

★ MÉTODO DE RELAJACIÓN

Hasta acá la rapidez de la convergencia de la sucesión $\{x^{(k)}\}_{k=1}^{\infty}$ generada por la iteración $x^{(k+1)} = Tx^{(k)} + c$ depende de $\rho(T)$. Los métodos de relajación buscan la rapidez de convergencia vista desde el problema equivalente original $Ax = b$, esto es, buscan acelerar la convergencia a 0 de la sucesión de residuales $\{r^{(k)}\}_{k=1}^{\infty}$, $r^{(k)} = Ax^{(k)} - b$.

Los métodos de relajación introducen un parámetro $\omega > 0$ que busca reducir la norma del vector residual y lograr una convergencia significativamente más rápida. Para ello, partimos de la ecuación de iteración escalar de Gauss-Seidel

$$x_i^{(k+1)} = \frac{b_i}{a_{i,i}} - \frac{1}{a_{i,i}} \sum_{j=1}^{i-1} a_{i,j} x_j^{(k+1)} - \frac{1}{a_{i,i}} \sum_{j=i+1}^n a_{i,j} x_j^{(k)}, \quad i = 1, \dots, n$$

e introducimos el parámetro ω generando la ecuación de **iteración escalar** de los métodos de relajación

$$x_i^{(k+1)} = (1 - \omega) x_i^{(k)} + \omega \left(\frac{b_i}{a_{i,i}} - \frac{1}{a_{i,i}} \sum_{j=1}^{i-1} a_{i,j} x_j^{(k+1)} - \frac{1}{a_{i,i}} \sum_{j=i+1}^n a_{i,j} x_j^{(k)} \right), \quad i = 1, \dots, n.$$

Los métodos de relajación para las elecciones de $0 < \omega < 1$ reciben el nombre de métodos de subrelajación, para las elecciones de $\omega > 1$ reciben el nombre de métodos de sobrerelajación y para $\omega = 1$ estamos en el método de Gauss-Seidel, así que el método de Gauss-Seidel es también un método de relajación. Estos métodos se designan con la abreviatura S.O.R. (Successive Over Relaxation) sobrerelajación sucesiva.

Vamos a obtener la matriz de iteración del método de S.O.R., partimos del sistema $Ax = b$ y escribimos $A = D - L - U$

$$\left. \begin{aligned} (D - L - U)x &= b \\ \omega(D - L - U)x &= \omega b \\ \omega(D - L - U)x + Dx - Dx &= \omega b \\ \omega(D - L - U)x + Dx - Dx &= \omega b \\ (D - \omega L)x &= [(1 - \omega)D + \omega U]x + \omega b \\ x &= (D - \omega L)^{-1} [(1 - \omega)D + \omega U]x + \omega(D - \omega L)^{-1} b \end{aligned} \right\} \Rightarrow \begin{aligned} T_\omega &:= (D - \omega L)^{-1} [(1 - \omega)D + \omega U] \\ c_\omega &:= \omega(D - \omega L)^{-1} b \\ &\Downarrow \\ x^{(k+1)} &= T_\omega x^{(k)} + c_\omega \end{aligned}$$

La pregunta ahora es: ¿cuál será el ω adecuado? Consideremos un primer resultado que nos limita la elección de ω .

Teorema. Si $a_{i,i} \neq 0$, $i = 1, \dots, n$, entonces $\rho(T_\omega) \geq |\omega - 1|$.

Para garantizar la convergencia de la sucesión generada por la iteración $x^{(k+1)} = T_\omega x^{(k)} + c_\omega$ necesitamos que $\rho(T_\omega) < 1$ y del teorema se sigue que $|\omega - 1| < 1$, esto es $0 < \omega < 2$. Esto significa que el método de S.O.R. puede converger sólo si $0 < \omega < 2$, ya que si $\omega \in [2, \infty)$, $\rho(T_\omega) \geq 1$. ¿Que significa la palabra puede en la frase anterior? Una primera condición sobre ω es que $0 < \omega < 2$, si $0 < \omega < 2$ la convergencia de la sucesión nuevamente recae sobre T_ω .

Ejemplo Considere el sistema de ecuaciones lineales

$$\begin{aligned} 4x_1 + 7x_2 - x_3 &= 1 \\ 2x_1 - 4x_2 - 3x_3 &= -1 \\ -x_1 + 3x_2 - 5x_3 &= 4 \end{aligned}$$

Analizar la convergencia de la sucesión generada por el método de S.O.R. con parámetros $\omega_1 = \frac{3}{2}$ y $\omega_2 = \frac{1}{2}$.

