ITERATIVE ESTIMATORS OF PARAMETERS IN A LINEAR MODEL WITH PARTIALLY VARIANT COEFFICIENTS

HU SHAOLIN[†‡], MEINKE KARL[‡], CHEN RUSHAN[†] and HUAJIANG OUYANG[§]

† Nanjing University of Science and Technology, 210071, Nanjing, China; hshaolin@ustc.edu.cn
‡ School of Computer Science and Communication, Royal Institute of Technology, Stockholm, 100-44, Sweden
§University of Liverpool, Liverpool, United Kingdom; 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 effectively overcome some difficulties in computation with multidimensional inputs and iteratively appended parameters. An important application is described at the end of this article, which shows that this new model is reasonable and applicable in practical situations.

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[1,2]

$$\mathbf{Y}_i = \mathbf{B}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \qquad (1)$$

where $\mathbf{Y}_i \in R^p, \mathbf{B}_i \in R^{p\times r}$ $(i = 1,2,\cdots,n)$ and the vector $\boldsymbol{\beta} \in R^r$ is a constant parameter vector to be estimated and $\boldsymbol{\varepsilon}_i \in R^p$ are errors arising from measurement or stochastic noises from disturbance.

Some excellent theories and practical results have been published for statistical inferences and stochastic decisions using this model. This model has also been successfully applied to many kinds of practical engineering problems (see Draper and Smith, 1981; Frank and Harrell, 2002; Graybill and Iyer, 1994; Hu and Sun, 2001).

Further research on this model shows that the limitation of constant coefficients in model (1) is quite restrictive and strong. In other words, there are some practical situations in which this linear model cannot be applied (see Brown, 1964; Hu and Sun, 2001). Although there has been a lot of further research to generalize or adapt this linear model (1), (see e.g. Fahrmeier and Tutz, 2001; Dodge and Kova, 2000) the constraint on constant coefficients has so far not been essentially relaxed.

In order to overcome this limitation on the generality of the model (1) we set up a new linear model with partially variant coefficients as follows:

$$\mathbf{Y}_i = \mathbf{A}_i\mathbf{X}_i + \mathbf{B}_i\boldsymbol{\beta} + \boldsymbol{\varepsilon}_i \qquad (2)$$

where $\mathbf{Y}_i \in R^p, \mathbf{A}_i \in R^{p\times q}, \mathbf{B}_i \in R^{p\times r}$ and $\{\mathbf{X}_i \in R^q\}$ is a variant vector series. Generally, the dimension p of the measurement output must be larger than the dimension q of the variant coefficients, so $p > q$, in order to make sure 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 components, i.e. $q = 0$, then the model (2) simplifies to the ordinary linear model (1).

In Section 2, we consider estimating all of the model coefficients in (2) under the Gauss-Markov assumptions (in Radoslaw and Krzysztof, 1988). A series of iterative algorithms are introduced that allow us to estimate the coefficients which include the constant parameters $\boldsymbol{\beta} \in R^r$ and the variant vector series $\{\mathbf{X}_i \in R^q\}$. In Section 3, a proof of the main result is given. In Section 4, a practical application is described and some computational results are presented, which show that this new model is useful in practise.

[*] Corresponding author: Hu Shaolin. Xi'an City, P.O.Box 505#16, Shannxi,China; Email: hshaolin@ustc.edu.cn

2. Iterative Estimators of Variant Coefficients

In order to make sure the results given in this section are universal, we first assume that the coefficient series elements $\{\mathbf{X}_i \in R^q\}$ are not related at different sampling points. In order to simplify the notation, we use the notation $\mathbf{\Phi}_n = (\boldsymbol{\beta}^\tau, \mathbf{X}_1^\tau, \cdots, \mathbf{X}_n^\tau)^\tau \in R^{r+nq}$, for vectors, where the superscript τ denotes the transpose of a matrix as well as a vector, and matrix

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

Under the following three well known Gauss-Markov assumptions (c.f. Radoslaw and Krzysztof, 1988 or Rencher, 2000) on random errors $\{\boldsymbol{\varepsilon}_i \in R^p\}$: (i) the error $\boldsymbol{\varepsilon}_i$ has expected value 0; (ii) the error series values $\{\boldsymbol{\varepsilon}_i, i=1,2,\cdots\}$ are uncorrelated, and (iii) the error series values $\{\boldsymbol{\varepsilon}_i, i=1,2,\cdots\}$ are homoscedastic, i.e. they all have the same variance. Then the least squared (LS-) estimators of coefficients in model (2) can be expressed as follows:

$$\hat{\mathbf{\Phi}}_n^{LS(n)} = \arg \underset{\boldsymbol{\beta}\in R^p, \{\mathbf{X}_i \in R^q\}}{\text{minimizing}} \sum_{i=1}^n \left\| \mathbf{Y}_i - \left(\mathbf{A}_i \mathbf{X}_i + \mathbf{B}_i \boldsymbol{\beta}\right) \right\|^2 \quad (4)$$

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

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

where $\overline{\mathbf{Y}}_n = (\mathbf{Y}_1^\tau, \cdots, \mathbf{Y}_n^\tau)^\tau \in R^{np}$.

In order to make sure the matrix $\mathbf{H}_n^\tau \mathbf{H}_n$ is reversible, the dimensions of the model (2) must satisfy the restriction $p > q$ and the sample capacity must satisfy $n > r/(p-q)$; otherwise, the reversion operator of matrix ¨$\mathbf{H}_n^\tau \mathbf{H}_n$£© in formula (5) must be substituted by the "+" reversion operator, namely ¨$\mathbf{H}_n^\tau \mathbf{H}_n$£©.

