Solutions to Homework 9 - Gaussian processes

1) Conditional Gaussian density. Consider a two-dimensional *standard* Gaussian random vector $\mathbf{X} = [U, V]^T$, with mean $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$ given by

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \qquad \mathbf{C} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

A) Notice that $\det(\mathbf{C}) = 1 - \rho^2$ and

$$\mathbf{C}^{-1} = \frac{1}{\det(\mathbf{C})} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}.$$

Substituting these expressions in the pdf

$$f_{UV}(u, v) = \frac{1}{2\pi\sqrt{\det(\mathbf{C})}} \exp\left(-\frac{1}{2}\mathbf{x}^T\mathbf{C}^{-1}\mathbf{x}\right)$$

where $\mathbf{x} = [u, v]^T$, readily yields the desired result

$$f_{UV}(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{u^2 - 2\rho uv + v^2}{2(1-\rho^2)}\right).$$

The Gaussian random variables (RVs) U and V are uncorrelated if $\text{cov}(U, V) = 0$. From the structure of $\mathbf{C}$, it follows that $\text{cov}(U, V) = \rho$. The conclusion is that U and V are uncorrelated if and only if $\rho = 0$. Since for Gaussian RVs uncorrelatedness implies independence, U and V are independent if and only if $\rho = 0$.

B) The conditional pdf $f_{U|V}(u \,|\, v)$ is defined as

$$f_{U|V}(u \,|\, v) = \frac{f_{UV}(u, v)}{f_V(v)}.$$

Since we obtained $f_{UV}(u, v)$ in part *A*, what is left to determine is the marginal pdf of V, i.e.,

$$f_V(v) = \int_{-\infty}^{\infty} f_{UV}(u, v)du.$$

Since it holds that $u^2 - 2\rho uv + v^2 = v^2(1-\rho^2) + (u - \rho v)^2$, one can rewrite the joint pdf as

$$\begin{aligned} f_{UV}(u, v) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{v^2(1-\rho^2) + (u-\rho v)^2}{2(1-\rho^2)}\right) \\ &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{v^2}{2}\right) \exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)}\right) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right) \times \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)}\right) \end{aligned}$$

where the first factor in the last line corresponds to the pdf of a one-dimensional standard Gaussian RV, and the second factor is the pdf of a Gaussian RV with mean ρv and variance $1 - \rho^2$. Carrying out the integration of $f_{UV}(u, v)$ to obtain the marginal pdf yields

$$\begin{aligned} f_V(v) &= \int_{-\infty}^{\infty} f_{UV}(u, v)du \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right) \times \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)}\right) du \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{v^2}{2}\right) \end{aligned}$$

where the last equality follows because every Gaussian pdf integrates to one over the whole real line. The factorization of the joint pdf derived earlier becomes quite handy towards obtaining the desired conditional pdf, since canceling out the factor $f_V(v)$ yields

$$f_{U|V}(u \mid v) = \frac{f_{UV}(u,v)}{f_V(v)} = \frac{1}{\sqrt{2\pi(1-\rho^2)}} \exp\left(-\frac{(u-\rho v)^2}{2(1-\rho^2)}\right).$$

It is thus apparent that given $V = v$, the RV U is conditionally Gaussian with mean $\mathbb{E}\left[U \,\middle|\, V = v\right] = \rho v$ and variance $\text{var}\left[U \,\middle|\, V = v\right] = 1 - \rho^2$.

2) Jointly Gaussian random variables. A random vector $\mathbf{X} = [X_1, \ldots, X_n]^T$ is said to be Gaussian if every linear combination of the entries of $\mathbf{X}$, i.e., $\sum_{i=1}^n a_i X_i$ is a scalar Gaussian RV. Equivalent terminology is that $X_1, \ldots, X_n$ are jointly Gaussian RVs.

A) To show that every subvector of a Gaussian vector $\mathbf{X}$ is also Gaussian, it suffices to set $a_i = 0$ for those entries i we want to leave out of the original random vector $\mathbf{X}$. Since by assumption all linear combinations yield a scalar RV, so will those particular cases where some of the coefficients are equal to zero. In particular, if all but the $i-$th entry of $\mathbf{a} := [a_1, \ldots, a_n]^T$ are equal to zero, then it follows that X_i is a Gaussian RV.

