Natural estimation of variances in a general finite discrete spectrum linear regression model

Martina Hančová
Institute of Mathematics, P. J. Šafárik University in Košice
Jesenná 5, 040 01 Košice, Slovakia
e-mail: martina.hancova@upjs.sk

Abstract

The method of "natural" estimation of variances in a general (orthogonal or nonorthogonal) finite discrete spectrum linear regression model of time series is suggested. Using geometrical language of the theory of projectors a form and properties of the estimators are investigated. Obtained results show that in describing the first and second moment properties of the new estimators the central role plays a matrix known in linear algebra as the Schur complement. Illustrative examples with particular regressors demonstrate direct applications of the results.

Key words: time series; finite discrete spectrum linear regression model; natural estimators of variance components; orthogonal and oblique projectors; the Schur complement

1 Introduction

In recent articles (Štulajter 2003, Štulajter & Witkovský 2004, Štulajter 2006) authors have introduced and investigated a time series model called *the finite discrete spectrum linear regression model* or shortly FDSLRM, which represents time series modeling and predicting by linear regression models (see Brockwell & Davis 1991, Christensen 2002, Štulajter 2002) — the alternative approach to the most popular and well-known Box-Jenkins methodology (e.g. Box et al 1994).

The class of FDSLRM models whose mean values are given by linear regression and error terms are characterized by a purely finite discrete spectrum and white noise, offer applications in a wide range of real situations. In practice we usually need to estimate not only mean value parameters, but also unknown parameters of the FDSLRM covariance function. One solution of this problem was just given in Štulajter & Witkovský (2004) who used the double ordinary least squares estimator (DOOLSE) and obtained invariant unbiased, not consistent quadratic estimators.

The used approach however works only for the orthogonal version of the FDSLRM and in some cases it can give negative estimates. There are also such cases that we have to use a numerical nonlinear constrain optimization procedure to compute the DOOLSE estimates. In general it means that we have no explicit expression for estimators, which causes a difficult theoretical study of their properties.

Because of these reasons we suggest in sec. 2 of the article an alternative method of estimating unknown variance parameters of the covariance function, called by us "natural" estimation, appropriate for the FDSLRM with and without the assumption of orthogonality and leading to estimations, which are always from the given parametric space. Moreover the method is based on the least square approach as DOOLSE so estimating does not require the normality assumption as it is in case of ML, REML or MINVAR estimation (Searle et al 1992). In sec. 3 using theory of projectors, summarized e.g. in recent works of Ben-Israel (2003) or Galántai (2004), we obtain the first and second moment properties of estimators. Final sec. 4 includes illustrative examples in which we apply developed theoretical results.

In the rest of the introduction we establish notation and recapitulate used model and basic results from Štulajter (2003), (2004) providing a starting point and assumptions for our considerations.

A model of time series $X(.)$ is said to be *the finite discrete spectrum linear regression model* (FDSLRM), if $X(.)$ satisfies

$$X(t) = \sum_{i=1}^{k} \beta_i f_i(t) + \sum_{j=1}^{l} Y_j v_j(t) + w(t); t = 1, 2, ..., \tag{1}$$

where

$\beta = (\beta_1, \beta_2, ..., \beta_k)' \in \mathbb{E}^k$ is a vector of unknown regression parameters;

$Y = (Y_1, Y_2, ..., Y_l)'$ is a $l \times 1$ random vector with zero mean value, $E[Y] = 0$, and with covariance matrix $Cov(Y) = diag(\sigma_j^2)$ of size $l \times l$, where unknown variances $\sigma_j^2 \geq 0$ for all $j = 1, 2, ..., l$;

$f_i(.); i = 1, 2, ..., k$ and $v_j(.); j = 1, 2, ..., l$ are known real functions defined on $\mathbb{E}$;

$w(.)$ is white noise time series with the variance $D[w(t)] = \sigma^2 > 0$ and it is uncorrelated with Y.

We denote the unknown variance parameters of Y and $w(.)$, which are also variance parameters of the FDSLRM, by $\nu = (\sigma^2, \sigma_1^2, ..., \sigma_l^2)'$. Under the FDSLRM assumptions direct computation applied to the definition of the time series covariance function $R(s,t)$ yields its expression $R_\nu(s,t) = \sigma^2 \delta_{s,t} + \sum_{j=1}^{l} \sigma_j^2 v_j(s) v_j(t)$;

$s, t = 1, 2, \ldots$ with the parameter ν belonging to the parametric space $\Upsilon = (0, \infty) \times \langle 0, \infty)^l$.

The basic result dealing with any finite observation of the FDSLRM time series — random vector $X = (X(1), \ldots, X(n))'$ — says that the observation X satisfies the following linear regression model (also called the FDSLRM model):

$$X = F\beta + \varepsilon, E(\varepsilon) = 0, Cov(\varepsilon) = \sigma^2 I_n + \sum_{j=1}^{l} \sigma_j^2 V_j \text{ is a p.d. matrix,} \qquad (2)$$

where

$F = (f_1\ f_2\ \ldots\ f_k) \in \mathbb{E}^{n \times k}$ is the design matrix of the model with columns $f_i = (f_i(1), \ldots, f_i(n))'; i = 1, 2, \ldots, k;$

$V_j = v_j v_j' \in \mathbb{E}^{n \times n}; v_j = (v_j(1), v_j(2), \ldots, v_j(n))'; j = 1, 2, \ldots, l$ are matrices describing the structure of covariance matrix $Cov(\varepsilon) \equiv \Sigma_\nu$.