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

$$\begin{aligned} \hat{\boldsymbol{\beta}}^{LS(n+1)} &= \hat{\boldsymbol{\beta}}^{LS(n)} + (\mathbf{L}_n + \mathbf{B}_{n+1}^\tau \mathbf{B}_{n+1})^{-1} \mathbf{B}_{n+1}^\tau (\mathbf{R}_{n+1}^{-1} - \mathbf{\Omega}_{n+1}) \\ &\quad \cdot \mathbf{R}_{n+1}(\mathbf{Y}_{n+1} - \mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}) \end{aligned}$$

$$\begin{cases} \hat{\mathbf{X}}_i^{LS(n+1)} = \hat{\mathbf{X}}_i^{LS(n)} + (\mathbf{A}_i^\tau \mathbf{A}_i)^{-1} \mathbf{A}_i^\tau \mathbf{B}_i (\hat{\boldsymbol{\beta}}^{LS(n)} - \hat{\boldsymbol{\beta}}^{LS(n+1)}) \\ \hat{\mathbf{X}}_{n+1}^{LS(n+1)} = (\mathbf{A}_{n+1}^\tau \mathbf{R}_{n+1} \mathbf{A}_{n+1})^{-1} \mathbf{A}_{n+1}^\tau \mathbf{R}_{n+1} (\mathbf{Y}_{n+1} - \mathbf{B}_{n+1} \hat{\boldsymbol{\beta}}^{LS(n)}) \end{cases} \quad (i = 1, 2, \cdots, n)$$

(6)

where
$$\begin{cases} \mathbf{L}_n = \sum_{i=1}^n \mathbf{B}_i^\tau [\mathbf{I} - \mathbf{A}_i (\mathbf{A}_i^\tau \mathbf{A}_i)^{-1} \mathbf{A}_i^\tau] \mathbf{B}_i \in R^{r\times r} \\ \mathbf{\Omega}_{n+1} = \mathbf{A}_{n+1} (\mathbf{A}_{n+1}^\tau \mathbf{R}_{n+1} \mathbf{A}_{n+1})^{-1} \mathbf{A}_{n+1}^\tau \in R^{p\times p} \\ \mathbf{R}_{n+1} = \mathbf{I} - \mathbf{B}_{n+1} (\mathbf{L}_n + \mathbf{B}_{n+1}^\tau \mathbf{B}_{n+1})^{-1} \mathbf{B}_{n+1}^\tau \in R^{p\times p} \end{cases}.$$

and the superscript in expression $\hat{\mathbf{X}}_i^{LS(n)}$ denotes the LS-estimators of $\mathbf{X}$£¨$i = 1, 2, \cdots, n$).

Proof The proof of Theorem 1 will be given in Appendix 1.

Obviously, the algorithm (6) is iterative and very wieldy in practical engineering applications. However, some obvious optimizations exist:

- $\hat{\boldsymbol{\beta}}^{LS(n+1)}$ is a linear combination of the estimator $\hat{\boldsymbol{\beta}}^{LS(n)}$ with innovation $\{\mathbf{Y}_{n+1} - \mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}\}$ from new sampling data.
- $\hat{\boldsymbol{\beta}}^{LS(n+1)}$ can be computed directly from estimator $\hat{\boldsymbol{\beta}}^{LS(n)}$ without directly involving old sampling data $\{\mathbf{Y}_i, i = 1, \cdots, n\}$ as well as estimations $\{\hat{\mathbf{X}}_i^{LS(n)}\}$;
- It is clear that the estimators $\hat{\mathbf{X}}_{n+1}^{LS(n+1)}$ are determined with innovation $\{\mathbf{Y}_{n+1} - \mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}\}$, which coincides with our understanding of the model (2);
- The estimation $\hat{\mathbf{X}}_i^{LS(n+1)}$ of $\mathbf{X}_i$ $(i \le n)$ can be adjusted accurately in succession with the estimator error of the constant parameter vector $\boldsymbol{\beta}$.

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

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

where $n_0 \in N$ must satisfy the constraint

$$n_0 > r/(p-q)$$

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

Theorem 2 Suppose the LS estimators (7) are chosen as initial estimates of the coefficients of model (2). If the disturbance $\{\boldsymbol{\varepsilon}_n \in R^p, n \in N\}$ is a stationary Gaussian white noise process with mean zero, then the iterative estimators (6) are unbiased.

Proof In order to prove Theorem 2, we just need to show that $E\hat{\boldsymbol{\beta}}^{LS(n)} = \boldsymbol{\beta}$, where the operator E denotes mathematical expectation or a stochastic variable. In fact,

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

where $\tilde{\mathbf{L}}_n = \sum_{i=1}^{n-1}\mathbf{B}_i^{\tau}[\mathbf{I}\text{-}\mathbf{A}_i(\mathbf{A}_i^{\tau}\mathbf{A}_i)^{-1}\mathbf{A}_i^{\tau}]\mathbf{B}_i + \mathbf{B}_n^{\tau}\mathbf{B}_n$ and $\boldsymbol{\Xi}_{n+1} = \mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}$. Using the equation (6), we get

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

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

Applying the principle of mathematical induction the result follows.

3 A Practical Application

Our new linear model (2) can be widely used in many different practical fields, e.g. in data fusion, in modeling and monitoring of computer controlled systems, in signal processing, and in spacecraft control engineering, etc.

In this section, we present an application of the model (2) to calculate the trajectory of rocket. Suppose that there are m transit instruments which are suitably located at different sites separately. These m transits are simultaneously used to track a payload carrying rocket M in space. Using these transits, we get a series of measurement data $\{(\mathrm{A}_j(\mathrm{t}_i), \mathrm{E}_j(\mathrm{t}_i)) \mid i = 1, 2, \cdots, n; j = 1, 2, \cdots, m\}$, where $\mathrm{A}_j(\mathrm{t}_i)$ denotes the azimuth and $\mathrm{E}_j(\mathrm{t}_i)$ denotes the elevation of the rocket M, at time t_i, with respect to a reference frame fixed at the center of the transit instrument j.

