EE226: Random Processes in Systems	**Fall'06**

Problem Set 5 — Due Oct, 24

Lecturer: Jean C. Walrand — *GSI: Assane Gueye*

This problem set essentially reviews estimation theory and the Kalman filter. Not all exercises are to be turned in. Only those with the sign ★ are due on *Tuesday, October* 24^{th} at the beginning of the class. Although the remaining exercises are not graded, you are encouraged to go through them.
We will discuss some of the exercises during discussion sections.
Please feel free to point out errors and notions that need to be clarified.

Exercise 5.1. *Let $Y = H \times X + Z$ where X is some Gaussian random vector in R^n with zero mean and covariance matrix K_X, H is a non-singular $n \times n$ known matrix, and Z is a gaussian random noise with zero mean and non-singular covariance matrix K_Z uncorrelated to X.*
(a) Find the MMSE estimator of X given Y.
(b) Explain what happens in the case where $|K_Z| = 0$.
(c) Repeat part (b) when both H and K_Z are singular.

Solution:Hint
(a) Since the matrix H is invertible, we can obtain a simpler equation $\tilde{Y} = H^{-1}Y = X + H^{-1}Z = X + \tilde{Z}$ where $\tilde{Z}$ is again Gaussian $\mathcal{N}(0, H^{-1}K_Z(H^{-1})^T)$ (note: the covariance matrix of $\tilde{Z}$ is not singular).
Since we have only Gaussian random vectors, the MMSE of X given $\tilde{Y}$ is given by (using formula)

$$\hat{X} = K_X \left(K_X + H^{-1}K_Z(H^{-1})^T\right)^{-1} H^{-1}Y$$

(b) Note that if $|K_Z| = 0$ then the covariance of $\tilde{Z}$ is singular. So, without loss of generality we can discuss the transformed observation $\tilde{Y}$.
The singularity of the covariance matrix indicates that the noise lies in a lower dimensional space then the signal X (which lies in R^n). The components of X that are in the subspace of the noise will be perturbed but the components that are orthogonal to the noise can be detected without error.
A more rigorous analysis using eigenvalue decomposition will involve projection into the subspace of the noise and into the orthogonal to that subspace.
(c) If H is singular we cannot make anymore transformation of the observation but still we can do the same analysis.
The effect of multiplying by H is to reduce the signal into a lower dimensional space. If this space is entirely inside the noise space, the observation is noisy. If some components are still in the orthogonal direction of the noise, those components can be detected without error.

Exercise 5.2. ★

The values of a random sample, $2.9, 0.5, -0.1, 1.2, 3.5,$ *and* 0*, are obtained from a random variable* X *uniformly distributed over the interval* $[a, b]$*. Find the maximum-likelihood estimates of* a *and* b *(assume that the samples are independent).*

Solution

The MLE is the value of (a, b) that maximizes the likelihood of observing the given sequence. In other terms, we are looking for $(\hat{a}, \hat{b})$ such that:

$$(\hat{a}, \hat{b}) = argmax_{(a,b)} f(x_1, \ldots, x_n; (a, b))$$

where x_i's are the observations.

The pdf $f(x_1, \ldots, x_n; (a, b))$ is equal to

$$f(x_1, \ldots, x_n; (a, b)) = \begin{cases} \frac{1}{b-a} & a \leq x_i \leq b, \forall i \\ 0 & \text{otherwise} \end{cases}$$

The choice of (a, b) that maximizes this expression is clearly $a = x_{min} = \min\{x_1, \ldots, x_n\}$ and $b = x_{max} = \max\{x_1, \ldots, x_n\}$. Thus the MLE of (a, b) is $(\hat{a}, \hat{b}) = (-0.1, 3.5)$.

Exercise 5.3. ★

For the estimation problem modeled by the equations:

$$\begin{aligned} &x_k = x_{k-1} + w_{k-1}, \quad w_k \sim \mathcal{N}(0, 30), \ \textit{white noise} \\ &z_k = x_k + v_k, \quad v_k \sim \mathcal{N}(0, 20), \ \textit{white noise} \\ &\sigma_0^2 = 150 \end{aligned}$$

find σ_k^2, s_k, *and* r_k *for* $k = 1, 2, 3, 4$ *and* σ_∞^2 *(the steady value).*

(σ_k^2, s_k *are the estimation and prediction square error updates, and* r_k *is the Kalman gain.)*

