1. Consider the general case of the Kalman filter:

\[
X_0 \sim \mathcal{N}(0, \sigma_0^2) \\
X_n = \alpha X_{n-1} + W_{n-1} \quad n = 1, 2, \ldots \\
Y_n = X_n + Z_n, \quad n = 0, 1, 2, \ldots
\]

with the \(W_n\)’s i.i.d. \(\mathcal{N}(0, \sigma_w^2)\) random variables and \(Z_n\)’s i.i.d. \(\mathcal{N}(0, \sigma_z^2)\) random variables, all independent of each other and independent of \(X_0\).

Consider the Kalman filter estimates \(\{\hat{X}_n\}\), where \(\hat{X}_n\) is the MMSE estimate of the state \(X_n\) based on the observations \(Y_0, \ldots, Y_n\). We wonder in class whether \(\{\hat{X}_n\}\) is a Markov process.

a) Define what it means by a Markov process.

b) Given \(\hat{X}_{n-1}\), show that \(Y_n\) is independent of \(Y_{n-1}, Y_{n-2}, \ldots, Y_0\).

c) Given \(\hat{X}_{n-1}\), show that \(\hat{X}_n\) is independent of \(\hat{X}_{n-2}, \hat{X}_{n-3}, \ldots, \hat{X}_0\), using part (b) or otherwise.

d) Using part (c) or otherwise, conclude whether \(\{\hat{X}_n\}\) is a Markov process or not.

**Solution:**

a) A random process \(\{X_n\}\) is Markov iff for all \(n\), \(X_n\) is independent of \(X_{n-2}, X_{n-3}, \ldots\) given \(X_{n-1}\).

b) There is no additional calculation required to show this. Since \(\hat{X}_{n-1}\) is the MMSE estimate of \(X_{n-1}\) given \(Y_{n-1}, \ldots, Y_0\), we know that \(X_{n-1}\) is independent of \(Y_{n-1}, \ldots, Y_0\) given \(\hat{X}_{n-1}\). Moreover, \(Y_n = X_n + Z_n = \alpha X_{n-1} + W_n + Z_n\) and \(W_n\) and \(Z_n\) are independent of both \(\hat{X}_{n-1}\) and \(Y_{n-1}, \ldots, Y_0\). Therefore, \(Y_n\) is also independent of \(Y_{n-1}, \ldots, Y_0\) given \(\hat{X}_{n-1}\).

c) From the recursion equations, we have the following relations:

\[
\hat{X}_n = \beta_n \hat{X}_{n-1} + \theta_n Y_n = f(\hat{X}_{n-1}, Y_n).
\]
We also have
\[ \hat{X}_{n-1} = \beta_{n-1} \hat{X}_{n-2} + \theta_{n-1} Y_{n-1} \implies \hat{X}_{n-2} = g_1(\hat{X}_{n-1}, Y_{n-1}), \]  
(2)
\[ \hat{X}_{n-2} = \beta_{n-2} \hat{X}_{n-3} + \theta_{n-2} Y_{n-2} \implies \hat{X}_{n-3} = g_2(\hat{X}_{n-1}, Y_{n-1}, Y_{n-2}), \]  
(3)

Thus \( \hat{X}_{n-2}, ..., \hat{X}_0 \) are functions of \( \hat{X}_{n-1} \) and \( Y_{n-1}, ..., Y_0 \) and \( \hat{X}_n \) is a function of \( \hat{X}_{n-1}, Y_n \). Since \( Y_n \) and \( Y_{n-1}, ..., Y_0 \) are independent given \( \hat{X}_{n-1} \), we can conclude that \( \hat{X}_n \) and \( \hat{X}_{n-2}, ..., \hat{X}_0 \) are independent given \( \hat{X}_{n-1} \).

