Lecture 7

Last updated by Serik Sagitov: December 10, 2012

Abstract

A course based on the book Probability and Random Processes by Geoffrey Grimmett and David Stirzaker. Chapter 9. Stationary processes.

1 Stationary processes

Definition 1 *The real-valued process $\{X(t), t \geq 0\}$ is called strongly stationary if the vectors $(X(t_1), \ldots, X(t_n))$ and $(X(t_1 + h), \ldots, X(t_n + h))$ have the same joint distribution for all $t_1, \ldots, t_n$ and $h > 0$.*

Definition 2 *The real-valued process $\{X(t), t \geq 0\}$ with $\mathbb{E}(X^2(t)) < \infty$ for all t is called weakly stationary if for all t_1, t_2 and $h > 0$*

$$\mathbb{E}(X(t_1)) = \mathbb{E}(X(t_2)), \qquad \mathbb{C}ov(X(t_1), X(t_2)) = \mathbb{C}ov(X(t_1 + h), X(t_2 + h)).$$

Its autocovariance and autocorrelation functions are

$$c(t) = \mathbb{C}ov(X(s), X(s + t)), \qquad \rho(t) = \frac{c(t)}{c(0)}.$$

Example 1. Consider an irreducible Markov chain $\{X(t), t \geq 0\}$ with countably many states and a stationary distribution $\boldsymbol{\pi}$ as the initial distribution. This is a strongly stationary process since

$$\mathbb{P}(X(h + t_1) = i_1, X(h + t_1 + t_2) = i_2, \ldots, X(h + t_1 + \ldots + t_n) = i_n) = \pi_{i_1} p_{i_1,i_2}(t_2) \ldots p_{i_{n-1},i_n}(t_n).$$

Example 2. The process $\{X_n, n = 1, 2, \ldots\}$ formed by iid Cauchy r.v is strongly stationary but not a weakly stationary process.

Example 3. Put $X(t) = \cos(t + Y)$ where Y is uniformly distributed over $[0, 2\pi]$. This is a strongly stationary process since $X(t+h) = \cos(t+Y')$, where Y' is uniformly distributed over $[h, 2\pi + h]$. Given an initial value this is a deterministic process and it is enough to show that $X(t) \stackrel{d}{=} X(0)$ for any t.

What is the distribution of $X = X(t)$? For an arbitrary bounded measurable function $\phi(x)$ we have

$$\begin{aligned}
\mathbb{E}(\phi(X)) = \mathbb{E}(\phi(\cos(t + Y))) &= \frac{1}{2\pi} \int_0^{2\pi} \phi(\cos(t + y))dy = \frac{1}{2\pi} \int_t^{t+2\pi} \phi(\cos(z))dz = \frac{1}{2\pi} \int_0^{2\pi} \phi(\cos(z))dz \\
&= \frac{1}{2\pi} \Big(\int_0^{\pi} \phi(\cos(z))dz + \int_{\pi}^{2\pi} \phi(\cos(z))dz \Big) = \frac{1}{2\pi} \Big(\int_0^{\pi} \phi(\cos(\pi - y))dy + \int_0^{\pi} \phi(\cos(\pi + y))dy \Big) \\
&= \frac{1}{\pi} \int_0^{\pi} \phi(-\cos(y))dy.
\end{aligned}$$

The change of variables $x = -\cos(y)$ yields $dx = \sin(y)dy = \sqrt{1 - x^2}dx$, hence

$$\mathbb{E}(\phi(X)) = \frac{1}{\pi} \int_{-1}^{1} \frac{\phi(x)dx}{\sqrt{1 - x^2}}.$$

Thus $X(t)$ has the so-called arcsine density $f(x) = \frac{1}{\pi\sqrt{1-y^2}}$ over the interval $[-1, 1]$. Notice that $Z = \frac{X+1}{2}$ has a Beta$(\frac{1}{2}, \frac{1}{2})$ distribution, since

$$\mathbb{E}(\phi(Z)) = \frac{1}{\pi} \int_{-1}^{1} \frac{\phi(\frac{x+1}{2})dx}{\sqrt{1 - x^2}} = \frac{1}{\pi} \int_0^1 \frac{\phi(z)dz}{\sqrt{z(1 - z)}}.$$

2 Linear combination of sinusoids

For a sequence of fixed frequencies $0 \leq \lambda_1 < \ldots < \lambda_k < \infty$ define a continuous time stochastic process by

$$X_t = \sum_{j=1}^{k} (A_j \cos(\lambda_j t) + B_j \sin(\lambda_j t)),$$

where $A_1, B_1, \ldots, A_k, B_k$ are uncorrelated r.v. with zero means and $\mathbb{V}ar(A_j) = \mathbb{V}ar(B_j) = \sigma_j^2$. Its mean is zero and its autocovariancies are

$$\begin{aligned}
\mathbb{C}ov(X_t, X_s) = \mathbb{E}(X_t X_s) &= \sum_{j=1}^{k} \mathbb{E}(A_j^2 \cos(\lambda_j t)\cos(\lambda_j s) + B_j^2 \sin(\lambda_j t)\sin(\lambda_j s)) \\
&= \sum_{j=1}^{k} \sigma_j^2 \cos(\lambda_j (s-t)), \\
\mathbb{V}ar(X_t) &= \sum_{j=1}^{k} \sigma_j^2.
\end{aligned}$$

