Revista Integración
Escuela de Matemáticas
Universidad Industrial de Santander
Vol. 32, No. 2, 2014, pág. 181–197

Sobre la convergencia de un método secante para ecuaciones matriciales no lineales

Mauricio Macías C.[a*], Héctor J. Martínez[b], Rosana Pérez[a]

[a] Universidad del Cauca, Departamento de Matemáticas, Popayán, Colombia.

[b] Universidad del Valle, Departamento de Matemáticas, Cali, Colombia.

Resumen. En este artículo desarrollamos una teoría general de convergencia de un método secante para resolver ecuaciones matriciales no lineales. Además, presentamos condiciones suficientes para que este método proporcione un algoritmo local y superlinealmente convergente.

Palabras clave: Función matricial, operador de Fréchet, Fréchet diferenciable, método secante, ecuación matricial no lineal, convergencia superlineal.
MSC2010: 65F10, 65N22, 65H10, 49M15, 49M37, 90C53.

On the convergence of a secant method for nonlinear matrix equations

Abstract. In this paper we develop a general theory of convergence of a secant method to solve nonlinear matrix equations. In addition, we give sufficient conditions in order to this method provide a local and superlinearly convergent algorithm.

Keywords: Matrix function, Fréchet operator, Fréchet differentiable, secant method, nonlinear matrix equation, superlinear convergence.

1. Introducción

Una función de variable y valor matricial $F\colon \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$ es llamada una *función de matrices*. Este tipo de funciones ha sido estudiado desde hace mucho tiempo; de hecho, Cayley (1858), en su obra *A Memoir on the Theory of Matrices,* investigó la raíz cuadrada de una matriz y, posteriormente, Sylvester y otros autores dieron una definición formal de una función de matrices [8]. Higham (2008), en su libro *Functions of matrices: theory and computations* [8], presenta una recopilación de su trabajo de investigación a lo largo de muchos años sobre funciones de matrices.

[*] E-mail: `mauromac@unicauca.edu.co`
Recibido: 20 de marzo de 2014, Aceptado: 13 de agosto de 2014.
Para citar este artículo: E.M. Macías, H.J. Martínez, R. Pérez, Sobre la convergencia de un método secante para ecuaciones matriciales no lineales, *Rev. Integr. Temas Mat.* 32 (2014), no. 2, 181-197.

Son numerosas las áreas de la ciencia e ingeniería en las que aparecen problemas relacionados con funciones de matrices, entre ellas, en ecuaciones diferenciales, modelos de Márkov, teoría de control y resonancia nuclear magnética [8].

En gran parte, el desarrollo de una teoría sobre funciones de matrices ha sido motivado por la necesidad de resolver *ecuaciones matriciales no lineales.* Dos ejemplos de este tipo de ecuaciones con una gran variedad de aplicaciones [8] son la llamada *ecuación de Riccati* ($XFX - A^*X - XA - G = \mathbf{O}$, donde A^* denota la transpuesta conjugada de A, y F y G son *hermitianas*), y la *ecuación cuadrática matricial* ($AX^2 + BX + C = \mathbf{O}$, donde $A, B, C \in \mathbb{C}^{n\times n}$) [7, 9]. Esta última ecuación es una generalización de la ecuación cuadrática escalar.

Entre los métodos numéricos que resuelven una ecuación matricial no lineal se destaca el amplio uso del *método de Newton* [8, 10, 16]. Recientemente fue propuesto un *método tipo secante* para resolver dicha ecuación [15], dejando un camino abierto para investigar al respecto. Esto, unido al hecho de que son numerosas las aplicaciones en las que es necesario resolver una ecuación matricial no lineal, nos motivó a desarrollar esta investigación.

El presente artículo está organizado de la siguiente forma. En la Sección **2** presentamos en forma descriptiva los sistemas de ecuaciones no lineales; inicialmente abordamos el caso *vectorial,* para luego extenderlo al *matricial.* En la Sección **3** demostramos que el algoritmo tipo secante propuesto en [15] para resolver problemas matriciales no lineales, es de *cambio mínimo* [13] y, usando las reglas generales de la teoría de convergencia local para métodos secantes [4], demostramos que dicho algoritmo converge local y superlinealmente. En la Sección **4** exploramos numéricamente el comportamiento local del *método secante* analizado teóricamente en la Sección **3**. Para ello, resolvemos una *ecuación cuadrática matricial.* En la Sección **5** hacemos algunos comentarios finales y propuestas para trabajos futuros sobre el tema.

2. Sistemas de ecuaciones no lineales

En muchos problemas en diferentes áreas de la investigación aplicada surge la necesidad de resolver un sistema de ecuaciones no lineales, que consiste en encontrar un vector $\boldsymbol{x} \in \mathbb{R}^n$ que satisfaga la ecuación,

$$F(\boldsymbol{x}) = \mathbf{0}, \tag{1}$$

donde $F : \mathbb{R}^n \to \mathbb{R}^n$ es una función no lineal y continuamente diferenciable [4, 18].

Al igual que en el caso de los sistemas de ecuaciones lineales, los no lineales pueden tener solución única, infinitas soluciones o no tener solución. En contraste con el caso de los sistemas de ecuaciones lineales, los métodos directos para la solución de sistemas no lineales son usualmente aplicables solamente para sistemas pequeños; así, todos los algoritmos prácticos para resolver (1) son iterativos [18].

Uno de los métodos más populares y frecuentemente usados para resolver el problema (1), por su convergencia cuadrática (bajo ciertas hipótesis), es el *método de Newton* [4, 18]. En este método, dado una aproximación inicial $\boldsymbol{x}_0 \in \mathbb{R}^n$ a la solución de (1), se considera en cada iteración un modelo de $F(\boldsymbol{x})$ de la forma $M_k(\boldsymbol{x}) \equiv F(\boldsymbol{x}_k) + F'(\boldsymbol{x}_k)(\boldsymbol{x} - \boldsymbol{x}_k)$, donde $F'(\boldsymbol{x}_k)$ denota la *matriz jacobiana* de F en $\boldsymbol{x}_k$, y se resuelve el problema

$M_k(\boldsymbol{x}) = 0$. Es decir, se encuentra la raíz de este modelo, la que se define como $\boldsymbol{x}_{k+1}$. Así, una iteración del método de Newton está dada por

$$\begin{aligned} F'(\boldsymbol{x}_k)\boldsymbol{s}_k &= -F(\boldsymbol{x}_k), \\ \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \boldsymbol{s}_k. \end{aligned} \tag{2}$$

En cada iteración el método de Newton debe resolver un sistema de ecuaciones lineales con matriz de coeficientes $F'(\boldsymbol{x}_k)$, lo cual, desde el punto de vista computacional, es costoso [1].

