EN 257: Applied Stochastic Processes
Problem Set 8

Douglas Lanman
dlanman@brown.edu
9 May 2007

Problem 8.2

Let $X(t)$ be a random process with constant mean $\mu_X \neq 0$ and covariance function

$$K_{XX}(t_1, t_2) = \sigma^2 \cos(\omega_0(t_1 - t_2)).$$

(a) Show that the mean-square (m.s.) derivative $X'(t)$ exists.
(b) Find the correlation function of the m.s. derivative $R_{X'X'}(t_1, t_2)$.
(c) Find the covariance function of the m.s. derivative $K_{X'X'}(t_1, t_2)$.

Part (a)

Recall, from Theorem 8.1-2 on page 490 in [4], that a random process $X(t)$ with autocorrelation function $R_{XX}(t_1, t_2)$ has a m.s. derivative at time t if $\partial^2 R_{XX}(t_1, t_2)/\partial t_1 \partial t_2$ exists at $t_1 = t_2 = t$. Furthermore, we recall that the correlation function $R_{XX}(t_1, t_2)$ is related to the covariance function $K_{XX}(t_1, t_2)$ as follows.

$$R_{XX}(t_1, t_2) = K_{XX}(t_1, t_2) + \mu_X(t_1)\mu_X^*(t_2) \tag{1}$$

For this problem we have a constant mean function $\mu_X \neq 0$ such that

$$R_{XX}(t_1, t_2) = \sigma^2 \cos(\omega_0(t_1 - t_2)) + \mu_X \mu_X^*.$$

Evaluating the mixed partial derivative of $R_{XX}(t_1, t_2)$ at $t_1 = t_2 = t$ yields the following result.

$$\left.\frac{\partial^2 R_{XX}(t_1, t_2)}{\partial t_1 \partial t_2}\right|_{(t_1,t_2)=(t,t)} = \sigma^2 \left\{ \left.\frac{\partial^2 \cos(\omega_0(t_1 - t_2))}{\partial t_1 \partial t_2}\right|_{(t_1,t_2)=(t,t)} \right\} = \sigma^2 \omega_0^2$$

By Theorem 8.1-2, the m.s. derivative $X'(t)$ exists since $\partial^2 R_{XX}(t_1, t_2)/\partial t_1 \partial t_2$ exists at $t_1 = t_2 = t$.

Part (b)

Recall, by Theorem 8.1-3 on page 494 in [4], that if a random process $X(t)$ with mean function $\mu_X(t)$ and correlation function $R_{XX}(t_1, t_2)$ has a m.s. derivative $X'(t)$, then the mean and correlation functions of $X'(t)$ are given by

$$\mu_{X'}(t) = \frac{d\mu_X(t)}{dt} \tag{2}$$

and

$$R_{X'X'}(t_1, t_2) = \frac{\partial^2 R_{XX}(t_1, t_2)}{\partial t_1 \partial t_2}. \tag{3}$$

As a result, we find that the correlation function $R_{X'X'}(t_1, t_2)$ is given by the following expression.

$$\boxed{R_{X'X'}(t_1, t_2) = \sigma^2 \omega_0^2 \cos(\omega_0(t_1 - t_2))}$$

Part (c)

From Equation 1 we conclude that the covariance function of the m.s. derivative $X'(t)$ is given by

$$K_{X'X'}(t_1, t_2) = R_{X'X'}(t_1, t_2) - \mu_{X'}(t_1)\mu^*_{X'}(t_2).$$

From Equation 2 we find that $\mu_{X'}(t) = \mu^*_{X'}(t) = 0$, since $X(t)$ has a constant mean $\mu_X \neq 0$. As a result, we conclude that the covariance function $K_{X'X'}(t_1, t_2)$ is equal to the correlation function $R_{X'X'}(t_1, t_2)$ for this example.

$$\boxed{K_{X'X'}(t_1, t_2) = \sigma^2\omega_0^2 \cos(\omega_0(t_1 - t_2))}$$

Problem 8.3

Let the random process $X(t)$ be wide-sense stationary with correlation function

$$R_{XX}(\tau) = \sigma^2 e^{-(\tau/T)^2}.$$

Let $Y(t) = 3X(t) + 2X'(t)$, where the derivative is interpreted in the mean-square sense.

(a) State conditions for the m.s. existence of $Y(t)$ in terms of $R_{XX}(\tau)$.

(b) Find the correlation function $R_{YY}(\tau)$ for the given $R_{XX}(\tau)$ in terms of σ^2 and T.

Part (a)

To begin our analysis we recall, from Theorem 8.1-4, that the m.s. derivative $X'(t)$ of a WSS random process $X(t)$ exists at time t if the autocorrelation function $R_{XX}(\tau)$ has up to second order derivatives at $\tau = 0$. In addition, from Definition 8.1-1 and Theorem 8.1-1, we note that the random process $Y(t)$ is m.s. continuous if

$$\lim_{\varepsilon \to 0} E\{|Y(t+\varepsilon) - Y(t)|^2\} = 0, \ \forall t.$$

Expanding this expression, we find

$$E\{|Y(t+\varepsilon) - Y(t)|^2\} = R_{YY}(t+\varepsilon, t+\varepsilon) - R_{YY}(t, t+\varepsilon) - R_{YY}(t+\varepsilon, t) + R_{YY}(t, t).$$

As a result, we find the $Y(t)$ exists (and is m.s. continuous) if $R_{YY}(t_1, t_2)$ is continuous at $t_1 = t_2 = t$. As we'll show in Part (b), the correlation function $R_{YY}(t_1, t_2)$ is WSS and has the following general form for $Y(t) = 3X(t) + 2X'(t)$.

$$R_{YY}(\tau) = 9R_{XX}(\tau) + 4R_{X'X'}(\tau)$$

By Corollary 8.1-1, we note that the WSS random process $X(t)$ is m.s. continuous if $R_{XX}(\tau)$ is continuous at $\tau = 0$. In conclusion, we can combine the previous results to arrive at the following conditions for the m.s. existence of $Y(t)$ in terms of $R_{XX}(\tau)$.

