

Investigación

Mejoras de convergencia de algoritmos

Algorithm convergence improvements

Federico Ruiz López

Revista de Investigación



Volumen XII, Número 1, pp. 017-055, ISSN 2174-0410
Recepción: 26 Sep'21; Aceptación: 27 Oct'21

1 de abril de 2022

Resumen

En este artículo se analiza el proceso de obtención de soluciones de ecuaciones no lineales, mediante procedimientos iterativos, basados en los teoremas de punto fijo. Se proponen algoritmos de mejora para ser implementados en un lenguaje de programación apropiado. Los resultados parten de la imposibilidad de resolver ecuaciones algebraicas de grado superior a cuatro mediante métodos puramente algebraicos. Este trabajo también se puede considerar una introducción a la teoría de sistemas dinámicos discretos.

Palabras Clave: Sistemas dinámicos, puntos fijos, algoritmos iterativos, convergencia.

Abstract

This article analyzes the process of obtaining solutions of nonlinear equations, through iterative procedures, based on the fixed point theorems. The article proposes some improvement algorithms to be implemented in an appropriate programming language. The results are based on the impossibility of solving algebraic equations of degree higher than four using purely algebraic methods. This work can also be considered an introduction to the theory of discrete dynamical systems.

Keywords: Dynamic systems, fixed points, iterative algorithms, convergence.

1. El cálculo de la raíz cuadrada de 2

Supóngase que se quiere determinar la raíz cuadrada de 2 sin una calculadora. Sólo se dispone de lápiz, papel y grandes dosis de paciencia. El problema se traduce en determinar la única solución positiva de la ecuación cuadrática:

$$x^2 - 2 = 0 \tag{1}$$

Como es sabido, la solución $\alpha = \sqrt{2}$, es un número real que trasciende al conjunto de los números racionales. Por tanto, debe existir una expresión decimal en forma de serie infinita,

$$\alpha = \sum_{n=0}^{+\infty} a_n \cdot 10^{-n} = a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \dots + \frac{a_n}{10^n} + \dots$$

donde $a_i \neq 0$, para una cantidad no finita de términos. El primer objetivo es determinar un algoritmo programable, con el fin de determinar la mayor cantidad de cifras del número α , recordando que:

$$a_i \in \{0, 1, 2, 3, \dots, 8, 9\}$$

son dígitos.

1.1. Ejemplo 1 (Dinámicas inútiles)

Se puede observar que cualquier número positivo que verifica la ecuación 1, también verifica la relación:

$$x = \frac{2}{x}$$

esto es, se trata de un punto fijo de la función $f(x) = \frac{2}{x}$. Gráficamente se tendría la configuración de la figura 1.

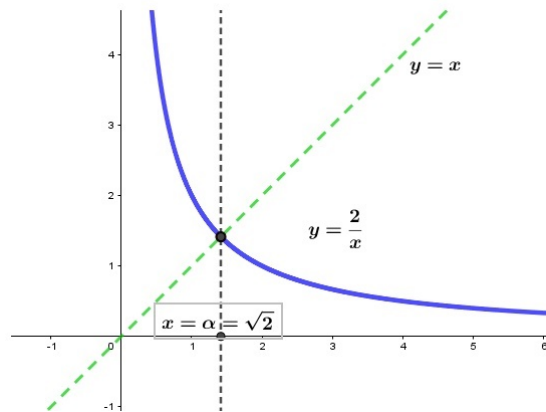


Figura 1. La solución buscada se encuentra en la intersección de la función $y = \frac{2}{x}$ con la bisectriz del primer cuadrante.

Esto motiva a definir la función real de variable real:

$$f : [1, 2] \rightarrow [1, 2]$$

$$x \rightarrow f(x) = \frac{2}{x}$$

Es claro que f conserva el intervalo al tratarse de una función decreciente, que verifica:

$$f(1) = 2, \quad f(2) = 1 \rightarrow f([1,2]) \subset [1,2]$$

Denotando por $I = [1,2]$ dicho intervalo, se puede considerar el par (I, f) , que a partir de ahora se denominará en este artículo «*sistema dinámico*».

Estudiar la dinámica generada por f en el intervalo I , consiste en determinar la estructura topológica de los conjuntos que definen las órbitas de cada punto, esto es,

$$Orb_f(x_0) = \left\{ x_0, f(x_0), f^2(x_0), f^3(x_0) \cdots \right\}$$

para cada $x_0 \in I$.

Para el caso mencionado, la dinámica generada es bastante simple, ya que todos los puntos generan órbitas de longitud dos (ciclos de orden 2) y por tanto todas las iteraciones son cerradas. En efecto, si $x_0 \in I = [1,2]$,

$$f(x_0) = \frac{2}{x_0} \rightarrow f^2(x_0) = f\left(\frac{2}{x_0}\right) = \frac{2}{\frac{2}{x_0}} = x_0$$

Por tanto,

$$Orb_f(x_0) = \left\{ x_0, \frac{2}{x_0} \right\}$$

Desde el punto de vista del *cálculo de la raíz* mediante iteraciones, este es el peor de los escenarios posibles, porque el punto fijo α no podrá ser alcanzado nunca, a partir de sucesivas iteraciones de valores racionales. En cambio desde el punto de vista de la *teoría de sistemas dinámicos*, es la mejor de las situaciones, ya que se conoce perfectamente la dinámica generada por cualquier punto y la estructura topológica del sistema dinámico.

¿Cuál es el objetivo? Establecer un sistema dinámico apropiado que encierre a la solución de la ecuación como punto fijo, y que pueda ser alcanzado rápidamente por sucesivas iteraciones iniciadas desde cualquier punto relativamente próximo a él. Esto es particularmente interesante desde una perspectiva computacional. En este sentido, se pueden encontrar muchos sistemas (I, F) diferentes que cumplan dicho cometido, como se muestra en el siguiente ejemplo.

1.2. Ejemplo 2 (Atractores y repulsores)

Partiendo de la ecuación 1, se pueden multiplicar por $x > 0$ los dos miembros de la ecuación para obtener la equivalente:

$$x^3 - 2x = 0 \tag{2}$$

La única raíz positiva de dicha ecuación cúbica es α . Dicha solución también deberá cumplir la relación:

$$x = \sqrt[3]{2x}, \quad x > 0 \tag{3}$$

Se transforma el problema de determinar α en un problema de punto fijo. En efecto, si se conviene en denotar por $J = [0,2]$, y:

$$F : J \rightarrow J$$

$$x \rightarrow F(x) = \sqrt[3]{2x}$$

se tiene un nuevo sistema dinámico (J, F) . Puesto que F es una función continua y creciente en J y:

$$F(0) = 0, \quad F(2) = \sqrt[3]{4} < \sqrt[3]{8} = 2$$

se conserva el intervalo,

$$F([0, 2]) \subset [0, 2]$$

y, consecuentemente, se garantiza la existencia de al menos un punto fijo. La unicidad se deriva de la monotonía de la función con valores positivos, aunque en realidad F tiene dos puntos fijos, situados en el origen de coordenadas y en $x = \sqrt{2}$.

En efecto:

$$F'(x) = \frac{2}{3} \cdot \frac{1}{\sqrt[3]{4x^2}} > 0, \quad \forall x > 0$$

Gráficamente, la situación aparece reflejada en la figura 2:

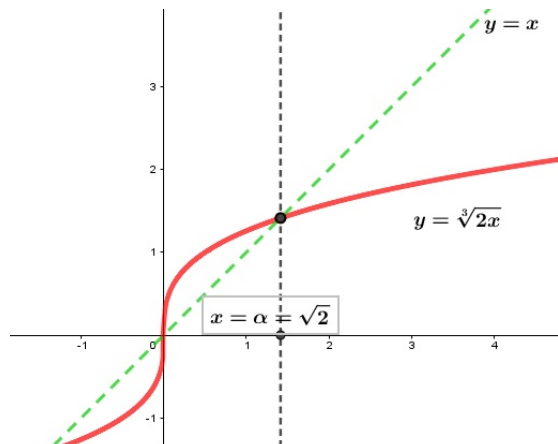


Figura 2. El valor de α se determina como intersección de la función $y = \sqrt[3]{2x}$ con la bisectriz del primer cuadrante.

Tanto de la expresión de la primera derivada, como de la gráfica de la figura 2, se observa que existe una singularidad en $x = 0$, donde la función derivada no está bien definida:

$$\lim_{x \rightarrow 0^+} F'(x) = \lim_{x \rightarrow 0^+} \frac{2}{3} \cdot \frac{1}{\sqrt[3]{4x^2}} = +\infty$$

Esto hace que el origen se convierta en un punto fijo imposible de alcanzar por iteración.

¿Cómo es la dinámica generada por el sistema (J, F) ?

Partiendo del valor inicial $x_0 = 1 \in J$, se obtiene la órbita asociada:

$$Orb_F(1) = \left\{ 1, \sqrt[3]{2}, \sqrt[3]{2 \cdot \sqrt[3]{2}}, \dots \right\} = \left\{ 1, 2^{\frac{1}{3}}, 2^{\frac{4}{9}}, 2^{\frac{13}{27}}, 2^{\frac{40}{81}}, \dots, 2^{\frac{3^n - 1}{2 \cdot 3^n}}, \dots \right\}$$

donde:

$$F^n(1) = 2^{\frac{3^n - 1}{2 \cdot 3^n}} = (\sqrt{2})^{1 - \frac{1}{3^n}}$$

Por tanto, la sucesión de iteraciones de F tienden al valor de α :

$$\lim_{n \rightarrow +\infty} F^n(1) = \lim_{n \rightarrow +\infty} (\sqrt{2})^{1 - \frac{1}{3^n}} = \sqrt{2}$$

De hecho, todas las órbitas que parten de valores estrictamente positivos, muestran el mismo comportamiento, i.e., tienden al punto fijo α . En efecto, partiendo de un valor cualquiera $x_0 \in (0, 2]$, es relativamente sencillo probar que:

$$F^n(x_0) = (\sqrt{2})^{1 - \frac{1}{3^n}} \cdot x_0^{\frac{1}{3^n}}$$

para $n = 1, 2, 3, \dots$. En consecuencia dicho punto fijo se convierte en un *atractor del sistema*, mientras que el origen se convierte en un **punto fijo repulsor**. Aunque los valores iniciales estén próximos al origen, las sucesivas iteraciones se van haciendo cada vez más próximas a $\alpha = \sqrt{2}$.

Esto se puede ver sin más que considerar la expresión de F en forma de potencia fraccionaria. De este modo:

$$F(x) = 2^{\frac{1}{3}} \cdot x^{\frac{1}{3}}, \quad \forall x \in J = [0, 2]$$

Definiendo la sucesión $x_n = F^n(x_0)$, los primeros valores son:

$$\begin{aligned} x_1 &= 2^{\frac{1}{3}} \cdot x_0^{\frac{1}{3}} \\ x_2 &= 2^{\frac{1}{3}} \cdot x_1^{\frac{1}{3}} = 2^{\frac{1}{3}} \cdot 2^{\frac{1}{9}} \cdot x_0^{\frac{1}{9}} = 2^{\frac{1}{3} + \frac{1}{9}} \cdot x_0^{\frac{1}{9}} \\ x_3 &= 2^{\frac{1}{3}} \cdot x_2^{\frac{1}{3}} = 2^{\frac{1}{3}} \cdot 2^{\frac{1}{9}} \cdot 2^{\frac{1}{27}} \cdot x_0^{\frac{1}{27}} = 2^{\frac{1}{3} + \frac{1}{9} + \frac{1}{27}} \cdot x_0^{\frac{1}{27}} \end{aligned}$$

y en general:

$$x_n = 2^{\frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots + \frac{1}{3^n}} \cdot x_0^{\frac{1}{3^n}}$$

Ahora bien, calculando la suma de los n primeros términos de una serie geométrica,

$$S_n = \frac{1}{3} + \frac{1}{9} + \frac{1}{27} + \dots + \frac{1}{3^n} = \frac{1}{3} \cdot \left(\frac{1 - \frac{1}{3^n}}{1 - \frac{1}{3}} \right) = \frac{1}{2} \cdot \left(1 - \frac{1}{3^n} \right)$$

se tiene que:

$$x_n = F^n(x_0) = 2^{\frac{1}{2} \cdot (1 - \frac{1}{3^n})} \cdot x_0^{\frac{1}{3^n}} = (\sqrt{2})^{1 - \frac{1}{3^n}} \cdot x_0^{\frac{1}{3^n}}$$

y por tanto:

$$\lim_{n \rightarrow +\infty} x_n = \lim_{n \rightarrow +\infty} F^n(x_0) = \sqrt{2}, \quad \forall x_0 \in J - \{0\}$$

La ventaja de este modelo frente al mostrado en el ejemplo 1, es evidente. Partiendo de cualquier valor inicial, e iterando suficientemente la función, se puede aproximar el valor de la raíz de 2, con tanta precisión como se quiera. No obstante, parece poco razonable que para determinar el valor de una raíz cuadrada, se precise el cálculo de infinitud de raíces cúbicas, cuyo algoritmo es, en principio, algo más complicado. En efecto, se desconoce un algoritmo simple que permita determinar la expresión decimal infinita del número:

$$2^{\frac{40}{81}} = \sqrt[3]{\sqrt[3]{\sqrt[3]{\sqrt[3]{2^{40}}}}}$$

Se debe recordar que lo que se pretende es obtener números irracionales mediante aproximaciones sucesivas cuyo cálculo sea relativamente simple, como es el caso de aproximaciones puramente racionales. Este planteamiento fuerza a buscar una tercera vía.

