Journal of Algebra, Number Theory: Advances and Applications
Volume 8, Number 1-2, 2012, Pages 19-39

THE LEAST SQUARES STOCHASTIC SOLUTIONS OF THE MATRIX EQUATION $AX = B$

FANGYING LI, JINGJING PENG and ZHENYUN PENG

Guilin University of Electronic Technology
Guilin 541004
P. R. China
e-mail: yunzhenp@163.com

Abstract

In this paper, we present an iteration method to compute the least squares stochastic solutions of the matrix equation $AX = B$. Numerical experiments are given to illustrate the usefulness of the proposed approach.

1. Introduction

Markov chains theory are widely used in the economic activities forecasting, queuing theory, and particle technology [2, 4, 6, 7]. The key problem of using Markov chains to predict the future state of a system is to compute the transition probability matrix. The transition probability matrix X may be obtained by solving the state matrix equations $AX = B$ with unknown matrix X satisfies $Xe = e$ and $X \geq 0$, where e is a vector

2010 Mathematics Subject Classification: 15A24, 15A39, 65F30.

Keywords and phrases: nonnegative matrix, iterative method, matrix equation, least squares problem.

Research supported by National Natural Science Foundation of China (10861005) and Provincial Natural Science Foundation of Guangxi (11261014, 11101100).

Received March 21, 2012

of all ones. However, the state matrices A and B, in general, are obtained from market statistics or experimental analysis, and it may be not satisfies the above matrix equation. In this case, one hopes to find the smallest correction stochastic solution X of the inconsistent matrix equation $AX = B$. This leads to consider the following constrained least squares problem:

$$\begin{aligned} &\text{minimize} \quad f(X) = \frac{1}{2}\|AX - B\|^2, \\ &\text{subject to} \quad Xe = e,\ X \geq 0, \end{aligned} \tag{1.1}$$

where $A, B \in R^{m\times n}$, $e = (1, 1, \dots, 1)^T \in R^{n\times n}$. The problem (1.1) is always solvable, and when the matrix A has full column rank, then the problem (1.1) has an unique solution for any matrix B. In this paper, we present an iteration method to compute the solutions of the problem (1.1). Numerical experiments are given to illustrate the usefulness of the proposed approach.

Throughout this paper, $R^{m\times n}$ denotes the set of $m \times n$ real matrices. A^T, $\|A\|$, and $tr(A)$ denote the transpose, the trace, and the Frobenius norm of the matrix A, respectively. For the matrices $A = (a_{ij})$, $B = (b_{ij})$, $A \otimes B$ denotes the Kronecker product defined as $A \otimes B = (a_{ij}B)$, $A \bullet B$ denotes the Hadamard product defined as $A \bullet B = (a_{ij}b_{ij})$, and the inequality $A \geq B\,(A > B)$ means that $a_{ij} \geq b_{ij}\,(a_{ij} > b_{ij})$ for all i and j. Defining the inner product in space $R^{m\times n}$ by

$$\langle A, B\rangle = tr(B^T A), \quad \forall A, B \in R^{m\times n}.$$

Obviously, $R^{m\times n}$ is a Hilbert inner product space and the norm of a matrix generated by this inner product space is the Frobenius norm.

2. Iterative Method to Solve Problem (1.1)

Note that the solution X^* of the problem (1.1) must satisfy the optimality Karush-Kuhn-Tucker (KKT) conditions

$$\begin{cases} A^T AX - A^T B - Ye^T - Z = 0, \\ Xe = e, \\ X \bullet Z = 0, \\ X \geq 0,\ Z \geq 0. \end{cases} \tag{2.1}$$

Hence, solving the problem (1.1) is equivalent to solve the optimality KKT conditions matrix equations (2.1) for X, Y, and Z. We use the predictor-corrector interior-point method (research results for this method, see references [1, 3, 5, 8-18]) to solve (2.1). The predictor step, we first solve the following matrix equations to obtain the affine scaling search direction $(\Delta X^a, \Delta Y^a, \Delta Z^a)$ for the current iterate (X_k, Y_k, Z_k):

$$\begin{cases} A^T A\Delta X^a - \Delta Y^a e^T - \Delta Z^a = A^T B - A^T AX_k + Y_k e^T + Z_k, \\ \Delta X^a e = e - X_k e, \\ Z_k \bullet \Delta X^a + X_k \bullet \Delta Z^a = -X_k \bullet Z_k, \end{cases} \tag{2.2}$$

and then we compute the maximum possible step length α_a by satisfying the following matrix equations:

$$X_k + \alpha_a \Delta X^a \geq 0, \quad Z_k + \alpha_a \Delta Z^a \geq 0. \tag{2.3}$$

The corrector step, we first compute the corrector search direction $(\Delta X, \Delta Y, \Delta Z)$ by solving the following matrix equations:

$$\begin{cases} A^T A\Delta X - \Delta Ye^T - \Delta Z = A^T B - A^T AX_k + Y_k e^T + Z_k, \\ \Delta Xe = e - X_k e, \\ Z_k \bullet \Delta X + X_k \bullet \Delta Z = \mu ee^T - X_k \bullet Z_k - \Delta X^a \bullet \Delta Z^a, \end{cases} \tag{2.4}$$

with

$$\mu = \left(\frac{\eta_a}{\eta}\right)^2 \frac{\eta_a}{n^2},$$

where $\eta_a = tr((X_k + \alpha_a \Delta X^a)^T (Z_k + \alpha_a \Delta Z^a))$ and $\eta = tr(X_k^T Z_k)$, then we choose the maximum possible step length α by satisfying the following matrix equations:

$$X_k + \alpha \Delta X \geq 0, \quad Z_k + \alpha \Delta Z \geq 0. \tag{2.5}$$

According to above discussions, the predictor-corrector interior-point method to solve constrained least squares problem (1.1) can be described as follows.

Algorithm 2.1. (Computing the solution of constrained least squares problem (1.1)):

(1) Input matrices A, B, e. Given initial matrices $X_0 > 0$, $Z_0 > 0$, Y_0, and tolerances $\varepsilon_1, \varepsilon_2, \varepsilon_3 \geq 0$.

(2) For $k = 1, 2, \ldots,$ until $\|A^T B - A^T A X_k + Y_k e^T + Z_k\| \leq \varepsilon_1$, $\|e - X_k e\| \leq \varepsilon_2$, and $\|X_k \bullet Z_k\| \leq \varepsilon_3$.

