On generalized inverses and Bayes estimators in balanced MANOVA models

Wolfgang Polasek and Shuangzhe Liu
Institute of Statistics and Econometrics
University of Basel
Holbeinstrasse 12, CH-4051 Basel
Switzerland

November 28, 2000

Abstract:

The one-way and two-way MANOVA (multivariate analysis of variance) models are formulated as Bayesian linear models with conjugate prior distributions. The classical case can be viewed as a special one in which matrix generalized inverses and the Moore-Penrose inverse are used to derive the so-called OLS^- and OLS^+ estimators for the rank deficient models. Two solutions of generalized inverses lead to two estimators, i.e., row-wise and column-wise averages as estimated effects. The Moore-Penrose inverse-based estimator is a mixture of the two estimators derived from the two generalized inverses above. These solutions are also special cases of the Bayesian analysis for the MANOVA models.

Keywords:

Conjugate priors, OLS^- and OLS^+ estimators, generalized and Moore-Penrose inverses, balanced MANOVA models.

1 Introduction

Analysis of variance (ANOVA) models are an important class of models in statistics, see, e.g. Scheffé (1959) and Rao (1973). As ANOVA models are linear models where a quantitative variable is explained by a series of qualitative variables, the use of recent developments in matrix algebra for estimation purposes should be encouraged. In Polasek and Liu (1997) we suggested a new parameterization for ANOVA models and showed that rank deficient matrices can be used to estimate with the concept of generalized inverses. This leads not to new tests (since the F-statistic is not changed by the parameterization) but to new ways of reporting the estimated effects. The so-called OLS^- estimators produce non-unique effect estimates, using the results of g-inverses of the cross product matrix. A unique effect can be estimated using the Moore-Penrose inverse (matrix g-inverses and the Moore-Penrose inverse are discussed in, e.g. Magnus and Neudecker, 1991). This parameterization also allows a generalization of the ANOVA model to the Bayesian linear model, see Polasek and Liu (1997). Bayesian methods in ANOVA models have been previously applied in, e.g. Press (1989) and Searle et al. (1992), and a Bayesian robustness approach can be found in Polasek and Poetzelberger (1994).

In this paper we extend the univariate ANOVA model to the multivariate MANOVA model. Again, the matrix formulation of the model makes the derivation of results more compact and allows a generalization to the multivariate Bayesian MANOVA model as well. The Bayesian analysis uses a conjugate Normal-gamma prior distribution for the univariate case and a conjugate Normal-Wishart prior distribution for the multivariate case. The multivariate OLS^- and OLS^+ estimators for the one-way and two-way balanced design models are shown to be matrix extensions of the univariate ANOVA models. The non-unique OLS^- estimators have two sets of solutions which are interpreted as row-wise or column-wise averages of decentered multivariate observations. The plan of the paper is as follows. In section 2 we derive the one-way MANOVA model for balanced design and section 3 describes the two-way MANOVA model. In section 4 we derive the OLS^- and OLS^+ estimators. A final section concludes.

2 One-way MANOVA for balanced design

The one-way MANOVA model for balanced design with the total number of observations $N = qn$ can be written as a multivariate regression system (see, e.g. Press, 1989)

$$\underset{K\times 1}{y_{ij}} = \underset{K\times 1}{\beta_i} + \underset{K\times 1}{u_{ij}}, \quad i = 1, \ldots, q; \qquad j = 1, \ldots n$$

where y_{ij} is a $K \times 1$ vector of the observations for the j^{th} replication in the i^{th} population, β_i is the main effect α_i due to population i plus the grand mean μ ($\beta_i = \alpha_i + \mu$), and u_{ij} is an error term. In matrix form it is

$$\underset{N\times K}{\mathbf{Y}} = \underset{(N\times q)(q\times K)}{\mathbf{X}\cdot\mathbf{B}} + \underset{N\times K}{\mathbf{U}}$$

or

$$\begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix} = (\mathbf{I}_q \otimes \mathbf{1}_n) \cdot \begin{pmatrix} \beta_1^\top \\ \vdots \\ \beta_q^\top \end{pmatrix} + \begin{pmatrix} \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_q \end{pmatrix}, \tag{1}$$

where $\mathbf{I}_q$ is a $q \times q$ identity matrix, $\mathbf{1}_n = (1, ..., 1)^\top$ is an $n \times 1$ vector, $\mathbf{Y}_i = (y_{i1}, \ldots, y_{in})^\top$ and $\mathbf{U}_i = (u_{i1}, \ldots, u_{in})^\top$, $i = 1, \ldots, n$ are $n \times K$ matrices.

