

# Apuntes de teórico - Métodos Numéricos



Instituto de Matemáticas y Estadística “Rafael Laguardia”  
Facultad de Ingeniería, Universidad de la República  
2016 - Montevideo, Uruguay

**Nota importante:** El presente material forma parte de una versión **en proceso de revisión** de un texto teórico para la asignatura “Métodos Numéricos”. Por lo tanto, para preparar los exámenes se debe utilizar el material de las clases teóricas y la **bibliografía recomendada**.



# Índice general

<b>3. Ecuaciones no lineales</b>	<b>1</b>
3.1. Ecuaciones no lineales en $\mathbb{R}$	1
3.1.1. Métodos de punto fijo	1
3.1.2. Método de Bipartición (o Bisección)	3
3.1.3. Método de Newton-Raphson	5
3.1.4. Método de la secante	7
3.1.5. Método de la regla falsa (o falsa posición)	8
3.2. Métodos Iterativos Generales	10
3.2.1. Condiciones de parada	14
3.3. Sistemas de Ecuaciones no lineales	14
3.3.1. Métodos Iterativos en las variables	14
3.3.2. Métodos Iterativos Generales	15
3.3.3. Newton-Raphson	15
3.3.4. Método de Newton amortiguado	17
3.3.5. Método de Newton modificado	17
3.3.6. Método de Steffensen	17
3.3.7. Método de Broyden*	18



## Capítulo 3

# Ecuaciones no lineales

### 3.1. Ecuaciones no lineales en $\mathbb{R}$

Consideramos  $f : \mathbb{R} \rightarrow \mathbb{R}$  función no lineal, queremos encontrar una raíz de  $f$ . O sea,  $x^* \in \mathbb{R}$  tal que  $f(x^*) = 0$ .

Si existe y conocemos la función inversa de  $f$ , la podemos aplicar a la ecuación no lineal y obtener la raíz:

$$f^{-1}(f(x^*)) = f^{-1}(0) \Rightarrow x^* = f^{-1}(0)$$

En general no hay manera de hallar la inversa, o es muy costoso. Por lo que buscamos una aproximación de la solución mediante algún método numérico.

#### 3.1.1. Métodos de punto fijo

Tenemos  $f(x^*) = 0$ ,  $f : \mathbb{R} \rightarrow \mathbb{R}$ .

Supongamos que existe y conozco  $g : \mathbb{R} \rightarrow \mathbb{R}$  tal que:

$$f(x) = x - g(x) \quad \forall x \in \mathbb{R}$$

entonces

$$f(x^*) = 0 \iff x^* - g(x^*) = 0 \iff x^* = g(x^*)$$

Genero el método iterativo que tiene a  $x^*$  como punto fijo:

$$\begin{cases} x_{k+1} = g(x_k) \\ x_0 \in \mathbb{R} \end{cases}$$

con  $x_0$  cercano a la raíz  $x^*$ .

**Ejemplo 3.1.1.** Estudiaremos el comportamiento de los métodos iterativos para la función  $f(x) = \log(x^2) - x/2$ . La solución más próxima a cero es:  $x^* = 1,4296\dots$

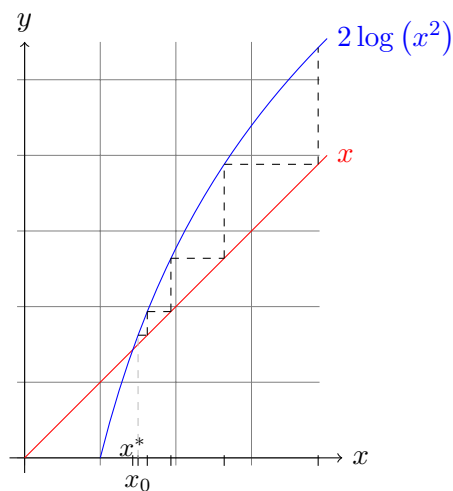


Figura 3.1: Primer método de punto fijo para la ecuación  $\log(x^2) = x/2$

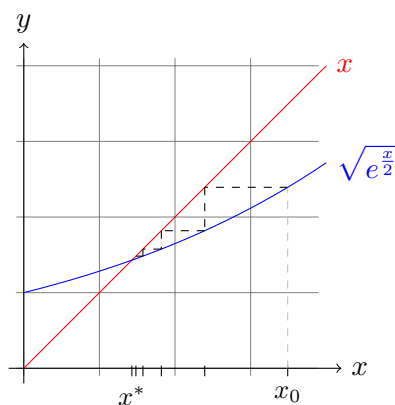


Figura 3.2: Segundo método de punto fijo para la ecuación  $\log(x^2) = x/2$

1.

$$\log(x^2) - x/2 = 0 \iff 2\log(x^2) = x$$

tomamos  $g_1(x) = 2\log(x^2)$  y obtenemos el método de punto fijo, que no converge, como se ve en la Figura 3.1:

$$\begin{cases} x_{k+1} &= 2\log(x_k^2) \\ x_0 &= 1,5 \end{cases}$$

2.

$$\log(x^2) - x/2 = 0 \iff 2\log(x^2) = x \iff x^2 = e^{\frac{x}{2}} \iff x = +\sqrt{e^{\frac{x}{2}}}$$

tomamos ahora la función  $g_2(x) = \sqrt{e^{\frac{x}{2}}}$  y obtenemos el método:

$$\begin{cases} x_{k+1} &= \sqrt{e^{x_k/2}} \\ x_0 &= 1,5 \end{cases}$$

Como se puede ver en la Figura 3.2,  $|g'(x^*)| < 1$ ,  $x^*$  es aproximadamente 1,4296, y se observa también que el método converge.

Profundizaremos sobre estos métodos en la Sección 3.2.

### 3.1.2. Método de Bipartición (o Bisección)

Antes de comentar el método recordemos el teorema de Bolzano:

**Teorema 3.1.1** (Teorema de Bolzano). *Sea  $f : [a, b] \rightarrow \mathbb{R}$  función continua tal que  $f(a) \cdot f(b) < 0$  entonces existe  $x^* \in [a, b]$  tal que  $f(x^*) = 0$ .*

El teorema nos da una idea para un algoritmo, supongamos que  $f$  es continua en un intervalo  $I_0 = [a_0, b_0]$  tal que  $f(a_0) \cdot f(b_0) < 0$ , esto significa que en los extremos del intervalo la función tiene distinto signo, por lo que debe haber una raíz de  $f$  en  $I_0$ . Recursivamente generamos sub intervalos de  $I_0$  tomando su punto medio en cada iteración que llamaremos  $m_k$ , tales que estén en la hipótesis del teorema de Bolzano y por lo tanto contienen siempre una raíz.

---

#### Algoritmo 1 Algoritmo de Bipartición

---

Sea  $f$  función continua en  $I_0 = [a_0, b_0]$ , tal que  $f(a_0) \cdot f(b_0) < 0$  y  $N$  una cantidad máxima de iteraciones a realizar.

```

for  $k = 0 \rightarrow N$  do
   $m_k \leftarrow (a_k + b_k)/2$ 
  if  $f(m_k) = 0$  then
    return  $m_k$ 
  else if  $f(m_k) \cdot f(a_k) > 0$  then
     $I_{k+1} = [a_{k+1}, b_{k+1}] \leftarrow [m_k, b_k]$ 
  else  $\{f(m_k) \cdot f(a_k) < 0\}$ 
     $I_{k+1} = [a_{k+1}, b_{k+1}] \leftarrow [a_k, m_k]$ 
  end if
end for

```

---

*Observación 3.1.1.* Para el algoritmo anterior:

$$|x^* - m_k| \leq |\max\_error(k)| = \frac{b_k - a_k}{2} = \frac{b_{k-1} - a_{k-1}}{2^2} = \dots = \frac{b_0 - a_0}{2^k} \quad \forall k > 0.$$

Por lo que  $\lim_k |x^* - m_k| \leq 0$  y así:

$$\lim_k m_k = x^*$$

y para un  $k$  grande,  $m_k$  es una buena aproximación de  $x^*$ .

