

Notes on Stochastic Network Modeling (SNM)

Llorenç Cerdà-Alabern

llorenc@ac.upc.edu

Barcelona, November 13, 2020

Contents

I Introduction 1

1 Probability Review 1

1.1	Ingredients of Probability	1
1.2	Expected value	1
1.3	Variance	2
1.4	Indicator Function	2
1.5	Expected value of non negative RVs	2
1.6	Wald's Equation	2
1.7	Probability in \mathbb{R}^k	2

2 Stochastic Process (SP) 3

2.1	Introduction	3
2.2	Analysis of Stochastic Processes	4

II Discrete Time Markov Chains 4

3 Definition of a DTMC 4

3.1	State Transition Diagram	4
3.2	Properties of a DTMC	4
3.3	Transition Matrix	4
3.4	Absorbing Chains	4
3.5	Sojourn or Holding Time	5
3.6	n-step transition probabilities	5
3.7	State Probabilities	5
3.8	Chapman-Kolmogorov Equations	5

4 Transient Solution 6

4.1	Close Form Solution	6
4.2	Eigenvalues	6
4.3	Determinant	6
4.4	Eigenvalues of a Stochastic Matrix	6
4.5	Chain with a Defective Matrix	7

5 Classification of States 7

5.1	Irreducibility	7
5.2	Transient and Recurrent	8
5.3	First Passage (Transition) Probabilities	8
5.4	Mean Recurrence Time	8
5.5	Property of States	8
5.6	Recursive Equation for the First Passage Probabilities	9
5.7	Recursive Equation for the Mean Recurrence Time	9
5.8	Periodic states	9

6 Steady State 10

6.1	Limiting Distribution	10
6.2	Stationary distribution	10
6.3	Numerical Solution	11
6.4	Global balance equations	11
6.5	Ergodic Chains	11
6.5.1	Theorems for ergodic chains	12

7 Reversed Chain 12

7.1	Computation of p_{ij}^r	12
-----	---------------------------	----

8 Reversible Chains 13

8.1	Kolmogorov Criteria	13
8.2	Product Form Solution	13
8.3	Birth and Death Chains	14

9 Research Example: Aloha 14

9.1	Analysis with finite population	14
9.1.1	Stationary distribution	15
9.2	Throughput	15
9.2.1	Dynamics	15
9.2.2	Stabilizing Aloha	15

10 Finite Absorbing Chains 15

10.1	Results	16
10.2	Extension of the Results	16
10.3	Inverse of a matrix	16

III Continous Time Markov Chains 17

11 Definition of a CTMC 17

11.1	Properties of a continuous time MC	17
11.2	Transition Matrix	17
11.3	State Transition Diagram	17
11.4	Sojourn Time	17
11.5	Exponential Jumps Description of a CTMC	18

12 Transient Solution 18

12.1	Chapman-Kolmogorov Equations	18
12.2	State Probabilities	19
12.3	Transient Solution	19
12.3.1	Eigenvalues of an Infinitesimal Generator	19
12.3.2	Chain with a Defective Matrix	20

13 Embedded MC of a CTMC 20

13.1	Definition	20
------	------------	----

14 Classification of States 20

14.1	Irreducibly	20
14.2	Transient and Recurrent	21
14.3	Mean recurrence time of the CTMC	21

15 Steady State 21

15.1	Limiting Distribution	21
15.2	Stationary Distribution	21
15.3	Numerical Solution	21
15.4	Global balance equations	22
15.4.1	Solving using flux balancing	22
15.5	Ergodic Chains	22
15.6	Reversible Chains	22
15.6.1	Example	23

16 Semi-Markov Process 23

16.1 Embedded MC (EMC) of a semi-Markov process . 23

16.1.1 Example 23

Solution 24

16.2 Embedded MC of a CTMC 24

17 Finite Absorbing Chains 24

17.1 Canonical Form 24

17.2 Results 24

17.3 Extension of the Results 25

Preface

These notes were prepared for the subject *Stochastic Network Modeling (SNM)* from the *Master in Innovation and Research in Informatics (MIRI)*, Universitat Politècnica de Catalunya.

Further reading: The book of Mor [3], Nelson [6] or Trivedi [7] are excellent books on Markov Chains and cover most of the course. The book of Kemeny [4] covers absorbing DTMC.

Part I

Introduction

Chapter 1

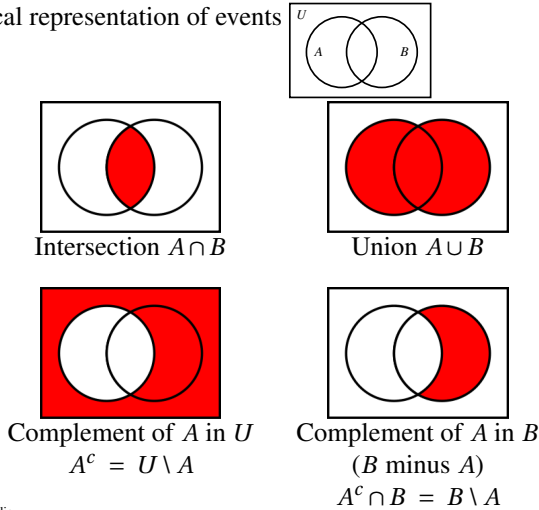
Probability Review

1.1 Ingredients of Probability

- **Random experiment**, e.g. toss a die.
- **Outcome**, ω , e.g. tossing a die can be $\omega = 2$, choosing a fruit can be $\omega = \text{orange}$.
- **Sample space or Universal set**, U , set of all possible outcomes. E.g. tossing a die $U = \{1, 2, 3, 4, 5, 6\}$.
- **Event**, A , any subset of U (e.g. tossing a die $A = \{1, 2, 3\}$). We say the event A occurs if the outcome of the experiment $\omega \in A$. U is the **sure event**, and we represent by the empty set \emptyset an **impossible outcome**.

Venn Diagrams

Graphical representation of events



source: wikipedia

Random Variable

- For simplicity it is defined a **random variable (RV)**, X as a function that assigns a real number to each outcome in the sample space U , i.e.:

$$X : U \rightarrow \mathbb{R} \quad (1.1)$$

- We will represent the experiment by a RV, X , and the possible outcomes by its values. $X = x_i$ is the **outcome** $X(\omega_i) = x_i$.
- Using RVs the sample space is mapped in a subset of \mathbb{R} . So, in terms of X , U is a set of points of \mathbb{R} . The same for any event.
- Normally the definition of X comes naturally from the experiment, e.g. tossing a die: $X = \{\text{number in the toss}\}$.
- RVs can be **discrete** (e.g. tossing a die) or **continuous** (e.g. waiting time of a packet in a queue).

Probability Measure

- If the sample space U of the RV X is **finite (discrete RV)**, $U = \{x_1, \dots, x_n\}$, a **probability measure** is an assignment of numbers $P(x_i)$, referred to as **probabilities**, to each **outcome** x_i such that:

$$0 \leq P(x_i) \leq 1$$

$$P(A) = \sum_{x_i \in A} P(x_i) \quad (1.2)$$

$$P(U) = 1$$

E.g. tossing a fair die,

$$P(x_i) = 1/6$$

$$P(X \in \{2, 4, 6\}) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2} \quad (1.3)$$

- If the sample space of the RV X is continuous (**continuous RV**), the events are intervals of \mathbb{R} . The probability measure is defined by means of the **cumulative distribution function, CDF**:

$$F(x) = P(X \in (-\infty, x]) = P(X \leq x) \quad (1.4)$$

- X is called absolutely continuous¹ if there exists the **probability density function, PDF**, such that for any interval $I = \{x \mid a \leq x \leq b\}$:

$$\int_a^b f(x) dx = P(X \in I) = F(b) - F(a) \quad (1.5)$$

Conditional Probability and Bayes Formula

- Given the the sample space U and the **events** $A, B \in U$ with $P(B) > 0$ the **probability of A conditioned by B** is defined as:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad (1.6)$$

NOTE: It's common to use commas to denote set intersection, and write $P(A \cap B)$ as $P(A, B)$.

- **Bayes Formula**

$$P(A|B) P(B) = P(B|A) P(A) \Rightarrow P(A|B) = \frac{P(B|A) P(A)}{P(B)} \quad (1.7)$$

Law of total probability

- Let B_i a **partition** of the sample space U ($\cup_i B_i = U$, $B_i \cap B_j = \emptyset, \forall i \neq j$), then

$$P(A) = \sum_i P(A|B_i) P(B_i) \quad (1.8)$$

- For **conditional probabilities**:

$$P(A|C) = \sum_i P(A|C \cap B_i) P(B_i|C) \quad (1.9)$$

- If C is **independent** of any of the B_i

$$P(A|C) = \sum_i P(A|C \cap B_i) P(B_i) \quad (1.10)$$

1.2 Expected value

- Given the discrete $N \in \mathbb{Z}$, respectively continuous $X \in \mathbb{R}$ RV, the **expected value** is:

$$E[N] = \sum_{k=-\infty}^{\infty} k P(N = k)$$

$$E[X] = \int_{-\infty}^{\infty} x f(x) dx \quad (1.11)$$

¹Some special distributions, called singular, do not have a PDF. One example is the Cantor distribution (see Wikipedia).

Example A number $X_1 \in \{1, 2, \dots, 6\}$ is obtained tossing a dice. Then, a number $X_2 \in [0, \infty]$ is obtained exponentially distributed with parameter X_1 . Compute $f(x_1, x_2)$, $f(x_2)$ and $E[X_2]$.
Note: Exponential distribution with parameter α :

$$f(x) = \alpha e^{-\alpha x}, x \in [0, \infty], E[X] = \frac{1}{\alpha}. \quad (1.12)$$

Solution:

$$f(x_1, x_2) = f(x_2|x_1) P(x_1) = x_1 e^{-x_1 x_2} \frac{1}{6}, \begin{cases} x_1 \in \{1, 2, \dots, 6\} \\ x_2 \in [0, \infty] \end{cases}$$

$$f(x_2) = \sum_{x_1} f(x_2|x_1) P(x_1) = \frac{1}{6} \sum_{n=1}^6 n e^{-n x_2}, x_2 \in [0, \infty]$$

$$E[X_2] = \frac{1}{6} \sum_{n=1}^6 \int_{x_2=0}^{\infty} x_2 n e^{-n x_2} = \frac{1}{6} \sum_{n=1}^6 \frac{1}{n} = \frac{49}{120}$$

1.3 Variance

- The amount of dispersion of a RV X with expected value $\mu = E[X]$ is measured by the **Variance**:

$$\text{Var}(X) = \sigma^2 = E[(X - \mu)^2] = E[X^2] - \mu^2 \quad (1.13)$$

- Often it is used the **standard deviation** $\sigma = \sqrt{\text{Var}(X)}$.

1.4 Indicator Function

$$I(A) = \begin{cases} 1, & \text{if } A \text{ occurs,} \\ 0, & \text{otherwise.} \end{cases} \quad (1.14)$$

Therefore:

$$E[I(A)] = 0 \times P(I(A) = 0) + 1 \times P(I(A) = 1) = P(A) \quad (1.15)$$

1.5 Expected value of non negative RVs

- For **non negative** RVs, $N \geq 0$ discrete and $X \geq 0$ continuous:

$$E[N] = \sum_{k=0}^{\infty} k P(N = k) = \sum_{k=0}^{\infty} P(N > k)$$

$$E[X] = \int_0^{\infty} x f(x) dx = \int_0^{\infty} P(X > x) dx = \int_0^{\infty} (1 - F(x)) dx \quad (1.16)$$

$$N = \sum_{k=0}^{N-1} 1 = \sum_{k=0}^{\infty} I(N > k) \quad (1.17)$$

$$X = \int_0^X dx = \int_0^{\infty} I(X > x) dx$$

and take expectations.

1.6 Wald's Equation

- Definition:** An positive integer RV $N > 0$ is a **stopping time** of a sequence X_1, X_2, \dots if the event $N = n$ is independent of X_{n+1}, X_{n+2}, \dots .
E.g. toss a die until you get 6. Let N be the number of tosses. N does not depend on the values obtained after getting 6.
- Wald's Equation** If X_1, X_2, \dots are independent and identically distributed and N is a stopping time:

$$E \left[\sum_{n=1}^N X_n \right] = E[X] E[N] \quad (1.18)$$

Proof.

$$\begin{aligned} E \left[\sum_{n=1}^N X_n \right] &= E \left[\sum_{n=1}^{\infty} X_n I(n \leq N) \right] = \\ &= \sum_{n=1}^{\infty} E[X_n] E[I(n \leq N)] = \\ &= E[X] \sum_{n=1}^{\infty} P(n \leq N) = \\ &= E[X] \sum_{n=0}^{\infty} P(N > n) = E[X] E[N] \quad \square \end{aligned}$$

1.7 Probability in \mathbb{R}^k

If we have a set of k RV $\mathbf{X} = (X_1, \dots, X_k)$ taking values in \mathbb{R}^k ($\mathbf{X} \in \mathbb{R}^k$), we define the **joint distribution**:

- Discrete RV

$$P(\mathbf{x}) = P(x_1, \dots, x_k) = P(X_1 = x_1, \dots, X_k = x_k) \quad (1.19)$$

- Continuous RV:

- **cumulative distribution function, CDF:**

$$F(\mathbf{x}) = F(x_1, \dots, x_k) = P(X_1 \in (-\infty, x_1], \dots, X_k \in (-\infty, x_k]) \quad (1.20)$$

- with **joint density** function $f(\mathbf{x}) = f(x_1, \dots, x_k)$ (if exists):

$$\begin{aligned} F(\mathbf{x}) &= F(x_1, \dots, x_k) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_k} f(x_1, \dots, x_k) dx_k \dots dx_1 \\ f(\mathbf{x}) &= f(x_1, \dots, x_k) = \frac{\partial^k F(x_1, \dots, x_k)}{\partial x_1 \dots \partial x_k} \end{aligned} \quad (1.21)$$

Marginal distributions Let $\mathbf{X} = (X_1, X_2)$, where $\mathbf{X} \in \mathbb{R}^k, X_1 \in \mathbb{R}^r, X_2 \in \mathbb{R}^{k-r}, 1 \leq r < k$:

- Discrete RV

$$P(\mathbf{x}_2) = \sum_{x_1} \dots \sum_{x_r} P(\mathbf{x}_1, \mathbf{x}_2) \quad (1.22)$$

- Continuous RV

$$f(\mathbf{x}_2) = \int_{x_1} \dots \int_{x_r} f(\mathbf{x}_1, \mathbf{x}_2) dx_1 \dots dx_r \quad (1.23)$$

Independent RV

- Discrete RV

$$\begin{aligned} P(\mathbf{x}) &= P(x_1, \dots, x_k) = \\ &= P(X_1 = x_1, \dots, X_k = x_k) = \\ &= P(X_1 = x_1) \dots P(X_k = x_k) \end{aligned} \quad (1.24)$$

- Continuous RV

$$\begin{aligned} F(\mathbf{x}) &= F(x_1, \dots, x_k) = F_{X_1}(x_1) \dots F_{X_k}(x_k) \\ f(\mathbf{x}) &= f(x_1, \dots, x_k) = f_{X_1}(x_1) \dots f_{X_k}(x_k) \end{aligned} \quad (1.25)$$

Conditional Distribution

- Let $\mathbf{X} = (X_1, X_2)$, where $\mathbf{X} \in \mathbb{R}^k$, $X_1 \in \mathbb{R}^r$, $X_2 \in \mathbb{R}^{k-r}$, the r -dimensional distribution of X_1 **conditioned by** $X_2 = \mathbf{x}_2$, $P(\{X_2 = \mathbf{x}_2\}) > 0$ is:

$$P(X_1 | X_2) = P(X_1 \leq \mathbf{x}_1 | X_2 = \mathbf{x}_2) = \frac{P(X_1 \leq \mathbf{x}_1, X_2 = \mathbf{x}_2)}{P(X_2 = \mathbf{x}_2)}.$$

If \mathbf{X} is **discrete** with probability $P(\mathbf{x}_1, \mathbf{x}_2)$ or absolutely **continuous** with density $f(\mathbf{x}_1, \mathbf{x}_2)$:

$$\begin{aligned} P(\mathbf{x}_1 | \mathbf{x}_2) &= \frac{P(\mathbf{x}_1, \mathbf{x}_2)}{P(\mathbf{x}_2)} \\ f(\mathbf{x}_1 | \mathbf{x}_2) &= \frac{f(\mathbf{x}_1, \mathbf{x}_2)}{f(\mathbf{x}_2)} \end{aligned} \quad (1.26)$$

Composition of marginals and conditionals Using the previous formulas we can compute (\mathbf{X} can be a mixture of discrete and continuous RV):

Law of total probability

- If $\mathbf{x}_1, \mathbf{x}_2$ are **discrete** RV: $P(\mathbf{x}_2) = \sum_{\mathbf{x}_1} P(\mathbf{x}_2 | \mathbf{x}_1) P(\mathbf{x}_1)$
- If \mathbf{x}_1 is **discrete** and \mathbf{x}_2 is **cont.**: $f(\mathbf{x}_2) = \sum_{\mathbf{x}_1} f(\mathbf{x}_2 | \mathbf{x}_1) P(\mathbf{x}_1)$
- If $\mathbf{x}_1, \mathbf{x}_2$ are **cont.**: $f(\mathbf{x}_2) = \int_{\mathbf{x}_1} f(\mathbf{x}_2 | \mathbf{x}_1) f(\mathbf{x}_1) d\mathbf{x}_1$
- If \mathbf{x}_1 is **cont.** and \mathbf{x}_2 is **discrete**: $P(\mathbf{x}_2) = \int_{\mathbf{x}_1} P(\mathbf{x}_2 | \mathbf{x}_1) f(\mathbf{x}_1) d\mathbf{x}_1$

Conditional expected value

- Given $X \in \mathbb{R}$, $Y \in \mathbb{R}^k$ with density $f(x, y)$:

$$\begin{aligned} E[X | Y = \mathbf{y}] &= \int_{\mathbb{R}} x f(x | \mathbf{y}) dx \\ E[X] &= \int_{\mathbb{R}^k} E[X | Y = \mathbf{y}] f(\mathbf{y}) d\mathbf{y} \end{aligned} \quad (1.27)$$

where the **marginal** $f(\mathbf{y}) = \int_{x=-\infty}^{\infty} f(x, y) dx$ and the **conditional** $f(x | \mathbf{y}) = f(x, y) / f(\mathbf{y})$.

