Copyright © 2007 by Karl Sigman

1 IEOR 4700: Introduction to stochastic integration

1.1 Riemann-Stieltjes integration

Recall from calculus how the Riemann integral $\int_a^b h(t)dt$ is defined for a continuous function h over the bounded interval $[a,b]$. We partition the interval $[a,b]$ into n small subintervals $a = t_0 < t_1 < \cdots < t_n = b$, and sum up the area of the corresponding rectangles to obtain a *Riemann sum*;

$$\int_a^b h(t)dt \approx \sum_{j=1}^{n} h(t_{j-1})(t_j - t_{j-1}).$$

($h(t_{j-1})$ can be replaced by $h(s_j)$, for any $s_j \in [t_{j-1}, t_j]$.) As n gets larger and larger while the partition gets finer and finer, the approximating sum to the true area under the function becomes exact. This "dt" integration can be generalized to increments "$dG(t)$" of (say) any monotone increasing function $G(t)$ by using $G(t_j) - G(t_{j-1})$ in place of $t_j - t_{j-1}$ yielding the so-called Riemann-Stieltjes integral

$$\int_a^b h(t)dG(t) \approx \sum_{j=1}^{n} h(t_{j-1})(G(t_j) - G(t_{j-1})).$$

G need not be monotone to define such integration. The needed condition on a function G ensuring the existence of such integrals (via a limit of Riemann-Stieltjes sums) is that its *variation* be bounded (finite) over $[a,b]$:

$$V(G)[a,b] \stackrel{\text{def}}{=} lim_{n\to\infty} \sum_{j=1}^{n} |G(t_j) - Gt_{j-1})| < \infty,$$

where by taking the limit we mean taking a finer and finer partition of $[a,b]$. (And the answer can not depend on how we partition, as long as the subinterval lengths tend to 0.) Examples include $t_j = a + j(b-a)/n,\ j = 0,1,2,\ldots,n$, in which case each subinterval is of length $(b-a)/n$. If we imagine a particle whose $y-$axis position at time t is given by $G(t)$, then this variation simply measures the total $y-$axis distance (e.g., up-down distance) travelled during the time interval $[a,b]$. Any differentiable function with continuous derivative $g(t) = G'(t)$ has finite variation and in fact then $V(G)[a,b] = \int_a^b |g(t)|dt$, and integration reduces to $\int_a^b h(t)dG(t) = \int_a^b h(t)g(t)dt$; we then write "$dG(t) = g(t)dt$".

In this differentiable case, the variation is related to the *arclength* over $[a,b]$ given by $\int_a^b \sqrt{1 + (G'(t))^2}dt$.

(It can be proved that a function G has finite variation if and only if $G(t) = G_1(t) - G_2(t)$ where both G_1 and G_2 are monotone.)

Assume that G is of bounded variation and continuous (so it has no jumps). Let us focus for simplicity on the interval $[a,b] = [0,t]$. If $f = f(x)$ is a differentiable function with a continuous derivative $f'(x)$, then the differential of $f(G(t))$ can be dealt with by calculus yielding

$$df(G(t)) = f'(G(t))dG(t), \text{ differential form} \quad (1)$$

$$f(G(t)) = f(G(0)) + \int_0^t f'(G(s))dG(s), \text{ integral form.} \quad (2)$$

More generally, if $f = f(t,x)$ is a real-valued function of two variables (t for time, x for space) with both partial derivatives continuous, $\frac{\partial f}{\partial t}(t,x)$, $\frac{\partial f}{\partial x}(t,x)$, then the differential of $f(t, G(t))$ can be dealt with by calculus yielding

$$
\begin{aligned}
df(t, G(t)) &= \frac{\partial f}{\partial x}(t, G(t))dG(t) + \frac{\partial f}{\partial t}(t, G(t))dt, \text{ differential form} &(3)\\
f(t, G(t)) &= f(0, G(0)) + \int_0^t \frac{\partial f}{\partial x}(s, G(s))dG(s) + \int_0^t \frac{\partial f}{\partial t}(s, G(s))ds, \text{ integral form.} &(4)
\end{aligned}
$$

1.2 Brownian motion has paths of unbounded variation

It should be somewhat intuitive that a typical Brownian motion path can't possibly be of bounded variation; although continuous the paths seem to exhibit rapid infinitesimal movement up and down in any interval of time. This is most apparent when recalling how BM can be obtained by a limiting (in n) procedure by scaling a simple symmetric random walk (scaling space by $1/\sqrt{n}$ and scaling time by n): In any time interval, no matter how small, the number of moves of the random walk tends to infinity, while the size of each move tends to 0. This intuition turns out to be correct (we state without proof):

Proposition 1.1 *With probability* 1*, the paths of Brownian motion* $\{B(t)\}$ *are not of bounded variation, in fact they are differentiable nowhere:* $P(V(B)[0,t] = \infty) = 1$ *for all fixed* $t > 0$*, and* $P(B'(t)$ *does not exist at any value of* $t) = 1$.

