⑥ Starting from (7.37) and (7.38)

$\boxed{\hat\beta \sim N_p(\beta, \sigma^2 \cdot S^{-1})}$ and $\boxed{\beta \sim N_p\left(0, \frac{\sigma^2}{\lambda} \cdot I\right)}$

want posterior of $\boxed{\beta \mid \hat\beta}$

$$\Rightarrow \quad g(\beta \mid \hat\beta) \propto f(\hat\beta \mid \beta) \cdot g(\beta)$$

$$g(\beta \mid \hat\beta) \propto \exp\left\{ -\frac{1}{2} \cdot \left[(\beta - \hat\beta)(\sigma^2 S^{-1})^{-1} \cdot (\beta - \hat\beta) + \beta^T \left(\frac{\sigma^2}{\lambda} \cdot I\right)^{-1} \cdot \beta \right] \right\}$$

$$g(\beta \mid \hat\beta) \propto \exp\left\{ -\frac{1}{2\sigma^2} \left[(\beta - \hat\beta)^T S (\beta - \hat\beta) + \beta^T \lambda I \beta \right] \right\}$$

$$\propto \exp\left\{ -\frac{1}{2\sigma^2} \cdot \left[\beta^T (S + \lambda I)\beta - \hat\beta^T S \beta - \beta^T S \hat\beta \right] \right\}$$

. Since we're working with expression that are proportional to $g(\beta \mid \hat\beta)$, we can multiply any expression (add to the argument inside of $\exp\{\cdot\}$) that does not contain β. We'll take advantage of this, to complete a square.

$$f(\beta \mid \hat\beta) \propto \exp\left\{ -\frac{1}{2\sigma^2} \left[\beta^T (S + \lambda I)\beta - \hat\beta^T S \beta - \beta^T S \hat\beta + \hat\beta^T S (S + \lambda I)^{-1} S \hat\beta \right] \right\}$$

$\vdots$

After some algebra

⋮

$$g(\beta \mid \hat{\beta}) \propto \exp\left\{-\frac{1}{2\sigma^2}\left[\left(\beta - (S+\lambda I)^{-1} S\hat{\beta}\right)^T \cdot (S+\lambda I)^{-1} \cdot \left(\beta - (S+\lambda I)^{-1} S\hat{\beta}\right)\right]\right\}$$

which is a $N_p\left((S+\lambda I)^{-1} S\hat{\beta},\ \sigma^2 (S+\lambda I)^{-1}\right)$

kernel

Thus, $\boxed{E(\beta \mid \hat{\beta}) = (S+\lambda I)^{-1} S\hat{\beta}}$.

- The expected value of your posterior belief on β is equal to the Ridge estimator value.

⑩ Lecture Notes 14.