\textbf{Note:} The steps shown in (2),(3) assume that the \( \beta_n \)'s are nonzero. We know that this is true as long as \( \alpha \neq 0 \). However, if \( \alpha = 0 \), then the \( X_n \)'s are i.i.d., therefore \( Y_n \)'s are also i.i.d. Finally, \( \hat{X}_n \) depends only on \( Y_n \), so \( \hat{X}_n \)'s are also i.i.d. Clearly, an i.i.d random process \( \{\hat{X}_n\} \) satisfies that \( \hat{X}_n \) and \( \hat{X}_{n-2}, ..., \hat{X}_0 \) are independent given \( \hat{X}_{n-1} \).

d) In part (c), we have shown that \( \{\hat{X}_n\} \) satisfies the definition of a Markov process.
2. Are the following processes stationary? For each, compute the mean function, covariance
function \( K_X(n_1, n_2) \) and sketch a typical sample path. If a process is stationary, compute
also its power spectral density. Each process is defined on \( n = 0, 1, 2, \ldots \).

a) \( X_n = A \) where \( A \sim \mathcal{N}(0, 1) \).

b) \( X_n \) i.i.d. \( \mathcal{N}(0, 1) \) random variables.

c) \( X_n = nA \), where \( A \sim \mathcal{N}(0, 1) \).

d) \( X_n = (-1)^n \).

e) \( X_n = (-1)^{n+U} \), where \( U \) is equally likely to be 0 or 1.

Solution:

a) This process is constant across time, i.e. \( X_n = A \) for all \( n \), but \( A \) is random.
\( \mu_X(n) = 0 \), and \( K_X(n_1, n_2) = 1 \) for all \( n_1, n_2 \). For any collection \( [X_n, X_{n+1}, \ldots, X_{n+k}] \),
its joint distribution is specified by \( X_n \sim \mathcal{N}(0, 1), X_{n+1} = \ldots = X_{n+k} = X_n \). Hence
it is stationary. The covariance function can be written as \( K_X(m) = 1 \) for all \( m \).
Taking its Fourier transform, \( S_X(f) = \delta(f) \).

b) This is the white Gaussian noise process. \( \mu_X(n) = 0 \) and \( K_X(n_1, n_2) = 1 \) if \( n_1 = n_2 \)
or 0 otherwise. For any collection \( [X_n, X_{n+1}, \ldots, X_{n+k}] \), its joint distribution is i.i.d.
\( \mathcal{N}(0, 1) \). Therefore this process is stationary. The covariance function is \( K_X(m) = 1 \)
if \( m = 0 \) or 0 otherwise, i.e. \( K_X(m) = \delta_m \). Taking its Fourier transform, \( S_X(f) = 1 \).

c) \( \mu_X(n) = 0 \) and \( K_X(n_1, n_2) = n_1n_2 \). In particular, \( \text{Var}(X_n) = n^2 \). So this is not
stationary.

d) This process is deterministic, but alternates between +1 and −1. Since it is deter-
ministic, the mean is \( \mu_X(n) = (-1)^n \) and the covariance is \( K_X(n_1, n_2) = 0 \). This is
not stationary since the mean can be either +1 or −1, depending on \( n \).

e) This process has two possible sample paths – \( X_n = (-1)^n \) with prob. \( \frac{1}{2} \), and \( X_n = 1 \)
for all \( n \) with prob. \( \frac{1}{2} \). Since \( X_n = \pm 1 \) with equal prob. for each \( n \), \( \mu_X(n) = 0 \).
Also,
\[
K_X(n_1, n_2) = \mathbb{E}[X_1X_2] = \mathbb{E}[(-1)^{n_1+n_2+2U}] = (-1)^{n_1+n_2} = (-1)^{n_1-n_2}.
\]

The joint distribution of \( [X_n, X_{n+1}, \ldots, X_{n+k}] \) can be described as \( X_n = \pm 1 \) with
equal prob. and \( X_{n+i} = (-1)^iX_n \). Therefore, this process is stationary. The
covariance function is \( K_X(m) = (-1)^m = e^{j\pi m} \). By using the properties of Fourier
transform, \( S_X(f) = \delta(f - \frac{1}{2}) \).
3. Let $A_n$ and $B_n$ be two independent stationary processes with means $\mu_A$ and $\mu_B$, respectively, and with covariance functions $K_A(\cdot)$ and $K_B(\cdot)$ respectively. Is the process \( \{C_n = A_n + B_n\} \) stationary? If so, compute the mean, covariance function and power spectral density of \( \{C_n\} \) in terms of the corresponding quantities for \( \{A_n\} \) and \( \{B_n\} \).

