Homework #5 Solutions

1. **Linear innovation sequence.** This comes from the example on lecture notes 4-21. Let the observation sequence be $Y_i = X + Z_i$ for $i = 1, 2, \ldots, n$, where $X, Z_1, \ldots, Z_n$ are zero mean, uncorrelated r.v.s with $E(X^2) = P$ and $E(Z_i^2) = N$ for $i = 1, 2, \ldots, n$. Find the linear innovation sequence of $Y$.

   **Solution** (10 points)

   The linear innovation sequence of $Y$ can be computed causally where
   \[
   \hat{Y}_{i+1} = Y_{i+1} - \hat{Y}_{i+1}(Y^i)
   \]
   Now we compute $\hat{Y}_{i+1}(Y^i) = \Sigma_{Y^i Y_{i+1}} \Sigma_{Y^i Y}^{-1} Y$. Since
   \[
   \Sigma_{Y^i Y_{i+1}} = \begin{bmatrix}
   P & \vdots \\
   \vdots & P \\
   \end{bmatrix}
   \]
   and
   \[
   \Sigma_{Y^i} = \begin{bmatrix}
   P + N & P & \cdots & P & P \\
   P & P + N & \cdots & P & P \\
   \vdots & \vdots & \ddots & \vdots & \vdots \\
   P & P & \cdots & P + N & P \\
   P & P & \cdots & P & P + N
   \end{bmatrix}
   \]
   we use the same argument on lecture notes 4-17 to obtain
   \[
   \hat{Y}_{i+1}(Y^i) = \frac{P}{iP + N} \sum_{k=1}^{i} Y_k
   \]
   Therefore, the linear innovation sequence of $Y$ is
   \[
   \hat{Y}_{i+1} = Y_{i+1} - \frac{P}{iP + N} \sum_{k=1}^{i} Y_k
   \]

2. **Cellphone.** We aim to design a cellphone which is able to denoise a signal modeled as $Y_1 = V + Z$, where $V$ is a random variable representing the user’s voice and $Z$ is a random variable representing background noise. An extra microphone measures the background. However this measurement also includes some distorted voice signal. This is taken into account by modeling it as $Y_2 = Z + U$, where $U$ is a random variable representing the distortion. Assume that $V, Z$ and $U$ are all zero mean, both $(V, Z)$ and $(U, Z)$ are uncorrelated and $\text{Corr}(U, V) = \rho$. We also know that $\text{Var}(V) = P$, $\text{Var}(Z) = N$ and $\text{Var}(U) = Q$. We decide to obtain a linear estimate of $V$ from $Y_1$ and $Y_2$.

   a. What is the innovation sequence $\hat{Y}_1$ and $\hat{Y}_2$ corresponding to $Y_1$ and $Y_2$?
b. What is the linear MMSE estimate of $V$ given the measurements expressed as a function of $\hat{Y}_1$, $\hat{Y}_2$, $P$, $Q$ and $N$?

c. What is the corresponding MSE in terms of $P$, $Q$ and $N$?

**Solution** (15 points)

a. The innovation sequence is equal to

\[
\hat{Y}_1 = Y_1
\]
\[
\hat{Y}_2 = Y_2 - \hat{Y}_2(Y_1) = Y_2 - \frac{\text{Cov}(Y_1, Y_2)}{\text{Var}(Y_1)} Y_1 = Y_2 - \frac{N + \rho \sqrt{PQ} N}{N + P} Y_1.
\]

b. The innovation sequence is uncorrelated by construction, so we can compute the linear MMSE estimates of $V$ given $\hat{Y}_1$ and $\hat{Y}_2$ separately and then add them up. It will be helpful to note that

