Advanced Control Systems
Detection, Estimation, and Filtering

Graduate Course on the
MEng PhD Program
Spring 2012/2013

Chapter 4
Linear Models in the Presence of Stochastic Signals

Instructor:
Prof. Paulo Jorge Oliveira
p.oliveira@dem.ist.utl.pt or pjcro @ isr.ist.utl.pt
Phone: 21 8419511 or 21 8418053 (3511 or 2053 inside IST)

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Syllabus:

Classical Estimation Theory

…

Chap. 3 - ***Cramer-Rao Lower Bound*** [1 week]
Estimator accuracy; Cramer-Rao lower bound (CRLB); CRLB for signals in white Gaussian noise; Examples;

Chap. 4 - ***Linear Models in the Presence of Stochastic Signals*** [1 week] Stationary and transient analysis; White Gaussian noise and linear systems; Examples; Sufficient Statistics; Relation with MVU Estimators;

Chap. 5 - ***Best Linear Unbiased Estimators*** [1 week]
Definition of BLUE estimators; White Gaussian noise and bandlimited systems; Examples; Generalized minimum variance unbiased estimation;

continues…

A very special class of systems:

FACT:

The determination of the MVU Estimator is in general a difficult task.

A class of systems that allows the determination of this estimator easily…

LINEAR SYSTEMS

The statistical performance is also easy to compute

and an efficient solution is obtained.

The key point is on the formulation of a problem as a linear one.

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

MVU Estimator for the Linear Model:

***Theorem 4.1** – If the data observed can be modeled as*

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

where $\mathbf{x}$ *is a N x 1 vector of observations,* $\mathbf{H}$ *is a known N x p observation matrix (with N>p) and rank p,* $\boldsymbol{\theta}$ *is a p x 1 vector of parameters to be estimated, and* $\mathbf{w}$ *is an N x 1 noise vector with PDF N(0,* $\sigma^2\mathbf{I}$*), then the MVU is*

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{x} \qquad (1)$$

and the covariance matrix of estimate is

$$C_{\hat{\theta}} = \sigma^2\left(\mathbf{H}^T\mathbf{H}\right)^{-1}. \qquad (2)$$

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

MVU Estimator for the Linear Model:

Proof outline:

As discussed in Chapter 3, it is possible to determine the MVU estimator if the equality constraints of the CRLB are satisfied.

From the signal model, it follows that the log-likelihood function is

$$\ln p(\mathbf{x};\boldsymbol{\theta}) = -\ln\left(2\pi\sigma^2\right)^{N/2} - \frac{(\mathbf{x}-\mathbf{H}\boldsymbol{\theta})^T(\mathbf{x}-\mathbf{H}\boldsymbol{\theta})}{2\sigma^2}$$

And

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

Using the relations (deduce them, good exercise…)

$$\frac{\partial \mathbf{b}^T\boldsymbol{\theta}}{\partial \boldsymbol{\theta}} = \mathbf{b} \qquad \frac{\partial \boldsymbol{\theta}^T\mathbf{A}\boldsymbol{\theta}}{\partial \boldsymbol{\theta}} = 2\mathbf{A}\boldsymbol{\theta}$$

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

MVU Estimator for the Linear Model:

Proof outline (cont):

It follows

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

Under the assumptions of the theorem, $\mathbf{H}^T\mathbf{H}$ is invertible

$$\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta} = \frac{\mathbf{H}^T\mathbf{H}}{\sigma^2}\left[\left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{x} - \theta\right]. \qquad \left(\frac{\partial \ln p(\mathbf{x};\theta)}{\partial \theta} = \mathbf{I}(\theta)\left[\mathbf{g}(\mathbf{x}) - \theta\right]\right)$$

Note that it is in the format introduced in the previous chapter, from where (1) and (2) follows immediately.

□

Major constraints:

what if $\mathbf{H}^T\mathbf{H}$ is not invertible?

what if $\mathbf{H}^T\mathbf{H}$ is ill-conditioned?

