Markov chains

Brief summary

David Manresa

Contents

Chapter 1

Brief introduction to Stochastic Processes

1.1 Introduction

Definition 1.1. ***Stochastic Process***
Let $(\Omega, \mathcal{F}, P)$ be a probability space, a stochastic process with set of parameters T and state space $(E, \mathcal{E})$ is a set of random variables $\{X_t\}_{t\in T}$

$$X_t : (\Omega, \mathcal{F}) \to (E, \mathcal{E}).$$

Therefore, for each $t \in T$, X_t is a measurable function, that is, $X_t^{-1}(B) \in \mathcal{F},\ \forall B \in \mathcal{E}$.

The most common approach is to interpret t as time and $E = \mathbb{R}$. We define X_t as a function from $(\Omega, \mathcal{F})$ to $(\mathbb{R}, \mathbb{B})$, where for each t, X_t is a random variable that describes the state of the system.

Definition 1.2. ***Sample path/Trajectory***
Let $\{X_t\}_{t\in T}$ be a stochastic process, then for each $\omega \in \Omega$, we obtain a function called sample path or trajectory

$$X(\cdot, \omega) : T \to E,$$

that represents the evolution of the process when the elementary event $\omega \in \Omega$ occurs.

If we denote E^T the set of all functions from T to E, then the stochastic process establishes an application

$$\mathcal{X} : \Omega \to E^T,$$

which associates each elementary event $\omega \in \Omega$ to $X(\cdot, \omega)$. We can construct the σ-algebra $\mathcal{E}^T$ engendered by the sets of the form

$$\{f \in E^T | f(t) \in B\}, \quad t \in T, \mathbb{B} \in \mathcal{E}.$$

Definition 1.3. ***Distribution of a stochastic process***
Let $\{X_t\}_{t\in T}$ be a stochastic process, the distribution of the process is the probability measure

$$P(A) = P(\mathcal{X}^{-1}(A)) \ \ \forall A \in \mathcal{E}^T.$$

We can also consider the **finite-dimensional distribution** by considering $t_1 \neq \ldots \neq t_k$ and

$$P_{t_1,\ldots,t_k}(B) = P((X_{t_1},\ldots,X_{t_k}) \in B),\ B \in \mathcal{E}^k.$$

Definition 1.4. ***Filtration***
A filtration is a family of sub-σ-algebras $\{\mathcal{F}_t\}_{t\in T}$ on Ω such that $\mathcal{F}_s \subset F_t \subset F$, whenever $0 \leq s < t$.

Definition 1.5. ***Stopping time***
Let τ be a random variable defined on the filtered probability space $(\Omega, \mathcal{F}, \{F_t\}_{t\in T}, P)$. Then τ is a stopping time if

$$\{\tau \leq t\} \in \mathcal{F}_t,\ \forall t \in T.$$

Intuitively, stopping times are those moments at which one can determine, by observing the evolution of the system up to time t, whether the stopping event has occurred. In other words, at any given time t, we can definitively say whether the stopping time τ has happened within the interval $[0, t]$.

Definition 1.6. ***Markov property***
Let $\{X_t\}_{t\in T}$ be a stochastic process with space of parameters $T = \mathbb{R}$ and state spaces $(E, \mathcal{E})$ defined in a probability space $(\Omega, \mathcal{F}, P)$. Then $\{X_t\}_{t\in T}$ satisfies the Markov property with respect of the filtration $\{\mathcal{F}_t\}_{t\in T}$ if

$$P(X_t \in B \mid \mathcal{F}_s) = P(X_t \in B \mid X_s),\ \forall s \leq t \in, B \in \mathcal{E}.$$

Intuitively, the probability of the future event $\{X_t \in B\}$ remains the same whether we only know the current state X_s or if we know the entire history up to time s, $\mathcal{F}_s$. Therefore, the entire evolution of the process up to time s is encapsulated by X_s.

1.2 Exercises

1. An athlete is going to run along a straight 100-meter track at a constant speed of 10 m/s. A spectator chooses a random point in the square stands and wonders about the distance that will separate them from the runner at any given moment. Model the problem with a stochastic process.

> Let $\Omega = \{(a, b) \in [0, 100] \times [0, 100]\}$, $\mathcal{F} = \mathbb{B}^2_\Omega$ and $P(B) = \lambda(B)/\lambda(\Omega)$, where λ is the Lebesgue measure on $\mathbb{B}^2_\Omega$. Then, when the spectator chooses the random point $\omega \in \Omega$, the distance between him and the athlete on each t is
>
> $$X_t(\omega) = \sqrt{(10t - a)^2 + b^2},\ t \in [0, 10].$$
>
> Therefore, $\{X_t\}_{t\in T}$ is a stochastic process on $(\Omega, \mathcal{F}, P)$ with set of parameters $T = [0, 10]$.
>
> Notice that the trajectories $X(\omega, t) : T \to \mathbb{R}$ are curves which represent the distance on each t after the spectator chooses the random point $(a, b) \in \Omega$ to watch the race.

