Topics in Computational Economics

Lecture 12

John Stachurski

NYU 2016

Preliminary Comments

HW: See

http://nbviewer.jupyter.org/github/jstac/quantecon_nyu_2016_homework/tree/master/hw_set_9/

The next slide gives a proof for Ex. 4

Show that $\mathcal{M}_+ :=$ all nonnegative definite $A \in \mathcal{M}(n \times n)$ is closed in $(\mathcal{M}, \| \cdot \|)$

Let $\{A_k\} \subset \mathcal{M}_+$ with $\|A_k - A\| \to 0$ for some $A \in \mathcal{M}(n \times n)$

We claim that $A \in \mathcal{M}_+$

To see this, pick any $x \in \mathbb{R}^n$ and observe that, for any $k \in \mathbb{N}$,

$$x'Ax = x'A_kx + x'Ax - x'A_kx$$

$$\therefore \quad x'Ax \geqslant \epsilon_k := x'Ax - x'A_kx$$

It suffices to show that $\epsilon_k \to 0$

This holds because, by Cauchy–Schwartz and the definition of the spectral norm,

$$|\epsilon_k| = |x'(A - A_k)x| \leqslant \|x\| \|(A - A_k)x\| \leqslant \|x\|^2 \|A - A_k\|$$

Today's Lecture

General state Markov processes — the density case

- Formulation
- Distribution dynamics
- Stability
- Ergodicity

References

- Stokey and Lucas (1989)
- Stachurski (2009)
- Lasota and Mackey. Chaos, Fractals and Noise (1998)

Stochastics on General Spaces

In economics we deal with many kinds of states and state spaces

Examples. S equals

- a discrete set (e.g., rich, middle, poor)
- a subset of $\mathbb{R}$ (e.g., one sector growth)
- $\mathbb{R}^n$ (e.g., growth with more state variables)
- a set of distributions (e.g., heterogeneous agent models)

Let's take $S :=$ a Borel subset of $\mathbb{R}^n$

Let $L_1(S) := L_1(S, \mathscr{B}, \lambda)$ be all $f \in m\mathscr{B}$ with

$$\int |f(x)| \, \mathrm{d}x := \int |f| \, \mathrm{d}\lambda < \infty$$

The L_1 norm is

$$\|f\| := \int |f(x)| \, \mathrm{d}x$$

Let $\mathscr{D}(S)$ be the set of **densities** on S

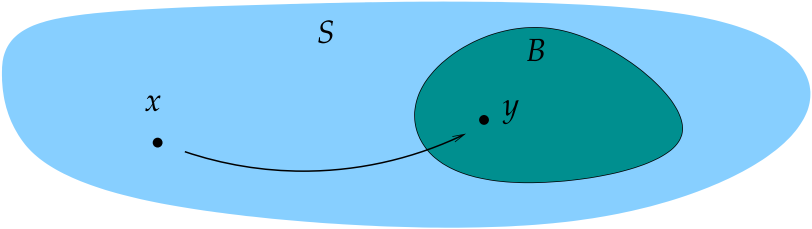
$$\mathscr{D}(S) := \{f \in L_1(S) : f \geqslant 0 \text{ and } \|f\| = 1\}$$

A **stochastic density kernel** on S is a function $p \colon S \times S \to \mathbb{R}$ such that

- $p \in m\mathscr{B}$
- $p(x, \cdot) \in \mathscr{D}(S)$ for all $x \in S$

Intuition:

1. One density $p(x, \cdot)$ for each $x \in S$
2. $\int_B p(x, y) \, \mathrm{d}y =$ prob of moving from x into B in one step



Example. Let $S = \mathbb{R}$ and let $p(x, y) = \phi(y)$ where ϕ is any density

This is the IID case

Example. Let $S = \mathbb{R}$ and let

$$p(x, y) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{(y - x)^2}{\sigma^2} \right\}$$

This stochastic kernel is from a **Gaussian random walk**

When current state is x, draw next state from $N(x, \sigma^2)$

From SDEs to SKs

Question. What is the stochastic kernel corresponding to

$$X_{t+1} = F(X_t, \xi_{t+1}), \qquad \{\xi_t\} \overset{\text{IID}}{\sim} \phi$$

