LONDON TAUGHT COURSE CENTRE

LTCC Basic Course Statistical Modelling and Estimation

Exercise Sheet 3: Solutions **February/March 2012**

---

1. Consider a $q \times r$ matrix $\mathbf{A}$ and an $r \times q$ matrix $\mathbf{B}$. Show that $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$.

   Denote the elements of $\mathbf{A}$ by a_{ij} and the element of $\mathbf{B}$ by b_{ji}, where $i = 1, \ldots, q$ and $j = 1, \ldots, r$. The kth diagonal element $(\mathbf{AB})_{kk}$, $k = 1, \ldots, q$, of $\mathbf{AB}$ is then equal to $(\mathbf{AB})_{kk} = \sum_{j=1}^{r} a_{kj} b_{jk}$ and so the trace of $\mathbf{AB}$ is

$$\mathrm{tr}(\mathbf{AB}) = \sum_{k=1}^{q} \sum_{j=1}^{r} a_{kj} b_{jk}.$$

   Similarly the ℓth diagonal element $(\mathbf{BA})_{\ell\ell}$, $\ell = 1, \ldots, r$, of $\mathbf{BA}$ is $(\mathbf{BA})_{\ell\ell} = \sum_{i=1}^{q} b_{\ell i} a_{i\ell}$ which implies that the trace of $\mathbf{BA}$ is equal to

$$\mathrm{tr}(\mathbf{BA}) = \sum_{\ell=1}^{r} \sum_{i=1}^{q} b_{\ell i} a_{i\ell}.$$

   Now changing the order of summation it can be seen that

$$\mathrm{tr}(\mathbf{BA}) = \sum_{\ell=1}^{r} \sum_{i=1}^{q} b_{\ell i} a_{i\ell} = \sum_{i=1}^{q} \sum_{\ell=1}^{r} a_{i\ell} b_{\ell i} = \mathrm{tr}(\mathbf{AB}).$$

2. Let $\mathbf{Y} = [Y_1, \ldots, Y_n]'$ be a random vector (not necessarily in a linear model) with expectation vector $E(\mathbf{Y}) = [E(Y_1), \ldots, E(Y_n)]'$. The variance-covariance matrix of $\mathbf{Y}$ is defined as

$$V(\mathbf{Y}) = E[(\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Y} - E(\mathbf{Y}))'].$$

   Show that if $\mathbf{A}$ is a fixed $m \times n$ matrix, then

$$V(\mathbf{AY}) = \mathbf{A}V(\mathbf{Y})\mathbf{A}'.$$

   By using the linearity of the expected value, it follows from the definition of the variance-covariance matrix that

$$\begin{aligned} V(\mathbf{AY}) &= E[(\mathbf{AY} - E(\mathbf{AY}))(\mathbf{AY} - E(\mathbf{AY}))'] = E[\mathbf{A}(\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Y} - E(\mathbf{Y}))'\mathbf{A}'] \\ &= \mathbf{A}E[(\mathbf{Y} - E(\mathbf{Y}))(\mathbf{Y} - E(\mathbf{Y}))']\mathbf{A}' = \mathbf{A}V(\mathbf{Y})\mathbf{A}'. \end{aligned}$$

3. For random vectors $\mathbf{U} = [U_1, \ldots, U_n]'$ with expectation $E(\mathbf{U}) = [E(U_1), \ldots, E(U_n)]'$ and $\mathbf{V} = [V_1, \ldots, V_m]'$ with expectation vector $E(\mathbf{V}) = [E(V_1), \ldots, E(V_m)]'$ the covariance matrix of $\mathbf{U}$ and $\mathbf{V}$ is defined as

$$Cov(\mathbf{U}, \mathbf{V}) = E[(\mathbf{U} - E(\mathbf{U}))(\mathbf{V} - E(\mathbf{V}))'].$$

   Show that if $\mathbf{A}$ is a fixed $a \times n$ matrix and $\mathbf{B}$ is a fixed $b \times m$ matrix, then

$$Cov(\mathbf{AU}, \mathbf{BV}) = \mathbf{A}Cov(\mathbf{U}, \mathbf{V})\mathbf{B}'.$$

