

State estimation under non-Gaussian Lévy noise:

A modified Kalman filtering method *

Xu Sun^{1,2}, Jinqiao Duan^{2,3}, Xiaofan Li² and Xiangjun Wang¹

1. Huazhong University of Science and Technology

Wuhan 430074, Hubei, China

E-mail: xjwang@hust.edu.cn

2. Illinois Institute of Technology

Chicago, IL 60616, USA

E-mail: lix@iit.edu

3. Institute for Pure and Applied Mathematics

University of California

Los Angeles, CA 90095, USA

E-mail: jduan@ipam.ucla.edu

March 12, 2013

Abstract

The Kalman filter is extensively used for state estimation for linear systems under Gaussian noise. When non-Gaussian Lévy noise is present, the conventional Kalman filter may fail to be effective due to the fact that the non-Gaussian Lévy noise may have infinite variance. A modified Kalman filter for linear systems with non-Gaussian Lévy noise is devised. It works effectively with reasonable computational cost. Simulation results are presented to illustrate this non-Gaussian filtering method.

*This work was partly supported by the NSF grant DMS-1025422.

Keywords: Kalman filter, modified Kalman filter, Non-Gaussian noise, Lévy noise, state estimation, data assimilation

1 Introduction and statement of the problem

The Kalman filter, or the Kalman filtering method, provides an efficient way to estimate the state of a linear dynamical system subject to Gaussian white noise [5, 6, 7]. It has been widely used in applications such as target tracking, parameter estimation, control theory, signal processing, and other data assimilation tasks.

The Kalman filter requires the noise be either Gaussian or with finite variance [5, 6], and thus it is not applicable to linear systems with non-Gaussian noise of infinite variance. As non-Gaussian Lévy noise with infinite variance exists ubiquitously [10, 9], it is desirable to study the Kalman filtering problems under Lévy noise. Very little work has been done for this issue. Breton and Musiela [4] presented a scheme for Kalman filtering with noise of infinite variance, while assuming the contribution of the jumps are exactly known. The filter in [4] is nonlinear and recursive, and thus may greatly limit its application in practice. Ahn and Feldman [1] proposed to minimize the difference between the true state and the filtered observation in the L^μ -norm. However, as pointed in [9], this method does not really address the Kalman filtering problem which consists of combining forecasts and observations. The method in [9] focuses on large errors and has a robust performance but high computational cost due to the matrix diagonalization and the operation of the fractional power in each step. This may also be an obstacle for real-time implementation in practical applications. Note that in practice, each iteration step must be completed during every sampling period, and it is greatly desirable to make the algorithm as fast as possible.

In this paper, we will present an algorithm, which has the similar computational cost as that of the Kalman filter, but can be applied to linear

systems with non-Gaussian Lévy noise of infinite variance.

We consider the following discrete time model with the state equation

$$\mathbf{x}_{k+1} = \mathbf{F}_k \mathbf{x}_k + \mathbf{w}_k, \quad (1)$$

and the observation equation

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{v}_k, \quad (2)$$

where \mathbf{x}_k , an n -by-1 vector, is the state variable, and \mathbf{z}_k , an m -by-1 vector, is the measurement (or observation) variable, \mathbf{w}_k represents the modeling error noise, \mathbf{v}_k the measurement error noise, and \mathbf{F}_k and \mathbf{H}_k are n -by- n and m -by- n matrices, respectively. We only consider the cases where \mathbf{w}_k is a Gaussian noise, and \mathbf{v}_k is a non-Gaussian Lévy noise.

This paper is arranged as follows. In section 2, the usual Kalman filter is briefly reviewed. The proposed modified Kalman filter is presented in section 3. A simulation example is provided in Section 4 to illustrate the effectiveness of the modified Kalman filter.

2 Review of the conventional Kalman filter

A derivation of the Kalman filter is briefly reviewed in this section. Some equations and ideas presented in this section will be used to present our proposed modified Kalman filter in the next section. Derivations of the Kalman filter can be found in many references [5, 6, 7].

