



# Identificación de Sistemas Dinámicos con un Enfoque Integrador Basado en Esparcimiento, Regularización y Aproximación de Bajo Rango.

System identification based on an integrated approach of sparsity, regularization and low rank approximation.

K. CARDONA<sup>1</sup>, F. VIDES<sup>2</sup>

Recibido: 11 de noviembre de 2023 / Aceptado: 2 de agosto de 2024

## RESUMEN

En este documento se presentan aplicaciones de algunas técnicas teóricas y computacionales para la aproximación estructurada de sistemas dinámicos basados en datos. La investigación realizada en este artículo está enfocada en modelos lineales con aplicaciones en ingeniería y ciencias. Específicamente, se integran las propiedades de regularización y esparcimiento en la aproximación de los parámetros con un enfoque en la aproximación de bajo rango. Los resultados son independientes de una representación en particular del sistema y además no se asume una partición de los datos como entrada-salida. Las técnicas antes mencionadas se comparan por medio de algunas simulaciones numéricas y la aplicación en datos reales.

## ABSTRACT

In this document applications of some theoretical and computational techniques are presented for the structured approximation of dynamic systems based on data. The research carried out in this article is focused on linear models with applications in engineering and science. Specifically, the regularization and sparsity properties are integrated into the parameter approximation with a focus on the low-rank approximation. The results are independent of a particular representation of the system and furthermore a partition of the data is not assumed as input-output. The aforementioned techniques are compared through some numerical simulations and the application in real data.

## PALABRAS CLAVES

Identificación de sistemas, esparcimiento y regularización, aproximación de bajo rango.

## KEYWORDS

System identification, sparsity and regularization, low rank approximation.

<sup>1</sup>Centro de Innovación y Computo Científico, Escuela de Matemáticas y Ciencias de la Computación. Universidad Nacional Autónoma de Honduras. [kerin.cardona@unah.edu.hn](mailto:kerin.cardona@unah.edu.hn)

<sup>2</sup>Centro de Innovación y Computo Científico, Escuela de Matemáticas y Ciencias de la Computación. Universidad Nacional Autónoma de Honduras. [fredy.vides@unah.edu.hn](mailto:fredy.vides@unah.edu.hn)

\* Esta obra está bajo una licencia Creative Commons Reconocimiento - NoComercial 4.0 Internacional

\* This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

## I | INTRODUCCIÓN

La construcción de modelos basados en datos es una pieza fundamental en aplicaciones en ciencias e ingeniería, la identificación de sistemas consiste en un conjunto de técnicas desarrolladas para construir modelos matemáticos de sistemas dinámicos a partir de datos observados.

Los modelos esparcidos o dispersos representan una de las mejores alternativas para tratar problemas de alta dimensión Vides (2021a, 2021b), la idea principal es asumir que el vector de parámetros contiene muchas componentes exactamente cero o lo suficientemente pequeñas. Dicha suposición es de gran importancia para lograr la identificación del verdadero modelo, especialmente si solo se cuenta con una muestra de datos pequeña.

Debido al gran avance en las ciencias computacionales, los datos de alta dimensión han surgido en muchos campos de la ciencia, ingeniería, humanidades y la economía. Diversos temas como son comercio electrónico, mercadeo, genética, imágenes biomédicas y muchos más, proporcionan datos a gran escala que requieren de nuevas técnicas teóricas y computacionales para un tratamiento más efectivo.

Por ejemplo, los sistemas dinámicos aplicados a los mercados financieros actualmente reciben mucha atención a medida que la información de los mercados se vuelve disponible Vides y Nogueira (2023). Construir modelos de los mercados financieros es un problema considerado complejo, dado que no se cuenta con principios fundamentales, como en el caso de la física e ingeniería.

El propósito de este documento es aplicar algunas técnicas teóricas y computacionales presentadas en Markovsky (2019) y Vides (2019, 2021a, 2021b), para la aproximación de sistemas basados en datos de alta dimensión. Nuestro enfoque es integrar los métodos y algoritmos en dichos trabajos, que nos permitan identificar, basados en datos, los modelos que posteriormente pueden ser usados para un análisis detallado de la estructura y dinámica del fenómeno.

En la Sección II se presenta la notación básica y los preliminares. Se definen los conceptos de sistema dinámico en tiempo discreto, el marco teórico general de identificación de sistemas, la estructura y representación de los modelos considerados en este trabajo.

En la Sección III se describen los métodos para resolver el problema principal de este estudio: identificación esparcida y regularizada de modelos lineales por medio de la aproximación de bajo rango.

En la Sección IV se validan las técnicas teóricas y computacionales con datos simulados y bases de datos reales DeMoor (1989). Esta es una parte fundamental del presente trabajo, poder mostrar como la aproximación de bajo rango ayudan a tratar el problema de reducción de dimensión. En la parte final V se plantean las conclusiones y bibliografía.

## II | PRELIMINARES

### 1 | Sistemas dinámicos

En términos clásicos un *sistema dinámico* es un objeto donde diferentes clases de variables se combinan para producir una *salida*, por ejemplo la Figura 1 muestra el esquema básico. Estas señales de salida

pueden ser afectadas por otras señales externas que pueden ser manipuladas por el observador llamadas *entradas*. También pueden ser señales no observables que distorsionan tanto las señales de entrada como de salida.

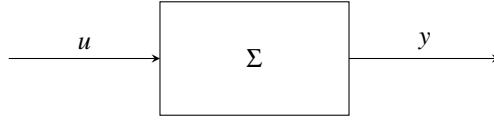


Figura 1: Bloque básico de un sistema dinámico.

El enfoque anterior de un sistema dinámico es ampliamente utilizado por la comunidad científica, sin embargo conduce hacia algunos inconvenientes detallados ampliamente en Willems (1986a, 1986b, 1987). Una de las desventajas es el hecho de asumir una partición de las variables en tipo de entrada y salida, lo cual puede conducir a resultados imprecisos. En el mismo trabajo mencionado anteriormente se presenta un enfoque alternativo conocido como *enfoque del comportamiento*, donde se corrige y define de manera más general y precisa un sistema dinámico.