Chapter 2

Discrete Markov chains

2.1 Introduction

Definition 2.1. ***Discrete Markov chain***
A discrete Markov chain is a stochastic process $\{X_t\}_{t\in T}$ that satisfies the Markov property, $T = \mathbb{N}$ and the space state E is discrete.

Notice that the Markov property for discrete Markov chains is

$$P(X_{n+1} = j \mid X_0 = i_0, \ldots, X_n = i_n) = P(X_{n+1} = j \mid X_n = i_n), \quad \forall n \in \mathbb{N},\ \forall i_0, \ldots, i_n, j \in E.$$

2.2 Transition probabilities

In a Markov chain, X_{n+1} only depends on X_n, so we denote the transition probabilities as

$$P(X_{n+1} = j \mid X_n = i) = p_{i,j}(n),$$

which can depend on i, j and n.

If we dote the state space E of an arbitrary order, then this transition probabilities can be expressed in a matrix called **transition matrix**

$$P(n) = (p_{i,j}(n))_{i,j\in E}.$$

The transition matrices are stochastic matrices, in a sense that

$$p_{i,j} \leq 0,\ \forall i, j \in E \qquad\qquad \text{and} \qquad\qquad \sum_{j\in E} p_{i,j}(n) = 1,\ \forall i \in E,$$

since the probabilities must be positive and if the process is in the state i, naturally, the probability of being in some of the other states $j \in E$ in the next step is 1.

Definition 2.2. ***Stationary Markov chain***
Let $\{X_t\}_{t\in T}$ be a Markov chain, if the transition matrices are all the same for all $n \in \mathbb{N}$ then the Markov chain is stationary.
In that case, the transition Matrix is unique and can be denoted as $P = (p_{i,j})_{i,j\in E}$.

Naturally, Markov chains are incomplete without specifying how the process initiates. We can do that by creating a initial random variable X_0 and atribute the probabilities for each state $i \in E$.

$$P(X_0 = i) = p_i, \ \forall i \in E.$$

The vector $p = (p_i)_{i \in E}$ is called **initial distribution**.

2.3 Marginal distributions

If we denote the marginal distribution of X_n by

$$P(X_n = j) = p_j^{(n)},$$

then

$$P(X_n = j) = \sum_{i \in E} P(X_{n-1} = i) P(X_n = j \mid X_{n-1} = i)$$

which can be expressed matricially as

$$p^{(n)} = p^{(n-1)} P,$$

therefore

$$p^{(n)} = p^{(n-1)} P = p^{(n-2)} P^2 = \cdots = pP^n.$$

Proposition 2.1. ***Marginal distribution***
Let $\{X_n\}_{n \in N}$ be a Markov chain, then the marginal distribution of X_n is pP^n, where p is the initial distribution and P the transition matrix.

2.4 Transition probabilities in n steps

If we are interested in computing the transition probabilities in n steps

$$P(X_{m+n} = j \mid X_n = i) = p_{i,j}^{(n)},$$

by the law of total probabilities

$$P(X_{m+r+n} = j \mid X_m = i) = \sum_{k \in E} P(X_{m+r} = k \mid X_m = i) P(X_{m+r+n} = j \mid X_{m+r} = k),$$

and by the Markov property

$$p_{i,j}^{(r+n)} = \sum_{k \in E} p_{i,k}^{(r)} p_{k,j}^{(n)}, \quad \text{or matricially} \quad P^{(n+r)} = P^{(r)} P^{(n)}.$$

Then,

$$P^{(n)} = PP^{(n-1)} = \cdots = P^n.$$

Proposition 2.2. ***Transition matrix in n steps***

Let $\{X_n\}_{n\in\mathbb{N}}$ be a Markov chain, then the transition matrix in n steps satisfies

$$P^{(n)} = P^n.$$

In practice, it is useful to diagonalize the transition matrix P in order to easily compute P^n since if $P = HJH^{-1}$, then $P^n = HJ^nH^{-1}$.

2.5 Distribution of arrival times

The probabilities $p_{i,j}^{(n)} = P_i(X_n = j)$ express the probability that after n steps, the process is on the state j if it starts on i, but they do not distinct if the process has achieved the state j before or not. We define the arrival time

$$\tau_j = \min(n > 0 \mid X_n = j),$$

and we are interested in computing the first arrival probabilities

$$f_{i,j}^{(n)} = P_i(\tau_j = n).$$

We can determine this probabilities by modifying the Markov chain and making the state j absorbent, that is, we define a new Markov chain

$$\tilde{X}_n = \begin{cases} X_n & \text{if } X_r \neq j \ \forall r < n, \\ j & \text{if } X_r = j \text{ for some } r < n. \end{cases}.$$

That way, the new Markov chain stays on the state j forever after reaching it. Notice that then

$$P_i(\tilde{X}_n = j) = P_i(\tau_j \leq n),$$

Therefore, if we denote $P_i(\tilde{X}_n = j) = \tilde{p}_{i,j}^{(n)}$ then

$$f_{i,j}^{(n)} = P_i(\tau_j = n) = P_i(\tau_j \leq n) - P_i(\tau_j \leq n-1) = \tilde{p}_{i,j}^{(n)} - \tilde{p}_{i,j}^{(n-1)}.$$

Proposition 2.3.
The probability of the first arrival time on the state j departing from i, occurring on step n is

$$f_{i,j}^{(n)} = \tilde{p}_{i,j}^{(n)} - \tilde{p}_{i,j}^{(n-1)},$$

where $\tilde{p}_{i,j}^{(n)}$ is the transition probability of the modified Markov chain after making the state j absorbent.