As a consequence, we can not naively define sample-path by sample-path an integral, $\int_0^t h(s)dB(s)$, in the Riemann-Stieltjes sense.

Of course, since BM has continuous sample paths, we can use it as an integrand; $\int_0^t B(s)ds$ exits in the Riemann sense for example, it is called *integrated BM*.

1.3 Ito integration, Ito Calculus

To see what we can do to remedy the problem, let us try to make sense of

$$\int_0^t B(s)dB(s),$$

using approximating sums of the form

$$\sum_{k=1}^{n} B(t_{k-1})(B(t_k) - B(t_{k-1})),$$

where $0 = t_0 < t_1 < \cdots < t_n = t$ denotes a partition that becomes finer and finer as $n \to \infty$.

Using the identity

$$B(t_{k-1})(B(t_k) - B(t_{k-1})) = \frac{1}{2}(B^2(t_k) - B^2(t_{k-1})) - \frac{1}{2}(B(t_k) - B(t_{k-1}))^2$$

yields

$$\sum_{k=1}^{n} B(t_{k-1})(B(t_k) - B(t_{k-1}) = \frac{1}{2}B^2(t) - \frac{1}{2}\sum_{k=1}^{n}(B(t_k) - B(t_{k-1})^2.$$

The sum of squares $\sum_{k=1}^{n}(B(t_k)-B(t_{k-1}))^2$ is the key to understanding what to do here. As the partition gets finer and finer, it turns out that the limiting so-called *squared variation* exists and equals t, that is

$$Q(B)[0,t] \stackrel{\text{def}}{=} \lim_{n\to\infty}\sum_{k=1}^{n}(B(t_k)-B(t_{k-1}))^2 = t.^1.$$

To get some intuition: Since $B(t_k)-B(t_{k-1}) \sim N(0, t_k - t_{k-1})$, it follows that $E(B(t_k)-B(t_{k-1}))^2 = Var(B(t_k)-B(t_{k-1})) = t_k - t_{k-1}$ and thus

$$E\Big[\sum_{k=1}^{n}(B(t_k)-B(t_{k-1}))^2\Big] = \sum_{k=1}^{n}(t_k - t_{k-1}) = t.$$

Since this is true for any partition, it holds in the limit too. Heuristically, we write "$(dB(t))^2 = dt$". (It is important to realize that for continuous functions G of bounded variation, a quadratic variation term does not arise, it is 0; $(dG(t))^2 = 0$. For example, consider the function $G(t)=t$, and use the standard partition $t_k = kt/n$. Then $\sum_{k=1}^{n}(G(t_k)-G(t_{k-1}))^2 = \sum_{k=1}^{n}(t/n)^2 = t^2/n \to 0$, as $n\to\infty$; in other words $(dt)^2 = 0$.)

Thus we conclude that we have derived an interesting formula:

$$\int_0^t B(s)dB(s) = \frac{1}{2}B^2(t) - \frac{1}{2}t. \tag{5}$$

The standard calculus in (1) applied to $f(G(t))$ with $f(x)=x^2$ (and assuming $G(0)=0$) yields $d(G^2(t)) = 2G(t)dG(t)$ or

$$\int_0^t G(s)dG(s) = \frac{1}{2}G^2(t).$$

For example, if G is a differentiable function with a continuous derivative and $G(0)=0$ then $\int_0^t G(s)dG(s) = \int_0^t G(s)G'(s)ds = \frac{1}{2}G^2(t)$. (5) is different due to the non-zero quadratic variation; a second term gets included.

