Tema 1: Resolución numérica de ecuaciones y sistemas no lineales

Métodos Numéricos II

Departamento de Matemática Aplicada
Universidad de Granada

Curso 2024/25

Métodos de iteración funcional

Muchos métodos para resolver

$$f(x) = 0$$

consisten en transformar el problema en uno equivalente de la forma

$$x = g(x)$$

y de ahí construir el método

$$x_{n+1} = g(x_n)$$

con una semilla x_0 adecuada.

Un valor s verificando $s = g(s)$ se denomina punto fijo de g.

Métodos de iteración funcional

Por ejemplo, la ecuación $x^3 + 9x + 9 = 0$ que tiene una única raíz real en el intervalo $[-1, 0]$ se puede reescribir de muy diversas formas:

$$x = g_1(x) = -\frac{x^3}{9} - 1, \qquad x = g_2(x) = -\frac{9}{x^2+9},$$

$$x = g_3(x) = -\sqrt{-\frac{9}{x} - 9}, \quad x = g_4(x) = \frac{2x^3 - 9}{3(x^2+3)}.$$

n	$x = g_1(x)$	$x = g_2(x)$	$x = g_3(x)$	$x = g_4(x)$
0	-1	-1	-1	-1
1	-0.888889	-0.9	0	-0.916667
2	-0.921963	-0.917431	$\nexists$	-0.914909
3	-0.912924	-0.914478		-0.914908
4	-0.915460	-0.914981		-0.914908
5	-0.914754	-0.914895		-0.914908

Métodos de iteración funcional

Función contráctil

Una función $g : [a, b] \to \mathbb{R}$ cumpliendo

$$|g(x) - g(y)| \leq L|x - y| \qquad \forall x, y \in [a, b]$$

se dirá lipschitziana.
Esta condición se denomina condición de Lipschitz y L es la constante de Lipschitz.

En el caso $L < 1$ la función g se dirá contráctil, contractiva o contracción y L la constante de contractibilidad.

Métodos de iteración funcional

Teorema del Punto Fijo

Sea $g : [a, b] \to [a, b]$ contráctil con constante de contractibilidad L.
Entonces:

1. Existe un único punto fijo s de g en $[a, b]$.
2. El método $x_{n+1} = g(x_n)$ genera una sucesión convergente a s $\forall x_0 \in [a, b]$
3. Los errores $e_n = x_n - s$ cumplen $|e_n| \leq \dfrac{L^n}{1-L}|x_1 - x_0|$, $n = 1, 2, \ldots$

Demostración:

1. La existencia de s se prueba aplicando Bolzano a $f(x) = g(x) - x$. La unicidad se deduce de la condición de contractibilidad pues si hubiera dos s_1 y s_2, entonces

$$|s_1 - s_2| = |g(s_1) - g(s_2)| \leq L|s_1 - s_2| < |s_1 - s_2|$$

Métodos de iteración funcional

Demostración (continuación):

2. Vemos que $\forall x_0 \in [a,b]$ la sucesión $\{x_n\}_{n\geq 0} \subset [a,b]$ converge a s:

$$|x_n - s| = |g(x_{n-1}) - g(s)| \leq L|x_{n-1} - s| \leq \cdots \leq L^n|x_0 - s|.$$

3. Como $|x_0 - s| \leq |x_1 - x_0| + |x_1 - s| \leq |x_1 - x_0| + L|x_0 - s|$, entonces

$$|x_0 - s| \leq \frac{1}{1-L}|x_1 - x_0| \;\text{ y por tanto }\; |x_n - s| \leq \frac{L^n}{1-L}|x_1 - x_0|.$$

Observaciones:

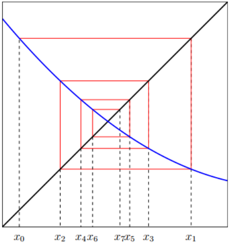
- Para funciones derivables, el teorema sigue siendo válido si se sustituye la contractibilidad por $|g'(x)| \leq L < 1 \; \forall x \in [a,b]$.
- Si $g : \mathbb{R} \to \mathbb{R}$ cumple la condición anterior, el teorema se sigue cumpliendo.