The FDSLRM model (2) is said to be *orthogonal*, if $f_i \perp v_j$ for all $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, l$ and $v_i \perp v_j$ for all $i, j = 1, 2, \ldots, l, i \neq j$. In this article we do not assume validity of the orthogonality conditions what will be also reminded by calling the model *a nonorthogonal* FDSLRM.

Model (2) is equivalent to a model belonging to the class of linear mixed models (see e.g. McCulloch & Searle 2001, Christensen 2002)[1]:

$$X = F\beta + VY + w, E(w) = 0, Cov(w) = \sigma^2 I_n, Cov(Y, w) = 0, \qquad (3)$$

where $V = (v_1\ v_2\ \ldots\ v_l) \in \mathbb{E}^{n \times l}$ and random vector $w = (w(1), \ldots, w(n))'$ is a finite observation of white noise $w(.)$. Symbols $F, \beta, Y, w(.)$ and $v_j; j = 1, 2, \ldots, l$ have the same meaning as above.

Since the observation of the FDSLRM is a special case of the linear mixed model, the problem of estimating ν is related to the problem of estimating variance-covariance components in linear mixed models studied besides Štulajter & Witkovský (2004) e.g. for more general cases in Rao & Kleffe (1988), Volaufová & Witkovský (1991) or Searle et al. (1992).

We shall assume that both matrices $F \in \mathbb{E}^{n \times k}$ and $V \in \mathbb{E}^{n \times l}$ are of full column rank[2], i.e. $r(F, V) = k + l$ and number $k + l + 1$ of unknown parameters β and ν, which arise in the FDSLRM (1), is smaller than length n of a realization $x = (x_1, x_2, \ldots, x_n) \in \mathbb{E}^n$ of finite observation X.

Finally we shall employ the following notation (Galántai 2004) resulting from using theory of projectors: $P_{\mathscr{N}} \equiv$ *the orthogonal projector onto some subspace* $\mathscr{N}$; $P_{\mathscr{N}^\perp} \equiv M_{\mathscr{N}} \equiv$ *the orthogonal projector onto the orthogonal complement of*

[1] In this case unobservable vector β is frequently called a vector of "fixed effects" and Y is an unobservable vector of "random effects".

[2] To have no problems in distinguishing between a matrix product (FV) and $\begin{pmatrix} F & V \end{pmatrix}$ as matrix F augmented by V, we will frequently write the matrix $\begin{pmatrix} F & V \end{pmatrix}$ as (F, V).

$\mathscr{N}$; $P_{\mathscr{N},\mathscr{O}} \equiv$ *the oblique projector onto* $\mathscr{N}$ *along* $\mathscr{O}$; $\mathscr{L}(H) = \{Hx | x \in \mathbb{E}^n\} \equiv$ *the column space of* $H \in \mathbb{E}^{m\times n}$; $\mathcal{N}(H) = \{x \in \mathbb{E}^n | Hx = 0\} \equiv$ *the null space of* $H \in \mathbb{E}^{m\times n}$.

2 "Natural" variance component estimation in the FDSLRM

2.1 Definition of the "natural" estimators

Variances of the FDSLRM $\sigma_j^2 = Cov(Y_j) = E(Y_j^2); j = 1, 2, \ldots, l$ so that if components Y_j were known, a "natural estimator" of σ_j^2 would be just Y_j^2. These unobservable "natural estimators" and their name are identical with those used in MINQUE estimating (Rao & Kleffe 1988, Searle et al 1992, Christensen 2002).

Although Y is a random vector, according to McCulloch & Searle (2001) it is convenient to consider the linear regression model conditional on unobservable realizations $y \in \mathbb{E}^l$ of Y. In such models mean values are $E(X|Y = y) = F\beta + Vy$ and realization y of Y can be understood as other unknown mean value parameters of model (3). From this viewpoint FDSLRM model (3) is nothing else than a regression model, linear with respect to unknown regression parameters $(\beta, y)' \in \mathbb{E}^{k+l}$ and with covariance matrix $Cov(X|Y = y) = Cov(w) = \sigma^2 I_n$.

These facts mean that for σ^2 it is natural to use the same unbiased double least square estimator as in classical linear regression models (Štulajter 2002). Moreover if we estimate y, then according to the above-mentioned idea of unobservable natural estimators a real estimator for σ_j^2 should be square of given estimator of y.

