ECON-GA 1025 Macroeconomic Theory I

Lecture 5

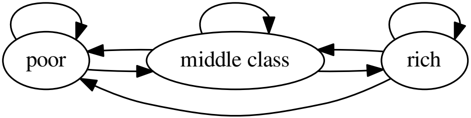
John Stachurski

Fall Semester 2018

Today's Lecture

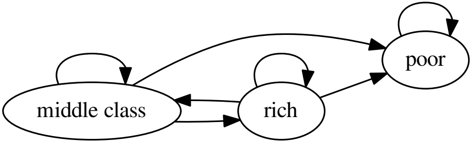
- Stochastic kernels and Markov chains on **finite** sets
- Distribution dynamics
- Aperiodicity and irreducibility
- Stability
- Ergodicity
- High performace computing in Python

A **directed graph** is a nonempty set of **nodes** $\mathsf{X} = \{x, y, \dots, z\}$ and a set of **arcs** $(x, y) \in \mathsf{X} \times \mathsf{X}$



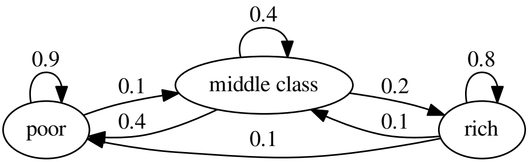
- y is called **accessible** from x if $y = x$ or $\exists$ a sequence of arcs leading from x to y
- The graph is called **strongly connected** if y is accessible from x for all $x, y \in \mathsf{X}$

Another example



- Strongly connected?

We can also attach numbers to the edges of a directed graph



The resulting graph is called a **weighted directed graph**

Interpretation will be given later

Prequel 2: Brouwer's Fixed Point Theorem

Theorem. (Brouwer–Hadamard, 1910) If

1. C is a convex compact subset of $\mathbb{R}^d$
2. T is a continuous self-map on C

then T has at least one fixed point in C

Prequel 3: The Space of Distributions

Let X be any finite set with elements $x_1, \ldots, x_n$

As usual, on $\mathbb{R}^{\mathsf{X}}$ we adopt the pointwise partial order

- $h \leqslant g$ if $h(x) \leqslant g(x)$ for all $x \in \mathsf{X}$

The set of **distributions** on X is denoted $\mathcal{P}(\mathsf{X})$ and defined as all $\varphi \in \mathbb{R}^{\mathsf{X}}$ such that

- $\varphi \geqslant 0$
- $\sum_{x \in \mathsf{X}} \varphi(x) = 1$

Think of $\varphi(x)$ as probability of hitting x

Metrizing $\mathcal{P}(\mathsf{X})$

As usual,

- $\|h\|_1 := \sum_{x \in \mathsf{X}} |h(x)|$
- $d_1(g,h) := \|g-h\|_1$

Thus,

$$\mathcal{P}(\mathsf{X}) = \{h \in \mathbb{R}^{\mathsf{X}} : h \geqslant 0 \text{ and } \|h\|_1 = 1\}$$

$\mathcal{P}(\mathsf{X})$ also called the **unit simplex** in $\mathbb{R}^n$

- A convex, compact subset of $\mathbb{R}^n$

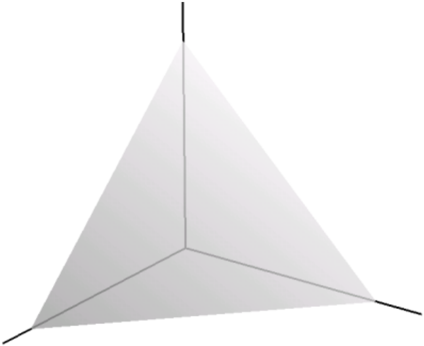
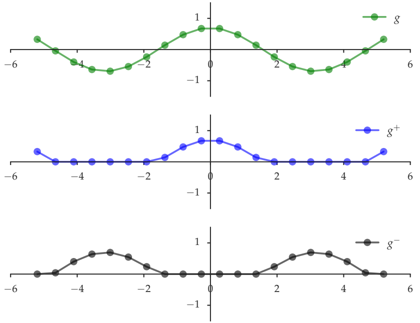


