8. GIRSANOV'S THEOREM

The stochastic calculus techniques we developed for continuous martingales did not make use any particular facts about their distributions. That being said, the connection with Brownian motion provided by Lévy characterization and other representation theorem revealed a prominent role of the Brownian motion, which is a Gaussian process. Here we will expand on the Gaussian nature of the standard Brownian motion further by proving Girsanov's theorem, which is a useful tool in stochastic analysis.

8.1 Exponential change of measure.

We begin by a list of motivational examples that will hopefully elucidate what Girsanov's theorem is really about. All of these will be based on a technique referred to as "exponential change of measure" which is a standard method to adjust the mean in large deviation theory. A simple form of that is the content of:

Lemma 8.1 (Exponential tilt) Let X_1, \ldots, X_n be independent and $\varphi_i(\lambda) := Ee^{\lambda X_i} < \infty$ for all $\lambda \in \mathbb{R}$ and all $i = 1, \ldots, n$. For $A \in \sigma(X_1, \ldots, X_n)$ and $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$, set

$$P_{\lambda}(A) := E\left(1_A \exp\left\{\sum_{i=1}^n \left[\lambda_i X_i - \log \varphi_i(\lambda_i)\right]\right\}\right)$$
(8.1)

Then (X_1, \ldots, X_n) remain to be independent under P_{λ} with expectation

$$E_{\lambda}(X_i) = \frac{\varphi_i'(\lambda_i)}{\varphi_i(\lambda_i)}$$
(8.2)

for all i = 1, ..., n.

Proof. The product structure of P_{λ} is verified directly from (8.1). For (8.2) the product structure shows $E_{\lambda}(X_i) = \varphi_i(\lambda_i)^{-1} E(X_i e^{\lambda X_i}) = \varphi'_i(\lambda_i) / \varphi_i(\lambda_i)$.

The point of the above lemma is that, since $\lambda \mapsto \varphi_i(\lambda)$ is strictly convex (under the assumption of having all exponential moments), its logarithmic derivative is strictly increasing. It follows that, by tuning λ_i appropriately, we can adjust the expectation of each X_i to whatever value in the interior of the convex hull of the support of X_i under P. The title of the lemma brings up another buzzword — exponential tilt — for the above technique as this is what the exponential factor — sometimes referred to as the "tilt" — does to the initial distribution.

For general underlying random variables, a drawback of the "tilting" technique is that is that varying λ_i is inevitably accompanied by changes to the whole distribution of the X_i 's. As our next lemma shows, this is not the case when Gaussian random variables are of concern as these are determined by only two parameters — the mean and covariance matrix.

Lemma 8.2 (Tilted Gaussian law) Let $X = (X_1, ..., X_n)$ be a multivariate normal with mean zero and covariance $C = {Cov(X_i, X_j)}_{i,j=1}^n$. For $\lambda = (\lambda_1, ..., \lambda_n) \in \mathbb{R}^n$ let

$$P_{\lambda}(A) := E\left(1_{A} e^{\lambda \cdot X - \frac{1}{2}\lambda \cdot C\lambda}\right)$$
(8.3)

Preliminary version (subject to change anytime!)

Typeset: April 18, 2024

$$X - C\lambda$$
 under $P_{\lambda} \stackrel{\text{law}}{=} X$ under P (8.4)

Here "·*" denotes the Euclidean inner product in* \mathbb{R}^{n} *.*

Proof. Let $t \in \mathbb{R}^n$. Then

$$E_{\lambda}(\mathbf{e}^{t\cdot(X-C\lambda)}) = E(\mathbf{e}^{t\cdot(X-C\lambda)+\lambda\cdot X-\frac{1}{2}\lambda\cdot C\lambda})$$

= $E(\mathbf{e}^{(\lambda+t)\cdot X-\frac{1}{2}(\lambda+t)\cdot C(\lambda+t)})\mathbf{e}^{\frac{1}{2}t\cdot Ct} = \mathbf{e}^{\frac{1}{2}t\cdot Ct}$ (8.5)

Since the latter is the Laplace transform of $\mathcal{N}(0, C)$, the claim follows using the Curtiss Theorem and the Cramér-Wold device.