In order to simplify the expressions below, we use some abbreviated notations such as $\mathrm{A}_{ij} = \mathrm{A}_j(\mathrm{t}_i)$ and $\mathrm{E}_{ij} = \mathrm{E}_j(\mathrm{t}_i)$, etc. So, the error decomposition models used in determining the location of the spacecraft M can be set up as follows (see Brown, 1964; Hu and Sun, 2001)

$$\begin{cases} \mathrm{A}_{ij} = \mathrm{tg}^{-1}\dfrac{\mathrm{x} - \mathrm{x}_{0j}}{\mathrm{y} - \mathrm{y}_{0j}} + \alpha_{j1} + \alpha_{j3}\mathrm{tgE}_{ij}\mathrm{sinA}_{ij} + \alpha_{j4}\mathrm{tgE}_{ij}\mathrm{cosA}_{ij} \\ \qquad + \alpha_{j5}\mathrm{tgE}_{ij} + \alpha_{j6}\mathrm{secE}_{ij} + \varepsilon_{\mathrm{A}ij} \\ \mathrm{E}_{ij} = \mathrm{tg}^{-1}\dfrac{\mathrm{z} - \mathrm{z}_{0j}}{[(\mathrm{x} - \mathrm{x}_{0j})^2 + (\mathrm{y} - \mathrm{y}_{0j})^2]^{1/2}} + \alpha_{j2} + \alpha_{j3}\mathrm{cosA}_{ij} \\ \qquad - \alpha_{j4}\mathrm{sinA}_{ij} + \varepsilon_{\mathrm{E}ij} \end{cases} \tag{11}$$

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

Assuming that we get a series of imprecise location data $\vec{\mathrm{p}}_i^* = (\mathrm{x}_i^*, \mathrm{y}_i^*, \mathrm{z}_i^*)$ for the spacecraft M at different sampling times $t_i \quad (i = 1, 2, ...)$, 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 ordinates and measurement data from radars, two functions can be set up as follows:

- $f_j(x, y, z) = tg^{-1}\dfrac{x - x_{0j}}{y - y_{0j}}$

$$g_j(x,y,z) = tg^{-1}\frac{\mathrm{z}-\mathrm{z}_{0\mathrm{j}}}{[(\mathrm{x}-\mathrm{x}_{0\mathrm{j}})^2+(\mathrm{y}-\mathrm{y}_{0\mathrm{j}})^2]^{1/2}}$$

and a design matrix can be defined as follows:

$$\Theta_{ij} = \begin{bmatrix} 1 & 0 & tgE_{ij}\sin A_{ij} & tgE_{ij}\cos A_{ij} & tgE_{ij} & \sec E_{ij} \\ 0 & 1 & \cos A_{ij} & -\sin A_{ij} & 0 & 0 \end{bmatrix}$$

Then we get the following linear model:

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

$$(j = 1,\cdots,m; i = 1,2,\cdots)$$

where $\Delta\tilde{A}_{ij} = A_{ij} - f_j(\vec{P}_i^*)$£¬$\Delta\tilde{E}_{ij} = E_{ij} - g_j(\vec{P}_i^*)$, and

$$J_1(\vec{P}) = \frac{\partial(f_j, g_j)}{\partial(x,y,z)} \quad (j = 1,\cdots,m; i = 1,2,\cdots)$$

Integrating all of the m instruments, we get an integrated error decomposition model as follows:

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

$$(i = 1,2,\cdots)$$

where $\overline{\mathbf{B}}_i = diag\{\Theta_{i1},\cdots,\Theta_{im}\}$.

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

In this case, there are four transit instruments tracking a rocket in space. Selecting the computation parameter $n_0 = 100$ (s), we use the formulae (6) to get the modification values.

Table 1 gives an estimation of the values of the error coefficients at 110 seconds; Table 2 gives the modification values of the rocket trajectories after $n_0 = 100$ (s).

Table 1 Estimation of Error Coefficients

[mrad]

Transit	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 Modification Values of Trajectories

[m]

=100+	Δx_i	Δy_i	Δz_i
1	-0.66.445	0.631514	0.16216E-2
2	-0.763551	0.687524	-0.654640E-2
3	-0.760541	0.677017	-0.127287E-1
4	-0.752472	0.673932	-0.171894E-1
5	-0.793005	0.699980	-0.238274E-1
6	0.835997	0.730210	-0.297603E-1
7	-0.832480	0.731190	-0.366932E-1
8	-0.802443	0.710644	-0.378403E-1
9	-0.739471	0.661947	-0.366333E-1
10	-0.739483	0.660732	-0.387798E-1

These computations, as well as the results given in Table 1 and Table 2, show that the iterative algorithms given in Section 2 not only decrease the time complexity of the computation but also efficiently improve the precision of the trajectory estimates for a rocket in space. What is more, the practical application shows that this new kind of linear model with variant coefficients is reasonable and useful not only in theory but also in different engineering fields.

4 Summary and Conclusions

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

As for our new algorithms, their advantage is evident. Obviously, if we use the old LS-algorithm (5), we must compute a very high dimensional reversion matrix $(\mathbf{H}_n^\tau \mathbf{H}_n)^{-1} \in R^{(r+nq)\times(r+nq)}$, what is more, the dimension of this matrix increases with the number n of samples as the process moves forwards in time. On the other hand, if we use our new iterative algorithm (6), we just need to deal with a series of low dimensional reversion matrices, the highest dimension of which is equal to $\max\{p,q,r\}$. In fact, the iterative algorithm (6) involves only the three reversion matrices $(\mathbf{A}_i^{\tau}\mathbf{A}_i)^{-1} \in R^{q\times q}$, $\mathbf{R}_{n+1}^{-1} \in R^{p\times p}$ and $(\mathbf{L}_n + \mathbf{B}_{n+1}^{\tau}\mathbf{B}_{n+1})^{-1} \in R^{r\times r}$.

