

# Alternative Admissible Biased Estimators

F. Gama, D. Casaglia, and B. C. Frías, *Senior Member, IEEE*

**Abstract**— In this work, a biased estimator obtained from an affine transformation of an unbiased estimator will be analyzed. In particular, the method proposed by Y. C. Eldar will be studied, and alternative approaches will be developed for the construction of biased estimators through transformations of unbiased ones. It is finally shown through simulations, that both estimators proposed in this work perform better than other estimators analyzed.

**Keywords**— Biased estimation, dominant estimators, admissibility, mean square error.

## I. INTRODUCCION

LA UTILIZACIÓN de modelos paramétricos para describir situaciones físicas de interés es esencial en todas las ramas de la ingeniería; esto ha convertido a la estimación de parámetros en un campo central de estudio.

Siendo  $\theta \in \mathfrak{R}$  el parámetro desconocido y siendo  $X = \{X_1, \dots, X_n\}$  la muestra, un estimador del parámetro,  $h(X) \in \mathfrak{R}$ , es una función de las muestras que intenta parecerse al verdadero valor del parámetro,  $\theta$ , en algún sentido. La práctica más habitual es buscar estimadores que tengan mínimo error cuadrático medio. En otras palabras, el objetivo es buscar una función  $h(X)$  que minimice

$$\begin{aligned} ECM(h(X)) &= E[(h(X) - \theta)^2] \\ &= Var(h(X)) + B(h(X))^2 \end{aligned} \quad (1)$$

donde  $Var(h(X)) = E[(h(X) - E[h(X)])^2]$  es la varianza del estimador y  $B(h(X)) = E[h(X)] - \theta$  es el sesgo. Los estimadores se definen como no sesgados cuando  $E[h(X)] = \theta$ , llevando a  $ECM(h(X)) = Var(h(X))$ .

En general, los estimadores más utilizados son los estimadores no sesgados. En este sentido, H. Cramér [1] y C. R. Rao [2] desarrollaron una cota, la CRLB, que impone un límite inferior a la mínima varianza que los estimadores no sesgados pueden tener, constituyendo un punto de referencia para el desempeño de todos los estimadores no sesgados.

Sin embargo, hay muchos casos en que se pueden encontrar estimadores sesgados que posean un mejor desempeño que el estimador no sesgado de mínima varianza (MVUE), al menos dentro de alguna región de interés [4], [5],

[6], [7]. Esta observación abrió el camino hacia el desarrollo de numerosas técnicas capaces de obtener y controlar estimadores sesgados que resulten mejores que los correspondientes estimadores no sesgados. En particular, estas técnicas suelen resultar útiles en situaciones de baja relación señal-ruido (SNR) o cuando hay poca cantidad de muestras disponibles. Se pueden usar, también, cuando existe cierta información acerca del parámetro que puede incluirse en la estimación para mejorar el desempeño.

En esta área de investigación, Y. C. Eldar publicó una serie de trabajos donde propuso una técnica específica para mejorar cualquier estimador no sesgado mediante la aplicación de una transformación afin particular. La idea principal consiste en elegir el sesgo apropiado (cuya forma funcional es conocida por ser una transformación afin de un estimador no sesgado) de tal forma que la cota de Van Trees [12] para estimadores sesgados,

$$ECM(h(X)) \geq \frac{\left(1 + \frac{dB}{d\theta}\right)^2}{E\left[\left(\frac{\partial}{\partial\theta}(\ln(p(X;\theta)))\right)^2\right]} + B(\theta)^2 \quad (2)$$

resulte inferior que la CRLB. Más aún, mediante el uso de una transformación afin, resulta cierto que si el estimador no sesgado alcanza la CRLB, entonces el estimador transformado alcanzará la cota dada por (2).

El objetivo de este trabajo es analizar la transformación particular utilizada por Y. C. Eldar e introducir dos nuevos estimadores sesgados que presentan un menor error cuadrático medio en ciertas regiones de interés. En la sección II se presenta el criterio bajo el cual los estimadores van a ser evaluados, en la sección III se explica y analiza la técnica desarrollada por Y. C. Eldar y en la sección IV se introducen dos nuevos estimadores sesgados. Finalmente, una comparación entre las distintas soluciones al problema se puede encontrar en la sección V, algunos ejemplos en la sección VI y las conclusiones en la sección VII.

Para el resto de este trabajo, se harán las siguientes consideraciones: el parámetro es unidimensional,  $\theta \in \mathfrak{R}$ , una muestra  $X = \{X_1, \dots, X_n\}$  de tamaño  $n$  está disponible, y se sabe que el parámetro yace en la región  $R = \{\theta \in \mathfrak{R} : \theta_A < \theta < \theta_B\}$  donde  $\theta_A, \theta_B$  son completamente conocidos.

## II. ADMISIBILIDAD Y DOMINANCIA

En el primer trabajo de la serie analizada, Yonina C. Eldar [8] desarrolla de forma exhaustiva los conceptos de admisibilidad y dominancia. Estos conceptos se utilizarán para

F. Gama, Universidad de Buenos Aires, Ciudad de Buenos Aires, Argentina, fgama@fi.uba.ar

D. Casaglia, Universidad de Buenos Aires, Ciudad de Buenos Aires, Argentina, fgama@fi.uba.ar

B. C. Frías, Universidad de Buenos Aires, Ciudad de Buenos Aires, Argentina, fgama@fi.uba.ar

determinar cuándo un estimador es mejor que otro. La necesidad de definir y usar estos conceptos surge del hecho de que, en general, el error cuadrático medio del estimador va a depender del verdadero valor del parámetro, que es desconocido. Este hecho vuelve inútil al ECM para comparar entre estimadores.

