ECON 2021 - Financial Economics I

Lecture 4 – Dynamic Programming

October 1, 2018

INTRODUCTION

- The basic problem in finance has two components:
  1. How much should I consume today vs. how much should I save?
  2. How should I invest my savings?

- These are inherently dynamic problems, so we must learn how to solve dynamic optimization problems.

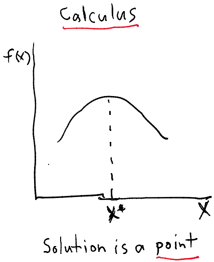
- We can interpret the solutions from either a normative perspective (ie, providing advice to investors), or from a positive perspective (ie, providing predictions about observed data).

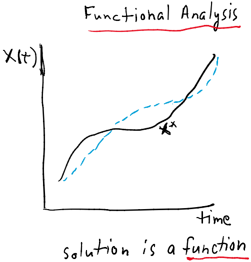
- As always, the solutions will depend on the interaction between preferences and market opportunities.

- There has been a lot of interesting work recently on the structure of intertemporal preferences.

Calculus vs. Functional Analysis

- Solving dynamic optimization problems is harder than solving static optimization problems. Static optimization problems involve finding a single optimal choice. They can be solved using calculus.
- Dynamic optimization problems involve finding an entire path. You can't choose myopically because today's choice has consequences for the future. These problems must be solved using functional analysis.





Why Dynamic Programming?

- There are 3 general approaches to these problems:
  1. Calculus of Variations $\Rightarrow$ Euler Eq.
  2. Optimal Control Theory $\Rightarrow$ Hamiltonian
  3. Dynamic Programming $\Rightarrow$ Bellman Eq.

- Dynamic Programming has distinct advantages in stochastic/uncertain settings, since it involves finding an optimal contingency plan. Future optimal choices are a function of the future state. You solve for the function rather than the choices themselves. (It clearly makes no sense to commit to a certain amount of future consumption until you know what your future income will be!)

- DP exploits the recursive structure of many dynamic optimization problems. When a problem is recursive, you can break it into 2 parts: Today & Tomorrow, exploiting the fact that Tomorrow's problem will be the same as Today's, except with new initial conditions determined by Today's decisions.

Principle of Optimality

- Bellman called this the Principle of Optimality:

Principle of Optimality

"Under certain conditions, optimal paths have the property that whatever the initial conditions and controls were over some initial period, the controls over the remaining periods must be optimal for the "remainder problem" given the state resulting from the earlier controls."

- Or, more succinctly:
  1. It is optimal to continue optimal paths.
  2. All parts of an optimal path are optimal.

- By construction, solutions produced by DP are dynamically consistent. It is not necessary to reconsider an optimal policy.

A Key Condition

Key condition for DP to produce the optimum:

- Today's actions influence current & future returns, but not past returns.

examples : · consumption determines current & future utility, but not past utility
· investment determines current & future profits, but not past profits

Or equivalently,

- Current returns depend on current & past actions, but not on future actions.

example : · current profits do not depend on future investment

- This condition is often violated in 2-agent/strategic settings, where agents often have incentives to makes promises or threats which they may not want to keep ex post.

HJB Equation

- Although DP is especially useful in stochastic settings, we start with the deterministic case, which is easier. As usual, we start in discrete-time, then take continuous-time limits.

Problem

$\max_u \int_0^T f(x,u,t)dt$ — $x =$ state (e.g., wealth)

$u =$ control (e.g., consumption)

subject to: (1) $\dot{x} = g(x,u,t)$ — $\dot{x} = \frac{dx}{dt}$

(2) $x(0), x(T)$ given — state transition eq. (eg, budget constraint)

- Start by defining the value function, $V(t_0, x_0)$, which is the optimized value of the above problem, given the state is x_0 and the time is t_0.

$$V(t_0, x_0) = \max_u \int_{t_0}^T f(x,u,t)dt \qquad \text{s.t.} \quad \dot{x} = g(x,u,t) \quad x(t_0) = x_0$$

