Estimador Estocástico *m*-dimensional

R. Palma Orozco[1] y J. Medel Juárez[1,2]

[1]*Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada del Instituto Politécnico Nacional, Legaria 694. Colonia Irrigación, 11500 México D. F.*
[2]*Centro de Investigación en Computación del Instituto Politécnico Nacional, Unidad Profesional Adolfo López Mateos*

Resumen
Este trabajo muestra el desarrollo del estimador óptimo para un sistema tipo caja negra en un espacio de Hilbert *m*-dimensional, sobre observaciones con ruido y con un modelo de la incertidumbre de la dinámica del mismo.

Introducción
En estadística la regresión lineal o ajuste lineal es un método matemático que modela la relación entre una variable dependiente Y_k, las variables independientes Y_{k-1} y un término aleatorio W_k, es decir, $Y_k \propto W_k$. Este modelo puede ser expresado como (1).

$$Y_k = A_k Y_{k-1} + W_k \quad (1)$$

El modelo lineal relaciona las variables Y_k yW_k o cualquier transformación de éstas, que generan un hiperplano de parámetros A_k desconocidos. Donde W_k es la perturbación aleatoria que recoge todos aquellos factores de la realidad no controlables u observables y que por tanto se asocian con el azar, y es la que confiere al modelo su carácter estocástico [1-4]. En el caso más sencillo, con una sola variable explicativa, el hiperplano es una recta. Los valores seleccionados como estimadores de los parámetros, $\tilde{A}_k$, son los coeficientes de regresión, sin que se pueda garantizar que coinciden con parámetros reales del proceso generador. Por tanto, se tiene (2).

$$\tilde{Y}_k = \tilde{A}_k \tilde{Y}_{k-1} + \tilde{W}_k \quad (2)$$

Resultados y Análisis
Así, de la relación (1), se define el modelo del sistema tipo caja negra que satisface el Teorema 1.

Teorema 1. *Sea la salida acotada Y_k en R^{1xm} con dominio en $N(\mu,\sigma^2<\infty)$. Existe un estimador estocástico $\tilde{A}_k$ dado por (3).*

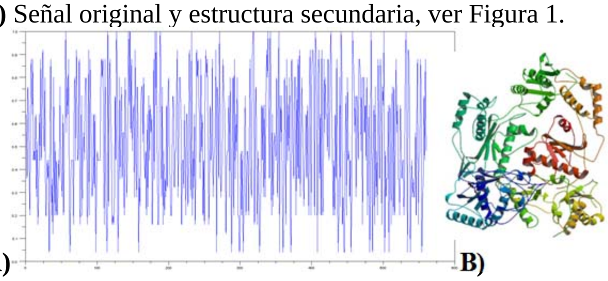
$$\tilde{A}_k = E\{Y_k M^+_{k-1}\} - E\{W_k M^+_{k-1}\} \quad (3)$$

Donde W_k en R^{1xm} y M^+_{k-1} en R^{mxm} son el ruido del proceso y la matriz de correlación, respectivamente.

Ejemplo. Identificación de la proteína 1RTH-A.

a) Secuencia de Aminoácidos:
PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPE
NPYNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKK
SVTVLDVGDAYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKG
SPAIFQSSMTKILEPFRKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQH
LLRWGLTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPEKDSWTVNDIQ
KLVGKLNWASQIYPGIKVRQLCKLLRGTKALTEVIPLTEEAELELAENREILK
EPVHGVYYDPSKDLIAEIQKQGQGQWTYQIYQEPFKNLKTGKYARMRGAH
TNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQKETWETWWTEYWQATWIP
EWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTNRG
RQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYALGIIQAQPDQSE
SELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKV

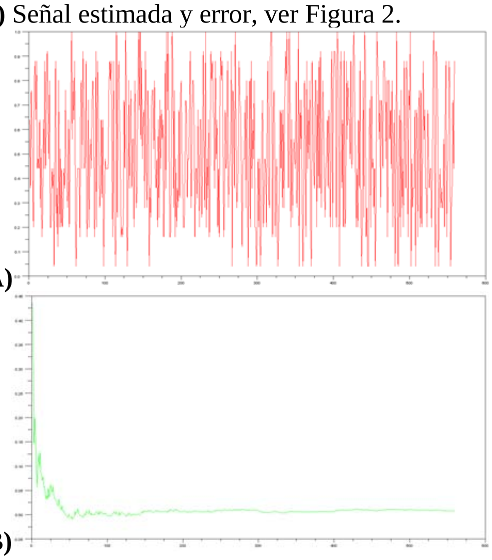
b) Señal original y estructura secundaria, ver Figura 1.



A) B)

Figura 1. A) Señal original codificada. **B)** Estructura secundaria de la proteína 1RTH-A.

c) Señal estimada y error, ver Figura 2.



A) B)

Figura 2. A) Señal estimada. **B)** Funcional de error.

Conclusiones
Un estimador para un sistema m-dimensional sin usar el cálculo de la matriz pseudoinversa fue transformado a una forma diagonal, obteniendo un sistema invertible válido y el estimador es definido mediante una matriz diagonal.

Agradecimientos
Agradecemos a CICATA-IPN por su apoyo.

Referencias
[1] J. Angrist, A. Krueger, *Instrumental variables and the search for identification: From supply and demand to natural experiments*, J. Econ. Perspect., 15(4) (**2001**) 69-85.
[2] J. Pearl, *Causality: Models, Reasoning and Inference*, (Cambridge University Press, **2000**).
[3] G. Imbens, J. Angrist, *Identification and estimation of local average treatment effects*, Econometrica 62 (**1994**) 467-476.
[4] A. Sinha, *Linear Systems: Optimal and Robust Control*, (CRC Press, **2007**).

 ISBN: 978-607-414-256-3