\[
\hat{Y}_2 = \frac{P - \rho \sqrt{PQ}}{N + P} Z + U - \frac{N + \rho \sqrt{PQ}}{N + P} V.
\]
Using this we obtain
\[
\hat{V}(\tilde{Y}_1) = \frac{\text{Cov}(\tilde{Y}_1, V)}{\text{Var}(\tilde{Y}_1)} \tilde{Y}_1
= \frac{\text{Cov}(\tilde{Y}_1, V)}{\text{Var}(\tilde{Y}_1)} \tilde{Y}_1
= \frac{P}{N + P} \tilde{Y}_1,
\]
\[
\hat{V}(\tilde{Y}_2) = \frac{\text{Cov}(\tilde{Y}_2, V)}{\text{Var}(\tilde{Y}_2)} \tilde{Y}_2
= \frac{\rho - \frac{(N + \rho \sqrt{PQ})P}{N + P}}{(\frac{P - \rho \sqrt{PQ}}{N + P})^2 N + Q + (\frac{N + \rho \sqrt{PQ}}{N + P})^2 P - 2 \frac{N + \rho \sqrt{PQ}}{N + P} \rho \sqrt{PQ}} \tilde{Y}_2
= \frac{\rho(N + P)^2 - (N + \rho \sqrt{PQ})P(N + P)}{N(P - \rho \sqrt{PQ})^2 + Q(N + P)^2 + (N + \rho \sqrt{PQ})^2 P - 2(N + \rho \sqrt{PQ})(N + P) \rho \sqrt{PQ}} \tilde{Y}_2.
\]
The final estimate is
\[
\hat{V}(\tilde{Y}_1, \tilde{Y}_2) = \frac{P}{N + P} \tilde{Y}_1
- \frac{\rho(N + P)^2 - (N + \rho \sqrt{PQ})P(N + P)}{N(P - \rho \sqrt{PQ})^2 + Q(N + P)^2 + (N + \rho \sqrt{PQ})^2 P - 2(N + \rho \sqrt{PQ})(N + P) \rho \sqrt{PQ}} \tilde{Y}_2.
\]
c. The MSE is
\[
\text{MSE} = \frac{\text{Var}(V)}{\text{Var}(\tilde{Y}_1)} - \frac{\text{Cov}^2(\tilde{Y}_1, V)}{\text{Var}(\tilde{Y}_1)} - \frac{\text{Cov}^2(\tilde{Y}_2, V)}{\text{Var}(\tilde{Y}_2)}
= P - \frac{p^2}{N + P} - \frac{\rho(N + P)^2 - (N + \rho \sqrt{PQ})P(N + P)}{N(P - \rho \sqrt{PQ})^2 + Q(N + P)^2 + (N + \rho \sqrt{PQ})^2 P - 2(N + \rho \sqrt{PQ})(N + P) \rho \sqrt{PQ}} \tilde{Y}_2.
\]
3. Vector Kalman filter experiment. Consider the state space model
\[
X_{i+1} = \begin{bmatrix} 1 & 1 \\ 0 & 0.9 \end{bmatrix} X_i + U_i, \quad \text{for } i = 1, \ldots, n.
\]
The first and second component of \(X_i\) are the one-dimensional position and velocity of a moving object. In each step, the position and velocity evolve according to Newton’s laws of physics. Due to friction, the velocity is dampened by a factor of 0.9 in each time step. The state is disturbed by independent noise vectors \(U_i\), distributed according to
\[
U_i \sim \mathcal{N} \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} , \begin{bmatrix} 0 & 0 \\ 0 & 10 \end{bmatrix} \right), \quad \text{for } i = 1, \ldots, n.
\]
The initial state is independent of the noise vectors \( U_i \) and distributed according to

\[
X_1 \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1000 & 0 \\ 0 & 0 \end{bmatrix}\right)
\]

The observations are

\[
Y_i = X_i + V_i, \quad \text{for } i = 1, \ldots, n + 1,
\]

where the noise \( V_i \) is distributed according to

\[
V_1 \sim \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1000 & 0 \\ 0 & 1 \end{bmatrix}\right), \quad \text{for } i = 1, \ldots, n + 1,
\]

independent of the initial state \( X_1 \) and the state noise vectors \( U_i \).

Download the posted MATLAB file `vectorKalman.m` and complete the code to compute the Kalman prediction filter \( \hat{X}_{i+1|i} \) from the observations \( Y^i \), for \( i = 0, \ldots, n + 1 \). For a single realization with time horizon \( n = 100 \), plot the true state and its prediction over the time index \( i \). In a separate figure, plot the prediction error over \( i \). Hand in your MATLAB code and the plots.

**Solution** (10 points)

```matlab
function [Xhatp,SigmaXhatp] = vectorKalmanPredictor(A,P1,Q,N,Y) % inputs:
% A(sxsxn)-systemmatrixfortimes1:n
% P1 (s x s) - covariance matrix of state X_1
% Q (s x s x n) - covariance matrix of state noise at times 1:n
% N (s x s x n) - covariance matrix of observation noise at times 1:n
% Y (s x n) - observations for times 1:n % outputs:
% Xhatp (s x n+1) - predicted state.
% hatXp(:,i+1) is Xhat(i+1|i)
% hatXp(:,1) is Xhat(1|0)
% hatXp(:,2) is Xhat(2|1), etc.
% SigmaXhatp (s x s x n) - MSE matrix of the prediction
% SigmaXhatp(:,:,i+1) is Sigma(i+1|i)

s = size(A,1);
n = size(A,3);
Xhatp = zeros(s,n+1); SigmaXhatp = zeros(s,s,n+1); K = zeros(s,s,n);
SigmaXhatp(:,:,1) = P1;
for i=1:n
    K(:,:,i) = A(:,:,i)*SigmaXhatp(:,:,i)*inv(SigmaXhatp(:,:,i)+N(:,:,i));
    SigmaXhatp(:,:,i+1) = A(:,:,i)*SigmaXhatp(:,:,i)*(eye(s) - inv(SigmaXhatp(:,:,i)+N(:,:,i)))*A(:,:,i)' + Q(:,:,i);
    Xhatp(:,i+1) = A(:,:,i)*Xhatp(:,i) + K(:,:,i)*(Y(:,i) - Xhatp(:,i));
end
end
```

The plots generated by the code are shown in Figures 2 and 2.

