Lecture 23

"Markov processes continued. <mark>Homogeneous Markov chains</mark>."

We know that a Wiener process $(W_t), t \geq 0$ is a Markov process (as a process with independent increments). It is also a zero-mean Gaussian process.

Theorem. (characterization of the class of zero-mean Gaussian Markov processes)

Let (X_t) be a zero-mean Gaussain process which is also a Markov process. Assume additionally that $R(t,t) > 0$ for all $t \geq 0$. Then, there are a deterministic function $\tau(t)$ being non-decreasing and a function $f(t)$ so that

$$X_t = f(t)W_{\tau(t)}, t \geq 0. \tag{1}$$

In other words, (1) means that every a zero-mean Gaussian Markov process is a rescaled Wiener process run according to the clock τ (function $f(t)$ is here a rescaling function and τ is the inner clock of (X_t)).

Proof. We first show that $R(s,t)$ can be factored in the form

$$R(s,t) = f(t)g(s)$$

for some functions f and g.

Without loss of generality we can assume that $s \leq t$ since $R(s,t)$ is symmetric.

Then, for all $s \leq t$ we can write

$$E(X_t|X_s = x) = R(s,t)R^{-1}(s,s)x \tag{2}$$

due to the formula for the best least squares estimate for Gaussian vectors.

From another side, let u be any intermediate state $s \leq u \leq t$. Then, using Chapman-Kolmogorov equation, we have:

$$E(X_t|X_s = x) = \int_{-\infty}^{\infty} zP(s,x;t,dz) =$$

$$\int_{-\infty}^{\infty} z \left(\int_{-\infty}^{\infty} P(s,x;u,dy)P(u,y;t,dz) \right) =$$

$$\int_{-\infty}^{\infty} \left(z \int_{-\infty}^{\infty} P(u,y;t,dz) \right) P(s,x;u,dy) =$$

$$\int_{-\infty}^{\infty} E(X_t|X_u = y)P(s,x;u,dy) = \int_{-\infty}^{\infty} \frac{R(u,t)}{R(u,u)} y P(s,x;u,dy) =$$

$$\frac{R(u,t)}{R(u,u)} \int_{-\infty}^{\infty} y P(s,x;u,dy) =$$

$$\frac{R(u,t)}{R(u,u)} \frac{R(s,u)}{R(s,s)} x. \tag{3}$$

By comparing (2) and (3), we obtain

$$R(s,t) = \frac{R(u,t)R(s,u)}{R(u,u)}, s \le u \le t. \tag{4}$$

In particular, for $s = 0$, we get

$$R(0,t) = \frac{R(u,t)R(0,u)}{R(u,u)}$$

so that switching $u \to s$, we can write that for all $0 < s < t$:

$$R(0,t) = \frac{R(s,t)R(0,s)}{R(s,s)}$$

or

$$R(s,t) = R(0,t)\frac{R(s,s)}{R(0,s)} = f(t)g(s)$$

with

$$f(t) = R(0,t) \text{ and } g(s) = \frac{R(s,s)}{R(0,s)}.$$

Define

$$\tau(t) := \frac{g(t)}{f(t)}.$$

Since the covariance function is positive semi-definite, it follows that

$$R(s,s)R(t,t) \ge R^2(s,t)$$

for all s,t (the determinant of the corresponding 2×2 matrix is non-negative).

In particular, for all $s \leq t$:

$$f(s)g(s)f(t)g(t) \geq (f(t)g(s))^2$$

so that

$$\tau(t) = \frac{g(t)}{f(t)} \geq \frac{g(s)}{f(s)} = \tau(s)$$

yielding that $\tau(t)$ is a non-decreasing function.

Define

$$Y_t := f(t)W_{\tau(t)}.$$

We see that Y is a zero-mean Gaussian process and

$$cov(Y_s, Y_t) = f(t)f(s)cov(W_{\tau(s)}, W_{\tau(t)}) =$$

$$f(t)f(s)\min\{\tau(s), \tau(t)\} = f(t)f(s)\tau(s) =$$

$$f(t)f(s)\frac{g(s)}{f(s)} = f(t)g(s) = R(s,t), s \leq t.$$

We know that any Gaussian process is determined by its mean and covariance function proving that X and Y are equivalent processes. $\square$

Example. Suppose X_t is a zero mean Gaussian process with covariance function $R(t,s) = e^{-\alpha|t-s|}$. The latter means that X_t is a WSS process with covariance function $R(\tau) = e^{-\alpha|\tau|}$.