**Definición 1.** Un sistema dinámico  $\Sigma$  se define como una tripleta  $\Sigma = (\mathbb{T}, \mathbb{W}, \mathcal{B})$ , donde  $\mathbb{T} \subset \mathbb{R}$  es el eje del tiempo,  $\mathbb{W}$  es el espacio de las señales y  $\mathcal{B} \subset \mathbb{W}^{\mathbb{T}}$  es el comportamiento. La forma usual de representar el comportamiento  $\mathcal{B} \subset \mathbb{W}^{\mathbb{T}}$  de un sistema dinámico es mediante ecuaciones, por ejemplo, el conjunto de funciones de la forma  $f : \mathbb{W}^{\mathbb{T}} \rightarrow \mathbb{R}^q$  o  $\mathcal{B} = \{w \in \mathbb{W}^{\mathbb{T}} \mid f(w) = 0\}$  Markovsky (2019).

En este estudio nos centraremos en el caso  $\mathbb{T} = \mathbb{Z}$ , que se conoce como *sistema dinámico discreto* y cada observación  $w \in \mathbb{W}^{\mathbb{T}}$  es llamada tradicionalmente como *serie de tiempo*.

Los sistemas *lineales invariantes en el tiempo* LTI (por sus siglas en inglés), forman la clase de sistemas dinámicos más utilizados en la práctica. Lo anterior se debe a los excelentes resultados en la mayoría de los casos aplicados a la vida real.

- Un modelo  $\mathcal{B}$  es *lineal* si es un subespacio del espacio de trayectorias  $(\mathbb{R}^q)^{\mathcal{T}}$ .
- Un modelo  $\mathcal{B}$  es *invariante en el tiempo* si  $\sigma^{\tau}\mathcal{B} = \mathcal{B}$ , donde

$$(\sigma^{\tau}w)(t) := w(t + \tau), \quad \text{para todo } t \in \mathcal{T}.$$

- Si el modelo  $\mathcal{B}$  es de dimensión finita, el comportamiento futuro del modelo es determinísticamente relacionado con un vector de *estado* de dimensión finita.
- Denotaremos por  $\mathcal{L}^w$  la clase de todos los sistemas LTI con  $w$  variables, es decir el espacio de señales  $\mathbb{W} = \mathbb{R}^w$ .

**Ejemplo 1.** Un sistema Dinámico Lineal Autónomo (sin variables de entrada) es un modelo simple para la serie  $\{x_t\} \subset \mathbb{R}^n$ , en el cual cada  $x_{t+1}$  está dado por

$$x_{t+1} = Ax_t, \quad t = 1, 2, \dots \quad (1)$$

donde  $A \in \mathbb{R}^{n \times n}$  es llamada la matriz dinámica del sistema.

**Ejemplo 2** (Modelo espacio de estados). *El sistema  $\mathfrak{S}$ , inducido por ecuaciones de entrada ( $u$ )/entrada ( $x$ )/salida ( $y$ ) con parámetros  $A, B, C, D$  se puede representar de la forma*

$$\mathfrak{S}_{i/s/o}(A, B, C, D) := \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in (\mathbb{R}^{m+n})^{\mathbb{N}} \mid \exists x \in (\mathbb{R}^n)^{\mathbb{N}} \right\} \quad (2)$$

tal que

$$\begin{aligned} x_{t+1} &= Ax_t + Bu_t, \\ y_t &= Cx_t + Du_t. \end{aligned}$$

**Ejemplo 3** (Función de Transferencia Markovsky (2006)). *El sistema  $\mathfrak{S}$ , con parámetros la matriz racional  $H \in \mathbb{R}^{n \times m}(z)$  se puede representar de la forma*

$$\mathfrak{S}_{i/o}(H) := \left\{ \begin{bmatrix} u \\ y \end{bmatrix} \in (\mathbb{R}^{m+n})^{\mathbb{N}} \mid \mathcal{F}(y) = H(z)\mathcal{F}(u) \right\} \quad (3)$$

donde  $\mathcal{F}$  es la transformada Z “18 - The z-transform” (2000).

## 2 | Identificación de Sistemas

El proceso de identificación de sistemas lo podemos resumir en tres componentes principales Soderstrom y Stoica (1989): el conjunto de datos obtenido a partir de mediciones de entrada-salida, el conjunto de modelos candidatos o estructura del modelo y el criterio de selección para el mejor modelo del conjunto de candidatos y una regla para evaluar los modelos candidatos basado en datos.

**Definición 2.** *Dada una serie de tiempo  $\Sigma = \{w_t\} \subset \mathbb{R}^n$ , un entero positivo  $L$ , denotaremos la matriz de trayectorias tipo Hankel en bloques correspondiente a  $\Sigma_T$*

$$\mathcal{H}_L(\Sigma_T) = \begin{bmatrix} w_1 & w_2 & w_3 & \cdots & w_{T-L+1} \\ w_2 & w_3 & w_4 & \cdots & w_{T-L+2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ w_L & w_{L+1} & w_{L+2} & \cdots & w_T \end{bmatrix}$$

**Teorema 1** (Descomposición en Valores Singulares Golub y Van Loan (2013)). *Sea  $A \in \mathbb{R}^{m \times n}$ , existen matrices ortogonales  $U \in \mathbb{R}^{m \times m}$  y  $V \in \mathbb{R}^{n \times n}$  tales que*

$$U^T A V = \Sigma = \text{diag}(\sigma_1, \dots, \sigma_p) \in \mathbb{R}^{m \times n},$$

donde  $p = \min\{m, n\}$ , además  $\sigma_1 \geq \sigma_2 \dots \geq \sigma_p$ .

En el resto de este documento, nos referiremos a la descomposición anterior como SVD (por sus siglas en inglés). Sea  $A = U \Sigma V^T \in \mathbb{R}^{m \times n}$  la SVD de  $A$  con  $m \geq n$ , entonces

$$A = U_1 \Sigma_1 V^T$$

donde  $U_1 = U(:, 1 : n) \in \mathbb{R}^{m \times n}$  y  $\Sigma_1 = \text{diag}(\sigma_1, \dots, \sigma_n) \in \mathbb{R}^{n \times n}$ . Nos referiremos a la versión truncada de la SVD como *SVD reducida*.

**Definición 3.** Dado  $\delta > 0$  y una matriz  $X \in \mathbb{R}^{m \times n}$ , denotaremos el rango de  $X$  por  $\text{rk}_\delta(X) > 0$  al valor determinado por

$$\text{rk}_\delta(X) = \sum_{i=1}^p \sigma_i, \quad \sigma_i > \delta,$$

donde los valores  $\sigma_i$  corresponden a los valores singulares de la SVD reducida de  $A$  y  $p = \min\{m, n\}$ .

