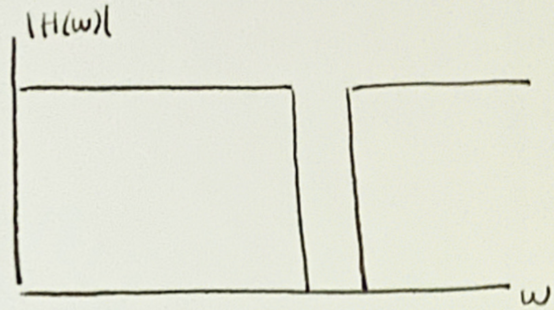
Today's Lecture

* Adaptive Filtering

$x[n] \longrightarrow \boxed{H} \longrightarrow y[n]$

LTI system

FIR filter with N taps



Noise

Goal: have an H that adapts to the input.

constantly update N taps $\longrightarrow$ adaptive filtering

* This topic is at the intersection of signal processing and stochastic processes

We can analyze

- Rates of convergence
- Quality of tracking
- Robustness
- Complexity

Applications

- Noise, interference cancellation
- Channel identification
- Prediction (future values)
- Echo cancellation

Biomedical (ECG interference removal)

Pattern recognition (back-propagation)

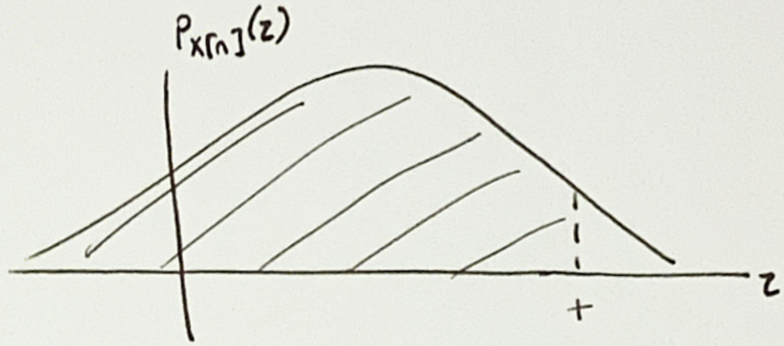
Array processing (beamforming)

Image processing (motion tracking)

$x[n]$ is stochastic. Hence, we can describe $x[n]$ with a cumulative distribution function (CDF):

$$P[x[n] < t] = \int_{-\infty}^{t} P_{x[n]}(z)\, dz$$ where $P_{x[n]}$ is the

probability density function (pdf) that also depends on n.



<u>Mean</u>

$$\mu[n] = E[x[n]] = \int_{-\infty}^{+\infty} z\, P_n(z)\, dz \qquad P_n = P_{x[n]}$$

<u>Auto covariance function</u>

$$c(n, n-k) = E\left[(x[n] - \mu[n])(x[n-k] - \mu[n-k]) \right]$$ if the signals are real.

<u>Auto correlation function</u>

$$r(n, n-k) = E[x[n]\, x[n-k]] = E[x[n-k]\, x[n]]$$

We are primarily interested in Wide Sense Stationary (WSS) processes:

1. $\mu[n] = \mu$ for all n.
2. $c(n, n-k) = c(n-m, n-m-k) = c(k)$
3. $r(n, n-k) = r(|k|)$
4. Finite second moment: $E[|x[n]|^2] < \infty$ for all n.

- WSS if and only if 1 - 4 is true.

Correlation Matrix

$$v[n] = [x[n] \quad x[n-1] \quad \cdots\cdots \quad x[n-M+1]]^T \Rightarrow M\times 1 \text{ vector}$$

$$R = E[v[n]v[n]^T] \Rightarrow M\times M \text{ matrix}$$

$$R_{ij} = E[x[i]x[j]] = E[x[j]x[i]] = R(|i-j|) = r(|i-j|)$$

$$R = \begin{bmatrix} r[0] & r[1] & r[2] & \cdots & r[M-1] \\ r[1] & r[0] & r[1] & & r[M-2] \\ r[2] & r[1] & r[0] & & \vdots \\ \vdots & & & & \\ \vdots & & & & r[1] \\ r[M-1] & & \cdots & r[1] & r[0] \end{bmatrix}$$

$M\times M$

Toeplitz matrix

- Symmetrix matrix
- Toeplitz (equal diagonals)
- Positive semidefinite ($x^T R x \geq 0$ for all x)

$$\begin{bmatrix} x^T \end{bmatrix} \begin{bmatrix} R \end{bmatrix} \begin{bmatrix} x \end{bmatrix} \geq 0$$

Special Models for Stochastic Signals

Parametric models : Moving Average (MA)
Auto Regressive (AR)

ARMA provides a parsimonious description for weakly stationary stochastic process in terms of 2 polynomials:

1 for AR, 1 for MA

Applications: Stock market

Weather forecast

developed by Peter Whittle in his thesis: "Hypothesis testing in time series analysis"

White Gaussian Noise Model

$v[n]$ has pdf $\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}}$ (zero mean)

$E[v[n]] = 0$

$$E[v[n]v[n-k]] = \begin{cases} \sigma_v^2, & k=0 \quad \text{Variance} \\ 0, & k \neq 0 \quad \text{Uncorrelated} \Rightarrow \text{independent?} \end{cases}$$

Moving Average (MA) Model

$$x[n] = v[n] + b_1 v[n-1] + b_2 v[n-2] + \dots + b_k v[n-k]$$

$\{b_i\}$ constants

$\{v[n]\}$ is a white Gaussian noise process

This is equivalent to filtering white Gaussian noise with an FIR filter:

$$x[n] = \sum_{l=0}^{k} v[n-l]\, b[l] \longrightarrow \text{linear function of random disturbances}$$

Auto Regressive (AR) Model

$$x[n] = -a_1 x[n-1] - a_2 x[n-2] - \dots - a_M x[n-M] + v[n]$$

$\{a_i\}$ constants

$\{v[n]\}$ white Gaussian noise process

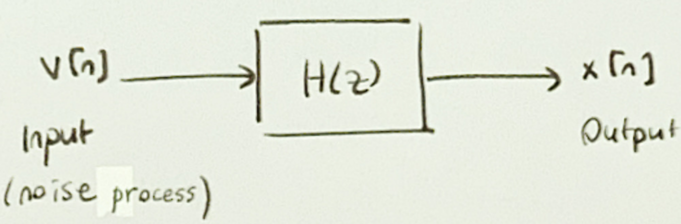
$$v[n] = \sum_{i=0}^{M} a[i]\, x[n-i] = x * a$$

$\rightarrow$ An auto regressive model can be viewed as the output of an IIR filter whose input is white noise.

Auto Regressive Moving Average (ARMA) Model

We generally have

$$x[n] + a_1 x[n-1] + \dots + a_M x[n-M] = v[n] + b_1 v[n-1] + \dots + b_k v[n-k]$$



$$H(z) = \frac{B(z)}{A(z)}$$ applied to noise process

How to estimate the parameters of an AR process?

$a_1, a_2, \dots, a_M, \sigma_v^2$: M+1 parameters

Yule-Walker Equations

$$x[n] + a_1 x[n-1] + \dots + a_M x[n-M] = v[n]$$

- Multiply both sides by $x[n-\ell]$:

$$x[n]x[n-\ell] + a_1 x[n-1]x[n-\ell] + \dots + a_M x[n-M]x[n-\ell] = v[n]x[n-\ell]$$

- Take statistical expectation of both sides:

  Recall that $E[x[n]x[n-\ell]] = r[\ell]$ is the autocorrelation function of a WSS process.

$$r[\ell] + a_1 r[\ell-1] + \dots + a_M r[\ell-M] = E[v[n]x[n-\ell]] = 0$$

  Note: $v[n]$ is statistically independent from $x[n]$ and is zero mean.

- Generate a system of equations for $\ell = 1, 2, \dots, M$

$\ell = 1$

$$r[1] + a_1 r[0] + a_2 r[-1] + \dots + a_M r[1-M] = 0$$

Note that $r[m] = r[-m]$ (symmetric)

$E[x[n+m]x[n]] = E[x[n]x[n+m]]$

$$\ell = 1 \Rightarrow r[1] + a_1 r[0] + a_2 r[1] + \dots + a_M r[M-1] = 0$$

$$\ell = 2 \Rightarrow r[2] + a_1 r[1] + a_2 r[0] + \dots + a_M r[M-2] = 0$$

$\vdots$

$$\ell = M \Rightarrow r[M] + a_1 r[M-1] + a_2 r[M-2] + \dots + a_M r[0] = 0$$

$$\underbrace{\begin{bmatrix} r[0] & r[1] & \dots & r[M-1] \\ r[1] & r[0] & \dots & r[M-2] \\ \vdots & & & \vdots \\ r[M-1] & r[M-2] & \dots & r[0] \end{bmatrix}}_{M \times M \text{ Autocorrelation matrix}} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_M \end{bmatrix}}_{M \times 1} = - \underbrace{\begin{bmatrix} r[1] \\ r[2] \\ \vdots \\ r[M] \end{bmatrix}}_{M \times 1}$$

Yule-Walker equations : $R.a = -r$ R is invertible

$$a = -R^{-1} r$$

How do we estimate correlations using empirical data?