1.3. Ejemplo 3: (El método de Newton)

Dada una función real de variable real $f : I \rightarrow I$, se llamará *transformada de Newton de f* a la nueva función:

$$F(x) = x - \frac{f(x)}{f'(x)}, \quad \forall x \in I$$

Se debe observar que, para que F esté bien definida en el intervalo, la función f debe ser al menos de clase $C^1(I)$, con derivada no nula. El comportamiento de la derivada, determina en cierto grado la dinámica topológica del nuevo sistema (I, F) .

Además se cumple que:

$$f(c) = 0 \leftrightarrow F(c) = c, \quad c \in I$$

Por tanto determinar los ceros de la función f en el intervalo I equivale a determinar los puntos fijos de F en dicho intervalo. Aplicado al caso que nos ocupa, para $f(x) = x^2 - 2$, la transformada de Newton será:

$$F(x) = x - \frac{x^2 - 2}{2x} = \frac{1}{2} \cdot \left(x + \frac{2}{x} \right)$$

cuya gráfica se puede apreciar en la figura 3.

Puesto que F no está bien definida en $x = 0$, se puede considerar un entorno de $x = 1,5$, por ejemplo el intervalo compacto

$$I = [1, 2]$$

No es complicado probar que

$$F([1, 2]) \subset [1, 2]$$

.

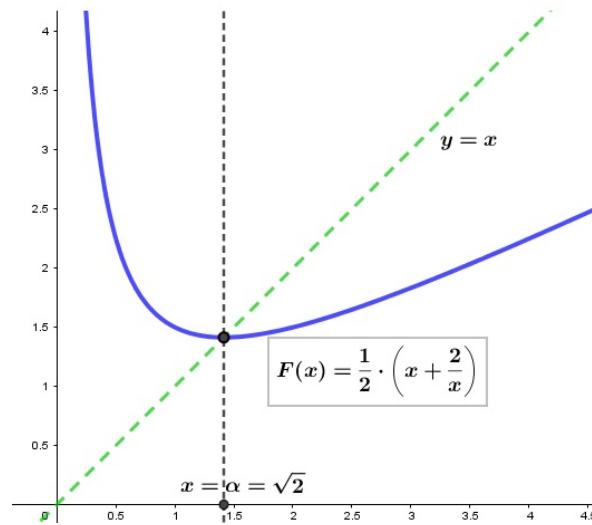


Figura 3. Transformada de Newton de la función $f(x) = x^2 - 2$

De hecho, al ser

$$F(1) = \frac{3}{2}, \quad F(2) = \frac{3}{2}$$

, se podría aplicar el *Teorema de Rolle* para probar que existe un valor $c \in (1, 2)$ donde la derivada se anula:

$$F(1) = F(2) \Rightarrow \exists c \in [1, 2] : F'(c) = 0$$

Lo curioso del caso, es que este valor es justamente el punto fijo del sistema. En efecto, derivando F se obtiene:

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

Así:

$$F'(x) = 0 \Leftrightarrow x^2 - 2 = 0 \Leftrightarrow x = \pm\sqrt{2}$$

Estos hechos son relevantes y paradigmáticos, porque el hecho de que el punto fijo tenga derivada nula, mejora la velocidad de convergencia, como se verá más adelante en su marco teórico.

Órbitas y convergencia

Si se toma como valor inicial $x_0 = 1$, se puede probar que todas las órbitas son convergentes al punto estacionario de la función F . Por tanto, la dinámica generada por este sistema es bastante útil. Lo interesante es que, mediante este procedimiento, se pueden ir obteniendo sucesivas aproximaciones racionales de $\alpha = \sqrt{2}$.

En efecto, cálculos más o menos sencillos, permiten escribir los primeros valores iterados:

$$Orb_F(1) = \left\{ 1, \frac{3}{2}, \frac{17}{12}, \frac{577}{408}, \frac{665857}{470832}, \dots \right\}$$

Se debe recordar que

$$\alpha = \sqrt{2} = 1,414213562373095\dots$$

, y que el valor aportado tras la cuarta iteración es una fracción cuya expresión decimal es:

$$\frac{665857}{470832} \approx 1,41421356237464\dots$$

Es decir, *con sólo cuatro iteraciones se obtiene una aproximación de ¡11 cifras decimales exactas!*, así como la curiosa aproximación racional:

$$\sqrt{2} \approx \frac{665857}{470832}$$

Se debe observar que la órbita del 2, enlaza con la del 1, a partir de la primera iteración.

¿Qué es lo que provoca esta velocidad tan rápida de convergencia?

Pues la cuestión es clara. Cuanto más plana sea la curva en torno al punto estacionario mayor será dicha velocidad de convergencia. Y esto sucede con la *transformada de Newton*. En efecto, si se deriva la expresión de F :

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

esto es:

$$F'(x) = \frac{f(x) \cdot f''(x)}{f'(x)^2}, \quad \forall x \in I$$

Por tanto

$$f(c) = 0 \leftrightarrow F'(c) = 0$$

, y el punto fijo de F es punto estacionario y por tanto atractor, en un entorno reducido del punto.

Estas consideraciones sugieren algoritmos que partan de transformadas de la forma

$$F(x) = \int \frac{f(x)}{g(x)} dx \leftrightarrow F'(x) = \frac{f(x)}{g(x)}$$

cuya derivada se anula precisamente donde la función inicial f posee ceros, esto es, los puntos estacionarios de F son los ceros de f . En estos casos la velocidad de convergencia mejora notablemente.

1.4. Ejemplo 4: (A vueltas con el método de Newton)

Se va a mostrar un último ejemplo, que puede servir para dar luz a estos asuntos. Se considera en este caso la función real de variable real

$$g(x) = x^3 - 2x$$

cuyos ceros son precisamente

$$Z(g) = \{-\sqrt{2}, 0, +\sqrt{2}\}$$

lo cual confirma que en cada iteración van surgiendo dos nuevas cifras decimales que coinciden con el valor esperado. Aquí la convergencia es mucho más fuerte que si se consideran órbitas alejadas del punto estacionario α .

Cuestiones abiertas El análisis de los ejemplos mostrados lleva a plantear una serie de cuestiones, en las que se tratará de profundizar:

- ¿Qué función es la más apropiada para iterar?
- ¿Cómo caracterizar la dinámica del sistema a partir de su derivada?
- ¿El valor inicial condiciona la velocidad de convergencia?
- ¿Cómo estimar el error cometido al realizar la aproximación?
- ¿Es posible mejorar el algoritmo aplicado?

El objetivo es demostrar que *cualquier algoritmo de iteración es mejorable*, i.e., se puede encontrar un nuevo proceso que permite determinar el punto fijo del sistema, con mucha más rapidez. Se comenzará exponiendo los conceptos fundamentales que permiten fundamentar el estudio de estas cuestiones para formalizar este notable resultado. Conceptos como *algoritmo*, *órbita* o *velocidad de convergencia*, serán definidos dentro de la topología de la recta real. Una vez precisados estos conceptos, se podrá dar una respuesta razonada a cada una de las cuestiones planteadas.

2. Conceptos básicos sobre sistemas dinámicos

Def.: Sea $I = [a, b] \subset \mathbb{R}$ un intervalo compacto (cerrado y acotado) de la recta real y $f : I \rightarrow I$ una función continua. Al par (I, f) se le llamará **sistema dinámico discreto**.

Dado un punto $x_0 \in I$, se define **la órbita del punto**, como el conjunto

$$\text{Orb}_f(x_0) = \{x_0, x_1, x_2, \dots, x_n\}$$

donde

$$x_n = f^n(x_0), \quad n = 1, 2, 3, \dots$$

Def.: Se define el conjunto **omega-límite** del sistema dinámico (X, f) como

$$\Omega(f) = \cup_{x \in I} \text{Orb}_f(x)$$

esto es, como la unión de las órbitas de cada punto. Determinar la estructura topológica de este conjunto es el principal objetivo en la teoría de sistemas dinámicos discretos.

Ejemplo: La función $f : [0, 1] \rightarrow [0, 1]$ definida por $f(x) = x^2$ define el sistema

$$([0, 1], f)$$

cuya dinámica es bastante simple. Posee dos puntos fijos,

$$f(0) = 0, \quad f(1) = 1$$

y la órbita de cualquier punto $x \in (0, 1)$, es convergente a cero. Por tanto $x = 0$ se convierte en un **punto fijo atractor** y $x = 1$ un **punto fijo repulsor**.

Observación 1: El concepto de sistema dinámico es mucho más general que el que se va a tratar aquí. Basta con extender la noción a espacios topológicos X en los que haya definida una

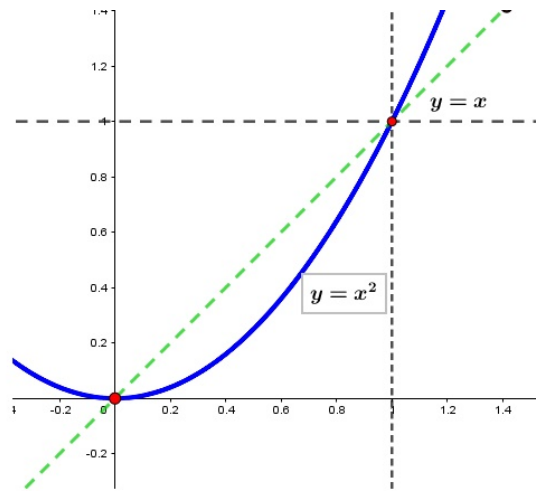


Figura 5. La función cuadrática $y = x^2$ genera una dinámica simple sin órbitas periódicas y dos puntos fijos.

métrica. En realidad toda aplicación continua de un espacio topológico en sí mismo, produce una transformación que se puede interpretar como una dinámica de deformación. La iteración de esa deformación del espacio, a intervalos discretos de tiempo, es lo que se entiende como **dinámica discreta**.

Observación 2: Si la órbita de un punto es cerrada, en el sentido de que exista un $k \in \mathbb{N}$, tal que

$$f^k(x_0) = x_0$$

entonces se dice que existe un **ciclo de orden k** , siendo k el mínimo de los números naturales que cumple tal condición. Cuando $k = 1$, se obtiene un *punto fijo* de f , Para $k = 2$ ciclos de orden 2, y así sucesivamente. Es importante recordar que cuando $X = \mathbb{R}$, la existencia de ciclos de cierto orden, fuerza la existencia de ciclos de otros órdenes, de acuerdo con el *teorema de Sarkovskii*.

Teorema de Sarkovskii

Sea una aplicación continua $f : \mathbb{R} \rightarrow \mathbb{R}$. Si esta función tiene un punto periódico de orden k , entonces tiene puntos periódicos de todos los períodos inferiores a k según el siguiente orden:

$$1 \ll 2 \ll 4 \ll 8 \ll \dots \ll 2^n \cdot 5 \ll 2^n \cdot 3 \ll \dots \ll 2 \cdot 5 \ll 2 \cdot 3 \ll \dots \ll 7 \ll 5 \ll 3$$

El orden debe ser interpretado de derecha a izquierda. En este sentido, si existe un ciclo de orden 8, existen ciclos de orden 4, 2 y 1, pero no necesariamente de orden 16. Si existe un ciclo de orden $6 = 2 \cdot 3$, existen ciclos de ordenes 10, 14, 18... pero no necesariamente de orden impar. En consecuencia, si existe un ciclo de orden 3, existen órbitas periódicas de todos los períodos (Teorema de Li-Yorke).

Este teorema es óptimo, es decir, si $m \ll k$ según el orden precedente, existen aplicaciones continuas con puntos periódicos de periodo m pero sin punto periódico de período k .

Este resultado implica que si se encuentra una función continua en un intervalo compacto, conteniendo un ciclo de orden tres, habrá órbitas de cualquier orden. Este hecho induce comportamientos caóticos en el sistema. Una de las expresiones conocidas por los teóricos de estas cuestiones, es que "*período tres implica caos*". Aunque realmente el concepto de caos, está más bien asociado a la existencia de órbitas densas, o cierta dependencia sensible a las condiciones iniciales.

Def.: Dado un sistema dinámico (I, f) , se dice que existe una **órbita densa**, si para cierto valor inicial $x_0 \in I$, y para cada $x \in I$

$$\forall \epsilon > 0, \exists k(\epsilon) \in \mathbb{N} : |f^k(x_0) - x| < \epsilon$$

Un sistema dinámico con una órbita densa se dice **transitivo**.

Def.: Se dice que el sistema (I, f) presenta **dependencia sensible a las condiciones iniciales** en el punto $x \in I$, si existe una constante positiva $\epsilon > 0$, (*constante de sensibilidad*) tal que

$$\forall \delta > 0, \exists y \in I : |y - x| < \delta \rightarrow |f^n(x) - f^n(y)| > \epsilon$$

para cierto $n \in \mathbb{N}$. Cuando esta propiedad es extensible a cualquier $x \in I$, se dice que el sistema tiene dependencia sensible a las condiciones iniciales.

Intuitivamente, esta noción sugiere que las órbitas generadas por puntos relativamente próximos se alejan de acuerdo con la constante de sensibilidad. Una primera intuición puede hacer pensar que las órbitas generadas por puntos próximos, deben ser muy similares, de manera que bajo iteración, la órbitas no se distancien demasiado. Esto no sucede en la realidad. Cuando un sistema presenta dependencia sensible a las condiciones iniciales, estas órbitas pueden diferenciarse tanto como se estime. Lo cual convierte este tipo de sistemas en impredecibles, desde el punto de vista del determinismo numérico. Y poco útiles para los propósitos que se pretende alcanzar.

En 1989, el matemático R. Devaney [7] formuló la siguiente definición, que trata de delimitar la noción de caos.

Sistemas caóticos

Sea X un espacio métrico, no finito y $f : X \rightarrow X$ una aplicación. Se dice que el sistema (X, f) es caótico si se verifican las siguientes condiciones:

1. El conjunto de puntos periódicos de la función f es denso en X .
2. El sistema es topológicamente transitivo (posee una órbita densa).
3. El sistema tiene dependencia sensible a las condiciones iniciales.

