The Cameron-Martin-Girsanov (CMG) Theorem

There are many versions of the CMG Theorem. In some sense, there are many CMG Theorems. The first version appeared in [1] in 1944. Here we present a standard version, which gives the "spirit" of these theorems.

1   Introductory material

Let $(\Omega, \mathcal{F}, P)$ be a probability space. We can introduce other probability measures on $(\Omega, \mathcal{F})$ and obtain other probability spaces with the same sample space Ω and the same σ-algebra of events $\mathcal{F}$. One "easy" way to do that is to start with a random variable Θ of $(\Omega, \mathcal{F})$ such that

$$\Theta \geq 0 \quad P\text{-a.s.} \qquad \text{and} \qquad E_P[\Theta] = 1, \qquad (1.1)$$

where $E_P[\,\cdot\,]$ denotes the expectation associated to the measure P. Then, we can obtain a probability measure Q on $(\Omega, \mathcal{F})$ by setting

$$Q(A) = E_P[\Theta \mathbf{1}_A] \qquad \text{for all} \ \ A \in \mathcal{F}, \qquad (1.2)$$

where $\mathbf{1}_A$ is the indicator (function) of the event A.
It is easy to see that Q is a probability measure on $(\Omega, \mathcal{F})$ and, furthermore, that

$$Q(A) = 0 \qquad \text{whenever} \qquad P(A) = 0. \qquad (1.3)$$

Recall that if two measures P and Q are related as in formula (1.3), we say that Q is *absolutely continuous with respect to P* and denote this as

$$Q \ll P. \qquad (1.4)$$

In fact, the Radon-Nicodym Theorem states that the above construction of Q is the *only* way to obtain probability measures, which are absolutely continuous with respect to P, namely: If Q is a probability measure on $(\Omega, \mathcal{F})$ and $Q \ll P$, then there is a random variable Θ on $(\Omega, \mathcal{F})$ satisfying (1.1), for which $Q(A) = E_P[\Theta \mathbf{1}_A]$ for all $A \in \mathcal{F}$. The random variable Θ is called the *Radon-Nicodym derivative of Q with respect to P*. Symbolically,

$$\Theta = \frac{dQ}{dP}$$

and it is not hard to see that Θ is unique P-a.s.
Furthermore, if we also have that $P \ll Q$, then

$$\frac{dP}{dQ} = \frac{1}{\Theta}.$$

If $E_Q[\,\cdot\,]$ is the expectation associated to the measure Q of (1.2), then

$$E_Q[\mathbf{1}_A] = Q(A) = E_P[\Theta \mathbf{1}_A] \qquad \text{for all} \ \ A \in \mathcal{F},$$

and hence

$$E_Q[X] = E_P[\Theta X] \qquad \text{for any r.v. } X \text{ of } (\Omega, \mathcal{F}) \qquad (1.5)$$

(if $E_P[\Theta X]$ does not exist, then so does $E_Q[X]$).

The following proposition tells us how we can express conditional expectations associated to the measure Q of (1.2) in terms of conditional expectations associated to P.

Proposition 1. Let $\mathcal{B}$ be a subalgebra of $\mathcal{F}$. Then

$$E_Q[X \,|\, \mathcal{B}] = \frac{E_P[\Theta X \,|\, \mathcal{B}]}{E_P[\Theta \,|\, \mathcal{B}]} \quad Q\text{-a.s.} \qquad (1.6)$$

Proof. For typographical convenience let us put

$$\Psi := E_P[\Theta \,|\, \mathcal{B}]. \qquad (1.7)$$

Of course, $\Psi \geq 0$ P-a.s., $\Psi \in \mathcal{M}(\mathcal{B})$ (i.e. Ψ is $\mathcal{B}$-measurable), and

$$E_P[\Psi] = E_P\big[E_P[\Theta \,|\, \mathcal{B}]\big] = E_P[\Theta] = 1. \qquad (1.8)$$

