Wiktor Oktaba

Characterization of the multivariate Gauss-Markoff model with singular covariance matrix and missing values

Applications of Mathematics, Vol. 43 (1998), No. 2, 119--131

Persistent URL: http://dml.cz/dmlcz/134380

CHARACTERIZATION OF THE MULTIVARIATE GAUSS-MARKOFF MODEL WITH SINGULAR COVARIANCE MATRIX AND MISSING VALUES

Wiktor Oktaba, Lublin

(Received April 10, 1997)

Abstract. The aim of this paper is to characterize the Multivariate Gauss-Markoff model (MGM) as in (2.1) with singular covariance matrix and missing values. $MGMDP2$ model and completed $MGMDP2Q$ model are obtained by three transformations D, P and Q (cf. (3.21)) of MGM. The unified theory of estimation (Rao, 1973) which is of interest with respect to MGM has been used.

The characterization is reached by estimation of parameters: scalar σ^2 and linear combination $\lambda'\overline{B}$ ($\overline{B} = vecB$) as in (4.8), (4.6), (4.7) as well as by the model of the form (5.1) (cf. Th. 5.1). Moreover, testing linear hypothesis in the available model $MGMDP2$ by test function F as in (6.3) and (6.4) is considered.

It is known (Oktaba 1992) that ten quantities in models $MGMDP2$ and $MGMDP2Q$ are identical (invariant). They permit to say that formulas for estimation and testing in both models are identical (Oktaba et al., 1988, Baksalary and Kala, 1981, Drygas, 1983).

An algorithm and the $UMGMBO$ program for calculations concerning estimation and testing in MGM have been presented by Oktaba and Osypiuk (1993).

Keywords: multivariate Gauss-Markoff model, missing value, developed model, available model, completed model, elementary transformation, BLUE, estimation, testing, consistency, invariant

MSC 2000: 62H05

1. Review of literature

The fundamental technique of estimation of a single missing value has been introduced by Allan and Wishart [1]. The following methods of missing data played general role: 1) Yates' [18] iteration method, 2) Wilkinson's one [17], 3) Biggers one [4] for experimental designs, 4) covariance analysis of Bartlett [3], 5) Hartley's and Hocking's [7] method of maximum likelihood and many others. Yates's technique is connected with minimalizing the error sum of squares. In this way BLUE's of missing data are found.

R.A. Fisher [6] is the author of the rule that the residual sum of squares for the model with missing data is equal to the corresponding sum of squares if the missing values are replaced by the least square estimators.

The statistical literature concerning the topic of missing data is very large particularly after 1970. Special attention should be paid to the monograph by Little and Rubin [8]. Oktaba and al. [13], [14] present sufficient conditions for a linear transformation of a univariate Gauss-Markoff model $\varepsilon(y) = X\beta$, $D(y) = \sigma^2 V$ preserving information needed for the estimation of the expected value, the scalar σ^2, an estimable parametric function $\lambda'\beta$ and the test function F for verifying the linear hypothesis. Oktaba and al. [12], [13] discuss estimation and verification of hypotheses in some Zyskind-Martin [19] models with missing values as well as estimation of missing values in the general Gauss-Markoff model. Oktaba [10] presents the solution of prediction of missing values in the case of the multivariate Gauss-Markoff model.

The problem considered in the present paper is given by the title.

2. A multivariate Gauss-Markoff model with missing data

Let us consider MGM model with missing values of the form

$$(U, XB, \sigma^2\Sigma \otimes V) \tag{2.1}$$

known matrices $\Sigma > 0$, $V \geqslant 0$, an unknown scalar $\sigma^2 > 0$, and

$$E(U) = XB, \quad Cov(U) = D(U) = \sigma^2\Sigma \otimes V.$$

U is an $n \times p$ known matrix of observations, X - an $n \times d$ known matrix, B—an $d \times p$ unknown matrix of parameters, $\otimes$—the Kronecker symbol of product of matrices.

Assume that m values are missing and $np - m$ are available in the matrix U. For each pair (j, j'), $j, j' = 1, 2, \ldots, p$ of p columns of U there are at least four available observations which guarantee the calculation of covariance between characters j and j', so $n - m_j \geqslant 2$, where m_j denotes the number of missing observations in the j th column of U; $m = \sum_{j=1}^{p} m_j$.

Model (2.1) is consistent under the condition

$$R(U) \subset R(T), \tag{2.2}$$

where $T = V + XMX'$ and $M = M'$ is such a matrix that $R(X) \subset R(T)$ (Oktaba and al., [14]).

