TTK4135 EXERCISE 0

1 ⓐ The gradient is defined as:

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

ⓑ The Jacobian is defined as:

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

when f maps $\mathbb{R}^n$ to $\mathbb{R}^m$.

ⓒ $f : \mathbb{R}^n \mapsto \mathbb{R} \Rightarrow \nabla f(x)$ is a $n \times 1$ vector.

ⓓ $f : \mathbb{R}^n \mapsto \mathbb{R}^m \Rightarrow \frac{\partial f}{\partial x}$ is a $m \times n$ matrix.

2 $f(x) = Ax$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

ⓐ $f(x) = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 \\ a_{21}x_1 + a_{22}x_2 \end{bmatrix}$

$$\underline{\underline{\frac{\partial f}{\partial x} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = A}}$$

(As is expected from the scalar case: $\frac{\partial}{\partial x} ax = a$)

This is the Jacobian, as $f : \mathbb{R}^2 \to \mathbb{R}^2$.

ⓑ $\underline{\frac{\partial Ax}{\partial x} = A}$

$\boxed{3}$ $f(x,y) = x^T G y$, $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $G = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$

ⓐ Since x^T is 1×2 and G is 2×3, the result of $x^T G$ is a 1×3 row vector. That means that the final result f is a ~~1x3 matrix~~ 1×3 row vector multiplied by a 3×1 vector, which is a scalar.

So f is a scalar.

That means that the Jacobian collapses into the gradient (transposed).

ⓑ $$\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix}$$

$$f(x,y) = [x_1 \ x_2]\begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \end{bmatrix} = [g_{11}x_1 + g_{21}x_2 \quad g_{12}x_1 + g_{22}x_2 \quad g_{13}x_1 + g_{23}x_2]$$

$$\cdot y = y_1(g_{11}x_1 + g_{21}x_2) + y_2(g_{12}x_1 + g_{22}x_2) + y_3(g_{13}x_1 + g_{23}x_2)$$

$$\underline{\nabla_x f(x,y)} = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \underline{\begin{bmatrix} y_1 g_{11} + y_2 g_{12} + y_3 g_{13} \\ y_1 g_{21} + y_2 g_{22} + y_3 g_{23} \end{bmatrix} = Gy}$$

ⓒ $$\nabla_y f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial y_1} \\ \frac{\partial f}{\partial y_2} \\ \frac{\partial f}{\partial y_3} \end{bmatrix} = \begin{bmatrix} g_{11}x_1 + g_{21}x_2 \\ g_{12}x_1 + g_{22}x_2 \\ g_{13}x_1 + g_{23}x_2 \end{bmatrix} = G^T x$$

ⓓ $f(x) = x^T H x$

By combining the two results we get:

$$\underline{\nabla f(x) = Hx + H^T x}$$

If H is symmetric we get: $\underline{\nabla f(x) = 2Hx}$

$\boxed{4}$ $\mathcal{L}(x,\lambda,\mu) = x^T G x + \lambda^T (Cx - d) + \mu^T (Ex - h)$

(a) $\nabla_x \mathcal{L}(x,\lambda,\mu) = Gx + G^T x + (\lambda^T C)^T + (\mu^T E)^T$

$$\underline{\underline{\nabla_x \mathcal{L}(x,\lambda,\mu) = Gx + G^T x + C^T \lambda + E^T \mu}}$$

$\cancel{(b)\ \nabla_\lambda \mathcal{L}(x,\lambda,\mu) = Cx - d}$

$\cancel{(c)\ \nabla_\mu \mathcal{L}(x,\lambda,\mu)}$

(b) $\underline{\underline{\nabla_\mu \mathcal{L}(x,\lambda,\mu) = Ex - h}}$

(c) $\underline{\underline{\nabla_\lambda \mathcal{L}(x,\lambda,\mu) = Cx - d}}$