These considerations motivate the following definition. Let us consider for a FDSLRM observation X of time series $X(.)$ the following classical linear regression model

$$X = \begin{pmatrix} F & V \end{pmatrix} \begin{pmatrix} \beta \\ y \end{pmatrix} + w;\ E(w) = 0,\ Cov(w) = E(ww') = \sigma^2 I_n, \tag{4}$$

where the known $n \times (k + l)$ design matrix $(F,\ V) = (f_1\ f_2\ ...\ f_k\ v_1\ v_2\ ...\ v_l)$ is of full column rank $(k + l)$ and $y \in \mathbb{E}^l$ is an unknown realization of random vector Y. Then estimators $\widetilde{\nu}(X) = (\widetilde{\sigma}^2(X), \widetilde{\sigma}_1^2(X), \ldots, \widetilde{\sigma}_l^2(X))'$ of ν are said to be *observable natural estimators* or shortly *natural estimators*, if

$$\widetilde{\sigma}^2(X) = \tfrac{1}{n-k-l}[X - F\widetilde{\beta}(X) - V\widetilde{y}(X)]'[X - F\widetilde{\beta}(X) - V\widetilde{y}(X)], \tag{5}$$

$$\widetilde{\sigma}_j^2(X) = \widetilde{y}_j^2(X), j = 1, 2, ..., l, \tag{6}$$

where $(\widetilde{\beta}(X), \widetilde{y}(X))' = (\widetilde{\beta}_1(X), \ldots, \widetilde{\beta}_k(X), \widetilde{y}_1(X), \ldots, \widetilde{y}_l(X))'$ is the ordinary least square estimator of $(\beta, y)' \in \mathbb{E}^{k+l}$.

From projection theory we know that the ordinary least square estimator $(\widetilde{\beta}(X), \widetilde{y}(X))'$ of $(\beta, y)'$ for linear regression model (4) is given by the equation:

$$\begin{pmatrix} \widetilde{\beta}(X) \\ \widetilde{y}(X) \end{pmatrix} = \begin{pmatrix} F'F & F'V \\ V'F & V'V \end{pmatrix}^{-1} \begin{pmatrix} F'X \\ V'X \end{pmatrix}, \tag{7}$$

where the $(k+l) \times (k+l)$ Gram matrix $G = \left(\begin{smallmatrix} F'F & F'V \\ V'F & V'V \end{smallmatrix}\right)$ is of full rank, since $r(G) = r[(F,\, V)'(F,\, V)] = r(F,\, V) = k + l$.

If $(F,\, V)$ is of full column rank, then F must be too, so making use of the so-called Banachiewicz formula for the inverse of a partitioned (block) matrix[3], see e.g. Zhang (2005), we can write

$$G^{-1} = \begin{pmatrix} * & * \\ -W^{-1}(V'F)(F'F)^{-1} & W^{-1} \end{pmatrix}, \tag{8}$$

where symbol * denotes blocks not interesting in deriving $\widetilde{y}(X)$ and where $W = V'V - V'F(F'F)^{-1}F'V \in \mathbb{E}^{l \times l}$ is called the Schur complement of $F'F$ in G. Substituting (8) to (7) and rearranging, we finally get the following form of estimator $\widetilde{y}(X)$ of y

$$\widetilde{y}(X) = W^{-1}V'(I - F(F'F)^{-1}F')X. \tag{9}$$

2.2 Geometrical interpretation of natural estimators

Now we use the geometrical language of projection theory for describing definition and properties of natural estimators. Such intermediate step provides us a powerful tool in easier establishing and understanding new features of given concepts. It will also shorten proofs, which would be long and tedious if we did them by "direct computations".

The orthogonal projector $P_{\mathscr{L}^\perp(F)} \equiv M_F = I - F(F'F)^{-1}F'$ offers the following simplification of (9):

$$\widetilde{y}(X) = TX; T \equiv W^{-1}V'M_F \in \mathbb{E}^{l \times n}, \tag{10}$$

where the Schur complement $W = V'M_FV$. Since every orthogonal projector M_F is a symmetric matrix, $W \in \mathbb{E}^{l \times l}$ has to be also symmetric.

Our natural estimators of ν can be effectively expressed by means of projectors:

$\widetilde{\sigma}^2(X) = \frac{1}{n-k-l}\, \|MX\|^2$, where M is $P_{\mathscr{L}^\perp(F,\, V)}$.

$\widetilde{\sigma}_j^2(X) = \widetilde{y}_j^2(X); j = 1, 2, ..., l$, where $V\widetilde{y}(X) = P_{\mathscr{L}(V),\, \mathscr{L}^\perp(M_FV)}X$.

[3]There exist two forms of inverses for given block matrix, which are mathematically equivalent. We have chosen the form providing simpler algebraic expressions.

The first expression for $\widetilde{\sigma}^2(X)$ is the standard result of the unbiased variance estimation in classical linear regression models (Štulajter 2002). Concerning $\widetilde{\sigma}_j^2(X)$ if H denotes matrix VT, it is obvious from properties of V and def. (10) of T that $H^2 = H$, $\mathscr{L}(H) = \mathscr{L}(V)$, $\mathcal{N}(H) = \mathcal{N}(T) = \mathcal{N}(V'M_F) = \mathscr{L}^{\perp}(M_F V)$. Then projection theory says that every idempotent matrix H is the oblique projector $P_{\mathscr{L}(H),\mathcal{N}(H)}$, so we have $H = P_{\mathscr{L}(V),\mathscr{L}^{\perp}(M_F V)}$.

