



Revista Integración

ISSN: 0120-419X

integracion@matematicas.uis.edu.co

Universidad Industrial de Santander

Colombia

Arenas, Favián; Pérez, Rosana; Vivas, Hevert
Un modelo de redes neuronales para complementariedad no lineal
Revista Integración, vol. 34, núm. 2, 2016, pp. 169-185
Universidad Industrial de Santander
Bucaramanga, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=327048838005>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

redalyc.org

Sistema de Información Científica

Red de Revistas Científicas de América Latina, el Caribe, España y Portugal
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

Un modelo de redes neuronales para complementariedad no lineal

FAVIÁN ARENAS*, ROSANA PÉREZ, HEVERT VIVAS

Universidad del Cauca, Departamento de Matemáticas, Popayán, Colombia.

Resumen. En este artículo presentamos un modelo de red neuronal para resolver el problema de complementariedad no lineal. Para ello, reformulamos este problema como uno de minimización sin restricciones usando una familia uniparamétrica de funciones de complementariedad. Demostramos resultados de existencia y convergencia de la trayectoria de la red neuronal, así como resultados de estabilidad en el sentido de *Lyapunov*, estabilidad asintótica y exponencial. Además, presentamos resultados numéricos preliminares que ilustran un buen desempeño práctico del modelo.

Palabras clave: Red neuronal, problema de complementariedad no lineal, estabilidad, reformulación.

MSC2010: 90C30, 90C33, 90C53, 90B10.

A neural network model for nonlinear complementarity problems

Abstract. In this paper we present a neural network model for solving the nonlinear complementarity problem. This model is derived from an equivalent unconstrained minimization reformulation of the complementarity problem, which is based on a one-parametric class of nonlinear complementarity functions. We establish the existence and convergence of the trajectory of the neural network, and we study its Lyapunov stability, asymptotic stability as well as exponential stability. Numerical tests verify the obtained theoretical results.

Keywords: Neural network, nonlinear complementarity problem, stability, reformulation.

*E-mail: farenas@unicauca.edu.co

Recibido: 22 de junio de 2016, Aceptado: 21 de noviembre de 2016.

Para citar este artículo: F. Arenas, R. Pérez, H. Vivas, Un modelo de redes neuronales para complementariedad no lineal, *Rev. Integr. Temas Mat.* 34 (2016), No. 2, 169–185.

1. Introducción

Dada $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ continuamente diferenciable, el Problema de Complementariedad No Lineal (PCNL) consiste en encontrar un vector $\mathbf{x} \in \mathbb{R}^n$ que satisfaga las tres condiciones siguientes,

$$\mathbf{x} \geq 0, \quad F(\mathbf{x}) \geq 0, \quad \mathbf{x}^T F(\mathbf{x}) = 0,$$

donde la expresión $\mathbf{y} \geq 0$, cuando $\mathbf{y} \in \mathbb{R}^n$, significa que $y_i \geq 0$ para todo $i = 1, \dots, n$.

Este problema surge naturalmente en ingeniería, física y economía, entre otros campos ([12], [16], [24]), por lo cual ha habido gran interés en el estudio y creación de métodos que permitan resolverlo. Entre dichos métodos están, por una parte, los de homotopía derivados de los de punto fijo ([11], [26]). Por otra parte, están los métodos (quizá, los más populares) de reformulación que plantean el PCNL como un problema de minimización sin restricciones usando una función de mérito ([9], [10], [15]). Este problema puede resolverse usando diferentes métodos de optimización, como los de tipo Newton ([9], [15]) o quasi-Newton ([1], [3]). En los de reformulación se utiliza una función $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}$ tal que $\varphi(a, b) = 0 \iff a \geq 0, b \geq 0, ab = 0$, conocida como *función de complementariedad* ([7], [13], [15], [22]).

No obstante, en algunas aplicaciones se necesitan soluciones del PCNL en tiempo real, para lo cual las técnicas estándares de optimización no son muy apropiadas por el alto costo computacional que demandan. En este caso, una técnica útil consiste en usar redes neuronales artificiales ([6], [14], [17]) como, por ejemplo, las llamadas redes de *Hopfield* para resolver el problema de optimización mencionado en el párrafo anterior. La idea principal de esta técnica consiste en construir una función no negativa, llamada función de energía, y un sistema dinámico que represente la red neuronal artificial, que normalmente se plantea en forma de una ecuación diferencial ordinaria de primer orden con una condición inicial. Se espera que la trayectoria solución del sistema se aproxime, a partir de un estado inicial, a un estado estático (o punto de equilibrio) que corresponde a la solución del problema de optimización subyacente ([14]).

En particular, en [6] y [17] se presentan dos propuestas que incorporan la técnica de solución al PCNL mediante redes neuronales artificiales. Para ello, se usan, por un lado, *funciones de complementariedad* para construir la función a minimizar, que es la función energía del modelo, y, por otro lado el método de máximo descenso ([10], [19]) para modelar la red neuronal (una método natural en un problema de minimización). En [17] se usa la función de complementariedad de *Fischer-Burmeinster* ([15]) definida por

$$\varphi(a, b) = \|(a, b)\|_2 - a - b,$$

y, en [6], se usa la función de complementariedad de *Fischer-Burmeinster generalizada* ([7]) definida por

$$\varphi_p(a, b) = \|(a, b)\|_p - a - b, \quad p > 1.$$

En ambos trabajos se analiza teórica y numéricamente el modelo propuesto.

Motivados por los resultados de los trabajos anteriores, y en especial por las propiedades de la *familia uniparamétrica de funciones de complementariedad* ([2], [15]) definida por

$$\varphi_\lambda(a, b) = \sqrt{(a - b)^2 + \lambda ab} - a - b, \quad \lambda \in (0, 4), \quad (1)$$

proponemos un modelo de red neuronal de Hopfield que usa, por primera vez, la familia uniparamétrica (1) para construir la función de energía del modelo. Tal como lo hacen los autores de [6] y [17], usamos el método de máximo descenso para modelar la red neuronal. Cabe mencionar que nuestro modelo es más general que el propuesto en [17], ya que la función que allí se usa es el miembro de la familia (1) correspondiente a $\lambda = 2$. Además, cuando λ tiende a cero, $\varphi_\lambda(a, b)$ tiende a un múltiplo de una de las funciones de complementariedad más utilizadas en el contexto de problemas de complementariedad no lineal mediante reformulación, la llamada *función mínimo* ([22]), definida por $\varphi(a, b) = \min\{a, b\}$. Algoritmos que usan esta función han reportado buenas propiedades de convergencia local ([20], [21], [22]), mientras que aquellos que usan la función de Fischer-Burmeister tienen buenas propiedades de convergencia global ([3], [15]). Aprovechamos estas ventajas incorporando un parámetro λ dinámico ([15]) en el algoritmo del modelo que proponemos, lo cual resulta más eficiente que usar un parámetro fijo, como se evidencia en los resultados numéricos que presentamos al final del artículo, donde comparamos la simulación numérica de nuestro modelo con la del modelo propuesto en [17].

Las funciones de complementariedad mencionadas son funciones no diferenciables. Por ello, para el análisis de estabilidad y convergencia de la trayectoria un modelo que use dichas funciones debe acudirse a teoría de funciones no diferenciables. En particular, la reformulación del PCNL basada en la familia (1) posee propiedades ([1], [15]) análogas a las de las reformulaciones utilizadas en [6] y [17], lo que fue de gran utilidad en el análisis de estabilidad y convergencia del modelo que proponemos en el presente artículo.

