KALMAN FILTERS

Mohinder Singh Grewal
California State University, Fullerton, USA

Keywords: Bierman, Thornton, Cholesky decomposition, continuous time optimal estimator, covariance, estimators, filters, GPS, implementation methods, INS, Kalman-Bucy filter, linear discrete optimal estimator, linearized and extended Kalman filter, nonlinear estimation, predictors, Riccati equation, WAAS, Wiener filter, white noise

Contents

Summary

The Kalman filter provides a means for estimating the parameters and states from indirect (and noisy) measurements in order to control complex, dynamic systems, and to predict the outcome of dynamic systems that people are not likely to control. Optimal Kalman filters are developed by assuming that plant and measurement noises are white noise processes, or sequences. The Kalman filter can be characterized as an algorithm for computing the conditional mean and covariance of the probability distribution of the state of a linear stochastic system with uncorrelated Gaussian process and measurement noise.

This article derives equations for the discrete-time estimator and the continuous-time optimal estimator (Kalman-Bucy filter) and its implementation. Kalman filters are developed for nonlinear discrete and continuous systems as well, including model

extensions, and approximation methods used for applying the methodology of Kalman filtering to "slightly nonlinear" problems.

The theoretical performance of Kalman filters has been shown to be characterized by the covariance matrix of estimation uncertainty that is computed as the solution of a matrix Riccati differential and difference equation. Current work on the Kalman filter primarily focuses on development of robust and numerically stable implementation methods. The article discusses the modified Cholesky (UD) decomposition algorithms implementation and other UD filtering modifications, including use of the Gram-Schmidt orthogonalizations. References are given to explore additional alternative Kalman filter implementations.

The Kalman filter is an observer, a parameter identifier in modeling, a predictor, a filter and a smoother in a variety of applications. It is an integral part of almost every onboard trajectory estimation and control system. Kalman filters are used in bioengineering, traffic systems, photogrammetry, global positioning systems, inertial navigation, guidance, and myriad process controls. The Kalman filter has become integral to 21st century technology.

1. Introduction

What is a Kalman filter? Theoretically, a Kalman filter is an estimator for what is called the linear-quadratic-Gaussian problem (LQG), which is the problem of estimating the instantaneous "state" of a linear dynamic system perturbed by Gaussian white noise, by using measurements linearly related to the state, but corrupted by Gaussian white noise. For any quadratic function of estimation error, an estimator can be designed to minimize it. R. E. Kalman introduced the "filter" in a 1960 paper, "A new approach to linear filtering and prediction problems."

Practically, the Kalman filter is certainly one of the greater discoveries in the history of statistical estimation theory and possibly the greatest discovery in the twentieth century. It has enabled humankind to do many things that could not have been done without it, and it has become as indispensable as silicon in the makeup of many electronic systems. Its most immediate applications have been for the control of complex dynamic systems such as continuous manufacturing processes, aircraft, ships, or spacecraft.

In order to control a dynamic system, it is necessary to first know what the system is doing. For these applications, it is not always possible or desirable to measure every variable that needs to be controlled. The Kalman filter provides a means for inferring the missing information from indirect (and noisy) measurements. In such situations, the Kalman filter is used to estimate the complete state vector from partial state measurements and is called an "observer". The Kalman filter is also used for predicting the outcome of dynamic systems that people are not likely to control, such as the flow of rivers during flood conditions, the trajectories of celestial bodies, or the prices of traded commodities.

From a practical standpoint, this article will present the following perspectives:

1. Kalman filtering is an algorithm made from mathematical models. The Kalman filter makes it easier to solve a problem, but it does not solve the problem all by itself. As with any algorithm, it is important to understand its use and function before you can apply it effectively. The purpose of this article is to make the use of the Kalman filter sufficiently familiar so that it can be correctly and efficiently applied.
2. The Kalman filter is a recursive algorithm. It is ideally suited to digital computer implementation, in part because it uses a finite representation of the estimation problem—by a finite number of variables. It does, however, assume that these variables are real numbers with infinite precision. Some of the problems encountered in its use arise from the distinction between the finite dimension and the finite information, and the distinction between finite and manageable problem sizes. These are all issues on the practical side of Kalman filtering that must be considered along with the theory.
3. It is a complete statistical characterization of an estimation problem. It is much more than an estimator, because it propagates the entire probability distribution of the variables that it is tasked to estimate. This is a complete characterization of the current state of knowledge of the dynamic system, including the influence of all past measurements. These probability distributions are also useful for statistical analysis and predictive design of sensor systems.
4. In a limited context, the Kalman filter is a learning process. It uses a model of the estimation problem which distinguishes between phenomena (what one is able to observe), noumena (what is really going on), and the state of knowledge about the noumena that one can deduce from the phenomena. This state of knowledge is represented by probability distributions. To the extent that those probability distributions represent knowledge of the real world, and the cumulative processing of knowledge is learning; this is a learning process. It is a fairly simple one, but quite effective in many applications.