Una alternativa al costo computacional que implica emplear la *matriz jacobiana* en cada iteración la representan los *métodos cuasi-Newton*, que usan una aproximación a la *matriz jacobiana* en lugar de ella misma. El objetivo con el que fueron creados estos métodos fue el de mantener las "bondades" del método de Newton pero con un menor costo computacional. En el desarrollo de estos métodos, en cada iteración, se considera un modelo de $F(\boldsymbol{x})$ de la forma $\hat{M}_k(\boldsymbol{x}) \equiv F(\boldsymbol{x}_k) + B_k(\boldsymbol{x} - \boldsymbol{x}_k)$, donde B_k es una matrix que aproxima a $F'(\boldsymbol{x}_k)$, y se resuelve el problema $\hat{M}_k(\boldsymbol{x}) = 0$. Esto es, se encuentra la raíz de este modelo, la que se define como $\boldsymbol{x}_{k+1}$. Una iteración de estos métodos está dada por [1, 4, 18]

$$\begin{aligned} B_k\boldsymbol{s}_k &= -F(\boldsymbol{x}_k), \\ \boldsymbol{x}_{k+1} &= \boldsymbol{x}_k + \boldsymbol{s}_k. \end{aligned}$$

El uso de una aproximación a la *matriz jacobiana* en lugar de ella misma se refleja en la disminución de la tasa de convergencia de estos métodos, la cual, en el mejor de los casos, es *superlineal,* en contraste con la del método de Newton que, en el mejor de los casos es *cuadrática* [4].

En algunos casos, la matriz B_k se encuentra interpolando el modelo $\hat{M}$ en $\boldsymbol{x}_{k+1}$; es decir, haciendo que $F(\boldsymbol{x}_{k+1})$ coincida con el modelo $\hat{M}(\boldsymbol{x}_{k+1})$, con lo cual se llega a la llamada *ecuación secante* [15]

$$B_{k+1}\boldsymbol{s}_k = \boldsymbol{y}_k, \tag{3}$$

donde $\boldsymbol{s}_k = \boldsymbol{x}_{k+1} - \boldsymbol{x}_k$, $\boldsymbol{y}_k = F(\boldsymbol{x}_{k+1}) - F(\boldsymbol{x}_k)$.

Si en el problema (1) la función F es una función no lineal, de *variable* y *valor matriciales* y *Fréchet diferenciable* $F\colon \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$ [8, 16], entonces el problema (1) se transforma en encontrar una matriz $X_* \in \mathbb{C}^{n\times n}$ tal que

$$F(X_*) = \mathbf{O}, \tag{4}$$

donde $\mathbf{O}$ denota la matriz cero de $\mathbb{C}^{n\times n}$. Dado que F es *Fréchet diferenciable* en X, existe una aplicación lineal y continua

$$\begin{aligned} L\colon\ \mathbb{C}^{n\times n} &\longrightarrow \mathbb{C}^{n\times n} \\ S\quad &\longmapsto L(X,S), \end{aligned}$$

tal que para todo $S \in \mathbb{C}^{n\times n}$,

$$F(X+S) = F(X) + L(X,S) + R(S), \tag{5}$$

con

$$\lim_{\|S\|\to 0} \frac{\|R(S)\|}{\|S\|} = 0, \tag{6}$$

lo cual significa que, para todo $\epsilon > 0$ existe un $\delta > 0$ tal que si $\|S\| < \delta$, entonces $\|R(S)\| < \epsilon \|S\|$ [15, 16].

Con el fin de identificar el operador derivada L en un caso particular, consideramos la *función cuadrática matricial,* la cual surge en numerosas aplicaciones [16], definida[1] por:

$$Q(X) = AX^2 + BX + C, \tag{7}$$

donde A, B y $C \in \mathbb{C}^{n\times n}$. Usando (5), tenemos que

$$Q(X+S) = A(X+S)^2 + B(X+S) + C = Q(X) + (ASX + (AX+B)S) + AS^2, \tag{8}$$

luego $L(X,S) = ASX + (AX+B)S$. En este caso, la expresión para $L(X,S)$ está relacionada con la ecuación del tipo $AXB - CXD = E$, conocida como la *ecuación de Silvester generalizada,* para la cual existen varios métodos de solución [6].

Por otro lado, volviendo al problema (4) tenemos que, para el caso matricial, el *método de Newton* es el más popular para resolver un sistema matricial. Él surge de manera natural al considerar la aproximación de *Taylor* de F alrededor de X dada por (5). Así, una iteración del método de Newton para resolver (4) está dada por

$$\begin{aligned} L(X_k, S_k) &= -F(X_k), \\ X_{k+1} &= X_k + S_k. \end{aligned} \tag{9}$$

Si denotamos $L(X_k, S_k) \equiv F'(X_k)S_k$, donde $F'(X_k) \in \mathbb{C}^{n^2\times n^2}$, y reescribimos las matrices X y S como vectores en $\mathbb{C}^{n^2}$, tenemos que la iteración (9) la podemos expresar de la forma

$$\begin{aligned} F'(X_k)S_k &= -F(X_k), \\ X_{k+1} &= X_k + S_k. \end{aligned} \tag{10}$$

Observamos que (10) está bien definida y nos permite identificar la aplicación de $F'(X)$ sobre S, que es lo que se necesita en una iteración de Newton para el problema (4). Así, una forma de abordar este problema mediante un *método cuasi-Newton* es usando matrices de tamaño $n^2 \times n^2$ para aproximar a $F'(X)$. Un problema que podríamos tener en este proceso es que no dispongamos de la forma explícita de la matriz $F'(X_k)$, con lo cual no sería posible encontrar una matriz que la aproxime. Por otro lado, si disponemos de la *matriz jacobiana*, el costo computacional al realizar dicha aproximación y resolver el sistema no lineal sería muy alto.

Históricamente, siempre que el *método de Newton* se ha propuesto para resolver un problema específico, también un *método secante* se ha podido desarrollar para el mismo problema, lo que sirvió como motivación a los autores en [15] para proponer uno de tales

[1] Esta no es la única forma de definirla; otras formas son $Q_1(X) = X^2A + XB + C$ y $Q_2(X) = XAX + BX + XC + D$, con $D \in \mathbb{C}^{n\times n}$; esta última conocida como la *función algebraica de Riccati* [2, 12].

métodos para el caso matricial. Su propuesta consiste en un método secante clásico con características interesantes, entre las que se destaca el uso de matrices de tamaño $n \times n$, lo que reduce significativamente el costo computacional asociado con el álgebra lineal del algoritmo.

3. Algoritmo y teoría de convergencia

En esta sección consideramos un algoritmo tipo *secante* propuesto en [15, 16] para resolver el problema (4). Demostramos que este algoritmo secante es *de cambio mínimo* y desarrollamos *la teoría general de convergencia local* para el mismo.

El análisis de convergencia local del *método secante* lo realizaremos siguiendo las reglas generales para convergencia de métodos secantes clásicos [4, 13]. Bajo hipótesis estándar demostramos que las matrices que aproximan a $L(X_*, S)$ se deterioran pero en forma controlada, con lo cual demostramos inicialmente la convergencia lineal del algoritmo, y posteriormente demostramos su convergencia local superlineal.

