Ridge-Type Kalman Filter and Its Algorithm

YONGMING LI
Institute of Science,
Information Engineering University
450001,Zhengzhou
CHINA
keepming@163.com

QINGMING GUI
Institute of Science,
Information Engineering University
450001,Zhengzhou
CHINA
guiqingming@126.com

YONGWEI GU
Institute of Science
Information Engineering University
450001,Zhengzhou
CHINA
gyw1019@sina.com

SONGHUI HAN
Institute of Science
Information Engineering University
450001,Zhengzhou
CHINA
hansonghui@126.com

KAI DU
College of Textiles & Fashion
Qingdao University
266071,Qingdao
CHINA
qddk2008@163.com

Abstract: Kalman filter is one of common ways of dealing with dynamic data under linear minimum variance estimation. However, in practice, the observation matrix maybe has multicollinearity. Therefore, the impact of multicollinearity on Kalman filter is studied in this paper. Firstly, by analyzing the normal equation of Kalman filter, we proposed sufficient conditions about how the multicollinearity affects the estimator. Secondly, a ridge-type Kalman filter algorithm is designed under the mean square error (MSE), and then the characters of the new algorithm are analyzed. Thirdly, six specific methods for determining the ridge parameters of ridge-type Kalman filter are proposed based on the canonical form of normal equation. Finally, examples illustrate the new algorithms can overcome the influence of the ill-condition on Kalman filter effectively, which improves the accuracy of the estimates of parameters.

Key–Words: Kalman filter, Ill-condition, Ridge-Type Kalman filter, Ridge parameter

1 Introduction

Kalman filter soon became one of most common ways of dealing with dynamic data since it was proposed in 1960[1]. The equations of Kalman filter are recursion formula and the calculation of it is a continuous process for predicting and revising. Kalman filter does not need to store a mass of data when calculating. The amount of calculation is greatly cut down, and it can handle the estimation problem in real time[2]. Therefore, Kalman filter has been widely used in engineering [3-7].

The discrete linear system of Kalman filter consists of state equation and observation equation [4,5]. When we only use observation equation to obtain the solutions under least square method(LS), the normal equation appears ill-conditioned if the observation matrix is multicollinearity. Then the estimators of the state parameters will be influenced by the ill-condition, leading to poor accuracy. However, the state equation in the discrete linear system supplies much more information for estimating parameters. Thus, it remains to demonstrate whether the information supplied by state equation can effectively control the ill-condition of normal equation so as to improve the estimating accuracy or not. Few approaches have been proposed in the literature for analyzing this problem, not mention to give a nature scheme.

In this paper, Kalman filter and its algorithm are introduced in Section 2. Section 3 analyzes the ill-condition of normal equation of Kalman filter caused by multicollinearity of observation matrix, and illustrates the effect of state equation in controlling the ill-condition. A new type of filterłridge-type Kalman filter is proposed by combining ridge regression estimation with Kalman filter in the sense of the MSE in Section 4. Besides, some characters of the new algorithm are given in this section. Six methods to determine the ridge parameter of ridge-type Kalman filter are given in Section 5. In Section 6, the new algorithm is simulated and analyzed. Finally, some brief conclusions are given in Section 7.

2 Kalman Filter and Its Algorithm

2.1 Discrete Linear System

It is assumed that state equation and observation equation are as follows:

$$X_k = \Phi_{k,k-1} X_{k-1} + W_{k-1}, \tag{1}$$

$$Z_k = H_k X_k + V_k, \tag{2}$$

respectively, where X_k denotes a $p \times 1$ vector describing the state of the system at time t_k and Z_k is the observation. The $p \times p$ matrix $\Phi_{k,k-1}$ is the state transition matrix at time at time t_k. The $n \times p$ matrix H_k is the observation matrix. The vectors W_{k-1} and V_k is the system noise and the observation noise, respectively.

It is assumed that

$$\begin{cases} E(W_k) = 0, Cov(W_k, W_j) = Q_k \delta_{k,j} \\ E(V_k) = 0, Cov(V_k, V_j) = R_k \delta_{k,j} \\ Cov(W_k, V_j) = 0 \end{cases}$$

where the system noise covariance Q_k is assumed to be a non-negative definite matrix and the observation noise covariance R_k is assumed to be a positive definite matrix. $\delta_{k,j}$ is the function of the Kronecher - δ.

2.2 Kalman Filter Algorithm

Kalman filter estimates a system by using a form of feedback control: the filter estimates the state parameters at some time and then obtains feedback in the form of (noisy) observations[1]. As such, the equations for Kalman filter fall into two groups: time update equations and observation update equations.

The specific equations for the time and observation updates are presented below in Table 1 and Table 2 respectively, where $\widehat{X}_{k,k-1}$ and $P_{k/k-1}$ are the prediction state and the prediction covariance at time t_k. K_k is the filter gain and P_k is the error covariance matrix of the estimator $\widehat{X}_k$.

Table 1: The time update equations of Kalman filter at time t_k

$$\widehat{X}_{k/k-1} = \Phi_{k,k-1} \widehat{X}_{k-1} \tag{3}$$

$$P_{k/k-1} = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + Q_{k-1} \tag{4}$$

Table 2: The observation update equations of Kalman filter at time t_k

$$K_k = P_{k/k-1} H_k^T (H_k P_{k/k-1} H_k^T + R_k)^{-1} \tag{5}$$

$$\widehat{X}_k = \widehat{X}_{k/k-1} + K_k (Z_k - H_k \widehat{X}_{k/k-1}) \tag{6}$$

$$\begin{aligned} P_k = {} & (I - K_k H_k) P_{k/k-1} (I - K_k H_k)^T \\ & + K_k R_k K_k^T \end{aligned} \tag{7}$$

Equations (3)-(7) are the basic equations of Kalman filter. Once the initial state value $\widehat{X}_0$ and error covariance matrix P_0 are given, the state estimator $\widehat{X}_k$ (k = 1,2,$\cdots$) at time t_k can be recursively obtained based on the observation Z_k.

