CALIFORNIA INSTITUTE OF TECHNOLOGY
Control and Dynamical Systems

CDS 110b

R. M. Murray **Kalman Filters** 14 January 2007

This set of lectures provides a brief introduction to Kalman filtering, following the treatment in Friedland.

Reading:

- Friedland, Chapter 11

1 Kalman Filters in Discrete Time

One of the principal uses of observers in practice is to estimate the state of a system in the presence of *noisy* measurements. We have not yet treated noise in our analysis and a full treatment of stochastic dynamical systems is beyond the scope of this text. In this section, we present a brief introduction to the use of stochastic systems analysis for constructing observers. We work primarily in discrete time to avoid some of the complications associated with continuous time random processes and to keep the mathematical pre-requisites to a minimum. This section assumes basic knowledge of random variables and stochastic processes.

Consider a discrete time, linear system with dynamics

$$\begin{aligned} x[k+1] &= Ax[k] + Bu[k] + Fv[k] \\ y[k] &= Cx[k] + w[k], \end{aligned} \tag{1}$$

where $v[k]$ and $w[k]$ are Gaussian, white noise processes satisfying

$$\begin{aligned} E\{v[k]\} &= 0 & E\{w[k]\} &= 0 \\ E\{v[k]v[j]^T\} &= \begin{cases} 0 & k \neq j \\ R_v & k = j \end{cases} & E\{w[k]w[j]^T\} &= \begin{cases} 0 & k \neq j \\ R_w & k = j \end{cases} \\ & & E\{v[k]w[j]^T\} = 0. \end{aligned} \tag{2}$$

We assume that the initial condition is also modeled as a Gaussian random variable with

$$E\{x_0\} = x_0 \qquad E\{x_0 x_0^T\} = P_0. \tag{3}$$

We wish to find an estimate $\hat{x}[k]$ that minimizes the mean square error $E\{(x[k]-\hat{x}[k])(x[k]-\hat{x}[k])^T\}$ given the measurements $\{y(\delta) : 0 \leq \tau \leq t\}$. We consider an observer in the same basic form as derived previously:

$$\hat{x}[k+1] = A\hat{x}[k] + Bu[k] + L[k](y[k] - C\hat{x}[k]). \tag{4}$$

The following theorem summarizes the main result.

Theorem 1. *Consider a random process $x[k]$ with dynamics* (1) *and noise processes and initial conditions described by equations* (2) *and* (3)*. The observer gain L that minimizes the mean square error is given by*

$$L[k] = A^T P[k] C^T (R_w + CP[k]C^T)^{-1},$$

where

$$\begin{aligned} P[k+1] &= (A - LC)P[k](A - LC)^T + R_v + LR_w L^T \\ P_0 &= E\{X(0)X^T(0)\}. \end{aligned} \tag{5}$$

Before we prove this result, we reflect on its form and function. First, note that the Kalman filter has the form of a *recursive* filter: given $P[k] = E\{E[k]E[k]^T\}$ at time k, can compute how the estimate and covariance *change*. Thus we do not need to keep track of old values of the output. Furthermore, the Kalman filter gives the estimate $\hat{x}[k]$ *and* the covariance $P[k]$, so we can see how reliable the estimate is. It can also be shown that the Kalman filter extracts the maximum possible information about output data. If we form the residual between the measured output and the estimated output,

$$e[k] = y[k] - C\hat{x}[k],$$

we can can show that for the Kalman filter the correlation matrix is

$$R_e(j,k) = W\delta_{jk}.$$

In other words, the error is a white noise process, so there is no remaining dynamic information content in the error.

In the special case when the noise is stationary (R_v, R_w constant) and *if* $P[k]$ converges, then the observer gain is constant:

$$K = A^T PC^T (R_w + CPC^T),$$

where

$$P = APA^T + R_v - APC^T \left(R_w + CPC^T\right)^{-1} CPA^T.$$

We see that the optimal gain depends on both the process noise and the measurement noise, but in a nontrivial way. Like the use of LQR to choose state feedback gains, the Kalman filter permits a systematic derivation of the observer gains given a description of the noise processes. The solution for the constant gain case is solved by the `dlqe` command in MATLAB.

Proof (of theorem). We wish to minimize the mean square of the error, $E\{(x[k]-\hat{x}[k])(x[k]-\hat{x}[k])^T\}$. We will define this quantity as $P[k]$ and then show that it satisfies the recursion given in equation (5). By definition,

$$\begin{aligned} P[k+1] &= E\{x[k+1]x[k+1]^T\} \\ &= (A-LC)P[k](A-LC)^T + R_v + LR_wL^T \\ &= AP[k]A^T - AP[k]C^TL^T - LCA^T + L(R_w + CP[k]C^T)L^T. \end{aligned}$$

Letting $R_\epsilon = (R_w + CP[k]C^T)$, we have

$$\begin{aligned} P[k+1] &= AP[k]A^T - AP[k]C^TL^T - LCA^T + LR_\epsilon L^T \\ &= AP[k]A^T + \big(L - AP[k]C^TR_\epsilon^{-1}\big)R_\epsilon\big(L - AP[k]C^TR_\epsilon^{-1}\big)^T \\ &\quad - AP[k]C^TR_\epsilon^{-1}CP[k]^TA^T + R_w. \end{aligned}$$

In order to minimize this expression, we choose $L = AP[k]C^TR_\epsilon^{-1}$ and the theorem is proven. □

2 Predictor-Corrector Form

The Kalman filter can be written in a two step form by separating the correction step (where we make use of new measurements of the output) and the prediction step (where we compute the expected state and covariance at the next time instant).

We make use of the notation $\hat{x}[k|j]$ to represent the estimated state at time instant k given the information up to time j (where typically $j = k-1$). Using this notation, the filter can be solved using the following algorithm:

Step 0: Initialization

$$\begin{aligned} k &= 0 \\ \hat{x}[0|-1] &= E\{x[0]\} \\ P[0|-1] &= E\{x^T[0]x[0]\} \end{aligned}$$

Step 1: Correction

$$\begin{aligned} \hat{x}[k|k] &= \hat{x}[k|k-1] + L[k](y[k] - C\hat{x}[k|k-1]) \\ P[k|k] &= P[k|k-1] - P[k|k-1]C^T(CP[k|k-1]C^T + R_w[k])^{-1}CP[k|k-1] \end{aligned}$$

Step 2: Prediction

$$\begin{aligned} \hat{x}[k+1|k] &= A\hat{x}[k|k] + Bu[k] \\ P[k+1|k] &= AP[k|k]A^T + FR_v[k]F^T \end{aligned}$$

Step 3: Interate Set k to $k + 1$ and repeat steps 1 and 2.

Note that the correction step reduces the covariance by an amount related to the relative accuracy of the measurement, while the prediction step increases the covariance by an amount related to the process disturbance.

This form of the discrete-time Kalman filter is convenient because we can reason about the estimate in the case when we do not obtain a measurement on every interation of the algorithm. In this case, we simply update the prediction step (increasing the covariance) until we receive new sensor data, at which point we call the correction step (decreasing the covariance).

3 Sensor Fusion

4 Information Filters