Thus X_t is weakly stationary with autocovariance and autocorrelation functions

$$\begin{aligned}
c(t) &= \sum_{j=1}^{k} \sigma_j^2 \cos(\lambda_j t) \\
\rho(t) &= \frac{c(t)}{c(0)} = \sum_{j=1}^{k} g_j \cos(\lambda_j t) = \int_0^\infty \cos(\lambda t) dG(\lambda),
\end{aligned}$$

where

$$g_j = \frac{\sigma_j^2}{\sigma_1^2 + \ldots + \sigma_k^2}, \qquad G(\lambda) = \sum_{j:\lambda_j \leq \lambda} g_j.$$

We can write

$$X(t) = \int_0^\infty \cos(t\lambda) dU(\lambda) + \int_0^\infty \sin(t\lambda) dV(\lambda),$$

where

$$U(\lambda) = \sum_{j:\lambda_j \leq \lambda} A_j, \qquad V(\lambda) = \sum_{j:\lambda_j \leq \lambda} B_j.$$

Example 4. Let $k = 1$, $\lambda_1 = \frac{\pi}{4}$, A_1 and B_1 be iid with

$$\mathbb{P}(A_1 = \frac{1}{\sqrt{2}}) = \mathbb{P}(A_1 = -\frac{1}{\sqrt{2}}) = \frac{1}{2}.$$

Then $X_t = \cos(\frac{\pi}{4}(t+\tau))$ with

$$\mathbb{P}(\tau = 1) = \mathbb{P}(\tau = -1) = \mathbb{P}(\tau = 3) = \mathbb{P}(\tau = -3) = \frac{1}{4}.$$

This stochastic process has only four possible trajectories. This is not a strongly stationary process since

$$\mathbb{E}(X^4(t)) = \frac{1}{2}\Big(\cos^4\left(\frac{\pi}{4}t + \frac{\pi}{4}\right) + \sin^4\left(\frac{\pi}{4}t + \frac{\pi}{4}\right)\Big) = \frac{1}{4}\Big(2 - \sin^2\left(\frac{\pi}{2}t + \frac{\pi}{2}\right)\Big) = \frac{1 + \sin^2(\frac{\pi}{2}t)}{2}.$$

3 The spectral representation

Any weakly stationary process $\{X(t) : -\infty < t < \infty\}$ with zero mean can be approximated by a linear combination of sinusoids. Indeed, its autocovariance function $c(t)$ is non-negative definite since for any $t_, \ldots, t_n$ and $z_, \ldots, z_n$

$$\sum_{j=1}^{n}\sum_{k=1}^{n} c(t_k - t_j) z_j z_k = \mathbb{V}ar\Big(\sum_{k=1}^{n} z_k X(t_k)\Big) \geq 0.$$

Thus due to the Bochner theorem, given that $c(t)$ is continuous at zero, there is a probability distribution function G such that

$$\rho(t) = \int_0^\infty \cos(t\lambda) dG(\lambda).$$

Definition 3 *The function G is called the spectral distribution function of the corresponding stationary random process, and the set of real numbers λ such that*

$$G(\lambda + \epsilon) - G(\lambda - \epsilon) > 0 \text{ for all } \epsilon > 0$$

is called the spectrum of the random process. If G has density it is called the spectral density function.

Theorem 4 *If $\{X(t) : -\infty < t < \infty\}$ is a weakly stationary process with zero mean, unit variance, continuous autocorrelation function and spectral distribution function G, then there exists a pair of uncorrelated zero mean random process $(U(\lambda), V(\lambda))$ with uncorrelated increments such that*

$$X(t) = \int_0^\infty \cos(t\lambda) dU(\lambda) + \int_0^\infty \sin(t\lambda) dV(\lambda)$$

and $\mathbb{V}ar(U(\lambda)) = \mathbb{V}ar(V(\lambda)) = G(\lambda)$.

Example 5. Consider an irreducible Markov chain $\{X(t), t \geq 0\}$ with two states $\{1, 2\}$ and generator

$$\mathbf{G} = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}.$$

Its stationary distribution is $\boldsymbol{\pi} = (\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta})$ will be taken as the initial distribution. From

$$p_{11}(t) = 1 - p_{12}(t) = \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} e^{-t(\alpha+\beta)},$$
$$p_{22}(t) = 1 - p_{21}(t) = \frac{\alpha}{\alpha+\beta} + \frac{\beta}{\alpha+\beta} e^{-t(\alpha+\beta)}$$

we find for $t \geq 0$

$$c(t) = \frac{\alpha\beta}{(\alpha+\beta)^2} e^{-t(\alpha+\beta)}, \qquad \rho(t) = e^{-t(\alpha+\beta)}.$$

Thus this process has a spectral density

$$g(\lambda) = \frac{2(\alpha+\beta)}{\pi((\alpha+\beta)^2 + \lambda^2)}, \quad \lambda \geq 0$$

corresponding to a scaled one-sided Cauchy distribution.

