Tatra Mt. Math. Publ. **22** (2001), 171–178

TATRA MOUNTAINS
Mathematical Publications

SOME ESTIMATION PROBLEMS IN A TWO-STAGE LINEAR MODEL – REVISITED

Júlia Volaufová

Dedicated to Lubomír Kubáček on the occasion of his 70th birthday

ABSTRACT. A two-stage linear model occurs in many different areas of statistical applications. It is defined as a set of two linear models connected with a common unknown subvector parameter of the mean. Estimators of linear functions of mean parameter that combine observations from both stages as well as estimators of linear functions of subvectors from each stage separately are investigated and their relations re-visited here.

1. Introduction

The two-stage linear model is a special case of a multistage linear model. It occurs in technical applications, notably in geodesic or geophysical setting, as well as in biomedical applications, as in meta analysis, when combining data from two clinical trials with non-identical experimental designs with a common treatment variable.

A two-stage linear model is defined as a set of linear models

$$
\begin{aligned}
Y_1 &= X_1\beta_1 + \varepsilon_1\,, \\
Y_2 &= C_{21}\beta_1 + X_2\beta_2 + \varepsilon_2\,,
\end{aligned}
\tag{1}
$$

where Y_i, $i = 1, 2$, are n_i-dimensional random vectors. The $n_i \times k_i$ full rank matrices X_i are known and the $n_2 \times k_1$-matrix C_{21} is also known. The random error vectors ε_i fulfil the distributional assumptions $E(\varepsilon_i) = 0$, $\operatorname{var}(\varepsilon_i) = V_i(\theta)$, $i = 1, 2$, $E(\varepsilon_1\varepsilon_2') = 0$, and all their moments up to the fourth order are finite. The vectors $\beta_i \in \mathbb{R}^{k_i}$, $i = 1, 2$, are unknown distributional parameters of the mean. The covariance matrices $V_i(\theta)$, $i = 1, 2$, are assumed to be nonnegative definite and depend on unknown vectors $\theta \in \Theta$. Assume that the parameter set Θ is a

2000 Mathematics Subject Classification: 62J05, 62F10.
Keywords: two-stage linear model, mixed linear model, minimum variance linear unbiased estimator.

product of two sets, $\Theta_1 \times \Theta_2$, where each Θ_i is a subset of an r_i dimensional space $\mathbb{R}^{r_i}$, $i = 1, 2$, and that Θ has an open interior. The functional dependence on θ is not specified, it may be linear or nonlinear as well. We assume that $V_i(\theta)$ are continuous in θ, $i = 1, 2$. Assume that there exists $\theta_0 \in \Theta$, such that $V_i(\theta_0)$ denoted by V_{0i}, $i = 1, 2$, are positive definite. The connection between the 1st and the 2nd model in (1) is the unknown vector parameter β_1. We shall refer to the two linear models in (1) as *two stages*. For the rest of the paper, we will denote by $R(A)$ the linear space spanned by columns of a matrix A. Vectors are considered to be columns, and the dimensions of matrices are noted as subscripts if necessary, otherwise they are omitted. An orthogonal projection onto $R(A)$ and onto the orthogonal complement of $R(A)$ is denoted by P_A and M_A, respectively.

2. Estimation

The concept of a multistage linear model was introduced and pursued in the 1980s by L. Kubáček. He devoted considerable attention to questions related to estimation of linear functions of β. Particularly, in a two-stage model, estimation of β_2 as a function of an available estimator of β_1 from the first stage only was investigated in great detail. Most of the results have been derived under special assumptions on covariance matrices $V_i(\theta)$. It was assumed either that they are fully known (see, e.g., [3] or [2]), or they depend linearly on a one dimensional parameter only. For the latter see, e.g., [7], [8], and [4].

Formally we may combine the two stages in (1) into a model that has the form:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} X_1 & 0 \\ C_{21} & X_2 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix} . \tag{2}$$

Model (2) represents a linear model with variance-covariance parameters, i.e.,

$$Y = X\beta + \varepsilon , \tag{3}$$

where X is the design matrix in (2) and ε is the error term in (2). Denote by $V(\theta) = \mathrm{Diag}\,\{V_i(\theta)\}$ the block diagonal covariance matrix of ε.

