Chapter 7
Adjustment and Filtering Methods

7.1 Introduction

In this chapter, we outline the most useful and necessary adjustment and filtering algorithms for statistical and kinematic as well as dynamic GPS data processing. We derive the necessary estimators, and provide a detailed discussion of the relationships between the methods presented.

The adjustment algorithms discussed here include least squares adjustment, sequential application of least squares adjustment via accumulation, sequential least squares adjustment, conditional least squares adjustment, a sequential application of conditional least squares adjustment, block-wise least squares adjustment, a sequential application of block-wise least squares adjustment, a special application of block-wise least squares adjustment for code–phase combinations, an equivalent algorithm to form the eliminated observation equation system, and an algorithm to diagonalise the normal and equivalent observation equations.

The filtering algorithms discussed here include the classic Kalman filter, the sequential least squares adjustment method as a special case of Kalman filtering, the robust Kalman filter, and the adaptively robust Kalman filter.

A priori constrained adjustment and filtering are discussed for solving rank-deficient problems. After a general discussion on a priori parameter constraints, a special case of the so-called a priori datum method is provided. A quasi-stable datum method is also discussed.

A summary is presented at the end of this chapter, and applications of the GPS data processing methods discussed are outlined.

7.2 Least Squares Adjustment

The principle of least squares adjustment can be summarised as outlined below (Gotthardt 1978; Cui et al. 1982):

G. Xu and Y. Xu, *GPS*, DOI 10.1007/978-3-662-50367-6_7

1. The linearised observation equation system can be represented by

$$V = L - AX,\ P \tag{7.1}$$

where
- L observation vector of dimension m,
- A coefficient matrix of dimension $m \times n$,
- X unknown parameter vector of dimension n,
- V residual vector of dimension m,
- n number of unknowns,
- m number of observations, and
- P symmetric and definite weight matrix of dimension $m \times m$.

2. The least squares criterion for solving the observation equations is well known as

$$V^{\mathrm{T}}PV = \min, \tag{7.2}$$

where
- V^{T} the transpose of the related vector V.

3. To solve X and compute V, a function F is set as

$$F = V^{\mathrm{T}}PV. \tag{7.3}$$

The function F reaches minimum value if the partial differentiation of F with respect to X equals zero, i.e.

$$\frac{\partial F}{\partial X} = 2V^{\mathrm{T}}P(-A) = 0$$

or

$$A^{\mathrm{T}}PV = 0, \tag{7.4}$$

where
- A^{T} transpose matrix of A.

4. Multiplying $A^{\mathrm{T}}P$ with Eq. 7.1, one has

$$A^{\mathrm{T}}PAX - A^{\mathrm{T}}PL = -A^{\mathrm{T}}PV. \tag{7.5}$$

Setting Eq. 7.4 into 7.5, one has

$$A^{\mathrm{T}}PAX - A^{\mathrm{T}}PL = 0. \tag{7.6}$$

5. For simplification, let $M = A^{\mathrm{T}}PA$, $Q = M^{-1}$, where superscript $^{-1}$ is an inverse operator, and M is usually called a normal matrix. The least squares solution of Eq. 7.1 is then

$$X = Q(A^{\mathrm{T}}PL). \tag{7.7}$$

6. The precision of the ith element of the estimated parameter is

$$p[i] = m_0\sqrt{Q[i][i]}, \tag{7.8}$$

where i is the element index of a vector or a matrix, m_0 is the so-called standard deviation (or sigma), $p[i]$ is the ith element of the precision vector, $Q[i][i]$ is the ith diagonal element of the cofactor matrix Q, and

$$m_0 = \sqrt{\frac{V^{\mathrm{T}}PV}{m-n}}, \quad \text{if} \quad (m > n). \tag{7.9}$$

7. For convenience of sequential computation, $V^{\mathrm{T}}PV$ can be calculated by using

$$V^{\mathrm{T}}PV = L^{\mathrm{T}}PL - (A^{\mathrm{T}}PL)^{\mathrm{T}}X. \tag{7.10}$$

This can be obtained by substituting Eq. 7.1 into $V^{\mathrm{T}}PV$ and considering Eq. 7.4. Thus far, we have derived the complete formulas of least squares adjustment.

7.2.1 Least Squares Adjustment with Sequential Observation Groups

Suppose one has two sequential observation equation systems

$$V_1 = L_1 - A_1X \quad \text{and} \tag{7.11}$$

$$V_2 = L_2 - A_2X, \tag{7.12}$$

with weight matrices P_1 and P_2. These two equation systems are uncorrelated or independent and have the common unknown vector X. The combined problem can be represented as

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} - \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} X \quad \text{and} \quad P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}. \tag{7.13}$$

The least squares normal equation can be formed then as

$$\begin{pmatrix} A_1^{\mathrm{T}} & A_2^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} A_1 \\ A_2 \end{pmatrix} X = \begin{pmatrix} A_1^{\mathrm{T}} & A_2^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}$$

or

$$(A_1^{\mathrm{T}}P_1A_1 + A_2^{\mathrm{T}}P_2A_2)X = (A_1^{\mathrm{T}}P_1L_1 + A_2^{\mathrm{T}}P_2L_2). \tag{7.14}$$

This is indeed the accumulation of the two least squares normal equations formed from Eqs. 7.11 and 7.12, respectively:

$$(A_1^{\mathrm{T}}P_1A_1)X = A_1^{\mathrm{T}}P_1L_1 \quad \text{and} \tag{7.15}$$

$$(A_2^{\mathrm{T}}P_2A_2)X = A_2^{\mathrm{T}}P_2L_2. \tag{7.16}$$

The solution is then

$$X = (A_1^{\mathrm{T}}P_1A_1 + A_2^{\mathrm{T}}P_2A_2)^{-1}(A_1^{\mathrm{T}}P_1L_1 + A_2^{\mathrm{T}}P_2L_2). \tag{7.17}$$

The precision of the ith element of the estimated parameter is

$$p[i] = m_0\sqrt{Q[i][i]}, \tag{7.18}$$

where

$$m_0 = \sqrt{\frac{V^{\mathrm{T}}PV}{m-n}}, \quad \text{if} \quad (m > n), \quad \text{and} \tag{7.19}$$

$$Q = (A_1^{\mathrm{T}}P_1A_1 + A_2^{\mathrm{T}}P_2A_2)^{-1}, \tag{7.20}$$

where m is the number of total observations and n is the number of unknowns. And $V^{\mathrm{T}}PV$ can be calculated by using

$$\begin{aligned} V^{\mathrm{T}}PV &= V_1^{\mathrm{T}}P_1V_1 + V_2^{\mathrm{T}}P_2V_2 \\ &= L_1^{\mathrm{T}}P_1L_1 + L_2^{\mathrm{T}}P_2L_2 - (A_1^{\mathrm{T}}P_1L_1)^{\mathrm{T}}X - (A_2^{\mathrm{T}}P_2L_2)^{\mathrm{T}}X. \\ &= (L_1^{\mathrm{T}}P_1L_1 + L_2^{\mathrm{T}}P_2L_2) - (A_1^{\mathrm{T}}P_1L_1 + A_2^{\mathrm{T}}P_2L_2)^{\mathrm{T}}X \end{aligned} \tag{7.21}$$

Equation 7.17 indicates that the sequential least squares problem can be solved by simply accumulating the normal equations of the observation equations. The weighted squares residuals can also be computed by accumulating the individual quadratic forms of the residuals using Eq. 7.21.

For further sequential and independent observation equation systems,

$$V_1 = L_1 - A_1X, \quad P_1, \tag{7.22}$$

$$V_2 = L_2 - A_2X, \quad P_2, \tag{7.23}$$

$$\cdots$$

$$V_i = L_i - A_i X, \quad P_i, \tag{7.24}$$

the solution can be similarly derived as

$$X = (A_1^{\mathrm{T}} P_1 A_1 + A_2^{\mathrm{T}} P_2 A_2 + \cdots + A_i^{\mathrm{T}} P_i A_i)^{-1} (A_1^{\mathrm{T}} P_1 L_1 + A_2^{\mathrm{T}} P_2 L_2 + \cdots + A_i^{\mathrm{T}} P_i L_i) \tag{7.25}$$

and

$$\begin{aligned} V^{\mathrm{T}} P V = {} & (L_1^{\mathrm{T}} P_1 L_1 + L_2^{\mathrm{T}} P_2 L_2 + \cdots + L_i^{\mathrm{T}} P_i L_i) \\ & - (A_1^{\mathrm{T}} P_1 L_1 + A_2^{\mathrm{T}} P_2 L_2 + \cdots + A_i^{\mathrm{T}} P_i L_i)^{\mathrm{T}} X. \end{aligned} \tag{7.26}$$

Obviously, if a solution is needed for every epoch, then the accumulated equation system must be solved at each epoch. The accumulations must always be made with the sequential normal equations. Of course, the solutions can be computed after a defined epoch or at the last epoch, which could be very useful if the solution to the problem is unstable at the beginning.

7.3 Sequential Least Squares Adjustment

Recalling the discussions in Sect. 7.2, one has sequential observation equation systems

$$V_1 = L_1 - A_1 X, \quad P_1 \quad \text{and} \tag{7.27}$$

$$V_2 = L_2 - A_2 X, \quad P_2. \tag{7.28}$$

These two equation systems are uncorrelated. The sequential problem can then be solved by accumulating the individual normal equations as discussed in Sect. 7.2:

$$(A_1^{\mathrm{T}} P_1 A_1 + A_2^{\mathrm{T}} P_2 A_2) X = (A_1^{\mathrm{T}} P_1 L_1 + A_2^{\mathrm{T}} P_2 L_2) \quad \text{or} \tag{7.29}$$

$$X = (A_1^{\mathrm{T}} P_1 A_1 + A_2^{\mathrm{T}} P_2 A_2)^{-1} (A_1^{\mathrm{T}} P_1 L_1 + A_2^{\mathrm{T}} P_2 L_2). \tag{7.30}$$

And $V^{\mathrm{T}} P V$ can be calculated by using

$$V^{\mathrm{T}} P V = (L_1^{\mathrm{T}} P_1 L_1 + L_2^{\mathrm{T}} P_2 L_2) - (A_1^{\mathrm{T}} P_1 L_1 + A_2^{\mathrm{T}} P_2 L_2)^{\mathrm{T}} X. \tag{7.31}$$

If Eq. 7.27 is solvable, the least squares solution can then be represented as

$$X = (A_1^{\mathrm{T}} P_1 A_1)^{-1} (A_1^{\mathrm{T}} P_1 L_1) \quad \text{and} \tag{7.32}$$

$$V^{\mathrm{T}} P V = L_1^{\mathrm{T}} P_1 L_1 - (A_1^{\mathrm{T}} P_1 L_1)^{\mathrm{T}} X. \tag{7.33}$$

For convenience, the estimated vector of X using the first group of observations is denoted by X_1 and the quadratic form of the residuals by $(V^{\mathrm{T}} P V)_1$ as well as $Q_1 = (A_1^{\mathrm{T}} P_1 A_1)^{-1}$.