Let us note that $R(X) \subset R(T) \Leftrightarrow R(V) \subset R(T)$. Symbol $R(A)$ is reserved for the vector space spanned on the columns of the matrix A.

We shall use notation given in the following relation (C.R. Rao, [16]):

$$\begin{bmatrix} V & X \\ X' & 0 \end{bmatrix}^{-} = \begin{bmatrix} C_1 & C_2 \\ C_3 & -C_4 \end{bmatrix},$$

where

$$\begin{cases} C_1 = T^- - T^- X(X'T^-X)^- X'T^-, \\ C_3 = C_2' = (X'T^-X)^- X'T^-, \\ C_4 = (X'T^-X)^- - M. \end{cases} \tag{2.3}$$

Symbol A^- denotes an arbitrary g-inverse of the matrix A, i.e., any solution to $AA^-A = A$.

3. Transformations D, P and Q of the model MGM. Completed model $MGMDP2Q$

We wish to obtain models $MGMD$, $MGMDP$, $MGMDP1$, $MGMDP2$ and $MGMDP2Q$ by applying three transformations D, P and Q with respect to the multivariate model (2.1) with missing data.

3.1. Model $MGMD$

By developing the matrix U as in (2.1) we get the univariate model $MGMD$ of the form (Oktaba [10])

$$(\overline{U}, \overline{XB}, \sigma^2 \Sigma \otimes V) \Leftrightarrow (Y_D, X_D\,\beta_D, \sigma^2 V_D) \tag{3.1}$$

where

$$Y_D = \overline{U}, \quad X_D = I \otimes X, \quad \beta_D = \overline{B}, \quad V_D = \Sigma \otimes V,$$
$$T_D = V_D + X_D M_D X_D' = \Sigma \otimes T, \quad M_D = M_D' = \Sigma \otimes M,$$
$$Y_D \in R(T_D), \quad R(X_D) \subset R(T_D) \Leftrightarrow R(X) \subset R(T).$$

Here Y_D, X_D and V_D are known, β_D and σ^2 are unknown. Symbol $\overline{A}$ denotes the development of the matrix A.

Lemma 3.1. *We have (Oktaba [9], p. 161, corollary 3.1)*

$$\begin{bmatrix} V_D & X_D \\ X_D' & 0 \end{bmatrix}^{-} = \begin{bmatrix} C_{1D} & C_{2D} \\ C_{3D} & -C_{4D} \end{bmatrix},$$

where

$$C_{1D} = \Sigma^{-1} \otimes C_1, \quad C'_{2D} = C_{3D} = I_p \otimes C_3, \quad C_{4D} = \Sigma \otimes C_4 \tag{3.2}$$

The condition of consistency of the model in the form

$$Y_D \subset R(T_D) \tag{3.3}$$

is equivalent to formula (2.2).

[The proof is given in Oktaba ([11], th.2.1, pp. 128–129)].

3.2. Model $MGMGD$

By applying the elementary transformation P (Rao, [16], p. 17) we collect m missing values together to form subsequently a subvector $Y_{DP\,1}$ with m missing observations and a vector of $(np-m)$ observed values $Y_{DP\,2}$, i.e.

$$Y_{DP} = \begin{bmatrix} Y_{DP1} \\ Y_{DP2} \end{bmatrix} = P\overline{U} = PY_D. \tag{3.4}$$

Thus we get the model $MGMDP$ of the form

$$(Y_{DP}, X_{DP}\beta_{DP}, \sigma^2 V_{DP}) \tag{3.5}$$

under the notation and some with relations among models MGM, $MGMD$ and $MGMDP$:

$$\begin{aligned} Y_{DP} &= PY_D = P\overline{U} \subset R(T_{DP}), \quad X_{DP} = PX_D = P(I_p \otimes X), \\ \beta_{DP} &= \beta_D = \overline{B}, \quad V_{DP} = PV_D P' = P(\Sigma \otimes V)P', \\ M_{DP} &= M_D = \Sigma \otimes M, \quad T_{DP} = PT_D P' = P(\Sigma \otimes T)P' \end{aligned}$$

and

$$\begin{cases} P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix}, \quad P'P = PP' = I_{np}, \quad |P| = \pm 1, \quad (PAP')^- = PA^-P', \\ P_1 P_1' = I_m, \quad P_2 P_2' = I_{np-m}, \quad P_1 P_2' = 0, \quad P_2 P_1' = 0 \end{cases} \tag{3.6}$$

where P_1 and P_2 are $m \times np$ and $(np-m) \times np$ matrices, respectively.