We can obtain a relation between probabilities of the first arrival time and transition probabilities as

$$p_{i,j}^{(n)} = \sum_{m=1}^{n} P_i(\tau_j = m) P_i(X_n = j \mid \tau_j = m) = \sum_{m=1}^{n} P_i(\tau_j = m) P_j(X_{n-m} = j) = \sum_{m=1}^{n} f_{i,j}^{(n)} p_{j,j}^{(n-m)}.$$

Furthermore, if we are interested on the total probability of arriving to the state j from i,

$f_{i,j} = P_i(\tau_j < \infty)$, we can decompose the expression as a function of X_1,

$$f_{i,j} = P_i(X_1 = j) + \sum_{k \neq j} P_i(X_1 = k) P_i(\tau_j < \infty \mid X_1 = k) = p_{i,j} + \sum_{k \neq j} p_{i,k} f_{k,j}.$$

Proposition 2.4. ***Total arrival probabilities***
The total arrival probabilities $f_{i,j} = P_i(\tau_j < \infty)$ can be computed solving the linear system of equations

$$\begin{cases} f_{i,j} = p_{i,j} + \sum_{k \neq j} p_{i,k} f_{k,j}, \\ \forall i \neq j. \end{cases}$$

2.6 Mean arrival times

If $f_{i,j} = 1$ for some $i, j \in E$, the mean arrival times are

$$e_{i,j} = E_i[\tau_j] = \sum_{n=1}^{\infty} n f_{i,j}^{(n)},$$

which, after some algebra, we obtain

$$e_{i,j} = 1 + \sum_{k \neq j} p_{i,k} e_{k,j}.$$

Proposition 2.5. ***Mean arrival times***
The mean arrival times $e_{i,j}$ can be computed solving the linear system of equations

$$\begin{cases} e_{i,j} = 1 + \sum_{k \neq j} p_{i,k} e_{k,j}, \\ \forall i \neq j. \end{cases}$$

2.7 Expected number of visits to a state

If we denote V_j as the total number of visits to the state j, the expected number of visits from i to j can be expressed using the indicator function

$$E_i[V_j] = \sum_{n=1}^{\infty} E_i[I_{(X_n = j)}] = \sum_{n=1}^{\infty} p_{i,j}^{(n)}.$$

However, there do exist a relationship with the total probabilities that is more useful in practice.

Proposition 2.6.
The expected number of visits from i to j is

$$E_i[V_j] = \frac{f_{i,j}}{1 - f_{j,j}}.$$

2.8 State classification

Definition 2.3. ***States communication***
Let $i, j \in E$ be states of a Markov chain, then we say that i communicates with j, and we denote it as $i \rightsquigarrow j$ if $p_{i,j}^{(n)} > 0$ for some $n \in \mathbb{N}$.
If $i \rightsquigarrow$ and $j \rightsquigarrow i$, then the states intercommunicate and it is denoted as $i \leftrightsquigarrow j$.

Definition 2.4. ***Essential state***
If $i \rightsquigarrow j$ implies that $j \rightsquigarrow i$, then the state i is essential.
The set of all essential states is denotes $E^\star$.

Notice that the intercommunication $\leftrightsquigarrow$ relation is a equivalence relation in $E^\star$, therefore this set can be decomposed in equivalences classes $E^\star = C_1 \cup C_2 \cup \dots$.

The sets C_r are called closed and irreducible sub-chains of the matrix P, since from states of C_r can only be achieved C_r states, and they do not have sub-chains in it.

Definition 2.5. ***Recurrent and transient states***
A state $i \in E$ is recurrent if $f_{i,i} = 1$, while if $f_{i,i} < 1$ is transient.

Proposition 2.7.
If $i \leftrightsquigarrow j$ then they are both recurrent of transient.

Definition 2.6. ***Positive and negative recurrent states***
A state $i \in E$ is positive recurrent if $e_{i,i} < \infty$ and negative recurrent if $e_{i,i} = \infty$.

Proposition 2.8.
If $i \leftrightsquigarrow j$ then they are both positive or negative recurrent.

If the initial state i is not an absorbing state, then it is likely that absorption into $E^\star$ will occur sooner or later. Let the absorption time be

$$\tau^\star = \min(n > 0 : X_n \in E^\star).$$

Then the absorption probabilities $f_i^\star = P_i(\tau^\star < \infty)$ are

$$f_i^\star = \begin{cases} f_i^\star = \sum_{j \in E^\star} p_{i,j} + \sum_{k \notin E^\star} p_{i,k} f_k^\star, \\ \forall i \notin E^\star. \end{cases}$$

2.9 Stationary measures

Definition 2.7. ***Stationary measure***
A measure $(w_j)_{j \in E}$, where $w_j \in [0, \infty)$ is called stationary if

$$\sum_{k \in E} w_k p_{k,j} = w_j, \quad \forall j \in E,$$

or matricially $wP = w$.

