Stochastic Control and Algorithmic Trading

Nicholas Westray

(nicholas.westray@db.com)

Deutsche Bank and Imperial College

RiO - Research in Options
Rio de Janeiro - 26th November 2011

Imperial College London

What is Stochastic Control?

- It is often the case that we as engineers are asked to control a system.

Example: Executing a risk trade.

- There is a trade off between executing a trade now and paying a high price and the risk of having that position on your book while you execute it.
- Stochastic control provides a mathematical formulation to tackle such problems.

There are two basic elements to any control problem,

- The controlled system, which we can affect (may be noisy).
- A cost function which depends upon our action and which we aim to minimise.

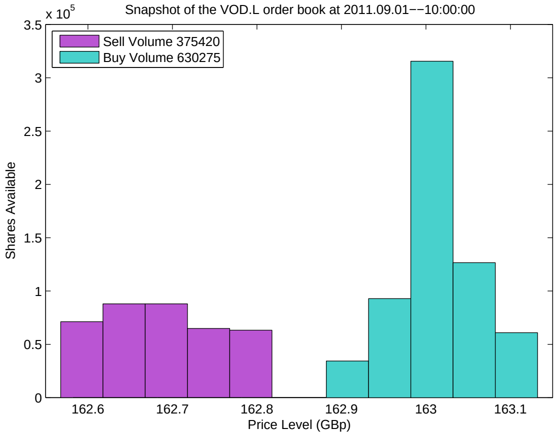
Motivating Example - Client Constrained Liquidation.

- Suppose we want to buy $Y = 10^7$ shares of VOD.L (approx 10% ADV) by T.
- The bid/offer is currently at 162.8p/162.9p and if we try to trade this all immediately, e.g as a market order, we would cross several levels of the order book and it would be expensive.
- Let us now formulate this more rigorously. The number of shares bought at time t is given by

$$X^u(t) = \sum_{i<t} u(t_i)\Delta t \approx \int_0^t u(s)ds, \quad X(T) = Y$$

Where we make a continuous time approximation, the function u is known as trading rate (the control).

VOD.L Sample Order Book



Motivating Example - Client Constrained Liquidation II.

- Let us assume a stock price equation of

$$S(t) = S(0) + \sigma_S B(t)$$

- However due to us crossing the spread we actually pay

$$\tilde{S}(t) = S(t) + \kappa u(t)$$

One can think of this as the average increase in price due to trading at rate u during the time interval (temporary impact).

- A calculation then shows that our costs at T of buying this position are given by

$$\sum_{i<T} u(t_i)\tilde{S}(t_i)\Delta t - YS(0) \approx \int_0^T u(s)\sigma_S B(s)ds + \int_0^T \kappa u^2(s)ds$$

- Taking expectation gives the final cost term as

$$\mathbb{E}\left[\int_0^T \kappa u^2(s)ds\right]$$

Motivating Example - Client Constrained Liquidation III

- Let us assume that in addition the client specifies a liquidation process $Z(s)$, possibly influenced by stochastic factors and that we are asked to track this.
- Choosing a quadratic penalty gives

$$\mathbb{E}\left[\int_0^T (u(s) - Z(s))^2 ds\right]$$

- This term can be thought of as the penalty arising due to not meeting client requirements.
- This leads us to minimize the functional

$$J(u) = \mathbb{E}\left[\int_0^T \kappa u^2(s) + \lambda (u(s) - Z(s))^2 ds\right].$$

  over trading strategies u, which have $X(T) = Y$.
- The aim of the first part of this presentation is to show you that stochastic control provides an analytic and tractable framework to study such problems.

Mathematical Preliminaries

Filtrations and Itô Integrals

- We will assume that we are working on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{\{0\leq t\leq T\}}, \mathbb{P})$ which satisfies the usual conditions. We denote by T the finite terminal time.
- We suppose that it supports a d-dimensional Brownian motion W.
- Given a left continuous (predictable) process H, the theory of stochastic integration allows us to define

$$M(t) = \int_0^t H(s)dW(s).$$

- When the function H is such that

$$\mathbb{E}\left[\int_0^T H^2(s)ds\right] < \infty,$$

the process M is a martingale, i.e. $\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s)$ for $s \leq t$.

