ACTA UNIVERSITATIS LODZIENSIS
FOLIA OECONOMICA 156, 2002

*Paweł R. Pordzik**

THE RECURSIVE LEAST SQUARES ESTIMATION OF PARAMETRIC FUNCTIONS IN THE GENERAL LINEAR MODEL

Abstract. The technique of recursive least squares estimation for the standard regression model is extended to the general linear model with possibly singular dispersion matrix of error term. It is shown how to update the minimum dispersion linear unbiased estimate of a given vector of parametric functions with respect to additional sample data which are to be successively incorporated to the inference base.

Key words: general linear model, recursive estimation, least squares estimation of parametric functions.

1. INTRODUCTION AND PRELIMINARIES

The technique of recursive estimation is commonly applied when estimates of parameters in the model have to be adjusted with respect to additional information contained in successively available sample data. Introduced by Placket (1950) for the standard linear regression model, this technique under different assumptions on design and/or dispersion matrix was developed in a series of papers, cf. McGilchrist and Sandland (1979), McGilchrist, Sandland and Hennessy (1983), Haslett (1985), Sant'anna (1989) and Liski (1990). Our aim is to extend the approach for the possibly singular linear model. Following the theory of least squares estimation in the general linear model, we shall derive recursively-oriented formula for the minimum dispersion linear unbiased estimate (MDLUE) of a given vector of parametric functions. This allows us to unify procedures

* Dr., Department of Mathematical and Statistical Methods, Agricultural University of Poznań.

for updating the estimates of parameters with respect to additional sample data as well as linear restrictions which are to be successively incorporated to the inference base. The diagnostic of recursive residuals, as a tool for monitoring stability of the singular linear model used for categorical data, cf. Bonett, Woodward and Bentler (1985), can be mentioned here as an example of application for our results.

We consider for an i-th step of analyzing data the general linear model

$$\mathcal{M}_i = \{y_i, X_i\boldsymbol{\beta}, \sigma^2 V_i\} \tag{1}$$

in which y_i is an observable random vector with expectation $E(y_i) = X_i\boldsymbol{\beta}$ and dispersion matrix $D(y_i) = \sigma^2 V_i$, where X_i and V_i are known fixed matrices of arbitrary rank, while a vector $\boldsymbol{\beta}$ and a positive scalar σ^2 are unknown parameters. We assume that the model $\mathcal{M}_i$ is consistent, that is

$$y_i \in \mathscr{C}(X_i : V_i)$$

where $\mathscr{C}(X_i : V_i)$ denotes the column space of the partitioned matrix $(X_i : V_i)$. Let the object of our interest be a given vector of parametric functions $\boldsymbol{\varkappa} = \boldsymbol{K}\boldsymbol{\beta}$ estimable in the model $\mathcal{M}_i$, i.e. $\mathscr{C}(\mathrm{K}') \subseteq \mathscr{C}(\mathrm{X}_i')$ and let $\hat{\boldsymbol{\varkappa}}_i = \boldsymbol{K}\hat{\boldsymbol{\beta}}_i$ stand for its MDLUE obtained under the model $\mathcal{M}_i$. Now suppose that, in addition to (1), successive data are of the form

$$\mathcal{M} = \{y, X\boldsymbol{\beta}, \sigma^2 V\}, \ y \in \mathscr{C}(X : V) \tag{2}$$

where X and V are known matrices; the latter being zero-matrix, if (2) represents a set of linear restrictions imposed on the vector $\boldsymbol{\beta}$. Pooling both sources of information, the inference base is $\mathcal{M}_{i+1}$, wherein $y_{i+1} = (y_i' : y')'$, $X_{i+1} = (X_i' : X')'$ and $V_{i+1} = Diag(V_i, V)$. In order to avoid recalculations while estimating $\boldsymbol{\varkappa}$ under the model $\mathcal{M}_{i+1}$, we show how to adjust $\hat{\boldsymbol{\varkappa}}_i'$ with respect to additional data given in $\mathcal{M}$. For shortening the formulae presented in this paper, we assume throughout that the vectors y_i and y are uncorrelated. If it is not the case and extra information, say, $\{y_0, X_0\boldsymbol{\beta}, \sigma^2 V_0\}$ is such that $Cov(y_i, y_0) = \sigma^2 V_{i0}$, then by $\mathcal{M}$ we mean transformed data $y = y_0 - V_{i0}' V_i^- y_i$, $X = X_0 - V_{i0}' V_i^- X_i$ and $V = V_0 - V_{i0}' V_i^- V_{i0}$ which one obtains while diagonalizing the blocks of dispersion matrix in the pooled data model; in subsequence, we denote by A^- and $r(A)$, respectively, a g-inverse and rank of a matrix A.