Organizamos la presentación de este artículo en la siguiente forma. En la Sección 2, reformulamos el problema de complementariedad como un problema de minimización usando la familia de funciones de complementariedad propuesta en [15], y proponemos un modelo de red neuronal artificial para resolverlo. En la Sección 3 presentamos, de manera preliminar, algunos conceptos necesarios para el desarrollo teórico posterior. En la Sección 4 realizamos el análisis de estabilidad y convergencia del modelo. En la Sección 5 analizamos el comportamiento numérico del modelo propuesto. Finalmente, en la Sección 6 hacemos algunos comentarios finales y propuestas de trabajos futuros sobre el tema.

2. Reformulación del PCNL como un problema de minimización

Usando la familia de funciones de complementariedad (1) podemos definir la función $\Phi_\lambda: \mathbb{R}^n \rightarrow \mathbb{R}^n$ y reformular el PCNL como un sistema de ecuaciones no lineales, el cual resulta no diferenciable debido a la no diferenciabilidad de φ_λ :

$$\Phi_\lambda(\mathbf{x}) = \begin{pmatrix} \varphi_\lambda(x_1, F_1(\mathbf{x})) \\ \vdots \\ \varphi_\lambda(x_n, F_n(\mathbf{x})) \end{pmatrix} = 0. \quad (2)$$

A partir de la definición de *función de complementariedad* tenemos que un vector \mathbf{x}_* es solución del PCNL si, y solo si, \mathbf{x}_* es solución del sistema de ecuaciones no lineales no diferenciable (2).

Si Φ_λ es continua según Lipschitz su jacobiano generalizado existe ([8], [15]) y está defi-

nido en un punto $\mathbf{x} \in \mathbb{R}^n$ como el conjunto,

$$\partial\Phi_\lambda(\mathbf{x}) = \text{conv} \left\{ \lim_{k \rightarrow \infty} \Phi'_\lambda(\mathbf{x}_k) \in \mathbb{R}^{n \times n} : \mathbf{x}_k \rightarrow \mathbf{x}, \mathbf{x}_k \in D_{\Phi_\lambda} \right\}, \quad (3)$$

donde D_{Φ_λ} es el conjunto de todos los puntos de \mathbb{R}^n en los que la función Φ_λ es diferenciable y $\text{conv}\{A\}$ representa la envolvente convexa del conjunto A .

Por otra parte, si definimos la *función de mérito* ([10]) $E_\lambda : \mathbb{R}^n \rightarrow \mathbb{R}^n$ como

$$E_\lambda(\mathbf{x}) = \frac{1}{2} \|\Phi_\lambda(\mathbf{x})\|^2, \quad (4)$$

tenemos que (2) es equivalente a $E_\lambda(\mathbf{x}_*) = 0$, lo que a su vez es equivalente a que \mathbf{x}_* es solución del PCNL. Por lo tanto, resolver el PCNL es equivalente a encontrar un minimizador global del problema de minimización sin restricciones:

$$\begin{aligned} &\text{Minimizar} \quad E_\lambda(\mathbf{x}), \\ &\mathbf{x} \in \mathbb{R}^n, \end{aligned} \quad (5)$$

el cual es un problema de optimización *diferenciable*. En efecto, en [15] se demuestra que la función E_λ es continuamente diferenciable y que, para cualquier $V \in \partial\Phi_\lambda(\mathbf{x})$,

$$\nabla E_\lambda(\mathbf{x}) = V^T \Phi_\lambda(\mathbf{x}). \quad (6)$$

Para resolver el problema de minimización (5) y, dado que E_λ es continuamente diferenciable, proponemos un modelo de red neuronal de la forma

$$\frac{d\mathbf{x}(t)}{dt} = -\rho \nabla E_\lambda(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0, \quad (7)$$

donde $\rho > 0$ es un factor de escala que por simplicidad en el análisis teórico del modelo lo tomaremos como 1. Este modelo es análogo al usado en [6] y [17], porque usa el método de máximo descenso para modelar la red neuronal; pero se diferencia en que su función de energía es la familia de funciones de complementariedad (1), lo que lo convierte en un modelo más general que el propuesto en [17].

3. Preliminares

En esta sección incluimos algunos conceptos que serán de utilidad para el desarrollo teórico posterior. Iniciamos con dos definiciones relacionadas con los conceptos de matriz y función P_0 ([6], [15], [17]).

Definición 3.1. Se dice que una matriz $M \in \mathbb{R}^{n \times n}$ es una matriz P_0 si todos sus menores principales son no negativos.

Definición 3.2. Una función $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función P_0 , si para todo $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ con $\mathbf{x} \neq \mathbf{y}$, se satisface que $\max(x_i - y_i)[F_i(\mathbf{x}) - F_i(\mathbf{y})] \geq 0$, $i = 1, \dots, n$.

En lo que sigue, presentamos conceptos relacionados con estabilidad de ecuaciones diferenciales ordinarias de primer orden (ver [27]).

Consideremos el siguiente sistema dinámico ([27]), donde $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ es una función continua

$$\frac{d\mathbf{x}(t)}{dt} = f(\mathbf{x}(t)), \quad \mathbf{x}(t_0) = \mathbf{x}_0. \quad (8)$$

Definición 3.3. Un vector $\mathbf{x}_* = \mathbf{x}(t_*) \in \mathbb{R}^n$ es llamado un punto de equilibrio o un estado estacionario de (8), si $f(\mathbf{x}_*) = 0$. Si existe una vecindad $\Omega^* \subseteq \mathbb{R}^n$ de \mathbf{x}_* tal que $f(\mathbf{x}_*) = 0$ y $f(\mathbf{x}) \neq 0$ para todo $\mathbf{x} \in \Omega^* \setminus \{\mathbf{x}_*\}$, entonces \mathbf{x}_* es llamado un punto de equilibrio aislado.

Si en (8), $f = \nabla E_\lambda$, obtenemos el modelo de red neuronal (7) propuesto. Así, de acuerdo con la Definición 3.3, un punto $\mathbf{x} \in \mathbb{R}^n$ es un punto de equilibrio de (7) si $\nabla E_\lambda(\mathbf{x}) = 0$. Es decir, un punto de equilibrio de (7) es un punto estacionario de ∇E_λ .

Teorema 3.4 (Existencia y unicidad). *Para cualesquiera $t_0 \geq 0$ y $\mathbf{x}_0 \in \mathbb{R}^n$ existe una solución local $\mathbf{x}(t)$, $t \in [t_0, \tau)$ de (8) para algún $\tau > t_0$. Además, si f es localmente continua según Lipschitz en \mathbf{x}_0 , entonces la solución es única; si f es continua según Lipschitz en \mathbb{R}^n , entonces τ puede extenderse a $+\infty$.*

Si una solución local definida en $[t_0, \tau)$ no se puede extender a una solución local en un intervalo más grande $[t_0, \tau_1)$, con $\tau_1 > \tau$, entonces es llamada una solución maximal, y el intervalo $[t_0, \tau)$ es el *intervalo maximal de existencia*. Una solución local arbitraria tiene una extensión a una maximal. El intervalo maximal de existencia asociado con \mathbf{x}_0 frecuentemente se denota por $[t_0, \tau(\mathbf{x}_0))$.

Teorema 3.5. *Si $\mathbf{x}(t)$ con $t \in [t_0, \tau(\mathbf{x}_0))$ es una solución maximal y $\tau(\mathbf{x}_0) < +\infty$, entonces*

$$\lim_{t \uparrow \tau(\mathbf{x}_0)} \|\mathbf{x}(t)\| = +\infty. \quad (9)$$

En otras palabras, si $\tau(\mathbf{x}_0) < +\infty$, la curva $\mathbf{x}(t)$ es inestable, como lo ilustra la Figura 1.

