MAX. ENTROPY ESTIMATION OF SPECTRUM

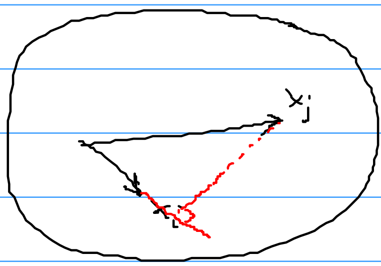
Stationary 2^{nd} order process $\{X_1, X_2, \dots\}$ $\quad E X_i = 0 \;;\; E|X_i|^2 < \infty$

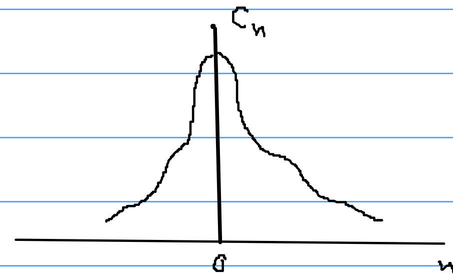
$$E[X_i X_j]^2 \leq E[X_i^2]\,E[X_j^2] \qquad (\text{Schwarz Inequality})$$

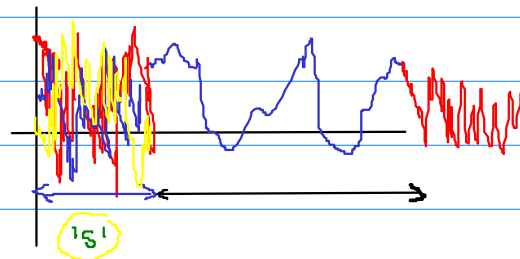
$< \infty$

$E[X_i\, X_{i+n}] = C_n$

$|C_n|^2 < \infty$







$\{X_1, \dots, X_N\}$

LPC

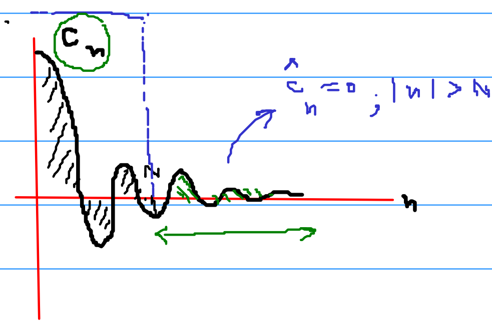
Reminder: $C_n = E[X_k X_{k+n}] \rightarrow \Gamma(\lambda) = \sum_{n=-\infty}^{\infty} C_n e^{-i2\pi\lambda n}$ = spectral power density or spectrum

$$\hat{\Gamma}(\lambda) = f(X_1, \dots, X_N)$$

$$\hat{\Gamma}(\lambda) = \sum_{n=-\infty}^{\infty} \hat{C}_n \cdot e^{-i2\pi\lambda n}.$$

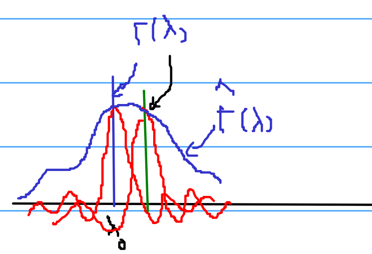
$$\hat{C}_n = \frac{1}{N-1}\sum_{k=1}^{N-1} X_k X_{k+n}$$

$n > N$



$$\hat{C}_n = C_n \cdot \phi_n \xrightarrow{\mathcal{F}} \Gamma(\lambda) \otimes \tilde{\phi}(\lambda)$$

Rect. w, d = N



$$\left.\begin{aligned} &f(\lambda) \geq 0 \\ &\begin{cases} \Phi_1[f] = 0 \\ \dots \\ \Phi_M[f] = 0 \end{cases} \end{aligned}\right\} \rightarrow \hat{f}(\lambda)$$

$$0 \leq E[(X - \alpha Y)^2] = E[X^2] - 2\alpha E[XY] + \alpha^2 E[Y^2]$$

$$= C - 2\alpha B + \alpha^2 A$$

$$\rightarrow B^2 - AC \geq 0$$

$$\rightarrow E[XY]^2 \leq E[X^2]\,E[Y^2]$$

Max. Entropy Spectrum Estimation

Inf. theory → Coding

$\{X_i\}_{i=1}^{\infty}$ stoch. proc.

