Stochastic Calculus, Courant Institute, Fall 2019
http://www.math.nyu.edu/faculty/goodman/teaching/StochCalc2019/StochCalc.html
Jonathan Goodman, September, 2019

Class 9, More differential equations, continued

1 Introduction

First, we finish the material on PDE. Time permitting, we pivot to the more theoretical material with more careful definitions of the Ito integral. [Alas, time did not permit.]

You can derive backward equations in many cases using a general generator formula. Recall that if $g(x)$ is a function that does not depend on t, then the generator determines the expected change in g in a time Δt, up to a "tiny" error term

$$\mathrm{E}_{x,t}[\, g(X_{t+\Delta t})] = g(x) + Lg(x)\Delta t + o(\Delta t) \; .$$

If a function depends on t explicitly, then the expected value can change either because $X_{t+\Delta t} \neq x$ or through the direct dependence of f on t. This explains the two terms in $[\cdots]$ here:

$$\mathrm{E}_{x,t}[\, f(X_{t+\Delta t}, t+\Delta t)] = f(x,t) + [\, \partial_t f(x,t) + Lf(x,t)]\, \Delta t + o(\Delta t) \; . \tag{1}$$

2 Backward equation for additive functionals

An *additive functional* is a function of the path of the form

$$G(X_{[0,T]}) = \int_0^T V(X_t)\, dt \; .$$

It is called *additive* because it comes from adding contributions for each time interval $(t, t+dt)$. It is a *path functional* because it depends on the whole path, not just the final value. Functionals like this arise in finance, where you get a "payment" for each time interval depending on the state during that time interval. For example, a floating rate loan results in payments that fluctuate with the short term interest rate. Engineers use functionals like this to evaluate the total cost corresponding to a stochastic process. For example, it might be at $V(x)$ is the fuel needed if the state is x.

There is a backward equation whose solution evaluates additive functionals like this. Starting at time t with state $X_t = x$, you can ask for the expected value of the integral starting at time t.

$$f(x,t) = \mathrm{E}_{x,t}\left[\int_t^T V(X_s)\, ds \right] \; . \tag{2}$$

We derive a backward equation for f using the general relation (1). Note that f at time $t + \Delta t$ (the left side if (1) involves only the integral in (2) starting at

$t + \Delta t$. Therefore, we write the integral as the sum of the first part, the part from t to $t + \Delta t$ and the rest. This gives

$$f(x,t) = \mathrm{E}_{x,t}\left[\int_t^{t+\Delta t} V(X_s)\,ds\right] + \mathrm{E}_{x,t}\left[\int_{t+\Delta t}^{T} V(X_s)\,ds\right] \ .$$

The first term on the right is $V(x)\Delta t + o(\Delta t)$. The second term is

$$\mathrm{E}_{x,t}[\,f(X_{t+\Delta t}, t+\Delta t)] \ .$$

We can use the general formula (1) and combine $o(\Delta t)$ terms, which gives

$$f(x,t) = V(x)\Delta t + f(x,t) + \partial_t f(x,t)\Delta t + Lf(x,t)\Delta t + o(\Delta t) \ .$$

Now cancel $f(x,t)$ from both sides, divide by Δt, and use $o(\Delta t)/\Delta t \to 0$ as $\Delta t \to 0$. The result is

$$0 = \partial_t f(x,t) + Lf(x,t) + V(x) \ . \tag{3}$$

This is the backward equation for an additive functional.

A variation on this theme is an additive functional with a discount rate, r. The value function for this is

$$f(x,t) = \mathrm{E}_{x,t}\left[\int_t^T e^{-r(s-t)} V(X_s)\,ds\right] \ . \tag{4}$$

If $s > t$, a payment at time s has value at time t that is discounted by a factor $e^{-r(s-t)}$. As with the simple additive functional, we find the backward equation using the general formula (1) and some analysis to relate the integral starting at time t to the integral starting at time $t + \Delta t$. For the simple functional, the only difference was the integral from t to $t + \Delta t$. Here, the discount factor also changes.

$$\int_t^T e^{-r(s-t)} V(X_s)\,ds = \int_t^{t+\Delta t} e^{-r(s-t)} V(X_s) ds + \int_{t+\Delta t}^T e^{-r(s-t)} V(X_s)\,ds \ .$$

