Statistics 212: Lecture February 24, 2025

Roughness of Brownian Motion

Instructor: Mark Sellke

Scribes: Alex Meterez, Clara Mohri

1 Last time

Last time, we constructed Brownian motion on $[0,1]$. For a random continuous function $B(t)$ (formally, we have a probability measure on $C([0,1])$ on the Borel σ-field), it is a Brownian motion if it satisfies the following two properties:

(a) $B(t) - B(s) \sim \mathcal{N}(0, t-s)$ for all $0 \le s \le t \le 1$.

(b) If $t_1, \ldots, t_k$ is an increasing sequence, then the increments

$$(B(t_1), B(t_2) - B(t_1), \ldots, B(t_k) - B(t_{k-1}))$$

are independent.

In this lecture, we show the uniqueness of Brownian motion, extend Brownian motion to $[0,\infty)$, and show further properties of Brownian motion.

2 Uniqueness

Theorem 2.1 (Uniqueness of Brownian Motion)**.** *There is only one probability measure on* $C([0,1])$ *which obeys the properties of Brownian motion.*

Proof. We give a proof by contradiction. Suppose there exist $\mu \neq \mu'$ which both obey the properties of Brownian motion. Let $S = \{\text{Borel } A \subseteq C([0,1]) : \mu(A) = \mu'(A)\}$. We claim that S contains all “finite-dimensional" cylinders. To show this, for all $k \ge 1$, $A_1, \ldots, A_k \subseteq \mathbb{R}$ Borel, we define

$$\mathscr{C}_{t_1,\ldots,t_k,A_1,\ldots,A_k} = \{B : B(t_1) \in A_1, \ldots, B(t_k) \in A_k\}.$$

Further, we define

$$\mathscr{C} = \{\mathscr{C}_{t_1,\ldots,t_k,A_1,\ldots,A_k} : k \ge 1, t_1, \ldots, t_k \in [0,1], A_1, \ldots, B(t_k) \in A_k \subseteq \mathbb{R} \text{ Borel}\}.$$

Then, $\mathscr{C}$ is a π-system and S is a λ-system because it is closed under disjoint unions. Hence, due to the $\pi - \lambda$ theorem, S contains the σ-field generated by $\mathscr{C}$. As such, we have obtained a contradiction and it must be the case that $\mu = \mu'$. $\square$

3 Extending Brownian Motion to Infinity

We would like to show that Brownian motion can be extended to $[0,\infty)$. It suffices to "concatenate" several independent and identically distributed copies of Brownian motion. Suppose that $B^{(0)}(t)$ is a Brownian motion defined on $t \in [0,1]$. Then, for $i \geq 1$, we define $B^{(i)}(t)$ to be an independent and identically distributed copy of $B^{(0)}(t)$. Then, we define:

$$B(t) = \begin{cases} B^{(0)}(t) & \text{if } t \in [0,1] \\ \sum_{i=0}^{n-1} B^{(i)}(1) & \text{if } n \in \mathbb{N} \\ B(n) + B^{(n)}(\alpha) & \text{if } t = n+\alpha, \alpha \in (0,1), n \in \mathbb{N}. \end{cases}$$

$B(t)$ is thus defined on $[0,\infty)$ and satisfies the properties of Brownian motion. An alternate characterization of Brownian motion is that is it a centered Gaussian process with $\mathbb{E}[B(t)B(s)] = \min(t,s)$ for all times s,t. A Gaussian process means that each of the finite dimensional marginals $(B(t_1),\ldots,B(t_k))$ are jointly Gaussian.

3.1 Defining the distance metric

On $C([0,1])$, we used the distance metric $d_{\sup}(f,g) = \sup_{x\in[0,1]} \left|f(x)-g(x)\right|$. However, for B, B' iid, it is possible that extending this to $[0,\infty)$ gives $d_{\sup}(B,B') = \infty$. Hence, we define a new metric, starting with $d_{\sup}^{(n)}(B,B') = \sup_{0\leq t\leq n} \left|B(t)-B'(t)\right|$. Then

$$d(B,B') = \sum_{n=1}^{\infty} 2^{-n} \left(\frac{d_{\sup}^{(n)}(B,B')}{1+d_{\sup}^{(n)}(B,B')} \right) \leq 1.$$

Note that $d(B,B^{(m)}) \to 0 \iff d_{\sup}^{(n)}(B,B^{(m)}) \to 0 \,\forall n$ as $m\to\infty$. Hence, $d(\cdot)$ generates a Borel σ-algebra, $\sigma(\{B(t)\})_{t\in[0,\infty)}$. Furthermore, we remark that $C([0,\infty))$ is complete and separable with respect to this metric.

4 Invariance Properties of Brownian motion

In this section, we discuss three invariances of Brownian motion. We check the covariance condition of Brownian motion to show that each invariance holds.

