EL7002 - Estimación y Detección

Clase No.1: Introducción

Patricio Parada

Departamento de Ingeniería Eléctrica
Universidad de Chile

Contenidos de la Clase I

Contenidos de la Clase II

Distribución de un Estimador

Resumen y Lecturas

Nota Inicial (1)

Este es el primer curso en este departamento que considera de manera formal y sistemática los problemas de Procesamiento Estadístico de Señales.

Nota Inicial (1)

Este es el primer curso en este departamento que considera de manera formal y sistemática los problemas de Procesamiento Estadístico de Señales.

Aunque los tópicos que cubriremos en este curso han sido tratados parcialmente en otros cursos, nos concentraremos en las técnicas y algoritmos de detección y estimación que puedan ser aplicados en problemas de ingeniería.

Nota Inicial (1)

Este es el primer curso en este departamento que considera de manera formal y sistemática los problemas de Procesamiento Estadístico de Señales.

Aunque los tópicos que cubriremos en este curso han sido tratados parcialmente en otros cursos, nos concentraremos en las técnicas y algoritmos de detección y estimación que puedan ser aplicados en problemas de ingeniería.

El procesamiento estadístico de señales es una subdisciplina del procesamiento de señales dedicado a dos objetivos:

- la extracción de información de señales, posiblemente ruidosas.
- la detección de señales - determinísticas o aleatorias - en presencia de ruido.

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes
- medicina: detección de patologías via señales bioléctricas o imágenes

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes
- medicina: detección de patologías via señales bioléctricas o imágenes
- minería: prospección no invasiva

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes
- medicina: detección de patologías via señales bioléctricas o imágenes
- minería: prospección no invasiva
- comunicaciones digitales

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes
- medicina: detección de patologías via señales bioléctricas o imágenes
- minería: prospección no invasiva
- comunicaciones digitales
- control automático

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes
- medicina: detección de patologías via señales bioléctricas o imágenes
- minería: prospección no invasiva
- comunicaciones digitales
- control automático
- geofísica

Nota Inicial (2)

La relevancia de los temas de este curso va más allá de los ámbitos de la teoría de información y comunicaciones, y abarca áreas de importancia estratégica como:

- defensa: radar y formación de imágenes
- medicina: detección de patologías via señales bioléctricas o imágenes
- minería: prospección no invasiva
- comunicaciones digitales
- control automático
- geofísica
- finanzas

Organización del Curso (1)

El curso está organizado en dos partes relacionadas pero independientes.

Organización del Curso (1)

El curso está organizado en dos partes relacionadas pero independientes.

La primera se focaliza en los problemas de estimación y profundiza en los temas de:

- estimación insesgada (mínima varianza)
- estimadores lineales
- estimación de máxima verosimilitud
- estimación bayesiana
- filtrado estadístico

Organización del Curso (2)

La segunda parte del curso se focaliza en los problemas de detección y modulación asociados a determinar si una señal dada o un patrón dado se encuentra presente en una determinada medición.

Organización del Curso (2)

La segunda parte del curso se focaliza en los problemas de detección y modulación asociados a determinar si una señal dada o un patrón dado se encuentra presente en una determinada medición.

Entre los temas que tratermos se encuentran los siguientes:

- Test de hipótesis (binario y múltiple)
- Detección de señales determinísticas en presencia de ruido
- Detección de señales aleatorias en presencia de ruido

Parte I

Estimación

1.1 Problema de Estimación

Contexto

Antes de comenzar es importante ser precisos respecto del tipo de hipótesis que vamos a asumir ciertas para nuestro análisis.

Contexto

Antes de comenzar es importante ser precisos respecto del tipo de hipótesis que vamos a asumir ciertas para nuestro análisis.

H1. Existe una variable aleatoria, vector aleatorio o proceso estocástico (de tiempo discreto o continuo) X cuya medida de probabilidad $\mathbf{p}(\mathbf{x};\boldsymbol{\theta})$ está parametrizada en términos de un término desconocido $\boldsymbol{\theta}$.