Solution

Recalling from class, the Kalman filter update equations are:

$$\begin{aligned} \hat{x}_k &= \hat{x}_{k-1} + r_k(z_k - \hat{x}_{k-1}) \\ r_k &= \frac{s_k}{s_k + \sigma_v^2} \\ s_k &= \sigma_{k-1}^2 + \sigma_w^2 \\ \sigma_k^2 &= (1 - r_k)s_k \\ \sigma_0^2 &= 150; \end{aligned}$$

The following matlab script gives the results.

```
n=10; sigma=[150 zeros(1,n-1)]; sigma_v=20; sigma_w=30;
s=zeros(1,n); r= zeros(1,n);
for i=2:n
    s(i)=sigma(i-1)+sigma_w;
    r(i)=s(i)/(s(i)+sigma_v);
    sigma(i)=(1-r(i))*s(i);
end
```

k	0	1	2	3	4	5	6	7	8	9
s_k	0	180.0000	48.0000	44.1176	43.7615	43.7266	43.7232	43.7229	43.7228	43.7228
r_k	0	0.9000	0.7059	0.6881	0.6863	0.6862	0.6861	0.6861	0.6861	0.6861
σ_k	150.0000	18.0000	14.1176	13.7615	13.7266	13.7232	13.7229	13.7228	13.7228	13.7228

Exercise 5.4. ★ *Parameter Estimation (recursive)*
Let x be a zero-mean Gaussian random variable with variance P_0, and let $z_k = x + v_k$ be an observation of x with white noise $v_k \sim \mathcal{N}(0, R)$.
(a) Find a recursive (MMSE) estimator of x given the observations z_k and compute the estimation error.
Hint: Example 4.3 Gallager's notes.
(b) What is the value of $\hat{x}_1$ if $R = 0$?
(c) What is the value of $\hat{x}_1$ if $R = \infty$?
(d) Explain the results of (b) and (c) in terms of measurement uncertainty.

Solution
(a) The solution of this problem is given in the Gallager's notes (eq. 4.33).

$$\hat{x}_k = \frac{P_0}{kP_0 + R} \sum_{i=1}^{k} z_k$$
$$\sigma_k^2 = \frac{P_0 R}{kP_0 + R}$$

(b)

$$\hat{x}_1 = \frac{P_0 z_1}{P_0 + R} \xrightarrow{R \to 0} z_1$$

(c)

$$\hat{x}_1 = \frac{P_0 z_1}{P_0 + R} \xrightarrow{R \to \infty} 0$$

(d) In part (b), the noise has zero mean and zero variance. Since it is a Gaussian random variable we can conclude that it is identically equal to zero. So the estimation is correct and has zero error.
In part (c) the noise has infinite variance so the observation becomes independent to the signal. Thus the best estimate given the observation is equal to the mean of x.

Exercise 5.5. ★ *Parameter Estimation (using KF)*
Let us consider the estimation of the value of an (unknown) constant x given measurements

$y_n = x + v_n$ that are corrupted (but uncorrelated) with a zero mean white noise v_n that has variance σ_v^2.
(a) Write the estimation problem as a Kalman filter problem and compute the Kalman gain (r_n) and the variance of the estimation error (e_n^2).
(You are asked to find close forms of e_n and r_n as functions of n, e_0^2, σ_v^2.
(b) What is the Kalman filter as $n \to \infty$?
(c) What is the Kalman filter as $\sigma_v^2 \to \infty$?
(d) Now suppose that we do not have no a priori information about x (i.e $\hat{x}_0 = 0$ and $e_0^2 \to \infty$).
Show that the Kalman filter simply becomes the sample mean

$$\hat{x} = \frac{1}{n}\sum_{i=1}^{n} y_i$$

Solution
(a) We can write the estimation problem into the following form:

$$\begin{aligned} x_{n+1} &= x_n + 0 \\ y_n &= x_n + v_n \end{aligned}$$

where v_n is a white noise.
Using the Kalman filter update equations we have

$$\begin{aligned} &\hat{x}_n = \hat{x}_{n-1} + r_n(y_n - \hat{x}_{n-1}) \\ &r_n = \frac{s_n}{s_n + \sigma_v^2} \\ &s_n = e_{n-1}^2 \\ &e_n^2 = (1 - r_n)s_n \end{aligned}$$

Substituting s_n and r_n in the last equation, we obtain

$$\begin{aligned} e_n^2 &= (1 - \frac{e_{n-1}}{e_{n-1} + \sigma_v^2})e_{n-1} \\ &= \frac{1}{\frac{1}{\sigma_v^2} + \frac{1}{e_{n-1}^2}} \end{aligned}$$

