Solutions to MATH3091 problem sheet 3

25 Feb 2022

1. See the proof in the lecture notes (Note this proof is not examinable).

2. (a) Since $\boldsymbol{\theta}$ is known, $\mathcal{G}$ is known. In Lecture 7-8, as a result of

$$\boldsymbol{Y}|\boldsymbol{\gamma} \sim N(\boldsymbol{X\beta} + \boldsymbol{U\gamma}, \sigma^2 \boldsymbol{I}_n) \quad \text{and} \quad \boldsymbol{\gamma} \sim N(0, \mathcal{G}_{\boldsymbol{\theta}}),$$

we derived the following conditional and marginal densities:

$$f(\boldsymbol{y}|\boldsymbol{\gamma}; \boldsymbol{\beta}) = (2\pi)^{-n/2}(\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma})^T(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma})\right\}$$

and

$$f(\boldsymbol{\gamma}; \boldsymbol{\beta}) = (2\pi)^{-mq/2}|\mathcal{G}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{\gamma}^T \mathcal{G}^{-1}\boldsymbol{\gamma}\right)$$

The joint likelihood ratio function is

$$f_{\boldsymbol{Y}}(\boldsymbol{y}, \boldsymbol{\gamma}; \boldsymbol{\beta}) \propto \quad \exp\left\{-\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma})^T(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma})\right\} \times \exp\left(-\frac{1}{2}\boldsymbol{\gamma}^T \mathcal{G}^{-1/2}\boldsymbol{\gamma}\right)$$

Denote the joint likelihood as a function of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, i.e.,

$$L(\boldsymbol{\beta}, \boldsymbol{\gamma}) = f_{\boldsymbol{Y}}(\boldsymbol{y}, \boldsymbol{\gamma}; \boldsymbol{\beta}).$$

Therefore we have the log-likelihood

$$\ell(\boldsymbol{\beta}, \boldsymbol{\gamma}) = -\frac{1}{2\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma})^T(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma}) - \frac{1}{2}\boldsymbol{\gamma}^T \mathcal{G}^{-1/2}\boldsymbol{\gamma} + C, \qquad (1)$$

where C is some constant term independent of $\boldsymbol{y}$ and $\boldsymbol{\gamma}$.

(b) We need to prove that $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ maximises $\ell(\boldsymbol{y}, \boldsymbol{\gamma})$. As a result, it is equivalent to minimize

$$Q(\boldsymbol{\beta}, \boldsymbol{\gamma}) = \frac{1}{\sigma^2}(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma})^T(\boldsymbol{y} - \boldsymbol{X\beta} - \boldsymbol{U\gamma}) + \boldsymbol{\gamma}^T \mathcal{G}^{-1/2}\boldsymbol{\gamma}$$

Using the formula for vector derivatives, we have

$$\begin{aligned}
\frac{\partial Q(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\beta}} &= -\frac{2}{\sigma^2}(\boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{X\beta} - \boldsymbol{X}^T\boldsymbol{U\gamma}) \\
\frac{\partial Q(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \boldsymbol{\gamma}} &= -\frac{2}{\sigma^2}(\boldsymbol{U}^T\boldsymbol{y} - \boldsymbol{U}^T\boldsymbol{X\beta} - \boldsymbol{U}^T\boldsymbol{U\gamma}) + 2\mathcal{G}^{-1}\boldsymbol{\gamma}
\end{aligned}$$

Therefore, the system of equations to solve $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ is

$$\begin{cases} \boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{X}^T\boldsymbol{U}\hat{\gamma} = 0 \\ -\dfrac{1}{\sigma^2}(\boldsymbol{U}^T\boldsymbol{y} - \boldsymbol{U}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} - \boldsymbol{U}^T\boldsymbol{U}\hat{\gamma}) + \mathcal{G}^{-1}\hat{\gamma} = 0 \end{cases} \tag{2}$$

(c) To show $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ are root of the equation system (2), we plug in the value of these two estimators, i.e.,

$$\hat{\boldsymbol{\beta}} = (\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{y}, \qquad \hat{\gamma} = \mathcal{G}\boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})$$

For the first equation in (2), we have the left hand side after plug-in $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ equals to

$$\begin{aligned} &\boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{U}\hat{\gamma} - \boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} \\ =&\boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{X}^T\boldsymbol{U}\mathcal{G}\boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) - \boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} \\ =&\boldsymbol{X}^T\boldsymbol{y} - \boldsymbol{X}^T(\boldsymbol{V} - \sigma^2\boldsymbol{I}_n)\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) - \boldsymbol{X}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} \\ =&\sigma^2\boldsymbol{X}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\ =&\sigma^2\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{y} - \sigma^2\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{y} \\ =&0 \end{aligned}$$

For the second equation in (2), we have the left hand side after plug-in $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ equals to

$$\begin{aligned} &-\frac{1}{\sigma^2}(-\boldsymbol{U}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{U}^T\boldsymbol{y} - \boldsymbol{U}^T\boldsymbol{U}\hat{\gamma}) + \mathcal{G}^{-1}\hat{\gamma} \\ =&-\frac{1}{\sigma^2}\left(-\boldsymbol{U}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{U}^T\boldsymbol{y} - \boldsymbol{U}^T\boldsymbol{U}\mathcal{G}\boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})\right) + \mathcal{G}^{-1}\mathcal{G}\boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) \\ =&-\frac{1}{\sigma^2}\left\{-\boldsymbol{U}^T\boldsymbol{X}\hat{\boldsymbol{\beta}} + \boldsymbol{U}^T\boldsymbol{y} - \boldsymbol{U}^T(\boldsymbol{V} - \sigma^2\boldsymbol{I}_n)\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}})\right\} + \boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \hat{\boldsymbol{\beta}}) \\ =&-\boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}) + \boldsymbol{U}^T\boldsymbol{V}^{-1}(\boldsymbol{y} - \hat{\boldsymbol{\beta}}) \\ =&0. \end{aligned}$$

Therefore, we have that $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\gamma}}$ are root of (2).

3. From the lecture notes, it is straightforward to obtain $\boldsymbol{C}\hat{\boldsymbol{\beta}} = \hat{\beta}_1$, $\boldsymbol{c} = 0$ and $r = 1$. Moreover $\boldsymbol{C}(\boldsymbol{X}^T\boldsymbol{V}^{-1}\boldsymbol{X})^{-1}\boldsymbol{C}^T$ is the covariance matrix of $\boldsymbol{C}\hat{\boldsymbol{\beta}}$, which equals to $\text{Var}(\hat{\beta}_1)$ here.

   As a result

$$W = \hat{\beta}_1^2/\text{Var}(\hat{\beta}_1) = 2.37$$

   For a size $\alpha = 0.05$ test, we should reject H_0 if $W > \chi^2_{1,0.95}$, where $\chi^2_{1,0.95}$ is the 95% point of the χ^2_1 distribution, which equals to 3.841459 as calculated by `R`, or 3.84, to two decimal places. So we do not reject H_0.