Consider a linear function of a parameter β presented in the form $P'\beta$ where P is a $(k_1 + k_2) \times q$-dimensional matrix, $q < k_1 + k_2$. Obviously, if $q = 1$ then we deal with a standard situation of $P'\beta$ being a scalar function of β. We deal here with a standard concept of estimability. A linear function $P'\beta$ is considered to be *estimable* in model (3) if and only if there exists a matrix L of appropriate dimensions such that $X'L = P$. An estimator $L^{*\prime}Y$ is said to be a Locally Minimum Variance Linear Unbiased Estimator (LMVLUE) of $P'\beta$ at θ_0 if it is unbiased, i.e., $E(L^{*\prime}Y) = P'\beta$ for all $\beta = (\beta_1', \beta_2')'$, and for any other unbiased estimator $L'Y$ of $P'\beta$ the difference of the covariance matrices $\mathrm{var}\,(L'Y) - \mathrm{var}\,(L^{*\prime}Y)$ is positive semidefinite at θ_0. If $L^{*\prime}Y$ and $L'\beta$ are both unbiased for $P'\beta$ and the difference in covariance matrices $\mathrm{var}\,(L'Y)-\mathrm{var}\,(L^{*\prime}Y)$ is positive semidefinite for all $\theta \in \Theta$, then $L^{*\prime}Y$ is said to be a Uniformly Minimum Variance Linear Unbiased Estimator (UMVLUE) of $P'\beta$. Under this setting there are two main questions that we are going to address here.

1. Under what conditions on the matrices X_i, $i = 1,2$, C_{21}, $V_1(\theta)$ and $V_2(\theta)$, do there exist UMVLUEs for all $P'\beta$ in (2), taking both vectors Y_1 and Y_2 into consideration?
2. Assume that from the first stage in (1) the UMVLUE of $P_1'\beta_1$ is available. Under what conditions do we not "gain" more information for estimating $P_1'\beta_1$ by additionally considering Y_2 in model (2)?

Since both matrices X_i, $i = 1,2$, are of full column rank, denote by Q_i, $i = 1,2$, the full row rank matrices that do not depend on the parameter θ and for which $Q_iX_i = I$ holds true. Consider a nonsingular matrix

$$T = \begin{pmatrix} I & 0 \\ -C_{21}Q_1 & I \end{pmatrix},$$

where the identity matrices have appropriate dimensions. Then, as it is easy to see, we can transform the model (3) by T to obtain

$$TY = TX\beta + T\varepsilon\,, \tag{4}$$

Notice that the equality $X'L = P$ can also be expressed as $X'T'T'^{-1}L = P$, which implies that all estimable functions of $P'\beta$ in model (3) are also estimable in model (4) and vice versa. Also notice that if $L'Y$ is a LMVLUE of $P'\beta$ in model (3) then it is LMVLUE in model (4) as well, and vice versa, hence the two models (3) and (4) are equivalent.

Now we can state a Theorem that generalizes the results presented in [7] and [4].

Theorem 1. *The* UMVLUE *for every estimable function $P'\beta$ of $\beta = (\beta_1', \beta_2')'$ in model* (2) *exists if and only if all three of the following conditions are met:*

1. $R\left(V_i(\theta)V_{0i}^{-1}X_i\right) \subset R(X_i)$ *for all* $\theta \in \Theta$, $i = 1,2$,
2. $R\big(V_2(\theta)V_{02}^{-1}C_{21} - C_{21}D_\theta\big) \subset R(X_2)$, *for all* $\theta \in \Theta$.

The matrix D_θ is defined by $V_1(\theta)V_{01}^{-1}X_1 = X_1D_\theta$.

The proof of the Theorem here follows the lines in [7] (or [4]) where special cases are investigated.