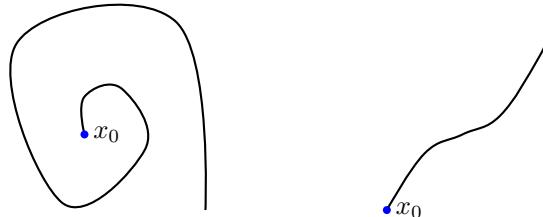


Figura 1. Situaciones que pueden ocurrir cuando $\lim_{t \uparrow \tau(\mathbf{x}_0)} \|\mathbf{x}(t)\| = +\infty$.

Definición 3.6 (Estabilidad en el sentido de Lyapunov). Sea $\mathbf{x}(t)$ una solución de (8). Un punto de equilibrio aislado \mathbf{x}_* es *estable según Lyapunov* si para cualquier $\mathbf{x}_0 = \mathbf{x}(t_0)$ y cualquier escalar $\epsilon > 0$, existe $\delta > 0$ tal que si $\|\mathbf{x}(t_0) - \mathbf{x}_*\| < \delta$, entonces $\|\mathbf{x}(t) - \mathbf{x}_*\| < \epsilon$, para $t \geq t_0$.

Definición 3.7 (Estabilidad asintótica). Un punto de equilibrio aislado \mathbf{x}_* es *estable asintóticamente* si, además de ser estable según Lyapunov, cumple que $\mathbf{x}(t) \rightarrow \mathbf{x}_*$ cuando $t \rightarrow \infty$ si $\|\mathbf{x}(t_0) - \mathbf{x}_*\| < \delta$.

Definición 3.8 (Función de Lyapunov). Sea $\Omega \subseteq \mathbb{R}^n$ una vecindad abierta de \mathbf{x}_* . Una función continuamente diferenciable $\xi: \mathbb{R}^n \rightarrow \mathbb{R}$ es una función de *Lyapunov* en el

estado $\bar{\mathbf{x}}$ (sobre el conjunto Ω) para (8), si satisface las dos condiciones siguientes:

$$\begin{aligned}\xi(\bar{\mathbf{x}}) &= 0, \quad \xi(\mathbf{x}) > 0, \quad \text{para todo } \mathbf{x} \in \Omega, \mathbf{x} \neq \bar{\mathbf{x}}; \\ \frac{d\xi(\mathbf{x}(t))}{dt} &= [\nabla_{\mathbf{x}(t)} \xi(\mathbf{x}(t))]^T f(\mathbf{x}(t)) \leq 0, \quad \text{para todo } \mathbf{x} \in \Omega.\end{aligned}\quad (10)$$

Una función de *Lyapunov* es llamada frecuentemente una *función de energía* para (8). El siguiente resultado relaciona estabilidad y función de *Lyapunov*.

Teorema 3.9. (i) *Un punto de equilibrio aislado \mathbf{x}_* es estable según Lyapunov si existe una función de Lyapunov sobre alguna vecindad abierta de \mathbf{x}_* .*

(ii) *Un punto de equilibrio aislado \mathbf{x}_* es estable asintóticamente si existe una función de Lyapunov sobre alguna vecindad Ω de \mathbf{x}_* que satisface*

$$\frac{d\xi(\mathbf{x}(t))}{dt} < 0 \quad \text{para todo } \mathbf{x}(t) \in \Omega, \mathbf{x}(t) \neq \mathbf{x}_*.$$

Una noción más fuerte que la estabilidad según *Lyapunov* es la llamada *estabilidad exponencial*.

Definición 3.10 (Estabilidad exponencial). Un punto de equilibrio aislado \mathbf{x}_* es *estable exponencialmente* para (8) si existen constantes $\omega < 0$, $\kappa > 0$ y $\delta > 0$ tales que una solución arbitraria $\mathbf{x}(t)$ de (8) con $\|\mathbf{x}(t_0) - \mathbf{x}_*\| < \delta$ está definida en $[0, \infty)$ y satisface $\|\mathbf{x}(t) - \mathbf{x}_*\| \leq \kappa e^{\omega t} \|\mathbf{x}(t_0) - \mathbf{x}_*\|$, $t \geq t_0$.

Observemos que todo punto exponencialmente estable es, a su vez, asintóticamente estable.

4. Análisis de estabilidad

Analizamos a continuación el comportamiento de la *trayectoria solución* del modelo de red neuronal propuesto (7). Estudiamos su existencia y unicidad. Además, para un punto de equilibrio aislado establecemos tres tipos de estabilidad. En primer lugar, analizamos la relación entre un punto de equilibrio de (7) y una solución del PCNL.

Proposición 4.1. (i) *Toda solución del PCNL es un punto de equilibrio de (7).*

(ii) *Si F es una función P_0 entonces todo punto de equilibrio de (7) es una solución del problema PCNL asociado a F .*

Demuestra (i) Si \mathbf{x} es una solución del PCNL, entonces $\Phi_\lambda(\mathbf{x}) = 0$, y por (6), $\nabla E_\lambda(\mathbf{x}) = 0$. Por lo tanto, \mathbf{x} es un punto de equilibrio de (7).

(ii) Si \mathbf{x} es un punto de equilibrio de (7), entonces $\nabla E_\lambda(\mathbf{x}) = 0$, y por tanto, \mathbf{x} es un punto estacionario de E_λ ; y dado que F es una función P_0 , se tiene que su matriz jacobiana $F'(\mathbf{x})$ es una matriz P_0 , lo cual garantiza que el punto estacionario \mathbf{x} es una solución de PCNL [15].



El siguiente es un lema técnico que usaremos en la prueba de existencia y unicidad de la trayectoria solución de (7).

Lema 4.2. *Si F es continua según Lipschitz en \mathbb{R}^n , entonces Φ_λ es continua según Lipschitz en \mathbb{R}^n .*

Demuestra. Usaremos las equivalencias entre las normas *euclíadiana* ($\|\cdot\|$) e *infinito* ($\|\cdot\|_\infty$) ([25]); la definición de norma *vectorial infinito* y la continuidad según Lipschitz de ϕ_λ (ver [1]). Sean \mathbf{x} y \mathbf{y} en \mathbb{R}^n . Para algún $1 \leq k \leq n$,

$$\begin{aligned}\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\| &\leq \sqrt{n} \|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\|_\infty = \sqrt{n} |\varphi_\lambda(x_k, F_k(\mathbf{x})) - \varphi_\lambda(y_k, F_k(\mathbf{y}))| \\ &\leq \sqrt{n} \mu \|(x_k - y_k, F_k(\mathbf{x}) - F_k(\mathbf{y}))\|_\infty \\ &\leq \sqrt{n} \mu \max \{|x_k - y_k|, |F_k(\mathbf{x}) - F_k(\mathbf{y})|\}.\end{aligned}$$

Si $\max \{|x_k - y_k|, |F_k(\mathbf{x}) - F_k(\mathbf{y})|\} = |x_k - y_k|$,

$$\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\| \leq \sqrt{n} \mu |x_k - y_k| \leq \sqrt{n} \mu \|\mathbf{x} - \mathbf{y}\|_\infty \leq \sqrt{n} \mu \|\mathbf{x} - \mathbf{y}\| = \gamma_1 \|\mathbf{x} - \mathbf{y}\|.$$

Si $\max \{|x_k - y_k|, |F_k(\mathbf{x}) - F_k(\mathbf{y})|\} = |F_k(\mathbf{x}) - F_k(\mathbf{y})|$,

$$\begin{aligned}\|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\| &\leq \sqrt{n} \mu |F_k(\mathbf{x}) - F_k(\mathbf{y})| \leq \sqrt{n} \mu \|F(\mathbf{x}) - F(\mathbf{y})\|_\infty \\ &\leq \sqrt{n} \mu \delta \|\mathbf{x} - \mathbf{y}\|_\infty \leq \sqrt{n} \mu \delta \|\mathbf{x} - \mathbf{y}\| = \gamma_2 \|\mathbf{x} - \mathbf{y}\|,\end{aligned}$$

donde δ es la constante de Lipschitz de la función F , con lo que se concluye la prueba. \checkmark

El siguiente resultado garantiza la existencia y unicidad de una trayectoria solución del modelo propuesto.