(i) Predictor step:

• Computing the predictor direction $(\Delta X^a, \Delta Y^a, \Delta Z^a)$ by solving the matrix equations (2.2).

• Computing the maximum possible step length α_a by solving the matrix inequalities (2.3).

(ii) Corrector step:

• Computing $\eta_a = tr((X_k + \alpha_a \Delta X^a)^T (Z_k + \alpha_a \Delta Z^a))$, $\eta = tr(X_k^T Z_k)$, and $\mu = \left(\frac{\eta_a}{\eta}\right)^2 \frac{\eta_a}{n^2}$.

• Computing the predictor-corrector direction $(\Delta X, \Delta Y, \Delta Z)$ by solving the matrix equations (2.4).

• Computing the maximum possible step length α by solving the matrix inequalities (2.5).

(iii) Update:

$$X_{k+1} = X_k + \alpha\Delta X, \quad Y_{k+1} = Y_k + \alpha\Delta Y, \quad Z_{k+1} = Z_k + \alpha\Delta Z.$$

In the implement of the Algorithm 2.1, the maximum possible step length α_a in (2.3) and α in (2.5) are usually chosen as

$$\alpha_a = \min\{1,\ c\cdot \min_{i,j:\Delta X_{ij}^a<0} \frac{-(X_k)_{ij}}{\Delta X_{ij}^a},\ c\cdot \min_{i,j:\Delta Z_{ij}^a<0} \frac{-(Z_k)_{ij}}{\Delta Z_{ij}^a}\},$$

$$\alpha = \min\{1,\ c\cdot \min_{i,j:\Delta X_{ij}<0} \frac{-(X_k)_{ij}}{\Delta X_{ij}},\ c\cdot \min_{i,j:\Delta Z_{ij}<0} \frac{-(Z_k)_{ij}}{\Delta Z_{ij}}\},$$

for some $c \in (0, 1)$, which is called as the step length parameter and, in practice, it is usually chosen as a fixed number from the interval (0.9, 1.0). To do this way is to avoid take a step all the way to the boundary.

The calculation of the search directions, that is solving linear matrix equations (2.2) and (2.4), are the most time-consuming and space-occupying steps in the implement of the Algorithm 2.1. In this paper, we propose the following iteration method to compute the search directions.

Algorithm 2.2. (Computing the solution of the matrix equations (2.2) and (2.4)):

(1) Input A, B, e, the k-th approximate X_k, Z_k, Y_k of Algorithm 2.1. Given initial matrices $\Delta X_0, \Delta Y_0, \Delta Z_0$, and a tolerance $\varepsilon > 0$.

Computing $F = -X_k \bullet Z_k$ when solving the matrix equations (2.2) and $F = \mu ee^T - X_k \bullet Z_k - \Delta X^a \bullet \Delta Z^a$ when solving the matrix equations (2.4). Computing

$$\begin{cases} R_{1,0} = (A^T B - A^T A X_k + Y_k e^T + Z_k) - (A^T A \Delta X_0 - \Delta Y_0 e^T - \Delta Z_0), \\ R_{2,0} = (e - X_k e) - \Delta X_0 e, \\ R_{3,0} = F - (Z_k \bullet \Delta X_0 + X_k \bullet \Delta Z_0), \end{cases}$$

$$\alpha_0 = \|R_{1,0}\|^2 + \|R_{2,0}\|^2 + \|R_{3,0}\|^2;$$

$$\begin{cases} P_{1,0} = A^T A R_{1,0} + R_{2,0} e^T + Z_k \bullet R_{3,0}, \\ P_{2,0} = -R_{1,0} e, \\ P_{3,0} = -R_{1,0} + X_k \bullet R_{3,0}, \end{cases}$$

$$\beta_0 = \|P_{1,0}\|^2 + \|P_{2,0}\|^2 + \|P_{3,0}\|^2.$$

(2) For $i = 0, 1, 2, \ldots,$ until $\beta_i \le \epsilon$

$$\begin{cases} \Delta X_{i+1} = \Delta X_i + \frac{\alpha_i}{\beta_i} P_{1,i}, \\ \Delta Y_{i+1} = \Delta Y_i + \frac{\alpha_i}{\beta_i} P_{2,i}, \\ \Delta Z_{i+1} = \Delta Z_i + \frac{\alpha_i}{\beta_i} P_{3,i}, \end{cases}$$

$$\begin{cases} R_{1,i+1} = (A^T B - A^T A X_k + Y_k e^T + Z_k) - (A^T A \Delta X_{i+1} - \Delta Y_{i+1} e^T - \Delta Z_{i+1}) \\ \qquad = R_{1,i} - \frac{\alpha_i}{\beta_i} (A^T A P_{1,i} - P_{2,i} e^T - P_{3,i}), \\ R_{2,i+1} = (e - X_k e) - \Delta X_{i+1} e \\ \qquad = R_{2,i} - \frac{\alpha_i}{\beta_i} P_{1,i} e, \\ R_{3,i+1} = F - (Z_k \bullet \Delta X_{i+1} + X_k \bullet \Delta Z_{i+1}) \\ \qquad = R_{3,i} - \frac{\alpha_i}{\beta_i} (Z_k \bullet P_{1,i} + X_k \bullet P_{3,i}), \end{cases}$$

$$\alpha_{i+1} = \|R_{1,i+1}\|^2 + \|R_{2,i+1}\|^2 + \|R_{3,i+1}\|^2;$$

$$\begin{cases} P_{1,i+1} = A^T A R_{1,i+1} + R_{2,i+1} e^T + Z_k \bullet R_{3,i+1} + \frac{\alpha_{i+1}}{\alpha_i} P_{1,i}, \\ P_{2,i+1} = -R_{1,i+1} e + \frac{\alpha_{i+1}}{\alpha_i} P_{2,i}, \\ P_{3,i+1} = -R_{1,i+1} + X_k \bullet R_{3,i+1} + \frac{\alpha_{i+1}}{\alpha_i} P_{3,i}, \end{cases}$$

$$\beta_{i+1} = \|P_{1,i+1}\|^2 + \|P_{2,i+1}\|^2 + \|P_{3,i+1}\|^2.$$

For the Algorithm 2.2, we have the following propositions:

Proposition 2.1. *For the sequences* $\{R_{1,i}\}, \{R_{2,i}\}, \{R_{3,i}\}, \{P_{1,i}\}, \{P_{2,i}\}$, *and* $\{P_{3,i}\}$ *generated by Algorithm* 2.2, *if there exists a positive number* k *such that* $\beta_i \neq 0$ *for all* $i = 0, 1, 2, \ldots, k$, *then we have* $\alpha_i = \|R_{1,i}\|^2 + \|R_{2,i}\|^2 + \|R_{3,i}\|^2 \neq 0$ *for all* $i = 0, 1, 2, \ldots, k$, *and the following two equalities hold*:

$$tr(P_{1,i}^T P_{1,j}) + tr(P_{2,i}^T P_{2,j}) + tr(P_{3,i}^T P_{3,j}) = 0, \ (i, j = 0, 1, 2, \cdots, k, i \neq j), \tag{2.6}$$

$$tr(R_{1,i}^T R_{1,j}) + tr(R_{2,i}^T R_{2,j}) + tr(R_{3,i}^T R_{3,j}) = 0, \ (i, j = 0, 1, 2, \cdots, k, i \neq j). \tag{2.7}$$

Proposition 2.2. *Suppose* $(\Delta\overline{X}, \Delta\overline{Y}, \Delta\overline{Z})$ *be arbitrary solution of the matrix equations* (2.2) *or* (2.4), *then we have*

$$tr[(\Delta\overline{X} - \Delta X_i)^T P_{1,i}] + tr[(\Delta\overline{Y} - \Delta Y_i)^T P_{2,i}] + tr[(\Delta\overline{Z} - \Delta Z_i)^T P_{3,i}]$$

$$= \|R_{1,i}\|^2 + \|R_{2,i}\|^2 + \|R_{3,i}\|^2, \quad i = 0, 1, 2, \ldots. \tag{2.8}$$

Proposition 2.3. *Algorithm* 2.2 *breaks down within finite iteration steps in the absence of roundoff errors. Furthermore, if Algorithm* 2.2 *breaks down at i-th iteration step, then, when* $\beta_i = 0$ *and* $\alpha_i = 0$, *the* *matrix equations* (2.2) (*or* (2.4)) *is solvable and* $(\Delta X_i, \Delta Y_i, \Delta Z_i)$ *is its a solution. When* $\beta_i = 0$ *and* $\alpha_i \neq 0$, *the matrix equations* (2.2) (*or* (2.4)) *has no solution.*

The proof of Propositions 2.1, 2.2, and 2.3 is given in the Appendix.

3. Numerical Examples

In this section, we give a numerical example to illustrate the application of the problem (1.1) and the efficiency of the methods proposed in this paper. Our computational experiments are done on a HP Compaq Presario CQ45-203TX with 2.0GHz and 2.0 ram. All the tests were performed by MATLAB 7.0, which runs on the operating system windows XP professional. In Algorithm 2.2, the initial iterative matrices ΔX_0, ΔY_0, and ΔZ_0 are chosen as zero matrices in suitable size, and the tolerance $\varepsilon = 10^{-10}$. In Algorithm 2.1, the initial iterative matrices X_0 and Z_0 are chosen as the matrices with all elements equal to one, Y_0 is chosen as zero matrix, and the tolerances $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 10^{-10}$.

Example 3.1. Assume that a goods is only made by three manufacturers. The market occupancy distribution of these three manufacturers in the past six months is given as follows Table 1. If, in nearly future time, people's consumption pattern and enterprise technology have not changed. There is also no other enterprise make this kind of goods. We are required giving the market occupancy distribution of these three manufacturers in the next six months.

Table 1. The market occupancy distribution in the past six months

	1	2	3	4	5	6	
Manufacturer 1	0.4666	0.4533	0.4551	0.4423	0.4113	0.4207	0.4087
Manufacturer 2	0.3633	0.3432	0.3312	0.3308	0.3211	0.3107	0.3074
Manufacturer 3	0.1701	0.2035	0.2137	0.2269	0.2676	0.2686	0.2839

Applying Markov chains theory to forecast this economic activities, we need first to find the transition probability matrix X by solving state matrix equation $AX = B$. From Table 1, the state matrices A and B are as follows:

$$A = \begin{pmatrix} 0.4666 & 0.3633 & 0.1701 \\ 0.4533 & 0.3432 & 0.2035 \\ 0.4551 & 0.3312 & 0.2137 \\ 0.4423 & 0.3308 & 0.2269 \\ 0.4113 & 0.3211 & 0.2676 \\ 0.4207 & 0.3107 & 0.2686 \end{pmatrix}, \quad B = \begin{pmatrix} 0.4533 & 0.3432 & 0.2035 \\ 0.4551 & 0.3312 & 0.2137 \\ 0.4423 & 0.3308 & 0.2269 \\ 0.4113 & 0.3211 & 0.2676 \\ 0.4207 & 0.3107 & 0.2686 \\ 0.4087 & 0.3074 & 0.2839 \end{pmatrix},$$

which do not satisfy $AX = B$, $Xe = e$, and $X \geq 0$. In other words, there is no stochastic matrix X such that $AX = B$. In this case, we hope to find the smallest correction stochastic solution X of the inconsistent matrix equation $AX = B$, that is, find the solution of the problem (1.1). Using Algorithms 2.1 and 2.2, we obtain the transition probability matrix X as follows:

$$X = \begin{pmatrix} 0.4346 & 0.4535 & 0.1118 \\ 0.6618 & 0.3382 & 0.0000 \\ 0.0860 & 0.0491 & 0.8649 \end{pmatrix}.$$

Using state distribution formula $w(t+1) = w(t)X$, $t = 1, 2, \ldots,$ where $w(t)^T \in R^3$ is the probability distribution vector in state t, we get the market occupancy distribution of these three manufacturers in the next six months as follows:

Table 2. The market occupancy distribution in the next six months

	7	8	9	10	11	12
Manufacturer 1	0.4055	0.4020	0.3993	0.3972	0.3954	0.3940
Manufacturer 2	0.3033	0.3008	0.2986	0.2969	0.2956	0.2945
Manufacturer 3	0.2913	0.2973	0.3021	0.3059	0.3090	0.3115

Note that the transition probability matrix X satisfies

$$\lim_{n \to \infty} X^n = \begin{pmatrix} 0.3884 & 0.2900 & 0.3216 \\ 0.3884 & 0.2900 & 0.3216 \\ 0.3884 & 0.2900 & 0.3216 \end{pmatrix},$$

we know that the fixed market occupancy distribution of these three manufacturers is $(0.3884, 0.2900, 0.3216)$.

