[AI2613 Lecture 12] Gaussian Processes, Brownian Bridge

June 9, 2023

1 Gaussian Processes and Brownian Motion

In the last lecture, we defined the standard Brownian motion:

Definition 1 (Standard Brownian Motion / Wiener Process)**.** *We say a stochastic process $\{W(t)\}_{t\geq 0}$ is a standard Brownian motion or Wiener process if it satisfies*

- $W(0) = 0$*;*
- **Independent increments***:* $\forall 0 \leq t_0 \leq t_1 \leq \cdots \leq t_n$, $W(t_1) - W(t_0)$, $W(t_2) - W(t_1), \ldots, W(t_n) - W(t_{n-1})$ *are mutually independent;*
- **Stationary increments***:* $\forall s, t > 0$, $W(s+t) - W(s) \sim \mathcal{N}(0, t)$*;*
- $W(t)$ *is continuous almost surely.*[1]

[1] Let Ω be the sample space. Then W can be viewed as a mapping from $\mathbb{R} \times \Omega$ to $\mathbb{R}$. The meaning of "$W(t)$ is continuous almost surely" is: $\exists \Omega_0 \subseteq \Omega$ with $\mathbf{Pr}[\Omega_0] = 1$ such that $\forall \omega \in \Omega_0$, $W(t, \omega)$ is continuous with regard to t.

Today we will give another characterization of Brownian motions in terms of the *Gaussian process*. First recall the notion of high dimensional Gaussian distribution. A vector of random variables $(X_1, X_2, \ldots, X_n)$ is said to be Gaussian iff $\forall a_1, a_2, \ldots, a_n$, $\sum_{i=1}^n a_i X_i$ is a one-dimensional Gaussian. Let $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ where $\mu_i = \mathbf{E}[X_i]$. Let $\Sigma = \left(\mathrm{Cov(X_i, X_j)}\right)_{i,j}$. Then the probability density function f of $(X_1, X_2, \ldots, X_n)$ is

$$\text{for } x = (x_1, x_2, \ldots, x_n),\ f(x) = (2\pi)^{-\frac{n}{2}} \cdot |\det\Sigma|^{-\frac{1}{2}} \cdot e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}.$$

Definition 2 (Gaussian Process)**.** *A stochastic process $\{X(t)\}_{t\geq 0}$ is called Gaussian process if* $\forall 0 \leq t_1 \leq t_2 \leq \cdots \leq t_n$, $(X(t_1), X(t_2), \ldots, X(t_n))$ *is Gaussian.*

Note that a Gaussian vector can be characterized by its mean vector and the covariance matrix. Standard Brownian motion is a special family of Gaussian processes where the covariance of $X(s)$ and $X(t)$ is $s \wedge t$.

Definition 3 (Standard Brownian Motion/Standard Wiener Process)**.** *We say a stochastic process $\{W(t)\}_{t\geq 0}$ is a standard Brownian motion or Wiener process if it satisfies*

- $\{W(t)\}_{t\geq 0}$ *is an almost surely continuous Gaussian Process;*
- $\forall s \geq 0$, $\mathbf{E}[W(s)] = 0$*;*
- $\forall 0 \leq s \leq t$, $\mathrm{Cov}(W(s), W(t)) = s$*.*

We will show that it is easier to use Definition 3 to verify that a certain stochastic process is a Brownian motion. Let us first verify that the two definitions are equivalent.

Proposition 4. *The two definitions of standard Brownian motions are equivalent.*

Proof. Given Definition 1, it is easy to know that $\mathbf{E}[W(s)] = 0$ for all $s \geq 0$ since $W(s) \sim \mathcal{N}(0, s)$. What we need is to verify that $\{W(t)\}_{t\geq 0}$ in Definition 1 is a Gaussian process and to compute the covariance of $W(s)$ and $W(t)$ in Definition 1.