3 Statistical properties of natural estimators

3.1 First and second moment properties

As we said the expressions of natural estimators through projectors give us a powerful tool in understanding and elegant proving properties of matrices T and M, which determine statistical properties of the estimators. It is easy to show next lemma:

Lemma 3.1. (basic properties of T and M)

(i) $TT' = W^{-1}$ *and* $r(T) = l$

(ii) $TF = 0$, $TV = I_l$ *and* $(VT)^2 = VT$

(iii) $T\Sigma T' = \sigma^2 W^{-1} + diag(\sigma_j^2)$

(iv) $[M_F(VT)]' = M_F(VT)$

(v) $M^2 = M$, $M' = M$, $tr(M) = n - k - l$

(vi) $M\,(F \quad V) = 0$ *and* $TM = 0$

(vii) $M\Sigma = \sigma^2 M$

(viii) $M = M_F - M_F VT$

Proof. (i) Employing (10), $M_F^2 = M_F, M_F' = M_F$ and symmetry of matrix $W = V'M_F V$ we can write $TT' = W^{-1}V'M_F(W^{-1}V'M_F)' = W^{-1}$. Since $r(T) = r(TT')$, we conclude that $r(T) = r(W^{-1}) = l$.

(ii) According to the standard properties of oblique projectors (Galántai 2004) projector $P_{\mathscr{L}(V),\mathscr{L}^{\perp}(M_F V)} = VT$ immediately gives $VTv_j = v_j$ and $VTf_i = 0$, where v_j and f_i are columns of V and F. It implies $VTV = V, VTF = 0$. Hence first two properties of (ii) are results of multiplying the equalities on the left by a left inverse to the full column rank V. The last property is simply a restatement of projector idempotentness.

(iii) Applying (i), (ii) and the expression for Σ_ν from (2)

$$T\Sigma_\nu T' = \sigma^2 TT' + \sum_{j=1}^{l} \sigma_j^2 T v_j\, v_j' T' = \sigma^2 W^{-1} + \sum_{j=1}^{l} \sigma_j^2 e_j\, e_j',$$

where e_j denotes the j^{th} unit vector with unity for its j^{th} element and zeros elsewhere. Since the last term is only another form of $diag(\sigma_j^2)$, (iii) is valid.

(iv) This property can be reached by a direct routine computation.

In a very similar way using properties of orthogonal projectors we can prove the basic properties (v)-(viii) dealing with matrix M. ■

The natural estimators of ν can be written as quadratic forms

$$\widetilde{\sigma}^2(X) = \tfrac{1}{n-k-l}(MX)'(MX) = \tfrac{1}{n-k-l}X'MX,$$

$$\widetilde{\sigma}_j^2(X) = (t_j'X)^2 = X't_jt_j'X, j = 1,2,...,l, \text{ where } t_j' \text{ are rows of } T.$$

so results $MF = 0$ and $TF = 0$ from lemma 3.1 lead to conclusion that natural estimators are invariant quadratic estimators.[4]

In the following theorem we summarize mean and covariance characteristics of natural estimators.

Theorem 3.2. *Natural estimators of ν have the following properties:*

(i) $E_\nu[\widetilde{\sigma}^2(X)] = \sigma^2$ *and* $E_\nu[\widetilde{\sigma}_j^2(X)] = \sigma_j^2 + \sigma^2(W^{-1})_{jj}; j = 1,2,\ldots,l.$

If $X \sim N_n(F\beta,\Sigma)$*, then*

(ii) $D_\nu[\widetilde{\sigma}^2(X)] = \frac{2\sigma^4}{n-k-l}$ *and* $Cov_\nu[\widetilde{\sigma}^2(X),\widetilde{\sigma}_j^2(X)] = 0; j = 1,2,\ldots,l,$

(iii) $D_\nu[\widetilde{\sigma}_j^2(X)] = 2[\sigma_j^2 + \sigma^2(W^{-1})_{jj}]^2; j = 1,2,\ldots,l,$

(iv) $Cov_\nu[\widetilde{\sigma}_i^2(X),\widetilde{\sigma}_j^2(X)] = 2[\sigma^2(W^{-1})_{ij}]^2; i,j = 1,2,...,l, i \neq j.$

Proof. Since used arguments are very similar in proofs of all items we show only proofs of (i) and (iv). Applying previous results with the aid of well-known expressions for mean values and covariances of invariant quadratic estimators (see e.g. Christensen 2002) $E_\nu(X'AX) = tr(A\Sigma_\nu)$ and if $X \sim N(F\beta,\Sigma_\nu)$,then $Cov_\nu(X'AX,X'BX) = 2tr(A\Sigma_\nu B\Sigma_\nu)$, we find

(i)

$$\begin{aligned} E_\nu[\widetilde{\sigma}_j^2(X)] &= tr(t_jt_j'\Sigma_\nu) = tr(t_j'\Sigma_\nu t_j) = t_j'\Sigma_\nu t_j \\ &= (T\Sigma T')_{jj} = \sigma_j^2 + \sigma^2(W^{-1})_{jj} \\ E_\nu[\widetilde{\sigma}^2(X)] &= \tfrac{1}{n-k-l}tr(M\Sigma_\nu) = \tfrac{\sigma^2}{n-k-l}tr(M) = \sigma^2 \end{aligned}$$

(iv)

$$\begin{aligned} Cov_\nu[\widetilde{\sigma}_i^2(X),\widetilde{\sigma}_j^2(X)] &= 2tr(t_it_i'\Sigma_\nu t_jt_j'\Sigma_\nu) = 2tr(t_i'\Sigma_\nu t_jt_j'\Sigma_\nu t_i) \\ &= 2(T\Sigma_\nu T')_{ij}(T\Sigma_\nu T')_{ji} = 2(\sigma^2(W^{-1})_{ij})^2 \quad ■ \end{aligned}$$

[4] We recall that if X satisfies a linear regression model with $E_\beta(X) = F\beta$ as it is in case of the FDSLRM observation, then quadratic form $X'AX$ is called *invariant quadratic estimator* (with respect to β), if $AF = 0$, which means that such quadratic form does not depend on the mean value parameter β.