Unlike Riemann-Stieltjes integration, however, the above derivation of (5) fails if we choose a different value for $B(t_{k-1})$ in our approximating sums, for example if we use

$$\sum_{k=1}^{n} B(t_k)(B(t_k)-B(t_{k-1})),$$

or

$$\sum_{k=1}^{n} B(t_k^M)(B(t_k)-B(t_{k-1})),$$

where $t_k^M = t_{k-1} + (1/2)(t_k - t_{k-1})$ denotes the midpoint of the interval $[t_{k-1}, t_k]$. Each of these other schemes leads to different answers. It is the one we derived using the values $B(t_{k-1})$ for the increments $B(t_k)-B(t_{k-1})$, that is called the *Ito* integral. It generalizes to integrals of the form $\int_0^t X(s)dB(s)$ for appropriate stochastic processes $\{X(t) : t \geq 0\}$. For example, we can

[1] The convergence here, in general, is *in probability* or *in* L^2

take any continuous function $f = f(x)$ and define $\int_0^t f(B(s))dB(s)$ using the same method we did for deriving $\int_0^t B(s)dB(s)$, using sums

$$\sum_{k=1}^{n} f(B(t_{k-1}))(B(t_k) - B(t_{k-1})).$$

The Ito integral leads to a nice *Ito calculus* so as to generalize (1) and (3); it is summarized by *Ito's Rule:*

Ito's Rule

Proposition 1.2 *If $f = f(x)$ is a twice differentiable function with a continuous second derivative $f''(x)$, then*

$$df(B(t)) = f'(B(t))dB(t) + \frac{1}{2}f''(B(t))dt, \text{ differential form} \quad (6)$$

$$f(B(t)) = f(B(0)) + \int_0^t f'(B(s))dB(s) + \frac{1}{2}\int_0^t f''(B(s))ds, \text{ integral form.} \quad (7)$$

Proposition 1.3 *If $f = f(t,x)$ is a function such that both $\frac{\partial f}{\partial t}(t,x)$ and $\frac{\partial^2 f}{\partial x^2}(t,x)$ are continuous, then*

$$df(t, B(t)) = \frac{\partial f}{\partial x}dB(t) + (\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2})dt,$$

which in integral form is given by

$$f(t, B(t)) = f(0,0) + \int_0^t \frac{\partial f}{\partial x}(s, B(s))dB(s) + \int_0^t \frac{\partial f}{\partial t}(s, B(s))ds + \frac{1}{2}\int_0^t \frac{\partial^2 f}{\partial x^2}(s, B(s))ds.$$

Examples

1. Take $f(x) = x^2$. Then $f'(x) = 2x$ and $f''(x) = 2$ yielding $d(B^2(t)) = 2B(t)dB(t) + dt$ which in integral form is precisely (5).

2. Consider geometric BM of the form
   $S(t) = e^{B(t)} = f(B(t))$, where $f(x) = e^x$. Then $f'(x) = f''(x) = e^x$ and thus Ito's formula yields

   $$d(e^{B(t)}) = e^{B(t)}dB(t) + \frac{1}{2}e^{B(t)}dt,$$

   or, as a *stochastic differential equation (SDE)*

   $$dS(t) = S(t)dB(t) + \frac{1}{2}S(t)dt.$$

   In integral form we have

   $$\int_0^t e^{B(s)}dB(s) = e^{B(t)} - 1 - \frac{1}{2}\int_0^t e^{B(s)}ds.$$

Notice how this generalizes the standard

$$\int_0^t e^s ds = e^t - 1.$$

3. *BM with drift and variance term:* As our next example, take $f(t,x) = \sigma x + \mu t$, and let $X(t) = \sigma B(t) + \mu t$. Then $\frac{\partial f}{\partial x} = \sigma$, $\frac{\partial f}{\partial t} = \mu$ and $\frac{\partial^2 f}{\partial x^2} = 0$ yielding (in differential form) $dX(t) = \sigma dB(t) + \mu dt$

4. *Geometric BM :* As our last example, consider geometric BM of the form

   $S(t) = S_0 e^{\sigma B(t) + \mu t}$, where $S_0 > 0$ is a constant initial value.

   We use $f(t,x) = S_0 e^{\sigma x + \mu t}$ and observing that $\frac{\partial f}{\partial x} = \sigma f$, $\frac{\partial f}{\partial t} = \mu f$ and $\frac{\partial^2 f}{\partial x^2} = \sigma^2 f$, obtain

$$dS(t) = \sigma S(t) dB(t) + (\mu + \frac{1}{2}\sigma^2) S(t) dt.$$

   This is the classic stochastic differential equation (SDE) for geometric BM.

