Lecture - 21

$\vec{s} = \vec{H}\vec{\theta}$

| θ : $p \times 1$
| s : $N \times 1$
| H : $N \times p$ $(N > p)$

$$J(\theta) = \| x - H\theta \|^2$$

$$\hat{\theta} = (H^T H)^{-1} H^T x$$

$$J_{min} = (x - H\hat{\theta})^T (x - H\hat{\theta})$$

$$= x^T \left(I - H (H^T H)^{-1} H^T \right) x$$

$$= x^T x - x^T \underbrace{H (H^T H)^{-1} H^T x}$$

$$J_{min} = x^T (x - H\hat{\theta})$$

$$J(\theta) = (x - H\theta)^T W (x - H\theta)$$

W : +ve definite matrix.

Constrained Least Square :-

Constraint :- $A\theta = b$

| A : $r \times p$
| b : known $r \times 1$ vector
| θ : $p \times 1$

If $p = 2$.

$$\theta_1 = -\theta_2$$

/

$$\theta_1 + \theta_2 = 0$$

$$A = \begin{bmatrix} 1 & 1 \end{bmatrix}$$

* $$J_c = (x - H\theta)^T (x - H\theta) + \lambda^T (A\theta - b)$$

[$\lambda : r \times 1$]

$$J_c = x^T x - 2\theta^T H^T x + \theta^T H^T H \theta + \lambda^T A \theta - \lambda^T b$$

$$\frac{\partial J_c}{\partial \theta} = -2H^T x + 2H^T H \theta + A^T \lambda \longrightarrow 0$$

$$\hat{\theta}_c = (H^T H)^{-1} H^T x - \frac{1}{2} (H^T H)^{-1} A^T \lambda.$$

$$\hat{\theta}_c = \hat{\theta} - \frac{1}{2} (H^T H)^{-1} A^T \lambda.$$

$$A\hat{\theta}_c = b \Rightarrow A\left(\hat{\theta} - \frac{1}{2} (H^T H)^{-1} A^T \lambda\right) = b$$

$$\frac{\lambda}{2} = \left[A (H^T H)^{-1} A^T\right]^{-1} (A\hat{\theta} - b)$$

$$\hat{\theta}_c = \hat{\theta} - (H^T H)^{-1} A^T \left[A (H^T H)^{-1} A^T\right)^{-1} (A\hat{\theta} - b)$$

$$\hat{\theta} = (H^T H)^{-1} H^T x.$$

1. $s[n] = \begin{cases} \theta_1 & n=0 \\ \theta_2 & n=1 \\ 0 & n=2 \end{cases}$

$s = H\theta$

$\hat{\theta} = (H^T H)^{-1} H^T x$

Data: $\{x[0], x[1], x[2]\}$

$$H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}_{3\times 2}$$

$$s = \begin{bmatrix} \theta_1 \\ \theta_2 \\ 0 \end{bmatrix} \qquad \hat{\theta} = \begin{bmatrix} \quad \end{bmatrix}_{2\times 1}$$

$$\hat{s} = H\hat{\theta} = \begin{bmatrix} \quad \end{bmatrix}_{3\times 1}$$

Constraint: $\theta_1 = \theta_2$

$A\theta = b$

$A = [\quad]$

$$\hat{\theta}_c = \hat{\theta} - A^T (AA^T)^{-1} A\hat{\theta}.$$

* Solving non-linear least squares:-

- Iterative methods are required.

(i) Transformation of parameter

(ii) Separability of parameter..

X = # of students got their own umbrella.

$$X_i = \begin{cases} 1 & \text{if } i \text{ get own umbrella..} \\ 0 & \text{else.} \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_n.$$

$$P(X_i = 1) = \frac{1}{n}.$$

$$E[X_i] = 1 \cdot \frac{1}{n} + 0 \cdot \left(1 - \frac{1}{n}\right) = \frac{1}{n}.$$

(i) $E[X] = E[X_1] + \cdots + E[X_n]$

$$= n \cdot \frac{1}{n} = 1.$$

(ii) $\mathrm{var}(X) = E[X^2] - (E[X])^2$

$$= E[X^2] - 1$$

$$X = \sum_i X_i$$

$$X^2 = \underbrace{\left(\sum_i X_i\right)^2}_{n^2 \text{ terms}} = \underbrace{\sum_i X_i^2}_{n \text{ terms}} + \underbrace{\sum_{\substack{i,j \\ (i \neq j)}} X_i X_j}_{(n^2 - n) \text{ terms}}$$

$$E[X^2] = \sum_i E[X_i^2] + \sum_i \sum_j E[X_i, X_j]$$

$$E[X_i^2] = \frac{1}{n}.$$

$i = 1, j = 2.$

$$P(X_1, X_2 = 1) = P(X_1 = 1)\, P(X_2 = 1 \mid X_1 = 1)$$

$$= \frac{1}{n} \cdot \frac{1}{n-1} = E[X_i X_j]$$

$$E[X^2] = n \cdot \frac{1}{n} + (n^2 - n)\, \frac{1}{n} \cdot \frac{1}{n-1}$$

$$= 2$$

$$\mathrm{var}(X) = 2 - 1 = 1$$

Lecture - 22

The Bayesian:-

θ : Random variable.

"not a deterministic unknown constant"

(i) Helpful to include the prior knowledge about θ

(ii) Apply it when we cannot find MVUE.

$$p(\theta)$$

$x[n] = A + w[n] \qquad n = 0, 1, \cdots, N-1$

$$\hat{A} = \bar{x} \quad (\text{MVUE}) \qquad -\infty < A < \infty$$

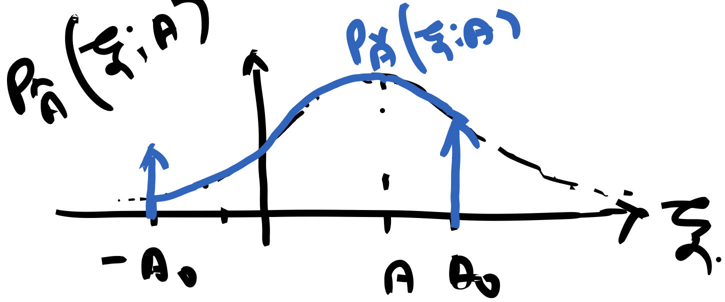
Prior info:- $\boxed{-A_0 \leq A \leq A_0}$

$$\check{A} = \begin{cases} -A_0 & \\ \bar{x} & -A_0 \leq \bar{x} \leq A_0 \\ +A_0 & \end{cases}$$

$$p_{\hat{A}}(\xi;A) = \Pr\{\bar{x} \le -A_0\}\,\delta(\xi + A_0)$$

$$+\, p_{\hat{A}}(\xi;A)\left[u(\xi + A_0) - u(\xi - A_0)\right]$$

$$+\, \Pr\{\bar{x} \ge A_0\}\,\delta(\xi - A_0)$$



$$\mathrm{mse}(\hat{A}) = \int_{-\infty}^{\infty} (\xi - A)^2\, p_{\hat{A}}(\xi;A)\, d\xi$$

$$= \int_{-\infty}^{-A_0} (\xi - A)^2\, p_{\hat{A}}(\xi;A)\, d\xi + \int_{-A_0}^{+A_0} (\xi - A)^2\, p_{\hat{A}}(\xi;A)\, d\xi$$

$$+ \int_{A_0}^{\infty} (\xi - A)^2\, p_{\hat{A}}(\xi;A)\, d\xi$$

$$> \int_{-\infty}^{-A_0} (-A_0 - A)^2\, p_{\hat{A}}(\xi;A)\, d\xi + \int_{-A_0}^{A_0} (\xi - A)^2\, p_{\hat{A}}(\xi;A)\, d\xi$$

$$+ \int_{A_0}^{\infty} (A_0 - A)\, p_{\hat{A}}(\xi;A)\, d\xi.$$

$$= mse(\check{A})$$

$$mse(\hat{A}) > mse(\check{A})$$

Classical Sense

$$mse(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$$

$$= \int (\hat{\theta} - \theta)^2 p(x;\theta)\, dx$$

Bayesian Sense

$$Bmse(\hat{\theta}) = E\left[(\hat{\theta} - \theta)^2\right]$$

$$= \iint (\hat{\theta} - \theta)^2 p(x,\theta)\, dx\, d\theta$$

$p(x,\theta)$ = Joint PDF

$$= p(\theta|x)\, p(x)$$

$$Bmse(\hat{\theta}) = \int \underbrace{\left[\int (\hat{\theta} - \theta)^2 p(\theta|x)\, d\theta\right]}\, p(x)\, dx.$$

$$\frac{\partial}{\partial \hat{\theta}} \int (\theta - \hat{\theta})^2 p(\theta|x)\, d\theta$$

$$= \int \frac{\partial}{\partial \hat{\theta}} (\theta - \hat{\theta})^2 p(\theta|x)\, d\theta$$

$$= \int -2(\theta - \hat{\theta})\, p(\theta|x)\, d\theta$$

$$= -2 \int \theta\, p(\theta|x)\, d\theta + 2\hat{\theta} \underbrace{\int p(\theta|x)\, d\theta}_{1}$$

$$\hat{\theta} = \int \theta \, p(\theta | x) \, d\theta$$

$$\hat{\theta}_{MMSE} = E[\theta | x].$$

$$p(\theta|x) = \frac{p(x|\theta)\, p(\theta)}{\int p(x|\theta)\, p(\theta)\, d\theta}$$

If $p(x, \theta) \rightarrow$ Gaussian

$p(\theta) \longrightarrow$ Gaussian (marginal PDF)

$p(\theta|x) \rightarrow$ Gaussian.

#1. $x[n] = A + w[n]$ $\quad n = 0, 1, \cdots, N-1$

A: R.V.

$$p(A) = \frac{1}{\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{1}{2\sigma_A^2}(A - \mu_A)^2\right]$$

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

$$= \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_n x^2[n]\right) \exp\left[-\frac{1}{2\sigma^2}(NA^2 - 2NA\bar{x})\right)$$

$$p(A|x) = \frac{p(x|A)\, p(A)}{\int p(x|A)\, p(A)\, dA}$$

$$= \frac{\frac{1}{(2\pi\sigma^2)^{N/2}\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{1}{2\sigma^2}\sum_n x^2[n]\right] \exp\left[-\frac{1}{2\sigma^2}\left(NA^2 - 2NA\bar{x}\right)\right] \cdot \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right]}{\int_{-\infty}^{\infty} \cdots \exp\left[-\frac{1}{2\sigma_A^2}(A-\mu_A)^2\right] dA}$$

$$= \frac{\exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2}\left(NA^2 - 2NA\bar{x}\right) + \frac{1}{\sigma_A^2}(A-\mu_A)^2\right)\right]}{\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}\left(\frac{1}{\sigma^2}\left(NA^2 - 2NA\bar{x}\right) + \frac{1}{\sigma_A^2}(A-\mu_A)^2\right)\right] dA}$$

$$= \frac{\exp\left[-\frac{1}{2}Q(A)\right]}{\int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}Q(A)\right] dA}$$

$$Q(A) = \frac{N}{\sigma^2}A^2 - \frac{2NA\bar{x}}{\sigma^2} + \frac{A^2}{\sigma_A^2} - \frac{2\mu_A A}{\sigma_A^2} + \frac{\mu_A^2}{\sigma_A^2}$$

$$= \left(\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}\right)A^2 - 2\left(\frac{N}{\sigma^2}\bar{x} + \frac{\mu_A}{\sigma_A^2}\right)A + \frac{\mu_A^2}{\sigma_A^2}$$