$f_n(X_{1:n})$, $n=1,2,\dots$ partial distributions

$$h_n(x_1,\dots,x_n) = h(X_{1:n}) = -\int f(x_{1:n}) \log f(x_{1:n})\, dx^n = h[f] \quad (1)$$

JAYNES
SHORE & JOHNSON

$$h = \lim_{n\to\infty} \frac{h_n(X_{1:n})}{n} \quad \text{Exists for stat. sources.} \quad (2)$$

$f(x_{1:n})$

Shannon-McMillan-Breiman

$\sim \exp -\frac{1}{2} x^T C^{-1} x$

* For gaussian process

$$h = \frac{1}{2}\log[2\pi e] + \frac{1}{4\pi}\int_{-\pi}^{\pi} \log[2\pi\Gamma(\lambda)]\, d\lambda \quad (3)$$

$C_{ij} = \mathbb{E}[x_i x_j]$

BURG'S ALGORITHM
↓
Cover, Thomas,
Haykin

$$\Gamma(\lambda) = \frac{1}{2\pi}\sum_{l=-\infty}^{\infty} \sigma(l)\, e^{-i\lambda l} = \text{spectral density}$$

$$\sigma(l) = \mathbb{E}\{x_n \cdot x_{n\pm l}\} = C_l \quad (4)$$

Problem: Estimate $\Gamma(\lambda)$ or $\sigma(l)$ for all $l = 0,1,2,\dots$

Subject to constraints:

$$\begin{cases} \sigma(0)=\alpha_0, \dots, \sigma(p)=\alpha_p, \quad \hat{\sigma}(p+1), \hat{\sigma}(p+2), \dots = ?? \\ \{\sigma(l)\}_{l=-\infty}^{\infty} \quad \text{Auto-Covariance} \end{cases}$$

THEOREM

The stoch. proc. $\{X_i\}$ that maximizes the diff Entropy subject to corr. constraint $\mathbb{E}[X_i X_{i+k}] = \alpha_k$, $k = 0:p$, $i = 1,2,\dots$ (5)

is the p^{th} order gauss. Markov process. satisfying these constraints

Rem. $\{X_i\}$ is not assumed gaussian, nor stationary

Proof.

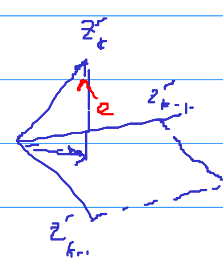
$X_{1:n}$ any coll. of rv's satisfy. (5)

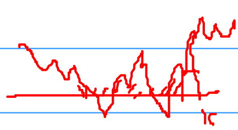
$Z_{1:n}$ zero mean, Normal with cov. given by (5)

$Z'_{1:n}$ p^{th} order gauss-Markov process. Satisfying (5)

$$z'_k = g_1 z'_{k-1} + \dots + g_p z'_{k-p} + v_k$$

iid





Then for $n \geq p$

$$h(X_{1:n}) \overset{(a)}{\leq} h(Z_{1:n})$$

$$\leq h(Z_{1:p}) + \sum_{k=p+1}^{n} h(Z_k \mid Z_{1:k-1}) \qquad (b) \qquad \text{Chain Rule for Entropy}$$

$$\leq h(Z_{1:p}) + \sum_{k=p+1}^{n} h(Z_k \mid Z_{k-p:k-1}) \qquad (c) \qquad h(A \mid B, C) \leq h(A \mid B)$$

$$= h(Z'_{1:p}) + \sum_{p+1}^{n} h(Z'_k \mid Z'_{k-p:k-1})$$

Markov property of $\{Z'_t\}$

$$\overset{(d)}{=} h(Z'_{1:p}) + \sum_{p+1}^{n} h(Z'_k \mid Z'_{1:k-1}) = h(Z'_{1:n})$$

(a) $D[f \| \varphi] \triangleq \int f \log \frac{f}{\varphi} \geq 0$ $\quad f$: density of $\{X\}$

$\phi(x) = N(X \mid 0, R)$, $R = E[x x^T]$

Conclusion: the p^{th} order Gauss-Markov Process with covariance $\sigma_{0:p}$ has higher entropy $h(Z'_{1:n})$ than any other process satisfying the auto-corr. constraints

$$\Rightarrow \lim_{n\to\infty} \frac{1}{n} h(X_{1:n}) \leq \lim_{n\to\infty} h(Z'_{1:n}) = h$$

for all stoch. process satisfying constraints.

