Lecture Notes for Math 210 – 28 September 2007

Shannon Starr

28 September 2007

Chapter 3: Deterministic and Stochastic Calculus

§4.3.2 The Stieltjes integral : aka change-of-variables

We were in the middle of *proving* the change-of-variables formula.

Let us begin by stating it (which we didn't do last time):

Suppose f is an integrable function on $[a, b]$.

Suppose $g : [c, d] \to [a, b]$ is an infinitely differentiable function,

so $g'(y), g''(y), \dots$ all exist,

and suppose the Taylor series converges.

Also suppose that g is increasing, and $g(c) = a$, $g(d) = b$.

Then

$$\int_a^b f(x)\, dx \,=\, \int_c^d f(g(y))\, g'(y)\, dy\,.$$

Sometimes people write $dg(y)$ in place of $g'(y)\, dy$ to say

$$\int_a^b f(x)\, dx \,=\, \int_c^d f(g(y))\, dg(y)\,.$$

With this notation, one could also calculate

$$\int_c^d h(y)\,dg(y)\,,$$

for any integrable function h.

Just take

$$\int_c^d h(y)\,dg(y) \;=\; \int_c^d h(y)\,g'(y)\,dy\,.$$

Now let us finish the proof.

We had already done a lot of work to get

$$\int_a^b f(x)\,dx \;=\; \lim_{n\to\infty}\left[\sum_{k=1}^{n} f(g(y_k))\,g'(y_k)\,\Delta y_k - \sum_{k=1}^{\infty}\frac{1}{2}f(g(y_k))\,g''(y_k)\,(\Delta y_k)^2 + \ldots\right]$$

Also, by the definition of the integral,

$$\lim_{n\to\infty}\sum_{k=1}^{n} f(g(y_k))\,g'(y_k)\,\Delta y_k \;=\; \int_c^d f(g(y))\,g'(y)\,dy\,.$$

So

$$\int_a^b f(x)\,dx \;=\; \int_c^d f(g(y))\,g'(y)\,dy - \lim_{n\to\infty}\left[\sum_{k=1}^{\infty}\frac{1}{2}f(g(y_k))\,g''(y_k)\,(\Delta y_k)^2 - \ldots\right].$$

So what we still have to do is to prove that

$$\lim_{n\to\infty}\left[\sum_{k=1}^{\infty}\frac{1}{2}f(g(y_k))\,g''(y_k)\,(\Delta y_k)^2 - \ldots\right] \;=\; 0\,.$$

Q: Suppose $\Delta y_k = \Delta y$ for all k. What is $\lim_{n\to\infty}\Delta y$?

A: 0 because then $\Delta y = (d-c)/n \to 0$ as $n\to\infty$.

Q: What is $\lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{1}{2} f(g(y_k))\, g''(y_k)\, (\Delta y_k)^2$ equal to?

A: It is 0 because you have a squared-power of Δy_k in the summation. We know if we only had one, we would get

$$\lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{1}{2} f(g(y_k))\, g''(y_k)\, \Delta y_k \;=\; \int_c^d \frac{1}{2} f(g(y))\, g''(y)\, dy\,.$$

But since we have two, we use the product rule for limits.

Suppose $\Delta y_k = \Delta y$ for all k:

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{1}{2} f(g(y_k))\, g''(y_k)\, (\Delta y_k)^2 \;&=\; \lim_{n\to\infty} (\Delta y) \sum_{k=1}^{n} \frac{1}{2} f(g(y_k))\, g''(y_k)\, (\Delta y_k)^2 \\
&= \left(\lim_{n\to\infty} \Delta y\right) \left(\lim_{n\to\infty} \sum_{k=1}^{\infty} \frac{1}{2} f(g(y_k))\, g''(y_k)\, \Delta y_k\right) \\
&= 0 \cdot \int_c^d \frac{1}{2} f(g(y))\, g''(y)\, dy \\
&= 0\,.
\end{aligned}$$

Also, if you ever have higher power of Δy than 2, you also get 0.

So all the ... in the Riemann sum also go to zero.

Conclusion:

$$\int_a^b f(x)\, dx \;=\; \int_c^d f(g(y))\, g'(y)\, dy\,.$$

Two points: First, we will abbreviate one of the previous facts as $(dy)^2 = 0$.