Example - Fourier Analysis:

Cyclic components in white Gaussian noise

Signal model:

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

Defining

$$\boldsymbol{\theta} = \begin{bmatrix} a_1 & \dots & a_M & b_1 & \dots & b_M \end{bmatrix}^T, \mathbf{w} = \begin{bmatrix} w_0 & & w_{N-1} \end{bmatrix}^T, \text{and}$$

$$\mathbf{H} = \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ c\left(\frac{2\pi}{N}\right) & \dots & c\left(\frac{2\pi M}{N}\right) & s\left(\frac{2\pi}{N}\right) & \dots & s\left(\frac{2\pi M}{N}\right) \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ c\left(\frac{2\pi(N-1)}{N}\right) & \dots & c\left(\frac{2\pi M(N-1)}{N}\right) & s\left(\frac{2\pi(N-1)}{N}\right) & \dots & s\left(\frac{2\pi M(N-1)}{N}\right) \end{bmatrix}$$

The model can be reformulated as a linear system, with solution if $M < N/2$

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

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Example - Fourier Analysis (cont.):

Important fact: The columns of $\mathbf{H}$ are orthogonal.

Define

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \cdots & \mathbf{h}_{2M} \end{bmatrix}, \qquad \text{it follows} \quad \mathbf{h}_i^T \mathbf{h}_j = 0, \qquad i \neq j \quad .$$

Moreover, the discrete Fourier Transform (DFT) relations can be applied, i.e.

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi in}{N}\right) \cos\left(\frac{2\pi jn}{N}\right) = \frac{N}{2}\delta_{ij}$$

$$\sum_{n=0}^{N-1} \sin\left(\frac{2\pi in}{N}\right) \sin\left(\frac{2\pi jn}{N}\right) = \frac{N}{2}\delta_{ij}$$

$$\sum_{n=0}^{N-1} \cos\left(\frac{2\pi in}{N}\right) \sin\left(\frac{2\pi jn}{N}\right) = 0, \qquad \text{for all} \quad i, j.$$

From where it follows

$$\mathbf{H}^T \mathbf{H} = \frac{N}{2}\mathbf{I}.$$

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Example - Fourier Analysis (cont.):

The MVU estimator is

§

$$\hat{\theta} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{x} = \frac{2}{N}\mathbf{H}^T\mathbf{x} = \begin{bmatrix} \frac{2}{N}h_1^T\mathbf{x} \\ \vdots \\ \frac{2}{N}h_{2M}^T\mathbf{x} \end{bmatrix},$$

or finally

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

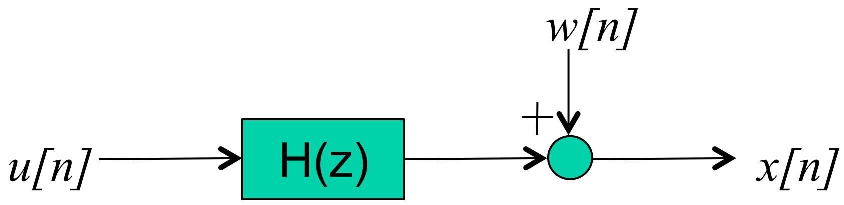
$$\hat{b}_k = \frac{2}{N}\sum_{n=0}^{N-1} x[n]\sin\left(\frac{2\pi kn}{N}\right).$$

with covariance

$$C_{\hat{\boldsymbol{\theta}}} = \frac{2\sigma^2}{N}\mathbf{I}.$$

Example: System Identification

Signal model, where the Finite Impulse Response (FIR) is to be estimated.



The user can apply the input signal u:

$$x[n] = \sum_{k=0}^{p-1} h[k]u[n-k] + w[n], \qquad n = 0,\ldots,N-1, \quad w[n] \sim N(0,\sigma^2) \ .$$

In matrix form, considering $\boldsymbol{x}=[x_0 \ \ldots \ x_{N-1}]^T$, the input/output relations of this linear system can be written as

$$\mathbf{x} = \begin{bmatrix} u[0] & 0 & \ldots & 0 \\ u[1] & u[0] & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ u[N-1] & u[N-2] & \ldots & u[N-p] \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ \vdots \\ h[p-1] \end{bmatrix} + \mathbf{w} \qquad \mathbf{w} = \begin{bmatrix} w_0 & & w_{N-1} \end{bmatrix}^T$$

Or in compact form, once again

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

Example - System Identification (cont.):

The MVU estimator is once again

$$\hat{\boldsymbol{\theta}} = \left(\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{x}, \quad \text{with covariance} \quad C_{\hat{\boldsymbol{\theta}}} = \sigma^2\left(\mathbf{H}^T\mathbf{H}\right)^{-1}.$$

Note that accuracy depends on the input signal applied. How to choose it?