4 The ergodic theorem

The following theorems are extensions of the Laws of Large Numbers.

Theorem 5 *Let $\{X_n, n = 1, 2, \ldots\}$ be a strongly stationary process with a finite mean. There exists a r.v. Y with the same mean such that*

$$\frac{X_1 + \ldots + X_n}{n} \to Y \text{ a.s. and in mean.}$$

Theorem 6 *Let $\{X_n, n = 1, 2, \ldots\}$ be a weakly stationary process. There exists a r.v. Y with the same mean such that*

$$\frac{X_1 + \ldots + X_n}{n} \to Y \text{ in square mean.}$$

Example 6. Let $Z_1, \ldots, Z_k$ be iid with mean μ and variance σ^2. Then the following cyclic process

$$X_1 = Z_1, \ldots, X_k = Z_k,$$
$$X_{k+1} = Z_1, \ldots, X_{2k} = Z_k,$$
$$X_{2k+1} = Z_1, \ldots, X_{3k} = Z_k, \ldots,$$

is a strongly stationary process. The corresponding limit in the ergodic theorem is not the constant μ like in the strong LLN but rather a random variable

$$Y = \frac{Z_1 + \ldots + Z_k}{k}.$$

Clearly, $\mathbb{E}(Y) = \mu$.

5 Gaussian processes

The covariance matrix of a random vector $(X_1, \ldots, X_n)$ with means $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$

$$\mathbf{V} = \mathbb{E}\big(\mathbf{X} - \boldsymbol{\mu}\big)^{\mathrm{t}}\big(\mathbf{X} - \boldsymbol{\mu}\big) = \|\mathrm{cov}(X_i, X_j)\|$$

is symmetric and nonnegative-definite. For any vector $\mathbf{a} = (a_1, \ldots, a_n)$ the r.v. $a_1X_1 + \ldots + a_nX_n$ has mean $\mathbf{a}\boldsymbol{\mu}^{\mathrm{t}}$ and variance

$$\mathrm{var}(a_1X_1 + \ldots + a_nX_n) = \mathbb{E}\big(\mathbf{a}\mathbf{X}^{\mathrm{t}} - \mathbf{a}\boldsymbol{\mu}^{\mathrm{t}}\big)\big(\mathbf{X}\mathbf{a}^{\mathrm{t}} - \boldsymbol{\mu}\mathbf{a}^{\mathrm{t}}\big) = \mathbf{a}\mathbf{V}\mathbf{a}^{\mathrm{t}}.$$

A multivariate normal distribution with mean vector $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)$ and covariance matrix $\mathbf{V}$ has density

$$f(\mathbf{x}) = \frac{1}{\sqrt{(2\pi)^n \det \mathbf{V}}}\, e^{-(\mathbf{x}-\boldsymbol{\mu})\mathbf{V}^{-1}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{t}}}$$

and moment generating function $\phi(\mathbf{t}) = e^{i\mathbf{t}\boldsymbol{\mu}^{\mathrm{t}} - \frac{1}{2}\mathbf{t}\mathbf{V}\mathbf{t}^{\mathrm{t}}}$. It follows that given a vector $\mathbf{X} = (X_1, \ldots, X_n)$ with a multivariate normal distribution any linear combination $\mathbf{a}\mathbf{X}^{\mathrm{t}} = a_1X_1 + \ldots + a_nX_n$ is normally distributed since

$$\mathbb{E}(e^{t\mathbf{a}\mathbf{X}^{\mathrm{t}}}) = \phi(t\mathbf{a}) = e^{it\mu - \frac{1}{2}t^2\sigma^2}, \quad \mu = \mathbf{a}\boldsymbol{\mu}^{\mathrm{t}}, \quad \sigma^2 = \mathbf{a}\mathbf{V}\mathbf{a}^{\mathrm{t}}.$$

Definition 7 *A random process $\{X(t), t \geq 0\}$ is called Gaussian if for any $(t_1, \ldots, t_n)$ the vector $(X(t_1), \ldots, X(t_n))$ has a multivariate normal distribution.*

A Gaussian random process is strongly stationary iff it is weakly stationary.

Theorem 8 *A Gaussian process $\{X(t), t \geq 0\}$ is Markov iff for any $0 \leq t_1 < \ldots < t_n)$*

$$\mathbb{E}(X(t_n)|X(t_1), \ldots, X(t_{n-1})) = \mathbb{E}(X(t_n)|X(t_{n-1})).$$

A stationary Gaussian Markov process is called the Ornstein-Uhlenbeck process. It is characterized by $\rho(t) = e^{-\alpha t}$, $t \geq 0$ with a positive α. This follows from the equation $\rho(t + s) = \rho(t)\rho(s)$.