Un estimador  $h_1(X)$  se dice que **domina** otro estimador  $h(X)$  en el conjunto  $R$ , si su error cuadrático medio nunca es mayor que el de  $h(X)$  para todos los valores de  $\theta$  en  $R$  y es menor para, al menos, un valor de  $\theta$  en  $R$ . Si  $ECM(h_1(X))$  es menor que  $ECM(h(X))$  para todos los valores de  $\theta$  en  $R$ , entonces  $h_1(X)$  **domina estrictamente**  $h(X)$  en la región  $R$ . Finalmente, un estimador que no es dominado por ningún otro estimador en el conjunto  $R$  se llama **admisibile**.

Es interesante notar que usar el concepto de admisibilidad como el criterio final para elegir una estrategia específica, da lugar a infinidad de estimadores admisibles. Esto es, obtener un estimador admisible es una tarea trivial, puesto que elegir un valor fijo de  $\theta \in R$  como estimador minimizará el ECM en un entorno de ese valor de  $\theta$ . Básicamente, la idea es que cualquier estimador que es mejor que un estimador admisible en alguna porción de la región  $R$ , también será admisible, dando lugar a una cantidad infinita de estimadores admisibles.

### III. EL ESTIMADOR AFÍN SEGÚN LA ESTRATEGIA MINIMAX.

Más tarde ese mismo año, Eldar [9] extiende la misma técnica utilizada para el modelo de regresión lineal, a un modelo paramétrico más general, construyendo un nuevo estimador sesgado mediante una transformación lineal a un estimador no sesgado. Luego, dos años después, la progresión natural se alcanza mediante la extensión de la idea a una transformación afín del estimador no sesgado, buscando mejorar aún más el desempeño del estimador sesgado obtenido en [10].

La idea esencial de estos dos trabajos ([9,10]) es la misma. Primero, como se mencionó en la introducción, se centra la atención en la cota (2) [12]: el sesgo que aparece en esta cota se torna conocido mediante la propuesta de un sesgo dentro de una determinada familia (en este caso, mediante el uso de transformaciones lineales o afines de un estimador no sesgado) determinando de forma unívoca la forma funcional de la ecuación (2). Segundo, se toma a la cota como una función de los parámetros de la transformación y se minimiza utilizando técnicas de optimización convexa, en particular, Semi Definite Programming (SDP).

La ventaja de esta estrategia es doble. Analíticamente, el problema es más sencillo (que otras formas del sesgo) y, en la práctica, como las transformaciones son afines, se vuelve trivial obtener estimadores que alcancen la cota (2) utilizando estimadores no sesgados que sean eficientes. Más aún, existen herramientas computacionales muy eficientes para resolver SDP.

#### A. El estimador lineal.

En el trabajo publicado en [9] un estimador de la forma

$$h_1(X) = (1+m)h(X) \quad (3)$$

es propuesto, donde  $h(X)$  es un estimador tal que  $E[h(X)] = \theta$  y  $m \in \mathfrak{R}$ . Estimadores de la forma (3) se llaman estimadores lineales. Para el estimador transformado,  $h_1(X)$ , se puede observar que

$$E[h_1(X)] = \theta + m\theta \Rightarrow B(\theta) = E[h_1(X)] - \theta = m\theta \quad (4)$$

$$\Rightarrow \frac{dB}{d\theta} = m$$

convirtiendo a la ecuación (2) en

$$ECM(h_1(X)) \geq \frac{(1+m)^2}{E\left[\left(\frac{\partial}{\partial\theta}(\ln(p(X;\theta)))\right)^2\right]} + m^2\theta^2 \quad (5)$$

Se puede mostrar que, si el estimador no sesgado  $h(X)$ , es, no solo de mínima varianza (MVUE), sino que además es eficiente Cramér-Rao, entonces el estimador transformado,  $h_1(X)$ , alcanza la cota dada por la ecuación (2). Esto es,

- Sea  $h(X)$  un estimador del parámetro  $\theta$  basado en la muestra  $X$ .
- $E[h(X)] = \theta$  y, más aún,  $h(X)$  es eficiente Cramér-Rao (alcanza la CRLB).
- $h_1(X) = (1+m)h(X)$ ,  $m \in \mathfrak{R}$ .

$$\Rightarrow Var(h_1(X)) = \frac{(1+m)^2}{E\left[\left(\frac{\partial}{\partial\theta}(\ln(p(X;\theta)))\right)^2\right]} \quad (6)$$

Ahora, la ventaja de usar una transformación lineal se puede ver claramente: permite trabajar tanto con la minimización de la cota como con toda la teoría de estimadores no sesgados de tal forma de obtener un estimador  $h_1(X)$  que tenga un error cuadrático medio menor que el de  $h(X)$  en la región de interés  $R$ .