It is worth noting that the sum of two Gaussians is not necessarily a Gaussian, unless they are joint Gaussian. Independence is just a special case of joint Gaussian (the covariance is zero).

Note that $\forall 0 \leq s < t$ and $\forall a, b$, we have

$$aW(s) + bW(t) = (a+b)W(s) + b(W(t) - W(s)).$$

Since $W(s)$ and $W(t)-W(s)$ are two independent Gaussian's, $aW(s)+bW(t)$ is still a Gaussian.

By the distributive law of covariance, for any $0 \leq s \leq t$, we have

$$\begin{aligned}
\mathrm{Cov}(W(s), W(t)) &= \mathrm{Cov}\left(W(s), W(t) - W(s) + W(s)\right) \\
&= \mathrm{Cov}(W(s), W(t) - W(s)) + \mathrm{Cov}(W(s), W(s)) \\
&= \mathbf{Var}\left[W(s)\right] = s.
\end{aligned}$$

Then we consider the counterpart. Given Definition 3, we can deduce the first and fourth property in Definition 1 directly. For any $0 \leq t_{i-1} \leq t_i \leq t_{j-1} \leq t_j$, we have

$$\begin{aligned}
&\mathrm{Cov}(W(t_i) - W(t_{i-1}), W(t_j) - W(t_{j-1})) \\
&= \mathrm{Cov}(W(t_i), W(t_j)) + \mathrm{Cov}(W(t_{i-1}), W(t_{j-1})) \\
&\quad - \mathrm{Cov}(W(t_i), W(t_{j-1})) - \mathrm{Cov}(W(t_{i-1}), W(t_j)) \\
&= t_i + t_{i-1} - t_i - t_{i-1} = 0,
\end{aligned}$$

which yields the independence of $W(t_i) - W(t_{i-1})$ and $W(t_j) - W(t_{j-1})$. Thus, the $\{W(t)\}_{t\geq 0}$ in Definition 3 satisfies independent increments.

It is easy to verify that $\forall s, t > 0$, $W(s+t) - W(s)$ is a Gaussian with mean 0. Note that

$$\begin{aligned}
\mathbf{Var}\left[W(t+s) - W(s)\right] &= \mathbf{E}\left[(W(t+s) - W(s))^2\right] \\
&= \mathbf{E}\left[W(t+s)^2\right] + \mathbf{E}\left[W(s)^2\right] - 2\mathbf{E}\left[W(t+s)W(s)\right] \\
&= \mathbf{Var}\left[W(t+s)^2\right] + \mathbf{Var}\left[W(s)^2\right] - 2\mathrm{Cov}\left(W(t+s), W(s)\right) \\
&= t + s + s - 2s = t.
\end{aligned}$$

Thus, the $\{W(t)\}_{t\geq 0}$ in Definition 3 satisfies stationary increments. □

Example 1. *Suppose $\{W(t)\}_{t\geq 0}$ is a standard Brownian motion. We claim that $\{X(t)\}_{t\geq 0}$ is also a standard Brownian motion where $X(0) = 0$ and $X(t) = t \cdot W(\frac{1}{t})$ for $t > 0$.*

We verify the three requirements in Definition 3.

Since $X(t) = t \cdot W(\frac{1}{t})$ which is the composition of two (almost surely) continuous function, $\{X(t)\}_{t\geq 0}$ is continuous almost surely as well. For any

$a_1, a_2, \ldots, a_n$ and $t_1, t_2, \ldots, t_n \geq 0$, $\sum_{i=1}^n a_i X(t_i) = \sum_{i=1}^n a_i t_i \cdot W(\frac{1}{t_i})$. Since $\{W(t)\}$ is standard Brownian motion, $\sum_{i=1}^n a_i t_i \cdot W(\frac{1}{t_i})$ is Gaussian. Thus, $\{X(t)\}_{t\geq 0}$ is a Gaussian process. For $0 \leq s < t$,

$$\begin{aligned}\mathrm{Cov}(X(s), X(t)) &= \mathrm{Cov}(sW(\frac{1}{s}), tW(\frac{1}{t})) \\ &= st \cdot \mathrm{Cov}(W(\frac{1}{s}), W(\frac{1}{t})) \\ &= st \cdot \frac{1}{t} = s.\end{aligned}$$

Thus, $\{X(t)\}_{t\geq 0}$ is a standard Brownian motion.