Using the formula (Cui et al. 1982; Gotthardt 1978)

$$(D + ACB)^{-1} = D^{-1} - D^{-1} AKB\, D^{-1}, \tag{7.34}$$

where A and B are any matrices, C and D are matrices that can be inverted and

$$K = (C^{-1} + B\, D^{-1} A)^{-1}, \tag{7.35}$$

the inversion of the accumulated normal matrix can be represented as Q:

$$\begin{aligned} Q &= (A_1^{\mathrm{T}} P_1 A_1 + A_2^{\mathrm{T}} P_2 A_2)^{-1} \\ &= (A_1^{\mathrm{T}} P_1 A_1)^{-1} - (A_1^{\mathrm{T}} P_1 A_1)^{-1} A_2^{\mathrm{T}} K A_2 (A_1^{\mathrm{T}} P_1 A_1)^{-1} \text{ and} \\ &= Q_1 - Q_1 A_2^{\mathrm{T}} K A_2 Q_1 \\ &= (E - Q_1 A_2^{\mathrm{T}} K A_2) Q_1 \end{aligned} \tag{7.36}$$

$$K = (P_2^{-1} + A_2 Q_1 A_2^{\mathrm{T}})^{-1}, \tag{7.37}$$

where E is an identity matrix. The total term in the parentheses on the right-hand side of Eq. 7.36 can be interpreted as a modifying factor for Q_1 matrix; in other words, due to the sequential Eq. 7.28, the Q matrix can be computed by multiplying a factor to the Q_1 matrix. Thus the sequential least squares solution of Eqs. 7.27 and 7.28 can be obtained:

$$\begin{aligned} X &= (Q_1 - Q_1 A_2^{\mathrm{T}} K A_2 Q_1)(A_1^{\mathrm{T}} P_1 L_1 + A_2^{\mathrm{T}} P_2 L_2). \\ &= (E - Q_1 A_2^{\mathrm{T}} K A_2) X_1 + Q(A_2^{\mathrm{T}} P_2 L_2) \end{aligned} \tag{7.38}$$

Mathematically, the solutions to the sequential problem of Eqs. 7.27 and 7.28 will be the same regardless of whether they are solved using accumulation of the least squares, as discussed in Sect. 7.2.1, or using sequential adjustment, as discussed above. However, in practice, the accuracy of the computation is always limited by the effective digits of the computer being used. Such limitations cause inaccuracy in numerical computation, and this inaccuracy will be accumulated and propagated in further computing processes. By comparing the results obtained with the above-mentioned methods, we note that the sequential method will produce a

drift in the results. This drift will increase with time and will generally become non-negligible after a long time interval.

7.4 Conditional Least Squares Adjustment

The principle of least squares adjustment with condition equations can be summarised as follows (Gotthardt 1978; Cui et al. 1982):

1. The linearised observation equation system can be represented by Eq. 7.1 (cf. Sect. 7.2).
2. The corresponding condition equation system can be written as

$$CX - W = 0, \tag{7.39}$$

where
C coefficient matrix of dimension $r \times n$,
W constant vector of dimension r, and
r number of conditions.

3. The least squares criterion for solving the observation equations with condition equations is well known as

$$V^{\mathrm{T}}PV = \min, \tag{7.40}$$

where V^{T} is the transpose of the related vector V.

4. To solve X and compute V, a function F can be formed as

$$F = V^{\mathrm{T}}PV + 2K^{\mathrm{T}}(CX - W), \tag{7.41}$$

where K is a gain vector (of dimension r) to be determined.
The function F reaches minimum value if the partial differentiation of F with respect to X equals zero, i.e.

$$\frac{\partial F}{\partial X} = 2V^{\mathrm{T}}P(-A) + 2K^{\mathrm{T}}C = 0;$$

then one has

$$-A^{\mathrm{T}}PV + C^{\mathrm{T}}K = 0 \tag{7.42}$$

or

$$A^{\mathrm{T}}PAX + C^{\mathrm{T}}K - A^{\mathrm{T}}PL = 0, \tag{7.43}$$

where A^{T}, C^{T} are transposed matrices of A and C, respectively.

5. Combining Eqs. 7.43 and 7.39, one has

$$A^{\mathrm{T}}PAX + C^{\mathrm{T}}K - A^{\mathrm{T}}PL = 0 \quad \text{and} \tag{7.44}$$

$$CX - W = 0. \tag{7.45}$$

6. For simplification, let $M = A^{\mathrm{T}}PA$, $W_1 = A^{\mathrm{T}}PL$, $Q = M^{-1}$, where superscript $^{-1}$ is an inverse operator. The solutions of Eqs. 7.44 and 7.45 are then

$$\begin{aligned} K &= (CQC^{\mathrm{T}})^{-1}(CQW_1 - W), \\ X &= -Q(C^{\mathrm{T}}K - W_1) \end{aligned} \tag{7.46}$$

or

$$\begin{aligned} X &= (A^{\mathrm{T}}PA)^{-1}(A^{\mathrm{T}}PL) - (A^{\mathrm{T}}PA)^{-1}C^{\mathrm{T}}K. \\ &= (A^{\mathrm{T}}PA)^{-1}(A^{\mathrm{T}}PL - C^{\mathrm{T}}K) \end{aligned} \tag{7.47}$$

7. The precision of the solutions is then

$$p[i] = m_0\sqrt{Q_{\mathrm{c}}[i][i]}, \tag{7.48}$$

where i is the element index of a vector or a matrix, $\sqrt{\ }$ is the square root operator, m_0 is the so-called standard deviation (or sigma), $p[i]$ is the ith element of the precision vector, $Q_{\mathrm{c}}[i][i]$ is the ith diagonal element of the quadratic matrix Q_{c}, and

$$Q_{\mathrm{c}} = Q - QC^{\mathrm{T}}Q_2CQ, \tag{7.49}$$

$$Q_2 = (CQC^{\mathrm{T}})^{-1} \quad \text{and} \tag{7.50}$$

$$m_0 = \sqrt{\frac{V^{\mathrm{T}}PV}{m - n + r}}, \quad \text{if} \quad (m > n - r). \tag{7.51}$$

8. For convenience of sequential computation, $V^{\mathrm{T}}PV$ can be calculated using

$$V^{\mathrm{T}}PV = L^{\mathrm{T}}PL - (A^{\mathrm{T}}PL)^{\mathrm{T}}X - W^{\mathrm{T}}K. \tag{7.52}$$

This can be obtained by substituting Eq. 7.1 into $V^{\mathrm{T}}PV$ and using the relations of Eqs. 7.39 and 7.42.

Thus far, we have derived the complete formulas of conditional least squares adjustment.

7.4.1 Sequential Application of Conditional Least Squares Adjustment

Recalling the least squares adjustment discussed in Sect. 7.2, the linearised observation equation system

$$V = L - AX, \quad P \tag{7.53}$$

has the solution

$$X = (A^{\mathrm{T}}PA)^{-1}(A^{\mathrm{T}}PL). \tag{7.54}$$

The precision of the solutions can be obtained by

$$p[i] = m_0\sqrt{Q[i][i]}, \tag{7.55}$$

where

$$m_0 = \sqrt{\frac{V^{\mathrm{T}}PV}{m-n}}, \quad \text{if} \quad (m > n), \tag{7.56}$$

and $V^{\mathrm{T}}PV$ can be calculated by using

$$V^{\mathrm{T}}PV = L^{\mathrm{T}}PL - (A^{\mathrm{T}}PL)^{\mathrm{T}}X. \tag{7.57}$$

For convenience, the least squares solution vector is denoted by X_0 and weighted residuals square by $(V^{\mathrm{T}}PV)_0$.

Similarly, in the conditional least squares adjustment discussed in Sect. 7.4, the linearised observation equation system and conditional equations read

$$V = L - AX \quad \text{and} \tag{7.58}$$

$$CX - W = 0; \tag{7.59}$$

the solution follows

$$X = (A^{\mathrm{T}}PA)^{-1}(A^{\mathrm{T}}PL - C^{\mathrm{T}}K), \tag{7.60}$$

where K is the gain, and

$$K = (CQC^{\mathrm{T}})^{-1}(CQW_1 - W). \tag{7.61}$$

The precision vector of the solution vector can be obtained by using Eqs. 7.48–7.52. Using the notations obtained in least squares solution, one has

$$X = X_0 - QC^{\mathrm{T}}K \tag{7.62}$$

and

$$V^{\mathrm{T}}PV = (V^{\mathrm{T}}PV)_0 + (A^{\mathrm{T}}PL)^{\mathrm{T}}QC^{\mathrm{T}}K - W^{\mathrm{T}}K. \tag{7.63}$$

Equation 7.62 indicates that the conditional least squares problem can be solved first without the conditions, and then through the gain K to compute a modification term. The change of the solution is caused by the conditions. For computing the weighted squares of the residuals, Eq. 7.63 can be used (by adding two modification terms to the weighted squares of residuals of the least squares solution). This property is very important for many practical applications such as ambiguity fixing or coordinates fixing. For example, after the least squares solution and fixing the ambiguity values, one needs to compute the ambiguity fixed solution. Of course, one can put the fixed ambiguities as known parameters and go back to solve the problem once again. However, using the above formulas, one can use the fixed ambiguities as conditions to compute the gain and the modification terms to get the ambiguity fixed solution directly. Similarly, this property can be also used for solutions with some fixed station coordinates.

7.5 Block-Wise Least Squares Adjustment

The principle of block-wise least squares adjustment can be summarised as follows (Gotthardt 1978; Cui et al. 1982):

1. The linearised observation equation system can be represented by Eq. 7.1 (cf. Sect. 7.2).
2. The unknown vector X and observable vector L are rewritten as two sub-vectors:

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} - \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}. \tag{7.64}$$

The least squares normal equation can then be formed as

$$\begin{aligned} &\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{\mathrm{T}} \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}. \end{aligned} \tag{7.65}$$

The normal equation can be denoted by

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{7.66}$$

or

$$M_{11}X_1 + M_{12}X_2 = B_1 \quad \text{and} \tag{7.67}$$

$$M_{21}X_1 + M_{22}X_2 = B_2, \tag{7.68}$$

where

$$M_{11} = A_{11}^{\mathrm{T}} P_1 A_{11} + A_{21}^{\mathrm{T}} P_2 A_{21}, \tag{7.69}$$

$$M_{12} = M_{21}^{\mathrm{T}} = A_{11}^{\mathrm{T}} P_1 A_{12} + A_{21}^{\mathrm{T}} P_2 A_{22}, \tag{7.70}$$

$$M_{22} = A_{12}^{\mathrm{T}} P_1 A_{12} + A_{22}^{\mathrm{T}} P_2 A_{22}, \tag{7.71}$$

$$B_1 = A_{11}^{\mathrm{T}} P_1 L_1 + A_{21}^{\mathrm{T}} P_2 L_2 \quad \text{and} \tag{7.72}$$

$$B_2 = A_{12}^{\mathrm{T}} P_1 L_1 + A_{22}^{\mathrm{T}} P_2 L_2. \tag{7.73}$$

3. Normal Eqs. 7.67 and 7.68 can be solved as follows: from Eq. 7.67, one has

$$X_1 = M_{11}^{-1}(B_1 - M_{12}X_2). \tag{7.74}$$

Substituting X_1 into Eq. 7.68, one gets a normal equation related to the second block of unknowns:

$$M_2 X_2 = R_2, \tag{7.75}$$

where

$$M_2 = M_{22} - M_{21}M_{11}^{-1}M_{12} \quad \text{and} \tag{7.76}$$

$$R_2 = B_2 - M_{21}M_{11}^{-1}B_1. \tag{7.77}$$

The solution of Eq. 7.75 is then

$$X_2 = M_2^{-1}R_2. \tag{7.78}$$

From Eqs. 7.78 and 7.74, the block-wise least squares solution of Eqs. 7.1 and 7.64 can be computed. For estimating the precision of the solved vector, one has (see discussion in Sect. 7.2):

$$p[i] = m_0 \sqrt{Q[i][i]} \tag{7.79}$$

where

$$m_0 = \sqrt{\frac{V^{\mathrm{T}} PV}{m-n}}, \quad \text{if} \quad (m > n). \tag{7.80}$$

Q is the inversion of the total normal matrix M. m is the number of total observations, and n is the number of unknowns.

Furthermore,

$$Q = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}^{-1} = \begin{pmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{pmatrix} \quad \text{is denoted,} \tag{7.81}$$

where (Gotthardt 1978; Cui et al. 1982)

$$Q_{11} = (M_{11} - M_{12} M_{22}^{-1} M_{21})^{-1}, \tag{7.82}$$

$$Q_{22} = (M_{22} - M_{21} M_{11}^{-1} M_{12})^{-1}, \tag{7.83}$$

$$Q_{12} = M_{11}^{-1}(-M_{12} Q_{22}), \text{ and} \tag{7.84}$$

$$Q_{21} = M_{22}^{-1}(-M_{21} Q_{11}). \tag{7.85}$$

And $V^{\mathrm{T}} PV$ can be calculated by using

$$V^{\mathrm{T}} PV = L^{\mathrm{T}} PL - (A^{\mathrm{T}} PL)^{\mathrm{T}} X. \tag{7.86}$$

One finds very important applications in GPS data processing by separating the unknowns into two groups, which will be discussed in the next sub-section.