3 Analysis of the Ill-condition

If the observation equation (2) is used only, the LSE [6,7] of state parameters at time t_k is

$$\widehat{X}_k(LS) = (H_k^T R_k^{-1} H_k)^{-1} H_k^T R_k^{-1} Z_k. \tag{8}$$

Concluded from perturbation analysis theory, the solution of the equation above will be greatly influenced by the ill-condition when the observation matrix H_k appears multicollinearity.

Equation (6) can be transformed as the following form,

$$\begin{aligned} \widehat{X}_k = {} & (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1})^{-1} (H_k^T R_k^{-1} Z_k \\ & + P_{k/k-1}^{-1} \widehat{X}_{k/k-1}), \end{aligned} \tag{9}$$

where $H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}$ is called the normal matrix. Equation (9) is the solution of the linear equation (10),

$$\begin{aligned} (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}) X_k = {} & H_k^T R_{k/k-1}^{-1} Z_k \\ & + P_{k/k-1}^{-1} \widehat{X}_{k/k-1} \end{aligned} \tag{10}$$

which is the normal equation of Kalman filter.

The solution of Kalman filter is expected to be better than LSE when the state equation supplies more information. The format of Equation (9) looks much similar to that of the solution of Tikhonov regularization [8]. Tikhonov regularization is the most commonly used method of regularization of ill-posed problems. In statistics, the method is known as ridge regression, and, with multiple independent discoveries, it is also variously known as the Tikhonov-Miller method, the Phillips-Twomey method, the constrained linear inversion method, and the method of linear regularization. Outwardly it seems that Equation (9) is improved from Equation (8) with Tikhonov regularization method, intending to improve the accuracy. But as a matter of fact, $P_{k/k-1}^{-1}$ may not improve the extent of normal matrix $H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}$ to be singular, because it is not a proper regularization matrix that fits Tikhonov regularization method, but one that has to be accepted in the solution process of Kalman filter. However, the ill-condition of normal equation is still influenced by $P_{k/k-1}^{-1}$.

The condition number as follows is commonly used to quantitatively describe the ill-condition of normal equation.

$$Cond(N_k) = \frac{\lambda_k^1}{\lambda_k^p},$$

where N_k denotes the normal matrix at time t_k, and its spectral decomposition is $U_k^T N_k U_k = diag(\lambda_k^1, \lambda_k^2, \cdots, \lambda_k^p)$. $\lambda_k^1 \geq \lambda_k^2 \geq \cdots \geq \lambda_k^p$ are the eigenvalues of N_k , and $U_k = (U_k^1, U_k^2, \cdots, U_k^p)$ is a orthogonal matrix formed by the orthogonal eigenvectors $U_k^1, U_k^2, \cdots, U_k^p$ which are corresponding to their eigenvalues. Obviously, the information supplied by state equation in normal matrix mainly embody on $P_{k/k-1}^{-1}$.

It is assumed that the spectral decomposition of the normal matrix $H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}$ is $U_k^T (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}) U_k = diag(\lambda_k^1, \lambda_k^2, \cdots, \lambda_k^p)$.

If $H_k^T R_k^{-1} H_k$ and $P_{k/k-1}^{-1}$ can simultaneously be diagonalized, that is

$$U_k^T (H_k^T R_k^{-1} H_k) U_k = diag(\lambda_k^{11}, \lambda_k^{12}, \cdots, \lambda_k^{1p}),$$

$$U_k^T (P_{k/k-1}^{-1}) U_k = diag(\lambda_k^{21}, \lambda_k^{22}, \cdots, \lambda_k^{2p}),$$

the relationship of their eigenvalues is $\lambda_k^i = \lambda_k^{1i} + \lambda_k^{2i}$.

When $Cond(N_k) \geq a$, we identify that normal equation is ill-conditioned. Therefore, if $N_k = H_k^T R_k^{-1} H_k$ is ill-conditioned, we can obtain $\frac{\lambda_k^{11}}{\lambda_k^{1p}} = c \geq a$. After adding to the matrix $P_{k/k-1}^{-1}$, it can be concluded that:

$$(1) \quad \frac{\lambda_k^1}{\lambda_k^p} = \frac{\lambda^{11} + \lambda^{21}}{\lambda^{1p} + \lambda^{2p}} > c, \quad for \ \frac{\lambda_k^{21}}{\lambda_k^{2p}} > c$$

$P_{k/k-1}^{-1}$ strengthens the ill-condition of normal equation in this case.

$$(2) \quad \frac{\lambda_k^1}{\lambda_k^p} = \frac{\lambda^{11} + \lambda^{21}}{\lambda^{1p} + \lambda^{2p}} = c, \quad for \ \frac{\lambda_k^{21}}{\lambda_k^{2p}} = c$$

$P_{k/k-1}^{-1}$ gets no impact on the ill-condition of normal equation in this case.

$$(3) \quad \frac{\lambda_k^1}{\lambda_k^p} = \frac{\lambda^{11} + \lambda^{21}}{\lambda^{1p} + \lambda^{2p}} < c, \quad for \ \frac{\lambda_k^{21}}{\lambda_k^{2p}} < c$$

$P_{k/k-1}^{-1}$ weakens the ill-condition of normal equation in this case.

From the basic equations of Kalman filter, it can be concluded that $P_{k/k-1}^{-1} = [E((X_k - \widehat{X}_{k/k-1})(X_k - \widehat{X}_{k/k-1})^T)]^{-1}$ is unlikely to have stronger multicollinearity. So the matrix $P_{k/k-1}^{-1}$ can weaken the ill-condition of normal equation to some extent in most cases. However, the state estimator is still greatly influenced by the ill-condition.

4 Ridge-Type Kalman Filter Algorithm and Its Characters

4.1 Ridge-Type Kalman Filter and Its Algorithm

To solve the problem that Kalman filter is easily influenced by the ill-condition, inspired by ridge regression estimation [9], we replace $(H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1})^{-1}$ in Equation (9) by the following formula in the new algorithm,

$$(H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1} + \lambda I)^{-1}, \qquad (11)$$

where $\lambda > 0$ is the ridge parameter and I is a $p \times p$ unit matrix.

The ridge-type Kalman filter algorithm is designed as follows.

The specific equations for the time and observation updates of ridge-type Kalman filter are presented below in table 3 and table 4, where the symbol RT means ridge-type Kalman filter.