Teorema 4.3. (i) *Para un estado inicial arbitrario \mathbf{x}_0 existe una solución maximal $\mathbf{x}(t)$, con $t \in [t_0; \tau(\mathbf{x}_0))$, de (7).*

(ii) *Si F es continua según Lipschitz, entonces la solución maximal es única y $\tau(\mathbf{x}_0) = +\infty$.*

(iii) *Si el conjunto de nivel asociado a \mathbf{x}_0 , $\mathcal{L}(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^n : E_\lambda(\mathbf{x}) \leq E_\lambda(\mathbf{x}_0)\}$ es acotado, entonces $\tau(\mathbf{x}_0) = +\infty$.*

Demuestra. (i) Sea $\mathbf{x}_0 \in \mathbb{R}^n$ un estado inicial arbitrario. Dado que E_λ es continuamente diferenciable, se tiene que ∇E_λ es una función continua. Luego, por el Teorema 3.4, existe una solución maximal de (7).

(ii) Supongamos que F es continua según Lipschitz en \mathbb{R}^n . Entonces por Lema 4.2 tenemos que Φ_λ es continua según Lipschitz en \mathbb{R}^n . Usando este resultado y (6); tenemos que para \mathbf{x} y \mathbf{y} en \mathbb{R}^n y $\|\cdot\|$ una norma inducida sobre \mathbb{R}^n ,

$$\begin{aligned}\|\nabla E_\lambda(\mathbf{x}) - \nabla E_\lambda(\mathbf{y})\| &= \|V^T \Phi_\lambda(\mathbf{x}) - V^T \Phi_\lambda(\mathbf{y})\| \leq \|V\| \|\Phi_\lambda(\mathbf{x}) - \Phi_\lambda(\mathbf{y})\| \\ &\leq \gamma \|V\| \|\mathbf{x} - \mathbf{y}\| = \eta \|\mathbf{x} - \mathbf{y}\|,\end{aligned}$$

donde γ es la constante de Lipschitz para Φ_λ ($\gamma = \gamma_1$ ó $\gamma = \gamma_2$), y $\eta = \gamma \|V\|$. Por lo tanto, E_λ también es continua según Lipschitz en \mathbb{R}^n . Por Teorema 3.4 concluimos que la solución maximal es única y $\tau(\mathbf{x}_0) = +\infty$.

- (ii) Supongamos ahora que el conjunto $\mathcal{L}(\mathbf{x}_0)$ es acotado y $\tau(\mathbf{x}_0) < +\infty$. Por el Teorema 3.5 tenemos que $\lim_{t \uparrow \tau(\mathbf{x})} \|\mathbf{x}(t)\| = +\infty$, lo cual implica que la trayectoria $\mathbf{x}(t)$ no converge a ningún punto de acumulación. Dado que $\mathcal{L}(\mathbf{x}_0)$ es acotado, podemos definir τ_0 de la siguiente forma (Figura 2):

$$\tau_0 = \inf \{s \geq 0 \mid s < \tau(\mathbf{x}_0), \mathbf{x}(s) \in \mathbb{R}^n - \mathcal{L}(\mathbf{x}_0)\} < +\infty;$$

por hipótesis, $\mathcal{L}(\mathbf{x}_0)$ es acotado, y por la continuidad de E_λ es un conjunto

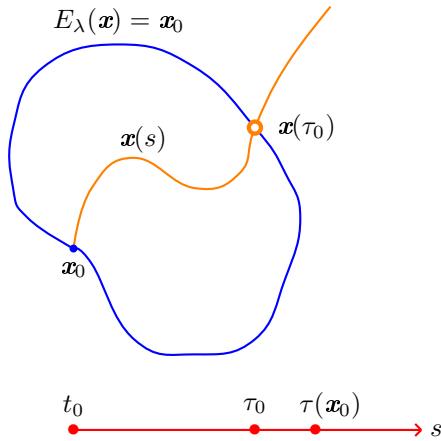


Figura 2. Trayectoria de la solución $\mathbf{x}(t)$ cuando se supone que $\tau(\mathbf{x}_0) < +\infty$.

cerrado; entonces $\mathbf{x}(\tau_0) \in \mathcal{L}(\mathbf{x}_0)$. Además, $\tau_0 < \tau(\mathbf{x}_0)$. Por lo tanto, para algunos s en el intervalo $(\tau_0, \tau(\mathbf{x}_0))$, se satisface que

$$E_\lambda(\mathbf{x}(s)) > E_\lambda(\mathbf{x}(\tau_0)). \quad (11)$$

Sin embargo, $\frac{dE_\lambda(\mathbf{x}(t))}{dt} = \nabla E_\lambda(\mathbf{x}(t))^T \frac{d\mathbf{x}}{dt} = -\rho \|\nabla E_\lambda(\mathbf{x}(t))\|^2 \leq 0$, lo cual significa que ∇E_λ es no creciente en $[\tau_0, \tau(\mathbf{x}_0))$, lo que contradice (11). Por lo tanto, $\tau(\mathbf{x}_0) = +\infty$.

✓

El siguiente resultado establece la convergencia de la trayectoria solución.

Corolario 4.4. *Sea $\mathbf{x}(t)$, con $t \in [t_0; \tau(\mathbf{x}_0))$, una solución maximal de (7).*

(i) *Si $\tau(\mathbf{x}_0) = +\infty$ y $\{\mathbf{x}(t)\}$ es acotada, entonces*

$$\lim_{t \rightarrow +\infty} \nabla E_\lambda(\mathbf{x}(t)) = 0. \quad (12)$$

(ii) *Si F es una función P_0 , entonces $\mathcal{L}(\mathbf{x}_0)$ es acotado y todo punto de acumulación de la trayectoria $\mathbf{x}(t)$ es solución del PCNL.*

*Demuestra*ción. (i) Demostramos anteriormente que $\frac{dE_\lambda(\mathbf{x}(t))}{dt} \leq 0$, lo que implica que $E_\lambda(\mathbf{x}(t))$ es decreciente y es acotada inferiormente, ya que $E_\lambda(\mathbf{x}(t)) \geq 0$. Luego, existe un punto de acumulación \mathbf{x}_* , en el cual E_λ alcanza un ínfimo; más aún, este punto es un minimizador de E_λ , lo cual implica que $\nabla E_\lambda(\mathbf{x}_*) = 0$. Además, dado que $\tau(\mathbf{x}_0) = +\infty$, podemos pasar al límite; en efecto,

$$0 = \nabla E_\lambda(\mathbf{x}_*) = \nabla E_\lambda \left(\lim_{t \rightarrow +\infty} \mathbf{x}(t) \right) = \lim_{t \rightarrow +\infty} \nabla E_\lambda(\mathbf{x}(t)).$$

La hipótesis de que $\{\mathbf{x}(t)\}$ esté acotada garantiza que el sistema dinámico (7) tiene solución. En [18] se muestra un ejemplo en el que la conclusión del corolario no se tiene, si $\{\mathbf{x}(t)\}$ no es acotada.

- (ii) Supongamos que F es una función P_0 ; entonces $\mathcal{L}(\mathbf{x}_0)$ es acotado ([15]) y por tanto $\{\mathbf{x}(t)\}$ también es acotada. Luego si \mathbf{x}_* es un punto de acumulación, $\nabla E_\lambda(\mathbf{x}_*) = 0$, lo cual implica que \mathbf{x}_* es un minimizador de E_λ y que F es una función P_0 ; \mathbf{x}_* también es solución del NCPL ([15]).

