EC316a: Advanced Scientific Computation, Fall 2003

Notes Section 3

Discrete time, continuous state dynamic models

We consider now discrete–time models in which decisions are made over continuous state variables. We may define three classes of such models. First, there are models of centralized decision–making in which either discrete or continuous actions may be taken. The former category includes, e.g., decisions to exercise an option (a binary choice) given the continuous underlying price process, or to enter (exit) an industry. The latter category would include models in which the action taken could take on any real value, not restricted to integer values.

As a second class, we might consider strategic games among a small group of individuals, firms or countries in which actions are taken in response to the movements of continuous state variables and other agents' choices. For instance, we could consider the actions of the members of the OPEC cartel, deciding how much crude oil to produce, in response to the continuous movements of the spot price of crude.

As a third class, we could consider partial and general equilibrium models of collective, decentralized economic behavior. These models characteristically consider the importance of agents' expectations about the future course of the state variables, i.e. what will happen at a future date. If it is assumed that agents' expectations are consistent with the implications of the model, the model is termed a rational expectations model. Rational expectations models have been widely used to study asset returns, futures prices in a commodity market, producer responses to price support programs and workers' responses to alterations in tax and transfer policies.

Dynamic optimization and equilibrium models are closely related. The latter possess intertemporal arbitrage conditions: that is, possibilities to unambiguously improve one's welfare, or wealth, or income by taking a certain action are systematically destroyed by the collective action of optiimizing agents. The solution to continuous state / continuous action dynamic optimization models may be characterized by first–order intertemporal equilibrium conditions obtained by differentiating the Bellman equation: for instance, a first–order condition (or Euler equation) that states that an individual should find the optimal tradeoff between consumption today and consumption tomorrow. These models generally pose infinite–dimensional fixed–point problems that lack a closed–form solution, so that numerical solutions are required.

Continuous state dynamic programming

The discrete time, continuous state (DTCS) Markov decision model has the following structure: in each period t, an agent observes the state of the economic system s_t, takes an action x_t, and earns a reward $f(s_t, x_t)$ contingent upon both the state of the system and the action taken. The state is a controlled Markov process, in which the state at time $t+1$ depends on the state and action in period t and an exogenous random shock ϵ_{t+1}, unforecastable in period t:

$$s_{t+1} = g(s_t, x_t, \epsilon_{t+1})$$

The agent seeks a sequence of policies $\{x_t^*\}$, or rules, that prescribe the action $x_t = x_t^*(s_t)$ that should be taken in any future state and period so as to maximize the present value of current and expected future rewards over a time horizon T, discounted with the per–period factor δ. The horizon may be either infinite or finite, and the model may be deterministic or stochastic. In a stochastic model, the shocks are assumed to be $i.i.d.$—independently and identically distributed over time—and independent of prior states and actions. The state space may contain both continuous and integer variables, or it may be comprised wholly of continuous variables. The action space may, conceptually, include either continuous variables or discrete variables. We will consider only models in which all action variables are one sort or the other.

Models of this sort can be handled with dynamic programming methods based on the Bellman principle. In a finite–horizon problem, we denote as $V_t(s)$ the maximum attainable sum of current and expected future rewards given that the system is in state s at time t. The principle of optimality implies that the value functions must satisfy

$$V_t(s) = \max_{x \in X(s)} \{f(s,x) + \delta E V_{t+1}(g(s,x,\epsilon))\},$$
$$s \in S, \; t = 1, 2, \ldots, T$$

If the decision problem has an infinite horizon, the value function will not depend on time t, so we may write the Bellman equation as a functional fixed–point equation in the value function:

$$V(s) = \max_{x \in X(s)} \{f(s,x) + \delta E V(g(s,x,\epsilon))\}, \; s \in S$$

If the discount factor δ is less than unity and the reward function f is bounded, the mapping implied by the Bellman equation is a strong contraction on the space of bounded continuous functions and will possess a unique solution.

Euler equations

If a model of this sort has purely continuous state and action spaces, its solution may be characterized by first–order equilibrium conditions: the so–called Euler conditions, defining intertemporal arbitrage opportunities. Let us consider these conditions for an infinite–horizon model with twice continuously differentiable reward (f) and state transition (g) functions, and a discount factor δ less than unity. The equilibrium conditions involve not the value function, but rather its derivative

$$\lambda(s) = V'(s)$$

which is termed the shadow price function, representing the marginal value of the state variable to the optimizer: that is, the price that the optimizer is willing to pay to relax the constraint.