Let $\sigma^2_{A|x} = \dfrac{1}{\dfrac{N}{\sigma^2} + \dfrac{1}{\sigma^2_A}}$; $\mu_{A|x} = \left(\dfrac{N}{\sigma^2}\bar{x} + \dfrac{\mu_A}{\sigma^2_A}\right)\sigma^2_{A|x}$

$$Q(A) = \frac{1}{\sigma^2_{A|x}}\left(A^2 - 2\mu_{A|x}A + \mu^2_{A|x}\right) - \frac{\mu^2_{A|x}}{\sigma^2_{A|x}} + \frac{\mu^2_A}{\sigma^2_A}$$

$$Q(A) = \frac{1}{\sigma^2_{A|x}}\left(A - \mu_{A|x}\right)^2 - \frac{\mu^2_{A|x}}{\sigma^2_{A|x}} + \frac{\mu^2_A}{\sigma^2_A}$$

$$p(A|x) = \frac{\exp\left[-\frac{1}{2\sigma^2_{A|x}}\left(A - \mu_{A|x}\right)^2\right]\exp\left[-\frac{1}{2}\left(\frac{\mu^2_A}{\sigma^2_A} - \frac{\mu^2_{A|x}}{\sigma^2_{A|x}}\right)\right]}{\int_{-\infty}^{\infty}\exp\left[-\frac{1}{2\sigma^2_{A|x}}\left(A - \mu_{A|x}\right)^2\right] - - - \, dA}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2_{A|x}}}\exp\left[-\frac{1}{2\sigma^2_{A|x}}\left(A - \mu_{A|x}\right)^2\right]$$

$$\hat{A} = E(A|x) = \mu_{A|x} = \frac{\frac{N}{\sigma^2}\bar{x} + \frac{\mu_A}{\sigma^2_A}}{\frac{N}{\sigma^2} + \frac{1}{\sigma^2_A}}$$

$$\hat{A}_{mmse} = \frac{\sigma^2_A}{\sigma^2_A + \frac{\sigma^2}{N}}\bar{x} + \frac{\sigma^2/N}{\sigma^2_A + \frac{\sigma^2}{N}}\mu_A$$

Lecture - 23

Bayesian Concept:-

$$\hat{\theta}_{mmse} = E(\theta \mid x)$$

$x[n] = A + w[n]$ $A \sim \mathcal{N}(\mu_A, \sigma_A^2)$

$w[n] \sim \mathcal{N}(0, \sigma^2)$

$$\hat{A} = E(A \mid x) = \frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma^2}{N}} \bar{x} + \frac{\sigma^2/N}{\sigma_A^2 + \sigma^2/N} \mu_A$$

(i) Little data:- $N \rightarrow$ very small.

$\frac{\sigma^2}{N} \rightarrow$ large $\left(\frac{\sigma^2}{N} >> \sigma_A^2\right)$

$$\hat{A} = \mu_A$$

mmse $\longrightarrow$ mean of the prior PDF.

ignores the contribution of the data

(ii) Large data:- $N \rightarrow$ Large.

$\frac{\sigma^2}{N} \rightarrow$ Small

$$\hat{A} \approx \bar{x}$$

mmse ignores the prior knowledge.

$$B_{mse}(\hat{A}) = E\left[(A-\hat{A})^2\right]$$

$$= \iint (A-\hat{A})^2 \, p(x, A) \, dx \, dA$$

$$= \iint (A-\hat{A})^2 \, p(A|x) \, dA \, p(x) \, dx$$

$$\therefore \quad \hat{A} = E[A|x]$$

$$B_{mse}(\hat{A}) = \iint \underbrace{\left[A - E(A|x)\right]^2 p(A|x) \, dA}\, p(x) \, dx$$

$$= \int \text{var}(A|x) \, p(x) \, dx$$

$$= \int \sigma^2_{A|x} \, p(x) \, dx$$

$$= \sigma^2_{A|x} \int p(x) \, dx$$

$$= \sigma^2_{A|x} = \frac{\sigma^2}{N}\left(\frac{\sigma^2_A}{\sigma^2_A + \frac{\sigma^2}{N}}\right)$$

$$B_{mse}(\hat{A}) < \frac{\sigma^2}{N}.$$

<u>Multivariate Gaussian</u>:- If x and y are jointly Gaussian, where x is $k \times 1$ vector

and y is $l \times 1$ vector with mean $[E(x)^T, E(y)^T]^T$

and covariance matrix C

$$C = \begin{bmatrix} [C_{xx}]_{k \times k} & [C_{xy}]_{k \times l} \\ [C_{yx}]_{l \times k} & [C_{yy}]_{l \times l} \end{bmatrix}$$

$$p(x,y) = \frac{1}{(2\pi)^{\frac{k+l}{2}} |C|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\begin{pmatrix} x - E(x) \\ y - E(y) \end{pmatrix}^T C^{-1} \begin{pmatrix} x - E(x) \\ y - E(y) \end{pmatrix}\right]$$

then the conditional PDF $p(y|x)$ is also Gaussian and

$$E[y|x] = E[y] + C_{yx} C_{xx}^{-1} (x - E[x])$$

$$C_{y|x} = C_{yy} - C_{yx} C_{xx}^{-1} C_{xy}$$

* **Bayesian Linear Model:-**

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

$$A \sim \mathcal{N}(\mu_A, \sigma_A^2)$$

$$\vec{x} = \vec{1} A + \vec{w} \qquad w[n] \sim \mathcal{N}(0, \sigma^2)$$

$$\vec{x} = \vec{H}\vec{\theta} + \vec{w}$$

$\vec{x}$: $N \times 1$

$\vec{H}$: $N \times p$

$\vec{\theta}$: $p \times 1$

$\vec{w}$: $N \times 1$ $\sim \mathcal{N}(0, C_w)$

$$z = \left[x^T \; \theta^T\right]^T$$

$$z = \begin{bmatrix} H\theta + w \\ \theta \end{bmatrix} = \begin{bmatrix} [H]_{N\times p} & [I]_{N\times N} \\ [I]_{p\times p} & [0]_{N\times N} \end{bmatrix} \begin{bmatrix} \theta \\ w \end{bmatrix}$$

$$E[x] = E[H\theta + w] = H E[\theta] = H \mu_\theta .$$

$$E[y] = E[\theta] = \mu_\theta \qquad [y = \theta]$$

$$C_{xx} = E\left[(x - E[x])(x - E[x])^T\right]$$

$$= E\left[(H\theta + w - H\mu_\theta)(H\theta + w - H\mu_\theta)^T\right]$$

$$= E\left[(H(\theta - \mu_\theta) + w)(H(\theta - \mu_\theta) + w)^T\right]$$

$$= H \underline{E\left[(\theta - \mu_\theta)(\theta - \mu_\theta)^T\right]} H^T + E\left[ww^T\right]$$

$$= H C_\theta H^T + C_w$$

$$C_{yx} = E\left[(y - E(y))(x - E(x))^T\right]$$

$$= E\left[(\theta - \mu_\theta)(H(\theta - \mu_\theta) + w)^T\right]$$

$$= E\left[(\theta - \mu_\theta)(H(\theta - \mu_\theta))^T\right]$$

$$= - C_\theta H^T$$

Bayesian General Linear Model:-

$$x = H\theta + w \qquad \theta \sim \mathcal{N}(\mu_\theta, C_\theta)$$

$$w \sim \mathcal{N}(0, C_w)$$

$p(\theta|x) \rightarrow$ Gaussian

$$E(\theta|x) = \mu_\theta + C_\theta H^T (H C_\theta H^T + C_N)^{-1} (x - H\mu_\theta)$$

$$C_{\theta|x} = C_\theta - C_\theta H^T (H C_\theta H^T + C_w)^{-1} H C_\theta$$

#. $\vec{x} = \vec{1} A + \vec{w} \qquad A \sim \mathcal{N}(\mu_A, \sigma_A^2)$

$$C_\theta = C_A = \sigma_A^2 I \qquad w \sim \mathcal{N}(0, \sigma^2)$$

$$E(A|x) = \mu_A + \underline{\underline{\sigma_A^2 \vec{1}^T}} (\vec{1} \sigma_A^2 \vec{1}^T + \sigma^2 I)^{-1} (x - \vec{1}\mu_A)$$

using woodbury's identity

$$\left(I + \frac{\sigma_A^2}{\sigma} 1 1^T\right)^{-1} = I - \frac{\frac{\sigma_A^2}{\sigma^2} 1 1^T}{1 + N \frac{\sigma_A^2}{\sigma^2}}$$

$$E(A|x) = \mu_A + \frac{\sigma_A^2}{\sigma^2} \vec{1}^T \left(\vec{I} - \frac{\vec{1}\vec{1}^T}{N + \frac{\sigma^2}{\sigma_A^2}}\right) (\vec{x} - \vec{1}\mu_A)$$

⋮

$$E(A|x) = \mu_A + \frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma^2}{N}} (\bar{x} - \mu_A)$$

$$\operatorname{var}(A|x) = \sigma_A^2 - \sigma_A^2 1^T (\vec{1} \sigma_A^2 \vec{1}^T + \sigma^2 I)^{-1} \vec{1} \sigma_A^2$$

$$= \quad \vdots$$

$$= \frac{\frac{\sigma^2}{N} \sigma_A^2}{\sigma_A^2 + \frac{\sigma^2}{N}}.$$

Lecture - 24

$$\hat{\theta}_{mmse} = E(\theta|x)$$

Nuisance Parameter:-

- Estimation problems are characterized by a set of parameters
- We are interested only on a subset of parameters
- The remaining ones are there to complicate the problem and known as nuisance parameters.

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

$$\sigma^2$$

$$p(\theta|x)$$

$$p(\theta, \alpha|x)$$

$$p(\theta|x) = \int p(\theta, \alpha|x)\, d\alpha$$

$$p(\theta|x) = \frac{p(x|\theta)\, p(\theta)}{\int p(x|\theta)\, p(\theta)\, d\theta}$$

$$p(x|\theta) = \int p(x|\theta,\alpha)\, p(\alpha|\theta)\, d\alpha$$

$$= \int p(x|\theta,\alpha)\, p(\alpha)\, d\alpha \quad \left[\begin{matrix}\text{if } \alpha, \theta \\ \text{indep.}\end{matrix}\right]$$

$*$ $\quad x[n] = A + w[n]$

$$\hat{A} = \underbrace{\frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma^2}{N}}}_{\alpha}\bar{x} + \frac{\sigma^2/N}{\sigma_A^2 + \frac{\sigma^2}{N}}\mu_A$$

$\hat{A} = \alpha\bar{x} + (1-\alpha)\mu_A \quad [0<\alpha<1]$

For deterministic A

$$\text{mse}(\hat{A}) = \text{var}(\hat{A}) + b^2(\hat{A}) \quad \left[\begin{matrix} b(\hat{A}) \\ = E(\hat{A}) - A\end{matrix}\right]$$

$$\text{mse}(\hat{A}) = \alpha^2 \text{var}(\bar{x}) + [E(\hat{A}) - A]^2$$

$$= \alpha^2 \text{var}(\bar{x}) + [\alpha \underbrace{E[\bar{x}]} + (1-\alpha)\mu_A - A]^2$$

$$= \alpha^2 \text{var}(\bar{x}) + [\alpha A + (1-\alpha)\mu_A - A]^2$$

$$= \alpha^2 \frac{\sigma^2}{N} + (1-\alpha)^2 (A - \mu_A)^2$$

if $\quad A \to \mu_A$

$$\text{mse}(\hat{A}) < \text{MVUE}.$$

$$\text{Bmse}(\hat{A}) = E_A[\text{mse}(\hat{A})]$$

$$= \alpha^2 \frac{\sigma^2}{N} + (1-\alpha)^2 E_A\left[(A - \mu_A)^2 \right]$$

$$= \alpha^2 \frac{\sigma^2}{N} + (1-\alpha)^2 \sigma_A^2$$

$$= \frac{\sigma^2}{N} \frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma^2}{N}} < \text{Bmse}(\bar{x}).$$

Bayesian MSE is less.

$$E\left[(\theta - \hat{\theta})^2 \right]$$

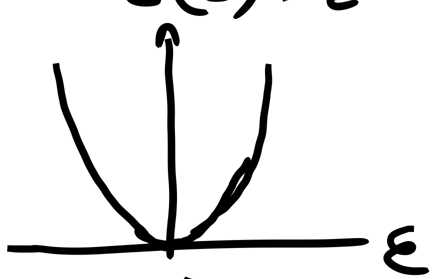
$$\epsilon = \theta - \hat{\theta}$$

$$C(\epsilon) = \epsilon^2 \qquad \left[C(\epsilon) = \text{cost fn} \right]$$

$$E[\epsilon^2] = E[C(\epsilon)]$$

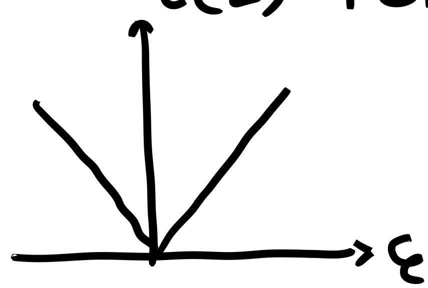
$E[C(\epsilon)]$ = Risk function.

