Stochastic signals and systems

Lecture 2.

Prediction, innovation and the Wold decomposition

REVIEW from last Lecture

Wide sense stationary process

Random variables: defined over a probability space $(\Omega, \mathcal{F}, P)$.

A discrete time stochastic process: $y = (y_n)$, with $-\infty < n < +\infty$.

$y = (y_n)$ is **wide sense stationary**, w.s.st. for short, if

$\mathrm{E} y_n = 0$, $\mathrm{E}(y_n^2) < +\infty$ and $r(\tau) = \mathrm{Cov}(y_{n+\tau}, y_n)$ is independent of n.

For any $p \in \mathbb{N}$ the autocovariance matrix R is a**Toeplitz matrix**:

$$R = \begin{pmatrix} r(0) & r(1) & \dots & \dots & r(p-1) \\ r(1) & r(0) & r(1) & \dots & r(p-2) \\ \vdots & \vdots & \ddots & \ddots & \\ r(p-1) & r(p-2) & \dots & \dots & r(0) \end{pmatrix}$$

Predition based on finite past

Problem: Predict y_n based on $y_{n-1}, \ldots, y_{n-p}$.

Proposition

*If R is **nonsingular**, then the LSQ linear predictor of y_n in terms of $y_{n-1}, \ldots, y_{n-p}$ is uniquely defined as*

$$\widehat{y}_n = \sum_{k=1}^{p} \alpha_k y_{n-k},$$

where $\alpha = (\alpha_1, \ldots, \alpha_p)^T$ is the solution of

$$R\alpha = r$$

with $r = (r(1), \ldots r(p))^T$.

PREDICTION II.

Prediction from the infinite past

Consider the "impractical" prediction problem with $p = \infty$.

Define the linear space of finite linear combinations of past values:

$$\mathcal{L}_{n-1}^{y} = \left\{ \sum_{i=1}^{k} \alpha_i y_{n-i} \; : \; \alpha_1, \ldots, \alpha_k \in \mathbb{R}, \quad k = 1, 2, \ldots \right\}.$$

The closure of $\mathcal{L}_{n-1}^{y}$ in $L_2(\Omega, \mathcal{F}, P)$ is denoted by H_{n-1}^{y}.

The best linear predictor of y_n in terms of the infinite past is defined as

$$\widehat{y}_n = (y_n | H_{n-1}^{y}).$$

It can be shown that

$$(y_n | H_{n-1}^{y}) = \lim_{p \to \infty} (y_n | H_{n-1,n-p}^{y}) \quad \text{in} \quad L_2(\Omega, \mathcal{F}, P).$$

The innovation process

A key object in the theory of w.s.st. processes is the following:

Definition

*The **innovation process** of (y_n) is defined as:*

$$e_n = y_n - (y_n | H^y_{n-1}).$$

e_n expresses the information of y_n not contained in $(y_{n-1}, y_{n-2}, ...)$.

Exercise. Prove that (e_n) is a w.s.st. orthogonal process.

Autoregressive processes

Assume that a **finite segment** of past values *is sufficient* to compute $\widehat{y}_n$:

$$\widehat{y}_n = \sum_{k=1}^{p} a_k y_{n-k}.$$

We have $y_n = \widehat{y}_n + e_n$, thus

$$y_n = \sum_{k=1}^{p} a_k y_{n-k} + e_n.$$

Autoregressive processes II.

Assume that a **finite segment** of past values is sufficient to compute $\widehat{y}_n$:

$$y_n = \sum_{k=1}^{p} a_k y_{n-k} + e_n. \tag{1}$$

Definition

A wide-sense stationary process (y_n) satisfying (1) is called an ***autoregressive*** *or* ***AR process****.*

If $a_p \neq 0$ then p is called the **order** of the process.

y is an AR(p) process.

SINGULAR PROCESSES

Singular processes

A "*truly random*" process has a non-trivial innovation, i.e. $e_n \neq 0$.

Thus $H_n^y \supset H_{n-1}^y$ in a strict sense.

A process with $y_n - \widehat{y}_n = e_n = 0$ is called a **singular process**.

Then have: $H_n^y = H_{n-1}^y \quad \text{for all} \quad n$.

It can be shown, that if $H_n^y = H_{n-1}^y$ for a single n, then it is true for all n.

Singular process, example

Consider the complex-valued process (y_n):

$$y_n = \xi e^{in\omega}, \qquad n = 0, \pm 1, \pm 2, \dots$$

where $\omega \in (0,\, 2\pi)$ is a fixed frequency, ξ is a complex-valued r.v. with

$$\mathrm{E}\xi = 0, \quad \mathrm{E}|\xi|^2 = \sigma^2 < +\infty.$$

(y_n) is w.s.st.: $\mathrm{E}y_n = 0$ and

$$\mathrm{E}y_{n+\tau}\overline{y_n} = \sigma^2 e^{i\tau\omega}.$$

Exercise. (**HW**) Show that $y_n = \xi e^{in\omega}$ is singular.

(*Hint:* If $\dim(H_n^y) < \infty$ for some n, then y is singular.)

Singular processes II.