Problem: Choose u[n] to minimize $\operatorname{var}\left(\hat{\theta}_i\right) = \left[\mathbf{C}_{\hat{\boldsymbol{\theta}}}\right]_{ii}, i = 1,\dots,p,$ subject to the constraint that $\sum_{n=0}^{N-1} u[n]$ is fixed.

Introducing the crosscorrelation (autocorrelation)

$$r_{ux}[i] = \frac{1}{N}\sum_{n=0}^{N-1-i} u[n]x[n+i] \qquad \mathbf{H}^T\mathbf{H} = \begin{bmatrix} r_{uu}[0] & r_{uu}[1] & \dots & r_{uu}[p-1] \\ r_{uu}[1] & r_{uu}[0] & \dots & r_{uu}[p-2] \\ \vdots & \vdots & \ddots & \vdots \\ r_{uu}[p-1] & r_{uu}[p-2] & \dots & r_{uu}[0] \end{bmatrix}$$

Choosing a Pseudorandom Noise (PRN) makes this last matrix diagonal

$$C_{\hat{\boldsymbol{\theta}}} = \sigma^2\left(r_{uu}[0]\mathbf{I}\right)^{-1} = \frac{\sigma^2}{r_{uu}[0]}\mathbf{I}.$$

Input Signal Energy

DEM
DEPARTAMENTO DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Extension to non-white Gaussian noise:

Theorem (Generalization of Theorem 4.1) *– If the data observed can be modeled as*

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

where $\mathbf{x}$ *is a N x 1 vector of observations,* $\mathbf{H}$ *is a known N x p observation matrix (with N>p) and rank p,* θ *is a p x 1 vector of parameters to be estimated, and* $\mathbf{w}$ *is an N x 1 colored noise vector with PDF* ***N(0, C)*** *(***C*** $\neq \sigma^2 \mathrm{I}$*), then the MVU is*

$$\hat{\theta} = \left(\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H}\right)^{-1} \mathbf{H}^T \mathbf{C}^{-1} \mathbf{x} \qquad (1bis)$$

and the covariance matrix of estimate is

$$C_{\hat{\theta}} = \left(\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H}\right)^{-1}. \qquad (2bis)$$

Extension to non-white Gaussian noise:

Proof: *The covariance matrix and its inverse are both positive semi-definite. Thus*

$$\mathbf{C}^{-1} = \mathbf{D}^T\mathbf{D}, \qquad \text{where} \qquad \mathbf{D} \in R^{NxN}$$

A noise whitening operation can be performed. For that purpose lets compute the covariance of

$$E\left[(\mathbf{D}w)(\mathbf{D}w)^T\right] = \mathbf{DCD}^T = \mathbf{DD}^{-1}\mathbf{D}^{T^{-1}}\mathbf{D}^T = I.$$

*If we define the new variable **x'** as*

$$\mathbf{x}' = \mathbf{Dx} = \mathbf{DH}\theta + \mathbf{Dw} = \mathbf{H}'\theta + \mathbf{w}'.$$

Applying the usual solution to this linear model (transformed) results in

$$\hat{\theta} = \left(\mathbf{H}'^T\mathbf{H}'\right)^{-1}\mathbf{H}'^T\mathbf{x}' = \left(\mathbf{H}^T\mathbf{D}^T\mathbf{DH}\right)^{-1}\mathbf{H}^T\mathbf{D}^T\mathbf{Dx} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x}$$

For a covariance

$$C_{\hat{\theta}} = \left(\mathbf{H}'^T\mathbf{H}'\right)^{-1} = \left(\mathbf{H}^T\mathbf{D}^T\mathbf{DH}'\right)^{-1} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}'\right)^{-1}.$$

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Sufficient Statistics:

General MVU Estimation:

Assume that the CRLB is not satisfied with equality!