Lemma 3.2. *(Oktaba, [10],* (3.2)*). Consistency in any one of the three models* MGM, $MGMD$ *and* $MGMDP$ *guarantees consistency in the other ones, i.e.*

$$R(U) \subset R(T) \Leftrightarrow \overline{U} \in R(T_D) \Leftrightarrow Y_{DP} \subset R(T_{DP}).$$

Lemma 3.3. *We have (Oktaba, [10])*

$$R(X_{DP}) \subset R(T_{DP}) \Leftrightarrow R(X_D) \subseteq R(T_D) \Leftrightarrow R(X) \subset R(T),$$

where

$$X_{DP} = \begin{bmatrix} X_{DP1} \\ X_{DP2} \end{bmatrix} = PX_D = P(I_p \otimes X) = \begin{bmatrix} P_1(I_p \otimes X) \\ P_2(I_p \otimes X) \end{bmatrix}$$

is an $np \times pd$ known matrix with

$$X_{DP1} = P_1(I_p \otimes X), \quad X_{DP2} = P_2(I_p \otimes X). \tag{3.7}$$

T_{DP2} is determined by the relation

$$T_{DP2} = PT_DP' = P(\Sigma \otimes T)P' \tag{3.8}$$

$$V_{DP} + X_{DP}M_{DP}X'_{DP} = \begin{bmatrix} T_{DP1} & T_{DP12} \\ T_{DP21} & T_{DP2} \end{bmatrix}$$

Lemma 3.4. *We have (Oktaba [10])*

$$\begin{aligned} C_{1DP} &= P(\Sigma^{-1} \otimes C_1)P', \\ C'_{2DP} &= C_{3DP} = (I_P \otimes C_3)P' \\ C_{4DP} &= \Sigma \otimes C_4 \end{aligned}$$

and

$$\begin{bmatrix} V_{DP} & X_{DP} \\ X'_{DP} & 0 \end{bmatrix}^{-} = \begin{bmatrix} C_{1DP} & C_{2DP} \\ C_{3DP} & -C_{4DP} \end{bmatrix}.$$

The available model $MGMDP2$ is of the form

$$(Y_{DP2}, X_{DP2}\beta_{DP2}, \sigma^2 V_{DP2}). \tag{3.9}$$

It is consistent under the condition $Y_{DP2} \in R(T_{DP2})$, where

$$T_{DP2} = V_{DP2} + X_{DP2}M_{DP2}X'_{DP2} = P_2T_DP'_2 = P_2(\Sigma \otimes T)P'_2$$

is a known matrix, since X_{DP2} is as in (3.7) and $M_{DP} = M_D = \Sigma \otimes M = M_{DP2}$ is known.

We have

$$R(X) \subset R(T), \quad R(X_D) \subset R(T_D), \quad R(X_{DP2}) \subset R(T_{DP2})$$

and

$$V_{DP} = PV_DP' = P(\Sigma \otimes V)P' = \begin{bmatrix} V_{DP1} & V_{DP12} \\ V_{DP21} & V_{DP2} \end{bmatrix} \tag{3.10}$$

where $V_{DP2} = P_2(\Sigma \otimes V)P_2'$.

Moreover,

$$\begin{bmatrix} V_{DP2} & X_{DP2} \\ X'_{DP2} & 0 \end{bmatrix}^{-} = \begin{bmatrix} C_{1DP2} & C_{2DP2} \\ C_{3DP2} & -C_{4DP2} \end{bmatrix}. \tag{3.11}$$

Hence we get

$$\begin{cases} C_{1DP2} = T^{-}_{DP2} - T^{-}_{DP2}X_{DP2}C_{3DP2}, \\ C_{3DP2} = C'_{2DP2} = (X'_{DP2}T^{-}_{DP2}X_{DP2})^{-}X'_{DP2}T^{-}_{DP2}, \\ C_{4DP2} = (X'_{DP2}T^{-}_{DP2}X_{DP2}) - M_{DP2}. \end{cases} \tag{3.12}$$

3.3. Model $MGMDP2Q$

The following known theorem (Oktaba, [10]) presents a predictor Y_{1DP} of missing values.