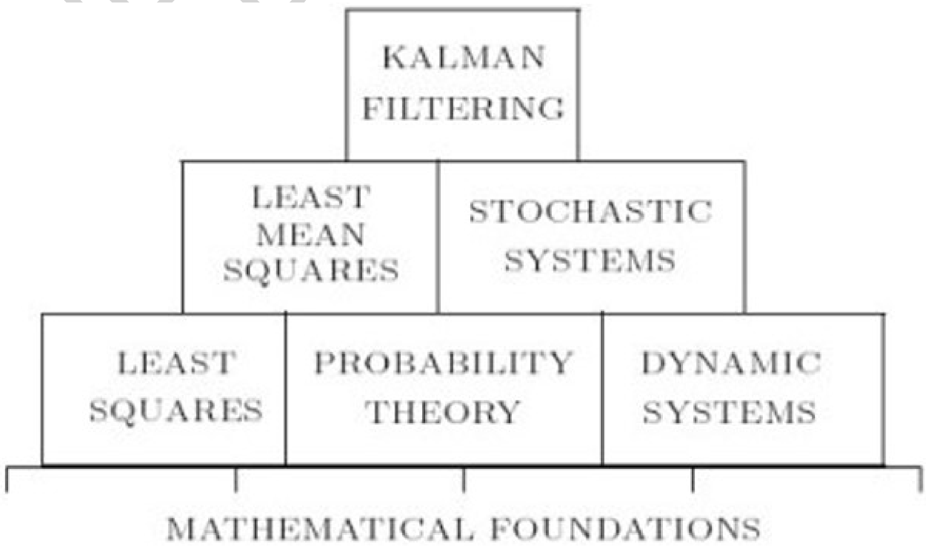

Figure 1: Foundational concepts in Kalman filtering

Figure 1 depicts the essential subjects forming the foundations for the Kalman filtering theory. Although this shows Kalman filtering as the apex of a pyramid, it is just a part of the foundations of another discipline—modern control theory—and a proper subset of statistical decision theory.

Applications of Kalman filtering encompass many fields. As a tool, the algorithm is used almost exclusively for estimation and performance analysis of estimators and as observers for control of a dynamical system. Except for a few fundamental physical constants, there is hardly anything in the universe that is truly constant. The orbital parameters of the asteroid Ceres are not constant, and even the "fixed" stars and continents are moving. Nearly all physical systems are dynamic to some degree. If one wants very precise estimates of their characteristics over time, then one has to consider their dynamics.

APPLICATION	DYNAMIC SYSTEM	SENSOR TYPES
Process control	Chemical plant	Pressure Temperature Flow rate Gas analyzer
Flood prediction	River system	Water level Rain gauge Weather radar
Tracking	Spacecraft	Radar Imaging system
Navigation	Ships Aircrafts, missiles Smart bombs Automobiles Golf carts	Sextant Log Gyroscope Accelerometer GPS receiver

Table 1: Examples of estimation problems

One does not always know the precise dynamics. Given this state of partial ignorance, the best that one can do is to express our ignorance more precisely—using probabilities. The Kalman filter allows us to estimate the state of such systems with certain types of random behavior by using statistical information. A few examples of such systems are listed in Table 1.

The third column of Table 1 lists some sensor types that one might use to estimate the state of the corresponding dynamic systems. The objective of design analysis is to determine how best to use these sensor types for a given set of design criteria. These criteria are typically related to estimation accuracy and system cost.

Because the Kalman filter uses a complete description of the probability distribution of its estimation errors in determining the optimal filtering gains, this probability distribution may be used in assessing its performance as a function of the "design parameters" of an estimation system, such as:

- The types of sensors to be used.
- The locations and orientations of the various sensor types with respect to the system to be estimated.
- The allowable noise characteristics of the sensors.
- The pre-filtering methods for smoothing sensor noise.
- The data sampling rates for the various sensor types.
- The level of model simplification to reduce implementation requirements.

