VE414 Lecture 12

Jing Liu

UM-SJTU Joint Institute

October 21, 2019

- Consider a small town, in which 30% of the married women get divorced each year and 20% of the single women get married each year.

$$\mathbf{w}_1 = \mathbf{A}\mathbf{w}_0 \qquad \text{where} \quad \mathbf{A} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix} \quad \text{and} \quad \mathbf{w}_0 = \begin{bmatrix} 800 \\ 200 \end{bmatrix}$$

Q: Consider the following Julia outputs, what do you notice?

```
julia> A = [0.7 0.2; 0.3 0.8]
```

```
2*2 Array{Float64,2}:
 0.7  0.2
 0.3  0.8
```

```
julia> w0 = [800; 200]
```

```
2-element Array{Int64,1}:
 800
 200
```

```
julia> A * w0
```

```
2-element Array{Float64,1}:
 600.0
 400.0
```

```
julia> A^2 * w0
```

```
2-element Array{Float64,1}:
 499.99999999999994
 500.0
```

```
julia> A^4 * w0
```

```
2-element Array{Float64,1}:
 425.0
 575.0
```

```
julia> A^8 * w0
```

```
2-element Array{Float64,1}:
 401.5625000000006
 598.4375000000002
```

```
julia> A^16 * w0
```

```
2-element Array{Float64,1}:
 400.0061035156251 7
 599.9938964843755
```

```
julia> A^20 * w0
```

```
2-element Array{Float64,1}:
 400.0003814697268
 599.9996185302739
```

```
julia> A^20 * w0
```

```
2-element Array{Float64,1}:
 400.0003814697268
 599.9996185302739
```

```
julia> A^40 * w0
```

```
2-element Array{Float64,1}:
 400.0000000003645
 599.9999999996372
```

- It seems the Markov Chain $\{\mathbf{w}_0, \mathbf{w}_1, \ldots, \}$ converges to $[400, 600]^{\mathrm{T}}$

```
julia> A^80 * w0
```

```
2-element Array{Float64,1}:
 400.00000000000136
 600.00000000002
```

```
julia> w0 = [ 123; 877]; A^80 * w0
```

```
2-element Array{Float64,1}:
 400.000000000014
 600.000000000023
```

```
julia> w0 = [ 877; 123]; A^80 * w0
```

```
2-element Array{Float64,1}:
 400.0000000000136
 600.000000000022
```

```
julia> w0 = [ 159; 841]; A^80 * w0
```

```
2-element Array{Float64,1}:
 400.000000000014
 600.000000000023
```

- And it seems it converges to the same limit independent of the initial $\mathbf{w}_0$.

- Of course, people get married and get divorced change from year to year

$$\mathbf{w}_k \to \begin{bmatrix} 400 \\ 600 \end{bmatrix} \qquad \text{as} \qquad k \to \infty$$

however, it seems the proportion/probability reminds the same, if we set

$$\mathbf{p}_k = \frac{1}{1000}\mathbf{w}_k$$

then $\mathbf{p}_k$ is essentially the pmf of being married or single at the kth year.

- Let $X = 0$ denote married and $X = 1$ as single, and

$$\pi_X(x) = \begin{cases} 0.4 & \text{for} \quad x = 0, \\ 0.6 & \text{for} \quad x = 1, \end{cases}$$

then $X_{k-1} \sim \pi_X$ implies $X_k \sim \pi_X$.

- Distributions π_X are called invariant, stationary or equilibrium distribution.

- For this simple model, where $\mathcal{D} = \{0, 1\}$, convergence is easy to show

$$\mathbf{p}_k = \mathbf{A}^k \mathbf{p}_0 = \mathbf{A}^k \left(\alpha_{10}\mathbf{v}_1 + \alpha_{20}\mathbf{v}_2\right)$$

where $\mathbf{v}_1$ and $\mathbf{v}_2$ are eigenvectors of $\mathbf{A}$ corresponding eigenvalues λ_1 and λ_2.

```
julia> eigvals(A)
```

