ECSE 6520: Estimation and Detection Theory

Kalman Filtering

Class Notes - 13

April 3rd, 2014

Contents

1 Introduction

- Wiener filtering assumes that the underlying processes are wide sense stationary, whereas in Kalman filtering the underlying processes are assumed to be Gauss-Markov.
- Wiener filtering uses data from infinite past and also from future (non-causal filter). Kalman filter, on the other hand, uses finite set of observations from past.
- Finally, in Kalman filtering the model (state and measurement models) can evolve over time whereas in Wiener filtering, the underlying model is fixed.

2 Evolution-Observation Model

- We consider two stochastic processes:
  - $\{X[k]\}_{k=0}^{\infty}$, where $X[k] = [X_1[k], X_2[k], ...X_p[k]]^T$ is a random vector, represents the primary quantities we wish to estimate. This random vector is called the *state vector*.
  - $\{Y[k]\}_{k=1}^{\infty}$, where $Y[k] = [Y_1[k], Y_2[k], ...Y_M[k]]^T$, represents the measurements. We refer to $Y[k]$ as the *observation* at the kth time instant.
- We postulate the following three properties for these processes:
  1. The process $\{X[k]\}_{k=0}^{\infty}$ is a Markov process, that is

$$f(x[k+1] \mid x[0], x[1], ..., x[k]) = f(x[k+1] \mid x[k]) \quad for\ all \quad k = 0, 1, ...$$

  2. The process $\{Y[k]\}_{k=1}^{\infty}$ is a Markov process with respect to the history of $\{X[k]\}_{k=0}^{\infty}$, that is

$$f(y[k] \mid x[0], x[1], ..., x[k]) = f(y[k] \mid x[k]) \quad for\ all \quad k = 0, 1, ...$$

3. The process $\{X[k]\}_{k=0}^{\infty}$ depends on the past observations only through its own history, that is,

$$f(x[k+1] \mid x[k], y[1], ..., y[k]) = f(x[k+1] \mid x[k]) \quad for\ all \quad k = 0, 1, ...$$

If the stochastic process $\{X[k]\}_{k=0}^{\infty}$ and $\{Y[k]\}_{k=1}^{\infty}$ satisfy conditions 1 to 3 above, the pair is called an *evolution-observation model*.

- However, for the evolution-observation model to be complete, we need to specify the following:
  - The pdf of the initial state $X[0]$.
  - The Markov transition pdf $f(x[k+1] \mid x[k])$, for $k = 0, 1, ...$
  - The likelihood function $f(y[k] \mid x[k])$, for $k = 1, 2, ...$

To better understand the assumptions above, consider the case that is often the starting point in practice:

- Assume that we have a Markov model describing the evolution of the states $X[k]$ and an observation model for vectors $Y[k]$ depending on the current state $X[k]$,

$$X[k+1] = F_{k+1}(X[k], W[k+1]), \qquad k = 0, 1, 2, \ldots \tag{1}$$

$$Y[k] = G_k(X[k], V[k]) \qquad k = 1, 2, \ldots \tag{2}$$

  - We assume the functions F_{k+1} and G_k are known.
  - The random vectors $W[k+1]$ and $V[k]$ are called the *state noise* and *observation noise*, respectively. The equation (1) is called the *state evolution equation* and (2) is called the *observation equation*.

- In order that the processes $\{X[k]\}_{k=0}^{\infty}$ and $\{Y[k]\}_{k=1}^{\infty}$ are an evolution-observation model, we make the following assumptions concerning the state noise and observation noise processes:

  1. For $k \neq l$, the noise vectors $W[k]$ and $W[l]$ as well as $V[k]$ and $V[l]$ are mutually independent and also mutually independent of the initial state $X[0]$.
  2. The noise vectors $W[k]$ and $V[l]$ are mutually independent for all k, l.

- Our objective is to estimate $X[k]$ based the measurements $Y[k]$. Let

$$D_k = \{y[1], y[2], \ldots, y[k]\}.$$

  In general, the problem of determining the pdf of

  - $X[k+1]$ given D_k is called the *prediction* problem;
  - $X[k]$ given D_k is called *filtering* problem and;
  - $X[k]$ given D_{k+p} $p \geq 1$ is called the *smoothing* problem.