**Solution:** For any $n, k$, the joint distribution of \([C_n, C_{n+1}, \ldots, C_{n+k}]\) can be obtained in terms of the joint distribution of \([A_n, A_{n+1}, \ldots, A_{n+k}]\) and the joint distribution of \([B_n, B_{n+1}, \ldots, B_{n+k}]\). If \( \{A_n\} \) and \( \{B_n\} \) are stationary, then these joint distributions are the same for all $n$, hence \( \{C_n\} \) is also stationary.

Due to linearity of expectation, \( \mu_C = \mu_A + \mu_B. \)

Let $\tilde{A}_n = A_n - \mu_A$ and $\tilde{B}_n = B_n - \mu_B$.

\[
K_C(m) = E[(C_0 - \mu_C)(C_m - \mu_C)] \\
= E[(\tilde{A}_0 + \tilde{B}_0)(\tilde{A}_m + \tilde{B}_m)] \\
= E[\tilde{A}_0 \tilde{A}_m] + E[\tilde{B}_0 \tilde{B}_m] \quad \text{due to independence} \\
= K_A(m) + K_B(m)
\]

Due to linearity of the Fourier transform,

\[
S_C(f) = S_A(f) + S_B(f).
\]
4. a) For Q. 2 and 3 in HW 7, compute the covariance functions and power spectral density for the processes \{X_n\}, \{Y_n\} and \{\hat{X}_n\} for the parameter values specified as in Q. 2(b) and 3(f) and for the 4 different values of \(\alpha\).

b) Comment on how the covariance functions you computed in part (a) depends on \(\alpha\). Are the dependencies consistent with your plots in Q. 2 and 3 in HW 7?

Solution:

a) i. For \(\alpha < 1\), these processes are stationary only in the steady state where the variance converges to a constant value. The system dynamics is given by

\[
X_{n+m} = \alpha X_{n+m-1} + W_{n+m-1} = \alpha (\alpha X_{n+m-2} + W_{n+m-2}) + W_{n+m-1} \\
= \alpha^2 X_{n+m-2} + \alpha W_{n+m-2} + W_{n+m-1} \\
= \ldots = \alpha^m X_n + \alpha^{m-1} W_n + \ldots + W_{n+m-1}.
\]

In this problem, we set \(\sigma_w^2 = 1 - \alpha^2\). At steady state, \(\sigma_n^2 = \sigma_w^2/(1 - \alpha^2) = 1\).

Since \(X_n\) is independent of \(W_n, \ldots, W_{n+m-1}\),

\[
K_X(n, n+m) = \mathbb{E}[X_n X_{n+m}] \\
= \alpha^m \mathbb{E}[X_n^2] \\
= \alpha^m
\]

Thus, \(K_X(m) = \alpha^{|m|}\). By taking the Fourier transform,

\[
S_X(f) = \frac{1 - \alpha^2}{1 + \alpha^2 - 2\alpha \cos(2\pi f)}
\]