Se debe observar que, mientras que la primera y segunda condición son topológicas (se pueden definir en espacios más generales), la tercera es métrica, por tanto, la noción de caos sólo tiene sentido en espacios métricos. Aunque existen otras definiciones sobre caos, esta es, hoy día, la más comúnmente aceptada.

Existen ciertas relaciones entre las condiciones. Banks [3] ha probado que cuando se trabaja en espacios métricos, la primera y segunda condición implican la tercera. Y cuando X es un espacio métrico completo con base numerable, entonces la primera y segunda condición son equivalentes. Y como éstas implican la tercera, bastaría analizar si la tercera condición implica la primera para tener la equivalencia entre todas ellas (una cuestión que se analizará más adelante en profundidad). De este modo, en el caso particular $X = \mathbb{R}$, la existencia de una órbita densa implica dependencia sensible a las condiciones iniciales, y por tanto la existencia de caos.

3. Métodos iterativos

Supóngase que se quiere resolver la ecuación $f(x) = 0$, donde $f : [a, b] \rightarrow \mathbb{R}$ es una función real de variable real. Siempre es posible encontrar una función g , sobre cierto intervalo compacto, de suerte que

$$f(\alpha) = 0 \leftrightarrow g(\alpha) = \alpha$$

En efecto basta definir $g(x) = x + f(x)$. De suerte que cualquier problema de determinar los ceros de una función, se transforma en un problema de punto fijo. Sin embargo este no es el único modo posible, lo que dota de interés nuestro estudio. Se necesita desarrollar herramientas que permitan aproximar las raíces de f por iteración de g , acotando los errores cometidos.

3.1. Ejemplo 1: (La transformada de Newton-Raphson)

Sea $f(x)$ una función de clase $C^2(I)$ y α una raíz en dicho intervalo, $f(\alpha) = 0$. Se supone, además, que $f'(\alpha) \neq 0$. Se define la *transformada de Newton-Raphson* de f como la función:

$$F(x) = x - \frac{f(x)}{f'(x)}, \quad x \in J \subset I$$

En tal caso,

$$f(\alpha) = 0 \leftrightarrow F(\alpha) = \alpha$$

como se puede comprobar fácilmente.

3.2. Ejemplo 2: (Generalizando)

Se puede tratar de generalizar el caso anterior y definir la transformada

$$F(x) = x + f(x) \cdot g(x)$$

de f , para cierta función derivable g en un entorno de $\alpha \in I$. Se supone, como antes, que $f(\alpha) = 0$. Si se quiere que dicho valor se convierta en punto crítico de F ,

$$F'(x) = 1 + f'(x) \cdot g(x) + f(x) \cdot g'(x)$$

Si se impone la condición $F'(\alpha) = 0$, se tendrá:

$$1 + f'(\alpha) \cdot g(\alpha) + \cancel{f(\alpha) \cdot g'(\alpha)} = 0$$

Esto es,

$$1 + f'(\alpha) \cdot g(\alpha) = 0 \rightarrow g(\alpha) = \frac{-1}{f'(\alpha)}$$

Esto motiva definir,

$$g(x) = -\frac{1}{f'(x)}, \quad \forall x \in J \subset I$$

y por tanto

$$F(x) = x - \frac{f(x)}{f'(x)}, \quad x \in J \subset I$$

En consecuencia, la *transformada de Newton-Raphson* es el prototipo de función que transforma los ceros de f en puntos críticos de F .

3.3. Ejemplo 3: (El caso polinómico)

Se sabe que toda ecuación cúbica se puede expresar en la forma

$$x^3 + px + q = 0$$

con $p, q \in \mathbb{R}$. Por comodidad en la notación, se la puede suponer escrita en la forma

$$x^3 = px + q, \quad p, q \in \mathbb{R}$$

Es preciso observar que las raíces del polinomio, son los puntos fijos de la función

$$g(x) = \sqrt[3]{px + q}$$

En general, cualquier ecuación algebraica en la forma

$$p_n(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_2x^2 + a_1x + a_0 \in \mathbb{R}[X]$$

se puede transformar en la expresión equivalente:

$$x = \sqrt[n]{P_{n-1}(x)} = \sqrt[n]{-a_{n-1}x^{n-1} - a_{n-2}x^{n-2} - \dots - a_2x^2 - a_1x - a_0}$$

y por tanto determinar la raíz de un polinomio, se convierte, bajo este prisma, en un problema de puntos fijos.

La cuestión que interesa es, de todas las transformaciones posibles de la ecuación $f(x) = 0$ en $g(x) = x$, discernir aquellas cuya dinámica sea lo más precisa, en el sentido de que el punto fijo sea un *punto fijo fuertemente atractor*. En este sentido, van orientadas las siguientes definiciones y resultados.

Def.: Dada la sucesión de números reales $\{x_n\}$, se dice que es **convergente** al número real $\alpha \in \mathbb{R}$, tal que

$$\forall \epsilon > 0, \quad \exists n_\epsilon \in \mathbb{N} : |x_n - \alpha| < \epsilon \quad \forall n \geq n_\epsilon$$

En tal caso, se escribe

$$\alpha = \lim_{n \rightarrow +\infty} x_n$$

Propiedad 1: El límite de una sucesión de números reales, de existir, es único. Basta aplicar un sencillo argumento de acotación para llegar a que

$$|\alpha - \beta| < \epsilon, \quad \forall \epsilon > 0$$

para α, β posibles límites distintos.

Propiedad 2: Toda sucesión convergente de números reales está acotada, i.e.,

$$|x_n| \leq M, \quad \forall n \in \mathbb{N}$$

Notación: Para sucesiones convergentes, se define el **error de aproximación** ϵ_n como

$$\epsilon_n = |d_n| = |x_n - \alpha| \geq 0 \quad \forall n \in \mathbb{N}$$

Teorema del punto fijo

Sea $g : [a, b] \rightarrow \mathbb{R}$ una función derivable que verifica:

1. $g([a, b]) \subset (a, b)$
2. $|g'(x)| < 1, \quad \forall x \in [a, b]$

Entonces, existe un único $\alpha \in [a, b]$ tal que $g(\alpha) = \alpha$. Además, para todo $x_0 \in [a, b]$, la sucesión $\{x_n\}$ generada por la iteración

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, 3, \dots$$

converge a α .

Demostración: Sea $h(x) = g(x) - x$, para todo $x \in [a, b]$. Entonces $h(x)$ es una función continua y derivable (aunque no se puede afirmar mucho más acerca de la continuidad de la derivada).

Por otro lado, es claro que

$$h(a) = g(a) - a > 0, \quad h(b) = g(b) - b < 0$$

por lo que se verifican las condiciones del Teorema de Bolzano. En consecuencia, existe un $\alpha \in (a, b)$ tal que

$$h(\alpha) = 0 \leftrightarrow g(\alpha) = \alpha$$

esto es, α es un punto fijo.

Para demostrar la unicidad, se supone que existen dos valores α, β , puntos fijos de g . En tal caso

$$g(\alpha) = \alpha, \quad g(\beta) = \beta$$

No es restrictivo suponer $\alpha < \beta$. Aplicando el Teorema del Valor Medio, existe un $c \in (\alpha, \beta)$, tal que

$$g(\beta) - g(\alpha) = g'(c) \cdot (\beta - \alpha)$$

Pero entonces

$$\beta - \alpha = g'(c) \cdot (\beta - \alpha) \leftrightarrow g'(c) = 1$$

en contra de la hipótesis planteada en la segunda condición.

Sea ahora $\{x_n\}$ la sucesión generada a partir del valor inicial $x_0 \in [a, b]$, mediante la iteración

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

Si se denota por

$$K = \sup_{x \in [a, b]} |g'(x)| < 1$$

y por $\epsilon_n = |x_n - \alpha|$, se puede escribir:

$$\begin{aligned} \epsilon_n = |x_n - \alpha| &= |g(x_{n-1}) - g(\alpha)| = |g'(c_n)| \cdot |x_{n-1} - \alpha| \leq K \cdot \epsilon_{n-1} \leq \\ &\leq K^2 \cdot \epsilon_{n-2} \leq K^3 \cdot \epsilon_{n-3} \leq \dots \leq K^n \cdot \epsilon_0 \end{aligned}$$

Tomando límites es claro que

$$\lim_{n \rightarrow +\infty} \epsilon_n = \lim_{n \rightarrow +\infty} K^n \cdot \epsilon_0 = 0$$

al ser $0 < K < 1$, lo que prueba la convergencia del método [cqd].

Del teorema expuesto, se deduce que interesa trabajar con funciones al menos de clase C^1 en el intervalo compacto, de suerte que la constante K sea lo más pequeña posible. De este modo se consigue acelerar la convergencia.

4. Orden de convergencia

Def.: Sea $\{x_n\}$ una sucesión de números reales convergente a $\alpha \in \mathbb{R}$, con $x_n \neq \alpha$, para todo $n \in \mathbb{N}$. Sean $p, \lambda \in \mathbb{R}^+$. Si

$$\lim_{n \rightarrow +\infty} \frac{\epsilon_{n+1}}{\epsilon_n^p} = \lambda \neq 0$$

entonces se dice que la sucesión $\{x_n\}$ tiene **orden de convergencia** p con *constante de error asintótica* λ .

- Si $p = 1$, se trata de **convergencia lineal**.
- Si $p = 2$, se trata de **convergencia cuadrática**.
- Si $p = 3$, se trata de **convergencia cúbica**.

y así sucesivamente. Es claro que si la constante asintótica λ es menor que 1,

$$0 < \lambda < 1 \rightarrow \epsilon_{n+1} \approx \lambda \cdot \epsilon_n^p, \quad n \rightarrow +\infty$$

El orden de convergencia marca de algún modo, la velocidad de convergencia de la sucesión. De manera que converger cuadráticamente es más potente que converger de manera lineal. Y en procesos iterativos de cálculo esto implica que en cada iteración se obtienen un mayor número de cifras decimales exactas. En efecto si

$$\epsilon_n < 10^{-(k+1)}$$

la aproximación obtenida tras la n -ésima iteración, proporciona al menos k cifras decimales exactas. Esto permite, conocida la expresión para el error de aproximación, determinar el número de iteraciones necesarias para acotar el error cometido.

4.1. Ejemplo 1: (convergencia lineal)

Sea $\lambda \in (0, 1)$ y se define por recurrencia la sucesión

$$x_0 = 1, \quad x_{n+1} = \lambda \cdot x_n, \quad n = 0, 1, 2, 3, \dots$$

Entonces x_n converge linealmente a cero.

4.2. Ejemplo 2: (convergencia cuadrática)

Sea $\lambda \in (0, 1)$ y se considera la sucesión definida por la recurrencia

$$x_0 = 1, \quad x_{n+1} = \lambda \cdot (x_n)^2, \quad n = 0, 1, 2, 3, \dots$$

Entonces x_n converge cuadráticamente a cero. De hecho, se puede comprobar que la sucesión tiene por término general

$$x_n = \lambda^{2^n - 1}$$

4.3. Ejemplo 3: (convergencia de orden p)

Sea $\lambda \in (0, 1)$, $p \geq 2$ y se considera la sucesión definida por la recurrencia

$$x_0 = 1, \quad x_{n+1} = \lambda \cdot (x_n)^p, \quad n = 0, 1, 2, 3, \dots$$

Entonces x_n converge a cero con orden de convergencia p . La expresión que determina el término general es de la forma

$$x_n = \lambda^{\frac{p^n - 1}{p - 1}}, \quad n = 1, 2, 3, \dots$$

que generaliza los ejemplos anteriores.

4.4. Velocidad de convergencia

Corolario: Las sucesiones geométricas convergentes, poseen un tipo de convergencia lineal.

Se supone que se tienen sucesiones convergentes al mismo número real. No es restrictivo suponer que dicho valor es nulo. Por ejemplo, se pueden considerar las sucesiones clásicas

$$a_n = \frac{1}{n}, \quad b_n = \frac{1}{n^2}, \quad c_n = \frac{1}{n^p}$$

con $p > 2$. Es claro que

$$\lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} b_n = \lim_{n \rightarrow +\infty} c_n = 0$$

Además en general, el tipo de convergencia que presentan es lineal, al ser

$$\lim_{n \rightarrow +\infty} \frac{c_{n+1}}{c_n} = \lim_{n \rightarrow +\infty} \frac{\frac{1}{(n+1)^p}}{\frac{1}{n^p}} = \lim_{n \rightarrow +\infty} \left(\frac{n}{n+1} \right)^p = 1$$

Pero es claro que dentro de los diferentes modos de aproximarse a cero, la aproximación más rápida es la generada por c_n , de suerte que

$$\lim_{n \rightarrow +\infty} \frac{c_n}{a_n} = \lim_{n \rightarrow +\infty} \frac{c_n}{b_n} = 0, \quad \forall p > 2$$

Esto motiva la siguiente definición.

Def.: (Velocidad de convergencia) Sean x_n e y_n dos sucesiones de números reales, convergentes a un mismo valor real α . Se denota por

$$\epsilon_n = |x_n - \alpha|, \quad \hat{\epsilon}_n = |y_n - \alpha|, \quad \forall n \in \mathbb{N}$$

los correspondientes errores de aproximación. Se dice que y_n **converge a α con mayor rapidez que x_n** si:

$$\lim_{n \rightarrow +\infty} \frac{\hat{\epsilon}_n}{\epsilon_n} = 0$$

Si existe una constante $\lambda > 0$ tal que

$$\lim_{n \rightarrow +\infty} \frac{\hat{\epsilon}_n}{\epsilon_n} = \lambda \neq 0$$

se dice que las sucesiones poseen el mismo carácter de convergencia. En el caso particular que $\lambda = 1$ las sucesiones son **convergentemente equivalentes**.