2. CONSISTENCY CONDITION FOR THE POOLED DATA MODEL

Due to possible singularity of the dispersion matrix in the model $\mathscr{M}_i$ and $\mathscr{M}$, the consistency of inference base (while pooling data) is to be considered. Clearly, obtaining a condition under which the model $\mathscr{M}_{i+1}$ is not self-contradictory poses no problem. Its derivation, however, allows us to introduced notation needed in the sequel. The pooled data model is consistent if and only if

$$(\boldsymbol{I}-\boldsymbol{T}_{i+1}\boldsymbol{T}_{i+1}^{-})\boldsymbol{y}_{i+1}=\boldsymbol{0} \tag{3}$$

where $\boldsymbol{T}_{i+1}=\boldsymbol{V}_{i+1}+\boldsymbol{X}_{i+1}\boldsymbol{X}'_{i+1}$. By the well known formula for inverting a nonnegative definite (n.n.d.) partitioned matrix, one of g-inverses of $\boldsymbol{T}_{i+1}$ takes the form

$$\begin{pmatrix} \boldsymbol{T}_i^- & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} + \begin{pmatrix} -\boldsymbol{T}_i^-\boldsymbol{X}_i\boldsymbol{X}' \\ \boldsymbol{I} \end{pmatrix} \boldsymbol{F}^-(-\boldsymbol{X}\boldsymbol{X}_i'\boldsymbol{T}_i^- : \boldsymbol{I}) \tag{4}$$

where $\boldsymbol{F}=\boldsymbol{V}+\boldsymbol{X}(\boldsymbol{I}-\boldsymbol{C}_i)\boldsymbol{X}'$ and

$$\boldsymbol{C}_i=\boldsymbol{X}_i'\boldsymbol{T}_i^-\boldsymbol{X}_i \tag{5}$$

Substituting (4) to (3), one easily shows that the consistency condition for $\mathscr{M}_{i+1}$ can be expressed as $(\boldsymbol{I}-\boldsymbol{F}\boldsymbol{F}^-)(\boldsymbol{y}-\boldsymbol{X}\boldsymbol{g}_i)=\boldsymbol{0}$, where

$$\boldsymbol{g}_i=\boldsymbol{X}_i'\boldsymbol{T}_i^-\boldsymbol{y}_i \tag{6}$$

In consequence, since the matrix $\boldsymbol{I}-\boldsymbol{C}_i$ is n.n.d., the following statement holds to be true.

Proposition 1. Let each of the models $\mathscr{M}_i$ and $\mathscr{M}$ be consistent. Then the pooled data model $\mathscr{M}_{i+1}$ is not self-contradictory if and only if

$$\boldsymbol{y}-\boldsymbol{X}\boldsymbol{g}_i\in\mathscr{C}(\boldsymbol{V}:\boldsymbol{X}(\boldsymbol{I}-\boldsymbol{C}_i)) \tag{7}$$

Note that if the model $\mathscr{M}_i$ is weakly singular, that is, $\mathscr{C}(\boldsymbol{X}_i)\subseteq\mathscr{C}(\boldsymbol{V}_i)$, then $\boldsymbol{I}-\boldsymbol{C}_i=(\boldsymbol{I}+\boldsymbol{X}_i'\boldsymbol{V}_i^-\boldsymbol{X}_i)^{-1}$ and thus (7) holds trivially. The same conclusion on consistency of $\mathscr{M}_{i+1}$ can be drawn when $\mathscr{C}(\boldsymbol{X})\subseteq\mathscr{C}(\boldsymbol{V})$.

3. THE RECURSIVE ESTIMATION OF PARAMETRIC FUNCTIONS

Let $\hat{\varkappa}_i = K\hat{\beta}_i$ denote the minimum dispersion linear unbiased estimate of $\varkappa = K\beta$ in the model $\mathcal{M}_i$. Our purpose is to adjust $\hat{\varkappa}_i$ with respect to additional information given in $\mathcal{M}$. By the theory of least squares estimation in the general linear model, cf. Rao (1971), the MDLUE of $\varkappa$ under $\mathcal{M}_{i+1}$ can be expressed as

