UNIVERSIDAD DE LA RIOJA

TRABAJO FIN DE ESTUDIOS

Título

Sobre la convergencia del método de Newton en la resolución de sistemas no lineales

Autor/es

BEATRIZ MARTÍNEZ MORAL

Director/es

JOSÉ ANTONIO EZQUERRO FERNÁNDEZ y MIGUEL ANGEL HERNÁNDEZ VERÓN

Facultad

Facultad de Ciencia y Tecnología

Titulación

Grado en Matemáticas

Departamento

MATEMÁTICAS Y COMPUTACIÓN

Curso académico

2019-20

publicaciones.unirioja.es
E-mail: publicaciones@unirioja.es

UNIVERSIDAD DE LA RIOJA

Facultad de Ciencia y Tecnología

TRABAJO FIN DE GRADO

Grado en Matemáticas

Sobre la convergencia del método de Newton en la resolución de sistemas no lineales

---

On the convergence of Newton's method for solving nonlinear systems

Realizado por:

Beatriz Martínez Moral

Tutelado por:

José Antonio Ezquerro Fernández

Miguel Ángel Hernández Verón

Logroño, febrero,2020

Resumen

El problema de resolver una ecuación no lineal aparece con frecuencia en numerosas áreas científicas. Las técnicas matemáticas para obtener una solución exacta de una ecuación no lineal son muy limitadas, por lo que con frecuencia este tipo de problemas se resuelven de forma aproximada mediante técnicas iterativas. De entre ellas, es ampliamente utilizado el método iterativo de Newton, debido a su eficacia y fácil aplicación.

El objetivo principal de este trabajo fin de grado es analizar la convergencia del método de Newton cuando éste se aplica a la resolución de sistemas de ecuaciones no lineales. Para ello, en primer lugar se muestra la procedencia y evolución histórica del método de Newton y su construcción a partir del Teorema del punto fijo. En segundo lugar, una vez presentado el método, se enuncian y demuestran algunos de los teoremas de convergencia para el método de Newton más importantes, abarcando los tres tipos de convergencia local, semilocal y global.

Como consecuencia de los teoremas de convergencia estudiados podremos obtener resultados de existencia y unicidad de solución para la resolución de sistemas de ecuaciones no lineales y, en algunas ocasiones, dominios de convergencia para el método de Newton.

Se incluyen también dos ejemplos prácticos consistentes en la resolución de ecuaciones integrales y, con el fin de realizar los cálculos necesarios, un ejemplo de programación del método de Newton elaborado en *Mathematica*.

Abstract

Solving a nonlinear equation is a mathematical problem that is frequently approached in multiple scientific areas. The mathematical techniques used to obtain an exact solution from a nonlinear equation are very limited, therefore this kind of equations are usually solved in an approximated approach using iterative techniques. One of those techniques that is commonly used is the Newton's iterative method, due to its simplicity and efficiency.

The main goal of this degree final project is to study the convergence of the Newton's method when it is applied to obtain the solution of nonlinear equation systems. First, the origin and historical evolution of the Newton's method, as well as its development from the Fixed Point Theorem are exposed. Secondly, after the introduction of the method, some of the most important convergence theorems for the Newton's method are given and proved, including the three types of convergence: local, semilocal and global convergence.

Moreover, as a consequence of the convergence theorems studied, we may obtain conclusions about the existence and uniqueness of solution of nonlinear equation systems and, in some cases, convergence domains for the Newton's method.

Finally, two practical examples based in the resolution of integral equations are included. With the objective of performing the needed calculations, a programmation example of the Newton's method using *Mathematica* is also presented.

Índice general

Introducción

La resolución de ecuaciones no lineales es uno de los problemas más estudiados por numerosos matemáticos a lo largo de la historia. Su importancia radica en que los resultados obtenidos no sólo gozan de gran interés teórico si no que además, en la práctica, proporcionan herramientas que son imprescindibles para la resolución de problemas en diversas áreas científicas.

Es bien conocido que, en general, no es posible hallar una solución exacta de una ecuación $f(x) = 0$ no lineal, por lo que es necesario recurrir a técnicas de análisis numérico para su resolución. Lo habitual en este tipo de problemas es utilizar técnicas iterativas en las que, empezando en un punto inicial, se van generando aproximaciones sucesivas de la solución.

Sin duda, dentro de estas técnicas iterativas, el método de Newton es el procedimiento más estudiado y empleado en la práctica. Con el objetivo de de aproximar una solución x^* de una ecuación no lineal $f(x) = 0$, el método de Newton consiste en construir, a partir de una aproximación inicial $x^{(0)}$ de x^* una sucesión de la forma

$$x^{(n+1)} = x^{(n)} - \frac{f(x^{(n)})}{f'(x^{(n)})}, \qquad n \geq 0.$$

En condiciones adecuadas, la sucesión anterior converge a la solución buscada x^*.

Así, los teoremas de convergencia para un método iterativo surgen de la necesidad de determinar cuáles son esas condiciones adecuadas que, en caso de verificarse, hacen que las sucesivas aproximaciones generadas por un método iterativo estén cada vez más próximas a la solución. Se clasifican en tres tipos:

- Teoremas de convergencia local: se imponen condiciones sobre la función f y la solución x^*.
- Teoremas de convergencia semilocal: se imponen condiciones sobre la función f y el punto de partida $x^{(0)}$ que inicia el método iterativo.
- Teoremas de convergencia global: se imponen condiciones sobre la función f.

El presente trabajo fin de grado se elabora con el objetivo de presentar los principales teoremas de convergencia para el método de Newton para los casos en los que la función f representa un sistema de n ecuaciones no lineales, siendo x un vector incógnita n dimensional. Para ello, previamente se estudia la procedencia y contrucción del método de Newton. Consta de 5 capítulos que se describen a continuación.

En el Capítulo 1 se relata una breve historia del método de Newton, que nos muestra el gran interés que ha suscitado a numerosos matemáticos a lo largo de los años. Abarca desde los primeros intentos por parte de las antiguas civilizaciones para hallar cantidades aproximadas, pasando por la primera definición formal del método por parte de Newton y todos los posteriores matemáticos que contribuyeron a desarrollar el teorema tal y como lo conocemos hoy en día, hasta los estudios de convergencia de Kantorovich. Cabe destacar que, además de Newton, Kantorovich representa un matemático de gran importancia dentro de este TFG, puesto que sus resultados de convergencia para el método de Newton son la base de los aquí expuestos.

El Capítulo 2, dedicado a los fundamentos teóricos necesarios para el desarrollo del TFG, se divide en dos partes. Dada la necesidad de trabajar con vectores y matrices, nuestro entorno de trabajo serán los espacios vectoriales normados. Estos espacios están dotados de una norma que es la herramienta que utilizamos para poder "medir" o "comparar" vectores y matrices y cuyas características y propiedades se exponen en la primera parte. A lo largo de la segunda parte se introducen los conceptos de aplicación Lipschitziana y aplicación contractiva, los cuales, como veremos a lo largo del trabajo, resultan imprescindibles para el estudio de la convergencia de un método iterativo, y culmina con el enunciado del Teorema del punto fijo, que constituye la base para la elaboración de los métodos iterativos expuestos en el Capítulo 3.

