Introduction to the Stochastic Growth Model

John Stachurski
Institute of Economic Research, Kyoto University
john@kier.kyoto-u.ac.jp

September 21, 2002

Structure of the Seminar

⋆ Day 1

1. Introduction to Discrete Stochastic Processes
2. Deterministic Dynamics
3. Stochastic Dynamics via Markov Operators

⋆ Day 2

4. Stability of Markov Processes
5. Applications
6. Empirics of Stochastic Growth

Back to top

Part 1: Discrete Stochastic Processes

- ⋆ Choice under uncertainty
- ⋆ Finite horizon control
- ⋆ Extending to the infinite horizon
- ⋆ Introduction to stochastic growth

Back to top

Choice under uncertainty

Let S be a set of outcomes.

Choice under uncertainty

Let S be a set of outcomes.

Outcomes are random. A typical distribution over S is F.

Choice under uncertainty

Let S be a set of outcomes.

Outcomes are random. A typical distribution over S is F.

Let $\mathscr{P}_S$ be the space of all distributions on S.

Choice under uncertainty

Let S be a set of outcomes.

Outcomes are random. A typical distribution over S is F.

Let $\mathscr{P}_S$ be the space of all distributions on S.

We cannot choose outcomes. Rather, our actions influence *probabilities* of outcomes.

Choice under uncertainty

Let S be a set of outcomes.

Outcomes are random. A typical distribution over S is F.

Let $\mathscr{P}_S$ be the space of all distributions on S.

We cannot choose outcomes. Rather, our actions influence *probabilities* of outcomes.

According to this we can rank actions.

A typical example of choice:

A typical example of choice:

Choose action g.

A typical example of choice:

Choose action g.

This affects probabilities of outcomes: selects $F_g \in \mathscr{P}_S$.

A typical example of choice:

Choose action g.

This affects probabilities of outcomes: selects $F_g \in \mathscr{P}_S$.

Assess value of action as

$$U(g) = \mathbb{E}(v_g | F_g) \equiv \int_S v_g(s) F_g(ds), \tag{1}$$

where v_g is a function from S to $\mathbb{R}$.

A typical example of choice:

Choose action g.

This affects probabilities of outcomes: selects $F_g \in \mathscr{P}_S$.

Assess value of action as

$$U(g) = \mathbb{E}(v_g|F_g) \equiv \int_S v_g(s)F_g(ds), \tag{1}$$

where v_g is a function from S to $\mathbb{R}$.

This induces an order on space of actions G.

Back to top

Finite horizon control

State variable is $x \in X$. Control is $u \in U$.

Finite horizon control

State variable is $x \in X$. Control is $u \in U$.

After observing x_t, controller chooses u_t, and then update by

$$x_{t+1} = h(x_t, u_t, \varepsilon_t).$$

Finite horizon control

State variable is $x \in X$. Control is $u \in U$.

After observing x_t, controller chooses u_t, and then update by

$$x_{t+1} = h(x_t, u_t, \varepsilon_t). \tag{2}$$

The process then repeats.

Finite horizon control

State variable is $x \in X$. Control is $u \in U$.

After observing x_t, controller chooses u_t, and then update by

$$x_{t+1} = h(x_t, u_t, \varepsilon_t). \tag{2}$$

The process then repeats.

Shocks ε_t are i.i.d.

Finite horizon control

State variable is $x \in X$. Control is $u \in U$.

After observing x_t, controller chooses u_t, and then update by

$$x_{t+1} = h(x_t, u_t, \varepsilon_t). \tag{2}$$

The process then repeats.

Shocks ε_t are i.i.d.

Total reward is

$$\mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, u_t) + w(x_T)\right]. \tag{3}$$

Back to top

The general control problem is to find a sequence of feasible control policies $g_t\colon x_t \mapsto u_t$ to solve

$$\max_{(g_t)} \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, u_t) + w(x_T)\right] \tag{4}$$

The general control problem is to find a sequence of feasible control policies $g_t \colon x_t \mapsto u_t$ to solve

$$\max_{(g_t)} \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, u_t) + w(x_T)\right] \tag{4}$$

subject to

$$x_{t+1} = h(x_t, u_t, \varepsilon_t), \quad t = 0, \ldots, T-1. \tag{5}$$

The general control problem is to find a sequence of feasible control policies $g_t \colon x_t \mapsto u_t$ to solve

$$\max_{(g_t)} \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, u_t) + w(x_T)\right] \tag{4}$$

subject to

$$x_{t+1} = h(x_t, u_t, \varepsilon_t), \quad t = 0, \ldots, T-1. \tag{5}$$

and

$$x_0 \sim \varphi_0. \tag{6}$$

The general control problem is to find a sequence of feasible control policies $g_t \colon x_t \mapsto u_t$ to solve

$$\max_{(g_t)} \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, u_t) + w(x_T)\right] \tag{4}$$

subject to

$$x_{t+1} = h(x_t, u_t, \varepsilon_t), \quad t = 0, \ldots, T-1. \tag{5}$$

and

$$x_0 \sim \varphi_0. \tag{6}$$

What is the relation to theory of choice under uncertainty?

The general control problem is to find a sequence of feasible control policies $g_t \colon x_t \mapsto u_t$ to solve

$$\max_{(g_t)} \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, u_t) + w(x_T)\right] \tag{4}$$

subject to

$$x_{t+1} = h(x_t, u_t, \varepsilon_t), \quad t = 0, \ldots, T-1. \tag{5}$$

and

$$x_0 \sim \varphi_0. \tag{6}$$

What is the relation to theory of choice under uncertainty? What is the meaning of $\mathbb{E}$ here?

Back to top

A two period example:

A two period example:

$$\max_g \mathbb{E}\left[\sum_{t=0}^{0} \beta^t r(x_t, g(x_t)) + w(x_1)\right] \tag{7}$$

subject to

$$x_1 = h(x_0, g(x_0), \varepsilon_0), \quad x_0 \sim \varphi_0. \tag{8}$$

A two period example:

$$\max_g \mathbb{E}\left[\sum_{t=0}^{0} \beta^t r(x_t, g(x_t)) + w(x_1)\right] \tag{7}$$

subject to

$$x_1 = h(x_0, g(x_0), \varepsilon_0), \quad x_0 \sim \varphi_0. \tag{8}$$

Since h and the distribution of ε_0 is known, for each g we can calculate the conditional distribution of x_1 given x_0.

A two period example:

$$\max_g \mathbb{E}\left[\sum_{t=0}^{0} \beta^t r(x_t, g(x_t)) + w(x_1)\right] \tag{7}$$

subject to

$$x_1 = h(x_0, g(x_0), \varepsilon_0), \quad x_0 \sim \varphi_0. \tag{8}$$

Since h and the distribution of ε_0 is known, for each g we can calculate the conditional distribution of x_1 given x_0. Call it $f_g(x_1|x_0)$.

Back to top

What is the joint distribution $F_g(x_0, x_1)$ of x_0 and x_1?

What is the joint distribution $F_g(x_0, x_1)$ of x_0 and x_1?

We have

$$f_g(x_1|x_0) = \frac{F_g(x_0, x_1)}{\varphi_0(x_0)}$$

What is the joint distribution $F_g(x_0, x_1)$ of x_0 and x_1?

We have

$$f_g(x_1|x_0) = \frac{F_g(x_0, x_1)}{\varphi_0(x_0)} \implies F_g(x_1, x_0) = f_g(x_0|x_1)\varphi_0(x_0).$$

What is the joint distribution $F_g(x_0, x_1)$ of x_0 and x_1?

We have

$$f_g(x_1|x_0) = \frac{F_g(x_0, x_1)}{\varphi_0(x_0)} \implies F_g(x_1, x_0) = f_g(x_0|x_1)\varphi_0(x_0).$$

Let $S = X \times X$.

What is the joint distribution $F_g(x_0, x_1)$ of x_0 and x_1?

We have

$$f_g(x_1|x_0) = \frac{F_g(x_0, x_1)}{\varphi_0(x_0)} \implies F_g(x_1, x_0) = f_g(x_0|x_1)\varphi_0(x_0).$$

Let $S = X \times X$.

Then $F_g \in \mathscr{P}_S$ for all g.

Also, let v_g be the real function on $S = X \times X$ defined by

$$v_g(x_0, x_1) = \sum_{t=0}^{0} \beta^t r(x_0, g(x_0)) + w(x_1) \tag{9}$$

Also, let v_g be the real function on $S = X \times X$ defined by

$$v_g(x_0, x_1) = \sum_{t=0}^{0} \beta^t r(x_0, g(x_0)) + w(x_1) \tag{9}$$

Problem is then

$$\max_g \mathbb{E}(v_g | F_g) = \max_g \int_S v_g(s) F_g(ds). \tag{10}$$

Same idea for any $T \in \mathbb{N}$.

Same idea for any $T \in \mathbb{N}$.

Each $(g_0, \ldots, g_{T-1})$ determines the conditional distributions via

$$x_{t+1} = h(x_t, g_t(x_t), \varepsilon_t) \implies f_{g_t}(x'|x). \tag{11}$$

Same idea for any $T \in \mathbb{N}$.

Each $(g_0, \ldots, g_{T-1})$ determines the conditional distributions via

$$x_{t+1} = h(x_t, g_t(x_t), \varepsilon_t) \implies f_{g_t}(x'|x). \tag{11}$$

From these we can build the joint distribution F_g over $x_0, \ldots, x_T$.

Same idea for any $T \in \mathbb{N}$.

Each $(g_0, \ldots, g_{T-1})$ determines the conditional distributions via

$$x_{t+1} = h(x_t, g_t(x_t), \varepsilon_t) \implies f_{g_t}(x'|x). \tag{11}$$

From these we can build the joint distribution F_g over $x_0, \ldots, x_T$.

The set of outcomes S is X^T, and $F_g \in \mathscr{P}_S$.

Same idea for any $T \in \mathbb{N}$.

Each $(g_0, \ldots, g_{T-1})$ determines the conditional distributions via

$$x_{t+1} = h(x_t, g_t(x_t), \varepsilon_t) \implies f_{g_t}(x'|x). \tag{11}$$

From these we can build the joint distribution F_g over $x_0, \ldots, x_T$.

The set of outcomes S is X^T, and $F_g \in \mathscr{P}_S$.

This is the meaning of

$$\max_g \mathbb{E}\left[\sum_{t=0}^{T-1} \beta^t r(x_t, g(x_t)) + w(x_T)\right] = \int_S v_g(s) F_g(ds). \tag{12}$$

Back to top

Extending to the infinite horizon

Same principle for $T = \infty$,

Extending to the infinite horizon

Same principle for $T = \infty$,

$$\max \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t r(x_t, u_t)\right] \tag{13}$$

subject to

$$x_{t+1} = h(x_t, u_t, \varepsilon_t). \tag{14}$$

Extending to the infinite horizon

Same principle for $T = \infty$,

$$\max \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t r(x_t, u_t) \right] \tag{13}$$

subject to

$$x_{t+1} = h(x_t, u_t, \varepsilon_t). \tag{14}$$

The problem is stationary, so choose just one g.

Back to top

Each g determines

$$x_{t+1} = h(x_t, g(x_t), \varepsilon_t), \tag{15}$$

Each g determines

$$x_{t+1} = h(x_t, g(x_t), \varepsilon_t), \tag{15}$$

which implies a stationary conditional distribution $f_g(x'|x)$.

Each g determines

$$x_{t+1} = h(x_t, g(x_t), \varepsilon_t), \tag{15}$$

which implies a stationary conditional distribution $f_g(x'|x)$.

Determines a unique joint distribution F_g over $(x_t)_{t=0}^{\infty}$.

Each g determines

$$x_{t+1} = h(x_t, g(x_t), \varepsilon_t), \tag{15}$$

which implies a stationary conditional distribution $f_g(x'|x)$.

Determines a unique joint distribution F_g over $(x_t)_{t=0}^{\infty}$.

Now $S = X^{\mathbb{N}}$,

Each g determines

$$x_{t+1} = h(x_t, g(x_t), \varepsilon_t), \tag{15}$$

which implies a stationary conditional distribution $f_g(x'|x)$.

Determines a unique joint distribution F_g over $(x_t)_{t=0}^{\infty}$.

Now $S = X^{\mathbb{N}}$, and $F_g \in \mathscr{P}_S$.

Each g determines

$$x_{t+1} = h(x_t, g(x_t), \varepsilon_t), \tag{15}$$

which implies a stationary conditional distribution $f_g(x'|x)$.

Determines a unique joint distribution F_g over $(x_t)_{t=0}^{\infty}$.

Now $S = X^{\mathbb{N}}$, and $F_g \in \mathscr{P}_S$.

Solve

$$\max_g \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t r(x_t, g(x_t))\right] = \int_S v_g(s) F_g(ds). \tag{16}$$

Introduction to stochastic growth

Modern deterministic growth theory begins with Solow (QJE, 1956), Swan (Economic Record, 1956).

Introduction to stochastic growth

Modern deterministic growth theory begins with Solow (QJE, 1956), Swan (Economic Record, 1956).

Optimal savings case treated by Cass (RE Stud., 1965) and Koopman (Pontif. Acad. Sci., 1965)

Introduction to stochastic growth

Modern deterministic growth theory begins with Solow (QJE, 1956), Swan (Economic Record, 1956).

Optimal savings case treated by Cass (RE Stud., 1965) and Koopman (Pontif. Acad. Sci., 1965)

Stochastic optimal savings case first treated by Brock and Mirman (JET, 1972).

Introduction to stochastic growth

Modern deterministic growth theory begins with Solow (QJE, 1956), Swan (Economic Record, 1956).

Optimal savings case treated by Cass (RE Stud., 1965) and Koopman (Pontif. Acad. Sci., 1965)

Stochastic optimal savings case first treated by Brock and Mirman (JET, 1972).

Making suitable analogous definitions, they show that all of the results of the deterministic case carry over.

Back to top

Basic model with representative agent:

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

The agent divides x_0 between consumption c_0 and savings.

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

The agent divides x_0 between consumption c_0 and savings. Consumption gives utility $u(c_0)$.

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

The agent divides x_0 between consumption c_0 and savings. Consumption gives utility $u(c_0)$. Consumption is the control.

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

The agent divides x_0 between consumption c_0 and savings. Consumption gives utility $u(c_0)$. Consumption is the control.

Savings equals both investment and capital stock $x_0 - c_0$ (depreciation is total).

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

The agent divides x_0 between consumption c_0 and savings. Consumption gives utility $u(c_0)$. Consumption is the control.

Savings equals both investment and capital stock $x_0 - c_0$ (depreciation is total).

Through investing $x_0 - c_0$, agent receives at the start of next period income $f(x_0 - c_0, \varepsilon_0)$.

Basic model with representative agent:

At time 0 income x_0 is drawn from φ_0 by nature and given to the agent.

The agent divides x_0 between consumption c_0 and savings. Consumption gives utility $u(c_0)$. Consumption is the control.

Savings equals both investment and capital stock $x_0 - c_0$ (depreciation is total).

Through investing $x_0 - c_0$, agent receives at the start of next period income $f(x_0 - c_0, \varepsilon_0)$.

The process now repeats.

The problem is to choose a consumption policy $g\colon x \longrightarrow c$ to maximize

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(g(x_t))\right] \tag{17}$$

The problem is to choose a consumption policy $g\colon x \longrightarrow c$ to maximize

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(g(x_t))\right] \tag{17}$$

subject to

$$x_{t+1} = f(x_t - g(x_t), \varepsilon_t). \tag{18}$$

The problem is to choose a consumption policy $g\colon x \to c$ to maximize

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(g(x_t))\right] \tag{17}$$

subject to

$$x_{t+1} = f(x_t - g(x_t), \varepsilon_t). \tag{18}$$

Each g selects a joint distribution F_g over $\mathbb{R}^{\mathbb{N}}$. Let v_g be the function $\sum_t \beta^t u(g(x_t))$ from $\mathbb{R}^{\mathbb{N}}$ to $\mathbb{R}$. Solve

$$\max_g \int_S v_g(s) F_g(ds). \tag{19}$$

Back to top

In practice, characterization of the optimal policy is as follows.

In practice, characterization of the optimal policy is as follows.

Let f, u be smooth, concave and have all other nice properties.

In practice, characterization of the optimal policy is as follows.

Let f, u be smooth, concave and have all other nice properties.

Theorem (Mirman–Zilcha). An optimal consumption policy g exists and is unique. It satisfies

$$u'(g(x)) = \beta \int u'[g(f(x - g(x))z)] f'(x - g(x)) z \psi(dz) \tag{20}$$

for all $x \geq 0$.

A simple example (Stokey, Lucas and Prescott, Sec. 2.2):

A simple example (Stokey, Lucas and Prescott, Sec. 2.2):

Let $u(c) = \ln c$, $f(k) = k^{\alpha}$, $\alpha < 1$,

A simple example (Stokey, Lucas and Prescott, Sec. 2.2):

Let $u(c) = \ln c$, $f(k) = k^{\alpha}$, $\alpha < 1$, ε lognormal with $\ln \varepsilon \sim N(0, 1)$.

A simple example (Stokey, Lucas and Prescott, Sec. 2.2):

Let $u(c) = \ln c$, $f(k) = k^\alpha$, $\alpha < 1$, ε lognormal with $\ln \varepsilon \sim N(0, 1)$.

In this case we can deduce that $g(x) = (1 - s)x$ for some constant s,

A simple example (Stokey, Lucas and Prescott, Sec. 2.2):

Let $u(c) = \ln c$, $f(k) = k^{\alpha}$, $\alpha < 1$, ε lognormal with $\ln \varepsilon \sim N(0, 1)$.

In this case we can deduce that $g(x) = (1 - s)x$ for some constant s, in which case

$$x_{t+1} = (sx_t)^{\alpha} \varepsilon_t = C x_t^{\alpha} \varepsilon_t. \tag{21}$$

A simple example (Stokey, Lucas and Prescott, Sec. 2.2):

Let $u(c) = \ln c$, $f(k) = k^\alpha$, $\alpha < 1$, ε lognormal with $\ln \varepsilon \sim N(0,1)$.

In this case we can deduce that $g(x) = (1-s)x$ for some constant s, in which case

$$x_{t+1} = (sx_t)^\alpha \varepsilon_t = C x_t^\alpha \varepsilon_t. \tag{21}$$

Taking logs gives the linear-Gaussian AR(1) system

$$\tilde{x}_{t+1} = a\tilde{x}_t + b + \tilde{\varepsilon}_t, \quad a < 1. \tag{22}$$

Back to top

Part 2: Deterministic Dynamics

- ⋆ Semidynamical systems
- ⋆ Stability and equilibrium
- ⋆ Brouwer-Schauder fixed point theorem
- ⋆ Banach contraction theorem
- ⋆ Contractions and compactness
- ⋆ Lagrange stability and contractions

Back to top

Semidynamical systems

Dynamic system evolves in a space X

Semidynamical systems

Dynamic system evolves in a space X

Time is discrete: $x_0 \in X$, $x_1 \in X$,...

Semidynamical systems

Dynamic system evolves in a space X

Time is discrete: $x_0 \in X$, $x_1 \in X$,...

The system is autonomous (a difference equation):

$$x_{t+1} = g(x_t), \quad x_0 \in X, \quad g\colon X \to X. \tag{23}$$

Semidynamical systems

Dynamic system evolves in a space X

Time is discrete: $x_0 \in X$, $x_1 \in X$,...

The system is autonomous (a difference equation):

$$x_{t+1} = g(x_t), \quad x_0 \in X, \quad g\colon X \to X. \tag{23}$$

Thus $x_1 = g(x_0)$, $x_2 = g(g(x_0)) \equiv g^2(x_0)$,..., $x_t = g^t(x_0)$.

Back to top

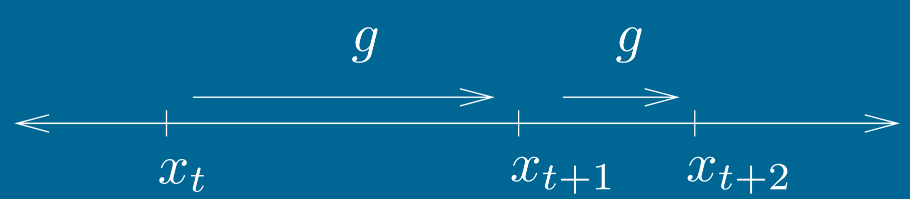
Definition. A *semidynamical system* is a space X and a map g sending X into X. Write (X, g).

Definition. A *semidynamical system* is a space X and a map g sending X into X. Write (X, g).

Definition. The *trajectory* of $x \in X$ under g is the sequence $\{g^t(x)\}_{t \geq 0}$.

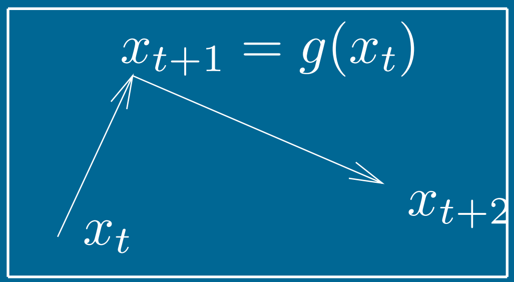
Definition. A *semidynamical system* is a space X and a map g sending X into X. Write (X, g).

