Ali Bakly
Student ID: 200109236935
al7765ba-s@student.lu.se

Bonus problems — week 1

Let $A \in \mathbb{R}^{n\times n}$ and define the function $f : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(x) = x^T Ax$$

for each $x \in \mathbb{R}^n$. Show that

$$\nabla f(x) = (A + A^T)x$$

for each $x \in \mathbb{R}^n$. If $A \in \mathbb{S}^n$, we conclude from (1) that

$$\nabla f(x) = 2Ax$$

for each $x \in \mathbb{R}^n$.

Proof. We write out the explicit matrices to find a closed form expression of the quadratic form $x^T Ax$:

$$\begin{aligned}
x^T Ax &= [x_1 \dots x_n] \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= [x_1 \dots x_n] \begin{bmatrix} \sum_{j=1}^{n} a_{1j}x_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj}x_j \end{bmatrix} \\
&= x_1 \sum_{j=1}^{n} a_{1j}x_j + x_2 \sum_{j=1}^{n} a_{2j}x_j + \cdots + x_n \sum_{j=1}^{n} a_{nj}x_j \\
&= \sum_{i=1}^{n} x_i \sum_{j=1}^{n} x_j a_{ij}.
\end{aligned}$$

For any index $k \in \mathbb{N}$ such that $1 \le k \le n$ we can find the derivative of $x^T Ax$ with respect to x_k using the product rule:

$$\begin{aligned}
\frac{\partial}{\partial x_k}(x^T Ax) &= \sum_{i=1}^{n} \left(\frac{\partial x_i}{\partial x_k} \sum_{j=1}^{n} x_j a_{ij} + x_i \sum_{j=1}^{n} \frac{\partial x_j}{\partial x_k} a_{ij} \right) \\
&= \sum_{j=1}^{n} x_j a_{kj} + \sum_{i=1}^{n} x_i a_{ik}.
\end{aligned}$$

Here we used the fact that $\dfrac{\partial x_i}{\partial x_k} = 1$ for $i = k$ and $\dfrac{\partial x_i}{\partial x_k} = 0$ for $i \neq k$. The same argument is used for $\dfrac{\partial x_j}{\partial x_k}$. For each $1 \leq k \leq n$ we can structure this in a column vector to find the gradient:

$$\nabla f(x) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} x_j a_{1j} + \sum_{i=1}^{n} x_i a_{i1} \\ \sum_{j=1}^{n} x_j a_{2j} + \sum_{i=1}^{n} x_i a_{i2} \\ \vdots \\ \sum_{j=1}^{n} x_j a_{nj} + \sum_{i=1}^{n} x_i a_{in} \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} x_j a_{1j} \\ \sum_{j=1}^{n} x_j a_{2j} \\ \vdots \\ \sum_{j=1}^{n} x_j a_{nj} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^{n} x_i a_{i1} \\ \sum_{i=1}^{n} x_i a_{i2} \\ \vdots \\ \sum_{i=1}^{n} x_i a_{in} \end{bmatrix} = Ax + A^T x = (A + A^T)x,$$

which is the main result.

Furthermore, it follows that $A = A^T$ if $A \in \mathbb{S}^n$. Then $A + A^T = 2A$ which in turn implies $\nabla f(x) = 2Ax$. This is maybe not so surprising since the quadratic function $g(t) = at^2$ (the case when $A \in \mathbb{R}^{1\times 1}$ and $x \in \mathbb{R}^1$) has the well known derivative $g'(t) = 2at$.

$\blacksquare$