The mean-zero restriction is made for convenience of expression; if the mean equals μ , then replace X by $X - \mu$ above. While the underlying Gaussian nature allowed us to treat a fully general case in one step, its special case of independent Gaussians could have also be dealt with inductively, by integrating out one variable at the time. This leads to another version of "exponential change of measure" which, this time, is very close to the one we are ultimately aiming for.

Lemma 8.3 (Discrete-time Girsanov Theorem) Let X_1, \ldots, X_n be independent with $X_i = \mathcal{N}(0, \sigma_i^2)$ for all $i = 1, \ldots, n$. For each $k = 1, \ldots, n$, let $\lambda_k \colon \mathbb{R}^{k-1} \to \mathbb{R}$ be a Borel-measurable function. (In particular, λ_1 is a constant.) Then

$$P_{\lambda}(A) := E\left(1_{A} \exp\left\{\sum_{k=1}^{n} \left[\lambda_{k}(X_{1}, \dots, X_{k-1})X_{k} - \frac{1}{2}\lambda_{k}(X_{1}, \dots, X_{k-1})^{2}\sigma_{k}^{2}\right]\right\}\right)$$
(8.6)

is a probability measure and

$$\left\{X_k - \sum_{j=1}^k \lambda_j(X_1, \dots, X_{j-1})\sigma_j^2\right\}_{k=1}^n \text{ under } P_\lambda \stackrel{\text{law}}{=} X \text{ under } P$$
(8.7)

Proof. For $k = 1, \ldots, n$, let

$$M_k := \prod_{j=1}^k e^{\lambda_j (X_1, \dots, X_{j-1}) X_j - \frac{1}{2} \lambda_j (X_1, \dots, X_{j-1})^2 \sigma_j^2}$$
(8.8)

and set $\mathcal{F}_k := \sigma(X_1, ..., X_k)$. We claim that $\{M_k\}_{k=1}^n$ is a martingale for filtration $\{\mathcal{F}_k\}_{k=1}^n$ under *P*. To see this note that

$$E(M_{k+1}|\mathcal{F}_k) = M_k E(e^{\lambda(X_1,\dots,X_k)X_{k+1} - \frac{1}{2}\lambda(X_1,\dots,X_k)^2 \sigma_{k+1}^2} | \mathcal{F}_k),$$
(8.9)

where the expectation on the right equals one by the fact that we can regard $\lambda(X_1, \ldots, X_k)$ as a constant under the conditional expectation. It follows that $E(M_k) = 1$ for all $k = 1, \ldots, n$ and since

$$\forall A \in \mathcal{F}_n: \quad P_\lambda(A) = E(1_A M_n), \tag{8.10}$$

we also get that P_{λ} is a probability measure.

Preliminary version (subject to change anytime!)

Typeset: April 18, 2024

MATH 285K notes

In order to prove (8.7), abbreviate

$$\widetilde{X}_k := X_k - \sum_{j=1}^k \lambda_j (X_1, \dots, X_{j-1}) \sigma_j^2$$
(8.11)

and, given any $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$, let

$$N_k := \prod_{j=1}^k e^{t_j \cdot \tilde{X}_j - \frac{1}{2}\sigma_j^2 t_j^2}$$
(8.12)

where, we note, we are using the shifted variables (8.11). A calculation shows that

$$N_k M_k = \prod_{j=1}^k e^{[t_j + \lambda_k(X_1, \dots, X_{j-1})]X_j - \frac{1}{2}[t_j + \lambda_j(X_1, \dots, X_{j-1})]^2 \sigma_j^2}$$
(8.13)

and so, by the same argument as above, also $\{N_k M_k\}_{k=1}^n$ is a martingale under *P*. It follows that

$$1 = E(N_0 M_0) = E(N_n M_n) = E_\lambda(N_n)$$
(8.14)

Using the explicit form of N_n , this can be written as

$$E_{\lambda} \mathbf{e}^{t \cdot \widetilde{X}} = \mathbf{e}^{\frac{1}{2}t \cdot \sigma^2 t} = E \mathbf{e}^{t \cdot X}$$
(8.15)

The Curtiss and Cramér-Wold theorems now imply the claim.

