Properties of Gauss-Newton filter in linear cases

BAO Zhichao* and JIANG Qiuxi

Institute of Electronic Countermeasure, National University of Defense Technology, Hefei 230037, China

Abstract: This paper presents the derivation of Gauss-Newton filter in linear cases and an analysis of its properties. Based on the minimum variance theorem, the Gauss-Newton filter is constructed and derived, including its state transition equation, observation equation and filtering process. Then, the delicate relationship between the Gauss-Aitken filter and the Kalman filter is discussed and it is verified that without process noise the two filters are equivalent. Finally, some simulations are conducted. The result shows that the Gauss-Aitken filter is superior to the Kalman filter in some aspects.

Keywords: Gauss-Newton filter, linear case, target tracking.

DOI: 10.21629/JSEE.2018.05.02

1. Introduction

Gauss invented the method of least squares and Gaussian probability density function at the beginning of the nineteenth century [1], which was a great break in estimation theory. Those great findings were generated by the tough problem of predicting the orbit of Ceres. When working on the orbit of Ceres, Gauss also made use of the differential calculus invented by Newton, thereby creating the Gauss-Newton algorithm. In 1958, Swerling [2] who was devoted to the radar field found that the non-recursive Gauss-Newton algorithm was unsuitable for tracking targets such as satellites due to the limited computation capacity of computers in that time. Thus he reworked the Gauss-Newton algorithm and proposed a new recursive algorithm which was later called the Swerling filter. The Swerling filter aimed to estimate orbital parameters without any delay. Although the accuracy was poor at first, it gradually converged to the Gauss-Newton filter. It was about two years later when Kalman and Bucy invented their algorithm, which is now known as Kalman filter [3,4]. The Kalman filter and the Swerling filter were derived using different approaches, but turned out to have the similar structure, and the Kalman filter can be derived from the Swerling filter as well as the Gauss-Newton filter. This

paper will examine and illustrate the relationship between the Gauss-Newton filter and the Kalman filter later.

Ever since the Kalman filter was put into use, its great superiority attracted many engineers, and algorithms based on the Kalman filter have been playing a leading role in tracking filter technique both in the academic field and in engineering [5–19]. Meanwhile, the computer technique [20–27] has improved in leaps since the 1960s, and computation is no longer a difficulty for the Gauss-Newton filter [28–31] now. Therefore, it is necessary to reconsider the properties of the Gauss-Newton filter and take full advantage of them.

2. Construction of Gauss-Newton filter in linear cases

The Gauss-Newton filter in linear cases is often called the Gauss-Aitken filter in memory of Gauss and Aitken's great contribution to the formulation of minimum variance algorithm (MVA), which is the basis of the Gauss-Newton filter.

2.1 The process to be estimated

The Gauss-Aitken filter addresses the general problem of trying to estimate the state $\mathbf{X}(t_n)$ of a discrete-time controlled process that is governed by the linear differential equation

$$d\mathbf{X}(t_n) = \mathbf{A}(t)\mathbf{X}(t_n) \quad (1)$$

in which $\mathbf{A}(t)$ is called the coefficient matrix, $\mathbf{X}(t_n) = (x, \dot{x}, \ddot{x}, \dots, D^m x)^T$ is the vector state of the process, and the entries in $\mathbf{X}(t_n)$ are the different derivatives of x which is defined as

$$d^j x_n \equiv d^j x(t)/dt^j|_{t=t_n}. \quad (2)$$

The subscript n signifies that it is evaluated at $t = t_n$.

2.2 Transition matrix and transition equations

In filter engineering, we use the transition matrix as implementation of process differential equation as in (1). It is often the case that we are required to evaluate $\mathbf{X}(t_n)$ at different instants. Let $\mathbf{X}(t_n)$ be evaluated at two instants

t_n and $t_n + \tau$, giving vectors $\mathbf{X}(t_n)$ and $\mathbf{X}(t_n + \tau)$. The two vectors are connected closely through the transition equation as

$$\mathbf{X}(t_n + \tau) = \Phi(\tau)\mathbf{X}(t_n) \quad (3)$$

in which $\Phi(\tau)$ is called the transition matrix of the selected differential equation, and τ is called the span of the transition.

Each of the linear differential equation gives rise to a unique transition matrix. If coefficient matrix $\mathbf{A}(t)$ is a matrix of constants, then the transition matrix can be obtained by

$$\Phi(\tau) = \exp(\tau\mathbf{A}) \quad (4)$$

$$\exp(\tau\mathbf{A}) \equiv \mathbf{I} + \tau\mathbf{A} + \tau^2\mathbf{A}^2/2! + \cdots + \tau^m\mathbf{A}^m/m! + \cdots \quad (5)$$

in which $\mathbf{I}$ is the identity matrix.

If at least one entry in coefficient matrix $\mathbf{A}(t)$ is a function of t , then (4) cannot apply here. Instead, the numerical integration method is required and the transition matrix Φ turns out to be a function of t , which makes (3) in the form of

$$\mathbf{X}(t_n + \tau) = \Phi(t_n, \tau)\mathbf{X}(t_n). \quad (6)$$

Since it is often the case that the coefficient matrix $\mathbf{A}$ is a matrix of constants, we assume $\mathbf{A}$ to be a matrix of constants throughout this paper.

2.3 Observation equation

Let the observation vector $\mathbf{Y}_n$ be a k -vector of observations that are received at time t_n , i.e.,

$$\mathbf{Y}_n = (y_1, y_2, \dots, y_k)_n^T. \quad (7)$$

Let the vector of errors in $\mathbf{Y}_n$ be

$$\mathbf{N}_n = (\nu_1, \nu_2, \dots, \nu_k)_n^T \quad (8)$$

where $\nu_1, \nu_2, \dots, \nu_k$ are independent and obey Gaussian distribution.