Solución: Notemos que los parámetros si pertenecen al intervalo donde podemos esperar obtener sucesiones convergentes a partir del método de S.O.R. Ahora, para analizar convergencia debemos obtener las matrices de iteración T_{ω_1} y T_{ω_2} y para ello, vamos a partir de la ecuación de iteración escalar del método de Gauss-Seidel

$$\begin{aligned}x_1^{(k+1)} &= \frac{1}{4} - \frac{7}{4}x_2^{(k)} + \frac{1}{4}x_3^{(k)} \\x_2^{(k+1)} &= \frac{1}{4} + \frac{1}{2}x_1^{(k+1)} - \frac{3}{4}x_3^{(k)} \\x_3^{(k+1)} &= -\frac{4}{5} - \frac{1}{5}x_1^{(k+1)} + \frac{3}{5}x_2^{(k+1)}\end{aligned}$$

de aquí que la iteración escalar del método de S.O.R. con parámetro ω es

$$\begin{aligned}x_1^{(k+1)} &= (1 - \omega)x_1^{(k)} + \omega\left(\frac{1}{4} - \frac{7}{4}x_2^{(k)} + \frac{1}{4}x_3^{(k)}\right) \\x_2^{(k+1)} &= (1 - \omega)x_2^{(k)} + \omega\left(\frac{1}{4} + \frac{1}{2}x_1^{(k+1)} - \frac{3}{4}x_3^{(k)}\right) \\x_3^{(k+1)} &= (1 - \omega)x_3^{(k)} + \omega\left(-\frac{4}{5} - \frac{1}{5}x_1^{(k+1)} + \frac{3}{5}x_2^{(k+1)}\right)\end{aligned}$$

$\omega_1 = \frac{3}{2}$ Reemplazamos ω por ω_1 y procedemos de manera similar a como se hizo para obtener la matriz de iteración de Gauss-Seidel en un ejemplo pasado.

$$\left. \begin{aligned}x_1^{(k+1)} &= -\frac{1}{2}x_1^{(k)} + \frac{3}{2}\left(\frac{1}{4} - \frac{7}{4}x_2^{(k)} + \frac{1}{4}x_3^{(k)}\right) \\x_2^{(k+1)} &= -\frac{1}{2}x_2^{(k)} + \frac{3}{2}\left(\frac{1}{4} + \frac{1}{2}x_1^{(k+1)} - \frac{3}{4}x_3^{(k)}\right) \\x_3^{(k+1)} &= -\frac{1}{2}x_3^{(k)} + \frac{3}{2}\left(-\frac{4}{5} - \frac{1}{5}x_1^{(k+1)} + \frac{3}{5}x_2^{(k+1)}\right)\end{aligned} \right\} \Rightarrow \begin{aligned}x_1^{(k+1)} &= \frac{3}{8} - \frac{1}{2}x_1^{(k)} - \frac{21}{8}x_2^{(k)} + \frac{3}{8}x_3^{(k)} \\x_2^{(k+1)} &= \frac{21}{32} - \frac{3}{8}x_1^{(k)} - \frac{79}{32}x_2^{(k)} - \frac{27}{32}x_3^{(k)} \\x_3^{(k+1)} &= -\frac{231}{320} - \frac{3}{16}x_1^{(k)} - \frac{459}{320}x_2^{(k)} - \frac{439}{320}x_3^{(k)}\end{aligned}$$

$$\text{así } T_{\omega_1} = \begin{bmatrix} -\frac{1}{2} & -\frac{21}{8} & \frac{3}{8} \\ -\frac{3}{8} & -\frac{79}{32} & -\frac{27}{32} \\ -\frac{3}{16} & -\frac{459}{320} & -\frac{439}{320} \end{bmatrix}.$$

Dado que hay más de una entrada en la matriz T_{ω_1} mayor a 1, se sigue que $\|T_{\omega_1}\|_1 > 1$ y $\|T_{\omega_1}\|_\infty > 1$ y no podemos concluir convergencia. Se puede mostrar que $\sigma(T_{\omega_1}) \approx \{-3.4209, -0.0416, -0.8781\}$ (ejercicio), así $\rho(T_{\omega_1}) \approx 3.4209 > 1$ y podemos concluir que la sucesión generada por la iteración $x^{(k+1)} = T_{\omega_1}x^{(k)} + c_{\omega_1}$ no converge para toda aproximación inicial $x^{(0)} \in \mathbb{R}^3$.

$\omega_2 = \frac{1}{2}$ Nuevamente, reemplazamos ω por ω_2 y procedemos como arriba.

$$\left. \begin{aligned}x_1^{(k+1)} &= \frac{1}{2}x_1^{(k)} + \frac{1}{2}\left(\frac{1}{4} - \frac{7}{4}x_2^{(k)} + \frac{1}{4}x_3^{(k)}\right) \\x_2^{(k+1)} &= \frac{1}{2}x_2^{(k)} + \frac{1}{2}\left(\frac{1}{4} + \frac{1}{2}x_1^{(k+1)} - \frac{3}{4}x_3^{(k)}\right) \\x_3^{(k+1)} &= \frac{1}{2}x_3^{(k)} + \frac{1}{2}\left(-\frac{4}{5} - \frac{1}{5}x_1^{(k+1)} + \frac{3}{5}x_2^{(k+1)}\right)\end{aligned} \right\} \Rightarrow \begin{aligned}x_1^{(k+1)} &= \frac{1}{8} - \frac{1}{2}x_1^{(k)} - \frac{7}{8}x_2^{(k)} + \frac{1}{8}x_3^{(k)} \\x_2^{(k+1)} &= \frac{5}{32} + \frac{1}{8}x_1^{(k)} + \frac{9}{32}x_2^{(k)} - \frac{11}{32}x_3^{(k)} \\x_3^{(k+1)} &= -\frac{117}{320} - \frac{1}{80}x_1^{(k)} + \frac{11}{64}x_2^{(k)} + \frac{123}{320}x_3^{(k)}\end{aligned}$$