**Definición 4** (Representación del Núcleo de un Sistema Dinámico Lineal Markovsky (2006); Vides (2021b)). Dada una serie de tiempo  $w \in \mathbb{R}^{q\mathbb{Z}}$  correspondiente a resultados de un experimento  $\mathbb{E}$ , denominaremos modelo dinámico lineal correspondiente al experimento  $\mathbb{E}$ , al conjunto  $\mathcal{B}_\mathbb{E}$  representado por la siguiente expresión

$$\mathcal{B}_\mathbb{E} := \left\{ w \in \mathbb{R}^{q\mathbb{Z}} : R_L(\sigma) = 0 \right\}, \quad (4)$$

donde  $R_L(\sigma) = R_0 + R_1\sigma + \dots + R_L\sigma^L$  es una matriz polinomial de grado  $L$  y las matrices  $R_0, R_1, \dots, R_L$  toman valores en  $\mathbb{R}^{q \times m \times n}$ .

**Lema 1** (Aproximación de Bajo Rango e Identificación de Sistemas Markovsky (2019)). La señal  $w$  es una trayectoria de un sistema LTI con  $m$  variables de entrada,  $n$  variables de salida y de orden  $L$  si y solo si

$$\text{rk}(\mathcal{H}_{L+1}(w)) \leq m(L+1) + nL.$$

**Problema 1** (Identificación de Sistemas). Dadas las mediciones de entrada  $u = \{u_0, \dots, u_N\} \subset \mathbb{R}^q$  y de salida  $y = \{y_0, \dots, y_N\} \subset \mathbb{R}^n$ , de un sistema desconocido  $\mathfrak{S}$ , encontrar un modelo (representación) razonable para el sistema basado en los datos de las mediciones  $u, y$ .

### 3 | Esparcimiento y Regularización

Con la disponibilidad de datos de alta dimensión, es indispensable contar con herramientas que permitan seleccionar las variables de mayor relevancia. En ese sentido, los métodos de regularización se han convertido en una poderosa herramienta para la aproximación y selección de modelos.

Dados  $X \in \mathbb{R}^{m \times n}$  y  $y \in \mathbb{R}^m$ , el problema de aproximación de mínimos cuadrados consiste en resolver

$$\beta = \arg \min_{\hat{\beta} \in \mathbb{R}^n} \|y - X\hat{\beta}\|_2, \quad (5)$$

Una solución conocida es utilizar la pseudo-inversa de  $X$ , para lo cual el vector solución de 5 es

$$\beta = X^\dagger b = (X^T X)^{-1} X^T y \quad (6)$$

En el caso que  $X$  esté mal condicionada, el cálculo de  $X^\dagger$  se vuelve inestable y esto puede resultar en problemas de sobre ajuste en el modelo. Para contrarrestar las situaciones de inestabilidad numérica en la práctica se utilizan técnicas de regularización.

Una técnica de regularización importante consiste en utilizar la descomposición truncada en valores singulares SVD, luego el problema resulta en

$$\beta = \min_{\hat{\beta} \in \mathbb{R}^n} \|y - X_\delta \hat{\beta}\|_2, \quad (7)$$

donde  $\delta < \text{rank}(X)$  es un parámetro y  $X_\delta \in \mathbb{R}^{m \times n}$  es la mejor aproximación de rango  $\delta$  de  $X$  calculada a partir de la SVD.

En este documento, utilizaremos resultados y algoritmos presentados en Vides (2021b), que nos permiten encontrar una solución a 7. La idea principal es utilizar métodos de aproximación de bajo rango y reducción de orden para la solución del problema lineal de mínimos cuadrados esparcido, esto es, soluciones con un número máximo de entradas diferentes de cero.

Una de las características más importantes de un modelo es la habilidad para generalizar, es decir, no solo lograr buen desempeño en un conjunto de datos dado, sino también obtener un desempeño razonable en datos nuevos. Un criterio para verificar una buena capacidad de generalizar el ajuste de un modelo es comparar predicciones de nuevos datos, con las predicciones logradas en los datos utilizados para construir el modelo.

### III | MÉTODOS

#### 1 | Aproximación de Bajo Rango

El objetivo de la aproximación de una matriz por otra matriz con rango menor o igual al original, es obtener un representación reducida de los datos con una mínima pérdida de información. Una colección de datos (trayectorias) es almacenada en  $X \in \mathbb{R}^{q \times N}$ , donde cada columna de  $X$  corresponde a un vector en  $\mathbb{R}^q$ .

**Problema 2** (Aproximación de bajo rango). *Dada  $X \in \mathbb{R}^{q \times N}$ , donde  $q \leq N$ , un entero  $r$ ,  $0 < r < n$ , encontrar*

$$\hat{X}^* := \arg \min_{\hat{X}} \|X - \hat{X}\|_F \quad \text{sujeto a: } \text{rank}(\hat{X}) \leq r.$$

**Teorema 2** (Eckart-Young-Mirsky). *Sea  $X = U\Sigma V^T$  la descomposición reducida en valores singulares de  $X$  y la siguiente partición en bloques*

$$U := \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{matrix} r & q-r \\ q & \end{matrix}, \Sigma := \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix} \begin{matrix} r & q-r \\ q-r & \end{matrix}, V := \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{matrix} r & q-r \\ N & \end{matrix},$$

entonces una solución al Problema 2 es  $\hat{X}^* = U_1 \Sigma_1 V_1^T$  y además el error de aproximación es

$$\|X - \hat{X}^*\|_F = \sqrt{\sigma_{r+1}^2 + \dots + \sigma_q^2}.$$

Lo anterior implica que es posible aproximar una matriz de datos de alta dimensión por  $r$  términos dominantes de las columnas de  $U_1$  y  $V_1$  de la forma

$$\hat{X}^* = \sum_{k=1}^r \sigma_k u_{1,k} v_{1,k}^T = \sigma_1 u_{1,1} v_{1,1}^T + \sigma_2 u_{1,2} v_{1,2}^T + \dots + \sigma_r u_{1,r} v_{1,r}^T. \quad (8)$$

El Teorema 2, también conocido simplemente como el teorema de Eckart-Young, es un resultado fundamental en álgebra lineal y teoría de aproximación. Este teorema se utiliza en el contexto de aproximación de matrices de rango bajo, lo que significa que se busca representar una matriz de manera eficiente utilizando una matriz de rango inferior.

