(a) Prove that

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

for any two n-dimensional column vectors x, a. Hint: differentiate w.r.t. each element of x, and then gather the partial derivatives into a column vector.

$$x^T a = \begin{bmatrix} x_1 \dots x_n \end{bmatrix} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \sum_{i=1}^{n} x_i a_i = \sum_{i=1}^{n} a_i x_i \tag{1}$$

$$a^T x = \begin{bmatrix} a_1 \dots a_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} a_i x_i \tag{2}$$

$$\sum_{i=1}^{n} a_i x_i = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} x \tag{3}$$

$$\nabla_x \left(\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} x \right) = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = a \tag{4}$$

(b) Prove that

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

for any column vector x, any $n \times n$ matrix A, and any constant n-dimensional vector b.

$$x^T A x = \begin{bmatrix} x_1 \dots x_n \end{bmatrix} \begin{bmatrix} A_{1,1} \dots A_{1,n} \\ \vdots \quad \ddots \quad \vdots \\ A_{n,1} \dots A_{n,n} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{5}$$

$$= \begin{bmatrix} (A_{1,1}x_1 + \dots + A_{n,1}x_n) \dots (A_{1,n}x_1 + \dots + A_{n,n}x_n) \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{6}$$

$$= \begin{bmatrix} \sum_{i=1}^n A_{i,1}x_i \dots \sum_{i=1}^n A_{i,n}x_i \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \tag{7}$$

$$= x_1 \sum_{i=1}^n A_{i,1}x_i + \dots + x_n \sum_{i=1}^n A_{i,n}x_i \tag{8}$$

$$= \sum_{j=1}^n x_j \sum_{i=1}^n A_{i,j}x_i \tag{9}$$

$$= \sum_{j=1}^n \sum_{i=1}^n A_{i,j}x_i x_j \tag{10}$$

$$\nabla_x(x^T A x) = \begin{bmatrix} \nabla_{x_1}(x^T A x) \\ \vdots \\ \nabla_{x_k}(x^T A x) \\ \vdots \\ \nabla_{x_n}(x^T A x) \end{bmatrix} \tag{11}$$

$$= \nabla_x(\sum_{j=1}^n \sum_{i=1}^n A_{i,j}x_i x_j) \tag{12}$$

$$\nabla_{x_k}(x^T A x) = \nabla_{x_k}(\sum_{j=1}^n \sum_{i=1}^n A_{i,j}x_i x_j) \tag{13}$$

$$= \nabla_{x_k}(x_1 \sum_{i=1}^n A_{i,j}x_i + \dots + x_k \sum_{i=1}^n A_{i,j}x_i + \dots + x_n \sum_{i=1}^n A_{i,j}x_i) \tag{14}$$

$$= x_1 A_{k,1} + \dots + (x_k \sum_{i=1}^n A_{i,k}x_i + A_{k,k}x_k) + \dots + x_n A_{k,n} \tag{15}$$

$$= (x_1 A_{k,1} + \dots + x_n A_{k,n}) + x_k \sum_{i=1}^n A_{i,k}x_i \tag{16}$$

$$= \sum_{i=1}^n A_{k,i}x_i + \sum_{i=1}^n A_{i,k}x_i \tag{17}$$

$$= (A_{k,.} + A^T_{.,k})x \tag{18}$$

$$\nabla_x(x^T A x) = \begin{bmatrix} (A_{1,.} + A^T_{.,1})x \\ \vdots \\ (A_{n,.} + A^T_{.,n})x \end{bmatrix} \tag{19}$$

$$= \left(\begin{bmatrix} (A_{1,.} \\ \vdots \\ A_{n,.} \end{bmatrix} + \begin{bmatrix} (A^T_{.,1} \\ \vdots \\ A^T_{.,n} \end{bmatrix} \right) x \tag{20}$$

$$= (A + A^T)x \tag{21}$$

(c) Based on the theorem above, prove that

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

for any n-dimensional column vector x and any symmetric $n \times n$ matrix A.

$$\begin{align}
\nabla_x(x^T Ax) &= (A + A^T)x \tag{22}\\
&= (A + A)x \tag{23}\\
&= 2Ax \tag{24}
\end{align}$$

(d) Based on the theorems above, prove that

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

for any n-dimensional column vector x, any symmetric $n \times n$ matrix A, and any constant n-dimensional column vector b.

$$\begin{align}
(Ax + b)^T + (Ax + b) &= \big((Ax)^T + b^T\big)(Ax + b) \tag{25}\\
&= (x^T A^T + b^T)(Ax + b) \tag{26}\\
&= x^T A^T(Ax + b) + b^T(Ax + b) \tag{27}\\
&= x^T A^T Ax + x^T A^T b + b^T Ax + b^T b \tag{28}\\
&= x^T A^2 x + x^T A^T b + b^T Ax + b^T b \tag{29}\\
\nabla_x(x^T A^2 x + x^T A^T b + b^T Ax + b^T b) &= \nabla_x(x^T A^2 x) + \nabla_x(x^T A^T b) + \nabla_x(b^T Ax) + \nabla_x(b^T b) \tag{30}\\
&= \nabla_x(x^T A^2 x) + Ab + Ab + 0 \tag{31}\\
&= (2A^2 x) + 2Ab \tag{32}\\
&= (2A^T Ax) + 2A^T b \tag{33}\\
&= 2A^T(Ax + b) \tag{34}
\end{align}$$