We can obtain a **stationary probability distribution** π by normalizing $(w_j)_{j\in E}$. Therefore, if we choose π as a initial distribution of the chain, then the marginal distribution of X_n is $\pi P^n = \pi$.

Theorem 2.1.
A irreducible Markov chain is positive recurrent if and only if there do exists a stationary distribution π. Furthermore, the stationary distribution satisfies

$$\pi_i = \frac{1}{e_{i,i}}.$$

2.10 Periodicity

If an irreducible chain have the structure $C = D_0 \cup \dots D_{r-1}$, where on each step n the chain can only reach states on $D_{(n+1) \mod r}$, then the chain has a period r.

Definition 2.8. ***Period***
The period of a state $i \in E$ is

$$d = gcd(n \geq 1 : p_{i,i}^{(n)} > 0).$$

If $d = 1$, the state is aperiodic.

Proposition 2.9.
If $i \leftrightsquigarrow j$ then both have the same period.

2.11 Chains conditioned on their final state

We already know that if $n > N$, then

$$P(X_{n+1} = j \mid X_n = i, X_N = k) = p_{i,j},$$

since the probabilities of the next step only depends on the current state on n. However, if $n < N$, then the probabilities are affected and cap be computed as

$$P(X_{n+1} = j \mid X_n = i, X_N = k) = \frac{P(X_{n+1} = j, X_N = k \mid X_n = i)}{P(X_N = k \mid X_n = i)} = \frac{p_{i,j} p_{j,k}^{(N-n-1)}}{p_{i,k}^{(N-n)}},$$

and conditioned by $\tau_k < \infty$ as

$$P(X_{n+1} = j \mid X_n = i, \tau_k < \infty) = \frac{p_{i,j} f_{j,k}}{f_{i,k}}.$$

2.12 Exercises

1. What are the long run proportions of the Markov chain with transition matrix

$$\begin{pmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}$$

The long run proportions of the Markov chain are the π_i such that $\pi P = \pi$, therefore

$$\begin{aligned}
\pi_0 &= 0.5\pi_0 + 0.3\pi_1 + 0.2\pi_2, \\
\pi_1 &= 0.4\pi_0 + 0.4\pi_1 + 0.3\pi_2, \\
\pi_2 &= 0.1\pi_0 + 0.3\pi_1 + 0.5\pi_2.
\end{aligned}$$

and $\pi_0 + \pi_1 + \pi_2 = 1$. The solution is $\pi_0 = \frac{21}{62},\ \pi_1 = \frac{23}{62},\ \pi_2 = \frac{18}{62}$.

2. Three people sitting around a table play as follows: the player whose turn it is throws three coins and wins if they get three heads; if they get two heads, they continue playing; if they get two tails, they pass the turn to the player on their right; and if they get three tails, they pass the turn to the player on their left. Compute:

1. The probability that the person to the right of the starting player wins the game.
2. The mean and distribution of the number of throws made by the person to the left of the starting player.
3. The distribution of the game's duration.

Let $\{X_n\}_{n\in\mathbb{N}}$ be the Markov chain with states $E = \{A, B, C, AW, BW, CW\}$, which denote the turn of the player A, B or C and if each players wins AW, BW or CW. The transition matrix is then

$$P = \begin{array}{c} \\ A \\ B \\ C \\ AW \\ BW \\ CW \end{array} \begin{array}{c} \begin{array}{cccccc} A & B & C & AW & BW & CW \end{array} \\ \begin{pmatrix} 3/8 & 3/8 & 1/8 & 1/8 & & \\ 1/8 & 3/8 & 3/8 & & 1/8 & \\ 3/8 & 1/8 & 1/8 & & & 1/8 \\ & & & 1 & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix} \end{array}$$

where empty spaces denote probability 0. We will consider the player A to be the first.

1.
We have to compute $f_{A,BW}$, then by the formula $f_{i,j} = p_{i,j} + \sum_{k\neq j} p_{i,k} f_{k,j}$, we obtain the system

$$\begin{cases}
f_{A,BW} = \frac{3}{8} f_{A,BW} + \frac{3}{8} f_{B,BW} + \frac{1}{8} f_{C,BW} \\
f_{B,BW} = \frac{1}{8} + \frac{1}{8} f_{A,BW} + \frac{3}{8} f_{B,BW} + \frac{3}{8} f_{C,BW} \\
f_{C,BW} = \frac{3}{8} f_{A,BW} + \frac{1}{8} f_{B,BW} + \frac{3}{8} f_{C,BW}
\end{cases}$$

with solutions $f_{A,BW} = \frac{8}{26}, f_{B,BW} = \frac{11}{26}, f_{C,BW} = \frac{7}{26}$.