Contexto

Antes de comenzar es importante ser precisos respecto del tipo de hipótesis que vamos a asumir ciertas para nuestro análisis.

H1. Existe una variable aleatoria, vector aleatorio o proceso estocástico (de tiempo discreto o continuo) X cuya medida de probabilidad $\mathbf{p}(\mathbf{x};\boldsymbol{\theta})$ está parametrizada en términos de un término desconocido $\boldsymbol{\theta}$.

H2. Tenemos acceso a realizaciones de esta variable $\{x[0], x[1], \dots, x[N-1]\}$ que siguen la ley de probabilidad $\mathbf{p}(\mathbf{x};\boldsymbol{\theta})$.

El problema de estimación

Problema de Estimación

El problema de estimación corresponde a **construir** una función $g(\mathbf{x})$, con $\mathbf{x} \in \mathbb{R}^N$, tal que

$$\hat{\theta} = g(x[0], x[1], \ldots, x[N-1]) \tag{1}$$

es un estimador de un parámetro desconocido θ, donde el argumento de $g(\cdot)$ es un conjunto de observaciones $\{x[0], x[1], \ldots, x[N-1]\}$ de proceso relacionado con el parámetro.

¿Cómo sé si un estimador es "bueno"? (1)

Este es un problema central en teoría de estimación: bajo que circunstancias un estimador es es bueno, regular o malo.

¿Cómo sé si un estimador es “bueno”? (1)

Este es un problema central en teoría de estimación: bajo que circunstancias un estimador es es bueno, regular o malo.

En general, la medida de desempeño que se emplea en este tipo de problemas es de cercanía entre el valor real del parámetro y el de la estimación.

¿Cómo sé si un estimador es “bueno”? (1)

Este es un problema central en teoría de estimación: bajo que circunstancias un estimador es es bueno, regular o malo.

En general, la medida de desempeño que se emplea en este tipo de problemas es de cercanía entre el valor real del parámetro y el de la estimación.

Por ejemplo, la Ley de los Grandes Números (versión débil) establece que “dado un conjunto de realizaciones de una variable aleatoria $X \sim P$, con $\mathbb{E}[X] = \mu$, y realizaciones i.i.d., $x_0, x_1, \ldots, x_{N-1}$, entonces

$$\hat{\mu}_N \equiv \frac{1}{N} \sum_{n=0}^{N-1} x_n \tag{2}$$

¿Cómo sé si un estimador es "bueno"? (2)

es un estimador de μ en el sentido que

$$\lim_{N\to\infty} \hat{\mu}_N = \mu$$

¿Cómo sé si un estimador es “bueno”? (2)

es un estimador de μ en el sentido que

$$\lim_{N\to\infty} \hat{\mu}_N = \mu$$

o bien

$$\lim_{N\to\infty} |\hat{\mu}_N - \mu| = 0.$$

Criterios de Calidad para la Estimación

En muchas ocasiones, la calidad de un estimador está íntimamente ligada a un criterio arbitrario de selección del estimador.

Criterios de Calidad para la Estimación

En muchas ocasiones, la calidad de un estimador está íntimamente ligada a un criterio arbitrario de selección del estimador.

Los criterios más empleados son los siguientes:

- En norma p:

$$\lim_{N\to\infty} \mathbb{E}[||\hat{\theta}_N - \theta||^p] = 0, \ p \in [1, \infty[. \tag{3a}$$

Criterios de Calidad para la Estimación

En muchas ocasiones, la calidad de un estimador está íntimamente ligada a un criterio arbitrario de selección del estimador.

Los criterios más empleados son los siguientes:

- En norma p:

$$\lim_{N\to\infty} \mathbb{E}[\|\hat{\theta}_N - \theta\|^p] = 0, \; p \in [1, \infty[. \tag{3a}$$

- En probabilidad (en medida)

$$\lim_{N\to\infty} \Pr[\|\hat{\theta}_N - \theta\| < \epsilon] = 1, \; \forall\epsilon > 0. \tag{3b}$$

Criterios de Calidad para la Estimación

En muchas ocasiones, la calidad de un estimador está íntimamente ligada a un criterio arbitrario de selección del estimador.