*Observación 3.1.2.* Si la función no es continua, el método puede converger a un punto de discontinuidad. (ej: piense en la función como en la Figura 3.3).

#### Ejercicio 3.1.1.

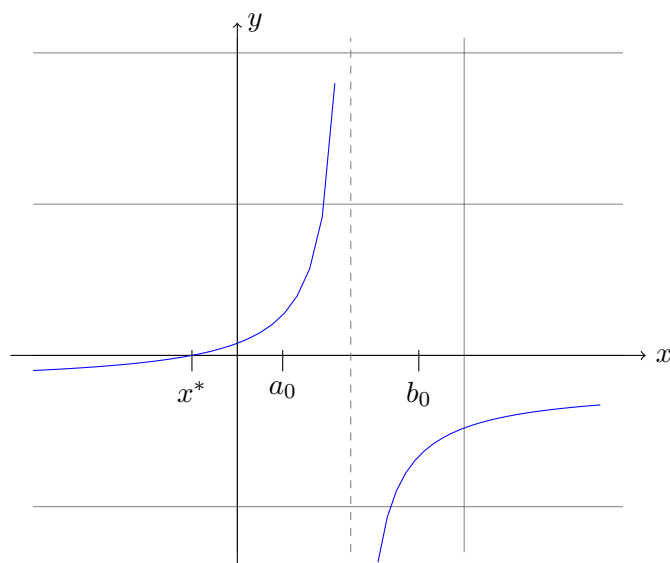


Figura 3.3: Ejemplo bipartición en función discontinua

- Sea  $x^* \in [a, b]$  raíz de  $f$ , hallar la cantidad de iteraciones  $\hat{k}$  necesarias para que el error usando el método de bipartición sea menor o igual que  $10^{-n}$ .
- Si  $[a, b] = [1, 2]$  y el error debe ser menor a  $10^{-5}$ , determine  $\hat{k}$ .
- ¿Puede deducir cuántas iteraciones se requieren para ganar un dígito decimal de precisión?

**Definición 3.1.1** (Orden y velocidad de convergencia). Sea  $\{x_k\}_{k \in \mathbb{N}} \subset \mathbb{R}^n$  una sucesión convergente a  $\alpha \in \mathbb{R}^n$ . Sean  $p > 0$  y  $\beta > 0$  tales que:

$$\lim_{k \rightarrow \infty} \frac{\|x_{k+1} - \alpha\|}{\|x_k - \alpha\|^p} = \beta$$

Decimos que

- $p$  es el *orden* de convergencia de la sucesión.
- $\beta$  es la *velocidad* de convergencia.
- Si  $p = 1$  decimos que la sucesión tiene convergencia lineal.
- Si  $p > 1$  decimos que la sucesión tiene convergencia supralineal.
- Si  $p = 2$  decimos que la sucesión tiene convergencia cuadrática, etc.

### Ejemplo 3.1.2.

1. Sea la sucesión  $a_k = \frac{1}{2^k}$ . Tiene convergencia lineal a 0, con velocidad 1/2.
2.  $b_0 = 1, b_1 = 1, b_2 = \frac{1}{4}, b_3 = \frac{1}{4}, b_4 = \frac{1}{16}, \dots, b_k = \frac{1}{4^{\text{floor}(k/2)}}$ . La sucesión  $b_k$  converge a 0, pero no entra dentro de la definición.



3.  $c_k = \frac{1}{2^{2^k}}$ , tiene convergencia cuadrática y velocidad 1.

*Observación 3.1.3.* Se demuestra que cuanto mayor es  $p$  mas rápido converge la sucesión y cuanto menor es  $\beta$  mas rápido converge la sucesión.

Observamos que para el método de bisección tenemos que en cada paso el máximo error se divide por dos y:

$$\lim_k \frac{|x^* - m_k|}{|x^* - m_{k-1}|} \leq \frac{1}{2}$$

por lo que el método tiene convergencia lineal y velocidad a lo sumo  $1/2$ .

### 3.1.3. Método de Newton-Raphson

Supongamos ahora que la función con la que estamos trabajando es diferenciable. Veamos la deducción geométrica del método de Newton-Raphson.

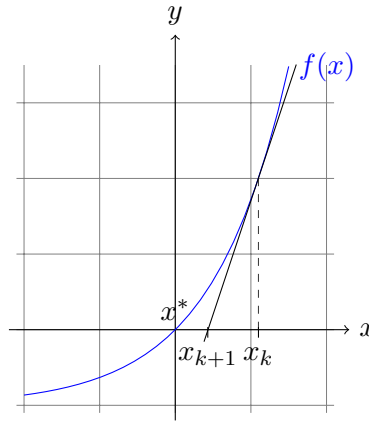


Figura 3.4: Método de Newton-Raphson

Dado  $x_k \in \mathbb{R}$ , hallamos la recta tangente a  $f$  en  $x_k$  y el punto de corte con el eje  $x$ , es  $x_{k+1}$ . La recta tangente a  $f$  en  $x_k$  es

$$r : y - f(x_k) = f'(x_k)(x - x_k)$$

si imponemos  $y = 0$ , se obtiene  $x = x_k - \frac{f(x_k)}{f'(x_k)}$ . Con lo que llegamos al método de Newton-Raphson:

$$(N-R): \begin{cases} x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \\ x_0 \in \mathbb{R} \end{cases}$$

que se puede ver en la Figura 3.4.

**Ejemplo 3.1.3.** Sea  $f(x) = x^2 - 2$ , una de las raíces de esta función es  $x^* = \sqrt{2}$ . Si aplicamos Newton-Raphson a  $f$ , con punto inicial 1, obtenemos la siguiente iteración:

$$\begin{cases} x_{k+1} = x_k - \frac{x_k^2 - 2}{2x_k} \\ x_0 = 1 \end{cases}$$

En la siguiente tabla vemos los valores de la iteración para ciertos  $k$  y sus respectivos errores.

$k$	$x_k$	orden del error
0	1	-1
1	$3/2 = 1.5$	-2
2	$17/2 \approx 1,41117$	-3
3	$577/408 \approx 1,41421$	-6
4	$665857/470832 \approx 1,41421$	-12
5	-	-25
6	-	-49

Veamos una teorema que nos garantiza la convergencia del método de Newton-Raphson bajo ciertas hipótesis.

**Teorema 3.1.2.** *Sea  $f : I \rightarrow \mathbb{R}$  función con derivada segunda continua, tal que  $f(x^*) = 0$ ,  $f'(x^*) \neq 0$  y  $f''(x^*) \neq 0$ . Entonces la sucesión generada por Newton-Raphson converge a  $x^*$  cuadráticamente y con velocidad  $\left| \frac{f''(x^*)}{2f'(x^*)} \right|$ , siempre que  $x_0$  se elija suficientemente próximo a  $x^*$ .*

*Demostración.*

- Postergaremos la demostración de la convergencia para más adelante, una vez hayamos desarrollado más la teoría. En particular, obtendremos este resultado como aplicación del Corolario 3.2.4. Ver Proposición 3.2.5.

