

IDENTIFICACIÓN DE PARÁMETROS DE SISTEMAS DINÁMICOS A TRAVÉS DE REDES NEURONALES ARTIFICIALES

Juan Giró^{a,b}, Alejandro García^b y José Stuardi^a

^a Departamento de Estructuras, Facultad de C. E. F. y N., Universidad Nacional de Córdoba, Casilla de Correo 916, 5000 Córdoba, Argentina, juanfgiro@gmail.com, <http://www.efn.uncor.edu>

^b Departamento de Ingeniería de Sistemas de Información, Facultad Regional Córdoba, Universidad Tecnológica Nacional, Córdoba, Argentina, secresis@bbs.frc.utn.edu.ar, <http://www.frc.utn.ar>

Keywords: Identificación de Parámetros, Sistemas Dinámicos, Modelos Regresivos, Redes Neuronales Artificiales, Redes Adaline, Redes Multicapa de Perceptrones.

Abstract. Las técnicas de identificación de parámetros de sistemas dinámicos lineales y no lineales muestran en la actualidad una clara orientación hacia los modelos de caja negra, ocupando allí un lugar destacado las Redes Neuronales Artificiales. En este trabajo se presenta un procedimiento para la identificación de parámetros de sistemas dinámicos lineales de dos etapas: en la primera se ajusta un modelo regresivo a partir de los registros temporales de excitación y respuesta, y en la segunda se identifican sus parámetros (matrices de rigidez y amortiguamiento) y características dinámicas (frecuencias y modos de vibración) en base al modelo anterior. Para la primera etapa se utilizan Redes Neuronales Artificiales, de tipo Adaline y Perceptrones multicapa. La segunda etapa es totalmente formulada a través del álgebra matricial, lo que facilita su implementación sistemática y la independiza de la complejidad o dimensión del sistema estudiado. El procedimiento propuesto está destinado a operar a partir de registros experimentales, por lo que se presta especial atención a la sensibilidad de los resultados al intervalo de los datos y a la presencia de ruido en las señales de entrada. Para esto último, a las respuestas correctas obtenidas en condiciones ideales se le incorporan diversos niveles de ruido, que responden a funciones de distribución Gaussiana, con media nula y desviación estándar especificada. Se presenta la justificación del procedimiento propuesto, los resultados obtenidos con los modelos regresivos y un estudio de sensibilidad de los resultados a la variación de la calidad de los datos disponibles.

1 INTRODUCCIÓN

El objetivo de la identificación de parámetros es desarrollar o mejorar la representación matemática de un sistema físico usando datos experimentales. Es la forma de establecer un puente entre el dominio de la realidad y el de los modelos que pretenden representarla, contribuyendo a entender mejor la primera y perfeccionar estos últimos. En este trabajo se centra la atención en la identificación de parámetros de sistemas mecánicos, determinando los valores característicos de sus modelos a partir de su respuesta a la acción de cargas que varían en el tiempo. Este enfoque constituye un planteo inverso al habitual, el analítico, que se presenta cuando se busca determinar la respuesta de cierta estructura a las cargas aplicadas. Por otra parte, existe una segunda forma de planteo inverso, cuya finalidad es identificar las cargas necesarias para que cierto sistema se comporte acorde a una respuesta previamente establecida, lo que representa un problema clásico de la teoría de control. Cabe aquí aclarar que cuando se habla de planteos inversos no se lo hace con un sentido estrictamente matemático, sino más bien con un sentido conceptual.

Naturalmente, para identificar los parámetros de un sistema dinámico primero hay que acotarlo dimensionalmente, razón por la cual este proceso implica la definición de un sistema equivalente, que es normalmente más simple que el sistema o estructura que se espera representar. Así, la identificación de parámetros resulta un medio adecuado para la obtención de sistemas equivalentes y los valores característicos obtenidos representan propiedades que podrían denominarse condensadas, tanto másicas como disipativas y elásticas. El calificativo de condensadas tiene su origen en el planteo directo del problema, donde a través del método de los elementos finitos se obtienen las matrices de inercia, amortiguamiento y rigidez del modelo y posteriormente la cantidad de grados de libertad es reducido en concordancia con el tipo de estudio que se quiere realizar.

Cabe acotar que en la actualidad se comprueba un creciente número de líneas de investigación que centran su atención en el modelado e identificación de parámetros de sistemas dinámicos. Además, los recientes desarrollos en tecnología computacional, en áreas tales como la adquisición de datos, el tratamiento de señales y el procesamiento masivamente paralelo de datos, contribuyen a acentuar esta tendencia.

Estas técnicas de identificación de parámetros reconocen una primer clasificación según se apoyen o no en modelos físicos de los sistemas representados. Para el primer caso se recurre al ya mencionado método de los elementos finitos y para el segundo a modelos denominados de “caja negra”. En este último caso la definición del modelo se basa en el comportamiento dinámico del sistema, es decir en los registros temporales de las excitaciones y sus respuestas, prescindiendo del planteo de la formulación matemática del fenómeno estudiado. Dentro de estos enfoques de “caja negra” están ocupando un lugar destacado los modelos autoregresivos y las redes neuronales artificiales. En particular, hay tres aspectos que justifican la creciente utilización de distintas variantes de redes neuronales en el modelado de sistemas dinámicos: 1) su aptitud para aproximar funciones complejas y altamente no lineales, lo que permite desarrollar modelos que servirán para la posterior identificación de parámetros, 2) su capacidad de aprendizaje, lo que hace posible ajustar el modelo a través de procesos que, con los algoritmos actualmente disponibles, han reducido mucho los clásicos problemas de convergencia y 3) la tolerancia de las redes neuronales a datos imprecisos o incompletos, los que las hace especialmente apropiadas para aplicaciones en las que se opera con datos originados en mediciones sobre el mismo sistema.

Con referencia al primer aspecto, Kolmogorov demostró que toda red que disponga de por lo menos una capa oculta y contenga en ella una adecuada cantidad de unidades convenientemente entrenadas adquirirá el carácter de aproximador universal (Valishevsky, 1998), es decir, será capaz de reproducir cualquier función continua y no lineal que pueda estar definida en cierto hiperespacio.

En segundo lugar, se debe observar que las redes neuronales artificiales son sistemas masivamente paralelos, que materializan su capacidad de aprendizaje a través de la modificación de los pesos de interconexión entre sus neuronas. Más aún, existe la posibilidad de alterar en forma dinámica la topología de una red, es decir, su número de capas, la cantidad de neuronas en cada capa o la manera como éstas son interconectadas, lo que ofrece otras formas menos convencionales de aprendizaje y brinda gran flexibilidad a este proceso. Todas estas técnicas tienen su antecedente en el método de Backpropagation (Rumelhart et al., 1986), que estimuló una verdadera revolución en el campo de las redes neuronales y se convirtió en un clásico.