**Algoritmo 1:** Aproximación de bajo rango  $\text{lra}$ **Datos:**  $X \in \mathbb{R}^{m \times n}$ ,  $r \in \mathbb{Z}^+$ **Resultado:**  $X_r = \text{lra}(X, r)$  $U, S, V \leftarrow \text{svd}(X)$ , descomposición reducida $s \leftarrow \min\{m, n\}$ 

$$U_r \leftarrow \sum_{j=1}^r U \hat{e}_{j,s} \hat{e}_{j,s}^T, S_r \leftarrow \sum_{j=1}^r \hat{e}_{j,s}^T S \hat{e}_{j,s}, V_r \leftarrow \sum_{j=1}^r \hat{e}_{j,s} \hat{e}_{j,s}^T V$$

$$X_r \leftarrow U_r S_r V_r^T$$

Es decir, dada una matriz  $A$  de tamaño  $n \times m$ , el teorema de Eckart-Young-Mirsky establece que para cualquier rango  $r \leq \min(n, m)$ , la mejor aproximación de rango  $r$  de  $A$  en términos de la norma de Frobenius (una medida de la magnitud de una matriz) se obtiene mediante la *SVD reducida*.

La aproximación de rango  $r$  de  $A$  se obtiene truncando la descomposición en valores singulares después de los  $r$  valores singulares más grandes, de la forma  $A_r = U_r \Sigma_r V_r^T$ , donde  $U_r, \Sigma_r, V_r^T$  son las primeras  $r$  columnas de  $U$ , las primeras  $r$  filas de  $\Sigma$  y las primeras  $r$  filas de  $V^T$ , respectivamente. Este resultado tiene aplicaciones en diversas áreas, incluyendo la reducción de dimensionalidad y la compresión de datos.

**2 | Identificación Esparcida con SpLra**

En esta sección se combinan los resultados del Teorema 2 y Teorema 3 con el objetivo de encontrar una solución aproximada al Problema 1 que contenga las propiedades de esparcimiento y regularización previamente establecidas. No se asume una partición en las variables asociadas al problema y se presenta el resultado en forma del núcleo  $R$ , lo cual permite luego escoger la representación del sistema que mejor convenga.

El siguiente resultado nos permiten encontrar soluciones regularizadas y esparcidas a tipos de problemas de la forma 7. En general, nos interesa el problema extendido cuando el vector  $y$  en 7 se convierte en una matriz. Es decir el problema matricial de mínimos cuadrados.

**Problema 3** (Vides (2021b)). *Dados  $\delta > 0$  y dos matrices  $X \in \mathbb{R}^{m \times n}$ ,  $Y \in \mathbb{R}^{m \times p}$ , denotaremos por  $Y \approx_\delta XA$  para representar el problema de encontrar  $A \in \mathbb{R}^{n \times p}$ ,  $\alpha, \beta \geq 0$  y un proyector ortogonal  $Q$  tal que*

$$\|Y - XA\|_F \leq \alpha\delta + \beta\|(I_m - Q)Y\|_F$$

La matriz  $A$  la llamaremos una solución del problema  $Y \approx_\delta XA$ .

**Teorema 3** (Vides (2021b)). *Dado  $\delta > 0$  y dos matrices  $X \in \mathbb{R}^{m \times n}$ ,  $Y \in \mathbb{R}^{m \times p}$ . Si  $r = \text{rk}(A)_\delta > 0$  entonces, existe una solución  $A$  al problema  $Y \approx_\delta XA$  con a lo sumo  $r$  entradas diferentes de cero.*

Para obtener información detallada sobre el Algoritmo 2, se recomienda revisar a Vides (2021c). Los detalles específicos del proceso y las operaciones están disponibles en dicha referencia, proporcionando una comprensión más profunda de la metodología empleada.

**Algoritmo 2:** Mínimos Cuadrados Lineal Esparcido Vides (2021b)**Datos:**  $X \in \mathbb{R}^{m \times n}$ ,  $Y \in \mathbb{R}^{m \times p}$ ,  $\delta > 0$ ,  $N \in \mathbb{Z}^+$ ,  $\varepsilon > 0$ **Resultado:**  $A = \text{Isspsolver}(X, Y, \delta, N, \varepsilon)$  $U, S, V \leftarrow \text{svd}(X)$ , descomposición reducida $s \leftarrow \min\{m, n\}$ ,  $r \leftarrow \text{rk}_\delta(X)$  $U_\delta \leftarrow \sum_{j=1}^r U \hat{e}_{j,s} \hat{e}_{j,s}^T$ ,  $T_\delta \leftarrow \sum_{j=1}^r (\hat{e}_{j,s}^T S \hat{e}_{j,s})^{-1} \hat{e}_{j,s} \hat{e}_{j,s}^T$ ,  $V_\delta \leftarrow \sum_{j=1}^r \hat{e}_{j,s} \hat{e}_{j,s}^T V$  $\hat{X} \leftarrow U_\delta^T X$ ,  $\hat{Y} \leftarrow U_\delta^T Y$  $A_0 \leftarrow V_\delta^T T_\delta \hat{Y}$ **para**  $j = 1, \dots, p$  **hacer** $K \leftarrow 1$ , error  $\leftarrow 1 + \delta$  $c \leftarrow A_0 \hat{e}_{j,p}$ ,  $x_0 = c$  $\hat{c} \leftarrow [\hat{c}_1 \ \dots \ \hat{c}_n]^T = [[\hat{e}_{1,n}^T c \ \dots \ |\hat{e}_{n,n}^T c]]^T$ Calcular la permutación  $\sigma : \{1, \dots, n\} \rightarrow \sigma : \{1, \dots, n\}$  tal que:  $\hat{c}_{\sigma(1)} \geq \hat{c}_{\sigma(2)} \geq \dots \geq \hat{c}_{\sigma(n)}$  $N_0 \leftarrow \max \left\{ \sum_{j=1}^n H_\varepsilon(\hat{c}_{\sigma(j)}), 1 \right\}$ **mientras**  $K \leq N$  **and** error  $> \delta$  **hacer**|  $x \leftarrow \mathbf{0}_{n,1}$ **fin****fin**

**Definición 5** (Representación del Núcleo de un Sistema Dinámico Lineal Markovsky (2006); Vides (2021b)). Dada una serie de tiempo  $w \in \mathbb{R}^{q\mathbb{Z}}$  correspondiente a resultados de un experimento  $\mathbb{E}$ , denominaremos modelo dinámico lineal correspondiente al experimento  $\mathbb{E}$ , al conjunto  $\mathcal{B}_{\mathbb{E}}$  representado por la siguiente expresión

$$\mathcal{B}_{\mathbb{E}} := \left\{ w \in \mathbb{R}^{q\mathbb{Z}} : R_L(\sigma) = 0 \right\}, \quad (9)$$

donde  $R_L(\sigma) = R_0 + R_1\sigma + \dots + R_L\sigma^L$  es una matriz polinomial de grado  $L$  y las matrices  $R_0, R_1, \dots, R_L$  toman valores en  $\mathbb{R}^{q \times m \times n}$ .