(1) $R_{XX}(\tau)$ must be continuous at $\tau = 0$.
(2) The derivatives of $R_{XX}(\tau)$ must exist up to second order.

Part (b)

Let's begin by expanding the expression for the correlation function $R_{YY}(\tau)$.

$$\begin{aligned}
R_{YY}(\tau) &= E\{Y(t+\tau)Y^*(t)\} \\
&= E\left\{\left[3X(t+\tau)+2X'(t+\tau)\right]\left[3X(t)+2X'(t)\right]^*\right\} \\
&= 9E\{X(t+\tau)X^*(t)\} + 6E\left\{X(t+\tau)X'^*(t)\right\} + 6E\left\{X'(t+\tau)X^*(t)\right\} + 4E\left\{X'(t+\tau)X'^*(t)\right\} \\
&= 9R_{XX}(\tau) + 6R_{XX'}(\tau) + 6R_{X'X}(\tau) + 4R_{X'X'}(\tau)
\end{aligned}$$

Recall from Equations 8.1-10 and 8.1-11 that, for a WSS random process $X(t)$, the cross-correlation functions are given by the following expressions.

$$R_{X'X}(\tau) = +\frac{dR_{XX}(\tau)}{d\tau} \quad \text{and} \quad R_{XX'}(\tau) = -\frac{dR_{XX}(\tau)}{d\tau}$$

In addition, we recall from Theorem 8.1-4 that the correlation function for the m.s. derivative $X'(t)$ is given by

$$R_{X'X'}(\tau) = -\frac{d^2R_{XX}(\tau)}{d\tau^2}$$

Substituting into the previous expression for $R_{YY}(\tau)$, we find that the correlation function for $Y(t)$ is given by the following expression.

$$\begin{aligned}
R_{YY}(\tau) &= 9R_{XX}(\tau) + 4R_{X'X'}(\tau) \\
&= 9R_{XX}(\tau) - 4\left\{\frac{d^2R_{XX}(\tau)}{d\tau^2}\right\} \\
&= 9\sigma^2 e^{-(\tau/T)^2} - 4\left\{\sigma^2\left(\frac{4\tau^2-2T^2}{T^4}\right)e^{-(\tau/T)^2}\right\}
\end{aligned}$$

In conclusion, the correlation function $R_{YY}(\tau)$ has the following solution.

$$\boxed{R_{YY}(\tau) = \sigma^2\left(\frac{9T^4-16\tau^2+8T^2}{T^4}\right)e^{-(\tau/T)^2}}$$

Problem 8.7

To estimate the mean of a stationary random process $X(t)$, we often consider an integral average

$$I(T) \triangleq \frac{1}{T}\int_0^T X(t)dt,\ T > 0.$$

(a) Find the mean of $I(T)$, denoted $\mu_I(T)$, in terms of the mean μ_X for $T > 0$.

(b) Find the variance of $I(T)$, denoted $\sigma_I^2(T)$, in terms of the covariance $K_{XX}(\tau)$ for $T > 0$.

Part (a)

The mean function $\mu_I(T)$ of the integral average $I(T)$ is given by the following expression.

$$\mu_I(T) = E\{I(t)\} = E\left\{\frac{1}{T}\int_0^T X(t)dt\right\} = \frac{1}{T}\int_0^T E\{X(t)\}dt = \frac{1}{T}\int_0^T \mu_X dt = \mu_X$$

Note that in the previous expression we have applied the linearity property of the expectation operator, as well as the condition that $E\{X(t)\} = \mu_X$ for a stationary random process $X(t)$. In conclusion, the mean function $\mu_I(T)$ is equal to μ_X – which implies that $X(t)$ is *ergodic in the mean* such that the time average equals the ensemble average.

$$\boxed{\mu_I(T) = \mu_X,\ \text{for } T > 0}$$

Part (b)

The variance function $\sigma_I^2(T)$ of the integral average $I(T)$ is given by the following expression.

$$\begin{aligned}
\sigma_I^2(T) &= E\left\{[I(T) - \mu_I(T)]^2\right\} = E\left\{[I(T) - \mu_I(T)]\,[I(T) - \mu_I(T)]^*\right\} \\
&= E\left\{\left[\frac{1}{T}\int_0^T X(t_1)dt_1 - \mu_X\right]\left[\frac{1}{T}\int_0^T X^*(t_2)dt_2 - \mu_X^*\right]\right\} \\
&= E\left\{\left[\frac{1}{T}\int_0^T (X(t_1) - \mu_X)\,dt_1\right]\left[\frac{1}{T}\int_0^T (X(t_2) - \mu_X)^*\,dt_2\right]\right\} \\
&= \frac{1}{T^2}\int_0^T\int_0^T E\left\{[X(t_1) - \mu_X]\,[X(t_2) - \mu_X]^*\right\} dt_1 dt_2 \\
&= \frac{1}{T^2}\int_0^T\int_0^T K_{XX}(t_1, t_2)dt_1 dt_2
\end{aligned}$$

