0 Outline Content

Motivating examples: simple control problem (Cramer?), quadratic regulator, utility maximisation, option pricing. (1 hour)

Technical Background: Stochastic integral, basic properties, Itô's Formula, martingale representation theorem, SDEs, weak & strong solutions, Diffusions: strong Markov property, Generators, Dynkin's formula, Feynman-Kac formula. Girsanov's theorem. Dirichlet-Poisson problem. (5 hours)

Stochastic Optimal Control 'Theory': Problem statement, Markov controls, value function, dynamic programming principle(?), characterisation of an optimal control — HJB equation, verification theorem, 'typical' approach: use intuition and HJB to guess a solution, then verify. (e.g. Oksendal, Ex. 11.2.5). (2 hours)

Applications: Option pricing: complete markets, trading and arbitrage, fundamental theorem, incomplete markets — upper and lower hedging price, dual representation. Utility indifference pricing. (4 hours) American options?

Utility maximisation: investment and consumption, Merton problem, dual representation of solutions, maximisation of growth rate(?), transaction costs (Davis & Norman). (4 hours)

Utility maximisation — inverse problem? (2 hours)

1 Motivating Examples

Example 1.1. Consider a company whose assets are valued according to a Brownian motion with diffusion coefficient σ and unit drift, and who can pay their shareholders dividends, so long as the company is not bankrupt (i.e. debts larger than assets). If, at time 0, the net value of the company is x, and they pay dividends at rate u_t at time t, the company value at time t is:

$$\begin{aligned} X_t^u &= x + \sigma B_t - \int_0^t u_s \, \mathrm{d}s + t \\ &= x + \sigma B_t + \int_0^t (1 - u_s) \, \mathrm{d}s. \end{aligned}$$

Adjusting for interest, the shareholders wish to maximise their average payout:

$$\mathbb{E}\left[\int_0^{\tau^u} \mathrm{e}^{-rt} u_t \, \mathrm{d}t\right]$$

where $r > 0$ is the interest rate, and $\tau^u = \inf\{t \geq 0 : X_t^u = 0\}$ is the time of bankruptcy. What should the company do — i.e. how should they choose the dividend payments — given that they are restricted to choosing $u_s \in [0,1] := U$?

Notes:

- The choice of $(u_s)_{s\geq 0}$ should clearly be allowed to depend on the behaviour of X_t^u up to the current time — e.g. require u_s adapted to $\mathcal{F}_s^B$.
- We're interested in two things: what is the best average payout, and also how do we achieve it.
- There is a sort of Markov property at play: set

$$V^u(x,t) := \mathbb{E}^{(x,t)}\left[\int_t^{\tau^u} \mathrm{e}^{-rs} u_s \, \mathrm{d}s\right]$$

  where $\mathbb{E}^{(x,t)}$ is the law of $(X_{t+s}^u)_{s\geq 0}$ given $X_t^u = x$, then $V^u(x,t) = V^{\tilde{u}}(x,0)\mathrm{e}^{rt}$ where $\tilde{u}_s = u_{s-t}\circ\theta_t$, and θ_t is the Markov shift operator. In particular, this suggests that the problem may be time-independent — i.e. if the optimal strategy exists, and is u_t^*, we might expect $u_t^* = u^*(X_t^u)$ for some function $u^*(x)$.

  This also suggests that we may often be reduced to strategies that are 'Markovian' in the sense that they are a function only of the current position, but we will usually not wish to exclude other strategies from our optimality statements.

Let's think about solving the problem. Let $\mathcal{U}$ be the set of possible **controls** and define the **value function**

$$V^*(x,t) = \sup_{u\in\mathcal{U}} V^u(x,t).$$

Then we might expect:

$$\begin{aligned} V^*(x,t) &\geq \mathbb{E}^{(x,t)}\left[\int_t^{\tau^u} \mathrm{e}^{-rs} u_s \, \mathrm{d}s\right] \\ &\geq \mathbb{E}^{(x,t)}\left[\int_t^{\tau^u\wedge\rho} \mathrm{e}^{-rs} u_s \, \mathrm{d}s + \int_{\rho\wedge\tau^u}^{\tau^u} \mathrm{e}^{-rs} u_s \, \mathrm{d}s\right] \end{aligned}$$

where $\rho \geq t$ is a stopping time. It then follows that:

$$V^*(x,t) \geq \mathbb{E}^{(x,t)}\left[\int_t^{\tau^u\wedge\rho} \mathrm{e}^{-rs} u_s \, \mathrm{d}s + V^u(X_{\rho\wedge\tau^u}^u, \rho\wedge\tau^u)\right].$$

