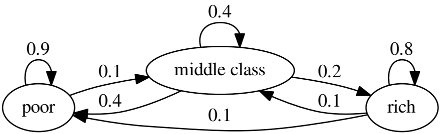
OSM Bootcamp

Lecture 5

John Stachurski

2018

Finite Markov Chains



$$\mathbb{P}\{X_{t+1} = \text{poor} \mid X_t = \text{rich}\} = 0.1$$

Distributions

We start with a **finite state space** $\mathbb{X} = \{x_1, \ldots, x_n\}$

Example. $x_1 =$ poor, $x_2 =$ middle class, $x_3 =$ rich

A **distribution** on $\mathbb{X}$ is a $\phi \colon \mathbb{X} \to \mathbb{R}$ such that

- $\phi(x) \geqslant 0$ for all $x \in \mathbb{X}$
- $\sum_{x \in \mathbb{X}} \phi(x) = 1$

Example. $\phi(x_1) = 1/2$, $\phi(x_2) = 1/4$, $x_3 = 1/4$

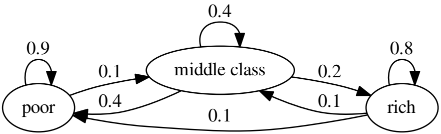
Let $\mathbb{D}$ be the set of distributions on $\mathbb{X}$

A **stochastic kernel** on $\mathbb{X}$ is a $P \colon \mathbb{X} \times \mathbb{X} \to \mathbb{R}_+$ such that

$$\sum_{y \in \mathbb{X}} P(x, y) = 1 \text{ for all } x \in \mathbb{X}$$

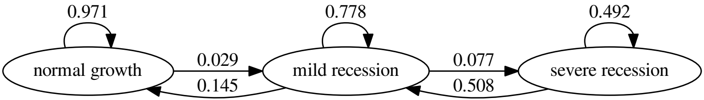
Interpretation: $P(x, y) =$ probability of moving $x \to y$ in one step

Example. $P(\text{rich}, \text{poor}) = 0.1$



Stochastic kernels can be represented by **weighted directed graphs**

Example. (Hamilton, 2005) estimates a statistical model of the business cycle based on US unemployment data

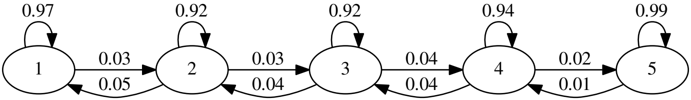


- set of nodes is $\mathbb{X}$
- no edge means $P(x, y) = 0$

QuantEcon

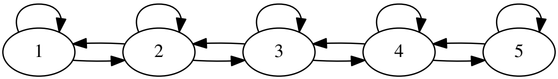
Example. International growth dynamics study of Quah (1993)

State = real GDP per capita relative to world average



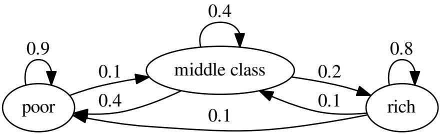
- state 1 means GDP per capita is $\leqslant 1/4$ of world ave
- state 2 means GDP per capita is $1/4 - 1/2$ of world ave
- ...

Dropping labels gives the directed graph



If P is a stochastic kernel, then

- $P(x, \cdot) \in \mathbb{D}$ for any x
- if at x today, then next period's state is drawn from $P(x, \cdot)$



If rich today, then next period is a draw from

$$P(\text{rich}, \cdot) = (0.1, 0.1, 0.8)$$

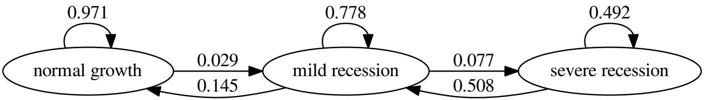
Matrix representation

We can represent any stochastic kernel P by a **Markov matrix**

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

- square
- nonnegative
- rows sum to one

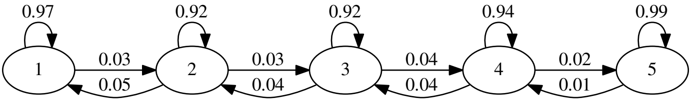
Example. (Hamilton, 2005)



Markov matrix:

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

Example. Quah (1993)



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

Let ψ be in $\mathbb{D}$ and let P be a stochastic kernel on $\mathbb{X}$

The corresponding **Markov chain** on $\mathbb{X}$ is generated as follows

set $t = 0$ and draw X_t from ψ ;
while $t < \infty$ **do**
    draw X_{t+1} from the distribution $P(X_t, \cdot)$;
    let $t = t + 1$;
end