Proof. The necessary and sufficient conditions for the existence of UMVLUE for all estimable functions of β in a linear model with variance-covariance components have been investigated in the literature extensively. Here we consider the condition derived by S e e l y and Z y s k i n d in [6] and later generalized by K l e f f e in [1]. The general form of a necessary and sufficient condition under which in model (3) there exists a UMVLUE for all estimable functions of β can be expressed as

$$M_X V(\theta) V_0^{-1} X = 0 \qquad \forall \theta \in \Theta . \tag{5}$$

Recall that $V_0 = \operatorname{Diag}\{V_{0_i}\} = \operatorname{Diag}\{V_i(\theta_0)\}$ is positive definite according to our assumption.

Since T is nonsingular, the equality (5) is equivalent to

$$M_{TX} T V(\theta) V_0^{-1} X = 0 \qquad \forall \theta \in \Theta . \tag{6}$$

Notice that

$$M_{TX} = \begin{pmatrix} M_{X_1} & 0 \\ 0 & M_{X_2} \end{pmatrix} .$$

Developing the left side of (6) we get the matrix

$$\begin{pmatrix} M_{X_1} V_1(\theta) V_{01}^{-1} X_1 & 0 \\ -M_{X_2} C_{21} Q_1 V_1(\theta) V_{01}^{-1} X_1 + M_{X_2} V_2(\theta) V_{02}^{-1} C_{21} & M_{X_2} V_2(\theta) V_{02}^{-1} X_2 \end{pmatrix} . \tag{7}$$

The upper left block matrix is equal to zero if and only if $R\big(V_1(\theta) V_{01}^{-1} X_1\big) \subset R(X_1)$ for all $\theta \in \Theta$. Similarly, the condition from the statement of the Theorem holds for the lower right block matrix, if and only if $R\big(V_2(\theta) V_{02}^{-1} X_2\big) \subset R(X_2)$ for all θ.

Since $R\big(V_1(\theta) V_{01}^{-1} X_1\big) \subset R(X_1)$ for all θ, there exists a matrix D_θ such that $V_1(\theta) V_{01}^{-1} X_1 = X_1 D_\theta$. Hence $Q_1 V_1(\theta) V_{01}^{-1} X_1 = Q_1 X_1 D_\theta = D_\theta$. The bottom left block matrix in (7) is a zero matrix for all θ if and only if $R\big(V_2(\theta) V_{02}^{-1} C_{21} - C_{21} D_\theta\big) \subset R(X_2)$ for all θ. This concludes the proof of the Theorem. □

The simplest linear model fulfilling the assumptions of Theorem 1 occurs in a special case when $V_1(\theta) = \vartheta_1 H_1$ and $V_2(\theta) = \vartheta_2 H_2$, the matrices H_1 and H_2 are known and the parameters ϑ_1 and ϑ_2 are unknown real positive parameters. Condition 1 in this case is automatically met by choosing $V_{0i} = H_i$, $i = 1, 2$, and the condition 2 reduces to $R(C_{21}) \subset R(X_2)$. This coincides with the result derived in [7] and [4].

COROLLARY 1. *If at each stage of the original model there exists a* UMVLUE *for each estimable function of the corresponding subvector of β vector, it does not imply that there exists a* UMVLUE *for each estimable function of β in a combined model* (3).

COROLLARY 2. *Let $V(\theta^*) = I$ for some $\theta^* \in \Theta$. In such setting the identity matrix can be chosen for V_0 and hence necessary and sufficient conditions for the existence of* UMVLUE *for all estimable functions of β in model* (3) *reduce to*

1. $R\big(V_i(\theta)X_i\big) \subset R(X_i)$, *for all* $\theta \in \Theta$, $i = 1, 2$,
2. $R\big(V_2(\theta)C_{21} - C_{21}D^*_\theta\big) \subset R(X_2)$,

where the matrix D^*_θ is defined by $V_1(\theta)X_1 = X_1 D^*_\theta$.

In this case the ordinary least squares estimator (OLS) *at each stage is the* UMVLUE *of every estimable function of a corresponding subvector of β.*

COROLLARY 3. *If $V(\theta^*) = I$ for some $\theta^* \in \Theta$ then under the conditions of Theorem 1 the* OLS *of every estimable function of β in model* (2) *coincides with the* LMVLUE *for all θ and hence it is* UMVLUE.