Let $\mathbf{X}_n$ be the true state vector of the observed trajectory at time t_n and let it be an m -vector. Then in linear cases the observation vector $\mathbf{Y}_n$ is related to $\mathbf{X}_n$ as follows:

$$\mathbf{Y}_n = \mathbf{M}\mathbf{X}_n + \mathbf{N}_n \quad (9)$$

in which $\mathbf{M}$ is a $k \times m$ matrix of constants and is referred to the observation matrix. Equation (9) is called the observation equation.

Since the Gauss-Newton filter is non-recursive, an estimate of the vector will always be derived from a sequence of $L + 1$ observation vectors at time $t_n, t_{n-1}, \dots, t_{n-L}$. Let those $L + 1$ observation vectors be called $\mathbf{Y}_n, \mathbf{Y}_{n-1}, \dots, \mathbf{Y}_{n-L}$ and let the associated $L + 1$ observation equations be as follows:

$$\mathbf{Y}_{n-k} = \mathbf{M}\mathbf{X}_{n-k} + \mathbf{N}_{n-k}, \quad 0 \leq k \leq L. \quad (10)$$

Rearranging (10) in the form of the block matrix gives

$$\begin{bmatrix} \mathbf{Y}_n \\ \mathbf{Y}_{n-1} \\ \vdots \\ \mathbf{Y}_{n-L} \end{bmatrix} = \begin{bmatrix} \mathbf{M}\mathbf{X}_n \\ \mathbf{M}\mathbf{X}_{n-1} \\ \vdots \\ \mathbf{M}\mathbf{X}_{n-L} \end{bmatrix} + \begin{bmatrix} \mathbf{N}_n \\ \mathbf{N}_{n-1} \\ \vdots \\ \mathbf{N}_{n-L} \end{bmatrix}. \quad (11)$$

Let $\tau = t_{n-k} - t_n$, then the vectors $\mathbf{X}_{n-k}$ and $\mathbf{X}_n$ will be related as follows:

$$\mathbf{X}_{n-k} = \Phi(t_{n-k} - t_n)\mathbf{X}_n. \quad (12)$$

Substituting $\mathbf{X}_{n-k}$ in (11) using (12) then gives

$$\begin{bmatrix} \mathbf{Y}_n \\ \mathbf{Y}_{n-1} \\ \vdots \\ \mathbf{Y}_{n-L} \end{bmatrix} = \begin{bmatrix} \mathbf{M}\mathbf{X}_n \\ \mathbf{M}\Phi(t_{n-1} - t_n)\mathbf{X}_n \\ \vdots \\ \mathbf{M}\Phi(t_{n-L} - t_n)\mathbf{X}_n \end{bmatrix} + \begin{bmatrix} \mathbf{N}_n \\ \mathbf{N}_{n-1} \\ \vdots \\ \mathbf{N}_{n-L} \end{bmatrix} \quad (13)$$

which can be reduced to

$$\begin{bmatrix} \mathbf{Y}_n \\ \mathbf{Y}_{n-1} \\ \vdots \\ \mathbf{Y}_{n-L} \end{bmatrix} = \begin{bmatrix} \mathbf{M} \\ \mathbf{M}\Phi(t_{n-1} - t_n) \\ \vdots \\ \mathbf{M}\Phi(t_{n-L} - t_n) \end{bmatrix} \mathbf{X}_n + \begin{bmatrix} \mathbf{N}_n \\ \mathbf{N}_{n-1} \\ \vdots \\ \mathbf{N}_{n-L} \end{bmatrix}. \quad (14)$$

In (14), we now define

$$\mathbf{Y}_n \equiv \begin{bmatrix} \mathbf{Y}_n \\ \mathbf{Y}_{n-1} \\ \vdots \\ \mathbf{Y}_{n-L} \end{bmatrix}, \quad \mathbf{N}_n \equiv \begin{bmatrix} \mathbf{N}_n \\ \mathbf{N}_{n-1} \\ \vdots \\ \mathbf{N}_{n-L} \end{bmatrix},$$

$$\mathbf{T}_n \equiv \begin{bmatrix} \mathbf{M} \\ \mathbf{M}\Phi(t_{n-1} - t_n) \\ \vdots \\ \mathbf{M}\Phi(t_{n-L} - t_n) \end{bmatrix} \quad (15)$$

in which sans-serif letters $\mathbf{Y}, \mathbf{T}, \mathbf{N}$ are used to represent the total observation vector, total observation matrix and total error vector respectively. The subscripts of the vectors $\mathbf{Y}_n$ and $\mathbf{N}_n$ are consistent with the most recent subvectors, namely $\mathbf{Y}_n$ and $\mathbf{N}_n$.

Equation (14) can now be written as

$$\mathbf{Y}_n = \mathbf{T}_n\mathbf{X}_n + \mathbf{N}_n \quad (16)$$

which is called the total observation equation.