Equivalent question: What is the density $p(x, \cdot)$ of

$$Y = F(x, \xi), \qquad \xi \sim \phi$$

Actually it might not exist

Example. F is a constant function or ϕ puts mass on a point

Fact. If $Y = b + a\xi$ with $\xi \sim \phi$ and $a \neq 0$, then the density of Y exists and equals

$$\phi_Y(y) = \phi\left(\frac{y-b}{a}\right)\frac{1}{|a|}$$

Proof for $a > 0$ case:

Letting F and F_Y be the CDFs of ξ and Y respectively,

$$F_Y(y) = \mathbb{P}\{Y \leqslant y\} = \mathbb{P}\{b + a\xi \leqslant y\} = \mathbb{P}\{\xi \leqslant (y-b)/a\}$$

$$\therefore \quad F_Y(y) = F((y-b)/a)$$

$$\therefore \quad \phi_Y(y) = \phi((y-b)/a)/a$$

Example. Consider the Solow–Swan model

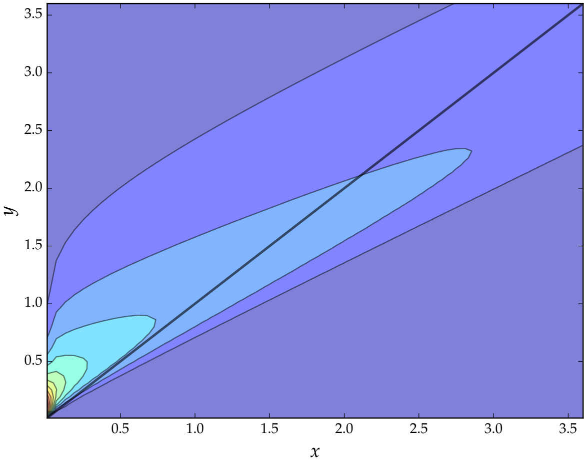
$$k_{t+1} = sf(k_t)\xi_{t+1} + (1-\delta)k_t \qquad \{\xi_t\}_{t \geqslant 1} \overset{\text{IID}}{\sim} \phi$$

Here

- k_t takes values in $S = (0, \infty)$
- $s, \delta \in (0,1)$ and $f(k) > 0$ when $k > 0$

The stochastic kernel is

$$p(x,y) = \phi\left(\frac{y - (1-\delta)x}{sf(x)}\right)\frac{1}{sf(x)}$$


3.5
3.0
2.5
2.0
1.5
1.0
0.5
y
0.5
1.0
1.5
2.0
2.5
3.0
3.5
x

Example. Consider the ARCH(1) model

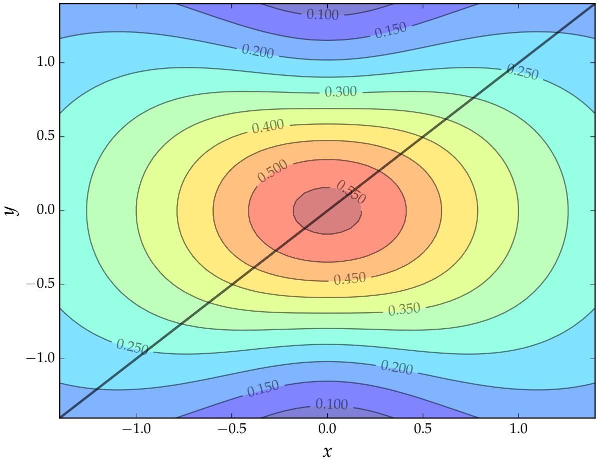
$$X_{t+1} = \sqrt{\alpha_0 + \alpha_1 X_t^2} \cdot \xi_{t+1} \qquad \{\xi_t\} \stackrel{\text{IID}}{\sim} N(0,1)$$

Here

- X_t takes values in $S = \mathbb{R}$
- $\alpha_0 > 0$, $\alpha_1 \geqslant 0$

The SK is

$$p(x,y) = \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 x^2)}} \exp\left\{ -\frac{y^2}{2(\alpha_0 + \alpha_1 x^2)} \right\}$$


