A^T$

$\therefore \nabla_x(x^TAx) = (A + A)x = 2Ax$

## d Proof: $\nabla_x(Ax + b)^T(Ax + b) = 2A^T(Ax + b)$

$\because (Ax + b)^T(Ax + b) = (x^TA^T + b^T)(Ax + b)$

$= x^TA^TAx + x^TA^Tb + b^TAx + b^Tb$

$\therefore \nabla_x(Ax + b)^T(Ax + b) = \nabla_x x^TA^TAx + x^TA^Tb + b^TAx + b^Tb$
$= 2A^TAx + A^Tb + b^TA$
$= 2A^TAx + A^Tb$
$= 2A^T(Ax + b)$