La iteración de un método secante general para resolver (4) puede expresarse como

$$X_{k+1} \; = \; X_k - A_k^{-1} F(X_k), \tag{11}$$

donde la actualización de A_k, denotada por $A_{k+1} \in \mathbb{C}^{n \times n}$, es una matriz que satisface la llamada *ecuación secante matricial*

$$A_{k+1} S_k = Y_k, \tag{12}$$

con $S_k = X_{k+1} - X_k$ y $Y_k = F(X_{k+1}) - F(X_k)$.

Es importante resaltar que la matriz $A_k \in \mathbb{C}^{n \times n}$ no es una aproximación de $F'(X_k)$, como se hace de manera natural en caso de *funciones de variable y valor vectorial.* Dado que, $A_k S_k$ y $F'(X_k) S_k$ son matrices de tamaño $n \times n$, los autores en [16] y [15] aproximan la matriz $F'(X_k) S_k$ por $A_k S_k$, lo cual constituye el aspecto novedoso de su propuesta.

Teniendo en cuenta que la iteración del *método de Newton* dada por la Ecuación (9) es equivalente a (10), y a partir de la forma general del operador L dada en (5), utilizaremos $L(X_k, S_k)$ en lugar de $F'(X_k) S_k$. Así, consideramos la iteración (11) con el supuesto de que en cada iteración $A_k S_k$ aproxima a $L(X_k, S_k)$, con lo cual, obtenemos el siguiente *algoritmo secante* para resolver el problema (4).

Algoritmo 3.1 (Algoritmo secante). Dadas las matrices iniciales X_{-1} y X_0, definimos $S_{-1} = X_0 - X_{-1}$, $Y_{-1} = F(X_0) - F(X_{-1})$ y A_0 como la solución del sistema matricial $A_0 S_{-1} = Y_{-1}$. Para $k = 0, 1, \ldots,$ las matrices X_{k+1}, S_k, Y_k y A_{k+1} son generadas como sigue:

$$\begin{aligned} A_k S_k &= -F(X_k), && (13)\\ X_{k+1} &= X_k + S_k, && (14)\\ Y_k &= F(X_{k+1}) - F(X_k), &&\\ A_{k+1} S_k &= Y_k. && (15) \end{aligned}$$

El Algoritmo 3.1 calcula la matriz S_k resolviendo el sistema de ecuaciones lineales dado por la ecuación (14). Observemos que, en cada iteración, la actualización de la matriz A_k, denotada A_{k+1}, debe satisfacer *la ecuación secante matricial,* caso en el cual la matriz S_k debe ser no singular. Además, dicha actualización se calcula en forma única mediante la Ecuación (16). Así, en cada iteración el conjunto de todas las matrices que satisfacen *la ecuación secante* (que en el caso vectorial es infinito $(n > 1)$)

$$V = V(X_k,\ X_{k+1}) = \left\{B \in \mathbb{C}^{n\times n} : B\,(X_{k+1} - X_k) = F(X_{k+1}) - F(X_k)\right\}, \tag{16}$$

es **unitario**: su único elemento es la matriz $A_{k+1} = Y_k S_k^{-1}$, como lo garantiza el siguiente teorema.

Teorema 3.2. *Si para todo* $k = 0, 1, 2, \ldots$, *la matriz* $X_{k+1} - X_k$ *es no singular, entonces el conjunto* V *definido por* (16) *es unitario.*

Demostración. Consideremos las matrices $S_k = X_{k+1} - X_k$ y $Y_k = F(X_{k+1}) - F(X_k)$ para todo $k = 0, 1, 2, \ldots$. Dado que la matriz S_k es no singular, entonces existe la matriz $Y_k\, S_k^{-1}$ para todo $k = 0, 1, 2, \ldots$; es decir, V es un conjunto no vacío.

Sean B_1 y B_2 matrices que pertenecen al conjunto V; entonces tenemos que $B_1 S_k = Y_k$ y $B_2 S_k = Y_k$, de ahí que

$$B_1 S_k = B_2 S_k, \qquad k = 0, 1, 2, \ldots; \tag{17}$$

como la matriz S_k es no singular, entonces multiplicando la ecuación (17) por S_k^{-1} se obtiene que $B_1 = B_2$. Por lo tanto, el conjunto V es unitario. ☑

En otras palabras, la matriz de *cambio mínimo* es la única solución de la *ecuación secante matricial* (16). Por lo tanto el Algoritmo 3.1 es *secante de cambio mínimo* [13].

A continuación presentamos las hipótesis locales bajo las cuales desarrollamos la teoría de convergencia local del Algoritmo 3.1. Estas hipótesis son análogas a las utilizadas para demostrar convergencia de métodos secantes clásicos en el caso vectorial [4].

3.1. Hipótesis

H1. $F\colon \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$ es ***Fréchet diferenciable*** en un conjunto $D \subset \mathbb{C}^{n\times n}$ abierto y convexo.

H2. Existe $X_* \in D$ tal que $F(X_*) = \mathbf{O}$.

H3. F' es *localmente Lipschitz continua* en un entorno de X_*.

H4. La matriz $L(X_*, S)$ es no singular y existe $\beta > 0$ tal que $\left\|L(X_*, S)^{-1}\right\| \leq \beta$.

3.2. Resultados de convergencia

A manera de preliminar, presentamos un resultado del álgebra lineal numérica que usaremos en la demostración del Teorema 3.6.

Lema 3.3 (***Lema de Banach***). *Sean* $\|\cdot\|$ *una norma matricial inducida en* $\mathbb{C}^{n\times n}$ *y* $A, B \in \mathbb{C}^{n\times n}$. *Si* A *es no singular y* $\left\|\mathrm{I}_n - A^{-1}B\right\| < 1$, *entonces* B *es no singular y*

$$\left\|B^{-1}\right\| \leq \frac{\left\|A^{-1}\right\|}{1 - \left\|\mathrm{I}_n - A^{-1}B\right\|}.$$

Demostración. Ver [4]. ☑

A continuación, presentamos algunos resultados útiles para probar la convergencia local y superlineal del Algoritmo 3.1.

Lema 3.4. *Sea* $F\colon \mathbb{C}^{n\times n} \longrightarrow \mathbb{C}^{n\times n}$ *una función que satisface las hipótesis* **H1** *a* **H3**; *entonces la aplicación* L *satisface la desigualdad*

$$\|L(X,S) - L(X_*,S)\| \leq \gamma \|X - X_*\| \|S\|.$$

Demostración. Si $X \in \mathcal{N}(X_*, r)$, $S \in \mathbb{C}^{n\times n}$ y $\|\cdot\|$ es una norma matricial, de la definición de la aplicación L dada por (5), teniendo en cuenta que $L(X,S) = F'(X)S$ y de la hipótesis **H3,** se obtiene

$$\begin{aligned} \|L(X,S) - L(X_*,S)\| &= \|F'(X)S - F'(X_*)S\| \leq \|F'(X) - F'(X_*)\| \|S\| \\ &\leq \gamma \|X - X_*\| \|S\|. \end{aligned}$$

☑

El siguiente lema garantiza que si las aproximaciones a $L(X_*,S)$ empeoran, entonces esto ocurre en forma controlada. Este resultado es análogo al de *deteriorización controlada* que ocurre en el caso vectorial [4].

