2.1

この節では，本章の準備として $W \in \mathbb{R}^{N\times N}$ を正定値行列として

$$L_0 := \frac{1}{2N}(y - \beta_0 \cdot \mathbb{1} - X\beta)^T W (y - \beta_0 \cdot \mathbb{1} - X\beta) \tag{1}$$

の線形回帰の Lasso 解を求める．そのために，第 1 章での線形回帰同様 X の各列及び y を中心化して，β_0 を 0 にする．具体的に $i \in \{1, \cdots N\}$ に対して重み $W = (w_{i,j})$ を考慮した中心化

$$\begin{aligned}
\bar{X}_k &:= \frac{\sum_{i=1}^N \sum_{j=1}^N w_{i,j} x_{j,k}}{\sum_{i=1}^N \sum_{j=1}^N w_{i,j}} \\
\bar{y} &:= \frac{\sum_{i=1}^N \sum_{j=1}^N w_{i,j} y_j}{\sum_{i=1}^N \sum_{j=1}^N w_{i,j}} \\
x_{i,k} &\leftarrow x_{i,k} - \bar{X}_k \\
y_i &\leftarrow y_i - \bar{y}
\end{aligned}$$

を行うと，各 k に対して

$$\begin{aligned}
\sum_{i=1}^N \sum_{j=1}^N w_{i,j}\left(x_{j,k} - \bar{X}_k\right) &= \sum_{i=1}^N \sum_{j=1}^N w_{i,j} x_{j,k} - \sum_{i=1}^N \sum_{j=1}^N w_{i,j} \bar{X}_k \\
&= \sum_{i=1}^N \sum_{j=1}^N w_{i,j} x_{j,k} - \sum_{i=1}^N \sum_{j=1}^N w_{i,j} \cdot \frac{\sum_{i=1}^N \sum_{j=1}^N w_{i,j} x_{j,k}}{\sum_{i=1}^N \sum_{j=1}^N w_{i,j}} = 0 \\
\sum_{i=1}^N \sum_{j=1}^N w_{i,j}(y_j - \bar{y}) &= \sum_{i=1}^N \sum_{j=1}^N w_{i,j} y_j - \sum_{i=1}^N \sum_{j=1}^N w_{i,j} \cdot \frac{\sum_{i=1}^N \sum_{j=1}^N w_{i,j} y_j}{\sum_{i=1}^N \sum_{j=1}^N w_{i,j}} = 0
\end{aligned}$$

が成立することから $\displaystyle\sum_{i=1}^{N}\sum_{j=1}^{N} w_{i,j}x_{j,k} = \sum_{i=1}^{N}\sum_{j=1}^{N} w_{i,j}y_j = 0$ が成立し，さらに正定値 W が特に対称行列であることを考えれば

$$\begin{aligned}
\frac{\partial L_0}{\partial \beta_0 \cdot \mathbb{1}} &= \frac{\partial (y - \beta_0 \cdot \mathbb{1} - X\beta)}{\partial \beta_0 \cdot \mathbb{1}} \frac{\partial L_0}{\partial (y - \beta_0 \cdot \mathbb{1} - X\beta)} \\
&= \frac{1}{2N} \cdot \frac{\partial (y - \beta_0 \cdot \mathbb{1} - X\beta)}{\partial \beta_0 \cdot \mathbb{1}} (W + W^T)(y - \beta_0 \cdot \mathbb{1} - X\beta) \\
&= -\frac{1}{N} W (y - \beta_0 \cdot \mathbb{1} - X\beta)
\end{aligned}$$

がわかり，

$$y - \beta_0 \cdot \mathbb{1} - X\beta = \begin{pmatrix} y_1 - \beta_0 - \displaystyle\sum_{k=1}^{p} x_{1,k}\beta_k \\ \vdots \\ y_N - \beta_0 - \displaystyle\sum_{k=1}^{p} x_{N,k}\beta_k \end{pmatrix}$$

であることから

$$-\frac{1}{N} W (y - \beta_0 \cdot \mathbb{1} - X\beta) = -\frac{1}{N} \begin{pmatrix} \displaystyle\sum_{j=1}^{N} w_{1,j}y_j - \beta_0 \sum_{j=1}^{N} w_{1,j} - \sum_{j=1}^{N} w_{1,j} \sum_{k=1}^{p} x_{j,k}\beta_k \\ \vdots \\ \displaystyle\sum_{j=1}^{N} w_{N,j}y_j - \beta_0 \sum_{j=1}^{N} w_{N,j} - \sum_{j=1}^{N} w_{N,j} \sum_{k=1}^{p} x_{j,k}\beta_k \end{pmatrix}$$

が成立する必要がある．(1) 式が $(\hat{\beta}, \hat{\beta}_0)$ で最小となるには $\dfrac{\partial L_0}{\partial \beta_0 \cdot \mathbb{1}} = \mathbf{0}$ となる必要があるので，最小値を与える $\hat{\beta}_0$ が満たす条件は

$$
\begin{aligned}
\hat{\beta}_0 \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} &= \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} y_j - \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \sum_{k=1}^{p} x_{j,k} \beta_k \\
\therefore \hat{\beta}_0 &= \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} y_j}{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j}} - \frac{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \sum_{k=1}^{p} x_{j,k} \beta_k}{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j}} \\
&= \bar{y} - \frac{\sum_{k=1}^{p} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} x_{j,k} \beta_k}{\sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j}} \\
&= \bar{y} - \sum_{k=1}^{p} \bar{X}_k \hat{\beta}_k = 0
\end{aligned}
$$

となることがわかる.

以上の中心化により，改めて $\beta_0 = 0$ とした上で

$$
L(\beta) := \frac{1}{2N} (y - X\beta)^T W (y - X\beta) + \lambda \|\beta\|_1
$$

の最小化を考えるが，これは $W = M^T M$ と Cholesky 分解すれば，$V := MX, u = My$ とおくことにより

$$
\begin{aligned}
L(\beta) &= \frac{1}{2N} (y - X\beta)^T (M^T M)(y - X\beta) + \lambda \|\beta\|_1 \\
&= \frac{1}{2N} ((y - X\beta)^T M^T)(M(y - X\beta)) + \lambda \|\beta\|_1 \\
&= \frac{1}{2N} (My - MX\beta)^T (My - MX\beta) + \lambda \|\beta\|_1 \\
&= \frac{1}{2N} \|u - V\beta\|_2^2 + \lambda \|\beta\|_1
\end{aligned}
$$

となるため，結局通常の線形回帰における Lasso

$$
L(\beta) = \frac{1}{2N} \|u - V\beta\|_2^2 + \lambda \|\beta\|_1
$$

に帰着することができる.