→ Start of lecture 11a ←

Continuous Time Finance

Stochastic Control Theory

Ch 19

Tomas Björk

→ Theory of stochastic control

Financial applications in lecture 12 (optimal investment and consumption)

Contents

1. Dynamic programming. (mathematical background)

2. Investment theory. (economic application, next lecture)

1. Dynamic Programming

- The basic idea.
- Deriving the HJB equation.
- The verification theorem.
- The linear quadratic regulator. (classic example from systems theory)

Problem Formulation

$$\max_{u} \quad E\left[\int_0^T F(t, X_t, u_t)dt + \Phi(X_T)\right]$$

subject to

Markov process for fixed u, $X_t = X_t^u$

$$\begin{aligned} dX_t &= \mu(t, X_t, u_t)\, dt + \sigma(t, X_t, u_t)\, dW_t \\ X_0 &= x_0, \\ u_t &\in U(t, X_t), \quad \forall t. \end{aligned}$$

often $\equiv U$ (for all t, x)

We will only consider **feedback control laws**, i.e. controls of the form

$$u_t = \mathbf{u}(t, X_t)$$

partly justified by previous Markov property

X^u will still be Markov

Terminology:

	X	= state variable	$X_t \in R^n$
	u	= control variable	$u_t \in R^k$
the set	U	= control constraint	$U \subset R^k$

Note: No state space constraints. (e.g. $X_t \geq 0$)

Main idea

- Embedd the problem above in a family of problems indexed by starting point in time and space. See p. 328

- Tie all these problems together by a PDE: the Hamilton Jacobi Bellman equation. See p. 338

HJB

- The control problem is reduced to the problem of solving the deterministic HJB equation.

can be a very complicated equation,

but it gives a way to "compute" the solution to the original problem, and in a way, it is an "easier" problem

Some notation

(looks messy, but try to see 'through')

- For any fixed vector $u \in R^k$, the functions μ^u, σ^u and C^u are defined by

$$
\left.\begin{aligned}
\mu^u(t,x) &= \mu(t,x,u), \\
\sigma^u(t,x) &= \sigma(t,x,u), \\
C^u(t,x) &= \sigma(t,x,u)\sigma(t,x,u)'.
\end{aligned}\right\} \text{functions of 3 variables}
$$

i.e. $u_t = \mathbf{u}(t,x)$

- For any control law $\mathbf{u}$, the functions $\mu^{\mathbf{u}}$, $\sigma^{\mathbf{u}}$, $C^{\mathbf{u}}(t,x)$ and $F^{\mathbf{u}}(t,x)$ are defined by ($\mathbf{u}$ is a function)

these are functions of two variables: t,x But there is also $\mathbf{u}$

$$
\begin{aligned}
\mu^{\mathbf{u}}(t,x) &= \mu(t,x,\mathbf{u}(t,x)), \\
\sigma^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x)), \\
C^{\mathbf{u}}(t,x) &= \sigma(t,x,\mathbf{u}(t,x))\sigma(t,x,\mathbf{u}(t,x))', \\
F^{\mathbf{u}}(t,x) &= F(t,x,\mathbf{u}(t,x)).
\end{aligned}
$$

Itô: $df(t,X_t) = f_t(t,X_t)dt + \underbrace{\mathcal{A}^u f(t,X_t)}dt + f_x(t,X_t)\sigma(t,X_t)dW_t$

More notation (to confuse you more)

- For any fixed vector $u \in R^k$, the partial differential operator $\mathcal{A}^u$ is defined by

$$\mathcal{A}^u = \sum_{i=1}^{n} \mu_i^u(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} C_{ij}^u(t,x)\frac{\partial^2}{\partial x_i \partial x_j}.$$

(Generator of X^u)

- For any control law $\mathbf{u}$, the partial differential operator $\mathcal{A}^{\mathbf{u}}$ is defined by

$$\mathcal{A}^{\mathbf{u}} = \sum_{i=1}^{n} \mu_i^{\mathbf{u}}(t,x)\frac{\partial}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{n} C_{ij}^{\mathbf{u}}(t,x)\frac{\partial^2}{\partial x_i \partial x_j}.$$

- For any control law $\mathbf{u}$, the process $X^{\mathbf{u}}$ is the solution of the SDE

notation!