Métodos de iteración funcional

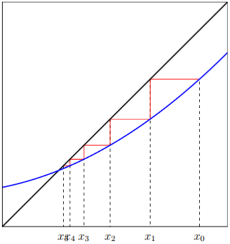
Teorema de convergencia local

Sea $g \in \mathcal{C}^1(I)$ siendo I un entorno abierto de un punto fijo s de g tal que $|g'(s)| < 1$. Entonces existe un subintervalo $I_\varepsilon \subseteq I$ en el que el método $x_{n+1} = g(x_n)$ converge a s $\forall x_0 \in I_\varepsilon$.

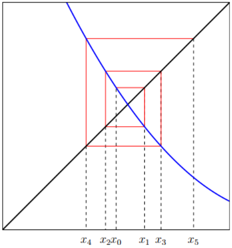
Interpretación geométrica:



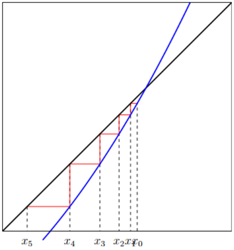
Convergencia en espiral



Convergencia monótona o en escalera



Divergencia en espiral



Divergencia monótona o en escalera

Ejemplo

Métodos de iteración funcional aplicados a la ecuación $e^x + x - 3 = 0$ con $x_0 = 1$,

$$g_1(x) = 3 - e^x, \qquad g_2(x) = \ln(3 - x), \qquad g_3(x) = \frac{xe^x - e^x + 3}{e^x + 1}.$$

n	g_1	g_2	g_3
0	1.000000	1.000000	1.000000
1	0.281718	0.693147	0.806824
2	1.674595	0.835884	0.792135
3	-2.336633	0.772012	0.792060
4	2.903347	0.801099	0.792060
5	-15.235084	0.787958	0.792060
6	3.000000	0.793916	0.792060
7	-17.085532	0.791219	0.792060
8	3.000000	0.792441	0.792060
9	-17.085536	0.791887	0.792060

Iteración funcional: error y orden de convergencia

En la práctica no es fácil obtener la constante de contractibilidad L.
Si g es suficientemente derivable podríamos escribir

$$x_{n+1} - s = g(x_n) - g(s) = g'(\xi_n)(x_n - s)$$

con ξ_n comprendido entre x_n y s. De esta forma,

$$e_{n+1} = g'(\xi_n)e_n,$$

y por tanto

$$\lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|} = |g'(s)|.$$

Si $g'(s) \neq 0$ se tendrá convergencia local lineal con constante $C = |g'(s)|$.
De lo contrario, si $g \in \mathcal{C}^2(I)$ y $g''(s) \neq 0$, por un razonamiento similar

$$e_{n+1} = \frac{1}{2}g''(\xi_n)e_n^2 \qquad \Rightarrow \qquad \lim_{n\to\infty} \frac{|e_{n+1}|}{|e_n|^2} = \frac{1}{2}|g''(s)|$$

con lo que la convergencia sería localmente cuadrática con constante asintótica del error $C = \frac{1}{2}|g''(s)|$.

Iteración funcional: error y orden de convergencia

De forma más general tenemos el siguiente

Teorema: Orden de convergencia

Sea $s = g(s)$ con $g \in \mathcal{C}^p(I)$ verificando

$$g'(s) = \cdots = g^{(p-1)}(s) = 0 \text{ y } g^{(p)}(s) \neq 0.$$

Entonces para x_0 suficientemente próximo a s el método iterativo $x_{n+1} = g(x_n)$ converge a s con orden de convergencia p.

