

# Estimación y Predicción en Series Temporales

Estimadores insesgados lineales óptimos (BLUE)

Departamento de Procesamiento de Señales

Instituto de Ingeniería Eléctrica  
Facultad de Ingeniería

2022

- 1 Repaso estimadores MVU en modelos lineales
- 2 Estimadores Lineales e Insesgados Óptimos (BLUE)

## Modelo lineal en ruido AWGN

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{w}$$

- $\mathbf{x} \in \mathbb{R}^N$ : datos observados.
- $\boldsymbol{\theta} \in \mathbb{R}^p$ :  $p$  parámetros desconocidos.
- $\mathbf{h} \in \mathbb{R}^{N \times p}$ : matriz de observación, con  $N > p$  y rango  $p$ .
- $\mathbf{w} \in \mathbb{R}^N$ : ruido en observación. Se asume  $\mathbf{w} \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$ .

$$p(\mathbf{x}; \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \exp \left[ -\frac{1}{2\sigma^2} (\mathbf{x} - \mathbf{h}\boldsymbol{\theta})^T (\mathbf{x} - \mathbf{h}\boldsymbol{\theta}) \right]$$

- Tomando el logaritmo y derivando con respecto a  $\boldsymbol{\theta}$  se obtiene,

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

# Modelo lineal + ruido blanco gaussiano

## CRLB en Modelos Lineales:

- El modelo cumple la condición del teorema de CRLB dada por,

$$\frac{\log p(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \underbrace{\frac{\mathbf{h}^T \mathbf{h}}{\sigma^2}}_{\mathbf{I}(\boldsymbol{\theta})} \left[ \underbrace{(\mathbf{h}^T \mathbf{h})^{-1} \mathbf{h}^T \mathbf{x}}_{\mathbf{g}(\mathbf{x})} - \boldsymbol{\theta} \right].$$

- Por lo que tenemos

Estimador MVU

$$\hat{\boldsymbol{\theta}} = \mathbf{g}(\mathbf{x}) = (\mathbf{h}^T \mathbf{h})^{-1} \mathbf{h}^T \mathbf{x}$$

Matriz de Covarianza

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

- Además, el estimador  $\hat{\boldsymbol{\theta}}$  es **eficiente**.
- Observación:** Problema equivalente al problema de mínimos cuadrados.

# Modelo lineal + ruido coloreado gaussiano

$$\mathbf{x} = \mathbf{h}\boldsymbol{\theta} + \mathbf{w}$$

- $\mathbf{x} \in \mathbb{R}^N$ : datos observados.
- $\boldsymbol{\theta} \in \mathbb{R}^p$ :  $p$  parámetros desconocidos.
- $\mathbf{h} \in \mathbb{R}^{N \times p}$ : matriz de observación, con  $N > p$  y rango  $p$ .
- $\mathbf{w} \in \mathbb{R}^N$ : Se asume  $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$  (ruido coloreado).

- En este caso vimos que aplicando el blanqueo a los datos:

$$\mathbf{x}' = \mathbf{D}\mathbf{x} = \mathbf{D}\mathbf{h}\boldsymbol{\theta} + \mathbf{D}\mathbf{w}$$

con  $\mathbf{D}$  matriz  $N \times N$  invertible t.q.  $\mathbf{C}^{-1} = \mathbf{D}^T \mathbf{D}$ .

- Se plantea un nuevo problema lineal  $\mathbf{x}' = \mathbf{h}'\boldsymbol{\theta} + \mathbf{w}'$  con  $\mathbf{h}' = \mathbf{D}\mathbf{h}$  y  $\mathbf{w}' = \mathbf{D}\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
- El estimador MVU (y eficiente) es en este caso,

$$\hat{\boldsymbol{\theta}} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \mathbf{h}^T \mathbf{C}^{-1} (\mathbf{x})$$

y la matriz de covarianza del estimador es

$$\mathbf{C}_{\theta} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1}.$$

- Qué pasa cuando no conocemos la PDF de  $\mathbf{w}$ ? (esta clase)