así

$$T_{\omega_2} = \begin{bmatrix} \frac{1}{2} & -\frac{7}{8} & \frac{1}{8} \\ \frac{1}{8} & \frac{9}{32} & -\frac{11}{32} \\ -\frac{1}{80} & \frac{11}{64} & \frac{123}{320} \end{bmatrix} \Rightarrow \begin{aligned}\|T_{\omega_2}\|_1 &= \max\left\{\frac{51}{80}, \frac{85}{64}, \frac{273}{320}\right\} = \frac{85}{64} \\ \|T_{\omega_2}\|_\infty &= \max\left\{\frac{3}{2}, \frac{3}{4}, \frac{91}{160}\right\} = \frac{3}{2}\end{aligned}$$

no podemos concluir convergencia. Se puede mostrar que $\sigma(T_{\omega_2}) \approx \{0.3730 \pm 0.3985i, 0.4196\}$ (ejercicio), así $\rho(T_{\omega_2}) \approx \max\{0.5458, 0.4196\} = 0.5458 < 1$ y podemos concluir que la sucesión generada por la iteración $x^{(k+1)} = T_{\omega_2}x^{(k)} + c_{\omega_2}$ converge para toda aproximación inicial $x^{(0)} \in \mathbb{R}^3$.

Para el método de S.O.R. también existe un caso especial donde podemos concluir convergencia basados en la matriz de coeficientes \mathbf{A} del sistema de ecuaciones $\mathbf{Ax} = \mathbf{b}$. Antes de ver el teorema de convergencia respectivo vamos a introducir/recordar unos conceptos y resultados.

Definición. Una matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ es definida positiva si es simétrica y si $\mathbf{y}^T \mathbf{A} \mathbf{y} > 0$, para todo $\mathbf{y} \in \mathbb{R}^n$, $\mathbf{y} \neq \mathbf{0}$.

Ejemplo Hallar los valores de α para los cuales la matriz $\mathbf{B} = \begin{bmatrix} 5 & \alpha \\ \alpha & 8 \end{bmatrix}$ es definida positiva.

Solución: Notemos primero que la matriz \mathbf{B} es simétrica. Sea $\mathbf{y} = [x, z]^T \in \mathbb{R}^2$ no nulo

$$\mathbf{y}^T \mathbf{B} \mathbf{y} = [x \ z] \begin{bmatrix} 5 & \alpha \\ \alpha & 8 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} = [x \ z] \begin{bmatrix} 5x + \alpha z \\ \alpha x + 8z \end{bmatrix} = 5x^2 + 2\alpha xz + 8z^2 = 5 \left(x + \frac{\alpha}{5}z\right)^2 + \left(8 - \frac{1}{5}\alpha^2\right)z^2$$

así $\mathbf{y}^T \mathbf{B} \mathbf{y} > 0$ siempre que $8 - \frac{1}{5}\alpha^2 > 0$, esto es, cuando $\alpha \in (-\sqrt{40}, \sqrt{40})$. Concluimos que \mathbf{B} es una matriz definida positiva siempre que $\alpha \in (-\sqrt{40}, \sqrt{40})$.

Notemos que analizar la condición $\mathbf{y}^T \mathbf{B} \mathbf{y} > 0$ no es inmediato, acá se pudo factorizar fácil, pero para una matriz de orden mayor no lo será.

Definición. Una primera submatriz principal de una matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ es la matriz que tiene la forma

$$\mathbf{A}_m := \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,m} \end{bmatrix} \in \mathbb{R}^{m \times m} \quad \text{para } 1 \leq m \leq n.$$

Teorema. Sea \mathbf{A} una matriz simétrica.

- ★ \mathbf{A} es definida positiva si y sólo si sus primeras submatrices principales tienen determinante positivo.
- ★ \mathbf{A} es definida positiva si y sólo si todos los valores propios son positivos.

Tenemos hasta acá tres formas de analizar cuándo una matriz es definida positiva. El siguiente teorema establece el caso especial de convergencia para el método de S.O.R.

Teorema. Si \mathbf{A} es una matriz definida positiva y $0 < \omega < 2$ entonces la sucesión generada por método de S.O.R. converge $\forall \mathbf{x}^{(0)} \in \mathbb{R}^n$.