Suppose $s \leq t$ so that $e^{-\alpha|t-s|} = e^{-\alpha t}e^{\alpha s} = f(t)g(s)$.

Define $\tau(t) = [: \frac{g(t)}{f(t)}] = e^{2\alpha t}$. The Theorem we proved implies then that

$$X_t = e^{-\alpha t}W_{e^{2\alpha t}}.$$

Recall: "random telegraph signal" has the same covariance function as process X_t, but is not a Gaussian process. This shows how **little** information is contained in covariance information in general.

Homogeneous Markov Chains

Definition 1. A (homogeneous) Markov chain is a discrete-time (homogeneous) Markov process with a *finite or countably infinite state space*. We normalize both the time steps and values to be $\mathbb{N} = \{0, 1, 2, ...\}$.

We use the following notations for transition probabilities:

$$p_{ij} = P(X_1 = j | X_0 = i)[=: P(1; i, \{j\})],$$

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i)[=: P(n; i, \{j\})].$$

p_{ij} are 1-step and $p_{ij}^{(n)}$ are n-step transition probabilities.

The matrix $P = (p_{ij})_{0 \le i,j < \infty}$ is called the *transition matrix.*

Note that all p_{ij} are non-negative and

$$\sum_{j=0}^{\infty} p_{ij} = 1$$

so that $P(1; i, \cdot)$ is a probability for all $i \in \mathbb{N}$. Matrices which have this property are called *stochastic matrices.* It implies that 1 is an *eigenvalue* with *eigenvector* $e = (1, ..., 1, ...)^T$. In fact, the spectral radius of P, $\rho(P) := \sup\{|\lambda| : Px = \lambda x \text{ for some } x \ne 0\}$, is 1, i.e. 1 is the largest eigenvalue of P.

Indeed: (for example, consider a finite chain) Suppose $Px = \lambda x$ implying $\lambda x_i = \sum_{j=0}^{n} p_{ij} x_j$. Clearly, $|\lambda||x_i| \le \sum_{j=0}^{n} p_{ij}|x_i| \le (\sum_{j=0}^{n} p_{ij})\|x\|_\infty \le \|x\|_\infty$.

Now pick i so that $|x_i| = \|x\|_\infty$. It follows then that $|\lambda| \le 1$.

The Chapman-Kolmogorov equation

$$P(s + t; i, j) = \sum_{h=0}^{\infty} P(s; i, h)P(t; h, j)$$

implies that, for all i and j,

$$p_{ij}^{(2)} = \sum_{k=0}^{\infty} p_{ik}p_{kj},$$

$$p_{ij}^{(3)} = \sum_{k=0}^{\infty} p_{ik}^{(2)} p_{kj} = \sum_{k=0}^{\infty} p_{ik}p_{kj}^{(2)}$$

so that n-step transition probabilities are the elements of the n-th power of the transition matrix:

$$P^n = (p_{ij}^{(n)})_{0 \le i,j < \infty}.$$

We are interested in the steady-state behavior of the chain, i.e. do the limits $\lim_{n\to\infty} p_{ij}^{(n)}$ exist? are they independent of the initial distribution?

Definition 2. (limiting distribution, ergodic) Let $X_n, n \in \mathbb{N}$, be a homogeneous Markov chain. We say X has a limiting distribution if for all states i, the limits $\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$ exist, are independent of i and form a probability distribution. If all π_j are positive, then we call the chain *ergodic* ("time-average" is equal to "space-average").

By taking the limit as $n \to \infty$ in

$$p_{ij}^{(n+1)} = \sum_{k=0}^{\infty} p_{ik}^{(n)} p_{kj},$$

we obtain

$$\pi_j = \sum_{k=0}^{\infty} \pi_k p_{kj}.$$

It follows that a limiting distribution is *stationary*, i.e. $\pi = \pi P, \pi = (\pi_0, \pi_1, ...)$ written as a row-vector. In fact, stationary distributions are left-eigenvectors to eigenvalue 1.

If the chain is finite, every sequence (π_j) will be a distribution, in the infinite case it is possible that $\pi_j = 0$ for all j.