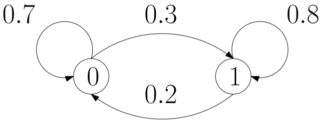
```
2-element Array{Float64,1}:
 0.5
 1.0
```

which leads to the following convergence result as $k \to \infty$,

$$\mathbf{p}_k = \alpha_{10}\mathbf{A}^k\mathbf{v}_1 + \alpha_{20}\mathbf{A}^k\mathbf{v}_2 = \alpha_{10}\left(\frac{1}{2}\right)^k \mathbf{v}_1 + \alpha_{20}\left(1\right)^k \mathbf{v}_2 \to a_{20}\mathbf{v}_2 = \begin{bmatrix} 0.4 \\ 0.6 \end{bmatrix}$$

Q: In this simple case, we can easily identify the conditions lead to convergence, but what is the general condition under which a Markov chain will converge?

- Note this example can be understood in terms of a different Markov chain,



$$X_k = \begin{cases} 0 & \text{Married,} \\ 1 & \text{Single.} \end{cases}$$

by thinking in terms of individual marriage status, then the matrix

$$\mathbf{A} = \begin{bmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{bmatrix}$$

is closely related to the transition matrix of this new Markov chain

$$\mathbf{P} = \mathbf{A}^{\mathrm{T}}$$

- Let $\boldsymbol{\pi}_k$ be the column vector consists of $f_{X_k}(j) = \Pr(X_k = j \mid X_0)$, then

$$\boldsymbol{\pi}_k^{\mathrm{T}} = \boldsymbol{\pi}_0^{\mathrm{T}} \mathbf{P}^k \qquad \text{where } \boldsymbol{\pi}_0 \text{ denote the initial condition}$$

Q: What does each component of $\boldsymbol{\pi}_k$ represent? How about $\lim\limits_{k\to\infty} \boldsymbol{\pi}_k$?

- The convergence result on the proportions,

$$\begin{aligned}\mathbf{p}_k = \mathbf{A}^k\mathbf{p}_0 = \alpha_{10}\left(\frac{1}{2}\right)^k\mathbf{v}_1 + \alpha_{20}\left(1\right)^k\mathbf{v}_2 \to \alpha_{20}\mathbf{v}_2 \quad \text{as} \quad k\to\infty \\ \to \begin{bmatrix}0.4\\0.6\end{bmatrix} = \mathbf{p}_\infty \quad \text{as} \quad k\to\infty\end{aligned}$$

i.e. 40% of women are married and 60% of women are single for a large k,

$$\boldsymbol{\pi}_k^{\mathrm{T}} = \boldsymbol{\pi}_0^{\mathrm{T}}\mathbf{P}^k = \mathbf{p}_0^{\mathrm{T}}\mathbf{P}^k = \left(\left(\mathbf{P}^k\right)^{\mathrm{T}}\mathbf{p}_0\right)^{\mathrm{T}} = \left(\mathbf{A}^k\mathbf{p}_0\right)^{\mathrm{T}} \to \mathbf{p}_\infty^{\mathrm{T}} \implies \boldsymbol{\pi}_\infty = \mathbf{p}_\infty$$

which means any individual has a 40% chance of being married for a large k.

- It is easier to grasp the concept of convergence and show convergence for a specific Markov chain using eigenvalues and eigenvectors. But doing so for a general Markov chain will either require too much maths or too little insight.
- We are interested in properties of Markov chains that lead to convergence in general rather than a tool to determine whether a specific chain converges.

Q: Would a Markov chain $\{X_k\}$ with the following transition matrix converge

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

having the above transition matrix means the two states are disconnected,

- In this case, the chain is trapped in one of the two absorbing states

$$X_k = X_0, \qquad \text{for all } k \geq 0.$$

so what it converges to depends on on the initial state/distribution.

- We do not include those like the one above as having a limiting distribution.

Definition

The distribution f_X is called the limiting distribution of the Markov chain $\{X_n\}$ if

$$f_X(j) = \lim_{k\to\infty} \Pr\left(X_k = j \mid X_0 = i\right)$$

for all i and j in the state space $\mathcal{S}$, and it is often denoted by

$$\pi_X = f_X$$

or in terms of a vector

$$\boldsymbol{\pi}^{\mathrm{T}} = \begin{bmatrix} \pi_0 & \pi_1 & \pi_2 & \cdots & \end{bmatrix}$$

where

$$\pi_j = \lim_{k\to\infty} \Pr\left(X_k = j \mid X_0 = i\right)$$

- When a limiting distribution exists, it does't depend on the initial state X_0.

Q: Would a Markov chain $\{X_k\}$ with the following transition matrix converge

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

having the above transition matrix means it has a periodic behaviour,

$$X_{k+2} = X_k, \qquad \text{for all } k \geq 0.$$

- In particular, $X_k = \begin{cases} X_0 & \text{if } n \text{ is even,} \\ X_1 & \text{if } n \text{ is odd,} \end{cases}$ so it will NOT converge to a single

$$\boldsymbol{\pi}_\infty$$

thus no limit, furthermore it also depends on the initial state/distribution.

- For example, consider the limit when $\boldsymbol{\pi}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, and then when $\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$.

Definition

Let gcd denote the greatest common divisor of the integers in the set, then

$$d(i) = \gcd\{n \geq 1 \mid \Pr(X_n = i \mid X_0 = i) > 0\}$$

is known as the period of a state i.

- If $d(i) > 1$, we say the state i is periodic.
- If $d(i) = 1$, we say the state i is aperiodic.

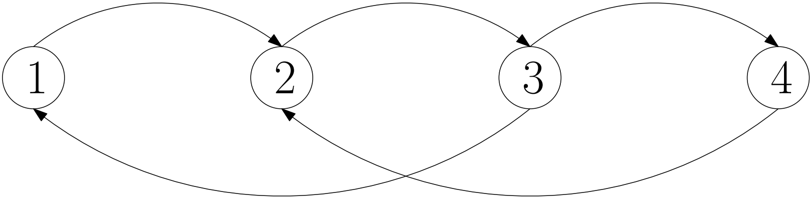
Q: Let a state be period k, is it always possible to return to the state in k steps?

- All states in the same communicating class have the same period.

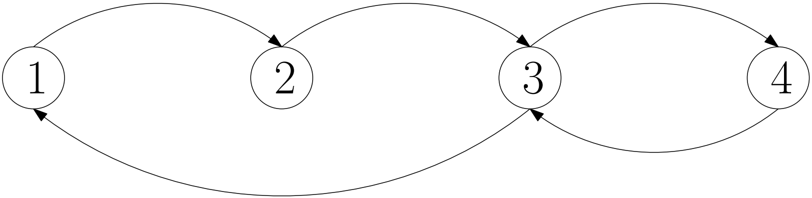
$$i \leftrightarrow j \iff d(i) = d(j)$$

- A class is said to be a/periodic if its states are a/periodic.
- Finally, a Markov chain is said to be aperiodic if all of its states are aperiodic.

Q: Is the following Markov chain periodic?



Q: How about the following?

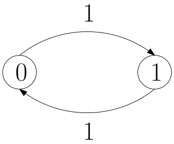


where having an arrow mean the transition probability is nonzero.

Finite Markov Chains

- We have seen no limiting distribution exist if there are two absorbing states

and no limiting distribution exist if there are two absorbing states



- It is not difficult to extend the above findings for classes rather than states.

Theorem

Suppose $\{X_k\}$ is a finite Markov chain, that is, the state space of $\{X_k\}$ is finite. If the chain is irreducible and aperiodic, then the limiting distribution exists.

- Consider a Markov chain $\{X_k, k \geq 0\}$ with two possible states $\mathcal{S} = \{0, 1\}$

$$\mathbf{P} = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}$$

where a and b are two real numbers in the interval $[0, 1]$ such that

$$0 < a + b < 2$$

Q: Does this Markov chain have a limiting distribution?

Q: What is the limiting distribution $\boldsymbol{\pi}$ of this Markov chain?

Q: What are the mean return times m_0 and m_1 for this Markov chain?

Q: Do you notice the connection between $\boldsymbol{\pi}$ and the mean return times?

Definition

The state i is known as recurrent if $h_{ii} = 1$, and transient if $h_{ii} < 1$, where

$$h_{ii} = \Pr\left(X_k = i \mid X_0 = i\right) \qquad \text{for some } k \geq 1.$$