En el Capítulo 3 se muestra la construcción, a partir del Teorema del punto fijo, de métodos iterativos para la resolución de ecuaciones no lineales. Si bien este TFG se centra el el método de Newton, también se incluye en este capítulo el método de aproximaciones sucesivas como su predecesor. Además, se enuncian algunos teoremas de convergencia en los que se exige, como condición principal, que la función a la que se le aplica el método iterativo sea contractiva. No obstante, la contractividad es una condición que rara vez se verifica en este tipo de funciones, por lo que estos teoremas son muy limitados y son necesarios teoremas más generales, en los que se exigen condiciones más suaves, como los que se exponen en los capítulos posteriores. Al final de este capítulo se muestra un ejemplo de programa para la resolución de sistemas de ecuaciones no lineales mediante el método de Newton realizado con el software *Mathematica*.

En los Capítulos 4 y 5 se enuncian y demuestran teoremas de convergencia para el método de Newton. El Capítulo 4, dada su importancia, se dedica exclusivamente a presentar una versión del Teorema de convergencia semilocal de Kantorovich. En el Capítulo 5 se enuncia y demuestra un teorema de convergencia global que, además de asegurar la convergencia del método, nos proporciona un dominio de convergencia restringida para la solución, esto es, una región en la que cualquier punto contenido en ella puede ser elegido como punto de inicio del método, asegurando la convergencia. Por último, mostramos el gran interés de este último teorema incluyendo dos teoremas de convergencia para el método de Newton, uno de convergencia semilocal y otro de convergencia local, que se obtienen como casos particulares del mismo. Cada uno de estos dos capítulos se acompaña de un ejemplo práctico, donde se analizará la convergencia y aplicación del método de Newton en la resolución de ecuaciones integrales. Servirán también estos ejemplos para mostrar el funcionamiento del programa elaborado el Capítulo 3. El resto de cálculos necesarios se realizarán también con *Mathematica*.

Los teoremas de convergencia escogidos gozan de gran interés ya que son fruto del trabajo desarrollado, a partir del Teorema de Kantorovich, por investigadores matemáticos en los últimos años, cuyo propósito se centra en construir resultados cada vez más eficaces y generales.

Capítulo 1

Breve historia del método de Newton

El método de Newton debe su nombre al científico británico Isaac Newton (1643-1727), quien sus trabajos de finales del siglo XVII muestra un método de aproximación a las raíces de una ecuación basado en repeticiones sucesivas de un mismo proceso. Aunque en estos escritos todavía no se hace referencia a procesos iterativos ni se utiliza el cálculo diferencial, son considerados el nacimiento del método de Newton.

No obstante, la idea general de resolver una ecuación mediante el uso de términos correctores que, sumados a una estimación de la solución, proporcionan una nueva solución aproximada mejor que la anterior ya había sido usada por antiguas civilizaciones para desarrollar métodos de aproximación de raíces.

Algunos métodos desarrollados en la Grecia clásica para aproximar números irracionales mediante números racionales o en la antigua mesopotamia para aproximar la raíz cuadrada de un número seguían este tipo de procedimientos. Por ejemplo, en la tablilla babilónica VAT6598, fechada en 2000-1700 a. C., se plantea, entre otros, el problema de encontrar la diagonal de un rectángulo de altura 40 y lado 10. Con la notación actual, el problema se traduciría en encontrar

$$\sqrt{40^2 + 10^2} = \sqrt{1700}.$$

En la misma tablilla se propone como aproximación el valor 41+15/60. Aunque no se conozca con certeza la forma en que los babilónicos obtenían este valor, algunos autores destacan el hecho de que esta cantidad coincida con la conocida aproximación para una raíz cuadrada

$$\sqrt{h^2 + l^2} \simeq h + \frac{l^2}{2h}$$

para $h = 40$ y $l = 10$.

La aproximación anterior se conoce como *fórmula de Herón*

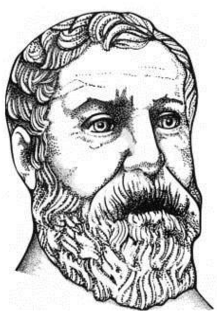

Figura 1.1: Herón de Alejandría (10-70 a. C aprox.)

para el cálculo de raíces cuadradas, y aparece por primera vez en el el primer tomo de la *Métrica* que Herón (10-70 d. C. aproximadamente) publicó en el siglo I. En este texto, descubierto por H. Schöne en 1896, se muestra cómo Herón calculaba el área de un rectángulo de lados 7, 8 y 9 unidades, es decir $\sqrt{720}$. Además, se refiere explícitamente a que una aproximación dada puede ser tomada como punto de partida para obtener mejores aproximaciones, siendo ésta la primera referencia de la utilización de un sistema iterativo para resolver un problema.

A partir de la fórmula de Herón, las técnicas para calcular la raíz cuadrada de un número (y, en general, las raíces $n-$ésimas) se fueron transmitiendo y/o redescubriendo a través de los siglos y de las civilizaciones hasta el siglo XVII. Aunque no hay muchas evidencias escritas de la que ocurrió durante este largo periodo de tiempo, sí que se pueden encontrar algunas referencias sobre métodos para el cálculo de raíces $n-$ésimas. Por ejemplo, en una versión del siglo III del libro matemático chino *Jiuzhang suansu* (*Nueve capítulos del arte matemático*) con comentarios de Liu Hui (220-280 d. C aproximadamente) puede encontrarse una colección de problemas que requieren el cálculo de raíces cuadradas y cúbicas. Posteriormente, en el siglo IV, Teón de Alejandría (335-405 aproximadamente) desarrolló un método totalmente geométrico para el cálculo aproximado de raíces cuadradas.
A partir del siglo XII, pueden destacarse los trabajos del matemático persa Sharaf al-Din al-Tusi (1135-1213) en los que da soluciones, tanto algebraicas como numéricas, de algunas ecuaciones cúbicas y los trabajos del matemático árabe Al-Marrakushi ibn Al-Banna (1256-1321)quien en su escrito *Raf al-Hijab* indica cómo calcular raíces cuadradas usando series y fracciones continuas. Parece ser que Al-Banna fue un gran recopilador de los conocimientos matemáticos de su época y que en sus escritos nos muestra su versión de los trabajos de matemáticos árabes anteriores.

Figura 1.2: François Viète (1540-1603)

El problema de encontrar la raíz $n-$ésima de un número fue evolucionando hacia el problema más general de encontrar las raíces de una ecuación polinómica e, incluso, de una ecuación trascendental. A partir del siglo XV, el problema se fue bifurcando en varias líneas: resolución algebraica de ecuaciones polinómicas, resolución aproximada usando iteraciones de punto fijo, aproximaciones mediante fracciones continuas, etc.

Es posible que algunos de los trabajos realizados por los algebristas árabes del sigo XII fueran conocidos por los principales algebristas del siglo XVI, entre los que se destaca el matemático francés François Viète (1540-1603). Viète centró su trabajo en la solución numérica de ecuaciones no lineales, desarrollando técnicas para hallar, de forma genérica, las soluciones positivas de polinomios mónicos de grados 2 a 6. Además, es el primero en utilizar las letras pera representar los parámetros de una ecuación (no sólo las incógnitas) y en darse cuenta de la relación existente entre las raíces y los coeficientes de un polinomio.

