

FUNCTIONS OF BROWNIAN MOTION

JOHN B. WALSH

ABSTRACT. The class of continuous functions h for which h (Brownian motion) is a Markov process is determined.

Let B be a one-dimensional Brownian motion. For what continuous real valued functions h is $\{h(B_t), t \geq 0\}$ again a Markov process? If we require that $h(B_t)$ be again a Brownian motion, Dudley [1] has shown that h must be either x or $-x$; there should be a larger class of functions which transform Brownian motion into a Markov process. What is perhaps surprising is that the class is still relatively small.

To be definite, suppose B is defined canonically on the space Ω of all continuous real valued functions on $[0, \infty)$, i.e. $B = (\Omega, F_t, B_t, P^x)$, where B_t is the canonical process, $F_t = \sigma\{B_s, s \leq t\}$, and P^x is the distribution of Brownian motion with initial value x . When we say $h(B_t)$ is a Markov process, we mean that there is a transition function $Q(t, x, A)$ on $h(\mathbf{R})$ such that for any initial distribution, $\{h(B_t), t \geq 0\}$ is a Markov process with transition function Q . We let \mathcal{H} be the class of all continuous functions with this property.

There is a general criterion, due to Rosenblatt for Markov chains [4] (see also [2, p. 325]). A continuous function h is in \mathcal{H} iff for all x and y , $h(x) = h(y)$ implies that $P_t(x, h^{-1}(A)) = P_t(y, h^{-1}(A))$ for all $t \geq 0$ and Borel sets A . Equivalently,

(1) $h \in \mathcal{H}$ iff $h(x) = h(y)$ implies that the process $\{h(B_t), t \geq 0\}$ has the same distribution under P^y as under P^x .

It is clear that \mathcal{H} contains constants and strictly monotone functions. What is more interesting is that the remainder of \mathcal{H} is composed of two classes of functions:

(I) functions h having a unique maximum or minimum, say at x_0 . Then h is symmetric about x_0 and strictly monotone on either side;

(II) nonconstant functions h having more than one local extremum. In

Received by the editors September 5, 1973 and, in revised form, February 4, 1974.

AMS (MOS) subject classifications (1970). Primary 60J60, 60J65.

this case h is periodic, symmetric about any local extremum, and strictly monotone between adjacent maxima and minima.

There is a more compact description of \mathcal{H} . Let

$$\begin{aligned} f_1(x) &\equiv 0, & f_2(x) &= x, \\ f_3(x) &= |x|, & f_4(x) &= d(x, 2\mathbb{N}), \end{aligned}$$

i.e. $f_4(x)$ is the distance from x to the nearest even integer.

Theorem. \mathcal{H} is the class of functions h of the form

$$(2) \quad h(x) = g \circ f_i(ax + b), \quad 1 \leq i \leq 4,$$

where a and b are constants and g is continuous and strictly monotone.

We can dispatch three-quarters of the theorem directly. If h is of the form (2), $h(B_t)$ is a Markov process. Indeed, $f_3(B_t)$ is reflecting Brownian motion and $f_4(B_t)$ is Brownian motion on $[0, 1]$, reflecting at both endpoints, while a strictly monotone (hence one-to-one) function of a Markov process is again a Markov process. Conversely, constants and strictly monotone functions can certainly be written in the form (2). It remains to deal with functions which are neither constant nor strictly monotone—and which are consequently not one-to-one.

Lemma 1. *Let $h \in \mathcal{H}$. Then any local maximum (resp. minimum) of h is also a global maximum (resp. minimum). If h is constant on any interval, it is a constant function.*

Proof. Suppose $h(x_0) = b$, a local maximum of h . If $h(x) = h(x_0)$, then by (1)

$$\begin{aligned} P^{x_0}\{h(B_t) \leq b \text{ for all small enough } t\} \\ = P^{x_0}\{h(B_t) \leq b \text{ for all small enough } t\} = 1. \end{aligned}$$

This implies that x is also a local maximum of h . As h is continuous, we can conclude $h(x) \leq b$ for all x . Finally, if h is constant on some interval, it has points which are simultaneously local (hence global) maxima and minima. Thus $\inf h = \sup h$, and h is constant. Q.E.D.

