Ecuaciones y Sistemas de Ecuaciones no Lineales

Estudiaremos algunos métodos básicos de resolución de ecuaciones o sistemas de ecuaciones no lineales.

El problema consiste en:

- dada $f : \mathbb{R} \to \mathbb{R}$ (no lineal), encontrar $x \in \mathbb{R}$ tal que $f(x) = 0$, para el caso de una sola ecuación, o bien
- dada $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ (no lineal), encontrar $\boldsymbol{x} = (x_1, \ldots, x_n)^{\mathrm{t}} \in \mathbb{R}^n$ tal que $\boldsymbol{f}(\boldsymbol{x}) = \mathbf{0}$, para el caso de un sistema de ecuaciones.

Para el caso escalar (una sola ecuación), la solución x se denomina **raíz** de la función f.

Métodos de convergencia garantizada

Existencia de raíces

Teorema del Valor Intermedio. *Sea $f : [a, b] \to \mathbb{R}$ una función continua en el intervalo $[a, b]$ y supongamos que $f(a) < f(b)$. Entonces, para cada valor intermedio z tal que $f(a) < z < f(b)$, existe $\alpha \in (a, b)$ tal que $f(\alpha) = z$. La misma conclusión se obtiene para el caso que $f(a) > f(b)$.*

En particular, si $f(a)$ y $f(b)$ tienen signos opuestos, entonces $z = 0$ es precisamente un valor intermedio y, por lo tanto, existe por lo menos una raíz α de f en el intervalo (a, b).

Método de Bisección. Algoritmo.

1. Dados a y b tales que $a < b$ y $f \in \mathcal{C}([a, b])$ tal que $f(a)f(b) < 0$, sean $x_a := a$ y $x_b := b$. Por lo tanto f tiene una raíz en el intervalo (x_a, x_b).
2. Sea $x_r := \dfrac{x_a + x_b}{2}$.
3. Forzosamente debemos caer en uno de los siguientes casos:
   a) $f(x_a)f(x_r) = 0$: en este caso se tiene que $f(x_r) = 0$; por lo tanto ya localizamos una raíz, x_r, y se finaliza el proceso;
   b) $f(x_a)f(x_r) < 0$: por lo tanto f tiene una raíz en el intervalo (x_a, x_r) y redefinimos entonces x_b como x_r;
   c) $f(x_a)f(x_r) > 0$: por lo tanto f tiene una raíz en el intervalo (x_r, x_b) y redefinimos entonces x_a como x_r.
4. En los casos (b) y (c) anteriores f tiene una raíz en el nuevo intervalo (x_a, x_b). Por lo tanto, el proceso se vuelve a repetir desde (2) con el nuevo intervalo (x_a, x_b), hasta que se satisfaga algún criterio de detención.

Es fácil comprobar a partir del algoritmo que, si α es la raíz de la ecuación, entonces los valores $x_k = x_r$ calculados en cada paso (donde k denota el número de paso) satisfacen

$$|\alpha - x_k| \leq \left(\frac{1}{2}\right)^k (b-a) \qquad \text{y, por lo tanto,} \qquad \alpha = \lim_{k\to\infty} x_k.$$

Observaciones.

1. La cota del error $\left(\frac{1}{2}\right)^k (b-a)$ en el método de bisección se reduce a la mitad en cada paso.
2. El método puede ser demasiado lento, pero al menos es un método en el que la convergencia está garantizada.
3. El método es sólo aplicable al caso escalar (de una sola ecuación), y no se generaliza al caso de sistemas de ecuaciones.

Métodos de convergencia veloz

Orden de convergencia de un método

Definición. *Una sucesión* $\{x_k\}_{k\in\mathbb{N}}$ *que converge a* α *se dice* **convergente con orden** $p \geq 1$, *si*

$$|\alpha - x_{k+1}| \leq C|\alpha - x_k|^p \qquad \forall k \in \mathbb{N},$$

para alguna constante $C > 0$.

