Lec 1

EE602 Overview

- EE 602 : STATISTICAL SIGNAL PROC.

SSP
- DETECTION ✓
- ESTIMATION ✓
- TIME SERIES ANALYSIS X
  - PROBABILITY/STATISTICS ✓
  - LINEAR ALGEBRA X
  - FOURIER ANALYSIS ✓
  - DSP ✓
  - ETC.... (RANDOM PROCESSES, ✓ VECTORS, ...)

WE WILL STICK TO '✓' and NOT TO 'X'
EACH WILL BE DEALT WITH AS PER COURSE CONTENT

-
-
-

- DETECTION THEORY (HYPOTHESIS TESTING)

  TAKE MEASUREMENTS AND ESTIMATE WHICH STATE THE UNDERLYING PROCESS RESIDES.

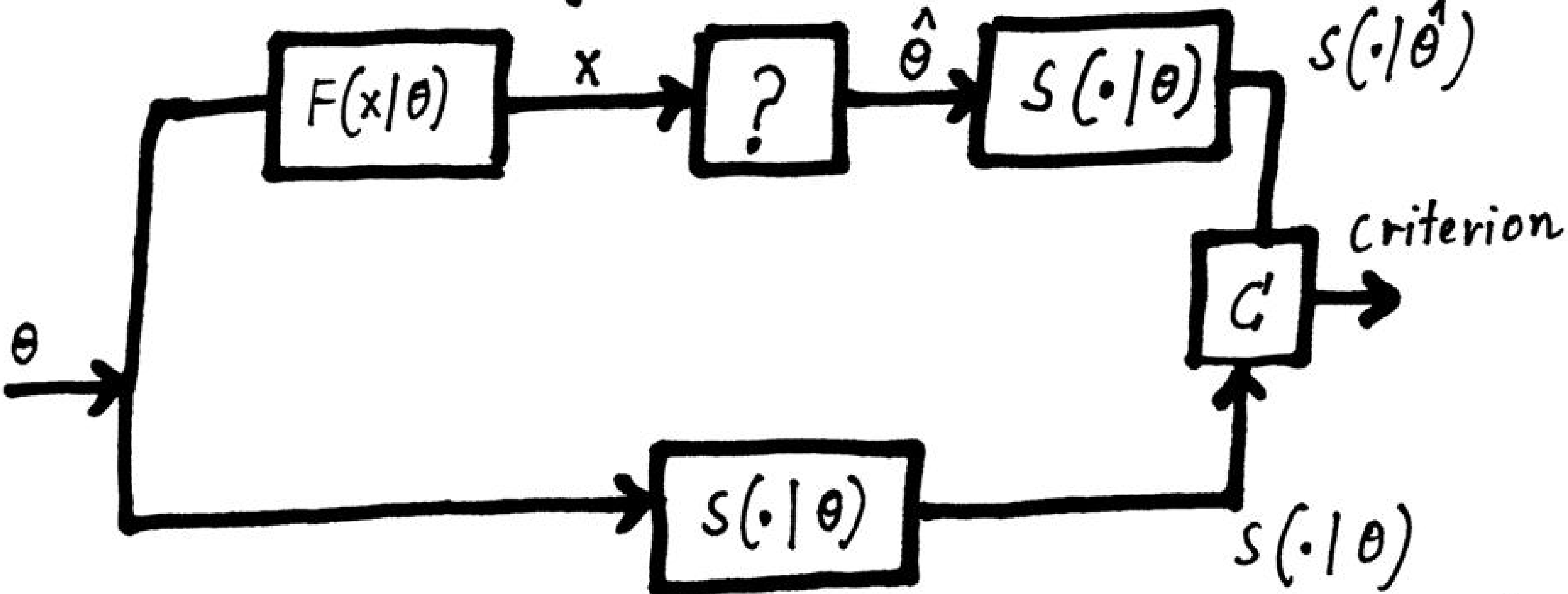
  DETECT SIGNALS IN RADAR/SONAR
  DECODE SYMBOLS IN COMMN
  RECOGNIZE SPEECH
  CLASSIFY PICTURES/IMAGES
  SEARCH VQ CODEBOOKS
  ETC...

- ESTIMATION THEORY (PARAMETER EST)

  TAKE MEASUREMENT & ESTIMATE NUMERICAL VALUE OF VECTOR

  IDENTIFY LINEAR/NON LINEAR S/M
  IDENTIFY COMMN. CX
  IDENTIFY PITCH PERIOD/FILTER IN SPEECH
  ESTIMATE FEATURES IN IMAGES
  ESTIMATE SOURCE DIRECTION IN ARR-AYS.

Structure of Statistical Reasoning



$S(\cdot|\theta)$: Determ. function $F(x|\theta)$: Distribution fn. that generates 'x'.

$S(\cdot|\hat{\theta})$: Estimate of $S(\cdot|\theta)$ C: Minimization Criterion

x: Measurement

EXAMPLE CASES

I. IF $S(\cdot|\theta) = m$; $\theta \in \Theta_m$, $m = 0, 1, \cdots, M-1$
Estimate which of the 'M' classes 'θ' lies?
DECISION/DETECTION THEORY
Criterion is MISCLASSIFICATION RATE/PROBABIL.

II IF $S(\cdot|\theta) = \theta$; THEN PROBLEM IS HOW TO ESTIMA-TE THE PARAMETER ITSELF?
$\Rightarrow$ PARAMETER ESTIMATION

III IF $S(\cdot|\theta) = F(x|\theta)$; Then problem is to est. the distribution function.

IV IF $S(\cdot|\theta) = H\theta$; Estimate the linear model $H\theta$ (MMSE)

V IF $S(\cdot|\theta) = S(e^{jw}|\theta)$; Then estimate PSD/MVE

Eg. Detection Problem

- Let $x_1, x_2, x_3, \cdots x_N$ denote 'N' scalar meas

  with $x_t = \theta s_t + n_t$; $t = 1, \cdots, N$

  s_t : sequence of numbers ; θ : Scalar paramtr.

  Define $\boxed{\begin{array}{l} H_0 : \theta \in \Theta_0 \\ H_1 : \theta \in \Theta_1 \end{array}}$ Observe x_t and decide between H_0 and H_1 is 'Detection pbm.'

  Eg. $\Theta_0 = -\mu$; Then
  $\Theta_1 = \mu$; signals $\pm \mu s_t$ are symbols in commn. s/m

  $t = 1, \cdots, N$ is baud rate

-

Correlation Statistic

- Let n_t be noise like, $t = 1, 2, \cdots, N$

  Define a correlation statistic

  $$C_N = \sum_{t=1}^{N} s_t x_t \; ;$$

  substituting x_t

  $$C_N = \theta \underbrace{\sum_{t=1}^{N} s_t^2}_{E_S} + \underbrace{\sum_{t=1}^{N} s_t n_t}_{\text{correlation term}}$$

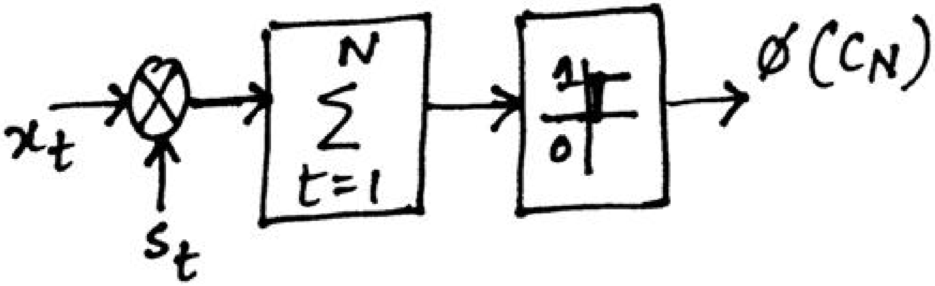
  $\therefore$ $\boxed{C_N \approx \theta E_S}$ and when $\theta = -\mu \Rightarrow C_N \approx -\mu E_S$

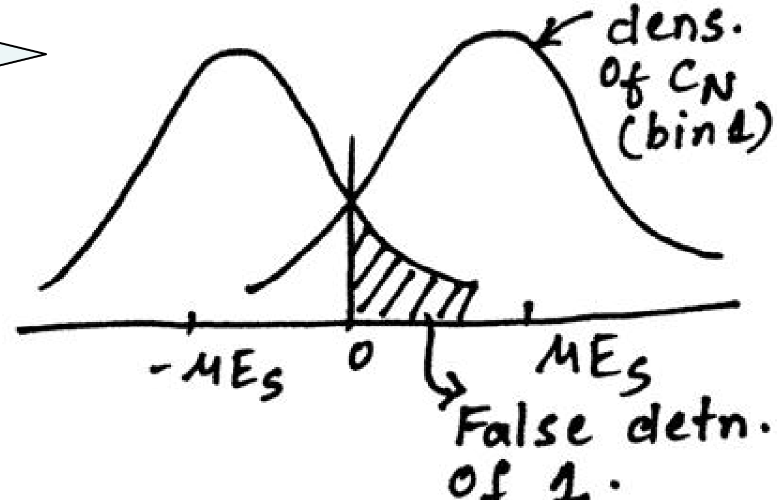
  $\theta = \mu \Rightarrow C_N \approx \mu E_S$

  $$\boxed{\phi(C_N) = \begin{cases} 1 \sim H_1, & C_N > 0 \\ 0 \sim H_0, & C_N \leq 0 \end{cases}} \rightarrow \text{STRATEGY}$$

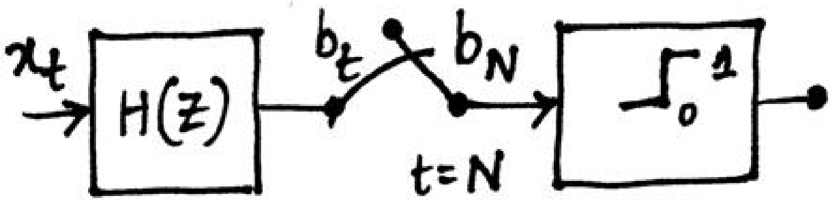
-

Correlation Detector (CD)





CD using a filter



$$\phi(b_N) = \begin{cases} 1 \sim H_1 & \text{if } b_N > 0 \\ 0 \sim H_1 & \text{if } b_N \leq 0 \end{cases}$$

$$H(Z) = \sum_{n=0}^{N-1} h_n z^{-n} \quad \text{and} \quad \underbrace{h_n = s_{N-n}}_{\text{Impulse response is time reversa of 'N'.}}$$

Estimation Problem

Let x_t, $t = 1, \dots, N-1$ are $N-1$ observations such that

$$x_t = \theta + n_t$$

Problem : Estimate θ

Plausible estimator is the Sample Mean

Let $\hat{\theta}_{N-1} = \frac{1}{N-1} \sum_{t=1}^{N-1} x_t$

For θ_N recursively $\sum_{t=1}^{N} x_t = N\theta_n$

But $\sum_{t=1}^{N} x_t = (N-1)\hat{\theta}_{N-1} + x_N = N\hat{\theta}_{N-1} + x_N - \hat{\theta}_{N-1}$

Recursive Estimate of $\hat{\theta}_N$

$$\hat{\theta}_N = \hat{\theta}_{N-1} + \frac{1}{N}\left(x_N - \hat{\theta}_{N-1}\right)$$

Now lets measure performance of the Estimator ϵ_N

$$\epsilon_N = \hat{\theta}_N - \theta = \frac{1}{N}\sum_{t=1}^{N}(x_t - \theta) = \frac{1}{N}\sum_{t=0}^{N-1} n_t$$

If errors n_t are i.i.d with mean $= 0$ var $= \sigma^2$ then

$$E(\epsilon_N) = 0$$

$$E(\hat{\theta}_N) = \theta$$

Mean of the squared Error

$$E(\epsilon_N^2) = E(\hat{\theta}_N - \theta)^2 = \text{var}\ \hat{\theta}_N = \frac{1}{N}\sigma^2$$

$\hat{\theta}_N$ is unbiased $\because$ its mean $= \theta$

Consistent $\because$ $\frac{\sigma^2}{N} \to 0$ as $N \to \infty$

$\therefore$ If n_t, $t = 0, 1, \cdots, N-1$ are iid $[N(0, \sigma^2)]$

then $\hat{\theta}_N$ is distributed as $N[\theta, \sigma^2/N]$

Notations and Terminology

Notations: $X = [x_0, x_1, \cdots, x_{M-1}]$; $x \in R^N$ means x is a point in an N-dimensional space R^N.