Note that we have applied the linearity of the expectation operator, as well as the condition that $K_{XX}(t_1, t_2) = E\{[X(t_1) - \mu_X]\,[X(t_2) - \mu_X]^*\}$ for a stationary random process $X(t)$. Furthermore, we recall that for a stationary random process the covariance function is only a function of the time shift $\tau = t_1 - t_2$ such that $K_{XX}(t_1, t_2) = K_{XX}(t_1 - t_2)$. In conclusion, the variance function $\sigma_I^2(T)$ is equal to the following expression in terms of the covariance function $K_{XX}(\tau)$.

$$\boxed{\sigma_I^2(T) = \frac{1}{T^2}\int_0^T\int_0^T K_{XX}(t_1 - t_2)dt_1 dt_2,\ \text{for } T > 0}$$

Problem 8.12

This problem concerns the mean-square derivative. Let the random process $X(t)$ be second order (i.e., $E\{|X(t)|^2\} < \infty$) with correlation function $R_{XX}(t_1, t_2)$. Let the random process $Y(t)$ be defined by the mean-square integral

$$Y(t) \triangleq \int_{-\infty}^{t} e^{-(t-s)} X(s) ds. \tag{4}$$

(a) State the condition for the existence of the m.s. integral $Y(t)$ in terms of $R_{XX}(t_1, t_2)$.

(b) Find the correlation function $R_{YY}(t_1, t_2)$ of $Y(t)$ in terms of $R_{XX}(t_1, t_2)$.

(c) Determine the condition on $R_{XX}(t_1, t_2)$ for the existence of the m.s. derivative $dY(t)/dt$.

Part (a)

Note that Equation 4 defines a weighted mean-square integral of the form

$$I \triangleq \int_{T_1}^{T_2} h(t) X(t) dt,$$

where $h(t) = e^{-(T_2 - t)}$ is the specific weighting function and $(T_1, T_2) = (-\infty, t)$. From pages 503 and 505 of [4], we recall that the weighted mean-square integral I is defined by

$$\lim_{n \to \infty} E\left\{ \left| I - \sum_{i=1}^{n} h(t_i) X(t_i) \Delta t_i \right|^2 \right\} = 0, \tag{5}$$

where the integral I is approximated by the following summation.

$$I_n \triangleq \sum_{i=1}^{n} h(t_i) X(t_i) \Delta t_i, \text{ for } \Delta t_i = (T_2 - T_1)/n$$

At this point we can apply the Cauchy criterion to determine the necessary conditions for the existence of the m.s. integral.

$$\lim_{m,n \to \infty} E\{|I_n - I_m|^2\} = 0$$

Expanding this expression yields the following condition for convergence.

$$\lim_{m,n \to \infty} E\{I_n I_n^*\} - 2\mathrm{Re}\left(E\{I_n I_m^*\}\right) + E\{I_m I_m^*\} = 0 \tag{6}$$

Focusing on the cross-term, we find the following result.

$$\begin{aligned} E\{I_n I_m^*\} &= \sum_{i=1}^{n} \sum_{j=1}^{m} h(t_i) h^*(t_j) E\{X(t_i) X^*(t_j)\} \Delta t_i \Delta t_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{m} h(t_i) h^*(t_j) R_{XX}(t_i, t_j) \Delta t_i \Delta t_j \end{aligned}$$

As a result, we conclude that the m.s. integral of $Y(t)$ will exist if and only if

$$\int_{-\infty}^{t_1} \int_{-\infty}^{t_2} e^{-(t_1 - s_1)} e^{-(t_2 - s_2)} R_{XX}(s_1, s_2) ds_1 ds_2$$

exists in the ordinary calculus. If this integral exists, then Equation 6 is satisfied and, via the Cauchy criterion, the weighted mean-square integral I must satisfy Equation 5.

Part (b)

The correlation function can be found by direct evaluation as follows.

$$
\begin{aligned}
R_{YY}(t_1,t_2) &= E\left\{Y(t_1)Y^*(t_2)\right\} \\
&= E\left\{\left[\int_{-\infty}^{t_1} e^{-(t_1-s_1)}X(s_1)ds_1\right]\left[\int_{-\infty}^{t_2} e^{-(t_2-s_2)}X(s_2)ds_2\right]^*\right\} \\
&= \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} e^{-(t_1-s_1)}e^{-(t_2-s_2)}E\left\{X(s_1)X^*(s_2)\right\}ds_1ds_2
\end{aligned}
$$

Since the correlation function of $X(t)$ satisfies $R_{XX}(t_1,t_2) = E\left\{X(t_1)X^*(t_2)\right\}$, we conclude that $R_{YY}(t_1,t_2)$ has the following solution.

$$
\boxed{R_{YY}(t_1,t_2) = \int_{-\infty}^{t_1}\int_{-\infty}^{t_2} e^{-(t_1-s_1)}e^{-(t_2-s_2)}R_{XX}(s_1,s_2)ds_1ds_2}
$$

Part (c)

From page 506 in [4], we recognize that the solution $Y(t)$ to the stochastic differential equation

$$
dY(t)/dt = X(t)
$$

is given by

$$
Y(t) = \int_{t_0}^{t} X(s)ds + Y(t_0), \text{ for } t \geq t_0.
$$

As a result, we note that the m.s. derivative $dY(t)/dt$ will exist if the weighted integral in Equation 4 exists and is bounded. From Equation 8.2-6 we recall that the following condition of the weighting kernel $h(t,s) = e^{-(t-s)}$ is required.

$$
\int_{-\infty}^{t} |e^{-(t-s)}|ds < \infty
$$

As before, this generalizes to a m.s. stochastic integral involving the correlation function $R_{XX}(t_1,t_2)$. In conclusion, we find that the m.s. derivative $dY(t)/dt$ will exist if and only if

$$
\int_{-\infty}^{t_1}\int_{-\infty}^{t_2} e^{-(t_1-s_1)}e^{-(t_2-s_2)}R_{XX}(s_1,s_2)ds_1ds_2
$$

exists in the ordinary calculus.