  Our objective is to develop a recursive scheme where we first solve the filtering problem and next the prediction problem by making use of the solution of the filtering problem and the Markov model (evolution model) and finally use the solution of the prediction problem to update the filtering problem when a new observation becomes available using the observation model. Therefore, we need to find the formulas for the following updating steps:

  1. Time evolution updating: Given $f(x[k] \mid D_k)$, find $f(x[k+1] \mid D_k)$ based on $f(x[k+1] \mid x[k])$.
  2. Observation updating: Given $f(x[k+1] \mid D_k)$, find $f(x[k+1] \mid D_{k+1})$ based on the new observation $y[k+1]$ and the likelihood function $f(y[k+1] \mid x[k+1])$.

These updating equations are given in the following theorem.

Theorem 1 *Assume that the pair* $\{X[k]\}_{k=0}^{\infty}$ *and* $\{Y[k]\}_{k=1}^{\infty}$ *of stochastic processes is an evolution-observation model. Then the following updating formulas apply:*

1. *Time evolution updating:*

$$f(x[k+1] \mid D_k) = \int f(x[k+1] \mid x[k]) f(x[k] \mid D_k) dx[k] \tag{3}$$

2. *Observation updating:*

$$f(x[k+1] \mid D_{k+1}) = \frac{f(y[k+1] \mid x[k+1]) f(x[k+1] \mid D_k)}{f(y[k+1] \mid D_k)} \tag{4}$$

where

$$f(y[k+1] \mid D_k) = \int f(y[k+1] \mid x[k+1]) f(x[k+1] \mid D_k) dx[k+1]$$

3 Kalman Filter for the Linear Gaussian Model

Consider the following model:

$$X[k+1] = F_{k+1}X[k] + W[k+1] \qquad k = 0, 1, ...$$

$$Y[k] = G_k X[k] + V[k] \qquad k = 1, 2, ...$$

where

1. $Y[k]$, $k = 1, 1, ...$ - observation/measurement vectors.
2. $X[k]$, $k = 0, 1...$ - unknown state vectors.
3. F_k and G_k are deterministic and known state and observation matrices, respectively.
4. $W[k]$ and $V[k]$ are Gaussian state noise and observation noise vectors with known mean and covariance matrices. Without loss of generality, we assume that they are

zero mean with covariance matrices, R_{W_k} and R_{V_k}, respectively. Furthermore, we assume that they are mutually statistically independent, i.e.,

$$E[W_k W_l^T] = R_{W_k}\delta_{kl}$$

$$E[V_k V_l^T] = R_{V_k}\delta_{kl}$$

$$E[W_k V_l^T] = 0.$$

Note that the covariance matrices of $W[k]$ and $V[k]$ may vary wrt to k.

4. Finally, the probability distribution of $X[0]$ is known and Gaussian and without loss of generality, $X[0]$ is zero mean with covariance matrix R_0.

- **Notation -** Define

- $x_{k|l} = E[X[k] \mid D_l]$.
- $R_{k|l} = Cov(X[k] \mid D_l)$.
- $f(X[0]) = f(X[0] \mid D_0)$.

Theorem 2 *Assume that the above assumptions are valid. Then the time evolution and observation updating formulas take the following forms:*

1. *Time evolution updating : Assume that we know the Gaussian distribution*

$$X[k] \mid D_k \sim \mathcal{N}(x_{k|k}, R_{k|k}).$$

Then,

$$X[k+1] \mid D_k \sim \mathcal{N}(x_{k+1|k}, R_{k+1|k}),$$

where

$$x_{k+1|k} = F_{k+1} x_{k|k}, \tag{5}$$

$$R_{k+1|k} = F_{k+1} R_{k|k} F_{k+1}^T + R_{W_{k+1}}. \tag{6}$$

2. *Observation updating : Assume that we know the Gaussian distribution*

$$X[k+1] \mid D_k \sim \mathcal{N}(x_{k+1|k}, R_{k+1|k}).$$

Then,

$$X[k+1] \mid D_{k+1} \sim \mathcal{N}(x_{k+1|k+1}, R_{k+1|k+1}),$$

where

$$x_{k+1|k+1} = x_{k+1|k} + K_{k+1}(y[k+1] - G_{k+1}x_{k+1|k}), \tag{7}$$

$$R_{k+1|k+1} = (1 - K_{k+1}G_{k+1})R_{k+1|k}, \tag{8}$$

and the matrix K_{k+1}*, known as the Kalman gain matrix, is given by*

$$K_{k+1} = R_{k+1|k}G_{k+1}^T(G_{k+1}R_{k+1|k}G_{k+1}^T + R_{V_{k+1}})^{-1}.$$