ECE-597: Probability, Random Processes, and Estimation
Homework # 8

Due: Friday May 15, 2015

Background Formulas

We begin with two general (and identical) relationships derived in class.

$$\hat{x}|_k = \hat{x}|_{k-1} + R_{x\epsilon_k} R^{-1}_{\epsilon_k \epsilon_k} \epsilon_k \tag{1}$$

$$\hat{x}|_k = \sum_{j=0}^{k} R_{x\epsilon_j} R^{-1}_{\epsilon_k \epsilon_j} \epsilon_j \tag{2}$$

where $\hat{x}|_{k-1}$ is the linear least square estimate of x given observations $\{y_0, y_1, ...y_{k-1}\}$, $\hat{x}|_k$ is the linear least square estimate of x given observations $\{y_0, y_1, ..., y_k\}$, and $\epsilon_k = y_k - \hat{y}_{k|k-1}$ is a measure of new information or *innovation* in y_k that could not have been predicted by the previous observations $\{y_0, y_1, ...y_{k-1}\}$. By construction, $\hat{x}|_k$ is orthogonal to $\{y_0, ...y_k\}$ and ϵ_k is orthogonal to $\{y_0, ..., y_{k-1}\}$.

Observations Linearly Related to the Unknown

Assume the observed sequence y_k is related to the x_k through

$$y_k = H_k x_k + v_k$$

where $E[v_k v_l^T] = R_k \delta_{kl}$ (white noise) and $E[x_k v_l^T] = 0$ (noise uncorrelated with signal.) Now define

$$\begin{aligned} \epsilon_k &= y_k - \hat{y}_{k|k-1} \\ &= y_k - H_k \hat{x}_{k|k-1} \end{aligned}$$

where $\hat{y}_k$ is the l.l.s.e. of y_k given the observations $\{y_0, ..., y_{k-1}\}$ and $\hat{x}_{k|k-1}$ is the l.l.s.e. of x_k given $\{y_0, ..., y_{k-1}\}$. Now define

$$\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$$

so that $\tilde{x}_{k|k-1}$ is the *prediction error.*

(1) Show that

$$\boxed{\epsilon_k = H_k \tilde{x}_{k|k-1} + v_k}$$

(2) Show that

$$\boxed{E[\epsilon_k \epsilon_k^T] = E[H_k \tilde{x}_{k|k-1} \tilde{x}^T_{k|k-1} H_k^T] + E[v_k \tilde{x}^T_{k|k-1} H_k^T] + E[H_k \tilde{x}_{k|k-1} v_k^T] + E[v_k v_k^T]}$$

(3) Defining

$$P_{k|k-1} = E[\tilde{x}_{k|k-1}\tilde{x}_{k|k-1}^T]$$

where $P_{k|k-1}$ is the *error covariance matrix*, show that the expression in part 2 reduces to

$$\boxed{E[\epsilon_k \epsilon_k^T] = H_k P_{k|k-1} H_k^T + R_k}$$

We will define this as

$$R_k^\epsilon = E[\epsilon_k \epsilon_k^T]$$

(4) Substituting x_k for x in equation 1 yields

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + R_{x_k\epsilon_k}(R_k^\epsilon)^{-1}\epsilon_k \\ &= \hat{x}_{k|k-1} + R_{x_k\epsilon_k}(R_k^\epsilon)^{-1}(y_k - H_k\hat{x}_{k|k-1}) \end{aligned}$$

Now,

$$\begin{aligned} R_{x_k\epsilon_k} &= E[x_k\epsilon_k^T] \\ &= E[x_k(H_k\tilde{x}_{k|k-1} + v_k)^T] \\ &= E[x_k\tilde{x}_{k|k-1}^T]H_k^T \end{aligned}$$

Show that

$$\boxed{P_{k|k-1} = E[x_k\tilde{x}_{k|k-1}^T]}$$

and hence that

$$\boxed{R_{x_k\epsilon_k} = P_{k|k-1}H_k^T}$$

We will define this as

$$K_k = R_{x_k\epsilon_k}$$

So far we have

$$\begin{aligned} \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(R_k^\epsilon)^{-1}(y_k - H_k\hat{x}_{k|k-1}) \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k(R_k^\epsilon)^{-1}\epsilon_k \end{aligned}$$

where $\hat{x}_{k|k}$ is the l.l.s.e. of x_k based on observations $\{y_0, ..., y_k\}$, $\hat{x}_{k|k-1}$ is the l.l.s.e. of x_k given observations $\{y_0, ..., y_{k-1}\}$, $K_k(R_k^\epsilon)$ is the gain, and $(y_k - H_k\hat{x}_{k|k-1})$ is the error between the prediction and the observation at time k. Now we will examine how the error covariance matrix, P_k evolves.

