U.T. Economics Summer 2013 Math Camp

Date: Friday, August 23 and Tuesday, August 27, 2013

Topics: Discrete time dynamic optimization, search theory based on value function iterations, the policy improvement algorithm, and the uniqueness of the solution to the Bellman/Pontryagin equation.
Readings: CSZ 4.11

1. Discrete Time, Finite Horizon, Deterministic

The main point is to keep track of the value of where you end up, and to use that to break a complicated dynamic problem into bite-sized pieces.

A. Do CSZ problem 4.11.12, being sure to keep track of the value of being in the various states.

2. Model: Random Job Offers

Now we move to an infinite horizon problem and add stochastics. To make up for having made our life more complicated, we make the problem stationary.

At times $t = 0, 1, 2, \ldots$, the DM (decision maker) receives a random wage offer W_t. the collection $\{W_t : t \geq 0\}$ is iid. If the DM accepts the offer $W_t = w$, they are out of the labor market and continue to receive w until the crack of doom. If the DM rejects the offer, they receive c for this period, and will have another chance in the next period. Utility in each period is given by consumption and consumption is either c or the accept w. [We can add a savings problem on top of this, and you will need to learn how to do that at some point in the future.] The utility of a consumption stream x_t, $t = 0, 1, 2, \ldots$ is $U(x_0, x_1, x_2, \ldots) = \sum_{t\geq 0} \beta^t u(x_t)$, $0 < \beta < 1$. Given a present offer of w, i.e. $W_0 = w$, following the optimal policy leads to the value function $V(w)$. To be a bit more specific:

a. accepting the offer $W_0 = w$ at $t = 0$ leads to utility $\sum_{t\geq 0} \beta^t u(w) = u(w)\frac{1}{1-\beta}$;
b. rejecting the first offer and accepting the offer $W_1 = w$ at $t = 1$ leads to utility $u(c) + \sum_{t\geq 1} \beta^t u(w) = u(c) + u(w)\frac{\beta}{1-\beta}$;
c. etc.

3. Value function analysis

If we suppose that starting tomorrow, at $t = 1$, the DM will start behaving optimally, then today the problem they face today, holding the offer $W_0 = w$, is

$$\max_{rej,\ acc}\{\underbrace{u(c) + \beta E\, V(W_1)}_{rej}, \underbrace{u(w)(1-\beta)^{-1}}_{acc}\}, \tag{1}$$

where the first term is the utility of rejecting the present offer, and the second term is the utility of accepting it. Behaving optimally today plus behaving optimally thereafter is behaving optimally at all points in time, hence the **Bellman equation** (or **Pontryagin equation**) for the **value function**,

$$V(w) = \max\{u(c) + \beta E\, V(W_1), u(w)\tfrac{1}{1-\beta}\}. \tag{2}$$

Note that this is an equation in the function $V(\cdot)$ that should be satisfied for each of infinitely many possible values of w. Such things are called **functional equations**. We will (later) be able to show that the solution to this functional equation is unique.

A **policy function** is a function from the present state, w, to the optimal action, reject or accept. Given the true value function, the *optimal* policy function rejects at w when $u(c) + \beta E\, V(W_1) \leq (<) u(w)\frac{1}{1-\beta}$, and accepts otherwise. From this, one can directly read off the dependence of the optimal cutoff wage on c.

4. Iterating on the value function

In this simple case, we might not proceed in the iterative fashion about to be described because there might be more direct methods, but we can see how it works out. Suppose that we are looking at the problem at $t = 0$ and we know that no utility will be delivered at $t \geq 1$ (think "Eat, drink, and be merry, for tomorrow we die"). Then the value of where we end up is $V_0(w_1) \equiv 0$. The value function for this one period problem is

$$V_1(w) = \max\{\underbrace{u(c) + \beta E\, V_0(W_1)}_{rej}, \underbrace{u(w) + \beta E\, V_0(w)}_{acc}\}. \tag{3}$$

Thus, $V_1(w) = \max\{u(c), u(w)\}$. The associated policy function is $a_0(w)$ =reject if $w < c$, accept else.

Now back up one period from the end of the world, suppose that we will receive utility in periods $t = 0$ and $t = 1$ but no utility for $t \geq 2$, and that at $t = 1$ we will follow the optimal policy for being in a situtation where there will be no future utility,

$$V_2(w) = \max\{\underbrace{u(c) + \beta E\, V_1(W_1)}_{rej}, \underbrace{(1+\beta)u(w)}_{acc}\}, \tag{4}$$

and the policy function, $a_1(w)$ is the argmax for this problem. The next equation is

$$V_3(w) = \max\{\underbrace{u(c) + \beta E\, V_2(W_1)}_{rej}, \underbrace{(1+\beta+\beta^2)u(w)}_{acc}\}, \tag{5}$$

and we just keep on going. The claim is that the limit function, $V(w) = \lim_n V_n(w)$ is the unique function satisfying eqn. (2).