Let $\mathbf{R}_{\mathbf{Y}_n}$ be the covariance matrix of $\mathbf{Y}_n$. Then $\mathbf{R}_{\mathbf{Y}_n}$ is called the total observation covariance matrix. According to the definition of covariance, we can obtain

$$R_{Y_n} = E\{N_n N_n^T\} = E \begin{bmatrix} N_n \\ N_{n-1} \\ \vdots \\ N_{n-L} \end{bmatrix} \begin{bmatrix} N_n^T & N_{n-1}^T & \cdots & N_{n-L}^T \end{bmatrix} =$$

$$\begin{bmatrix} E(N_n N_n^T) & E(N_n N_{n-1}^T) & \cdots & E(N_n N_{n-L}^T) \\ E(N_{n-1} N_n^T) & E(N_{n-1} N_{n-1}^T) & \cdots & E(N_{n-1} N_{n-L}^T) \\ \vdots & \vdots & \ddots & \vdots \\ E(N_{n-L} N_n^T) & E(N_{n-L} N_{n-1}^T) & \cdots & E(N_{n-L} N_{n-L}^T) \end{bmatrix}. \quad (17)$$

It is often the case that the error vectors are uncorrelated when they are made at different instants, i.e.,

$$E(N_n N_m^T) = \mathbf{0}, \quad n \neq m. \quad (18)$$

Then (17) can be reduced to the block diagonal matrix

$$R_{Y_n} = \begin{bmatrix} E(N_n N_n^T) & & & \\ & E(N_{n-1} N_{n-1}^T) & & \\ & & \ddots & \\ & & & E(N_{n-L} N_{n-L}^T) \end{bmatrix}. \quad (19)$$

We can now write (3) and (16) together, which gives

$$X(t_n + \tau) = \Phi(\tau) X(t_n) \quad (20)$$

$$Y_n = T_n X_n + N_n. \quad (21)$$

Observe the two equation carefully, and you will realize that the Gauss-Newton filter's process and observation equation are similar to the Kalman filter's in linear cases. Now we will give the structure of the Gauss-Aitken filter first in Fig. 1, and the derivation will be given in the next subsection.

Fig. 1 Gauss-Aitken filter

3. Derivation of Gauss-Aitken filter

3.1 Residuals

If we replace X_n in the total observation equation (16) by any estimate vector $X_{n,n}^*$, we obtain

$$Y_n = T_n X_{n,n}^* + N_n^*. \quad (22)$$

Leaving N_n^* on one side, we obtain

$$N_n^* = Y_n - T_n X_{n,n}^* \quad (23)$$

in which the vector N_n^* is called the residual vector with respect to $X_{n,n}^*$, and $T_n X_{n,n}^*$ is called the fitted observation vector.

3.2 Finding X_{minvar}^*

According to the minimum variance theorem proposed by Gauss, among all the linear combinations of measurements for estimating an unknown parameter, the weighted least squares estimate has the greatest precision, regardless of distribution of observation errors. Professor Morrison [32] gives the definition of the sum of the weighted square residuals as follows:

$$e(X_{n,n}^*) = (Y_n - T_n X_{n,n}^*)^T R_{T_n}^{-1} (Y_n - T_n X_{n,n}^*). \quad (24)$$

By differentiating $e(X_{n,n}^*)$ with respect to $X_{n,n}^*$ and setting result equal to zero matrix $\mathbf{0}$, we obtain

$$X_{n,n}^* = (T_n^T R_{Y_n}^{-1} T_n)^{-1} T_n^T R_{Y_n}^{-1} Y_n. \quad (25)$$

We now restate it as

$$X_{minvar}^* = W_{minvar} Y_n \quad (26)$$

in which

$$W_{minvar} = (T_n^T R_{Y_n}^{-1} T_n)^{-1} T_n^T R_{Y_n}^{-1}. \quad (27)$$

For this we note that X_{minvar}^* is a linear transformation of Y_n and so we can obtain

$$S_{minvar}^* = W_{minvar} R_{Y_n} W_{minvar}^T \quad (28)$$

and S_{minvar}^* in (28) can be reduced as

$$S_{minvar}^* = (T_n^T R_{Y_n}^{-1} T_n)^{-1}. \quad (29)$$

Thus, the Gauss-Aitken filter algorithm has been proved in Fig. 1 above.

4. Delicate relationship between Gauss-Aitken filter and Kalman filter

The relationship between the Gauss-Aitken filter and the Kalman filter is really delicate, and it is mainly shown in two aspects. First, as we can see from (20) and (21) above, the two filters' process and observation equations have similar formation, and we can derive the Kalman filter algorithm from the Gauss-Aitken's. Second, we will prove that the two filters are equivalent if there is no process noise. However, when process noise is present, the equivalence does not exist anymore.

4.1 Derivation of Kalman filter algorithm from Gauss-Aitken filter

We will see that, in order to carry out the derivation, it is necessary that the observation errors must be uncorrelated, and if this condition is not present, then the Kalman filter cannot be used.

Let's assume that we have obtained $X_{n-1,n-1}^*$ with its filter covariance matrix $S_{n-1,n-1}^*$, at time t_n when the observation vector Y_n and its covariance matrix are received, and we wish to combine them with $X_{n-1,n-1}^*$ and $S_{n-1,n-1}^*$ using MVA.

At the first step, we forward-project $X_{n-1,n-1}^*$ and $S_{n-1,n-1}^*$ from time t_{n-1} to t_n using (20).

$$X_{n,n-1}^* = \Phi(n, n-1)X_{n-1,n-1}^* \quad (30)$$

$$S_{n,n-1}^* = \Phi(n, n-1)S_{n-1,n-1}^*\Phi(n, n-1)^T \quad (31)$$

We now write the following two equations:

$$X_{n,n-1}^* = X_n + N_{n,n-1}^* \quad (32)$$

$$Y_n = MX_n + N_n. \quad (33)$$

In (32) the elements of vector $N_{n,n-1}^*$ are the errors in $X_{n,n-1}^*$, and in (33) the elements of N_n are errors in Y_n .