Figure: The unit simplex in $\mathbb{R}^3$

Notation: $g^+ := g \vee 0$ and $g^- := (-g) \vee 0$



Ex. Show that, for all $g \in \ell_1(\mathsf{X})$ we have

$$g = g^+ - g^- \quad \text{and} \quad |g| = g^+ + g^-$$

Stochastic Kernels

A **stochastic kernel** on X is a function $\Pi \colon \mathsf{X} \times \mathsf{X} \to \mathbb{R}$ such that

$$\Pi(x, \cdot) \in \mathcal{P}(\mathsf{X}) \quad \text{for all} \quad x \in \mathsf{X}$$

In other words,

1. $\Pi(x, y) \geqslant 0$ for all $(x, y) \in \mathsf{X} \times \mathsf{X}$
2. $\sum_{y \in \mathsf{X}} \Pi(x, y) = 1$ for all $x \in \mathsf{X}$

Intuition:

1. We have one distribution $\Pi(x, \cdot)$ for each point $x \in \mathsf{X}$
2. $\Pi(x, y)$ is the probability of moving from x to y in one step

Matrix Representation

There are some alternative representations of stochastic kernels

When X is finite, we can represent Π by a matrix

$$\Pi = \begin{pmatrix} \Pi(x_1, x_1) & \cdots & \Pi(x_1, x_n) \\ \vdots & & \vdots \\ \Pi(x_n, x_1) & \cdots & \Pi(x_n, x_n) \end{pmatrix}$$

Note: this is a Markov matrix / stochastic matrix

1. Square, nonnegative, rows sum to one
2. Distributions are rows, stacked vertically

Example. (Hamilton, 2005)

Estimates a statistical model of the business cycle based on US unemployment data

Markov matrix:

$$P_H := \begin{pmatrix} 0.971 & 0.029 & 0 \\ 0.145 & 0.778 & 0.077 \\ 0 & 0.508 & 0.492 \end{pmatrix}$$

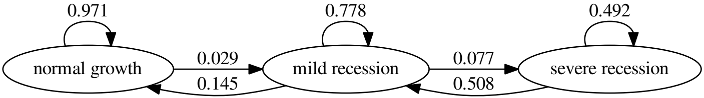
- state 1 $=$ normal growth
- state 2 $=$ mild recession
- state 3 $=$ severe recession

Length of the period $=$ one month

Digraph Representation

Another way to represent a finite stochastic kernel is by a weighted directed graph

Example. Here's Hamilton's business cycle model as a digraph



- set of nodes is X
- no edge means $\Pi(x, y) = 0$

Example. International growth dynamics study of Quah (1993)

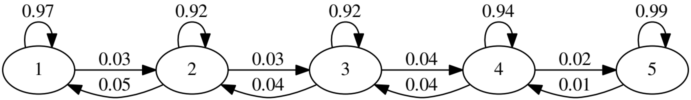
State = real GDP per capita relative to world average

States are 0–1/4, 1/4–1/2, 1/2–1, 1–2 and 2–∞

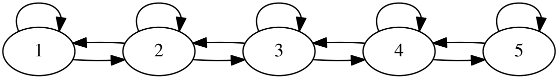
$$\Pi_Q = \begin{pmatrix} 0.97 & 0.03 & 0.00 & 0.00 & 0.00 \\ 0.05 & 0.92 & 0.03 & 0.00 & 0.00 \\ 0.00 & 0.04 & 0.92 & 0.04 & 0.00 \\ 0.00 & 0.00 & 0.04 & 0.94 & 0.02 \\ 0.00 & 0.00 & 0.00 & 0.01 & 0.99 \end{pmatrix}$$

The transitions are over a one year period

Quah's income dynamics model as a weighted directed graph:



Dropping labels gives the directed graph



From Stochastic Kernels to Markov Chains

Let

1. X be a finite set
2. $\{X_t\}_{t=0}^{\infty}$ be an X-valued stochastic process

$\{X_t\}_{t=0}^{\infty}$ is called a **Markov chain** on X if there exists a stochastic kernel Π on X such that

$$\mathbb{P}\{X_{t+1} = y \mid X_0, X_1, \ldots, X_t\} = \Pi(X_t, y) \quad \text{for all} \quad t \geqslant 0, \; y \in \mathsf{X}$$

In this case we say that $\{X_t\}_{t=0}^{\infty}$ is **generated by** Π

If $X_0 \sim \psi$, then ψ is called the **initial condition**

Simulation

One technique for generating $\{X_t\}$ from a given kernel Π

For $x \in \mathsf{X}$ and $u \in (0, 1)$, let

$$F(x, u) := \sum_{i=1}^{n} y_i \mathbb{1}\{q_{i-1}(x) < u \leqslant q_i(x)\}$$

where $\{y_1, \ldots, y_n\} = \mathsf{X}$ and

$$q_i(x) := \sum_{j=1}^{i} \Pi(x, y_j) \quad \text{with } q_0 = 0$$

Now X_0 is drawn from $\psi_0 \in \mathcal{P}(\mathsf{X})$ and then

$$X_{t+1} = F(X_t, U_{t+1}) \quad \text{where } \{U_t\} \overset{\text{IID}}{\sim} U(0, 1) \tag{1}$$

Generates a Markov chain with stochastic kernel Π

The next exercise asks you to verify this

Ex. Conditional on $X_t = x$, show that, for each i in $1, \ldots, n$,

1. $X_{t+1} = y_i$ if and only if

$$q_{i-1}(x) < U_{t+1} \leqslant q_i(x)$$

2. This event has probability $\Pi(x, y_i)$

Conclude that X_{t+1} in (1) is a draw from $\Pi(x, \cdot)$

Linking Marginals

By the law of total probability we have

$$\mathbb{P}\{X_{t+1} = y\} = \sum_{x \in \mathsf{X}} \mathbb{P}\{X_{t+1} = y \mid X_t = x\} \cdot \mathbb{P}\{X_t = x\}$$

Letting ψ_t be the distribution of X_t, this becomes

$$\psi_{t+1}(y) = \sum_{x \in \mathsf{X}} \Pi(x, y) \psi_t(x) \qquad (y \in \mathsf{X})$$

Regarding distributions as **row** vectors, we can write this as

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

The map $\psi \mapsto \psi \Pi$ **updates the distribution** of the state

Dynamical System Representation

Think of $(\mathcal{P}(\mathsf{X}), \Pi)$ as a dynamical system

- Π is identified with the map $\psi \mapsto \psi\Pi$,

$$(\Pi\psi)(y) := \sum_{x \in \mathsf{X}} \Pi(x, y)\psi(x)$$

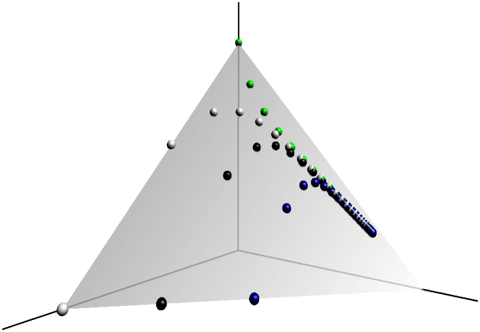
- Also called the **Markov operator** generated by kernel Π

Ex. Show that Π is a self-mapping on $\mathcal{P}(\mathsf{X})$

Interpretation of trajectories:

- $X_0 \sim \psi \implies X_t \sim \psi\Pi^t$
- $X_0 = x \implies X_t \sim \delta_x\Pi^t$

Some of trajectories in $\mathcal{P}(X)$ under Hamilton's business cycle model:



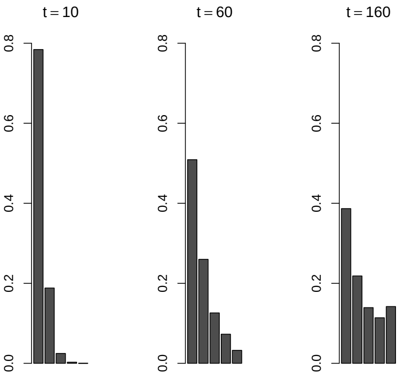


Figure: Distributions from Quah's stochastic kernel, $X_0 = 1$

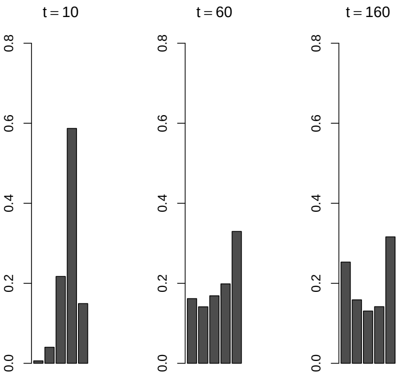


Figure: Distributions from Quah's stochastic kernel, $X_0 = 4$

Higher Order Kernels

Let Π be a stochastic kernel and let $\{\Pi^k\}$ be defined inductively by

$$\Pi^1 := \Pi \quad \text{and} \quad \Pi^{k+1}(x,y) := \sum_{z \in \mathsf{X}} \Pi(x,z)\Pi^k(z,y)$$

- Called the **k-step stochastic kernel**
- We are just taking matrix powers (finite case)

Ex. Show that if Π is a stochastic matrix, then so is Π^k for all k

If $\{X_t\}$ is generated by Π, then, for any $k \in \mathbb{N}$, we have

$$\Pi^k(x, y) = \mathbb{P}\{X_k = y \mid X_0 = x\} \qquad (x, y \in \mathsf{X})$$

To see why, recall that

$$\{X_t\} \text{ generated by } \Pi \text{ and } X_0 = x \implies X_k \sim \delta_x \Pi^k$$

Hence

$$\mathbb{P}\{X_k = y \mid X_0 = x\} = (\delta_x \Pi^k)(y)$$

But

$$(\delta_x \Pi^k)(y) = \Pi^k(x, y)$$

Chapman-Kolmogorov Equations

The kernels $\{\Pi^k\}$ satisfy the **Chapman–Kolmogorov relation**

$$\Pi^{j+k}(x,y) = \sum_{z \in \mathsf{X}} \Pi^k(x,z)\Pi^j(z,y) \qquad ((x,y) \in \mathsf{X} \times \mathsf{X})$$

Proof: Let $X_0 = x$ and let $y \in \mathsf{X}$ be given

By the law of total probability, we have

$$\begin{aligned}
\Pi^{j+k}(x,y) &= \mathbb{P}\{X_{j+k} = y\} \\
&= \sum_{z \in \mathsf{X}} \mathbb{P}\{X_{j+k} = y \mid X_k = z\}\mathbb{P}\{X_k = z\} \\
&= \sum_{z \in \mathsf{X}} \Pi^k(x,z)\Pi^j(z,y)
\end{aligned}$$

Expectations

Given stochastic kernel Π and h in $\mathbb{R}^{\mathsf{X}}$, consider

$$(\Pi h)(x) = \sum_{y \in \mathsf{X}} h(y)\Pi(x, y) \qquad (x \in \mathsf{X})$$

($\Pi h =$ the product of matrix Π and column vector h)

Interpretation

$$(\Pi h)(x) = \mathbb{E}[h(X_{t+1}) \mid X_t = x]$$

More generally,

$$(\Pi^k h)(x) = \sum_y h(y)\Pi^k(x, y) = \mathbb{E}[h(X_{t+k}) \mid X_t = x]$$

Stationary Distributions

Let Π be a stochastic kernel on X

If $\psi^* \in \mathcal{P}(\mathsf{X})$ satisfies

$$\psi^*(y) = \sum_{x \in \mathsf{X}} \Pi(x, y) \psi^*(x) \quad \text{for all} \quad y \in \mathsf{X}$$

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

Equivalent to the above:

- $\psi^* = \psi^* \Pi$
- ψ^* is a steady state of $(\mathcal{P}(\mathsf{X}), \Pi)$

Interpretation:

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

Existence

Theorem (Krylov–Bogolyubov). If X is finite then Π has at least one stationary distribution

Proof: Think of distribution $\varphi \in \mathcal{P}(\mathsf{X})$ as vector $(\varphi(x_i))_{i=1}^n$

- Π is a continuous map (just matrix multiplication)
- Π maps $\mathcal{P}(\mathsf{X})$ into itself
- $\mathcal{P}(\mathsf{X})$ is a closed, bounded subset of $\mathbb{R}^n$
- $\mathcal{P}(\mathsf{X})$ is also convex in $\mathbb{R}^n$

Existence of a fixed point follows from Brouwer

Computing the Stationary Distribution

Consider solving $\psi^* \Pi = \psi^*$ for ψ^*

Problem: there are trivial solutions, such as $\psi^* = 0$

To force our solution to be in $\mathcal{P}(\mathsf{X})$, let

- I be the $n \times n$ identity matrix
- $\mathbb{1}_n$ be the $1 \times n$ vector of ones, $\mathbb{1}_{n \times n}$ be the $n \times n$ matrix of ones

Ex. Show that $\psi \in \mathcal{P}(\mathsf{X})$ is stationary for Π if and only if

$$\mathbb{1}_n = \psi(I - \Pi + \mathbb{1}_{n \times n}) \tag{2}$$

Now transpose and solve for ψ' — requires that $I - \Pi + \mathbb{1}_{n \times n}$ is nonsingular

Probabilistic Properties

Let Π be a stochastic kernel on X and let x, y be states

We say that y is **accessible** from x if $x = y$ or

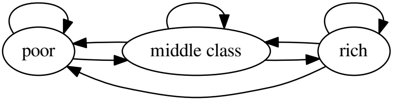
$$\exists\, k \in \mathbb{N} \text{ such that } \Pi^k(x, y) > 0$$

Equivalent: Accessible in the induced directed graph

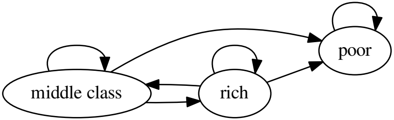
A stochastic kernel Π on X is called **irreducible** if every state is accessible from any other

Equivalent: The induced directed graph is strongly connected

Irreducible:



Not irreducible:



Aperiodicity

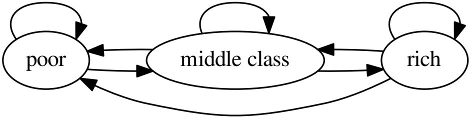
Let Π be a stochastic kernel on X

State $x \in$ X is called **aperiodic** under Π if

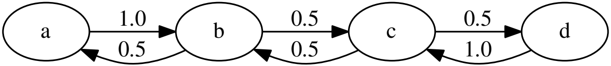
$$\exists\, i \in \mathbb{N} \text{ such that } k \geqslant i \implies \Pi^k(x, x) > 0$$

A stochastic kernel Π on X is called **aperiodic** if every state in X is aperiodic under Π

Aperiodic?

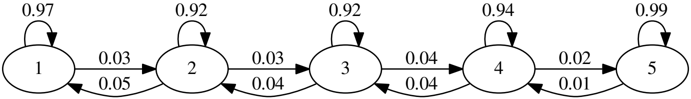


Aperiodic?



Stability of Markov Chains

Recall the distributions generated by Quah's model



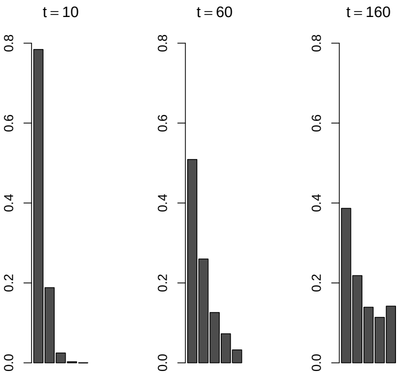


Figure: $X_0 = 1$

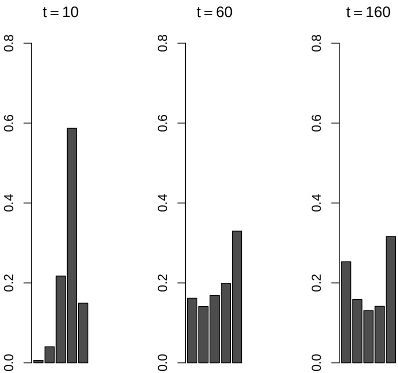


Figure: $X_0 = 4$

What happens when $t \to \infty$?