- We can break this integral into 2 pieces:

$$V(t_0, x_0) = \max_{t_0 \le t \le t_0+\Delta t} \left\{ \int_{t_0}^{t_0+\Delta t} f\,dt + \int_{t_0+\Delta t}^{T} f\,dt \right\}$$

- If the principle of optimality applies,

$$V(t_0, x_0) = \max_{t_0 \le t \le t_0+\Delta t} \left\{ \int_{t_0}^{t_0+\Delta t} f\,dt + \overbrace{\max_{t_0+\Delta t \le t \le T} \left[\int_{t_0+\Delta t}^{T} f\,dt \right]}^{\text{when choosing } u \text{ you can neglect past } f\text{'s}} \right\}$$

$$\Rightarrow \quad V(t_0, x_0) = \max_{t_0 \le t \le t_0+\Delta t} \left\{ \int_{t_0}^{t_0+\Delta t} f\,dt + V(t_0 + \Delta t, x_0 + \Delta x) \right\}$$

For Δt small,

1. $\int_{t_0}^{t_0+\Delta t} f\,dt \approx f \cdot \Delta t$
2. $V(t_0 + \Delta t, x_0 + \Delta x) \approx V(t_0, x_0) + V_t \cdot \Delta t + V_x \cdot \Delta x$ } 1st-order Taylor series

Sub-in, divide by Δt, and letting $\Delta t \to 0$ we get the HJB equation

$$-V_t = \max_u \{ f(x, u, t) + V_x(t, x) \cdot g(x, u, t) \}$$

- Performing the maximization (which is a calculus problem) gives us the policy function, which provides a recursive representation of the optimal path.

$$u^* = h(x; V)$$

- Subbing u^* back in gives us an alternative representation of the HJB eq.

$$-V_t = f(x, h(x; V), t) + V_x(t, x) \cdot g(x, h(x; V), t)$$

- Note the HJB equation is a nonlinear partial differential equation. PDEs are the bread & butter of science. Unfortunately, they are notoriously difficult to solve. For linear PDEs, Fourier/Laplace transform methods provide a general strategy. For nonlinear PDEs, the only hope is a separation-of-variables/guess-and-verify approach.

- In general, there are many solutions to a PDE. Unique solutions are pinned down by boundary conditions.

Example:

$$\min_u \int_0^\infty e^{-rt}(ax^2 + bu^2)dt$$

$$\text{s.t.} \qquad \dot{x} = c \cdot x + u$$

$$\text{HJB:} \qquad -V_t = \min_u \left\{ e^{-rt}(ax^2 + bu^2) + V_x \cdot (cx + u) \right\}$$

FOC:(u): $\qquad 2e^{-rt}bu + V_x = 0$

$\Rightarrow \qquad \boxed{u = -\frac{V_x}{2b}e^{rt}}$

Sub the optimal u back into the HJB eq.

$$-V_t = e^{-rt}(ax^2 + \frac{1}{4b}V_x^2 e^{2rt}) + V_x \left(cx - \frac{V_x}{2b}e^{rt} \right)$$

Guess: $\qquad V(t, x) = e^{-rt}Ax^2$

$\Rightarrow \qquad V_t = -re^{-rt}Ax^2 \qquad V_x = 2e^{-rt}Ax$

- sub these into the HJB equation

$$re^{-rt}Ax^2 = ae^{-rt}x^2 + \frac{1}{b}e^{-rt}A^2x^2 + 2ce^{-rt}Ax^2 - \frac{2}{b}e^{-rt}A^2x^2$$

Note: There is common $e^{-rt}x^2$ term, which can be cancelled out. (This ability to cancel defines a correct guess!)