✓

En lo que sigue, consideramos una solución \mathbf{x}_* del PCNL, la cual por la Proposición 4.1 es también un *punto de equilibrio* de (7), y supondremos que \mathbf{x}_* es un *punto de equilibrio aislado* de (7).

Teorema 4.5. *Si \mathbf{x}_* es un punto de equilibrio aislado de (7), \mathbf{x}_* es estable asintóticamente para (7).*

*Demuestra*ción. Sean $\Omega_* \subseteq \mathbb{R}^n$ una vecindad abierta del *punto de equilibrio aislado* \mathbf{x}_* y E_λ la función definida en (4), la cual es continuamente diferenciable y no negativa para todo $\mathbf{x} \in \mathbb{R}^n$. Debido a que \mathbf{x}_* es una solución del PCNL,

$$E_\lambda(\mathbf{x}_*) = 0. \quad (13)$$

Supongamos que existe un $\mathbf{x} \in \Omega_*$ con $\mathbf{x} \neq \mathbf{x}_*$ tal que $E_\lambda(\mathbf{x}) = 0$. Por (6) tenemos que $\nabla E_\lambda(\mathbf{x}) = V^T \Phi_\lambda(\mathbf{x}) = 0$, para toda $V \in \partial \Phi_\lambda(\mathbf{x})$, lo que implica que \mathbf{x} es un punto de equilibrio de (7), contradiciendo la hipótesis de que \mathbf{x}_* es un punto de equilibrio aislado de (7). Luego

$$E_\lambda(\mathbf{x}) > 0, \quad \text{para todo } \mathbf{x} \in \Omega_*, \mathbf{x} \neq \mathbf{x}_*. \quad (14)$$

Por otra parte tenemos que

$$\frac{dE_\lambda(\mathbf{x}(t))}{dt} = \nabla_{\mathbf{x}(t)} E_\lambda(\mathbf{x}(t))^T \frac{d\mathbf{x}}{dt} = -\|\nabla_{\mathbf{x}(t)} E_\lambda(\mathbf{x}(t))\|^2 \leq 0. \quad (15)$$

Así, de (13), (14), (15) y de la Definición 3 tenemos que E_λ es una *función de Liapunov* para (7) sobre Ω_* .

Dado que \mathbf{x}_* es un punto de equilibrio aislado del modelo (7), de (15) concluimos que

$$\frac{dE_\lambda(\mathbf{x}(t))}{dt} < 0, \quad \text{para todo } \mathbf{x}(t) \in \Omega_*, \mathbf{x}(t) \neq \mathbf{x}_*.$$

Por lo tanto, el Teorema 3.9 garantiza que \mathbf{x}_* es estable asintóticamente para (7). ✓

Diremos que una solución \mathbf{x}_* de PCNL se llama *regular*, si todas las matrices del $\partial\Phi_\lambda(\mathbf{x}_*)$ son no singulares ([22]). Con frecuencia la *regularidad* conduce a convergencia superlineal de algunos métodos de optimización.

Un aspecto interesante del siguiente teorema es la relación del concepto de *regularidad* de una solución del PCNL con el de *estabilidad exponencial*. La demostración del teorema es análoga a la del Teorema 4.6 en [17] para la función de complementariedad de *Fischer-Burmeister* (que corresponde a φ_λ con $\lambda = 2$).

Teorema 4.6. *Si \mathbf{x}_* es una solución regular del PCNL, entonces \mathbf{x}_* es estable exponencialmente para (7).*

*Demuestra*o. Dado que \mathbf{x}_* es una solución regular del PCNL, existen ([23]) una vecindad $B(\mathbf{x}_*; \delta) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_*\| < \delta\}$ de \mathbf{x}_* y una constante C tal es que, para todo \mathbf{x} en dicha vecindad y cualquier $V \in \partial\Phi_\lambda(\mathbf{x})$, la matriz V es no singular y

$$\max\{\|V\|, \|V^{-1}\|\} \leq C. \quad (16)$$

Sea $\delta > 0$ suficientemente pequeño tal que, para cualquier $\mathbf{x}(t_0) \in B(\mathbf{x}_*; \delta)$, se tiene que $\mathbf{x}(t) \rightarrow \mathbf{x}_*$ siempre que $t \rightarrow \infty$, y de modo que el resultado del párrafo anterior se satisfaga. Por tanto, existen κ_1 y κ_2 tales que para todo $\mathbf{x} \in B(\mathbf{x}_*; \delta)$,

$$\kappa_1 \|\mathbf{v}\|^2 \leq \mathbf{v}^T V^T V \mathbf{v} \leq \kappa_2 \|\mathbf{v}\|^2. \quad (17)$$

Por otra parte, si \mathbf{x}_* es una solución regular del PCNL, entonces es una solución aislada de $\Phi_\lambda(\mathbf{x}) = 0$ ([22], [23]). Luego, de (9), \mathbf{x}_* es un punto de equilibrio aislado de (7), lo cual por el Teorema 4.5 implica que \mathbf{x}_* es estable asintóticamente.

Dado que la función Φ_λ es semisuave ([15]), se tiene que para todo $V \in \partial\Phi_\lambda(\mathbf{x})$

$$\Phi_\lambda(\mathbf{x}) = \Phi_\lambda(\mathbf{x}_*) + V(\mathbf{x} - \mathbf{x}_*) + o(\|\mathbf{x} - \mathbf{x}_*\|). \quad (18)$$

Observemos que, $o(\|\mathbf{x} - \mathbf{x}_*\|) = H(\mathbf{x})$, donde H es un campo vectorial sobre \mathbb{R}^n . Esto significa que

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_*} \frac{\|H(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_*\|} = 0,$$

lo cual implica que, para todo $\epsilon > 0$, existe $\bar{\delta} > 0$ tal que

$$\frac{\|H(\mathbf{x})\|}{\|\mathbf{x} - \mathbf{x}_*\|} \leq \epsilon, \quad \text{para todo } \mathbf{x} \in B(\mathbf{x}_*; \bar{\delta}).$$

Equivalentemente,

$$o(\|\mathbf{x} - \mathbf{x}_*\|) \leq \epsilon \|\mathbf{x} - \mathbf{x}_*\| \quad \text{para todo } \mathbf{x} \in B(\mathbf{x}_*; \bar{\delta}). \quad (19)$$

En particular, la condición (19) se satisface para algún $0 < \epsilon < \kappa_1$ y para todo \mathbf{x} en $B(\mathbf{x}_*; \delta)$. Aquí hemos usado el hecho de que (ver Figura 3), si $\bar{\delta} \geq \delta$, entonces la condición (19) se satisface para todo $\mathbf{x} \in B(\mathbf{x}_*; \delta)$; y si $\bar{\delta} < \delta$, entonces reducimos el tamaño de δ para que (19) se satisfaga.

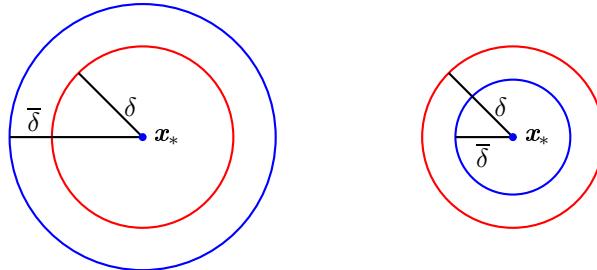


Figura 3. Reduciendo el tamaño de δ si fuese necesario.