Si $p = 1$ se dice que la sucesión **converge linealmente** a α; si $p = 2$, que **converge cuadráticamente**; etc. Cuanto mayor es p, más velozmente se reduce el error.

1) Iteraciones de punto fijo.

Definición. Dado $x_0 \in \mathbb{R}$ y una función g, diremos que x_0 es un **punto fijo** de g si $g(x_0) = x_0$

Nuestro problema original es:

$$\text{Hallar } x \in \mathbb{R} \text{ tal que } f(x) = 0, \text{ para } f \text{ dada.} \tag{1}$$

Definimos el problema de punto fijo como:

$$\text{Hallar } x \in \mathbb{R} \text{ tal que } g(x) = x, \text{ para } g \text{ dada.} \tag{2}$$

Los problemas de búsqueda de raíces (1) y los de punto fijo (2) son **equivalentes**.

El siguiente teorema entrega las condiciones que nos permitan asegurar la **existencia y unicidad** de un punto fijo

Teorema.

a) Si $g \in \mathbb{C}[a,b]$ y $g(x) \in [a,b]$, para todo $x \in [a,b]$, entonces g **tiene un punto fijo** en $[a,b]$.

b) Si además $g'(x)$ existe en (a,b) y existe una constante positiva $k<1$ con

$$|g'(x)| \leq k, \quad \forall x \in (a,b)$$

entonces el punto fijo en $[a,b]$ es **único**.

Lo que buscamos ahora es generar un **algoritmo** que permita aproximar un punto fijo. Supongamos que tenemos el problema: Hallar $x \in \mathbb{R}$ tal que $g(x)=x$, para $g \in [a,b]$ y, existe $k>0$ tal que $|g'(x)|<k<1$ $\forall x \in (a,b)$.

Para aproximar supondremos que tenemos x_0 dado, con $x_0 \in [a,b]$. Entonces definimos la sucesión $\{p_n\}_{n=0}^{\infty}$ de la siguiente forma:

$$p_{n+1} = g(p_n) \quad n \geq 0\,.$$

Podemos demostrar que la **solución converge** y el **límite** será el punto fijo de la función g.

Teorema de punto fijo. Sea $g \in \mathbb{C}[a,b]$ tal que $g(x) \in [a,b]$, $\forall x \in [a,b]$. Además supongamos que existe $k>0$ tal que $|g'(x)|<k<1$ $\forall x \in (a,b)$. Entonces para cualquier p_0 en $[a,b]$, la sucesión definida por

$$p_{n+1} = g(p_n) \quad n \geq 0$$

converge al único punto fijo p en $[a,b]$.

El siguiente colorario, nos entrega un resultado de **velocidad de convergencia**.

Corolario. Bajo las hipótesis del teorema anterior, se tiene que

i) $|p_n - p| \leq k^n \max\{p_0 - a, b - p_0\}$

ii) $|p_n - p| \leq \dfrac{k^n}{1-k}|p_1 - p_0|$ para todo $n \geq 1$.

Nuestro interés ahora es caracterizar el **orden de convergencia** de los esquemas de punto fijo. Consideremos el problema: Hallar $x \in \mathbb{R}$ tal que $g(x)=x$.

Teorema. Sea $g \in \mathbb{C}[a,b]$ tal que

i) $g(x) \in [a,b]$, $\forall x \in [a,b]$

i) Existe $k>0$ tal que $|g'(x)|<k<1$ $\forall x \in (a,b)$

Si además denotamos por p el **único punto fijo** y se cumple que $g'(p) \neq 0$ entonces el orden de convergencia de la sucesión $p_{n+1} = g(p_n) \quad n \geq 0$ es **lineal**.

A continuación se presenta un resultado que asegura la convergencia en **orden cuadrático**.

Teorema. Sea $g \in \mathbb{C}^2[a,b]$ tal que

i) $g(x) \in [a,b]$, $\forall x \in [a,b]$

i) Existe $k>0$ tal que $|g'(x)|<k<1$ $\forall x \in (a,b)$

Si además se cumple que $g'(p) = 0$ con $p = g(p)$ y $g''(p) \neq 0$ entonces el **orden de convergencia de la sucesión** $p_{n+1} = g(p_n)$ $n \geq 0$ es **cuadrático**.

Observación. Podemos estudiar como caso particular

$$g(x) = x - d(x)f(x)$$

y encontrar el valor de $d(x)$ para asegurar un orden cuadrático. Cuando $d(x) = 1/f'(x)$ esto genera el método de Newton-Raphson.

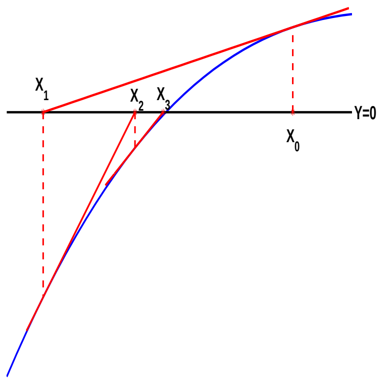
2) El método de Newton–Raphson.

Se basa en usar una **recta tangente** a la gráfica de f para aproximar esta gráfica, cerca del punto donde la función se anula.

Supongamos que tenemos la aproximación x_k a la raíz α de $f(x)$. Trazamos la recta tangente a la curva en el punto $(x_k, f(x_k))$:

$$y = f(x_k) + f'(x_k)(x - x_k).$$

Esta recta cruza al eje de abscisas en un punto x_{k+1} que será nuestra siguiente aproximación a la raíz α.



El punto x_{k+1} donde la recta tangente

$$y = f(x_k) + f'(x_k)(x - x_k)$$

corta al eje de abscisas queda determinado por

$$f(x_k) + f'(x_k)(x_{k+1} - x_k) = 0.$$

El método de Newton–Raphson (o de la tangente) define entonces la sucesión de aproximaciones a α de la manera siguiente:

$$x_{k+1} := x_k - \frac{f(x_k)}{f'(x_k)}, \qquad k = 0, 1, 2, \ldots$$

a partir de una aproximación inicial x_0 dada y siempre que $f'(x_k) \neq 0$.

Teorema 1. *Sea $f \in \mathcal{C}^2([a,b])$ con una raíz $\alpha \in (a,b)$ y sean m_1 y M_2 tales que*

$$m_1 \leq \min_{x\in[a,b]} \left|f'(x)\right| \qquad y \qquad \max_{x\in[a,b]} \left|f''(x)\right| \leq M_2.$$

Supongamos que $m_1 > 0$. Dado $x_0 \in [a,b]$, sea $\{x_k\}_{k\in\mathbb{N}}$ la sucesión obtenida por el método de Newton–Raphson. Supongamos que $x_k \in [a,b] \quad \forall k \in \mathbb{N}$. Entonces

$$|\alpha - x_{k+1}| \leq \frac{M_2}{2m_1} |\alpha - x_k|^2 .$$

Por lo tanto, si x_0 se escoge suficientemente cercano a α, se tiene la convergencia

$$\lim_{k\to\infty} x_k = \alpha,$$

con orden $p = 2$. Además,

$$|\alpha - x_{k+1}| \leq \frac{M_2}{2m_1} |x_{k+1} - x_k|^2 , \qquad k = 0, 1, 2, \ldots$$

Observaciones.

1. De acuerdo al Teorema 1, si el método de Newton–Raphson converge, lo hace **cuadráticamente** (es decir, con **orden** $p = 2$).
2. La convergencia está asegurada por el Teorema 1, bajo la hipótesis de que x_0 esté **suficientemente cerca de la solución** α:
$$|\alpha - x_0| < \frac{2m_1}{M_2}.$$
Sin embargo, no hay una forma práctica de verificar esto.

Criterio de detención.

Si el método de Newton–Raphson converge, cuando

$$|x_{k+1} - x_k| \leq \texttt{tol},$$

con `tol` suficientemente pequeño como para que $\dfrac{M_2\texttt{tol}}{2m_1} \ll 1$, entonces se tiene que

$$\begin{aligned} |\alpha - x_k| &\leq |\alpha - x_{k+1}| + |x_{k+1} - x_k| \\ &\leq \frac{M_2}{2m_1} |x_{k+1} - x_k|^2 + |x_{k+1} - x_k| \approx |x_{k+1} - x_k| . \end{aligned}$$

Como además $|\alpha - x_{k+1}| \ll |\alpha - x_k|$, es razonable estimar

$$|\alpha - x_{k+1}| \leq |x_{k+1} - x_k| \leq \texttt{tol}.$$

En consecuencia, si se desea calcular la raíz α con error menor que `tol` mediante este método, puede usarse como criterio de detención confiable

$$|x_{k+1} - x_k| \leq \texttt{tol}.$$

Ejemplo: Cálculo de $\sqrt{2}$.

Resolución de la ecuación $f(x) = x^2 - 2 = 0$ con error menor que $\texttt{tol} = 10^{-5}$.

Resultados obtenidos por los métodos de *Bisección* y *Newton–Raphson*.

	Bisección	Newton–Raphson
1	1.50000000000000	2.00000000000000
2	1.25000000000000	1.50000000000000
3	1.37500000000000	1.41666666666667
4	1.43750000000000	1.41421568627451
5	1.40625000000000	1.41421356237469
6	1.42187500000000	
7	1.41406250000000	
⋮	⋮	
15	1.41421508789063	
16	1.41419982910156	
17	1.41420745849609	

3) El método de la Secante.

Cuando la derivada de la función f es difícil de evaluar, conviene utilizar el **método de la secante** en lugar del de Newton–Raphson.

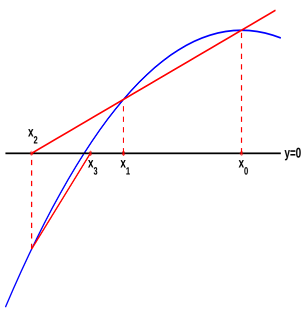
Éste simplemente consiste en reemplazar la derivada $f'(x_k)$ por el cociente incremental

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

Es decir,

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

para $k = 1, 2, \ldots,$ con x_0, x_1 dados.



Teorema. *Sea $f \in \mathcal{C}^2([a,b])$ con una raíz $\alpha \in (a,b)$ y tal que $f'(x) \neq 0$ $\forall x \in [a,b]$. Dado $x_0, x_1 \in [a,b]$, sea $\{x_k\}_{k\in\mathbb{N}}$ la sucesión obtenida por el método de la secante. Supongamos que $x_k \in [a,b]$ $\forall k \in \mathbb{N}$. Entonces*

$$|\alpha - x_{k+1}| \leq C\,|\alpha - x_k|^p,$$

con $p = \dfrac{1+\sqrt{5}}{2}$ *y* $C > 0$. *Por lo tanto, si* x_0 *se escoge suficientemente cercano a* α*, se tiene la convergencia*

$$\lim_{k\to\infty} x_k = \alpha,$$

con orden $p = \dfrac{1+\sqrt{5}}{2} \approx 1{,}618...$

Observación. Como en el método de Newton–Raphson, la convergencia del método de la secante no está siempre garantizada, pero cuando tiene lugar es bastante veloz, con un orden levemente inferior al de Newton–Raphson.

Sistemas de Ecuaciones no Lineales

Sea $\boldsymbol{f}:\mathbb{R}^n\to\mathbb{R}^n$ una función de varias variables, no-lineal. Se quiere resolver el sistema de ecuaciones $\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{0}$, donde $\boldsymbol{x}=(x_1,\dots,x_n)^{\mathrm{t}}\in\mathbb{R}^n$ representa al vector de incógnitas.

El método de Newton.

Una de las ventajas del método de Newton–Raphson además de su velocidad de convergencia, es que se puede generalizar fácilmente a sistemas de ecuaciones no lineales. Esta generalización se conoce como **método de Newton**.

