Sakshi Suman (01/29/2022) suman.sak@northeastern.edu

Problem 1. *Assume $\vec{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $\vec{b} \in \mathbb{R}^m$. Let $f(\vec{x}) = \vec{b}^T A\vec{x}$. Find ∇f.*

Solution. Given, $\vec{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $\vec{b} \in \mathbb{R}^m$, $f(\vec{x}) = \vec{b}^T A\vec{x}$. Let,

$$\vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \text{and } A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}$$

$$\vec{b}^T A = \left[\sum_{i=1}^m b_i \cdot a_{i1} \cdots \sum_{i=1}^m b_i \cdot a_{in}\right]$$

Therefore,

$$f(\vec{x}) = \vec{b}^T A\vec{x} = \sum_{j=1}^n \left(\left(\sum_{i=1}^m b_i \cdot a_{ij}\right) x_j\right)$$

$$\begin{aligned} \nabla f &= \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^m b_i \cdot a_{i1} \\ \vdots \\ \sum_{i=1}^m b_i \cdot a_{in} \end{bmatrix} \\ &= (\vec{b}^T A)^T \\ &= \boxed{A^T \vec{b}} \end{aligned}$$

□

Problem 2. *Assume $\vec{x} \in \mathbb{R}^n$. Find $\frac{\partial \vec{x}^T \vec{x}}{\partial \vec{x}}$.*

Solution. Let, $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \implies \vec{x}^T \vec{x} = \sum_{i=1}^n x_i^2$.

Therefore,

$$\begin{aligned} \frac{\partial \vec{x}^T \vec{x}}{\partial \vec{x}} &= \begin{bmatrix} 2 \cdot x_1 \\ \vdots \\ 2 \cdot x_n \end{bmatrix} \\ &= \boxed{2 \cdot \vec{x}} \end{aligned}$$

□

Problem 3. *Assume $\vec{x}$ and $\vec{a} \in \mathbb{R}^n$. Find $\frac{\partial (\vec{x}^T \vec{a})^2}{\partial \vec{x}}$*

Solution. Let, $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \vec{a} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \implies \vec{x}^T \vec{a} = \sum_{i=1}^{n} x_i \cdot a_i \implies (\vec{x}^T \vec{a})^2 = (\sum_{i=1}^{n} x_i \cdot a_i)^2$

Therefore,

$$\begin{aligned} \frac{\partial(\vec{x}^T \vec{a})^2}{\partial \vec{x}} &= \begin{bmatrix} \frac{\partial(\vec{x}^T \vec{a})^2}{\partial x_1} \\ \vdots \\ \frac{\partial(\vec{x}^T \vec{a})^2}{\partial x_n} \end{bmatrix} \\ &= \begin{bmatrix} 2 \cdot (\sum_{i=1}^{n} x_i \cdot a_i) \cdot a_1 \\ \vdots \\ 2 \cdot (\sum_{i=1}^{n} x_i \cdot a_i) \cdot a_n \end{bmatrix} \\ &= \boxed{2 \cdot (\vec{x}^T \vec{a}) \cdot \vec{a}} \end{aligned}$$

□

Problem 4. *Suppose $\vec{x} : \mathbb{R}^n \to \mathbb{R}^m$ is a map sending $\vec{z} \in \mathbb{R}^n$ to $\vec{x}(\vec{z}) \in \mathbb{R}^m$. Similarly, suppose $\vec{y} : \mathbb{R}^n \to \mathbb{R}^m$ and A is an $m \times m$ constant matrix. Prove that $\frac{\partial(\vec{y}^T A \vec{x})}{\partial \vec{z}} = \frac{\partial \vec{y}}{\partial \vec{z}} A \vec{x} + \frac{\partial \vec{x}}{\partial \vec{z}} A^T \vec{y}$*

Solution. Let, $\vec{z} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \vec{x}(\vec{z}) = \begin{bmatrix} x_1 \\ \vdots \\ x_m \end{bmatrix}, \vec{y}(\vec{z}) = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}, A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix}$

Also let, $G = A^T \vec{y}$, $F = \vec{x}$ and $H = G^T F = \vec{y}^T A \vec{x}$. As per lecture notes,

$$\frac{\partial H}{\partial \vec{z}} = \frac{\partial G}{\partial \vec{z}} F + \frac{\partial F}{\partial \vec{z}} G \tag{1}$$

Therefore,

$$\frac{\partial F}{\partial \vec{z}} = \frac{\partial \vec{x}}{\partial \vec{z}} \tag{2}$$

