Chapter 11

Markov Chains

This section briefly presents some fundamental results concerning the theory of Markov chains with a finite number of states. These results will be used in the following chapter. We will use the usual terminology introduced by Chung (1960) and Parzen (1962).

11.1. Definitions

Let us consider an economic or physical system S with m possible states, represented by the set I :

$$I = \{1, \ 2, \dots, \ m\} . \tag{11.1}$$

Let the system S evolve randomly in discrete time $(t = 0, \ 1, \ 2, \dots, \ n, \dots)$, and let J_n be the r.v. representing the state of the system S at time n.

Definition 11.1 *The random sequence* $(J_n, \ n \in \mathrm{N})$ *is a Markov chain if and only if for all* $j_0, \ j_1, \dots, j_n \in I$:

$$P(J_n = j_n \mid J_0 = j_0, \ J_1 = j_1, \dots, \ J_{n-1} = j_{n-1}) = P(J_n = j_n \mid J_{n-1} = j_{n-1}) \tag{11.2}$$

(*provided this probability has meaning*).

Definition 11.2 *A Markov chain* $(J_n, \ n \geq 0)$ *is homogenous if and only if probabilities* (1.2) *do not depend on* n *and are non-homogenous in the other cases.*

For the moment, we will only consider the homogenous case for which we write:

$$P(J_n = j \mid J_{n-1} = i) = p_{ij}, \tag{11.3}$$

and we introduce matrix **P** defined as:

$$\mathbf{P} = [\, p_{ij} \,]. \tag{11.4}$$

The elements of matrix **P** have the following properties:

(i) $p_{ij} \geq 0,\ $ for all $i, j \in I,$ (11.5)

(ii) $\sum_{j \in I} p_{ij} = 1,\ $ for all $i \in I.$ (11.6)

A matrix **P** satisfying these two conditions is called a *Markov matrix* or a *transition matrix*.

With every transition matrix, we can associate a *transition graph* where vertices represent states. There exists an *arc* between vertices *i* and *j* if and only if $p_{ij} > 0.$

To fully define the evolution of a Markov chain, it is also necessary to fix an *initial distribution* for state J_0, i.e. a vector

$$\mathbf{p} = (p_1, \ldots, p_m), \tag{11.7}$$

such that:

$$p_i \geq 0, \qquad i \in I, \tag{11.8}$$

$$\sum_{i \in I} p_i = 1. \tag{11.9}$$

For all i, p_i represents the *initial probability* of starting from i:

$$p_i = P(J_0 = i). \tag{11.10}$$

For the rest of this chapter we will consider homogenous Markov chains as being characterized by the couple $(\mathbf{p}, \mathbf{P})$.

If $J_n = i$ a.s., that is, if the system starts with probability equal to 1 from state i, then the components of vector **p** will be:

$$p_j = \delta_{ij}. \tag{11.11}$$

We now introduce the *transition probabilities of order* $p_{ij}^{(n)}$, defined as:

$$p_{ij}^{(n)} = P(J_{\nu+n} = j \mid J_\nu = i). \tag{11.12}$$

From the Markov property (11.2), it is clear that conditioning with respect to $J_{\nu+1}$, and we obtain

$$p_{ij}^{(2)} = \sum_k p_{ik} p_{kj}. \tag{11.13}$$

Using the following matrix notation:

$$\mathbf{P}^{(2)} = \left[p_{ij}^{(2)} \right], \tag{11.14}$$

we find that relation (11.13) is equivalent to

$$\mathbf{P}^{(2)} = \mathbf{P}^2. \tag{11.15}$$

Using induction, it is easy to prove that if

$$\mathbf{P}^{(n)} = \left[p_{ij}^{(n)} \right], \tag{11.16}$$

then we obtain for all $n \geq 1$:

$$\mathbf{P}^{(n)} = \mathbf{P}^n. \tag{11.17}$$

Note that (11.17) implies that the transition probability matrix in n steps is equal to the *n*th power of matrix **P**.

For the marginal distributions related to J_n, we define for $i \in I$ and $n \geq 0$:

$$p_i(n) = P(J_n = i). \tag{11.18}$$

These probabilities may be calculated as follows:

$$p_i(n) = \sum_j p_j p_{ji}^{(n)}, \qquad i \in I. \tag{11.19}$$

If we write:

$$p_{ji}^{(0)} = \delta_{ji} \text{ or } \mathbf{P}^{(0)} = \mathbf{I}, \tag{11.20}$$

then relation (11.19) is true for all $n \geq 0$.