Ahora consideremos la función Γ definida por $\Gamma(t) = \|\mathbf{x}(t) - \mathbf{x}_*\|^2$, para todo $t \in [t_0, \infty)$. Derivando Γ respecto a t y usando (6), tenemos que, para todo $V \in \partial\Phi_\lambda(\mathbf{x}(t))$,

$$\frac{d\Gamma(t)}{dt} = 2[\mathbf{x}(t) - \mathbf{x}_*]^T \frac{d\mathbf{x}}{dt} = -2[\mathbf{x}(t) - \mathbf{x}_*]^T V^T \Phi_\lambda(\mathbf{x}(t)). \quad (20)$$

Supongamos que $T = \{t \in [t_0, \infty) : \|\mathbf{x}(t) - \mathbf{x}_*\| \geq \delta\} \neq \emptyset$. El conjunto T es acotado inferiormente, luego existe $\bar{\tau} = \inf T$, que corresponde al primer valor de t para el cual $\mathbf{x}(t) \notin B(\mathbf{x}_*; \delta)$.

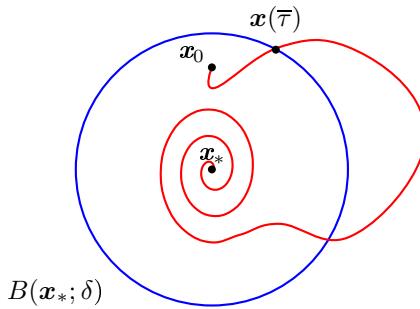


Figura 4. \mathbf{x}_* es asintóticamente estable, pero no exponencialmente estable.

Sustituyendo (18) en (21) y teniendo en cuenta que $\Phi_\lambda(\mathbf{x}_*) = 0$, tenemos que para todo $t \in \bar{I} = [t_0, \bar{\tau})$,

$$\frac{d\Gamma(t)}{dt} = -2(\mathbf{x}(t) - \mathbf{x}_*)^T V^T V(\mathbf{x}(t) - \mathbf{x}_*) + (-2)(\mathbf{x}(t) - \mathbf{x}_*)^T V^T o(\|\mathbf{x}(t) - \mathbf{x}_*\|). \quad (21)$$

Usando (17) se tiene que

$$-2(\mathbf{x}(t) - \mathbf{x}_*)^T V^T V(\mathbf{x}(t) - \mathbf{x}_*) \leq -2\kappa_1 \|\mathbf{x}(t) - \mathbf{x}_*\|^2. \quad (22)$$

El segundo sumando de (21) satisface la desigualdad

$$(-2)(\mathbf{x}(t) - \mathbf{x}_*)^T V^T o(\|\mathbf{x}(t) - \mathbf{x}_*\|) \leq \epsilon \|\mathbf{x}(t) - \mathbf{x}_*\|^2, \quad (23)$$

porque $g(\mathbf{x}) = (-2)(\mathbf{x}(t) - \mathbf{x}_*)^T V^T o(\|\mathbf{x}(t) - \mathbf{x}_*\|) = o(\|\mathbf{x}(t) - \mathbf{x}_*\|^2)$, ya que

$$\lim_{\mathbf{x}(t) \rightarrow \mathbf{x}_*} \frac{|g(\mathbf{x})|}{\|\mathbf{x}(t) - \mathbf{x}_*\|^2} = 0.$$

En efecto,

$$\begin{aligned} \lim_{\mathbf{x}(t) \rightarrow \mathbf{x}_*} \frac{|g(\mathbf{x})|}{\|\mathbf{x}(t) - \mathbf{x}_*\|^2} &\leq \lim_{\mathbf{x}(t) \rightarrow \mathbf{x}_*} \frac{2\|V\| \|(\mathbf{x}(t) - \mathbf{x}_*)\| o(\|\mathbf{x}(t) - \mathbf{x}_*\|)}{\|\mathbf{x}(t) - \mathbf{x}_*\| \|\mathbf{x}(t) - \mathbf{x}_*\|} \\ &\leq 2\|V\| \lim_{\mathbf{x}(t) \rightarrow \mathbf{x}_*} \frac{o(\|\mathbf{x}(t) - \mathbf{x}_*\|)}{\|\mathbf{x}(t) - \mathbf{x}_*\|} \leq 2(C)(0) = 0, \end{aligned}$$

donde C es la constante dada en (16).

Luego, por (21), (22) y (23) tenemos que

$$\frac{d\Gamma(t)}{dt} \leq (-2\kappa_1 + \epsilon) \|\mathbf{x}(t) - \mathbf{x}_*\|^2 = (-2\kappa_1 + \epsilon) \Gamma(t).$$

Usando el Corolario 2.1 en [27] tenemos que $\Gamma(t) \leq e^{(-2\kappa_1 + \epsilon)t} \Gamma(t_0)$, para todo $t \in \bar{I}$, o equivalentemente,

$$\|\mathbf{x}(t) - \mathbf{x}_*\| \leq e^{\omega t} \|\mathbf{x}(t_0) - \mathbf{x}_*\|, \quad t \in \bar{I}, \quad (24)$$

con $\omega = -\kappa_1 + \epsilon/2 < 0$. Tomando el límite superior en (24), tenemos que

$$\delta \leq \limsup_{t \uparrow \bar{T}} \|\mathbf{x}(t) - \mathbf{x}_*\| \leq e^{\omega \bar{T}} \|\mathbf{x}(t_0) - \mathbf{x}_*\| < \delta,$$

lo cual es una contradicción. Luego, $T = \emptyset$ e $\bar{I} = [0, +\infty)$, y por la desigualdad (24) se completa la prueba de la estabilidad exponencial de \mathbf{x}_* . \checkmark

Del teorema anterior tenemos que para todo $t \in [0, +\infty)$ se satisface la desigualdad (24). En particular, para $t = k \in \mathbb{N}$ con $\mathbf{x}(k) := \mathbf{x}_k$:

$$\|\mathbf{x}_k - \mathbf{x}_*\| \leq e^{\omega k} \|\mathbf{x}_0 - \mathbf{x}_*\|;$$

esto garantiza la convergencia de cualquier sucesión que aproxime la trayectoria solución del modelo (7), como se ilustra en la simulación numérica que presentamos en la siguiente sección.

5. Pruebas numéricas

En esta sección analizamos numéricamente el algoritmo propuesto. Para ello usamos siete problemas de complementariedad asociados a algunas funciones de prueba ampliamente utilizadas en la literatura sobre complementariedad, las cuales definimos a continuación (entre paréntesis aparece la abreviatura que utilizaremos en las tablas de resultados para hacer referencia a cada problema), con su solución (\mathbf{x}_*) y el punto inicial usado en los algoritmos (\mathbf{x}_0) .