Theorem 3.5. *In the model* $MGMDP$ *as in* (3.5) *a predictor of the vector* Y_{DP1} *of* m *missing values is*

$$\hat{Y}_{DP1} = -(K_1 + K_1')^{-}(K_2 + K_3')Y_{DP2} = ZY_{DP2} \tag{3.13}$$

under the following two conditions:

$$\hat{Y}_{DP} = \begin{bmatrix} \hat{Y}_{DP1} \\ Y_{DP2} \end{bmatrix} \in R(T_{DP}) = R[P(\Sigma \otimes T)P'], \tag{3.14}$$

$$(K_2 + K_3')Y_{DP2} \in R(K_1 + K_1') \tag{3.15}$$

where K_1, K_2, K_3 *and* K_4 *are submatrices of the matrix*

$$C_{1DP} = PC_{1D}P' = P(\Sigma^{-1} \otimes C_1)P' = \begin{bmatrix} K_1 & K_2 \\ K_3 & K_4 \end{bmatrix}$$

and the $m \times (np - m)$ *matrix* Z *is of the form*

$$Z = -(K_1 + K_1')^{-}(K_2 + K_3') \tag{3.16}$$

where C_{1D} *is as in* (3.2).

Let us note that $K_1 = P_1 C_{1D} P_1'$, $K_2 = P_1 C_{1D} P_2'$, $K_3 = P_2 C_{1D} P_1'$, $K_4 = P_2 C_{1D} P_2'$. The predictor $\hat{Y}_{DP1}$ as in (3.14) is unbiased iff

$$X_{DP1} = Z X_{DP2}, \tag{3.17}$$

where $X_{DP} = [X_{DP1}' \vdots X_{DP2}']'$, $X_{DP2} = P_2(I_p \otimes X)$.

Premultiplying the vector Y_{DP2} as in (3.9) by

$$Q = \begin{bmatrix} Z \\ I_{np-m} \end{bmatrix}$$

we get the vector

$$Y_{DP2Q} = Q Y_{DP2} = \begin{bmatrix} Z \\ I \end{bmatrix} Y_{DP2} = \begin{bmatrix} \hat{Y}_{DP1} \\ Y_{DP2} \end{bmatrix} = \begin{bmatrix} Z Y_{DP2} \\ Y_{DP2} \end{bmatrix}$$

and the completed model $MGMDP2Q$ of the form

$$(Y_{DP2Q}, X_{DP2Q} \beta_{DP2Q}, \sigma^2 V_{DP2Q}). \tag{3.18}$$

We have

$$X_{DP2Q} = \begin{bmatrix} Z X_{DP2} \\ X_{DP2} \end{bmatrix} = Q X_{DP2} = \begin{bmatrix} X_{DP2Q1} \\ X_{DP2Q2} \end{bmatrix}$$

$$\begin{aligned} \beta_{DP2Q} &= \beta_{DP2} = \beta_{DP1} = \beta_{DP} = \overline{B}, \\ V_{DP2Q} &= Q V_{DP2} Q' \\ X_{DP2Q1} &= Z X_{DP2}, \quad X_{DP2Q2} = X_{DP2} \end{aligned} \tag{3.19}$$

The interpretation of the transformation Q is as follows. By $MGMDP2Q$ we define the model in which the missing values are completed by their predictors.

The model $MGMDP1$ of missing observations is

$$(Y_{DP1}, X_{DP1} \beta_{DP1}, \sigma^2 V_{DP1}). \tag{3.20}$$

We have two conditions for the model $MGMDP2Q$:

1. consistency of the form $Y_{DP2Q} \in R(T_{DP2Q})$
2. solution of the equations with respect to the predictor of missing values as in (3.15) where T_{DP} is as in (3.8).

By applying three transformations D, P and Q we get six models given in the following scheme:

$$MGM \xrightarrow{D} MGMD \xrightarrow{P} MGMDP \to \begin{bmatrix} MGMDP1 \\ MGMDP2 \end{bmatrix} \to MGMDP2Q \tag{3.21}$$

They are as in (2.1), (3.1), (3.5), (3.20), (3.9) and (3.18), respectively.

4. BLUE's of the estimable parametric function $\lambda'\beta$ in the models $MGMDP1$ of missing values and $MGMDP2$ of available values. Estimator of the scalar σ^2

Theorem 4.1. *In the models $MGMDP1$ and $MGMDP2$ as in* (3.20) *and* (3.9)*, respectively, we have*

$$\begin{cases} C_{1DPi} = P_i C_{1D} P_i', \\ C_{3DPi} = C_{2DPi}' = (I \otimes C_3) P_i' \\ C_{4DPi} = \Sigma \otimes C_4, \quad i = 1, 2, \end{cases} \tag{4.1}$$

where

$$\begin{bmatrix} V_{DPi} & X_{DPi} \\ X_{DPi}' & 0 \end{bmatrix}^{-} = \begin{bmatrix} C_{1DPi} & C_{2DPi} \\ C_{3DPi} & -C_{4DPi} \end{bmatrix}, \quad i = 1, 2, \tag{4.2}$$