- Veamos el orden y velocidad de convergencia:

Hallamos el desarrollo de Taylor de  $f$  en un entorno de  $x_k$  y evaluando en  $x^*$ , obtenemos

$$0 = f(x^*) = f(x_k) + f'(x_k)(x^* - x_k) + \frac{f''(\xi_k)}{2}(x^* - x_k)^2, \quad \xi_k \in (x^*, x_k)$$

por lo que, dividiendo entre  $f'(x_k)$

$$x^* - x_{k+1} = x^* - \left( x_k - \frac{f(x_k)}{f'(x_k)} \right) = -\frac{f''(\xi_k)}{2f'(x_k)}(x^* - x_k)^2, \quad \xi_k \in (x^*, x_k)$$

y

$$\frac{|x^* - x_{k+1}|}{|x^* - x_k|^2} = \left| \frac{f''(\xi_k)}{2f'(x_k)} \right|, \quad \xi_k \in (x^*, x_k)$$

finalmente, tomando limite

$$\lim_k \frac{|x^* - x_{k+1}|}{|x^* - x_k|^2} = \lim_k \left| \frac{f''(\xi_k)}{2f'(x_k)} \right| = \left| \frac{f''(x^*)}{2f'(x^*)} \right|$$

ya que  $\lim_k f''(\xi_k) = f''(x^*)$ ,  $\lim_k f'(x_k) = f'(x^*)$  por continuidad de  $f''$  y  $f'$ , y  $\lim_k \xi_k = x^*$ .

□

Se deja como ejercicio probar:

- Si  $f'(x^*) = 0$  el orden es 1.
- Si  $f'(x^*) \neq 0$  y  $f''(x^*) = 0$  el orden es mayor o igual a 3.

### 3.1.4. Método de la secante

El método de la secante es una modificación del método de Newton-Raphson, en vez de hallar la recta tangente a un punto, aproximamos la misma por una recta secante. Ver Figura 3.5.

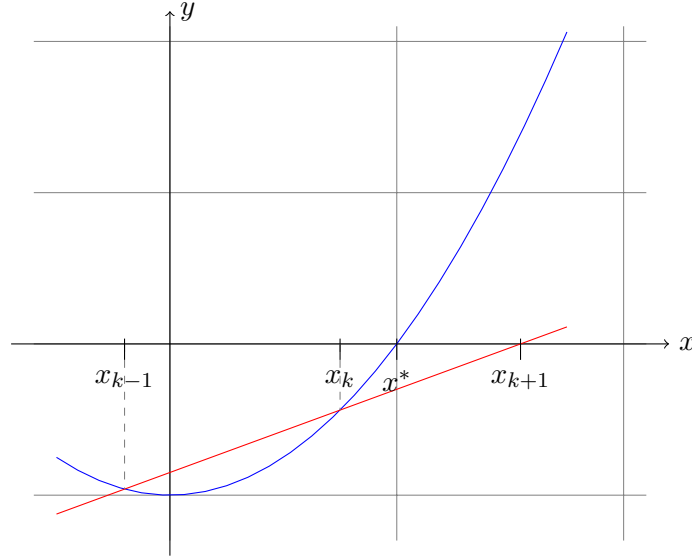


Figura 3.5: Método de la secante

Geométricamente, dados dos puntos de aproximación  $x_{k-1}$ ,  $x_k$ , se halla la intersección de la recta que pasa por los puntos  $(x_{k-1}, f(x_{k-1}))$  y  $(x_k, f(x_k))$  con el eje  $x$ .

La ecuación de la recta es:

$$r : y - f(x_k) = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}(x - x_k)$$

cuya intersección con el eje  $x$  es:  $x = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k)$ .

Obtenemos así el método de la secante:

$$\begin{cases} x_{k+1} = x_k - \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})} f(x_k) \\ x_0, x_1 \in \mathbb{R} \end{cases}$$

*Observación 3.1.4.*

- El método de la secante genera una sucesión de orden 2, es decir, que para el cálculo de  $x_{k+1}$  que requieren los valores de dos pasos anteriores ( $x_k$  y  $x_{k-1}$ ).

- Reescribiendo el método:

$$x_{k+1} = x_k - \frac{f(x_k)}{\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}}$$

Si  $x_k \approx x_{k-1} \Rightarrow$ , es posible pensarlo como una variación del método de N-R, aproximando la derivada por medio de diferencias finitas, y esto es especialmente útil cuando el costo

computacional de derivar la función es elevado. Por tanto, si bien el orden de convergencia es menor que el de Newton, al considerar el costo de evaluar  $f'(x)$  puede resultar ventajoso usar la secante.

- No se asegura convergencia si los valores iniciales son demasiado lejanos o la raíz no es simple.

Vemos ahora un teorema sobre la convergencia del método de la secante, que no demostraremos.

**Teorema 3.1.3.** *Sea  $f : I \rightarrow \mathbb{R}$  función con derivada segunda continua, tal que  $f(x^*) = 0$ ,  $f'(x^*) \neq 0$ . Entonces la sucesión generada por el método de la secante converge a  $x^*$  siempre que  $x_0$  se elija suficientemente próximo a  $x^*$  y en ese caso la convergencia es supralineal con  $p = \frac{1+\sqrt{5}}{2} \approx 1,618$ .*

*Observación 3.1.5.* Destacamos como curiosidad que  $p = \frac{1+\sqrt{5}}{2}$  es la razón áurea.

### 3.1.5. Método de la regla falsa (o falsa posición)

El método de la Regla Falsa (o Falsa posición) requiere que la función  $f$  sea continua y requiere dos puntos iniciales  $a$  y  $b$  tal que sus valores funcionales son de distinto signo.

Es similar al de la secante o al de bipartición en el sentido de que el intervalo  $[a_n, b_n]$  se va actualizando.

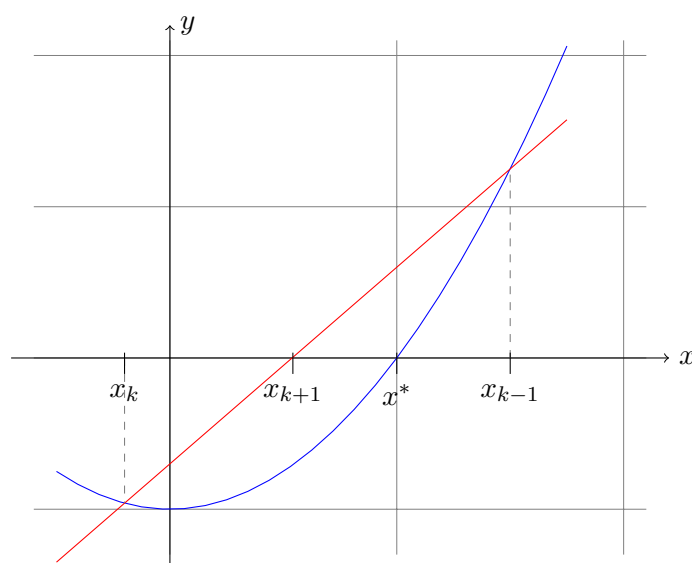


Figura 3.6: Método de la regla falsa

El procedimiento es el siguiente (ver Figura 3.6):

Se toma la cuerda dada por los puntos  $(a, f(a))$  y  $(b, f(b))$ . Se interseca con el eje  $x$ , es decir, que se hace  $y = 0$  para obtener el punto  $x_1$ . Se genera un nuevo segmento  $[a, b]$ , sustituyendo o  $a$  o  $b$  por  $x_1$  según sea el signo de  $f(x_1)$  de forma que siempre  $f(a)$  y  $f(b)$  tengan signos opuestos.