If:

$$\mathbf{p}(n) = (p_1(n), \ldots, p_m(n)), \tag{11.21}$$

then relation (11.19) can be expressed, using matrix notation, as:

$$\mathbf{p}(n) = \mathbf{p}\mathbf{P}^n. \tag{11.22}$$

Definition 11.3 *A Markov matrix* **P** *is regular if there exists a positive integer* k*, such that all the elements of matrix* $\mathbf{P}^{(k)}$ *are strictly positive.*

From relation (11.17), **P** is regular if and only if there exists an integer $k > 0$ such that all the elements of the *k*th power of **P** are strictly positive.

Example 11.1

(i) If:

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 \\ 1 & 0 \end{bmatrix} \tag{11.23}$$

we have:

$$\mathbf{P}^2 = \begin{bmatrix} 0.75 & 0.25 \\ 0.5 & 0.5 \end{bmatrix} \tag{11.24}$$

so that **P** is regular.

The transition graph associated with **P** is given in Figure 11.1.

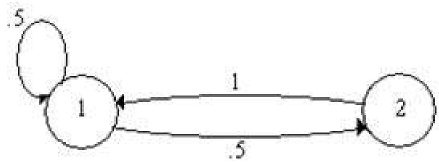

Figure 11.1. *Transition graph of matrix (11.23)*

(ii) If:

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0.75 & 0.25 \end{bmatrix}, \tag{11.25}$$

P is not regular, because for any integer k,

$$p_{12}^{(k)} = 0. \tag{11.26}$$

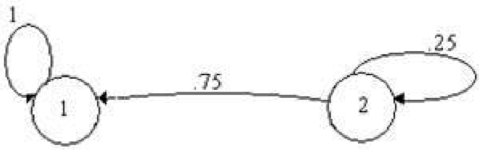

Figure 11.2. *Transition graph for matrix (11.25)*

The transition graph in this case is depicted in Figure 11.2.

The same is true for the matrix:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{11.27}$$

(iii) Any matrix **P** whose elements are all strictly positive is regular.

For example:

$$\begin{bmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix} \qquad \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.6 & 0.2 & 0.2 \\ 0.4 & 0.1 & 0.5 \end{bmatrix}. \tag{11.28}$$

11.2. State classification

Let $i \in I$, and let $d(i)$ be the greatest common divisor of the set of integers n, such that

$$p_{ii}^{(n)} > 0. \tag{11.29}$$

Definition 11.4 *If $d(i) > 1$, the state i is said to be periodic with period $d(i)$. If $d(i) = 1$, then state i is aperiodic.*

Clearly, if $p_{ii} > 0$, then i is aperiodic. However, the converse is not necessarily true.

Remark 11.1 If **P** is regular, then all the states are aperiodic.

Definition 11.5 *A Markov chain whose states are all aperiodic is called an aperiodic Markov chain.*

From now on, we will have only Markov chains of this type.

Definition 11.6 *A state i is said to lead to state j (written $i \rhd j$) if and only if there exists a positive integer n such that*

$$p_{ij}^{n} > 0. \tag{11.30}$$

i C j means that i does not lead to j.

Definition 11.7 *States i and j are said to communicate if and only if $i \rhd j$ and $j \rhd i$, or if $j = i$. We write $i \lhd\rhd j$.*

Definition 11.8 *A state i is said to be essential if and only if it communicates with every state it leads to; otherwise it is called inessential.*

Relation $\lhd\rhd$ defines an equivalence relation over the state space I resulting in a partition of I. The equivalence class containing state i is represented by $C(i)$.

Definition 11.9 *A Markov chain is said to be irreducible if and only if there exists only one equivalence class.*

Clearly, if **P** is regular, the Markov chain is both irreducible and aperiodic. Such a Markov chain is said to be *ergodic*.

It is easy to show that if the state i is essential (inessential), then all the elements of class $C(i)$ are essential (inessential) (see Chung (1960)).

We can thus speak of essential and inessential classes.

Definition 11.10 *A subset* E *of the state space* I *is said to be closed if and only if:*

$$\sum_{j\in E} p_{ij} = 1 \text{, for all } i \in E . \tag{11.31}$$

It can be shown that every essential class is minimally closed; see Chung (1960).