There is no efficient estimator.

How do we find the MVU estimator (if it exists)?

Use the concept of Sufficient Statistics.

Example: To compute the value of a DC signal in noise, given n samples, i=0,…,N-1.

Consider

$$S_1 = \{x[0], x[1], \ldots, x[N-1]\}$$

$$S_2 = \{x[0] + x[1], \ldots, x[N-1]\}$$

$$S_3 = \left\{\sum_{n=0}^{N-1} x[n]\right\}$$

All sets are sufficient since the unknown parameter can be found. S_3 *is the minimal one.*

Sufficient Statistics:

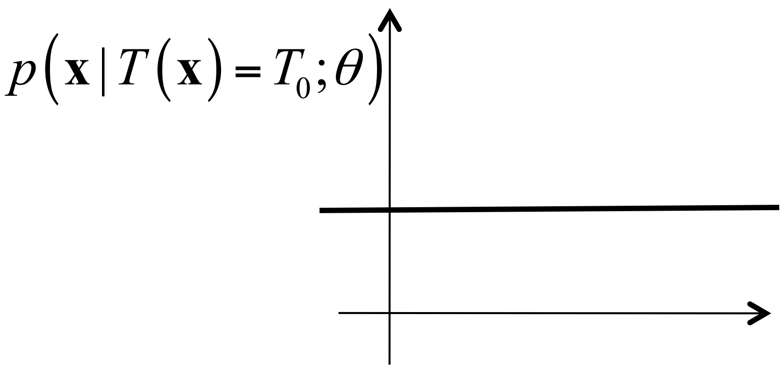
Theorem 5.1 (Neyman-Fisher Factorization) *– If we can factor the PDF p ($\mathbf{x};\boldsymbol{\theta}$) as*

$$p(\mathbf{x};\theta) = g(T(\mathbf{x}),\theta)h(\mathbf{x}) \qquad (3)$$

*where $g(.)$ is a function depending on **x** only through $T(\mathbf{x})$ and $h\,(.)$ is a function depending only on **x**, then $T(\mathbf{x})$ is sufficient statistic for θ. Conversely, if $T\,(\mathbf{x})$ is a sufficient statistic for θ then the PDF can be factored as in (3).*

Proof outline (=>):

- *$p(\mathbf{x},T(\mathbf{x});\theta)$ must have a minimum at $\mathbf{x}=\mathbf{x}_0$, denoted as $T(\mathbf{x}_0)=T_0$;*
- *If $\mathbf{y}=g(\mathbf{x})$, for the vector random variable **x**,* $p(y) = \int p(\mathbf{x})\delta(y - g(\mathbf{x}))d\mathbf{x}.$



- *Knowledge of the value of a sufficient statistics makes the conditional PDF not to depend on the parameters*

Sufficient Statistics:

Proof outline (cont):

Using conditional probability definition:

$$p\left(\mathbf{x} \mid T(\mathbf{x})=T_0;\theta\right)=\frac{p\left(\mathbf{x},T(\mathbf{x})=T_0;\theta\right)}{p\left(T(\mathbf{x})=T_0;\theta\right)}=\frac{p(\mathbf{x};\theta)\delta\left(T(\mathbf{x})-T_0\right)}{p\left(T(\mathbf{x})=T_0;\theta\right)}$$

$$=\frac{g\left(\mathbf{x},T(\mathbf{x})=T_0,\theta\right)h(\mathbf{x})\delta\left(T(\mathbf{x})-T_0\right)}{p\left(T(\mathbf{x})=T_0;\theta\right)}.$$