Therefore, since A player starts the game, the probability is $f_{A,BW} = \dfrac{8}{26}$.

2.
The number of throws made by C is the number of visits to the state C, then we have to

compute $E_A[V_C]$, which with the formula $E_i[V_j] = \dfrac{f_{i,j}}{1-f_{j,j}}$.

Analogously as the previous exercise section

$$\begin{cases} f_{A,C} = \frac{3}{8} f_{A,C} + \frac{3}{8} f_{B,C} + \frac{1}{8} \\ f_{B,C} = \frac{1}{8} f_{A,C} + \frac{3}{8} f_{B,C} + \frac{3}{8} \\ f_{C,C} = \frac{3}{8} f_{A,C} + \frac{1}{8} f_{B,C} + \frac{3}{8} \end{cases}$$

with solutions $f_{A,C} = \frac{7}{11}, f_{B,C} = \frac{8}{11}, f_{C,C} = \frac{31}{44}$.

Therefore,

$$\begin{aligned} E_A[V_C] &= \frac{f_{A,C}}{1-f_{C,C}} = \frac{28}{13} \\ E_B[V_C] &= \frac{f_{B,C}}{1-f_{C,C}} = \frac{32}{13} \\ E_C[V_C] &= \frac{f_{C,C}}{1-f_{C,C}} = \frac{31}{13} \end{aligned}$$

Thus, the mean of the number of throws made by C is $E_A[V_C] = 31/13$.

The distribution can be computed as

$$P_A(V_C = k) = \begin{cases} \dfrac{4}{11} & \text{if } k = 0, \\ \dfrac{7}{11}\dfrac{31}{44}^{k-1}\dfrac{13}{44} & \text{if } k = 1, 2, \ldots. \end{cases}$$

Since visiting k times the state C means, visiting it one time, returning to it $k-1$ times, and not returning more, with probabilities $f_{A,C}, (f_{C,C})^{k-1}$ and $1 - f_{C,C}$ respectively.

3.

We can now group the states AW, BW, CW to a new state F and A, B, C to ABC, since now we are interested when the Markov chain reaches the state F and the game finishes. The new transition matrix is

$$Q = \begin{array}{c c} & \begin{array}{cc} ABC & F \end{array} \\ \begin{array}{c} ABC \\ F \end{array} & \begin{pmatrix} 7/8 & 1/8 \\ 0 & 1 \end{pmatrix} \end{array},$$

which we can diagonalize and obtain

$$Q = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & \\ & 7/8 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Therefore, if Y is the number of throws till the game finishes its distribution is

$$P(Y \leq n) = q_{J,F}^{(n)} = 1 - (7/8)^n.$$

Chapter 3

Continuous Markov chains

3.1 Introduction

Definition 3.1. ***Continuous Markov chain***
A continuous Markov chain is a stochastic process $\{X_t\}_{t\geq 0}$ *on a probability space* $(\Omega, \mathcal{F}, P)$ *where the state space* E *is discrete, the parameter set is* $T = [0, \infty)$*, and satisfies*

$$P(X_{t_n+1} = t_{n+1} \mid X_1 = i_1, \ldots, X_n = i_n) = P(X_{t_{n+1}} = t_{n+1} \mid X_{t_n} = i_n),$$

for every $0 \leq t_1 < \cdots < t_n$ *and* $i_1, \ldots, i_{n+1} \in E$.

Therefore, we keep the condition that E is discrete, where we will define the σ-algebra $\mathcal{E} = \mathcal{P}(E)$, but the parameter set T is now continuous. Thus, we do not consider steps anymore but continuous time.

The Markov chain is specified by

1. The state space, E.
2. The initial distribution $p_i(0) = P(X_0 = i)$ $\forall i \in E$.
3. The transition probabilities $P_{i,j}(s,t) = P(X_t = j \mid X_s = s)$, $\forall s < t \in [0, \infty)$ and $\forall i, j \in E$.

Proposition 3.1.
Let $\{X_t\}_{t\geq 0}$ *be a Markov chain with state space* E *and transition functions* $\{P_{i,j}(t)\}_{i,j\in E}$ *then*

1. $P_{i,j}(t) \geq 0$ *and* $\sum_{k\in E} P_{i,k}(t) = 1$ $\forall i, j \in E, t \geq 0$.
2. $P_{i,j}(0) = \delta_{i,j}$. *(Kronecker delta).*
3. $P_{i,j}(t+s) = \sum_{k\in E} P_{i,k}(t) P_{k,j}(s)$. *(Chapman-Kolmogorov equation),*

3.2 Transition and infinitesimal matrices

We can express all the transition functions in a transition matrix $P(t) = (P_{i,j}(t))_{i,j\in E}$. Therefore the properties of the last proposition are summarized as

1. $P(t)$ is a stochastic matrix $\forall t \geq 0$.
2. $P(0) = I$.

3. $P(t+s) = P(t)P(s)\ \forall t \geq 0$.