Thus we have:

$$\begin{aligned} \frac{1}{e_n^2} &= \frac{1}{\sigma_v^2} + \frac{1}{e_{n-1}^2} \\ &= \frac{1}{\sigma_v^2} + \frac{1}{\sigma_v^2} + \frac{1}{e_{n-1}^2} \\ &= \frac{1}{\sigma_v^2} + \frac{1}{\sigma_v^2} + \cdots + \frac{1}{e_0^2} \\ &= \frac{n}{\sigma_v^2} + \frac{1}{e_0^2} \end{aligned}$$

This gives:

$$e_n^2 = \frac{1}{\frac{n}{\sigma_v^2} + \frac{1}{e_0^2}}$$
$$s_n = \frac{1}{\frac{n-1}{\sigma_v^2} + \frac{1}{e_0^2}}$$
$$r_n = \frac{1}{n + \frac{\sigma_v^2}{e_0^2}}$$
$$\hat{x}_n = \hat{x}_{n-1} + \frac{1}{n + \frac{\sigma_v^2}{e_0^2}}(y_n - \hat{x}_{n-1})$$

(b) When $n \to \infty$, then

$$e_n^2 \to 0; \quad s_n \to 0; \quad r_n \to 0$$
$$\hat{x}_n = \hat{x}_{n-1} \quad \text{Stationary!}$$

(c) When $\sigma_v^2 \to \infty$ (infinite noise!), then

$$e_n^2 = s_n \to e_0^2$$
$$r_n \to 0$$
$$\hat{x}_n = \hat{x}_{n-1} = \hat{x}_{n-2} = \cdots = \hat{x}_0 \quad \text{thus we can ignore the obeservation}$$

(c) If $\hat{x}_0 = 0$ and $e_0^2 \to \infty$ then $r_n \to \frac{1}{n}$. So we have that

$$\begin{aligned}
\hat{x}_n &= \frac{n-1}{n}\hat{x}_{n-1} + \frac{1}{n}y_n \\
&= \frac{n-1}{n}\left(\frac{n-2}{n-1}\hat{x}_{n-2} + \frac{1}{n-1}y_{n-1}\right) + \frac{1}{n}y_n \\
&= \frac{n-2}{n}\hat{x}_{n-2} + \frac{1}{n-1}y_{n-1} + \frac{1}{n}y_n \\
&= \ldots \\
&= \frac{n-n}{n}\hat{x}_0 + \frac{1}{n}y_1 + \cdots + \frac{1}{n}y_{n-1} + \frac{1}{n}y_n \\
&= \frac{1}{n}\sum_{i=1}^{n} y_i
\end{aligned}$$

Exercise 5.6. *Consider a system consisting of two sensors, each making a simple measurement of an unknown constant x. Each measurement is noisy and may be modeled as follows*

$$\begin{aligned}
y(1) &= x + v(1) \\
y(2) &= x + v(2)
\end{aligned}$$

where $v_i, i = 1, 2$'s are zero mean, uncorrelated random variables with variances $\sigma_i^2, i = 1, 2$.
(a) We want to compute the best linear estimate of x of the form

$$\hat{x} = k_1 y(1) + k_2 y(2).$$

Find the values of k_1 and k_2 that produce an unbiased estimate of x that minimizes the mean-square error $E[(x - \hat{x})^2]$.
(b) Repeat part (a) for the case where the measurement are correlated,

$$E[v(1)v(2)] = \rho\sigma_1\sigma_2$$

where ρ is the correlation coefficient.
(c) Repeat part (a) in the framework of a Kalman filtering, treating the measurements $y(1)$ and $y(2)$ sequentially.

Solution: Hint
(a) Note that since we want an unbiased estimator, $\hat{x}$ should verify $E[\hat{x}] = x$. Since $E[y(i)] = x, i = 1, 2$, k_1 and k_2 should satisfy $k_1 + k_2 = 1$ or $k_2 = 1 - k_1$.
The mean square error is given by:

$$E[x - \hat{x}]^2 = E[x - k_1 y_1 + k_2 y_2 =]^2 = E[x - k_1 y_1 + (1 - k_1) y_2 =]^2$$

Developing and taking the derivative with respect to k_1 gives

$$\hat{k}_1 = \frac{\sigma_2^2}{\sigma_2^2 + \sigma_1^2}, \quad \hat{k}_2 = \frac{\sigma_1^2}{\sigma_2^2 + \sigma_1^2}$$

(b) This case does not differ the previous a lot...just take into account the cross-correlation. We will obtain:

$$\hat{k}_1 = \frac{\sigma_2^2 + \rho\sigma_1\sigma_2}{\sigma_2^2 + \sigma_1^2 + \rho\sigma_1\sigma_2}$$

(c) The Kalman filter formulation is similar to exercises 5.4 and 5.5.