Consider mutually i.i.d. normal u_{ij}s with a $K \times 1$ mean vector equal to zero and a $K \times K$ variance matrix $\mathbf{\Sigma} > \mathbf{0}$. The likelihood function is proportional to

$$L(\mathbf{Y} \mid \mathbf{X}, \mathbf{B}, \mathbf{\Sigma}) \ \propto \ \mid \mathbf{\Sigma} \mid^{-\frac{N}{2}} \exp(-\tfrac{1}{2}\mathrm{tr}\mathbf{\Sigma}^{-\mathbf{1}}\mathbf{U}^\top\mathbf{U}).$$

If the prior distribution for $\mathbf{B} = (\beta_1, ..., \beta_q)^\top$ and $\mathbf{\Sigma}^{-\mathbf{1}}$ is of the conjugate normal-Wishart density family

$$(\mathbf{B}, \mathbf{\Sigma}^{-1}) \sim \mathcal{NW}_{q\times K}[\mathbf{B}_*, \mathbf{H}_*, \mathbf{\Sigma}_*, n_*],$$

i.e.

$$\mathbf{\Sigma}^{-1} \sim \mathcal{W}_K(\mathbf{\Sigma}_*, n_*)$$

and

$$\mathbf{B} \mid \mathbf{\Sigma}^{-1} \sim \mathcal{N}_{q\times K}(\mathbf{B}_*, \mathbf{H}_*),$$

then the posterior distribution is again normal-Wishart distributed:

$$(\mathbf{B}, \mathbf{\Sigma}^{-1} \mid \mathbf{X}, \mathbf{Y}) \ \sim \ \mathcal{NW}_{q\times K}[\mathbf{B}_{**}, \mathbf{H}_{**}, \mathbf{\Sigma}_{**}, n_{**}].$$

Let $\mathbf{H}_*^{-1} = \frac{1}{n'}\mathbf{I}_q \otimes \mathbf{\Sigma}^{-\mathbf{1}}$, where n' is the hypothetical sample size with respect to the prior information $\mathbf{B}_*$, then the posterior parameters are (see Polasek, 1995)

$$\begin{aligned} \mathbf{H}_{**}^{-1} &= \mathbf{G}_{**} \otimes \mathbf{\Sigma}, \\ \mathbf{B}_{**} &= \mathbf{G}_{**}(n'\mathbf{B}_* + \mathbf{X}^\top\mathbf{Y}), \\ \mathbf{G}_{**}^{-1} &= n'\mathbf{I}_q + \mathbf{X}^\top\mathbf{X}, \\ n_{**} &= n_* + N. \end{aligned}$$

Now for model (1) with $\mathbf{X} = \mathbf{I}_q \otimes \mathbf{1}_n$ we get $\mathbf{G}_{**}^{-1} = n''\mathbf{I}_q$ with the posterior degrees of freedom $n'' = n' + n$. If $\mathbf{B}_* = \mathbf{1}_q\mu_*^\top$, then the posterior mean matrix is

$$\begin{aligned} \mathbf{B}_{**} &= (\mathbf{b}_{1**}, \ldots, \mathbf{b}_{q**})^\top \\ &= \frac{1}{n''}(n'\mathbf{1}_q\mu_*^\top \ + \ n\overline{\mathbf{Y}}^*). \end{aligned} \tag{2}$$

This follows from the cross product term

$$\begin{aligned}\mathbf{X}^\top\mathbf{Y} &= \underset{q\times N}{(\mathbf{I}_q \otimes \mathbf{1}_n^\top)} \underset{N\times K}{\mathbf{Y}} \\ &= \underbrace{\begin{pmatrix} \mathbf{1}_n^\top & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{1}_n^\top \end{pmatrix}}_{q} \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix} \\ &= n \cdot \underset{q\times K}{\begin{pmatrix} \bar{\mathbf{y}}_1 \\ \vdots \\ \bar{\mathbf{y}}_q \end{pmatrix}} = n\bar{\mathbf{Y}}^*, \end{aligned} \tag{3}$$

where $\bar{\mathbf{Y}}^*$ is a $q \times K$ matrix of sample means.

Note that $\mathbf{B}_{**}$ in (2) is a simple average between the prior location $\mathbf{B}_*$ and the maximum likelihood (ML) location $\overline{\mathbf{Y}}^*$. An alternative way to obtain the result (3) using a compact matrix notation is as follows.

$$\begin{aligned}\mathbf{X}^\top\mathbf{Y} &= (\mathbf{I}_q \otimes \mathbf{1}_n^\top)\mathbf{Y} \\ &= (\mathbf{1}_n^\top \otimes \mathbf{I}_q)\mathbf{L}_{nq}\mathbf{Y} \\ &= (\mathbf{1}_n^\top \otimes \mathbf{I}_q)\mathbf{Y}^* \\ &= n\bar{\mathbf{Y}}^*,\end{aligned}$$

where $\mathbf{L}_{nq}$ is the so-called $N \times N$ commutation matrix (see, e.g. Magnus and Neudecker, 1991) with the following properties

$$\begin{aligned}\mathbf{I}_q \otimes \mathbf{1}_n^\top &= (\mathbf{1}_n^\top \otimes \mathbf{I}_q)\mathbf{L}_{nq}^\top, \\ \mathbf{L}_{nq}^\top\mathbf{L}_{nq} &= \mathbf{I}_N, \\ \mathbf{L}_{nq}\mathbf{Y} &= \mathbf{Y}^*,\end{aligned}$$

and $\mathbf{Y}^*$ is a reordering of the dependent variable matrix $\mathbf{Y}$. Also, the mean matrix can be calculated as an average of empirical observation matrices:

$$\begin{aligned}\bar{\mathbf{Y}}^* &= \frac{1}{n}\sum_{j=1}^{n} \mathbf{Y}_j^* : \qquad q \times K, \\ \mathbf{Y}^* &= (\mathbf{Y}_1^{*\top}, \dots \mathbf{Y}_n^{*\top})^\top : \qquad N \times K, \\ \mathbf{Y}_j^* &= (y_{1j}, \dots, y_{qj})^\top, \quad j = 1, \dots, n.\end{aligned}$$

We easily see that the i^{th} row of $\bar{\mathbf{Y}}^*$ is $\bar{\mathbf{y}}_i = \frac{1}{n}(y_{i1}^\top + \ldots + y_{in}^\top)$, $i = 1, \ldots, q$.