For a discrete time, continuous state, continuous action Markov decision problem, we apply the Kuhn–Tucker and Envelope theorems to the optimization problem. Assuming actions are unconstrained, the Kuhn–Tucker conditions imply that the optimal action x, given state s, satisfies the "equimarginality" condition

$$f_x(s,x) + \delta E[\lambda(g(s,x,\epsilon))g_x(s,x,\epsilon)] = 0$$

while the Envelope Theorem applied to this problem implies

$$f_s(s,x) + \delta E[\lambda(g(s,x,\epsilon))g_s(s,x,\epsilon)] = \lambda(s)$$

Here f_x (and similar expressions) represents the partial derivative of f with respect to that argument. Note that in the Kuhn–Tucker condition, the derivatives are taken with respect to the action variable x, denoting that one cannot improve by altering x. In the Envelope Theorem, the derivatives are taken with respect to the continuous state variable s, and the right hand side reflects the shadow price of that value of the state. In certain applications, the state transition depends only on the action taken by the agent, and not upon the prior state, so that $g_s = 0$, and one may substitute the second condition into the first, which defines the Euler equation as a single functional equation in a single unknown, the optimal policy x.

An Euler equation problem may also be constrained, by defining bounds on the policy variable; for instance

$$a(s) \leq x \leq b(s)$$

where a and b are differentiable functions of the state s. In these instances, the Euler conditions reflect the degree to which the constraint binds. For instance, the Kuhn–Tucker conditions are modified to

$$f_x(s, x) + \delta E[\lambda(g(s, x, \epsilon))g_x(s, x, \epsilon)] = \mu$$

where x and μ satisfy the complementarity conditions

$$x_i > a_i(s) \rightarrow \mu_i \geq 0$$
$$x_i < b_i(s) \rightarrow \mu_i \leq 0$$

where μ is a vector, the i^{th} element of which (μ_i) measures the current and expected future reward from a marginal increase in the i^{th} action variable x_i. At the optimum, μ_i must be nonpositive if x_i is less than its upper bound, or rewards could be increased by raising x_i. Similarly, μ_i must be nonnegative if x_i exceeds its lower bound, or rewards could be increased by costlessly decreasing x_i. Thus, models may contain constraints upon actions—such as a limit on the amount of borrowing against future income that a consumer might make–which would prevent them from reaching their unconstrained optimum.

For a deterministic model, we may seek to examine the steady–state properties of a model, which if it exists is the solution to a nonlinear equation. For an unconstrained deterministic problem, the steady state consists of a state s^*, an action x^* and a shadow price λ^* that satisfy the Euler and state stationarity conditions

$$\begin{aligned} f_x(s^*, x^*) + \delta\lambda^* g_x(s^*, x^*) &= 0 \\ \lambda^* = f_s(s^*, x^*) + \delta\lambda^* g_s(s^*, x^*) \\ s^* &= g(S^*, x^*) \end{aligned}$$

The steady–state conditions (for this problem or its constrained counterpart) pose a finite–dimensional problem which can usually be solved by standard numerical methods. In simpler applications, the steady–state conditions can often be solved analytically, even when the Bellman and Euler equations do not possess closed–form solutions. Many stochastic economic models are solved by a method of linearization around a trial solution (as we will discuss later when considering DYNARE).

Continuous State Discrete Choice models

We return to the asset replacement example considered in discrete state modeling, considering that now to be a continuous–state process with stochastic components. At the beginning of each year, the manufacturer must decide whether to continue to use an asset or replace it. An a–year–old asset produces $q(a)$ units of output up to $\bar{a}$ years, when it becomes unsafe and must be replaced at a cost of c. The profit contribution of one unit of output is an exogenous continuous–valued Markov process

$$p_{t+1} = h(p_t, \epsilon_{t+1})$$

We seek the profit–maximizing replacement policy. This is an infinite–horizon stochastic model with time t measured in years. Although the asset has maximum life $\bar{a}$ years, the firm is infinitely–lived, and maximizes profit over the infinite horizon. There are two state variables, $p \in (0, \infty)$ and $a \in \{1, 2, 3, \ldots, \bar{a}\}$. The action variable $x \in \{keep, replace\}$ is discrete, so that the state transition function is

$$g(p,a,x,\epsilon) = \begin{cases} (h(p,\epsilon), a+1), & x = keep \\ (h(p,\epsilon), 1), & x = replace \end{cases}$$

With the reward function

$$f(p,a,x) = \begin{cases} (pq(a), & x = keep \\ (pq(0) - c, & x = replace \end{cases}$$

The value of an asset of age a satisfies the Bellman equation

$$V(p,a) = max\{pq(a) + \delta EV(h(p,\epsilon), a+1), pq(0) - c \\ + \delta EV(h(p,\epsilon), 1)\}$$

Even though the profit contribution in future years is unknown, the optimization problem may still be solved by taking expectations, since the ϵ process has expected value zero. A numerical implementation of this model is contained in `demdp01`.