Concatenating of the above equations gives

$$\begin{bmatrix} X_{n,n-1}^* \\ Y_n \end{bmatrix} = \begin{bmatrix} I \\ M_n \end{bmatrix} X_n + \begin{bmatrix} N_{n,n-1}^* \\ N_n \end{bmatrix}. \quad (34)$$

We now define the three quantities appearing in (34) as

$$\begin{bmatrix} X_{n,n-1}^* \\ Y_n \end{bmatrix} \equiv Y_n, \quad \begin{bmatrix} I \\ M_n \end{bmatrix} \equiv T_n, \quad \begin{bmatrix} N_{n,n-1}^* \\ N_n \end{bmatrix} \equiv N_n. \quad (35)$$

Thus, (35) can be written as

$$Y_n = T_n X_n + N_n. \quad (36)$$

We can see (36) is the same form as the equation which forms the starting point for the derivation of the MVA. Making use of what we did there, we can accordingly write the following:

$$X_{n,n}^* = (T_n^T R_n^{-1} T_n)^{-1} T_n^T R_n^{-1} Y_n \quad (37)$$

$$S_{n,n}^* = (T_n^T R_n^{-1} T_n)^{-1}. \quad (38)$$

In the above two equations, R_n is in the covariance matrix of N_n in (36) and can be expanded as follows:

$$R_n = E(N_n N_n^T) = E \begin{bmatrix} N_{n,n-1}^* \\ N_n \end{bmatrix} \begin{bmatrix} N_{n,n-1}^{*T} & N_n^T \end{bmatrix} = \begin{bmatrix} S_{n,n-1}^* & EN_{n,n-1}^* N_n^T \\ EN_n N_{n,n-1}^{*T} & R_n \end{bmatrix} \quad (39)$$

since we make the assumption that the observation errors are uncorrelated, i.e.,

$$EN_n N_m = 0, \quad n \neq m. \quad (40)$$

It then follows that the two off-diagonal matrices in (39) satisfy

$$EN_{n,n-1}^* N_n^T = 0, \quad EN_n N_{n,n-1}^{*T} = 0. \quad (41)$$

Then (39) reduces to

$$R_n = \begin{bmatrix} S_{n,n-1}^* & 0 \\ 0 & R_n \end{bmatrix}. \quad (42)$$

Returning to (38), we expand the terms, obtaining

$$S_{n,n}^* = (T_n^T R_n^{-1} T_n)^{-1} = (S_{n,n-1}^{*-1} + M_n^T R_n^{-1} M_n)^{-1} \quad (43)$$

$$\begin{aligned} X_{n,n}^* &= (T_n^T R_n^{-1} T_n)^{-1} T_n^T R_n^{-1} Y_n = \\ &= S_{n,n}^* (S_{n,n-1}^{*-1} X_{n,n-1}^* + M_n^T R_n^{-1} Y_n) = \\ &= S_{n,n}^* (S_{n,n-1}^{*-1} X_{n,n-1}^* + M_n^T R_n^{-1} M_n X_{n,n-1}^* - \\ &= M_n^T R_n^{-1} M_n X_{n,n-1}^* + M_n^T R_n^{-1} Y_n) = \\ &= S_{n,n}^* (S_{n,n-1}^{*-1} + M_n^T R_n^{-1} M_n) X_{n,n-1}^* + \\ &= S_{n,n}^* M_n^T R_n^{-1} (Y_n - M_n X_{n,n-1}^*). \end{aligned} \quad (44)$$

However, the two quantities $S_{n,n}^*$ and $(S_{n,n-1}^{*-1} + M_n^T R_n^{-1} M_n)$ are each other's inverse, and we continue as follows:

$$X_{n,n}^* = X_{n,n-1}^* + S_{n,n}^* M_n^T R_n^{-1} (Y_n - M_n X_{n,n-1}^*). \quad (45)$$

We define the weight matrix as

$$H_n = S_{n,n}^* M_n^T R_n^{-1}. \quad (46)$$

Then (45) becomes

$$\mathbf{X}_{n,n}^* = \mathbf{X}_{n,n-1}^* + \mathbf{H}_n(\mathbf{Y}_n - \mathbf{M}_n \mathbf{X}_{n,n-1}^*). \quad (47)$$

We use what is known as Woodbury matrix identity to rearrange (43) thereby obtaining

$$\begin{aligned} \mathbf{S}_{n,n}^* &= (\mathbf{S}_{n,n-1}^{*-1} + \mathbf{M}_n^T \mathbf{R}_n^{-1} \mathbf{M}_n)^{-1} = \\ \mathbf{S}_{n,n-1}^* - \mathbf{H}_n \mathbf{M}_n \mathbf{S}_{n,n-1}^* &= (\mathbf{I} - \mathbf{H}_n \mathbf{M}_n) \mathbf{S}_{n,n-1}^*. \end{aligned} \quad (48)$$

Thus, the derivation is complete.

Assembling the equations for the Kalman filter, we have the following:

$$\mathbf{X}_{n,n-1}^* = \Phi(n, n-1) \mathbf{X}_{n-1,n-1}^* \quad (49)$$

$$\mathbf{S}_{n,n-1}^* = \Phi(n, n-1) \mathbf{S}_{n-1,n-1}^* \Phi(n, n-1)^T \quad (50)$$

$$\mathbf{S}_{n,n}^* = (\mathbf{S}_{n,n-1}^{*-1} + \mathbf{M}_n^T \mathbf{R}_n^{-1} \mathbf{M}_n)^{-1} \quad (51)$$

$$\mathbf{H}_n = \mathbf{S}_{n,n}^* \mathbf{M}_n^T \mathbf{R}_n^{-1} \quad (52)$$

$$\mathbf{X}_{n,n}^* = \mathbf{X}_{n,n-1}^* + \mathbf{H}_n(\mathbf{Y}_n - \mathbf{M}_n \mathbf{X}_{n,n-1}^*). \quad (53)$$