Sketch of a proof of Ito's formula $f = f(x)$ case:

The proof of Ito's formula relies on using the Taylor's series expansion

$$f(x+h) - f(x) = f'(x)h + \frac{1}{2}h^2 f''(x) + R(x,h),$$

where $R(x,h)$ is the remainder.We partition $[0,t]$ into n subintervals and use Taylor's formula on each subinterval;

$$\begin{aligned} f(B(t)) - f(0)) &= \sum_{k=1}^{n} (f(B(t_k) - f(B(t_{k-1})) \\ &= \sum_{k=1}^{n} f'(B(t_{k-1}))(B(t_k) - B(t_{k-1})) + \frac{1}{2}\sum_{k=1}^{n} f''(B(t_{k-1}))(B(t_k) - B(t_{k-1}))^2 + ERROR. \end{aligned}$$

As $n \to \infty$, the first sum converges to $\int_0^t f'(B(s))dB(s)$, the second to $\frac{1}{2}\int_0^t f''(B(s))ds$ (recall "$(dB(s))^2 = ds$ ") while the ERROR tends to 0; we have derived Ito's formula.

1.4 Ito processes

If we replace BM with another process $Y(t)$ and try to define yet again such integration as

$$\int_0^t Y(s) dY(s),$$

we once again can use the identity

$$Y(t_{k-1})(Y(t_k) - Y(t_{k-1})) = \frac{1}{2}(Y^2(t_k) - Y^2(t_{k-1})) - \frac{1}{2}(Y(t_k) - Y(t_{k-1}))^2$$

yielding

$$\sum_{k=1}^{n} Y(t_{k-1})(Y(t_k) - Y(t_{k-1})) = \frac{1}{2}Y^2(t) - \frac{1}{2}\sum_{k=1}^{n}(Y(t_k) - Y(t_{k-1}))^2.$$

As with BM, we thus need to be able to compute the squared variation:

$$Q(Y)[0,t] \stackrel{\text{def}}{=} \lim_{n\to\infty}\sum_{k=1}^{n}(Y(t_k) - Y(t_{k-1}))^2.$$

Then we would have a formula for $(dY(t))^2$, and obtain a "calculus"

$$df(Y(t)) = f'(Y(t))dY(t) + \frac{1}{2}f''(Y(t))(dY(t))^2.$$

If $Y(t)$ itself is related to BM via $dY(t) = H(t)dB(t) + K(t)dt$, we call it an *Ito process*, in which case it can be shown that $(dY(t))^2 = H^2(t)dt$.

A heuristic way to see this: squaring both sides of $dY(t) = H(t)dB(t) + K(t)dt$ and recalling that $(dB(t))^2 = dt$, while $(dt)^2 = 0$, yields
$(dY(t))^2 = H^2(t)dt + K^2(t)(dt)^2 + 2K(t)H(t)dB(t)dt$
$= H^2(t)dt + 2K(t)H(t)dB(t)dt$. But it can be shown that $dB(t)dt = 0$, and the result follows. Here, the point is that the *cross variation*

$$\lim_{n\to\infty}\sum_{k=1}^{n}(B(t_k) - B(t_{k-1}))(t_k - t_{k-1}) = 0.$$

We summarize with a general result:

Theorem 1.1 *Suppose $dY(t) = H(t)dB(t) + K(t)dt$ is an Ito process. If $f = f(t,x)$ is a function such that both $\frac{\partial f}{\partial t}(t,x)$ and $\frac{\partial^2 f}{\partial x^2}(t,x)$ are continuous, then for $Z(t) = f(t, Y(t))$,*

$$\begin{aligned} dZ(t) &= \frac{\partial f}{\partial x}dY(t) + \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}H^2(t)dt \\ &= \Big[\frac{\partial f}{\partial x}K(t) + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}H^2(t)\Big]dt + \frac{\partial f}{\partial x}H(t)dB(t). \end{aligned}$$

Application to geometric BM

As an application of Theorem 1.1, recall that geometric BM is indeed an Ito process: $dS(t) = \sigma S(t)dB(t) + (\mu + \frac{1}{2}\sigma^2)S(t)dt$; $H(t) = \sigma S(t)$ and $K(t) = (\mu + \frac{1}{2}\sigma^2)S(t)$; thus for $Z(t) = f(t, S(t))$, we have

$$dZ(t) = \frac{\partial f}{\partial x}dS(t) + \frac{\partial f}{\partial t}dt + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 S^2(t)dt \tag{8}$$

$$= \Big[\frac{\partial f}{\partial x}(\mu + \frac{1}{2}\sigma^2)S(t) + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 S^2(t)\Big]dt + \frac{\partial f}{\partial x}\sigma S(t)dB(t). \tag{9}$$

1.5 Derivation of the Black-Scholes PDE using Ito calculus

Theorem 1.2 (Black-Scholes Partial Differential Equation (PDE)) *Let $f(t,x)$ denote the price at time t of a European style derivative of stock (such as a European call option) with expiration date T, when $S(t) = x$, where $S(t) = S_0 e^{\sigma B(t)+\mu t}$ is geometric Brownian motion, modeling the price per share of the stock. Then f must satisfy the partial differential equation:*

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} rx + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 x^2 = rf.$$

Proof :

As in our proof under the binomial lattice model, we will replicate our option payoff by creating a portfolio of stock and risk-free asset, readjusting the portfolio over time. By doing this we will see that f must satisfy the PDE in question.

