

Estimación diagonal y recursiva de parámetros para sistemas tipo caja negra con entradas y salidas acotadas

Rosaura Palma Orozco¹, José de Jesús Medel Juárez²

¹ Instituto Politécnico Nacional, Escuela Superior de Cómputo, Ciudad de México, México

² Instituto Politécnico Nacional, Centro de Investigación en Computación, Ciudad de México, México

rpalma@ipn.mx, jjmedelj@yahoo.com.mx

Resumen. La Teoría de la Estimación Estocástica se emplea para obtener información de la operación interna con respecto a la respuesta observable de un sistema tipo caja negra. Un problema por resolver es describir a los parámetros internos, a partir de un modelo de referencia. Se ha considerado que las dinámicas de los parámetros en un sistema estocástico está descrita por la relación de la varianza y covarianza de la señal observable. El método de los momentos de probabilidad permite obtener resultados que convergen a la respuesta deseada en un sentido de probabilidad. La estimación para sistemas MIMO (Multiple Input, Multiple Output) requiere del cálculo de la matriz pseudoinversa aunque se considere que es óptimo el modelo por el método del gradiente, al aplicar esa técnica se propone un vector propio y valores propios afines para la selección de los parámetros, haciendo que la estimación pierda gran parte de sus propiedades de convergencia. Este artículo presenta el desarrollo de un estimador estocástico óptimo para un modelo de sistemas tipo caja negra con ruido en un espacio m -dimensional. Se describe un algoritmo para evaluar y construir la forma diagonal del sistema en un espacio de estados con el propósito de estimar las ganancias internas. Los resultados presentan una solución sin pérdida de generalidad de las características del modelo de referencia. La técnica de estimación usada se basa en el gradiente estocástico junto con la variable instrumental para eficientar su nivel de convergencia. Este tipo de matriz de contribución es óptima en un sentido de probabilidad. El algoritmo permite eliminar el cálculo de matrices pseudoinversas que tiene una complejidad computacional de orden no lineal. La propuesta de la matriz diagonal sugiere una menor complejidad que los métodos utilizados tradicionalmente, ya que es de orden lineal, $O(j)$

donde $j \in N$, es la dimensión de la matriz. Los resultados muestran que es posible reconstruir la señal observable con una buena aproximación en un sentido de probabilidad, basado en la estimación por diagonales.

Palabras clave. Estimación recursiva, gradiente estocástico, sistema tipo caja negra.

Diagonal and Recursive Parameter Estimation for Black-box Systems with Bounded Inputs and Outputs

Abstract. Estimation theory is a branch of stochastic and signal processing that deals with estimating the parameter values based on an observable known signal as a random variable. The parameters describe an underlying physical setting in such a way that their value affects the distribution of the observable known signal. An estimator attempts to approximate the unknown parameters using the stochastic signal. In the estimation theory it is assumed that the output signal is random with the probability depending on the interest parameters. The estimation takes the measured observable signal as an input and produces an estimation of internal unknown gains. It is also preferable to derive an estimation that exhibits optimality, achieving minimum average error over some class, for example, an unbiased minimum variance as estimation. This paper presents the development of an optimal stochastic estimator for a black-box system in a m -dimensional space, observing noise with an unknown dynamics model. The results are described in a state space, with a discrete stochastic estimator and noise characterization. The results are obtained by an algorithm to construct the diagonal form for the state space system. Thus,

the matrix is estimated in probability considering the distribution function. The estimation technique is used on the instrumental variable based on a gradient stochastic matrix. This kind of matrix contribution is optimal in the probability sense. This is a new technique for an instrumental variable tool, and a diagonalization process avoiding the calculation of pseudo-inverse matrices is presented with a linear computational complexity $O(j)$ and j as the diagonal matrix dimension. The results show that it is possible to reconstruct the observable signal with a probability approximation. The advantages with respect to traditional solutions are focused on estimating the matrix contribution on line with a linear complexity.

Keywords. Recursive estimator, stochastic gradient, black-box system.