$$\hat{\varkappa}_{i+1} = KC_{i+1}^{-}g_{i+1} \tag{8}$$

where C_{i+1} and g_{i+1} are defined as in (5) and (6), respectively. We begin by noting that, in view of (4), it follows

$$C_{i+1} = C_i + N'F^{-}N \tag{9}$$

and

$$g_{i+1} = g_i + N'F^{-}(y - Xg_i) \tag{10}$$

where $N = X(I - C_i)$. Since the matrix $I - C_i$ is n.n.d., it holds $\mathscr{C}(N) \subset \mathscr{C}(F)$ and, by Lemma 1 in Kala and Kłaczyński (1988), one of g-inverses of C_{i+1} takes the form

$$C_i^{\sim} - C_i^{\sim}GC_i^{\sim} + (C_i^{\sim}G - I)Q(Q'GQ)^{-}Q'(GC_i^{\sim} - I) \tag{11}$$

where $C_o^{\sim}$ is a nonnegative definite g-inverse of C_i, $G = N'H^{-}N$, $H = NC_i^{\sim}N' + F$ and $Q = I - C_i^{\sim}C_i$. Substituting (10) and (11) into (8) and making use of the assumption (7) and equalities $(K:g_i)Q = 0$ and $GC_i^{\sim}N' = N' - N'H^{-}F$ (the latter being a consequence of $\mathscr{C}(N) \subseteq \mathscr{C}(H)$) we obtain

$$\hat{\varkappa}_{i+1} = KC_i^{\sim}g_i - KC_i^{\sim}N'S(NC_i^{\sim}g_i - y + Xg_i) \tag{12}$$

where

$$S = H^{-} - H^{-}NQ(Q'N'H^{-}NQ)^{-}Q'N'H^{-} \tag{13}$$

To link recursively $\hat{\varkappa}_i$ and $\hat{\varkappa}_{i+1}$, let us now observe that $N'SNQ = 0$ which, together with $g_i \in \mathscr{C}(C)$, and $y - Xg_i \in \mathscr{C}(H)$, implies that the formula (12) is invariant with respect to the choice of a g-inverse of C_i and, consequently, it can be written as

$$K\hat{\beta}_{i+1} = K\hat{\beta}_i + KZ'S(y - X\hat{\beta}_i) \tag{14}$$

where $\hat{\beta}_i = C_i^{-}g_i$, $Z = X(C_i^{-} - I)$. Furthermore, note that the formula (14) is invariant with respect to the choice of a g-inverse of H and $Q'N'H^{-}NQ$;

putting $C_i^{\sim} = C_i^- C_i C_i^{-\prime}$ we get $NQ = X(I - C_i^- C_i)$ and $H = ZCZ' + F$. Similar arguments applied for the dispersion matrix $D(\hat{\varkappa}_{i+1}) = \sigma^2 K(C_{i+1}^- - I)K'$ lead to the following equation

$$D(\hat{\varkappa}_{i+1}) = \mathrm{D}(\hat{\varkappa}_i) - \sigma^2 KZ'SZK' \tag{15}$$

For the special case, where $V = \mathbf{0}$, the problem of adjusting $\hat{\varkappa}_i$ with respect to linear restrictions imposed on $\boldsymbol{\beta}$ was discussed by Kala and Kłaczyński (1988). Extending their approach, by the formula (14), we bring together procedures for updating the least squares estimates of parametric functions with respect to additional sample data as well as non-stochastic information on parameters in the model.

Let us now consider a situation when additional set of data $\mathcal{M}$ is such that either (i) $\mathscr{C}(X') \subseteq \mathscr{C}(X_i')$ or (ii) $\mathscr{C}(X') \cap \mathscr{C}(X_i') = \{\mathbf{0}\}$; this takes place, e.g.,when one by one observation and/or linear restriction is to be incorporated to the model $\mathcal{M}_i$. We first focus our attention on updating $\hat{\varkappa}_i$ under the assumption (i); this certainly is the case in the linear regression model $\mathcal{M}_i$ wherein the matrix X_i is of the full column rank. Since $NQ = \mathbf{0}$, simplifying the formulae (14) and (15) leads to the following statement.

Proposition 2. The MDULE of $\varkappa = K\boldsymbol{\beta}$ under the pooled data linear model $\mathcal{M}_{i+1}$ wherein $\mathscr{C}(X') \subseteq \mathscr{C}(X_i')$, can be expressed as

$$K\hat{\boldsymbol{\beta}}_{i+1} = K\hat{\boldsymbol{\beta}}_i + KZ'L^-(y - X\hat{\boldsymbol{\beta}}_i) \tag{16}$$

where $\hat{\boldsymbol{\beta}}_i = C_i^- g_i$, $Z = X(C_i^- - I)$ and $\mathrm{L} = X(C_i^- - I)X' + V$. Moreover, it holds $D(K\hat{\boldsymbol{\beta}}_{i+1}) = D(K\hat{\boldsymbol{\beta}}_i) - \sigma^2 KZ'L^- ZK'$.