$$dX_t^{\mathbf{u}} = \underbrace{\mu\left(t, X_t^{\mathbf{u}}, \mathbf{u}_t\right)}_{\mu^{\mathbf{u}}(t,X_t^{\mathbf{u}})} dt + \sigma\left(t, X_t^{\mathbf{u}}, \mathbf{u}_t\right) dW_t,$$

where

$$\mathbf{u}_t = \mathbf{u}(t, X_t^{\mathbf{u}})$$

So: $dX_t^{\mathbf{u}} = \mu(t, X_t^{\mathbf{u}}, \mathbf{u}(t,X_t^{\mathbf{u}}))dt + \sigma(\cdot\cdot)dW_t$

is of the form

$= \tilde{\mu}(t, X_t^{\mathbf{u}})dt + \sigma(t, X_t^{\mathbf{u}})dW_t \Rightarrow$ Markov solution

Embedding the problem of p.324 into a family of problems $\mathcal{P}_{t,x}$

For every fixed (t, x) the control problem $\mathcal{P}_{t,x}$ is defined as the problem to maximize [from initial time t, and initial value x: $X_t^u = x$]

$$E_{t,x}\left[\int_t^T F(s, X_s^{\mathbf{u}}, u_s)ds + \Phi\left(X_T^{\mathbf{u}}\right)\right],$$

$= E\left[\int_t^T F ds + \Phi \,\middle|\, \mathcal{F}_t\right]_{X_t^u = x} = E\left[\int_t^T F ds + \Phi \,\middle|\, X_t^u = x\right]$

Markov

given the dynamics

$$\begin{aligned}
dX_s^{\mathbf{u}} &= \mu\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right)ds + \sigma\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right)dW_s, \\
X_t^{u} &= x,
\end{aligned}$$

and the constraints

$$\mathbf{u}(s, y) \in U, \quad \forall (s, y) \in [t, T] \times R^n.$$

The original problem was $\mathcal{P}_{0,x_0}$, as special instance of $\mathcal{P}_{t,x}$ with $t=0$, $x = x_0$

The optimal value function

- The **value function**

$$\mathcal{J} : R_+ \times R^n \times \mathcal{U} \to R$$

is defined by (recall x, the initial value at time t)

$$\mathcal{J}(t, x, \mathbf{u}) = E\left[\int_t^T F(s, X_s^{\mathbf{u}}, \mathbf{u}_s)ds + \Phi\left(X_T^{\mathbf{u}}\right)\right]$$

given the dynamics above. (and $X_t^{\mathbf{u}} = x$)

Note: in fact $X_s^{\mathbf{u}}$ also depends on x for all $s \geq t$; should write $X_s^{\mathbf{u},x}$, but more cumbersome

- The **optimal value function**

$$V : R_+ \times R^n \to R$$

is defined by (recall we want to maximize)

$$V(t, x) = \sup_{\mathbf{u} \in \mathcal{U}} \mathcal{J}(t, x, \mathbf{u}).$$ Note: $V(T,x) = \Phi(x)$

Our aim:

- We want to derive a PDE for V.

If sup is attained, then there is some $\mathbf{u} = \hat{\mathbf{u}} = \hat{\mathbf{u}}(t,x)$

Assumptions

We assume:

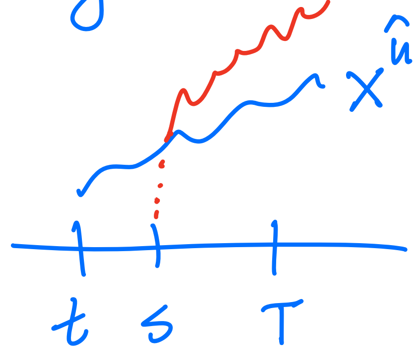
- There exists an optimal control law $\hat{\mathbf{u}}$. ($\rightarrow \hat{u}(t,x)$, $\hat{u}(t, X_t)$)
- The optimal value function V is regular in the sense that $V \in C^{1,2}$.
- A number of limiting procedures in the following arguments can be justified. We will make big steps and ignore many mathematical details that would require a finer analysis; beyond the scope and aims of this course.

Bellman Optimality Principle

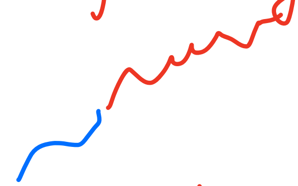
Theorem: If a control law $\hat{\mathbf{u}}$ is optimal for the time interval $[t, T]$ then it is also optimal for all smaller intervals $[s, T]$ where $s \geq t$.

Proof: Exercise.