- we are left with the following quadratic eq. for A. We must select the positive root.

$$\frac{1}{b}A^2 + (r - 2c)A - a = 0$$

- We then get the optimal feedback policy

$$\boxed{u = -\frac{V_x}{2b}e^{rt} = -\frac{A}{b}x}$$

A Useful Shortcut

- Many econ/finance problems feature an *infinite* horizon (as a useful approximation), and time only enters as an exponential discount factor in the objective function. In these problems, time per se doesn't matter. (There is always tomorrow). As a result, the HJB partial diff. eq. reduces to a much easier ordinary diff. eq.
- Here's how it works

$$V(x) = \max_u \left\{ f(x,u) \cdot \Delta t + e^{-r\Delta t} V(x + \Delta x) \right\}$$

Note: $e^{-r\Delta t} \approx \frac{1}{1+r\cdot\Delta t}$

- Multiply both sides by $(1 + r \cdot \Delta t)$ and expand $V(x + \Delta x)$.

$$V(x)(1 + r \cdot \Delta t) = f(x,u) \cdot \Delta t + f(x,u) \cdot r(\Delta t)^2 + V(x) + V_x \cdot \Delta x$$

Simplify, divide by Δt, let $\Delta t \to 0$, and drop higher order terms

$$\underbrace{rV(x)}_{\text{riskless return}} = \max_u \{ \underbrace{f(x,u)}_{\text{dividends}} + \underbrace{V'(x) \cdot g(x,u)}_{\text{capital gain/loss}} \} \quad \} \text{ Stationary HJB Eq.}$$

Stochastic Dynamic Programming

- Let's now consider a more realistic situation, where the future evolution of the state is uncertain.

- The previous analysis goes through, with two exceptions:
  1. Since we don't know the future, we can only optimize expected returns/utility
  2. Doing the Taylor series approx. of $V(x)$ when obtaining the HJB eq. is a bit tricky when x is a function of Brownian motion. Since $dx \sim \sqrt{dt}$, we must expand to 2nd-order to get all the dt terms.

- Suppose

$$dx = \mu(x,u) \cdot dt + \sigma(x) dW$$

Given that x follows an Ito process, let's use Ito's lemma to derive the stochastic HJB eq. The value function is

$$V(x_t) = E_t \int_t^{\infty} e^{-r(s-t)} f(x,u) ds$$

- Following the same steps as before,

$$V(x) = \max_u \left\{ f(x,u)\Delta t + \frac{1}{1 + r \cdot \Delta t} E[V(x + \Delta x)|x,u] \right\}$$

- Multiply by $1 + r \cdot \Delta t$, subtract $V(x)$ from both sides, divide by Δt, and let $\Delta t \to 0$

$$rV(x) = \max_u \left\{ f(x,u) + \frac{1}{dt} E[dV] \right\}$$

- From Ito's lemma,

$$\begin{aligned} dV &= V_x dx + \frac{1}{2}\sigma^2(x) V_{xx} dt \\ &= V_x[\mu(x,u)dt + \sigma(x)dW] + \frac{1}{2}\sigma^2(x) V_{xx} dt \end{aligned}$$

$$\Rightarrow \quad E[dV] = \left(V_x \mu(x,u) + \tfrac{1}{2}\sigma^2(x) V_{xx} \right) dt$$

since $E[dW] = 0$

- Hence, we get the following stochastic HJB eq.

$$rV = \max_u \left\{ f(x,u) + \mu(x,u) \cdot V_x + \frac{1}{2}\sigma^2(x)V_{xx} \right\}$$

Note, this is a 2nd-order ODE.

Example 1: The Stochastic Linear-Quadratic Regulator

- Let's return to our earlier example, but now suppose the state evolves randomly,

$$\min_u E \int_0^\infty e^{-rt}(ax^2 + bu^2)dt$$

s.t. $dx = (cx + u)dt + \sigma dW$

- The stationary HJB eq. is

$$rV = \min_u \left\{ (ax^2 + bu^2) + (cx + u) \cdot V_x + \frac{1}{2}\sigma^2 V_{xx} \right\}$$

- FOC (u): $2bu + V_x = 0 \Rightarrow u = \frac{-1}{2b} V_x$
- Sub u back into HJB

$$rV = ax^2 + \frac{1}{4b} V_x^2 + cx \cdot V_x - \frac{1}{2b} V_x^2 + \frac{1}{2}\sigma^2 V_{xx}$$