El objetivo, entonces, está puesto en encontrar el valor de  $m$  que minimice (5). Derivando respecto de  $m$ , igualando a cero y resolviendo para  $m$  (que es el mínimo dado que la ecuación es cuadrática y positiva), de la ecuación (5) se obtiene

$$0 = \frac{d}{dm} \left( \frac{(1+m)^2}{E\left[\left(\frac{\partial}{\partial\theta}(\ln(p(X;\theta)))\right)^2\right]} + m^2\theta^2 \right) \Bigg|_{m=m^*} \quad (7)$$

$$\Rightarrow m^* = \frac{-1}{1 + \theta^2 E\left[\left(\frac{\partial}{\partial\theta}(\ln(p(X;\theta)))\right)^2\right]}$$

De donde se observa que, desafortunadamente, el valor

óptimo de  $m$  que minimiza el ECM depende, en general, de  $\theta$ .

La estrategia adoptada por Eldar [9] para resolver este problema consiste en buscar el valor de  $m$  que ocasione que el ECM sea menor que la CRLB para todos los valores de  $\theta \in R$ . Se aplica el criterio minimax, buscando maximizar la mínima diferencia entre la cota de Van Trees (5) y la CRLB.

En su trabajo [9], página 5, se puede encontrar el Teorema 2 que afirma,

- Sean  $X$  las muestras obtenidas de una variable aleatoria con densidad de probabilidades  $p(X; \theta)$ , donde  $\theta$  es un parámetro determinístico y desconocido que desea ser estimado.
- Sea  $h(X)$  un estimador no sesgado, y sea  $h_1(X) = (1+m)h(X)$ .
- Sea  $MSE_B(m, \theta)$  la cota dada por la ecuación (5).
- Sea  $R$  el conjunto de valores que el parámetro  $\theta$  puede tomar.
- Sea

$$m^* = \arg \min_m \sup_{\theta \in R} \{MSE_B(m, \theta) - MSE_B(0, \theta)\} \quad (8)$$

Entonces,

- $m^*$  es único.
- $h_1(X)$  es admisible en  $R$ .
- Si  $m^* \neq 0$  entonces  $MSE_B(m^*, \theta) < MSE_B(0, \theta)$  para todo  $\theta \in R$ . En otras palabras,  $h_1(X)$  es estrictamente dominante sobre la CRLB.

Se encuentra que la prueba de este Teorema es la misma que la del teorema 1 en [8].

Analizando el teorema, surgen varias observaciones que pueden hacerse.

En primer lugar, al afirmar que el valor óptimo,  $m^*$ , puede obtenerse como resultado de un problema minimax es lo que le permite a Eldar utilizar la teoría de optimización convexa para resolver (8).

En segundo lugar, al decir que  $h_1(X)$  es única, se refiere a que existe un único  $m^*$  que puede ser obtenido mediante la solución del problema (8).

En tercer lugar, la afirmación de admisibilidad de  $h_1(X)$  en  $R$  ilustra que el estimador es aceptable en términos de los criterios de decisión que se explicaron en la sección II. Sin embargo, es absolutamente necesario señalar que no es el único estimador admisible, ni siquiera dentro de la clase de estimadores lineal. En otras palabras, Eldar encuentra un método sistemático para construir estimadores lineales admisibles, pero no es la única estrategia que puede ser utilizada.

Finalmente, resulta conveniente observar que la ecuación

(8), en un caso útil, lleva a  $MSE_B(m, \theta) < MSE_B(0, \theta)$ , por lo que  $MSE_B(m, \theta) - MSE_B(0, \theta) < 0$  y, entonces, lo que la ecuación (8) dice explícitamente es que hay que buscar el valor de  $\theta$  que ocasiona que  $MSE_B(m, \theta)$  esté lo más cerca posible de  $MSE_B(0, \theta)$  (que es el supremo dado que el intervalo es negativo), y luego llevar ese valor lo más lejos posible de cero (minimizarlo), finalmente haciendo la diferencia lo más grande posible.

La estrategia minimax puede ser fácilmente observada en el caso particular en que  $Var(h(X)) = V = const.$  [11]. El valor óptimo  $m$  resulta ser

$$m^* = \frac{-1}{1 + \theta^2 / V} \quad (9)$$

que depende de  $\theta$  como era esperado.  $MSE(h_1(X))$  es

$$MSE(h_1(X)) = (1+m)^2 V + m^2 \theta^2 \quad (10)$$

Que es una ecuación cuadrática positiva en  $\theta$ , mientras que  $MSE(h(X))$  es constante en  $\theta$ . Todo esto se ilustra en la Fig. 1.

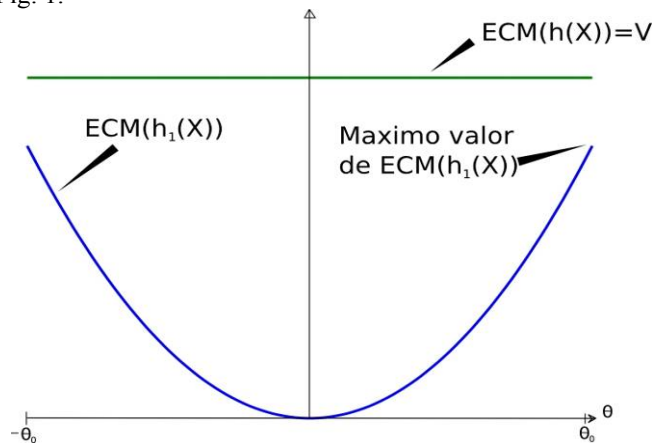


Figura 1. Situación minimax. Buscar el valor de  $\theta$  que hace mínima la diferencia.