Table 3: The time update equations of Kalman filter at time t_k

$\widehat{X}_{k/k-1} = \Phi_{k,k-1} \widehat{X}_{k-1}$
$P_{k/k-1} = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + Q_{k-1}$

Table 4: The observation update equations of Kalman filter at time t_k

$$\widehat{X}_k(RT) = (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1} + \lambda I)^{-1} (H_k^T R_k^{-1} Z_k + P_{k/k-1}^{-1} \widehat{X}_{k/k-1}) \quad (12)$$
$$P_k(RT) = (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1} + \lambda I)^{-1} (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1} + \lambda^2 X_k X_k^T) (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1} + \lambda I)^{-1}$$

The matrix $X_k X_k^T$ is replaced by $\widehat{X}_k(RT)\widehat{X}_k^T(RT)$ if it is unknown.

The state estimator varies if we choose different λ. Therefore, ridge-type Kalman filter is a class of estimators.

Kalman filter is a periodic recursion process. The state estimator and the error covariance matrix will be used as updates for the next period. The state estimator at time t_k becomes more precise when it is improved by the new algorithm. When working on state estimator at time t_{k+1}, it is necessary to replace the original updates of Kalman filter with that of the new algorithm. Then the state estimator at time t_{k+1} will be better than that of Kalman filter estimate.

The algorithm of Kalman filter during a filter period is a closed system under linear minimum variance. That is to say, unbiasedness of the state estimator at time t_{k+1} will not be influenced by the state estimator at time t_k. Therefore, the specific equations for the time and observation updates are presented below in Table 5 and Table 6.

Table 5: The time update equations of ridge-type Kalman filter at time t_{k+1}

$$\widehat{X}_{k+1/k}(RT) = \Phi_{k+1,k} \widehat{X}_k(RT)$$
$$P_{k+1/k} = \Phi_{k+1,k} P_k(RT) \Phi_{k+1,k}^T + Q_k$$

Table 6: The observation update equations of ridge-type Kalman filter at time t_{k+1}

$$\widehat{X}_{k+1} = (H_{k+1}^T R_{k+1}^{-1} H_{k+1} + P_{k+1/k}^{-1})^{-1} (H_{k+1}^T R_{k+1}^{-1} Z_{k+1} + P_{k+1/k}^{-1} \widehat{X}_{k+1/k})$$
$$P_{k+1} = (H_{k+1}^T R_{k+1/k}^{-1} H_{k+1} + P_{k+1/k}^{-1})$$

If the observation matrix at time still t_{k+1} has stronger multicollinearity, it has to be improved again with the new algorithm. The new algorithm sacrifices the unbiasedness of the estimator to obtain a smaller MSE. So if the observation matrix at time t_{k+1} has no or less multicollinearity, the ill-condition of the normal equation has less influence on Kalman filter. Then there is no need to sacrifice the unbiasedness of the estimator for minor MSE.

4.2 Characters of the New Algorithm

1) The estimator of ridge-type Kalman filter is the linear transformation of that of Kalman filter.

Proof: From equation (9) and (12) it is readily established that

$$\widehat{X}_k(RT) = B_k \widehat{X}_k, \quad (13)$$

where $B_k = (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1} + \lambda I)^{-1} (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1})$.

2) The estimator of ridge-type Kalman filter is a biased estimator of the state parameter X_k.

Proof: Concluded from equation (13), the expectation of the state estimator $\widehat{X}_k(RT)$ is

$$E(\widehat{X}_k(RT)) = E(B_k \widehat{X}_k) = B_k E(\widehat{X}_k) = B_k X_k.$$

It is biased for $\lambda > 0$.

3) There always exists $\lambda > 0$ such that $MSE(\widehat{X}_k(RT)) < MSE(\widehat{X}_k)$. That is to say, there always exists $\widehat{X}_k(RT)$ that is more accurate than $\widehat{X}_k$ in the sense of the MSE.

Proof: The canonical form of the linear equation (10) is introduced in order to compute the MSE easier[10].

$$H_k^T R_k^{-1} Z_k + P_{k/k-1}^{-1} \widehat{X}_{k/k-1} = A_k \alpha_k,$$

where $A_k = (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}) U_k$; $\alpha_k = (\alpha_k^1, \alpha_k^2, \cdots, \alpha_k^p)^T = U_k^T X_k$ is called the canonical parameter. $U_k = (u_k^1, u_k^2, \cdots, u_k^p)^T$ is a orthogonal matrix formed by the orthogonal eigenvectors $u_k^1, u_k^2, \ldots, u_k^p$ which are corresponding to their eigenvalues. Therefore,

$$U_k^T (H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}) U_k = \Lambda_k,$$

where $\Lambda_k = diag(\lambda_k^1, \lambda_k^2, \cdots, \lambda_k^p)$ is a diagonal matrix containing the eigenvalues of $H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}$.

The LSE of the canonical parameter α_k is

$$\begin{aligned}\widehat{\alpha_k} &= (\widehat{\alpha}_k^1, \widehat{\alpha}_k^2, \cdots, \widehat{\alpha}_k^p)^T \\ &= \Lambda_k^{-1} U_k^T (H_k^T R_k^{-1} Z_k + P_{k/k-1}^{-1} \widehat{X}_{k/k-1}).\end{aligned}$$

From the canonical form, it can be concluded that $\widehat{X}_k = U_k\widehat{\alpha}_k$, and $MSE(\widehat{X}_k) = MSE(\widehat{\alpha}_k)$, and they are correct in the general situation. That is, $MSE(\tilde{X}_k) = MSE(\tilde{\alpha}_k)$ for any estimator of α_k. The ridge-type Kalman filter estimator of the canonical parameter α_k is $\widehat{\alpha}_k(RT) = (\Lambda_k + \lambda I)^{-1}\Lambda_k\widehat{\alpha}_k$. Hence,

$$\begin{aligned} & MSE(\widehat{X}_k(RT)) = MSE(\widehat{\alpha}_k(RT)) \\ &= trace(Cov(\widehat{\alpha}_k(RT))) + \| E(\widehat{\alpha}_k(RT)) - \alpha_k \|^2 \\ &= \sum_{i=1}^{p} \frac{\lambda_k^i}{(\lambda_k^i + \lambda)^2} + \lambda^2 \sum_{i=1}^{p} \frac{(\alpha_k^i)^2}{(\lambda_k^i + \lambda)^2} \\ &= g(\lambda) \end{aligned}$$

