Discrete-time Markov chains

Stochastic algorithms

Francesco Corona

Department of Computer Science
Federal University of Ceará, Fortaleza

Stochastic processes and Markov chains

We shall describe the behaviour of a system by describing all the different states the system may occupy and by indicating how it moves among them

- The number of states is possibly infinite

We assume that the system occupies one and only one state at any time

We also assume that the system's evolution is represented by transitions

- Transitions occur from state to state
- Transitions occur instantaneously

Stochastic processes and Markov chains (cont.)

If the future evolution of the system depends only on its current state and not on its history, then the system may be represented by a **Markov process**

Possible even when the system does not possess this property explicitly

- We can construct a corresponding implicit representation

A Markov process is a special case of a **stochastic process**

Stochastic processes and Markov chains (cont.)

We define a stochastic process as a family of random variables $\{X(t), t \in T\}$

- Each $X(t)$ is a random variable (on some probability space)
- Parameter t can be understood as time

Thus, $x(t)$ is the value assumed by the random variable $X(t)$ at time t

T is called the **index** or **parameter set**

- It is a subset of $(-\infty, +\infty)$

Stochastic processes and Markov chains (cont.)

Continuous-time parameter stochastic process

⇝ Index set is continuous

$$T = \{t | 0 \leq t < +\infty\}$$

Discrete-time parameter stochastic process

⇝ Index set is discrete

$$T = \{0, 1, 2, \dots\}$$

Stochastic processes and Markov chains (cont.)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

The values assumed by the random variables $X(t)$ are called **states**

- The space of all possible states is called **state-space**

When the state-space is discrete, the process is often called a chain

- To denote states, we use a subset of natural numbers

$$\rightsquigarrow \quad \{0, 1, 2, \dots\}$$

Stochastic processes and Markov chains (cont.)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Two important features of a stochastic process

⇝ Discrete/continuous time-evolution

⇝ Discrete/continuous states

Stochastic processes and Markov chains (cont.)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

A process whose evolution depends on the time it is initiated

⇝ **Non-stationary**

A process whose evolution is invariant under arbitrary shifts

⇝ **Stationary**

Stochastic processes and Markov chains (cont.)

Stationary random process

A random process is said to be a **stationary random process** if its joint distribution function is invariant to time shifts

For any constant α, we have

$$\begin{aligned} \rightsquigarrow \quad & \text{Prob}\{X(t_1) \leq x_1, X(t_2) \leq x_2, \cdots, X(t_n) \leq x_n\} \\ &= \text{Prob}\{X(t_1+\alpha) \leq x_1, X(t_2+\alpha) \leq x_2, \cdots, X(t_n+\alpha) \leq x_n\}, \\ & \qquad \text{for all } n \text{ and all } t_i \text{ and } x_i \text{ with } i = 1, 2, \dots, n \end{aligned}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Stochastic processes and Markov chains (cont.)

Stationarity is not that transitions probabilities cannot depend on time

- Transition may depend on the amount of elapsed time
- The process is said to be **non-homogeneous**

Whether homogeneous or not the process can/cannot be stationary

- If it is stationary, its evolution may change over time
- This evolution is the same, irrespective of initial
- (stationary and non-homogeneous)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Stochastic processes and Markov chains (cont.)

We are interested in some statistical characteristics of the dynamical system

- They may depend on time t at which the system is initiated

A process whose evolution depends on the time in which is is started

- **Non-stationary random process**

A process that is invariant under an arbitrary shift of the time origin

- **Stationary random process**

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Stochastic processes and Markov chains (cont.)

A Markov process is a stochastic process with a conditional probability distribution function that satisfies the **Markov** or **memoryless property**

We focus on discrete-state processes in both discrete and continuous time

Discrete-time Markov chains

Discrete-time Markov chains

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities

We shall consider a discrete-time Markov chain (discrete-state)
$\rightsquigarrow$ We observe its state at a discrete, but infinite, set of times

We assume that state transitions either can or cannot occur

- Transitions take place only at those abstract time instants
- (We can take time instants to be one time-unit apart)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

We consider the discrete time index set T to be the set of natural numbers

$$\rightsquigarrow \quad T = \{0, 1, \cdots, n, \cdots\}$$

Successive observations define the random variables $X_0, X_1, \cdots, X_n, \cdots$

A discrete-time Markov chain $\{X_n, n = 0, 1, 2, \cdots\}$ is a Markov process

As such, it satisfies the **Markov property**

$$\begin{aligned}\text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0\} \\ \rightsquigarrow \quad = \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n\}\end{aligned} \quad (1)$$

The state at time step $n+1$ will depend only on the state at time step n

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Markov property

For all natural numbers n and for all states x_n,

$$\begin{aligned}\text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0\} \\ = \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n\}\end{aligned}$$

Generalities (cont.)

$$\text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0\} = \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n\} \quad (2)$$

The future evolution of the system depends only on its current state

It is irrelevant whether the system is in state $X_0 = x_0$ at time step 0, in state $X_1 = x_1$ at time 1, and so on up to state $X_{n-1} = x_{n-1}$ at time $n-1$

State x_n the sum total of all the information concerning the history

- This is all is relevant to the future evolution

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

To simplify notation, we shall not use x_i, x_j and x_k to denote states

⇝ We shall use i, j and k, ...

So, we write the conditional probabilities

$$\text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n\} \quad \rightsquigarrow \text{Prob}\{X_{n+1} = j | X_n = i\} \quad \rightsquigarrow p_{ij}(n)$$

The conditional probability $p_{ij}(n)$ of performing a transition from state $x_n = i$ to state $x_{n+1} = j$ when the time parameter changes from n to $n+1$

⇝ **Single-step transition probabilities**

⇝ (**Transition probabilities**)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Single-step transition probabilities

They are conditional probabilities of making a transition from state $x_n = i$ to state $x_{n+1} = j$, when the time parameter is increased from n to $n+1$

$$\rightsquigarrow \quad p_{ij}(n) = \text{Prob}\{X_{n+1} = j | X_n = i\} \quad (3)$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Transition probability matrix or **chain matrix**

Let $P(n)$ be a matrix with $p_{ij}(n)$ in row i and column j, for all i and j

$$P(n) = \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \vdots \end{matrix} \begin{pmatrix} p_{00}(n) & p_{01}(n) & p_{02}(n) & \cdots & p_{0j}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & p_{12}(n) & \cdots & p_{1j}(n) & \cdots \\ p_{20}(n) & p_{21}(n) & p_{22}(n) & \cdots & p_{2j}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i0}(n) & p_{i1}(n) & p_{i2}(n) & \cdots & p_{ij}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

The elements of matrix $P(n)$ satisfy two properties

$$\rightsquigarrow \quad 0 \leq p_{ij}(n) \leq 1$$

$$\rightsquigarrow \quad \sum_{\text{all } j} p_{ij}(n)$$

A matrix that satisfies this properties is a **Markov** or **stochastic matrix**

Generalities (cont.)

(Time-) homogeneous Markov chain

A Markov chain is **time-homogeneous**, if for all states i and j, we have

$$\rightsquigarrow \quad \text{Prob}\{X_{n+1} = j | X_n = i\} = \text{Prob}\{X_{n+m+1} = j | X_{n+m} = i\}$$

- For all $n = 0, 1, 2, \dots$
- For any $m \geq 0$

Generalities (cont.)

That is, we can write the transition probabilities

$$\begin{aligned} p_{ij} &= \text{Prob}\{X_1 = j | X_0 = i\} \\ &= \text{Prob}\{X_2 = j | X_1 = i\} \\ &= \text{Prob}\{X_3 = j | X_2 = i\} \\ &= \cdots \end{aligned}$$

$p_{ij}(n)$ can be (has been) replaced by p_{ij}

$\rightsquigarrow$ (Transitions no longer depend on n)

Generalities (cont.)

(Time-) non-homogeneous Markov chain

A Markov chain is **time-non-homogeneous**, if for all states i and j,

$$\rightsquigarrow \quad p_{ij}(0) = \text{Prob}\{X_1 = j | X_0 = i\} \neq \text{Prob}\{X_2 = j | X_1 = i\} = p_{ij}(1)$$

Generalities (cont.)

Consider a homogeneous discrete-time Markov chain

$$\begin{aligned} \rightsquigarrow \quad p_{ij}(n) &= \text{Prob}\{X_{n+1} = j | X_n = i\} \\ &= \text{Prob}\{X_{n+1} = j | X_n = i\} = p_{ij}, \\ &\qquad \text{for all } n = 0, 1, 2, \dots \text{ (and for all } i \text{ and } j) \end{aligned}$$

$\rightsquigarrow$ Transition probabilities are independent of n

Thus, matrix $P(n)$ can be replaced with matrix P

$$\rightsquigarrow \quad P = \begin{matrix} 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \vdots \end{matrix} \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0j} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1j} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2j} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i0} & p_{i1} & p_{i2} & \cdots & p_{ij} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Generalities (cont.)

We now consider the **time-evolution of the Markov chain**

We are given a (potentially infinite) set of time steps, $0, 1, \cdots$

We assume that at some initial time 0 the chain is in state i

- ⇝ The chain may change state at each step
- ⇝ (But, only at those time steps)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Discrete-time Markov chains (cont.)

$$P(n) = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \vdots \end{array} \begin{pmatrix} p_{00}(n) & p_{01}(n) & p_{02}(n) & \cdots & p_{0j}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & p_{12}(n) & \cdots & p_{1j}(n) & \cdots \\ p_{20}(n) & p_{21}(n) & p_{22}(n) & \cdots & p_{2j}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i0}(n) & p_{i1}(n) & p_{i2}(n) & \cdots & p_{ij}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Suppose that at time step n the chain is in state i

- At step $n+1$, the chain will be in state j
- ⇝ with probability $p_{ij}(n)$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

$$P(n) = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \vdots \end{array} \begin{pmatrix} p_{00}(n) & p_{01}(n) & p_{02}(n) & \cdots & p_{0j}(n) & \cdots \\ p_{10}(n) & p_{11}(n) & p_{12}(n) & \cdots & p_{1j}(n) & \cdots \\ p_{20}(n) & p_{21}(n) & p_{22}(n) & \cdots & p_{2j}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i0}(n) & p_{i1}(n) & p_{i2}(n) & \cdots & p_{ij}(n) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Suppose that $p_{ii}(n) > 0$, the chain is said to have a **self-loop**

- Suppose that at time step n the chain is in state i
- At step $n+1$, the chain will remain in its state i
- ⇝ with probability $p_{ii}(n)$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

The probability of being in state j at step $n+1$ and in state k at step $n+2$

- Given that the state at time step n is i

We have,

$$\begin{aligned} \rightsquigarrow \quad & \text{Prob}\{X_{n+2} = k, X_{n+1} = j | X_n = i\} \\ &= \text{Prob}\{X_{n+2} = k | X_{n+1} = j, X_n = i\}\text{Prob}\{X_{n+1} = j | X_n = i\} \\ &= \text{Prob}\{X_{n+2} = k | X_{n+1} = j\}\text{Prob}\{X_{n+1} = j | X_n = i\} \\ &= p_{jk}(n+1)p_{ij}(n) \end{aligned}$$

We used the definition of conditional probability and then Markovianity

Generalities (cont.)

