International Journal of Control Science and Engineering 2021, 11(1): 1-8
DOI: 10.5923/j.control.20211101.01

Stochastic Dynamic Systems' State Estimation Based on Mean Squared Error Minimizing and Kalman Filtering

Bukhar Kussainov

Institute of Heat Power Engineering and Control Systems, Almaty University of Power Engineering and Telecommunications, Almaty, Republic of Kazakhstan

Abstract In automatic control systems, telecommunications and information systems subjected to impact of random disturbances and measurement inaccuracies, there is the problem of estimating the state vector of observed stochastic system. With the aim to solve the problem the state space system model is described and the problem statement is given. To solve the problem it's used the discrete Kalman filter (KF) presenting itself the recurrent procedure in the form of the set of the difference vector-matrix equations. In the paper the way of deriving the equations of KF on the basis of the procedure of minimization of the mean-squared error of estimation based on a method of the least squares is considered. Using this procedure the discrete analog of the Wiener-Hopf equation as well as Gaussian and Gaussian-Markov estimates of the state vector of linear stochastic system are received satisfying to a minimum of the mean-squared error in the estimate. On the basis of the received estimates and the discrete equation of Wiener-Hopf the equations of the KF is derived, the theorem of the KF with the minimum mean-squared error is formulated, the sequence of using the equations of KF making up the recursive algorithm of KF for computer program realization is explained.

Keywords Stochastic systems, State estimation, Kalman filter

1. Introduction

Modern automatic control systems as well as telecommunications and information systems transmitting and processing signals and subjected to impact of random perturbations and uncertainties of system parameters and to influence of random external disturbances and measurement noises can be considered in the form of the state space models of non-stationary linear dynamical stochastic systems [1]. In stochastic systems, to realize modern control algorithms or to separate a useful signal from its mixture with noise, there is a problem of estimating the entire state vector of the dynamical system based on measured values of system's output signal [2]. The solution of this problem in real time is a filtering problem, for which the classical and most popular solution algorithm is the discrete Kalman filter (KF) [3,4], using both in control theory and in the theory of signal transmission and processing [5]. The KF has the same structure as the considered dynamical system, is the mathematical model and consists of a set of difference vector-matrix equations for calculating estimates of the state of a stochastic system, estimates of the error covariance matrices and the filter gain. The difference between it and the system is that at any given time, the filter gain is optimal relative to the specified statistical properties of disturbances and measurement errors [2-8]. The computational algorithm of the KF is a recursive procedure that is convenient for program realization using programming languages as well as MATLAB [3,9-11] and other computer programs for system modeling.

The article presents a mathematical model of a discrete non-stationary linear stochastic dynamical system, the formulation of the problem of estimating the vector of the system state, the derivation of equations and the formulation of the KF theorem, as well as an algorithm for using the equations of the discrete KF. As it's known, the estimation of the state of a dynamical system (the solution of the filtration problem), as well as the derivation of the KF equations can be carried out using the Bayesian approach, maximum likelihood estimation or the least squares method [12]. Here we follow the already known path and consider the filtration problem as a generalization of the Gaussian least squares method, described in detail in [13]. Based on the least squares method and the procedure for minimizing the mean-squared error of estimation, a discrete analog of the Wiener-Hopf equation is obtained, as well as Gaussian and Gaussian-Markov estimates (and estimates of their error covariance matrices) of the state vector of the observed system, which are linear unbiased and satisfy the minimum value of the mean-squared error of estimation [13]. The discrete Wiener-Hopf equation, Gaussian and Gaussian-Markov estimates with a minimum mean-squared

* Corresponding author:
bukhar41@mail.ru (Bukhar Kussainov)
Received: Aug. 9, 2021; Accepted: Aug. 25, 2021; Published: Sep. 15, 2021
Published online at http://journal.sapub.org/control

error are used later to derive the discrete KF equations, that are a recurrent procedure in which at a discrete time k of the extrapolation (prediction) stage, based on the difference equations of the dynamics of the observed system, the estimate of the state vector is calculated for the next $k+1$ moment of time, and then, at the time $k+1$ of the correction stage, based on new measurement of the system output signal and the changed value of the KF gain, the estimate of the state vector of the system calculated at the time k of extrapolation of the KF procedure is corrected [2-13].

From the first application of the KF in the airspace the KF was a part of the Apollo onboard guidance [3] and to our days the KF has been demonstrating its usefulness in many various applications in different areas of technology and economics [14-16]. However, it is still not easy for people who are not familiar with the estimation theory to understand and implement the vector-matrix equations of the KF. Whereas there is a large number of excellent introductory materials and literature on the KF the purpose of this paper is to remind one simple method for deriving and explain the recursive algorithm for using the equations of the KF.

