ECON 260A: Stochastic Resource Economics

Professor Christopher Costello

November 6, 2018

1 Stochastic Dynamic Programming (discrete-time)

Consider the general dynamic optimization problem:

$$\max_{\{u_t\}} \sum_{t=0}^{T} \delta^t \pi(x_t, u_t) \tag{1}$$
$$\text{s.t. } x_{t+1} = G(x_t, u_t) \tag{2}$$

Here δ is the discount factor, u_t is the control variable (or vector) in period t, x_t is the state variable (or vector), $\pi(x, u)$ is the current period payoff conditional on the current state and control, and the function $G(x, u)$ defines the equation of motion (Equation 4). The time horizon is finite T, but could also be infinite.

We can include a random shock to the payoff function, the equation of motion, or both:

$$\max_{\{u_t\}} \sum_{t=0}^{T} \delta^t E\pi(x_t, u_t, \epsilon_t) \tag{3}$$
$$\text{s.t. } x_{t+1} = G(x_t, u_t, \theta_t) \tag{4}$$

where the distributions of the shocks are assumed to be known. Notice that the objective is now to maximize the expected value of utility over the time horizon.

Now that we have added stochasticity, it helps to be clear about what the analyst knows, and when she knows it. What she knows should be included in the state vector, if it is pertinent to the solution. Solving these things numerically is, in principle, no more difficult than solving the deterministic version. The practical challenge is that the state space now expands to account for the possible future realizations of the random variable.

A more specific natural resource problem is: A resource stock x_t is harvested to maximize profit. Harvest in period t is h_t and the current period profit from harvest is $\pi(x,h)$. The discount factor is δ and the equation of motion (i.e. the growth of the resource stock) is:

$$x_{t+1} = z_t f(x_t - h_t) \tag{5}$$

where z_t is an i.i.d. random variable with mean 1. In this setup, z_{t-1} is known before h_t is chosen, but z_t is unknown when h_t is chosen. Why don't we include z_{t-1} as a state variable? The dynamic programming equation is:

$$V_t(x_t) = \max_{h_t} \pi(x_t, h_t) + \delta E_t V_{t+1}(z_t f(x_t - h_t)) \tag{6}$$

Under some special circumstances, this can be solved analytically. Under lots of circumstances, it cannot, but certain properties can be derived. But problems of this nature can always be solved numerically.

Some interesting applications of this model are:

- Reed solved this model analytically for certain forms of the profit function
- Clark and Kirkwood solved a similar model, but where the previous period shock (and therefore the current resource stock) were unknown prior to the harvest decision.
- Weitzman looked at a similar problem to CK, but compared taxes vs. quantities. He found that taxes were always superior.
- McGough et al. incorporated a general form for $\pi(x,h)$ and solved the problem using log-linearization techniques used in macroeconomics.
- Costello and Polasky incorporated space into the stochastic problem
- Sethi et al. examined the implications of including several forms of uncertainty into the same model

1.1 Auto-correlated shocks

Suppose now that $z_t = \alpha z_{t-1} + \varepsilon_t$, for $0 \leq \alpha \leq 1$ and $\varepsilon_t \sim iid(0)$. How does this affect the problem? Now, because z_{t-1} is known prior to the decision of h_t, we can use z_{t-1} as a state variable (why do so?). Now, the DPE is:

$$V_t(x_t, z_{t-1}) = \max_{h_t} \Pi(x_t, h_t) + \delta E_t V_{t+1}((\alpha z_{t-1} + \varepsilon_t) f(x_t - h_t)) \tag{7}$$

If the error process was AR(2), what would be the state variables?

2 Regime shift or Catastrophe

Consider the traditional Faustmann model of forest rotation, but where the forest faces a risk of fire (which would completely decimate the forest). Let λdt be the probability that a fire occurs within the interval $t+dt$; λ is called a "hazard rate" (here it is constant, but other models could allow it to be endogenous). Let τ_i be the time between the $i-1$st and ith clearing of the forest (which could occur via fire or via logging). We seek the optimal age of the forest T at which to harvest the trees. The cdf of τ_i is given as follows:

$$Pr(\tau \leq t) = \begin{cases} 1 - e^{-\lambda t} \quad \text{if} \quad t < T \\ 1 \quad \text{if} \quad t \geq T \end{cases} \tag{8}$$

If the fire occurs prior to logging, the forester must pay a cost (replanting plus other costs associated with fire cleanup), c_2. If the fire does not occur, she earns $V(T)$ and must pay a replanting cost c_1. which implies that the ith economic return is:

$$Y_i = \begin{cases} -c_2 \quad \text{if} \quad \tau_i < T \\ V(T) - c_1 \quad \text{if} \quad \tau_i = T \end{cases} \tag{9}$$

Proceeding in this manner, it can be shown that the optimal rotation interval satisfies:

$$\frac{V'(T)}{V(T)-c_1} = \frac{\delta+\lambda}{1-e^{-(\delta+\lambda)T}} \tag{10}$$

(see Conrad and Clark for a proof). This implies that the forest fire risk is mathematically identical to an increase in the discount rate, δ.[1]

Another prominent type of regime shift problem involves a decision maker (or non-cooperative agents) who emits a stock pollutant that causes environmental damage. The pollution stock can trigger a regime shift (that either causes more severe damage (carbon shuts down thermohaline circulation) or changes the benefits of polluting (air pollutants reduce demand)). For example, see the current homework challenge.