The derivation of $g(\lambda)$ is

$$g'(\lambda) = -2\sum_{i=1}^{p} \frac{\lambda_k^i}{(\lambda_k^i + \lambda)^3} + 2\lambda \sum_{i=1}^{p} \frac{\lambda_k^i(\alpha_k^i)^2}{(\lambda_k^i + \lambda)^3}.$$

Because $g'(\lambda)$ is a continuous function of λ when $\lambda > 0$, besides $\lim_{\lambda\to 0} g'(\lambda)$ exists and it is negative, there always exists $\delta > 0$ making $g'(\lambda) < 0$ for $0 < \lambda < \delta$ [11]. That is to say, $g(\lambda) = MSE(\widehat{X}_k(RT))$ is a decreasing function of λ in the interval $(0, \delta)$. Therefore, there always exists $\lambda \in (0, \delta)$ making $g(\lambda) < \lim_{\lambda\to 0} g(\lambda)$ which means $MSE(\widehat{X}_k(RT)) < MSE(\widehat{X}_k)$.

4) The estimator $\widehat{X}_k(RT)$ is the compression of the estimator $\widehat{X}_k$ to the origin.

Proof: By calculating, it is clear that

$$\begin{aligned} \| \widehat{X}_k(RT) \|^2 &= \| (\Lambda_k + \lambda I)^{-1}\Lambda_k U_k^T \widehat{X}_k \|^2 \\ &< \| U_k^T \widehat{X}_k \|^2 = \| \widehat{X}_k \|^2 \end{aligned}$$

5 Choosing the Ridge Parameter

It is important to choose an appropriate ridge parameter in practice. Six specific methods are proposed based on the analysis of the MSE in Section 4.

It is clear that

$$g'(\lambda) < 0 \ , \ \text{for} \ \ 0 < \lambda < \frac{1}{(\widehat{\alpha}_k^i)^2_{max}} \ ,$$

and

$$g'(\lambda) > 0 \ , \ \text{for} \ \ \frac{1}{(\widehat{\alpha}_k^i)^2_{min}} < \lambda \ ,$$

where $(\widehat{\alpha}_k^i)^2_{max}$ is the maximum element of $\widehat{\alpha}_k^2$ and $(\widehat{\alpha}_k^i)^2_{min}$ is the minimum element of $\widehat{\alpha}_k^2$.

Thus, there exists a unique minimum of MSE for $\lambda > 0$, which lies in the interval

$$[\frac{1}{(\widehat{\alpha}_k^i)^2_{max}} \ , \ \frac{1}{(\widehat{\alpha}_k^i)^2_{min}}].$$

It is expected to get a ridge parameter making use of the known data, which makes $MSE(\widehat{X}_k(RT)) < MSE(\widehat{X}_k)$. Therefore, six specific methods to obtain the ridge parameter are proposed based on different criteria.

Method(1):

$$\lambda = \frac{1}{(\widehat{\alpha}_k^i)^2_{max}}$$

Method(2):

$$\lambda = \frac{p}{\sum_{i=1}^{p} (\widehat{\alpha}_k^i)^2} = \frac{p}{\widehat{\alpha}_k^T \widehat{\alpha}_k} = \frac{p}{\widehat{X}_k^T \widehat{X}_k}$$

Method(3)

$$\lambda = \frac{\sum_{i=1}^{p} (\lambda_k^i \widehat{\alpha}_k^i)^2}{[\sum_{i=1}^{p} \lambda_k^i (\widehat{\alpha}_k^i)^2]^2}$$

Method(4)

$$\lambda = \frac{1}{p} \sum_{i=1}^{p} \frac{1}{(\widehat{\alpha}_k^i)^2}$$

Method(5)

$$\lambda = \frac{1}{[\prod_{i=1}^{p} (\widehat{\alpha}_k^i)^2]^{\frac{1}{p}}}$$

Method(6)

$$\lambda = \underset{1\le i\le p}{Median} \mid \frac{1}{(\widehat{\alpha}_k^i)^2} \mid$$

Compared with methods to obtain the ridge parameter in ridge regression estimate [9,12,13], the formulas mentioned above do not need to compute the residual sum of squares, so it is more convenient.

However, no one can be comprehensively better than the others, even though there are many methods to obtain the ridge parameter.

6 Simulations and Analysis

Computer simulations are used to verify the validity of the new algorithm described in the previous sections. We consider a discrete linear system described by the state equation (1) and observation equation (2), where state $X_k \in R^{4\times1}$ is estimated. The state transition matrix $\Phi_{k,k-1}$, observation matrix H_k, system noise covariance Q_{k-1} and observation noise covariance R_k are set as follows:

$$\Phi_{k,k-1} = I_4 \ ,$$

$$H_k = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -2 & 4.2 & 0 & 1 \\ 3 & -2 & 4 & 8 \\ -1 & 2 & 0 & 0 \\ 0 & 1 & 1 & 2 \end{bmatrix} ,$$

$$Q_{k-1} = 0.1^2 \times I_4 \ ,$$
$$R_k = 0.5^2 \times I_5 \ ,$$

The initial value is $\widehat{X}_0 = X_0 + 0.001 \times [1\ \ 1\ \ 1\ \ 1]^T$ where $X_0 = [2\ \ 4\ \ 6\ \ 8]^T$ and the initial error covariance matrix is $P_0 = I_4$.

The condition number of the normal matrix $H_k^T R_k^{-1} H_k$ is 5.5×10^4, which means that normal equation is ill-conditioned seriously. Choosing the ridge parameter λ by means of methods (1-6) and compared the new algorithm proposed in this paper with Kalman filter, the results are described as Fig.1-Fig.24.