2. Notational Preliminaries

All vectors and matrices are time-varying quantities are treated at the discrete time instants $k = 0, 1, 2, \dots, N$. By convention, the argument k of vectors (e.g., $x(k)$, $v(k)$, …) and matrices (e.g., $P(k)$, $R(k)$, …) denotes the fact that the values of these variables correspond to the kth step of time. The notation $\hat{x}(k|N)$ designates that the value of the estimation vector $\hat{x}$ at the time instant k conditioned on N time instant measurements. If $N < k$, we are estimating a future value of $\hat{x}(k)$, and we refer to this as a predicted estimate. The case $N = k$ is referred to as a filtered estimate. Prediction and filtering make up the algorithm of KF and can be done in real time [6-8].

The list of notations used through the paper is summarized in the Table 1.

Table 1. List of notations

Symbol	Meaning
Vectors	
x	$n \times 1$ state vector of the system or of the generic linear observation model
$\hat{x}$	$n \times 1$ estimate of a state vector of the system or of the generic linear observation model
$\tilde{x}$	$n \times 1$ error in estimation of a state vector
u	$l \times 1$ input signal of the system
y	$m \times 1$ measurement signal of the system
$\hat{y}$	$m \times 1$ estimate of the measurement signal
$\tilde{y}$	$m \times 1$ error in estimation of the measurement signal
w	$p \times 1$ Gaussian white noise sequence of model uncertainties and disturbances
v	$m \times 1$ Gaussian white noise sequence of measurement inaccuracies
z	a known $m \times 1$ measurement vector of the generic linear observation model
s	an unknown $m \times 1$ vector of measurement errors of the generic linear observation model
g	$n \times 1$ vector to be determined
g^i	ith row of matrix G, a row vector
Matrices	
A	$n \times n$ system matrix
B	$m \times 1$ input matrix
C	$m \times n$ measurement matrix
Q	$p \times p$ covariance matrix of Gaussian white noise sequence of model uncertainties and disturbances
R	$m \times m$ covariance matrix of Gaussian white noise sequence of measurement inaccuracies
P	$n \times n$ covariance matrix of a state vector x
$\tilde{P}$	$n \times n$ error-covariance matrix of an estimation error $\tilde{x}$
D	a known $m \times n$ measurement matrix of the generic linear observation model
S	$m \times m$ covariance matrix of a vector s of the generic linear observation model
G	$n \times m$ matrix to be determined
K	$n \times n$ gain matrix

3. The Basic Model and the Problem of the State Estimation

Consider the basic linear, time-varying (nonstationary), discrete-time state variable model of dynamical systems [5,6] as:

$$x(k+1) = A(k)x(k) + B(k)u(k) + w(k), \tag{1}$$

$$y(k) = C(k)x(k) + v(k),\ k = 0, 1, 2, \dots, N, \tag{2}$$

where $x(k)$ is a $n \times 1$ state vector; $y(k)$ is a $m \times 1$ measurement vector; $u(k)$ is $l \times 1$ input vector; $A(k)$ is $n \times n$ system matrix; $B(k)$ is $n \times l$ input matrix; $C(k)$ is $m \times n$ measurement matrix. $A(k)$, $B(k)$, $C(k)$ matrices and $u(k)$ vector are known.

Additionally, $w(k)$ is $p \times 1$ Gaussian white noise sequence of model uncertainties and disturbances and $v(k)$ is $m \times 1$ Gaussian white noise sequence of measurement inaccuracies, i.e.,

$$E\{w(k)\} = 0,\ E\{w(k)w^T(j)\} = Q(k)\delta_{kj}, \tag{3}$$

$$E\{v(k)\} = 0,\ E\{v(k)v^T(j)\} = R(k)\delta_{kj},\ \forall\, j, k, \tag{4}$$

respectively, where superscript T denotes the matrix transposition.

$Q(k)$, $R(k)$ are $p \times p$ and $m \times m$ covariance matrices, respectively, δ_{kj} is the Dirac delta function, i.e., $\delta_{kj} = 1$ for $k = j$ and $\delta_{kj} = 0$ for $k \neq j$. Supposed that $w(k)$ and $v(k)$ are mutually uncorrelated, i.e.,

$$E\{w(k)v(j)\} = 0,\ \forall\, j, k. \tag{5}$$

The state vector $x(k)$ is zero mean and has a $n \times n$

covariance matrix $P(k)$, i.e.,

$$E\{x(k)\}=0,\ E\{x(k)x^T(j)\}=P(k),\ \forall\, j,k. \quad (6)$$

Initial state vector $x(0)$ and its covariance matrix $P(0)$ are known and $x(0)$ is uncorrelated with $w(k)$ and $v(k)$, i.e.,

$$E\{x(0)w'(k)\}=0,$$
$$E\{x(0)v'(k)\}=0,\ k=0,1,2,\dots,N. \quad (7)$$

The objective is to estimate the $n\times 1$ unknown state vector $x(k)$ at $k=1,2,\dots,N$ from the $m\times 1$ noisy measurement vector $y(k)$, where $k=1,2,\dots,N$.

The estimate $\hat{x}(k)$ of a state vector $x(k)$ must be: 1) linear, 2) unbiased, i.e., $E\{\hat{x}(k)\}=E\{x(k)\}$ and must have 3) a minimum value of the mean of the squared error $E\{\tilde{x}(k)\}^2$, i.e.,

$$E(\tilde{x}(k)\tilde{x}(k)^T)=minimum, \quad (8)$$

where $\tilde{x}(k)=x(k)-\hat{x}(k)$ is the error in the estimate.