- Guess: $V(x) = Ax^2 + B \Rightarrow V_x = 2Ax \qquad V_{xx} = 2A$
- Sub guess in HJB, $r(Ax^2 + B) = ax^2 - \frac{1}{b} A^2 x^2 + 2Acx^2 + \sigma^2 A$
- Match Coefficients
  1. $rA = a - \frac{1}{b} A^2 + 2Ac \Rightarrow \frac{1}{b} A^2 + (r - 2c)A - a = 0$ } same as before
  2. $rB = \sigma^2 A \Rightarrow B = \frac{\sigma^2 A}{r}$

Comments:

1. Note, the optimal policy is the same as before, $u = -\frac{A}{b} x$. Why? Why doesn't 'risk' matter? Does this mean risk is irrelevant?
2. Verify that if instead, $dx = (cx + u)dt + \sigma x dW$ (note, x now affects the variance of the shocks) then σ would influence behavior.

Example 2: A Stochastic Growth Model

- Consider the following growth model featuring a linear/stochastic technology and CRRA preferences:

$$\max_c E \int_0^\infty \frac{c^{1-\gamma}}{1-\gamma} e^{-\rho t} dt$$

s.t. $$dk = (\mu k - c) \cdot dt + \sigma k dW$$

- The stationary HJB eq. is,

$$\rho V = \max_c \left\{ \frac{c^{1-\gamma}}{1-\gamma} + (\mu k - c) \cdot V_k + \frac{1}{2}\sigma^2 k^2 V_{kk} \right\}$$

- FOC(u): $c^{-\gamma} = V_k \Rightarrow c = (V_k)^{-1/\gamma}$

- Guess: $V(k) = \frac{A}{1-\gamma} k^{1-\gamma} \Rightarrow V_k = Ak^{-\gamma} \qquad V_{kk} = -\gamma A k^{-\gamma-1}$

- Sub into HJB eq.

$$\rho \frac{A}{1-\gamma} k^{1-\gamma} = \frac{1}{1-\gamma} A^{1-1/\gamma} k^{1-\gamma} + \mu A k^{1-\gamma} - A^{1-1/\gamma} k^{1-\gamma} - \frac{1}{2}\sigma^2 \gamma A k^{1-\gamma}$$

Cancel $Ak^{1-\gamma}$ from both sides,

$$\rho = A^{-1/\gamma} + \mu(1-\gamma) - (1-\gamma)A^{-1/\gamma} - \frac{1}{2}\sigma^2\gamma(1-\gamma)$$

- Note, $c = (V_k)^{-1/\gamma} = A^{-1/\gamma}k$. So let's solve for $A^{-1/\gamma}$

$$A^{-1/\gamma} = \frac{1}{\gamma}\left[\rho - (1-\gamma)\mu + \frac{1}{2}\sigma^2\gamma(1-\gamma)\right]$$

Therefore,

$$c = \left[\mu + \frac{1}{\gamma}(\rho - \mu) - \frac{1}{2}\sigma^2(\gamma - 1)\right] k \quad \} \text{ Policy Function}$$

$$dk = \left[\frac{1}{\gamma}(\mu - \rho) + \frac{1}{2}\sigma^2(\gamma - 1)\right] \cdot k dt + \sigma k dW$$

Comments

1. The 2nd term in the consumption function is an intertemporal substitution effect. The 3rd term is a precautionary savings effect.

2. If $\gamma = 1$ (log utility), then we get Friedman's Permanent Income Hypothesis., $c = \rho k$.

3. The usual condition for 'endogenous growth' is that $\mu > \rho$. However, notice with uncertainty there can be sustained growth even if $\mu < \rho$, as long as $\gamma > 1$ and σ^2 is big enough.