Es decir:

Cuerda  $(a, f(a)) - (b, f(b))$ :

$$y = mx + d \quad \begin{cases} m = \frac{f(b)-f(a)}{b-a} \\ d = \frac{af(b)-bf(a)}{a-b} \end{cases}$$

Haciendo  $y = 0$ :

$$x = \frac{-d}{m} = \frac{af(b) - bf(a)}{f(a) - f(b)} = x_1$$

Si  $\begin{cases} f(a)f(x_1) < 0 \Rightarrow \text{Repetimos en } [a, x_1] \\ f(x_1)f(b) < 0 \Rightarrow \text{Repetimos en } [x_1, b] \end{cases}$

A diferencia del método de la secante, en que se debería utilizar la cuerda formada por  $x_k$  y  $x_{k+1}$  para construir el siguiente punto, en el método de la falsa regla debemos utilizar siempre puntos tal que sus valores funcionales sean de distinto signo. Por lo que, si tomamos como ejemplo la Figura 3.6, como  $x_k$  y  $x_{k+1}$  toman valores funcionales negativos, debemos buscar un punto anterior de la sucesión cuyo valor funcional sea positivo. En el ejemplo, utilizamos  $x_{k+1}$  con  $x_{k-1}$  y con ellos construimos la cuerda que generará el punto  $x_{k+2}$

*Observación 3.1.6.*

- El método es siempre convergente (para funciones continuas).
- Sin embargo es de orden 1.
- Es aconsejable su uso mientras se está lejos de la raíz pero conviene cambiar a otro cerca de ella.
- Pese a ser similar al método de la secante, es más complejo, tiene menor orden, pero es globalmente convergente.

**Ejercicio 3.1.2.** Sea la función  $f(x) = x^2 - 1$ ;  $x_1 = 1, 5$ ;  $x_2 = 0, 2$ .

- Hallar  $x_3$  y  $x_4$  usando los 4 métodos vistos.
- Determinar si todos ellos convergen.
- Fundamentar cuál es el más conveniente.
- Varía de algún modo lo anterior si en lugar de  $x_2 = 0, 2$  se toma  $x_2 = -0, 2$ .

Para concluir remarcamos que existe toda una galería de métodos aplicables a funciones. Aquí solamente se han presentado las más generales y/o sencillas.

## 3.2. Métodos Iterativos Generales

Sea  $g : \mathbb{R} \rightarrow \mathbb{R}$  función, y el método:

$$(M) \begin{cases} x_{k+1} = g(x_k) \\ x_0 \in \mathbb{R} \end{cases}$$

que genera una sucesión  $\{x_k\}_{k \in \mathbb{Z}}$ .

**Ejemplo 3.2.1.** En N-R,  $g(x) = x - \frac{f(x)}{f'(x)}$ .

**Definición 3.2.1.** Decimos que  $\alpha \in \mathbb{R}$  es punto fijo de (M) si  $g(\alpha) = \alpha$ .

*Observación 3.2.1.* Si  $\alpha \in \mathbb{R}$  es punto fijo de  $g(x)$ , entonces es raíz de  $f(x) = g(x) - x$ .

**Ejercicio 3.2.1.** Hallar los puntos fijos de  $g(x) = 2 - x^2$ .

**Definición 3.2.2.** La función  $g$  es contractiva en  $I \subset \mathbb{R}$  si existe  $0 \leq m < 1$  tal que

$$|g(x) - g(y)| \leq m \cdot |x - y|$$

para todos  $x, y \in I$ .

**Ejemplo 3.2.2.** La función  $x^2$  no es contractiva en  $\mathbb{R}$  pues tomando  $x = 3$ ,  $y = 2$ , no es posible hallar  $m < 1$  que verifique la definición.

Sin embargo, sí lo es en el intervalo  $(-\frac{1}{4}, \frac{1}{4})$ :

$|x^2 - y^2| = |x + y||x - y|$  y como  $|x + y| < \frac{1}{2}$ , basta tomar  $m = \frac{1}{2}$ .

**Ejercicio 3.2.2.** Dar otro ejemplo de una función contractiva en algún intervalo y probarlo.

*Observación 3.2.2.*

- Toda contracción es continua.
- Más aún, toda contracción es Lipschitz (con constante de Lipschitz  $m$ ), y por tanto, uniformemente continua.

**Teorema 3.2.1** (Punto Fijo). *Si  $X$  es un espacio métrico completo y  $\varphi : X \rightarrow X$  una  $m$ -contracción con  $m < 1$ , entonces la sucesión  $\{x_k\}_{k \in \mathbb{N}}$  tal que  $x_{k+1} = \varphi(x_k)$  converge al punto fijo de  $\varphi$ , que además es único. El resultado no depende del elemento inicial  $x_0 \in X$ .*

*Demostración.* Vamos a probar el Teorema 3.2.1 en cuatro etapas:

- i Toda  $m$ -contracción es continua: sea  $f : X \rightarrow X$  una  $m$ -contracción y  $\{x_k\}_{k \in \mathbb{N}}$  una sucesión en  $X$  tal que  $x_k \rightarrow x$ . Basta ver que  $f(x_k) \rightarrow f(x)$ . Efectivamente, como el espacio admite una métrica  $d$  tenemos que  $d(f(x_k), f(x)) \leq md(x_k, x) \rightarrow 0$ . En particular, tenemos por hipótesis que  $\varphi$  es continua.

- ii Probemos ahora que la sucesión  $\{x_k\}_{k \in \mathbb{N}}$  es de Cauchy. Sea  $\varepsilon > 0$  arbitrario. Veamos que existe  $k_0$  tal que si  $l \geq k \geq k_0$  entonces  $d(x_l, x_k) < \varepsilon$ . Por definición de la sucesión y la desigualdad triangular de una métrica, tenemos que:

$$\begin{aligned} d(x_l, x_k) &= d(\varphi^{(l)}(x_0), \varphi^{(k)}(x_0)) \leq \sum_{i=0}^{l-k-1} d(\varphi^{(l-i)}(x_0), \varphi^{(l-i-1)}(x_0)) \\ &\leq \sum_{i=0}^{l-k-1} m^{l-i-1} d(x_0, x_1) = d(x_0, x_1) \sum_{j=k}^{l-1} m^j = d(x_0, x_1) \frac{m^l - m^k}{m - 1} \\ &= d(x_0, x_1) \frac{m^k}{1 - m} (1 - m^{l-k}) < \varepsilon, \end{aligned}$$

eligiendo  $l \geq k$  suficientemente grande. Luego,  $\{x_k\}_{k \in \mathbb{N}}$  es de Cauchy. Como  $X$  es completo, converge a cierto elemento  $\alpha$  perteneciente al espacio  $X$ :  $\lim_k x_k = \alpha \in X$ .

