Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea Triennale in Matematica

FINITE STATE MARKOV CHAINS AND PREDICTION OF MARKET TRENDS USING REAL DATA

Relatore:
Chiar.mo Prof.
Nicola Arcozzi

Presentata da:
Francesca Tiozzo Gobetto

♯ Sessione Unica
2019/2020

Introduction

In this thesis we discuss finite state Markov chains, which are a way to model stochastics processes without memory in time. Thus, we assume that the past does not influence the future. Finite chains are interesting for their properties. We examine some of them in order to study how we can use these chains to model long-term processes. In the end, we discuss one application to finance.

In the first chapter we introduce some basic definitions and results of probability. The concepts of measurable space, measure and of space of probability are presented on the base of the course pf Probability of A. Pascucci. After the definition of a stochastic process, we introduce the reader to Markov chains. They are stochastic processes in discrete time where the past has no influence on the future behavior of the process. Two useful ways to represent them are matrices and directed graphs, which are a set of nodes, each of which corresponds to a certain state, connected by directed arrows.

In the second chapter we classify Markov chains in the same way as MIT open courses present them. First we define what classes are and then we divide them into recurrent and transient. We then conclude dealing with the periodicity of classes.

Previous chapters lay the basis to explain when $[P^n]$, the matrix obtained by taking the n^{th} power of $[P]$, converges as n approaches infinity. We focus on the case of ergodic unichains. We give some preliminary results, using the Chapman-Kolmogorov equation, that will be used to prove convergence of this particular type of Markov chains.
We first prove the convergence in the case of $[P] > 0$ and then we generalize to the case of ergodic unichains. We conclude we a comment on other types of Markov chains.
Last chapter deals with the application of Markov chains in the prediction of fluctuations in prices of stocks. We first give a brief insight of how stock

markets work.
Then we carry on an empirical analysis in order to predict how prices will change in the future. Since the model we develop is based on ergodic unichains, the theory we introduced in previous chapters shows that we should be able to predict prices in the long run. We use data on the FITSE-MIB index from Borsa Italiana to build Markov chains. The index is based on a basket of forty relevant firms belonging to different sectors of the Italian economy. We use the fluctuations in prices between July the 29^{th} and July the 31^{st} of each component of the basket the calculate the percentage of increase, decrease or stationarity. We then use the matrix form to predict the prices in the future. A Python code is also used to make predictions in the far future.
Comparing the theoretical results to the real fluctuations, we see that results were not reliable. We explain what are the difficulties in developing a mathematical model for the financial markets.

1 Preliminary knowledge

We start by giving some basic definitions.

Definition 1.1. A *measurable space* is a couple (Ω, F), where:
i) Ω is a non-empty set
ii) F is a σ-algebra, meaning that F is a non-empty family of subsets of Ω that satisfies the following properties:
ii-a) if $A \in F$ then $A^c := \Omega \setminus A \in F$
ii-b) the countable union of the elements of F belongs to F.

Definition 1.2. A *measure* on the measurable space (Ω, F) is a function $\mu : F \longrightarrow [0, +\infty]$ such that:
i) $\mu(\emptyset) = 0$
ii) μ is σ-summable on F, meaning that for every sequence $(A_n)_{n\in\mathbb{N}}$ of disjoint elements of F it holds that

$$\mu(\bigcup_{n=1} A_n) = \sum_{n=1} \mu(A_n).$$

Definition 1.3. Let us consider as measurable space a generic metric space (M, ρ).
The *Borel* $\sigma - algebra$ B_ρ is the smallest $\sigma - algebra$ that contains the open sets of (M, ρ).

Definition 1.4. A measure space (Ω, F, μ) in which $\mu(\Omega)$=1 is called a *space of probability*. In this such case, let us use the letter P instead of μ and let us call P a measure of probability.

In a probability space (Ω, F, P), each element $\omega \in \Omega$ is defined as *outcome*; each $A \in F$ is defined as an event and $P(A)$ is the probability of A. If Ω is finite or countable, we always assume that $F = P(\Omega)$ and we say that $(\Omega, P(\Omega), P)$ is a *discrete probability space*. If Ω is not countable, then we are in the case of *continuous probability space*.

Definition 1.5. Let B_ρ be the *Borel* $\sigma - algebra$ on a metric space (M, ρ). A *distribution* is a probability measure on (M, B_ρ).

Definition 1.6. In a probability space (Ω, F, P), let B be an event such that $P(B) > 0$. The *probability of A given B* is defined by:

$$P(A|B) := \frac{P(A \cap B)}{P(B)}, \quad A \in F.$$

Definition 1.7. Let us a consider a given probability space (Ω, F, P) and $d \in \mathbb{N}$. For a given $H \in \mathbb{R}^d$ and a given function $X : \Omega \longrightarrow \mathbb{R}^d$, let us denote with

$$(X \in H) := \{\omega \in \Omega \mid X(\omega) \in H\} = X^{-1}$$

the inverse image of H.
Intuitively $(X \in H)$ is the set of ω such that $X(\omega) \in H$.

Definition 1.8. A random variable on (Ω, F, P) with values in $\mathbb{R}^d$ is a function $X : \Omega \longrightarrow \mathbb{R}^d$ such that $(X \in H) \in F$ for each $H \in B^d$. In this such case, we say that X is $F - measurable$.

Definition 1.9. A *discrete distribution* is of the form

$$\mu(H) := \sum_{n=1} p_n \delta_{x_n}(H), \quad H \in B_d$$

where (x_n) is a sequence of points in $\mathbb{R}^d$ and (p_n) is a sequence that satisfies the following property: $\sum_{n=1}^{\infty} p_n = 1$ and $p_n \geq 0$, $n \in \mathbb{N}$.

Definition 1.10. Let X be a random variable on (Ω, F, P) with values in $\mathbb{R}^d$. The *distribution of X conditional on B* is the distribution of X relatively to the conditional probability $P(\bullet \mid B)$. It is defined as

$$\mu_{X|B}(H) := P(X \in H \mid B), \quad H \in B_d.$$

Definition 1.11. A *stochastic process* is a family of random variables X_θ, indexed by a parameter θ, which belongs to some index set Θ.
If this set is a set of integers, we have stochastic processes in discrete time. In this such a case we write $\Theta = \{t_0,\ t_1,\ ,...,\ t_k,\ t_{k+1},...\}$ for $k \in \mathbb{N}$.

Definition 1.12. A discrete stochastic process is said to be *homogeneous* (or time invariant) if the transition probability between any two state values at two given times depends only on the difference between those times.
In formulas, we have that the conditional probabilities must satisfy

$$P(X_t|X_{t-a}) = P(X_s|X_{s-a})\ for\ all\ t,\ s \in \Theta\ and\ for\ all\ a \in \mathbb{N}$$

We are often interested in conditional distributions of the form

$$P(X_{t_k}|X_{t_{k-1}},\ X_{t_{k-2}},...,\ X_{t_1},\ X_{t_0})$$

for a certain set of times $t_k > t_{k-1} > t_{k-2} > ... > t_1 > t_0$ where $t_i \in \Theta$ for all $i \in \mathbb{N}$. In discussing them, we will focus on a specific kind of stochastic processes that satisfy the Markov property.

Definition 1.13. The *Markov property* states that

$$P_{ij} = P(X_t = j | X_{t-1} = i) \tag{1.1}$$

where P_{ij} is the probability of transition from state i to state j. In other words, the distribution of X_t only depends on the position of the process at time $t-1$ and not on what happened at times $s \leq t-2$.

Definition 1.14. A *finite Markov chain* is a homogeneous discrete stochastic process with a finite number of states.

Let M be the number of states of a finite Markov chain: $[M] = \{1, 2, ..., M\}$. The previous definitions state that Markov chains are stochastic processes in which changes can happen only for integer-times. Due to the Markov property, the effect of the past on the future is totally summarized by the previous state. Markov chains are called homogeneous because of their time independence. In fact, the conditional probability of the Markov property states that all the history depends only on the last step, independently on when that step happens.
Thus, finite state Markov chains are a tool to model any discrete integer time process.

For every Markov chain, we have an initial probability, which is the probability distribution of the states at time 0. Using the iteration process, just knowing the transition probabilities and initial probabilities is enough to find the probability distribution of the states at any instant time.

1.1 Representation of finite state Markov chains

We have different ways to represent a finite Markov chain: we can use a matrix, such as part (b) of figure (1.1), or we can use a directed graph, such as part (a).

The $M \times M$ transition matrix of a Markov chain is $[P] = (P_{ij})_{ij}$, where $P_{ij} = P\left[x_{n+1} = j | x_n = i\right] i, j \in [M]$. Associated to the matrix $[P]$ there is an initial distribution $q = (q_1, ..., q_M)$, where $q_i = P(x_0 = i)$. This $[P]$ matrix is *stochastic* if its entries are non-negative real numbers and the sum of each row is 1.