Without worrying too much about the technical details, we could imagine if u^1 and u^2 are two strategies, then we could construct a new strategy

$$\tilde{u}_t = \begin{cases} u_t^1 & t \leq \rho \\ u_t^2 & t \geq \rho \end{cases},$$

and this would suggest that

$$V^*(x,t) \geq \mathbb{E}^{(x,t)}\left[\int_t^{\tau^{u^1}\wedge\rho} \mathrm{e}^{-rs} {u^1}_s \, \mathrm{d}s + V^{u_2}(X_{\rho\wedge\tau^{u^1}}^{u^1}, \rho\wedge\tau^{u^1})\right].$$

But u^2 is arbitrary, so if we take the supremum over u^2, and add $\int_0^t \mathrm{e}^{-rs} u_s^1 \, \mathrm{d}s$ to both sides, and take $x = X_t^u$ to get

$$V^*(X_t^{u^1}, t) + \int_0^t \mathrm{e}^{-rs} u_s^1 \, \mathrm{d}s \geq \mathbb{E}^{(X_t^{u^1},t)}\left[\int_0^{\tau^{u^1}\wedge\rho} \mathrm{e}^{-rs} {u^1}_s \, \mathrm{d}s + V^*(X_{\rho\wedge\tau^{u^1}}^{u^1}, \rho\wedge\tau^{u^1})\right].$$

Assuming some sort of Markovianity, this suggests that

$$V^*(X^u_{t\wedge\tau^u}, t\wedge\tau^u) + \int_0^{t\wedge\tau^u} \mathrm{e}^{-rs} u_s \,\mathrm{d}s \text{ is a supermartingale for all } u \in \mathcal{U}. \tag{1}$$

Moreover, we should have equality throughout if we can find a strategy $u^* \in \mathcal{U}$ which is optimal, which suggests that

$$V^*(X^{u^*}_{t\wedge\tau^u}, t\wedge\tau^u) + \int_0^{t\wedge\tau^u} \mathrm{e}^{-rs} u^*_s \,\mathrm{d}s \text{ is a martingale for some } u^* \in \mathcal{U}. \tag{2}$$

This principle of representing the value function for the problem in terms of itself is known as the Bellman principle, or the principle of dynamic programming.

Let us now try to solve the problem based on these ideas, and assuming everything is 'well-behaved'.

First, let's suppose that the value function exists, and is differentiable. Then (and this is essentially an argument that we will formalise as Itô's Lemma) by Taylor's Theorem:

$$\begin{aligned} V^*(X^u_{t+\delta t}, t+\delta t) - V^*(X^u_t, t) &\approx \frac{\partial V^*}{\partial x}(X^u_{t+\delta t} - X^u_t) + \frac{\partial V^*}{\partial t}\delta t + \frac{1}{2}\frac{\partial^2 V^*}{\partial x^2}(X^u_{t+\delta t} - X^u_t)^2 \\ &\approx \frac{\partial V^*}{\partial x}(\sigma\,(B_{t+\delta t} - B_t) + \delta t(1-u_t)) + \frac{\partial V^*}{\partial t}\delta t \\ &\quad + \frac{1}{2}\frac{\partial^2 V^*}{\partial x^2}(\sigma\,(B_{t+\delta t} - B_t) + \delta t(1-u_t))^2 \\ &\approx \sigma\frac{\partial V^*}{\partial x}\delta B_t + \left((1-u_t)\frac{\partial V^*}{\partial x} + \frac{\partial V^*}{\partial t}\right)\delta t + \frac{1}{2}\frac{\partial^2 V^*}{\partial x^2}\sigma^2(\delta B_t)^2 + \text{ h.o.t.} \end{aligned}$$

But $\mathbb{E}(\delta B_t)^2 = \delta t$, and in fact, for small δt, we will be able to justify the substitution $(\delta B_t)^2 = \delta t$ to get

$$V^*(X^u_{t+\delta t}, t+\delta t) - V^*(X^u_t, t) \approx \sigma\frac{\partial V^*}{\partial x}\delta B_t + \left((1-u_t)\frac{\partial V^*}{\partial x} + \frac{\partial V^*}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V^*}{\partial x^2}\right)\delta t.$$

Since B_t is zero on average, the (super-)martingale conditions (1) and (2) suggest we need:

$$(1-u_t)\frac{\partial V^*}{\partial x} + \frac{\partial V^*}{\partial t} + \frac{1}{2}\sigma^2\frac{\partial^2 V^*}{\partial x^2} \leq -\mathrm{e}^{-rt}u_t \tag{3}$$

for all $u_t \in U$. Recalling that our behaviour should (intuitively) be independent of time, we conjecture that in fact $V^*(x,t) = \mathrm{e}^{-rt}V(x)$, for some function V. Then (3) implies:

$$(1-u_t)V_x(x) - rV + \frac{1}{2}\sigma^2 V_{xx} \leq -u_t,$$

for all $u_t \in U$, with equality for some $ut \in U$, or equivalently:

$$V_x + \frac{1}{2}\sigma^2 V_{xx} - rV = \inf_{w\in[0,1]}[w(V_x(x) - 1)] = -(1 - V_x(x))_+. \tag{4}$$