VERIFICATION THEOREMS

- You may have noticed a chicken-or-the-egg aspect of DP. We use V to compute u, but u determines V! In practice, we start with a guess (of either V or u), and then iterate till convergence.
- Is this always valid? How do we know this produces a result that is independent of the initial guess?
- In discrete-time, one can derive existence and uniqueness results (Stokey-Lucas-Prescott). These conditions are rarely met in practice, and do not apply in continuous-time settings.
- In continuous-time, one can show that the HJB eq. is necessary, but you need to do extra work to show that it is sufficient. That is the role of a Verification Theorem.

Example:

- Let's go back to our doubling strategy. Consider an agent who can continuously bet on the instantaneous outcomes of a Brownian motion during the interval $[0, 1]$. Letting θ denote the size of his bet, his wealth evolves as

$$dW = \theta dB$$

  (Note, I've now used B to denote Brownian motion, since W is used to denote wealth). His portfolio is of course adapted to the filtration generated by B, so it is a random variable. Define the set $\mathcal{L}^2 = \{\theta : \int_0^1 \theta_s^2 ds < \infty \quad w.p.1\}$. Thus, his bets must be bounded w.p.1.

- Suppose his objective is to maximize a quadratic function of his expected terminal wealth,

$$\max_{\theta} E[-(\bar{W} - W(1))^2 | W(0) < \bar{W}]$$

- Since there is no discounting, drift, or flow payoff, the HJB equation is trivial

$$0 = \max_{\theta} \left\{ \frac{1}{2}\theta^2 V_{ww} \right\} \qquad V(W_1) = -(\bar{W} - W(1))^2$$

  Since $V_{ww} < 0$ (ie, the agent is risk averse), the FOC clearly implies that it is optimal to set $\theta = 0$. (By definition, a risk averter will never take a fair bet).

- However, we know this is not the optimal strategy! By engaging in a doubling strategy he can (w.p.1) guarantee himself $\bar{W}$.

- To see this, define the stopping time $\tau = \inf\{t > 0 : B_t = b\}$ where b is some threshold. First, note that $\tau < \infty$ w.p.1 and $E(\tau) = \infty$.

As an illustration of the power of martingale reasoning, let's prove this. Suppose $B_0 = 0$ and let $A = (-a, b)$ be some interval containing B_0. Let $\tau = \inf\{t : B_t = -a \quad \text{or} \quad B_t = b\}$ be the escape time from the interval. We know B_t and $B_t^2 - t$ are martingales. Therefore, letting $p = \text{prob}(B_\tau = b)$, we have $E(B_\tau) = pb + (1-p)(-a) = 0$. Thus, $p = \frac{a}{a+b}$. From the 2nd martingale property, $E(\tau) = E(B_\tau^2) = pb^2 + (1-p)a^2$. Subbing in for p and solving we get, $E(\tau) = ab$. Finally, note that $p \to 1$ and $E(\tau) \to \infty$ as $a \to \infty$. In other words, B_t will hit b w.p.1., but the mean time until it does so is infinite!

- Let α be the stopping time till hitting the wealth threshold. Consider the portfolio strategy,

$$\begin{aligned} \theta_t &= \frac{1}{\sqrt{1-t}} \qquad t < \alpha \\ &= 0 \qquad t \geq \alpha \end{aligned}$$

Notice that bets get larger and larger until the threshold is hit. An important result in the theory of continuous time martingales is that a change of variance is equivalent to a change in time-scale. In this case, the betting strategy induces a standardized Brownian motion, Y_t, over a transformed, infinite horizon, time-scale

$$Y_t = \hat{B}(\beta_t) \qquad \beta_t = \int_0^t \frac{ds}{1-s} = \log\left(\frac{1}{1-t}\right) \qquad 0 \leq t < 1$$

- We can now appeal to our earlier hitting time results for standard Brownian motion. First, the fact that we hit our threshold w.p.1 tells us that our $\mathcal{L}^2$ constraint on portfolios is satisfied,

$$\int_0^1 \theta_s^2 ds = \int_0^\alpha \frac{ds}{1-s} = \log\left(\frac{1}{1-\alpha}\right) = \tau < \infty \qquad \text{w.p.1}$$

- However, the fact that the mean hitting time (on the elongated, infinite horizon, time scale) is infinite tells us that the variance of our portfolio strategy (and wealth) is infinite,

$$E\left[\int_0^1 \theta_s^2 ds\right] = E\left[\log\left(\frac{1}{1-\alpha}\right)\right] = E(\tau) = \infty$$

Hence, if the agent cared about the variance of his interim wealth, he would never engage in this strategy.