Thus, there is the mean-squared estimation problem: given the noisy measurements $y(1)$, $y(2)$, …, $y(k)$, determine a linear unbiased estimator of the entire state vector $x(k)$, $\hat{x}(k)$, such that the conditional mean-squared error in the estimate

$$E\{\tilde{x}(k)\tilde{x}(k)^T|y(1),y(2),\dots,y(k)\} \quad (9)$$

is minimized [5].

This mean-squared estimator, $\hat{x}(k)$, can also be called as the minimum variance estimator, since

$$\sigma_{\tilde{x}}^2=E\{\tilde{x}(k)-E(\tilde{x}(k))\}^2=$$
$$E\{\tilde{x}(k)\}^2=E\{\tilde{x}(k)\tilde{x}(k)^T\}=minimum. \quad (10)$$

Note that here the $\sigma_{\tilde{x}_i}^2$ $(i=\overline{1,n})$ variances are diagonal elements of $n\times n$ error-covariance matrix defined by [6]:

$$\tilde{P}(k)=E(\tilde{x}(k)\tilde{x}(k)^T)=$$
$$E\left\{\left(x(k)-\hat{x}(k)\right)\left(x(k)-\hat{x}(k)\right)^T\right\}. \quad (11)$$

4. The Method of Minimizing of the Mean-Squared Error of Estimation

To obtain the expressions for estimates $\hat{x}$ of unknown vector x from the measurement vector y in conditions of measurement noises v according to the Eq. (2) let's consider the generic linear observation model [5]:

$$z=Dx+s, \quad (12)$$

where x is an $n\times 1$ unknown vector, z is a known $m\times 1$ measurement vector, D is a known $m\times n$ measurement matrix, s is an unknown $m\times 1$ vector of measurement errors.

The unknown quantities x and s are random variables with the following expectations and covariance matrices and they are mutually uncorrelated, i.e.:

$$E(x)=0,\ E(xx^T)=P,\ E(s)=0,$$
$$E(ss^T)=S,\ E(xs^T)=0. \quad (13)$$

The assumption of a linearity leads to the following expression for the estimate of x, $\hat{x}$:

$$\hat{x}=g+Gz, \quad (14)$$

where $n\times 1$ vector g and $n\times m$ matrix G must determine, however, the request of the unbiased estimation means that:

$$E\{\hat{x}\}=g+GE\{z\}=E\{x\},$$

hence $g=0$, since $E\{x\}=E\{z\}=0$.

Thus the Eq. (14) for $\hat{x}$ becomes as:

$$\hat{x}=Gz. \quad (15)$$

The matrix G will be determined from the condition that the variance of estimation error $\tilde{x}=x-\hat{x}$ is minimum. According to Eq. (15) every component $\hat{x}_i$ of $\hat{x}$ is depended on vector z via an ith row of matrix G which is denoted as g^i. Thus

$$\hat{x}_i=g^i z, \quad (16)$$

where g^i is the row vector.

Mentioned request about a minimum variance of estimation error signifies that

$$E\{\tilde{x}_i-E(\tilde{x}_i)\}^2=E\{\tilde{x}_i\}^2\to \min_{g^i}\tilde{x}_i,\ i=1,2,\dots,n. \quad (17)$$

Hence

$$E\{\tilde{x}_i\}^2=E\{x_i-\hat{x}_i\}^2=E\{x_i-g^i z\}^2=$$
$$E\{x_i^2\}-2E\{x_i z^T\}(g^i)^T+g^i E\{zz^T\}(g^i)^T. \quad (18)$$

Thus, the variance of ith error is the sum in which the first term don't depend on g^i, the second and third terms are linear and quadratic forms of g^i ($(g^i)^T$ is the column vector). A necessary condition of minimum of Eq. (18) is that all its partial derivatives with respect to $g^i(k)$ must be equal to zero. In other words, taking the gradient of $E\{\tilde{x}_i\}^2$ with respect to $(g^i)^T$ must be equal to zero, i.e.,

$$\frac{\partial}{\partial (g^i)^T}E\{\tilde{x}_i\}^2=0. \quad (19)$$

Applying the rule of the gradient calculation to the right hand side of Eq. (18) yields [13]:

$$E\{x_i z^T\}-g^i E\{zz^T\}=0,\ i=1,2,\dots,n. \quad (20)$$

This Eq. (20) is regarded as the Wiener-Hopf equation in the discrete form, let's rewrite it in the compact vector-matrix form:

$$E\{xz^T\}-GE\{zz^T\}=0. \quad (21)$$

The necessary condition for the fairness of the system of linear algebraic equations (21) with unknown weighting matrix G is that the variance $E\{\tilde{x}_i\}^2$ of each i th estimation error must be extremum. The sufficient condition of it is the positive definiteness of the matrix formed by the second derivatives of the function $E\{\tilde{x}_i\}^2$ with respect to $g^i(k)$. In other words, the Hessian matrix with respect to $(g^i)^T$ must be positive definite for all i, i.e.:

$$\frac{\partial^2}{\partial (g^i)'^2}E\{\tilde{x}_i\}^2 \text{ is positive definite.} \quad (22)$$

Recalculation of partial derivatives for the left hand side of the Eq. (20) yields the required Hessian matrix. Hence the condition above turns into the following [13]:

$$E\{zz^T\} \text{ is positive definite.} \tag{23}$$

This condition is sufficient that the extremum values of variance obtained by using Eq. (21) will be really minimum. Thus the requirement of Eq. (23) is necessary and sufficient condition that the Eq. (21) will have only one solution for G.

