amcs

ITERATIVE ESTIMATORS OF PARAMETERS IN LINEAR MODELS WITH PARTIALLY VARIANT COEFFICIENTS

SHAOLIN HU [*,**],   KARL MEINKE [**],   RUSHAN CHEN [*],   OUYANG HUAJIANG [***]

[*] Nanjing University of Science and Technology
210071, Nanjing, China
e-mail: hshaolin@ustc.edu.cn

[**] Department of Computer Science, Royal Institute of Technology
Stockholm, 100-44, Sweden
e-mail:karlm@csc.kth.se

[***] University of Liverpool, Liverpool, United Kingdom
L69 3 GH, Liverpool, UK
e-mail: h.ouyang@liverpool.ac.uk

A new kind of linear model with partially variant coefficients is proposed and a series of iterative algorithms are introduced and verified. The new generalized linear model includes the ordinary linear regression model as a special case. The iterative algorithms efficiently overcome some difficulties in computation with multidimensional inputs and incessantly appending parameters. An important application is described at the end of this article, which shows that this new model is reasonable and applicable in practical fields.

Keywords: linear model, parameter estimation, iterative algorithms, variant coefficients

1. Introduction

In the last centuries, many statisticians and mathematicians have considered the following kind of linear regression model:

$$\boldsymbol{Y}_i = \boldsymbol{B}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i, \tag{1}$$

where $\boldsymbol{Y}_i \in \mathbb{R}^p, \boldsymbol{B}_i \in \mathbb{R}^{p\times r}$ $(i = 1, 2, \ldots, n)$ and the vector $\boldsymbol{\beta} \in \mathbb{R}^r$ is a constant parameter vector to be estimated while $\boldsymbol{\varepsilon}_i \in \mathbb{R}^p$ are errors from measurements or stochastic noise from disturbances.

Some excellent theories and practical results were published for statistical inference and stochastic decisions under this model. This model was also successfully used in many different kinds of practical fields (see Draper and Smith, 1981; Frank and Harrell, 2002; Graybill and Iyer, 1994; Hu and Sun, 2001).

Further research into this model shows that the limitation of constant coefficients in (1) is very rigorous and critical. In other words, there are some practical situations in which this linear model cannot be applied (see Brown, 1964; Hu and Sun, 2001). Although statisticians

and researchers (see Fahrmeier and Tutz, 2001; Dodge and Kova, 2000) have done a lot to generalize and/or to adapt the linear model (2), the limitation on constant coefficients has not been essentially untied.

In order to overcome this limitation on the model structure, we set up a new linear model with partially variant coefficients as follows:

$$\boldsymbol{Y}_i = \boldsymbol{A}_i\boldsymbol{X}_i + \boldsymbol{B}_i\boldsymbol{f}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i, \tag{2}$$

where $\boldsymbol{Y}_i \in \mathbb{R}^p, \boldsymbol{A}_i \in \mathbb{R}^{p\times q}, \boldsymbol{B}_i \in \mathbb{R}^{p\times r}$, and a variant vector series $\{\boldsymbol{X}_i \in \mathbb{R}^q\}$. Generally, the dimension p of the measurement output must be greater than the dimension q of the variant coefficient, i.e., $p > q$, so as to assure that the structure of the time-variant multidimensional linear system is identifiable.

Obviously, the ordinary linear regression model (1) is just a special case of the generalized model (2). If there are not any time-variant parts, namely, $q = 0$, the model (2) degenerates to the ordinary linear model (1).

In Section 2, much attention is paid to estimate all the model coefficients in (2) under the Gauss-Markov as-

sumptions (in Kala R. and Kłaczyński K., 1988). A series of iterative algorithms are built to estimate the coefficients which include the constant parameters $\boldsymbol{\beta} \in \mathbb{R}^r$ and the variant vector series $\{\boldsymbol{X}_i \in \mathbb{R}^q\}$. In Section 3, a practical application is described and computation results show that this new model is valuable.

2. Iterative Estimators of Variant Coefficients

In order to make the results of this section universal, we assume first that the coefficient series $\{\boldsymbol{X}_i \in \mathbb{R}^q\}$ are not related at different sampling points. For simplicity, we adapt the notation of stacked vectors $\boldsymbol{\Phi}_n = (\boldsymbol{\beta}^\tau, \boldsymbol{X}_1^\tau, \ldots, \boldsymbol{X}_n^\tau)^\tau \in \mathbb{R}^{r+nq}$, where the superscript τ denotes the matrix transpose. Moreover,

$$
\boldsymbol{H}_n = \begin{pmatrix} \boldsymbol{B}_1 & \boldsymbol{A}_1 & \ldots & \boldsymbol{0} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{B}_n & \boldsymbol{0} & \cdots & \boldsymbol{A}_n \end{pmatrix} \in \mathbb{R}^{np \times (r+nq)}. \quad (3)
$$

Under the following famous Gauss-Markov assumptions (Kala R. and Kłaczyński K., 1988; Rencher, 2000) on random errors $\{\boldsymbol{\varepsilon}_i \in \mathbb{R}^p\}$:

(i) the error $\boldsymbol{\varepsilon}_i$ has expected value 0;

(ii) the errors series $\{\boldsymbol{\varepsilon}_i, i = 1, 2, \ldots\}$ are uncorrelated, and

(iii) the error series $\{\boldsymbol{\varepsilon}_i, i = 1, 2, \ldots\}$ are homoscedastic, i.e.,

they all have the same variance, the least-squares (LS) estimators of coefficients in the model (2) can be expressed as follows:

$$
\begin{aligned}
&\hat{\boldsymbol{\Phi}}_n^{LS(n)} \\
&= \arg \min_{\{\boldsymbol{\beta} \in \mathbb{R}^p, \boldsymbol{X}_i \in \mathbb{R}^q\}} \sum_{i=1}^{n} \|\boldsymbol{Y}_i - (\boldsymbol{A}_i \boldsymbol{X}_i + \boldsymbol{B}_i \boldsymbol{\beta})\|^2. \quad (4)
\end{aligned}
$$

We can directly deduce a compact formula, which is very similar to the LS estimators of the model (1). The compact formula of the LS estimator of the coefficients in the model (2) is as follows:

$$
\hat{\boldsymbol{\Phi}}_n^{LS(n)} = (\boldsymbol{H}_n^\tau \boldsymbol{H}_n)^{-1} \boldsymbol{H}_n^\tau \bar{Y}_n, \quad (5)
$$

where $\bar{Y}_n = (\boldsymbol{Y}_1^\tau, \ldots, \boldsymbol{Y}_n^\tau)^\tau \in \mathbb{R}^{np}$.

In order to guarantee that the matrix $\boldsymbol{H}_n^\tau \boldsymbol{H}_n$ is invertible, the dimensions of the model (2) must satisfy the requirement $p > q$ and the sample cardinality must satisfy $n > r/(p-q)$. Otherwise, the matrix inversion in (5) must be replaced by the pseudo-inverse operator "+", namely, $(\boldsymbol{H}_n^\tau \boldsymbol{H}_n)^+$.