EQUIVALENT CHARACTERISATION:

$$\left\{ X_n = \sum_{i=1}^{p} a_i X_{n-i} + u_n \right. \qquad \{u_n\} \text{ iid}, \quad u_n \sim N(0, \sigma^2)$$

$$\sigma_\ell = E[X_n \cdot X_{n-\ell}]$$

$$\sigma_\ell = \overline{X_n X_{n-\ell}} = \sum_{j=1}^{p} a_j \cdot \underbrace{\overline{X_{n-j} X_{n-\ell}}}_{\sigma_{\ell-j}} + \underbrace{\overline{U_n X_{n-\ell}}}_{\overline{U_n} \cdot \overline{X_{n-\ell}} = 0} \qquad \ell > 0$$

$U_{n-\ell}, U_{n-\ell-1}, \ldots$

$$\begin{cases} \sigma_\ell = \sum_{j=1}^{p} a_j \sigma_{\ell-j} & \ell > 0 \\ \sigma_0 = \sum_{i,j} a_i a_j \sigma_{i-j} + \sigma_u^2 \end{cases}$$

YULE-WALKER EQ.

$$H(\lambda) = \frac{1}{1 - \sum_{j=1}^{p} a_j e^{-i2\pi j\lambda}}$$

$$\Gamma(\lambda) = \frac{\sigma_u^2}{\left|1 - \sum_{j=1}^{p} a_j e^{-i2\pi j\lambda}\right|^2}$$

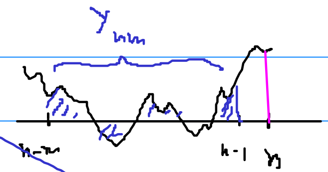
$$\begin{cases} u_n \to \boxed{H(\lambda)} \to x_n \\ \Gamma(\lambda) = \sigma_u^2 \cdot |H(\lambda)|^2 \end{cases}$$

→ Fitting a Gauss-Markov model to data (Linear Prediction)

$\{X_n\}$ stat. process.

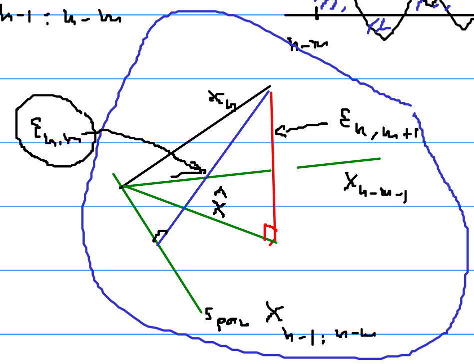
$L = E[X \cdot Y^T] E[Y Y^T]^{-1}$

$\varepsilon_n = X_n - L_{nm} \underbrace{X_{n-1:n-m}}_{Y_{nm}}$



L_{nm} = opt. projection of X_n / $X_{n-1:n-m}$

$\left(L_{nm} = E[X_n \cdot Y_{nm}^T] \cdot R_{Y_{nm}}^{-1} \right.$

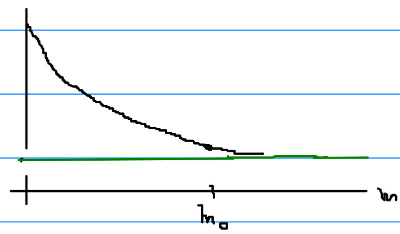


Var(ε_n) decreases with m

$\varepsilon_{nm} = X_n - L_{nm} Y_n$

Var $\varepsilon_{n,m+1} \leq$ Var $\varepsilon_{n,m}$

Wold's theorem $= \sigma^2_{min}$



$m_0 \simeq$ order at which

$X_n \simeq$ Gauss. Markov

↓

$\left\{ \begin{array}{l} \text{Estimate parameters } \hat{a}_k \\ \text{By M.S. method} \end{array} \right.$

Good Ref for Stoch. Process.

1/ GIKHMAN, SKOROKHOD RANDOM PROCESSES (Dover)

2/ TODOROVIC intro St. proc. & Applic (Springer)

(libgen ...) § 6-7

Next Lecture

21/04 $9^H - 11^H$

22/04 $8^H 45 - 10^H 45$