Además, la constante asintótica del error es

$$C = \frac{1}{p!}|g^{(p)}(s)|.$$

Ejemplo

Volvemos al ejemplo anterior en el que veíamos tres métodos de iteración funcional aplicados a la ecuación $e^x + x - 3 = 0$ con $x_0 = 1$. Sabemos que tiene una raíz en el intervalo $[0, 1]$ y que:

$$g_1(x) = 3 - e^x, \qquad g_2(x) = \ln(3 - x), \qquad g_3(x) = \frac{xe^x - e^x + 3}{e^x + 1}.$$

Observamos que

$$g_1'(x) = -e^x, \qquad g_2'(x) = \frac{-1}{3 - x}, \qquad g_3'(x) = \frac{(x + e^x - 3)e^x}{(e^x + 1)^2}$$

Se cumple que:

$$|g_1'(s)| > 1, \qquad |g_2'(s)| < 1, \qquad g_3'(s) = 0$$

Aceleración Δ^2 de Aitken

Se trata de un método que permite acelerar la convergencia de una sucesión que converge linealmente.

A partir de la sucesión $\{x_n\}_{n\geq 0}$ se construye la sucesión acelerada $\{\hat{x}_n\}_{n\geq 0}$ como

$$\hat{x}_n = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} = x_n - \frac{(x_{n+1}-x_n)^2}{x_{n+2}-2x_{n+1}+x_n} = \frac{x_{n+2}x_n - x_{n+1}^2}{x_{n+2}+x_n-2x_{n+1}}$$

siendo Δ el operador diferencia progresiva

$$\Delta x_n = x_{n+1} - x_n \qquad \Delta^2 x_n = \Delta(\Delta x_n) = x_{n+2} - 2x_{n+1} + x_n.$$

En algunos textos (Atkinson, Ralston, Wikipedia,...) se presenta la formulación del método Δ^2 de Aitken como

$$\hat{x}_{n+2} = x_{n+2} - \frac{(\Delta x_{n+1})^2}{\Delta^2 x_n}$$

que es equivalente, salvo por la numeración de subíndices de la sucesión acelerada. Aquí usaremos la primera (Gasca, Burden).

Aceleración Δ^2 de Aitken

Teorema: aceleración de Aitken

Sea la sucesión $\{x_n\}_{n\geq 0}$ convergente a s al menos linealmente. Entonces la sucesión acelerada $\hat{x}_n = x_n - \dfrac{(\Delta x_n)^2}{\Delta^2 x_n}$ converge a s más rápidamente en el sentido $\lim\limits_{n\to\infty} \dfrac{\hat{e}_n}{e_n} = 0$ siendo $\hat{e}_n = \hat{x}_n - s$ y $e_n = x_n - s$.

Demostración:
Supongamos $\dfrac{e_{n+1}}{e_n} = \lambda_n$ con $\lim\limits_{n\to\infty} |\lambda_n| = C < 1$. Como $\Delta e_n = \Delta x_n$,

$$\begin{aligned} \hat{e}_n &= e_n - \frac{(\Delta e_n)^2}{\Delta^2 e_n} = e_n - \frac{e_n^2(\lambda_n - 1)^2}{e_n(\lambda_{n+1}\lambda_n - 2\lambda_n + 1)} \\ &= e_n \frac{\lambda_{n+1}\lambda_n - 2\lambda_n + 1 - \lambda_n^2 + 2\lambda_n - 1}{\lambda_{n+1}\lambda_n - 2\lambda_n + 1} = e_n \frac{\lambda_n \Delta\lambda_n}{\lambda_{n+1}\lambda_n - 2\lambda_n + 1}, \end{aligned}$$

luego $\lim_{n\to\infty} \frac{\hat{e}_n}{e_n} = 0$.

Aceleración de Steffensen

Cuando la sucesión $\{x_n\}_{n\geq 0}$ convergente a s proviene de un método de iteración funcional $x_{n+1} = g(x_n)$, se puede combinar éste con el procedimiento de aceleración de Aitken para aumentar la velocidad de convergencia.

El procedimiento de aceleración de Steffensen que parte de x_0 consiste en usar g hasta tener tres términos, acelerar mediante Aitken y volver a generar tres términos mediante g para volver a acelerar.