(5) Starting from the definition

$$P_{k|k} = E[\tilde{x}_{k|k}\tilde{x}_{k|k}^T]$$

show that

$$\begin{aligned} P_{k|k} = \; & E[x_k(x_k - \hat{x}_{k|k-1})^T] - E[x_k\epsilon_k^T](R_k^\epsilon)^{-1}K_k^T \\ & -E[\hat{x}_{k|k-1}(x_k - \hat{x}_{k|k-1})^T] + E[\hat{x}_{k|k-1}\epsilon_k^T](R_k^\epsilon)^{-1}K_k^T \\ & -K_k(R_k^\epsilon)^{-1}E[\epsilon_k x_k^T] + K_k(R_k^\epsilon)^{-1}E[\epsilon_k\hat{x}_{k|k-1}^T] \\ & +K_k(R_k^\epsilon)^{-1}E[\epsilon_k\epsilon_k^T](R_k^\epsilon)^{-1}K_k^T \end{aligned}$$

and show that, by analyzing each term, this can be reduced to

$$\boxed{P_{k|k} = P_{k|k-1} - K_k (R_k^\epsilon)^{-1} K_k^T}$$

Hints: (1) use the result from part 1, (2) the results for the second, fifth, and seventh term are the same, (3) most of the remaining terms can be shown to be zero using orthogonality.

State Space Signal Models

Now we assume we know some dynamics:

$$x_{k+1} = \Phi_k x_k + \Gamma_k w_k$$

and the observation part remains the same:

$$y_k = H_k x_k + v_k$$

We assume here that $E[w_k v_k^T] = 0$, $R_{ww}(k,l) = Q_k \delta_{kl}$, $E[w_k x_0^T] = 0$,$E[w_k] = 0$, $E[X_0 X_0^T] = \Pi_0$, $R_{vv}(k,l) = R_k \delta_{kl}$, $E[x_k v_l^T] = 0$, and and $\Phi_k, \Gamma_k, Q_k, \Pi_0, R_k$, and H_k are *known* matrices.

(6) Argue that

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k} + \Gamma_k \hat{w}_{k|k}$$

and, hence

$$\boxed{\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k}}$$

Specifically, why is $\hat{w}_{k|k}$ 0? *Hint: $\hat{w}_{k|k}$ is construncted from the observations* $\{y_0, ...y_k\}$. Hence, since we have already shown that

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k (R_k^\epsilon)^{-1} \left\{ y_k - H_k \hat{x}_{k|k-1} \right\}$$

we have

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k-1} + \Phi_k K_k (R_k^\epsilon)^{-1} \left\{ y_k - H_k \hat{x}_{k|k-1} \right\}$$

where $x_{0|-1}$ is defined to be zero.

(7) Now define

$$\Sigma_{k+1|k} = E[\hat{x}_{k+1|k} \hat{x}_{k+1|k}^T]$$

Show that

$$\boxed{\Sigma_{k+1|k} = \Phi_k \Sigma_{k|k-1} \Phi_k^T + \Phi_k K_k (R_k^\epsilon)^{-1} K_k^T \Phi_k^T}$$

where $\Sigma_{0|-1}$ is defined to be zero. *: Hints: (1) R_k^ϵ is symmetric, and (2) $\epsilon_k = y_k - H_k \hat{x}_{k|k-1}$ is the innovation, and is orthogonal to* $\{y_0, ..., y_{k-1}\}$.

(8) Define

$$\Pi_{k+1} = E[x_{k+1} x_{k+1}^T]$$

and show that

$$\boxed{\Pi_{k+1} = \Phi_k \Pi_k \Phi_k^T + \Gamma_k Q_k \Gamma_k^T}$$

(9) Recall the error covariance matrix is given by

$$P_{k+1|k} \quad = \quad E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T]$$

show that

$$\begin{aligned} P_{k+1|k} &= \Pi_{k+1} - \Sigma_{k+1|k} \\ &= \Phi_k \left\{ \Pi_k - \Sigma_{k|k-1} \right\} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T - \Phi_k K_k (R_k^\epsilon)^{-1} K_k^T \Phi_k^T \end{aligned}$$

and, finally,

$$\boxed{P_{k+1|k} = \Phi_k P_{k|k-1} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T - \Phi_k K_k (R_k^\epsilon)^{-1} K_k^T \Phi_k^T}$$