El siguiente resultado establece algo esperado: a mayor orden mayor rapidez de convergencia.

Proposición Sean $\{x_n\}$ e $\{y_n\}$ dos sucesiones de números reales convergentes a un mismo valor $\alpha \in \mathbb{R}$. Se supone que (x_n) tiene orden de convergencia p e (y_n) orden de convergencia q , con $0 < p < q$. Si existe el límite

$$\lim_{n \rightarrow +\infty} \frac{|y_n - \alpha|}{|x_n - \alpha|} = \lambda < +\infty$$

entonces $\lambda = 0$, es decir, la sucesión (y_n) converge al límite más rápido que la sucesión (x_n) (donde son considerados solamente los valores de n para los que $|x_n - \alpha| \neq 0$).

Demostración: Se definen, como antes, los errores de aproximación

$$\epsilon_n = |x_n - \alpha|, \quad \hat{\epsilon}_n = |y_n - \alpha|, \quad \forall n \in \mathbb{N}$$

y la sucesión

$$z_n := \frac{\hat{\epsilon}_n}{\epsilon_n}, \quad \text{con } \epsilon_n \neq 0, \quad \forall n \in \mathbb{N}$$

En tal caso, se puede escribir:

$$z_{n+1} = \frac{\hat{\epsilon}_{n+1}}{\epsilon_{n+1}} = \frac{\hat{\epsilon}_{n+1}}{\hat{\epsilon}_n} \cdot \frac{\hat{\epsilon}_n^q}{\epsilon_n^q} \cdot \frac{\epsilon_n^q}{\epsilon_n^p} \cdot \frac{\epsilon_n^p}{\epsilon_{n+1}} = \frac{\hat{\epsilon}_{n+1}}{\hat{\epsilon}_n} \cdot \left(\frac{\hat{\epsilon}_n}{\epsilon_n}\right)^q \cdot \epsilon_n^{q-p} \cdot \frac{\epsilon_n^p}{\epsilon_{n+1}}$$

esto es,

$$z_{n+1} = \frac{\hat{\epsilon}_{n+1}}{\hat{\epsilon}_n^q} \cdot z_n^q \cdot \epsilon_n^r \cdot \frac{\epsilon_n^p}{\epsilon_{n+1}}, \quad r = q - p > 0$$

Si la sucesión z_n es convergente a un valor λ , en particular debe estar acotada por un cierto valor M ,

$$|z_n| \leq M, \quad \forall n \in \mathbb{N}$$

entonces es convergente con valor nulo, ya que se tiene la acotación

$$|z_{n+1}| \leq \frac{\hat{\epsilon}_{n+1}}{\hat{\epsilon}_n^q} \cdot M^q \cdot \epsilon_n^r \cdot \frac{\epsilon_n^p}{\epsilon_{n+1}}, \quad \forall n \in \mathbb{N}$$

y teniendo garantizada la existencia de los límites,

$$\lim_{n \rightarrow +\infty} \frac{\hat{\epsilon}_{n+1}}{\hat{\epsilon}_n} = \lambda_2, \quad \lim_{n \rightarrow +\infty} \frac{\epsilon_{n+1}}{\epsilon_n^p} = \lambda_1, \quad \lim_{n \rightarrow +\infty} \epsilon_n^r = 0$$

se deduce por la propiedad de monotonía, que

$$0 \leq \lim_{n \rightarrow +\infty} |z_{n+1}| \leq \lambda_2 \cdot M^q \cdot \frac{1}{\lambda_1} \cdot \lim_{n \rightarrow +\infty} \epsilon_n^r = 0$$

esto es,

$$\lambda = \lim_{n \rightarrow +\infty} z_n = 0$$

como se quería demostrar.

4.4.1. Ejemplo 1:

Las sucesiones

$$x_n = \frac{1}{n^p}, \quad y_n = e^{-n}, \quad p > 0$$

convergen linealmente hacia cero (ambas son *infinitésimos*). En cambio la velocidad de convergencia de la segunda es mucho mayor que la primera, al ser

$$\lim_{n \rightarrow +\infty} \frac{y_n}{x_n} = \lim_{n \rightarrow +\infty} \frac{e^{-n}}{\frac{1}{n^p}} = \lim_{n \rightarrow +\infty} \frac{n^p}{e^n} = 0, \quad \forall p > 0$$

4.4.2. Ejemplo 2:

Las sucesiones

$$x_n = 2^{-3^n}, \quad y_n = 3^{-2^n}, \quad \forall n \in \mathbb{N}$$

convergen a cero. Pero mientras que la convergencia en la primera es cúbica, la segunda es cuadrática. Por consiguiente,

$$\lim_{n \rightarrow +\infty} \frac{x_n}{y_n} = \lim_{n \rightarrow +\infty} \frac{2^{-3^n}}{3^{-2^n}} = \lim_{n \rightarrow +\infty} \frac{3^{2^n}}{2^{3^n}} = 0$$

la velocidad de convergencia de la primera es mucho mayor que la segunda. Esto también se puede comprobar observando que

$$\left(\frac{3}{2}\right)^n > \frac{\log(3)}{\log(2)}, \quad \forall n > 1$$

En consecuencia

$$\frac{3^n}{2^n} > \frac{\log(3)}{\log(2)} \rightarrow 3^n \cdot \log(2) > 2^n \cdot \log(3) \rightarrow \log(2^{3^n}) > \log(3^{2^n})$$

i.e.,

$$2^{3^n} > 3^{2^n} \Leftrightarrow 2^{-3^n} < 3^{-2^n}, \quad \forall n > 1$$

tal como se deduce del orden de convergencia.

4.4.3. Ejemplo 3:

Dado cualquier *infinitésimo* ϵ_n , entendido como sucesión de números reales que tiende a cero, siempre es posible encontrar otro infinitésimo que mejora la velocidad de convergencia. En efecto basta considerar

$$y_n = \epsilon_n^p, \quad \forall n \in \mathbb{N}, \quad p > 1$$

y se prueba de manera trivial que

$$\lim_{n \rightarrow +\infty} \frac{y_n}{\epsilon_n} = \lim_{n \rightarrow +\infty} \frac{\epsilon_n^p}{\epsilon_n} = \lim_{n \rightarrow +\infty} \epsilon_n^{p-1} = 0$$

5. Multiplicidad del cero de una función

Def.: Se dice que $\alpha \in \mathbb{R}$ es un **cero de multiplicidad** m ($m \in \{1, 2, 3, \dots\}$) de la función $f : [a, b] \rightarrow \mathbb{R}$ si la función se puede expresar en la forma

$$f(x) = (x - \alpha)^m \cdot g(x), \quad \forall x \neq \alpha$$

con $x \in I(\alpha)$ un entorno de α y

$$\lim_{x \rightarrow \alpha} g(x) \neq 0$$

O también, si existe un entero positivo m , tal que verifica:

$$\lim_{x \rightarrow \alpha} \frac{f(x)}{(x - \alpha)^k} = 0, \quad \forall k < m, \quad y \quad \lim_{x \rightarrow \alpha} \frac{f(x)}{(x - \alpha)^m} \neq 0$$

Cuando $m = 1$ se dice que α es un *cero simple* de f . Para funciones suficientemente derivables, los siguientes criterios, obtenidos por aplicación de los Teoremas de Taylor, permiten caracterizar los ceros múltiples de una función.

Criterio de cero simple Sea $f \in C^1(a, b)$ y $\alpha \in (a, b)$. Entonces f tiene un cero simple en α si y sólo si

$$f(\alpha) = 0 \quad y \quad f'(\alpha) \neq 0$$

Criterio de cero múltiple Sea $f \in C^n(a, b)$ y $\alpha \in (a, b)$. Entonces f tiene un cero de multiplicidad m en α si y sólo si

$$f(\alpha) = f'(\alpha) = f''(\alpha) = \dots = f^{m-1}(\alpha) = 0 \quad \text{y} \quad f^m(\alpha) \neq 0$$

Ejemplo:

Se considera la función $f(x) = e^x - x - 1$ definida en un entorno del cero I . Se calculan las dos primeras derivadas:

$$f'(x) = e^x - 1, \quad f''(x) = e^x$$

Ahora si se evalúan las funciones f, f', f'' en el punto $\alpha = 0$:

$$f(0) = 0, \quad f'(0) = 0, \quad f''(0) = 1 \neq 0$$

Por los criterios establecidos, $\alpha = 0$ es un cero de f de multiplicidad $m = 2$.

6. Teoremas sobre convergencia lineal y cuadrática

Teorema 1 (condición suficiente para convergencia lineal)

Sea $g \in C^1[a, b]$ tal que $g([a, b]) \subset [a, b]$ y $|g'(x)| \leq k$ para todo $x \in [a, b]$, donde $k \in (0, 1)$. Se denota por α un punto fijo de g en el intervalo $[a, b]$. Si $g'(\alpha) \neq 0$, entonces para cualquier $x_0 \in [a, b] - \{\alpha\}$, la sucesión (x_n) definida por la fórmula recursiva

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, 3, \dots$$

converge linealmente al punto fijo α .

Demostración: Por la definición de la sucesión x_n , se pueden evaluar los errores de aproximación sucesivos, expresados en la forma:

$$\epsilon_{n+1} = |x_{n+1} - \alpha| = |g(x_n) - g(\alpha)|$$

Aplicando el *Teorema del Valor Medio* a g en el intervalo de extremos x_n y α , existe un número real $c_n \in [x_n, \alpha]$, tal que:

$$|g(x_n) - g(\alpha)| = |g'(c_n)| \cdot |x_n - \alpha|$$

Esto es,

$$\frac{\epsilon_{n+1}}{\epsilon_n} = |g'(c_n)|$$

Además es claro que $\lim_{n \rightarrow +\infty} c_n = \alpha$, al verificarse la desigualdad

$$0 \leq |c_n - \alpha| \leq |x_n - \alpha|, \quad \forall n \in \mathbb{N}$$

Como g' es una función continua, se tendrá entonces que

$$\lim_{n \rightarrow +\infty} g'(c_n) = g'(\alpha)$$

Por tanto, en el paso al límite, se deduce que

$$\lim_{n \rightarrow +\infty} \frac{\epsilon_{n+1}}{\epsilon_n} = \lim_{n \rightarrow +\infty} |g'(c_n)| = |g'(\alpha)| \neq 0$$

Por consiguiente la convergencia (garantizada por el teorema del punto fijo) es lineal, como se quería demostrar.

Teorema 2 (condición suficiente para convergencia cuadrática)

Sea $g \in C^2[a, b]$ tal que $g([a, b]) \subset [a, b]$ y $|g''(x)| \leq M$ para todo $x \in [a, b]$. Se denota por α , un punto fijo de g en el intervalo $[a, b]$. Si $g'(\alpha) = 0$, entonces existe un $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta]$, la sucesión (x_n) definida por la fórmula recursiva

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, 3, \dots$$

converge al menos cuadráticamente al punto fijo α . Además, para valores suficientemente grandes de n ,

$$\epsilon_{n+1} < \frac{M}{2} \cdot \epsilon_n^2$$

Demostración: Al ser $g \in C^2[a, b]$, se puede aproximar g por su *polinomio de Taylor* de segundo orden, en torno al valor fijo α :

$$g(x) = g(\alpha) + g'(\alpha)(x - \alpha) + \frac{g''(\xi)}{2}(x - \alpha)^2, \quad \alpha < \xi < x$$

Como $g(\alpha) = \alpha$ y $g'(\alpha) = 0$, poniendo $x = x_n$, se obtiene:

$$x_{n+1} = g(x_n) = \alpha + \frac{g''(\xi_n)}{2}(x_n - \alpha)^2, \quad \alpha < \xi_n < x_n$$

donde ξ_n está comprendido entre α y x_n . Por consiguiente, el cociente de aproximaciones

$$\frac{\epsilon_{n+1}}{\epsilon_n^2} = \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^2} = \frac{|g''(\xi_n)|}{2}$$

Suponiendo que la segunda derivada de g es continua, es claro que

$$\lim_{n \rightarrow +\infty} \frac{|g''(\xi_n)|}{2} = \frac{|g''(\alpha)|}{2}$$

Por lo tanto,

$$\lim_{n \rightarrow +\infty} \frac{\epsilon_{n+1}}{\epsilon_n^2} = \lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^2} = \frac{|g''(\alpha)|}{2} < +\infty$$

lo que implica convergencia al menos cuadrática. La acotación dada como última tesis del teorema, es consecuencia de la acotación de la segunda derivada en cierto entorno del punto fijo.