Más concretamente:

$$\left.\begin{array}{l} x_0 \\ x_1 = g(x_0) \\ x_2 = g(x_1) \end{array}\right\} \longrightarrow x_0' = x_0 - \frac{(\Delta x_0)^2}{\Delta^2 x_0}$$

$$\left.\begin{array}{l} x_0' \\ x_1' = g(x_0') \\ x_2' = g(x_1') \end{array}\right\} \longrightarrow x_0'' = x_0' - \frac{(\Delta x_0')^2}{\Delta^2 x_0'}$$

y así sucesivamente.

La sucesión $\{x_0^{(n)}\}$ convergerá a s aún más rápidamente que la de Aitken.

Aceleración de Steffensen

Aceleraciones de Aitken y Steffensen aplicadas a la función $g(x) = e^{-x}$ con $x_0 = 0.5$.

ACELERACIONES DE AITKEN Y STEFFENSEN		
g	Aitken	Steffensen
0.5000000000		
0.6065306597		
0.5452392119	0.5676238764	0.5676238764
0.5797030949	0.5672989893	
0.5600646279	0.5671931424	
0.5711721490	0.5671593645	0.5671433141
0.5648629470	0.5671484533	
0.5684380476	0.5671449524	
0.5664094527	0.5671438247	0.5671432904
0.5675596343	0.5671434623	
0.5669072129	0.5671433457	
0.5672771960	0.5671433082	0.5671432904

Ecuaciones polinómicas

Se pretende resolver $P(x) = 0$ siendo $P(x)$ un polinomio en x.
Evidentemente son válidos todos los métodos de aplicación general para ecuaciones (Bisección, Newton-Raphson, Secante, etc.).
El tratamiento específico de una ecuación polinómica pasa por tres etapas:

1. Determinación del número de raíces reales.
2. Localización y separación en intervalos disjuntos de las raíces reales.
3. Estimación precisa de cada raíz mediante un método numérico adecuado.

Las dos primeras etapas se abordan con la teoría de Sturm (véase Gasca, Cálculo Numérico, UNED, 1996).
Para la tercera es muy adecuado adaptar el método de Newton-Raphson.

Ecuaciones polinómicas: localización de raíces

Teorema (acotación de raíces)

Sea $P(x) = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0$. Toda raíz s de P verifica:

$$|s| \leq 1 + \alpha \quad \text{donde} \quad \alpha = \max_{0 \leq i < k} \left| \frac{a_i}{a_k} \right|.$$

Demostración:

$$P(s) = 0 \Rightarrow s^k = -\sum_{i=0}^{k-1} \frac{a_i}{a_k} s^i \Rightarrow |s|^k \leq \alpha \sum_{i=0}^{k-1} |s|^i = \alpha \frac{|s|^k - 1}{|s| - 1}.$$

Si fuese $|s| > \alpha + 1$ sería $\alpha < |s| - 1$ y entonces $|s|^k < |s|^k - 1$, absurdo.

Ecuaciones polinómicas: sucesión de Sturm

Definición

$\{f_0, f_1, \ldots, f_m\} \subset \mathcal{C}[a, b]$ es una sucesión de Sturm en $[a, b]$ si $\forall x \in [a, b]$:

1. $f_0 \in \mathcal{C}^1[a, b]$,
2. $f_0(x) = 0 \Rightarrow f_0'(x) f_1(x) > 0$,
3. $f_j(x) = 0 \Rightarrow f_{j-1}(x) f_{j+1}(x) < 0 \quad j = 1, \ldots, m - 1.$
4. $f_m(x) \neq 0.$

Ecuaciones polinómicas: sucesión de Sturm

Teorema de Sturm

Si $\{f_0, f_1, \ldots, f_m\}$ es una sucesión de Sturm en $[a, b]$, entonces la cantidad de raíces reales que f_0 tiene en $[a, b]$ es la diferencia entre el número de cambios de signo que hay en la secuencia $\{f_0(a), f_1(a), \ldots, f_m(a)\}$ respecto del que hay en la secuencia $\{f_0(b), f_1(b), \ldots, f_m(b)\}$

Demostración:

Supongamos que x recorre el intervalo desde a hasta b.

