1 Ex.3.5

 β^c 与 β 之间的关系可以写成

$$eta_0^c = eta_0 + \sum_{j=1}^p ar{x}_j eta_j \ eta_i^c = eta_j, j = 1, 2, \cdots, p$$

将 $\sum_{i=1}^{N} [y_i - \beta_0^c - \sum_{j=1}^{p} (x_{ij} - \bar{x}_j)\beta_j^c]^2 + \lambda \sum_{j=1}^{p} (\beta_j^c)^2$ 关于 β_0^c 求导,有

$$eta_0^c - rac{1}{N}\sum_{i=1}^N y_i = 0$$

于是 $\beta_0^c = \bar{y}$ 。令

$$egin{aligned} ilde{y}_i &= y_i - eta_0^c, \ ilde{x}_{ij} &= x_{ij} - ar{x}_j, \end{aligned} \ \hat{eta}^c &= rg\min_{eta^c} \left\{ \sum_{i=1}^N [y_i - eta_0^c - \sum_{j=1}^p (x_{ij} - ar{x}_j) eta_j^c]^2 + \lambda \sum_{j=1}^p (eta_j^c)^2
ight\} \end{aligned}$$

可以写为

$$\hat{eta}^c = rg\min_{eta^c} \left\{ \sum_{i=1}^N [ilde{y}_i - \sum_{j=1}^p (ilde{x}_{ij})eta_j^c]^2 + \lambda \sum_{j=1}^p (eta_j^c)^2
ight\}$$

上式形式与原始ridge回归问题一致,于是该问题的解为

$$\hat{eta}_c = (\mathbf{ ilde{X}}^T\mathbf{ ilde{X}} + \lambda \mathbf{I})^{-1}\mathbf{ ilde{X}}^T\mathbf{y}$$

将上述对应关系代入lasso,

$$egin{aligned} \hat{eta}^{ ext{lasso}} &= rg\min_{eta^c} \left\{ \sum_{i=1}^N [y_i - eta^c_0 - \sum_{j=1}^p (x_{ij} - ar{x}_j)eta^c_j]^2
ight\} \ s.\,t. \sum_{j=1}^p |eta^c_j| \leq t \end{aligned}$$

与原先lasso的形式完全一致。

2 Ex.3.6

根据高斯分布概率密度函数的定义,有

$$P(\mathbf{y} \mid \mathbf{X}, eta) = rac{1}{(2\pi)^{rac{N}{2}} \sigma^N} \exp\left(-rac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}eta)^T (\mathbf{y} - \mathbf{X}eta)
ight)$$

$$P(eta) = rac{1}{(2\pi)^{rac{p}{2}} au^{rac{p}{2}}} \exp\left(-rac{1}{2 au}eta^Teta
ight)$$

由贝叶斯公式

$$egin{aligned} P(eta \mid \mathbf{y}, \mathbf{X}) &= rac{P(\mathbf{y} \mid eta, \mathbf{X}) P(eta)}{P(\mathbf{y})} \ \log P(eta \mid \mathbf{y}, \mathbf{X}) &= -rac{1}{2\sigma^2} (\mathbf{y} - \mathbf{X}eta)^T (\mathbf{y} - \mathbf{X}eta) - rac{1}{2 au} eta^T eta + C \end{aligned}$$

 $\diamondsuit \lambda = \frac{\sigma^2}{\tau}$, β 的均值与众数对应着ridge的解。即

$$\min_{eta} (\mathbf{y} - \mathbf{X}eta)^T (\mathbf{y} - \mathbf{X}eta) + \lambdaeta^Teta,$$

3 Ex.3.7

根据高斯分布概率密度函数的定义,有

$$P(\mathbf{y} \mid \beta) = \frac{1}{(2\pi)^{\frac{N}{2}} \sigma^N} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^N \left(y_i - \beta_0 - \sum_j x_{ij} \beta_j\right)^2\right)$$
$$P(\beta) = \frac{1}{(2\pi)^{\frac{p}{2}} \sigma^p} \exp\left(-\frac{1}{2\tau^2} \sum_{j=1}^p \beta_j^2\right)$$

由贝叶斯公式,

$$P(\beta \mid \mathbf{y}) = \frac{P(\mathbf{y} \mid \beta)P(\beta)}{P(\mathbf{y})}$$

代入概率密度, 并取负对数

$$egin{aligned} -\ln(P(eta \mid \mathbf{y})) &= rac{1}{2\sigma^2} \sum_{i=1}^N (y_i - eta_0 - \sum_j x_{ij} eta_j)^2 + rac{1}{2 au^2} \sum_{j=1}^p eta_j^2 + C \ &= rac{1}{2\sigma^2} \Biggl(\sum_{i=1}^N (y_i - eta_0 - \sum_j x_{ij} eta_j)^2 + \lambda \sum_{j=1}^p eta_j^2 \Biggr) + C \end{aligned}$$

其中 $\lambda = \frac{\sigma^2}{\tau^2}$,所以 $P(\beta \mid \mathbf{y})$ 的负对数与 $\sum_{i=1}^N (y_i - \beta_0 - \sum_j x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^p \beta_j^2$ 呈正比例关系。