Los criterios más empleados son los siguientes:

- En norma p:

$$\lim_{N\to\infty} \mathbb{E}[\|\hat{\theta}_N - \theta\|^p] = 0, \ p \in [1, \infty[. \tag{3a}$$

- En probabilidad (en medida)

$$\lim_{N\to\infty} \Pr[\|\hat{\theta}_N - \theta\| < \epsilon] = 1, \ \forall \epsilon > 0. \tag{3b}$$

- En discriminante de Kullback-Leibler

$$\lim_{N\to\infty} D(p(\mathbf{x};\theta)\|p(\mathbf{x};\hat{\theta})) = 0. \tag{3c}$$

Modelos para los Datos

El primer paso para construir un estimador es conocer, aplicar o forzar un modelo a los datos observados.

Modelos para los Datos

El primer paso para construir un estimador es conocer, aplicar o forzar un modelo a los datos observados.

Dado que los datos son naturalmente aleatorios (en el caso determinístico el problema es trivial), vamos a utilizar un modelo de la densidad de probabilidad

$$p(x[0], x[1], \ldots, x[N-1]; \theta). \tag{4}$$

Modelos para los Datos

El primer paso para construir un estimador es conocer, aplicar o forzar un modelo a los datos observados.

Dado que los datos son naturalmente aleatorios (en el caso determinístico el problema es trivial), vamos a utilizar un modelo de la densidad de probabilidad

$$p(x[0], x[1], \ldots, x[N-1]; \theta). \tag{4}$$

Al parametrizar la función de densidad de probabilidad (f.d.p.) en términos de θ lo que estamos haciendo es considerar la familia de distribuciones con $\theta \in \Theta$.

Modelos para los Datos

El primer paso para construir un estimador es conocer, aplicar o forzar un modelo a los datos observados.

Dado que los datos son naturalmente aleatorios (en el caso determinístico el problema es trivial), vamos a utilizar un modelo de la densidad de probabilidad

$$p(x[0], x[1], \ldots, x[N-1]; \theta). \tag{4}$$

Al parametrizar la función de densidad de probabilidad (f.d.p.) en términos de θ lo que estamos haciendo es considerar la familia de distribuciones con $\theta \in \Theta$.

Ejemplo

$$p(x[0]; \theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(x[0]-\theta)^2\right)$$

Ejemplo: Actividad Bursátil (1)

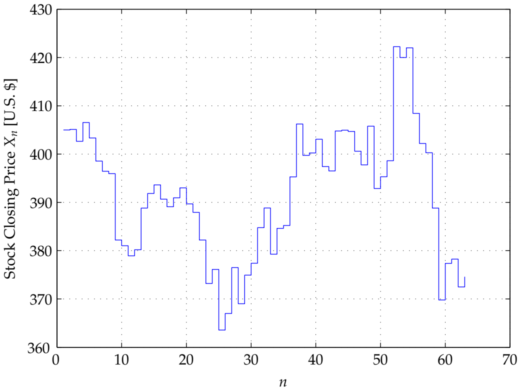

Figura: Precios al cierre del día de la acción de Apple Inc., código AAPL en NYS Exchange. Período: 01.Oct.11 - 31.Dic.11.

Ejemplo: Actividad Bursátil (2)

Podemos aplicar el siguiente modelo a los datos:

$$X_n = a + bn + W_n \tag{5}$$

donde a y b son constantes (desconocidas) y $\{W_n\}$ es un proceso gaussiano i.i.d.

Ejemplo: Actividad Bursátil (2)

Podemos aplicar el siguiente modelo a los datos:

$$X_n = a + bn + W_n \tag{5}$$

donde a y b son constantes (desconocidas) y $\{W_n\}$ es un proceso gaussiano i.i.d.