Teorema 3 (condición suficiente para convergencia de orden p)

Sea $g \in C^p[a, b]$ tal que $g([a, b]) \subset [a, b]$. Se denota por α un punto fijo de g en el intervalo $[a, b]$. Si

$$g'(\alpha) = g''(\alpha) = \dots = g^{(p-1)}(\alpha) = 0, \quad g^{(p)}(\alpha) \neq 0$$

entonces existe un $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta]$, la sucesión (x_n) definida por la fórmula recursiva

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, 3, \dots$$

es convergente al punto fijo α , con orden de convergencia p . Además, si $|g^{(p)}(x)| < M$, para todo $x \in [a, b]$, para valores suficientemente grandes de n ,

$$\epsilon_{n+1} < \frac{M}{p!} \cdot \epsilon_n^p$$

Demostración: Al ser $g \in C^p[a, b]$, se puede aproximar g por su *polinomio de Taylor* de orden p en torno al valor fijo α :

$$g(x) = g(\alpha) + g'(\alpha)(x - \alpha) + \frac{g''(\alpha)}{2}(x - \alpha)^2 + \dots + \frac{g^{(p)}(\xi)}{p!}(x - \alpha)^p, \quad \alpha < \xi < x$$

Como $g(\alpha) = \alpha$ y $g'(\alpha) = g''(\alpha) = \dots = g^{(p-1)}(\alpha) = 0$, poniendo $x = x_n$, se obtiene:

$$x_{n+1} = g(x_n) = \alpha + \frac{g^{(p)}(\xi_n)}{p!} (x_n - \alpha)^p, \quad \alpha < \xi_n < x_n$$

Por tanto,

$$|x_{n+1} - \alpha| = \frac{|g^{(p)}(\xi_n)|}{p!} |x_n - \alpha|^p, \quad \alpha < \xi_n < x_n$$

esto es,

$$\frac{\epsilon_{n+1}}{\epsilon_n^p} = \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \frac{|g^{(p)}(\xi_n)|}{p!}$$

Suponiendo que la derivada p -ésima de g es continua, es claro que

$$\lim_{n \rightarrow +\infty} \frac{|g^{(p)}(\xi_n)|}{p!} = \frac{|g^{(p)}(\alpha)|}{p!}$$

Por lo tanto,

$$\lim_{n \rightarrow +\infty} \frac{\epsilon_{n+1}}{\epsilon_n^p} = \lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = \frac{|g^{(p)}(\alpha)|}{p!} \neq 0$$

lo que implica convergencia de orden p .

Ejemplo: (Aproximaciones al número de oro)

Se considera la sucesión de números reales, definida por:

$$x_0 = 1, \quad x_{n+1} = \sqrt{1 + x_n}, \quad n = 0, 1, 2, 3, \dots$$

La sucesión converge a la **constante de oro**,

$$\lim_{n \rightarrow +\infty} x_n = \Phi = \frac{1 + \sqrt{5}}{2} \approx 1,6180\dots$$

siendo además esta convergencia de tipo lineal. En efecto, basta considerar la función $g : [1, 2] \rightarrow \mathbb{R}$ definida por:

$$g(x) = +\sqrt{1 + x}$$

y observar que $g([1, 2]) \subset [1, 2]$. Además, $g(\Phi) = \Phi$. La primera derivada es

$$g'(x) = \frac{1}{2\sqrt{1 + x}} \neq 0, \quad \forall x \in [1, 2]$$

Por aplicación del *Teorema 1*, se concluye que la convergencia es lineal.

Ejemplo: (Aproximaciones de \sqrt{c})

Sea $c > 1$. Se considera la sucesión de números reales, definida por:

$$x_0 = 1, \quad x_{n+1} = \frac{1}{2} \cdot \left(x_n + \frac{c}{x_n} \right), \quad n = 0, 1, 2, 3, \dots$$

La sucesión converge cuadráticamente a \sqrt{c} .

Para demostrar esto, se considera la función

$$g(x) = \frac{1}{2} \cdot \left(x + \frac{c}{x} \right), \quad x \in I$$

siendo I un intervalo tal que $[1, c] \subset I$. Hay que observar que

$$1 < g(1) = g(c) = \frac{c+1}{2} < c, \quad c > 1$$

Además g tiene un único punto fijo dentro de este intervalo.

$$g(x) = x \leftrightarrow x = \frac{1}{2} \cdot \left(x + \frac{c}{x}\right) \leftrightarrow 2x = x + \frac{c}{x} \rightarrow x^2 = c \rightarrow x = \pm\sqrt{c}$$

Al considerar sólo valores positivos,

$$g(x) = x \leftrightarrow x = +\sqrt{c}$$

Derivando esta función una vez,

$$g'(x) = \frac{1}{2} \cdot \left(1 - \frac{c}{x^2}\right) = \frac{x^2 - c}{2x^2}$$

se comprueba que $g'(+\sqrt{c}) = 0$. En consecuencia, aplicando el *Teorema 2*, se puede afirmar que existe un entorno del punto fijo, donde la convergencia es al menos cuadrática.

Para ilustrar la velocidad de convergencia, se puede tomar para el último ejemplo, $c = 3$. Los valores numéricos obtenidos para las primeras iteraciones son:

$$\begin{aligned} x_0 = 1, \quad x_2 = 2, \quad x_3 = 1,732, \quad x_4 = 1,73205 \\ x_5 = 1,7320504, \quad x_6 = 1,732050806, \quad x_7 = 1,732050808\dots \end{aligned}$$

Puesto que $\sqrt{3} = 1,7320508075688772\dots$ (*Wolfram alpha*), ya en la séptima iteración se obtienen 8 cifras decimales exactas!

7. Análisis del método de Newton-Raphson

Teorema de convergencia local

Sea $f : [a, b] \rightarrow$ una función de clase C^2 en un entorno de α , tal que

$$f(\alpha) = 0, \quad f'(\alpha) \neq 0$$

Entonces existe un número real $\delta > 0$ tal que para todo $x_0 \in (\alpha - \delta, \alpha + \delta)$ la sucesión definida por la *transformada de Newton-Raphson*

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, 3\dots$$

converge a α . Además si f es de clase $C^3[a, b]$ la convergencia es al menos cuadrática.

Este resultado local permite diseñar algoritmos para calcular raíces de polinomios con bastante eficiencia. No obstante, queda englobado en un resultado más general, como el que se expone a continuación.

Teorema de convergencia global

Sea $f : [a, b] \rightarrow$ una función de clase $C^2[a, b]$ verificando:

1. $f(a) \cdot f(b) < 0$.

2. $f'(x) \neq 0, \quad \forall x \in [a, b]$
3. $f''(x) \cdot f''(y) \geq 0, \quad \forall x, y \in [a, b]$
4. $mx \left\{ \frac{|f(a)|}{|f'(a)|}, \frac{|f(b)|}{|f'(b)|} \right\} \leq b - a$

Entonces existe un único $\alpha \in [a, b]$ tal que $f(\alpha) = 0$ y para cada $x_0 \in [a, b]$ la sucesión del método de Newton-Raphson, definida por

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad n = 0, 1, 2, 3, \dots$$

converge a α . Además si f es de clase $C^3[a, b]$ la convergencia es al menos cuadrática.

Demostración: Por la primera condición, aplicando el *Teorema de Bolzano*, existe al menos una raíz α para f . De la segunda condición, aplicando el *Teorema de Rolle* se deduce que dicha raíz debe ser única (de lo contrario la derivada tendría una raíz).

De las condiciones 2. y 3. se obtiene que, tanto f' como f'' , conservan su signo en $[a, b]$. Esto es, f mantiene la monotonía (o bien es creciente o bien siempre decreciente) y la curvatura (cóncava o convexa, y carente de puntos de inflexión). Se supone que $f'(x) > 0$ y $f''(x) \geq 0$. El análisis del resto de casos sería similar.

Se define la función $g(x) = x - \frac{f(x)}{f'(x)}$ en $[a, b]$. Derivando

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

Por ser f creciente, con $f(\alpha) = 0$ es claro que

$$f(x) < 0, \quad \forall x \in [a, \alpha)$$

$$f(x) > 0 \quad \forall x \in (\alpha, b]$$

Por consiguiente, g es decreciente en $[a, \alpha)$ y creciente en $(\alpha, b]$. Entonces

$$\text{Si } x \in [a, \alpha) \rightarrow \alpha = g(\alpha) \leq g(x) \leq g(a)$$

Ahora bien, por la condición (4),

$$\frac{|f(a)|}{|f'(a)|} \leq b - a \rightarrow a - \frac{f(a)}{f'(a)} \leq b$$

$$g(a) = a - \frac{f(a)}{f'(a)} \leq b$$

Por tanto

$$x \in [a, \alpha) \rightarrow g(x) \in [\alpha, b]$$

Análogamente, como g es creciente en $(\alpha, b]$,

$$\text{Si } x \in (\alpha, b] \rightarrow \alpha = g(\alpha) \leq g(x) \leq g(b) < b$$

Por tanto

$$x \in (\alpha, b] \rightarrow g(x) \in [\alpha, b]$$

Es decir, en cualquiera de los dos casos, la sucesión generada por el método de Newton-Raphson verifica

$$x_n \in [\alpha, b], \quad \forall n \in \mathbb{N}$$

Además

$$x_{n+1} = g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)} < x_n$$

luego se trata de una sucesión estrictamente decreciente y acotada inferiormente. En consecuencia debe ser convergente. Si se denota por β a dicho límite es claro que

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

y por la unicidad del punto fijo, debe ser

$$\beta = \lim_{n \rightarrow +\infty} g(x_n) = \alpha$$

lo que prueba la tesis del teorema. Además, como $g'(\alpha) = 0$, por los resultados previos (sección 6, teorema 2), la convergencia es al menos cuadrática.

Aplicación: (Cálculo de la raíz positiva de un número)

Para hallar la raíz cuadrada positiva del número real c , se considera la función $f(x) = x^2 - c$. Sean a, b números reales, tales que

$$0 < a^2 < c < b^2$$

de suerte que $a^2 - c < 0$ y $b^2 - c > 0$.

En tal caso se verifican las cuatro condiciones del teorema de convergencia global anterior. En efecto,

1. $f(a) \cdot f(b) = (a^2 - c) \cdot (b^2 - c) < 0$.
2. $f'(x) = 2x > 0, \quad \forall x \in [a, b]$.
3. $f''(x) = 2 > 0, \quad \forall x \in [a, b]$.
4. $\frac{|f(b)|}{|f'(b)|} = \frac{b^2 - c}{2b} \leq \frac{b^2 - a^2}{2b} = \frac{(b+a)(b-a)}{2b} \leq b - a$

En consecuencia, el método de *Newton-Raphson* aplicado a $f(x) = x^2 - c$ en un intervalo $[a, b]$ en las condiciones dadas, converge a la raíz positiva de c .

Aplicación: (Cálculo de la raíz p -ésima de un número)

Para hallar la raíz p -ésima positiva del número real c ($p > 2$), se considera la función $f(x) = x^p - c$. Sean a, b números reales, tales que

$$0 < a^p < c < b^p$$

Siguiendo un razonamiento similar al caso previo, es fácil demostrar las hipótesis del teorema de convergencia global.

1. $f(a) \cdot f(b) = (a^p - c) \cdot (b^p - c) < 0$.
2. $f'(x) = px^{p-1} > 0, \quad \forall x \in [a, b]$.
3. $f''(x) = p(p-1)x^{p-2} > 0, \quad \forall x \in [a, b]$.
4. $\frac{|f(b)|}{|f'(b)|} = \frac{b^p - c}{p \cdot b^{p-1}} \leq \frac{b^p - a^p}{p \cdot b^{p-1}} \leq b - a$

Para probar la última condición, se parte de la desigualdad

$$1 \leq 1 + x + x^2 + \dots + x^{p-1} \leq p, \quad \forall x \in [0, 1]$$

o, de manera equivalente

$$1 \leq \frac{x^p - 1}{x - 1} \leq p, \quad \forall x \in [0, 1]$$

Al ser $0 < a < b$, se puede considerar $x = \frac{a}{b} < 1$, y se tendrá:

$$\frac{\left(\frac{a}{b}\right)^p - 1}{\left(\frac{a}{b}\right) - 1} \leq p \rightarrow \frac{\frac{a^p - b^p}{b^p}}{\frac{a - b}{b}} \leq p \rightarrow \frac{b^p - a^p}{(b - a) \cdot b^{p-1}} \leq p$$

Esto es,

$$\frac{b^p - a^p}{p \cdot b^{p-1}} \leq b - a$$

lo que equivale a afirmar

$$0 < \frac{|f(b)|}{|f'(b)|} = \frac{b^p - c}{p \cdot b^{p-1}} \leq \frac{b^p - a^p}{p \cdot b^{p-1}} \leq b - a$$

siempre que $0 < a < b$, con $a^p < c < b^p$. Por tanto, el algoritmo de iteración

$$x_{n+1} = \frac{1}{p} \cdot \left((p - 1) \cdot x_n + \frac{c}{x_n^{p-1}} \right), \quad n = 0, 1, 2, 3, \dots$$

con $x_0 \in [a, b]$, converge cuadráticamente al número $\sqrt[p]{c}$.

Cálculo de $\sqrt[5]{2}$.

Por ejemplo, para $p = 5$ y $c = 2$, se obtendría la función

$$g(x) = \frac{1}{5} \cdot \left(4x + \frac{2}{x^4} \right)$$

que se podría considerar definida en el intervalo $[1, 2]$, al ser $1^5 < 2 < 2^5$. Partiendo del valor inicial $x_0 = 1$, se obtendrían los valores numéricos

$$x_0 = 1, \quad x_1 = g(1) = 1,2, \quad x_2 = g(1,2) = 1,15290, \quad x_3 = g(1,15290) = 1,14872$$

$$x_4 = g(1,1487) = 1,148675, \quad x_5 = g(1,148675) = 1,14869835, \quad x_6 = 1,148698355$$

Ya en la séptima iteración se observa que las primeras ocho cifras significativas no varían,

$$x_7 = g(1,148698355) = 1,148698355\dots$$

aproximándose sucesivamente al valor real

$$\sqrt[5]{2} = 1,1486983549970350067986269467779275894438508890977975055137111184\dots$$

obtenido mediante el algoritmo de computación (*Wolfram Alpha*).

8. Aceleración de convergencia

La cuestión que se plantea ahora es la siguiente: *dada una sucesión de números reales (x_n) que converge con cierto orden al número real α , encontrar un algoritmo iterativo que mejore la velocidad de convergencia, i.e.*

$$\lim_{n \rightarrow +\infty} \frac{|y_n - \alpha|}{|x_n - \alpha|} = 0$$

para una cierta sucesión (y_n) con límite α .