Thus, the law of total probability also applies to expected value, and it is known as **law of total expectation**.

Chapter 2

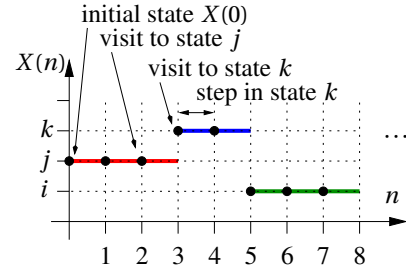
Stochastic Process (SP)

2.1 Introduction

- Sequence of RVs** $\{X(t)\}_{t \geq 0}$.
- $X(t)$ is the **state** at time t .
- The **state** $X(t)$ can be **continuous** or **discrete**.
- The **index** can be **continuous** or **discrete**. We shall use n for the **index**, and refer to it as **steps** when it is **discrete**, and t and refer to it as **time** when it is **continuous**.
- We call a possible sequence of states of the SP the **sample function** (or sample path) of the SP.

Sample Path

- Possible evolution (**sample path**) of a **discrete state, discrete time** SP $\{X(n)\}_{n \geq 0}$:



- To characterize the stochastic process we would need the distribution and **joint probabilities** of the $\{X(n)\}_{n \geq 0}$ RVs:

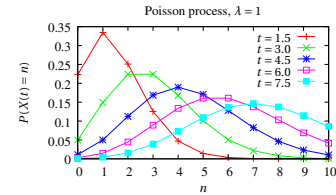
$$P(X(n) = i, X(n-1) = k, \dots, X(0) = j) \quad (2.1)$$

Example 2: Poisson Process

- It is a discrete state continuous time SP.
- It counts the number of events occurred in a time interval.
- Often used to build models of other stochastic processes.
- Definition:** The number of “events” in any interval of length t , $X(t)$, is **Poisson distributed** with mean λt , i.e.

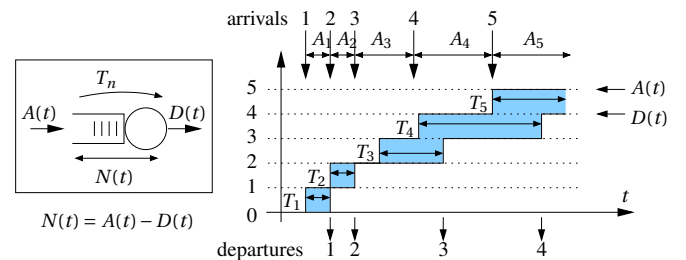
$$\begin{aligned} P(X(t+s) - X(s) = n) &= P(X(t) - X(0) = n) = \\ P(X(t) = n) &= e^{-\lambda t} \frac{(\lambda t)^n}{n!} \end{aligned} \quad (2.2)$$

where we assume $X(0) = 0$.



Example 3: Queue with Poisson Arrivals

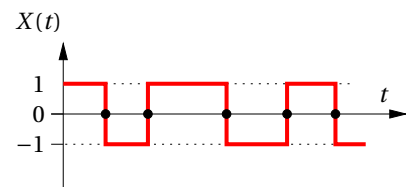
- The queue arrivals, $A(t)$, are modeled as a **Poisson process** with mean λt . Each event model an arrival.



Example 4: Telegraph signal

- The signal is modeled as a **Poisson process** with mean λt such that $X(0) = 1$ or $X(0) = -1$ with equal probability of $1/2$ and:

$$X(t) = \begin{cases} 1 & \text{if the number of events in } (0, t] \text{ is even} \\ -1 & \text{if the number of events in } (0, t] \text{ is odd} \end{cases} \quad (2.3)$$



2.2 Analysis of Stochastic Processes

- **Signal Theory:** Normally interested in the **spectral analysis** of the signal. The basic tool is the **Fourier transform** of the **auto-correlation function** of the process (**energy spectral density**). We will not do this analysis.

$$R(t) = E[X(\tau)X(\tau - t)] \quad \text{autocorrelation}$$

$$F(f) = \mathcal{F}[R(t)] = \int_{-\infty}^{\infty} R(t) e^{-j2\pi f t} dt \quad \text{Fourier transform (energy spectral density)}$$
(2.4)

- **Computer Networks:** Normally interested in probabilistic models using **Markov Chains** and **Queueing Theory**.

Part II

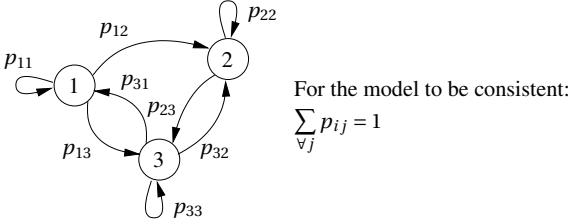
Discrete Time Markov Chains

Chapter 3

Definition of a DTMC

3.1 State Transition Diagram

- We are interested in a **process that evolve in stages**.
- For the model to be tractable, it is convenient to represent the SP by giving all **possible states** (there may be ∞), and the **possible transitions** between them:



- Mathematically:

$$p_{ij} = P(X(n) = j | X(n-1) = i) \quad (3.1)$$

3.2 Properties of a DTMC

- The event $X(n) = i$ (at step n the system is in state i) must satisfy (**memoryless property**):

$$P(X(n) = j | X(n-1) = i, X(n-2) = k, \dots) = P(X(n) = j | X(n-1) = i) \quad (3.2)$$

- If $P(X(n) = j | X(n-1) = i) = P(X(1) = j | X(0) = i)$ for any n we have an **homogeneous** DTMC. We shall only consider homogeneous DTMC.
- We call **one-step transition probabilities** to:

$$p_{ij} = P(X(n) = j | X(n-1) = i) \quad (3.3)$$

- The SP is called a Markov Process (MP) or Markov Chain (MC) depending on the state being continuous or discrete.

3.3 Transition Matrix

- Transition probabilities:

$$p_{ij} = P(X(n) = j | X(n-1) = i) \quad (3.4)$$

- In matrix form:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \dots \\ p_{21} & p_{22} & \dots \\ \dots & \dots & \dots \end{bmatrix} \quad (3.5)$$

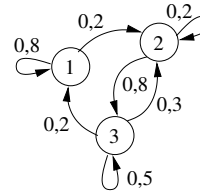
- For the model to be consistent, the probability to move from i to any state must be 1. Mathematically:

$$\sum_j p_{ij} = \sum_j P(X(n) = j | X(n-1) = i) = \sum_j \frac{P(X(n-1) = i | X(n) = j) P(X(n) = j)}{P(X(n-1) = i)} = \frac{P(X(n-1) = i)}{P(X(n-1) = i)} = 1 \quad (3.6)$$

- \mathbf{P} is a **stochastic matrix**, i.e. a matrix which rows sum 1.

Example

- Assume a terminal can be in **3 states**:
 - State 1: Idle.
 - State 2: Active without sending data.
 - State 3: Active and sending data at a rate v bps.



	to state			
	1	2	3	
$\mathbf{P} =$	0,8	0,2	0	from state
	0	0,2	0,8	
	0,2	0,3	0,5	
	1	2	3	

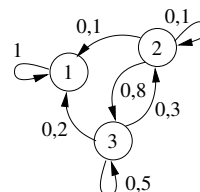
(3.7)

- The **average transmission rate** (throughput), v_a , is:

$$v_a = P(\text{the terminal is in state 3}) \times v \quad (3.8)$$

3.4 Absorbing Chains

- It is possible to have chains with **absorbing states**.
- A state i is absorbing if $p_{ii} = 1$.
- Example: State 1 is absorbing.

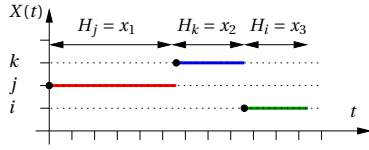


	to state			
	1	2	3	
$\mathbf{P} =$	1	0	0	from state
	0,1	0,1	0,8	
	0,2	0,3	0,5	
	1	2	3	

(3.9)

3.5 Sojourn or Holding Time

- **Sojourn or holding time** in state k : Is the RV H_k equal to the number of steps that the chain remains in state k before leaving to a different state:



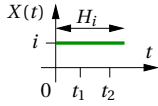
- The Markov property implies:

$$H_i(n) = P(H_i = n) = p_{ii}^{n-1} (1 - p_{ii}), n \geq 1 \quad (3.10)$$

- Which is a **geometric** distribution with mean:

$$E[H_i] = \sum_{n=1}^{\infty} n P(H_i = n) = \frac{1}{1 - p_{ii}}. \quad (3.11)$$

The geometric distribution satisfies the Markov property



Proof.

- Markov property:
 $P(X(n_2) = i | X(n_1) = i, X(n_0) = i) = P(X(n_2) = i | X(n_1) = i)$
- Thus, the Markov property **in terms of the sojourn time** can be written as:

$$P(H_i > n_2 - n_0 | H_i > n_1 - n_0) = P(H_i > n_2 - n_1) \quad (3.12)$$

- Since

$$P(H_i > k) = 1 - P(H_i \leq k) = 1 - \sum_{n=1}^k p^{n-1} (1 - p) = 1 - (1 - p) \frac{1 - p^k}{1 - p} = p^k \quad (3.13)$$

- We have:

$$\begin{aligned} P(H_i > n_2 - n_0 | H_i > n_1 - n_0) &= \frac{P(H_i > n_2 - n_0, H_i > n_1 - n_0)}{P(H_i > n_1 - n_0)} = \\ &= \frac{P(H_i > n_2 - n_0)}{P(H_i > n_1 - n_0)} = \frac{p^{n_2 - n_0}}{p^{n_1 - n_0}} = p^{n_2 - n_1} = P(H_i > n_2 - n_1) \end{aligned} \quad (3.14)$$

3.6 n-step transition probabilities

- Transition probabilities: $p_{ij} = P(X(n) = j | X(n-1) = i)$
- In matrix form:

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \quad (3.15)$$

- We define the **n-step** transition probabilities:

$$p_{ij}(n) = P(X(n) = j | X(0) = i) \quad (3.16)$$

$$\mathbf{P}(n) = \begin{bmatrix} p_{11}(n) & p_{12}(n) & \cdots \\ p_{21}(n) & p_{22}(n) & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix} \quad (3.17)$$

- \mathbf{P} and $\mathbf{P}(n)$ are **stochastic matrices**: Their rows sum 1.

3.7 State Probabilities

- Define the probability of being in state i at step n :

$$\pi_i(n) = P(X(n) = i) \quad (3.18)$$

- In vector form (row vector)

$$\boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \dots) = (P(X(n) = 1), P(X(n) = 2), \dots). \quad (3.19)$$

- Thus, the vector $\boldsymbol{\pi}(n)$ is the distribution of the random variable $X(n)$, and it is called the **state probability at step n** .

- Law of total prob. $P(A) = \sum_n P(A \cap B_n) = \sum_n P(A | B_n) P(B_n)$:

$$\begin{aligned} \pi_i(n) &= \sum_k P(X(n-1) = k) P(X(n) = i | X(n-1) = k) = \sum_k \pi_k(n-1) p_{ki} \\ \pi_i(n) &= \sum_k P(X(0) = k) P(X(n) = i | X(0) = k) = \sum_k \pi_k(0) p_{ki}(n) \end{aligned} \quad (3.20)$$

- In matrix form:

$$\begin{aligned} \boldsymbol{\pi}(n) &= \boldsymbol{\pi}(n-1) \mathbf{P} \\ \boldsymbol{\pi}(n) &= \boldsymbol{\pi}(0) \mathbf{P}(n) \end{aligned} \quad (3.21)$$

where $\boldsymbol{\pi}(0)$ is the **initial distribution**.

- Iterating

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n-1) \mathbf{P} = \boldsymbol{\pi}(n-2) \mathbf{P} \mathbf{P} = \boldsymbol{\pi}(n-3) \mathbf{P} \mathbf{P} \mathbf{P} = \dots = \boldsymbol{\pi}(0) \mathbf{P}^n \quad (3.22)$$

- Thus:

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0) \mathbf{P}(n) = \boldsymbol{\pi}(0) \mathbf{P}^n \quad (3.23)$$

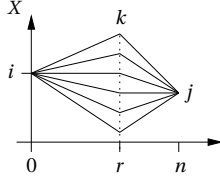
3.8 Chapman-Kolmogorov Equations

$$p_{ij}(n) = \sum_k p_{ik}(r) p_{kj}(n-r) \quad (3.24)$$

Proof.

$$\begin{aligned} p_{ij}(n) &= P(X(n) = j | X(0) = i) = \\ &= \sum_k P(X(n) = j, X(r) = k | X(0) = i) \\ &= \sum_k \frac{P(X(n) = j, X(r) = k, X(0) = i)}{P(X(0) = i)} \times \\ &= \sum_k \frac{P(X(r) = k, X(0) = i)}{P(X(r) = k, X(0) = i)} \times \\ &= \sum_k P(X(n) = j | X(r) = k, X(0) = i) \times \\ &= \sum_k P(X(r) = k | X(0) = i) \times \\ &= \sum_k P(X(n) = j | X(r) = k) P(X(r) = k | X(0) = i) \\ &= \sum_k p_{ik}(r) p_{kj}(n-r) \quad \square \end{aligned}$$

- Graphical interpretation:



- In matrix form:

$$\mathbf{P}(n) = \mathbf{P}(r) \mathbf{P}(n-r) \quad (3.25)$$

- Particularly:

$$\mathbf{P}(n) = \mathbf{P}(1) \mathbf{P}(n-1) = \mathbf{P} \mathbf{P}(n-1) = \mathbf{P}(n-1) \mathbf{P} \quad (3.26)$$

- Iterating:

$$\mathbf{P}(n) = \mathbf{P}^n \quad (3.27)$$

- Thus:

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0) \mathbf{P}(n) = \boldsymbol{\pi}(0) \mathbf{P}^n \quad (3.28)$$

Chapter 4

Transient Solution

4.1 Close Form Solution

- If we are interested in the **transient evolution** we shall study $\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0) \mathbf{P}^n$.
- If we can **diagonalize** \mathbf{P} , we can obtain the transient evolution in **close form**.
- \mathbf{P} can be **diagonalized** if \mathbf{P} can be decomposed as:

$$\mathbf{P} = \mathbf{L}^{-1} \boldsymbol{\Lambda} \mathbf{L} \quad (4.1)$$

where \mathbf{L} is some invertible matrix and $\boldsymbol{\Lambda}$ is the diagonal matrix

$$\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N) = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_N \end{bmatrix} \quad (4.2)$$

with λ_l , $l = 1, \dots, N$ the **eigenvalues** of \mathbf{P} .

- Assume a **finite DTMC** with N states. Then $\mathbf{P} = \mathbf{P}^{N \times N}$.
- Assume that \mathbf{P} can be **diagonalized**: $\mathbf{P} = \mathbf{L}^{-1} \boldsymbol{\Lambda} \mathbf{L}$, where $\boldsymbol{\Lambda}$ is the diagonal matrix $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$, with λ_l , $l = 1, \dots, N$ the eigenvalues of \mathbf{P} .
- But $\mathbf{L}^{-1} \text{diag}(\lambda_1^n, \dots, \lambda_N^n) \mathbf{L}$ are linear combinations of $\lambda_1^n, \dots, \lambda_N^n$. Thus, the probability of being in state i is given by:

$$\pi_i(n) = (\boldsymbol{\pi}(n))_i = \sum_{l=1}^N a_i^{(l)} \lambda_l^n \quad (4.3)$$

where the **unknown coefficients** $a_i^{(l)}$ can be obtained solving the system of equations:

$$\sum_{l=1}^N a_i^{(l)} \lambda_l^n = (\boldsymbol{\pi}(n))_i = (\boldsymbol{\pi}(0) \mathbf{P}^n)_i, \quad n = 0, \dots, N-1 \quad (4.4)$$

4.2 Eigenvalues

- The **eigenvalues** λ_l of a matrix \mathbf{A} are scalars that satisfy: $\mathbf{A} \mathbf{l} = \lambda_l \mathbf{l}$ (or $\mathbf{A} \mathbf{r} = \lambda_l \mathbf{r}$) for some row vectors \mathbf{l} (column vectors \mathbf{r}), referred to as **left and right eigenvectors**, respectively.

$$\begin{aligned} \mathbf{A} \mathbf{l} = \lambda_l \mathbf{l} &\Rightarrow \mathbf{l}(\mathbf{A} - \lambda_l \mathbf{I}) = 0 \Rightarrow \det(\lambda_l \mathbf{I} - \mathbf{A}) = 0 \\ \mathbf{A} \mathbf{r} = \lambda_l \mathbf{r} &\Rightarrow (\mathbf{A} - \lambda_l \mathbf{I}) \mathbf{r} = 0 \Rightarrow \det(\lambda_l \mathbf{I} - \mathbf{A}) = 0 \end{aligned} \quad (4.5)$$

- Thus, λ_l solve the **characteristic polynomial** $\det(\lambda \mathbf{I} - \mathbf{A}) = 0$.
- Note that, in general, **left and right eigenvectors** are different, but eigenvalues are the same (they solve the same **characteristic polynomial**).
- A matrix can be **diagonalized** if all eigenvalues are single (multiplicity = 1). If a matrix cannot be diagonalized it is called **defective**.

4.3 Determinant

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11} a_{22} - a_{12} a_{21} \quad (4.6)$$

$$\det \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{aligned} &+a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{21} a_{32} a_{13} \\ &-a_{31} a_{22} a_{13} - a_{12} a_{21} a_{33} - a_{23} a_{32} a_{11} \end{aligned} \quad (4.7)$$

- Cofactor Formula**: expanding along a row i :

$$\det \mathbf{A} = \sum_{j=1}^N a_{ij} (-1)^{i+j} \det M_{ij}, \quad (4.8)$$

where the **minor matrices** M_{ij} are obtained removing the row i and column j from \mathbf{A} . $(-1)^{i+j} \det M_{ij}$ is called the **cofactor** of a_{ij} .

$$\det \mathbf{A} = \prod \text{eigenvalues of } \mathbf{A} \quad (4.9)$$

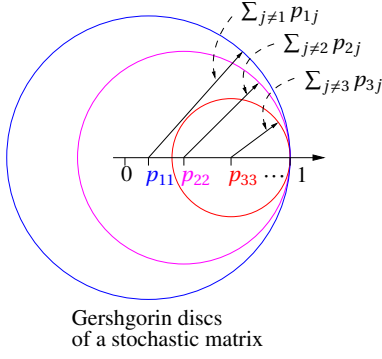
$$\text{trace } \mathbf{A} = \sum \text{eigenvalues of } \mathbf{A} \quad (4.10)$$

where $\text{trace } \mathbf{A} = \sum$ elements of the diagonal of \mathbf{A} .