Definition 11.11 *For given states* i *and* j*, with* $J_0 = i,$ *we can define the r.v.* τ_{ij} *called the first passage time to state j as follows:*

$$\tau_{ij} = \begin{cases} n & \text{if } J_\nu \neq j, \quad 0 < \nu < n, \quad J_n = j, \\ \infty & \text{if } J_\nu \neq j, \quad \text{for all } \nu > 0. \end{cases} \tag{11.32}$$

τ_{ij} is said to be the *hitting time* of the singleton $\{ j \}$, starting from state i at time 0.

Assuming:

$$f_{ij}^{(n)} = P\left(\tau_{ij} = n \mid J_0 = i \right), \qquad n \in \mathbb{N}_0 \tag{11.33}$$

and

$$f_{ij} = P\left(\tau_{ij} < \infty \mid J_0 = i \right), \tag{11.34}$$

we can see that for $n > 0$:

$$f_{ij}^{(n)} = P\left(J_n = j, \quad J_\nu \neq j, \quad 0 < \nu < n \mid J_0 = i \right), \tag{11.35}$$

$$= \sum_{S'_{n,i,j}} \prod_{k=0}^{n-1} p_{\alpha_k \alpha_{k+1}}, \tag{11.36}$$

where the summation set $S'_{n,i,j}$ is defined as:

$$S'_{n,i,j} = \{(\alpha_0, \alpha_1, \ldots, \alpha_n) : \alpha_0 = i, \alpha_n = j, \alpha_k \in I, \alpha_k \neq j, k = 1, \ldots, n-1\}. \tag{11.37}$$

We also have:

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}, \tag{11.38}$$

$$1 - f_{ij} = P(\tau_{ij} = \infty \mid J_0 = i). \tag{11.39}$$

Elements $f_{ij}^{(n)}$ can readily be calculated by induction, using the following relations:

$$p_{ij} = f_{ij}^{(1)}, \tag{11.40}$$

$$p_{ij}^{(n)} = \sum_{\nu=1}^{n-1} f_{ij}^{(\nu)} p_{jj}^{(n-\nu)} + f_{ij}^{(n)}, \quad n \geq 2. \tag{11.41}$$

Let:

$$m_{ij} = E(\tau_{ij} \mid J_0 = i), \tag{11.42}$$

with the possibility of an infinite mean. The value of m_{ij} is given by:

$$m_{ij} = \sum_{n=1}^{\infty} n f_{ij}^{(n)} - \infty(1 - f_{ij}).^{(*)} \tag{11.43}$$

If $i = j$, then m_{ij} is called the *first passage time mean* or the *mean recurrence* time of state i.

For every j, we define the sequence of successive return times to state j $\left(r_n^{(j)}, \quad n \geq a\right)$ as follows:

(*) Using the following conventions: $\infty + a = \infty, \quad a \in \mathbb{R}, \quad \infty \cdot a = \infty, \quad (a > 0)$, and in this particular case, $\infty \cdot 0 = 0$.

$$r_0^{(j)} = 0, \tag{11.44}$$

$$r_n^{(j)} = \sup_k \left\{ k \in \mathrm{N}_{0,} \quad k > r_{n-1}^{(j)}, \quad J_\nu \neq j, \quad r_{n-1}^{(j)} < \nu < k \right\}, \quad n > 0. \tag{11.45}$$

Using the Markov property and supposing $J_0 = j$, the sequence of return times to state j is a renewal sequence with the r.v.

$$r_n^{(j)} - r_{n-1}^{(j)}, \quad n \geq 1 \tag{11.46}$$

is a sequence of independent r.v. all distributed according to τ_{jj} .

If $J_0 = i, \quad i \neq j$, then the first time of hitting *j* is

$$r_1^{(j)} = \tau_{ij}, \tag{11.47}$$

and

$$r_n^{(j)} - r_{n-1}^{(j)} \sim \tau_{jj}, \quad n > 1. \tag{11.48}$$

Definition 11.12 *A state i is*

$$i \quad \text{transient} \quad \Leftrightarrow \quad f_{ii} < 1, \tag{11.49}$$

$$i \quad \text{recurrent} \quad \Leftrightarrow \quad f_{ii} = 1. \tag{11.50}$$

A recurrent state i is said to be *zero* (*positive*) if $m_{ii} = \infty \quad (m_{ii} < \infty)$. It can be shown that if $m_{ii} < \infty$, then we can only have positive recurrent states.

This classification leads to the decomposition theorem (see Chung (1960)).