Obtained results show that unlike DOOLSE estimators used in orthogonal FDSLRMs, our natural estimators $\widetilde{\nu}(X) = (\widetilde{\sigma}^2(X), \widetilde{\sigma}_1^2(X), \ldots, \widetilde{\sigma}_l^2(X))'$ of ν with the exception $\widetilde{\sigma}^2(X)$ are *not consistent* and *biased* and a bias is determined by elements of the inverse of the Schur complement $W = V'V - V'F(F'F)^{-1}F'V = (V'M_FV) \in \mathbb{E}^{l\times l}$ of $F'F$ in Gram matrix $G = \left(\begin{smallmatrix} F'F & F'V \\ V'F & V'V \end{smallmatrix}\right)$ for linear regression model (4). The consistent and unbiased estimator is only estimator $\widetilde{\sigma}^2(X)$ of σ^2, which is also uncorrelated with remaining estimators.

Generally it is clear that if we have only data x, one realization of a finite length n of a time series $X(.)$ given by the FDSLRM, then for every $j = 1, 2, ..., l$ we have only one realization y_j of the random variable Y_j and it is impossible to find a consistent estimator of the variance $\sigma_j^2 = D_\nu[Y_j]$ based only on one value (estimate) of the random variable Y_j.

The same reason causes that in an orthogonal FDSLRM the DOOLSE $\hat{\nu}(X)$ used by Štulajter & Witkovský (2004) is also not a consistent estimator of ν. In general, in any FDSLRM there is no consistent estimator of the variances parameter ν.

3.2 Further asymptotic properties

Now we will study further asymptotic properties and find conditions of asymptotic unbiasedness. Let symbol X_n denote the finite observation X of time series $X(.)$, if that observation has size $n \times 1$. Then natural estimators of ν, matrices F, V, W and G also depend on n, so that we will use the more specific notation $\widetilde{\nu}(X) = \widetilde{\nu}(X_n), F = F_n, V = V_n, W = W_n, G = G_n$.

If we apply the concept of the order $O(1/n)$ of a real matrix sequence[5] to the sequence of inverses of the Schur complements W_n and combine it with the well-known fact for any matrix sequence $\{A_n\}$: $\lim\limits_{n\to\infty} A_n = 0 \in \mathbb{E}^{r\times s}$, if $A_n = O(1/n)$, then theorem 3.2 yields the following result showing a sufficiency for asymptotic unbiasedness of $\sigma_j^2(X_n)$ and corresponding asymptotic second-order properties in case of normality of observation X_n.

Theorem 3.3. *Let us consider the nonorthogonal FDSLRM*

$$X_n = F_n\beta + \varepsilon_n,\ E(\varepsilon_n) = 0,\ Cov_\nu(\varepsilon_n) = \Sigma_n = \sigma^2 I + \sum_{j=1}^{l} \sigma_j^2 v_{n,j} v_{n,j}',$$

where $v_{n,j}; j = 1, 2, \ldots, l$ are columns of V_n and $\begin{pmatrix} F_n & V_n \end{pmatrix} \in \mathbb{E}^{n\times(k+l)}$ are of full rank. Let $W_n^{-1} = O(1/n)$, where $W_n \in \mathbb{E}^{l\times l}$ are Schur complements of $(F_n'F_n)$ in the $(k+l)\times(k+l)$ partitioned Gram matrices $G_n = \begin{pmatrix} F_n'F_n & F_n'V_n \\ V_n'F_n & V_n'V_n \end{pmatrix}$. Then natural estimators $\widetilde{\sigma}_j^2(X_n)$ of variances $\sigma_j^2; j = 1, 2, \ldots, l$ are:

[5] A sequence $\{A_n\}$ of $r \times s$ matrices is said to be of the order $O(1/n)$, if for any fixed pair of i and j $(i = 1, .., r; j = 1, ..., s)$ a real sequence $\{|(A_n)_{ij}/(1/n)|\}$ formed by matrix elements $(A_n)_{ij}$ is bounded. In such case we write $A_n = O(1/n)$.

(i) *asymptotically unbiased, i.e.* $\lim_{n\to\infty} E[\widetilde{\sigma}_j^2(X_n)] = \sigma_j^2$.

If $X_n \sim N_n(F_n\beta, \Sigma_n)$, *then the estimators are also*

(ii) *mutually asymptotically uncorrelated, i.e.* $\lim_{n\to\infty} Cov[\widetilde{\sigma}_i^2(X_n), \widetilde{\sigma}_j^2(X_n)] = 0$ *for* $i \neq j; i,j = 1,2,\ldots,l$

(iii) *with asymptotic dispersions* $2\sigma_j^4$, *i.e.* $\lim_{n\to\infty} D[\widetilde{\sigma}_j^2(X_n)] = 2\sigma_j^4$.

4 Illustrations

In the following three examples we illustrate theoretical results obtained in previous sections. Our concern is primarily to show different forms of Schur complements which play the central role in establishing properties of natural estimators.[6]

Example 1. Let $X(.)$ be a time series given by the model

$$X(t) = \beta_1 + Y_1 t + w(t); t = 1,2,...$$

It means that the FDSLRM has the mean value as an unknown constant and the errors are given by a random linear trend plus a white noise term.