4.4 Eigenvalues of a Stochastic Matrix

- \mathbf{P} has an **eigenvalue equal to 1** ($\mathbf{P} \mathbf{x} = \lambda \mathbf{x}$, for $\lambda = 1$).
- Proof**: $\mathbf{P} \mathbf{e} = \mathbf{e}$, where $\mathbf{e} = \begin{bmatrix} 1 & 1 & \dots \end{bmatrix}^T$ is a column vector of 1 (all rows of \mathbf{P} add to 1).
- All eigenvalues of \mathbf{P} are $|\lambda_l| \leq 1$.

Proof. Using Gerschgorin's theorem *The eigenvalues of a matrix $\mathbf{P}_{n \times n}$ lie within the union of the n circular disks with center p_{ii} and radius $\sum_{j \neq i} |p_{ij}|$ in \mathbb{C} .* Since $\sum_j p_{ij} = 1$, the property is proved.



Proof. of Gerschgorin's theorem

From $\mathbf{P}\mathbf{x} = \lambda\mathbf{x}$ we have

$$\sum_j p_{ij} x_j = \lambda x_i \quad \forall i \in \{1, \dots, n\}. \quad (4.11)$$

We choose i such that $|x_i| = \max_j |x_j|$. Thus, $\sum_{j \neq i} p_{ij} x_j = \lambda x_i - p_{ii} x_i$, and

$$|\lambda - p_{ii}| = \left| \sum_{j \neq i} p_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} \left| p_{ij} \frac{x_j}{x_i} \right| \leq \sum_{j \neq i} |p_{ij}| \quad (4.12)$$

and the equation $|\mathbf{x} - \mathbf{c}| \leq \mathbf{r}$, $\mathbf{x}, \mathbf{c} \in \mathbb{C}, \mathbf{r} \in \mathbb{R}$ is a disk of center \mathbf{c} and radius \mathbf{r} in \mathbb{C} .

Example

- Assume a DTMC with

$$\mathbf{P} = \begin{bmatrix} 4/5 & 1/5 \\ 2/5 & 3/5 \end{bmatrix}$$

- We want the probability of being in state 2 in n steps starting from state 1: $\pi_2(n)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

Solution

- It can be easily found that the **eigenvalues** of \mathbf{P} are $\lambda_1 = 1$ and $\lambda_2 = 2/5$.

$$\pi_2(n) = \lambda_1^n a + b \lambda_2^n = a + b(2/5)^n \quad (4.13)$$

- Imposing the **boundary conditions** $\pi_i(n) = (\boldsymbol{\pi}(0) \mathbf{P}^n)_i$:

$$\begin{aligned} \pi_2(0) &= a + b = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{P}^0 = (\mathbf{P}^0)_{12} = 0 \\ \pi_2(1) &= a + b(2/5) = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{P}^1 = (\mathbf{P})_{12} = 1/5 \end{aligned} \quad (4.14)$$

we have that $a = 1/3$, $b = -1/3$, thus:

$$\begin{aligned} \pi_2(n) &= 1/3 - 1/3(2/5)^n, \quad n \geq 0 \\ \pi_1(n) &= 1 - \pi_2(n) = 2/3 + 1/3(2/5)^n, \quad n \geq 0 \end{aligned} \quad (4.15)$$

4.5 Chain with a Defective Matrix

- What if \mathbf{P} cannot be diagonalized? (**defective** matrix).
- Let λ_l , $l = 1, \dots, L$ be the eigenvalues of $\mathbf{P}^{N \times N}$, each with multiplicity k_l ($k_l \geq 1$, $\sum_l k_l = N$), and a possible eigenvalue $\lambda_1 = 0$ with multiplicity k_1 . Then [1]:

$$\pi_j(n) = \sum_{m=0}^{k_1-1} a_j^{(1,m)} I(n=m) + \sum_{l=2}^L \lambda_l^n \sum_{m=0}^{k_l-1} a_j^{(l,m)} n^m, \quad (4.16)$$

$$1 \leq j \leq N, n \geq 0$$

$I(n=m)$ is the indicator func.: $I(n) = 1$ if $n=m$, $I(n) = 0$ if $n \neq m$.

Example

- Assume a DTMC with

$$\mathbf{P} = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 0 & 3/4 & 1/4 \\ 1 & 0 & 0 \end{bmatrix} \quad (4.17)$$

- We want the probability of being in state 1 in n steps starting from state 1: $\pi_1(n)$ with $\pi_1(0) = 1$.
- It can be easily found that the **eigenvalues** of \mathbf{P} are $\lambda_1 = 1$ and $\lambda_2 = 1/4$ with multiplicity 2. We guess:

$$\pi_1(n) = a + 1/4^n(b + cn) \quad (4.18)$$

- Imposing $\pi_1(0) = 1$, $\pi_1(1) = 3/4$, $\pi_1(2) = (3/4)^2$, we have:

$$\pi_1(n) = \frac{4}{9} + \frac{1}{4^n} \left(\frac{5}{9} + \frac{2}{3} n \right) \quad (4.19)$$

Chapter 5

Classification of States

Objective

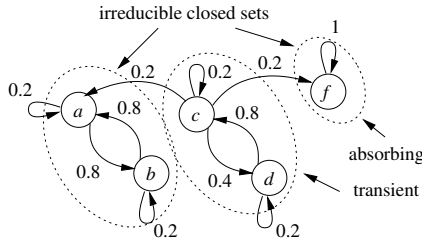
- Identify the different **types of behavior** that the chain can have.
- Introduce the concepts of **first passage probability** and **mean recurrence time**.

5.1 Irreducibility

- A state j is said to **communicate** with i , $i \leftrightarrow j$, if $p_{ij}(m_1) > 0$, $p_{ji}(m_2) > 0$ for some $m_1, m_2 \geq 0$.
- We define an **irreducible closed set, ICS** C_k as a set where all states communicate with each other, and have no transitions to other states out of the set:
 $i \leftrightarrow j, \forall i, j \in C_k$ and $p_{ij} = 0, \forall i \in C_k, j \notin C_k$
 (note that for $i \in C_k, j \notin C_k$ we have: $p_{ij}(2) = \sum_k p_{ik} p_{kj} = 0$, since $p_{ik} = 0$ if $k \notin C_k$, and $p_{kj} = 0$ if $k \in C_k$. Thus, $p_{ij}(n) = 0, \forall n$.)
- An **absorbing state** form an ICS of only one element. This state, i , must have $p_{ii} = 1$, $p_{ij} = 0 \forall j \neq i$.
- Transient states** do not belong to any ICS.
- A MC is **irreducible** if all the states form a unique ICS.
- Assume a MC has M ICSs: By properly numbering the states, we can write \mathbf{P} as an M block diagonal matrix with the probabilities of the transient states in the last rows.
- Example**, if $M = 3$:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2 \\ \text{at least one } > 0 & \mathbf{T} \end{bmatrix} \Rightarrow \boldsymbol{\pi}(n) = \boldsymbol{\pi}(0) \mathbf{P}^n = \boldsymbol{\pi}(0) \begin{bmatrix} \mathbf{P}_1^n & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_2^n \\ \text{at least one } > 0 & \mathbf{T}^n \end{bmatrix}$$

- Note that the M sub-matrices are **stochastic** (their rows sum 1).

Example

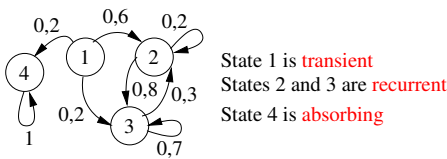
$$\mathbf{P} = \begin{array}{c|cc|cc|c} & a & b & f & c & d \\ \hline a & 0,2 & 0,8 & 0 & 0 & 0 \\ b & 0,8 & 0,2 & 0 & 0 & 0 \\ f & 0 & 0 & 1,0 & 0 & 0 \\ \hline c & 0,2 & 0 & 0,2 & 0,2 & 0,4 \\ d & 0 & 0 & 0 & 0,8 & 0,2 \end{array}$$

$$\mathbf{P}^\infty = \begin{array}{c|cc|cc|c} & a & b & f & c & d \\ \hline a & 0,5 & 0,5 & 0 & 0 & 0 \\ b & 0,5 & 0,5 & 0 & 0 & 0 \\ f & 0 & 0 & 1,0 & 0 & 0 \\ \hline c & 0,25 & 0,25 & 0,5 & 0 & 0 \\ d & 0,25 & 0,25 & 0,5 & 0 & 0 \end{array}$$

- What is the meaning of the probabilities in \mathbf{P}^∞ ? (recall that $(\mathbf{P}^n)_{ij} = p_{ij}(n) = P(X(n) = j | X(0) = i)$).

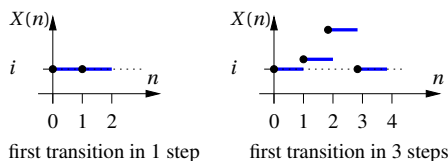
5.2 Transient and Recurrent

- Recurrent:** States that, being visited, they are visited again with probability 1. They are visited an infinite number of times when $n \rightarrow \infty$.
- Transient:** States that, being visited, have a probability > 0 of never being visited again. They are visited a finite number of times when $n \rightarrow \infty$.
- Absorbing:** A single (recurrent) state where the chain remains with probability = 1.

**5.3 First Passage (Transition) Probabilities**

- To derive a classification criteria, we shall study the distribution of the number of steps to **go for the first time from a state i another state j** . Definition:

$$f_{ii}(n) = P \left(\begin{array}{l} \text{first transition into state } i \\ \text{in } n \text{ steps starting from } i \end{array} \right) \quad (5.1)$$



- Do **not confuse** with the n -step transition probability $p_{ii}(n)$, where the state i can be visited in the intermediate states.

Relation between $f_{ii}(n)$ and $p_{ii}(n)$

- $f_{ii}(n)$ and $p_{ii}(n)$ satisfy:

$$\begin{aligned} f_{ii}(1) &= p_{ii}(1) \\ p_{ii}(n) &= \sum_{l=1}^n f_{ii}(l) p_{ii}(n-l), \quad n \geq 1 \end{aligned} \quad (5.2)$$

- The probability that the MC **eventually enters state i starting from i** is given by:

$$f_{ii} = \sum_{n=1}^{\infty} f_{ii}(n) \quad (5.3)$$

- If $f_{ii} = 1$ we say i is a **recurrent state**.
- If $f_{ii} < 1$ we say i is a **transient state**.

5.4 Mean Recurrence Time

- When $f_{ii} = 1$, we define the **mean recurrence time m_{ii}** as:

$$m_{ii} = \sum_{n=1}^{\infty} n f_{ii}(n) \quad (5.4)$$

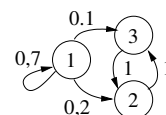
- m_{ii} is the **average number of steps to eventually reach i starting from i** . If $f_{ii} < 1$ (**transient state**) then we define $m_{ii} = \infty$.
- Classification of **recurrent states** ($f_{ii} = 1$):
 - If $m_{ii} = \infty$ the state is **null recurrent**: it takes an ∞ time to reach the state after leave it. Can only happen in chains with an infinite number of states.
 - If $m_{ii} < \infty$ the state is **positive recurrent**: the state is reached in a finite time after leave it.

5.5 Property of States

In **finite MC**:

- States can be only of type positive recurrent or transient.
- At least one state must be positive recurrent.
- There are not null recurrent states.

- Example:**



- State 1 is transient. States 2 and 3 are positive recurrent.

Generalization to Any State Pair

- Analogously to $f_{ii}(n)$, we define the probability of the **first passage to state j starting from any state i** in n steps: $f_{ij}(n)$.
- $f_{ij}(n)$ and $p_{ij}(n)$ satisfy:

$$p_{ij}(n) = \sum_{l=1}^n f_{ij}(l) p_{ij}(n-l), n \geq 1 \quad (5.5)$$

- When $f_{ij} = 1$, the average number of steps to eventually reach j starting from i , m_{ij} is given by:

$$m_{ij} = \sum_{n=1}^{\infty} n f_{ij}(n) \quad (5.6)$$

- If state j can not be reached starting from state i with probability one (if $f_{ij} < 1$), then we define $m_{ij} = \infty$.

5.6 Recursive Equation for the First Passage Probabilities

- Recall that the The probability that the MC **eventually enters state j starting from i** is given by: $f_{ij} = \sum_{n=1}^{\infty} f_{ij}(n)$
- f_{ij} can be computed as follows: Assume we are in i . With probability p_{ij} we will go to j in one step. Otherwise, we will go to k , $k \neq j$, and then we will reach j with probability f_{kj} . Thus:

$$f_{ij} = p_{ij} + \sum_{k \neq j} p_{ik} f_{kj} \quad (5.7)$$

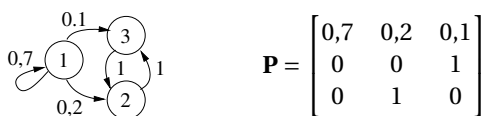
- If there are more than 1 **absorbing states**, we can compute the probability to reach them using this method (if there is only 1, say j , then $f_{ij} = 1, \forall i$).

5.7 Recursive Equation for the Mean Recurrence Time

- Recall that the **mean recurrence time** $m_{ij} = \sum_{n=1}^{\infty} n f_{ij}(n)$ is the average number of steps to eventually reach j starting from i , i.e. it is the mean first passage time from state i to j .
- When $f_{ij} = 1$, m_{ij} can be computed as follows: Assume we are in i . With probability p_{ij} we will go to j in one step. Otherwise, we will go to k , $k \neq j$, and then it will take m_{kj} steps to reach j . Thus:

$$m_{ij} = p_{ij} + \sum_{k \neq j} p_{ik} (1 + m_{kj}) = 1 + \sum_{k \neq j} p_{ik} m_{kj} \quad (5.8)$$

since $\sum_j p_{ij} = 1$.

Example: Recurrence Times Using the Definition

$$\begin{aligned} f_{21}(n) &= f_{31}(n) = 0 \\ f_{11}(n) &= 0,7 I(n=1) \\ f_{22}(n) &= f_{33}(n) = I(n=2) \\ f_{23}(n) &= f_{32}(n) = I(n=1) \end{aligned}$$

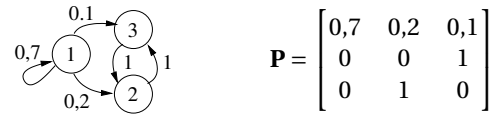
$$f_{12}(n) = \begin{cases} 0,2, & n=1 \\ 0,7^{n-1} 0,2 + 0,7^{n-2} 0,1, & n>1 \end{cases}$$

$$f_{13}(n) = \begin{cases} 0,1, & n=1 \\ 0,7^{n-1} 0,1 + 0,7^{n-2} 0,2, & n>1 \end{cases}$$

$$\begin{aligned} f_{11} &= 0,7 \\ f_{12} &= f_{13} = 1 \\ f_{32} &= f_{33} = 1 \end{aligned} \quad \begin{aligned} f_{22} &= f_{23} = 1 \\ f_{21} &= f_{31} = 0 \end{aligned}$$

$$M = (m_{ij}) = \begin{bmatrix} \infty & 11/3 & 12/3 \\ \infty & 2 & 1 \\ \infty & 1 & 2 \end{bmatrix}$$

- State 1 is **transient**. States 2 and 3 are **recurrent**.

Example: First Passage Probability Using Recursion

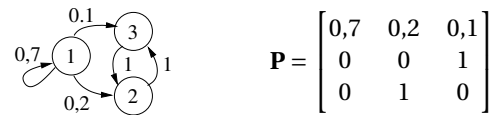
- We have:

$$f_{12} = p_{11} f_{12} + p_{12} + p_{13} f_{32} \quad (5.9)$$

- Clearly $f_{32} = 1$, thus:

$$f_{12} = 0,7 f_{12} + 0,2 + 0,1 \times 1 \Rightarrow f_{12} = 1 \quad (5.10)$$

as before.

Example: Mean Recurrence Time Using Recursion

- We have:

$$m_{12} = p_{12} + p_{11} (1 + m_{12}) + p_{13} (1 + m_{32}) = 1 + p_{11} m_{12} + p_{13} m_{32} \quad (5.11)$$

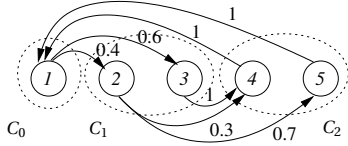
- Clearly $m_{32} = 1$, thus:

$$m_{12} = 1 + 0,7 m_{12} + 0,1 \times 1 \Rightarrow m_{12} = 11/3. \quad (5.12)$$

5.8 Periodic states

- A recurrent state j is **periodic** with period $d > 1$ if j can only be reached after leaving it with a multiple of d steps.
- If $d = 1$ the state is aperiodic.
- Any **periodic irreducible chain** can be partitioned in d **cyclic classes** C_0, \dots, C_{d-1} such that at each step a transition occur from class C_i to $C_{(i+1) \bmod d}$.
- By properly numerating the states, the transition matrix can be written as (the sub-matrices A_i may not be square):

$$P = \begin{matrix} & C_0 & C_1 & C_2 & \dots & C_{d-1} \\ \begin{matrix} C_0 \\ C_1 \\ \dots \\ C_{d-1} \end{matrix} & \begin{bmatrix} 0 & A_1 & 0 & \dots & 0 \\ 0 & 0 & A_2 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ A_{d-1} & 0 & 0 & \dots & 0 \end{bmatrix} \end{matrix} \quad (5.13)$$

Example

$$\mathbf{P} = \begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.3 & 0.7 \\ 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{P}^2 = \begin{bmatrix} 0 & 0 & 0 & 0.72 & 0.28 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \end{bmatrix},$$

$$\mathbf{P}^3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.72 & 0.28 \\ 0 & 0 & 0 & 0.72 & 0.28 \end{bmatrix},$$

$$\mathbf{P}^4 = \begin{bmatrix} 0 & 0.4 & 0.6 & 0 & 0 \\ 0 & 0 & 0 & 0.72 & 0.28 \\ 0 & 0 & 0 & 0.72 & 0.28 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \dots$$

- In periodic chains \mathbf{P}^n does not converge.