Se puede observar que, aplicar la estrategia minimax significa encontrar el valor de  $\theta \in R$  para el cual la diferencia es mínima (supremo de  $ECM_B(m, \theta) - ECM_B(0, \theta)$ ) y luego buscar el valor de  $m$  que maximice esta diferencia (minimice  $ECM_B(m, \theta) - ECM_B(0, \theta)$  como función de  $m$ ) tal y como se observa en la Fig. 2. De esta forma, se obtiene el estimador sesgado que minimiza el máximo valor del ECM en el conjunto  $R$ ,

$$h_1(X) = \frac{\theta_0^2}{\theta_0^2 + V} h(X) \quad (11)$$

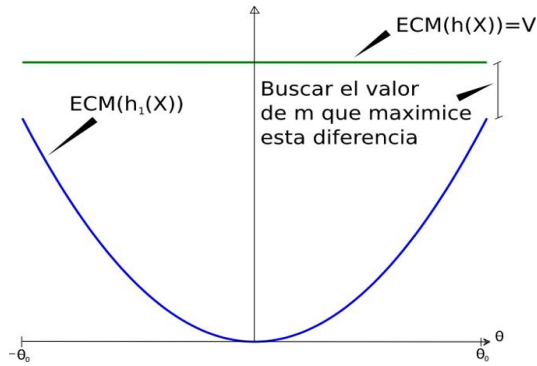


Figura 2. Situación minimax. Buscar el valor de  $m$  que maximice la mínima diferencia.

**B. El estimador afín.**

Considerando los resultados acerca de transformaciones lineales obtenidos en la sección previa, resulta natural extender la idea al caso de transformaciones afines; después de todo, si el estimador lineal es óptimo, entonces el *offset* de la transformación afín será cero, caso contrario, se esperarán mejoras en el desempeño.

Observando la obtención del estimador (11), se puede notar que, si el conjunto  $R$  no está centrado, entonces  $\theta_0 = \sup\{\theta \in R\}$ , ocasionando que el máximo valor del  $ECM(h_1(X))$  sea innecesariamente alto, degradando el desempeño del estimador. En otras palabras, si el parámetro pertenece a un intervalo no centrado  $R = \{\theta \in \mathfrak{R} : \theta_A < \theta < \theta_B\}$ , es intuitivo pensar que, si se puede obtener un estimador  $h_2(X)$ , que sea un desplazamiento del vértice del estimador  $h_1(X)$  al centro del intervalo  $R$  entonces el máximo valor del ECM se verá considerablemente reducido, como se muestra en la Fig. 3.

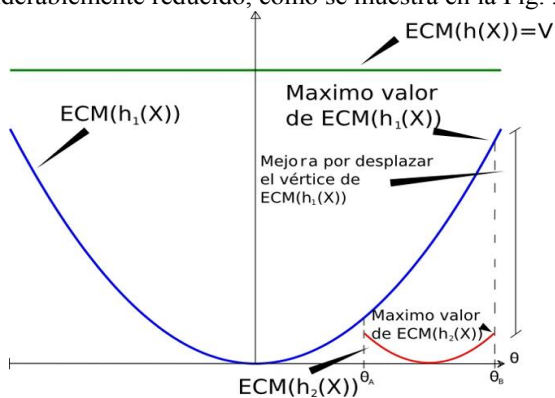


Figura 3. Desplazamiento del vértice del  $ECM(h_1(X))$  al centro del intervalo  $R$ .

Este desplazamiento se puede obtener sencillamente como un cambio de coordenadas,

$$\theta' = \theta - \frac{\theta_A + \theta_B}{2} \tag{12}$$

$$h_2(X) = (1 + m^*)h(X) - m^* \frac{\theta_A + \theta_B}{2} \tag{13}$$

$$1 + m^* = \frac{\left(\frac{\theta_B - \theta_A}{2}\right)^2}{\left(\frac{\theta_B - \theta_A}{2}\right)^2 + V}$$

La ecuación (13) muestra que la corrección realizada sobre el estimador lineal (11) para poder adaptarlo a cualquier segmento  $R$ , resulta ser una transformación afín del estimador no sesgado  $h(X)$ .

Eldar [10] publicó un trabajo acerca de transformaciones afines, extendiendo todas las ideas desarrolladas en la subsección anterior, al caso de una transformación afín sobre un estimador no sesgado.

Del mismo modo que en su trabajo anterior, primero se propone un estimador  $h_2(X)$  de la forma

$$h_2(X) = (1 + m)h(X) + u \tag{14}$$

donde  $h(X)$  es un estimador no sesgado. Luego, se obtiene que,

$$E[h_2(X)] = (1 + m)\theta + u \Rightarrow B(\theta) = m\theta + u \tag{15}$$

$$\Rightarrow \frac{dB}{d\theta} = m$$

Y la cota en (2) resulta ser

$$MSE(h_2(X)) \geq \frac{(1 + m)^2}{E\left[\left(\frac{\partial}{\partial \theta}(\ln(p(X; \theta)))\right)^2\right]} + (m\theta + u)^2 \tag{16}$$

Además, es inmediato que si  $h(X)$  es un estimador eficiente Cramér-Rao, entonces  $h_2(X)$  alcanza la cota dada por (16).