El último aspecto señalado es la tolerancia a datos imprecisos o incompletos. En efecto, las redes neuronales son capaces de reconocer patrones a partir de señales con ruido, distorsionadas o incompletas, debido a que la información está distribuida en las conexiones entre neuronas y además en este tipo de almacenamiento existe cierto grado de redundancia. Estudios recientes (Chandra y Singh, 2003) han demostrado que la redundancia es el factor condicionante tanto de la tolerancia a ruido como también de la capacidad de generalización, y que ambas son a su vez propiedades mutuamente relacionadas entre sí.

Cuando se mencionan los buenos desempeños que exhiben las redes neuronales artificiales como aproximadores de funciones desconocidas, también deben reconocerse sus debilidades. En su mayoría, éstas se refieren a aspectos específicos de su configuración y entrenamiento, donde se presentan múltiples posibilidades y se carece de recomendaciones definitivas que permitan seleccionar las más convenientes para cada caso. Aquí deben mencionarse: a) la arquitectura de la red, en lo que se refiere a la cantidad de capas, las unidades por capa y los vínculos entre ellas, b) las funciones de activación, que pueden ser lineales, hiperbólicas, sigmoidales o una combinación de ellas, c) los pesos iniciales más apropiados y d) las técnicas para el proceso de entrenamiento. A pesar de que estos aspectos son objeto de intenso estudio, por el momento no puede evitarse una tarea de exploración hasta encontrar la combinación más conveniente para cada caso.

Sin embargo, en los casos en que es posible anticipar la forma general de las expresiones diferenciales que gobiernan el problema, como aquí ocurre, se puede aplicar este conocimiento para definir modelos lo más específicos posible. De esta manera, al representarse fenómenos conocidos se reducen drásticamente las opciones que presenta el modelo neuronal. La idea de que la arquitectura de las redes neuronales debe estar inspirada en el conocimiento sobre la naturaleza de los fenómenos representados fue postulada por numerosos autores (Haecker y Rudolph, 2001), quienes aseguraron que ese es el camino para obtener modelos simples, eficientes y previsibles.

Lo anteriormente mencionado permite afirmar que las redes neuronales artificiales ofrecen una opción válida para implementar modelos inversos de sistemas dinámicos y en el procedimiento que aquí se presenta se las adopta como base para la posterior identificación de sus parámetros. En este caso el tipo de redes empleadas son las llamadas “hacia adelante” o “feedforward”, en el sentido que establecen una relación directa entre entrada y salida que emula el comportamiento del sistema representado, sin ciclos de realimentación o procesos recurrentes. En particular, se emplean unidades Adaline, redes Madaline y de Perceptrones.

En este trabajo se propone un procedimiento para la identificación de parámetros de sistemas dinámicos que reconoce dos etapas, tal como se muestra en el esquema de la **Figura 1**. La primera etapa está destinada a ajustar un modelo regresivo a partir de los registros temporales de excitación y respuesta del sistema real. Este modelo tiene la finalidad de filtrar los ruidos de las señales de la respuesta y de reproducir el comportamiento del sistema ante condiciones iniciales que no serían factibles de imponer en el sistema real. Se trata de un modelo que es implementado con algunos de los tipos de redes neuronales ya mencionados.

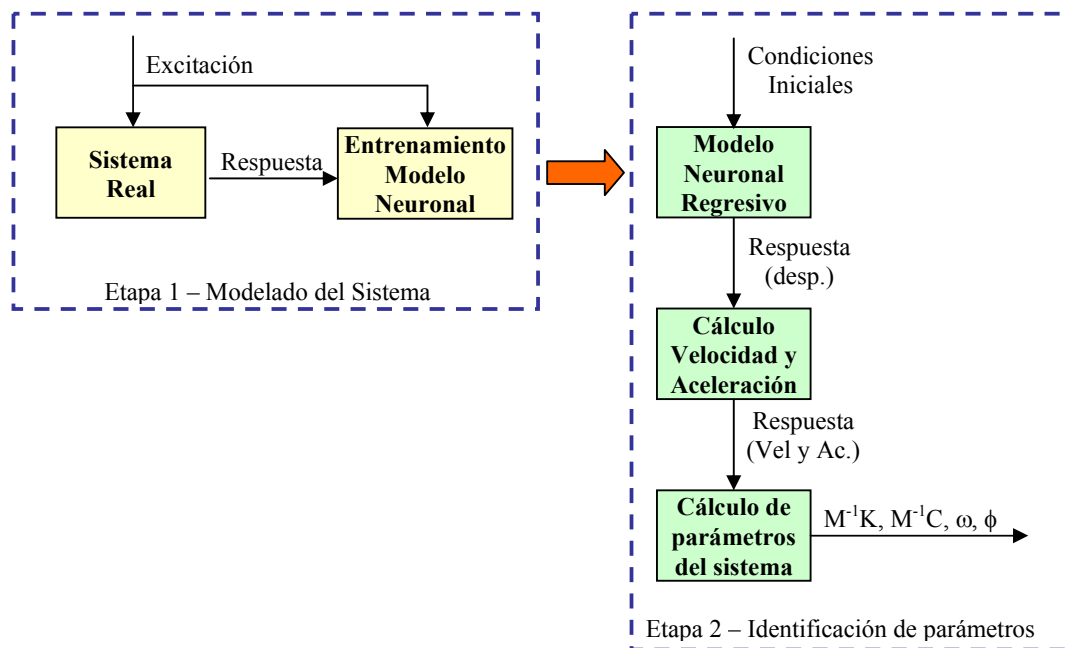


Figura 1: Modelado e identificación de parámetros

En la segunda etapa se identifican los parámetros del sistema (matrices de rigidez y amortiguamiento) y sus características dinámicas (frecuencias y modos de vibración) a partir de la respuesta (desplazamientos) obtenida con el modelo regresivo. Para ello se obtienen la velocidad y aceleración de la respuesta, que son luego incorporados a un procedimiento de álgebra matricial.

El resto de este trabajo se organiza de la siguiente manera, en la Sección 2 se justifica el modelo neuronal regresivo que es adoptado para representar el sistema, y luego, en la Sección 3 se formula el procedimiento que se apoya en ese modelo para determinar sus parámetros y características dinámicas. La sensibilidad del modelo es estudiada en la Sección 4 y finalmente, en la Sección 5 se presentan las conclusiones de este trabajo e ideas para futuros desarrollos.

