

**75.12/95.04 ANÁLISIS NUMÉRICO I
95.10 MODELACIÓN NUMÉRICA
95.13 MÉTODOS MATEMÁTICOS Y NUMÉRICOS
SISTEMAS DE ECUACIONES NO LINEALES**

Ing. Rodolfo A. Schwarz

Año 2021

Índice

① INTRODUCCIÓN

② MÉTODO DE NEWTON

③ MÉTODO DE BROYDEN

④ BIBLIOGRAFÍA

Introducción

- Supongamos ahora que tenemos el siguiente sistema de ecuaciones:

$$\begin{aligned}f_1(x_1, x_2) &= x_1^2 + x_2^2 - 1 = 0, \\f_2(x_1, x_2) &= x_2^2 - 2x_1 - x_2 + 1 = 0.\end{aligned}\tag{1}$$

- Se trata de un sistema no lineal.
- Los métodos vistos para resolver *Sistemas de Ecuaciones Lineales* no nos sirven.
- Debemos adaptar alguno de los métodos vistos.
- Un modo para resolver el sistema es hacer:

$$\begin{aligned}x_1^{(j+1)} &= \pm \sqrt{1 - x_2^{(j)2}}, \\x_2^{(j+1)} &= x_1^{(j)2} - 2x_1^{(j)} + 1.\end{aligned}\tag{2}$$

Introducción

- Al modelo anterior podemos describirlos de dos maneras:
 - Es la aplicación del *Método de las Aproximaciones Sucesivas*,
 - Es una forma de aplicar el *Método de Jacobi*.
- Si nos quedamos con la segunda descripción, podemos aplicar también el *Método de Gauss-Seidel*:

$$\begin{aligned}x_1^{\langle j+1 \rangle} &= \pm \sqrt{1 - x_2^{\langle j \rangle 2}}, \\x_2^{\langle j+1 \rangle} &= x_1^{\langle j+1 \rangle 2} - 2x_1^{\langle j+1 \rangle} + 1.\end{aligned}\tag{3}$$

- Este método debe converger más rápido.
- Pero nada nos asegura que ambos métodos sean convergentes.

Introducción

- Si nos quedamos con la primera descripción, la convergencia, en uno u otro caso, es lineal.
- ¿Cómo podríamos desarrollar un método cuya convergencia fuere mejor?
- A partir de la primera descripción del modelo, podemos analizar el problema como los hicimos para el caso de *Ecuaciones No Lineales*.
- Vimos que la de mejora del *Método de las Aproximaciones Sucesivas* es el *Método de Newton–Raphson*.
- Propongamos algo similar en esta situación.
- Para ello vamos a adaptar este método para dos o más variables.

Método de Newton

- Empecemos por desarrollar las dos funciones mediante series de Taylor, en este caso para dos variables:

$$\begin{aligned}
 f_1(\hat{x}_1, \hat{x}_2) &= f_1(x_1^{(j)}, x_2^{(j)}) + \left. \frac{\partial f_1(x_1, x_2)}{\partial x_1} (\hat{x}_1 - x_1^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \\
 &\quad + \left. \frac{\partial f_1(x_1, x_2)}{\partial x_2} (\hat{x}_2 - x_2^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \dots, \\
 f_2(\hat{x}_1, \hat{x}_2) &= f_2(x_1^{(j)}, x_2^{(j)}) + \left. \frac{\partial f_2(x_1, x_2)}{\partial x_1} (\hat{x}_1 - x_1^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \\
 &\quad + \left. \frac{\partial f_2(x_1, x_2)}{\partial x_2} (\hat{x}_2 - x_2^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \dots.
 \end{aligned} \tag{4}$$

- En este caso, \hat{x}_1 y \hat{x}_2 representan a las soluciones «exactas».

Método de Newton

- Entonces tenemos:

$$\begin{aligned}f_1(\hat{x}_1, \hat{x}_2) &= 0, \\f_2(\hat{x}_1, \hat{x}_2) &= 0.\end{aligned}\tag{5}$$

y nos queda:

$$\begin{aligned}0 &= f_1(x_1^{(j)}, x_2^{(j)}) + \left. \frac{\partial f_1(x_1, x_2)}{\partial x_1} (\hat{x}_1 - x_1^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \\&\quad + \left. \frac{\partial f_1(x_1, x_2)}{\partial x_2} (\hat{x}_2 - x_2^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \dots, \\0 &= f_2(x_1^{(j)}, x_2^{(j)}) + \left. \frac{\partial f_2(x_1, x_2)}{\partial x_1} (\hat{x}_1 - x_1^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \\&\quad + \left. \frac{\partial f_2(x_1, x_2)}{\partial x_2} (\hat{x}_2 - x_2^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} + \dots.\end{aligned}\tag{6}$$

