MARTINGALE REPRESENTATION THEOREM, AND THE LÉVY CHARACTERIZATION

NATE ELDREDGE

Header

General setup: Let B be Brownian motion on $(\Omega, \mathcal{F}, P)$ and $\{\mathcal{F}_t\}$ the filtration associated to B.

General note: We are proving one-dimensional versions of all these theorems. Multi-dimensional versions are also true and can usually be proved by similar arguments.

1. Martingale Convergence Theorem

Lemma 1.1. *Fix $T > 0$. Then the collection of random variables of the form*

$$\phi(B_{t_1}, \dots, B_{t_n}), \quad \phi \in C_C^\infty(\mathbb{R}^n), \quad n = 1, 2, \dots$$

is dense in $L^2(\mathcal{F}_T, P)$.

Proof. Choose a countable dense subset $\{t_1, t_2, \dots\}$ of $[0, T]$. Let $\mathcal{G}_n = \sigma(B_{t_1}, \dots, B_{t_n})$; then $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \dots$ and $\bigcup_n \mathcal{G}_n = \mathcal{F}_T = \sigma(B_t : t \in [0, T])$ by continuity of the filtration. So fix any $X \in L^2(\mathcal{F}_T, P)$. Then $M_n = E[X|\mathcal{G}_n]$ is a discrete time martingale, and $M_n \to X$ a.s. and in L^2. Now $M_n \in \mathcal{G}_n$; by the Doob-Dynkin lemma (Lemma 1.2 below), $M_n = g_n(B_{t_1}, \dots, B_{t_n})$ for some measurable $g_n : \mathbb{R}^n \to \mathbb{R}$. Moreover $g_n \in L^2(\mathbb{R}^n, \mu)$ where μ is the density of the random vector $(B_{t_1}, \dots, B_{t_n})$. So we can approximate g_n in L^2 by some $\phi_n \in C_C^\infty(\mathbb{R}^n)$, and then M_n is equally well approximated by $\phi_n(B_{t_1}, \dots, B_{t_n})$. Now we are done. □

Lemma 1.2 (Doob-Dynkin lemma)**.** *Let $\mathbf{X}$ be an n-dimensional random vector defined on $(\Omega, \mathcal{F}, P)$ and let Y be a $\sigma(\mathbf{X})$-measurable random variable defined on the same space. Then $Y = g(\mathbf{X})$ for some Borel measurable $g : \mathbb{R}^n \to \mathbb{R}$.*

Proof. Start with a characteristic function. Let $A \in \sigma(\mathbf{X})$; then $A = \mathbf{X}^{-1}(B)$ for some Borel set B. So $1_A(\omega) = 1_{\mathbf{X}^{-1}(B)}(\omega) = 1_B(\mathbf{X}(\omega))$ and the lemma holds for 1_A. So it also holds for simple r.v.s, and thus, by the monotone class theorem, for all $\sigma(\mathbf{X})$-measurable r.v.s. □

Lemma 1.3. *Random variables of the form*

$$\exp\left(\int_0^T h(t)\, dB_t - \frac{1}{2} \int_0^T h^2(t)\, dt \right), \tag{1}$$

where $h : [0, T] \to \mathbb{R}$ is a non-random step function of the form $h(t) = \sum_{i=1}^n \lambda_i 1_{[0,t_i]}$, $\lambda_i \in \mathbb{R}$, span a dense subspace of $L^2(\mathcal{F}_T, P)$.

Proof. First, observe that random variables of the form (1) are in fact in $L^2(\mathcal{F}_T, P)$. If $h(t) = \sum_{i=1}^n \lambda_i 1_{[0,t_i]}$, we have

$$E\left[\exp\left(\int_0^T h(t)\,dB_t - \frac{1}{2}\int_0^T h^2(t)\,dt\right)^2\right]$$
$$= \exp\left(-\int_0^T h^2(t)\,dt\right) E\left[\exp\left(2\int_0^T h(t)\,dB_t\right)\right].$$

However

$$E\left[\exp\left(2\int_0^T h(t)\,dB_t\right)\right] = E\left[\exp\left(2\int_0^T \sum_{i=1}^n \lambda_i 1_{[0,t_i]}(t)\,dB_t\right)\right]$$
$$= E\left[\exp\left(2\sum_{i=1}^n \lambda_i B_{t_i}\right)\right]$$

Since the B_{t_i} are joint normal random variables, a linear combination of them is normal, and so exp of it has finite expectation.