In the process of derivation, we do not allow for process noise. However, if the process noise is present, we assume the process noise is $\omega_1, \omega_2, \dots, \omega_k$ and their covariance matrix is $\mathbf{Q}$, and all we need to do is adding $\mathbf{Q}$ up to the right side of (50). Thus we obtain

$$\mathbf{S}_{n,n-1}^* = \Phi(n, n-1) \mathbf{S}_{n-1,n-1}^* \Phi(n, n-1)^T + \mathbf{Q}. \quad (54)$$

4.2 Equivalence of the two filters without process noise

Assume that we are observing the coordinates of a fixed point in 2-space. Let $\mathbf{X} = (x_1, x_2)^T$ be the state vector and the those observation vectors be $\mathbf{Y}_m = (y_1, y_2)_m^T$ ($m = 1, 2, \dots$), where

$$\mathbf{Y}_m = \mathbf{X} + \mathbf{N}_m. \quad (55)$$

Let the covariance matrices of $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_m$ be $\mathbf{R}_1, \mathbf{R}_2, \dots, \mathbf{R}_m$, $\mathbf{N}_m$ is observation error at the m th sample. The derivation of equivalence of the two filters equals to prove that the Kalman filter's output $\mathbf{X}_{m,m}^*|_K$, $\mathbf{S}_{m,m}^*|_K$ are the same with Gauss-Aitken's output $\mathbf{X}_{m,m}^*|_G$, $\mathbf{S}_{m,m}^*|_G$. We choose to use the principle of mathematical induction.

Step 1 When $m = 2$, it is easy to see that the state vector satisfies the differential equation

$$d\mathbf{X} = \mathbf{0}. \quad (56)$$

Therefore, it follows that the transition matrix is

$$\Phi(\tau) = \mathbf{I} \quad (57)$$

in which $\mathbf{I}$ is the 2×2 identity matrix. Then processing the data vectors $\mathbf{Y}_1, \mathbf{Y}_2$ using the Kalman filter, let

$$\mathbf{X}_{1,1}^*|_K = \mathbf{Y}_1, \quad \mathbf{S}_{1,1}^{2*}|_K = \mathbf{R}_{Y_1}. \quad (58)$$

Making use of the Kalman filter algorithm (49)–(53), we obtain

$$\begin{aligned} \mathbf{X}_{2,2}^*|_K &= \mathbf{Y}_1 + (\mathbf{R}_{Y_1}^{-1} + \mathbf{R}_{Y_2}^{-1})^{-1} \mathbf{R}_{Y_2}^{-1} (\mathbf{Y}_2 - \mathbf{Y}_1) = \\ &= \mathbf{S}_{2,2}^*|_K (\mathbf{R}_{Y_1}^{-1} \mathbf{Y}_1 + \mathbf{R}_{Y_2}^{-1} \mathbf{Y}_2) \\ \mathbf{S}_{2,2}^*|_K &= (\mathbf{R}_{Y_1}^{-1} + \mathbf{R}_{Y_2}^{-1})^{-1}. \end{aligned} \quad (59)$$

Processing the data vectors $\mathbf{Y}_1, \mathbf{Y}_2$ using the Gauss-Aitken filter, let

$$\mathbf{T} = \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \end{bmatrix} \quad \mathbf{R} = \begin{bmatrix} \mathbf{R}_{Y_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{Y_2} \end{bmatrix}. \quad (60)$$

Making use of the Gauss-Aitken algorithm shown in Fig. 1, we obtain

$$\begin{cases} \mathbf{S}_{2,2}^*|_G = (\mathbf{T}^T \mathbf{R} \mathbf{T})^{-1} = \\ \left(\begin{bmatrix} \mathbf{I} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{R}_{Y_1} & \mathbf{0} \\ \mathbf{0} & \mathbf{R}_{Y_2} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{I} \\ \mathbf{I} \end{bmatrix} \right)^{-1} = \\ (\mathbf{R}_{Y_1}^{-1} + \mathbf{R}_{Y_2}^{-1})^{-1} \\ \mathbf{X}_{2,2}^*|_G = \mathbf{S}_{2,2}^*|_G \mathbf{T}^T \mathbf{R}^{-1} \mathbf{T} = \\ \mathbf{S}_{2,2}^*|_G (\mathbf{R}_{Y_1}^{-1} \mathbf{Y}_1 + \mathbf{R}_{Y_2}^{-1} \mathbf{Y}_2) \end{cases} \quad (61)$$

Comparison of (59) and (61) shows that the equivalence is satisfied when $m = 2$.

Step 2 Assume that when $m = k$ ($k \geq 2$), the equivalence is satisfied, i.e.,

$$\mathbf{X}_{k,k}^*|_K = \mathbf{X}_{k,k}^*|_G, \quad \mathbf{S}_{k,k}^*|_K = \mathbf{S}_{k,k}^*|_G. \quad (62)$$

When $m = k + 1$, making use of the Kalman filter algorithm (49)–(53), we obtain

$$\begin{aligned} \mathbf{X}_{k+1,k+1}^*|_K &= \mathbf{S}_{k+1,k+1}^*|_K (\mathbf{S}_{k,k}^{*-1} \mathbf{X}_{k,k}^* + \mathbf{R}_{k+1}^{-1} \mathbf{Y}_{k+1}) \\ \mathbf{S}_{k+1,k+1}^*|_K &= (\mathbf{S}_{k,k}^{*-1} + \mathbf{R}_{k+1}^{-1})^{-1}. \end{aligned} \quad (63)$$

