# An Introduction to the Kalman Filter

**Greg Welch[1] and Gary Bishop[2]**

**TR 95-041**
**Department of Computer Science**
**University of North Carolina at Chapel Hill**
**Chapel Hill, NC 27599-3175**

**Updated: Monday, March 11, 2002**

## Abstract

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem. Since that time, due in large part to advances in digital computing, the Kalman filter has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation.

The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) solution of the least-squares method. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown.

The purpose of this paper is to provide a practical introduction to the discrete Kalman filter. This introduction includes a description and some discussion of the basic discrete Kalman filter, a derivation, description and some discussion of the extended Kalman filter, and a relatively simple (tangible) example with real numbers & results.

1. welch@cs.unc.edu, http://www.cs.unc.edu/~welch
2. gb@cs.unc.edu, http://www.cs.unc.edu/~gb

# 1 The Discrete Kalman Filter

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem [Kalman60]. Since that time, due in large part to advances in digital computing, the *Kalman filter* has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation. A very "friendly" introduction to the general idea of the Kalman filter can be found in Chapter 1 of [Maybeck79], while a more complete introductory discussion can be found in [Sorenson70], which also contains some interesting historical narrative. More extensive references include [Gelb74; Grewal93; Maybeck79; Lewis86; Brown92; Jacobs93].

### *The Process to be Estimated*

The Kalman filter addresses the general problem of trying to estimate the state $x \in \Re^n$ of a discrete-time controlled process that is governed by the linear stochastic difference equation

$$x_k = Ax_{k-1} + Bu_k + w_{k-1}, \tag{1.1}$$

with a measurement $z \in \Re^m$ that is

$$z_k = Hx_k + v_k. \tag{1.2}$$

The random variables $w_k$ and $v_k$ represent the process and measurement noise (respectively). They are assumed to be independent (of each other), white, and with normal probability distributions

$$p(w) \sim N(0, Q), \tag{1.3}$$

$$p(v) \sim N(0, R). \tag{1.4}$$

In practice, the *process noise covariance* $Q$ and *measurement noise covariance* $R$ matrices might change with each time step or measurement, however here we assume they are constant.

The $n \times n$ matrix $A$ in the difference equation (1.1) relates the state at the previous time step $k-1$ to the state at the current step $k$, in the absence of either a driving function or process noise. Note that in practice $A$ might change with each time step, but here we assume it is constant. The $n \times l$ matrix $B$ relates the optional control input $u \in \Re^l$ to the state $x$. The $m \times n$ matrix $H$ in the measurement equation (1.2) relates the state to the measurement $z_k$. In practice $H$ might change with each time step or measurement, but here we assume it is constant.

### *The Computational Origins of the Filter*

We define $\hat{x}_k^- \in \Re^n$ (note the "super minus") to be our *a priori* state estimate at step $k$ given knowledge of the process prior to step $k$, and $\hat{x}_k \in \Re^n$ to be our *a posteriori* state estimate at step $k$ given measurement $z_k$. We can then define *a priori* and *a posteriori* estimate errors as

$$e_k^- \equiv x_k - \hat{x}_k^-, \text{ and}$$

$$e_k \equiv x_k - \hat{x}_k.$$

The *a priori* estimate error covariance is then

$$P_k^- = E[e_k^- e_k^{-T}], \tag{1.5}$$

and the *a posteriori* estimate error covariance is

$$P_k = E[e_k e_k^T]. \tag{1.6}$$

In deriving the equations for the Kalman filter, we begin with the goal of finding an equation that computes an *a posteriori* state estimate $\hat{x}_k$ as a linear combination of an *a priori* estimate $\hat{x}_k^-$ and a weighted difference between an actual measurement $z_k$ and a measurement prediction $H\hat{x}_k^-$ as shown below in (1.7). Some justification for (1.7) is given in "The Probabilistic Origins of the Filter" found below.

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \tag{1.7}$$

The difference $(z_k - H\hat{x}_k^-)$ in (1.7) is called the measurement *innovation*, or the *residual*. The residual reflects the discrepancy between the predicted measurement $H\hat{x}_k^-$ and the actual measurement $z_k$. A residual of zero means that the two are in complete agreement.

The $n \times m$ matrix $K$ in (1.7) is chosen to be the *gain* or *blending factor* that minimizes the *a posteriori* error covariance (1.6). This minimization can be accomplished by first substituting (1.7) into the above definition for $e_k$, substituting that into (1.6), performing the indicated expectations, taking the derivative of the trace of the result with respect to *K*, setting that result equal to zero, and then solving for *K*. For more details see [Maybeck79; Brown92; Jacobs93]. One form of the resulting *K* that minimizes (1.6) is given by[1]

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ &= \frac{P_k^- H^T}{H P_k^- H^T + R} \end{aligned} \tag{1.8}$$

Looking at (1.8) we see that as the measurement error covariance $R$ approaches zero, the gain $K$ weights the residual more heavily. Specifically,

$$\lim_{R_k \to 0} K_k = H^{-1}.$$

On the other hand, as the *a priori* estimate error covariance $P_k^-$ approaches zero, the gain *K* weights the residual less heavily. Specifically,

$$\lim_{P_k^- \to 0} K_k = 0.$$

1. All of the Kalman filter equations can be algebraically manipulated into to several forms. Equation (1.8) represents the Kalman gain in one popular form.