The corresponding model of FDSLRM observation (2) has the form $X = F\beta + \varepsilon; E[\varepsilon] = 0, Cov_\nu(X) = \sigma^2 I + \sigma_1^2 v_1 v_1'$, where $F = (1,1,...,1)' \equiv j_n, V = v_1 = (1,2,...,n)'$.

Then we get $M_F = I - F(F'F)^{-1}F = I_n - \frac{1}{n}J_n = C_n$, where $J_n = j_n j_n'$ is a matrix whose every element is unity and C_n is the well-known *centering matrix* having these elementary properties: $C_n' = C_n, C_n^2 = C_n, C_n J_n = J_n C_n = 0, C_n x = x - \bar{x}j_n, x'C_n y = x'y - n\bar{x}\bar{y}; x,y \in E^n$.

After that a routine computation with the aid of the properties of C_n leads to the following results

$$\begin{aligned} &W_n = V'M_F V = v_1'v_1 - n\bar{v}_1^2, \\ &T = t_1' = W_n^{-1}V'M_F = W_n^{-1}(v_1' - \bar{v}_1 j_n'), \\ &M = M_F - M_F V T = C_n - W_n^{-1}(v_1 - \bar{v}_1 j_n)(v_1' - \bar{v}_1 j_n'), \end{aligned}$$

where $\bar{v}_1 = 1/n \sum_{j=1}^{n} (v_1)_j = 1/n \sum_{t=1}^{n} t$ and $v_1'v_1 = \|v_1\|^2 = \sum_{t=1}^{n} t^2$.

Evaluating Schur complements W_n after substitution of $\bar{v}_1 = n(n+1)/2$ and v_1, we get $W_n = (n^3 - n)/12$, so the inverses $W_n^{-1} = 12/(n^3 - n) = O(1/n)$.

[6] We also added some numerical results for particular values of model parameters, since evaluating with aid of computer we find them very quickly and effectively. We have used advanced mathematics and computer algebra software package Maple.

According to theorem 3.3 $\widetilde{\sigma}_1^2(X)$ is an asymptotically unbiased estimator with asymptotic covariance $2\sigma_1^4$. The $\widetilde{\sigma}_1^2(X)$ is given by

$$\widetilde{\sigma}_1^2(X) = \left(\tfrac{12}{n^3-n}\right)^2 \sum_{s=1}^{n}\sum_{t=1}^{n} \left(s - \tfrac{n+1}{2}\right)\left(t - \tfrac{n+1}{2}\right) X(s)X(t)$$

and is uncorrelated with the consistent unbiased estimator $\widetilde{\sigma}^2(X)$ with $D_\nu[\widetilde{\sigma}^2(X)] = \frac{2\sigma^4}{n-2}$, whose explicit form after calculating element of M is

$$\widetilde{\sigma}^2(X) = \tfrac{1}{n-2}\left(\sum_{t=1}^{n} X(t)^2 - \sum_{s=1}^{n}\sum_{t=1}^{n}\left[\tfrac{1}{n} + \tfrac{12}{n^3-n}\left(s - \tfrac{n+1}{2}\right)\left(t - \tfrac{n+1}{2}\right)\right] X(s)X(t)\right).$$

Example 2. Let $X(.)$ be a time series given by the model

$$X(t) = \beta_1 + \beta_2 t + Y_1 \cos\lambda t + Y_2 \sin\lambda t + w(t); t = 1, 2, ..., n$$

where $\lambda \in \langle 0, \pi\rangle$ is some non-fourier frequency and Y_1, Y_2 are uncorrelated random variables with zero mean values and variances $\sigma_j^2 = D[Y_j]; j = 1, 2$. Since in this case we have

$$F = (f_1\ f_2) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 2 & \dots & n \end{pmatrix}',$$
$$V = (v_1\ v_2) = \begin{pmatrix} \cos\lambda & \cos 2\lambda & \dots & \cos n\lambda \\ \sin\lambda & \sin 2\lambda & \dots & \sin n\lambda \end{pmatrix}',$$

the orthogonal projection matrix M_F onto $\mathscr{L}^\perp(F)$ is identical with orthogonal projection matrix M onto $\mathscr{L}^\perp(F, V)$ in the previous example. The only difference are other symbols. The role of column v_1 takes column f_2, so

$$M_F = C_n - (f_2' f_2 - n\bar{f}_2^2)(f_2 - \bar{f}_2 j_n)(f_2' - \bar{f}_2 j_n').$$

This expression yields the Schur complement with components

$$\begin{aligned}(W_n)_{ij} &= v_i' C_n v_j - (f_2' f_2 - n\bar{f}_2^2)\, v_i'(f_2 - \bar{f}_2 j_n)(f_2' - \bar{f}_2 j_n')v_j \\ &= v_i' v_j - n\bar{v}_i\bar{v}_j - (f_2' f_2 - n\bar{f}_2^2)\,(v_i' f_2 - n\bar{f}_2\bar{v}_i)(v_j' f_2 - n\bar{f}_2\bar{v}_j);\ i, j = 1, 2.\end{aligned}$$