Definition 3.2. ***Standard transition matrix***
A family of transition matrices $\{P(t)\}_{t\geq 0}$ *is standard if* $\lim_{t\to 0} P(t) = P(0) = I$.

Proposition 3.2.
Let $\{P(t)\}_{t\geq 0}$ *be a family of transition matrices with space state* E*, then*

1. *For every* $i \in E$ *the following limit exists.*

$$q_i = \lim_{t\to 0} \frac{1 - P_{i,i}(t)}{t}.$$

2. *For every* $i \neq j$ *the following limit exists.*

$$q_{i,j} = \lim_{t\to 0} \frac{P_{i,j}(t)}{t} \geq 0.$$

3. *For every* $i \in E$, $\sum_{i\neq j} q_{i,j} \leq q_i$.

Notice that $q_{i,j} = P'_{i,j}(0)$ and $q_i = -P'_{i,i}(0)$.

Definition 3.3. ***Infinitesimal Matrix***
Let $\{P(t)\}_{t\geq 0}$ *be a family of transition matrices of continuous Markov chain. Then, the infinitesimal matrix of* $\{P(t)\}_{t\geq 0}$ *is defined as*

$$Q = (q_{i,j})_{i,j\in E},$$

where $q_{i,j} = P'_{i,j}(0)$ *and* $q_i = -P'_{i,i}(0)$.

In practice, sometimes it is useful to express $P_{i,i}(\Delta t) = 1 - q_i\Delta t + o(\Delta t)$ and $P_{i,j} = q_{i,j}\Delta t + o(\Delta t)$, and solve for the infinitesimals, knowing that $\lim_{t\to 0} = o(\Delta t)/\Delta t = 0$.

3.3 Kolmogorov differential equations

Let $Q = (q_{i,j})_{i,j\in E}$ be a real matrix and $\{P_{i,j}(t)\}_{t\geq 0}$ be a family of functions, then satisfies the **Kolmogorov backward equations**

$$P'_{i,j}(t) = \sum_{k\in E} q_{i,k} P_{k,j}(t)\ \forall i,j \in E, t \geq 0,$$

and the **Kolmogorov forward equations**

$$P'_{i,j}(t) = \sum_{k\in E} P_{i,k}(t) q_{k,j}\ \forall i,j \in E, t \geq 0.$$

which can be expressed matricially as
Kolmogorov backward equations: $P'(t) = QP(t)$.
Kolmogorov forward equations: $P'(t) = P(t)Q$.

Proposition 3.3.
Let $\{P_{i,j}(t)\}_{t\geq 0}$ *be a family of transition matrices on the finite state space* E *with infinitesimal matrix*

Q. Then $P(t)$ satisfies the backward and forward Kolmogorov equations, therefore

$$P(t) = e^{tQ} = \sum_{n=0}^{\infty} \frac{1}{n!} Q^n, \quad \forall t \geq 0.$$

The idea is to diagonalize $Q = HJH^{-1}$, then $e^{tJ} = \operatorname{diag}(e^{t\lambda_1}, \ldots, e^{t\lambda_n})$ is a diagonal matrix. If Q has eigenvectors associated with an invariant space of dimension $m > 1$, then it can be proven that e^{tJ} is block diagonal, where each block is of the form

$$e^{\lambda t} \begin{pmatrix} 1 & t & t^2/2! & t^3/3! & \ldots & t^{m-1}/(m-1)! \\ & 1 & t & t^2/2! & \ldots & t^{m-2}/(m-2)! \\ & & 1 & t & \ldots & t^{m-3}/(m-3)! \\ & & & 1 & \ldots & t^{m-4}/(m-4)! \\ & & & & \ddots & \vdots \\ & & & & & 1 \end{pmatrix}.$$

3.4 Jump chains

Proposition 3.4.
If $i \in E$ is not absorbent, then X_a is independent of a and has distribution

$$P_i(X_a = j) = \frac{q_{i,j}}{q_i}.$$

Furthermore, if the Markov chain is on the state $i \in E$, it remains on it a random time with exponential distribution with parameter q_i, after that, it jumps to another state $j \neq i$ with probability $q_{i,j}/q_i$.

Therefore, we can define a new discrete Markov chain $\{Y_n\}_{n\in\mathbb{N}}$ that tracks the states visited by $\{X_t\}_{t\geq 0}$, with probabilities

$$P(Y_{n+1} = j \mid Y_n = i) = \begin{cases} q_{i,j}/q_i & \text{if } j \neq i, \\ 0 & \text{if } j = i. \end{cases}$$

whenever $q_i > 0$. Naturally, if $q_i = 0$, the state is absorbent and $P(Y_{n+1} = i \mid Y_n = i) = 1$. These probabilities can be expressed as a transition matrix of $\{Y_n\}_{n\in\mathbb{N}}$ called jump matrix

$$\Sigma = \begin{pmatrix} 0 & \frac{q_{i_0} q_{i_1}}{q_{i_0}} & \frac{q_{i_0} q_{i_2}}{q_{i_0}} & \frac{q_{i_0} q_{i_3}}{q_{i_0}} & \cdots \\ \frac{q_{i_1} q_{i_0}}{q_{i_1}} & 0 & \frac{q_{i_1} q_{i_2}}{q_{i_1}} & \frac{q_{i_1} q_{i_3}}{q_{i_1}} & \cdots \\ \frac{q_{i_2} q_{i_0}}{q_{i_2}} & \frac{q_{i_2} q_{i_1}}{q_{i_2}} & 0 & \frac{q_{i_2} q_{i_3}}{q_{i_2}} & \cdots \\ \vdots & & & & \ddots \end{pmatrix}$$