## Estimadores Lineales e Insesgados Óptimos (BLUE)

# Estimadores Lineales e Insesgados Óptimos (BLUE)

- En la práctica ocurre frecuentemente que aunque el estimador MVU existe, no puede ser encontrado.
  - No se conoce la PDF o no se puede asumir un modelo de los datos
  - No puede aplicarse CRLB o teoría de estadísticos suficientes
- En estos casos es razonable recurrir a estimadores subóptimos
  - No se conoce la pérdida de desempeño del estimador porque no se conoce la varianza del MVU.
  - Estimador puede usarse si cumple requisitos de varianza del problema.
- Una estrategia es restringir el estimador a ser **lineal con los datos**, y buscar el **estimador lineal insesgado y de varianza mínima (BLUE)**
- Virtud del estimador BLUE: solo se necesita conocer el **primer y segundo momento** de la PDF, por lo que es apropiado en aplicaciones prácticas.

# Mejor Estimador Lineal Insesgado

*Best Linear Unbiased Estimator (BLUE)*

Se observa el conjunto de datos  $\{x[0], x[1], \dots, x[N-1]\}$  cuya PDF  $p(\mathbf{x}; \theta)$  depende del parámetro desconocido  $\theta$  que se quiere estimar.

- Se dice que un estimador es **lineal** si se restringe a ser lineal con los datos,

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

- **Estimador BLUE:** estimador lineal, insesgado y tiene varianza mínima entre todos los estimadores lineales.
- Hay que determinar los coeficientes  $a_n$  para que el estimador cumpla estas condiciones.



# Búsqueda del BLUE

Para determinar el BLUE se impone que el estimador  $\hat{\theta}$  sea lineal e insesgado y se determinan los coeficientes  $a_n$  que minimizan la varianza.

- Estimador lineal

$$\hat{\theta} = \sum_{n=0}^{N-1} a_n x[n] = \mathbf{a}^T \mathbf{x}$$

- Condición de estimador insesgado

$$\mathbb{E}(\hat{\theta}) = \sum_{n=0}^{N-1} a_n \mathbb{E}(x[n]) = \theta, \quad \forall \theta.$$

- Para satisfacer la condición de insesgado,  $\mathbb{E}(x[n])$  tiene que ser lineal con el parámetro desconocido  $\theta$ ,

$$\mathbb{E}(x[n]) = s[n]\theta,$$

con  $s[n]$  conocido.

**Nota.** Si esto no se cumple, es imposible satisfacer la condición de insesgado.

**Ejemplo.** si  $\mathbb{E}(x[n]) = \cos \theta$ , la condición de insesgado sería  $\sum_{n=0}^{N-1} a_n \cos \theta = \theta$ . No existen coeficientes  $a_n$  que cumplan esto para todo  $\theta$ .

# Búsqueda del BLUE

- Condición necesaria para ser insesgado, implica que el BLUE solo es aplicable en estimación de amplitud de señales conocidas en ruido,

$$x[n] = \theta s[n] + w[n].$$

- Se puede generalizar mediante transformaciones no lineales de los datos (e.g.  $y[n] = x[n]^2$  para estimar  $\sigma^2$  en WGN).
- Continuando con la condición de no sesgado,

$$\sum_{n=0}^{N-1} a_n \mathbb{E}(x[n]) = \theta$$

$$\sum_{n=0}^{N-1} a_n s[n] \theta = \theta$$

$$\sum_{n=0}^{N-1} a_n s[n] = 1$$

es decir,

$$\mathbf{a}^T \mathbf{s} = 1,$$

con  $\mathbf{s} = [s[0], s[1], \dots, s[N-1]]^T$ . (Notar que hay infinitos  $\mathbf{a}^T$  posibles.)

# Búsqueda del BLUE

- Por otro lado, la varianza del estimador es:

$$\begin{aligned}\text{var}(\hat{\theta}) &= \mathbb{E} \left[ \left( \hat{\theta} - \mathbb{E}(\hat{\theta}) \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \mathbf{a}^T \mathbf{x} - \mathbb{E}(\mathbf{a}^T \mathbf{x}) \right)^2 \right] \\ &= \mathbb{E} \left[ \left( \mathbf{a}^T (\mathbf{x} - \mathbb{E}(\mathbf{x})) \right)^2 \right] \\ &= \mathbb{E} \left[ \mathbf{a}^T (\mathbf{x} - \mathbb{E}(\mathbf{x})) (\mathbf{x} - \mathbb{E}(\mathbf{x}))^T \mathbf{a} \right] \\ &= \mathbf{a}^T \mathbb{E} \left[ (\mathbf{x} - \mathbb{E}(\mathbf{x})) (\mathbf{x} - \mathbb{E}(\mathbf{x}))^T \right] \mathbf{a} \\ &= \mathbf{a}^T \mathbf{C} \mathbf{a}\end{aligned}$$