Sample path

A sequence of states visited by the chain is called a **sample path**

The probability of the sample path $i \to j \to k$, given state i at step n

$$\rightsquigarrow \quad p_{jk}(n+1)p_{ij}(n)$$

More generally, for a sample path $i \to j \to k \to \cdots \to b \to a$

$$\begin{aligned} &\text{Prob}\{X_{n+m} = a, X_{n+m-1} = b, \cdots, X_{n+2} = k, X_{n+1} = j | X_n = i\} \\ &= \text{Prob}\{X_{n+m} = a | X_{n+m-1} = b\}\text{Prob}\{X_{n+m-1} = b | X_{n+m-2} = c\} \\ &\cdots \text{Prob}\{X_{n+2} = k | X_{n+1} = j\}\text{Prob}\{X_{n+1} = j | X_n = i\} \\ &= p_{ba}(n+m-1)p_{cb}(n+m-2) \\ &\cdots p_{jk}(n+1)p_{ij}(n) \end{aligned} \quad (4)$$

Generalities (cont.)

$$\begin{aligned} &\text{Prob}\{X_{n+m} = a, X_{n+m-1} = b, \cdots, X_{n+2} = k, X_{n+1} = j | X_n = i\} \\ &= \text{Prob}\{X_{n+m} = a | X_{n+m-1} = b\}\text{Prob}\{X_{n+m-1} = b | X_{n+m-2} = c\} \\ &\cdots \text{Prob}\{X_{n+2} = k | X_{n+1} = j\}\text{Prob}\{X_{n+1} = j | X_n = i\} \\ &= p_{ba}(n+m-1)p_{cb}(n+m-2) \\ &\cdots p_{jk}(n+1)p_{ij}(n) \end{aligned}$$

Consider the case of a chain that is homogeneous

We have,

$$\begin{aligned} &\text{Prob}\{X_{n+m} = a, X_{n+m-1} = b, \cdots, X_{n+2} = k, X_{n+1} = j | X_n = i\} \\ &= p_{ij}p_{jk}\cdots p_{cb}p_{ba}, \quad \text{(for all possible values of } n) \end{aligned}$$

Generalities (cont.)

Markov chains are commonly depicted using some graphical device

- ⇝ **Transition diagram**
- ⇝ Nodes are used to represent states
- ⇝ Directed edges represent single-state transitions
- ⇝ Edges can be labelled to show transition probabilities

The absence of an edge indicates no single-step transition

Generalities (cont.)

Example

A weather model

Consider an application of a homogenous discrete-time Markov chain

- We use a Markov chain to describe the weather in some place

We simplify weather, three types of weather only

- ⇝ Rainy (R), Cloudy (C) and Sunny (S)
- ⇝ The (three) states of the Markov chain

We assume that the weather is observed daily

We assume the chain is time-homogeneous

Generalities (cont.)

We are given values for the transition probabilities

We have,

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix} \qquad (5)$$

$P(i,j) = p_{ij}$ is the conditional probability that given that the chain (weather) is in state i in one time it will be found in state j after one time step

$$\rightsquigarrow \quad \text{Prob}\{X_{n+1} = j | X_n = i\} = p_{ij}$$

Generalities (cont.)

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

We can calculate tomorrow's weather, given today's weather

$$\rightsquigarrow \quad \text{Prob}\{X_{n+1} = C | X_n = S\} = p_{SC} = 0.30$$

$\rightsquigarrow$ (The probability of the sample path $S \rightarrow C$)

Generalities (cont.)

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

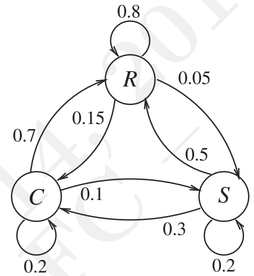
We can calculate the weather in the next two days, given today's weather

$$\begin{aligned} \rightsquigarrow \quad & \text{Prob}\{X_{n+2} = R, X_{n+1} = C | X_n = S\} \\ &= \text{Prob}\{X_{n+2} = R | X_{n+1} = C, X_n = S\}\text{Prob}\{X_{n+1} = C | X_n = S\} \\ &= \underbrace{\text{Prob}\{X_{n+2} = R | X_{n+1} = C\}}_{p_{CR}} \underbrace{\text{Prob}\{X_{n+1} = C | X_n = S\}}_{p_{SC}} \\ &= p_{SC}p_{CR} = 0.30 \cdot 0.70 = 0.21 \end{aligned}$$

$\rightsquigarrow$ (The probability of the sample path $S \rightarrow C \rightarrow R$)

Generalities (cont.)

The transition diagram



The transition probability matrix

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

Generalities (cont.)

Example

The fate of data scientists

The career destiny of a data scientist across years of work

Three levels of competence/status were established

- ⇝ Wizard (W)
- ⇝ Regular (R)
- ⇝ Poser (P)

Career can be modelled as a discrete-time Markov chain $\{X_n, n \geq 0\}$

- RV X_n models the status at the n-th year

Transition probabilities from year to year (one-step) were estimated

$$P = \begin{matrix} W \\ R \\ P \end{matrix} \begin{pmatrix} p_{WW} & p_{WR} & p_{WP} \\ p_{RW} & p_{RR} & p_{RP} \\ p_{PW} & p_{PR} & p_{PP} \end{pmatrix} = \begin{pmatrix} 0.85 & 0.14 & 0.01 \\ 0.05 & 0.85 & 0.10 \\ 0.00 & 0.20 & 0.80 \end{pmatrix}$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

$$P = \begin{matrix} W \\ R \\ P \end{matrix} \begin{pmatrix} p_{WW} & p_{WR} & p_{WP} \\ p_{RW} & p_{RR} & p_{RP} \\ p_{PW} & p_{PR} & p_{PP} \end{pmatrix} = \begin{pmatrix} 0.85 & 0.14 & 0.01 \\ 0.05 & 0.85 & 0.10 \\ 0.00 & 0.20 & 0.80 \end{pmatrix}$$

The probability that a poser at time n becomes a wizard in 2-year time

- After becoming a regular after one year

This is the probability of the sample path $P \rightarrow R \rightarrow W$

$$\begin{aligned} \rightsquigarrow \quad & \text{Prob}\{X_{n+2} = W, X_{n+1} = R | X_n = P\} \\ &= \text{Prob}\{X_{n+2} = W | X_{n+1} = R, X_n = P\}\text{Prob}\{X_{n+1} = R | X_n = P\} \\ &= \underbrace{\text{Prob}\{X_{n+2} = W | X_{n+1} = R\}}_{p_{RW}} \underbrace{\text{Prob}\{X_{n+1} = R | X_n = P\}}_{p_{PR}} \\ &= p_{PR}p_{RW} = 0.20 \cdot 0.05 = 0.01 \end{aligned}$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

$$P = \begin{matrix} W \\ R \\ P \end{matrix} \begin{pmatrix} p_{WW} & p_{WR} & p_{WP} \\ p_{RW} & p_{RR} & p_{RP} \\ p_{PW} & p_{PR} & p_{PP} \end{pmatrix} = \begin{pmatrix} 0.85 & 0.14 & 0.01 \\ 0.05 & 0.85 & 0.10 \\ 0.00 & 0.20 & 0.80 \end{pmatrix}$$

According to this (any?) model, a poser cannot move in one year to wizard

■

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Example

The Ehrenfest model

Suppose that you have two boxes containing a total of N small balls in it

At each time instant, a ball is chosen at random from one of the boxes

- Then, the ball is moved into the other box

The state of the system is the number of balls X_n in the first box

- After n selections

This is a Markov chain, X_{n+1} depends on $X_n = x_n$ only

$$\{X_n, n = 1, 2, \dots\}$$

Generalities (cont.)

X_n is the random variable 'balls in first box, after n selections'

Let $k < N$ be the number of balls in the first box at step n

The probability of $k+1$ balls after the next step

$$\rightsquigarrow \quad \text{Prob}\{X_{n+1} = k+1 | X_n = k\} = \frac{N-k}{N}$$

To increase the number of balls in the first box by one, one of the $N-k$ balls in the second box must be selected, at random, with probability $(N-k)/N$

The probability of $k-1$ balls after the next step

$$\rightsquigarrow \quad \text{Prob}\{X_{n+1} = k-1 | X_n = k\} = \frac{k}{N} \text{ (for } k \geq 1)$$

To decrease by one the number of balls in the first box, one of the k balls must be selected, at random, with probability k/N

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Suppose that $N = 6$

$$P = \begin{array}{c} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{pmatrix} 0 & 6/6 & 0 & 0 & 0 & 0 & 0 \\ 1/6 & 0 & 5/6 & 0 & 0 & 0 & 0 \\ 0 & 2/6 & 0 & 4/6 & 0 & 0 & 0 \\ 0 & 0 & 3/6 & 0 & 3/6 & 0 & 0 \\ 0 & 0 & 0 & 4/6 & 0 & 2/6 & 0 \\ 0 & 0 & 0 & 0 & 5/6 & 0 & 1/6 \\ 0 & 0 & 0 & 0 & 0 & 6/6 & 0 \end{pmatrix}$$

The transition probability matrix has seven rows and seven columns

- ⇝ They correspond to the states
- ⇝ (Zero through six)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Example

A two-state non-homogeneous Markov chain

Consider a two-state discrete-time Markov chain $\{X_n, n = 1, 2, \dots\}$

- Let a and b be the states X_n can take on

Let $p_{aa}(n) = p_{bb}(n)$ be the probability to keep the current state (time n)

$$p_{aa}(n) = p_{bb}(n) = 1/n$$

The probability $p_{ab}(n) = p_{ba}(n)$ to change state is given by the complement

$$p_{ab}(n) = p_{ba}(n) = (n-1)/n$$

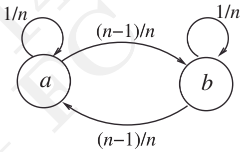
Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Thus, the n-th transition matrix

$$p(n) = \begin{array}{c} a \\ b \end{array} \begin{pmatrix} p_{aa}(n) & p_{ab}(n) \\ p_{ba}(n) & p_{bb}(n) \end{pmatrix} = \begin{pmatrix} 1/n & (n-1)/n \\ (n-1)/n & 1/n \end{pmatrix}$$

The transition diagram for this non-homogeneous Markov chain



The probability of changing state increases at each time step

- (That of remaining, decreases)

Generalities (cont.)

$$p(n) = \begin{matrix} a \\ b \end{matrix}\begin{pmatrix} p_{aa}(n) & p_{ab}(n) \\ p_{ba}(n) & p_{bb}(n) \end{pmatrix} = \begin{pmatrix} 1/n & (n-1)/n \\ (n-1)/n & 1/n \end{pmatrix}$$

The first four transition probability matrices

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

$$P(2) = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

$$P(3) = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}$$

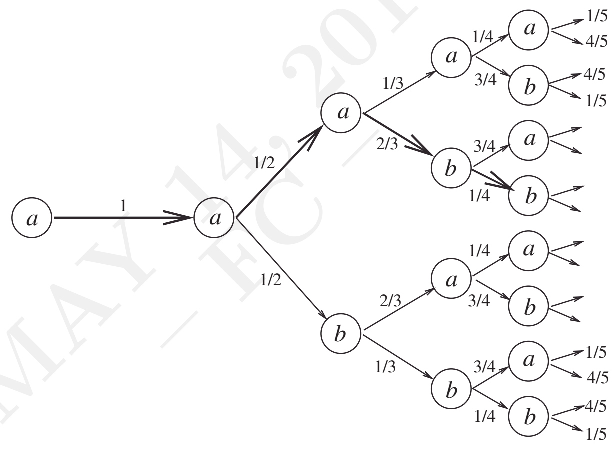
$$P(4) = \begin{pmatrix} 1/4 & 3/4 \\ 3/4 & 1/4 \end{pmatrix}$$

Probabilities change with time, yet the process is Markovian

- At any step, future evolution only depends on present

Generalities (cont.)