0.100
0.150
0.200
0.250
0.300
0.400
0.500
0.550
0.450
0.350
0.250
0.200
0.150
0.100
1.0
0.5
0.0
−0.5
−1.0
y
−1.0
−0.5
0.0
0.5
1.0
x

Higher Order Kernels

Let p be an SK on S and let $\{p^k\}$ be defined by

$$p^1 := p \quad \text{and} \quad p^{k+1}(x, y) := \int p(x, z) p^k(z, y) \, \mathrm{d}z$$

Called the k**-step stochastic kernel**

Fact. If p is an SK on S, then so is p^k for all k

Fact. The kernels $\{p^k\}$ satisfy the **Chapman–Kolmogorov relation**

$$p^{j+k}(x, y) = \int p^k(x, z) p^j(z, y) \, \mathrm{d}z \qquad ((x, y) \in S \times S)$$

Markov Operators on $L_1(S)$

A linear map $P \colon L_1(S) \to L_1(S)$ satisfying

1. $g \geqslant 0 \implies gP \geqslant 0$
2. $g \geqslant 0 \implies \|gP\| = \|g\|$

is called a **Markov operator** on S

In other words, $P \colon L_1(S) \to L_1(S)$ is a Markov operator if it is

1. linear
2. positive (i.e., invariant on the positive cone) and
3. norm preserving on the positive cone

Properties of Markov Operators

Fact. If P is a Markov operator on S, then P is invariant on $\mathscr{D}(S)$

In other words,

$$\phi \in \mathscr{D}(S) \quad \Longrightarrow \quad \phi P \in \mathscr{D}(S)$$

Proof: Fix $\phi \in \mathscr{D}(S)$ and let P be a Markov operator

Since $\phi \in \mathscr{D}(S)$, we have $\phi \geqslant 0$ and hence $\phi P \geqslant 0$

Also, $\phi \geqslant 0$ and $\|\phi\| = 1$, so $\|\phi P\| = \|\phi\| = 1$

Hence $\phi P \in \mathscr{D}(S)$ as claimed

Fact. If $g \in L_1(S)$, then $\|gP\| \leqslant \|g\|$

Proof: For $g \in L_1(S)$, we have

$$
\begin{aligned}
\|gP\| &= \|(g^+ - g^-)P\| \\
&= \|g^+P - g^-P\| \\
&\leqslant \|g^+P\| + \|g^-P\| \\
&= \|g^+\| + \|g^-\| \\
&= \int g^+(x)\,\mathrm{d}x + \int g^-(x)\,\mathrm{d}x \\
&= \int |g(x)|\,\mathrm{d}x = \|g\|
\end{aligned}
$$

Fact. If P is a Markov operator on S, then

$$\|gP - hP\| \leqslant \|g - h\| \quad \text{for all} \quad g, h \in L_1(S)$$

We say that P is **nonexpansive** on $L_1(S)$

Proof:

$$\|gP - hP\| = \|(g-h)P\| \leqslant \|(g-h)\| = \|g - h\|$$

One implication is that P is continuous on $L_1(S)$

Indeed, if $g_n \to g$ in $L_1(S)$, then

$$\|g_n P - gP\| \leqslant \|g_n - g\| \to 0$$

Markov Operator Representation

Given stochastic kernel p on S, let $P\colon L_1(S) \to L_1(S)$ be defined by

$$(gP)(y) = \int p(x,y)g(x)\,\mathrm{d}x \qquad (y \in S)$$

Fact. P is a Markov operator on $L_1(S)$

For example, if $g \in L_1^+(S)$, then

$$\|gP\| = \int\int p(x,y)g(x)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int\int p(x,y)\,\mathrm{d}y\, g(x)\,\mathrm{d}x = \int g(x)\,\mathrm{d}x = \|g\|$$