- Hay que encontrar  $\mathbf{a}$  de forma de minimizar la varianza manteniendo la restricción impuesta de estimador insesgado,

$$\mathbf{a}_* = \arg \min_{\mathbf{a}} \mathbf{a}^T \mathbf{C} \mathbf{a} \quad \text{sujeto a} \quad \mathbf{a}^T \mathbf{s} = 1$$

Se puede resolver utilizando multiplicadores de Lagrange.

# Búsqueda del BLUE

- El Lagrangiano es:

$$\mathcal{L}(\mathbf{a}, \lambda) = \mathbf{a}^T \mathbf{C} \mathbf{a} + \lambda(\mathbf{a}^T \mathbf{s} - 1)$$

- El gradiente respecto a  $\mathbf{a}$  es:

$$\frac{\partial \mathcal{L}(\mathbf{a}, \lambda)}{\partial \mathbf{a}} = 2\mathbf{C} \mathbf{a} + \lambda \mathbf{s}$$

- Imponiendo  $\frac{\partial \mathcal{L}(\mathbf{a}, \lambda)}{\partial \mathbf{a}} = \mathbf{0}$ , se tiene que

$$\mathbf{a} = -\frac{\lambda}{2} \mathbf{C}^{-1} \mathbf{s}. \quad (1)$$

- Sustituyendo el valor de  $\mathbf{a}$  en la restricción, se encuentra el multiplicador de Lagrange,

$$\mathbf{a}^T \mathbf{s} = -\frac{\lambda}{2} \mathbf{s}^T \mathbf{C}^{-1} \mathbf{s} = 1,$$

con lo cual

$$-\frac{\lambda}{2} = \frac{1}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}$$

- Sustituyendo el valor del multiplicador en (1), se obtiene el valor de  $\mathbf{a}$  óptimo,

$$\mathbf{a}_* = \frac{\mathbf{C}^{-1} \mathbf{s}}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}.$$

# Búsqueda del BLUE

- El estimador BLUE es entonces

$$\hat{\theta} = \mathbf{a}_*^T \mathbf{x} = \frac{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}$$

- Su varianza es,

$$\text{var}(\hat{\theta}) = \mathbf{a}_*^T \mathbf{C} \mathbf{a}_* = \frac{1}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}}$$

- Notar que el estimador es insesgado (usando que  $\mathbb{E}(x[n]) = s[n]\theta$ ),  $\forall \theta$ ,

$$\begin{aligned}\mathbb{E}(\hat{\theta}) &= \frac{\mathbf{s}^T \mathbf{C}^{-1} \mathbb{E}(\mathbf{x})}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}} \\ &= \frac{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s} \theta}{\mathbf{s}^T \mathbf{C}^{-1} \mathbf{s}} \\ &= \theta.\end{aligned}$$

- Observación: Para determinar el BLUE solo se requiere conocer:
  - $\mathbf{s}$  (media escalada) y  $\mathbf{C}$  (matrix de covarianza de  $\mathbf{x}$ )

Es decir, los dos primeros momentos de  $p(\mathbf{x})$  en lugar de la PDF completa.

# Ejemplo I

## Ejemplo: Nivel de DC en ruido blanco

Se quiere estimar  $A$  a partir de las observaciones

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

donde  $w[n]$  es ruido blanco con varianza  $\sigma^2$  y **PDF no está especificada**.