Consider the model as given in (1) and (2). The Kalman filtering assumes that both the modeling error noise \mathbf{w}_k and the measurement disturbance \mathbf{v}_k are Gaussian with the following covariance matrix,

$$E [\mathbf{w}_i \mathbf{w}_k^T] = \begin{cases} \mathbf{Q}_k, & \text{for } i = k, \\ 0, & \text{for } i \neq k. \end{cases} \quad (3)$$

$$E[\mathbf{v}_i \mathbf{v}_k^T] = \begin{cases} \mathbf{R}_k, & \text{for } i = k, \\ 0, & \text{for } i \neq k. \end{cases} \quad (4)$$

Let $\bar{\mathbf{x}}_k$ be the priori estimate, which is the estimate of \mathbf{x}_k given $\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_{k-1}$, and let $\hat{\mathbf{x}}_k$ be the posterior estimate, which is the estimate of \mathbf{x}_k given $\mathbf{z}_0, \mathbf{z}_1, \dots, \mathbf{z}_k$. It is known that

$$E\{\bar{\mathbf{x}}_k\} = E\{\mathbf{x}_k\} \quad (5)$$

and

$$\bar{\mathbf{x}}_{k+1} = \mathbf{F}_k \hat{\mathbf{x}}_k, \quad (6)$$

where $E\{\cdot\}$ represents expectation and \mathbf{F}_k is from (1).

The Kalman filter assumes that the posterior estimate is expressed as the prior estimate corrected by the measurement data,

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H}_k \bar{\mathbf{x}}_k), \quad (7)$$

for some n -by- m matrix \mathbf{K}_k (so called Kalman gain). Note that \mathbf{H}_k is from (2). The Kalman gain \mathbf{K}_k is solved by minimizing $E[(\hat{\mathbf{x}}_k - \mathbf{x}_k)^2]$. Note that

$$E[(\hat{\mathbf{x}}_k - \mathbf{x}_k)^T (\hat{\mathbf{x}}_k - \mathbf{x}_k)] = \text{Tr}\{\mathbf{P}_k\}, \quad (8)$$

where $\text{Tr}\{\cdot\}$ represents the trace operator, and the n -by- n covariance matrix \mathbf{P}_k is defined as follows

$$\mathbf{P}_k = E[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T]. \quad (9)$$

Define

$$\bar{\mathbf{P}}_k = E[(\mathbf{x}_k - \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T]. \quad (10)$$

It follows from (3), (4), (9) and (10) that

$$\bar{\mathbf{P}}_{k+1} = \mathbf{F}_k \mathbf{P}_k \mathbf{F}_k^T + \mathbf{Q}_k. \quad (11)$$

Substituting (7) into (9), we get

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \bar{\mathbf{P}}_k (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T. \quad (12)$$

It follows from (12) that

$$\frac{d}{d\mathbf{K}_k} \text{Tr}\{\mathbf{P}_k\} = -2(\mathbf{H} - \mathbf{K} \bar{\mathbf{P}}_k)^T + 2\mathbf{K}_k (\mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T + \mathbf{R}_k). \quad (13)$$

Solve \mathbf{K}_k by letting $\frac{d}{d\mathbf{K}_k} \text{Tr}\{\mathbf{P}_k\} = 0$, we get

$$\mathbf{K}_k = \bar{\mathbf{P}}_k \mathbf{H}_k^T (\mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T + \mathbf{R}_k)^{-1} \quad (14)$$

By (12) and (14), \mathbf{P}_k can be rewritten as

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \bar{\mathbf{P}}_k. \quad (15)$$

Combining (6), (14) and (15), we thus have the conventional Kalman filter. This algorithm is shown in Figure 1.