By this we mean the following.

dy is really Δy, in the Riemann sum, in the limit that $\Delta y \to 0$.

Even though $\Delta y \to 0$, the number of terms in the sum is n and $n \to \infty$.

So the total Riemann sum does not go to zero:

$$\lim_{n\to\infty} \sum_{i=1}^{n} h(y_i)\,\Delta y \;=\; \int_c^d h(y)\,dy \;\neq\; 0\,,$$

generically.

But if you try to take something like $\int_c^d h(y)\,(dy)^2$,

then it is like you are putting $(\Delta y)^2$ in the Riemann sum and that does go to 0.

Formally,

$$\begin{aligned}
\int_c^d h(y)\,(dy)^2 \;&=\; (dy)\int_c^d h(y)\,dy && \text{because you only need 1 power of } dy \text{ in the integral} \\
& && \text{to cancel out the ``infinite sum'':} \int \\
&= 0\cdot\int_c^d h(y)\,dy && \text{because } dy \text{ is infinitesimal,} \\
& && \text{which means it is 0 when it appears by itself} \\
&= 0\,.
\end{aligned}$$

The formula I just wrote is *not rigorous.*

But algebraically, it mimics the Riemann sum calculation we just did.

For ordinary variables such as x and t, we have similar formulas: $(dx)^2 = (dt)^2 = 0$.

Second point: If we did not have $(dy)^2 = 0$, just theoretically, we would have had to have kept more than just the first term in the Taylor series.

We would have also obtained at least the term like $\int f(g(y))\,g''(y)\,(dy)^2$.

Of course this is ridiculous because we do know $(dy)^2 = 0$.

But, when we cover Brownian motion, we will have B_t being a non-differentiable function of t, and we will not have $(dB_t)^2 = 0$.

Just as a preview, in that calculation, we will have $(dB_t)^3 = 0$.

So we will have, for Brownian motion,

$$\int_{t_0}^{t_1} f(B_t)\, dg(B_t) \;=\; \int_{t_0}^{t_1} \left[f(B_t)\, g'(B_t) dt + \frac{1}{2} f(B_t)\, g''(B_t)\, (dB_t)^2 \right]$$

(The fact that we have a $+$ instead of a $-$ in front of the second term is not a typo, even though it seems to contradict our earlier calculation. It is because of a subtle but important point about Brownian motion, and stocks: we always have to choose $t_i^* \in [t_{i-1}, t_i]$ to be the left endpoint, $t_i^* = t_{i-1}$ instead of t_i, because that is the way to have a "non-anticipating" function. I.e., in the stock market, we cannot predict the future, not even an infinitesimal distance Δt into the future, so we have to take t_i^* to be the smallest time in the interval, t_{i-1}. Because of this, we take the Taylor series for $f(t_i)$ about t_{i-1}, instead of the other way around, and this gives the $+ \ldots (dB_t)^2$ instead of $-$.)

We have talked a lot about Brownian motion even though I still haven't defined what it is.

I am not expecting this to make perfect sense.

But I am hoping that some of the ideas will be familiar when we return to Brownian motion in several weeks.

§5. *Partial Derivatives*

Suppose you have a function $f(x, y)$.

The partial derivative $\frac{\partial f}{\partial x}(x, y)$ is the derivative with respect to x holding y fixed.

Formally,

$$\frac{\partial f}{\partial x}(x,y) \;=\; \lim_{h\to 0} \frac{f(x+h,y)-f(x,y)}{h}\,.$$

In other words, for each fixed value of y, define a new function $g(x) = f(x,y)$ for that fixed value of y.

Then $\frac{\partial f}{\partial x}(x,y) = g'(x)$ for that fixed value of y.

E.g.,

$$f(x,y) = x^2 + y^2 + \sin(xy) \qquad \Longrightarrow \qquad \frac{\partial f}{\partial x}(x,y) \;=\; 2x + \cos(xy)\cdot y\,.$$

Note: Unfortunately, a common notation is to write f_x for $\frac{\partial f}{\partial x}$.

But this does not work well with other notations such as B_t, S_t and C_t.

In class, I will try to use the longer notation $\frac{\partial f}{\partial x}$.

Of course $\frac{\partial f}{\partial y}(x,y)$ is defined similarly: the derivative with respect to y, holding x fixed (as a constant).