- En este caso,  $\mathbb{E}(x[n]) = A$ , con lo cual  $s[n] = 1$ , para todo  $n$ . Es decir  $\mathbf{s} = \mathbf{1}$ .
- La matriz de covarianza de  $x[n]$  es  $\mathbf{C} = \sigma^2 \mathbf{I}$ . Con lo cual  $\mathbf{C}^{-1} = \frac{1}{\sigma^2} \mathbf{I}$ .
- Usando la fórmula para el estimador BLUE vista anteriormente se obtiene,

$$\hat{A} = \frac{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{x}}{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{1}} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] = \bar{x}.$$

- La varianza del estimador BLUE es,

$$\text{var}(\hat{A}) = \frac{1}{\mathbf{1}^T \frac{1}{\sigma^2} \mathbf{I} \mathbf{1}} = \frac{\sigma^2}{N}.$$

- Media muestral es el BLUE independientemente de PDF del ruido.

# Ejemplo II

## Ejemplo: Nivel de DC en ruido blanco de varianza no constante

Se quiere estimar  $A$  a partir de las observaciones

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

donde  $w[n]$  es ruido blanco con  $\text{var}(w[n]) = \sigma_n^2$  y **PDF no está especificada**.

- Como en el caso anterior,  $s = 1$ .
- La matriz de covarianza es ahora

$$\mathbf{C} = \begin{bmatrix} \sigma_0^2 & 0 & \cdots & 0 \\ 0 & \sigma_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{N-1}^2 \end{bmatrix} \quad \text{con lo cual, } \mathbf{C}^{-1} = \begin{bmatrix} \frac{1}{\sigma_0^2} & 0 & \cdots & 0 \\ 0 & \frac{1}{\sigma_1^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\sigma_{N-1}^2} \end{bmatrix}.$$

### Ejemplo: Nivel de DC en ruido blanco de varianza no constante

- El estimador BLUE es,

$$\hat{A} = \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{x}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} = \frac{\sum_{n=0}^{N-1} \frac{x[n]}{\sigma_n^2}}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}.$$

- Y su varianza es,

$$\text{var}(\hat{A}) = \frac{1}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} = \frac{1}{\sum_{n=0}^{N-1} \frac{1}{\sigma_n^2}}$$

- El estimador BLUE da más peso a las muestras con menor varianza (para igualar contribución de ruido en cada muestra).
- El estimador encontrado es idéntico al estimador MVU encontrado en el caso de ruido Gaussiano no correlacionado.



# Estimador BLUE: Extensión a vector de parámetros

- Se quiere encontrar el BLUE para el caso de un vector  $\boldsymbol{\theta}$  de  $p$  parámetros

$$\boldsymbol{\theta} = [\theta_1, \theta_2, \dots, \theta_p]^T.$$

- Para que el estimador sea lineal con los datos, se requiere

$$\hat{\theta}_i = \sum_{n=0}^{N-1} a_{in} x[n] = \mathbf{a}_i^T \mathbf{x}, \quad i = 1, 2, \dots, p,$$

donde  $a_{in}$  son los coeficientes a estimar.

- En notación matricial,

$$\hat{\boldsymbol{\theta}} = \mathbf{a}\mathbf{x}, \quad \mathbf{a} \text{ matrix } p \times N \text{ con filas } \mathbf{a}_i^T.$$

- La condición para que el estimador  $\hat{\boldsymbol{\theta}}$  sea insesgado es,

$$\mathbb{E}(\hat{\theta}_i) = \sum_{n=0}^{N-1} a_{in} \mathbb{E}(x[n]) = \theta_i, \quad \forall \theta_i, \quad i = 1, 2, \dots, p$$

que en notación matricial es

$$\mathbb{E}(\hat{\boldsymbol{\theta}}) = \mathbf{a}\mathbb{E}(\mathbf{x}) = \boldsymbol{\theta}, \quad \forall \boldsymbol{\theta}.$$

# Estimador BLUE: Extensión a vector de parámetros

- Análogamente al caso escalar, para satisfacer la condición de estimador insesgado,  $\mathbb{E}(\mathbf{x})$  tiene que ser lineal con  $\boldsymbol{\theta}$ ,

$$\mathbb{E}(\mathbf{x}) = \mathbf{h}\boldsymbol{\theta}, \quad \text{con } \mathbf{h} \text{ matriz } N \times p$$