Acknowledgements

We gratefully acknowledge partial financial support from the National Nature Science Fund of China (NSFC 90305007), the SI Project (SI-210 -05483) of the Swedish Institute, the Jiangsu Nature Science Fund (BK-06200) and the NSFC-RS Joint Project (NSFC-RS/0510881-207043).

References

Brown D.C. (1964): *The Error Model Best Estimation Trajectory.* -- AD 602799

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

Draper N.R. and Smith H. (1981): *Applied Regression Analysis*. --John Wiley & Sons, Inc.

Eubank R.,Chunfeng H and Maldonado Y.,et al (2004): *Smoothing spline estimation in varying -oefficient models.* --Journal of the Royal Statistical Society B,Vol.66,No.3, pp:653-667

Fahrmeier L. and Tutz G.(2001): *Multivariate Statistical Modeling Based on Generalized Linear Models*.--Springer series in statistics, Berlin: Springer-Verlag

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

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

Hu Shaolin and Sun Guoji (2001): *Process Monitoring Technique and Applications.* Bejing: National Defense Industry Press, pp:68-103

Radoslaw K. and Krzysztof K. (1988): *Recursive Improvement of Estimates in a Gauss-Markov Model with Linear Restrictions*. The Canadian Journal of Statistics, Vol.16, No.3, 301-305

Rencher A. (2000): *Linear Models in Statistics.* --John Wiley & Sons, Inc. pp:121-256

Appendix 1: Proof of Theorem 1

In order to prove Theorem 1, we make use of two lemmas, that we shall state without proof, which are fundamental in linear algebra:

Lemma 1 If a block matrix $\mathbf{A}$ and the block $\mathbf{A}_{11}$ in matrix $\mathbf{A}$, defined as follows, are reversible, then we have

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} \mathbf{A}_{11}^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \mathbf{A}_{11}^{-1}\mathbf{A}_{12} \\ -\mathbf{I} \end{pmatrix} \cdot (\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12})^{-1}(\mathbf{A}_{21}\mathbf{A}_{11}^{-1} \vdots -\mathbf{I}) \quad \text{(a-1)}$$

Similarly, if the matrix A and the block $\mathbf{A}_{22}$ are reversible, then we have

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{A}_{22}^{-1} \end{pmatrix} + \begin{pmatrix} -\mathbf{I} \\ \mathbf{A}_{22}^{-1}\mathbf{A}_{21} \end{pmatrix} \cdot (\mathbf{A}_{11} - \mathbf{A}_{12}\mathbf{A}_{22}^{-1}\mathbf{A}_{21})^{-1}(-\mathbf{I} \vdots \mathbf{A}_{12}\mathbf{A}_{22}^{-1}) \quad \text{(a-2)}$$

Lemma 2 If two matrices $\mathbf{F}$ and $\mathbf{G}$ are reversible and the inverse matrix $(\mathbf{F} - \mathbf{H}\mathbf{G}^{-1}\mathbf{K})^{-1}$ exists, then

$$(\mathbf{F} - \mathbf{H}\mathbf{G}^{-1}\mathbf{K})^{-1} = \mathbf{F}^{-1} + \mathbf{F}^{-1}\mathbf{H}(\mathbf{G} - \mathbf{K}\mathbf{F}^{-1}\mathbf{H})^{-1}\mathbf{K}\mathbf{F}^{-1} \quad \text{(a-3)}$$

The proofs of these two lemmas can be found in reference [5]. Now, let us prove Theorem 1 in detail.

Proof of Theorem 1

With model (2) and n samples, the equation (5) shows that the LS estimators are

$$\hat{\mathbf{\Phi}}_n^{LS(n)} = (\mathbf{H}_n^\tau \mathbf{H}_n)^{-1}\mathbf{H}_n^\tau \overline{\mathbf{Y}}_n$$

If there is another sampling datum

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

which is added into the sampling set, then the LS estimators of all the coefficients in model (2) must be modified according to the following expressions:

$$\hat{\mathbf{\Phi}}_{n+1}^{LS(n+1)} = \left\{\mathbf{\Psi}_n^{\ \tau}\mathbf{\Psi}_n\right\}^{-1}\mathbf{\Psi}_n^{\ \tau}\begin{bmatrix} \overline{\mathbf{Y}}_n \\ \mathbf{Y}_{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{C}_{n+1}^\tau \mathbf{A}_{n+1} \\ \mathbf{A}_{n+1}^\tau \mathbf{C}_{n+1} & \mathbf{A}_{n+1}^\tau \mathbf{A}_{n+1} \end{bmatrix}^{-1}\mathbf{\Psi}_n^{\ \tau}\begin{bmatrix} \overline{\mathbf{Y}}_n \\ \mathbf{Y}_{n+1} \end{bmatrix} \quad \text{(a-4)}$$

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

and $\mathbf{C}_{n+1} = (\mathbf{B}_{n+1}, 0) \in R^{p\times(q+nq)}$£¬$\mathbf{D}_{11} = \mathbf{H}_n^\tau \mathbf{H}_n + \mathbf{C}_{n+1}^\tau \mathbf{C}_{n+1}$