To obtain the matrix G we must find the covariance matrices in Eq. (21), so taking into account the Eqs. (12), (13) we have:

$$E\{xz^T\} = E\{x(x^T D^T + s^T)\} = PD^T, \tag{24}$$

$$E\{zz^T\} = E\{(Dx + s)(x^T D^T + s^T)\}$$
$$= DPD^T + S. \tag{25}$$

Now rewrite the Eq. (21) in the following form:

$$G(DPD^T + S) = PD^T. \tag{26}$$

Here in parentheses is $m \times m$ matrix. If m measurements are less than n unknowns, then from the Eq. (26) we can find the matrix G:

$$G = PD^T(DPD^T + S)^{-1}. \tag{27}$$

Substitute the Eq. (27) into the Eq. (15) we receive the first form of the linear unbiased estimate with a minimum value of its mean-squared error:

$$\hat{x} = PD^T(DPD^T + S)^{-1}z. \tag{28}$$

In order to obtain the covariance matrix of estimation error consider following equations:

$$E\{\tilde{x}\tilde{x}^T\} = E\{\tilde{x}(x - \hat{x})^T\} = E\{\tilde{x}x^T\} = \tag{29}$$

$$E\{(x - Gz)x^T\} = E\{xx^T\} - GE\{zx^T\} \tag{30}$$

Taking into account the Eqs. (15), (24) we have from the Eq. (30):

$$E\{\tilde{x}\tilde{x}^T\} = P - GDP. \tag{31}$$

Substitute here the Eq. (27) to the Eq. (31) we receive the covariance matrix of the first form estimate (28):

$$E\{\tilde{x}\tilde{x}^T\} = P - PD^T(DPD^T + S)^{-1}DP. \tag{32}$$

For the case, if m measurements are more than n unknowns, the Eq. (26) can be transformed to the following form [13]:

$$(D^T S^{-1} D + P^{-1})G = D^T S^{-1}. \tag{33}$$

Here in parentheses is $n \times n$ matrix. Using matrix G obtained from the Eq. (33) the second form of the linear unbiased estimate with a minimum value of mean-squared estimation error is:

$$\hat{x} = (D^T S^{-1} D + P^{-1})^{-1} D^T S^{-1} z. \tag{34}$$

According to the Eq. (31) the covariance matrix of the second form estimate (34) is given by:

$$E\{\tilde{x}\tilde{x}^T\} = [I - GD]P$$
$$= [I - (D^T S^{-1} D + P^{-1})^{-1} D^T S^{-1} D]P$$
$$= (D^T S^{-1} D + P^{-1})^{-1}[(D^T S^{-1} D + P^{-1}) - D^T S^{-1} D]P,$$
$$E\{\tilde{x}\tilde{x}^T\} = (D^T S^{-1} D + P^{-1})^{-1}. \tag{35}$$

In the estimate (34), if $P \to \infty$, that is $P^{-1} = 0$, and rank of matrix D is n, then we have the Gaussian-Markov estimate [13]:

$$\hat{x} = (D^T S^{-1} D)^{-1} D^T S^{-1} z. \tag{36}$$

According to the Eq. (35) the covariance matrix of the Gaussian-Markov estimate is given by:

$$E\{\tilde{x}\tilde{x}^T\} = (D^T S^{-1} D)^{-1}. \tag{37}$$

5. Deriving the Equations of the Kalman Filter

The Kalman filter operates in a predict-correct manner [5].

5.1. Prediction

At the initial observation moment according to the Eqs. (2), (6) the following measuring is obtained:

$$y(0) = C(0)x(0) + v(0),$$
$$E\{x(0)x^T(0)\} = P(0). \tag{38}$$

Compare these Eqs. (38) with the Eqs. (12), (13) and substitute the corresponding quantities to the Eq. (28) for a linear unbiased estimate with a minimum of a mean-squared error, we have:

$$\hat{x}(0) = K(0)y(0), \tag{39}$$

where

$$K(0) = P(0)C^T(0)[C(0)P(0)C^T(0) + R(0)]^{-1}. \tag{40}$$

The error-covariance matrix according to the Eq. (32) is given by:

$$\tilde{P}(0) = P(0) - K(0)C(0)P(0). \tag{41}$$

The solution for all subsequent moments of time is obtained by moving from k to $k+1$. With this aim consider the discrete Wiener-Hopf equation (21) which is the necessary and sufficient condition that estimate will have a minimum mean-squared error. Rewrite Eq. (21) in more short form:

$$E\{(x - Gz)z^T\} = E\{(x - \hat{x})z^T\} = E\{\tilde{x}z^T\} = 0. \tag{42}$$

Suppose observations $y(0),\ y(1),\ \ldots,\ y(k)$ are already done and the estimate $\hat{x}(k)$ with the minimum of the mean-squared error is obtained. The latter means that the Eq. (42) is satisfied, i.e.:

$$E\{[x(k) - \hat{x}(k)][y^T(0), y^T(1), \ldots, y^T(k)]\} = 0. \tag{43}$$

We point out that this Eq. (43) is already satisfied.