3 Two-way MANOVA for balanced design

Consider the following two-way MANOVA model for balanced design with the total number of observations $N = qn$

$$\underset{K\times 1}{\mathbf{y}_{ij}} = \alpha_i + \beta_j + \underset{K\times 1}{\mathbf{u}_{ij}}, \quad i = 1,\ldots,q; \qquad j = 1,\ldots,n$$

which can be written as a multivariate regression system

$$\underset{N\times K}{\begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix}} = \underset{N\times q}{(\mathbf{I}_q \otimes \mathbf{1}_n)} \begin{pmatrix} \alpha_1^\top \\ \vdots \\ \alpha_q^\top \end{pmatrix} + \underset{N\times n}{(\mathbf{1}_q \otimes \mathbf{I}_n)} \cdot \begin{pmatrix} \beta_1^\top \\ \vdots \\ \beta_n^\top \end{pmatrix} + \begin{pmatrix} \mathbf{U}_1 \\ \vdots \\ \mathbf{U}_q \end{pmatrix}$$

or compactly as

$$\begin{aligned} \underset{N\times K}{\mathbf{Y}} &= \underset{(N\times q)}{\mathbf{X}_1} \cdot \underset{(q\times K)}{\mathbf{A}} + \underset{(N\times n)}{\mathbf{X}_2} \cdot \underset{(n\times K)}{\mathbf{B}} + \underset{N\times K}{\mathbf{U}} \\ &= (\mathbf{X}_1 : \mathbf{X}_2) \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} + \mathbf{U} \\ &= \underset{(N\times m)}{\tilde{\mathbf{X}}} \cdot \underset{(m\times K)}{\theta} + \underset{N\times K}{\mathbf{U}}, \end{aligned} \tag{4}$$

with the design block matrices

$$\mathbf{X}_1 = \mathbf{I}_q \otimes \mathbf{1}_n : N \times q,$$

$$\mathbf{X}_2 = \mathbf{1}_q \otimes \mathbf{I}_n : N \times n$$

and the stacked regression matrix

$$\tilde{\mathbf{X}} = (\mathbf{X}_1 : \mathbf{X}_2) : N \times m,$$

with $N = qn$ and $m = q + n$.

Also, we define a multivariate normal distribution for the error matrix $\mathbf{U}$, i.e.

$$\mathbf{U} \sim \mathcal{N}[\mathbf{0}, \mathbf{I}_N \otimes \mathbf{\Sigma}],$$

where $\mathbf{\Sigma}$ is a $K \times K$ matrix and $\mathbf{I}_N$ an $N \times N$ identity matrix. For the prior distribution we can assume exchangeability and conjugacy, i.e., conditional on the variance matrix $\mathbf{\Sigma}$

$$\begin{aligned} \mathbf{A} &\sim \mathcal{N}_{q\times K}[\mathbf{A}_*, \mathbf{H}_{\mathbf{A}*} \otimes \mathbf{\Sigma}], \\ \mathbf{B} &\sim \mathcal{N}_{n\times K}[\mathbf{B}_*, \mathbf{H}_{\mathbf{B}*} \otimes \mathbf{\Sigma}], \\ \mathbf{\Sigma}^{-1} &\sim \mathcal{W}_K[\mathbf{\Sigma}_*, n_*]. \end{aligned}$$

For simplicity we let for the prior locations and variance matrices

$$\begin{aligned}
\mathbf{A}_* &= \mathbf{1}_q \mathbf{a}_*^\top, && (5)\\
\mathbf{B}_* &= \mathbf{1}_n \mathbf{b}_*^\top, && (6)\\
\mathbf{H}_{\mathbf{A}*} &= \frac{1}{n'_a}\mathbf{I}_q, && (7)\\
\mathbf{H}_{\mathbf{B}*} &= \frac{1}{n'_b}\mathbf{I}_n, && (8)
\end{aligned}$$

where $\mathbf{a}_*$ and $\mathbf{b}_*$ are $K \times 1$ vectors and n'_a and n'_b can be viewed as values of the prior information expressed in hypothetical sample sizes.

The joint prior distribution of $\mathbf{A}$ and $\mathbf{B}$ given $\mathbf{\Sigma}$ is then

$$\begin{aligned}
\underset{mK\times 1}{\text{vec } \theta^\top = \begin{pmatrix} vec\ \mathbf{A}^\top \\ vec\ \mathbf{B}^\top \end{pmatrix}} &\sim \mathcal{N}\left[\begin{pmatrix} \mathbf{1}_q \otimes \mathbf{a}_* \\ \mathbf{1}_n \otimes \mathbf{b}_* \end{pmatrix}, \begin{pmatrix} \mathbf{H}_{\mathbf{A}*} \otimes \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{\mathbf{B}*} \otimes \mathbf{\Sigma} \end{pmatrix}\right] \\
&= \mathcal{N}[vec\ \theta_*^\top, \tilde{\mathbf{H}}_*]. \qquad (9)
\end{aligned}$$