Local Martingales

- Recall that a stopping time is a random variable in $[0, \infty]$, such that $\{\tau \leq t\} \in \mathcal{F}(t)$.

Definition

- A process M is a local martingale if there exists an increasing sequence of stopping times τ_n converging (up) to T such that

$$M(t \wedge \tau_n) \text{ is a martingale for each } n.$$

- To help visualise the difference between the two types of martingale, one should think of bubbles.
- Every stochastic integral with respect to Brownian motion is a local martingale.

SDEs - Existence and Uniqueness

- Consider the following SDE

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW(t), \quad X(0) = x \in \mathbb{R}^n$$

where $\mu : [0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ and $\sigma : [0, T] \times \mathbb{R}^n \to \mathbb{R}^{n \times d}$.

Theorem (Existence and Uniqueness of Strong Solution)

- *If the continuous functions μ and σ are such that*

$$\sup_{0 \le t \le T} (|\mu(t, 0)| + |\sigma(t, 0)|) \le C$$

$$|\mu(t, x) - \mu(t, y| + |\sigma(t, x) - \sigma(t, y)| \le C|x - y| \text{ for all } x, y \in \mathbb{R}^n$$

i.e. Lipschitz continuity uniformly in t, plus growth conditions.

- *Then the above SDE has a unique strong solution, which satisfies ($p \ge 1$)*

$$\mathbb{E}\left[\sup_{0 \le t \le T} |X(t)|^p\right] < \infty$$

Markov Property and the Connection with PDEs

- The solution is a strong Markov process, that is to say for bounded measurable g and a stopping time τ we have

$$\mathbb{E}\left[g\big(X(t+\tau)\big)|\mathcal{F}(\tau)\right] = \mathbb{E}\left[g\big(X(t+\tau)\big)|X(\tau)\right].$$

- A consequence of the Markov property is that for $t \in [0, T]$ and bounded smooth measurable g we have

$$\mathbb{E}[g\big(X(T)\big)|\mathcal{F}(t)] = \varphi(t, X(t)),$$

where φ solves the PDE on [0,T)

$$\varphi_t(t,x) + \mu(t,x)^\intercal D_x\varphi(t,x) + \frac{1}{2}\mathrm{Tr}\left(\sigma\sigma^\intercal D^2\varphi(t,x)\right) = 0,$$
$$\varphi(T,x) = g(x).$$

Generator of an SDE

- We call the differential operator $\mathcal{L}_X$, defined via

$$\mathcal{L}_X := \mu(t,x)^\intercal D + \frac{1}{2}\mathrm{Tr}\left(\sigma\sigma^\intercal(t,x)D^2\right).$$

  the generator for X.
- This allows us to write Itô's formula in condensed form, namely for a $\mathcal{C}^2$ function $\varphi(x)$ we have

$$d\varphi(X(t)) = \mathcal{L}_X\varphi(X(t))\,dt + \sigma D\varphi(X(t))\,dW(t)$$

- Given an SDE for the process X the generator is such that for any $\mathcal{C}^2$ function φ the process

$$\varphi(X(t)) - \mathcal{L}_X\varphi(X(t))$$

  is a local martingale.

Markov SDEs - The Flow Property

- A key idea about Lipschitz SDEs is that we can connect solutions to the SDE starting at different points. This is called the flow property. Consider a time point $0 \leq t_0 < T$ and consider the SDE started at (t_0, x_0)

$$dX^{t_0,x_0}(t) = \mu(t, X^{t_0,x_0}(t))dt + \sigma(t, X^{t_0,x_0}(t))dW(t), \quad X^{t_0,x_0}(t_0) = x_0 \in \mathbb{R}^n$$

- It follows from uniqueness that for a second time point t_1 and any $s \in [0, T]$ such that $t_0 < t_1 < s < T$ we have

$$X^{t_0,x_0}(s) = X^{t_1, X^{t_0,x_0}(t_1)}(s)$$

- This may look complicated, but the following diagram helps illustrate the intuitive property clearly.

