Homework - 1

Exercise 1.1

→ A set A is defined as an open set if $\forall x \in A$, we can find an open neighborhood of x that is contained within A. i.e. $\forall x \in A, \exists \epsilon > 0$ s.t. $N_\epsilon(x) \subset A$

→ Closed set: A set X is said to be closed if every limit point of X is a point of X. i.e. X^c shouldn't contain any limit ~~of~~ points of X. (or) set X is closed iff X^c is open

→ Bounded set: A set X in a metric space (X, d) is defined as bounded if $\exists r \in \mathbb{R}^+$ s.t. $\forall x, y \in X$, $d(x, y) \leq r$

1) For $X = \mathbb{R}^n$, we have $X^c = \phi$

- ϕ is open, $\because \forall x \in \phi$, x is an interior point
  - $\Rightarrow$ X is ~~an~~ a closed set ($\because X^c$ is open)
- ϕ is also closed, $\because$ it ~~sets~~ contains all its limit points
  - $\Rightarrow$ X is an open set ($\because X^c$ is closed)

$\therefore X = \mathbb{R}^n$ is both open and closed set.

- $X = \mathbb{R}^n$ is not bounded

Proof by contradiction: Assume that X is bounded.

- $\Rightarrow \exists$ a finite $r \in \mathbb{R}$ s.t. $\forall x, y \in \mathbb{R}^n$ s.t. $d(x, y) < r$
- $\because r \in \mathbb{R}$, $(r+1) \in \mathbb{R}$. Let $x = (r+1, 0, \ldots, 0)$ & $y = (0, \ldots, 0)$. Clearly $d(x, y) = r+1 > r$, leading to a contradiction

2) We have that $B_n = \{x \in \mathbb{R}^n \mid \|x\| < 1\}$

$\Rightarrow B_n = \{x \in \mathbb{R}^n \mid d(x, 0) < 1\} = B(0, 1)$. ↳ center ↳ radius

- $\forall x \in B_n$, consider $a = 1 - d(0, x)$

Clearly $B(x, a) \subset B(0, 1)$.

$\Rightarrow x$ is an interior point of B_n $\forall x \in B_n$

$\therefore B_n$ is an open set

- B_n is bounded; $\because \forall x, y$, $d(x, 0) < 1$ & $d(y, 0) < 1$ & $d(x, y) \leq d(x, 0) + d(y, 0)$
  $\Rightarrow d(x, y) < 2$ i.e. $\exists r \in \mathbb{R}$ s.t. $d(x, y) < r$ $\forall x, y \in B_n$

3) * $B_n = \{x \in \mathbb{R}^n \mid \|x\| \leq 1\} = B_n[0,1] = \{x \in \mathbb{R}^n \mid d(0,1) \leq 1\}$.

* Let x be a limit point of B_n. If B_n is closed, we need to show that $\forall\ r > 0$, $N_r(x) \setminus \{x\} \cap B_n[0,1] \neq \phi$., or equivalently for any $x \notin B_n[0,1]$ should not be a limit point of $B_n[0,1]$.
* $\therefore\ \forall\ x \notin B_n[0,1]$, consider $r = \frac{d(x,0) - 1}{2}$ $\quad (r > 0 \because d(x,0) > 1)$

Clearly $N_r(x) \cap B_n[0,1] = \phi$.

$\Rightarrow$ x is not a limit point.

$\Rightarrow B_n$ is a closed set.

* B_n is bounded, as similarly to prev q, $\forall\ x, y \in B_n$, $d(x,y) \leq 2$

4) * We have $X = \{x \in [0,1] : f(x) \leq 1\}$, where $f : \mathbb{R} \to \mathbb{R}$ is a continuous f.
* If $\forall\ x \in [0,1]$, $f(x) > 1$, $\Rightarrow X = \phi$, which is both open and closed.
* Else, since f is continuous, the range where $f(x) \leq 1$ will be a closed interval, or the union of multiple closed ranges, which is closed.
* $\because [0,1]$ is bounded, $\Rightarrow X \subseteq [0,1]$ is also bounded.