- Sustituyendo en la condición de insesgado del estimador lineal  $\mathbb{E}(\hat{\boldsymbol{\theta}}) = \mathbf{a}\mathbb{E}(\mathbf{x}) = \mathbf{a}\mathbf{h}\boldsymbol{\theta} = \boldsymbol{\theta} \quad (\forall \boldsymbol{\theta})$ ,

$$\mathbf{a}\mathbf{h} = \mathbf{I}.$$

- Denotando la fila  $i$ -ésima de  $\mathbf{a}$  como  $\mathbf{a}_i^T$  y la columna  $j$ -ésima de  $\mathbf{h}$  como  $\mathbf{h}_j$ , tenemos

$$\mathbf{a} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_p^T \end{bmatrix} \quad \mathbf{h} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_p],$$

podemos re-escribir la condición para estimador insesgado como,

$$\mathbf{a}_i^T \mathbf{h}_j = \delta_{ij}, \quad i = 1, 2, \dots, p; \quad j = 1, 2, \dots, p.$$

# Estimador BLUE: Extensión a vector de parámetros

- Realizando un razonamiento similar al caso escalar, la varianza es

$$\text{var}(\hat{\theta}_i) = \mathbb{E} \left[ \left( \hat{\theta}_i - \mathbb{E}(\hat{\theta}_i) \right)^2 \right] = \mathbb{E} \left[ \left( \mathbf{a}_i^T \mathbf{x} - \mathbb{E}(\mathbf{a}_i^T \mathbf{x}) \right)^2 \right] = \mathbf{a}_i^T \mathbf{C} \mathbf{a}_i$$

- Hay que encontrar  $\mathbf{a}$  de forma de minimizar la varianza manteniendo la condición de insesgado. Para encontrar la fila  $i$ -ésima de  $\mathbf{a}$  hay que resolver,

$$\mathbf{a}_{i*} = \arg \min_{\mathbf{a}_i} \mathbf{a}_i^T \mathbf{C} \mathbf{a}_i \quad \text{sujeto a} \quad \mathbf{a}_i^T \mathbf{h}_j = \delta_{ij} \quad \text{con } (i, j) \in [1, 2, \dots, p]^2.$$

- En este caso tenemos  $p$  restricciones y la función  $i$ -ésima de Lagrange queda,

$$J_i = \mathbf{a}_i^T \mathbf{C} \mathbf{a}_i + \sum_{j=1}^p \lambda_j^{(i)} \left( \mathbf{a}_i^T \mathbf{h}_j - \delta_{ij} \right).$$

- Resolviendo (Kay 1993, Apéndice 6B), se obtienen los coeficientes óptimos,

$$\mathbf{a}_{i*} = \mathbf{C}^{-1} \mathbf{h} (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \mathbf{e}_i,$$

donde  $\mathbf{e}_i = [0, 0, 1, 0, \dots, 0]$ , donde el 1 está en la posición  $i$ -ésima,

- y además,

$$\mathbf{a}_* = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \mathbf{h}^T \mathbf{C}^{-1}.$$

# Estimador BLUE: Extensión a vector de parámetros

- El estimador BLUE en el caso vectorial queda,

$$\hat{\boldsymbol{\theta}} = \mathbf{a}_* \mathbf{x} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \mathbf{h}^T \mathbf{C}^{-1} \mathbf{x}$$

y su matriz de covarianza es

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1}.$$

## Observación

- El estimador BLUE es el mismo que el estimador MVU para el modelo lineal general

$$\mathbf{x} = \mathbf{h}\boldsymbol{\theta} + \mathbf{w}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}).$$

¿Por qué?

# Estimador BLUE: Extensión a vector de parámetros

**Teorema de Gauss-Markov.** Si los datos observados  $\mathbf{x}$  tienen la forma del modelo lineal general

$$\mathbf{x} = \mathbf{h}\boldsymbol{\theta} + \mathbf{w},$$

donde

- $\mathbf{x}$ :  $N \times 1$  vector de observaciones
- $\mathbf{h}$ :  $N \times p$  matriz de observación conocida, con  $N \geq p$  y rango  $p$ .
- $\boldsymbol{\theta}$ :  $p \times 1$  vector de parámetros a estimar
- $\mathbf{w}$ :  $N \times 1$  vector de ruido con PDF arbitraria, media nula y covarianza  $\mathbf{C}$ .