Al atravesar x un cero de f_j, $j = 1, \ldots m-1$ no se altera el número de cambios de signo en la secuencia $\{f_0(x), f_1(x), \ldots, f_m(x)\}$.

Por otro lado, al atravesar x un cero de f_0 hay dos casos posibles:

1. pasar de $-$ a $+$ con lo que $f_1 > 0$, o
2. pasar de $+$ a $-$ con lo que $f_1 < 0$.

En cualquier caso se pierde un cambio de signo en la secuencia.

Ecuaciones polinómicas: sucesión de Sturm

Procedimiento de obtención de una sucesión de Sturm para un polinomio

Sea $P(x)$ un polinomio.

1. $f_0 = P$, $f_1 = P'$.
2. $f_{j-2} = q_{j-1}f_{j-1} - f_j$, $j = 2, \dots, m$
   (algoritmo de Euclides de división, con el resto cambiado de signo, es decir, f_j es el resto cambiado de signo de la división de f_{j-2} entre f_{j-1}).
3. $f_m =$ último resto no nulo cambiado de signo.
4. Si f_m es constante, entonces $\{f_0, \dots, f_m\}$ es una sucesión de Sturm $\forall [a, b]$; de lo contrario lo es $\left\{\frac{f_0}{f_m}, \frac{f_1}{f_m}, \dots, \frac{f_m}{f_m}\right\}$.

Ejemplo: $P(x) = 2x^5 - x^4 - 4x^3 + 2x^2 - 6x + 3$

Partimos de

$$\begin{aligned} f_0 &= P = 2x^5 - x^4 - 4x^3 + 2x^2 - 6x + 3 \\ f_1 &= P' = 10x^4 - 4x^3 - 12x^2 + 4x - 6. \end{aligned}$$

Para calcular f_2, dividimos f_0 entre f_1 y nos queda cociente $\frac{1}{5}x - \frac{1}{50}$ y de resto $-\frac{42}{25}x^3 + \frac{24}{25}x^2 - \frac{118}{25}x + \frac{72}{25}$. Cambiamos el signo al resto y esa podría ser f_2. Multiplicamos por $25/2$ para simplificar. Entonces:

$$f_2 = 21x^3 - 12x^2 + 59x - 36$$

Para calcular f_3, dividimos f_1 entre f_2 y nos queda cociente $\frac{10}{21}x + \frac{12}{147}$ y de resto $-\frac{5750}{147}x^2 + \frac{800}{49}x - \frac{150}{49}$. Cambiamos el signo al resto y esa podría ser f_3. Multiplicamos por $147/50$ para simplificar. Entonces:

$$f_3 = 115x^2 - 48x + 9$$

Ejemplo (continuación)

Para calcular f_4, dividimos f_2 entre f_3 y nos queda cociente $\frac{21}{115}x - \frac{372}{13225}$ y de resto $-\frac{740684}{13225}x - \frac{472752}{13225}$. Cambiamos el signo al resto y esa podría ser f_4. Multiplicamos por $13225/196$ para simplificar. Entonces:

$$f_4 = 3779x - 2412$$

Por último, dividimos f_3 entre f_4 y nos queda cociente $\frac{115}{3779}x - \frac{95988}{14280841}$ y de resto $\frac{360050625}{14280841}$. Podemos entonces tomar $f_5 = -1$

La sucesión de Sturm entonces sería:

$$\begin{array}{rcl} f_0 & = & 2x^5 - x^4 - 4x^3 + 2x^2 - 6x + 3 \\ f_1 & = & 10x^4 - 4x^3 - 12x^2 + 4x - 6 \\ f_2 & = & 21x^3 - 12x^2 + 59x - 36 \\ f_3 & = & 115x^2 - 48x + 9 \\ f_4 & = & -185171x + 118188 \\ f_5 & = & -1 \end{array}$$

Ejemplo (continuación)

Para conocer el número de raíces reales de $P = 2x^5 - x^4 - 4x^3 + 2x^2 - 6x + 3$ utilizando la sucesión de Sturm, primero utilizamos que todas las raíces cumplen:

$$|s| \leq 1 + \alpha \quad \text{donde} \quad \alpha = \max_{0 \leq i < k} \left| \frac{a_i}{a_k} \right|.$$

En este caso $\alpha = \max\{1/2, 4/2, 2/2, 6/2, 3/2\} = 3$ por lo que las raíces están en el intervalo $[-4, 4]$.