Let $Y \in L^2(\mathcal{F}_T, P)$ be orthogonal to all random variables of the form (1). Then if $h = \sum_{i=1}^n \lambda_i 1_{[0,t_i]}$, we have

$$G(\lambda) := E[\exp(\lambda_1 B_{t_1} + \cdots + \lambda_n B_{t_n})Y] = 0$$

for all λ, where we divided out the constant $\exp(-\int_0^T h^2(t)\,dt/2)$. I claim this equation also holds if I replace λ with $z = (z_1, \ldots, z_n) \in \mathbb{C}^n$. For observe that $G(z)$ is holomorphic (differentiate under the integral sign to verify the Cauchy-Riemann equations) and is zero on $\mathbb{R}^n$. Then it must be identically zero [1]. Now let $\phi \in C_C^\infty(\mathbb{R}^n)$; by Fourier inversion $\phi(x) = \int_{\mathbb{R}^n} \hat{\phi}(y) e^{ix\cdot y}\,dy$. So

$$E[Y\phi(B_{t_1}, \ldots, B_{t_n})] = E\left[Y \int_{\mathbb{R}^n} \hat{\phi}(y) \exp(i(B_{t_1}y_1 + \cdots + B_{t_n}y_n))\,dy\right]$$
$$= \int_{\mathbb{R}^n} \hat{\phi}(y) E[Y \exp(i(B_{t_1}y_1 + \cdots + B_{t_n}y_n))]\,dy$$
$$= 0$$

where Fubini's theorem applies because $Y \in L^1(\Omega), \hat{\phi} \in L^1(\mathbb{R}^n)$ and thus $Y\hat{\phi} \in L^1(\Omega \times \mathbb{R}^n)$. Therefore Y is orthogonal to all random variables of the form $\phi(B_{t_1}, \ldots, B_{t_n})$. These are dense by Lemma 1.1, so $Y = 0$. □

Theorem 1.4 (Itô Representation Theorem). *Let $F \in L^2(\mathcal{F}_T, P)$. Then we can write*

$$F = E[F] + \int_0^T f(t,\omega)dB_t$$

for some unique process f for which the integral makes sense.

[1] Suppose G is holomorphic on $\mathbb{C}^n$ and vanishes on $\mathbb{R}^n$. Write $z = x + iy$. Now for $x_0 \in \mathbb{R}^n \subset \mathbb{C}^n$, we have $\frac{\partial G}{\partial x}(x_0) = 0$. By the Cauchy-Riemann equations, we also have $\frac{\partial G}{\partial y}(x_0) = 0$. That is, $G'(x_0) = 0$. This is true for all $x_0 \in \mathbb{R}^n$, so G' vanishes on $\mathbb{R}^n$, likewise G'' and so on. Thus G vanishes with all its derivatives on $\mathbb{R}^n$; by expanding G in a Taylor series we see it must be identically zero on $\mathbb{C}^n$.

Proof. The integral makes sense for f with $E\left[\int_0^T f^2\,dt\right] < \infty$, i.e. $f \in L^2(P \times dt)$, which are adapted (or a.e. equal to such a process). Let $\mathcal{H}$ be the set of such f. I claim this is a closed subspace of $L^2(P \times dt)$. For say $f_n \in \mathcal{H}$, $f_n \to f$ in $L^2(P \times dt)$. Then there is a subsequence with $f_{n_k}(t, \omega) \to f(t, \omega)$ for a.e. t, ω. Specifically, for a.e. t, $f_{n_k}(t, \cdot) \to f(t, \cdot)$ a.s.; since each $f_{n_k}(t, \cdot)$ is in $\mathcal{F}_t$, so is $f(t, \cdot)$. Thus by fudging f on a dt-null set we can make it adapted, without changing its equivalence class in $L^2(P \times dt)$. So $\mathcal{H}$ is a closed subspace of the Hilbert space $L^2(P \times dt)$, hence a Hilbert space in its own right.