Theorem 1. *Assume that the number of sampling points is $n > r/(p-q)$, where the LS estimators of the coefficients in the model (2) can be iteratively expressed by*

$$
\begin{cases}
\hat{\boldsymbol{\beta}}^{LS(n+1)} = \hat{\boldsymbol{\beta}}^{LS(n)} + (\boldsymbol{L}_n + \boldsymbol{B}_{n+1}^\tau \boldsymbol{B}_{n+1})^{-1} \\
\quad \times \boldsymbol{B}_{n+1}^\tau (\boldsymbol{R}_{n+1}^{-1} - \boldsymbol{\Xi}_{n+1}) \\
\quad \times \boldsymbol{R}_{n+1}(\boldsymbol{Y}_{n+1} - \boldsymbol{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}), \\
\hat{\boldsymbol{X}}_i^{LS(n+1)} = \hat{\boldsymbol{X}}_i^{LS(n)} + (\boldsymbol{A}_i^\tau \boldsymbol{A}_i)^{-1} \boldsymbol{A}_i^\tau \boldsymbol{B}_i \\
\quad \times (\hat{\boldsymbol{\beta}}^{LS(n)} - \hat{\boldsymbol{\beta}}^{LS(n+1)}), \\
\hat{\boldsymbol{X}}_{n+1}^{LS(n+1)} = (\boldsymbol{A}_{n+1}^\tau \boldsymbol{R}_{n+1} \boldsymbol{A}_{n+1})^{-1} \boldsymbol{A}_{n+1}^\tau \boldsymbol{R}_{n+1} \\
\quad \times (\boldsymbol{Y}_{n+1} - \boldsymbol{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}),
\end{cases} \quad (6)
$$

where

$$
\boldsymbol{L}_n = \sum_{i=1}^{n} \boldsymbol{B}_i^\tau [\boldsymbol{I} - \boldsymbol{A}_i(\boldsymbol{A}_i^\tau \boldsymbol{A}_i)^{-1} \boldsymbol{A}_i^\tau] \boldsymbol{B}_i \in \mathbb{R}^{r \times r},
$$

$$
\boldsymbol{\Xi}_{n+1} = \boldsymbol{A}_{n+1}(\boldsymbol{A}_{n+1}^\tau \boldsymbol{R}_{n+1} \boldsymbol{A}_{n+1})^{-1} \boldsymbol{A}_{n+1}^\tau \in \mathbb{R}^{p \times p},
$$
$$
\boldsymbol{R}_{n+1} = \boldsymbol{I} - \boldsymbol{B}_{n+1}(\boldsymbol{L}_n + \boldsymbol{B}_{n+1}^\tau \boldsymbol{B}_{n+1})^{-1} \boldsymbol{B}_{n+1}^\tau \in \mathbb{R}^{p \times p}.
$$

Here the superscript in $\hat{\boldsymbol{X}}_i^{LS(n)}$ denotes the LS estimators of $\boldsymbol{X}_i$, $i = 1, 2, \ldots, n$.

The proof of Theorem 1 is given in Appendix.

Obviously, the algorithm (6) is iterative and very common in practical engineering fields. Some obvious advantages are the following:

- $\hat{\boldsymbol{\beta}}^{LS(n+1)}$ is a linear combination of the estimator $\hat{\boldsymbol{\beta}}^{LS(n)}$ with innovation $\boldsymbol{Y}_{n+1} - \boldsymbol{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}$ from new sampling data.

- $\hat{\boldsymbol{\beta}}^{LS(n+1)}$ can be computed from the estimator $\hat{\boldsymbol{\beta}}^{LS(n)}$ without directly involving old sampling data $\{\boldsymbol{Y}_i, i = 1, \ldots, n\}$ as well as the estimates $\{\hat{\boldsymbol{X}}_i^{LS(n)}\}$.

- It is reasonable that the estimators $\hat{\boldsymbol{X}}_{n+1}^{LS(n+1)}$ are determined with innovation $\{\boldsymbol{Y}_{n+1} - \boldsymbol{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}\}$.

- The estimate $\hat{\boldsymbol{X}}_i^{LS(n+1)}$ of $\boldsymbol{X}_i$ $(i \leq n)$ can be adjusted accurately in succession with the estimator error of the constant parameter vector $\boldsymbol{\beta}$.

In order to use these iterative algorithms to effectively solve practical problems, the initial estimates of the iterative algorithms (6) must be carefully selected. Generally, the initial estimators can be selected as the LS estimators processed in batch as follows:

$$
\hat{\boldsymbol{\Phi}}_{n_0}^{LS(n_0)} = (\boldsymbol{H}_{n_0}^\tau \boldsymbol{H}_{n_0})^{-1} \boldsymbol{H}_{n_0}^\tau \bar{Y}_{n_0}, \quad (7)
$$

where $n_0 \in N$ must satisfy the constraint

$$
n_0 > \frac{r}{(p-q)}.
$$

If the disturbance $\{\varepsilon_n \in \mathbb{R}^p, n \leq n_0\}$ is a stationary Gaussian white noise process with zero mean, then it can be easily shown that the ordinary LS estimators given by (4) are unbiased.

Theorem 2. *Given the LS estimators (7) as an initial estimate of the coefficients of (2), if the disturbance $\{\varepsilon_n \in \mathbb{R}^p, n \in N\}$ is a stationary Gaussian white noise process with zero mean, then the iterative estimators (6) are unbiased.*

Proof. We just have to show that $E\hat{\boldsymbol{\beta}}^{LS(n)} = \boldsymbol{\beta}$. We get

$$
\begin{aligned}
E\{\hat{\boldsymbol{\beta}}^{LS(n)}\} &= E\{\hat{\boldsymbol{\beta}}^{LS(n-1)}\} + \tilde{\boldsymbol{L}}_n^{-1}\boldsymbol{B}_n^\tau[\boldsymbol{R}_n^{-1} - \boldsymbol{\Xi}_n]\boldsymbol{R}_n \\
&\quad \times \left(\boldsymbol{A}_n\boldsymbol{X}_n + \boldsymbol{B}_n(\boldsymbol{\beta} - E\{\hat{\boldsymbol{\beta}}^{LS(n-1)}\})\right) \\
&= \boldsymbol{\beta} + \tilde{\boldsymbol{L}}_n^{-1}\boldsymbol{B}_n^\tau\Big[\boldsymbol{I} - \boldsymbol{A}_n(\boldsymbol{A}_n^\tau \boldsymbol{R}_n \boldsymbol{A}_n)^{-1} \\
&\quad \times \boldsymbol{A}_n^\tau \boldsymbol{R}_n\Big]\boldsymbol{A}_n\boldsymbol{X}_n = \boldsymbol{\beta},
\end{aligned} \tag{8}
$$

where

$$
\tilde{\boldsymbol{L}}_n = \sum_{i=1}^{n-1} \boldsymbol{B}_i^\tau\Big[\boldsymbol{I} - \boldsymbol{A}_i(\boldsymbol{A}_i^\tau \boldsymbol{A}_i^+)\boldsymbol{A}_i^\tau\Big]\boldsymbol{B}_i + \boldsymbol{B}_n^\tau\boldsymbol{B}_n.
$$