7.5.1 Sequential Solution of Block-Wise Least Squares Adjustment

Suppose one has two sequential observation equation systems

$$V_{t1} = L_{t1} - A_{t1} Y_{t1} \quad \text{and} \tag{7.87}$$

$$V_{t2} = L_{t2} - A_{t2} Y_{t2}, \tag{7.88}$$

with weight matrices P_{t1} and P_{t2}. The unknown vector Y can be separated into two sub-vectors; one is sequence-dependent and the other is time-independent. Let us assume

$$Y_{t1} = \begin{pmatrix} X_{t1} \\ X_2 \end{pmatrix} \quad \text{and} \quad Y_{t2} = \begin{pmatrix} X_{t2} \\ X_2 \end{pmatrix}, \tag{7.89}$$

where X_2 is the common unknown vector, and X_{t1} and X_{t2} are sequential (time) independent unknowns (i.e. they are different from each other).

Equations 7.87 and 7.88 can be solved separately using the block-wise least squares method, as follows (cf. Sect. 7.5):

$$X_{t1} = (M_{11})_{t1}^{-1}(B_1 - M_{12}X_2)_{t1}, \tag{7.90}$$

$$(M_2)_{t1}X_2 = (R_2)_{t1} \quad \text{and} \tag{7.91}$$

$$X_2 = (M_2)_{t1}^{-1}(R_2)_{t1}, \tag{7.92}$$

and

$$X_{t2} = (M_{11})_{t2}^{-1}(B_1 - M_{12}X_2)_{t2}, \tag{7.93}$$

$$(M_2)_{t2}X_2 = (R_2)_{t2} \quad \text{and} \tag{7.94}$$

$$X_2 = (M_2)_{t2}^{-1}(R_2)_{t2}, \tag{7.95}$$

where indices $t1$ and $t2$ outside of the parenthesis indicate that the matrices and vectors are related to Eqs. 7.87 and 7.88, respectively.

The combined solution of Eqs. 7.87 and 7.88 then can be derived as

$$X_{t1} = (M_{11})_{t1}^{-1}((B_1)_{t1} - (M_{12})_{t1}(X_2)_{ta}), \tag{7.96}$$

$$X_{t2} = (M_{11})_{t2}^{-1}((B_1)_{t2} - (M_{12})_{t2}(X_2)_{ta}), \tag{7.97}$$

$$((M_2)_{t1} + (M_2)_{t2})(X_2)_{ta} = (R_2)_{t1} + (R_2)_{t2} \quad \text{and} \tag{7.98}$$

$$(X_2)_{ta} = ((M_2)_{t1} + (M_2)_{t2})^{-1}((R_2)_{t1} + (R_2)_{t2}), \tag{7.99}$$

where index ta means that the solution is related to all equations. The normal equations related to the common unknowns are accumulated and solved for. The solved common unknowns are used for computing sequentially different unknowns.

In the case of many sequential observations, a combined solution could be difficult or even impossible because of the large number of unknowns and the requirement of the computing capacities. Therefore, a sequential solution could be a good alternative. For the sequential observation equations

$$V_{t1} = L_{t1} - A_{t1}Y_{t1}, \quad P_{t1}, \tag{7.100}$$

$$V_{ti} = L_{ti} - A_{ti}Y_{ti}, \quad P_{ti}, \tag{7.101}$$

the sequential solutions are

$$X_{t1} = (M_{11})_{t1}^{-1}(B_1 - M_{12}X_2)_{t1}, \tag{7.102}$$

$$(M_2)_{t1}X_2 = (R_2)_{t1}, \tag{7.103}$$

$$X_2 = (M_2)_{t1}^{-1}(R_2)_{t1}, \tag{7.104}$$

$$X_{ti} = (M_{11})_{ti}^{-1}((B_1)_{ti} - (M_{12})_{ti}X_2), \tag{7.105}$$

$$((M_2)_{t1} + \cdots + (M_2)_{ti})X_2 = (R_2)_{t1} + \cdots + (R_2)_{ti}, \quad \text{and} \tag{7.106}$$

$$X_2 = ((M_2)_{t1} + \cdots + (M_2)_{ti})^{-1}((R_2)_{t1} + \cdots + (R_2)_{ti}). \tag{7.107}$$

It is notable that the sequential solution of the second unknown sub-vector X_2 is exactly the same as the combined solution at the last step. The only difference between the combined solution and the sequential solution is that the X_2 used are different. In the sequential solution, only the up-to-date X_2 is used. Therefore, at end of the sequential solution (Eq. 7.107), the last obtained X_2 has to be substituted into all X_{tj} computing formulas, where $j < i$. This can be done in two ways. The first way is to remember all formulas for computing X_{tj}, after X_2 is obtained from Eq. 7.107, using X_2 to compute X_{tj}. The second way is to go back to the beginning after the X_2 is obtained, and use X_2 as the known vector to solve X_{tj} once again. In these ways, the combined sequential observation equations can be solved exactly in a sequential way.

7.5.2 Block-Wise Least Squares for Code–Phase Combination

Recalling the block-wise observation equations discussed in Sect. 7.5, one has

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} - \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix}. \tag{7.108}$$

Such an observation equation can be used for solving the problem of code-phase combination. Supposing L_1 and L_2 are phase and code observation vectors, respectively, and they have the same dimensions, then X_2 is a sub-vector that only exists in phase observation equations. Then one has $A_{22} = 0$, and $A_{11} = A_{21}$, as well as $P_1 = w_p P_0$, $P_2 = w_c P_0$, where P_0 is the weight matrix, and w_p and w_c are weight

factors of phase and code observables. In order to keep the coefficient matrices $A_{11} = A_{21}$, the observable vectors L_1 and L_2 must be carefully scaled. Equation 7.108 can be rewritten as

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \end{pmatrix} - \begin{pmatrix} A_{11} & A_{12} \\ A_{11} & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \quad \text{and} \quad P = \begin{pmatrix} w_\mathrm{p} P_0 & 0 \\ 0 & w_\mathrm{c} P_0 \end{pmatrix}. \tag{7.109}$$

The least squares normal equation can then be formed as

$$\begin{aligned} &\begin{pmatrix} A_{11} & A_{12} \\ A_{11} & 0 \end{pmatrix}^\mathrm{T} \begin{pmatrix} w_\mathrm{P} P_0 & 0 \\ 0 & w_\mathrm{c} P_0 \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ A_{11} & 0 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \\ &= \begin{pmatrix} A_{11} & A_{12} \\ A_{11} & 0 \end{pmatrix}^\mathrm{T} \begin{pmatrix} w_\mathrm{P} P_0 & 0 \\ 0 & w_\mathrm{c} P_0 \end{pmatrix} \begin{pmatrix} L_1 \\ L_2 \end{pmatrix}. \end{aligned} \tag{7.110}$$

The normal equation can be denoted by

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \tag{7.111}$$

where

$$M_{11} = (w_\mathrm{p} + w_\mathrm{c}) A_{11}^\mathrm{T} P_0 A_{11}, \tag{7.112}$$

$$M_{12} = M_{21}^\mathrm{T} = w_\mathrm{p} A_{11}^\mathrm{T} P_0 A_{12}, \tag{7.113}$$

$$M_{22} = w_\mathrm{p} A_{12}^\mathrm{T} P_0 A_{12}, \tag{7.114}$$

$$B_1 = A_{11}^\mathrm{T} P_0 (w_\mathrm{p} L_1 + w_\mathrm{c} L_2), \quad \text{and} \tag{7.115}$$

$$B_2 = w_\mathrm{p} A_{12}^\mathrm{T} P_0 L_1. \tag{7.116}$$

Normal Eq. 7.111 can be solved using the general formulas derived in Sects. 7.2 and 7.5.

7.6 Zhou's Theory: Equivalently Eliminated Observation Equation System

In least squares adjustment, the unknowns can be divided into two groups and then solved in a block-wise manner, as discussed in Sect. 7.5. In practice, sometimes only one group of unknowns is of interest, and it is better to eliminate the other

group of unknowns (called nuisance parameters) because of its size, for example. In this case, using the so-called equivalently eliminated observation equation system could be very beneficial (Wang et al. 1988; Xu and Qian 1986; Zhou 1985). The nuisance parameters can be eliminated directly from the observation equations instead of from the normal equations.

The linearised observation equation system can be represented by

$$V = L - \begin{pmatrix} A & B \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad P. \tag{7.117}$$

where

L	observation vector of dimension m,
A, B	coefficient matrices of dimension $m \times (n - r)$ and $m \times r$,
X_1, X_2	unknown vectors of dimension n–r and r,
V	residual vector of dimension m,
n	number of total unknowns,
m	number of observations, and
P	symmetric and definite weight matrix, of dimension $m \times m$.

The least squares normal equation can then be formed by

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \tag{7.118}$$

where

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} = \begin{pmatrix} A^{\mathrm{T}}PA & A^{\mathrm{T}}PB \\ B^{\mathrm{T}}PA & B^{\mathrm{T}}PB \end{pmatrix}, \tag{7.119}$$

$$B_1 = A^{\mathrm{T}}PL, \; B_2 = B^{\mathrm{T}}PL. \tag{7.120}$$

The elimination matrix

$$\begin{pmatrix} E & 0 \\ -Z & E \end{pmatrix} \quad \text{is formed,} \tag{7.121}$$

where E is the identity matrix, 0 is a zero matrix, and $Z = M_{21}M^{-1}_{11}$. M^{-1}_{11} is the inversion of M_{11}. Multiplying the elimination matrix Eq. 7.121 to the normal Eq. 7.118 one has

$$\begin{pmatrix} E & 0 \\ -Z & E \end{pmatrix} \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} E & 0 \\ -Z & E \end{pmatrix} \begin{pmatrix} B_1 \\ B_2 \end{pmatrix},$$

or

$$\begin{pmatrix} M_{11} & M_{12} \\ 0 & M_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ R_2 \end{pmatrix} \tag{7.122}$$

where

$$\begin{aligned} M_2 &= -M_{21}M_{11}^{-1}M_{12} + M_{22} \\ &= B^{\mathrm{T}}PB - B^{\mathrm{T}}PAM_{11}^{-1}A^{\mathrm{T}}PB = B^{\mathrm{T}}P(E - AM_{11}^{-1}A^{\mathrm{T}}P)B. \end{aligned} \tag{7.123}$$

$$R_2 = B_2 - M_{21}M_{11}^{-1}B_1 = B^{\mathrm{T}}P(E - AM_{11}^{-1}A^{\mathrm{T}}P)L. \tag{7.124}$$

If we are interested only in the unknown vector X_2, then only the second equation of Eq. 7.122 needs to be solved. The solution is identical to that obtained by solving all of Eq. 7.122. The above elimination process is similar to the Gauss-Jordan algorithm, which has often been used for the inversion of the normal matrix (or for solving linear equation systems). Indeed, the second equation of Eq. 7.122 is identical to Eq. 7.75 derived in the block-wise least squares adjustment (cf. Section 7.5).

Letting

$$J = AM_{11}^{-1}A^{\mathrm{T}}P, \tag{7.125}$$

one has properties of

$$\begin{aligned} &J^2 = (AM_{11}^{-1}A^{\mathrm{T}}P)(AM_{11}^{-1}A^{\mathrm{T}}P) = AM_{11}^{-1}A^{\mathrm{T}}PAM_{11}^{-1}A^{\mathrm{T}}P = AM_{11}^{-1}A^{\mathrm{T}}P = J, \\ &(E - J)(E - J) = E^2 - 2EJ + J^2 = E - 2J + J = E - J \quad \text{and} \\ &[P(E - J)]^{\mathrm{T}} = (E - J^{\mathrm{T}})P = P - (AM_{11}^{-1}A^{\mathrm{T}}P)^{\mathrm{T}}P = P - PAM_{11}^{-1}A^{\mathrm{T}}P = P(E - J), \end{aligned}$$

i.e. matrices J and $(E–J)$ are idempotent and $(E–J)^{\mathrm{T}}P$ is symmetric, or

$$J^2 = J, \quad (E - J)^2 = E - J \quad \text{and} \quad (E - J)^{\mathrm{T}}P = P(E - J). \tag{7.126}$$

Using the above derived properties, M_2 in Eq. 7.123 and R_2 in Eq. 7.124 can be rewritten as

$$M_2 = B^{\mathrm{T}}P(E - J)B = B^{\mathrm{T}}P(E - J)(E - J)B = B^{\mathrm{T}}(E - J)^{\mathrm{T}}P(E - J)B \quad \text{and} \tag{7.127}$$

$$R_2 = B^{\mathrm{T}}P(E - J)L = B^{\mathrm{T}}(E - J)^{\mathrm{T}}PL. \tag{7.128}$$

Denoting

$$D_2 = (E - J)B, \tag{7.129}$$

then the eliminated normal equation (the second equation of Eq. 7.122) can be rewritten as

$$B^{\mathrm{T}}(E - J)^{\mathrm{T}}P(E - J)BX_2 = B^{\mathrm{T}}(E - J)^{\mathrm{T}}PL \quad \text{or} \tag{7.130}$$

$$D_2^{\mathrm{T}}PD_2X_2 = D_2^{\mathrm{T}}PL. \tag{7.131}$$

This is the least squares normal equation of the following linear observation equation:

$$U_2 = L - D_2X_2, \quad P \tag{7.132}$$

or

$$U_2 = L - (E - J)BX_2, \quad P, \tag{7.133}$$

where L and P are the original observation vector and weight matrix, and U_2 is the residual vector, which has the same property as V in Eq. 7.117.