Lema 3.5. *Supongamos que las hipótesis* **H1** *a* **H3** *se verifican. Sean* $A_+ \in \mathbb{C}^{n\times n}$ *la única matriz en el conjunto* $V(X,Y)$ *y* $A \in \mathbb{C}^{n\times n}$ *tal que* $AS = -F(X)$, *con* $S = Y - X$. *Entonces existe una constante positiva* c *tal que*

$$\|A_+S - L(X_*,S)\| \leq \|AS - L(X_*,S)\| + c \|Y - X_*\|. \tag{18}$$

Demostración. Usando las hipótesis **H2**, $A_+ \in V(X,Y)$, $AS = -F(X)$ y la desigualdad triangular, obtenemos

$$\begin{aligned} \|A_+S - L(X_*,S)\| &= \|A_+(Y-X) - L(X_*,Y-X)\| = \|F(Y) - F(X) - L(X_*,Y-X)\| \\ &= \|F(Y) + A(Y-X) - L(X_*,Y-X)\| \\ &\leq \|F(Y) - F(X_*)\| + \|A(Y-X) - L(X_*,Y-X)\|. \end{aligned} \tag{19}$$

- Por la hipótesis **H1**, tenemos que

$$F(Y) = F(X_*) + L(X_*, Y - X_*) + R(Y - X_*), \tag{20}$$

donde

$$\lim_{\|Y-X_*\|\to 0} \frac{\|R(Y - X_*)\|}{\|Y - X_*\|} = 0, \tag{21}$$

es decir, para $\rho > 0$ existe $\epsilon > 0$ tal que si $\|Y - X_*\| < \epsilon$, entonces

$$\|R(Y - X_*)\| < \rho \|Y - X_*\|. \tag{22}$$

- La aplicación L es acotada, ya que su dominio es de dimensión finita [11]; por lo tanto, existe $k_1 > 0$ tal que

$$\|L(X_*, Y - X_*)\| \leq \ k_1 \, \|Y - X_*\| \, . \tag{23}$$

Usando (20), (22) y (23) en (19), obtenemos

$$\begin{aligned} \|A_+ S - L(X_*, S)\| &\leq \|L(X_*, Y - X_*)\| + \|R(Y - X_*)\| + \|A(Y - X) - L(X_*, Y - X)\| \\ &\leq c \, \|Y - X_*\| + \|A(Y - X) - L(X_*, Y - X)\| \, , \end{aligned} \tag{24}$$

donde $c = k_1 + \rho$. Por lo tanto, existe una constante positiva c tal que

$$\|A_+ S - L(X_*, S)\| \ \leq \ \|AS - L(X_*, S)\| + c \, \|Y - X_*\| \, . \qquad \boxtimes$$

El lema anterior es útil para demostrar que la sucesión $\{X_k\}$ generada por el Algoritmo 3.1 converge local y linealmente a X_*, lo cual se garantiza en el siguiente teorema.

Teorema 3.6. *Supongamos las hipótesis* **H1** *a* **H4**. *En esas condiciones existen constantes positivas* ϵ *y* δ *tales que si* $\|X_0 - X_*\| \leq \epsilon$ *y* $\|A_0 S_0 - L(X_*, S_0)\| \leq \delta$, *entonces la sucesión* $\{X_k\}$ *generada por el Algoritmo 3.1 está bien definida y converge linealmente a* X_*.

Demostración. Sean $\|\cdot\|$ una norma matricial, β la constante dada por la hipótesis **H4,** c la constante dada en (18), $L_* = L(X_*, S_k)$, $E_k = X_k - X_*$, y sean $\overline{\epsilon}$ y δ tales que

$$12 \, \beta \, \delta \ \leq \ 1, \tag{25}$$

$$2 \, c \, \overline{\epsilon} \ \leq \ \delta. \tag{26}$$

Para la escogencia de ϵ hacemos las siguientes consideraciones. Para $i \in \mathbb{N}$,

- De (6), tenemos que para $\rho_0 < \dfrac{\delta}{\|E_0\|}$, existe $\epsilon_0 > 0$ tal que si $\|S_i\| < \epsilon_0$,

$$\|R(S_i)\| \ < \ \rho_0 \, \|S_i\| \, . \tag{27}$$

- Por la *continuidad* de la función F, para $\rho_1 < \delta$, existe $\epsilon_1 > 0$ tal que si $\|S_i\| < \epsilon_1$,

$$\|F(X_* + S_i) - F(X_*)\| \ \leq \ \rho_1. \tag{28}$$

  En forma análoga, para $\rho_2 < \delta$, existe $\epsilon_2 > 0$ tal que si $\|E_i\| < \epsilon_2$,

$$\|F(X_i) - F(X_*)\| \ \leq \ \rho_2. \tag{29}$$

Sea

$$\epsilon \ = \text{mín} \left\{ \overline{\epsilon}, \, \epsilon_0, \, \epsilon_1, \, \epsilon_2 \right\} . \tag{30}$$

Para demostrar que la sucesión $\{X_k\}$ generada por el Algoritmo 3.1 converge linealmente a X_*, demostraremos por inducción que para todo $k = 0, 1, 2, \ldots$

$$\|A_k S_k - L(X_*, S_k)\| \ \leq \ \left(2 - 2^{-k}\right) \delta, \tag{31}$$

$$\|X_{k+1} - X_*\| \ \leq \ \frac{1}{2} \, \|X_k - X_*\| \, . \tag{32}$$

1. Para $k = 0$, la desigualdad (31) es inmediata, ya que
$$\|A_0 S_0 - L(X_*, S_0)\| \;\leq\; \delta \;=\; \left(2 - 2^0\right)\delta. \tag{33}$$
La demostración de (32) es idéntica a la prueba en el *paso de inducción,* por lo cual la omitimos aquí.

2. *Hipótesis inductivas.* Supongamos que las desigualdades (31) y (32) se cumplen para $k = 0, 1, \ldots, i-1$.