Now, using the notations $\mathbf{D}_{22} = \mathbf{A}_{n+1}^\tau \mathbf{A}_{n+1}$ and $\mathbf{D}_{12} = \mathbf{C}_{n+1}^\tau \mathbf{A}_{n+1}$ as well as $\mathbf{\Omega} = \mathbf{D}_{22} - \mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{D}_{12}$, the following formula (a-5) can be directly derived from lemma 1:

$$\begin{pmatrix} \hat{\mathbf{\Phi}}_n^{LS(n+1)} \\ \hat{\mathbf{X}}_{n+1}^{LS(n+1)} \end{pmatrix} = \left\{\begin{pmatrix} \mathbf{D}_{11}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{D}_{11}^{-1}\mathbf{D}_{12} \\ -\mathbf{I} \end{pmatrix}\mathbf{\Omega}^{-1}\left(\mathbf{D}_{21}\mathbf{D}_{11}^{-1} \vdots -\mathbf{I}\right)\right\} \cdot \begin{pmatrix} \mathbf{H}_n^\tau \overline{\mathbf{Y}}_n + \mathbf{C}_{n+1}^\tau \mathbf{Y}_{n+1} \\ \mathbf{A}_{n+1}^\tau \mathbf{Y}_{n+1} \end{pmatrix} = \mathbf{E}_n \mathbf{H}_n^\tau \overline{\mathbf{Y}}_n + \mathbf{F}_n \mathbf{Y}_{n+1} \quad \text{(a-5)}$$

where

$$\mathbf{F}_n = \begin{pmatrix} \mathbf{D}_{11}^{-1}[\mathbf{C}_{n+1}^\tau + \mathbf{D}_{12}\mathbf{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^\tau - \mathbf{D}_{12}\mathbf{\Omega}^{-1}\mathbf{A}_{n+1}^\tau] \\ -\mathbf{\Omega}^{-1}[\mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^\tau - \mathbf{A}_{n+1}^\tau] \end{pmatrix}$$

and $\mathbf{E}_n = \begin{pmatrix} \mathbf{D}_{11}^{-1} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1} \\ -\mathbf{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1} \end{pmatrix}$.

[Step 1] We first analyze the expression $\mathbf{E}_n$. From the expression for the block matrix $\mathbf{D}_{11}$ and Lemma 1 we have

$$\mathbf{D}_{11}^{-1} = \begin{bmatrix} \sum_{i=1}^{n+1}\mathbf{B}_i^\tau \mathbf{B}_i & \mathbf{B}_1^\tau \mathbf{A}_1 & \cdots & \mathbf{B}_n^\tau \mathbf{A}_n \\ \mathbf{A}_1^\tau \mathbf{B}_1 & \mathbf{A}_1^\tau \mathbf{A}_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{A}_n^\tau \mathbf{B}_n & 0 & \cdots & \mathbf{A}_n^\tau \mathbf{A}_n \end{bmatrix}^{-1}$$

It can be shown that the first r rows of the matrix $\mathbf{D}_{11}^{-1}$ can be expressed as follows

$$\mathbf{D}_{11}^{-1} = \begin{bmatrix} \left(\sum_{i=1}^{n+1}\mathbf{B}_i^\tau \mathbf{B}_i - \sum_{i=1}^{n+1}\mathbf{B}_i^\tau \mathbf{U}_i \mathbf{A}_i^\tau \mathbf{B}_i\right)^{-1} & * \\ -\mathbf{T}_n \cdot \left(\sum_{i=1}^{n+1}\mathbf{B}_i^\tau \mathbf{B}_i - \sum_{i=1}^{n+1}\mathbf{B}_i^\tau \mathbf{U}_i \mathbf{A}_i^\tau \mathbf{B}_i\right)^{-1} & * \end{bmatrix} \quad \text{(a-6)}$$

where $\mathbf{U}_i = \mathbf{A}_i(\mathbf{A}_i^\tau \mathbf{A}_i)^{-1}$£¬$\mathbf{T}_n = [\mathbf{B}_1^\tau \mathbf{U}_1, \cdots, \mathbf{B}_n^\tau \mathbf{U}_n]^\tau$ and the asterisk "*" denotes an omitted matrix block which is very complicated and has no effect on the following deduction process.

After analyzing the formulae for the matrix block $\mathbf{D}_{22}$ and matrix block $\mathbf{D}_{12} = \mathbf{D}_{21}$, we have the following equation

$$\mathbf{D}_{11}^{-1} + \mathbf{D}_{11}^{-1}\mathbf{D}_{12}\mathbf{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1} = \{\mathbf{I} + \mathbf{D}_{11}^{-1}\mathbf{C}_n^\tau \mathbf{V}_{n+1}\mathbf{C}_{n+1}\}\mathbf{D}_{11}^{-1}$$

(a-7)

where

$$\mathbf{V}_{n+1}=\mathbf{A}_{n+1}(\mathbf{A}^{\tau}_{n+1}\mathbf{A}_{n+1}-\mathbf{A}^{\tau}_{n+1}\mathbf{C}_{n+1}\mathbf{D}^{-1}_{11}\mathbf{C}^{\tau}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}^{\tau}_{n+1}$$

and the equation

$$\mathbf{D}^{-1}_{11}\mathbf{C}^{\tau}_{n}\mathbf{V}_{n+1}\mathbf{C}_{n+1} = \mathbf{D}^{-1}_{11}\begin{pmatrix}\mathbf{B}^{\tau}_{n+1}\mathbf{A}_{n+1}\mathbf{W}^{-1}_{k+1}\mathbf{A}^{\tau}_{n+1}\mathbf{B}_{n+1} & \mathbf{0}\\ \mathbf{0} & \mathbf{0}\end{pmatrix} \quad \text{(a-8)}$$

$$=\begin{pmatrix}\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{B}^{\tau}_{n+1}\mathbf{A}_{n+1}\mathbf{W}^{-1}_{k+1}\mathbf{A}^{\tau}_{n+1}\mathbf{B}_{n+1} & \mathbf{0}\\ -\mathbf{T}_n\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{B}^{\tau}_{n+1}\mathbf{A}_{n+1}\mathbf{W}^{-1}_{k+1}\mathbf{A}^{\tau}_{n+1}B_{n+1} & \mathbf{0}\end{pmatrix}$$

holds, where

- $\mathbf{W}_{k+1}=\mathbf{A}^{\tau}_{n+1}(\mathbf{I}\text{-}\mathbf{B}_{n+1}\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{B}^{\tau}_{n+1})\mathbf{A}_{n+1}$
- $\tilde{\mathbf{L}}_{n+1}=\sum_{i=1}^{n}\mathbf{B}^{\tau}_{i}[\mathbf{I}\text{-}\mathbf{U}_i\mathbf{A}^{\tau}_{i}]\mathbf{B}_i+\mathbf{B}^{\tau}_{n+1}\mathbf{B}_{n+1}$