El trabajo de Vietè caería en poco tiempo en el olvido, debido sobre todo a la aparición de la geometría cartesiana. Sin embargo, varios años despúes, fueron publicadas dos obras, *Francisci Vieate Opera Mathematica* por Frans van Schooten en 1646 y *Clavis Mathematicae* por William Oughtred en 1647, en las que sus autores se ocuparon de recopilar y reescribir los trabajos de Viète. A partir de estas obras, y por tanto familiarizado con el trabajo de Vietè, Newton realizó extensas anotaciones que muestran los primeros signos del interés de Newton en la solución numérica de ecuaciones no lineales.

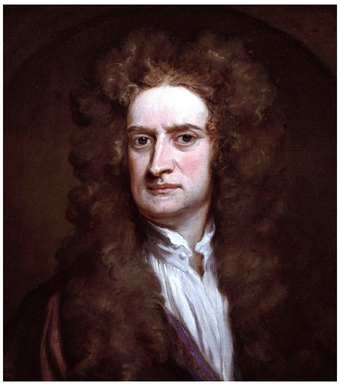

Figura 1.3: Isaac Newton (1643-1727)

La primera referencia escrita sobre el método de Newton se encuentra en *De analysi per aequationes numero terminorum infinitas*, en una carta escrita a sus colegas Barrow y Collins en 1669, pero que no fue publicada hasta 1711. A los dos años de escribir la citada carta, en 1671, Newton desarrollaba su método en *De metodis fluxionium et serierum infinitarum*. De nuevo, esta obra tardó en publicarse y no es hasta 1736 cuando se publica una traducción de la misma bajo el título de *Method of fluxions*.

Para hacernos una idea de cómo trabajaba Newton, podemos ilustrar su método con el mismo ejemplo que él consideró, la ecuación $x^3 - 2x - 5 = 0$. Newton argumentaba de la siguiente manera:

> Por tanteo, se ve que la solución está cerca de 2. Haciendo $x = 2 + \epsilon$ y sustituyendo en la ecuación se obtiene:
>
> $$\epsilon^3 + 6\,\epsilon^2 + 10, \epsilon - 1 = 0 \tag{1.1}$$
>
> Ignorando los términos $\epsilon^3 + 6\,\epsilon^2$ con el pretexto de que ϵ es pequeño, se llega a que $10\ \epsilon - 1 \simeq 0$ ó $\epsilon = 0,1$. Entonces, $x = 2,1$ es una aproximación de la solución mejor que la inicial.

> Haciendo ahora $\epsilon = 0,1 + v$ y sustituyendo en 1.1 se sigue que
>
> $$v^3 + 6,3\,v^2 + 11,23\,v + 0,061 = 0.$$
>
> Ignorando de nuevo los términos en v de grado mayor o igual que dos, se llega a que $v \simeq -0,054$ y, por tanto, $x = 2,046$ es una aproximación que mejora las anteriores.

Newton indicaba que el proceso se puede repetir tantas veces como sean necesarias.

Notemos que, llamando $f(x) = x^3 - 2x - 5$, tenemos $f'(x) = 3x^2 - 2$ y, aplicando la fórmula del método de Newton tomando como punto de partida la aproximación inicial $x_0 = 2$, la siguiente aproximación es

$$x_1 = 2 - \frac{f(2)}{f'(2)} = 2 + \frac{1}{10} = 2,1,$$

valor que coincide con el obtenido por Newton.

Newton considera la aplicación de su método tan sólo en ecuaciones polinómicas. Además, no se tiene constancia de que utilizara el cálculo ni el concepto de derivada. Pese a ello, el método que describe marca el punto de inicio para el posterior desarrollo, por parte de numerosos matemáticos, de su formulación tal y como la consideramos hoy en día.

El primer matemático en desarrollar la idea de iteración fue Joseph Raphson (1648-1715). Como hemos visto, el proceso que sigue Newton consiste en añadir un término corrector ϵ a una aproximación dada y truncar el binomio de Newton en el segundo término, en expresiones de tipo

$$(a + \epsilon)^n \simeq a^n + na^{n-1}\epsilon,$$

de forma que puede obtenerse el valor aproximado de ϵ mediante la resolución de una ecuación lineal.

Éste no puede considerarse un método iterativo, en el sentido de que pueda hallarse una aproximación partiendo de la anterior, puesto que cada paso requiere del cálculo de un polinomio en función del término corrector y el valor estimado final no es calculado hasta final del proceso como la suma de los términos correctores ($x = \epsilon + v + \dots$) en lugar de obtener sucesivas estimaciones x_i. Sin embargo, Raphson publica en 1690 un tratado, *Analysis Aequationum Universallis*, en el que, además de presentar una versión mejorada del método de Newton, muestra fórmulas explícitas para el término corrector para algunos casos particulares de ecuaciones. En concreto, calcula los términos correctores para las ecuaciones $f(x) = x^3 - r = 0$ y $g(x) = x^3 - px - q = 0$ indicando que, si x_0 es una estimación de la solución, puede obtenerse una estimación mejor $x_0 + c$, donde

$$c = \frac{r - x_0^3}{3x_0^2} \qquad y \qquad c = \frac{q + px_0 - x_0^3}{3x_0^2 - p},$$

respectivamente.

Notemos que

$$-\frac{f(x)}{f'(x)} = \frac{r - x^3}{3x^2} \quad y \quad -\frac{g(x)}{g'(x)} = \frac{q + px - x^3}{3x^2 - p},$$

por lo que los cálculos realizados por Raphson coinciden nuevamente con los que se realizan utilizando la formulación actual del método.

Pese a no haber sido publicado hasta 1736, el proceso desarrollado por Newton era bien conocido por los científicos de la época. El hecho de que Raphson publicara su obra 46 años antes que el *Método de las fluxiones* de Newton propició que su trabajo gozara de gran relevancia, llegando incluso a denominarse al proceso como el método de Newton-Raphson por numerosos autores.
No obstante, en los trabajos de Raphson no se aprecia todavía la conexión existente entre el término corrector, la función que define la ecuación y su derivada.

La incorporación del cálculo diferencial se debe a Thomas Simpson (1710-1761). En 1740 publica su obra *Essays on mathematics*, donde establece el método tal y como lo conocemos ahora salvo aspectos notacionales, puesto que describe el método de forma retórica.

Figura 1.4: Thomas Simpson (1710-1761)

Otro de los grandes aportes de Simpson fue el de extender el método a funciones cualesquiera, no solamente polinomios. Además, es el primero en aplicar el método a un sistema de ecuaciones no lineales $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. En concreto, muestra el proceso a seguir para la resolución de un sistema de dos ecuaciones con dos incógnitas, en el cual describe la forma de calcular los elementos de la matriz jacobiana $J_{\mathbf{f}}(\mathbf{x})$ y , mediante el uso de la regla de Cramer, muestra la fórmula explícita de la solución del correspondiente sistema de ecuaciones lineales

$$J_{\mathbf{f}}(\mathbf{x})(\mathbf{x}^{(\mathbf{n+1})} - \mathbf{x}^{(\mathbf{n})}) = -\mathbf{f}(\mathbf{x}^{(\mathbf{n})}).$$

Pese a su carácter realmente innovador, las contribuciones de Simpson no gozaron de gran reconocimiento en su época.