Use a concatenation argument: other trajectory, that is supposed to have a better performance on $[s, T]$ than $\hat{u}$.



But then combined trajectory



would have a better performance on $[t, T]$: contradiction

Basic strategy

HJB

To derive the PDE do as follows:

- Fix $(t, x) \in (0, T) \times R^n$.
- Choose a real number h (interpreted as a "small" time increment).
- Choose an arbitrary control law $\mathbf{u}$ on the time inerval $[t, t+h]$.

Now define the control law $\mathbf{u}^\star$ by

$$\mathbf{u}^\star(s, y) = \begin{cases} \mathbf{u}(s, y), & (s, y) \in [t, t+h] \times R^n \\ \hat{\mathbf{u}}(s, y), & (s, y) \in (t+h, T] \times R^n. \end{cases}$$

In other words, if we use $\mathbf{u}^\star$ then we use the arbitrary control $\mathbf{u}$ during the time interval $[t, t+h]$, and then we switch to the optimal control law during the rest of the time period.

Note that u^* is worse than $\hat{u}$ on $[t, T]$

Basic idea

The whole idea of DynP boils down to the following procedure.

- Given the point (t, x) above, we consider the following two strategies over the time interval $[t, T]$:

  I: Use the optimal law $\hat{\mathbf{u}}$.

  II: Use the control law $\mathbf{u}^{\star}$ defined above on p. 332

- [if you can] Compute the expected utilities obtained by the respective strategies.

- Using the obvious ! fact that $\hat{\mathbf{u}}$ is at least as good as $\mathbf{u}^{\star}$, and letting h tend to zero, we obtain our fundamental PDE. (in this step we will reason rather heuristically)

→ End of lecture 11a ←

→ Start of lecture 11 b ←

Strategy values

I: **Expected utility for $\hat{\mathbf{u}}$:**

$$\mathcal{J}(t, x, \hat{\mathbf{u}}) = V(t, x)$$

(p. 329, definition of V)

II: **Expected utility for $\mathbf{u}^{\star}$:** Split the time interval $[t, T]$:

- The expected utility for $[t, t+h)$ is given by

$$E_{t,x}\left[\int_t^{t+h} F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right) ds\right].$$

- Conditional expected utility over $[t+h, T]$, given (t, x):

$$E_{t,x}\left[V(t+h, X_{t+h}^{\mathbf{u}})\right].$$

starting point at time t+h reached from $x_t = x$ by applying $\mathbf{u}$

- Total expected utility for Strategy II is

$$V_{II} = E_{t,x}\left[\int_t^{t+h} F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right) ds + V(t+h, X_{t+h}^{\mathbf{u}})\right]. \quad (*)$$

Comparing strategies

(the math here is a bit sloppy)

We have trivially (V results from optimal $\hat{u}$, strategy I is optimal)

$$(**) \quad V(t,x) \geq E_{t,x}\left[\int_t^{t+h} F\left(s, X_s^{\mathbf{u}}, \mathbf{u}_s\right) ds + V(t+h, X_{t+h}^{\mathbf{u}})\right] = \; = V_{\mathrm{II}}$$

Remark (trivial)

We have equality above if and only if the control law $\mathbf{u}$ is the optimal law $\hat{\mathbf{u}}$.

Now use Itô to obtain

$$V(t+h, X_{t+h}^{\mathbf{u}}) = V(t,x)$$

$$+ \int_t^{t+h} \left\{ \frac{\partial V}{\partial t}(s, X_s^{\mathbf{u}}) + \mathcal{A}^{\mathbf{u}} V(s, X_s^{\mathbf{u}}) \right\} ds$$

$$+ \underbrace{\int_t^{t+h} \nabla_x V(s, X_s^{\mathbf{u}}) \sigma^{\mathbf{u}} dW_s,}_{\text{has conditional expectation } E[\cdot\cdot|\mathcal{F}_t]=0}$$

and plug into the formula above.

if a true martingale (needs conditions)

(*) on p. 334

We obtain, also using (**),

$$E_{t,x}\left[\int_t^{t+h}\left\{F\left(s,X_s^{\mathbf{u}},\mathbf{u}_s\right)+\frac{\partial V}{\partial t}(s,X_s^{\mathbf{u}})+\mathcal{A}^{\mathbf{u}}V(s,X_s^{\mathbf{u}})\right\}ds\right]\le 0.$$

Going to the limit:
Divide by h, move h within the expectation and let h tend to zero.
We get, with $X_t^u = x$:

$$F(t,x,u)+\frac{\partial V}{\partial t}(t,x)+\mathcal{A}^u V(t,x)\le 0,$$

$\left(\frac{1}{h}\int_t^{t+h} g(s)ds \to g(t) \text{ for continuous } g \right)$

Recall from previous slide:

$$F(t,x,u)+\frac{\partial V}{\partial t}(t,x)+\mathcal{A}^u V(t,x)\leq 0,$$

This holds for all $u=\mathbf{u}(t,x)$, with equality if and only if $\mathbf{u}=\hat{\mathbf{u}}$.