Where the factorization was used in the last step. The denominator can be written as

$$p\left(T(\mathbf{x})=T_0;\theta\right)=\int p(\mathbf{x};\theta)\delta\left(T(\mathbf{x})-T_0\right)d\mathbf{x}=$$

$$=\int g\left(T(\mathbf{x})=T_0,\theta\right)h(\mathbf{x})\delta\left(T(\mathbf{x})-T_0\right)d\mathbf{x}=g\left(T(\mathbf{x})=T_0,\theta\right)\int h(\mathbf{x})\delta\left(T(\mathbf{x})-T_0\right)d\mathbf{x}.$$

The integral is zero in R^n except over the surface where $T(\mathbf{x})=T_0$. where it is constant.

$$p\left(\mathbf{x} \mid T(\mathbf{x})=T_0;\theta\right)=\frac{h(\mathbf{x})\delta\left(T(\mathbf{x})-T_0\right)}{\int h(\mathbf{x})\delta\left(T(\mathbf{x})-T_0\right)d\mathbf{x}},$$

Which does not depend on θ. Hence, we conclude that $T(\mathbf{x})$ is a sufficient statistic. ■

Sufficient Statistics:

Proof outline (<=):

Consider the joint PDF

$$p\left(\mathbf{x}, T(\mathbf{x}) = T_0; \theta\right) = p\left(\mathbf{x} \mid T(\mathbf{x}) = T_0; \theta\right) p\left(T(\mathbf{x}) = T_0; \theta\right) = p(\mathbf{x}; \theta)\,\delta\left(T(\mathbf{x}) - T_0\right).$$

*Because T(**x**) is a sufficient statistic, the conditional PDF does not depend on θ. We can let*

$$p\left(\mathbf{x} \mid T(\mathbf{x}) = T_0\right) = w(\mathbf{x})\,\delta\left(T(\mathbf{x}) - T_0\right)$$

Substituting in the previous expression

$$p(\mathbf{x}; \theta)\,\delta\left(T(\mathbf{x}) - T_0\right) = w(\mathbf{x})\,\delta\left(T(\mathbf{x}) - T_0\right) p\left(T(\mathbf{x}) = T_0; \theta\right)$$

Setting w(x) to

$$w(\mathbf{x}) = \frac{h(\mathbf{x})}{\int h(\mathbf{x})\,\delta\left(T(\mathbf{x}) - T_0\right) d\mathbf{x}},$$

Allows one to write

$$p(\mathbf{x}; \theta)\,\delta\left(T(\mathbf{x}) - T_0\right) = \frac{h(\mathbf{x})\,\delta\left(T(\mathbf{x}) - T_0\right)}{\int h(\mathbf{x})\,\delta\left(T(\mathbf{x}) - T_0\right) d\mathbf{x}}\, p\left(T(\mathbf{x}) = T_0; \theta\right)$$

Thus based on the factorization a sufficient statistic can be found

$$p(\mathbf{x}; \theta) = g\left(T(\mathbf{x}) = T_0; \theta\right) h(\mathbf{x})$$

Motivating Example: illustration

DC Level in WGN:

The RBLS can be used to find the MVU estimator in two different ways:

1) Find any unbiased estimator of A, say $\breve{A} = x[0]$, and determine $\hat{A} = E[\breve{A} \mid T]$.

The expectation is taken with respect to $p(\breve{A} \mid T)$.

2) Find some function *g(.)* so that $\hat{A} = g(T)$ is an unbiased estimator of *A.*

First approach:

Let $\breve{A} = x[0]$ and determine $\hat{A} = E\left[x[0] \mid \sum_{n=0}^{N-1} x[n]\right]$

We need auxiliary results for $[x\ y]^T$ a Gaussian random vector with mean $\boldsymbol{\mu}=[E[x]\ E[x]]^T$

$$E[x \mid y] = E[x] + \frac{\operatorname{cov}(x, y)}{\operatorname{var}(y)}\left(y - E[y]\right)$$

(see Appendix 10A for details.)

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Motivating Example:

DC Level in WGN (cont.):

1) Find any unbiased estimator of A, say $\breve{A} = x[0]$, and determine $\hat{A} = E[\breve{A} \mid T]$.

The expectation is taken with respect to $p(\breve{A} \mid T)$.