Método de Newton

- Si reordenamos (6), tenemos:

$$\left| \frac{\partial f_1(x_1, x_2)}{\partial x_1} (\hat{x}_1 - x_1^{(j)}) + \frac{\partial f_1(x_1, x_2)}{\partial x_2} (\hat{x}_2 - x_2^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} = -f_1(x_1^{(j)}, x_2^{(j)}),$$
$$\left| \frac{\partial f_2(x_1, x_2)}{\partial x_1} (\hat{x}_1 - x_1^{(j)}) + \frac{\partial f_2(x_1, x_2)}{\partial x_2} (\hat{x}_2 - x_2^{(j)}) \right|_{x_1^{(j)}, x_2^{(j)}} = -f_2(x_1^{(j)}, x_2^{(j)}). \quad (7)$$

- Este mismo sistema lo podemos escribir en forma matricial.

Método de Newton

- De esta forma, nos queda:

$$\begin{bmatrix} \frac{\partial f_1(x_1, x_2)}{\partial x_1} & \frac{\partial f_1(x_1, x_2)}{\partial x_2} \\ \frac{\partial f_2(x_1, x_2)}{\partial x_1} & \frac{\partial f_2(x_1, x_2)}{\partial x_2} \end{bmatrix}_{(x_1^{(j)}, x_2^{(j)})} \cdot \begin{bmatrix} \hat{x}_1 - x_1^{(j)} \\ \hat{x}_2 - x_2^{(j)} \end{bmatrix} = \begin{bmatrix} -f_1(x_1^{(j)}, x_2^{(j)}) \\ -f_2(x_1^{(j)}, x_2^{(j)}) \end{bmatrix} \quad (8)$$

- Si lo generalizamos, queda:

$$\mathbf{J}(\mathbf{x}^{(j)}) \cdot (\hat{\mathbf{x}} - \mathbf{x}^{(j)}) = -\mathbf{F}(\mathbf{x}^{(j)}). \quad (9)$$

- Si operamos algebraicamente, obtenemos:

$$\begin{aligned} \hat{\mathbf{x}} - \mathbf{x}^{(j)} &= -\mathbf{J}(\mathbf{x}^{(j)})^{-1} \cdot \mathbf{F}(\mathbf{x}^{(j)}) \\ \hat{\mathbf{x}} &= \mathbf{x}^{(j)} - \mathbf{J}(\mathbf{x}^{(j)})^{-1} \cdot \mathbf{F}(\mathbf{x}^{(j)}) \Rightarrow \\ \mathbf{x}^{(j+1)} &= \mathbf{x}^{(j)} - \mathbf{J}(\mathbf{x}^{(j)})^{-1} \cdot \mathbf{F}(\mathbf{x}^{(j)}) \end{aligned} \quad (10)$$

Método de Newton

- Es evidente que tenemos algunos puntos que resolver:
 - Cálculo del *Jacobiano*: hay que hacerlo para cada iteración.
 - Invertir el *Jacobiano*: también hay que hacerlo en cada iteración.
- El método tendrá una mejor convergencia pero el costo operativo es muy alto.
- Hemos de buscar la forma de evitar la inversión del *Jacobiano* como primera opción.
- Si tenemos:

$$\mathbf{J}(\mathbf{x}^{(j)}) \cdot (\mathbf{x}^{(j+1)} - \mathbf{x}^{(j)}) = -\mathbf{F}(\mathbf{x}^{(j)}), \quad (11)$$

podemos definir

$$\mathbf{s}^{(j)} = \mathbf{x}^{(j+1)} - \mathbf{x}^{(j)}. \quad (12)$$

- Con esta definición, resulta:

$$\mathbf{J}(\mathbf{x}^{(j)}) \cdot \mathbf{s}^{(j)} = -\mathbf{F}(\mathbf{x}^{(j)}), \quad (13)$$

y no necesitamos invertir $\mathbf{J}(\mathbf{x}^{(j)})$ pues nos queda un *Sistema de Ecuaciones Lineales*.