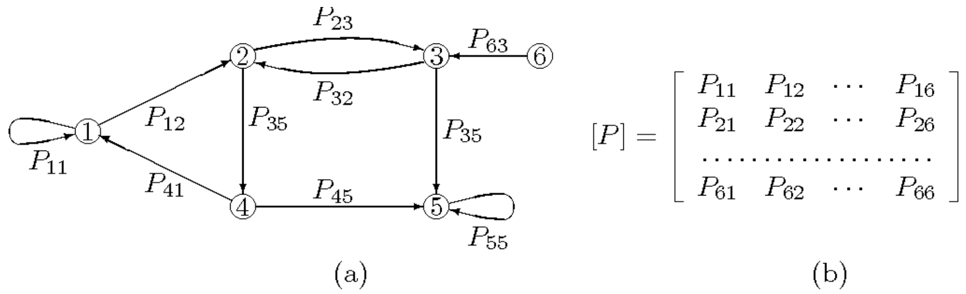


Figure 1.1: Representation of Markov chains

On the other side we have the graphical model, in which each node corresponds to a state and each arc corresponds to a transition probability. It is really useful because it makes a clear distinction between the zero and non-zero transition probabilities: if there is a positive probability of going from one of the M states to another we have an arc, otherwise we don't.

2 Classification of states

We now introduce some other useful definitions.

Definition 2.1. A *walk* is an ordered string of nodes, where the probability of going from each node to the next one is non-zero.

There is no kind of constraint in the definition of a walk. For example, we can have repetition of states. The minimum number of states in a walk is one and there is no limit for the maximum number.

Definition 2.2. A *path* is a walk where there are no repeated nodes.

Since we never go through one node twice, the maximum number of steps for a path in an M finite state Markov chain is $M - 1$.
Walks can be joined to form a longer walk. If there's a walk from i to j and from j to k, then there is a walk also from i to k that is found by concatenation.
We say that state i communicates with state j if it exists a walk from i to j and one from j to i.

Looking at figure (1.1), we see that node 3 is accessible from node 1 because there's a walk made by nodes 1, 2, 3. So there's a positive probability of going to state 3 from node 1. But if we want to calculate probability P_{13}, we should also look at the fact that we can have cycles. For example 1, 1, 2, 3 and 1, 2, 3, 2, 3 are also walks.
Node 6 is not accessible. So, if we are in node 6, we always go away from it. More formally, we say that state j is *not accessible* from i if P_{ij}^n is equal to 0 for all n.

We now want to group states on the basis of how they communicate with each other.

Definition 2.3. A *class* of states is a non-empty set of states, where all the pairs of states in a class communicate with each other, and none of them communicate with any other state in the Markov chain.

For finding or naming a class, we can have a representative state. We can just pick one of the states in a certain class and find all the states that communicate with this single state. This is due to the fact that if two states communicate with one state, then these two states communicate with each other. And if there's a state that doesn't communicate with the chosen state, it doesn't communicate with anybody else whom the chosen state communicates with.

We can partition a Markov chain using classes, which means that we want to cover all the finite space of the states without intersections between classes. To do so, we call a single state a class even if it does not communicate with itself.
We have to show that classes do not intersect: if a state i belongs to both $class\ 1$ and $class\ 2$, this means that i communicates with all the states in $class\ 1$ and in $class\ 2$, so all states in $class\ 1$ communicate with the states in $class\ 2$.

In the Markov chain of figure (2.1), we notice that there are four communicating classes. Nodes 1 and 2 communicate with each other, but they do not communicate with any other nodes in the graph. Similarly, states 3 and 4 communicate with each other, but with none of the others. Node 5 does not communicate with any other nodes, so it by itself is a class. Finally, states 6, 7, and 8 form another class. Thus, here are the classes:

$$Class\ 1 = \{state\ 1,\ state\ 2\},$$
$$Class\ 2 = \{state\ 3,\ state\ 4\},$$
$$Class\ 3 = \{state\ 5\},$$
$$Class\ 4 = \{state\ 6,\ state\ 7,\ state\ 8\}.$$

Definition 2.4. We say that state i is *recurrent* if for all the states j that

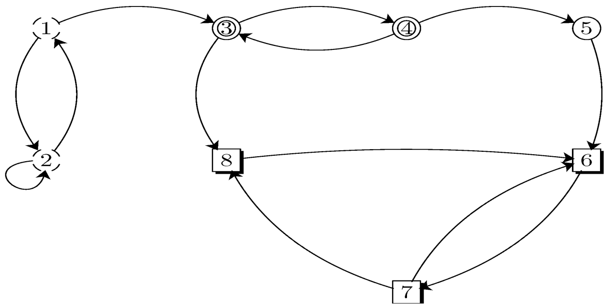


Figure 2.1: Classification of states

are accessible from i, we also have that i is accessible from j.
Using symbols: $i \rightarrow j \Rightarrow j \rightarrow i$.

If a state is not recurrent, we call it *transient*.

If i is recurrent and there is a walk from i to any other state j, there should be a walk from j to i.
On the other side, a transient state i says that there might be some kind of walk from i to some k but it's impossible to go back.

In figure (2.1), we have that states $\{6, \ 7, \ 8\}$ are recurrent, while states $\{1, \ 2, \ 3, \ 4, \ 5\}$ are transient.

Theorem 2.1. *If we partition a Markov chain using classes, then the states in the class are all recurrent or all transient.*

Proof: Let's assume that state i is recurrent and let's define S_i as the set of all the states that communicate with i. Since i is recurrent, if j is accessible from i, state i is also accessible from j. We know that i and j communicate with each other if and only if j is in set S_i.
Let's now consider a state k, that is accessible from j, and j is accessible from i. So k is accessible from i. But k accessible from i implies that i is also accessible from k, because i is recurrent. We also have that j is accessible from i and this implies that j is also recurrent. So if k is accessible from j, then j is also accessible from k for any k. This means that if one state

in a class is transient, then all the states in the class are transient too. We conclude that if a state in a class is recurrent, then also all the other states in that class are recurrent.
Thus, the states in a class are either or all recurrent or all transient.

Definition 2.5. The *periodicity* of state i is the greatest common divisor of the number of steps needed to go from state i to state i, under the hypothesis that there's a positive probability that we can go from state i to state i.

If i's period is a number greater than 1, then i is said to be *periodic*, otherwise it's called *aperiodic*.

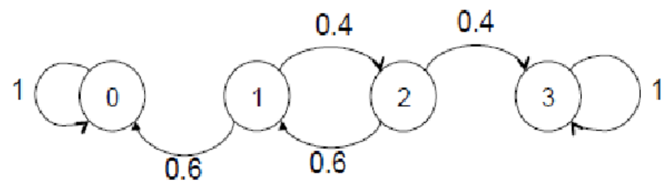


Figure 2.2: Periodicity of states

Consider the Markov chain shown in figure (2.2). Starting from state 1, we only return to 1 at times n = 2, 4, 6, 8, ..., *i.e.* at times $n = 2^i$, where i is a positive integer. We have periodicity since the greatest common divisor of 2^i is 2. Thus, state 1 is called a periodic state with period $d(1) = 2$. If n is not divisible by 2 we have that $P^n_{1,1} = 0$.
We always say that a state i is aperiodic if there's a walk from i to i, and in this walk, there is a loop. In figure (2.2), state 0 has a period of 1 and is called aperiodic.

Theorem 2.2. *All the states in the same class have the same period.*

Proof: Let P^k_{ij} denote the k-step transition probability from state i to j. Let $d(i)$ denote i's period. Suppose that i and j communicate, which means

that $P_{ij} > 0$ and $P_{ji} > 0$. Let m and n be some positive integers such that $P_{ij}^n > 0$ and $P_{ji}^m > 0$. Then we have that $P_{ii}^{n+m} \geq P_{ij}^n P_{ji}^m > 0$. This implies that $n + m$ is a multiple of $d(i)$.
Let's consider $q \in \{k \mid P_{jj}^k > 0\}$. We have that $P_{ii}^{n+q+m} > 0$, which means that we can go from i to j, then from j to some other node and back to j and finally returne to i. Therefore $n+q+m$ is a multiple of $d(i)$. Thus, also q is a multiple of $d(i)$. As a consequence, $d(j) \geq d(i)$. If we reverse the roles of i and j, we get $d(i) \geq d(j)$.

Since all the representatives of a certain class have the same period, we say that *the class* they belong to has that period.
Similarly, since all the states in a class are or all recurrent or all transient, we say that *the class* is recurrent or transient.

