STA 4103/5107 Computational Methods in Statistics II

Department of Statistics
Florida State University

Class 7
January 31, 2017

Review: Finite-State Space Case

$$X_{t_i} \in \{x_1, x_2, \ldots, x_m\}.$$

- The Markov property implies:

$$P\{X_{t_n} = a_n \mid X_{t_{n-1}} = a_{n-1}, \ldots, X_{t_1} = a_1\} = P\{X_{t_n} = a_n \mid X_{t_{n-1}} = a_{n-1}\}.$$

- This transition probability in a **homogeneous Markov chain** is denoted by an mxm matrix $\Pi = \{\Pi_{i,j}\}$, where

$$\Pi_{i,j} = P\{X_{t_n} = x_j \mid X_{t_{n-1}} = x_i\}.$$

- **The probability of transition from x_i to x_j in n ($n \geq 1$) steps is given by the (i, j)-th entry in the matrix Π^n. That is,**

$$\boxed{P\{X_{t_{n+1}} = x_j \mid X_{t_1} = x_i\} = \{\Pi^n\}_{i,j}}$$

Review: Probability Transition

- Let $P[n] = (P\{X_{t_n} = x_1\}, P\{X_{t_n} = x_2\}, \ldots, P\{X_{t_n} = x_m\})$.
- Then $\boldsymbol{P[n] = P[n-1]\ \Pi = \ldots = P[1]\ \Pi^{n-1}}$.
- If $P[1] = P$ such that $P\,\Pi = P$, then $P[n] = P$ for all n and the resulting Markov process is not only homogeneous but also stationary.
- P is called the stationary probability distribution associated with the Markov chain.
- Main question: Under what conditions on Π does the resulting Markov chain converge to a stationary process?

Review: Peron-Frobenius Theorem

- **Theorem 7 (Peron-Frobenius)** If $\Pi^n >> 0$ for some $n \geq 1$, then

  1. there exists an $X >> 0$ such that $X\Pi = X$, and

  2. if λ is any other eigenvalue of Π, then $|\lambda| < 1$.

- Main Results:

  1. The resulting Markov chain has a unique stationary probability vector P ($P = X/\text{sum}(X)$).

  2. Irrespective of the starting condition, the Markov chain converges to a stationary process with stationary probability is P, and the chain samples from P for t very large.

- Difficult to establish the condition in the theorem. Alternative way is sought that can be checked easily.

Characterizing Markov Chains

- Let $x_1, x_2, \ldots, x_m$ be the states in a finite state space and Π be an $m \times m$ transition matrix. We start with a few definitions.

- **Definition 19** Two states are said to **communicate** if it is possible to go from either one to the other in a finite number of steps. In other words, x_i and x_j are said to communicate if there exist positive integer m and n such that $\Pi^m_{i,j} > 0$ and $\Pi^n_{j,i} > 0$.

- For example,

all states communicate: $\begin{bmatrix} .1 & .3 & .4 & .2 \\ .2 & .4 & .0 & .4 \\ .0 & .3 & .5 & .2 \\ .5 & .3 & .2 & .0 \end{bmatrix}$ x_1 and x_3 do not: $\begin{bmatrix} .5 & .5 & .0 & .0 \\ .5 & .5 & .0 & .0 \\ .0 & .0 & .5 & .5 \\ .0 & .0 & .5 & .5 \end{bmatrix}$

Irreducibility

- **Definition 20** A Markov chain is said to be **irreducible** if all states communicate with each other for the corresponding transition matrix Π.

- For example, in the Markov chain resulting from the above two matrices, the first one will be irreducible.

- The chain resulting from the second matrix will be reducible into two clusters: one including states x_1 and x_2, and the other including the states x_3 and x_4.

Aperiodicity

- **Definition 21** For a state x_i and a given transition matrix Π, define the period of x_i as:

$$d(x_i) = GCD\{n : \Pi^n_{i,i} > 0\},$$

where GCD implies the **greatest common divisor**.

- For the following transition matrices:

$$\begin{bmatrix} .1 & .3 & .4 & .2 \\ .2 & .4 & .0 & .4 \\ .0 & .3 & .5 & .2 \\ .5 & .3 & .2 & .0 \end{bmatrix} \qquad \begin{bmatrix} 0.0 & 1.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 1.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 1.0 \\ 1.0 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

$$d(x_1) = 1 \qquad d(x_1) = 4$$

Aperiodicity

- Do we have $GCD\{n : \Pi_{i,i}^n > 0\} = \min\{n : \Pi_{i,i}^n > 0\}$?
- No. For example,

$$\Pi = \begin{bmatrix} 0 & 0.5 & 0.5 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad \Pi^2 = \begin{bmatrix} 0.5 & 0 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 \end{bmatrix}$$

$$\Pi^3 = \begin{bmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.5 & 0 & 0 & 0.5 \\ 0 & 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0 & 0.5 \end{bmatrix}$$

$$GCD\{n : \Pi_{1,1}^n > 0\} = 1$$

$$\min\{n : \Pi_{1,1}^n > 0\} = 2$$

Period

- **Proposition 5** If two states communicate, then their periods are same.

- Therefore, in an irreducible Markov chain where all the states communicate, they all have the same period. This is called the **period of an irreducible Markov chain**.

- **Definition 22** An irreducible Markov chain is called **aperiodic** if its period is one.

- Irreducibility and aperiodicity are sufficient to show that the Markov chain converges to a stationary process and it samples from a (unique) stationary probability as the chain is long.

Main Theorem

- **Theorem 8**

  An irreducible, aperiodic, homogeneous Markov chain on a finite state space has the property that $\Pi^n >> 0$ for some $n > 0$.

  Furthermore, this Markov chain has a unique probability distribution P such that $P\,\Pi = P$.

  For any arbitrary starting condition, this Markov chain converges to a stationary process and generates samples from P as the time t goes to infinity.

Equivalence

- From Theorem 8, for a homogeneous Markov chain with transition matrix Π, we have

$$\text{irreducible \& aperiodic} \Rightarrow \Pi^n >> 0 \text{ for some } n > 0.$$

- Is the inverse true? That is,

$$\Pi^n >> 0 \text{ for some } n > 0 \Rightarrow \text{irreducible \& aperiodic?}$$

Yes.

The proof is straightforward (note: $\Pi^n >> 0 \Rightarrow \Pi^{n+1} >> 0$).

Therefore, these two conditions are equivalent.

Summary

- In summary, given a probability distribution P on a finite state space, we construct a Markov chain using a transition matrix Π that satisfies three conditions:

- 1. The resulting Markov chain is irreducible.

- 2. The resulting Markov chain is aperiodic.

- 3. P is a stationary probability of that Markov chain, i.e. $P\,\Pi = P$.