Obviously, the matrix $\mathbf{D}^{-1}_{11}$ can be expressed as follows:

$$\mathbf{D}^{-1}_{11}=(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}-(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}\mathbf{C}^{\tau}_{n+1}\cdot[\mathbf{I}+\mathbf{C}^{\tau}_{n+1}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}\mathbf{C}^{\tau}_{n+1}]^{-1}\mathbf{C}_{n+1}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1} \quad \text{(a-9)}$$

Furthermore, using the notation $\mathbf{L}_n=\sum_{i=1}^{n}\mathbf{B}^{\tau}_{i}[\mathbf{I}\text{-}\mathbf{A}_i(\mathbf{A}^{\tau}_{i}\mathbf{A}_i)^{-1}\mathbf{A}^{\tau}_{i}]\mathbf{B}_i$, we have

$$(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}=\begin{bmatrix}\sum_{i=1}^{n}\mathbf{B}^{\tau}_{1}\mathbf{B}_i & \mathbf{B}^{\tau}_{1}\mathbf{A}_1 & \cdots & \mathbf{B}^{\tau}_{n}\mathbf{A}_n\\ \mathbf{A}^{\tau}_{1}\mathbf{B}_1 & \mathbf{A}^{\tau}_{1}\mathbf{A}_1 & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{A}^{\tau}_{n}\mathbf{B}_n & \mathbf{0} & \cdots & \mathbf{A}^{\tau}_{n}\mathbf{A}_n\end{bmatrix}^{-1}$$

$$=\begin{pmatrix}\mathbf{L}^{-1}_{n} & -\mathbf{L}^{-1}_{n}\mathbf{T}^{\tau}_{n}\\ -\mathbf{T}_n\mathbf{L}^{-1}_{n} & *\end{pmatrix} \quad \text{(a-10)}$$

so we get

$$\mathbf{D}^{-1}_{11}=\left\{\mathbf{I}-\begin{pmatrix}\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1} & \mathbf{0}\\ -\mathbf{T}_n\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1} & \mathbf{0}\end{pmatrix}\right\}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1} \quad \text{(a-11)}$$

where $\tilde{\mathbf{Z}}_{n+1}=\mathbf{B}^{\tau}_{n+1}(\mathbf{I}+\mathbf{B}_{n+1}\mathbf{L}^{-1}_{n}\mathbf{B}^{\tau}_{n+1})^{-1}\mathbf{B}_{n+1}$.

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

$$\mathbf{\Omega}^{-1}\mathbf{D}_{21}=\{\mathbf{A}^{\tau}_{n+1}(\mathbf{I}-\mathbf{C}_{n+1}\mathbf{D}^{-1}_{11}\mathbf{C}^{\tau}_{n+1})\mathbf{A}_{n+1}\}^{-1}\mathbf{A}^{\tau}_{n+1}\mathbf{C}_{n+1}$$
$$=\{\mathbf{A}^{\tau}_{n+1}(\mathbf{I}-\mathbf{B}_{n+1}\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{B}^{\tau}_{n+1})\mathbf{A}_{n+1})^{-1}\cdot\mathbf{A}^{\tau}_{n+1}(\mathbf{B}_{n+1},\mathbf{0}\cdots\mathbf{0}) \quad \text{(a-12)}$$

Letting $\mathbf{Q}_{n+1}=\{\mathbf{A}^{\tau}_{n+1}(\mathbf{I}-\mathbf{B}_{n+1}\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{B}^{\tau}_{n+1})\mathbf{A}_{n+1}\}^{-1}\mathbf{A}^{\tau}_{n+1}\mathbf{B}_{n+1}$ and $\mathbf{P}_{n+1}=\mathbf{B}^{\tau}_{n+1}\mathbf{A}_{n+1}\mathbf{Q}_{\text{n+1}}$ then the following formula can be obtained:

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

(a-13)

and

$$\mathbf{D}^{-1}_{11}+\mathbf{D}^{-1}_{11}\mathbf{D}_{12}\mathbf{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}^{-1}_{11}$$
$$=\left\{\mathbf{I}+\begin{pmatrix}\mathbf{L}^{-1}_{n+1}\mathbf{P}_{n+1} & 0\\ \mathbf{T}_n\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{P}_{n+1} & 0\end{pmatrix}\right\}\cdot\left\{\mathbf{I}-\begin{pmatrix}\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1} & \mathbf{0}\\ -\mathbf{T}_n\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1} & \mathbf{0}\end{pmatrix}\right\}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}$$
$$=\begin{pmatrix}\mathbf{I}+\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{P}_{n+1}\text{-}(\mathbf{I}+\tilde{\mathbf{L}}_{n+1}\mathbf{P}_{n+1})\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1} & \mathbf{0}\\ -\mathbf{T}_n\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{P}_{n+1}+\mathbf{T}_n(\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{P}_{n+1}+\mathbf{I})\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1} & \mathbf{I}\end{pmatrix}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}$$

(a-14)

Combining the equations (a-5), (a-13) and (a-14), we get an expression for the matrix $\mathbf{E}_n$ as follows:

$$\mathbf{E}_n=\begin{pmatrix}\mathbf{I}+\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{P}_{n+1}(\mathbf{I}-\tilde{\mathbf{E}}_n)-\tilde{\mathbf{E}}_n & 0\\ -\mathbf{T}_n[\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{P}_{n+1}(\mathbf{I}-\tilde{\mathbf{E}}_n)-\tilde{\mathbf{E}}_n] & \mathbf{I}\\ -\mathbf{Q}_{n+1}(\mathbf{I}-\tilde{\mathbf{E}}_{(n)}) & \mathbf{0}\end{pmatrix}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1}$$

