Bonus Question 2.1 Solutions

- $f_1 = \sin(x_1)\cos(x_2), \quad x \in \mathbb{R}^2$

$$\frac{\partial f_1}{\partial x_1} = \cos(x_1)\cos(x_2)$$

$$\frac{\partial f_1}{\partial x_2} = -\sin(x_1)\sin(x_2)$$

$$\implies J = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \end{bmatrix} = \begin{bmatrix} \cos(x_1)\cos(x_2) & -\sin(x_1)\sin(x_2) \end{bmatrix} \in \mathbb{R}^{1\times 2}$$

BQ 2.2 solutions

- $f_2 = x^\top y \ , \ x, y \in \mathbb{R}^n$

$$\boldsymbol{x}^\top \boldsymbol{y} = \sum_i x_i y_i$$

$$\frac{\partial f_2}{\partial \boldsymbol{x}} = \begin{bmatrix} \frac{\partial f_2}{\partial x_1} & \cdots & \frac{\partial f_2}{\partial x_n} \end{bmatrix} = \begin{bmatrix} y_1 & \cdots & y_n \end{bmatrix} = \boldsymbol{y}^\top \in \mathbb{R}^n$$

$$\frac{\partial f_2}{\partial \boldsymbol{y}} = \begin{bmatrix} \frac{\partial f_2}{\partial y_1} & \cdots & \frac{\partial f_2}{\partial y_n} \end{bmatrix} = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} = \boldsymbol{x}^\top \in \mathbb{R}^n$$

$$\Longrightarrow J = \begin{bmatrix} \frac{\partial f_2}{\partial \boldsymbol{x}} & \frac{\partial f_2}{\partial \boldsymbol{y}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y}^\top & \boldsymbol{x}^\top \end{bmatrix} \in \mathbb{R}^{1 \times 2n}$$

BQ 2.3 Solutions

- $f_3 : \mathbb{R}^n \to \mathbb{R}^{n\times n}$

$$\boldsymbol{x}\boldsymbol{x}^\top = \begin{bmatrix} x_1\boldsymbol{x}^\top \\ x_2\boldsymbol{x}^\top \\ \vdots \\ x_n\boldsymbol{x}^\top \end{bmatrix} = \begin{bmatrix} \boldsymbol{x}x_1 & \boldsymbol{x}x_2 & \cdots & \boldsymbol{x}x_n \end{bmatrix} \in \mathbb{R}^{n\times n}$$

$$\implies \frac{\partial f_3}{\partial x_1} = \underbrace{\begin{bmatrix} \boldsymbol{x}^\top \\ \boldsymbol{0}_n^\top \\ \vdots \\ \boldsymbol{0}_n^\top \end{bmatrix}}_{\in\mathbb{R}^{n\times n}} + \underbrace{\begin{bmatrix} \boldsymbol{x} & \boldsymbol{0}_n & \cdots & \boldsymbol{0}_n \end{bmatrix}}_{\in\mathbb{R}^{n\times n}} \in \mathbb{R}^{n\times n}$$

$$\implies \frac{\partial f_3}{\partial x_i} = \underbrace{\begin{bmatrix} \boldsymbol{0}_{(i-1)\times n} \\ \boldsymbol{x}^\top \\ \boldsymbol{0}_{(n-i+1)\times n} \end{bmatrix}}_{\in\mathbb{R}^{n\times n}} + \underbrace{\begin{bmatrix} \boldsymbol{0}_{n\times(i-1)} & \boldsymbol{x} & \boldsymbol{0}_{n\times(n-i+1)} \end{bmatrix}}_{\in\mathbb{R}^{n\times n}} \in \mathbb{R}^{n\times n}$$

To get the Jacobian, we need to concatenate all partial derivatives $\frac{\partial f_3}{\partial x_i}$ and obtain

$$J = \begin{bmatrix} \frac{\partial f_3}{\partial x_1} & \cdots & \frac{\partial f_3}{\partial x_n} \end{bmatrix} \in \mathbb{R}^{(n\times n)\times n}$$

$f_3(x) = x x^T$, $x \in \mathbb{R}^n$

Eg. $x = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

$x^T = \begin{bmatrix} 2 & 3 \end{bmatrix}$

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix}\begin{bmatrix} 2 & 3 \end{bmatrix} = \begin{bmatrix} 2\,[2 \;\; 3] \\ 3\,[2 \;\; 3] \end{bmatrix}$$

$$= \begin{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} 2 & \begin{bmatrix} 2 \\ 3 \end{bmatrix} 3 \end{bmatrix}$$

BQ 3.1 solution

$$f(t) = \sin(\log(t^T t)) \qquad t \in \mathbb{R}^D$$

$$\frac{\partial f}{\partial t} = \cos(\log(t^T t)) \cdot \frac{1}{t^T t} \cdot 2t^T$$

BQ 3.2 solution

- The trace for $\boldsymbol{T} \in \mathbb{R}^{D\times D}$ is defined as

$$\operatorname{tr}(\boldsymbol{T}) = \sum_{i=1}^{D} T_{ii}$$

A matrix product $\boldsymbol{ST}$ can be written as

$$(\boldsymbol{ST})_{pq} = \sum_i S_{pi}T_{iq}$$

The product $\boldsymbol{AXB}$ contains the elements

$$(\boldsymbol{AXB})_{pq} = \sum_{i=1}^{E}\sum_{j=1}^{F} A_{pi}X_{ij}B_{jq}$$

When we compute the trace, we sum up the diagonal elements of the matrix. Therefore we obtain,

$$\operatorname{tr}(\boldsymbol{AXB}) = \sum_{k=1}^{D}(\boldsymbol{AXB})_{kk} = \sum_{k=1}^{D}\left(\sum_{i=1}^{E}\sum_{j=1}^{F} A_{ki}X_{ij}B_{jk}\right)$$

$$\frac{\partial}{\partial X_{ij}}\operatorname{tr}(\boldsymbol{AXB}) = \sum_k A_{ki}B_{jk} = (\boldsymbol{BA})_{ji}$$

We know that the size of the gradient needs to be of the same size as $\boldsymbol{X}$ (i.e., $E \times F$). Therefore, we have to transpose the result above, such that we finally obtain

$$\frac{\partial}{\partial \boldsymbol{X}}\operatorname{tr}(\boldsymbol{AXB}) = \underbrace{\boldsymbol{A}^{\top}}_{E\times D}\underbrace{\boldsymbol{B}^{\top}}_{D\times F} = (BA)^{T}$$

$$g = \operatorname{tr}(\underbrace{A}_{p\times E}\,\underbrace{X}_{E\times F}\,\underbrace{B}_{F\times q}) \qquad p = q = D$$

Eg

$$\begin{array}{ccc} A & X & B \\ \begin{bmatrix} 2 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}_{3\times 2} & \begin{bmatrix} 2 & 3 & 1 & 4 \\ 1 & 4 & 1 & 6 \end{bmatrix}_{2\times 4} & \begin{bmatrix} 1 & 1 & 0 \\ 2 & 1 & 1 \\ 3 & 1 & 1 \\ 4 & 0 & 1 \end{bmatrix}_{4\times 3} \end{array}$$

$\frac{\partial g}{\partial X}$ has to be of same dimension as X

ie 2×4.