The collection of sample paths, beginning from state a



Generalities (cont.)

Consider paths that begins in state a, stay in a after the first and second time steps, move to state b on third step and then remain in b on the fourth

$$a \to a \to a \to b \to b$$

The probability of the path is the product of the probabilities of the segments

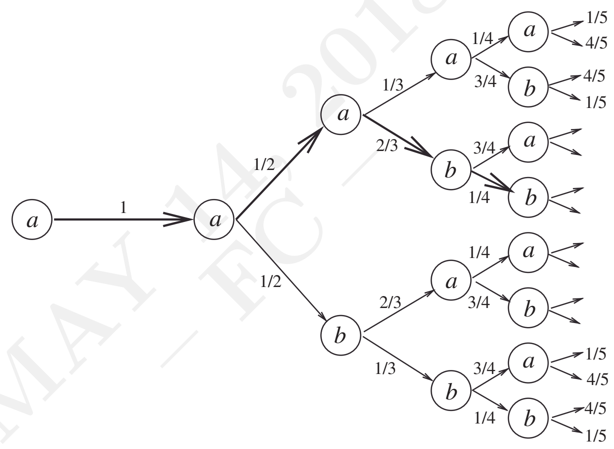
$$\text{Prob}\{X_5 = b, X_4 = b, X_3 = a, X_2 = a | X_1 = a\} = \underbrace{p_{aa}(1)}_{1}\underbrace{p_{aa}(2)}_{1/2}\underbrace{p_{ab}(3)}_{2/3}\underbrace{p_{bb}(4)}_{1/4} = 1/12$$

There exist other paths that take the chain from state a to b in four steps

$\rightsquigarrow$ They are assigned different probabilities

Generalities (cont.)

No matter which path is chosen, once the chain arrives to state b after four steps, the future evolution is specified by $P(5)$ and not any other $P(i), i \leq 4$



All transition probabilities leading out of b are the same

Generalities (cont.)

k-dependent Markov chains

A process is not Markovian if evolution depends on more than current state

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Example

A weather model

Consider the simplified weather model and suppose the following

- Transition at $n+1$ depends on state at time n and $n-1$

We have been given one-step probabilities given two rainy days in a row

- The probabilities the next day be rainy, cloudy or sunny

$$\rightsquigarrow \quad (0.6, 0.3, 0.1)$$

And, one-step probabilities given a sunny day followed by a rainy day

- The probabilities the next day be rainy, cloudy or sunny

$$\rightsquigarrow \quad (0.80, 0.15, 0.05)$$

And, one-step probabilities given a cloudy day followed by a rainy day

- The probabilities the next day be rainy, cloudy or sunny

$$\rightsquigarrow \quad (0.80, 0.15, 0.05)$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

The transition probabilities depend on today's and also yesterday's weather

$\rightsquigarrow$ The process is not a (first-order) Markovian process

We can still transform this process into a (first-order) Markov chain

$\rightsquigarrow$ We must increase the number of states

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

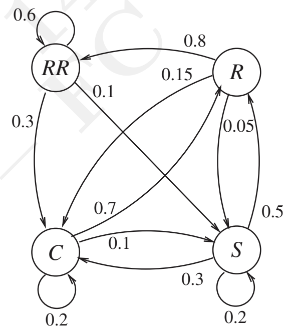
The probability transition matrix of the original process

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

Consider the case in which we add a single extra state

- State RR, two consecutive days of rain

We assumed original probabilities remain unchanged



Discrete-time Markov chains (cont.)

This device converts non-Markovian processes into Markovian ones

⇝ It can be generalised

Consider a process with s states with dependence on two prior steps

We can define a new (now first-order) process with s^2 states

- Each new state characterise the weather two days back

For the simplified weather model

⇝ RR, RC and RS

⇝ CR, CC and CS

⇝ SR, SC and SS

Discrete-time Markov chains (cont.)

Consider a process that has s states and k-step back dependencies

- We can build a first-order Markov process with s^k states

Let $\{X_n, n \geq 0\}$ be a stochastic process and let k be an integer

$$\begin{aligned}\text{Prob}\{&X_{n+1} = x_{n+1} | \\ &X_n = x_n, \cdots, X_{n-k+1} = x_{n-k+1}, X_{n-k} = x_{n-k}, \cdots, X_0 = x_0\} \\ &= \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, \cdots, X_{n-k+1} = x_{n-k+1}\}\end{aligned}$$

(for all $n \geq k$)

The evolution of the process depends on the k previously visited states

⇝ **k-dependent process** or **k-order process**

Discrete-time Markov chains (cont.)

$$\begin{aligned}\text{Prob}\{&X_{n+1} = x_{n+1} | \\ &X_n = x_n, \cdots, X_{n-k+1} = x_{n-k+1}, X_{n-k} = x_{n-k}, \cdots, X_0 = x_0\} \\ &= \text{Prob}\{X_{n+1} = x_{n+1} | X_n = x_n, \cdots, X_{n-k+1} = x_{n-k+1}\}\end{aligned}$$

(for all $n \geq k$)

For $k = 1$, $\{X_n\}$ is a Markov process

For $k > 1$, $\{Y_n\}$ is Markov process

$$\rightsquigarrow \quad Y_n = (X_n, X_{n+1}, \cdots, X_{n+k-1})$$

The states Y_n are elements of the Cartesian-product of states

$$\rightsquigarrow \quad \underbrace{\mathcal{S} \times \mathcal{S} \times \cdots \times \mathcal{S}}_{k \text{ terms}}$$

$\mathcal{S}$ denotes the original set of states X_n

Discrete-time Markov chains (cont.)

Holding/sojourn time

Consider the diagonal elements of the transition probability matrix P

Suppose that they are all non-zero and strictly smaller than one

$$\rightsquigarrow \quad p_{ii} \in (0, 1)$$

⇝ At any step the chain may remain in its current state

Discrete-time Markov chains (cont.)

We define the number of consecutive steps a chain remains in a state

- $\rightsquigarrow$ **Sojourn time**, or **holding time** of that state

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

At each time, the probability of leaving state i is past-independent

For a homogeneous Markov chain, we have

$$\rightsquigarrow \quad \sum_{i \neq j} p_{ij} = 1 - p_{ii}$$

The process may be understood as a sequence of Bernoulli trials

- $\rightsquigarrow$ The probability of success is $(1 - p_{ii})$
- $\rightsquigarrow$ (Success means 'exit from i')

MAY 14, 2018 _ FC

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Suppose that the probability that the holding time is equal to k steps

- $\rightsquigarrow$ $(k-1)$ consecutive Bernoulli failures, and one success

The probability R_i of the holding time of state i,

$$\rightsquigarrow \quad \text{Prob}\{R_j = k\} = (1 - p_{ii}) p_{ii}^{k-1}$$

(for $k = 1, 2, \dots,$ and 0 elsewhere)

This corresponds to the geometric distribution with parameter $(1 - p_{ii})$

- $\rightsquigarrow$ Its distinguished feature is the memoryless property
- $\rightsquigarrow$ (No other discrete RVs has it)

MAY 14, 2018 _ FC

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Mean and variance of holding time of a homogeneous chain in state i

$$\rightsquigarrow \quad \text{E}[R_i] = \frac{1}{1 - p_{ii}}$$

$$\rightsquigarrow \quad \text{Var}[R_i] = \frac{p_{ii}}{(1 - p_{ii})^2}$$

This is only valid for homogeneous Markov processes (it is important)

MAY 14, 2018 _ FC

Generalities (cont.)

Memoryless property

A sequence of $j-1$ unsuccessful trials has no effect on success at j-th trial

$\rightsquigarrow$ (On the probability of success, to be exact)

Discrete-time Markov chains

UFC/DC
SA (CK0191)
2018.1

Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Consider a non-homogeneous Markov chains in state i at time step n

Let $R_i(n)$ be the RV representing the remaining number of steps in i

$R_i(n)$ is not distributed according to a geometric distribution

$$\rightsquigarrow \quad \text{Prob}\{R_i(n) = k\} = p_{ii}(n)p_{ii}(n+1)\cdots p_{ii}(n+k-2)\big[1 - p_{ii}(n+k-1)\big]$$

(It becomes geometric for $p_{ii}(n) = p_{ii}$, for all n)

Discrete-time Markov chains

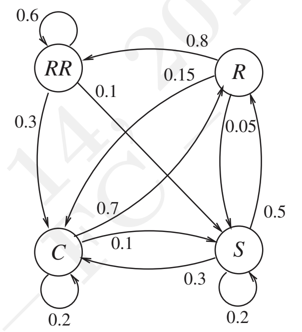
UFC/DC
SA (CK0191)
2018.1

Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Generalities (cont.)

Embedded Markov chains

Consider the transition diagram of the extended weather model



There are no self-loops on state R (other states have them)

- A rainy (R) day is followed by a rainy day (state RR)
- A rainy (R) day is followed by a cloudy day (C)
- A rainy (R) day is followed by a sunny day (S)

Discrete-time Markov chains

UFC/DC
SA (CK0191)
2018.1

Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Discrete-time Markov chains (cont.)

We consider an alternative way of consider this homogeneous Markov chain

So far, we only considered the process evolution at each time steps

- (Including self-transitions, at any step)

Consider evolution at those steps in which there is an actual change

- We neglect (postpone) self-transitions

Generalities (cont.)

To embed this interpretation we modify the transition probability matrix

For all rows i for which $p_{ii} \in (0,1)$, we remove the self loop (set $p_{ii} = 0$)

- Then, we must replace p_{ij} with conditional probabilities
- Condition of the fact that transitions are out of state i

For $i \neq j$, we have

$$\rightsquigarrow \quad p_{ij} = \frac{p_{ij}}{(1-p_{ii})}$$

This keeps the transition matrix row-stochastic, for all i

$$\sum_j \frac{p_{ij}}{1-p_{ii}} = \frac{p_{ii}}{1-p_{ii}} + \sum_{j \neq i} \frac{p_{ij}}{1-p_{ij}}$$
$$= \sum_{j\neq i} \frac{p_{ij}}{1-p_{ii}} = \frac{1}{1-p_{ii}} \sum_{j\neq i} p_{ij}$$
$$= \frac{1}{1-p_{ii}}(1-p_{ii}) = 1$$

Generalities (cont.)

The new process is called the **embedded process**

The embedding points (of time) are those at which transitions occur

- (Transitions out of state, that is)

It is fully characterised by adding the holding times for all states

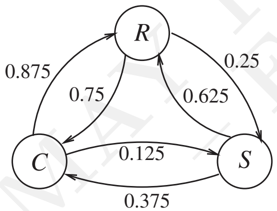
Generalities (cont.)