Utilizando la sucesión de Sturm y evaluando en -4 y 4 tenemos:

$$\begin{aligned}
\{f_0(-4), &f_1(-4), f_2(-4), f_3(-4), f_4(-4), f_5(-4)\} \\
&= \{-1989, 2602, -1808, 2041, 858872, -1\} \\
\{f_0(4), &f_1(4), f_2(4), f_3(4), f_4(4), f_5(4)\} \\
&= \{1547, 2122, 1352, 1657, -622496, -1\}
\end{aligned}$$

Con lo que hay 4 cambios de signo en la primera lista y solo 1 cambio en la segunda.

El polinomio tiene entonces 3 raíces reales y están en $[-4, 4]$.

Ejemplo (continuación)

Para localizar las raíces en intervalos de amplitud 1 vamos evaluando en diferentes puntos del intervalo:

$$\{f_0(0), f_1(0), f_2(0), f_3(0), f_4(0), f_5(0)\} = \{3, -6, -36, 9, 118188, -1\}$$

Hay 3 cambios de signo por lo que hay una raíz en $[-4, 0]$ y dos en $[0, 4]$

$$\{f_0(1), f_1(1), f_2(1), f_3(1), f_4(1), f_5(1)\} = \{-4, -8, 32, 76, -66983, -1\}$$

Hay dos cambios de signo por lo que hay una raíz en $[0, 1]$ y otra en $[1, 4]$. Se suele continuar hasta localizar las raíces en intervalos de longitud 1.

Ejercicio (raíces dobles)

Localizar las raíces de

$$P(x) = 9x^4 - 12x^3 + 13x^2 - 12x + 4$$

Método de Newton-Raphson para ecuaciones polinómicas

Se puede mejorar la eficiencia del método mediante algoritmos de evaluación para P y P'.
Para evaluar $P(x) = a_k x^k + a_{k-1} x^{k-1} + \cdots + a_1 x + a_0$ en el punto $x = t$ se puede escribir

$$P(x) = (x - t)Q(x) + b_0$$

donde $Q(x) = b_k x^{k-1} + \cdots + b_2 x + b_1$ es el cociente de dividir $P(x)$ entre $(x - t)$ y b_0 es el valor de $P(t)$.
El método para obtener estos coeficientes, que no es más que la regla de Ruffini, se conoce como algoritmo de Horner:

$$\begin{aligned} b_k &= a_k \\ b_j &= a_j + t \cdot b_{j+1}, \quad j = k-1, \ldots, 0 \end{aligned}$$

Método de Newton-Raphson para ecuaciones polinómicas

La misma idea se puede aprovechar para el cálculo eficiente de $P'(t)$. Podemos escribir

$$P'(x) = Q(x) + (x - t)Q'(x).$$

y por tanto $P'(t) = Q(t)$ con lo que para obtener $P'(t)$ basta con aplicar el algoritmo de Horner a $Q(x)$

$$\begin{aligned} c_k &= b_k \\ c_j &= b_j + t \cdot c_{j+1}, \quad j = k-1, \ldots, 1. \end{aligned}$$

Métodos de resolución para sistemas de ecuaciones no lineales

Consideraremos aquí el problema de resolver el sistema de ecuaciones $F(X) = 0$ donde $F : D \subseteq \mathbb{R}^k \to \mathbb{R}^k$ es una función vectorial de las variables reales $X = (x_1, \ldots, x_k)$ y componentes $f_1, \ldots, f_k : D \subseteq \mathbb{R}^k \to \mathbb{R}$.