This analytical capability of the Kalman filter enables system designers to assign “error budgets” to subsystems of an estimation system, and to trade off the budget allocations to optimize costs or other measures of performance while achieving a required level of estimation accuracy. Further, it acts like an observer by which all of the states not measured by the sensors can be constructed for use in the control system applications. In the following sections, we will be developing the following Kalman filters:

- Linear Kalman filters for discrete and continuous systems.
- Linearized and extended Kalman filters for discrete and continuous nonlinear systems.
- Implementation methods to overcome the shortcomings of the above two methods.

2. White Noise

Optimal Kalman filters are developed by assuming that the plant and measurement noises are white noise processes, or sequences. It is a common engineering practice to model: *uncertainty* in terms of Gaussian probability distributions and *dynamic uncertainty* in terms of linear dynamic systems disturbed by uncorrelated (white noise) processes—even though empirical analysis may indicate that the probability distributions are not truly Gaussian, and that the random processes are not truly white, or the relationships are not truly linear. Although this approach may be discarding useful information, we continue the practice for the following reasons:

1. **Approximation:** Probability distributions may not be precisely Gaussian, but close enough. Nonlinear systems are often smooth enough that local linearization is adequate. Even though the “flicker” noise observed in electronic systems cannot be modeled precisely using only white noise, it can often be done closely enough for practical purposes.
2. **Simplicity:** These models have few parameters to be estimated. Gaussian distributions are characterized by their means and variances, and white noise processses are characterized by their variances.
3. **Consistency:** Linearity preserves Gaussianity. That is, Gaussian probability distributions remain Gaussian under linear transformations of the variates.
4. **Tractability:** These models allow us to derive estimators minimizing expected squared errors.
5. **Good performance:** The resulting estimators have performed well for many important applications, despite apparent discrepancies between models and reality.
6. **Adaptability:** These estimators can often be extended to estimate parameters of the model or to track slow random variations in parameters.

7. **Extendability:** The variances used for calculating feedback gains can also be used for comparing performance to modeled performance, detecting anomalous behavior, and rejecting anomalous sensor data.

Vector valued random processes $x(t)$ and $y(t)$ are called uncorrelated, if their cross-covariance matrix is identically zero for all times, t_1 and t_2.

$$\mathrm{E}\left\{\left[x(t_1)-\mathrm{E}\left\{x(t_1)\right\}\right]\left[y(t_2)-\mathrm{E}\left\{y(t_2)\right\}\right]^T\right\}=0 \qquad (1)$$

where E is the expected value operator and T is the transpose of the vector.
The random process $x(t)$ is called uncorrelated if

$$\mathrm{E}\left\{\left[x(t_1)-\mathrm{E}\left\{x(t_1)\right\}\right]\left[x(t_2)-\mathrm{E}\left\{x(t_2)\right\}\right]^T\right\}=Q_1(t_1)\delta(t_1-t_2) \qquad (2)$$

where $\delta(t)$ is the Dirac delta "function" (actually, a generalized function), defined by

$$\int_a^b \delta(t)\,dt=\begin{cases}1 & \text{if } a\le 0\le b\\ 0 & \text{otherwise}\end{cases}. \qquad (3)$$

Similarly, a random sequence x_k in discrete time is called uncorrelated if

$$\mathrm{E}\left\{\left[x_k-\mathrm{E}x_k\right]\left[x_j-\mathrm{E}x_j\right]^T\right\}=Q_2(k)\Delta(k-j) \qquad (4)$$

where $\Delta(\cdot)$ is the Kronecker delta function, defined by

$$\Delta(k)=\begin{cases}1 & \text{if } k=0\\ 0 & \text{otherwise}\end{cases}. \qquad (5)$$

$Q_1(t)$ and $Q_2(k)$ are the intensity matrices of the white noise process and sequence, respectively. If $Q_1(t)$ and $Q_2(t)$ are constant, then the processes and sequences are stationary. If the probability distribution of a white noise process at each instant of time is Gaussian, then it is completely defined by its first two moments, mean and covariance. If E $x(t)$ = 0, the Gaussian process is called "zero mean."

A white noise process or sequence is an example of an uncorrelated process or sequence. Generally, a white noise process has no time structure. In other words, knowledge of the white process value at one instant of time provides no knowledge of what its value will be (or was) at any other point in time.