Here ψ is called the **initial condition**

Linking Marginals

By the law of total probability we have

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

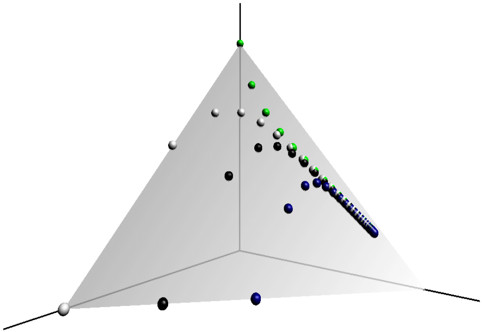
Letting ψ_t be the distribution of X_t, this becomes

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

In matrix form, with ψ_i as **row** vectors, this becomes

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

We can view $\psi_{t+1} = \psi_t P$ as a **dynamical system** $(\mathbb{D}, P)$



Trajectories in $\mathbb{D}$ under Hamilton's business cycle model

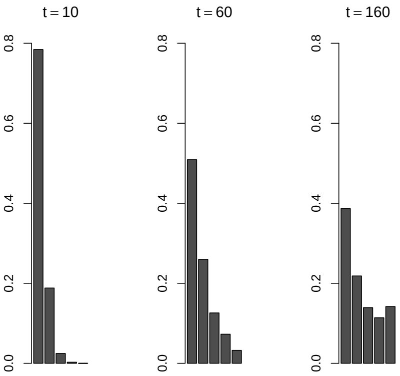


Figure: Distributions from Quah's stochastic kernel, $X_0 = 1$

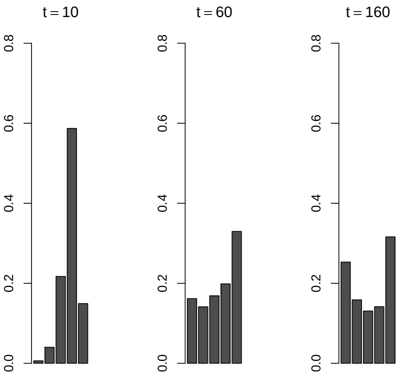


Figure: Distributions from Quah's stochastic kernel, $X_0 = 4$

Stationary Distributions

Let P be a stochastic kernel on $\mathbb{X}$

If $\psi^* \in \mathbb{D}$ satisfies

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

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

Equivalent: $\psi^* P = \psi^*$

Equivalent: ψ^* is a steady state of $(\mathbb{D}, P)$

Theorem. Every finite state Markov chain has at least one stationary distribution (see Brouwer fixed point theorem)

Probabilistic Properties

Let P be a stochastic kernel on $\mathbb{X}$ and let x, y be states

- $P^k(x, y) =$ probability of moving $x \to y$ in k steps

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

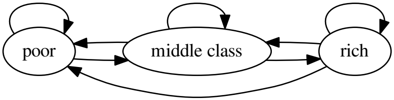
$$\exists \, k \in \mathbb{N} \text{ such that } P^k(x, y) > 0$$

Equivalent: Accessible in the induced directed graph

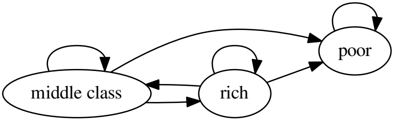
A stochastic kernel P on $\mathbb{X}$ is called **irreducible** if every state is accessible from any other

QuantEcon

Irreducible?



Irreducible?



Aperiodicity

Let P be a stochastic kernel on $\mathbb{X}$

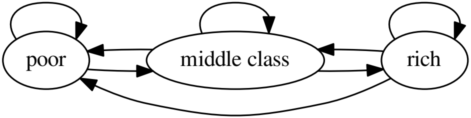
State $x \in \mathbb{X}$ is called **aperiodic** under P if

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

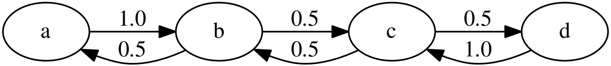
A stochastic kernel P on $\mathbb{X}$ is called **aperiodic** if every state in $\mathbb{X}$ is aperiodic under P

Remark. If $P(x,y) > 0$ for every $x, y \in \mathbb{X}$, then P is both aperiodic and irreducible

Aperiodic?

