Lecture Notes 6
Stochastic Dynamic Programming

Assaf Patir

21.11.2013

1 Mathematical Preliminaries

1.1 Disclaimer

Probability theory relies heavily on the concept of measurability: given a probability space $(\Omega, \mathfrak{F}, P)$ and a measurable space $(S, \mathfrak{S})$, a function $f : \Omega \to S$ is **measurable** if for all $B \in \mathfrak{S}$, the pre-image $f^{-1}(B)$ is measurable in Ω, i.e. $\{\omega | \, f(\omega) \in B\} \in \mathfrak{F}$.[1]

In order to extend the formalism of dynamic programming to include uncertainty, we will need to require that various functions – e.g. the target function F, the constraint Γ, etc. – are measurable functions, and we will need to prove that the functions we construct are measurable as well. All these issues will be completely ignored in what follows. The issue of measurability has virtually no implications for macroeconomic applications, but it does mean that all our proofs here are actually incomplete. A more rigorous treatment can be found in Stokey-Lucas' textbook.

1.2 Stochastic Processes

Definition 1. Given a a probability space $(\Omega, \mathfrak{F}, P)$, a **stochastic process** is a set $\{X_t \,|\, t \in \mathbb{N}\}$ such that for each $t \in \mathbb{N}$, $X_t : \Omega \to \mathbb{R}^n$ is a random variable (for some fixed n).

Since each one of the X_t's is a random variable, we can define cumulative probability functions and so on as usual. In particular, we define a **transition**

[1] For example, we can define a *random variable* as a function $X : \Omega \to \mathbb{R}$ such that for all $c \in \mathbb{R}$, $\{\omega | \, X(\omega) \leq c\}$ is measurable.

function as

$$T_t(x_t; x_{t-1}) = F_{X_t|X_{t-1}=x_{t-1}}(x_t),$$

i.e. the conditional distribution of $X_t|X_{t-1}$. If the random variables $\{X_t\}$ are all discrete and only take values in a finite set $\{a_1, \ldots, a_q\}$ then every transition function can be described by a q-by-q matrix:

$$T_t^{i,j} = P(X_t = a_i \,|\, X_{t-1} = a_j).$$

Definition 2. Given a a probability space $(\Omega, \mathfrak{F}, P)$, a stochastic process $\{X_t \,|\, t \in \mathbb{N}\}$ is said to be a **Markov process** if for all $t \geq 1$, and for all A measurable in $\mathbb{R}^n$,

$$P(X_t \in A | X_{t-1}, X_{t-2}, \ldots, X_0) = P(X_t \in A | X_{t-1}),$$

and, in addition, the transition function is independent of t.

In other words, for a Markov process, if one is trying to predict X_t and knows X_{t-1}, the rest of the history of the process is irrelevant, and it is also irrelevant what t is. For example, if $\{\varepsilon_t\}_{t=1}^{\infty}$ are a set independent and identically distributed random variables, then X_t defined by

$$\forall t \geq 1, \qquad X_t = \rho X_{t-1} + \varepsilon_t, \tag{1}$$

is a Markov process, where ρ is a constant. The transition function is

$$T(x_t; x_{t-1}) = P(X_t \leq x_t | X_{t-1} = x_{t-1}) = P(\varepsilon_t \leq x_t - \rho x_{t-1}) = F_\varepsilon(x_t - \rho x_{t-1}).$$

The stochastic process defined by

$$X_t = \rho_1 X_{t-1} + \rho_2 X_{t-2} + \varepsilon_t$$

is *not* a Markov process because $X_t|X_{t-1}$ and $X_t|X_{t-1}, X_{t-2}$ do not have the same distribution because the latter depends on the realization of X_{t-2}. However, if we define $Y_t = (X_t, X_{t-1})'$, then Y_t is Markov since

$$Y_t = \begin{pmatrix} \rho_1 & \rho_2 \\ 1 & 0 \end{pmatrix} Y_{t-1} + \begin{pmatrix} \varepsilon_t \\ 0 \end{pmatrix}.$$

For a Markov process, the transition function is all that one needs to know in order to describe the entire process (and also X_0). For example, in the discrete case that we discussed above, the transition matrix is now just T^{ij} without an index t, and we have:

$$
\begin{aligned}
&P(X_{t+k} = a_i | X_t = a_j) = \\
&= \sum_{j_1} P(X_{t+k} = a_i | X_{t+k-1} = a_{j_1}, X_t = a_j) P(X_{t+k-1} = a_{j_1} |\, X_t = a_j) = \\
&= \sum_{j_1} P(X_{t+k} = a_i | X_{t+k-1} = a_{j_1}) P(X_{t+k-1} = a_{j_1} |\, X_t = a_j) = \\
&= \sum_{j_1} T^{i,j_1} P(X_{t+k-1} = a_{j_1} |\, X_t = a_j) = \\
&= \sum_{j_1} \sum_{j_2} T^{i,j_1} T^{j_1,j_2} P(X_{t+k-2} = a_{j_2} |\, X_t = a_j) = \\
&\qquad\qquad = \sum_{j_1} \cdots \sum_{j_{k-1}} (T^{i,j_1} \cdots T^{j_{k-1},j}) = (T^k)^{i,j}.
\end{aligned}
$$

Definition 3. Given a a probability space $(\Omega, \mathfrak{F}, P)$, a stochastic process $\{X_t \,|\, t \in \mathbb{N}\}$ is said to be a **stationary** if for all $k, t, \tau \geq 0$, $(X_t, X_{t+1}, \ldots, X_{t+k})$ and $(X_{t+\tau}, X_{t+1+\tau}, \ldots, X_{t+k+\tau})$ are equal in distribution.