En este caso, los parámetros del modelo son $\boldsymbol{\theta} = [a;\ b]^T$ y

$$p(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{N/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x_n - a - bn)^2\right) \tag{6}$$

donde $\mathbf{x} = (x_0, x_1, \ldots, x_{N-1})$ es un conjunto de observaciones (precios de cierre en este caso).

Algo más sobre modelos

En general, la calidad o flexibilidad de un estimador estará intimamente ligada a la familia de f.d.p.'s que uno haya considerado.

Algo más sobre modelos

En general, la calidad o flexibilidad de un estimador estará intimamente ligada a la familia de f.d.p.'s que uno haya considerado.

En el caso del ejemplo, sólo consideramos distribuciones gaussianas con un valor medio dependiente del tiempo.

Algo más sobre modelos

En general, la calidad o flexibilidad de un estimador estará intimamente ligada a la familia de f.d.p.'s que uno haya considerado.

En el caso del ejemplo, sólo consideramos distribuciones gaussianas con un valor medio dependiente del tiempo.

Importante

El tipo de estimadores que uno desea construir es robusto, en el sentido que pequeños cambios en la f.d.p. generan pequeñas degradaciones en el desempeño y no cambios abruptos.

Estimación clásica (1)

La estimación basada en el cálculo de parámetros de un modelo recibe el nombre de **estimación clásica**.

Estimación clásica (1)

La estimación basada en el cálculo de parámetros de un modelo recibe el nombre de **estimación clásica**.

En ella, $\boldsymbol{\theta}$ es un vector de parámetros determinístico pero desconocido.

Estimación clásica (1)

La estimación basada en el cálculo de parámetros de un modelo recibe el nombre de **estimación clásica**.

En ella, $\boldsymbol{\theta}$ es un vector de parámetros determinístico pero desconocido.

Cuando se adquiere información adicional sobre el parámetro (via observaciones), se puede asumir que $\boldsymbol{\theta}$ es aleatorio y que los parámetros que estamos buscando son una realización particular de un determinado proceso.

Estimación clásica (1)

La estimación basada en el cálculo de parámetros de un modelo recibe el nombre de **estimación clásica**.

En ella, $\boldsymbol{\theta}$ es un vector de parámetros determinístico pero desconocido.

Cuando se adquiere información adicional sobre el parámetro (via observaciones), se puede asumir que $\boldsymbol{\theta}$ es aleatorio y que los parámetros que estamos buscando son una realización particular de un determinado proceso.

Notemos que en este caso, $\boldsymbol{\theta}$ y $\mathbf{x}$ se relacionan mediante la ecuación

$$p(\mathbf{x}, \boldsymbol{\theta}) = p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \tag{7}$$

Estimación clásica (2)

donde $p(\boldsymbol{\theta})$ es la información *a priori* que se conoce de $\boldsymbol{\theta}$ y $p(\mathbf{x}|\boldsymbol{\theta})$ es la verosimilitud de $\mathbf{x}$ condicional al valor de $\boldsymbol{\theta}$.

Estimación clásica (2)

donde $p(\boldsymbol{\theta})$ es la información *a priori* que se conoce de $\boldsymbol{\theta}$ y $p(\mathbf{x}|\boldsymbol{\theta})$ es la verosimilitud de $\mathbf{x}$ condicional al valor de $\boldsymbol{\theta}$.

Este tipo de enfoque se conoce como **estimación bayesiana**.

Estimación clásica (2)

donde $p(\boldsymbol{\theta})$ es la información *a priori* que se conoce de $\boldsymbol{\theta}$ y $p(\mathbf{x}|\boldsymbol{\theta})$ es la verosimilitud de $\mathbf{x}$ condicional al valor de $\boldsymbol{\theta}$.

Este tipo de enfoque se conoce como **estimación bayesiana**.

Notar que

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int\limits_{\Theta} p(\mathbf{x}|\boldsymbol{\theta}')p(\boldsymbol{\theta}')d\boldsymbol{\theta}'} \tag{8}$$

Estimación clásica (2)

donde $p(\boldsymbol{\theta})$ es la información *a priori* que se conoce de $\boldsymbol{\theta}$ y $p(\mathbf{x}|\boldsymbol{\theta})$ es la verosimilitud de $\mathbf{x}$ condicional al valor de $\boldsymbol{\theta}$.