Ejemplo Considere el sistema de ecuaciones lineales

$$\begin{aligned} 3x_1 - x_2 + 2x_3 &= -21 \\ -x_1 + 3x_2 - 2x_3 &= 13 \\ 2x_1 - 2x_2 + 4x_3 &= 14 \end{aligned}$$

Analizar la convergencia de la sucesión generada por el método de S.O.R. ¿Qué puede decir sobre la convergencia de la sucesión generada por el método de Gauss-Seidel?

Solución: Si escribimos el sistema en forma matricial $\mathbf{Ax} = \mathbf{b}$, la matriz de coeficientes del sistema es

$$\mathbf{A} = \begin{bmatrix} 3 & -1 & 2 \\ -1 & 3 & -2 \\ 2 & -2 & 4 \end{bmatrix}.$$

Notemos que \mathbf{A} es simétrica, veamos que se cumple una de las siguientes opciones: $\mathbf{y}^T \mathbf{Ay} > 0$ para todo vector no nulo $\mathbf{y} \in \mathbb{R}^3$ (**ejercicio**), \mathbf{A} tiene todos sus valores propios positivos (**ejercicio**) o las primeras submatrices principales de \mathbf{A} tienen determinante positivo. Las primeras submatrices principales de \mathbf{A} son

$$\mathbf{A}_1 = [3], \quad \mathbf{A}_2 = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \quad \text{y} \quad \mathbf{A}_3 = \mathbf{A}$$

así $\det(\mathbf{A}_1) = 3$, $\det(\mathbf{A}_2) = 8$ y $\det(\mathbf{A}_3) = 16$. Podemos concluir que \mathbf{A} es una matriz definida positiva y gracias al teorema, la sucesión generada por el método de S.O.R. converge para cualquier $\omega \in (0, 2)$ y $\forall \mathbf{x}^{(0)} \in \mathbb{R}^3$. Dado que al tomar $\omega = 1$ se obtiene el método de Gauss-Seidel, podemos también concluir que la sucesión generada por el método de Gauss-Seidel converge $\forall \mathbf{x}^{(0)} \in \mathbb{R}^3$.

Una pregunta que surge en el caso de matriz \mathbf{A} definida positiva es ¿cuál parámetro ω hará que la sucesión converja más rápido?

Definición. Una matriz $\mathbf{A} \in \mathbb{R}^{n \times n}$ recibe el nombre de matriz banda si existen los enteros p y q con $1 < p, q < n$ que tienen la propiedad que $a_{i,j} = 0$ siempre que $p \leq j - i$ o $q \leq i - j$. El ancho de banda se define como $p + q - 1$ donde p describe el número de diagonales superiores no nulas (incluyendo la diagonal principal) y q el número de diagonales no nulas inferiores (incluyendo la diagonal principal).

Ejemplo Hallar el ancho de banda de las siguientes matrices banda.

$$\mathbf{A} = \begin{bmatrix} 5 & 0 & 8 & 0 & 0 \\ -1 & 15 & 0 & 2 & 0 \\ -4 & -1 & -3 & 0 & 0 \\ 0 & 6 & -1 & 5 & 0 \\ 0 & 0 & 7 & -1 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 5 & 3 & 0 & 0 & 0 \\ 4 & 5 & 4 & 0 & 0 \\ -3 & 14 & 5 & 7 & 0 \\ 0 & 8 & 4 & 5 & 0 \\ 0 & 0 & -7 & 4 & 5 \end{bmatrix}, \quad \mathbf{M} = \begin{bmatrix} 4 & -3 & 0 & 0 \\ -3 & 5 & 1 & 0 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{bmatrix}.$$

Solución: Notemos que todas las matrices tienen algunas diagonales nulas en los extremos, por esto decimos que son matrices banda. Analicemos primero el ancho de banda de la matriz \mathbf{B}

$$\mathbf{B} = \begin{bmatrix} 5 & 3 & 0 & 0 & 0 \\ 4 & 5 & 4 & 0 & 0 \\ -3 & 14 & 5 & 7 & 0 \\ 0 & 8 & 4 & 5 & 0 \\ 0 & 0 & -7 & 4 & 5 \end{bmatrix}$$

Tenemos tres diagonales nulas superiores (azul) y tenemos dos diagonales nulas inferiores (rojo), luego $p = 2$, $q = 3$ y el ancho de banda es 4.

Si analizamos \mathbf{M} , esta tiene dos diagonales superiores e inferiores nulas, luego $p = 2$, $q = 2$ y el ancho de banda es 3. Finalmente, la matriz \mathbf{A} tiene dos diagonales superiores e inferiores nulas, luego $p = 3$, $q = 3$ y el ancho de banda es 5.

Las matrices de ancho de banda 3 con $p = q = 2$ reciben el nombre de matriz tridiagonal y tienen la forma

$$\begin{bmatrix} a_{1,1} & a_{1,2} & 0 & \cdots & 0 \\ a_{2,1} & a_{2,2} & a_{2,3} & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & 0 & a_{n,n-1} & a_{n,n} \end{bmatrix}.$$

Finalizamos con el caso más especial de todos.