Chapter 6

Steady State

6.1 Limiting Distribution

- Probability of being in state i at step n :

$$\pi_i(n) = P(X(n) = i). \quad (6.1)$$

In vector form (row vector)

$$\boldsymbol{\pi}(n) = (\pi_1(n), \pi_2(n), \dots). \quad (6.2)$$

- The evolution of the chain depends on the initial distribution $\boldsymbol{\pi}(0)$.
- If we are interested in the **transient** evolution we shall study

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0) \mathbf{P}^n. \quad (6.3)$$

- If we are interested the **steady state** we shall be interested in the **limiting distribution** (if the limit exists):

$$\boldsymbol{\pi}(\infty) = (\pi_1(\infty), \pi_2(\infty), \dots) \quad (6.4)$$

Assume an **irreducible** chain with **positive recurrent** states.

- With infinite steps, we look for a probability converging to a value that depends only on the final state:

$$\pi_j(\infty) = \sum_i \pi_i(0) \lim_{n \rightarrow \infty} p_{ij}(n), \forall j \text{ and for any } \boldsymbol{\pi}(0), \quad (6.5)$$

which implies:

$$\pi_j(\infty) = \lim_{n \rightarrow \infty} p_{ij}(n) \sum_i \pi_i(0) = p_{ij}(\infty), \forall j \Rightarrow$$

$$\mathbf{P}(\infty) = \lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{bmatrix} \boldsymbol{\pi}(\infty) \\ \dots \\ \boldsymbol{\pi}(\infty) \end{bmatrix} \quad (6.6)$$

- If this limit exists, we call $\mathbf{P}(\infty)$ the **limiting matrix**, and $\boldsymbol{\pi}(\infty)$ the **limiting distribution**.

Example

$$\mathbf{P} = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$$

$$\mathbf{P}^2 = \begin{bmatrix} 0.764 & 0.168 & 0.068 \\ 0.760 & 0.170 & 0.070 \\ 0.752 & 0.174 & 0.074 \end{bmatrix}$$

$$\mathbf{P}^4 = \begin{bmatrix} 0.7628 & 0.1686 & 0.0686 \\ 0.7620 & 0.1690 & 0.0690 \\ 0.7604 & 0.1698 & 0.0698 \end{bmatrix}$$

$$\mathbf{P}^8 = \begin{bmatrix} 0.762500 & 0.168750 & 0.068750 \\ 0.762499 & 0.168750 & 0.068750 \\ 0.762497 & 0.168752 & 0.068752 \end{bmatrix}$$

$$\dots$$

$$\Rightarrow \boldsymbol{\pi}(\infty) = (0.76250, 0.16875, 0.06875)$$

6.2 Stationary distribution

- We have:

$$\pi_i(n) = P(X(n) = i) = \sum_k P(X(n-1) = k) P(X(n) = i | X(n-1) = k) = \sum_k \pi_k(n-1) p_{ki} \quad (6.7)$$

- In matrix form: $\boldsymbol{\pi}(n) = \boldsymbol{\pi}(n-1) \mathbf{P}$
- If $\pi_i(n) = \pi_i(n-1) = \pi_i \forall i$, we call π_i the **stationary probability of state i** , and $\boldsymbol{\pi} = (\pi_1, \pi_2, \dots)$, the **stationary distribution** of the chain.
- In matrix form (**Global balance equations**):

$$\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$$

$$\boldsymbol{\pi} \mathbf{e} = 1, \mathbf{e} = \begin{bmatrix} 1 & 1 & \dots \end{bmatrix}^T \quad (6.8)$$

- Thus, the stationary distribution is the left-hand eigenvector corresponding to the unit eigenvalue of \mathbf{P} .
- $\boldsymbol{\pi}(n) = \boldsymbol{\pi} \Rightarrow \boldsymbol{\pi}(n+1) = \boldsymbol{\pi}(n) \mathbf{P} = \boldsymbol{\pi} \mathbf{P} = \boldsymbol{\pi} \Rightarrow \boxed{\boldsymbol{\pi}(k) = \boldsymbol{\pi}, k \geq n}$
- Do not confuse the **limiting distribution** $\boldsymbol{\pi}(\infty)$ and the **stationary distribution** $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$.
- $\boldsymbol{\pi}(\infty)$ and $\boldsymbol{\pi}$ may not be the same, e.g. in **periodic chains** $\boldsymbol{\pi}(\infty)$ does not exist (\mathbf{P} does not converge), but we can compute the stationary distribution.

- **Example:** the periodic chain

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad (6.9)$$

has the stationary distribution

$$\boldsymbol{\pi} = \left[1/3 \quad 1/3 \quad 1/3 \right]. \quad (6.10)$$

6.3 Numerical Solution

Replace one equation method

$$\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}$$

$$\boldsymbol{\pi} \mathbf{e} = 1, \mathbf{e} = \begin{bmatrix} 1 & 1 & \dots \end{bmatrix}^T$$

We solve the equation $\boldsymbol{\pi}(\mathbf{I} - \mathbf{P}) = 0$ replacing the last equation by $\boldsymbol{\pi} \mathbf{e} = 1$:

$$\boldsymbol{\pi} \begin{bmatrix} p_{11}-1 & p_{12} & \dots & p_{1n-1} & 1 \\ p_{21} & p_{22}-1 & \dots & p_{2n-1} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn-1} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 & 1 \end{bmatrix} \quad (6.11)$$

Examples

- **Replace one equation method:** $\mathbf{P} = \begin{bmatrix} 0.8 & 0.15 & 0.05 \\ 0.7 & 0.2 & 0.1 \\ 0.5 & 0.3 & 0.2 \end{bmatrix}$
- With **octave** (matlab clone):

```
octave:1> P
      =[0.8,0.15,0.05;0.7,0.2,0.1;0.5,0.3,0.2];
octave:2> s=size(P,1); # number of rows.
octave:3> [zeros(1,s-1),1] / ...
> [eye(s,s-1)-P(1:s,1:s-1), ones(s,1)]
ans =
0.762500 0.168750 0.068750
```

- With **R**

```
> P <- matrix(nc=3, byr=T,
+ c(0.8,0.15,0.05,0.7,0.2,0.1,0.5,0.3,0.2))
> s <- nrow(P)
> solve(t(cbind(P[,1:(s-1)]-diag(nr=s,nc=s-1),
+ rep(1,s))), c(rep(0,s-1),1))
[1] 0.76250 0.16875 0.06875
```

NOTE: $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P} \Rightarrow \boldsymbol{\pi}^T = \mathbf{P}^T \boldsymbol{\pi}^T$. The transpose operator in R is `t()`.

6.4 Global balance equations

- Why are they called Global balance equations?

$$\left. \begin{aligned} \boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P} &\Rightarrow \pi_j = \sum_{i=0}^{\infty} \pi_i p_{ij} \\ \sum_{i=0}^{\infty} p_{ji} = 1 &\Rightarrow \pi_j \sum_{i=0}^{\infty} p_{ji} = \pi_j \end{aligned} \right\} \Rightarrow \sum_{i=0}^{\infty} \pi_i p_{ij} = \pi_j \sum_{i=0}^{\infty} p_{ji} \quad (6.12)$$

$$\sum_{i=0}^{\infty} \pi_i p_{ij} \Rightarrow \text{Frequency of transitions entering state } j$$

$$\pi_j \sum_{i=0}^{\infty} p_{ji} \Rightarrow \text{Frequency of transitions leaving state } j \quad (6.13)$$

- In **stationary regime**, the frequency of transitions leaving state j is equal to the frequency of transitions entering state j .

Flux Balancing

- Define the **flux** F_{uv} from state u to v :

$$F_{uv} = \pi_u p_{uv} \quad (6.14)$$

- and the flux from set of states U to V :

$$F(U,V) = \sum_{u \in U} \sum_{v \in V} F_{uv} \quad (6.15)$$

- From the Global balance equations we have:

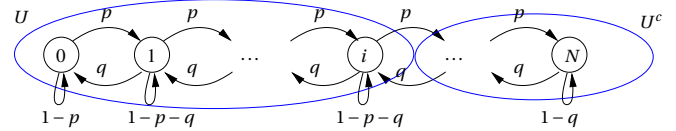
$$\sum_{i=0}^{\infty} \pi_i p_{ij} = \pi_j \sum_{i=0}^{\infty} p_{ji} \Rightarrow \sum_{i \in U} F_{ij} + \sum_{i \notin U} F_{ij} = \sum_{i \in U} F_{ji} + \sum_{i \notin U} F_{ji} \quad (6.16)$$

- Adding for $j \in U$:

$$\sum_{j \in U} \sum_{i \in U} F_{ij} + \sum_{j \in U} \sum_{i \notin U} F_{ij} = \sum_{j \in U} \sum_{i \in U} F_{ji} + \sum_{j \in U} \sum_{i \notin U} F_{ji} \Rightarrow$$

$$\sum_{j \in U} \sum_{i \notin U} F_{ij} = \sum_{j \in U} \sum_{i \notin U} F_{ji} \Rightarrow F(U, U^c) = F(U^c, U)$$

Solution Using Flux Balancing



- Flux balancing $\Rightarrow p \pi_i = q \pi_{i+1}$
- Iterating: $\pi_1 = \rho \pi_0, \pi_2 = \rho \pi_1 = \rho^2 \pi_0, \dots, \Rightarrow$

$$\pi_i = \rho^i \pi_0, i = 1, \dots, N \quad \text{where: } \rho = \frac{p}{q},$$

- Normalizing: $\sum_{i=0}^N \pi_i = 1$

$$\pi_0 = \frac{1-\rho}{1-\rho^{N+1}}, \quad p \neq q$$

$$\pi_0 = \frac{1}{N+1}, \quad p = q$$

6.5 Ergodic Chains

Ergodic state positive recurrent and aperiodic state.

Ergodic chain if all states are ergodic.

Theorem: All states of an irreducible Markov chain are of the same type: Transient or positive/null recurrent, and aperiodic/periodic with the same period [2, chapter XV].

Consequences:

- **Finite aperiodic and irreducible** chains are **ergodic** (since all states are positive recurrent).
- **Infinite aperiodic and irreducible** chains can be:
 - **Ergodic:** all the states are positive recurrent (stable chains).
 - **Non ergodic:** all states are null recurrent or transient (unstable chains).

6.5.1 Theorems for ergodic chains

- $\pi = \pi(\infty)$

Proof. For an **aperiodic irreducible** chain with **positive recurrent** states:

$$\begin{cases} \pi(\infty) &= \pi(0) \mathbf{P}(\infty) \\ \mathbf{P}(\infty) &= \lim_{n \rightarrow \infty} \mathbf{P}^n = \begin{bmatrix} \pi(\infty) \\ \dots \\ \pi(\infty) \end{bmatrix} \end{cases} \Rightarrow$$

$$\pi(\infty) \mathbf{P} = (\pi(0) \lim_{n \rightarrow \infty} \mathbf{P}^n) \mathbf{P} = \pi(0) \mathbf{P}(\infty) = \pi(\infty)$$

$$\Rightarrow \begin{cases} \pi(\infty) \mathbf{P} = \pi(\infty) \\ \pi(\infty) \mathbf{e} = 1 \end{cases} \quad \pi(\infty) \text{ satisfies the GBE} \Rightarrow \pi = \pi(\infty)$$
(6.17) \square

- In stationary regime (when $\pi(n) \mathbf{P} = \pi(n)$), the **mean number of steps the system remains in state j** during k steps is given by

$$k \pi_j. \quad (6.18)$$

Proof. Assume the chain in stationary regime at time $t = 0$ ($\pi(0) \mathbf{P} = \pi(0)$), and let $j(k)$ be the number of visits to j in k steps: $j(k) = \sum_{i=0}^{k-1} I(X(i) = j)$ ($I(A)$ is the indicator function: $I(A) = 1$ if A occurs, $I(A) = 0$ otherwise):

$$E[j(k)] = \sum_{i=0}^{k-1} E[I(X(i) = j)] = \sum_{i=0}^{k-1} P(X(i) = j) = k \pi_j \quad (6.19) \quad \square$$

- In stationary regime the **mean recurrence time** (mean number of steps between two consecutive visits to state j) is given by

$$m_{jj} = 1/\pi_j \quad (6.20)$$

Proof. Let $j(k)$ be the number of visits to j in k steps:

$$\pi_j = \lim_{k \rightarrow \infty} \frac{j(k)}{k} = \lim_{k \rightarrow \infty} \frac{1}{k/j(k)} = 1/m_{jj} \quad (6.21) \quad \square$$

Chapter 7

Reversed Chain

Definition

- Let $X(n)$ be an **ergodic** MC. The chain $X^r(n) = X(-n)$ is referred to as the **time reversal chain** of $X(n)$.
- **Example**, consider a possible sample path of $X(n)$:

$$\dots (i_0, n_0), (i_1, n_1), (i_2, n_2), \dots \quad (7.1)$$

The same path in the time reversal chain $X^r(n)$ would be:

$$\dots (i_2, -n_2), (i_1, -n_1), (i_0, -n_0), \dots \quad (7.2)$$

Properties

- Let p_{ij} , p_{ij}^r be the transition probabilities of $X(n)$ respectively $X^r(n)$, and π_i , π_i^r the stationary distributions of $X(n)$ respectively $X^r(n)$, then:

$$\pi_i = \pi_i^r \quad (7.3)$$

- **Proof:** the mean time in each state is the same for both chains.
- However, **in general** $p_{ij} \neq p_{ij}^r$. For example, $X(n)$ may be able to jump from state i to j , but not from j to $i \Rightarrow X^r(n)$ can jump from j to i , but not from i to j .
- But it must be $p_{ii} = p_{ii}^r$, since self-state transitions are the same in the direct and reversed chains.

7.1 Computation of p_{ij}^r

The transition probabilities in the time reversal chain (p_{ij}^r) satisfy:

$$\pi_i p_{ij} = \pi_j p_{ji}^r \quad (7.4)$$

Proof. Assume the chain in **steady state**. We have:

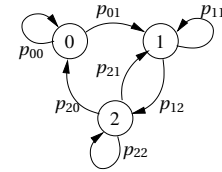
$$\begin{aligned} P\{X(n+1) = j, X(n) = i\} &= \\ P\{X^r(-n) = i, X^r(-n-1) = j\} &= \\ P\{X^r(n+1) = i, X^r(n) = j\} &\Rightarrow \\ P\{X(n) = i\} P\{X(n+1) = j \mid X(n) = i\} &= \\ P\{X(n) = j\} P\{X^r(n+1) = i \mid X^r(n) = j\} &\Rightarrow \\ \pi_i p_{ij} &= \pi_j p_{ji}^r. \quad \square \end{aligned}$$

We can **compute** p_{ji}^r using the **reversed balance equations**:

$$\pi_i p_{ij} = \pi_j p_{ji}^r \Rightarrow \sum_{i \in U} \sum_{j \in V} \pi_i p_{ij} = \sum_{i \in U} \sum_{j \in V} \pi_j p_{ji}^r \Rightarrow$$

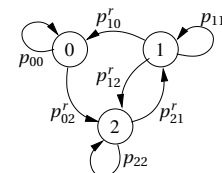
$$F(U, V) = F^r(V, U) \quad (7.5)$$

Example



$$\Rightarrow \begin{cases} \pi_0 = \frac{p_{12} p_{20}}{p_{12} p_{20} + p_{01} (p_{20} + p_{21}) + p_{01} p_{12}} \\ \pi_1 = \frac{p_{01} (p_{20} + p_{21})}{p_{12} p_{20} + p_{01} (p_{20} + p_{21}) + p_{01} p_{12}} \\ \pi_2 = \frac{p_{01} p_{12}}{p_{12} p_{20} + p_{01} (p_{20} + p_{21}) + p_{01} p_{12}} \end{cases}$$

Time reversal chain:



$$\Rightarrow \begin{cases} \pi_0 p_{01} = \pi_1 p_{10}^r \\ \pi_1 p_{12} = \pi_2 p_{21}^r \\ \pi_2 p_{21} = \pi_1 p_{12}^r \\ \pi_2 p_{20} = \pi_0 p_{02}^r \end{cases} \Rightarrow \begin{cases} p_{10}^r = \frac{p_{12} p_{20}}{p_{20} + p_{21}} \\ p_{21}^r = \frac{p_{20} + p_{21}}{p_{12} p_{21}} \\ p_{12}^r = \frac{p_{12} p_{21}}{p_{20} + p_{21}} \\ p_{02}^r = p_{01} \end{cases}$$

Chapter 8

Reversible Chains

Definition

- A chain is reversible if:

$$p_{ij} = p_{ij}^r \quad (8.1)$$

- This equality implies the **reversibility balance equations**:

$$\pi_i p_{ij} = \pi_j^r p_{ji}^r \Rightarrow F(U, V) = F^r(U, V) \quad (8.2)$$

- Using both reversed ($F^r(U, V) = F(V, U)$) and reversibility balance equations, the following relation holds for a reversible chain (**detailed balance equations**):

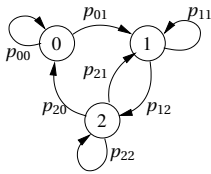
$$F(U, V) = F(V, U) \quad (8.3)$$

- NOTE: Compare with the **global balance equations**: $F(U, U^C) = F(U^C, U)$.

8.1 Kolmogorov Criteria

Definition of path

- Define a **path** as a possible sequence of transitions of the chain. For example, in the figure it could be $0 \rightsquigarrow 0 \rightsquigarrow 1 \rightsquigarrow 2$.



- We denote the **sequence of states** of one path l as:

$$(l, 1) \rightsquigarrow (l, 2) \rightsquigarrow \dots \rightsquigarrow (l, m) \quad (8.4)$$

- For instance, if l is $0 \rightsquigarrow 0 \rightsquigarrow 1 \rightsquigarrow 2$, then $(l, 1) = 0$, $(l, 2) = 0$, $(l, 3) = 1$, $(l, 4) = 2$.
- We say that a path is **closed** if it starts and ends in the same state. For instance, a path starting and ending in state $(l, 1)$:

$$(l, 1) \rightsquigarrow (l, 2) \rightsquigarrow \dots \rightsquigarrow (l, m) \rightsquigarrow (l, 1) \quad (8.5)$$

Kolmogorov Criteria

- Take a **closed path** l with $m \geq 0$ transitions, i.e.:

$$(l, 1) \rightsquigarrow (l, 2) \rightsquigarrow \dots \rightsquigarrow (l, m) \rightsquigarrow (l, 1), m \geq 0 \quad (8.6)$$

- The chain is **reversible iff for all** l :

$$p_{(l,1)(l,2)} p_{(l,2)(l,3)} \dots p_{(l,m)(l,1)} = p_{(l,1)(l,m)} p_{(l,m)(l,m-1)} \dots p_{(l,2)(l,1)} \quad (8.7)$$

• Proof:

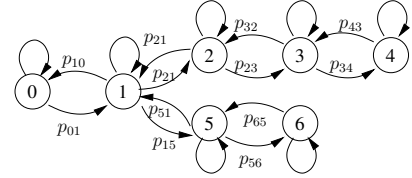
- If the chain is reversible $\pi_i p_{ij} = \pi_j p_{ji}$ (detailed balance equations): $\Rightarrow \pi_{(l,k)} p_{(l,k)(l,k+1)} = \pi_{(l,k+1)} p_{(l,k+1)(l,k)}$
- Multiplying for $k = 1, 2, \dots, m$ and simplifying we obtain the previous relation.