Consider,

$$
\begin{aligned}
\frac{\partial G}{\partial \vec{z}} &= \frac{\partial (A^T \vec{y})}{\partial \vec{z}} \\
&= \frac{\partial}{\partial \vec{z}}\left(\begin{bmatrix} \sum_{i=1}^m a_{i1} \cdot y_i \\ \vdots \\ \sum_{i=1}^m a_{im} \cdot y_i \end{bmatrix}\right) \\
&= \begin{bmatrix} \frac{\partial}{\partial z_1}\left(\sum_{i=1}^m a_{i1} \cdot y_i\right) & \cdots & \frac{\partial}{\partial z_1}\left(\sum_{i=1}^m a_{im} \cdot y_i\right) \\ \vdots & \ddots & \vdots \\ \frac{\partial}{\partial z_n}\left(\sum_{i=1}^m a_{i1} \cdot y_i\right) & \cdots & \frac{\partial}{\partial z_n}\left(\sum_{i=1}^m a_{im} \cdot y_i\right) \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^m \left(a_{i1} \cdot \frac{\partial y_i}{\partial z_1}\right) & \cdots & \sum_{i=1}^m \left(a_{im} \cdot \frac{\partial y_i}{\partial z_1}\right) \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^m \left(a_{i1} \cdot \frac{\partial y_i}{\partial z_n}\right) & \cdots & \sum_{i=1}^m \left(a_{im} \cdot \frac{\partial y_i}{\partial z_n}\right) \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial y_1}{\partial z_1} & \cdots & \frac{\partial y_m}{\partial z_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial z_n} & \cdots & \frac{\partial y_m}{\partial z_n} \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{bmatrix} \\
&= \frac{\partial \vec{y}}{\partial \vec{z}} A
\end{aligned}
$$

Therefore,

$$
\frac{\partial G}{\partial \vec{z}} = \frac{\partial \vec{y}}{\partial \vec{z}} A \tag{3}
$$

From equations (1), (2) and (3) above,

$$
\boxed{\frac{\partial (\vec{y}^T A \vec{x})}{\partial \vec{z}} = \frac{\partial \vec{y}}{\partial \vec{z}} A \vec{x} + \frac{\partial \vec{x}}{\partial \vec{z}} A^T \vec{y}}
$$

□

Problem 5. *Suppose $A(x) : \mathbb{R} \to \mathbb{R}^{n \times n}$ is a map from $\mathbb{R}$ to $\mathbb{R}^{n \times n}$.*

Show that if $A(x)$ is invertible, then $\dfrac{dA^{-1}}{dx} = -A^{-1} \dfrac{dA}{dx} A^{-1}$

Solution. Given $A(x) : \mathbb{R} \to \mathbb{R}^{n \times n}$. We have, $AA^{-1} = I_n$. Differentiating both sides w.r.t x gives,

$$
\begin{aligned}
\frac{d}{dx}\left(AA^{-1}\right) &= \frac{dI_n}{dx} \\
\implies \frac{dA}{dx}A^{-1} + A\frac{dA^{-1}}{dx} &= 0 \\
\implies \boxed{\frac{dA^{-1}}{dx} = -A^{-1}\frac{dA}{dx}A^{-1}}
\end{aligned}
$$

□

Problem 6. *Let $\vec{x}$ and $\beta \in \mathbb{R}^p$. Prove that* $\dfrac{\partial \vec{x}^T \beta}{\partial \vec{x}} = \beta$

Solution.

$$
\begin{aligned}
\frac{\partial \vec{x}^T \beta}{\partial \vec{x}} &= \frac{\partial\left(\sum_{i=1}^{p} x_i\beta_i\right)}{\partial \vec{x}} \\
&= \begin{bmatrix} \dfrac{\partial\left(\sum_{i=1}^{p} x_i \cdot \beta_i\right)}{\partial x_1} \\ \vdots \\ \dfrac{\partial\left(\sum_{i=1}^{p} x_i \cdot \beta_i\right)}{\partial x_p} \end{bmatrix} \\
&= \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \\
&= \boxed{\beta}
\end{aligned}
$$

□

Problem 7. *Assume that Y is an n vector but assume that Y depends on X and X depends on some $Z \in \mathbb{R}^q$. Show that*

$$
\frac{\partial Y}{\partial Z} = \frac{\partial X}{\partial Z}\frac{\partial Y}{\partial X}
$$

Does the order matter?