Proposition 11.1 (*Decomposition theorem*) *The state space I of any Markov chain can be decomposed into* r $(r \geq 1)$ *subsets* $C_1, \ldots, C_r$ *forming a partition, such that each subset* C_i *is one and only one of the following types*:

(i) an essential recurrent positive closed set;

(ii) an inessential transient non-closed set.

Remark 11.2

(1) If an inessential class reduces to a singleton $\{i\}$, there are two possibilities:

a) there exists a positive integer *N* such that:

$$0 < p_{ii}^N < 1. \tag{11.51}$$

b) the N in a) does not exist. In this case, state i is said to be a *non-return state*.

(2) If singleton $\{i\}$ forms an essential class, then

$$p_{ii} = 1 \tag{11.52}$$

and state i is said to be an *absorbing state*.

(3) If $m = \infty$, there may be two other types of classes in the decomposition theorems:

a) essential transient closed;

b) essential recurrent non-closed classes.

Other works on Markov chains give the following necessary and sufficient conditions for recurrence and transience.

Proposition 11.2

(i) State *i* is transient if and only if

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty. \tag{11.53}$$

In this case, for all $k \in I$:

$$\sum_{n=1}^{\infty} p_{ki}^{(n)} < \infty, \tag{11.54}$$

and in particular:

$$\lim_{n\to\infty} p_{ki}^{(n)} = 0, \quad \forall k \in I. \tag{11.55}$$

(ii) State *i* is recurrent if and only if

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty. \tag{11.56}$$

In this case:

$$k \lhd\rhd i \Rightarrow \sum_{n=1}^{\infty} p_{ki}^{(n)} = \infty, \tag{11.57}$$

and

$$k \,\mathrm{C}\, i \Rightarrow \sum_{n=1}^{\infty} p_{ki}^{(n)} = 0. \tag{11.58}$$

11.3. Occupation times

For any state j, and for $n \in \mathbb{N}_0$, we define the r.v. $N_j(n)$ as the number of times state j is occupied in the first n transitions:

$$N_j(n) = \#\{k \in \{1,\ldots,n\} : J_k = j\}. \tag{11.59}$$

By definition, the r.v. $N_j(n)$ is called the *occupation time of state j in the first n transitions.*

The r.v.

$$N_j(\infty) = \lim_{n\to\infty} N_j(n) \tag{11.60}$$

is called the *total occupation time of state j.*

For any state j and $n \in \mathbb{N}_0$ let us define:

$$Z_j(n) = \begin{cases} 1 & \text{if } J_n = j, \\ 0 & \text{if } J_n \neq j. \end{cases} \tag{11.61}$$

We may write:

$$N_j(n) = \sum_{\nu=1}^{n} Z_j(\nu). \tag{11.62}$$

We have from relation (11.34):

$$\mathrm{P}\left(N_j(\infty) > 0 \mid J_0 = i\right) = f_{ij}. \tag{11.63}$$

Let g_{ij} be the conditional probability of an infinite number of visits to state j, starting with $J_0 = i$; that is:

$$g_{ij} = P\left(N_j(\infty) = \infty \mid J_0 = i \right). \tag{11.64}$$

It can be shown that:

$$g_{ii} = \lim_{n\to\infty} f_{ii}{}^{(n)}, \tag{11.65}$$

$$g_{ij} = f_{ij} \cdot g_{jj}, \tag{11.66}$$

$$g_{ii} = 1 \Leftrightarrow f_{ii} = 1 \Leftrightarrow i \text{ is recurrent,} \tag{11.67}$$

$$g_{ii} = 0 \Leftrightarrow f_{ii} < 1 \Leftrightarrow i \text{ is transient.} \tag{11.68}$$

Results (11.67) and (11.68) can be interpreted as showing that system S will visit a recurrent state an infinite number of times, and that it will visit a transient state a finite number of times.

11.4. Absorption probabilities

Proposition 11.3

(i) If i is recurrent and if $j \in C(i)$, then $f_{ij} = 1$.

(ii) If i is recurrent and if $j \notin C(i)$, then $f_{ij} = 0$.