Corollary

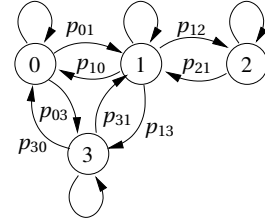
- A reversible chain must satisfy:

$$\begin{aligned} p_{ij} > 0 &\Rightarrow p_{ji} > 0 \\ p_{ij} = 0 &\Rightarrow p_{ji} = 0 \end{aligned} \quad (8.8)$$

- An **ergodic tree chain is always reversible**. We define a tree chain as chain having no cycles, i.e. the only possible transitions are bidirectional arcs between states, and self transitions.



Example



- The chain is **reversible iff**:

$$p_{01} p_{13} p_{30} = p_{10} p_{03} p_{31}$$

8.2 Product Form Solution

- Let $X(n)$ be a reversible MC with space state $S \Rightarrow$ the **stationary probabilities** of $X(n)$ can be computed as follows:
- Choose a state $\mathbf{s} \in S$,
- For every other state $\mathbf{i} \in S$, $\mathbf{i} \neq \mathbf{s}$ look for a possible path l_i from state \mathbf{s} to state \mathbf{i} :

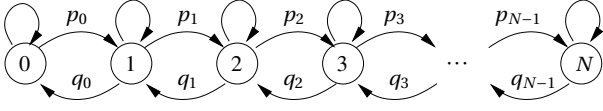
$$\mathbf{s} = (l_i, 1) \rightsquigarrow (l_i, 2) \rightsquigarrow \dots \rightsquigarrow (l_i, m_{l_i}) = \mathbf{i}, m_{l_i} \geq 1 \quad (8.9)$$

- The stationary probabilities are given by:

$$\pi_i = \frac{\psi_i}{\sum_{j \in S} \psi_j}, i \in S \quad \text{where } \psi_i = \begin{cases} 1, & i = \mathbf{s} \\ \prod_{k=1}^{m_{l_i}-1} \frac{p_{(l_i,k)(l_i,k+1)}}{p_{(l_i,k+1)(l_i,k)}}, & i \neq \mathbf{s} \end{cases} \quad (8.10)$$

- Proof** Use the detailed balance equations $\pi_i p_{ij} = \pi_j p_{ji}$.

8.3 Birth and Death Chains



- Birth and death chains are reversible.
- Applying the product form solution for reversible chains we obtain the general solution of birth death chains. Choosing $s = 0$:

$$\pi_i = \frac{\psi_i}{\sum_{j=0}^N \psi_j}, i \geq 0 \quad \text{where } \psi_i = \begin{cases} 1, & i = 0 \\ \prod_{k=0}^{i-1} \frac{p_k}{q_k}, & i = 1, \dots, N \end{cases} \quad (8.11)$$

Truncated Reversible Chain

- Consider a reversible MC X with a stationary distribution π_i .
- Suppose that we **truncate the chain** X and we obtain another irreducible chain X' .
- Then, X' is also reversible with stationary distribution:

$$\pi'_i = \frac{\pi_i}{G}, \quad \sum_k \pi'_k = 1 \quad (8.12)$$

Chapter 9

Research Example: Aloha

Access Protocol (see the paper of Kleinrock and Lam [5]).

• Pure Aloha:

- Broadcast radio system.
- **Single hop** system (all stations are in coverage).
- Whenever a station has a frame ready, it is transmitted.
- If two or more frames Tx overlap in time there is a **collision**, otherwise the frame is received correctly.
- Colliding frames are reTx after a **random time (backoff)**.

• Slotted Aloha:

- Time is slotted.
- Tx can only occur at the beginning of a slot.
- Collisions occur when 2 or more stations Tx in the same slot.

9.1 Analysis with finite population

Assumptions

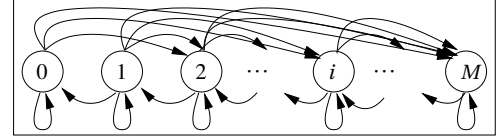
- **Slotted Aloha**.
- **Acks** are sent immediately after the reception of a frame, and are never lost.
- M nodes with a **buffer** of 1 frame.
- The **nodes** can be in 2 states:
 - **Thinking**: when the buffer is empty
 - **Backlogged**: when there is a frame in the buffer.

- A thinking node generate one frame in each slot with probability σ . When a frame collides, the frame is stored and the node becomes backlogged.
- A backlogged node ReTx the frame in each slot with probability v .

Markov Chain

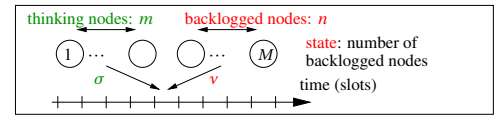
- The system **state**, $X(n)$, is the **number of backlogged nodes**:

$$p_{ij} = P(X(n) = j \text{ baklogged} | X(n-1) = i \text{ baklogged}) \quad (9.1)$$



Transition probabilities

- 0 for $j < i - 1$.
- for $j = i - 1$: no thinking Tx and only 1 backlogged Tx.
- for $j = i$:
 1. no thinking Tx and none or more than 1 backlogged Tx,
 2. only 1 thinking Tx and no backlogged Tx.
- for $j = i + 1$: 1 thinking and 1 or more backlogged Tx.
- for $j > i + 1$: $j - i$ thinking Tx, regardless of backlogged Tx.



In order to compute the previous events, define the probabilities:

- **Arrivals:** $Q_a(m, n)$, Probability of m thinking nodes Tx in a slot given that n nodes are backlogged:

$$Q_a(m, n) = P\left(\begin{matrix} m \text{ think.} \\ \text{nodes Tx} \end{matrix} \middle| \begin{matrix} n \text{ nodes are} \\ \text{backlogged} \end{matrix}\right) = \binom{M-n}{m} \sigma^m (1-\sigma)^{M-n-m} \quad (9.2)$$

- **Retransmissions:** $Q_r(m, n)$, Probability of m backlogged nodes Tx in a slot given that n nodes are backlogged:

$$Q_r(m, n) = P\left(\begin{matrix} m \text{ backl.} \\ \text{nodes Tx} \end{matrix} \middle| \begin{matrix} n \text{ nodes are} \\ \text{backlogged} \end{matrix}\right) = \binom{n}{m} v^m (1-v)^{n-m} \quad (9.3)$$

- and we have:

$$p_{ij} = \begin{cases} 0, & j < i - 1 \\ Q_a(0, i) Q_r(1, i), & j = i - 1 \\ Q_a(0, i) (1 - Q_r(1, i)) + Q_a(1, i) Q_r(0, i), & j = i \\ Q_a(1, i) (1 - Q_r(0, i)), & j = i + 1 \\ Q_a(j - i, i), & j > i + 1 \end{cases} \quad (9.4)$$

9.2. Throughput

9.1.1 Stationary distribution

- Solving the global balance equations:

$$\begin{aligned}\pi &= \pi \mathbf{P} \\ \pi \mathbf{e} &= 1\end{aligned}\quad (9.5)$$

- We obtain the probability of having i backlogged nodes:

$$\pi_i = P(i \text{ backlogged nodes}) \quad (9.6)$$

NOTE: there is **no closed form solution** of the chain. The matrix \mathbf{P} must be constructed using the expression of p_{ij} , and solved numerically.

9.2 Throughput

- Define the probabilities:

$$P_{succ}(i) = P(\text{successful Tx} \mid i \text{ backlogged}) \quad (9.7)$$

- The **normalized throughput**, i.e. proportion of steps with a successful transmission) is:

$$S = \sum_{i=0}^M P(\text{successful Tx} \mid i \text{ backlogged}) \times P(i \text{ backlogged}) = \sum_{i=0}^M P_{succ}(i) \pi_i \quad (9.8)$$

- For a slot to be successful: (i) 1 thinking Tx and no backlogged Tx, or (ii) no thinking Tx and 1 backlogged Tx:

$$P_{succ}(i) = Q_a(1, i) Q_r(0, i) + Q_a(0, i) Q_r(1, i) \quad (9.9)$$

Notes on the throughput

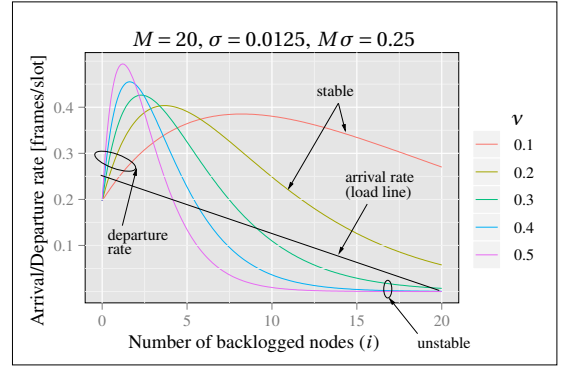
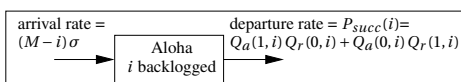
$$S = \sum_{i=0}^M P_{succ}(i) \pi_i \quad (9.10)$$

- For the **special case** $\sigma = \nu$ (thinking Tx with the same probability as backlogged): $P_{succ}(i) = M\sigma(1-\sigma)^{M-1}$, which does not depend on i , thus: $S = M\sigma(1-\sigma)^{M-1}$.
- The **offered load** (i.e. proportion of arrivals per slot) G is now: $G = M\sigma$, thus:

$$S = G \left(1 - \frac{G}{M}\right)^{M-1} \Rightarrow \lim_{M \rightarrow \infty} S = G e^{-G} \quad (9.11)$$

- We conclude that the **infinite population model** is the limit of the finite population if backlogged Tx with the same probability as thinking, and $M \rightarrow \infty$.

9.2.1 Dynamics



Note on the arrival rate (expected value of a binomial distribution):

$$\sum_{k=0}^{M-i} k \binom{M-i}{k} \sigma^k (1-\sigma)^{M-i-k} = (M-i) \sigma$$

Solving the chain: $S = \sum_{i=0}^M P_{succ}(i) \pi_i$

ν	S
0.1	2.38e-01
0.2	2.42e-01
0.3	1.30e-02
0.4	4.98e-04
0.5	1.90e-05

9.2.2 Stabilizing Aloha

- The **retransmission probabilities** must adapt in accordance with the state of the system.
- Example: **binary exponential backoff** (ethernet). The retransmission rate at retransmission i is adapted as $\nu = 2^{-i}$. Thus, the higher are the number of retransmission trials i , the lower (exponentially) is the retransmission rate.

Chapter 10

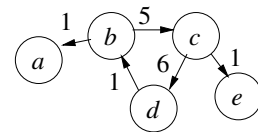
Finite Absorbing Chains

Canonical Form

- Let \mathbf{P}^{rxr} be the transition probability matrix of a chain with r states: s **transient** states and $r-s$ **absorbing** states. We can write \mathbf{P}^{rxr} in the **canonical form**:

$$\mathbf{P}^{rxr} = \begin{bmatrix} \mathbf{Q}^{s \times s} & \mathbf{R}^{s \times r-s} \\ \mathbf{0}^{r-s \times s} & \mathbf{I}^{r-s \times r-s} \end{bmatrix} \quad (10.1)$$

Example



$$\mathbf{P} = \begin{array}{c|cc|cc} & b & c & d & a & e \\ \hline b & 0 & 0.1 & 0 & 0.9 & 0 \\ c & 0 & 0 & 0.8 & 0 & 0.2 \\ d & 0.3 & 0 & 0.7 & 0 & 0 \\ \hline a & 0 & 0 & 0 & 1 & 0 \\ e & 0 & 0 & 0 & 0 & 1 \end{array}$$

10.1 Results

- Define:

$$\begin{aligned} \mathbf{n}_{ij} &= \begin{cases} \text{number of steps in state } j \text{ before} \\ \text{absorption, starting from state } i \end{cases}, \\ \mathbf{t}_i &= \begin{cases} \text{number of steps in transient states before} \\ \text{absorption, starting from state } i \end{cases}, \\ \mathbf{b}_{ij} &= P(\text{probability to be absorbed } j \text{ starting } i) \end{aligned} \quad (10.2)$$

- Then:

$$\begin{aligned} \{E[n_{ij}]\} &= \mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1} \\ \{\text{Var}[n_{ij}]\} &= \mathbf{N}(2\mathbf{N}_{\text{diag}} - \mathbf{I}) - \mathbf{N}\mathbf{sqr} \\ \{E[t_i]\} &= \boldsymbol{\tau} = \mathbf{N}\mathbf{e} \\ \{\text{Var}[t_i]\} &= (2\mathbf{N} - \mathbf{I})\boldsymbol{\tau} - \boldsymbol{\tau}_{\text{sqr}} \\ \{b_{ij}\} &= \mathbf{B} = \mathbf{N}\mathbf{R}. \end{aligned} \quad (10.3)$$

where $\{a_{ij}\}$ is a matrix with a_{ij} as element ij and \mathbf{e} is a column vector of 1s. \mathbf{N} is called the **fundamental matrix**.

Proofs

- $\{E[n_{ij}]\} = \mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}$

Proof.

$$\begin{aligned} E[n_{ij}] &= \sum_{k \in A} p_{ik} \delta_{ij} + \sum_{k \in T} p_{ik} E[n_{kj} + \delta_{ij}] = \\ &\quad \delta_{ij} + \sum_{k \in T} p_{ik} E[n_{kj}] \\ \Rightarrow \{E[n_{ij}]\} &= \mathbf{N} = \mathbf{I} + \mathbf{Q}\mathbf{N} \Rightarrow \mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1} \end{aligned}$$

where A is the set of absorbing states and T is the set of transient states.

$$\text{Notation: } \delta_{ij} = I(i = j) = \begin{cases} 1, & i = j, \\ 0, & \text{otherwise.} \end{cases} \quad \square$$

- $\{\text{Var}[n_{ij}]\} = \mathbf{N}(2\mathbf{N}_{\text{diag}} - \mathbf{I}) - \mathbf{N}\mathbf{sqr}$

Proof.

$$\begin{aligned} \text{Var}[n_{ij}] &= E[n_{ij}^2] - E[n_{ij}]^2 \Rightarrow \\ &\quad \{\text{Var}[\mathbf{n}_{ij}]\} = \{E[\mathbf{n}_{ij}^2]\} - \mathbf{N}\mathbf{sqr} \\ E[\mathbf{n}_{ij}^2] &= \sum_{k \in A} p_{ik} \delta_{ij}^2 + \sum_{k \in T} p_{ik} E[(n_{kj} + \delta_{ij})^2] = \\ &\quad \sum_{k \in A} p_{ik} \delta_{ij} + \sum_{k \in T} p_{ik} (E[n_{kj}^2] + 2E[n_{kj}] \delta_{ij} + \delta_{ij}) = \\ &\quad \delta_{ij} + \sum_{k \in T} (p_{ik} E[n_{kj}^2] + 2p_{ik} E[n_{kj}] \delta_{ij}) \Rightarrow \\ \{E[\mathbf{n}_{ij}^2]\} &= \mathbf{I} + \mathbf{Q}\{E[\mathbf{n}_{ij}^2]\} + 2(\mathbf{Q}\mathbf{N})_{\text{diag}} = \\ &\quad (\mathbf{I} - \mathbf{Q})^{-1} (\mathbf{I} + 2(\mathbf{Q}\mathbf{N})_{\text{diag}}) = \\ &\quad \mathbf{N}(\mathbf{I} + 2(\mathbf{N} - \mathbf{I})_{\text{diag}}) = \mathbf{N}(2\mathbf{N}_{\text{diag}} - \mathbf{I}) \quad \square \end{aligned}$$

- $\{E[t_i]\} = \boldsymbol{\tau} = \mathbf{N}\mathbf{e}$

Proof.

$$E[t_i] = \sum_{k \in T} E[n_{ik}] \Rightarrow \{E[t_i]\} = \boldsymbol{\tau} = \mathbf{N}\mathbf{e} \quad \square$$

- $\{\text{Var}[t_i]\} = (2\mathbf{N} - \mathbf{I})\boldsymbol{\tau} - \boldsymbol{\tau}_{\text{sqr}}$

Proof.

$$\begin{aligned} \text{Var}[t_i] &= E[t_i^2] - E[t_i]^2 \Rightarrow \{\text{Var}[\mathbf{t}_i]\} = \{E[\mathbf{t}_i^2]\} - \boldsymbol{\tau}_{\text{sqr}} \\ E[\mathbf{t}_i^2] &= \sum_{k \in A} p_{ik} + \sum_{k \in T} p_{ik} E[(t_k + 1)^2] = \\ &\quad \sum_{k \in A} p_{ik} + \sum_{k \in T} p_{ik} (E[t_k^2] + 2E[t_k] + 1) = \\ &\quad 1 + \sum_{k \in T} (p_{ik} E[t_k^2] + 2p_{ik} E[t_k]) \Rightarrow \\ \{E[\mathbf{t}_i^2]\} &= \mathbf{e} + \mathbf{Q}\{E[\mathbf{t}_i^2]\} + 2\mathbf{Q}\boldsymbol{\tau} = \\ &\quad (\mathbf{I} - \mathbf{Q})^{-1} (\mathbf{e} + 2\mathbf{Q}\boldsymbol{\tau}) = \\ &\quad \mathbf{N}(\mathbf{e} + 2\mathbf{Q}\boldsymbol{\tau}) = \boldsymbol{\tau} + 2\mathbf{N}\mathbf{Q}\boldsymbol{\tau} = \\ &\quad \boldsymbol{\tau} + 2(\mathbf{N} - \mathbf{I})\boldsymbol{\tau} = (2\mathbf{N} - \mathbf{I})\boldsymbol{\tau} \quad \square \end{aligned}$$

- $\{b_{ij}\} = \mathbf{B} = \mathbf{N}\mathbf{R}$

Proof.

$$\begin{aligned} b_{ij} &= p_{ij} + \sum_{k \in T} p_{ik} b_{kj}, \quad j \in A \Rightarrow \\ \{b_{ij}\} &= \mathbf{B} = \mathbf{R} + \mathbf{Q}\mathbf{B} \\ \Rightarrow \mathbf{B} &= (\mathbf{I} - \mathbf{Q})^{-1} \mathbf{R} = \mathbf{N}\mathbf{R}. \quad \square \end{aligned}$$

10.2 Extension of the Results

- The previous results can be generalized to any group of states of \mathbf{P} :
- A set S is referred to as **open** if the chain can reach some state of S^c starting from any state of S . Let

$$\mathbf{Q} = \{p_{ij}, i \in S, j \in S\} \quad (10.4)$$

$$\mathbf{R} = \{p_{ij}, i \in S, j \in S^c\} \quad (10.5)$$

Let assume that the process starts from $i \in S$. Define:

$$\begin{aligned} \mathbf{n}_{ij} &= \begin{cases} \text{number of steps in state } j \text{ before} \\ \text{leaving } S, \text{ starting from state } i \end{cases}, \\ \Rightarrow \{E[n_{ij}]\} &= \mathbf{N} = (\mathbf{I} - \mathbf{Q})^{-1}. \end{aligned}$$

- Similarly for the other results, e.g. $\boldsymbol{\tau} = \{E[t_i]\} = \mathbf{N}\mathbf{e}$ and $\mathbf{B} = \{b_{ij}\} = \mathbf{N}\mathbf{R}$.

