

FINANCIAL OPTIMIZATION

Lecture 5: Dynamic Programming and a Visit to the Soft Side

Dynamic Programming

All situations in practice are more complex than the simple examples we been analyzing. One important feature in almost every practical situation is the time element complicated by the arrival of information. Typically, this makes the choice problem very complicated, and we must do something to make the problem simple enough to solve. Some possibilities:

- Use a restricted list of choice variables
- Limit the amount of information used
- Assume a limited form for strategies (e.g. linear)
- Use functional forms that simplify analysis
- Use mathematics to simplify the problem
 - one-shot valuation
 - maximum principle
 - solve for some choice analytically, leaving fewer variables

In-class Exercise: Investments Problem with Consumption Withdrawal

Suppose we take a very simple investments setting of investing for consumption over the next five years. A relative of yours has accepted a retirement package from his firm and is asking your help in planning consumption and investments over the next five years, under the assumption that this is a good time to retire. Five years from now, your relative will start to receive retirement income much larger than what is available now. Assuming that borrowing against the retirement money is not possible, that can be viewed as the decision horizon.

Without writing down any formulas, discuss what might be included in the choice problem? Address the following:

- choice variables
- objective function
- constraints

Investments Problem: Simplification

- look at overall consumption and looked at two categories (fixed and variable) rather than details of consumption
- restrict asset classes, possible targets
- limit the type of information we are using
- abstract from taxes or choose outside the optimization the form of investment accounts (401K, Keogh, etc.) and optimize conditional on that
- assume preferences (e.g. quadratic or log utility) known to simplify the analysis if that seems consistent with your relative's preferences
- discretize time, perhaps annual or monthly. or use a fiction of continuous expenditure and solve a continuous-time model
- leave out or simplify tax considerations and transaction costs
- constant interest rate or other returns

Investments Problem: Extremely Simple Formulation

Given w_t at time t ,

choose adapted risky holdings h_s , consumption c_s , and wealth plan w_s to

maximize $E[\sum_{s=t}^N e^{-\rho(s-t)} u(c_s)]$, subject to

$w_{s+1} = (w_s - c_s)(1 + r) + h_s(R_{s+1} - r)$, and

$(\forall s)(w_s \geq 0)$.

What is the problem over the next period? We can substitute the value function $V(w_t, t)$ for this problem for the actual continuation, and we have:

$$V(w, N) = u(w)$$

(assumes u is increasing, so it is optimal to consume the remaining wealth in the last period), and for $s < N$

$$V(w, s) = \max_{c, h} u(c) + E_s[e^{-\rho} V((w - c)(1 + r) + (R_{s+1} - r)h, s + 1)]$$

Solving for the function V can be a lot easier than trying to think about the original optimization directly.

Solution with Quadratic Utility: Solve Back from the End

Take $u(w) = w - aw^2/2$ and assume $E[R_s] = \mu$ (a constant), $\text{var}(R_s) = \sigma^2$ (also constant), and $\rho = 0$ (not impatient). Also, take $r = 0$ to keep the algebra under control. We have

$$V(w, N) = w - \frac{a}{2}w^2$$

and

$$\begin{aligned} V(w, N - 1) &= \max_{c,h} c - \frac{a}{2}c^2 + E[V(w - c + R_N h, N)] \\ &= \max_{c,h} c - \frac{a}{2}c^2 + E[w - c + R_N h - \frac{a}{2}(w - c + R_N h)^2] \\ &= \max_{c,h} c - \frac{a}{2}c^2 + w - c + \mu h \\ &\quad - \frac{a}{2}(w^2 + c^2 + (\mu^2 + \sigma^2)h^2 - 2wc + 2w\mu h - 2c\mu h) \end{aligned}$$

First-order conditions:

$$\begin{aligned} 1 - ac - 1 + aw - ac + ah\mu &= 0 & c &= \frac{a\sigma^2 w + \mu^2}{a(2\sigma^2 + \mu^2)} \\ \mu - a(w - c)\mu - a(\sigma^2 + \mu^2)h &= 0 & h &= \frac{(2 - aw)\mu}{a(2\sigma^2 + \mu^2)} \end{aligned}$$

Solution

Solution with quadratic Utility: Earlier Periods

Note that both c and h are affine in w . Therefore, it is easy to see that $V(w, N - 1)$ is quadratic in w (although I am not eager to do the algebra to compute the coefficients). As a result, the step from period $N - 1$ to $N - 2$ will be similar to the step from period N to period $N - 1$. And, while the algebra may not be fun, it is possible to write down a general formula relating the quadratic coefficients and optimal strategy in terms of the optimal coefficients the following period. This represents a complete solution of the optimization problem using the maximum principle.

State Variables in Dynamic Programming

Central to the maximum principle approach to dynamic programming is the value function. The thing that helps to make the value function simple and tractable is that it is a function of what is important looking forward, not of everything known so far. In our portfolio choice example, the value function depends on two state variables, time to maturity and wealth. One could imagine a value function that also depends on past wealth, portfolio composition at the beginning of the period, past portfolio choices, and past stock returns. However, given current wealth and time-to-maturity these things do not affect the payoffs or constraints looking forward, and therefore the expected value of the objective function depends only on these two state variables. The reason for this comes from assumptions, implicit and explicit, in the optimization. For example, it is assumed that no information arrives that tells about mean and variance of stock returns: μ and σ are parameters (fixed constants) that are known from the beginning and do not need to be included explicitly as arguments to the value function. Another assumption is that preferences are additive over time, so preferences about future consumption do not depend on past consumption.

