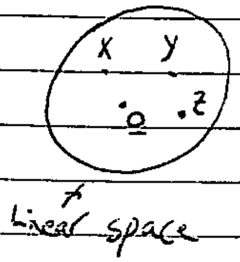
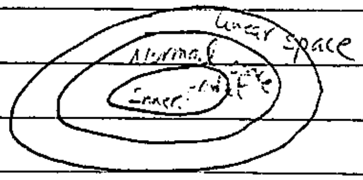
Notes by Pezman Mohammadi

EE 503 (Fall 2010-11)   __/__/__   ①

① Inner Products:





Finite Dimensions: (N. Dim) & R^N

For complex value.

$x, y \in R^N$, $\langle \underline{x}, \underline{y} \rangle = (\underline{x}, \underline{y}) = \underline{x}^T \underline{y}$ or $(\underline{x}^H \underline{y})$

$= \sum_{k=1}^{N} x_k y_k$ SUm(x.*y)

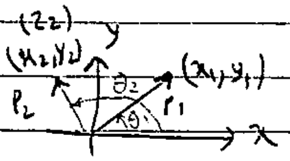
MATLAB

$\|\underline{x}\|^2 = x^T . x$ نرم

$(\|x\| > 0$ if $\|x\| = 0 \longleftrightarrow \underline{x} = \underline{0}$

$\|\underline{x} + y\| \leq \|x\| + \|y\|$; $\|\alpha \underline{x}\| = |\alpha| . |\underline{x}|$

Ex: 2-D



$z_1 = x_1 + j y_1$ $\|z_1\|^2 = \underline{z}_1^T . z_1 = x_1^2 + y_1^2 \Rightarrow$ distanse to origin

$z_2 = x_2 + j y_2$ Caushy _ Schuartz unequality:

$z_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$ $\langle \underline{x}, \underline{y} \rangle \leq \|\underline{x}\| \|\underline{y}\|$

$\langle \underline{z}_1, \underline{z}_2 \rangle = x_1 x_2 + y_1 y_2 = (\|z_1\| \cos\theta_1)(\|z_2\| \cos\theta_2) + (\|z_1\| \sin\theta_1)(\|z_2\| \sin\theta_2)$

$= \|z_1\| \|z_2\| \cos(\theta_1 - \theta_2)$

$\Rightarrow \dfrac{\langle z_1, z_2 \rangle}{\|z_1\| \|z_2\|} = \cos(\theta_1 - \theta_2)$

$z_1 \perp z_2 \Rightarrow \langle z_1, z_2 \rangle = 0 \Rightarrow \theta_1 - \theta_2 = 90^\circ$

Ex. r.v $E\{\underset{\sim}{x}\underset{\sim}{y}\} = 0$ $\underset{\sim}{x} \perp \underset{\sim}{y}$ → orthogonal

$E\{(x - \bar{x})(y - \bar{y})\} = 0$ → Uncorrelated

MATLAB

Ex. Gram - Schmidt Orthogonalization (orth.m)
(orthonormalization)

$\underline{\underline{A}} = [\underline{a}_1 \ \underline{a}_2 \ \ldots \ \underline{a}_N]$

$\underline{e}_1, \ldots, \underline{e}_N$ $\{\underline{e}_1, \ldots, \underline{e}_N\}$ and $\{\underline{a}_1, \ldots, \underline{a}_N\}$

span the same space

$a_k \perp a_\ell \quad k \neq \ell$

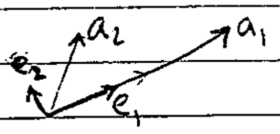
$\|e_k\|^2 = 1 \quad \forall k$

steps:

① $\hat{e}_1 = \underline{a}_1$

$e_1 = \cancel{\frac{a_1}{\|a_1\|}} \ \frac{\hat{e}_1}{\|\hat{e}_1\|}$; $\|e_1\| = 1$

② $\hat{e}_2 = \underline{a}_2 - \underbrace{\langle a_2, e_1 \rangle}_{\|a_2\| \|e_1\| \mathrm{Cs}(\theta_{a_2} - \theta_{e_1})} e_1$



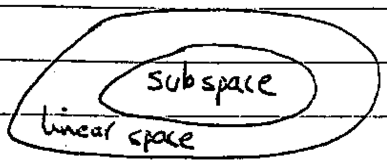
1

$e_2 = \frac{\hat{e}_2}{\|\hat{e}_2\|}$

③ $\hat{e}_3 = a_3 - \langle a_3, e_2 \rangle e_2 - \langle a_3, e_1 \rangle e_1$

$\hat{e}_3 = \frac{\hat{e}_3}{\|\hat{e}_3\|}$

Projection Matrices: (Projection operators)



subspace satisfies all properties of linear space but it is embeded in a larger space.

Projection matrices $\hat{mp}$ element of linear space $\xrightarrow[\text{to}]{}$ subspace

$\underline{\underline{P}}$: Projection matrices $\longleftrightarrow$ ① $\underline{\underline{P}}^2 = \underline{\underline{P}}$

② $\underline{\underline{P}}^T = \underline{\underline{P}}$ (Symmetric)

$$\underset{\substack{\uparrow \\ \text{element of} \\ \text{smaller space}}}{\underline{s}} = \underline{\underline{P}}\, \underline{z} \rightarrow \text{biger space}$$

$$\underline{s} = \underline{\underline{P}} (\underbrace{\underline{\underline{P}}\,\underline{z}}_{\underline{s}}) = \underline{\underline{P}}^2 \underline{z} \qquad ; \underline{\underline{P}}^2 = \underline{\underline{P}}$$

$\underline{z} \longrightarrow \underline{s}$ e.g $\begin{bmatrix}1\\0\\\vdots\end{bmatrix}$, $\begin{bmatrix}1\\\vdots\\1\end{bmatrix}$

Let's say that sub-space $\{\underline{a}_1, \underline{a}_2\}$

step 1: Find an Orthonormal basis for $\{\underline{a}_1, \underline{a}_2\}$

$$\underline{e}_1 = \begin{bmatrix}1\\1\\0\\\vdots\\0\end{bmatrix}\frac{1}{\sqrt{2}}$$

$$\hat{e}_2 = \begin{bmatrix}1\\0\\\vdots\\1\end{bmatrix} - \Big(\underbrace{\begin{bmatrix}1 & 0 & 0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\1\\0\\\vdots\\0\end{bmatrix}\frac{1}{\sqrt{2}}}_{\langle a_2, e_1\rangle}\Big)\overset{e_1}{\begin{bmatrix}1\\1\\0\\\vdots\\0\end{bmatrix}\frac{1}{\sqrt{2}}} = \begin{bmatrix}1\\0\\0\\0\\1\end{bmatrix} - \frac{1}{\sqrt{2}}\begin{bmatrix}1\\1\\0\\0\\0\end{bmatrix} = \begin{bmatrix}\frac{1}{2}\\-\frac{1}{2}\\0\\0\\1\end{bmatrix}$$

$$e_2 = \frac{\hat{e}_2}{\|\hat{e}_2\|} = \frac{\hat{e}_2}{\sqrt{1/2+1}} = \frac{\sqrt{2}}{\sqrt{3}}\hat{e}_2$$

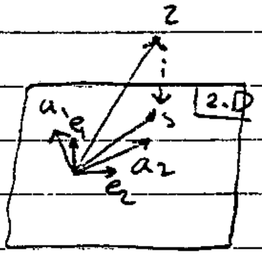
step ②: $\underline{S} = \langle \underline{z}, \underline{e}_1 \rangle \underline{e}_1 + \langle \underline{z}, \underline{e}_2 \rangle \underline{e}_2$

$$= (\underline{e}_1^T \underline{z}) \underline{e}_1 + (\underline{e}_2^T \cdot \underline{z}) \underline{e}_2$$

$$= (\underline{e}_1 \cdot \underline{e}_1^T) \underline{z} + (\underline{e}_2 \cdot \underline{e}_2^T) \cdot \underline{z}$$

$$= \left(\begin{bmatrix} e_1 \end{bmatrix} \begin{bmatrix} e_1^T \end{bmatrix} \right) \underline{z} + \left(\begin{bmatrix} e_2 \end{bmatrix} \begin{bmatrix} e_2^T \end{bmatrix} \right) \underline{z}$$

$$= (\underline{e}_1 \underline{e}_1^T + \underline{e}_2 \underline{e}_2^T) \underline{z} = \underline{\underline{P}} \, \underline{z}$$

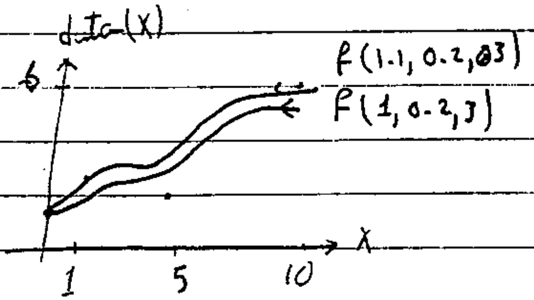


$\underline{e} \perp S$

Optimization:

$f(\alpha_1, \alpha_2, \alpha_3) = \alpha_1 + \alpha_2 x + \alpha_3 \tanh x$

Data	Value
X=0	1
X=2	1.5
⋮	⋮
X=10	6



Cost function $\Rightarrow \sum |d(x_k) - \underset{\alpha_1, \alpha_2, \alpha_3}{f(x_k)}|^2 = C(\alpha_1, \alpha_2, \alpha_3)$

$$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\alpha}_3) = \underset{\alpha_1, \alpha_2, \alpha_3}{\arg\min} \, C(\alpha_1, \alpha_2, \alpha_3)$$

$\Rightarrow \frac{\partial C}{\partial \alpha_1} = 0$

$\frac{\partial C}{\partial \alpha_2} = 0$

$\frac{\partial C}{} = 0$

Ex: $f(x,y) = x^2 + y^2 + 4xy + \overset{2}{\cancel{3}}x + \overset{5}{\cancel{4}}y + \cancel{2}1$

cost fn.

$$f(x,y) = (x \;\vdots\; y)\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix} + (2 \;\vdots\; 5)\begin{pmatrix} x \\ y \end{pmatrix} + 1$$

Quadratic form

$$\boxed{f(\underline{X}) = \underline{X}^T \underline{\underline{A}}\,\underline{X} + \underline{b}^T \underline{X} + c}$$ Quadratic cost function

$$\underline{X} = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\nabla_{\underline{X}} f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \underline{0} \rightarrow \begin{pmatrix} 0 \\ 0 \end{pmatrix} \rightarrow \begin{bmatrix} 2x + 4y + 2 \\ 2y + 4x + 5 \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Ex.

$$f(\underline{X}) = \underline{X}^T \underline{\underline{A}}\,\underline{X} + \underline{b}^T \underline{X} + c$$

$$\nabla_{\underline{X}}(\underline{b}^T \underline{X}) = \overset{b}{\underline{X}} \qquad \nabla_{\underline{X}}(\underline{X}^T \underline{\underline{A}} X) = (\underline{\underline{A}} + \underline{\underline{A}}^T)\underline{X} = \left(\begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}^T\right)\begin{pmatrix} x \\ y \end{pmatrix}$$

$$= \begin{bmatrix} 2 & 4 \\ 4 & 2 \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

$$f(x) = \underline{X}^T \underline{\underline{A}}\,\underline{X} + \underline{b}^T \underline{X} + c \; ; \; \nabla_{\underline{X}} f(x) = (\underline{\underline{A}} + \underline{\underline{A}}^T)\underline{X} + \underline{b} = \underline{0}$$

$$\Rightarrow (\underline{\underline{A}} + \underline{\underline{A}}^T)\underline{X} = -\underline{b}$$

$$\Rightarrow \boxed{\underline{X}_* = -(\underline{\underline{A}} + \underline{\underline{A}}^T)^{-1}\underline{b}}$$

$\underline{X}_*$: has the values (x,y) minimizing $f(x,y)$

Ex. $f(z_1, z_2) = \|z_1\|^2 + \|z_2\|^2 + Re(z_1 z_2)$

z_1 & $z_2 \in \mathbb{C}$.
Complex

Alternative method: ↓

$$f(z_{1R}, z_{1i}, z_{2R}, z_{2i}) \; ; \begin{cases} \frac{\partial f(z_1,z_2)}{\partial z_1} = 0 \\ \frac{\partial f(z_1,z_2)}{\partial z_1^*} = 0 \end{cases} \begin{cases} \frac{\partial f}{\partial z_2} = 0 \\ \frac{\partial f}{\partial z_2^*} = 0 \end{cases}$$

(Hayes' 2 chapter)

$$f(z_1, z_2) = z_1 z_1^* + z_2 z_2^* + \frac{z_1 z_2 + z_1^* z_2^*}{2}$$

$$\frac{\partial f}{\partial z_1} = 0 \longrightarrow z_1^* + \frac{z_2^{\cancel{*}}}{2} = 0$$

$$\frac{\partial f}{\partial z_1^*} = 0 \longrightarrow \left(z_1^* + \frac{z_2}{2}\right)^* = 0 \quad \Rightarrow \quad z_1^* + \frac{z_2}{2} = 0 \quad \Rightarrow z_1^* = z_2^* = 0$$

$$\frac{\partial f}{\partial z_2} = 0 \longrightarrow z_2^* + \frac{z_1}{2} = 0 \quad\quad z_2^* + \frac{z_1}{2} = 0 \quad \Rightarrow z_1 = z_2 = 0$$

$$\frac{\partial f}{\partial z_2^*} = 0 \longrightarrow \left(z_2^* + \frac{z_1}{2}\right)^* = 0$$

$$f(z) = \underline{z}^H \underline{\underline{A}}\, \underline{z} + \underline{b}^H \underline{z} + c$$

$$\nabla_{z^*} \underline{z}^H \underline{\underline{A}}\, \underline{z} = (\underline{\underline{A}} + \underline{\underline{A}}^H) z$$

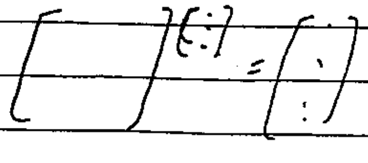
$$\nabla_{z^*} b^H z = \underline{b}$$

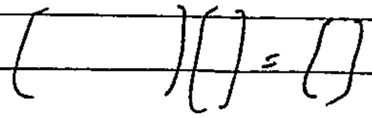
Linear Equation Systems:

$$\underset{M\times N}{\underline{\underline{A}}}\; \underset{M\times 1}{\underline{x}} = \underset{N\times 1}{\underline{b}} \qquad (\underline{a}_1 \;\ldots\; \underline{a}_N)\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \underline{b}$$

$M > N$ Overdetermined $\qquad\qquad$ $M < N$ Underdetermined

$\underline{\underline{A}}$; tall – long matrix $\qquad\qquad$ $\underline{\underline{A}}$: fat. short matrix





→ Uniqe solution $\sum \gamma_k \underline{a}_K = \underline{0} \Longleftrightarrow \gamma_k = 0 \Longleftrightarrow$ Columns are linear independence

→ No solution $\begin{cases} x+y=1 \\ x+y=2 \end{cases}$ $\underline{b}$ is not Range space of $\underline{\underline{A}}$

→ Infinite # of ~ $(x+y=1)$ $\underline{b}$ is in Column space

$\{\underline{a}_1, \dots, \underline{a}_N\}$ linearly independent

$\underline{\underline{A}} = (\underline{a}_1 \; \dots \; \underline{a}_N)$

$\underline{\underline{A}}\,\underline{x} = \underline{a}_1 x_1 + \underline{a}_2 x_2 + \dots + \underline{a}_N x_N \Longrightarrow$ a linear Combination of $\{\underline{a}_1, \dots, \underline{a}_N\}$

$\Downarrow$

Vectors in column $\overset{\wedge}{}$ space of $\underline{\underline{A}}$ matrix.

(Range space = Column space)

Least square Solution for $\underline{\underline{A}}\,\underline{x} = \underline{b}$

$\underline{\underline{A}}\,\underline{x} = \underline{b}_{M \times N}$ $\qquad \underset{\text{error vector}}{\underline{\epsilon}} = \underline{\underline{A}}\,\underline{x} - \underline{b}$

$$\underset{\text{Cost fn}}{J(\underline{x})} = \underline{\epsilon}^T \underline{\epsilon} = \sum_{k=1}^{M} \left| (\underline{\underline{A}}\,\underline{x} - \underline{b})_k \right|^2 \qquad (\cdot)_k \leftarrow \text{K'th element of vector}$$

$$= (\underline{\underline{A}}\,\underline{x} - \underline{b})^T (\underline{\underline{A}}\,\underline{x} - \underline{b}) = (\underline{x}^T \underline{\underline{A}}^T - \underline{b}^T)(\underline{\underline{A}}\,\underline{x} - \underline{b})$$

$$= \underline{x}^T \underline{\underline{A}}^T \underline{\underline{A}}\,\underline{x} - \underline{x}^T \underline{\underline{A}}^T \underline{b} - \underline{b}^T \underline{\underline{A}}\,\underline{x} + \underline{b}^T \underline{b}$$

Transpose of each other and scaler s.s are the sames

$$= \underline{x}^T \underline{\underline{A}}^T \underline{\underline{A}}\,\underline{x} - 2\underline{x}^T \underline{\underline{A}}^T \underline{b} + \underline{b}^T \underline{b}$$

$$J(\underline{x}) = \underline{x}^T \underline{\underline{A}}^T \underline{\underline{A}}\,\underline{x} - 2\underline{x}^T \underline{\underline{A}}^T \underline{b} + \underline{b}^T \underline{b}$$

$$\nabla_{\underline{x}} J = (A^T A + A^T A)\underline{x} - 2\underline{\underline{A}}^T \underline{b} = \underline{0}$$

$$\Rightarrow (\underline{\underline{A}}^T\underline{\underline{A}})\underline{x} = \underline{\underline{A}}^T\underline{b} \Rightarrow \underline{x}_{LS} = (\underline{\underline{A}}^T\underline{\underline{A}})^{-1}\underline{\underline{A}}^T\underline{b}$$

For Complex value: $X_{LS} = (\underline{\underline{A}}^H\underline{\underline{A}})^{-1}\underline{\underline{A}}^H\underline{b}$

$$\underline{\underline{A}}\,\underline{x} = \underline{b}$$

① Unique Sol. $\Longleftarrow$ $\underline{\underline{A}}$: $N \times N$ system ; $\underline{\underline{A}}^{-1}$ exist

$$X_{LS} = (\underline{\underline{A}}^T\underline{\underline{A}})^{-1}A^Tb = \underline{\underline{A}}^{-1}(\underline{\underline{A}}^T)^{-1}\underline{\underline{A}}^T\underline{b} = \underline{\underline{A}}^{-1}\underline{b}$$

order change

② Inconsistent equation systems:

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 4 \end{bmatrix}}_{n\times m}\underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{m\times 1} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$$

over-determined

$$\underline{\underline{A}}\,\underline{x} = \underline{b} \xrightarrow{\& A^T} \underline{\underline{A}}^T\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}^T\underline{b} \Rightarrow X_{LS} = (\underline{\underline{A}}^T\underline{\underline{A}})^{-1}\underline{\underline{A}}^T\underline{b}$$

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix}}_{\underline{\underline{A}}^T}\underbrace{\begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 4 \end{bmatrix}}_{\underline{\underline{A}}}\begin{bmatrix} x \\ y \end{bmatrix} = \underbrace{\begin{bmatrix} 20 \\ 24 \end{bmatrix}}_{\underline{\underline{A}}^T\underline{b}}$$

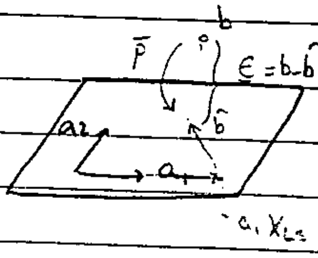
$$\begin{bmatrix} 14 & 17 \\ 17 & \overset{}{\cancel{18}}\,21 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}_{LS} = \begin{bmatrix} 20 \\ 24 \end{bmatrix}$$

→ for invertable eigenvalue $\neq 0$

$\underline{\underline{A}}^T\underline{\underline{A}}$: Grammian matrix

$|A^TA| \neq 0 \Longleftrightarrow$ full column rank

$\Rightarrow \{a_1, \ldots, a_N\} \rightarrow$ linearly independent



$$\hat{b} = A X_{LS} = \underbrace{A(A^T A)^{-1} A^T}_{} b$$

P : Projection matrix

① $P^2 = P$ ② $P^T = P \rightarrow (A(A^TA)^{-1}A^T)^T$

$$(A(A^TA)^{-1}A^T)^2 = A(A^TA)^{-1}A^T \cdot A(A^TA)^{-1}A^T = A(A^TA)^{-1}\cancel{(A^TA)}\cancel{(A^TA)^{-1}}A^T$$

$$= A(A^TA)^{-1}A^T = P \quad ?$$

$$\epsilon = b - \hat{b} = b - Pb = \underbrace{(I - P)}_{P_{A\perp}} b$$

$P_{A\perp}$: Orthogonal Projector to $\{a_1, \cdots, a_N\}$ space

chek: ① $P_{A\perp}^2 \overset{?}{=} P_{A\perp}$

② $P_{A\perp}^T = P_{A\perp}$ $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_N v_N = 0$

$\alpha_i \neq 0$ v are linearly dependent

$\underset{n\times m}{A}\underset{m\times 1}{X} = b$ $m = n$ $X = A^{-1}b$

$n < m \rightarrow$ Underdetermined $X = A^H (AA^H)^{-1} b$ min. norm | IF A is orthogonal:

$A^TA = I \Rightarrow A^T = A^{-1}$

$n > m \rightarrow$ Overdetermined. $X_{LS} = (A^HA)^{-1}A^Hb$. Least square sol. | For Complex matrix

$A^HA = I \Rightarrow A^{-1} = A^H$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & a_{23} \cdots & a_{2n} \\ 0 & 0 & a_{33} & \cdots \end{bmatrix} \qquad \det(A) = \prod_{i=1}^{n} a_{ii}$$

Upper triangular Lower or Upper triangular

A is said to Unitary matrix

$A = \text{Teop}\{1, j, 1-j\} = \begin{bmatrix} 1 & -j & 1+j \\ j & 1 & -j \end{bmatrix}$ A : Hermitian Toeplitz

Under determined Eq. system

$$\underset{M\times N}{\underline{\underline{A}}}\,\underset{N\times 1}{\underline{X}} = \underset{M\times 1}{\underline{b}} \qquad M < N$$

There are infinite No. of sol. we try to minimize:

$$J(\underline{x}) = \|\underline{\underline{A}}\underline{x} - \underline{b}\|^2 + \cancel{\lambda\|x\|^2} \quad (\text{infinite No. of sol. } \underline{x}_k \text{ satisfies } J(\underline{x}_k) = 0)$$

minimize $\|X\|^2$ such that $\underbrace{\underline{\underline{A}}\underline{X} = \underline{b}}_{\text{Constraint of optimization}}$ (called min. norm sol.)

$$\overset{\text{min=norm}}{J(\underline{x})} = X^T X + \underset{\substack{\downarrow \\ \text{Lagrange multiplier}}}{\lambda_1} [\underline{a}_1^T \underline{x} - \underline{b}_1] + \lambda_2 [\underline{a}_2^T \underline{x} - b_2] + \dots + \lambda_m (\underline{a}_m^T \underline{x} - \underline{b}_m)$$

$$\begin{bmatrix} a_1^T \\ \vdots \\ a_M^T \end{bmatrix} \begin{bmatrix} \underline{x} \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_M \end{bmatrix} \quad ; \quad J_\lambda(x) = \underline{X}^T \underline{X} + \underline{\lambda}^T (\underline{\underline{A}}\,\underline{X} - \underline{b})$$

?

$$\Rightarrow \quad \nabla_{\underline{x}} J(x) = 0 \qquad \frac{\partial J(x)}{\partial \lambda_k} = 0 \qquad k = \{1, \dots, M\}$$

$$J_\lambda(\underline{x}) = \underline{X}^T \underline{X} + \underline{\lambda}^T (\underline{\underline{A}}\underline{X} - \underline{b}) \qquad \underline{\lambda} = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_2 \end{pmatrix}$$

$$\left.\begin{aligned} &\nabla_{\underline{x}} J(x) = 0 \\ &\nabla_\lambda J_\lambda(x) = 0 \end{aligned}\right\} \longrightarrow \underline{X}$$

Pseudo-inverse (Pinv.m) and SVD

$$\underline{\underline{A}}\underline{X} = \underline{b} \xrightarrow{\text{Pseudo.inverse}} A^+ \longrightarrow A^+A = I \longrightarrow \begin{matrix} \text{LS sol.} \\ \text{or} \\ \text{min. norm. sol.} \end{matrix}$$

$\searrow$ SVD

SVD $\rightarrow$ $\underline{\underline{A}} = \underline{\underline{U}}\,\underline{\underline{\Sigma}}\,\underline{\underline{V}}^H$ $\rightarrow$ diagonal, Singular vectors

$$\underline{\underline{A}}^{-1} = \underline{\underline{V}}\,\underline{\underline{\hat{\Sigma}}}^{-1}\,\underline{\underline{U}}^H \rightarrow \hat{\Sigma} = \begin{pmatrix} \frac{1}{\Sigma_{11}} & & 0 \\ & \frac{1}{\Sigma_{22}} & \\ 0 & & \frac{1}{\Sigma_{kk}} \end{pmatrix}$$ k non zero singular values

$\underline{\underline{M}}$: Positive definite Matrix

$$\underline{\underline{M}}\,\underline{x} = \underline{b} \longrightarrow \underline{\underline{M}}\,\underline{x}_0 = \underline{0} \qquad \underline{x}_0 \neq \underline{0}$$

$$\begin{bmatrix} 1 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \end{bmatrix} \rightarrow x_0 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

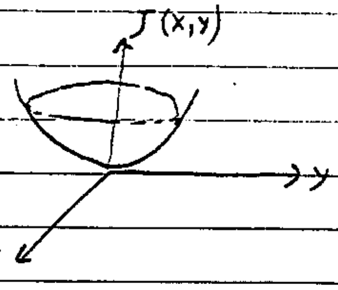
$$\underline{\underline{M}}\,\underline{x}_s = \underline{b} \quad \text{then} \quad \underline{\underline{M}}\,(\underline{x}_s + \underset{\uparrow}{\alpha}\,\underline{x}_0) = \underline{b}$$

by changing α $\longrightarrow$ infinite No. of sol.