Substituting columns of F and V for our model to these components we get

$$(W_n)_{11} = \sum_{t=1}^{n} \cos^2\lambda t - \tfrac{1}{n}\left(\sum_{t=1}^{n}\cos\lambda t\right)^2 - \tfrac{12}{n^3-n}\left[\sum_{t=1}^{n}\left(t - \tfrac{n+1}{2}\right)\cos\lambda t\right]^2,$$

$$(W_n)_{22} = \sum_{t=1}^{n} \sin^2\lambda t - \tfrac{1}{n}\left(\sum_{t=1}^{n}\sin\lambda t\right)^2 - \tfrac{12}{n^3-n}\left[\sum_{t=1}^{n}\left(t - \tfrac{n+1}{2}\right)\sin\lambda t\right]^2,$$

$$(W_n)_{12} = \sum_{t=1}^{n} \cos\lambda t \sin\lambda t - \tfrac{1}{n}\sum_{t=1}^{n}\cos\lambda t \ \sum_{t=1}^{n}\sin\lambda t -$$
$$- \tfrac{12}{n^3-n}\sum_{t=1}^{n}\left(t-\tfrac{n+1}{2}\right)\cos\lambda t \ \cdot \sum_{t=1}^{n}\left(t-\tfrac{n+1}{2}\right)\sin\lambda t.$$

Now we demonstrate that 2×2 inverse W_n^{-1} given by the well-known expression

$$W_n^{-1} = \frac{1}{D}\begin{pmatrix} (W_n)_{22} & -(W_n)_{12} \\ -(W_n)_{12} & (W_n)_{11} \end{pmatrix}; D = (W_n)_{11}(W_n)_{22} - (W_n)_{12}^2,$$

is of the order $O(1/n)$, so according to theorem 3.3 this property identifies our natural estimators $\widetilde{\sigma}_j^2(X); j = 1,2$ as asymptotically unbiased with asymptotic dispersions $2\sigma_j^2$. They also become mutually asymptotically uncorrelated.

By boundedness of trigonometric functions and the well-known trigonometric identities we can derive for W_n elements that $(W_n)_{11} = O(n), (W_n)_{22} = O(n), (W_n)_{12} = O(1), D_n = O(n^2)$, so

$$W_n^{-1} = \begin{pmatrix} O(1/n) & O(1/n^2) \\ O(1/n^2) & O(1/n) \end{pmatrix},$$

or in other words matrix sequence $\{W_n^{-1}\}$ is of the order $O(1/n)$.

We illustrate mentioned properties for given numerical value of parameter $\lambda = 0.32\pi$ and $n = 7, 15, 101$.

For $n = 7$

$$W_7^{-1} = \begin{pmatrix} 0.439267 & -0.146941 \\ -0.146941 & 0.383206 \end{pmatrix}.$$

It means that $E[\tilde{\sigma}_1^2(X)] = \sigma_1^2 + 0.439\sigma^2$, $D[\tilde{\sigma}_1^2(X)] = 2[\sigma_1^2 + 0.439\sigma^2]^2$, $E[\tilde{\sigma}_2^2(X)] = \sigma_2^2 + 0.383\sigma^2$, $D[\tilde{\sigma}_2^2(X)] = 2[\sigma_2^2 + 0.383\sigma^2]^2$, $Cov[\tilde{\sigma}_1^2(X); \tilde{\sigma}_2^2(X)] = 2\sigma^4(0.147)^2$.

We can determine these characteristics by elements W_n^{-1} in remaining cases by the same way

for $n = 15$

$$W_{15}^{-1} = \begin{pmatrix} 0.140156 & 0.001597 \\ 0.001597 & 0.132087 \end{pmatrix},$$

for $n = 101$

$$W_{101}^{-1} = \begin{pmatrix} 0.019916 & -0.0001932 \\ -0.0001932 & 0.0197342 \end{pmatrix},$$

where there is apparent that elements of W_n^{-1} tend to zero as n increases which is in accordance with the result of our proof.

Example 3. Let $X(.)$ be a time series given by the model

$$X(t) = \beta_1 + \beta_2 \ln t + Y_1 \exp(-\gamma_1 t) + Y_2 \exp(-\gamma_2 t) + w(t); t = 1, 2, ..., n$$

where $\gamma_1, \gamma_2 \in (0, \infty)$. Matrices F and V are

$$F = (f_1\ f_2) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \ln 1 & \ln 2 & \dots & \ln n \end{pmatrix}',$$
$$V = (v_1\ v_2) = \begin{pmatrix} \exp(-\gamma_1) & \exp(-2\gamma_1) & \dots & \exp(-n\gamma_1) \\ \exp(-\gamma_2) & \exp(-2\gamma_2) & \dots & \exp(-n\gamma_2) \end{pmatrix}',$$

hence for a full rank FDSLRM observation model the condition $\gamma_1 \neq \gamma_2$ must be satisfied.