Teorema. Si \mathbf{A} es una matriz definida positiva y tridiagonal entonces $\rho(\mathbf{T}_{\text{gs}}) = [\rho(\mathbf{T}_{\text{j}})]^2 < 1$ y la elección óptima de ω para el método de S.O.R. es

$$\omega_{\text{opt}} := \frac{2}{1 + \sqrt{1 - [\rho(\mathbf{T}_{\text{j}})]^2}}$$

y en este caso $\rho(\mathbf{T}_{\omega_{\text{opt}}}) = \omega_{\text{opt}} - 1$.

El teorema nos garantiza la convergencia de los métodos de Jacobi, Gauss-Seidel y S.O.R. con $0 < \omega < 2$, para cualquier aproximación inicial $\mathbf{x}^{(0)} \in \mathbb{R}^n$.

Ejemplo Aproximar la solución del sistema

$$\begin{aligned} 4x_1 - x_2 &= 5 \\ -x_1 + 4x_2 - x_3 &= -3 \\ -x_2 + 4x_3 - x_4 &= 8 \\ -x_3 + 4x_4 &= 13 \end{aligned}$$

usando el método de S.O.R. con parámetro óptimo ω_{opt} y $\mathbf{x}^{(0)} = [0, 3, 2, -1]^T$.

Solución: La matriz de coeficientes del sistema \mathbf{A} esta dada por

$$\mathbf{A} = \begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \quad \begin{array}{l} \text{la matriz } \mathbf{A} \text{ es simétrica, } \det(\mathbf{A}_1) = 4, \det(\mathbf{A}_2) = 15, \det(\mathbf{A}_3) = 56 \text{ y } \det(\mathbf{A}_4) = 209, \text{ luego} \\ \mathbf{A} \text{ es una matriz definida positiva, así que podemos garantizar convergencia del método de} \\ \text{S.O.R. para todo } \omega \in (0, 2). \text{ Adicionalmente, la matriz } \mathbf{A} \text{ es tridiagonal, así que también} \\ \text{podemos garantizar convergencia de Jacobi (Gauss-Seidel ya se tenía al tomar } \omega = 1). \end{array}$$

Y por ser este caso especial, podemos obtener el parámetro óptimo ω_{opt} . Se puede mostrar que $\sigma(\mathbf{T}_{\text{j}}) \approx \{\pm 0.1545, \pm 0.4045\}$ ([ejercicio](#)), así

$$\omega_{\text{opt}} \approx \frac{2}{1 + \sqrt{1 - [0.4045]^2}} \approx 1.04464$$

y según el teorema $\rho(\mathbf{T}_{\omega_{\text{opt}}}) \approx 1.04464 - 1 = 0.04464$. Para obtener las aproximaciones a la solución del sistema, basta emplear las ecuaciones escalares

$$\begin{aligned} x_1^{(\mathbf{k}+1)} &\approx -0.04464x_1^{(\mathbf{k})} + 1.04464 \left(\frac{5}{4} + \frac{1}{4}x_2^{(\mathbf{k})} \right) \\ x_2^{(\mathbf{k}+1)} &\approx -0.04464x_2^{(\mathbf{k})} + 1.04464 \left(-\frac{3}{4} + \frac{1}{4}x_1^{(\mathbf{k}+1)} + \frac{1}{4}x_3^{(\mathbf{k})} \right) \\ x_3^{(\mathbf{k}+1)} &\approx -0.04464x_3^{(\mathbf{k})} + 1.04464 \left(2 + \frac{1}{4}x_2^{(\mathbf{k}+1)} + \frac{1}{4}x_4^{(\mathbf{k})} \right) \\ x_4^{(\mathbf{k}+1)} &\approx -0.04464x_4^{(\mathbf{k})} + 1.04464 \left(\frac{13}{4} + \frac{1}{4}x_3^{(\mathbf{k}+1)} \right) \end{aligned}$$

Así, tomando $\mathbf{x}^{(0)} = [0, 3, 2, -1]^T$

$$\left. \begin{aligned} x_1^{(\mathbf{1})} &\approx -0.04464x_1^{(\mathbf{0})} + 1.04464 \left(\frac{5}{4} + \frac{1}{4}x_2^{(\mathbf{0})} \right) \\ x_2^{(\mathbf{1})} &\approx -0.04464x_2^{(\mathbf{0})} + 1.04464 \left(-\frac{3}{4} + \frac{1}{4}x_1^{(\mathbf{1})} + \frac{1}{4}x_3^{(\mathbf{0})} \right) \\ x_3^{(\mathbf{1})} &\approx -0.04464x_3^{(\mathbf{0})} + 1.04464 \left(2 + \frac{1}{4}x_2^{(\mathbf{1})} + \frac{1}{4}x_4^{(\mathbf{0})} \right) \\ x_4^{(\mathbf{1})} &\approx -0.04464x_4^{(\mathbf{0})} + 1.04464 \left(\frac{13}{4} + \frac{1}{4}x_3^{(\mathbf{1})} \right) \end{aligned} \right\} \Rightarrow \begin{aligned} x_1^{(\mathbf{1})} &\approx 2.0893 \\ x_2^{(\mathbf{1})} &\approx 0.1506 \\ x_3^{(\mathbf{1})} &\approx 1.7782 \\ x_4^{(\mathbf{1})} &\approx 3.9041 \end{aligned} \Rightarrow \begin{aligned} \|\mathbf{x}^{(\mathbf{1})} - \mathbf{x}^{(\mathbf{0})}\|_1 &\approx 10.0647 \\ \|\mathbf{r}^{(\mathbf{1})}\|_1 &\approx 9.2521 \end{aligned}$$