Hints: (1) Use the identities

$$E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T] \quad = \quad E[x_{k+1} x_{k+1}^T] - E[\hat{x}_{k+1|k}(x_{k+1} - \hat{x}_{k+1|k})^T] - E[x_{k+1} \hat{x}_{k+1|k}^T]$$

(2) Use orthogonality for the second term, and then use the trick

$$E[x_{k+1} \hat{x}_{k+1|k}^T] \quad = \quad E[(x_{k+1} - \hat{x}_{k+1|k} + \hat{x}_{k+1|k}) \hat{x}_{k+1|k}^T]$$

Summary of Equations

The recursive equations for the Kalman filter are:

$$\hat{x}_{k+1|k} = \Phi_k \hat{x}_{k|k-1} + \Phi_k K_k (R_k^\epsilon)^{-1} \left\{ y_k - H_k \hat{x}_{k|k-1} \right\}$$

where

$$\begin{aligned} \hat{x}_{0|-1} &= 0 \\ P_{0|-1} &= \Pi_0 \\ R_k^\epsilon &= H_k P_{k|k-1} H_k^T + R_k \\ K_k &= P_{k|k-1} H_k^T \\ P_{k+1|k} &= \Phi_k P_{k|k-1} \Phi_k^T + \Gamma_k Q_k \Gamma_k^T - \Phi_k K_k (R_k^\epsilon)^{-1} K_k^T \Phi_k^T \end{aligned}$$

There are, of course, alternative forms of the filter and alternative derivations.

(10) Using the Kalman filter equations, show that

$$\boxed{\hat{x}_{1|0} = \Phi_0 \Pi_0 H_0^T (H_0 \Pi_0 H_0^T + R_0)^{-1} y_0}$$

(11) Using the orthogonality condition determine the least squares estimator

$$\boxed{\hat{x}_{1|0} = \alpha y_0}$$

directly, i.e. determine the optimal α.

Computer Assignment

This part of the assignment can be done independently of the derivation. You mostly just play around with code. In all of your computer runs, look at 500 and 1500 time steps. You can get the code from the class website.

(1) We will now use the Kalman filter as an to estimate the parameters of a simple autoregressive model with **fixed** coefficients (i.e., a_1 and a_2 are fixed). Specifically, assume

$$y_k = -a_1 y_{k-1} + -a_2 y_{k-2} + v_k$$

were v_k is a white noise sequence with variance R.

Here x_k is our estimate of the coefficient for the AR process, $H_k = [y_{k-1}\ y_{k-2}]$,and Q is noise in the model. We will then set Φ and Γ equal to the identity matrix. Note that with this formulation, we are assuming the estimates do not change from one time instant to the next, which is a good assumption for constant coefficients, but not so good if they are time varying.

2) For the problem stated in (1), simulate the AR process with $a_1 = 0.7$, $a_2 = 0.12$, and $R = 0.05$ and $Q_k = 0.05$ (assume Q is diagonal), and determine the estimate of the coefficients as a function of time. Plot these estimates versus the true values for 500 and 1500 time steps. Turn in your graphs.

3) Modify Q (leave R at 0.05) to try and get a good estimate of the coefficients. Run your simulations for 500 and 1500 time steps, and turn in your graphs.

4) Now we will assume our parameters are changing as a function of time. Specifically, assume

$$\begin{aligned} a_1(k) &= 0.4 + 0.5\cos(3\pi k/200) \\ a_2(k) &= 0.5 + 0.3\sin(2\pi k/200) \end{aligned}$$

5) For the problem stated in (4), simulate the AR process using and $R = 0.05$ and $Q_k = 0.05$ (assume Q is diagonal), and determine the estimate of the coefficients as a function of time. Plot these estimates versus the true values for 500 and 1500 time steps. Turn in your graphs.

6) Modify R and Q to try and get a good estimates of the coefficients, and turn in your graphs.

7) Now we will again assume our parameters are changing as a function of time. Specifically, assume

$$\begin{aligned} a_1(k) &= 0.4 + 0.5\cos(3\pi k/200) \\ a_2(k) &= 1.5 + 0.3\sin(2\pi k/200) \end{aligned}$$

8) For the problem stated in (7), simulate the AR process using and $R = 0.05$ and $Q_k = 0.05$ (assume Q is diagonal), and determine the estimate of the coefficients as a function of time.

Plot these estimates versus the true values for 500 and 1500 time steps. Turn in your graphs.

9) Modify R and Q to try and get good estimates of the coefficients, and turn in your graphs.