Applying the previous results to *x=x[0]* and $y = \sum_{n=0}^{N-1} x[n]$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x[0] \\ \sum_{n=0}^{N-1} x[n] \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 1 \end{bmatrix}}_{\mathbf{L}} \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix}$$

Hence the PDF of [x y]^T is $N(\boldsymbol{\mu}, C)$, where

$$\boldsymbol{\mu} = \mathbf{L}E[\mathbf{x}] = \mathbf{L}A\mathbf{1} = \begin{bmatrix} A \\ NA \end{bmatrix},$$

$$C = \sigma^2 \mathbf{L}\mathbf{L}^T = \sigma^2 \begin{bmatrix} 1 & 1 \\ 1 & N \end{bmatrix}.$$

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Motivating Example:

DC Level in WGN (cont.):

1) Find any unbiased estimator of A, say $\breve{A} = x[0]$, and determine $\hat{A} = E[\breve{A} \mid T]$.

The expectation is taken with respect to $p(\breve{A} \mid T)$.

Hence we have finally

$$\hat{A} = E[x \mid y] = A + \frac{\sigma^2}{N\sigma^2}\left(\sum_{n=0}^{N-1} x[n] - NA\right) = \frac{1}{N}\sum_{n=0}^{N-1} x[n] = \frac{1}{N}\sum_{n=0}^{N-1} x[n].$$

Which is the MVU estimator. Usually this option is mathematically intractable.

2) Find some function *g(.)* so that $\hat{A} = g(T)$ is an unbiased estimator of *A.*

We need to find some function $\hat{A} = g\left(\sum_{n=0}^{N-1} x[n]\right)$ *so that it is an unbiased estimator.*

That is the case of

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

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

RBLS Theorem:

Definition: a statistic is complete if there is only one function of the statistic that is unbiased.

Theorem 5.1 (Rao-Blackwell-Lehmann-Scheffe) *– If $\breve{\theta}$ is an unbiased estimator of θ and T(**x**) is a sufficient statistic for θ, then $\hat{\theta} = E\left[\breve{\theta} \mid T\left(\mathbf{x}\right)\right]$ is*

1. A valid estimator for θ

2. Unbiased

3. Of lesser or equal variance than that of $\breve{\theta}$, for all θ.

Additionally, if the sufficient statistic is complete, then $\hat{\theta}$ is the MVU estimator.

To validate that a statistic is complete is in general very difficult, (see examples 5.6 and 5.7). It must verify

$$\int_{-\infty}^{+\infty} v\left(T\right) p\left(T;\theta\right) dT = 0, \qquad \text{for all} \quad \theta. \qquad (5.8)$$

Only for the zero function and for *v(T).*

Note: - For an example of an incomplete statistic check Example 5.7

Methodology:

Use Neyman-Fisher factorization theorem (5.1) to find sufficient statistic

↓ $T(\mathbf{x})$

Determine if $T(\mathbf{x})$ is complete see (5.8)

↓

Find function of $T(\mathbf{x})$ that is unbiased

↓

$\hat{\theta} = g\left(T\left(\mathbf{x}\right)\right) = \text{MVU Estimator}$

DEM
DEPARTAMENTO DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Example:

Mean of Uniform Noise:

Data model: *x[n]=w[n], n=0,1,…,N-1*

Where *w[n]* is IID noise with *PDF U[0,β], for β>0.*

We wish to find the MVU estimator for the mean *θ=β/2.*

The approach to find the CRLB can not be followed as the PDF does not satisfy the regularity conditions. A natural estimator is

$$\hat{\theta} = \frac{1}{N}\sum_{n=0}^{N-1} x[n], \qquad with \quad \operatorname{var}(\hat{\theta}) = \frac{1}{N}\operatorname{var}(x[n]) = \frac{\beta^2}{12N}.$$

To determine if the sample mean is the MVU we will follow the methodology previously presented.