Diferencias progresivas

Def.: Dada la sucesión $(x_n)_{n=0}^{+\infty}$, se define la *diferencia progresiva de primer orden* Δx , por:

$$(\Delta x)_n := x_{n+1} - x_n, \quad n = 0, 1, 2, 3, \dots$$

La *diferencia progresiva de segundo orden*, $\Delta^2 x$ es:

$$(\Delta^2 x)_n = (\Delta x)_{n+1} - (\Delta x)_n = x_{n+2} - 2x_{n+1} + x_n, \quad n = 0, 1, 2, 3, \dots$$

En general, la *diferencia progresiva de orden k* , $\Delta^k x$ se define por la expresión recurrente:

$$\boxed{(\Delta^k x)_n = (\Delta^{k-1} x)_{n+1} - (\Delta^{k-1} x)_n, \quad n = 0, 1, 2, 3, \dots}$$

Así, la diferencia progresiva de orden 3, se obtiene del siguiente modo:

$$\begin{aligned} (\Delta^3 x)_n &= (\Delta^2 x)_{n+1} - (\Delta^2 x)_n = (x_{n+3} - 2x_{n+2} + x_{n+1}) - (x_{n+2} - 2x_{n+1} + x_n) = \\ &= x_{n+3} - 3x_{n+2} + 3x_{n+1} - x_n = \binom{3}{0}x_{n+3} - \binom{3}{1}x_{n+2} + \binom{3}{2}x_{n+1} - \binom{3}{3}x_n \end{aligned}$$

No es complicado observar que, en general, estas diferencias obedecen a expresiones de la forma:

$$(\Delta^k x)_n = \sum_{i=0}^k (-1)^i \binom{k}{i} \cdot x_{n+k-i}, \quad n = 0, 1, 2, 3, \dots$$

Por ejemplo, si se considera la sucesión $x_n = \frac{1}{n}$, la diferencia de primer orden es una nueva sucesión, de la forma:

$$(\Delta x)_n = x_{n+1} - x_n = \frac{1}{n+1} - \frac{1}{n} = \frac{-1}{n(n+1)} < 0$$

Esta nueva sucesión converge con mayor rapidez a cero. Las diferencias de segundo orden serían:

$$\begin{aligned} (\Delta^2 x)_n &= (\Delta x)_{n+1} - (\Delta x)_n = x_{n+2} - 2x_{n+1} + x_n = \\ &= \frac{1}{n+2} - \frac{2}{n+1} + \frac{1}{n} = \frac{n(n+1) - 2n(n+2) + (n+1)(n+2)}{n(n+1)(n+2)} = \\ &= \frac{n^2 + n - 2n^2 - 4n + n^2 + 3n + 2}{n(n+1)(n+2)} = \frac{2}{n(n+1)(n+2)} \end{aligned}$$

que aumenta la convergencia de la diferencia progresiva de primer orden. Esto es una propiedad general.

Deducción de la fórmula de Aitken

Sea $(x_n)_{n=0}^{+\infty}$ una sucesión de números reales que converge linealmente a α . Se supone que las diferencias

$$e_n = x_n - \alpha$$

tienen el mismo signo (se puede suponer que son positivas). Entonces para valores suficientemente grandes de n ,

$$\frac{x_{n+2} - \alpha}{x_{n+1} - \alpha} \approx \frac{x_{n+1} - \alpha}{x_n - \alpha}$$

ya que

$$\lim_{n \rightarrow +\infty} \frac{e_{n+1}}{e_n} = \lim_{n \rightarrow +\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = \lambda \neq 0$$

al suponer convergencia lineal. En este caso, multiplicando en cruz,

$$\begin{aligned} x_{n+1}^2 - 2\alpha x_{n+1} + \alpha^2 &\approx x_{n+2}x_n - \alpha x_{n+2} - \alpha x_n + \alpha^2 \\ \alpha \cdot (x_{n+2} - 2x_{n+1} + x_n) &\approx x_{n+2}x_n - x_{n+1}^2 \\ \alpha \cdot (x_{n+2} - 2x_{n+1} + x_n) &\approx (x_{n+2}x_n - 2x_n x_{n+1} + x_n^2) - (x_{n+1}^2 - 2x_n x_{n+1} + x_n^2) \\ \alpha \cdot (x_{n+2} - 2x_{n+1} + x_n) &\approx x_n \cdot (x_{n+2} - 2x_{n+1} + x_n) - (x_{n+1} - x_n)^2 \end{aligned}$$

esto es,

$$\alpha \cdot (\Delta^2 x)_n \approx x_n \cdot (\Delta^2 x)_n - ((\Delta x)_n)^2$$

y por tanto

$$\alpha \approx x_n - \frac{((\Delta x)_n)^2}{(\Delta^2 x)_n}$$

Esto motiva definir la **transformada de Aitken** de la sucesión x_n ,

$$y_n = x_n - \frac{((\Delta x)_n)^2}{(\Delta^2 x)_n}, \quad n = 0, 1, 2, 3...$$

Por lo anterior, es claro que la *transformada de Aitken* de una sucesión convergente, es una nueva sucesión que converge al mismo límite. Lo interesante es que esta transformada acelera la velocidad de convergencia, tal y como se prueba en el siguiente teorema.

Teorema sobre aceleración de convergencia (Aitken, 1926)

Sea $(x_n)_{n=0}^{+\infty}$ una sucesión de números reales que converge linealmente al límite α , con constante de convergencia $0 < \lambda < 1$. Se supone que para todos los valores de n suficientemente grandes, se tiene

$$(x_n - \alpha) \cdot (x_{n+1} - \alpha) > 0$$

Entonces la sucesión (y_n) , definida por

$$y_n = x_n - \frac{((\Delta x)_n)^2}{(\Delta^2 x)_n}, \quad n = 0, 1, 2, 3...$$

converge a α con mayor rapidez que x_n , en el sentido de que

$$\lim_{n \rightarrow +\infty} \frac{|y_n - \alpha|}{|x_n - \alpha|} = 0$$

Demostración: Se denota

$$k_n = \frac{\epsilon_{n+1}}{\epsilon_n}, \quad \epsilon_n = |x_n - \alpha| \neq 0$$

Por hipótesis, se sabe que

$$\lim_{n \rightarrow +\infty} k_n = \lambda \neq 0, \quad |\lambda| < 1$$

Sea y_n la sucesión definida por el *algoritmo de Aitken*. Las aproximaciones sucesivas, en este caso, son:

$$\begin{aligned} \hat{\epsilon}_n &= y_n - \alpha = x_n - \alpha - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n} = \\ &= (x_n - \alpha) - \frac{((x_{n+1} - \alpha) - (x_n - \alpha))^2}{(x_{n+2} - \alpha) - 2(x_{n+1} - \alpha) + (x_n - \alpha)} = \\ &= \epsilon_n - \frac{(\epsilon_{n+1} - \epsilon_n)^2}{\epsilon_{n+2} - 2\epsilon_{n+1} + \epsilon_n} = \epsilon_n - \frac{(k_n \epsilon_n - \epsilon_n)^2}{k_{n+1} k_n \epsilon_n - 2\epsilon_n k_n + \epsilon_n} = \\ &= \epsilon_n - \frac{\epsilon_n^2 \cdot (k_n - 1)^2}{\epsilon_n \cdot (k_{n+1} k_n - 2k_n + 1)} = \epsilon_n \cdot \frac{k_n(k_{n+1} - k_n)}{k_{n+1} k_n - 2k_n + 1} \end{aligned}$$

Suponiendo que $\lambda \neq 1$, se pueden tomar límites en la expresión previa y obtener:

$$\lim_{n \rightarrow +\infty} \frac{\hat{\epsilon}_n}{\epsilon_n} = \lim_{n \rightarrow +\infty} \frac{k_n(k_{n+1} - k_n)}{k_{n+1} k_n - 2k_n + 1} = 0$$

Observación: Nótese que en la demostración anterior, es relevante la hipótesis sobre la constante asintótica

$$0 < \lambda < 1$$

En efecto, se va a demostrar con un par de ejemplos que, si $\lambda = 1$, el resultado no es válido, ni se mejora la velocidad de convergencia del algoritmo.

Ejemplo: El caso $x_n = \frac{1}{n}$. Para la sucesión $x_n = \frac{1}{n}$, que es estrictamente decreciente y convergente a $\alpha = 0$, se dispone de las diferencias progresivas,

$$\begin{aligned} (\Delta x)_n &= \frac{1}{n+1} - \frac{1}{n} = -\frac{1}{n(n+1)} \\ (\Delta^2 x)_n &= \frac{1}{n+2} - \frac{2}{n+1} + \frac{1}{n} = \frac{2}{n(n+1)(n+2)} \end{aligned}$$

La sucesión generada por el *algoritmo de Aitken*, es:

$$\begin{aligned} y_n &= x_n - \frac{((\Delta x)_n)^2}{(\Delta^2 x)_n} = \frac{1}{n} - \frac{\left(-\frac{1}{n(n+1)}\right)^2}{\frac{2}{n(n+1)(n+2)}} = \frac{1}{n} - \frac{\frac{1}{n^2(n+1)^2}}{\frac{2}{n(n+1)(n+2)}} = \\ &= \frac{1}{n} - \frac{n(n+1)(n+2)}{2n^2(n+1)^2} = \frac{1}{n} - \frac{n+2}{2n(n+1)} = \frac{2n+2-n-2}{2n(n+1)} = \frac{n}{2n(n+1)} = \frac{1}{2(n+1)} \end{aligned}$$

Es decir, la sucesión generada para acelerar la convergencia es

$$y_n = \frac{1}{2(n+1)}$$

que si bien es convergente a cero, verifica

$$\lim_{n \rightarrow +\infty} \frac{y_n}{x_n} = \frac{1}{2} \neq 0$$

esto es, poseen la misma velocidad de convergencia y no se cumple la tesis establecida por el teorema.

Ejemplo: El caso $x_n = \frac{1}{n^2}$.

El caso anterior es paradigmático de lo que sucede cuando $\lambda = 1$. Para la sucesión $x_n = \frac{1}{n^2}$, que converge a cero con mayor velocidad que $\frac{1}{n}$, se puede comprobar que

$$\lim_{n \rightarrow +\infty} k_n = \lim_{n \rightarrow +\infty} \frac{x_{n+1}}{x_n} = 1$$

El algoritmo ideado no mejora la velocidad de convergencia.

En efecto, cálculos relativamente sencillos, permiten obtener en este caso, las expresiones:

$$\begin{aligned} (\Delta x)_n &= \frac{1}{(n+1)^2} - \frac{1}{n^2} = -\frac{2n+1}{n^2(n+1)^2} \\ (\Delta^2 x)_n &= \frac{1}{(n+2)^2} - \frac{2}{(n+1)^2} + \frac{1}{n^2} = \frac{6n^2+12n+4}{n^2(n+1)^2(n+2)^2} \end{aligned}$$

Por consiguiente,

$$\begin{aligned} y_n &= x_n - \frac{((\Delta x)_n)^2}{(\Delta^2 x)_n} = \frac{1}{n^2} - \frac{\left(-\frac{2n+1}{n^2(n+1)^2}\right)^2}{\frac{6n^2+12n+4}{n^2(n+1)^2(n+2)^2}} = \frac{1}{n^2} - \frac{(2n+1)^2 n^2 (n+1)^2 (n+2)^2}{n^4 (n+1)^4 (6n^2+12n+4)} = \\ &= \frac{1}{n^2} - \frac{(2n+1)^2 (n+2)^2}{n^2 (n+1)^2 (6n^2+12n+4)} = \frac{(n+1)^2 (6n^2+12n+4) - (2n+1)^2 (n+2)^2}{n^2 (n+1)^2 (6n^2+12n+4)} = \\ &= \frac{(n^2+2n+1)(6n^2+12n+4) - (4n^2+4n+1)(n^2+4n+4)}{n^2 (n+1)^2 (6n^2+12n+4)} = \\ &= \frac{6n^4+12n^3+4n^2+12n^3+24n^2+8n+6n^2+12n+4}{n^2 (n+1)^2 (6n^2+12n+4)} = \\ &= \frac{4n^4+16n^3+16n^2+4n^3+16n^2+16n+n^2+4n+4}{n^2 (n+1)^2 (6n^2+12n+4)} = \\ &= \frac{2n^4+4n^3+n^2}{n^2 (n+1)^2 (6n^2+12n+4)} = \frac{2n^2+4n+1}{(n+1)^2 (6n^2+12n+4)} \end{aligned}$$

Nótese que para n suficientemente grande, se puede tomar en buena aproximación,

$$\begin{aligned} y_n &\approx \frac{2n^2+4n+1}{(n+1)^2 (6n^2+12n+4)} = \frac{2(n^2+2n+1)}{(n+1)^2 \cdot 2(3n^2+6n+2)} = \\ &= \frac{2(n+1)^2}{2(n+1)^2 (3n^2+6n+2)} = \frac{1}{3n^2+6n+2} \end{aligned}$$

Aunque bien es cierto que

$$y_n = \frac{1}{3n^2+6n+2} < \frac{1}{n^2} = x_n, \quad \forall n \in \mathbb{N}$$

no es cierto que aumente la velocidad de convergencia. La sucesión y_n posee el mismo carácter de convergencia lineal que x_n , al verificar

$$\lim_{n \rightarrow +\infty} \frac{y_n}{x_n} = \lim_{n \rightarrow +\infty} \frac{n^2}{3n^2 + 6n + 2} = \frac{1}{3} \neq 0$$

Los ejemplos anteriores muestran las debilidades del algoritmo de Aitken. Lo cual insta a buscar nuevos procesos de mejora, que suplan los casos más restrictivos.

9. Método de Steffensen

El *método de Steffensen* se puede considerar como una combinación del método del punto fijo y del método de Aitken. Para construir la sucesión de las aproximaciones x_n , en todo tercer paso se usa la fórmula de Aitken, y en el resto se va aplicando la iteración

$$\begin{aligned} x_n &= g(x_{n-1}), & n \neq 3 \\ x_{3n+3} &= x_{3n} - \frac{(\Delta x)_{3n}^2}{(\Delta^2 x)_{3n}}, & n = 0, 1, 2, 3, \dots \end{aligned}$$

Esto aumenta considerablemente la velocidad de convergencia. Concretamente se probará el siguiente resultado.