We try to rewrite the second integral to put it into the form of the right side of (4) for $t+\Delta t$. We want $r(s-(t+\Delta t))$ instead of $r(s-t)$. For this, we calculate the exponent:

$$-r(s-t) = -r(s-(t+\Delta t)+\Delta t) = -r(s-(t+\Delta t)) - r\Delta t \ .$$

Therefore, the second integral is

$$\int_{t+\Delta t}^T e^{-r(s-t)} V(X_s)\,ds = e^{-r\Delta t} \int_{t+\Delta t}^T e^{-r(s-(t+\Delta t))} V(X_s)\,ds \ .$$

The factor $e^{-r\Delta t}$ represents the fact that on the left a payout is discounted from time s to time $t+\Delta t$, while on the right it is discounted to time t, which is more discounting. These preliminaries allow us to repeat the calculation that led to (3)

$$\begin{aligned}
f(x,t) &= V(x)\Delta t + e^{-r\Delta t}\mathrm{E}_{x,t}[\,f(x+\Delta X, t+\Delta t] \\
&= V(x)\Delta t + e^{-r\Delta t} f(x,t) + \Delta t \partial_t f(x,t) + \Delta t L f(x,t) + o(\Delta t) \\
&= V(x)\Delta t + (1 - r\Delta t + o(\Delta t)) f(x,t) + \Delta t \partial_t f(x,t) + \Delta t L f(x,t) + o(\Delta t) \\
0 &= V(x)\Delta t - r\Delta t f(x,t) + \Delta t \partial_t f(x,t) + \Delta t L f(x,t) + o(\Delta t) \\
0 &= \partial_t f(x,t) + L f(x,t) + V(x) - r f(x,t) \; . \qquad (5)
\end{aligned}$$

This is the backward equation for the discounted additive functional (4).

There may be a final payout in addition to the "running" payout. If so, that is the final condition for f. Otherwise, the final condition is $f(x,T) = 0$ because $f(x,T) = \int_T^T (\cdots) ds = 0$. If the final payout is $f(x,T) = 1$, and if the running payout $V(x) = 0$, then the value function is $f(x,t) = e^{-r(T-t)}$. This satisfies the backward equation (5) and has the right final condition. It corresponds to discounting the payment at time T by the amount of time until you reach T. To be sure, $\partial_t e^{-r(T-t)} = \partial_t e^{-rT} e^{rt} = r e^{-r(T-t)}$. We have $\partial_t f > 0$, which means that the number we give as the price of a discounted asset (payout) is an increasing function of t because increasing t means less discounting.

The integral in the discounted additive functional (4) represents the expected discounted present value of all future payments $V(X_s)ds$ for $s > t$. If the discount rate is positive ($r > 0$), the present value integral may converge as $T \to \infty$. We define the limit value with the hope that it exists

$$\begin{aligned}
g(x) &= \lim_{T\to\infty} f(x,t) \\
&= \lim_{T\to\infty} \mathrm{E}_{x,t}\left[\int_t^T e^{-r(s-t)} V(X_s)\, ds \right] \\
g(x) &= \mathrm{E}_{x,t}\left[\int_t^\infty e^{-r(s-t)} V(X_s)\, ds \right] \; . \qquad (6)
\end{aligned}$$

In finance, an instrument that remains in force forever is called *perpetual*. The integral (6) could represent paying a variable rate interest on a loan but never paying back the "principal" (the money borrowed). The left side is written as independent of t. The right side "clearly" is independent of t, which you see by substituting $s' = s - t$ in the integral. The variable s in the integral represents "time in the future from now". With $T < \infty$, it matters how long until the integral ends. But here, the integral never ends.

The infinite "time horizon" value function satisfies the PDE

$$L g(x) = -V(x) + r g(x) \; . \qquad (7)$$

You can derive this by setting $f(x,t) = g(x)$ in (5). The term $\partial_t f$ is zero because f does not depend on t. Here is a simple example. The process is one dimensional Brownian motion with generator $L = \frac{1}{2}\partial_x^2$. The running payout is $V(x) = 1$ if $|x| < 1$ and $V(x) = 0$ if $|x| > 1$.

We analyze the three regions $x > 1$, $-1 < x < 1$ and $x < -1$. We construct the solution "up to unknown constants", then find the constants to fit the pieces of the solution together at $x = \pm 1$. For $x > 1$, the value function differential equation (7) becomes

$$\frac{1}{2}\partial_x^2 g(x) = rg(x) \; .$$

In a differential equations class, you would learn that there are solutions of the form $g(x) = e^{\mu x}$. This *ansatz* can be substituted into the differential equation:

$$\frac{1}{2}\mu^2 e^{\mu x} = re^{\mu x} \; .$$

This leads to the algebraic equation for the growth/decay rate

$$\mu^2 = 2r \; , \quad \mu = \pm\sqrt{2r} \; .$$

The plus sign gives $e^{\mu x}$, which grows exponentially as $x \to +\infty$. On the contrary, we expect $g(x) \to 0$ as $x \to \infty$, because if you start at $X_0 = x$ that is a large positive number, then it takes a long time before X_s is "in the money" (has $|X_s| < 1$ so $V(X_s) \neq 0$). With discounting, this makes the instrument worth little. Therefore, we expect

$$g(x) = Ce^{-\mu x} \; , \quad \text{for } x > 1 \; .$$

The problem is symmetric, in that $+x$ is "the same" as $-x$. Therefore the value function should be symmetric in the sense that $g(-x) = g(x)$. This gives

$$g(x) = Ce^{\mu x} \; , \quad \text{for } x < -1 \; .$$

In the middle region, the differential equation is $\partial_x^2 g = -2 + 2rg$. The strategy from differential equations class is to combine a simple specific solution to the "inhomogeneous" problem, which is the full differential equation, with the "general solution" (a parameterized representation of all solutions) to the "homogeneous problem" (the differential equation terms involving g). The specific solution can be a constant, which makes the second derivative equal to zero. When you put a constant in for $g(x)$, you get

$$-2 + 2r\,const = 0 \; , \quad const = \frac{1}{r} \; .$$

The "general solution" to the homogeneous problem, as we have seen, is $C_2 e^{-\mu x} + C_3 e^{\mu x}$. But our solution is symmetric, with $g(-x) = g(x)$. This implies that the $e^{-\mu x}$ and $e^{\mu x}$ terms have the same weight, which is the same as $C_3 = C_2$. Therefore, the solution for $|x| < 1$ has the form

$$g(x) = \frac{1}{r} + C_2\left[e^{-\mu x} + e^{\mu x}\right] \; .$$

Finally, we "glue" the solutions together at $x = \pm 1$. There are two continuity conditions, one for g and one for $\partial_x g$. Each gives one equation involving the two constants C and C_2. Two conditions will determine the two constants completely. Our solution has a limit as $x \uparrow 1$, which we write as

$$g(1-) = \lim_{x \uparrow 1} g(x) = \lim_{x \to 1, x<1} g(x) \; .$$

The limit from above is

$$g(1+) = \lim_{x \downarrow 1} g(x) = \lim_{x \to 1, x>1} g(x) \; .$$

Our g is continuous at $x = 1$ if $g(1-) = g(1+)$. Given our formulas, this is

$$\frac{1}{r} + C_2 \left[e^{-\mu} + e^{\mu} \right] = C e^{-\mu} \; .$$

We are going to use a system of two linear equations for C and C_2, so we rewrite this in a convenient form for that

$$e^{-\mu} C - \left[e^{-\mu} + e^{\mu} \right] C_2 = \frac{1}{r} \; .$$

Similarly, we calculate

$$\begin{aligned} \partial_x g(1-) &= \lim_{x \uparrow 1} \partial_x g(x) \\ &= \lim_{x \uparrow 1} C_2 \left[-\mu e^{-\mu x} + \mu e^{\mu x} \right] \\ &= C_2 \left[-\mu e^{-\mu} + \mu e^{\mu} \right] \; . \end{aligned}$$

The calculation from the other side is simpler

$$\partial_x g(1+) = -C \mu e^{-\mu} \; .$$

The derivative continuity condition is $\partial_x g(1-) = \partial_x g(1+)$. With the calculated values, this is

$$C_2 \left[-\mu e^{-\mu} + \mu e^{\mu} \right] = -C \mu e^{-\mu} \; .$$

We cancel the common μ factor and write this as the second of two linear equations

$$e^{-\mu} C + \left[-e^{-\mu} + e^{\mu} \right] C_2 = 0 \; .$$

We assemble these equations into a system of two equations:

$$\begin{array}{ccccc} e^{-\mu} C & - & \left[-e^{-\mu} + e^{\mu} \right] C_2 & = & \dfrac{1}{r} \\ e^{-\mu} C & + & \left[-e^{-\mu} + e^{\mu} \right] C_2 & = & 0 \end{array}$$

This is too easy. We subtract the second equation from the first and get

$$-2e^{\mu} C_2 = \frac{1}{r} \implies C_2 = -\frac{e^{-\mu}}{2r} \; .$$

We find C by putting the C_2 formula into the second equation and calculating

$$e^{-\mu} C + e^{-\mu} \frac{e^{-\mu}}{2r} - e^{\mu} \frac{e^{-\mu}}{2r} = 0 \; .$$

