Stat873 **Exam 2** **Mar. 6, 2024**

Name:

1. In model $Y = X\beta + \epsilon$, $\epsilon \sim N(0, \sigma^2\Sigma)$, $X \in R^{n\times p}$ has full column rank.

   (1) Write out $\widehat{\beta}$, the MVUE for β, and its distribution. (5 points)

   $$\widehat{\beta} = (\Sigma^{-1/2}X)^{+}\Sigma^{-1/2}Y = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y \sim N\left(\beta,\, \sigma^2(X'\Sigma^{-1}X)^{-1}\right).$$

   (2) Write out the two risks, $\mathrm{MSEM}(\widehat{\beta}, \beta)$ and $\mathrm{MSE}(\widehat{\beta}, \beta)$. (10 points)

   $\mathrm{MSEM}(\widehat{\beta}, \beta) = \mathrm{Cov}(\widehat{\beta}) = \sigma^2(X'\Sigma^{-1}X)^{-1}$.
   $\mathrm{MSE}(\widehat{\beta}, \beta) = \mathrm{tr}\left(\mathrm{Cov}(\widehat{\beta})\right) = \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)$.

2. Let $\widehat{\beta}(s) = s\widehat{\beta}$ where $0 < s < 1$ and $\widehat{\beta}$ is in 1 (1).

   (1) Show that $\widehat{\beta}(s)$ is a linear biased estimator for β. (10 points)

   $\widehat{\beta}(s) = s\widehat{\beta} = s(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$ is a linear function of Y.
   $\widehat{\beta}(s) \sim N\left(s\beta,\, \sigma^2 s^2(X'\Sigma^{-1}X)^{-1}\right)$.
   $E(\widehat{\beta}(s)) = s\beta \neq \beta$. So $\widehat{\beta}(s)$ is a biased estimator for β.
   Hence $\widehat{\beta}(s)$ is a linear biased estimator for β.

   (2) Find the risk $\mathrm{MSE}(\widehat{\beta}(s), \beta)$. (10 points)

   $$\begin{aligned}
   \mathrm{MSE}(\widehat{\beta}(s), \beta) &= E[(\widehat{\beta}(s) - \beta)'(\widehat{\beta}(s) - \beta)] = \mathrm{tr}E[(\widehat{\beta}(s) - \beta)(\widehat{\beta}(s) - \beta)'] \\
   &= \mathrm{tr}E[(\widehat{\beta}(s) - s\beta + (s-1)\beta)(\widehat{\beta}(s) - s\beta + (s-1)\beta)'] \\
   &= \mathrm{tr}[\mathrm{Cov}(\widehat{\beta}(s)) + (s-1)^2\beta\beta'] \\
   &= (s-1)^2\beta'\beta + s^2\sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right).
   \end{aligned}$$

   (3) Find a sufficient and necessary condition on s for $\widehat{\beta}(s)$ to dominate $\widehat{\beta}$ by $\mathrm{MSE}(\cdot,\cdot)$.
   Hint: Write $\mathrm{MSE}(\widehat{\beta}(s), \beta) - \mathrm{MSE}(\widehat{\beta}, \beta)$ as a function of s. (20 points)

   $$\begin{aligned}
   \mathrm{MSE}(\widehat{\beta}(s), \beta) - \mathrm{MSE}(\widehat{\beta}, \beta) &= (s-1)^2\beta'\beta + (s^2-1)\sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right) \\
   &= as^2 + bs + c
   \end{aligned}$$
   where $a = \beta'\beta + \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)$, $b = -2\beta'\beta$ and $c = \beta'\beta - \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)$.
   With $a > 0$, $s_1 = \frac{-b-\sqrt{b^2-4ac}}{2a} = \frac{\beta'\beta - \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)}{\beta'\beta + \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)}$ and

   $$s_2 = \frac{-b+\sqrt{b^2-4ac}}{2a} = \frac{\beta'\beta + \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)}{\beta'\beta + \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right)} = 1,$$

   $$\mathrm{MSE}(\widehat{\beta}(s), \beta) - \mathrm{MSE}(\widehat{\beta}, \beta) < 0 \Longleftrightarrow \begin{cases} 0 < s < 1, & s_1 \le 0 \Longleftrightarrow \beta'\beta \le \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right) \\ s_1 \le s < 1, & s_1 > 0 \Longleftrightarrow \beta'\beta > \sigma^2\mathrm{tr}\left((X'\Sigma^{-1}X)^{-1}\right) \end{cases}.$$