Note that, instead of (5) and (6), if the prior locations are

$$\begin{aligned}
\mathbf{A}_* &= \mathbf{a}_* \mathbf{1}_K^\top, && (10)\\
\mathbf{B}_* &= \mathbf{b}_* \mathbf{1}_K^\top, && (11)
\end{aligned}$$

where $\mathbf{a}_*$ and $\mathbf{b}_*$ are vectors of $q \times 1$ and $n \times 1$ respectively, then the joint prior distribution of $\mathbf{A}$ and $\mathbf{B}$ given $\mathbf{\Sigma}$ is

$$\begin{aligned}
\underset{mK\times 1}{\text{vec } \theta^\top = \begin{pmatrix} vec\ \mathbf{A}^\top \\ vec\ \mathbf{B}^\top \end{pmatrix}} &\sim \mathcal{N}\left[\begin{pmatrix} \mathbf{a}_* \otimes \mathbf{1}_K \\ \mathbf{b}_* \otimes \mathbf{1}_K \end{pmatrix}, \begin{pmatrix} \mathbf{H}_{\mathbf{A}*} \otimes \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{H}_{\mathbf{B}*} \otimes \mathbf{\Sigma} \end{pmatrix}\right] \\
&= \mathcal{N}[vec\ \theta_*^\top, \tilde{\mathbf{H}}_*]. \qquad (12)
\end{aligned}$$

Using conjugate Bayesian inference (Polasek, 1995) yields the posterior normal-Wishart distribution

$$\theta | \mathbf{Y} \sim \mathcal{N}W[\theta_{**}, \tilde{\mathbf{H}}_{**}, \mathbf{\Sigma}_{**}, n_{**}], \qquad (13)$$

$$\begin{aligned}
\tilde{\mathbf{H}}_{**}^{-1} &= \mathbf{H}_{**} \otimes \mathbf{\Sigma}, \\
\theta_{**} &= \mathbf{H}_{**}(\mathbf{H}_*^{-1}\ \theta_* + \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}} \hat{\theta}) = \mathbf{H}_{**}(\mathbf{H}_*^{-\mathbf{1}}\ \theta_* + \tilde{\mathbf{X}}^\top \mathbf{Y}), && (14)\\
\mathbf{H}_{**}^{-1} &= \mathbf{H}_*^{-1} + \tilde{\mathbf{X}}^\top \tilde{\mathbf{X}}, && (15)\\
\hat{\theta} &= (\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}})^{-1} \tilde{\mathbf{X}}^\top \mathbf{Y}, \\
n_{**} &= n_* + N.
\end{aligned}$$

We can also consider the following two-way MANOVA model for balanced design with the total number of observations $N = qn$, which is a multivariate generalization of the two-way ANOVA model in Polasek and Liu (1997, p. 162)

$$\mathbf{y}_k = (\mathbf{I}_q \otimes \mathbf{1}_n)\alpha_k + (\mathbf{1}_q \otimes \mathbf{I}_n)\beta_k + \mathbf{u}_k, \quad k = 1, \ldots, K$$

or compactly as

$$\begin{aligned}
\underset{N\times K}{\mathbf{Y}} &= \underset{(N\times q)}{\mathbf{X}_1} \cdot \underset{(q\times K)}{\mathbf{A}} + \underset{(N\times n)}{\mathbf{X}_2} \cdot \underset{(n\times K)}{\mathbf{B}} + \underset{N\times K}{\mathbf{U}} \\
&= (\mathbf{X}_1 : \mathbf{X}_2) \begin{pmatrix} \mathbf{A} \\ \mathbf{B} \end{pmatrix} + \mathbf{U} \\
&= \underset{(N\times m)}{\tilde{\mathbf{X}}} \cdot \underset{(m\times K)}{\theta} + \underset{N\times K}{\mathbf{U}}, \qquad (16)
\end{aligned}$$

where

$$\begin{aligned}
\mathbf{X}_1 &= \mathbf{I}_q \otimes \mathbf{1}_n : N \times q, \\
\mathbf{X}_2 &= \mathbf{1}_q \otimes \mathbf{I}_n : N \times n, \\
\tilde{\mathbf{X}} &= (\mathbf{X}_1 : \mathbf{X}_2) : N \times m, \\
\mathbf{Y} &= (\mathbf{y}_1 : \ldots : \mathbf{y}_K) : N \times K, \\
\mathbf{U} &= (\mathbf{u}_1 : \ldots : \mathbf{u}_K) : N \times K, \\
\mathbf{A} &= (\alpha_1 : \ldots : \alpha_K) : q \times K, \\
\mathbf{B} &= (\beta_1 : \ldots : \beta_K) : n \times K,
\end{aligned}$$

with $N = qn$ and $m = q + n$.

We see that the matrix form in (16) is the same as in (4) and (13) holds for model (16).
Now we compute the posterior mean θ_{**} with the prior precision matrix

$$\begin{aligned}
\mathbf{H}_*^{-1} &= \begin{pmatrix} \mathbf{H}_{\mathbf{A}_*}^{-1} & 0 \\ 0 & \mathbf{H}_{\mathbf{B}_*}^{-1} \end{pmatrix} \\
&= \begin{pmatrix} n'_a \mathbf{I}_q & 0 \\ 0 & n'_b \mathbf{I}_n \end{pmatrix}, \qquad (17)
\end{aligned}$$

the cross-product term $\tilde{\mathbf{X}}^\top \mathbf{Y}$ and matrix $\mathbf{H}_{**}^{-1}$.