Problem 8.17

Consider the m.s. differential equation

$$\frac{dY(t)}{dt} + 2Y(t) = X(t),$$

for $t \geq 0$, subject to the initial condition $Y(0) = 0$. Let the input be given by

$$X(t) = 5\cos(2t) + W(t), \tag{7}$$

where $W(t)$ is a mean-zero Gaussian noise process with covariance function $K_{WW}(\tau) = \sigma^2\delta(\tau)$.

(a) Find the mean function $\mu_Y(t)$ for $t \geq 0$.

(b) Find the covariance function $K_{YY}(t_1, t_2)$ for $t_1 \geq 0$ and $t_2 \geq 0$.

(c) What is the maximum value of σ such that $P[|Y(t) - \mu_Y(t)| < 0.1] > 0.99$, for all $t > 0$?.

Part (a)

Note that in the following analysis we will follow the general approach outlined in Example 8.3-1. Let's begin by taking the expectation of both sides of Equation 7.

$$\frac{dE\{Y(t)\}}{dt} + 2E\{Y(t)\} = E\{X(t)\}, \text{ for } E\{Y(0)\} = 0 \text{ and } t \geq 0$$

$$\Rightarrow \mu_Y'(t) + 2\mu_Y(t) = \mu_X(t) = 5\cos(2t), \text{ for } \mu_Y(0) = 0 \text{ and } t \geq 0$$

In conclusion, the solution to this ordinary differential equation is given by the following expression.

$$\boxed{\mu_Y(t) = \frac{5}{4}\left(\cos(2t) + \sin(2t) - e^{-2t}\right), \text{ for } t \geq 0}$$

Part (b)

For brevity, we recall that the derivation of the covariance function $K_{YY}(t_1, t_2)$ is presented on pages 506-511 in [4]. From that section we recall that the following expression defines the cross-covariance function $K_{XY}(t_1, t_2)$ for $t_1 \geq 0$ and $t_2 \geq 0$.

$$\frac{\partial K_{XY}(t_1, t_2)}{\partial t_2} + 2K_{XY}(t_1, t_2) = K_{XX}(t_1, t_2) = \sigma^2\delta(t_1 - t_2), \text{ for } K_{XY}(t_1, 0) = 0$$

Note that the initial condition is given by $K_{XY}(t_1, 0) = 0$ since $Y(0) = 0$. Also recognize that the covariance of the input function is given by $K_{XX}(t_1, t_2) = \sigma^2\delta(t_1 - t_2)$. As described in Example 8.3-1, this ordinary differential equation has the following solution.

$$K_{XY}(t_1, t_2) = \begin{cases} 0, & \text{for } 0 \leq t_2 < t_1, \\ \sigma^2 e^{-2(t_2 - t_1)}, & \text{for } t_2 \geq t_1 \end{cases}$$

Continuing with our analysis, we recall that Equation 8.3-4 yields the following expression for the output covariance $K_{YY}(t_1, t_2)$ in terms of the cross-covariance $K_{XY}(t_1, t_2)$.

$$\frac{\partial K_{YY}(t_1, t_2)}{\partial t_1} + 2K_{YY}(t_1, t_2) = K_{XY}(t_1, t_2), \text{ for } K_{YY}(0, t_2) = 0$$

In conclusion, we find that covariance function $K_{YY}(t_1, t_2)$ is given by the following expression.

$$\boxed{K_{YY}(t_1,t_2) = \begin{cases} \frac{\sigma^2}{4} e^{-2t_2}\left(e^{2t_1} - e^{-2t_1}\right), & \text{for } 0 < t_1 \leq t_2, \\ \frac{\sigma^2}{4}\left(1 - e^{-4t_2}\right) e^{-2(t_1 - t_2)}, & \text{for } t_1 \geq t_2 \end{cases}}$$

Part (c)

As discussed on page 511, the random process $Y(t)$ has *asymptotic wide-wense stationarity* such that covariance $K_{YY}(t_1, t_2)$ tends to the constant $\sigma^2/4$ as t_1 and t_2 become large. As a result, let's assume that the random process $Y(t) - \mu_Y(t)$ is modeled by a white Gaussian random process noise with mean zero and variance $\sigma^2/4$. Under these circumstances we find that the maximum value of σ can be found using the following constraint.

$$P[|Y(t) - \mu_Y(t)| < 0.1] = P[-0.1 < Y(t) - \mu_Y(t) < 0.1] > 0.99$$

$$\Rightarrow \frac{2}{\sqrt{2\pi\sigma^2}} \int_{-0.1}^{0.1} \exp\left(\frac{-2x^2}{\sigma^2}\right) dx > 0.99$$

Recall that the error function has the following definition.

$$\text{erf}(z) \triangleq \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

Substituting this expression in the previous result yields the following condition.

$$\text{erf}\left(\frac{1}{5\sqrt{2}\sigma}\right) > 0.99$$

In conclusion, σ must satisfy the following inequality in order for $P[|Y(t) - \mu_Y(t)| < 0.1] > 0.99$, for all $t > 0$.

$$\boxed{\sigma < 0.07765}$$

Problem 8.22

To detect a constant signal of amplitude A in white Gaussian noise of variance σ^2 and mean zero, we consider two hypotheses (i.e., events):

$$\left.\begin{array}{ll} H_0: & R(t) = W(t) \\ H_1: & R(t) = A + W(t) \end{array}\right\} \text{ for } t \in [0, T].$$

It can be shown that the optimal detector, to decide between hypotheses, first computes the integral

$$\Lambda \triangleq \int_0^T R(t)dt$$

and then performs a threshold test.