En el teorema 1 de su trabajo [10], Eldar desarrolla un resultado que es análogo al del caso lineal,

- Sean  $X$  las muestras obtenidas de una variable aleatoria con densidad de probabilidades  $p(X; \theta)$ , donde  $\theta$  es un parámetro determinístico y desconocido que desea ser estimado.
- Sea  $h(X)$  un estimador no sesgado, y sea  $h_2(X) = (1 + m)h(X) + u$ .
- Sea  $ECM_B(m, u, \theta)$  la cota dada por (16).
- Sea  $R$  el conjunto de valor al que pertenece  $\theta$ .
- Sea

$$(m^*, u^*) = \arg \min_{m, u} \sup_{\theta \in R} \{ECM_B(m, u, \theta) - ECM_B(0, 0, \theta)\}$$

(17)

Entonces,

- $m^*$  y  $u^*$  son únicas.
- $h_2(X)$  es admisible en  $R$ .

- Si  $m^* \neq 0$  o  $u^* \neq 0$  entonces  $MSE_B(m^*, u^*, \theta) < MSE_B(0, 0, \theta)$  para todo  $\theta \in R$ . Esto significa que  $h_2(X)$  es estrictamente dominante sobre la CRLB.

Básicamente, Eldar propone encontrar el par  $(m^*, u^*)$  que es solución de (17), asegurando que el estimador resultante  $h_2(X)$  construido a partir de  $h(X)$  (estimador no sesgado) usando (14), posee menor ECM para todos los valores de  $\theta \in R$  que el correspondiente a  $ECM(h(X))$ .

#### IV. DESARROLLO DE ESTIMADOR SESGADOS ALTERNATIVOS.

En esta sección se desarrollan dos estimadores admisibles. El primero es una transformación no lineal (truncamiento) de un estimador cualquiera. El segundo es una transformación afín de un estimador no sesgado pero, en este caso, construido utilizando un criterio diferente al criterio minimax adoptado por Eldar.

##### A. El estimador truncado.

Para este nuevo estimador, la idea es natural: sabiendo que el parámetro  $\theta$  no puede yacer fuera del segmento  $R$ , entonces es intuitivo pensar que, si cualquier estimador  $h(X)$  da un estimado afuera de esta región, entonces el seleccionar el valor más cercano que sí pertenece a  $R$  debería resultar en un estimador mejor.

Es en este espíritu que un lema general que prueba este resultado fue desarrollado,

- Sean  $X$  las muestras obtenidas de una variable aleatoria con densidad de probabilidades  $p(X; \theta)$ , donde  $\theta$  es un parámetro determinístico y desconocido que desea ser estimado.
- $\theta \in R$  donde  $R = \{\theta \in \mathfrak{R} : \theta_A < \theta < \theta_B\}$  y  $\theta_A$  y  $\theta_B$  son conocidos.
- Sea  $h(X)$  un estimador *cualquiera* de  $\theta$ .
- Sea  $\tilde{h}(X)$  un nuevo estimador tal que

$$\tilde{h}(X) = \begin{cases} \theta_A & \text{si } h(X) < \theta_A \\ h(X) & \text{si } \theta_A < h(X) < \theta_B \\ \theta_B & \text{si } \theta_B < h(X) \end{cases} \quad (18)$$

- $E\left[\left(\tilde{h}(X) - \theta\right)^2 \middle| h(X)\right]$  existe.

$$\text{Entonces,} \quad ECM(\tilde{h}(X)) \leq ECM(h(X)) \quad (19)$$

para cualquier estimador  $h(X)$ .

El ECM para el estimador truncado es

$$ECM(\tilde{h}(X)) = (\theta_A - \theta)^2 p_{h_A} + (\theta_B - \theta)^2 p_{h_B} + \int_{\theta_A}^{\theta_B} (\tilde{h}(X) - \theta)^2 p_{h(X)}(X; \theta) dX \quad (20)$$

donde  $p_{h_A} = \int_{-\infty}^{\theta_A} p_{h(X)}(X; \theta) dX$  es la probabilidad de que un estimado caiga debajo de  $\theta_A$  y, análogamente,

$$p_{h_B} = \int_{\theta_B}^{\infty} p_{h(X)}(X; \theta) dX.$$

Como puede observarse, (20) depende fuertemente del estimador  $h(X)$  que se elija. En el caso más general, la integral (20) no tendrá una expresión cerrada analítica. Sin embargo, las probabilidades de caer fuera del intervalo (outliers) se pueden estimar fácilmente. Más aún, es interesante notar que, incluso para el caso más general, las probabilidades  $p_{h_A}$  y  $p_{h_B}$  tenderán a aumentar cuando aumente el ruido (varianza del estimador  $h(X)$ ), acotando superiormente al máximo  $ECM(\tilde{h}(X))$  en, aproximadamente,  $(\theta_A - (\theta_A + \theta_B)/2)^2$ .

##### B. Estimador Afín-RB.

En esta sección se considerarán transformaciones afines a estimador no sesgados. La técnica presentada en este caso es bastante diferente de aquella utilizada en la sección III. Ideas de este tipo se pueden encontrar en [13], [14].