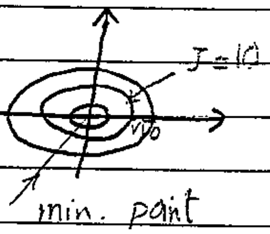
$\underline{\underline{M}}$: Positive definite matrix $\longrightarrow$ $\underline{x}^T\underline{\underline{M}}\,\underline{x} > 0 \quad \forall\,\underline{x} \neq \underline{0}$

$$M = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} x & y \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x^2 + 2y^2 > 0 \quad \text{For } \forall\,(x,y) \neq (0,0)$$

$J(x,y) = x^2 + 2y^2$



$$\left.\begin{aligned} J(x,y) &= 10 \\ x^2 + 2y^2 &= 10 \end{aligned}\right\}$$ level curve.



$\det(A) = \prod_{i=1}^{n} \lambda_i$ λ_i are eigenvalues

___/___/___

positive definite

$\underline{J}(\underline{x}) = \underline{x}^T \underline{\underline{M}} \underline{x} + \underline{b}^T \underline{x} + \underline{c}$

Unique solution if $\underline{\underline{M}} > 0$ ($\underline{\underline{M}}$: P.d.)

Many Solutions if $\underline{\underline{M}} \geq 0$ ($\underline{\underline{M}}$: positive semi definite)

$\underline{\underline{M}}\, \underline{x}_0 = \underline{0}$

non-zero vector

if $\underline{\underline{M}}$ is not P.d or P.S.d → There is not min. for Cost Fn.

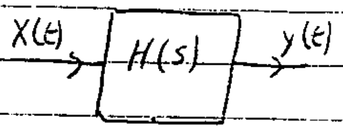
$\underline{\underline{A}}\underline{x} = \underline{b} \rightarrow J(\underline{x}) = \|\underline{\underline{A}}\underline{x} - \underline{b}\|^2 = \underline{x}^T \underbrace{\underline{\underline{A}}^T \underline{\underline{A}}} \underline{x} + 2\underline{x}^T \underline{\underline{A}}^T \underline{b} + \underline{b}^T \underline{b}$

$\underline{\underline{M}} = \underline{\underline{A}}^T \underline{\underline{A}} \longrightarrow M > 0$ or $M \geq 0$

$x^T M x = x^T A^T A x = (\underline{\underline{A}}\underline{x})^T (\underline{\underline{A}}\underline{x}) = \|Ax\|^2 \geq 0$

IF A is Hermition → all e.values are real

DSP Review:

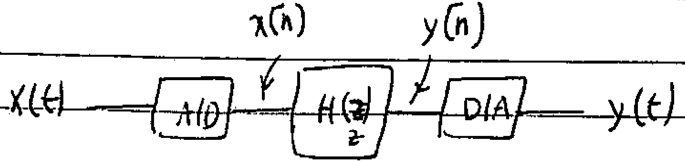


LTI system

$H(s) = \frac{Y(s)}{X(s)} = \frac{s+2}{s^2+2s+4} \Rightarrow Y(s)(s^2+2s+4) = (s+2)X(s)$

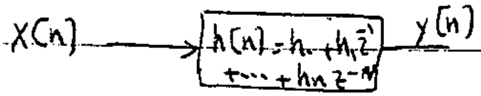
$\Rightarrow \frac{d^2}{dt^2} y(t) + 2\frac{dy}{dt} + 4y(t) = \frac{dx(t)}{dt} + 2x(t)$

MIT 8.02 video
open physic



$H(z) = 1 + z^{-1} \Rightarrow y(n) = x(n) + x(n-1)$

$X(t) = \sum_{n=-\infty}^{+\infty} x(n)\,\mathrm{Sinc}(t-nT)$; $\mathrm{Sinc}(x) = \frac{\sin(\pi x)}{\pi x}$



$y(n) = \sum_{k=0}^{N} h(k)\,x(n-k)$

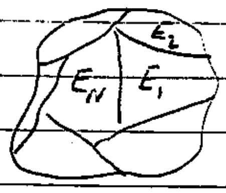
$$\begin{bmatrix} y(0) \\ \vdots \\ y(n) \end{bmatrix} = \underbrace{\begin{bmatrix} h(0) & 0 & 0 & \cdots \\ h(1) & h(0) & 0 & \cdots \\ h(2) & h(1) & h(0) & \cdots \\ \vdots\ h(N) & h(N-1) & \cdots & \\ \vdots & & & h(N) \end{bmatrix}}_{\text{teoplitz Matrix or Convolution Matrix}} \begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(n) \end{bmatrix}$$

P.15 Probability Theory:

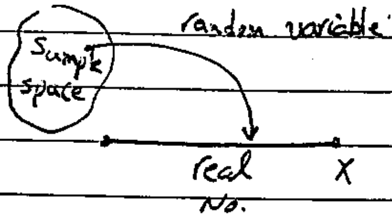
① $P(A) \geq 0$ ② $P(\text{Universal set}) = 1$ ③ $P(A \cup B) = P(A) + P(B)$

$\mathbb{Z}(A \cap B) = \emptyset$
intersection

Events:



mapping event to → a scalar (a random variable)
→ a vector (~ ~ vector)
→ a function (~ ~ process)

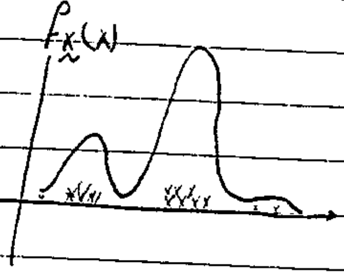


$P(A/B) = \dfrac{P(A \wedge B)}{P(B)}$ Conditional Prob.

Expectation:

$$E\{\underset{\sim}{X}\} = \int_{-\infty}^{+\infty} x \, \overset{\text{pdf}}{f_{\underset{\sim}{X}}}(x)\, dx \quad ; \quad \underset{\sim}{\bar{X}} = \frac{1}{N}\sum_{k=1}^{N} x_k$$

$$P(X=1) = \frac{\#\text{ times } 1 \text{ is observed out of } N \text{ experiment}}{N}$$



$$E\{g(x)\} = \int g(x)\, f_{\underset{\sim}{X}}(x)\, dx \quad ; \quad E\{\hat{g}(x)\} = \frac{1}{N}\sum_{k=1}^{N} g(X_k) \cong E\{g(x)\}$$

Moments: $E\{x\} = \mu_x = \bar{X}$

k^{th} moment $E\{X^k\} = m_k \qquad E\{(X-\bar{X})^2\} = \sigma_X^2 \rightarrow$ varience

$E\{(X-\bar{X})^k\}$ = Central k^{th} moment

Characteristic function:

$\phi(s) = E\{e^{sX}\} \rightarrow$ moment generating function

$$E\{e^{s\underset{\sim}{X}}\} = \int_{-\infty}^{+\infty} e^{sx} f_{\underset{\sim}{X}}(x)\, dx \quad ; \quad \int_{-\infty}^{+\infty} 1\, f_{\underset{\sim}{X}}(x)\, dx = 1$$

$$\frac{d\phi(s)}{ds} = E\left\{\frac{d}{ds} e^{sX}\right\} = E\{X e^{sX}\}$$

$$\phi^{(1)}(\cdot) = E\{X\} \qquad \phi^{2}(\cdot) = E\{X^2\}$$

$$\phi(s) = E\left\{\sum_{k=0}^{\infty} \frac{(sX)^k}{k!}\right\} = \sum_{k=0} \frac{E\{X^k\} s^k}{k!} = \sum_{k=0}^{\infty} \frac{\phi^{k}(0)\, s^k}{k!}$$

___/___/___

$$\phi(s) = \sum_{k=0}^{\infty} \underbrace{\phi^{(k)}(0)}_{E\{X^k\}} \frac{s^k}{k!}$$ (Taylor Series expansion)

Two Random Variables: $(\underset{\sim}{X}, \underset{\sim}{y})$

joint density $\swarrow f_{\underset{\sim}{x},\underset{\sim}{y}}(x,y)$, $F_{\underset{\sim}{x}\underset{\sim}{y}}(x,y) = P\{\underset{\sim}{X} \le x, \underset{\sim}{y} \le y\}$

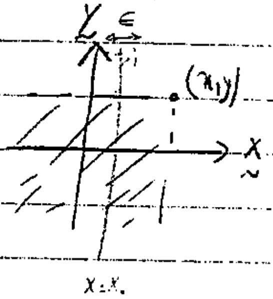
joint PDF

marginal density

$$f_{\underset{\sim}{X}}(x_0) = \int_{-\infty}^{+\infty} f_{x,y}(x_0, y)\, dy$$

$$\downarrow$$

$$\simeq P\{\, |X - x_0| < \epsilon \}$$



Functions of Random variables

$$\underset{\sim}{Z} = f(x,y) \qquad f(x,y) = ax + by$$

$$\underset{\sim}{W} = g(x,y) \qquad g(x,y) = cx + dy$$

$$\begin{bmatrix} \underset{\sim}{Z} \\ \underset{\sim}{W} \end{bmatrix} = \begin{bmatrix} \underset{=}{A} \end{bmatrix} \begin{bmatrix} \underset{\sim}{x} \\ \underset{\sim}{y} \end{bmatrix}$$

↑ 2 fun. of 2 r.v.

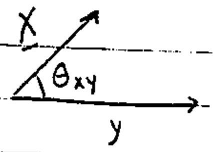
$\underset{\sim}{Z} = a\underset{\sim}{X} + b\underset{\sim}{y}$ ← 1 fun. of 2 r.v.

$\phi(s_1, s_2) = E\{e^{s_1 \underset{\sim}{x} + s_2 \underset{\sim}{y}}\}$ moment generating fun.

$E_{\underset{\sim}{X}/\underset{\sim}{y}}\{X\} = E_{\underset{\sim}{y}}\{E_{\underset{\sim}{x}}\{\underset{\sim}{x}/\underset{\sim}{y}\}\}$ itearated expectation

$E\{g(x,y)\} = E_y\{E_x\{g(x,y)/y\}\}$?

How to measure the similarity of $\underset{\sim}{x}$ and $\underset{\sim}{y}$:



Finite dime $\underline{x}$ and $\underline{y}$ vectors
(N dimension)

$$\cos(\theta_{x,y}) = \frac{\underline{x}^T \underline{y}}{\|x\| \cdot \|y\|} = \frac{\sum_{k=1}^{N} x_k y_k}{\sqrt{\sum x_k^2}\sqrt{\sum y_k^2}}$$

$$r_{xy} = \frac{E\{xy\}}{\sqrt{E\{x^2\}E\{y^2\}}} = \frac{\iint xy\, f_{\underset{\sim}{x}\underset{\sim}{y}}(x,y)\,dx\,dy}{\sqrt{\iint x^2 f_{\underset{\sim}{x}\underset{\sim}{y}}(x,y)\,dx\,dy}\sqrt{\iint y^2 f_{xy}(x,y)\,dx\,dy}}$$

Properties:

① $|r_{xy}| \leq 1$ ② if $r_{xy} = \pm 1 \xleftrightarrow[\text{iff}]{} \underset{\sim}{x} = a\underset{\sim}{y}$ (linear Combination)

Here we assume $E\{x\} = E\{y\} = 0$

Corr. Coef. :

$$C_{xy} = \frac{E\{(x-\bar{x})(y-\bar{y})\}}{\sqrt{E\{(x-\bar{x})^2\} \cdot E\{(y-\bar{y})^2\}}}$$

$C_{xy} = r_{xy}$ if $\bar{x} = \bar{y} = 0$

① $|C_{xy}| \leq 1$ ② $C_{xy} = \pm 1 \longleftrightarrow \underset{\sim}{y} = a\underset{\sim}{x} + b$ (affine Combination) but not linear

Ex: $\underset{\sim}{y} = a\underset{\sim}{x} + \underset{\sim}{n}$

$\underset{\sim}{x}$: r.v
n: r.v independent of x

Note that I can assume $\underset{\sim}{x}$ and $\underset{\sim}{n}$ are Zero mean.
(Since Cxy is a not a fcn. of mean $(\bar{x}, \bar{y})$)

$$r_{xy} = \frac{R_{xy}}{\cancel{\sigma_x\sigma_y}\ \sqrt{E(x^2)E(y^2)}} \qquad C_{xy} = \frac{Cor(x,y)}{\sigma_x\sigma_y}$$

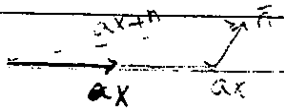
$$\downarrow \rho_{xy}$$

$$C_{xy} = \frac{E\{xy\}}{\sqrt{E\{x^2\}E\{y^2\}}} = \frac{E\{xy\}}{\sigma_x . \sigma_y} = \frac{E\{x(ax+n)\}}{\sqrt{E\{x^2\}E\{(ax+n)^2\}}}$$

$$= \frac{a\sigma_x^2 + E\{x\}E\{n\}}{\sigma_x\sqrt{\sigma_x^2a^2 + \sigma_n^2 + 2\bar{x}\bar{n})}} = \frac{a\sigma_x^2}{\sigma_x . \sigma_x\sqrt{a^2 + \frac{\sigma_n^2}{\sigma_x^2}}} = \frac{a}{\sqrt{\cdots}}$$

$$\Rightarrow C_{xy} = \frac{1}{\sqrt{1 + \frac{\sigma_n^2}{a^2\sigma_x^2}}} = \frac{1}{\sqrt{1 + \frac{1}{SNR}}} \quad < 1$$

$$SNR = \frac{\text{Variance of Signal Component in } y}{\text{Variance of noise " "}} = \frac{E\{(ax)^2\}}{E\{n^2\}} = \frac{a^2\sigma_x^2}{\sigma_n^2}$$



In high SNR → x, y ⇒ approach the same vector

Ex: $E\{(x-\bar{x})(y-\bar{y})\} = Cov(x,y) = 0 \longleftrightarrow$ Uncorrelated r.v.'s

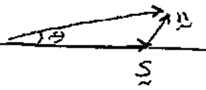
A, B are two events

$$x = \begin{cases} 1 & A \text{ occurs} \\ 0 & \text{others} \end{cases} \qquad y = \begin{cases} 1 & B \text{ occurs} \\ 0 & \text{others} \end{cases}$$

$$Cov(x,y) = E\{(x-\bar{x})(y-\bar{y})\} \qquad C_{xy} = \frac{Cov(x,y)}{\sigma_x\sigma_y}$$

$$= E\{xy\} - \bar{x}\bar{y}$$

$\underset{\sim}{y} = \underset{\sim}{s} + \underset{\sim}{n}$



$s \,\&\, \underset{\sim}{n}$ independent ___/___/___

$$C_{xy} = \frac{E\{(x-\bar{x})(y-\bar{y})\}}{\sqrt{E\{(x-\bar{x})^2\}}\ \sqrt{E\{(y-\bar{y})^2\}}} = \frac{\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle}{\|\underset{\sim}{x}\| \cdot \|\underset{\sim}{y}\|}$$

$|C_{xy}| \leqslant 1 \rightarrow C_\theta \dashrightarrow C_{xy}$

$C_{xy} = Cov(x,y)$ if $\sigma_x^2 = \sigma_y^2 = 1$ (Unit length vectors)

Ex:

$$\underset{\sim}{X} = \begin{cases} 1 & A \text{ occurs} \\ 0 & \text{others} \end{cases} \qquad \underset{\sim}{y} = \begin{cases} 1 & B \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

$Cov(\underset{\sim}{X}, \underset{\sim}{y}) = E\{xy\} - \bar{x}\bar{y} =$

$E\{\underset{\sim}{x}\}E\{\underset{\sim}{y}\}$

event

$E\{xy\} = E\{\underset{\sim}{x}\underset{\sim}{y} \mid \underset{\sim}{x} = 1\}\, P(X=1, y=1) + E\{xy \mid (x,y) = \{(0,0), (1,0), (0,1)\}\}\, P(\text{event}))$

$E\{x\} = \int \overset{}{\underset{1}{x}} \underset{P(A)}{f_x(x)} dx = P(A)$ $\quad Cov(x,y) = P\{x=1, y=1\} - P\{x=1\}P\{y=1\} = E(xy) - \bar{x}\bar{y}$

$Cov(x,y) = 0 \rightarrow P(A,B) = P(A)P(B)$

$\Rightarrow Cov^0(\underset{\sim}{X}, \underset{\sim}{y}) = P(\underset{\sim}{X} = 1, y=1) - P(X=1)P(y=1) = P(A \text{ and } B) - P(A \text{ occurs}) \cdot P(B \text{ o}\_)$

$Cov(x,y) > 0 \quad P(x=1, y=1) > P(x=1)P(y=1)$ occurs together

$P(x=1, y=1) / P(x=1)$

Case 1: $Cov(\underset{\sim}{x}, \underset{\sim}{y}) > 0$ $\quad P(A|B) > P(A)$ or $P(B|A) > P(B)$

$\Rightarrow P(A \text{ and } \underset{\text{occurs}}{B}) \text{ Jointly} > P(A \text{ occurs}) P(B \text{ occurs})$

$$= \frac{P(A \text{ and } B)}{P(A)} > P(B) \quad = P(A/B) > P(B)$$

Ross' Intr. Prob. Models 7^{th} ed.

$\Rightarrow P(A/B) > P(A)$

Case 2: $Cov(\underset{\sim}{X}, \underset{\sim}{y}) = 0 \longrightarrow P(A \text{ and } B) = P(A) \cdot P(B)$

Looks like A & B are independent.

$\underset{\sim}{x}, \underset{\sim}{y}$ $\underrightarrow{\text{ind.}}$ $f_{\underset{\sim}{x}\underset{\sim}{y}}(x,y) = f_{\underset{\sim}{x}}(x) \cdot f_{\underset{\sim}{y}}(y)$

$Cov(\underset{\sim}{x}, \underset{\sim}{y}) = 0 \longrightarrow E\{xy\} = E\{x\}E\{y\}$

$\underset{\sim}{X}$ and $\underset{\sim}{y}$ are orthogonal if $E\{\underset{\sim}{x}\underset{\sim}{y}\}=0$

if $\bar{x}=\bar{y}=0$ orth. r.v's = Uncorrelated r.v.s

Properties of $Cov(X,y)$:

$Cov(X,X) = Var(X)$ $\qquad$ $Cov(c\underset{\sim}{X},\underset{\sim}{y}) = c\,Cov(\underset{\sim}{X},\underset{\sim}{y})$

$Cov(X,y) = Cov(y,X)$ $\qquad$ $Cov(\underset{\sim}{X},\underset{\sim}{y}+\underset{\sim}{z}) = Cov(\underset{\sim}{X},\underset{\sim}{y}) + Cov(\underset{\sim}{X},\underset{\sim}{z})$

Ex: $Var\left(\sum_{k=1}^{N}\underset{\sim}{X_k}\right) = Cov\left(\sum_{k=1}^{N}\underset{\sim}{X_k}, \sum_{l=1}^{N}X_l\right)$

$$\Rightarrow \sum_{k=1}^{N}\sum_{l=1}^{N} Cov(X_k,X_l) = \sum_{\substack{k=1\\(k=l)}}^{N}\sum_{l=1}^{N} Cov(X_k,X_l) + \sum_{\substack{k=1\\k\neq l}}^{N}\sum_{l=1}^{N} Cov(X_k,X_l)$$

$$= \sum_{k=1}^{N} Cov(X_L,X_k) + 2\sum_{\substack{k=1\\(k>l)}}^{N}\sum_{l=1}^{N} Cov(X_k,X_l) = \sum_{k=1}^{N} Var(X_k) + 2\sum_{\substack{k=1\\(k>l)}}^{N}\sum_{l=1}^{N} Cov(X_k,X_l)$$

Note, If X_k and X_l are uncorrelated for all $k\neq l$

then: $Var\sum_{k=1}^{N}X_k = \sum_{k=1}^{N}Var(X_k)$

Later: $\underset{\sim}{z} = \underbrace{(1\ 1\ \dots\ 1)}_{1^T}\begin{bmatrix}X_1\\ \vdots\\ X_N\end{bmatrix}$; $Var(z) = E\{z^2\}$ $\quad \underset{\bar{z}=0}{\downarrow}$

$$= E\{(1^TX)(1^TX)\} = E\{(1^TX)(X^T1)\} = 1^TE\{XX^T\}.1 = 1^TR_x.1$$

$$= (1\ 1\ \dots\ 1)\begin{pmatrix}\sigma_{X_1}^2 & E(X_1X_2) & \cdots \\ & \ddots & \\ & & \sigma_{X_N}^2\end{pmatrix}\begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$$

If A is Hermition ($A^H = A$) → eigenvalue λ is real

$$\boxed{A v_i = \lambda_i v_i \Rightarrow \underset{①}{v_i^H A v_i = v_i^H \lambda_i v_i} \xRightarrow{*T} v_i^H A^H v_i = v_i^H \lambda_i^* v_i \Rightarrow \boxed{v_i^H A v_i = v_i^H \lambda_i^* v_i \quad ②}}$$

42
Hayes

$$①, ② \Rightarrow \lambda_i = \lambda_i^* \Rightarrow \lambda \text{ is real}$$

Random Vectors:

$$\underline{X} = \begin{bmatrix} \underset{\sim}{X_1} \\ \underset{\sim}{X_2} \\ \vdots \\ \underset{\sim}{X_N} \end{bmatrix} \rightarrow f_{\underset{\sim}{X_1}, \underset{\sim}{X_2}, \ldots, \underset{\sim}{X_N}}$$

Joint density

If A Hermitian and $\lambda_i \neq \lambda_j \Rightarrow \langle v_i, v_j \rangle = 0$ orthogonal

$A v_i = \lambda_i v_i \longrightarrow v_j^H A v_i = v_j^H \lambda_i v_i$ ①

$A v_j = \lambda_j v_j \longrightarrow v_i^H A v_j = v_i^H \lambda_j v_j$

$v_j^H A^H v_i = v_j^H \lambda_j^* v_i$

$v_j^H A v_i = v_j^H \lambda_j^* v_i$ ②

$\Rightarrow (\lambda_i - \lambda_j^*) v_j^H v_i = 0 \Rightarrow v_j^H v_i = 0 \Rightarrow$ orthogonal

$$\underset{=}{R}_X = E\{X X^T\} = E\left\{ \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix} (X_1 \cdots X_N) \right\} = \begin{bmatrix} E(X_1^2) & E(X_1 X_2) & \cdots & E(X_1 X_N) \\ E(X_2 X_1) & E(X_2^*) & & E(X_2 X_N) \\ \vdots & & \ddots & \vdots \\ E(X_N X_1) & & \cdots & E(X_N^2) \end{bmatrix}$$

↓
Colleltion matrix

$$\underset{=}{C}_X = E\{(\underline{X} - E(X))(\underline{X} - E\{X\})^T\}$$

[illegible] Hayes
If $A > 0 \Rightarrow \lambda_i > 0$
positive definite
$\det(A) = \prod_{j=1}^{n} \lambda_j$

$$\underset{=}{C}_X = \begin{bmatrix} E\{(X_1 - \bar{X}_1)^2 & E\{(X_1 - \bar{X}_1)(X_2 - \bar{X}_2)^T\} \\ E\{(X_2 - \bar{X}_2)(X_1 - \bar{X}_1)^T & E\{(X_2 - \bar{X}_2)^2\} \end{bmatrix}$$

Note: $\bar{X} = E\{\underline{X}\} = E\{\underline{y}\} = 0 \rightarrow R_X = C_X$

Hayes' book and in this course, all r.v's (from now on) are assumed to be zero mean, since we consider mean value as a DC component of the signal (Unless otherwise stated)

Properties of $\underset{=}{R}_X$:

$\underset{=}{R}_X = E[X X^T]$ $\left(\underset{=}{R}_X = E[X X^H] \right)$

$\underline{R_X} \geq 0 \xleftrightarrow{iff} \lambda_k \geq 0 \quad \forall k$

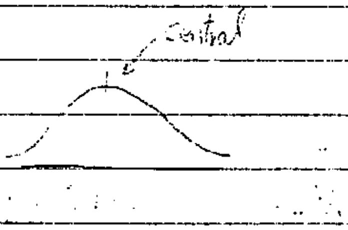
① $\underline{\underline{R}}_X^H = \underline{\underline{R}}_X$ (Hermition matrix or symmetric (real value r.v.'s))

Since $\left(E\left[x\cancel{\;}x^{H}\right]\right)^{H} = E\{\underline{x}\,\underline{x}^{H}\}$

② $\underline{\underline{R}}_X \geq 0$ ($\underline{\underline{R}}_X$ is semipositive definite) $\underline{a}^H \underline{\underline{R}}_X \underline{a} \geq 0 \quad \forall\, \underline{a} \neq 0$

$$\underline{a}^H \underline{\underline{R}}_X \underline{a} = E\{\underbrace{(a^H x)}_{z}\underbrace{(x^H a)}_{\overset{\text{scalar}}{z^*}}\} = E\{\|\underset{\overset{\uparrow}{\underline{a}^H \underline{x}}}{z}\|^2\} \geq 0$$

Gaussian distributions:



Central limit theoriem :

$$\underset{\sim}{\Sigma} = \sum_{k=1}^{N} \underset{\sim}{x}_k \longrightarrow N(\mu_\Sigma, \sigma_\Sigma^2)$$

① $\underset{\sim}{x}_k$'s i.i.d (independent & identically distribution) ② Finite variance then as $N \to \infty$ $\Rightarrow F_{\underset{\sim}{\Sigma}}(\Sigma) \to$ c.d.f of Gaussian dist.