Este tipo de enfoque se conoce como **estimación bayesiana**.

Notar que

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{x})} = \frac{p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{\int\limits_{\Theta} p(\mathbf{x}|\boldsymbol{\theta}')p(\boldsymbol{\theta}')d\boldsymbol{\theta}'} \tag{8}$$

Luego, podemos decir que un estimador es una regla que asigna un valor $\boldsymbol{\theta}$ por cada realización $\mathbf{x}$.

Desempeño de un Estimador (1)

La calidad de un estimador es extremadamente importante para determinar el mecanismo de cálculo de

$$\hat{\boldsymbol{\theta}} = g(x[0], x[1], \ldots, x[N-1]), \tag{9}$$

así como su calidad generalizadora.

Desempeño de un Estimador (1)

La calidad de un estimador es extremadamente importante para determinar el mecanismo de cálculo de

$$\hat{\boldsymbol{\theta}} = g(x[0], x[1], \ldots, x[N-1]), \tag{9}$$

así como su calidad generalizadora.

La calidad de un estimador la vamos a medir definiendo una función de cercanía entre $\boldsymbol{\theta}$ y $\hat{\boldsymbol{\theta}}$, lo que nos permitirá en algunos casos prácticos, establecer reglas para construir mejores estimadores.

Desempeño de un Estimador (2)

Por ejemplo, si un conjunto de valores puede ser descrito mediante el modelo

$$x[n] = A + w[n] \tag{10}$$

con $\{w[n]\}$ un proceso guassiano, entonces es posible determinar A como:

$$\hat{A}_1(N) \equiv \frac{1}{N}\sum_{n=0}^{N-1} x[n]. \tag{11}$$

Desempeño de un Estimador (2)

Por ejemplo, si un conjunto de valores puede ser descrito mediante el modelo

$$x[n] = A + w[n] \tag{10}$$

con $\{w[n]\}$ un proceso guassiano, entonces es posible determinar A como:

$$\hat{A}_1(N) \equiv \frac{1}{N}\sum_{n=0}^{N-1} x[n]. \tag{11}$$

Pero puedo emplear un segundo estimador

$$\hat{A}_2(N) = x[0]. \tag{12}$$

Desempeño de un Estimador (3)

Si $\hat{A}_2$ es más cercano que $\hat{A}_1$ en un caso particular, ¿es suficiente para concluir que $\hat{A}_2$ es mejor estimador que $\hat{A}_1$?

Desempeño de un Estimador (3)

Si $\hat{A}_2$ es más cercano que $\hat{A}_1$ en un caso particular, ¿es suficiente para concluir que $\hat{A}_2$ es mejor estimador que $\hat{A}_1$?

No: debemos conocer el desempeño en un conjunto de realizaciones, no una o unas pocas.

Desempeño de un Estimador (3)

Si $\hat{A}_2$ es más cercano que $\hat{A}_1$ en un caso particular, ¿es suficiente para concluir que $\hat{A}_2$ es mejor estimador que $\hat{A}_1$?

No: debemos conocer el desempeño en un conjunto de realizaciones, no una o unas pocas.

Por ejemplo, si $A = 1$ y $w[n]$ sigue una distribución normal $\mathcal{N}(0, 1)$ entonces podemos calcular varias realizaciones de la secuencia $\{x[0], \ldots, x[N-1]\}$ y determinar la dispersión de ambos estimadores para determinar cual de los dos entrega, en promedio, una mejor estimación.