Problema de Billups (Billups): $F: \mathbb{R} \rightarrow \mathbb{R}$ está definida como

$$F(\mathbf{x}) = (x - 1)^2 - 1.1;$$

$$\mathbf{x}^* = 2.0488 \text{ y } \mathbf{x}_0 = 0.$$

Problema de Kojima–Shindo (Koj-Shi): $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$ está definida como

$$F(\mathbf{x}) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_2^2 + x_1 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 1 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix};$$

$$\mathbf{x}^* = (\sqrt{6}/2 \ 0 \ 0 \ 0.5)^T \text{ y } \mathbf{x}_0 = (0 \ 0 \ 0 \ 0)^T.$$

Problema de Kojima–Shindo modificado (Koj-Shi-mod): $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$,

$$F(\mathbf{x}) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_2^2 + x_1 + 10x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 9x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix};$$

$$\mathbf{x}^* = (\sqrt{6}/2 \ 0 \ 0 \ 0.5)^T \text{ y } \mathbf{x}_0 = (0 \ 0 \ 0 \ 0)^T.$$

Problema de Kojima–Josephy (Koj-Jo): $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$, definida como

$$F(\mathbf{x}) = \begin{pmatrix} 3x_1^2 + 2x_1x_2 + 2x_2^2 + x_3 + 3x_4 - 6 \\ 2x_1^2 + x_2^2 + x_1 + 3x_3 + 2x_4 - 2 \\ 3x_1^2 + x_1x_2 + 2x_2^2 + 2x_3 + 3x_4 - 9 \\ x_1^2 + 3x_2^2 + 2x_3 + 3x_4 - 3 \end{pmatrix};$$

$$\mathbf{x}^* = (1 \ 0 \ 3 \ 0)^T \text{ y } \mathbf{x}_0 = (0 \ 0 \ 0 \ 0)^T.$$

Problema de Mathiesen Modificado (Mathiesen): $F: \mathbb{R}^4 \rightarrow \mathbb{R}^4$,

$$F(\mathbf{x}) = \begin{pmatrix} -x_2 + x_3 + x_4 \\ x_1 - \frac{4.5x_3 + 2.7x_4}{x_2 + 1} \\ 5 - x_1 - \frac{0.5x_3 + 0.3x_4}{x_3 + 1} \\ 3 - x_1 \end{pmatrix};$$

$$\mathbf{x}^* = (a \ 0 \ 0 \ 0)^T, \ a \in [0, 3] \text{ y } \mathbf{x}_0 = (1 \ 1 \ 1 \ 1)^T.$$

Problema tomado de [17] (cuadrático): $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ definida como

$$F(\mathbf{x}) = \begin{pmatrix} 2x_1 + 4x_2 \\ 2x_2 + 4x_1 \end{pmatrix};$$

$$\mathbf{x}^* = (0 \ 0)^T \text{ y } \mathbf{x}_0 = (1 \ 1)^T.$$

A continuación presentamos dos algoritmos que componen el modelo de red neuronal propuesto para resolver el PCNL. El primero describe un procedimiento para calcular una matriz $V \in \partial\Phi_\lambda(\mathbf{x})$ análogo al usado en [9] para la *función de Fischer*, y cuya idea general consiste en considerar la sucesión de vectores $\{\mathbf{y}_k = \mathbf{x} + \varepsilon_k \mathbf{z}\}$, donde $\{\varepsilon_k\}$ es una sucesión de números positivos que converge a cero y \mathbf{z} es un vector tal que $z_i = 1$

si $F_i(\mathbf{x}) = x_i = 0$, y $z_i = 0$ en otro caso. La sucesión así construida converge a \mathbf{x} y las matrices jacobianas $\Phi'_\lambda(\mathbf{y}_k)$ existen. Esta deducción se encuentra detalladamente en [1] y [15].

El segundo algoritmo constituye propiamente el *modelo de red neuronal propuesto* para resolver el PCNL. Dentro de este algoritmo implementamos una versión matricial del método de *Runge-Kutta* de cuarto orden ([5]), cuya iteración básica es de la forma

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{h}{6} M [1 \ 2 \ 2 \ 1]^T,$$

donde M es una matriz de tamaño $n \times 4$, siendo n la dimensión del problema.

Algoritmo 1 Cálculo de las matrices en $\partial\Phi_\lambda(\mathbf{x})$

Entrada: Dado \mathbf{x} , se construye $\beta = \{i \mid x_i = 0 = F_i(\mathbf{x})\}$.

Salida: $V \in \partial\Phi_\lambda(\mathbf{x})$.

```

1: para  $i = 1$  hasta  $n$  hacer
2:   si  $i \in \beta$  entonces
3:      $S \leftarrow \sqrt{(z_i - \nabla F_i(\mathbf{x})^T \mathbf{z})^2 + z_i \lambda \nabla F_i(\mathbf{x})^T \mathbf{z}}$ .
4:      $V_i \leftarrow \left( \frac{2(z_i - \nabla F_i(\mathbf{x})^T \mathbf{z}) + \lambda \nabla F_i(\mathbf{x})^T \mathbf{z}}{S} - 1 \right) \mathbf{e}_i^T + \left( \frac{-2(z_i - \nabla F_i(\mathbf{x})^T \mathbf{z}) + \lambda z_i}{S} - 1 \right) \nabla F_i(\mathbf{x})^T$ 
5:   si no
6:      $S \leftarrow \sqrt{(x_i - F_i(\mathbf{x}))^2 + x_i \lambda F_i(\mathbf{x})}$ .
7:      $V_i \leftarrow \left( \frac{2(x_i - F_i(\mathbf{x})) + \lambda F_i(\mathbf{x})}{S} - 1 \right) \mathbf{e}_i^T + \left( \frac{2(x_i - F_i(\mathbf{x})) + \lambda x_i}{S} - 1 \right) \nabla F_i(\mathbf{x})^T$ 
8:   fin si
9: fin para
10: Salida  $V$ 

```

Algoritmo 2 Modelo de red neuronal para resolver el PCNL

Entrada: $t_0 = 0$ y $\mathbf{x}_0 = \mathbf{x}(t_0)$.

Salida: Solución numérica de $\mathbf{x}(t)$.

- 1: Se calcula $V \equiv V(\mathbf{x}_0)$ mediante el algoritmo 1.
- 2: Se calcula $\Phi_\lambda(\mathbf{x}_0)$.
- 3: **mientras** $\frac{1}{2} \|\Phi_\lambda(\mathbf{x}_k)\|^2 \geq 10^{-10}$ y $\|V^T \Phi_\lambda(\mathbf{x}_k)\| \geq 10^{-5}$ **hacer**
- 4: Actualizamos \mathbf{x}_{k+1} , utilizando la forma matricial de *Runge-Kutta*, es decir,

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{h}{6} [M^1 \ M^2 \ M^3 \ M^4] [1 \ 2 \ 2 \ 1]^T$$
, donde
$$\begin{aligned} M^1 &= G(\mathbf{x}_k), \\ M^2 &= G(\mathbf{x}_k + \frac{1}{2}h G(\mathbf{x}_k)), \\ M^3 &= G(\mathbf{x}_k + \frac{1}{2}h G(\mathbf{x}_k + \frac{1}{2}h G(\mathbf{x}_k))), \\ M^4 &= G(\mathbf{x}_k + h G(\mathbf{x}_k + \frac{1}{2}h G(\mathbf{x}_k + \frac{1}{2}h G(\mathbf{x}_k)))), \end{aligned}$$
con $G(\mathbf{x}_k) = -\rho V^T \Phi_\lambda(\mathbf{x}_k)$.
- 5: Se calcula $V \equiv V(\mathbf{x}_k)$ mediante el algoritmo 1
- 6: Se calcula $\Phi_\lambda(\mathbf{x}_k)$.
- 7: **fin mientras**
- 8: **Salida** \mathbf{x}_k .

Para escribir los códigos de los algoritmos utilizamos el *software* MATLAB®. Realizamos los experimentos numéricos en un computador con procesador Intel(R) Core(TM) i5-3450S CPU de 2,80 GHz en un sistema operativo de 64 bits.

La Tabla 1 contiene los resultados obtenidos con el Algoritmo 2, el cual usa un parámetro λ dinámico ([15]) y los obtenidos con un algoritmo que usa un modelo análogo al propuesto en [17], en el cual $\lambda = 2$. Estos últimos resultados se indican en la tabla usando un asterisco después de la abreviatura del problema. Para todos los experimentos, el valor del parámetro ρ , en el modelo, es $\rho = 10^3$, y el tamaño de paso escogido en el método de *Runge-Kutta* es $h = 10^{-5}$.