The advantage in using Eq. 7.133 is that the unknown vector X_1 has been eliminated; however, L vector and P matrix remain the same as the originals. Applications of this theory can be found in Sect. 6.8, 8.3, and 9.2. The theory was proposed by Jiangwen Zhou in 1985.

7.6.1 Zhou–Xu's Theory: Diagonalised Normal Equation and the Equivalent Observation Equation

In least squares adjustment, the unknowns can be divided into two groups. One group of unknowns can be eliminated by matrix partitioning to obtain an equivalently eliminated normal equation system of the other group of unknowns. Using the elimination process twice for the two groups of unknowns respectively, the normal equation can be diagonalised. The algorithm can be outlined as follows.

A linearised observation equation and the normal equations can be represented by Eqs. 7.117 and 7.118. From the first equation of 7.118, one has

$$X_1 = M_{11}^{-1}(B_1 - M_{12}X_2). \tag{7.134}$$

Setting X_1 into the second equation of 7.118, one gets an equivalently eliminated normal equation of X_2:

$$M_2 X_2 = R_2, \tag{7.135}$$

where

$$\begin{aligned} M_2 &= M_{22} - M_{21} M_{11}^{-1} M_{12}. \\ R_2 &= B_2 - M_{21} M_{11}^{-1} B_1 \end{aligned} \tag{7.136}$$

Similarly, from the second equation of 7.118, one has

$$X_2 = M_{22}^{-1}(B_2 - M_{21} X_1). \tag{7.137}$$

Setting X_2 into the first equation of 7.118, one gets an equivalently eliminated normal equation of X_1:

$$M_1 X_1 = R_1, \tag{7.138}$$

where

$$\begin{aligned} M_1 &= M_{11} - M_{12} M_{22}^{-1} M_{21}. \\ R_1 &= B_1 - M_{12} M_{22}^{-1} B_2 \end{aligned} \tag{7.139}$$

Combining Eqs. 7.138 and 7.135, one has

$$\begin{pmatrix} M_1 & 0 \\ 0 & M_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \end{pmatrix}, \tag{7.140}$$

where (cf., e.g., Cui et al. 1982; Gotthardt 1978)

$$\begin{aligned} Q_{11} &= M_1^{-1}, & Q_{22} &= M_2^{-1} \\ Q_{12} &= -M_{11}^{-1}(M_{12} Q_{22}), & Q_{21} &= -M_{22}^{-1}(M_{21} Q_{11}) \end{aligned}. \tag{7.141}$$

It is obvious that Eqs. 7.118 and 7.140 are two equivalent normal equations. The solutions of the both equations are identical. Equation 7.140 is a diagonalised normal equation related to X_1 and X_2. The process of forming Eq. 7.140 from Eq. 7.118 is called the diagonalisation process of a normal equation.

As discussed in Sect. 7.6, the equivalently eliminated observation equation of the second equation of Eq. 7.140 is Eq. 7.133. Similarly, if

$$\begin{aligned} I &= B M_{22}^{-1} B^{\mathrm{T}} P \text{ and} \\ D_1 &= (E - I) A, \end{aligned}$$

then the equivalently eliminated observation equation of the first normal equation of Eq. 7.140 has the form

$$U_1 = L - (E - I)AX_1, \quad P,$$

where U_1 is a residual vector that has the same property as V in Eq. 7.117. L and P are the original observation vector and weight matrix.

The above equation and Eq. 7.133 can be written together as

$$\begin{pmatrix} U_1 \\ U_2 \end{pmatrix} = \begin{pmatrix} L \\ L \end{pmatrix} - \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix}. \tag{7.142}$$

Equation 7.142 is derived from the normal Eq. 7.140; therefore, it is true inversely, i.e. Equation 7.140 is the least squares normal equation of the observation Eq. 7.142. Equations 7.118 and 7.140 are normal equations of the observation Eqs. 7.117 and 7.142. Thus, Eq. 7.142 is an equivalent observation equation of Eq. 7.117. Equations 7.140 and 7.142 are called diagonalised equations of 7.118 and 7.117, respectively. This diagonalised normal equation and the equivalent observation equation could be called Zhou–Xu diagonalisation and equivalent theory (Xu 2003).

7.7 Kalman Filter

7.7.1 Classic Kalman Filter

The principle of the classical Kalman filter can be summarised as follows (Yang et al. 1999):

The linearised observation equation system can be represented by

$$V_i = L_i - A_i X_i, \quad P_i, \tag{7.143}$$

where

L observation vector of dimension m,
A coefficient matrix of dimension $m \times n$,
X unknown vector of dimension n,
V residual vector of dimension m,
n number of unknowns,
m number of observations,
i sequential index, i = 1,2,3,…, and
P_i weight matrix of index i.

Suppose the system equations are known and can be presented as

$$U_i = X_i - F_{i,i-1}X_{i-1}, \quad i = 2, 3, \ldots, \tag{7.144}$$

where
F transition matrix of dimension $n \times n$, and
U residual vector of dimension n.

U and V are uncorrelated and have zero expectations. Using the covariance propagation law, one has from Eq. 7.144

$$Q(X_i) = F_{i,i-1}Q(X_{i-1})(F_{i,i-1})^{\mathrm{T}} + Q_U. \tag{7.145}$$

The normal Eq. 7.143 can be formed as

$$M_i X_i = B_i. \tag{7.146}$$

For the initial step or epoch, i.e. $i = 1$, Eq. 7.146 has the solution under the least squares principle

$$\tilde{X}_i = Q_i B_i, \quad \text{where} \quad Q_i = M_i^{-1}, \tag{7.147}$$

and here one will assume

$$\tilde{Q}_i = Q_i, \tag{7.148}$$

where $\tilde{X}_i$ and $\tilde{Q}_i$ are called estimated values. Using the estimated values and transition matrix, one can predict the unknown values and covariance matrix of the next epoch (say $i = 2$):

$$\underline{X}_i = F_{i,i-1}\tilde{X}_{i-1} \quad \text{and} \tag{7.149}$$

$$\underline{Q}_i = F_{i,i-1}\tilde{Q}_{i-1}(F_{i,i-1})^{\mathrm{T}} + Q_U, \tag{7.150}$$

where $\underline{X}_i$ and $\underline{Q}_i$ are called predicted values (vector and matrix). Then estimated values of this epoch can be calculated by

$$\tilde{X}_i = \underline{X}_i + K(L_i - A_i\underline{X}_i), \tag{7.151}$$

$$\tilde{Q}_i = (E - KA_i)\underline{Q}_i, \quad \text{and} \tag{7.152}$$

$$K = \underline{Q}_i A_i^{\mathrm{T}}(A_i\underline{Q}_i A_i^{\mathrm{T}} + Q_V)^{-1}, \tag{7.153}$$

where K is the gain matrix.

For the next sequential step i, the predicted values must be computed by using Eqs. 7.149 and 7.150, and the estimated values can be computed by using Eqs. 7.151 and 7.152. This iterative process is called Kalman filtering.

In classical Kalman filtering, it is assumed that for the problem of Eq. 7.143 there exists a system transition matrix $F_{i,i-1}$ in Eq. 7.144 and the cofactor Q_{U}. Therefore, the estimated values in the Kalman filter process are dependent on $F_{i,i-1}$ and Q_{U}. The transition matrix will be based on strengthened physical models, and the cofactor will be well known or reasonably given. If the system description is accurate enough, of course Kalman filtering will lead to a more precise solution. However, if the system is not sufficiently well known, the results of Kalman filter will sometimes not converge to the true values (divergence). Furthermore, a kinematic process is generally difficult to be precisely represented by theoretical system equations. However, for a dynamic process (such as onboard GPS for satellite to satellite tracking or orbit determination) the system equation can be well formulated (by an orbital equation of motion). Another problem of Kalman filtering is the strong dependency of the given initial values. Many studies have been made in this area to overcome the above-mentioned shortages.

7.7.2 Kalman Filter: A General Form of Sequential Least Squares Adjustment

The sequential least squares problem is a special case of the classic Kalman filter. If one lets

$$F_{i,i-1} = E, \tag{7.154}$$

then the system Eq. 7.144 in Sect. 7.7.1 turns out to be

$$X_i = X_{i-1}, \quad U = 0 \quad \text{and} \quad Q_U = 0. \tag{7.155}$$

The Kalman filter process is then as follows, for the initial step or epoch, i.e. $i = 1$, Eq. 7.27 in Sect. 7.3 has the solution under the least squares principle:

$$\tilde{X}_i = Q_i B_i, Q_i = M_i^{-1}, \tag{7.156}$$

with

$$\tilde{Q}_i = Q_i, \tag{7.157}$$

where $\tilde{X}_i$ and $\tilde{Q}_i$ are called estimated values. The predicted unknown values and covariance matrix of the next epoch (say $i = 2$) of Eqs. 7.149 and 7.150 in Sect. 7.7.1 are then

$$\underline{X}_i = \tilde{X}_{i-1} \text{ and} \tag{7.158}$$

$$\underline{Q}_i = \tilde{Q}_{i-1}. \tag{7.159}$$

The estimated values of Eqs. 7.151, 7.152 and 7.153 in Sect. 7.7.1 can be simplified as

$$\tilde{X}_i = \tilde{X}_{i-1} + G(L_i - A_i\tilde{X}_{i-1}), \tag{7.160}$$

$$\tilde{Q}_i = (E - GA_i)\tilde{Q}_{i-1}, \quad \text{and} \tag{7.161}$$

$$G = \tilde{Q}_{i-1}A_i^{\mathrm{T}}(A_i\tilde{Q}_{i-1}A_i^{\mathrm{T}} + Q_V)^{-1}, \tag{7.162}$$

where G denotes the gain matrix. If one notices that $Q_V = (P_i)^{-1}$ and applies the formula of Bennet (Cui et al. 1982; Koch 1986), one has

$$\tilde{Q}_{i-1}A_i^{\mathrm{T}}(A_i\tilde{Q}_{i-1}A_i^{\mathrm{T}} + Q_V)^{-1} = \tilde{Q}_{i-1}A_i^{\mathrm{T}}P_i. \tag{7.163}$$

Equation 7.160 can then be rewritten as

$$\begin{aligned}\tilde{X}_i &= (E - GA_i)\tilde{X}_{i-1} + GL_i. \\ &= (E - GA_i)\tilde{X}_{i-1} + \tilde{Q}_iA_i^{\mathrm{T}}P_iL_i\end{aligned} \tag{7.164}$$

Comparing the derived Eqs. 7.161 and 7.164 with the Eqs. 7.36 and 7.38 derived in Sect. 7.3, one can easily determine that they are identical. Therefore, the sequential least squares adjustment is a special case of Kalman filtering.

7.7.3 Robust Kalman Filter

The classical Kalman filter is suitable for real-time applications. The chief problem in Kalman filtering is the divergence caused by the inexact descriptions of system equations and its statistical properties, as well as the divergence caused by data with inhomogeneous precision.