Convergence is in distribution

1-D Gaussions:

$$f_{\underset{\sim}{x}}(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{(x-\mu_x)^2}{2\sigma_x^2}}$$

2-D Gaussians

$$f_{\underset{\sim}{x}_1 \underset{\sim}{x}_2}(x_1, x_2) = \frac{1}{(\sqrt{2\pi})^2 |\underline{\underline{C}}_X|^{1/2}} e^{-\frac{1}{2}[(x_1-\mu_1),(x_2=\mu_2)]\,\underline{\underline{C}}_X^{-1}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}}$$

$|\underline{\underline{C}}_x| \to \det(\underline{\underline{C}}_x)$

$$\underline{\underline{C}}_X = E\left\{\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix}[x_1-\mu_1 \quad x_2-\mu_2]\right\}$$

$\uparrow$ cov. matrix of x_1, x_2

diagonal ↗ $C_x^{-1} = \begin{pmatrix} \frac{1}{\sigma_{x_1}^2} & 0 \\ 0 & \frac{1}{\sigma_{x_2}^2} \end{pmatrix}$

Let's check whether two ind. Gauss. r.v. can be expressed as clai? in 2-D Gauss. form

Then $\underline{\underline{C}}_X = \begin{bmatrix} \sigma_{x_1}^2 & 0 \\ 0 & \sigma_{x_2}^2 \end{bmatrix} \rightarrow Cov(x_1, x_2) = 0 \rightarrow x_1 \,\&\, x_2$ are ind.

$$f_{\underset{\sim}{x_1}\underset{\sim}{x_2}}(x_1,x_2) = \frac{1}{(\sqrt{2\pi})^2}\frac{1}{\sqrt{\sigma_{x_1}^2\sigma_{x_2}^2}} e^{-\frac{1}{2}\left(\frac{(x_1-\mu_1)^2}{\sigma_{x_1}^2} + \frac{(x_2-\mu_2)^2}{\sigma_{x_2}^2}\right)}$$

$$= \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma_{x_1}} e^{-\frac{1}{2}\frac{(x-\mu_1)^2}{\sigma_{x_1}^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_{x_2}^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma_{x_2}^2}} = f_{\underset{\sim}{x_1}}(x_1) \cdot f_{\underset{\sim}{x_2}}(x_2)$$

N-D Gaussians:

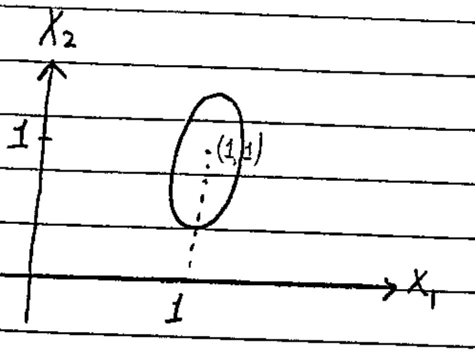
$$f_{\underset{\sim}{x_1},\dots \underset{\sim}{x_N}}(x_1,\dots,x_N) = \frac{1}{(\sqrt{2\pi})^N |\underline{\underline{C}}_x|^{1/2}} e^{-\frac{1}{2}(\underline{x}-\underline{\mu}_x)^T \underline{\underline{C}}_x^{-1} [\underline{x}-\underline{\mu}_x]}$$

Level Curves for 2-D Gaussian P.d.f

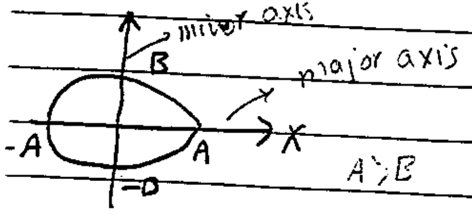
$\mu_{x_1} = 1, \quad \mu_{x_2} = 1 \qquad \underline{\underline{C}}_x = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$

$$f(x_1, x_2) = K e^{-\left(\left(\frac{x_1-1}{2.2}\right)^2 + \frac{(x_2-1)^2}{2.5}\right)}$$

$f(x_1, x_2) = Ke^{-10}$



$+\frac{y^2}{B^2} = 1$



A > B

$$\frac{1}{1-\rho^2}\Big\{ \frac{x_1^2}{\sigma_{x_1}^2} - \frac{2x_1x_2\rho}{\sigma_{x_1}\sigma_{x_2}} + \frac{x_2^2}{\sigma_{x_2}^2}\Big)$$

② $\mu_{x_1}=0$, $\mu_{x_2}=0$ $\underline{C}_x = \begin{bmatrix}\sigma_{x_1}^2 & Cov(x_1,x_2)\\ Cov(x_2,x_1) & \sigma_{x_2}^2\end{bmatrix}$

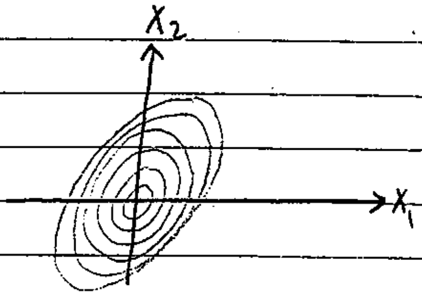
$\rho_{xy} = \frac{Cov(x,y)}{\sigma_x\sigma_y} \Rightarrow \underline{C}_x = \begin{bmatrix}\sigma_{x_1}^2 & \rho\sigma_{x_1}\sigma_{x_2}\\ \rho\sigma_{x_2}\sigma_{x_1} & \sigma_{x_2}^2\end{bmatrix}$

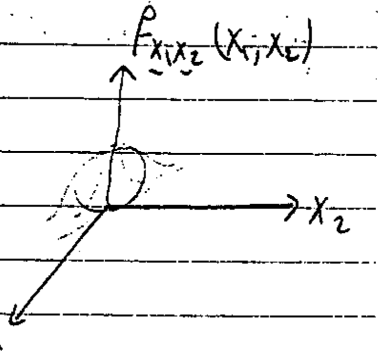
Correlation Coef.

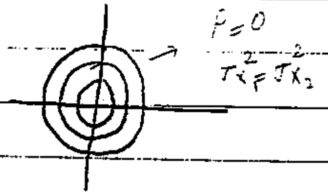
$e^{-\frac{1}{2}\underline{x}^T C_x^{-1}\underline{x}} \Rightarrow \frac{1}{\Delta}\begin{bmatrix}x_1 & x_2\end{bmatrix}\begin{bmatrix}\sigma_{x_2}^2 & -\rho\sigma_{x_2}\sigma_{x_1}\\ -\rho\sigma_{x_1}\sigma_{x_2} & \sigma_{x_1}^2\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} =$

$\Delta = \sigma_{x_1}^2\sigma_{x_2}^2(1-\rho^2)$

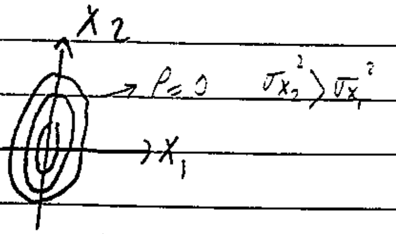
When $\rho>0$

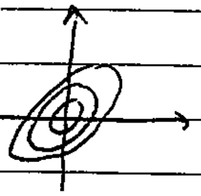






Circular level curves (zero mean, i.i.d variables)
$x_1^2+x_2^2=k$ independent

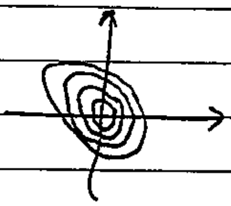




Correlated r.v's!
$\rho>0$

$\frac{x_1^2}{1}+\frac{x_2^2}{10}=k$ independent

x_2 is larger variance



" "
$\rho<0$

Random Vectors:

$$\begin{bmatrix} \underset{\sim}{X_1} \\ \vdots \\ \underset{\sim}{X_N} \end{bmatrix} = \underset{N\times 2}{\underline{\underline{H}}}\,\underline{S} + W \quad \leftarrow N(0, \underline{\underline{R_W}})$$

channel coefficient

equal likly

$$\underset{2\times 1}{\underline{S_k}} = \begin{bmatrix} \pm 1 \\ \pm 1 \\ \pm 1 \end{bmatrix} \qquad E(\underline{S}) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\underline{\underline{R_s}} = E\{\underline{S}\,\underline{S}^H\} \overset{?}{=} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Linear Transformation of Random Vectors: (P.23)

$$\underline{X} \qquad\qquad \underline{y} = \underline{\underline{A}}\,\underline{X}$$

random vectors

deterministic matrix

① $E\{\underline{y}\} = \underline{\underline{A}}\, E\{\underline{X}\} = \underline{\underline{A}}\,\underline{\mu_x}$

② $\underline{\underline{R_y}} = E\{\underline{y}\,\underline{y}^H\} = E\{\underline{\underline{A}}\,\underline{X}\,\underline{X}^H \underline{\underline{A}}^H\} = \underline{\underline{A}}\, E\{\underline{X}.\underline{X}^H\}\underline{\underline{A}}^H = \underline{\underline{A}}\, R_x\, \underline{\underline{A}}^H$

Auto correlation matrix of $\underline{y}$

Diagonalization of Auto-Corr matrices through linear Combining.

Problem: Given $\underline{\underline{R_x}}$ find $\underline{\underline{T}}$ s.t. $\underline{y} = \underline{\underline{T}}\,\underline{X}$ and $\underline{R_y} = \underline{T} R_x \underline{T}^H$ is diagonal.

Methods: ① Diagonalization by Unitary (Orthogonal) Transformation

Approach: $\underline{\underline{R_x}} = \underline{\underline{E}}\,\underline{\underline{\Lambda}}\,\underline{\underline{E}}^H$ (eig. value, eig vec. decomp. of $\underline{\underline{R_x}}$)

$$\underline{\underline{A}} = \underline{\underline{E}}\,\underline{\underline{\Lambda}}\,\underline{\underline{E}}^{-1}$$

$$\uparrow$$

$$\underline{\underline{E}} = [\underline{e_1}, \dots, \underline{e_N}]$$

$R_x e_k = \lambda_k e_k$?

$\underline{\underline{A}}\,\underline{e_k} = \lambda_k \underline{e_k} \quad \Leftarrow \quad \Lambda = \mathrm{dig}(\lambda_1, \dots, \lambda_N)$

$\underline{\underline{R}}_x$: Hermition(symmetric)

→ $\underline{e}_k$'s are orthogonal $\underline{e}_k^{+} \underline{e}_k = 0 \quad k \neq e$

→ λ_k's are real

$$E = [\underline{e}_1 \; .. \; \underline{e}_N]$$

$$\underline{\underline{E}}^H \cdot \underline{\underline{E}} = \begin{bmatrix} \overset{1}{e_1^H e_1} & \overset{0}{e_1^H e_2} & \dots & e_1^H e_N \\ \vdots & & & \vdots \\ e_N^H e_1 & \dots & & \overset{1}{e_N^H e_N} \end{bmatrix}$$

if eig. vectors are orthonormalized

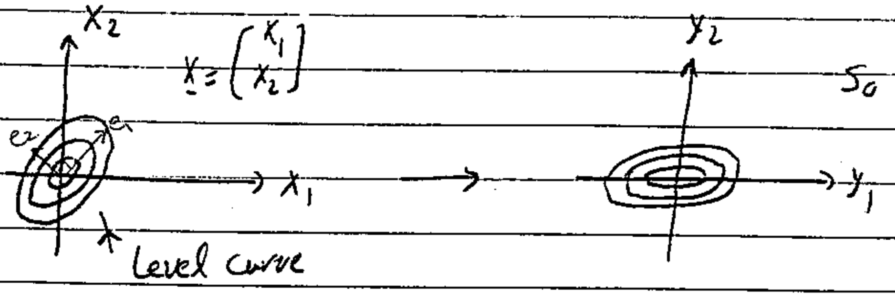
$$\Rightarrow \underline{\underline{E}}^{-1} = \underline{\underline{E}}^H$$

Let's select $\underline{\underline{T}} = \underline{\underline{E}}^H = \begin{bmatrix} e_1^H \\ \vdots \\ e_N^H \end{bmatrix}$ then $\underline{\underline{R}}_y = \underline{\underline{T}}\,\underline{\underline{R}}_x \underline{\underline{T}}^H = \underline{\underline{E}}^H (\underline{\underline{E}} \Lambda \underline{\underline{E}}^H) \underline{\underline{E}}$

$$= \underline{\underline{\Lambda}}$$

$$\underline{X} \xrightarrow{\underline{\underline{T}}\,\underline{X} = \underline{y}} \begin{bmatrix} e_1^H \cdot \underline{X} \\ \vdots \\ e_N^H \cdot \underline{X} \end{bmatrix} = \underline{y}$$

$$\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$



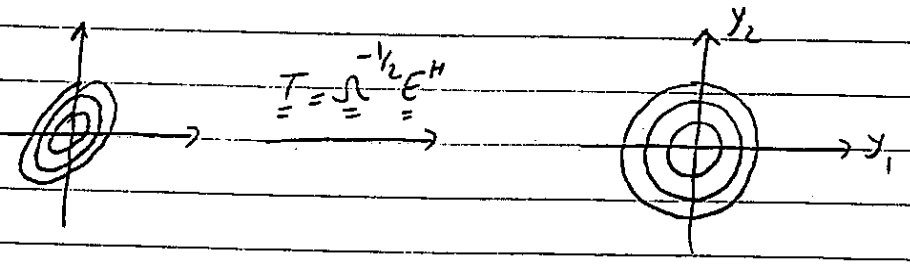
So $R_y = \begin{bmatrix} \sigma_{y_1}^2 & 0 \\ 0 & \sigma_{y_2}^2 \end{bmatrix}$

② ~~Whitening by~~ Unitary Transformation followed by scaling:

Let's select $\underline{\underline{T}} = \underline{\Lambda}^{-1/2} \underline{\underline{E}}^H$

$$\downarrow \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \ddots & \\ & & \frac{1}{\sqrt{\lambda_N}} \end{bmatrix}$$

then $\underline{y} = \underline{\underline{T}}\,\underline{x} \Rightarrow \underline{\underline{R}}_y = \underline{\underline{\Lambda}}^{-1/2} E^H (E \Lambda E^H) E \Lambda^{-1/2} = \underline{\underline{I}}$



③ Diganaliation by LU decomposition:

$$\underline{\underline{A}} = \underset{\substack{\uparrow \\ \text{Lower triangular}}}{\underline{\underline{L}}}\ \overset{\swarrow\ \text{Upper matrix}}{\underline{\underline{U}}}$$

$$\underline{\underline{A}} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ -\frac{c}{a} & 1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ 0 & d - \frac{bc}{a} \end{bmatrix}$$

$$\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 & 0 \\ -\frac{d}{a} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} a & b & c \\ 0 & c - \frac{db}{a} & f - \frac{dc}{a} \\ g & h & i \end{bmatrix}$$

$$\searrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\frac{g}{a} & 0 & 1 \end{bmatrix} \quad '' \quad = \begin{bmatrix} a & b & c \\ 0 & * & * \\ 0 & * & * \end{bmatrix}$$

$$\underbrace{\underline{\underline{L}}_4 \underline{\underline{L}}_3 \underline{\underline{L}}_2 \underline{\underline{L}}_1}_{\hat{\underline{\underline{L}}}} \underline{\underline{A}} = \underline{\underline{U}}$$

$$\Rightarrow \underline{\underline{A}} = (\hat{\underline{\underline{L}}})^{-1} \underline{\underline{U}} \quad \swarrow \text{inverse of lower } \Delta \text{ is also lower } \Delta$$

$$\Rightarrow \boxed{\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}}$$

Unit Lower triangular matrix:

if $\underline{\underline{A}} = \underline{\underline{A}}^H \Rightarrow \underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{L}}^H$

10 20

$$\underline{\underline{L}}_K = \begin{pmatrix} 10 & 0 \\ 5 & 20 \end{pmatrix} = \underbrace{\overset{\underline{\underline{L}}_U}{\begin{pmatrix} 1 & 0 \\ 5/10 & 1 \end{pmatrix}}}_{\substack{\text{Unit lower } \Delta \\ \text{matrix. } \underline{\underline{L}}_U}} \cdot \underbrace{\overset{\underline{\underline{D}}}{\begin{pmatrix} 10 & 0 \\ 0 & 20 \end{pmatrix}}}_{D}$$

Midterm 6 Nov.

Given $\underline{\underline{R}}_x$, I can apply $\underline{\underline{L}}\,\underline{\underline{U}}$ decomp. that is $\underline{\underline{R}}_x = \underline{\underline{L}}\,\underline{\underline{L}}^H$

(Since Hermision or symmetric) $\boxed{\underline{\underline{R}}_x = \underline{\underline{L}}_u DD^H \underline{\underline{L}}_u^H}$

in matlab LU.m, cholesky.m

$\underline{y} = \underline{\underline{T}}\,\underline{x} \longrightarrow$ Select $\underline{\underline{T}} = \underline{\underline{L}}_u^{-1}$ $|\underline{\underline{L}}_u| = 1$

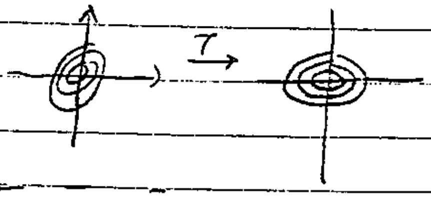
$\underline{\underline{R}}_y = \underline{\underline{T}}\,\underline{\underline{R}}_x T^H = \underline{\underline{D}}\,\underline{\underline{D}}^{*H} = \underline{\underline{D}}$ ← diagonal matrix

If $T = \begin{bmatrix} 1 & & 0 \\ a & 1 & \\ b & c & 1 \end{bmatrix}$; $\underline{\underline{T}}\,\underline{x} = y \Rightarrow T \begin{bmatrix} x(0) \\ x(1) \\ x(2) \end{bmatrix} = \begin{bmatrix} x(0) \\ a x(0) + x(1) \\ b x(0) + c x(1) + x(2) \end{bmatrix}$

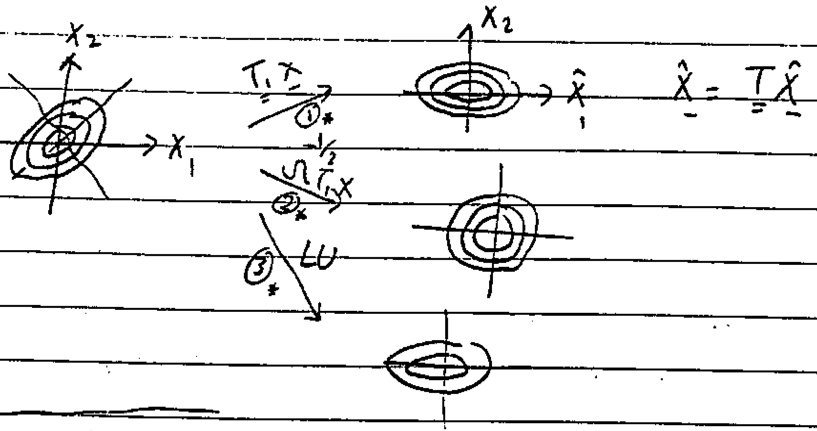
format T ⇒

$= \begin{bmatrix} y(0) \\ y(1) \\ y(2) \end{bmatrix}$

causal system
(″ decomp.)



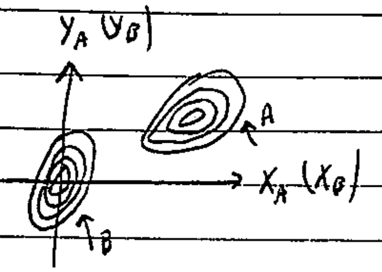
④ Jointly Diagonalization of two Cov. Matrices: (P.2..)



$\underset{\sim}{z} = \underset{\sim}{x} + i\underset{\sim}{y}$ $\underset{\sim}{x}$ & $\underset{\sim}{y}$ are iid and zero mean Gaussian rv
then $E\{xy\} = 0$, $E\{x^2\} = E\{y^2\}$

___/___/______

④



$B \xrightarrow[\text{with sca.}]{\text{Unitary tran.}}$

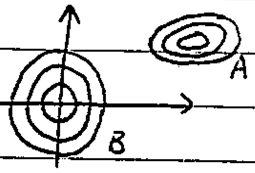
after s1

$A \xrightarrow[\text{Tran.}]{\text{Unit.}}$

step ① , $R_B = E_B \Lambda_B E_B^H$

$$X_B^{\text{after step 1}} = T_B X_B$$

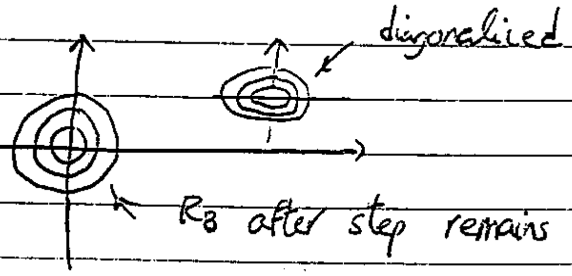
$$\uparrow \quad T_B = \Lambda_B^{-1/2} E_B^H \qquad (\text{as in } 2_*)$$



Note: $R_A^{\text{after s1}} = T_B R_A T_B^H$

step ② : $R_A^{\text{after s}_1} e_k = \lambda_k e_k \leftarrow$ eig. vector of $R_A^{\text{after s1}}$

If I use e_i's to diagonaliz R_A of s1 then I have.



$\nwarrow$ R_B after step remains as it is before $(R_B^{\text{after s1}} = I)$

$$R_A^{\text{after s1}} e_k = \lambda_k e_k \Rightarrow (T_B R_A T_B^H) e_k = \lambda_k e_k$$

$$\Rightarrow (R_A T_B^H) e_k = \lambda_k T_B^{-1} e_k$$

$$R_A T_B^H e_k = \lambda_k \underbrace{\overbrace{(E_B \Lambda^{+1/2})}^{T_B^{-1}} \Lambda^{+1/2} E_B^H}_{R_B} \underbrace{E_B \Lambda^{-1/2}}_{I} \overset{T_B^H}{\nearrow} e_k = \lambda_k R_B T_B^H e_k$$

$$\Longrightarrow R_A \underbrace{T_B^H e_k}_{f_k} = \lambda_k R_B T_B^H e_k$$

$\Rightarrow$ $\boxed{\underline{\underline{R}}_A \underline{f}_k = \lambda_k \underline{\underline{R}}_B \underline{f}_k}$ ← generalized eigenvector of $(\underline{\underline{R}}_A, \underline{\underline{R}}_B)$

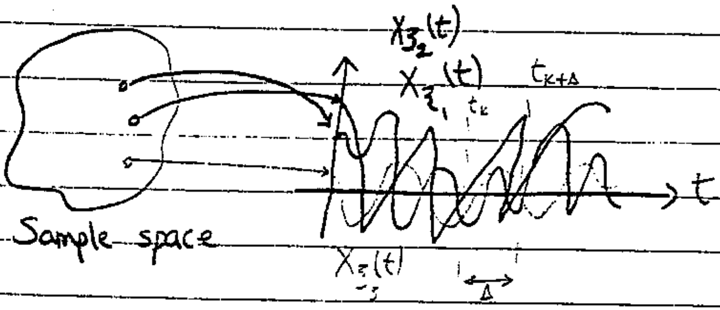
After S_1 and S_2

$$\underline{\underline{T}}_{s_1 \textcircled{s_2}} = \begin{bmatrix} e_1^H \\ \vdots \\ e_N^H \end{bmatrix} \quad \underline{\underline{T}}_B = \begin{bmatrix} f_1^H \\ \vdots \\ f_N^H \end{bmatrix}$$

$\because$ e_k's are orthogonal : $T_{s_1 s_2} \Rightarrow$ On B vector

$$R_B^{\text{after } s_1 s_2} = T_{s_1 s_2} \underline{\underline{R}}_B T_{s_1 s_2}^H = \underline{\underline{E}}^H \underbrace{\underline{\underline{T}}_B \underline{\underline{R}}_B \underline{\underline{T}}_B^H}_{\underline{\underline{I}} \leftarrow \text{identity matrix by step①}} \underline{\underline{E}}^H = \underline{\underline{E}}^H \overset{e_k = (e_1, \dots e_N)}{\underline{\underline{E}}} = I$$

Random processes:



Note: If $\underset{\sim}{X}(t)$ is a random process~~s~~es then:

① $\underset{\sim}{X}(t)\Big|_{t=t_k}$ is a r.v ($E\{X(t_k)\}$ or any fun. of $X(t_k)$ r.v. can be calculated as it is ordinary r.v.

