Linear Models
in
Statistical Signal Processing

Anima Mishra Sharma, Y8104001

Bishwajit Sharma, Y8104015

Contents

- Definition & Properties
- Linear Model Examples
  - Curve Fitting
  - Fourier Analysis

- Generalized Linear Model
  - Whitening of colored noise
  - Known signal components
- MVUE for General Linear Model

Definition

- $x[n] = A + Bn + w[n] \qquad n = 0, 1, \ldots, N-1$
- Matrix notation for Linear Model: $\mathbf{x} = \mathcal{H}\boldsymbol{\theta} + \mathbf{w}$,

$$\begin{aligned}
\mathbf{x} &= [x[0]\ x[1] \ldots x[N-1]]^T \\
\mathbf{w} &= [w[0]\ w[1] \ldots w[N-1]]^T \\
\boldsymbol{\theta} &= [A\ B]^T
\end{aligned}$$

$$\mathbf{H} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ \vdots & \vdots \\ 1 & N-1 \end{bmatrix}$$

- H: Nx2 observation matrix
- Note: H^TH is invertible
- $\mathbf{w} \sim \mathcal{N}(0, \sigma^2 I)$

MVUE: Linear Model

$$\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \mathbf{I}(\boldsymbol{\theta})(\mathbf{g}(\mathbf{x}) - \boldsymbol{\theta})$$

$$\begin{aligned}\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} &= \frac{\partial}{\partial \boldsymbol{\theta}}\left[-\ln(2\pi\sigma^2)^{\frac{N}{2}} - \frac{1}{2\sigma^2}(\mathbf{x}-\mathbf{H}\boldsymbol{\theta})^T(\mathbf{x}-\mathbf{H}\boldsymbol{\theta})\right] \\ &= -\frac{1}{2\sigma^2}\frac{\partial}{\partial \boldsymbol{\theta}}\left[\mathbf{x}^T\mathbf{x} - 2\mathbf{x}^T\mathbf{H}\boldsymbol{\theta} + \boldsymbol{\theta}^T\mathbf{H}^T\mathbf{H}\boldsymbol{\theta}\right].\end{aligned}$$

$$\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \frac{\mathbf{H}^T\mathbf{H}}{\sigma^2}[(\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x} - \boldsymbol{\theta}]$$

$$\begin{aligned}\hat{\boldsymbol{\theta}} &= (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x} \\ \mathbf{I}(\boldsymbol{\theta}) &= \frac{\mathbf{H}^T\mathbf{H}}{\sigma^2}\end{aligned}$$

$\partial\theta$

Linear transformation of Gaussian vector remains Gaussian,

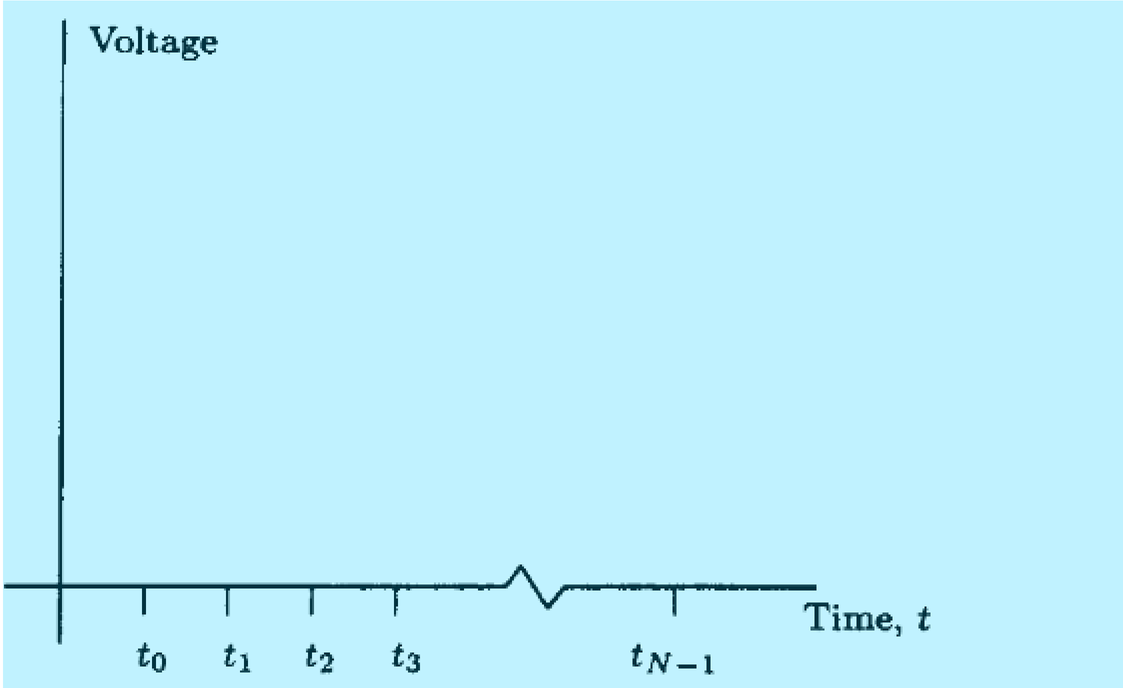
$$\hat{\boldsymbol{\theta}} \sim \mathcal{N}(\boldsymbol{\theta}, \sigma^2(\mathbf{H}^T\mathbf{H})^{-1})$$