5. Policy function iterations

Again, in this simple case, we might not proceed in the iterative fashion about to be described in favor of more direct methods, but we can see how it works out. Propose a policy, say $a_0(w)$ =reject if $w < w_\circ$, accept else. There is some expected value to following this policy. Find it, and call it $V_0(\cdot)$. Now consider the problem in which we commit to using $a_0(\cdot)$ for $t \geq 2$, but are considering doing the best we can subject to this limitation. The problem is

$$V_1(w) = \max\{u(c) + \beta E\, V_0(W_1), u(w)\tfrac{1}{1-\beta}\}. \tag{6}$$

Call the argmax from the above policy $a_1(\cdot)$, there is some expected value to following this policy, find it, call it V_1 and the new problem is

$$V_2(w) = \max\{u(c) + \beta E\, V_1(W_1), u(w)\tfrac{1}{1-\beta}\}. \tag{7}$$

This delivers a sequence of policy functions, each one of which is better than the previous one (a pretty easy argument).

6. Direct calculation methods

Suppose that we are consider a policy reject if $W \leq w_0$, accept otherwise. Let T be the random time until the first $W_t > w$. If $T = \tau$, the sequence of realized W_t's satisfies $W_0, W_1, \ldots, W_{\tau-1} \leq w_0$, $W_\tau > w_0$, and the associated utility is

$$\sum_{t=0}^{\tau-1} \beta^t u(c) + \sum_{t\geq\tau} \beta^t u(W_\tau). \tag{8}$$

The probability that $\tau = 0$ is $(1 - F(w_0))$, $P(\tau = 0) = F(w_0)^0(1 - F(w_0))$, $P(\tau = 1) = F(w_0)^1(1 - F(w_0))$, $P(\tau = 2) = F(w_0)^2(1 - F(w_0))$, etc. In other words, τ has a geometric distribution. From these, we can calculate the utility associated with following this policy from $t = 0$ onwards,

$$E\,U(w_0) = E\left[\sum_{t=0}^{\tau-1} \beta^t u(c) + \sum_{t\geq\tau} \beta^t u(W_\tau)\right] = \frac{1}{1-\beta F(w_0)}\left[u(c)\beta F(w_0) + E\left(u(W)|W > w_0\right)\frac{1-F(w_0)}{1-\beta}\right], \tag{9}$$

where W is a random variable independent of everything in sight and having the same distribution as all of the W_t. The resulting direct calculation problem is $\max_{w_0} E\,U(w_0)$.

7. Problems

B. From the observation that there is a unique solution to $V(w) = \max\{u(c) + \beta E\, V(W_1), u(w)\frac{1}{1-\beta}\}$ and that the optimal policy function solves it, give the dependence of the optimal cutoff on c.

C. [Discrete time discounting] Suppose that the net benefits for a project are $B_0, B_1, \ldots$ where: net benefits per period are bounded, i.e. there exists a (perhaps quite large) number $\overline{B}$ such that for all t, $|B_t| \leq \overline{B}$; there exists a time period T such that $B_t < 0$ for $t = 1, \ldots, T$ and $B_t > 0$ for $t = T+1, T+2, \ldots$; and for a sufficiently large T', $\sum_{t=0}^{T'} B_t > 0$, i.e. that if we do not discount, then the net benefits outweight the net costs.

1. Show that for ρ sufficiently close to but still strictly less than 1, $\sum_{t=0}^{\infty} \rho^t B_t > 0$.

One interpretation of discounting has to do with the ability to invest resources at t and receive a larger set of resources at $t+1$ or later. If this is true, receiving an amount x at t is worth less than receiving the same x at $t+1$ or later. This has to do with $\rho = \frac{1}{1+r}$ where r is the interest rate. Another interpretation of discounting has to do with the uncertainty of future benefits.

Let τ be a random variable taking on the values $0, 1, \ldots$, and having the **geometric distribution**, i.e. $P(\tau = t) = \delta^t(1-\delta)$.

2. Give the expectation of τ, $E\,\tau$, as a function of δ.
3. Show that τ is "memoryless," i.e. that $P(\tau = t+s|\tau \geq t) = P(\tau = s)$ for $s, t \geq 0$.
4. Give $E\sum_{t=0}^{\tau} B_t$. ($\sum_{t=0}^{\tau} B_t$ is called, for pretty straightforward reasons, a "random sum." Random sums turn out to be very important in many contexts beyond discounting.)