The Control Problem and Value Function

Controlled SDEs

- We shall be interested in SDEs which are influenced by a control process u.
- Consider the following SDE on [0,T]

$$dX^u(t) = \mu(t, X^u(t), u(t))dt + \sigma(t, X^u(t), u(t))dW(t), \quad X^u(0) = x \in \mathbb{R}^n.$$

Definition

We call a control process u admissible and write $u \in \mathcal{U}_0$ if u is adapted, valued in $U \subset \mathbb{R}^k$ and the above equation has a unique strong solution.

- The process u describes our interaction with the system and may depend on the information up to t.

Markov Controls

- We shall be especially interested in those controls which can be written as

$$u(t) = \bar{u}(t, X^u(t))$$

for some deterministic function $\bar{u} : [0, T] \times \mathbb{R}^n \to U$. These are called Markov controls.

- These play an especially important role since they are of a much reduced form and can be simulated easily.

The Value Function - I

- We are asked to design a control to minimise some cost functional, we thus take functions f and g continuous in all their arguments and satisfying, for some $\eta > 1$,

$$\sup_{0 \le t \le T} |f(t,x,u)| + |g(x)| \le C(1 + |x|^\eta)$$

i.e. they have polynomial growth in x uniformly in (t,u).

- We then define

$$\widetilde{J}(u) := \mathbb{E}\left[\int_0^T f(t, X^u(t), u(t))dt + g(X^u(T))\right].$$

- In my initial example we had $f(t,(x,z),u) = \kappa u^2 + \lambda(u-z)^2$, $g(x) = \delta_Y(x)$.
- One typically calls f the running costs and g the terminal cost.

The Value Function - II

- We will not only be interested in the time point $t = 0$, but rather think about the problem started at a general point $(t_0, x_0) \in [0, T] \times \mathbb{R}^n$. We thus think about the function

$$J(t_0, x_0, u) := \mathbb{E}\left[\int_{t_0}^{T} f(t, X^{t_0,x_0}(t), u(t))dt + g(X^{t_0,x_0}(T))\middle|\mathcal{F}(t_0)\right].$$

- The controlled process is given by

$$dX^{t_0,x_0}(t) = \mu(t, X^{t_0,x_0}(t), u(t))dt + \sigma(t, X^{t_0,x_0}(t), u(t))dW(t),$$
$$X^{t_0,x_0}(t_0) = x_0 \in \mathbb{R}^n$$

- Our goal is to analyse the value function given by

$$v(t_0, x_0) = \inf_{u \in \mathcal{U}_{t_0}} J(t_0, x_0, u).$$

Markov Nature of the Value Function

- Observe that the function J may depend on the past history (it is $\mathcal{F}(t_0)$ measurable). Indeed the control is only supposed to be adapted.
- In contrast the value function only depends on the current values (t_0, x_0).
- In summary our main object of study will be the function

$$v(t_0, x_0) = \inf_{u \in \mathcal{U}_{t_0}} \mathbb{E}\left[\int_{t_0}^{T} f(t, X^{t_0,x_0}(t), u(t))dt + g(X^{t_0,x_0}(T)) \middle| X(t_0) = x_0 \right].$$

- Note that we are looking at a family of problems, indexed by the starting point of the SDEs (t_0, x_0).

Bellman's Optimality Principle, Dynamic Programming and The Hamilton Jacobi Bellman Equation

Motivation in Discrete Time

- Suppose we are given a discrete time grid, $[0, 1, \dots, T]$. At each stage we choose a control $u(i)$ an $\mathcal{F}(i)$ measurable random variable.
- The discrete time analogue of our controlled system is the following

$$X(i+1) = F(X(i), u(i)), \quad X(t_0) = x_0$$

- We want to study the cost and value functions

$$J(t_0, x_0, u) = \mathbb{E}\left[\sum_{i=t_0}^{T-1} f\left(X^u(i), u(i)\right) + g(X^u(T)) \middle| \mathcal{F}(t_0)\right],$$

$$v(t_0, x_0) = \inf_{u \in \mathcal{U}_{t_0}} J(t_0, x_0, u)$$

- Note the similarity with the previous formula.