We thus obtain the **HJB equation**

$$\frac{\partial V}{\partial t}(t,x)+\sup_{u\in U}\left\{F(t,x,u)+\mathcal{A}^u V(t,x)\right\}=0.$$

The HJB equation

Theorem:
Under suitable regularity assumptions the follwing hold:

I: V satisfies the Hamilton–Jacobi–Bellman equation

$$\begin{aligned}
\frac{\partial V}{\partial t}(t,x) + \sup_{u \in U} \left\{ F(t,x,u) + \mathcal{A}^u V(t,x) \right\} &= 0, \\
V(T,x) &= \Phi(x),
\end{aligned}$$

II: For each $(t,x) \in [0,T] \times R^n$ the supremum in the HJB equation above is attained by $u = \hat{\mathbf{u}}(t,x)$, i.e. by the optimal control.

Logic and problem

Note: We have shown that **if** V is the optimal value function, and **if** V is regular enough, **then** V satisfies the HJB equation. The HJB eqn is thus derived as a **necessary** condition, and requires strong *ad hoc* regularity assumptions, alternatively the use of viscosity solutions techniques.

Problem: Suppose we have solved the HJB equation. Have we then found the optimal value function and the optimal control law? In other words, is HJB a **sufficient** condition for optimality.

Answer: Yes! This follows from the **Verification Theorem**.

The Verification Theorem

Suppose that we have two functions $H(t,x)$ and $g(t,x)$, such that

- H is sufficiently integrable, and solves the HJB equation

$$\begin{cases} \frac{\partial H}{\partial t}(t,x) + \sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u H(t,x)\} & = & 0, \\ H(T,x) & = & \Phi(x), \end{cases}$$

- For each fixed (t,x), the supremum in the expression

$$\sup_{u \in U} \{F(t,x,u) + \mathcal{A}^u H(t,x)\}$$ ← static problem for each (t,x)

  is attained by the choice $u = g(t,x)$.

Then the following hold.

1. The optimal value function V to the initial control problem is given by

$$V(t,x) = H(t,x),$$ the function that solves HJB

2. There exists an optimal control law $\hat{\mathbf{u}}$, and in fact

$$\hat{\mathbf{u}}(t,x) = g(t,x)$$

Proof perhaps (see book pp 291-293)

Handling the HJB equation (Section 19.4)

1. Consider the HJB equation for V.
2. Fix $(t, x) \in [0, T] \times R^n$ and solve, the static optimization problem

(maximizer exists?) $$\max_{u \in U} \quad [F(t, x, u) + \mathcal{A}^u V(t, x)]$$ of slide 340

   Here u is the only variable, whereas t and x are fixed parameters. The functions F, μ, σ and V are considered as given.
3. The optimal $\hat{u}$, will depend on t and x, and on the function V and its partial derivatives. We thus write $\hat{u}$ as

$$\hat{\mathbf{u}} = \hat{\mathbf{u}}\,(t, x; V)\,. \tag{4}$$

4. The function $\hat{\mathbf{u}}\,(t, x; V)$ is our candidate for the optimal control law, but since we do not know V this description is incomplete. Therefore we substitute the expression for $\hat{u}$ into the PDE , giving us the highly nonlinear (why?) PDE

$$\begin{aligned} \frac{\partial V}{\partial t}(t, x) + F^{\hat{\mathbf{u}}}(t, x) + \mathcal{A}^{\hat{\mathbf{u}}}\,(t, x)\, V(t, x) &= 0, \\ V(T, x) &= \Phi(x). \end{aligned}$$

   if you can!

5. Now we solve the PDE above! Then we put the solution V into expression (4). Using the verification theorem we can identify V as the optimal value function, and $\hat{u}$ as the optimal control law.

Does this work in concrete situations?

Making an Ansatz

- The hard work of dynamic programming consists in solving the highly nonlinear HJB equation

- There are no general analytic methods available for this, so the number of known optimal control problems with an analytic solution is very small indeed.