B) Next we show that an affine transformation of a Gaussian random vector $\mathbf{X}$ is Gaussian, i.e., for arbitrary (deterministic) matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and column vector $\mathbf{b} \in \mathbf{R}^m$, then $\mathbf{Y} = \mathbf{AX} + \mathbf{b}$ is a Gaussian random vector. Notice that it suffices to show that $\mathbf{AX}$ is Gaussian, because the addition of $\mathbf{b}$ only affects the centering (mean) of $\mathbf{Y}$, but not the type of distribution. For an arbitrary vector of coefficients $\mathbf{a} := [a_1, \ldots, a_n]^T$ the linear combination

$$Y = \mathbf{a}^T \mathbf{AX} = (\mathbf{a}^T \mathbf{A})\mathbf{X} = \mathbf{c}^T \mathbf{X}$$

is a Gaussian RV because $\mathbf{c} := \mathbf{A}^T \mathbf{a}$ is just another vector of coefficients, and by assumption $\mathbf{X}$ is a Gaussian vector completing the proof.

Suppose that $\mathbf{X}$ has mean vector $\boldsymbol{\mu}_X$ and covariance matrix $\mathbf{C}_X$. Then the mean vector $\boldsymbol{\mu}_Y$ of $\mathbf{Y}$ is given by

$$\mu_Y = \mathbb{E}\left[\mathbf{Y}\right] = \mathbb{E}\left[\mathbf{AX} + \mathbf{b}\right] = \mathbf{A}\mu_X + \mathbf{b}$$

while the covariance matrix $\mathbf{C}_Y$ is

$$\mathbf{C}_Y = \mathbb{E}\left[(\mathbf{Y} - \mu_Y)(\mathbf{Y} - \mu_Y)^T\right] = \mathbb{E}\left[\mathbf{A}(\mathbf{X} - \mu_X)(\mathbf{X} - \mu_X)^T \mathbf{A}^T\right] = \mathbf{AC}_X\mathbf{A}^T.$$

Given μ_Y and $\mathbf{C}_Y$, the joint pdf of $\mathbf{Y}$ is

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{(2\pi)^{n/2}\sqrt{\det(\mathbf{AC}_X\mathbf{A}^T)}} \exp\left(-\frac{1}{2}(\mathbf{y} - \mathbf{A}\mu_X - \mathbf{b})^T \left(\mathbf{AC}_X\mathbf{A}^T\right)^{-1} (\mathbf{y} - \mathbf{A}\mu_X - \mathbf{b})\right).$$

C) Finally we show that if $Z_1, \ldots, Z_n$ are mutually independent Gaussian RVs, then they are also jointly Gaussian. To that end, we will argue by induction on n. Clearly, the claim is true for $n = 1$ since aZ_i is a Gaussian RV for each a and all $i = 1, \ldots, n$. Next we suppose the claim holds true for $n - 1$ (meaning that if $Z_1, \ldots, Z_{n-1}$ are mutually independent Gaussian RVs, then they are also jointly Gaussian), and need to show it also holds for n. To that end, we have to show that for arbitrary $a_1, \ldots, a_n$ then

$$Y^{(n)} = \sum_{i=1}^{n} a_i Z_i = Y^{(n-1)} + a_n Z_n$$

is a Gaussian RV. But by assumption $Y^{(n-1)}$ is Gaussian and independent of the Gaussian RV $a_n Z_n$. So the pdf of $Y^{(n)}$ is given by the convolution of the pdfs of each of the Gaussian summands. We will not prove it here, but either a direct calculation of the convolution integral or a Fourier transform argument can be used to establish that the convolution of two Gaussian pdfs is itself a Gaussian pdf. This shows that $Y^{(n)}$ is a Gaussian RV, completing the proof.

3) Decorrelation of a Gaussian random vector. Suppose the random vector $\mathbf{X} = [X_1, \ldots, X_n]^T$ has zero mean and covariance matrix $\mathbf{C}_X$. Notice that since covariance matrices are symmetric, by the spectral theorem they are diagonalizable and there exists an orthonormal basis of eigenvectors. Specifically, there exists a diagonal matrix

$\mathbf{\Lambda} \in \mathbf{R}^{n\times n}$ and a unitary matrix $\mathbf{U} \in \mathbf{R}^{n\times n}$ (meaning $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$) containing the eigenvectors of $\mathbf{C}_X$ as its rows, such that

$$\mathbf{U}\mathbf{C}_X\mathbf{U}^T = \mathbf{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_n)$$

where $\lambda_1, \dots, \lambda_n$ are the eigenvalues of $\mathbf{C}_X$.