10.3 Inverse of a matrix

Cofactors

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \mathbf{C}^T \quad (10.6)$$

where \mathbf{C}^T is the transposed cofactor matrix: $c_{ij} = (-1)^{i+j} \det \mathbf{M}_{ij}$, and \mathbf{M}_{ij} are the minor matrices obtained removing the row i and column j from \mathbf{A} .

Gaussian Elimination

Do the transformation:

$$[\mathbf{A} | \mathbf{I}] \rightarrow [\mathbf{I} | \mathbf{A}^{-1}] \quad (10.7)$$

using the elementary row operations:

- Swapping two rows.
- Multiplying a row by a nonzero number.
- Adding a multiple of one row to another row.

Part III

Continuous Time Markov Chains

Chapter 11

Definition of a CTMC

11.1 Properties of a continuous time MC

- The states must be a numerable set.
- Let $X(t)$ be the event {at time t the system is in state i }, then it must hold the **memoryless property**:

$$P(X(t_n) = i | X(t_1) = j, X(t_2) = k, \dots) = P(X(t_n) = i | X(t_1) = j) \text{ for any } t_n > t_1 > t_2 > t_3 \dots \quad (11.1)$$

11.2 Transition Matrix

- **Transition probabilities:**

$$p_{ij}(t_1, t_2) = P(X(t_2) = j | X(t_1) = i) \quad (11.2)$$

- For an **homogeneous chain**:

$$p_{ij}(t) = P(X(t_1 + t) = j | X(t_1) = i) = P(X(t) = j | X(0) = i), \forall t_1 \quad (11.3)$$

- In matrix form (**transition probability matrix**):

$$\mathbf{P}(t) = \begin{bmatrix} p_{11}(t) & p_{12}(t) & \dots \\ p_{21}(t) & p_{22}(t) & \dots \\ \dots & \dots & \dots \end{bmatrix}, t \geq 0 \quad (11.4)$$

- Notes:

- Compare with the n-step prob. matrix of a DTMC: $\mathbf{P}(n)$.
- $\mathbf{P}(t)$ must be a **stochastic matrix** (all rows add to 1).

- We look for an equivalent 1-step prob. matrix \mathbf{P} of DTMCs.

- For consistency: $\lim_{t \rightarrow 0} p_{ij}(t) = \delta_{ij}$. In matrix form:

$$\lim_{t \rightarrow 0} \mathbf{P}(t) = \mathbf{I}. \quad (11.5)$$

- And assume that the following **transition rates** exist:

$$q_{ij} = \lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t}, i \neq j; \quad q_{ii} = \lim_{t \rightarrow 0} \frac{p_{ii}(t) - 1}{t} \quad (11.6)$$

- In matrix form: $\mathbf{Q} = \lim_{t \rightarrow 0} \frac{\mathbf{P}(t) - \mathbf{I}}{t}$

- Note that $\sum_j p_{ij}(t) = 1 \Rightarrow p_{ii}(t) = 1 - \sum_{j \neq i} p_{ij}(t)$, thus:

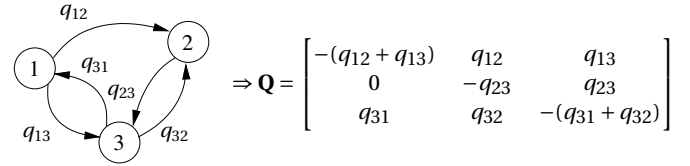
$$q_{ii} = \lim_{t \rightarrow 0} \frac{p_{ii}(t) - 1}{t} = \lim_{t \rightarrow 0} \frac{-\sum_{j \neq i} p_{ij}(t)}{t} = -\sum_{j \neq i} q_{ij} \quad (11.7)$$

- The matrix \mathbf{Q} is called the **transition rate or infinitesimal generator** of the chain.

- Since $q_{ii} = -\sum_{j \neq i} q_{ij}$, **all the rows of \mathbf{Q} add to 0**.
- The rate q_{ij} , $i \neq j$ measures “how fast” the chain moves from state i to j : the higher is q_{ij} , the faster it moves from i to j .
- For $q_{ii} = -\sum_{j \neq i} q_{ij}$, the higher $-q_{ii}$ is, the faster the chain leaves state i .
- If $q_{ij} = 0, \forall j \Rightarrow q_{ii} = 0$, then i is an **absorbing state**: the chain “moves with rate 0 from i to other states”, i.e. never leaves i .

11.3 State Transition Diagram

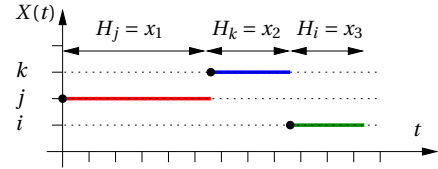
- A continuous MC is characterized by the transition rate or infinitesimal generator: the \mathbf{Q} -matrix.
- The state transition diagram is now represented as:



- Note that now we have **transition rates** ($0 \leq q_{ij} < \infty, i \neq j$) and **not probabilities**.
- The rates q_{ii} are **not written** in the diagram, **no self transitions**.

11.4 Sojourn Time

- Sojourn or holding time: Is the RV H_k equal to the sojourn time in state k :



- The Markov property implies that **the sojourn time is exponentially distributed with parameter q_{ii}** :

$$P(H_i \leq x) = 1 - e^{-q_{ii}x} \Rightarrow P(H_i > x) = e^{-q_{ii}x}, q_{ii} = -\sum_{j \neq i} q_{ij}, x \geq 0 \quad (11.8)$$

The exponential distribution satisfies the Markov property

- Markov property (**memoryless**):

$$P(X(t_2) = i | X(t_1) = i, X(0) = i) = P(X(t_2) = i | X(t_1) = i), t_2 > t_1 > 0 \quad (11.9)$$

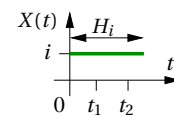


Figure 11.1: Sojourn time.

- In terms of the sojourn time:

$$P(H_i > t_2 | H_i > t_1) = P(H_i > t_2 - t_1)$$

- But:

$$P(H_i > t_2 | H_i > t_1) = \frac{P(H_i > t_2, H_i > t_1)}{P(H_i > t_1)} = \frac{P(H_i > t_2)}{P(H_i > t_1)} = \frac{e^{q_{ii} t_2}}{e^{q_{ii} t_1}} = e^{q_{ii} (t_2 - t_1)} = P(H_i > t_2 - t_1)$$

- The **exponential distribution** is the only one satisfying the **memoryless property**.

11.5 Exponential Jumps Description of a CTMC

Assume a process built as follows:

- Upon reaching a state i
 1. the process can jump to a state $j \in \{1, 2, \dots, l\}$.
 2. A set of **independent exponential RVs**, $\{H_{i1}, H_{i2}, \dots, H_{il}\}$, with parameters $\{q_{i1}, q_{i2}, \dots, q_{il}\}$ are triggered. That is, $P(H_{ij} \leq t) = 1 - e^{-q_{ij} t}$, $t \geq 0$.
- If $\min\{H_{i1}, H_{i2}, \dots, H_{il}\} = H_{ij} \Rightarrow$ the process jumps to the state j . In other words, a transition occurs to state j if the RV H_{ij} is the minimum of $\{H_{i1}, H_{i2}, \dots, H_{il}\}$ (see figure 11.2).

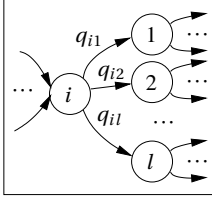


Figure 11.2: Jumps in a CTMC.

Theorem: This process is a CTMC with transition rates q_{ij} .

$$P(H_{ij} \leq t) = 1 - e^{-q_{ij} t}. \quad (11.10)$$

Proof. • The RV $H_i = \min\{H_{i1}, H_{i2}, \dots, H_{il}\}$ (sojourn time in state i) is **exponentially distributed** with parameter $q_i = \sum_j q_{ij}$: $P(H_i \leq t) = 1 - e^{-q_i t}$.

- $P(\min\{H_{i1}, H_{i2}, \dots, H_{il}\} = H_{ij}) = q_{ij} / \sum_j q_{ij}$. Thus, the **transition rate to state j** is:

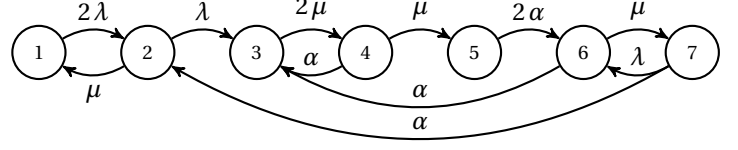
$$\lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t} = \lim_{t \rightarrow 0} \frac{P(\min\{H_{i1}, H_{i2}, \dots, H_{il}\} = H_{ij}) \times P(H_i \leq t)}{t} = \frac{q_{ij}}{\sum_j q_{ij}} \frac{\partial P(H_i \leq t)}{\partial t} \Big|_{t=0} = \frac{q_{ij}}{\sum_j q_{ij}} \sum_j q_{ij} = q_{ij} \quad (11.11)$$

□

Example: Pure Aloha System

- Consider a **Pure Aloha System** with **2 nodes**:
 - Nodes in **thinking state** Tx a packet in a time exponentially distributed with rate λ .
 - **Transmission time** is exponentially distributed with rate μ .

- If two transmissions overlap, the packet is lost and stations become backlogged (after the packet transmission) until the packet is successfully transmitted.
- Nodes in **backlogged state** Tx a packet in a time exponentially distributed with rate α .



State	Condition	Legend
1	T, T	T Thinking
2	X, T	X Transmitting
3	C, C	C Collided transmission
4	B, C	B Backlogged
5	B, B	
6	X, B	
7	T, B	

Chapter 12

Transient Solution

12.1 Chapman-Kolmogorov Equations

- **Chapman-Kolmogorov:** $p_{ij}(t) = \sum_k p_{ik}(t - \alpha) p_{kj}(\alpha)$, $0 \leq \alpha \leq t$

- Thus:

$$\frac{p_{ij}(t + \Delta t) - p_{ij}(t)}{\Delta t} = \sum_k \left\{ \frac{p_{ik}(t + \Delta t - \alpha) - p_{ik}(t - \alpha)}{\Delta t} p_{kj}(\alpha) \right\}$$

- Taking the limit

$$\alpha \rightarrow t, \Delta t \rightarrow 0 \Rightarrow \begin{cases} p_{ik}(t - \alpha) \rightarrow 0, & i \neq k \\ p_{ik}(t - \alpha) \rightarrow 1, & i = k \end{cases}$$

and using:

$$\begin{cases} q_{ij} = \lim_{t \rightarrow 0} \frac{p_{ij}(t)}{t}, & i \neq j \\ q_{ij} = \lim_{t \rightarrow 0} \frac{p_{ij}(t) - 1}{t}, & i = j \end{cases}$$

we have:

$$\frac{\partial p_{ij}(t)}{\partial t} = \sum_k q_{ik} p_{kj}(t), \quad t \geq 0, \forall i, j \quad (12.1)$$

- In matrix form:

$$\frac{\partial \mathbf{P}(t)}{\partial t} = \mathbf{Q} \mathbf{P}(t), \quad t \geq 0 \quad (12.2)$$

known as the master equations of a CTMC.

- The solution of the previous matrix differential equation is the **exponential matrix**:

$$\mathbf{P}(t) = e^{\mathbf{Q}t} = \sum_{i=0}^{\infty} \frac{(\mathbf{Q}t)^i}{i!} = \mathbf{I} + \mathbf{Q}t + \frac{\mathbf{Q}^2 t^2}{2!} + \frac{\mathbf{Q}^3 t^3}{3!} + \dots, \quad t \geq 0 \quad (12.3)$$

- Due to rounding errors, the previous series is difficult to compute numerically (the powers of \mathbf{Q} have positive and negative entries).

12.2 State Probabilities

- Define the probability of being in state i at time t :

$$\pi_i(t) = P(X(t) = i) \quad (12.4)$$

- In vector form (row vector)

$$\boldsymbol{\pi}(t) = (\pi_1(t), \pi_2(t), \dots). \quad (12.5)$$

- Clearly:

$$\pi_i(t) = \sum_k P(X(0) = k) P(X(t) = i | X(0) = k) = \sum_k \pi_k(0) p_{ki}(t) \quad (12.6)$$

- In matrix form:

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) \mathbf{P}(t) = \boldsymbol{\pi}(0) e^{\mathbf{Q}t}, t \geq 0 \quad (12.7)$$

where $\boldsymbol{\pi}(0)$ is the **initial distribution**.

- NOTE:** Compare with DTMC

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0) \mathbf{P}^n, n \geq 0 \quad (12.8)$$

12.3 Transient Solution

- If we are interested in the **transient evolution** we shall study $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) \mathbf{P}(t) = \boldsymbol{\pi}(0) e^{\mathbf{Q}t}, t \geq 0$.
- Assume a **finite CTMC** with N states (infinitesimal generator $\mathbf{Q}^{N \times N}$).
- Assume that \mathbf{Q} can be **diagonalized**: $\mathbf{Q} = \mathbf{L}^{-1} \boldsymbol{\Lambda} \mathbf{L}$, where $\boldsymbol{\Lambda}$ is the diagonal matrix $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \dots, \lambda_N)$, with $\lambda_l, l = 1, \dots, N$ the eigenvalues of \mathbf{Q} .
- NOTE:** the **eigenvalues** λ_l of a matrix \mathbf{A} are scalars that satisfy: $\mathbf{I}\mathbf{A} = \lambda_l \mathbf{I}$ (or $\mathbf{A}\mathbf{r} = \lambda_l \mathbf{r}$) for some row vectors \mathbf{I} (column vectors \mathbf{r}), referred to as **left and right eigenvectors**, respectively. Thus, solve the **characteristic polynomial** $\det(\lambda \mathbf{I} - \mathbf{A}) = 0$.
- Since

$$\mathbf{P}(t) = e^{\mathbf{Q}t} = \sum_{i=0}^{\infty} \frac{(\mathbf{Q}t)^i}{i!} = \sum_{i=0}^{\infty} \frac{(\mathbf{L}^{-1} \boldsymbol{\Lambda} \mathbf{L}t)^i}{i!} = \mathbf{L}^{-1} \text{diag}\left(\sum_{i=0}^{\infty} \frac{(\lambda_1 t)^i}{i!}, \dots\right) \mathbf{L} = \mathbf{L}^{-1} \text{diag}(e^{\lambda_1 t}, \dots) \mathbf{L} \quad (12.9)$$

we have that

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) e^{\mathbf{Q}t} = \boldsymbol{\pi}(0) \mathbf{L}^{-1} \text{diag}(e^{\lambda_1 t}, e^{\lambda_2 t}, \dots, e^{\lambda_L t}) \mathbf{L} \quad (12.10)$$

- Thus, the probability of being in state i is given by:

$$\pi_i(t) = (\boldsymbol{\pi}(t))_i = \sum_{l=1}^N a_i^{(l)} e^{\lambda_l t}, t \geq 0 \quad (12.11)$$

where the **unknown coefficients** $a_i^{(l)}$ can be obtained solving the system of equations:

$$\left. \frac{\partial^n \pi_i(t)}{\partial t^n} \right|_{t=0} = (\boldsymbol{\pi}(0) \mathbf{Q}^n)_i, n = 0, \dots, N-1 \quad (12.12)$$

NOTE: Compare with $(\boldsymbol{\pi}(n))_i = (\boldsymbol{\pi}(0) \mathbf{P}^n)_i, n = 0, \dots, N-1$

12.3.1 Eigenvalues of an Infinitesimal Generator

- \mathbf{Q} has an **eigenvalue equal to 0** ($\mathbf{Q}\mathbf{x} = \lambda\mathbf{x}$, for $\lambda = 0, \mathbf{x} \neq \mathbf{0}$).
Proof: $\mathbf{Q}\mathbf{e} = \mathbf{0}$, where $\mathbf{e} = (1, 1, \dots)^T$ is a column vector of 1 (all rows of \mathbf{Q} add to 0).
- The eigenvalue $\lambda = 0$ is **single** if \mathbf{Q} is **irreducible** (Perron-Frobenius theorem). \mathbf{Q} is irreducible if all states communicate: for $t > 0, p_{ij}(t) > 0, \forall i, j$.
- All eigenvalues of \mathbf{Q} are $\lambda_l \leq 0$.
Proof: Using Gerschgorin's theorem (see figure 12.1) and the fact that the rows of \mathbf{Q} add to 0.

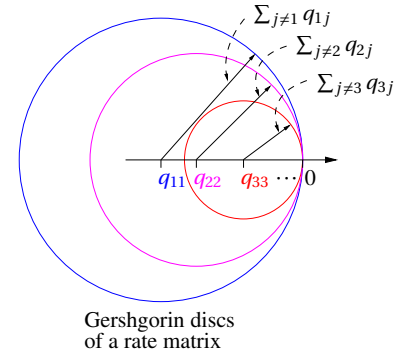


Figure 12.1: Gerschgorin discs of a CTMC

Example

- Assume a CTMC with

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 \\ 1/2 & -1/2 \end{bmatrix}$$

- We want the probability of being in state 2 at time t starting from state 1: $\pi_2(t)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 \end{bmatrix}$.