Processing the data vectors $\mathbf{Y}_1, \mathbf{Y}_2, \dots, \mathbf{Y}_{k+1}$ using the Gauss-Aitken filter, let

$$\begin{aligned} \mathbf{T} &= [\mathbf{I} \quad \mathbf{I} \quad \dots \quad \mathbf{I}]_{k+1}^T \\ \mathbf{R} &= \begin{bmatrix} \mathbf{R}_{Y_1} & & & \\ & \ddots & & \\ & & \mathbf{R}_{Y_k} & \\ & & & \mathbf{R}_{Y_{k+1}} \end{bmatrix} \end{aligned} \quad (64)$$

in which $\mathbf{T}$ is made up of $k + 1$ identity matrices $\mathbf{I}$, and $\mathbf{R}$ is the block diagonal matrix which is made up of $\mathbf{R}_{Y_1}, \mathbf{R}_{Y_2}, \dots, \mathbf{R}_{Y_{k+1}}$.

Making use of the Gauss-Aitken algorithm shown in Fig. 1, we obtain

$$\left\{ \begin{array}{l} S_{k+1,k+1}^*|_G = (T^T R T)^{-1} = \\ \left(\begin{bmatrix} I & I & \cdots & I \end{bmatrix} \begin{bmatrix} R_{Y_1} & & & \\ & \ddots & & \\ & & R_{Y_{k+1}} & \end{bmatrix}^{-1} \begin{bmatrix} I \\ I \\ \vdots \\ I \end{bmatrix} \right)^{-1} = \\ (R_{Y_1}^{-1} + \cdots + R_{Y_{k+1}}^{-1})^{-1} \\ X_{k+1,k+1}^*|_G = S_{k+1,k+1}^*|_G T^T R^{-1} T = \\ S_{k+1,k+1}^*|_G (R_{Y_1}^{-1} Y_1 + \cdots + R_{Y_{k+1}}^{-1} Y_{k+1}) \end{array} \right. \quad (65)$$

We assume that

$$\left\{ \begin{array}{l} X_{k,k}^*|_K = X_{k,k}^*|_G = S_{k,k}^*|_G (R_{Y_1}^{-1} Y_1 + \cdots + R_{Y_k}^{-1} Y_k) \\ S_{k,k}^*|_K = S_{k,k}^*|_G = (R_{Y_1}^{-1} + \cdots + R_{Y_k}^{-1})^{-1} \end{array} \right. \quad (66)$$

Substituting (65) and (66) in (63), we obtain

$$\begin{aligned} X_{k+1,k+1}^*|_K &= X_{k+1,k+1}^*|_G = \\ S_{k+1,k+1}^*|_G (R_{Y_1}^{-1} Y_1 + \cdots + R_{Y_{k+1}}^{-1} Y_{k+1}) \\ S_{k+1,k+1}^*|_K &= S_{k+1,k+1}^*|_G = \\ (R_{Y_1}^{-1} + \cdots + R_{Y_{k+1}}^{-1})^{-1}. \end{aligned} \quad (67)$$

Thus, for $m = k+1$, the equivalent is also satisfied, and the verification is complete.

At this point it is worthwhile to review our results. The Kalman filter is derived by converting the Gauss-Aitken filter from a non-recursive form to a recursive form. The Kalman filter in linear cases is therefore the recursive dual of the non-recursive Gauss-Aitken filter without process noise.

4.3 The Q matrix

Since both of the two filters are model-based filters, these models inevitably exhibit structural and parametric mismatch with the process, resulting from fitting a model of limited complexity to a finite set of data for a complex process [33]. Thus, the Kalman filter uses process noise to represent the mismatch and uncertainty. However, in reality the covariance of process noise which is called the Q matrix is hard to achieve and a complicated process of “tuning” is required. If the Q matrix matches well with the actual mismatch and uncertainty, the Kalman filter performs well; otherwise, good performance cannot be expected and filter becomes divergent, as is often the case. What’s more, in order to keep the Kalman filter stable, the Q matrix is always involved in filter even though the mismatch and uncertainty can be neglected, which never happens in the Gauss-Aitken filter.

As for the Gauss-Aitken filter, we assume the model matches the process well. If the mismatch and uncertainty cannot be neglected, then we cannot use the Gauss-Aitken filter in those circumstances, which is also a tough job for the Kalman filter.

5. Simulation and results analysis

In the previous sections we present the basic form of the Gauss-Aitken filter, and the relationship between the Gauss-Aitken filter and the Kalman filter. To help developing a better understanding of the operation and properties of the Gauss-Aitken filter, we present a very simple example below.

Assume that a target is moving along x -axis with a constant velocity, we need to estimate the target’s location and velocity. Let the state vector be $X = [x \ \dot{x}]^T$ which satisfies the differential equation

$$d \begin{bmatrix} x \\ \dot{x} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ \dot{x} \end{bmatrix}. \quad (68)$$

Resolve it using (4), we can get the transition matrix

$$\Phi(\tau) = \begin{bmatrix} 1 & \tau \\ 0 & 1 \end{bmatrix}. \quad (69)$$

Assume that we can only observe the location of the target with the observation error v_k ($k = 0, 1, \dots, 99$) which is zero-mean Gaussian distribution and uncorrected with each other. Let all the variances be equal to 4. Thus, when $t = t_k$ the total observation vector Y_k , the total error vector N , the observation matrix M , and the total covariance matrix R_{Y_k} equal

$$\begin{aligned} Y_k &= \begin{bmatrix} Y_0 \\ Y_1 \\ \vdots \\ Y_k \end{bmatrix}, \quad N_k = \begin{bmatrix} \nu_0 \\ \nu_1 \\ \vdots \\ \nu_k \end{bmatrix}, \quad M = \begin{bmatrix} 1 & 0 \end{bmatrix}, \\ R_{Y_k} &= \begin{bmatrix} 4 & & & \\ & 4 & & \\ & & \ddots & \\ & & & 4 \end{bmatrix}_{(k+1) \times (k+1)}. \end{aligned} \quad (70)$$