4.1 Scale invariance

Fix $a > 0$. If B is a Brownian motion on $[0,\infty)$, then $X(t) = B(a^2 t/a)$ is a Brownian motion on $[0,\infty)$. We check the covariance condition:

$$\begin{aligned} \mathbb{E}[X(t)X(s)] &= \frac{1}{a^2}\mathbb{E}[B(a^2 t)B(a^2 s)] \\ &= \frac{1}{a^2}\min(a^2 t, a^2 s) \\ &= \min(t,s). \end{aligned}$$

4.2 Shift invariance

Fix $s > 0$. If B is a Brownian motion on $[0,\infty)$, then $X(t) = B(t+s) - B(s)$, $t \geq s$ is a Brownian motion on $[0,\infty)$. We check the covariance condition:

$$\begin{aligned} \mathbb{E}[X(t)X(r)] &= \mathbb{E}[(B(t+s)-B(s))(B(r+s)-B(s))] \\ &= \mathbb{E}[B(t+s)B(r+s) - B(t+s)B(s) - B(r+s)B(s) + B(s)B(s)] \end{aligned}$$

$$
\begin{aligned}
&= \min(t+s, r+s) - \min(t+s, s) - \min(r+s, s) + \min(s, s) \\
&= (\min(t, r) + s) - s - s + s \\
&= \min(t, r).
\end{aligned}
$$

Note: this further justifies the concatenation of Brownian motion to extend from $[0,1]$ to $[0,\infty)$. By starting the next Brownian motion interval at the place where the former interval ended, we are shifting the iid copy of Brownian motion. This shows that this concatenation is also a Brownian motion.

4.3 Time inversion

If B is a Brownian motion on $[0,\infty)$, then

$$
X(t) = \begin{cases} 0, & \text{if } t = 0 \\ tB(1/t) & \text{if } t > 0 \end{cases}
$$

is a Brownian motion on $[0,\infty)$. We check the covariance condition:

$$
\begin{aligned}
\mathbb{E}[X(t)X(s)] &= \mathbb{E}[tB(1/t) \cdot sB(1/s)] \\
&= ts \cdot \mathbb{E}[B(1/t)B(1/s)] \\
&= ts \cdot \min(1/t, 1/s) \\
&= \frac{ts}{\max(t,s)} \\
&= \min(t, s).
\end{aligned}
$$

The above holds because $\min(1/t, 1/s) = 1/t \iff t > s$. This implies that X and B have the same law as continuous functions $f : (0,\infty) \to \mathbb{R}$. Since continuity at 0 is a measurable event for such functions (e.g. it is equivalent to $\max_{q\in(0,1/n)\cap\mathbb{Q}} |f(q)| \to 0$ as $n \to \infty$), we also retain continuity at zero.

5 Roughness of Brownian Motion

Theorem 5.1 (Paley-Wiener-Zygmund, 1933)**.** *Almost surely, there does not exist $t \in [0,\infty)$ where $B'(t)$ exists. In fact, define*

$$
\overline{D}f(t) = \limsup_{h\downarrow 0} \frac{f(t+h) - f(t)}{h},
$$
$$
\underline{D}f(t) = \liminf_{h\downarrow 0} \frac{f(t+h) - f(t)}{h}.
$$

Then, almost surely, for all t, $\overline{D}B(t) = +\infty$ or $\underline{D}B(t) = -\infty$ or both.

Proof. We give a proof by contradiction. Suppose there exists t such that $\left|\overline{D}B(t)\right|, \left|\underline{D}B(t)\right| \le M < \infty$ for some constant M. Then,

$$
\overline{M} = \sup_{0\le n\le 1} \left| \frac{B(t+h) - B(t)}{h} \right| < \infty. \tag{1}
$$

This holds for small h because the term is less than $2M$, and for large h because we are locally bounded. In fact, we will show that (1) has probability zero to hold for any finite $\overline{M}$ simultaneously in t. More precisely, letting $A(\overline{M})$ be the event that (1) holds for at least 1 value of $t \in [0,1]$, we'll show that $\mathbb{P}[A(\overline{M})] = 0$. This implies the desired result by countable exhaustion over a sequence $\overline{M} \to \infty$, and the same argument for $t \in [1,2]$, $t \in [2,3]$, etc.

Note that by bundling all t into the single event $A(\overline{M})$, we avoid having to union-bound over uncountably many values of t in the latter exhaustion arguments.

