Math 481: Stochastic Models
Topic 1a): Discrete time Markov Chains (Part II)

Instructor: Prof. Tomasz Bielecki

1 Long time (or long range) behavior and invariant probability (finite chains)

By the long time behavior of a Markov chain we mean behavior of the conditional probabilities P_n and the unconditional probabilities $\bar{\phi}_n$ for large n. In view of the fact that $\bar{\phi}_n = \bar{\phi}_0 P_n = \bar{\phi}_0 P^n$ this essentially boils down to the behavior of the powers P^n of the transition matrix for large n. That really means that we want to analyze the limit, when $n \to \infty$ of the powers P^n.

Understanding the long time behavior of a Markov chains that model real systems is crucial for various applications in operations research and engineering [manufacturing, investment, scheduling etc.].

The analysis of the long time behavior of Markov chains is relatively simple for finite Markov chains. Let us start with several simple examples that, in a sense, illustrate all possible long time behavior patterns of finite Markov chains.

Example 1.1 Let

$$P = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & 1 \end{array} \\ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array} = I.$$

Then, of course, $\lim_{n\to\infty} P^n = P = I$. If the chain starts from the state $i = -1, 1$ then the chain stays in this state for ever. This is an example of a chain whose state space contains distinct *ergodic* classes [$\{-1\}$ is one class and $\{1\}$ is the other one].

To complete the example let us denote $\bar{\pi} = (\pi(-1), \pi(1))$. It will become clear soon why we are interested in finding these vectors $\bar{\pi}$ that satisfy the equation

$$\bar{\pi} = \bar{\pi} P \tag{1}$$

subject to

$$\pi(-1) + \pi(1) = 1, \quad \pi(-1), \pi(1) \geq 0. \tag{2}$$

It is clear that there are infinitely many solutions to the above system (1), (2) [this is because P is the identity matrix here]. This means that the Markov chain considered here has infinitely many *invariant probability distributions* $\bar{\pi}$.

In particular, the vectors $\bar{\pi}^1 = (1,0)$ and $\bar{\pi}^2 = (0,1)$ satisfy the system (1), (2). Observe that any other solution to (1), (2) is a convex combination of these to particular solutions. In addition, observe that

$$\lim_{n\to\infty} P^n = \begin{pmatrix} \bar{\pi}^1 \\ \bar{\pi}^2 \end{pmatrix}.$$

Example 1.2 Let

$$P = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & 1 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{array} = \textit{permutation matrix}.$$

Then, of course, $\lim_{n\to\infty} P^n$ does not exist. The chain alternates between its two states. This is an example of a chain whose state space contain one *ergodic* class [$S = \{-1, 1\}$ is the only ergodic class].

Let us remark here that even though the ordinary limit $\lim_{n\to\infty} P^n$ does not exist in this case, the so called Cesaro limit exists:

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} P^k = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & 1 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \end{array}.$$

Consider the system (1), (2) in the present case. There exist a unique solution $\bar{\pi} = (1/2, 1/2)$. The underlying Markov chain posses only one invariant distribution.

Example 1.3 Let

$$P = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & 1 \end{array} \\ \begin{pmatrix} p & 1-p \\ 0 & 1 \end{pmatrix} \end{array}, \quad 0 < p < 1.$$

Then, of course, $\lim_{n\to\infty} P^n = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & 1 \end{array} \\ \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \end{array}$. If the chains starts from the state $i = -1$ then the chain eventually gets absorbed in state $i = 1$ [$\lim_{n\to\infty} p_n(-1, 1) = 1$]. If the chain starts from the state $i = 1$ then the chain stays in this state for ever. This is an example of a chain whose state space contains one class of *transient* states [the state $i = -1$ is the only transient state], and one class of ergodic states [the state $i = 1$ is the only ergodic state].

Consider the system (1), (2) in the present case. There exist a unique solution $\bar{\pi} = (0, 1)$. The underlying Markov chain posses only one invariant distribution.