Definition. The *trajectory* of $x \in X$ under g is the sequence $\{g^t(x)\}_{t \geq 0}$.



Definition. A *semidynamical system* is a space X and a map g sending X into X. Write (X, g).

Definition. The *trajectory* of $x \in X$ under g is the sequence $\{g^t(x)\}_{t \geq 0}$.



Back to top

Stability and equilibrium

Let a semidynamical system (X, g) be given.

Stability and equilibrium

Let a semidynamical system (X, g) be given.

Definition. A *steady state* or *equilibrium* for (X, g) is a fixed point of g on X. That is, $g(x^*) = x^*$.

Stability and equilibrium

Let a semidynamical system (X, g) be given.

Definition. A *steady state* or *equilibrium* for (X, g) is a fixed point of g on X. That is, $g(x^*) = x^*$.

Definition. For fixed point x^* of g on X, the *stable set* $S_g(x^*)$ of x^* is all points $x \in X$ such that the trajectory of x converges to x^*.

Stability and equilibrium

Let a semidynamical system (X, g) be given.

Definition. A *steady state* or *equilibrium* for (X, g) is a fixed point of g on X. That is, $g(x^*) = x^*$.

Definition. For fixed point x^* of g on X, the *stable set* $S_g(x^*)$ of x^* is all points $x \in X$ such that the trajectory of x converges to x^*.

Definition. Equilibrium x^* is called *stable* if $S_g(x^*)$ contains an open neighborhood of x^*.

Stability and equilibrium

Let a semidynamical system (X, g) be given.

Definition. A *steady state* or *equilibrium* for (X, g) is a fixed point of g on X. That is, $g(x^*) = x^*$.

Definition. For fixed point x^* of g on X, the *stable set* $S_g(x^*)$ of x^* is all points $x \in X$ such that the trajectory of x converges to x^*.

Definition. Equilibrium x^* is called *stable* if $S_g(x^*)$ contains an open neighborhood of x^*.

Definition. It is *globally stable* if $S_g(x^*) = X$.

Back to top

Brouwer-Schauder fixed point theorem

Theorem. Let X be a subset of a normed linear space. If X is compact and convex, and, in addition, g is continuous on X, then (X, g) has at least one equilibrium.

Back to top

Banach contraction theorem

Definition. Let (X, g) be a semidynamical system, with X a subset of a normed linear space. Map g is called a *Banach contraction* if there exists a $\lambda < 1$ such that

$$\|g(x) - g(y)\| \leq \lambda \|x - y\|. \tag{24}$$

Banach contraction theorem

Definition. Let (X, g) be a semidynamical system, with X a subset of a normed linear space. Map g is called a *Banach contraction* if there exists a $\lambda < 1$ such that

$$\|g(x) - g(y)\| \le \lambda \|x - y\|. \tag{24}$$

Note that along the trajectory $\{g^t(x)\}$ from x,

$$\|g^{t+1}(x) - g^t(x)\| \le \lambda \|g^t(x) - g^{t-1}(x)\|$$

Banach contraction theorem

Definition. Let (X, g) be a semidynamical system, with X a subset of a normed linear space. Map g is called a *Banach contraction* if there exists a $\lambda < 1$ such that

$$\|g(x) - g(y)\| \leq \lambda \|x - y\|. \quad (24)$$

Note that along the trajectory $\{g^t(x)\}$ from x,

$$\begin{aligned} \|g^{t+1}(x) - g^t(x)\| &\leq \lambda \|g^t(x) - g^{t-1}(x)\| \\ &\leq \lambda^2 \|g^{t-1}(x) - g^{t-2}(x)\| \end{aligned}$$

Banach contraction theorem

Definition. Let (X, g) be a semidynamical system, with X a subset of a normed linear space. Map g is called a *Banach contraction* if there exists a $\lambda < 1$ such that

$$\|g(x) - g(y)\| \leq \lambda \|x - y\|. \tag{24}$$

Note that along the trajectory $\{g^t(x)\}$ from x,

$$\begin{aligned} \|g^{t+1}(x) - g^t(x)\| &\leq \lambda \|g^t(x) - g^{t-1}(x)\| \\ &\leq \lambda^2 \|g^{t-1}(x) - g^{t-2}(x)\| \\ &\leq \lambda^t \|g(x) - x\|. \end{aligned}$$

Back to top

Theorem. If X is complete and g a Banach contraction, then (X, g) has a unique and globally stable equilibrium.

Theorem. If X is complete and g a Banach contraction, then (X, g) has a unique and globally stable equilibrium.

Proof:

Theorem. If X is complete and g a Banach contraction, then (X, g) has a unique and globally stable equilibrium.

Proof:

The sequence is Cauchy and hence convergent to some x^*, because of

$$\|g^{t+1}(x) - g^t(x)\| \leq \lambda^t \|g(x) - x\|.$$

Theorem. If X is complete and g a Banach contraction, then (X, g) has a unique and globally stable equilibrium.

Proof:

The sequence is Cauchy and hence convergent to some x^*, because of

$$\|g^{t+1}(x) - g^t(x)\| \leq \lambda^t \|g(x) - x\|.$$

The point x^* is a fixed point.

Theorem. If X is complete and g a Banach contraction, then (X, g) has a unique and globally stable equilibrium.

Proof:

The sequence is Cauchy and hence convergent to some x^*, because of

$$\|g^{t+1}(x) - g^t(x)\| \leq \lambda^t \|g(x) - x\|.$$

The point x^* is a fixed point.

The fixed point is unique.

Back to top

Contractions and compactness

Definition. Call g *strongly contracting* on X if

$$\|g(x) - g(y)\| < \|x - y\|, \quad \forall x \neq y \tag{25}$$

Contractions and compactness

Definition. Call g *strongly contracting* on X if

$$\|g(x) - g(y)\| < \|x - y\|, \quad \forall x \neq y \tag{25}$$

which is weaker than

$$\|g(x) - g(y)\| \leq \lambda \|x - y\|, \quad \lambda < 1. \tag{26}$$

Contractions and compactness

Definition. Call g *strongly contracting* on X if

$$\|g(x) - g(y)\| < \|x - y\|, \quad \forall x \neq y \tag{25}$$

which is weaker than

$$\|g(x) - g(y)\| \leq \lambda \|x - y\|, \quad \lambda < 1. \tag{26}$$

Note that strongly contracting does not imply existence of f.p. (e.g. speed is $1 + 1/n$).

Back to top

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$,

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Existence:

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Existence: Define $r\colon X \to \mathbb{R}$ by $r(x) = \|g(x) - x\|$.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Existence: Define $r \colon X \to \mathbb{R}$ by $r(x) = \|g(x) - x\|$. Evidently r continuous.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Existence: Define $r \colon X \to \mathbb{R}$ by $r(x) = \|g(x) - x\|$. Evidently r continuous. Since X is compact, r has a minimizer x^*.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Existence: Define $r\colon X \to \mathbb{R}$ by $r(x) = \|g(x) - x\|$. Evidently r continuous. Since X is compact, r has a minimizer x^*. But then $g(x^*) = x^*$, because

$$g(x^*) \neq x^* \implies r(g(x^*)) = \|g^2(x^*) - g(x^*)\| < \|g(x^*) - x^*\| = r(x^*),$$

in which case $g(x^*)$ is the minimizer.

Theorem. Let (X, g) be strongly contracting. If X is compact, then the system has a unique and globally stable equilibrium.

Uniqueness: Let $x \neq y$ both be equilibria. Then $\|g(x) - g(y)\| = \|x - y\|$, and also $\|g(x) - g(y)\| < \|x - y\|$. Contradiction.

Existence: Define $r \colon X \to \mathbb{R}$ by $r(x) = \|g(x) - x\|$. Evidently r continuous. Since X is compact, r has a minimizer x^*. But then $g(x^*) = x^*$, because

$$g(x^*) \neq x^* \implies r(g(x^*)) = \|g^2(x^*) - g(x^*)\| < \|g(x^*) - x^*\| = r(x^*),$$

in which case $g(x^*)$ is the minimizer.

Convergence: Not proved, but true.

Back to top

Lagrange stability and contractions

What if X is not compact?

Lagrange stability and contractions

What if X is not compact?

Definition. Semidynamical system (X, g) is called *Lagrange stable* if the trajectory of x is precompact for every $x \in X$.

Lagrange stability and contractions

What if X is not compact?

Definition. Semidynamical system (X, g) is called *Lagrange stable* if the trajectory of x is precompact for every $x \in X$.

Example: $x_{t+1} = ax_t$ depends on $a \leq 1$ or $a > 1$.

Lagrange stability and contractions

What if X is not compact?

Definition. Semidynamical system (X, g) is called *Lagrange stable* if the trajectory of x is precompact for every $x \in X$.

Example: $x_{t+1} = ax_t$ depends on $a \leq 1$ or $a > 1$.

Theorem. If semidynamical system (X, g) is both Lagrange stable and strongly contracting, then it has a unique and globally stable equilibrium.

Lagrange stability and contractions

What if X is not compact?

Definition. Semidynamical system (X, g) is called *Lagrange stable* if the trajectory of x is precompact for every $x \in X$.

Example: $x_{t+1} = ax_t$ depends on $a \leq 1$ or $a > 1$.

Theorem. If semidynamical system (X, g) is both Lagrange stable and strongly contracting, then it has a unique and globally stable equilibrium.

Note that Lagrange stability substitutes for compactness.

Sketch of proof:

Sketch of proof:

Pick $x \in X$. Let $\gamma(x)$ be the closure of the trajectory.

Sketch of proof:

Pick $x \in X$. Let $\gamma(x)$ be the closure of the trajectory.

Note that $\gamma(x)$ is compact.

Sketch of proof:

Pick $x \in X$. Let $\gamma(x)$ be the closure of the trajectory.

Note that $\gamma(x)$ is compact.

Also, g maps $\gamma(x)$ into $\gamma(x)$.

Sketch of proof:

Pick $x \in X$. Let $\gamma(x)$ be the closure of the trajectory.

Note that $\gamma(x)$ is compact.

Also, g maps $\gamma(x)$ into $\gamma(x)$.

Note that g is contracting on $\gamma(x)$.

Sketch of proof:

Pick $x \in X$. Let $\gamma(x)$ be the closure of the trajectory.

Note that $\gamma(x)$ is compact.

Also, g maps $\gamma(x)$ into $\gamma(x)$.

Note that g is contracting on $\gamma(x)$.

Hence a fixed point exists in $\gamma(x)$.

Part 3: Stochastic Dynamics via Markov Operators

- ⋆ The general Markov operator
- ⋆ Construction from perturbed systems
- ⋆ Equilibrium and stability
- ⋆ Example: the AR(1) model

Back to top

The general Markov operator

We want to study something like this:

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{27}$$

The general Markov operator

We want to study something like this:

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{27}$$

How can we apply techniques for deterministic systems?

The general Markov operator

We want to study something like this:

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{27}$$

How can we apply techniques for deterministic systems?

The method: transform this into deterministic system on an infinite dimensional space called $\mathscr{L}_1$.

As before, X is some space. Transform X into a measure space $(X, \mathscr{B}_X, \mu)$.

As before, X is some space. Transform X into a measure space $(X, \mathscr{B}_X, \mu)$.

In general, X a subset of $\mathbb{R}$, $\mathscr{B}_X$ is the Borel sets, and μ is Lebesgue measure.

As before, X is some space. Transform X into a measure space $(X, \mathscr{B}_X, \mu)$.

In general, X a subset of $\mathbb{R}$, $\mathscr{B}_X$ is the Borel sets, and μ is Lebesgue measure.

Let $\mathscr{L}_1(\mu)$ be all of the integrable functions on X.

As before, X is some space. Transform X into a measure space $(X, \mathscr{B}_X, \mu)$.

In general, X a subset of $\mathbb{R}$, $\mathscr{B}_X$ is the Borel sets, and μ is Lebesgue measure.

Let $\mathscr{L}_1(\mu)$ be all of the integrable functions on X.

The deterministic model is a semidynamical system in X, but the stochastic version is a semidynamical system in $\mathscr{L}_1(\mu)$!

So what is $\mathscr{L}_1(\mu)$ like?

So what is $\mathscr{L}_1(\mu)$ like?

As usual, $\mathscr{L}_1(\mu)$ is a normed linear space (Banach lattice).

So what is $\mathscr{L}_1(\mu)$ like?

As usual, $\mathscr{L}_1(\mu)$ is a normed linear space (Banach lattice).

The norm is $\|f\| \equiv \int_X |f| d\mu = \int |f(x)| dx$.

So what is $\mathscr{L}_1(\mu)$ like?

As usual, $\mathscr{L}_1(\mu)$ is a normed linear space (Banach lattice).

The norm is $\|f\| \equiv \int_X |f| d\mu = \int |f(x)| dx$.

Let $\mathscr{D}(\mu)$ be all the f in $\mathscr{L}_1(\mu)$ such that $f \geq 0$ and integral is 1. What are these functions called...?

So what is $\mathscr{L}_1(\mu)$ like?

As usual, $\mathscr{L}_1(\mu)$ is a normed linear space (Banach lattice).

The norm is $\|f\| \equiv \int_X |f| d\mu = \int |f(x)| dx$.

Let $\mathscr{D}(\mu)$ be all the f in $\mathscr{L}_1(\mu)$ such that $f \geq 0$ and integral is 1. What are these functions called...? The density functions on X!

So what is $\mathscr{L}_1(\mu)$ like?

As usual, $\mathscr{L}_1(\mu)$ is a normed linear space (Banach lattice).

The norm is $\|f\| \equiv \int_X |f| d\mu = \int |f(x)| dx$.

Let $\mathscr{D}(\mu)$ be all the f in $\mathscr{L}_1(\mu)$ such that $f \geq 0$ and integral is 1. What are these functions called...? The density functions on X!

A *Markov operator* is an operator $P \colon \mathscr{L}_1(\mu) \to \mathscr{L}_1(\mu)$ such that

$$f \in \mathscr{D}(\mu) \implies Pf \in \mathscr{D}(\mu). \tag{28}$$

Construction from perturbed systems

Consider the (very common macroeconomic) model

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{29}$$

Construction from perturbed systems

Consider the (very common macroeconomic) model

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{29}$$

Assume that

Construction from perturbed systems

Consider the (very common macroeconomic) model

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{29}$$

Assume that

1. ε_t an r.v. taking values in X.

Construction from perturbed systems

Consider the (very common macroeconomic) model

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{29}$$

Assume that

1. ε_t an r.v. taking values in X.

2. uncorrelated over time

Construction from perturbed systems

Consider the (very common macroeconomic) model

$$x_{t+1} = g(x_t, \varepsilon_t). \tag{29}$$

Assume that

1. ε_t an r.v. taking values in X.

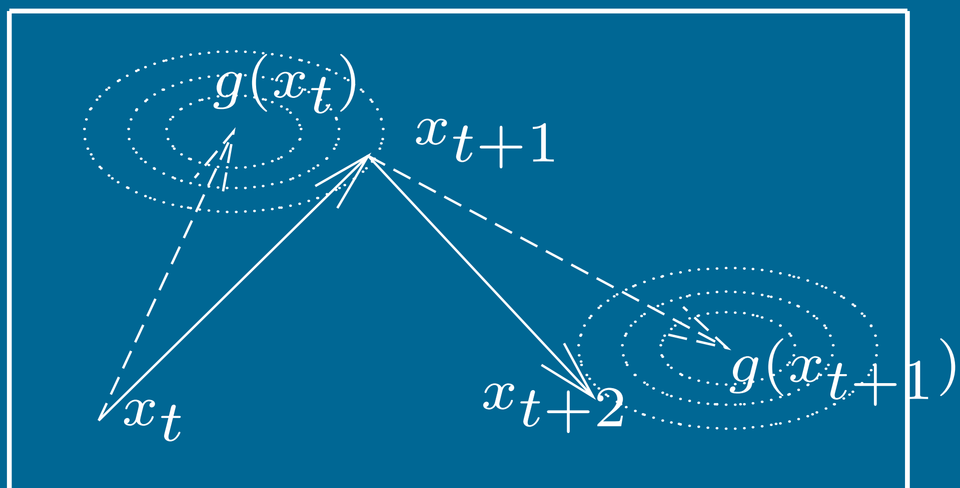
2. uncorrelated over time

1. identically distributed by $\psi \in \mathscr{D}(\mu)$.

Back to top

We can construct a *conditional distribution* $\Gamma(x_{t+1}, x_t)$ for x_{t+1} given x_t and knowledge of g, ψ.

We can construct a *conditional distribution* $\Gamma(x_{t+1}, x_t)$ for x_{t+1} given x_t and knowledge of g, ψ.



Example: AR(1).

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

In this case how is Γ?

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

In this case how is Γ?

Clearly

$$\Gamma(x_t, x_{t+1}) = (2\pi)^{-1/2} \exp\left[\frac{-(x_{t+1} - (a \cdot x_t + b))^2}{2}\right]$$

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

In this case how is Γ?

Clearly

$$\Gamma(x_t, x_{t+1}) = (2\pi)^{-1/2} \exp\left[\frac{-(x_{t+1} - (a \cdot x_t + b))^2}{2}\right]$$

Assume throughout that $\Gamma(x, \cdot)$ always a density.

Assume now that the initial condition x_0 is drawn from $\varphi_0 \in \mathscr{D}(\mu)$.

Assume now that the initial condition x_0 is drawn from $\varphi_0 \in \mathscr{D}(\mu)$.

What is the marginal distribution φ_1 of x_1?

Assume now that the initial condition x_0 is drawn from $\varphi_0 \in \mathscr{D}(\mu)$.

What is the marginal distribution φ_1 of x_1?

By the law of total probability,

$$\varphi_1(x_1) = \int \Gamma(x_0, x_1)\varphi_0(x_0)dx_0. \tag{30}$$

Assume now that the initial condition x_0 is drawn from $\varphi_0 \in \mathscr{D}(\mu)$.

What is the marginal distribution φ_1 of x_1?

By the law of total probability,

$$\varphi_1(x_1) = \int \Gamma(x_0, x_1)\varphi_0(x_0)dx_0. \tag{30}$$

And similarly,

$$\varphi_2(x_2) = \int \Gamma(x_1, x_2)\varphi_1(x_1)dx_1. \tag{31}$$

More generally,

$$\varphi_{t+1}(x_{t+1}) = \int \Gamma(x_t, x_{t+1})\varphi_t(x_t)dx_t. \tag{32}$$

More generally,

$$\varphi_{t+1}(x_{t+1}) = \int \Gamma(x_t, x_{t+1})\varphi_t(x_t)dx_t. \tag{32}$$

So now define an operator P mapping $\mathscr{L}_1(\mu)$ into itself, where if f is in $\mathscr{L}_1(\mu)$ then $Pf \in \mathscr{L}_1(\mu)$ is defined by

$$Pf(y) = \int \Gamma(x, y)f(x)dx. \tag{33}$$

More generally,

$$\varphi_{t+1}(x_{t+1}) = \int \Gamma(x_t, x_{t+1})\varphi_t(x_t)dx_t. \tag{32}$$

So now define an operator P mapping $\mathscr{L}_1(\mu)$ into itself, where if f is in $\mathscr{L}_1(\mu)$ then $Pf \in \mathscr{L}_1(\mu)$ is defined by

$$Pf(y) = \int \Gamma(x, y)f(x)dx. \tag{33}$$

Then $\varphi_{t+1} = P\varphi_t$.

More generally,

$$\varphi_{t+1}(x_{t+1}) = \int \Gamma(x_t, x_{t+1})\varphi_t(x_t)dx_t. \tag{32}$$

So now define an operator P mapping $\mathscr{L}_1(\mu)$ into itself, where if f is in $\mathscr{L}_1(\mu)$ then $Pf \in \mathscr{L}_1(\mu)$ is defined by

$$Pf(y) = \int \Gamma(x, y)f(x)dx. \tag{33}$$

Then $\varphi_{t+1} = P\varphi_t$. Alternatively, $\varphi_t = P^t\varphi_0$.

Back to top

Note that $\varphi \in \mathscr{D}(\mu) \implies P\varphi \in \mathscr{D}(\mu)$:

Note that $\varphi \in \mathscr{D}(\mu) \implies P\varphi \in \mathscr{D}(\mu)$:

$$\int P\varphi(y)dy = \int \left[\int \Gamma(x,y)\varphi(x)dx \right] dy \tag{34}$$

Note that $\varphi \in \mathscr{D}(\mu) \implies P\varphi \in \mathscr{D}(\mu)$:

$$\int P\varphi(y)dy = \int \left[\int \Gamma(x,y)\varphi(x)dx\right] dy \tag{34}$$

$$= \int \left[\int \Gamma(x,y)dy\right] \varphi(x)dx \tag{35}$$

Note that $\varphi \in \mathscr{D}(\mu) \implies P\varphi \in \mathscr{D}(\mu)$:

$$\int P\varphi(y)dy = \int \left[\int \Gamma(x,y)\varphi(x)dx\right] dy \tag{34}$$

$$= \int \left[\int \Gamma(x,y)dy\right] \varphi(x)dx \tag{35}$$

$$= \int \varphi(x)dx = 1.$$

Note that $\varphi \in \mathscr{D}(\mu) \implies P\varphi \in \mathscr{D}(\mu)$:

$$\int P\varphi(y)dy = \int \left[\int \Gamma(x,y)\varphi(x)dx \right] dy \tag{34}$$

$$= \int \left[\int \Gamma(x,y)dy \right] \varphi(x)dx \tag{35}$$

$$= \int \varphi(x)dx = 1. \tag{36}$$

Therefore $(\mathscr{D}(\mu), P)$ is a semidynamical system!!