**Problema 4** (Identificación exacta). Dada la serie de tiempo  $w \in \mathbb{R}^{m+n}$ , generada a partir de un sistema desconocido  $\mathfrak{S}$  y el orden del sistema  $\text{lag}=L$ . Sea  $\mathcal{H}_{L+1}(w)$  la matriz de trayectorias tipo Hankel por bloques de  $w$ , podemos expresar una representación del modelo  $\mathfrak{S}$  de la forma

$$\tilde{\mathcal{B}} := \left\{ R \in \widehat{\mathcal{U}} : R \mathcal{H}_{L+1}(w) = 0 \right\} \quad (10)$$

**Problema 5** (Identificación aproximada). Una representación alternativa aproximada del Problema 4 se puede escribir de la forma

$$\tilde{\mathcal{B}} := \left\{ R \in \mathbb{R}^{n \times (L+1)(m+n)} := \arg \min_R \left\| \widehat{R} \mathcal{H}_{L+1}(w) \right\|^2 \right\} \quad (11)$$

El algoritmo propuesto  $\text{SpLra}$ , para resolver el Problema 5, se puede resumir en los siguientes pasos:

1. Construir la matriz de trayectorias tipo Hankel por bloques  $\mathcal{H}_{L+1}(w)$ .



2. Aplicar el Algoritmo 1 (lra) a  $\mathcal{H}_{L+1}(w)$  para calcular

$$D(\mathcal{H}_{L+1}(w)) = \text{lra}(\mathcal{H}_{L+1}(w), r),$$

donde  $r = (L+1)(m+n) - n$ .

3. Hacer la partición en bloques

$$D(\mathcal{H}_{L+1}(w)) = \begin{bmatrix} D_0 \\ D_1 \end{bmatrix},$$

donde  $D_0 \in \mathbb{R}^{r \times N}$  y  $D_1 \in \mathbb{R}^{n \times N}$ .

4. Calcular  $x_{sp}$ , la solución esparcida del sistema  $D_1 = x_{sp}D_0$

$$x_{sp} = \text{lsspsolver}(D_0, D_1, \varepsilon, \delta),$$

5. Hacer  $R_{sp} = \begin{bmatrix} -x_{sp} & I_n \end{bmatrix}$ .

---

**Algoritmo 3:** SpLra.

---

**Datos:**  $w \in \mathbb{R}^{(m+n) \times N}$ ,  $\delta > 0$ ,  $L \in \mathbb{Z}^+$ ,  $\varepsilon > 0$

**Resultado:**  $R_{sp} = \text{SpLra}(w, L, \delta, \varepsilon)$

$\mathcal{H}_{L+1}(w) \leftarrow \text{blkhank}(w, L+1)$ , matriz de trayectorias tipo hankel en bloques.

$r \leftarrow (L+1)(m+n) - n$

$D \leftarrow \text{lra}(\mathcal{H}_{L+1}(w), r)$ , aproximación de bajo rango

$$D_0 \in \mathbb{R}^{r \times N}, D_1 \in \mathbb{R}^{n \times N} \leftarrow D = \begin{bmatrix} D_0 \\ D_1 \end{bmatrix}$$

$x_{sp} \leftarrow \text{lsspsolver}(D_0, D_1, \delta, N, \varepsilon)$

$R_{sp} \leftarrow \begin{bmatrix} x_{sp} & -I_w \end{bmatrix}$

---

La capacidad de generalización es crucial en un modelo, destacando su capacidad para rendir bien no solo en un conjunto de datos específico, sino también al enfrentarse a nuevos datos. Para evaluar esta capacidad, es común comparar las predicciones del modelo en un conjunto de datos de prueba con aquellas realizadas en los datos utilizados para construir el modelo.

Idealmente, se busca un rendimiento similar en ambos conjuntos; de lo contrario, se concluye que el modelo carece de habilidad de generalización. Una práctica común para validar el modelo implica dividir los datos en un conjunto de entrenamiento (*id*) entre 50-90% y un conjunto de prueba (*val*), siendo este último completamente oculto para evaluar el rendimiento del modelo en datos nuevos.

En resumen, se busca identificar modelos que ofrezcan buenas predicciones tanto en el conjunto de entrenamiento como en el de prueba, utilizando métodos como la identificación por mínimos cuadrados con diferentes valores de  $L$  y seleccionando aquel modelo que presente el menor error en la predicción, siendo coherente en ambos conjuntos de datos.

## IV | APLICACIONES

En este documento, se destaca la importancia de asegurar la aplicabilidad de las técnicas propuestas en el proceso de investigación. Se llevaron a cabo experimentos replicables, que abarcaron tanto datos simulados como reales. La evaluación de la efectividad y rendimiento de las técnicas comenzó con datos simulados para comprender su comportamiento en condiciones ideales. Además, se realizaron experimentos con datos reales de diversas áreas, permitiendo evaluar la eficiencia de las técnicas en situaciones más desafiantes y complejas del mundo real.

Es importante destacar que todos los experimentos realizados están completamente disponibles para su replicación. La información detallada, los datos utilizados y los procedimientos empleados se encuentran accesibles en el repositorio Cardona (2022), permitiendo a otros investigadores replicar y validar los resultados de manera integral.