Example 1.4 Let

$$P = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & \quad 1 \end{array} \\ \begin{pmatrix} p & 1-p \\ 1-q & q \end{pmatrix} \end{array}, \quad 0 < p, q < 1.$$

Then [see Lawler, pages 13-14, for derivation] $\lim_{n\to\infty} P^n = \begin{array}{c} \\ -1 \\ 1 \end{array} \begin{array}{c} \begin{array}{cc} -1 & \quad 1 \end{array} \\ \begin{pmatrix} \frac{1-q}{2-p-q} & \frac{1-p}{2-p-q} \\ \frac{1-q}{2-p-q} & \frac{1-p}{2-p-q} \end{pmatrix} \end{array}$. Independent of the starting position of the chain, the long time probability that the chain is at state $i = -1$ is $\frac{1-q}{2-p-q}$, and the long time probability that the chain is at state $i = 1$ is $\frac{1-p}{2-p-q}$. This is an example of a chain whose state space contains one *ergodic* classes [$S = \{-1, 1\}$ is the only ergodic class].

Consider the system (1), (2) in the present case. There exist a unique solution $\bar{\pi} = (\frac{1-q}{2-p-q}, \frac{1-p}{2-p-q})$. The underlying Markov chain has only one invariant distribution.

In the examples 1.3 and 1.4 the limiting transition matrix Π exists and satisfies:

$$\Pi = \lim_{n\to\infty} P^n = \begin{pmatrix} \bar{\pi} \\ \bar{\pi} \end{pmatrix}.$$

In the example 1.1 the limiting transition matrix Π exists and satisfies:

$$\Pi = \lim_{n\to\infty} P^n = \begin{pmatrix} \bar{\pi}^1 \\ \bar{\pi}^2 \end{pmatrix}.$$

In these three cases the matrix Π contains long range conditional probabilities (or long range transition probabilities) for the underlying Markov chain. In examples 1.3 and 1.4 these probabilities do not depend on the initial state of the chain. This is not the case in example 1.1.

In the example 1.2 there is no limiting transition matrix in the ordinary sense, but the limiting matrix in the Cesaro sense [also denoted by Π].

Remark 1.1 Suppose $\bar{\pi}^i$ is the ith row of the limiting transition matrix Π. Since the matrix Π is stochastic, the elements of $\bar{\pi}^i$ are non-negative and add up to one. Thus, $\bar{\pi}^i$ is a *probability vector*.

Now, suppose that $\bar{v}^i$ is a [row] probability vector such that

$$\lim_{n\to\infty} \bar{v}^i P^n = \bar{\pi}^i.$$

Then we have

$$\bar{\pi}^i = \lim_{n\to\infty} \bar{v}^i P^n = \lim_{n\to\infty} \bar{v}^i P^{(n+1)} = (\lim_{n\to\infty} \bar{v}^i P^n)P = \bar{\pi}^i P.$$

Thus we see that the vector $\bar{\pi}^i$ satisfies the conditions

$$\bar{\pi} = \bar{\pi} P, \tag{3}$$

$$\bar{\pi}\mathbf{1}^T = 1, \quad \bar{\pi} \geq \mathbf{0}, \tag{4}$$

where $\mathbf{1}$ is the [row] vector of 1's and $\mathbf{0}$ is the [row] vector of 0's. The conditions (1), (2) are special cases of the conditions (3), (4).

Any vector $\bar{\pi}$ that satisfies (3), (4) is called an *invariant probability distribution vector* for P.

Note: Let $\bar{\phi} = \bar{v}^i$ be an initial probability distribution vector such that

$$\lim_{n\to\infty} \bar{v}^i P^n = \bar{\pi}^i.$$

Thus, the elements of $\bar{\pi}^i$ are the unconditional long range probabilities associated with the initial distribution $\bar{v}^i$ [that is, $\bar{\pi}^i$ is the limiting probability vector corresponding to the initial distribution $\bar{v}^i$].

Every limiting distribution is an invariant distribution.