Because of the above we can define another probability measure on $(\Omega, \mathcal{F})$ as

$$\tilde{Q}(A) = E_P[\Psi \mathbf{1}_A]. \qquad (1.9)$$

Observe that, for $A \in \mathcal{B}$ we have

$$Q(A) = E_P[\Theta \mathbf{1}_A] = E_P\big[E_P[\Theta \mathbf{1}_A \,|\, \mathcal{B}]\big] = E_P\big[\mathbf{1}_A E_P[\Theta \,|\, \mathcal{B}]\big] = E_P[\Psi \mathbf{1}_A],$$

that is (in view of (1.9))

$$Q(A) = \tilde{Q}(A) \qquad \text{for all } A \in \mathcal{B}, \qquad (1.10)$$

which also implies immediately that

$$E_Q[X] = E_{\tilde{Q}}[X] \qquad \text{for all } X \in \mathcal{M}(\mathcal{B}). \qquad (1.11)$$

We continue by setting

$$Y_1 := E_P[\Theta X \,|\, \mathcal{B}] \qquad \text{and} \qquad Y_2 := \Psi E_Q[X \,|\, \mathcal{B}]. \qquad (1.12)$$

Then, formula (1.6) is equivalent to

$$Y_1 = Y_2 \quad Q\text{-a.s.} \qquad (1.13)$$

Notice that $Y_1, Y_2 \in \mathcal{M}(\mathcal{B})$. Hence, in view of (1.10), to establish (1.13) it suffices to show that

$$E_{\tilde{Q}}[Y_1 \mathbf{1}_A] = E_{\tilde{Q}}[Y_2 \mathbf{1}_A] \qquad \text{for all } A \in \mathcal{B}. \qquad (1.14)$$

2

Now

$$E_{\tilde{Q}}[Y_1 \mathbf{1}_A] = E_P[\Psi Y_1 \mathbf{1}_A] = E_P\big[\Psi E_P[\Theta X \,|\, \mathcal{B}]\mathbf{1}_A\big]$$
$$= E_P\big[E_P[\Psi \Theta X \mathbf{1}_A \,|\, \mathcal{B}]\big] = E_P[\Psi \Theta X \mathbf{1}_A] = E_Q[\Psi X \mathbf{1}_A]$$

and (with the help of (1.11))

$$E_{\tilde{Q}}[Y_2 \mathbf{1}_A] = E_{\tilde{Q}}\big[\Psi E_Q[X \,|\, \mathcal{B}]\mathbf{1}_A\big] = E_{\tilde{Q}}\big[E_Q[\Psi X \mathbf{1}_A \,|\, \mathcal{B}]\big]$$
$$= E_Q\big[E_Q[\Psi X \mathbf{1}_A \,|\, \mathcal{B}]\big] = E_Q[\Psi X \mathbf{1}_A].$$

Therefore (1.14) is true and the proof is completed. ∎

Why do we need to consider more than one probability measures? First of all for educational purposes, since it helps to clarify certain concepts and issues. But, also, because such situations arise in applications. For example, two persons want to divide a pie (or a property) into two pieces so that each of them feels (s)he got a fair share. However, each person may have a different measure of fairness.

2 An example

A random variable X of $(\Omega, \mathcal{F})$ is a $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}$. Thus, X depends on Ω and $\mathcal{F}$. However, X *does not depend on the probability measure put on* $(\Omega, \mathcal{F})$. It is the *distribution* of X which depends on the measure. For instance, if P and Q are two probability measures on $(\Omega, \mathcal{F})$, then the distribution functions of X with respect to P and Q respectively are

$$F_P(x) = P\{X \le x\} \qquad \text{and} \qquad F_Q(x) = Q\{X \le x\}.$$

Of course, in general $F_P(x) \ne F_Q(x)$. Thus, e.g., the phrase "X is a normal random variable" (i.e. "X is a normally distributed random variable") may be misleading (or, at least, equivocal), if we work with more than one probability measures, since the distribution of X depends on the measure. The same remarks apply, of course, to random vectors and stochastic processes. For example, a process which is a Brownian motion with respect to a measure P, it will probably not be a Brownian motion with respect to another measure Q. Furthermore, if X and Y are random variables which are independent with respect to P, they may not be independent with respect to another measure Q. These issues do not come up if there is only one measure, but if there are more than one measures around, then they may become very important.