Alternatively, recall from the solution to HW 7 Q. 2(c), that the impulse response of the LTI system from \(W_n\) to \(X_n\) is \(h_n = \alpha^{n-1} u_{n-1}\). Since \(W_n\) is an i.i.d. process, \(K_W(m) = \sigma_w^2 \delta_m = (1 - \alpha^2) \delta_n\). Then by passing \(W_n\) through this
LTI system,

\[ K_X(m) = (h * h_R * K_W)(m) \]
\[ = (1 - \alpha^2) (h * h_R)(m) \]
\[ = (1 - \alpha^2) \sum_{n=-\infty}^{\infty} h_n(h_R)_{m-n} \]
\[ = (1 - \alpha^2) \sum_{n=-\infty}^{\infty} \alpha^{n-1}u_{n-1} \alpha^{n-m-1}u_{n-m-1} \]
\[ = (1 - \alpha^2) \sum_{n=m+1}^{\infty} \alpha^{2n-m-2} \text{ for } m \geq 0 \]
\[ = (1 - \alpha^2) \frac{\alpha^m}{1 - \alpha^2} = \alpha^m, \]

\[ K_X(m) = \alpha^{|m|} \text{ for all } m \text{ by symmetry,} \]

\[ S_X(f) = |H(f)|^2S_W(f) \]
\[ = \left| \frac{e^{-j2\pi f}}{1 - \alpha e^{-j2\pi f}} \right|^2 (1 - \alpha^2) \]
\[ = \frac{1 - \alpha^2}{1 + \alpha^2 - 2\alpha \cos(2\pi f)} \]

ii. \( Y_n = X_n + Z_n \) where \( Z_n \) is independent of \( X_n \). \( Z_n \) is i.i.d. with \( \sigma_z^2 = 1 \), so \( K_Z(m) = \delta_m \). Using the answer from Q. 3 above,

\[ K_Y(m) = K_X(m) + K_Z(m) \]
\[ = \alpha^{|m|} + \delta_m, \]

\[ S_Y(f) = S_X(f) + S_Z(f) \]
\[ = \frac{1 - \alpha^2}{1 + \alpha^2 - 2\alpha \cos(2\pi f)} + 1. \]

iii. Recall from the solution to Q. 3(c) in HW 7, that \( v_n^2 \) converges to a value \( v_\infty^2 \) which satisfies

\[ \alpha^2 v_\infty^4 + 2(1 - \alpha^2)v_\infty^2 - (1 - \alpha^2) = 0 \]
\[ \implies v_\infty^2 = \frac{\sqrt{1 - \alpha^2} - (1 - \alpha^2)}{\alpha^2}. \]

The steady state recursion is given by

\[ \hat{X}_n = \alpha(1 - \beta_\infty)\hat{X}_{n-1} + \beta_\infty Y_n \]
Thus, by substituting the value of $\beta_\infty = \frac{\alpha^2 v_\infty^2 + 1 - \alpha^2}{\alpha^2 v_\infty^2 + 1 - \alpha^2 + 1} = \frac{\sqrt{1 - \alpha^2}}{1 + \sqrt{1 - \alpha^2}}$.

Expanding the recursion, we have

$$
\hat{X}_{n+m} = \alpha (1 - \beta_\infty) \hat{X}_{n+m-1} + \beta_\infty Y_{n+m} \\
= \alpha (1 - \beta_\infty) \left( \alpha (1 - \beta_\infty) \hat{X}_{n+m-2} + \beta_\infty Y_{n+m-1} \right) + \beta_\infty Y_{n+m} \\
= \alpha^2 (1 - \beta_\infty)^2 \hat{X}_{n+m-2} + \alpha (1 - \beta_\infty) \beta_\infty Y_{n+m-1} + \beta_\infty Y_{n+m} \\
= \cdots = \alpha^m (1 - \beta_\infty)^m \hat{X}_n + \alpha^{m-1} (1 - \beta_\infty)^{m-1} \beta_\infty Y_{n+1} + \cdots + \beta_\infty Y_{n+m}.
$$