2 MODELOS REGRESIVOS

Como ya fue mencionado, en la primera etapa se emula el comportamiento de una función desconocida a partir de los valores de su respuesta a la acción de cierta condición de excitación. Este método suele ser denominado **ARX** (**A**uto **R**egresivo y **E**xógeno) en alusión a que: 1) cada valor de respuesta “y” es determinado a partir de otros valores obtenidos por el mismo modelo, 2) estos corresponden a instantes anteriores y 3) representan la respuesta de un sistema a una acción externa “u”. La expresión general para la respuesta es la siguiente:

$$y_k = \sum_{j=1}^n \alpha_j y_{k-j} + \sum_{j=1}^m \beta_j u_{k-j} + e_k \quad (1)$$

donde “n” es la cantidad de valores de la respuesta correspondientes a instantes anteriores y “m” es la cantidad de valores de la excitación. Notar que el término e_k representa el error del modelo y “n” representa su orden. También hay que tener presente que normalmente se emplean valores de $m \leq n$, y que en el caso de tratarse de un sistema de “q” grados de libertad, con $q > 1$, tanto la respuesta “y” como la acción “u” quedan representadas por vectores, mientras que α y β quedan representados por matrices.

En la literatura es frecuente encontrar referencias sobre la imposibilidad de predecir la cantidad de valores de “n” y “m” que son más convenientes para resolver cada problema específico (Aguado y Martínez, 2003). Sin embargo, debe considerarse que en la representación de funciones se distinguen dos casos que son esencialmente distintos, como es el de representar fenómenos absolutamente desconocidos y el de modelar sistemas que responden a principios físicos claramente establecidos. En el primer caso se trata de verdaderos modelos de cajas negras, solo expresables a través de sus relaciones de entrada-salida, mientras que en el segundo es posible anticipar, como mínimo, la forma general de las expresiones diferenciales que gobiernan el problema. Como ya fue anticipado, es recomendable aplicar este conocimiento para guiar la definición del modelo y hacerlo tan específico como sea posible.

Considerando el caso de un sistema elástico lineal de varios grados de libertad, su equilibrio dinámico queda representado por un sistema de ecuaciones diferenciales que tienen la siguiente forma general:

$$M \ddot{y} + C \dot{y} + K y = u \quad (2)$$

donde M , C y K representan las matrices de inercia, amortiguamiento y rigidez respectivamente. Para resolver numéricamente este problema se recurre a fórmulas de diferencias finitas de 4º orden que expresen la velocidad y aceleración en función de los desplazamientos en cuatro intervalos anteriores:

$$\begin{aligned} \dot{y}_{t+\Delta t} &= (11y_{t+\Delta t} - 18y_t + 9y_{t-\Delta t} - 2y_{t-2\Delta t}) / (6\Delta t) \\ \ddot{y}_{t+\Delta t} &= (2y_{t+\Delta t} - 5y_t + 4y_{t-\Delta t} - y_{t-2\Delta t}) / (\Delta t^2) \end{aligned} \quad (3)$$

Luego se considera la ecuación de equilibrio dinámico (2) en el instante $t+\Delta t$, se reemplazan la aceleración y velocidad por las expresiones (3), se agrupan finalmente los términos y se llega a una ecuación que es atribuida a Houbolt (1950) que tiene la siguiente forma general:

$$y_{t+\Delta t} = A y_t + B y_{t-\Delta t} + D y_{t-2\Delta t} + E u_{t+\Delta t} \quad (4)$$

donde las matrices A , B , D y E quedan expresadas como una combinación lineal de las matrices M , C y K , tal como lo muestran en las ecuaciones siguientes:

$$\begin{aligned} A &= H^{-1} [(5/\Delta t^2)M + (3/\Delta t)C] \\ B &= -H^{-1} [(4/\Delta t^2)M + (3/2\Delta t)C] \\ D &= H^{-1} [(1/\Delta t^2)M + (1/3\Delta t)C] \\ E &= H^{-1} \\ H &= (2/\Delta t^2)M + (11/6\Delta t)C + K \end{aligned} \quad (5)$$

Al adoptarse este planteo se persiguieron dos objetivos: el primero fue acotar la cantidad de términos de la expresión general (1) a partir de un criterio racional, eliminando la mencionada incertidumbre respecto de los valores más convenientes para “n” y “m”, que en

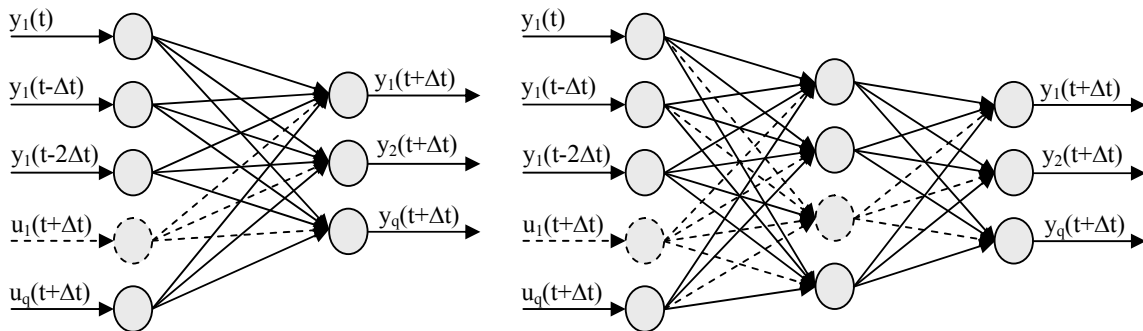
este caso son tres y uno respectivamente. El segundo objetivo fue disponer de información sobre el comportamiento que podía esperarse de un modelo de estas características. Para esto último se recurrió a algunos de los numerosos trabajos que evaluaron la propuesta de Houbolt, varios de los cuales se encuentran resumidos en un estudio de Semler et al. (1996), de donde se pudieron extraer valiosos antecedentes, tales como los siguientes:

- Este planteo demostró ser uno de los más eficientes para integrar ecuaciones correspondientes a problemas de valor inicial de sistemas elastoplásticos dinámicos.
- El proceso numérico es incondicionalmente estable en la resolución de sistemas lineales.
- Introduce un leve amortiguamiento y distorsión en las frecuencias, tal como ocurre con la mayoría de los métodos numéricos que integran ecuaciones de movimiento.

Estas consideraciones permitieron anticipar que el modelo propuesto en (4) era apto para representar la solución del sistema de ecuaciones diferenciales mostradas en (2), a condición de que los elementos de las matrices A , B , D y E pudiesen ser determinados.