2 Brownian Bridge

In the last lecture, we already caluclated the distribution of $W(t)$ conditioned on $W(u) = x$ for some $u \geq t$. We use $X(t)$ to denote this process, and $X(t)$ is usually called a *Brownian bridge.*

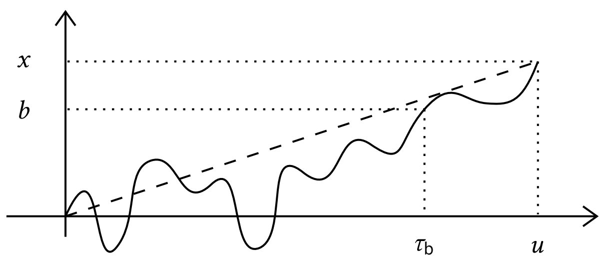

Figure 1: A Brownian bridge

We know from previous calculations that $X(t) \sim N(\frac{t}{u}x, \frac{t(u-t)}{u})$ is a Gaussian. Since the *conditional distribution of a multidimensional Gaussian distribution is Gaussian as well*, $X(t)$ is a Gaussian process. As a result, it is useful to compute the covariance of this process.

Recall $W(t)$ is a standard Brownian motion. For any $s \leq t$, we have

$\mathbf{E}\left[W(t)^2 \,\middle|\, W(u) = x\right] = \mathbf{Var}\left[X(t)\right] + \mathbf{E}\left[X(t)\right]^2$

$$\begin{aligned}&\mathtt{Cov}(X(s), X(t)) \\ &= \mathtt{Cov}\left(W(s), W(t) \mid W(u) = x\right) \\ &= \mathbf{E}\left[W(s) \cdot W(t) \mid W(u) = x\right] - \mathbf{E}\left[W(s) \mid W(u) = x\right] \cdot \mathbf{E}\left[W(t) \mid W(u) = x\right] \\ &= \int_{-\infty}^{\infty} y\mathbf{E}\left[W(s) \mid W(t) = y, W(u) = x\right] \cdot f_{W(t) \mid W(u)}\left(y \mid x\right) \mathrm{d}y - \frac{st}{u^2}x^2. \\ &= \frac{s}{t}\mathbf{E}\left[W(t)^2 \,\middle|\, W(u) = x\right] - \frac{st}{u^2}x^2 \\ &= \frac{s(u-t)}{u}.\end{aligned}$$

Definition 5 (Standard Brownian Bridge)**.** *A standard Brownian motion ending at $W(1) = 0$ is called a standard Brownian bridge.*

We can verify that $X(t) = W(t) - tW(1)$ is a standard Brownian bridge by calculating its mean and covariances.

Again like we did in the last lecture, we can compute the hitting time of a standard Brownian bridge using the principle of reflection.

Example 2 (Hitting Time in a Brownian Bridge)**.** *Let $\{W(t)\}_{t\ge 0}$ be a standard Brownian motion. Let $\tau_b \triangleq \inf\{t \ge 0 \mid W(t) > b\}$. Then we compute $\mathbf{Pr}[\tau_b < u \mid W(u) = x]$. Note that if $b < x$, $\mathbf{Pr}[\tau_b < u \mid W(u) = x] = 1$. Let ψ be the probability density function of standard Gaussian distribution, that is, $\psi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{x^2}{2}}$. If $b > x$, letting $\mathrm{d}x$ be an infinitesimal change, we have*

$$\begin{aligned}\mathbf{Pr}[\tau_b < u \mid W(u) = x] &= \frac{\mathbf{Pr}[\tau_b < u \land W(u) \in [x, x+\mathrm{d}x]]}{\mathbf{Pr}[W(u) \in [x, x+\mathrm{d}x]]}\\ &= \frac{\mathbf{Pr}[\tau_b < u] \cdot \mathbf{Pr}[W(u) \in [x, x+\mathrm{d}x] \mid \tau_b < u]}{f_u(x)\,\mathrm{d}x}.\end{aligned}$$