3. *Paso de inducción.* Demostremos que (31) y (32) se cumplen para $k = i$.
De la propiedad de *deteriorización* (18) y de las *hipótesis inductivas,* tenemos que
$$\|A_i S_i - L(X_*, S_i)\| \;\leq\; \left(2 - 2^{-(i-1)}\right)\delta + c\,\|X_{i-1} - X_*\|\,, \tag{34}$$
$$\|X_{i-1} - X_*\| \;\leq\; \frac{1}{2}\,\|X_{i-2} - X_*\| \;\leq\; \left(\frac{1}{2}\right)^{i-1}\|X_0 - X_*\| \;\leq\; 2^{-(i-1)}\,\epsilon. \tag{35}$$
Sustituyendo (35) en (34), usando (30) y (26), tenemos
$$\begin{aligned}\|A_i S_i - L(X_*, S_i)\| &\leq \left(2 - 2^{-(i-1)}\right)\delta + c\,2^{-(i-1)}\,\epsilon \leq \left(2 - 2^{-(i-1)}\right)\delta + c\,2^{-(i-1)}\,\overline{\epsilon}\\ &= \left(2 - 2^{-(i-1)}\right)\delta + \,2^{-i}\,2\,c\,\overline{\epsilon} \;\leq\; \left(2 - 2^{-(i-1)} + 2^{-i}\right)\delta\\ &= \left(2 - 2^{-i}\right)\delta.\end{aligned}$$
Por lo tanto,
$$\|A_i S_i - L(X_*, S_i)\| \;\leq\; \left(2 - 2^{-i}\right)\delta. \tag{36}$$
Para demostrar (32) debemos mostrar primero que la matriz $A_k S_k$ es no singular, con lo cual el Algoritmo 3.1 está bien definido. Para ello, por la hipótesis **H4,** (25) y (36), tenemos
$$\begin{aligned}\left\|\mathrm{I}_n - L(X_*, S_i)^{-1}A_i S_i\right\| \;&=\; \left\|L(X_*, S_i)^{-1}L(X_*, S_i) - L(X_*, S_i)^{-1}A_i S_i\right\|\\ &\leq\; \left\|L(X_*, S_i)^{-1}\right\|\,\|L(X_*, S_i) - A_i S_i\|\\ &\leq\; \beta\left(2 - 2^{-i}\right)\delta \;\leq\; 2\,\beta\delta \;\leq\; \frac{1}{6}\,\cdot\end{aligned} \tag{37}$$
Por el Lema de Banach, concluimos que la matriz $A_i S_i$ es no singular, y por lo tanto las matrices A_i y S_i son no singulares. Además,
$$\left\|(A_i S_i)^{-1}\right\| \leq \quad \frac{\beta}{1 - \frac{1}{6}} \;=\; \frac{6}{5}\,\beta. \tag{38}$$
Así, X_{i+1} está bien definido. Sumando a ambos lados de la ecuación (38) la matriz $-X_*$, sustituyendo S_i por $-A_i^{-1}F(X_i)$ y mediante algunas operaciones algebraicas, obtenemos,
$$\begin{aligned}E_{i+1} &= X_i - X_* - A_i^{-1}F(X_i) = \mathrm{I}_n E_i - A_i^{-1}F(X_i)\\ &= (A_i S_i)^{-1}A_i S_i(E_i) - (A_i S_i)^{-1}L_* E_i + (A_i S_i)^{-1}L_* E_i\\ &\qquad - (A_i S_i)^{-1}F(X_* + S_i)E_i + (A_i S_i)^{-1}F(X_* + S_i)E_i - A_i^{-1}F(X_i).\end{aligned} \tag{39}$$

En la igualdad anterior aplicamos una norma matricial y algunas de sus propiedades, sumamos $F(X_*) = \mathbf{O}$, usamos la desigualdad: $\|S_i\| \leq \|E_{i+1}\| + \|E_i\|$ y obtenemos la siguiente desigualdad:

$$\begin{aligned}
\|E_{i+1}\| &\leq \left\|(A_iS_i)^{-1}\right\| \left[\, \|A_iS_i - L_*\| + \|L_* - F(X_* + S_i) + F(X_*)\| \,\right] \|E_i\| \\
&\qquad + \left\|(A_iS_i)^{-1} F(X_* + S_i)E_i - S_iS_i^{-1}A_i^{-1}F(X_i)\right\| \\
&\leq \left\|(A_iS_i)^{-1}\right\| \left[\|A_iS_i - L_*\| + \|R(S_i)\| + \|F(X_* + S_i) - F(X_*)\|\right] \|E_i\| \\
&\qquad + \left\|(A_iS_i)^{-1}\right\| \|F(X_i) - F(X_*)\| \|S_i\| \qquad (40) \\
&\leq \left\|(A_iS_i)^{-1}\right\| \left[\|A_iS_i - L_*\| + \|R(S_i)\| + \|F(X_* + S_i) - F(X_*)\|\right] \|E_i\| \\
&\quad + \left\|(A_iS_i)^{-1}\right\| \|F(X_i) - F(X_*)\| \|E_i\| + \left\|(A_iS_i)^{-1}\right\| \|F(X_i) - F(X_*)\| \|E_{i+1}\| .
\end{aligned}$$

Usando (27), (28), (29), (30), las hipótesis (31) y (32) en (40) tenemos:

$$\begin{aligned}
\|E_{i+1}\| &\leq \frac{6}{5}\beta\left[\left(2 - 2^{-i}\right)\delta + \rho_0 \|S_i\| + \rho_1 + \rho_2\right] \|E_i\| + \frac{6}{5}\beta\,\rho_2 \|E_{i+1}\| \\
&\leq \frac{6}{5}\beta\left[2\,\delta - 2^{-i}\delta + \rho_0 \|E_i\| + \rho_1 + \rho_2\right] \|E_i\| + \frac{6}{5}\beta\left[\rho_0 \|E_i\| + \rho_2\right] \|E_{i+1}\| \\
&\leq \frac{6}{5}\beta\left[2\delta - 2^{-i}\delta + \rho_0\, 2^{-i} \|E_0\| + \rho_1 + \rho_2\right] \|E_i\| + \frac{6}{5}\beta\left[\rho_0\, 2^{-i} \|E_0\| + \rho_2\right] \|E_{i+1}\| \\
&\leq \frac{6}{5}\beta\,\delta\,[2 + 1 + 1] \|E_i\| + \frac{6}{5}\beta\,\delta\,[2] \|E_{i+1}\| \;\leq\; \frac{2}{5}\|E_i\| + \frac{1}{5}\|E_{i+1}\| \qquad (41)
\end{aligned}$$

haciendo operaciones algebraicas en la desigualdad (41), se concluye que $\|E_{i+1}\| \leq \frac{1}{2}\|E_i\|$, y equivalentemente, $\|X_{i+1} - X_*\| \;\leq\; \dfrac{1}{2}\|X_i - X_*\|$.

Con lo cual la sucesión $\{X_k\}$ converge linealmente a X_*. ☑

A continuación, presentamos un teorema análogo al Teorema condición de Dennis-Moré [4], el cual da una condición suficiente para convergencia local superlineal de un algoritmo cuasi-Newton.

Teorema 3.7 (Condición tipo Dennis-Moré). *Sea* $D \in \mathbb{C}^{n\times n}$ *un conjunto abierto y convexo para el cual son válidas las hipótesis* **H1** *a* **H4.** *Sea* $\{A_k\}$ *la sucesión de matrices generadas por* (16), *y supongamos que para alguna matriz* $X_0 \in D$ *la sucesión de matrices* $\{X_k\}$ *generada por* (15) *satisface que* $\lim_{k\to\infty} X_k = X_*$. *Si*

$$\lim_{k\to\infty} \frac{\|A_kS_k - L(X_*, S_k)\|}{\|S_k\|} = 0, \qquad (42)$$

donde $S_k = X_{k+1} - X_k$, *entonces la sucesión* $\{X_k\}$ *converge superlinealmente a* X_*.