Given the aforementioned matrix of eigenvectors $\mathbf{U}$, define the random vector $\mathbf{Y} := \mathbf{U}\mathbf{X}$. Clearly,

$$\mathbb{E}[\mathbf{Y}] = \mathbb{E}[\mathbf{U}\mathbf{X}] = \mathbf{U}\mathbb{E}[\mathbf{X}] = \mathbf{0}$$

and the covariance matrix of $\mathbf{Y}$ is

$$\mathbf{C}_Y = \mathbb{E}\left[\mathbf{Y}\mathbf{Y}^T\right] = \mathbb{E}\left[\mathbf{U}\mathbf{X}\mathbf{X}^T\mathbf{U}^T\right] = \mathbf{U}\mathbb{E}\left[\mathbf{X}\mathbf{X}^T\right]\mathbf{U}^T = \mathbf{U}\mathbf{C}_X\mathbf{U}^T = \mathbf{\Lambda}.$$

The conclusion is that $\mathbf{Y}$ has uncorrelated entries $Y_1, \dots, Y_n$, since its covariance matrix $\mathbf{C}_Y = \mathbf{\Lambda}$ is diagonal. In this context, the linear transformation effected by $\mathbf{U}$ is known as *decorrelating transformation*.

If we also assume that $\mathbf{X}$ is a Gaussian random vector, then because a linear transformation of a Gaussian vector is Gaussian it follows that $Y_1, \dots, Y_n$ are jointly Gaussian. Moreover, since uncorrelatedness implies independence for Gaussian RVs, then $Y_1, \dots, Y_n$ are also mutually independent.

4) Difference of two independent Brownian motion processes. Let $X_1(t)$ and $X_2(t)$ be two Brownian motion processes defined for $t \geq 0$, with respective variance parameters σ_1^2 and σ_2^2. Suppose that $X_1(t)$ and $X_2(t)$ are independent. Let the random process $D(t)$ be defined as $D(t) = X_1(t) - X_2(t)$, $t \geq 0$.

A) From the definition of autocorrelation function $R_D(t_1, t_2) = \mathbb{E}[D(t_1)D(t_2)]$, and because $X_1(t)$ and $X_2(t)$ are zero-mean and independent Brownian motion processes it follows that

$$\begin{aligned}
R_D(t_1, t_2) &= \mathbb{E}[(X_1(t_1) - X_2(t_1))(X_1(t_2) - X_2(t_2))] \\
&= R_{X_1}(t_1, t_2) + R_{X_2}(t_1, t_2) - \mathbb{E}[X_1(t_1)X_2(t_2)] - \mathbb{E}[X_2(t_1)X_1(t_2)] \\
&= \sigma_1^2 \min(t_1, t_2) + \sigma_2^2 \min(t_1, t_2) - \mathbb{E}[X_1(t_1)]\,\mathbb{E}[X_2(t_2)] - \mathbb{E}[X_2(t_1)]\,\mathbb{E}[X_1(t_2)] \\
&= (\sigma_1^2 + \sigma_2^2) \min(t_1, t_2).
\end{aligned}$$

B) Since a linear transformation of GPs is a GP, then $D(t)$, $t \geq 0$, is a Gaussian RV with mean

$$\mathbb{E}[D(t)] = \mathbb{E}[X_1(t)] - \mathbb{E}[X_2(t)] = 0$$

and variance

$$\text{var}[D(t)] = \mathbb{E}\left[D^2(t)\right] - \mathbb{E}[D(t)]^2 = R_D(t, t) = (\sigma_1^2 + \sigma_2^2)t.$$

Thus the desired pdf is given by

$$f_{D(t)}(x) = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)t}} \exp\left(-\frac{x^2}{2(\sigma_1^2 + \sigma_2^2)t}\right).$$

5) A Brownian motion processes is a Martingale. Consider a Brownian motion process $X(t)$ defined for $t \geq 0$. Then for $0 \leq s < t$, one has

$$\begin{aligned}
\mathbb{E}\left[X(t) \,\middle|\, X(u), 0 \leq u \leq s\right] &= \mathbb{E}\left[X(s) + X(t) - X(s) \,\middle|\, X(u), 0 \leq u \leq s\right] \\
&= \mathbb{E}\left[X(s) \,\middle|\, X(u), 0 \leq u \leq s\right] + \mathbb{E}\left[X(t) - X(s) \,\middle|\, X(u), 0 \leq u \leq s\right] \\
&= X(s) + \mathbb{E}[X(t) - X(s)] = X(s)
\end{aligned}$$

which establishes that $X(t)$ is a Martingale. Notice that the last equality follows because $X(t)$ has zero mean, while the third equality follows since $\mathbb{E}\left[X(t) - X(s) \,\middle|\, X(u), 0 \leq u \leq s\right] = \mathbb{E}[X(t) - X(s)]$ by the independent increments property of $X(t)$.