Entonces, el estimador BLUE de  $\boldsymbol{\theta}$  es

$$\hat{\boldsymbol{\theta}} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \mathbf{h}^T \mathbf{C}^{-1} \mathbf{x},$$

y la varianza es

$$\text{var}(\hat{\theta}_i) = \left[ (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \right]_{ii}.$$

Además, la matriz de covarianza del estimador es

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1}.$$

# Consideraciones sobre la optimalidad del BLUE

## Modelo lineal con ruido gaussiano

- En el modelo lineal general con ruido gaussiano se encontró que el estimador MVU es

$$\hat{\boldsymbol{\theta}} = (\mathbf{h}^T \mathbf{C}^{-1} \mathbf{h})^{-1} \mathbf{h}^T \mathbf{C}^{-1} \mathbf{x}.$$

**Obs.** El estimador MVU es lineal en los datos.

- Restringir el estimador a ser lineal no conduce a un estimador subóptimo, ya que el MVU pertenece a la clase de operadores lineales.

## Ejemplo: Estimación de nivel de DC en WGN

En ese caso, el MVU es la media muestral

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

que es lineal en los datos.

- Si el modelo es lineal y los datos son Gaussianos, el BLUE es también el MVU.

# Consideraciones sobre la optimalidad del BLUE

## Modelo lineal con ruido no gaussiano

- Si el modelo es lineal pero los datos no son Gaussianos (ruido no Gaussiano), el estimador MVU no es lineal con los datos.

**Ejemplo:** Estimación de DC en ruido blanco **uniforme**. Asumimos,

$$x[n] \sim \mathcal{U}(0, \beta),$$

el parámetro que quiero estimar es  $\theta = \frac{\beta}{2}$  (nivel de DC).

- El MVU, encontrado mediante estadísticos suficientes es (Kay 1993, Cap. 5)

$$\hat{\theta}_{\text{MVU}} = \frac{N+1}{2N} \max x[n],$$

que no es lineal en los datos.

- Restringiendo el estimador a ser lineal, el BLUE es la media muestral.
- La varianza de los estimadores es:

$$\text{var}(\hat{\theta}_{\text{MVU}}) = \frac{\beta^2}{4N(N+2)}$$

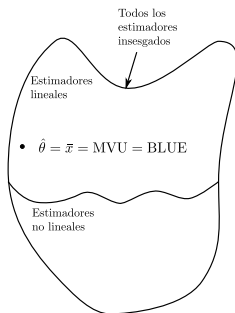
$$\text{var}(\hat{\theta}_{\text{BLUE}}) = \frac{\beta^2}{12N}$$

- Si el modelo es lineal y los datos no son Gaussianos, el BLUE es subóptimo.

# Consideraciones sobre la optimalidad del BLUE

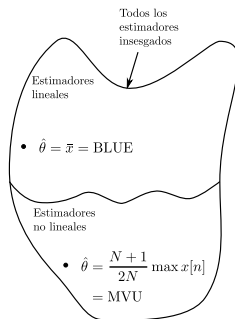
Nivel de DC en WGN

El BLUE es óptimo



Media en ruido uniforme

El BLUE es subóptimo



- Si el MVU pertenece a la clase de estimadores lineales, no se pierde desempeño con el BLUE.
- Si el MVU pertenece a la clase de estimadores no lineales, hay pérdida de desempeño (puede ser significativa).



# Ejemplo III: Transformación de los datos

## Ejemplo: Estimación de la varianza en WGN

Se tienen las observaciones,

$$x[n] \sim \mathcal{N}(0, \sigma^2) \quad n = 0, 1, \dots, N-1,$$

y se quiere estimar  $\sigma^2$ .

- Empleando Cramér-Rao, se puede demostrar que el estimador MVU (que además es estimador eficiente) es:

$$\sigma_{\text{MVU}}^2 = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n].$$

- **Obs.** El MVU no es lineal en los datos.
- Se intenta calcular el BLUE en este caso, se tiene que

$$\widehat{\sigma^2} = \sum_{n=0}^{N-1} a_n x[n],$$

con lo cual

$$\mathbb{E}(\widehat{\sigma^2}) = \sum_{n=0}^{N-1} a_n \mathbb{E}(x[n]) = 0$$

- Cualquier estimador lineal es totalmente inapropiado en este caso (estimador lineal siempre es sesgado).