Linking Marginals

By the definition of conditional densities, we always have

$$p_Y(y) = \int p_{Y|X}(y \,|\, x) p_X(x) \, \mathrm{d}x$$

Letting ψ_t be the distribution of X_t, this becomes

$$\psi_{t+1}(y) = \int p(x, y) \psi_t(x) \, \mathrm{d}x \qquad (y \in S)$$

If P is the Markov operator induced by p, we can write this as

$$\psi_{t+1} = \psi_t P$$

Thus the Markov operator **updates the distribution** of the state

Stationary Distributions

Let p be a stochastic kernel on S

If $\psi^* \in \mathscr{D}(S)$ satisfies

$$\psi^*(y) = \int p(x,y)\psi^*(x)\,\mathrm{d}x \quad \text{for all} \quad y \in S$$

then ψ^* is called **stationary** or **invariant** for p

Equivalent: ψ^* is a fixed point of the induced Markov operator

Interpretation:

$$X_t \sim \psi^* \implies X_{t+1} \sim \psi^*$$

Stability

A stochastic kernel on S is called **globally stable** if the corresponding Markov operator is globally stable on $\mathscr{D}(S)$

Fact. If P is a Markov operator and P^k is globally stable on $\mathscr{D}(S)$ for some $k \in \mathbb{N}$, then P is also globally stable on $\mathscr{D}(S)$

Ex. Prove it (or see Stachurski EDTC (2009), lemma 4.1.21)

We will use the following strategy to obtain stability conditions

- Give conditions under which P is strictly contracting
- Combine this with some kind of compactness

Strict Contraction Property

Theorem. If $p(x,y) > 0$ for all $x, y \in S$, then the corresponding Markov operator P is a strict contraction on $(\mathscr{D}(S), \| \cdot \|)$

That is,

$$\|\phi P - \psi P\| < \|\phi - \psi\| \quad \text{whenever} \quad \psi \neq \phi$$

An immediate implication is uniqueness:

If ϕ, ψ are distinct fixed points of P, then

$$\|\phi P - \psi P\| = \|\phi - \psi\|$$

Contradiction

Proof of the theorem: Under the stated conditions, if $\phi \neq \psi$, then

$$
\begin{aligned}
\|\phi P - \psi P\| &= \int \left| \int p(x,y)\phi(x)\,\mathrm{d}x - \int p(x,y)\psi(x)\,\mathrm{d}x \right| \mathrm{d}y \\
&= \int \left| \int p(x,y)[\phi(x) - \psi(x)]\,\mathrm{d}x \right| \mathrm{d}y \\
&< \int \int |p(x,y)[\phi(x) - \psi(x)]|\,\mathrm{d}x\,\mathrm{d}y \\
&= \int \int p(x,y)\,\mathrm{d}y |\phi(x) - \psi(x)|\,\mathrm{d}x \\
&= \|\phi - \psi\|
\end{aligned}
$$

Example. Recall the ARCH(1) model

$$X_{t+1} = (\alpha_0 + \alpha_1 X_t^2)^{1/2} \xi_{t+1} \qquad \{\xi_t\} \overset{\text{IID}}{\sim} N(0,1)$$

with SK

$$p(x,y) = \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 x^2)}} \exp\left\{ -\frac{y^2}{2(\alpha_0 + \alpha_1 x^2)} \right\}$$

Here the state space is $S = \mathbb{R}$ and $\alpha_0 > 0$, $\alpha_1 \geqslant 0$

Evidently $p(x,y) > 0$ for all $x, y \in S$

The Issue of Compactness

In the finite case we used the following theorem

Theorem. If (M, ρ) is compact and $T \colon M \to M$ is a strict contraction, then T is globally stable on (M, ρ)

But when S is infinite, $\mathscr{D}(S)$ is not compact in $L_1(S)$

Example. If $S = \mathbb{R}$, then $\phi_k := N(k, 1)$ has no convergent subsequence in $\mathscr{D}(S)$

However, recall it suffices that T is **Lagrange stable** on M

Meaning: $\{T^k x\}_{k \in \mathbb{N}}$ is precompact for all $x \in M$

So when is $\{\psi P^k\}$ precompact in $\mathscr{D}(S)$?