3 A modified Kalman filter

It is known that the discrete time Gaussian white noise can be approximated by the increments of Brownian motion per time step, and the non-Gaussian Lévy noise can be approximated by the increments of the corresponding Lévy process per time step. By Lévy-Ito theorem [2], a Lévy process can be decomposed into the sum of a Gaussian process and a pure jump process. It is shown in [3] that the small jumps of a Lévy process can be approximated by a Gaussian process. Therefore, we can approximately regard a Lévy process as combination of a Gaussian process and a process with big jumps. For more information about decomposition of a Lévy processes, see [3, 2]. These results enable us to decompose a non-Gaussian Lévy noise into a Gaussian white noise plus some extremely large values.

In our proposed filtering method, we convert the original Lévy noise into a Gaussian white noise by clipping off its extremely large values.

Let $\tilde{\mathbf{v}}_k$ represent the clipped version of the Lévy measurement disturbance \mathbf{v}_k , and let $\tilde{\mathbf{z}}_k$ represent the corresponding clipped observation. Thus

$$\tilde{\mathbf{z}}_k = \mathbf{H}_k \mathbf{x}_k + \tilde{\mathbf{v}}_k. \quad (16)$$

In practice, since the measurement noise, \mathbf{v}_k , is unknown, we propose to clip the observation \mathbf{z}_k instead of \mathbf{v}_k in a component-wise way by the following operation:

$$\tilde{\mathbf{z}}_k^i = \begin{cases} \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j + C \cdot \text{sign} \left(\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j \right) & \text{if } |\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j| \geq C, \\ \mathbf{z}_k^i & \text{if } |\mathbf{z}_k^i - \sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j| < C, \end{cases} \quad (17)$$

where C is some positive threshold value, \mathbf{z}_k^i and $\bar{\mathbf{x}}_k^i$ represent the i -th components of the vectors \mathbf{z}_k and $\bar{\mathbf{x}}_k$, respectively, and $\sum_j \mathbf{H}_k^{i,j} \bar{\mathbf{x}}_k^j$ is the i -th component of the vector $\mathbf{H}_k \bar{\mathbf{x}}_k$. Note that C is determined by the statistical properties of the measurement noise \mathbf{v}_k . Replacing the observation value \mathbf{z}_k in (7) with its clipped value, we get

$$\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k + \mathbf{K}_k (\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k). \quad (18)$$

Repeating the same procedure in Section 2 to solve the Kalman gain \mathbf{K}_k by minimizing $E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)^2\}$, we get

$$\mathbf{K}_k = \bar{\mathbf{P}}_k \mathbf{H}_k^T (\mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T + \tilde{\mathbf{R}}_k)^{-1}, \quad (19)$$

where $\tilde{\mathbf{R}}_k$ is the covariance matrix of $\tilde{\mathbf{v}}_k$ defined as

$$\tilde{\mathbf{R}}_k = E \{ \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^T \}. \quad (20)$$

In the conventional Kalman filter, \mathbf{Q}_k and \mathbf{R}_k are assumed to be known, and as noted in [6], it is often a difficult task to estimate the covariance matrices \mathbf{Q}_k and \mathbf{R}_k .

In the modified Kalman filter here, we only assume \mathbf{Q}_k is known and suggest $\tilde{\mathbf{R}}_k$ be estimated as follows. It follows from (16) and (17) that

$$\begin{aligned}\tilde{\mathbf{R}}_k &= E \{ \tilde{\mathbf{v}}_k \tilde{\mathbf{v}}_k^T \} \\ &= E \left\{ [(\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k) - \mathbf{H}_k (\mathbf{x}_k - \bar{\mathbf{x}}_k)] [(\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k) - \mathbf{H}_k (\mathbf{x}_k - \bar{\mathbf{x}}_k)]^T \right\} \\ &= (\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k)(\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k)^T + \mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T.\end{aligned}\quad (21)$$