- iii Veamos ahora  $\alpha$  es además punto fijo de  $\varphi$ , es decir que  $\varphi(\alpha) = \alpha$ . De hecho, por continuidad de  $\varphi$  tenemos que:

$$\varphi(\alpha) = \varphi(\lim_k x_k) = \lim_k \varphi(x_k) = \lim_k x_{k+1} = \alpha.$$

- iv Finalmente, probemos la unicidad del punto fijo: si  $\beta = \varphi(\beta)$ , entonces:

$$d(\alpha, \beta) = d(\varphi(\alpha), \varphi(\beta)) \leq m d(\alpha, \beta),$$

o equivalentemente:

$$d(\alpha, \beta)(1 - m) \leq 0.$$

Como  $1 - m > 0$  y  $d(\alpha, \beta) \geq 0$ , la única posibilidad es que  $d(\alpha, \beta) = 0$ , y como  $d$  es una métrica entonces  $\alpha = \beta$ .

Hemos probado así que en un espacio métrico completo  $X$ , la reiterada aplicación de una  $m$ -contracción a un punto inicial cualquiera  $x_0 \in X$  genera una sucesión que converge siempre al único punto fijo  $\alpha$  de la contracción, siempre que  $m < 1$ .  $\square$

**Teorema 3.2.2** (Convergencia). *Supongamos que existe  $\alpha$  punto fijo de  $(M)$  y  $\delta > 0$  tal que  $g$  es contractiva en  $B_{\alpha, \delta}$  con constante  $m$ . Luego, para todo  $x_0 \in B_{\alpha, \delta}$  se cumple:*

1.  $x_k \in B_{\alpha, \delta} \ \forall k \in \mathbb{N}$ .
2.  $\lim_k x_k = \alpha$ .
3.  $\alpha$  es el único punto fijo de  $(M)$  en  $B_{\alpha, \delta}$ .

*Demostración.*

1. Hacemos inducción en  $k$ , sabemos que  $x_0 \in B_{\alpha, \delta}$ . Si  $x_k \in B_{\alpha, \delta}$

$$|\alpha - x_{k+1}| = |g(\alpha) - g(x_k)| \leq m \cdot |\alpha - x_k| < 1 \cdot \delta = \delta$$

y  $x_{k+1} \in B_{\alpha, \delta}$ .

2. Sabemos que  $|\alpha - x_k| \leq m \cdot |\alpha - x_{k-1}| \leq m^2 \cdot |\alpha - x_{k-2}| \leq \dots \leq m^k \cdot |\alpha - x_0|$  por lo que  $\lim_k |\alpha - x_k| \leq \lim_k m^k \cdot |\alpha - x_0| = 0$ , ya que  $m < 1$ . Concluimos que  $\lim_k x_k = \alpha$ .

3. Supongo que existe  $\beta \in B_{\alpha, \delta}$  otro punto fijo de (M), vemos que

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| \leq m \cdot |\alpha - \beta| < |\alpha - \beta|$$

que es absurdo salvo que  $\alpha = \beta$ .

□

**Lema 3.2.3.** Si  $g$  es derivable en un entorno de  $\alpha$  y  $|g'(x)| \leq m$ , con  $m < 1$ , para todo  $x \in B_{\alpha, \delta}$  entonces  $g$  es contractiva en  $B_{\alpha, \delta}$ .

*Demostración.* Sean  $x, y \in B_{\alpha, \delta}$ , por el teorema del valor medio, existe  $\xi \in (x, y)$  tal que  $|g(x) - g(y)| = |g'(\xi)| \cdot |x - y|$ . Y como  $|g'(\xi)| \leq m$  queda demostrado el lema. □

Veamos una consecuencia directa del lema anterior.

**Corolario 3.2.4.** Si  $\alpha$  es punto fijo de (M),  $g \in \mathcal{C}^1$  y  $|g'(x)| \leq m < 1$  para todo  $x \in B_{\alpha, \delta}$  con  $\delta > 0$ , entonces:

1.  $x_k \in B_{\alpha, \delta}$  para todo  $k \in \mathbb{N}$ .
2.  $\lim_k x_k = \alpha$ .
3.  $\alpha$  es el único punto fijo de (M) en  $B_{\alpha, \delta}$ .

**Ejemplo 3.2.3.** Volvamos al ejemplo 3.1.1, tenemos la ecuación

$$\log(x^2) - \frac{x}{2} = 0$$

con solución aproximada  $x^* \approx 1,4296$ .

En el primer método iterativo usábamos la función  $g_1(x) = 2 \log(x^2)$ . Vemos que  $|g'_1(x^*)| = 4/x^* \approx 2,79796 > 1$ , por lo cual el método no converge.

En el segundo método iterativo usamos la función  $g_2(x) = \sqrt{e^{x/4}}$ , con  $|g'_2(x^*)| \approx 0,35740 < 1$  por lo que vemos que el método converge.

**Proposición 3.2.5.** La sucesión generada por Newton-Raphson en las hipótesis del Teorema 3.1.2 converge a  $x^*$ .

*Demostración.* Si denotamos  $g(x) = x - \frac{f(x)}{f'(x)}$  entonces el método de Newton-Raphson es un método de punto fijo con función  $g$ . Derivando  $g$ :

$$g'(x) = 1 - \frac{f'(x)f'(x) - f''(x)f(x)}{f'(x)^2}$$

y vemos que  $g'(x^*) = 0$ . Entonces,  $x^*$  es un atractor para  $g$ , por lo que existe  $\varepsilon > 0$  tal que si  $x_0 \in (x^* - \varepsilon, x^* + \varepsilon)$  la sucesión generada por el método converge. □



**Teorema 3.2.6.** Si  $g \in \mathcal{C}^p$ ,  $p \in \mathbb{N}^+$ .  $(M)$  es de orden  $p$  y velocidad  $\frac{|g^{(p)}(\alpha)|}{p!}$  si y solo si  $g^{(i)}(\alpha) = 0$  para  $i = 1, \dots, p-1$  y  $g^{(p)}(\alpha) \neq 0$ .

*Demostración.*  $(\Rightarrow)$  Desarrollamos Taylor a  $g$  en  $\alpha$ :

$$g(x) = g(\alpha) + \overbrace{\sum_{i=1}^{p-1} \frac{g^{(i)}(\alpha)}{i!} (x-\alpha)^i}^{=0} + \frac{g^{(p)}(\xi)}{p!} (x-\alpha)^p = \alpha + \frac{g^{(p)}(\xi)}{p!} (x-\alpha)^p$$

con  $\xi \in (\alpha, x)$ . Evaluando en  $x_k$  vemos que  $x_{k+1} = g(x_k) = \alpha + \frac{g^{(p)}(\xi_k)}{p!} (\alpha - x_k)^p$ ,  $\xi_k \in (\alpha, x_k)$ . Luego

$$\frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^p} = \frac{|g^{(p)}(\xi_k)|}{p!} \rightarrow_k \frac{|g^{(p)}(\alpha)|}{p!} \neq 0$$

ya que  $\lim_k g^{(p)}(\xi_k) = g^{(p)}(\alpha)$  por continuidad.