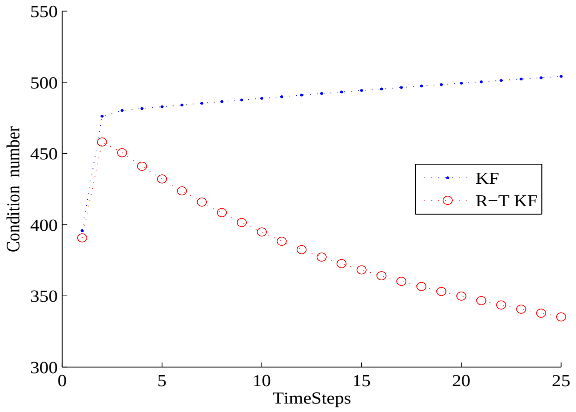

Fig. 1: Comparison of condition number between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(1)

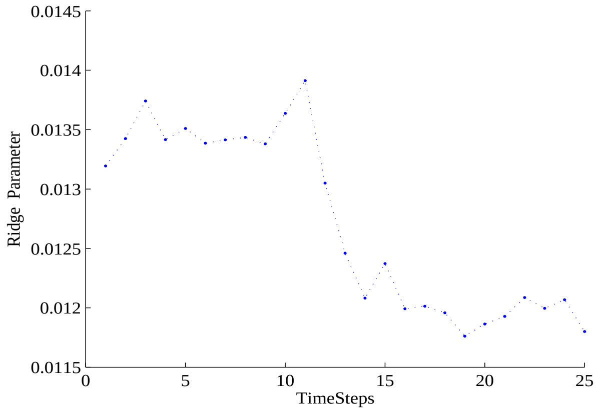

Fig. 2: Choosing the ridge parameter by means of method(1)

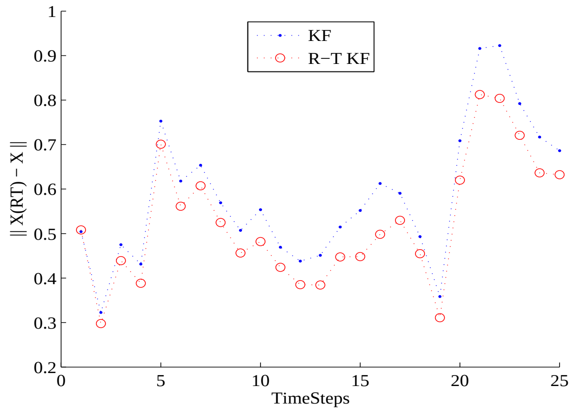

Fig. 3: Comparison of Euclidean distance between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(1)

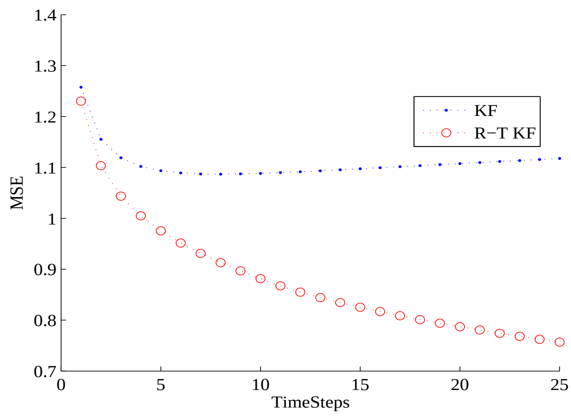

Fig. 4: Comparison of MSE between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(1)

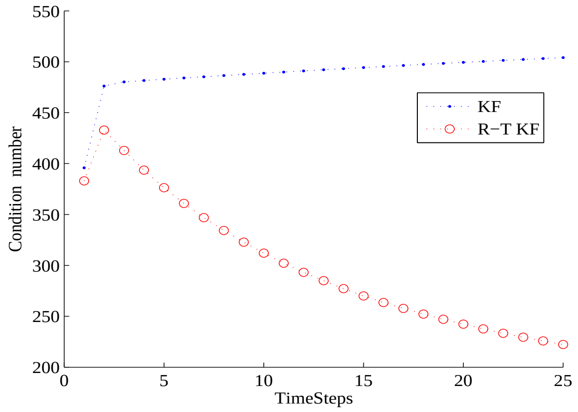

Fig. 5: Comparison of condition number between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(2)

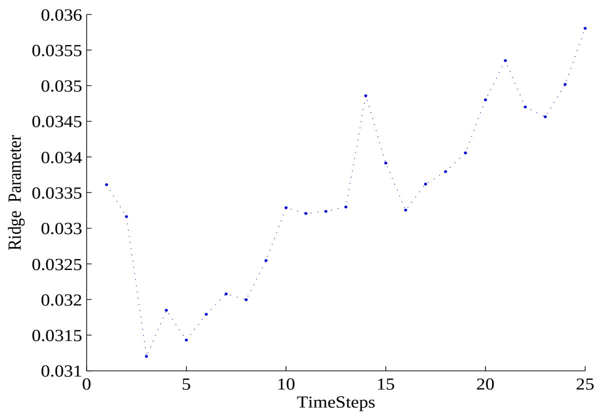

Fig. 6: Choosing the ridge parameter by means of method(2)

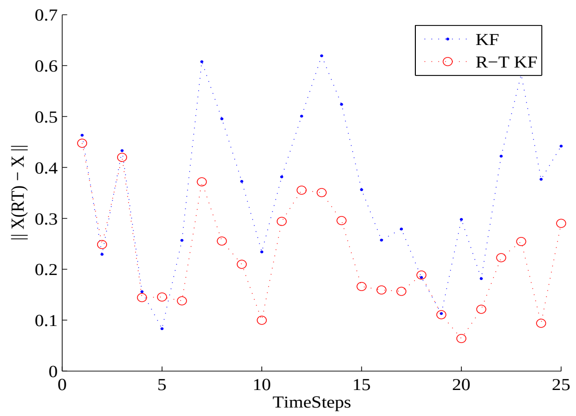

Fig. 7: Comparison of Euclidean distance between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(2)

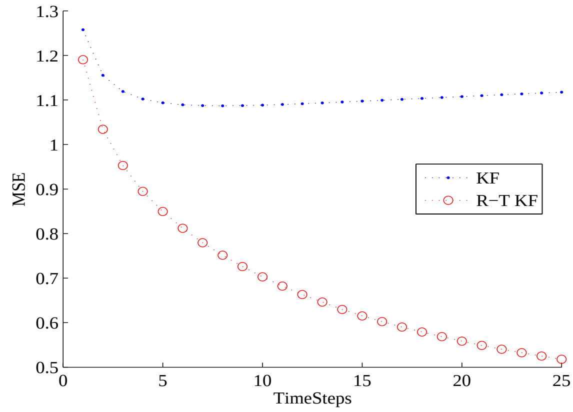

Fig. 8: Comparison of MSE between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(2)