Demostración. Denotemos $L_* = L(X_*, S_k)$, $L_k = L(X_k, S_k)$ y $E_k = X_k - X_*$. De (14) tenemos que $\mathbf{O} = A_k S_k + F(X_k) = A_k S_k - L_* + F(X_k) + L_*$; sumando $-F(X_{k+1})$ en ambos lados de la igualdad anterior, obtenemos

$$-F(X_{k+1}) = A_k S_k - L_* + [-F(X_{k+1}) + F(X_k) + L_*]\,; \tag{43}$$

aplicando la desigualdad triangular, algunas operaciones algebraicas y el Lema 3.4, obtenemos

$$\begin{aligned}
\frac{\|F(X_{k+1})\|}{\|S_k\|} &\leq \frac{\|A_k S_k - L_*\|}{\|S_k\|} + \frac{\|-F(X_{k+1}) + F(X_k) + L_*\|}{\|S_k\|} \\
&= \frac{\|A_k S_k - L_*\|}{\|S_k\|} + \frac{\|-L_k + L_* - R(S_k)\|}{\|S_k\|} \\
&\leq \frac{\|A_k S_k - L_*\|}{\|S_k\|} + \frac{\|L_* - L_k\| + \|R(S_k)\|}{\|S_k\|} \\
&\leq \frac{\|A_k S_k - L_*\|}{\|S_k\|} + \gamma\,\|E_k\| + \frac{\|R(S_k)\|}{\|S_k\|}.
\end{aligned} \tag{44}$$

Analizando el lado de derecho de la desigualdad (44) tenemos que, por (42), el primer término tiende a cero. Por hipótesis, $\lim\limits_{k\to\infty} X_k = X_*$, entonces el segundo término tiende a cero y, por (6), el último término también tiende a cero, con lo cual

$$\lim_{k\to\infty} \frac{\|F(X_{k+1})\|}{\|S_k\|} \;=\; 0. \tag{45}$$

Por la continuidad de la función F tenemos que $\lim\limits_{k\to\infty} F(X_{k+1}) = F(X_*) = \mathbf{O}$.

Después de realizar algunas operaciones algebraicas, usando propiedades de la norma matricial y de (5), tenemos

$$\begin{aligned}
\|F(X_{k+1})\| &= \|F(X_{k+1}) - F(X_*)\| = \|L(X_*, E_{k+1}) + F(X_{k+1}) - F(X_*) - L(X_*, E_{k+1})\| \\
&\geq \|L(X_*, E_{k+1})\| - \|F(X_{k+1}) - F(X_*) - L(X_*, E_{k+1})\| \\
&\geq \frac{1}{\|L(X_*, E_{k+1})^{-1}\|} - \|R(E_{k+1})\|\,.
\end{aligned} \tag{46}$$

Por la hipótesis **H4** tenemos que $\left\|L(X_*, E_{k+1})^{-1}\right\| < \beta$. Por otro lado, en el Teorema 3.6 demostramos que la sucesión $\{X_k\}$ satisface que $\|E_{k+1}\| \leq \frac{1}{2}\|E_k\|$ para todo $k = 0, 1, \ldots$. Así que

$$\|E_{k+1}\| \leq \|E_k\| \leq \cdots \leq \|E_0\|\,. \tag{47}$$

Equivalentemente,

$$\frac{1}{\|E_{k+1}\|} \geq \frac{1}{\|E_k\|} \geq \cdots \geq \frac{1}{\|E_0\|}, \tag{48}$$

y de otro lado, por (6), tenemos que, para $\rho < \dfrac{1}{\beta\,\|E_0\|}$, existe $\epsilon > 0$ tal que si $\|E_{k+1}\| < \epsilon$, entonces $\|R(E_{k+1})\| < \rho\,\|E_{k+1}\|$. Así, usando (48) en (46), tenemos

$$\begin{aligned} \|F(X_{k+1})\| \;&\geq\; \frac{1}{\beta} - \rho\,\|E_{k+1}\| \;=\; \left[\frac{1}{\beta\,\|E_{k+1}\|} - \rho\right]\|E_{k+1}\| \\ &\geq\; \left[\frac{1}{\beta\,\|E_0\|} - \rho\right]\|E_{k+1}\| \;=\; \alpha\,\|E_{k+1}\|, \end{aligned} \tag{49}$$

donde $\alpha = \left(\dfrac{1}{\beta\,\|E_0\|} - \rho\right) > 0$. Combinando (45) y (49),

$$\begin{aligned} 0 \;&=\; \lim_{k\to\infty} \frac{\|F(X_{k+1})\|}{\|S_k\|} \;\geq\; \lim_{k\to\infty} \frac{\alpha\,\|E_{k+1}\|}{\|S_k\|} \;=\; \alpha \lim_{k\to\infty} \frac{\|E_{k+1}\|}{\|E_{k+1} - E_k\|} \\ &\geq\; \alpha \lim_{k\to\infty} \frac{\|E_{k+1}\|}{\|E_{k+1}\| + \|E_k\|} \;=\; \alpha \lim_{k\to\infty} \frac{\dfrac{\|E_{k+1}\|}{\|E_k\|}}{\dfrac{\|E_{k+1}\|}{\|E_k\|} + 1}, \end{aligned}$$

lo cual implica que $\lim\limits_{k\to\infty} \dfrac{\|E_{k+1}\|}{\|E_k\|} = 0$. Por lo tanto, si (42) se satisface, entonces la sucesión $\{X_k\}$ converge superlinealmente a X_*. ☑

A continuación presentamos un resultado útil para demostrar que el Algoritmo 3.1 satisface la condición tipo Dennis-Moré dada por (42).

Lema 3.8. *Supongamos las hipótesis* **H1** *a* **H4,** *que las sucesiones* $\{X_k\}$ *y* $\{A_k\}$ *son generadas por* (15) *y* (16)*, respectivamente, y que* $\lim\limits_{k\to\infty} X_k = X_*$. *Entonces*

$$\lim_{k\to\infty} \frac{\|A_{k+1}S_k - L(X_*, S_k)\|}{\|S_k\|} = 0. \tag{50}$$

Demostración. Por el Teorema 3.2, la matriz A_{k+1} es la única matriz que satisface *la ecuación secante matricial*

$$A_{k+1}S_k = Y_k = F(X_{k+1}) - F(X_k). \tag{51}$$

Por la hipótesis **H1,**

$$F(X_{k+1}) = F(X_k) + L(X_k, S_k) + R(S_k), \tag{52}$$

donde

$$\lim_{\|S_k\|\to 0} \frac{\|R(S_k)\|}{\|S_k\|} = 0. \tag{53}$$

Usando las ecuaciones (51), (52) y (53), algunas manipulaciones algebraicas, la desigualdad triangular y el Lema 3.4, obtenemos

$$\begin{aligned} \|A_{k+1}S_k - L(X_*, S_k)\| &= \|F(X_{k+1}) - F(X_k) - L(X_*, S_k)\| \\ &= \|L(X_k, S_k) - L(X_*, S_k) + R(S_k)\| \\ &\leq \|L(X_k, S_k) - L(X_*, S_k)\| + \|R(S_k)\| \\ &\leq \gamma \|X_k - X_*\| \|S_k\| + \|R(S_k)\| ; \end{aligned} \tag{54}$$

multiplicando la desigualdad (54) por $\|S_k\|^{-1}$ tenemos

$$\frac{\|A_{k+1}S_k - L(X_*, S_k)\|}{\|S_k\|} \leq \gamma \|X_k - X_*\| + \frac{\|R(S_k)\|}{\|S_k\|}. \tag{55}$$