X_m : random variable where $m = 0, 1, \cdots, M-1$

x_m : realization of a random variable

$F_\theta(x)$ or $F(x|\theta)$: Distribution of random vector X

θ is a $p \times 1$ vector that parameterizes the distn.

$X : (m, R) \rightarrow X$ has mean m and cov. R

$X : N(m, R) \rightarrow X$ is normally distributed

When θ is random and jointly distd. with x

$$f(x|\theta) = f(x, \theta) / f(\theta) \text{ where } f(\theta) = f(x, \theta)\, dx$$

Quick Look at linear models

$$x = H\theta \; ; \quad H = \underset{[N \times P]}{[h_1, h_2, \cdots h_p]} \cdot \underset{[p \times 1]}{\theta}$$

where

$$H \in R^{N \times P} \text{ and } \theta \in R^{P} \;\therefore\; x \in R^{N}$$

These form basis of Signal Processing

(I) H as a row matrix

$$\underbrace{x_n}_{N \times 1} = \underset{[N \times P]}{c_n^T} \underset{[p \times 1]}{\theta} \; ; \quad c_n^T \to H$$

system matrix is a set of correlators $[c_n^T]_1^N$

n^{th} entry is a correlation of vector c_n with θ

II H as column matrix

$$x = [h_1, h_2, \cdots, h_p] \begin{bmatrix} \theta_1 \\ \vdots \\ \theta_p \end{bmatrix}$$

$$x = \sum_{i=1}^{p} \theta_i h_i$$

Each 'x' is linear combn. of h_i by their co.eff. θ

Each 'h_i' is

eg. $$h_i = \underbrace{[h_1 \cdots\cdots h_p]}_{\text{System matrix}} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = H \delta_i$$

i^{th} pos.

Linear Model Eg. ARMA Impulse Response

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_{p-1} z^{-(p-1)}}{1 + a_1 z^{-1} + \cdots + a_p z^{-p}}$$

Partial fraction expansion yeilds

$$H(z) = \sum_{i=1}^{p} A_i \frac{1}{1 - z^{i} z^{-1}}$$

Corresponding $h(t) = \sum_{i=1}^{p} A_i z_i^{t} \; ; \; t = 0, 1, 2, \ldots$

$$= 0 \, , \; t < 0$$

$\therefore$

$\therefore \quad h = \begin{bmatrix} h_0 \\ h_1 \\ \vdots \\ h_{N-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_p \\ \vdots & \vdots & \cdots & \vdots \\ z_1^{N-1} & z_2^{N-1} & \cdots & z_p^{N-1} \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_p \end{bmatrix}$

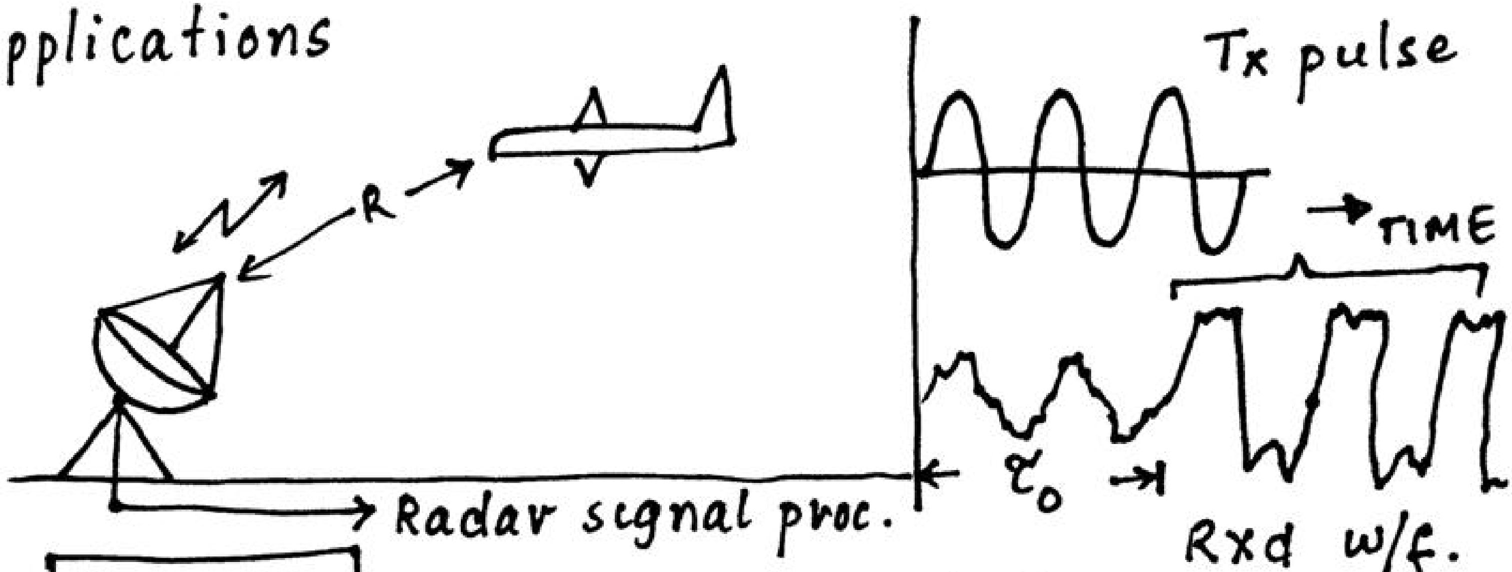
$h = H\theta$ $\qquad\qquad$ $\delta_t \rightarrow \boxed{H(z)} \rightarrow h_t$

h : First N values of $h(n)$ $\quad$ θ : mode weights

H : Vandermonde Matrix $\quad$ θ_i

Introduction to Estimation

- Eg. Applications
- <u>Radar</u>

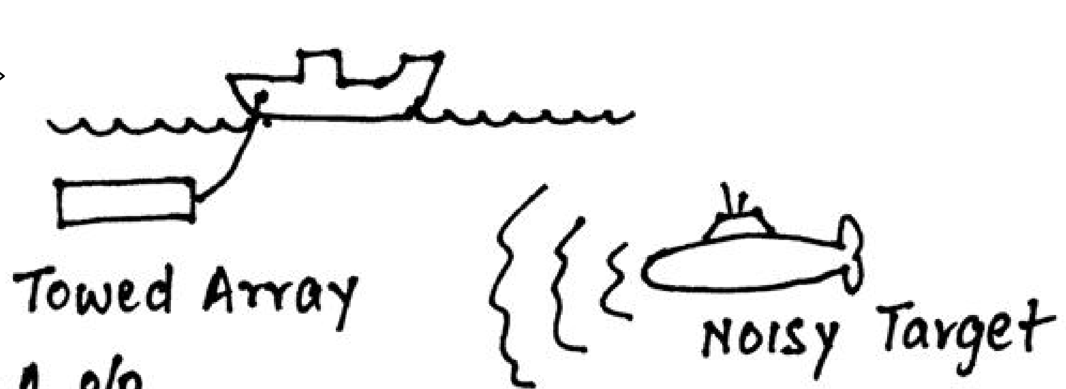


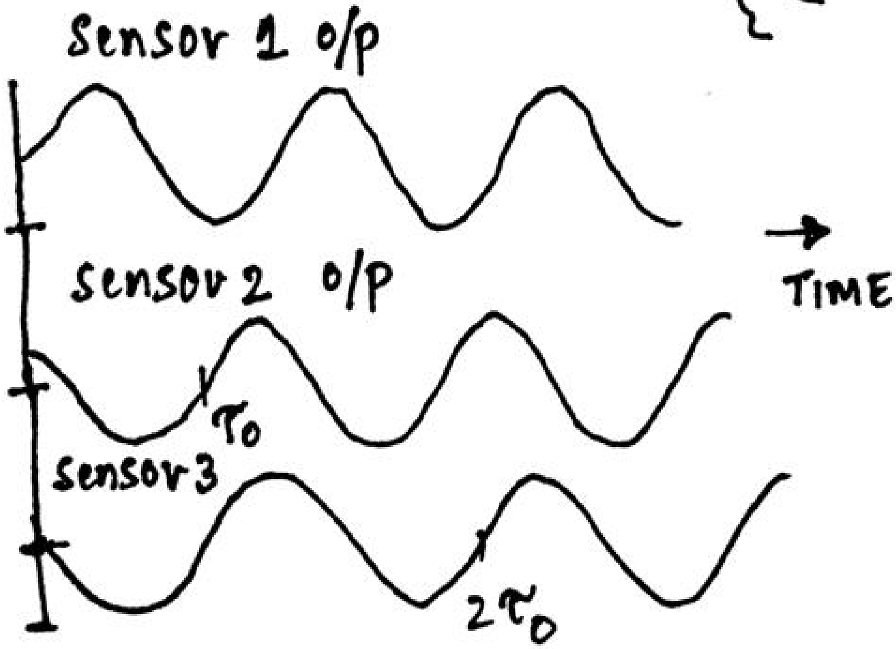
$$\boxed{\tau_0 = \frac{2R}{c}}$$; c : Speed of Electromagnetic propogation