1. Introducción

La Teoría de Filtrado es el estudio de una variedad de problemas en la estimación de procesos estocásticos. Un ejemplo típico concierne a la estimación de una señal con ruido aditivo, en esta situación, son de interés las propiedades del proceso estocástico. El propósito del filtrado es estimar el estado de un sistema dinámico el cual es gobernado por perturbaciones aleatorias (ruidos), dando un proceso observado con ruido [1]. El procedimiento paso a paso para la construcción de un filtro es el siguiente. Primero, modelar los estados del sistema y su evolución dinámica más la dinámica de los ruidos y entradas desconocidas. Segundo, analizar el modelo si es lineal o no, si los ruidos son gaussianos o no y la observabilidad del sistema. Tercero, el diseño de un filtro adaptándose al modelo, donde la minimización del error cuadrático medio es el criterio general elegido. Así, en la Teoría de Filtrado Lineal el propósito es producir la mejor estimación que es un funcional lineal de las salidas del sistema [2].

El interés por los problemas de filtrado se debe a su papel central en varios temas de aplicación. Por ejemplo, en la Teoría de Comunicaciones, es común el modelo para una señal enviada en un canal con ruido; la transmisión exitosa de información requiere extraer la señal del ruido [3]. Problemas de Control Estocástico, en el cual un control debe ser elegido con el fin de influir en

el comportamiento de la señal, también puede implicar el filtrado, si el control permite depender de observaciones ruidosas o parciales de la señal.

La literatura moderna del filtrado comienza con la contribución de Kalman y Bucy, quienes formularon y resolvieron el modelo para el caso en el cual $x(t)$ es una distribución gaussiana y $w(t)$ es browniano. Su resultado principal prueba que la función de densidad condicional de $x(t)$ es gaussiana y provee un método para calcular la esperanza condicional y la covarianza recursivamente. Para pocos casos es una solución completa y fácilmente construible [4]. Sin embargo, dos caracterizaciones muy poderosas de filtros óptimos son conocidas y se cumplen en algunas situaciones generales. La primera es una fórmula tipo Bayes, la cual es debida a Kallianpur y Striebel, ésta es válida para ruido browniano con mínimas restricciones. Cuando la señal es markoviana, puede ser caracterizada como la solución de una ecuación diferencial estocástica (Fujisaki, Kallianpur y Kunita). En general, no puede ser encontrada de este resultado porque los coeficientes de la ecuación del filtro involucran estimaciones óptimas para diferentes funciones. Así, es que se requieren ecuaciones adicionales las cuales se necesitan estimar de otras funciones. El sistema resultante de ecuaciones es en general infinito-dimensional [5]. La formulación más limpia de esta dimensionalidad infinita es la ecuación diferencial parcial estocástica de Zakai, para una versión no normalizada de la función de densidad condicional, asumiendo que esta función de densidad existe.

Finalmente, muchos recientes desarrollos prometen nuevas ideas. V. Benes ha desarrollado ejemplos de problemas de filtrado explícitamente resueltos y Brockett, Clark y Mitter han aplicado técnicas y operadores de Lie al estudio de ecuaciones de densidad condicional [6]. Lo anterior, es una breve descripción de los resultados que constituyen las vetas principales de la Teoría de Filtrado, pero a pesar de su profundidad matemática, ellas permanecen incompletamente desarrolladas [6, 7].

El artículo está organizado de la siguiente forma. La Sección 1, con el propósito de poner en contexto el trabajo realizado, introduce los trabajos

de la Teoría de Filtrado. La Sección 2, se dedica íntegramente a la deducción teórica del modelo del estimador estocástico m -dimensional descrito en el Resultado 2.1. Los resultados obtenidos se presentan al final de la sección resaltando el proceso de transformación diagonal del espacio de estados para el filtrado. La Sección 3 presenta una aplicación del estimador para dos señales, verificando su optimalidad. Por último, se concluye con sugerencias para trabajo a futuro.

2. Principales resultados

Dentro del desarrollo de los estimadores es común usar el gradiente para la obtención del modelo de construcción de parámetros; pero al hacer uso de la pseudoinversa la optimización se transforma en una región, la cual está acotada por el grado de exactitud del algoritmo de inversión.