For the main proof, fix n, and consider the 2^{-n} scale discretization of the real line. Then, we consider the nearby times $\frac{k-1}{2^n}, \frac{k}{2^n}, \frac{k+1}{2^n}, \frac{k+2}{2^n}$ where $t \in \left[\frac{k-2}{2^n}, \frac{k-1}{2^n}\right]$. Define the increments:

$$I_1 = B\left(\frac{k}{2^n}\right) - B\left(\frac{k-1}{2^n}\right), \qquad I_2 = B\left(\frac{k+1}{2^n}\right) - B\left(\frac{k}{2^n}\right), \qquad I_3 = B\left(\frac{k+2}{2^n}\right) - B\left(\frac{k+1}{2^n}\right).$$

Note that $I_1, I_2, I_3 \sim \mathcal{N}(0, 2^{-n})$ are IID. Given the constant $\overline{M}$, we have via Triangle Inequality:

$$\begin{aligned} |I_3| = \left|B\left(\frac{k+2}{2^n}\right) - B\left(\frac{k+1}{2^n}\right)\right| &\le \left|B\left(\frac{k+2}{2^n}\right) - B(t)\right| + \left|B\left(\frac{k+1}{2^n}\right) - B(t)\right| \\ &\le \overline{M}\left(\left|\frac{k+2}{2^n} - t\right| + \left|\frac{k+1}{2^n} - t\right|\right) \\ &\le 10\overline{M}2^{-n}. \end{aligned}$$

Using similar reasoning, we have that I_1, I_2 are also bounded above by $10\overline{M}2^{-n}$. Next, fixing k, n (denote that the defintions of I_1, I_2, I_3 depend on k, n):

$$\Pr[|I_1| \le 10\overline{M}2^{-n}] \le 100\overline{M}2^{-n/2},$$

as $I_1 \sim \mathcal{N}(0, 2^{-n})$ has standard deviation $2^{-n/2}$. Hence, by independence:

$$\Pr[|I_1|, |I_2|, |I_3| \le 10\overline{M}2^{-n}] \le 10^6\overline{M}^3 2^{-3n/2}.$$

We define $I_{k,n} = B\left(\frac{k}{2^n}\right) - B\left(\frac{k-1}{2^n}\right)$. Then, by a union bound over k:

$$\begin{aligned} \Pr[E_n(\overline{M})] &= \Pr[\exists\, k \text{ s.t. } \left|I_{k,n}\right|, \left|I_{k+1,n}\right|, \left|I_{k+2,n}\right| \le 10\overline{M}2^{-n}] \\ &\le 10^6\overline{M}^3 2^{-n/2}. \end{aligned}$$

Now, we have seen that **if** the event $A(\overline{M})$ defined above holds, **then** $E_n(\overline{M})$ holds for all n. However, for all $\overline{M} < \infty$, we have $\lim_{n\to\infty} \Pr[E_n(\overline{M})] = 0$. Hence, $A(\overline{M})$ has probability zero for any fixed $\overline{M}$, which completes the proof. □

Note that by considering more than 3 consecutive intervals, the same proof implies stronger "uniform local roughness" properties of Brownian motion.

6 Additional facts

At the end of class, we also mentioned some more difficult facts about the exact roughness of Brownian motion. There has been a lot of work on this (e.g. computing fractal dimensions of the sets of special points including the ones below).

(a) At a typical point, WLOG $t = 0$, the roughness is described by the law of the iterated logarithm:

$$\limsup_{\varepsilon \downarrow 0} \frac{|B(\varepsilon)|}{\sqrt{2\varepsilon \log\log 1/\varepsilon}} = 1 \iff \limsup_{t\to\infty} \frac{|B(t)|}{\sqrt{2t\log\log t}} \text{ (by inversion).}$$

See [JMNB14, KCG16, HRMS21] for some interesting applications of the second statement in statistics and machine learning.

In lecture, it was stated that $t \in \left[\frac{k-1}{2^n}, \frac{k}{2^n}\right]$ is contained in the first interval. However in the definition of $\overline{D}, \underline{D}$ we were only considering derivatives from the right with $h > 0$, so we actually need $t < \frac{k-1}{2^n}$. Note that requiring $h > 0$ just makes the divergence of $\max(\overline{D}, \underline{D})$ we showed slightly stronger.

(b) There exist fast points: there exists $t \in [0,1]$ such that

$$\limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{h \log(1/h)}} \in (0, \infty).$$

However no points are faster, i.e. the LHS is never infinity (see closely related extra credit problem on homework).

(c) There exist slow points: there exists $t \in [0,1]$ such that

$$\limsup_{h \downarrow 0} \frac{|B(t+h) - B(t)|}{\sqrt{h}} \in (0, \infty).$$

However no points are slower, i.e. the LHS is never zero.

References

[HRMS21] Steven R Howard, Aaditya Ramdas, Jon McAuliffe, and Jasjeet Sekhon. Time-uniform, nonparametric, nonasymptotic confidence sequences. *The Annals of Statistics*, 49(2):1055–1080, 2021. 4

[JMNB14] Kevin Jamieson, Matthew Malloy, Robert Nowak, and Sébastien Bubeck. lil'UCB: An Optimal Exploration Algorithm for Multi-Armed Bandits. In *Conference on Learning Theory*, pages 423–439. PMLR, 2014. 4

[KCG16] Emilie Kaufmann, Olivier Cappé, and Aurélien Garivier. On the complexity of best-arm identification in multi-armed bandit models. *The Journal of Machine Learning Research*, 17(1):1–42, 2016. 4