Si se tiene un estimador no sesgado  $h(X)$  y se aplica una transformación afín

$$h_1(X) = ah(X) + b \quad (21)$$

donde  $a, b \in \mathfrak{R}$  entonces el ECM será

$$ECM(h_1(X)) = (a-1)^2 \theta^2 + 2b(a-1)\theta + b^2 + a^2 Var(h(X)) \quad (22)$$

Si, además, se supone que  $Var(h(X)) = V$  es conocida y constante, entonces la ecuación (22) será una ecuación cuadrática positiva para todo  $\theta \in R$ .

Es deseable, para que el estimador  $h_1(X)$  sea mejor (en términos del MSE) que el estimador  $h(X)$ , que

$$(a-1)^2 \theta^2 + 2b(a-1)\theta + b^2 + (a-1)^2 V < 0 \quad (23)$$

Si la parábola (23) es igual a cualquier otra parábola con raíces  $\theta_A$  y  $\theta_B$ , el objetivo de  $ECM(h_1(X)) \leq ECM(h(X))$  para todo  $\theta \in R$  será alcanzado.

Se puede observar que el vértice de la función  $ECM(h_1(X)) = (a-1)^2 \theta^2 + 2b(a-1)\theta + b^2 + a^2 V$  se encuentra en el valor  $\theta_V = -b/(a-1)$  y el ECM en ese punto es  $ECM(h_1(X))_V = a^2 V$  de donde resulta que, si las

raíces de la ecuación (23) son reales, entonces debe darse que  $a^2V < V$ , y, por lo tanto,  $-1 < a < 1$ .

El valor del vértice,  $a^2V$ , puede tomar cualquier valor desde 0 hasta  $V$ . En el caso donde  $0 < a < 1$ , se puede observar que la parábola se abre y el vértice *sube* a medida que  $a$  va de cero a uno; si  $-1 < a < 0$ , entonces el vértice *sube* pero la parábola continúa cerrándose, cuando  $a$  toma valores desde 0 hasta  $-1$ . Por lo tanto, el caso límite es cuando  $a = 0$ ,  $ECM(h_1(X))_V = V_0 = (\theta_B - \theta_A)^2 / 4$ . Se puede observar que para valores de  $V$  menores que  $V_0$ , la parábola tiene que abrirse para lograr cortar la constante en los puntos  $\theta_A$  y  $\theta_B$ . Esto se logra incrementando el valor de  $a$  hacia 1 (observar que para  $a = 1$  el ECM del estimador es una constante). Por otro lado, si el valor de  $V$  es mayor que  $V_0$ , entonces la parábola puede únicamente cerrarse, a medida que  $a$  se reduce hacia  $-1$ , pero esto implica levantar el vértice de tal forma que ya no es posible cortar la constante  $ECM(h(X))=V$  en el lugar correcto. Por lo tanto, se debe satisfacer la condición  $V \leq (\theta_B - \theta_A)^2 / 4$ . Todo esto se observa en la Fig. 4.

Finalmente, se puede mostrar que los valores de  $a$  y  $b$  que hacen que la parábola (23) coincida con la que tiene raíces en  $\theta_A$  y  $\theta_B$ , asegurando  $ECM(h_1(X)) \leq ECM(h(X))$  para todo  $\theta \in R$ , son

$$a_{RB} = \frac{\gamma - 1}{\gamma + 1} \quad ; \quad b_{RB} = \frac{2}{\gamma + 1} \frac{\theta_A + \theta_B}{2} \tag{24}$$

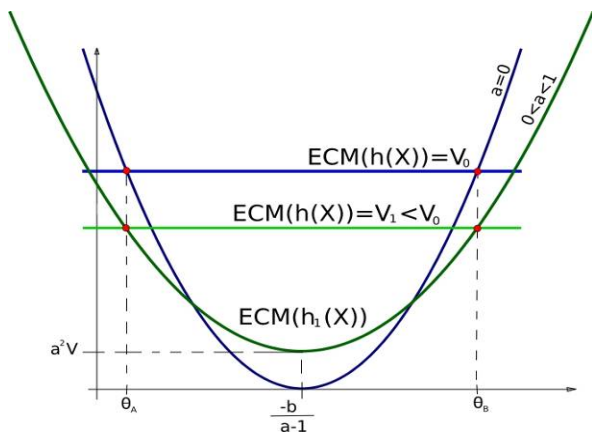


Figura 4. Parábolas del  $ECM(h_1(X))$  para diferentes valores de  $a$  según el valor del  $ECM(h(X))$ .

donde se ha definido el factor de mérito  $\gamma$  (que puede ser visto como una relación señal a ruido),

$$\gamma = \frac{\left(\frac{\theta_B - \theta_A}{2}\right)^2}{V} \tag{25}$$

### V. COMPARACIÓN.

En la Fig. 5 se presentan los ECMs para  $ECM(h_E(X))$  y  $ECM(h_{RB}(X))$ , siendo  $h_E(X)$  el estimador de Eldar, descrito por las ecuaciones (13), y  $h_{RB}(X)$  el estimador Afín-RB, desarrollado en las ecuaciones (21, 24, 25).