Recieved data can be analyzed using "time Series Analysis"

-

SONAR





Bearing β

$$\beta = \arccos\left[\frac{c\tau_0}{d}\right]$$

c : speed of sound in water

τ_0 : Delay between sensor

d : dist. betn. sensors

Speech Processing

I. Speech sounds (discussion)

II. Spectral envelope modeling (LPC, FFT)

Discuss with examples and plots

Image Processing, Biomedicine, Commns.
Seismology.

Mathematical Estimation Problem

If N point data set $\{x[0], x[1], \ldots, x[N-1]\}$ which depends on 'θ'

then parameter estimation is determining

parameter $\hat{\theta} = g\{x[0], x[1], \ldots, x[N-1])$ where 'g' is some function

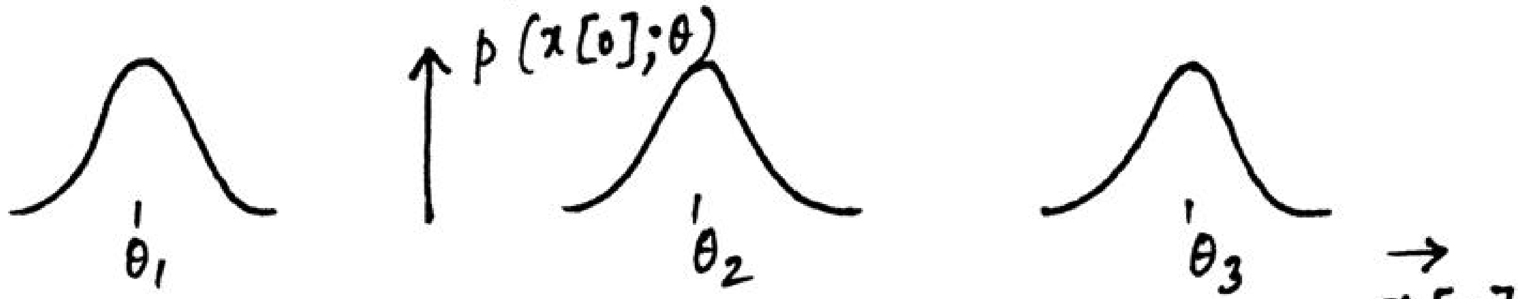
$g(\cdot|\theta)$ can be taken as $p(\cdot|\theta)$

$\therefore \quad p(x[0], x[1], \ldots, x[N-1])$ is like p.d.f.

Mathematical Estimation Problem (Contd.)

If $N = 1$ and θ denotes 'mean' then

$$p(x[0];\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}\left(x[0]-\theta\right)^2\right]$$



if $x[0]$ is $\ominus$ve, then $\theta = \theta_1$ is most likely.

Selection of θ : (a) consistent with constraints.

(b) Mathematically tractable.

Eg: Straight line embedded in random noise

$$x[n] = A + Bn + w[n]; \quad n = 0, 1, \ldots, N-1$$

Reasonable model for $w[n]$ is $\mathcal{N}(0, \sigma^2)$ } → Gaussian

Then: $\theta = [A\ B]^T$ and $x[n] = [x[0], x[1], \ldots, x[N-1]]^T$

$$p(x;\theta) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A-Bn)^2\right].$$

Eg: Dow Jones : A models constant hovering

$B > 0$ models increase in index.

Types of Estimators

Classical Estimation : 'A' is 'deterministic'.

Bayesian Estimation (BE) : 'B' is 'Random', described by a pdf.