Efforts have been made to modify the performance of Kalman filtering. In the classical Kalman filter, the weight matrix P of the observables is static, i.e. P is assumed to be a definite matrix. Taking the residuals of Kalman filtering into account, one may adjust the weight P of the observables accordingly. This process is called robust Kalman filtering (Koch and Yang 1998a, b; Yang 1999).

Generally, observations are either accepted or rejected in least squares adjustment and the classical Kalman filter. In other words, the weight is either set as 1

(accepted) or zero (rejected). In the robust Kalman filter, a continuous weight between 1 and zero is introduced.

Originally, one has $P = (Q_V)^{-1}$, the adjusted P is denoted by $\overline{P}$; then the Eq. 7.153 in the classical Kalman filter can be rewritten as

$$K = \underline{Q}_i A_i^{\mathrm{T}} (A_i \underline{Q}_i A_i^{\mathrm{T}} + \bar{P}_i^{-1})^{-1}. \tag{7.165}$$

In the case of independent observations, P_i is a diagonal matrix. Taking the residuals into account, P_i may be adjusted as (Huber 1964; Yang et al. 2000)

$$\bar{P}_i(k) = \begin{cases} P_i(k) \\ P_i(k)\frac{c}{|V_i(k)/\sigma_i|} \end{cases}, \quad \text{if} \quad \begin{matrix} |V_i(k)/\sigma_i| \le c \\ |V_i(k)/\sigma_i| > c \end{matrix}, \tag{7.166}$$

where $V_i(k)$ is the kth element of the vector V, $P_i(k)$ is the diagonal element of matrix P_i, and c is a constant, which is usually chosen as 1.3–2.0 (Yang et al. 2000). V_i is the residual of the observation L_i, σ_i is the standard deviation of the ith epoch, and $P_i = 1/\sigma_i$. In this way, the weight of the observation L_i is adjusted due to the related residual.

If the observations are correlated with each other, the weight matrix may be given by (Yang et al. 2000)

$$\bar{P}_{kj} = \begin{cases} P_{kj} \\ P_{kj}\frac{c}{\max\{|V_i(k)/\sigma_i|,|V_i(j)/\sigma_i|\}} \end{cases}, \quad \text{if} \quad \begin{matrix} |V_i(k)/\sigma_i| \le c & \text{and} & |V_i(j)/\sigma_i| \le c \\ |V_i(k)/\sigma_i| > c & \text{or} & |V_i(j)/\sigma_i| > c \end{matrix}. \tag{7.167}$$

It is obvious that an adjusted weight matrix can better reflect the different data quality and can better fit the reality of the observations.

Usually the outlier will be rejected if the absolute value of the residual is greater than $e\sigma_i$, i.e. $|V_i| > e\sigma_i$, where e is a constant, e may be selected as 3–4, σ_i is the standard deviation, and i is the iterative calculation index. That is, $P_{-i} = 0$ if $|V_i/\sigma_i| \ge e$. Setting $|V_i/\sigma_i| = e$ into Eq. 7.166 one gets $\overline{P}_i = (c/e)P_i$. In other words, the weight definitions of Eqs. 7.166 and 7.167 are not continuous at point e. A modification (Xu 2003) of Eq. 7.166 can be made by defining

$$\bar{P}_i(k) = \begin{cases} p_i(k) \\ y_1 P_i(k) \\ y_2 P_i(k) \\ 0 \end{cases}, \quad \text{if} \quad \begin{cases} |V_i(k)/\sigma_i| \le c \\ c < |V_i(k)/\sigma_i| \le d \\ d < |V_i(k)/\sigma_i| \le e \\ |V_i(k)/\sigma_i| \le e \end{cases}, \tag{7.168}$$

where

$$y_1 = 1 - \frac{1-b}{(d-c)^2}\left(\left|\frac{V_i(k)}{\sigma_i}\right| - c\right)^2 \quad \text{and} \tag{7.169}$$

$$y_2 = \frac{b}{(e-d)^2}\left(e - \left|\frac{V_i(k)}{\sigma_i}\right|\right)^2, \tag{7.170}$$

where b is the value of y_1 if $|V_i(k)/\sigma_i| = d$. c, d, e are constants, and $0 < c < d < e$. For simplification, if one lets $b = (e–d)/(e–c)$, then one has $1 - b = (d - c)/(e - c)$. One may let $d = (e + c)/2$ for further simplification and have

$$y_1 = 1 - \frac{2}{(e-c)^2}\left(\left|\frac{V_i(k)}{\sigma_i}\right| - c\right)^2 \quad \text{and}$$

$$y_2 = \frac{2}{(e-c)^2}\left(e - \left|\frac{V_i(k)}{\sigma_i}\right|\right)^2.$$

By selecting $c = 1$, $e = 3$, and using the above assumptions, the weight functions of Eqs. 7.166 and 7.168 are shown in Fig. 7.1 with broken and continuous lines. It is obvious that Eq. 7.168 is a more reasonable weight function, which may make the Kalman filter more robust.

Similar determinations can be made similarly for correlated cases. Denoting $|V_i(k)/\sigma_i|$ as $v(k)$, a modification of Eq. 7.167 can be rewritten as (Xu 2007)

$$\bar{P}_i(k,j) = \begin{cases} p_i(k,j) \\ z_1 P_i(k,j) \\ z_2 P_i(k,j) \\ 0 \end{cases}, \quad \text{if} \quad \begin{cases} \max\{v(k), v(j)\} \le c \\ c < \max\{v(k), v(j)\} \le d \\ d < \max\{v(k), v(j)\} \le e \\ \max\{v(k), v(j)\} > e \end{cases}, \tag{7.171}$$

where

$$z_1 = 1 - \frac{1-b}{(d-c)^2}(\max\{v(k), v(j)\} - c)^2 \quad \text{and} \tag{7.172}$$

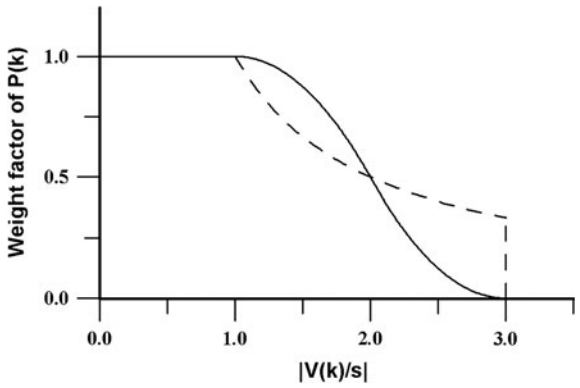

Fig. 7.1 Weight functions

$$z_2 = \frac{b}{(e-d)^2}(e - \max\{v(k), v(j)\})^2, \tag{7.173}$$

where b is the value of z_1 if $\max\{v(k), v(j)\} = d$. For simplification, if one lets $b = (e - d)/(e - c)$, then one has $1 - b = (d - c)/(e - c)$. Further if one lets $d = (e - c)/2$, then one has

$$z_1 = 1 - \frac{2}{(e-c)^2}(\max\{v(k), v(j)\} - c)^2 \quad \text{and}$$
$$z_2 = \frac{2}{(e-c)^2}(e - \max\{v(k), v(j)\})^2.$$

7.7.4 Yang's Filter: Adaptively Robust Kalman Filtering

The reliability of the linear filtering results, however, will degrade when the noise of the kinematic model is not accurately modelled in filtering or the measurement noise at any measurement epoch is not normally distributed. In this section, we introduce a new adaptively robust filtering technique proposed by Yang et al. (2001a, b) based on the robust M (maximum likelihood type) estimation. It consists in weighting the influence of the updated parameters in accordance with the magnitude of the discrepancy between the updated parameters and the robust estimates obtained from the kinematic measurements, and in weighting individual measurements at each discrete epoch. The new procedure is different from functional model error compensation; it changes the covariance matrix or, equivalently, changes the weight matrix of the predicted parameters to cover the model errors. A general estimator for an adaptively robust filter is presented, which includes the estimators of the classical Kalman filter, adaptive Kalman filter, robust filter, sequential least squares (LS) adjustment, and robust sequential adjustment. The procedure not only resists the influence of outlying kinematic model errors, but also controls the effects of measurement outliers. In addition to the robustising properties, feasibility in implementation of the new filter is achieved through the equivalent weights of the measurements and the predicted state parameters.

Applications of the Kalman filter in dynamic or kinematic positioning have sometimes encountered difficulties, which have been referred to as divergences. These divergences can often be traced to three factors: (1) insufficient accuracy in modelling the dynamics or kinematics (functional model errors of the state equations); (2) insufficient accuracy in modelling the observations (functional model errors of observation equations); and (3) insufficient accuracy in modelling the distributions or the priori covariance matrices of the measurements and the updated parameters (stochastic model errors).

The current basic procedure for the quality control of a Kalman filter consists of the following:

- Functional model compensation for model errors by introducing uncertain parameters into the state and/or the observation equations. Any model error term can be arbitrarily introduced into the models, and the state can then be augmented (Jazwinski 1970, p. 308). A similar approach was developed by Schaffrin (1991, pp. 32–34). Here, the state vector is partitioned into h groups, each affected by a common scale error, and $h \times 1$ vectors of scale parameters are then introduced into the models. This type of approach may, of course, lead to a high-dimensional state vector, which in turn greatly increases the filter computational load (Jazwinski 1970, p. 305).
- Stochastic model compensation by introducing a variance–covariance matrix of the model errors. In taking this approach to prevent divergence, one must determine which covariance matrix to add. A reasonable covariance matrix may compensate for the model errors. An ineffective covariance matrix, however, adds to the model divergence. For instance, when the model is accurate in some dynamic or kinematic periods, an unsuitable increase of the covariance matrix of model error will degrade the state estimators. Thus an effective covariance matrix for model errors can be determined only by trial and error.
- The DIA procedure—detection, identification, and adaptation (Teunissen 1990). This approach employs a recursive testing procedure to eliminate outliers. In the detection step, one looks for unspecified model errors, and in the identification step, one tries to find the cause of the model error and its most likely starting time. After a model error has been detected and identified, the bias in the state estimate caused by the model error must be eliminated as well. This model recovery from errors is called adaptation (Salzmanm 1995). The identification of the model, however, is quite difficult, especially when the measurements are not accurate enough to detect the unspecified model errors.
- The sequential least squares procedure. A rather different procedure frequently used for kinematic positioning does not use the dynamic model information at all, but determines discrete positions at the measurement epochs (Cannon et al. 1986). In this case, there is no assumption made on a dynamic model, and only the measurements at the discrete epoch are employed to estimate the state parameters. The model error, therefore, does not affect the estimates of new state parameters. This method is typically presented as a sequential least squares algorithm (Schwarz et al. 1989). The current limitation to this approach is that it wastes the good information of the state model in cases when the model accurately describes the dynamic process.
- Adaptive Kalman filtering. An innovation-based adaptive Kalman filter for an integrated INS/GPS was developed by Mohamed and Schwarz (1999), based on the maximum likelihood criterion by proper choice of the filter weight. Another adaptive Kalman filter algorithm to directly estimate the variance and covariance components for the measurements was studied by Wang et al. (1999). Both

algorithms need to collect the residuals of the measurements or the updated series to calculate the state variance–covariance matrices.

- A robust filter based on the min–max robust theory. The deviation of observation error distribution from Gaussian distribution may also seriously degrade the performance of Kalman filtering. Thus, there appears to be considerable motivation for considering filters which are robustised to perform fairly well in non-Gaussian environments. To address this problem, Masreliez and Martin (1977) applied the influence function of the min–max robust theory to replace the score function of the classical Kalman filter. The key disadvantages with this kind of robust filter are that the estimator requires symmetric distribution of the unknown contamination, and this filter does not work as well as the standard Kalman filter in Gaussian noise.
- A robust filter based on M estimation theory (Huber 1964) and Bayesian statistics. To resist the negative influence of both state model errors and measurement outliers, a robust M–M filter was developed (Yang 1991, 1997a, b; Zhou et al. 1997, p. 299). Here, measurement outliers are controlled by robust equivalent weights of the measurements, and the model errors are resisted by the equivalent weights of the updated parameters according to the divergence of the predicted parameters from the estimated parameters. In addition, a robust filter for rank-deficient observation models was developed by Koch and Yang (1998a, b), using Bayesian statistics and applying the robust M estimate.