$C(\epsilon) = \epsilon^2$



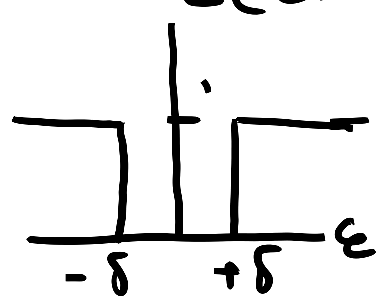
Quadratic error.

$C(\epsilon) = |\epsilon|$



Absolute Error.

$C(\epsilon)$



Hit or Miss error.

$$C(\epsilon) = 0; |\epsilon| < \delta$$
$$= 1; |\epsilon| > \delta$$

Absolute Error: $R = E[C(\epsilon)]$

$$= \iint C(\theta - \hat{\theta})\, p(x, \theta)\, dx\, d\theta$$

$$= \int \Big[\underbrace{\int C(\theta - \hat{\theta})\, p(\theta|x)\, d\theta}_{g(\hat{\theta})} \Big] p(x)\, dx$$

$$g(\hat{\theta}) = \int_{-\infty}^{\infty} |\theta - \hat{\theta}|\, p(\theta|x)\, d\theta$$

$$g(\hat{\theta}) = \int_{-\infty}^{\hat{\theta}} (\hat{\theta} - \theta)\, p(\theta|x)\, d\theta + \int_{\hat{\theta}}^{\infty} (\theta - \hat{\theta})\, p(\theta|x)\, d\theta$$

we need to minimize $g(\hat{\theta})$.

Leibnitz's rule

$$\frac{\partial}{\partial u} \int_{\phi_1(u)}^{\phi_2(u)} h(u, v)\, dv$$

$$= \int_{\phi_1(u)}^{\phi_2(u)} \frac{\partial h(u, v)}{\partial u}\, dv + \frac{d\phi_2(u)}{du}\, h(u, \phi_2(u)) - \frac{d\phi_1(u)}{du}\, h(u, \phi_1(u))$$

$$h(\hat{\theta}, \int \theta) = (\hat{\theta} - \theta)\, p(\theta | x)$$

$$h(u, \phi_2(u)) = (\hat{\theta} - \hat{\theta})\, p(\theta | x) \quad \left[\begin{array}{l} u = \hat{\theta} \\ \phi_2(u) = \hat{\theta} \end{array}\right]$$

$$= 0$$

and $\frac{d\phi_1(u)}{du} = 0$ as $\phi_1(u) = -\infty$

$$\frac{dg(\hat{\theta})}{d\hat{\theta}} = \int_{-\infty}^{\hat{\theta}} p(\theta|x)\, d\theta - \int_{\hat{\theta}}^{\infty} p(\theta|x)\, d\theta = 0$$

$$\int_{-\infty}^{\hat{\theta}} p(\theta|x)\, d\theta = \int_{\hat{\theta}}^{\infty} p(\theta|x)\, d\theta$$

$\hat{\theta}$ is the median of posterior PDF.

$$\Pr\{\theta \le \hat{\theta} | x\} = \frac{1}{2}.$$

Hit - or - miss:

$$c(\epsilon) = 1 \quad \text{for } \epsilon > \delta \text{ and } \epsilon < -\delta$$

$$\epsilon > \delta = \theta > \hat{\theta} + \delta$$

$$\epsilon < \delta = \theta < \hat{\theta} - \delta$$

$$g(\hat{\theta}) = \int_{-\infty}^{\hat{\theta}-\delta} 1 \cdot p(\theta|x)\, d\theta + \int_{\hat{\theta}+\delta}^{\infty} 1 \cdot p(\theta|x)\, d\theta$$

$$\int_{-\infty}^{\infty} p(\theta | x)\, d\theta = 1$$

$$g(\hat{\theta}) = 1 - \int_{\hat{\theta}-\delta}^{\hat{\theta}+\delta} p(\theta | x)\, d\theta$$

$\hat{\theta}$ = mode of $p(\theta | x)$.

Lecture - 25

Quadratic Error $\longrightarrow$ Mean.

Absolute Error $\longrightarrow$ Median.

Hit - or - Miss " $\longrightarrow$ mode.

(i) **MMSE Estimator:-** Mean of the posterior PDF.

$$\hat{\theta} = \int \theta \, p(\theta \mid x) \, d\theta = \int \theta \, p(x \mid \theta) p(\theta) \, d\theta$$

- Requires integration

(ii) **MAP Estimator:-** mode of the posterior PDF.

$$\hat{\theta} = \arg\max_{\theta} p(\theta \mid x)$$

$$= \arg\max_{\theta} \frac{p(x \mid \theta) p(\theta)}{\int p(x \mid \theta) p(\theta) \, d\theta}$$

$$= \arg\max_{\theta} p(x \mid \theta) p(\theta)$$

- No integration

(iii) **Bayesian ML Estimator:**

$$\hat{\theta} = \arg\max_{\theta} p(x \mid \theta) \qquad \left[p(x; \theta) \right]$$

Linear MMSE Estimation :- θ

$$\{x[0], x[1], \dots, x[N-1]\}.$$

$$\vec{x} = [x[0], x[1], \dots, x[N-1]]^T$$

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N \quad — ①$$

Bayesian MSE :- $B_{mse}(\hat{\theta}) = E[(\theta - \hat{\theta})^2]$

\# θ $\quad x[0] \sim \mathcal{N}(0, \sigma^2)$

$$\theta = x^2[0] \quad [\text{Power of } x[0]]$$

$$\hat{\theta} = x^2[0]$$

$$\hat{\theta} = a_0 x[0] + a_1$$

$$B_{mse}(\hat{\theta}) = E[(\theta - \hat{\theta})^2] = E[(\theta - a_0 x[0] + a_1)^2]$$

Differentiation w.r.t a_0 :- $E[(\theta - a_0 x[0] + a_1) x[0]] = 0$

" " a_1 :- $E[(\theta - a_0 x[0] + a_1)] = 0$

$$a_0 E[x^2[0]] + a_1 E[x[0]] = E[\theta x[0]]$$

$$a_0 E[x[0]] + a_1 = E[\theta].$$

$\therefore E[x[0]] = 0 \rightarrow E[\theta\, x[0]] = E[x^3[0]] = 0$

$$a_0 = 0$$

$$a_1 = E[\theta] = E[x^2[0]] = \sigma^2$$

$$\hat{\theta} = a_1 = \sigma^2 \rightarrow \text{does not depend on data.}$$

The minimum MSE

$$B_{mse}(\hat{\theta}) = E[(\theta - \hat{\theta})^2] = E[(\theta - \sigma^2)^2]$$

$$= E[(x^2[0] - \sigma^2)^2] = E[x^4[0]] - 2\sigma^2 E[x^2[0]] + \sigma^4$$

$$\boxed{E(x^4) = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4} \qquad = 3\sigma^4 - 2\sigma^4 + \sigma^4 = 2\sigma^4$$

minimum MSE of $\theta = x^2[0]$ is zero.
So LMMSE is not appropriate for this problem.

- -

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] + a_N$$

$$B_{mse}(\hat{\theta}) = E\left[\left(\theta - \sum_{n=0}^{N-1} a_n x[n] - a_N\right)^2\right]$$

$$\frac{\partial}{\partial a_N} E\left[\left(\theta - \sum_n a_n x[n] - a_N\right)^2\right] = 0$$

$$-2E\left[\theta - \sum_n a_n x[n] - a_N\right] = 0$$

$$a_N = E[\theta] - \sum_n a_n E[x[n]]$$

$$B_{mse}(\hat{\theta}) = E\left\{\left[\sum_n a_n\left(x[n] - E[x[n]]\right) - \left(\theta - E(\theta)\right)\right]^2\right\}$$

$$= E\left\{\left[\vec{a}^T\left(\vec{x} - E[\vec{x}]\right) - \left(\theta - E(\theta)\right)\right]^2\right\}$$

$$= E\left[\vec{a}^T\left(\vec{x} - E[\vec{x}]\right)\left(\vec{x} - E[\vec{x}]\right)^T\vec{a}\right]$$

$$- E\left[\vec{a}^T\left(\vec{x} - E(\vec{x})\right)\left(\theta - E(\theta)\right)\right]$$

$$- E\left[\left(\theta - E(\theta)\right)\left(\vec{x} - E(\vec{x})\right)^T\vec{a}\right] + E\left[\left(\theta - E(\theta)\right)^2\right]$$

$$= \vec{a}^T\vec{C}_{xx}\vec{a} - \vec{a}^T\vec{C}_{x\theta} - \vec{C}_{\theta x}\vec{a} + \vec{C}_{\theta\theta}$$

$$\boxed{\vec{C}_{\theta x} = \vec{C}_{x\theta}^T}$$

$[\vec{C}_{xx}]_{N\times N}$: Covariance matrix of $\vec{x}$

$[\vec{C}_{\theta x}]_{1\times N}$: Cross covariance vector.

$\vec{C}_{\theta\theta}$: variance of θ.

$$\frac{\partial B_{mse}(\hat{\theta})}{\partial \vec{a}} = 2\vec{C}_{xx}\vec{a} - 2\vec{C}_{x\theta} = 0$$

$$\vec{a} = \vec{C}_{xx}^{-1}\vec{C}_{x\theta} \quad — ②$$

$$\hat{\theta} = \vec{a}^T\vec{x} + a_N$$

$$= \left(\vec{C}_{xx}^{-1} \vec{C}_{x\theta}\right)^T \vec{x} + E(\theta) - \vec{C}_{x\theta}^T \vec{C}_{xx}^{-1} E(\vec{x})$$

$$\hat{\theta}_{LMMSE} = E(\theta) + \vec{C}_{\theta x} \vec{C}_{xx}^{-1} \left(\vec{x} - E(\vec{x})\right)$$

$$\boxed{\begin{array}{l} E(\theta) \rightarrow 0 \quad \& \quad E(\vec{x}) \rightarrow 0 \\ \hat{\theta}_{LMMSE} = \vec{C}_{\theta x} C_{xx}^{-1} \vec{x} \end{array}}$$

$$B_{MSE}(\hat{\theta}) = \vec{C}_{\theta\theta} - \vec{C}_{\theta x} \vec{C}_{xx}^{-1} \vec{C}_{x\theta}$$

* $x[n] = A + w[n]$ $\qquad n = 0, 1, \dots, N-1$

$A \sim \mathcal{U}[-A_0, A_0]$

$w[n] \sim \mathcal{N}(0, \sigma^2)$

$$\boxed{\begin{array}{l} \hat{A} = E[A \mid x] \\ \quad = \int A\, p(A \mid x)\, dA \end{array}}$$

$E(\theta) = E(A) = 0$

$E[x[n]] = 0$

$E[\vec{x}] = 0$

$$\vec{C}_{xx} = E\left[x x^T\right] = E\left[(A\vec{1} + \vec{w})(A\vec{1} + \vec{w})^T\right]$$

$$= E[A^2]\, \vec{1}\vec{1}^T + \sigma^2 \vec{I}$$

$$\vec{C}_{\theta x} = E\left[A\vec{x}^T\right] = E\left[A\left(A\vec{1} + \vec{w}\right)^T\right]$$

$$= E\left[A^2\right]\vec{1}^T$$

$$\hat{A} = \vec{C}_{\theta x} C_{xx}^{-1}\vec{x} = \sigma_A^2 1^T\left(\sigma_A^2 \underline{1}\,\underline{1}^T + \sigma^2 \underline{I}\right)^{-1}\bar{x}$$

$$\left[E\left(A^2\right) = \sigma^2\right]$$

$$\hat{A} = \frac{\sigma_A^2}{\sigma_A^2 + \frac{\sigma^2}{N}}\bar{x}$$

Find σ_A^2 and $U[-A_0, A_0]$

Lecture - 28

Detection Theory

Deals with decision making.