The Itô isometry says that the map

$$I : \mathcal{H} \to L^2(P)$$
$$I(f) = \int_0^T f(t, \omega)\,dB_t$$

is a (linear) isometry of Hilbert spaces. The claim at hand is essentially that I is onto. Well, almost; we will always have $E[I(f)] = 0$. But I claim $\operatorname{ran} I = \ker E$.

First, take F of the form

$$F = \exp\left(\int_0^T h(t)\,dB_t - \frac{1}{2}\int_0^T h^2(t)\,dt\right).$$

where $h : [0, T] \to \mathbb{R}$ is a step function, as in Lemma 1.3. Let

$$X_s = \int_0^s h(t)\,dB_t - \frac{1}{2}\int_0^s h^2(t)\,dt$$

and $Y_s = \exp(X_s)$, so that $Y_T = F$. Now

$$\begin{aligned} dY_s &= \exp(X_s)dX_s + \exp(X_s)(dX_s)^2 \\ &= Y_s h(s) dB_s - \frac{1}{2} Y_s h^2(s) ds + \frac{1}{2} Y_s h^2(s) ds \end{aligned}$$

Thus $Y_s = Y_0 + \int_0^s Y_t h(t)\,dB_t$, where $Y_0 = 1$. Why does this integral make sense? As previously argued, X_s is normal for each s, and thus Y_s is L^2 for each s. In fact, it is easily shown that $E[Y_s^2]$ is a continuous function of s, hence in $L^1([0, T])$. So $\int_0^T E[Y_s^2]\,ds < \infty$, i.e. $Y \in L^2(P \times dt)$. This suffices, since h is bounded.

We now have $I(Y) = \int_0^T Y_t h(t)\,dB_t = Y_T - Y_0 = F - 1$. Since $E[I(Y)] = 0$, we must have $E[F] = 1$.

Now let us think of expectation E as a bounded linear functional $E : L^2(P) \to \mathbb{R}$. Then $F \mapsto F - E[F]$ is orthogonal projection onto $\ker E := \{F : E[F] = 0\}$. So as F ranges over a dense subset of $L^2(P)$, $F - E[F]$ ranges over a dense subset of $\ker E$. Thus we have shown $\operatorname{ran} I$ is dense in $\ker E$. But an isometry of Hilbert spaces has closed range, thus $\operatorname{ran} I = \ker E$. Thus for any $F \in L^2(P)$, $F - E[F] \in \ker E$ and so $F - E[F] = I(f)$ for some f. Uniqueness follows because an isometry of Hilbert spaces is one-to-one. □

Theorem 1.5 (Martingale Representation Theorem)**.** *Suppose M is a continuous square-integrable $\{\mathcal{F}_t\}$-martingale. Then*

$$M_t = E[M_0] + \int_0^t g(s,\omega)\,dB_s, \quad t > 0. \tag{2}$$

for some adapted process g such that $\int_0^t E[g^2]\,ds < \infty$ for every $t > 0$. (In this way the integral in (2) makes sense for each t.)

Proof. By the Itô representation theorem, for each t there exists a unique $f^{(t)} \in L^2([0,t] \times \Omega, dt \times P)$ with

$$M_t = E[M_t] + \int_0^t f^{(t)}(s,\omega)\,dB_s.$$

Now $E[M_t] = E[M_0]$ for all t. Moreover, for $u \le t$ we have

$$M_u = E[M_0] + \int_0^u f^{(u)}(s,\omega)\,dB_s$$

and also

$$M_u = E[M_t|\mathcal{F}_u] = E[M_0] + \int_0^u f^{(t)}(s,\omega)\,dB_s$$

since we know $\int_0^t \cdots dB_s$ is an $\{\mathcal{F}_t\}$-martingale. By the Itô isometry, the stochastic integral is 1-1, and so we have $f^{(t)}(s,\omega) = f^{(u)}(s,\omega)$ for a.s. ω and a.e. $s \in [0,u]$. Thus let $g(s,\omega) = f^{(t)}(s,\omega)$ for any $t \ge s$; it is well defined, and we get the desired result. □

Using this theorem, we can characterize Brownian motion as a martingale with a specific quadratic variation.