Desempeño de un Estimador (4)

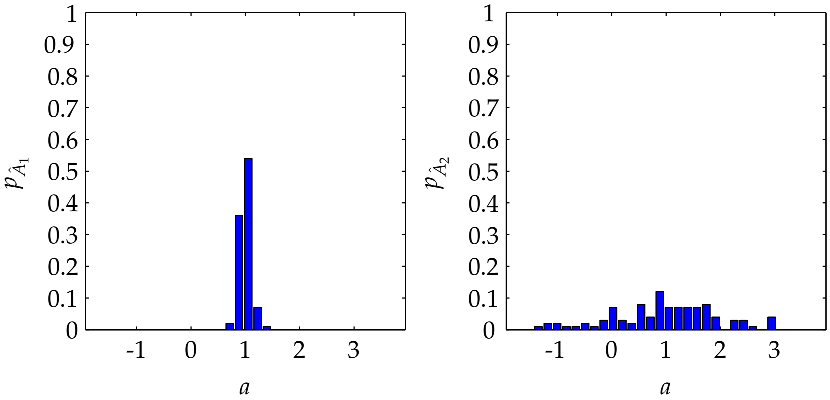

Figura: Dispersión de los estimadores $\hat{A}_1$ y $\hat{A}_2$ asumiendo que $w[n] \sim \mathcal{N}(0,1)$, $N = 100$ y con un total de $S = 100$ simulaciones.

Varianza del Estimador (1)

En muchas ocasiones, resulta conveniente estudiar la varianza de un estimador como una manera de comparar su grado de cercanía al parámetro bajo estudio.

Varianza del Estimador (1)

En muchas ocasiones, resulta conveniente estudiar la varianza de un estimador como una manera de comparar su grado de cercanía al parámetro bajo estudio.

En el caso del estimador $\hat{A}_1$ tenemos que su valor esperado es

$$\begin{aligned}\mathbb{E}[\hat{A}_1] &= \frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}[x[n]] \\ &= \frac{1}{N}\sum_{n=0}^{N-1}(A + \underbrace{\mathbb{E}[w[n]]}_{0}) \\ &= \frac{NA}{N} \\ &= A.\end{aligned}$$

Varianza del Estimador (2)

Similarmente, en el caso del estimador $\hat{A}_2$ su valor esperado es

$$\begin{aligned}\mathbb{E}[\hat{A}_2] &= \mathbb{E}[x[n]] \\ &= A + \underbrace{\mathbb{E}[w[n]]}_{0}) \\ &= A.\end{aligned}$$

Varianza del Estimador (2)

Similarmente, en el caso del estimador $\hat{A}_2$ su valor esperado es

$$\begin{aligned}\mathbb{E}[\hat{A}_2] &= \mathbb{E}[x[n]] \\ &= A + \underbrace{\mathbb{E}[w[n]]}_{0}) \\ &= A.\end{aligned}$$

Ambos estimadores se dicen que son **insesgados**.

Varianza del Estimador (2)

Similarmente, en el caso del estimador $\hat{A}_2$ su valor esperado es

$$\begin{aligned}
\mathbb{E}[\hat{A}_2] &= \mathbb{E}[x[n]] \\
&= A + \underbrace{\mathbb{E}[w[n]])}_{0} \\
&= A.
\end{aligned}$$

Ambos estimadores se dicen que son **insesgados**.

¿Qué hay de sus varianzas?:

$$\begin{aligned}
\mathrm{Var}(\hat{A}_1) &= \mathrm{Var}(\frac{1}{N}\sum_{n=0}^{N-1} x[n]) = \frac{1}{N^2}\sum_{n=0}^{N-1}\mathrm{Var}(x[n]) \\
&= \frac{\sigma^2}{N}
\end{aligned}$$

Varianza del Estimador (3)

Por otro lado,

$$\mathrm{Var}(\hat{A}_2) = \mathrm{Var}(x[0]) = \sigma^2 > \mathrm{Var}(\hat{A}_1)$$

Este resultado es válido independiente del tipo de distribución que siga $\{w[n]\}$.

Varianza del Estimador (3)

Por otro lado,

$$\mathrm{Var}(\hat{A}_2) = \mathrm{Var}(x[0]) = \sigma^2 > \mathrm{Var}(\hat{A}_1)$$

Este resultado es válido independiente del tipo de distribución que siga $\{w[n]\}$.