Suppose the quantities $\hat{x}$ and $\tilde{P}$ at the time k, $\hat{x}(k)$ and $\tilde{P}(k)$, are known. According to the Eq. (1) let's find the predicted value of $\hat{x}$, $\hat{x}(k+1)$, herewith uncertainties and disturbances $w(k)$ (with zero expectations) are not taken into account:

$$\hat{x}(k+1|k) = A(k)\hat{x}(k) + B(k)u(k). \tag{I}$$

This is the first equation of the Kalman filtering procedure.

Let's show that this prediction $\hat{x}(k+1|k)$ is optimal. According to Eq. (42) we have:

$$E\{x(k+1)-\hat{x}(k+1|k)\}\{y^T(0),y^T(1),\dots,y^T(k)\}=0 \quad (44)$$

Here in Eq. (44) replace $x(k+1)$ with $A(k)x(k)+B(k)u(k)$ from the Eq. (1), herewith disturbances $w(k)$ cannot be taken into account because they are not correlated with $y(0)$, $y(1)$, …, $y(k)$. We have the following equation:

$$E\{A(k)x(k)+B(k)u(k)-\hat{x}(k+1|k)\} \times\{y^T(k_0),y^T(k_0+1),\dots,y^T(k)\}=0. \quad (45)$$

From the Eq. (45) we can see that the optimal prediction is corresponded to the Eq. (I).

Prediction error is equal to:

$$\begin{aligned} &x(k+1)-\hat{x}(k+1|k)\\ &=A(k)x(k)+B(k)u(k)+w(k)\\ &\quad -A(k)\hat{x}(k|k)-B(k)u(k)\\ &=A(k)\tilde{x}(k|k)+w(k). \end{aligned} \quad (46)$$

The error-covariance matrix of prediction is given by:

$$\begin{aligned} &\tilde{P}(k+1|k)\\ &=E[x(k+1)-\hat{x}(k+1|k)][x(k+1)-\hat{x}(k+1|k)]^T\\ &=E[A(k)\tilde{x}(k|k)+w(k)][\tilde{x}^T(k|k)A^T(k)+w^T(k)], \end{aligned}$$

however, the noises $w(k)$ are not correlated with the estimation errors $\tilde{x}(k)$, so

$$\tilde{P}(k+1|k)=A(k)\tilde{P}(k)A^T(k)+Q(k). \quad (II)$$

This is the second equation of the Kalman filtering procedure. By this the prediction is done.

5.2. Correction

The estimate of the predicted state vector x at the $k+1$ instant of time, $\hat{x}(k+1|k)$, obtained with the available $y(k)$ measurement (Eq. (I)), after the next $y(k+1)$ measurement must be corrected to the value $\hat{x}(k+1|k+1)$.

In order to find $\hat{x}(k+1|k+1)$ let's replace $\hat{x}(k|k)$ in Eq. (I) with $x(k)-\tilde{x}(k|k)$ and substitute instead of $A(k)x(k)$ the corresponding expression from the Eq. (1):

$$\hat{x}(k+1|k)=A(k)\big(x(k)-\tilde{x}(k|k)\big)+B(k)u(k),$$
$$\hat{x}(k+1|k)=x(k+1)-A(k)\tilde{x}(k|k)-w(k). \quad (47)$$

Taking into account the Eq. (2) we can write:

$$\begin{bmatrix}\hat{x}(k+1|k)\\ y(k+1)\end{bmatrix}=\begin{bmatrix}1\\ \mathrm{C}(k+1)\end{bmatrix}x(k+1)+ \begin{bmatrix}-A(k)\tilde{x}(k|k)-w(k)\\ v(k+1)\end{bmatrix} \quad (48)$$

or

$$z=Dx+s. \quad (49)$$

Here the quantities z, D, s are defined through comparing the Eqs. (48), (49). The covariance matrix of the upper part of vector s was earlier denoted as $\tilde{P}(k+1|k)$, the covariance matrix of the lower part $v(k+1)$ is equal to $R(k+1)$. Covariance between $v(k+1)$ and $\tilde{x}(k|k)$ and between $v(k+1)$ and $w(k)$ are equal to zero. So, we have:

$$E\{ss^T\}=S=\begin{bmatrix}\tilde{P}(k+1|k) & 0\\ 0 & R(k+1)\end{bmatrix}. \quad (50)$$