en la novena iteración tenemos una aproximación $\mathbf{x}^{(\mathbf{9})}$ tal que $\|\mathbf{r}^{(\mathbf{9})}\|_1 \approx 6.7588e - 09$.

MÉTODO DE NEWTON PARA SISTEMAS NO LINEALES

Consideremos un sistema de ecuaciones no lineales

$$\left. \begin{array}{l} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{array} \right\} \iff \mathbf{F}(\mathbf{x}) = \mathbf{0}$$

donde $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 1, \dots, n$, campos escalares y $\mathbf{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ campo vectorial dado por

$$\mathbf{F}(\mathbf{x}) := \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix} = \begin{bmatrix} f_1(x_1, x_2, \dots, x_n) \\ f_2(x_1, x_2, \dots, x_n) \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) \end{bmatrix}$$

Nuevamente, vamos a construir una sucesión de aproximaciones $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ de \mathbf{x} , \mathbf{x} solución de $\mathbf{F}(\mathbf{x}) = \mathbf{0}$. Estamos ante un problema de ‘ceros’ y en vista que el método más completo que se estudio, fue el método de Newton, vamos a extender la idea para este sistema de ecuaciones no lineales $\mathbf{F}(\mathbf{x}) = \mathbf{0}$. Analíticamente, el método de Newton se baso en el Teorema de Taylor, así que vamos a partir recordando el Teorema de Taylor para una función en varias variables.

Teorema. Sea D una bola en \mathbb{R}^n centrada en \mathbf{y} y $f \in C^{n+1}(D)$. Para cualquier $\mathbf{x} \in D$ existe $\boldsymbol{\beta}$ entre \mathbf{x} y \mathbf{y} tal que

$$f(\mathbf{x}) = f(\mathbf{y}) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{y})(x_i - y_i) + \sum_{i=1}^n \sum_{j=1}^n \frac{1}{2!} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{y})(x_i - y_i)(x_j - y_j) + \sum_{|\boldsymbol{\alpha}|=3}^n \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\mathbf{y})(\mathbf{x} - \mathbf{y})^{\boldsymbol{\alpha}} + \dots + R_n(\mathbf{x})$$

donde $\boldsymbol{\alpha} = [\alpha_1, \dots, \alpha_n] \in \mathbb{N}^n$ es una n -tupla, $|\boldsymbol{\alpha}| = \alpha_1 + \dots + \alpha_n$, $\boldsymbol{\alpha}! = \alpha_1! \dots \alpha_n!$ y $R_n(\mathbf{x})$ es el residuo (o error de truncamiento)

$$R_n(\mathbf{x}) = \sum_{|\boldsymbol{\alpha}|=n+1}^n \frac{1}{\boldsymbol{\alpha}!} \frac{\partial^{|\boldsymbol{\alpha}|} f}{\partial \mathbf{x}^{\boldsymbol{\alpha}}}(\boldsymbol{\beta})(\mathbf{x} - \mathbf{y})^{\boldsymbol{\alpha}}$$

Básicamente, si f es suficientemente suave, esto es, sus derivadas parciales en todas las variables son continuas, podemos aproximar el valor de $f(\mathbf{x})$ por medio de un polinomio de grado n en todas sus variables, con un residuo $R_n(\mathbf{x})$ en términos del vector $\boldsymbol{\beta}$. La notación en términos de la n -tupla es para evitar escribir tres o mas sumatorias, lo que buscar es tener todos los posibles combinaciones de los α_i que suman el grado de la derivada.

Para obtener la ecuación de iteración para el método de Newton, vamos a suponer que tenemos un sistema no lineal con $n = 2$, esto es, consideremos el sistema

$$\begin{array}{l} f_1(x, y) = 0 \\ f_2(x, y) = 0 \end{array}$$

y supongamos que $f_1, f_2 \in C^2(D)$, $D \subseteq \mathbb{R}^2$ y $\mathbf{p} = [p, q]^T$ es tal que $\mathbf{F}(\mathbf{p}) = \mathbf{0}$. Sea $\mathbf{p}_0 = [p_0, q_0]^T$ una aproximación de \mathbf{p} . Sea $\mathbf{x} = [x, y]^T \in D$, por teorema existen $\boldsymbol{\beta}, \boldsymbol{\xi} \in D$ tales que