An important fact is the chain rule for partial derivatives.

Suppose $x = x(t)$ and $y = y(t)$ are differentiable curves.

Then $f(x(t), y(t))$ is a differentiable function of t, and

$$\frac{d}{dt}\, f(x(t),y(t)) \;=\; \frac{\partial f}{\partial x}(x(t),y(t))\,\frac{dx}{dt} + \frac{\partial f}{\partial y}(x(t),y(t))\,\frac{dy}{dt}\,.$$

This is particularly important for a special case.

Suppose X_t is an asset price which is differentiable in time t.

Then, thinking of $f(X_t, t)$ as a function $f(x, t)$ evaluated at $x = X_t$, we have

$$\frac{d}{dt} f(X_t, t) \;=\; \frac{\partial f}{\partial x}(X_t, t)\, \frac{dX_t}{dt} + \frac{\partial f}{\partial t}(X_t, t)\,.$$

We often write this as

$$df(X_t, t) \;=\; \frac{\partial f}{\partial x}(X_t, t)\, dX_t + \frac{\partial f}{\partial t}(X_t, t)\, dt\,,$$

which is what we get by multiplying everything through by dt.

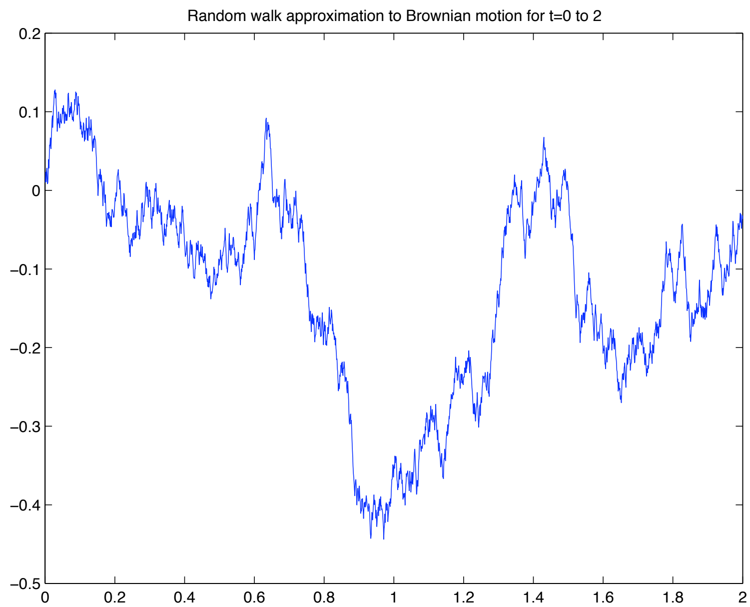
This is called a “total derivative”.

One useful thing you can do with this is integrate it:

$$f(X_T, T) - f(X_0, 0) \;=\; \int_0^T df(X_t, t) \;=\; \int_0^T \left[\frac{\partial f}{\partial x}(X_t, t)\, dX_t + \frac{\partial f}{\partial t}(X_t, t)\, dt \right] .$$

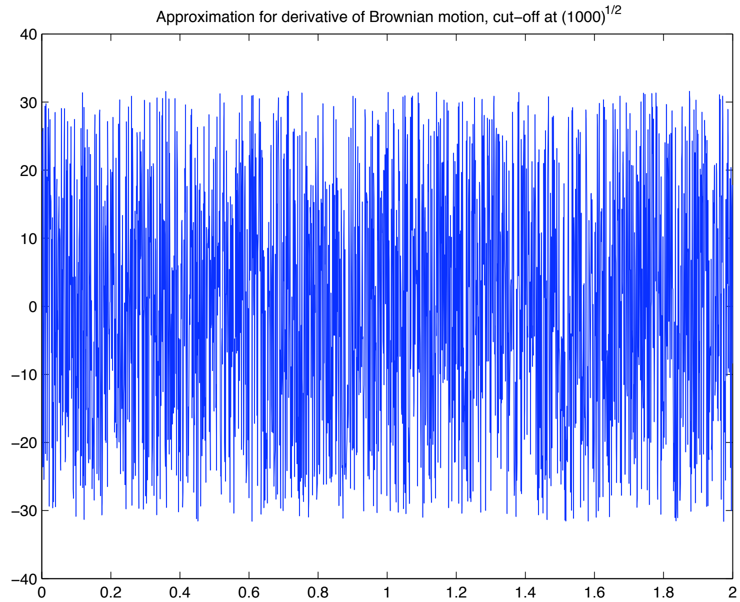
However, once again, beware: if the asset X_t is not differentiable in t, this simple formula does not work. In that case we need more terms of the Taylor series of f.

