Lineaariset mallit, Spring 2012, Exercise 4, week 16

1. Consider the linear model $\mathbf{Y} \sim \mathsf{N}\left(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n\right)$ ($\boldsymbol{\beta} \in \mathbb{R}^p$, $\sigma^2 > \mathbf{0}$, $\mathsf{r}(\mathbf{X}) = p$) and recall that the log–likelihood function of the parameter vector $\left[\boldsymbol{\beta}' \quad \sigma^2\right]'$ is

$$l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{y}\right) = -\frac{n}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\mathsf{S}(\boldsymbol{\beta}),$$

where

$$\mathsf{S}(\boldsymbol{\beta}) = (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) = \sum_{i=1}^{n}\left(y_i - \mathbf{x}_i'\boldsymbol{\beta}\right)^2.$$

Show that the Fisher information matrix, that is, the matrix

$$\mathbf{i}\left(\boldsymbol{\beta}, \sigma^2\right) = \mathsf{E}\begin{bmatrix} -\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{Y}\right)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}' & -\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{Y}\right)/\partial\boldsymbol{\beta}\partial\sigma^2 \\ -\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{Y}\right)/\partial\sigma^2\partial\boldsymbol{\beta}' & -\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{Y}\right)/\partial\sigma^2\partial\sigma^2 \end{bmatrix}$$

is

$$\mathbf{i}\left(\boldsymbol{\beta}, \sigma^2\right) = \begin{bmatrix} \sigma^{-2}\mathbf{X}'\mathbf{X} & \mathbf{0} \\ \mathbf{0} & n/2\sigma^4 \end{bmatrix}.$$

(*Hint*: It may be useful to calculate first $\partial l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{y}\right)/\partial\boldsymbol{\beta}$ as $\partial\mathsf{S}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}$ in Exercise 2.3 and after that $\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{y}\right)/\partial\boldsymbol{\beta}\partial\sigma^2$ and further $E\left[\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{Y}\right)/\partial\boldsymbol{\beta}\partial\sigma^2\right]$. $\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{y}\right)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ can be calculated as $\partial^2\mathsf{S}(\boldsymbol{\beta})/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'$ (cf. Exercise 2.5) and $\partial^2 l\left(\boldsymbol{\beta}, \sigma^2; \mathbf{y}\right)/\partial\sigma^2\partial\sigma^2$ only requires differentiation with respect to the real variable σ^2. Note also that symmetry.)

2. Let the correct (full rank) model be $\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \boldsymbol{\varepsilon}$ ($\boldsymbol{\varepsilon} \sim \mathsf{N}\left(\mathbf{0}, \sigma^2\mathbf{I}_n\right)$, $\boldsymbol{\beta}_1 \in \mathbb{R}^{p_1}$, $\boldsymbol{\beta}_2 \in \mathbb{R}^{p_2}$, $\sigma^2 > \mathbf{0}$). Suppose that $\boldsymbol{\beta}_1$ is estimated by using the model from which $\mathbf{X}_2$ has been omitted (in other words, the model equation is $\mathbf{Y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \boldsymbol{\varepsilon}_*$). Derive the expectation of the least squares estimator of the parameter $\boldsymbol{\beta}_1$ and its probability distribution. When is this estimator unbiased?

3. Consider two independent linear models

$$\mathbf{Y}_i = \mathbf{X}_i\boldsymbol{\beta}_i + \boldsymbol{\varepsilon}_i, \quad \boldsymbol{\varepsilon}_i \sim \mathsf{N}_{n_i}\left(\mathbf{0}, \sigma^2\mathbf{I}_{n_i}\right), \quad \boldsymbol{\varepsilon}_1 \perp\!\!\!\perp \boldsymbol{\varepsilon}_2, \quad \boldsymbol{\beta}_i \in \mathbb{R}^p, \ \sigma^2 > \mathbf{0}, \ \mathsf{r}(\mathbf{X}_i) = p, \ i = 1, 2.$$

Form a single model from these two models by using matrices and derive the expression of the least squares estimator of the parameter vector $\boldsymbol{\beta} = \left[\boldsymbol{\beta}_1' \quad \boldsymbol{\beta}_2'\right]'$. What is the probability distribution of this least squares estimator?

An auxiliary result: If $\mathbf{A}$ and $\mathbf{B}$ are nonsingular square matrices, then

$$\begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}^{-1} \end{bmatrix}.$$

4. Consider the linear model of two independent normal samples, that is,

$$Y_1, ..., Y_n \ \underline{\perp\!\!\!\perp}\ , \ Y_i \sim \begin{cases} \mathsf{N}\left(\mu_1, \sigma^2\right), & \text{as } i = 1, ..., n_1 \\ \mathsf{N}\left(\mu_2, \sigma^2\right), & \text{as } i = n_1 + 1, ..., n_1 + n_2 = n \end{cases}$$

(μ_1, $\mu_2 \in \mathbb{R}$, $\sigma^2 > 0$, n_1, $n_2 > 1$). Estimate the parameters μ_1 and μ_2 under the constraint $\mu_1 = \mu_2$ by using (i) the general formula for the constrained least squares estimate and (ii) taking the constraint $\mu_1 = \mu_2$ into account in the model and estimating the parameters of the resulting model.

Note: Let the model be $\mathbf{Y} \sim \mathsf{N}\left(\mathbf{X}\boldsymbol{\beta}, \sigma^2\mathbf{I}_n\right)$ ($\boldsymbol{\beta} \in \mathbb{R}^p$, $\sigma^2 > \mathbf{0}$, $\mathsf{r}(\mathbf{X}) = p$) and consider the general linear constraints $\mathbf{A}\boldsymbol{\beta} = \mathbf{c}$ where the matrix $\mathbf{A}$ $(q \times p)$ and the vector $\mathbf{c}$ $(q \times 1)$ are known and $\mathsf{r}(\mathbf{A}) = q$. Then the general formula for the constrained least squares estimate is

$$\hat{\boldsymbol{\beta}}_H = \hat{\boldsymbol{\beta}} - (\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}'(\mathbf{A}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{A}')^{-1}(\mathbf{A}\hat{\boldsymbol{\beta}} - \mathbf{c}),$$