En el caso SISO (Single Input Single Output) esta situación no es observada. Para los casos MISO (Multiple Input Single Output), SIMO (Single Input Multiple Output) y MIMO (Multiple Input Multiple Output), el problema de encontrar una buena convergencia se limita al proceso de Penrose. La problemática principal se encuentra en la elección del vector propio que permita la inversión del vector seleccionado. Se observa que una inversión por pseudoinversa no permite una buena convergencia [8].

Para que la estimación no pierda sus propiedades óptimas se propone una transformación que permite el desarrollo de un algoritmo sin pérdida de generalidad. Los principales resultados de la investigación son presentados a continuación.

2.1. Estimador estocástico

El espacio de estados es un espacio de probabilidades, con espacio de medida Euclideo. El operador E , esperanza matemática, es un operador lineal en el espacio de estados.

Los vectores $Y_k, W_k \in \mathbb{R}^{1 \times m}$ son variables aleatorias linealmente independientes (v.a.l.i.) $N(\mu, \sigma^2 < \infty)$ y se corresponden con las matrices diagonales $\mathcal{Y}_k, \mathcal{W}_k \in \mathbb{R}^{m \times m}$ con todas sus entradas distintas de cero.

Toda matriz diagonal es simétrica, triangular y normal si sus entradas provienen de un cuerpo \mathbb{R} ó \mathbb{C} . Sea $D = \text{diag}(n_i) \in \mathbb{R}^{1 \times m}$, con $i = 1, 2, \dots, m$. El $\det(D) = \prod_i n_i$, es invertible si y sólo si cada $n_i \neq 0$, por lo tanto existen $\mathcal{Y}_k^{-1}, \mathcal{W}_k^{-1} \in \mathbb{R}^{1 \times m}$.

Sean Y_k y W_k , variedades diferenciables afines en el espacio de estados. Por lo tanto, $\{Y_k\} \propto \{W_k\}$ si y sólo si existe una relación lineal afín $Y_k \in \mathbb{R}^{m \times m}$ tal que $\mathcal{A}, \mathcal{I} \in \mathbb{R}^{m \times m}$ constantes, satisfacen la combinación lineal $Y_k = \mathcal{A}Y_{k-1} + \mathcal{I}W_k$.

Si $\mathcal{I} = I$, entonces

$$Y_k = \mathcal{A}Y_{k-1} + W_k. \quad (1)$$

Así, de la relación (1), se define el modelo del sistema tipo caja negra que satisface el Resultado 2.1.

Resultado 2.1 *Sea la salida acotada $\mathcal{Y}_k \in \mathbb{R}^{1 \times m}$ con dominio en $N(\mu, \sigma^2 < \infty)$. Existe un estimador estocástico $\tilde{\mathcal{A}}_k$ dado por (2)*

$$\tilde{\mathcal{A}}_k = E\{\mathcal{Y}_k \mathcal{M}_{k-1}^+\} - E\{\mathcal{W}_k \mathcal{M}_{k-1}^+\}, \quad (2)$$

donde $\mathcal{W}_k \in \mathbb{R}^{1 \times m}$ y $\mathcal{M}_{k-1}^+ \in \mathbb{R}^{m \times m}$ son el ruido del proceso y la matriz de correlación.

Se verifica por inducción sobre la dimensión del espacio de estados.

$$\begin{aligned} \text{diag}(Y_k) &= \text{diag}(A_{1k})\text{diag}(Y_{1k-1}) + \\ &\quad \text{diag}(A_{2k})\text{diag}(Y_{2k-1}) + \text{diag}(W_k), \\ \text{diag}(Y_k) - \text{diag}(W_k) &= \text{diag}(A_{1k})\text{diag}(Y_{1k-1}) + \\ &\quad \text{diag}(A_{2k})\text{diag}(Y_{2k-1})\text{diag}(Y_k - W_k) \\ &= \text{diag}(A_{1k}Y_{1k-1}) + \text{diag}(A_{2k}Y_{2k-1}) \\ &= \text{diag}(A_{1k}Y_{1k-1} + A_{2k}Y_{2k-1}) \\ &= \text{diag}(A_{nk})\text{diag}(Y_{nk-1})\text{diag}(Y_k - W_k)\text{diag}(Y_{nk-1})^T \\ &= \text{diag}(A_{nk})\text{diag}(Y_{nk-1})\text{diag}((Y_{nk-1})^T) \\ &= \text{diag}(A_{nk})\text{diag}(Y_{nk-1}(Y_{nk-1})^T) \\ &= \text{diag}(A_{nk})\text{diag}(M_{k-1}). \\ \therefore \text{diag}(A_{nk}) &= \\ \text{diag}(Y_k - W_k)\text{diag}(Y_{nk-1})^T \text{diag}(M_{k-1}^{-1}) \\ &= \text{diag}(Y_k - W_k)\text{diag}(Y_{nk-1}^T M_{k-1}^{-1}) \end{aligned}$$

$$\begin{aligned}
&= \text{diag}(Y_k - W_k) \text{diag}(M_{k-1}^+) \\
&= [\text{diag}(Y_k) - \text{diag}(W_k)] \text{diag}(M_{k-1}^+) \\
\implies \mathcal{A}_k &= [\mathcal{Y}_k - \mathcal{W}_k] \mathcal{M}_{k-1}^+.
\end{aligned}$$

Para el caso del modelo de sistema (1) se tiene una representación estimada dada por (2). El estimador \tilde{A}_k queda definido por (3).

$$\begin{aligned}
\tilde{A}_k &= E\{[\mathcal{Y}_k - \mathcal{W}_k] \mathcal{M}_{k-1}^+\} \\
&= E\{\mathcal{Y}_k \mathcal{M}_{k-1}^+\} - E\{\mathcal{W}_k \mathcal{M}_{k-1}^+\}. \quad (3)
\end{aligned}$$

2.2. Error y funcional de error

En principio, el espacio de estados es una variedad diferenciable y el objetivo es hacer que el error de identificación sea cero. Cada una de las formas de minimizar ese error es un método de implementar los filtros adaptativos. Aquí, se propone minimizar el funcional del error $\mathcal{J}_k = E\{\mathcal{E}_k \mathcal{E}_k^T\}$ aplicando el gradiente estocástico que proporciona la dirección de máximo descenso en la superficie de error. Esto se establece en los Resultados 2.2 y 2.3.

Resultado 2.2 Sea \mathcal{E}_k el error de identificación, definido en forma diagonal como $\mathcal{E}_k := \mathcal{Y}_k - \tilde{\mathcal{Y}}_k$. El funcional del error \mathcal{J}_k , está dado de manera recursiva por (4)

$$\mathcal{J}_k = \frac{1}{k^2} [(k-1)^2 \mathcal{J}_{k-1} + \mathcal{E}_k^2]. \quad (4)$$

Para verificarlo, supóngase que si $\mathcal{E}_k = \mathcal{Y}_k - \tilde{\mathcal{Y}}_k$, entonces se cumple (5) y si es estacionario se tiene (6).

$$E\{\mathcal{E}_k^2\} = \frac{1}{k^2} \sum_{i=1}^k \mathcal{E}_i^2 \quad (5)$$

$$= \frac{1}{k^2} \left[\mathcal{E}_k + \frac{(k-1)^2}{(k-1)^2} \sum_{i=1}^{k-1} \mathcal{E}_i^2 \right]. \quad (6)$$

Resultado 2.3 Sea $\nabla_{\tilde{A}_k}$ el operador gradiente estocástico respecto a la matriz de parámetros estimada \tilde{A}_k . Entonces $\nabla_{\tilde{A}_k} \mathcal{J}_k = 0$, minimiza el funcional del error y existe una vecindad de convergencia óptima dada por (7) en casi todos los puntos (c.t.p.).

$$\tilde{A}_k \xrightarrow[\text{c.t.p.}]{} \mathcal{A}_k \pm \epsilon. \quad (7)$$

Para verificarlo, se evalúa el gradiente estocástico con respecto a la matriz de parámetros estimada \tilde{A}_k .

$$\begin{aligned}
0 &= \nabla_{\tilde{A}_k} \mathcal{J}_k, \\
0 &= 2\tilde{A}_k \tilde{\mathcal{Y}}_{k-1}^2 - 2\tilde{\mathcal{Y}}_{k-1} \mathcal{A}_k \mathcal{Y}_{k-1} \\
&\quad - 2\tilde{\mathcal{Y}}_{k-1} \tilde{\mathcal{W}}_k^2, \\
0 &= \tilde{A}_k \tilde{\mathcal{Y}}_{k-1}^2 - \tilde{\mathcal{Y}}_{k-1} \mathcal{A}_k \mathcal{Y}_{k-1} - \tilde{\mathcal{Y}}_{k-1} \tilde{\mathcal{W}}_k^2, \\
\tilde{A}_k \tilde{\mathcal{Y}}_{k-1}^2 &= \tilde{\mathcal{Y}}_{k-1} \mathcal{A}_k \mathcal{Y}_{k-1} + \tilde{\mathcal{Y}}_{k-1} \tilde{\mathcal{W}}_k^2, \\
\tilde{A}_k &= \tilde{\mathcal{Y}}_{k-1} \mathcal{A}_k \mathcal{Y}_{k-1} (\tilde{\mathcal{Y}}_{k-1}^2)^{-1} \\
&\quad + \tilde{\mathcal{Y}}_{k-1} \tilde{\mathcal{W}}_k^2 (\tilde{\mathcal{Y}}_{k-1}^2)^{-1}, \\
\tilde{A}_k &\sim \mathcal{A}_k \mathcal{Y}_{k-1}^2 (\tilde{\mathcal{Y}}_{k-1}^2)^{-1} + \tilde{\mathcal{W}}_k^2 (\tilde{\mathcal{Y}}_{k-1}^2)^{-1} \\
&\sim \mathcal{A}_k + \tilde{\mathcal{W}}_k^2 \tilde{\mathcal{Y}}_{k-1}^{-1}. \\
\therefore \tilde{A}_k &\xrightarrow[\text{c.t.p.}]{} \mathcal{A}_k \pm \epsilon; \quad \epsilon = \tilde{\mathcal{W}}_k^2 \tilde{\mathcal{Y}}_{k-1}^{-1}.
\end{aligned}$$

El principal resultado obtenido en la tesis resuelve este problema, mediante la evaluación de la matriz inversa \mathcal{M}_{k-1}^+ , con una descomposición de matrices diagonales en forma de Jordan, que sin necesidad del cálculo de los valores propios es posible realizar de forma directa la inversión de la matriz. Y computacionalmente tiene muchas ventajas, ya que su complejidad es de orden lineal, dependiente de la dimensión del espacio de estados, como se muestra en la figura 1.

Finalmente, el resultado de la minimización del funcional del error garantiza la existencia de una vecindad de convergencia óptima, la cual está completamente determinada por $\pm \tilde{\mathcal{W}}_k^2 \tilde{\mathcal{Y}}_{k-1}^{-1}$.

3. Ejemplo de aplicación

Esta sección presenta un ejemplo con $m = 2$. Donde, las señales observables, y_{i1k} , son números reales aleatorios con una función de distribución normal $N(0, 1)$ y el ruido, w_{i1k} , está acotado por

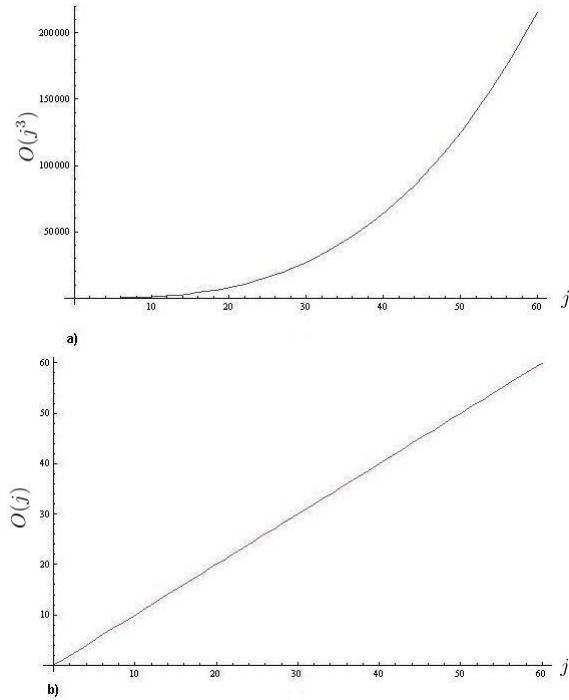


Fig. 1. Comparación de la complejidad computacional: **a)** Método de la pseudoinversa, **b)** Método de descomposición en diagonales