References

[1] F. Bastos and A. Teixeira, An extension of variant of a predictor-corrector primal-dual method from linear programming to semidefinite programming, Investação Operacional 25 (2005), 253-276.

[2] Henri Berthiaux, Vadim Mizonov and Vladimir Zhukov, Application of the theory of Markov chains to model different processes in particle technology, Powder Technology 157 (2005), 128-137.

[3] Coralia Cartis, Some disadvantages of a Mehrotra-type primal-dual corrector interior point algorithm for linear programming, Applied Numerical Mathematics 59(5) (2009), 1110-1119.

[4] X. F. Cha, Application of Markov chain to market forecasting, Journal of Jiangsu University (Social Science Edition) 5(1) (2003), 110-113.

[5] Bo Kyung Choi and Gue Myung Lee, On complexity analysis of the primal-dual interior-point method for semidefinite optimization problem based on a new proximity function, Nonlinear Analysis: Theory, Methods & Applications 71(12) (2009), 2628-2640.

[6] Richard G. Clegg, A discrete-time Markov-modulated queuing system with batched arrivals, Performance Evaluation 67 (2010), 376-385.

[7] R. L. Feng, D. F. Ou and Z. H. Peng, Analysis of market share: A practical research of Markov theory, Journal of University of Electric Power (Natural Science) 20(4) (2005), 85-88.

[8] Tibor Ills and Marianna Nagy, A Mizuno-Todd-Ye type predictor-corrector algorithm for sufficient linear complementarity problems, European Journal of Operational Research 181(3) (2007), 1097-1111.

[9] N. M. S. Karmitsa, M. M. Mkel and M. M. Ali, Limited memory interior point bundle method for large inequality constrained nonsmooth minimization, Applied Mathematics and Computation 198(1) (2008), 382-400.

[10] L. X. Liu, S. Y. Liu and H. W. Liu, A predictor-corrector smoothing Newton method for symmetric cone complementarity problems, Applied Mathematics and Computation 217(7) (2010), 2989-2999.

[11] J. Peng, T. Terlaky and Y. B. Zhao, A predictor-corrector algorithm for linear optimization based on a specific self-regular proximity function, SIAM Journal on Optimization 15(4) (2005), 1105-1127.

[12] J. J. F. A. Potra and S. Huang, A predictor-corrector method for linear complementarity problems with polynomial complexity and superlinear convergence, Journal of Optimization Theory and Applications 84(1) (1995), 187-199.

[13] Florian A. Potra, The Mizuno-Todd-Ye algorithm in a large neighbourhood of the central path, European Journal of Operation Research 143 (2002), 257-267.

[14] Florian A. Potra, A superlinearly convergent predictor-corrector method for degenerate LCP in a wide neighbourhood of the central path with $O(\sqrt{n}L)$ -iteration complexity, Mathematical Programming 100(2) (2004), 317-337.

[15] V. Rico-Ramrez and A. W. Westerberg, Interior point methods for the solution of conditional models, Computers & Chemical Engineering 26(3) (2002), 375-383.

[16] Maziar Salahi and Nezam Mahdavi-Amiri, Polynomial time second order Mehrotra-type predictor-corrector algorithms, Applied Mathematics and Computation 183(1) (2006), 646-658.

[17] M. Salahi, J. Peng and T. Terlaky, On Mehrotra-type predictor-corrector algorithms, SIAM Journal on Optimization 18(4) (2007), 1377-1397.

[18] Maziar Salahi, A finite termination Mehrotra-type predictor-corrector algorithm, Applied Mathematics and Computation 190(2) (2007), 1740-1746.

Appendix

Proof of Proposition 2.1. If there exists a positive number k such that $\beta_i \neq 0$ for all $i = 0, 1, 2, \cdots, k,$ the conclusion $\alpha_i \neq 0$ obviously holds. Since $\langle A, B\rangle = \langle B, A\rangle$ holds for all matrices A and B in $R^{m\times n}$, we only need to prove the conclusions (2.6) and (2.7) hold for all $0 \leq j < i \leq k.$ Using induction and the following two steps are required:

Step 1. Show that

$$tr(P_{1,i}^T P_{1,j}) + tr(P_{2,i}^T P_{2,j}) + tr(P_{3,i}^T P_{3,j}) = 0, \ (j = 0, \ i = 1, 2, \cdots, k), \quad (3.1)$$

and

$$tr(R_{1,i}^T R_{1,j}) + tr(R_{2,i}^T R_{2,j}) + tr(R_{3,i}^T R_{3,j}) = 0, \ (j = 0, \ i = 1, 2, \cdots, k). \quad (3.2)$$

To prove these conclusions, we also using induction.

For $j = 0, \ i = 1,$ we have

$$\begin{aligned} tr(R_{1,1}^T R_{1,0}) &= \langle R_{1,0}, \ R_{1,0} - \frac{\alpha_0}{\beta_0}(A^T A P_{1,0} - P_{2,0} e^T - P_{3,0}) \rangle \\ &= \|R_{1,0}\|^2 - \frac{\alpha_0}{\beta_0}\langle R_{1,0}, \ A^T A P_{1,0} - P_{2,0} e^T - P_{3,0} \rangle \\ &= \|R_{1,0}\|^2 - \frac{\alpha_0}{\beta_0}[\langle A^T A R_{1,0}, \ P_{1,0} \rangle - \langle R_{1,0} e, \ P_{2,0} \rangle \\ &\quad - \langle R_{1,0}, \ P_{3,0} \rangle], \end{aligned}$$

$$\begin{aligned} tr(R_{2,1}^T R_{2,0}) &= \langle R_{2,0}, \ R_{2,0} - \frac{\alpha_0}{\beta_0} P_{1,0} e \rangle \\ &= \|R_{2,0}\|^2 - \frac{\alpha_0}{\beta_0}\langle R_{2,0} e^T, \ P_{1,0} \rangle, \end{aligned}$$

$$\begin{aligned} tr(R_{3,1}^T R_{3,0}) &= (R_{3,0}, \ R_{3,0} - \frac{\alpha_0}{\beta_0}(Z_k \bullet P_{1,0} + X_k \bullet P_{3,0}) \rangle \\ &= \|R_{3,0}\|^2 - \frac{\alpha_0}{\beta_0}\langle R_{3,0}, \ Z_k \bullet P_{1,0} + X_k \bullet P_{3,0} \rangle \\ &= \|R_{3,0}\|^2 - \frac{\alpha_0}{\beta_0}[\langle Z_k \bullet R_{3,0}, \ P_{1,0} \rangle + \langle X_k \bullet R_{3,0}, \ P_{3,0} \rangle]. \end{aligned}$$