Backward Induction

- First note that we have

$$J(T, x_0, u) = v(T, x_0) = g(x_0) \text{ for all } x_0 \in \mathbb{R}^n$$

- At T we can describe v perfectly, in particular v describes the minimum costs we can achieve if we start at time T with initial value x_0.
- Let us look next at time $T-1$ and suppose we start at $X_{T-1} = x_0$. We have

$$\begin{aligned} v(T-1, x_0) &= \inf_{u_{T-1} \in U} \mathbb{E}\left[f(X_{T-1}, u_{T-1}) + g(X_T^u) \middle| \mathcal{F}_{T-1} \right] \\ &= \inf_{u_{T-1} \in U} \mathbb{E}\left[f(X_{T-1}, u_{T-1}) + v(T, X_T^u) \middle| \mathcal{F}_{T-1} \right] \\ &= \inf_{u_{T-1} \in U} \mathbb{E}\left[f(X_{T-1}, u_{T-1}) + v\big(T, F(X_{T-1}, u_{T-1})\big) \middle| \mathcal{F}_{T-1} \right] \end{aligned}$$

Backward Induction - II

- If our optimal control exists we should be able to do this minimisation to find $\hat{u}(X_{T-1}) = \hat{u}(x_0)$.
- This then leads to an expression for $v(T-1, x_0)$.
- We now proceed by backward induction, which leads to the formula

$$v(T-2, x_0) = \inf_{u_{T-2} \in U} \mathbb{E}\left[f(X_{T-2}, u_{T-2}) + v\big(T-1, F(X_{T-2}, u_{T-2})\big) \Big| \mathcal{F}_{T-2} \right]$$

- This then leads to the formula for general t_0,

$$v(t_0, x_0) = \inf_{u_{t_0} \in U} \mathbb{E}\left[f(X_{t_0}, u_{t_0}) + v\big(t_0+1, F(X_{t_0}, u_{t_0})\big) \Big| \mathcal{F}_{t_0} \right]$$

- This is known as the Dynamic Programming Principle.

Bellman's Optimality Principle

- This property of the value function was first observed by Richard Bellman in the 1960s as his Principle of Optimality

  An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

- You see the graphical interpretation of his principle in the diagrams.
- The dynamic programming principle is the basis for solving control problems, (in discrete time) it is a functional equation for v.

Dynamic Programming - Continuous Time

- In continuous time the DPP takes the following form

Theorem (Dynamic Programming Principle)

For any stopping time τ valued in $[t_0, T]$ we have

$$v(t_0, x_0) = \inf_{u \in \mathcal{U}_{t_0,\tau}} \mathbb{E}\left[\int_{t_0}^{\tau} f\big(t, X^u(t), u(t)\big)\, dt + v\big(\tau, X^u(\tau)\big) \Big| \mathcal{F}(t_0) \right],$$

where $\mathcal{U}_{t_0,\tau}$ denotes those controls admissible for X^{t_0,x_0} on $[t_0, \tau]$.

- Whereas before the DPP led to a functional equation at distinct time points, now we choose $\tau = t_0 + h$ and the above theorem implies a PDE, the Hamilton-Jacobi-Bellman (HJB) PDE which characterises the local form of the value function.
- Then we want to show that together with boundary conditions, $v(T, x) = g(x)$, this describes v uniquely.