Thus,

$$
K_X(n, n + m) = \mathbb{E}[\hat{X}_n \hat{X}_{n+m}] \\
= \alpha^m (1 - \beta_\infty)^m \mathbb{E}[\hat{X}_n^2] + \alpha^{m-1} (1 - \beta_\infty)^{m-1} \beta_\infty \mathbb{E}[\hat{X}_n Y_{n+1}] + \cdots + \beta_\infty \mathbb{E}[\hat{X}_n Y_{n+m}] \\
= \alpha^m (1 - \beta_\infty)^m \mathbb{E}[\hat{X}_n^2] + \alpha^{m-1} (1 - \beta_\infty)^{m-1} \beta_\infty \alpha \mathbb{E}[-\hat{X}_n^2] + \cdots + \beta_\infty \alpha^m \mathbb{E}[\hat{X}_n^2] \\
= \alpha^m (1 - \beta_\infty)^m \mathbb{E}[\hat{X}_n^2] + \alpha^{m-1} \beta_\infty \frac{1 - (1 - \beta_\infty)^m}{1 - (1 - \beta_\infty)} \mathbb{E}[\hat{X}_n^2] \\
= \alpha^m (1 - v_\infty^2).
$$

The third line above is because

$$
\mathbb{E}[Y_{n+i} \hat{X}_n] = \mathbb{E} \left[ \alpha^i X_n \hat{X}_n \right] = \mathbb{E} \left[ \alpha^i \hat{X}_n \right] = \alpha^i \mathbb{E} \left[ \hat{X}_n^2 \right].
$$

Thus, $K_X(m) = \alpha^{|m|}(1 - v_\infty^2)$. Taking the Fourier transform, we get

$$
S_X(f) = \frac{(1 - \alpha^2)(1 - v_\infty^2)}{1 + \alpha^2 - 2 \alpha \cos(2\pi f)}.
$$

However, the easier way to solve this is using the impulse response of the LTI system from $Y_n$ to $X_n$, $h_m = \beta_\infty \alpha^m (1 - \beta_\infty)^m u_m$. We have already calculated that $K_Y(m) = \alpha^{|m|} + \delta_m$. Then by passing $Y_n$ through this LTI system,

$$
S_X(f) = |H(f)|^2 S_Y(f) \\
= \frac{\beta_\infty}{1 - \alpha (1 - \beta_\infty) e^{-j2\pi f}} \left( \frac{1 - \alpha^2}{1 + \alpha^2 - 2 \alpha \cos(2\pi f)} + 1 \right) \\
= \frac{\beta_\infty^2}{1 + \alpha^2 (1 - \beta_\infty)^2 - 2 \alpha (1 - \beta_\infty) \cos(2\pi f)} \cdot \frac{1}{1 + \alpha^2 - 2 \alpha \cos(2\pi f)}.
$$

By substituting the value of $\beta_\infty$,

$$
S_X(f) = \frac{1 - \alpha^2}{(1 + \sqrt{1 - \alpha^2})^2 + \alpha^2 - 2 \alpha (1 + \sqrt{1 - \alpha^2}) \cos(2\pi f)} \cdot \frac{2 - 2 \alpha \cos(2\pi f)}{1 + \alpha^2 - 2 \alpha \cos(2\pi f)} \\
= \frac{1 - \alpha^2}{(1 + \frac{1 - \alpha^2}{2 \alpha \cos(2\pi f)}) (2 - 2 \alpha \cos(2\pi f))} \cdot \frac{1}{1 + \alpha^2 - 2 \alpha \cos(2\pi f)} \\
= \frac{1}{1 + \sqrt{1 - \alpha^2}} \cdot \frac{1 - \alpha^2}{1 + \alpha^2 - 2 \alpha \cos(2\pi f)}.
$$
We can then take the inverse Fourier transform to get

\[ K_X(m) = \frac{\alpha^{|m|}}{1 + \sqrt{1 - \alpha^2}} = (1 - v^2_{\infty})\alpha^{|m|} \]