Método de Newton

- Seguimos con la necesidad de calcular el *Jacobiano en cada iteración*.
- Una forma de evitar este cálculo es aproximarlos así:

$$\begin{aligned}\frac{\partial f_k(x_1, x_2)}{\partial x_1} &= \frac{f_k(x_1 + h_1, x_2) - f_k(x_1, x_2)}{h_1}, \\ \frac{\partial f_k(x_1, x_2)}{\partial x_2} &= \frac{f_k(x_1, x_2 + h_2) - f_k(x_1, x_2)}{h_2}.\end{aligned}\tag{14}$$

- Esta aproximación evita el cálculo del *Jacobiano* pero seguimos con el problema de calcular la aproximación en cada iteración.
- Ahora la idea es encontrar una forma de no tener que aproximar explícitamente el *Jacobiano*.

Método de Broyden

- Esta idea se logra con el *Método de Broyden*.
- Vamos a suponer que tenemos el *Jacobiano* e iteramos una vez:

$$\mathbf{x}^{(0)} \rightarrow \mathbf{F}(\mathbf{x}^{(0)}), \mathbf{J}(\mathbf{x}^{(0)}) \rightarrow \mathbf{x}^{(1)}. \quad (15)$$

- En primera instancia, vamos a evitar las aproximaciones en las iteraciones siguientes, proponiendo que:

$$\mathbf{A}^{(1)} \cdot (\mathbf{x}^{(1)} - \mathbf{x}^{(0)}) = \mathbf{F}(\mathbf{x}^{(1)}) - \mathbf{F}(\mathbf{x}^{(0)}), \quad (16)$$

con una condición:

$$\mathbf{A}^{(1)} \cdot \mathbf{z} = \mathbf{J}(\mathbf{x}^{(0)}) \cdot \mathbf{z} \text{ con } (\mathbf{x}^{(1)} - \mathbf{x}^{(0)})^T \cdot \mathbf{z} = 0, \quad (17)$$

donde \mathbf{z} es un vector adicional.

Método de Broyden

- Con esta nueva idea, obtenemos esta aproximación del *Jacobiano*:

$$\mathbf{A}^{(1)} = J(\mathbf{x}^{(0)}) + \frac{[\mathbf{F}(\mathbf{x}^{(1)}) - \mathbf{F}(\mathbf{x}^{(0)}) - \mathbf{J}(\mathbf{x}^{(0)}) \cdot (\mathbf{x}^{(1)} - \mathbf{x}^{(0)})] \cdot (\mathbf{x}^{(1)} - \mathbf{x}^{(0)})^T}{\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_2^2}. \quad (18)$$

- Con esta aproximación, $\mathbf{J}(\mathbf{x}^{(1)}) \approx \mathbf{A}^{(1)}$, podemos hacer lo siguiente:

$$\begin{aligned} \mathbf{A}^{(1)} \cdot \mathbf{s}^{(1)} &= -\mathbf{F}(\mathbf{x}^{(1)}), \\ \mathbf{x}^{(2)} &= \mathbf{x}^{(1)} + \mathbf{s}^{(1)}. \end{aligned} \quad (19)$$

- También podemos expresarlo así:

$$\mathbf{x}^{(2)} = \mathbf{x}^{(1)} - \mathbf{A}^{(1)^{-1}} \cdot \mathbf{F}(\mathbf{x}^{(1)}). \quad (20)$$

Método de Broyden

- Esto último parece una total contradicción pues hemos desarrollado todo el método anterior para no tener que invertir una matriz.
- Primero generalicemos todo lo hecho. Para ello, si $\mathbf{s}^{(j)} = \mathbf{x}^{(j+1)} - \mathbf{x}^{(j)}$ podemos partir de lo visto y hacer:

$$\Delta \mathbf{F}(\mathbf{x}^{(0)}) = \mathbf{y}^{(0)} = \mathbf{F}(\mathbf{x}^{(1)}) - \mathbf{F}(\mathbf{x}^{(0)}), \quad (21)$$

y al generalizar queda

$$\mathbf{y}^{(j)} = \mathbf{F}(\mathbf{x}^{(j+1)}) - \mathbf{F}(\mathbf{x}^{(j)}). \quad (22)$$

- Además, vamos a definir $\mathbf{A}^{(0)} = \mathbf{J}(\mathbf{x}^{(0)})$ y queda:

$$\mathbf{A}^{(1)} = \mathbf{A}^{(0)} + \frac{[\mathbf{y}^{(0)} - \mathbf{A}^{(0)} \cdot \mathbf{s}^{(0)}] \cdot \mathbf{s}^{(0)T}}{\|\mathbf{s}^{(0)}\|_2^2}. \quad (23)$$