Obsérvese que, si se tratara de en un planteo directo, las matrices M , C y K serían todas conocidas y la solución del problema quedaría expresada por las ecuaciones (5).

Por el contrario, este es un planteo inverso en el que M , C y K no son conocidas y por esta razón se recurre a las redes neuronales artificiales para desarrollar un modelo de caja negra que represente la expresión (4). Por tratarse de un problema lineal, la opción más conveniente es un conjunto de unidades *Adaline* (Adaptive Linear Element), o más precisamente tantas unidades *Adaline* como grados de libertad tenga el sistema representado. Es decir que, para un sistema de “q” grados de libertad dinámicos habrá “4q” entradas que incidirán sobre “q” unidades *Adaline*, como muestra la Figura 2a. Las alternativas de una red *Madaline* (Multiple *Adaline*) o una red Multicapa de Perceptrones (mostrada en la Figura 2b) pueden ofrecer algunas ventajas, que serán discutidas más adelante, pero impiden una relación directa con la expresión (4).



Figuras 2a y 2b: Esquemas de un conjunto de unidades *Adaline* (izquierda) y una red de Perceptrones (derecha)

Para el entrenamiento de estas redes se emplea un proceso de ajustes sucesivos de los pesos sinápticos denominado genéricamente “del gradiente descendente”, para el cual es necesario disponer de registros de las respuestas del sistema a ciertas condiciones de excitación. En el caso de las redes multicapa los errores de la salida son proyectados hacia atrás con el ya mencionado método de “backpropagation”. Estos procesos buscan encontrar el mínimo de la función error ε para determinar los pesos sinápticos w_i de las redes, que en el caso de las unidades *Adaline* son los elementos de las matrices A , B , D y E . Para ello se define ε a partir de la ecuación (4):

$$\varepsilon = y'_{t+\Delta t} - (Ay'_t + By'_{t-\Delta t} + Dy'_{t-2\Delta t} + Eu'_{t+\Delta t}) \quad (6)$$

donde los conjuntos de pares de entrenamiento $\{u', y'\}$ están formados por los vectores de valores conocidos de la excitación y los vectores de respuesta del sistema. Luego se define el escalar J que representa el error global de la red:

$$J = \varepsilon^T \varepsilon, \quad \frac{\partial J}{\partial w} = 0 \quad (7)$$

cuyo mínimo es buscado en el hiperespacio w de los pesos sinápticos de la red neuronal a través del mencionado método del gradiente descendente.

3 IDENTIFICACIÓN DE PARÁMETROS

3.1 Falta de una solución única

Para el caso de utilizarse unidades Adaline parecería que el problema ha quedado resuelto. Una vez conocidas las matrices A , B , D y E no debería haber mayores inconvenientes para deducir los parámetros representados por las matrices M , C y K , ya que ambos grupos de matrices están vinculados por las expresiones (5). Sin embargo, se presenta aquí una dificultad inesperada. Ocurre que A , B , D y E no representan una solución única para el problema formulado en (6) y (7). En otras palabras, hay numerosas matrices A , B , D y E diferentes que son capaces de representar con muy buena aproximación la función desconocida a través de la ecuación (4), pero cuyos valores no conducirán a valores correctos para las matrices M , C y K . Esto se debe a que el funcional J que se pretende minimizar presenta múltiples mínimos locales, que tienen su origen en la inadecuada excitación empleada en la determinación de la respuesta del sistema.

Con el fin de tipificar los requerimientos que deben cumplir las señales de excitación a fin de que un sistema sea identificable, [Aström y Eykhof \(1971\)](#) introdujeron el concepto de excitación persistente. Esta condición establece que las señales de excitación deben tener un espectro de frecuencias suficientemente amplio, de manera de perturbar al sistema en forma apropiada, y es reconocida como un requerimiento indispensable para que sea posible alcanzarse el mínimo absoluto, al minimizarse J . Sin embargo, no es fácil asegurar tal condición y mucho menos si se trata de la excitación de estructuras reales, por lo que la viabilidad práctica de identificar los parámetros por este camino debió ser descartada.

3.2 Determinación de valores característicos

A raíz del inconveniente antes mencionado, fue necesario buscar otras alternativas para la identificación de los parámetros del sistema y con este fin se desarrolló el procedimiento que a continuación se describe. Utilizando el modelo neuronal regresivo obtenido en la etapa anterior, se lo somete a condiciones iniciales cuidadosamente elegidas a efectos de obtener una respuesta “ y ” en la que estén presentes sus principales modos naturales de vibración. Una vez conocido el vector respuesta “ y ”, se emplean las expresiones (3) para calcular los vectores de velocidad \dot{y} y de aceleración \ddot{y} , todos ellos de dimensión “ q ”. Estos desplazamientos y velocidades son agrupados en un vector de estado de dimensión $2q$ denominado z :

$$z = \begin{Bmatrix} \dot{y} \\ y \end{Bmatrix} \quad (8)$$

Luego, se reformula la ecuación (2) en función de z :

$$\dot{z} = F z + G u \quad (9)$$

donde

$$F = \begin{bmatrix} -M^{-1}C & -M^{-1}K \\ I & 0 \end{bmatrix}, \quad G = \begin{bmatrix} -M^{-1} \\ 0 \end{bmatrix} \quad (10)$$

Los “p” vectores columna z y u , correspondientes a otros tantos instantes “p” de la respuesta del sistema, pueden agruparse en las matrices Z y U de forma tal que:

$$Z = [z_1 z_2 z_3 \dots z_p], \quad U = [u_1 u_2 u_3 \dots u_p] \quad (11)$$

A partir de (9) y (11) se obtiene la siguiente expresión que considera el comportamiento del sistema en todos los instantes “p”:

$$\dot{Z} = F Z + G U \quad (12)$$

Reordenando y postmultiplicando ambos miembros por la pseudoinversa de la matriz Z se obtiene una expresión que permite determinar F :

$$F = (\dot{Z} - G U) Z^+ \quad (13)$$

donde

$$Z^+ = (Z^T Z)^{-1} Z^T \quad (14)$$

Considerando ahora que la excitación U es nula, es decir que el sistema solo responde a ciertas condiciones iniciales en ausencia de cargas exteriores, se tiene:

$$F = \dot{Z} Z^+ \quad (15)$$

y teniendo en cuenta la definición de F de la ecuación (10) se deduce que:

$$F_{11} = -M^{-1}C, \quad F_{12} = -M^{-1}K \quad (16)$$

Aquí puede comprobarse que para implementar el cálculo de la matriz F no es conveniente usar la expresión (15), sino hacer un análisis similar a partir de la transpuesta de la expresión (12), que conduce a una expresión equivalente:

$$F^T = (Z^T)^+ \dot{Z}^T \quad (17)$$

Obsérvese que, si bien las ecuaciones (15) y (17) son algebraicamente equivalentes, esta última es ventajosa desde el punto de vista numérico. En efecto, el cálculo de la pseudoinversa de la primera ecuación involucra la inversión de una matriz de orden “p”, mientras que la determinación de la pseudoinversa de la ecuación (17) involucra la inversión de una matriz de orden $2q$, y siempre $q \ll p$, ya que la cantidad de grados de libertad “q” del modelo es siempre muy inferior a la cantidad de intervalos de tiempo “p” en el que se considera la respuesta del sistema.