(a) Find the mean value of the integral Λ under each hypothesis.

(b) Find the variance of Λ under each hypothesis.

(c) An optimal detector would compare Λ to the threshold $\Lambda_0 \triangleq AT/2$ when each hypothesis is equally likely (i.e., $P[H_0] = P[H_1] = 1/2$). Under these conditions, find $P[\Lambda \geq \Lambda_0 | H_0]$ and express your result in terms of the error function.

Part (a)

Let's begin by evaluating the mean value of the integral Λ under hypothesis H_0.

$$\mu_{\Lambda|H_0}(T) = E\left\{\int_0^T W(t)dt\right\} = \int_0^T E\{W(t)\}dt = 0$$

Note that we have applied the linearity property of the expectation operator, as well as the mean-zero condition $E\{W(t)\} = 0$ for the white Gaussian noise process. Similarly, the mean value of the integral Λ under hypothesis H_1 is given by the following expression.

$$\mu_{\Lambda|H_1}(T) = E\left\{\int_0^T (A + W(t))\,dt\right\} = AT + \int_0^T E\{W(t)\}dt = AT$$

In conclusion, we find that the mean value function has the following values under each hypothesis.

$$\boxed{\begin{array}{ll} H_0: & \mu_{\Lambda|H_0}(T) = 0 \\ H_1: & \mu_{\Lambda|H_1}(T) = AT \end{array}} \tag{8}$$

Part (b)

Following the derivation in Problem 8.7(b), we conclude that the general solution for the variance function $\sigma_\Lambda^2(T)$ is given by the following expression.

$$\begin{aligned} \sigma_\Lambda^2(T) &= E\left\{[\Lambda(T) - \mu_\Lambda(T)]^2\right\} = E\left\{[\Lambda(T) - \mu_\Lambda(T)]\,[\Lambda(T) - \mu_\Lambda(T)]^*\right\} \\ &= E\left\{\left[\int_0^T R(t_1)dt_1 - \mu_\Lambda(T)\right]\left[\int_0^T R^*(t_2)dt_2 - \mu_\Lambda^*(T)\right]\right\} \\ &= \int_0^T\int_0^T R_{RR}(t_1, t_2)dt_1dt_2 - \mu_\Lambda(T)\mu_\Lambda^*(T) \end{aligned} \tag{9}$$

Note that we have substituted for the correlation function $R_{RR}(t_1, t_2) = E\{R(t_1)R^*(t_2)\}$. At this point, we require closed-form expressions for the correlation function under each hypothesis. Let's begin by evaluating the correlation under hypothesis H_0.

$$R_{RR|H_0}(t_1, t_2) = E\{W(t_1)W^*(t_2)\} = \sigma^2\delta(t_1 - t_2) \tag{10}$$

Note that, by Equation 7.3-6 on page 436 in [4], we conclude that the correlation function for mean zero white Gaussian noise is given by the previous expression. Now let's evaluate the correlation function under the hypothesis H_1.

$$R_{RR|H_1}(t_1, t_2) = E\{[A + W(t_1)][A + W(t_2)]^*\} = A^2 + \sigma^2\delta(t_1 - t_2) \tag{11}$$

In conlcusion, substituting Equations 12, 10, and 11 into Equation 9 yields the following solution for the variance function under each hypothesis (which, as should be expected, turns out to be identical under either hypothesis).

$$\boxed{\begin{array}{ll} H_0: & \sigma^2_{\Lambda|H_0}(T) = T\sigma^2 \\ H_1: & \sigma^2_{\Lambda|H_1}(T) = T\sigma^2 \end{array}} \tag{12}$$

Part (c)

First, by Problem 8.13(c), we conclude that Λ is a Gaussian random variance. Under hypothesis H_0, Λ is a white Gaussian random noise processes with mean zero and variance $T\sigma^2$. As a result, the false alarm probability (i.e., the probability of incorrectly identifying a noise sequence as containing the target signal) is given by the following expression.

$$\begin{aligned} P[\Lambda \geq \Lambda_0|H_0] &= 1 - P[\Lambda < \Lambda_0|H_0] \\ &= 1 - \frac{1}{\sqrt{2\pi T\sigma^2}} \int_{-\infty}^{AT/2} \exp\left(\frac{-x^2}{2T\sigma^2}\right) dx \end{aligned}$$

Recall that the error function has the following definition.

$$\operatorname{erf}(z) \triangleq \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$$

In conclusion, we find that the false alarm probability has the following simple form in terms of the error function.

$$\boxed{P[\Lambda \geq \Lambda_0|H_0] = \frac{1}{2}\left[1 - \operatorname{erf}\left(\frac{AT}{2\sqrt{2T\sigma^2}}\right)\right]}$$

Briefly, we note that this function has several properties which must logically follow from the detection criterion. If $A = 0$, then the hypotheses are equal and we obtain $P[\Lambda \geq \Lambda_0|H_0] = 1/2$ – corresponding to equal detection likelihoods. Similarly, we recall that the error function has an output on the interval $(-1, 1)$. As a result, the false alarm probability $P[\Lambda \geq \Lambda_0|H_0]$ must be within the interval $(0, 1)$ depending on the value of parameters $\{A, T, \sigma\}$.