$$\mathbf{C}_{\hat{\theta}} = \sigma^2(\mathbf{H}^T\mathbf{H})^{-1}$$

CRLB is achieved

Curve Fitting

- Data model: $x(t_n) = \theta_1 + \theta_2 t_n + \theta_3 t_n^2 + w(t_n) \qquad n = 0, 1, \ldots, N-1.$
- $w(t_n)$ are IID, $\boldsymbol{w} \sim \mathcal{N}(0, \sigma^2 \boldsymbol{I})$, ie WGN
- Linear model: $\boldsymbol{x} = \mathcal{H}\boldsymbol{\theta} + \boldsymbol{w}$



$$\begin{aligned}\mathbf{x} &= [x(t_0)\, x(t_1) \ldots x(t_{N-1})]^T \\ \boldsymbol{\theta} &= [\theta_1\, \theta_2\, \theta_3]^T\end{aligned}$$

$$\mathbf{H} = \begin{bmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{N-1} & t_{N-1}^2 \end{bmatrix}$$

Curve fitting: General case

- Fitting a polynomial of order (p-1)
- Data model

$$x(t_n) = \theta_1 + \theta_2 t_n + \cdots + \theta_p t_n^{p-1} + w(t_n) \qquad n = 0, 1, \ldots, N-1.$$

- MVU estimator of linear model: $\hat{\boldsymbol{\theta}} = (\mathbf{H}^T\mathbf{H})^{-1}\mathbf{H}^T\mathbf{x}$
- Linear model with $\mathbf{x} = [x(t_0)\, x(t_1) \ldots x(t_{N-1})]^T$

$$\mathbf{H} = \begin{bmatrix} 1 & t_0 & \ldots & t_0^{p-1} \\ 1 & t_1 & \ldots & t_1^{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & t_{N-1} & \ldots & t_{N-1}^{p-1} \end{bmatrix}$$

- Result of curve-fitting: $\hat{s}(t) = \sum_{i=1}^{p} \hat{\theta}_i t^{i-1}$

Fourier Analysis

- Data model:

$$x[n] = \sum_{k=1}^{M} a_k \cos\left(\frac{2\pi kn}{N}\right) + \sum_{k=1}^{M} b_k \sin\left(\frac{2\pi kn}{N}\right) + w[n] \qquad n = 0, 1, \ldots, N-1$$

- Noise w[n] is WGN, $a_k\ b_k$ are to be determined

$$\boldsymbol{\theta} = [a_1\, a_2 \ldots a_M\, b_1\, b_2 \ldots b_M]^T$$

- Formulating as linear model $\mathcal{H}$ (Nx2M)

$$\mathbf{H} = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ \cos\left(\frac{2\pi}{N}\right) & \cdots & \cos\left(\frac{2\pi M}{N}\right) & \sin\left(\frac{2\pi}{N}\right) & \cdots & \sin\left(\frac{2\pi M}{N}\right) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \cos\left[\frac{2\pi(N-1)}{N}\right] & \cdots & \cos\left[\frac{2\pi M(N-1)}{N}\right] & \sin\left[\frac{2\pi(N-1)}{N}\right] & \cdots & \sin\left[\frac{2\pi M(N-1)}{N}\right] \end{bmatrix}$$