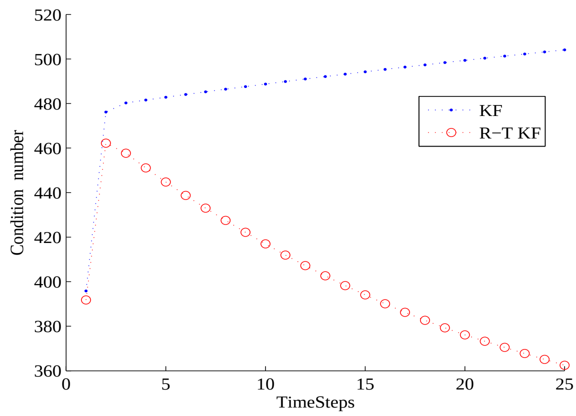

Fig. 9: Comparison of condition number between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(3)

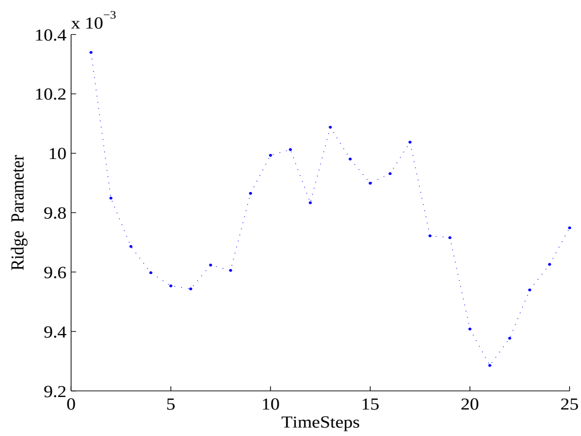

Fig. 10: Choosing the ridge parameter by means of method(3)

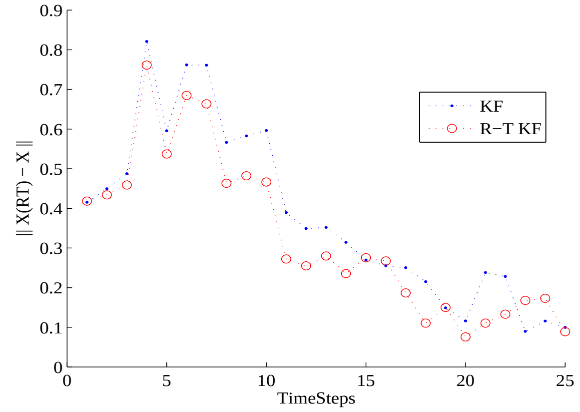

Fig. 11: Comparison of Euclidean distance between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(3)

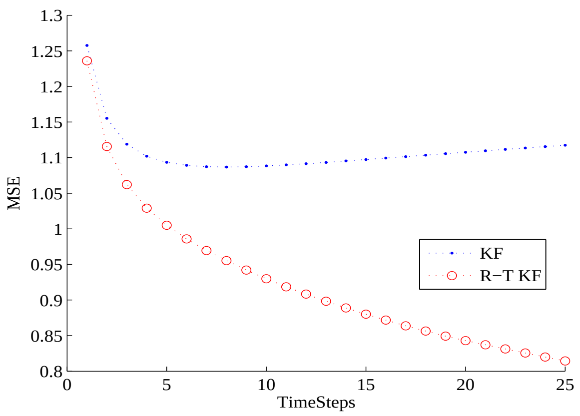

Fig. 12: Comparison of MSE between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(3)

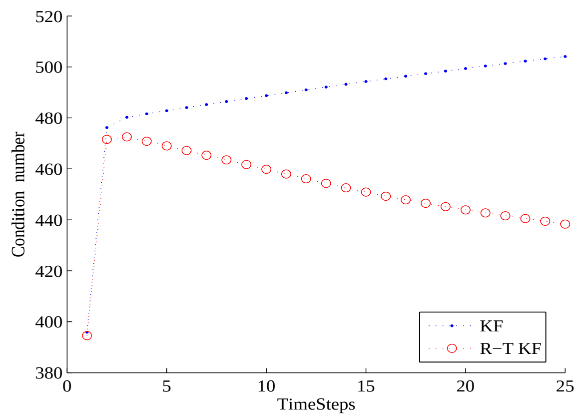

Fig. 13: Comparison of condition number between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(4)

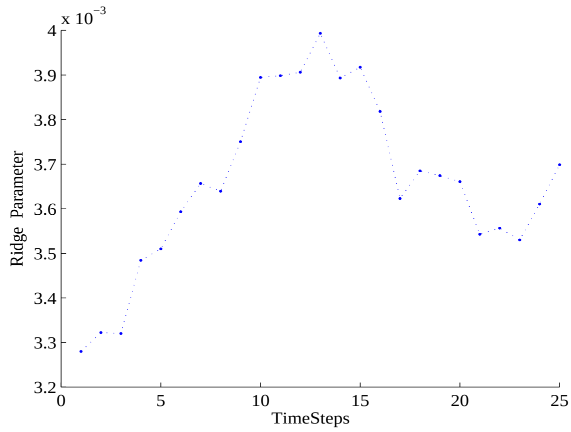

Fig. 14: Choosing the ridge parameter by means of method(4)

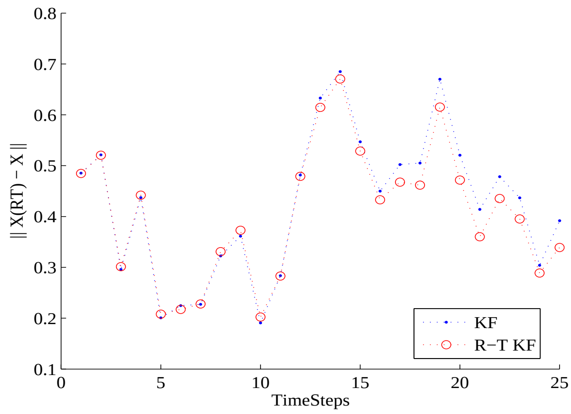

Fig. 15: Comparison of Euclidean distance between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(4)