D. Let $f(c, w_0) = \frac{1}{1-\beta F(w_0)}\left[u(c)\beta F(w_0) + E\left(u(W)|W > w_0\right)\frac{1-F(w_0)}{1-\beta}\right]$ and consider the problem $\max_{w_0} f(c, w_0)$. How does the optimal w_0 depend on c? [We've seen this result before.]

E. For each of the following cases, find $E\left(W|W > w_\circ\right)$ and $E\left(u(W)|W > w_\circ\right)$ as a function of $w_\circ$.

1. $W \sim$ Unif$[80, 220]$, $u(w) = w + 2\sqrt{w}$.
2. $W \sim$ Unif$[80, 220]$, $u(w) = \frac{w^{1-a}}{1-a}$, $0 < a < 1$.
3. $W \sim$ Unif$[80, 220]$, $u(w) = 1 - e^{-\gamma w}$, $\gamma > 0$.
4. $W \sim$ neg. exp.(λ), $u(w) = 1 - e^{-\gamma w}$, $\lambda, \gamma > 0$. [An especially interesting case has $\gamma = \lambda$.]

F. Set $c = 90$ and $\beta = 0.95$. For each of the following cases, perform the first two iterations on the policy function.

1. $W \sim$ Unif$[80, 220]$, $u(w) = w + 2\sqrt{w}$.
2. $W \sim$ Unif$[80, 220]$, $u(w) = \frac{w^{1-a}}{1-a}$, $0 < a < 1$.
3. $W \sim$ Unif$[80, 220]$, $u(w) = 1 - e^{-\gamma w}$, $\gamma > 0$.
4. $W \sim$ neg. exp.$(\frac{1}{150})$, $u(w) = 1 - e^{-\gamma w}$, $\gamma > 0$.

G. Set $c = 90$ and $\beta = 0.95$. For each of the following cases, find the optimal job search policy if it is possible to give it analytically, otherwise specify how you would find it numerically.

1. $W \sim$ Unif$[80, 220]$, $u(w) = w + \sqrt{w}$.
2. $W \sim$ Unif$[80, 220]$, $u(w) = \frac{w^{1-a}}{1-a}$, $0 < a < 1$.
3. $W \sim$ Unif$[80, 220]$, $u(w) = 1 - e^{-\gamma w}$, $\gamma > 0$.
4. $W \sim$ neg. exp.$(\frac{1}{150})$, $u(w) = 1 - e^{-\gamma w}$, $\gamma > 0$.

H. [An extended example due to Eyal Winter] In principal-agent models, there is a principal (think boss) who offers rewards that are contingent on outcomes (assumed observable), but who cannot observe the effort (assumed un-observable, or at least assumeb to be something on which one cannot write a legally binding contract). In this problem, there is a team of two people who need to be motivated, and there is randomness involved – effort leads to success, but lack of effort does not doom the an agents part of the project.

Two agents form a team to manage a joint project. Each agent is in charge of a different task. Each agent can either shirk or exert effort. If an agent exerts effort, then he performs his task successfully with certainty. If he shirks, his task succeeds with probability $\alpha < 1$. The common cost of effort is c. The joint project will succeed if and only if both tasks end successfully. The principal, who can neither monitor agents' effort nor the outcome of individual tasks, offers the agents rewards that are contingent on the projects outcome. If the project suceeds, agent 1 gets v_1 and agent 2 gets v_2. They both get zero if the project fails. Assume, now, that agents move sequentially. Agent 1 acts first. Agent 2 observes agent 1's effort decision, but does not observe the the outcome of his task, and then makes his own effort decision. We raise the following question: Is it possible that higher rewards for both agents will generate less effort in equilibrium?

Set $\alpha = 0.9$, $c = 1$, and assume, first, that $v_1 = 5.5$ and $v_2 = 11$.

1. Show that 2's optimal response to observing effort is to put in effort, and that 2's best response to observing shirking is to not put in effort. Suppose that 1 understands this, and believes that 2 will act accordingly. Show that 1's best choice is to exert effort. [Think one stage reward plus expected valued of where he ends up.]
2. Suppose now that the principal raises the rewards of both agents by 15 percent, thereby yielding $v_1 = 6.33$ and $v_2 = 12.66$. Show that 2's optimal response has changed, that he puts in effort no matter what 1 does. Show that 1's best choice, given how 2 will behave, is now to shirk.

This is called *incentive reversal*, the principal spent more money and got less effort. It turns out that incentive reversal is a possiblity under a broad range of conditions on the relations between effort and probability of success in multi-stage projects.