We may also consider the option pricing model developed in discrete state modeling. We now relax the assumption of a binomial pricing process and allow the price to follow a continuous stochastic Markov process

$$p_{t+1} = h(p_t, \epsilon_{t+1})$$

The American put option gives the purchaser the right (but not the obligation) to sell a specified quantity of a commodity at strike price K on or before the expiration date T. This is a finite horizon stochastic model with state variables $p \in (0, \infty)$ and $d \in \{0, 1\}$, where d is a discrete variable expressing the exercise status of the option: 1 if it has been exercised, 0 otherwsise. The state transition function is

$$g(p, d, x, \epsilon) = (h(p, \epsilon), x)$$

and the reward function is

$$f(p, d, x) = \begin{cases} K - p, & d = 0,\ x = 1 \\ 0, & otherwise \end{cases}$$

The value of an unexercised option in period t, given that the commodity price is p, satisfies the Bellman equation

$$V_t(p,0) = max\{K - p, \delta E V_{t+1}(h(p,\epsilon), 0\}$$

subject to the terminal condition $V_{T+1}(p,0) = 0$. The value of a previously exercised option is zero, regardless of the price of the commodity: $V_t(p,1) = 0$. A numerical implementation of this model is contained in `demdp04`. Note that there are four errata in the description of this model on p.197 of the text.

Let us now consider some continuous state, continuous choice models. An economy produces a single composite good. Each year, there is a predetermined amount of the good s in stock, of which x is invested and the remainder, $s-x$, is consumed, yielding a social benefit $u(s-x)$. The amount of the good available is a controlled continuous–valued Markov process:

$$s_{t+1} = \gamma x_t + \epsilon_{t+1} h(x_t)$$

where the first term represents depreciation (γ is the capital survival rate), h is the aggregate production function, and ϵ is a positive production shock with mean of unity. What balance of consumption and investment will maximize the social planner's objective of current and expected future social benefits?

This is an infinite horizon stochastic model, with the state variable s representing the amount of the good available each year, and the action variable x being the amount invested. The state transition function is

$$g(s, x, \epsilon) = \gamma x + \epsilon h(x)$$

with reward function

$$f(s, x) = u(s - x)$$

Given a stock s, the sum of current and expected future benefits satisfies the Bellman equation

$$V(s) = \max_{0 \le x \le s} \{u(s - x) + \delta E V(\gamma x + \epsilon h(x))\}$$

If we assume that $u'(0) = \infty$ and $h(0) = 0$, the constraints will never be binding at an optimum: that is, there will be an interior solution to the problem, satisfying the Euler conditions

$$u'(s - x) - \delta E[\lambda(\gamma x + \epsilon h(x))(\gamma x + \epsilon h'(x))] = 0$$

$$\lambda(s) = u'(s - x)$$

which taken together imply that along the optimal path

$$u'_t = \delta E_t[u'_{t+1}(\gamma + \epsilon_{t+1} h'(x))]$$

where u'_t is the current marginal utility of consumption, and $\epsilon_{t+1} h'(x)$ is the following period's marginal product of capital. The utility derived from consuming one unit of the good today must equal the discounted expected utility derived from investing one unit of the good and consuming its yield tomorrow.

A certainty–equivalent steady state may be derived by setting ϵ to its mean of unity and solving the nonlinear equation system

$$\begin{aligned} u'(s^* - x^*) &= \delta\lambda^*(\gamma + h'(x^*)) \\ \lambda^* &= u'(s^* - x^*) \\ s^* &= \gamma x^* + h(x^*) \end{aligned}$$

where the starred variables are the steady–state values. These conditions imply the "golden rule" of economic growth models: that $1-\gamma+r = h'(x^*)$, where $\delta = 1/(1+r)$. This states that the marginal product of capital must equal the capital depreciation rate plus the discount rate.

A numerical example of this model is presented in demo `demdp07`.

Let us now consider an example based on the mine management problem. Each year begins with a stock of ore s and an extraction amount

x, which involves a total cost $c(s, x)$ and a revenue $p(x)$, where $c_s \leq 0$, $c_x \geq 0$, $c_s(s, 0) = 0$ and $p' < 0$. Given that the current stock of ore is $\bar{s}$, what is the profit–maximizing extraction policy?

