3(a) To prove that $\nabla_x (x^T a) = \nabla_x (a^T x) = a$

~~Proof~~

Given that x & a are column vectors -

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} \quad \& \quad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

$$x^T a = [x_1\ x_2\ x_3 \cdots x_n] \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

$$x^T a = \sum_{i=1}^{n} a_i x_i$$

$$\nabla_x (x^T a) = \nabla_x \sum_{i=1}^{n} a_i x_i = \begin{bmatrix} \frac{\partial (a_1 x_1 + a_2 x_2 + \cdots a_n x_n)}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} (a_1 x_1 + \cdots a_n x_n) \end{bmatrix}$$

$$= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a \text{ — (1)} \quad \boxed{\text{Proved}}$$

For the other ~~far~~ expression, $\left[\nabla_x (a^T x)\right]$ we follow the same method

$$a^T x = [a_1\ a_2\ a_3 \cdots a_n] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1 x_1 + \cdots a_n x_n$$

$$a^T x = \nabla_x \sum_{i=1}^{n} a_i x_i$$

Can be proved in the same way as (1).

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

3(b) To prove : $\nabla_x(x^T A x) = (A + A^T)x$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & a_{22} & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \quad x^T = [x_1 x_2 \cdots x_n]$$

$$x^T A x = [x_1\ x_2 \cdots x_n] \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \sum_j^n x_j \sum_i^n x_i a_{ij} = \sum_j^n \sum_i^n x_i x_j a_{ij}$$

$$\frac{\partial}{\partial x_k} (x^T A x) = \frac{\partial}{\partial x_k}\left(\left[x_1 \sum_i x_i a_{i1} + \cdots + \sum_i x_k a_{ik} + \cdots \sum_i x_n a_{in} \right] \right)$$

$$\Delta x_k (x^T A x) = a_1 x_n + \cdots \left(\sum_i x_i a_{ik} + \sum x_k a_{kk} \right) + \cdots a_n x_n$$

$$= \cancel{\sum_i x_i (a_{ij} + a_{ji})} \quad \sum_j x_j x_{jk} + \sum_k x_i x_{ki}$$

$$= \sum_i x_i (a_{ij} + a_{ji})$$

$$= \begin{pmatrix} [1^{st} \text{ row of } A + (\text{transpose of } 1^{st} \text{ col. of } A)]\, x \\ \vdots \\ [n^{th} \text{ row of } A + (\text{transpose of } n^{th} \text{ col. of } A)]\, x \end{pmatrix}$$

$$= \begin{pmatrix} 1^{st} \text{ row of } A + (\text{transpose of } 1^{st} \text{ col of } A) \\ \vdots \\ n^{th} \text{ row of } A + (\text{transpose of } n^{th} \text{ col of } A) \end{pmatrix} X$$

$$= (A_{ij} + A_{ji})\, x = (A + A^T)\, x$$

Hence proved: $\nabla_x (x^T A x) = (A + A^T)\, x$

3 (c) To prove:

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

for symmetric 'A'.

Using proofs from 3(b) we know that,

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

When A is symmetric, $A_{ij} = A_{ji} \Rightarrow A = A^T$

Hence

$$\nabla_x (x^T A x) = (2A)\, x$$

3 (d) To prove, $\nabla_x \left[(Ax+b)^T (Ax+b) \right] = 2A^T (Ax+b)$

~~Differentiating $\left[(Ax+b)^T (Ax+b) \right]$:-~~

$$\cancel{\nabla_{x_i} \left(\sum_j^n \left[\left(\sum_i^n A_{ij}\, x_i + b \right)^2 \right] \right)}$$

$$\cancel{2 \sum_j^n A_{ij}^2\, x_i + A_{ij} b}$$

Simplifying $(Ax+b)^T (Ax+b)$

$$= \left((Ax)^T + b^T \right) (Ax+b)$$

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

$$= x^T A^T A x + b^T A x + b^T b + \cancel{x^T A^T A x}\ x^T A^T b$$

$$\nabla_x\left(x^T A^T A x + b^T A x + x^T A^T b + b^T b \right)$$

$$= \left((A^T A) + A^T A \right) x + A^T b + 0$$

$$= 2 A^T A x + A^T b$$

$$= 2 A^T (A x + B)$$

Hence proved.