In deriving the last identity of (21), we have used the fact that $\tilde{\mathbf{z}}_k$ and $\bar{\mathbf{x}}_k$ are known values and

$$E \{ (\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T \mathbf{H}_k^T \} = (\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k) E \{ (\mathbf{x}_k - \bar{\mathbf{x}}_k)^T \} \mathbf{H}_k^T = 0. \quad (22)$$

With (21), (19) can be rewritten as

$$\mathbf{K}_k = \bar{\mathbf{P}}_k \mathbf{H}_k^T (2 \cdot \mathbf{H}_k \bar{\mathbf{P}}_k \mathbf{H}_k^T + \tilde{\mathbf{R}}_k)^{-1}, \quad (23)$$

where

$$\tilde{\mathbf{R}}_k = (\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k)(\tilde{\mathbf{z}}_k - \mathbf{H}_k \bar{\mathbf{x}}_k)^T. \quad (24)$$

Combining equations (17), (18), (23), and (24), we obtain the modified Kalman filter, as shown graphically in Figure 2.

Comparing with the conventional Kalman filter, the proposed filter has an moderately increased computational cost due to the following two operations: i) the clipping operation for \mathbf{z}_k ; ii) the computation of $\tilde{\mathbf{R}}_k$. The former operation is implemented by IF-ELSE sentence, and the latter is simply a vector-vector outer product.

4 Simulation results

Consider a particle moving in the plane at some velocity subject to random perturbations in its trajectory. The new position at time $k + 1$ is equal to

the old position at time k plus the velocity and noise. The model can be expressed in form of (1) and (2) as

$$\begin{pmatrix} x_{k+1}^1 \\ x_{k+1}^2 \\ u_{k+1}^1 \\ u_{k+1}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_k^1 \\ x_k^2 \\ u_k^1 \\ u_k^2 \end{pmatrix} + \begin{pmatrix} w_k^1 \\ w_k^2 \\ w_k^3 \\ w_k^4 \end{pmatrix}, \quad (25)$$

and

$$\mathbf{z}_k = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \mathbf{x}_k + \begin{pmatrix} v_k^1 \\ v_k^2 \end{pmatrix}, \quad (26)$$

where (x_k^1, x_k^2) is the position at time k , u_k^1, u_k^2 the velocity, w_k^1, w_k^2, w_k^3 , and w_k^4 are all Gaussian white noises with zero mean and unit variance, and v_k^1 and v_k^2 are independent and identically distributed noises consist of two components: i) a symmetric α -stable Lévy noises with the index of stability $\alpha = 1.3$ and the scale parameter $\sigma = 10$ (see [8]); ii) a Gaussian white noise with variance of 5. Since the measurement noises, v_k^1 and v_k^2 , have infinite variances, the conventional Kalman filter can not be applied to estimate the position x_k^1 and x_k^2 . So we apply the modified Kalman filter proposed in the previous section.

Take $x_0^1 = 10, x_0^2 = 10, u_0^1 = 1, u_0^2 = 0$, and we apply the modified Kalman filtering method to estimate x_k^1 and x_k^2 . In the simulation, the initial a priori estimate of the state, $(\bar{x}_0^1, \bar{x}_0^2)$, is set to be equal to the observation at time 0, its error covariance, \bar{P}_0 , is set to be unit matrix, and the threshold value C is set to be 40. The simulation results are shown in Figure 3, where the estimate position error, ER , defined by

$$ER_k = \sqrt{(z_k^1 - x_k^1)^2 + (z_k^2 - x_k^2)^2}, \quad (27)$$

is compared with the observed position error, OR , defined by

$$OR_k = \sqrt{(\bar{x}_k^1 - x_k^1)^2 + (\bar{x}_k^2 - x_k^2)^2}. \quad (28)$$

The results in Figure 3 are calculated by averaging 10,000 times of simulations. It is seen from this figure that the position estimation error is significantly improved by using our modified Kalman filter.

In the simulation, we select C by the method of trial and error, and it is found that the threshold value C is not very picky and C can vary from 30 to 100 without significant effects on the performance of the modified Kalman filter. Determining the optimal C , which is crucial for the modified Kalman filter, deserves further research and will be left for our future work.