Now we have M , $\Phi(\tau)$, $L = k$, and we can construct the total observation matrix T_k as follows:

$$T_k = \begin{bmatrix} M \Phi(0) \\ M \Phi(1) \\ \vdots \\ M \Phi(k) \end{bmatrix} = \begin{bmatrix} 1 & -k \\ 1 & -(k-1) \\ \vdots & \vdots \\ 1 & 0 \end{bmatrix}. \quad (71)$$

Then all we need to carry out the Gauss-Aitken filter is available. In simulation we assume that the real state vector is $X(0) = [2 \ 15]^T$. In order to illustrate the properties

of the Gauss-Newton filter, exactly the same data is processed by the Kalman filter. The results are as follows.

Fig. 2 presents that the two filters can trace the target position successfully; the Kalman estimation trajectory and the Gauss-Aitken estimation trajectory both diverge to the real trajectory.

Fig. 2 Trajectories of position

Fig. 3 is the local zoom of Fig. 2. It shows that the Gauss-Aitken estimation trajectory is closer to the real trajectory at some random time. We analyze the phenomenon and believe that since the Kalman filter is recursive and every state estimate needs many times inverse of the matrix, as a result the round-off error is accumulative.

Fig. 3 Local zoom of Fig. 2

In contrast, the Gauss-Aitken filter is non-recursive and every state estimate needs only one-time inverse of the matrix, so the round-off error is trivial. Therefore, the Kalman filter is more sensitive to round-off error which makes it

potentially unstable. Meanwhile the Gauss-Aitken filter is more stable and its robustness is better.

Fig. 4 displays the process of tracking and filtering the target's velocity. It shows clearly that Gauss-Aitken estimation velocity has converged to the real velocity much more quickly. In addition, while the Kalman filter needs initialization which in turn affects the filter's convergence speed, the Gauss-Aitken filter does not need initialization at all.

Fig. 4 Trajectories of velocity

Fig. 5 and Fig. 6 illustrate the estimation error variance changing over time. They shows that the two filters have the same accuracy: the estimation error converges to zero with time passing by.

Another important index is the running time. We may have the intuition that the Gauss-Aitken filter will need much more time than the Kalman filter to finish the same filtering, but that is not true.

Fig. 5 Position estimation error

Fig. 6 Velocity estimation error

Running the two filters on a laptop whose configuration is shown in Table 1, it turns out that the running time of the Kalman filter is 0.002 0 s and that of the Gauss-Aitken filter is 0.002 4 s, which are in the same order of magnitude.

Table 1 Configuration of laptop

Brand	Hp TPN-Q117
Processor	Inter(R) Core(TM) i5-4200M CPU@2.5 GHz
RAM	4.00 GB

6. Conclusions

This article presents the construction and verification of the Gauss-Aitken filter, and provides a new way to derive the Kalman filter. The equivalence between the Gauss-Aitken filter and the Kalman filter is verified, and a simple example is provided to illustrate the properties and attributes of the Gauss-Aitken filter which can be concluded as follows:

- (i) The Gauss-Aitken filter is more stable and its robustness is better than the Kalman filter;
- (ii) The Gauss-Aitken filter and the Kalman filter have the same accuracy;
- (iii) The Gauss-Aitken filter can process the correlated observations, but the Kalman filter cannot, because the total covariance matrix R_{Y_n} is not limited to be diagonal;
- (iv) The Gauss-Aitken filter does not need initialization while the Kalman filter does;
- (v) The running time of the Gauss-Aitken filter and that of the Kalman filter are in the same order of magnitude.

Acknowledgment

The authors would like to thank Dr Norman Morrison for his contribution to the Gauss-Newton filter and providing insights into the fundamentals of filter engineering.

References

- [1] GAUSS C F. Theory of the combination of observations least subject to errors. California, USA: SIAM Press, 1995.