$$\begin{array}{l} f_1(\mathbf{x}) = f_1(\mathbf{p}_0) + \frac{\partial f_1}{\partial x}(\mathbf{p}_0)(x - p_0) + \frac{\partial f_1}{\partial y}(\mathbf{p}_0)(y - q_0) + \frac{1}{2} \left[\frac{\partial^2 f_1}{\partial x^2}(\boldsymbol{\beta})(x - p_0)^2 + \frac{\partial^2 f_1}{\partial y^2}(\boldsymbol{\beta})(y - q_0)^2 + 2 \frac{\partial^2 f_1}{\partial x \partial y}(\boldsymbol{\beta})(x - p_0)(y - q_0) \right] \\ f_2(\mathbf{x}) = f_2(\mathbf{p}_0) + \frac{\partial f_2}{\partial x}(\mathbf{p}_0)(x - p_0) + \frac{\partial f_2}{\partial y}(\mathbf{p}_0)(y - q_0) + \frac{1}{2} \left[\frac{\partial^2 f_2}{\partial x^2}(\boldsymbol{\xi})(x - p_0)^2 + \frac{\partial^2 f_2}{\partial y^2}(\boldsymbol{\xi})(y - q_0)^2 + 2 \frac{\partial^2 f_2}{\partial x \partial y}(\boldsymbol{\xi})(x - p_0)(y - q_0) \right] \end{array}$$

si evaluamos en \mathbf{p}

$$\begin{aligned} f_1(\mathbf{p}) &= f_1(\mathbf{p}_0) + \frac{\partial f_1}{\partial x}(\mathbf{p}_0)(p - p_0) + \frac{\partial f_1}{\partial y}(\mathbf{p}_0)(q - q_0) + \frac{1}{2} \left[\frac{\partial^2 f_1}{\partial x^2}(\boldsymbol{\beta})(p - p_0)^2 + \frac{\partial^2 f_1}{\partial y^2}(\boldsymbol{\beta})(q - q_0)^2 + 2 \frac{\partial^2 f_1}{\partial x \partial y}(\boldsymbol{\beta})(p - p_0)(q - q_0) \right] \\ f_2(\mathbf{p}) &= f_2(\mathbf{p}_0) + \frac{\partial f_2}{\partial x}(\mathbf{p}_0)(p - p_0) + \frac{\partial f_2}{\partial y}(\mathbf{p}_0)(q - q_0) + \frac{1}{2} \left[\frac{\partial^2 f_2}{\partial x^2}(\boldsymbol{\xi})(p - p_0)^2 + \frac{\partial^2 f_2}{\partial y^2}(\boldsymbol{\xi})(q - q_0)^2 + 2 \frac{\partial^2 f_2}{\partial x \partial y}(\boldsymbol{\xi})(p - p_0)(q - q_0) \right] \end{aligned}$$

$f_1(\mathbf{p}) = f_2(\mathbf{p}) = 0$, dado que \mathbf{p}_0 es una aproximación de \mathbf{p} , $(q - q_0)^2 \approx 0$, $(p - p_0)^2 \approx 0$ y $(p - p_0)(q - q_0) \approx 0$, así

$$\begin{aligned} 0 &\approx f_1(\mathbf{p}_0) + \frac{\partial f_1}{\partial x}(\mathbf{p}_0)(p - p_0) + \frac{\partial f_1}{\partial y}(\mathbf{p}_0)(q - q_0) \\ 0 &\approx f_2(\mathbf{p}_0) + \frac{\partial f_2}{\partial x}(\mathbf{p}_0)(p - p_0) + \frac{\partial f_2}{\partial y}(\mathbf{p}_0)(q - q_0) \end{aligned}$$

escrito matricialmente

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} \approx \begin{bmatrix} f_1(\mathbf{p}) \\ f_2(\mathbf{p}) \end{bmatrix} + \begin{bmatrix} \frac{\partial f_1}{\partial x}(\mathbf{p}_0) & \frac{\partial f_1}{\partial y}(\mathbf{p}_0) \\ \frac{\partial f_2}{\partial x}(\mathbf{p}_0) & \frac{\partial f_2}{\partial y}(\mathbf{p}_0) \end{bmatrix} \begin{bmatrix} p - p_0 \\ q - q_0 \end{bmatrix}$$

si la matriz $\begin{bmatrix} \frac{\partial f_1}{\partial x}(\mathbf{p}_0) & \frac{\partial f_1}{\partial y}(\mathbf{p}_0) \\ \frac{\partial f_2}{\partial x}(\mathbf{p}_0) & \frac{\partial f_2}{\partial y}(\mathbf{p}_0) \end{bmatrix}$ es invertible

$$\begin{bmatrix} p \\ q \end{bmatrix} \approx \begin{bmatrix} p_0 \\ q_0 \end{bmatrix} - \begin{bmatrix} \frac{\partial f_1}{\partial x}(\mathbf{p}_0) & \frac{\partial f_1}{\partial y}(\mathbf{p}_0) \\ \frac{\partial f_2}{\partial x}(\mathbf{p}_0) & \frac{\partial f_2}{\partial y}(\mathbf{p}_0) \end{bmatrix}^{-1} \begin{bmatrix} f_1(\mathbf{p}) \\ f_2(\mathbf{p}) \end{bmatrix}.$$