Theorem 2.3. *Assume that all the states in a certain class have period $d = n$. Then the considered class can be partitioned into n subclasses S_1, S_2, ..., S_n. For each subclass, we only have transitions from it to another one. There's no transition in a subclass to itself and the only possible transitions are from S_1 to S_2, from S_2 to S_3,..., from S_{n-1} to S_n and finally from S_n to S_1.*

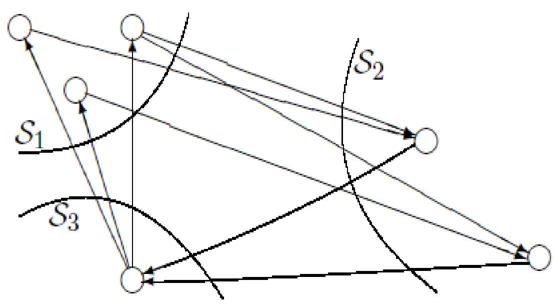


Figure 2.3: Subclasses

Figure (2.3) gives an illustration of the theorem.

Proof: We choose a certain state in the class, for example state 1. Then, we define the subclasses as the sets $S_1, ..., S_n$, such that

$$S_m = \left\{j : P_{1j}^{dn+m} > 0 \; for \; some \; d \geq 0\right\}; \; 1 \leq m \leq n. \tag{2.1}$$

For every state j in the class, we show that there is one and only one value of m such that $j \in S_m$.
Since 1 and j communicate, it exists some r such that $P_{1j}^r > 0$ and some s such that $P_{j1}^s > 0$. Thus, we can build a walk of length $r + s$ from state 1 to state 1 by joining the walks of length r from 1 to j and of length s from j to 1. As a consequence, $r + s$ is divisible by n.
Dividing r by n, we define m, $1 \leq m \leq n$, by $r = m + nd$, where d is an integer.
From (2.1) we have that $j \in S_m$.
We define r' as any other integer such that $P_{1j}^{r'} > 0$. Then $r' + s$ is also divisible by n, so that $r' - r$ is divisible by n. Thus $r' = m + d'n$ for some integer d' and the same m. Since r' is any integer such that $P_{1j}^{r'} > 0$, then $j \in S_m$ for only that one value of m. The fact that j was chosen as any of the states of a given class shows that the sets S_m are disjoint and partition the class.
Finally, suppose $j \in S_m$ and $P_{ik} > 0$. Given a walk of length $r = nd + m$ from state 1 to j, there is a walk of length $nd + m + 1$ from state 1 to k. It follows that if $m < n$, then $k \in S_{m+1}$ and if $m = n$, then $k \in S_1$. The proof in complete.

3 Convergence of $[P^n]$

The initial state distribution can be deterministic, which means either that we start from some specific state all the time or that there is a fixed distribution at the initial state. Knowing the transition probabilities and the initial state, we can find the distribution of states at each time instant. So, for characterizing a Markov chain, we just need to know the transition probabilities and the initial distribution.
However, this data is not sufficient to determine the behavior in the very far future. There are very interesting questions that can be asked about it, such as: is there any kind of stable behavior in the very far future? If we can state something about this behavior, then what kind of statements can we have? Under which hypotheses can we have these statements? If there is a pattern of future behavior, is it unique? Is the initial state relevant in determining what happens very far away from the present?
To answer these questions, we see that there is a certain kind of Markov chains that loses memory as n goes to infinity, meaning that whatever distribution it has for the initial state, it will lose memory of that.
This kind of Markov chains is called ergodic.

Definition 3.1. An *ergodic Markov chain* is a Markov chain that has a single recurrent class and is aperiodic.

To study ergodic chains, we use the Champan-Kolmogorov equation, which needs some preliminary knowledge to be introduced.

Definition 3.2. A *stochastic matrix* is a square matrix of non-negative terms in which the elements in each row sum to 1.

An example of stochastic matrix is $[P]$, the matrix of transition probabilities of a Markov chain. We denote by $[P^n]$ the product of n terms: $[P][P]...[P]$. If $n = 2$, $[P^2]$ is the matrix with entries:

$$P^2_{ij} = \sum_{k=1}^{M} P_{ik} P_{kj}. \tag{3.1}$$

We already know that P_{ij} is the probability to go from state i to state j in one step, while P^n_{ij} is the probability that at time n the state is equal to j, given that the initial state is equal to i. Using the mathematical notation we have:

$$P^n_{ij} = \Pr\{X_n = j | X_0 = i\}$$

For example, when $n = 2$, we have to sum over all the possible values 1 to M that x_1 can take to get the desired probability, as we do in equation (3.1). The same reasoning can be iterated to find P^n_{ij}. This idea is made more general by the $Chapman - Kolmogorov$ equation: since $[P^{m+n}] = [P^m][P^n]$, we have that:

$$P^{m+n}_{ij} = \sum_{k=1}^{M} P^m_{ik} P^n_{kj}. \tag{3.2}$$

This implies that, when we want to go from step i to step j, we can go to an intermediate state and sum over all the possible intermediate states we can have. So in this case, if the step is $m + n$ transition, we can slit it up into m and n.

We now ask: does $[P^n]$ converge? In other words, does

$$\lim_{n\to\infty} P^n_{ik} = \pi_{ij} \tag{3.3}$$

exist for all i and j?

Let's assume that this limit exists for all i and j, which means that for each column j of $[P^n]$, all the elements of the column should converge towards the same value. If this limit does exist and is equal to π_j for each column j, we can first multiply both sides of (3.3) by P_{ik} and then sum over all j:

$$\lim_{n\to\infty}\sum_j P_{ij}^n P_{jk} = \sum_j \pi_j P_{jk}. \tag{3.4}$$

We can now use the $Chapman - Kolmogorov$ equation and the left side of (3.4) becomes $\lim_{n\to\infty} P_{ik}^{n+1}$. We define:

$$\lim_{n\to\infty} P_{ik}^{n+1} = \pi_k$$

So (3.4) becomes:

$$\pi_k = \sum_j \pi_j P_{jk}$$

In vector form we write: $\vec{\pi} = \vec{\pi}[P]$. We define the probability vector.

Definition 3.3. A *probability vector* is a vector $\vec{\pi} = (\pi_1, ..., \pi_m)$ for which each π_i is non-negative and $\sum_i \pi_i = 1$. A probability vector $\vec{\pi}$ is called a $steady - state\ vector$ for the transition matrix $[P]$ if

$$\vec{\pi} = \vec{\pi}[P] \;\; where \;\; \sum_i \pi_i = 1, \;\; \pi_i \leq 0, \;\; 0 \leq i \leq M \tag{3.5}$$

If $\vec{\pi}$ satisfies (3.5), it is a probability vector and it holds that $\vec{\pi}[P] = \vec{\pi}[P^2]$. Iterating this, we get $\vec{\pi} = \vec{\pi}[P^2] = \vec{\pi}[P^3] = ... = \vec{\pi}[P^n] = ...$ Thus, for each positive integer n it holds that $\vec{\pi}[P] = \vec{\pi}[P^n]$.
This shows it is sufficient for $\vec{\pi}$ to be a steady state vector for $[P^n]$ to converge to a matrix whose rows are π. However, it is only a sufficient and not a necessary condition. If $\vec{\pi}$ satisfies (3.5), it does not imply that $[P^n]$ converges. We remark this concept with an example: let $\vec{\pi}$ be such that $\pi_2 = \pi_3 = 1/2$ and $\pi_i = 0$ otherwise. Then it satisfies (3.5). This means that if the chain starts at time 0 at states 2 or 3, then it oscillates between the two for the rest of the time.

However, as we see in figure (3.1), the $\vec{\pi}$ we just defined is not a satisfying form of steady state.

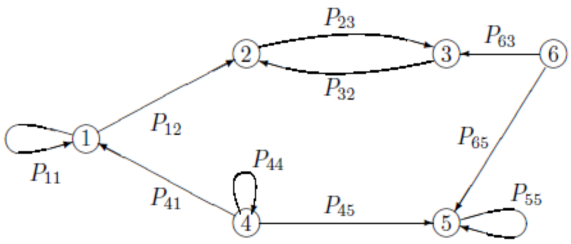


Figure 3.1: Steady state

In order to study the steady-state, it is useful to introduce the following definitions:

Definition 3.4. A *unichain* is a finite state Markov chain that contains a single recurrent class and, perhaps, some transient states. An *ergodic unichain* is a unichain for which the recurrent class is ergodic.

A unichain is a recurrent chain that allows for some transient initial behaviors, that do not affect the long term behavior of the chain.