Sum up above three equalities, we have

$$tr(R_{1,1}^T R_{1,0}) + tr(R_{2,1}^T R_{2,0}) + tr(R_{3,1}^T R_{3,0})$$

$$= \alpha_0 - \frac{\alpha_0}{\beta_0}(\langle A^T AR_{1,0} + R_{2,0}e^T + Z_k \bullet R_{3,0},\ P_{1,0}\rangle$$

$$- \langle R_{1,0}e,\ P_{2,0}\rangle + \langle - R_{1,0} + X_k \bullet R_{3,0},\ P_{3,0}\rangle)$$

$$= \alpha_0 - \frac{\alpha_0}{\beta_0}[\|P_{1,0}\|^2 + \|P_{2,0}\|^2 + \|P_{3,0}\|^2] = 0.$$

Analogously, we have

$$tr(P_{1,1}^T P_{1,0}) = \langle P_{1,0},\ A^T AR_{1,1} + R_{2,1}e^T + Z_k \bullet R_{3,1} + \frac{\alpha_1}{\alpha_0}P_{1,0}\rangle$$

$$= \frac{\alpha_1}{\alpha_0}\|P_{1,0}\|^2 + \langle P_{1,0},\ A^T AR_{1,1} + R_{2,1}e^T + Z_k \bullet R_{3,1}\rangle$$

$$= \frac{\alpha_1}{\alpha_0}\|P_{1,0}\|^2 + \langle A^T AP_{1,0},\ R_{1,1}\rangle + \langle P_{1,0}e,\ R_{2,1}\rangle$$

$$+ \langle Z_k \bullet P_{1,0},\ R_{3,1}\rangle,$$

$$tr(P_{2,1}^T P_{2,0}) = \langle P_{2,0},\ - R_{1,1}e + \frac{\alpha_1}{\alpha_0}P_{2,0}\rangle$$

$$= \frac{\alpha_1}{\alpha_0}\|P_{2,0}\|^2 - \langle P_{2,0}e^T,\ R_{1,1}\rangle,$$

$$tr(P_{3,1}^T P_{3,0}) = \langle P_{3,0},\ - R_{1,1} + X_k \bullet R_{3,1} + \frac{\alpha_1}{\alpha_0}P_{3,0}\rangle$$

$$= \frac{\alpha_1}{\alpha_0}\|P_{3,0}\|^2 + \langle P_{3,0},\ - R_{1,1} + X_k \bullet R_{3,1}\rangle$$

$$= \frac{\alpha_1}{\alpha_0}\|P_{3,0}\|^2 - \langle P_{3,0},\ R_{1,1}\rangle + \langle X_k \bullet P_{3,0},\ R_{3,1}\rangle.$$

And sum up above three equalities, we have

$$tr(P_{1,1}^T P_{1,0}) + tr(P_{2,1}^T P_{2,0}) + tr(P_{3,1}^T P_{3,0})$$

$$= \frac{\alpha_1}{\alpha_0}\beta_0 + \langle A^T A P_{1,0} - P_{2,0}e^T - P_{3,0},\, R_{1,1}\rangle + \langle P_{1,0}e,\, R_{2,1}\rangle$$

$$+ \langle Z_k \bullet P_{1,0} + X_k \bullet P_{3,0},\, R_{3,1}\rangle$$

$$= \frac{\alpha_1}{\alpha_0}\beta_0 + \frac{\beta_0}{\alpha_0}\langle R_{1,0} - R_{1,1},\, R_{1,1}\rangle + \frac{\beta_0}{\alpha_0}\langle R_{2,0} - R_{2,1},\, R_{2,1}\rangle$$

$$+ \frac{\beta_0}{\alpha_0}\langle R_{3,0} - R_{3,1},\, R_{3,1}\rangle$$

$$= \frac{\beta_0}{\alpha_0}(tr(R_{1,1}^T R_{1,0}) + tr(R_{2,1}^T R_{2,0}) + tr(R_{3,1}^T R_{3,0}))$$

$$= 0.$$

Assume that the conclusions (3.1) and (3.2) hold for all $j = 0,\ i = 1, 2, \cdots, s,$ then for $j = 0,\ i = s+1,$ we have

$$tr(R_{1,s+1}^T R_{1,0}) = \langle R_{1,0},\, R_{1,s} - \frac{\alpha_s}{\beta_s}(A^T A P_{1,s} - P_{2,s}e^T - P_{3,s})\rangle$$

$$= \langle R_{1,0},\, R_{1,s}\rangle - \frac{\alpha_s}{\beta_s}\langle R_{1,0},\, A^T A P_{1,s} - P_{2,s}e^T - P_{3,s}\rangle$$

$$= \langle R_{1,0},\, R_{1,s}\rangle - \frac{\alpha_s}{\beta_s}[\langle A^T A R_{1,0},\, P_{1,s}\rangle - \langle R_{1,0}e,\, P_{2,s}\rangle$$

$$- \langle R_{1,0},\, P_{3,s}\rangle],$$

$$tr(R_{2,s+1}^T R_{2,0}) = \langle R_{2,0},\, R_{2,s} - \frac{\alpha_s}{\beta_s}P_{1,s}e\rangle$$

$$= \langle R_{2,0},\, R_{2,s}\rangle - \frac{\alpha_s}{\beta_s}\langle R_{2,0}e,\, P_{1,s}\rangle,$$

$$tr(R_{3,s+1}^T R_{3,0}) = \langle R_{3,0},\, R_{3,s} - \frac{\alpha_s}{\beta_s}(Z_k \bullet P_{1,s} + X_k \bullet P_{3,s})\rangle$$

$$= \langle R_{3,0},\ R_{3,s}\rangle - \frac{\alpha_s}{\beta_s}\langle R_{3,0},\ Z_k \bullet P_{1,s} + X_k \bullet P_{3,s}\rangle$$

$$= \langle R_{3,0},\ R_{3,s}\rangle - \frac{\alpha_s}{\beta_s}[\langle Z_k \bullet R_{3,0},\ P_{1,s}\rangle + \langle X_k \bullet R_{3,0},\ P_{3,s}\rangle];$$