Exercise 1.2

i) $\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$

* Converting it to Row Reduced Echlon Form, we have

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} 1 & 3 \\ 0 & 0 \end{bmatrix}$$

$\Rightarrow$ Rank = no. of non zero Row = $\underline{\underline{1}}$

* Using Rank-Nullity Th^m, we have

Nullity = Dimension - Rank := $2 - 1 = \underline{1}$

* Column space = $a\begin{bmatrix} 1 \\ 2 \end{bmatrix} + b\begin{bmatrix} 3 \\ 6 \end{bmatrix} = (a + 3b)\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$

* Right null space is the set of all X st. $AX = 0$.

$$\Rightarrow \begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow x_1 + 3x_2 = 0 = 2x_1 + 6x_2$$

$$\Rightarrow x_1 = -3x_2$$

$$\Rightarrow \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = k \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

$\therefore$ Right Null space $= \text{span}\left(\begin{bmatrix} -3 \\ 1 \end{bmatrix} \right)$

ii) $\begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 2 & 4 & 3 \end{bmatrix}$

* Converting it to Row Reduced Echlon Form, we have

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 2 & 4 & 3 \end{bmatrix} \xrightarrow[R_1 - R_2]{R_3 - R_1} \begin{bmatrix} 0 & 2 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \xrightarrow{R_3 - R_2} \begin{bmatrix} 0 & 2 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \xrightarrow{R_2 \leftrightarrow R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$\Rightarrow$ Rank $= \underline{\underline{2}}$, Nullity $= 3 - 2 = \underline{\underline{1}}$

* ~~Column space~~ $\cancel{= [a\ b\ c] \begin{bmatrix} 1 & 3 & 2 \\ 1 & 1 & 1 \\ 2 & 4 & 3 \end{bmatrix} = [a\ b\ c] \begin{bmatrix} 1 & 2 & 1 \\ 1 & 0 & 0 \\ 2 & 0 & 1 \end{bmatrix}}$

$$\cancel{= a \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} + c \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}}$$

$$\cancel{= (c-a) \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} +}$$

* The basis of the column space are the pivot columns of the reduced matrix from the original matrix.

$\therefore$ column space $= \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 4 \end{bmatrix} \right)$

* For the right null space,

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & -2 & -1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = k \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} = \text{span}\left(\begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \\ 1 \end{bmatrix} \right)$$

(iii) $\begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 3 \end{bmatrix}$

* Row-Reduced form:

$$\begin{bmatrix} 1 & 3 & 2 \\ 3 & 1 & 1 \\ 2 & 1 & 3 \end{bmatrix} \xrightarrow{R_2-3R_1} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -8 & -5 \\ 2 & 1 & 3 \end{bmatrix} \xrightarrow{R_3-2R_1} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -8 & -5 \\ 0 & -5 & -1 \end{bmatrix} \xrightarrow{R_3-\frac{5R_2}{8}} \begin{bmatrix} 1 & 3 & 2 \\ 0 & -8 & -5 \\ 0 & 0 & \frac{17}{8} \end{bmatrix}$$

* $\Rightarrow$ Rank $= \underline{3}$, Nullity $= 3-3 = \underline{0}$
* Column space $=$ span $\left\{ \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} \right\}$ (space corresponding to pivot columns of reduced matrix)

* Right null space:

$$\begin{bmatrix} 1 & 3 & 2 \\ 0 & -8 & -5 \\ 0 & 0 & \frac{17}{8} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$\Rightarrow x_1 = 0, x_2 = 0, x_3 = 0$

$\Rightarrow$ Right null space $= \underline{0}$.

Exercise 1.4

A) * For a symmetric matrix, $A = A^T$.

$\Rightarrow$ A is of the form

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}, \quad a, b, c \in \mathbb{R}$$

* The characteristic eq^n of A is given by

$$\begin{vmatrix} a-\lambda & b \\ b & c-\lambda \end{vmatrix} = 0$$

$\Rightarrow (a-\lambda)(c-\lambda) - b^2 = 0$

$\Rightarrow \lambda^2 - (a+c)\lambda + (ac-b^2) = 0$

$\Rightarrow$ Det $= (a+c)^2 - 4(1)(ac-b^2) = a^2+c^2-2ac+4b^2 = (a-c)^2+(2b)^2$

$\therefore$ Det(characteristic eq.) $= (a-c)^2 + (2b)^2 \geq 0$

* $\therefore$ The roots of the characteristic eq^n, are real.
  $\Rightarrow$ $\therefore$ All the eigen values are real.