Since

$$\tilde{\mathbf{X}}^\top \mathbf{Y} = \begin{pmatrix} \mathbf{X}_1^\top \mathbf{Y} \\ \mathbf{X}_2^\top \mathbf{Y} \end{pmatrix},$$

where the design matrix is

$$\begin{aligned}\tilde{\mathbf{X}} &= (\mathbf{X}_1 : \mathbf{X}_2)\\ &= (\mathbf{I}_q \otimes \mathbf{1}_n : \mathbf{1}_q \otimes \mathbf{I}_n),\end{aligned}$$

we can calculate the two components of $\tilde{\mathbf{X}}^\top \mathbf{Y}$ separately. For the first component we find

$$\begin{aligned}\mathbf{X}_1^\top \mathbf{Y} = & \underset{q\times N}{(\mathbf{I}_q \otimes \mathbf{1}_n^\top)} \quad \underset{N\times K}{\mathbf{Y}}\\ = & \underbrace{\begin{pmatrix} \mathbf{1}_n^\top & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathbf{1}_n^\top \end{pmatrix}}_{q} \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix}\\ = & \; n \cdot \begin{pmatrix} \bar{\mathbf{y}}_1 \\ \vdots \\ \bar{\mathbf{y}}_q \end{pmatrix} = \underset{q\times K}{n \cdot \bar{\mathbf{Y}}^*_{.q}}\end{aligned}$$

and for the second component

$$\begin{aligned}\mathbf{X}_2^\top \mathbf{Y} &= \underset{n\times N}{(\mathbf{1}_q^\top \otimes \mathbf{I}_n)} \quad \underset{N\times K}{\mathbf{Y}}\\ &= \underbrace{(\mathbf{I}_n, \dots, \mathbf{I}_n)}_{q} \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix}\\ &= \sum_{i=1}^q \mathbf{Y}_i = \underset{n\times K}{q \cdot \bar{\mathbf{Y}}_{n.}},\end{aligned}$$

where $\bar{\mathbf{Y}}^*_{.q}$ is a $q \times K$ matrix of averages over n groups, and $\bar{\mathbf{Y}}_{n.}$ an $n \times K$ matrix over q groups.

Noting that

$$\begin{aligned}\tilde{\mathbf{X}}^\top \tilde{\mathbf{X}} &= \begin{pmatrix} \mathbf{X}_1^\top \mathbf{X}_1 & \mathbf{X}_1^\top \mathbf{X}_2 \\ \mathbf{X}_2^\top \mathbf{X}_1 & \mathbf{X}_2^\top \mathbf{X}_2 \end{pmatrix}\\ &= \begin{pmatrix} \mathbf{I}_q \otimes \mathbf{1}_n^\top \mathbf{1}_n & \mathbf{I}_q \mathbf{1}_n^\top \otimes \mathbf{1}_q \\ \mathbf{I}_q^\top \otimes \mathbf{1}_n & \mathbf{1}_q^\top \mathbf{1}_q \otimes \mathbf{I}_n \end{pmatrix}\\ &= \begin{pmatrix} n\mathbf{I}_q & \mathbf{1}_q \mathbf{1}_n^\top \\ \mathbf{1}_n \mathbf{1}_q^\top & q\mathbf{I}_n \end{pmatrix},\end{aligned}$$

and based on (15) and (17) we get

$$\mathbf{H}_{**}^{-1} = \begin{pmatrix} n_a''\mathbf{I}_q & \mathbf{1}_q \mathbf{1}_n^\top \\ \mathbf{1}_n \mathbf{1}_q^\top & n_b''\mathbf{I}_n \end{pmatrix},$$

and (see Lemma A.1 in the appendix)

$$\mathbf{H}_{**} = \begin{bmatrix} \frac{1}{n''_a}(\mathbf{I}_q - \frac{n}{d}\mathbf{1}_q\mathbf{1}_q^\top) & \frac{1}{d}\mathbf{1}_q\mathbf{1}_n^\top \\ \frac{1}{d}\mathbf{1}_n\mathbf{1}_q^\top & \frac{1}{n''_b}(\mathbf{I}_n - \frac{q}{d}\mathbf{1}_n\mathbf{1}_n^\top) \end{bmatrix},$$

where $n''_a = n'_a + n, n''_b = n'_b + q$ and $d = N - n''_a n''_b$.