$[0, 0,01]$ con $i \in [1, m]$. El sistema para $m = 2$ está dado por (8). La figura 2, muestra una simulación para $m = 2$. La señal observable está en azul y la señal estimada en rojo. La señal estimada tiene una buena aproximación en probabilidad con $\{ |a_{ij}| \} \leq 1$ and $\sigma^2\{w_{ij}\} < \infty$.

El modelo para el caso $m = 2$ es escrito como (8), donde $Y_k, Y_{k-1}, W_k \in \mathbb{R}^{[2 \times 1]}$ y $A_k \in \mathbb{R}^{[2 \times 2]}$:

$$\begin{bmatrix} y_{11k} \\ y_{21k} \end{bmatrix} = \begin{bmatrix} a_{11k} & a_{12k} \\ a_{21k} & a_{22k} \end{bmatrix} \begin{bmatrix} y_{11k-1} \\ y_{21k-1} \end{bmatrix} + \begin{bmatrix} w_{11k} \\ w_{21k} \end{bmatrix}. \quad (8)$$

La forma diagonal para (8) está dada por (10), donde $d_1 = y_{11k} - w_{11k}$, $d_2 = y_{21k} - w_{21k}$, $d_3 = y_{21k} - w_{21k}$, $d_4 = y_{11k} - w_{11k}$.

Para el caso $m = 2$, la figura 2 muestra separadamente cada una de las componentes del sistema. La figura 2a) muestra en azul la

primera señal observable, y_{11k} , y en rojo la primera identificación \hat{y}_{11k} .

$$\begin{bmatrix} d_1 & 0 & | & 0 & 0 \\ 0 & d_2 & | & 0 & 0 \\ \hline 0 & 0 & | & d_3 & 0 \\ 0 & 0 & | & 0 & d_4 \end{bmatrix} = \begin{bmatrix} a_{11k} & 0 & | & 0 & 0 \\ 0 & a_{22k} & | & 0 & 0 \\ \hline 0 & 0 & | & a_{12k} & 0 \\ 0 & 0 & | & 0 & a_{21k} \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} y_{11k-1}^2 & 0 & | & 0 & 0 \\ 0 & y_{21k-1}^2 & | & 0 & 0 \\ \hline 0 & 0 & | & y_{21k-1}^2 & 0 \\ 0 & 0 & | & 0 & y_{11k-1}^2 \end{bmatrix}$$

La figura 2b) muestra en azul la segunda señal observable, y_{21k} , y en rojo la segunda identificación, \hat{y}_{21k} . Como se muestra, el estimador es muy preciso, porque mantiene la razón de cambio de cada señal. La figura 2c) muestra los valores estimados para \hat{A}_k . Finalmente, la figura 2d) muestra la superficie correspondiente a cada uno de los datos $\hat{a}_{11k}, \hat{a}_{12k}, \hat{a}_{21k}, \hat{a}_{22k}$.

4. Conclusiones y trabajo a futuro

Se consideró un sistema de entradas y salidas acotadas BIBO (Bounded Input Bounded Output) tipo MIMO estacionario de forma rígida; es decir, que toda entrada acotada produce una salida acotada, además la función de distribución del proceso permanece invariante con respecto al tiempo y así sus dos primeros momentos de probabilidad. Dada la respuesta de un sistema tipo caja negra, se propuso un modelo autorregresivo de media móvil de orden uno con una matriz de parámetros desconocidos, con base a esto, se presentó de manera recursiva el estimador matricial de parámetros.