La primera vez que aparece la discusión de la convergencia del método de Newton es en 1768, en el *Traité de la résolution des équations en general* de Jean Raymond Mourraille (1720-1808), en el que utiliza la representación geométrica para explicar el comportamiento de la sucesión iterativa producida por el algoritmo de Newton. Mourraille observa por primera vez que, dependiendo del punto de salida elegido, la sucesión generada por el método puede converger a alguna de las raíces de la ecuación, oscilar, tender a infinito, o acercarse a un límite que no es solución de la ecuación. Muestra, además, que la convergencia puede ser más o menos rápida.

Posteriormente, Joseph-Louis Lagrange (1736-1813) en su *Traité de la résolution des équations numériques de tous les degrés* publicado en 1808, afirma que el método de Newton es el que se emplea habitualmente para resolver ecuaciones numéricas. Sin embargo, advierte que el método sólo se puede usar en ecuaciones "casi resueltas", en el sentido de que para aplicarlo se necesita una buena aproximación de la solución. Además, plantea dudas sobre la exactitud de cada nueva iteración y observa que el método puede tener problemas para el caso de raíces múltiples o muy próximas entre sí.

Jean Baptiste Joseph Fourier (1768-1830) fue el primero en analizar la convergencia del método de Newton en una nota titulada *Question dánalyse algébraique* (1818). Fourier expresa el método con la notación actual y lo bautiza como "la métode newtoniense", haciendo referencia a las obras de Newton, Raphson y Lagrange pero no realizando ningún tipo de mención al trabajo de Simpson.

En 1829, tras varios años dedicándose al estudio del método de Newton, el matemático Augustin Louis Cauchy (1789-1857) publica en su obra *Leçons sur le Calcul différentiel* los primeros resultados de convergencia global del método. Cauchy da condiciones, en términos de las derivadas f' y f'', para asegurar que el método converge a una solución de $f(x) = 0$ para todo punto de partida x_0 perteneciente a un intervalo determinado.

Figura 1.5: Augustin Louis Cauchy (1789-1857)

El estudio de la extensión del método de Newton al caso multidimensional, iniciado por Simpson para el caso $n = 2$ en 1740, cayó en el olvido hasta comienzos del siglo XX. En 1916, Henry B. Fine (1858-1928) y Albert A. Bennet (1888-1971), ambos profesores de la Universidad de Princeton, publican sus artículos *On Newton's Method of Approximation* y *Newton's method in general analysis* para sistemas de ecuaciones y para ecuaciones funcionales respectivamente. En el trabajo de Fine se da un resultado de convergencia para el método de Newton aplicado a un sistema de ecuaciones en el cual no se asume la existencia de solución, si no que sólo se exigen condiciones sobre el punto de partida. En el trabajo de Bennet se justifica el empleo del método de Newton también en el caso de ecuaciones funcionales, extendiendo el trabajo de Fine. Ambos trabajos resultaron muy innovadores en una época en la que el análisis funcional era una rama de las Matemáticas que estaba dando sus primeros pasos. De hecho, sólo unos años más tarde, en 1932, Stefan Banach (1892-1945) introdujo la noción de espacios de Banach en su famoso libro *Theorie des opérations linéaires.*

A partir de ese momento, varios investigadores muestran interés en la extensión del método de Newton para sistemas de ecuaciones y las publicaciones sobre el tema empiezan a ser numerosas. Entre ellas que puede destacarse las realizadas por Alexander Ostrowski (1893-1986) *Über die Konvergenz und die Abrundungsfestigkeit des Newtonschen* (1937) y *Über einen Fall der Konvergenz des Newtonschen Näherungsverfahrens* (1938), en las que se estudian y comparan distintas condiciones de convergencia dadas previamente por otros autores. Además se dan estimaciones del error cometido al aplicar el método de Newton para resolver un sistema de ecuaciones no lineales.

Es en este contexto cuando el matemático soviético Leonid V. Kantorovich (1912-1986) se plantea en *On Newton's method for functional equations* (1948) [7] la extensión del método de Newton para resolver ecuaciones funcionales del tipo

$$F(x) = 0,$$

Figura 1.6: Leonid V. Kantorovich (1912-1986)

donde F es un operador no lineal diferenciable definido entre dos espacios de Banach. De esta forma, si $F'(x^{(n)})$ es la derivada de Frèchet del operador $F(x)$ en el punto $x^{(n)}$ y $\left(F'(x^{(n)})\right)^{-1}$ es su operador inverso, Kantorovich reescribe el método de Newton como

$$x^{(n+1)} = x^{(n)} - \left(F'(x^{(n)})\right)^{-1} F(x^{(n)}), \qquad n \geqslant 0,$$

y elabora, además, un resultado de convergencia semilocal en el que se exigen condiciones al punto de partida pero no a la solución.

La gran importancia del trabajo de Kantorovich no se debe a la demostración en sí de sus resultados, sino al hecho de apoyarse en técticas de análisis funcional para demostrar resultados de análisis numérico. Esta nueva formulación del método de Newton permite que puedan resolverse numerosos problemas no lineales, tales como la resolución de ecuaciones integrales, ecuaciones diferenciales ordinarias y en derivadas parciales, o problemas de cálculo variacional.

Con Kantovorich se inicia el estudio "moderno"del método de Newton, siendo numerosas las publicaciones en las que se exponen variantes o nuevos campos de aplicación de los resultados de Kantovorich. De entre sus obras, además de *On Newton's method for functional equations* puede destacarse *The majoral principle and Newton's method* (1951) [8].

Capítulo 2

Fundamentos teóricos

2.1. Espacios vectoriales normados

Sea $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ con funciones componentes $f_i : \mathbb{R}^n \longrightarrow \mathbb{R}, i = 1, 2, \ldots, n$. Dado un sistema de ecuaciones no lineales

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \qquad \Leftrightarrow \qquad \begin{cases} f_1(x_1, x_2, \ldots, x_n) = 0 \\ f_2(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0 \end{cases}$$

el objetivo será desarrollar métodos iterativos que, a partir de un vector inicial $\mathbf{x^{(0)}}$, generen una sucesión de vectores $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ que vayan aproximándose a la solución $\mathbf{x}^*$ mediante el siguiente esquema de cálculo

$$\begin{cases} \mathbf{x^{(0)}} & \text{dado} \\ \mathbf{x^{(n+1)}} = g\left(\mathbf{x^{(n)}}\right) & n = 0, 1, 2, \ldots \end{cases}$$

Para el buen funcionamiento de este tipo de métodos deberá escogerse una función $\mathbf{g}(\mathbf{x})$ adecuada, ya que de ella dependerá que los valores obtenidos en las sucesivas iteraciones estén cada vez más cerca de la solución. Más adelante estudiaremos la forma de definir esta función $\mathbf{g}(\mathbf{x})$ ya que, en primer lugar, debemos establecer la forma en que medimos la distancia entre los elementos del espacio al que pertenecen los elementos de la sucesión y que nos permita garantizar la proximidad entre ellos.

Definición 2.1 *Sea un espacio euclídeo no vacío E. Una* ***distancia*** *o* ***métrica*** *de E es una función*

$$d : E \times E \longrightarrow \mathbb{R}^+$$

tal que

1. $d(x, y) = 0 \Leftrightarrow x = y$,
2. $d(x, y) = d(y, x) \qquad \forall\, x, y \in E$ *(simetría)*,
3. $d(x, y) \leqslant d(x, z) + d(z, y) \qquad \forall\, x, y, z \in E$ *(desigualdad triangular)*.

*Un **espacio métrico** (E, d) es un par formado por un conjunto no vacío E y una distacia d.*

Ejemplo 2.1 *En $\mathbb{R}^n$ se pueden considerar las distancias*

$$d_1(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2}$$

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max_{1 \leqslant i \leqslant n} \{|x_i - y_i|\}$$

Siendo así $(\mathbb{R}^n, d_1)$, $(\mathbb{R}^n, d_2)$ y $(\mathbb{R}^n, d_\infty)$ espacios métricos.

Definición 2.2 *Dada una sucesión infinita de elementos $\{x^{(n)}\}_{n=0}^{\infty}$ del espacio métrico (E, d) se dice que es **convergente** hacia el elemento $x^* \in E$, si para cualquier valor $\epsilon > 0$ siempre se puede encontrar un número natural N tal que para todo índice $i > N$ se verifica que $d(x^{(i)}, x^*) < \epsilon$. Al elemento x^* anterior, si existe, se le denomina **límite** de la sucesión $\{x^{(n)}\}_{n=0}^{\infty}$.*

Teorema 2.1 ***(Unicidad del límite)*** *El límite de una sucesión $\{x^{(n)}\}_{n=0}^{\infty}$ en un espacio métrico (E, d), si existe, es único. (Ver [2]).*

Definición 2.3 *Dada una sucesión infinita de elementos $\{x^{(n)}\}_{n=0}^{\infty}$ del espacio métrico (E, d) se dice que es una **sucesión de Cauchy**, si para cualquier valor $\epsilon > 0$ siempre se puede encontrar un número natural N tal que para todo par de índices $i > N$ y $j > N$ se verifica que $d(x^{(i)}, x^{(j)}) < \epsilon$.*

Para evitar el problema de que el supuesto límite de una sucesión no pertenezca al espacio considerado, cuando sea posible, se trabajará con espacios que incluyan entre sus elementos a todos los posibles límites de sus sucesiones. Estos espacios reciben el nombre de espacios métricos completos. Más rigurosamente:

Definición 2.4 *Se dice que el espacio métrico (E, d) es un **espacio métrico completo** si toda sucesión de Cauchy de elementos de E es una sucesión convergente en (E, d).*

Ejemplo 2.2 *En $\mathbb{R}^n$ se pueden considerar las distancias d_1, d_2 y d_∞, descritas en el Ejemplo 2.1, entonces los espacios métricos $(\mathbb{R}^n, d_1)$, $(\mathbb{R}^n, d_2)$ y $(\mathbb{R}^n, d_\infty)$ son espacios métricos completos. Además, siendo C un conjunto cerrado de $\mathbb{R}^n$, los espacios métricos (C, d_1), (C, d_2) y (C, d_∞) son también espacios métricos completos.*

Lo dicho hasta aquí es aplicable a espacios métricos en general. No obstante, puesto que los elementos que deberemos tratar son vectores y matrices, necesitaremos trabajar en espacios que tengan la estructura de espacios vectoriales. En ellos la forma de medir distancias se asocia al concepto de norma de un vector (que, a su vez, generaliza el concepto de módulo de un vector). De forma más concreta puede darse la siguiente definición:

Definición 2.5 *Siendo E un espacio vectorial definido sobre un cuerpo K se denomina **norma** sobre E, y se representa por $\|\cdot\|$, a toda aplicación*

$$\|\cdot\| : E \longrightarrow \mathbb{R}^+$$

que verifica las condiciones siguientes:

i) $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$,

ii) $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|, \qquad \forall\, \lambda \in K, \quad \forall\, \mathbf{x} \in E.$

iii) $\|\mathbf{x}+\mathbf{y}\| \leqslant \|\mathbf{x}\| + \|\mathbf{y}\| \qquad \forall\, \mathbf{x}, \mathbf{y} \in E.$

*A todo espacio vectorial E sobre el que se defina una norma $\|\cdot\|$ se le denomina **espacio vetorial normado** y se representará por $(E, \|\cdot\|)$.*

Ejemplo 2.3 *En $\mathbb{R}^n$ son normas vectoriales las siguientes:*

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|, \quad \|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2}, \quad \|\mathbf{x}\|_\infty = \max_{1\leqslant i \leqslant n} \{|x_i|\},$$

*a las que se les llama, respectivamente, **norma-1**, **norma-2** y **norma infinito**.*

Definición 2.6 *Siendo $\|\cdot\|$ y $\|\cdot\|'$ dos normas definidas sobre un mismo espacio vectorial E, se dice que ambas normas son **equivalentes** si existen dos constantes K_1 y K_2 tales que se verifica:*

$$K_1\|\mathbf{x}\| \leqslant \|\mathbf{x}\|' \leqslant K_2\|\mathbf{x}\|, \qquad \forall\, \mathbf{x} \in E.$$

Ejemplo 2.4 *Si n es finito las normas sobre $\mathbb{R}^n$ introducidas en el Ejemplo 2.3 son equivalentes.*

En un espacio vectorial normado $(E, \|\cdot\|)$ podrían definirse muy diferentes distancias. Entre todas ellas es habitual trabajar con la distancia que induce la norma vectorial. Esta, de forma más precisa, se puede definir como sigue:

Proposición 2.2 *Siendo $(E, \|\cdot\|)$ un espacio vectorial normado se verifica que la aplicación $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$ es una distancia denominada **distancia asociada a la norma** $\|\cdot\|$. (Ver [3]).*

Ejemplo 2.5 *A las normas $\|\cdot\|_1$, $\|\cdot\|_2$ y $\|\cdot\|_\infty$ se les asocian respectivamente las distancias d_1, d_2 y d_∞.*

Una conclusión de lo anterior es que los espacios vectoriales normados $(\mathbb{R}^n, \|\cdot\|_1)$, $(\mathbb{R}^n, \|\cdot\|_2)$ y $(\mathbb{R}^n, \|\cdot\|_\infty)$ son espacios métricos completos.

El que dos normas sean equivalentes no quiere decir que al aplicarlas a un mismo vector tomen el mismo valor, pero sí que indica que si una de ellas toma un valor "elevado" al aplicarla a un cierto valor de E, la otra también tomará un valor "elevado" al aplicarla al mismo vector. Y si el valor es "pequeño" para una de ellas, también lo será para la otra. En ese sentido las distancias que están asociadas a dichas normas también serán equivalentes. Y por ello si una sucesión es convergente con la distancia asociada a una de las normas, también lo será con la otra. Como en el caso de trabajar en $\mathbb{R}^n$ todas las normas son equivalentes, a efectos de analizar la convergencia de una sucesión de vectores será equivalente hacerlo con cualquier norma que se considere en $\mathbb{R}^n$.