Exercise 5.7. *A vector discrete-time random sequence x_k is given by*

$$x_k = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} x_{k-1} + w_{k-1}$$
$$w_k \sim \mathcal{N}(0, 1), \textit{white noise}$$

The observation equation is given by

$$z_k = [1|0]x_k + v_k$$
$$v_k \sim \mathcal{N}(0, 2 + (-1)^k), \textit{white noise}$$

Compute the Kalman filter updates for the recursive estimation of the process x_k given the observation z_k.
Assume that

$$\Sigma_0 = \begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$$

Solution: Hint
This is just a standard Kalman filter derivation. The only (maybe) new thing is that the variance of v_k depends on time. But we know from the generalization of the Kalman filter that the update equations will keep the same form.

Exercise 5.8. ★
Estimation of an autoregressive process
An autoregressive process of order 1 is described by the difference equation

$$x_n = 0.5x_{n-1} + w_n$$

where w_n is zero-mean white noise with a variance $\sigma_w^2 = 0.64$. The observed process y_n is given by

$$y_n = x_n + v_n$$

where v_n is zero-mean white noise with a variance $\sigma_v^2 = 1$.
(a) Write the Kalman filter equations to find the LLSE estimate $\hat{x}_n(y_1^n)$ of x_n given the observations $y_i, i = 1, \ldots, n$.
The initial conditions are $\hat{x}_0(\cdot) = 0, \sigma_0^2 = 1, (\sigma_i^2 = E[(x_i - \hat{x}_i)^2])$
(b) Assuming that the filter reaches a steady state solution, find the steady Kalman gain and the limiting form of the estimation equation for $\hat{x}_n(y_1^n)$.

Solution: Hint
This exercise is, up some coefficients, identical to 5.3 (Why did I assign it!)
Using the same matlab script (with minor changes) gives:

k	0	1	2	3	4	5	6	7	8	9
s_k	0	0.8900	0.7577	0.7478	0.7470	0.7469	0.7469	0.7469	0.7469	0.7469
r_k	0	0.4709	0.4311	0.4278	0.4276	0.4276	0.4276	0.4276	0.4276	0.4276
σ_k	1.0000	0.4709	0.4311	0.4278	0.4276	0.4276	0.4276	0.4276	0.4276	0.4276

And the steady filer will be:

$$\begin{aligned} \hat{x}_n &= 0.5\hat{x}_{n-1} + 0.4276(y_n - \hat{x}_{n-1}) \\ &= 0.0724\hat{x}_{n-1} + 0.4276y_n \end{aligned}$$

Exercise 5.9. ★
In class we have derived the Kalman filter equations of the estimation problem,

$$X_{n+1} = AX_n + V_n \quad \text{and} \quad Y_n = CX_n + W_n, n \geq 1$$

where $\{X_1, V_n, W_n, n \geq 1\}$ are all orthogonal and are zero-mean with cov$(V_n) = K_V$ and cov$(W_n) = K_W$. (ref. Theorem 10.2 of the course notes for R_n, S_n, Σ_n)
In this exercise, we will derive the following expression for the Kalman gain,

$$R_n = \Sigma_n C^T K_W^{-1} \tag{5.1}$$

(a) By substituting the expression given for the Kalman gain R_n into that of the estimation error covariance matirx Σ_n, show that

$$\Sigma_n = S_n - S_n C^T \times \left[C S_n C^T + K_W \right]^{-1} C S_n^T$$

(b) Using the matrix inversion Lemma, show that that inverse covariance matrix can be written as

$$\Sigma_n^{-1} = S_n^{-1} + C^T K_W^{-1} C$$

Theorem: *Matrix Inversion Lemma*
Suppose A is $n \times n$, B is $n \times m$, and D is $m \times n$ with A and C nonsingular matrices. Then

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$$

(c) By using the results in part (b) show that the expression of R_n given in the note can be written as eq. 5.1.

Solution:
(a) Substituting $R_n = S_n C^T (C S_n C^T + K_W)^{-1}$ into the expression of Σ_n we obtain:

$$\Sigma_n = (I - R_n C) S_n = S_n - S_n C^T \left[C S_n C^T + K_W \right]^{-1} C S_n^T$$