Dado que, $\lim_{k\to\infty} X_k = X_*$, entonces $\lim_{k\to\infty} \|X_k - X_*\| = 0$ y $\lim_{k\to\infty} \|S_k\| = 0$. Por lo tanto,

$$\lim_{k\to\infty} \frac{\|A_{k+1}S_k - L(X_*, S_k)\|}{\|S_k\|} = 0. \qquad ☑$$

Con el resultado anterior se puede derivar una condición suficiente para garantizar convergencia superlineal del Algoritmo 3.1, como se prueba en el siguiente teorema.

Teorema 3.9. *Supongamos las hipótesis* **H1** *a* **H4**, *que las sucesiones* $\{X_k\}$ *y* $\{A_k\}$ *son generadas por* (15) *y* (16), *respectivamente, y que* $\lim_{k\to\infty} X_k = X_*$. *Si*

$$\lim_{k\to\infty} \frac{\|A_kS_k - A_{k+1}S_k\|}{\|S_k\|} = 0, \tag{56}$$

entonces la sucesión $\{X_k\}$ *converge superlinealmente a* X_*.

Demostración. Adicionando $A_{k+1}S_k$ y usando la desigualdad triangular, obtenemos

$$\begin{aligned} \lim_{k\to\infty} \frac{\|A_kS_k - L(X_*, S_k)\|}{\|S_k\|} &= \lim_{k\to\infty} \frac{\|A_kS_k - A_{k+1}S_k + A_{k+1}S_k - L(X_*, S_k)\|}{\|S_k\|} \\ &\leq \lim_{k\to\infty} \frac{\|A_kS_k - A_{k+1}S_k\|}{\|S_k\|} + \lim_{k\to\infty} \frac{\|A_{k+1}S_k - L(X_*, S_k)\|}{\|S_k\|}. \end{aligned} \tag{57}$$

Por el Lema 3.8,

$$\lim_{k\to\infty} \frac{\|A_{k+1}S_k - L(X_*, S_k)\|}{\|S_k\|} = 0,$$

y el segundo término del lado derecho de la desigualdad (57) es cero por la hipótesis (56). Luego

$$\lim_{k\to\infty} \frac{\|A_kS_k - L(X_*, S_k)\|}{\|S_k\|} = 0, \tag{58}$$

esta es la condición suficiente del Teorema 3.7. Por lo tanto, la sucesión $\{X_k\}$ generada por el Algoritmo 3.1 converge superlinealmente a X_*. ☑

Con los resultados demostrados en esta sección tenemos que bajo las hipótesis **H1** a **H4**, la sucesión $\{X_k\}$ generada por el Algoritmo 3.1 converge localmente y superlinealmente a X_*, lo cual resumimos en el siguiente teorema.

Teorema 3.10. *Supongamos las hipótesis* **H1** *a* **H4** *y que la sucesión* $\{A_k\}$ *está definida por* (16). *En estas condiciones, existen constantes positivas* ϵ *y* δ *tales que si* $\|X_0 - X_*\| \leq \epsilon$ *y* $\|A_0 S_0 - L(X_*, S_0)\| \leq \delta$, *entonces la sucesión* $\{X_k\}$ *generada por*

$$X_{k+1} = X_k - A_k^{-1} F(X_k)$$

está bien definida y converge linealmente a X_*. *Además, si*

$$\lim_{k \to \infty} \frac{\|A_k S_k - A_{k+1} S_k\|}{\|S_k\|} = 0,$$

entonces la sucesión $\{X_k\}$ *converge superlinealmente a* X_*.

Demostración. Es una aplicación directa de los Teoremas 3.6, 3.7 y 3.9. ☑

4. Pruebas numéricas

Como complemento al estudio teórico del Algoritmo 3.1 realizado en la Sección **3**, presentamos a continuación la solución de una *ecuación cuadrática matricial* utilizando el Algoritmo 3.1. Este tipo de problemas aparece con frecuencia en el contexto del *problema cuadrático de los valores propios* [9, 10, 19].

Para mayor claridad en la lectura de este documento, incluimos a continuación la estructura general del algoritmo.

Algoritmo 4.1. Dadas las matrices X_{-1}, $X_0 \in \mathbb{C}^{n\times n}$.

P.1: Calculamos S_{-1}.

P.2: Encontramos la solución W_0 del sistema $W_0 S_{-1} = A\left(X_0^2 - X_{-1}^2\right)$.

P.3: Calculamos $A_0 = W_0 + B$.

P.4: Para $k = 0, 1, \ldots,$

Mientras $Res(X_k) \leq n * eps$ y $k < N$

P.5: Resolvemos $A_k S_k = -F(X_k)$ para S_k.

P.6: Actualizamos X_k usando $X_{k+1} = X_k + S_k$.

P.7: Encontramos la solución W_{k+1} del sistema $W_{k+1} S_k = A\left(X_{k+1}^2 - X_k^2\right)$.

P.8: Actualizamos A_k utilizando $A_{k+1} = W_{k+1} + B$.

P.9 $k \leftarrow k + 1$

Fin.

En el paso **P.2**, encontramos la matriz W_0 que resulta de lo siguiente:

$$A_0 S_{-1} = F(X_0) - F(X_{-1}) = A(X_0^2 - X_{-1}^2) + B(X_0 - X_{-1});$$

pero $S_{-1} = X_0 - X_{-1}$, entonces $A_0 = A(X_0^2 - X_{-1}^2)(S_{-1})^{-1} + B$. Así, definiendo la matriz $W_0 = A(X_0^2 - X_{-1}^2)(S_{-1})^{-1}$ tenemos que $W_0 S_{-1} = A\left(X_0^2 - X_{-1}^2\right)$, y por lo anterior, en el paso **P.3** definimos $A_0 = W_0 + B$. En forma análoga se obtienen las variables W_{k+1} y A_{k+1}, respectivamente, en los pasos **P.7** y **P.8**.