$(\Leftarrow)$  Si el orden de convergencia del método es  $p$ , entonces si  $0 \leq i < p$  vemos que

$$\lim_k \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^i} = \lim_k \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^p} |x_k - \alpha|^{p-i} = 0$$

ya que  $(p-i) > 0$  y  $\lim_k \frac{|x_{k+1} - \alpha|}{|x_k - \alpha|^p} < \infty$ . Veamos ahora por inducción que  $g^{(i)}(\alpha) = 0$ . Para  $i = 1$ , desarrollamos Taylor a  $g$  en  $\alpha$  y evaluamos en  $x_k$  y obtenemos  $x_{k+1} - \alpha = g'(\xi_k)(x_k - \alpha)$ ,  $\xi_k \in (\alpha, x_k)$ . Luego  $\frac{|x_{k+1} - \alpha|}{|x_k - \alpha|} = |g'(\xi_k)|$ . Y tomando limite probamos que  $g'(\alpha) = 0$ . Usando inducción y Taylor se prueba el resto de las igualdades.  $\square$

A modo de resumen, para una función  $f(x)$  es importante elegir adecuadamente la función  $g(x)$  y el valor inicial  $x_0$ . Si se tienen dos métodos iterativos, serán convergentes si  $|g'(x)| < 1$  en un entorno de la raíz  $x^*$ . El orden de convergencia viene dado por el estudio de la anulación de sus derivadas, elegimos el que tenga más derivadas nulas en  $x^*$ . Si ambos métodos tienen derivada primera no nula, la velocidad es lineal y convendrá usar la iteración en la cual el valor absoluto de  $g(x^*)$  es menor.

**Ejemplo 3.2.4.** Sea  $f(x) = \sin(x) - x^2$  que tiene raíz  $x^* \approx 0,87$ .

Tenemos las siguientes opciones:

- $g_1(x) = \sqrt{\sin(x)} \Rightarrow |g'_1(x)| = \left| \frac{\cos(x)}{2\sqrt{\sin(x)}} \right| < 0,45$  en  $(0,8;0,9)$ . Sirve.
- $g_2(x) = \arcsin(x^2) \Rightarrow |g'_2(x)| = \left| \frac{2x}{\sqrt{1-(x^2)^2}} \right| > 2$  en  $(0,8;0,9)$ . No sirve.
- $g_3(x) = \sin(x) - x^2 + x \Rightarrow |g'_3(x)| = |\cos(x) - 2x + 1| < 0,18$  en  $(0,8;0,9)$ . Sirve.

Es así que tanto  $g_1$  como  $g_3$  son métodos convergentes con velocidad lineal. Ahora  $g'_1(x^*) \approx 0,37$  y  $g'_3(x^*) \approx 0,10$ , por lo que convendrá usar  $g_3$ .

### 3.2.1. Condiciones de parada

Entre las condiciones de parada de los métodos iterativos es posible considerar:

1. Parada por número de iteraciones ( $k > k_{max}$ ). El algoritmo se detiene cuando el número de iteraciones llega a una cantidad fijada de antemano. Es adecuado cuando no se conoce el comportamiento del método para la función  $f$  considerada o también en conjunción con otras condiciones de parada que pudieran no satisfacerse a lo largo de la recursión. La condición evita que el algoritmo entre en un ciclo sin fin.
2. Parada por proximidad a la raíz  $x^*$ . Es adecuada cuando lo que se desea es hallar, con determinada precisión, la raíz  $x^*$ . Se busca detener el algoritmo cuando el error  $|x^{(k)} - x^*|$  o bien el error relativo  $\frac{|x^{(k)} - x^*|}{|x^*|}$ , es menor que cierta tolerancia  $\varepsilon > 0$  fijada de antemano. Como no se conoce a priori la raíz  $x^*$ , se hace la estimación  $x^{(k+1)} \approx x^*$  con lo que el método se detiene si:

$$|x^{(k)} - x^{(k+1)}| < \varepsilon, \quad \text{o bien si} \quad \frac{|x^{(k)} - x^{(k+1)}|}{|x^{(k+1)}|} < \varepsilon.$$

3. Parada por proximidad a la anulación de  $f$ . Se utiliza cuando lo que se pretende es hallar valores de las variables que hagan que  $f(x)$  sea pequeño. La condición de parada es  $|f(x^{(k)})| < \varepsilon$ , donde  $\varepsilon > 0$  es una tolerancia fijada de antemano.

## 3.3. Sistemas de Ecuaciones no lineales

Consideramos ahora un sistema de  $n$  ecuaciones no lineales en  $n$  variables,  $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ ,

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \dots, x_n) = 0 \end{cases}$$

que vectorialmente lo podemos expresar como  $F(\mathbf{x}) = 0$  con  $F = (f_1, f_2, \dots, f_n)$ ,  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$  y  $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ .

### 3.3.1. Métodos Iterativos en las variables

Podemos definir métodos iterativos para resolver sistemas de ecuaciones con las ideas de los métodos de Jacobi, Gauss-Seidel y SOR, dado un método para resolver sistemas no lineales en  $\mathbb{R}$ . En el Algoritmo 2 vemos el método de Jacobi no lineal.

Se puede aplicar la misma idea para los métodos de Gauss-Seidel y SOR.

**Algoritmo 2** Método de Jacobi no lineal

Sea  $(M)$  un método para resolver ecuaciones del tipo  $f(x) = 0$  con  $f : \mathbb{R} \rightarrow \mathbb{R}$ , y  $\mathbf{x}^{(0)} \in \mathbb{R}^n$ .

```

 $k \leftarrow 0$ 
repeat
  for  $j = 1 \rightarrow n$  do
     $x_j^{(k+1)} \leftarrow$  solución mediante el método (M) de  $f_j \left( x_1^{(k)}, x_2^{(k)}, \dots, x_{j-1}^{(k)}, x_j^{(k+1)}, x_{j+1}^{(k)}, \dots, x_n^{(k)} \right) =$ 
    0
  end for
   $k \leftarrow k + 1$ 
until convergencia

```

**3.3.2. Métodos Iterativos Generales**

Generalizamos aquí la idea de los MIG para funciones reales a funciones en  $\mathbb{R}^n$ . Es posible probar los mismos resultados para  $\mathbf{x}^{(k+1)} = g(\mathbf{x}^{(k)})$ , con  $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ , convirtiendo valores absolutos en normas, derivadas en jacobiana. En particular, la convergencia requerirá las mismas hipótesis de continuidad y de dominio para  $g$  y que  $\|J_g(\mathbf{x}^*)\| < 1$ .

**Ejemplo 3.3.1.**

$$f(x_1, x_2) = \begin{pmatrix} e^{x_1+x_2^2} + x_1^2 \\ x_1^2 e^{-x_2} - x_2 x_1^2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Entonces podemos generar un problema de punto fijo mediante:

$$g(x_1, x_2) = \begin{pmatrix} e^{x_1+x_2^2} + x_1^2 + x_1 \\ x_1^2 e^{-x_2} - x_2 x_1^2 + x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

**3.3.3. Newton-Raphson**

Sea  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ,  $F(\mathbf{x}^*) = 0$ ,  $\mathbf{x}^* \in \mathbb{R}^n$ .

Se puede generalizar el método de Newton-Raphson a sistemas de ecuaciones no lineales. La idea es resolver la ecuación linealizada en un punto próximo a la solución.

Si suponemos  $F$  diferenciable en el punto  $\mathbf{x}^{(k)}$ , por el teorema de Taylor:

$$F(\mathbf{x}) = F(\mathbf{x}^{(k)}) + J_F(\mathbf{x}^{(k)})(\mathbf{x} - \mathbf{x}^{(k)}) + O(\|\mathbf{x} - \mathbf{x}^{(k)}\|^2)$$

trunco el termino  $O(\|\mathbf{x} - \mathbf{x}^{(k)}\|^2)$  y resuelvo suponiendo  $F(x) \approx 0$ , obteniendo el método:

$$\begin{cases} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - J_F(\mathbf{x}^{(k)})^{-1} F(\mathbf{x}^{(k)}) \\ x^{(0)} \in \mathbb{R}^n \end{cases}$$

Notamos que el determinante de  $J_F(\mathbf{x}^{(k)})$  debe de ser no nulo para todo  $k$  para poder definir el método. Vemos también que el caso en que  $n = 1$  del método coincide con el método presentado en la Subsección 3.1.3.