Solution

- It can be easily found that the **eigenvalues** of \mathbf{Q} are $\lambda_1 = 0$ and $\lambda_2 = -3/2$.

$$\pi_2(t) = a e^{\lambda_1 t} + b e^{\lambda_2 t} = a + b e^{-(3/2)t} \quad (12.13)$$

- Imposing the **boundary conditions**:

$$\pi_2(0) = a + b = (\boldsymbol{\pi}(0) \mathbf{Q}^0)_2 = (\boldsymbol{\pi}(0) \mathbf{I})_2 = (\boldsymbol{\pi}(0))_2 = 0$$

$$\left. \frac{\partial \pi_2(t)}{\partial t} \right|_{t=0} = b(-3/2) = (\boldsymbol{\pi}(0) \mathbf{Q})_2 = \mathbf{Q}_{12} = 1 \quad (12.14)$$

we have that $a = 2/3, b = -2/3$, thus:

$$\pi_2(t) = 2/3(1 - e^{-(3/2)t}), t \geq 0 \quad (12.15)$$

12.3.2 Chain with a Defective Matrix

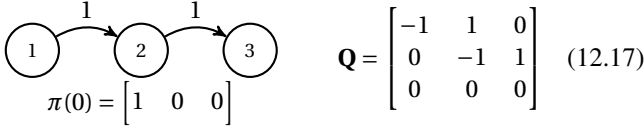
- What if \mathbf{Q} cannot be diagonalized? (**defective** matrix).
- Let λ_l , $l = 1, \dots, L$ be the eigenvalues of $\mathbf{Q}^{N \times N}$, each with multiplicity k_l ($k_l \geq 1$, $\sum_l k_l = N$). Then [1]:

$$\pi_j(t) = \sum_{l=1}^L e^{\lambda_l t} \sum_{m=0}^{k_l-1} a_j^{(l,m)} t^m \quad (12.16)$$

where $a_j^{(l,m)}$ are constants. So, exponentials associated with eigenvalues λ_l of multiplicity $k_l > 1$ are multiplied by polynomials in t of degree $k_l - 1$.

Example

- Assume the CTMC:



- We have $\lambda_1 = 0$ and $\lambda_2 = -1$ with multiplicity 2. Thus:

$$\pi_3(t) = a + e^{-t}(b + c t) \quad (12.18)$$

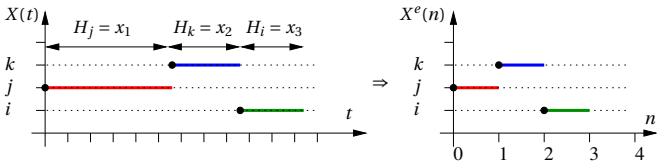
- We have that $a = 1$, because state 3 is absorbing. Imposing $\pi_3(0) = 0$ and $\pi'_3(0) = 0$, we have $b = c = -1$, and

$$\pi_3(t) = 1 - e^{-t}(1 + t), t \geq 0 \quad (12.19)$$

Chapter 13

Embedded MC of a CTMC

13.1 Definition



- We form a discrete time process $X^e(n)$, called the **Embedded MC (EMC)**, by looking at a CTMC at the transition time instants.

Theorem: This process is a **DTMC** with transition probabilities:

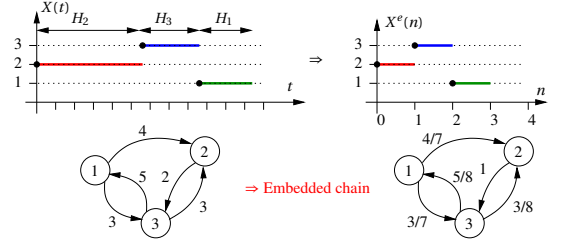
$$p_{ij}^e = \begin{cases} 0, & i = j \\ \frac{q_{ij}}{\sum_{j \neq i} q_{ij}}, & i \neq j \end{cases} \quad (13.1)$$

- NOTE:** If i is **absorbing** ($q_{ii} = 0$), we define $p_{ii}^e = 1$.

Proof.

- The EMC satisfies the **memoryless** property.
- Since we look the system only upon transition to a different state: $p_{ii}^e = 0$. NOTE: it might be $p_{ii}^e \neq 0$ if we look at transitions that end up in the same state.
- The probability that there is a transition from state i to j in the CTMC is the probability that the exponentially distributed RV with parameter q_{ij} is the **minimum from the independent exponentially distributed RVs** with parameters $\{q_{ik}\}_{k \neq i}$. This probability is $q_{ij} / \sum_{k \neq i} q_{ik}$.

Example



$$\mathbf{Q} = \begin{bmatrix} -7 & 4 & 3 \\ 0 & -2 & 2 \\ 5 & 3 & -8 \end{bmatrix} \Rightarrow \mathbf{P}_e = \begin{bmatrix} 0 & 4/7 & 3/7 \\ 0 & 0 & 1 \\ 5/8 & 3/8 & 0 \end{bmatrix} \quad (13.2)$$

- Each **transition** in the CTMC is a transition in the EMC.
- One step in i in the EMC is a **sojourn time** H_i in the CTMC.

Chapter 14

Classification of States

14.1 Irreducibly

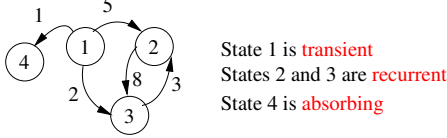
- A state j is said to **communicate** with i , $i \leftrightarrow j$, if $p_{ij}(t_1) > 0$, $p_{ji}(t_2) > 0$ for some $t_1 \geq 0$, $t_2 \geq 0$.
- We define an **irreducible closed set, ICS** C_k as a set where all states communicate with each other, and have no transitions to other states out of the set:
 $i \leftrightarrow j, \forall i, j \in C_k$ and $q_{ij} = 0, \forall i \in C_k, j \notin C_k$
- An **absorbing state** form an ICS of only one element. This state, i , must have $q_{ij} = 0 \forall i, j$.
- Transient states** do not belong to any ICS.
- A MC is **irreducible** if all the states form a unique ICS.
- Assume a MC has M **ICSs**: By properly numbering the states, we can write \mathbf{P} as an M block diagonal matrix with the probabilities of the transient states in the last rows.
- Example**, if $M = 3$:

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}_2 & \\ & & \mathbf{Q}_3 \\ \text{at least one } > 0 & & \mathbf{T} \end{bmatrix} \Rightarrow \pi(t) = \pi(0) e^{\mathbf{Q}t} = \pi(0) \begin{bmatrix} e^{\mathbf{Q}_1 t} & & \mathbf{0} \\ \mathbf{0} & e^{\mathbf{Q}_2 t} & \\ & & e^{\mathbf{Q}_3 t} \\ \text{at least one } > 0 & & e^{\mathbf{T}t} \end{bmatrix}$$

- Note that the M **sub-matrices are infinitesimal generators** (their rows add to 0).

14.2 Transient and Recurrent

- **Recurrent:** States that, being visited, have a probability > 0 of being visited again. They are visited an infinite number of times when $t \rightarrow \infty$.
- **Transient:** States that, being visited, have a probability > 0 of never being visited again. They are visited a finite number of times when $t \rightarrow \infty$.
- **Absorbing:** A single (recurrent) state where the chain remains with probability $= 1$.



- To derive a classification criteria, we shall study the **embedded MC (EMC)**, and proceed as in DTMC: Let $f_{ij}^e(n)$ the first passage prob. of the EMC, and $f_{ij}^e = \sum_{n=1}^{\infty} f_{ij}^e(n)$.
- If $f_{ii}^e = 1$ we say i is a **recurrent state**.
- If $f_{ii}^e < 1$ we say i is a **transient state**.
- When $f_{ij}^e = 1$, we define the **mean recurrence time of the EMC** $m_{ij}^e = \sum_{n=1}^{\infty} n f_{ij}^e(n)$. **NOTE:** in **steps**, not time units.
- If $m_{ii}^e = \infty$ the state is **null recurrent**.
- If $m_{ii}^e < \infty$ the state is **positive recurrent**.
- **NOTES:** (i) Even if the EMC is periodic, **there are not periodic CTMC** (it has no sense). (ii) f_{ij}^e and m_{ij}^e can be computed using the **recursive equations** for DTMC.

14.3 Mean recurrence time of the CTMC

- If the chain is in i at a time t , it takes an **expected time to leave i** equal to $1/(-q_{ii}) = 1/\sum_{j \neq i} q_{ij}$ (**sojourn time exponentially distributed** with rate $q_i = -q_{ii} = \sum_{j \neq i} q_{ij}$).
- Thus, if the chain is **in state i** , it takes a **mean time to enter state j** (**mean first passage time**):

$$m_{ij} = \frac{1}{q_i} + \sum_{k \neq i} p_{ik}^e m_{kj} \quad (14.1)$$

- Since: $p_{ij}^e = \begin{cases} 0, & i = j \\ \frac{q_{ij}}{\sum_{j \neq i} q_{ij}} = \frac{q_{ij}}{q_i}, & i \neq j \end{cases}$ we have:

$$m_{ij} = \frac{1}{q_i} + \sum_{k \neq i} p_{ik}^e m_{kj} = \frac{1}{q_i} + \sum_{k \neq i} \frac{q_{ik}}{q_i} m_{kj} \text{ [time units]} \quad (14.2)$$

Chapter 15

Steady State

15.1 Limiting Distribution

- The probability to be in state i at time t is:

$$\pi_i(t) = P(X(t) = i) = \sum_k \pi_k(0) p_{ki}(t), t \geq 0 \quad (15.1)$$

- In matrix form:

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) \mathbf{P}(t) = \boldsymbol{\pi}(0) e^{\mathbf{Q}t}, t \geq 0 \quad (15.2)$$

- Assume that the following limit exists:

$$\boldsymbol{\pi}(\infty) = \lim_{t \rightarrow \infty} \boldsymbol{\pi}(t) = \lim_{t \rightarrow \infty} \boldsymbol{\pi}(0) \mathbf{P}(t) = \boldsymbol{\pi}(0) \lim_{t \rightarrow \infty} e^{\mathbf{Q}t} \quad (15.3)$$

- for any $\boldsymbol{\pi}(0)$, which implies

$$\lim_{t \rightarrow \infty} e^{\mathbf{Q}t} = \mathbf{P}(\infty) = \begin{bmatrix} \boldsymbol{\pi}(\infty) & \cdots & \boldsymbol{\pi}(\infty) \end{bmatrix}^T \quad (15.4)$$

- If this limit exists, we call $\mathbf{P}(\infty)$ the **limiting matrix**, and $\boldsymbol{\pi}(\infty)$ the **limiting distribution**.
- $\mathbf{P}(\infty) = \begin{bmatrix} \boldsymbol{\pi}(\infty) & \cdots & \boldsymbol{\pi}(\infty) \end{bmatrix}^T$ does not exist if the CTMC has more than one irreducible closed set (each ICS will converge to a diagonal block, and $\boldsymbol{\pi}(\infty)$ will depend on $\boldsymbol{\pi}(0)$).

15.2 Stationary Distribution

- We have: $\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0) e^{\mathbf{Q}t}, t \geq 0$.
- In steady state the probabilities do not change. We look for a probability vector $\boldsymbol{\pi} = \boldsymbol{\pi}(t_1)$ satisfying: $\boldsymbol{\pi}(t_1) e^{\mathbf{Q}t} = \boldsymbol{\pi}(t_1)$. In other words, for $t \geq t_1$ the probability vector reach the steady state $\boldsymbol{\pi}$, and do not change anymore. Thus:

$$\boldsymbol{\pi} \frac{\partial e^{\mathbf{Q}t}}{\partial t} = \boldsymbol{\pi} \mathbf{Q} e^{\mathbf{Q}t} = \mathbf{0} \quad (15.5)$$

- and we obtain that the **stationary distribution $\boldsymbol{\pi}$** can be **computed with the Global balance equations**:

$$\begin{aligned} \boldsymbol{\pi} \mathbf{Q} &= \mathbf{0} \\ \boldsymbol{\pi} \mathbf{e} &= 1, \mathbf{e}^T = (1, 1, \dots) \end{aligned} \quad (15.6)$$

- **NOTE:** Compare with DTMC $\boldsymbol{\pi} = \boldsymbol{\pi} \mathbf{P}, \boldsymbol{\pi} \mathbf{e} = 1$.

15.3 Numerical Solution

- **Replace one equation method:**

$$\begin{aligned} \boldsymbol{\pi} \mathbf{Q} &= \mathbf{0} \\ \boldsymbol{\pi} \mathbf{e} &= 1, \mathbf{e}^T = (1, 1, \dots) \end{aligned} \quad (15.7)$$

- We solve the equation $\boldsymbol{\pi} \mathbf{Q} = \mathbf{0}$ replacing the last equation by $\boldsymbol{\pi} \mathbf{e} = 1$:

$$\boldsymbol{\pi} \begin{bmatrix} q_{11} & q_{12} & \cdots & q_{1n-1} & 1 \\ q_{21} & q_{22} & \cdots & q_{2n-1} & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ q_{n1} & q_{n2} & \cdots & q_{nn-1} & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \quad (15.8)$$

- **Replace one equation method:**

$$\mathbf{Q} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \quad \begin{aligned} \boldsymbol{\pi} \mathbf{Q} &= \mathbf{0} \\ \boldsymbol{\pi} \mathbf{e} &= 1 \end{aligned}$$

- Solving with **octave** (matlab clone):

```
octave:1> Q=[ -2,1,1;1,-2,1;1,1,-2];
octave:2> s=size(Q,1); # number of rows.
octave:3> [zeros(1,s-1),1] / ...
> [Q(1:s,1:s-1), ones(s,1)]
ans =
0.33333 0.33333 0.33333
```

- With **R**

```
> Q <- matrix(nc=3, byr=T, c
(-2,1,1,1,-2,1,1,1,-2))
> s <- nrow(Q)
> solve(t(cbind(Q[,1:(s-1)], rep(1,s))), c(rep
(0,s-1),1))
[1] 0.3333333 0.3333333 0.3333333
```

15.4 Global balance equations

- Why are they called Global balance equations?

$$\left. \begin{aligned} \pi Q = 0 &\Rightarrow \sum_{i=0}^{\infty} \pi_i q_{ij} = 0 \\ \sum_{i=0}^{\infty} q_{ji} = 0 &\Rightarrow \pi_j \sum_{i=0}^{\infty} q_{ji} = 0 \end{aligned} \right\} \Rightarrow \pi_j \sum_{i=0}^{\infty} q_{ji} = \sum_{i=0}^{\infty} \pi_i q_{ij} \quad (15.9)$$

$$\sum_{i=0}^{\infty} \pi_i q_{ij} \Rightarrow \text{Frequency of transitions entering state } j$$

$$\pi_j \sum_{i=0}^{\infty} q_{ji} \Rightarrow \text{Frequency of transitions leaving state } j \quad (15.10)$$

- In **stationary regime**, the frequency of transitions leaving state j is equal to the frequency of transitions entering state j .

15.4.1 Solving using flux balancing

- Define the **flux** F_{uv} from state u to v :

$$F_{uv} = \pi_u q_{uv} \quad (15.11)$$

- and the flux from set of states U to V :

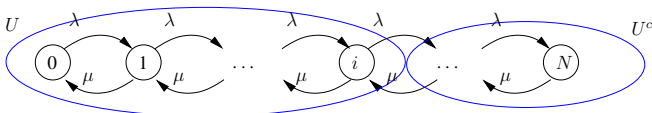
$$F(U,V) = \sum_{u \in U} \sum_{v \in V} F_{uv} \quad (15.12)$$

- From the Global balance equations, and reasoning exactly as in DTMC:

$$F(U,U^c) = F(U^c,U) \quad (15.13)$$

- **NOTE:** Same equation as in DTMC, changing p_{ij} by q_{ij} .

Example: Birth-dead Process



- Flux balancing $\Rightarrow \lambda \pi_i = \mu \pi_{i+1}$
- Iterating:

$$\pi_i = \pi_0 \rho^i, i = 0, 1, \dots, N-1, \rho = \frac{\lambda}{\mu} \quad (15.14)$$

- Normalizing:

$$\pi_0 = \frac{1-\rho}{1-\rho^N} \quad (15.15)$$

15.5 Ergodic Chains

- **Ergodic state:** positive recurrent ($f_{ii}^e = 1, m_{ii}^e < \infty$).
- **Ergodic chain** if all states are ergodic.
- **Theorem:** All states of an irreducible Markov chain are of the same type: Transient or positive/null recurrent (see [2, chapter XV]).
- **Consequences:**

- **Finite irreducible chains are ergodic** (since all states are positive recurrent).
- **Infinite irreducible chains can be:**
 - * **Ergodic:** all the states are positive recurrent (stable chains).
 - * **Non ergodic:** all states are null recurrent or transient (unstable chains).

Theorems for ergodic chains

- $\pi = \pi(\infty)$. Proof: $\pi(\infty)$ satisfies the GBE.
- In stationary regime (when $\pi = \pi e^{Q^t}$), the **mean number of time the system remains in state j** during T time units is given by

$$T \pi_j \quad (15.16)$$

thus, π_j is the fraction of time the chain remains in state j . The proof is analogous to DTMC.

- **NOTE:** The relation of DTMC between **mean recurrence time** and stationary probabilities does not hold for CTMC. I.e., the mean number of time units between two consecutive visits to state j , m_{jj} , **cannot be computed as $1/\pi_j$** . It must be computed with the **recursive equations** (slide 21).