- What does this have to do with DP and Verification Theorems? Remember that when deriving the HJB eq. we used Ito's lemma to calculate $E[dV]$. When doing this, we assumed $E\int \sigma(x)V_x dW = 0$, appealing to the result that Ito integrals are martingales. However, remember that without restrictions on the integrand, Ito integrals are only local martingales. To be a martingale, we must have $E\int \sigma^2(x)V_x^2 ds < \infty$.

- Since $V_w \sim |W| \sim \theta$ and $\sigma^2(x) \sim$ constant on the transformed time scale, this martingale restriction is equivalent to $E\left[\int_0^1 \theta_s^2 ds\right] < \infty$, which the doubling strategy violates.
- Define $\mathcal{H}^2 = \{\theta : E\int \theta_s^2 ds < \infty\}$. Note, $\mathcal{H}^2 \subset \mathcal{L}^2$, since $\mathcal{L}^2$ permits infinite outcomes with vanishingly small probability. These infinitely large, measure zero portfolios are essential to the success of the doubling strategy. Hence, most applications of continuous-time finance constrain admissible control policies to be in $\mathcal{H}^2$. Even then one must engage in an ex post verification that the induced state process satisfies the martingale requirement $E\int \sigma^2(x_s)V_x^2 ds < \infty$ for all admissible controls.
- Assuming this martingale restriction applies, the logic of the verification theorem is as follows. Assume that V is sufficiently smooth to apply Ito's lemma, and that V satisfies the boundary condition

$$f_T(X_T) = V_T(X_T) = V_0(X_0) + \int_0^T dV(X_s)$$

Using Ito's lemma and the state transition eq. $dX = \mu(X,u)dt + \sigma(X)dB$

$$f_T(X_T) = V_T(X_T) = V_0(X_0) + \int_0^T [V_t + \mu(X,u)V_x + \frac{1}{2}\sigma^2(x)V_{xx}]dt + \int_0^T \sigma(X)V_x dB$$

Taking expectations of both sides, adding $\int_0^T e^{-\rho t} f(x,u)dt$ to both sides, and then imposing the martingale assumption to drop the last integral, we get

$$E\left[\int_0^T e^{-\rho t} f(X,u)dt + f(X_T)\right] = V_0(X_0) + E\left[\int_0^T \left(e^{-\rho t} f(X,u) + V_t + \mathcal{D}[V]\right) dt\right]$$

where $\mathcal{D}[V] = \mu(X,u)V_x + \frac{1}{2}\sigma^2(x)V_{xx}$ is the Dynkin operator. According to the HJB eq., the integrand on the r.h.s achieves a maximum at zero. For nonoptimal controls it is negative. Hence,

$$E\left[\int_0^T e^{-\rho t} f(X,u)dt + f(X_T)\right] \le V_0(X_0)$$

with equality for controls satisfying the HJB equation.

Viscosity Solutions

- Notice that we assumed V was sufficiently smooth to apply Ito's lemma. How are we supposed to know this unless we know what V is (which begs the question)?
- Leading sources of potential nondifferentiability in economics are: (1) binding state/control constraints, and (2) nonconvexities.
- In practice, Inada conditions are typically assumed so that constraints never bind along the optimal path.
- If nondifferentiability is a potential issue, one can appeal to a weaker solution concept of the HJB eq., based on subgradients, known as a viscosity solution.
- Ben Moll has a useful exposition aimed at economists posted on his website [*Viscosity Solutions for Dummies (Including Economists)*].