This is an infinite horizon deterministic model. The state variable $s \in [0, \bar{s}]$ is the stock of ore at the beginning of the year, and the action variable $x \in [0, s]$ is the amount of ore extracted. The state transition function is merely

$$g(s, x) = s - x$$

and the reward function is

$$f(s, x) = p(x)x - c(s, x)$$

The value of a mine with ore stock s must satisfy the Bellman equation

$$V(s) = \max_{0 \leq x \leq s} \{p(x)x - c(s, x) + \delta V(s - x)\}$$

At some stock level, it may be optimal strategy to abandon the mine if it is not possible to earn a profit extracting the remaining ore. The Euler conditions take the form of a complementarity condition, with the shadow price $\lambda(s)$ of the resource derived from

$$p(x) + p'(x)x - c_x(s, x) - \delta\lambda(s - x) = \mu$$
$$\lambda(s) = c_s(s, x) + \delta\lambda(s - x) + max(\mu, 0)$$

where the ore extracted x and the long–run marginal profit of extraction μ must satisfy the complementarity condition

$$0 \leq x \leq s$$
$$x > 0 \rightarrow \mu \geq 0$$
$$x < s \rightarrow \mu \leq 0$$

Thus, in every period, either ore is extracted until the long–run marginal profit is driven to zero, or the mine is abandoned because it is not possible to do so. There will be a unique steady state in which the mine will be abandoned when the ore stock reaches the critical level s^*, derived from $c_s(s^*, 0) = 0$. Until the mine is abandoned, it will be operated such that the marginal revenue of extracted ore equals the shadow price of unextracted ore plus the marginal cost of extraction:

$$p_t + p_t' x_t = c_{x_t} + \delta \lambda_{t+1}$$

and the shadow price of unextracted ore will rise, at the same rate at which the cost of extraction rises as a function of the remaining stock.

A numerical example of this model is presented in demo `demdp09`.

As the last model in this category, let us consider a production–inventory model. The firm chooses to maximize long–run profit by managing its levels of production and inventories. Each period the firm has a predetermined stock of inventory s and decides how much to produce (q) and how much to store (x), buying or selling the resulting difference $s + q - x$ on the open market at the price p. The firm's production and storage costs are given by functions $c(q)$ and $k(x)$ respectively, and the market price follows an exogenous Markov process

$$p_{t+1} = h(p_t, \epsilon_{t+1})$$

This is an infinite horizon stochastic model with two state variables: $s \in [0, \infty)$ and $p \in [0, \infty)$ measuring beginning inventories and the current market price, respectively. There are two action variables, $q \in [0, \infty)$ and $x \in [0, \infty)$, with state transition function

$$g(s, p, q, x, \epsilon) = (x, h(p, \epsilon))$$

and reward function

$$f(s, p, q, x) = p(s + q - x) - c(q) - k(x)$$

The value of the firm, given the initial conditions, satisfies the Bellman equation

$$V(s, p) = \max_{0 \leq q, 0 \leq x} \{p(s + q - x) - c(q) - k(x) + \delta E V(x, h(p, \epsilon))\}$$

If production is subject to increasing marginal costs and $c'(0)$ is sufficiently small, production will be positive in all states, and the shadow price of beginning inventories $\lambda(s,p)$ will satisfy the Euler conditions

$$\begin{aligned} p &= c'(q) \\ \delta E\lambda(x, h(p,\epsilon)) - p - k'(x) &= \mu \\ \lambda(s,p) &= p \\ x \geq 0,\ \ \mu \leq 0,\ \ x > 0 &\rightarrow \mu = 0 \end{aligned}$$

It follows that along the optimal path

$$\begin{aligned} p_t &= c'_t \\ x_t &\geq 0 \\ \delta E p_{t+1} - p_t - k'_t &\leq 0 \\ x > 0 \rightarrow \delta E p_{t+1} - p_t - k'_t &= 0 \end{aligned}$$

Implying that the firm's production and storage decisions are independent. Production is governed by the familiar short–run condition for profit maximization, that marginal revenue be set equal to marginal cost. Storage is entirely driven by intertemporal arbitrage opportunities. If the expected marginal profit from storage is negative, no storage is undertaken. Otherwise, stocks are accumulated up to the point at which the marginal cost of storage equals the expected appreciation in the market price, in present value terms.

A numerical example of this model is presented in demo `demdp13`.

In the remaining section of the module, we will discuss the computational methodologies for solution of these discrete time, continuous state dynamic models.