6) Conditional density of a Brownian motion process with drift. Let $X(t)$ for $t \geq 0$ be a Brownian motion process with drift parameter μ and variance parameter σ^2, i.e., $X(t)$ is Gaussian distributed with mean μt and variance $\sigma^2 t$. The goal is to specify the conditional distribution of $X(t)$ given that $X(s) = c$, for $s < t$. Writing $X(t) = X(s) + X(t) - X(s)$ and using the independent increments property of the Brownian motion process with drift, we obtain that given $X(s) = c$, $X(t)$ has the same distribution as $c + X(t) - X(s)$. By the stationary increments property of the Brownian motion process with drift, $X(t)$ given $X(s)=c$ is thus distributed as $c+X(t-s)$ which is Gaussian with mean $\mathbb{E}\left[X(t)\,\middle|\,X(s)=c\right]=c+\mu(t-s)$, and variance var $\left[X(t)\,\middle|\,X(s)=c\right]=\sigma^2(t-s)$.

7) White Gaussian noise.

A) Independent values. The autocorrelation function of the white Gaussian noise (WGN) process is

$$R_W(t_1,t_2)=\sigma^2\delta(t_1-t_2)$$

where $\delta(t)$ is the Dirac delta (generalized) function that is infinite at $t=0$, and zero everywhere else. This means that for different times t_1 and t_2, $R_W(t_1,t_2)=0$. Therefore, the random variables $W(t_1)$ and $W(t_2)$ are uncorrelated. But since $W(t)$ is a Gaussian process (GP) then uncorrelatedness implies independence, so that $W(t_1)$ and $W(t_2)$ are also independent.

B) Integral of WGN. The Brownian motion process

$$X(t)=\int_0^t W(u)du$$

is a GP, because it is a linear functional of a GP (WGN specifically). Using the linearity of the expectation operator and that $\mu_W(t)=0$, the mean function $\mu_X(t)$ of $X(t)$ is

$$\mu_X(t)=\mathbb{E}\left[\int_0^t W(u)du\right]=\int_0^t \mathbb{E}\left[W(u)\right]du=\int_0^t \mu_W(u)du=0.$$

Likewise, since $R_W(t_1,t_2)=\sigma^2\delta(t_1-t_2)$ the autocorrelation function $R_X(t_1,t_2)$ of $X(t)$ is for $t_1<t_2$

$$\begin{aligned}
R_X(t_1,t_2)&=\mathbb{E}\left[\left(\int_0^{t_1} W(u)du\right)\left(\int_0^{t_2} W(v)dv\right)\right]\\
&=\mathbb{E}\left[\int_0^{t_1}\int_0^{t_2} W(u)W(v)dvdu\right]\\
&=\int_0^{t_1}\int_0^{t_2}\mathbb{E}\left[W(u)W(v)\right]dvdu\\
&=\int_0^{t_1}\int_0^{t_2}\sigma^2\delta(u-v)dvdu\\
&=\int_0^{t_1}\int_0^{t_1}\sigma^2\delta(u-v)dvdu+\int_0^{t_1}\int_{t_1}^{t_2}\sigma^2\delta(u-v)dvdu\\
&=\int_0^{t_1}\int_0^{t_1}\sigma^2\delta(u-v)dvdu=\int_0^{t_1}\sigma^2du=\sigma^2t_1.
\end{aligned}$$

Arguing in the exactly same way for $t_2<t_1$ yields $R_X(t_1,t_2)=\sigma^2t_2$, so that all in all

$$R_X(t_1,t_2)=\sigma^2\min(t_1,t_2).$$

Since the Brownian motion process is a GP, then for each $t\geq 0$, $X(t)$ is a zero-mean Gaussian random variable with variance

$$\text{var}\left[X(t)\right]=\mathbb{E}\left[X^2(t)\right]-\mathbb{E}\left[X(t)\right]^2=R_X(t,t)-\mu_X^2(t)=\sigma^2t.$$