Drift and Compactness

Let $W \colon S \to \mathbb{R}_+$ be a given function and let

$$C_\alpha := \{x \in S : W(x) \leqslant \alpha\}$$

W is called **coercive** on S if C_α is compact for every $\alpha \geqslant 0$

Example. $W(x) = \sqrt{\langle x, x \rangle}$ is coercive on $\mathbb{R}^n$ because C_α is a closed ball about the origin

Example. $W(x) = x^2$ is coercive on $\mathbb{R}$ because $C_\alpha = [-\sqrt{\alpha}, \sqrt{\alpha}]$

Let p be a stochastic kernel on S and let P be its Markov operator

Theorem. If there exists a continuous function $h \colon S \to \mathbb{R}$, a coercive function W on S, and constants $\gamma \in (0,1)$ and $L \geqslant 0$ such that

1. $p(x,y) \leqslant h(y)$ for all $(x,y) \in S \times S$
2. p satisfies

$$\int W(y) p(x,y) \, \mathrm{d}y \leqslant \gamma W(x) + L \qquad (x \in S)$$

then P is Lagrange stable

Proof: See §8.2.3 of Stachurski (2009)

Example. Recall the ARCH(1) model on $S = \mathbb{R}$ with SK

$$p(x, y) = \frac{1}{\sqrt{2\pi(\alpha_0 + \alpha_1 x^2)}} \exp\left\{-\frac{y^2}{2(\alpha_0 + \alpha_1 x^2)}\right\}$$

By assumption, $\alpha_0 > 0$, $\alpha_1 \geqslant 0$

Evidently $\exists\, M < \infty$ such that

$$p(x, y) \leqslant M \qquad \forall\, x, y \in S$$

Setting $h(y) = M$ establishes part 1

Can we find coercive W and $\gamma \in (0,1)$, $L \geqslant 0$ such that

$$\int W(y)p(x,y)\,\mathrm{d}y \leqslant \gamma W(x) + L\,?$$

Letting $W(x) = x^2$, we have

$$\int W(y)p(x,y)\,\mathrm{d}y = \mathbb{E}\,W[(\alpha_0 + \alpha_1 x^2)^{1/2}\xi_{t+1}]$$

$$= (\alpha_0 + \alpha_1 x^2)\mathbb{E}\,[\xi_{t+1}^2] = \alpha_0 + \alpha_1 W(x)$$

Since W is coercive on $\mathbb{R}$, it suffices that $\alpha_1 < 1$

Under this condition, the ARCH(1) model is Lagrange stable (and also strongly contracting, and hence globally stable)

LLN

Let $h \in m\mathscr{B}$ and let $\{X_t\}$ be Markov with stochastic kernel p

Theorem. If P is globally stable with stationary density ψ^*, then

$$\mathbb{P}\left\{ \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} h(X_t) = \int h(x) \psi^*(x) \, \mathrm{d}x \right\} = 1$$

Example. If $h(x) = \mathbb{1}_B(x)$ for some $B \in \mathscr{B}(S)$, then

$$\frac{1}{n} \sum_{t=1}^{n} \mathbb{1}_B(X_t) \to \int \mathbb{1}_B(x) \psi^*(x) \, \mathrm{d}x = \int_B \psi^*(x) \, \mathrm{d}x$$

Hence

$$\int_B \psi^*(x) \, \mathrm{d}x \approx \text{ fraction of time that } \{X_t\} \text{ spends in set } B$$

Application: Look Ahead Estimation

Suppose we have a globally stable model

$$X_{t+1} = F(X_t, \xi_{t+1}), \qquad \{\xi_t\} \overset{\text{IID}}{\sim} \phi$$

with stochastic kernel p

Let ψ^* be the stationary density, suppose we want to compute it

Bad option: Compute a histogram

Good option: Use the **look ahead estimator** ψ_n^* defined by

$$\psi_n^*(y) = \frac{1}{n} \sum_{t=1}^{n} p(X_t, y) \qquad (y \in S)$$

By the LLN, we have

$$\psi_n^*(y) = \frac{1}{n} \sum_{t=1}^{n} p(X_t, y)$$

$$\to \int p(x,y) \psi^*(x) \, \mathrm{d}x = \psi^*(y)$$

In fact one can show that

$$\mathbb{P} \left\{ \lim_{n \to \infty} \| \psi_n^* - \psi^* \| = 0 \right\} = 1$$

Hence globally convergent with probability one

Good finite sample properties can also be established