Let us look at a specific example (taken from [3]). Suppose we have a probability space $(\Omega, \mathcal{F}, P)$ and a random variable Z of this space, which

has the standard normal distribution (with respect to P). That is

$$P\{Z \le z\} = \Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\xi^2/2} d\xi, \qquad z \in \mathbb{R}. \qquad (2.1)$$

For typographical convenience we write $Z \underset{P}{\leadsto} N(0,1)$. Of course, there are many random variables of $(\Omega, \mathcal{F}, P)$ having the standard normal distribution (e.g., due to the symmetry of $\Phi(z)$, if Z is such a variable, then so is $-Z$). Recall that

$$E_P[Z] = 0, \qquad V_P[Z] = 1 \qquad (2.2)$$

(where $V_P[\,\cdot\,]$ is the variance with respect to P), and

$$E_P\left[e^{\mu Z}\right] = e^{\mu^2/2} \qquad \text{for any } \mu \in \mathbb{C}. \qquad (2.3)$$

Equation (2.3) can be written as

$$E_P\left[e^{\mu Z - \frac{1}{2}\mu^2}\right] = 1. \qquad (2.4)$$

Thus, for a Z as above and a $\mu \in \mathbb{R}$ we can introduce the probability measure Q defined by

$$Q(A) = E_P\left[e^{\mu Z - \frac{1}{2}\mu^2} \mathbf{1}_A\right], \qquad A \in \mathcal{F}. \qquad (2.5)$$

A question arising here is: What is the distribution of Z with respect to Q? We have

$$Q\{Z \le x\} = E_P\left[e^{\mu Z - \frac{1}{2}\mu^2} \mathbf{1}_{\{Z \le x\}}\right] = \int_{-\infty}^{\infty} e^{\mu z - \frac{1}{2}\mu^2} \mathbf{1}_{\{z \le x\}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\mu)^2}{2}} dz,$$

i.e., with respect to Q, Z is a normal random variable with mean μ and variance 1; symbolically, $Z \underset{Q}{\leadsto} N(\mu, 1)$, which can be written equivalently as $(Z - \mu) \underset{Q}{\leadsto} N(0,1)$, i.e. $Z - \mu$ has the standard normal distribution with respect to Q.

To extend this example, let the random variables $Z_1, \ldots, Z_n$ be independent and identically distributed with respect to P, all having the standard normal distribution. We now set

$$Q(A) = E_P\left[e^{\sum_{j=1}^{n} \mu_j Z_j - \frac{1}{2} \sum_{j=1}^{n} \mu_j^2} \mathbf{1}_A\right], \qquad A \in \mathcal{F}, \qquad (2.6)$$

where $\mu_1, \ldots \mu_n \in \mathbb{R}$ are given constants. It is not hard to see that Q is a probability measure on $(\Omega, \mathcal{F})$. Again, we would like to find the (joint) distribution of $Z_1, \ldots, Z_n$ with respect to Q.

For a "reasonable" (say bounded and continuous) function $g : \mathbb{R}^n \to \mathbb{R}$ we have (recall (1.5))

$$E_Q[g(Z_1, \ldots, Z_n)] = E_P\left[e^{\sum_{j=1}^n \mu_j Z_j - \frac{1}{2}\sum_{j=1}^n \mu_j^2} g(Z_1, \ldots, Z_n)\right]$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(z_1, \ldots, z_n) e^{\sum_{j=1}^n \mu_j z_j - \frac{1}{2}\sum_{j=1}^n \mu_j^2} \frac{1}{(2\pi)^{n/2}} e^{-\left(\frac{z_1^2}{2} + \cdots + \frac{z_n^2}{2}\right)} dz_1 \cdots dz_n$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(z_1, \ldots, z_n) \frac{1}{(2\pi)^{n/2}} e^{-\frac{(z_1-\mu)^2}{2}} \cdots e^{-\frac{(z_n-\mu)^2}{2}} dz_1 \cdots dz_n.$$

It follows that the joint probability density function of $Z_1, \ldots, Z_n$ (with respect to Q) is

$$f_Q(z_1, \ldots, z_n) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{(z_1-\mu)^2}{2}} \cdots e^{-\frac{(z_n-\mu)^2}{2}}, \qquad (2.7)$$

which tells us that, with respect to Q, the variables $Z_1, \ldots, Z_n$ are independent and furthermore $Z_j \overset{Q}{\rightsquigarrow} N(\mu_j, 1)$, i.e. $(Z_j - \mu_j) \overset{Q}{\rightsquigarrow} N(0, 1)$, for $j = 1, \ldots, n$. Thus, the random variable

$$\sum_{j=1}^n Z_j - \sum_{j=1}^n \mu_j \qquad (2.8)$$

is, with respect to Q, normally distributed with mean 0 and variance n.