Note that the fact that we only need consider $u = 0$ or $u = 1$ suggests that the optimal behaviour will only be not to pay dividends, or to pay as much as possible. There is no (substantial) middle ground. We can also try to consider some properties of the solution. Clearly, we expect $V(x) \downarrow 0$ as $x \downarrow 0$, and also $V(x) \to r^{-1}$ as $x \to \infty$, since $u_t \leq 1$ always, and as $x \to \infty$, we can keep paying out at rate 1 for arbitrarily long on average (this could be proved using properties of Brownian motion). Also, we are better off if we start with bigger x, so $V_x(x) \geq 0$.

Suppose that we can ignore the $(\cdot)_+$, so that we want to solve

$$V_x + \frac{1}{2}\sigma^2 V_{xx} - rV = V_x(x) - 1,$$

then this has solutions of the form $V(x) = r^{-1} + \alpha \mathrm{e}^{-\sqrt{\frac{2r}{\sigma^2}}x} + \beta \mathrm{e}^{\sqrt{\frac{2r}{\sigma^2}}x}$. Assuming that, for large x, we want $V(x) \to r^{-1}$, we must have $\beta = 0$. In fact, if we now choose $V(x) = r^{-1}\left(1 - \mathrm{e}^{-\sqrt{\frac{2r}{\sigma^2}}x}\right)$, we see also that $V(0) = 0$, and $V_x(x) \leq r^{-1}\sqrt{\frac{2r}{\sigma^2}}$. In particular, if the parameters are such that $\frac{2}{\sigma^2} \leq r$, then $V_x(x) \leq 1$ always, and (4) holds.

Now, going back through this argument, with a little bit of extra machinery, we can check that (1) and (2) hold for $V^*(x,t) = \mathrm{e}^{-rt}V(x)$, with $u_t^* \equiv 1$ for all $t \geq 0$, and hence that:

$$\begin{aligned} V^*(x,0) &= \mathbb{E}^{(x,0)}\left[V^*(X^{u^*}_{t\wedge\tau^{u^*}}, t \wedge \tau^{u^*}) + \int_0^{t\wedge\tau^{u^*}} \mathrm{e}^{-rs}u_s^*\,\mathrm{d}s\right] \\ &\to \mathbb{E}^{(x,0)}\left[\int_0^{\tau^{u^*}} \mathrm{e}^{-rs}u_s^*\,\mathrm{d}s\right] \end{aligned}$$

as $t \to \infty$, by bounded convergence, the fact that $\tau^{u^*} < \infty$ a.s. and monotone convergence. Using the super-martingale property, we can show a similar inequality when the RHS involves an arbitrary u rather than u^*. Hence we have found the optimal u^* and the corresponding value, V^*.

In the case where $r < \frac{2}{\sigma^2}$, things are a bit different. Essentially, what we want to do is combine solutions to the two equations. We guess that there is some x_0 such that $V_x(x) \geq 1$ on $[0, x_0]$ and $V_x(x) \leq 1$ on $[x_0, \infty)$. We then solve $V_x(x)$ on $[0,\infty)$ with the boundary conditions $V(0) = 0, V(\infty) = r^{-1}$ and (let's say) with the additional condition that $V'(x_0+) = 1$, and $V(x_0-) = V(x_0+)$, which is certainly necessary for the martingale/supermartingale conditions to hold. Together, for fixed x_0, these are enough to specify a unique solution to the pair of equations, but in general, we will not have $V'(x_0-) = 1$, and this will cause difficulty in the martingale/supermartingale conditions. To rectify this, we can try to vary x_0 (the remaining degree of freedom) to get $V'(x_0-) = 1$, and (hopefully, if this is the right general solution), we will see that the resulting solution is of this form: particularly, the resulting $V(X_t^u, t) + \int_0^t \mathrm{e}^{-rs}u_s\,\mathrm{d}s$ is a supermartingale up to τ_0, and a martingale under the optimal control. Once we have a candidate solution, then this is normally not too hard to check.

Notes:

- The procedure here is a sort of informed 'guess and verify.' We have used a mix of martingale-type properties, specific intuition about the problem, and informed guessing to find a candidate value function. Once we have this, then we are able to show that it is optimal via a martingale argument, but we need to have a specific guess first.