This simplifies to

$$C = \frac{e^{\mu} - e^{-\mu}}{2r} \; .$$

3 Multiplicative functionals

A simple multiplicative functional is a function of the form

$$G(X_{[0,T]}) = e^{\int_0^T V(X_s)\,ds} \ . \tag{8}$$

These are called *multiplicative* because they represent products of individual factors coming from times s and lasting for duration ds. A discrete approximation would be

$$\exp\left[\sum_{t<s_k<T} \Delta s V(X_{s_k})\right] = \prod_{t<s_k<T} e^{\Delta s V(X_{s_k})} \ .$$

The contributions from times s_k are multiplied together, making the functional literally multiplicative. The limit $\Delta s \to 0$ would be a sort of multiplicative integral, like the ordinary integral might be called an additive integral. The limit may be written using an ordinary (additive) integral because it uses the exponential function.

Functionals like this arise in finance as models of accumulation of stochastic interest rates. If X_s is the state of some model of a relevant part of "the market", then the interest rate for that state is $V(X_s)$. The value of an asset changes by the factor $e^{V(X_s)ds} = 1 + V(X_s)ds$ during the time interval $s, s+ds$.

By now, we have a good guess how to derive a backward equation to evaluate $E[G]$. The value function is

$$f(x,t) = \mathrm{E}_{x,t}\left[e^{\int_t^T V(X_s)ds}\right] \ . \tag{9}$$

We want to relate $f(x,t)$ to $f(\cdot, t+\Delta t)$, so we separate the integral into the parts before and after $t+\Delta t$. The exponential of the sum of the two integrals is the product of the two exponentials.

$$\begin{aligned}
e^{\int_t^T V(X_s)\,ds} &= e^{\int_t^{t+\Delta t} V(X_s)\,ds}\, e^{\int_{t+\Delta t}^T V(X_s)\,ds} \\
&= e^{V(X_t)\Delta t + o(\Delta t)}\, e^{\int_{t+\Delta t}^T V(X_s)\,ds} \\
&= \left[1 + V(X_t)\Delta t + o(\Delta t)\right]\, e^{\int_{t+\Delta t}^T V(X_s)\,ds} \ .
\end{aligned}$$

This gets inserted into the right side of the value function definition (9). For that purpose, $X_t = x$ is not random, so that fact comes out of the expectation. The calculations are similar to the calculations for additive functionals

$$\begin{aligned}
f(x,t) &= \left[1 + V(X_t)\Delta t + o(\Delta t)\right] \mathrm{E}_{x,t}\left[e^{\int_{t+\Delta t}^T V(X_s)\,ds}\right] \\
&= \left[1 + V(X_t)\Delta t + o(\Delta t)\right] \mathrm{E}_{x,t}\left[f(X_{t+\Delta t}, t+\Delta t)\right] \\
&= \left[1 + V(X_t)\Delta t + o(\Delta t)\right]\left[f(x,t) + \Delta t \partial_t f(x,t) + \Delta t L f(x,t) + o(\Delta t)\right] \\
&= f(x,t) + \Delta t\left[V(x)f(x,t) + \partial_t f(x,t) + Lf(x,t)\right] + o(\Delta t) \\
0 &= V(x)f(x,t) + \partial_t f(x,t) + Lf(x,t) + \frac{o(\Delta t)}{\Delta t} \ .
\end{aligned}$$

The last part goes to zero as $\Delta t \to 0$ (which is the definition of $o(\Delta t)$, and we are left with

$$0 = \partial_t f(x,t) + Lf(x,t) + V(x) f(x,t) \; . \tag{10}$$

This is the backward equation that computes expected values of multiplicative functionals.

Someone interested in modeling physical systems using differential equations might derive an equation of the form (10). A later class will give an example of this. The mathematician Marc Kac discovered that the solution may be expressed as an expected value (9). [Americans pronounce his name "cats". The final "c" in Polish is pronounced "ts" in English, and often written "tz". For example, you can visit the famous "Katz Deli" on Houston Street with its huge pastrami sandwiches and long tourist lines.] Kac was led to this formula in an attempt to understand a related formula by the physicist Richard Feynman (author of the *Feynman Lectures on Physics* and *Surely You're Joking, Mr. Feynman* and enough important physics to earn a Nobel Prize). Feynman's formula was for the *Schrödinger equation*, which is related to the backward equation (10). A later class will explain the reasoning Kac used and why he was troubled by Feynman's formula. The value function formula (9) is called the *Feynman Kac* formula for the solution of the partial differential equation (10). But you should be prepared for any formula or equation in these notes to be called "the Feynman Kac formula". People can be careless with names of equations.