Fourier Analysis

- System is solvable if N > 2M
- H is orthogonal: computations can be simplified
- Orthogonal columns: $\boldsymbol{h}_i^T \boldsymbol{h}_j = 0$, if i not equal to j
- $\mathcal{H}^T\mathcal{H} = (N/2)I$, diagonal matrix
- MVU Estimator is given by

$$\hat{\boldsymbol{\theta}} = \frac{2}{N}\mathbf{H}^T\mathbf{x} = \frac{2}{N}\begin{bmatrix} \mathbf{h}_1^T \\ \vdots \\ \mathbf{h}_{2M}^T \end{bmatrix}\mathbf{x} = \begin{bmatrix} \frac{2}{N}\mathbf{h}_1^T\mathbf{x} \\ \vdots \\ \frac{2}{N}\mathbf{h}_{2M}^T\mathbf{x} \end{bmatrix}$$

- Solution:

$$\hat{a}_k = \frac{2}{N}\sum_{n=0}^{N-1} x[n]\cos\left(\frac{2\pi kn}{N}\right) \qquad \hat{b}_k = \frac{2}{N}\sum_{n=0}^{N-1} x[n]\sin\left(\frac{2\pi kn}{N}\right)$$

General Linear model: 1

- Colored Gaussian noise: $\boldsymbol{w} \sim \mathcal{N}(0, \mathcal{C})$,
- General linear model: $\mathcal{C}$ may not be scaled $\boldsymbol{I}$
- Whitening approach
  - C and C^{-1} are positive definite
  - Factorize C^{-1} as $C^{-1} = D^T D$
  - $E[(Dw)(Dw)^T] = DCD^T = DD^{-1}(D^T)^{-1}D^T = \boldsymbol{I}$
  - Hence D is a Whitening transformation

General Linear model: 1

- $x' = Dx = DH\theta + Dw = H'\theta + w'$
  - $w' = Dw \sim \mathcal{N}(0, I)$
- The MVU estimator of θ
  - $\theta = (H'^T H')^{-1} H'^T x' = (H^T C^{-1} H)^{-1} H^T C^{-1} x$
  - $C_\theta = (H'^T H')^{-1} = (H^T C^{-1} H)^{-1}$

Example: Colored Noise

- Case: DC level in Colored Noise
- $x[n] = A + w[n]$ for $n = 0,1\ldots..N - 1$
- $w[n]$: Colored Gaussian noise with covariance matrix C [NxN]
- Observation matrix $H = \mathbf{1} = [1\ 1\ldots1]$
- MVUE A: $(\mathbf{1}^T C^{-1} x)/(\mathbf{1}^T C^{-1} \mathbf{1})$
  - Interpreted as whitened data $x'[n]$ averaged with weights d_n

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

- var(A): $1/(1^T c^{-1} 1)$

General Linear model: 2

- s: Known signal contained in the data
- Data model: $x = H\theta + s + w$
- Choose $x' = x - s$
- Linearized model $x' = H\theta + w$
- MVU estimator $\theta = (H^T H)^{-1} H^T (x - s)$
- Covariance $C_\theta = \sigma^2 (H^T H)^{-1}$

Example: DC + Exponential

- DC level and Exponential in White Noise
- $x[n] = A + r^n + w[n]$ $n = 0,1\ldots..N-1$
  - r is known
  - w[n] is WGN
  - A is to be estimated
- Model: $\mathbf{x} = A[1\ 1\ \ldots 1]^T + \mathbf{s} + \mathbf{w}$,
  - $\mathbf{s} = [1\ r\ \ldots\ r^{N-1}]^T$
- MVUE: $$\hat{A} = \frac{1}{N}\sum_{n=0}^{N-1}(x[n] - r^n)$$
- Variance: $$\mathrm{var}(\hat{A}) = \frac{\sigma^2}{N}.$$

MVUE for General Linear Model

- Theorem: If data can be modeled as
  - $\mathbf{x} = \mathbf{H}\theta + \mathbf{s} + \mathbf{w}$
    - **x**: (Nx1) vector of observations
    - **H**: (Nxp) observation matrix, rank p
    - θ: (px1) vector of parameters to be estimated,
    - **s**: (Nx1) vector of known signal samples
    - **w**: (NX1) noise vector with PDF $\mathcal{N}(0, C)$
- Then
  - MVUE: $\theta = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}(\mathbf{x} - \mathbf{s})$, attains CRLB
  - Covariance: $\mathbf{C}_\theta = \sigma^2 (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}$

Thankyou!