최소제곱법을 이용한 선형 회귀에 대해 다음 표와 주어진 식을 보고 물음에 답하시오.

x	x_0	x_1	$\cdots$	x_m
y	y_0	y_1	$\cdots$	y_m

$$S = \underbrace{\begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_r & 0 & \cdots & 0 \end{bmatrix}}_{n} \ (m)$$

$$S^+ = \underbrace{\begin{bmatrix} 1/\sigma_1 & & & & & \\ & 1/\sigma_2 & & & & \\ & & \ddots & & & \\ & & & 1/\sigma_r & 0 & \cdots \end{bmatrix}}_{m} \ (n)$$

$$\begin{aligned} ax_0 + b &= y_0 \\ ax_1 + b &= y_1 \\ ax_2 + b &= y_2 \\ \vdots \quad & \quad \vdots \\ ax_m + b &= y_m \end{aligned} \quad \Rightarrow \quad \underbrace{\begin{bmatrix} x_0 & 1 \\ x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{\boldsymbol{b}}$$

(1) A의 열 벡터가 독립이면 $A^T A$는 symmetric positive definite 임을 보이시오.

(2) $A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$, $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$ 일 때, $F(\boldsymbol{x}) = \boldsymbol{x}^T A \boldsymbol{x}$라 하자. $\dfrac{\partial F}{\partial x}$ 와 $\dfrac{\partial F}{\partial y}$ 를 구하시오.

(3) $\|\boldsymbol{v}\|_2^2 = \boldsymbol{v}^T \boldsymbol{v}$라 할 때, ℓ_2 오차함수는 $\|A\boldsymbol{x} - \boldsymbol{b}\|^2$ 으로 나타낼 수 있다. 이때 최소제곱해 $\boldsymbol{x}^*$는 $A^T A \boldsymbol{x}^* = A^T \boldsymbol{b}$를 만족함을 보이시오.

(4) A는 특잇값 분해(SVD)에 의해 $A = USV^T$ 이고 $A^+ = VS^+U^T$ 라 하자. 여기서 $(S)_{ii} \neq 0$일 때 $(S^+)_{ii} = 1/(S)_{ii}$를 만족하고 S가 $m \times n$ 행렬일 때 S^+는 $n \times m$ 행렬이다. 그러면 $\boldsymbol{x}^+ = A^+ \boldsymbol{b}$는 최소제곱해임을 보이시오.

(1) if $A^T A$ is positive definite, then

$x^T A^T A x > 0$ for all $x \neq 0$

$\Downarrow$

$(Ax)^T (Ax) > 0$

$\Downarrow$

$\|Ax\|^2 > 0$ $\qquad Ax = \begin{bmatrix} ax_1 + b \\ ax_2 + b \\ \vdots \\ ax_m + b \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$

$\Downarrow$

$(ax_1+b)^2 + (ax_2+b)^2 + \cdots + (ax_m+b)^2 > 0$

$y_1^2 + y_2^2 + \cdots + y_m^2 > 0$ by independent

$$A^T A = \begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix} = \begin{bmatrix} x_1^2 + x_2^2 + \cdots + x_m^2 & x_1 + x_2 + \cdots + x_m \\ x_1 + x_2 + \cdots + x_m & 1 \end{bmatrix}$$

symmetric

(2) $F(x) = x^T A x = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

$= \begin{bmatrix} ax + by & bx + cy \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax^2 + 2bxy + cy^2 \end{bmatrix}$

$\dfrac{\partial F}{\partial x} = 2ax + 2by = \begin{bmatrix} 2a & 2b \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

$\dfrac{\partial F}{\partial y} = 2cy + 2bx = \begin{bmatrix} 2b & 2c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

(3) $\ell_2 = \|Ax - b\|^2$

$\nabla \ell_2 = 2A^T (Ax - b) = 0$

$2A^T A x - 2A^T b = 0$

$2A^T A x = 2A^T b$

$A^T A x^* = A^T b$

(4) $A = USV^T$

$Ax = b$

$USV^T x = b$

$x = VS^{-1}U^T b$

$x = VS^+U^T b$

$x^+ = A^+ b$

by (3) $A^T A A^+ b = A^T b$

$(USV^T)^T (USV^T)(VS^+U^T) b = (USV^T)^T b$

$(VS^T \underbrace{U^T)(U}_{I} S \underbrace{V^T)(V}_{I} S^+ U^T) b = (VS^T U^T) b$

$(VS^T \underbrace{S S^+}_{I} U^T) b = (VS^T U^T) b$

$(VS^T U^T) b = (VS^T U^T) b$

성립