All of the methods described above require knowledge of the dynamic model errors, with which the functional or stochastic models to compensate for the model errors and the equivalent weights for the robust filter are constructed. In practical applications, it is very difficult to predict the error distribution or the error type of the updated parameters or the dynamic model errors, and thus it is very difficult to construct functional and stochastic models. Furthermore, when a moving vehicle accelerates from zero or decelerates to a stop, the acceleration profile is discontinuous. If this discontinuity falls between two measurement epochs, the dynamics cannot be accurately modelled or predicted by state equations; in this case, one should not rely too heavily on the information predicted from the dynamic model. Thus, the filtering procedure should weaken the effects of the updated parameters. In addition, if the updated parameter vector is contaminated by model error, it is generally distorted in its entirety. Therefore, it is not necessary to consider the error influence of the individual element of the updated parameter vector as is done with the robust M–M filter. In this case, an adaptive filter is suitable for balancing the dynamic model information and the measurements.

1. ***General Estimator of Adaptively Robust Filtering***

An adaptively robust filter is constructed as (cf. Yang et al. 2001a, b)

$$\tilde{X}_i = (A_i^{\mathrm{T}} \bar{P}_i A_i + \alpha P_{\underline{X}_i})^{-1} (A_i^{\mathrm{T}} \bar{P}_i L_i + \alpha P_{\underline{X}_i} \underline{X}_i) \quad \text{and} \tag{7.174}$$

$$Q_{\tilde{X}_i} = (A_i^{\mathrm{T}} \bar{P}_i A_i + \alpha P_{\underline{X}_i})^{-1} \sigma_0^2, \tag{7.175}$$

where $\bar{P}_i$ is the equivalent weight matrix of the observation vector, $P_{\underline{X}_i}$ is the weight matrix of the predicted vector $\underline{X}_i$, $Q_{\tilde{X}_i}$ is the covariance matrix of the estimated state vector, σ_0^2 is a scale factor, and α is an adaptive factor, which can be chosen as

$$\alpha = \begin{cases} 1 & \left|\Delta\tilde{X}_i\right| \le \mathrm{c}_0 \\ \frac{c_0}{\left|\Delta\tilde{X}_i\right|}\left(\frac{c_1 - \left|\Delta\tilde{X}_i\right|}{c_1 - c_0}\right)^2 & \mathrm{c}_0 < \left|\Delta\tilde{X}_i\right| \le c_1 \\ 0 & \left|\Delta\tilde{X}_i\right| > c_1 \end{cases}, \tag{7.176}$$

where c_0 and c_1 are constants that are experienced, valued as c_0 = 1.0–1.5, c_1 = 3.0–4.5,

$$\Delta\tilde{X}_i = \frac{\left\|\hat{X}_i - \underline{\hat{X}}_i\right\|}{\sqrt{tr\{Q_{\underline{\hat{X}}_i}\}}}, \tag{7.177}$$

and $\hat{X}_i$ is a robust estimate of state vector (state position), which is evaluated only by new measurements at epoch i, and the raw velocity observations are not included in it. $\underline{\hat{X}}_i$ is a predicted position from Eq. 7.149 in which the a priori velocity components are not included. The change in the position expressed by Eq. 7.177 can also reflect the stability of the velocity (cf. Yang et al. 2001a, b).

Expression 7.174 is a general estimator of an adaptively robust filter. In the case of $\alpha \neq 0$, Eq. 7.174 is changed, using the matrix identities (Koch 1988, p. 40), into

$$\tilde{X}_i = \underline{X}_i + Q_{\underline{X}_i} A_i^{\mathrm{T}} (A_i Q_{\underline{X}_i} A_i^{\mathrm{T}} + \alpha Q_V)^{-1} (L_i - A_i \underline{X}_i). \tag{7.178}$$

2. ***Special Estimators***

The adaptive factor α changes between zero and one, which balances the contribution of the new measurements and the updated parameters to the new estimates of state parameters.

Case 1: If α = 0 and $\overline{P}_i = P_i$, then

$$\tilde{X}_i = (A_i^{\mathrm{T}} P_i A_i)^{-1} A_i^{\mathrm{T}} P_i L_i, \tag{7.179}$$

which is an LS estimator by using only the new measurements at epoch i. This estimator is suitable in the case where the measurements are not contaminated by outliers and the updated parameters are biased to such a degree that $\Delta\tilde{X}_i$ in Eq. 7.177 is larger than c_1 (rejecting point), and the information of updated parameters is completely forgotten.

Case 2: If $\alpha = 1$ and $\overline{P}_i = P_i$, then

$$\tilde{X}_i = (A_i^{\mathrm{T}} P_i A_i + P_{\underline{X}_i})^{-1} (A_i^{\mathrm{T}} P_i L_i + P_{\underline{X}_i} \underline{X}_i), \tag{7.180}$$

which is a general estimator of the classical Kalman filter.

Case 3: If α is determined by Eq. 7.177 and $\bar{P}_i = P_i$, then

$$\tilde{X}_i = (A_i^{\mathrm{T}} P_i A_i + \alpha P_{\underline{X}_i})^{-1} (A_i^{\mathrm{T}} P_i L_i + \alpha P_{\underline{X}_i} \underline{X}_i), \tag{7.181}$$

which is an adaptive LS estimator of the Kalman filter. It balances the contribution of the updated parameters and the measurements. The only difference between Eqs. 7.174 and 7.181 is the weight matrix of L_i. The former uses the equivalent weights and the latter uses the original weights of L_i.

Case 4: If $\alpha = 0$, we obtain

$$\tilde{X}_i = (A_i^{\mathrm{T}} \bar{P}_i A_i)^{-1} A_i^{\mathrm{T}} \bar{P}_i L_i, \tag{7.182}$$

which is a robust estimator by using only the new measurements at epoch i.

Case 5: If $\alpha = 1$, then

$$\tilde{X}_i = (A_i^{\mathrm{T}} \bar{P}_i A_i + P_{\underline{X}_i})^{-1} (A_i^{\mathrm{T}} \bar{P}_i L_i + P_{\underline{X}_i} \underline{X}_i), \tag{7.183}$$

which is an M–LS filter estimator (Yang 1997a, b).

Further Development of the Theory

The adaptive factor α was considered a diagonal matrix by Ou (2004) and grouped by the physical meaning of the parameters by Yang and Xu (2004). Since then, several advances have been made (cf. Yang and Cui 2006; Yang and Gao 2005a, b, 2006a, b, c, Yang et al. 2006).

7.7.5 Progress in Adaptively Robust Filter Theory and Application

A new adaptively robust filtering technique for use in kinematic navigation and positioning has been systematically established and developed in recent years (Yang et al. 2013). The adaptively robust filter applies a robust estimation principle to resist the effects of measurement outliers, and introduces an adaptive factor to control the influence of dynamic model disturbances. It can thus balance the contribution of the dynamic model information and the measurements in accordance with the magnitude of their discrepancy (Yang et al. 2001a). In this section, we introduce the major advancements in the theory and application of the adaptively robust filter.

Following the development of adaptively robust filtering, four learning statistics and four adaptive factors were established based on experiences, and these have been proven effective in practical applications. An accompanying adaptive factor was created that features a three-segment descending function and a learning statistic constructed using the discrepancy between the predicted state from the kinematic model and the state estimated from the measurements. Three other types of adaptive factors have been developed: a two-segment descending function (Yang et al. 2001b), an exponential function (Yang and Gao 2005), and a zero/one function for state component adaptation (Ou et al. 2004; Ren et al. 2005). Three additional learning statistics have also been set up, which include a predicted residual statistic (Xu and Yang 2000; Yang and Gao 2006b), a variance component ratio statistic from both the measurements and the predicted states (Yang and Xu 2003), and a velocity discrepancy between the predicted velocity from the kinematic model and the velocity evaluated from the measurements (Cui and Yang 2006).

A key problem has been in constructing an adaptive factor suitable for balancing the contribution of the measurements and the predicted dynamic model information. Two optimal adaptive factors have been established that satisfy the conditions that the theoretical uncertainty of the predicted state outputted from the adaptive filtering is equal or nearly equal to its actual estimated uncertainty, or that the theoretical uncertainty of the predicted residual vector is equal or nearly equal to its actual estimated uncertainty (Yang and Gao 2006a). An adaptively robust filter with classified adaptive factors (Cui and Yang 2006) was also developed, which is more effective in tracking the disturbances of the vehicle movements. In addition, an adaptively robust filter with multi-adaptive factors (Yang and Cui 2008) was created, which is more general in theory and contains adaptively robust filters with single and classified adaptive factors.

To control the influence of the measurement outliers and disturbances of the dynamic model, an adaptively robust filter based on the current statistical model (Gao et al. 2006b) was developed. In addition, an adaptively robust filter based on a neural network (Gao et al. 2007a, b) was studied to solve the construction of the dynamic model. The adaptively robust filter can also be integrated with error detection, identification, and application (DIA). To control the nonlinear disturbances of the dynamic model, an adaptive unscented Kalman filter (UKF) algorithm for improving the generalization of neural networks (Gao et al. 2008) and an adaptively robust filter based on the Bancroft algorithm (Zhang et al. 2007) have been derived.

In terms of applications, the adaptively robust filter has been successfully applied to satellite orbit determination (Yang and Wen 2004) and data processing in repeated observations of geodetic networks (Sui et al. 2007). An adaptively robust filter with constraints has also been studied for navigation applications (Yang et al. 2011). In integrated navigation applications, an adaptive Kalman filtering algorithm for the IMU/GPS integrated navigation system (Gao et al. 2006a) and a two-step adaptively robust Kalman filtering algorithm for a GPS/INS integrated navigation system (Wu and Yang 2010) have been developed. A comparison of several adaptive filtering algorithms for controlling the influence of coloured noise was analysed in order to simultaneously control the influence of coloured noise and

dynamic model disturbances (Cui et al. 2006). In research on the estimation and prediction of the satellite clock offset, an adaptively robust sequential adjustment with opening window classified adaptive factors (Huang et al. 2011) and an adaptively robust Kalman filter with classified adaptive factors for real-time estimation of satellite clock offset (Huang and Zhang 2012) were derived. Improvements in adaptive filtering have also been made with regard to estimation of deformation parameters in relation to geometric measurements and geophysical models (Yang and Zeng 2009).

7.7.6 A Brief Introduction to the Intelligent Kalman Filter

Considering the filtering methods applied in kinematic navigation, the motion state models of the moving vehicles are described and set empirically without exception. However, the actual motion rules of the moving carriers are unpredictable. In Kalman filtering, the unknown motion is described by an a priori empirical model, while GNSS observations are used to obtain the unknown motion. For this case, a method called intelligent Kalman filtering is proposed in this section, for the purpose of upgrading and extending the adaptive filter theory. The original concept of the intelligent Kalman filter was introduced in 2007 by Guochang Xu, and was funded for study by the Chinese Natural Science Foundation in 2012.

The purpose of this so-called intelligent Kalman filter is to apply the Doppler observation information in constructing the system equation. The system descriptions—which until now, without exception, have used a few empiric system equations—will be upgraded using co-determined Doppler measurements, thus providing more realistic descriptions. Because of the additional velocity information (nearly as much as the positioning information), the much more objective description of the system, and the more reasonable and precise estimation of the error disturbances, intelligent Kalman filtering can provide for greater stability and can yield more accurate results. Furthermore, the additional velocity information will be considered in determining a more reasonably adaptive factor, which is a new and advanced extension in adaptive filtering. Application of the intelligent Kalman filter in kinematic GNSS navigation and positioning is ongoing, and its application for autonomous orbit determination and manoeuvring in particular is expected to yield outstanding results.

7.8 A Priori Constrained Least Squares Adjustment

Thus far in the chapter, we have discussed several adjustment and filtering methods, all of which are suitable for full-rank linear equation problems. A full-rank quadratic matrix can be inverted to obtain its inversion. A rank-deficient linear equation

system is sometimes referred to as an over-parameterised problem. Except for the conditional least squares adjustment method, none of the methods discussed above can be directly used for solving a rank-deficient problem. The conditional least squares adjustment method with extra conditions can make the problem solvable. The conditions, of course, should be mathematically well formulated and physically well reasoned. In other words, the conditions are considered as exactly known. In practice, the conditions are quite often known with certain a priori precision. Adjustment that uses such a priori information as constraints is called an a priori constrained adjustment, which will be discussed in this section.