Since the measurement vector z contains the component $y(k+1)$ we can calculate Gaussian-Markov estimate $\hat{x}(k+1)$ at the $(k+1)$ instant of time according to the Eqs. (36), (37), i.e.:

$$\hat{x}(k+1|k+1)=\tilde{P}(k+1|k+1)D^TS^{-1}z, \quad (51)$$

where

$$\tilde{P}(k+1|k+1)=(D^TS^{-1}D)^{-1}. \quad (52)$$

Here

$$\begin{aligned} D^TS^{-1}&=[I \quad C^T(k+1)]\begin{bmatrix}\tilde{P}^{-1}(k+1|k) & 0\\ 0 & R^{-1}(k+1)\end{bmatrix}\\ &=[\tilde{P}^{-1}(k+1|k) \quad C^T(k+1)R^{-1}(k+1)]. \end{aligned} \quad (53)$$

$$D^TS^{-1}D=[\tilde{P}^{-1}(k+1|k) \quad C^T(k+1)R^{-1}(k+1)] \times\begin{bmatrix}1\\ \mathrm{C}(k+1)\end{bmatrix} \quad (54)$$

$$D^TS^{-1}z=[\tilde{P}^{-1}(k+1|k) \quad C^T(k+1)R^{-1}(k+1)] \times\begin{bmatrix}\hat{x}(k+1|k)\\ y(k+1)\end{bmatrix}. \quad (55)$$

Substitute the expression for $D^TS^{-1}z$ from the Eq. (55) into the Eq. (51) of Gaussian-Markov estimate:

$$\hat{x}(k+1|k+1)=\tilde{P}(k+1|k+1)\big(\tilde{P}^{-1}(k+1|k)\times \hat{x}(k+1|k)+C^T(k+1)R^{-1}(k+1)y(k+1)\big). \quad (56)$$

In Eq. (56) the factor before $y(k+1)$ is the gain matrix K:

$$K(k+1)=\tilde{P}(k+1|k+1)C^T(k+1)R^{-1}(k+1). \quad (57)$$

Using the Eq. (54) for $D^TS^{-1}D$ and taking into account the Eq. (52) we receive:

$$\tilde{P}^{-1}(k+1|k+1)=\tilde{P}^{-1}(k+1|k)+ C^T(k+1)R^{-1}(k+1)C(k+1). \quad (58)$$

Let's multiply by $\tilde{P}(k+1|k+1)$ on the left the Eq. (58) and taking into account the Eq. (57) we receive:

$$I=\tilde{P}(k+1|k+1)\tilde{P}^{-1}(k+1|k)+K(k+1)C(k+1). \quad (59)$$

Let's find the expression for $\tilde{P}(k+1|k+1)\tilde{P}^{-1}(k+1|k)$ from the Eq. (59) and substitute it into the Eq. (56):

$$\hat{x}(k+1|k+1)=\hat{x}(k+1|k)+K(k+1) \times\big(y(k+1)-\mathrm{C}(k+1)\hat{x}(k+1|k)\big). \quad (III)$$

This is the third equation of the Kalman filter. It corresponds to the observer model equation of the observed object (system) with feedback equaled a difference of weighted output signals [17,18].

Let's multiply by $\tilde{P}(k+1|k)$ on the right the Eq. (59):

$$\tilde{P}(k+1|k)=\tilde{P}(k+1|k+1)+ K(k+1)C(k+1)\tilde{P}(k+1|k),$$

and from this equation we can receive the error-covariance matrix:

$$\tilde{P}(k+1|k+1) = \tilde{P}(k+1|k)$$
$$-K(k+1)C(k+1)\tilde{P}(k+1|k). \qquad \text{(IV)}$$

This is the forth equation of the Kalman filter. Let's multiply this equation by C^T on the right and ascribe to the first tirm in the right hand side the factor $R^{-1}R$. From the obtained equation taking into account the Eq. (57) we can find the gain matrix K:

$$K(k+1) = \tilde{P}(k+1|k)C^T(k+1) \times$$
$$\left(C(k+1)\tilde{P}(k+1|k)C^T(k+1) + R(k+1)\right)^{-1}. \qquad \text{(V)}$$

This is the fifth equation of the Kalman filter.

6. The Theorem of the Kalman Filter

The Eqs. (39)-(41) and (I)-(V) make up together the Kalman filter which is usually formulated as the theorem. Let's present the formulation of the theorem of Kalman filter satisfying the minimum of mean-squared error of estimation.

Theorem (The Kalman Filter). Let given a discrete stochastic system defined by the Eqs. (1)-(7) and considered at $k = 0, 1, 2, \dots, N$ instants of time. The linear unbiased estimate with the minimum mean-squared error in the estimation of the state vector of this system at any time instant $k > 0$ is obtained by the recursive equations (I)-(V) the initial state of which at $k = 0$ is determined by the equations (39)-(41).