3 LQ problems with (and without) uncertainty

A special class of dynamic optimization problems is the Linear Quadratic Control class. Any dynamic optimization problem with the following features fits into this class:

- The period-t payoff is:
  - (at most) Quadratic in the state vector
  - (at most) Quadratic in the control vector

[1]Recall the deterministic Faustmann problem. There, with replanting cost c, the optimal rotation age T is given implicitly by: $\frac{V'(T)}{V(T)-c} = \frac{\delta}{1-e^{-\delta T}}$.

- (at most) Bilinear in the state and control

- The Equation of Motion is:
  - (at most) Linear in the state vector
  - (at most) Linear in the control vector

3.1 One state, one control

[Note: In these notes, I will jump directly to the case in which the parameters are time-dependent. The simpler case is where the parameters are constant over time]. Consider the one state (x) and one control (u) version of this problem. The payoff function is:

$$\pi_t(x,u) = \frac{1}{2}q_t x^2 + r_t ux + \frac{1}{2}s_t u^2 \tag{11}$$

The equation of motion is:

$$x_{t+1} = a_t x_t + b_t u_t \tag{12}$$

The Dynamic Programming Equation (Bellman Equation) is:

$$V_t(x_t) = \max_{u_t} \frac{1}{2}q_t x_t^2 + r_t u_t x_t + \frac{1}{2}s_t u_t^2 + \beta V_{t+1}(a_t x_t + b_t u_t) \tag{13}$$

As usual, the problem is that the form of the function V is unknown.

The magic of LQ problems is that, in fact, the form of V is always the same (up to a constant). In particular, it can be shown that any problem with the above specification has the following properties:

- The Value Function is a quadratic function of x and
- The Policy Function is linear in x

In particular, we will "guess" the form of the value function, and then confirm that it solves our fixed point problem above. Guess that:

$$V_t(x_t) = \frac{1}{2}w_t x_t^2 \tag{14}$$

for some constant w_t. This gives a DPE of:

$$\frac{1}{2}w_t x_t^2 = \max_{u_t} \frac{1}{2}q_t x_t^2 + r_t u_t x_t + \frac{1}{2}s_t u_t^2 + \frac{1}{2}\beta w_{t+1}(a_t x_t + b_t u_t)^2 \tag{15}$$

Taking the first order conditions, gives:

$$0 = r_t x_t + s_t u_t + \beta w_{t+1} b_t (a_t x_t + b_t u_t) \tag{16}$$

Solving for u_t gives:

$$u_t = \underbrace{-\frac{r + \beta w_{t+1} b_t a_t}{s_t + \beta w_{t+1} b_t^2}}_{U_t} x_t = U_t x_t \tag{17}$$

Plugging back into the DPE, and noticing that U_t contains w_{t+1} gives the "Ricatti Equation":

$$w_t = q_t + \beta a_t^2 w_{t+1} + (\beta a_t w_{t+1} b_t + r_t) U_t \tag{18}$$

If we knew w_t for all t then we'd be done: this would give us both the value function and the optimal control rule (i.e. the policy function). If time is finite T, the we can simply iterate backwards, starting with w_{T+1}, which is closely related to the terminal value (the terminal value is: $\frac{1}{2} w_{T+1} x^2$...e.g. if there is no salvage value, then $w_{T+1} = 0$.). If time is infinite, we just iterate until convergence of w.

3.2 Generalizing L-Q control to n states & k controls

Following Judd's notation, let the nx1 state vector be given by x, the kx1 control vector be given by u, and the discount factor be given by β. The period t payoff function is:

$$\pi_t(x, u) = \frac{1}{2} x' Q_t x + u' R_t x + \frac{1}{2} u' S_t u \tag{19}$$

where Q_t is nxn, R_t is kxn, and S_t is kxk. The Equation of Motion is:

$$x_{t+1} = A_t x_t + B_t u_t \tag{20}$$

where A_t is nxn and B_t is nxk.

The Dynamic Programming Equation (Bellman Equation) is:

$$V_t(x) = \max_{u_t} \frac{1}{2} x' Q_t x + u_t' R_t x + \frac{1}{2} u_t' S_t u_t + \beta V_{t+1}(A_t x + B_t u_t) \tag{21}$$

Guess that:

$$V_t(x) = \frac{1}{2} x' W_t x \tag{22}$$

This produces the "Matrix Ricatti Equation":

$$W_t = Q_t + \beta A_t' W_{t+1} A_t + (\beta A_t' W_{t+1} B_t + R_t') U_t \tag{23}$$

which can be solved via backward induction (value function iteration) or via policy function iteration. In infinite time, the subscript t's drop out and we obtain the "Algebraic Matrix Ricatti Equation":

$$W = Q + \beta A' W A - (\beta A' W B + R')(S + \beta B' W B)^{-1}(\beta B' W A + R) \tag{24}$$

which itself can be solved by iterating over W (see Judd for details).