Theorem 3.1.
Let Q be the infinitesimal matrix of a continuous Markov chain, then if Q is regular, a state $i \in E$ is recurrent or transient for $\{X_t\}_{t\geq 0}$ if and only if is recurrent or transient for the jump chain $\{Y_n\}_{n\in\mathbb{N}}$.

3.5 Stationary distribution

Theorem 3.2.
Let $\{P(t)\}_{t\geq 0}$ be a family of standard transition matrices, then $\forall i, j \in E$,

$$\pi_{i,j} = \lim_{t\to\infty} P_{i,j}(t) \quad \textit{exists.}$$

Analogously to the discrete Markov chain, if $\pi = (\pi_i)_{i\in E}$ is a stationary distribution, then $\pi P(t) = \pi \ \forall t \geq 0$, and the marginal distributions of X_t are, precisely, π. However, in the continuous case, in practice, is useful to compute π using the infinitesimal matrix Q.

Proposition 3.5.
Let Q be the infinitesimal matrix of a continuous Markov chain $\{X_t\}_{t\geq 0}$ and let $\Pi = \lim_{t\to\infty} P(t)$, then $Q\Pi = 0$, that is

$$\sum_{k\in E} q_{i,k}\pi_{k,j} = 0, \quad \forall i, j \in E.$$

Proposition 3.6.
Let $\{X_t\}_{t\geq 0}$ be an irreducible Markov chain with all states being recurrent, and let the initial state be $i \in E$. Define S_1 as the time of the first state change

$$S_1 = \min(t \geq 0 : X_t \neq i).$$

Let τ_i be the first time of return to state i after the initial state change, that is,

$$\tau_i = \min(t \geq S_1 : X_t = i).$$

Then,

$$E_i[\tau_i] = \frac{1}{q_i \pi_i},$$

3.6 Expected values

Let $\mathcal{T}$ and $\mathcal{R}$ be the sets of transient and recurrent states respectively. Then, the **mean arrival times** can be computed solving the linear system

$$\begin{cases} \displaystyle\sum_{j\in\mathcal{T}} q_{i,j} M_j = -1, \\ \forall i \in \mathcal{T}. \end{cases}$$

Suppose that $f(i)$ is the gain that we obtain for reaching the state i, then the **expected reward for the visited states** satisfy

$$\begin{cases} M_i = f(i) + \displaystyle\sum_{j\neq i} \frac{q_{i,j}}{q_i} M_j, \\ \forall i \in E, \quad \text{such that } q_i \neq 0. \end{cases}$$

Notice that if $f(i) = 1, \ \forall i \in \mathcal{T}$ and $f(i) = 0, \ \forall i \in \mathcal{R}$, then M_i is the expected number of jumps of the chain before being absorbed for some recurrent state.

Analogously, if $g(i,j)$ is the gain obtained for jumping from i to j, then **expected reward per jumps** satisfy

$$\begin{cases} M_i = \sum_{j\neq i} \frac{q_{i,j}}{q_i}(g(i,j)+M_j), \\ \forall i \in E, \quad \text{such that } q_i \neq 0. \end{cases}$$

Notice that if $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$ and $g(i,j) = 1$ if $i \in \mathcal{T}_1,\ j \in \mathcal{T}_2$, and $g(i,j) = 0$ otherwise, then M_i is the expected number of jumps from states of $\mathcal{T}_1$ to $\mathcal{T}_2$ before the absorption of some recurrent state.

3.7 Exercises

1. In a factory, there are two machines simultaneously in operation. Each machine takes a random time with an exponential distribution with a mean of two weeks to break down, independently of each other. When a machine breaks down, the company's sole operator repairs it, taking a time with an exponential distribution with a mean of one week. Initially, both machines are operational. Compute:

1. The distribution of the number of operative machines.
2. The long run proportion of time that there is no machine working and the long run proportion of time that the operator is working.
3. The density function of the first time T in which all machines are broke down.
4. The mean number of breakdowns till T.
5. The average time it takes, from the first breakdown, until both machines are operational again.

Let S_1, S_2 be the breakdowns times of the machines, both with exponential distribution with parameter $1/2$ and T_1 the repair time of the operator with exponential distribution with parameter 1.

If the chain is at the state 0, then the operator repairs the machine with probability $1-e^{-t}$, therefore $P'_{0,1}(0) = q_{0,1} = 1$ and $-P'_{0,0}(0) = q_0 = -1$.