Using (6), we get

$$
\begin{aligned}
E\{\hat{\boldsymbol{X}}_i^{LS(n)}\} &= E\{\hat{\boldsymbol{X}}_i^{LS(n-1)}\} + (\boldsymbol{A}_i^\tau \boldsymbol{A}_i)^{-1}\boldsymbol{A}_i^\tau \boldsymbol{B}_i \\
&\quad \times \left(E\{\hat{\boldsymbol{\beta}}^{LS(n-1)}\} - E\{\hat{\boldsymbol{\beta}}^{LS(n)}\}\right) \\
&= \boldsymbol{X}_i + (\boldsymbol{A}_i^\tau \boldsymbol{A}_i)^{-1}\boldsymbol{A}_i^\tau \boldsymbol{B}_i(\boldsymbol{\beta} - \boldsymbol{\beta}) = \boldsymbol{X}_i, \\
&\quad i = 1, 2, \ldots, n-1,
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
E\{\hat{\boldsymbol{X}}_n^{LS(n)}\} &= (\boldsymbol{A}_n^\tau \boldsymbol{R}_n \boldsymbol{A}_n)^{-1} \boldsymbol{A}_n^\tau \boldsymbol{R}_n \\
&\quad \times \left(E\{\boldsymbol{Y}_n\} - \boldsymbol{B}_n E\{\hat{\boldsymbol{\beta}}^{LS(n-1)}\}\right) \\
&= (\boldsymbol{A}_n^\tau \boldsymbol{R}_n \boldsymbol{A}_n)^{-1} \boldsymbol{A}_n^\tau \boldsymbol{R}_n \boldsymbol{A}_n \boldsymbol{X}_n = \boldsymbol{X}_n.
\end{aligned} \tag{10}
$$

Mathematical induction is implicitly used in the above proof, which justifies its correctness. ∎

3. Applications

The new linear model (2) can be widely used in many different fields, e.g., in data fusion, in modeling and identifying a computer controlled system, in signal processing, in spacecraft control engineering, etc. In this section, we present an application of the model (2) to the trajectory of a rocket.

Suppose that there are m transducers which are suitably located at different sites. These m devices are simultaneously used to track a carrying rocket M in space. Using these transducers, we get a series of measurement data $\{(A_j(t_i), E_j(t_i)) \mid i = 1, 2, \ldots, n; j = 1, 2, \ldots, m\}$, where $A_j(t_i)$ denotes the azimuth and $E_j(t_i)$ denotes the elevation of the rocket M at time t_i with respect to a reference frame fixed at the center of Transducer j.

In order to simplify the expressions below, we use a simplified notation such as $A_{ij} = A_j(t_i)$ and $E_{ij} = E_j(t_i)$, etc. Accordingly, the error decomposition models used in determining the location of the spacecraft M can be set up as follows (Brown, 1964; Hu and Sun, 2001):

$$
\begin{cases}
A_{ij} = \tan^{-1}\left(\dfrac{x - x_{0j}}{y - y_{0j}}\right) + \alpha_{j1} + \alpha_{j3}\tan(E_{ij})\sin(A_{ij}) \\
\qquad + \alpha_{j4}\tan(E_{ij})\cos(A_{ij}) + \alpha_{j5}\tan(E_{ij}) \\
\qquad + \alpha_{j6}\sec(E_{ij}) + \varepsilon_{A_{ij}}, \\
E_{ij} = \tan^{-1}\left(\dfrac{z - z_{0j}}{[(x - x_{0j})^2 + (y - y_{0j})^2]^{1/2}}\right) + \alpha_{j2} \\
\qquad + \alpha_{j3}\cos(A_{ij}) - \alpha_{j4}\sin(A_{ij}) + \varepsilon_{E_{ij}},
\end{cases} \tag{11}
$$

where the coefficients $(\alpha_{j1}, \alpha_{j2})$ are non-zero errors of Transducer j used to measure the azimuth and elevation of the spacecraft, the vectors $(\alpha_{j3}, \ldots, \alpha_{j6})$ are non-orthogonal coefficients representing measurement errors arising from departures from right angles between each pair of axes in measurement equipment (the mechanical axis, the laser axis and the electrical axis) separately, $(\varepsilon_A, \varepsilon_E)$ are stochastic errors included in the measurement data.

Assuming that we get a series of imprecise location data

$$
\vec{p}_i^* = (x_i^*, y_i^*, z_i^*)
$$

for the spacecraft M at different sampling times t_i $(i = 1, 2, \ldots)$, what we want to do is to estimate all of the instrument error coefficients as well as the precise location of the spacecraft M.

According to the geometrical relationship between the ordinates and measurement data from radars, two functions are defined as follows:

$$
f_j(x, y, z) = \tan^{-1}\left(\frac{x - x_{0j}}{y - y_{0j}}\right),
$$

$$
g_j(x, y, z) = \tan^{-1}\left(\frac{z - z_{0j}}{[(x - x_{0j})^2 + (y - y_{0j})^2]^{1/2}}\right),
$$

and the design matrix is

$$
\begin{aligned}
&\Theta_{ij} \\
&= \begin{bmatrix}
1 & 0 & \tan(E_{ij})\sin(A_{ij}) & \tan(E_{ij})\cos(A_{ij}) \\
0 & 1 & \cos(A_{ij}) & -\sin(A_{ij})
\end{bmatrix} \\
&\qquad \begin{matrix}
\tan(E_{ij}) & \sec(E_{ij}) \\
0 & 0
\end{matrix} \Bigg].
\end{aligned}
$$

Then we get the following linear model:

$$
\begin{bmatrix} \Delta \tilde{A}_{ij} \\ \Delta \tilde{E}_{ij} \end{bmatrix} = J_j(\vec{P_i})\big|_{\vec{P}=\vec{P_i^*}} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \end{pmatrix} \\
+ \Theta_{ij} \begin{pmatrix} a_{j1} \\ \vdots \\ a_{j6} \end{pmatrix} + \begin{bmatrix} \varepsilon_{Aij} \\ \varepsilon_{Eij} \end{bmatrix}, \quad (12)
$$

for $j = 1, \ldots$ and $i = 1, 2, \ldots$, where $\Delta \tilde{A}_{ij} = A_{ij} - f_j(\vec{P_i^*}) \Delta \tilde{E}_{ij} = E_{ij} - g_j(\vec{P_i^*})$,

$$
J_1(\vec{P}) = \frac{\partial(f_j, g_j)}{\partial(x, y, z)},
$$

Integrating all m instruments, we get the following integrated error decomposition model:

$$
\begin{bmatrix} \Delta \tilde{A}_{i1} \\ \Delta \tilde{E}_{i1} \\ \vdots \\ \Delta \tilde{A}_{im} \\ \Delta \tilde{E}_{im} \end{bmatrix} = \begin{pmatrix} J_1(\vec{P}) \\ \vdots \\ J_m(\vec{P}) \end{pmatrix}\Bigg|_{\vec{P}=\vec{P_i^*}} \begin{pmatrix} \Delta x_i \\ \Delta y_i \\ \Delta z_i \end{pmatrix} \\
+ \bar{B}_i \begin{pmatrix} \alpha_{11} \\ \vdots \\ \alpha_{m6} \end{pmatrix} + \begin{bmatrix} \varepsilon_{Ai1} \\ \varepsilon_{Ei1} \\ \vdots \\ \varepsilon_{Aim} \\ \varepsilon_{Eim} \end{bmatrix}, \\
i = 1, 2, \ldots \quad (13)
$$

where $\bar{B}_i = \mathrm{diag}\{\Theta_{i1}, \ldots, \Theta_{im}\}$.