3. Linear Estimation

Linear estimation addresses the problem of estimating the state of a linear stochastic system by using measurements or sensor outputs that are linear functions of the state. We suppose that the stochastic systems can be represented by the types of plant and measurement models (for continuous and discrete time) shown as equations in Table 2, with dimensions of the vector and matrix quantities. The measurement and plant noise v_k and w_k, respectively, are assumed to be zero-mean Gaussian processes, and the initial value x_0 is a Gaussian random variable with known mean x_0 and known covariance matrix P_0. Although the noise sequences w_k and v_k are assumed to be uncorrelated, this restriction can be removed by modifying the estimator equations accordingly.

<table>
<tr><td>MODEL</td><td colspan="2">CONTINUOUS TIME</td><td colspan="2">DISCRETE TIME</td></tr>
<tr><td>Plant</td><td colspan="2">$\dot{x}(t) = F(t)x(t) + w(t)$</td><td colspan="2">$x_k = \Phi_{k\text{-}1}x_{k\text{-}1} + w_{k\text{-}1}$</td></tr>
<tr><td>measurement</td><td colspan="2">$z(t) = H(t)x(t) + v(t)$</td><td colspan="2">$z_k = H_k x_k + v_k$</td></tr>
<tr><td>Plant noise</td><td colspan="2">$\mathrm{E}\langle w(t)\rangle = 0$
$\mathrm{E}\left\langle w(t)w^T(s)\right\rangle = \delta(t-s)Q(t)$</td><td colspan="2">$\mathrm{E}\langle w_k\rangle = 0$
$\mathrm{E}\left\langle w_k w_i^T\right\rangle = \Delta(k-i)Q_k$</td></tr>
<tr><td>Observation</td><td colspan="2">$\mathrm{E}\langle v(t)\rangle = 0$</td><td colspan="2">$\mathrm{E}\langle v_k\rangle = 0$</td></tr>
<tr><td>noise</td><td colspan="2">$\mathrm{E}\left\langle v(t)v^T(s)\right\rangle = \delta(t\text{-}s)R(t)$</td><td colspan="2">$\mathrm{E}\left\langle v_k v_i^T\right\rangle = \Delta(k\text{-}i)R_k$</td></tr>
<tr><td colspan="5"></td></tr>
<tr><td>(LINEAR MODEL)</td><td>SYMBOL</td><td>DIMENSIONS</td><td>SYMBOL</td><td>DIMENSIONS</td></tr>
<tr><td rowspan="3">Dimensions of vectors and matrices</td><td>x, w</td><td>$n\times 1$</td><td>Φ, Q</td><td>$n\times n$</td></tr>
<tr><td>z, v</td><td>$l\times 1$</td><td>H</td><td>$l\times n$</td></tr>
<tr><td>R</td><td>$l\times l$</td><td>Δ, δ</td><td>SCALAR</td></tr>
</table>

Table 2: Linear plant and measurement models

The objective of statistical optimization is to find an estimate of the n state vector x_k represented by $\hat{x}_k$, a linear function of the measurements $z_i, \ldots, z_k$, that minimizes the weighted mean-squared error

$$\mathrm{E}\left[x_k - \hat{x}_k\right]^T M \left[x_k - \hat{x}_k\right], \tag{6}$$

where M is any symmetric non-negative definite weighting matrix.

We will now derive the mathematical form of an optimal linear estimator for the states of linear stochastic systems given in Table 2. This is called the linear quadratic Gaussian (LQG) estimation problem. The dynamic systems are linear, the performance cost functions are quadratic, and the random processes are Gaussian.
Let us consider similar types of estimators for the LQG problem:

- **Filters** use observations up to the time that the state of the dynamic system is to be estimated:

$$t_{\text{obs.}} \le t_{\text{est.}} \tag{7}$$

- **Predictors** estimate the state of the dynamic system beyond the time of the observations:

$$t_{\text{obs.}} < t_{\text{est.}} \tag{8}$$

This is a relatively minor distinction, and the differences between the respective estimators are correspondingly slight.

A straightforward and simple approach using orthogonality principles is used in the derivation of estimators. These estimators will have minimum variance and be unbiased and consistent.

The Kalman filter can be characterized as an algorithm for computing the conditional mean and covariance of the probability distribution of the state of a linear stochastic system with uncorrelated Gaussian process and measurement noise. The conditional mean is the unique unbiased estimate. It is propagated in feedback form by a system of linear differential equations, or by the corresponding discrete-time equations. The conditional covariance is propagated by a nonlinear differential equation, or its discrete-time equivalent. This implementation automatically minimizes any quadratic loss function of the estimation error.