3.3 Exploring the generality of LQ control

By cleverly defining the state and control vectors in your problem, you can often transform a “complicated” problem into a “simple” LQ control problem. The main advantage of LQ control is that you can solve any (reasonably) sized problem. Using value function iteration on large problems is not possible. Here are a couple of examples of problems from the literature that use LQ control

- Karp and Hoel (among others) have adapted Weitzmans classic prices vs. quantities problem to a dynamic setting. One can define the stock of pollution (e.g. CO2 concentration in the atmosphere) as the state, and the pollution level as the control (under a quota policy). Then, if environmental damage is quadratic in the stock, and abatement cost is quadratic in abatement, you can write the payoff as quadratic in the state and control, as required. The state transition equation is simply: $S_{t+1} = \delta S_t + x_t$, where $1-\delta$ measures the per-period “decay" of the stock and x_t measures the emissions. This now fits within our requirements for an LQ control problem.
- Blackwood et al. address a problem of optimally controlling an invasive species across space. In their problem, the species spreads in a known way across space, and patches are connected via dispersal of the species. The species imposes environmental costs and costs are incurred in its control. The vector N_t gives the stock of species in each patch (the state vector). Removal of species in each patch is given by H_t (the control vector). They argue that (1) invasive species populations grow approximately exponentially, so the state equation is linear in the state and control, and (2) environmental costs of the species are quadratic and the removal costs are quadratic. So the requirements for an LQ problem are fulfilled.

3.4 LQ with Additive Error

It turns out that we can further generalize the LQ problem by adding an iid random disturbance. In particular, suppose we wish to include an additive, mean zero shock to the equation of motion, as follows:

$$x_{t+1} = A_t x_t + B_t u_t + \varepsilon_t \tag{25}$$

Here the random variable $\varepsilon_t \sim iid(0)$. It can be shown that the optimal policy function that results from this problem, $u^*(x)$, is independent of the higher order properties of the stochastic process, ϵ_t. This implies that from a policy perspective, one can proceed as if there was no uncertainty and use the Ricatti Equation above to calculate the optimal policy function. However, higher order properties of the distribution will affect the value function itself.

4 Appendix: Time-constant parameters

4.1 One state, one control

Consider the one state (x) and one control (u) version of this problem. The payoff function is:

$$\pi_t(x,u) = \frac{1}{2}qx^2 + rux + \frac{1}{2}su^2 \tag{26}$$

The equation of motion is:

$$x_{t+1} = ax_t + bu_t \tag{27}$$

The Dynamic Programming Equation (Bellman Equation) is:

$$V_t(x_t) = \max_{u_t} \frac{1}{2}qx_t^2 + ru_tx_t + \frac{1}{2}su_t^2 + \beta V_{t+1}(ax_t + bu_t) \tag{28}$$

As usual, the problem is that the form of the function V is unknown.

The magic of LQ problems is that, in fact, the form of V is always the same (up to a constant). In particular, it can be shown that any problem with the above specification has the following properties:

- The Value Function is a quadratic function of x and
- The Policy Function is linear in x

In particular, we will “guess" the form of the value function, and then confirm that it solves our fixed point problem above. Guess that:

$$V_t(x_t) = \frac{1}{2}w_tx_t^2 \tag{29}$$

for some constant w_t. This gives a DPE of:

$$\frac{1}{2}w_tx_t^2 = \max_{u_t} \frac{1}{2}qx_t^2 + ru_tx_t + \frac{1}{2}su_t^2 + \frac{1}{2}\beta w_{t+1}(ax_t + bu_t)^2 \tag{30}$$

Taking the first order conditions, gives:

$$0 = rx_t + su_t + \beta w_{t+1}b(ax_t + bu_t) \tag{31}$$

Solving for u_t gives:

$$u_t = \underbrace{-\frac{r + \beta w_{t+1}ba}{s + \beta w_{t+1}b^2}}_{U_t} x_t = U_tx_t \tag{32}$$

Plugging back into the DPE, and noticing that U_t contains w_{t+1} gives the “Ricatti Equation":

$$w_t = q + \beta a^2 w_{t+1} + (\beta a w_{t+1}b + r)U_t \tag{33}$$

If we knew w_t for all t then we'd be done: this would give us both the value function and the optimal control rule (i.e. the policy function). If time is finite T, the we can simply iterate backwards, starting with w_{T+1}, which is closely related to the terminal value (the terminal value is: $\frac{1}{2}w_{T+1}x^2$...e.g. if there is no salvage value, then $w_{T+1} = 0$.). If time is infinite, we just iterate until convergence of w.