We have two continuous time processes.

1. The stock price per share: $S(t) = S_0 e^{\sigma B(t)+\mu t}$; the SDE for this geometric BM is $dS(t) = \overline{\mu} S(t)dt + \sigma S(t)dB(t)$, where $\overline{\mu} = \mu + \sigma^2/2$.
2. The risk-free asset (bond (say)) per share: $b(t) = e^{rt}$, a deterministic function of t where $r > 0$ is the interest rate; the differential equation for this asset is $db(t) = rb(t)dt$.

We wish to construct a portfolio $(\alpha(t), \beta(t))$, denoting the number of shares of stock and risk-free asset we have at any time $t \in [0,T]$, the value of which matches the option payoff. At any given time $t \in [0,T]$ the value of our portfolio is $V(t) = \alpha(t)S(t) + \beta(t)b(t)$. Our objective is to buy some initial shares at time $t = 0$, (α_0, β_0), and then continuously readjust our portfolio in such a way that $V(t) = f(t, S(t))$ for all $t \in [0,T]$, in particular the price of the option (at time $t = 0$) will be $C_0 = V(0) = \alpha_0 S_0 + \beta_0$.

We are not allowed to either insert or remove funds along the way, that is, we assume a *self-financing* strategy, mathematically enforced by assuming that

$$dV(t) = \alpha(t)dS(t) + \beta(t)db(t), \ \textit{self-financing condition.} \tag{10}$$

In integral form, this condition is given by

$$V(t) = \alpha_0 S_0 + \beta_0 + \int_0^t \alpha(s)dS(s) + \int_0^t \beta(s)db(s).$$

It simply ensures that any change in $V(t)$ must equal the profit or loss due to changes in the price of the two assets themselves. For example, if at time t we have $\alpha(t)S(t) = \$500$, then we can exchange $\alpha(t)/2$ shares of stock for \$250 worth of risk-free asset (so $V(t)$ does not change), but we can't just sell the $\alpha(t)/2$ shares and remove this money from the portfolio.

Using our differentials $dS(t)$ and $db(t)$ together with the self-financing condition yields

$$\begin{aligned} dV(t) &= \alpha(t)[\overline{\mu} S(t)dt + \sigma S(t)dB(t)] + \beta(t)rb(t)dt & (11)\\ &= (\alpha(t)\overline{\mu} S(t) + \beta(t)rb(t))dt + \alpha(t)\sigma S(t)dB(t). & (12)\end{aligned}$$

We are looking for factors $\alpha(t)$ and $\beta(t)$ such that $V(t) = \alpha(t)S(t)+\beta(t)b(t)) = f(t, S(t))$, $t \in [0,T]$. We will thus equate the two differentials $dV(t) = df(t, S(t))$ so as to figure out what these factors must be, and also see what properties f must satisfy.

To this end, using Ito's formula on $f(t, S(t))$ exactly as in (8) yields

$$
\begin{aligned}
df(t,S(t)) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dS(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(dS(t))^2 &(13)\\
df(t,S(t)) &= \frac{\partial f}{\partial t}dt + \frac{\partial f}{\partial x}dS(t) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 S^2(t)dt &(14)\\
&= \Big[\frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}\sigma^2 S^2(t) + \frac{\partial f}{\partial x}\overline{\mu}S(t)\Big]dt + \frac{\partial f}{\partial x}\sigma S(t)dB(t). &(15)
\end{aligned}
$$

Equating the $dB(t)$ coefficients from (12) and (15) yields

$$
\alpha(t) = \frac{\partial f}{\partial x}(t,S(t)).
$$

Similarly equating the two dt coefficients yields

$$
\beta(t) = \frac{1}{rb(t)}\Big(\frac{\partial f}{\partial t}(t,S(t)) + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(t,S(t))\sigma^2 S^2(t)\Big).
$$

Finally, plugging in these values for $\alpha(t)$ and $\beta(t)$ while equating $\alpha(t)S(t) + \beta(t)b(t) = f(t,S(t))$ then replacing $S(t)$ with x yields the Black-Scholes PDE as was to be shown. ■