② ξ_k is given $\underset{\sim}{X}_{\xi_k}(t)$ is a fun. of time

③ ξ_k is given at $t=t_k$ $\longrightarrow$ $\underset{\approx}{X}(t_k)$ is a deterministic scalar

Ex: $x[n] = \underset{\sim}{A} \cos(\omega_0 n + \underset{\sim}{\theta})$; $x(t) = \underset{\sim}{A} \cos(\omega_0 t + \underset{\sim}{\theta})$

$\uparrow$ (Gaussian) $N(0, \sigma_A^2)$; $\leftarrow$ Unif. $[0, 2\pi)$

$\underset{\sim}{A}$ & $\underset{\sim}{\theta}$ are independent

Description of Random processes:

① Joint Pdf Description

Assuming we have N samples of $x(t)$

That is $\underset{\sim}{X}(t_1)$, $\underset{\sim}{X}(t_2)$, ——— $\underset{\sim}{X}(t_N)$

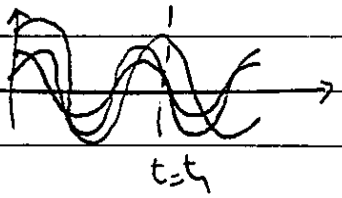
$f_{\underset{\sim}{x_1}(t_1), \dots, \underset{\sim}{x_N}(t_N)}(x_1, x_2, \dots, x_N)$ $\leftarrow$ Joint Pdf of r.p at give t_i's

1st order Pdf : $f_{\underset{\sim}{X}(t_1)}(x_1)$ $\forall t_1$

2nd order Pdf $f_{\underset{\sim}{x_1}(t_1) \underset{\sim}{x_2}(t_2)}(x_1, x_2)$ $\forall (t_1, t_2)$

Ex : $x(t) = A \cos(2\pi f_0 t + \theta_0)$ $\leftarrow$ Unif in $[0, 2\pi)$

1st Order Description:



$$f_{\underset{\sim}{X}(t_1)}(x_1) = \begin{cases} \dfrac{1}{\sqrt{A^2 - x_1^2}} & |x_1| < A \\ 0 & |x_1| > A \end{cases}$$

$\underset{\sim}{\theta} \longrightarrow \underset{\sim}{X}(t_1)$
1 func. of 1 random v.r

$$f_{\underset{\sim}{y}}(y) = \frac{f_x(x_0)}{|g'(x_0)|} \qquad g(x_0) = y$$

2nd order description: $f_{\underset{\sim}{X}(t_1),\, \underset{\sim}{X}(t_2)}(x_1, x_2) = ?$

$$f_{\underset{\sim}{X}(t_1),\underset{\sim}{X}(t_2)}(x_1,x_2) = f_{\underset{\sim}{X}(t_2)|\underset{\sim}{X}(t_1)}(x_2|x_1)\, f_{\underset{\sim}{X}(t_1)}(x_1)$$

$$X(t_2) = A\cos(2\pi f t_2 + \theta) = A\cos(\underbrace{2\pi f_2 t_1 + \theta}_{(I)} + \underbrace{2\pi f(t_2 - t_1)}_{II})$$

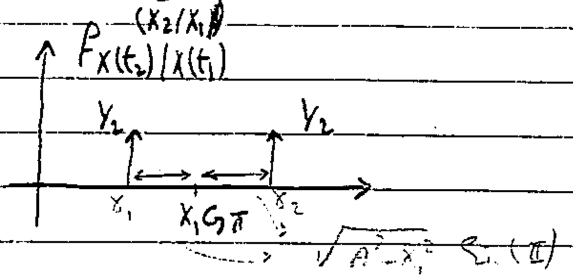
$$= A(\cos(I)\cos II - \sin I \sin II)$$

$\therefore A\cos(I) = x_1 \quad ; \quad A\cos(II) = \gamma$ * deterministic

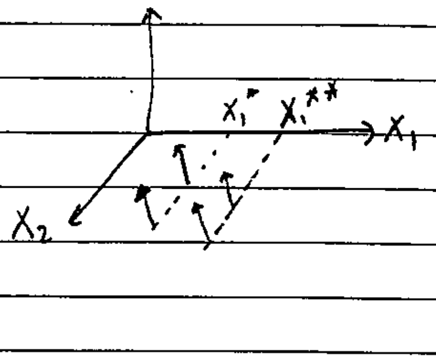
$\Rightarrow X(t_2) = x_1\gamma_1 - A\sin I \sin II = x_1 \cos II - A\sin I \sin II \quad ; \quad A\sin I = \mp\sqrt{A^2 - x_1^2}$

$= x_1 \cos I \pm \sqrt{A^2 - x_1^2}\, \sin II \quad ; \quad \cos(II),\ \sin(II)$ non-random

$$f_{X(t_2)/X(t_1)}(x_2/x_1) = \frac{1}{2}\delta(x_2 - \gamma_1) + \frac{1}{2}\delta(x_2 - \gamma_2)$$



$$f_{\underset{\sim}{X}(t_2),\underset{\sim}{X}(t_1)}(x_2,x_1) = \left(\frac{1}{2}\delta(x_2 - \gamma_1) + \frac{1}{2}(x_2 - \gamma_1)\right) f_{X(t_1)}x_1$$

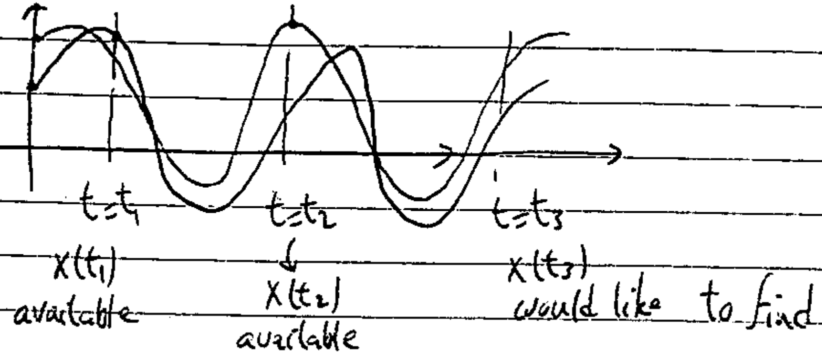


Hay's 2.4, 2.5, 2.17, 3.9 Due: 2^{nd} Nov.

Therian: 2.22, 2.23

3rd order Description

$$f_{\substack{X_1X_2X_3\\ X(t_1)X(t_2)X(t_3)}} = f_{X(t_3)/X(t_1),X(t_2)}(X_3/X_1,X_2)\; f_{X(t_1)X(t_2)}(X_1,X_2)$$



X(t2) given X(t1) Can be related only two different real values.

② Description of R.P's by Moments: (partial description)

1st order $\mu_{X(t_\circ)} = E\{X(t_\circ)\}$ $\forall\ t_\circ$

2nd order $R_X(t_1,t_\circ) = E\{X(t_1)X(t_\circ)\}$, $\forall\ (t_1,t_\circ)$

2nd order P.d.f for Gaussian Process

$$f_{\underset{\sim}{X}(t_1),\underset{\sim}{X}(t_2)}(X_1,X_2) = \frac{1}{(\sqrt{2\pi})^2\,|\underset{\sim}{C}_X|^{1/2}}\, e^{-[X_1-\mu_X(t_1)\,;\,X_2-\mu_X(t_2)]\begin{bmatrix}\sigma^2_{X(t_1)} & \rho\sigma_{X(t_1)}\sigma_{X(t_2)}\\ \rho\sigma_X(t_1)\sigma_X(t_2) & \sigma^2_X(t_2)\end{bmatrix}^{-1}\begin{bmatrix}X_1-\mu_X(t_1)\\ X_2-\mu_X(t_2)\end{bmatrix}}$$

So moment desc. for Gaussian Pr is equivalent to Joint P.d.f description)

Optimization: If $f(z,z^*)$ is real valued fn. of z, z^* and f is analytic

$\Rightarrow$ For Finding stationary point $\frac{\partial f}{\partial z}=0$ or $\frac{\partial f}{\partial z^*}=0$

If $F(z,z^*)$ is real valued Fn. of complex vector z, z^*. F ~ ~

$\Rightarrow$ For Finding stationary point : $\nabla_{z^*}F(z,z^*)=0$

Ex: Let $\underset{\sim}{x}(t)$ be a r.p with $\mu_x(t) = 3$ and $R_x(t_1,t_2) = 9 + 4e^{-0.2(t_1-t_2)}$

$\swarrow$ $E\{x(t_1)x^*(t_2)\}$

Let $Z = \underset{\sim}{x}(\overset{5}{8})$, $\omega = \underset{\sim}{x}(8)$

Find $E\{Z\}$, $E\{\omega\}$, $E\{Z^2\}$, $E\{\omega^2\}$, $E\{Z\omega\}$

$E\{Z\} = E\{\underset{\sim}{x}(5)\} = 3$; $E\{W\} = 3$ $\Bigg\} \Rightarrow \sigma_Z^2 = 13 - 3^2 = 4$

$E\{\underset{\sim}{Z}^2\} = E\{x^2(3)\} = R_x(3,3) = 13$

$E\{W^2\} = R_x\{8,8\} = 13$

$E\{ZW\} = R_x(5,8) = 9 + 4e^{-0.6}$

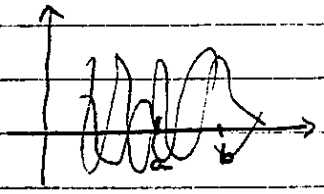
average power

Ex: $\underset{\sim}{Z} = \underset{\sim}{x}(t_1) + \underset{\sim}{x}(t_2)$; $E\{Z^2\} = ?$

$E\{Z^2\} = \underset{\substack{\swarrow \\ \text{av power} \\ \text{at } t_1}}{R_x(t_1,t_1)} + 2\underset{\substack{\swarrow \\ E\{x(t_1)x(t_2)\}}}{R_x(t_1,t_2)} + R_x(t_2,t_2)$

Ex:

$$S = \int_a^b x(t)dt \quad \text{(stockastic Integral)}$$



a) $E\{S\} = \int_a^b E\{x(t)\}dt = \int_a^b \mu_x(t)dt$

b) $E\{S^2\} = E\{(\int_a^b x(t_1)dt_1)(\int_a^b x(t_2)dt_2)\} = \int_a^b\int_a^b E\{x(t_1)x(t_2)\}dt_1dt_2$

$= \iint R_x(t_1,t_2)dt_1dt_2$

Ex: $x(t) = \underset{\sim}{r}\cos(\omega t + \underset{\sim}{\theta})$; $(\underset{\sim}{r}, \underset{\sim}{\theta})$ are independent $\underset{\sim}{\theta}$ is Uniform $[-\pi,\pi]$

$\underset{\sim}{r}$: pdf not given

a) $\mu_x(t)$

$E\{x(t)\} = E\{\underset{\sim}{r}\cos(\omega t + \underset{\sim}{\theta})\} = E\{r\}E\{\cos(\omega t + \underset{\sim}{\theta})\} = 0$

$\int^{2\pi} \frac{1}{\ }\cos(\omega t + \theta)\,d\theta = 0$

b) $R_x(t_1,t_2) = E\{x(t_1)x(t_2)\} = E\{r\cos(\omega t_1+\theta_1)\, r\cos(\omega_2 t_2+\theta)\}$

$= E_r\{r^2\}\, E_\theta\{\underbrace{\cos(\omega t_1+\theta)\cos(\omega t_2+\theta)}_{\frac{1}{2}\cos(\omega(t_1-t_2))+\frac{1}{2}\cos(\omega(t_1+t_2)+2\theta)}\} = \frac{1}{2}E\{r^2\}\cos(\omega(t_1-t_2))$

Note: $R_x(t_1,t_2)$ is a fn. of (t_1-t_2) ; stationary

White noise: $\underset{\sim}{w}(t)$ is called white noise if

$E\{w(t_1)w(t_2)\}=0 \quad t_1 \neq t_2 \qquad E\{w(t)\}=0 \qquad E\{w^2(t)\} = \sigma^2_{w(t)}$

↓ Variance can be different at different time

~~at stationary white noise~~

$\Rightarrow E\{w(t_1)w(t_2)\} = \sigma^2_{w(t)}\,\delta(t_1-t_2)$

Note: stationary white noise has:

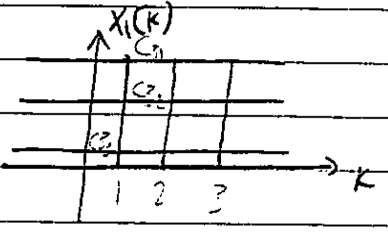
$E\{w(t_1)w(t_2)\} = \sigma_w^2\,\delta(t_1-t_2)$ variance is constant not a fn. of time

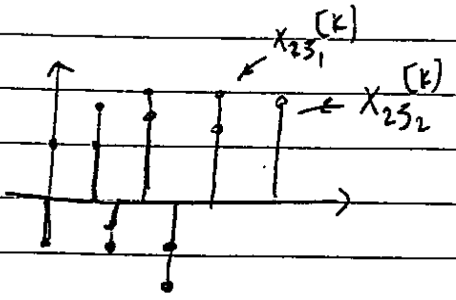
Ex: Two random process: $\underset{\sim}{X}_1(k)$, $\underset{\sim}{X}_2(k)$

$X_1(k) : N(0, \sigma_w^2)$

$X_2(k) = \underset{\sim}{w}_k$

where w_k is i.i.d $N(0,\sigma_w^2)$





1^{st} order P.d.f

$X_1[k]$	$X_2[k]$
$f_{X_1(k)}^{(x_1)} = \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-x_1^2/2\sigma_w^2}$	$f_{X_2(k)}^{(x_2)} = \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-\frac{x_1^2}{2\sigma_w^2}}$

2^{nd} order P.d.f

$X_1[k], X_1[k_2]$	$X_2[k_1], X_2[k_2]$
$f_{X_1(k_1)X_2(k_2)}^{x_1^{k_1}, x_2^{k_2}} = \ldots$	$f_{X_2(k_2)X_2(k_2)}^{x^{k_1} x^{k_2}} = \left(\frac{1}{(\sqrt{2\pi\sigma_w^2})} e^{-\frac{x^{k_1^2}}{2\sigma_w^2}}\right)\left(\frac{1}{\sqrt{\ldots}} e^{-\frac{x^{k_2^2}}{2\sigma_w^2}}\right)$
$= \frac{1}{\sqrt{2\pi\sigma_w^2}} e^{-k_1^2/2\sigma_w^2} \underbrace{\delta(x^{k_1} - x^{k_2})}_{\text{Conditional density}}$	

Moment description

$X_1[k]$	$X_2[k]$
$E\{X_1[k]\} = 0$	$E\{X_2[k]\} = 0$
$E\{X_1(k)X_2[k]\} = \sigma_w^2$	$E\{X_1[k]X_2[k]\} = \sigma_w^2 \delta(k_1 - k_2)$

Gaussian dis. stationary white noise

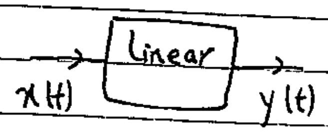
Roles for valid Autocorrelation:

$r_x(\cdot) \geq r_x(k)$

$R_x \geq 0$

$S_x(e^{j\omega}) = S_x(z) \geq 0$

Linear System with stachastics Inputs: (Page 31)



$$y(t) = \int_{-\infty}^{+\infty} h(\underset{\tau}{\overset{\uparrow}{t,\tau}})\, x(\underset{\tau}{\overset{\uparrow}{\tau}})\, d\underset{\tau}{\overset{\uparrow}{\tau}}$$

What are the mean and $R_y(t_1,t_2)$ after linear processing?

We make use of $E\{L\{(\cdot)\}\} = L\{E\{(\cdot)\}\}$ interchange two linear operators

1) $\mu_y(t) \rightarrow E\{y(t)\} = E\{L\{x(t)\}\} = L\{E\{x(t)\}\} = L\{\mu_x(t)\} = \int_{-\infty}^{+\infty} h(t,\tau)\mu_x(\tau)\, d\tau$

2) $R_y(t_1,t_2) \Rightarrow$

$$R_{xy}(t_1,t_2) = E\{x(t_1)\,y(t_2)\} = E\{\overset{\text{Fixed Value}}{x(t_1)}\, L\{x(t)\}\big|_{t=t_2}\}$$

$$= E\{L\{x(t_1)\,x(t)\}\big|_{t=t_2}\} = L\{E\{x(t_1)\,x(t)\big|_{t=t_2}\}\} = L\{R_x(t_1,t)\big|_{t=t_2}\}$$

$$y(t_1,t_2) = \int h(t,\tau)\, R_x(t_1,\tau)\, d\tau \Big|_{t=t_2}$$

If system is LTI, $y(t) = \int \hat{h}(t-\tau)\,x(\tau)\,d\tau$

For LTI system: $R_{xy}(t_1,t_2) = \int \hat{h}(t-\tau)\,R_x(t_1,\tau)\,d\tau \Big|_{t=t_2}$

...

$$R_y(t_1,t_2) = E\{y(t_1)\,y(t_2)\} = E\{L\{x(t)\}\big|_{t=t_1} \overset{\text{Fixed}}{y(t_2)}\} = L\{E\{x(t)\,y(t_2)\}\}\Big|_{t=t_1}$$

$$= L\{R_{xy}(t_1,t_2)\}\big|_{t=t_1} = \int h(t,\beta)\, R_{xy}(\beta,t_2)\, d\beta \Big|_{t=t_1}$$

$$R_y(t_1,t_2) = \iint_{-\infty}^{+\infty} h(t_1,\beta)\, h(t_2,\tau)\, R_x(\beta,\tau)\, d\tau\, d\beta$$

If LTI $\rightarrow R_y(t_1,t_2) = \iint \hat{h}(t_1-\beta)\,\hat{h}(t_2-\tau)\, R_x(\beta,\tau)\, d\tau\, d\beta$

Let's remember that $\underline{y} = \underline{\underline{A}}\,\underline{X}$ is the discrete time version of the pre. results:

$$\underline{\underline{R_y}} = \underline{\underline{A}}\,\underline{\underline{R_X}}\,\underline{\underline{A}}^T$$

$$(R_y)_{k,l} = R_y(k,l) = (\underline{\underline{A}}\,\underline{\underline{R_X}}\,\underline{\underline{A}}^T)_{k,l} = \sum_{\beta=1}^{N}\sum_{\tau=1}^{N} A_{k,\tau} R_X(\tau,\beta) (A^T)_{(\beta,l)}$$

$$(\underline{\underline{A}}\,\underline{\underline{B}})_{k,l} = \sum_{\tau=1}^{N} A_{k,\tau} B_{\tau,l}$$

Gaussian processes when operated under linear operator the resulting process is also Gaussian! So partial description $\mu_y(t)$, $R_y(t_1,t_2)$ is sufficient to give p.d.f description of the output.

Stationarity of Random Processes

If $X(t)$ is a stationary r.p. then, if the time shifted processes that is $\underset{\sim}{X}(t-\Delta)$, results in the same processes. The stationary processes has no time origion!

Description for stationary Processes:

Ⓐ P.d.f Descriptions:

$$f_{X(t_1)X(t_2)\cdots X(t_N)}(x_1,\cdots,x_N) = f_{X(t_1+\Delta)X(t_2+\Delta)\cdots X(t_N+\Delta)}(x_1,x_2,\cdots,x_N)$$

N^{th} order stationarity, if satisfied for $\forall (t_1,\cdots,t_N)$ and $\forall \Delta$ scalars

Spacial Cases:

1^{st} order stationarity: $f_{X(t_1+\Delta)}(x_1) = f_{X(t_1)}(x_1)$

2^{nd} " " $f_{(t_1+\Delta)(t_2+\Delta)}(x_1,x_2) = f_{X(t_1)X(t_2)}(x_1,x_2)$

Ⓑ Moment Descriptions: $\mu_X(t_1)$; $R_X(t_1,t_2)$

If process is stationary then: $E\{\underset{\sim}{x}(t_1)\} = E\{\underset{\sim}{x}(t_1+\Delta)\} \Rightarrow \mu_X(t_1) = \mu_X(t_1+\Delta)$ for all $\forall\Delta$

$\Rightarrow \mu_{X_1}(t) = C = \mu_X$ mean of the process is not fn. of time

$R_X(t_1,t_2) = E\{X(t_1)X(t_2)\} = E\{X(t_1+\Delta)X(t_2+\Delta)\} = R_X(t_1+\Delta, t_2+\Delta) \quad \forall\Delta$

then : $R_X(t_1,t_2) = R_X(t_1+\Delta, t_2+\Delta) \quad \forall\Delta$

$$\underset{\substack{\uparrow\\ \text{double var.}\\ \text{fn.}}}{R_X(t_1,t_2)} = \underset{\substack{\nwarrow\\ \text{Single variable fn.}}}{f(t_1-t_2)}$$

So $R_X(t_1,t_2) = f(t_1-t_2)$ (for $f(t)$ fn.)

then $R_X(t_1,t_2) = R_X(t_1+\Delta, t_2+\Delta)$ for $\forall\Delta$

So this is a sufficient Condition

If $x(t)$ is only stationary in moment : $\begin{cases}\mu_X(t) = \mu_X\\ R_X(t_1,t_2) = R_X(t_1-t_2)\end{cases}$

we sade it Wide Sense stationary (WSS)

$(N^{th}$ stationarity$) \longrightarrow$ WSS

$N \geq 2$

Ex. $\underset{\sim}{x}(t) = \underset{\sim}{a}\cos\omega t + \underset{\sim}{b}\sin\omega t$ Find condition on $\underset{\sim}{a}, \underset{\sim}{b}$ s.t $\underset{\sim}{x}(t)$ is WSS

Ⓐ $E\{x(t)\} = C \Rightarrow E\{a\}\cos\omega t + E\{b\}\sin\omega t = C \quad \forall t$

Only satisfied $E\{a\} = E\{b\} = 0 \Rightarrow C = 0$

⑧ $R_X(t_1,t_2) = R_X(t,t-\tau)$

$R_X(t,t)\big|_{\tau=0}$ should be idepedent of time ① $\rightarrow E\{(a\cos\omega t + b\sin\omega t)^2\}$

$= E\{a^2\}\cos^2\omega t + E\{b^2\}\sin^2\omega t + E\{ab\}\sin 2\omega t = E$

① $\Rightarrow R_X(0,0) = R_X(\frac{\pi}{2\omega}, \frac{\pi}{2\omega}) \Rightarrow E\{x(\cdot)x(\cdot)\} = E\{x(\frac{\pi}{2\omega})x(\frac{\pi}{2\omega})\}$

$\rightarrow E\{a^2\} = E\{b^2\}$ (2i)

$R_X(t,t-\tau) \Rightarrow E\{(a\cos\omega t + b\sin\omega t)(a\cos\omega(t-\tau) + b\sin(\omega(t-\tau)))\}$

$= \underbrace{E\{a^2\}}\cos\omega t\cos\omega(t-\tau) + \underbrace{E\{b^2\}}\sin\omega t\sin\omega(t-\tau) + E\{ab\}\{\sin\omega t\cos(\omega(t-\tau)) + \sin\omega(t-\tau)\cos\omega t\}$

→ time dependence

$= E\{a^2\}\cos\omega\tau + E(ab)\{\ldots\}$

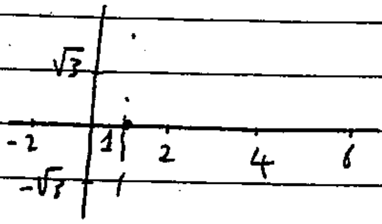
↗ $E(ab) = 0$ (2ii)

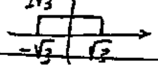
The same problem can be posed as finding condition for strict sense stationarity. The solution is given in Papoulis p. 301 Ex 10_13

Ex: $x(n)$ is a discrete time r.p. At even samples, $x[n]$ is Unif. $(-\sqrt{3}, \sqrt{3})$

At odd samples, $x[n]$ is $N(0,1)$

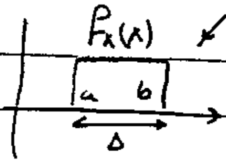
All samples are independent from each other, Comment on the stationary on the processes





$E(x) = \int_{\text{odd}} x f_x(x) dx = 0$

Unif.