$$tr(R_{1,s+1}^T R_{1,0}) + tr(R_{2,s+1}^T R_{2,0}) + tr(R_{3,s+1}^T R_{3,0})$$

$$= -\frac{\alpha_s}{\beta_s}(\langle A^T AR_{1,0} + R_{2,0}e^T + Z_k \bullet R_{3,0},\ P_{1,s}\rangle$$

$$- \langle R_{1,0}e,\ P_{2,s}\rangle + \langle - R_{1,0} + X_k \bullet R_{3,0},\ P_{3,s}\rangle)$$

$$= -\frac{\alpha_s}{\beta_s}(\langle P_{1,0},\ P_{1,s}\rangle + \langle P_{2,0},\ P_{2,s}\rangle + \langle P_{3,0},\ P_{3,s}\rangle)$$

$$= 0.$$

$$tr(P_{1,s+1}^T P_{1,0}) = \langle P_{1,0},\ A^T AR_{1,s} + R_{2,s}e^T + Z_k \bullet R_{3,s} + \frac{\alpha_{s+1}}{\alpha_s} P_{1,s}\rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s}\langle P_{1,0},\ P_{1,s}\rangle + \langle P_{1,0},\ A^T AR_{1,s} + R_{2,s}e^T + Z_k \bullet R_{3,s}\rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s}\langle P_{1,0},\ P_{1,s}\rangle + \langle A^T AP_{1,0},\ R_{1,s}\rangle + \langle P_{1,0}e,\ R_{2,s}\rangle$$

$$+ \langle Z_k \bullet P_{1,0},\ R_{3,s}\rangle,$$

$$tr(P_{2,s+1}^T P_{2,0}) = \langle P_{2,0},\ - R_{1,s}e + \frac{\alpha_{s+1}}{\alpha_s} P_{2,0}\rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s}\langle P_{2,0},\ P_{2,s}\rangle - \langle P_{2,0}e^T,\ R_{1,s}\rangle,$$

$$tr(P_{3,s+1}^T P_{3,0}) = \langle P_{3,0},\ Z_k \bullet R_{1,s} + X_k \bullet R_{3,s} + \frac{\alpha_{s+1}}{\alpha_s} P_{3,0}\rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s}\langle P_{3,0},\ P_{3,s}\rangle + \langle P_{3,0},\ - R_{1,s} + X_k \bullet R_{3,s}\rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s} \langle P_{3,0}, P_{3,s} \rangle - \langle P_{3,0}, R_{1,s} \rangle + \langle X_k \bullet P_{3,0}, R_{3,s} \rangle;$$

$$tr(P_{1,s+1}^T P_{1,0}) + tr(P_{2,s+1}^T P_{2,0}) + tr(P_{3,s+1}^T P_{3,0})$$

$$= \langle A^T A P_{1,0} - P_{2,0} e^T - P_{3,0}, R_{1,s} \rangle + \langle P_{1,0} e, R_{2,s} \rangle$$

$$+ \langle Z_k \bullet P_{1,0} + X_k \bullet P_{3,0}, R_{3,s} \rangle$$

$$= \frac{\beta_0}{\alpha_0} \langle R_{1,0} - R_{1,1}, R_{1,s} \rangle + \frac{\beta_0}{\alpha_0} \langle R_{2,0} - R_{2,1}, R_{2,s} \rangle$$

$$+ \frac{\beta_0}{\alpha_0} \langle R_{3,0} - R_{3,1}, R_{3,s} \rangle$$

$$= 0.$$

By the principle of induction, the conclusions (3.1) and (3.2) hold for all $i = 0, 1, 2, \cdots, k$.

Step 2. Assume that

$$tr(P_{1,i}^T P_{1,s}) + tr(P_{2,i}^T P_{2,s}) + tr(P_{3,i}^T P_{3,s}) = 0,$$

$$tr(R_{1,i}^T R_{1,s}) + tr(R_{2,i}^T R_{2,s}) + tr(R_{3,i}^T R_{3,s}) = 0,$$

hold for all $0 \leq i \leq k$ and $1 < s+1 < i$, show that

$$tr(P_{1,i}^T P_{1,s+1}) + tr(P_{2,i}^T P_{2,s+1}) + tr(P_{3,i}^T P_{3,s+1}) = 0,$$

$$tr(R_{1,i}^T R_{1,s+1}) + tr(R_{2,i}^T R_{2,s+1}) + tr(R_{3,i}^T R_{3,s+1}) = 0.$$

The proof are following:

$$tr(R_{1,i}^T R_{1,s+1}) = \langle R_{1,i}, R_{1,s} - \frac{\alpha_s}{\beta_s} (A^T A P_{1,s} - P_{2,s} e^T - P_{3,s}) \rangle$$

$$= \langle R_{1,i}, R_{1,s} \rangle - \frac{\alpha_s}{\beta_s} \langle R_{1,i}, A^T A P_{1,s} - P_{2,s} e^T - P_{3,s} \rangle$$

$$= \langle R_{1,i}, R_{1,s} \rangle - \frac{\alpha_s}{\beta_s} [\langle A^T A R_{1,i}, P_{1,s} \rangle - \langle R_{1,i} e, P_{2,s} \rangle$$

$$- \langle R_{1,i}, P_{3,s} \rangle],$$

$$tr(R_{2,i}^T R_{2,s+1}) = \langle R_{2,i}, R_{2,s} - \frac{\alpha_s}{\beta_s} P_{1,s} e \rangle$$

$$= \langle R_{2,i}, R_{2,s} \rangle - \frac{\alpha_s}{\beta_s} \langle R_{2,i} e, P_{1,s} \rangle,$$

$$tr(R_{3,i}^T R_{3,s+1}) = \langle R_{3,i}, R_{3,s} - \frac{\alpha_s}{\beta_s} (Z_k \bullet P_{1,s} + X_k \bullet P_{3,s}) \rangle$$

$$= \langle R_{3,i}, R_{3,s} \rangle - \frac{\alpha_s}{\beta_s} \langle R_{3,i}, Z_k \bullet P_{1,s} + X_k \bullet P_{3,s} \rangle$$

$$= \langle R_{3,i}, R_{3,s} \rangle - \frac{\alpha_s}{\beta_s} [\langle Z_k \bullet R_{3,i}, P_{1,s} \rangle + \langle X_k \bullet R_{3,i}, P_{3,s} \rangle];$$