7.8.1 A Priori Parameter Constraints

1. A linearised observation equation system can be represented by

$$V = L - AX, \quad P_L, \tag{7.184}$$

where
P_L symmetric and definite weight matrix of dimension $m \times m$.

2. The corresponding a priori condition equation system can be written as

$$U = W - BX, \quad P_W, \tag{7.185}$$

where
B coefficient matrix of dimension $r \times n$,
W constant vector of dimension r,
U residual vector of dimension r,
P_W a priori (symmetric and definite) weight matrix of dimension $r \times r$, and
r number of condition equations; $r < n$.

3. One may interpret the constraints of Eq. 7.185 as additional pseudo-observations or as fictitious observations. This leads to the total observation equations

$$\begin{pmatrix} V \\ U \end{pmatrix} = \begin{pmatrix} L \\ W \end{pmatrix} - \begin{pmatrix} A \\ B \end{pmatrix} X, \quad P = \begin{pmatrix} P_L & 0 \\ 0 & P_W \end{pmatrix}. \tag{7.186}$$

The least squares normal equations are then well known, as (see, e.g., Sect. 7.2.1)

$$\begin{pmatrix} A^{\mathrm{T}} & B^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} P_L & 0 \\ 0 & P_W \end{pmatrix} \begin{pmatrix} A \\ B \end{pmatrix} X = \begin{pmatrix} A^{\mathrm{T}} & B^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} P_L & 0 \\ 0 & P_W \end{pmatrix} \begin{pmatrix} L \\ W \end{pmatrix}$$

or

$$(A^{\mathrm{T}}P_L A + B^{\mathrm{T}}P_W B)X = (A^{\mathrm{T}}P_L L + B^{\mathrm{T}}P_W W). \tag{7.187}$$

For convenience, a factor k (here $k = 1$) is introduced in Eq. 7.187:

$$(A^{\mathrm{T}}P_L A + kB^{\mathrm{T}}P_W B)X = (A^{\mathrm{T}}P_L L + kB^{\mathrm{T}}P_W W). \tag{7.188}$$

Equation 7.188 shows that the a priori information constraints can be added to the original least squares normal equations. In other words, the a priori information can be used for solving the rank-deficient problem and makes it possible to invert the normal matrix. Of course, these a priori information constraints should be reasonable and realistic; otherwise, the solutions could be disturbed by more serious a priori constraints. In the case of $k = 0$, the normal Eq. 7.188 turns out to be the original one, and will yield the free solution (without any a priori constraints).

The solution to the a priori constrained least squares solution is then

$$X = (A^{\mathrm{T}}P_L A + kB^{\mathrm{T}}P_W B)^{-1}(A^{\mathrm{T}}P_L L + kB^{\mathrm{T}}P_W W), \tag{7.189}$$

where $k = 1$. Generally, the a priori weight matrix is given by covariance matrix Q_W and

$$P_W = Q_W^{-1}. \tag{7.190}$$

The a priori constraints cause only two additional terms in both sides of the normal equations; therefore, all the adjustment and filtering methods discussed above can be directly used for solving the a priori constrained problem.

7.8.2 A Priori Datum

Suppose the B matrix in the a priori constraints of Eq. 7.185 is an identity matrix, and the parameter vector W is just a coordinate sub-vector of the total parameter vector. This results in a special case called a priori datum. The observation equations and a priori constraints may be rewritten as

$$V = L - (A_1 \quad A_2)\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad P_L \quad \text{and} \tag{7.191}$$

$$U = \bar{X}_2 - X_2, \quad P_2, \tag{7.192}$$

where $\overline{X}_2$ is the "observed" parameter sub-vector, P_2 is the weight matrix with respect to the parameter sub-vector X_2 and is generally a diagonal matrix, and U is a residual vector that has the same property as V. Generally, $\overline{X}_2$ is "observed"

independently, so P_2 is a diagonal matrix. If X_2 is a sub-vector of station coordinates, then the constraint of Eq. 7.192 is called the datum constraint (this is also the reason for the name "a priori datum").

The least squares normal equation of problems 7.191 and 7.192 can then be formed (similar to what discussed in Sect. 7.8.1) as

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix} \tag{7.193}$$

or

$$M_{11}X_1 + M_{12}X_2 = B_1 \quad \text{and} \tag{7.194}$$

$$M_{21}X_1 + M_{22}X_2 = B_2, \tag{7.195}$$

where

$$M_{11} = A_1^{\mathrm{T}} P_L A_1, \tag{7.196}$$

$$M_{12} = M_{21}^{\mathrm{T}} = A_1^{\mathrm{T}} P_L A_2, \tag{7.197}$$

$$M_{22} = A_2^{\mathrm{T}} P_L A_2 + P_2, \tag{7.198}$$

$$B_1 = A_1^{\mathrm{T}} P_L L, \quad \text{and} \tag{7.199}$$

$$B_2 = A_2^{\mathrm{T}} P_L L + P_2 \bar{X}_2. \tag{7.200}$$

The least squares principle used here is

$$V^{\mathrm{T}} P_L V + U^{\mathrm{T}} P_2 U = \min. \tag{7.201}$$

The normal Eq. 7.193 can be also derived by differentiating Eq. 7.201 with respect to X, and then letting it equal zero and taking Eq. 7.192 into account. In practice, the sub-vector $\overline{X}_2$ is usually a zero vector; this can be achieved through careful initialisation by forming the observation Eq. 7.191. Comparing the normal equation system of the a priori datum problem of Eqs. 7.191 and 7.192 with the normal equation of Eq. 7.191, the only difference is that the a priori weight matrix P_2 has been added to M_{22}. This indicates that the a priori datum problem can be dealt with simply by adding P_2 to the normal equation of the observation Eq. 7.191.

If some diagonal components of the weight matrix P_2 are set to zero, then the related parameters (X_2) are free parameters (or free datum) of the adjustment problem (without a priori constraints). Otherwise, parameters with a priori constraints are called a priori datum. Large weight values indicate strong constraint and small weight values indicate soft constraint. The strongest constraint is keeping the datum fixed.

7.8.3 Zhou's Theory: Quasi-Stable Datum

The quasi-stable datum method was proposed by Zhou et al. (1997). Its basic premise is that the network is dynamic, i.e. most parameters are changing all the time. However, a few points are relatively stable, or their geometric centre is relatively stable. All assumptions and observation equations are the same as in Sect. 7.8.2:

$$V = L - (A_1 \quad A_2)\begin{pmatrix} X_1 \\ X_2 \end{pmatrix}, \quad P_L \quad \text{and} \tag{7.202}$$

$$U = \bar{X}_2 - X_2, \quad P_2. \tag{7.203}$$

The least squares principles for the quasi-stable datum are

$$V^{\mathrm{T}} P_L V = \min \tag{7.204}$$

and

$$U^{\mathrm{T}} P_2 U = \min. \tag{7.205}$$

Equation 7.204 is the same as the original least squares principle. From Eq. 7.204, one has the normal equation

$$\begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = \begin{pmatrix} B_1 \\ B_2 \end{pmatrix}, \tag{7.206}$$

where

$$\begin{aligned} M_{11} &= A_1^{\mathrm{T}} P_L A_1, \\ M_{12} &= M_{21}^{\mathrm{T}} = A_1^{\mathrm{T}} P_L A_2, \\ M_{22} &= A_2^{\mathrm{T}} P_L A_2, \\ B_1 &= A_1^{\mathrm{T}} P_L L, \quad \text{and} \\ B_2 &= A_2^{\mathrm{T}} P_L L. \end{aligned} \tag{7.207}$$

Even if Eq. 7.206 is a rank-deficient equation, one may first solve Eq. 7.206 to get an explicit expression for X_2. Recalling the discussion in Sect. 7.5, one gets a normal equation related to X_2:

$$M_2 X_2 = R_2, \tag{7.208}$$

where

$$\begin{aligned} M_2 &= M_{22} - M_{21}M_{11}^{-1}M_{12} \quad \text{and} \\ R_2 &= B_2 - M_{21}M_{11}^{-1}B_1^{\mathrm{T}}. \end{aligned} \tag{7.209}$$

The new condition can be considered by forming

$$F = U^{\mathrm{T}}P_2U + 2K^{\mathrm{T}}(M_2X_2 - R_2)$$

and

$$\frac{\partial F}{\partial X} = 2U^{\mathrm{T}}P_2 + 2K^{\mathrm{T}}M_2 = 0.$$

Considering the symmetry of M_2, we have

$$U = -P_2^{-1}M_2K. \tag{7.210}$$

Substituting Eq. 7.210 into 7.203, one gets

$$X_2 = \bar{X}_2 + P_2^{-1}M_2K \tag{7.211}$$

or

$$M_2X_2 = M_2\bar{X}_2 + M_2P_2^{-1}M_2K. \tag{7.212}$$

Substituting Eq. 7.208 into 7.212, one has

$$K = (M_2P_2^{-1}M_2)^{-1}(M_2\bar{X}_2 - R_2). \tag{7.213}$$

Thus,

$$X_2 = \bar{X}_2 + P_2^{-1}M_2K, \tag{7.214}$$

$$X_1 = M_{11}^{-1}(A_1^{\mathrm{T}}P_LL - M_{12}X_2), \quad \text{and} \tag{7.215}$$

$$m_0 = \sqrt{\frac{V^{\mathrm{T}}P_LV}{n-r}}, \tag{7.216}$$

where m_0 is the standard deviation, n is the number of observations, and r is the summation of the both ranks of the matrices A_1 and A_2.

7.9 Summary

In this chapter, we have outlined the most applicable and necessary algorithms for static and kinematic as well as dynamic GPS data processing.

Least squares adjustment is the most basic adjustment method. It starts by establishing observation equations and forming normal equations, and then solves the unknowns. The sequential application of least squares adjustment by accumulating the sequential normal equations makes applications of least squares adjustment more effective. Normal equations can be formed epoch-wise and then accumulated. This method can be used not only for ultimately solving the problem, but also for obtaining epoch-wise solutions. It is suitable for static GPS data processing. The equivalent sequential least squares adjustment, which can be found in various publications, was also derived. This is an epoch-wise solving method and thus is generally not suitable for static GPS data processing. Xu (author) and Morujao (Coimbra University, Portugal) have independently reported that results obtained by applying such an algorithm will differ from those obtained by the accumulation method. The differences increase with time and are generally non-negligible. Therefore, when this method is used, the numerical process must be carefully examined to avoid the accumulation of numerical errors.

If there are constraints that have to be taken into account, a conditional least squares adjustment is needed. The commonly used least squares ambiguity search criterion is derived from this principle (cf. Sect. 8.3.4), and the general criterion of integer ambiguity search is also based on this theory (cf. Sect. 8.3.5). This method is typically applied in GPS data processing to take into account the known distance of multiple kinematic antennas. The sequential application of conditional least squares adjustment was discussed here in terms of practical needs. The problem may be solved first without conditions, after which conditions may be applied. Constraints such as the known distances of multiple antennas fixed on an aircraft must be considered for every epoch.

We also discussed lock-wise least squares adjustment for separating the unknowns into two groups—for example, one group of time-dependent parameters such as kinematic coordinates, and the other a group of time-independent parameters such as ambiguities. The sequential application of block-wise least squares adjustment makes it possible to give up some unknowns (say, out-of-date unknowns, such as past coordinates) and to keep the information related to the common unknowns during processing. This method avoids problems that may be caused by a rapid increase in the number of unknowns. There are two ways to keep the solution equivalent to a solution that is not sequential. One is to use the time-independent unknowns at the end of data processing as known, and to then go back to process the data again. The other is to remember all sequential normal equations until the best solution of the time-independent unknowns is obtained, after which the coordinates can be recomputed. A special application of block-wise

least squares adjustment was discussed for a code–phase combination model. Of course, the two observables must be suitably scaled and weighted.