P r o o f. One of possible necessary and sufficient conditions for OLS being the same as LMVLUE, expressed in terms of model (3), is $R\big(V(\theta)X\big) \subset R(X)$ for all θ. This is equivalent to the existence of a matrix, say A_θ, such $V(\theta)X = XA_\theta$. Assume A_θ to be partitioned in the same way as the matrix X with the blocks $(A_\theta)_{ij}$, $i = 1, 2$, of appropriate dimensions. Such matrix A_θ exists if and only if, in terms of model (2), the following equality holds true. For simplicity we omit the parenthesis and the subscripts θ.

$$\begin{aligned} V(\theta)X &= \begin{pmatrix} V_1(\theta)X_1 & 0 \\ V_2(\theta)C_{21} & V_2(\theta)X_2 \end{pmatrix} \\ &= \begin{pmatrix} X_1A_{11} & X_1A_{12} \\ C_{21}A_{11} + X_2A_{21} & C_{21}A_{12} + X_2A_{22} \end{pmatrix}. \end{aligned} \tag{8}$$

Now it is obvious that the conditions of the Theorem 1 assure the existence of A_{ij}, $i = 1, 2$. □

COROLLARY 4. *Sufficient conditions for $R\big(V_2(\theta)C_{21} - C_{21}D_\theta\big) \subset R(X_2)$ for all $\theta \in \Theta$ are*

1. $R(C_{21}) \subset R(X_2)$ *and*
2. $R\left(V_2(\theta)C_{21}\right) \subset R(C_{21})$ *for all* $\theta \in \Theta$

or

1. $R(C_{21}) \subset R(X_2)$ *and*
2. $R\big(V_2(\theta)C_{21}\big) \subset R(X_2)$ *for all* $\theta \in \Theta$

Comment 1. Notice that under the condition $R(C_{21}) \subset R(X_2)$, the unknown parameter $\beta = (\beta_1', \beta_2')'$ is not identifiable in the second stage of model (1) but it is identifiable in model (2).

Theorem 2. *Under the conditions of Corollary 2 the* UMVLUE *of vector parameters β_1 and β_2 in model* (2) *as well as in model* (4) *are $\hat{\beta}_1$ and $\hat{\beta}_2$ given by*

$$\hat{\beta}_1 = \left(X_1'X_1 + C_{21}'M_{X_2}C_{21}\right)^{-1}\left(X_1'Y_1 + C_{21}'M_{X_2}Y_2\right) \tag{9}$$

and

$$\begin{aligned}\hat{\beta}_2 = {} & (X_2'X_2)^{-1}X_2'\left(Y_2 + C_{21}\left(X_1'X_1 + C_{21}'M_{X_2}C_{21}\right)^{-1}C_{21}'M_{X_2}Y_2\right) \\ & - (X_2'X_2)^{-1}X_2'C_{21}\left(X_1'X_1 + C_{21}'M_{X_2}C_{21}\right)^{-1}X_1'Y_1 \,. \end{aligned} \tag{10}$$

Comment 2. Note that under a stronger condition $R(C_{21}) \subset R(X_2)$, $\hat{\beta}_1$ from Theorem 2 takes the form

$$\hat{\beta}_1 = (X_1'X_1)^{-1}X_1'Y_1 \,, \tag{11}$$

what is the OLS estimator of β_1 in the first stage of model (1), and

$$\hat{\beta}_2 = (X_2'X_2)^{-1}X_2'Y_2 - (X_2'X_2)^{-1}X_2'C_{21}(X_1'X_1)^{-1}X_1'Y_1 \,. \tag{12}$$

Hence, the condition $R(C_{21}) \subset R(X_2)$ establishes a sufficient condition under which the answer to question 2 on page 3 is that the UMVLUE of β_1 from the first stage of model (1) is UMVLUE also in the combined model (2).