$C_1, C_3 = C_2'$, C_4 *are as in* (2.3) *and* P_1 *and* P_2 *are as in (3.6),* $T_D = \Sigma \otimes T$, $C_{1D} = \Sigma^{-1} \otimes C_1$ *(cf. Oktaba, [9],* (3.2) *and* (3.4)*).*

P r o o f. From the definition we have

$$C_{1DPi} = T_{DPi}^{-} - T_{DPi}^{-} X_{DPi} (X_{DPi}' T_{DPi}^{-} X_{DPi})^{-} X_{DPi}' T_{DPi}^{-}, \quad i = 1, 2. \tag{4.3}$$

In virtue of (3.6) and

$$\begin{cases} X_{DPi} = P_i (I_p \otimes X) \\ T_{DPi} = P_i T_D P_i' = P_i (\Sigma \otimes T) P_i', \quad i = 1, 2 \end{cases} \tag{4.4}$$

we obtain

$$\begin{aligned} X_{DiPi}' T_{DPi}^{-} &= (I_p \otimes X') P_i^{-} [P_i (\Sigma \otimes T) P_i']^{-} = (I_p \otimes X') P_i' P_i (\Sigma^{-1} \otimes T^{-}) P_i' \\ &= (I_p \otimes X')(\Sigma^{-1} \otimes T^{-}) P_i' = (\Sigma^{-1} \otimes X' T^{-}) P_i', \end{aligned}$$

$$T_{DPi}^{-} X_{DPi} = P_i (\Sigma^{-1} \otimes T^{-} X)$$

and

$$\begin{aligned} (X_{DiPi}' T_{DPi}^{-} X_{DPi})^{-} &= [(\Sigma^{-1} \otimes X' T^{-}) P_i' P_i (I_p \otimes X)]^{-} \\ &= [(\Sigma^{-1} \otimes X' T^{-})(I_p \otimes X)]^{-} = (\Sigma^{-1} \otimes X' T^{-} X)^{-}. \end{aligned} \tag{4.5}$$

Thus in virtue of (4.3) and (4.4)

$$\begin{aligned} C_{1DPi} &= P_i (\Sigma^{-1} \otimes T^{-}) P_i' - P_i (\Sigma^{-1} \otimes T^{-} X)(\Sigma^{-1} \otimes X' T^{-} X)^{-} (\Sigma^{-1} \otimes X' T^{-}) P_i' \\ &= P_i \{ (\Sigma^{-1} \otimes T^{-}) - (\Sigma^{-1} \otimes T^{-} X)[\Sigma \otimes (X' T^{-} X)^{-}](\Sigma^{-1} \otimes X' T^{-}) \} P_i' \\ &= P_i \{ (\Sigma^{-1} \otimes T^{-}) - [\Sigma^{-1} \otimes T^{-} X (X' T^{-} X)^{-} X' T^{-}] \} P_i' = P_i [\Sigma^{-1} \otimes C_1] P_i' \\ &= P_i C_{1D} P_i' \end{aligned}$$

where $T_D = \Sigma \otimes T$, $C_{1D} = \Sigma^{-1} \otimes C_1 \quad i = 1, 2$.

Now we calculate C_{4DPi} when $M_{DPi} = M_{DP} = M_D = \Sigma \otimes M$, $i = 1, 2$. Since (4.5) holds, we get

$$\begin{aligned} C_{4DP} &= (X'_{DPi} T^-_{DPi} X_{DPi})^- - M_{DPi} = (\Sigma^{-1} \otimes X'T^-X)^- - \Sigma \otimes M \\ &= \Sigma \otimes (X'T^-X)^- - \Sigma \otimes M = \Sigma \otimes [(X'T^-X)^- - M] = \Sigma \otimes C_4. \end{aligned}$$

We prove that $C_{3DPi} = (I \otimes C_3)P'_i$, $i = 1, 2$. In fact, for $i = 1, 2$ we have

$$C_{3DPi} = (X'_{DPi} T^-_{DPi} X_{DPi})^- X'_{DPi} T^-_{DPi}.$$

Let us note that

$$\begin{aligned} X'_{DPi} T^-_{DPi} &= (I_p \otimes X')P'_i \, [P_i(\Sigma \otimes T)P'_i]^- = (I_p \otimes X')P'_i P_i (\Sigma \otimes T)^- P'_i \\ &= (I \otimes X')(\Sigma^{-1} \otimes T^-)P'_i = (\Sigma^{-1} \otimes X'T^-)P'_i \end{aligned}$$

and

$$\begin{aligned} (X'_{DPi} T^-_{DPi} X_{DPi})^- &= [(\Sigma^{-1} \otimes X'T^-)P'_i P_i (I_p \otimes X)] \\ &= [(\Sigma^{-1} \otimes X'T^-)(I_p \otimes X)]^- = (\Sigma^{-1} \otimes X'T^-X)^- \\ &= \Sigma \otimes (X'T^-X)^-. \end{aligned}$$