$E[x[n]x[n-k]]$ might not be possible.

$$\hat{r}[i] = \frac{1}{N} \sum_{j=1}^{N} x[j]\,x[j-i]$$ for given realization of data.

Yule-Walker equation with $l=0$

$$x[n]\,x[n] + a_1 x[n]\,x[n-1] + \dots + a_M x[n]\,x[n-M] = v[n]\,x[n]$$

↳ E of both sides

$$r[0] + a_1 r[1] + \dots + a_M r[M] = E[v[n]\,x[n]]$$

$$E[v[n]\,x[n]] = E\left[v[n]\left(-\sum_{i=1}^{M} a[i]\,x[n-i] + v[n]\right)\right]$$

$$= E\left[-\sum_{i=1}^{M} a[i]\,v[n]\,x[n-i] + v^2[n]\right]$$

$$= -\sum_{i=1}^{M} a[i]\,\underbrace{E[v[n]\,x[n-i]]}_{0} + E[v^2[n]]$$

$v[n]$ is additive White Gaussian noise (AWGN) with zero mean:

$$E[v[n]] = 0, \quad E[v^2[n]] = \sigma_v^2$$

$$r[0] + a_1 r[1] + \dots + a_M r[M] = \sigma_v^2$$

$$a_0 = 1, \quad \sum_{k=1}^{M} a_k\, r[k] = \sigma_v^2$$

Other ways to determine/estimate AR parameters

- Estimation of autocorrelation and autocovariances
- Least squares regression problem
- Maximum entropy spectral estimation
- Maximum likelihood estimation (MLE)

Log-likelihood function

$$L(\theta_m) = \log\left(p(\{x[i]\} \mid \theta_m)\right)$$

↳ estimated parameters from Yule-Walke equations (assuming a model order M)

Wiener Filters

Wiener filters are optimal 'linear discrete-time filters' in the mean-square error sense.

$$\underset{\text{Random process}}{x[n]} \longrightarrow \boxed{H} \longrightarrow \underset{\text{Random process}}{y[n]}$$

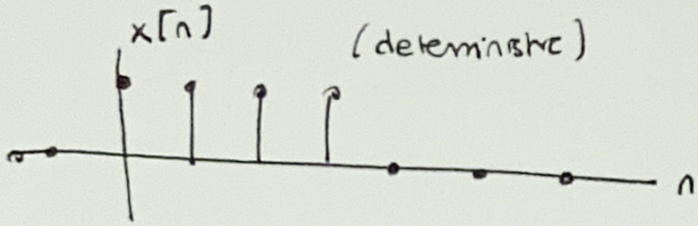
Wiener filter is used to produce an estimate of a desired random process by LTI filtering of an observed noisy process, assuming known WSS signal and noise processes and noise is additive.

Goal : Design H to drive $y[n]$ to a desired output $d[n]$.

The error : $e[n] = d[n] - y[n]$

We want to minimize $E[(e[n])^2]$

Deterministic versus Random Process



(deterministic)

$x[n] \longrightarrow \boxed{H} \longrightarrow y[n]$

$\downarrow$ deterministic $\quad$ $\downarrow$ deterministic $\quad$ $\downarrow$ deterministic

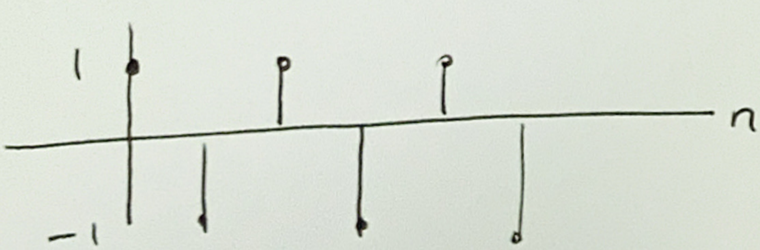
$$X[n] * h[n] = y[n]$$

Random process

$x[n]$ will be sampled from $P_{x[n]}$

For example $\quad P_{x[n]} = \begin{cases} \frac{1}{2} & x[n] = 1 \\ \frac{1}{2} & x[n] = -1 \end{cases}$

Realization of $x[n]$:



$$E[x[n]] = 0$$

$$E[x[n]x[n-k]]$$

$$= 0$$

Another Realization