Problem 9.9

In this problem we will derive the Kalman filter under a Gauss-Markov signal model with nonzero mean. In the general case we consider a vector-valued discrete random process $\mathbf{X}[n]$ with nonzero mean. Let the Gauss-Markov signal model be

$$\mathbf{X}[n] = \mathbf{A}\mathbf{X}[n-1] + \mathbf{B}\mathbf{W}[n],\ n \geq 0$$

where $\mathbf{X}[-1] = \mathbf{0}$ and the centered noise process $\mathbf{W}_c[n] \triangleq \mathbf{W}[n] - \mu_{\mathbf{W}}[n]$ is white Gaussian with variance $\sigma^2_{\mathbf{W}}$ and $\mu_{\mathbf{W}}[n] \neq \mathbf{0}$. Note that $\mathbf{V} \perp \mathbf{W}_c$ and that the observation equation is given by

$$\mathbf{Y}[n] = \mathbf{X}[n] + \mathbf{V}[n],\ n \geq 0.$$

(a) Find expressions for $\mu_{\mathbf{X}}[n]$ and $\mu_{\mathbf{Y}}[n]$.

(b) Show that the MMSE estimate of $\mathbf{X}[n]$ equals the sum of $\mu_{\mathbf{X}}[n]$ and the MMSE estimate of $\mathbf{X}_c[n] \triangleq \mathbf{X}[n] - \mu_{\mathbf{X}}[n]$ based on the centered observations $\mathbf{Y}_c[n] \triangleq \mathbf{Y}[n] - \mu_{\mathbf{Y}}[n]$.

(c) Extend the Kalman filter Equation 9.2-16 to the nonzero mean case using the result of (b).

Part (a)

Let's begin by evaluating the mean function for $\mathbf{X}[n]$.

$$\mu_{\mathbf{X}}[n] = E\{\mathbf{A}\mathbf{X}[n-1] + \mathbf{B}\mathbf{W}[n]\} = \mathbf{A}\mu_{\mathbf{X}}[n-1] + \mathbf{B}\mu_{\mathbf{W}}[n]$$

Substituting the initial condition $\mathbf{X}[-1] = \mathbf{0}$, we find that $\mu_{\mathbf{X}}[0]$ is given by

$$\mu_{\mathbf{X}}[0] = \mathbf{A}\mu_{\mathbf{X}}[-1] + \mathbf{B}\mu_{\mathbf{W}}[0] = \mathbf{B}\mu_{\mathbf{W}}[0].$$

Iterating, we find the following expressions for $\mu_{\mathbf{X}}[1]$ and $\mu_{\mathbf{X}}[2]$.

$$\mu_{\mathbf{X}}[1] = \mathbf{A}\mu_{\mathbf{X}}[0] + \mathbf{B}\mu_{\mathbf{W}}[1] = \mathbf{A}\mathbf{B}\mu_{\mathbf{W}}[0] + \mathbf{B}\mu_{\mathbf{W}}[1]$$

$$\mu_{\mathbf{X}}[2] = \mathbf{A}\mu_{\mathbf{X}}[1] + \mathbf{B}\mu_{\mathbf{W}}[2] = \mathbf{A}^2\mathbf{B}\mu_{\mathbf{W}}[0] + \mathbf{A}\mathbf{B}\mu_{\mathbf{W}}[1] + \mathbf{B}\mu_{\mathbf{W}}[2]$$

By induction we conclude that $\mu_{\mathbf{X}}[n]$ and $\mu_{\mathbf{Y}}[n]$ are given by the following expression.

$$\boxed{\mu_{\mathbf{X}}[n] = \mu_{\mathbf{Y}}[n] = \sum_{m=0}^{n} \mathbf{A}^m\mathbf{B}\mu_{\mathbf{W}}[n-m],\ n \geq 0}$$

Note that $\mu_{\mathbf{X}}[n] = \mu_{\mathbf{Y}}[n]$ by substituting $\mu_{\mathbf{V}}[n] = \mathbf{0}$ in the following expression.

$$\mu_{\mathbf{Y}}[n] = E\{\mathbf{X}[n] + \mathbf{V}[n]\} = \mu_{\mathbf{X}}[n] + \mu_{\mathbf{V}}[n] = \mu_{\mathbf{X}}[n]$$

Part (b)

First, we recall that the MMSE estimate of $\mathbf{X}[n]$ is defined on page 576 in [4] as follows.

$$\hat{\mathbf{X}}[n] \triangleq E\{\mathbf{X}[n]|\mathbf{Y}[n-1], \mathbf{Y}[n-2], \ldots, \mathbf{Y}[0]\}$$

Note that the state estimate $\hat{\mathbf{X}}[n]$ is the conditional expectation of $\mathbf{X}[n]$ given the set of prior observations $\{\mathbf{Y}[n-1], \mathbf{Y}[n-2], \ldots, \mathbf{Y}[0]\}$. By the problem statement we wish to show that the following equality holds.