Hint: This means that $X : \mathbb{R}^q \to \mathbb{R}^p$ and $Y : \mathbb{R}^p \to \mathbb{R}^n$.

Solution. Given, $Y(X) \in \mathbb{R}^n$ depends on $X(Z) \in \mathbb{R}^p$ for some $Z \in \mathbb{R}^q$.

$$
\begin{aligned}
\frac{\partial Y}{\partial Z} &= \begin{bmatrix} \frac{\partial\left(y_1(x_1,\cdots x_p)\right)}{\partial z_1} & \cdots & \frac{\partial\left(y_m(x_1,\cdots x_p)\right)}{\partial z_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial\left(y_1(x_1,\cdots x_p)\right)}{\partial z_n} & \cdots & \frac{\partial\left(y_m(x_1,\cdots x_p)\right)}{\partial z_n} \end{bmatrix} \\
&= \begin{bmatrix} \sum_{i=1}^{p} \frac{\partial x_i}{\partial z_1} \cdot \frac{\partial y_1}{\partial x_i} & \cdots & \sum_{i=1}^{p} \frac{\partial x_i}{\partial z_1} \cdot \frac{\partial y_m}{\partial x_i} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{p} \frac{\partial x_i}{\partial z_n} \cdot \frac{\partial y_1}{\partial x_i} & \cdots & \sum_{i=1}^{p} \frac{\partial x_i}{\partial z_n} \cdot \frac{\partial y_m}{\partial x_i} \end{bmatrix} \\
&= \begin{bmatrix} \frac{\partial x_1}{\partial z_1} & \cdots & \frac{\partial x_p}{\partial z_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial z_n} & \cdots & \frac{\partial x_p}{\partial z_n} \end{bmatrix} \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \cdots & \frac{\partial y_m}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial y_1}{\partial x_p} & \cdots & \frac{\partial y_m}{\partial x_p} \end{bmatrix} \\
&= \frac{\partial X}{\partial Z} \frac{\partial Y}{\partial X}
\end{aligned}
$$

Yes, the ordering matters as the dimensions won't allow multiplication if placed in the other way.

□

Problem 8. *Let $z : \mathbb{R}^p \to \mathbb{R}$ be a function that depends on $\vec{x} \in \mathbb{R}^p$ and let Y be a n-vector that depends on $\vec{x} \in \mathbb{R}^p$. Prove that*

$$\frac{\partial}{\partial \vec{x}}(zY) = z\frac{\partial Y}{\partial \vec{x}} + \frac{\partial z}{\partial \vec{x}} Y^T$$

Solution. Given, $\vec{x} \in \mathbb{R}^p$, $z(\vec{x}) \in \mathbb{R}$ and $Y(\vec{x}) \in \mathbb{R}^n$. Therefore,

$$
\begin{aligned}
\frac{\partial}{\partial \vec{x}}(zY) &= \frac{\partial}{\partial \vec{x}}\left(\begin{bmatrix} z \cdot y_1 \\ \vdots \\ z \cdot y_n \end{bmatrix}\right) \\
&= \begin{bmatrix} \dfrac{\partial (z \cdot y_1)}{\partial x_1} & \cdots & \dfrac{\partial (z \cdot y_n)}{\partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial (z \cdot y_1)}{\partial x_p} & \cdots & \dfrac{\partial (z \cdot y_n)}{\partial x_p} \end{bmatrix} \\
&= \begin{bmatrix} z \cdot \dfrac{\partial y_1}{\partial x_1} + y_1 \cdot \dfrac{\partial z}{\partial x_1} & \cdots & z \cdot \dfrac{\partial y_n}{\partial x_1} + y_n \cdot \dfrac{\partial z}{\partial x_1} \\ \vdots & \ddots & \vdots \\ z \cdot \dfrac{\partial y_1}{\partial x_p} + y_1 \cdot \dfrac{\partial z}{\partial x_p} & \cdots & z \cdot \dfrac{\partial y_n}{\partial x_p} + y_n \cdot \dfrac{\partial z}{\partial x_p} \end{bmatrix} \\
&= z\frac{\partial Y}{\partial \vec{x}} + \begin{bmatrix} \dfrac{\partial z}{\partial x_1} \\ \vdots \\ \dfrac{\partial z}{\partial x_p} \end{bmatrix} \begin{bmatrix} y_1 \cdots y_n \end{bmatrix} \\
&= z\frac{\partial Y}{\partial \vec{x}} + \frac{\partial z}{\partial \vec{x}} Y^T
\end{aligned}
$$

□