(a-15)

where $\tilde{\mathbf{E}}_n=\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1}$. Using the following formula

$$\tilde{\mathbf{L}}^{-1}_{n+1}=(\mathbf{L}_n+\mathbf{B}^{\tau}_{n+1}\mathbf{B}_{n+1})^{-1}=\mathbf{L}^{-1}_{n}-\mathbf{L}^{-1}_{n}\tilde{\mathbf{Z}}_{n+1}\mathbf{L}^{-1}_{n}$$

we get

$$\mathbf{I}-\tilde{\mathbf{E}}_n=\mathbf{L}^{-1}_{n+1}\mathbf{L}_n \quad \text{(a-16)}$$

So, the equation (25) can be expressed as

$$\mathbf{E}_n=\begin{pmatrix}\mathbf{I}+\mathbf{L}^{p}_{n+1}+\tilde{\mathbf{L}}^{-1}_{n+1}[\mathbf{L}_n-\tilde{\mathbf{L}}^{-1}_{n+1}] & \mathbf{0}\\ -\mathbf{T}_n[\mathbf{L}^{p}_{n+1}+\tilde{\mathbf{L}}^{-1}_{n+1}(\mathbf{L}_n-\tilde{\mathbf{L}}^{-1}_{n+1})] & \mathbf{I}\\ -\mathbf{Q}_{n+1}\tilde{\mathbf{L}}^{-1}_{n+1}\mathbf{L}_n & \mathbf{0}\end{pmatrix}(\mathbf{H}^{\tau}_{n}\mathbf{H}_n)^{-1} \quad \text{(a-17)}$$

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

[Step 2] We next analyze the expression for $\mathbf{F}_n$. From equation (a-6), we have

$$\mathbf{D}_{12}\mathbf{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}^{-1}_{11}\mathbf{C}^{\tau}_{n+1}=\mathbf{C}^{\tau}_{n+1}\mathbf{A}_{n+1}\mathbf{\Omega}^{-1}\mathbf{A}^{\tau}_{n+1}\mathbf{C}_{n+1}\mathbf{D}^{-1}_{11}\mathbf{C}^{\tau}_{n+1}$$

namely,

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

(a-18)

So we have the following four equations:

- $\mathbf{D}_{11}^{-1}[\mathbf{D}_{12}\boldsymbol{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^{\tau}]$

$$=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1} & -\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{T}_n^{\tau}\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1} & *\end{pmatrix}\begin{pmatrix}\mathbf{M}_{n+1}\\ 0\end{pmatrix}=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{M}_{n+1}\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{M}_{n+1}\end{pmatrix} \quad \text{(a-19)}$$

- $$\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^{\tau}=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1} & -\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{T}_n^{\tau}\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1} & *\end{pmatrix}\begin{pmatrix}\mathbf{B}_{n+1}^{\tau}\\ 0\end{pmatrix}=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\end{pmatrix} \quad \text{(a-20)}$$

- $$\mathbf{D}_{11}^{-1}\mathbf{D}_{12}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau}=\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^{\tau}\mathbf{A}_{n+1}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau}=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau}\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau}\end{pmatrix} \quad \text{(a-21)}$$

- $$\boldsymbol{\Omega}^{-1}\mathbf{D}_{21}\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^{\tau}=\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau} \quad \text{(a-22)}$$

where $\mathbf{M}_{n+1}=\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}$.

Inserting equations (a-19), … , (a-22) into the expression for $\mathbf{F}_n$ which was defined in equation (a-5), we immediately have that

$$\mathbf{F}_n=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}(\mathbf{B}_{n+1}^{\tau}+\mathbf{M}_{n+1}-\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau})\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1}(\mathbf{B}_{n+1}^{\tau}+\mathbf{M}_{n+1}-\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}\boldsymbol{\Omega}^{-1}\mathbf{A}_{n+1}^{\tau})\\ -\mathbf{Q}^{-1}(\mathbf{A}_{n+1}^{\tau}\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}-\mathbf{A}_{n+1}^{\tau})\end{pmatrix} \quad \text{(a-23)}$$

And, after considering the definition of $\boldsymbol{\Omega}$ and equation (a-16), we have

$$\begin{aligned}\boldsymbol{\Omega}&=\mathbf{A}_{n+1}^{\tau}\mathbf{A}_{n+1}-\mathbf{A}_{n+1}^{\tau}\mathbf{C}_{n+1}\mathbf{D}_{11}^{-1}\mathbf{C}_{n+1}^{\tau}\mathbf{A}_{n+1}\\ &=\mathbf{A}_{n+1}^{\tau}\mathbf{A}_{n+1}-\mathbf{A}_{n+1}^{\tau}\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}\\ &=\mathbf{A}_{n+1}^{\tau}(\mathbf{I}-\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau})\mathbf{A}_{n+1}\end{aligned}$$

so

$$\boldsymbol{\Omega}^{-1}=[\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1}]^{-1} \quad \text{(a-24)}$$

where $\mathbf{R}_{n+1}=\mathbf{I}-\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}$. Inserting equations (a-18) and (a-24) into equation (a-23), we have

$$\mathbf{F}_n=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}[\mathbf{I}-\mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}]\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}[\mathbf{I}-\mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}]\\ (\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\end{pmatrix} \quad \text{(a-25)}$$