- In an actual case one usually tries to **guess** a solution, i.e. we typically make a parameterized **Ansatz** for V then use the PDE in order to identify the parameters.

- **Hint:** V often inherits some structural properties from the boundary function Φ as well as from the instantaneous utility function F. (this is experience)

- Most of the known solved control problems have, to some extent, been "rigged" in order to be analytically solvable.

EXAMPLE

standard, classical problem in systems & contr

The Linear Quadratic Regulator

(section 19.5)

Now: min!

$$\min_{u \in R} \quad E\left[\int_0^T \left\{QX_t^2 + Ru_t^2\right\} dt + HX_T^2\right],$$

$X_t \in R^1$

with dynamics

$$dX_t = \{AX_t + Bu_t\} dt + CdW_t.$$

multidimensional

for each fixed u_t, this gives Gaussian X_t, OU process, LQG control problem

Example: We want to control a vehicle in such a way that it stays close to the origin (the terms Qx^2 and Hx^2) while at the same time keeping the "energy" Ru^2 small.

Here $X_t \in R$ and $\mathbf{u}_t \in R$, and we impose no control constraints on u.

The real numbers Q, R, H, A, B and C are assumed to be known. We assume that R is strictly positive.

Handling the Problem

The HJB equation becomes (use the generator: $Af = f_x \mu + \frac{1}{2} f_{xx} \sigma^2$ for $dX = \mu\, dt + \sigma\, dW$)

$$\begin{cases} \dfrac{\partial V}{\partial t}(t,x) & + \quad \inf_{u \in R} \left\{ Qx^2 + Ru^2 + V_x(t,x)\left[Ax + Bu\right] \right\} \\ & + \quad \frac{1}{2}\frac{\partial^2 V}{\partial x^2}(t,x)C^2 = 0, \\ V(T,x) & = \quad Hx^2. \end{cases}$$

$\frac{\partial V}{\partial x}$

For each fixed choice of (t,x) we now have to solve the static unconstrained optimization problem to minimize

$$Qx^2 + Ru^2 + V_x(t,x)\left[Ax + Bu\right].$$

The problem was:

$$\min_u \quad Qx^2 + Ru^2 + V_x(t,x)\left[Ax + Bu\right].$$

Since $R > 0$ we set the u-derivative to zero and obtain

$$2Ru = -V_x B,$$

which gives us the optimal u as

$$\hat{u} = -\frac{1}{2}\frac{B}{R}V_x.$$

Note: This is our candidate of optimal control law, but it depends on the unkown function V.

We now make an educated guess about the structure of V.

From the boundary function Hx^2 and the term Qx^2 in the cost function we make the Ansatz

Note $V(T,x) = Hx^2$

$$V(t,x) = P(t)x^2 + q(t),$$

where $P(t)$ and $q(t)$ are deterministic functions. to be found later

With this trial solution we have,

$$\begin{aligned}
\frac{\partial V}{\partial t}(t,x) &= \dot{P}x^2 + \dot{q}, \\
V_x(t,x) &= 2Px, \qquad (P = P(t) \text{ etc.}) \\
V_{xx}(t,x) &= 2P \\
\hat{u} &= -\frac{B}{R}Px.
\end{aligned}$$

(see p. 345)

Inserting these expressions into the HJB equation we get

$$x^2\left\{\dot{P} + Q - \frac{B^2}{R}P^2 + 2AP\right\}$$

$$\cancel{+\dot{q}PC^2 = 0.} \qquad + \dot{q} + PC^2 = 0, \ \forall x!$$

We thus get the following ODE for P

$$\begin{cases} \dot{P} &= \frac{B^2}{R}P^2 - 2AP - Q, \\ P(T) &= H. \end{cases}$$

from $V(T,x) = Hx^2$

and we can integrate directly for q:

$$\begin{cases} \dot{q} &= -C^2P, \\ q(T) &= 0. \end{cases}$$

The ODE for P is a **Riccati equation**. The equation for q can then be integrated directly, once you have P

Final Result for LQ: (note that P is not given explicitely)

$$\begin{aligned} V(t,x) &= P(t)x^2 + \int_t^T C^2P(s)ds, \\ \hat{\mathbf{u}}(t,x) &= -\frac{B}{R}P(t)x, \end{aligned}$$

← verify!

this is a linear feedback law

→ End of lecture 11b ←