$$tr(R_{1,i}^T R_{1,s+1}) + tr(R_{2,i}^T R_{2,s+1}) + tr(R_{3,i}^T R_{3,s+1})$$

$$= -\frac{\alpha_s}{\beta_s} (\langle A^T A R_{1,i} + R_{2,i} e^T + Z_k \bullet R_{3,i}, P_{1,s} \rangle$$

$$- \langle R_{1,i} e, P_{2,s} \rangle + \langle - R_{1,i} + X_k \bullet R_{3,i}, P_{3,s} \rangle)$$

$$= -\frac{\alpha_s}{\beta_s} (\langle P_{1,i+1} - \frac{\alpha_{i+1}}{\alpha_i} P_{1,i}, P_{1,s} \rangle$$

$$+ \langle P_{2,i+1} - \frac{\alpha_{i+1}}{\alpha_i} P_{2,i}, P_{2,s} \rangle + \langle P_{3,i+1} - \frac{\alpha_{i+1}}{\alpha_i} P_{3,i}, P_{3,s} \rangle)$$

$$= -\frac{\alpha_s}{\beta_s} [tr(P_{1,i+1}^T P_{1,s}) + tr(P_{2,i+1}^T P_{2,s}) + tr(P_{3,i+1}^T P_{3,s})]$$

$$+ \frac{\alpha_s \alpha_{i+1}}{\beta_s \alpha_i} [tr(P_{1,i}^T P_{1,s}) + tr(P_{2,i}^T P_{2,s}) + tr(P_{3,i}^T P_{3,s})] = 0.$$

$$tr(P_{1,i}^T P_{1,s+1}) = \langle P_{1,i},\ A^T A R_{1,s} + R_{2,s} e^T + Z_k \bullet R_{3,s} + \frac{\alpha_{s+1}}{\alpha_s} P_{1,s} \rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s} \langle P_{1,i},\ P_{1,s} \rangle + \langle P_{1,i},\ A^T A R_{1,s} + R_{2,s} e^T + Z_k \bullet R_{3,s} \rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s} \langle P_{1,i},\ P_{1,s} \rangle + \langle A^T A P_{1,i},\ R_{1,s} \rangle + \langle P_{1,i} e,\ R_{2,s} \rangle$$

$$+ \langle Z_k \bullet P_{1,i},\ R_{3,s} \rangle,$$

$$tr(P_{2,i}^T P_{2,s+1}) = \langle P_{2,i},\ - R_{1,s} e + \frac{\alpha_{s+1}}{\alpha_s} P_{2,s} \rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s} \langle P_{2,i},\ P_{2,s} \rangle - \langle P_{2,i} e^T,\ R_{1,s} \rangle,$$

$$tr(P_{3,i}^T P_{3,s+1}) = \langle P_{3,i},\ Z_k \bullet R_{1,s} + X_k \bullet R_{3,s} + \frac{\alpha_{s+1}}{\alpha_s} P_{3,s} \rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s} \langle P_{3,i},\ P_{3,s} \rangle + \langle P_{3,i},\ - R_{1,s} + X_k \bullet R_{3,s} \rangle$$

$$= \frac{\alpha_{s+1}}{\alpha_s} \langle P_{3,i},\ P_{3,s} \rangle - \langle P_{3,i},\ R_{1,s} \rangle + \langle X_k \bullet P_{3,i},\ R_{3,s} \rangle;$$

$$tr(P_{1,i}^T P_{1,s+1}) + tr(P_{2,i}^T P_{2,s+1}) + tr(P_{3,i}^T P_{3,s+1})$$

$$= \langle A^T A P_{1,i} - P_{2,i} e^T - P_{3,i},\ R_{1,s} \rangle + \langle P_{1,i} e,\ R_{2,s} \rangle$$

$$+ \langle Z_k \bullet P_{1,0i} + X_k \bullet P_{3,i},\ R_{3,s} \rangle$$

$$= \frac{\beta_i}{\alpha_i} \langle R_{1,i} - R_{1,i+1},\ R_{1,s} \rangle + \frac{\beta_i}{\alpha_i} \langle R_{2,i} - R_{2,i+1},\ R_{2,s} \rangle$$

$$+ \frac{\beta_i}{\alpha_i} \langle R_{3,i} - R_{3,i+1},\ R_{3,s} \rangle$$

$$= \frac{\beta_i}{\alpha_i}[tr(R_{1,i}^T R_{1,s}) + tr(R_{2,i}^T R_{2,s}) + tr(R_{3,i+1}^T R_{3,s})]$$

$$- \frac{\beta_i}{\alpha_i}[tr(R_{1,i+1}^T R_{1,s}) + tr(R_{2,i+1}^T R_{2,s}) + tr(R_{3,i+1}^T R_{3,s}) = 0.$$

From Steps 1 and 2, we have by principle induction that the conclusions (2.6) and (2.7) hold for all $i, j = 0, 1, 2, \cdots, k, i \neq j$. □

Proof of Proposition 2.2. We use induction to prove this conclusion. For $i = 0$, we have

$$tr[(\Delta\overline{X} - \Delta X_0)^T P_{1,0}] + tr[(\Delta\overline{Y} - \Delta Y_0)^T P_{2,0}] + tr[(\Delta\overline{Z} - \Delta Z_0)^T P_{3,0}]$$

$$= \langle \Delta\overline{X} - \Delta X_0,\ A^T A R_{1,0} + R_{2,0}e^T + Z_k \bullet R_{3,0} \rangle$$

$$+ \langle \Delta\overline{Y} - \Delta Y_0,\ - R_{1,0}e \rangle + \langle \Delta\overline{Z} - \Delta Z_0,\ - R_{1,0} + X_k \bullet R_{3,0} \rangle$$

$$= \langle A^T A(\Delta\overline{X} - \Delta X_0),\ R_{1,0} \rangle + \langle (\Delta\overline{X} - \Delta X_0)e,\ R_{2,0} \rangle$$

$$+ \langle Z_k \bullet (\Delta\overline{X} - \Delta X_0),\ R_{3,0} \rangle + \langle (\Delta\overline{Y} - \Delta Y_0)e^T,\ - R_{1,0} \rangle$$

$$+ \langle (\Delta\overline{Z} - \Delta Z_0),\ - R_{1,0} \rangle + \langle X_k \bullet (\Delta\overline{Z} - \Delta Z_0),\ R_{3,0} \rangle$$

$$= \langle A^T A(\Delta\overline{X} - \Delta X_0) - (\Delta\overline{Y} - \Delta Y_0)e^T - (\Delta\overline{Z} - \Delta Z_0),\ R_{1,0} \rangle$$