The HJB Equation

Theorem (Hamilton Jacobi Bellman Equation)

Assume that the value function $v \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^n)$. Then it solves the following PDE.

$$\inf_{u_0 \in U} \{v_t(t,x) + \mathcal{L}_X^{u_0} v(t,x) + f(t,x,u_0)\} = 0,$$

$$v(T,x) = g(x),$$

$$\mathcal{L}_X^{u_0} := \mu(t,x,u_0)^\intercal D + \frac{1}{2}\mathrm{Tr}\left(\sigma\sigma^\intercal(t,x,u_0)D^2\right).$$

- If we perform the minimisation in the above analytically we will have

$$\hat{u}(t,x) = H(t,x,Dv(t,x),D^2v(t,x)).$$

- Intuitively we should then have that the optimal control is Markov and (provided it is admissible) is given by

$$\hat{u}(t,\hat{X}(t)) = H(t,\hat{X}(t),Dv(t,\hat{X}(t)),D^2v(t,\hat{X}(t))).$$

Verification Theorems

Uniqueness - Verification theorem

- We have shown that (when sufficiently smooth) the value function satisfies the HJB equation. Let us now address the question of uniqueness,

Theorem (Verification Theorem)

Let $\varphi \in \mathcal{C}^{1,2}([0,T] \times \mathbb{R}^n)$ be a solution to the HJB PDE. Suppose that the functions f and φ have quadratic growth, i.e

$$|\varphi(t,x)| + |f(t,x,u)| \leq C(1+|x|^2), \text{ for all } (t,x,u).$$

In addition suppose that there exists a unique minimiser $\hat{u}(t,x)$ of the function $u \mapsto \{\varphi_t + \mathcal{L}_X^u \varphi(t,x) + f(t,x,u)\}$ and that for all initial data $(t_0,x_0) \in [0,T] \times \mathbb{R}^n$ the equation

$$d\hat{X}(t) = \mu(t,\hat{X}(t),\hat{u}(t,\hat{X}(t)))dt + \sigma(t,\hat{X}(t),\hat{u}(t,\hat{X}(t))dW(t),$$
$$\hat{X}^{t_0,x_0}(t_0) = x_0 \in \mathbb{R}^n$$

has a unique strong solution (i.e. the control is admissible). Then $v = \varphi$ and the control $\hat{u}$ is an admissible Markov control.

Solution technique for control problems

A plan of attack is now clear.

- Formulate the control problem.
- Write down the HJB PDE.
- Apply PDE theory/ansatz to find a solution.
- Use the verification theorem to confirm that you indeed have the solution.

This is the standard recipe and works well in a few famous scenarios.

Existence of Classical Solutions

- The theory of second order nonlinear PDEs can provide existence in certain cases, I give an example of an appropriate result here, note that the conditions are restrictive.

Theorem (Existence)

Suppose that the matrix σ is uniformly elliptic, i.e there exists $\delta > 0$ such that for any $\xi \in \mathbb{R}^n$,

$$\xi^\intercal \sigma\sigma^\intercal (t, x, u)\xi \geq \delta|\xi|^2 \text{ for all } (t, x, u).$$

In addition suppose that, U is compact, μ, σ and f are $\mathcal{C}_b^{1,2}([0, T] \times \mathbb{R}^n)$ and $g \in \mathcal{C}_b^3(\mathbb{R}^n)$.
Then the HJB PDE has a unique $\mathcal{C}^{1,2}$ solution.

- Unfortunately it is in general very very hard to prove apriori regularity of the value function to illustrate I give now a deterministic example.

Nonsmooth value function

Example

We set $\sigma = 0, \mu(t,x,u) = u$ *and* $U = [-1,1]$. *The dynamics then become*

$$X(t) = x_0 + \int_{t_0}^{t} u(s)ds, X(t_0) = x_0.$$

We choose $g(x) = -x^2$ *thus we aim to minimise*

$$v(t_0,x_0) = \inf_{u \in \mathcal{U}_{t_0}} -\left(x_0 + \int_{t_0}^{t} u(s)ds\right)^2.$$

It is clear we have

$$v(t_0,x_0) = \begin{cases} -(x_0 + T - t_0)^2, & x \geq 0, \hat{u} = 1 \\ -(x_0 - T + t_0)^2, & x < 0, \hat{u} = -1. \end{cases}$$

This is not smooth at $x = 0$.

