Discrete-Time Stochastic Dynamic Programming

These notes extract results from the previously distributed dynamic programming notes. The main difference is just that proofs, remarks, and discussion have been taken out to allow listing the results in compact form. Theorem 2 below uses a strengthened regularity condition (requiring that every solution to the Bellman equation be an attainable maximum, rather than just a least upper bound) that allows the result to be stated more simply, while also relaxing its condition (iii). The sections of the previous notes not summarized here, on value iteration, constraints at infinity, and exogenous states, were already in compact form in the previous notes.

I. Notation and basic assumptions

We consider a problem defined in terms of

- t : a time index, with integer values
- C : a *k*x1 vector, called the control vector
- S : an *n*x1 vector, called the state vector
- $\Gamma(\cdot)$: a mapping from state vector values to subsets of R^k, defining constraints on the choice of C
- $\Im_t$: the information set at *t*, consisting of $\{C(s), S(s), \varepsilon(s), \text{all } s \le t\}$
- ε_t : a *p*x1 random vector of disturbances at time t .

The objective is to maximize

$$E\left[\sum_{t=0}^{\infty} \beta^t U(C_t, S_t)\right] \tag{1}$$

by choice of $\{C_t, S_{t+1}, t = 0, \ldots, \infty\}$. We assume that the infinite sum inside the brackets in (1) is well-defined for each choice of *C*'s that satisfies the constraints below and that the expectation of the sum is well-defined for each such choice of *C*'s.

A) S_0 is given, not subject to choice;

B) for each *t*=1,...,∞, S_t is determined from past history and current ε_t according to

$$S_t = f(C_{t-1}, S_{t-1}, \varepsilon_t) \ . \tag{2}$$

C) for each *t*, C_t is constrained to lie in the set $\Gamma(S_t)$;

D) for each *t*, C_t is allowed to depend only on information in $\Im_t$, and only in such a way that C_t and $U(C_t, S_t)$ are well-defined random variables.

E) for each t, ε_{t+1} is independent of C_t and of all the random variables in $\mathfrak{I}_t$[1] and that

F) the random variables $\{\varepsilon_t, t=1,\ldots,\infty\}$ are mutually independent and identically distributed (i.i.d.).

We denote by $\mathcal{S}$ the set of all possible values of S. We denote by $V(\cdot)$ the function mapping S's in $\mathcal{S}$ into the least upper bound of achievable values of the objective function. If the problem is well defined, $V(S)$ exists for each S in $\mathcal{S}$, though it is important in practice to check that the infinite sum in (1) indeed converges for all feasible choices of actions. V is called the **value function**.

II. The principle of optimality: necessity and sufficiency

Theorem 1: (**The Principle of Optimality)** Suppose that V is the value function for the problem of maximizing (1) subject to (A)-(F). Then for each S in $\mathcal{S}$, $E\big[V\big(f(C,S,\varepsilon)\big)\big]$ exists for all C in $\Gamma(S)$ (with ±infinity allowable values), and

$$V(S) = \underset{C \text{ in } \Gamma(S)}{\text{l.u.b.}} \left\{U(C,S) + \beta E\big[V\big(f(C,S,\varepsilon)\big)\big]\right\} \tag{3}$$

Theorem 2: Suppose that V is the value function for the problem of maximizing (1) subject to (A)-(F). Then for every $\delta > 0$, it is possible to choose a policy function $C_\delta^*(\cdot)$ such that, for each S in $\mathcal{S}$ with $V(S)$ finite, the value of the objective function attained using $C_\delta^*(\cdot)$ is at least $V(S)-\delta$, and the sequence of $S_t, t=1,\ldots,\infty$ generated by setting $S_0 = S$ and

$$S_t = f(C_\delta^*(S_{t-1}), S_{t-1}, \varepsilon_t),\ t=1,\ldots,\infty \tag{4}$$

satisfies

$$\beta^t E_0 V(S_t) \xrightarrow[t\to\infty]{} 0\,. \tag{5}$$

Theorem 3: Suppose there is a function V^* that satisfies (3) for every S in $\mathcal{S}$ and that in addition

i) for every S in $\mathcal{S}$ there is a value $C^*(S)$ for C that attains the maximum on the right-hand side of (3) with $V = V^*$;

ii) for every S_0 in $\mathcal{S}$, if S_t, $t=1,\ldots,\infty$, is generated from

$$S_{t+1} = f(C^*(t), S_t, \varepsilon_{t+1}) \tag{6}$$

then

$$\beta^t E_0\big[V^*(S_t)\big] \xrightarrow[t\to\infty]{} 0 \text{ ; and} \tag{7}$$

---

[1] Since C_t is required by (D) above to depend only on random variables in $\mathfrak{I}_t$, (E) could omit the "of C_t and" phrase.

iii) for any $V^{**} \neq V^{*}$ that solves (3) for every S in $\mathcal{S}$, there is an associated policy function C^{**} satisfying

$$U(C^{**}(S),S)+\beta E\Big[V^{**}\Big(f(C^{**}(S),S,\varepsilon\Big)\Big]=V^{**}(S)\,, \tag{8}$$

and for every S_0 in $\mathcal{S}$, if S_t, $t=1,...,\infty$, are generated from

$$S_t = f(C^{**}(S_{t-1}),S_{t-1},\varepsilon_t),\; t=1,...,\infty \text{ , then} \tag{9}$$

$$\liminf_{t\to\infty} \beta^t E_0 V^{*}(S_t) \geq 0\,. \tag{10}$$

Then V^{*} is the value function for the problem.

Corollary: If U is bounded below, the necessary conditions of Theorems 1 and 2 are also sufficient.

III. First Order Conditions

Sometimes the following equation is called, by itself, "the" first order condition,

$$\frac{\partial U(C,S)}{\partial C}+\beta E\left[V'(f(C,S,\varepsilon)\cdot\frac{\partial f}{\partial C}\right]=\lambda(S)\frac{\partial H(C,S)}{\partial C}\;, \tag{11}$$

while this next one,

$$V'=\frac{\partial U}{\partial S}+\beta E\left[V'\big(f(C,S,\varepsilon)\big)\cdot\frac{\partial f}{\partial S}\right]-\lambda(S)\cdot\frac{\partial H(C,S)}{\partial S}\;, \tag{12}$$

is called the "envelope condition". Note that both emerge simply as first order conditions if we apply to the standard dynamic programming setup of objective function and constraints the usual rules for generating first order conditions from stochastic Lagrange multipliers, with V' playing the role of the vector of Lagrange multipliers.

Equations (11) and (12) together are sometimes called the Euler equations for the problem. Note that one way to remember them is to form the "Hamiltonian-like" expression

$$V(S)-U(C,S)-\beta E\big[V(f(C,S,\varepsilon))\big]+\lambda H(C,S) \tag{13}$$

Equations (11) and (12) are then the partial derivatives of (13) with respect to C and S, respectively.