$\sigma_x^2 = \frac{\Delta^2}{12}$

WSS:

$$E\{x[n]\} = \begin{cases} 0 & n = \text{even} \\ 0 & n = \text{odd} \end{cases}$$

$$R_x(n, n-k) = \begin{cases} E\{x[n]x[n-k]\} = \sigma^2_{x[n]} \delta[k] & n = \text{even} \\ \sigma^2_{x(\text{odd})} \delta[k] & n = \text{odd} \end{cases} = \delta[k] \quad (\text{white noise})$$

n or $2n$

$1 = \Delta$

$= 1$

SSS

1^{st} order : $f^{x_{even}}_{x(even)} = f^{x_{odd}}_{x(odd)}$

Unif.   Gaussian

Not even 1^{st} order stationary

Jointly WSS Processes and LTI systems:

Jointly WSS

) $x(n), y(n)$ are WSS R.P's

) $R_{xy}(k_1, k_2) = E\{x(k_1) y^*(k_2)\} = f_{unc}(k_1 - k_2)$

$$x(n) \rightarrow \boxed{LTI} \rightarrow y[n]$$

known to be WSS

Is $y(n)$ WSS?

$$E\{y(n)\} = E\{\sum h(k) x(n-k)\} = E\{\sum h(k) \mu_x\} = \mu_x \sum_{k=-\infty}^{+\infty} h(k)$$

$$= \mu_x H(e^{j\omega})\Big|_{\omega=0} = \mu_x H(0)$$

3) $R_{yy}[n, n-k] = ?$ $R_{xy}[n, n-k] = E\{x(n) y^*(n-k)\}$

$$= \sum_{k_1} h^*[k_1] E\{x(n) x^*(n-k-k_1)\} = \sum_{k_1} h^*(k_1) R_x(k+k_1)$$

the process $x(t)$ is SSS if the joint density $f(a,b)$ of the RVs a and b has circular symmetry, that is if

$$f(a,b) = f(\sqrt{a^2+b^2})$$

$$k_2 = -k_1 \Rightarrow \; = \sum_{k_2} h^*[-k_2] R_x(k-k_2) = g(k) * R_x(k) = h^*(-k) * R_x(k)$$

$$R_{yy}(n, n-k) = E\{y(n) y^*[n-k]\} = E\{\sum h(k_1) x(n-k_1) y^*[n-k]\}$$

$$= \sum h(k_1) R_{xy}(k-k_1) = h(k) * R_{xy}(k)$$

$$\Rightarrow R_{yy}(k) = h(k) * h^*(-k) * R_{xx}(k)$$

?

$\Rightarrow y[n]$: WSS and $x(n), y(n)$ Jointly WSS.

Power spectral Density (P.S.D):

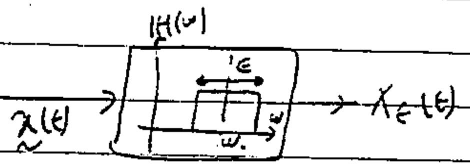
$x(t) = \underset{\sim}{a} \cos\omega t + \underset{\sim}{b} \sin\omega t$ (WSS process)

Definition: Let $h_\epsilon(t)$ be the impulse response of a filter with

$$H_\epsilon(j\omega) = \begin{cases} \sqrt{\frac{2\pi}{\epsilon}} & |\omega-\omega_0| < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

Then P.S.D is:

$$S_{xx}(j\omega_0) = \lim_{\epsilon \to 0} Var(X_\epsilon(t))$$



Wineer Khinchine Theorim:

$$S_{xx}(j\omega_0) = \int \underbrace{C_{xx}(\tau)}_{\uparrow \text{ auto-covariance func.}} e^{-j\omega_0\tau} d\tau$$

Proof: $Var(X_\epsilon(t)) = C_{x_\epsilon x_\epsilon}(0)$

$$C_{x_\epsilon x_\epsilon}(\tau) = h_\epsilon(\tau) * h_\epsilon^*(-\tau) * C_{xx}(\tau) \Rightarrow F\{C_{x_\epsilon x_\epsilon}(\tau)\} = |H_\epsilon(j\omega)|^2 F\{C_{xx}(\tau)\}$$

$$C_{x_\epsilon x_\epsilon}(0) = F^{-1}\left\{ |H_\epsilon(j\omega)|^2 F\{C_x(\omega)\}\right\}\Big|_{\tau=0} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |H_\epsilon(j\omega)|^2 F\{C_x(\omega)\} \overset{\tau=0\Rightarrow 1}{\cancel{e^{j\omega\tau}}} d\omega$$

$$= \frac{1}{2\pi}\int_{|\omega-\omega_0|<\epsilon} \frac{2\pi}{\epsilon} F\{C_x(\tau)\}\, d\omega = \frac{1}{\epsilon}\int_{|\omega-\omega_0|<\epsilon} F\{C_x(\tau)\} d\tau$$

So as $\epsilon \to 0$ then Proof of the claimed relation easily follows.

Properties:

1. P.S.D is real
2. P.S.D is positive
3. $\sigma^2_{X(t)} \Rightarrow$ Area under PSD $\quad \frac{1}{2\pi}\int S_X(j\omega)d\omega$

* 4. If $r_X(K)$ is a valid autocorrelation then:

$$\text{Valid auto-Corr} \xleftrightarrow[\text{iff}]{} \underline{\underline{R}}_X \geq 0 \xleftrightarrow[\text{iff}]{} S_X(e^{j\omega}) \geq 0 \quad \forall \omega$$

Note on 3. $\quad S_X(j\omega) = \int_{-\infty}^{+\infty} C_X(\tau) e^{-j\omega\tau} d\tau$

$$\underset{\tau=0}{\underbrace{C_X(\tau)}} = \frac{1}{2\pi}\int_{-\infty}^{+\infty} S_X(j\omega) e^{j\omega\tau} d\omega \Big|_{\tau=0}$$

$$Var(x(t)) = \frac{1}{2\pi}\int S_X(j\omega)d\omega$$

Ex 10-24 of Papulis P322

Ex: $s(t) = a e^{j\omega_0 (t - \frac{r(t)}{c})}$ $\qquad r(t) = r_0 + \underset{\sim}{v} t$ ← velocity

$$r_s(\tau) = E\{s(t)s^*(t-\tau)\} = a^2 E\{e^{j\omega_0(\tau - \frac{r(t)-r(t-\tau)}{c})}\}$$

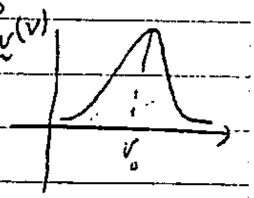
$$= a^2 E\{e^{j\omega_0\tau(1-v/c)}\} = a^2 e^{j\omega_0\tau} E\{e^{-j\omega_0\tau v/c}\} = a^2 e^{j\omega_0\tau} E\{e^{-j\underset{\sim}{\bar\omega}\tau}\}$$

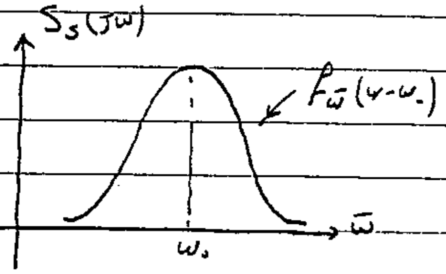
$$= a^2 e^{j\omega_0\tau}\int e^{+j\bar\omega\tau} f_{\bar\omega}(\bar\omega)d\bar\omega = a^2 e^{j\omega_0\tau} F^{-1}\{f_{\bar\omega}(\bar\omega)\} = a^2 e^{j\omega_0\tau}\Phi(\tau)$$

$\bar\omega = -\omega v/c \Rightarrow$ $\qquad$ ↑ characteristic fn. of $\bar\omega$

PSD. $F\{r_s(\tau)\} = S_s(j\omega)$

$\Rightarrow$ $= a^2 F\{e^{j\omega_0\tau}\} f_{\underset{\sim}{\bar\omega}}(\bar\omega) = a^2 f_{\bar\omega}(\omega-\omega_0)$



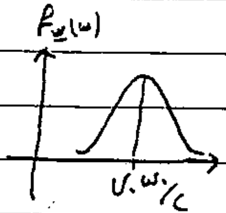


$\bar\omega = \frac{-\omega_0 v}{c}$

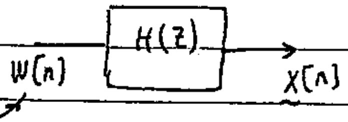
$$f_{\bar\omega}(\omega) = \frac{f_{\underset{\sim}{v}}(\frac{c}{\omega_0}\bar\omega)}{|\frac{c}{\omega_0}|}$$

Ex

$\Rightarrow$



Types of Random Processes:



white noise

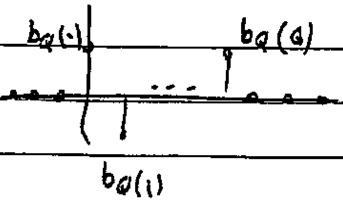
$r_w(k) = \sigma_w^2 \delta(k)$

① H(z); All zero system (Moving Average system)

$H(z) = \sum_{k=0}^{Q} b_Q(k) z^{-k}$ ← FIR with Q+1 Coefficient (Q+1 Taps)

Q^{th} order MA Filter

Causal



For MA: $r_x(k) = \sigma_0^2\, h(k) * h^*(-k)$ / For A.R

$$\underset{\sigma^2}{w} \rightarrow \boxed{h(k)} \rightarrow x$$

$$r_x(k) + \sum_{l=1}^{p} a_l r_x(k-l) = \sigma^2 |b(0)|^2 \delta(k)$$

$$r_x[k] = r_w(k) * h(k) * h^*(-k) = \sigma_w^2 h[k] * h(-k)$$

Ex.

$$w(n) \rightarrow \boxed{h_0 + h_1 z^{-1} + h_2 z^{-2}} \rightarrow x(n)$$

h_0, h_1, h_2 are reals

$$r_x(k) = \sigma_w^2 h(k) * h^*(-k) = \sigma_w^2 \sum_{k'=-\infty}^{+\infty} h^*(-k') h(k-k') = \sigma_w^2 \sum_{k''=-\infty}^{+\infty} h^*(k'') h(k+k'')$$

$k'' = -k'$

$$\therefore \sum_{k'=-\infty}^{+\infty} a_{k'} = a_{-\infty} + \dots + a_{-1} + a_0 + a_1 + \dots + a_{\infty}$$

$k'' = -k'$;

$$r_x(k) = \sigma_w^2 \sum_{k''=-}^{\infty} h^*(k'') h(k+k'') \quad ; \quad r_x(0) = \sigma_w^2 (h_0^2 + h_1^2 + h_2^2)$$

$r_x(-1) = \sigma_w^2 (h_1 h_0 + h_2 h_1)$ $\qquad$ $r_x(1) = \sigma_w^2 (h_0 h_1 + h_1 h_2)$

$r_x(-2) = \sigma_w^2 (h_2 h_0)$ $\qquad$ $r_x(2) = \sigma_w^2 (h_0 h_2)$

$$r_x(k) = \sigma_w^2 h(k) * h(-k)$$

$\longrightarrow$ even Fn.

$$r_x(-k) = \sigma_w^2 h(-k) * h(k)$$

MA process has an autocorrelation which is a non-linear Fn. of filter Coefficents.

$$S_x(e^{j\omega}) = DTFT\{\underbrace{r_x[k]}_{r_w(k) * h(k) * h^*(-k)}\}$$

$$= S_w(e^{j\omega})\, H(e^{j\omega})\, H^*(e^{j\omega}) = S_w(e^{j\omega})\, |H(e^{j\omega})|^2$$

$$S_x(z) = Z\{r_x[k]\} = S_w(z)\, H(z)\, \underbrace{H^*\left(\frac{1}{z^*}\right)}_{Z\{h^*(-k)\}}$$

From Hayes 26 / 11 / 2010

Signal Modeling: (Hayes)

Given $\underset{\underset{LTI}{\uparrow}}{h[n]}$ $\longrightarrow$ Goal: Finding $H(z) = \dfrac{b_0 + b_1 z^{-1} + \cdots + b_Q z^{-Q}}{1 + a_1 z^{-1} + \cdots + a_P z^{-P}}$ Q zeros, P Poles

s.t $\underbrace{Z^{-1}\{H(z)\}}_{\hat{h}(n)} \approx h[n]$

Cost $\Rightarrow \min\{\max\{\underbrace{\hat{h}[n] - h[n]}_{error}\}\}$

Cost $\longrightarrow J_L(\underline{a}, \underline{b}) = \sum_{n=-\infty}^{+\infty} |h(n) - \hat{h}(n)|^L$ L-norm

Pade Approximations:

Pade Approximation matches first $(P+Q+1)$ Samples of $\hat{h}[n]$ and $h[n]$. (Does not say anything about the rest)

Assume that $x[n]$ is the given sequence and $h[n]$ is the designed response.

Ex: $H(z) = \dfrac{b_0}{1 + a_1 z^{-1}}$, $x(0), x[1]$ given (causal sequence)

$h(n) = b_0 (-a_1)^n u(n) \longrightarrow \overset{b_0}{h(0)} = x(0) \Longrightarrow a_1, b_0$ Find

$\underset{\underset{b_0(-a_1)}{\uparrow}}{h(1)} = x(1)$

Pades' method:

Assume $x(n) = h(n) \longrightarrow X(z) = \dfrac{B_Q(z)}{A_P(z)}$ $\begin{matrix} \leftarrow b_0 + b_1 z^{-1} + \cdots + b_Q z^{-Q} \\ \leftarrow 1 + a_1 z^{-1} + \cdots + a_P z^{-P} \end{matrix}$

$\downarrow$

$$X(z) A_P(z) = B_Q(z)$$

$$x(n) * a_P(n) = b_Q(n)$$

$\uparrow$ pade wants equality for $n = \{0, \cdots, Q+P\}$

$b_Q(n) = x(n) * a_p(n) = \sum_{k=0} a_p(k)\, x(n-k) \rightarrow n=0 \quad \sum_{k=0}^{p} a_p(k) x(-k) = \underset{1}{a_p(0)} x(0) + \underset{x(n) \text{ Causal}}{0}$

$= X(0)$

Convolution matrix

$n=1 \quad \sum_{k=0}^{p} a_p(k) x(1-k) = a_p(0) x(1) + a_p(1) x(0) = X(1) + X(0)\, a_p(1)$

$n=Q \quad \sum_{k=0}^{p} a_p(k) x(Q-k) = x(Q) + x(Q-1) a_p(1) + \dots + X(Q-p) a(p)$

1+P Columns

$$\begin{array}{c} \\ n=0 \\ n=1 \\ n=2 \\ \\ n=Q \\ n=Q+1 \\ n=Q+P \\ \\ \end{array}
\begin{bmatrix} X(0) & 0 & \cdots & 0 \\ X(1) & X(0) & 0 \cdots & 0 \\ X(2) & X(1) & X(0) & 0 \cdots 0 \\ & & & \\ X(Q) & X(Q-1) & \cdots & X(Q-P) \\ X(Q+1) & \cdots & & X(Q-P+1) \\ X(Q+P) & \cdots & & X(Q) \\ \vdots & & & \\ X(N) & & & X(N-P) \end{bmatrix}
\begin{bmatrix} 1 \\ a_p(1) \\ \vdots \\ a_p(P) \end{bmatrix} =
\begin{bmatrix} b_Q(0) \\ b_Q(1) \\ \vdots \\ b_Q(Q) \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$Q+P+1 \times 1$

$\leftarrow b_Q(n) \rightarrow n > Q \quad b_Q(n) = 0$

↓ Prony (LS solution of bottom part to find a_p's)

Pade's interest

$$\begin{bmatrix} \underline{U} \\ ---- \\ \underline{\underline{L}} \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ \vdots \\ a_p(P) \end{bmatrix} = \begin{bmatrix} b_U \\ \hline b_L \end{bmatrix} \leftarrow \underline{0}$$

Solve the bottom part first

$$\underline{\underline{L}} \begin{bmatrix} 1 \\ \vdots \\ a_p(P) \end{bmatrix} = \underline{0}$$

$$\Rightarrow \begin{bmatrix} X(q+1) & X(Q) & \cdots \\ & & \\ X(Q+P) & X(Q+P-1) & \cdots \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ a_p(P) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\overset{P+1}{\begin{bmatrix} X(0) & \cdots \\ \vdots & \\ X(q) & \\ X(q+1) & \\ & \\ X(q+p) & \end{bmatrix}} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ \\ a_p \end{bmatrix} = \begin{bmatrix} b_0 \\ \vdots \\ b_q \\ 0 \\ 0 \end{bmatrix}$$

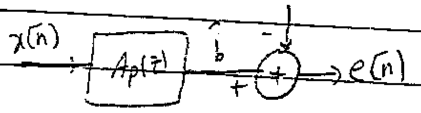
P terms

$$\Rightarrow \begin{bmatrix} X(Q+1) \\ \vdots \\ X(Q+P) \end{bmatrix} + \begin{bmatrix} X(Q) & \cdots & \underset{X(Q+1-P)}{X(Q+P)} \\ \vdots & & \\ X(Q+P-1) & & X(Q) \end{bmatrix} \begin{bmatrix} a_p(1) \\ \vdots \\ a_p(Q) \end{bmatrix} = \bar{0} \quad \Rightarrow a_p\text{'s are found}$$

Substitute a_p in the equation $\underline{\underline{U}} a_p = \underline{b_U} \Rightarrow b_q$'s → get

Prony's method:

Cost: $J^{\text{Prony}}(a_p) = \sum_{n=Q+1}^{\infty} |e(n)|^2$



$E'(z) = X(z) - H(z) \leftarrow \frac{B_Q(z)}{A_p(z)}$

$= X(z) - \frac{B_Q(z)}{A_p(z)} \Rightarrow \underbrace{A_p(z) E'(z)}_{E(z)} = X(z) A_p(z) - B_Q(z)$

$$\Rightarrow e(n) = x(n) * a_p(n) - b_Q(n)$$

$$e(n) = x(n) * a_p(n) \qquad n \geq Q+1 \qquad (b_Q(n) = 0 \quad n \geq Q+1)$$

$$= \sum_{k=0}^{P} a_p(k) x[n-k]$$

$$J^{\text{Prony}}(\underline{a}_P) = \sum_{n=Q+1}^{\infty} \Big(\underbrace{\sum_{k=0}^{P} a_p(k) x(n-k)}_{e(n)} \Big)^2$$

$$\frac{\partial J^{\text{Prony}}}{\partial a_p(\ell)} = 2 \sum_{n=Q+1}^{\infty} e(n) . \underbrace{\frac{\partial e(n)}{\partial a_p^*(\ell)}}_{x^*(n-\ell)} = 0 \qquad \ell = \{1, \ldots, P\}$$

$$\Rightarrow \sum_{n=Q+1}^{\infty} e(n) x(n-\ell) = 0 \qquad \sum_{n=Q+1}^{\infty} \Big(\sum_{k=0}^{P} a_p(k) x(n-k) \Big) x^*(n-\ell) = 0$$

$$\sum_{k=0}^{P} \Big(\sum_{n=Q+1}^{\infty} x(n-k) x(n-\ell) \Big) a_p(k) = 0 \quad : \boxed{r_x(\ell,k) = \sum_{n=Q+1}^{\infty} x(n-k) x^*(n-\ell)}$$

$$\Rightarrow \sum_{k=0}^{P} r_x(\ell,k) \underset{\uparrow\, a_p(0)=1}{a_p(k)} = 0 \qquad \Rightarrow r_x(\ell,0) + \sum_{k=1}^{P} r_x(\ell,k) a_p(k) = 0$$

$$(1 + a_p(1) z^{-1} + \cdots + a_p(P) z^{-P})$$

$$\begin{array}{c} \ell=1 \\ | \\ | \\ | \\ \ell=P \end{array} \underset{P \times P}{\begin{bmatrix} r_x(1,1) & \cdots & r_x(1,P) \\ | & & \\ | & & \\ | & & \\ r_x(P,1) & \cdots & r_x(P,P) \end{bmatrix}} = \underset{P \times 1}{\begin{bmatrix} a_p(1) \\ \vdots \\ \vdots \\ a_p(P) \end{bmatrix}} = - \underset{P \times 1}{\begin{bmatrix} r_x(1,0) \\ r_x(2,0) \\ \vdots \\ \vdots \\ r_x(P,0) \end{bmatrix}}$$

Solv for a_p's from ↑ then insert in top part of Pade's equation system to find $\underline{b}_Q$

$$\begin{bmatrix} r_x(0,0) & r_x(0,1) & r_x(0,2) & \cdots & r_x(0,p) \\ r_x(1,0) & r_x(1,1) & & \cdots & r_x(1,p) \\ r_x(2,0) & & & & \\ \vdots & & & & \\ r_x(P,0) & r_x(P,1) & \cdots & & r_x(P,P) \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ a_p(2) \\ \vdots \\ a_p(P) \end{bmatrix} = \begin{bmatrix} \epsilon_{P,q} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow \text{Hays P-104}$$

$$\begin{bmatrix} \underline{\underline{X}} \\ \hline \underline{\underline{Y}} \end{bmatrix} \begin{bmatrix} 1 \\ \underline{a}_p \end{bmatrix} = \begin{bmatrix} b_Q \\ \hline 0 \end{bmatrix} \qquad \text{Prony}: \quad \underline{\underline{Y}} \begin{bmatrix} 1 \\ a_p \end{bmatrix} = \underline{0} \Rightarrow \begin{bmatrix} \underline{y}_1 & \vdots & \underline{\underline{Y}}_P \\ & \vdots & \end{bmatrix} \begin{bmatrix} 1 \\ a_p \end{bmatrix} = 0$$

First Column of Y & the rest of

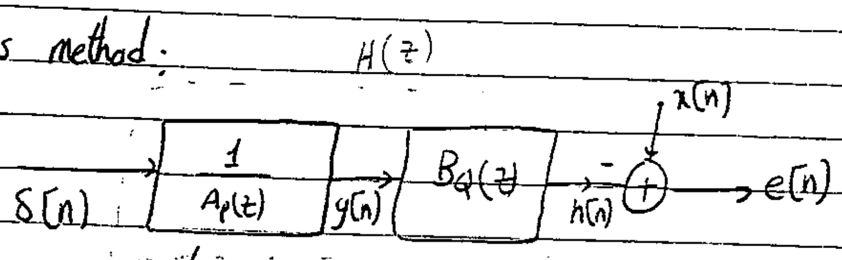
$$\longrightarrow \underset{N\times P}{\underline{\underline{Y}}_P}\, \underset{P\times 1}{\underline{a}_P} = -\underset{N\times 1}{\underline{y}_1} \qquad \Rightarrow \quad \underline{a}_P^{LS} = -\left(\underline{\underline{Y}}_P^T \underline{\underline{Y}}_P\right)^{-1} \underline{\underline{Y}}_P^T \underline{y}_1$$

exactly the same as previously find sol. by differentiation

$$\longrightarrow \begin{bmatrix} r_x(1,1) & r_x(1,2) & \cdots \\ & & \end{bmatrix} \begin{bmatrix} a_p(1) \\ \vdots \\ a_p(P) \end{bmatrix} = - \begin{bmatrix} r_x(1,0) \\ \vdots \\ r_x(P,0) \end{bmatrix}$$

$$\underline{\underline{Y}}_P^T \underline{\underline{Y}}_P \begin{bmatrix} a_p(1) \\ \vdots \\ a_p(P) \end{bmatrix} = -\underline{\underline{Y}}_P^T \begin{bmatrix} x(Q+1) \\ x(Q+2) \\ \vdots \end{bmatrix}$$

Shank's method:



Assume fixed (or calculated by Proney's method)

$$e(n) = -g(n) * b_Q[n] + x(n) = x(n) - \sum_{k'=0}^{Q} b_Q(k') g(n-k')$$

$$J = \sum_{n=0}^{\infty} (|e(n)|)^2 \quad ; \quad \frac{\partial J}{\partial b_Q^*(k)} = \sum_{n=0}^{\infty} e(n) \underbrace{\frac{\partial e^*(n)}{\partial b_Q^*(k)}}_{g^*(n-k)} = 0 \qquad k = \{0, 1, \ldots, Q\}$$

$$J(\underline{b}_Q) = \sum_{n=0}^{\infty} e(n) e^*(n)$$

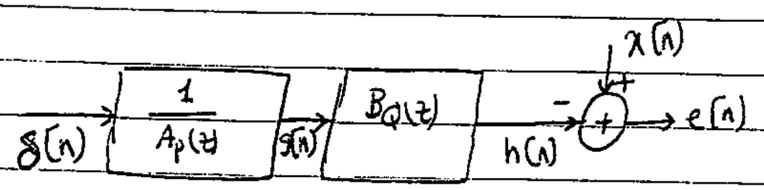
$$\Longrightarrow \sum_{n=0}^{\infty} e(n) g^*(n-k) = 0 \Rightarrow \underbrace{\sum_{n=0}^{\infty} x(n) g^*(n-k)}_{r_{xg}(k,0)} - \underbrace{\sum_{n=0}^{\infty} \left(\sum_{k'=0}^{Q} b_Q(k') g(n-k') \right) g^*(n-k)}_{\sum_{k'=0}^{Q} \underbrace{\sum_{n=0}^{\infty} g(n-k') g^*[n-k]}_{r_g(k,k')} b_Q(k')}$$

$$r_{xg}(k,0) - \sum_{k'=0}^{Q} r_g(k,k')\, b_Q(k') = 0$$

$$\begin{array}{l} k=0 \rightarrow \\ k=1 \rightarrow \\ \\ \\ k=Q \rightarrow \end{array} \begin{bmatrix} r_g(0,0) & r_g(0,1) & \dots & r_g(0,Q) \\ r_g(1,0) & r_g(1,1) & \dots & r_g(1,Q) \\ | & & & | \\ . & & & \\ r_g(Q,0) & & - \; - & r_g(Q,Q) \end{bmatrix} \begin{bmatrix} b_Q(0) \\ \vdots \\ \\ b_Q(Q) \end{bmatrix} = \begin{bmatrix} r_{xg}(0,0) \\ | \\ | \\ r_x(Q,0) \end{bmatrix}$$

Solve this to find $\underline{b}_Q$ in LS sense. (minimizing $\sum_{n=}^{\infty} (e(n))^2$)

Revisiting the same sol. through matrix algebra



$g(n) * b_Q(n) \simeq x(n)$ (everything is causal)

$$\begin{array}{c} n=0 \\ \\ \\ \downarrow \\ \infty \end{array} \underbrace{\begin{bmatrix} g(0) & 0 & 0 & \dots & 0 \\ g(1) & g(0) & \dots & & 0 \\ & & & & \\ g(N) & \dots & & & g(N-Q) \end{bmatrix}}_{\underline{\underline{G}}} \underbrace{\begin{bmatrix} b_Q(0) \\ b_Q(1) \\ \vdots \\ b_Q(Q) \end{bmatrix}}_{\underline{b}} = \underbrace{\begin{bmatrix} X(0) \\ X(1) \\ \vdots \\ X(N) \end{bmatrix}}_{\underline{\underline{X}}} \Rightarrow \underline{\underline{G}}\,\underline{b} = \underline{\underline{X}}$$

$$\rightarrow \underline{b}^{LS} = (\underline{\underline{G}}^H \underline{\underline{G}})^{-1} \underline{\underline{G}}^H \underline{X}$$

$$\Rightarrow (\underline{\underline{G}}^H \underline{\underline{G}})\, \underline{b}^{LS} = \underline{\underline{G}}^H \underline{X}$$

$$\begin{bmatrix} r_g(0,0) & r_g(0,1) & \dots & r_g(0,Q) \\ . & & & \\ & & & \\ & - & - & . \end{bmatrix} \underline{\underline{b}}^{LS} = \begin{bmatrix} \sum_{k=}^{\infty} g(k)X(k) \\ \sum_{k=}^{\infty} g(k)X(k+1) \\ \vdots \\ \underbrace{\sum_{k=}^{\infty} g(k)X(k+Q)}_{r_{xG}(Q,0)} \end{bmatrix} =$$

All-pole modeling

$$H(z) = \frac{b_0}{1 + a_1 z^{-1} + \dots + a_p z^{-p}} \quad ; \quad \begin{matrix} P \text{ poles} \\ \text{no zeros} \end{matrix}$$

$$J_{prony} = \sum_{n=Q+1}^{\infty} |e(n)|^2$$

$$e(n) = a_p(n) * x(n) - b_Q(n) = \begin{cases} x(n) + \sum_{k=1}^{P} a_p(k) x(n-k) - b_Q(n) & n = \{0, \dots, Q\} \\ x(n) + \sum_{k=1}^{P} a_p(k) x(n-k) & n \geq Q+1 \end{cases}$$

For all-pole modeling $\rightarrow Q = 0$ $\quad J_{prony} = \sum_{n=1}^{\infty} |e(n)|^2$

$$J'_{prony} = \sum_{n=1}^{\infty} |e(n)|^2 + |e(0)|^2 \quad \Big| \; \underset{\uparrow \text{does not depend on } a_p}{e(0) = x(0) - b_0}$$

Minimizing J_{proney} or $J'_{proney}(a_p)$ give the same optimal set of a_p coeff.

$$J'_{prony} = \sum_{n=0}^{\infty} |e(n)|^2$$ will use $J'_{p..}$ for all-pole modeling

$$r_x(k,l) = \sum_{k=Q+1}^{\infty} x(n-l) x^*(n-k)$$

In classical prony ; $J = \sum_{k=Q+1}^{\infty} (\dots) \longrightarrow r_x(k,l) = \sum_{k=Q+1}^{\infty} (x(\dots))(x(\dots))$

" Primed " ; $J' = \sum_{k=0}^{\infty} (\dots) \longrightarrow r_x(k,l) = \sum_{n=0}^{\infty} x(n-l) x^*(n-l)$

$$r_x(k,l) = \sum_{n=0}^{\infty} x(n-l) x^*(n-k)$$

$$\downarrow$$

$$r_x(k+\Delta, l+\Delta) = r_x(k,l) \Rightarrow \sum_{n=0}^{\infty} x(n-\Delta-l) x(n-\Delta-k) = \sum_{n=-\Delta}^{\infty} x(n-l) x(n-k)$$

$$= \sum_{n=0}^{\infty} x(n-l) x(n-k)$$

Causal $\quad l > 0, \; k > 0$

$\Leftarrow$ for negative argument if is zero

$r_x(k-l) = \sum_{n=0}^{\infty} x(n-l)x(n-k) \triangleq r_x(k,l)$ ← deterministic auto correlation

$$\begin{bmatrix} r_x(0) & r_x^*(1) & \cdots & r_x^*(p-1) \\ r_x(1) & r_x(0) & & \\ & & \ddots & \\ r_x(p-1) & & & r_x(0) \end{bmatrix} \begin{bmatrix} a_p(1) \\ \\ \\ a_p(p) \end{bmatrix} = \begin{bmatrix} r_x(1) \\ r_x(2) \\ \vdots \\ r_x(p) \end{bmatrix}$$

↑ all-pole modeled $x(n)$ satisfied above

↑ Teoplitz & Hermition matrix

Modeling with finite data records:

Previously we have assumed that $x(n)$ is given for all $n \geq 0$

In practice $x(n)$ is given for $n = 0, \ldots, N$

① Auto-Correlation Method:

Assumption: $x(n) = 0$ whenever it is not avilable.

All-pole modeling $r_x(k-l) = \sum_{n=0}^{\infty} x(n-l)x^*(n-k) = \begin{cases} \sum_{n=0}^{N} x(n-l)x^*(n-k) & k=0,\ l=0 \\ \sum_{n=1}^{N} x(n)\, x^*(n-1) & k=1,\ l=0 \end{cases}$

Advantage: of

$\underline{\underline{R}}\, \underline{a_p} = -r_x$

① $\underline{\underline{R}}$: Teoplitz Hermition ← Fast method to sol. of such sys.

② The resulting all-pole system is guaranted to be stable. (Proof in Hsys)

Disadvantges: Possible mismatch between true $x(n)$ and assumed $x(n)$ for $n > N$

② Covariance Method: Do not any assumption and calculate $r_x(k-l)$ Using only available data:

$$r_x(k,l) = \sum_{n=p}^{N} x(n-l)x^*(n-k) \qquad l,k \in \{1, —, p\}$$

$$\begin{bmatrix} r_x(1,1) & r_x(1,2) & \cdots & r_x(p,1) \\ \vdots & & & \\ r_x(p,1) & \cdots & & r_x(p,p) \end{bmatrix} \begin{bmatrix} a_p(1) \\ \vdots \\ a_p(p) \end{bmatrix} = \begin{bmatrix} r_x(1,0) \\ \vdots \\ r_x(p,0) \end{bmatrix}$$

Advantages: Better modeling (without any assumptions)

Dis ~ : all-pole filter is not gouvaranties to be stable.

All-pole model

$e[n] = a_p[n] * x[n] - b_q(n) \qquad Q=0$ for all-pole

$$e[n] = a_p(n) * x[n] = [\overset{b_0}{1} \;\; a_p(1) \; --- \; a_p(p)] * x[n]$$

$$\begin{array}{l} n=1 \\ n=2 \\ \\ \\ \\ \end{array} \begin{bmatrix} x(1) & x(0) & \cdots & 0 \\ x(2) & x(1) \; x(0) & — & 0 \\ & & & \\ & & & \\ x(N) & x(N-1) & -- & x(N-p) \end{bmatrix} \begin{bmatrix} 1 \\ a_p(1) \\ \vdots \\ a_p(p) \end{bmatrix} = \underline{0}$$

$x(-1) \rightarrow$ assumption

$$\begin{array}{c|cccc} & X(0) & 0 & \cdots & 0 \\ T & X(1) & X(0) & & \\ & X(2) & X(1) & X(0) & \\ & \vdots & & & \\ & \underline{X(p-1)} & \underline{X(p-2)} & \underline{---} & \underline{X(0)} \\ M_{min} & X(N-1) & --- & & X(N-p) \\ & X(N) & -- & & X(N-p+1) \\ 3 & X(N+1) & & & \end{array} \begin{bmatrix} a_p(1) \\ \vdots \\ a_p(p) \end{bmatrix} = - \begin{bmatrix} \cancel{X(1)} \\ \cancel{X(2)} \\ \vdots \\ \overline{X(p)} \\ \vdots \\ \underline{X(N)} \\ X(N+1) \end{bmatrix}$$

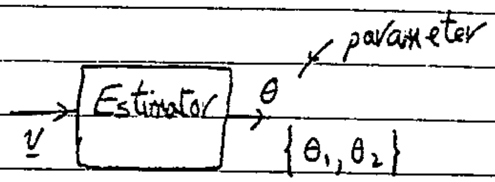
53

07 / 12 / 10

$\underline{\underline{A}}\,\underline{a}_p = \underline{b}$ ① Auto-Correlation ; uses whole $\underline{\underline{A}}$ matrix ($\underline{\underline{T}}, \underline{\underline{M}}, \underline{\underline{B}}$) and finds Least-square sol. of $\underline{\underline{A}}\,\underline{a}_p = \underline{b}$

② Covariance Use only middle part and solves $\underline{\underline{M}}\,\underline{a}_p = \underline{b}_m$ in LS sense

Estimation:



$\underline{X}(n) = \underline{\underline{A}}\,\underline{X}(n) + \underline{\underline{B}}\,\underline{U}(n) + C w(n)$

state

input deterministic

$\underline{g}(n) = \underline{\underline{C}}\,\underline{X}(n) + \underline{V}(n)$ ← state Est.

observation

Parameter Estimation:

Parameter can be non-random	or Random
θ deterministic	θ Bayesin Est
	θ has an a priori pdf ?
Maximum likelihood (ML)	
least square sol. LS	minimum mean square error Est. (MSE)
	minimum Absolute error ~ mean
min. Var. Unbiased Est.	min. Var. Unbiased Est.
Bounds: (Estimation error)	Bounds:
Lower Bounds on Estimation Error	Bayesin Bounds, Payesin
Cramer-Rao Bound and other bounds	Cramer-Rao

Ex: ML Estimation: (Non-random parameter)

$$\underline{r} = C + \underline{w} \rightarrow N(0, \sigma_w^2 \underline{\underline{I}}) \qquad \underline{r} = \begin{pmatrix} r(1) \\ r(2) \\ \vdots \\ r(N) \end{pmatrix}$$

C ← non-random unknown constant

$$f_{\underline{r}}(r; \overset{c}{\theta}) \Rightarrow N(C\underline{1}, \sigma_w^2 I) \qquad \underline{1} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

$$\hat{\theta} = \text{~~Max~~} \arg\max_{\theta} \underbrace{f_{\underline{r}}(\underline{r}; \theta)}_{\text{likelihood fn.}} = \arg\max_{\theta} \underbrace{\log(f_{\underline{r}}(\underline{r}; \theta))}_{\text{log-likehood}}$$

$$f_{\underline{r}}(\underline{r}; \theta) = \left(\frac{1}{\sqrt{2\pi}}\right)^N \cdot \underbrace{\frac{1}{(\sigma_w^2 \underline{\underline{I}})^{N/2}}}_{\alpha} e^{-(\underline{r} - \underline{1}\theta)^T (\sigma_w^2 I)^{-1} (\underline{r} - \underline{1}\theta)}$$

$$\log_2 f_r(\underline{r}; \theta) = \ln f_r(r; \theta) = \ln(\alpha) - \frac{1}{\sigma_w^2} \underbrace{(\underline{r} - \underline{1}\theta)^T (r - \underline{1}\theta)}$$

$$\Rightarrow \frac{d}{d\theta} \ln f_r(r; \theta) = -\frac{1}{\sigma_w^2}(-2\underline{1}^T r + 2N\theta) = 0 \qquad = r^T r - \underline{1} r \theta - r^T \underline{1} \theta + \underbrace{\underline{1}^T \underline{1}}\theta^2$$

$$= r^T r - 2 \underline{1}^T r \theta + N\theta^2$$

$$\Rightarrow \underset{\substack{\uparrow \\ \text{satisfies} \\ \text{ML eq.}}}{\hat{\theta}} = \frac{\underline{1}^T r}{N} = \frac{1}{N} \sum_{k=1}^{N} r(k)$$

Ex: $$\underline{r} = \underset{\substack{\uparrow \\ N(0, \sigma_c^2)}}{C} + \underline{w} \leftarrow N(0, \sigma_w^2 \underline{\underline{I}})$$ (Random par. Est.)

$$r_n = r(n) = C + w(n)$$

$$\underset{\substack{\uparrow \\ \text{Estimator}}}{\hat{C}} = \sum_{k=1}^{N} h_k r_k \quad , \quad E\{(C - \hat{C})^2\} \leftarrow (\text{MSE}) \rightarrow \text{minimized}$$

$$J(\underline{h}) = E\left\{(C - \sum h_k r_k)^2\right\} \quad ; \quad \underset{\substack{\downarrow \\ k = \{1, \dots, N\}}}{\frac{\partial J}{\partial h_k}} = 0 \Rightarrow E\left\{\frac{\partial}{\partial h_k}(C - \sum_{k'=1}^{N} h_{k'} r_{k'})^2\right\} = 0$$

$$E\left\{-2\left(c-\sum_{k'=1}^{N} h_{k'} r_{k'}\right) r_k\right\} = 0 \Rightarrow E\{c\, r_k\} = \sum_{k'=1}^{N} h_{k'}\, E\{r_{k'} r_k\}$$

$\underset{\sim}{c}$ and $\underset{\sim}{w}$ are indp.

$$= E\{c^2\} + E\{c\,w\} = \sum_{k'=1}^{N} h_{k'}\, E\{r_{k'} r_k\}$$

$$E\{r_k r_{k'}\} = E\{(c+w_{k'})(c+w_k)\} = E\{c^2\} + E\{w_{k'} w_k\} = \begin{cases} \sigma_c^2 & k' \neq k \\ \sigma_c^2+\sigma_w^2 & k'=k \end{cases}$$

$$= \sigma_c^2 + \sigma_w^2\, \delta[k-k']$$

$$\begin{array}{l} \frac{\partial J}{\partial h_k} = 0 \Rightarrow \\ \\ \frac{\partial J}{\partial h_N} = 0 \longrightarrow \end{array} \begin{bmatrix} \sigma_c^2+\sigma_w^2 & & & & \\ \sigma_c^2 & \sigma_c^2 & \cdots & \sigma_c^2+\sigma_w^2 \cdots & \sigma_c^2 \\ & & & & \\ \sigma_c^2 & \sigma_c^2 & \cdots & & \sigma_c^2+\sigma_w^2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_k \\ \vdots \\ h_N \end{bmatrix} = \begin{bmatrix} \\ \sigma_c^2 \\ \\ \end{bmatrix}$$

For $N=2 \Rightarrow \begin{bmatrix} \sigma_c^2+\sigma_w^2 & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2+\sigma_w^2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} \sigma_c^2 \\ \sigma_c^2 \end{bmatrix}$

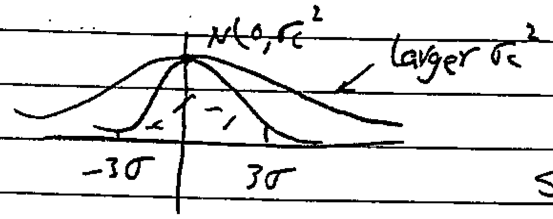
$$\Rightarrow \begin{bmatrix} 1+\frac{1}{SNR} & 1 \\ 1 & 1+\frac{1}{SNR} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \longrightarrow h_1 = h_2 = \frac{1}{\left(1+\frac{1}{SNR}\right)+1}$$

$$= \frac{1}{2+1/SNR}$$

$$SNR = \frac{\sigma_c^2}{\sigma_w^2} = \frac{E\{c^2\}}{E\{w^2\}}$$

$$\boxed{\hat{c} = \frac{1}{2+1/SNR}(r_1+r_2)} \leftarrow \text{MSE sense optimal Est. for } \underset{\sim}{c}$$

$SNR \rightarrow \infty \Rightarrow \hat{c} \rightarrow$ ML Est. $\qquad \hat{c}_{MSE} \rightarrow \hat{c}_{ML}$



So if $\sigma_c^2 \rightarrow \infty$ then the problem ? becomes non-random pr. Est.

Estimation:

① non-random par. est. (ML, LS, etc...)

② Random " " (MSE, wiener Filter, LMin MSE)
?

Properties of Estimator:

① Bias (θ: parameter to be estimated)

$E\{\hat{\theta}\} = E\{\theta\} \rightarrow$ Unbiased

From ex. $E\{C_{ML}\} = \underset{\uparrow \text{non-random}}{C} \longrightarrow$ Unbiased

$$\underset{c,w}{E}\{\hat{C}_{LMMSE}\} = E\left\{\frac{1}{N+\frac{1}{SNR}} \underset{E(x)=0}{\sum X_k}\right\} = 0 = E\{\underset{\uparrow N(0,\sigma_c^2)}{C}\} \rightarrow \text{Unbiased}$$

② Consistency : $\epsilon = \theta - \hat{\theta} \rightarrow \underbrace{E\{\epsilon^2\}}_{\text{MSE (mean square error)}} \xrightarrow[N\to\infty]{} 0 \Rightarrow$ estimator is Consistent

Gaussian Linear Combination Gaussian

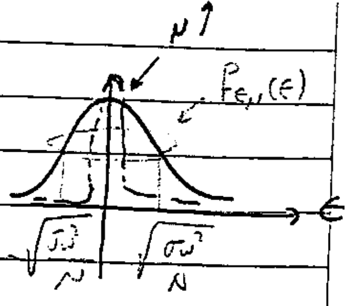
Ex. $\hat{C}_{ML} = \frac{1}{N}\sum_{k=1}^{N} X_k$

$$\epsilon = \hat{C}_{ML} - C_{ML} = \frac{1}{N}\sum_{k=1}^{N} X_k - \frac{\overset{\sum_{k=1}^{N}1}{N}}{N} C = \frac{1}{N}\left(\sum_{k=1}^{N} X_k - \sum_{k=1}^{N} C\right)$$

$$= \frac{1}{N}\left(\sum_{k=1}^{N} X_k - C\right) = \frac{1}{N}\sum_{k=1}^{N} w_k \Rightarrow E(\epsilon^2) = \frac{1}{N^2} E\left\{\left(\sum_{k=1}^{N} w_k\right)^2\right\}$$

$\because$ w_k's are i.i.d
zero mean

$$= \frac{1}{N^2}\sum_{k=1}^{N}\sigma_w^2 = \frac{\sigma_w^2}{N}$$



$\hat{C}_{ML}$ is also consistent!

$$\sum_{K=1}^{N} X_K = X_1+\dots+X_N = 1^T \underline{X} = [1 \cdots 1]_{1\times N} \begin{pmatrix} X_1\\ \vdots \\ X_N\end{pmatrix}_{N\times 1} = (X_1+X_2+\dots+X_N)$$

③ Efficiency: A non-random par. est. is called efficient if its error variance is equal to Crame Rao Lower band.

(Similar bounds and efficiency concepts also available for random par. est.)

Estimation of the mean value of random process from a Single realization:

$x(n) \Rightarrow$ WSS random process ($\mu_X(n)$ is Constant)

$$\hat{\mu}_X(n) = \frac{1}{N}\sum_{K=1}^{N} X(k) \longrightarrow \text{Unbiased}$$

Consistency ; $\epsilon = \hat{\mu}_X - \mu_X = \frac{1}{N}\sum_{k}(X_k - \hat{\mu}_X) = \frac{1}{N}\underline{\underline{1}}^T \cdot \underline{X}_m$; $X_m = \begin{bmatrix} X_1-\mu_X \\ X_2-\mu_X \\ \vdots \\ X_N-\mu_X \end{bmatrix}$

$$\epsilon^2 \longrightarrow E\{\epsilon^2\} = \frac{1}{N^2} E\{(\underline{\underline{1}}^T \cdot \underline{X}_m)(\underline{X}_m^T \cdot \underline{\underline{1}})\}$$

$$E\{(X_k-\mu_X)(X_m-\mu_X)\}$$

$$= \frac{1}{N^2}\, \underline{1}^T \underbrace{E\{\underline{X}_m \underline{X}_m^T\}}_{\substack{C_X \\ Cov(X)}} \underline{1} = \frac{1}{N^2}\, \underline{1}^T \begin{bmatrix} C(1,1), C(1,2) \dots C(1,N) \\ \vdots \\ C(K,1) \cdots C(K,m) \\ \vdots \\ C(N,1) \qquad C(N,N) \end{bmatrix} \underline{\underline{1}}$$

$$= \frac{1}{N^2}\, \underline{\underline{1}}^T \begin{bmatrix} C(0) & C(1) & \cdots & C(N-1) \\ C(1) & C(0) & & C(N-2) \\ C(2) & & \ddots & \vdots \\ \vdots & & \ddots & \vdots \\ C(N-1) & & & C(0) \end{bmatrix} \cdot \underline{\underline{1}}$$

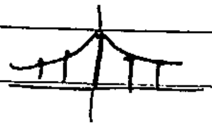
$$\Rightarrow E\{\epsilon^2\} = \frac{1}{N^2}\left(NC(0) + 2(N-1)C(1) + 2(N-2)C(2) + \dots + 2C(N-1)\right)$$

$$= \frac{1}{N}\sum_{K=-(N-1)}^{N-1}\left(1 - \frac{|l|}{N}\right) C(l) \quad ; \quad C(-l) = C(l)$$

Iff $\frac{1}{N}(\dots) \xrightarrow[\text{as } N\to\infty]{} 0 \overset{\text{iff}}{\longleftrightarrow} \hat{\mu}_X$ is Consistent est.

A sufficient Cond. If $(C_x(\ell) \to 0)$ $\longrightarrow$ Consistent Est.
as $\ell \to \infty$

$C_x(\ell) = \alpha^{|\ell|}$



$\Rightarrow$ Ergodicity (Time average = Ensemble average)

Processes with $\hat{\mu}_x = \frac{1}{N} \sum X_t \longrightarrow \mu_x$ is called ergodic in the mean

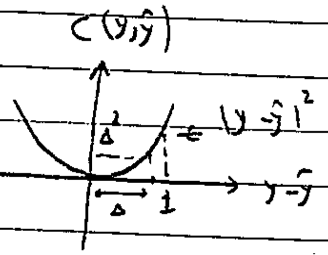
[illegible]

There is ~~a~~ ergodocity in auto-correlation

Bayes Estimation of Random variables:

$\underline{X} \leftarrow$ random vector
$\quad\nwarrow$ observation ; $y \leftarrow$ would like to estimate
$\quad\nwarrow$ random

$C(y, \hat{y}) \to$ Cost fn.



$$J = \underset{\underline{x},y}{E}\{C(y,\hat{y})\} = \iint_{\underline{x}\ y} c(y, f(\underline{x}))\, \overset{\nearrow\, f_{\underline{x},y}(y/\underline{x})\, f_{\underline{x}}(\underline{x})}{f_{\underline{x},y}(\underline{x},y)}\, dy\, d\underline{x}$$

$$= \int_{\underline{x}} \Big(\underbrace{\int c(y, \hat{f}(x))\, f_{\underline{x},y}(y/x)\, dy}_{I(\hat{y}) \to f(\underline{x})} \Big) f_{\underline{x}}(x)\, dx$$

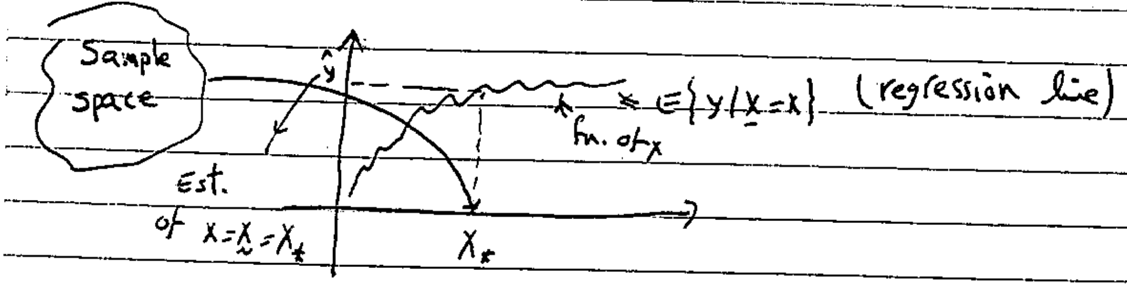
$J = \int_{-\infty}^{+\infty} I(f(x))\, f_{\underline{x}}(x)\, dx$, Minimizing J with a proper choice of $f(x)$ is equivalent to minimizing $I(f(x))$ when $f(x)$ is Fixed.