La primera columna de la tabla indica el problema utilizado; la segunda, da el valor del parámetro t en el cual declaramos convergencia, y las tres columnas siguientes contienen los valores finales de $\|\mathbf{x}(t) - \mathbf{x}^*\|$, $\|\nabla E_\lambda(\mathbf{x}(t))\|$ y $|E_\lambda(\mathbf{x}(t))|$, respectivamente.

Problema	t	$\ \mathbf{x}(t) - \mathbf{x}^*\ $	$\ \nabla E_\lambda\ $	$ E_\lambda $
Billups	1,270000e-03	2,398465e-06	5,675707e-05	9,151717e-11
Billups*	3,010000e-03	2,161289e-06	2,942230e-05	9,837168e-11
Koj-Shi	4,931000e-03	2,113732e-05	1,888083e-05	9,977221e-11
Koj-Shi*	2,199000e-02	2,229580e-05	9,957821e-06	1,110084e-10
Koj-Shi-mod	5,249000e-03	2,010010e-05	1,892561e-05	9,977325e-11
Koj-Shi-mod*	2,327000e-02	2,061545e-05	9,987753e-06	1,029506e-10
Koj-Jos	9,847000e-03	2,761663e-05	1,447005e-05	9,990373e-11
Koj-Jos*	4,337000e-02	3,812819e-05	9,988814e-06	1,904285e-10
Mathiesen	6,669000e-02	3,005689e-02	9,999414e-06	3,586985e-10
Mathiesen*	2,600000e-02	2,990021e-02	9,976923e-06	1,569724e-09
Cuadrático	2,860000e-03	1,525697e-05	2,798854e-05	9,791978e-11
Cuadrático*	1,359000e-02	1,530965e-05	1,288038e-05	9,859705e-11

Cuadro 1. Resultados con λ dinámico y con $\lambda = 2$.

Esta tabla muestra que el modelo propuesto es competitivo, y en el caso del problema cuadrático es mejor. Cabe mencionar que el problema de *Billups* fue construido por *Billups* ([4]) con el objetivo de hacer que los métodos usados para resolverlo fallen ([15]); es decir, es un problema de alta exigencia en cuanto a su convergencia, pero observamos que los algoritmos lo resolvieron.

6. Conclusiones

En este artículo proponemos un modelo de red neuronal basado en el método de máximo descenso para resolver problemas de complementariedad no lineal, el cual usa por primera vez una familia uniparamétrica de funciones de complementariedad como función energía. Establecemos resultados de existencia, estabilidad y de convergencia de la trayectoria

de la red neuronal. Además, presentamos algunas pruebas numéricas preliminares que muestran que el modelo propuesto es competitivo.

Finalmente, en búsqueda de ampliar el espectro de trabajo con respecto a redes neuronales para complementariedad no lineal, consideramos conveniente realizar más pruebas numéricas del algoritmo propuesto y modificar el modelo propuesto para resolver problemas con restricciones.

Agradecimientos

Los autores agradecen a la Universidad del Cauca por el tiempo concedido para esta investigación mediante el Proyecto de investigación VRI ID 4189, y a un árbitro anónimo por sus sugerencias y comentarios que ayudaron a mejorar la presentación y contenido de este artículo.

Referencias

- [1] Arenas F., Martínez H.J. and Pérez R., "Least change secant update methods for nonlinear complementarity problem", *Ingeniería y Ciencia* 11 (2015), 11-36.
- [2] Arenas F., Martínez H.J. & Pérez R., "Redefinición de la función de complementariedad de Kanzow", *Revista de Ciencias* 18 (2014), No. 2, 111-122.
- [3] Arias C.A., *Un algoritmo cuasi Newton global para problemas de complementariedad no lineal*, Tesis de Maestría, Universidad del Cauca, 2014, 44 p.
- [4] Billups S. C., *Algorithms for complementarity problems and generalized equations*, Thesis (Ph.D.), University of Wisconsin, 1995.
- [5] Burden R.L. & Douglas Faires J., *Análisis Numérico*, 7ma ed., Thomson Learning, 2002.
- [6] Chen J-S., Ko C-H. and Pan S., "A Neural network based on the generalized Fisher-Burmeister function for nonlinear complementarity problems", *Information Sciences* 180 (2010), No. 5, 697-711.
- [7] Chen J-S. and Pan S., "A family of NCP functions and a descent method for the nonlinear complementarity problem", *Comput. Optim. Appl.* 40 (2008), No. 3, 389-404.
- [8] Clarke F.H., *Optimization and nonsmooth analysis*, John Wiley & Sons, Inc., New York, 1983.
- [9] De Luca T., Facchinei F. and Kanzow C., "A semismooth equation approach to the solution of nonlinear complementarity problems", *Math. Programming* 75 (1996), No. 3, Ser. A, 407-439.
- [10] Dennis J.E., Jr. and Schnabel R.B., *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice Hall, Inc., Englewood Cliffs, NJ, 1983.
- [11] Ding J. and Yin H., "A new homotopy method for solving non-linear complementarity problems", *Numer. Math. J. Chi. Univ. (Engl. Ser.)* 16 (2007), No. 2, 155-163.
- [12] Ferris M.C. and Pang J.S., "Engineering and economic applications of complementarity problems", *SIAM Rev.* 39 (1997), No. 4, 669-713.

- [13] Fischer A., "A special Newton-type optimization method", *Optimization* 24 (1992), No. 3-4, 269–284.
- [14] Hopfield J. and Tank D.W., "Neural computation of decisions in optimization problems", *Biol. Cybernet.* 52 (1985), No. 3, 141–152.
- [15] Kanzow C.J. and Kleinmichel H., "A new class of semismooth Newton-type methods for nonlinear complementarity problems", *Computat. Optim. Appl.* 11 (1998), No. 3, 227–251.
- [16] Kostreva M., "Elasto-hydrodynamic lubrication: An non-linear complementarity problem", *Internat. J. Numer. Methods Fluids* 4 (1984), No. 4, 377–397.
- [17] Liao L-Z, Qi H. and Qi L., "Solving nonlinear complementarity problems with neural networks: a reformulation method approach", *J. Comput. Appl. Math.* 131 (2001), No. 1-2, 343–359.
- [18] Lillo W.E., Loh M.H., Hui S. and Zak S.H., "On solving constrained optimization problems with neural networks: a penalty method approach", *IEEE Trans. Neural Networks* 4 (1993), No. 6, 931–940.
- [19] Nocedal J. and Wright S.J., *Numerical optimization*, Second ed., Springer, New York, 2006.
- [20] Pang J-S. and Qi L.Q., "Nonsmooth equations: motivation and algorithms", *SIAM J. Optim.* 3 (1993), No. 3, 443–465.
- [21] Pérez R., Lopes V.L.R. and Martínez J.M., "On the local convergence of quasi-Newton methods for solving non linear complementarity problems", *Appl. Numer. Math.* 30 (1999), No. 1, 3–22.
- [22] Qi L.Q., "Convergence analysis of some algorithms for solving nonsmooth equations", *Math. Oper. Research* 18 (1993), No. 1, 227–244.
- [23] Qi L.Q. and Sun J., "A nonsmooth version of Newton method", *Math. Programming*, 58 (1993), No. 3, Ser. A, 353–368.
- [24] Wardrop J.G., "Some theoretical aspects of road traffic research", *Proceedings of the ICE* 1 (1952), No. 3, 325–362.
- [25] Watkins D.S., *Fundamentals of matrix computations*, Second ed., Wiley-Interscience, New York, 2002.
- [26] Xu Q. and Dang C., "A new homotopy method for solving non-linear complementarity problems", *Optimization* 57 (2008), No. 5, 681–689.
- [27] Zabczyk J., *Mathematical control theory: an introduction*, Birkhäuser, Boston Inc., Boston, MA, 1992.