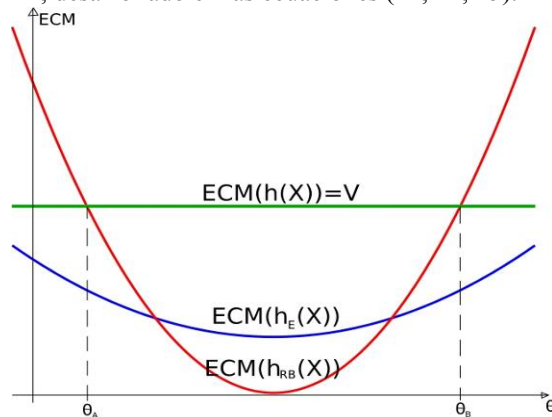


Figura 5.  $ECM(h(X))$ ,  $ECM(h_E(X))$ ,  $ECM(h_{RB}(X))$  como función de  $\theta$ .

El aspecto más destacable que se puede apreciar en la Fig. 5 es que el estimador de Eldar,  $h_E(X)$ , *no domina* al estimador Afín-RB,  $h_{RB}(X)$  (que también es una transformación afín de un estimador no sesgado pero construido de forma diferente). Más aún, si el estimador  $h_E(X)$  es admisible, como se mostró en los teoremas de la sección III, entonces el estimador  $h_{RB}(X)$  también será admisible, dado que no es dominado por un estimador admisible. Además, es interesante mencionar que el estimador  $h_{RB}(X)$  es estrictamente dominante sobre  $h(X)$  en  $R$  si el conjunto es abierto.

Finalmente, se puede encontrar una relación entre las longitudes de los intervalos donde cada estimador domina,

$$\xi = \frac{\text{longitud}\{RB\}}{\text{longitud}\{E\}} = \left(\sqrt{\frac{3}{2 - \gamma^{-1}} - 1}\right)^{-1} \tag{26}$$

que sería el parámetro esencial a tener en cuenta a la hora de decidir qué estimador utilizar (si  $\xi > 1$ , elegir  $h_{RB}(X)$ ).

### VI. EJEMPLOS.

Sea  $X = \{X_1, \dots, X_n\}$  una muestra de tamaño  $n$ , obtenida de variables aleatorias independientes, idénticamente distribuidas con distribución  $N(\theta, \sigma^2)$  donde  $\sigma^2 = 6,25$  es conocido y  $\theta$  es el parámetro a ser estimado. Más aún, se sabe que el parámetro  $\theta$  yace en la región  $R = \{\theta \in \mathfrak{R} : \theta_A < \theta < \theta_B\}$  con  $\theta_A = 5$  y  $\theta_B = 10$ . El verdadero valor del parámetro es  $\theta = 7$ . En la Fig. 6 se

presentan los ECMs para todos los estimadores desarrollados en este trabajo, como función del tamaño de la muestra

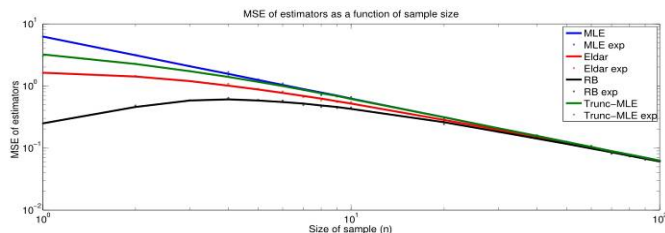


Figura 6. Comparación entre los ECMs de los diferentes estimador, para  $\theta = 7$ , y como función del tamaño de la muestra.

Se observa que  $ECM(h_{RB}(X)) \leq ECM(h_E(X))$ , para  $\theta = 7$ , y, por lo tanto, el estimador  $h_{RB}(X)$  resulta ser mejor que el estimador  $h_E(X)$ . Sin embargo, ambos estimadores son mejores que el estimador truncado,  $h_T(X)$ , que es, a su vez, mejor que el estimador de máxima verosimilitud (MLE, la media muestral).

En la Fig. 7 hay otro ejemplo, bajo las mismas condiciones que el anterior, pero, en este caso, se considera que  $\theta = 5,5$  es el verdadero valor del parámetro. Además, ahora, el estimador truncado se obtiene como una transformación del estimador de Eldar en lugar del MLE.

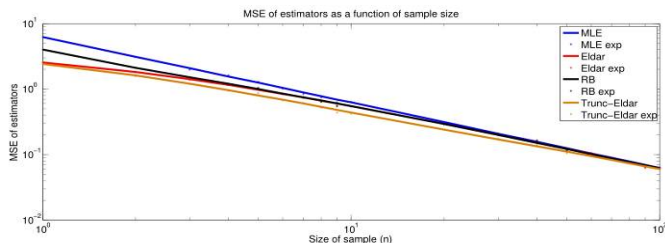


Figura 7. Comparación entre los ECMs de los diferentes estimadores para  $\theta = 5,5$ , como función del tamaño de la muestra.

Sucede ahora que, como para  $\theta = 5,5$ ,  $ECM(h_E(X)) \leq ECM(h_{RB}(X))$  el estimador de Eldar es mejor que el estimador Afin-RB. Sin embargo, es muy interesante observar que el estimador truncado es el mejor estimador de todos, como predecía la ecuación (19),  $ECM(h_T(X)) \leq ECM(h_E(X))$ .

Finalmente, para ambos problemas, el valor de  $\xi$  es 1,366, lo que implica que el estimador Afin-RB debiera ser elegido.

Estos estimadores son aplicables a cualquier problema de estimación donde el parámetro está sujeto a una región  $R$ . Aplicaciones típicas son estimación de la frecuencia de un tono puro [15] y filtro de Kalman [16].