?

$$I(f(x)) = \int_{-\infty}^{+\infty} (y-\underset{\underset{f(\underline{x})}{\uparrow}}{\hat{y}})^2 f_{\underline{x},y}(y/x)\,dy \Rightarrow \frac{\partial I(f(x))}{\partial \hat{y}} = 2\int_{-\infty}^{+\infty} (y-\hat{y})\, f(y/\underline{x})\,dy = 0$$

$$\Rightarrow \int y\, f(y/x)\,dy = \int \underset{\underset{\text{~~not~~ fn of } x}{\downarrow}}{\hat{y}}\, f(y/\underline{x})\,dy = \hat{y}\underbrace{\int f(y/x)\,dy}_{1} = \hat{y}$$

$$\Rightarrow \boxed{\hat{y} = E\{y/\underline{x}\}} \leftarrow \text{Conditional mean Est.}$$



Note: ① $E\{y/\underset{\sim}{x}=x\}$ calculation requires joint pdf of y and x

$$(x,y) \rightarrow \underset{\sim}{y}/\hat{x} = \underset{\sim}{z} \qquad \frac{f(x,y)}{f(x)}$$

② $\underset{x}{E}\{\underset{y}{E}\{y/\underline{x}=x\} = \underset{x,y}{E}\{y\} \longrightarrow$ Unbiased Est. $\leftarrow E\{\hat{y}\} = E\{y\}$

③

$$\begin{tikzcd} & y_1 \\ \underline{x} \arrow[ur] \arrow[dr] & \\ & y_2 \end{tikzcd}$$

to be estimated

$$J = \sum_{k=1}^{2} (y_k - \hat{y}_k)^2 = (y - \underset{\underset{f_1(x)}{\uparrow}}{\hat{y}_1})^2 + (y_1 - \underset{\underset{f_2(x)}{\uparrow}}{\hat{y}_2})^2$$

So, $f_1(x) = E\{y_1/\underline{x}\}$

$f_2(x) = E\{y_2/\underline{x}\}$

$$\rightarrow \underline{\hat{y}} = E\{\underline{y}/\underline{x}\}$$

$$E\left\{\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} / \underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix}\right\}$$

④ Orthogonality results: $\underline{\hat{y}} = E\{\underline{y}/\underline{x}\}$ then $E\{\underbrace{(\underline{\hat{y}}-\underline{y})}_{\underline{\epsilon}}(g(\underline{x}))^T\} = \underline{0}$ For any $g(x)$

$$\rightarrow E\left\{\begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_m \end{bmatrix}[g_1(x) \;-\!-\; g_m(x)]\right\} = \underline{0}$$

$E\{g(x,y)\} = E_X\{E_y\{g(x,y)/x = x_1\}\}$ ← Iterated expectation

___/___/___

$$E\{\hat{\underline{y}}\,\underline{g}^T(x)\} \overset{?}{=} E\{\underline{y}\,\underline{g}^T(x)\} = E_x\{E_y\{\underline{y}\,\underline{g}^T(x)/\underline{x}\}\} = E_x\{E\{y/x\}\,\underline{g}^T(x)\}$$

$$= E_X\{\hat{\underline{y}}\;g^T(x)\}$$

Converse:

$$E\{\overbrace{(\underline{y}-\hat{\underline{y}}(x))}^{\epsilon}\,g^T(\underline{x})\} = \underline{0} \Rightarrow \hat{\underline{y}}(x) = E\{\underline{y}/\underline{x}\}$$

if error orthogonal with observation ⇒ Estimator is $\hat{y}(x)$ (mean-Conditional Est.)

See P.59 for Proof

Ex: $\underset{\wedge}{x}$ and $\underset{\wedge}{y}$ Jointly Gaussian, find $f_{y/x}(y/x) = ?$

$$f_{x,y}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-r^2}}\,e^{-\left(-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_x^2} - \frac{2r\sigma xy}{\sigma_x\sigma_y} + y^2/\sigma_y^2\right)\right)}$$

$$*\;\therefore \frac{1}{(\sqrt{2\pi})^2\,|C|^{1/2}}\,e^{-1/2\,(x\;\;y)\begin{pmatrix}\sigma_x^2 & r\sigma_x\sigma_y\\ r\sigma_x\sigma_y & \sigma_y^2\end{pmatrix}^{-1}\begin{pmatrix}x\\ y\end{pmatrix}} \quad \& \quad |C| = (\sigma_x\sigma_y)^2 - (r\sigma_x\sigma_y)^2 = (\sigma_x\sigma_y)^2(1-r^2)$$

$$\frac{f_{x,y}(x,y)}{f_x(x)} = ? = \frac{C_1}{C_2}\,\frac{e^{-\left(\frac{(y - r\frac{\sigma_y}{\sigma_x}x)^2}{2\sigma_y^2(1-r^2)} + x^2/2\sigma_x^2\right)}}{e^{-x^2/2\sigma_x^2}}$$

$$\frac{1}{\sqrt{2\pi\sigma_x^2}}\,e^{-\frac{1}{2}x/\sigma_x^2}$$

$$f_{y/x} \longrightarrow N\left(\underbrace{r\frac{\sigma_y}{\sigma_x}x}_{\text{Conditional mean}},\;\sigma_y^2(1-r^2)\right)$$

$$\hat{y} = E\{y/x\} = \underset{\substack{\uparrow\\ \text{Correlation Coefficient of } x \& y}}{r}\,\frac{\sigma_y}{\sigma_x}x \;\leftarrow \text{linear fn. (for Gaussian dist.)}$$

In general: $E\{y/x\}$ is linear fn. of observation

$$* \to \left(\frac{1}{(\sqrt{2\pi})^2\,\sigma_x\sigma_y\sqrt{1-y^2}}\right)e^{-1/2\,\frac{1}{(\sigma_x\sigma_y)^2(1-r^2)}\left(x^2\sigma_y^2 - 2rxy\sigma_x\sigma_y + y^2\sigma_x^2\right)} \quad \& \;\sigma_x^2 \text{ [illegible]}$$

$$\frac{-1}{2}\,\frac{1}{\sigma_y^2(1-r^2)}\left\{\left(\frac{x\sigma_y}{\sigma_x}\right)^2 - \frac{2rxy\sigma_y}{\sigma_x} + y^2\right\} = \frac{-1}{2}\,\frac{1}{\sigma_y^2(1-r^2)}\left\{\left(y - \frac{rx\sigma_y}{\sigma_x}\right)^2 + \left(\frac{x\sigma_y}{\sigma_x}\right)^2 - \left(\frac{rx\sigma_y}{\sigma_x}\right)^2\right\}$$

$$= \frac{-1}{2}\,\frac{1}{\sigma_y^2(1-r^2)}\left\{\left(y - \frac{rx\sigma_y}{\sigma_x}\right)^2 + \left(\frac{x\sigma_y}{\sigma_x}\right)^2(1-r^2)\right\} = -\left\{\frac{\left(y - \frac{rx\sigma_y}{\sigma_x}\right)^2}{2\sigma_y^2(1-r^2)} + \frac{x^2}{2\sigma_x^2}\right\}$$

$\hat{y} = E\{y/\underset{\sim}{X}=x\}$ ← Conditional

optimal in MSE

Linear MSE Estimation:

$E\{y|x\}$ is a non-linear fn and

Then as an alternative the estim

and parameters of this structure is

$$\hat{y} = \underset{\text{weight}}{W_1}\overset{\text{observations}}{X_1} + W_2X_2 + \cdots + W_NX_N =$$

Ex. (Sec. 3.2.6 of Hayes)

$\hat{y} = ax+b$, Goal: minimize $\underset{a,b}{E}\{(y$

$J = E\{e^2\} = E\{(y-\hat{y})^2\}$, $\frac{\partial J}{\partial a}=0$,

$\Rightarrow E\{2e.\frac{\partial e}{\partial a}\}=0 \Rightarrow E\{e.x\}=0$

$\frac{\partial J}{\partial b}=0 \Rightarrow E\{2e(-1)\}=0 \Rightarrow E\{e\}=0$

$$\begin{bmatrix} E\{x^2\} & E\{x\} \\ E\{x\} & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{Bmatrix} E(xy) \\ E\{y\} \end{Bmatrix} \Rightarrow$$

From 2^{nd} Eq. $\Rightarrow b = m_y - m_x a \Rightarrow \hat{y}$

$(\sigma_x^2+m_x^2)a + m_x(m_y - m_x a) = \overset{E\{xy\}}{m_y} \Rightarrow a=$

$$= \frac{\overset{\text{Correlation Coefficient}}{\rho}\sigma_x\sigma_y}{\sigma_x^2} = \rho\frac{\sigma_y}{\sigma_x} \Rightarrow a = \rho\frac{\sigma_y}{\sigma_x}$$

$$\hat{y} = \rho\frac{\sigma_y}{\sigma_x}(x-m_x)+m_y$$

Note that the last formula on $\hat{y}$ is exactly equal to $E\{x|y\}$ for the Gaussian dis.
$\llcorner m_x = m_y = 0$

$$J_{min} = E\{(y-\hat{y})^2\}$$

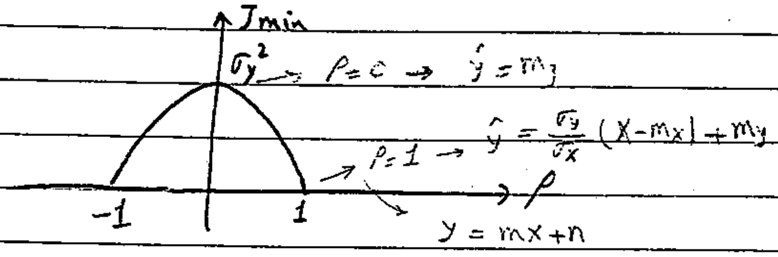
$$J_{min} = E\{e_{opt}(y-\hat{y})\} = E\{e_{opt}\, y\} - E\{e_{opt}(a_{opt}x + b_{opt})\}$$

$$= E\{e_{opt}\, y\} - a_{opt}\underbrace{E\{e_{opt}X\}}_{0} - b_{opt}\cancel{E\{e_{opt}\}} = E\{(y-(a_{opt}X + b_{opt}))y\}$$

$$= E\{y^2\} - a_{opt}E\{xy\} - b_{opt}E\{y\} = \sigma_y^2 + m_y^2 - \rho\frac{\sigma_y}{\sigma_x}(\rho\sigma_x\sigma_y + m_x m_y)$$

$$-(m_y - m_x\rho\frac{\sigma_y}{\sigma_x})m_y$$

$$= \sigma_y^2 - \rho^2\sigma_y^2 = \sigma_y^2(1-\rho^2)$$



$\sigma_y^2 \rightarrow \rho = 0 \rightarrow \hat{y} = m_y$

$\rho = 1 \rightarrow \hat{y} = \frac{\sigma_y}{\sigma_x}(X - m_x) + m_y$

$y = mx + n$

Multiple observation for the estimation of y:

$\underline{X}$: observation vector $\Rightarrow \hat{y} = w_1x_1 + w_2x_2 + \dots + w_Nx_N = \underline{w}^T\underline{X}$

$J(\omega) = E\{(y-\hat{y})^2\}$, $\frac{\partial J(\underline{\omega})}{\partial \omega_k} = 0 \quad K = \{1, -, N\}$

$\frac{\partial}{\partial w^T}$

$$\underline{\nabla}_\omega J = \begin{bmatrix} \frac{\partial J}{\partial \omega_1} \\ \vdots \\ \vdots \\ \frac{\partial J}{\partial \omega_N} \end{bmatrix} = \underline{0} \quad ; \nabla_\omega J(\omega) = E\{\nabla_\omega e^2\} = E\{2e\nabla_\omega e\} ; \; e = y - \underline{w}^T\underline{X}$$

$$= E\{2e(-\underline{X})\} \Rightarrow \nabla_\omega J = 0 \Rightarrow E\{e\underline{X}\} = 0$$

orthogonality

$$E\{e\underline{X}\} = 0 \rightarrow E\{(y - w^T\underline{X})\underline{X}\} = 0 \Rightarrow E\{\underline{X}y\} - E\{\underline{X}(\underline{X}^T\omega)\} = 0$$

↑ optimality condition for $\underline{\omega}$

$$\Rightarrow E\{\underline{X}\underline{X}^T\}\omega = E\{y\underline{X}\} \Rightarrow \boxed{\underline{\underline{R}}_X\, \omega = r_{y\underline{X}}}$$

↑ cross corr. vector between each obs. and desied

$$\Rightarrow \underline{w} = \underline{\underline{R}}_x^{-1} r_{y\underline{x}} \quad , \quad \underline{w} = R_x^{-1} \underset{\nwarrow \text{ desired Quantity}}{r_{d\underline{x}}}$$

$$J_{min}(\underset{\uparrow\, R_x^{-1} r_{yx}}{w_{opt}}) = E\{e^2\} = E\{e(y - \underline{w}_{opt}^T \underline{x})\} = E\{ey\} - \underline{w}_{opt}^T \overset{\text{by orthogonality}\ 0}{\cancel{E\{e\underline{x}\}}}$$

$$= E\{(y - \underline{w}_{opt}^T x) y\} = E\{y^2\} - \underline{w}_{opt}^T E\{xy\} = r_y(0) - \underline{w}_{opt}^T r_{y\underline{x}} \quad \text{for zero-mean } y$$

$$= \sigma_y^2 - r_{y\underline{x}}^T R_x^{-1} r_{y\underline{x}}$$

Ex. $\hat{y} = ax + b$ (Previous example)

$$\hat{y} = (a \ \ b)\underbrace{\begin{pmatrix} x \\ 1 \end{pmatrix}}_{\underline{X}} \quad ; \quad R_x \begin{pmatrix} a \\ b \end{pmatrix} = r_{y\underline{x}} \qquad R_x = E\{XX^T\} = E\left\{\begin{pmatrix} x \\ 1 \end{pmatrix}(x \ \ 1)\right\}$$

$$\Rightarrow R_x = \begin{bmatrix} \sigma_x^2 + m_x^2 & m_x \\ m_x & 1 \end{bmatrix} \quad ; \quad r_{yx} = E\left\{y \begin{pmatrix} x \\ 1 \end{pmatrix}\right\} = \begin{bmatrix} E(xy) \\ \mu_y \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \sigma_x^2 + \mu_x^2 & \mu_x \\ \mu_x & 1 \end{bmatrix}\begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} E(xy) \\ \mu_y \end{bmatrix} \quad \leftarrow \text{Same eq. for optimal } (a,b)$$

Ex. $X_n = \underset{\nwarrow\, N(0,\sigma_c^2)}{C} + \underset{\nwarrow\, N(0,\sigma_{w_n}^2)}{W_n}$; w_n's are independent. ; $\underset{\sim}{c}$ and $\underset{\sim}{w}$ are independent.

$\underset{\uparrow\, \text{Estimator}}{\hat{C}} = \underline{w}^T \underline{X}$; X_n is noisy "c" but noise variance change by n

Find optimal $\underline{w}$ s.t. $E\{(c - \hat{c})^2\}$ is minimized $\qquad (SNR)_k = \dfrac{E\{c^2\}}{E\{w_k^2\}} = \dfrac{\sigma_c^2}{\sigma_{w_k}^2}$

$$\underline{\underline{R}}_x \underline{w} = \overset{\text{obs.}}{\underset{\text{desired}}{r_{c\underline{x}}}} \quad ; \quad R_x = E\{(\underline{w} + c\underline{1})(\underline{w} + c\underline{1})^T\} = E\{ww^T\} + \underline{1}\underset{= E\{c\}E\{w^T\}\ 0\ 0}{\cancel{E\{cw^T\}}} +$$

$$\cancel{E\{\underline{w}c\}}\underline{\underline{1}}^T + E\{c^2\}\underline{1}.1^T$$

$$= \begin{bmatrix} E\{w_1^2\} & E\{w_1 w_2\} & \cdots & \\ \cdots & E\{w_2^2\} & & \\ & & \ddots & \\ & & & E\{w_N^2\} \end{bmatrix} + E\{c^2\} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ \vdots & & \ddots & \\ \vdots & & & \\ 1 & & & 1 \end{bmatrix} = \begin{bmatrix} \sigma_c^2+\sigma_w^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ \sigma_c^2 & \sigma_c^2+\sigma_w^2 & \cdots & \sigma_c^2 \\ & & & \\ \sigma_c^2 & \sigma_c^2 & \cdots & \sigma_c^2+\sigma_w^2 \end{bmatrix}$$

$$r_{cx} = E\{c\underline{x}\} = \sigma_c^2 \underline{1}$$

$$\begin{bmatrix} \sigma_c^2+\sigma_{w_1}^2 & \sigma_c^2 & \cdots & \sigma_c^2 \\ | & \sigma_c^2+\sigma_{w_2}^2 & & \vdots \\ & & \ddots & \sigma_c^2 \\ \sigma_c^2 & & \sigma_c^2 & \sigma_c^2+\sigma_{w_N}^2 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} \sigma_c^2 \\ \vdots \\ \vdots \\ \sigma_c^2 \end{bmatrix} \quad ; \quad SNR = \frac{E\{c^2\}}{E\{w_k^2\}} = \frac{\sigma_c^2}{\sigma_{w_k}^2}$$

$$\begin{bmatrix} 1+\frac{1}{SNR_1} & 1 & \cdots & 1 \\ 1 & 1+\frac{1}{SNR_2} & & 1 \\ \vdots & & & \vdots \\ 1 & & \ddots & 1+\frac{1}{SNR_k} \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{bmatrix}$$

K^{th} Equation in the system: $\underbrace{\sum_{K=1}^{N} w_k}_{S} + \frac{w_k}{(SNR)_k} = 1 \Rightarrow S + \frac{w_k}{(SNR)_k} = 1$

$$\Rightarrow w_k = (SNR)_k (1-S) \quad ; \quad S = \sum_K w_k = \left(\sum (SNR)_k\right)(1-S)$$

$$\Rightarrow S = \frac{\sum_K (SNR)_K}{1+\sum_K (SNR)_k} \rightarrow \overset{(SNR)_k(1-S)}{w_k} = \frac{(SNR)_k}{1+\sum (SNR)_k}$$

if: $SNR_1 = 100$, $S = 103$, $\underline{w} = \frac{1}{103}\begin{bmatrix} 100 \\ 1 \\ 1 \end{bmatrix}$
$SNR_{2,3} = 1$

if all SNR are the same $\Rightarrow w_k = \frac{1}{N + 1/SNR}$

Properties of Linear $\overset{\text{min}}{\text{MMSE}}$ Estimators:

① $E\{e\underline{x}\}=0$ Orthogonality relation

error with optimal coef is uncorrelated with observations.

② Let $y_1, y_2, \dots, y_N$ are N parameters to be estimated from the same observation $\underline{x}$.

$$\hat{\underline{y}} = \underline{\underline{W}}^T\cdot\underline{x} \longrightarrow \begin{array}{l} \hat{y}_1 = w_1^T\cdot x \\ \hat{y}_2 = w_2^T\cdot x \\ \vdots \quad \vdots \\ y_N = w_N^T\cdot \underline{x} \end{array} \longrightarrow \hat{\underline{y}} = \underbrace{\begin{bmatrix} \underline{w}_1^T \\ \underline{w}_2^T \\ \vdots \\ \underline{w}_N^T \end{bmatrix}}_{\underline{\underline{W}}^T}\cdot\underline{x}$$

$$J(\underline{w}) = J(\underline{w}_1, \underline{w}_2, \dots, \underline{w}_N) = E\left\{\sum_{k=1}^{N}(y_k - \underline{w}_k^T\underline{x})^2\right\} = E\{\|\underset{\downarrow}{\underline{y}} - \hat{\underline{y}}\|^2\}$$

To minimize J; we note that, $\underline{w}_k$ only effect $\hat{y}_k$. $\left(\begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} - \underline{\underline{W}}^T\underline{x}\right)$

So individual estimators to estimate $\hat{y}_k$ (i.e. $\underline{\underline{R}}_x\underline{w}_k = r_{y_k\underline{x}}$)

are the optimal ones to min $J(\underline{w}_1, \dots, \underline{w}_N)$

$$\underline{\underline{R}}_x\underline{w}_k = r_{\underline{y_k}\underline{x}} \qquad k = \{1, \dots, N\}$$

$$\underline{\underline{R}}_x[\underline{w}_1, \dots, \underline{w}_N] = [\underset{\uparrow \atop E\{xy_1\}}{r_{y_1\underline{x}}} \quad \underset{\uparrow \atop E\{xy_2\}}{r_{y_2\underline{x}}} \quad \cdots \quad r_{y_N\underline{x}}] = E\{\underline{x}\,\underline{y}^T\}$$

$$\Rightarrow \boxed{\underset{N\times N}{\underline{\underline{R}}_x}\,\underset{N\times N}{\underline{\underline{W}}} = \underset{N\times N}{R_{xy}}} \quad ; \quad \hat{\underline{y}}_{opt} = \underline{\underline{W}}^T\underline{x}$$

$$E\{\underline{e}\,\underline{e}^T\}_{\text{with opt. filter}} = E\{e(\underline{y} - \underline{\underline{W}}^T\underline{x})^T\} = E\{\underline{e}\,\underline{y}^T\} - \cancelto{0}{E\{\underline{e}\,\underline{x}^T\}}\,\underline{\underline{W}}$$

$$= E\{(\underline{y} - \underline{\underline{W}}^T\underline{x})\underline{y}^T\} = \underline{\underline{R}}_y - \underline{\underline{W}}^T\underline{\underline{R}}_{xy} \Rightarrow E\{\underline{e}\,\underline{e}^T\} = \underline{\underline{R}}_y - W^T\underline{\underline{R}}_{xy}$$ (extension of 1-D estimation to N-D)

$e = y - \hat{y}$ $E\{e.x^T\} = 0$ $_{N\times1}$ $_{1\times N}$

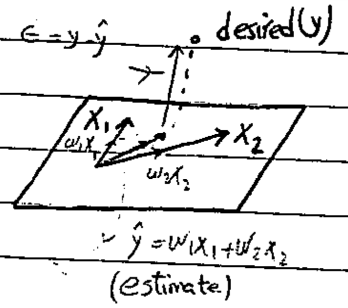
$_{N\times1}$ $e = y - \hat{y} \rightarrow E\{e.x\} = 0$
non vector

66

③ $\underline{e} = \underline{y} - \underline{\hat{y}}$ is orthogonal to a linear a Combination of observation

$E\{\underline{e}.(\underline{\underline{M}}\,\underline{x})^T\} = \underline{\underline{0}} \longrightarrow E\{\underline{e}\,\underline{x}^T\}\underline{\underline{M}}^T = \underline{\underline{0}}$ we have equality

an orbitary matrix



e orthogonal to X_1 and X_2 and its linear Combination (plane)

$E\{xy\} = \langle x, y\rangle$; $E\{x^2\} = \|x\|^2$

④ $\underline{\hat{y}}_{opt} = \underline{\underline{W}}^T_{opt}\underline{X}$; (optimal estimator for $\underline{y}$ is given)

but we would like to estimate $\underline{\overset{u}{\theta}} = \underline{\underline{M}}\,\underline{y}$ not $\underline{y}$

Question: $\underline{\hat{\theta}} \overset{?}{=} \underline{\underline{M}}\,\underline{\hat{y}}$

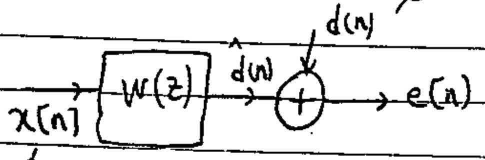
to estimate $\underline{\hat{\theta}} = \underline{\underline{W}}_\theta^T\underline{X} \longrightarrow \underline{\underline{R}}_x\underline{\underline{W}}_\theta = \underline{\underline{R}}_{x\theta} \Rightarrow \underline{\underline{R}}_x\underline{W}_\theta = E\{\underline{X}(\underline{\underline{M}}\,\underline{y})^T\}$