Example

Consider the transition probability matrix of the simplified weather model

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

The transition diagram and the transition matrix of the embedded chain



$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} 0.000 & 0.750 & 0.250 \\ 0.875 & 0.000 & 0.125 \\ 0.625 & 0.375 & 0.000 \end{pmatrix}$$

The probabilities of success for the geometric distribution

- The holding times, 0.2, 0.8 and 0.8

Chapman-Kolmogorov equations

Discrete-time Markov chains

Chapman-Kolmogorov equations

Suppose that we have a discrete-time Markov chain

We are interested in computing path probabilities
⇝ Starting from some given state, onward
We derived a general expression for that

For a sample path $i \rightarrow j \rightarrow k \rightarrow \cdots \rightarrow b \rightarrow a$,

$$\begin{aligned} &\text{Prob}\{X_{n+m} = a, X_{n+m-1} = b, \cdots, X_{n+2} = k, X_{n+1} = j | X_n = i\} \\ &= \text{Prob}\{X_{n+m} = a | X_{n+m-1} = b\}\text{Prob}\{X_{n+m-1} = b | X_{n+m-2} = c\} \cdots \\ &\quad \cdots \text{Prob}\{X_{n+2} = k | X_{n+1} = j\}\text{Prob}\{X_{n+1} = j | X_n = i\} \\ &\quad = p_{ba}(n+m-1)p_{cb}(n+m-2)\cdots \\ &\quad\quad \cdots p_{jk}(n+1)p_{ij}(n) \end{aligned}$$

Chapman-Kolmogorov equations (cont.)

For the homogeneous case, the expression simplifies

For a sample path $i \rightarrow j \rightarrow k \rightarrow \cdots \rightarrow b \rightarrow a$

$$\begin{aligned} &\text{Prob}\{X_{n+m} = a, X_{n+m-1} = b, \cdots, X_{n+2} = k, X_{n+1} = j | X_n = i\} \\ &\quad = p_{ij}p_{jk}\cdots p_{cb}p_{ba}, \quad \text{(for all possible values of } n) \end{aligned}$$

Chapman-Kolmogorov equations (cont.)

Example

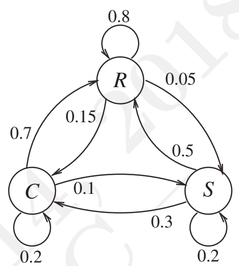
A weather model

Consider the simplified weather model for some location
- Daily observations, time-homogenous Markov chain
- Rainy (R), Cloudy (R) and Sunny S

The transition probability matrix

$$P = \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

Chapman-Kolmogorov equations (cont.)



Probability of rain tomorrow and clouds the day after, given today is sunny

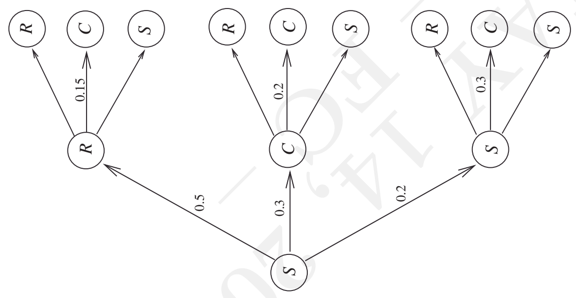
The probability of the sample path $S \rightarrow R \rightarrow C$

$$\begin{aligned} \rightsquigarrow \quad &\text{Prob}\{X_{n+2} = C, X_{n+1} = R | X_n = S\} \\ &= \text{Prob}\{X_{n+2} = C | X_{n+1} = R, X_n = S\}\text{Prob}\{X_{n+1} = R | X_n = S\} \\ &\quad = \text{Prob}\{X_{n+2} = C | X_{n+1} = R\}\text{Prob}\{X_{n+1} = R | X_n = S\} \\ &\quad\quad = p_{SR}p_{RC} = 0.50 \cdot 0.15 = 0.075 \end{aligned}$$

Chapman-Kolmogorov equations (cont.)

What the probability that it will be cloudy in two days, given today is sunny

$\rightsquigarrow$ The probability of the sample path $S \to \times \to C$



We must examine all sample paths (those that include intermediate weather)

$\rightsquigarrow$ The sample path $S \to S \to C$

$\rightsquigarrow$ The sample path $S \to C \to C$

$\rightsquigarrow$ The sample path $S \to R \to C$

Chapman-Kolmogorov equations (cont.)

$$P = \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

The weather on the next (2) day can be through any of those three possibilities

- The events are all mutually exclusive
- Probabilities can be summed up

Thus, the probability that it will be C in two days, given that today is S

$$\rightsquigarrow \quad p_{SR}p_{RC} + p_{SC}p_{CC} + p_{SS}p_{SC} = \sum_{w \in \{R,C,S\}} p_{Sw}p_{wC}$$

$$= \underbrace{0.5 \cdot 0.15}_{0.075} + \underbrace{0.3 \cdot 0.2}_{0.06} + \underbrace{0.2 \cdot 0.3}_{0.06} = 0.195$$

Chapman-Kolmogorov equations (cont.)

$$p_{SR}p_{RC} + p_{SC}p_{CC} + p_{SS}p_{SC} = \sum_{w \in \{R,C,S\}} p_{Sw}p_{wC}$$

$$= \sum_{w \in \{R,C,S\}} \underbrace{\text{Prob}\{X_{n+2} = C | X_{n+1} = w\}}_{p_{wC}} \underbrace{\text{Prob}\{X_{n+1} = w | X_n = S\}}_{p_{Sw}}$$

This result can be readily obtained from the single-step transition matrix

$$P = \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

It is an inner product of one row and one column of P

Row 3, transitions from a sunny S day into a w-day, Sw

$\rightsquigarrow$ $r_3 = (0.50, 0.30, 0.20)$

Column 2, transitions from a w-day into a cloudy day C, wC

$\rightsquigarrow$ $c_2 = (0.15, 0.20, 0.30)^T$

$$\rightsquigarrow \quad r_3 c_2 = 0.5 \cdot 0.15 + 0.3 \cdot 0.2 + 0.2 \cdot 0.3 = 0.195$$

Chapman-Kolmogorov equations (cont.)

Let today's weather be given by w_1

$\rightsquigarrow$ With $w_1 \in \{R, C, S\}$

Let the weather in two days be w_2

$\rightsquigarrow$ With $w_2 \in \{R, C, S\}$

The probability that the weather in two days is w_2

- The product between row w_1 and column w_2

All possibilities ($\{R, C, S\} \times \{R, C, S\}$) make the elements of matrix $P^2 = PP$

$$\rightsquigarrow \quad P^2 = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix} \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

$$= \begin{pmatrix} 0.770 & 0.165 & 0.065 \\ 0.750 & 0.175 & 0.075 \\ 0.710 & 0.195 & 0.095 \end{pmatrix}$$

Chapman-Kolmogorov equations (cont.)

$$P^2 = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} 0.770 & 0.165 & 0.065 \\ 0.750 & 0.175 & 0.075 \\ 0.710 & 0.195 & 0.095 \end{pmatrix}$$

Element $(2,3)$ gives the probability that it will be sunny (3) in two days

- Given that today is cloudy (2)

More generally, element (i,j) is the probability of state j in two steps

- Given the current state i

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

In the same way, we can construct matrix $P^3 = P(PP)$

$$\rightsquigarrow P^3 = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix} \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix} \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

$$= \begin{pmatrix} 0.764 & 0.168 & 0.068 \\ 0.760 & 0.170 & 0.070 \\ 0.752 & 0.174 & 0.074 \end{pmatrix}$$

It gives conditional probabilities three steps from now

And, so on with higher powers of P

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

For this particular example, this sequence converges to a particular matrix

$\rightsquigarrow$ All rows become identical

$$\rightsquigarrow \quad \lim_{n\to\infty} P^n = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} 0.7625 & 0.16875 & 0.06875 \\ 0.7625 & 0.16875 & 0.06875 \\ 0.7625 & 0.16875 & 0.06875 \end{pmatrix} \qquad (6)$$

■

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

The results we obtained are mere applications of the laws of probability

$\rightsquigarrow$ We summed over all sample paths of length n, from i to k

Chapman-Kolmogorov equations (cont.)

$$P = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ i \\ \vdots \end{array} \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots & p_{0j} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots & p_{1j} & \cdots \\ p_{20} & p_{21} & p_{22} & \cdots & p_{2j} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{i0} & p_{i1} & p_{i2} & \cdots & p_{ij} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

For a homogeneous discrete-time Markov chain, for any $n = 0, 1, 2, \dots$

$$\text{Prob}\{X_{n+2} = k, X_{n+1} = j | X_n = i\} = p_{ij} p_{jk}$$

From the laws of probability,

$$\rightsquigarrow \quad \text{Prob}\{X_{n+2} = k | X_n = i\} = \sum_{\text{all } j} p_{ij} p_{jk} = p_{ik}^{(2)}$$

The RHS defines element (i, k) of matrix $P^2 = PP$

$$\rightsquigarrow \quad p_{ik}^{(2)} \equiv (P^2)_{ik}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

We may proceed in the same fashion for sample path of arbitrary length

$$\text{Prob}\{X_{n+3} = l, X_{n+2} = k, X_{n+1} = j | X_n = i\} = p_{ij} p_{jk} p_{kl}$$

Again, from the laws of probability,

$$\rightsquigarrow \quad \text{Prob}\{X_{n+3} = l | X_n = i\} = \sum_{\text{all } j} \sum_{\text{all } k} p_{ij} p_{jk} p_{kl}$$

$$= \sum_{\text{all } j} p_{ij} \sum_{\text{all } k} p_{jk} p_{kl} = \sum_{\text{all } j} p_{ij} p_{jl}^{(2)} = p_{il}^{(3)}$$

The RHS is the (i, l)-th element of matrix $P^3 = PPP$

$$\rightsquigarrow \quad p_{il}^{(3)} \equiv (P^3)_{il}$$

Above, we had that $p_{jl}^{(2)}$ is the (i, l)-th element of P^2

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

We can fully generalise the single-step transition probability matrix

For a homogeneous discrete-time Markov chain, the m-step matrix

$$p_{ij}^{(m)} = \text{Prob}\{X_{n+m} = j | X_n = i\}$$

$p_{ij}^{(m)}$ is obtained from the single-step transition probability matrix

$$\rightsquigarrow \quad p_{ij}^{(m)} = \sum_{\text{all } k} p_{ik}^{(l)} p_{kj}^{(m-l)}, \quad (\text{for } 0 < l < m)$$

The recursive formula is the **Chapman-Kolmogorov equation**

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

Consider a homogenous discrete-time Markov chain

We have,

$$p_{ij}^{(m)} = \text{Prob}\{X_m = j | X_0 = i\}$$

$$= \sum_{\text{all } k} \text{Prob}\{X_m = j, X_l = k | X_0 = i\}, \quad (\text{for } 0 < l < m)$$

$$= \sum_{\text{all } k} \text{Prob}\{X_m = j | X_l = k, X_0 = i\} \text{Prob}\{X_l = k | X_0 = i\}, \quad (\text{for } 0 < l < m)$$

By using the Markov property, we get

$$\rightsquigarrow \quad p_{ij}^{(m)}$$

$$= \sum_{\text{all } k} \text{Prob}\{X_m = j | X_l = k\} \text{Prob}\{X_l = k | X_0 = i\}, \quad (\text{for } 0 < l < m)$$

$$= \sum_{\text{all } k} p_{kj}^{(m-l)} p_{ik}^{(l)}, \quad (\text{for } 0 < l < m)$$

Chapman-Kolmogorov equations (cont.)