- Presence or absence of aircraft.
- Detecting '0' or '1' in BPSK system
- Detecting a spoken word among the group of possibilities. "0", "1", "2", . . . , "9"
- Submarine present?
- Object detection
- Detection of presence of cardiac arrythmia.
-

$$K = T\{x[0], x[1], \ldots, x[N-1]\}$$

- Determine T.
- Make decision based on the value of T.

Hypothesis testing problem

* First two problems are binary hypothesis testing problems
* Third one is multiple hypothesis testing problem

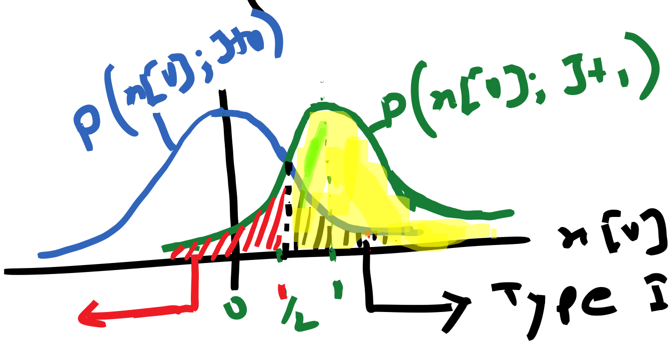
— Simple hypothesis testing

- Classical — $\boxed{\text{Sonar / Radar}}$
  - Neyman-Pearson
- Bayesian
  - minimizing Bayes Risk (Comm & Pattern recognition)

* $x[0] \sim \mathcal{N}(0,1)$ or $\mathcal{N}(1,1)$

$H_0 : \mu = 0$ (Null hypothesis)

$H_1 : \mu = 1$ (Alternate ")



Type II error $P(H_0; H_1)$

Type I error $P(H_1; H_0)$

Th: $\frac{1}{2}$.

if $x[0] > \frac{1}{2}$ choose H_1

if $x[0] < \frac{1}{2}$ choose H_0

$$p(x[0]; H_1) > p(x[0]; H_0)$$

choose H_1

— Type I error can be decreased at the cost of increasing Type II error.

— It is not possible to reduce both kind of errors at the same time.

* $\mathcal{H}_0 : \quad x[0] = w[0]$

$\mathcal{H}_1 : \quad x[0] = s[0] + w[0]$

Signal $s[0] = 1$ $\qquad w[0] \sim \mathcal{N}(0, 1)$

$P(\mathcal{H}_1 ; \mathcal{H}_0)$ = Probability of deciding $\mathcal{H}_1$ when $\mathcal{H}_0$ is true. = Probability of False alarm = P_{FA}

Keep P_{FA} very small : 10^{-8}

Optimal detector requires minimizing

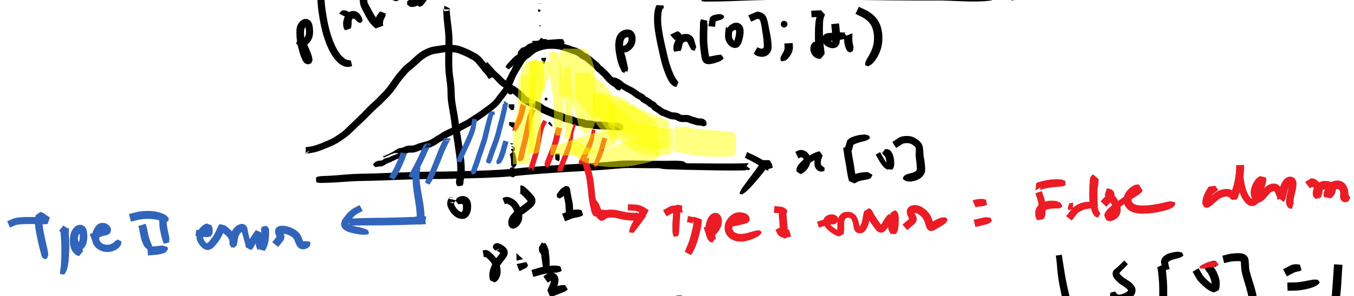
minimizing $P(\mathcal{H}_0 ; \mathcal{H}_1)$ or maximize $(1 - P(\mathcal{H}_0 ; \mathcal{H}_1))$

$= P(\mathcal{H}_1 ; \mathcal{H}_1)$

= probability of detection

$= P_D$

Lecture - 27



$$H_0 : \quad x[0] = w[0]$$

$$H_1 : \quad x[0] = s[0] + w[0]$$

$s[0] = 1$

$w[0] \sim \mathcal{N}(0, 1)$

$$x[0] > \gamma \longrightarrow H_1$$

$$x[0] < \gamma \longrightarrow H_0$$

$$P(H_1 ; H_0) = P_{FA} \quad (\text{keep it small})$$

$$\text{minimize } P(H_0 ; H_1) = \text{maximize } [1 - P(H_0 ; H_1)]$$

$$= \text{maximize } P(H_1 ; H_1)$$

$$= \text{maximize } P_D$$

Neyman-Pearson Principle

* We wish to maximize P_D s.t. $P_{FA} = \alpha$

$$P_{FA} = P(H_1 ; H_0) = \Pr\{x[0] > \gamma ; H_0\}$$

$$* \quad P_x = \frac{1}{\sqrt{2\pi\sigma_x^2}} \exp\left[-\frac{1}{2\sigma_x^2}(x-\mu)^2\right]$$

$\mu = 0, \quad \sigma_x^2 = 1$

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}t^2\right] dt$$

$$Q(x) = 1 - \Phi(x) = \int_{x}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}t^2\right] dt$$

$$Q(x) \approx \frac{1}{\sqrt{2\pi}\,x} \exp\left[-\frac{1}{2}x^2\right]$$

$$P_{FA} = \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}t^2\right) dt = Q(\gamma)$$

$$1 - 10^{-3} = 0.99$$

$$P_{FA} = 10^{-3} \Rightarrow \gamma = 3$$

we decide H_1 if $x[0] > 3$.

so, $P_D = P(H_1; H_1) = Pr\{x[0] > 3; H_1\}$

$$= \int_{\gamma}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(t-1)^2\right] dt$$

$$= Q(\gamma - 1)$$

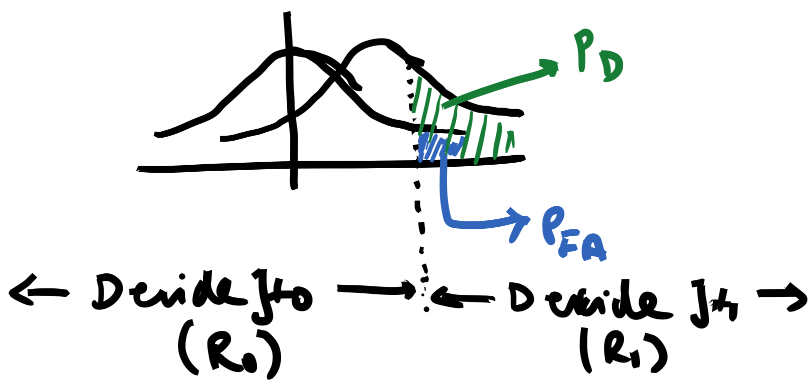
$$= Q(2) = 1 - \Phi(2)$$

$$= 1 - 0.977$$

$$= 0.023.$$

* $\{x[0], x[1], \ldots, x[N-1]\}$

$\mathcal{H}_0$ or $\mathcal{H}_1$



$R_1 = \{\vec{x} : \text{decide } \mathcal{H}_1 \text{ or reject } \mathcal{H}_0\}$

↳ critical region.

$$P_{FA} = \int_{R_1} p(x; \mathcal{H}_0)\, dx = \alpha. \quad \text{— (1)}$$

$$P_D = \int_{R_1} p(x; \mathcal{H}_1)\, dx \quad \text{— (2).}$$

NP-theorem tells us how to choose R_1 if we are given $p(x; \mathcal{H}_0)$, $p(x; \mathcal{H}_1)$ and α.

* <u>Neyman-Pearson Theorem:-</u>

To maximize P_D for a given $P_{FA} = \alpha$ decide $\mathcal{H}_1$ if

$$L(x) = \frac{p(x; \mathcal{H}_1)}{p(x; \mathcal{H}_0)} > \gamma \quad \text{— (3)}$$

where the threshold γ is found from

$$P_{FA} = \int_{\{\bar{x} : L(x) > \gamma\}} p(x; \mathcal{H}_0)\, dx = \alpha \quad \text{— (4)}$$

$L(x)$: Likelihood ratio.

Likelihood Ratio Test (LRT).

#. $\mathcal{H}_0: x[0] = w[0]$
$\mathcal{H}_1: x[0] = s[0] + w[0]$ | $s[0] = 1$, $w[0] \sim \mathcal{N}(0,1)$

$$P_{FA} = 10^{-3}$$

we decide $\mathcal{H}_1$ if $\dfrac{p(x; \mathcal{H}_1)}{p(x; \mathcal{H}_0)} > \gamma$

$$\frac{\frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(x[0]-1\right)^2\right]}{\frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} x^2[0]\right]} > \gamma$$

$$\exp\left[x[0] - \frac{1}{2}\right] > \gamma$$

$$P_{FA} = P(\mathcal{H}_1; \mathcal{H}_0) = \Pr\left\{\exp\left[x[0]-\frac{1}{2}\right] > \gamma; \mathcal{H}_0\right\}$$
$$= 10^{-3}$$

$$\exp\left[x[0] - \frac{1}{2}\right] > \gamma = \exp(\beta)$$

$$x[0] - \frac{1}{2} > \beta$$

$$x[0] > \beta + \frac{1}{2} = \ln\gamma + \frac{1}{2}.$$

$\gamma' = \ln \gamma + \frac{1}{2}$.

we decide $\mathcal{H}_1$ if $x[0] > \gamma'$

$$P_{FA} = \Pr\{x[0] > \gamma'; \mathcal{H}_0\} = 10^{-3}$$

$$\int_{\gamma'}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}t^2\right] dt = 10^{-3}$$

$$\gamma' = 3$$

$$P_D = \Pr\{x[0] > 3; \mathcal{H}_1\}$$

$$= \int_3^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}(t-1)^2\right] dt$$

$$= 0.023$$

$P_{FA} = 0.5 \longrightarrow \gamma' = 0$

$P_D = ?$ 0.84

Lecture - 28

NP Theorem:-

maximize P_D for $P_{FA} = \alpha$

Decide $\mathcal{H}_1$ if

$$L(x) = \frac{p(x; \mathcal{H}_1)}{p(x; \mathcal{H}_0)} > \gamma$$

$$P_{FA} = \int_{\{\vec{x} : L(x) > \gamma\}} p(x; \mathcal{H}_0)\, dx = \alpha .$$

#. $\mathcal{H}_0 : x[n] = w[n]$ $\quad n = 0, 1, \dots, N-1$

$\mathcal{H}_1 : x[n] = A + w[n]$ $\quad n = 0, 1, \dots, N-1$

$[A > 0]$

$w[n] \sim \mathcal{N}(0, \sigma^2)$

under $\mathcal{H}_0$, $\vec{x} \sim \mathcal{N}(0, \sigma^2 \vec{I})$

under $\mathcal{H}_1$, $\vec{x} \sim \mathcal{N}(A\vec{1}, \sigma^2 \vec{I})$

$\mathcal{H}_0 : \vec{\mu} = 0$

$\mathcal{H}_1 : \vec{\mu} = A\vec{1}$

Decide H_1 if

$$\frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2\right]}{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}x^2[n]\right]} > \gamma$$

$$-\frac{1}{2\sigma^2}\left(-2A\sum_{n=0}^{N-1}x[n] + NA^2\right) > \ln\gamma.$$

$$\Rightarrow \frac{A}{\sigma^2}\sum_{n=0}^{N-1}x[n] > \ln\gamma + \frac{NA^2}{2\sigma^2}.$$

$$\Rightarrow \frac{1}{N}\sum_{n=0}^{N-1}x[n] > \boxed{\frac{\sigma^2}{NA}\ln\gamma + \frac{A}{2} = \gamma'}$$

$$\boxed{\bar{x} > \gamma'} \quad [\text{Decide } H_1]$$

$$T(x) = \bar{x} = \frac{1}{N}\sum_n x[n] \quad (\text{Test statistic})$$

$$E[T(x); H_0] = 0$$

$$E[T(x); H_1] = A$$

$$\text{var}(T(x); H_0) = \frac{\sigma^2}{N}.$$

$$\text{var}\left(\frac{1}{N}\sum_n w[n]\right)$$

$$\frac{1}{N^2}\sum_n \text{var}(w[n])$$

$$\text{var}(T(x); H_1) = \frac{\sigma^2}{N}.$$

$$T(x) \sim \begin{cases} \mathcal{N}\left(0, \frac{\sigma^2}{N}\right) & \text{under } \mathcal{H}_0 \\ \mathcal{N}\left(A, \frac{\sigma^2}{N}\right) & \text{under } \mathcal{H}_1 \end{cases}$$

$$P_{FA} = \Pr\{T(x) > \gamma' ; \mathcal{H}_0\}.$$

$$P_{FA} = Q\left(\frac{\gamma'}{\sqrt{\sigma^2/N}}\right)$$

$$P_D = \Pr\{T(x) > \gamma' ; \mathcal{H}_1\}$$

$$P_D = Q\left(\frac{\gamma' - A}{\sqrt{\sigma^2/N}}\right)$$

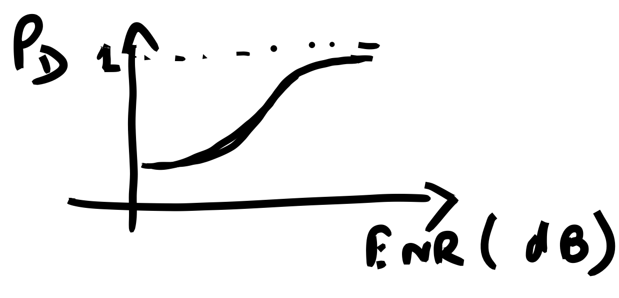
$$\gamma' = \sqrt{\frac{\sigma^2}{N}}\, Q^{-1}(P_{FA})$$

[Q fn is monotonically decreasing]

$$P_D = Q\left(\frac{\sqrt{\sigma^2/N}\, Q^{-1}(P_{FA}) - A}{\sqrt{\sigma^2/N}}\right)$$

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{NA^2}{\sigma^2}}\right)$$

$\frac{NA^2}{\sigma^2}$ = signal energy to noise ratio. (ENR)



Mean-Shifted Gauss Gauss Problem

we decided $\mathcal{H}_1$ if $T(x) > \gamma'$

$$T \sim \begin{cases} \mathcal{N}(\mu_0, \sigma^2) & \text{under } \mathcal{H}_0 \\ \mathcal{N}(\mu_1, \sigma^2) & \text{" } \mathcal{H}_1 \end{cases}$$

$[\mu_1 > \mu_0]$

* $T = \bar{x}$

Deflection Coefficient. (d^2)

$$d^2 = \frac{\left[E[T; \mathcal{H}_1] - E[T; \mathcal{H}_0]\right]^2}{\mathrm{var}(T; \mathcal{H}_0)}$$

$$d^2 = \frac{(\mu_1 - \mu_0)^2}{\sigma^2}$$

$\mu_0 \to 0$: $d^2 \to \mu_1^2/\sigma^2$ (SNR)

$$P_{FA} = \Pr\{T > \gamma' ; \mathcal{H}_0\} = Q\left(\frac{\gamma' - \mu_0}{\sigma}\right)$$

$$\gamma' = \mu_0 + \sigma Q^{-1}(P_{FA})$$

$$P_D = Pr\{T > \gamma' ; \mathcal{H}_1\} = Q\left(\frac{\gamma' - \mu_1}{\sigma}\right)$$

$$= Q\left(\frac{\mu_0 + \sigma Q^{-1}(P_{FA}) - \mu_1}{\sigma}\right)$$

$$= Q\left(Q^{-1}(P_{FA}) - \left(\frac{\mu_1 - \mu_0}{\sigma}\right)\right)$$

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{d^2}\right) \quad [\because \mu_1 > \mu_0]$$

* let $\vec{x} = [x[0] \; x[1] \cdots, x[N-1]]^T$ form a PDF, parameterized by θ. $p(\vec{x}; \theta)$

$$\mathcal{H}_0 : \theta = \theta_0$$

$$\mathcal{H}_1 : \theta = \theta_1$$

<u>NF Theorem</u>: $p(\vec{x}; \theta) = g(T(\vec{x}), \theta)\, h(\vec{x})$

$T(x)$: sufficient statistic for θ.

$$\frac{p(\vec{x}; \theta_1)}{p(\vec{x}; \theta_0)} > \gamma \Rightarrow \frac{g(T(\vec{x}), \theta_1)}{g(T(\vec{x}), \theta_0)} > \gamma.$$

Receiver Operating Characteristic :- (ROC)

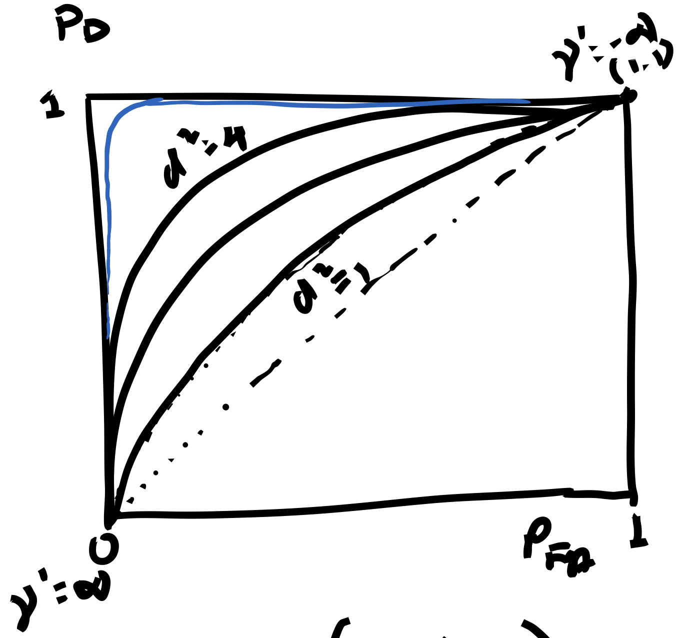
DC-level detection problem

$$P_{FA} = Q\left(\frac{\gamma'}{\sqrt{\sigma^2/N}}\right)$$

$$P_D = Q\left(\frac{\gamma' - A}{\sqrt{\sigma^2/N}}\right)$$

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{d^2}\right)$$

$$d^2 = \frac{NA^2}{\sigma^2}$$



$$P(x \leq x)\Big|_{x=\infty} = 1$$

γ' increases $\longrightarrow$ P_{FA} as well as P_D decrease

This type of performance summary is called ROC.

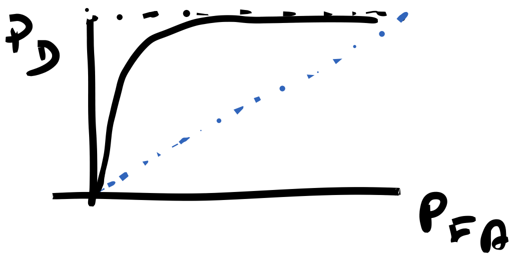
The ROC should always be above $45°$ line.

$d^2 \longrightarrow \infty$: $P_D = 1$ for any P_{FA}

$d^2 \longrightarrow 0$: lower bound is obtained.

Lecture - 29

ROC :-



$P(H) = p$

$P(H \text{ under } \mathcal{H}_1) =$

$P(H \text{ under } \mathcal{H}_0)$

$P_{FA} = P_D$.

Irrelevant Data :-

$$\mathcal{H}_0 : \quad x[n] = w[n] \qquad n = 0, 1, \dots, N-1$$

$$\mathcal{H}_1 : \quad x[n] = A + w[n]$$

$$w_R[n] \qquad n = 0, 1, \dots, N-1$$

$$\{x[0], x[1], \dots, x[N-1], w_R[0], w_R[1], \dots w_R[N-1]\}$$

$$[\vec{x}^T \; \vec{w}_R^T]^T$$

$$\mathcal{H}_0 : \quad x[n] = w[n]$$

$$\mathcal{H}_1 : \quad x[n] = A + w[n] \qquad \text{for } A > 0$$

$$w_R[n] = w[n]$$

Decide $\mathcal{H}_1$ if

$$T = x[0] - w_R[0] > \frac{A}{2}.$$

under $\mathcal{H}_0$: $T = 0$ | Extreme case of
" $\mathcal{H}_1$: $T = A$ | statistical dependence.

* $\vec{w}_R$ is independent of $\vec{x}$ under either hypotheses. $\vec{w}_R$ will be irrelevant.

$$\{x[0], x[1], \dots, x[N-1], x[N], \dots, x[2N-1]\}.$$

$$\vec{x} : [\vec{x}_1^T \ \vec{x}_2^T]^T$$

$$\mathcal{H}_0 : x[n] = w[n] \qquad n = 0, 1, \dots, 2N-1$$

$$\mathcal{H}_1 : x[n] = \begin{cases} A + w[n] & n = 0, \dots, N-1 \\ w[n] & n = N, \dots, 2N-1 \end{cases}$$

$$\frac{p(\vec{x}_1, \vec{x}_2 ; \mathcal{H}_1)}{p(\vec{x}_1, \vec{x}_2 ; \mathcal{H}_0)} > \gamma.$$

$$\frac{\prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}(x[n]-A)^2\right] \prod_{n=N}^{2N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} x^2[n]\right]}{\prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2} x^2[n]\right] \prod_{n=N}^{2N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} x^2[n]\right)}$$

Or $$\frac{p(\vec{x}_1 ; H_1)}{p(\vec{x}_1 ; H_0)} > \gamma$$

So, $\vec{x}_2$ is irrelevant to the detection problem

$$* \quad L(\vec{x}_1, \vec{x}_2) = \frac{p(\vec{x}_1, \vec{x}_2 ; H_1)}{p(\vec{x}_1, \vec{x}_2 ; H_0)} = \frac{p(\vec{x}_2 | \vec{x}_1 ; H_1)\, p(\vec{x}_1 ; H_1)}{p(\vec{x}_2 | \vec{x}_1 ; H_0)\, p(\vec{x}_1 ; H_0)}$$

if $p(\vec{x}_2 | \vec{x}_1 ; H_1) = p(\vec{x}_2 | \vec{x}_1 ; H_0)$

$L(\vec{x}_1, \vec{x}_2) = L(\vec{x}_1)$ and $\vec{x}_2$ is irrelevant to the detection problem.

Minimum Probability of Error:-

$P_e = \Pr\{\text{decide } H_0, H_1 \text{ is true}\} + \Pr\{\text{decide } H_1, H_0 \text{ is true}\}$

$$P_e = P(H_0 | H_1)\, P(H_1) + P(H_1 | H_0)\, P(H_0)$$

$P(H_i | H_j)$: Conditional probability that indicate the probability of deciding H_i when H_j is true.

NP-approach

$P(H_i; H_j)$

Probability of deciding H_i when H_j is true with no probabilistic meaning assigned that H_j is true

Bayesian

$P(H_i | H_j)$

The outcome of a probabilistic experiment is observed to be H_j and the probability of deciding H_i is conditional on that outcome.

Decide H_1 if

$$\frac{P(x|H_1)}{P(x|H_0)} > \frac{P(H_0)}{P(H_1)} = \gamma$$

$$P(x|H_1) > P(x|H_0) \quad \left[\begin{array}{l} \text{if } P(H_0) \\ = P(H_1) \end{array}\right]$$

This is called ML detection.

#. $H_0 : x[n] = w[n]$ $\quad n = 0, 1, \ldots, N-1$

$H_1 : x[n] = A + w[n]$ $\quad n = 0, 1, \ldots, N-1$

$A > 0, \; w[n] \sim \mathcal{N}(0, \sigma^2)$

OOK system.