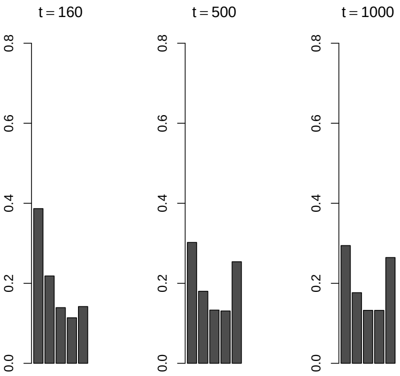


Figure: $X_0 = 1$

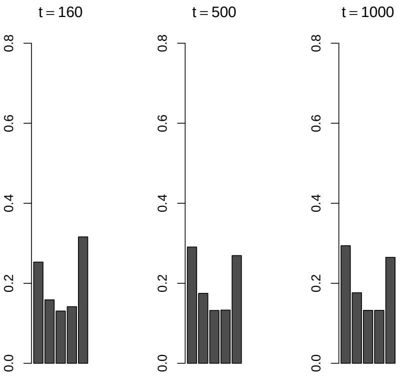


Figure: $X_0 = 4$

At $t = 1000$, the distributions are almost the same for both starting points

This suggests we are observing a form of stability

- is $(\mathcal{P}(\mathrm{X}), \Pi)$ globally stable?

Not all stochastic kernels are globally stable

Example. Let $\mathsf{X} = \{1, 2\}$ and consider the periodic Markov chain

$$\Pi = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Ex. Show $\psi^* = (0.5, 0.5)$ is stationary for Π

Ex. Show that

$$\delta_0 \Pi^t = \begin{cases} \delta_1 & \text{if } t \text{ is odd} \\ \delta_0 & \text{if } t \text{ is even} \end{cases}$$

Conclude that global stability fails

Proving Stability

Fact. The operator Π is always nonexpansive:

$$\|\varphi\Pi - \psi\Pi\|_1 \leqslant \|\varphi - \psi\|_1 \quad \forall\, \varphi, \psi \in \mathcal{P}(\mathsf{X})$$

Proof:

$$\begin{aligned}
\|\varphi\Pi - \psi\Pi\|_1 &= \sum_y \left| \sum_x \Pi(x,y)[\varphi(x) - \psi(x)] \right| \\
&\leqslant \sum_y \sum_x \Pi(x,y)|\varphi(x) - \psi(x)| \\
&= \sum_x \sum_y \Pi(x,y)|\varphi(x) - \psi(x)| = \|\varphi - \psi\|_1
\end{aligned}$$

With some more conditions we might be able to apply this result:

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

- **strict contraction** means $\rho(Tx, Ty) < \rho(x, y)$ when $x \neq y$
- a variation on the Banach CMT

X is finite, so $\mathcal{P}(\mathsf{X})$ is compact

We just need to boost nonexpansiveness to strict contractivity

Lemma. If $\Pi(x,y) > 0$ for all x, y, then Π is a strict contraction on $\mathcal{P}(\mathsf{X})$ under the metric d_1

The proof uses two lemmas:

Fact. If $\varphi, \psi \in \mathcal{P}(\mathsf{X})$ and $\varphi \neq \psi$, then

$$\exists\, x, x' \in \mathsf{X} \text{ such that } \varphi(x) > \psi(x) \text{ and } \varphi(x') < \psi(x')$$