Se llega así a la determinación de la submatriz F_{12} , que abre las puertas al cálculo de las frecuencias y modos normales de vibración del sistema. En efecto, volviendo a la ecuación (2), omitiendo las fuerzas de amortiguamiento y suponiendo una respuesta armónica, se plantea el clásico problema de autovalores que queda expresado como:

$$(M^{-1}K - \lambda I) \bar{y} = 0, \quad \text{donde: } \lambda = \omega^2 \quad (18)$$

lo que equivale a decir que

$$(-F_{12} - \omega^2 I) \bar{y} = 0 \quad (19)$$

en donde ω representa las frecuencias e \bar{y} los modos normales de vibración.

Para los casos en que se conozca la matriz de inercia M , que habitualmente es diagonal, las matrices de rigidez K y amortiguamiento C pueden determinarse a partir de las ecuaciones (16). En efecto:

$$\begin{aligned} K &= -M F_{12} \\ C &= -M F_{11} \end{aligned} \quad (20)$$

Más aún, cuando sea posible asumir que el amortiguamiento es proporcional a la inercia y rigidez del sistema (amortiguamiento de Rayleigh), se define una matriz Φ que contiene los modos de vibración \bar{y} como sus columnas y se determina la matriz Γ de amortiguamiento generalizado del sistema. Así se pueden conocer los factores de amortiguamiento ξ_i correspondientes a cada modo de vibración:

$$\Gamma = \Phi^T C \Phi = -\Phi M A_{11} \Phi \quad (21)$$

donde

$$\Gamma = \begin{bmatrix} 2\xi_1\omega_1 & \dots & 0 \\ \dots & \ddots & \dots \\ 0 & \dots & 2\xi_n\omega_n \end{bmatrix} \quad (22)$$

3.3 Procedimiento

Una vez formulada la solución del problema se resume el procedimiento propuesto:

- a) Desarrollar un modelo neuronal capaz de reproducir el comportamiento del sistema estudiado en un cierto intervalo de tiempo. Este modelo estará formado por tantas unidades Adaline como grados de libertad “q” tenga el sistema, con un total de 4q entradas y “q” salidas. Las entradas correspondientes a cada grado de libertad son los desplazamientos en tres intervalos de tiempo sucesivos y la acción externa en el último intervalo, tal como indica la ecuación (4). La salida de cada unidad será el desplazamiento en el intervalo de tiempo siguiente.
- b) Usar el modelo anterior para determinar la respuesta “y” del sistema a ciertas condiciones iniciales, en ausencia de cargas exteriores, que aseguren que en esta respuesta habrá componentes de todas las frecuencias de interés.
- c) Obtener mediante las ecuaciones (3) los vectores velocidad y aceleración en los mismos intervalos en que se determinaron los desplazamientos.
- d) Seleccionar desplazamientos, velocidades y aceleraciones correspondientes a cierto número de puntos “p”, representativos de la respuesta del sistema.
- e) Con estos vectores armar las matrices Z y \dot{Z} . Luego calcular la matriz F según (17).
- f) Formular el problema de autovalores (19) y calcular frecuencias y modos normales de vibración.
- g) En el caso que pueda conocerse la matriz de inercia M , calcular K y C con las ecuaciones (20).
- h) Determinar los factores de amortiguamiento con las ecuaciones (21) y (22) en los casos en que el modelo de amortiguamiento de Rayleigh sea válido.

4 RESULTADOS

4.1 Modelos Neuronales

Para el ajuste de los modelos neuronales se implementó una versión básica del método de Backpropagation, en la que se evitó el uso de algoritmos más sofisticados para tener la mayor sensibilidad a las diferentes condiciones de las señales de entrada. La única mejora fue permitir el ajuste del factor de aprendizaje η durante el proceso de entrenamiento (Plagianakos et al., 2002), mientras que se mantuvo constante el factor de momentum θ .

Los modelos regresivos reprodujeron el comportamiento de los sistemas estudiados a través de unidades Adaline y redes de Perceptrones con una capa oculta. Si bien en todos los casos estas redes demostraron un excelente desempeño, los resultados obtenidos confirmaron la conveniencia de recurrir a los modelos neuronales más simples, ya que estas permiten establecer un paralelismo con los modelos numéricos que representan el problema estudiado. Esto ya fue motivo de un trabajo anterior de los mismos autores (Giró y otros, 2006).

Se utilizaron modelos de sistemas mecánicos lineales de entre uno y tres grados de libertad, con una única señal de excitación, y períodos de las frecuencias naturales superiores a 0,2 segundos. Los conjuntos de entrenamiento correspondieron a segmentos de respuesta de entre 10 y 20 segundos, con intervalos entre $\Delta t=0,1$ y $\Delta t=0,01$ segundos, los que dieron lugar a una cantidad de puntos conocidos de la excitación y respuesta del sistema que varió entre 100 y 2000.

En todos los casos se utilizó una condición de carga para entrenar el modelo y una condición de carga diferente para su validación. Así pudo comprobarse, como era de esperar, que los modelos neuronales no se limitan a memorizar patrones de conducta del sistema real, sino que verdaderamente capturan sus propiedades, por lo que una vez entrenados reproducen correctamente su respuesta a cualquier condición de excitación y condiciones iniciales. Para ello, es necesario asegurar que la excitación del sistema provoque la presencia de todos los modos de vibración en su respuesta.

4.2 Identificación de parámetros

Para validar el procedimiento presentado se utilizaron los mismos sistemas mecánicos con los que se desarrollaron los modelos neuronales, de uno a tres grados de libertad. La diferencia fue que, en este caso, las respuestas correspondieron a condiciones iniciales de desplazamiento y velocidad, sin cargas exteriores, a fin de hacer aplicable la ecuación (17). Como ya fue comentado, estas condiciones iniciales debieron ser cuidadosamente definidas de manera de asegurar la presencia de todos los modos de vibración en las respuestas de los sistemas. Teniendo en cuenta esta importante condición, en todos los casos pudo comprobarse que el método propuesto permitió obtener muy buenos resultados.