Definición 2.7 *Siendo M_n el espacio de las matrices cuadradas de orden n definidas sobre un cuerpo K (con $K = \mathbb{R}$ o $K = \mathbb{C}$) se denomina **norma matricial** definida sobre M_n y se representará por $\|\cdot\|$, a toda aplicación*

$$\|\cdot\| : M_n \longrightarrow \mathbb{R}^+$$

que verifica las propiedades siguientes:

i) $\|\mathbf{A}\| = 0 \Leftrightarrow \mathbf{A} = \mathbf{0}$,

ii) $\|\lambda \mathbf{A}\| = |\lambda| \|\mathbf{A}\|, \qquad \forall\, \lambda \in K, \quad \forall\, \mathbf{A} \in M_n$,

iii) $\|\mathbf{A} + \mathbf{B}\| \leqslant \|\mathbf{A}\| + \|\mathbf{B}\|, \qquad \forall\, \mathbf{A}, \mathbf{B} \in M_n$,

iv) $\|\mathbf{A}\mathbf{B}\| \leqslant \|\mathbf{A}\| \|\mathbf{B}\|, \qquad \forall\, \mathbf{A}, \mathbf{B} \in M_n$.

Ejemplo 2.6 *En M_n son normas matriciales las siguientes:*

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

$$\|\mathbf{A}\|_1 = \max_{1 \leqslant j \leqslant n} \left\{ \sum_{i=1}^{n} |a_{ij}| \right\},$$

$$\|\mathbf{A}\|_\infty = \max_{1 \leqslant i \leqslant n} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\},$$

donde $\|\mathbf{A}\|_F$ se conoce como la norma de Fröbenius.

Entre todas las normas matriciales que se pueden definir en M_n pueden considerarse aquellas que están definidas a partir de una norma vectorial y, por tanto, están vinculadas a normas vectoriales. Este grupo de normas matriciales se conocen con el nombre de normas matriciales subordinadas a una norma vectorial.

Definición 2.8 *Sea $\|\cdot\|$ una norma vectorial definida sobre K^n (con $K = \mathbb{R}$ o $K = \mathbb{C}$) y sea M_n el espacio de las matrices cuadradas de orden n definidas sobre K. La aplicación $\|\cdot\|$ definida de cualquiera de las formas (equivalentes entre sí) siguientes:*

$$\|\mathbf{A}\| = \sup_{\mathbf{v} \in K^n - \mathbf{0}} \left(\frac{\|\mathbf{A}\mathbf{v}\|}{\|\mathbf{v}\|} \right) = \sup_{\mathbf{v} \in K^n - \mathbf{0} \,|\, \|\mathbf{v}\| \leqslant 1} (\|\mathbf{A}\mathbf{v}\|) = \sup_{\mathbf{v} \in K^n \,|\, \|\mathbf{v}\| = 1} (\|\mathbf{A}\mathbf{v}\|)$$

*es una norma matricial que se denomina **norma matricial subordinada** a la norma vectorial $\|\cdot\|$.*

Observación 2.1 *En la definición anterior se ha utilizado el mismo símbolo ($\|\cdot\|$) para referirse a la norma matricial y a la norma vectorial. Fácilmente se distinguirá entre una y otra por el tipo de elemento al que se aplica.*

Ejemplo 2.7 *Las normas matriciales $\|\cdot\|_1$, $\|\cdot\|_2$ y $\|\cdot\|_\infty$ son normas matriciales subordinadas a las normas vectoriales de $\mathbb{R}^n$ definidas con los mismos subíndices. Sin embargo, la norma de Fröbenius no es una norma matricial subordinada.*

Las normas matriciales subordinadas permiten trabajar con formas de medir "coherentes" entre los vectores y las matrices cuando estos aparecen mezclados en los problemas que deban abordarse. Es importante en este sentido tener siempre presente la siguiente propiedad que relaciona el valor de una norma vectorial con la norma matricial subordinada a ella.

Proposición 2.3 *Siendo $\|\cdot\|$ una norma matricial definida sobre M^n (subordinada a la norma vectorial $\|\cdot\|$) se verifica que:*

$$\|\mathbf{A}\mathbf{v}\| \leqslant \|\mathbf{A}\|\|\mathbf{v}\| \qquad \forall\, \mathbf{A} \in M_n, \quad \forall\, \mathbf{v} \in K^n.$$

Demostración.
Si $\mathbf{v} = \mathbf{0}$ entonces $\|\mathbf{A}\mathbf{v}\| = 0$ y $\|\mathbf{A}\|\|\mathbf{v}\| = 0$ para cualquier $\mathbf{A} \in M_n$, por lo que en este caso se verifica el teorema.

Si $\mathbf{v} \neq \mathbf{0}$ se tiene que $\|\mathbf{v}\| \neq \mathbf{0}$ y por tanto:

$$\|\mathbf{A}\mathbf{v}\| = \left\|\mathbf{A}\frac{\mathbf{v}}{\|\mathbf{v}\|}\|\mathbf{v}\|\right\| = \left\|\mathbf{A}\frac{\mathbf{v}}{\|\mathbf{v}\|}\right\|\|\mathbf{v}\| = \frac{\|\mathbf{A}\mathbf{v}\|}{\|\mathbf{v}\|}\|\mathbf{v}\| \leqslant$$

$$\leqslant \sup_{\mathbf{u} \in K^n - \mathbf{0}} \left(\frac{\|\mathbf{A}\mathbf{u}\|}{\|\mathbf{u}\|}\right)\|\mathbf{v}\| = \|\mathbf{A}\|\,\|\mathbf{v}\|$$

y esto se tiene para cualesquiera $\mathbf{A} \in M_n$ y $\forall\, \mathbf{v} \in K^n - \mathbf{0}$. □

También pueden darse resultados que relacionan el valor de normas matriciales con la norma de la inversa de una matriz. Con este propósito se enuncian los dos siguientes lemas que serán frecuentemente utilizados en los Capítulos 4 y 5.

Lema 2.4 *Sea $\|\cdot\|$ una norma matricial subordinada y $\mathbf{A} \in M_n$. Si $\|\mathbf{A}\| < 1$, entonces existe $(\mathbf{I} - \mathbf{A})^{-1}$ y se tiene que*

$$\|(\mathbf{I} - \mathbf{A})^{-1}\| \leqslant \frac{1}{1 - \|\mathbf{A}\|}.$$

Demostración.
Para la matriz $\mathbf{A}$ se cumple que $\|\mathbf{A}^2\| = \|\mathbf{A}\,\mathbf{A}\| \leqslant \|\mathbf{A}\|\|\mathbf{A}\| = \|\mathbf{A}\|^2$.
Reiteradamente, se tiene

$$\|\mathbf{A}^n\| \leqslant \|\mathbf{A}\|^n, \qquad n = 1, 2, \ldots$$

Consideramos la serie

$$\mathbf{S} = \mathbf{I} + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^n + \cdots ,$$

que converge, ya que $\|\mathbf{A}\| < 1$ y está mayorada por la serie numérica convergente

$$1 + \|\mathbf{A}\| + \|\mathbf{A}\|^2 + \cdots + \|\mathbf{A}\|^n + \cdots$$

Ahora como

$$\mathbf{S}(\mathbf{I} - \mathbf{A}) = (\mathbf{I} + \mathbf{A} + \cdots + \mathbf{A}^n + \cdots)(\mathbf{I} - \mathbf{A}) = \mathbf{I}$$

y, análogamente, $(\mathbf{I} - \mathbf{A})\mathbf{S} = \mathbf{I}$, entonces $\mathbf{S} = (\mathbf{I} - \mathbf{A})^{-1}$ y se tiene

$$\|(\mathbf{I} - \mathbf{A})^{-1}\| = \|\mathbf{S}\| \leqslant 1 + \|\mathbf{A}\| + \|\mathbf{A}\|^2 + \cdots + \|\mathbf{A}\|^n + \cdots = \frac{1}{1 - \|\mathbf{A}\|}.$$