If the chain has a unique recurrent class, *i.e.* it is a unichain, then the steady state vector $\vec{\pi}$ is unique. If a chain has c recurrent classes, then (3.5) has c linearly independent solutions, each nonzero only over the elements of the corresponding recurrent class.
Since the limit exists for ergodic chains, it holds that each row of $[P^n]$ converges to a unique probability vector solution if the chain is an ergodic unichain.
On the other side, $[P^n]$ does not converge if the Markov chain has one or more periodic recurrent classes.

To have convergence, the elements of each column j must be the same for each row i. Therefore, we are now going to study the difference between the largest and the smallest value of a column j and see how this difference varies as n grows large.

Lemma 3.1. *Let $[P]$ be the $M \times M$ transition matrix of an arbitrary finite state Markov chain. Let $[P^n]$ be the matrix with entries P_{ij}^n. Then for each state j and for all integers $n \geq 1$*

$$\max_i P_{ij}^{n+1} \leq \max_l P_{lj}^n \qquad \min_i P_{ij}^{n+1} \geq \min_l P_{lj}^n. \tag{3.6}$$

So for each j, $\max_i P_{ij}^n$, which is the most probable path from i to j in n steps is non increasing in n, and $\min_i P_{ij}^n$ is non-decreasing in n.
This theorem implicitly says that for an ergodic finite state Markov chain the limit of equation (3.3) converges, as we will later see.

Proof: For any states i, j and any step n,

$$P_{ij}^{n+1} = \sum_{l=1}^{M} P_{il} P_{lj}^n \leq \sum_{l=1}^{M} P_{il} \max_l P_{lj}^n = \max_l P_{lj}^n$$

This is true for all states i, so it is true also for the maximizing i. The same holds when we substitute the max with the min and we reverse the inequality:

$$P_{ij}^{n+1} = \sum_{l=1}^{M} P_{il} P_{lj}^n \geq \sum_{l=1}^{M} P_{il} \min_l P_{lj}^n = \min_l P_{lj}^n$$

Example 1 Figure (3.2) shows a two-states chain where $P_{12} = 1$ and

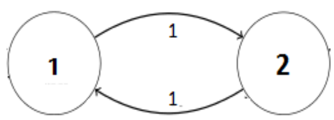


Figure 3.2: Example 1

$P_{21} = 1$. Then, P_{12}^n and P_{21}^n alternates between 1 and 0. So the maximum is 1, which is non-increasing, and the minimum is 0, which is non-decreasing in n.

Example 2

Consider the two-states Markov chain with $P_{12} = 3/4$ and $P_{21} = 3/4$ of figure (3.3).

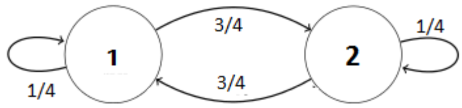


Figure 3.3: Example 2

Then, as n increases, $P_{12}^n = 3/4,\ 3/8,\ 9/16, ...$ and $P_{22}^n = 1/4,\ 5/8,\ 7/16, ...$ Each sequence oscillates while approaching 1/2. $\max(P_{12}^n,\ P_{22}^n) = 3/4,\ 5/8,\ 9/16, ...,$ which is decreasing towards 1/2, while $min(P_{12}^n,\ P_{22}^n) = 1/4,\ 3/8,\ 7/16, ...$

We now show how $max(P_{12}^n,\ P_{22}^n)$ and $min(P_{12}^n,\ P_{22}^n)$ vary with n.

First, we consider the case where end up in state 2 in exactly n steps with $n = 1$. We have two alternatives: either we are in state 1 and we go to state 2, or we are in state 2 and we stay in state 2. The maximum is going to be 3/4 and the minimum is going to be 1/4.

Now suppose $n = 2$, so we want to end up in state 2 in 2 steps. We are going to have the maximum if we start in state 2, then we go with the first step to state 1 and then we go back to state 2. We get that $max(P_{12}^2,\ P_{22}^2) = P_{22}^2 = 5/8$.

We can repeat the same reasoning for various values of n.

Notice that, since the chain of example 2 has loops, its periodicity is completely destroyed.

3.1 Steady state for $[P] > 0$

The previous lemma (3.1) is true for any finite state Markov chain. Now we are going to focus on the case in which $[P] > 0$ and then we will prove the results for an arbitrary finite Markov chain.

$[P] > 0$ means that every entry in this matrix is greater than 0 for all i and j, which means that the graph is fully connected. In this such case, we can get from i to j in one step with nonzero probability.

We now want to prove that, as n goes to infinity, the state at time n is independent of the initial state at time 0.

Theorem 3.2. *Let P be the transition matrix of a finite state Markov chain where $[P] > 0$ and let $\alpha = \min_{i,j} P_{ij}$. Then for all states j and all $n \geq 1$:*

$$\max_i P_{ij}^{n+1} - \min_i P_{ij}^{n+1} \leq (\max_l P_{lj}^n - \min_l Plj^n)(1 - 2\alpha) \tag{3.7}$$

$$(\max_l P_{lj}^n - \min_l P_{lj}^n) \leq (1 - 2\alpha)^n \tag{3.8}$$

$$\lim_{n\to\infty} \max_l P_{lj}^n = \lim_{n\to\infty} \min_l P_{lj}^n = \pi_j > 0 \tag{3.9}$$

Proof: For given j and n, let l_{min} be the value that minimizes P_{lj}^n. Then:

$$P_{ij}^{n+1} = \sum_k P_{ik} P_{kj}^n$$

$$\leq \sum_{k \neq l_{min}} P_{ik} \max_l P_{lj}^n + P_{il_{min}} \min_l P_{lj}^n$$

$$= \sum_{k \neq l_{min}} P_{ik} \max_l P_{lj}^n + P_{il_{min}} \min_l P_{lj}^n + P_{il_{min}} \max_l \ Plj^n - P_{il_{min}} \max_l Plj^n$$

$$= \max_l P_{lj}^n - P_{il_{min}} (\max_l P_{lj}^n - \min_l P_{lj}^n)$$

$$\leq \max_l P_{lj}^n - \alpha(\max_l P_{lj}^n - \min_l P_{lj}^n) \tag{3.10}$$

In the last step we used the fact that $\alpha \leq P_{il_{min}}$ and that the terms in parenthesis are positive.

We can repeat the same argument switching min and max, as follows:

$$P_{ij}^{n+1} = \sum_k P_{ik} P_{kj}^n$$

$$\geq \sum_{k \neq l_{min}} P_{ik} \min_l P_{lj}^n + P_{il_{min}} \max_l P_{lj}^n$$

$$= \sum_{k \neq l_{min}} P_{ik} \min_l P_{lj}^n + P_{il_{min}} \max_l P_{lj}^n + P_{il_{min}} \min_l Plj^n - P_{il_{min}} \min_l Plj^n$$

$$= \min_l P_{lj}^n + P_{il_{min}} (\max_l P_{lj}^n - \min_l P_{lj}^n)$$

$$\geq \min_l P^n_{lj} + \alpha(\max_l P^n_{lj} - \min_l P^n_{lj}) \tag{3.11}$$

We then substitute the left side of (3.10) with $\min_i P^{n+1}_{lj}$ instead of P^{n+1}_{lj}, and the left side of (3.11) with $\max_i P^{n+1}_{lj}$ instead of P^{n+1}_{lj}. Doing so we obtain the following inequalities:

$$\min_i P^{n+1}_{lj} \leq \max_l P^n_{lj} - \alpha(\max_l P^n_{lj} - \min_l P^n_{lj})$$

$$\max_i P^{n+1}_{lj} \geq \min_l P^n_{lj} + \alpha(\max_l P^n_{lj} - \min_l P^n_{lj})$$

Subtracting (3.12) from (3.13), we obtain:

$$\max_i P^{n+1}_{ij} - \min_i P^{n+1}_{ij} \leq (\max_l P^n_{lj} - \min_l Plj^n)(1 - 2\alpha)$$

(3.7) is proved.

To prove (3.8), we observe that:

$$\min_l P_{ij} \geq \alpha > 0 \tag{3.12}$$

$$\max_l P_{ij} \leq 1 - \alpha \tag{3.13}$$

Subtracting (3.12) from (3.13), we obtain: $\max_l P_{ij} - \min_l P_{ij} \leq 1 - 2\alpha$, which can be iterated over n to get (3.8).

Taking into account lemma (3.1), we can conclude that the limits of (3.9) exist and are non negative.

3.2 Ergodic Markov chains

So far, we have proved results for $[P] > 0$. However the previous theorem (3.2) can be extended to ergodic finite state Markov chains. To do this, we first prove that $P^n_{ij} > 0$ for all i and j when n is sufficiently large. We quantify "n sufficiently large" in the following theorem.

Theorem 3.3. *For an ergodic M state Markov chain,* $P_{ij}^m > 0$ *for all i, j and* $m \geq (M-1)^2 + 1$.