$P(H_0) = P(H_1) = \frac{1}{2}.$

$$\gamma = \frac{P(\mathcal{H}_0)}{P(\mathcal{H}_1)} = 1.$$

we decide $\mathcal{H}_1$ if

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

Taking logarithm:-

$$-\frac{1}{2\sigma^2}\left(-2A\sum_n x[n] + NA^2\right) > 0$$

$$\frac{A}{\sigma^2}\sum_n x[n] - \frac{NA^2}{2\sigma^2} > 0$$

$$\frac{1}{N}\sum_n x[n] > \frac{A}{2} \Rightarrow \bar{x} > \frac{A}{2}.$$

$$P_e = P(\mathcal{H}_0|\mathcal{H}_1)P(\mathcal{H}_1) + P(\mathcal{H}_1|\mathcal{H}_0)P(\mathcal{H}_0)$$

$$= ?$$

$$\Rightarrow \quad \frac{P(x \mid J{=}1)}{P(x \mid J{=}0)} > \frac{P(J{=}0)}{P(J{=}1)}$$

$$\Rightarrow \quad P(x \mid J{=}1)\,P(J{=}1) > P(x \mid J{=}0)\,P(J{=}0)$$

$$\Rightarrow \quad \frac{P(x \mid J{=}1)\,P(J{=}1)}{P(x)} > \frac{P(x \mid J{=}0)\,P(J{=}0)}{P(x)}$$

$$\Rightarrow \quad P(J{=}1 \mid x) > P(J{=}0 \mid x)$$

MAP - Detection.

Lecture-30

$$P_e = P(H_0 | H_1) P(H_1) + P(H_1 | H_0) P(H_0)$$

Decide H_1 if $$\frac{P(x | H_1)}{P(x | H_0)} > \frac{P(H_0)}{P(H_1)} = \gamma$$

H_0 : Part is defective

H_1 : Part is satisfactory.

C_{ij} : Cost when we decide H_i but H_j is true.

$$C_{10} > C_{01}$$

Bayes Risk $$R = \sum_{i=0}^{1} \sum_{j=0}^{1} C_{ij} \, P(H_i | H_j) \, P(H_j)$$

If $C_{ij} = 1$ when $i \neq j$

$C_{ij} = 0$ when $i = j$

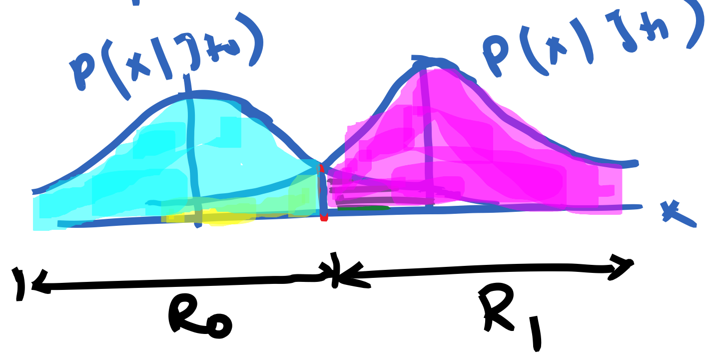
$R = P_e$

$$C_{10} > C_{00} \quad \& \quad C_{01} > C_{11}$$

Decide H_1 if

$$\frac{p(x|H_1)}{p(x|H_0)} > \frac{(C_{10}-C_{00})\,P(H_0)}{(C_{01}-C_{11})\,P(H_1)} = \gamma$$

$$R = \sum_{i=0}^{1}\sum_{j=0}^{1} C_{ij}\, P(H_i|H_j)\, P(H_j)$$



$R_1 = \{x : \text{decide } H_1\}$

$R_0 = \{x : \text{decide } H_0\}$

$$R = C_{00}\,P(H_0)\int_{R_0} p(x|H_0)\,dx + C_{01}\,P(H_1)\int_{R_0} p(x|H_1)\,dx$$
$$+ C_{10}\,P(H_0)\int_{R_1} p(x|H_0)\,dx + C_{11}\,P(H_1)\int_{R_1} p(x|H_1)\,dx$$

$$\boxed{\int_{R_0} p(x|H_i)\,dx = 1 - \int_{R_1} p(x|H_i)\,dx}$$

$$R = C_{01}\,P(H_1) + C_{00}\,P(H_0) + \int_{R_1}\Big[\{C_{10}\,P(H_0) - C_{00}\,P(H_0)\}\,p(x|H_0)$$
$$- \big(-C_{11}\,P(H_1) + C_{01}\,P(H_1)\big)\,p(x|H_1)\Big]\,dx$$

Include x in R_1 if the integration is negative.

$$-(C_{10}-C_{00})P(H_0)p(x|H_0) < (C_{01}-C_{11})P(H_1)p(x|H_1)$$

$$C_{10} > C_{00}, \quad C_{01} > C_{11}$$

$$\frac{p(x|H_1)}{p(x|H_0)} > \frac{(C_{10}-C_{00})P(H_0)}{(C_{01}-C_{11})P(H_1)} = \gamma.$$

Multiple Hypothesis Testing:-

M

$$\{H_0, H_1, H_2, \ldots, H_m\}$$

$$R = \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} C_{ij}\, P(H_i|H_j)\, P(H_j)$$

$$C_{ij} = \begin{cases} 0 & \text{when } i=j \\ 1 & i \neq j \end{cases}$$

$R = P_e.$

$$R = \sum_{i=0}^{m-1}\sum_{j=0}^{m-1} C_{ij} \int_{R_i} p(x|H_j)\, P(H_j)\, dx$$

$$= \sum_i \int_{R_i} \sum_j c_{ij}\, P(x \mid H_j)\, P(H_j)\, dx$$

$$= \sum_i \int_{R_i} \underbrace{\sum_j c_{ij}\, P(H_j \mid x)}_{c_i(x)}\, P(x)\, dx$$

$$R = \sum_i \int_{R_i} c_i(x)\, P(x)\, dx.$$

$$c_1(x)\, P(x)\, dx$$

$$c_0(x)\, P(x)\, dx$$

We need to minimize $c_i(x)$.

$$c_i(x) = \sum_{j=0}^{m-1} c_{ij}\, P(H_j \mid x)$$

over $i = 0, 1, \ldots, m-1$

$$c_i(x) = \sum_{\substack{j=0 \\ j \neq i}}^{m-1} P(H_j \mid x)$$

$$c_i(x) = \sum_{j=0}^{m-1} P(H_j \mid x) - P(H_i \mid x)$$

We shall maximize $P(H_i \mid x)$ to minimize $c_i(x)$.

choose H_k if $P(H_k|x) > P(H_i|x)$ $i \neq k$.

$$P(H_i|x) = \frac{P(x|H_i)\,P(H_i)}{P(x)}$$

$$= \frac{P(x|H_i)\,\frac{1}{M}}{P(x)}$$

To maximize $P(H_i|x)$, maximize $P(x|H_i)$

$$P(x|H_k) > P(x|H_i) \quad i \neq k.$$

m-ary ML detector.

If not uniform distribution for prior knowledge

$$P(H_k|x) > P(H_i|x) \quad i \neq k$$

m-ary MAP detector.

Composite Hypothesis Testing:-

$$P(x; A, H_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2\right]$$

$\mathcal{H}_0: x[n] = w[n] \qquad n = 0, 1, \ldots, N-1$

$\mathcal{H}_1: x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1$

A is unknown $(A > 0)$

NP-test

$$\frac{p(x; A, \mathcal{H}_1)}{p(x; \mathcal{H}_0)} = \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2\right]}{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_n x^2[n]\right]}$$

Decide $\mathcal{H}_1$ if $\quad \dfrac{p(x; A, \mathcal{H}_1)}{p(x; \mathcal{H}_0)} > \gamma.$

$$-\frac{1}{2\sigma^2}\left(-2A\sum_n x[n] + NA^2\right) > \ln\gamma.$$

$$\Rightarrow A\sum_n x[n] > \sigma^2 \ln\gamma + \frac{NA^2}{2}$$

$$\Rightarrow \frac{1}{N}\sum_n x[n] > \frac{\sigma^2}{NA}\ln\gamma + \frac{A}{2} = \gamma'$$

doesn't depend on A

$$T(x) > \gamma$$

$$T(x) = \bar{x} \sim \mathcal{N}(0, \sigma^2/N)$$

$$P_{FA} = \Pr\{T(x) > \gamma'; \mathcal{H}_0\}$$

$$P_{FA} = Q\left(\frac{\gamma'}{\sqrt{\sigma^2/N}}\right) \Rightarrow \gamma' = \sqrt{\frac{\sigma^2}{N}}\, Q^{-1}(P_{FA})$$

doesn't depend on A

Lecture - 31

Composite Hypothesis Testing:

$\mathcal{H}_0: \; x[n] = w[n] \qquad n = 0, 1, \ldots, N-1$

$\mathcal{H}_1: \; x[n] = A + w[n] \qquad n = 0, 1, \ldots, N-1$

A is unknown but $A > 0$.

$w[n] \sim \mathcal{N}(0, \sigma^2)$

We will decide $\mathcal{H}_1$ if

$$T(x) = \bar{x} > \gamma' = \sqrt{\frac{\sigma^2}{N}}\, Q^{-1}(P_{FA})$$

$$P_D = \Pr\{T(x) > \gamma'; \mathcal{H}_1\} \qquad \left[\begin{array}{l} T(x) \sim \mathcal{N}\left(0, \frac{\sigma^2}{N}\right) \text{ under } \mathcal{H}_0 \\ T(x) \sim \mathcal{N}\left(A, \frac{\sigma^2}{N}\right) \text{ under } \mathcal{H}_1 \end{array}\right.$$

$$= Q\left(\frac{\gamma' - A}{\sqrt{\sigma^2/N}}\right)$$

$$P_D = Q\left(Q^{-1}(P_{FA}) - \sqrt{\frac{NA^2}{\sigma^2}}\right)$$

We may say that over all possible detectors that have a given P_{FA} the one that decides $\mathcal{H}_1$ if $\bar{x} > \sqrt{\frac{\sigma^2}{N}}\, Q^{-1}(P_{FA})$ yields the highest P_D for any value of A ($A > 0$). This type of test is known as Uniformly Most

Powerful (UMP) Test.

* If $-a < A < a$, UMP test does not exist.

If $A < 0$, we decide H_1 if

$$\bar{x} < -\sqrt{\frac{\sigma^2}{N}}\, Q^{-1}(P_{FA})$$

The hypothesis testing problem:

$$\left.\begin{array}{l} H_0: \quad A = 0 \\ H_1: \quad A > 0. \end{array}\right\} \text{One-sided test}$$

$$\left.\begin{array}{l} H_0: \quad A = 0 \\ H_1: \quad A \neq 0 \end{array}\right\} \text{Two-sided test.}$$

* UMP test must be One-sided.

<u>Composite Hypothesis Testing approaches</u>

1. Bayesian approach.
2. Generalized Likelihood Ratio Test (GLRT)

① <u>Bayesian Approach:-</u>

- Requires prior knowledge about the parameter.
- " multi-dimensional integration

$$\vec{\theta}_0 \qquad \vec{\theta}_1$$

need to know :- $p(\vec{\theta}_0) \quad p(\vec{\theta}_1)$

Independent of θ_0 & θ_1

$$p(\vec{x}; \mathcal{H}_0) = \int p(\vec{x} \mid \vec{\theta}_0; \mathcal{H}_0)\, p(\vec{\theta}_0)\, d\vec{\theta}_0$$

$$p(\vec{x}; \mathcal{H}_1) = \int p(\vec{x} \mid \vec{\theta}_1; \mathcal{H}_1)\, p(\vec{\theta}_1)\, d\vec{\theta}_1$$

$p(\vec{x} \mid \vec{\theta}_i; \mathcal{H}_i)$: Conditional PDF.