A consequence of Lemma 1 is that if $h \in \mathcal{H}$ is nonconstant and has no global extremum in some interval, it is strictly monotone there.

Define $A_a = \{x: h(x) = a\}$, and let $d(x, A_a)$ be the distance from x to A_a .

Lemma 2. Let $h \in \mathcal{H}$. If $A_a \neq \emptyset$, there is a function u on $h(\mathbb{R})$ such that $d(x, A_a) = u \circ h(x)$ for all x .

Proof. Suppose $h(x) = h(x')$ and let I and I' be the maximal open subintervals of $\mathbb{R} - A_a$ which contain x and x' respectively. Write $I = (x - c, x + d)$ and $I' = (x' - c', x' + d')$. We define $T_a = \inf\{t: b(B_t) = a\}$. If $B_0 = x$, then $T_a = \inf\{t: B_t - B_0 = -c \text{ or } d\}$, and if $B_0 = x'$, then $T_a = \inf\{t: B_t - B_0 = -c' \text{ or } d'\}$. Since $h(x) = h(x')$, T_a has the same distribution under P^x as under $P^{x'}$. By Lemma 3 below, we conclude that

$$d(x, A_a) = \min(c, d) = \min(c', d') = d(x', A_a).$$

Thus we can set $u(h(x)) = \min(c, d)$. This proves the lemma.

Lemma 3. Let $0 < c \leq d$ be (possibly infinite) reals, and let $T = \inf\{t: B_t = -c \text{ or } d\}$. Suppose $B_0 = 0$. Then both c and d are determined by the distribution of T .

Proof. First, $c = d = \infty$ iff $P^0\{T = \infty\} = 1$, so we may as well suppose $c < \infty$. For the moment, we also suppose $d < \infty$. Let

$$f_\alpha(x) = E^x\{e^{-\alpha T}\}, \quad -c \leq x \leq d.$$

Then [3, p. 129] f_α satisfies

$$(3) \quad \frac{1}{2}f''_\alpha = \alpha f_\alpha, \quad f_\alpha(-c) = f_\alpha(d) = 1.$$

Solving this and setting $x = 0$:

$$(4) \quad f_\alpha(0) = (\sinh \sqrt{2\alpha}c + \sinh \sqrt{2\alpha}d) / \sinh \sqrt{2\alpha}(c + d).$$

Now the distribution of T determines $f_\alpha(0)$, and this in turn determines both c and d . Indeed

$$(5) \quad \begin{aligned} c &= - \lim_{\alpha \rightarrow \infty} (2\alpha)^{-1/2} \log f_\alpha(0), \\ d &= - \lim_{\alpha \rightarrow \infty} (2\alpha)^{-1/2} \log [f_\alpha(0) - \exp(-\sqrt{2\alpha}c)]. \end{aligned}$$

In fact, (5) holds even if $d = \infty$, for then the condition $f_\alpha(d) = 1$ is replaced by $0 \leq f_\alpha \leq 1$, and we get $f_\alpha(0) = \exp(-\sqrt{2\alpha}c)$.

We can now complete the proof of the Theorem. Suppose $h \in \mathcal{H}$ has a single local (hence global) extremum at x_0 . We may as well suppose x_0 is a minimum. Let $a = h(x_0)$. By Lemma 2 applied to $A_a = \{x_0\}$:

$$(6) \quad |x - x_0| = u \circ h(x).$$

But h is continuous and strictly monotone on $(-\infty, x_0]$ and on $[x_0, \infty)$ (for it is nowhere constant) so that u must also be continuous and strictly increasing on the range of h . In particular, it is invertible, so (6) gives us