References

- [1] A. Ahn and R. Feldman. Optimal filtering of a gaussian signal in the presence of Lévy noise. *SIAM Journal of Applied Mathematics*, 60:359–369, 1999.
- [2] D. Applebaum. *Lévy Processes and Stochastic Calculus*. Cambridge University Press, 2nd Edition, 2009.
- [3] S. Asmussen and J. Rosinski. Approximation of small jumps of Lévy processes with a view towards simulation. *Journal of Applied Probability*, 38:482–493, 2001.
- [4] A. Breton and M. Musiela. A generation of the kalman filter to models with infinite variance. *Stochastic Processes and Applications*, 47:75–94, 1993.
- [5] R. Brown and P. Hwang. *Introduction to Random Signals and Applied Kalman Filtering*. John Wiley & Sons, 3rd Edition, 1997.
- [6] M. Grewal and A. Andrews. *Kalman Filtering: Theory and Practice Using Matlab*. John Wiley & Sons, 3rd Edition, 2008.
- [7] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York, 1970.

- [8] A. Janicki and A. Weron. *Simulation and Chaotic Behavior of α -Stable Stochastic Processes*, Marcel Dekker, Inc., 1994.
- [9] D. Sornette and K. Ide. The Kalman-Lévy filter. *Physica D*, 151:142–174, 2001.
- [10] W. A. Woyczynski. Lévy processes in the physical sciences. In *Lévy Processes: Theory and Applications*, O. E. Barndorff-Nielsen, T. Mikosch and S. I. Resnick (Eds.), 241-266, Birkhäuser, Boston, 2001.