Then

$$\begin{aligned} C_{3DPi} &= [\Sigma \otimes (X'T^-X)^-](\Sigma^{-1} \otimes X'T^-)P'_i \\ &= (I \otimes (X'T^-X)^- X'T^-)P'_i = (I \otimes C_3)P'_i \end{aligned}$$

□

Theorem 4.2. *BLUE's of the estimable parametric function $\lambda'\beta$ in model MGMDPi $(i = 1, 2)$ are of the form (cf. (3.19))*

$$\lambda' \hat{\beta}_{DPi} = \lambda' \widehat{\widehat{B}} \tag{4.6}$$

where

$$\hat{\beta}_{DPi} = C_{3DPi} Y_{DPi} = (I \otimes C_3) P'_i Y_{DPi}. \tag{4.7}$$

P r o o f. The formulae as in (4.6) and (4.7) are obtained directly from (4.1) for C_{3DP1} and C_{3DP2} if we use the result for $\hat{\beta}$ from the theory of unified estimation (Rao, [16], p. 298; (4i, 3.2)). □

R e m a r k 4.1. Let us note that the dispersion matrices of $\lambda'\hat{\beta}_{DP1}$ and $\lambda'\hat{\beta}_{DP2}$ (Rao, loc.cit.) are the same:

$$V(\lambda'\hat{\beta}_{DP1}) = V(\lambda'\hat{\beta}_{DP2}) = \sigma^2\lambda' C_{4DP1}\lambda = \sigma^2\lambda' C_{4DP2}\lambda = \sigma^2\lambda'(\Sigma \otimes C_4)\lambda.$$

Theorem 4.3. *The unbiased estimator of* σ^2 *in the available model* $MGMDP2$ *is of the form*

$$\hat{\sigma}^2_{eDP2} = \frac{Y'_{DP2}C_{1DP2}Y_{DP2}}{tr(C_{1DP2}V_{DP2})}, \tag{4.8}$$

where $\mathrm{tr}(C_{1DP2}V_{DP2}) = r(V_{DP2}\vdots X_{DP2}) - r(X_{DP2})$ *denotes the number of degrees of freedom.* C_{1DP2} *is as in* (4.1), $V_{DP2} = P_2(\Sigma \otimes V)P'_2$; Y_{DP2} *is as in* (3.4).

P r o o f. To prove it, it is sufficient to use (4.1) and apply in $MGMDP2$ the formula for the estimator $\hat{\sigma}^2$ given by Rao ([16], p. 298, 4i, 3.4). □

5. The completed matrix model $MGMDP2QP'D^{-1}$ of the form $[\hat{U}, E(\hat{U}), \mathbf{D}(\hat{U})]$

Applying two transformations $P' = P^{-1}$ and D^{-1} with respect to the complete vector model $MGMDP2Q$ (Oktaba, [11], 140–156) we obtain the completed matrix model $MGMDP2QP'D^{-1}$ of the form

$$[\hat{U}, E(\hat{U}), \mathbf{D}(\hat{U})]. \tag{5.1}$$

$\hat{U}$ is the $n \times p$ matrix obtained from the matrix U with missing values by replacing the vector Y_{DP1} of missing values by the predictor $\hat{Y}_{DP1}$ as in (3.13).

The transformations of the model $MGMDP2Q$ into $MGMDP2QP'D^{-1}$ are presented in the following scheme:

$$MGMDP2Q \xrightarrow{P'} MGMDP2QP' \xrightarrow{D^{-1}} MGMDP2QP'D^{-1} \Leftrightarrow (5.1). \tag{5.2}$$

The symbol D^{-1} denotes the transformation of a column vector into a matrix; this is the inverse transformation with respect to D (development of matrix).

Theorem 5.1. *In the multivariate Gauss-Markoff model* (2.1) *with missing values under the condition*

$$X_{DP1} = ZX_{DP2} \tag{5.3}$$

with Z *as in* (3.16), X_{DP1} *and* X_{DP2} *as in* (3.7), *we have for* (5.1)*:*

$$E(\hat{U}) = XB, \tag{5.4}$$

$$\mathbf{D}(\hat{U}) = \sigma^2 (P_1' Z + P_2') V_{DP2} (P_1' Z + P_2')' \tag{5.5}$$

where P_1 *and* P_2 *are as in* (3.6), *and* V_{DP2} *is as in* (3.10).