$\square$

Lema 2.5 *Sean* $\mathbf{A}$, $\mathbf{B} \in M_n$*. Si* $\mathbf{A}$ *es no singular y* $\|\mathbf{I} - \mathbf{A}^{-1}\mathbf{B}\| < 1$*, entonces* $\mathbf{B}$ *es no singular y*

$$\|\mathbf{B}^{-1}\| \leqslant \frac{\|\mathbf{A}^{-1}\|}{1 - \|\mathbf{I} - \mathbf{A}^{-1}\mathbf{B}\|}.$$

Demostración.
Tomando $\mathbf{I} - \mathbf{A}^{-1}\mathbf{B}$ en el Lema 2.4, se tiene que existe $(\mathbf{I} - (\mathbf{I} - \mathbf{A}^{-1}\mathbf{B}))^{-1} = \mathbf{A}^{-1}\mathbf{B}$ y

$$\|(\mathbf{A}^{-1}\mathbf{B})^{-1}\| \leqslant \frac{1}{1 - \|\mathbf{I} - \mathbf{A}^{-1}\mathbf{B}\|}.$$

Entonces

$$\|\mathbf{B}^{-1}\| = \|\mathbf{B}^{-1}\mathbf{A}\mathbf{A}^{-1}\| = \|(\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{A}^{-1}\| \leqslant$$
$$\leqslant \|(\mathbf{A}^{-1}\mathbf{B})^{-1}\|\|\mathbf{A}^{-1}\| = \frac{\|\mathbf{A}^{-1}\|}{1 - \|\mathbf{I} - \mathbf{A}^{-1}\mathbf{B}\|}.$$

$\square$

2.2. Aplicaciones Lipschitzianas

Con los resultados expuestos hasta el momento, pueden comenzarse a analizar las características que hacen que una aplicación $\mathbf{g}(\mathbf{x})$ sea una buena candidata para llevar a cabo un método iterativo

$$\left\{ \begin{array}{cl} \mathbf{x}^{(\mathbf{0})} & \text{dado} \\ \mathbf{x}^{(\mathbf{n+1})} = \mathbf{g}(\mathbf{x}^{(\mathbf{n})}) & n = 0, 1, 2, \ldots \end{array} \right. \tag{2.1}$$

Así, nos interesará trabajar, cuando ello sea posible, con aplicaciones $\mathbf{g}(\mathbf{x})$ para las que se pueda asegurar que la sucesión, que con ellas se genere, sea convergente hacia alguna solución del sistema que se quiere resolver.

Si $\mathbf{x}^*$ fuese el límite de la sucesión generada por el método iterativo y, además, la aplicación $\mathbf{g}(\mathbf{x})$ fuese continua se verificará que

$$\lim_{n\to\infty}(\mathbf{x}^{(\mathbf{n+1})}) = \lim_{n\to\infty}\mathbf{g}(\mathbf{x}^{(\mathbf{n})}) \Leftrightarrow \mathbf{x}^* = \mathbf{g}(\mathbf{x}^*).$$

Los puntos $\mathbf{x}^*$ del dominio sobre el que está definida una aplicación $\mathbf{g}(\mathbf{x})$, que verifican $\mathbf{x}^* = \mathbf{g}(\mathbf{x}^*)$, reciben el nombre de puntos fijos de la aplicación. Más concretamente:

Definición 2.9 *Sean (E, d) un espacio métrico y $\mathbf{g} : E \longrightarrow E$ una aplicación. Se denomina **punto fijo** de la aplicación $\mathbf{g}$ a todo elemento $\mathbf{x}^*$ de E tal que $\mathbf{x}^* = \mathbf{g}(\mathbf{x}^*)$.*

Interesará por tanto trabajar con funciones $\mathbf{g}(\mathbf{x})$ que posean un punto fijo. Un tipo de tales funciones son las que se denominan contracciones y que pasamos a definir a continuación.

Definición 2.10 *Sean (E, d) y (V, d') dos espacios métricos y $\mathbf{g} : E \longrightarrow V$ una aplicación . Se dice que $\mathbf{g}$ es una **aplicación Lipschitziana** cuando existe una constante real $L > 0$ tal que:*

$$d'(\mathbf{g}(\mathbf{x}), \mathbf{g}(\mathbf{y})) \leqslant L\, d(\mathbf{x}, \mathbf{y}) \qquad \forall\, \mathbf{x}, \mathbf{y} \in E.$$

*A la menor constante L que verifica la condición anterior se la denomina **constante de Lipschitz (o razón)** de la aplicación.*

Observación 2.2 *En el caso de que se esté trabajando sobre los espacios vectoriales normados $(E, \|\cdot\|)$ y $(V, \|\cdot\|')$, toda aplicación $\mathbf{g} : E \longrightarrow V$ Lipschitziana de razón L verificará:*

$$\|g(\mathbf{x}) - g(\mathbf{y})\|' \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad \forall\, \mathbf{x}, \mathbf{y} \in E.$$

Proposición 2.6 *Sean (E, d) y (V, d') espacios métricos. Toda aplicación Lipschitziana definida en (E, d) y con valores en (V, d') es una aplicación continua en todo E. (Ver [3]).*

Definición 2.11 *Se denomina **contracción** sobre E a toda aplicación $\mathbf{g}$ Lipschitziana que verifique las dos condiciones siguientes:*

i) Estar definida en un espacio métrico (E, d) sobre sí mismo: $\mathbf{g} : E \longrightarrow E$.

ii) Tener una constante de Lipschitz $L < 1$.

El hecho de que para las contracciones se garantice la convergencia de la sucesión generada mediante el esquema iterativo se debe al teorema que se expone a continuación y cuya demostración, realizada mediante la técnica de aproximaciones sucesivas, recoge las bases de los métodos de resolución de sistemas de ecuaciones.

