Quiz02

1. Vector Caculus

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

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

(a) $\vec{g}(\vec{x}) = A\vec{x}$.

Solution: We compute each partial derivative and reconstitute $D\vec{g}$ at the end. That is,

$$
\begin{aligned}
[D\vec{g}(\vec{x})]_{jk} &= \frac{\partial g_j}{\partial x_k}(\vec{x}) && (1)\\
&= \frac{\partial (A\vec{x})_j}{\partial x_k} && (2)\\
&= \frac{\partial (\vec{a}_j^\top \vec{x})}{\partial x_k} && (3)\\
&= (\vec{a}_j)_k && (4)\\
&= a_{jk} && (5)\\
&= [A]_{jk}, && (6)
\end{aligned}
$$

where $[A]_{jk}$ is the (j,k)th entry of A. Thus $D\vec{g}(\vec{x}) = A$.

(b) $\vec{g}(\vec{x}) = f(\vec{x})\vec{x}$ where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is differentiable.

Solution: We again compute each partial derivative using the scalar product rule, obtaining

$$
\begin{aligned}
[D\vec{g}(\vec{x})]_{jk} &= \frac{\partial g_j}{\partial x_k}(\vec{x}) && (7)\\
&= \frac{\partial (f(\vec{x})x_j)}{\partial x_k} && (8)\\
&= f(\vec{x})\frac{\partial x_j}{\partial x_k} + x_j \frac{\partial f}{\partial x_k}(\vec{x}) && (9)\\
&= x_j[\nabla f(\vec{x})]_k + \begin{cases} f(\vec{x}), & \text{if } j = k \\ 0, & \text{if } j \neq k. \end{cases} && (10)
\end{aligned}
$$

This gives a Jacobian whose entries are

$$
\begin{aligned}
[D\vec{g}(\vec{x})]_{jk} &= x_j[\nabla f(\vec{x})]_k, \qquad \forall j \neq k && (11)\\
[D\vec{g}(\vec{x})]_{jj} &= x_j[\nabla f(\vec{x})]_k + f(\vec{x}), \qquad \forall j. && (12)
\end{aligned}
$$

For the purpose of self-grades it is fine to stop here. But one can also write the Jacobian as

$$D\vec{g}(\vec{x}) = \vec{x}[\nabla f(\vec{x})]^\top + f(\vec{x})I. \tag{13}$$

(c) $\vec{g}(\vec{x}) = f(A\vec{x} + \vec{b})\vec{x}$ where $f : \mathbb{R}^n \mapsto \mathbb{R}$ is differentiable.

Solution: We compute each partial derivative using the scalar product rule and chain rule, obtaining

$$
\begin{aligned}
[D\vec{g}(\vec{x})]_{jk} &= \frac{\partial g_j}{\partial x_k}(\vec{x}) && (14)\\
&= \frac{\partial f(A\vec{x} + \vec{b})x_j}{\partial x_k} && (15)
\end{aligned}
$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j \frac{\partial f(A\vec{x}+\vec{b})}{\partial x_k} \tag{16}$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j \sum_{\ell=1}^{n} \frac{\partial f(A\vec{x}+\vec{b})}{\partial (A\vec{x}+\vec{b})_\ell} \cdot \frac{\partial (A\vec{x}+\vec{b})_\ell}{\partial x_k} \tag{17}$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j \sum_{\ell=1}^{n} [\nabla f(A\vec{x}+\vec{b})]_\ell \cdot \frac{\partial (\vec{a}_\ell^\top \vec{x} + b_\ell)}{\partial x_k} \tag{18}$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j \sum_{\ell=1}^{n} [\nabla f(A\vec{x}+\vec{b})]_\ell \cdot (\vec{a}_\ell)_k \tag{19}$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j \sum_{\ell=1}^{n} [\nabla f(A\vec{x}+\vec{b})]_\ell \cdot a_{\ell k} \tag{20}$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j \sum_{\ell=1}^{n} [\nabla f(A\vec{x}+\vec{b})]_\ell \cdot [A^\top]_{k\ell} \tag{21}$$

$$= f(A\vec{x}+\vec{b})\frac{\partial x_j}{\partial x_k} + x_j [A^\top \nabla f(A\vec{x}+\vec{b})]_k \tag{22}$$

$$= x_j [A^\top \nabla f(A\vec{x}+\vec{b})]_k + \begin{cases} f(A\vec{x}+\vec{b}), & \text{if } j = k \\ 0 & \text{if } j \neq k. \end{cases} \tag{23}$$

This gives a Jacobian whose entries are

$$[D\vec{g}(\vec{x})]_{jk} = x_j [A^\top \nabla f(A\vec{x}+\vec{b})]_k, \qquad \forall j \neq k \tag{24}$$

$$[D\vec{g}(\vec{x})]_{jj} = x_j [A^\top \nabla f(A\vec{x}+\vec{b})]_k + f(A\vec{x}+\vec{b}), \qquad \forall j. \tag{25}$$

For the purpose of self-grades it is fine to stop here. But one can also write the Jacobian as

$$D\vec{g}(\vec{x}) = \vec{x}[A^\top \nabla f(A\vec{x}+\vec{b})]^\top + f(A\vec{x}+\vec{b})I \tag{26}$$

$$= \vec{x}[\nabla f(A\vec{x}+\vec{b})]^\top A + f(A\vec{x}+\vec{b})I. \tag{27}$$