Lo único que necesitamos es que

$$\mathbb{E}[w[n]] = 0$$
$$\mathrm{Cov}(w[n], w[m]) = \sigma^2 \delta_{n,m}.$$

1.2 Estimador de Mínima Varianza

Hoja de Ruta

A partir de este momento comenzaremos a ver criterios y técnicas para determinar buenos estimadores de un parámetro (o conjunto de ellos) que son desconocidos.

Hoja de Ruta

A partir de este momento comenzaremos a ver criterios y técnicas para determinar buenos estimadores de un parámetro (o conjunto de ellos) que son desconocidos.

El criterio que veremos en esta clase se denomina de varianza mínima, ya que utiliza como función discriminadora la varianza del estimador bajo estudio.

Hoja de Ruta

A partir de este momento comenzaremos a ver criterios y técnicas para determinar buenos estimadores de un parámetro (o conjunto de ellos) que son desconocidos.

El criterio que veremos en esta clase se denomina de varianza mínima, ya que utiliza como función discriminadora la varianza del estimador bajo estudio.

Vamos a considerar una clase particular de estimadores denominados **insesgados**, que tienen un buen desempeño y son realizables (causales).

Estimadores Insesgados (1)

Un estimador es insesgado si - en promedio - entrega el valor correcto del parámetro desconocido.

Estimadores Insesgados (1)

Un estimador es insesgado si - en promedio - entrega el valor correcto del parámetro desconocido.

Si $\boldsymbol{\theta} \in \Theta$ entonces podemos restringir nuestra atención al caso en que

$$\mathbb{E}[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}, \ \boldsymbol{\theta} \in \Theta. \tag{13}$$

Estimadores Insesgados (1)

Un estimador es insesgado si - en promedio - entrega el valor correcto del parámetro desconocido.

Si $\boldsymbol{\theta} \in \Theta$ entonces podemos restringir nuestra atención al caso en que

$$\mathbb{E}[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}, \ \boldsymbol{\theta} \in \Theta. \tag{13}$$

Ejemplo 1: consideremos el problema de estimar un parámetro desconocido $A \in \mathbb{R}$ a partir del modelo

$$x[n] = A + w[n] \tag{14}$$

donde $\{w[n]\}$ es i.i.d. $\sim \mathcal{N}(0, 1)$.

Estimadores Insesgados (2)

Asumamos que tenemos N observaciones de $x[n]$ de forma tal que definimos el estimador

$$\hat{A} = \frac{1}{N} \sum_{n=0}^{N-1} x[n] \tag{15}$$

Es fácil mostrar que $\hat{A}$ es insesgado, ya que

$$\begin{aligned} \mathbb{E}[\hat{A}] &= \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}[x[n]] \\ &= \frac{1}{N} \sum_{n=0}^{N-1} A \\ &= A. \end{aligned}$$

Estimadores Insesgados (3)

Ejemplo 2: Considere $X \sim \mathcal{N}(0, \sigma^2)$ y las realizaciones $x[0], x[1], \ldots, x[N-1]$. Definimos el estimador de la varianza de X como:

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{n=0}^{N-1} x^2[N] \tag{16}$$

Luego,

$$\begin{aligned} \mathbb{E}[\hat{\sigma}^2] &= \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}[x^2[N]] \\ &= \sigma^2. \end{aligned}$$

Estimadores Insesgados (4)

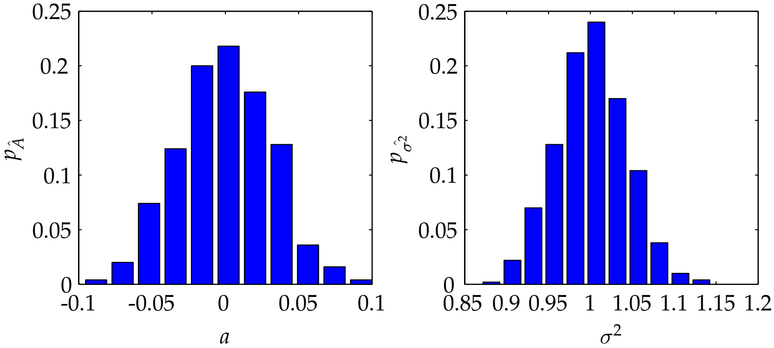

Figura: Distribución empírica (histograma) de los estimadores $\hat{A}$ y $\hat{\sigma^2}$ producidos via simulación con $N = 1000$ y 500 simulaciones; los parámetros reales son $A = 0$ y $\sigma^2 = 1$.