$$\hat{\mathbf{X}}[n] \stackrel{?}{=} \mu_{\mathbf{X}}[n] + \hat{\mathbf{X}}_c[n] = \mu_{\mathbf{X}}[n] + E\{\mathbf{X}_c[n]|\mathbf{Y}_c[n-1], \mathbf{Y}_c[n-2], \ldots, \mathbf{Y}_c[0]\}$$

This can be shown by applying the linearity property of the conditional expectation operator.

$$\begin{aligned}\hat{\mathbf{X}}[n] &= E\{\mathbf{X}[n]|\mathbf{Y}[n-1], \mathbf{Y}[n-2], \ldots, \mathbf{Y}[0]\} \\ &= E\{\mu_{\mathbf{X}}[n] + (\mathbf{X}[n] - \mu_{\mathbf{X}}[n])|\mathbf{Y}[n-1], \mathbf{Y}[n-2], \ldots, \mathbf{Y}[0]\} \\ &= \mu_{\mathbf{X}}[n] + E\{\mathbf{X}_c[n]|\mathbf{Y}[n-1], \mathbf{Y}[n-2], \ldots, \mathbf{Y}[0]\}\end{aligned}$$

Note that, as shown in the previous part, $\mu_{\mathbf{X}}[n]$ is independent of the observation sequence, so $E\{\mu_{\mathbf{X}}[n]|\mathbf{Y}[n-1], \mathbf{Y}[n-2], \ldots, \mathbf{Y}[0]\} = \mu_{\mathbf{X}}[n]$. At this point we can define the following innovations sequence for $\mathbf{Y}[n]$ for the noiseless centered observations $\mathbf{X}_c[n]$. As shown on pages 576-577, such a sequence must be an orthogonal (or white) random sequence which consists of a causal, linear transformation of $\mathbf{Y}[n]$. By the previous part we recall that $\mu_{\mathbf{X}}[n] = \mu_{\mathbf{Y}}[n]$. As a result, we find the the innovations sequence $\tilde{\mathbf{Y}}_c[n]$ is defined on the centered observations $\mathbf{Y}_c[n]$ as follows.

$$\tilde{\mathbf{Y}}_c[0] \triangleq \mathbf{Y}_c[0]$$

$$\tilde{\mathbf{Y}}_c[n] \triangleq \mathbf{Y}_c[n] - E\{\mathbf{Y}_c[n]|\mathbf{Y}_c[n-1], \mathbf{Y}_c[n-2], \ldots, \mathbf{Y}_c[0]\}, \text{ for } n \geq 1$$

Since the innovations sequence $\tilde{\mathbf{Y}}_c[n]$ and $\mathbf{Y}_c[n]$ are equivalent, we conclude that the equality holds.

$$\boxed{\hat{\mathbf{X}}[n] = \mu_{\mathbf{X}}[n] + E\{\mathbf{X}_c[n]|\mathbf{Y}_c[n-1], \mathbf{Y}_c[n-2], \ldots, \mathbf{Y}_c[0]\} = \mu_{\mathbf{X}}[n] + \hat{\mathbf{X}}_c[n]}$$

Part (c)

The Kalman filter, providing an optimal estimate of the system state $\mathbf{X}[n]$ given the observations $\{\mathbf{Y}[n], \mathbf{Y}[n-1], \ldots, \mathbf{Y}[0]\}$, is defined for mean zero sequences by Equation 9.2-16 as

$$\hat{\mathbf{X}}[n|n] = \mathbf{A}\hat{\mathbf{X}}[n-1|n-1] + \mathbf{G}_n(\mathbf{Y}[n] - \mathbf{A}\hat{\mathbf{X}}[n-1|n-1]),$$

where $\hat{\mathbf{X}}[n|m] \triangleq E\{\mathbf{X}[n]|\mathbf{Y}[m], \mathbf{Y}[m-1], \ldots, \mathbf{Y}[0]\}$ and $\hat{\mathbf{X}}[-1|-1] \triangleq \mathbf{0}$. From the previous part, we conclude that the Kalman filter for nonzero mean sequences has a similar form for the centered sequences.

$$\hat{\mathbf{X}}_c[n|n] = \mathbf{A}\hat{\mathbf{X}}_c[n-1|n-1] + \mathbf{G}_{c_n}(\mathbf{Y}_c[n] - \mathbf{A}\hat{\mathbf{X}}_c[n-1|n-1])$$

Note that the Kalman gain matrix $\mathbf{G}_{c_n}$ for the centered sequences may not correspond to that in the previous expression. Finally, we add the mean function to obtain the desired expression for the Kalman filter.

$$\boxed{\hat{\mathbf{X}}[n|n] = \mu_{\mathbf{X}}[n] + \mathbf{A}\hat{\mathbf{X}}_c[n-1|n-1] + \mathbf{G}_{c_n}(\mathbf{Y}_c[n] - \mathbf{A}\hat{\mathbf{X}}_c[n-1|n-1])}$$

Problem 2.3-4 [Larson and Shubert, p. 130]

A Gaussian random sequence $X[n]$, for $n = 0, 1, 2, \ldots$, is defined as

$$X[n] = -\sum_{k=1}^{n} \binom{k+2}{2} X[n-k] + W[n], \tag{13}$$

where $X[0] = W[0]$ and $W[n]$ is a Gaussian white noise sequence with zero mean and unity variance.

(a) Show that $W[n]$ is the innovations sequence for $X[n]$.

(b) Show that $X[n] = W[n]-3W[n-1]+3W[n-2]-W[n-3]$, for $W[-1] = W[-2] = W[-3] = 0$.