Proposition 11.4 *Let T be the set of all transient states of I, and let C be a recurrent class.*

For all $j,\ k \in C$,

$$f_{ij} = f_{ik}. \tag{11.69}$$

Labeling this common value as f_{iC}, *the probabilities* $\left(f_{i,C},\ \ i \in T \right)$ *satisfy the linear system*:

$$f_{i,C} = \sum_{k\in T} p_{ik} f_{k,C} + \sum_{k\in C} p_{ik}, \qquad i \in T. \tag{11.70}$$

Remark 11.3 Parzen (1962) proved that under the assumption of Proposition 11.4, the linear system (11.70) has a unique solution. This shows, in particular, that if there is only one irreducible class C, then for all $i \in T$:

$$f_{i,C} = 1. \tag{11.71}$$

Definition 11.13 *The probability* $f_{i,C}$ *introduced in Proposition 11.4 is called absorption probability in class C, starting from state i.*

If class C is recurrent:

$$f_{i,C} = \begin{cases} 1 & \text{if } i \in C, \\ 0 & \text{if } i \text{ is recurrent, } i \notin C. \end{cases} \tag{11.72}$$

11.5. Asymptotic behavior

Consider an irreducible aperiodic Markov chain which is positive recurrent.

Suppose that the following limit exists:

$$\lim_{n\to\infty} p_j(n) = \pi_j, \quad j \in I \tag{11.73}$$

starting with $J_0 = i$.

The relation

$$p_j(n+1) = \sum_{k\in I} p_k(n) p_{kj} \tag{11.74}$$

becomes:

$$p_{ij}^{(n+1)} = \sum_{k\in I} p_{ik}^{(n)} p_{kj}, \tag{11.75}$$

because

$$p_j(n) = p_{ij}^{(n)}. \tag{11.76}$$

Since the state space I is finite, we obtain from (11.73) and (11.75):

$$\pi_j = \sum_{k\in I} \pi_k p_{kj}, \tag{11.77}$$

and from (11.76):

$$\sum_{i\in I} \pi_i = 1. \tag{11.78}$$

The result:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \tag{11.79}$$

is called an *ergodic result*, since the value of the limit in (11.79) is independent of the initial state i.

From result (11.79) and (11.19), we see that for any initial distribution **p**:

$$\lim_{n\to\infty} p_i(n) = \lim_{n\to\infty} \sum_j p_j p_{ji}^{(n)}, \tag{11.80}$$

$$= \sum_j p_j \pi_i , \tag{11.81}$$

so that:

$$\lim_{n\to\infty} p_i(n) = \pi_i . \tag{11.82}$$

This shows that the asymptotic behavior of a Markov chain is given by the existence (or non-existence) of the limit of matrix $\mathbf{P}^n$.

A standard result concerning the asymptotic behavior of $\mathbf{P}^n$ is given in the next proposition. The proof can be found in Chung (1960), Parzen (1962) or Feller (1957).

Proposition 11.5 *For any aperiodic Markov chain of transition matrix* **P** *and having a finite number of states, we have*:

a) if state j is recurrent (necessarily positive), then

(i) $$i \in C(j) \Rightarrow \lim_{n\to\infty} p_{ij}^{(n)} = \frac{1}{m_{jj}}, \tag{11.83}$$

(ii) i is recurrent and $\notin C(j) \Rightarrow \lim_{n\to\infty} p_{ij}^{(n)} = 0,$ (11.84)

(iii) i is transient and $\lim_{n\to\infty} p_{ij}^{(n)} = \dfrac{f_{i,C(j)}}{m_{jj}}.$ (11.85)

b) If j is transient, then for all $i \in I$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0. \tag{11.86}$$

Remark 11.4 Result (ii) of part a) is trivial since in this case:

$$p_{ij}^{(n)} = 0 \text{ for all positive } n.$$

From Proposition 11.5, the following corollaries can be deduced.

Corollary 11.1 (*Irreducible case*) *If the Markov chain of transition matrix* **P** *is irreducible, then for all* $i, j \in I$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \tag{11.87}$$

with

$$\pi_j = \frac{1}{m_{jj}}. \tag{11.88}$$

It follows that for all j :

$$\pi_j > 0. \tag{11.89}$$

If we use Remark 11.4 in the particular case where we have only one recurrent class and where the states are transient (the *uni-reducible* case), then we have the following corollary.