Distribución de un Estimador (1)

En general, si conocemos el modelo paramétrico que relaciona las observaciones con el parámetro desconocido, podemos hacer algo de álgebra y determinar en forma analítica la distribución de probabilidad del parámetro en cuestión.

Distribución de un Estimador (1)

En general, si conocemos el modelo paramétrico que relaciona las observaciones con el parámetro desconocido, podemos hacer algo de álgebra y determinar en forma analítica la distribución de probabilidad del parámetro en cuestión.

Si el estimador es insesgado, es habitual que la distribución del parámetro sea simétrica y centrada en torno al valor real del parámetro.

Distribución de un Estimador (1)

En general, si conocemos el modelo paramétrico que relaciona las observaciones con el parámetro desconocido, podemos hacer algo de álgebra y determinar en forma analítica la distribución de probabilidad del parámetro en cuestión.

Si el estimador es insesgado, es habitual que la distribución del parámetro sea simétrica y centrada en torno al valor real del parámetro.

Sin embargo, existen casos donde esto no se cumple:

$$\bar{\sigma}^2 = \frac{1}{2}(x^2[0] + x^2[1]) \tag{17}$$

se distribuye en forma exponencial con parámetro $1/\sigma^2$:

$$f(\bar{\sigma}^2) = \frac{1}{\sigma^2} \exp\left(- \frac{\bar{\sigma}^2}{\sigma^2} \right). \tag{18}$$

Distribución de un Estimador (2)

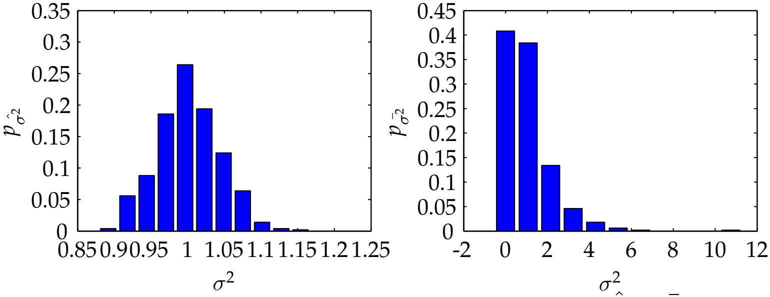

Figura: Distribución empírica (histograma) de los estimadores $\hat{\sigma}^2$ y $\bar{\sigma}^2$ producidos via simulación con $N = 1000$ y 500 simulaciones; los parámetros reales son $A = 0$ y $\sigma^2 = 1$.

Ideas para madurar

- Un estimador es una variable aleatoria. Ello requiere que utilicemos mediciones estadísticas de desempeño.

Ideas para madurar

- Un estimador es una variable aleatoria. Ello requiere que utilicemos mediciones estadísticas de desempeño.
- Aunque el uso de simulación computacional resulta muy útil al momento de calcular el desempeño de un estimador, en muchos casos no es sufciente.

Ideas para madurar

- Un estimador es una variable aleatoria. Ello requiere que utilicemos mediciones estadísticas de desempeño.
- Aunque el uso de simulación computacional resulta muy útil al momento de calcular el desempeño de un estimador, en muchos casos no es sufciente.
- La complejidad computacional asociada al cálculo de un determinado estimador puede ser un factor en contra que uno debe tomar en cuenta al momento de seleccionar una determinada función.

Lecturas

- Steven M. Kay. *Fundamentals of Statistical Signal Processing: Volume 1 - Estimation Theory*, Capítulo 1.