El filtro de estimación se contruyó con base al segundo momento de probabilidad aplicado al modelo de diagonales extendido para describir la matriz de parámetros con la variable instrumental

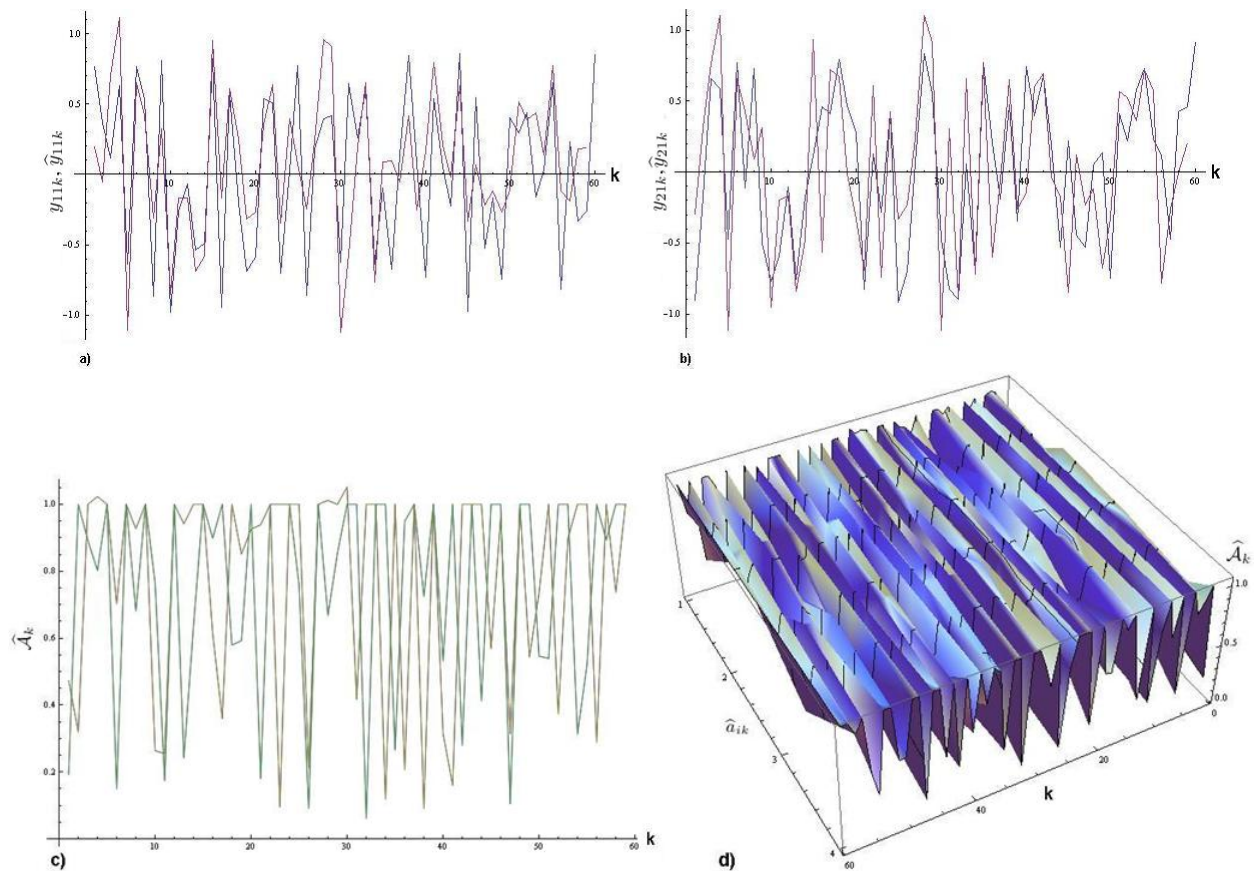


Fig. 2. Caso $m=2$. La señal observable y la identificación de sus componentes. **a)** Primera componente, y_{11k} , \hat{y}_{11k} , **b)** Segunda componente, y_{21k} , \hat{y}_{21k} , **c)** Valores estimados para \hat{A}_k , **d)** Gráfica 3-dimensional de una superficie representando \hat{A}_k

en diagonales. Este tipo de matriz de contribución es óptima en un sentido de probabilidad. Los resultados presentaron una solución sin pérdida de generalidad de las características del modelo de referencia. La técnica de estimación que se usó se basó en el gradiente estocástico junto con la variable instrumental para eficientar su nivel de convergencia.