We can then use $\mathbf{H}_{**}$ to derive θ_{**} in (14):

$$\begin{aligned} \theta_{**} &= \mathbf{H}_{**}\left[\begin{pmatrix} n'_a\mathbf{A}_* \\ n'_b\mathbf{B}_* \end{pmatrix} + \begin{pmatrix} \mathbf{I}_q \otimes \mathbf{1}_n^\top \\ \mathbf{1}_q^\top \otimes \mathbf{I}_n \end{pmatrix}\mathbf{Y}\right] \\ &= \mathbf{H}_{**}\begin{pmatrix} n'_a\mathbf{A}_* + n\bar{\mathbf{Y}}^*_{.q} \\ n'_b\mathbf{B}_* + q\bar{\mathbf{Y}}_{n.} \end{pmatrix}. \end{aligned} \quad (18)$$

4 OLS^- and OLS^+ estimators

For model (4) we see that $n'_a = n'_b = 0$, i.e. $\mathbf{H}_*^{-1} = 0$ in (14) leads to the classical case in which $(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})$ is singular and the usual OLS estimator $\hat{\theta} = (\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^{-1}\tilde{\mathbf{X}}^\top\mathbf{Y}$ does not exist. For studies on such problems, see, e.g. Zellner (1971) and Dodge and Majumdar (1979). We proceed along the lines of Polasek and Liu (1997). For the stacked regression matrix $\tilde{\mathbf{X}} = (\mathbf{X}_1, \mathbf{X}_2)$ we use two expressions of $(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^-$, instead of $(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^{-1}$ for $\hat{\theta}$ above, to give two OLS^- estimators, and use the unique $(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^+$ to give the OLS^+ estimator.

Theorem 4.1 *Consider the two-way rank deficient MANOVA model (4)*

$$\mathbf{Y} = \mathbf{X}_1\mathbf{A} + \mathbf{X}_2\mathbf{B} + \mathbf{U}.$$

Two OLS^- estimators are given by (a) and (b):

(a)

$$\hat{\mathbf{A}} = \bar{\tilde{\mathbf{Y}}}^* = \frac{1}{n}\sum_{j=1}^{n}\tilde{\mathbf{Y}}_j^*,$$

$$\hat{\mathbf{B}} = \bar{\mathbf{Y}} = \frac{1}{q}\sum_{j=1}^{q}\mathbf{Y}_j;$$

and

(b)

$$\hat{\mathbf{A}} = \bar{\mathbf{Y}}^* = \frac{1}{n}\sum_{j=1}^{n} \mathbf{Y}_j^*,$$

$$\hat{\mathbf{B}} = \bar{\tilde{\mathbf{Y}}} = \frac{1}{q}\sum_{i=1}^{q} \mathbf{M}_n \mathbf{Y}_i,$$

where

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix}, \qquad \mathbf{Y}^* = \begin{pmatrix} \mathbf{Y}_1^* \\ \vdots \\ \mathbf{Y}_n^* \end{pmatrix} = \mathbf{L}_{nq}\mathbf{Y},$$

$$\tilde{\mathbf{Y}} = \begin{pmatrix} \tilde{\mathbf{Y}}_1 \\ \vdots \\ \tilde{\mathbf{Y}}_q \end{pmatrix}, \qquad \tilde{\mathbf{Y}}^* = \begin{pmatrix} \tilde{\mathbf{Y}}_1^* \\ \vdots \\ \tilde{\mathbf{Y}}_n^* \end{pmatrix},$$

$$\underset{n\times K}{\tilde{\mathbf{Y}}_i} = \mathbf{M}_n \mathbf{Y}_i, \qquad \underset{q\times K}{\tilde{\mathbf{Y}}_j^*} = \mathbf{M}_q \mathbf{Y}_j^*, \qquad i = 1, \ldots, q, \quad j = 1, \ldots, n.$$

The projection matrices $\mathbf{M}_n = \mathbf{I}_n - \mathbf{1}_n\mathbf{1}_n^\top/n$ *and* $\mathbf{M}_q = \mathbf{I}_q - \mathbf{1}_q\mathbf{1}_q^\top/q$ *are two centering matrices, and* $\mathbf{L}_{nq}$ *is the commutation matrix which reorders the dependent variable matrix* $\mathbf{Y}$*.*

Proof 4.1 *: The formula for* OLS^- *estimators in the model is*

$$\hat{\theta} = (\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^- \tilde{\mathbf{X}}^\top \mathbf{Y}.$$

(a) Using the g-inverse given in Lemma A.2(a) of the appendix we find

$$\begin{aligned} \hat{\theta} &= \begin{pmatrix} \hat{\mathbf{A}} \\ \hat{\mathbf{B}} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{n}\mathbf{M}_q & 0 \\ 0 & \frac{1}{q}\mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_q \otimes \mathbf{1}_n^\top \\ \mathbf{1}_q^\top \otimes \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix}. \end{aligned}$$

This can be split up into two equations and the OLS^- *estimator of the first component is*

$$\hat{\mathbf{A}} = \frac{1}{n}(\mathbf{M}_q \otimes \mathbf{1}_n^\top)\mathbf{Y}$$

$$
\begin{aligned}
&= \frac{1}{n}(\mathbf{1}_n^\top \otimes \mathbf{M}_q)\mathbf{Y}^* \\
&= \frac{1}{n}\sum_{j=1}^{n} \tilde{\mathbf{Y}}_j^* = \bar{\tilde{\mathbf{Y}}}^*,
\end{aligned}
$$

and the second component is

$$
\begin{aligned}
\hat{\mathbf{B}} &= \frac{1}{q}(\mathbf{1}_q^\top \otimes \mathbf{I}_n)\mathbf{Y} \\
&= \frac{1}{q}\sum_{j=1}^{q} \mathbf{Y}_j \\
&= \bar{\mathbf{Y}}.
\end{aligned}
$$