$= E\{\underline{x}\,\underline{y}^T\}\underline{\underline{M}}^T \Rightarrow \underline{\underline{R}}_x\underline{\underline{W}}_\theta = E\{\underline{x}\,\underline{y}^T\}\underline{\underline{M}}^T \Rightarrow \underline{\underline{W}}_\theta = \underline{\underline{R}}_x^{-1}\underline{\underline{R}}_{xy}\underline{\underline{M}}^T$

$\Rightarrow \hat{\theta} = (\underline{\underline{W}}_\theta^T\underline{X}) = \underline{\underline{M}}\,\underline{\underline{R}}_{xy}^T\underline{\underline{R}}_x^{-1}X = \underline{\underline{M}}\,\underline{\hat{y}}$ ✓

FIR Wiener Filter

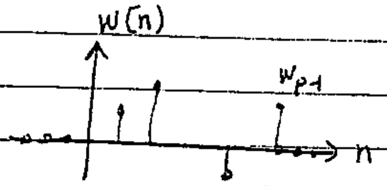
WSS and Joint WSS with x(n)



WSS process

$\hat{d}(n) = x(n) * w(n)$ ← FIR $\quad W(z) = w_0 + w_1 z^{-1} + \dots + W_{P-1} z^{-(P-1)}$

$$J = E\{\underbrace{(d(n) - \hat{d}(n))^2}_{e^2(n)}\}$$



$$\hat{d}(n) = \underline{W}^T \underline{x}(n) = [w_0 \;\; w_1 \;\; \dots \;\; w_{P-1}]^T \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-(P-1)) \end{bmatrix} = \sum_{k=0}^{P-1} W_k x(n-k)$$

$$\underline{\nabla}_w J = \underline{0} \rightarrow \underline{\nabla}_w E\{(d(n) - W^T x(n))^2\} = \underline{0} \Rightarrow 2E\{e \nabla_w e\} = 0$$

$$-2E\{e \underline{x}(n)\} = \underline{0} \Rightarrow E\{e \underline{x}(n)\} = \underline{0} \quad E\{(d(n) - \underbrace{\underline{W}^T \underline{x}(n)}_{\underline{x}^T(n) W}) \underline{x}(n)\} = 0$$

$$\Rightarrow E\{\underline{x}(n) \underline{x}^T(n)\} \underline{W} = E\{d(n) x(n)\}$$

$$\begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-(P-1)) \end{bmatrix} [x(n) \;\; x(n-1) \;\; \dots \;\; x(n-(P-1))] \Rightarrow \underline{\underline{R}}_x \underline{W} = \underline{r}_{dx}$$

$$E\{d(n) x(n)\} = E\{d(n) \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \end{bmatrix}\} = \begin{bmatrix} r_{dx}(0) \\ r_{dx}(1) \\ \vdots \\ r_{dx}(P-1) \end{bmatrix}$$

OPt. Coef. Satisfies: $\underline{\underline{R}}_x \underline{W} = \underline{r}_{dx}$

$$J_{min}(\underline{W}_{opt}) = E\{e(d(n) - W_{opt}^T \underline{x}(n))\} = E\{e d(n)\} - W_{opt}^T \cancel{E\{e \underline{x}(n)\}} =$$

$$E\{(d(n) - W_{opt}^T x(n)) d(n)\} = r_d(0) - W_{opt}^T r_{dx}$$

$$\boxed{J_{min} = r_d(0) - \underline{r}_{dx}^T \underline{\underline{R}}_x^{-1} \underline{r}_{dx}}$$

Using last P samples (Causal FIR Filter)

Filtering Application of wiener Filtering:

$x(n) = d(n) + v(n)$

observation desired noise

Typically $d(n)$ & $v(n)$ are indepedent but for wiener filtering they are uncorrelatedness $\Rightarrow$ is sufficient to find the opt. wiener

Then $r_d(k)$ and $r_v(k)$ is given; $r_x(k) = r_d(k) + r_v(k)$ due to uncorrelatenes of $d(n)$ & $v(n)$ & we assume that these are zero mean processes

Ex. Two Tap wiener filter

$d(n): \quad r_d(k) = \alpha^{|k|}$

$v(n): \quad r_v(k) = \sigma_v^2 \delta(k)$

} Zero mean

$x(n) = d(n) + v(n)$

$\hat{d}(n) = w_0 x(n) + w_1 x(n-1)$; $\underline{\underline{R}}_x \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = r_{d\underline{x}}$

$r_x(k) = E\{x(n) x(n-k)\} = \alpha^{|k|} + \sigma_v^2 \delta(k)$

$r_{d\underline{x}}(k) = E\{d(n) x(n-k)\} = E\{d(n)\,(d(n-k) + v(n-k))\} = \alpha^{|k|} = r_d(k)$

$\underline{\underline{R}}_x \underline{\underline{w}} = r_{d\underline{x}}$ $\quad E\left\{\begin{pmatrix} x(n) \\ x(n-1) \end{pmatrix} (x(n) \;\; x(n-1))\right\} = \begin{bmatrix} r_x(0) & r_x(1) \\ r_x(-1) & r_x(0) \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} r_d(0) \\ r_d(1) \end{bmatrix}$

$\alpha = 0.8$
$\sigma_v^2 = 1$ $\Rightarrow$ $\begin{bmatrix} w_0 = 0.4048 \\ w_1 = 0.2381 \end{bmatrix}$ $\quad W(z) = 0.4048 + 0.2381 z^{-1}$

$J_{min} = r_d(0) - w^T \cancel{R_x^{-1}} r_{d\underline{x}} = 1 \;\; (0.4048 \;\; 0.2381) \begin{pmatrix} 1 \\ 0.8 \end{pmatrix} = 0.4048$

Comparisons:

① $x(n) = d(n) + v(n)$ $\qquad \hat{d}(n) = 1 . x(n)$

$J_1 = E\{(d(n) - \hat{d}(n))^2\} = E\{(+v(n))^2\} = \sigma_v^2 = 1$ (in this example)

② Let's find optimal 1 tap filter

$\hat{d}(n) = \beta_0 x(n)$

$\underline{\underline{R}}_x \beta = \underline{r}_{dx} \Rightarrow r_x(\cdot)\beta = r_{dx}(\cdot) \Rightarrow \beta = \frac{r_{dx}(\cdot)}{r_x(\cdot)} = \frac{1}{1+\sigma_v^2} = 1/2$

$J_{(2)} = r_d(\cdot) - w^T r_{dx} = 1 - (1/2)(1) = 1/2$

SNR Improvement through filtering

Befor

$X(n) = d(n) + v(n)$

$SNR = \frac{E\{(d(n))^2\}}{E\{(v(n))^2\}} = \frac{1}{1} = 1$

$\uparrow$ observation

$\alpha = 0.8 \rightarrow$

$\sigma^2 = 1$

$10 \log_{10} SNR = 0 \text{ dB}$

$\uparrow$ input SNR

After

$\hat{d}(n) = \underline{w}^T x(n) = \underbrace{\underline{w}^T \underline{d}(n)}_{\text{signal}} + \underbrace{\underline{w}^T \underline{v}(n)}_{\text{noise}}$

$$SNR_{output} = \frac{E\{(\text{signal})^2\}}{E\{(\text{noise})^2\}} = \frac{E\{(w^T d(n))^2\}}{E\{(w^T v(n))^2\}}$$

$$= \frac{E\{w^T d(n) d^T(n) w\}}{E\{w^T v(n) v^T(n) w\}} = \frac{\underline{w}^T \underline{\underline{R}}_d \underline{w}}{= \underline{w}^T \underline{\underline{R}}_v w}$$

$v(n) = \begin{pmatrix} v(n) \\ v(n-1) \end{pmatrix} \quad d(n) = \begin{bmatrix} d(n) \\ d(n-1) \\ \vdots \end{bmatrix}$

$\downarrow$ two Tap

$$= \frac{(0.4048 \;\; 0.2381)\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}\begin{bmatrix} 0.4048 \\ 0.2381 \end{bmatrix}}{(0.4048 \;\; 0.2381)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{bmatrix} 0.4048 \\ 0.2381 \end{bmatrix}}$$

$\Rightarrow (SNR)_{dB} \cong 2 \text{ dB}$

Question; Is SNR improvement of 2dB's the max. that we can acheive ?

$(SNR)_{output} = \frac{W^T \underline{R}_d W}{W^T \underline{\underline{R}}_v W}$ max $(SNR)_{output}$ is achived when w = eigenvector of $(\underline{\underline{R}}_v^{-1} \underline{\underline{R}}_d)$ with max eigenvalue.

eigenvalue

$So, R_v = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$; $R_d = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$; $eig \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix} \nearrow \begin{pmatrix} 1 \\ 1 \end{pmatrix} \rightarrow \overset{\lambda_1}{1.8} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\searrow \begin{pmatrix} 1 \\ -1 \end{pmatrix} \rightarrow \underset{\lambda_2}{0.2} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Then $(SNR)_{output}$ maximizing two tap filter is $(1 \quad 1)$

$(\lambda-1)^2 = (0.8)^2$

$\lambda - 1 = \pm 0.8$ / $\lambda_1 = 1.8$, \ $\lambda_2 = 0.2$

$\max (SNR)_{output} = (2.55) dB$

Wiener ~~w~~ filter and SNR maximizing filter are identical if $\underline{\underline{R}}_d$ is rank = 1 $\Rightarrow \underline{\underline{R}}_d = \lambda_x e_x e_x^T$

Wiener Filtering $\equiv$ Optimal Linear min. MSE estimator

$E\{e(n)\, x(n-k)\} = 0$ ← k^{th} sample in history

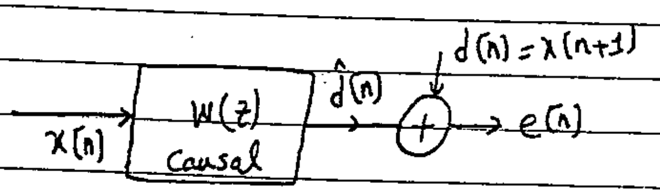
$e(n)$: $d - \underline{w}^T x$

$= d - \underline{w}^T \begin{bmatrix} x(n) \\ x(n-1) \\ \vdots \\ x(n-(P-1)) \end{bmatrix}$

$E\{e\, o_k\} = 0$ orthogonality

e: $d - \underline{w}^T \underline{o}$; o_k: k^{th} observation

2nd Application Area: Lined prediction



$\hat{d}(n) = \sum_{k=0}^{P-1} W_k\, x(n-k)$ ← FIR with P Tap ($(P-1)^{th}$ order ← number of zeros)

$E\{e(n)\, x(n-k)\} = 0 \quad k = 0, \ldots, P-1$

$d(n) - w_{opt}^T \begin{pmatrix} x(n) \\ \vdots \\ x(n-(P-1)) \end{pmatrix}$

$\downarrow$

$\underline{\underline{R}}_x W_{opt} = \underline{r}_{dx}$ $\quad \underline{\underline{R}}_x = E\left\{ \begin{pmatrix} x(n) \\ \vdots \\ x(n-(P-1)) \end{pmatrix} \begin{pmatrix} x(n) & \ldots & x(n-(P-1)) \end{pmatrix} \right\}$ ↘ not a fn of n (WSS)

$r_{dx} = E\{d(n)\, x(n)\}$ ↘ not fn of since JWSS

For this application: $d(n) = x[n+1] \Rightarrow R_x W_{opt} = E\left\{ x[n+1] \begin{bmatrix} x[n] \\ \vdots \\ x[n-(p-1)] \end{bmatrix} \right\} = \begin{bmatrix} r_x(1) \\ r_x(2) \\ \vdots \\ r_x(p) \end{bmatrix}$

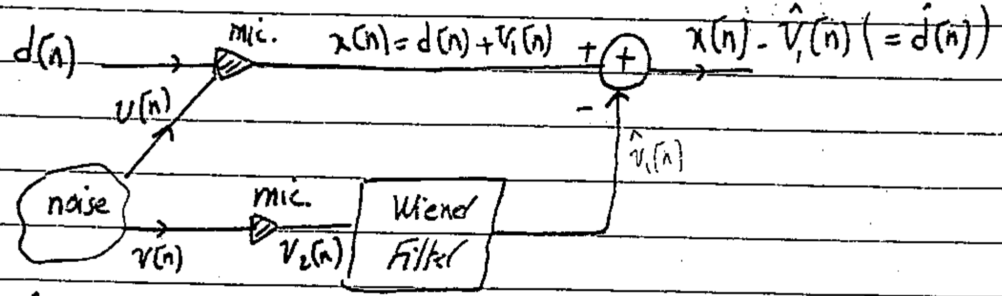
$$\Rightarrow \quad \underline{\underline{R}}_x W_{opt} = \begin{bmatrix} r_x(1) \\ \vdots \\ r_x(p) \end{bmatrix}$$

a Single step linear predictor

if : $d(n) = x(n+\Delta) \rightarrow \underline{\underline{R}}_x W_{opt} = \begin{bmatrix} r_x(\Delta) \\ r_x(\Delta+1) \\ \vdots \\ r_x(\Delta+p-1) \end{bmatrix}$

multi step predictor

③ Noise cancellation: (Hayes) ?



Goal is predict $v_1(n)$ from $v_2(n)$

desired ($v_1(n)$ is tried to predict)

$$\underline{\underline{R}}_{v_2} \underline{w} = r_{\underline{v_1 v_2}}$$

← observation

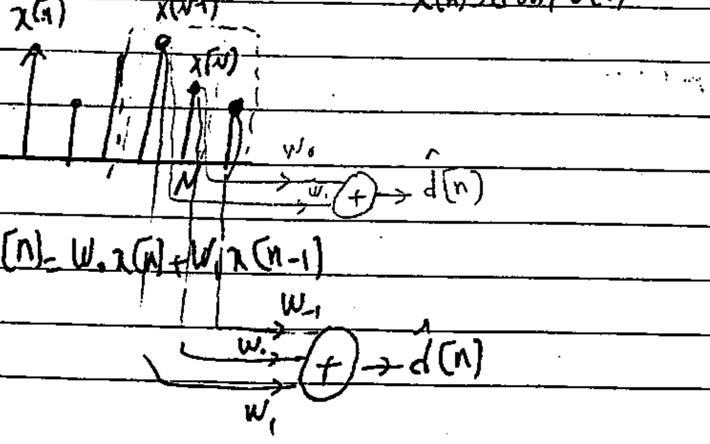
$r_{v_1v_2} = E(v_1 v_2) = E((x-d)v_2) = E(x v_2)$
$= r_{xv_2}$

$$r_{v_1 \underline{v_2}} = E\left\{ v_1(n) \begin{bmatrix} v_2(n) \\ \vdots \\ v_2(n-(p-1)) \end{bmatrix} \right\} = E\left\{ \underset{d(n)+v_1(n)}{x(n)} \begin{bmatrix} v_2(n) \\ \vdots \\ v_2(n-(p-1)) \end{bmatrix} \right\}$$

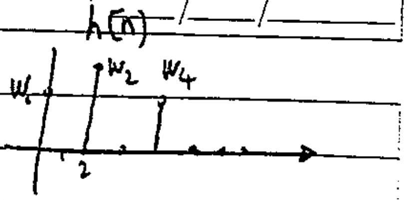
if $d(n)$ and $v_2(n)$ independent and $v_2(n)$ is zero mean

④ Smoothing, Filtering with non-causal filter ?

$x(n) = d(n) + v(n)$



$\hat{d}(n) = W_0 x(n) + W_1 x(n-1)$



Note: $\hat{d}(n) = \sum_{k=0}^{2} w_{2k}\, x(n-2k) = x(n) * h(n)$

$= (w_1 \ w_2 \ w_4) \begin{pmatrix} x(n) \\ x(n-2) \\ x(n-4) \end{pmatrix}$

$\underline{\underline{R_x}}\, W = \underline{r_{dx}} \longrightarrow E\left\{ \begin{pmatrix} x(n) \\ x(n-2) \\ x(n-4) \end{pmatrix} (x(n) \ \ x(n-2) \ \ x(n-4)) \right\} = \underline{\underline{R_x}}$

$E\left\{ \begin{pmatrix} x(n) \\ x(n-2) \\ x(n-4) \end{pmatrix} d(n) \right\} = \begin{pmatrix} r_{dx}(0) \\ r_{dx}(2) \\ r_{dx}(4) \end{pmatrix}$ $\qquad \underline{\underline{R_x}} = \begin{bmatrix} r_x(0) & r_x(2) & r_x(4) \\ r_x(2) & r_x(0) & r_x(-2) \\ r_x(4) & r_x(2) & r_x(0) \end{bmatrix}$

IIR Wiener Filtering:

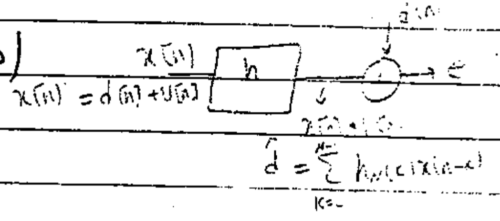
① non-causal Wiener Filter

$J = E\{e^2\}$; $e = d(n) - \sum_k h(k) x(n-k)$

$\hat{d}(n) = \sum_{k=-\infty}^{+\infty} h_k x(n-k)$ $\qquad \frac{\partial J}{\partial h_k} = 0 \Rightarrow E\{ e \cdot x(n-k) \} = 0$

$E\{ \underset{\uparrow}{e(n)}\, x(n-k) \} = 0 \quad k \in (-\infty, +\infty)$

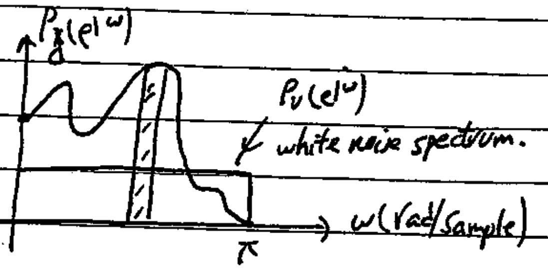
$d(n) - \sum_{k'} h(k') x(n-k')$



$\Rightarrow r_{dx}(k) - \sum_{k'} h(k') r_x(k-k') = 0 \quad \forall k$

$r_{dx}(k) - h(k) * r_x(k) = 0 \quad \forall k \quad \Rightarrow \text{DTFT} \Rightarrow$

$P_{dx}(e^{j\omega}) - H(e^{j\omega}) P_x(e^{j\omega}) = 0 \Rightarrow \underset{\substack{\uparrow \\ \text{non-Causal} \\ \text{IIR Filter}}}{H(e^{j\omega})} = \frac{P_{dx}(e^{j\omega})}{P_x(e^{j\omega})}$



Periodic (2π), $P_x(e^{j\omega})$ is even Fn of ω

$\hat{\sigma}_y^2 = \frac{1}{2\pi} \int P_x(e^{j\omega}) |H(e^{j\omega})|^2 d\omega$

$x(n) = d(n) + v(n)$

$H(e^{j\omega}) = \frac{P_{dx}}{P_x} \Rightarrow H(e^{j\omega}) = \frac{P_d(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})}$ uncorrelated

$E(dx) = E[(d+v) d] = E[d^2 + dv] = E(d^2)$

IIR: $J_{min} = r_d(0) - \sum h_k r_{dx}(k) = \frac{1}{2\pi}\int_{-\pi} P_v(e^{j\omega}) H(e^{j\omega}) d\omega$

Ex: $x(n) = d(n) + v(n)$ $r_d(k) = 0.8^{|k|}$ $r_v(k) = \delta(k)$

$\rightarrow$ 2 Tap Wiener Filter: $J_{min} = 0.4048$

1 " " " : $J_{min} = 0.5$

What is the min error for non causal linear filter?

$$H^{NC}(z) = \frac{P_{dx}(z)}{\underset{\uparrow P_d(z)+P_v(z)}{P_x(z)}} = \frac{P_d(z)}{P_d(z)+1} = \left.\frac{\frac{1-\alpha^2}{(1-\alpha z^{-1})(1-\alpha z)}}{\frac{1-\alpha^2}{(1-\alpha z^{-1})(1-\alpha z)}+1}\right|_{\alpha=0.8} = 0.3\,\frac{(1-1/4)}{(1-\frac{1}{2}z)(1-\frac{1}{2}z^{-1})}$$

?

$\Rightarrow h^{NC}(n) = 0.8\,\frac{1}{2}^{|n|}$

J_{min} Calculation: FIR $\rightarrow$ $J_{min} = r_d(0) - w^T \underline{r}_{dx}$

IIR: $J_{min}^{w_{opt}} = E\{e(n)\,(d(n) - \sum h_k x(n-k))^*\} = E\{\underset{\downarrow d(n)}{d^*(n)}(-\sum h_k^* x^*(n-k))\}$

$$= r_d(0) - \sum_{k=-\infty}^{+\infty} h_k r_{dx}^*(k) = r_d(0) \cancel{+} - \left[h(-\infty) \;\cdots\; h(\infty)\right]\begin{bmatrix} r_{dx}(-\infty) \\ \vdots \\ r_{dx}(\infty)\end{bmatrix}$$ non-Causal IIR

$\Rightarrow J_{min} = r_d(0) - \sum h_k r_{dx}(k)$ for real process $\rightarrow$ Parseval's Thm.

$\int x(t) g^*(t) dt = \frac{1}{2\pi}\int X(e^{j\omega}) G^*(e^{j\omega}) d\omega$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} P_d(e^{j\omega}) d\omega - \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{j\omega}) P_{dx}(e^{j\omega}) d\omega$$

$\uparrow$ For Filtering $P_{dx}(e^{j\omega}) = P_d(e^{j\omega})$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \underbrace{(1 - H(e^{j\omega}))}_{1 - \frac{P_d(e^{j\omega})}{P_d + P_v}} P_d(e^{j\omega}) d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{P_v(e^{j\omega})}{P_d(e^{j\omega}) + P_v(e^{j\omega})} P_d(e^{j\omega}) d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \underset{\uparrow \sigma_v^2 = 1}{P_v(e^{j\omega})} H(e^{j\omega}) d\omega = h[0] \Rightarrow J_{min} = h[0] = 0.3$$

best performance with ∞ Taps

___/___/______

The causal IIR Linear filter:

$$\hat{d}[n] = \sum_{k=0}^{\infty} h_k x[n-k] \leftarrow \text{IIR \& Causal}$$

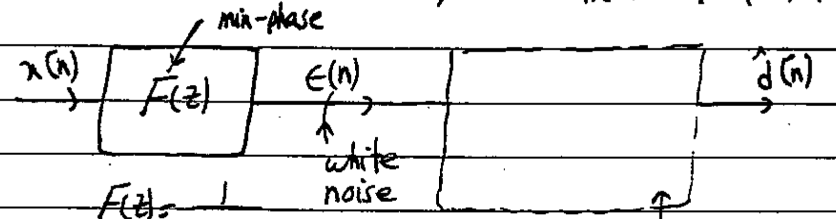
$$E\{e[n]x[n-k]\}=0 \quad \text{orthogonality} \quad k \geq 0$$

If $x[n]$ is white noise ; $E\{e(n)x[n-k]\} = E\{(d(n) - \sum_{k'} h_{k'} x[n-k'])\, x(n-k)\} = 0$

$$\Rightarrow \sum_{k'=0}^{\infty} h_{k'} \underset{\uparrow \atop \delta[k-k']}{r_x(k-k')} = r_{dx}(k) \quad k \geq 0$$

$$\Rightarrow \boxed{h_k = r_{dx}(k) \qquad k \geq 0}$$

If $x[n]$ is not white, but $P_x(z) = \sigma_0^2 Q(z) Q^*(\frac{1}{z^*})$ (has spectral factorization)



$F(z) = \dfrac{1}{\sigma_0 Q(z)}$ ← Causal & invertable

G(z) ≡ Causal IIR filter

$G(z)$; is a IIR filter with coeff. $g(k) = r_{d\epsilon}(k)$

$$r_{d\epsilon}(k) = ? \quad E\{d(n)(\underbrace{\sum_{k'} f_{k'} x(n-k-k')}_{e[n-k]})\} = \sum_{k'=0}^{\infty} f_{k'} r_{dx}(k+k') = f_k * r_{dx}(-k)$$

$$k'' = -k'$$

$$\sum_{k'=-\infty}^{+\infty} f[-k''] u[-k''] r_{dx}(k-k'') = f(-k) * r_{dx}(+k)$$

$$P_{d\epsilon}(z) = F(\tfrac{1}{z}) P_{dx}(z)$$

$$G(z) = [P_{d\epsilon}(z)]_+ = [F(\tfrac{1}{z}) P_{dx}(z)]_+ = \left[\frac{P_{dx}(z)}{\sigma_0 Q(1/z)}\right]_+$$

$$[-5z^2 + 3z + 3 + 10z^{-1} + 20z^{-2}]_+ = 3 + 10z^{-1} + 20z^{-2}$$

$]_+$ take the causal part of input

$$\underset{\text{Causal IIR}}{H(z)} = F(z) G(z) = \frac{1}{\sigma_0 Q(z)} \left[\frac{P_{dx}(z)}{\sigma_0 Q(\frac{1}{z})}\right]_+$$