* If det $= 0$ (i.e. eigen values are not distinct), $\Rightarrow a = c$ & $b = 0$
  $\Rightarrow A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, which is trivially true

Else det > 0.
  $\Rightarrow$ Eigen values are distinct.
  $\Rightarrow$ A is diagonalizable as all eigenvalues are real & distinct.
(Sir's notes)

Exercise 1.5

i) $A = \begin{bmatrix} 3 & 2 & -1 \\ 2 & 4 & -2 \\ -1 & -2 & 5 \end{bmatrix}$

* To solve for the eigen values,
  $|A - \lambda I| = 0$
  $\Rightarrow \lambda^3 - 12\lambda^2 + 38\lambda - 32 = 0$
  $\Rightarrow$ Eigen values are: $\lambda_1 = 1.36$, $\lambda_2 = 3.135$, $\lambda_3 = 7.505$

* Solving for eigen values x, we have
  $Ax = \lambda x \quad \Rightarrow (A - \lambda I)x = 0$

i) $\Rightarrow \begin{bmatrix} 3-1.36 & 2 & -1 \\ 2 & 4-1.36 & -2 \\ -1 & -2 & 5-1.36 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} -4.125 \\ 3.682 \\ 1 \end{bmatrix}\right\}$

ii) $\begin{bmatrix} 3-3.135 & 2 & -1 \\ 2 & 4-3.135 & -2 \\ -1 & -2 & 5-3.135 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} 0.762 \\ 8.552 \\ 1 \end{bmatrix}\right\}$

iii) $\begin{bmatrix} 3-7.505 & 2 & -1 \\ 2 & 4-7.505 & -2 \\ -1 & -2 & 5-7.505 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} -0.637 \\ -6.934 \\ 1 \end{bmatrix}\right\}$

* Since A is symmetric and positive definite (all eigenvalues are +ve)
we can spectral decomposition to find $A^{\frac{1}{2}}$.

$$A^{\frac{1}{2}} = P \begin{bmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \sqrt{\lambda_2} & 0 \\ 0 & 0 & \sqrt{\lambda_3} \end{bmatrix} P^{-1}$$

$$= \begin{bmatrix} -4.125 & 0.762 & -0.637 \\ 3.882 & 0.552 & -0.934 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1.166 & 0 & 0 \\ 0 & 1.770 & 0 \\ 0 & 0 & 2.739 \end{bmatrix} \begin{bmatrix} -4.125 & 3.882 & 1 \\ 0.762 & 0.552 & 1 \\ -0.637 & -0.934 & 1 \end{bmatrix}^{-1}$$

$$\Rightarrow A^{\frac{1}{2}} = \begin{bmatrix} 1.632 & 0.545 & -0.195 \\ 0.545 & 1.886 & -0.468 \\ -0.195 & -0.468 & 2.177 \end{bmatrix}$$

ii) $A = \begin{bmatrix} 3 & -2 & -1 \\ -2 & 4 & -2 \\ -1 & -2 & 3 \end{bmatrix}$

* Solving for eigenvalues, we have

$$\det(A - \lambda I) = 0$$

$$\Rightarrow -\lambda^3 + 10\lambda^2 - 24\lambda = 0$$

$$\Rightarrow \lambda_1 = 0, \quad \lambda_2 = 4, \quad \lambda_3 = 6$$

* Solving for eigenvectors, we have

i) $\begin{bmatrix} 3-0 & -2 & -1 \\ -2 & 4-0 & -2 \\ -1 & -2 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}\right\}$

ii) $\begin{bmatrix} 3-4 & -2 & -1 \\ -2 & 4-4 & -2 \\ -1 & -2 & 3-4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right\}$

iii) $\begin{bmatrix} 3-6 & -2 & -1 \\ -2 & 4-6 & -2 \\ -1 & -2 & 3-6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \quad \Rightarrow \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \text{span}\left\{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right\}$

* Since A is symmetric and positive semi-definite, we can use spectral decomposition to find $A^{\frac{1}{2}}$.

$$A^{\frac{1}{2}} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{0} & 0 & 0 \\ 0 & \sqrt{4} & 0 \\ 0 & 0 & \sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}^{-1}$$

$$\Rightarrow A^{\frac{1}{2}} = \cancel{\begin{bmatrix} -0.225 & 2.45 & -2.225 \\ -2.45 & 4.89 & -2.45 \\ -2.225 & 2.45 & -0.226 \end{bmatrix}} \begin{bmatrix} 1.41 & -0.816 & -0.591 \\ -0.816 & 1.633 & -0.816 \\ -0.591 & -0.816 & 1.41 \end{bmatrix}$$