### 1 | Simulación e Identificación de un Sistema SISO

El siguiente experimento está basado en algunas ideas presentadas en Armenise, Vaccari, Di Capaci, y Pannocchia (2018), los datos del proceso de identificación son simulados utilizando un sistema de una entrada y una salida (SISO por sus siglas en inglés) con la siguiente función de transferencia

$$G(z) = \frac{1.5z^2 - 2.07z + 1.315}{z^{14} - 2.21z^{13} + 1.7497z^{12} - 0.5843z^{11} + 0.0684z^{10}} \quad (12)$$

La señal de entrada es una secuencia pseudo-aleatoria binaria (PRBS por sus siglas en inglés), con una probabilidad de cambio de 8% y un rango entre  $[-1, 1]$ . La Figura 2a muestra la salida gráfica de las señales de entrada y salida  $u_t, y_t$ , respectivamente. La Figura 2b es una representación gráfica de los polos y ceros de la función de transferencia en el plano complejo. Los polos y ceros determinan el comportamiento dinámico y la estabilidad de un sistema, por lo que es crucial transferir estas propiedades al sistema identificado en nuestro análisis.

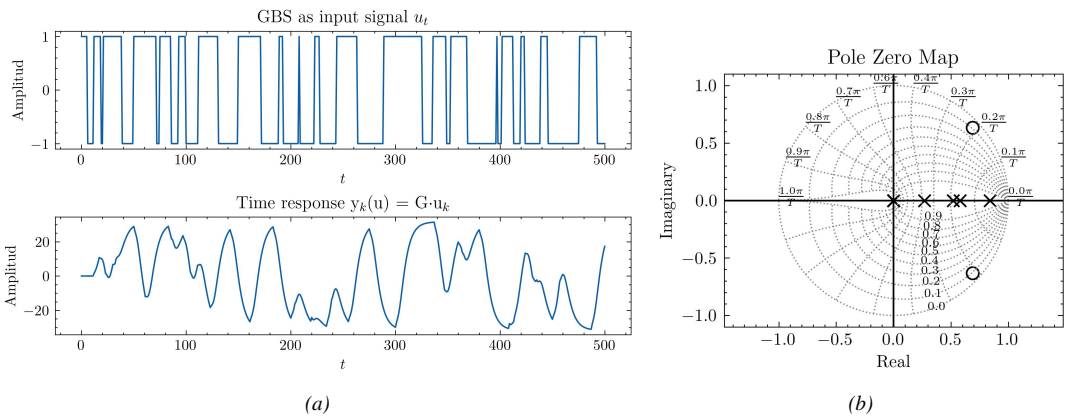


Figura 2: (a) Salida gráfica de la serie de tiempo simulada. (b) Polos y ceros del sistema simulado original, figura generada por el paquete *Control* de Python.

Con el objetivo de validar los diferentes modelos identificados, se generan nuevos datos a partir de otra señal PRBS. Estos datos generados no participan en el proceso de identificación. La Figura 3 muestra la salida gráfica de los datos perturbados utilizando una estructura ARMAX, generando cuatro señales independientes de ruido blanco,  $\sigma_1^2 = 0.000$ ,  $\sigma_2^2 = 0.001$ ,  $\sigma_3^2 = 0.010$  y  $\sigma_4^2 = 0.1000$ . Lo anterior con el objetivo de elevar el grado de dificultad del proceso de identificación. Las señales de ruido blanco son filtradas por la siguiente función de transferencia

$$H(z) = \frac{z^{14} + 0.3z^{13} + 0.2z^{12}}{z^{14} - 2.21z^{13} + 1.7497z^{12} - 0.5843z^{11} + 0.0684z^{10}} \quad (13)$$

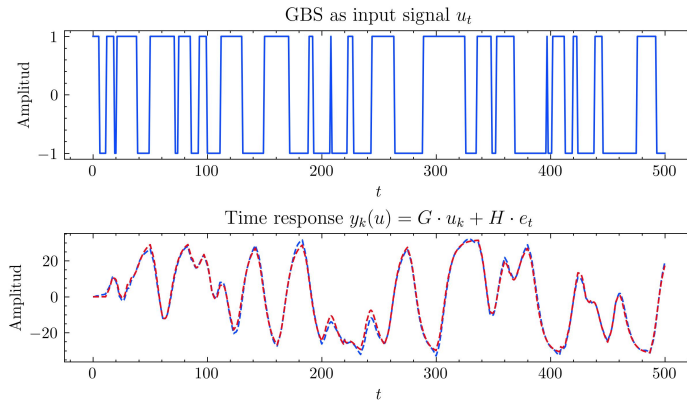


Figura 3: Serie de tiempo del sistema perturbado con una estructura ARMAX ( $y_k = Gu_k + He_k$ ) con  $\sigma^2 = 0.010$ .

Tres tipos de modelos son identificados, los modelos ARX y ARMAX son calculados utilizando el paquete System Identification Package for Python (SIPPY) Armenise y cols. (2018) y el modelo SpLra con los algoritmos presentados en Cardona (2022). La dificultad de este experimento radica en seleccionar el orden adecuado del sistema, esto debido a un nivel alto de desfase en la función de transferencia  $G(z)$ . El paquete SIPPY cuenta con tres criterios de selección del orden de los modelos, AIC, BIC y AcIC, los cuales son los criterios mayormente utilizados. En el método SpLra se utiliza un método gráfico que consiste en definir un número máximo para  $L$  y graficar los diferentes valores con respecto al desajuste producido por cada modelo.

La Figura 4a presenta una comparación de los resultados obtenidos durante el proceso de validación. Es importante destacar que los datos de validación no se emplearon en el proceso de identificación, con el propósito de demostrar la capacidad del modelo identificado para simular datos nuevos. Por otro lado, la Figura 4b ilustra la gráfica de polos y ceros del modelo identificado, permitiendo su comparación con la Figura 2b.

La Tabla 1 proporciona un resumen detallado de los resultados obtenidos para diferentes modelos identificados, utilizando  $\text{nnz}(\%)$  para evaluar el porcentaje de elementos distintos de cero en los parámetros del modelo y  $\%fit$  para cuantificar el porcentaje de ajuste del modelo de la forma

$$\%fit = 100 \left( 1 - \frac{\|y - y_h\|}{\|y - \bar{y}\|} \right) \quad (14)$$

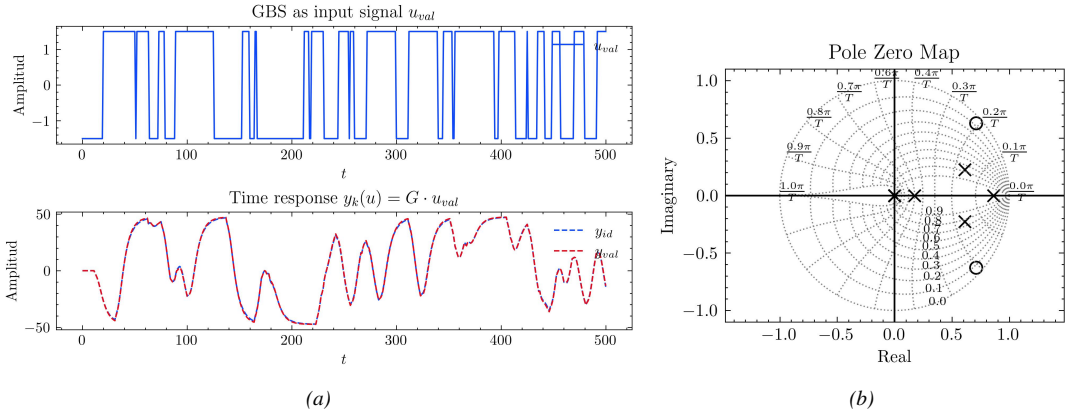


Figura 4: (a) Serie de tiempo para la fase de validación. (b) Polos y ceros del sistema identificado.

	Modelo					
	Arx		Armax		SpLra	
nnz (%)	50		45		26	
$\sigma^2$	%fit(id)	%fit(val)	%fit(id)	%fit(val)	%fit(id)	%fit(val)
0.000	81.97	80.67	81.96	80.67	99.99	99.99
0.001	81.81	80.47	$-\infty$	$-\infty$	99.28	99.33
0.010	83.13	82.12	$-\infty$	$-\infty$	96.55	96.37
0.100	78.05	77.47	$-\infty$	$-\infty$	91.81	92.64

Tabla 1: Resumen de resultados.

Notablemente, el método propuesto SpLra destaca al lograr un rendimiento superior en el ajuste del modelo, y lo hace con un menor número de parámetros en comparación con los modelos Arx y Armax en todos los casos examinados.

Es crucial destacar que los valores de %fit(id) y %fit(val) indican el ajuste en conjuntos de datos de entrenamiento y validación, respectivamente. La consistencia en el rendimiento entre estos conjuntos sugiere que el método SpLra no solo se adapta bien a los datos utilizados para construir el modelo, sino que también generaliza de manera efectiva a nuevos datos, lo cual es fundamental para la capacidad de un modelo de aplicarse en situaciones del mundo real.

La presencia de  $\%fit = -\infty$  en los resultados para algunos modelos indica un fallo en la estabilidad del modelo ajustado, subrayando la importancia de la estabilidad en la identificación de sistemas. Este hallazgo refuerza la idea de que, además del ajuste, la estabilidad del modelo es un factor crítico a considerar en la evaluación de la calidad de los modelos identificados.

## 2 | Aplicaciones en Procesos Industriales

La base de datos para la identificación de sistemas DAISY DeMoor (1989) contiene diferentes conjuntos de datos reales en varias disciplinas, por ejemplo, sistemas en procesos industriales, sistemas eléctricos y electrónicos, sistemas mecánicos, entre otros. A continuación una breve descripción de los datos seleccionados.

**Brazo robótico flexible:** Estos datos son el resultado del movimiento de un brazo robótico conectado a un motor eléctrico. Los datos modelan la función de transferencia a partir de la reacción por el esfuerzo de torsión en la aceleración de la estructura. Los datos de entrada son el esfuerzo de torsión de reacción de la estructura y la salida la aceleración del brazo robótico.

**Secador de pelo:** Experimento de laboratorio que funciona como un secador de pelo. El aire es ventilado por medio de un tubo y calentado en su entrada. La temperatura del aire es medida cuando sale del tubo. Como entrada se mide el voltaje sobre la calefacción, la cual es una malla de cables de resistencia. La entrada es el voltaje del dispositivo de calentamiento y la salida la temperatura del aire de salida.

**Reactor de un tanque:** Este proceso modela el flujo continuo de un reactor de tanque de agitación, donde la reacción exotérmica y la concentración son regulados mediante un refrigerante. La entrada es el flujo del refrigerante ( $l/min$ ) y las salidas son la concentración del flujo ( $mol/l$ ) y la temperatura en (grados  $K$ ).

**Secador industrial:** Datos de un secador industrial por Cambridge Control Ltd. Las entradas son la tasa de flujo de combustible, velocidad del extractor de gas caliente y tasa de flujo de materia prima. Como salidas la temperatura de bulbo seco, temperatura de bulbo húmedo y el contenido de humedad de la materia prima.

La Tabla 2 presenta la cantidad de datos  $N$ , el número de variables de entrada  $m$ , cantidad de variables de salida  $n$  y el número de observaciones pasadas  $lag$  a considerar para los conjuntos de datos seleccionados. Para el proceso de validación se separaron los datos en 50% (id) y el resto para (val).

No.	Nombre	N	$m$	$n$	lag
1	Brazo robótico flexible	1024	1	1	4
2	Secador de pelo	1000	1	1	5
3	Reactor de un tanque	7500	3	3	1
4	Secador industrial	867	3	3	1

Tabla 2: Selección de conjuntos de datos DAISY.

La Tabla 3 resume los resultados obtenidos en este experimento. Para los datos 1 y 2 (sistemas SISO) se consideraron los métodos  $ARX$  y  $N4SID$ , el último es un algoritmo de subespacio no iterativo implementado en el paquete  $SIPPY$  que calcula directamente una representación del modelo en variables de espacio de estados. Para los datos 3 y 4 (sistemas MIMO) se consideran los métodos  $N4SID$  y  $MOESP$ , el segundo, también implementado en el paquete  $SIPPY$  es una versión mejorada del primero.

Los resultados muestran que el método propuesto  $SpLra$  tiene un desempeño sobresaliente en los diferentes conjuntos de datos seleccionados. Este método, que incorpora esparcimiento, regularización y aproximación de bajo rango, ha demostrado ser efectivo en la selección del orden del modelo, logran-

No.	Nombre	Método	%fit(id)	%fit(val)	nnz(%)
1	Brazo robótico	Arx	85.27	82.39	100
		SpLra	96.99	92.64	38.78
		N4SID	81.60	91.48	97.96
2	Secador de pelo	Arx	87.45	82.39	66.66
		SpLra	93.34	92.24	36.11
		N4SID	87.59	82.11	97.22
3	Reactor de un tanque	SpLra	84.11	82.76	75.00
		N4SID	84.22	86.25	83.33
		MOESP	84.16	86.17	83.33
4	Secador industrial	SpLra	61.84	53.87	75.00
		N4SID	58.27	36.01	75.00
		MOESP	58.32	35.97	75.00

Tabla 3: Resultados base de datos Daisy.

do un ajuste significativamente mejor y una reducción en el número de parámetros aproximados en comparación con los métodos Arx, N4SID y MOESP.

En particular, al observar los resultados para los conjuntos de datos específicos, se destaca que SpLra ha superado a los métodos tradicionales como Arx y N4SID en términos de porcentaje de ajuste tanto en los conjuntos de entrenamiento (*id*) como en los conjuntos de validación (*val*). Además, el análisis del porcentaje de parámetros no nulos (%nnz) revela que SpLra logra una mayor eficiencia al mantener un número reducido de parámetros sin comprometer significativamente la calidad del ajuste.

En resumen, los resultados respaldan la eficacia del método SpLra en la identificación de modelos, destacando su capacidad para lograr un equilibrio óptimo entre la complejidad del modelo y su capacidad predictiva. Estos hallazgos tienen implicaciones significativas en el diseño y la optimización de modelos en una variedad de campos, desde el control de procesos hasta la predicción de comportamientos dinámicos en sistemas físicos complejos.

## V | CONCLUSIONES

Los experimentos realizados, tanto en datos simulados como en datos reales, revelan resultados altamente prometedores del método propuesto SpLra. La integración de técnicas de esparcimiento, regularización y aproximación de bajo rango ha demostrado ser efectiva en diversos aspectos del modelado de sistemas, proporcionando un enfoque sólido y versátil.

Una ventaja distintiva de SpLra es su independencia del método de representación del sistema, lo que le permite adaptarse sin problemas a una variedad de escenarios, ya sean sistemas SISO o MIMO, y

modelos expresados en variables de espacio de estados o función de transferencia. Esta flexibilidad lo convierte en una herramienta versátil y adaptable en el ámbito del modelado de sistemas.

Los resultados también indican que SpLra logra una aproximación precisa de las propiedades del sistema original en datos simulados y destaca en la identificación de modelos estables en datos reales, lo cual es crucial en aplicaciones del mundo real. Estos logros respaldan la utilidad y la efectividad del enfoque SpLra en la identificación y modelado de sistemas complejos.

En resumen, este estudio subraya el potencial y las ventajas significativas del método SpLra en la selección de órdenes de modelo, el ajuste de datos y la aproximación de parámetros. Su versatilidad y capacidad para mantener una representación precisa independientemente del tipo de sistema hacen que SpLra sea una herramienta valiosa en la investigación y aplicación de modelos de sistemas dinámicos.

## I REFERENCIAS

- 18 - the z-transform. (2000). En A. Jeffrey y D. Zwillinger (Eds.), *Table of integrals, series, and products (sixth edition)* (Sixth Edition ed., p. 1127-1132). San Diego: Academic Press. Descargado de <https://www.sciencedirect.com/science/article/pii/B9780122947575500222>
- Armenise, G., Vaccari, M., Di Capaci, R. B., y Pannocchia, G. (2018). An open-source system identification package for multivariable processes. En *2018 ukacc 12th international conference on control (control)* (p. 152-157).
- Cardona, K. (2022). *Sparse low rank approximation*. Descargado de <https://github.com/kerin84/SpLra>
- DeMoor, B. (1989). *Daisy: Database for the identification of systems*. Department of Electrical Engineering, ESA-T/STADIUS, KU Leuven, Belgium. Descargado de <https://homes.esat.kuleuven.be/~smc/daisy/daisydata.html>
- Fan, J., y Kim, D. (2017). Structured Volatility Matrix Estimation for Non-Synchronized High-Frequency Financial Data. *SSRN Electronic Journal*. Descargado 2021-07-26, de <https://www.ssrn.com/abstract=3085737> doi:
- Golub, G. H., y Van Loan, C. F. (2013). *Matrix computations* (Fourth edition ed.). Baltimore: The Johns Hopkins University Press. (OCLC: ocn824733531)
- Markovsky, I. (Ed.). (2006). *Exact and approximate modeling of linear systems: a behavioral approach*. Philadelphia: Society for Industrial and Applied Mathematics. (OCLC: ocm62525241)
- Markovsky, I. (2019). *Low-Rank Approximation: Algorithms, Implementation, Applications*. Cham: Springer International Publishing. Descargado 2022-04-06, de <http://link.springer.com/10.1007/978-3-319-89620-5> doi:
- Soderstrom, T., y Stoica, P. (1989). *System identification*. New York: Prentice Hall.
- Vides, F. (2019, noviembre). On Cyclic Finite-State Approximation of Data-Driven Systems. En *2019 IEEE 39th Central America and Panama Convention (CONCAPAN XXXIX)* (pp. 1–6). Guatemala City, Guatemala: IEEE. Descargado 2021-07-26, de <https://ieeexplore.ieee.org/document/8977039/> doi:
- Vides, F. (2021a, octubre). Computing Semilinear Sparse Models for Approximately Eventually Periodic Signals. *arXiv:2110.08966 [cs, eess, math]*. Descargado 2022-03-18, de <http://arxiv.org/abs/2110.08966> (arXiv: 2110.08966)
- Vides, F. (2021b). *Sparse system identification by low-rank approximation*. <https://arxiv.org/abs/2105.07522>.
- Vides, F. (2021c). *A toolset with matlab, python and julia programs for approximate sparse system identification*. Descargado de <https://github.com/FredyVides/SDSI>
- Vides, F., y Nogueira, I. B. R. (2023, 9). *Dynamic financial processes identification using sparse regressive reservoir computers*. <https://arxiv.org/pdf/2310.12144.pdf>.
- Willems, J. C. (1986a). From time series to linear system—part i. finite dimensional linear time invariant systems. *Automatica*, 22(5), 561-580. Descargado de <https://www.sciencedirect.com/science/article/pii/000510988690066X>

- Willems, J. C. (1986b). From time series to linear system—part ii. exact modelling. *Automatica*, 22(6), 675-694.  
Descargado de <https://www.sciencedirect.com/science/article/pii/0005109886900051>
- Willems, J. C. (1987, enero). From time series to linear system—part iii: Approximate modelling. *Automatica*, 23(1), 87-115.