Método de Broyden

- Al generalizar esta última expresión, nos queda:

$$\mathbf{A}^{(j+1)} = \mathbf{A}^{(j)} + \frac{(\mathbf{y}^{(j)} - \mathbf{A}^{(j)} \cdot \mathbf{s}^{(j)}) \cdot \mathbf{s}^{(j)T}}{\|\mathbf{s}^{(j)}\|_2^2}. \quad (24)$$

- Facilita el cálculo de las matrices $\mathbf{A}^{(j)}$ pero mantiene la necesidad de calcular la inversa de la matriz $(\mathbf{A}^{(j)})^{-1}$.
- El álgebra numérica nos trae una solución a esta situación, si definimos $\mathbf{H}^{(j)} = \mathbf{A}^{(j)-1}$:

$$\mathbf{H}^{(j+1)} = \mathbf{H}^{(j)} + \frac{(\mathbf{s}^{(j)} - \mathbf{H}^{(j)} \cdot \mathbf{y}^{(j)}) \cdot \mathbf{s}^{(j)T} \cdot \mathbf{H}^{(j)}}{\mathbf{s}^{(j)T} \cdot \mathbf{H}^{(j)} \cdot \mathbf{y}^{(j)}}, \quad (25)$$

el algoritmo de Sherman y Morrison para obtener la inversa de una matriz.

Método de Broyden





- Con este desarrollo obtenemos el *Método de Broyden de primer orden*. El algoritmo del método es el siguiente:

$$\begin{aligned}
\mathbf{x}^{(0)} &\rightarrow \mathbf{F}(\mathbf{x}^{(0)}) \rightarrow \mathbf{J}(\mathbf{x}^{(0)}) = \mathbf{A}^{(0)} \\
\mathbf{x}^{(1)} &= \mathbf{x}^{(0)} - \mathbf{A}^{(0)^{-1}} \cdot \mathbf{F}(\mathbf{x}^{(0)}) \rightarrow \mathbf{F}(\mathbf{x}^{(1)}) \\
\mathbf{s}^{(j-1)} &= \mathbf{x}^{(j)} - \mathbf{x}^{(j-1)} \\
\mathbf{y}^{(j-1)} &= \mathbf{F}(\mathbf{x}^{(j)}) - \mathbf{F}(\mathbf{x}^{(j-1)}) \\
\mathbf{H}^{(j)} &= \mathbf{A}^{(j)^{-1}} \\
\mathbf{H}^{(j)} &= \mathbf{H}^{(j-1)} + \frac{[\mathbf{s}^{(j-1)} - \mathbf{H}^{(j-1)} \cdot \mathbf{y}^{(j-1)}] \cdot \mathbf{s}^{(j-1)^T} \cdot \mathbf{H}^{(j-1)}}{\mathbf{s}^{(j-1)^T} \cdot \mathbf{H}^{(j-1)} \cdot \mathbf{y}^{(j-1)}} \\
\mathbf{x}^{(j+1)} &= \mathbf{x}^{(j)} - \mathbf{H}^{(j)} \cdot \mathbf{F}(\mathbf{x}^{(j)}),
\end{aligned} \tag{26}$$

Método de Broyden

- Este método se clasifica también como *Método Quasi-Newton*.
- Ventajas
 - No requiere invertir matrices en las iteraciones, salvo al iniciar el proceso ($\mathbf{x}^{(0)}$).
 - Tampoco es necesario calcular la matriz jacobiana.
- Desventajas
 - Hay que invertir la matriz jacobiana al empezar.
 - La convergencia es supralineal, en el mejor de los casos.
- Una ventaja adicional es que al aplicar el algoritmo de Sherman y Morrison no es necesario calcular la matriz $\mathbf{J}(\mathbf{x})$ ni calcular su inversa.
- Como $\mathbf{H}^{(j)} = \mathbf{A}^{(j)^{-1}}$, podemos definir la primera matriz como $\mathbf{A}^{(0)} = \mathbf{I}$, y $\mathbf{H}^{(0)} = \mathbf{I}$.
- El costo de esta aproximación es reducir la velocidad de convergencia.

Bibliografía

-  Broyden, C.G.
A class of methods for solving Nonlinear Simultaneous Equations.
Math. Comp. 19, 577-593, 1965.
-  Burden, R. L., Faires, J. D. & Burden, A. M.
Análisis Numérico.
Décima Edición. CENGAGE Learning, 2016.
-  Dennis Jr., J. E. & Moré, J.
Quasi-Newton methods, motivation and theory.
SIAM Review, Vol 19, No 1, pp. 46-89. January 1977.
-  Schwarz, R.
Resumen de clases.