Corollary 11.2 (*Uni-reducible case*) *If the Markov chain of transition matrix* **P** *has one essential class C* (*necessarily recurrent positive*) *and T as transient set, then we have*:

(i) for all $i, \quad j \in C$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j, \tag{11.90}$$

with $\left\{\pi_j, j \in C\right\}$ *being the unique solution of the system*:

$$\pi_j = \sum_{i\in C} \pi_i p_{ij}, \tag{11.91}$$

$$\sum_{j\in C} \pi_j = 1; \tag{11.92}$$

(ii) for all $j \in T$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0 \text{ for all } i \in I; \tag{11.93}$$

(iii) for all $j \in C$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \text{ for all } i \in T. \tag{11.94}$$

Remark 11.5 Relations (11.91) and (11.92) are true because the set C of recurrent states can be seen as a Markov sub-chain of the initial chain.

If the ℓ transient states belong to the set $\{1,\ldots,\ell\}$, using a permutation of the set I, if necessary, then matrix **P** takes the following form:

$$\begin{array}{c} \\ 1 \\ \vdots \\ \ell \\ \ell+1 \\ \vdots \\ m \end{array} \begin{array}{c} 1 \quad \cdots \quad \ell \quad \ell+1 \quad \cdots \quad m \\ \left[\begin{array}{cc} \mathbf{P}_{11} & \mathbf{P}_{12} \\ \mathbf{O} & \mathbf{P}_{22} \end{array}\right] \end{array} = \mathbf{P}. \tag{11.95}$$

This proves that the sub-matrix $\mathbf{P}_{22}$ is itself a Markov transition matrix.

Let us now consider a Markov chain of matrix **P**. The general case is given by a partition of I:

$$I = T \cup C_1 \cup \cdots \cup C_r, \tag{11.96}$$

where T is the set of transient states and $C_1,\ldots,C_r$ the r positive recurrent classes.

By reorganizing the order of the elements of I, we can always suppose that

$$T = \{1,\ldots,\ell\}, \tag{11.97}$$

$$C_1 = \{\ell+1,\ldots,\ell+\nu_1\}, \tag{11.98}$$

$$C_2 = \{\ell+\nu_1+1,\ldots,\ell+\nu_1+\nu_2\}, \tag{11.99}$$

$$\vdots$$

$$C_r = \left\{\ell + \sum_{j=1}^{r-1} \nu_j + 1,\ldots,\ m\right\}, \tag{11.100}$$

where ν_j is the number of elements in C_j, $(j = 1,\ldots,r)$ and

$$\ell + \sum_{j=1}^{r} \nu_j = m. \tag{11.101}$$

This results from the following block partition of matrix **P**:

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_{\ell\times\ell} & \mathbf{P}_{\ell\times\nu_1} & \mathbf{P}_{\ell\times\nu_2} & \cdots & \mathbf{P}_{\ell\times\nu_r} \\ \mathbf{0} & \mathbf{P}_{\nu_1\times\nu_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_{\nu_2\times\nu_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{P}_{\nu_r\times\nu_r} \end{bmatrix} \tag{11.102}$$

where, for $j = 1,\ldots, \quad r$:

– $\mathbf{P}_{\ell\times\ell}$ is the transition sub-matrix for T ;

– $\mathbf{P}_{\ell\times\nu_j}$ is the transition sub-matrix from T to C_j ;

– $\mathbf{P}_{\nu_j\times\nu_j}$ is the transition sub-matrix for the class C_j .

From Proposition 11.1, we obtain the following corollary.

Corollary 11.3 *For a general Markov chain of matrix* **P**, *given by* (11.102), *we have*:

(i) *for all* $i \in I$ *and all* $j \in T$ *:* $\lim_{n\to\infty} p_{ij}^{(n)} = 0;$ (11.103)

(ii) *for all* $j \in C_\nu \quad (\nu = 1,\ldots, \quad r)$:

$$\lim_{n\to\infty} p_{ij}^{(n)} = \begin{cases} \pi_j & \text{if } i \in C_\nu, \\ 0 & \text{if } i \in C_{\nu'} \quad \nu' \neq \nu, \\ f_{i,C_\nu} \pi_j^\nu & \text{if } i \in T, \end{cases} \tag{11.104}$$

moreover, for all $\nu = 1,\ldots, \quad r$:

$$\sum_{j\in C_\nu} \pi_j^\nu = 1. \tag{11.105}$$

This last result allows us to calculate the limit values quite simply.

For $\left(\pi_j^\nu, \quad j \in C_\nu\right), \quad \nu = 1,\ldots, \quad r$, it suffices to solve the linear systems for each fixed ν :

$$\begin{cases} \pi_j^{\nu} = \sum_{k\in C_\nu} \pi_k^{\nu} p_{kj}, & j \in C_\nu, \\ \sum_{i\in C_\nu} \pi_i^{\nu} = 1. \end{cases} \tag{11.106}$$

Indeed, since each C_ν is itself a space set of an irreducible Markov chain of matrix $\mathbf{P}_{\nu\times\nu}$, the above relations are none other than (11.77) and (11.78).