$$(7) \quad h(x) = u^{-1} \circ f_3(x - x_0).$$

This leaves the case where h is nonconstant and has more than one local extremum. Let $a = \inf h < \sup h = b$. Noting that there is a local (hence global) minimum between any two local maxima, we may assume there exist x_0 and x_1 such that $h(x_0) = a$, $h(x_1) = b$, and h is strictly monotone between x_0 and x_1 . To fix our ideas, suppose $x_0 < x_1$. By Lemma 2

$$(8) \quad d(x, A_a) = u \circ h(x).$$

Consider A_a . $A_a \neq \{x_0\}$, for, as x_1 is a strict local maximum, we can find z and z' in some neighborhood $(x_1 - \epsilon, x_1 + \epsilon)$ such that $x_0 < z < x_1 < z'$ and $h(z) = h(z')$. Then $d(z', A_a) = d(z, A_a)$, which requires a point $x' > x_1$ in A_a . Evidently, $x' = z' + (z - x_0)$. Letting $\epsilon \rightarrow 0$ we see that $x' = x_0 + d$, where $d = 2|x_1 - x_0|$. Similarly, $x_1 - d \in A_b$. Repeating this argument, one sees that

$$A_a = \{x_0 \pm nd, n = 0, 1, 2, \dots\}, \quad A_b = \{x_1 \pm nd, n = 0, 1, 2, \dots\}.$$

Thus

$$(9) \quad d(x, A_a) = f_4((x - x_0)/d).$$

But both $h(x)$ and $f_4((x - x_0)/d)$ are strictly monotone on x_0, x_1 , hence u is strictly monotone on $[a, b]$, the range of h . In particular, u^{-1} exists, and (8) and (9) give

$$(10) \quad h(x) = u^{-1} \circ f_4((x - x_0)/d),$$

which completes the proof.

One can ask the more general question: "What Borel measurable functions h transform Brownian motion into a strong Markov process?" The function h may no longer have local maxima, but Lemma 2 remains valid (the continuity of h did not enter into its proof), and one can use it and arguments similar to those in its proof to determine the sets A_a directly. For example, if A_a is a singleton for all a , h must be one-to-one. If $A_a = \{x_1, x_2\}$ for some a , one argues using Lemma 2 that if $x_0 = (x_1 + x_2)/2$,

then $A_{h(x_0)} = \{x_0\}$, and that for any other c for which A_c is nonempty, $A_c = \{x_0 - y, x_0 + y\}$ for some $y > 0$. This would imply that h is of the form $g \circ f_3(x - x_0)$, where g is one-to-one. Similar arguments can be used to establish that if A_a contains more than two points, its elements form an arithmetic progression.

Thus no new ideas are involved in this more general question, but the details are both more complicated and less interesting than before. The answer is not, however. If we add the function $f_5(x) = x - [x]$ to our list, then if h is a real valued Borel measurable function, $h(B_t)$ will be a strong Markov process iff $h(x) = g \circ f_i(ax + b)$, $1 \leq i \leq 5$, where a and b are constants and g is one-to-one Borel measurable.

REFERENCES

1. R. M. Dudley, *Non-linear equivalence transformations of Brownian motion*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 20 (1971), 249–258. MR 46 #6485.
2. E. B. Dynkin, *Markov processes*, Fizmatgiz, Moscow, 1963; English transl., Vols. I, II, Die Grundlehren der math. Wissenschaften, Bände 121, 122, Academic Press, New York; Springer-Verlag, Berlin, 1965. MR 33 #1886; #1887.
3. K. Itô and H. P. McKean, *Diffusion processes and their sample paths*, Die Grundlehren der math. Wissenschaften, Band 125, Academic Press, New York; Springer-Verlag, Berlin, 1965. MR 33 #8031.
4. M. Rosenblatt, *Functions of a Markov process that are Markovian*, J. Math. Mech. 8 (1959), 665–681.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF BRITISH COLUMBIA, VANCOUVER 8, BRITISH COLUMBIA, CANADA