Supongamos que $\boldsymbol{\alpha}=(\alpha_1,\dots,\alpha_n)^{\mathrm{t}}\in\mathbb{R}^n$ es la solución del sistema de ecuaciones, y que $\boldsymbol{f}=(f_1,\dots,f_n)^{\mathrm{t}}$ es dos veces diferenciable. Entonces, aplicando el desarrollo de Taylor para funciones de varias variables de $\boldsymbol{f}$ en torno a una aproximación de la raíz $\boldsymbol{x}^{(k)}=(x_1^{(k)},\dots,x_n^{(k)})^{\mathrm{t}}$, se tiene que:

$$\boldsymbol{0}=\boldsymbol{f}(\boldsymbol{\alpha})=\boldsymbol{f}(\boldsymbol{x}^{(k)})+\boldsymbol{Df}(\boldsymbol{x}^{(k)})\,(\boldsymbol{\alpha}-\boldsymbol{x}^{(k)})+\mathcal{O}\left(\|\boldsymbol{\alpha}-\boldsymbol{x}^{(k)}\|^2\right),$$

donde $\boldsymbol{Df}(\boldsymbol{x}^{(k)})$ es la **matriz Jacobiana** de $\boldsymbol{f}$ en $\boldsymbol{x}^{(k)}$:

$$\boldsymbol{Df}(\boldsymbol{x}^{(k)}):=\begin{pmatrix}\dfrac{\partial f_1}{\partial x_1}(\boldsymbol{x}^{(k)}) & \cdots & \dfrac{\partial f_1}{\partial x_n}(\boldsymbol{x}^{(k)})\\ \vdots & & \vdots\\ \dfrac{\partial f_n}{\partial x_1}(\boldsymbol{x}^{(k)}) & \cdots & \dfrac{\partial f_n}{\partial x_n}(\boldsymbol{x}^{(k)})\end{pmatrix}.$$

Notamos que cuando $\|\boldsymbol{\alpha}-\boldsymbol{x}^{(k)}\|$ es pequeño, el término $\mathcal{O}\left(\|\boldsymbol{\alpha}-\boldsymbol{x}^{(k)}\|^2\right)$ es mucho más pequeño aún y puede despreciarse en el desarrollo de Taylor anterior:

$$\begin{aligned}\boldsymbol{0}&=\boldsymbol{f}(\boldsymbol{x}^{(k)})+\boldsymbol{Df}(\boldsymbol{x}^{(k)})\,(\boldsymbol{\alpha}-\boldsymbol{x}^{(k)})+\mathcal{O}\left(\|\boldsymbol{\alpha}-\boldsymbol{x}^{(k)}\|^2\right)\\ &\approx\boldsymbol{f}(\boldsymbol{x}^{(k)})+\boldsymbol{Df}(\boldsymbol{x}^{(k)})\,(\boldsymbol{\alpha}-\boldsymbol{x}^{(k)}).\end{aligned}$$

Si además la matriz $\boldsymbol{Df}(\boldsymbol{x}^{(k)})$ es invertible, entonces podemos aproximar la raíz $\boldsymbol{\alpha}$ despejándola en la ecuación anterior:

$$\boldsymbol{\alpha}\approx\boldsymbol{x}^{(k)}-\boldsymbol{Df}(\boldsymbol{x}^{(k)})^{-1}\boldsymbol{f}(\boldsymbol{x}^{(k)}).$$

El **método de Newton** consiste en, dada la aproximación de la solución $\boldsymbol{x}^{(k)}$, tomar como nueva aproximación $\boldsymbol{x}^{(k+1)}$ el valor de la expresión anterior:

$$\boldsymbol{x}^{(k+1)}:=\boldsymbol{x}^{(k)}-\boldsymbol{Df}(\boldsymbol{x}^{(k)})^{-1}\boldsymbol{f}(\boldsymbol{x}^{(k)}),\qquad k=0,1,2,\dots$$

donde $\boldsymbol{x}^{(0)}$ es la aproximación inicial.