Finally, in what follows we denote the history of the process up until time t as $s^t = (X_0, X_1, \ldots, X_t)$, and s^∞ as the entire process.

2 Sequential Approach

Let $\beta \in (0,1)$, X be the state space, $\{a_t\}$ be a stochastic process that takes values in a set S, $F : X \times X \times S \to \mathbb{R}$, and $\Gamma : X \times S \to 2^X$. In the sequential approach to dynamic programming we seek to solve

$$
\max_{\{x_{t+1}(s^t)\}_{t=0}^\infty} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t F(x_t, x_{t+1}, a_t), \tag{2}
$$

subject to $x_{t+1} \in \Gamma(x_t, a_t)$ and x_0 and a_0 given. $\mathbb{E}_t[\cdots]$ is shorthand for $\mathbb{E}[\cdots | s^t]$.

In the deterministic case, we imagined the agent choosing at $t = 0$ the entire sequence $(x_1, x_2, \ldots)$. In the stochastic case, we imagine the agent making a big list of contingency plans $x_{t+1}(s^t)$, that is, at time 0 the agent plans that if at time t the history of the world is s^t, then she will choose $x_{t+1}(s^t)$.

Now, let's pause for a moment to think about the $\mathbb{E}_0$ operator in (2). Schematically, what we are doing is summing over all possible histories $\{a_t\}_{t=0}^\infty$, for each history we are evaluating the utility along that path, and weigh it by the probability that it occurs. If there was only a finite number of possible histories, then the utility could be written explicitly:

$$U = \sum_{(a_1,a_2,\dots)} P\Big(s^\infty = (a_0, a_1, a_2, \dots) |\, s^0\Big) \sum_{t=0}^{\infty} \beta^t F(x_t(s^{t-1}), x_{t+1}(s^t), a_t).$$

The choice of $x_{t+1}(s^t)$ affects only histories which start with s^t whatever they continue with, and for each such history affect two terms in the second sum. If we ignore the constraint Γ for a moment, we have

$$\begin{aligned}
0 = \frac{\partial U}{\partial x_{t+1}(s^t)} &= \sum_{(a_{t+1},a_{t+2},\dots)} P\Big(s^\infty = (s^t, a_{t+1}, a_{t+2}, \dots) |\, s^0\Big) \times \\
&\times [F_2(x_t(s^{t-1}), x_{t+1}(s^t), a_t) + \beta F_1(x_{t+1}(s^t), x_{t+2}(s^{t+1}), a_{t+1})] = \\
&= \sum_{a^{t+1}} [F_2(x_t(s^{t-1}), x_{t+1}(s^t), a_t) + \beta F_1(x_{t+1}(s^t), x_{t+2}(s^{t+1}), a_{t+1})] \times \\
&\times \sum_{(a_{t+2},\dots)} P\Big(s^\infty = (s^t, a_{t+1}, a_{t+2}, \dots) |\, s^0\Big) = \\
&= \sum_{a^{t+1}} [F_2(x_t(s^{t-1}), x_{t+1}(s^t), a_t) + \beta F_1(x_{t+1}(s^t), x_{t+2}(s^{t+1}), a_{t+1})] \times \\
&\qquad \times P\Big(s^{t+1} = (s^t, a_{t+1}) |\, s^0\Big).
\end{aligned}$$

Dividing both sides by $P(s^t|s^0)$,

$$\begin{aligned}
0 &= \sum_{a^{t+1}} [F_2(x_t(s^{t-1}), x_{t+1}(s^t), a_t) + \beta F_1(x_{t+1}(s^t), x_{t+2}(s^{t+1}), a_{t+1})] \times \\
&\times P\Big(s^{t+1} |\, s^t\Big) = \\
&= F_2(x_t(s^{t-1}), x_{t+1}(s^t), a_t) + \beta \mathbb{E}_t[F_1(x_{t+1}(s^t), x_{t+2}(s^{t+1}), a_{t+1}) | s^t].
\end{aligned}$$

This is the stochastic version of the Euler equation:

$$\boxed{0 = F_2(x_t, x_{t+1}, a_t) + \beta \mathbb{E}_t[F_1(x_{t+1}, x_{t+2}, a_{t+1})]}\,. \tag{3}$$

Of course, in the typical application, the number of possible histories is infinite, and the probability of each possible history is zero, so the above derivation will not work. One has to carefully define $\{x^t\}$ as a sequence of random variable and be careful about issues of measurability. However, the spirit of the derivation still works, and the result (3) holds.