Note that $r(\tau) = \sigma^2 e^{i\tau\omega}$ does not decay in absolute value, as τ increases.

Consider now a finite sum of complex-valued singular processes:

$$y_n = \sum_{k=1}^{m} \xi_k e^{in\omega_k}. \tag{2}$$

Here $\omega_k \neq \omega_j$ for $k \neq j$, and $\mathrm{E}\xi_k = 0$, $\mathrm{E}|\xi_k|^2 = \sigma_k^2 < +\infty$. Assume also, that $\mathrm{E}\xi_k \bar{\xi}_j = 0$ for $k \neq j$.

Singular processes, III.

The auto-covariance function of $y_n = \sum_{k=1}^{m} \xi_k e^{in\omega_k}$ is:

$$\mathrm{E} y_{n+\tau} \bar{y}_n = \sum_{k=1}^{m} \sigma_k^2 e^{i\tau\omega_k} = r(\tau).$$

It follows that the process (y_n) is w.s.st.

The variance of y_n is obtained by setting $\tau = 0$:

$$\mathrm{E}|y_n|^2 = \sum_{k=1}^{n} \sigma_k^2.$$

The values σ_k^2 show how the energy of y_n is spread among frequencies.

Singular processes, example

A simple example for a *real-valued* singular process is:

$$y_n = \cos(\omega n + \varphi) \qquad \omega \neq 0,$$

where φ is a random phase with uniform distribution on $[0, 2\pi]$.

Exercise. Show that (y_n) is a wide sense stationary singular process.

(*Hint*: Apply the identity $\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta)$).

WOLD DECOMPOSITION

Wold decomposition, I.

Let us consider a process y which is *not singular*, and write

$$y_n = e_n + (y_n | H^y_{n-1}).$$

We have $H^y_n = H^y_{n-1} \oplus e_n$ for all n, thus decompose

$$(y_n | H^y_{n-1}) = v_{n-1} + (y_n | H^y_{n-2}), \qquad v_{n-1} = c_1 e_{n-1}.$$

Iterating this decomposition we get, with $c_0 = 1$,

$$y_n = \sum_{k=0}^{p} c_k e_{n-k} + (y_n | H^y_{n-p-1}).$$

$$y_n = \sum_{k=0}^{p} c_k e_{n-k} + (y_n | H^y_{n-p-1}).$$

To deal with the residual term define the prehistory of y as:

$$H^y_{-\infty} = \bigcap_{m \geq 0} H^y_{-m}.$$

Definition

The process (y_n) is ***completely regular****, if*

$$H^y_{-\infty} = \{0\}.$$

Define the w.s.st. processes

$$y^s_n = (y_n | H^y_{-\infty}) \qquad \text{and} \qquad y^r_n = \sum_{k=0}^{\infty} c_k e_{n-k}.$$

Decomposition: $y_n = y_n^r + y_n^s$.

$$y_n = \sum_{k=0}^{\infty} c_k e_{n-k} + (y_n | H_{-\infty}^{y}) = y_n^r + y_n^s.$$

The processes (y_n^s) and (y_n^r) are orthogonal, $y^s \perp y^r$ meaning that

$$y_n^s \perp y_m^r \qquad \text{for all} \quad n,\ m.$$

The process (y_n^s) is singular and $H_{-\infty}^{y^s} = H_{-\infty}^{y}$.

For the process $y^r = (y_n^r)$ we have $H_{-\infty}^{y^r} = \{0\}$.

Completely regular process

Let (y_n) be a completely regular process, i.e.

$$H_{-\infty}^{y} = \{0\}.$$

Then for (y_n) we have:

$$y_n = \sum_{k=0}^{\infty} c_k e_{n-k}$$

with $c_0 = 1$, and $\sum_{k=0}^{\infty} c_k^2 < \infty$. Moreover, we have

$$H_n^e = H_n^y \qquad \text{for all} \quad n.$$

Wold decomposition. Result

Proposition (Wold decomposition of a w.s.st. process)

Any wide sense stationary process can be decomposed as

$$y_n = y_n^r + y_n^s,$$

where (y_n^r) *is* ***completely regular****,* (y_n^s) *is* ***singular*** *and* $y^s \perp y^r$*.*
Moreover

$$H_{-\infty}^{y} = H_{-\infty}^{y^s}.$$

The singular component of the process contains 'a priori' randomness, or randomness in the distant past or prehistory of y.

Corollary

Proposition

Let (e_n) *be the innovation process of a completely regular process* (y_n)*. Then*

$$y_n = \sum_{k=0}^{\infty} c_k e_{n-k} \qquad \text{with} \quad c_0 = 1,$$

with $\sum_{k=0}^{\infty} c_k^2 < \infty$*, and*

$$H_n^e = H_n^y \qquad \text{for all} \quad n.$$

A note on singular processes

Remark. Not all singular processes are of the form given above.

A fascinating result: any process passed through a so-called bandpass filter becomes singular !

Mixtures

A mixture is defined as $y_n(\xi)$, where ξ is a pre-selected r.v. representing distant past.

Example: Select a gambling machine randomly in Las Vegas.

Non-trivial mixtures are not completely regular:

$$H^y_{-\infty} \neq \{0\}.$$