Teorema sobre convergencia de Steffensen

Sea $g : [a, b] \rightarrow \mathbb{R}$ una función de clase $C^3[a, b]$, con un punto fijo α , tal que

$$g(\alpha) = \alpha, \quad g'(\alpha) \neq 1$$

Entonces existe un $\delta > 0$, tal que el *método de Steffensen* converge cuadráticamente para cualquier $x_0 \in (\alpha - \delta, \alpha + \delta)$.

Aplicación

Se considera la función $g : [1, 2] \rightarrow \mathbb{R}$ definida por

$$g(x) = \sqrt[3]{1+x}$$

Al ser una función continua con $g([1, 2]) \subset [1, 2]$, existe al menos un punto fijo en el intervalo. Además, la derivada

$$g'(x) = \frac{1}{3\sqrt[3]{(1+x)^2}} > 0, \quad \forall x > 0$$

por lo que en el intervalo considerado se trata de una función estrictamente creciente. Esto garantiza la unicidad del punto fijo. Por otra parte es claro que

$$g'(x) < 1, \quad \forall x > 0$$

En consecuencia, se puede aplicar el algoritmo de iteración

$$x_{n+1} = g(x_n), \quad x_0 = 1$$

que tenderá al único punto fijo $\alpha \in [1, 2]$. Nótese que, en este caso,

$$\alpha = \sqrt[3]{1+\alpha} \Leftrightarrow \alpha^3 = 1 + \alpha$$

i.e., se está obteniendo indirectamente la única raíz real positiva del polinomio

$$p(x) = x^3 - x - 1$$

en el intervalo $[1, 2]$.

Partiendo del valor inicial $x_0 = 1$, se obtienen los primeros valores de la sucesión x_n :

$$\begin{aligned} x_0 &= 1, & x_1 &= g(1) = \sqrt[3]{2} = 1,2599\dots, & x_2 &= g(1,2599) = 1,312293\dots \\ x_3 &= g(1,3123) = 1,322353, & x_4 &= g(1,3223) = 1,324268, & x_5 &= g(1,32468) = 1,324632 \\ x_6 &= g(1,324632) = 1,32470, & x_7 &= g(1,32470) = 1,3247148, \\ x_8 &= g(1,3247148) = 1,324717372, & x_9 &= g(1,324717372) = 1,324717846 \\ x_{10} &= g(1,324717846) = 1,324717936\dots \end{aligned}$$

En la décima iteración se obtienen 7 cifras decimales exactas, aproximándose cada vez más al valor

$$\alpha = 1,32471795724474609085\dots$$

obtenido con *Wolfram Alpha*.

Si se aplica el *algoritmo de Steffensen*, se deben fijar los tres primeros valores por iteración:

$$x_0 = 1, \quad x_1 = g(1) = 1,2599, \quad x_2 = g(1,2599) = 1,3123$$

Para el cálculo de x_3 , se aplica el *algoritmo de Aitken*,

$$x_3 = x_0 - \frac{(\Delta x)_0^2}{(\Delta^2 x)_0} = x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0} = 1,32553$$

Continuando con el proceso, se itera el valor obtenido:

$$\begin{aligned} x_4 &= g(x_3) = g(1,32553) = 1,324872 \\ x_5 &= g(x_4) = g(1,324872) = 1,324747 \\ x_6 &= x_3 - \frac{(\Delta x)_3^2}{(\Delta^2 x)_3} = x_3 - \frac{(x_4 - x_3)^2}{x_5 - 2x_4 + x_3} = 1,324717685\dots \end{aligned}$$

y se sigue un nuevo ciclo,

$$\begin{aligned} x_7 &= g(x_6) = g(1,324717685) = 1,324717906 \\ x_8 &= g(x_7) = g(1,324717906) = 1,324717947 \\ x_9 &= x_6 - \frac{(\Delta x)_6^2}{(\Delta^2 x)_6} = x_6 - \frac{(x_7 - x_6)^2}{x_8 - 2x_7 + x_6} = 1,324717956\dots \end{aligned}$$

obteniendo 8 cifras decimales exactas en la novena iteración. No obstante, este método puede ser mejorado, utilizando un mayor número de veces el algoritmo de Aitken en el proceso.

Si se considera un proceso de iteración de punto fijo, $x_{n+1} = f(x_n)$, que se supone convergente a un cierto valor α , el método de Aitken se vuelve poco útil, ya que se conoce la sucesión recurrente, pero no se dispone, en principio, de la expresión general que permite determinar el término n-ésimo de la sucesión

$$x_n = F(n)$$

Es por esto, que se puede realizar una pequeña variación de la expresión discreta del algoritmo y suponer la siguiente transformada de f ,

$$g(x) = x - \frac{(f(x) - x)^2}{f(f(x)) - 2f(x) + x}, \quad x \in I$$

conocida como *transformada de Steffensen*. De este modo, se dispone de una nueva sucesión de iteraciones,

$$y_{n+1} = g(y_n) = y_n - \frac{(f(y_n) - y_n)^2}{f(f(y_n)) - 2f(y_n) + y_n}, \quad n = 0, 1, 2, 3, \dots$$

con $y_0 \in I$. Lo interesante de esta sucesión es que mejora la velocidad de convergencia, pasando a ser al menos cuadrática, tal como se prueba en el siguiente teorema.

Teorema local de convergencia de Steffensen

Sea $f : [a, b] \rightarrow \mathbb{R}$ una función de clase C^2 en (a, b) , con un punto fijo $\alpha \in (a, b)$, tal que

$$f(\alpha) = \alpha, \quad f'(\alpha) \neq 1$$

Entonces, existe un $\delta > 0$ tal que $(\alpha - \delta, \alpha + \delta) \subset (a, b)$ y para cualquier $y_0 \in (\alpha - \delta, \alpha + \delta)$, el algoritmo de Steffensen,

$$y_{n+1} = y_n - \frac{(f(y_n) - y_n)^2}{f(f(y_n)) - 2f(y_n) + y_n}, \quad n = 0, 1, 2, 3, \dots$$

está bien definido y converge hacia α con orden de convergencia al menos $p = 2$.

Demostración: La idea fundamental consiste en expresar la diferencia $|y_{n+1} - \alpha|$ en términos de $|y_n - \alpha|$, utilizando los desarrollos de Taylor de f en un entorno del punto α .

Nótese que, en tal caso, si g es la *transformada de Steffensen* de f , se tiene:

$$\lim_{n \rightarrow +\infty} \frac{|y_{n+1} - \alpha|}{|y_n - \alpha|^2} = \lim_{x \rightarrow \alpha} \frac{g(x) - x}{(x - \alpha)^2}$$

Se probará que dicho límite tiene por valor una expresión de la forma,

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

lo que garantiza una convergencia, al menos cuadrática.

En lo sucesivo, se dirá que $\epsilon_p(x) \approx o(x - \alpha)^p$, es una *función infinitesimal* en torno al número real α , si verifica:

$$\lim_{x \rightarrow \alpha} \frac{\epsilon_p(x)}{(x - \alpha)^p} = 0$$

Nótese además que se verifican las propiedades:

$$[P1] : \quad o((x - \alpha)^p) + o((x - \alpha)^q) = o((x - \alpha)^{\min(p,q)})$$

$$[P2] : \quad o((x - \alpha)^p) \cdot o((x - \alpha)^q) = o((x - \alpha)^{p+q})$$

para todo $p, q \in \mathbb{N}$.

El polinomio de Taylor de segundo orden de la función f en torno al punto fijo α , es:

$$f(x) = f(\alpha) + \frac{f'(\alpha)}{1!}(x - \alpha) + \frac{f''(\alpha)}{2!}(x - \alpha)^2 + o((x - \alpha)^2)$$

Puesto que $f(\alpha) = \alpha$,

$$f(x) - \alpha = \frac{f'(\alpha)}{1!}(x - \alpha) + \frac{f''(\alpha)}{2!}(x - \alpha)^2 + o((x - \alpha)^2)$$

para x en un cierto entorno de α . Elevando al cuadrado la expresión, se obtendrá:

$$(f(x) - \alpha)^2 = f'(\alpha)^2(x - \alpha)^2 + o((x - \alpha)^2), \quad \forall x \in (\alpha - \delta, \alpha + \delta)$$

Por otro lado,

$$(\Delta f)(x) = f(x) - x = (f'(\alpha) - 1)(x - \alpha) + \frac{f''(\alpha)}{2!}(x - \alpha)^2 + o((x - \alpha)^2)$$

$$(\Delta f)^2(x) = (f(x) - x)^2 = (f'(\alpha) - 1)^2(x - \alpha)^2 + (f'(\alpha) - 1) \cdot f''(\alpha)(x - \alpha)^3 + o((x - \alpha)^3)$$

que proporciona la expresión para el numerador de la función g . Para el denominador, se realizan los siguientes cálculos:

$$f(f(x)) - f(x) = (f'(\alpha) - 1)(f(x) - \alpha) + \frac{f''(\alpha)}{2!}(f(x) - \alpha)^2 + o((x - \alpha)^2)$$

Restando las expresiones obtenidas,

$$\begin{aligned} (\Delta^2 f)(x) &= f(f(x)) - 2f(x) + x = (f(f(x)) - f(x)) - (f(x) - x) = \\ &= (f'(\alpha) - 1)(f(x) - x) + \frac{f''(\alpha)}{2!} \cdot [(f(x) - \alpha)^2 - (x - \alpha)^2] + o((x - \alpha)^2) = \\ &= (f'(\alpha) - 1)^2(x - \alpha) + (f'(\alpha) - 1) \frac{f''(\alpha)}{2!}(x - \alpha)^2 + \frac{f''(\alpha)}{2!}(f'(\alpha)^2 - 1)(x - \alpha)^2 + o((x - \alpha)^2) = \\ &= (f'(\alpha) - 1)^2(x - \alpha) + \frac{f''(\alpha)}{2!}(f'(\alpha) - 1)(1 + f'(\alpha) + 1)(x - \alpha)^2 + o((x - \alpha)^2) = \end{aligned}$$

esto es,

$$(\Delta^2 f)(x) = (f'(\alpha) - 1)^2(x - \alpha) + \frac{f''(\alpha)}{2!}(f'(\alpha) - 1)(f'(\alpha) + 2)(x - \alpha)^2 + o((x - \alpha)^2)$$

Considerando la función que define el algoritmo de Steffensen,

$$g(x) = x - \frac{(f(x) - x)^2}{f(f(x)) - 2f(x) + x} = x - \frac{(\Delta f)^2(x)}{(\Delta^2 f)(x)}$$

se tendrá,

$$\begin{aligned} g(x) - \alpha &= (x - \alpha) - \frac{(\Delta f)^2(x)}{(\Delta^2 f)(x)} = \frac{(\Delta^2 f)(x)(x - \alpha) - (\Delta f)^2(x)}{(\Delta^2 f)(x)} = \\ &= \frac{\frac{f''(\alpha)}{2!}(f'(\alpha) - 1)f'(\alpha)(x - \alpha)^3 + o((x - \alpha)^3)}{(f'(\alpha) - 1)^2(x - \alpha) + \frac{f''(\alpha)}{2!}(f'(\alpha) - 1)(f'(\alpha) + 2)(x - \alpha)^2 + o((x - \alpha)^2)} = \\ &= \frac{\frac{f''(\alpha)}{2!}(f'(\alpha) - 1)f'(\alpha)(x - \alpha)^3 + o((x - \alpha)^3)}{(f'(\alpha) - 1)^2(x - \alpha) + \frac{f''(\alpha)}{2!}(f'(\alpha) - 1)(f'(\alpha) + 2)(x - \alpha)^2 + o((x - \alpha)^2)} = \end{aligned}$$

$$= \frac{\frac{f''(\alpha)}{2!} f'(\alpha)(x - \alpha)^2 + o((x - \alpha)^2)}{(f'(\alpha) - 1) + \frac{f''(\alpha)}{2!} (f'(\alpha) + 2)(x - \alpha) + o((x - \alpha))} =$$

En consecuencia,

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

Tomando límites, cuando $x \rightarrow \alpha$,

$$\lim_{x \rightarrow \alpha} \frac{g(x) - \alpha}{(x - \alpha)^2} = \lim_{x \rightarrow \alpha} \frac{\frac{f''(\alpha)}{2!} f'(\alpha) + \frac{o((x - \alpha)^2)}{(x - \alpha)^2}}{(f'(\alpha) - 1) + \frac{f''(\alpha)}{2!} (f'(\alpha) + 2)(x - \alpha) + o((x - \alpha))} = \frac{f'(\alpha) f''(\alpha)}{2(f'(\alpha) - 1)}$$

En definitiva, si $f'(\alpha) \neq 1$, dicho límite existe con valor finito, como se quería demostrar.

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

Por tanto, ya sea un punto fijo *atractor* ($f'(\alpha) < 1$), como repulsor ($f'(\alpha) > 1$), el teorema sigue siendo válido.

Observación 1: Nótese que el método de Steffensen permite trabajar directamente con la función, con lo que mejora el algoritmo de Aitken, en el que hay que determinar los primeros valores por iteración antes de aplicar el algoritmo acelerado.

Observación 2: Si la sucesión determinada por iteración converge a un punto fijo linealmente, la aceleración de Steffensen lo hace al menos cuadráticamente. Si las iteraciones de punto fijo convergen con orden de convergencia $p > 1$, las de Steffensen convergen con convergencia de orden al menos $2p - 1$.