Exercise 1.6

1) $f(\underline{x}) = \underline{a}^T \underline{x}$, $f: \mathbb{R}^n \to \mathbb{R}$

* $\Rightarrow f(\underline{x}) = \sum_{i=1}^{n} a_i x_i$

$\Rightarrow \cancel{\frac{df(x)}{dx}} \; \frac{\delta f(x)}{\delta x_i} = a_i$

* And since

$$\left[\nabla f(x) = \left(\frac{\delta f(x)}{\delta x_1}, \dots, \frac{\delta f(x)}{\delta x_n}\right)^T\right] \quad \text{— ①}$$

$$\Rightarrow \nabla f(x) = (a_1, a_2, \dots, a_n)^T = \underline{a}$$

2) $f: \mathbb{R}^n : \mathbb{R}^m$, $f(\underline{x}) = Ax$

* $\Rightarrow f(\underline{x}) = \begin{bmatrix} \sum_{i=1}^{n} a_{1i} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi} x_i \end{bmatrix}$

$$\therefore \frac{\delta f(x)}{\delta (x_i)} = \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix}$$

* $\therefore \nabla f(\underline{x}) = \begin{bmatrix} a_{1i} & a_{12} & \dots & a_{1n} \\ \vdots & | & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}^T = \underline{A^T}$ (Using ①)

3) $f: \mathbb{R}^n \to \mathbb{R}$, $f(\underline{x}) = \underline{x}^T A \underline{x}$.

* $\Rightarrow f(x) = \left(\sum_{j=1}^{n} x_j\right)\left(\sum_{i=1}^{n} a_{ji} x_i\right)$

$$= \sum_{j=1}^{n} \sum_{i=1}^{n} x_j a_{ji} x_i$$

$$= \sum_{j=1}^{n} \left(x_j a_{j1} x_1 + \sum_{i=2}^{n} x_j a_{ji} x_i\right)$$

$$= x_1 \left(\sum_{j=1}^{n} x_j a_{j1}\right) + \sum_{i=2}^{n} x_1 a_{1i} x_i + \sum_{i=2}^{n} \sum_{j=2}^{n} x_j a_{ji} x_i$$

$$\Rightarrow f(x) = x_1^2 a_{11} + x_1 \left(\sum_{j=2}^{n} x_j a_{j1} + \sum_{i=2}^{n} x_i a_{1i}\right) + \sum_{i=2}^{n} \sum_{j=2}^{n} x_j a_{ji} x_i$$

* $\therefore \dfrac{\delta f(\underline{x})}{\delta x_1} = 2x_1 a_{11} + \sum_{j=2}^{n} x_j a_{j1} + \sum_{i=2}^{n} x_i a_{1i}$

$$= \sum_{i=1}^{n} x_i a_{1i} + \sum_{j=1}^{n} a_{j1} x_j = \sum_{i=1}^{n} x_i (a_{1i} + a_{i1})$$

* $\therefore$ Using ①

$$\nabla f(x) = \begin{bmatrix} \sum_{i=1}^{n} x_i a_{1i} + \sum_{j=1}^{n} a_{j1} x_j \\ \vdots \\ \sum_{i=1}^{n} x_i a_{ni} + \sum_{j=1}^{n} a_{jn} x_j \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^{n} x_i (a_{1i} + a_{i1}) \\ \vdots \\ \sum_{i=1}^{n} x_i (a_{ni} + a_{in}) \end{bmatrix} = \underline{\underline{(A + A^T)\underline{x}}}$$

4) $f: \mathbb{R}^n \to \mathbb{R}$, $f(\underline{x}) = \dfrac{e^{a^T x}}{1 + e^{a^T \underline{x}}}$

* Let $g(\underline{x}) = a^T \underline{x}$ & $h(x) = \dfrac{e^x}{1+e^x}$ $\left(g'(\underline{x}) = a \;,\; h'(x) = \dfrac{e^x}{(1+e^x)^2}\right)$

$\Rightarrow f(\underline{x}) = h(g(\underline{x}))$

$\Rightarrow f'(x) = h'(g(x)) \times g'(x)$

$$\underline{\underline{\nabla f(x) = \left(\frac{e^{a^T x}}{1 + e^{a^T x}}\right) \times a}}$$

5) $f : \mathbb{R}^n \to \mathbb{R}$, $f(\underline{x}) = \|A\underline{x} - \underline{b}\|^2$

* $\Rightarrow f(\underline{x}) = (\sum_{i=1}^{n} a_{1i} x_i - b_1)^2 + (\sum_{i=1}^{n} a_{2i} x_i - b_2)^2 + \dots + (\sum_{i=1}^{n} a_{ni} x_i - b_n)^2$

$\Rightarrow \frac{\delta(f(\underline{x}))}{\delta(x_i)} = 2(a_{1i})(\sum_{i=1}^{n} a_{1i} x_i - b_1) + \dots + 2(a_{ni})(\sum_{i=1}^{n} a_{ni} x_i - b_n)$

$= 2(A_i)^T(Ax - b)$, where $A_i = \begin{bmatrix} a_{1i} \\ a_{2i} \\ a_{ni} \end{bmatrix}$

* $\therefore$ From ①,

$\nabla f(x) = [\, 2(A_1)^T(Ax-b) \quad 2A_2^T(Ax-b) \quad \dots \quad 2A_n^T(Ax-b) \,]^T$

$= 2\left((Ax-b)^T A\right)^T$

$= 2A^T(Ax-b)$

$\Rightarrow \nabla f(x) = \cancel{\quad}\ \underline{\underline{2\|A\|^2 x - A^T b}}$

6) $f : \mathbb{R}^n \to \mathbb{R}^n$, $f_i(x) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}}$

* $f(x) = \left[\frac{e^{x_1}}{e^{x_1} + e^{x_2} + \dots + e^{x_n}}, \frac{e^{x_2}}{\sum_{j=1}^{n} e^{x_j}}, \dots, \frac{e^{x_n}}{\sum_{j=1}^{n} x_j} \right]^T$

$\Rightarrow \frac{\delta f(\underline{x})}{\delta x_1} = \left[\frac{e^{x_1}(e^{x_2} + e^{x_3} + \dots + e^{x_n})}{(\sum_{j=1}^{n} e^{x_j})^2}, \frac{-e^{x_1} e^{x_2}}{(\sum_{j=1}^{n} e^{x_j})^2}, \frac{-e^{x_1} e^{x_3}}{(\sum_{j=1}^{n} e^{x_j})^2}, \dots, \frac{-e^{x_1} e^{x_n}}{(\sum_{j=1}^{n} e^{x_j})^2} \right]^T$

* $\therefore$ from ①,

$$\nabla f(x) = \begin{bmatrix} \frac{e^{x_1}(e^{x_2} + \dots + e^{x_n})}{(\sum_{j=1}^{n} e^{x_j})^2} & \frac{-e^{x_1} e^{x_2}}{(\sum_{j=1}^{n} e^{x_j})^2} & \dots & - \\ \frac{-e^{x_1} e^{x_2}}{(\sum_{j=1}^{n} e^{x_j})^2} & \frac{e^{x_2}(e^{x_1} + e^{x_3} + \dots + e^{x_n})}{(\sum_{j=1}^{n} e^{x_j})^2} & \ddots & \\ \vdots & \vdots & & \\ \frac{-e^{x_1} e^{x_n}}{(\sum_{j=1}^{n} e^{x_j})^2} & \vdots & & \frac{e^{x_n}(e^{x_1} + \dots + e^{x_{n-1}})}{(\sum_{j=1}^{n} e^{x_j})^2} \end{bmatrix}^T$$

7) $f : \mathbb{R}^n_{>0} \to \mathbb{R}$, $f(\underline{x}) = \sum_{i=1}^{n} x_i \ln \frac{x_i}{\alpha_i}$.

* $\Rightarrow f(\underline{x}) = \sum_{i=1}^{n} (x_i (\ln(x_i) - \ln(\alpha_i)))$

$$\Rightarrow \frac{\partial(f(x))}{\partial x_i} = \ln(x_i) + 1 - \ln(\alpha_i)$$
$$= \ln\left(\frac{e\, x_i}{\alpha_i}\right)$$

* $\therefore \nabla f(\underline{x}) = \left[\ln\left(\frac{e\, x_1}{\alpha_1}\right), \ln\left(\frac{e\, x_2}{\alpha_2}\right), \ldots \ldots \ln\left(\frac{e\, x_n}{\alpha_n}\right)\right]^T$