If the chain is at the state 1, it remains on it with a exponential distribution $\min(S_1, T_1)$ with parameter $1/2+1 = 3/2$. Therefore $q_1 = -3/2$. The other probabilities are $P_{1,0} = e^{-t}(1-e^{-\frac{1}{2}t})$ and $P_{1,2} = (1-e^{-t})e^{-\frac{1}{2}t})$, then $q_{1,0} = 1/2$ and $q_{1,2} = 1$.

Finally, if the chain is at the state 2, then the first breakdown is produced with distribution $\min(S_1, S_2)$ which have exponential distribution with mean 1. Therefore $P'_{2,1}(0) = q_{2,1} = 1$ and $-P'_{2,2}(0) = q_2 = -1$.

The infinitesimal matrix is then

$$Q = \begin{array}{c|c} & \begin{matrix} 0 & 1 & 2 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \end{matrix} & \begin{pmatrix} -1 & 1 & \\ 1/2 & -3/2 & 1 \\ & 1 & -1 \end{pmatrix} \end{array},$$

1.

If we diagonalize Q, we obtain

$$Q = \frac{1}{15}\begin{pmatrix}1 & 2 & 2\\ 1 & 0 & -3\\ 1 & -1 & 2\end{pmatrix}\begin{pmatrix}0 & & \\ & -1 & \\ & & -5/2\end{pmatrix}\begin{pmatrix}3 & 6 & 6\\ 5 & 0 & -5\\ 1 & -3 & 2\end{pmatrix}.$$

Therefore

$$P(t) = \frac{1}{15}\begin{pmatrix}1 & 2 & 2\\ 1 & 0 & -3\\ 1 & -1 & 2\end{pmatrix}\begin{pmatrix}1 & & \\ & e^{-t} & \\ & & e^{-5t/2}\end{pmatrix}\begin{pmatrix}3 & 6 & 6\\ 5 & 0 & -5\\ 1 & -3 & 2\end{pmatrix}.$$

Which, after multiplying the matrices, the last row (since we begin the process on the state 2), gives us the distribution of the number of operative machines

$$\begin{aligned} P_{2,0}(t) &= \frac{1}{5} - \frac{1}{3}e^{-t} + \frac{2}{15}e^{-5t/2},\\ P_{2,1}(t) &= \frac{2}{5} - \frac{2}{5}e^{-5t/2},\\ P_{2,2}(t) &= \frac{2}{5} + \frac{1}{3}e^{-t} + \frac{4}{15}e^{-5t/2}. \end{aligned}$$

2.
The long run proportions can be computed as $\pi_i = \lim_{t\to\infty} P_{2,i}(t)$ or as the solutions of $\pi Q = 0$. Naturally, we obtain the same solution with both approaches which are $\pi_0 = 1/5, \pi_1 = 2/5$ and $\pi_3 = 2/5$. Thus, the long run proportion where there is no machine working is $\pi_0 = 1/5$, while the long run proportion of the operator working is $\pi_0 + \pi_1 = 3/5$.

3.
We modify the Markov chain making the state 0 absorbent. The new infinitesimal matrix is then

$$\bar{Q} = \begin{pmatrix}0 & & \\ 1/2 & -3/2 & 1\\ & 1 & -1\end{pmatrix},$$

which we can diagonalize, to obtain

$$\bar{P}(t) = \frac{1}{34}\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 1\\ 1 & \frac{1+\sqrt{17}}{4} & \frac{1-\sqrt{17}}{4}\end{pmatrix}\begin{pmatrix}1 & & \\ & e^{(-5+\sqrt{17})t/4} & \\ & & e^{(-5-\sqrt{17})t/4}\end{pmatrix}\begin{pmatrix}34 & 0 & 0\\ -17-3\sqrt{17} & 17-\sqrt{17} & 4\sqrt{17}\\ -17+3\sqrt{17} & 17+\sqrt{17} & -4\sqrt{17}\end{pmatrix}.$$

Thus,

$$f_T(t) = \frac{d}{dt}\bar{P}_{2,0}(t) = \frac{\sqrt{17}}{17}\left(e^{(-5+\sqrt{17})t/4} - e^{(-5-\sqrt{17})t/4}\right).$$

4.

Using the formula $M_i = \sum_{j\neq i} \frac{q_{i,j}}{q_i}(g(i,j) + M_j)$, where M_j is the expected number of of breakdowns from j, and $g(1,0) = 1,\ g(2,1) = 1$ and 0 in the other cases (we only count the breakdowns). We obtain

$$M_1 = \frac{1}{3} + \frac{2}{3} M_2 \qquad \text{and} \qquad M_2 = 1 + M_1.$$

Solving the equations we obtain that $M_2 = 4$.

5.

The expected recurrence time of the state 2 is $1/(\pi_2 q_2)$, therefore the average time it takes until both machines are operational again is $5/2$ weeks.

References

[1] Sheldon M.Ross. *Introduction to probability models*. Academic Press, 2019.

[2] Ricardo Vélez Ibarrola y Tomás Prieto Rumeau. *Procesos Estocásticos*. UNED, 2013.