For the absorption probabilities $\left(f_{i,C_\nu}, \quad i\in T \right), \quad \nu = 1,\ldots, \quad r$, it suffices to solve the following linear system for each fixed ν. Using Proposition 11.4, we have:

$$f_{i,C_\nu} = \sum_{k\in T} p_{ik} f_{i,C_\nu} + \sum_{k\in C_\nu} p_{ik}, \qquad i\in T. \tag{5.35}$$

An algorithm, given in De Dominicis and Manca (1984b) and very useful for the classification of the states of a Markov chain, is fully developed in Janssen and Manca (2006).

11.6. Examples

Markov chains appear in many practical problems in fields such as operations research, business, social sciences, etc.

To give an idea of this potential, we will present some simple examples followed by a fully developed case study in the domain of social insurance.

11.6.1. *A management problem in an insurance company*

A car insurance company classifies its customers in three groups:
– G_0: those having no accidents during the year;
– G_1: those having one accident during the year;
– G_2: those having more than one accident during the year.

The statistics department of the company observes that the annual transition between the three groups can be represented by a Markov chain with state space $\{G_0, G_1, G_2\}$ and transition matrix **P**:

$$\mathbf{P} = \begin{bmatrix} 0.85 & 0.10 & 0.05 \\ 0 & 0.80 & 0.20 \\ 0 & 0 & 1 \end{bmatrix}. \tag{11.108}$$

We assume that the company produces 50,000 new contracts per year and wants to know the distribution of these contracts for the next four years.

After one year, we have, on average:

– in group G_0 : $50,000 \times .85 = 42,500$;

– in group G_1 : $50,000 \times .10 = 5,000$;

– in group G_2 : $50,000 \times .05 = 2,500$.

These results are simply the elements of the first row of **P**, multiplied by 50,000. After two years, multiplying the elements of the first row of $\mathbf{P}^{(2)}$ by 50,000, we obtain:

– in group G_0 : $36,125$;

– in group G_1 : $8,250$;

– in group G_2 : $5,625$.

A similar calculation gives:

	After three years	After four years
G_0	30,706	26,100
G_1	10,213	11,241
G_3	9,081	12,659

To find the type of the Markov chain with transition matrix (11.108), the simple graph of possible transitions given in Figure 11.3 shows that class $\{1,\ 2\}$ is transient and class $\{3\}$ is absorbing. Thus, using Corollary 11.2 we obtain the limit matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}. \tag{11.109}$$

The limit matrix can be interpreted as showing that regardless of the initial composition of the group the customers will finish by having at least two accidents.

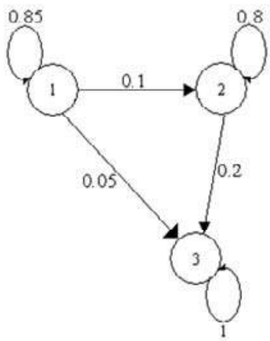

Figure 11.3. *Transition graph of matrix (11.108)*

Remark 11.6 If we want to know the situation after one or two changes, we can use relation (1.19) with $n = 1,2,3$ and with **p** given by:

$$\mathbf{p} = (0.26, 0.60, 0.14). \tag{11.110}$$

We obtain the following results:

$$\begin{array}{lll} p_1^{(1)} = 0.257 & p_2^{(1)} = 0.597 & p_3^{(1)} = 0.146 \\ p_1^{(2)} = 0.255 & p_2^{(2)} = 0.594 & p_3^{(2)} = 0.151 \\ p_1^{(3)} = 0.254 & p_2^{(3)} = 0.590 & p_3^{(3)} = 0.156. \end{array}$$

These results show that the convergence of $\mathbf{p}^{(n)}$ to π is relatively fast.

11.6.2. *A case study in social insurance (Janssen (1966))*

To calculate insurance or pension premiums for occupational diseases such as silicosis, we need to calculate the average (mean) degree of disability at pre-assigned time periods. Let us suppose we retain m degrees of disability:

$S_1, \ldots, \; S_m$, the last being 100% and including the pension paid out at death.