BE :

$$\underbrace{p(x;\theta)}_{\text{Family of pdf's.}} = \underbrace{p(x|\theta)}_{\text{conditional pdf.}} \underbrace{p(\theta)}_{\text{prior pdf.}}$$

$\therefore$ Estimate of 'θ' is the value of θ, given a realization of 'x'

Lets assess Estimator performance

Assessing Estimator Performance

- Consider $x[n] = A + w[n]$

  Estimate $\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1} x[n]$; → 'Sample Mean'

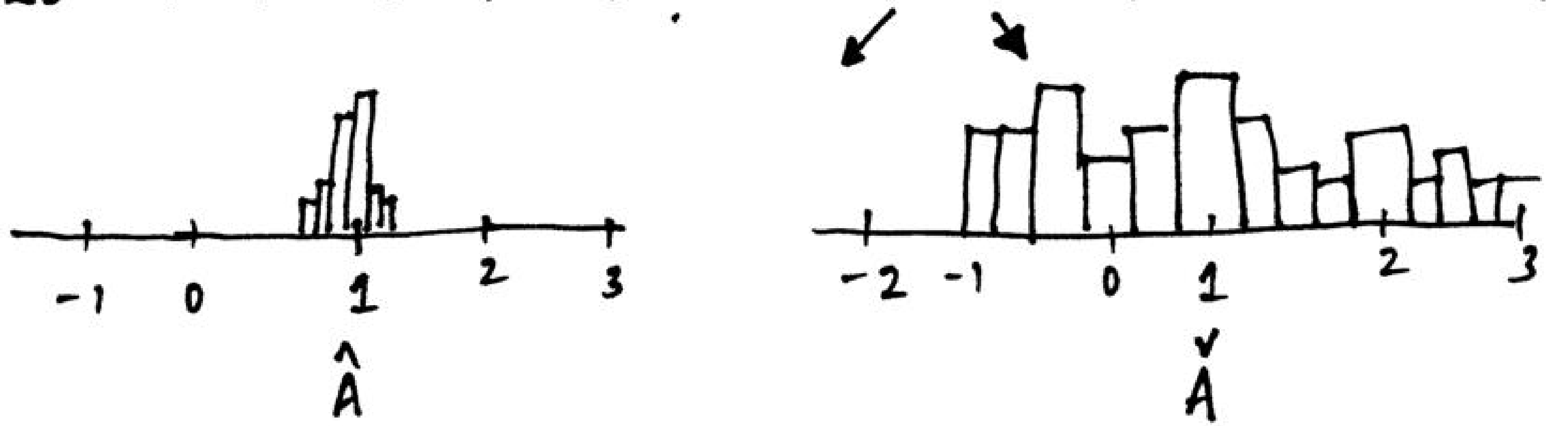
Another estimate $\check{A} = x[0]$, OR USE A HISTOGRAM

HISTOGRAM : No. of times the estimator produces a given range of values

- AND ∴ An approx to the PDF

Assessing Estimator Performance

Is $\hat{A}$ or $\check{A}$ better? 100 realizations with diff. $w[n]$



$\hat{A}$ better ∵ Hist more Concentrated?

Better way: Show variance is less.

HENCE

$$E(\hat{A}) = E\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right)$$

$$= \frac{1}{N}\sum_{n=0}^{N-1} E(x[n]) = \underline{A}$$

$$E(\check{A}) = E(x[0])$$

$$= \underline{A}$$

Assessing Estimator Performance

- $$\mathrm{var}(\hat{A}) = \mathrm{var}\left(\frac{1}{N}\sum_{n=0}^{N-1} x[n]\right) = \cancel{va}\ \frac{1}{N^2}\sum_{n=0}^{N-1}\mathrm{var}(x[n])$$

$$= \frac{1}{N^2} N\sigma^2 = \underline{\frac{\sigma^2}{N}}$$

$$\mathrm{var}(\check{A}) = \mathrm{var}(x[0]) = \underline{\sigma}^2$$

$$\therefore \quad \mathrm{var}(\check{A}) > \mathrm{var}(\hat{A})$$

Hence $\hat{A}$ better estimator

* Performance and computational complexity tradeoff - discuss.

$\hat{A}$ $\check{A}$