Se estableció que el estimador es óptimo porque el gradiente estocástico del funcional del error de la señal observable correspondió a la relación descrita por el modelo de estimación mediante el segundo momento de probabilidad. El algoritmo permitió eliminar el cálculo de matrices pseudo-inversas que tiene una complejidad computacional de orden no lineal. La propuesta de la matriz

diagonal sugirió una menor complejidad que los métodos utilizados tradicionalmente, siendo de orden lineal, $O(j)$ donde $j \in N$, es la dimensión de la matriz. Los resultados mostraron que es posible reconstruir la señal observable con una buena aproximación en un sentido de probabilidad, basado en la estimación por diagonales.

El método de diagonalización propuesto ha demostrado ser efectivo en resolver problemas de identificación y estimación de parámetros internos en sistemas tipo caja negra. De forma inmediata, se puede utilizar en la estimación de sistemas de redes neuronales artificiales para determinar los pesos de la red de forma dinámica. También, posteriormente será muy interesante aplicarlo como predictor de la evolución de sistemas

dinámicos mediante la implementación de un filtrado adaptivo.

Agradecimientos

Este trabajo fue financiado por los proyectos SIP-IPN con registros 20150462 y 20150184.

Referencias

1. **Kalman, R.E. (1960).** A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME—Journal of Basic Engineering*, Vol. 82, Series D, pp. 35–45.
2. **Haykin, S. (2001).** *Kalman filtering and neural networks*. John Wiley and Sons Inc.
3. **Nelson, C.R. & Startz, R. (1990).** Some further results on the small sample properties of the instrumental variable estimator. *Econometrica*, Vol. 58(4), pp. 967–976.
4. **Angrist, J. & Krueger, A. (2001).** Instrumental variables and the search for identification: From supply and demand to natural experiments. *J. Econ. Perspect.*, Vol. 15(4), pp. 69–85.
5. **Aitken, A.C. (1935).** On Least Squares and Linear Combinations of Observations. *Proc. R. Soc. Edinburgh*, Vol. 55, pp. 42–48.
6. **Diniz, P.S.R. (2010).** *Adaptive Filtering: Algorithms and Practical Implementation*. Springer-Verlag.
7. **Liu, W., Principe, J.C., & Haykin, S. (2010).** *Kernel Adaptive Filtering: A Comprehensive Introduction*. John Wiley and Sons Inc.
8. **Palma, R., (2012).** *Estimación recursiva de parámetros para sistemas tipo caja negra con entradas y salidas acotadas*. Tesis doctoral, CICATA-Legaria, IPN.

Rosaura Palma Orozco es profesora de la Escuela Superior de Cómputo del IPN, egresada del Doctorado en Tecnología Avanzada por el CICATA Legaria. Es miembro del Sistema Nacional de Investigadores. Cuenta con varias publicaciones en el área de Filtrado lineal estocástico y Filtrado recursivo, específicamente en estimación dinámica. Su interés se centra en las aplicaciones computacionales y matemáticas de la Teoría de Filtrado para análisis y estudio de sistemas complejos representados en espacio de estados.

José de Jesús Medel Juárez es profesor de tiempo completo del Centro de Investigación en Computación del IPN, egresado del Doctorado en Ciencias en Control Automático del CINVESTAV. Es miembro del Sistema Nacional de Investigadores, así como de la Academia Mexicana de Ciencias y Who's who in Sciences and Engineering. Ha desarrollado y sometido diversas patentes, y graduado tanto en doctorado como en maestría a más de 35 profesionistas. Cuenta con más de 8 libros publicados por editoriales nacionales como internacionales. Con sus diversas colaboraciones y en la etapa de formación de sus estudiantes ha logrado artículos indizados en JCR, Scopus, CONACyT, entre otros. Su interés se centra en el estudio de los sistemas de control usando Lógica Difusa.

*Artículo recibido 13/05/2015; aceptado 26/11/2015.
Author de correspondencia es Rosaura Palma Orozco.*