If we have known the value of τ_b and $\tau_b < u$, we can regard $\{W(u)\}_{t \ge \tau_b}$ as a Brownian motion starting from b. Then we have

$$\begin{aligned}\mathbf{Pr}[\tau_b < u] \cdot \mathbf{Pr}[W(u) \in [x, x+\mathrm{d}x] \mid \tau_b < u] &= \mathbf{Pr}[\tau_b < u] \cdot \mathbf{Pr}[W(u) \in [2b-x-\mathrm{d}x, 2b-x] \mid \tau_b < u]\\ &= \mathbf{Pr}[\tau_b < u \land W(u) \in [2b-x-\mathrm{d}x, 2b-x]]\\ &= \mathbf{Pr}[W(u) \in [2b-x-\mathrm{d}x, 2b-x]]\\ &= f_u(2b-x)\,\mathrm{d}x\end{aligned}$$

Thus, when $b > x$, $\mathbf{Pr}[\tau_b < u \mid W(u) = x] = \frac{f_u(2b-x)}{f_u(x)} = e^{-\frac{2b(b-x)}{u}}$.

When $b = x$, we have

$$\mathbf{Pr}[\tau_b < u \mid W(u) = b] = \frac{\mathbf{Pr}[\tau_b < u \land W(u) \in [b, b+\mathrm{d}b]]}{\mathbf{Pr}[W(u) \in [b, b+\mathrm{d}b]]}.$$

Note that

$$\mathbf{Pr}[\tau_b < u \land W(u) \in [b, b+\mathrm{d}b]] = \mathbf{Pr}[\tau_b < u] - \mathbf{Pr}[\tau_b < u \land W(u) > b+\mathrm{d}b] - \mathbf{Pr}[\tau_b < u \land W(u) < b]. \tag{1}$$

We know that $\mathbf{Pr}[\tau_b < u] = 2\left(1 - \Phi\left(\frac{b}{\sqrt{u}}\right)\right)$. Note that

$$\begin{aligned}\mathbf{Pr}[\tau_b < u \land W(u) > b+\mathrm{d}b] &= \mathbf{Pr}[W(u) > b+\mathrm{d}b]\\ &= 1 - \Phi\left(\frac{b}{\sqrt{u}}\right) - \mathbf{Pr}[W(u) \in [b, b+\mathrm{d}b]].\end{aligned}$$

And

$$\begin{aligned}\mathbf{Pr}[\tau_b < u \land W(u) < b] &= \mathbf{Pr}[\tau_b < u] \cdot \mathbf{Pr}[W(u) < b \mid \tau_b < u]\\ &= \frac{1}{2} \cdot \mathbf{Pr}[\tau_b < u] = 1 - \Phi\left(\frac{b}{\sqrt{u}}\right).\end{aligned}$$

Thus, Equation (1) equals to $\mathbf{Pr}[W(u) \in [b, b+\mathrm{d}b]]$ and

$$\mathbf{Pr}[\tau_b < u \mid W(u) = b] = 1.$$

3 Kolmogorov-Smirnov Test

In this section, we introduce an application of Brownian Bridge, the Kolmogorov-Smirnov test.

Suppose that $U_1, U_2, \dots, U_n$ are independently sampled from some distribution $[0,1]$ with CDF F. We would like to check if it is a uniform distribution, i.e., if the F satisfies $F(t) = t$ for every $t \in [0,1]$.