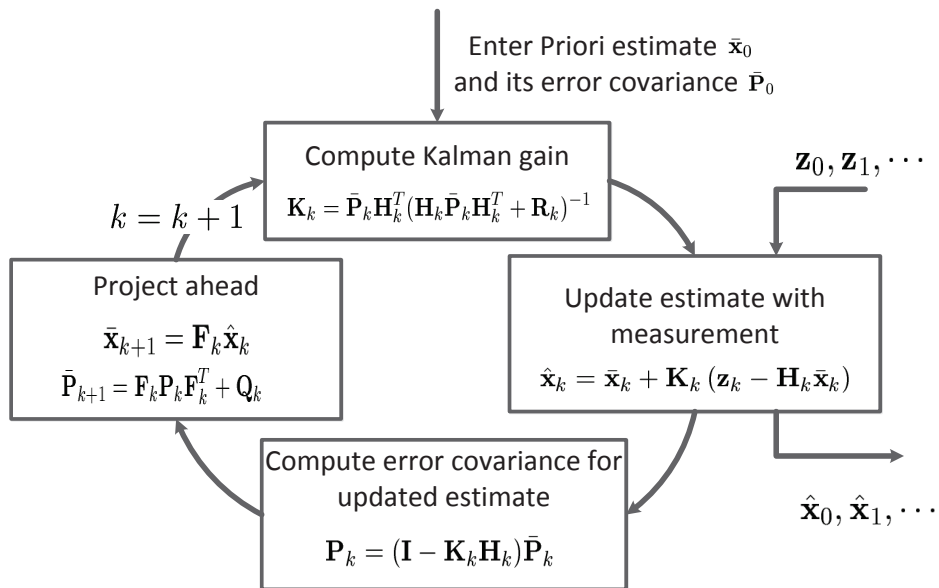


Figure 1: The usual Kalman filtering algorithm

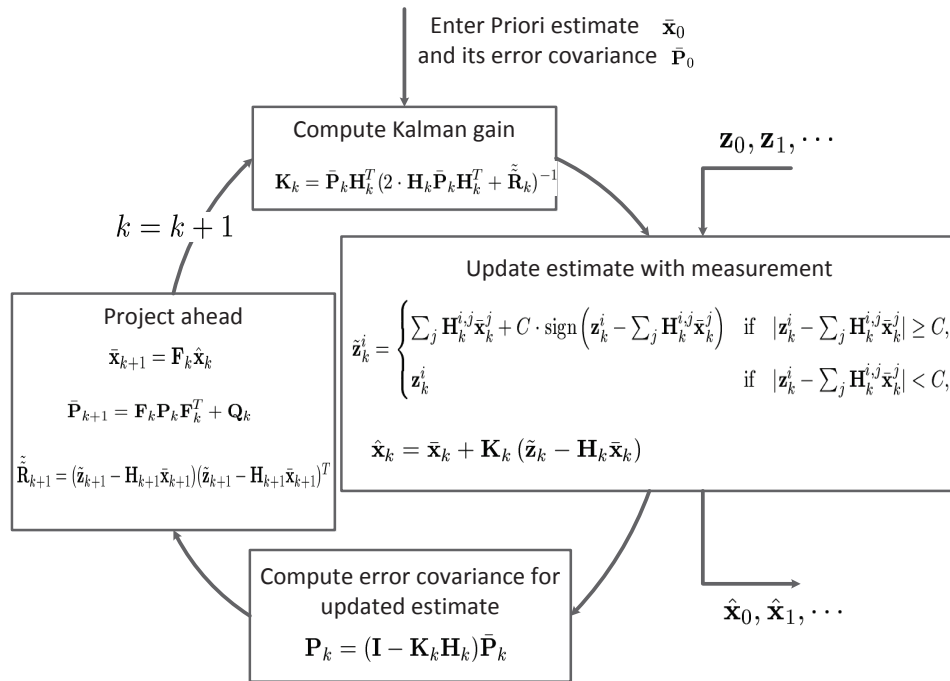


Figure 2: The modified Kalman filtering algorithm

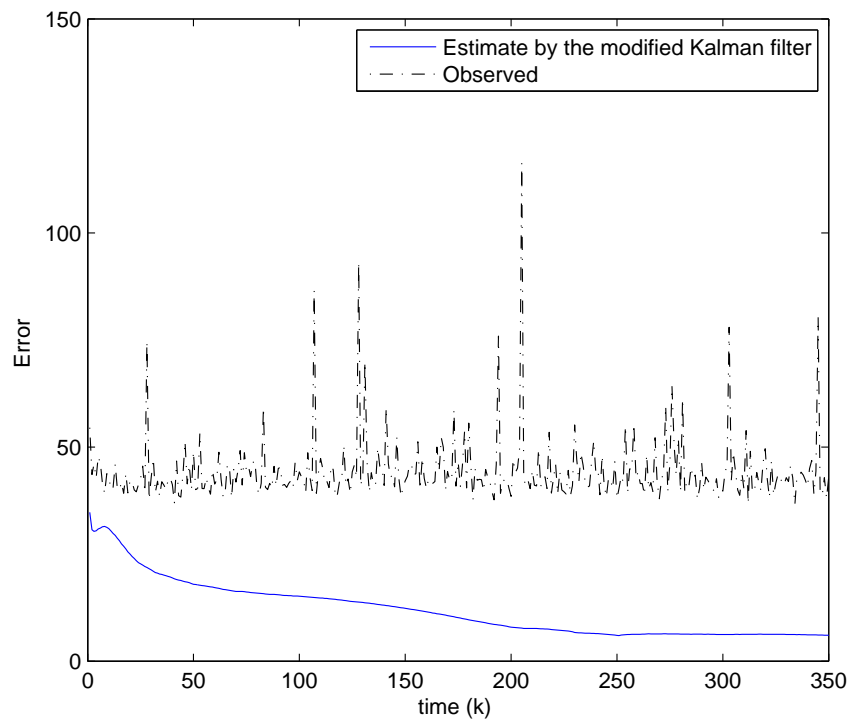


Figure 3: The error of the modified Kalman filtering algorithm