Distribution of Y

$$X_n = Y_n - \sum_{i=1}^{n-1} a_{ni} Y_i$$

- Using vector notation

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & & 0 \\ -a_{21} & 1 & 0 & \cdots & 0 \\ & & \ddots & & \\ -a_{n1} & \cdots & & -a_{n\,n-1} & 1 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}$$

$$X = \underset{\uparrow}{A} Y$$

↑ lower triangular matrix

$$X = f(Y)$$

$$p_y(y) = p_x(f(y)) \underbrace{\left| \frac{\partial f}{\partial y}(y) \right|}_{\text{Jacobian}}$$

$$\left| \frac{\partial f}{\partial y} \right| = \left| \frac{\partial A y}{\partial y} \right| = \underbrace{|A|}_{\text{determinant of } A} = 1$$

$$p_y(y) = \left(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi \sigma_i^2}} \right) \exp\left\{ -\frac{1}{2} Y^T A^T \Lambda A Y \right\}$$

$$\Lambda = \mathrm{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\}$$

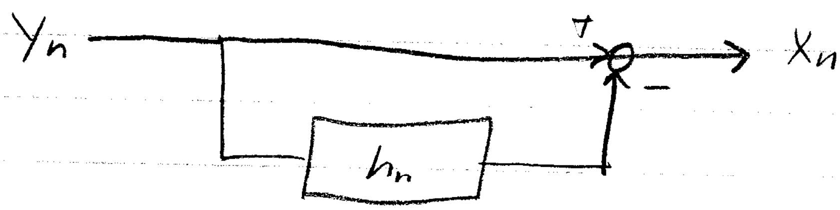
$$R_y^{-1} = A^T \Lambda A$$

Y-Stationary

$$X_n = Y_n - \sum_{i=0}^{n-1} a_{ni} Y_i = Y_n - E[Y_n \mid Y_i \; i<n]$$

Stationary $\Rightarrow a_{ni} = h_{n-i}$

$\Rightarrow E[X_i^2] = \sigma^2$



prediction filter

Properties of X:

1) iid $N(0,\sigma^2)$

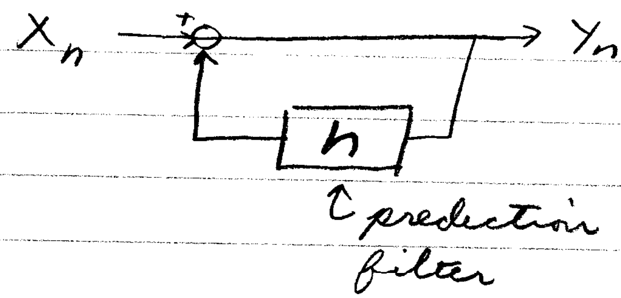
2) $R_X(k) = E[X_n X_{n+k}] = \sigma^2 \delta_k$

3) $S_X(\omega) = \underset{\text{DTFT}}{\mathcal{F}}\{\sigma^2 \delta_k\} = \sigma^2$

power spectrum

Properties of Y:

$$Y_n = X_n + \sum_{i=0}^{n-1} h_{n-i} Y_i \quad \text{(IIR filter)}$$



$$S_Y(\omega) = \left| \frac{1}{1 - H(\omega)} \right|^2 \underbrace{S_X(\omega)}_{= \sigma^2}$$

$$= \frac{\sigma^2}{|1 - H(\omega)|^2}$$

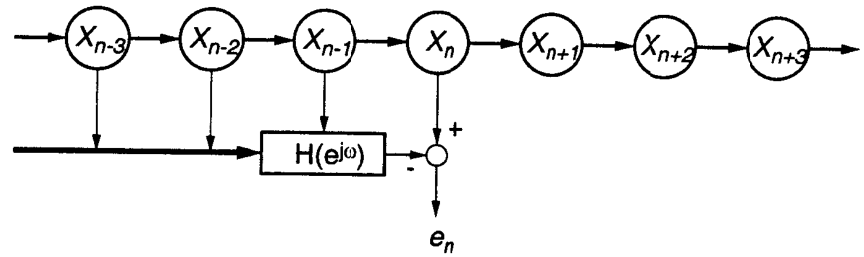
$$H(\omega) = \underbrace{\mathcal{F}}_{\text{DTFT}}\{h\}$$

$$= \sum_{n=-\infty}^{\infty} h_n e^{j\omega n} = H(z)\Big|_{z = e^{j\omega}}$$

Definition: If h_n is FIR, then Y_n is called a Gaussian autoregressive (AR) process.