P r o o f. a) Let us note that the predictor

$$\hat{Y}_{DP1} = ZY_{DP2} \tag{5.6}$$

is unbiased under condition (5.3).

In fact, in virtue of (3.9), (5.6) and $\beta_{DP1} = \beta_{DP2}$ we have $EY_{DP1} = X_{DP1}\beta_{DP1}$ and $E(\hat{Y}_{DP1}) = ZEY_{DP2} = ZX_{DP2}\beta_{DP2} = ZX_{DP2}\beta_{DP1} = X_{DP1}\beta_{DP1}$.

b) We shall show that (5.3) is a consequence of $\hat{Y}_{DP1} = ZY_{DP2}$. In fact, we assume that $E\hat{Y}_{DP1} = E(Y_{DP1})$. It means that $ZE(Y_{DP2}) = X_{DP1}\beta_{DP1}$ and in virtue of (3.9) $ZX_{DP2}\beta_{DP2} = X_{DP1}\beta_{DP1} = X_{DP1}\beta_{DP2}$ for each vector β_{DP2}, so we get (5.3).

c) We prove that $\hat{U}$ is unbiased under (5.4). Now as a result of unbiasedness of predictor $\hat{Y}_{DP1}$ under condition (5.3) we obtain

$$E(\hat{Y}_{DP1}) = X_{DP1}\beta_{DP1}.$$

Since $E(Y_{DP2}) = X_{DP2}\beta_{DP2}$ we get

$$\begin{aligned} E(\hat{\overline{U}}) &= P_1' E(\hat{Y}_{DP1}) + P_2' E(Y_{DP2}) = P_1' X_{DP1}\beta_{DP1} + P_2' X_{DP2}\beta_{DP2} \\ &= (P_1' \vdots P_2') \begin{bmatrix} X_{DP1} \\ X_{DP2} \end{bmatrix} \beta_{DP1} = P' X_{DP} \overline{B} = P' P (I_p \otimes X) \overline{B} = \overline{XB}. \end{aligned}$$

Thus $E(\hat{Y}) = XB$.

d) Since $\mathbf{D}(\hat{Y}_{DP2Q}) = \sigma^2 Q V_{DP2} Q'$ we have $\mathbf{D}(\hat{U}) = \mathbf{D}(\hat{\overline{U}}) = \mathbf{D}(P'\hat{Y}_{DP2Q}) = P'\mathbf{D}(\hat{Y}_{DP2Q})P' = \sigma^2 P' Q V_{DP2} Q' P = \sigma^2 (P_1' \vdots P_2') \begin{bmatrix} Z \\ I \end{bmatrix} V_{DP2} (Z' \vdots I) \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \sigma^2 (P_1' Z + P_2') V_{DP2} (Z' P_1 + P_2)$. □

6. Estimation and testing in the available model $MGMDP2$

In the paper of Oktaba ([11], pp. 135–6, Th. 2.6) ten quantities in the available $MGMDP2$ and completed $MGMDP2Q$ models are presented. These quantities are invariants with respect to the predictor of the vector of missing values and g-inverses. So we get the following results:

1. BLUE for a parametric function $\lambda'\overline{B}$
2. sufficient and necessary conditions for estimability of this function
3. variance and covariance of BLUE's
4. unbiased estimators $\hat{\sigma}^2_{eDP2}$ and $\hat{\sigma}^2_{eDP2Q}$ of the scalar $\sigma^2 > 0$ (cf. (4.11)),
5. sufficient and necessary conditions of the consistency of the linear hypothesis

$$H_0\colon\ N\beta_{DP2} = \varphi^0 \tag{6.1}$$

6. test functions F_{DP2} and F_{DP20} for H_0.

Theorem 6.1. *In the $MGMDP2$ model under the assumption of normality $Y_{DP2} \sim N(X\beta_{DP2}, \sigma^2 V_{DP2})$ we have*

1° BLUE of $\lambda'\beta_{DP2}$ is equal to(6.2)

$$\lambda'\hat{\beta}_{DP2} = \lambda' C_{3DP2} Y_{DP2}, \tag{6.2}$$

where $\hat{\beta}_{DP2}$ and C_{3DP2} are as in (4.7) *and* (4.1)*, respectively, in $MGMDP2$.*

2° The formula for the unbiased estimator $\hat{\sigma}^2_{eDP2}$ is given as in (4.8)*.*