(b) Using the g-inverse given in Lemma A.2(b) in the appendix yields the following OLS^- *estimator:*

$$
\begin{aligned}
\hat{\theta} &= \begin{pmatrix} \hat{\mathbf{A}} \\ \hat{\mathbf{B}} \end{pmatrix} \\
&= \begin{pmatrix} \frac{1}{n}\mathbf{I}_q & 0 \\ 0 & \frac{1}{q}\mathbf{M}_n \end{pmatrix} \begin{pmatrix} \mathbf{I}_q \otimes \mathbf{1}_n^\top \\ \mathbf{1}_q^\top \otimes \mathbf{I}_n \end{pmatrix} \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_q \end{pmatrix}.
\end{aligned}
$$

This splits into two equations; the OLS^- *estimator of the first component is*

$$
\begin{aligned}
\hat{\mathbf{A}} &= \frac{1}{n}(\mathbf{I}_q \otimes \mathbf{1}_n^\top)\mathbf{Y} \\
&= \frac{1}{n}(\mathbf{1}_n^\top \otimes \mathbf{I}_q)\mathbf{Y}^* \\
&= \frac{1}{n}\sum_{j=1}^{n} \mathbf{Y}_j^* \\
&= \bar{\mathbf{Y}}^*,
\end{aligned}
$$

and the second component is

$$
\begin{aligned}
\hat{\mathbf{B}} &= \frac{1}{q}(\mathbf{1}_q^\top \otimes \mathbf{M}_n)\mathbf{Y} \\
&= \frac{1}{q}\sum_{i=1}^{q} \mathbf{M}_n \mathbf{Y}_i \\
&= \bar{\tilde{\mathbf{Y}}}.
\end{aligned}
$$

Theorem 4.2 *Consider the two-way rank deficient MANOVA model (4)*

$$\mathbf{Y} = \mathbf{X}_1\mathbf{A} + \mathbf{X}_2\mathbf{B} + \mathbf{U}.$$

The two components of the OLS$^+$ estimator are as follows:

$$\begin{aligned}\hat{\mathbf{A}} &= \frac{n}{m}\bar{\mathbf{Y}}^* + \frac{q}{m}\bar{\tilde{\mathbf{Y}}}^*,\\ \hat{\mathbf{B}} &= \frac{q}{m}\bar{\mathbf{Y}} + \frac{n}{m}\bar{\tilde{\mathbf{Y}}},\end{aligned}$$

where $m = q + n$.

Proof 4.2 *Using the Moore-Penrose inverse*
$(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^+$
given in Lemma A.2 in the appendix we obtain

$$\begin{aligned}\hat{\theta} &= (\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^+\tilde{\mathbf{X}}^\top\mathbf{Y}\\ &= \begin{pmatrix} \frac{1}{n}\mathbf{M}_q + \frac{n}{m^2}\mathbf{N}_q & \frac{1}{m^2}\mathbf{1}_q\mathbf{1}_n^\top \\ \frac{1}{m^2}\mathbf{1}_n\mathbf{1}_q^\top & \frac{1}{q}\mathbf{M}_n + \frac{q}{m^2}\mathbf{N}_n \end{pmatrix}\begin{pmatrix} \mathbf{I}_q\otimes\mathbf{1}_n^\top \\ \mathbf{1}_q^\top\otimes\mathbf{I}_n \end{pmatrix}\mathbf{Y}\\ &= \begin{bmatrix} (\frac{1}{m}\mathbf{I}_q + \frac{q}{mn}\mathbf{M}_q)\otimes\mathbf{1}_n^\top \\ \mathbf{1}_q^\top\otimes(\frac{1}{m}\mathbf{I}_n + \frac{n}{mq}\mathbf{M}_n) \end{bmatrix}\mathbf{Y},\end{aligned}$$

where $\mathbf{N}_q = \mathbf{I}_q - \mathbf{M}_q$ and $\mathbf{N}_n = \mathbf{I}_n - \mathbf{M}_n$.

The OLS$^+$ estimator for the component $\mathbf{A}$ is then

$$\begin{aligned}\hat{\mathbf{A}} &= [(\frac{1}{m}\mathbf{I}_q + \frac{q}{mn}\mathbf{M}_q)\otimes\mathbf{1}_n^\top]\mathbf{Y}\\ &= \frac{1}{m}(\mathbf{I}_q\otimes\mathbf{1}_n^\top)\mathbf{Y} + \frac{q}{mn}(\mathbf{M}_q\otimes\mathbf{1}_n^\top)\mathbf{Y}\\ &= \frac{n}{m}\bar{\mathbf{Y}}^* + \frac{q}{m}\bar{\tilde{\mathbf{Y}}}^*,\end{aligned}$$

and is

$$\begin{aligned}\hat{\mathbf{B}} &= [\mathbf{1}_q^\top\otimes(\frac{1}{m}\mathbf{I}_n + \frac{n}{mq}\mathbf{M}_n)]\mathbf{Y}\\ &= \frac{q}{m}\bar{\mathbf{Y}} + \frac{n}{m}\bar{\tilde{\mathbf{Y}}}\end{aligned}$$

for $\mathbf{B}$.