The NP-detector decides $\mathcal{H}_1$ if

$$\frac{p(\vec{x}; \mathcal{H}_1)}{p(\vec{x}; \mathcal{H}_0)} = \frac{\int p(\vec{x} \mid \vec{\theta}_1; \mathcal{H}_1)\, p(\vec{\theta}_1)\, d\vec{\theta}_1}{\int p(\vec{x} \mid \vec{\theta}_0; \mathcal{H}_0)\, p(\vec{\theta}_0)\, d\vec{\theta}_0} > \gamma$$

<u>Choice of Prior PDF:-</u>

→ Given → use it

→ Not given → Assume depending on the application.

(i) Detecting phase angle of a Sinusoid, $\phi \sim \mathcal{U}[0, 2\pi]$

(ii) DC level detection $(-\infty < \theta < \infty)$ Gaussian PDF $A \sim \mathcal{N}(0, \sigma_\theta^2)$

$\sigma_A^2 \to \infty$: Lack of ${}_{\wedge}^{\text{prior}}$ knowledge

* $\mathcal{H}_0$: $x[n] = w[n]$ $\quad n = 0, 1, \ldots, N-1$

$\mathcal{H}_1$: $x[n] = A + w[n]$ $\quad n = 0, 1, \ldots, N-1$

A is unknown and $-\infty < A < \infty$

$w[n] \sim \mathcal{N}(0, \sigma^2)$

Choose $A \sim \mathcal{N}(0, \sigma_A^2)$ [A is independent of $w[n]$]

$$p(\vec{x} \mid A; \mathcal{H}_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2\right]$$

$$\frac{p(\vec{x}; \mathcal{H}_1)}{p(\vec{x}; \mathcal{H}_0)} = \frac{\int_{-\infty}^{\infty} p(\vec{x} \mid A; \mathcal{H}_1)\, p(A)\, dA}{p(x; \mathcal{H}_0)} > \gamma \quad \left[\begin{matrix}\text{Decide} \\ \mathcal{H}_1\end{matrix}\right]$$

$$p(\vec{x}; \mathcal{H}_1) = \int_{-\infty}^{\infty} p(\vec{x} \mid A; \mathcal{H}_1)\, p(A)\, dA$$

$$= \int_{-\infty}^{\infty} \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_{n=0}^{N-1} (x[n] - A)^2\right] \frac{1}{\sqrt{2\pi\sigma_A^2}} \exp\left[-\frac{A^2}{2\sigma_A^2}\right] dA$$

Let $Q(A) = \frac{1}{\sigma^2} \sum_n x^2[n] - \frac{2N}{\sigma^2} \bar{x} A + \frac{N}{\sigma^2} A^2 + \frac{A^2}{\sigma_A^2}$

$$= \left(\frac{N}{\sigma^2} + \frac{1}{\sigma_A^2}\right) A^2 - \frac{2N}{\sigma^2} \bar{x} A + \frac{1}{\sigma^2} \sum_{n=0}^{N-1} x^2[n]$$

$$= \frac{A^2}{\sigma^2_{A|x}} - \frac{2N \sigma^2_{A|x} \bar{x} A}{\sigma^2 \sigma^2_{A|x}} + \frac{1}{\sigma^2} \sum_n x^2[n]$$

$$Q(A) = \frac{1}{\sigma^2_{A|x}} \left(A - \frac{N \bar{x} \sigma^2_{A|x}}{\sigma^2} \right) - \frac{N^2 \bar{x}^2}{\sigma^4} \sigma^2_{A|x} + \frac{1}{\sigma^2} \sum_n x^2[n]$$

$$\frac{p(x; \mathcal{H}_1)}{p(x; \mathcal{H}_0)} = \frac{\frac{1}{(2\pi\sigma^2)^{N/2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2_A}} \int_{-\infty}^{A} \exp\left[-\frac{1}{2} Q(A)\right] dA}{\frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_n x^2[n]\right)}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2_A}} \sqrt{2\pi\sigma^2_{A|x}} \exp\left(\frac{N^2 \bar{x}^2 \sigma^2_{A|x}}{2\sigma^4}\right) > \gamma$$

$\bar{x}^2 > \gamma'$ [After taking logarithm]

$|\bar{x}| > \sqrt{\gamma'}$ decide $\mathcal{H}_1$

(ii) Generalized Likelihood Ratio Test :-

Replace θ with $\hat{\theta}_{ML}$ (ML estimation of θ).

GLRT decides $\mathcal{H}_1$ if

$$L_G(\vec{x}) = \frac{p(\vec{x}; \hat{\theta}_1, \mathcal{H}_1)}{p(\vec{x}; \hat{\theta}_0, \mathcal{H}_0)} > \gamma$$

$\hat{\theta}_i$ is the MLE of θ_i assuming H_i is true.

Lecture - 32

Hypothesis Testing

Simple
(PDFs under the hypotheses are known)

NP
- Classical detector
- Optimizes P_D for P_{FA}
- Does not require prior info
- It can be applied for all kinds of problem.

Bayesian
- Considers prior information
- Assigns costs of each decision
- Minimizes the avg. cost of decision making.

Composite
(Do not have complete knowledge about PDF)

Bayesian
- The LRT is performed by integrating out the unknown parameters using prior PDFs.
- Closed form integration may not be possible.
- Optimal

GLRT
- Most commonly used
- Unknown parameters are replaced with MLE
- Optimality is not guaranteed
- Practically useful

GLRT:-

$$L_G(\bar{x}) = \frac{p(x;\hat{\theta}_1,\mathcal{H}_1)}{p(x;\hat{\theta}_0,\mathcal{H}_0)} > \gamma_1 \Rightarrow \mathcal{H}_1$$

$\hat{\theta}_i$ = MLE of θ.

#1. $\mathcal{H}_0$: $A = 0$ (DC level detection problem)

$\mathcal{H}_1$: $A \neq 0$.

$$L_G(\vec{x}) = \frac{p(\vec{x};\hat{A},\mathcal{H}_1)}{p(\vec{x};\mathcal{H}_0)} > \gamma.$$

$$p(x;A,\mathcal{H}_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A)^2\right]$$

$$\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1} x[n] = \bar{x}$$

$$L_G(x) = \frac{\frac{1}{(2\pi\sigma^2)^{N/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_n (x[n]-\bar{x})^2\right]}{\frac{1}{(2\pi\sigma^2)^{N/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_n x^2[n]\right]}$$

$$\ln L_G(x) = -\frac{1}{2\sigma^2}\left(-2N\bar{x}^2 + N\bar{x}^2\right) = \frac{N\bar{x}^2}{2\sigma^2}.$$

Decide for $\mathcal{H}_1$ if $\frac{N\bar{x}^2}{2\sigma^2} > \gamma$

$$|\bar{x}| > \gamma'$$

$$L_G(x) = \frac{\max_{\theta_1} p(x; \theta_1, H_1)}{\max_{\theta_0} p(x; \theta_0, H_0)}$$

Non-Parametric Detection:-

CFAR detector.

Constant False Alarm Rate (CFAR).

(i) Sign Detector

(ii) Semiversial Detector.

Sign Detector:-

$$\frac{1}{N}\sum_n x[n] > \gamma \quad \text{Decide } H_1$$

$$\sum_n x[n] > \gamma' \quad \text{"} \quad H_1$$

$$\sum_{n=0}^{N-1} u(x[n]) > \gamma_u.$$

$u(\cdot)$ unit step fn.

$$u(x[n]) = \begin{cases} 1 & x[n] > 0 \\ 0 & x[n] < 0 \end{cases}$$

Performance analysis:-

$$\mathcal{H}_0: \Pr\{x[n] \geq 0\} = 0.5 \qquad n = 0, 1, \ldots, N-1$$

$$\mathcal{H}_1: \Pr\{x[n] \geq 0\} = p > 0.5 \qquad n = 0, 1, \ldots, N-1$$

$d[n]$ denotes sign of $x[n]$

$$d[n] = \begin{cases} 1 & ; \; x[n] > 0 \\ 0 & ; \; x[n] < 0 \end{cases}$$

$$\Pr\{d[n] \mid \mathcal{H}_0\} = \begin{cases} 0.5 & d[n] = 0 \; [x[n] < 0] \\ 0.5 & d[n] = 1 \; [x[n] > 0]. \end{cases}$$

$$\Pr\{d[n] \mid \mathcal{H}_1\} = \begin{cases} 1-p & d[n] = 0 \; [x[n] < 0] \\ p & d[n] = 1 \; [x[n] > 0]. \end{cases}$$

$L(d) > \gamma \rightarrow$ Decide $\mathcal{H}_1$

$$L(d) = \frac{p(x; \mathcal{H}_1)}{p(x; \mathcal{H}_0)} = \frac{p^{\sum_n d[n]} (1-p)^{N - \sum_n d[n]}}{(0.5)^N}$$

$\sum_n d[n] = N^+$ denotes the no. of +ve observations

$$L(d) = \frac{p^{N^+}(1-p)^{N-N^+}}{(0.5)^N} = [2(1-p)]^N \left(\frac{p}{1-p}\right)^{N^+} > \gamma$$

$$N^+ > \log_{p/1-p}\gamma - N\log_{p/1-p}[2(1-p)] = \gamma'$$

$$Pr(N^+ = n \mid H_0) = \binom{N}{n}(0.5)^n(1-0.5)^{N-n}$$

$$= \binom{N}{n}(0.5)^N$$

$$P_{FA} = \sum_{n=\gamma'+1}^{N}\binom{N}{n}(0.5)^N$$

$$Pr(N^+ = n \mid H_1) = \binom{N}{n}p^n(1-p)^{N-n}$$

$$P_D = \sum_{n=\gamma'+1}^{N}\binom{N}{n}p^n(1-p)^{N-n}$$

#. Derive a sign detector that uses nine observations and ensures a probability of false alarm of 0.1 for detecting a positive signal A in presence of 0 mean Gaussian noise and analyze its performance for probability of data being positive (i) 0.75 and (ii) 0.99.

Lecture - 33

Sign Detector:-

$$L(x) = \frac{p^{N^+}(1-p)^{N-N^+}}{(0.5)^N} > \gamma \quad \text{Decide } \mathcal{H}_1$$

N^+ : no. of positive observations

$N^+ > \gamma'$: Decide $\mathcal{H}_1$

$$\boxed{\gamma' = \log_{p/1-p}\gamma - N\log_{p/1-p}[2(1-p)]}$$

$$P_{FA} = \sum_{n=\gamma'+1}^{N}\binom{N}{n}(0.5)^N$$

$$P_D = \sum_{n=\gamma'+1}^{N}\binom{N}{n}p^n(1-p)^{N-n}$$

* Derive a sign detector that uses nine observations and ensures a P_{FA} of 0.1 for detecting a positive signal A in presence of 0 mean Gaussian noise and analyze its performance for probability of data being positive (i) 0.75 (ii) 0.99.