# Ejemplo III: Transformación de los datos

## Ejemplo: Estimación de la varianza en WGN

- Sin embargo, el BLUE puede calcularse a partir de los datos transformados como  $y[n] = x^2[n]$ ,

$$\widehat{\sigma^2} = \sum_{n=0}^{N-1} a_n y[n] = \sum_{n=0}^{N-1} a_n x^2[n].$$

- La restricción de insesgado puede cumplirse, ya que,

$$\mathbb{E}(\widehat{\sigma^2}) = \sum_{n=0}^{N-1} a_n \mathbb{E}(x^2[n]) = \sum_{n=0}^{N-1} a_n \sigma^2 = \sigma^2.$$

- Teniendo en cuenta que  $\mathbb{E}(x^2[n]) = \sigma^2$ , el modelo con los datos transformados puede expresarse como

$$\begin{aligned} y[n] = x^2[n] &= \underbrace{\mathbb{E}(x^2[n])}_{\sigma^2} + \underbrace{(x^2[n] - \mathbb{E}(x^2[n]))}_{w[n]} \\ &= \sigma^2 + w[n] \end{aligned}$$

# Ejemplo III: Transformación de los datos

## Ejemplo: Estimación de la varianza en WGN

- Ahora el modelo es,

$$y[n] = \sigma^2 + w[n], \quad \text{con } w[n] = x^2[n] - \sigma^2.$$

- La transformación de los datos hace que el problema se reduzca a un problema de estimación de amplitud de una señal en ruido.
- Usando las ecuaciones derivadas para el estimador BLUE con  $s = 1$ , se obtiene,

$$\widehat{\sigma^2} = \frac{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{y}}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}} \quad \text{var}(\widehat{\sigma^2}) = \frac{1}{\mathbf{1}^T \mathbf{C}^{-1} \mathbf{1}}$$

- Debemos calcular la matriz de covarianza del ruido  $w[n]$ .
- Por construcción del modelo transformado, la media  $w[n]$  es nula,

$$\mathbb{E}(w[n]) = \mathbb{E}(x^2[n]) - \sigma^2 = 0.$$

- El elemento (i,j) de la matriz de covarianza es

$$\mathbf{C}_{ij} = \mathbb{E}(w[i]w[j]) = \begin{cases} \mathbb{E}(w^2[i]) & i = j \\ \mathbb{E}(w[i])\mathbb{E}(w[j]) = 0 & i \neq j \end{cases}$$

# Ejemplo III: Transformación de los datos

## Ejemplo: Estimación de la varianza en WGN

- La matriz de covarianza es diagonal, y sus elementos de diagonal valen

$$\begin{aligned}[\mathbf{C}]_{ii} &= \mathbb{E}(w^2[i]) \\&= \mathbb{E}((x^2[n] - \sigma^2)^2) \\&= \mathbb{E}(x^4[n]) - 2\sigma^2\mathbb{E}(x^2[n]) + \mathbb{E}(x^2[n]) \\&= 3\sigma^4 - 2\sigma^4 + \sigma^4 \\&= 2\sigma^4.\end{aligned}$$

Con lo cual

$$\mathbf{C} = 2\sigma^4\mathbf{I} \quad \text{y su inversa, } \mathbf{C}^{-1} = \frac{1}{2\sigma^4}\mathbf{I}.$$

- El estimador y su varianza son,

$$\begin{aligned}\widehat{\sigma^2} &= \frac{\mathbf{1}^T \frac{1}{2\sigma^4} \mathbf{I} \mathbf{y}}{\mathbf{1}^T \frac{1}{2\sigma^4} \mathbf{I} \mathbf{1}} = \frac{1}{N} \sum_{n=0}^{N-1} y[n] = \frac{1}{N} \sum_{n=0}^{N-1} x^2[n] \\ \text{var}(\widehat{\sigma^2}) &= \frac{1}{\mathbf{1}^T \frac{1}{2\sigma^4} \mathbf{I} \mathbf{1}} = \frac{2\sigma^4}{N}.\end{aligned}$$

- Alcanza la CRLB (Kay 1993, Cap. 3).

- **Kay, S. M.** (1993)  
*Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory*, Capítulo 6.