(c) Use the preceding result to obtain the best two-step predictor of $X[12]$ as a linear combination of $X[0], \ldots, X[10]$. Also calculate the resulting mean-square prediction error.

Part (a)

Recall, from Definition 9.2-1 on page 571 in [4], that the innovations sequence for a random sequence $X[n]$ is defined to be a white random sequence which is a casual and causally-invertible linear transformation of the sequence $X[n]$. From Equation 13 we find that $W[n]$ is a causal linear transformation of $\{X[0], X[1], \ldots, X[n]\}$ such that

$$W[n] = X[n] + \sum_{k=1}^{n} \binom{k+2}{2} X[n-k].$$

In addition, we note that each $X[n]$ is composed of a linear combination of zero-mean Gaussian random variables and, as a result, must also be a white random sequence. In conclusion, we find that $W[n]$ is a white random sequence that is causally equivalent to $X[n]$. Similarly, as we'll show in Part (b), $X[n]$ can be expressed as a causal linear combination of $\{W[n-3], W[n-2], W[n-1], W[n]\}$. As a result, we find that $W[n]$ is the innovations sequence for $X[n]$ since it satisfies Definition 9.2-1. In other words, $W[n]$ contains the new information obtained when we observe $X[n]$ given the past observations $\{X[n-1], X[n-2], \ldots, X[0]\}$.

Part (b)

Let's begin by evaluating $X[1]$ by direct evaluation of Equation 13.

$$\begin{aligned} X[1] &= W[1] - \sum_{k=1}^{1} \binom{k+2}{2} X[1-k] \\ &= W[1] - 3X[0] = W[1] - 3W[0] \end{aligned}$$

Similarly, for $X[2]$ we find the following result.

$$\begin{aligned} X[2] &= W[2] - \sum_{k=1}^{2} \binom{k+2}{2} X[2-k] \\ &= W[2] - 3X[1] - 6X[0] = W[2] - 3W[1] + 3W[0] \end{aligned}$$

Continuing our analysis we find that $X[3]$ has the following solution.

$$
\begin{aligned}
X[3] &= W[3] - \sum_{k=1}^{3} \binom{k+2}{2} X[3-k] \\
&= W[3] - 3X[2] - 6X[1] - 10X[0] = W[3] - 3W[2] + 3W[1] - W[0]
\end{aligned}
$$

By induction we conclude that the general solution for $X[n]$, for $n = 0, 1, 2, \ldots$, is given by the following expression.

$$
\boxed{X[n] = W[n] - 3[n-1] + 3W[n-2] - W[n-3], \text{ for } W[-1] = W[-2] = W[-3] = 0}
$$

Part (c)

Recall that the best two-step predictor $\hat{X}[12]$ of $X[12]$ will be given by the following conditional expectation.

$$
\hat{X}[12] = E\{X[12]|X[10], \ldots, X[0]\}
$$

Note that $W[n]$, the innovations sequence, is causally equivalent to $X[n]$. As a result, we can also express the two-step predictor as follows.

$$
\begin{aligned}
\hat{X}[12] &= E\{X[12]|W[10], \ldots, W[0]\} \\
&= E\{W[12] - 3W[11] + 3W[10] - W[9]|W[10], \ldots, W[0]\} \\
&= 3W[10] - W[9]
\end{aligned}
$$

Note that we substituted for $X[n]$ using the result found in Part (b). Since $W[n]$ is a white random process, we also conclude that $E\{W[12]|W[10], \ldots, W[0]\} = E\{W[11]|W[10], \ldots, W[0]\} = 0$. As a result, the best two-step predictor of $X[12]$ is given by the following expression.

$$
\boxed{\hat{X}[12] = 3W[10] - W[9] = 3\left\{X[10] + \sum_{k=1}^{10} \binom{k+2}{2} X[10-k]\right\} - \left\{X[9] + \sum_{k=1}^{9} \binom{k+2}{2} X[9-k]\right\}}
$$

Finally, we note that the mean-square prediction error ε^2 is given by the following expression.

$$
\begin{aligned}
\varepsilon^2 &= E\{(X[12] - \hat{X}[12])^2\} = E\{(W[12] - 3W[11])^2\} \\
&= E\{W[12]^2\} - 6E\{W[12]W[11]\} + 9E\{W[11]^2\} \\
&= E\{W[12]^2\} - 6E\{W[12]\}E\{W[11]\} + 9E\{W[11]^2\} = 10
\end{aligned}
$$

Since $W[n]$ is a mean-zero white random process we conclude that $E\{W[12]^2\} = E\{W[11]^2\} = 1$ and $E\{W[12]W[11]\} = E\{W[12]\}E\{W[11]\} = 0$. In conclusion, the mean-square prediction error ε^2 for $X[12]$ is given by the following equation.

$$
\boxed{\varepsilon^2 = E\{(X[12] - \hat{X}[12])^2\} = 10}
$$

References

[1] Geoffrey Grimmett and David Stirzaker. *Probability and Random Processes (Third Edition).* Oxford University Press, 2001.

[2] Harold J. Larson and Bruno O. Shubert. *Probabilistic Models in Engineering Sciences: Random Variables and Stochastic Processes.* John Wiley & Sons, 1979.

[3] Michael Mitzenmacher and Eli Upfal. *Probability and Computing: Randomized Algorithms and Probabilistic Analysis.* Cambridge University Press, 2005.

[4] Henry Stark and John W. Woods. *Probability and Random Processes with Applications to Signal Processing (Third Edition).* Prentice-Hall, 2002.