For the chain in figure (3.4), this theorem implies that if $m \geq 26$, then $[P^m] > 0$.
Notice that P_{11}^m is possible for $m = 6,\ 11,\ 16,\ 17,\ 18, \ldots$. In other terms, $P_{11}^m > 0$ for $m = 6x + 5y,\ x, y \in \mathbb{N}$. So, if for example $7 \leq m \leq 10$ or $m = 25$, we cannot go back to state 1 starting from 1. However, if $m \geq 26$, we can always start from state 1 and arrive in state 1.

To have an intuition of the reason why theorem (3.3) holds, we introduce the following lemma:

Lemma 3.4. *Let a and b be two positive relatively prime integers, then the greatest number that cannot be written as a combination of a and b, i.e.* $m \neq xa + yb,\ x, y \in \mathbb{N},\ x \geq 1,\ y \geq 0$, *is* $m = ab - a - b$. [1]

If we set $a = M$ and $b = (M-1)$, we obtain theorem (3.3). In fact, $M(M-1) - M - (M-1) + M = (M-1)^2$ is the largest m that cannot be obtained.

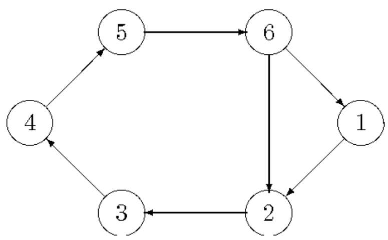


Figure 3.4: Cycles in ergodic Markov chains

Proof of theorem (3.3): First, we notice that the chain must contain a cycle with fewer than M nodes. The definition of ergodic chain states that the chain must have cycles. If the chain has only an M-nodes cycle, then in would be periodic, which is in contrast with the definition of ergodic Markov chain.

[1] See the appendix for the proof. Note that the conditions of lemma (6.1) include that $x, y \geq 0$. On the contrary, in lemma (3.4) we account for $x \geq 1$.

We conclude that some nodes must be on smaller cycles. Let us consider the cycle with the smallest number of nodes. Let ψ be such number. From what we just said, it must be that $\psi \leq M - 1$. Let i be any of the states in the considered cycle. We denote with $\Psi(m) \geq 1$ the set of states that are accessible from state i in m steps. Using mathematical notations, this is equivalent to:

$$\Psi(m) = \left\{j : P_{ij}^{m} > 0\right\}. \tag{3.14}$$

Since we have a cycle, we have that $P_{ii}^{\psi} > 0$. Then it is possible to construct an $m + \psi$ walk from i to j, joining a ψ step walk from i to i with an m step walk from i to j. We have that

$$\Psi(m) \subseteq \Psi(m + \psi). \tag{3.15}$$

We denote the singleton $\{i\}$ as $\Psi(0)$. We set $m = 0$ in (3.15) and we iterate to obtain:

$$\Psi(0) \subseteq \Psi(\psi) \subseteq \Psi(2\psi) \subseteq ... \subseteq \Psi(n\psi) \subseteq ... \tag{3.16}$$

Now we want to show that if one inclusion is satisfied with an equality, then the subsequent inclusions are all equalities. More generally, we want to show that for some given $m \geq 0$ and $s \geq 1$:

$$if\ \Psi(m) = \Psi(m + s),\ then\ \Psi(n) = \Psi(n + s)\ for\ all\ n \geq m. \tag{3.17}$$

By definition, $\Psi(m + 1)$ is the set of states that can be reached in one step from the states in $\Psi(m)$, while $\Psi(m + s + 1)$ is the set of states reachable from $\Psi(m + s)$ in one step. By hypotheses, we have that $\Psi(m) = \Psi(m + s)$, so we can conclude that $\Psi(m + 1) = \Psi(m + s + 1)$. By iteration, we obtain (3.17). The set Ψ can have at maximum M members. Therefore, in (3.16) we can have at most $M - 1$ strict inclusions. So, we have that

$$\Psi((M - 1)\psi) = \Psi(n\psi)\ for\ all\ integers\ n \geq M - 1.$$

Setting $k = (M-1)\psi$, the previous equality can be rewritten as

$$\Psi(k) = \Psi(k + j\psi) \; for \; all \; j \geq 1. \tag{3.18}$$

So we have that $\Psi(k)$ consists of all M nodes of the chain. We now want to show that $\Psi(k) = \Psi(k+1)$. Let us consider an integer t such that $t \neq \psi$ and $P_{ii}^t > 0$. We rewrite (3.15) using k instead of m and t instead of ψ:

$$\Psi(k) \subseteq \Psi(k+t) \subseteq \Psi(k+2t) \subseteq ... \subseteq \Psi(k+\psi t) \tag{3.19}$$

Since $\Psi(k) = \Psi(k+\psi t)$, we obtain that

$$\Psi(t) = \Psi(k+t). \tag{3.20}$$

Let us define s as the smallest integer such that

$$\Psi(k) = \Psi(k+s). \tag{3.21}$$

We want to show that $s = 1$ using the indirect proof. Let's suppose $s \neq 1$. We know that $s = \psi$, so we must have $1 < s \leq \psi$. Since the chain is ergodic, which implies that it is also aperiodic, it exists t such that s does not divide t and $P_{ij}^t > 0$. We can write $t = js + l$, where $1 \leq l < s \;\; and \;\; j \geq 0$. We use iteration on (3.21) to get $\Psi(k) = \Psi(k + js)$. We then apply (3.17) to this,

$$\begin{aligned} \Psi(k+l) &= \Psi(k+js+l) \\ &= \Psi(k+t) \\ &= \Psi(k). \end{aligned} \tag{3.22}$$

We have used that $t = js+l$ followed by (3.20). Here we have a contradiction, since $l < s$. Thus $s = 1$ and $\Psi(k) = \Psi(k+1)$. Iterating this,

$$\Psi(k) = \Psi(k+n) \; for \; all \; n \geq 0. \tag{3.23}$$

Since in an ergodic chain it is possible to go from every state to every other

state, each state j continues to be accessible after k steps. Therefore it must be that, for every state j, $j \in \Psi(k+n)$. Since from (3.22) $\Psi(k+n) = \Psi(k)$, this implies that $j \in \Psi(k)$. This holds for $\forall j$, so we must have that $\Psi(k)$ must be the entire set of states. Thus, $P_{ij}^n > 0$ $\forall n$ and $\forall j$.
This same argument can be applied to any state i on the given cycle with ψ nodes.
Any state m not on this cycle has a path to the cycle using at most $M - \psi$ steps. This path can be used to reach a node i on the cycle. We connect this walk this with all the walks from i of length $k = (M-1)\psi$. We have that:

$$P_{mj}^{M-\psi+(M-1)\psi} > 0 \; for \; all \; j, \; m. \tag{3.24}$$

The proof is complete, since $M - \psi + (M-1)\psi \le (M-1)2 + 1$ for all ψ, $1 \le \psi \le M-1$, with equality when $\psi = M-1$.

If $[P]$ is the matrix of an M state ergodic Markov chain, the previous theorem implies that for each h such that $h \ge (M-1)^2 + 1$, $[P^h]$ is a positive matrix. We choose $h = (M-1)^2 + 1$, so that we can apply theorem (3.3) to $[P^h]$. So, from equation (3.9) of theorem (3.2), it holds that:

$$\lim_{n\to\infty} \max_l P_{lj}^{hm} = \lim_{n\to\infty} \min_l P_{lj}^{hm} = \pi_j > 0$$

Let $\gamma \in \mathbb{N}$ such that $\gamma \ge 1$ and $\beta = \min_{i,j} P_{ij}^h$, then:

$$\max_i P_{ij}^{h(\gamma+1)} - \min_i P_{ij}^{h(\gamma+1)} \le (\max_l P_{lj}^{h\gamma} - \min_l Plj^{h\gamma})(1-2\beta) \tag{3.25}$$

$$(\max_l P_{lj}^{h\gamma} - \min_l P_{lj}^{h\gamma}) \le (1-2\beta)^\gamma \tag{3.26}$$

$$\lim_{\gamma\to\infty} \max_l P_{lj}^{h\gamma} = \lim_{\gamma\to\infty} \min_l P_{lj}^{h\gamma} > 0. \tag{3.27}$$