Let $\widehat{F_n}$ be the empirical cumulative distribution function, that is, for $t \in [0,1]$, $\widehat{F_n}(t) = \frac{1}{n}\sum_{i=1}^n \mathbf{1}[U_i \le t]$. It then follows from the law of large numbers that

$$\widehat{F_n}(t) \overset{n\to\infty}{\longrightarrow} \mathbf{E}\left[\widehat{F_n}(t)\right] = \frac{1}{n}\sum_{i=1}^n \mathbf{Pr}\left[U_i \le t\right] = F(t).$$

The idea of Kolmogorov-Smirnov test is to monitor the variable $\widehat{F_n}(t) - t$ for every $t \in [0,1]$ and reject the uniformity hypothesis if there exists some t that $\left|\widehat{F_n}(t) - t\right|$ is large. Then our goal is to find a suitable rejection threshold b such that if F is indeed a uniform distribution, the failure probability $\lim_{n\to\infty} \mathbf{Pr}\left[\max_{t\in[0,1]}\left|\widehat{F_n}(t) - t\right| \ge b\right]$ is sufficiently small (i.e., $\le \frac{1}{100}$). If F is a uniform distribution, for a fixed t, we have

$$\begin{aligned}
\mathbf{E}\left[\widehat{F_n}(t)\right] &= F(t) = t; \\
\mathbf{Var}\left[\widehat{F_n}(t)\right] &= \frac{1}{n^2}\sum_{i=1}^n \mathbf{Var}\left[\mathbf{1}[U_i \le t]\right] = \frac{1}{n}\cdot t(1-t).
\end{aligned}$$

Let $X_n(t) \triangleq \sqrt{n}\cdot(\widehat{F_n}(t) - t)$ for $t \in [0,1]$. By the Central Limit Theorem, we have $X_n(t) \sim \mathcal{N}(0, t(1-t))$ when $n \to \infty$. For any $0 \le s \le t \le 1$,

$$\begin{aligned}
\mathrm{Cov}(X_n(s), X_n(t)) &= n\cdot \mathrm{Cov}\left(\widehat{F_n}(s) - s, \widehat{F_n}(t) - t\right) \\
&= \frac{1}{n}\mathrm{Cov}\left(\sum_{i=1}^n \mathbf{1}[U_i \le s], \sum_{i=1}^n \mathbf{1}[U_i \le t]\right) \\
&= \mathrm{Cov}\left(\mathbf{1}[U_1 \le s], \mathbf{1}[U_1 \le t]\right) \\
&= \mathbf{Pr}\left[U_1 \le s, U_1 \le t\right] - \mathbf{Pr}\left[U_1 \le s\right]\mathbf{Pr}\left[U_1 \le t\right] \\
&= s(1-t).
\end{aligned}$$

For any $0 \le t_1 \le t_2 \le \dots \le t_k \le 1$, let $\Sigma = \left(\mathrm{Cov}\left(X_n(t_i), X_n(t_j)\right)\right)_{i,j}$. It follows from the high-dimensional Central Limit Theorem that

$$(X_n(t_1), X_n(t_2), \dots, X_n(t_k))^T \overset{D}{\longrightarrow} \mathcal{N}(\mathbf{0}, \Sigma) \sim (X(t_1), X(t_2), \dots, X(t_k))^T,$$

where $\{X(t)\}$ is a standard Brownian Bridge. Then using the result in Example 2, we have

$$\begin{aligned}
\lim_{n\to\infty} \mathbf{Pr}\left[\max_{t\in[0,1]} \widehat{F_n}(t) - t \ge b\right] &= \mathbf{Pr}\left[\max_{t\in[0,1]} X(t) \ge \sqrt{n}b\right] \\
&= \mathbf{Pr}\left[\tau_{\sqrt{n}b} < 1 \,\middle|\, W(1) = 0\right] = \exp\left\{-2nb^2\right\}.
\end{aligned}$$