The Chapman-Kolmogorov equations can be compacted, in matrix notation

$$\rightsquigarrow \quad P^{(m)} = P^{(l)} P^{(m-l)}$$

By definition, $P^{(0)} = I$

It is possible to write any m-step homogeneous transition probability matrix

- $\rightsquigarrow$ The sum of products of l-step and $(m-l)$-step transition matrices

Chapman-Kolmogorov equations (cont.)

Consider how we travelled the space-time, from i to j in m steps

- $\rightsquigarrow$ First, go from i to whatever intermediate state k in l steps
- $\rightsquigarrow$ Then, bang from k to j in the remaining $(m-l)$ steps

We considered all possible distinct paths from i to j in m steps

- $\rightsquigarrow$ By summing over all intermediate states k

Chapman-Kolmogorov equations (cont.)

Consider a non-homogeneous discrete-time Markov chain

Matrices $P(n)$ depend on time step n

- $\rightsquigarrow$ P^2 must be replaced by $P(n)P(n+1)$
- $\rightsquigarrow$ P^3 must be replaced by $P(n)P(n+1)P(n+2)$
- $\rightsquigarrow$ $\cdots$

Thus, we have

$$\rightsquigarrow \quad P^{(m)}(n, n+1, \cdots, n+m-1) = P(n)P(n+1)\cdots P(n+m-1)$$

Element (i,j) is $\text{Prob}\{X_{n+m} = j | X_n = i\}$

Chapman-Kolmogorov equations (cont.)

Let $\pi_i^{(0)}$ be the probability that the Markov chain begins in state i

- $\rightsquigarrow$ $\pi^{(0)}$ is a row-vector, the i-th element is $\pi_i^{(0)}$

The probabilities of being in any state j after the first time-step

$$\rightsquigarrow \quad \pi^{(1)} = \pi^{(0)} P(0)$$

- $\rightsquigarrow$ The j-th element $\pi_j^{(1)}$ of $\pi^{(1)}$

Chapman-Kolmogorov equations (cont.)

Consider the homogeneous discrete-time Markov chain

We have,

$$\rightsquigarrow \quad \pi^{(1)} = \pi^{(0)} P$$

The elements of vector $\pi^{(1)}$, probabilities after first step

- For all the various states (the probability distribution)

Chapman-Kolmogorov equations (cont.)

Example

A weather model

Consider the simplified weather model for some location

- Daily observations, time homogenous Markov chain
- Rainy (R), Cloudy (R) and Sunny S

The transition probability matrix

$$P = \begin{matrix} R \\ C \\ S \end{matrix} \begin{pmatrix} p_{RR} & p_{RC} & p_{RS} \\ p_{CR} & p_{CC} & p_{CS} \\ p_{SR} & p_{SC} & p_{SS} \end{pmatrix} = \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

Chapman-Kolmogorov equations (cont.)

Assume that at time 0, we begin with a cloudy weather, $\pi^{(0)} = (0, 1, 0)$

$$\rightsquigarrow \quad \pi^{(1)} = \pi^{(0)} P(0)$$

$$= (0, 1, 0) \begin{pmatrix} 0.80 & 0.15 & 0.05 \\ 0.70 & 0.20 & 0.10 \\ 0.50 & 0.30 & 0.20 \end{pmatrix}$$

$$= (0.7, 0.2, 0.1)$$

This result corresponds to the second row of matrix P (unsurprisingly)

Starting from a cloudy day 0, the probability that day 1 is cloudy is 0.2

Chapman-Kolmogorov equations (cont.)

We can compute the probability of being in any state after two steps

Consider the non-homogeneous case

We have,

$$\rightsquigarrow \pi^{(2)} = \pi^{(1)} P(1) = \underbrace{\pi^{(0)} P(0)}_{\pi^{(1)}} P(1)$$

Consider the homogeneous case

We have,

$$\rightsquigarrow \pi^{(2)} = \pi^{(1)} P = \pi^{(0)} P^2$$

In computing the j-th component of $\pi^{(2)}$, we sum over all sample paths of length 2 that begin, with probability $\pi^{(0)}$, from state i and finish at state j

Chapman-Kolmogorov equations (cont.)

More specifically, in the case of the weather example

For the probability of being in any state after two steps, we have

$$\rightsquigarrow \quad \pi^{(2)} = \pi^{(1)} P = \pi^{(0)} P^2$$
$$= (0, 1, 0) \begin{pmatrix} 0.770 & 0.165 & 0.065 \\ 0.750 & 0.175 & 0.075 \\ 0.710 & 0.195 & 0.095 \end{pmatrix}$$
$$= (0.750, 0.175, 0.075)$$

We summed over all sample paths of length two that begin and end in C

The probability that it will be cloudy after two days, from cloudy, 0.175

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

We can compute the full state probability distribution after any n steps

Consider the non-homogeneous case

We have,

$$\rightsquigarrow \quad \pi^{(n)} = \pi^{(n-1)} P(n-1) = \pi^{(0)} P(0) P(1) \cdots P(n-1)$$

Consider the homogeneous case

We have,

$$\rightsquigarrow \quad \pi^{(n)} = \pi^{(n-1)} P = \pi^{(0)} P^n$$

In computing the j-th component of $\pi^{(n)}$, we sum over all sample paths of length n that begin, with probability $\pi^{(0)}$, from state i and finish at state j

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

Again, for the weather example

Let $n \rightarrow \infty$

$$lim_{n \rightarrow \infty} \pi^{(n)} = \pi^{(0)} \lim_{n \rightarrow \infty} P^n$$
$$= (0, 1, 0) \begin{pmatrix} 0.7625 & 0.16875 & 0.06875 \\ 0.7625 & 0.16875 & 0.06875 \\ 0.7625 & 0.16875 & 0.06875 \end{pmatrix}$$
$$= (0.7625, 0.16875, 0.06875)$$

■

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Chapman-Kolmogorov equations (cont.)

The limit $\lim_{n \rightarrow \infty} \pi^{(n)}$ does not necessarily exist for all Markov chains

$\rightsquigarrow$ Not even for finite-state ones

Classification of states

Discrete-time Markov chains

Classification of states

We shall provide some important definitions regarding the individual states

⇝ We focus on (homogeneous) discrete-time Markov processes

- (We discuss the classification of groups of states later on)

We distinguish between two main types of individual states

⇝ **Recurrent states**

⇝ **Transient states**

Classification of states (cont.)

Informally first,

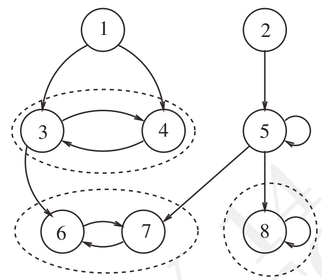
⇝ **Recurrent states**

The Markov chain is guaranteed to return to these states infinitely often

⇝ **Transient states**

The Markov chain has a nonnull probability to never return to such states

Classification of states (cont.)



Some **transient** states

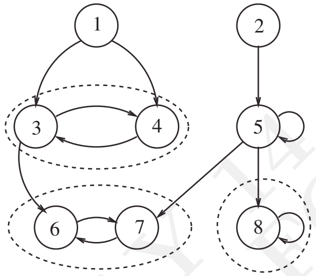
State 1 and 2 are transient states

- The chain can be in state 1 or 2 only at the first step
- States that exist at the first step are **ephemeral**

State 3 and 4 are transient states

- The chain can enter either of these states, move from one to the other
- Eventually, the chain will exit the loops, from state 3, to enter state 6

Classification of states (cont.)

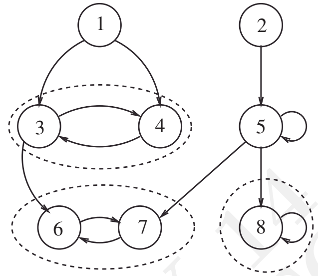


Another **transient** state

State 5 is a transient state

- State 5 can be entered from state 2, at first step if 2 is occupied
- Once in 5, the chain remains in it for a finite number of steps

Classification of states (cont.)



Some **recurrent** states

State 6 and 7 are recurrent states

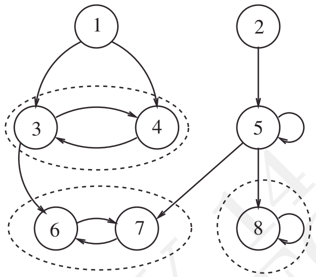
If one of these states is reached, subsequent transitions will start alternating

- When in state 6, the chain returns to state 6 every other transition
- (The same is true for state 7)

Returns to states in this group occur at time steps that are multiples of 2

- Such states are said to be **periodic** (the period is two)

Classification of states (cont.)



Some **recurrent** states

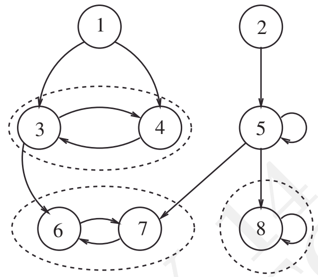
State 6 and 7 are recurrent states (cont.)

Recurrent states can have a finite or an infinite mean recurrence time

- Finite recurrence time, **positive recurrent**
- Infinite recurrence time, **null recurrent**

Infinite recurrence time occurs only in infinite-state Markov chains

Classification of states (cont.)



Another **recurrent** state

State 8 is a recurrent state

- When the chain reaches this state, it will stay there
- Such states are said to be **absorbing**

A state i is an absorbing state if and only if $p_{ii} = 1$

- For non-absorbing states, we have $p_{ii} < 1$
- (Either transient or plain recurrent)

Classification of states (cont.)

We can define the return/non-return properties more formally

Let $p_{jj}^{(n)}$ be the probability that the process is again in state j after n steps

- The process (may have) visited many states (including state j)

$$j \rightarrow \times \rightarrow \cdots \rightarrow \times \rightarrow j$$

We know how to compute this quantity

$$p_{jj}^{(n)} = \text{Prob}\{\text{a return to state } j \text{ occurs } n \text{ steps after leaving it}\}$$
$$= \text{Prob}\{X_n = j, X_{n-1} = \times, \ldots, X_1 = \times | X_0 = j\}$$
$$(\text{for } n = 1, 2, \ldots)$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

We now define/introduce a new conditional probability

Let $f_{jj}^{(n)}$ be probability that the **first-return** to state j occurs in n steps

- This probability is defined on leaving state j

That is,

$$f_{jj}^{(n)} = \text{Prob}\{\text{first return to state } j \text{ occurs } n \text{ steps after leaving it}\}$$
$$= \text{Prob}\{X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j | X_0 = j\}$$
$$(\text{for } n = 1, 2, \ldots)$$

This probability is NOT probability $p_{jj}^{(n)}$ of returning to state j in n steps

- (There, state j may be visited at intermediate steps)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

We relate $p_{jj}^{(n)}$ and $f_{jj}^{(n)}$, then construct a recursive relation to compute $f_{jj}^{(n)}$

We get $p_{jj}^{(n)}$ from powers of the single-step probability transition matrix P

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

Consider the probability $f_{jj}^{(1)}$ of first return to j in one step after leaving it

$\rightsquigarrow$ It is equal to the single-step-probability p_{jj}

$$\rightsquigarrow \quad f_{jj}^{(1)} = p_{jj}^{(1)} = p_{jj}$$

For $n = 1$, compare the two definitions

$$f_{jj}^{(n)} = \text{Prob}\{\text{first return to state } j \text{ occurs } n \text{ steps after leaving it}\}$$
$$= \text{Prob}\{X_n = j, X_{n-1} \neq j, \ldots, X_1 \neq j | X_0 = j\}$$
$$p_{jj}^{(n)} = \text{Prob}\{\text{a return to state } j \text{ occurs } n \text{ steps after leaving it}\}$$
$$= \text{Prob}\{X_n = j | X_0 = j\}$$

Since $p_{jj}^{(0)} = 1$, we can write

$$p_{jj}^{(1)} = f_{jj}^{(1)} p_{jj}^{(0)}$$

Classification of states (cont.)

Consider $p_{jj}^{(2)}$, the probability of being in j, two steps after leaving it

We have two ways of getting there

⇝ The process does not move from state j at either time step

$$j \rightarrow j \rightarrow j$$

⇝ The process leaves j on step 1 and returns on step 2

$$j \rightarrow \times \rightarrow j$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

We can interpret these two possibilities

Case 1 ($j \rightarrow j \rightarrow j$)

The process leaves j and returns to it for the first time after one step (probability $f_{jj}^{(1)}$) and then returns to it at the second step (probability $p_{jj}^{(1)}$)

Case 2 ($j \rightarrow \times \rightarrow j$)

The process leaves j and does not return for the first time until two steps later (probability $f_{jj}^{(2)}$)

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

Thus, by combining these (mutually exclusive) possibilities

$$\rightsquigarrow \quad p_{jj}^{(2)} = \underbrace{f_{jj}^{(1)} p_{jj}^{(1)}}_{jjj} + \underbrace{f_{jj}^{(2)} p_{jj}^{(0)}}_{j \times j}$$

Then, we can compute $f_{jj}^{(2)}$,

$$\rightsquigarrow \quad f_{jj}^{(2)} = p_{jj}^{(2)} - f_{jj}^{(1)} p_{jj}^{(1)}$$

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

In a similar manner, we can write an expression for $p_{jj}^{(3)}$

- Probability of state j, three steps after leaving

The three ways that this may occur

This occurs if the first return to j is after one step and in the two next steps the process may have been elsewhere but has returned to state j after that

$$j \rightarrow j \rightarrow \times \rightarrow j$$

Or, this occurs if the first return to state j is two steps after leaving it

$$j \rightarrow \times \rightarrow j \rightarrow j$$

Or, this occurs if the first return to state j is three steps after leaving it

$$j \rightarrow \times \rightarrow \times \rightarrow j$$

Classification of states (cont.)

Again, by combining these possibilities,

$$p_{jj}^{(3)} = \underbrace{f_{jj}^{(1)} p_{jj}^{(2)}}_{jj \times j} + \underbrace{f_{jj}^{(2)} p_{jj}^{(1)}}_{j \times jj} + \underbrace{f_{jj}^{(3)} p_{jj}^{(0)}}_{j \times \times j}$$

We can then compute $f_{jj}^{(3)}$,

$$\rightsquigarrow \quad f_{jj}^{(3)} = p_{jj}^{(3)} - f_{jj}^{(1)} p_{jj}^{(2)} - f_{jj}^{(2)} p_{jj}^{(1)}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

Summarising,

$$\begin{aligned} p_{jj}^{(1)} &= f_{jj}^{(1)} p_{jj}^{(0)} \\ p_{jj}^{(2)} &= f_{jj}^{(1)} p_{jj}^{(1)} + f_{jj}^{(2)} p_{jj}^{(0)} \\ p_{jj}^{(3)} &= f_{jj}^{(1)} p_{jj}^{(2)} + f_{jj}^{(2)} p_{jj}^{(1)} + f_{jj}^{(3)} p_{jj}^{(0)} \end{aligned}$$

We can continue by applying the laws of probability and using $p_{jj}^{(0)} = 1$

We get,

$$\rightsquigarrow \quad p_{jj}^{(n)} = \sum_{l=1}^{n} f_{jj}^{(l)} p_{jj}^{(n-l)}, \quad (\text{for } n \geq 1) \tag{7}$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

Similarly,

$$\begin{aligned} f_{jj}^{(1)} &= p_{jj}^{(1)} \\ f_{jj}^{(2)} &= p_{jj}^{(2)} - f_{jj}^{(1)} p_{jj}^{(1)} \\ f_{jj}^{(3)} &= p_{jj}^{(3)} - f_{jj}^{(1)} p_{jj}^{(2)} - f_{jj}^{(2)} p_{jj}^{(1)} \end{aligned}$$

Hence, $f_{jj}^{(n)}$ can be recursively computed for $n \geq 1$

$$\rightsquigarrow \quad f_{jj}^{(n)} = p_{jj}^{(n)} - \sum_{l=1}^{n-1} f_{jj}^{(l)} p_{jj}^{(n-l)}, \quad (\text{for } n \geq 1)$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

Consider the probability (denoted f_{jj}) of ever returning to state j

$$f_{jj} = \sum_{n=1}^{\infty} f_{jj}^{(n)}$$

If $f_{jj} = 1$, then we say that state j is a **recurrent** state

State j is recurrent IFF, starting in j, the probability of returning to j is 1

$\rightsquigarrow$ (The process is guaranteed to return to j)

In this case, we must have that $p_{jj}^{(n)} > 0$, for some $n > 0$

- The process returns to j infinitely often

Classification of states (cont.)

When $f_{jj} = 1$, we can define the **mean recurrence time** M_{jj} of state j

$$M_{jj} = \sum_{n=1}^{\infty} n f_{jj}^{(n)}$$

The expected number of steps till first-return to state j after leaving

- A recurrent state j for which M_{jj} is finite is **positive recurrent**

If $M_{jj} = \infty$, we say that state j is **null recurrent**

Classification of states (cont.)

Consider the probability f_{jj} of ever returning to state j

$$f_{jj} = \sum_{n=1}^{\infty} f_{jj}^{(n)}$$

If $f_{jj} < 1$, there is a non-zero probability the process will never return to j

⇝ We say that the state j is a **transient** state

Each time the chain is in state j, the probability it will never return is $1 - f_{jj}$

Classification of states (cont.)

Theorem

Consider a finite Markov chain

We have that

1. *No state is null recurrent ($M_{ii} \neq \infty$, for all i)*
2. *At least one state must be positive recurrent*

⇝ *(Not all states can be transient)*

⇝ *($f_{ii} \neq 1$ for some i)*

Suppose that all states are transient ($f_{ii} < 1$, for all i)

The process would spend some finite amount of time in each of them

- After that time, the process would have nowhere to go

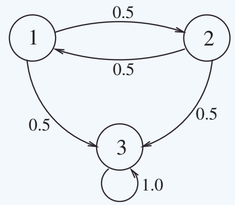
This is impossible, there must be at least one positive-recurrent state

■

Classification of states (cont.)

Example

Consider the following discrete-time Markov chain



$$P = \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 0 & 1 \end{pmatrix}$$

We are interested in the probability of first-return to state j after leaving it

$$f_{jj}^{(n)} = p_{jj}^{(n)} - \sum_{l=1}^{n-1} f_{jj}^{(l)} p_{jj}^{(n-l)}, \ \text{(for } n \geq 1)$$

Classification of states (cont.)

The sequence of powers of P

$$P^k = \begin{cases} \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \begin{pmatrix} 0 & (1/2)^k & 1-(1/2)^k \\ (1/2)^k & 0 & 1-(1/2)^k \\ 0 & 0 & 1 \end{pmatrix}, & \text{if } k = 1,3,5,\dots \\ \begin{matrix} 1 \\ 2 \\ 3 \end{matrix} \begin{pmatrix} (1/2)^k & 0 & 1-(1/2)^k \\ 0 & (1/2)^k & 1-(1/2)^k \\ 0 & 0 & 1 \end{pmatrix}, & \text{if } k = 2,4,6,\dots \end{cases}$$

Classification of states (cont.)

$$f_{jj}^{(n)} = p_{jj}^{(n)} - \sum_{l=1}^{n-1} f_{jj}^{(l)} p_{jj}^{(n-l)}, \text{ (for } n \geq 1)$$

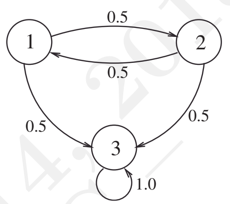
For state $j = 1$, we have

$$\begin{aligned} f_{11}^{(1)} &= p_{11}^{(1)} = 0 \\ f_{11}^{(2)} &= p_{11}^{(2)} - f_{11}^{(1)} p_{11}^{(1)} \\ &= (1/2)^2 - 0 = (1/2)^2 \\ f_{11}^{(3)} &= p_{11}^{(3)} - f_{11}^{(2)} p_{11}^{(1)} - f_{11}^{(1)} p_{11}^{(2)} \\ &= 0 - (1/2)^2 \cdot 0 - 0 \cdot (1/2)^2 = 0 \\ f_{11}^{(4)} &= p_{11}^{(4)} - f_{11}^{(3)} p_{11}^{(1)} - f_{11}^{(2)} p_{11}^{(2)} - f_{11}^{(1)} p_{11}^{(3)} \\ &= (1/2)^4 - 0 - (1/2)^2 \cdot (1/2)^2 - 0 = 0 \end{aligned}$$

In general, we get

$$\rightsquigarrow \quad f_{11}^{(k)} = 0, \text{ (for all } k \geq 3)$$

Classification of states (cont.)



The first return to state 1 must occur after 2 steps (or never again)

State 1 must therefore be a transient state (not recurrent, $f_{11} \neq 1$)

$$f_{11} = \sum_{n=1}^{\infty} f_{11}^{(n)} = 0 + 1/4 + 0 + \cdots \neq 1$$

A similar result applies to state 2 (transient, not recurrent, $f_{22} \neq 1$)

Classification of states (cont.)

The Markov chain has three states, of which two are transient (1 and 2)

$\rightsquigarrow$ The third state (3) must be positive recurrent ($M_{33} \neq \infty$)

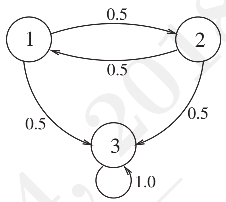
More explicitly,

$$\begin{aligned} f_{33}^{(1)} &= p_{33}^{(1)} = 1 \\ f_{33}^{(2)} &= p_{33}^{(2)} - f_{33}^{(1)} p_{33}^{(1)} = 1 - 1 = 0 \\ f_{33}^{(3)} &= p_{33}^{(3)} - f_{33}^{(1)} p_{33}^{(2)} - f_{33}^{(2)} p_{33}^{(2)} = 1 - 1 - 0 = 0 \end{aligned}$$

In general, we get

$$\rightsquigarrow \quad f_{33}^{(k)} = 0, \text{ (for all } k \geq 2)$$

Classification of states (cont.)



State 3 is therefore a recurrent state ($f_{33} = 1$)

$$f_{33} = \sum_{n=1}^{\infty} f_{33}^{(n)} = 1 + 0 + 0 + \cdots = 1$$

To see that, consider the mean recurrence time of state 3

$$M_{33} = \sum_{n=1}^{\infty} n f_{33}^{(n)} = 1 \cdot 1 + 2 \cdot 0 + 3 \cdot 0 + \cdots = 1$$

M_{33} is finite, 3 is a positive-recurrent state

■

Classification of states (cont.)

We considered only transitions from any state back again to that same state

Now, we consider also transitions between two different states

Classification of states (cont.)

Let $f_{ij}^{(n)}$ for $i \neq j$ be the **first-passage** probability to state j in n steps

- Conditioned on the fact that we started from state i

We have that $f_{ij}^{(1)} = p_{ij}$

We derive a recursive expression to compute $p_{ij}^{(n)}$

$$\rightsquigarrow \quad p_{ij}^{(n)} = \sum_{l=1}^{n} f_{ij}^{(l)} p_{jj}^{(n-l)}, \quad (\text{for } n \geq 1)$$

After rearranging the terms, we get an expression to compute $f_{ij}^{(n)}$

$$\rightsquigarrow \quad f_{ij}^{(n)} = p_{ij}^{(n)} - \sum_{l=1}^{n-1} f_{ij}^{(l)} p_{jj}^{(n-l)}$$

Again, this recursive relation is a more convenient

Classification of states (cont.)

Let f_{ij} be the probability that state j is ever visited from state i

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$$

If $f_{ij} < 1$, a process starting from state i might never visit state j

If $f_{ij} = 1$, the expected value of sequence $f_{ij}^{(n)}, n = 1, 2, \ldots$ of first-passage probabilities for i, j ($j \neq i$) is the **mean first-passage time** M_{ij} of state j

$$M_{ij} = \sum_{n=1}^{\infty} n f_{ij}^{(n)}, \quad (\text{for } i \neq j)$$

The expected number of steps to first-passage to state j after leaving i

Classification of states (cont.)

The M_{ij} uniquely satisfies the equation

$$M_{ij} = p_{ij} + \sum_{k \neq j} p_{ik}(1 + M_{kj}) = 1 + \sum_{k \neq j} p_{jk} M_{kj} \tag{8}$$

Consider a process that is in some state i

The chain can either go to state j in one step (probability p_{ij}), or go to some intermediate state k in one step (probability p_{ik}), then onto state j

- This will require an extra (expected) M_{kj} steps

(Let $i = j$, then M_{ij} corresponds to the mean recurrence time of state j)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

$$M_{ij} = p_{ij} + \sum_{k \neq j} p_{ik}(1 + M_{kj}) = 1 + \sum_{k \neq j} p_{jk} M_{kj}$$

We have,

$$\rightsquigarrow \quad M_{ij} = 1 + \sum_{k \neq j} p_{ik}(1 + M_{kj}) = 1 + \sum_{k \neq j} p_{jk} M_{kj} - p_{ij} M_{jj}$$

Let e be a column vector whose components are all ones

Let E be a square matrix whose elements are all ones

$$\rightsquigarrow \quad E = ee^T$$

Let $\text{diag}(M)$ be a diagonal matrix whose i-th column is M_{ii}

Thus, in matrix form we obtain

$$\rightsquigarrow \quad M = E + P\big[M - \text{diag}(M)\big]$$

MAY 14, 2018 — FC

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

$$M = E + P\big[M - \text{diag}(M)\big]$$

- ⇝ The diagonal elements of M are mean recurrence (first-return) times
- ⇝ The off-diagonal entries are expected first-passage times

Matrix M can be built iteratively, starting from $M^{(0)} = E$

$$M^{(k+1)} = E + P\big[M^{(k)} - \text{diag}(M^{(k)})\big] \tag{9}$$

MAY 14, 2018 — FC

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Classification of states (cont.)

Matrix F whose elements are f_{ij} is referred to as **reacheability matrix**

- ⇝ Probabilities to ever visit state j after leaving state i
- ⇝ (Alternative ways of calculating probabilities f_{ij})
- ⇝ We will examine this matrix later on

We also study how to get the expected number of visits to j on leaving i

MAY 14, 2018 — FC

Classification of states (cont.)

Periodicity

A state j is said **periodic** or **cyclic**, with **period** p, if on leaving state j a return is only possible in a number of step that is multiple of integer $p > 1$

⇝ The period of state j is therefore defined as the largest common divisor of the set of integers n for which we have that $p_{jj}^{(n)} > 0$

A state for which $p = 1$ is **aperiodic**

Classification of states (cont.)

A state that is positive-recurrent and aperiodic is said to be **ergodic**

⇝ If all states are ergodic, the chain is ergodic

Classification of states (cont.)

Some limit results on the behaviour of P^n as $n \rightarrow \infty$

Classification of states (cont.)

Theorem

Consider a homogeneous discrete-time Markov chain

Let state j be a null-recurrent[1] or transient[2] state and let i be any state

We have,

$$\rightsquigarrow \quad \lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$$

Let j be a positive-recurrent[3] and aperiodic (i.e., ergodic) state

We have,

$$\rightsquigarrow \quad \lim_{n \rightarrow \infty} p_{jj}^{(n)} > 0$$

Let j be a positive-recurrent and aperiodic (i.e., ergodic) state and let i be a any positive-recurrent, transient, or otherwise state

We have

$$\rightsquigarrow \quad \lim_{n \rightarrow \infty} p_{ij}^{(n)} = f_{ij} \lim_{n \rightarrow \infty} p_{jj}^{(n)}$$

[1] Recurrent, $f_{jj} = 1$, with mean recurrence time $M_{jj} = \infty$.
[2] Non-zero probability of never returning, $f_{jj} < 1$.
[3] Recurrent, $f_{jj} = 1$, with finite mean recurrence time $M_{jj} \neq \infty$.

Classification of states (cont.)

Example

Consider the Markov chain with transition probability P and $\lim_{n\to\infty} P^n$

$$P = \begin{pmatrix} 0.4 & 0.5 & 0.1 & 0 \\ 0.3 & 0.7 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0.8 & 0.2 \end{pmatrix}, \quad \lim_{n\to\infty} P^n = \begin{pmatrix} 0 & 0 & 4/9 & 5/9 \\ 0 & 0 & 4/9 & 5/9 \\ 0 & 0 & 4/9 & 5/9 \\ 0 & 0 & 4/9 & 5/9 \end{pmatrix}$$

The process has two transient states 1 and 2 and two ergodic states 3 and 4

State 1 and 2 are transient

$$\rightsquigarrow \quad \lim_{n\to\infty} p_{ij}^{(n)} = 0 \text{ (for } i = 1,2,3,4 \text{ and } j = 1,2)$$

States 3 and 4 are ergodic

$$\rightsquigarrow \quad \lim_{n\to\infty} p_{jj}^{(n)} > 0 \text{ (for } j = 3,2)$$

As $f_{ij} = 1$, for $i = 1,2,3,4$, $j = 3,4$ and $i \neq j$, we have

$$\rightsquigarrow \quad \lim_{n\to\infty} p_{ij}^{(n)} = f_{ij} \lim_{n\to\infty} p_{jj}^{(n)} = \lim_{n\to\infty} p_{jj}^{(n)} > 0$$

Irreducibility

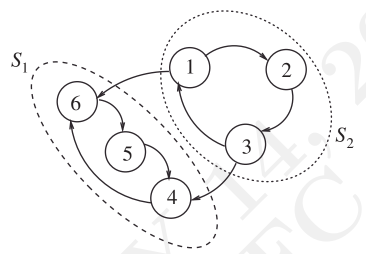
Discrete-time Markov chains

Irreducibility

We discuss the classification of groups of states

Irreducibility (cont.)

Let S be the complete set of states in a discrete-time Markov chain



Two subsets partition S

- S_1 and S_2

No one-step transition from any state in S_1 to any state in S_2

$\rightsquigarrow$ $S_1 = \{4,5,6\}$ is said to be **closed**

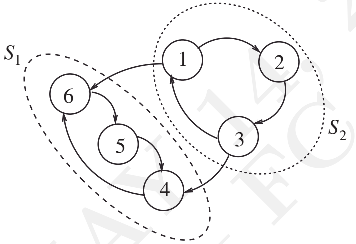
There exist transitions from states in S_2 to states in S_1

$\rightsquigarrow$ $S_2 = \{1,2,3\}$ is said to be **open**

(The set of all subsets, $S = \{1,2,3,4,5,6\}$, is closed)

Irreducibility (cont.)

S_1 is said to be closed if no state in S_1 leads to any state outside S_1



$$p_{ij}^{(n)} = 0, \quad (i \in S_1, j \notin S_1)$$

This must be true for any number of steps ($n \geq 1$)

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

These concepts can be generalised to any non-empty subset S_1 of S

Let the closed subset S_1 consists of a single state

⇝ Any such state is an absorbing state

Any finite set of transient states must define an open set

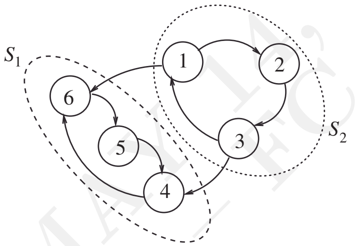
⇝ Any non-absorbing state is an open set

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Suppose that S does not contain any proper subset that is closed

- The Markov chain is said to be **irreducible**
- (There are no absorbing states)



This chain is not irreducible

- $S_1 = \{4, 5, 6\}$ is closed

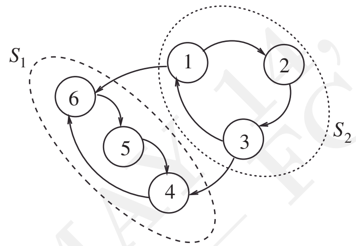
If S contains proper subsets that are closed, then the chain is **reducible**

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Consider a closed subset of states with no proper subset that is closed

⇝ This is called an **irreducible subset**



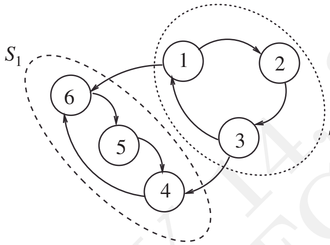
$S_1 = \{4, 5, 6\}$ is irreducible

- No closed sub-subsets

S_1, only irreducible subset

Any proper subset of an irreducible subset makes an open set of states

Irreducibility (cont.)



Matrix of transition probabilities

$$P = \begin{array}{c} 1\\2\\3\\4\\5\\6 \end{array}\begin{pmatrix} 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \end{pmatrix} = \begin{pmatrix} D_{11} & U_{12} \\ L_{21} & D_{22} \end{pmatrix}$$

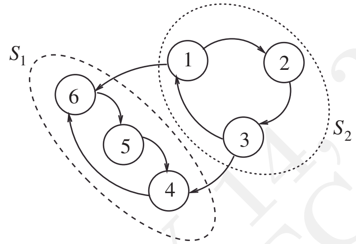
$*$ indicates non-zero probabilities

The matrix is decomposed according to the partition $\{1,2,3\}$ and $\{4,5,6\}$

- Two off-diagonal blocks U_{12} and L_{21}
- Two diagonal blocks D_{11} and D_{22}

Each block has size (3×3)

Irreducibility (cont.)