The statistical performance of the estimator can be predicted *a priori* (that is, before it is actually used) by solving the nonlinear differential (or difference) equations used in computing the optimal feedback gains of the estimator. These are called "Riccati equations," named in 1763 by Jean le Rond D'Alembert (1717-1783) for Count Jacopo Francesco Riccati (1676-1754), who had studied a second-order scalar differential equation, although not the form that we have here. Kalman gives credit to Richard S. Bucy for the discovery that the Riccati differential equation serves the same role as the Wiener-Hopf integral equation in defining optimal gains. The Riccati equation also arises naturally in the problem of separation of variables in ordinary differential equations, and in the transformation of two-point boundary value problems to initial value problems. The behavior of their solutions can be shown analytically in trivial cases. These equations also provide a means for verifying the proper performance of the actual estimator when it is running.

-
-
-

Bibliography

Athans M. (1971). The role and use of the stochastic linear-quadratic-Gaussian problem in control system design. *IEEE Transactions on Automatic Control* **16** (6), 529-552. [This article shows a relationship between estimation theory and control systems. It is part of a special issue on the linear-quadratic-Gaussian estimation and control problem.]

Bierman G. J. (1977). *Factorization Methods for Discrete Sequential Estimation*, 241pp. New York: Academic Press. [Bierman derives the modified Cholesky decomposition, referred to as UDU transpose.]

Bittanti S., Laub A. J., and Willems J. C. (eds.). (1991). *The Riccati Equation*. New York: Springer-Verlag. [This book describes different methods for solving the Riccati equation.]

Grewal M. S. and Andrews A. P. (2001). *Kalman Filtering Theory & Practice Using MATLAB,* Second Edition, 401 pp. New York, Chichester, Brisbane: John Wiley & Sons. [Grewal and Andrews examine the subtleties, common problems, and limitations of estimation theory as it applies to real-world situations. It contains a complete history of Kalman filtering, gives a thorough description of square root algorithms, illustrative practical examples, and companion software in MATLAB.]

Grewal M.S., Weill L. R., and Andrews A. P. (2001). *Global Positioning Systems, Inertial Navigation and Integration*, 392 pp. New York, Chichester, Brisbane: John Wiley & Sons. [This book gives the application of Kalman filtering to GPS, INS, and describes the Kalman filtering engineering. It includes a companion disk in MATLAB.]

Kalman R. E. (1960). A new approach to linear filtering and prediction problems. *American Society of Mechanical Engineers Journal of Basic Engineering* **82**, 34-45. [This is the publication in which Kalman first introduced the filtering and prediction algorithms which became known as “Kalman filtering”.]

Kalman R. E. (1963). New methods in Wiener filtering. (*Proceedings of First Symposium on Engineering Applications of Random Function Theory and Probability*). New York, Chichester, Brisbane: John Wiley & Sons. [Here Kalman describes the relationship between his filter and the Wiener filter.]

Kalman R. E. and Bucy Richard S. (1961). New results in linear filtering and prediction theory. *American Society of Mechanical Engineers Journal of Basic Engineering, Series D* **83**, 95-108. [This is a continuation of the original results of Kalman’s 1960 filter and prediction.]

Poubelle M. A., Petersen I. R., Gevers M. R., and Bitmead R. R. (1986). A miscellany of results on an equation of Count J. F. Riccati. *IEEE Transactions on Automatic Control* **31**, 651-654. [This article derives the solution of the Riccati equation.]

Thornton C. L. (1976). *Triangular Covariance Factorizations for Kalman Filtering*. Doctoral Thesis, School of Engineering, University of California, Los Angeles, California. [This is a UDU transpose temporal update of the Kalman filter covariance equation.]

Biographical Sketch

M. S. Grewal received his Ph.D. in Electrical Engineering from the University of Southern California. He co-authored *Global Positioning Systems*, *Inertial Navigation, and Integration*, Wiley and Sons, 2001; and *Kalman Filtering Theory* and *Practice,* Second Edition, Wiley and Sons, 2001; both with implementation software in MATLAB. He has published over 55 papers in refereed journals and proceedings and over 215 technical reports on applications in guidance, navigation, control and modeling. Dr. Grewal has been recognized as Engineer of the Year by the aerospace community. He is internationally known for his expertise in implementation of the Kalman filter. Presently, Dr. Grewal is Professor of Electrical Engineering at California State University, Fullerton. He is a registered Professional Engineer in the state of California, a Senior Member of the Institute of Electrical and Electronics Engineers, a Fellow of the Institute for the Advancement of Engineering, and a member of Institute of Navigation (ION).