4. *The filtering version of the Kalman filter* We have studied the derivation of scalar Kalman filter to predict the next state in class. Now we are interested in estimating the current state. Derive
Figure 2: Vector Kalman filter experiment: True and predicted state.

Figure 3: Vector Kalman filter experiment: Prediction error
the update equations and the error:
\[
\hat{X}_{i+1|i+1} = a_i(1 - k_i)\hat{X}_{i|i} + k_iY_{i+1}
\]
where
\[
k_i = \frac{a_i^2\sigma^2_{ii} + Q_i}{a_i^2\sigma^2_{ii} + Q_i + N_{i+1}}
\]
and
\[
\sigma^2_{i+1|i+1} = (1 - k_i)(a_i^2\sigma^2_{ii} + Q_i).
\]

**Solution** (15 points)

Assume we know the estimate \(\hat{X}_{i|i}\) and its MSE \(\sigma^2_{i|i}\). Now we receive a new observation \(Y_{i+1}\) and would like to compute the updated estimate \(\hat{X}_{i+1|i+1}\) and its MSE \(\sigma^2_{i+1|i+1}\).

First, we consider the estimate (prediction) of the state at time \(i + 1\) without new observations. We have
\[
\hat{X}_{i+1|i} = a_i\hat{X}_{i|i},
\]
\[
\sigma^2_{i+1|i} = a_i^2\sigma^2_{i|i} + Q_i.
\]

When we observe \(Y_{i+1}\), we are only interested in its innovation \(\tilde{Y}_{i+1}\), i.e., the part that was not predictable from the previous observations. It is given as
\[
\tilde{Y}_{i+1} = Y_{i+1} - \hat{Y}_{i+1|i}.
\]

The predictable part \(\hat{Y}_{i+1|i}\), in turn, is simply
\[
\hat{Y}_{i+1|i} = \hat{X}_{i+1|i} = a_i\hat{X}_{i|i},
\]
as computed above. The new state estimate is
\[
\hat{X}_{i+1|i+1} = \hat{X}_{i+1|i} + k_i\tilde{Y}_{i+1},
\]
because \(X_{i+1|i}\) is a function of \(Y^i\) and therefore orthogonal to \(\tilde{Y}_{i+1}\). Substituting, we obtain
\[
\hat{X}_{i+1|i+1} = a_i\hat{X}_{i|i} + k_i(Y_{i+1} - a_i\hat{X}_{i|i})
= a_i(1 - k_i)\hat{X}_{i|i} + k_iY_{i+1}.
\]

The coefficient \(k_i\) is computed as
\[
k_i = \frac{\text{Cov}(X_{i+1}, \tilde{Y}_{i+1})}{\text{Var}(\tilde{Y}_{i+1})},
\]
where
\[
\text{Cov}(X_{i+1}, \tilde{Y}_{i+1}) = \text{Cov}(a_iX_i + U_i, Y_{i+1} - a_i\hat{X}_{i|i})
= \text{Cov}(a_iX_i + U_i, a_iX_i + U_i + V_{i+1} - a_i\hat{X}_{i|i})
= a_i^2\text{Cov}(X_i, X_i - \hat{X}_{i|i}) + Q_i
= a_i^2\sigma^2_{i|i} + Q_i.
\]
and

\[ \text{Var}(\tilde{Y}_{i+1}) = \text{Var}(X_{i+1} + V_{i+1} - a_i \tilde{X}_{j|i}) \]
\[ = \text{Var}(aX_i + U_i + V_{i+1} - a_i \tilde{X}_{j|i}) \]
\[ = Q_i + N_{i+1} + a_i^2 \sigma_{ij|i}^2. \]

Substituting, we have

\[ k_i = \frac{a_i^2 \sigma_{ij|i}^2 + Q_i}{a_i^2 \sigma_{ij|i}^2 + Q_i + N_{i+1}}. \tag{2} \]

Finally, the updated MSE is

\[ \sigma_{i+1|i+1}^2 = \sigma_{i+1|i}^2 - \frac{\text{Cov}(X_{i+1}, \tilde{Y}_{i+1})^2}{\text{Var}(\tilde{Y}_{i+1})} \]
\[ = \sigma_{ij|i}^2 + Q_i - \frac{(a_i^2 \sigma_{ij|i}^2 + Q_i)^2}{a_i^2 \sigma_{ij|i}^2 + Q_i + N_{i+1}} \]
\[ = (1 - k_i) \left( a_i^2 \sigma_{ij|i}^2 + Q_i \right). \tag{3} \]

Equations (1), (2), and (3) were what we set out to show.