Obviously, the model (13) is very similar to the linear model (2) with partially variant parameters. Consequently, we can use the iterative algorithm (6) to calibrate the error coefficients in the transducers and, at the same time, to accurately determine the rocket trajectory.

In our simulations, there are four transducers tracking a rocket in space. Selecting the computation parameter $n_0 = 100[\mathrm{s}]$, we use (6) to get the modification values.

Table 1 contains the estimated values of the error coefficients estimated at 110 seconds. Table 2 includes inerements in the rocket trajectories after $n_0 = 100[\mathrm{s}]$.

The computation results given in Tables 1 and 2 show that the iterative algorithms given in Section 2 not only decrease the computation time, but also efficiently modify the precision of the rocket trajectory. What is more, this practical application shows that this new kind of linear model with variant coefficients is reasonable and valuable not only in theory, but also in various engineering fields.

Table 1. Estimation of error coefficients.

[mrad]

Transducer	α_{j1}	α_{j2}	α_{j3}	α_{j4}	α_{j5}	α_{j6}
$i = 1$	1.203	0.686	0.018	−0.006	0.006	0.003
$i = 2$	−0.058	−0.070	−0.000	0.000	−0.000	0.001
$i = 3$	1.087	−1.837	0.041	−0.016	−0.019	−0.004
$i = 4$	−0.514	1.449	−0.024	0.009	0.015	0.001

Table 2. Increments in values of trajectories.

[m]

$i = 100+$	Δx_i	Δy_i	Δz_i
1	−0.662445	0.631514	0.001622
2	−0.763551	0.687524	−0.006546
3	−0.760541	0.677017	−0.012729
4	−0.752472	0.673932	−0.017189
5	−0.793005	0.699980	−0.023827
6	0.835997	0.730210	−0.029760
7	−0.832480	0.731190	−0.036693
8	−0.802443	0.710644	−0.037840
9	−0.739471	0.661947	−0.036633
10	−0.739483	0.660732	−0.038780

4. Discussion

The paper not only presents a new kind of linear model, but also builds a series of convenient algorithms. This new model usefully generalized the widely used ordinary linear regression model. The new model can be used in many different kinds of fields, e.g., in data fusion, process monitoring, control engineering, etc.

As for the new algorithms, their advantages are evident. Obviously, if we use the old LS algorithm (5), we must compute the inverse of a large matrix $(\boldsymbol{H}_n^\tau \boldsymbol{H}_n)^{-1} \in \mathbb{R}^{(r+nq) \times (r+nq)}$. What is more, its dimension is continually increased as long as the number n of samples increases or the process moves on. On the other hand, if we use the new iterative algorithm (6), we just need to deal with a series of lower-dimensional inverse matrices, the highest dimension of which is equal to $\max\{p, q, r\}$. In fact, the iterative algorithm (6) involves three inverse matrices $(\boldsymbol{A}_i^\tau \boldsymbol{A}_i)^{-1} \in \mathbb{R}^{q \times q}$, $\boldsymbol{R}_{n+1}^{-1} \in \mathbb{R}^{p \times p}$ and $(\boldsymbol{L}_n + \boldsymbol{B}_{n+1}^\tau \boldsymbol{B}_{n+1})^{-1} \in \mathbb{R}^{r \times r}$.

References

Brown D.C. (1964): *The Error Model Best Estimation Trajectory.* — Tech. Rep. AD 602799:
`http://stinet.dtic.mil/oai/oai?&verb=getRecorg&metadataPrefix=html&identifier=AD0602799`

Dodge Y. and Kova J. (2000): *Adaptive Regression.* – Berlin: Springer.

Draper N.R. and Smith H. (1981): *Applied Regression Analysis.* – New York: Wiley.

Eubank R., Chunfeng H., Maldonado Y., Naisyin W., Suojin W., Buchanan R.J. (2004): *Smoothing spline estimation in varying coefficient models.* – J. Roy. Stat. Soc. B, Vol. 66, No. 3, pp. 653–667.

Fahrmeier L. and Tutz G.(2001): *Multivariate Statistical Modeling Based on Generalized Linear Models.* — Berlin: Springer.

Frank E. and Harrell J.(2002): *Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis.* — New York: Springer.

Graybill F.A. and Iyer H.K. (1994): *Regression Analysis: Concepts and Applications.* — Massachusetts: Duxbury Press.

Hu Shaolin and Sun Guoji (2001): *Process Monitoring Technique and Applications.* — Bejing: National Defense Industry Press.

Kala R. and Kłaczyński K. (1988): *Recursive improvement of estimates in a Gauss-Markov model with linear restrictions.* Canad. J. Stat., Vol. 16, No. 3, pp. 301–305.

Rencher A. (2000): *Linear Models in Statistics.* – New York: Wiley.

Appendix

In order to prove Theorem 1, recall two lemmas without proofs. They are fundamental in linear algebra.

Lemma 1. *If a block matrix $\boldsymbol{A}$ and its submatrix $\boldsymbol{A}_{11}$ in $\boldsymbol{A}$ are invertible, then we have*

$$
\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}^{-1}
$$

$$
= \begin{pmatrix} \boldsymbol{A}_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix}
$$

$$
+ \begin{pmatrix} \boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12} \\ -\boldsymbol{I} \end{pmatrix} (\boldsymbol{A}_{22} - \boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1}\boldsymbol{A}_{12})^{-1} (\boldsymbol{A}_{21}\boldsymbol{A}_{11}^{-1} \vdots - \boldsymbol{I}).
$$

$$\tag{A1}$$

Similarly, if a matrix $\boldsymbol{A}$ and its submatrix $\boldsymbol{A}_{22}$ are invertible, then we have

$$
\begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} \end{pmatrix}^{-1}
$$

$$
= \begin{pmatrix} 0 & 0 \\ 0 & \boldsymbol{A}_{22}^{-1} \end{pmatrix}
$$

$$
+ \begin{pmatrix} -\boldsymbol{I} \\ \boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21} \end{pmatrix} (\boldsymbol{A}_{11} - \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1}\boldsymbol{A}_{21})^{-1} (-\boldsymbol{I} \vdots \boldsymbol{A}_{12}\boldsymbol{A}_{22}^{-1})
$$

$$\tag{A2}$$

Lemma 2. *If matrices $\boldsymbol{F}$ and $\boldsymbol{G}$ are invertible and an inverse matrix $(\boldsymbol{F} - \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{K})^{-1}$ exists, then*

$$
(\boldsymbol{F} - \boldsymbol{H}\boldsymbol{G}^{-1}\boldsymbol{K})^{-1}
$$
$$
= \boldsymbol{F}^{-1} + \boldsymbol{F}^{-1}\boldsymbol{H}(\boldsymbol{G} - \boldsymbol{K}F^{-1}\boldsymbol{H})^{-1}\boldsymbol{K}\boldsymbol{F}^{-1}. \tag{A3}
$$

The proofs of these two lemmas can be found in the references (Draper and Smith, 1981).