State Variables: Investments Examples for Discussion

- Frictionless investment
- Transaction costs
- Taxes
- Predictable returns

Maximum Principle: Martingale Perspective

Recall the general version of the investment problem given earlier:

Given w_t at time t ,

choose adapted risky holdings h_s , consumption c_s , and wealth plan w_s to

maximize $E[\sum_{s=t}^N e^{-\rho(s-t)} u(c_s)]$, subject to

$w_{s+1} = (w_s - c_s)(1 + r) + h_s(R_{s+1} - r)$, and

$(\forall s)(w_s \geq 0)$.

Consider the following process, where V is the correct objective function:

$$\mathcal{M}_t = \left(\sum_{s=0}^t e^{-\rho s} u(c_s) \right) + e^{-\rho(t+1)} V(w_{t+1}, t + 1)$$

and the following Bellman equation:

$$V(w, t) = \max_{c, h} u(c) + e^{-\rho} E[V((w - c)(1 + r) + h(R - r), t + 1)]$$

Connection between the Bellman equation and the Martingale \mathcal{M}

Given the optimal strategy, \mathcal{M} is the conditional expectation of the objective function given what is known at time t . At time 0, this is therefore the value of the problem, and over time, because this is a conditional expectation, it is a martingale (by the law of iterated expectations).

On the other hand, if we have the right value function, but the wrong strategy, \mathcal{M} starts at the optimal value but declines on average over time as mistakes are made. In other words, choosing the optimal strategy maximizes the expected change in \mathcal{M} at zero. This is the source of the Bellman equation (sometimes with other names added, e.g. Hamilton-Jacobi-Bellman equation).

For teaching about the Bellman equation and verification theorems (proofs that the solution to the first-order condition is a solution to the choice problem) in continuous time, some PhD programs use my paper "Duesenberry's Ratcheting of Consumption: Optimal Dynamic Consumption and Investment Given Intolerance for any Decline in Standard of Living," *Review of Economic Studies* 62, 1995, 287–313.

Continuous-time Bellman Equation

An investments example:

Given w_t at time t ,
choose adapted risky holdings h_s , consumption c_s , and wealth plan w_s to
maximize $E[\int_{s=t}^T e^{-\rho s} u(c_s) ds]$, subject to
 $dw_{s+1} = (w_s r - c_s) dt + h_{s-1}((\mu - r) dt + \sigma dZ_t)$. ($\forall s$)($w_s \geq 0$).

$$\mathcal{M}_t = \int_{s=0}^t e^{-\rho s} u(c_s) ds + e^{-\rho t} V(w, t)$$

Bellman equation:

$$E \max_{c,h} [d\mathcal{M}_t] = 0.$$

Using Itô's lemma, can be computed as

$$\max_{c,h} u(c) - \rho V + V_t + (wr - c + h(\mu - r))V_w + \frac{1}{2}h^2\sigma^2V_{ww} = 0$$

Option Pricing Application

One of the most successful application of dynamic programming in practice is to option pricing, which is complicated enough not to be obvious and simple enough (under reasonable assumptions) to analyze. For example, perpetual put (in risk-neutral probabilities):

Given S_0 , choose a stopping time τ to maximize $E[(X - S_\tau)exp(-r\tau)]$, subject to $\frac{dS_t}{S_t} = rdt + \sigma dZ_t$

A particular American lookback call:

Given S_0 , choose a stopping time $\tau \leq T$ to maximize $E[(\max\{S_t|0 \leq t \leq \tau\} - X)^+e^{-r\tau}]$, subject to $\frac{dS_t}{S_t} = rdt + \sigma dZ_t$

Reload options, used in executive compensation, see my paper

Dybvig, Philip H. and Mark Loewenstein, 2003, Employee Reload Options: Pricing Hedging, and Optimal Exercise, *Review of Financial Studies* 16, 145–171.

Option Pricing Applications: State Variables

American put: stock price, time (if not perpetual), whether exercised

Lookback call on previous slide: stock price, time, max price so far

Option on average: stock price, time, average price so far

the lookback call described earlier (all in r-n probs):

Given S_0 , choose an exercise indicator $x_t \in \{0, 1\}$ to maximize $E[\sum_{t=0}^T x_t (\max\{S_s | 0 \leq s \leq t\} - X) \frac{1}{R^t}]$, subject to

$$S_{t+1} = \begin{cases} US_t & \text{with probability } \pi_u^* \\ DS_t & \text{with probability } \pi_d^* \end{cases}$$

and $\sum_{t=0}^T x_t \leq 1$

State variables: S_t , t , $\sum_{s=0}^{t-1} x_s$, and $\max\{S_s | 0 \leq s \leq t\}$ (running maximum stock price)

where π_u^* and π_d^* come from the binomial model.

In-class exercise: lookback call

Write down the Bellman equation for the lookback call.

Which is probably faster numerically, using the running maximum price or the whole price history as a state variable?

A Visit to the Soft Side

The optimization paradigm we have been studying is probably just as useful for organizing thinking as for explicitly computing what to do.

- models
- empirics
- intuition

As hard-core quants, we would like to think most of the action is in the models, but intuition is maybe most important and empirics (including institutions) are also very important.

In-class Exercise: A Visit to the Soft Side

You are working in a corporate treasury office and your firm is planning to raise some money. Being a good scientific problem-solver, you want to think about this using decision theory. What are the choice variables, objective function, and constraints?