Al estudiarse la calidad de los resultados obtenidos se consideró conocida la matriz de inercia y se determinaron las matrices de rigidez y amortiguamiento con las expresiones (20). Luego, se tomaron como valores de comparación los errores medios cuadráticos de los elementos no nulos de estas matrices y los mayores errores individuales en cada una. Los resultados que se reproducen a continuación corresponden a un sistema de tres grados de libertad, con períodos naturales de vibración $T = 1.708, 0.654$ y 0.293 seg. respectivamente.

En uno de los estudios realizados se evaluó la sensibilidad de los resultados a los diferentes intervalos Δt que son utilizados en el entrenamiento del modelo neuronal y en el cálculo

posterior de los vectores de velocidad y aceleración, a partir de las derivadas de su respuesta. En la [Figura 3](#) se muestra la evolución de los errores en la matriz de rigidez con intervalos de tiempo que dan lugar a valores de $\Delta t/T$ entre 0,035 y 0,35 (línea azul), tomando $T = 0.293$ seg. Aquí pudo comprobarse que con un valor $\Delta t/T$ del orden de 0,1 (valor recomendado por la literatura para procesos de integración numérica de ecuaciones diferenciales) los errores alcanzan el 2%, y que para $\Delta t/T > 0,1$ crecen en forma sostenida.

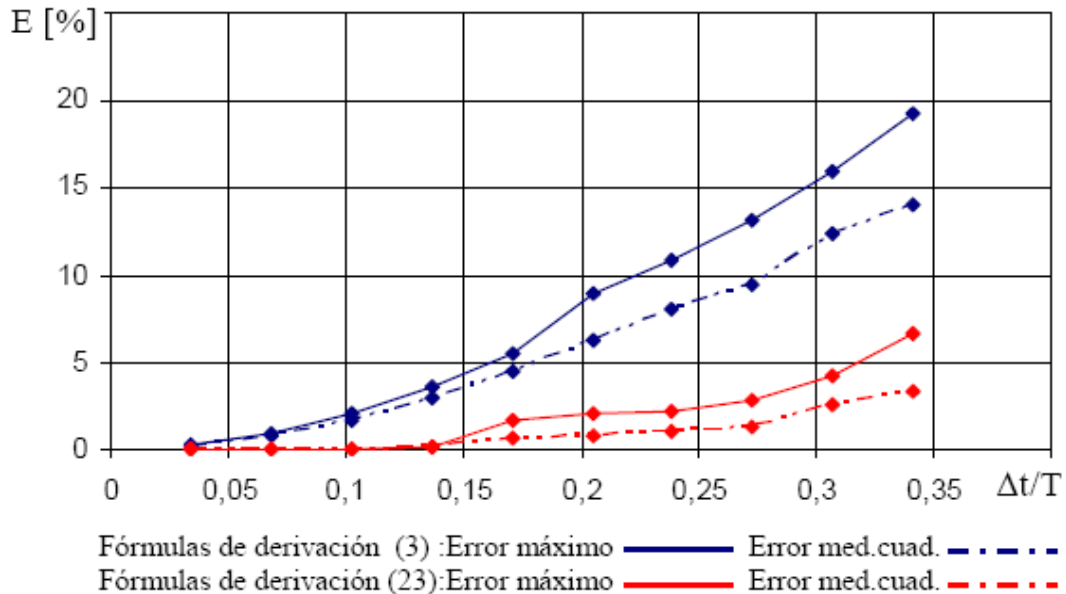


Figura 3: Evolución del error de la matriz de rigidez al incrementarse $\Delta t/T$

Para procurar mejorar estos resultados se exploraron varias opciones, las que condujeron a comprobar la necesidad de incrementar el orden de las fórmulas de derivación numérica para la obtención de los vectores velocidad y aceleración. Así se adoptaron las ecuaciones (23) propuestas por [McCormick y Salvadori \(1976\)](#) que produjeron una drástica mejora en los resultados obtenidos (línea roja). En efecto, puede observarse en la [Figura 3](#) que con valores de $\Delta t/T$ tan elevados como 0,2, el máximo error en la matriz de rigidez es del orden de 2% y su error medio cuadrático está por debajo de 0,8 %. Se reconoce así la sensibilidad del método a la calidad de los vectores velocidad y aceleración, y las nuevas fórmulas de derivación numérica son:

$$\begin{aligned} \dot{y}_{t+\Delta t} &= (25y_{t+\Delta t} - 48y_t + 36y_{t-\Delta t} - 16y_{t-2\Delta t} + 3y_{t-3\Delta t}) / (12\Delta t) \\ \ddot{y}_{t+\Delta t} &= (35y_{t+\Delta t} - 104y_t + 114y_{t-\Delta t} - 56y_{t-2\Delta t} + 11y_{t-3\Delta t}) / (12\Delta t^2) \end{aligned} \quad (23)$$

Aquí es necesario destacar que la cantidad de puntos “p” de la respuesta de un sistema que se usan en esta etapa de identificación de parámetros, no guarda ninguna relación con los puntos utilizados en la etapa anterior, de entrenamiento del modelo. Al respecto se hicieron estudios utilizando 500, 200, 100 y 50 puntos, siempre igualmente espaciados, obteniéndose idénticos resultados. Al aumentar la complejidad del sistema y/o disminuir la calidad de los datos empleados al entrenar el modelo neuronal, es probable que sea beneficiosa una mayor cantidad de puntos “p”, pero esto todavía no ha sido estudiado suficientemente como para poder hacer recomendaciones al respecto.

A continuación, con las mismas matrices de rigidez utilizadas para estudiar la evolución de

los errores, se calcularon las frecuencias naturales de vibración del sistema, planteando el problema de autovalores representado por la ecuación (19). Aquí pudo comprobarse que, a pesar que los errores de la matriz de rigidez muestran una importante diferencia según se empleen las fórmulas de derivación numérica (3) o (23), su incidencia en la determinación de las frecuencias se reduce notablemente. Tal como puede observarse en la [Figura 4](#), los valores de los períodos naturales de vibración se mantuvieron prácticamente estacionarios para valores de $\Delta t/T$ entre 0,05 y 0,35. Los valores reales son los que corresponden a $\Delta t/T=0$.