P r o o f o f T h e o r e m 2. Under the conditions of Corollary 2 and from the Corollary 3 it follows that it suffices to express the OLS in model (2). □

Comment 3. Denote by $\widetilde{\beta_1} = (X_1'X_1)^{-1}X_1'Y_1$, the OLS from the first stage of model (1). Then both the estimators $\widehat{\beta_1}$ and $\widehat{\beta_2}$ in Theorem 2 can be expressed as functions of $\widetilde{\beta_1}$ and Y_2 only. One possible expression is

$$\begin{aligned}\widehat{\beta_1} = {} & \widetilde{\beta_1} - (X_1'X_1)^{-1}C_{21}'M_{X_2}\left(I + M_{X_2}C_{21}(X_1'X_1)^{-1}C_{21}'M_{X_2}\right)^{-1}M_{X_2}C_{21}\widetilde{\beta_1} \\ & + \left(X_1'X_1 + C_{21}'M_{X_2}C_{21}\right)^{-1}C_{21}'M_{X_2}Y_2 \end{aligned} \tag{13}$$

and

$$\begin{aligned}\widehat{\beta_2} = {} & (X_2'X_2)^{-1}X_2'Y_2 + (X_2'X_2)^{-1}X_2'C_{21}\left(X_1'X_1 + C_{21}'M_{X_2}C_{21}\right)^{-1}C_{21}'M_{X_2}Y_2 \\ & - (X_2'X_2)^{-1}X_2'C_{21} \\ & \times \left(\widetilde{\beta_1} - (X_1'X_1)^{-1}C_{21}'M_{X_2}\left(I + M_{X_2}C_{21}(X_1'X_1)^{-1}C_{21}'M_{X_2}\right)^{-1}M_{X_2}C_{21}\widetilde{\beta_1}\right). \end{aligned} \tag{14}$$

Again, it is obvious that if the stronger condition $R(C_{21}) \subset R(X_2)$ is met then $\widehat{\beta_1}$ reduces to $\widetilde{\beta_1}$ and $\widehat{\beta_2} = (X_2'X_2)^{-1}X_2'Y_2 - (X_2'X_2)^{-1}X_2'C_{21}\widetilde{\beta_1}$.

3. Examples

Example 1. For stage 1, consider a balanced one-way ANOVA model given by

$$Y_{ij} = \mu_i + e_{ij} \tag{15}$$

with μ_i, $i = 1, \ldots, a$, fixed unknown parameters, $j = 1, \ldots, l_1$, and $\operatorname{var}(e) = \sigma_1^2 I$. The second stage model is a two-way balanced fixed effects model given by

$$Y_{ijk} = \mu_i + \alpha_j + \varepsilon_{ijk}\,, \tag{16}$$

where again, $i = 1, \ldots, a$, $j = 1, \ldots, b$, $k = 1, \ldots, l_2$, and $\operatorname{var}(\varepsilon_{ijk}) = \sigma_2^2 I$. The parameters α_j are assumed to be fixed and unknown.

The two-stage model can be expressed as

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = \begin{pmatrix} I_a \otimes \mathbf{1}_{l_1} & 0 \\ I_a \otimes \mathbf{1}_b \otimes \mathbf{1}_{l_2} & \mathbf{1}_a \otimes I_b \otimes \mathbf{1}_{l_2} \end{pmatrix} \beta + \varepsilon^*, \tag{17}$$

where $\beta = (\mu', \alpha')'$, with $\mu = (\mu_1, \ldots, \mu_a)'$ and $\alpha = (\alpha_1, \ldots, \alpha_b)'$, $\varepsilon^* = (e', \varepsilon')'$ with e and ε defined analogously. The subscripts at I, the identity matrix, and at $\mathbf{1}$, the vector of all ones, indicate their dimensions. The covariance matrices for each stage are $\sigma_1^2 I$ and $\sigma_2^2 I$, respectively. It is easy to see that in the model (16) the parameters μ and α are not identifiable while in the combined model (17) they are. At the same time, while the first two conditions of Theorem 1 are met, the Condition 2, that reduces to

$$R\big((\sigma_2^2 - \sigma_1^2)(I_a \otimes \mathbf{1}_b \otimes \mathbf{1}_{l_2})\big) \subset R(\mathbf{1}_a \otimes I_b \otimes \mathbf{1}_{l_2})\,, \tag{18}$$

is fulfilled if and only $\sigma_1^2 = \sigma_2^2$, hence the conditions of the Theorem 1 are not met and the UMVLUEs for all estimable functions of β in this model do not exist.

Example 2. For the stage one, consider now a random one-way balanced ANOVA model given by

$$Y_{ij} = \mu + \alpha_i + e_{ij}\,, \qquad i = 1, \ldots, a_1\,,\ j = 1, \ldots, l_1\,, \tag{19}$$

with $\alpha = (\alpha_1, \ldots, \alpha_{a_1})'$ and $e = (e_{11}, \ldots, e_{al_1})'$ both random, fulfilling the distributional assumptions $E(\alpha) = 0$, $E(e) = 0$, $E(\alpha e') = 0$, $\operatorname{var}(\alpha) = \sigma_a^2 I$, and $\operatorname{var}(e) = \sigma_e^2 I$. Let the stage two be a mixed two-way balanced model,

$$Y_{ijk} = \mu + \alpha_i^* + \beta_j + \varepsilon_{ijk}\,, \qquad i = j \ldots, a_2\,,\ j = 1, \ldots, b,\ k = 1, \ldots l_2\,. \tag{20}$$

Assume μ and β to be fixed and, α^* and ε random. Let $E(\alpha^*) = 0$, $E(\varepsilon) = 0$, $E(\alpha^*\alpha') = 0$, $E(e\varepsilon') = 0$, $E(\alpha^*\varepsilon') = 0$, and $E(e\varepsilon') = 0$. The variance components of the second stage are defined by $\operatorname{var}(\alpha^*) = \sigma_{\alpha^*}^2 I$ and $\operatorname{var}(\varepsilon) = \sigma_\varepsilon^2 I$. We can use the notation of Theorem 1: $X_1 = \mathbf{1}_{a_1 l_1} = \mathbf{1}_{a_1} \otimes 1 \otimes \mathbf{1}_{l_1}$,

$X_2 = \mathbf{1}_{a_2} \otimes I_b \otimes \mathbf{1}_{l_2}$, $C_{21} = \mathbf{1}_{a_2 b l_2} = \mathbf{1}_{a_2} \otimes \mathbf{1}_b \otimes \mathbf{1}_{l_2}$, and the covariance matrices $V_1(\theta)$ and $V_2(\theta)$ are $\sigma_a^2(I_{a_1} \otimes \mathbf{1}_{l_1}\mathbf{1}_{l_1}') + \sigma_e^2 I$ and $\sigma_{\alpha^*}^2(I_{a_2} \otimes \mathbf{1}_b\mathbf{1}_b' \otimes \mathbf{1}_{l_2}\mathbf{1}_{n_2}') + \sigma_\varepsilon^2 I$, respectively. In this case, as it is easy to verify, the conditions of Theorem 1 are fulfilled and hence the UMVLUE in a model combining both stages exists for all estimable functions of $(\mu, \beta')'$.

REFERENCES

[1] KLEFFE, J.: *Simultaneous estimation of expectation and covariance matrix in linear models*, Math. Operationsforsch. Stat., Ser. Stat., No 3, Vol. 9, 1978, pp. 443–478.

[2] KUBÁČEK, L.: *Multistage regression model*, Apl. Mat. **31** (1986), 89–96.

[3] KUBÁČEK, L.: *Two-stage regression model*, Math. Slovaca **38** (1988), 383–383.

[4] KUBÁČEK, L.—KUBÁČKOVÁ, L.—VOLAUFOVÁ, J.: *Statistical models with linear structures*, Veda, Bratislava, 1995.

[5] RAO, C. R.: *Linear Statistical Inference and its Applications*, 2nd ed., J. Wiley, New York, 1973.

[6] SEELY, J.—ZYSKIND, G.: *Linear spaces and minimum variance unbiased estimation*, Ann. Math. Stat. **42** (1971), 691–703.

[7] VOLAUFOVÁ, J.: *Estimation of parameters of mean and variance in two-stage linear models*, Apl. Mat. **32** (1987), 1–8.

[8] VOLAUFOVÁ, J.: *Note on the estimation of parameters of the mean and the variance in n-stage linear models*, Apl. Mat. **33** (1988), 41–48.

Received February 22, 2001

Biostatistic and Data Management Core
Pennington Biomedical Research Center
6400 Perkins Road
Baton Rouge, LA 70808
U. S. A.

E-mail: volaufj@pbrc.edu