Fact. If $g \in \mathbb{R}^{\mathsf{X}}$ and $\exists\, x, x' \in \mathsf{X}$ s.t. $g(x) > 0$ and $g(x') < 0$, then

$$|\sum_{y \in \mathsf{X}} g(y)| < \sum_{y \in \mathsf{X}} |g(y)|$$

Ex. Prove both

Under the conditions of the theorem, if $\varphi \neq \psi$, then

$$\begin{aligned}
\|\varphi\Pi - \psi\Pi\|_1 &= \sum_y \left| \sum_x \Pi(x,y)\varphi(x) - \sum_x \Pi(x,y)\psi(x) \right| \\
&= \sum_y \left| \sum_x \Pi(x,y)[\varphi(x) - \psi(x)] \right| \\
&< \sum_y \sum_x |\Pi(x,y)[\varphi(x) - \psi(x)]| \\
&= \sum_y \sum_x \Pi(x,y)|\varphi(x) - \psi(x)| \\
&= \sum_x \sum_y \Pi(x,y)|\varphi(x) - \psi(x)| = \|\varphi - \psi\|_1
\end{aligned}$$

We have prove the following:

Proposition. If $\Pi \gg 0$, then $(\mathcal{P}(\mathsf{X}), \Pi)$ is globally stable

But this condition is rather strict

- Quah's matrix fails it
- Hamilton's matrix fails it

Let's see if we can do better

Fact. If $(\mathcal{P}(\mathsf{X}), \Pi^i)$ is globally stable for some $i \in \mathbb{N}$, then $(\mathcal{P}(\mathsf{X}), \Pi)$ is also globally stable

Recall: If

1. dynamical system (M, g^i) is globally stable for some $i \in \mathbb{N}$
2. g is continuous at the fixed point of g^i

then (M, g) is also globally stable

Moreover, $\psi \mapsto \psi\Pi$ is everywhere continuous as already discussed

Theorem. If X is finite and Π is both aperiodic and irreducible, then Π is globally stable

Proof: It suffices to show that

$$\forall\, x, y \in \mathsf{X} \times \mathsf{X}, \quad \exists\, i_{x,y} \in \mathbb{N} \text{ s.t. } k \geqslant i_{x,y} \implies \Pi^k(x, y) > 0$$

Indeed, if this statement holds, then

$$i := \max\{i_{x,y}\} \quad \implies \quad \Pi^i(x, y) > 0 \quad \text{for all} \quad (x, y) \in \mathsf{X} \times \mathsf{X}$$

Implies that

- $(\mathcal{P}(\mathsf{X}), \Pi^i)$ is globally stable
- and hence $(\mathcal{P}(\mathsf{X}), \Pi)$ is globally stable

So fix $x, y \in \mathsf{X} \times \mathsf{X}$ and let's try to show that

$$\exists\, i = i_{x,y} \in \mathbb{N} \text{ s.t. } k \geqslant i \implies \Pi^k(x, y) > 0$$

Since Π is irreducible, $\exists\, j \in \mathbb{N}$ such that $\Pi^j(x, y) > 0$

Since Π is aperiodic, $\exists\, m \in \mathbb{N}$ such that

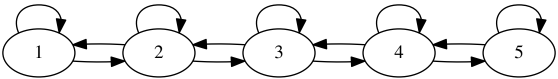
$$\ell \geqslant m \implies \Pi^\ell(y, y) > 0$$

Picking $\ell \geqslant m$ and applying the Chapman–Kolmogorov equation, we have

$$\Pi^{j+\ell}(x, y) = \sum_{z \in \mathsf{X}} \Pi^j(x, z)\Pi^\ell(z, y) \geqslant \Pi^j(x, y)\Pi^\ell(y, y) > 0$$