En la práctica no es necesario (ni conveniente) invertir la matriz $\boldsymbol{Df}(\boldsymbol{x}^{(k)})$, sino que se utiliza un método menos costoso: resolver en cada iteración el sistema de ecuaciones lineales siguiente:

$$\boldsymbol{Df}(\boldsymbol{x}^{(k)})\,(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) = -\boldsymbol{f}(\boldsymbol{x}^{(k)}).$$

Así, llamando $\boldsymbol{\delta x}^{(k)} := \boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}$, se obtiene el siguiente **algoritmo**:

> Dado $\boldsymbol{x}^{(0)} \in \mathbb{R}^n$,
> para $k = 0, 1, 2, \ldots$
>   resolver el SEL: $\boldsymbol{Df}(\boldsymbol{x}^{(k)})\,\boldsymbol{\delta x}^{(k)} = -\boldsymbol{f}(\boldsymbol{x}^{(k)})$,
> $\boldsymbol{x}^{(k+1)} := \boldsymbol{x}^{(k)} + \boldsymbol{\delta x}^{(k)}$,
> hasta que se satisfaga algún criterio de detención.

Observaciones.

1. Los teoremas de convergencia, y estimación del error del método de Newton–Raphson se pueden generalizar al caso de sistemas, reemplazando el valor absoluto por una norma vectorial. Así se obtiene que
$$\|\boldsymbol{\alpha} - \boldsymbol{x}^{(k+1)}\| \leq C\|\boldsymbol{\alpha} - \boldsymbol{x}^{(k)}\|^2, \qquad k = 0, 1, 2, \ldots$$
donde C es una constante positiva que depende de las derivadas primeras y segundas de $\boldsymbol{f}$.

2. Al igual que en el método de Newton–Raphson, si se desea calcular la solución $\boldsymbol{\alpha}$ del sistema con error menor que `tol`, puede usarse
$$\|\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}\| \leq \texttt{tol}$$
como criterio de detención.

Ejemplo: Resolver el siguiente sistema de ecuaciones con error menor que $\texttt{tol} = 10^{-5}$:

$$\begin{cases} y^2 + x^2 = 1 \\ y = x^2 \end{cases}$$

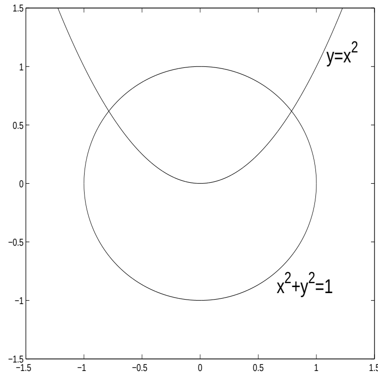
Solución: Las funciones a utilizar son:

$$\boldsymbol{f}(x, y) = \begin{pmatrix} x^2 + y^2 - 1 \\ y - x^2 \end{pmatrix}$$

$$\boldsymbol{Df}(x, y) = \begin{pmatrix} 2x & 2y \\ -2x & 1 \end{pmatrix}$$

Localización de las raíces:

Algoritmo:



(x_0, y_0): datos iniciales.
Para $k = 0, 1, 2, \ldots$

$$\text{resolver} \begin{pmatrix} 2x_k & 2y_k \\ -2x_k & 1 \end{pmatrix} \begin{pmatrix} \delta x_k \\ \delta y_k \end{pmatrix} = \begin{pmatrix} 1 - x_k^2 - y_k^2 \\ x_k^2 - y_k \end{pmatrix},$$

$x_{k+1} := x_k + \delta x_k,$

$y_{k+1} := y_k + \delta y_k,$

hasta que $\sqrt{\delta x_k^2 + \delta y_k^2} < \texttt{tol}$.

Resultados obtenidos:

x	y
1.00000000000000	1.00000000000000
0.83333333333333	0.66666666666667
0.78809523809524	0.61904761904762
0.78615406630609	0.61803444782168
0.78615137776208	0.61803398874999

VAD-DMH.