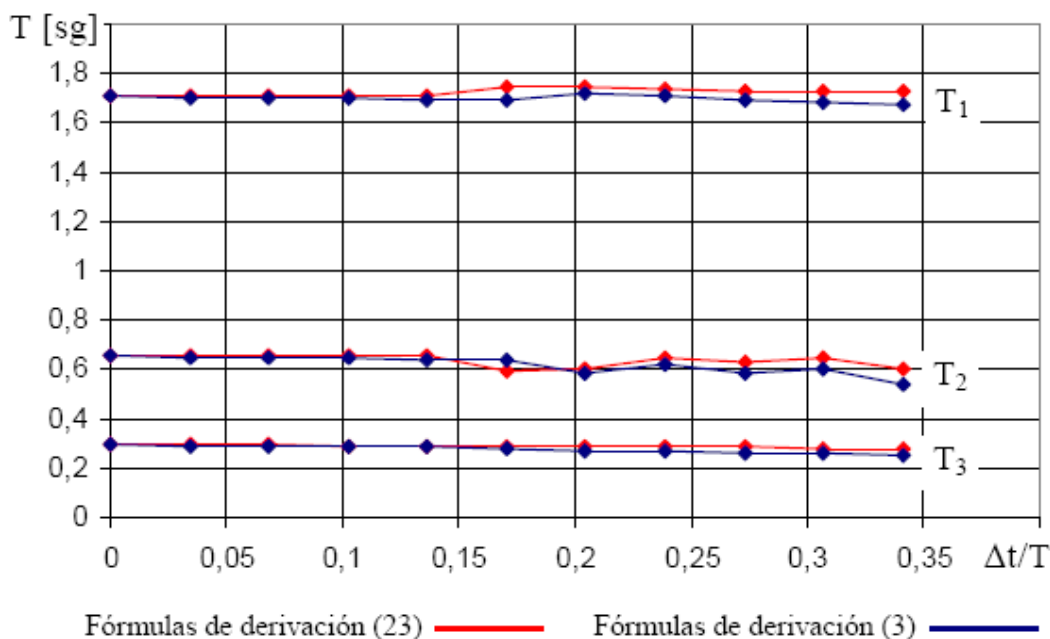


Figura 4: Evolución de los períodos naturales de vibración al incrementarse $\Delta t/T$

Estos resultados indicarían que los elementos de la matriz de rigidez sufren distorsiones a medida que se incrementa el intervalo de tiempo entre los datos, pero lo hacen de forma tal que se preservan las características dinámicas del modelo.

También se determinaron las matrices de amortiguamiento del sistema en las mismas condiciones en que se determinaron las matrices de rigidez. Los resultados obtenidos se presentan en la [Figura 5](#), y como puede observarse, la calidad de los resultados es muy buena para valores bajos de $\Delta t/T$, pero el error crece rápidamente cuando el intervalo de tiempo se incrementa. También pudo comprobarse que la calidad de estos resultados es muy sensible a las condiciones iniciales impuestas al sistema, mucho más que lo que demostró ser la matriz de rigidez. En este estudio se utilizaron las fórmulas de diferenciación numérica de mayor orden.

En razón de que éste procedimiento será utilizado para identificar parámetros de sistemas reales, a partir de mediciones sobre su respuesta, se consideró muy importante conocer su sensibilidad a eventuales errores en las señales de entrada. Para ello se usaron respuestas obtenidas numéricamente, a las que se les sumó un ruido aleatorio, con distribución Gaussiana, media nula y desviación estándar especificada. Así, usando diferentes amplitudes de ruido en los datos de entrenamiento de la red, se estudió la sensibilidad del procedimiento para determinar las matrices de rigidez y amortiguamiento.

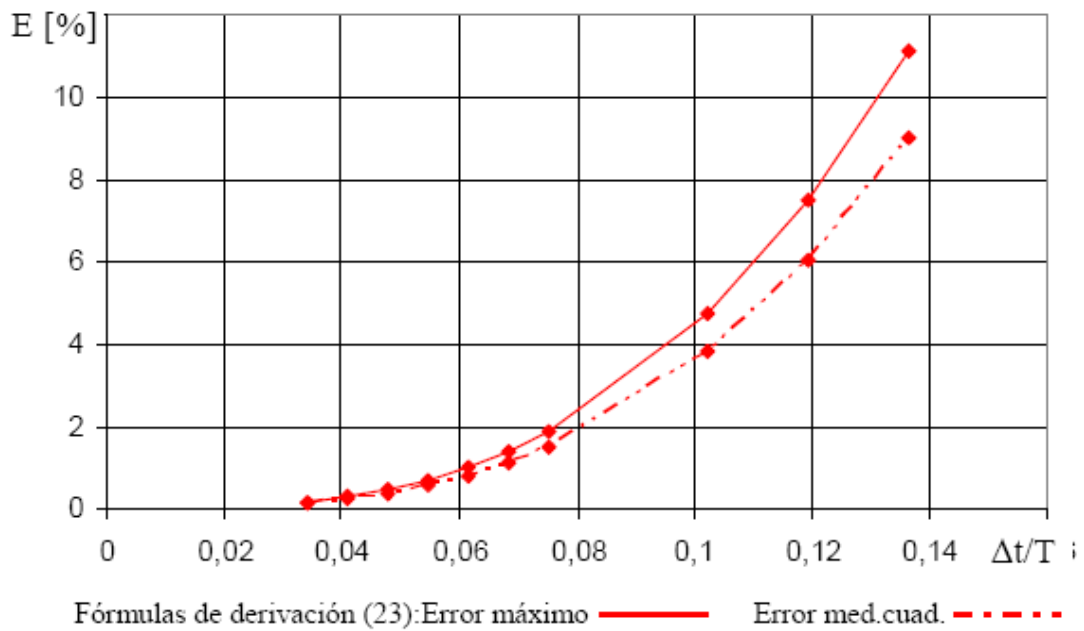


Figura 5: Evolución del error de la matriz de amortiguamiento al incrementarse $\Delta t/T$

En la [Figura 6](#) se representa la evolución de los errores en la matriz de rigidez, en función de la amplitud del error en los datos de entrenamiento de la red, que es expresado como un porcentaje de la máxima amplitud de la oscilación.