As in the previous question the result follows from the linearity of the expected value, since

$$\begin{aligned} Cov(\mathbf{AU}, \mathbf{BV}) &= E[(\mathbf{AU} - E(\mathbf{AU}))(\mathbf{BV} - E(\mathbf{BV}))'] \\ &= E[\mathbf{A}(\mathbf{U} - E(\mathbf{U}))(\mathbf{V} - E(\mathbf{V}))'\mathbf{B}'] \\ &= \mathbf{A}E[(\mathbf{U} - E(\mathbf{U}))(\mathbf{V} - E(\mathbf{V}))']\mathbf{B}' = \mathbf{A}Cov(\mathbf{U}, \mathbf{V})\mathbf{B}'. \end{aligned}$$

4. Let $\mathbf{Y} = [Y_1, \ldots, Y_n]'$ be a random vector (not necessarily in a linear model) with $E(\mathbf{Y}) = \boldsymbol{\mu}$ and $V(\mathbf{Y}) = \mathbf{V}$, where $\mathbf{V}$ is a fixed matrix. Prove that if $\mathbf{A}$ is a known $n \times n$ matrix, then

$$E(\mathbf{Y}'\mathbf{AY}) = \mathrm{tr}(\mathbf{AV}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}.$$

The calculation

$$\begin{aligned} E(\mathbf{Y}'\mathbf{AY}) &= E(\mathrm{tr}(\mathbf{Y}'\mathbf{AY})) = E(\mathrm{tr}(\mathbf{AYY}')) = \mathrm{tr}(E(\mathbf{AYY}')) = \mathrm{tr}(\mathbf{A}E(\mathbf{YY}')) \\ &= \mathrm{tr}(\mathbf{A}(V(\mathbf{Y}) + E(\mathbf{Y})E(\mathbf{Y})')) = \mathrm{tr}(\mathbf{A}(\mathbf{V} + \boldsymbol{\mu}\boldsymbol{\mu}')) = \mathrm{tr}(\mathbf{AV}) + \mathrm{tr}(\mathbf{A}\boldsymbol{\mu}\boldsymbol{\mu}') \\ &= \mathrm{tr}(\mathbf{AV}) + \mathrm{tr}(\boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu}) = \mathrm{tr}(\mathbf{AV}) + \boldsymbol{\mu}'\mathbf{A}\boldsymbol{\mu} \end{aligned}$$

proves the result. Note that in this calculation the result about the trace in Question 1 has been applied several times. Further the linearity of the expected value and of the trace, as well as the fact that the order of the trace and the expected value can be interchanged have been used. Finally, the identity $V(\mathbf{Y}) = E(\mathbf{YY}') - E(\mathbf{Y})E(\mathbf{Y})'$ has been employed, which is the multivariate analogue of the well-known relationship $Var(Y) = E(Y^2) - [E(Y)]^2$ for a univariate random variable Y.

5. Consider the one-way analysis of variance model in Example 1.3 for $t = 3$ treatments with design matrix $\mathbf{X}$ and parameter vector $\boldsymbol{\beta} = [\mu\ \alpha_1\ \alpha_2\ \alpha_3]'$. Assume that each treatment has replication $r = 3$ and that the data vector is given by

$$\mathbf{y} = [6.48, 5.55, 5.05, 7.07, 6.88, 8.36, 3.84, 4.43, 4.20]'$$

where the first three responses are for treatment 1, the next three responses are for treatment 2 and the final three responses are for treatment 3.

   (a) Write down the design matrix $\mathbf{X}$ and find its rank.
   The design matrix is

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

   and obviously $\mathrm{rank}(\mathbf{X}) = 3$.

(b) Is the parameter μ estimable?

The parameter μ corresponds to the linear function $\boldsymbol{\lambda}'\boldsymbol{\beta}$ where $\boldsymbol{\lambda}' = [1\ 0\ 0\ 0]$. For $\boldsymbol{\lambda}'\boldsymbol{\beta}$ to be estimable we would need to find a vector $\mathbf{a} \in \mathbb{R}^9$ such that $\boldsymbol{\lambda}' = \mathbf{a}'\mathbf{X}$. Obviously, such a vector $\mathbf{a}$ does not exist and so $\boldsymbol{\lambda}'\boldsymbol{\beta}$ or equivalently μ is not estimable.

(c) Is the linear function $\alpha_1 - \frac{1}{2}(\alpha_2 + \alpha_3)$ estimable?