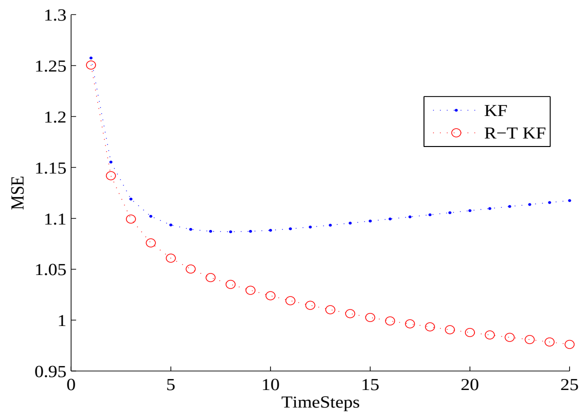

Fig. 16: Comparison of MSE between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(4)

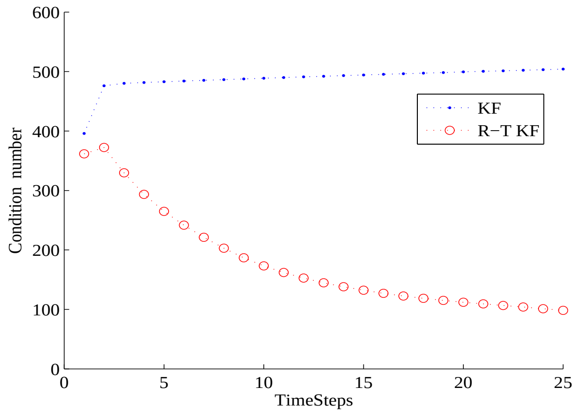

Fig. 17: Comparison of condition number between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(5)

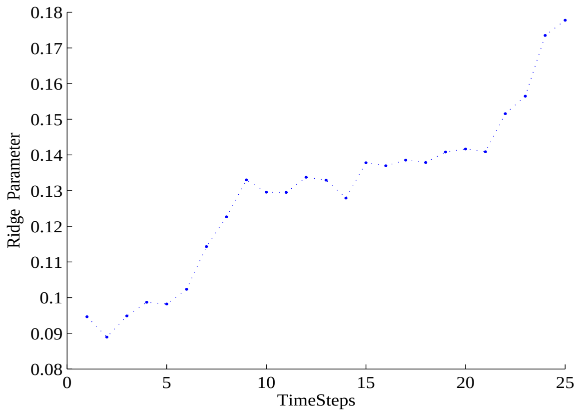

Fig. 18: Choosing the ridge parameter by means of method(5)

It can be concluded from Fig.1-Fig.24 that:

1) $P_{k/k-1}^{-1}$ reduces the condition number of $H_k^T R_k^{-1} H_k + P_{k/k-1}^{-1}$ to some extent, compared with that of $H_k^T R_k^{-1} H_k$. That is to say, information supplied by state equation controls the ill-condition of normal equation. However, normal equation still has stronger ill-condition.

2) New algorithm weakens the ill-condition of normal equation all the time, no matter which way is chosen to obtain the ridge parameter. The new algorithm works better than Kalman filter in the sense of the MSE. The advantage becomes more evident when time lapse.

3) The new algorithm has better controlled on the ill-condition of normal equation and it decreases the MSE well using Method(5) or Method(6) to determine the ridge parameter, compared within Fig.1-Fig.24. That is because ridge parameters determined

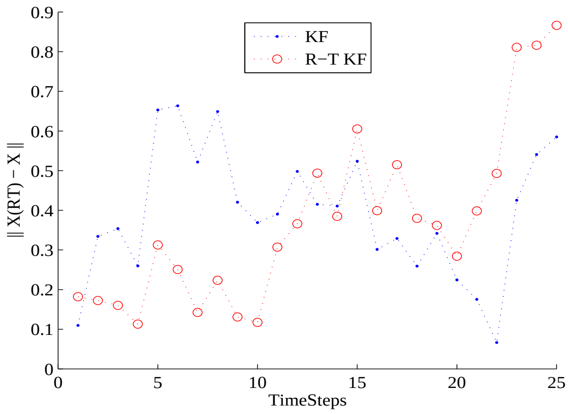

Fig. 19: Comparison of Euclidean distance between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(5)

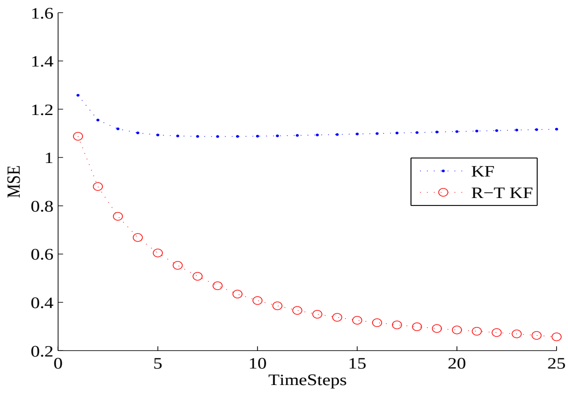

Fig. 20: Comparison of MSE between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(5)

by these two methods is larger. At the same time, larger ridge parameter enlarges the Euclidean distance between the state estimator $\widehat{X}_k(RT)$ and the truth-value X_k. It is not appropriate to seek smaller MSE which leads to the Euclidean distance between the state estimator and the truth-value becoming too large.