The condition (3) means that if the Markov chain starts with the initial distribution equal to the invariant distribution $\bar{\pi}^i$, then this distribution is preserved over time [because then $\bar{\pi}^i = \bar{\pi}^i P^n$]. That is the reason why the invariant distributions are also called the *stationary distributions.* □

Let us consider another example:

Example 1.5 Let

$$P = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/6 & 5/6 & 0 & 0 & 0 \\ 0 & 0 & 3/4 & 1/4 & 0 \\ 0 & 0 & 1/8 & 2/3 & 5/24 \\ 0 & 0 & 0 & 1/6 & 5/6 \end{pmatrix} \end{array}.$$

Then,

$$\Pi = \lim_{n\to\infty} P^n = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{pmatrix} .25 & .75 & 0 & 0 & 0 \\ .25 & .75 & 0 & 0 & 0 \\ 0 & 0 & .182 & .364 & .454 \\ 0 & 0 & .182 & .364 & .454 \\ 0 & 0 & .182 & .364 & .454 \end{pmatrix} \end{array} = \begin{pmatrix} \bar{\pi}^1 \\ \bar{\pi}^1 \\ \bar{\pi}^2 \\ \bar{\pi}^2 \\ \bar{\pi}^2 \end{pmatrix}.$$

Every invariant probability distribution is a convex combination of $\bar{\pi}^1$ and $\bar{\pi}^2$.

Suppose the Markov chain starts with the initial distribution $\bar{\phi}^1 = (1/3, 2/3, 0, 0, 0)$. The limiting distribution is $\bar{\pi}^1 = (.25, .75, 0, 0, 0)$. So, for example, $lim_{n\to\infty} P(X_n = 2) = lim_{n\to\infty} P(X_n = 2|X_0 = 2) = lim_{n\to\infty} P(X_n = 2|X_0 = 1) = .75$, whereas $lim_{n\to\infty} P(X_n = 2|X_0 = i)$ is not defined for $i = 3, 4, 5$.

Suppose the Markov chain starts with the initial distribution $\bar{\phi}^2 = (0, 0, 1/3, 1/3, 1/3)$. The limiting distribution is $\bar{\pi}^2 = (0, 0, .182, .364, .454)$. So, for example, $lim_{n\to\infty} P(X_n = 2) = lim_{n\to\infty} P(X_n = 2|X_0 = i) = 0$ for $i = 3, 4, 5$, whereas $lim_{n\to\infty} P(X_n = 2|X_0 = i)$ is not defined for $i = 1, 2$.

Suppose the Markov chain starts with the initial distribution $\bar{\phi}^3 = (1/2)\bar{\phi}^1 + (1/2)\bar{\phi}^2 = (1/6, 2/6, 1/6, 1/6, 1/6)$. The limiting distribution is $\bar{\pi}^3 = (1/2)\bar{\pi}^1 + (1/2)\bar{\pi}^2$. So, for example, $lim_{n\to\infty} P(X_n = 2) = lim_{n\to\infty} P(X_n = 2|X_0 = i) = 0.375$ for $i = 1, 2, 3, 4, 5$

Consider now the three questions 1)-3), Lawler p. 13. Our examples indicate that:

- The answer to question 1) is positive (in the finite case).
- The invariant probability may not be unique.
- This happens if the chain does not alternate between states in a *periodic* way.

Discussion on pages 13-15 in Lawler employs the so called Perron-Frobenius theory to imply that positive answers to questions 2) and 3) are obtained under conditions (1.9) and (1.10) [in Lawler, page 14]. We have the following sufficient criterion for these conditions to be satisfied:

Fact If P is a stochastic matrix such that for some n, P^n has all entries strictly positive, then the conditions (1.9) and (1.10) in Lawler are satisfied for P. [This is a sufficient, but by no means a necessary condition. Compare example 1.3.]

We shall next classify states of a Markov chain according to the arithmetic properties of the entries of the powers P^n of the corresponding transition matrix.

HW

1) Recall our introductory example and use the same data as in HW problem 1) in Topic 1a) part I. What is the probability that the interest rate will be high in the long range?

2) Lawler, Ex. 1.3.