It is interesting to compare the *OLS* estimators for the model. We see that the two components $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ of the OLS^+ estimator are compromises of those of the two OLS^- estimators, respectively. The
projection matrices for the OLS^- and OLS^+ estimators are identical with the relation

$$\begin{aligned}\tilde{\mathbf{X}}(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^-\tilde{\mathbf{X}}^\top &= \tilde{\mathbf{X}}(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^+\tilde{\mathbf{X}}^\top \\ &= \mathbf{I}_q \otimes \mathbf{I}_n - \mathbf{M}_q \otimes \mathbf{M}_n,\end{aligned}$$

and we can obtain the unique *OLS* estimator $\tilde{\mathbf{X}}\hat{\theta}$ for $\tilde{\mathbf{X}}\theta$ by using any OLS^- or OLS^+ estimator $\hat{\theta}$.

A Appendix

Several results on matrix inverse and generalized inverses have been used to derive the normal OLS^{-1} estimator for the Bayesian two-way MANOVA model when it is nonsingular, and the OLS^- and OLS^+ estimators for the two-way MANOVA model which is rank deficient. These results are presented as follows:

Lemma A.1 *The posterior covariance matrix is*

$$\begin{aligned}\mathbf{H}_{**} &= \begin{pmatrix} n''_a\mathbf{I}_q & \mathbf{1}_q\mathbf{1}_n^\top \\ \mathbf{1}_n\mathbf{1}_q^\top & n''_b\mathbf{I}_n \end{pmatrix}^{-1} \\ &= \begin{bmatrix} \frac{1}{n''_a}(\mathbf{I}_q - \frac{n}{d}\mathbf{1}_q\mathbf{1}_q^\top) & \frac{1}{d}\mathbf{1}_q\mathbf{1}_n^\top \\ \frac{1}{d}\mathbf{1}_n\mathbf{1}_q^\top & \frac{1}{n''_b}(\mathbf{I}_n - \frac{q}{d}\mathbf{1}_n\mathbf{1}_n^\top) \end{bmatrix},\end{aligned}$$

where $n''_a = n'_a + n, n''_b = n'_b + q$ *and* $d = N - n''_a n''_b$.

Proof A.1 *Using the rule of the inverse for a partitioned matrix we can easily verify the above equality.*

Lemma A.2 *The two chosen generalized inverses of* $\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}}$ *used to give generalized inverses* $\tilde{\mathbf{X}}^- = (\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^-\tilde{\mathbf{X}}^\top$ *for* $\tilde{\mathbf{X}} = (\mathbf{I}_q \otimes \mathbf{1}_n\ ,\ \mathbf{1}_q \otimes \mathbf{I}_n)$ *are*

(a)

$$(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^- = \begin{pmatrix} \frac{1}{n}\mathbf{M}_q & 0 \\ 0 & \frac{1}{q}\mathbf{I}_n \end{pmatrix},$$

(b)

$$(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^- = \begin{pmatrix} \frac{1}{n}\mathbf{I}_q & 0 \\ 0 & \frac{1}{q}\mathbf{M}_n \end{pmatrix}.$$

The Moore-Penrose inverse of $\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}}$ *used in deriving the unique Moore-Penrose inverse* $\tilde{\mathbf{X}}^+ = (\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^+\tilde{\mathbf{X}}^\top$ *is*

$$(\tilde{\mathbf{X}}^\top\tilde{\mathbf{X}})^+ = \begin{pmatrix} \frac{1}{n}\mathbf{M}_q + \frac{n}{m^2}\mathbf{N}_q & \frac{1}{m^2}\mathbf{1}_q\mathbf{1}_n^\top \\ \frac{1}{m^2}\mathbf{1}_n\mathbf{1}_q^\top & \frac{1}{q}\mathbf{M}_n + \frac{q}{m^2}\mathbf{N}_n \end{pmatrix}.$$

Here,

$$\begin{aligned} \mathbf{M}_q &= \mathbf{I}_q - \frac{1}{q}\mathbf{1}_q\mathbf{1}_q^\top, \\ \mathbf{N}_q &= \frac{1}{q}\mathbf{1}_q\mathbf{1}_q^\top, \\ \mathbf{M}_n &= \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top, \\ \mathbf{N}_n &= \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top. \end{aligned}$$

Proof A.2 *Using the definitions of a generalized inverse and the Moore-Penrose inverse we can easily verify the above equalities.*

References

Dodge, Y. and D. Majumdar (1979). An algorithm for finding least square generalized inverses for classification models with arbitrary patterns. *J. Statist. Comput. Simul.* 9, 1-17.

Magnus, J.R. and H. Neudecker (1991). *Matrix Differential Calculus with Applications in Statistics and Econometrics.* Wiley, Chichester.

Polasek, W. (1995). *Theoretical Bayesian Statistics.*
University of Basel, Switzerland.

Polasek, W. and S. Liu (1997). On generalized inverses and Bayesian analysis in simple ANOVA models. *Student*, 2(2), 159-168.

Polasek, W. and Poetzelberger, K. (1994) Robust Bayesian methods in simple ANOVA Models. *J. of Planning and Inference*, 295-311.

Press, S.J. (1989). *Bayesian Statistics: Principles, Models and Applications.* Wiley, New York.

Rao, C.R. (1973). *Linear Statistical Inference and Applications*, 2nd ed. Wiley, New York.

Scheffé, H. (1959). *The Analysis of Variance.* Wiley, New York.

Searle, S.R., Casella, G. and C.E. McCulloch (1992). *Variance Components.* Wiley, New York.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics.* Wiley, New York.