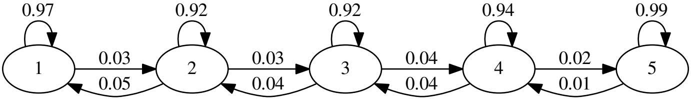


Aperiodic?



Stability of Markov Chains

Recall the distributions generated by Quah's model



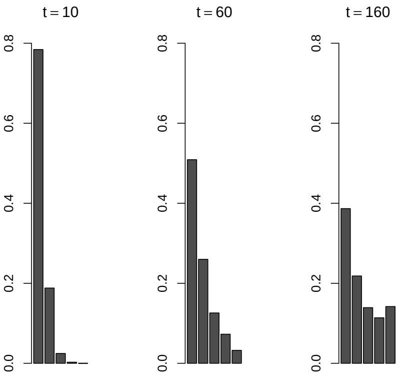


Figure: $X_0 = 1$

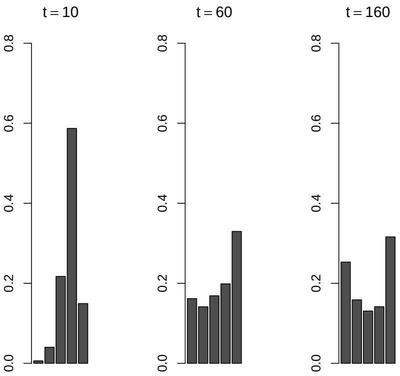


Figure: $X_0 = 4$

What happens when $t \to \infty$?

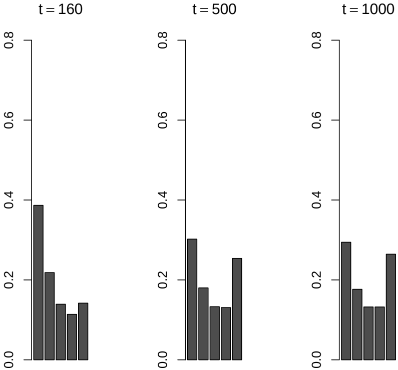


Figure: $X_0 = 1$

QuantEcon

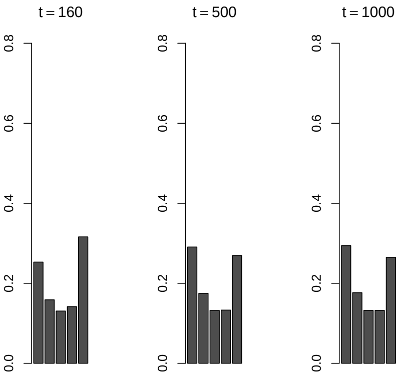


Figure: $X_0 = 4$

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

This suggests we are observing a form of stability

But how to define stability of Markov chains?

A stochastic kernel P on $\mathbb{X}$ is called **globally stable** if the dynamical sytem $(\mathbb{D}, P)$ is globally stable

Example. Let $\mathbb{X} = \{1, 2\}$ and consider

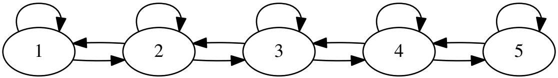
$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

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

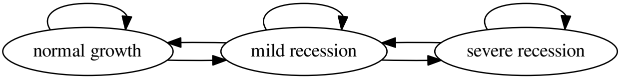
So is this P globally stable?

Theorem. If P is aperiodic and irreducible, then $(\mathbb{D}, P)$ is globally stable

Example. Quah's stochastic kernel is globally stable



Same with Hamilton's business cycle model



QuantEcon

```
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]])
```

Discretization

We can approximate continuous state Markov processes with finite state Markov chains

This is called **discretization** of the process

A common task: discretize the Gaussian AR(1) process

$$X_{t+1} = \rho X_t + \sigma \xi_{t+1} \quad \text{where} \quad \{\xi_t\} \overset{\text{IID}}{\sim} N(0,1)$$

We need a function that maps (ρ, σ, n) to a discrete Markov chain with n states

QuantEcon

A common algorithm in economics is **Tauchen's** method:

```
In [10]: import quantecon as qe

In [11]: mc = qe.tauchen(0.9, 0.1, n=2)

In [12]: mc.state_values
Out[12]: array([-0.6882472,  0.6882472])

In [13]: mc.P
Out[13]:
array([[  1.00000000e+00,   2.92862845e-10],
       [  2.92862879e-10,   1.00000000e+00]])
```

QuantEcon