- Note that our final proofs of optimality do not require any Markovianity of a sub-optimal strategy, they will hold for essentially any choice of the function u.
- In the more complicated case, even in this relatively easy problem, finding the actual solution is not straightforward, and would in practice probably need a computer or similar. This is a fairly common situation — our answers are very likely to be only partially computable, but we can at least derive some qualitative features of the solution without calculation (that $u(X_t) = \mathbf{1}_{X_t \geq \alpha}$ for some constant α).
- In this example, the final value function is nice, but candidate value functions may not be — in particular, they may not be nice and differentiable, or bounded, etc., and in fact, even for nice problems, the value function is often not suitably differentiable. This can cause substantial technical difficulties.
- In this case, our choice of control only affected the drift of our process. In many circumstances, we may want to affect both the drift and the diffusion coefficient.

Example 1.2 (Stochastic Linear-Quadratic Regulator)*.* Consider a stochastic process whose change over a small time is given by:

$$\mathrm{d}X_t^u = (\alpha X_t^u + u_t)\,\mathrm{d}t + \sigma\,\mathrm{d}B_t$$

so if $\alpha = 0$, we just have $X_t^u = X_0 + \int_0^t u_s\,\mathrm{d}s + \sigma B_t$, but we also allow a more complicated local change which can depend on the process itself.

Consider the problem of minimising

$$\mathbb{E}\left[\int_0^T \beta(X_t^u)^2 + \gamma u_t^2\,\mathrm{d}t + \xi(X_T^u)^2\right].$$

Clearly, we should push towards the origin, but the question is how fast? In fact, the optimal control in this problem is independent of σ, so this problem can be used to investigate the impact of noise in controlled systems.

Example 1.3 (Utility Maximisation)*.* Consider the problem of an investor who wishes to maximise their wealth at retirement (say) but is risk-averse. They might choose to maximise:

$$\mathbb{E}U\left(X_T\right)$$

where $U(x)$ is a function known as the **utility** function, and is generally supposed to be concave and increasing, and $(X_t)_{t\in[0,T]}$ is the investor's wealth at time T. Note that concavity implies that $U(\mathbb{E}X) \geq \mathbb{E}U(X)$ for all random variables X, and this has an economic interpretation of risk-aversion.

The investor can affect X_t by choosing how to invest: suppose that she can invest some proportion of her wealth in the stock-market, and some proportion in a risk-free bank account, where she receives a fixed interest rate r. If the stock-market follows a Geometric Brownian motion (**Black-Scholes model**) then $S_t = S_0 \exp \mu t + \sigma B_t$ and we have the representation:

$$\mathrm{d}S_t = S_t\mu\,\mathrm{d}t + S_t\sigma\,\mathrm{d}B_t.$$

If we invest an amount u_t in the stock market at time t (so we have u_t/S_t units of the stock, and $X_t - u_t$ in the bank), over a small time period, our wealth changes by:

$$\begin{aligned} \mathrm{d}X_t &= \frac{u_t}{S_t}\,\mathrm{d}S_t + r(X_t - u_t)\,\mathrm{d}t \\ &= rX_t\,\mathrm{d}t + u_t(\mu - r)\,\mathrm{d}t + u_t\sigma\,\mathrm{d}B_t. \end{aligned} \tag{5}$$

A natural problem (the utility maximisation problem) is then:

$$\sup_{u_t} \mathbb{E}U(X_T^u)$$

where the control now affects the reward solely through the final value of wealth.

Depending on the situation, we may allow $u_t < 0$ (short selling), or require $u_t \geq 0$, or include other constraints, such as requiring our wealth X_t to remain positive, or above some fixed level $-\alpha$. In addition, we could include the option for the investor to *consume* their wealth before time t, and we then need to optimise over both the investment and the consumption strategy.

Example 1.4 (Option Pricing). Finally, consider the problem of a bank who has sold a derivative contract: that is, if S_t is the price of some financial asset, they have contracted to pay another party an amount $f(S_t)$ at some future date. To reduce their risk, they wish to super-hedge their exposure: that is, they want to find a trading strategy u_t and an initial wealth x_0 such that if $X_0^u = x_0$, and (5) holds, then $X_T^u \geq f(S_t)$ almost surely.

Solving this problem does not really need the full stochastic control machinery (as we will see, it can be handled in a slightly simpler manner), but it has much of the 'flavour' of a stochastic control problem, and can be generalised to some natural stochastic control problems: consider the problem

$$UIP(f) = \inf \left\{ p \in \mathbb{R} : \sup_{u_t} \mathbb{E}U\left(X_T^u + (p - f(S_T))\right) \geq \sup_{u_t} \mathbb{E}U\left(X_T^u\right) \right\}$$

that is, p is the smallest price at which I would prefer (greater utility) to sell the contract $f(S_T)$ for price p, and hedge using the asset, than to simply invest without selling the contract. $UIP(f)$ is the **Utility Indifference Price**, and clearly the UIP will be below the superhedging price, but taking $x_0 = 0$ and $U(x) = \mathbf{1}_{\{[0,\infty)\}}(x)$, we can recover the superhedging price as the UIP price.