Para mayor facilidad supondremos $D = \displaystyle\prod_{i=1}^{k} [a_i, b_i]$, que es un dominio compacto y convexo de $\mathbb{R}^k$.

De forma desarrollada el sistema se escribiría

$$\left.\begin{array}{c} f_1(x_1, \ldots, x_k) = 0 \\ \vdots \\ f_k(x_1, \ldots, x_k) = 0 \end{array}\right\}.$$

Métodos de iteración funcional para sistemas

Siguiendo la idea de la iteración funcional de punto fijo, transformaremos el sistema original en otro equivalente en la forma

$$X = G(X),$$

a partir del cual se construirá el método iterativo

$$X_{n+1} = G(X_n)$$

que parte de la semilla X_0.

De forma desarrollada, el sistema transformado sería

$$\left.\begin{array}{c} x_1 = g_1(x_1, \ldots, x_k) \\ \vdots \\ x_k = g_k(x_1, \ldots, x_k) \end{array}\right\}.$$

Ejemplo

Consideremos el sistema

$$\left.\begin{array}{r} x^2 + xy - 10 = 0 \\ 3xy^2 + y - 57 = 0 \end{array}\right\} \text{ en } D = [1,3] \times [2,4],$$

que tiene una solución en $s = (2,3) \in D$. Una posible transformación podría ser

$$\left.\begin{array}{l} x = g_1(x,y) = \frac{10}{x+y} \\ y = g_2(x,y) = 57 - 3xy^2 \end{array}\right\}.$$

Partiendo de $(x_0, y_0) = (2.1, 2.9) \approx s$ se obtienen los resultados de la tabla

n	x_n	y_n
0	2.1	2.9
1	2.00	4.02
2	1.66	-39.82
3	-0.26	-7847.86

por lo que está claro que el método no es adecuado.

Métodos de iteración funcional para sistemas

Teorema (Convergencia global)

Sea $G : D \to D$ verificando para alguna norma $||\cdot||$ en $\mathbb{R}^k$

$$||G(X) - G(Y)|| \leq L||X - Y|| \quad \forall X, Y \in D \quad \text{con } 0 \leq L < 1.$$

Entonces,

1. Existe un único punto fijo $S = (s_1, \ldots, s_k) \in D$ de G.
2. El método $X_{n+1} = G(X_n)$ converge a S para todo $X_0 \in D$, es decir, $\lim\limits_{n\to\infty} ||X_n - S|| = 0$.
3. Los errores $E_n = X_n - S$ cumplen $||E_n|| \leq \dfrac{L^n}{1-L}||X_1 - X_0|| \; \forall n \geq 1$.

Métodos de iteración funcional para sistemas

Observación: Si existen todas las derivadas parciales y son continuas en D, la condición de contractibilidad puede sustituirse por

$$||G'(X)|| \leq L < 1$$

o también por

$$\left|\frac{\partial g_i(X)}{\partial x_j}\right| \leq \frac{L}{k}, \qquad \forall i, j \quad \forall X \in D$$

donde $G'(X)$ indica la matriz jacobiana:

$$G' = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_k}{\partial x_1} & \cdots & \frac{\partial g_k}{\partial x_k} \end{pmatrix}$$

y k indica el número de variables.

Métodos de iteración funcional para sistemas

Teorema (Convergencia local)

Sea I un entorno de $S = G(S)$ y $G \in \mathcal{C}^1(I)$ tal que $||G'(S)|| < 1$ para alguna norma matricial $||\cdot||$. Entonces existe un subentorno $I_\varepsilon \subseteq I$ en el que el método $X_{n+1} = G(X_n)$ converge a S para todo $X_0 \in I_\varepsilon$.

Observaciones:

- Se puede sustituir la condición $||G'(S)|| < 1$ por $\rho(G'(S)) < 1$ (radio espectral).
- Se puede sustituir la condición $||G'(S)|| < 1$ por $\left|\dfrac{\partial g_i(S)}{\partial x_j}\right| \leq \dfrac{1}{k}\ \forall i, j$.