$$+ \langle (\Delta\overline{X} - \Delta X_0)e,\ R_{2,0} \rangle + \langle Z_k \bullet (\Delta\overline{X} - \Delta X_0)$$

$$+ X_k \bullet (\Delta\overline{Z} - \Delta Z_0),\ R_{3,0} \rangle$$

$$= \|R_{1,0}\|^2 + \|R_{2,0}\|^2 + \|R_{3,0}\|^2.$$

Assume (2.8) holds for $i = s$. Since

$$tr[(\Delta\overline{X} - \Delta X_{s+1})^T P_{1,s+1}] + tr[(\Delta\overline{Y} - \Delta Y_{s+1})^T P_{2,s+1}]$$

$$+ \; tr[(\Delta\overline{Z} - \Delta Z_{s+1})^T P_{3,s+1}]$$

$$= \langle \Delta\overline{X} - \Delta X_{s+1}, \; A^T A R_{1,s+1} + R_{2,s+1} e^T + Z_k \bullet R_{3,s+1} \rangle$$

$$+ \frac{\alpha_{s+1}}{\alpha_s} \langle \Delta\overline{X} - \Delta X_{s+1}, \; P_{1,s} \rangle$$

$$+ \langle \Delta\overline{Y} - \Delta Y_{s+1}, \; -R_{1,s+1} e \rangle + \frac{\alpha_{s+1}}{\alpha_s} \langle \Delta\overline{Y} - \Delta Y_{s+1}, \; P_{2,s} \rangle$$

$$+ \langle \Delta\overline{Z} - \Delta Z_{s+1}, \; -R_{1,s+1} + X_k \bullet R_{3,s+1} \rangle$$

$$+ \frac{\alpha_{s+1}}{\alpha_s} \langle \Delta\overline{Z} - \Delta Z_{s+1}, \; P_{3,s} \rangle$$

$$= \langle A^T A(\Delta\overline{X} - \Delta X_{s+1}), \; R_{1,s+1} \rangle + \langle (\Delta\overline{X} - \Delta X_{s+1}) e, \; R_{2,s+1} \rangle$$

$$+ \langle Z_k \bullet (\Delta\overline{X} - \Delta X_{s+1}), \; R_{3,s+1} \rangle + \langle (\Delta\overline{Y} - \Delta Y_{s+1}) e^T, \; -R_{1,s+1} \rangle$$

$$+ \langle (\Delta\overline{Z} - \Delta Z_{s+1}), \; -R_{1,s+1} \rangle + \langle X_k \bullet (\Delta\overline{Z} - \Delta Z_{s+1}), \; R_{3,s+1} \rangle$$

$$+ \frac{\alpha_{s+1}}{\alpha_s} [\langle \Delta\overline{X} - \Delta X_s - \frac{\alpha_s}{\beta_s} P_{1,s}, \; P_{1,s} \rangle$$

$$+ \langle \Delta\overline{Y} - \Delta Y_s - \frac{\alpha_s}{\beta_s} P_{2,s}, \; P_{2,s} \rangle + \langle \Delta\overline{Z} - \Delta Z_s - \frac{\alpha_s}{\beta_s} P_{3,s}, \; P_{3,s} \rangle]$$

$$= \langle A^T A(\Delta\overline{X} - \Delta X_{s+1}) - (\Delta\overline{Y} - \Delta Y_{s+1}) e^T - (\Delta\overline{Z} - \Delta Z_{s+1}), \; R_{1,s+1} \rangle$$

$$+ \langle (\Delta\overline{X} - \Delta X_{s+1}) e, \; R_{2,s+1} \rangle + \langle Z_k \bullet (\Delta\overline{X} - \Delta X_{s+1})$$

$$+ X_k \bullet (\Delta\overline{Z} - \Delta Z_{s+1}), \; R_{3,s+1} \rangle + \frac{\alpha_{s+1}}{\alpha_s} [\langle \Delta\overline{X} - \Delta X_s, \; P_{1,s} \rangle$$

$$+ \langle \Delta\overline{Y} - \Delta Y_s, \; P_{2,s} \rangle + \langle \Delta\overline{Z} - \Delta Z_s, \; P_{3,s} \rangle]$$

$$-\frac{\alpha_{s+1}}{\alpha_s}\frac{\alpha_s}{\beta_s}[\langle P_{1,s}, P_{1,s}\rangle + \langle P_{2,s}, P_{2,s}\rangle + \langle P_{3,s}, P_{3,s}\rangle]$$

$$= \|R_{1,s+1}\|^2 + \|R_{2,s+1}\|^2 + \|R_{3,s+1}\|^2.$$

Hence the conclusion (2.8) holds by the principle of induction. □

Proof of Proposition 2.3. Let $R_i = \operatorname{diag}(R_{1,i}, R_{2,i}, R_{3,i})$, $P_i = \operatorname{diag}(P_{1,i}, P_{2,i}, P_{3,i})$, then the conclusions (2.6) and (2.7) in Proposition 2.1 can be rewritten as $tr(R_i^T R_j) = 0, tr(P_i^T P_j) = 0$ hold for all $i, j = 0, 1, 2, \cdots, k, i \neq j$, which imply that the matrix sequences $\{R_i\}$ and $\{P_i\}$ are F-orthogonal sequences in the finite dimension matrix space $R^{3n\times(2n+1)}$. Hence, it is certainly there exists a positive number $i \leq 3n\ (2n+1)$ such that $P_i = 0$ and $R_i = 0$, (or $P_i = 0$ and $R_i \neq 0$). These conclusions imply that Algorithm 2.1 will break down within finite iteration steps in the absence of roundoff errors. Noting that $R_{1,i}$, $R_{2,i}$, and $R_{3,i}$ are, respectively, the residual of the first equation, the second equation, and the third equation of the matrix equations (2.2) (or (2.4)) at step i, then, if $\beta_i = 0$ and $\alpha_i = 0$, the matrix equations (2.2) (or (2.4)) is certainly solvable and $(\Delta X_i, \Delta Y_i, \Delta Z_i)$ is its a solution. If $\beta_i = 0$ and $\alpha_i \neq 0$, then the matrix equations (2.2) (or (2.4)) has no solution. Otherwise, if the matrix equations (2.2) (or (2.4)) is solvable and $\alpha_i \neq 0$, then we know from Proposition 2.2 that $\beta_i \neq 0$. □

■