8.2 Statement and proof.

The ultimate result of this lecture is a continuous-time version of Lemma 8.3. The main difference is that here we have to assume that the tilted measure is a probability, rather than derive it as part of the proof.

Theorem 8.4 (Girsanov 1960) Assume a Brownian motion B adapted to a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ with \mathcal{F}_0 containing all P-null sets. For $Y \in \mathcal{V}_B^{\text{loc}}$ and $t \geq 0$ set

$$M_t := \exp\left\{\int_0^t Y_s dB_s - \frac{1}{2} \int_0^t Y_s^2 ds\right\}$$
(8.16)

and let

$$\forall A \in \mathcal{F}_t: \quad \widetilde{P}(A) := E(1_A M_t) \tag{8.17}$$

If $EM_t = 1$ *, then* \tilde{P} *is a probability measure and*

$$\left\{B_s - \int_0^s Y_u \mathrm{d}u \colon s \in [0, t]\right\} \text{ under } \widetilde{P}$$
(8.18)

is a standard Brownian motion.

We start by a lemma that reveals why the assumption $EM_t = 1$ is important:

Lemma 8.5 For *M* as in Theorem 8.4, if $EM_t = 1$ then $\{M_{s \wedge t} : s \ge 0\}$ is a martingale.

Preliminary version (subject to change anytime!)

Typeset: April 18, 2024

Proof. First note that, by the Itô formula,

$$dM_t = M_t (Y_t dB_t - \frac{1}{2}Y_t^2 dt) + \frac{1}{2}M_t Y_t^2 dt = M_t Y_t dB_t$$
(8.19)

and so $\{M_s : s \ge 0\}$ is a local martingale. Letting

$$\tau_n := \inf \left\{ u \ge 0 \colon \left| \int_0^u Y_s \mathrm{d}B_s \right| \ge n \right\}$$
(8.20)

the process { $M_{s \wedge \tau_n}$: $s \ge 0$ } is bounded by e^n and is thus a martingale with $EM_{s \wedge \tau_n} = EM_0 = 1$. The Optional Stopping Theorem then shows

$$\forall s \ge 0: \quad E(M_{t \wedge \tau_n} \mid \mathcal{F}_s) = M_{t \wedge s \wedge \tau_n} \quad \text{a.s.}$$
(8.21)

As $\tau_n \to \infty$ *P*-a.s. as $n \to \infty$ by the fact that $Y \in \mathcal{V}^{\text{loc}}$, we have $M_{u \wedge \tau_n} \to M_u$ pointwise a.s. Fatou's lemma then shows

$$\forall s \ge 0: \quad E(M_{t \land s}) \le 1 \tag{8.22}$$

while its conditional version gives

$$\forall s \ge 0: \quad E(M_t | \mathcal{F}_s) \le M_{t \land s} \quad \text{a.s.}$$
(8.23)

But the expectation of the left hand side equals $EM_t = 1$ which in conjunction with (8.22) forces $EM_{t \wedge s} = 1$ for all $s \ge 0$. It follows that $M_{t \wedge s \wedge \tau_n} \rightarrow M_{t \wedge s}$ both pointwise and in the mean, which implies that the convergence takes place in L^1 . Using this (8.21) shows that $\{M_{s \wedge t}: s \ge 0\}$ is a martingale, as claimed.