2. Lévy Characterization of Brownian Motion

Definition 1. Let M_t be a continuous square-integrable martingale. Then $[M]_t$ is the (unique) process such that $M_t^2 - [M]_t$ is also a martingale.

By the Martingale Representation Theorem, we can write

$$M_t = M_0 + \int_0^t v_s\,dB_s.$$

Then by Itô's formula,

$$M_t^2 = \underbrace{M_0^2 + 2\int_0^t M_s v_s\,dB_s}_{\text{martingale}} + \int_0^t |v_s|^2\,ds.$$

Thus $[M]_t = \int_0^t |v_s|^2\,ds$.

Lemma 2.1. *Let W be a one-dimensional Itô process of the form*

$$W_t = W_0 + \int_0^t v_s \, dB_s, \quad t \geq 0 \tag{3}$$

such that

$$[W]_t := \int_0^t |v_s|^2 \, dt = t, \quad \forall t \geq 0. \tag{4}$$

Then W is one-dimensional Brownian motion.

Proof. First note that W_0 must obviously be a constant for the lemma to hold. Take $W_0 = 0$ for simplicity.

Observe for reference that by the fundamental theorem of calculus for L^1 functions, we can differentiate (4) to see that $|v_t|^2 = 1$ for a.e. t.

By definition of the Itô integral, W_t is adapted and has continuous sample paths. We will show that $W_t \sim N(0, t)$ for each t and that W has independent increments.

In fact, we can show both facts simultaneously. Fix $T < U$; we compute the joint distribution of $(W_T, W_U - W_T)$ via the characteristic function (Fourier transform). Set $X_t = W_{t \wedge T}$, $Y_t = W_t - W_{t \wedge T}$. Then

$$X_t = \int_0^{t \wedge T} v_s \, dB_s = \int_0^t 1_{\{s \leq T\}} v_s \, dB_s,$$
$$Y_t = \int_{t \wedge T}^t v_s \, dB_s = \int_0^t 1_{\{s > T\}} v_s \, dB_s$$

so that X, Y are Itô processes. Note that $X_U = W_T$, $Y_U = W_U - W_T$.

Now fix $\xi, \eta \in \mathbb{R}$ and let

$$g(x, y) = e^{i(\xi x + \eta y)}.$$

Let g_t be the process $g(X_t, Y_t)$. g_t is bounded, in fact $|g_t| = 1$ for all t, and $g_0 = 1$ since $X_0 = Y_0 = 0$. We will denote the partial derivatives of g with superscripts, so $g_t^x = \frac{\partial g}{\partial x}(X_t, Y_t) = i\xi g(X_t, Y_t) = i\xi g_t$. Likewise $g_t^{xx} = -\xi^2 g_t$, $g_t^{xy} = -\xi\eta g_t$, etc.

We will compute $\phi(t) := E[g(X_t, Y_t)]$. When we fix t and allow ξ, η to vary, $\phi(t)$ is the characteristic function of (X_t, Y_t), which is the characteristic function of (X_t, Y_t). Note that by path continuity of X, Y, continuity of g and the bounded convergence theorem, ϕ is a continuous function of t.

By Itô's formula, applied to the real and imaginary parts of g, we have

$$\begin{aligned}
dg(X_t, Y_t) &= g_t^x v_t\, dX_t + g_t^y v_t\, dY_t \\
&\quad + \frac{1}{2}\left(g_t^{xx}(dX_t)^2 + g_t^{xy} dX_t dY_t + g_t^{yx} dY_t dX_t + g_t^{yy}(dY_t)^2\right) \\
&= i\xi g_t 1_{\{t\le T\}} v_t dB_t + i\eta g_t 1_{\{t>T\}} v_t dB_t \\
&\quad + \frac{1}{2}\Bigg(-\xi^2 g_t (1_{\{t\le T\}})^2 \cancelto{1}{|v_t|^2} dt - \xi\eta g_t \cancelto{0}{1_{\{t\le T\}} 1_{\{t>T\}}} |v_t|^2\, dt \\
&\qquad - \eta\xi g_t \cancelto{0}{1_{\{t>T\}} 1_{\{t\le T\}}} |v_t|^2\, dt - \eta^2 g_t (1_{\{t>T\}})^2 \cancelto{1}{|v_t|^2} dt\Bigg) \\
&= i(\xi 1_{\{t\le T\}} + \eta 1_{\{t>T\}}) g_t v_t\, dB_t - \frac{1}{2}(\xi^2 1_{\{t\le T\}} + \eta^2 1_{\{t>T\}}) g_t\, dt
\end{aligned}$$