Let us suppose, as Yntema (1962) did, that an insurance policy holder can go from degree S_i to degree S_j with a probability p_{ij}. This strong assumption leads to the construction of a Markov chain model in which the $m \times m$ matrix:

$$\mathbf{P} = [\, p_{ij} \,] \tag{11.111}$$

is the transition matrix related to the degree of disability.

For individuals starting at time 0 with S_i as the degree of disability, the mean degree of disability after the *n*th transition is:

$$\overline{S}_i(n) = \sum_{j=1}^{m} p_{ij}^{(n)} S_j . \tag{11.112}$$

To study the financial equilibrium of the funds, we must calculate the limiting value of $\overline{S}_i(n)$:

$$\overline{S}_i = \lim_{n\to\infty} \overline{S}_i(n), \tag{11.113}$$

or

$$\overline{S}_i = \lim_{n\to\infty} \sum_{j=1}^{m} p_{ij}^{(n)} S_j . \tag{11.114}$$

This value can be found by applying Corollary 11.3 for $i = 1,\dots,\ m$.

Numerical example

Using real-life data for silicosis, Yntema (1962) began with the following intermediate degrees of disability:

$$\begin{aligned} S_1 &= 10\% \\ S_2 &= 30\% \\ S_3 &= 50\% \\ S_4 &= 70\% \\ S_5 &= 100\% \end{aligned}$$

Using real observations recorded in the Netherlands, he considered the following transition matrix **P**:

$$\mathbf{P} = \begin{bmatrix} 0.90 & 0.10 & 0 & 0 & 0 \\ 0 & 0.95 & 0.05 & 0 & 0 \\ 0 & 0 & 0.90 & 0.05 & 0.05 \\ 0 & 0 & 0 & 0.90 & 0.10 \\ 0 & 0 & 0.05 & 0.05 & 0.90 \end{bmatrix}; \tag{11.115}$$

the transition graph associated with matrix (11.115) being given in Figure 11.4. This immediately shows that:

(i) all states are aperiodic;

(ii) the set $\{S_3, S_4, S_5\}$ is an essential class (positive recurrent);

(iii) the singletons $\{1\}$ and $\{2\}$ are two inessential transient classes.

Thus a uni-reducible Markov chain can be associated with matrix **P**. We can thus apply Corollary 11.2. It follows from relation (11.114) that:

$$\overline{S_i} = \lim_{n\to\infty}\sum_{j=3}^{5}\pi_j S_j\ , \tag{11.116}$$

where (π_3, π_4, π_5) is the unique solution of the linear system:

$$\begin{array}{ccccccc} \pi_3 & = & 0.9\cdot\pi_3 & + & 0\cdot\pi_4 & + & 0.05\cdot\pi_5, \\ \pi_5 & = & 0.05\cdot\pi_3 & + & 0.9\cdot\pi_4 & + & 0.05\cdot\pi_5, \\ \pi_4 & = & 0.05\cdot\pi_3 & + & 0.05\cdot\pi_4 & + & 0.9\cdot\pi_5, \\ 1 & = & \pi_3 & + & \pi_4 & + & \pi_5. \end{array} \tag{11.117}$$

The solution is:

$$\pi_3 = \frac{2}{9},\quad \pi_4 = \frac{3}{9},\quad \pi_5 = \frac{4}{9}\,. \tag{11.118}$$

Therefore:

$$\overline{S_i} = \left(\frac{2}{9}50 + \frac{3}{9}70 + \frac{4}{9}100\right)\% \tag{11.119}$$

or

$$\overline{S_i} = 79\% \tag{11.120}$$

which is the result obtained by Yntema.

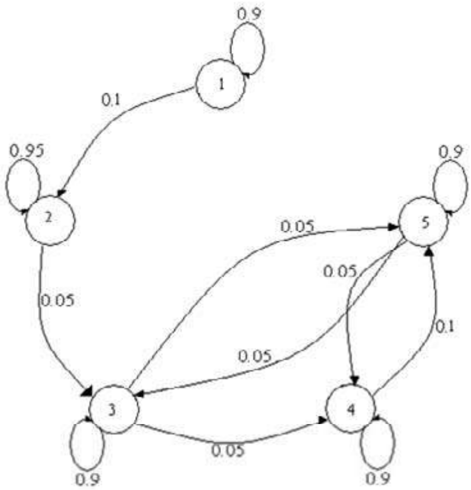

Figure 11.4. *Transition graph of matrix (11.115)*

The last result proves that the mean degree of disability is, at the limit, independent of the initial state i.