We discussed the equivalently eliminated observation equation system for eliminating some nuisance parameters. This method is nearly the same as block-wise least squares adjustment if one carefully compares the normal equations of the second group of unknowns (see Sect. 7.5) and the eliminated normal equations (see Sect. 7.6). However, the most important point here is that the equivalently eliminated observation equations have been derived. Instead of solving the original problem, one may directly solve the equivalently eliminated observation equations, where the unknowns are greatly reduced, whereas the original observation vector and weight matrix remain (i.e. the problem remains uncorrelated). The precision estimation can also be made more easily by using the formulas derived in least squares adjustment. The derivation of such an equivalent observation equation was first described by Zhou (1985) and was then applied in GPS theory by Xu (2002). The unified GPS data processing method is derived using this principle (cf. Sect. 6.8). Based on the derivation of the equivalent equation, a diagonaliation algorithm of the normal equation and the observation equation was presented. The diagonalisation algorithm can be used for separating one adjustment problem into two sub-problems.

The classic Kalman filter was also discussed. It is suitable for real-time applications. A key problem of the classic Kalman filter is the divergence caused by the inexact description of system equations and its statistical properties as well as the inhomogeneity of the data. Furthermore, the solutions can be strongly dependent on the given initial values. The sequential least squares adjustment method as a special case of Kalman filtering was outlined.

Efforts have been made to modify the performance of classic Kalman filtering. In the classic Kalman filter, the weight matrix P of observables is static, i.e. P is assumed to be a definitive defined matrix. Taking the residuals of Kalman filtering into account, one may adjust the weight P of the observables accordingly; this process is called robust Kalman filtering (Koch and Yang 1998a, b). This principle can be also used for controlling the outliers of observations (Yang 1999). This idea indeed can be also used in all of the adjustment methods. The weight of an observation is usually either one (be accepted) or zero (be rejected). In robust Kalman filtering, a continuous weight between one and zero was defined and introduced. A modified weight function was also discussed and given for use. Generally speaking, the robust weighting method may modify the convergence process of the filtering procedure.

As soon as the system is defined, the Kalman filter also obtains memory abilities. However, if the system makes a discontinuous change (for example, aircraft that is static begins to run), the Kalman filter should be able to forget a part of the updated parameters. A robust Kalman filter with the addition of this ability is called an adaptively robust Kalman filter (Yang et al. 2001a, b), and was discussed in detail.

A priori constrained least squares adjustment was discussed in Sect. 7.8 for solving the rank-deficient problems, and a general discussion on the a priori parameter constraints was provided. This method makes it possible to form the

observation equations in a general way, and then a priori information can be added to keep some references fixed, such as the clock error of the reference satellite and the coordinates of the reference station. As a special case of the a priori parameter constraints, a so-called a priori datum method was discussed. The advantage of this method is that the a priori constraints just change the normal equation by adding a term (the a priori weight matrix), so that all discussed least squares adjustment and filtering methods can be directly used for solving the rank-deficient problems. Linear conditions related to the coordinate parameters can be introduced using this method. A quasi-stable datum method was also discussed. From the point of view of the dynamic earth, none station is fixed. The quasi-stable datum method takes such dynamic behaviour of the stations into account.

References

Cui X, Yang Y (2006) Adaptively Robust Filtering with Classified Adaptive Factors. Progress in Natural Science 16(8):846–851

Cui X, Yang Y, Gao W (2006) Comparison of Adaptive Filter Arithmetics in Controlling Influence of Colored Noises. Geomatics and Information Science of Wuhan University 31 (8):731-735

Cui X, Yu Z, Tao B, Liu D (1982) Adjustment in surveying. Surveying Press, Peking, (in Chinese)

Gao W, Feng X, Zhu D (2007a) GPS/INS Adaptively Integrated Navigation Algorithm Based on Neural Network. Journal of Geodesy and Geodynamics 27(2):64-67

Gao W, Yang Y, Cui X, Zhang S (2006a) Application of Adaptive Kalman Filtering Algorithm in IMU/GPS Integrated Navigation System. Geomatics and Information Science of Wuhan University 31(5):466-469

Gao W, Yang Y, Zhang S (2006b) Adaptive Robust Kalman Filter Based on the Current Statistical Model. Acta Geodaetica et Cartographica Sinica 35(1):15-18

Gao W, Yang Y, Zhang T (2007b) Neural Network Aided Adaptive Filtering for GPS/INS Integrated Navigation. Acta Geodaetica et Cartographica Sinica 36(1):26-30

Gao W, Yang Y, Zhang T (2008) An Adaptive UKF Algorithms for Improving the Generalizaiton of Neural Network. Geomatics and Information Science of Wuhan University 33(5):500-503

Gotthardt E (1978) Einführung in die Ausgleichungsrechnung. Herbert Wichmann Verlag, Karlsruhe

Huang G, Yang Y, Zhang Q (2011) Estimate and Predict Satellite Clock Error Using Adaptively Robust Sequential Adjustment with Classified Adaptive Factors Based on Opening Windows. Acta Geodaetica et Cartographica Sinica 40(1):15-21

Huang G, Zhang Q (2012) Real-time estimation of satellite clock offset using adaptively robust Kalman filter with classified adaptive factors. GPS Solutions 16(4):531-539

Huber PJ (1964) Robust estimation of a location parameter. Ann Math Stat 35:73–101

Jazwinski AH (1970) Stochastic processes and filtering theory. In: Mathematics in science and engineering, Vol. 64. Academic Press, New York and London

Koch KR (1986) Maximum likelihood estimate of variance components. Bulletin Géodésique, 60:329–338

Koch KR (1988) Parameter estimation and hypothesis testing in linear models. Springer-Verlag, Berlin

Koch KR, Yang Y (1998a) Konfidenzbereiche und Hypothesentests für robuste Parameterschätzungen. ZfV 123(1):20–26

Koch KR, Yang Y (1998b) Robust Kalman filter for rank deficient observation model. J Geodesy 72: 436–441

Masreliez CJ, Martin RD (1977) Robust Bayesian estimation for the linear model and robustifying the Kalman filter. IEEE T Automat Contr AC-22:361–371

Mohamed AH, Schwarz KP (1999) Adaptive Kalman filtering for INS/GPS. J Geodesy 73: 193–203

Ou J, Chai Y, Yuan Y (2004) Adaptive filtering for kinematic positioning by selection of the parameter weights. In: Zhu, Y. and Sun, H. (eds) Progress in Geodesy and Geodynamics. Hubei Science & Technology Press, Hubei, 816–823 (in Chinese)

Ou JK (2004) Private communication

Ren C, Ou J, Yuan Y (2005) Application of adaptive filtering by selecting the parameter weight factor in precise kinematic GPS positioning. Prog. Nat. Sci., 15(1), 41–46

Salzmann M (1995) Real-time adaptation for model errors in dynamic systems. B Geod 69:81–91

Schaffrin B (1991) Generating robustified Kalman filters for the integration of GPS and INS. Techni-cal Report, No. 15, Institute of Geodesy, University of Stuttgart

Schwarz K-P, Cannon ME, Wong RVC (1989) A Comparison of GPS kinematic models for the determination of position and velocity along a trajectory. Manuscr Geodaet 14:345–353

Sui L, Liu Y, Wang W (2007) Adaptive Sequential Adjustment and Its Application. Geomatics and Information Science of Wuhan University 32(1):51-54

Teunissen P (1990) An integrity and quality control procedure for use in multi sensor integration. In: Proceedings ION GPS90, pp. 513–522

Wang G, Chen Z, Chen W, Xu G (1988) The principle of GPS precise positioning system. Surveying Press, Peking, ISBN 7-5030-0141-0/P.58, 345 p, (in Chinese)

Wu F, Yang Y (2010) A New Two-Step Adaptive Robust Kalman Filtering in GPS/INS Integrated Navigation System. Acta Geodaetica et Cartographica Sinica 39(5):522-533

Xu G (2002) GPS data processing with equivalent observation equations, GPS Solutions, Vol. 6, No. 1-2, 6:28-33

Xu G (2003) GPS – Theory, Algorithms and Applications, Springer Heidelberg, ISBN 3-540-67812-3, 315 pages, in English

Xu G (2007) GPS – Theory, Algorithms and Applications, 2nd Ed. Springer Heidelberg, ISBN 978-3-540-72714-9, 350 pages

Xu G, Qian Z (1986) The application of block elimination adjustment method for processing of the VLBI Data. Crustal Deformation and Earthquake, Vol. 6, No. 4, (in Chinese)

Xu T, Yang Y (2000) The Improved Method of Sage Adaptive Filtering. Science of Surveying and Mapping 25(3):22-24

Yang M, Tang CH, Yu TT (2000) Development and assessment of a medium-range real-time kinematic GPS algorithm using an ionospheric information filter. Earth Planets Space 52 (10):783–788

Yang Y (1991) Robust Bayesian estimation. B Geod 65:145–150

Yang Y (1997a) Estimators of covariance matrix at robust estimation based on influence functions. ZfV 122(4):166–174

Yang Y (1997b) Robust Kalman filter for dynamic systems. Journal of Zhengzhou Institute of Surveying and Mapping 14:79–84

Yang Y (1999) Robust estimation of geodetic datum transformation. J Geodesy 73:268–274

Yang Y, Chai H, Song L (1999) Approximation for Contaminated Distribution and Its Applications. Acta Geodaetica et Cartographic Sinica 28(3):209–214

Yang Y, Cui X (2006) Adaptively Robust Filter with Multi Adaptive Factors. J. Surv. Eng.

Yang Y, Cui X (2008) Adaptively Robust Filter with Multi Adaptive Factors. Survey Review 40 (309):260-270

Yang Y, Gao W (2006a) A New Learning Statistic for Adaptive Filter Based on Predicted Residuals. Progress in Natural Science 16(8):833-837

Yang Y, Gao W (2006b) An Optimal Adaptive Kalman Filter. Journal of Geodesy 80(4):177-183

Yang Y, Gao W (2005) Comparison of Adaptive Factors on Navigation Results. The J. Navigation, 2005, 58: 471-478.

Yang Y, Gao W (2005) Influence comparison of adaptive factors on navigation results. Journal of Navigation 58, 471–478

Yang Y, Gao W (2006c) Optimal Adaptive Kalman Filter with Applications in Navigation. J Geodesy

Yang Y, He H, Xu G (2001a) Adaptively robust filtering for kinematic geodetic positioning. J Geodesy 75:109–116

Yang Y, Ren X, Xu Y (2013) Main Progress of Adaptively Robust Filter with Application in Navigation. Journal of Navigation and Positioning 1(1):9-15

Yang Y, Tang Y, Li Q and Zou Y (2006) Experiments of Adaptive Filters for Kinemetic GPS Positioning Applied in Road Information Updating in GIS. J. Surv. Eng. (in press)

Yang Y, Wen Y (2004) Synthetically adaptive robust filtering for satellite orbit determination. Science in China Series D Earth Sciences 47(7):585-592

Yang Y, Xu T (2003) An Adaptive Kalman Filter Based on Sage Windowing Weights and Variance Components. Journal of Navigation 56(2):231-240

Yang Y, Xu T (2004) An Adaptively Regularization Method with Combination of Priori and Posterior Information. In: Zhu, Y. and Sun, H. (eds) Progress in Geodesy and Geodynamics. Hubei Science & Technology Press, Hubei

Yang Y, Xu T, He H (2001b) On adaptively kinematic filtering. Selected Papers for English of Acta Geodetica et Cartographica Sinica, pp. 25–32

Yang Y, Zeng A (2009) Adaptive Filtering for Deformation Parameter Estimation in Consideration of Geometrical Measurements and Geopgysical Models. Science in China Series D Earth Sciences 52(8):1216-1222

Yang Y, Zhang X, Xu J (2011) Adaptively Constrained Kalman Filtering for Navigation Applications. Survey Review 43(322):370-381

Zhang S, Yang Y, Zhang Q (2007) An Adaptively Robust Filter Based on Bancroft Algorithm in GPS Navigation. Geomatics and Information Science of Wuhan University 32(4):309-311

Zhou J (1985) On the Jie factor. Acta Geodaetica et Geophysica 5 (in Chinese)

Zhou J, Huang Y, Yang Y, Ou J (1997) Robust least squares method. Publishing House of Huazhong University of Science and Technology, Wuhan