Thus

$$\begin{aligned}
\phi(t) = E[g_t] = E[g_0] &+ \cancelto{0}{E\left[\int_0^t i(\xi 1_{\{s\le T\}} + \eta 1_{\{s>T\}}) g_s v_s\, dB_s\right]} \\
&- \frac{1}{2} E \int_0^t (\xi^2 1_{\{s\le T\}} + \eta^2 1_{\{s>T\}}) g_s\, ds
\end{aligned}$$

where the second term vanishes because the integrand is bounded, thus in $L^2(P \times dt)$, and so the integral has zero expectation (Theorem 3.2.1 (iii) of Øksendal)

$$\begin{aligned}
&= 1 - \frac{1}{2}\int_0^t (\xi^2 1_{\{s\le T\}} + \eta^2 1_{\{s>T\}}) E[g_s]\, ds \qquad \text{by Fubini's theorem} \\
&= 1 - \frac{1}{2}\int_0^t (\xi^2 1_{\{s\le T\}} + \eta^2 1_{\{s>T\}}) \phi(s)\, ds
\end{aligned}$$

ϕ is continuous, as claimed above, so we can differentiate this using the (classical) fundamental theorem of calculus to obtain the ODE initial value problem

$$\begin{cases} \phi'(s) = -\frac{1}{2}(\xi^2 1_{\{s\le T\}}, \quad s \neq T + \eta^2 1_{\{s>T\}})\phi(s) \\ \phi(0) = 1 \end{cases} \tag{5}$$

(The derivative clearly does not exist at $s = T$, but this is not an obstruction.) For $s < T$ this becomes

$$\begin{cases} \phi'(s) = -\frac{1}{2}\xi^2 \phi(s), \quad s < T \\ \phi(0) = 1 \end{cases} \tag{6}$$

which we solve to find $\phi(s) = e^{-\xi^2 s/2}$. By continuity of ϕ the same formula is valid for $s = T$, so $\phi(T) = e^{-\xi^2 T/2}$. Then for $s > T$ (and, by continuity of ϕ, also $s \ge T$)

we have

$$\begin{cases} \phi'(s) = -\frac{1}{2}\eta^2\phi(s), & s > T \\ \phi(T) = e^{-\xi^2 T/2} \end{cases} \tag{7}$$

whose solution is

$$\phi(s) = e^{-(\xi^2-\eta^2)T/2}e^{-\eta^2 s/2} = e^{-(\xi^2 T+\eta^2(s-T))/2}. \tag{8}$$

Taking $s = U$ we find the joint characteristic function of $(W_T, W_U - W_T) = (X_U, Y_U)$ to be

$$E[e^{i(\xi W_T+\eta(W_U-W_T))}] = E[e^{i(\xi X_U+\eta Y_U)}] = \phi(U) = e^{-(\xi^2 T+\eta^2(U-T))/2}. \tag{9}$$

This is the chf of a joint normal with means 0, variances $(T, U - T)$ and covariance 0. Since W_T and $W_U - W_T$ are joint normal random variables which are uncorrelated, they are independent. Thus W satisfies $W_T \sim N(0, T)$ and has independent increments. This completes the proof that W is a Brownian motion. □

Corollary 2.2 (A form of the Lévy martingale characterization of Brownian motion)**.** *W_t is Brownian motion with respect to the filtration $\{\mathcal{F}_t\}$ iff W_t is a continuous square-integrable $\{\mathcal{F}_t\}$-martingale with $[W]_t = t$.*

As mentioned in class, this fact remains true with most of these conditions dropped. We can take $\{F_t\}$ to be any filtration, not necessarily the one associated to some pre-existing Brownian motion B_t; and we can assume that W_t is merely a local martingale which need not be square-integrable (suitably defining $[W]_t$).