Escribimos los códigos del algoritmo y de las funciones de prueba en MATLAB® y realizamos los experimentos numéricos en un computador Intel (R) Core (TM) i5-3450 de 2.8 GHz. Consideramos las matrices iniciales y el criterio de parada utilizados en [9, 15, 19], a saber: $X_{-1} = 0.1\mathrm{I}_n$ y $X_0 = \beta \mathrm{I}_n$; adicionalmente, usamos $X_0 = 10^r \mathrm{I}_n$ para algunos valores particulares de r. En este contexto, I_n denota la matriz identidad de orden n y β es una constante dada por

$$\beta = \frac{\|B\|_F + \sqrt{\|B\|_F^2 + 4\,\|A\|_F\,\|C\|_F}}{2\,\|A\|_F},$$

donde $\|\cdot\|_F$ denota la norma matricial de *Frobenius* [9, 15]. Para el criterio de parada, definimos

$$Res(X_k) = \frac{\|Q(X_k)\|}{\|A\|_F\,\|X_k\|_F^2 + \|B\|_F\,\|X_k\|_F + \|C\|_F} \tag{59}$$

y declaramos convergencia si $Res(X_k) < n * eps$, donde eps denota el épsilon de la máquina, que en nuestro caso, corresponde a $eps = 2{,}22044604925031 \times 10^{-16}$. Declaramos divergencia si el número de iteraciones es mayor que 200.

Con el fin de analizar el desempeño numérico del Algoritmo 4.1 resolvemos dos *ecuaciones cuadráticas matriciales.*

Presentamos los resultados obtenidos en la Tabla 1, cuyas tres columnas contienen la siguiente información: la primera indica la matriz inicial utilizada (X_0); la segunda indica el *número de iteraciones* utilizado por el Algoritmo 4.1 $(No.)$ y la tercera columna hace referencia al *criterio de parada* $(Res(X_k))$.

Problema [9]. Encontrar una solución de *la ecuación cuadrática matricial,*

$$Q(X) = X^2 + X + C = \mathbf{O}, \tag{60}$$

donde

$$C = \begin{bmatrix} -2 & -1 \\ 0 & -2 \end{bmatrix}.$$

Las soluciones de la ecuación (60) son las matrices

$$X_*^1 = \begin{bmatrix} -2 & -1/3 \\ 0 & -2 \end{bmatrix} \qquad \text{y} \qquad X_*^2 = \begin{bmatrix} 1 & 1/3 \\ 0 & 1 \end{bmatrix}.$$

En la Tabla 1 observemos que para las diferentes matrices iniciales la convergencia a X_*^2 se obtiene prácticamente en el mismo número de iteraciones.

X_0	N_0.	$Res(X_k)$
βI_2	8	2,76162288449527e-017
$10I_2$	10	0,0
10^4I_2	11	2,76162288449527e-017
10^5I_2	11	8,28486865348581e-017
$10^{18}I_2$	11	3,13660064934264e-016
$10^{20}I_2$	11	3,13660064934264e-016

Cuadro 1. Resultados para el Problema usando el Algoritmo 4.1.

5. Comentarios finales

Las funciones matriciales y *las ecuaciones no lineales matriciales* surgen en numerosos contextos de la ciencia y de la ingeniería. Este gran número de aplicaciones ha motivado el desarrollo de una teoría sobre *funciones matriciales* que ayude en la solución de *ecuaciones matriciales no lineales.* Hasta hace poco tiempo, el único método numérico para resolver ecuaciones matriciales no lineales era el método de Newton. Recientemente, se propone un *método tipo secante* [16] para resolver problemas matriciales no lineales, y se presentan algunas pruebas numéricas que muestran un buen desempeño del método y dejan el camino abierto para investigar más al respecto [16].

En este artículo consideramos dicho método secante, demostramos que es un método *secante de cambio mínimo* y, bajo hipótesis estándar, desarrollamos una *teoría general de convergencia* para él. Demostramos que este método proporciona un *algoritmo local y superlinealmente convergente.* Complementamos el análisis teórico del método resolviendo numéricamente una *ecuación cuadrática matricial*.

Finalmente, en búsqueda de ampliar el espectro de trabajo con respecto a problemas matriciales no lineales, pensamos que sería conveniente incorporar estrategias de globalización al algoritmo secante y realizar pruebas numéricas del algoritmo globalizado donde se involucren funciones matriciales y aplicaciones a problemas reales.

Agradecimientos. Agradecemos a la Universidad del Cauca por el tiempo concedido para esta investigación, mediante el Proyecto de investigación VRI ID 3908, y a los evaluadores de este artículo por sus acertadas observaciones.

Referencias

[1] Acevedo R., Pérez R. y Arenas F., “El método DL para resolver sistemas de ecuaciones no lineales”, *Matemáticas: Enseñanza Universitaria* 16 (2008), 23-36.

[2] Bittani S., Laub A.J. and Willems C., *The Riccati equation*, Springer-Verlag, Berlin, 1991.

[3] Davis J.G., “Numerical solution of a quadratic matrix equation”, *SIAM J Sci and Stat. Comput.* 2 (1981), no. 2, 164-175.

[4] Dennis J.E. and Schnabel R.B., *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall, New Jersey, 1983.

[5] Dennis J.E. and Walker H.F., "Convergence theorems for least-change secant update methods", *SIAM J. Numer. Anal.* 18 (1981), no. 6, 949-987.

[6] Golub H.G., Nash S. and Van Loan C., "A Hessenberg-Schur method for the problem $AX + XB = C$", *IEEE Trans. Automat. Control* 24 (1979), no. 6, 909-913.

[7] Hashemi B. and Dehghan M., "Efficient computation of enclosures for the exact solvents of a quadratic matrix equation", *Electron. J. Linear Algebra* 20 (2010), 519-536.

[8] Higham N.J., *Functions of matrices theory and computations*, SIAM, 2008.

[9] Higham N.J. and Kim H., "Numerical analysis of a quadratic matrix equation", *IMA J. Numer. Anal.* 20 (2000), no. 4, 499-519.

[10] Higham N.J. and Kim H., "Solving a quadratic matrix equation by Newton's methods with exact line searches", *SIAM J. Matrix Anal. Appl.* 23 (2001), no. 2, 303-316.

[11] Kreyszig E., *Introductory functional analysis with applications*, Wiley & Sons, Canada, 1978.

[12] Lancaster P. and Rodman L., *Algebraic Riccati equations*, The Clarendon Press, Oxford University Press, New York, 1995.

[13] Martínez J.M., "On the relation between two local convergence theories of least-change secant updates methods", *Math. Comp.* 59 (1992), no. 200, 457-481.

[14] Martínez J.M. and Santos S.A., "Métodos computacionais de otimização", 20 Colóquio Brasileiro de Matemática, IMPA 1995, p. 87.

[15] Monsalve M. and Raydan M., "Newton's method and secant methods: A long-standing relationship from vectors to matrices", *Port. Math.* 68 (2011), no. 4, 431-475.

[16] Monsalve M. and Raydan M., "A secant method for nonlinear matrix problems", Chapter 18 *of Numerical Linear Algebra in Signals, Systems and Control,* P. Van Dooren et al. (eds.), Springer Verlag. 80, 2011, 387-412.

[17] Parks P.C., "AM Lyapunov's stability theory–100 years on", *IMA J. Math. Control Inform.* 9 (1992), no. 4, 275-303.

[18] Pérez R. y Díaz T., *Minimización sin restricciones*, Editorial Universidad del Cauca, 2010.

[19] Tisseur F. and Meerbergen K., "The quadratic eigenvalue problem", *SIAM Rev.* 43 (2001), no. 2, 235-286.