- [2] SWERLING P. A proposed stagewise differential correction procedure for satellite tracking and prediction. Santa Monica, CA: Rand Corp, 1958.
- [3] KALMAN R E. A new approach to linear filtering and prediction problems. Journal of Basic Engineering, 1960, 82(1): 33–45.
- [4] KALMAN R E, BUCY R S. New results in linear filtering and prediction theory. Journal of Basic Engineering, 1961, 83(1): 95–108.
- [5] YIN J, ZHAN X. Impact of bias-correction methods on effectiveness of assimilating SMAP soil moisture data into NCEP global forecast system using the ensemble Kalman filter. IEEE Geoscience & Remote Sensing Letters, 2018, 15(5): 659–663.
- [6] GUO L N, DING Y, WANG Z, et al. A dynamic load estimation method for nonlinear structures with unscented Kalman filter. Mechanical Systems & Signal Processing, 2018, 39(1): 254–273.
- [7] ALMOBAIED M, EKSIN I, GUZELKAYA M. Inverse optimal controller based on extended Kalman filter for discrete-time nonlinear systems. Optimal Control Applications & Methods, 2017, 23(1): 19–34.
- [8] MEI Z, WU B, BURSI O S, et al. Hybrid simulation of structural systems with online updating of concrete constitutive law parameters by unscented Kalman filter. Structural Control & Health Monitoring, 2018, 25(7): 2–10.
- [9] ATRSAEI A, SALARIEH H, ALASTY A, et al. Human arm motion tracking by inertial/magnetic sensors using unscented Kalman filter and relative motion constraint. Journal of Intelligent & Robotic Systems, 2017, 90(1/2): 161–170.
- [10] HUANG Y, ZHANG Y, WU Z, et al. A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices. IEEE Trans. on Automatic Control, 2017, 63(2): 594–601.
- [11] RONG H, PENG C, CHEN Y, et al. Adaptive gain regulation of extended Kalman filter for use in inertial and magnetic units based on hidden Markov model. IEEE Sensors Journal, 2018, 18(2): 3016–3027.
- [12] ZHANG C, JIA Y. Distributed Kalman consensus filter with event-triggered communication: formulation and stability analysis. Journal of the Franklin Institute, 2017, 354(13): 5486–5502.
- [13] HU J, XIONG R. Contact force estimation for robot manipulator using semi-parametric model and disturbance Kalman filter. IEEE Trans. on Industrial Electronics, 2018, 65(4): 3365–3375.
- [14] LEE K T, DAI M J, CHUANG C C. Temperature-compensated model for lithium-ion polymer batteries with extended Kalman filter state-of-charge estimation for an implantable charger. IEEE Trans. on Industrial Electronics, 2017, 65(1): 589–596.
- [15] BARRAU A, BONNABEL S. The invariant extended Kalman filter as a stable observer. IEEE Trans. on Automatic Control, 2017, 62(4): 1797–1812.
- [16] ZHAO J. Dynamic state estimation with model uncertainties using H-infinity extended Kalman filter. IEEE Trans. on Power Systems, 2017, 33(2): 1099–1100.
- [17] KORDESTANI M, SAMADI M F, SAIF M, et al. A new fault prognosis of MFS system using integrated extended Kalman filter and Bayesian method. IEEE Trans. on Industrial Informatics, 2018, 1(1): 1.
- [18] LIN H, HUANG Y, TANG X, et al. Robust multiple update-rate Kalman filter for new generation navigation signals carrier tracking. GPS Solutions, 2018, 22(1): 1–10.
- [19] PAN Y, YE H, HE K. Spherical simplex unscented Kalman filter-based jumping and static interacting multiple model. In-

- ternational Journal of Pattern Recognition & Artificial Intelligence, 2018, 32(2): 4–16.
- [20] THEIS T N, WONG H S P. The end of Moore's law: a new beginning for information technology. *Computing in Science & Engineering*, 2017, 19(2): 41–50.
- [21] FLEETWOOD D M. Evolution of total ionizing Dose effects in MOS devices with Moore's law scaling. *IEEE Trans. on Nuclear Science*, 2017, 65(8): 1465–1481.
- [22] GREENSTEIN S. Moore's law and economic architectures. *IEEE Micro*, 2017, 37(4): 82–84.
- [23] GARGINI P A. How to successfully overcome inflection points, or long live Moore's law. *Computing in Science & Engineering*, 2017, 19(2): 51–62.
- [24] GLENN I C, ABDULHAI S, PONSKY T A. Role of new media for the young pediatric surgeon: fighting exponential knowledge growth with Moore's law. *European Journal of Pediatric Surgery*, 2017, 27(3): 218–222.
- [25] TURKOT B, CARSON S, LIO A. Continuing Moore's law with EUV lithography. *Proc. of the Electron Devices Meeting*, 2018: 346–347.
- [26] SHALF J M, LELAND R. Computing beyond Moore's law. *Computer*, 2015, 48(12): 14–23.
- [27] CUSUMANO M A, YOFFIE D B. Extrapolating from Moore's law. *Communications of the ACM*, 2015, 59(1): 33–35.
- [28] NADJIASNGAR R, INGGS M. Gauss-Newton filtering incorporating Levenberg-Marquardt methods for tracking. *Digital Signal Processing*, 2013, 23(5): 1662–1667.
- [29] NADJIASNGAR R, MIDDLETON S, INGGS M. Doppler-only tracking with the recursive Gauss-Newton filter. *Proc. of the IET International Conference on Radar Systems*, 2013: 1–5.
- [30] MORRISON N, LORD R T, INGGS M R. The Gauss-Newton algorithm in passive aircraft tracking using Doppler and bearings. *Proc. of the IET International Conference on Radar Systems*, 2009: 1–5.
- [31] MORRISON N, LORD R T, INGGS M R. The Gauss-Newton algorithm applied to track-while-scan radar. *Proc. of the IET International Conference on Radar Systems*, 2009: 6–10.
- [32] MORRISON N. Tracking filter engineering: the Gauss-Newton and polynomial filters. London: The Institution of Engineering and Technology Press, 2013.
- [33] VALAPPIL J, GEORGAKIS C A. Systematic tuning approach for the use of extended Kalman filters in batch processes. *Proc. of American Control Conference*, 1999, 2: 1143–1147.

Biographies

BAO Zhichao was born in 1991. He received his B.S. and M.S. degrees from Electronic Engineering Institute, Hefei, Anhui, China, in 2014 and 2016 respectively. He is currently pursuing his Ph.D. degree in National University of Defense Technology, Hefei, Anhui, China. His current research interests mainly focus on radar target tracking.
E-mail: baozhichao520@sina.com

JIANG Qiuxi was born in 1960. He graduated from Xidian University, Xi'an, Shannxi, China. He is a professor and doctoral supervisor at National University of Defense Technology University. From 2010 to 2015, he has authored two books: *Network Radar Countermeasure Systems* (National Defence Industry Press, 2010) and *Introduction to Innovative Engineering* (National Defence Industry Press, 2014). His current research interests include signal and data processing, and radar countermeasure technology.
E-mail: jsc2013@sina.com