El método de Newton-Raphson para sistemas

Para resolver el sistema $F(X) = 0$ el método de Newton-Rahpson consiste en

$$X_{n+1} = X_n - J_n^{-1} F(X_n), \qquad n = 0, 1, \ldots,$$

donde J_n es la matriz jacobiana

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots & \frac{\partial f_k}{\partial x_k} \end{pmatrix}$$

evaluada en la aproximación X_n.
El método estará bien definido si J_n admite inversa, es decir, si $\det(J_n) \neq 0 \ \forall n$.

El método de Newton-Raphson para sistemas

En la práctica no hay que invertir el jacobiano en cada paso, sino resolver un sistema de ecuaciones lineales. La deducción es la siguiente:

$$\begin{aligned} X_{n+1} &= X_n - J_n^{-1} F(X_n); \\ X_{n+1} - X_n &= -J_n^{-1} F(X_n); \\ \Delta X_n &= -J_n^{-1} F(X_n) \quad \text{siendo} \quad \Delta X_n = X_{n+1} - X_n; \\ J_n \Delta X_n &= -F(X_n) \end{aligned}$$

luego ΔX_n es la solución del sistema de ecuaciones lineales cuya matriz de coeficientes es J_n y cuyo vector de términos independientes es $-F(X_n)$. Una vez resuelto, se tomará $X_{n+1} = X_n + \Delta X_n$.

El método de Newton-Raphson para sistemas

Ejemplo

En el caso del ejemplo anterior:

$$\left.\begin{array}{r} x^2 + xy - 10 = 0 \\ 3xy^2 + y - 57 = 0 \end{array}\right\} \text{ en } D = [1,3] \times [2,4],$$

la matriz jacobiana asociada es

$$J = \begin{pmatrix} 2x + y & x \\ 3y^2 & 6xy + 1 \end{pmatrix}$$

que ha de ser invertible en cada aproximación calculada.
Esto es cierto en la región D considerada, ya que
$\det(J) = 12x^2y + 3xy^2 + 2x + y > 0$ en D.

El método de Newton-Raphson para sistemas

Ejemplo (continuación)

El método de Newton-Raphson quedaría:

$$X_0 = \begin{pmatrix} x_0 \\ y_0 \end{pmatrix} = \begin{pmatrix} 1.5 \\ 3.5 \end{pmatrix} \text{ (semilla no muy próxima a la solución } s\text{)}$$

$$X_{n+1} = \begin{pmatrix} x_{n+1} \\ y_{n+1} \end{pmatrix} = \begin{pmatrix} x_n \\ y_n \end{pmatrix} - \begin{pmatrix} 2x_n + y_n & x_n \\ 3y_n^2 & 6x_n y_n + 1 \end{pmatrix}^{-1} \begin{pmatrix} x_n^2 + x_n y_n - 10 \\ 3x_n y_n^2 + y_n - 57 \end{pmatrix}$$

$$n = 0, 1, 2, \ldots$$

Los resultados numéricos con tolerancia 10^{-6} se muestran en el cuadro

n	x_n	y_n	$\|\|x_{n+1} - x_n\|\|_\infty$
0	1.50000000	3.50000000	
1	2.03602882	2.84387510	6.56e-01
2	1.99870061	3.00228856	1.58e-01
3	1.99999998	2.99999941	2.29e-03
4	2.00000000	3.00000000	5.87e-07

El método de Newton-Raphson para sistemas

Teorema

Sea el sistema de ecuaciones $F(X) = 0$ con $F \in \mathcal{C}^2(I)$ siendo I un entorno de una solución S con $\det(J(F)|_{X=S}) \neq 0$. Entonces existe un entorno $I_\varepsilon \subseteq I$ para el cual la sucesión generada por el método de Newton Rapshon converge a la solución con orden de convergencia, al menos, cuadrático $\forall X_0 \in I_\varepsilon$.