But what about the values of n that are not multiples of h? We want to write equation (3.27) with the limit in n rather than in γ. We can do such a replacement thanks to lemma (3.1). Since $\max_i P_{ij}^n$ is non-increasing in n, we have that it must have the same limit as $\max_i P_{ij}^{h\gamma}$. Similarly, $\min_i P_{ij}^n$ must have the same limit as $\max_i P_{ij}^{h\gamma}$. In particular, setting $n = h\gamma$, then (3.26) and (3.27) become:

$$(\max_l P_{lj}^n - \min_l P_{lj}^n) \leq (1-2\beta)^{n/h}$$

$$\lim_{n\to\infty} \max_l P_{lj}^n = \lim_{n\to\infty} \min_l P_{lj}^n > 0.$$

We define $\vec{\pi} > 0$ such that for each column j of $\vec{\pi}$ we have:

$$\pi_j = \lim_{n\to\infty} \max_l P_{lj}^n = \lim_{n\to\infty} \min_l P_{lj}^n$$

We see that π_j is independent from both n and the rows of the j^{th} column of $[P]$, so we can simply write:

$$\pi_j = \lim_{n\to\infty} P_{lj}^n \ for\ each\ l, j.$$

We can now state the following theorem:

Theorem 3.5. *Let $[P]$ be the matrix of an ergodic finite state Markov chain. Than there is a unique steady state vector $\vec{\pi}$, which is positive and it is such that:*

$$\pi_j = \lim_{n\to\infty} P_{lj}^n \ \ for\ \ each\ \ l, j$$

and, given $\vec{e} = [1,\ 1,\ 1,\ ...,\ 1]^T$

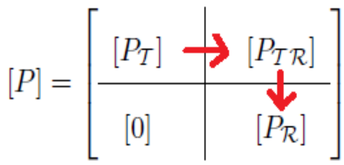
$$\lim_{n\to\infty} P^n = \vec{e}\,\vec{\pi} = \begin{bmatrix} 1 \\ 1 \\ ... \\ 1 \end{bmatrix} * [\pi_1,\ \pi_2,\ ...,\ \pi_M] = \begin{bmatrix} \pi_1 & \pi_2 & ... & \pi_M \\ \pi_1 & \pi_2 & ... & \pi_M \\ ... & ... & ... & ... \\ \pi_1 & \pi_2 & ... & \pi_M \end{bmatrix}$$

Proof: We have to prove the uniqueness of such limit. Let $\vec{\mu}$ be another steady state vector. Therefore $\vec{\mu}$ must be such that $\vec{\mu} = \vec{\mu}[P]$ and for all $n > 1$ it must hold that $\vec{\mu} = \vec{\mu}[P^n]$. Taking the limit, we obtain:

$$\vec{\mu} = \vec{\mu} \lim_{n\to\infty} P^n = \vec{\mu}\, \vec{e}\, \vec{\pi} = \vec{\pi}.$$

3.3 Ergodic unichains

We just saw that for ergodic chains, $\lim_{n\to\infty} P_{ij}^n = \pi_j$ for all i, where $\vec{\pi}$ is a probability vector. Moreover, $\vec{\pi}$ is a steady-state vector and unique solution to $\vec{\pi}[P] = \vec{\pi}$ (theorem (3.5)).
This result can be extended to ergodic unichains, which are ergodic Markov chains with the addition of some transient states (definition (3.4)).



where

$$[P_{\mathcal{T}}] = \begin{bmatrix} P_{11} & \cdots & P_{1t} \\ \cdots & \cdots & \cdots \\ P_{t1} & \cdots & P_{tt} \end{bmatrix}$$

$$[P_{\mathcal{TR}}] = \begin{bmatrix} P_{1,t+1} & \cdots & P_{1,t+r} \\ \cdots & \cdots & \cdots \\ P_{t,t+1} & \cdots & P_{t,t+r} \end{bmatrix} \qquad [P_{\mathcal{R}}] = \begin{bmatrix} P_{t+1,t+1} & \cdots & P_{t+r,t+1} \\ \cdots & \cdots & \cdots \\ P_{t+r,t+1} & \cdots & P_{t+r,t+r} \end{bmatrix}$$

Figure 3.5: The transition matrix of an ergodic unichain. $[P_R]$ is the matrix associated with the ergodic class, while $[P_T]$ is the one associated with the transient class. Notice that the matrix $[P]$ has a block of zeros, since we do not go from a recurrent class to a transient one.

First, suppose that there is one transient class that contains just one state. If a state is in a singleton transient class, then there is a fixed probability, say α, of leaving the class at each step. The probability of remaining in the class for more than n steps in $(1-\alpha)^n$, where $(1-\alpha) < 1$.

As n grows larger, we see that it is very unlikely to remain in the transient state since $\lim_{n\to\infty}(1-\alpha)^n = 0$.

Now we extend the same reasoning also to the case of not just a singleton transient state but of a set of transient states. Notice that each transient state has at least one path to a recurrent state and is very likely that one of those paths will be taken. Also in this scenario, the probability of remaining in a set of transient states goes to 0 as n goes to infinity.
Since we are dealing with ergodic unichains, this means that the ergodic class will be reached. From then on, we have already shown the results about convergence.
For every state i then,

$$\lim_{n\to\infty} \max_i P_{ij}^n = \lim_{n\to\infty} \min_i P_{ij}^n = \pi_j$$

where $\pi_j = 0$ for each transient state and $\pi_j > 0$ for each recurrent state.

Let T denote the set of transient states and assume that these states are numbered $1, 2, ..., t$. Let R denote the recurrent class, whose states are numbered $t+1, ..., t+r$. As we will now show, there is a tendency to move from the transient to the recurrent states. Thus,

$$\lim_{n\to\infty} P_{ij}^n = 0 \; for \; i, \; j \in T. \tag{3.28}$$

For each transient state, there must be a walk to a recurrent state. Since there are only t transient states, the longest path from a transient state to a recurrent one must have at most t steps. Each path has a positive probability, thus for each $i \in T$, $\sum_{j\in R} P_{ij}^t > 0$. This implies that for each $i \in T$, $\sum_{j\in T} P_{ij}^t < 1$. Let γ denote the maximum of these probabilities over $i \in T$. In other terms, let $\gamma = \max_{i\in T} \sum_{i\in T} P_{ij}^t < 1$.

Lemma 3.6. *Let $[P]$ be the transition matrix of a unichain with a set T of* *t transient states. Then*

$$\max_{l \in T} \sum_{j \in T} P_{lj}^{n} \leq \gamma^{n/t}. \tag{3.29}$$

Proof: For each integer δt multiple of t and each $i \in T$:

$$\sum_{j \in T} P_{ij}^{(\delta+1)t} = \sum_{k \in T} P_{ik}^{t} \sum_{j \in T} P_{kj}^{\delta t} \leq \sum_{k \in T} P_{ik}^{t} \max_{l \in T} \sum_{j \in T} P_{lj}^{\delta t} \leq \gamma \max_{l \in T} \sum_{j \in T} P_{lj}^{\delta t}.$$

This is valid for all $i \in T$, which includes also the maximum over i. Thus, it holds:

$$\max_{l \in T} \sum_{j \in T} P_{lj}^{t} \leq \gamma \tag{3.30}$$

Iterating (3.29), we obtain:

$$\max_{l \in T} \sum_{j \in T} P_{lj}^{\delta t} \leq \gamma^{\delta}.$$

Setting $n = \delta t$ and recalling that the maximum in non-increasing in n, (3.28) follows.

We now deal with the case in which the initial state in $i \in T$ and the final state is $j \in R$. Let us define $m = n/2$. For each $i \in T$ and $j \in R$, the Chapman-Kolmogorov equation says that:

$$P_{ij}^{n} = \sum_{k \in T} P_{ik}^{m} P_{kj}^{n-m} + \sum_{k \in R} P_{ik}^{m} P_{kj}^{n-m}.$$

From what we proved in section 3.2, for every $j \in R$ it exists $\pi_j = \lim_{n \to \infty} P_{kj}^{n}$. Thus, for each $i \in T$, it holds that:

$$|P_{ij}^{n} = \pi_j| = |\sum_{k \in T} P_{ik}^{m} (P_{kj}^{n-m} - \pi_j) + \sum_{k \in R} P_{ik}^{m} (P_{kj}^{n-m} - \pi_j)|$$

$$\leq \sum_{k \in T} P_{ik}^{m} |P_{kj}^{n-m} - \pi_j| + \sum_{k \in R} P_{ik}^{m} |P_{kj}^{n-m} - \pi_j|$$

$$\leq \sum_{k\in T} P_{ik}^m + \sum_{k\in R} P_{ik}^m |P_{kj}^{n-m} - \pi_j|. \tag{3.31}$$

We can take the limit of n to infinity in both sides of equation (3.31) and we obtained the desired result.
This is summarized by the following theorem:

Theorem 3.7. *Let $[P]$ be the matrix of an ergodic finite state unichain. Then $\lim_{n\to\infty}[P^n] = \vec{e}\ \vec{\pi}$, where $\vec{e} = [1,\ 1,\ 1,\ ...,\ 1]^T$ and $\vec{\pi}$ is the steady-state vector with entries zero for each transient state and π_j for each recurrent state.*

3.4 Other finite state Markov chains

First consider a Markov chain with several ergodic classes, $C_1, ..., C_m$. The classes don't communicate and should be considered separately.
If this is the case, $[P]$ will have m independent steady state vectors, one nonzero on each class. $[P^n]$ will then converge, but the rows will not be all the same.
There will be m sets of rows, one for each class, and the row for class k will be nonzero for the elements of that class. So the steady-state vector will be in blocks.