3. Consider Bayesian approach for estimating β in 1.

(1) Find likelihood function $L(\beta)$ and $f_1(\beta)$ such that $L(\beta) \propto f_1(\beta)$. (10 points)

$$\begin{aligned} L(\beta) &= \frac{1}{(2\pi)^{n/2}|\sigma^2\Sigma|^{1/2}} \exp\left[-\frac{1}{2\sigma^2}(Y - X\beta)'\Sigma^{-1}(Y - X\beta)\right] \\ &= \frac{1}{(2\pi)|\sigma^2\Sigma|^{1/2}} \exp\left(-\frac{1}{2\sigma^2}Y'\Sigma^{-1}Y\right) \exp\left(-\frac{\beta' X'\Sigma^{-1}X\beta - 2\beta' X'\Sigma^{-1}Y}{2\sigma^2}\right) \\ &\propto f_1(\beta) \qquad \text{where} \\ f_1(\beta) &= \exp\left(-\frac{\beta' X'\Sigma^{-1}X\beta - 2\beta' X'\Sigma^{-1}Y}{2\sigma^2}\right). \end{aligned}$$

(2) With prior $\beta \sim N(\beta_0, \Sigma_0)$, its pdf $f_\beta(\beta) \propto f_0(\beta)$. This $f_0(\beta)$ was given in the lecture. Based on the relation $f_{\beta|Y}(\beta) \propto f_0(\beta)f_1(\beta)$, find the posterior distribution of β given Y. (20 points)

With $f_0(\beta) = \exp\left(\frac{\beta'\Sigma_0^{-1}\beta - 2\beta'\Sigma_0^{-1}\beta_0}{-2}\right)$ from class, and $f_1(\beta)$ in (1),

$$f_0(\beta)f_1(\beta) = \exp\left[\frac{\beta'\left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)\beta - 2\beta'\left(\Sigma_0^{-1}\beta_0 + \frac{X'\Sigma^{-1}Y}{\sigma^2}\right)}{-2}\right]$$

$$\begin{aligned} &= \exp\left[\frac{\beta'\left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)\beta - 2\beta'\left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)\left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + \frac{X'\Sigma^{-1}Y}{\sigma^2}\right)}{-2}\right] \\ &\propto \exp\left\{\frac{\left[\beta - \left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + \frac{X'\Sigma^{-1}Y}{\sigma^2}\right)\right]'\left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)\left[\beta - \left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + \frac{X'\Sigma^{-1}Y}{\sigma^2}\right)\right]}{-2}\right\}. \end{aligned}$$

Hence $\beta|Y \sim N\left(\left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + \frac{X'\Sigma^{-1}Y}{\sigma^2}\right),\ \left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)^{-1}\right)$

(3) Find the Bayesian estimator $\widehat{\beta}_B$ for β and express it as a weighted average of β_0 and $\widehat{\beta}$ in 1. (15 points)

Let $W_1 = \Sigma_0^{-1} > 0$, $W_2 = \frac{X'\Sigma^{-1}X}{\sigma^2} > 0$ and $W = W_1 + W_2 > 0$.
With the BLUE for β, $\widehat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$,

$$\begin{aligned} \widehat{\beta}_B &= E(\beta|Y) = \left(\Sigma_0^{-1} + \frac{X'\Sigma^{-1}X}{\sigma^2}\right)^{-1}\left(\Sigma_0^{-1}\beta_0 + \frac{X'\Sigma^{-1}Y}{\sigma^2}\right) \\ &= (W_1 + W_2)^{-1}\left(W_1\beta_0 + \frac{X'\Sigma^{-1}X}{\sigma^2}(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y\right) \\ &= W^{-1}(W_1\beta_0 + W_2\widehat{\beta}\ ) \end{aligned}$$

is a weighted average of β_0 and $\widehat{\beta}$ with weight matrices W_1 and W_2.