$$\lim_{x \downarrow 0^+} v_x(t_0,x) = -2(T - t_0) \neq -2(t_0 - T) = \lim_{x \uparrow 0^-} v_x(t_0,x).$$

Linear Quadratic Regulator Problems

Riccati Equations

- To solve these problems we shall need some elements of the theory of ODEs, in particular Riccati equations
- Recall that a general ODE has the form

$$\frac{dy}{dt} = F(t, y).$$

- Fixing t we approximate this as

$$\frac{dy}{dt} = F_1(t) + F_2(t)y + F_3(t)y^2.$$

- The above is called a Riccati Equation and you may think of such equations as one step up from linear ODEs. They were studied by the mathematician Jacopo Riccati in the 17^{th} century.

Riccati Equations - II

- Such equations always have a closed form solution, in particular given a particular solution $y_0(t)$, one can find the general solution as

$$y(t) = y_0(t) + \frac{1}{z(t)},$$

  where the function z solves

$$\frac{dz}{dt} = -(F_2(t) + 2y_0(t)F_3(t))z - F_3(t).$$

- One should think of these functions (indeed all quadratic ODEs) as being like tan and tanh, this follows because you have

$$D\tan = 1 + \tan^2, \quad D\tanh = 1 - \tanh^2$$

- Such equations are key to the solution of the famous Linear Quadratic Regulator Problem.

Formulation

- One of the oldest problems in stochastic control and very popular.
- The name arises since the underlying SDE has linear dynamics and the cost functions are quadratic.
- To fix ideas, we assume that $n = 1$, the X dynamics are

$$dX^{t_0,x_0}(t) = (AX^{t_0,x_0}(t) + Bu(t))dt + \sigma dW(t), \quad X^{t_0,x_0}(t_0) = x_0.$$

- The cost functions are given by $f(t,x,u) = F(x^2 + u^2)$, $g(t,x,u) = Gx^2$, we assume that $A, B \neq 0$ and that $\sigma, F, G > 0$.
- This leads us to the stochastic control problem

$$v(t_0,x_0) = \inf_{u \in \mathcal{U}_{t_0}} \mathbb{E}\left[\int_{t_0}^{T} FX^2(t) + Fu^2(t)dt + GX^2(T)\middle| X(t_0) = x_0\right].$$

Intuitive interpretation

- Our costs are quadratic in the process X, thus ideally we want "small" X.
- To achieve this we use the process u, however when we use u this also has associated costs $\sim Fu^2$.
- The aim is therefore to balance these two conflicting objectives.
- On top of this there is noise in the system σdW which affects any input u we might make.
- One could think of X as the deviation of a rocket from its target, with the u being fuel, or perhaps even an idealised stock holding trajectory with u being trading rate.
- It is important to think about the coefficients and their values, we should certainly ask $F > 0$ else we have no costs, additionally $\sigma > 0$ so we guarantee a stochastic problem.

Problem Solution

- We proceed through the standard steps, the generator for our controlled SDE is given, for a $\mathcal{C}^2$ function φ, by

$$\mathcal{L}^u\varphi = (Ax + Bu)\varphi_x + \frac{1}{2}\sigma^2\varphi_{xx}.$$

- This leads to the HJB PDE, which the value function should satisfy, assuming it is sufficiently smooth.

$$\inf_{u\in\mathbb{R}}\left\{v_t + (Ax + Bu)v_x + Fu^2 + Fx^2 + \frac{1}{2}\sigma^2 v_{xx}\right\} = 0, \quad v(T, x) = Gx^2.$$

- Performing the minimization, we see that the candidate optimal control should be

$$\hat{u}(t,x) = \frac{-Bv_x(t,x)}{2F}.$$

- Note that here we need $F > 0$ else the minimum is not defined. Since $F > 0$ we have a convex function in u and so it has a unique minimum (for fixed (t, x)).