Next consider a periodic recurrent chain of period d. This can be separated into d subclasses with a cyclic rotation between them.
If we look at $[P^d]$, we see that each subclass becomes an ergodic class, say $C_1, , ..., C_d$. Thus, $\lim_{n\to\infty}[P^{nd}]$ exists.
A steady state is reached within each subclass, but the chain rotates from one subclass to another.

4 Applications

Markov chains are widely used to model many real-world situations that involve randomness. They are a tool for a variety of different fields, that range from search engines' page ranking to the study of genes in biology. In economics and finance, they are often used to model randomness and to predict the value of assets and the evolution of macroeconomic situations, such as cycles between recession and expansion.

4.1 Markov chains to predict stock market trends

I am going to find a model to predict stock market fluctuations using Markov chains. I start by giving a brief insight to financial markets and I continue by using some data provided by Borsa Italiana to apply the theoretical model.

A *stock* is the ownership of a corporation indicated by shares, which represent a piece of the corporation's assets and earnings.
The are different *stock indexes* that track a portfolio of stocks.
I am going to consider FTSE MIB Index. It is "the primary benchmark Index for the Italian equity markets. Capturing approximately 80% of the domestic market capitalization, the Index is comprised of highly liquid, leading companies across ICB sectors in Italy. The FTSE MIB Index measures the performance of 40 Italian equities and seeks to replicate the broad sector weights of the Italian stock market" [2].

As stock prices increase and decrease, financial markets *trends* can identified. They can be grouped into three categories:

1. *Bull markets*: are characterized by a general rise of prices of financial activities and by optimistic expectations.

[2] Borsa Italiana

Figure 4.1: FTSE-MIB Index graph on August the 1st, 2019

2. *Bear markets*: are characterized by a decline in prices and by a pessimistic view of the future.

3. *Stagnant markets*: are characterized by neither a decline nor rise in general prices.

Normally, it is assumed that all the actors in stock markets have access to the same information and that prices fluctuate randomly. For this particular model, I am going to make some further assumptions:

1. The trend of a certain stock today depends only the state of the stock on the day before yesterday. It has little to do with the past. I make this assumption so that I can apply the Markov property.

2. The probability of a stock to go from state i to state j has nothing to do with the time at which we consider state i. This assumption makes it possible to use homogeneous Markov chains.

4.1.1 Empirical analysis

In order to develop this model, I am going to divide each day's closing price of each considered stock into three states: up, down and zero. The up state

corresponds to a positive increase of the price from the day before yesterday to today's price. The down state corresponds to a decrease in price and the zero state corresponds neither to an increase nor a decrease.
I use the following notation: $x_1 = up$, $x_2 = down$, $x_3 = zero$. I denote with $\eta_i = [p_1,\ p_2,\ p_3]$ the vector of probabilities at time $i = 0,\ 1,\ 2,\ ...,\ n$. Then I proceed by determining the probability of each state.
From table 1, we see that: $x_1 = 9$, $x_2 = 29$, $x_3 = 2$. Thus, the initial state distribution is: $p_1 = 9/40 \approx 0.225$, $p_2 = 29/40 \approx 0.725$, $p_3 = 2/40 \approx 0.005$ and the initial state distribution vector is: $\vec{\eta_0} = [0.225,\ 0.725,\ 0.005]$.

Now I want to establish the transition probability matrix. For this purpose, I use table 2, that contains variations of prices of the stocks in the Index on July the 31^{st}. In order to find P_{ij} for $i,\ j$=1, 2, 3, I need to count how many times the stocks in the basket change from up to down, from up to zero, from up to up, from down to up, ... and so on. Then I divide each sum by the values of x_1, x_2 and x_3 previously found. So I find that:
$P_{11} = 3/9 \approx 0.333$, $P_{12} = 5/9 \approx 0.555$, $P_{13} = 1/9 \approx 0.111$,
$P_{21} = 20/29 \approx 0.690$, $P_{22} = 9/29 \approx 0.310$, $P_{23} = 0/29 = 0$,
$P_{31} = 2/2 = 1$, $P_{32} = 0/2 = 0$, $P_{33} = 0/2 = 0$.

Thus, the transition matrix is: $P = \begin{bmatrix} 0.333 & 0.555 & 0.111 \\ 0.690 & 0.310 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.
Notice that we have a stochastic matrix since the entries of each row sum up to 1.
Moreover, the associated chain is ergodic since there is just one class, whose states are recurrent and aperiodic, as it is made clear by the graphical model of figure (4.2). Thus, it exists the steady-state vector.

Now I calculate the state of the Index in the future, using the following equation:

$$\vec{\eta_{i+1}} = \vec{\eta_i} * P,\ where\ i \in \mathbb{N}. \tag{4.1}$$

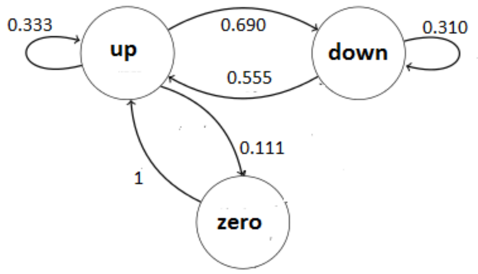


Figure 4.2: Graphical model

Using data from tables 1 and 2 in equation (4.1), I get:

$$\vec{\eta_1} = \vec{\eta_0} * P = [0.225,\ 0.725,\ 0.005] * \begin{bmatrix} 0.333 & 0.555 & 0.111 \\ 0.690 & 0.310 & 0 \\ 1 & 0 & 0 \end{bmatrix} = [0.58,\ 0.35,\ 0.025]$$

As we saw during the previous sections, under the hypothesis of stable conditions, the state probability vector is independent from the initial distribution as n grows larger.
So, iterating equation (4.1) over i, I find that:

$$\vec{\eta_2} = \vec{\eta_1} * P = [0.58,\ 0.35,\ 0.025] * \begin{bmatrix} 0.333 & 0.555 & 0.111 \\ 0.690 & 0.310 & 0 \\ 1 & 0 & 0 \end{bmatrix} = [0.460,\ 0.431,\ 0.064]$$

and so on.
I used the following Python code in order to calculate $\vec{\eta_n}$ with a large $n = 100$:

```
n = 1
e = [0.225, 0.725, 0.005]
while n < 101:
     a = e[0]
    b = e[1]
    c = e[2]
    e[0] = a * (1/3) + b * (20/29) + c * 1
```

```
    e[1] = a * (5/9) + b * (9/29) + c * 0
    e[2]= a * (1/9) + b * 0 + c * 0
    n = n + 1
print(e)
```

After the code run, Python found the solution vector: [0.498, 0.401, 0.055]. Thus I can conclude that the predictions for the future are optimistic, since the probability of stocks in the Index is up for 49.8% of cases. However the probability of a decrease in prices is also quite high and it is equal to 40.1%.

4.1.2 Comments

Are predictions reliable? On the second of August we have that prices decreased. Thus, predictions were not very reliable.
There are many factors that I did not take into account while developing this model. It does account for the magnitude of price changes and it does not consider possible market caps or market floors.
Moreover, the main difficulty for building a decent model lies on the fact that stock prices are really volatile. They change every day as a result of market forces. If the demand for a certain stock rises, then the price moves up. On the contrary, if supply increases because investors want to sell a stock, the price falls. The primary challenge is to understand why people want to buy or sell a stock. The main theory is that prices reflect how investors perceive the worth of the company that issues those stocks. Not only the current value of the company is taken into account, but also how investors expect the company to grow in the future.
The value of a company is affected mostly by its earnings. Public companies, which are companies whose securities can be bought by the general public, are required to report their earnings four times a year. If a company's report is better than expected, the price of its stock rises. On the other side, if earnings are not as high as prospected, the price of stocks falls down. Thus, in this model, the matrix P should be changed every quarter of a year to reflect the effects of the disclosure of companies' earnings. Moreover, good or bad

news about a company immediately influence the price of its stocks. Thus, the matrix P should also be modified every time some information about the companies in the Index or about the general sociopolitical economical situation is made public. However, this would still not be enough to face the complexity of the real world. As it happened with financial bubbles, earnings and expectations are not enough to predict the behavior of investors. Hundreds of variables, ratios and indicators have been developed in order to predict fluctuations. Still, some believe that it is impossible to make forcasts.