Proof of Theorem 1. With the model (2) and n samples, Eqn. (5) shows that the LS estimator is

$$
\hat{\boldsymbol{\Phi}}_n^{LS(n)} = (\boldsymbol{H}_n^\tau \boldsymbol{H}_n)^{-1} \boldsymbol{H}_n^\tau \bar{\boldsymbol{Y}}_n.
$$

If there is another sampling datum

$$
\boldsymbol{Y}_{n+1} = \boldsymbol{A}_{n+1}\boldsymbol{X}_{n+1} + \boldsymbol{B}_{n+1}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_{n+1}
$$

which has been added into the sampling set, the LS estimators of all the coefficients in the model (2) must be changed in accordance with the following expressions:

$$
\hat{\boldsymbol{\Phi}}_{n+1}^{LS(n+1)} = (\boldsymbol{\Psi}_n^\tau \boldsymbol{\Psi}_n)^{-1} \boldsymbol{\Psi}_n^\tau \begin{bmatrix} \bar{\boldsymbol{Y}}_n \\ \boldsymbol{Y}_{n+1} \end{bmatrix}
$$

$$
= \begin{bmatrix} \boldsymbol{D}_{11} & \boldsymbol{C}_{n+1}^\tau \boldsymbol{A}_{n+1} \\ \boldsymbol{A}_{n+1}^\tau \boldsymbol{C}_{n+1} & \boldsymbol{A}_{n+1}^\tau \boldsymbol{A}_{n+1} \end{bmatrix}^{-1} \boldsymbol{\Psi}_n^\tau
$$

$$
\times \begin{bmatrix} \bar{\boldsymbol{Y}}_n \\ \boldsymbol{Y}_{n+1} \end{bmatrix}, \tag{A4}
$$

where

$$
\hat{\boldsymbol{\Phi}}_{n+1}^{LS(n+1)} = \begin{bmatrix} \hat{\boldsymbol{\Phi}}_n^{LS(n+1)} \\ \hat{\boldsymbol{X}}_{n+1}^{LS(n+1)} \end{bmatrix}, \quad \boldsymbol{\Psi}_n = \begin{bmatrix} \boldsymbol{H}_n & 0 \\ \boldsymbol{C}_{n+1} & \boldsymbol{A}_{n+1} \end{bmatrix},
$$

$$
\boldsymbol{C}_{n+1} = (\boldsymbol{B}_{n+1}, 0) \in \mathbb{R}^{p \times (q+nq)}
$$

$$
\boldsymbol{D}_{11} = \boldsymbol{H}_n^\tau \boldsymbol{H}_n + \boldsymbol{C}_{n+1}^\tau \boldsymbol{C}_{n+1}.
$$

Now, using the notation

$$D_{22} = A_{n+1}^\tau A_{n+1}, \quad D_{12} = C_{n+1}^\tau A_{n+1},$$

$$\Omega = D_{22} - D_{21} D_{11}^{-1} D_{12},$$

the following formula can be directly derived from Lemma 1:

$$\begin{pmatrix} \hat{\boldsymbol{\Phi}}_n^{LS(n+1)} \\ \hat{\boldsymbol{X}}_{n+1}^{LS(n+1)} \end{pmatrix}$$

$$= \left\{ \begin{pmatrix} D_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} D_{11}^{-1} D_{12} \\ -I \end{pmatrix} \Omega^{-1} \right.$$

$$\left. \begin{pmatrix} D_{21} D_{11}^{-1} \vdots -I \end{pmatrix} \right\} \begin{pmatrix} H_n^\tau \bar{Y}_n + C_{n+1}^\tau Y_{n+1} \\ A_{n+1}^\tau Y_{n+1} \end{pmatrix}$$

$$= E_n H_n^\tau \bar{Y}_n + F_n Y_{n+1}, \qquad (A5)$$

where

$$F_n = \begin{pmatrix} D_{11}^{-1}[C_{n+1}^\tau + D_{12}\Omega^{-1} D_{21} D_{11}^{-1} C_{n+1}^\tau \\ - D_{12}\Omega^{-1} A_{n+1}^\tau] \\ -\Omega^{-1}[D_{21} D_{11}^{-1} C_{n+1}^\tau - A_{n+1}^\tau] \end{pmatrix}$$

and

$$E_n = \begin{pmatrix} D_{11}^{-1} + D_{11}^{-1} D_{12}\Omega^{-1} D_{21} D_{11}^{-1} \\ -\Omega^{-1} D_{21} D_{11}^{-1} \end{pmatrix}.$$

Step 1. We analyze the expression E_n. From the expression for the block matrix D_{11} and Lemma 1 we have

$$D_{11}^{-1} = \begin{bmatrix} \sum_{i=1}^{n+1} B_l^\tau B_i & B_1^\tau A_1 & \cdots & B_n^\tau A_n \\ A_1^\tau B_1 & A_1^\tau A_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_n^\tau B_n & 0 & \cdots & A_n^\tau A_n \end{bmatrix}^{-1}.$$

It can be shown that the first r rows of the matrix D_{11}^{-1} can be expressed as follows:

$$D_{11}^{-1} = \begin{bmatrix} \left(\sum_{i=1}^{n+1} B_i^\tau B_i - \sum_{i=1}^{n+1} B_i^\tau U_i A_i^\tau B_i \right)^{-1} & * \\ -T_n \left(\sum_{i=1}^{n+1} B_i^\tau B_i - \sum_{i=1}^{n+1} B_i^\tau U_i A_i^\tau B_i \right)^{-1} & * \end{bmatrix},$$

(A6)

where

$$U_i = A_i (A_i^\tau A_i)^{-1}, \ T_n = \begin{bmatrix} B_1^\tau U_1, \ldots, B_n^\tau U_n \end{bmatrix}^\tau,$$

and the asterisk '*' denotes an omitted matrix block which is rather complicated and does not have any effect on the following deduction process.