3° The test function F for the hypothesis (6.1) *is of the form*

$$F_{DP2} = \frac{u'[\mathbf{D}(u)]^- u}{r[\mathbf{D}(u)] \cdot \hat{\sigma}^2_{eDP2}} \tag{6.3}$$

with

$$\nu_{DP2} = r[D(u)] \quad \text{and} \quad \nu_{eDP2} = tr(C_{1DP2} V_{DP2}) \tag{6.4}$$

degrees of freedom where $\hat{\sigma}^2_{eDP2}$ is as in (4.8)*. The covariance matrix of the vector*

$$u = N\hat{\beta}_{DP2} - \varphi$$

is

$$Cov(u) = D(u) = NC_{4DP2}N',$$

where C_{4DP2} is as in (4.1)*, $\hat{\beta}_{DP2}$ is as in* (4.7)*.*

P r o o f. To prove the theorem it is enough to apply the unified theory of estimation (Rao, [16]) to the available model $MGMDP2$. □

References

[1] *Allan, F.E., Wishart, J.*: A method of estimating the yield of a missing plot in field experimental work. Jour. Agr. Sci. *20* (1930), 399–406.

[2] *Baksalary, J.K., Kala, R.*: Linear transformations preserving best linear unbiased estimators in a general Gauss-Markoff model. Ann. St. *9* (1981), no. 4, 913–916.

[3] *Bartlett, M.S.*: Some examples of statistical methods of research in agriculture and applied biology. J.R. Statist. Soc. Suppl. *4* (1937), 137–170.

[4] *Biggers, J.D.*: The estimation of missing and mixed-up observations in several experimental designs. Biometrika *16* (1959), 91–105.

[5] *Drygas, H.*: Sufficiency and completeness in the general Gauss-Markoff model. Sankhyā 45A, (1983), 88–99.

[6] *Fisher, R.A.*: The Design of Experiments. Edinburgh (1960).

[7] *Hartley, H.O., Hocking, R.R.*: The analysis of incomplete data. Biometrics *4* (1971), 783–823.

[8] *Little, R.J.A., Rubin, D.B.*: Statistical Analysis with Missing Data. J. Wiley, New York, 1987.

[9] *Oktaba, W.*: Estimation in the general multivariate Gauss-Markoff model with the known covariance matrix. XIX Colloquium Metodol. Agrobiom. Warsaw, Polish Academy of Sciences, 1989, pp. 156–169. (In Polish.)

[10] *Oktaba, W.*: Predictor of the vector of missing observations in the general multivariate Gauss-Markoff model. Probastat'91, Proceedings of the 10th Conference on Probability and Mathematical Statistics, Bratislava, Czechoslovakia. 1991, pp. 51–67.

[11] *Oktaba, W.*: Invariants in estimation and testing in the available and completed multivariate generalized Gauss-Markoff models with missing values. XXII Colloquium Metodol. Agrobiom.. Warsaw, Polish Academy of Sciences, 1992, pp. 125–139. (In Polish.)

[12] *Oktaba, W., Kornacki, A. and Wawrzosek, J.*: Estimation and verification of hypotheses in some Zyskind-Martin models with missing values. Biom. J. *27, 7* (1985), 733–740.

[13] *Oktaba, W., Kornacki, A. and Wawrzosek, J.*: Estimation of missing values in the general Gauss-Markoff model. Statistics *17, 2* (1986), 167–177.

[14] *Oktaba, W., Kornacki, A. and Wawrzosek, J.*: Invariants linearly sufficient transformations of the general Gauss-Markoff model. Estimation and testing. Scand. J. Statist. *15* (1988), 117–124.

[15] *Oktaba, W. and Osypiuk, Z.*: Program for estimation and testing in the multivariate generalized Gauss-Markoff model with missing values and known covariance matrix. XXIII Colloquium Metodol. Agrobiom.. Warsaw, Polish Academy of Sciences, 1993, pp. 192–211. (In Polish.)

[16] *Rao, C.R.*: Linear Statistical Inference and its Applications. sec. ed. New York, 1973.

[17] *Wilkinson, G. N.*: A general recursive procedure for analysis of variance. Biometrika *57* (1970), 19–46.

[18] *Yates, F.*: The analysis of replicated experiments when the fields results are incomplete. Emp. Jour. Exp. Agr. *1* (1933), 235–244.

[19] *Zyskind, G., Martin, F.B.*: On best linear estimation and a general Gauss-Markoff theorem in linear models with arbitrary nonnegative covariance structure. SIAM J. Appl. Math. *17* (1969), 1190–1202.

Author's address: *Wiktor Oktaba*, Institute of Applied Mathematics, Department of Mathematical Statistics, Agricultural University, Akademicka 13, 20-934 Lublin, Poland, e-mail `sek314@ursus.ar.lublin.pl`.