Back to top

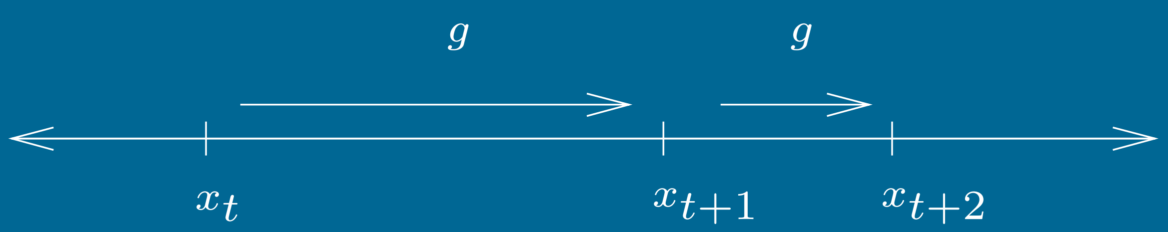
To recap, take a deterministic system $x_{t+1} = g(x_t)$, $x \in X$.

To recap, take a deterministic system $x_{t+1} = g(x_t)$, $x \in X$.

This is a semidynamical system (X, g).

To recap, take a deterministic system $x_{t+1} = g(x_t)$, $x \in X$.

This is a semidynamical system (X, g).



Back to top

Now consider $x_{t+1} = g(x_t, \varepsilon_t)$.

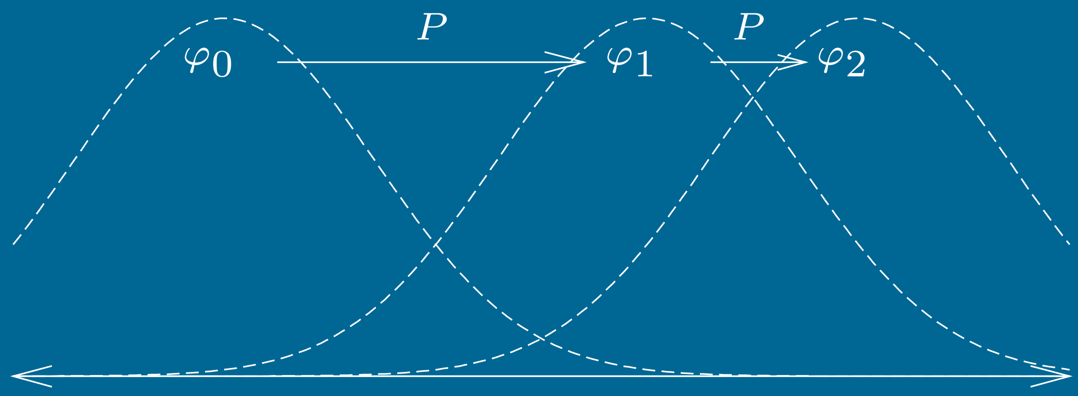
Now consider $x_{t+1} = g(x_t, \varepsilon_t)$. Calculate Γ, and hence P:
$Pf(y) = \int \Gamma(x, y) f(x) dx$.

Now consider $x_{t+1} = g(x_t, \varepsilon_t)$. Calculate Γ, and hence P:
$Pf(y) = \int \Gamma(x, y) f(x) dx$.

This returns us to a semidynamical system $(\mathscr{D}(\mu), P)$.

Now consider $x_{t+1} = g(x_t, \varepsilon_t)$. Calculate Γ, and hence P:
$Pf(y) = \int \Gamma(x,y) f(x) dx$.

This returns us to a semidynamical system $(\mathscr{D}(\mu), P)$.



Back to top

Equilibrium and stability

Traditional definition of stochastic equilibrium is just our definition for semidynamical system $(\mathscr{D}(\mu), P)$.

Equilibrium and stability

Traditional definition of stochastic equilibrium is just our definition for semidynamical system $(\mathscr{D}(\mu), P)$.

For example, Brock and Mirman (JET, 1972) study growth model generating sequence of r.v.s $\{k_t\}$.

Equilibrium and stability

Traditional definition of stochastic equilibrium is just our definition for semidynamical system $(\mathscr{D}(\mu), P)$.

For example, Brock and Mirman (JET, 1972) study growth model generating sequence of r.v.s $\{k_t\}$. $k_t \sim \varphi_t$.

Equilibrium and stability

Traditional definition of stochastic equilibrium is just our definition for semidynamical system $(\mathscr{D}(\mu), P)$.

For example, Brock and Mirman (JET, 1972) study growth model generating sequence of r.v.s $\{k_t\}$. $k_t \sim \varphi_t$.

They define stochastic equilibrium as a φ^* with

$$\varphi^*(x_{t+1}) = \int \Gamma(x_t, x_{t+1}) \varphi^*(x_t) dx_t. \tag{37}$$

Equilibrium and stability

Traditional definition of stochastic equilibrium is just our definition for semidynamical system $(\mathscr{D}(\mu), P)$.

For example, Brock and Mirman (JET, 1972) study growth model generating sequence of r.v.s $\{k_t\}$. $k_t \sim \varphi_t$.

They define stochastic equilibrium as a φ^* with

$$\varphi^*(x_{t+1}) = \int \Gamma(x_t, x_{t+1}) \varphi^*(x_t) dx_t. \tag{37}$$

But this just means a $\varphi^* \in \mathscr{D}(\mu)$ s.t. $P\varphi^* = \varphi^*$.

Back to top

Part 4: Stability of Markov Processes

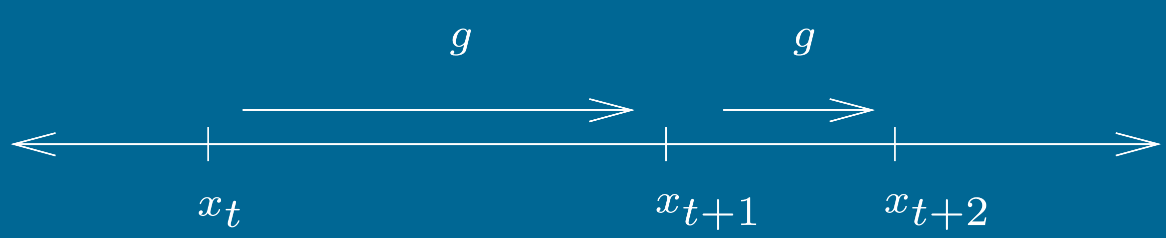
To recap, take a deterministic system $x_{t+1} = g(x_t)$, $x \in X$.

To recap, take a deterministic system $x_{t+1} = g(x_t)$, $x \in X$.

This is a semidynamical system (X, g).

To recap, take a deterministic system $x_{t+1} = g(x_t)$, $x \in X$.

This is a semidynamical system (X, g).



Back to top

Now consider $x_{t+1} = g(x_t, \varepsilon_t)$.

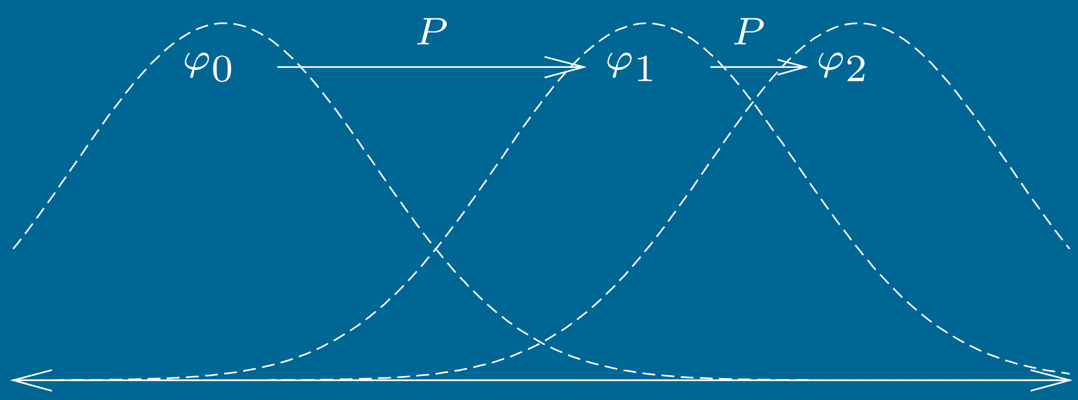
Now consider $x_{t+1} = g(x_t, \varepsilon_t)$. Calculate Γ, and hence P:
$Pf(y) = \int \Gamma(x, y) f(x) dx$.

Now consider $x_{t+1} = g(x_t, \varepsilon_t)$. Calculate Γ, and hence P:
$Pf(y) = \int \Gamma(x, y) f(x) dx$.

This returns us to a semidynamical system $(\mathscr{D}(\mu), P)$.

Now consider $x_{t+1} = g(x_t, \varepsilon_t)$. Calculate Γ, and hence P: $Pf(y) = \int \Gamma(x,y) f(x) dx$.

This returns us to a semidynamical system $(\mathscr{D}(\mu), P)$.



Analysis of $(\mathscr{D}(\mu), P)$.

Analysis of $(\mathscr{D}(\mu), P)$.

Recall from Brouwer/Shauder that if there exists a $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ which is convex and compact

Analysis of $(\mathscr{D}(\mu), P)$.

Recall from Brouwer/Shauder that if there exists a $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ which is convex and compact and satisfies $P\colon \mathscr{D}_0 \to \mathscr{D}_0$,

Analysis of $(\mathscr{D}(\mu), P)$.

Recall from Brouwer/Shauder that if there exists a $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ which is convex and compact and satisfies $P\colon \mathscr{D}_0 \to \mathscr{D}_0$, then P has a fixed point in $\mathscr{D}_0$.

Analysis of $(\mathscr{D}(\mu), P)$.

Recall from Brouwer/Shauder that if there exists a $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ which is convex and compact and satisfies $P\colon \mathscr{D}_0 \to \mathscr{D}_0$, then P has a fixed point in $\mathscr{D}_0$.

Recall from Lasota that if P satisfies

$$\|P\varphi' - P\varphi''\| < \|\varphi' - \varphi''\|, \quad \forall \varphi' \neq \varphi''$$

Analysis of $(\mathscr{D}(\mu), P)$.

Recall from Brouwer/Shauder that if there exists a $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ which is convex and compact and satisfies $P\colon \mathscr{D}_0 \to \mathscr{D}_0$, then P has a fixed point in $\mathscr{D}_0$.

Recall from Lasota that if P satisfies

$$\|P\varphi' - P\varphi''\| < \|\varphi' - \varphi''\|, \quad \forall \varphi' \neq \varphi''$$

and $\{P^t \varphi\}$ is precompact, $\forall \varphi \in \mathscr{D}(\mu)$, then exists unique, globally stable equilibrium

Back to top

Strongly contracting Markov operators

Proposition. If $\Gamma(x, y) > 0$ for all x, y, then P is strongly contracting.

Intuition...

Strongly contracting Markov operators

Proposition. If $\Gamma(x, y) > 0$ for all x, y, then P is strongly contracting.

Intuition... I don't know

Proof...

Back to top

Back to top

$$\|P\varphi' - P\varphi''\| = \|P(\varphi' - \varphi'')\|$$

$$\|P\varphi' - P\varphi''\| = \|P(\varphi' - \varphi'')\|$$
$$= \int |\int \Gamma(x, y)[\varphi'(x) - \varphi''(x)]dx|dy$$

$$
\begin{aligned}
\|P\varphi' - P\varphi''\| &= \|P(\varphi' - \varphi'')\| \\
&= \int | \int \Gamma(x, y)[\varphi'(x) - \varphi''(x)]dx|dy \\
&< \int \int |\Gamma(x, y)[\varphi'(x) - \varphi''(x)]|dx\,dy
\end{aligned}
$$

$$
\begin{aligned}
\|P\varphi' - P\varphi''\| &= \|P(\varphi' - \varphi'')\| \\
&= \int | \int \Gamma(x,y)[\varphi'(x) - \varphi''(x)]dx|dy \\
&< \int \int |\Gamma(x,y)[\varphi'(x) - \varphi''(x)]|dx\,dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dx\,dy
\end{aligned}
$$

$$
\begin{aligned}
\|P\varphi' - P\varphi''\| &= \|P(\varphi' - \varphi'')\| \\
&= \int | \int \Gamma(x,y)[\varphi'(x) - \varphi''(x)]dx|dy \\
&< \int \int |\Gamma(x,y)[\varphi'(x) - \varphi''(x)]|dx\,dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dx\,dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dydx
\end{aligned}
$$

$$
\begin{aligned}
\|P\varphi' - P\varphi''\| &= \|P(\varphi' - \varphi'')\| \\
&= \int | \int \Gamma(x,y)[\varphi'(x) - \varphi''(x)]dx|dy \\
&< \int \int |\Gamma(x,y)[\varphi'(x) - \varphi''(x)]|dx\,dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dx\,dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dydx \\
&= \int \int \Gamma(x,y)dy|\varphi'(x) - \varphi''(x)|dx
\end{aligned}
$$

$$
\begin{aligned}
\|P\varphi' - P\varphi''\| &= \|P(\varphi' - \varphi'')\| \\
&= \int |\int \Gamma(x,y)[\varphi'(x) - \varphi''(x)]dx|dy \\
&< \int \int |\Gamma(x,y)[\varphi'(x) - \varphi''(x)]|dx\, dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dx\, dy \\
&= \int \int \Gamma(x,y)|\varphi'(x) - \varphi''(x)|dydx \\
&= \int \int \Gamma(x,y)dy|\varphi'(x) - \varphi''(x)|dx \\
&= \|\varphi' - \varphi''\|.
\end{aligned}
$$

Back to top

Application to AR(1)

Example: AR(1).

Application to AR(1)

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

Application to AR(1)

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

Recall that

$$\Gamma(x_t, x_{t+1}) = (2\pi)^{-1/2} \exp\left[\frac{-(x_{t+1} - (a \cdot x_t + b))^2}{2}\right] > 0.$$

Application to AR(1)

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

Recall that

$$\Gamma(x_t, x_{t+1}) = (2\pi)^{-1/2} \exp\left[\frac{-(x_{t+1} - (a \cdot x_t + b))^2}{2}\right] > 0.$$

Hence the AR(1) process is strongly contracting.

Application to AR(1)

Example: AR(1). $X = \mathbb{R}$, $x_{t+1} = a \cdot x_t + b + \varepsilon_t$, where ψ standard normal.

Recall that

$$\Gamma(x_t, x_{t+1}) = (2\pi)^{-1/2} \exp\left[\frac{-(x_{t+1} - (a \cdot x_t + b))^2}{2}\right] > 0.$$

Hence the AR(1) process is strongly contracting.

Corollary. At most one equilibrium

Lagrange stability of Markov operators

When is the collection $\{P^t\varphi\}$ precompact for all $\varphi \in \mathscr{D}(\mu)$?

Lagrange stability of Markov operators

When is the collection $\{P^t\varphi\}$ precompact for all $\varphi \in \mathscr{D}(\mu)$?

More generally, when is $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ precompact?

Lagrange stability of Markov operators

When is the collection $\{P^t\varphi\}$ precompact for all $\varphi \in \mathscr{D}(\mu)$?

More generally, when is $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ precompact?

Precompact is in a sense like "bounded"...

Lagrange stability of Markov operators

When is the collection $\{P^t\varphi\}$ precompact for all $\varphi \in \mathscr{D}(\mu)$?

More generally, when is $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ precompact?

Precompact is in a sense like "bounded"...

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(n, 1)$, $n \in \mathbb{N}$.

Lagrange stability of Markov operators

When is the collection $\{P^t\varphi\}$ precompact for all $\varphi \in \mathscr{D}(\mu)$?

More generally, when is $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ precompact?

Precompact is in a sense like "bounded"...

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(n, 1)$, $n \in \mathbb{N}$.

Rule this out by tightness:

Lagrange stability of Markov operators

When is the collection $\{P^t\varphi\}$ precompact for all $\varphi \in \mathscr{D}(\mu)$?

More generally, when is $\mathscr{D}_0 \subset \mathscr{D}(\mu)$ precompact?

Precompact is in a sense like "bounded"...

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(n, 1)$, $n \in \mathbb{N}$.

Rule this out by tightness:

$\mathscr{D}_0$ is *tight* if $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} \varphi < \varepsilon$ whenever $\varphi \in \mathscr{D}_0$.

Back to top

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

Rule this out by a kind of upper bound.

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

Rule this out by a kind of upper bound.

Lasota-Stachurski Theorem.

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

Rule this out by a kind of upper bound.

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ.

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

Rule this out by a kind of upper bound.

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x, y) \leq h(y)$,

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

Rule this out by a kind of upper bound.

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x, y) \leq h(y)$, and

2. $\{P^t \varphi\}$ is tight,

This set is *not* precompact: $\{\varphi_n\}$, $\varphi_n \sim N(0, 1/n)$, $n \in \mathbb{N}$.

Rule this out by a kind of upper bound.

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x, y) \leq h(y)$, and

2. $\{P^t \varphi\}$ is tight,

then $\{P^t \varphi\}$ is precompact

Application to AR(1)

Recall again the example where $X = \mathbb{R}$,

$$x_{t+1} = a \cdot x_t + b + \varepsilon_t, \quad \varepsilon \sim N(0, 1). \tag{38}$$

Application to AR(1)

Recall again the example where $X = \mathbb{R}$,

$$x_{t+1} = a \cdot x_t + b + \varepsilon_t, \quad \varepsilon \sim N(0,1). \tag{38}$$

In this case

$$\Gamma(x,y) = (2\pi)^{-1/2} \exp\left[\frac{-(y-(a \cdot x + b))^2}{2}\right]$$

Application to AR(1)

Recall again the example where $X = \mathbb{R}$,

$$x_{t+1} = a \cdot x_t + b + \varepsilon_t, \quad \varepsilon \sim N(0,1). \tag{38}$$

In this case

$$\Gamma(x,y) = (2\pi)^{-1/2} \exp\left[\frac{-(y-(a \cdot x + b))^2}{2}\right]$$

But this is $\leq (2\pi)^{-1/2}$, so...

Application to AR(1)

Recall again the example where $X = \mathbb{R}$,

$$x_{t+1} = a \cdot x_t + b + \varepsilon_t, \quad \varepsilon \sim N(0,1). \tag{38}$$

In this case

$$\Gamma(x,y) = (2\pi)^{-1/2} \exp\left[\frac{-(y-(a \cdot x + b))^2}{2}\right]$$

But this is $\leq (2\pi)^{-1/2}$, so... exists continuous h such that $\Gamma(x,y) \leq h(y)$

Back to top

How about tightness?

How about tightness?

Definition: $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

How about tightness?

Definition: $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

Consider again AR(1) example.

How about tightness?

Definition: $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

Consider again AR(1) example.

Proposition. Let $X = \mathbb{R}$, and let

$$x_{t+1} = a \cdot x_t + b + \varepsilon_t, \quad \varepsilon \sim N(0,1). \tag{39}$$

How about tightness?

Definition: $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

Consider again AR(1) example.

Proposition. Let $X = \mathbb{R}$, and let

$$x_{t+1} = a \cdot x_t + b + \varepsilon_t, \quad \varepsilon \sim N(0,1). \tag{39}$$

Let P be the associated Markov operator. If $|a| < 1$, then $P^t \varphi$ is tight for every Gaussian $\varphi \in \mathscr{D}(\mu)$.

Back to top

Chebychev inequality: If ξ is an r.v. on $\mathbb{R}$, then

$$\mathrm{Prob}(|\xi| \geq r) \leq \frac{\mathbb{E}|\xi|}{r}, \quad \forall r > 0. \tag{40}$$

Chebychev inequality: If ξ is an r.v. on $\mathbb{R}$, then

$$\mathrm{Prob}(|\xi| \geq r) \leq \frac{\mathbb{E}|\xi|}{r}, \quad \forall r > 0. \tag{40}$$

Therefore, for each t, x_t satisfies

$$\int_{[-r,r]^c} P^t \varphi(x) dx \leq \frac{\mathbb{E}_\varphi |x_t|}{r}, \quad \forall r > 0. \tag{41}$$

Chebychev inequality: If ξ is an r.v. on $\mathbb{R}$, then

$$\mathrm{Prob}(|\xi| \geq r) \leq \frac{\mathbb{E}|\xi|}{r}, \quad \forall r > 0. \tag{40}$$

Therefore, for each t, x_t satisfies

$$\int_{[-r,r]^c} P^t \varphi(x) dx \leq \frac{\mathbb{E}_\varphi |x_t|}{r}, \quad \forall r > 0. \tag{41}$$

If $\exists M < \infty$ s.t. $\mathbb{E}_\varphi |x_t| \leq M$ for all t then we are done.

Back to top

Since $|x| \leq 1 + x^2$,

$$\mathbb{E}|x_t| \leq 1 + \mathbb{E}x_t^2$$

Since $|x| \leq 1 + x^2$,

$$\begin{aligned} \mathbb{E}|x_t| &\leq 1 + \mathbb{E}x_t^2 \\ &= 1 + (\mathbb{E}x_t)^2 + \mathbb{V}x_t \qquad [\mathbb{V}x_t = \mathbb{E}x_t^2 - (\mathbb{E}x_t)^2] \end{aligned}$$

Since $|x| \leq 1 + x^2$,

$$\begin{aligned} \mathbb{E}|x_t| &\leq 1 + \mathbb{E}x_t^2 \\ &= 1 + (\mathbb{E}x_t)^2 + \mathbb{V}x_t \qquad [\mathbb{V}x_t = \mathbb{E}x_t^2 - (\mathbb{E}x_t)^2] \\ &= 1 + \mu_t^2 + \sigma_t^2. \end{aligned}$$

Since $|x| \leq 1 + x^2$,

$$\begin{aligned} \mathbb{E}|x_t| &\leq 1 + \mathbb{E}x_t^2 \\ &= 1 + (\mathbb{E}x_t)^2 + \mathbb{V}x_t \qquad [\mathbb{V}x_t = \mathbb{E}x_t^2 - (\mathbb{E}x_t)^2] \\ &= 1 + \mu_t^2 + \sigma_t^2. \end{aligned}$$

But in this case

$$\mu_{t+1} = b + a\mu_t, \quad \mu_0 = \mathbb{E}x_0. \tag{42}$$

Since $|x| \leq 1 + x^2$,

$$\begin{aligned} \mathbb{E}|x_t| &\leq 1 + \mathbb{E}x_t^2 \\ &= 1 + (\mathbb{E}x_t)^2 + \mathbb{V}x_t \qquad [\mathbb{V}x_t = \mathbb{E}x_t^2 - (\mathbb{E}x_t)^2] \\ &= 1 + \mu_t^2 + \sigma_t^2. \end{aligned}$$

But in this case

$$\mu_{t+1} = b + a\mu_t, \quad \mu_0 = \mathbb{E}x_0. \tag{42}$$

$$\sigma_{t+1}^2 = 1 + a^2\sigma_t^2, \quad \sigma_0^2 = \mathbb{V}x_0. \tag{43}$$

Since $|x| \leq 1 + x^2$,

$$\begin{aligned} \mathbb{E}|x_t| &\leq 1 + \mathbb{E}x_t^2 \\ &= 1 + (\mathbb{E}x_t)^2 + \mathbb{V}x_t \qquad [\mathbb{V}x_t = \mathbb{E}x_t^2 - (\mathbb{E}x_t)^2] \\ &= 1 + \mu_t^2 + \sigma_t^2. \end{aligned}$$

But in this case

$$\mu_{t+1} = b + a\mu_t, \quad \mu_0 = \mathbb{E}x_0. \tag{42}$$

$$\sigma_{t+1}^2 = 1 + a^2\sigma_t^2, \quad \sigma_0^2 = \mathbb{V}x_0. \tag{43}$$

Therefore, $a < 1$ implies tightness

Back to top

Part 5: Applications

- ⋆ The stochastic growth model
- ⋆ Stability in the increasing returns model

Back to top

The stochastic growth model

Recall that problem is to choose a consumption policy $g\colon x \longrightarrow c$ to maximize

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(g(x_t))\right] \tag{44}$$

The stochastic growth model

Recall that problem is to choose a consumption policy $g\colon x \longrightarrow c$ to maximize

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(g(x_t))\right] \tag{44}$$

subject to

$$x_{t+1} = f(x_t - g(x_t))\varepsilon_t. \tag{45}$$

The stochastic growth model

Recall that problem is to choose a consumption policy $g\colon x \to c$ to maximize

$$\mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t u(g(x_t))\right] \tag{44}$$

subject to

$$x_{t+1} = f(x_t - g(x_t))\varepsilon_t. \tag{45}$$

Recall that if optimal g exists then

$$u'(g(x)) = \beta \int u'[g(f(x - g(x))z)] f'(x - g(x)) z \psi(dz). \tag{46}$$

Back to top

The state space is $X = (0, \infty)$

The state space is $X = (0, \infty)$

We assume that...

The state space is $X = (0, \infty)$

We assume that...

The utility function u satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$,

The state space is $X = (0, \infty)$

We assume that...

The utility function u satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$, preferably nonnegative and bounded.

The state space is $X = (0, \infty)$

We assume that...

The utility function u satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$, preferably nonnegative and bounded.

The production function satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$,

The state space is $X = (0, \infty)$

We assume that...

The utility function u satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$, preferably nonnegative and bounded.

The production function satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

The state space is $X = (0, \infty)$

We assume that...

The utility function u satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$, preferably nonnegative and bounded.

The production function satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

The shocks are i.i.d. and have density function ψ.

The state space is $X = (0, \infty)$

We assume that...

The utility function u satisfies $u' > 0$, $u'' < 0$, $u'(0) = \infty$, preferably nonnegative and bounded.

The production function satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

The shocks are i.i.d. and have density function ψ.

Under these assumptions g exists and is unique.

Back to top

Prove this model has a unique, globally stable (stochastic) equilibrium.

Prove this model has a unique, globally stable (stochastic) equilibrium.

Case of shock with compact support: $\varepsilon \in [a, b]$ with prob 1:

Prove this model has a unique, globally stable (stochastic) equilibrium.

Case of shock with compact support: $\varepsilon \in [a, b]$ with prob 1:

- Brock and Mirman (JET, 1972)

Prove this model has a unique, globally stable (stochastic) equilibrium.

Case of shock with compact support: $\varepsilon \in [a, b]$ with prob 1:

- Brock and Mirman (JET, 1972)
- Razin and Yahav (IER, 1979)

Prove this model has a unique, globally stable (stochastic) equilibrium.

Case of shock with compact support: $\varepsilon \in [a, b]$ with prob 1:

- Brock and Mirman (JET, 1972)
- Razin and Yahav (IER, 1979)
- Lucas, Stokey and Prescott (RMED, 1989)

Prove this model has a unique, globally stable (stochastic) equilibrium.

Case of shock with compact support: $\varepsilon \in [a, b]$ with prob 1:

- Brock and Mirman (JET, 1972)
- Razin and Yahav (IER, 1979)
- Lucas, Stokey and Prescott (RMED, 1989)
- Hopenhayn and Prescott (Econometrica, 1992)

Prove this model has a unique, globally stable (stochastic) equilibrium.

Case of shock with compact support: $\varepsilon \in [a, b]$ with prob 1:

- Brock and Mirman (JET, 1972)
- Razin and Yahav (IER, 1979)
- Lucas, Stokey and Prescott (RMED, 1989)
- Hopenhayn and Prescott (Econometrica, 1992)

Unbounded shock: Stachurski (JET, 2002)

With restrictions on u, f and hence g, we analyze

$$x_{t+1} = f(x_t - g(x_t))\varepsilon_t. \tag{47}$$

With restrictions on u, f and hence g, we analyze

$$x_{t+1} = f(x_t - g(x_t))\varepsilon_t. \tag{47}$$

From this we obtain $\Gamma(x, y)$.

With restrictions on u, f and hence g, we analyze

$$x_{t+1} = f(x_t - g(x_t))\varepsilon_t. \tag{47}$$

From this we obtain $\Gamma(x, y)$. In fact by a c-o-v argument,

$$\Gamma(x, y) = \psi\left(\frac{y}{f(x - g(x))}\right)\frac{1}{f(x - g(x))}. \tag{48}$$

With restrictions on u, f and hence g, we analyze

$$x_{t+1} = f(x_t - g(x_t))\varepsilon_t. \tag{47}$$

From this we obtain $\Gamma(x, y)$. In fact by a c-o-v argument,

$$\Gamma(x, y) = \psi\left(\frac{y}{f(x - g(x))}\right)\frac{1}{f(x - g(x))}. \tag{48}$$

And hence P as per usual

$$Pf(y) = \int \Gamma(x, y) f(x) dx.$$

Back to top

So now we have a semidynamical system $(\mathscr{D}(\mu), P)$.

So now we have a semidynamical system $(\mathscr{D}(\mu), P)$.

Recall the general method of Lasota: Existence, uniqueness and global stability if $(\mathscr{D}(\mu), P)$ is strongly contracting and Lagrange stable.

So now we have a semidynamical system $(\mathscr{D}(\mu), P)$.

Recall the general method of Lasota: Existence, uniqueness and global stability if $(\mathscr{D}(\mu), P)$ is strongly contracting and Lagrange stable.

Thus if P satisfies

$$\|P\varphi' - P\varphi''\| < \|\varphi' - \varphi''\|, \quad \forall \varphi' \neq \varphi''$$

So now we have a semidynamical system $(\mathscr{D}(\mu), P)$.

Recall the general method of Lasota: Existence, uniqueness and global stability if $(\mathscr{D}(\mu), P)$ is strongly contracting and Lagrange stable.

Thus if P satisfies

$$\|P\varphi' - P\varphi''\| < \|\varphi' - \varphi''\|, \quad \forall \varphi' \neq \varphi''$$

and also $\{P^t\varphi\}$ is precompact for all $\varphi \in \mathscr{D}(\mu)$,

So now we have a semidynamical system $(\mathscr{D}(\mu), P)$.

Recall the general method of Lasota: Existence, uniqueness and global stability if $(\mathscr{D}(\mu), P)$ is strongly contracting and Lagrange stable.

Thus if P satisfies

$$\|P\varphi' - P\varphi''\| < \|\varphi' - \varphi''\|, \quad \forall \varphi' \neq \varphi''$$

and also $\{P^t\varphi\}$ is precompact for all $\varphi \in \mathscr{D}(\mu)$,

then we are done.

Check first strong contractiveness.

Check first strong contractiveness.

Recall that a sufficient condition is $\Gamma(x, y) > 0$ everywhere.

Check first strong contractiveness.

Recall that a sufficient condition is $\Gamma(x,y) > 0$ everywhere.

But

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{1}{f(x-g(x))}, \tag{49}$$

Check first strong contractiveness.

Recall that a sufficient condition is $\Gamma(x, y) > 0$ everywhere.

But

$$\Gamma(x, y) = \psi\left(\frac{y}{f(x - g(x))}\right)\frac{1}{f(x - g(x))}, \tag{49}$$

Therefore $\psi > 0$ everywhere on $(0, \infty)$ implies strong contractiveness.

Check first strong contractiveness.

Recall that a sufficient condition is $\Gamma(x,y) > 0$ everywhere.

But

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{1}{f(x-g(x))}, \tag{49}$$

Therefore $\psi > 0$ everywhere on $(0,\infty)$ implies strong contractiveness.

Examples: lognormal, gamma, wiebull, χ-squared, etc.

Check first strong contractiveness.

Recall that a sufficient condition is $\Gamma(x, y) > 0$ everywhere.

But

$$\Gamma(x, y) = \psi\left(\frac{y}{f(x - g(x))}\right)\frac{1}{f(x - g(x))}, \tag{49}$$

Therefore $\psi > 0$ everywhere on $(0, \infty)$ implies strong contractiveness.

Examples: lognormal, gamma, wiebull, χ-squared, etc.

Note: This is all we need for uniqueness.

Check next Lagrange stability.

Check next Lagrange stability.

Recall the

Lasota-Stachurski Theorem.

Check next Lagrange stability.

Recall the

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x, y) \leq h(y)$, and

2. $\{P^t\varphi\}$ is tight,

then $\{P^t\varphi\}$ is precompact

Check next Lagrange stability.

Recall the

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x, y) \leq h(y)$, and
2. $\{P^t \varphi\}$ is tight,

then $\{P^t \varphi\}$ is precompact

We check Condition 1 first.

Show exists continuous h such that $\Gamma(x, y) \leq h(y)$.

Show exists continuous h such that $\Gamma(x, y) \leq h(y)$.

We have

$$\Gamma(x, y) = \psi\left(\frac{y}{f(x - g(x))}\right)\frac{1}{f(x - g(x))}, \tag{50}$$

Show exists continuous h such that $\Gamma(x,y) \leq h(y)$.

We have

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{1}{f(x-g(x))}, \tag{50}$$

Suppose that $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$.

Show exists continuous h such that $\Gamma(x,y) \leq h(y)$.

We have

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{1}{f(x-g(x))}, \tag{50}$$

Suppose that $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$.

Then on $(0,\infty)$,

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{y}{f(x-g(x))}\frac{1}{y}$$

Show exists continuous h such that $\Gamma(x,y) \le h(y)$.

We have

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{1}{f(x-g(x))}, \tag{50}$$

Suppose that $\exists M < \infty$ with $\psi(z)z \le M$, $\forall z$.

Then on $(0,\infty)$,

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{y}{f(x-g(x))}\frac{1}{y} \le \frac{M}{y}$$

Show exists continuous h such that $\Gamma(x,y) \leq h(y)$.

We have

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{1}{f(x-g(x))}, \tag{50}$$

Suppose that $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$.

Then on $(0,\infty)$,

$$\Gamma(x,y) = \psi\left(\frac{y}{f(x-g(x))}\right)\frac{y}{f(x-g(x))}\frac{1}{y} \leq \frac{M}{y} = h(y). \tag{51}$$

Back to top

Regarding the condition $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$,

Regarding the condition $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$,

Suppose for example that ε is lognormal.

Regarding the condition $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$,

Suppose for example that ε is lognormal. Then

$$\psi(z) = (2\pi\sigma^2)^{-1/2}\frac{1}{z}\exp\left[-\frac{(\ln z - \mu)^2}{2\sigma^2}\right]. \tag{52}$$

Regarding the condition $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$,

Suppose for example that ε is lognormal. Then

$$\psi(z) = (2\pi\sigma^2)^{-1/2}\frac{1}{z}\exp\left[-\frac{(\ln z - \mu)^2}{2\sigma^2}\right]. \tag{52}$$

In this case the condition clearly holds.

Regarding the condition $\exists M < \infty$ with $\psi(z)z \leq M$, $\forall z$,

Suppose for example that ε is lognormal. Then

$$\psi(z) = (2\pi\sigma^2)^{-1/2} \frac{1}{z} \exp\left[-\frac{(\ln z - \mu)^2}{2\sigma^2}\right]. \tag{52}$$

In this case the condition clearly holds.

Same for many common shocks.

Back to top

Regarding Condition 2 (tightness),

Regarding Condition 2 (tightness), pick any $\varphi \in \mathscr{D}(\mu)$. We need to show that $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

Regarding Condition 2 (tightness), pick any $\varphi \in \mathscr{D}(\mu)$. We need to show that $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

Evidently following two conditions are sufficient

Regarding Condition 2 (tightness), pick any $\varphi \in \mathscr{D}(\mu)$. We need to show that $\forall \varepsilon > 0$, $\exists K$ compact such that $\int_{K^c} P^t \varphi < \varepsilon$, $\forall t \in \mathbb{N}$.

Evidently following two conditions are sufficient

$$\forall \varepsilon > 0,\ \exists r > 0 \ \text{ s.t. } \left\{ \int_r^\infty P^t \varphi(x) dx < \varepsilon, \quad \forall t \in \mathbb{N}_0 \right\} \tag{53}$$

and

$$\forall \varepsilon > 0,\ \exists r > 0 \ \text{ s.t. } \left\{ \int_0^{1/r} P^t \varphi(x) dx < \varepsilon \quad \forall t \in \mathbb{N}_0 \right\} \tag{54}$$

if we set $K = [1/r, r]$, r sufficiently large.

Back to top

We prove only the first condition:

$$\forall \varepsilon > 0,\ \exists r > 0 \ \text{ s.t. } \left\{ \int_r^{\infty} P^t \varphi(x) dx < \varepsilon, \quad \forall t \in \mathbb{N}_0 \right\} \tag{55}$$

We prove only the first condition:

$$\forall \varepsilon > 0,\ \exists r > 0 \ \text{ s.t. } \left\{ \int_r^{\infty} P^t \varphi(x) dx < \varepsilon, \quad \forall t \in \mathbb{N}_0 \right\} \tag{55}$$

In view of the Chebychev inequality

$$\int_r^{\infty} P^t \varphi(x) dx \leq \frac{\mathbb{E}_{\varphi} x_t}{r}, \quad \forall r > 0,$$

We prove only the first condition:

$$\forall \varepsilon > 0,\ \exists r > 0 \ \text{ s.t. } \left\{ \int_r^\infty P^t \varphi(x) dx < \varepsilon, \quad \forall t \in \mathbb{N}_0 \right\} \tag{55}$$

In view of the Chebychev inequality

$$\int_r^\infty P^t \varphi(x) dx \leq \frac{\mathbb{E}_\varphi x_t}{r}, \quad \forall r > 0,$$

need show only that the sequence of real numbers $\mathbb{E}_\varphi x_t$ is bounded.

Back to top

$$\mathbb{E}_{\varphi} x_t = \int_0^\infty \mathbb{E}(x_t | x_{t-1} = x) \,\mathrm{Prob}(x_{t-1} = x) dx$$

$$\begin{aligned}
\mathbb{E}_\varphi x_t &= \int_0^\infty \mathbb{E}(x_t | x_{t-1} = x) \, \mathrm{Prob}(x_{t-1} = x) dx \\
&= \int_0^\infty [f(x - g(x)) \mathbb{E}\varepsilon_t] P^{t-1} \varphi(x) dx \\
&\leq \int_0^\infty f(x) \mathbb{E}\varepsilon_t P^{t-1} \varphi(x) dx
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{\varphi} x_t &= \int_0^{\infty} \mathbb{E}(x_t | x_{t-1} = x) \,\mathrm{Prob}(x_{t-1} = x) dx \\
&= \int_0^{\infty} [f(x - g(x)) \mathbb{E}\varepsilon_t] P^{t-1} \varphi(x) dx \\
&\leq \int_0^{\infty} f(x) \mathbb{E}\varepsilon_t P^{t-1} \varphi(x) dx \\
&= \mathbb{E}\varepsilon_t \int_0^{\infty} f(x) P^{t-1} \varphi(x) dx
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}_{\varphi} x_t &= \int_0^{\infty} \mathbb{E}(x_t | x_{t-1} = x)\, \mathrm{Prob}(x_{t-1} = x) dx \\
&= \int_0^{\infty} [f(x - g(x)) \mathbb{E}\varepsilon_t] P^{t-1} \varphi(x) dx \\
&\leq \int_0^{\infty} f(x) \mathbb{E}\varepsilon_t P^{t-1} \varphi(x) dx \\
&= \mathbb{E}\varepsilon_t \int_0^{\infty} f(x) P^{t-1} \varphi(x) dx \\
&\leq \mathbb{E}\varepsilon_t f \left[\int_0^{\infty} x P^{t-1} \varphi(x) dx \right]
\end{aligned}$$

$$
\begin{aligned}
\mathbb{E}_\varphi x_t &= \int_0^\infty \mathbb{E}(x_t | x_{t-1} = x)\,\mathrm{Prob}(x_{t-1} = x) dx \\
&= \int_0^\infty [f(x - g(x))\mathbb{E}\varepsilon_t] P^{t-1}\varphi(x) dx \\
&\leq \int_0^\infty f(x)\mathbb{E}\varepsilon_t P^{t-1}\varphi(x) dx \\
&= \mathbb{E}\varepsilon_t \int_0^\infty f(x) P^{t-1}\varphi(x) dx \\
&\leq \mathbb{E}\varepsilon_t f\left[\int_0^\infty x P^{t-1}\varphi(x) dx\right] \\
&= C f[\mathbb{E}_\varphi x_{t-1}].
\end{aligned}
$$

Back to top

Therefore $\mathbb{E}_\varphi x_t \leq C f[\mathbb{E}_\varphi x_{t-1}]$,

Therefore $\mathbb{E}_\varphi x_t \leq Cf[\mathbb{E}_\varphi x_{t-1}]$, and the sequence $\mathbb{E}_\varphi x_t$ is bounded, as was to be proved.

Therefore $\mathbb{E}_\varphi x_t \leq C f[\mathbb{E}_\varphi x_{t-1}]$, and the sequence $\mathbb{E}_\varphi x_t$ is bounded, as was to be proved.

This proves tightness...

Therefore $\mathbb{E}_\varphi x_t \leq Cf[\mathbb{E}_\varphi x_{t-1}]$, and the sequence $\mathbb{E}_\varphi x_t$ is bounded, as was to be proved.

This proves tightness...

which proves Lagragne stability...

Therefore $\mathbb{E}_\varphi x_t \leq Cf[\mathbb{E}_\varphi x_{t-1}]$, and the sequence $\mathbb{E}_\varphi x_t$ is bounded, as was to be proved.

This proves tightness...

which proves Lagragne stability...

This completes proof of global stability!

Back to top

Incidentally, why did everyone else use $\varepsilon \in [a, b]$ with prob 1?

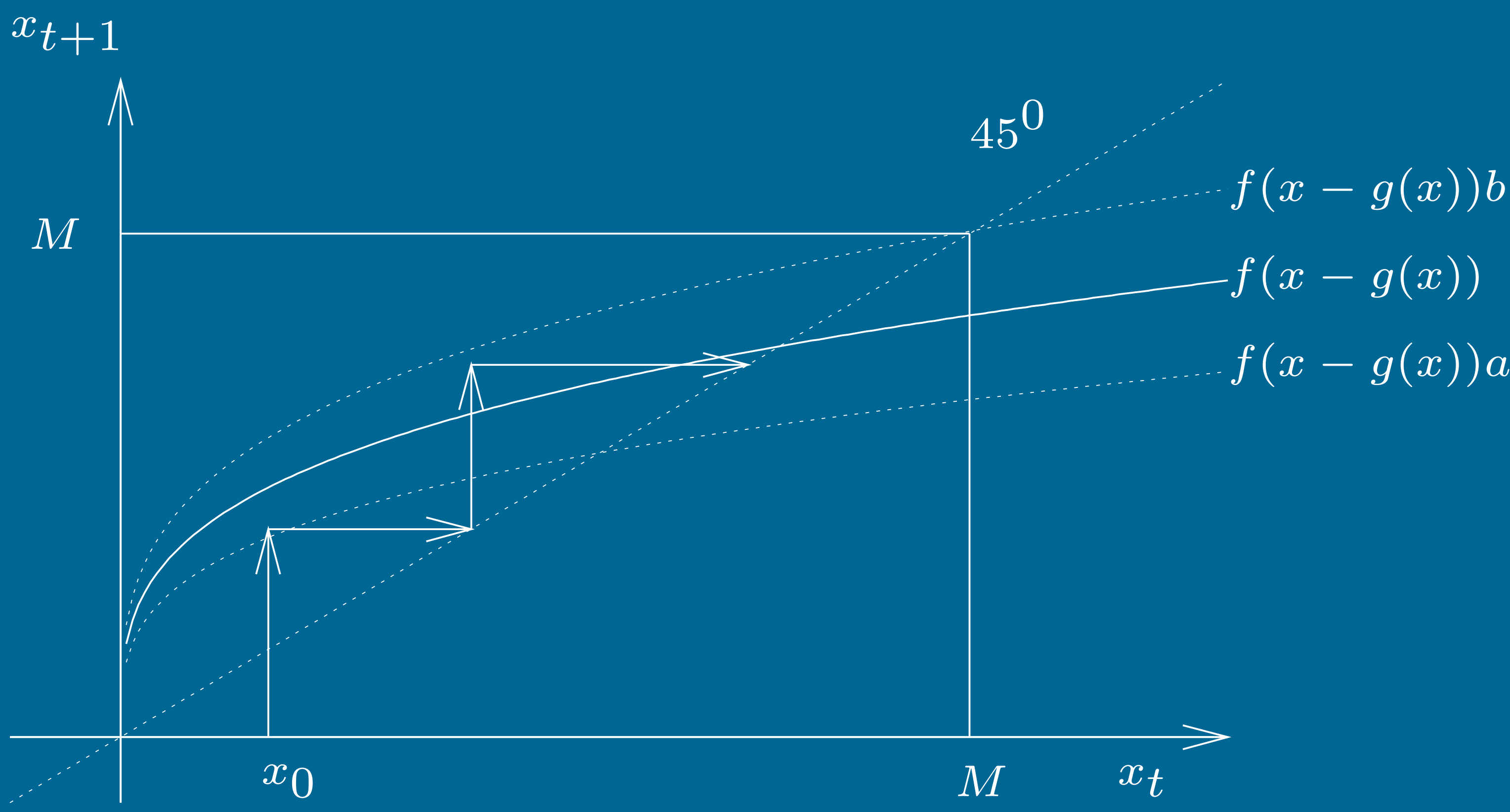
Incidentally, why did everyone else use $\varepsilon \in [a, b]$ with prob 1?

The reason is that the state space is can be taken to be compact, and hence tightness is trivial. (Intuition next slide.)

Incidentally, why did everyone else use $\varepsilon \in [a, b]$ with prob 1?

The reason is that the state space is can be taken to be compact, and hence tightness is trivial. (Intuition next slide.)

By working a bit harder for tightness we can incorporate models with unbounded state.



Back to top

Stability in the increasing returns model

Stability in the increasing returns model

The model is a stochastic verion of Azariadis and Drazen (QJE, 1990). overlapping generations with spillover.

Stability in the increasing returns model

The model is a stochastic verion of Azariadis and Drazen (QJE, 1990). overlapping generations with spillover.

Agents live for two periods, working in the first and living off savings in the second.

Stability in the increasing returns model

The model is a stochastic verion of Azariadis and Drazen (QJE, 1990). overlapping generations with spillover.

Agents live for two periods, working in the first and living off savings in the second.

Savings in first period in the following period is combined with the labor of new generation.

Stability in the increasing returns model

The model is a stochastic verion of Azariadis and Drazen (QJE, 1990). overlapping generations with spillover.

Agents live for two periods, working in the first and living off savings in the second.

Savings in first period in the following period is combined with the labor of new generation.

Depreciation is total.

The production function is

$$y_t = A_t F(k_t, \ell_t)\varepsilon_t^{\sigma}.$$

The production function is

$$y_t = A_t F(k_t, \ell_t) \varepsilon_t^{\sigma}.$$

Here ε is uncorrelated and identically distributed as usual.

The production function is

$$y_t = A_t F(k_t, \ell_t)\varepsilon_t^\sigma.$$

Here ε is uncorrelated and identically distributed as usual.

As before, ε has density ψ.

The production function is

$$y_t = A_t F(k_t, \ell_t) \varepsilon_t^{\sigma}.$$

Here ε is uncorrelated and identically distributed as usual.

As before, ε has density ψ.

The exponent σ is ≥ 0.

The production function is

$$y_t = A_t F(k_t, \ell_t) \varepsilon_t^{\sigma}.$$

Here ε is uncorrelated and identically distributed as usual.

As before, ε has density ψ.

The exponent σ is ≥ 0.

Note nonstationarity of technology as a result of A_t.

Back to top

Let $f(k) = F(k, 1)$.

Let $f(k) = F(k, 1)$.

Assumption. The function $f\colon [0, \infty) \to [0, \infty)$ satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

Let $f(k) = F(k, 1)$.

Assumption. The function $f\colon [0, \infty) \to [0, \infty)$ satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

In the overlapping generations model, additional assumptions are required to prevent the economy collapsing to zero output.

Let $f(k) = F(k, 1)$.

Assumption. The function $f\colon [0, \infty) \to [0, \infty)$ satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

In the overlapping generations model, additional assumptions are required to prevent the economy collapsing to zero output.

Assumption. The inequality $kf'(k) \leq \lambda f(k)$ holds everywhere for some $\lambda < 1$.

Let $f(k) = F(k,1)$.

Assumption. The function $f\colon [0,\infty) \to [0,\infty)$ satisfies $f(0) = 0$, $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$.

In the overlapping generations model, additional assumptions are required to prevent the economy collapsing to zero output.

Assumption. The inequality $kf'(k) \leq \lambda f(k)$ holds everywhere for some $\lambda < 1$.

This says that the capital share of income cannot become arbitrarily close to total income.

Back to top

Optimization:

Optimization:

Assumption. consumer maximize

$$U(c_t, c'_{t+1}) = \ln c_t + \beta \mathbb{E} \ln c'_{t+1},$$

Optimization:

Assumption. consumer maximize

$$U(c_t, c'_{t+1}) = \ln c_t + \beta \mathbb{E} \ln c'_{t+1},$$

subject to

$$c'_{t+1} = (w_t - c_t)(1 + r_{t+1}).$$

Optimization:

Assumption. consumer maximize

$$U(c_t, c'_{t+1}) = \ln c_t + \beta \mathbb{E} \ln c'_{t+1},$$

subject to

$$c'_{t+1} = (w_t - c_t)(1 + r_{t+1}).$$

Implies total savings is $\beta/(1+\beta)$ times w_t.

Optimization:

Assumption. consumer maximize

$$U(c_t, c'_{t+1}) = \ln c_t + \beta \mathbb{E} \ln c'_{t+1},$$

subject to

$$c'_{t+1} = (w_t - c_t)(1 + r_{t+1}).$$

Implies total savings is $\beta/(1+\beta)$ times w_t.

Competitive factor markets imply that $w_t = A_t[f(k_t) - k_t f'(k_t)]\varepsilon_t^\sigma$.

Back to top

Regarding process (A_t), technology depends on economy-wide aggregates;

Regarding process (A_t), technology depends on economy-wide aggregates;

Thus $A_t = A(k_t)$.

Regarding process (A_t), technology depends on economy-wide aggregates;

Thus $A_t = A(k_t)$.

In development, influence of spillovers may be local (Azariadis and Drazen; Murphy, Shleifer and Vishny, JPE 1989; etc.).

Regarding process (A_t), technology depends on economy-wide aggregates;

Thus $A_t = A(k_t)$.

In development, influence of spillovers may be local (Azariadis and Drazen; Murphy, Shleifer and Vishny, JPE 1989; etc.).

Assumption. $\operatorname{rng} A(k)$ compact.

Regarding process (A_t), technology depends on economy-wide aggregates;

Thus $A_t = A(k_t)$.

In development, influence of spillovers may be local (Azariadis and Drazen; Murphy, Shleifer and Vishny, JPE 1989; etc.).

Assumption. $\operatorname{rng} A(k)$ compact.

Law of motion is

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{56}$$

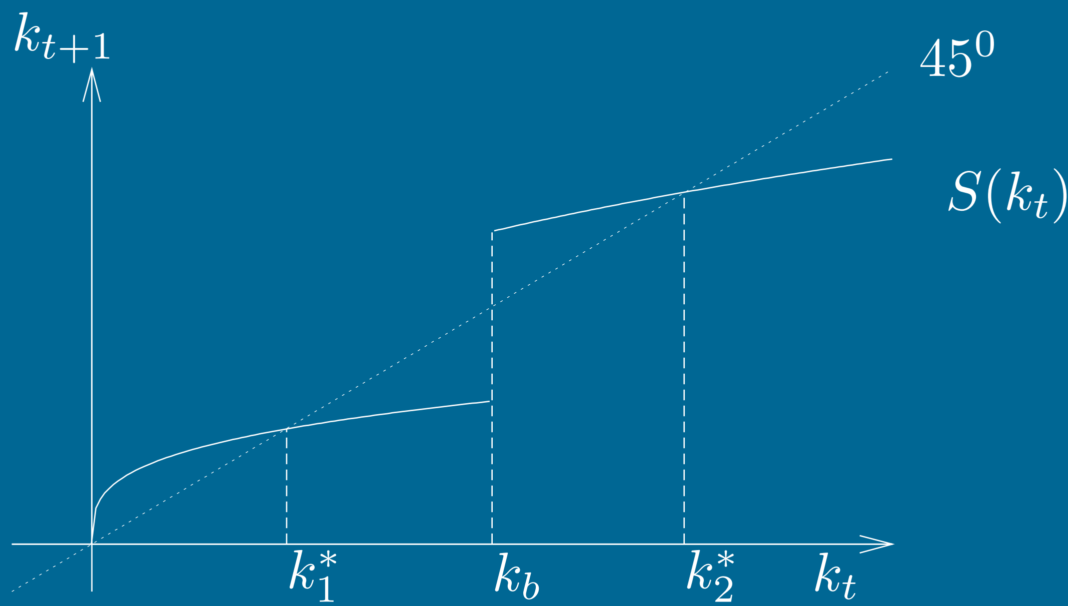
where $S(k) = DA(k)[f(k) - kf'(k)]$.

For example, the “critical mass” form

$$A(k) = A_1 \cdot \mathbf{1}_{[0,k_b)}(k) + A_2 \cdot \mathbf{1}_{[k_b,\infty)}(k),$$

For example, the "critical mass" form

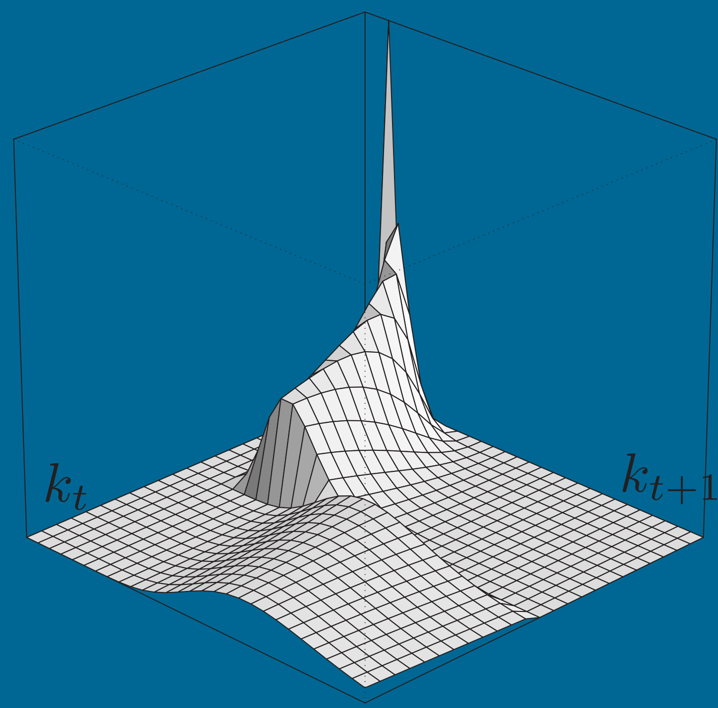
$$A(k) = A_1 \cdot \mathbf{1}_{[0,k_b)}(k) + A_2 \cdot \mathbf{1}_{[k_b,\infty)}(k),$$



Back to top

For lognormal shock, $\Gamma(k, k')$ looks like...

For lognormal shock, $\Gamma(k, k')$ looks like...



Back to top

Regarding the dynamics of

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{57}$$

Regarding the dynamics of

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{57}$$

Proposition. If $\sigma = 0$ then there may be multiple, locally stable equilibria.

Regarding the dynamics of

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{57}$$

Proposition. If $\sigma = 0$ then there may be multiple, locally stable equilibria. However, for every $\sigma > 0$, there is a single, unique (stochastic) equilibrium.

Regarding the dynamics of

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{57}$$

Proposition. If $\sigma = 0$ then there may be multiple, locally stable equilibria. However, for every $\sigma > 0$, there is a single, unique (stochastic) equilibrium.

Proof: we will show that for any $\sigma > 0$, $(\mathscr{D}(\mu), P_\sigma)$ is strongly contracting and Lagrange stable.

Back to top

We prove first strong contractiveness.

We prove first strong contractiveness.

Recall that a sufficent condition is $\Gamma(k, k') > 0$ everywhere.

We prove first strong contractiveness.

Recall that a sufficent condition is $\Gamma(k, k') > 0$ everywhere.

But if $\varepsilon \sim \psi$, then

$$\Gamma_\sigma(k, k') = \psi \left[\left(\frac{k'}{S(k)} \right)^{\frac{1}{\sigma}} \right] \left(\frac{k'}{S(k)} \right)^{\frac{1}{\sigma}} \frac{1}{\sigma k'}, \tag{58}$$

whence $\psi > 0 \implies \Gamma > 0$.

We prove first strong contractiveness.

Recall that a sufficent condition is $\Gamma(k, k') > 0$ everywhere.

But if $\varepsilon \sim \psi$, then

$$\Gamma_\sigma(k, k') = \psi\left[\left(\frac{k'}{S(k)}\right)^{\frac{1}{\sigma}}\right]\left(\frac{k'}{S(k)}\right)^{\frac{1}{\sigma}}\frac{1}{\sigma k'}, \tag{58}$$

whence $\psi > 0 \implies \Gamma > 0$.

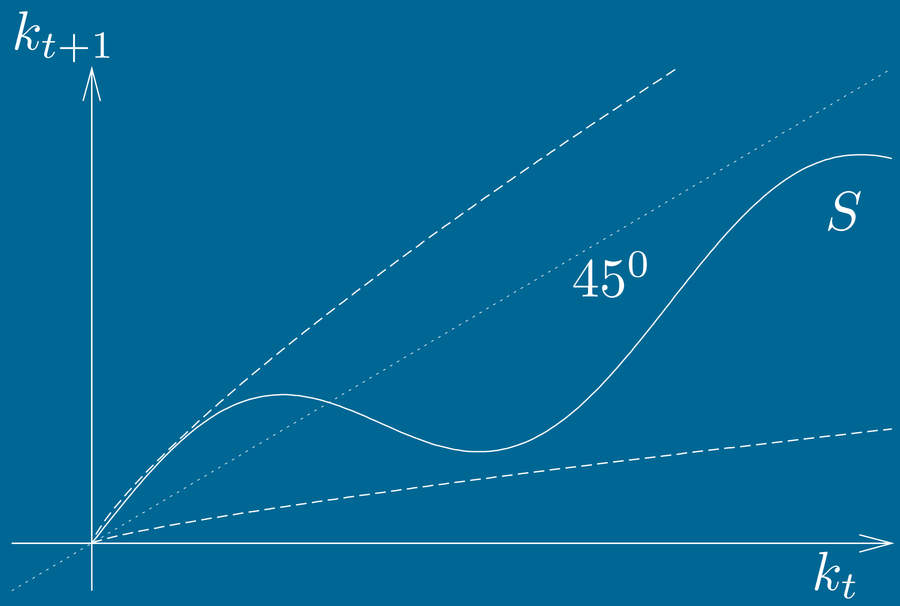
This is all we need for uniqueness!

Back to top

As for Lagrange stability, we prove it using stronger assumptions.

As for Lagrange stability, we prove it using stronger assumptions.

Let us assume that



More formally, given

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{59}$$

assume that exists α, β_i positive with $\alpha < 1$,

$$\beta_1 k^{\alpha} \leq S(k) \leq \beta_2 k^{\alpha}.$$

More formally, given

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{59}$$

assume that exists α, β_i positive with $\alpha < 1$,

$$\beta_1 k^{\alpha} \leq S(k) \leq \beta_2 k^{\alpha}.$$

Then Lagrange stable.

Back to top

Recall the

Lasota-Stachurski Theorem.

Recall the

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x, y) \leq h(y)$, and

2. $\{P^t \varphi\}$ is tight,

then $\{P^t \varphi\}$ is precompact

Recall the

Lasota-Stachurski Theorem. Let P be a M.O. on $\mathscr{D}(\mu)$ with kernel Γ. Let $\varphi \in \mathscr{D}(\mu)$. If

1. exists continuous h such that $\Gamma(x,y) \leq h(y)$, and

2. $\{P^t \varphi\}$ is tight,

then $\{P^t \varphi\}$ is precompact

Condition 1 holds again for the lognormal shock (we will not check it).

Back to top

Regarding condition 2 (tightness):

Regarding condition 2 (tightness):

Chebychev inequality:

$$\text{Prob}(|\ln k_t| \geq r) \leq \frac{\mathbb{E}|\ln k_t|}{r}, \quad \forall r > 0. \tag{60}$$

Regarding condition 2 (tightness):

Chebychev inequality:

$$\text{Prob}(|\ln k_t| \geq r) \leq \frac{\mathbb{E}|\ln k_t|}{r}, \quad \forall r > 0. \tag{60}$$

Therefore, for each t,

$$\int_{[\exp(-r),\exp(r)]^c} P^t \varphi(k) dk \leq \frac{\mathbb{E}_\varphi |\ln k_t|}{r}, \quad \forall r > 0. \tag{61}$$

Regarding condition 2 (tightness):

Chebychev inequality:

$$\text{Prob}(|\ln k_t| \geq r) \leq \frac{\mathbb{E}|\ln k_t|}{r}, \quad \forall r > 0. \tag{60}$$

Therefore, for each t,

$$\int_{[\exp(-r),\exp(r)]^c} P^t \varphi(k) dk \leq \frac{\mathbb{E}_\varphi |\ln k_t|}{r}, \quad \forall r > 0. \tag{61}$$

If $\exists M < \infty$ s.t. $\mathbb{E}_\varphi |\ln k_t| \leq M$ for all t then we are done.

We have

$$\mathbb{E}_{\varphi}|\ln k_t| = \int_0^{\infty} \mathbb{E}(|\ln k_t| \text{ given } \; k_{t-1} = k)\,\mathrm{Prob}(k_{t-1} = k)dk$$

We have

$$\mathbb{E}_\varphi|\ln k_t| = \int_0^\infty \mathbb{E}(|\ln k_t| \text{ given } \ k_{t-1} = k)\,\mathrm{Prob}(k_{t-1} = k)dk$$
$$= \int_0^\infty \mathbb{E}_\varepsilon|\ln[S(k)\varepsilon^\sigma]|P^{t-1}\varphi(k)dk$$

We have

$$\begin{aligned}\mathbb{E}_{\varphi}|\ln k_t| &= \int_0^{\infty} \mathbb{E}(|\ln k_t| \text{ given } \; k_{t-1} = k)\,\mathrm{Prob}(k_{t-1} = k)dk \\ &= \int_0^{\infty} \mathbb{E}_{\varepsilon}|\ln[S(k)\varepsilon^{\sigma}]|P^{t-1}\varphi(k)dk \\ &\leq \int_0^{\infty} \{|\ln S(k)| + \sigma\mathbb{E}_{\varepsilon}|\ln\varepsilon|\}P^{t-1}\varphi(k)dk.\end{aligned}$$

We have

$$\mathbb{E}_\varphi |\ln k_t| = \int_0^\infty \mathbb{E}(|\ln k_t| \text{ given } \; k_{t-1} = k)\,\mathrm{Prob}(k_{t-1} = k) dk$$
$$= \int_0^\infty \mathbb{E}_\varepsilon |\ln[S(k)\varepsilon^\sigma]| P^{t-1}\varphi(k) dk$$
$$\leq \int_0^\infty \{|\ln S(k)| + \sigma \mathbb{E}_\varepsilon |\ln \varepsilon|\} P^{t-1}\varphi(k) dk.$$
$$\leq \int_0^\infty |\ln S(k)| P^{t-1}\varphi(k) dk + \sigma \mathbb{E}_\varepsilon |\ln \varepsilon|.$$

We have

$$\begin{aligned}\mathbb{E}_{\varphi}|\ln k_t| &= \int_0^{\infty} \mathbb{E}(|\ln k_t| \text{ given } \; k_{t-1} = k)\,\mathrm{Prob}(k_{t-1} = k)dk \\ &= \int_0^{\infty} \mathbb{E}_{\varepsilon}|\ln[S(k)\varepsilon^{\sigma}]|P^{t-1}\varphi(k)dk \\ &\leq \int_0^{\infty} \{|\ln S(k)| + \sigma\mathbb{E}_{\varepsilon}|\ln\varepsilon|\}P^{t-1}\varphi(k)dk. \\ &\leq \int_0^{\infty} |\ln S(k)|P^{t-1}\varphi(k)dk + \sigma\mathbb{E}_{\varepsilon}|\ln\varepsilon|.\end{aligned}$$

Note also that

Back to top

$$\beta_1 k^\alpha \leq S(k) \leq \beta_2 k^\alpha \implies |\ln S(k)| \leq \alpha |\ln k| + M.$$

Back to top

To repeat,

$$\mathbb{E}_{\varphi}|\ln k_t| \leq \int |\ln S(k)|P^{t-1}\varphi(k)dk + \sigma \mathbb{E}_{\varepsilon}|\ln \varepsilon|.$$

To repeat,

$$\mathbb{E}_\varphi |\ln k_t| \leq \int |\ln S(k)| P^{t-1}\varphi(k)dk + \sigma \mathbb{E}_\varepsilon |\ln \varepsilon|.$$

Hence $|\ln S(k)| \leq \alpha |\ln k| + M$ implies

$$\mathbb{E}_\varphi |\ln k_t| \leq \int \alpha |\ln k| P^{t-1}\varphi(k)dk + C$$

To repeat,

$$\mathbb{E}_\varphi |\ln k_t| \leq \int |\ln S(k)| P^{t-1}\varphi(k)dk + \sigma \mathbb{E}_\varepsilon |\ln \varepsilon|.$$

Hence $|\ln S(k)| \leq \alpha |\ln k| + M$ implies

$$\begin{aligned}\mathbb{E}_\varphi |\ln k_t| &\leq \int \alpha |\ln k| P^{t-1}\varphi(k)dk + C \\ &= \alpha \mathbb{E}_\varphi |\ln k_{t-1}| + C.\end{aligned}$$

To repeat,

$$\mathbb{E}_\varphi |\ln k_t| \le \int |\ln S(k)| P^{t-1} \varphi(k) dk + \sigma \mathbb{E}_\varepsilon |\ln \varepsilon|.$$

Hence $|\ln S(k)| \le \alpha |\ln k| + M$ implies

$$\begin{aligned} \mathbb{E}_\varphi |\ln k_t| &\le \int \alpha |\ln k| P^{t-1} \varphi(k) dk + C \\ &= \alpha \mathbb{E}_\varphi |\ln k_{t-1}| + C. \end{aligned}$$

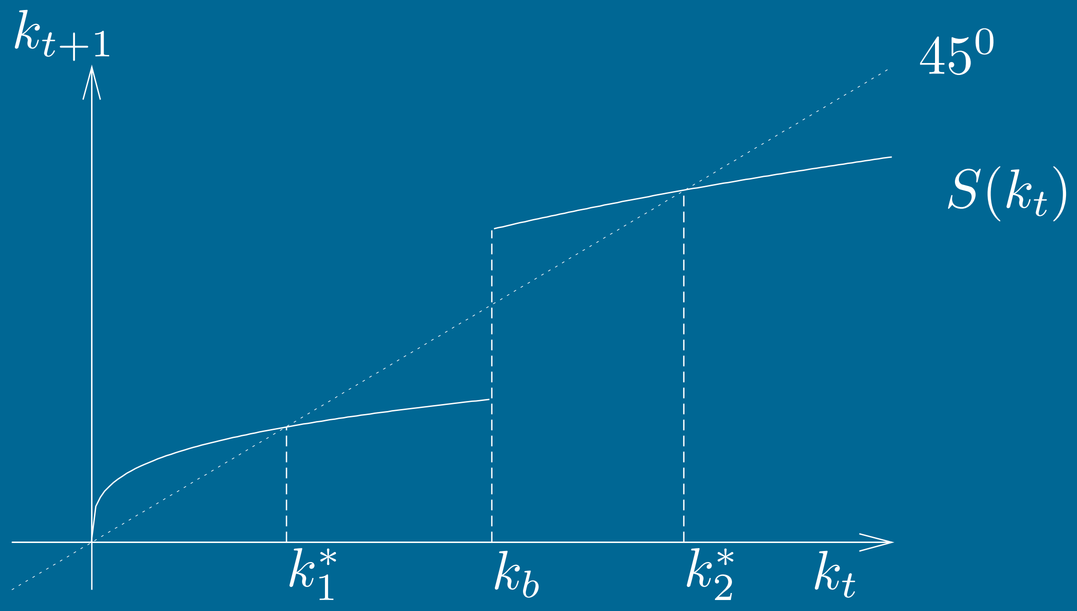
Which is sufficient for the proof.

Back to top

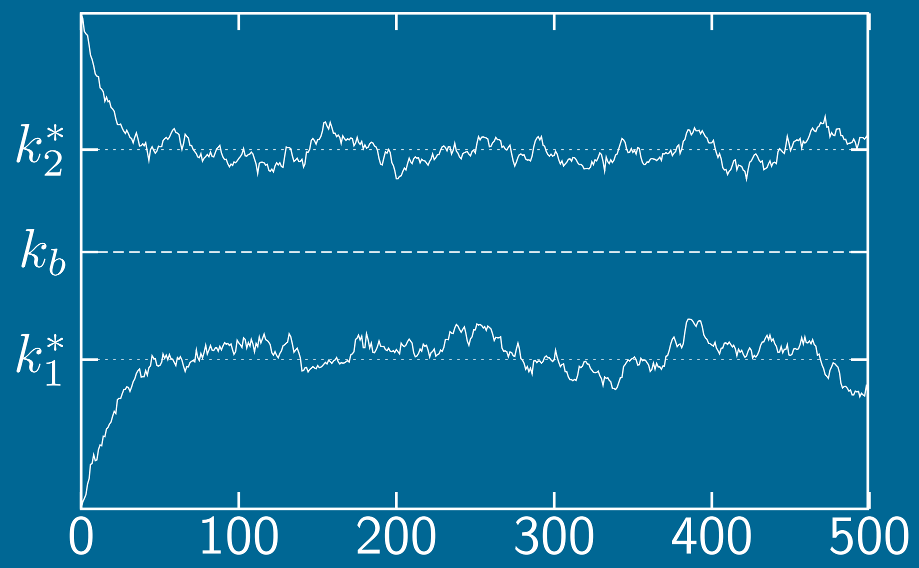
Notice our proof works for any noise level σ.

Notice our proof works for any noise level σ.

This is suprising, because when noise level is nearly zero, behavior must be similar to



Indeed this is the case: History does not matter, but



Back to top

Part 6: Empirics of Stochastic Growth

Back to top

Some stylized facts

In the theoretical part of this seminar we have looked at evolution of densities generated by growth models.

Some stylized facts

In the theoretical part of this seminar we have looked at evolution of densities generated by growth models.

In this section we look at the actual evolution of densities, and try to predict future evolution.

Some stylized facts

In the theoretical part of this seminar we have looked at evolution of densities generated by growth models.

In this section we look at the actual evolution of densities, and try to predict future evolution.

To do so we must choose an appropriate model.

Some stylized facts

In the theoretical part of this seminar we have looked at evolution of densities generated by growth models.

In this section we look at the actual evolution of densities, and try to predict future evolution.

To do so we must choose an appropriate model.

The convex neoclassical model is rejected.

Some stylized facts

In the theoretical part of this seminar we have looked at evolution of densities generated by growth models.

In this section we look at the actual evolution of densities, and try to predict future evolution.

To do so we must choose an appropriate model.

The convex neoclassical model is rejected.

Instead, we fit and predict with the increasing returns Azariadis-Drazen model.

Back to top

Begin with stylized facts: (Lant Pritchett JEP, 11(3), 1997)

Begin with stylized facts: (Lant Pritchett JEP, 11(3), 1997)

1. What accounts for continued per capita growth and technological progress of those leading countries at the frontier?

Begin with stylized facts: (Lant Pritchett JEP, 11(3), 1997)

1. What accounts for continued per capita growth and technological progress of those leading countries at the frontier?

2. What accounts for the few countries that are able to initiate and sustain periods of rapid growth in which they gain significantly on the leaders?

Begin with stylized facts: (Lant Pritchett JEP, 11(3), 1997)

1. What accounts for continued per capita growth and technological progress of those leading countries at the frontier?

2. What accounts for the few countries that are able to initiate and sustain periods of rapid growth in which they gain significantly on the leaders?

3. What accounts for why some countries fade and lose the momentum of rapid growth?

Begin with stylized facts: (Lant Pritchett JEP, 11(3), 1997)

1. What accounts for continued per capita growth and technological progress of those leading countries at the frontier?

2. What accounts for the few countries that are able to initiate and sustain periods of rapid growth in which they gain significantly on the leaders?

3. What accounts for why some countries fade and lose the momentum of rapid growth?

4. What accounts for why some countries remain in low growth for very long periods?

Also, we can look at the data: (countries with per capita GDP data for both 1960 and 1990 in the Summers and Heston (1991) PWT5.6)

Also, we can look at the data: (countries with per capita GDP data for both 1960 and 1990 in the Summers and Heston (1991) PWT5.6)

1. 12 countries that saw income *decline* between 1960 and 1990.

Also, we can look at the data: (countries with per capita GDP data for both 1960 and 1990 in the Summers and Heston (1991) PWT5.6)

1. 12 countries that saw income *decline* between 1960 and 1990. Of these 11 poor–very poor (ave 1960 per capita GDP \$968 1985 US dollars).

Also, we can look at the data: (countries with per capita GDP data for both 1960 and 1990 in the Summers and Heston (1991) PWT5.6)

1. 12 countries that saw income *decline* between 1960 and 1990. Of these 11 poor–very poor (ave 1960 per capita GDP \$968 1985 US dollars).

2. 12 richest countries ($\geq$ \$6,000 in 1960) generally grew steadily, increasing output per person in a narrow band around the average of just under factor of 2.

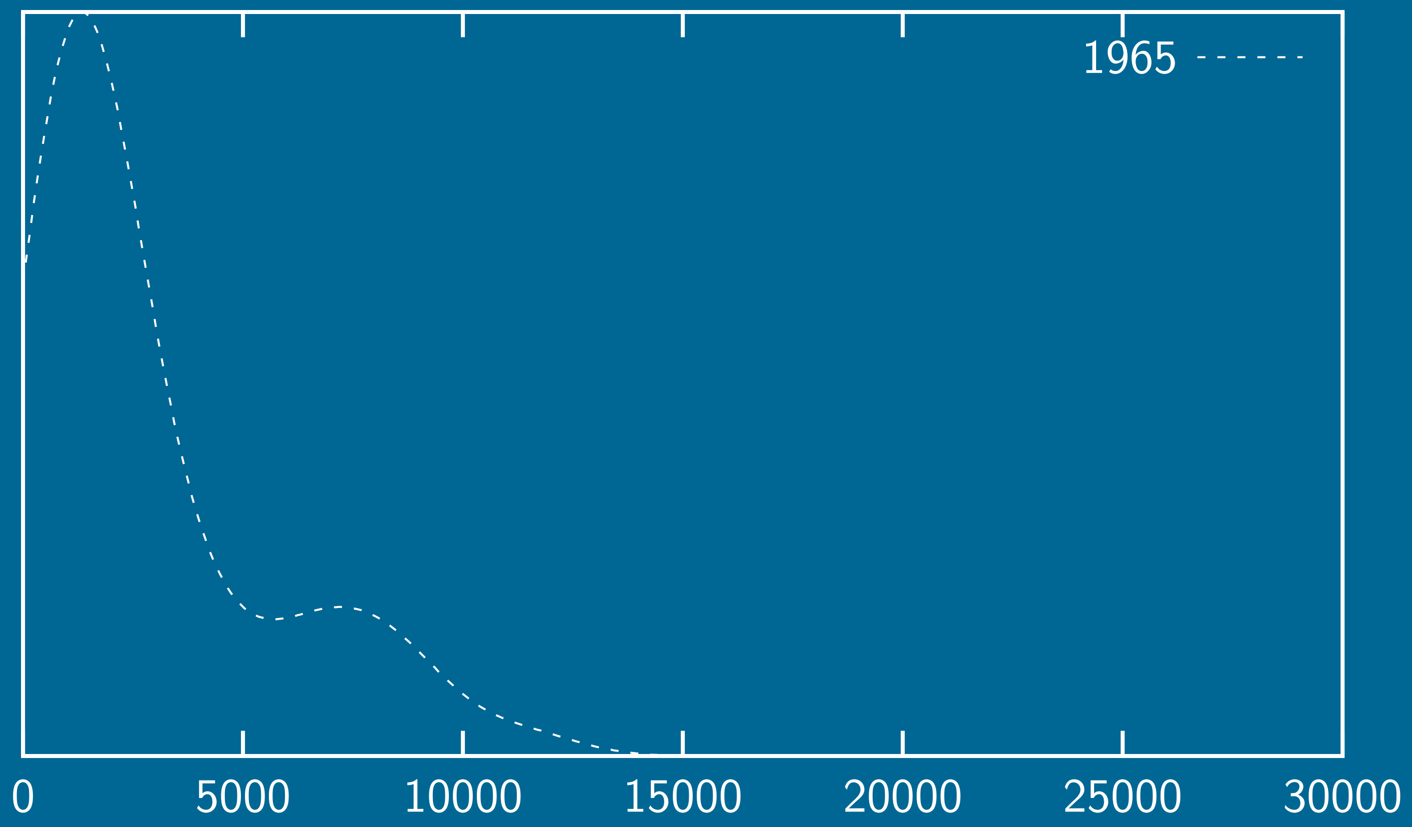
Also, we can look at the data: (countries with per capita GDP data for both 1960 and 1990 in the Summers and Heston (1991) PWT5.6)

1. 12 countries that saw income *decline* between 1960 and 1990. Of these 11 poor–very poor (ave 1960 per capita GDP \$968 1985 US dollars).

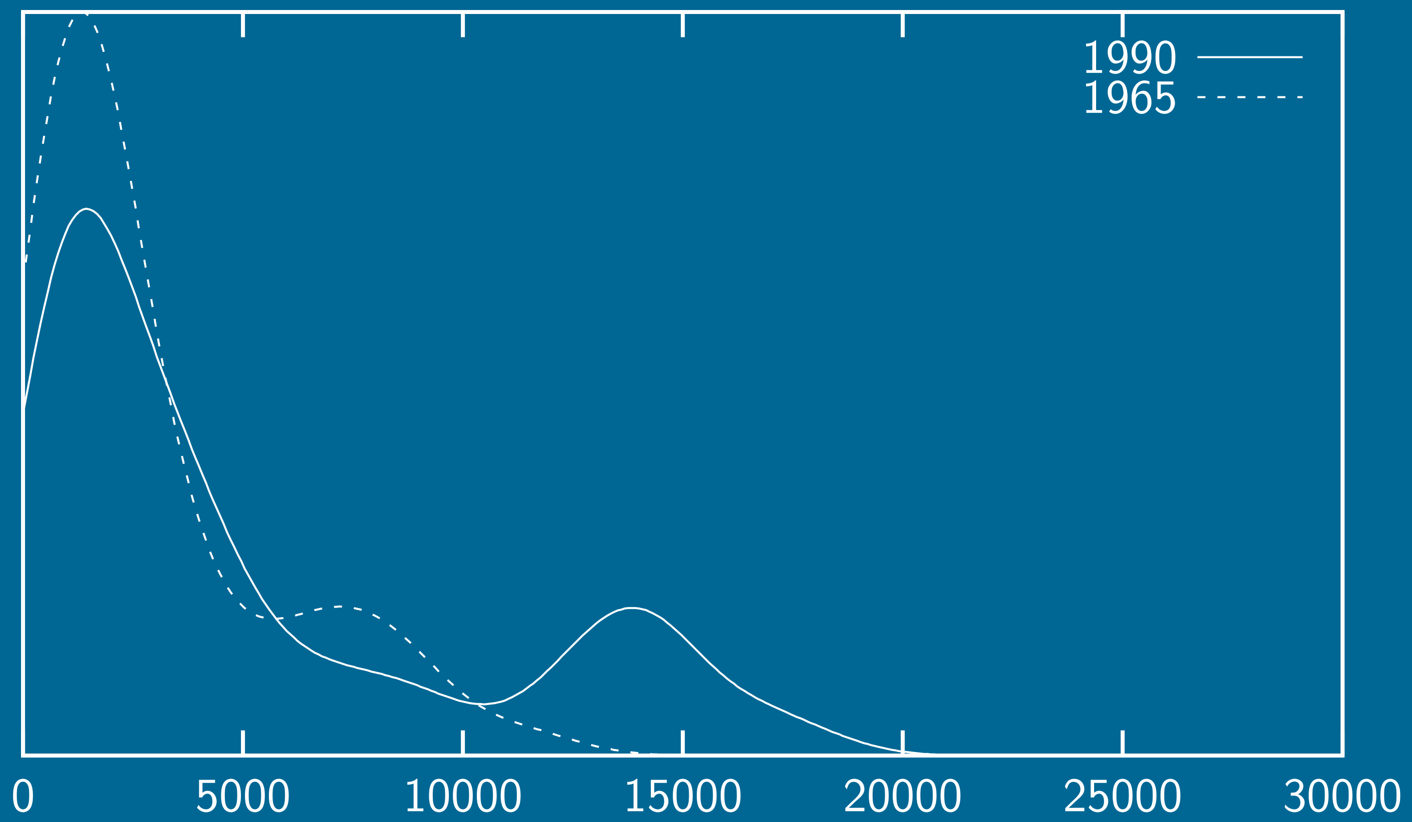
2. 12 richest countries ($\geq$ \$6,000 in 1960) generally grew steadily, increasing output per person in a narrow band around the average of just under factor of 2.

3. Regarding rapid take-off, some poor–middle income countries grew *much* faster than rich countries (S. Korea $\uparrow$ by factor of 7.4; Singapore, 7.1; Hong Kong, 6.6; Taiwan, 6.4; Japan, 4.9; Portugal, 4.2).

Back to top



Back to top



Back to top

Fitting the convex model

Fitting the convex model

In the simplest possible model,

$$k_{t+1} = s k_t^{\alpha} \varepsilon_t.$$

Fitting the convex model

In the simplest possible model,

$$k_{t+1} = s k_t^{\alpha} \varepsilon_t. \tag{62}$$

Suppose further that $k_0 \sim \varphi_0$, where φ_0 is lognormal $(0, \sigma_\varepsilon^2)$.

Fitting the convex model

In the simplest possible model,

$$k_{t+1} = sk_t^{\alpha}\varepsilon_t. \tag{62}$$

Suppose further that $k_0 \sim \varphi_0$, where φ_0 is lognormal $(0, \sigma_\varepsilon^2)$.

Then *all of trajectory is lognormal*

Fitting the convex model

In the simplest possible model,

$$k_{t+1} = s k_t^{\alpha} \varepsilon_t. \tag{62}$$

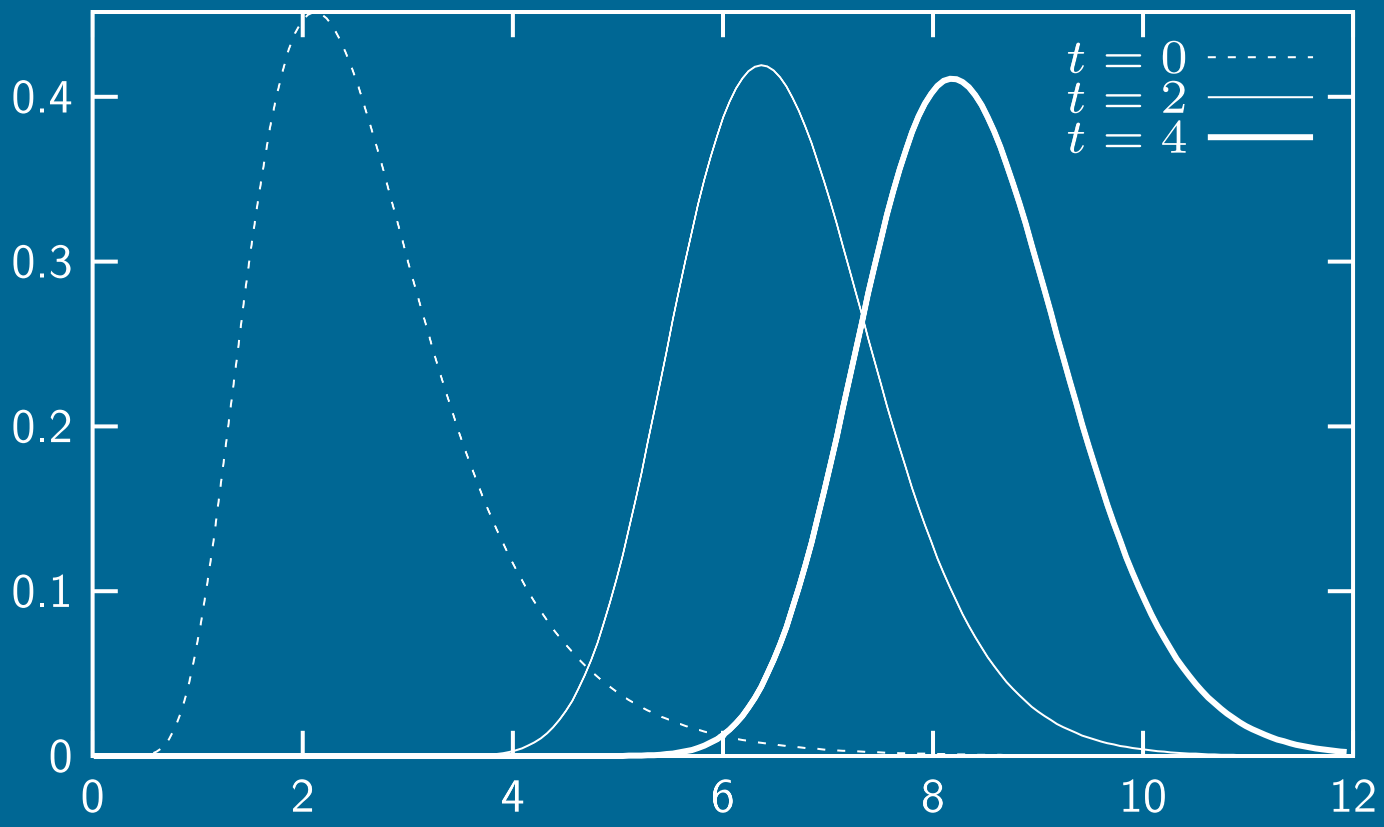
Suppose further that $k_0 \sim \varphi_0$, where φ_0 is lognormal $(0, \sigma_\varepsilon^2)$.

Then *all of trajectory is lognormal* with parameters

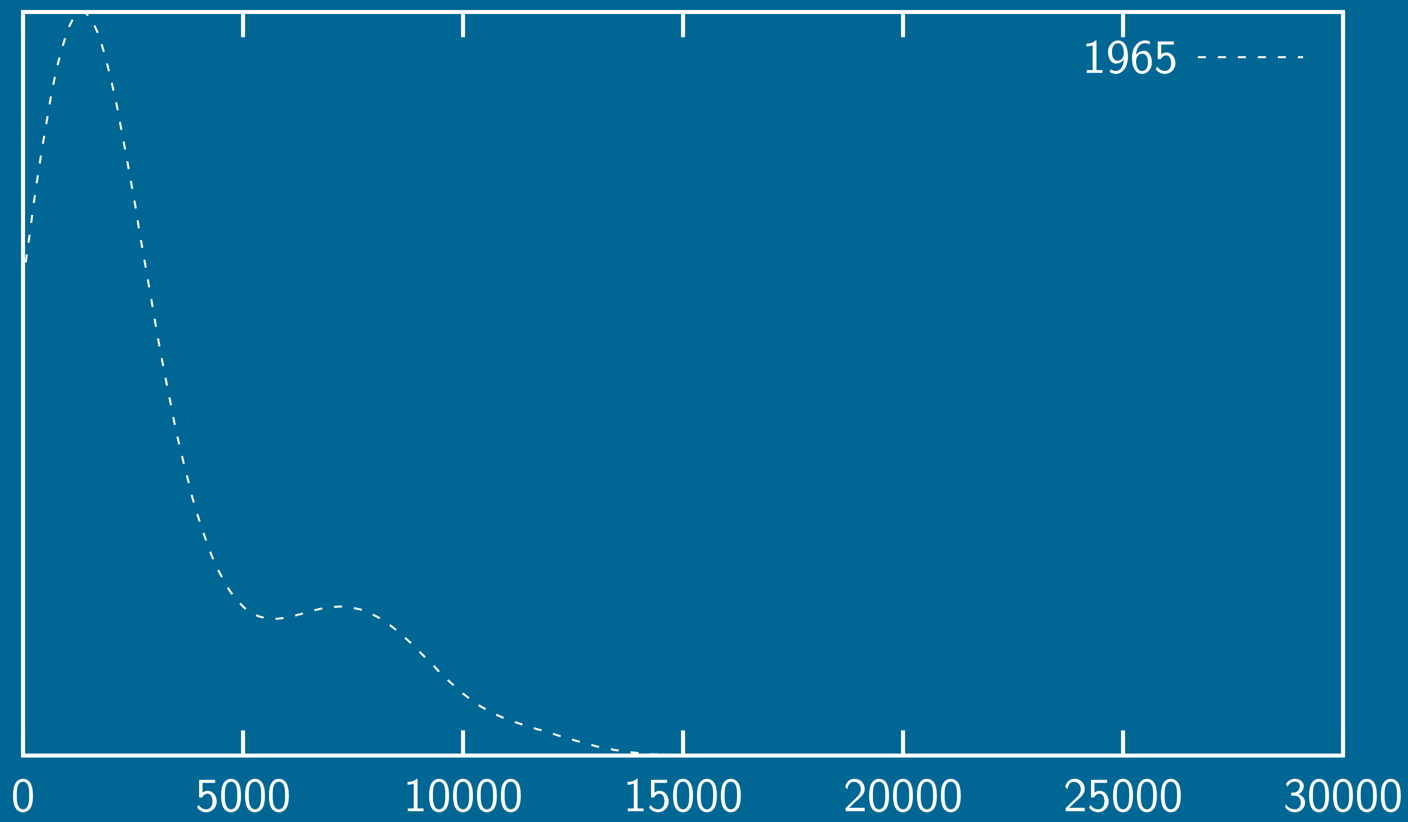
$$\mu_{t+1} = \ln s + \alpha \mu_t, \quad \mu_0 = \mathbb{E} k_0. \tag{63}$$

$$\sigma_{t+1}^2 = \sigma_\varepsilon^2 + \alpha^2 \sigma_t^2, \quad \sigma_0^2 = \mathbb{V} k_0. \tag{64}$$

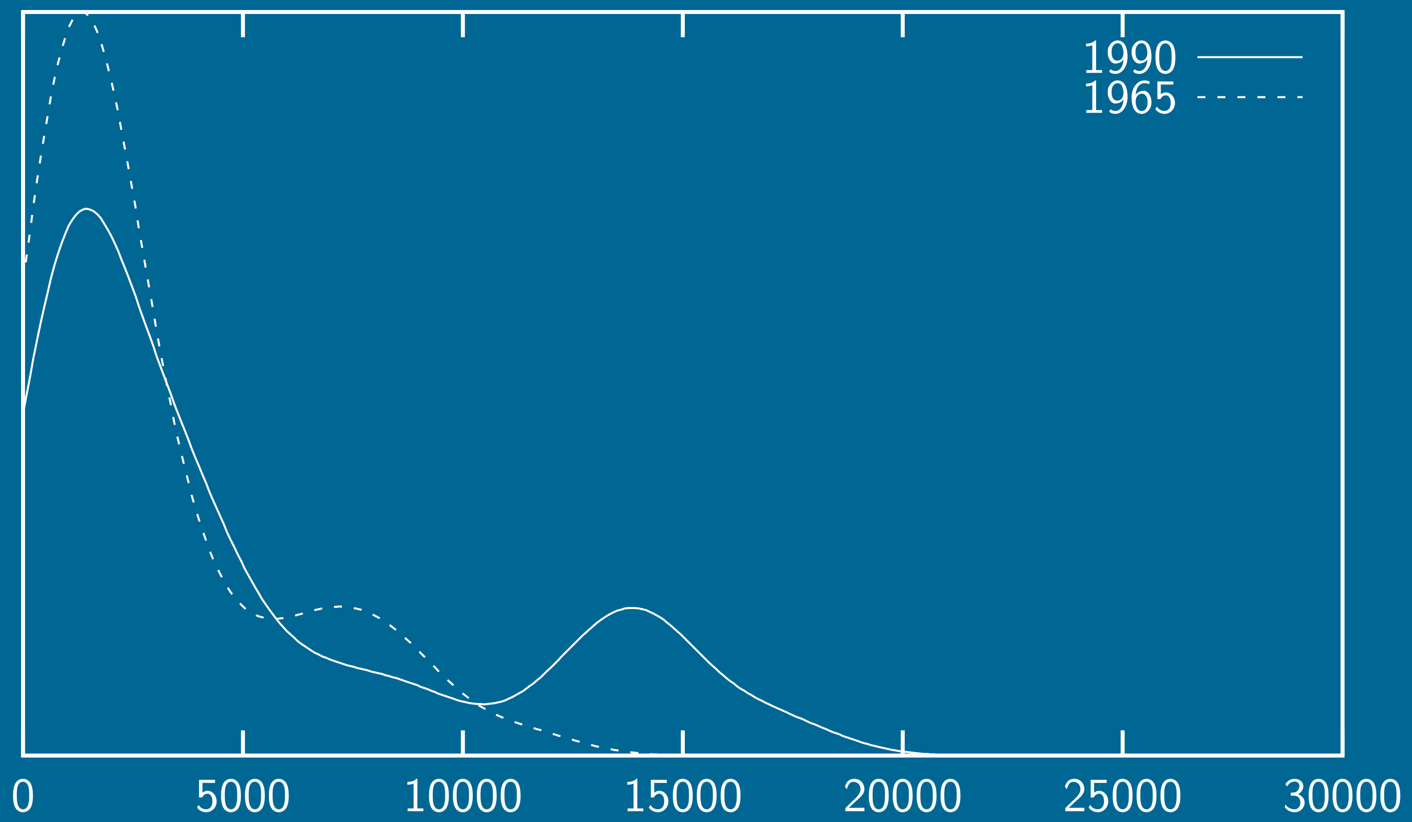
Back to top



Back to top



Back to top



Back to top

Fitting the Azariadis-Drazen model

Fitting the Azariadis-Drazen model

Why does the neoclassical model fail?

Fitting the Azariadis-Drazen model

Why does the neoclassical model fail?

The main reason is that poverty exhibits severe persistence.

Fitting the Azariadis-Drazen model

Why does the neoclassical model fail?

The main reason is that poverty exhibits severe persistence.

Also, fastest growth rates are found in the middle income countries.

Fitting the Azariadis-Drazen model

Why does the neoclassical model fail?

The main reason is that poverty exhibits severe persistence.

Also, fastest growth rates are found in the middle income countries.

For the same reasons, the Azariadis-Drazen model might be a better fit.

Fitting the Azariadis-Drazen model

Why does the neoclassical model fail?

The main reason is that poverty exhibits severe persistence.

Also, fastest growth rates are found in the middle income countries.

For the same reasons, the Azariadis-Drazen model might be a better fit.

Here we fit, and then use the operator P to project densities into the future.

Back to top

Recall the Azariadis-Drazen model.

Recall the Azariadis-Drazen model.

1. overlapping generations,

Recall the Azariadis-Drazen model.

1. overlapping generations,

2. threshold externalities.

Recall the Azariadis-Drazen model.

1. overlapping generations,

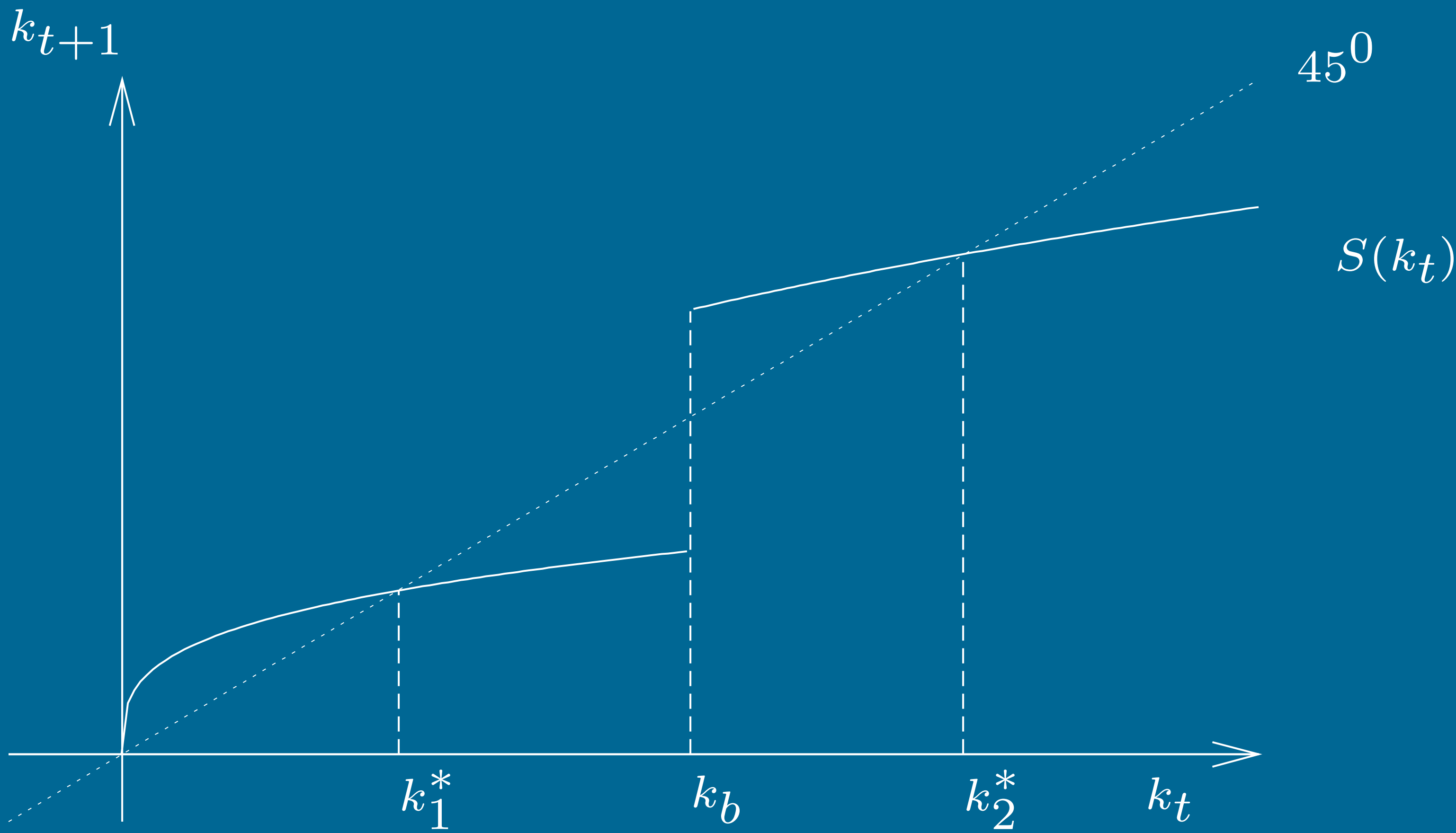
2. threshold externalities.

In the Cobb-Douglas case,

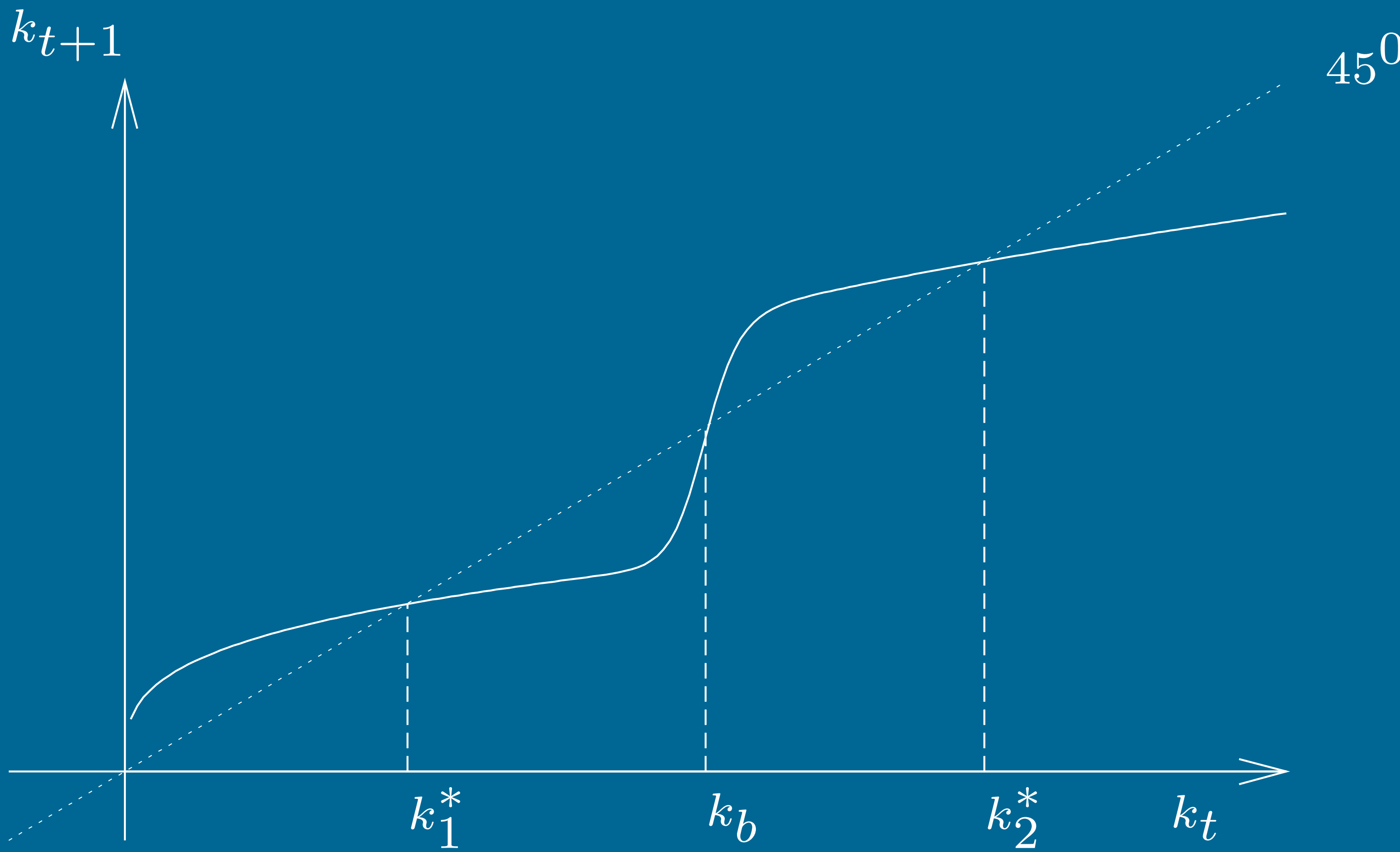
$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{65}$$

where, $S(k) = D(1-\alpha)A(k)k^{\alpha}$.

Back to top



Back to top



k_{t+1}
45^0
k_1^*
k_b
k_2^*
k_t

This models can be fitted using nonlinear regression.

This models can be fitted using nonlinear regression.

The data is Penn World Tables version 5.6, GDP per capita.

This models can be fitted using nonlinear regression.

The data is Penn World Tables version 5.6, GDP per capita.

We assume that all countries have the same model

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{66}$$

This models can be fitted using nonlinear regression.

The data is Penn World Tables version 5.6, GDP per capita.

We assume that all countries have the same model

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{66}$$

We let $t = 1965$ and $t + 1 = 1990$ (25 years for OLG).

Back to top

Data is pooled and we regress across all countries i:

$$k^i_{1990} = S(k^i_{1965})\varepsilon^\sigma_{1965}, \quad i = 1, \ldots 105. \tag{67}$$

Data is pooled and we regress across all countries i:

$$k^i_{1990} = S(k^i_{1965})\varepsilon^\sigma_{1965}, \quad i = 1, \ldots 105. \tag{67}$$

Shock is lognormal and we use maximum likelihood.

Data is pooled and we regress across all countries i:

$$k^i_{1990} = S(k^i_{1965})\varepsilon^\sigma_{1965}, \quad i = 1, \dots 105. \tag{67}$$

Shock is lognormal and we use maximum likelihood.

The discontinuous version uses the TAR procedure.

Data is pooled and we regress across all countries i:

$$k^i_{1990} = S(k^i_{1965})\varepsilon^\sigma_{1965}, \quad i = 1, \ldots 105. \tag{67}$$

Shock is lognormal and we use maximum likelihood.

The discontinuous version uses the TAR procedure.

The smooth transition version uses STAR.

Data is pooled and we regress across all countries i:

$$k^i_{1990} = S(k^i_{1965})\varepsilon^{\sigma}_{1965}, \quad i = 1, \ldots 105. \tag{67}$$

Shock is lognormal and we use maximum likelihood.

The discontinuous version uses the TAR procedure.

The smooth transition version uses STAR.

Implementation of the procedure is in Java.

Back to top

Once the model

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{68}$$

is known, we also know the conditional density:

$$\Gamma_{\sigma}(k, k') = \psi\left[\left(\frac{k'}{S(k)}\right)^{\frac{1}{\sigma}}\right]\left(\frac{k'}{S(k)}\right)^{\frac{1}{\sigma}}\frac{1}{\sigma k'}, \tag{69}$$

Once the model

$$k_{t+1} = S(k_t)\varepsilon_t^{\sigma}, \tag{68}$$

is known, we also know the conditional density:

$$\Gamma_{\sigma}(k, k') = \psi\left[\left(\frac{k'}{S(k)}\right)^{\frac{1}{\sigma}}\right]\left(\frac{k'}{S(k)}\right)^{\frac{1}{\sigma}}\frac{1}{\sigma k'}, \tag{69}$$

and hence the Markov operator

$$Pf(k') = \int \Gamma(k, k')f(k)dk.$$

Imagine the following thought experiment.

Imagine the following thought experiment.

There are M countries. All of them have the same dynamics, which are known to us.

Imagine the following thought experiment.

There are M countries. All of them have the same dynamics, which are known to us.

At time 0, the “supervisor” picks the initial income for each country independently from φ_0,

Imagine the following thought experiment.

There are M countries. All of them have the same dynamics, which are known to us.

At time 0, the "supervisor" picks the initial income for each country independently from φ_0, which is unknown to us.

Imagine the following thought experiment.

There are M countries. All of them have the same dynamics, which are known to us.

At time 0, the "supervisor" picks the initial income for each country independently from φ_0, which is unknown to us.

The M initial values y_m^0 are revealed to us.

Imagine the following thought experiment.

There are M countries. All of them have the same dynamics, which are known to us.

At time 0, the "supervisor" picks the initial income for each country independently from φ_0, which is unknown to us.

The M initial values y_m^0 are revealed to us.

The supervisor then updates each country independently to get y_m^1.

Imagine the following thought experiment.

There are M countries. All of them have the same dynamics, which are known to us.

At time 0, the "supervisor" picks the initial income for each country independently from φ_0, which is unknown to us.

The M initial values y_m^0 are revealed to us.

The supervisor then updates each country independently to get y_m^1.

How should we guess the distribution of y_m^1?

Back to top

What is the probability that at time 1 a country has income y?

What is the probability that at time 1 a country has income y?

By the law of total probability,

$$\varphi_1(y) = \int \Gamma(x, y)\varphi_0(x)dx.$$

What is the probability that at time 1 a country has income y?

By the law of total probability,

$$\varphi_1(y) = \int \Gamma(x, y)\varphi_0(x)dx.$$

Thus $\varphi_1 = P\varphi_0$.

What is the probability that at time 1 a country has income y?

By the law of total probability,

$$\varphi_1(y) = \int \Gamma(x, y)\varphi_0(x)dx.$$

Thus $\varphi_1 = P\varphi_0$.

But φ_0 is unknown to us...

Back to top

However, we know the initial distribution of income, which was drawn from φ_0.

However, we know the initial distribution of income, which was drawn from φ_0.

Hence we can estimate φ_0 with $\hat{\varphi}_0$.

However, we know the initial distribution of income, which was drawn from φ_0.

Hence we can estimate φ_0 with $\hat{\varphi}_0$.

Suppose that $\hat{\varphi}_0 \to \varphi_0$ in $\mathscr{L}_1(\mu)$ norm as sample size becomes large.

However, we know the initial distribution of income, which was drawn from φ_0.

Hence we can estimate φ_0 with $\hat{\varphi}_0$.

Suppose that $\hat{\varphi}_0 \to \varphi_0$ in $\mathscr{L}_1(\mu)$ norm as sample size becomes large.

Then by norm-continuity of P,

However, we know the initial distribution of income, which was drawn from φ_0.

Hence we can estimate φ_0 with $\hat{\varphi}_0$.

Suppose that $\hat{\varphi}_0 \to \varphi_0$ in $\mathscr{L}_1(\mu)$ norm as sample size becomes large.

Then by norm-continuity of P,

$$P\hat{\varphi}_0 \to P\varphi_0 = \varphi_1.$$

Back to top

In the next two slides we present $P^3\varphi_{1990}$ and $P^5\varphi_{1990}$.

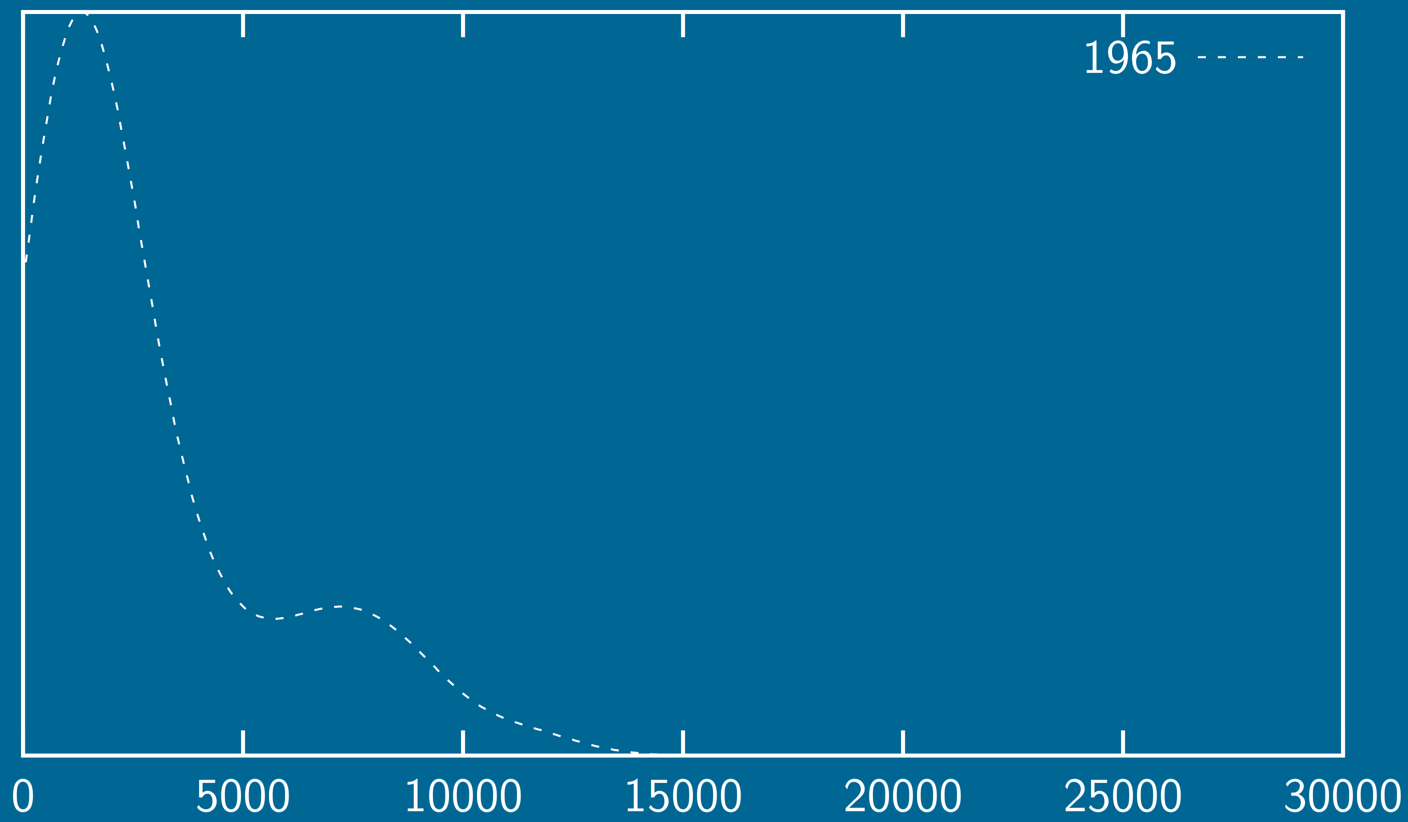
In the next two slides we present $P^3\varphi_{1990}$ and $P^5\varphi_{1990}$.

This gives 75 year and 125 year projections for the cross-country income distribution.

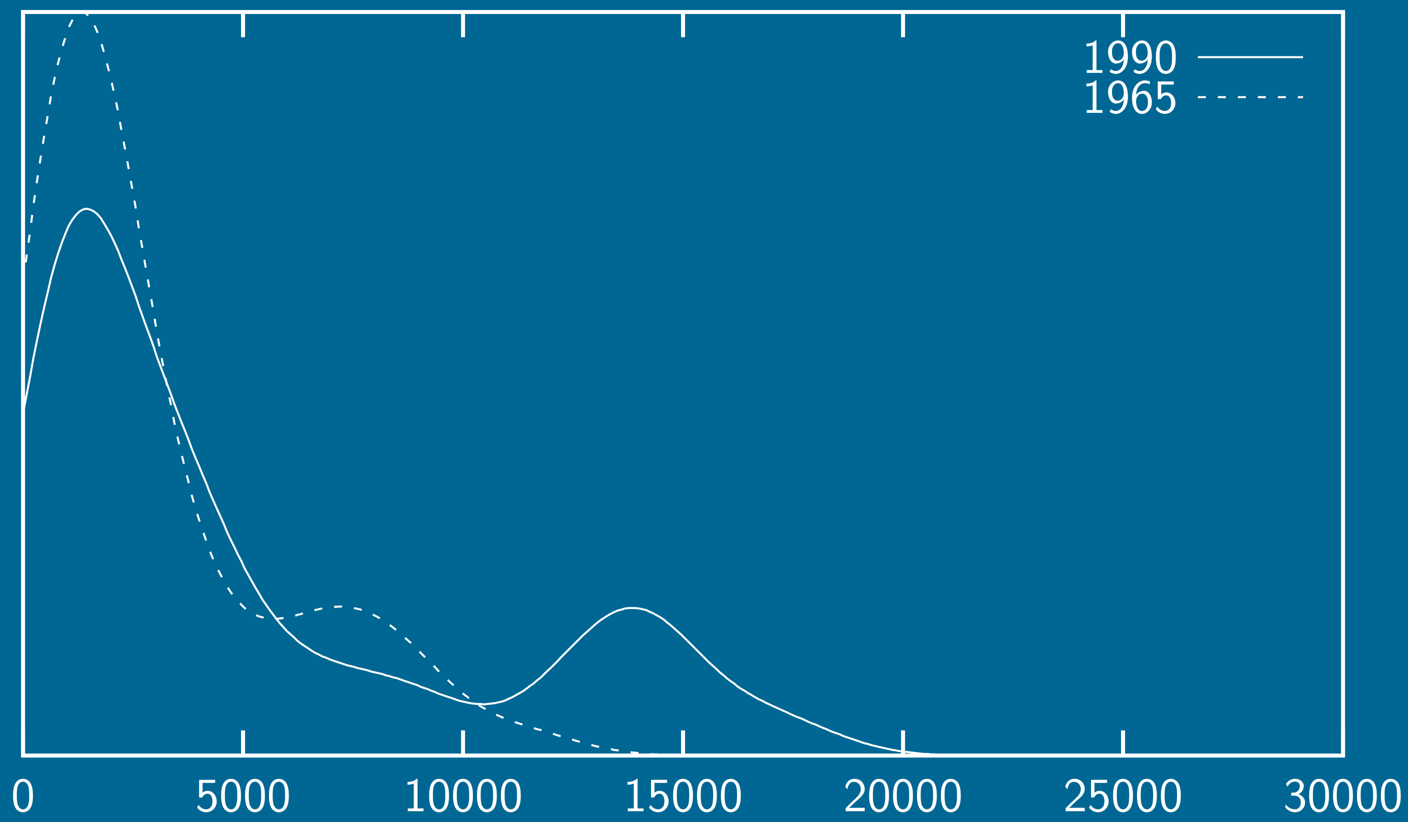
In the next two slides we present $P^3\varphi_{1990}$ and $P^5\varphi_{1990}$.

This gives 75 year and 125 year projections for the cross-country income distribution.

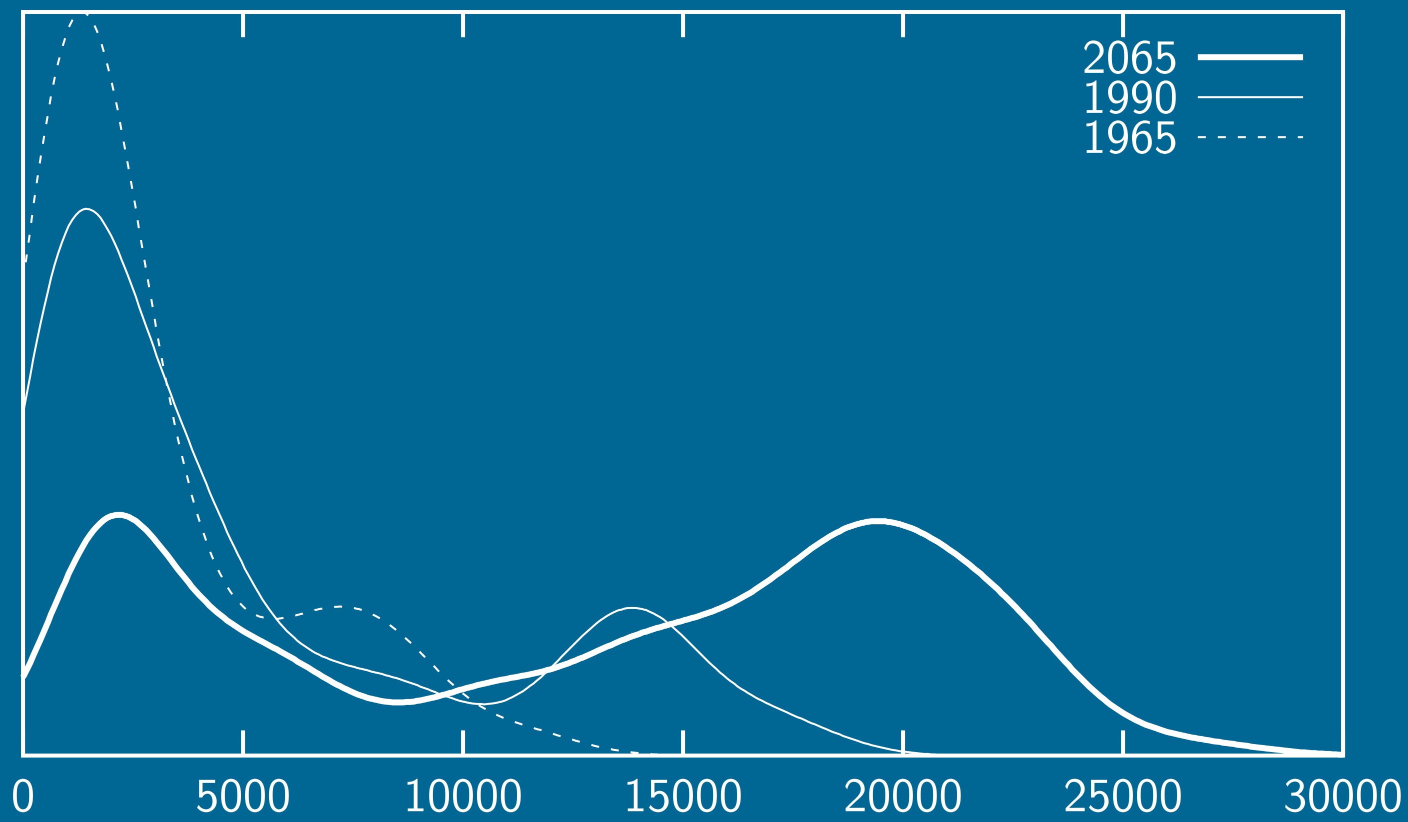
Back to top



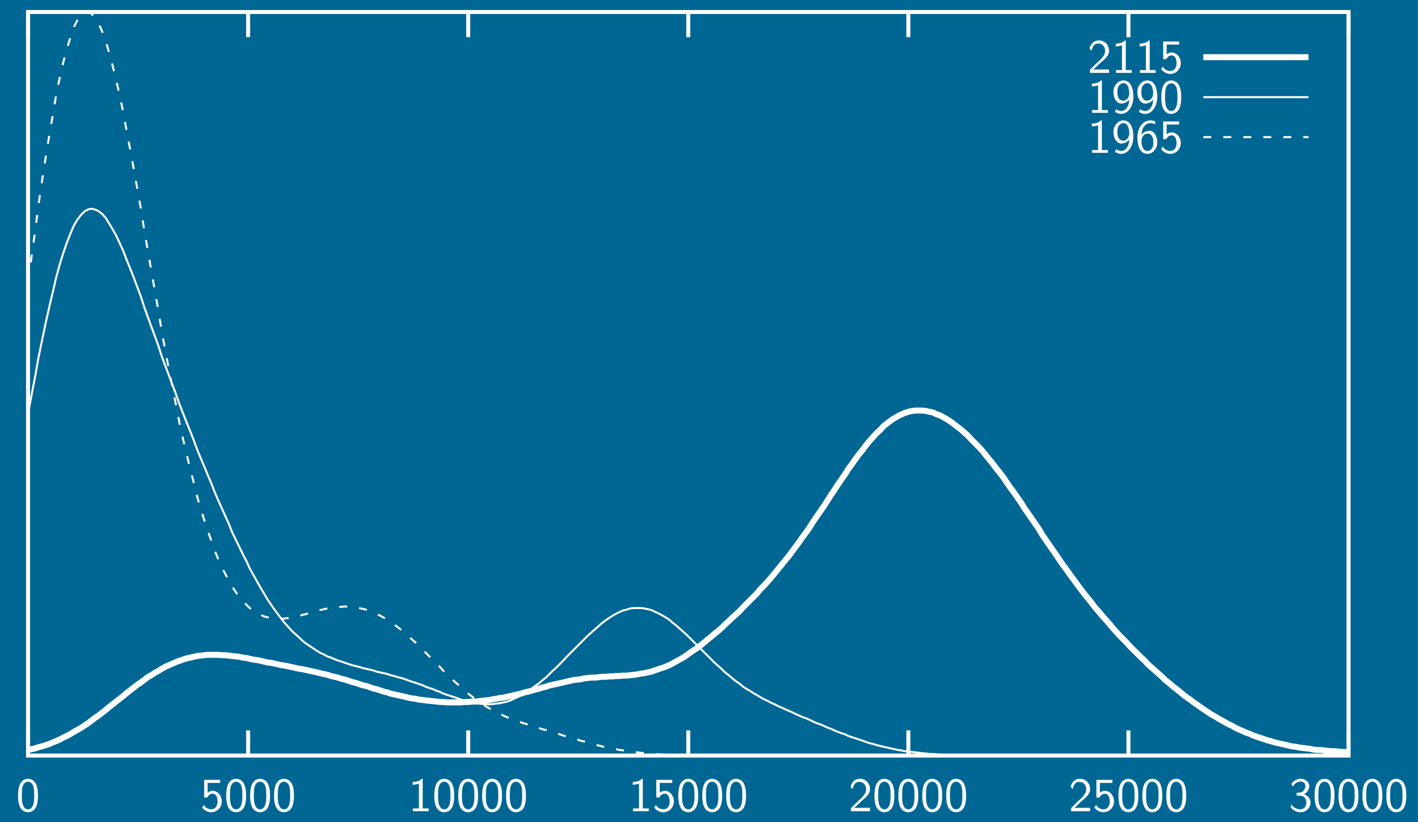
Back to top



Back to top



Back to top



Back to top

Kuznets hypothesis: inequality $\uparrow$ and then $\downarrow$ over time as industrialization occurs.

Kuznets hypothesis: inequality $\uparrow$ and then $\downarrow$ over time as industrialization occurs.

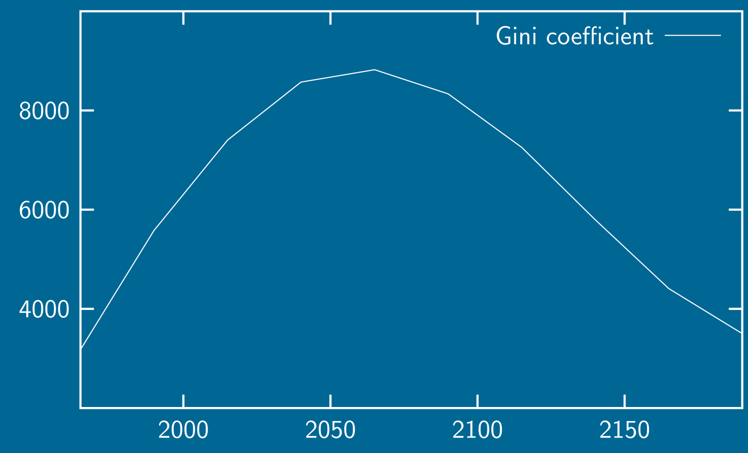
Gini coefficient:

$$g(\varphi) = \int\int |x - y|\varphi(x)\varphi(y)dxdy.$$

Kuznets hypothesis: inequality ↑ and then ↓ over time as industrialization occurs.

Gini coefficient:

$$g(\varphi) = \int \int |x - y| \varphi(x) \varphi(y) dx dy.$$



Back to top