Assignment 1

Question 3

Keshav Bimbraw, Vignesh K, Bryan Gass

February 2021

1 Question 3a

We need to prove -

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

Proof -

Since it is a scalar product,

$$x^T a = a^T x$$

Therefore,

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

$$x^T a = xa^T = x_1a_1 + x_2a_2 + ... + x_na_n$$

The partial derivatives can be expanded as follows -

$$\nabla_x(x^T a) = \begin{bmatrix} \nabla_{x_1}(x_1a_1 + x_2a_2 + ... + x_na_n) \\ \nabla_{x_2}(x_1a_1 + x_2a_2 + ... + x_na_n) \\ . \\ . \\ . \\ \nabla_{x_n}(x_1a_1 + x_2a_2 + ... + x_na_n) \end{bmatrix}$$

Therefore,

$$\nabla_x(x^T a) = \nabla_x(a^T x) = \begin{bmatrix} a_1 \\ a_2 \\ . \\ . \\ . \\ a_n \end{bmatrix} = a$$

2 Question 3b

We need to prove -

$$\nabla_x(x^T Ax) = (A + A^T)x$$

Proof -

Using chain rule,

$$\nabla_x(x^T Ax) = ((\nabla_x(x^T A))x) + (x^T A(\nabla_x x))$$
$$\nabla_x(x^T Ax) = Ax + x^T A$$

Since it is a scalar product,

$$x^T A = A^T x$$

Therefore,

$$\nabla_x(x^T Ax) = Ax + A^T x$$
$$\nabla_x(x^T Ax) = (A + A^T)x$$

3 Question 3c

We need to prove -

$$\nabla_x(x^T Ax) = 2Ax$$

Since A is a square matrix,

$$A = A^T$$

From question 3b,

$$\nabla_x(x^T Ax) = (A + A^T)x$$

Therefore,

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

4 Question 3d

$$\nabla_x[(Ax + b)^T (Ax + b)] = 2A^T (Ax + b)$$

$$\nabla_x[(Ax + b)^T (Ax + b)] = ((\nabla_x(Ax + b)^T)(Ax + b)) + ((Ax + b)^T (\nabla_x(Ax + b)))$$

Using the matrix first-order derivative property -

$$\nabla_x[(Ax + b)^T (Ax + b)] = ((\nabla_x(Ax + b))^T (Ax + b)) + ((Ax + b)^T (\nabla_x(Ax + b)))$$

It must be noted that -

$$\nabla_x(Ax + b) = A$$

So,

$$(\nabla_x (Ax+b))^T = A^T$$

Therefore,

$$\nabla_x [(Ax+b)^T (Ax+b)] = A^T (Ax+b) + (Ax+b)^T A$$

Since -

$$(Ax+b)^T A = A^T (Ax+b)$$

The final equation becomes -

$$\nabla_x [(Ax+b)^T (Ax+b)] = A^T (Ax+b) + A^T (Ax+b)$$

$$\nabla_x [(Ax+b)^T (Ax+b)] = 2A^T (Ax+b)$$