Applying the same argument and way of calculations as in example 2 (the form of M_F is analogical) we conclude that

$$(W_n)_{jj} = \sum_{t=1}^{n} \exp(-2\gamma_j t) - \tfrac{1}{n}\left(\sum_{t=1}^{n} \exp(-\gamma_j t)\right)^2 - U_n^{-1}\left[\sum_{t=1}^{n}\left(\ln t - \tfrac{\ln n!}{n}\right)\exp(-\gamma_j t)\right]^2 ; j = 1, 2,$$

$$(W_n)_{12} = \sum_{t=1}^{n} \exp[-(\gamma_1+\gamma_2)t] - \tfrac{1}{n}\sum_{t=1}^{n}\exp(-\gamma_1 t)\sum_{t=1}^{n}\exp(-\gamma_2 t) - U_n^{-1}\sum_{t=1}^{n}\left(\ln t - \tfrac{\ln n!}{n}\right)\exp(-\gamma_1 t)\cdot\sum_{t=1}^{n}\left(\ln t - \tfrac{\ln n!}{n}\right)\exp(-\gamma_2 t),$$

where $U_n \equiv f_2' f_2 - n\bar{f}_2^2 = \sum_{t=1}^{n} \ln^2 t - \frac{\ln^2 n!}{n}$. By a routine computation, making use of concept of the order, elementary formulas for sums of geometrical sequences and Stirling's formula $n! = (n/e)^n\sqrt{2\pi n}\exp(\Theta_n/12n)$; $0 < \Theta_n < 1$, we can show that for any $\gamma_1, \gamma_2 \in (0, \infty), \gamma_1 \neq \gamma_2$

$$\lim_{n\to\infty} W_n^{-1} = \frac{1}{D}\begin{pmatrix} (\exp 2\gamma_2 - 1)^{-1} & -(\exp(\gamma_1+\gamma_2) - 1)^{-1} \\ -(\exp(\gamma_1+\gamma_2) - 1)^{-1} & (\exp 2\gamma_1 - 1)^{-1} \end{pmatrix},$$

where $1/D = (\exp 2\gamma_1 - 1)(\exp 2\gamma_2 - 1)(\exp(\gamma_1+\gamma_2) - 1)^2/(\exp\gamma_1 - \exp\gamma_2)^2$.

On the basis of that result we observe that W_n^{-1} is of the order $O(1)$ and natural estimators cannot be asymptotically unbiased estimators.

As a numerical illustration let us consider the FDSLRM in case $\gamma_1 = 2, \gamma_2 = 5$. If $n = 20, 100, 500, 1000$, then we observe that

$$W_{20}^{-1} = \begin{pmatrix} 5704.07 & -106560 \\ -106560 & 2.0281 \times 10^6 \end{pmatrix},$$

$$W_{100}^{-1} = \begin{pmatrix} 3891.25 & -76361.8 \\ -76361.8 & 1.5250 \times 10^6 \end{pmatrix},$$

$$W_{500}^{-1} = \begin{pmatrix} 3446.42 & -68720.4 \\ -68720.4 & 1.3937 \times 10^6 \end{pmatrix},$$

$$W_{1000}^{-1} = \begin{pmatrix} 3365.29 & -67311.7 \\ -67311.7 & 1.3692 \times 10^6 \end{pmatrix}$$

are really approaching values obtained from the derived exact limit expression

$$\lim_{n\to\infty} W_n^{-1} = \begin{pmatrix} 3235.14 & -65035.9 \\ -65035.9 & 1.3329 \times 10^6 \end{pmatrix}.$$

Acknowledgement. The author is especially indebted to František Štulajter for useful discussions and helpful suggestions. The presented work is supported by the grant 1/3001/06 of the Slovak Scientific Grant Agency (VEGA).

References

Ben-Israel, A. and Greville, T.N.E. (2003) *Generalized Inverses: Theory and Applications.* 2nd ed., Springer–Verlag, New York.

Box, G.E.P., Jenkins, G.M. and Reinsel G.C. (1994) *Time Series Analysis, Forecasting and Control.* 3rd ed., Prentice Hall, New Jersey.

Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods.* 2nd ed., Springer–Verlag, New York.

Christensen, R. (1991) *Linear Models for Multivariate, Time Series and Spatial Data.* Springer–Verlag, New York.

Christensen, R. (2002) *Plane Answers to Complex Questions: The Theory of Linear Models.* 3rd ed., Springer–Verlag, New York.

Galántai, A. (2003) *Projectors and Projection Methods.* Kluwer Academic Publishers, Boston, 2004.

McCulloch, Ch. E. and Searle, S.R. (2001) *Generalized, Linear, and Mixed Models.* Wiley–Interscience, New York.

Rao, C.R. and Kleffe, J. (1988) *Estimation of Variance Components and Applications.* North–Holland, Amsterdam.

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

Štulajter, F. (2002) *Predictions in Time Series Using Regression Models.* Springer–Verlag, New York.

Štulajter, F. (2003) The MSE of the BLUP in a finite discrete spectrum LRM. *Tatra Mt. Math. Publ.* **26**: 125–131.

Štulajter, F. and Witkovský, V. (2004) Estimation of variances in orthogonal finite discrete spectrum linear regression models. *Metrika* **60**: 105–118.

Štulajter, F. (2006) Mean squared error of the empirical best linear unbiased predictor in an orthogonal finite discrete spectrum linear regression model, *Metrika*, 23 August 2006, published online <http://www.springerlink.com/content/1435-926X/>.

Volaufová, J. and Witkovský, V. (1991) Least squares and minimum MSE estimators of variance components in mixed linear models. *Biometr. J.* **33**: 923–936.

Zhang, F. et al (2005) *The Schur Complement and Its Applications.* Springer Science+Business Media, New York.