En el Algoritmo 3 se puede ver el método de Newton-Raphson general (en el que no se calcula la inversa de  $J_F$ ).

---

**Algoritmo 3** Método de Newton-Raphson
 

---

Sea  $\mathbf{x}^{(0)} \in \mathbb{R}^n$ .

$k \leftarrow 0$

**repeat**

$\mathbf{s}^{(k)} \leftarrow$  solución del sistema lineal  $J_F(\mathbf{x}^{(k)})\mathbf{s}^{(k)} = -F(\mathbf{x}^{(k)})$

$\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} + \mathbf{s}^{(k)}$

$k \leftarrow k + 1$

**until** convergencia

---

**Ejemplo 3.3.2.** Sea  $F(x, y, z) = (xy - z^2, y^2 - xz, 2x^2 - yxz - 1)$  con  $\mathbf{x}^{(0)} = (1, 1, 0)$ . Calculemos  $\mathbf{x}^{(1)}$  aplicando Newton-Raphson.

$J_F(x, y, z) = \begin{pmatrix} y & x & -2z \\ -z & 2y & -x \\ 4x - yz & -xz & -yx \end{pmatrix}$ . Por otro lado  $F(1, 1, 0) = (1, 1, 1)^t$ . Para calcular  $\mathbf{x}^{(1)}$  tenemos el sistema:

$$J_F(1, 1, 0)\mathbf{s}^{(0)} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & -1 \\ 4 & 0 & -1 \end{pmatrix} \begin{pmatrix} s_1^{(0)} \\ s_2^{(0)} \\ s_3^{(0)} \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix} = -F(1, 1, 0).$$

Resolviéndolo tenemos  $\mathbf{s}^{(0)} = (s_1^{(0)}, s_2^{(0)}, s_3^{(0)})^t = (-1/3, -2/3, -1/3)^t$ , y

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s}^{(0)} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -\frac{1}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{pmatrix} = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \\ -\frac{1}{3} \end{pmatrix}.$$

Se deja como ejercicio computar de forma análoga  $\mathbf{x}^{(2)}$  planteando y resolviendo el sistema:  $J_F(\mathbf{x}^{(1)})\mathbf{s}^{(1)} = -F(\mathbf{x}^{(1)})$  y luego realizar  $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{s}^{(1)}$ .

*Observación 3.3.1.*

1. El número de evaluaciones de funciones reales en cada caso es:  $O(n^2)$  para  $J_F$ , y  $O(n)$  para  $F$ .
2. Por cada paso necesito resolver un sistema lineal para el cual tiene un costo  $O(2/3n^3)$  de flops.
3. Hay que calcular  $n^2$  derivadas analíticamente por paso.

Es posible demostrar que el método es de orden 2.

Veremos a continuación algunas variantes.

### 3.3.4. Método de Newton amortiguado

Se varía la iteración y se convierte en:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - \alpha J_F(\mathbf{x}^{(k)})^{-1} F(\mathbf{x}^{(k)})$$

### 3.3.5. Método de Newton modificado

El método de Newton-Raphson es computacionalmente costoso y  $J_F(\mathbf{x}^{(k)})$  puede estar mal condicionada, lo que hace difícil tener una buena aproximación de  $\mathbf{x}^*$ . Por estas razones se pueden definir ciertas modificaciones para evitar dichos problemas. El método de Newton modificado evita calcular la matriz Jacobiana y los  $O(2/3n^3)$  flops por resolución del sistema, dejando constante  $J_F(\mathbf{x}^{(p)})$ . Es decir que  $J_F$  solo se recalcula para algunos pasos. O sea dejamos la matriz Jacobiana, más precisamente su descomposición, constante para una cantidad de pasos  $p \geq 2$ , luego se computa  $J_F(\mathbf{x}^{(2p)})$  y se prosigue con el método sin variar la matriz por otras  $p$  iteraciones.

---

**Algoritmo 4** Método de Newton modificado
 

---

Sea  $\mathbf{x}^{(0)} \in \mathbb{R}^n$ .

$k \leftarrow 0$

**repeat**

$LU \leftarrow J_F(\mathbf{x}^{(kp)})$

**for**  $j = 0 \rightarrow p - 1$  **do**

$\mathbf{s}^{(kp+j)} \leftarrow$  solución del sistema lineal, usando descomposición LU:  $LU \cdot \mathbf{s}^{(kp+j)} = -F(\mathbf{x}^{(kp+j)})$

$\mathbf{x}^{(kp+j+1)} \leftarrow \mathbf{s}^{(kp+j)} + \mathbf{x}^{(kp+j)}$

**end for**

$k \leftarrow k + 1$

**until** convergencia

---

*Observación 3.3.2.*

1. Aunque el método de Newton-Raphson converja, el método de Newton modificado puede no converger.
2. El método de Newton modificado es más lento que el de Newton-Raphson. Esto disminuye el orden de convergencia si  $J_F$  varía mucho (caso en que  $F$  es fuertemente no lineal).

### 3.3.6. Método de Steffensen

El método de Steffensen evita calcular  $J_F$  y la obtiene como cociente incremental:

$$\frac{\partial f_i(\mathbf{x}^{(k)})}{\partial x_j} \approx \frac{f_i(\mathbf{x}^{(k)} + h_j \mathbf{e}_j) - f_i(\mathbf{x}^{(k)})}{h_j}$$

con  $\mathbf{e}_j = [0, 0, \dots, 1, \dots, 0]^t$ , el  $j$ -ésimo vector de la base canónica.

Steffensen toma  $h_j = f_j(\mathbf{x}^{(k)})$ . Así:

$$\frac{\partial f_i(\mathbf{x}^{(k)})}{\partial x_j} \approx \frac{f_i(\mathbf{x}^{(k)} + f_j(\mathbf{x}^{(k)})\mathbf{e}_j) - f_i(\mathbf{x}^{(k)})}{f_j(\mathbf{x}^{(k)})}$$

De esta forma no es necesario conocer en forma explícita las derivadas parciales y se logra una velocidad supralineal.

### 3.3.7. Método de Broyden\*

El método de Broyden mejora las desventajas computacionales de Newton-Raphson, que vimos en la Observación 3.3.1, pero converge más lento. La idea general del método es usar el paso de Newton-Raphson pero no usar  $J_F(\mathbf{x}^{(k)})$  sino una aproximación.

$$\begin{cases} \mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - B_k^{-1}F(\mathbf{x}^{(k)}) \\ \mathbf{x}^{(0)} \in \mathbb{R}^n \end{cases}$$

donde  $B_k$  es la aproximación de Broyden de  $J_F(\mathbf{x}^{(k)})$ .