Roughly speaking, the Cameron-Martin-Girsanov Theorem is a "continuous version" of the above simple example. In fact, having this example in mind, one can guess the statement of the CMG Theorem (see the remark after Theorem 1 in the next section).

3 The Cameron-Martin-Girsanov Theorem

3.1 CMG Theorem in $\mathbb{R}^1$

Consider a probability space $(\Omega, \mathcal{F}, P)$ and an one-dimensional Brownian motion process $\{B_t = B(t)\}_{t \geq 0}$ (with respect to P) with associated filtration $\{\mathcal{F}_t\}_{t \geq 0}$. Of course $\mathcal{F}_t$ is a subalgebra of $\mathcal{F}$ for every $t \geq 0$.

Next, let $\{X_t = X(t)\}_{t \geq 0}$ be a (real-valued) stochastic process, adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. We assume that for some constant $T > 0$ we have

$$\int_0^T X_t^2 dt < \infty \quad P\text{-a.s.} \qquad (3.1)$$

Having X_t we introduce

$$M_t := e^{Y_t}, \qquad \text{where} \quad Y_t := -\frac{1}{2}\int_0^t X_s^2 ds + \int_0^t X_s dB_s \qquad (3.2)$$

(notice that $M_t > 0$ P-a.s.). In differential notation we have

$$dY_t = -\frac{1}{2}X_t^2 dt + X_t dB_t, \qquad Y_0 = 0. \tag{3.3}$$

Applying Itô Calculus to (3.2) yields

$$dM_t = e^{Y_t} dY_t + \frac{1}{2}e^{Y_t}X_t^2 dt = M_t X_t dB_t, \qquad M_0 = 1. \tag{3.4}$$

Of course, (3.4) can be written equivalently as an integral equation

$$M_t = 1 + \int_0^t M_s X_s dB_s. \tag{3.5}$$

It follows that the Itô process M_t (being a stochastic integral) is an $\mathcal{F}_t$-martingale (with respect to P). In particular,

$$E_P[M_t] = E_P[M_0] = 1 \qquad \text{for all } t \in [0, T]. \tag{3.6}$$

Having M_t, and in view of (3.6), we introduce the probability measures on $(\Omega, \mathcal{F})$

$$Q(A) = E_P[M_T \mathbf{1}_A]; \qquad Q_t(A) = E_P[M_t \mathbf{1}_A], \quad 0 < t < T. \tag{3.7}$$

Observe that, for $0 \leq t \leq T$ we have $M_t = E_P[M_T \,|\, \mathcal{F}_t]$ (since M_t is a martingale). Hence, as in (1.10), we have

$$Q_t(A) = Q(A) \qquad \text{for all } A \in \mathcal{F}_t. \tag{3.8}$$

Theorem 1 (CMG in $\mathbb{R}^1$). Let $(\Omega, \mathcal{F}, P)$, B_t, X_t, $\mathcal{F}_t$, and T be as above. Set

$$W_t := B_t - \int_0^t X_s ds, \qquad t \in [0, T]. \tag{3.9}$$

Then, for any fixed $T > 0$ the process W_t, $0 \leq t \leq T$ is an $\mathcal{F}_t$-Brownian motion on $(\Omega, \mathcal{F}, Q)$ (i.e. with respect to Q).

Proof. We will prove the theorem with the help of the Lévy characterization of Brownian motion in $\mathbb{R}^1$:

A continuous stochastic process $W = \{W_t\}$ in a probability space $(\Omega, \mathcal{F}, Q)$ is an one-dimensional Brownian motion if and only if
(i) W_t is a martingale with respect to $\mathcal{F}_t^W$ and
(ii) $W_t^2 - t$ is a martingale with respect to $\mathcal{F}_t^W$.

The proof of this characterization will not be given here (it can be found, e.g., in [3]). If we accept this characterization of Brownian motion, then, in order to prove Theorem 1 we only need to check that W_t of (3.9) satisfies (i) and (ii).