## VII. CONCLUSIONES.

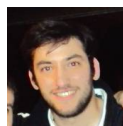
En el presente trabajo se desarrollaron y analizaron dos estimadores sesgados que explotan cierta información acerca

del parámetro ( $\theta \in R$ ) para lograr reducir el ECM. Más aún, el estimador  $h_{RB}(X)$ , ilustra una técnica distinta para construir estimadores sesgados como transformación afín de estimador no sesgados que es bastante distinta de la que utiliza Eldar. Se mostró que este estimador,  $h_{RB}(X)$ , también es admisible en el conjunto  $R$ .

La extensión a  $n$  dimensiones es bastante directa y es el tema de nuestro próximo trabajo.

## REFERENCIAS

- [1] H. Cramér, "Mathematical Methods of Statistics", First edition. Princeton Landmarks in Mathematics, Princeton University Press, United States, 1946.
- [2] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters", in Bulletin of Cal. Math. Soc, vol. 37, no. 3, pp. 81-91, 1945.
- [3] E. W. Barankin, "Locally best unbiased estimates", Ann. Math. Stat., vol. 20, 1946.
- [4] L. Le Cam, "On some asymptotic properties of maximum likelihood estimates and related Bayes estimates", University of California Publications in Statistics 1, pp. 277-330, 1953.
- [5] A. W. van der Vaart, "Asymptotic statistics", Cambridge University Press, 1998.
- [6] W. James and C. Stein, "Estimation with quadratic loss", Proc. Fourth Berkeley Symp. Math. Statist. Prob., vol. 1, pp. 361-379, University of California Press, 1961.
- [7] P. Stoica and R. Moses, "Introduction to spectral analysis", Prentice Hall, United States of America, 1997.
- [8] Y. C. Eldar, "Comparing Between Estimation Approaches: Admissible and Dominating Linear Estimators", IEEE Transactions on Signal Processing, Vol. 54, No. 5, pp. 1689-1702, May 2006.
- [9] Y. C. Eldar, "Uniformly Improving the Cramér-Rao Bound and Maximum Likelihood Estimation", IEEE Transactions on Signal Processing, Vol. 54, No. 8, pp. 2943-2956, August 2006.
- [10] Y. C. Eldar, "MSE Bounds With Affine Bias Dominating the Cramér-Rao Bound", IEEE Transactions on Signal Processing, Vol. 56, No. 8, pp. 3824-3836, August 2008.
- [11] S. Kay and Y. C. Eldar, "Rethinking Biased Estimation", Lecture Notes, IEEE Signal Processing Magazine, Vol. 25, No. 3, pp. 133-136, May 2008.
- [12] H. L. Van Trees, "Detection, Estimation and Linear Modulation Theory", Part I of Detection, Estimation and Modulation Theory, First Edition, John Wiley and Sons, United States of America, 1968.
- [13] G. A. Roitman and B. Cernuschi-Frías, "Data fusion best affine unbiased estimation of a deterministic vector, applications to image fusion". XXII Congreso Argentino de Control Automático, (AF-002313), Sep. 2010.
- [14] G. A. Roitman and B. Cernuschi-Frías, "Data fusion baue estimation of a deterministic vector, applications to image noise and blur reduction". 39 Jornadas Argentinas de Informática (JAIIO), Argentine Symposium on Technology, pages 1551-1562, Sep. 2010. ISSN:1850-2806.
- [15] F. Gama, D. Casaglia and B. Cernuschi-Frías, "Application of Affine Estimators to Single Tone Frequency Estimation", 41 Jornadas Argentinas de Informática (JAIIO), 13<sup>th</sup> Argentine Symposium on Technology, pages 121-131, Aug. 2012. ISSN: 1850-2806.
- [16] J. Tan, D. Li, J. Zhang, B. Hu and Q. Lu, "Biased Kalman Filter", 2011 Fifth International Conference on Sensing Technology, 2011, 581-584.



**Fernando Gama** is studying Electronic Engineering at the School of Engineering from University of Buenos Aires (FIUBA). He works as a teaching assistant in courses in Probability and Statistics and Stochastic Processes. He is currently working on his final research project in Estimation Theory at the Laboratory of Investigation in Signal and Image Processing and Neural Networks (LIPSIRN), from the Electronics Department, FIUBA. His areas of interest are Signal Processing in general, Telecommunications, Detection and Estimation Theory, Probability, among others.



**Daniel C. Casaglia**, received the Ingeniero Electromecánico Orientación Electrónica degree from the Universidad de Buenos Aires in 1981. In november of 1981 he joined the Instituto de Ingeniería Biomédica FIUBA as a young researcher, where he is now member of the Biological Signal Processing Group. He awarded three fellowships from CONICET for working in biomedical signal processing. From 1989 to 2003 he was member of the Scientific Research Career of the CONICET. He is member of the IEEE and is currently teaching Signal and Systems at Universidad de Buenos Aires.



**Bruno Cernuschi-Frias** (S'77–M'78–SM'01) was born in 1952, Montevideo, Uruguay. He received the Ph.D. degree in electrical engineering from Brown University, Providence, RI, in 1984. He is currently a full tenured Professor with the Department of Electronics, Faculty of Engineering, University of Buenos Aires, Argentina, and Principal Researcher at the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET).