7. $K=2$, $y_n \to y_n' = 2y_n - 3$ logistic regression $\to \vec{w}_{lr} \Rightarrow$ function of $(\vec{w}_1^*, \vec{w}_2^*)$

↑ opt. sol from MLR

$\because K=2 \therefore y_n \in \{1,2\}$

$$\begin{cases} \text{if } y_n = 1, \text{ then } y_n' = 2-3 = -1 \\ \text{if } y_n = 2, \text{ then } y_n' = 4-3 = 1 \end{cases}$$

Thus, we form the binary classification as:

$$\begin{cases} y_n = 1 \longrightarrow y_n' = -1 \\ y_n = 2 \longrightarrow y_n' = 1 \end{cases}$$

Since $P(y=k \mid \vec{x}) = \dfrac{e^{\vec{w}_k^{*T}\vec{x}}}{e^{\vec{w}_1^{*T}\vec{x}} + e^{\vec{w}_2^{*T}\vec{x}}}$ (MLR)

and

$P(y'=1 \mid \vec{x}) = \dfrac{1}{1+e^{-\vec{w}_{lr}^T\vec{x}}}$, $P(y'=-1 \mid \vec{x}) = \dfrac{1}{1+e^{\vec{w}_{lr}^T\vec{x}}}$ (logistic regression)

$\to$ from $\begin{cases} P(y=1 \mid \vec{x}) = P(y'=-1 \mid \vec{x}) \\ P(y=2 \mid \vec{x}) = P(y'=1 \mid \vec{x}) \end{cases}$

we can get:

$$\begin{cases} \dfrac{e^{\vec{w}_1^{*T}\vec{x}}}{e^{\vec{w}_1^{*T}\vec{x}} + e^{\vec{w}_2^{*T}\vec{x}}} = \dfrac{1}{1+e^{\vec{w}_{lr}^T\vec{x}}} \quad \text{— ①} \\[2ex] \dfrac{e^{\vec{w}_2^{*T}\vec{x}}}{e^{\vec{w}_1^{*T}\vec{x}} + e^{\vec{w}_2^{*T}\vec{x}}} = \dfrac{1}{1+e^{-\vec{w}_{lr}^T\vec{x}}} \quad \text{— ②} \end{cases}$$

①: $e^{\vec{w}_1^{*T}\vec{x}}(1+e^{\vec{w}_{lr}^T\vec{x}}) = e^{\vec{w}_1^{*T}\vec{x}} + e^{\vec{w}_2^{*T}\vec{x}}$

$\to \cancel{e^{\vec{w}_1^{*T}\vec{x}}} + e^{\vec{w}_1^{*T}\vec{x}} e^{\vec{w}_{lr}^T\vec{x}} = \cancel{e^{\vec{w}_1^{*T}\vec{x}}} + e^{\vec{w}_2^{*T}\vec{x}}$

$\to e^{\vec{w}_1^{*T}\vec{x} + \vec{w}_{lr}^T\vec{x}} = e^{\vec{w}_2^{*T}\vec{x}}$

$\to \vec{w}_1^{*T}\vec{x} + \vec{w}_{lr}^T\vec{x} = \vec{w}_2^{*T}\vec{x}$

$\to \vec{w}_{lr}^T\vec{x} = -\vec{w}_1^{*T}\vec{x} + \vec{w}_2^{*T}\vec{x} = (-\vec{w}_1^* + \vec{w}_2^*)^T\vec{x}$

Thus, $\vec{w}_{lr} = -\vec{w}_1^* + \vec{w}_2^*$ □