Analyzing the formulas for the matrix blocks D_{22} and $D_{12} = D_{21}$, we get

$$D_{11}^{-1} + D_{11}^{-1} D_{12}\Omega^{-1} D_{21} D_{11}^{-1}$$

$$= \{ I + D_{11}^{-1} C_n^\tau V_{n+1} C_{n+1} \} D_{11}^{-1}, \quad (A7)$$

where

$$V_{n+1} = A_{n+1}(A_{n+1}^\tau A_{n+1}$$

$$- A_{n+1}^\tau C_{n+1} D_{11}^{-1} C_{n+1}^\tau A_{n+1})^{-1} A_{n+1}^\tau,$$

and

$$D_{11}^{-1} C_n^\tau V_{n+1} C_{n+1}$$

$$= D_{11}^{-1} \begin{pmatrix} B_{n+1}^\tau A_{n+1} W_{k+1}^{-1} A_{n+1}^\tau B_{n+1} & 0 \\ 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \tilde{L}_{n+1}^{-1} B_{n+1}^\tau A_{n+1} W_{k+1}^{-1} A_{n+1}^\tau B_{n+1} & 0 \\ -T_n \tilde{L}_{n+1}^{-1} B_{n+1}^\tau A_{n+1} W_{k+1}^{-1} A_{n+1}^\tau B_{n+1} & 0 \end{pmatrix},$$

(A8)

where

$$W_{k+1} = A_{n+1}^\tau (I - B_{n+1} \tilde{L}_{n+1}^{-1} B_{n+1}^\tau) A_{n+1},$$

$$\tilde{L}_{n+1} = \sum_{i=1}^{n} B_i^\tau [I - U_i A_i^\tau] B_i + B_{n+1}^\tau B_{n+1}.$$

The matrix D_{11}^{-1} can be expressed as follows:

$$D_{11}^{-1} = (H_n^\tau H_n)^{-1} - (H_n^\tau H_n)^{-1} C_{n+1}^\tau$$

$$\times \left[I + C_{n+1}^\tau (H_n^\tau H_n)^{-1} C_{n+1}^\tau \right]^{-1}$$

$$\times C_{n+1}(H_n^\tau H_n)^{-1}. \qquad (A9)$$

On second thoughts, using the notation

$$L_n = \sum_{i=1}^{n} B_i^\tau \left[I - A_i (A_i^\tau A_i)^{-1} A_i^\tau \right] B_i,$$

we have

$$(H_n^\tau H_n)^{-1} = \begin{bmatrix} \sum_{i=1}^{n} B_1^\tau B_i & B_1^\tau A_1 & \ldots & B_n^\tau A_n \\ A_1^\tau B_1 & A_1^\tau A_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ A_n^\tau B_n & 0 & \cdots & A_n^\tau A_n \end{bmatrix}^{-1}$$

$$= \begin{pmatrix} L_n^{-1} & -L_n^{-1} T_n^\tau \\ -T_n L_n^{-1} & * \end{pmatrix}, \qquad (A10)$$

so we get

$$D_{11}^{-1} = \left\{ I - \begin{pmatrix} L_n^{-1} \tilde{Z}_{n+1} & 0 \\ -T_n L_n^{-1} \tilde{Z}_{n+1} & 0 \end{pmatrix} \right\} (H_n^\tau H_n)^{-1},$$

(A11)

where $\tilde{Z}_{n+1} = B_{n+1}^\tau (I + B_{n+1} L_n^{-1} B_{n+1}^\tau)^{-1} B_{n+1}.$

On the other hand, the matrix $\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21} = (\boldsymbol{D}_{22} - \boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1}\boldsymbol{D}_{12})^{-1}\boldsymbol{D}_{21}$ can be simplified as follows:

$$
\begin{aligned}
\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21} &= \left\{ \boldsymbol{A}_{n+1}^{\tau}(\boldsymbol{I} - \boldsymbol{C}_{n+1}\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau})\boldsymbol{A}_{n+1} \right\}^{-1} \\
&\quad \times \boldsymbol{A}_{n+1}^{\tau}\boldsymbol{C}_{n+1} \\
&= \left\{ \boldsymbol{A}_{n+1}^{\tau}(\boldsymbol{I} - \boldsymbol{B}_{n+1}\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau})\boldsymbol{A}_{n+1})^{-1} \right\} \\
&\quad \times \boldsymbol{A}_{n+1}^{\tau}(\boldsymbol{B}_{n+1}, \boldsymbol{0}\cdots\boldsymbol{0}).
\end{aligned} \tag{A12}
$$

Write

$$
\begin{aligned}
\boldsymbol{Q}_{n+1} &= \{ \boldsymbol{A}_{n+1}^{\tau}(\boldsymbol{I} - \boldsymbol{B}_{n+1}\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau})\boldsymbol{A}_{n+1} \}^{-1}, \\
&\quad \times \boldsymbol{A}_{n+1}^{\tau}\boldsymbol{B}_{n+1} \\
\boldsymbol{P}_{n+1} &= \boldsymbol{B}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{Q}_{n+1}.
\end{aligned}
$$

Then

$$
\begin{aligned}
&\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1} \\
&= (\boldsymbol{Q}_{n+1}, \boldsymbol{0}_{p\times nq}) \left\{ \boldsymbol{I} - \begin{pmatrix} \boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1} & \boldsymbol{0} \\ -\boldsymbol{T}_n\boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1} & \boldsymbol{0} \end{pmatrix} \right\} \\
&\quad \times (\boldsymbol{H}_n^{\tau}\boldsymbol{H}_n)^{-1} \\
&= (\boldsymbol{Q}_{n+1}[\boldsymbol{I} - \boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1}] \,\vdots\, \boldsymbol{0})(\boldsymbol{H}_n^{\tau}\boldsymbol{H}_n)^{-1} \tag{A13}
\end{aligned}
$$

and

$$
\begin{aligned}
&\boldsymbol{D}_{11}^{-1} + \boldsymbol{D}_{11}^{-1}\boldsymbol{D}_{12}\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1} \\
&= \left\{ \boldsymbol{I} + \begin{pmatrix} \boldsymbol{L}_{n+1}^{-1}\boldsymbol{P}_{n+1} & \boldsymbol{0} \\ \boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1} & \boldsymbol{0} \end{pmatrix} \right\} \\
&\quad \times \left\{ \boldsymbol{I} - \begin{pmatrix} \boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1} & \boldsymbol{0} \\ -\boldsymbol{T}_n\boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1} & \boldsymbol{0} \end{pmatrix} \right\} (\boldsymbol{H}_n^{\tau}\boldsymbol{H}_n)^{-1} \\
&= \begin{pmatrix} \boldsymbol{I}+\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1}-(\boldsymbol{I}+\tilde{\boldsymbol{L}}_{n+1}\boldsymbol{P}_{n+1})\boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1} & \boldsymbol{0} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1}+\boldsymbol{T}_n(\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1}+\boldsymbol{I})\boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1} & \boldsymbol{I} \end{pmatrix} \\
&\quad \times (\boldsymbol{H}_n^{\tau}\boldsymbol{H}_n)^{-1}. \tag{A14}
\end{aligned}
$$

Combining (A5), (A13) and (A14), we get

$$
\begin{aligned}
\boldsymbol{E}_n &= \begin{pmatrix} \boldsymbol{I}+\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1}(\boldsymbol{I}-\tilde{\boldsymbol{E}}_n)-\tilde{\boldsymbol{E}}_n & \boldsymbol{0} \\ -\boldsymbol{T}_n[\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1}(\boldsymbol{I}-\tilde{\boldsymbol{E}}_n)-\tilde{\boldsymbol{E}}_n] & \boldsymbol{I} \\ -\boldsymbol{Q}_{n+1}(\boldsymbol{I}-\tilde{\boldsymbol{E}}_{(n)}) & \boldsymbol{0} \end{pmatrix} \\
&\quad \times (\boldsymbol{H}_n^{\tau}\boldsymbol{H}_n)^{-1}, \tag{A15}
\end{aligned}
$$

where $\tilde{\boldsymbol{E}}_n = \boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1}$. Using

$$
\begin{aligned}
\tilde{\boldsymbol{L}}_{n+1}^{-1} &= (\boldsymbol{L}_n + \boldsymbol{B}_{n+1}^{\tau}\boldsymbol{B}_{n+1})^{-1} \\
&= \boldsymbol{L}_n^{-1} - \boldsymbol{L}_n^{-1}\tilde{\boldsymbol{Z}}_{n+1}\boldsymbol{L}_n^{-1},
\end{aligned}
$$

we get

$$
\boldsymbol{I} - \tilde{\boldsymbol{E}}_n = \boldsymbol{L}_{n+1}^{-1}\boldsymbol{L}_n. \tag{A16}
$$