For completing the recurrence, note that one of g-inverses of the matrix C_{i+1} takes the form

$$C_{i+1}^- = C_i^- - (C_i^- - I)X'(X(C_i^- - I)X' + V)^- X(C_i^- - I) \tag{17}$$

The formula (16) was stated by Liski (1990) under the additional assumption $\mathscr{C}(X_i) \subseteq \mathscr{C}(V_i)$, which would clearly not allow for more than one recursive adjusting of $\hat{\varkappa}_i$ with respect to linear restrictions. At the end of this section, concerning with the weakly singular linear model, we shall present its alternative form. We now turn to the case (ii) where $\mathscr{C}(X') \cap \mathscr{C}(X_i') = \{\mathbf{0}\}$. Taking into account the equality $r(NQ) = r(N) -$ - $\dim\{\mathscr{C}(N') \cap \mathscr{C}(C_i)\}$ and the fact that (ii) implies $\mathscr{C}(N') \cap \mathscr{C}(C_i) = \{\mathbf{0}\}$, we have $r(Q'N'H^- NQ) = r(N'H^- N)$. Hence by Lemma 2.2.5 in Rao and Mitra (1971), the matrix $Q(Q'N'H^- NQ)^- Q'$ is g-inverse of $N'H^- N$ and, consequently, $N'SH = \mathbf{0}$. Assuming consistency of the model $\mathcal{M}_{i+1}$ which implies $y - X\hat{\boldsymbol{\beta}}_i \in \mathscr{C}(H)$, the following conclusion can be drawn.

Proposition 3. For the MDULE of $\varkappa = \boldsymbol{K\beta}$ under the model $\mathcal{M}_{i+1}$, wherein $\mathscr{C}(\boldsymbol{X}') \cap \mathscr{C}(\boldsymbol{X}_i') = \{\boldsymbol{0}\}$, it holds

$$\boldsymbol{K}\hat{\boldsymbol{\beta}}_{i+1} = \boldsymbol{K}\hat{\boldsymbol{\beta}}_i \tag{18}$$

The procedure stated above enables us to apply the well known techniques of recursive residuals for monitoring stability of the general linear models (here, let us only mention that a singular dispersion matrix of error term occurs naturally, if finite randomization processes are included in the model construction).

At the end, let us restate the solution to the problem of recursive estimation under the weakly singular model $\mathcal{M}_{i+1}$ where $\mathscr{C}(\boldsymbol{X}_{i+1}) \subseteq \mathscr{C}(\boldsymbol{V}_{i+1})$. It is known that the MDLUE of $\varkappa = \boldsymbol{K\beta}$ in the model $\mathcal{M}_{i+1}$ can be expressed as $\hat{\varkappa}_{i+1} = \boldsymbol{K}\boldsymbol{C}_{i+1}^{-}\boldsymbol{g}_{i+1}$ where, from now to the end of this paper, $\boldsymbol{C}_{i+1} = \boldsymbol{X}_{i+1}'\boldsymbol{V}_{i+1}^{-}\boldsymbol{X}_{i+1}$ and $\boldsymbol{g}_{i+1} = \boldsymbol{X}_{i+1}'\boldsymbol{V}_{i+1}^{-}\boldsymbol{y}_{i+1}$. Following much in the same way as before (14), with $\boldsymbol{N}$ and $\boldsymbol{F}$ replaced by $\boldsymbol{X}$ and $\boldsymbol{V}$, respectively, we obtain

$$\boldsymbol{K}\hat{\boldsymbol{\beta}}_{i+1} = \boldsymbol{K}\hat{\boldsymbol{\beta}}_i + \boldsymbol{K}\boldsymbol{Z}_w'\boldsymbol{S}_w'(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_i) \tag{19}$$

and

$$D(\boldsymbol{K}\hat{\beta}_{i+1}) = D(\boldsymbol{K}\hat{\beta}_i) - \sigma^2\boldsymbol{K}\boldsymbol{Z}_w'\boldsymbol{S}_w^{-}\boldsymbol{K}' \tag{20}$$