An example of an asset that is differentiable in t is the integral of a stock price over some window.



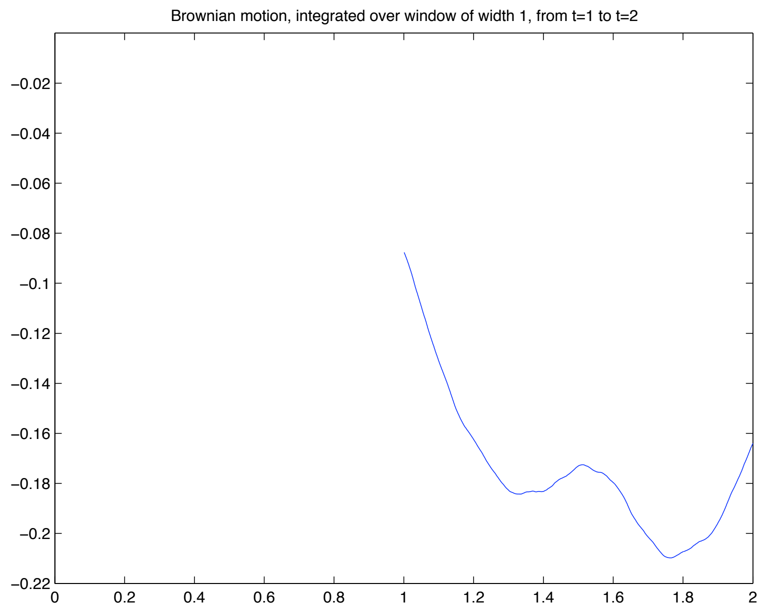
We have plotted a numerical approximation to Brownian motion from time $t = 0$ to $t = 2$. This has some of the same qualitative features of a stock prices S_t, except that Brownian motion, itself, can be negative (which no stock price can ever be). Although it is not 100% obvious from the picture, the graph of Brownian motion is continuous, but not differentiable. (Remember that a loose definition of continuity says that you can draw it without lifting your pen/pencil.)

Next we show a plot of the finite-difference approximation to the derivative of Brownian motion.



We cut this off at $\Delta t = 1000$, which results in a cut-off of the derivative at $\approx \sqrt{1000}$. (We made the approximation for Brownian motion out of a random walk, so that in our approximation, the absolute max and min for this approximate derivative are at $\pm\sqrt{1000} \approx \pm 31.62$.) The feature you see is that this graph fills out space up to the cut-off. The derivative of Brownian motion is white noise, and one of its features is that it has fluctuations on all scales. It is not a function. It is a generalized function or distribution, like the Dirac-delta function that you learn about in physics.

Finally, we show a plot for an asset $X_t = \int_{t-1}^{t} S_t \, dt$, where we are modelling S_t right now by Brownian motion. (This is wrong because Brownian motion can be negative, whereas S_t cannot be. But qualitative features other than that are similar.)



This graph is much different than the others. Not only is it continuous, it is also differentiable. Therefore, even though this is also a random function, we can take the derivative of X_t. (We would get $\frac{d}{dt}X_t = S_t - S_{t-1}$ by the Fundamental Theorem of Calculus.) So if we had some function $f(X_t, t)$ for this asset, the total derivative would be

$$df(X_t, t) \;=\; \frac{\partial f}{\partial x}(X_t, t)\, dX_t + \frac{\partial f}{\partial t}(X_t, t)\, dt\,.$$

However, since S_t is not differentiable the same type of formula would *not* work for dS_t. In fact, as we will see much later, we will need to add an extra term to the right-hand-side

$$df(S_t, t) \;=\; \frac{\partial f}{\partial x}(S_t, t)\, dS_t + \frac{\partial f}{\partial t}(S_t, t)\, dt + \frac{1}{2} \cdot \frac{\partial^2 f}{\partial x^2}(S_t, t)\, (dS_t)^2\,.$$