7 Conclusion

It has been shown that when the observation matrix H_k has severe multicollinearity, state equation supplies much more information for the estimate of the state parameter so that the ill-condition of normal equation gets controlled. However, normal equation still presents stronger ill-condition, leading to poor accuracy. The role of $P_{k/k-1}^{-1}$ played in controlling the ill-condition of normal equation is also analyzed. We

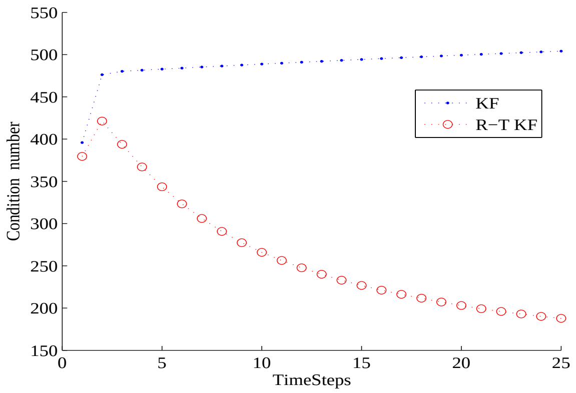

Fig. 21: Comparison of condition number between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(6)

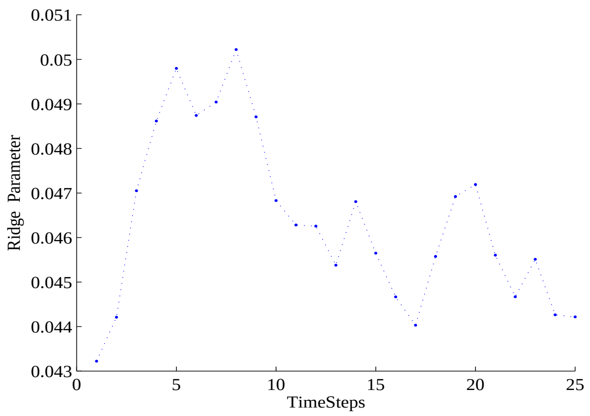

Fig. 22: Choosing the ridge parameter by means of method (6)

proposed a ridge-type Kalman filter under the MSE by analyzing normal equation. Furthermore, some characters of the new algorithm are given. Six methods for determining the ridge parameter of the new algorithm are given which are more efficient than those of ridge regression estimation for it is unnecessary to compute the residual sum of squares.

At last, the new algorithm is proved to be effective to weaken the ill-condition of normal equation by simulations. The new algorithm is superior to Kalman filter in the sense of the MSE. Meanwhile, though the new algorithm weakens the ill-condition of normal equation effectively, it enlarges the Euclidean distance between the state estimator and the truth-value when the ridge parameter is too larger. It needs to be further investigated that the ill-condition of normal equation becomes much severer with time going that appears in the simulations mentioned above.

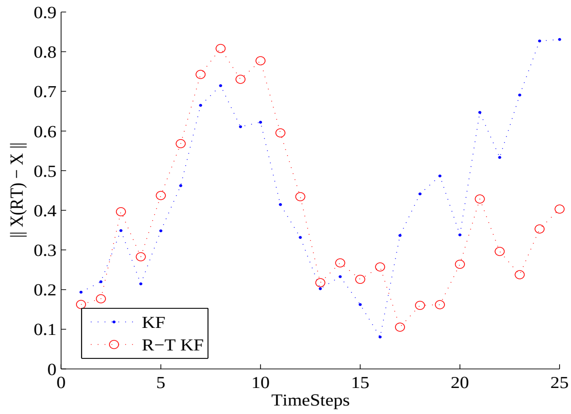

Fig. 23: Comparison of Euclidean distance between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(6)

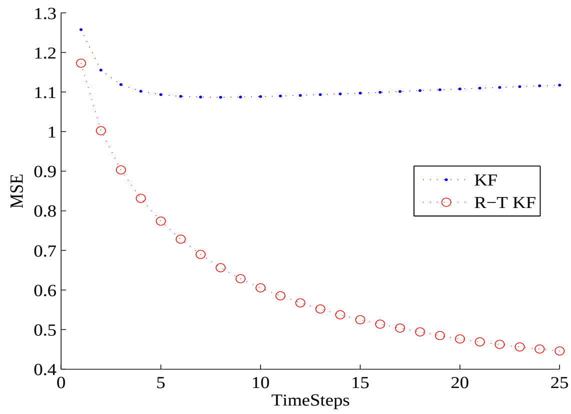

Fig. 24: Comparison of MSE between Kalman filter and ridge-type Kalman filter when the ridge parameter is determined by method(6)

Acknowledgements: This research was supported jointly by National Science Foundation of China (No.40974009, No.41174005, No.41474009) and Funded Project with youth of Annual Meeting of China's satellite navigation.

References:

[1] R. E. Kalman, A New Approach to Linear Filtering and Prediction Problems, *Transaction of the ASME-Journal of Basic Engineering*, Vol.82, No.1, 1960, pp. 35-45.

[2] C. K. Chui and G. R. Chen, *Kalman Filtering: With Real-Time Applications*, Springer, 2009.

[3] Y. Y. Qin and H. Y. Zhang and S. H. Wang, *Theory of Kalman Filter and Integrated Navigation*

Principles, Northwestern Polytechnical University Press, 2012.

[4] D. E. Catlin, *Estimation, Control, and the Discrete Kalman Filter*, Springer, 1989.

[5] C. X. Zhang and J. J Yue, Application of an improved adaptive chaos prediction model in aeroengine performance parameters, *WSEAS Transactions on Mathematics*. Vol.11, 2012, pp. 114-124.

[6] A. A. Keller, l_p-Norm Minimization Method for Solving Nonlinear Systems of Equations, *WSEAS Transactions on Mathematics*. Vol.13, 2014, pp. 654-665.

[7] N. A. Nechval and G. Berzins and M. Purgailis, Improved Estimation of State of Stochastic Systems via Invariant Embedding Technique, *WSEAS Transactions on Mathematics*. Vol.7, 2008, pp. 141-159.

[8] A. N. Tikhonov and V. I. A. Arsenin, *Solutions of ill-posed problems*, John Wiley and Sons, 1977.

[9] A. E. Hoerl and R. W. Kennard, Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*. Vol.12, 1970, pp. 55-67.

[10] S. C. Chow and S. G. Wang, *Advanced Linear Models: Theory and Applications*, CRC Press, 1993.

[11] W. Rudin, *Principles of Mathematical Analysis*, McGRAW-Hill Book Co., 1976.

[12] A. E. Hoerl and R. W. Kennard and K. F. Baldwin, Ridge regression: some simulations, *Communications in Statistics-Theory and Methods*,Vol.4, No.2, 1975, pp. 105-123.

[13] J. F. Lawless and P. Wang, A simulation study of ridge and other regression estimators, *Communications in Statistics-Theory And Methods*, Vol.5, No.4, 1975, pp. 307-323.