Matrix of transition probabilities

$$P = \begin{array}{c} 1\\2\\3\\4\\5\\6 \end{array}\begin{pmatrix} 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \end{pmatrix} = \begin{pmatrix} D_{11} & U_{12} \\ L_{21} & D_{22} \end{pmatrix}$$

The lower off-diagonal block L_{21} is identically equal to zero

- No transitions from states in D_{22} to states in D_{11}

The upper off-diagonal block U_{12} contains non-zero elements

- Transitions from states in D_{11} to states in D_{22}

Irreducibility (cont.)

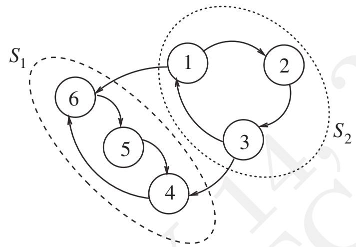
We consider an alternative definition of irreducibility of a Markov chain

- We define it in terms of **reachability** of the states
- (We defined already the reachability matrix)

State j is **reachable**/**accessible** from state i if there is a i to j path

- The path need not be unit-length, we write $i \rightarrow j$

Irreducibility (cont.)



Matrix of transition probabilities

$$P = \begin{array}{c} 1\\2\\3\\4\\5\\6 \end{array}\begin{pmatrix} 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \end{pmatrix} = \begin{pmatrix} D_{11} & U_{12} \\ L_{21} & D_{22} \end{pmatrix}$$

There is a path from state 3 to state 5, through states 4 and 6

⇝ State 5 is reachable from state 3

⇝ Path probability, $p_{34}p_{46}p_{65} > 0$

There exists no path from state 5 to state 3

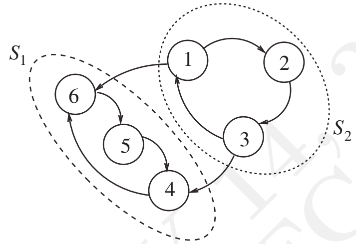
⇝ State 3 is not reachable from state 5

Irreducibility (cont.)

A chain is **irreducible** if every state is reachable from every other state

- That is, there exists an integer n for which $p_{ij}^{(n)} > 0$
- This must be true for every pair of states i and j

Irreducibility (cont.)



Matrix of transition probabilities

$$P = \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{pmatrix} 0 & * & 0 & 0 & 0 & * \\ 0 & 0 & * & 0 & 0 & 0 \\ * & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & * \\ 0 & 0 & 0 & * & 0 & 0 \\ 0 & 0 & 0 & 0 & * & 0 \end{pmatrix} = \begin{pmatrix} D_{11} & U_{12} \\ L_{21} & D_{22} \end{pmatrix}$$

It is not possible for any of the states 4, 5 and 6 to reach states 1, 2 and 3

- (Though the converse is possible)

We could make this Markov chain irreducible by adding a single path

- From any state of $\{4, 5, 6\}$ to whatever state of $\{1, 2, 3\}$

Irreducibility (cont.)

Consider the case of a state j that is reachable from state i

$$\rightsquigarrow \quad i \rightarrow j$$

Consider the case of a state i that is reachable from state j

$$\rightsquigarrow \quad j \rightarrow i$$

States i and j are **communicating states**

$$\rightsquigarrow \quad i \longleftrightarrow j$$

The communication property sets an equivalence relationship

- ⇝ Symmetry
- ⇝ Reflexivity
- ⇝ Transitivity

Irreducibility (cont.)

Consider any three states i, j and k

$$i \longleftrightarrow j \Longrightarrow j \longleftrightarrow i$$
$$i \longleftrightarrow j \text{ and } j \longleftrightarrow i \Longrightarrow i \longleftrightarrow i$$
$$i \longleftrightarrow j \text{ and } j \longleftrightarrow k \Longrightarrow i \longleftrightarrow k$$

⇝ The first relation holds by definition

⇝ The third relation holds because $i \longleftrightarrow j$ implies $i \rightarrow j$

- There is a $n_1 > 0$ for which $p_{ij}^{(n_1)} > 0$

Similarly, we also have that $j \longleftrightarrow k$ implies $j \rightarrow k$

- There is a $n_2 > 0$ for which $p_{jk}^{(n_2)} > 0$

Set $n = n_1 + n_2$, then use Chapman-Kolmogorov equation to get $i \rightarrow k$

$$\rightsquigarrow \quad p_{ik}^{(n)} = \sum_{\text{all } l} p_{il}^{(n_1)} p_{lk}^{(n_2)} \geq p_{ij}^{(n_1)} p_{jk}^{(n_2)} > 0$$

It may be similarly shown that $k \rightarrow i$

⇝ The second relation follows from the third one by transitivity

$$i \longleftrightarrow j \text{ and } j \longleftrightarrow i \Longrightarrow i \longleftrightarrow i$$

Irreducibility (cont.)

A state that communicates with itself in this way is a **return state**

A **non-return state** is one that does not communicate with itself

- Once gone, the Markov chain never returns to that state

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Consider the set of all states that communicate with state i

- ⇝ The set is a called a **communicating class**, $C(i)$
- ⇝ (The communicating class can be an empty set)

It is possible that a state communicates with no other state

- ⇝ This is the case of the ephemeral states
- ⇝ They can only be occupied initially

On the other hand, any state i for which $p_{ii}^{(n)} > 0$ is a return state

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

This suggests an alternative partitioning of the states of a Markov chain

- ⇝ Communicating classes
- ⇝ Non-return states

Moreover, we have that communicating classes may or may not be closed

- ⇝ A recurrent state belongs to a closed communicating class
- ⇝ Only transient states can be members of non-closed classes

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

If state i is recurrent and $i \rightarrow j$ then state j must communicate with state i

$$\rightsquigarrow \quad i \longleftrightarrow j$$

There is a path from i to j and, since i is recurrent, after leaving j we must eventually return to i, which implies that there is also a path from j to i

- In this case, j must be also recurrent

We return to i infinitely often and j is reachable from i

- ⇝ Then, we can also return to j infinitely often

Irreducibility (cont.)

Since i and j communicate, for some $n_1, n_2 > 0$ we have

$$p_{ij}^{(n_1)} > 0$$
$$p_{ji}^{(n_2)} > 0$$

Since i is recurrent, for some integer $n > 0$ we have

$$\rightsquigarrow \quad p_{jj}^{(n_2+n+n_1)} \geq p_{ji}^{(n_2)} p_{ii}^{(n)} pij^{(n_1)} > 0$$

State i is recurrent ($\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$) then j is recurrent ($\sum_{m=1}^{\infty} p_{jj}^{(m)} = \infty$)

This is because we have

$$\sum_{n=1}^{\infty} p_{ji}^{(n_2)} p_{ii}^{(n)} p_{ij}^{(n_1)} = p_{ji}^{(n_2)} p_{ij}^{(n_1)} \sum_{n=1}^{(\infty)} p_{ii}^{(n)} = \infty$$

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Recurrent states can only reach other recurrent states

- $\rightsquigarrow$ Transient states cannot be reached from them
- $\rightsquigarrow$ The set of recurrent states must be closed

If state i is a recurrent state, then $C(i)$ is an irreducible closed set

- And, it contains only recurrent states

All states must be positive recurrent or they must be null recurrent

Consider a chain whose states are in the same communicating class

- We say that that Markov chain is irreducible

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Some theorems concerning irreducible discrete-time Markov chains

Discrete-time Markov chains
UFC/DC
SA (CK0191)
2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Theorem

Consider an irreducible discrete-time Markov chain

The process is positive-recurrent or null-recurrent or it is transient

That is,

- $\rightsquigarrow$ *All states are positive-recurrent, or*
- $\rightsquigarrow$ *All states are null-recurrent, or*
- $\rightsquigarrow$ *All states are transient*

Moreover, all states are periodic, with same period p

- $\rightsquigarrow$ *Or, else, they are aperiodic*

■

Irreducibility (cont.)

Theorem

In a finite, irreducible Markov chain, all states are positive-recurrent

In finite Markov chain, no states are null-recurrent

⇝ At least one state must be positive recurrent

Adding irreducibility means all states must be positive recurrent ■

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

Theorem

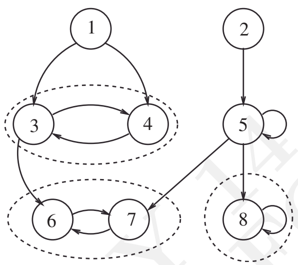
The states of an aperiodic, finite, irreducible Markov chain are ergodic

The conditions are only sufficient conditions

⇝ (Not a definition of ergodicity) ■

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)



State 1 and 2 are non-return states

State 3 and 4, communicating class

- Not closed

State 5, a communicating class

- It is a return state
- Non-closed

(Non-return if without self-loop)

State 7 and 6 together form a closed communicating class

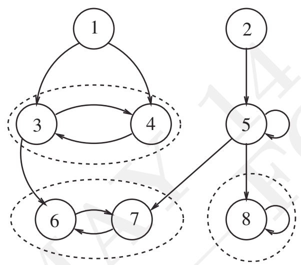
State 8 is an return state, a closed communicating class

- An absorbing state

Discrete-time Markov chains
UFC/DC SA (CK0191) 2018.1
Generalities
Chapman-Kolmogorov
Classification of states
Irreducibility

Irreducibility (cont.)

We can partition the state-space of the Markov chain



We consider two main subsets

1. Transient states

2. Recurrent states

- Irreducible
- Closed communicating classes

Irreducibility (cont.)

It is possible to bring the transition matrix to a **normal form**

- (Potentially, re-ordering of the states is needed)

$$\rightsquigarrow \quad P = \begin{pmatrix} T_{11} & T_{12} & T_{13} & \cdots & T_{1N} \\ 0 & R_2 & 0 & \cdots & 0 \\ 0 & 0 & R_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R_N \end{pmatrix}$$

$\rightsquigarrow$ $(N-1)$ closed communicating classes R_k $(k = 2, \dots, N)$
$\rightsquigarrow$ A set containing all transient states T

The set of transient states may contain multiple communicating classes

Suppose that some transient state T_{1k} is not identically null, transitions from at least one of the transient states into the closed set R_k are possible

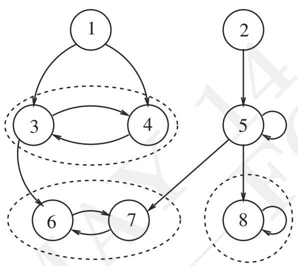
Irreducibility (cont.)

$$\rightsquigarrow \quad P = \begin{pmatrix} T_{11} & T_{12} & T_{13} & \cdots & T_{1N} \\ 0 & R_2 & 0 & \cdots & 0 \\ 0 & 0 & R_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & R_N \end{pmatrix}$$

Each of the R_k may be considered as a irreducible Markov chain
$\rightsquigarrow$ Many properties of R_k are independent of other states

Irreducibility (cont.)

The transition matrix in partitioned form and the transition diagram



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

$$= \begin{pmatrix} T_{11} & T_{12} & T_{13} \\ 0 & R_2 & 0 \\ 0 & 0 & R_3 \end{pmatrix}$$