[Step 3] From step 1 and step 2, substituting equations (a-18) and (a-26) into the equation (a-6), we immediately get the following expression

$$\begin{pmatrix}\hat{\boldsymbol{\Phi}}_n^{LS(n+1)}\\ \hat{\mathbf{X}}_{n+1}^{LS(n+1)}\end{pmatrix}=\mathbf{M}_{n+1}^{a}\hat{\boldsymbol{\Phi}}_n^{LS(n)}+\mathbf{M}_{n+1}^{b}\mathbf{Y}_{n+1} \quad \text{(a-26)}$$

where $\mathbf{M}_{n+1}^{a}=\begin{pmatrix}\mathbf{I}+\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{P}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n-\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\mathbf{B}_{n+1} & \mathbf{0}\\ -\mathbf{T}_n(\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{P}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n-\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\mathbf{B}_{n+1}) & \mathbf{I}\\ -\mathbf{Q}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n & \mathbf{0}\end{pmatrix}$,

and

$$\mathbf{M}_{n+1}^{b}=\begin{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}[\mathbf{I}-\mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}]\\ -\mathbf{T}_n\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}[\mathbf{I}-\mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}]\\ (\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\end{pmatrix}$$

If we let $\boldsymbol{\Xi}_{n+1}=\mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}$, then we have

$$\begin{aligned}&\mathbf{P}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)}-\mathbf{B}_{n+1}^{\tau}\mathbf{A}_{n+1}(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{Y}_{n+1}\\ &=\mathbf{B}_{n+1}^{\tau}\boldsymbol{\Xi}_{n+1}\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)}-\mathbf{B}_{n+1}^{\tau}\boldsymbol{\Xi}_{n+1}\mathbf{R}_{n+1}\mathbf{Y}_{n+1}\\ &=\mathbf{B}_{n+1}^{\tau}\boldsymbol{\Xi}_{n+1}[\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)}-\mathbf{R}_{n+1}\mathbf{Y}_{n+1}]\end{aligned}$$

and

$$\begin{aligned}&\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)}-\mathbf{R}_{n+1}\mathbf{Y}_{n+1}\\ &=\mathbf{B}_{n+1}(\mathbf{I}-\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\mathbf{B}_{n+1})\hat{\boldsymbol{\beta}}^{LS(n)}-(\mathbf{I}-\mathbf{B}_{n+1}^{\tau}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1})\mathbf{Y}_{n+1}\\ &=(\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}-\mathbf{I})(\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)})\\ &=-\mathbf{R}_{n+1}(\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)})\end{aligned}$$

So, decomposing the matrix equation (a-26) into some appropriate blocks, we get

$$\begin{aligned}\hat{\boldsymbol{\beta}}^{LS(n+1)}&=\hat{\boldsymbol{\beta}}^{LS(n)}+\tilde{\mathbf{L}}_{n+1}^{-1}[\mathbf{P}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)}-\mathbf{B}_{n+1}^{\tau}\boldsymbol{\Xi}_{n+1}\mathbf{R}_{n+1}\mathbf{Y}_{n+1}]\\ &\quad+\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}(\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)})\\ &=\hat{\boldsymbol{\beta}}^{LS(n)}+[\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}-\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\boldsymbol{\Xi}_{n+1}\mathbf{R}_{n+1}](\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)})\\ &=\hat{\boldsymbol{\beta}}^{LS(n)}+\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}[\mathbf{R}_{n+1}^{-1}-\boldsymbol{\Xi}_{n+1}]\mathbf{R}_{n+1}(\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)})\end{aligned} \quad \text{(a-27)}$$

and

$$\begin{pmatrix}\hat{\mathbf{X}}_1^{LS(n+1)}\\ \vdots\\ \hat{\mathbf{X}}_n^{LS(n+1)}\end{pmatrix}=\begin{pmatrix}\hat{\mathbf{X}}_1^{LS(n)}\\ \vdots\\ \hat{\mathbf{X}}_n^{LS(n)}\end{pmatrix}-\begin{pmatrix}(\mathbf{A}_1^{\tau}\mathbf{A}_1)^{-1}\mathbf{A}_1^{\tau}\mathbf{B}_1\\ \vdots\\ (\mathbf{A}_n^{\tau}\mathbf{A}_n)^{-1}\mathbf{A}_n^{\tau}\mathbf{B}_n\end{pmatrix}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{B}_{n+1}^{\tau}\cdot\left(\mathbf{R}_{n+1}^{-1}-\boldsymbol{\Xi}_{n+1}\right)\mathbf{R}_{n+1}(\mathbf{Y}_n-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}) \quad \text{(a-28)}$$

From the formula (a-28), for $i=1,2,\cdots,n$, we have

$$\hat{\mathbf{X}}_i^{LS(n+1)}=\hat{\mathbf{X}}_i^{LS(n)}+(\mathbf{A}_i^{\tau}\mathbf{A}_i)^{-1}\mathbf{A}_i^{\tau}\mathbf{B}_i(\hat{\boldsymbol{\beta}}^{LS(n)}-\hat{\boldsymbol{\beta}}^{LS(n+1)}) \quad \text{(a-29)}$$

and

$$\begin{aligned}\hat{\mathbf{X}}_{n+1}^{LS(n+1)}&=-\mathbf{Q}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)}+(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1})^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1}\\ &=\left(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1}\right)^{-1}\mathbf{A}_{n+1}^{\tau}(\mathbf{R}_{n+1}\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\tilde{\mathbf{L}}_{n+1}^{-1}\mathbf{L}_n\hat{\boldsymbol{\beta}}^{LS(n)})\\ &=\left(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1}\right)^{-1}\mathbf{A}_{n+1}^{\tau}\left(\mathbf{R}_{n+1}\mathbf{Y}_{n+1}-\mathbf{R}_{n+1}\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}\right)\\ &=\left(\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\mathbf{A}_{n+1}\right)^{-1}\mathbf{A}_{n+1}^{\tau}\mathbf{R}_{n+1}\left(\mathbf{Y}_{n+1}-\mathbf{B}_{n+1}\hat{\boldsymbol{\beta}}^{LS(n)}\right)\end{aligned} \quad \text{(a-30)}$$

Combining the equations (a-27) ,…, (a-30), the result of Theorem 1 is obtained.