So any desired probability can be obtained by suitably integrating the Gaussian pdf, for instance

$$\mathrm{P}\left[X(t)>a\right]=\int_a^\infty\frac{1}{\sqrt{2\pi\sigma^2t}}\exp\left(-\frac{x^2}{2\sigma^2t}\right)dx.$$

C) Discrete time representation of WGN. For the discrete-time process $W_h(n)$ defined by

$$W_h(n)=\int_{nh}^{(n+1)h}W(u)du$$

the mean function is

$$\mu_{W_h}(n)=\mathbb{E}\left[W_h(n)\right]=\mathbb{E}\left[\int_{nh}^{(n+1)h}W(u)\,du\right]=\int_{nh}^{(n+1)h}\mathbb{E}\left[W(u)\right]\,du=\int_{nh}^{(n+1)h}\mu_W(u)du=0.$$

The autocorrelation function $R_{W_h}(n_1, n_2)$ follows from the definition, namely

$$\begin{aligned}
R_{W_h}(n_1,n_2) &= \mathbb{E}\left[\left(\int_{n_1 h}^{(n_1+1)h} W(u)\,du\right)\left(\int_{n_2 h}^{(n_2+1)h} W(v)\,dv\right)\right] \\
&= \mathbb{E}\left[\int_{n_1 h}^{(n_1+1)h}\int_{n_2 h}^{(n_2+1)h} W(u)W(v)\,du\,dv\right] \\
&= \int_{n_1 h}^{(n_1+1)h}\int_{n_2 h}^{(n_2+1)h} \mathbb{E}\left[W(u)W(v)\right]\,du\,dv \\
&= \int_{n_1 h}^{(n_1+1)h}\int_{n_2 h}^{(n_2+1)h} \sigma^2\delta(u-v)\,du\,dv = \begin{cases} \sigma^2 h, & n_1 = n_2 \\ 0, & n_1 \neq n_2 \end{cases} .
\end{aligned}$$

So upon defining the discrete-time Dirac delta sequence $\delta_d(n)$ as

$$\delta_d(n) = \begin{cases} 1, & n = 0 \\ 0, & n \neq 0 \end{cases}$$

the autocorrelation function can be expressed as $R_{W_h}(n_1, n_2) = \sigma^2 h \delta_d(n_1 - n_2)$.

D) Simulation of Brownian motion process. The Matlab code to perform a discrete-time simulation $X_h(n) = \sum_{i=0}^{n} W_h(i)$ of the Brownian motion process $X(t)$ follows. The choice of parameters is $h = 0.01$, $\sigma^2 = 1$, and $t_{max} = 10$. Based on our previous calculations the mean is $\mu_{X_h}(n) = 0$, and the variance is $\text{var}[X_h(n)] = R_{X_h}(n, n) = \sigma^2 h$.

```
clear all; close all; clc;
h=0.01; sigma=1; t_MAX=10;

W_vector=normrnd(0,sigma*sqrt(h),1,t_MAX/h);
X_vector=cumsum(W_vector);

plot(h:h:t_MAX,X_vector,'r','Linewidth',1);
xlabel('time');title(['Weiner Process Simulated, h=',num2str(h)]);
grid on; axis([0 t_MAX -5 5])
```

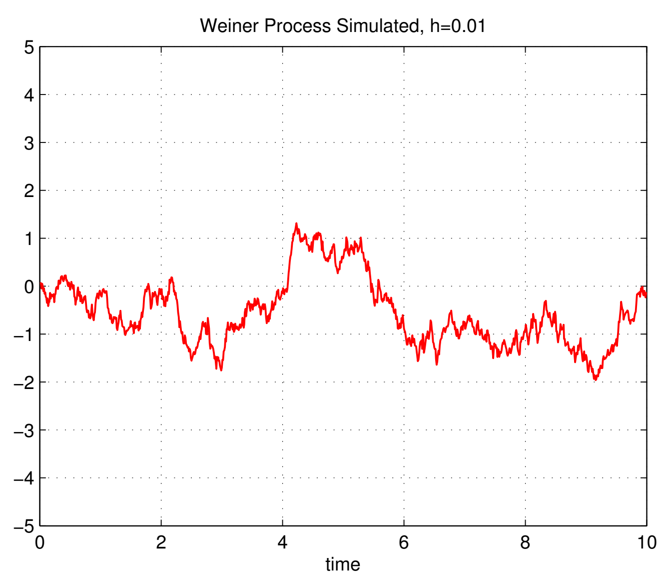

Fig. 1. A simulated sample path of the Weiner process $X(t)$ using its discrete approximation. (Part *D*.)