(b) Using the Matrix Inversion Lemma in the RHS ($A \leftrightarrow S_n^{-1}, C \leftrightarrow K_W^{-1}, D = B^T \leftrightarrow C$) we obtain:

$$\Sigma_n = [S_n^{-1} + C^T K_W^{-1} C]^{-1}$$

which gives the result.
(c) To get equation 5.1 observe that

$$\begin{aligned}
R_n &= S_n C^T (C S_n C^T + K_W)^{-1} \\
&= S_n C^T \left[K_W^{-1} - K_W^{-1} C [S_n^{-1} + C^T K_W^{-1} C]^{-1} C_T K_W^{-1} \right]; \quad \text{using Lemma} \\
&= S_n C^T K_W^{-1} - S_n C^T K_W^{-1} C \Sigma_n C_T K_W^{-1}; \quad \text{using part (b)} \\
&= \left[S_n \Sigma_n^{-1} - S_n C^T K_W^{-1} C \right] \Sigma_n C_T K_W^{-1} \\
&= \Sigma_n C_T K_W^{-1}
\end{aligned}$$

where in the last step we have used part (b) again.

Exercise 5.10. ★ ★ ***Bonus***
Derive the Kalman filter equations for the general case

$$X_{n+1} = A_n X_n + V_n \quad \text{and} \quad Y_n = C_n X_n + W_n, n \geq 1$$

where $E[V_n V_k^T] = K_V(n)\delta(n-k)$ *and* $E[W_n W_k^T] = K_W(n)\delta(n-k)$*, with the usual orthogonality assumptions.*

Solution: Hint
This exercise was assigned to encourage you to go through the Kalman filter derivation at least once. It is important to know the different steps of the derivation.

Exercise 5.11. ★ *In the Kalman filter setting, we are always interested in the limiting behavior of the updates.*
Let's consider the KF problem

$$X_{n+1} = AX_n + V_n \quad \text{and} \quad Y_n = CX_n + W_n$$

with the usual orthogonality assumptions.
Assume that $K_V = QQ^T$ *and that* (A, Q) *is reachable and* (A, C) *is observable.*
(a) For $A = I_n$ *and* $cov(X_1 - L(X_1|Y_0)) = 0$*, show that there exists a positive semi-definite matrix* M *such that*

$$MC^T(CMC^T + K_W)CM - QQ^T = 0$$

(b) Now suppose that A *is general.*
Show that if the prediction error covariance matrix S_n *converges to some limit* S*, then* S *must satisfy*

$$S = (A - ARC)S(A - ARC)^T + ARK_W R^T A^T + K_V$$

where R *is some matrix that you should specify.*

Solution
(a) Observe that if $A = I$ we can rewrite the prediction error covariance matrix S_n as

$$\begin{aligned} S_n &= \Sigma_{n-1} + K_V \\ &= (I - R_{n-1C})S_{n-1} + K_V; \quad \text{using the expression of } \Sigma_n \\ &= S_{n-1} - S_{n-1}C^T[CS_{n-1}C^T + K_W]^{-1} + K_V; \quad \text{substituting } R_{n-1} \end{aligned}$$

Since $K_V = QQ^T$, (A, Q) is reachable, and (A, C) is observable, Theorem 11.2 applies and there exist limiting matrices

$$\Sigma_n \to \Sigma, R_n \to R, S_n \to S$$

At the limit, S verifies

$$S = S - SC^T[CSC^T + K_W]^{-1} + K_V$$

Rearranging and observing that $K_V = QQ^T$ we get the result for $M = S$.
(b) The key point in this part is to show that

$$S_{n+1} = (A - ARC)S_n(A - ARC)^T + ARK_W R^T A^T + K_V \tag{5.2}$$

To show that, recall that $S_{n+1} = cov(X_{n+1} - \hat{X}_n)$. Now observe that

$$\begin{aligned} X_{n+1} - \hat{X}_n &= AX_n + V_n - A\left[A\hat{X}_{n-1} + R_N(Y_n - R_n C\hat{X}_{n-1})\right] \\ &= AX_n + V_n - A\left[A\hat{X}_{n-1} + R_n(CX_n + W_n - R_n C\hat{X}_{n-1})\right] \\ &= (A - AR_nC)(X_n - \hat{X}_{n-1}) + V_n + AR_nW_n \end{aligned}$$

Take covariance in both sides and use the fact that $X_n - \hat{X}_{n-1}$, W_n, and V_n are pairwise uncorrelated to get equation 5.2.
Now assuming that S_n converges and going to the limit gives the result with R being the limiting matrix of R_n.

Exercise 5.12. ★
Give an example of (A, B) not reachable with state space in R^2.

Solution

$$A = I_2, B = (1, 0)^T$$

Then,

$$rank[B \quad AB] = rank \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 1 < 2$$

Exercise 5.13. ★
Give an example of (A, C) not observable with state space in R^2.

Solution

$$A = I_2, C = (1, 0)$$

Then,

$$rank[C^T \quad A^TC^T] = rank \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 1 < 2$$