5 Tables with data

Number	1	2	3	4	5	6	7	8	9	10
State	up; +0.22	zero	zero	down; -0.64	down; -1.73	down; -1.37	down; -0.34	down; -0.94	down; -0.96	up; +0.49

Number	11	12	13	14	15	16	17	18	19	20
State	up; +0.58	down; 0.63	down; -0.06	down; -0.64	down; -0.97	down; -1.67	up; +0.24	up; +0.18	down; -0.94	down; -0.10

Number	21	22	23	24	25	26	27	28	29	30
State	down; -1.06	down; -2.10	down; -0.21	down; -0.97	down; -0.46	down; -2.50	up; +1.93	down; -3.30	down; -0.12	down; -3.27

Number	31	32	33	34	35	36	37	38	39	40
State	down; -0.28	up; +0.53	down; -1.87	up; +1.35	down; -1.24	up; +0.90	down; -0.50	down; -2.80	down; -1.31	down; -1.21

Table 1: FTSE-MIB from Borsa Italiana on July the 29^{th}, 2019. In the row 'Number', each entry corresponds to one of the companies that compose the basket of the Index. To each company, it is associated a state (up, zero or down) and the magnitude of the change in percentage points.

Number	1	2	3	4	5	6	7	8	9	10
State	down; -0.54	up; +5.11	up; +0.35	up; +1.14	up; +0.87	down; -0.32	up; +1.31	down; -1.65	up; +1.29	up; +3.36

Number	11	12	13	14	15	16	17	18	19	20
State	down; -0.44	up; +0.64	up; +2.23	up; +1.07	up; +4.10	down; -1.06	zero	down; -0.54	down; -0.01	down; -1.32

Number	21	22	23	24	25	26	27	28	29	30
State	up; +1.31	up; +0.36	down; -0.42	down; -1.17	up; +0.25	up; +1.54	up; +3.23	up; +1.57	up; +1.08	up; +2.81

Number	31	32	33	34	35	36	37	38	39	40
State	down; -7.36	down; -0.63	up; +0.93	up; +0.75	up; +2.58	down; -1.41	up; +0.26	down; -0.11	up; +0.69	up; +0.21

Table 2: FTSE-MIB from Borsa Italiana on July the 31^{st}, 2019.

1. A2a
2. Amplifon
3. Atlantia
4. Azimut Holding
5. Banco Bpm
6. Bper Banca
7. Buzzi Unicem
8. Campari
9. Cnh Industrial
10. Diasorin
11. Enel
12. Eni
13. Exor
14. Ferrari
15. Fiat Chrysler Automobiles
16. Finecobank
17. Generali
18. Hera
19. Intesa Sanpaolo
20. Italgas
21. Juventus Football Club
22. Leonardo
23. Mediobanca
24. Moncler
25. Nexi
26. Pirelli e Co
27. Poste Italiane
28. Prysmian
29. Recordati
30. Saipem
31. Salvatore Ferragamo
32. Snam
33. Stmicroelectronics
34. Telecom Italia
35. Tenaris
36. Terna - Rete Elettrica Nazionale
37. Ubi Banca
38. Unicredit
39. Unipol
40. Unipolsai

Figure 5.1: List of companies used in the FTSE-MIB Index

6 First Appendix

Lemma 6.1. *If* $a,\ b \in \mathbb{N} \setminus \{0\}$ *are relatively prime, then*

$$ax + by = ab - a - b \tag{6.1}$$

has no solutions $x,\ y \in \mathbb{N}$.

Proof Equation (6.1) can be rewritten as:

$$b(y+1) = a(b-1-x) \tag{6.2}$$

or

$$a(x+1) = b(a-1-y). \tag{6.3}$$

Considering equation (6.2), since a and b have no common factor, it exists $k \in \mathbb{N},\ k \geq 1$ such that: $y+1 = ka$ and $b-1-x = kb$. This is equivalent to

$$y + 1 = ka \tag{6.4}$$

and

$$x + 1 = b(1-k). \tag{6.5}$$

Equation (6.5) can hold only if $k = 0$. In this such case, equation (6.4) is equivalent to $y = -1$, so we found a contradiction.
Starting from equation (6.3), similar steps also lead to a contradiction.

Lemma 6.2. *(Euclidean Algorithm). If* $a, b \in \mathbb{N} \setminus \{0\}$*, then* $\exists x, y \in \mathbb{Z}$ *such that* $ax + by = 1$.

Proof We set $a = a_0 > b = a_1$ and we perform the division with reminder. Thus, we obtain: $a_0 = a_1 q_1 + a_2$, where $a_1,\ a_2 \in \mathbb{N}$ *and* $0 \leq a_2 < a_1$. Iterating this, we obtain that: $a_{j-1} = a_j q_j + a_{j+1}$. For a certain $j = n$, one will have

that $a_{n+1} = 0$. If $q \in \mathbb{N} \setminus \{0\}$ divides a_n then, iterating backward, it divides $a_{n-1},\ a_{n-2},\ ...,\ a_1 = b,\ a_0 = a$. Thus, $a_n = 1$, since G.D.C.(a,b)=1.
Thus, we have that:

$$\begin{cases} a_0 & = a_1 q_1 + a_2 \\ a_1 & = a_2 q_2 + a_3 \\ \cdots & \\ a_{n-2} & = a_{n-1} q_{n-1} + 1 \end{cases} \Leftrightarrow \begin{cases} 1 & = a_{n-2} - a_{n-1} q_{n-1} \\ a_{n-1} & = a_{n-3} - a_{n-2} q_{n-2} \\ \cdots & \\ a_3 & = a_1 - a_2 q_2 \\ a_2 & = a_0 - a_1 q_1 \end{cases}$$

Starting from the bottom and substituting in the equation above, inductively, we see that we have a linear combination of $a_0 = a$ and $a_1 = b$ with integer coefficients.
Let's observe that the same argument shows, more generally, that G.D.C.(a,b)=ax+by $\exists x,\ y \in \mathbb{Z}$.

Theorem 6.3. *Let $a,\ b \in \mathbb{N}$ such that the greatest common divisor between a and b is 1, i.e. $G.D.C.(a,\ b) = 1$.*
If $m \in \mathbb{N},\ m \geq ab - a - b + 1$, then $\exists x,\ y \in \mathbb{N}$ such that $m = ax + by$ and there are no $x,\ y \in \mathbb{N}$ such that $ax - by = ab - a - b$.

Proof Lemma (6.1) proves the last assertion in the statement.
By lemma (6.2), there are $x',\ y' \in \mathbb{Z}$ such that $1 = ax' + by'$. Clearly x' and y' have opposite signs and we can assume that $1 = ax_1 - by_1$, with $x_1,\ y_1 \in \mathbb{N} \setminus \{0\}$. For a certain $j \in \mathbb{Z}$ and a certain $n \in \mathbb{N} \setminus \{0\}$, we have that: $0 < n = anx_1 - bny_1 = anx_1 + jab - bny_1 - jab = ax_n - by_n$
with $\begin{cases} x_n & = nx_1 + jb \\ y_n & = ny_1 + ja. \end{cases}$

We can find j so that it holds: $\begin{cases} anx_1 + jab > 0 \\ bny_1 + jab < ab. \end{cases}$

Let's consider the minimal $j \in \mathbb{Z}$ with the property that $anx_1 + jab > 0$. Then $anx_1 + jab \leq ab$, so that $bny_1 + jab < anx_1 + jab < ab$. Thus $x_n \geq 1$ and $y_n \leq a - 1$.
Finally, it holds that: $b(a - 1 - y_n) + a(x_n - 1) = ba - b - a + ax_n - by_n =$

$ba-b-a+n = m$. If we choose $n = m-(ba-b-a) \geq 1$, which is a consequence of the hypothesis. The theorem is proved because $y = a - 1 - y_n \geq 0$ and $x = x_n - 1 \geq 0$.

7 References

1. R.G. Gallager, “finite state Markov chains” in stochastic processes: theory for applications. United Kingdom: Cambridge university press, 2013.

2. F.H.C. Marriott, "A dictionary of statistical terms", 5^{th} edition, Longman Scientific Technical, 1989.

3. D. Zhang, X. Zhang, “Study on forecasting the stock market trend based on stochastic analysis method”, International Journal of Business and Management, June 2009, vol. 4, nr. 6.

4. A. Pascucci, "‘Dispense del corso Probabilità e Statistica"’, 2019. http://www.dm.unibo.it/ pascucci

5. Borsa Italiana. https://www.borsaitaliana.it/borsa/azioni/ftse-mib/lista.html?page=1