- If $h_n = 0$ for $n > p$ then Y has order p

Autoregressive (AR) Models



$$e_n = x_n - \sum_{k=1}^{\infty} x_{n-k} h_k$$

- $H(e^{j\omega})$ is an optimal predictor $\Rightarrow \epsilon(n)$ is white noise.
- The density for the N point vector X is given by

$$p_x(x) = \frac{1}{Z} \exp\left\{ -\frac{1}{2} x^t \mathbf{A}^t \mathbf{A} x \right\}$$

where

$$\mathbf{A} = \begin{bmatrix} 1 & & -h_{m-n} \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

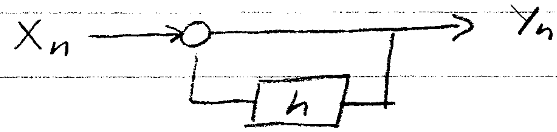
$$Z = (2\pi)^{N/2} |\mathbf{A}|^{-1} = (2\pi)^{N/2}$$

- The power spectrum of X is

$$S_x(e^{j\omega}) = \frac{\sigma_e^2}{|1 - H(e^{j\omega})|^2}$$

Lecture 4

Example:



$h_n = 0 \quad n > p$ (order p)

$X_n = 0$ for $n \le 0$
X_n i i d $N(0, \sigma^2)$ for $n > 0$

- Compute ML estimate of $(\sigma^2; h)$ from $(Y_1, \dots, Y_N)$

$$X_n = Y_n - \sum_{i=n-p}^{n-1} h_{n-i} Y_i$$

$$\begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & & & 0 \\ h_1 & \ddots & & \\ h_2 & \ddots & \ddots & \\ \vdots & & & \\ 0 & \dots \; h_2 & h_1 & 1 \end{bmatrix}}_{\text{Toeplitz}} \begin{bmatrix} Y_1 \\ \vdots \\ Y_N \end{bmatrix}$$

$|A| = 1$

$$p(y) = \prod_{n=2}^{N} \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left\{ \frac{-1}{2\sigma^2} \left(y_n - \sum_{i=n-p}^{n-1} h_{n-i} y_i \right)^2 \right\}$$

Define $$z_n = \begin{bmatrix} y_{n-1} \\ \vdots \\ y_{n-p} \end{bmatrix}$$

$$h = \begin{bmatrix} h_1 \\ \\ h_p \end{bmatrix}$$

$$\log p(y) = -\frac{1}{2\sigma^2} \sum_{n=2}^{N} (y_n - h^t z_n)^2 - \frac{N}{2} \log(2\pi\sigma^2)$$

$$= -\frac{1}{2\sigma^2} \sum_{n=2}^{N} \left(y_n^2 - 2\, y_n h^t z_n + (h^t z_n)^2 \right)$$
$$- \frac{N}{2} \log(2\pi\sigma^2)$$

Note: $(h^t z_n)^2 = h^t z_n z_n^t h$

$$= -\frac{N}{2\sigma^2} \left(\hat{\sigma}_y^2 - 2\, h^t b + h^t R h \right)$$
$$- \frac{N}{2} \log(2\pi\sigma^2)$$

where $\hat{\sigma}_y^2 = \frac{1}{N} \sum_{n=2}^{N} y_n^2$

$$b = \frac{1}{N} \sum_{n=2}^{N} y_n z_n$$

$$R = \frac{1}{N} \sum_{n=2}^{N} z_n z_n^t$$

$$(\hat{\sigma}^2_{ML}, \hat{h}_{ML}) = \underset{(\sigma^2, h)}{\operatorname{argmax}} \log p(y \mid \sigma^2, h)$$

$$= \underset{(\sigma^2, h)}{\operatorname{argmax}} \left\{ -\frac{N}{2\sigma^2}\left(\hat{\sigma}_y - 2h^* b + h^* R h\right) - \frac{N}{2}\log(2\pi\sigma^2) \right\}$$

∘ Maximize with respect to h first.

$$\nabla_h \log p(y) = \frac{-1}{N}\nabla_h\left(h^* R h - 2b^* h\right)$$

$$= \frac{-1}{N}\left(2h^* R - 2b^*\right) = 0$$

$$\hat{h}_{ML} = R^{-1} b$$

∘ Compute ML estimate of σ^2

$$\log p(y \mid \hat{h}_{ML}) = -\frac{1}{2\sigma^2}\sum_{n=2}^{N}\left(y_n - \hat{h}^*_{ML} z_n\right)^2 - \frac{N}{2}\log(2\pi\sigma$$

$$= -\frac{N}{2\sigma^2}\hat{\sigma}_x^2 - \frac{N}{2}\log(2\pi\sigma^2)$$

$$\text{where } \hat{\sigma}_x^2 \triangleq \frac{1}{N}\sum_{n=2}^{N}\underbrace{\left(y_k - \hat{h}^* z_k\right)^2}_{x_n^2}$$

$$\frac{d}{d\sigma^2} \log p(y \mid \sigma, \hat{h}) = \frac{N}{2\sigma^4} \hat{\sigma}_x^2 - \frac{N}{2} \frac{2\pi}{2\pi\sigma^2} = 0$$

$$\left. \frac{\hat{\sigma}_x}{\sigma^2} - 1 \right|_{\sigma^2 = \hat{\sigma}^2_{ML}} = 0$$

$$\hat{\sigma}^2_{ML} = \hat{\sigma}_x$$