STA 4103/5107 Computational Methods in Statistics II

Department of Statistics
Florida State University

Class 6
January 26, 2017

Review: Discrete Time Stochastic Processes

- Denote a discrete time stochastic process as: $\{X_t, t = t_1, t_2, \ldots\}$. Such a process can be characterized by n^{th}-order joint probability density function,

$$f_{X_{t_1}, X_{t_2}, \cdots, X_{t_n}}(x_1, x_2, \cdots, x_n), \text{ or simply } f(x_1, x_2, \cdots, x_n).$$

- **Definition 16** A stochastic process is called a **Markov process** if

$$f(x_n \mid x_{n-1}, \cdots, x_2, x_1) = f(x_n \mid x_{n-1})$$

- This definition implies that joint density function can be written as a product of one-step conditional densities as follows:

$$f(x_1, x_2, \cdots, x_n) = f(x_1) f(x_2 \mid x_1) \cdots f(x_n \mid x_{n-1})$$

Review: Stationarity and Homogeneity

- **Definition 17** A stochastic process is called **stationary** if its n^{th}-order joint density function is translation invariant, for all $n \geq 1$. That is, for any collection of times $\{t_1, t_2, \ldots, t_n\}$, we have

$$f_{X_{t_1}, X_{t_2}, \cdots, X_{t_n}}(x_1, x_2, \cdots, x_n) = f_{X_{t_{1+k}}, X_{t_{2+k}}, \cdots, X_{t_{n+k}}}(x_1, x_2, \cdots, x_n),$$

for all $n > 0$ and k.

- **Definition 18** A Markov process is called **homogeneous** if the conditional density is invariant to a time shift. That is, for all n,

$$f_{X_{t_n} | X_{t_{n-1}}}(x_n \mid x_{n-1}) = f_{X_{t_2} | X_{t_1}}(x_n \mid x_{n-1}),$$

- A stationary Markov process is always homogeneous. However, a homogeneous process, in general, is not stationary.

Review: Stationary Probability Density

- If there exists a density function g such that:

$$g(y) = \int f_{X_{t_2}|X_{t_1}}(y \mid x) g(x) dx,$$

then the resulting Markov chain is a stationary process. The density function g is called the **stationary probability density** of that Markov chain.

- Our goal is to construct such homogeneous Markov processes that are not stationary to start with, but converge to stationary processes as the process is followed for a long time.

7.2 Markov Chains for Sampling from Probabilities

Framework

- In the next few sections, we develop a framework for using Markov chains to sample from given probability distributions.
- We start with a probability distribution on a finite set.
- This analysis can be broken into two distinct issues:

1. When does a given homogeneous Markov chain, with a given transition function, converge to a stationary process?

2. For a given probability distribution, how to construct a homogeneous Markov chain that samples from that distribution asymptotically?

Finite-State Space Case

- We will consider a discrete time Markov chain that takes values only in a finite set. That is, for any time

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

- For this finite state setup, the n-th order probability mass function is given by:

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

where $a_i \in \{x_1, x_2, \ldots, x_m\}, i = 1, \ldots, n.$

- The Markov property implies:

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

Homogeneity

- We assume that the Markov chain is homogeneous, i.e. its one-step transition probability distribution does not change in time.

- This transition probability is denoted by an $m \times m$ matrix $\Pi = \{\Pi_{i,j}\}$, where
$$\Pi_{i,j} = P\{X_{t_n} = x_j \mid X_{t_{n-1}} = x_i\}.$$

- Note that each row of the matrix adds up to one.

- **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,**

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

Probability Vector

- Let $P[n]$ be the probability vector at time $t = t_n$, that is,
$$P[n] = (P\{X_{t_n} = x_1\}, P\{X_{t_n} = x_2\}, \ldots, P\{X_{t_n} = x_m\}).$$

- If the transitions are made according to Π, then the probability (row) vector associated with time $t = t_n$ is given by
$$\boldsymbol{P[n] = P[n-1]\,\Pi = \ldots = P[1]\,\Pi^{n-1}}.$$

- A special case arises when $P[1] = P$ such that $P\Pi = P$, i.e. P is the (row) eigenvector of the matrix Π with the corresponding eigenvalue given by one.

- In this situation, $P[n] = P$ for all n and the resulting Markov process is not only homogeneous but also stationary.

Stationary Probability Distribution

- P is called the stationary probability distribution associated with the Markov chain.

- If $P[1] \neq P$, then the resulting chain is, in general, not stationary, and $P[n] \neq P$ for all n.

- But it is possible to construct a Markov process, with a transition matrix Π, in such a way that $P[n] \rightarrow P$ as $n \rightarrow \infty$.

- Main questions: Under what conditions on Π does the resulting Markov chain converge to a stationary process? Alternatively, under what conditions on Π does the probability $P[n]$ converge to a stationary probability P?

Peron-Frobenius Theorem

- The symbol $A >> 0$ implies that all elements of that array (vector or matrix) are strictly positive.

- **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 $|\lambda| < 1$.

- This theorem states that if there exists an n such that we can go from any state to any other state in n steps with positive probability, then the resulting Markov chain has a unique stationary probability vector P (normalized vector of X).

Peron-Frobenius Theorem

- Furthermore, it states that **irrespective of the starting condition, the resulting Markov chain converges to a stationary process whose stationary probability is *P*.**

- According to the theorem, if the transition matrix Π^n has all positive elements for some $n > 0$ and if $P\Pi = P$, then the Markov chain samples from P for $t \to \infty$. For a large T, the process at times $t > T$ approximately sample from P.

- The condition in Peron-Frobenious theorem is difficult to establish in practice. We seek another way of characterizing the convergence of a homogeneous Markov chain with conditions that can be checked easily.