# Quiz02

**1. Gradients, Jacobians, and Hessians**

Suppose $A \in \mathbb{R}^{n \times n}$ is a square matrix whose entries are denoted $a_{ij}$ and whose rows are denoted $\vec{a}_1^\top, \ldots, \vec{a}_n^\top$, and $\vec{b} \in \mathbb{R}^n$ is a vector whose entries are denoted $b_i$.

Compute the gradients for the following functions. Show your work.

(a) $g_1(\vec{x}) = \vec{x}^\top A \vec{x}$.

**Solution:** We have

$$
\begin{aligned}
g_1(\vec{x}) &= \vec{x}^\top A \vec{x} && (1)\\
&= \vec{x}^\top \begin{bmatrix} \vec{a}_1^\top \\ \vdots \\ \vec{a}_n^\top \end{bmatrix} \vec{x} && (2)\\
&= \vec{x}^\top \begin{bmatrix} \vec{a}_1^\top \vec{x} \\ \vdots \\ \vec{a}_n^\top \vec{x} \end{bmatrix} && (3)\\
&= \sum_{i=1}^{n} x_i (\vec{a}_i^\top \vec{x}). && (4)
\end{aligned}
$$

Taking the derivative with respect to any $x_j$ yields

$$
\begin{aligned}
\frac{\partial g_1}{\partial x_j}(\vec{x}) &= \frac{\partial}{\partial x_j} \sum_{i=1}^{n} x_i (\vec{a}_i^\top \vec{x}) && (5)\\
&= \frac{\partial}{\partial x_j} \left[ x_j (\vec{a}_j^\top \vec{x}) + \sum_{\substack{i=1 \\ i \neq j}}^{n} x_i (\vec{a}_i^\top \vec{x}) \right] && (6)\\
&= \frac{\partial}{\partial x_j} x_j (\vec{a}_j^\top \vec{x}) + \sum_{\substack{i=1 \\ i \neq j}}^{n} \frac{\partial}{\partial x_j} x_i (\vec{a}_i^\top \vec{x}) && (7)\\
&= x_j \frac{\partial}{\partial x_j} (\vec{a}_j^\top \vec{x}) + \vec{a}_j^\top \vec{x} + \sum_{\substack{i=1 \\ i \neq j}}^{n} x_i \frac{\partial}{\partial x_j} (\vec{a}_i^\top \vec{x}) && (8)\\
&= x_j (\vec{a}_j)_j + \vec{a}_j^\top \vec{x} + \sum_{\substack{i=1 \\ i \neq j}}^{n} x_i (\vec{a}_i)_j && (9)\\
&= \vec{a}_j^\top \vec{x} + \sum_{i=1}^{n} x_i (\vec{a}_i)_j && (10)\\
&= \sum_{i=1}^{n} a_{ji} x_i + \sum_{i=1}^{n} a_{ij} x_i && (11)\\
&= (A\vec{x})_j + (A^\top \vec{x})_j && (12)\\
&= [(A + A^\top)\vec{x}]_j. && (13)
\end{aligned}
$$

Thus we have

$$
\nabla g_1(\vec{x}) = (A + A^\top)\vec{x}. \qquad (14)
$$

(b) $g_6(\vec{x}) = e^{\|\vec{x}\|_2^2}$.

**Solution:** We use the multivariable chain rule, obtaining

$$\nabla g_6(\vec{x}) = (Dg_6(\vec{x}))^\top \tag{15}$$
$$= ([D\exp(g_2(\vec{x}))][Dg_2(\vec{x})])^\top \tag{16}$$
$$= [Dg_2(\vec{x})]^\top [D\exp(g_2(\vec{x}))]^\top \tag{17}$$
$$= [\nabla g_2(\vec{x})][e^{g_2(\vec{x})}]^\top \tag{18}$$
$$= [2\vec{x}][e^{g_2(\vec{x})}] \tag{19}$$
$$= 2e^{\|\vec{x}\|_2^2}\vec{x}, \tag{20}$$

since the transpose of a scalar is a scalar.

(c) $g_7(\vec{x}) = e^{\|A\vec{x}-\vec{b}\|_2^2} = g_6(A\vec{x} - \vec{b})$.

**Solution:** Notice that in part **??**, we did not use any specific functional properties of $g_4$, $\nabla g_4$, or $\nabla^2 g_5$ to determine $\nabla g_5$ and $\nabla^2 g_5$. Using the exact same analysis, we can conclude that

$$\nabla g_7 = A^\top [\nabla g_6(A\vec{x} - \vec{b})] \tag{21}$$

 