Desarrollamos  $F$  por Taylor en  $\mathbf{x}^{(k)}$  y evaluando en  $\mathbf{x}^{(k-1)}$  obtenemos la siguiente aproximación:

$$F(\mathbf{x}^{(k-1)}) - F(\mathbf{x}^{(k)}) \approx J_F(\mathbf{x}^{(k)})(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)})$$

Buscamos entonces  $B_k$  tal que  $\mathbf{y}^{(k-1)} = B_k \mathbf{s}_{k-1}$ , donde  $\mathbf{y}^{(k-1)} = F(\mathbf{x}^{(k)}) - F(\mathbf{x}^{(k-1)})$  y  $\mathbf{s}^{(k-1)} = \mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}$ , esta condición para  $B_k$  es llamada *condición Cuasi-Newton*. En el paso  $k$  conocemos  $\mathbf{y}^{(k-1)}$  y  $\mathbf{s}^{(k-1)}$ .  $B_k$  tiene  $n^2$  incógnitas y  $n$  ecuaciones, por lo que es un sistema indeterminado.

Broyden propone que  $B_k$  y  $B_{k-1}$  no difieran en direcciones ortogonales al paso  $k-1$ , o sea

$$(\mathbf{s}^{(k-1)})^T \cdot p = 0 \text{ implica } B_k \cdot p = B_{k-1} \cdot p \quad (3.1)$$

que es llamada la *condición de Broyden*.

La condición (3.1) es equivalente a que  $B_k$  minimice  $\|\tilde{B} - B_{k-1}\|_F$  para  $\tilde{B}$  solución del sistema  $\tilde{B}\mathbf{s}^{(k-1)} = \mathbf{y}^{(k-1)}$ . Donde  $\|\cdot\|_F$  es la norma de Frobenius y esta definida por

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}, \quad A = \{a_{ij}\}_{i=1,j=1}^{n,n}$$

**Proposición 3.3.1.** *La matriz  $B_k$  dada por*

$$B_k = B_{k-1} + \frac{(\mathbf{y}^{(k-1)} - B_{k-1}\mathbf{s}^{(k-1)})(\mathbf{s}^{(k-1)})^T}{(\mathbf{s}^{(k-1)})^T \mathbf{s}^{(k-1)}}$$

*cumple con la condición Cuasi-Newton y la condición de Broyden en las hipótesis hasta ahora mencionadas.*

*Demostración.* Veamos la condición Cuasi-Newton,

$$\begin{aligned} B_k \mathbf{s}^{(k-1)} &= B_{k-1} \mathbf{s}^{(k-1)} + \left( \frac{(\mathbf{y}^{(k-1)} - B_{k-1} \mathbf{s}^{(k-1)})(\mathbf{s}^{(k-1)})^T}{(\mathbf{s}^{(k-1)})^T \mathbf{s}^{(k-1)}} \right) \mathbf{s}^{(k-1)} \\ &= B_{k-1} \mathbf{s}^{(k-1)} + \mathbf{y}^{(k-1)} - B_{k-1} \mathbf{s}^{(k-1)} \\ &= \mathbf{y}^{(k-1)} \end{aligned}$$

La condición de Broyden,

$$\begin{aligned} (\mathbf{s}^{(k-1)})^T \cdot p = 0 \implies B_k \cdot p &= B_{k-1} \cdot p + \frac{(\mathbf{y}^{(k-1)} - B_{k-1} \mathbf{s}^{(k-1)})(\mathbf{s}^{(k-1)})^T \cdot p}{(\mathbf{s}^{(k-1)})^T \mathbf{s}^{(k-1)}} \\ &= B_{k-1} \cdot p \end{aligned}$$

□

---

**Algoritmo 5** Método de Broyden

---

Dados  $x_0 \in \mathbb{R}^n$  y  $B_0$  matriz no singular  $n \times n$ .

```

for  $k \geq 0$  do
   $\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} - B_k^{-1} \cdot F(\mathbf{x}^{(k)})$ 
   $\mathbf{y}^{(k)} \leftarrow F(\mathbf{x}^{(k+1)}) - F(\mathbf{x}^{(k)})$ 
   $\mathbf{s}^{(k)} \leftarrow \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$ 
   $B_{k+1} \leftarrow B_k + \frac{(\mathbf{y}^{(k)} - B_k \mathbf{s}^{(k)})(\mathbf{s}^{(k)})^T}{(\mathbf{s}^{(k)})^T \mathbf{s}^{(k)}}$ 
end for

```

---

*Observación 3.3.3.* En el Algoritmo 5:

1.  $B_0$  se elije de forma que sea fácil de invertir, por ejemplo una matriz diagonal.
2. Sigue haciendo  $O(n^3)$  flops por paso como Newton-Raphson, al invertir la matriz  $B_k$ .
3. Solo hace  $n$  llamadas a funciones contra las  $n^2$  de Newton-Raphson.

Veamos ahora una manera de solucionar el problema de computar  $B_k^{-1}$  usando la formula de actualización de la inversa de Sherman Morrison,

**Proposición 3.3.2.** Sea  $A \in \mathcal{M}_{n \times n}(\mathbb{R})$ ,  $u, v \in \mathbb{R}^n$ , definimos  $M = A + uv^T$ , tal que  $\det(M) \neq 0$ . Luego

$$M^{-1} = (A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{\alpha}$$

donde  $\alpha = 1 + v^T A^{-1}u \neq 0$ .

*Demostración.*

$$\begin{aligned}
 (A + uv^T) \left( A^{-1} - \frac{A^{-1}uv^T A^{-1}}{\alpha} \right) &= \\
 I - \frac{uv^T A^{-1}}{\alpha} + uv^T A^{-1} - \frac{\overbrace{u(v^T A^{-1}u)}^{\in \mathbb{R}} v^T A^{-1}}{\alpha} &= \\
 I + uv^T A^{-1} \left( -\frac{1}{\alpha} + 1 - \frac{v^T A^{-1}u}{\alpha} \right) &= I
 \end{aligned}$$

□

Aplicando la proposición 3.3.2 a  $B_{k+1} = B_k + uv^T$ , con  $u = \frac{(\mathbf{y}^{(k)} - B_k \mathbf{s}^{(k)})}{(\mathbf{s}^{(k)})^T \mathbf{s}^{(k)}}$  y  $v = \mathbf{s}^{(k)}$ , podemos definir el Algoritmo 6.

---

**Algoritmo 6** Algoritmo de Broyden usando Sherman Morrison

---

Dados  $\mathbf{x}^{(0)} \in \mathbb{R}^n$  y  $B_0 \in M^{n \times n}(\mathbb{R})$ .

**for**  $k \geq 0$  **do**

$$\mathbf{x}^{(k+1)} \leftarrow \mathbf{x}^{(k)} - B_k^{-1} F(\mathbf{x}^{(k)})$$

$$\mathbf{y}^{(k)} \leftarrow F(\mathbf{x}^{(k+1)}) - F(\mathbf{x}^{(k)})$$

$$\mathbf{s}^{(k)} = \mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}$$

$$B_{k+1}^{-1} = B_k^{-1} + \frac{(\mathbf{s}^{(k)} - B_k^{-1} \mathbf{y}^{(k)})(\mathbf{s}^{(k)})^T B_k^{-1}}{(\mathbf{s}^{(k)})^T B_k^{-1} \mathbf{y}^{(k)}}$$

**end for**

---

*Observación 3.3.4.* En el Algoritmo 6:

1. Hacemos  $n$  evaluaciones de funciones por paso como en Newton-Raphson y Broyden.
2. Tenemos  $O(n^2)$  de flops por paso al hallar  $B_k^{-1}$ , que es mejor que Newton-Raphson y Broyden.
3. El orden de convergencia es supralineal.