**Note:** Useful Fourier transforms used in this solution:

\[ x(m) = a^m u[m] \quad \leftrightarrow \quad X(f) = \frac{1}{1 - ae^{-j2\pi f}} \quad \text{for } |a| < 1 \]

\[ x(m) = a^{|m|} \quad \leftrightarrow \quad X(f) = \frac{1 - a^2}{1 + a^2 - 2a \cos(2\pi f)} \quad \text{for } |a| < 1 \]

\[ x(m) = \delta[m] \quad \leftrightarrow \quad X(f) = 1 \]

\[ y(m) = x(m - m_0) \quad \leftrightarrow \quad Y(f) = X(f) e^{-j2\pi fm_0} \]

b) The covariance functions decay as \( \alpha^{|m|} \). As \( \alpha \) becomes close to 0, the covariance function becomes an impulse, so adjacent values are fully uncorrelated. As \( \alpha \) becomes close to 1, the covariance functions become a constant, so there is very high correlation between adjacent samples.

In the plots in Q. 2 and Q. 3 of HW 7, adjacent values of \( X_n \) and of \( \hat{X}_n \) are highly uncorrelated when \( \alpha \) is small and much more correlated when \( \alpha \) is close to 1. Hence our conclusions are consistent with those plots.
5. a) Suppose there is a zero-mean stationary random process \( \{ X_n : n \geq 0 \} \) but you don’t know its covariance function \( K_X(m) \). You observe a sample path of the process up to time \( N \). Find a reasonable estimator \( \hat{K}_X^{(N)}(m) \) of the covariance function \( K_X(m) \) for every \( m \).

b) Extend part (a) to the non-zero mean case.

c) The Madden Julian oscillation (MJO) is the largest element of the intraseasonal variability in the tropical atmosphere. It is a traveling pattern that propagates eastward, and each cycle lasts approximately 30-60 days. The MJO is most obvious in the variation of outgoing longwave radiation (OLR), a satellite-derived measure of tropical convection and rainfall.

The data package at [http://ee278.stanford.edu/homeworks/hw8_data.mat](http://ee278.stanford.edu/homeworks/hw8_data.mat) consists of one time series \( \{ X_n \} \) of OLR in the Indian ocean that contains 1752 points. Each point in the time series records a 5-day average of OLR over the region specified, from the beginning of 1979 through the end of 2002. Suppose that the time series is generated by a stationary process.

Estimate the covariance function of the time series and plot it for time lags 0 through 30. Describe briefly on any cyclic characteristics that the covariance function might present. What can you say about the periodicity of the MJO?

**Solution:**

a) You can estimate the covariance function as

\[
\hat{K}_X^{(N)}(m) = \frac{1}{N-m} \sum_{i=1}^{N-m} X_i X_{i+m}
\]

for every \( m \geq 0 \). For \( m < 0 \), you can set \( \hat{K}_X^{(N)}(m) = \hat{K}_X^{(N)}(-m) \).

b) First estimate the mean of the process as

\[
\bar{X}^{(N)} = \frac{1}{N} \sum_{i=1}^{N} X_i
\]

Then estimate the covariance function as

\[
\hat{K}_X^{(N)}(m) = \frac{1}{N-m} \sum_{i=1}^{N-m} \left( X_i - \bar{X}^{(N)} \right) \left( X_{i+m} - \bar{X}^{(N)} \right)
\]

for every \( m \geq 0 \). For \( m < 0 \), you can set \( \hat{K}_X^{(N)}(m) = \hat{K}_X^{(N)}(-m) \).

c) We can estimate the covariance function using the estimator defined above. The estimates are plotted in Fig. I. From the plot, we can see that the fluctuation of the covariance function seems to exhibit some periodicity, i.e. the period of peaks
is approximately 8 - 10 points, corresponding to the cycle of 40 - 50 days in the MJO data.

Practically, the uncertainty and noise often make it difficult to spot oscillatory behavior in a random signal, even if such behavior is expected. The covariance function of a random periodic signal has the same cyclic characteristics as the signal itself, and can be estimated in a more robust fashion. Thus, the covariance function is a more robust measure for detecting the presence of periodicity and determining their durations.