15.6 Reversible Chains

- Let $X(t)$ be an **ergodic MC**. The chain $X^r(t) = X(-t)$ is referred to as the **time reversal chain** of $X(t)$.
- The same results obtained for DTMC reversed chains apply to CTMC, changing p_{ij} by q_{ij} :
 - The reversed chain transition rates q_{ij}^r , given by:

$$\pi_i q_{ij} = \pi_j q_{ji}^r \quad (15.17)$$

satisfy the **reversed balance equations**: $F(U,V) = F^r(V,U)$

- A chain is **reversible** if:

$$q_{ij} = q_{ji}^r \quad (15.18)$$

- Reversible chains satisfy the **detailed balance equations**:

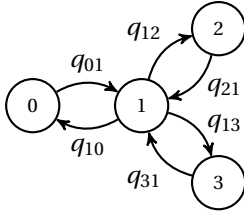
$$F(U, V) = F(V, U), \forall (V, U), V \cap U = \emptyset \quad (15.19)$$

- The same results obtained for DTMC Reversible Chains apply to CTMC: **Kolmogorov Criteria** and **Product Form Solution** for the stationary distribution (changing p_{ij} by q_{ij}).
- E.g. the **stationary probabilities** are given by:
 - Choose a state $\mathbf{s} \in S$,
 - For every other state $\mathbf{i} \in S$, $\mathbf{i} \neq \mathbf{s}$ look for a possible path l_i from state \mathbf{s} to state \mathbf{i} :

$$\mathbf{s} = (l_i, 1) \rightsquigarrow (l_i, 2) \rightsquigarrow \dots \rightsquigarrow (l_i, m_{l_i}) = \mathbf{i}, m_{l_i} \geq 1$$

$$\pi_i = \frac{\psi_i}{\sum_{j \in S} \psi_j}, i \in S \quad \text{where } \psi_i = \begin{cases} 1, & i = s \\ \prod_{k=1}^{m_{l_i}-1} \frac{q_{(l_i, k)(l_i, k+1)}}{q_{(l_i, k+1)(l_i, k)}}, & i \neq s \end{cases} \quad (15.20)$$

15.6.1 Example



- An ergodic tree is always reversible, thus

$$\pi_0 = \frac{1}{G}, \pi_1 = \frac{1}{G} \frac{q_{01}}{q_{10}}, \pi_2 = \frac{1}{G} \frac{q_{01}}{q_{10}} \frac{q_{12}}{q_{21}}, \pi_3 = \frac{1}{G} \frac{q_{01}}{q_{10}} \frac{q_{13}}{q_{31}}. \quad (15.21)$$

- Normalizing:

$$G = 1 + \frac{q_{01}}{q_{10}} + \frac{q_{01}}{q_{10}} \frac{q_{12}}{q_{21}} + \frac{q_{01}}{q_{10}} \frac{q_{13}}{q_{31}} \quad (15.22)$$

Chapter 16

Semi-Markov Process

- Define the continuous RV H_i equal to the **sojourn time** in state i .
- In a **semi-Markov process** we leave the H_i distribution to be **arbitrary**. If H_i is exponentially distributed, we have a CTMC.
- **NOTE: If H_i is not exponentially distributed**, considering only the current state does not satisfy the Markov property (memoryless) since **the evolution of the process depends on the current state and the sojourn time in the state**: (i, t_i) .
- If we consider (i, t_i) as the state, the state would satisfy the Markov property, but we would have a **Markov process** (since t_i is not a discrete RV).

16.1 Embedded MC (EMC) of a semi-Markov process

- **Embedded MC (EMC) of the process: We only look at the state transition instants.**

- **The EMC is a DTMC** with transition probabilities p_{ij}^e .
- The **time step is variable**.
- There are not self transitions ($p_{ii}^e = 0$), unless we look at some memoryless event that produce a self transition.
- **Theorem:** let π_i^e and π_i be the stationary distribution of the EMC and the semi-Markov process respectively. Let $E[H_i]$ be the mean sojourn time in state i , then:

$$\pi_i = \frac{\pi_i^e E[H_i]}{\sum_j \pi_j^e E[H_j]} \quad (16.1)$$

NOTE: By *stationary distribution* for the semi-Markov process we mean to the long-run proportion of time that the process is in each state.

Proof.

- **For n steps of the EMC**, define:

- $f_i(n)$: **proportion of time** the process is in state i .
- $N_i(n)$: **number of visits** to state i .
- $H_i(l)$: **sojourn time** in state i in the **visit number** l .

$$f_i(n) = \frac{\sum_{l=1}^{N_i(n)} H_i(l)}{\sum_j \sum_{l=1}^{N_j(n)} H_j(l)} = \frac{\frac{N_i(n)}{n} \sum_{l=1}^{N_i(n)} \frac{H_i(l)}{N_i(n)}}{\sum_j \frac{N_j(n)}{n} \sum_{l=1}^{N_j(n)} \frac{H_j(l)}{N_j(n)}} \Rightarrow \pi_i = \frac{\pi_i^e E[H_i]}{\sum_j \pi_j^e E[H_j]}$$

- since:

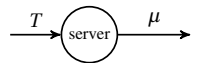
$$\lim_{n \rightarrow \infty} f_i(n) = \pi_i$$

$$\lim_{n \rightarrow \infty} \sum_{l=1}^{N_i(n)} \frac{H_j(l)}{N_i(n)} = E[H_i]$$

$$\lim_{n \rightarrow \infty} \frac{N_i(n)}{n} = \pi_i^e. \quad \square$$

16.1.1 Example

Suppose the system:



- Packets arrive deterministically every T time units.
- Upon a packet arrival it goes immediately into service if the server is empty, and it is lost if the server is busy.
- Services are exponentially distributed with rate μ .

Define a semi-Markov process with states $\begin{cases} 0 & \text{server empty,} \\ 1 & \text{server busy.} \end{cases}$ (16.2)

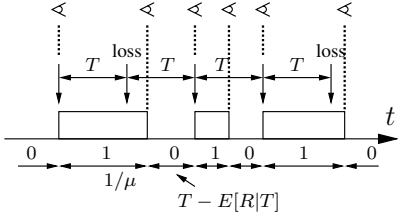
1. Derive the EMC and stationary distribution of the EMC and continuous time process.
2. Compute the throughput and loss probability.

Hint: The distribution of an event R exponentially distributed with rate μ , given that occurs in an interval $t \in [0, T]$, is

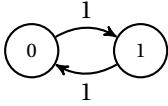
$$F_R(t|T) = \frac{P(R \leq t, R \leq T)}{P(R \leq T)} = \frac{P(R \leq t)_{t \in [0, T]}}{P(R \leq T)} = \frac{1 - e^{-\mu t}}{1 - e^{-\mu T}}, \quad t \in [0, T], \text{ and}$$

$$E[R|T] = \int_0^T (1 - F_R(t|T)) dt = \frac{1 - \alpha - \alpha \mu T}{\mu(1 - \alpha)}, \text{ where } \alpha = e^{-\mu T}.$$

Solution



EMC:



- $\pi_0^e = \pi_1^e = 1/2$,
- $E[H_0] = T - E[R|T] = \frac{T\mu - (1 - \alpha)}{\mu(1 - \alpha)}, E[H_1] = \frac{1}{\mu}.$

And the continuous time process:

$$G = \pi_0^e E[H_0] + \pi_1^e E[H_1] = \frac{T}{2(\alpha - 1)} \left[\frac{\text{time units}}{\text{step}} \right] \quad (16.3)$$

$$\pi_0 = \frac{\pi_0^e E[H_0]}{G} = \frac{1 - \alpha}{\mu T}, \pi_1 = \frac{\pi_1^e E[H_1]}{G} = \frac{\mu T - (1 - \alpha)}{\mu T}$$

$$\text{Throughput: } S = \mu \pi_1 = \frac{1 - \alpha}{T} \quad (\text{check } S = \frac{1}{E[H_0] + E[H_1]})$$

$$\text{Loss probability: } S = \frac{1}{T} (1 - p_L), p_L = 1 - S T = \alpha = e^{-\mu T}. \quad (16.4)$$

16.2 Embedded MC of a CTMC

- Assume that the semi-Markov process is a **CTMC** (sojourn times are exponentially distributed).
- The transition probabilities p_{ij}^e of the EMC, and the stationary distribution π_i of the CTMC are given by:

$$\begin{cases} p_{ij}^e = q_{ij}/q_i, & i \neq j \\ p_{ij}^e = 0, & i = j \end{cases} \quad \pi_i = \frac{\pi_i^e/q_i}{\sum_k \pi_k^e/q_k} \quad (16.5)$$

$$\text{where: } q_i = \sum_{k \neq i} q_{ik} = -q_{ii}.$$

Proof.

- The equations for p_{ij}^e was proven in a previous section.

- The distribution of the **sojourn time in state i in the CTMC** is the distribution of the **minimum of independent exponentially distributed RV** with parameters $\{q_{ik}\}_{k \neq i}$. This distribution is **exponentially distributed with parameter $\sum_{k \neq i} q_{ik}$** . Thus $E[H_i] = 1/\sum_{k \neq i} q_{ik} = 1/q_i$, $q_i = \sum_{k \neq i} q_{ik}$.
- Substituting:

$$\pi_i = \frac{\pi_i^e E[H_i]}{\sum_j \pi_j^e E[H_j]} = \frac{\pi_i^e/q_i}{\sum_j \pi_j^e/q_j} \quad (16.6)$$

$$\text{where: } q_i = \sum_{k \neq i} q_{ik} = -q_{ii}. \quad \square$$

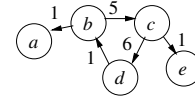
Chapter 17

Finite Absorbing Chains

17.1 Canonical Form

- Let $\mathbf{Q}^{r \times r}$ be the transition probability matrix of a chain with a set S of s **transient** states and a set A of $r - s$ **absorbing** states. We can write $\mathbf{Q}^{r \times r}$ in the **canonical** form:

$$\mathbf{Q}^{r \times r} = \begin{bmatrix} \mathbf{T}^{s \times s} & \mathbf{R}^{s \times r-s} \\ \mathbf{0}^{r-s \times s} & \mathbf{0}^{r-s \times r-s} \end{bmatrix} \quad (17.1)$$



$$\mathbf{P} = \begin{array}{c} \begin{matrix} b & c & d & a & e \end{matrix} \\ \begin{matrix} b \\ c \\ d \\ a \\ e \end{matrix} \left[\begin{array}{ccccc} -6 & 5 & 0 & 1 & 0 \\ 0 & -7 & 6 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right] \end{array} \quad (17.2)$$

17.2 Results

- Define:

$$\mathbf{n}_{ij} = \left\{ \begin{array}{l} \text{time in state } j \text{ before absorption,} \\ \text{starting from state } i \end{array} \right\}, \quad (17.3)$$

$$\mathbf{t}_i = \left\{ \begin{array}{l} \text{time in transient states before} \\ \text{absorption, starting from state } i \end{array} \right\},$$

$$\mathbf{b}_{ij} = P(\text{probability to be absorbed } j \text{ starting } i).$$

- Then:

$$\begin{aligned} \{E[n_{ij}]\} &= \mathbf{N} = -\mathbf{T}^{-1} \\ \{E[t_i]\} &= \mathbf{\tau} = \mathbf{N}\mathbf{e} \\ \{b_{ij}\} &= \mathbf{B} = \mathbf{N}\mathbf{R} \end{aligned} \quad (17.4)$$

where $\{a_{ij}\}$ is a matrix with a_{ij} as element ij and \mathbf{e} is a column vector of 1s. \mathbf{N} is called the **fundamental matrix**.

Proof.

Let S be the set of **transient** states and A the set of **absorbing** states.

- $\{\mathbf{E}[n_{ij}]\} = \mathbf{N} = -\mathbf{T}^{-1}$

$$\begin{aligned} \mathbf{E}[n_{ij}] &= \frac{\delta_{ij}}{-q_{ii}} + \sum_{\substack{k \neq i \\ k \in S}} p_{ik}^e \mathbf{E}[n_{kj}] = \\ &= \frac{\delta_{ij}}{-q_{ii}} + \sum_{\substack{k \neq i \\ k \in S}} \frac{q_{ik}}{-q_{ii} - q_{ii}} \mathbf{E}[n_{kj}] \Rightarrow \\ &= - \sum_{k \in S} q_{ik} \mathbf{E}[n_{kj}] = \delta_{ij} \Rightarrow \\ &= -\mathbf{T} \{\mathbf{E}[n_{ij}]\} = \mathbf{I} \Rightarrow \{\mathbf{E}[\mathbf{n}_{ij}]\} = -\mathbf{T}^{-1} = \mathbf{N} \end{aligned}$$

- $\{\mathbf{E}[t_i]\} = \boldsymbol{\tau} = \mathbf{N}\mathbf{e}$

$$\mathbf{E}[t_i] = \sum_{k \in S} \mathbf{E}[n_{ik}] \Rightarrow \{\mathbf{E}[\mathbf{t}_i]\} = \boldsymbol{\tau} = \mathbf{N}\mathbf{e} \quad (17.5)$$

- $\{b_{ij}\} = \mathbf{B} = \mathbf{N}\mathbf{R}$

$$\begin{aligned} b_{ij} &= p_{ij}^e + \sum_{\substack{k \neq i \\ k \in S}} p_{ik}^e b_{kj} = \\ &= \frac{q_{ij}}{-q_{ii}} + \sum_{\substack{k \neq i \\ k \in S}} \frac{q_{ik}}{-q_{ii} - q_{ii}} b_{kj}, i \in S, j \in \mathcal{A} \Rightarrow \\ &= - \sum_{k \in S} q_{ik} b_{kj} = q_{ij} \Rightarrow -\mathbf{T} \{b_{ij}\} = \mathbf{R} \Rightarrow \\ &= \{\mathbf{b}_{ij}\} = \mathbf{B} = -\mathbf{T}^{-1} \mathbf{R} = \mathbf{N}\mathbf{R}. \quad \square \end{aligned}$$

17.3 Extension of the Results

- The previous results can be generalized to any group of states of \mathbf{Q} :
- A set S is referred to as **open** if the chain can reach some state of S^c starting from any state of S . Let

$$\begin{aligned} \mathbf{T} &= \{q_{ij}, i \in S, j \in S\} \\ \mathbf{R} &= \{q_{ij}, i \in S, j \in S^c\} \end{aligned} \quad (17.6)$$

Let assume that the process starts from $i \in S$. Define:

$$\begin{aligned} \mathbf{n}_{ij} &= \left\{ \begin{array}{l} \text{number of steps in state } j \text{ before} \\ \text{leaving } S, \text{ starting from state } i \end{array} \right\}, \\ &\Rightarrow \boxed{\{\mathbf{E}[n_{ij}]\} = \mathbf{N} = -\mathbf{T}^{-1}}. \end{aligned} \quad (17.7)$$

- Similarly for the other results, e.g. $\boldsymbol{\tau} = \{\mathbf{E}[t_i]\} = \mathbf{N}\mathbf{e}$ and $\mathbf{B} = \{b_{ij}\} = \mathbf{N}\mathbf{R}$.

Bibliography

- [1] Llorenç Cerdà-Alabern. *Transient Solution of Markov Chains Using the Uniformized Vandermonde Method*. Tech. rep. UPC-DAC-RR-XCSD-2010-2. Universitat Politècnica de Catalunya, Dec. 2010. URL: https://www.ac.upc.edu/app/research-reports/html/research_center_index-XCSD-2010,en.html.
- [2] William Feller. *An Introduction to Probability Theory and Its Applications: Volume One*. 3rd ed. John Wiley & Sons, 1968.
- [3] Mor Harchol-Balter. *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge University Press, 2013.
- [4] John G Kemeny and James Laurie Snell. *Finite markov chains*. Vol. 356. van Nostrand Princeton, NJ, 1960.
- [5] Leonard Kleinrock and Simon Lam. “**Packet Switching in a Multiaccess Broadcast Channel: Performance Evaluation**”. In: *Communications, IEEE Transactions on* 23.4 (1975), pp. 410–423.
- [6] Randolph Nelson. *Probability, stochastic processes, and queueing theory: the mathematics of computer performance modeling*. Springer, 1995.
- [7] Kishor S Trivedi. *Probability & statistics with reliability, queueing and computer science applications*. John Wiley & Sons, 2008.

Table with some distributions

Distribution	Parametres	Density	Mean	Variance	Characteristic Function
Bernoulli	$0 \leq p \leq 1$ $q = 1 - p$	$p^k (1 - p)^{1-k}$ $k = 0, 1$	p	$p(1 - p)$	$q + pe^{it}$
Binomial	$0 \leq p \leq 1$ $q = 1 - p$	$\binom{n}{k} p^k (1 - p)^{n-k}$ $k = 0, 1, \dots, n$	np	$np(1 - p)$	$(q + pe^{it})^n$
Geometric	$0 \leq p \leq 1$ $q = 1 - p$	$p(1 - p)^k$ $k \geq 0$	$\frac{1 - p}{p}$	$\frac{1 - p}{p^2}$	$\frac{p}{1 - qe^{it}}$
Negative binomial	$r > 0$ $0 \leq p \leq 1$ $q = 1 - p$	$\binom{k+r-1}{k} p^r q^k$ $k \geq 0$	$r \frac{1 - p}{p}$	$r \frac{1 - p}{p^2}$	$\left(\frac{p}{1 - qe^{it}} \right)^r$
Poisson	$\lambda > 0$	$\frac{\lambda^k}{k!} e^{-\lambda}, \quad k \geq 0$	λ	λ	$\exp \{ \lambda (e^{it} - 1) \}$
Normal $N(\mu, \sigma)$	$\mu \in \mathbb{R}$ $\sigma > 0$	$\frac{1}{\sqrt{2\pi}\sigma^2} \exp \left\{ \frac{-(x - \mu)^2}{2\sigma^2} \right\}$ $x \in \mathbb{R}$	μ	σ^2	$\exp \left\{ \mu it - \frac{t^2 \sigma^2}{2} \right\}$
Uniform	$a < b$	$\frac{1}{b - a}, \quad a \leq x \leq b$	$\frac{a + b}{2}$	$\frac{(b - a)^2}{12}$	$\frac{e^{itb} - e^{ita}}{it(b - a)}$
Exponential	α	$\alpha e^{-\alpha x}, \quad x \geq 0$	$\frac{1}{\alpha}$	$\frac{1}{\alpha^2}$	$\left(1 - \frac{it}{\alpha} \right)^{-1}$
Gamma $\gamma(n, \alpha)$	$\alpha > 0,$ $n > 0$	$\frac{\alpha^n x^{n-1} e^{-\alpha x}}{\Gamma(n)}, \quad x \geq 0$	$\frac{n}{\alpha}$	$\frac{n}{\alpha^2}$	$\left(1 - \frac{it}{\alpha} \right)^{-n}$
Beta $\beta(p, q)$	$p > 0,$ $q > 0$	$\frac{x^{p-1} (1 - x)^{q-1}}{B(p, q)},$ $0 \leq x \leq 1$	$\frac{p}{p + q}$	$\frac{pq}{(p + q)^2 (p + q + 1)}$	

$$\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt, \Gamma(n) = (n - 1)! \quad B(p, q) = \int_0^1 t^{p-1} (1 - t)^{q-1} dt, B(p, q) = \frac{\Gamma(p) \Gamma(q)}{\Gamma(p + q)}$$