Notemos que la matriz anterior, no es más que la **matriz Jacobiana** asociada al campo vectorial $\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \end{bmatrix}$ que

$$\text{denotaremos } \mathbf{JF}(\mathbf{x}) := \begin{bmatrix} \frac{\partial f_1}{\partial x}(\mathbf{x}) & \frac{\partial f_1}{\partial y}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x}(\mathbf{x}) & \frac{\partial f_2}{\partial y}(\mathbf{x}) \end{bmatrix}.$$

Procediendo de manera similar para n , la ecuación de **iteración de Newton para sistemas no lineales** esta dada por

$$\mathbf{x}^{(\mathbf{k}+\mathbf{1})} = \mathbf{x}^{(\mathbf{k})} - \left[\mathbf{JF}(\mathbf{x}^{(\mathbf{k})}) \right]^{-1} \mathbf{F}(\mathbf{x}^{(\mathbf{k})}), \quad k = 0, 1, \dots$$

$$\text{donde } \mathbf{F}(\mathbf{x}) := \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix}, \mathbf{JF}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{x}) & \cdots & \frac{\partial f_n}{\partial x_n}(\mathbf{x}) \end{bmatrix} \text{ y } \mathbf{JF}(\mathbf{x}^{(\mathbf{k})}) \text{ invertible, } k = 0, 1, \dots$$

Se suele realizar la iteración en pos pasos

$$\mathbf{x}^{(\mathbf{k}+\mathbf{1})} = \mathbf{x}^{(\mathbf{k})} + \mathbf{y}^{(\mathbf{k})} \quad \text{donde } \mathbf{y}^{(\mathbf{k})} \text{ es solución de } \mathbf{JF}(\mathbf{x}^{(\mathbf{k})})\mathbf{y}^{(\mathbf{k})} = -\mathbf{F}(\mathbf{x}^{(\mathbf{k})}).$$

Ejemplo Aproximar la solución del sistema

$$\begin{aligned} \ln(x^2 + y^2) - \sin(xy) &= \ln 2 + \ln \pi \\ e^{x-y} + \cos(xy) &= 0 \end{aligned}$$

tomando $\mathbf{x}^{(0)} = [2, 2]^T$ hasta que $\|\mathbf{F}(\mathbf{x}^{(k)})\|_\infty < 10^{-3}$.

Solución: El campo vectorial del sistema esta dado por

$$\mathbf{F}(\mathbf{x}) = \begin{bmatrix} f_1(x, y) \\ f_2(x, y) \end{bmatrix} = \begin{bmatrix} \ln(x^2 + y^2) - \sin(xy) - \ln 2 - \ln \pi \\ e^{x-y} + \cos(xy) \end{bmatrix}$$

así la matriz Jacobiana asociada a \mathbf{F} es

$$\mathbf{JF}(\mathbf{x}) = \begin{bmatrix} \frac{2x}{x^2 + y^2} - y \cos(xy) & \frac{2y}{x^2 + y^2} - x \cos(xy) \\ e^{x-y} - y \sin(xy) & -e^{x-y} - x \sin(xy) \end{bmatrix}$$

Tomando $\mathbf{x}^{(0)} = [2, 2]^T$

$$\mathbf{F}(\mathbf{x}^{(0)}) \approx \begin{bmatrix} 0.9964 \\ 0.3464 \end{bmatrix} \quad \text{y} \quad \mathbf{JF}(\mathbf{x}^{(0)}) \approx \begin{bmatrix} 1.8073 & 1.8073 \\ 2.5136 & 0.5136 \end{bmatrix}.$$

Recordemos que $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. Luego $\mathbf{y}^{(0)}$ es la solución del sistema

$$\begin{bmatrix} 1.8073 & 1.8073 \\ 2.5136 & 0.5136 \end{bmatrix} \mathbf{y}^{(0)} = - \begin{bmatrix} 0.9984 \\ 0.3464 \end{bmatrix} \quad \Rightarrow \quad \mathbf{y}^{(0)} \approx \begin{bmatrix} -0.03134 \\ -0.5211 \end{bmatrix}$$

por lo tanto

$$\mathbf{x}^{(1)} \approx \begin{bmatrix} 2 \\ 2 \end{bmatrix} + \begin{bmatrix} -0.03134 \\ -0.5211 \end{bmatrix} \approx \begin{bmatrix} 1.96866 \\ 1.4789 \end{bmatrix}$$

se tiene que $\|\mathbf{F}(\mathbf{x}^{(1)})\|_{\infty} \approx 0.6583$ y se puede mostrar que la cuarta iteración $\mathbf{x}^{(4)} \approx \begin{bmatrix} 1.7725 \\ 1.7725 \end{bmatrix}$ y se cumple que $\|\mathbf{F}(\mathbf{x}^{(4)})\|_{\infty} \approx 2.6875 \times 10^{-5}$ (**ejercicio**).