The function $\alpha_1 - \frac{1}{2}(\alpha_2 + \alpha_3)$ can be written as $\boldsymbol{\lambda}'\boldsymbol{\beta}$ where $\boldsymbol{\lambda}' = [0\ 1\ -\frac{1}{2}\ -\frac{1}{2}]$. Now it can easily be verified that $\boldsymbol{\lambda}' = \mathbf{a}'\mathbf{X}$ for the vector $\mathbf{a} = \frac{1}{6}[2\ 2\ 2\ -1\ -1\ -1\ -1\ -1\ -1]'$ and so $\boldsymbol{\lambda}'\boldsymbol{\beta}$ or equivalently $\alpha_1 - \frac{1}{2}(\alpha_2 + \alpha_3)$ is estimable.

(d) For whichever of the functions in (b) and (c) is estimable find its least squares estimate using the data given and also estimate the variance of the estimator.

Of the two functions in parts (b) and (c) only $\alpha_1 - \frac{1}{2}(\alpha_2 + \alpha_3)$ is estimable. By using the first method in Section 1.3 of the lecture notes with

$$\mathbf{X}^* = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

a generalized inverse of $\mathbf{X}'\mathbf{X}$ is

$$(\mathbf{X}'\mathbf{X})^- = (\mathbf{X}'\mathbf{X} + \mathbf{X}^{*\prime}\mathbf{X}^*)^{-1} = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & \frac{4}{3} & 1 & 1 \\ -1 & 1 & \frac{4}{3} & 1 \\ -1 & 1 & 1 & \frac{4}{3} \end{bmatrix}.$$

Further we have that

$$\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6.48 \\ 5.55 \\ 5.05 \\ 7.07 \\ 6.88 \\ 8.36 \\ 3.84 \\ 4.43 \\ 4.20 \end{bmatrix} = \begin{bmatrix} 51.86 \\ 17.08 \\ 22.31 \\ 12.47 \end{bmatrix}$$

and so by using equation (2.3) from the lecture notes a least squares estimate of $\boldsymbol{\beta}$ is

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^-\mathbf{X}'\mathbf{y} = \begin{bmatrix} 1 & -1 & -1 & -1 \\ -1 & \frac{4}{3} & 1 & 1 \\ -1 & 1 & \frac{4}{3} & 1 \\ -1 & 1 & 1 & \frac{4}{3} \end{bmatrix} \begin{bmatrix} 51.86 \\ 17.08 \\ 22.31 \\ 12.47 \end{bmatrix} = \begin{bmatrix} 0 \\ 5.69 \\ 7.44 \\ 4.16 \end{bmatrix}.$$

The unique least squares estimate of the estimable function $\boldsymbol{\lambda}'\boldsymbol{\beta} = \alpha_1 - \frac{1}{2}(\alpha_2 + \alpha_3)$ where $\boldsymbol{\lambda}' = [0\ 1\ \ -\frac{1}{2}\ \ -\frac{1}{2}]$ is then

$$\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}} = [0\ 1\ \ -\tfrac{1}{2}\ \ -\tfrac{1}{2}] \begin{bmatrix} 0 \\ 5.6933 \\ 7.4367 \\ 4.1567 \end{bmatrix} = -0.1033.$$

Note that this is equal to $\bar{y}_{1.} - \frac{1}{2}(\bar{y}_{2.} + \bar{y}_{3.})$, where $\bar{y}_{1.} = 5.6933$, $\bar{y}_{2.} = 7.4367$ and $\bar{y}_{3.} = 4.1567$ are the means for treatments 1, 2 and 3.
The variance of $\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}}$ is $Var(\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}}) = \sigma^2\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{\lambda}$ (see Section 2.3) and this is estimated unbiasedly by MSE $\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{\lambda}$ where MSE is the residual mean square. Using the estimate $\hat{\boldsymbol{\beta}}$ we have found before the residual sum of squares is equal to

$$(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}) = 2.527$$

and so the residual mean square is

$$\text{MSE} = \frac{(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})'(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})}{n-r} = \frac{2.527}{6} = 0.4212$$

where $n = 9$ is the number of observations and $r = \text{rank}(\mathbf{X}) = 3$ is the rank of $\mathbf{X}$. Thus the estimated variance of $\boldsymbol{\lambda}'\hat{\boldsymbol{\beta}}$ is equal to

$$\text{MSE}\,\boldsymbol{\lambda}'(\mathbf{X}'\mathbf{X})^{-}\boldsymbol{\lambda} = 0.4212 \times \frac{1}{2} = 0.2106.$$