Ans:- $N = 9$

$\mathcal{H}_0 : x[n] = w[n] \qquad n = 1, \cdots, 9$

$\mathcal{H}_1 : x[n] = A + w[n] \qquad n = 1, \cdots, 9$

$P_{FA} = 0.1$

$$P_{FA} = \sum_{n=\gamma'+1}^{N} \binom{9}{n} (0.5)^9 \leq 0.1$$

$$\sum_{n=\gamma'+1}^{N} \binom{9}{n} \leq 51.2$$

$\binom{9}{9} = 1$, $\binom{9}{8} = 9$, $\binom{9}{7} = 36$, $\binom{9}{6} = 84$

$\gamma' + 1 = 7$

$\gamma' = 6.$

H_1 : $N^+ = \sum_{n=1}^{9} u(x[n]) > \gamma = 6$

$$P_{FA} = \sum_{n=7}^{9} \binom{9}{n} (0.5)^n$$

$$= \frac{1+9+36}{512} = 0.089 < 0.1$$

$$P_D = \sum_{n=7}^{9} \binom{9}{n} p^n (1-p)^{9-n}$$

(i) For $p = 0.75$

$$P_D = \sum_{n=7}^{9} \binom{9}{n} (0.75)^n (0.25)^{9-n}$$

$$= \frac{9 \times 8}{2} (0.75)^7 (0.25)^2 + 9 (0.75)^8 (0.25)^1 + (0.25)^9$$

$$P_D = 0.601$$

(ii) P_D when $p = 0.99$.

$$P_D = 0.923 \,?$$

Sequential Detection

$L(x) > \eta_1$ Decide H_1

$L(x) < \eta_0$ Decide H_0

Determine η_1 from $P_{FA} = \alpha$

" η_0 from $P_m = \beta$.

$$x_n : \{ x_1, x_2, \dots, x_N \}$$

$$\Lambda(x_N) = \frac{P_1(x_N)}{P_0(x_N)} = \prod_{n=1}^{N} \frac{P_1(x_n)}{P_0(x_n)} = \prod_{n=1}^{N-1} \frac{P_1(x_n)}{P_0(x_n)} \cdot \frac{P_1(x_N)}{P_0(x_N)}$$

$$\Lambda(x_N) = \Lambda(x_{N-1}) \, \Lambda(x_N)$$

$$\alpha = P_{FA} = \int_{R_1} \cdots \int P_0(x_n)\, dx_1 \cdots dx_n \quad [n=1,\dots,N]$$

$$\beta = P_m = 1 - P_D \Rightarrow \quad P_D = 1 - \beta.$$

$$P_D = \int_{R_1} \cdots \int P_1(x_n)\, dx_1 \cdots dx_n \quad [n=1,\dots,n]$$

$$= \int_{R_1} \cdots \int \frac{P_1(x_n)}{P_0(x_n)} \cdot P_0(x_n)\, dx_1 \cdots dx_n.$$

$$\Lambda(x_N) = \Lambda(x_{N-1})\,\Lambda(x_{N-2}) \cdots \Lambda(x_1) \geq \eta_1$$

Decide H_1

$$\Lambda(x_n) \geq \eta_1$$

$$\Lambda(x_N)\, P_0(x_n) \geq \eta_1\, P_0(x_n)$$

$$\int_{R_1} \cdots \int \Lambda(x_N)\, P_0(x_n)\, dx_1 \cdots dx_n \geq \eta_1 \int_{R_1} \cdots \int P_0(x_n)\, dx_1 \cdots dx_n$$

$$P_D \geq \eta_1 P_{FA} = \eta_1 \alpha.$$

Since $P_D = 1 - P_m = 1 - \beta.$

$$1 - \beta \geq \eta_1 \alpha.$$

$$\eta_1 \leq \frac{1-\beta}{\alpha}.$$

Similarly $P_m = \int \cdots \int_{R_0} P_1(x_n)\, dx_1 \cdots dx_n$

$$\Lambda(x_N) \leq \eta_0$$

$$\beta \leq \eta_0 \int \cdots \int_{R_0} P_0(x_n)\, dx_n$$

$$\beta \leq \eta_0 (1-\alpha)$$

$$\eta_0 \geq \frac{\beta}{1-\alpha}.$$

We decide H_1 if $\Lambda(x_N) > \eta_1$

H_0 if $\Lambda(x_N) < \eta_0$

If $\Lambda(x_N) > \eta_0$ but $\Lambda(x_N) < \eta_1$ then another sample is taken to form $\Lambda(x_{N+1})$ and the test is continued.

<u>Advantage</u>:- You can terminate the test earlier than the conventional NP-test once the presence or absence of a target has been determined with an acceptable level of error (P_{FA} or P_m).

<u>Disadvantages</u>:- (i) Samples are assumed to be IID.

(ii) P_{FA} & P_m ... constant

Consider the detection of a DC level A in AWGN with 0 mean and σ^2 variance. Conduct a sequential LRT to detect the presence and absence of the signal. It is desired to terminate the test when $P_{FA} \leq \alpha$ or $P_m \leq \beta$.

$$\mathcal{H}_0 : x[n] = w[n] \qquad n = 1, \ldots, N.$$

$$\mathcal{H}_1 : x[n] = A + w[n] \qquad n = 1, \ldots, N.$$

$$\eta_1 \leq \frac{1 - P_m}{P_{FA}} \quad ; \quad \eta_0 \geq \frac{P_m}{1 - P_{FA}} = \frac{\beta}{1-\alpha}.$$

$$= \frac{1-\beta}{\alpha}$$

$$\Lambda(x_N) = \exp\left(\frac{\sum_{n=1}^{N} x_n^2 - \sum_{n=1}^{N} (x_n - A)^2}{2\sigma^2} \right) \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \begin{matrix} \eta_1 \\ \eta_0 \end{matrix}$$

$$z = \sum_{n=1}^{N} x_n \quad \begin{matrix} \overset{\mathcal{H}_1}{>} \dfrac{\sigma^2 \ln(\eta_1)}{A} + \dfrac{AN}{2}. \\ \underset{\mathcal{H}_0}{<} \dfrac{\sigma^2 \ln(\eta_0)}{A} + \dfrac{AN}{2} \end{matrix}$$

$n = 1, \cdots, 7$

$n = 8 \rightarrow \mathcal{H}_1$

Lecture - 34

Detection of Deterministic Signals in WGN:-

Replica - Correlator Detection:

$$\mathcal{H}_0: \quad x[n] = w[n] \qquad n = 0, \ldots, N-1$$

$$\mathcal{H}_1: \quad x[n] = s[n] + w[n] \qquad n = 0, \ldots, N-1$$

$s[n]$: Deterministic known signal.

$w[n] \sim \mathcal{N}(0, \sigma^2)$

Decide $\mathcal{H}_1$ if $\quad L(x) = \dfrac{p(\vec{x}; \mathcal{H}_1)}{p(\vec{x}; \mathcal{H}_0)} > \gamma$

$$\vec{x} = [x[0], x[1], \ldots, x[N-1]]^T.$$

$$p(x; \mathcal{H}_0) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_n x^2[n]\right]$$

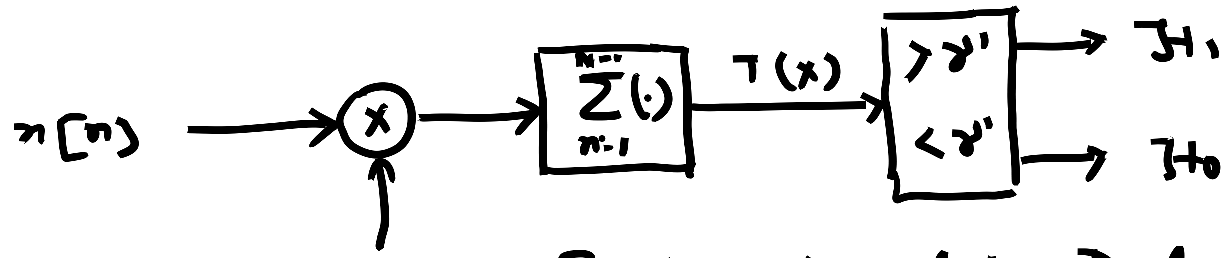
$$p(x; \mathcal{H}_1) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left[-\frac{1}{2\sigma^2} \sum_n (x[n] - s[n])^2\right]$$

$$l(x) = \ln L(x) = -\frac{1}{2\sigma^2}\left(\sum_{n=0}^{N-1} (x[n] - s[n])^2 - \sum_n x^2[n]\right) > \ln\gamma.$$

$$T(x) = \sum_{n=0}^{N-1} x[n]\, s[n] > \sigma^2 \ln\gamma + \frac{1}{2}\sum_n s^2[n] = \gamma'$$

$$\sum_{n=0}^{N-1} x[n]\, s[n] > \gamma' \quad : \quad \mathcal{H}_1$$

we can find γ' from $P_{FA} = \alpha$



<u>Replica-Correlator Detector</u> (RCD)

<u>Matched Filter Detector:-</u>

RCD ≡ FIR filtering on the data.

$h[n] \neq 0$ for $n = 0, 1, \cdots, N-1$

$$y[n] = \sum_{k=0}^{N-1} h[n-k]\, x[k]$$

$$s[n] \xrightarrow{\text{Flip}} s[-n] \xrightarrow{\text{Shift}} s[N-1-n]$$

If $\quad h[n] = s[N-1-n]$

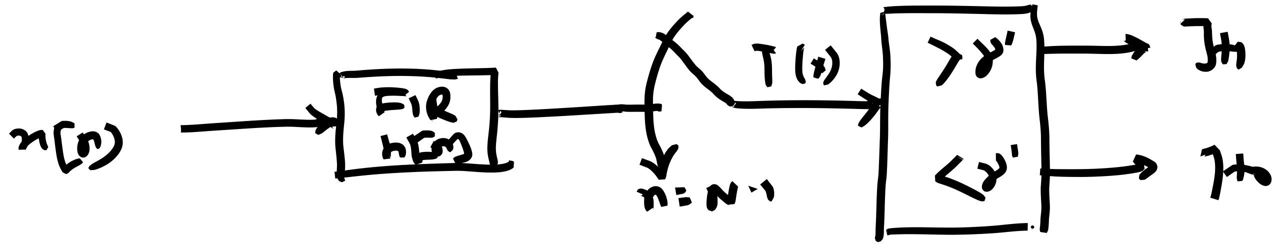
$n = 0, 1, \cdots, N-1$

$$y[n] = \sum_{k=0}^{N-1} s[N-1-(n-k)]\, x[k]$$

Output of the FIR filter at time $n = N-1$

1 $y[N-1] = \sum_{k=0}^{N-1} s[k]\, x[k]$

There is a change of summation variable, Otherwise it is identical to Replica-Correlator.



The filter's impulse response is matched to the signal. The detector is known as Matched Filter detector.

Frequency domain interpretation of Matched Filter:-

$$y[n] = \sum_{k=0}^{N-1} h[n-k]\, x[k]$$

$\omega \rightarrow \pi$

$2\pi f$

$f \rightarrow \frac{1}{2}$

$$Y(f) = X(f)\, H(f)$$

$$y[n] = \int_{-\frac{1}{2}}^{\frac{1}{2}} Y(f) \exp(j2\pi f n)\, df$$

$$= \int_{-1/2}^{1/2} X(f) H(f) \exp(j2\pi f n)\, df$$

$$H(f) = \mathcal{F}\{ s[N-1-n] \}$$

$$\left[\begin{array}{l} s[n] \longleftrightarrow S(f) \\ s[-n] \longleftrightarrow S^*(f) \end{array} \right.$$

$$H(f) = S^*(f) \exp[-j2\pi f (N-1)]$$

$$y[n] = \int_{-1/2}^{1/2} S^*(f) X(f) \exp[j2\pi f (n-(N-1))] \, df$$

Sample the output at $n = N-1$

$$y[N-1] = \int_{-1/2}^{1/2} S^*(f) X(f) \, df.$$

[correlator implementation using Parseval's Theorem].

Matched filter as maximizer of SNR.

$h[n]$ for over $n = 0, \cdots, N-1$

Output SNR $\eta = \dfrac{E^2(y[N-1]; H_1)}{\mathrm{Var}(y[N-1]; H_1)}$

$$\eta = \frac{\sum_{k=0}^{N-1} h[N-1-k] \, s[k]}{E\left[\left(\sum_k h[N-1-k] \, w[k]\right)^2\right]}$$

?

$h[n] = s[N-1-n] \Rightarrow$ maximize η.

$\vec{s} = [s[0] \; s[1] \; \cdots \; s[N-1]]^T$

$\vec{h} = [h[N-1] \; h[N-2], \cdots, h[0]]^T$

$\vec{w} = [w[0] \; w[1] \; \cdots \; w[N-1]]^T$

$$\eta = \frac{(\vec{h}^T \vec{s})^2}{E\left[(\vec{h}^T \tilde{w})^2\right]} = \frac{(h^T s)^2}{h^T E(w w^T) h} = \frac{(h^T s)^2}{h^T \sigma^2 I h}$$

$$= \frac{1}{\sigma^2} \frac{(h^T s)^2}{h^T h}.$$

CS inequality $\quad (h^T s)^2 \le (h^T h)(s^T s)$

[$h = cs$ For equality]

$$\frac{1}{\sigma^2} \frac{(h^T s)^2}{h^T h} \le \frac{s^T s}{\sigma^2}$$

$$\eta \le \frac{s^T s}{\sigma^2}$$

If $c = 1$

$$h = s$$

$$h[N-1-n] = s[n]$$

or $\quad h[n] = s[N-1-n]$

matched filter.