Consequently, Eqn. (25) can be expressed as

$$
\begin{aligned}
\boldsymbol{E}_n &= \begin{pmatrix} \boldsymbol{I}+\boldsymbol{L}_{n+1}^p+\tilde{\boldsymbol{L}}_{n+1}^{-1}[\boldsymbol{L}_n-\tilde{\boldsymbol{L}}_{n+1}^{-1}] & \boldsymbol{0} \\ -\boldsymbol{T}_n[\boldsymbol{L}_{n+1}^p+\tilde{\boldsymbol{L}}_{n+1}^{-1}(\boldsymbol{L}_n-\tilde{\boldsymbol{L}}_{n+1}^{-1})] & \boldsymbol{I} \\ -\boldsymbol{Q}_{n+1}\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{L}_n & \boldsymbol{0} \end{pmatrix} \\
&\quad \times (\boldsymbol{H}_n^{\tau}\boldsymbol{H}_n)^{-1}, \tag{A17}
\end{aligned}
$$

where $\boldsymbol{L}_{n+1}^p = \tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{P}_{n+1}\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{L}_n$.

Step 2. Let us analyze the expression for $\boldsymbol{F}_n$. From (A 6), we have

$$
\begin{aligned}
&\boldsymbol{D}_{12}\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau} \\
&\quad = \boldsymbol{C}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau}\boldsymbol{C}_{n+1}\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau}.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\boldsymbol{D}_{12}\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau} \\
&= \begin{pmatrix} \boldsymbol{B}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau}\boldsymbol{B}_{n+1} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \\ \vdots & & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \cdots & \boldsymbol{0} \end{pmatrix} \\
&\quad \times \begin{pmatrix} \tilde{\boldsymbol{L}}_{n+1}^{-1} & -\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{T}_n^{\tau} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1} & * \end{pmatrix} \begin{pmatrix} \boldsymbol{B}_{n+1}^{\tau} \\ \boldsymbol{0} \end{pmatrix} \\
&= \begin{pmatrix} \boldsymbol{B}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau}\boldsymbol{B}_{n+1}\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1} \\ \boldsymbol{0} \end{pmatrix} \tag{A18}
\end{aligned}
$$

and we have the following four equations:

$$
\begin{aligned}
&\boldsymbol{D}_{11}^{-1}\left[\boldsymbol{D}_{12}\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau}\right] \\
&= \begin{pmatrix} \tilde{\boldsymbol{L}}_{n+1}^{-1} & -\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{T}_n^{\tau} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1} & * \end{pmatrix} \begin{pmatrix} \boldsymbol{M}_{n+1} \\ \boldsymbol{0} \end{pmatrix} \\
&= \begin{pmatrix} \tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{M}_{n+1} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{M}_{n+1} \end{pmatrix}, \tag{A19}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau} &= \begin{pmatrix} \tilde{\boldsymbol{L}}_{n+1}^{-1} & -\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{T}_n^{\tau} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1} & * \end{pmatrix} \begin{pmatrix} \boldsymbol{B}_{n+1}^{\tau} \\ \boldsymbol{0} \end{pmatrix} \\
&= \begin{pmatrix} \tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau} \end{pmatrix}, \tag{A20}
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{D}_{11}^{-1}\boldsymbol{D}_{12}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau} &= \boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau} \\
&= \begin{pmatrix} \tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau} \\ -\boldsymbol{T}_n\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau}\boldsymbol{A}_{n+1}\boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau} \end{pmatrix}, \tag{A21}
\end{aligned}
$$

$$
\boldsymbol{\Omega}^{-1}\boldsymbol{D}_{21}\boldsymbol{D}_{11}^{-1}\boldsymbol{C}_{n+1}^{\tau} = \boldsymbol{\Omega}^{-1}\boldsymbol{A}_{n+1}^{\tau}\boldsymbol{B}_{n+1}\tilde{\boldsymbol{L}}_{n+1}^{-1}\boldsymbol{B}_{n+1}^{\tau}, \tag{A22}
$$

where

$$M_{n+1} = B_{n+1}^\tau A_{n+1}\Omega^{-1}A_{n+1}^\tau B_{n+1}\tilde{L}_{n+1}^{-1}B_{n+1}.$$

Inserting (A19)–(A22) into the expression for F_n which was defined in (A 5), we immediately have
$$F_n$$
$$= \begin{pmatrix} \tilde{L}_{n+1}^{-1}(B_{n+1}^\tau + M_{n+1} - B_{n+1}^\tau A_{n+1}\Omega^{-1}A_{n+1}^\tau) \\ -T_n\tilde{L}_{n+1}^{-1}(B_{n+1}^\tau + M_{n+1} - B_{n+1}^\tau A_{n+1}\Omega^{-1}A_{n+1}^\tau) \\ -Q^{-1}(A_{n+1}^\tau B_{n+1}\tilde{L}_{n+1}^{-1}B_{n+1}^\tau - A_{n+1}^\tau) \end{pmatrix}$$
$$\tag{A23}$$

From the definition of Ω and Eqn. (A16), we have

$$\begin{aligned}\Omega &= A_{n+1}^\tau A_{n+1} - A_{n+1}^\tau C_{n+1}D_{11}^{-1}C_{n+1}^\tau A_{n+1}\\ &= A_{n+1}^\tau A_{n+1} - A_{n+1}^\tau B_{n+1}\tilde{L}_{n+1}^{-1}B_{n+1}^\tau A_{n+1}\\ &= A_{n+1}^\tau(I - B_{n+1}\tilde{L}_{n+1}^{-1}B_{n+1}^\tau)A_{n+1}.\end{aligned}$$

Hence

$$\Omega^{-1} = [A_{n+1}^\tau R_{n+1}A_{n+1}]^{-1}, \tag{A24}$$

where $R_{n+1} = I - B_{n+1}\tilde{L}_{n+1}^{-1}B_{n+1}^\tau$. Inserting (A18) and (A24) into (A23), we have
$$F_n$$
$$= \begin{pmatrix} \tilde{L}_{n+1}^{-1}B_{n+1}^\tau[I - A_{n+1}(A_{n+1}^\tau R_{n+1}A_{n+1})^{-1} \\ A_{n+1}^\tau R_{n+1}] \\ -T_n\tilde{L}_{n+1}^{-1}B_{n+1}^\tau[I - A_{n+1}(A_{n+1}^\tau R_{n+1}A_{n+1})^{-1} \\ A_{n+1}^\tau R_{n+1}] \\ (A_{n+1}^\tau R_{n+1}A_{n+1})^{-1}A_{n+1}^\tau R_{n+1} \end{pmatrix}$$
$$\tag{A25}$$