Let us check (i). Set

$$K_t := M_t W_t, \tag{3.10}$$

where M_t is given by (3.2). Then, Itô Calculus, (3.4), and (3.9) give

$$dK_t = W_t dM_t + M_t dW_t + dW_t dM_t$$

$$= \left(B_t - \int_0^t X_s ds \right) M_t X_t dB_t + M_t \left(dB_t - X_t dt \right) + M_t X_t dt$$

$$= \left[\left(B_t - \int_0^t X_s ds \right) M_t X_t + M_t \right] dB_t, \tag{3.11}$$

which implies that $K_t = M_t W_t$ is an $\mathcal{F}_t$-martingale with respect to P. Now, for $0 \leq s \leq t \leq T$, by (3.8) and Proposition 1 we get

$$E_Q[W_t \mid \mathcal{F}_s] = E_{Q_t}[W_t \mid \mathcal{F}_s] = \frac{E_P[M_t W_t \mid \mathcal{F}_s]}{E_P[M_t \mid \mathcal{F}_s]} = \frac{M_s W_s}{M_s} = W_s, \quad Q\text{-a.s.}$$

which says that W_t is an $\mathcal{F}_t$-martingale with respect to Q.

Checking of (ii), namely that $W_t^2 - t$ is an $\mathcal{F}_t$-martingale with respect to Q, can be done in a similar way and is left as an exercise (**EXERCISE 1**). ∎

Remark. The quantity M_t of (3.2) can be viewed (at least in the case where X_t is a deterministic function of t) as a continuous analog of the quantity

$$e^{\sum_{j=1}^n \mu_j Z_j - \frac{1}{2} \sum_{j=1}^n \mu_j^2},$$

which appears in (2.6). Likewise, by writing $B_t = \int_0^t dB_s$ (assuming $B_0 = 0$) we can view the quantity W_t of (3.9) as a continuous analog of the quantity displayed in (2.8).

Theorem 2 (CMG in $\mathbb{R}^d$). Let $\{B(t)\}_{t \geq 0}$, be a d-dimensional Brownian motion on $(\Omega, \mathcal{F}, P)$, with associated filtration $\{\mathcal{F}_t\}_{t \geq 0}$. Also, let $\{X(t)\}_{t \geq 0}$ be a stochastic process with values in $\mathbb{R}^d$, adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$. We assume that for some constant $T > 0$ we have

$$\int_0^T |X_t|^2 dt < \infty \quad P\text{-a.s.} \tag{3.12}$$

Set

$$M_t := e^{Y_t}, \qquad \text{where} \quad Y_t := -\frac{1}{2} \int_0^t |X_s|^2 ds + \int_0^t X_s \cdot dB_s \tag{3.13}$$

(notice that $M_t > 0$ P-a.s.). It is not hard to see that M_t is an $\mathcal{F}_t$-martingale (with respect to P) and $E_P[M_t] = 1$ for all $t \in [0, T]$ (**EXERCISE 2**). If Q is the probability measure on $(\Omega, \mathcal{F})$ defined by

$$Q(A) = E_P[M_T \mathbf{1}_A], \tag{3.14}$$

7

then, for any fixed $T > 0$, the process

$$W_t := B_t - \int_0^t X_s ds, \qquad t \in [0, T] \qquad (3.15)$$

is a d-dimensional $\mathcal{F}_t$-Brownian motion on $(\Omega, \mathcal{F}, Q)$.

Proof. The proof is similar to the proof of Theorem 1 and it is left as an exercise (**EXERCISE 3**).

You may use (without proof) the Lévy characterization of Brownian motion in $\mathbb{R}^d$ [3]:

A continuous stochastic process $W = \{W(t)\}$ in a probability space $(\Omega, \mathcal{F}, Q)$ is an d-dimensional Brownian motion if and only if
(i) $W(t)$ is a martingale with respect to $\mathcal{F}_t^W$ (meaning that each component $W_j(t)$, $j = 1, \ldots, d$, is a martingale) and
(ii) for any $j, k \in \{1, \ldots, d\}$ we have that $W_j(t)W_k(t) - \delta_{jk}t$ is a martingale with respect to $\mathcal{F}_t^W$ (here δ_{jk} is the Kronecker delta). ∎

References

[1] R.H. Cameron and W.T. Martin, Transformation of Wiener integrals under translations, *Annals of Mathematics*, **45**, 386–396, 1944.

[2] R. Durrett, *Brownian Motion and Martingales in Analysis*, Wadsworth, Inc. Belmont, CA, USA, 1984.

[3] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Second Edition, Springer, New York, 1991.

[4] B. Øksendal, *Stochastic Differential Equations. An Introduction with Applications*, Sixth Edition, Springer-Verlag, Berlin, 2003.