where $\hat{\boldsymbol{\beta}}_i = \boldsymbol{C}_i^{-}\boldsymbol{g}_i$, $\boldsymbol{Z}_w = \boldsymbol{X}\boldsymbol{C}_i^{-}$, $\boldsymbol{S}_w = \boldsymbol{H}_w(\boldsymbol{I} - \boldsymbol{X}_w(\boldsymbol{X}_w'\boldsymbol{H}_w'\boldsymbol{X}_w)^{-}\boldsymbol{X}_w'\boldsymbol{H}_w^{-})$ with $\boldsymbol{H}_w = \boldsymbol{Z}_w\boldsymbol{C}_i\boldsymbol{Z}_w' + \boldsymbol{V}$ and $\boldsymbol{X}_w = \boldsymbol{X}(\boldsymbol{I} - \boldsymbol{C}_i^{-}\boldsymbol{C}_i)$. As previously, we may conclude that $\hat{\varkappa}_{i+1} = \hat{\varkappa}_i$, if $\mathscr{C}(\boldsymbol{X}') \cap \mathscr{C}(\boldsymbol{X}_i') = \{\boldsymbol{0}\}$, and

$$\boldsymbol{K}\hat{\boldsymbol{\beta}}_{i+1} = \boldsymbol{K}\hat{\boldsymbol{\beta}}_i + \boldsymbol{K}\boldsymbol{C}_i^{-}\boldsymbol{X}'(\boldsymbol{V} + \boldsymbol{X}\boldsymbol{C}_i^{-}\boldsymbol{X}')^{-}(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{\beta}}_i) \tag{21}$$

if $\mathscr{C}(\boldsymbol{X}') \subseteq \mathscr{C}(\boldsymbol{X}_i')$; this extends for the case of weakly singular linear model the procedure stated by McGilchrist, Sandland and Hennessy (1983). Furthermore, assuming a full column rank of the matrix $\boldsymbol{X}_{i+1}$ and non-singularity of $\boldsymbol{V}_{i+1}$ the formula given in (19) coincides with the result by McGilchrist and Sandland (1979) and Hennessy (1985). It is to be emphasized that extending the approach of recursive least squares estimation for possibly singular model enables us to unify procedures for updating the estimates of parametric functions with respect to both additional sample data and linear restrictions imposed on the location parameters of the model. The problem of recursive adjusting the scale parameter is not considered in this note and still remains to be solved in a context of the general linear model.

REFERENCES

Bonett D. G., Woodward J. A., Bentler P. M. (1985), *Some Extensions of a Linear Model For Categorical Variables*, „Biometrics", **41**, 745–750.

Haslett S. (1985), *Recursive Estimation of the General Linear Model with Dependent Errors and Multiple Additional Observations*, Austral. J. Statist, **27**., 183–188.

Liski E. P. (1990), *Comparing Generalized Mixed Estimators with Respect to Covariance Matrix in Linear Regression Model*, „Statistics", **21**, 1–6.

McGilchrist C. A. Sandland R. L. (1979), *Recursive Estimation of the General Linear Model with Dependend Errors*, J. Roy. Statistical Soc., **B41**, 65–68.

McGilchrist C. A., Sandland R. L., Hennessy J. L. (1983), *Generalized Inverses Used in Recursive Estimation of the General Linear Model*, Austral. J. Statist, **25**, 321–328.

Plackett R. L. (1950), *Some Theorems in Least Squares*, „Biometrika", **37**, 149–157.

Rao C. R. (1971), *Unified Theory of Linear Estimation*. „Sankhya", **A33**, 371–394; *Corrigendum, ibidem*, **34**(1972), 194, 477.

Rao C. R., Mitra S. K. (1971), *Generalized Inverse of Matrices and Its Applications*, Wiley, New York.

Sant'anna A. P. (1989), *Generalized Inverses in Recursive Least Squares Estimation with Multiple Additional Observations*, J. Statist. Comput. Simul., **33**, 53–63.

Paweł R. Pordzik

REKURENCYJNA ESTYMACJA FUNKCJI PARAMETRYCZNYCH METODĄ NAJMNIEJSZYCH KWADRATÓW

(Streszczenie)

W pracy uogólniona została technika rekurencyjnej estymacji funkcji parametrycznych metodą najmniejszych kwadratów w ogólnym modelu liniowym. Proponowana procedura umożliwia aktualizację estymatorów zarówno ze względu na dodatkową stochastyczną, jak i niestochastyczną informację o parametrach modelu.