Homework of Chapter 3

Cheng Chen, 1130339005

December 26, 2014

Ex. 3.5

Solution.

$$\begin{aligned}
&\underset{\beta^c}{argmin}\left\{\sum_{i=1}^{N}[y_i-\beta_0^c-\sum_{j=1}^{p}(x_{ij}-\overline{x}_j)\beta_j^c]^2+\lambda\sum_{j=1}^{p}\beta_j^{c2}\right\}\\
=&\underset{\beta^c}{argmin}\left\{\sum_{i=1}^{N}(y_i-\beta_0^c-\sum_{j=1}^{p}x_{ij}\beta_j^c+\sum_{j=1}^{p}\overline{x}_j\beta_j^c)^2+\lambda\sum_{j=1}^{p}\beta_j^{c2}\right\}\\
=&\underset{\beta^c}{argmin}\left\{\sum_{i=1}^{N}[y_i-(\beta_0^c-\sum_{j=1}^{p}\overline{x}_j\beta_j^c)-\sum_{j=1}^{p}x_{ij}\beta_j^c]^2+\lambda\sum_{j=1}^{p}\beta_j^{c2}\right\}
\end{aligned}$$

By comparing the above expression with (3.41), we can get the correspondence between β^c and the original β:

$$\beta_0=\beta_0^c-\sum_{j=1}^{p}\overline{x}_j\beta_j^c$$

$$\beta_j=\beta_j^c \quad j=1,2,...,p$$

So, the new problem is equivalent to the ridge regression problem (3.41). For the lasso, we have the same result:

$$\beta_0=\beta_0^c-\sum_{j=1}^{p}\overline{x}_j\beta_j^c$$

$$\beta_j=\beta_j^c \quad j=1,2,...,p$$

and

$$
\begin{aligned}
&\underset{\beta^c}{argmin}\left\{\sum_{i=1}^{N}[y_i-\beta_0^c-\sum_{j=1}^{p}(x_{ij}-\overline{x}_j)\beta_j^c]^2+\lambda\sum_{j=1}^{p}|\beta_j^c|\right\}\\
=&\underset{\beta}{argmin}\left\{\sum_{i=1}^{N}[y_i-\beta_0-\sum_{j=1}^{p}x_{ij}]^2+\lambda\sum_{j=1}^{p}|\beta_j^c|\right\}
\end{aligned}
$$

$\square$

Ex. 3.7

Solution. As $\beta_j \sim N(0,\tau^2), j = 1, ..., p$ and $y_i \sim N(\beta_0 + x_i^T\beta, \sigma^2), i = 1, 2, ..., N$, we can get:

$$
\begin{aligned}
P\left(y|\beta\right)=&\prod_{i=1}^{N}P\left(y_i|\beta\right)\\
=&\prod_{i=1}^{N}\frac{1}{\sqrt{2\pi}\sigma}\exp\left\{-\frac{(y_i-(\beta_0+x_i^T\beta))^2}{2\sigma^2}\right\}\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^N}\exp\left\{\sum_{i=1}^{N}-\frac{(y_i-(\beta_0+x_i^T\beta))^2}{2\sigma^2}\right\}
\end{aligned}
$$

and

$$
\begin{aligned}
P\left(\beta\right)=&\prod_{j=1}^{p}P(\beta_j)\\
=&\prod_{j=1}^{p}\frac{1}{\sqrt{2\pi}\tau}\exp\left\{-\frac{\beta_j^2}{2\tau^2}\right\}\\
=&\frac{1}{(\sqrt{2\pi}\tau)^p}\exp\left\{\sum_{j=1}^{p}-\frac{\beta_j^2}{2\tau^2}\right\}
\end{aligned}
$$

By using Bayes Rule, we have:

$$
\begin{aligned}
P(\beta|y)=&\frac{P(y|\beta)P(\beta)}{P(y)}\\
=&\frac{\exp\left\{-\sum_{i=1}^{N}\frac{(y_i-(\beta_0+x_i^T\beta))^2}{2\sigma^2}-\sum_{j=1}^{p}\frac{\beta_j^2}{2\tau^2}\right\}}{(\sqrt{2\pi}\sigma)^N(\sqrt{2\pi}\tau)^pP(y)}
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
-logP(\beta|y) =& -log\frac{P(y|\beta)P(\beta)}{P(y)} \\
=& -log\frac{\exp\left\{-\sum_{i=1}^{N}\frac{(y_i-(\beta_0+x_i^T\beta))^2}{2\sigma^2}-\sum_{j=1}^{p}\frac{\beta_j^2}{2\tau^2}\right\}}{(\sqrt{2\pi}\sigma)^N(\sqrt{2\pi}\tau)^pP(y)} \\
=& \sum_{i=1}^{N}\frac{(y_i-(\beta_0+x_i^T\beta))^2}{2\sigma^2}+\sum_{j=1}^{p}\frac{\beta_j^2}{2\tau^2}-log\{(\sqrt{2\pi}\sigma)^N(\sqrt{2\pi}\tau)^pP(y)\} \\
=& \frac{1}{2\sigma^2}\left\{\sum_{i=1}^{N}(y_i-(\beta_0+x_i^T\beta))^2+\sum_{j=1}^{p}\frac{\sigma^2}{\tau^2}\beta_j^2\right\}+constant \\
\propto& \sum_{i=1}^{N}(y_i-\beta_0-\sum_{j=1}^{p}x_{ij}\beta_j)^2+\lambda\sum_{j=1}^{p}\beta_j^2
\end{aligned}
$$

The proportional holds if we ignore the constant. $\square$