The stochastic version of the transversality condition is

$$\lim_{\tau\to\infty} \mathbb{E}_t[\beta^\tau F_1(x_\tau, x_{\tau+1})x_\tau] = 0.$$

We allow paths where the non-stochastic transversality condition is violated, but these must have measure zero, and furthermore, this has to be true when evaluating these paths at any point in time.

3 Recursive Approach

In the recursive approach we define the value function

$$v(x_0, a_0) = \max_{\{x_{t+1}(s^t)\}_{t=0}^\infty} \mathbb{E}_0 \sum_{t=0}^\infty \beta^t F(x_t, x_{t+1}, a_t). \tag{4}$$

Following in the same steps as we did in the non-stochastic case

$$\begin{aligned} v(x_0, a_0) &= \max_{\{x_{t+1}(s^t)\}_{t=0}^\infty} \left\{ F(x_0, x_1, a_0) + \beta \mathbb{E}_0 \sum_{t=1}^\infty \beta^{t-1} F(x_t, x_{t+1}, a_t) \right\} = \\ &= \max_{\{x_1(s^0)\}} \max_{\{x_{t+1}(s^t)\}_{t=1}^\infty} \left\{ F(x_0, x_1, a_0) + \beta \mathbb{E}_0 \sum_{t=0}^\infty \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) \right\} = \\ &= \max_{x_1 \in \Gamma(x_0, a_0)} \left\{ F(x_0, x_1, a_0) + \beta \max_{\{x_{t+1}(s^t)\}_{t=1}^\infty} \mathbb{E}_0 \sum_{t=0}^\infty \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) \right\}. \end{aligned}$$

On the last line, the second term within the braces calls for us to find the sequence from a given x_1 that maximizes the expected value of the discounted sum given a_0. It is similar to (4), but different in that the definition in (4) has the realization of the first element in the stochastic process a_t given, and here we are only given a_0, but the sum starts from a_1. Without further assumptions, this is as far as we can go. If we assume that $\{a_t\}$ is a Markov process, then we can use the law of total probability to write

$$
\begin{aligned}
&\max_{\{x_{t+1}(s^t)\}_{t=1}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) = \\
&= \max_{\{x_{t+1}(s^t)\}_{t=1}^{\infty}} \mathbb{E}_{\{a_t\}_{t=1}^{\infty}} \left[\sum_{t=0}^{\infty} \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) \middle| a_0 \right] = \\
&= \max_{\{x_{t+1}(s^t)\}_{t=1}^{\infty}} \mathbb{E}_{a_1|a_0} \mathbb{E}_{\{a_t\}_{t=2}^{\infty}} \left[\sum_{t=0}^{\infty} \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) \middle| a_0, a_1 \right] = \\
&= \max_{\{x_{t+1}(s^t)\}_{t=1}^{\infty}} \mathbb{E}_{a_1|a_0} \mathbb{E}_{\{a_t\}_{t=2}^{\infty}} \left[\sum_{t=0}^{\infty} \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) \middle| a_1 \right] = \\
&= \mathbb{E}_{a_1|a_0} \left[\max_{\{x_{t+1}(s^t)\}_{t=1}^{\infty}} \mathbb{E}_{\{a_t\}_{t=2}^{\infty}} \left[\sum_{t=0}^{\infty} \beta^t F(x_{t+1}, x_{t+2}, a_{t+1}) \middle| a_1 \right] \right] = \mathbb{E}_0 v(x_1, a_1).
\end{aligned}
$$

Therefore,

$$
v(x, a) = \max_{x' \in \Gamma(x,a)} \left\{ F(x, x', a) + \beta\, \mathbb{E}[v(x', a') |\, a] \right\}. \tag{5}
$$

This is the stochastic version of the functional equation, which is also called the Bellman equation. If we manage to solve this equation we end up with a policy function $x'(x, a)$.

3.1 Generalization

We have so far considered a case where the agent chooses at time t the state variable for the next period x_{t+1} based on the current realization of uncertainty a_t. One example where this is appropriate is as follows: consider a variant of the neoclassical growth model where production depends not only on the capital stock (e.g. seeds) but also on the realization of some uncertainty (e.g. rainfall), and is given by $y_t = a_t f(k_t)$. The functional equation for this problem is

$$
v(k, a) = \max_{0 \le k' \le af(k)} \{u(af(k) - k') + \beta \mathbb{E}[v(k', a') |\, a]\}.
$$

A different class of problems are ones where the value of the state variable in the next period is itself stochastic. For example, consider the neoclassical model again, but not with a deterministic production function $f(k)$. Imagine instead that when the agent saves k' the actual amount of capital that survives to the next period is stochastic and given by $a'k'$ (e.g. some seed gets eaten by mice). This adjustment can easily be done in the recursive formalism, we simply write

$$
v(k, a) = \max_{0 \le k' \le f(k)} \{u(f(k) - k') + \beta \mathbb{E}[v(a'k', a') |\, a]\}.
$$