In addition to the proof of the theorem considered above to check the correctness of the Eqs. (IV), (V). With this aim let's make the expression for $\hat{x}(k+1|k+1)$ according to the Eq. (42):

$$E\{x(k+1) - \hat{x}(k+1|k+1)\} \times$$
$$\{y^T(0), y^T(1), \dots, y^T(k) | y^T(k+1)\} = 0. \qquad (60)$$

In this Eq. (60) the estimation error according to the Eq. (III) and taking into account the Eq. (2) is equal to:

$$x(k+1) - \hat{x}(k+1|k+1)$$
$$= [I - K(k+1)C(k+1)]$$
$$\times [x(k+1) - \hat{x}(k+1|k)]$$
$$-K(k+1)v(k+1). \qquad (61)$$

To check the correctness of the Eq. (IV) multiply on the right the both parts of the Eq. (61) by $x^T(k+1)$ and calculate expectation. Still, using the Eq. (42) will allow us to obtain the Eq. (IV).

To check the correctness of the deriving the Eq. (V), consider the first part of the mathematical expectation in Eq. (60), containing values from $y(0)$ to $y(k)$ and located to the left of the vertical line for $y(k+1)$. The new measurement error $v(k+1)$ is uncorrelated with the old observations from $y(k_0)$ to $y(k)$. The product of two expressions in square brackets in (61), correlated with the set of observations from $y(0)$ to $y(k)$, means zero mathematical expectation according to equation (44). This means that expression (III) satisfies the part of the requirement (42) that is to the left of the vertical line. The remaining part of the requirement (60) allows us to determine the undefined gain matrix $K(k+1)$. On the basis of (61), the following equality must be valid:

$$E\{[I - K(k+1)C(k+1)]$$
$$\times [x(k+1) - \hat{x}(k+1|k)] - K(k+1)v(k+1)\}$$
$$\times \{x^T(k+1)C^T(k+1) + v^T(k+1)\} = 0. \qquad (62)$$

The quantities $x(k+1)$ and $\hat{x}(k+1|k)$ are not correlated with $v(k+1)$. The rest of the mathematical expectations can be represented in a simpler form. To do this, we use Eq. (29) with respect to $\hat{x}(k+1|k)$, taking into account that the covariance between $x(k+1) - \hat{x}(k+1|k)$ and $x(k+1)$ can be replaced by $\tilde{P}(k+1|k)$. As a result, we have:

$$\left(I - K(k+1)C(k+1)\right)\tilde{P}(k+1|k)$$
$$\times C^T(k+1) - K(k+1)R(k+1) = 0. \qquad (63)$$

Solving this Eq. (63) with respect to $K(k+1)$ we'll receive the Eq. (V).

7. The Algorithm of Using the Equations of the Kalman Filter

The Kalman filter is a recursive procedure that is convenient for program realization on computers. The algorithm of using the Eqs. (39)-(41) and (I)-(V) of KF is the following as:

1) At the initial state $k = 0$ the initial estimate of the state vector $\hat{x}(0|0)$ and the initial error-covariance matrix $\tilde{P}(0)$ are built according to the Eqs. (39)-(41):

$$\hat{x}(0|0) = \hat{x}(0) = K(0)y(0),$$

where

$$K(0) = P(0)C^T(0)[C(0)P(0)C^T(0) + R(0)]^{-1}$$

and

$$\tilde{P}(0) = P(0) - K(0)C(0)P(0).$$

Prediction:

2) The estimate and its error-covariance matrix are extrapolated to the next $(k+1)$ observation instant of time according to the Eqs. (I), (II):

$$\hat{x}(k+1|k) = A(k)\hat{x}(k|k) + B(k)u(k),$$
$$\tilde{P}(k+1|k) = A(k)\tilde{P}(k)A^T(k) + Q(k).$$

Correction:

3) The optimal gain matrix $K(k+1)$ is calculated according to the Eq. (V) and extrapolated (predicted) estimate $\hat{x}(k+1|k)$ is improved to the value $\hat{x}(k+1|k+1)$ according to the Eq. (III) using the new measurement $y(k+1)$:

$$K(k+1) = \tilde{P}(k+1|k)C^T(k+1) \times$$
$$\left(C(k+1)\tilde{P}(k+1|k)C^T(k+1) + R(k+1)\right)^{-1},$$
$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)\tilde{y}(k+1|k),$$

where

$$\tilde{y}(k+1|k) = \left(y(k+1) - \hat{y}(k+1|k)\right)$$

is called the innovation process,

$$\hat{y}(k+1|k) = \mathrm{C}(k+1)\hat{x}(k+1|k)$$

is called the predicted value of the new measurement.

4) The error-covariance matrix $\tilde{P}(k+1|k+1)$ of the new modified estimate $\hat{x}(k+1|k+1)$ is calculated according to the Eq. (IV):

$$\tilde{P}(k+1|k+1) = \tilde{P}(k+1|k)$$
$$-K(k+1)C(k+1)\tilde{P}(k+1|k).$$

5) If the next k is $k \le N$ then the current time instant $(k+1)$ should be considered as k. For the estimate of the state calculated at the step 3 and now considered as $\hat{x}(k)$, for the error-covariance matrix calculated at the step 4 and now considered as $\tilde{P}(k)$ should be carried out the steps 2, 3 and 4 of the algorithm. If $k > N$ then the procedure is ended.