Observación 3: Si además se cumple que el punto fijo es estacionario, el algoritmo de Steffensen tienen una convergencia de orden $p \geq 3$.

Métodos combinados

Dado un sistema dinámico de punto fijo α , (I, f) , siempre es posible encontrar otro sistema dinámico (I, g) con el mismo punto fijo, y que mejora la velocidad de convergencia. Por tanto, se podría acelerar un algoritmo diseñando una función compuesta con otra de mayor orden de convergencia. Esto presenta ciertas limitaciones como muestra la siguiente proposición.

Proposición

Sean $f, g : [a, b] \rightarrow \mathbb{R}$ dos funciones con un mismo punto fijo α . Se supone que el orden de convergencia de f es p y el orden de convergencia g es q , esto es,

$$\lim_{x \rightarrow \alpha} \frac{f(x) - \alpha}{(x - \alpha)^p} = \lambda_1 \neq 0, \quad \lim_{x \rightarrow \alpha} \frac{g(x) - \alpha}{(x - \alpha)^q} = \lambda_2 \neq 0, \quad p, q \in \mathbb{N}$$

Entonces el orden de convergencia de la función compuesta $f \circ g$ es q .

Demostración: Hay que observar en primer lugar que, efectivamente, la composición tiene el mismo punto fijo.

$$(f \circ g)(\alpha) = f(g(\alpha)) = f(\alpha) = \alpha$$

Sea $n \in \mathbb{N}$ el orden de convergencia de la composición considerada. En tal caso,

$$\begin{aligned} \frac{f(g(x)) - x}{(x - \alpha)^n} &= \frac{f(g(x)) - g(x)}{(x - \alpha)^n} + \frac{g(x) - x}{(x - \alpha)^n} = \\ &= \frac{f(g(x)) - g(x)}{(g(x) - \alpha)^p} \cdot \frac{(g(x) - \alpha)^p}{(x - \alpha)^n} + \frac{g(x) - x}{(x - \alpha)^q} \cdot \frac{(x - \alpha)^q}{(x - \alpha)^n} \\ &= \frac{f(g(x)) - g(x)}{(g(x) - \alpha)^p} \cdot \frac{(g(x) - \alpha)^p}{(x - \alpha)^{pq}} \cdot \frac{(x - \alpha)^{pq}}{(x - \alpha)^n} + \frac{g(x) - x}{(x - \alpha)^q} \cdot \frac{(x - \alpha)^q}{(x - \alpha)^n} \end{aligned}$$

Se denota por [EQ] la última expresión obtenida. Se va a analizar lo que sucede con cada uno de los sumandos al pasar al límite:

$$\begin{aligned} \lim_{x \rightarrow \alpha} \frac{f(g(x)) - g(x)}{(g(x) - \alpha)^p} &= \lim_{y=g(x) \rightarrow \alpha} \frac{f(y) - y}{(y - \alpha)^p} = \lambda_1 \neq 0 \\ \lim_{x \rightarrow \alpha} \frac{(g(x) - \alpha)^p}{(x - \alpha)^{pq}} &= \lim_{x \rightarrow \alpha} \left(\frac{g(x) - \alpha}{(x - \alpha)^q} \right)^p = \lambda_2^p \neq 0 \end{aligned}$$

Si se supone que $q > n$, entonces $pq > n$, y en consecuencia

$$\lim_{x \rightarrow \alpha} \frac{(x - \alpha)^{pq}}{(x - \alpha)^n} = \lim_{x \rightarrow \alpha} (x - \alpha)^{pq-n} = 0$$

por lo que el primer sumando se anula. De igual modo, los límites que figuran en segundo término deben valer:

$$\begin{aligned} \lim_{x \rightarrow \alpha} \frac{g(x) - x}{(x - \alpha)^q} &= \lambda_2 \neq 0 \\ \lim_{x \rightarrow \alpha} \frac{(x - \alpha)^q}{(x - \alpha)^n} &= \lim_{x \rightarrow \alpha} (x - \alpha)^{q-n} \end{aligned}$$

Para $n < q$ este último límite también es nulo. En cambio, si $n = q$, se obtiene para la expresión general [EQ], un valor λ no nulo. En consecuencia,

$$\lim_{x \rightarrow \alpha} \frac{(f \circ g)(x) - x}{(x - \alpha)^n} = \lim_{x \rightarrow \alpha} \frac{f(g(x)) - x}{(x - \alpha)^n} = \begin{cases} 0 & \text{si } n < q \\ \lambda & \text{si } n = q \\ \infty & \text{si } n > q \end{cases}$$

siendo por tanto el orden de convergencia de la composición igual a q .

La proposición muestra que el orden en la composición condiciona el orden de convergencia. Se denota por $Ord(f)$ el orden de convergencia de f , de suerte que, en la nueva notación

$$\boxed{Ord(f) = p, \quad Ord(g) = q \quad \Rightarrow \quad Ord(f \circ g) = Ord(g) = q}$$

Aplicación numérica

Las ideas anteriormente expuestas pueden servir para acelerar considerablemente los algoritmos diseñados con el fin de determinar aproximaciones numéricas de ceros de una función.

La idea resulta bastante sencilla: partiendo de la ecuación funcional

$$f(x) = 0$$

se obtiene la transformada de Newton F ,

$$F(x) = N_f(x) = x - \frac{f(x)}{f'(x)}, \quad \forall x \in I$$

de suerte que

$$f(\alpha) = 0 \leftrightarrow F(\alpha) = \alpha$$

en un determinado intervalo $I \subset \mathbb{R}$ donde está garantizada la convergencia (*verificar condiciones del teorema del punto fijo*).

A continuación se aplica el algoritmo de Steffensen a esta función F . Si se denota, como antes

$$S_f(x) = g(x) = x - \frac{(f(x) - x)^2}{f(f(x)) - 2f(x) + x}, \quad \forall x \in I$$

es claro que se puede considerar la composición

$$G(x) = (S_f \circ N_f)(x) = g(F(x)) = x - \frac{(F(x) - x)^2}{F(F(x)) - 2F(x) + x}$$

Al iterar esta función, a partir de un valor inicial $x_0 \in I$, la velocidad de convergencia del método de Newton, que se sabe que es al menos cuadrática, es acelerada por el método de Steffensen, por lo que se aumenta considerablemente la velocidad del proceso.

Para mostrar la eficiencia del proceso, considérese el problema de determinar un buen número de cifras decimales de $\alpha = +\sqrt{2}$, raíz positiva de la ecuación cuadrática

$$f(x) = x^2 - 2 = 0$$

La transformada de Newton es

$$F(x) = N_f(x) = \frac{x}{2} + \frac{1}{x}, \quad x \in [1, 2]$$

Además

$$\begin{aligned} F(x) - x &= \frac{1}{x} - \frac{x}{2} \\ F^2(x) = F(F(x)) &= \left(\frac{\frac{x}{2} + \frac{1}{x}}{2} + \frac{1}{\frac{x}{2} + \frac{1}{x}} \right) = \frac{(x^2 + 2)^2 + 8x^2}{4x(x^2 + 2)} \end{aligned}$$

La transformada de Steffensen de F es

$$\begin{aligned} G(x) = S_f(F(x)) &= x - \frac{\left(\frac{x}{2} + \frac{1}{x}\right)^2}{\frac{(x^2+2)^2+8x^2}{4x(x^2+2)} - x - \frac{x}{2} + x} = x - \frac{\left(\frac{2-x^2}{2x}\right)^2}{\frac{(x^2+2)^2+8x^2}{4x(x^2+2)} - \frac{x}{2}} = \\ &= x - \frac{\left(\frac{2-x^2}{2x}\right)^2}{\frac{(x^2+2)^2+8x^2-8(x^2+2)}{4x(x^2+2)}} = x - \frac{4x(x^2+2) \cdot \left(\frac{2-x^2}{2x}\right)^2}{(x^2+2)^2 - 16} = \\ &= \frac{x(x^2+2)^2 - 16x - \frac{4x}{4x^2} \cdot (2+x^2)(2-x^2)^2}{(x^2+2)^2 - 16} = \frac{6x^4 - 8x^2 - 8}{x \cdot (x^4 + 4x^2 - 12)} \end{aligned}$$

La expresión racional obtenida para $G = S_f \circ N_f$,

$$G(x) = \frac{6x^4 - 8x^2 - 8}{x^5 + 4x^3 - 12x}, \quad x \in [1, 2]$$

proporciona un algoritmo eficiente de aproximación del punto fijo $\alpha = +\sqrt{2}$, con valores racionales.

En efecto, partiendo del valor inicial $x_0 = 1$, se tiene:

$$G(1) = \frac{10}{7} \approx 1,4285714\dots$$

$$G\left(\frac{10}{7}\right) = \frac{11144}{7880} = \frac{1393}{985} \approx 1,414213198\dots$$

que da una aproximación en la segunda iteración de 6 cifras decimales exactas. Iterando una vez más,

$$G\left(\frac{1393}{985}\right) = \frac{15290740090}{10812186007} \approx 1,4142135623730950488077373\dots$$

Hay que recordar que

$$\sqrt{2} = 1,4142135623730950488016887242096980785696718753769480731766797379\dots$$

por lo que en la tercera iteración se cuenta con **¡20 cifras decimales exactas!** Esto constata la potencia del método.

Idéntico proceso se podría seguir para determinar las raíces cúbicas o n-ésimas de un número real $c > 0$. Basta considerar la función

$$f(x) = x^n - c$$

y construir la composición $G = S_f \circ N_f$, como se ha explicado. De hecho, se podría aplicar sucesivamente el proceso

$$S_f \circ S_f \circ \dots \circ N_f$$

de suerte que se vaya acelerando progresivamente la convergencia. En el fondo lo que se está haciendo en esta composición es agrupar en una función, aquella que se obtendría de ir iterando la misma n pasos. Pero simplemente el hecho de combinar estos dos algoritmos, permite obtener importantes aceleraciones en la convergencia al punto fijo.

10. Conclusiones

El análisis de los métodos mostrados permite responder adecuadamente aquellas cuestiones planteadas al final de la sección primera.

¿Qué función es la más apropiada para iterar? Aquellas funciones de clase al menos C^2 en intervalos cerrados por la composición, que contienen un punto fijo α , siendo deseable que dicho punto fijo sea atractor, i.e., $f'(\alpha) < 1$. Cuanto más próximo el valor de esta derivada se acerca a cero, mayor velocidad de convergencia.

¿Cómo caracterizar la dinámica del sistema a partir de su derivada? En el caso real, si $|f'(\alpha)| < 1$ el algoritmo converge al menos linealmente, para casi todo valor inicial (*punto fijo atractor*). Si $|f'(\alpha)| > 1$ el algoritmo diverge para casi todo valor inicial (*punto fijo repulsor*). En los casos en los que $|f'(\alpha)| = 1$, no se puede asegurar nada.

¿El valor inicial condiciona la velocidad de convergencia? En general el valor inicial no influye en la velocidad de convergencia, cuando se encuentra relativamente próximo al punto fijo, y el sistema es atractor. En cambio depende en buena medida de dicho valor, para sistemas

dinámicos caóticos. Discernir entre sistemas dinámicos que se comportan caóticamente, en el sentido de dependencia sensible a las condiciones iniciales, abre las puertas de la moderna teoría de sistemas dinámicos.

¿Cómo estimar el error cometido al realizar la aproximación? Las diferencias definidas

$$\epsilon_n = |x_n - \alpha|, \quad \forall n \in \mathbb{N}$$

permiten estimar el error de aproximación. El comportamiento de esta secuencia de errores absolutos marca en cierto sentido la velocidad de convergencia. El análisis de la potencia de este tipo de infinitésimos permite establecer el orden de convergencia de la sucesión iterada.

¿Es posible mejorar el algoritmo aplicado? En efecto, este modelo es una propuesta sencilla, que sugiere la búsqueda de procesos de mejora, tipo *Newton-Steffensen*, donde en pocas iteraciones se consigan aproximaciones numéricas con un error relativamente pequeño.

Un objetivo deseable sería el determinar nuevos algoritmos (*tipo Aitken-Steffensen*) que resuelvan y superen las debilidades del método expuesto. El estudio de mejoras de algoritmos es un problema matemáticamente interesante, de igual modo que lo es su desarrollo práctico dentro de un lenguaje de programación adecuado, tipo *Python* o *C++*. Nuestra línea de trabajo va orientada en este sentido.

Referencias

- [1] BURDEN, R.L.; FAIRES, J.D., *Análisis numérico.*, International Thomson, México, 1998.
- [2] PALLARÉS RUIZ, A. ; JIMÉNEZ LÓPEZ, V., *Métodos numéricos.*, Publicaciones Universidad de Murcia (UMU), Murcia 2008.
- [3] J. BANKS, J. BROOKS, G. CAIRNS, G. DAVIS Y P. STACEY., *On Devaney's definition of chaos*, Amer.Math. Monthly, Vol. 99 (1992) pp. 332–334.
- [4] L. ALSEDA, S. KOLYADA, J. LLIBRE Y L. SNOHA., *Entropy and periodic points for transitive maps.*, Trans. Amer. Math. Soc., Vol. 351 (1999) pp. 1551–1573.
- [5] N.C. BERNARDES, A. BONILLA, V. MULLER Y A.PERIS., *Li-Yorke chaos in linear dynamics.*, Institute of Mathematics, Academic of Sciences of the Czech Republic, preprint no. 22–2012.
- [6] B. CASCALES, J. M. MIRA, J. ORIHUELA Y M. RAJA., *Análisis Funcional.*, Electolibris. Murcia, 2012.
- [7] R. L. DEVANEY. , *An introduction to chaotic dynamical systems.*, Benjamin-Cummings. Menlo Park, 1986.

Sobre el autor:

Federico Ruiz López
 Email: federico.ruiz2011@gmail.com
 IES Jaime II (Alicante)