Example:

Lets define the unit step function:

$$u(x) = \begin{cases} 1 & for \quad x > 0 \\ 0 & for \quad x < 0 \end{cases} .$$

Then,

$$p(x[n];\beta) = \frac{1}{\beta}\left[u(x[n]) - u(x[n]-\beta)\right], \quad where \quad \beta = 2\theta.$$

and the PDF is

$$p(x[n];\beta) = \frac{1}{\beta^N}\prod_{n=0}^{N-1}\left[u(x[n]) - u(x[n]-\beta)\right] = \begin{cases} \dfrac{1}{\beta^N} & 0 < x[n] < \beta \quad n = 0,1,\ldots,N-1 \\ 0 & otherwise \end{cases} .$$

Alternative, we can write

$$p(x[n];\beta) = \begin{cases} \dfrac{1}{\beta^N} & \max(x[n]) < \beta, \min(x[n]) > 0 \\ 0 & otherwise \end{cases} ,$$

So that

$$p(x[n];\beta) = \frac{1}{\beta^N}u(\beta - \max(x[n]))u(\min(x[n]))$$

Example:

Note that can be identified

$$p(x[n];\beta)=\underbrace{\frac{1}{\beta^N}u(\beta-\max(x[n]))}_{g(T(x),\beta)}\underbrace{u(\min(x[n]))}_{h(x)}$$

By the Neyman-Fisher factorization theorem, *T(**x**)=max(x[n])* is a sufficient statistic for θ. Furthermore, it can be shown that the sufficient statistic is complete. We need next to find a function of *T(**x**)* that is not biased (denominated as order statistics). Lets write the cumulative distribution function

$$\Pr\{T\leq\xi\}=\Pr\{x[0]\leq\xi,x[1]\leq\xi,\ldots,x[N-1]\leq\xi,\}=\prod_{n=0}^{N-1}\Pr\{x[n]\leq\xi\}=\Pr\{x[n]\leq\xi\}^N.$$

The PDF follows as

$$p_T(\xi)=\frac{d\Pr\{T\leq\xi\}}{d\xi}=N\Pr\{x[n]\leq\xi\}^{N-1}\frac{d\Pr\{x[n]\leq\xi\}}{d\xi}.$$

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA

Example:

But
$$\frac{d\Pr\left\{x\left[n\right]\leq\xi\right\}}{d\xi}=p_{x[n]}\left(\xi\right)=\begin{cases}\frac{1}{\beta} & 0<\xi<\beta\\ 0 & otherwise\end{cases},$$

Integrating we obtain

$$p_T\left(\xi\right)=\begin{cases}0 & \xi<0\\ N\left(\frac{\xi}{\beta}\right)^{N-1}\frac{1}{\beta} & 0<\xi<\beta\\ 0 & \xi>\beta\end{cases},\quad and\quad E\left[T\right]=\int_0^{\beta}\xi N\left(\frac{\xi}{\beta}\right)^{N-1}\frac{1}{\beta}d\xi$$

From where it results

$$E\left[T\right]=\frac{N}{N+1}\beta=\frac{2N}{N+1}\theta,\qquad thus\quad \hat{\theta}=\frac{N+1}{2N}T$$ makes the expected value unbiased.

The MVU estimator is
$$\hat{\theta}=\frac{N+1}{2N}\max\left(x[n]\right)$$

with a variance... $\operatorname{var}\left(\hat{\theta}\right)=\dfrac{\beta^2}{4N(N+2)}<<\dfrac{\beta^2}{12N}$ (sample mean var) for large N!

Bibliography:

Further reading

- Thomas Kailath, ***Linear Systems,*** Prentice Hall, 1980.
- Thomas Kailath, Ali Sayed, and Babak Hassibi, ***Linear Estimation,*** Prentice Hall, 2000.
- Harry L. Van Trees, ***Detection, Estimation, and Modulation Theory, Parts I to IV,*** John Wiley, 2001.
- J. Bibby, H. Toutenburg, ***Prediction and Improved Estimation in Linear Models,*** John Wiley, 1977.
- C.Rao, ***Linear Statistical Inference and Its Applications,*** John Wiley, 1973.

DEM
DEPARTAMENTO
DE ENGENHARIA MECÂNICA
TÉCNICO LISBOA