Therefore, the best estimate of $x(k+1)$, $\hat{x}(k+1)$, using all observations up to and including $k+1$, is obtained by a predictor step, $\hat{x}(k+1|k)$, and a corrector step, $K(k+1)\tilde{y}(k+1|k)$. The predictor step uses information from the state equation (1). The corrector step uses the new measurement available at $k+1$. The correction is the error (difference) between new measurement, $y(k+1)$, and its best predicted value, $\hat{y}(k+1|k)$, multiplied by weighting (or gain) factor $K(k+1)$. The factor K determines how much we will alter (change) the best estimate $\hat{x}$ based on the new observation, i.e., 1) if the elements of $K(k+1)$ are small, we have considerable confidence in our model, and 2) if they are large, we have considerable confidence in our observation measurements. Thus, the KF is a dynamical feedback system, its gain matrix and predicted- and filtering-error covariance matrices comprise a matrix feedback system operating within the KF [5,6].

8. Conclusions

The discrete Kalman filter, developed by R. Kalman back in 1960 [19], is currently a classic result of the theory of control systems and the theory of signal processing, as well as the most popular filtering algorithm using in automatic control systems, telecommunications and information systems subjected to random disturbances and measurement inaccuracies. The Kalman filter is a recursive procedure consisting of difference vector-matrix equations for calculating estimates of the state of a stochastic system, the estimates of the error covariance matrices and the filter gain. A common approach to the derivation of the KF equations is the Bayesian approach [11]. The paper describes the simplest way to obtain the KF equations, based on the use of the procedure for minimizing the mean-squared error of estimation, which is a further generalization of the least squares method [12]. As a result of this procedure, we obtained: a discrete analog of the Wiener-Hopf equation, as well as Gaussian and Gaussian-Markov estimates (and their error covariance matrices), which are linear and unbiased and satisfy the minimum value of the mean-squared error of the estimation. Based on the discrete Wiener-Hopf equation, Gaussian and Gaussian-Markov estimates, the KF equations are obtained using simple algebraic transformations and reasoning. The KF theorem is formulated, which satisfies the minimum of mean-squared error of estimation, and the algorithm for using the KF equations, which is convenient for program realization, is also explained.

REFERENCES

[1] P. R. Kumar, Pravin Varaiya, Stochastic systems: Estimation, Identification and Adaptive Control, New Jersey, Prentice Hall, 1986.

[2] K. J. Åström, R. M. Murray, Feedback systems: An Introduction for Scientist and Engineers, New Jersey, Oxford, Princeton University Press, 2008.

[3] Mohinder S. Grewal, Angus P. Andrews, Kalman Filtering: Theory and Praxis Using MATLAB, 4th ed., New Jersey, John Wiley & Sons, 2015.

[4] Greg Welch, Gary Bishop, An Introduction to the Kalman Filter, UNC-Chapel Hill, TR 95-041, July 24, 2006.

[5] Jerry, M., Mendel, Lessons in Estimation Theory for Signal Processing, Communications, and Control, New Jersey, Prentice Hall, 1995.

[6] George, M., Siouris, An Engineering Approach to Optimal Control and Estimation Theory, New York, John Wiley & Sons, 1996.

[7] Lennart, Ljung, System Identification. Theory for the User, 2nd ed., New Jersey, Prentice Hall, 1999.

[8] Karel, J., Keesman, System Identification. An Introduction, London, Springer-Verlag, 2011.

[9] Robert G. Brown, Patrick Y.C. Hwang, Introduction to random signals and applied Kalman filtering: with MATLAB exercises, 4th ed., New Jersey, John Wiley & Sons, 2012.

[10] Lennart, Ljung, System Identification Toolbox™. User's Guide, 3 Apple Hill Drive Natick, The Mathworks, Inc., 2015.

[11] Armando Barreto, Malek Adjouadi, Intuitive Understanding of Kalman Filtering with MATLAB, CRC Press, 2021.

[12] P. Eykhoff, Bases of identification of control systems: parameter and state estimation, Moscow, Mir, 1975. (In Russ.)

[13] K. Brammer, G. Ziffling, Kalman-Bucy Filter: Deterministic Observation and Stochastic Filtering (Translation from German), Moscow, Nauka, 1984. (In Russ.)

[14] Vedran Cordić (editor), Kalman filter, Vukovar, Croatia, Intech, 2010.

[15] Bruce P. Gibbs, Advanced Kalman filtering, least squares and modeling, New Jersey, John Wiley & Sons, 2011.

[16] Charles K. Chui, Guangrong Chen, Kalman filtering with real-time applications, 5th ed., Springer International Publishing AG, 2017.

[17] R.C. Dorf, R.H. Bishop, Modern control systems, 12th ed., New Jersey, Prentice Hall, 2011.

[18] K. Ogata, Modern control engineering, 3rd ed., New Jersey, Prentice Hall, 1997.

[19] R.E. Kalman, A new approach to linear filtering and prediction problems, ASME Transactions of the ASME – Journal of Basic Eng. 82(1), pp. 35-45, March, 1960.