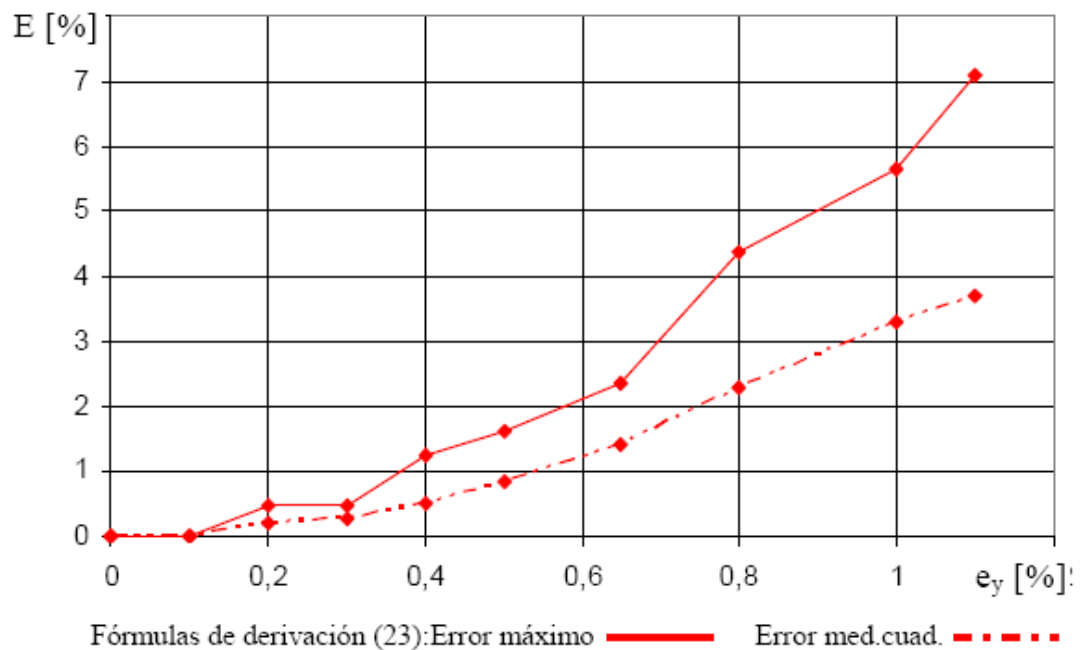


Figura 6: Evolución del error de la matriz de rigidez al incrementarse el error en la entrada

Un aspecto que merece destacarse es que el procedimiento presentado para la segunda etapa es formulado en su totalidad a través del álgebra matricial, lo que lo independiza de la complejidad o dimensión del sistema estudiado, facilitándose su implementación sistemática.

Esto lo distingue de otros procedimientos que han sido presentados con la misma finalidad, como es el caso de los trabajos de Xu et al (2004), de Javanović (1997) y de Trivailo y otros (2004), donde no parece que las propuestas sean fácilmente extensibles al tratamiento de sistemas más complejos que los de los ejemplos allí mostrados.

5 CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se presentó un procedimiento que permite obtener los parámetros (matrices de rigidez y amortiguamiento) y características dinámicas (frecuencias y modos de vibración) de sistemas mecánicos lineales. El método reconoce dos etapas; en la primera se ajusta un modelo mediante redes neuronales artificiales y en la segunda se utiliza este modelo neuronal para la identificación de los parámetros del sistema real. En esta última etapa el procedimiento queda completamente formulado a través del álgebra matricial, lo que lo hace apropiado para abordar problemas complejos de numerosos grados de libertad.

Los resultados obtenidos con sistemas simples permitieron confirmar las ventajas del procedimiento presentado, y también su robustez tanto ante intervalos de tiempo relativamente grandes como ante señales de entrada ruidosas. Se espera que este método sea de utilidad en la predicción de daños estructurales, y en la linealización e identificación de parámetros de mecanismos complejos en el área de la robótica.

Como trabajo futuro se continuará estudiando la sensibilidad de los parámetros calculados a diferentes condiciones iniciales y a la calidad de las señales de entrada. Progresivamente, se tratarán problemas de mayor dimensión y complejidad. También se estudiarán variantes en los modelos neuronales a los efectos de disponer de mayor capacidad de filtrado de ruidos y otras perturbaciones.

REFERENCIAS

- Aguado, A. y Martínez, M., Identificación y Control Adaptativo. *Prentice Hall*, 2003.
- Aström, K. y Eykhof, P., System Identification: a survey. *Second IFAC Symposium on Identification and Process Parameter Identification, Praga*, 1971.
- Chandra, P. y Singh, Y., Fault tolerance of feedforward artificial neural networks- a framework of study. *Proceedings of the International Joint Conference on Neural Networks, School of Information Technology, Indraprastha University, Delhi, India*.1:489-494, 2003.
- Giró, J., Olariaga, S., Paez, N., García, A. y Vuirli, A., Definición y evaluación de Modelos Neuronales para Sistemas de Control. *Jornadas Argentinas de Robótica (JAR). Universidad Tecnológica Nacional. Córdoba, Argentina*, 2006.
- Haecker, J. y Rudolph, S., On neural network topology design for nonlinear control. *Proceedings SPIE Aerosense 2001 Conference On Applications and Science of Computational Intelligence IV, Orlando, Florida*, 2001.
- Houbolt, J., A recurrent matrix solution for the dynamic response of elastic aircraft. *Journal of Aeronautical Science*, 17:540-550, 1950.
- Javanović, O., Identification of Dynamic System using Neural Network. *The Scientific Journal, Series Architecture and Civil Engineering. University of Nis, Yugoslavia*. 1-4:525-532, 1997.
- McCormick, J. y Salvadori, M., Numerical Methods in Fortran, *Prentice-Hall*, 1964.

- Plagianakos, Magoulas y Vrahatis. Deterministic nonmonotone strategies for effective training of multilayer perceptrons. *IEEE Transactions on Neural Networks*, 13(6):1268-1284, 2002.
- Rumelhart, D., Hinton, G. y Williams, R., Learning internal representations by error propagation. *Parallel distributed processing: Explorations in the microstructures of cognition*, MIT Press, 1:318-362, 1986.
- Semler, C., Gentleman, C. y Paidoussis, M., Numerical solutions of second order implicit non-linear ordinary differential equations. *Journal of Sound and Vibration*, 195(4):553-574, 1996.
- Trivailo, P., Dulikravich, G., Sgarioto, D. y Gilbert, T., Inverse problem of Aircraft Parameter Estimation: application of Neural Networks. *Inverse Problems, Design and Optimization Symposium. Rio de Janeiro, Brazil*, 2004.
- Valishevsky, A., Comparative analysis of different approaches towards multilayer perceptron training , Department of Computer Science, University of Latvia, Riga, 1998.
- Xu, B., Wu, Z., Chen G. y Yokohama K., Direct identification of structural parameters from dynamic responses with Neural Networks. *Engineering Applications of Artificial Intelligence*, Elsevier, 931-943, 2004.