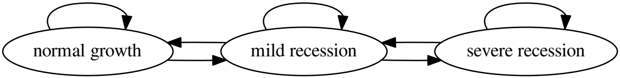
QED

Example. Quah's stochastic kernel is both irreducible and aperiodic



And therefore globally stable

Same with Hamilton's business cycle model



```
In [1]: import quantecon as qe

In [2]: P = [[0.971 , 0.029 , 0],
   ...:      [0.145 , 0.778 , 0.077],
   ...:      [0 , 0.508 , 0.492]]

In [3]: mc = qe.MarkovChain(P)

In [4]: mc.is_aperiodic
Out[4]: True

In [5]: mc.is_irreducible
Out[5]: True

In [6]: mc.stationary_distributions
Out[6]: array([[ 0.8128 ,  0.16256,  0.02464]])
```

A Weaker Set of Conditions

Let Π be a stochastic kernel on (finite set) X

Theorem. The following statements are equivalent:

1. Π^k has a strictly positive column for some $k \in \mathbb{N}$
2. For any $x, x' \in$ X, there exists a $k \in \mathbb{N}$ and a $y \in$ X such that

$$\Pi^k(x, y) > 0 \text{ and } \Pi^k(x', y) > 0$$

3. $(\mathcal{P}(\mathsf{X}), \Pi)$ is globally stable

Intuition for sufficiency

We know a stationary distribution exists, just need to prove convergence

Suppose that, for any $x, x' \in \mathsf{X}$, there exists a $k \in \mathbb{N}$ and a $y \in \mathsf{X}$ such that

$$\Pi^k(x, y) > 0 \text{ and } \Pi^k(x', y) > 0$$

Wherever we are now, we can meet up again

Hence no one is stuck at a local attractor

Initial conditions don't matter in the long run

Hence $(\mathcal{P}(\mathsf{X}), \Pi)$ is globally stable

Application: Inventory Dynamics

Let X_t = inventory of a product, obeys

$$X_{t+1} = \begin{cases} (X_t - D_{t+1})^+ & \text{if } X_t > s \\ (S - D_{t+1})^+ & \text{if } X_t \leqslant s \end{cases}$$

Assume $\{D_t\} \overset{\text{IID}}{\sim}$ the geometric distribution, say

A Markov chain on $\mathsf{X} := \{0, 1, \ldots, S\}$ with kernel

$$\Pi(x, y) = \begin{cases} \mathbb{P}\{(x - D_{t+1})^+ = y\} & \text{if } x > s \\ \mathbb{P}\{(S - D_{t+1})^+ = y\} & \text{if } x \leqslant s \end{cases}$$

Proposition The pair $(\mathcal{P}(\mathsf{X}), \Pi)$ is globally stable

Proof: Suppose that $D_{t+1} \geqslant S$

Then

$$0 \leqslant X_{t+1} \leqslant (S - D_{t+1})^+ = 0$$

Hence $\mathbb{P}\{D_{t+1} \geqslant S\} > 0$ implies $\Pi(x, 0) > 0$ for all x

Moreover $\mathbb{P}\{D_{t+1} \geqslant S\} > 0$ holds for the geometric distribution

Hence $(\mathcal{P}(\mathsf{X}), \Pi)$ is globally stable

The Law of Large Numbers

Let $h \in \mathbb{R}^{\mathsf{X}}$ and let $\{X_t\}$ be a Markov chain on generated by stochastic kernel p

Theorem. If X is finite and Π is globally stable with stationary distribution ψ^*, then

$$\mathbb{P}\left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} h(X_t) = \sum_{x \in \mathsf{X}} h(x)\psi^*(x) \right\} = 1$$

Intuition: $\{X_t\}$ "almost" identically distributed for large t

Also, stability means that initial conditions die out — a form of long run independence

An approximation of the IID property used in the classical LLN

LLN provides a new **interpretation** for the stationary distribution

Using the LLN with $h(x) = \mathbb{1}\{x = y\}$, we have

$$\frac{1}{n} \sum_{t=1}^{n} \mathbb{1}\{X_t = y\} \to \sum_{x \in \mathsf{X}} \mathbb{1}\{x = y\} \psi^*(x) = \psi^*(y)$$

Turning this around,

$$\psi^*(y) \approx \text{ fraction of time that } \{X_t\} \text{ spends in state } y$$

This is **not** always valid **unless** the chain in question is stable