Problem Solution - II

- Inserting this back in again we are lead to our candidate PDE

$$\left\{ v_t + Axv_x + Fx^2 - \frac{(Bv_x)^2}{4F} + \frac{1}{2}\sigma^2 v_{xx} \right\} = 0, \quad v(T,x) = Gx^2.$$

- There is no obvious solution to this. However if we look at the formulation of the running and cost functions a first suggestion would be that value function should be quadratic (in x), fitting the terminal condition of the PDE.
- Taking this a stage further, for a function quadratic in x the derivative would be linear. This would then separate the PDE.
- We thus propose the ansatz φ defined by

$$\varphi(t,x) = \frac{1}{2}a(t)x^2 + c(t).$$

Reduction to Riccati equations

- Upon substitution this leads to the system of ODEs

$$a'(t) + 2Aa(t) - \frac{B^2}{2F}a^2(t) + 2F = 0, \quad a(T) = 2G$$
$$c'(t) + \frac{1}{2}\sigma^2 a(t) = 0, \quad c(T) = 0.$$

- We see that a is a Riccati equation, which we know we can solve in closed form. Moreover, given a, c is a simple ODE which can just be solved by integration.
- Now we turn to verification, the quadratic growth of φ comes from the ansatz and noting that we deal with continuous functions a and c on the compact set $[0, T]$.
- The candidate optimal control has explicit form given by

$$\hat{u}(t,x) = \frac{-Ba(t)x}{2F}.$$

Verification

- Thus the optimal system has dynamics given by

$$d\hat{X}(t) = \left(A - \frac{B^2 a(t)}{2F}\right)\hat{X}(t)dt + \sigma dW(t), \quad \hat{X}(t_0) = x_0.$$

- This is a linear SDE and so we know that it satisfies our Lipschitz assumptions. This means

$$\mathbb{E}\left[\sup_{t_0 \le t \le T} \hat{X}^2(t)\right] < \infty,$$

so that the candidate control is admissible.

- Using the positivity of the running and terminal cost functions we can now verify the optimality of φ.

The Value Function of a Linear Quadratic Regulator Problem

Theorem

- *For the linear quadratic regulator problem, the value function v is quadratic and given by*

$$v(t_0, x_0) = \frac{1}{2}a(t_0)x_0^2 + c(t_0).$$

where a and c solve the system of ODEs given above.

- *There is an optimal Markov control and it is given by the function*

$$\hat{u}(t, x) = \frac{-Ba(t)x}{2F}.$$

- To understand the solution, we look at the value function (costs) for different starting values (t_0, x_0).
- When $x_0 = 0$ we are lead to the interpretation of c as the cost of noise in the system.

Interpretation of solution

- Viewed as a function of x_0 the initial point which gives the minimal cost trajectory is $x_0 = 0$.
- Thus intuitively we should control the system to move toward this point. Writing the control as

$$\hat{u}(t, \hat{X}(t)) = \frac{a(t)}{2F}\left(-B\hat{X}(t) - 0\right),$$

  we see that this is indeed the case, noting that the sign of B influences which direction the control is applied.
- We then see that the control is proportional to the displacement from this point of minimum cost.
- The proportion is decided by the ratio of the time varying function a (discount factor) and the cost of using the control F.

Stochastic Control Summary

- Stochastic optimal control is the study of optimisation problems where we can influence the a set of state variables which evolve stochastically.
- The key idea in attacking these problems is dynamic programming. Connecting the value function at different starting points allows us to characterise its local behaviour.
- This led us to the HJB equation, a second order nonlinear PDE. When one can find a solution to this equation one can typically characterise the value function.
- This led to a 4-step plan for solving stochastic control problems which was illustrated via the Linear Quadratic Regulator.

The Stochastic Control Literature

- For the topics covered thus far I recommend the three articles

  Stochastic calculus, filtering and stochastic control - van Handel, R.

  Continuous time stochastic control with applications - Pham, H.

  Stochastic control, viscosity solutions and applications to finance - Touzi, N.
- These provide rigorous proofs of all of the theorems you have seen here as well as providing an introduction to the more advanced topics
- If one is interested in a nice overview of all of stochastic control, the following article is nice

  On some recent applications of stochastic control and their applications - Pham, H. (Probability Surveys).

Disclaimer