Note that the argument (8.22–8.23) shows that any non-negative supermartingale with constant expectation is a martingale. We are now ready to give:

Proof of Theorem 8.4. In order to prove (8.18), we will invoke a continuous-time version of the argument (8.12–8.15) but, since integrability is no longer automatic, we will rely on characteristic functions instead of Laplace transforms. Abbreviate

$$\widetilde{B}_s := B_s - \int_0^s Y_u \mathrm{d}u \tag{8.24}$$

where our assumptions allow us to take a continuous version of the stochastic integral on the right. Given any $Z \in \mathcal{V}$ bounded, set

$$N_s := \exp\left\{i\int_0^s Z_s d\widetilde{B}_s + \frac{1}{2}\int_0^s Z_s^2 ds\right\}.$$
(8.25)

Putting $M_s N_s$ together into one exponential as in (8.13), the Itô formula gives

$$\mathrm{d}M_s N_s = M_s N_s (Y_s + \mathrm{i}Z_s) \mathrm{d}B_s \tag{8.26}$$

and so $\{M_s N_s : s \ge 0\}$ is a local martingale.

Let τ_n be as in (8.20). Then for all $s \ge 0$ and $n \ge 1$,

$$E(M_{t \wedge \tau_n} N_{t \wedge \tau_n} | \mathcal{F}_s) = M_{s \wedge t \wedge \tau_n} N_{s \wedge t \wedge \tau_n} \quad \text{a.s.}$$
(8.27)

Since the proof of Lemma 8.5 shows that $\{M_{s \wedge t \wedge \tau_n} : n \ge 1\}$ is uniformly integrable, the fact that $\{N_s : s \le t\}$ is bounded also implies that $\{M_{s \wedge t \wedge \tau_n}, N_{s \wedge t \wedge \tau_n} : n \ge 1\}$ is uniformly

Preliminary version (subject to change anytime!)

Typeset: April 18, 2024

integrable. This upgrades pointwise convergence $M_{t \wedge \tau_n} \rightarrow M_t$ to L^1 -convergence and so we get

$$\forall s \ge 0: \quad E(M_t N_t \,|\, \mathcal{F}_s) = M_{s \wedge t} N_{s \wedge t} \quad \text{a.s.}$$
(8.28)

Applying this to s = 0 shows

$$1 = E(M_0 N_0) = E(M_t N_t) = \tilde{E}(N_t)$$
(8.29)

For the choice

$$Z_t := \sum_{j=1}^n \lambda_j \mathbf{1}_{(t_{j-1}, t_j]}(s)$$
(8.30)

where $0 = t_0 < t_1 < \cdots < t_n = t$ and $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, this becomes

$$\widetilde{E}\left(\exp\left\{i\sum_{j=1}^{n}\lambda_{j}(\widetilde{B}_{t_{j}}-\widetilde{B}_{t_{j-1}})\right\}\right)=\exp\left\{-\frac{1}{2}\sum_{j=1}^{n}\lambda_{j}^{2}(t_{j}-t_{j-1})\right\}$$
(8.31)

Using the Cramér-Wold device we get that, under \tilde{P} , the process \tilde{B} has the same finitedimensional distributions as B under P. Since \tilde{B} is continuous, it is a standard Brownian motion as claimed.

The above proof is somewhat subtle because it works under the minimal possible conditions. Indeed, a slight upgrade of Lemma 8.3 gives the statement for all Y simple without the proviso $EM_t = 1$ (which comes automatically in this case) so another strategy to address the general case could rely on approximation of Y by simple processes. The problem here is that, while the $L^2([0, t])$ -convergence $\int_0^t [Y_s^{(n)} - Y_s]^2 ds \rightarrow 0$ in probability implies that the associated exponential martingales $M_t^{(n)}$ converge to M_t in probability, we need $M_t^{(n)} \rightarrow M_t$ in L^1 which entails uniform integrability for which we have no direct argument.

Further reading: Karatzas-Shreve, Section 3.5