Step 3. From Steps 1 and 2, integrating (A18) and (A26) into (A6), we immediately get the following expression:

$$\begin{pmatrix} \hat{\Phi}_n^{LS(n+1)} \\ \hat{X}_{n+1}^{LS(n+1)} \end{pmatrix} = M_{n+1}^a\hat{\Phi}_n^{LS(n)} + M_{n+1}^b Y_{n+1}, \quad \text{(A26)}$$

where

$$M_{n+1}^a$$
$$= \begin{pmatrix} I + \tilde{L}_{n+1}^{-1}P_{n+1}\tilde{L}_{n+1}^{-1}L_n - \tilde{L}_{n+1}^{-1}B_{n+1}^\tau B_{n+1} & 0 \\ -T_n(\tilde{L}_{n+1}^{-1}P_{n+1}\tilde{L}_{n+1}^{-1}L_n - \tilde{L}_{n+1}^{-1}B_{n+1}^\tau B_{n+1}) & I \\ -Q_{n+1}\tilde{L}_{n+1}^{-1}L_n & 0 \end{pmatrix},$$

$$M_{n+1}^b$$
$$= \begin{pmatrix} \tilde{L}_{n+1}^{-1}B_{n+1}^\tau[I - A_{n+1}(A_{n+1}^\tau R_{n+1}A_{n+1})^{-1} \\ A_{n+1}^\tau R_{n+1}] \\ -T_n\tilde{L}_{n+1}^{-1}B_{n+1}^\tau[I - A_{n+1}(A_{n+1}^\tau R_{n+1}A_{n+1})^{-1} \\ A_{n+1}^\tau R_{n+1}] \\ (A_{n+1}^\tau R_{n+1}A_{n+1})^{-1}A_{n+1}^\tau R_{n+1} \end{pmatrix}.$$

Setting

$$\Xi_{n+1} = A_{n+1}(A_{n+1}^\tau R_{n+1}A_{n+1})^{-1}A_{n+1}^\tau,$$

we have

$$\begin{aligned} &P_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)} \\ &\quad - B_{n+1}^\tau A_{n+1}(A_{n+1}^\tau R_{n+1}A_{n+1})^{-1}A_{n+1}^\tau R_{n+1}Y_{n+1} \\ &= B_{n+1}^\tau \Xi_{n+1}B_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)} \\ &\quad - B_{n+1}^\tau \Xi_{n+1}R_{n+1}Y_{n+1} \\ &= B_{n+1}^\tau \Xi_{n+1}\Big[B_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)} - R_{n+1}Y_{n+1} \Big] \end{aligned}$$

and

$$\begin{aligned} &B_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)} - R_{n+1}Y_{n+1} \\ &\quad = B_{n+1}(I - \tilde{L}_{n+1}^{-1}B_{n+1}^\tau B_{n+1})\hat{\beta}^{LS(n)} \\ &\quad\quad - (I - B_{n+1}^\tau \tilde{L}_{n+1}^{-1}B_{n+1})Y_{n+1} \\ &\quad = (B_{n+1}\tilde{L}_{n+1}^{-1}B_{n+1}^\tau - I)(Y_{n+1} - B_{n+1}\hat{\beta}^{LS(n)}) \\ &\quad = -R_{n+1}(Y_{n+1} - B_{n+1}\hat{\beta}^{LS(n)}). \end{aligned}$$

Consequently, decomposing the matrix equation (A26) into appropriate blocks, we get
$$\hat{\beta}^{LS(n+1)}$$

$$\begin{aligned} &= \hat{\beta}^{LS(n)} + \tilde{L}_{n+1}^{-1}\Big[P_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)} \\ &\quad - B_{n+1}^\tau \Xi_{n+1}R_{n+1}Y_{n+1} \Big] \\ &\quad + \tilde{L}_{n+1}^{-1}B_{n+1}(Y_{n+1} - B_{n+1}\hat{\beta}^{LS(n)}) \\ &= \hat{\beta}^{LS(n)} + [\tilde{L}_{n+1}^{-1}B_{n+1}^\tau - \tilde{L}_{n+1}^{-1}B_{n+1}^\tau \Xi_{n+1}R_{n+1}] \\ &\quad \times (Y_{n+1} - B_{n+1}\hat{\beta}^{LS(n)}) \\ &= \hat{\beta}^{LS(n)} + \tilde{L}_{n+1}^{-1}B_{n+1}^\tau[R_{n+1}^{-1} - \Xi_{n+1}] \\ &\quad \times R_{n+1}(Y_{n+1} - B_{n+1}\hat{\beta}^{LS(n)}) \end{aligned} \tag{A27}$$

and
$$\begin{aligned} &\begin{pmatrix} \hat{X}_1^{LS(n+1)} \\ \vdots \\ \hat{X}_n^{LS(n+1)} \end{pmatrix} = \begin{pmatrix} \hat{X}_1^{LS(n)} \\ \vdots \\ \hat{X}_n^{LS(n)} \end{pmatrix} \\ &\quad - \begin{pmatrix} (A_1^\tau A_1)^{-1}A_1^\tau B_1 \\ \vdots \\ (A_n^\tau A_n)^{-1}A_n^\tau B_n \end{pmatrix} \tilde{L}_{n+1}^{-1}B_{n+1}^\tau \\ &\quad \times \left(R_{n+1}^{-1} - \Xi_{n+1} \right) R_{n+1}\Big(Y_n - B_{n+1}\hat{\beta}^{LS(n)} \Big). \end{aligned} \tag{A28}$$

From (A28), for $i = 1, 2, \ldots, n$, we have

$$\begin{aligned} \hat{X}_i^{LS(n+1)} &= \hat{X}_i^{LS(n)} \\ &\quad + (A_i^\tau A_i)^{-1}A_i^\tau B_i(\hat{\beta}^{LS(n)} - \hat{\beta}^{LS(n+1)}) \end{aligned} \tag{A29}$$

and

$$\hat{X}_{n+1}^{LS(n+1)} = -Q_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)}$$
$$+ (A_{n+1}^{\tau}R_{n+1}A_{n+1})^{-1}A_{n+1}^{\tau}R_{n+1}A_{n+1}$$
$$= (A_{n+1}^{\tau}R_{n+1}A_{n+1})^{-1}A_{n+1}^{\tau}$$
$$\times (R_{n+1}Y_{n+1} - B_{n+1}\tilde{L}_{n+1}^{-1}L_n\hat{\beta}^{LS(n)})$$
$$= (A_{n+1}^{\tau}R_{n+1}A_{n+1})^{-1}A_{n+1}^{\tau}$$
$$\times (R_{n+1}Y_{n+1} - R_{n+1}B_{n+1}\hat{\beta}^{LS(n)})$$
$$= (A_{n+1}^{\tau}R_{n+1}A_{n+1})^{-1}A_{n+1}^{\tau}R_{n+1}$$
$$\times (Y_{n+1} - B_{n+1}\hat{\beta}^{LS(n)}). \qquad \text{(A30)}$$

Integrating (A27)–(A30), the results of Theorem 1 are obtained. ■