Teorema 2.7 ***(Teorema del punto fijo)**. Toda contracción definida sobre un espacio métrico completo admite un único punto fijo.*

Demostración.
a) **Existencia.** Sea $\mathbf{g} : E \longrightarrow E$ una contracción, con constante de Lipschitz L definida en el espacio métrico (E, d) y sea $\mathbf{x^{(0)}}$ un elemento cualquiera de E. Considérese entonces la sucesión formada a partir de $\mathbf{x^{(0)}}$ mediante:

$$\mathbf{x}^{(n+1)} = \mathbf{g}\left(\mathbf{x}^{(n)}\right), \qquad n = 0, 1, 2, \ldots$$

Para la sucesión $\left\{\mathbf{x}^{(n)}\right\}_{n=0}^{\infty}$ se verificará:

$$\begin{gathered}
d(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) = d(g(\mathbf{x}^{(0)}), g((\mathbf{x}^{(1)})) \leqslant L\ d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) \\
d(\mathbf{x}^{(2)}, \mathbf{x}^{(3)}) = d(g(\mathbf{x}^{(1)}), g(\mathbf{x}^{(2)})) \leqslant L\, d(\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \leqslant L^2\ d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) \\
\vdots \\
d(\mathbf{x}^{(i)}, \mathbf{x}^{(i+1)}) \leqslant L^i\ d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) \\
\vdots
\end{gathered}$$

De estas desigualdades, aplicando la desigualdad triangular de las distancias, resultará que:

$$\begin{gathered}
d(\mathbf{x}^{(i)}, \mathbf{x}^{(i+p)}) \leqslant d(\mathbf{x}^{(i)}, \mathbf{x}^{(i+1)}) + d(\mathbf{x}^{(i+1)}, \mathbf{x}^{(i+p)}) \leqslant \\
\leqslant d(\mathbf{x}^{(i)}, x^{(i+1)}) + d(\mathbf{x}^{(i+1)}, x^{(i+2)}) + d(\mathbf{x}^{(i+2)}, \mathbf{x}^{(i+p)}) \leqslant \\
\leqslant d(\mathbf{x}^{(i)}, \mathbf{x}^{(i+1)}) + d(\mathbf{x}^{(i+1)}, \mathbf{x}^{(i+2)}) + \cdots + d(\mathbf{x}^{(i+p-1)}, \mathbf{x}^{(i+p)}) \leqslant \\
\leqslant L^i\, d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) + L^{(i+1)}\, d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) + \cdots + L^{(i+p-1)}\, d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) = \\
= L^i\, d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})\, (1 + L + \cdots + L^{(p-1)}) \leqslant L^i\, d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)}) \left(\sum_{i=0}^{\infty} L^n\right).
\end{gathered}$$

En la expresión anterior el sumatorio que aparece representa la suma de una progresión geométrica de razón $L < 1$, por tanto:

$$\sum_{n=0}^{\infty} L^n = \frac{1}{1-L}$$

Lo que nos conduce a que:

$$d(\mathbf{x}^{(i)}, \mathbf{x}^{(i+p)}) \leqslant \frac{L^i}{1-L} d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})$$

Veremos ahora que la sucesión $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ es de Cauchy.
Al ser $\mathbf{g}(\mathbf{x})$ una contracción, se tiene $L < 1$. Por lo tanto, para cualquier valor $\epsilon > 0$ bastará con considerar $N \in \mathbb{N}$ de forma que

$$N \geqslant \left[\frac{\log\left(\frac{\epsilon\cdot(1-L)}{d(\mathbf{x}^{(0)}, \mathbf{x}^{(1)})}\right)}{\log(L)}\right] + 1$$

para que se verifique que $d(\mathbf{x}^{(i)}, \mathbf{x}^{(j)}) < \epsilon$ para todo par de índices i y j mayores que N, donde $[\,\cdot\,]$ representa la función parte entera de un número.

La sucesión $\{\mathbf{x}^{(n)}\}_{n=0}^{\infty}$ es de Cauchy y, por hipótesis del teorema, se está trabajando en un espacio métrico completo. En consecuencia, la sucesión es convergente y admitirá límite $\mathbf{x}^*$. Puesto que al ser $\mathbf{g}(\mathbf{x})$ una contracción es continua, se verificará que

$$\mathbf{g}(\mathbf{x}^*) = \lim_{n\to\infty} \mathbf{g}(\mathbf{x}^{(n)}) = \lim_{n\to\infty} \mathbf{x}^{(n+1)} = \mathbf{x}^*.$$

Por lo tanto $\mathbf{g}(\mathbf{x})$ admite un punto fijo que es el límite de la sucesión generada mediante:

$$\mathbf{x}^{(n+1)} = \mathbf{g}(\mathbf{x}^{(n)}), \qquad n = 0, 1, 2, \ldots$$

a partir de cualquier elemento $\mathbf{x}^{(0)}$ perteneciente a E.

b) **Unicidad:** Demostraremos la unicidad del punto fijo mediante reducción al absurdo. Consideremos que, en las condiciones del teorema, existen $\mathbf{a}, \mathbf{b} \in E$ tales que $\mathbf{a} \neq \mathbf{b}$. De esta forma, debe verificarse que

$$d(\mathbf{a}, \mathbf{b}) = d(\mathbf{g}(\mathbf{a}), \mathbf{g}(\mathbf{b})) \leqslant L\, d(\mathbf{a}, \mathbf{b}) < d(\mathbf{a}, \mathbf{b}),$$

lo cual es absurdo. □

Notemos que el teorema anterior establece condiciónes suficientes para la existencia de un punto fijo, pero no necesarias. Puede ocurrir que otras aplicaciones que no sean contracciones, o que estén definidas en espacios que no sean completos, puedan tener uno o varios puntos fijos.

La demostración de la existencia de punto fijo nos indica, además, que el punto fijo puede hallarse como el límite de la sucesión $\mathbf{x}^{(n+1)} = \mathbf{g}(\mathbf{x}^{(n)})$ generada a partir de cualquier $\mathbf{x}^{(\mathbf{0})} \in E$. Este hecho hace que hablemos de que hemos obtenido un resultado de convergencia global, ya que no se establece ninguna restricción para el punto de salida $\mathbf{x}^{(\mathbf{0})}$. Este tipo de resultados de convergencia global son poco habituales dado que la condición exigida a la aplicación $\mathbf{g}$, ser una contracción en el espacio E, resulta ser muy estricta en general.

Capítulo 3

Métodos iterativos

A partir del Teorema del punto fijo, pueden desarrollarse esquemas iterativos para obtener una solución aproximada de un sistema de ecuaciones no lineales. El más utilizado de todos ellos, y en cuyo análisis se centra este trabajo, es el método de Newton que, como veremos más adelante, se trata de un caso particular del método de aproximaciones sucesivas.

3.1. Método de aproximaciones sucesivas

El método de aproximaciones sucesivas para la resolución de un sistema de n ecuaciones no lineales con n incógnitas se basa en el Teorema del punto fijo. A partir de un sistema de ecuaciones $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ puede hallarse un sistema de la forma $\mathbf{g}(\mathbf{x}) = \mathbf{x}$ que sea equivalente, esto es, que tenga las mismas soluciones. De esta forma, transformamos el problema de hallar las raíces de $\mathbf{f}(\mathbf{x})$ al problema equivalente de hallar los puntos fijos de $\mathbf{g}(\mathbf{x})$.

La forma de llevar a cabo esta transformación no es única. Por ejemplo, si se suma el vector $\mathbf{x}$ a ambos lados del sistema se tendrá

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \Leftrightarrow \mathbf{x} + \mathbf{f}(\mathbf{x}) = \mathbf{x} \Leftrightarrow \mathbf{x} = \mathbf{x} - \mathbf{f}(\mathbf{x})$$

y podría tomarse tanto $\mathbf{g}(\mathbf{x}) = \mathbf{x} + \mathbf{f}(\mathbf{x})$ como $\mathbf{g}(\mathbf{x}) = \mathbf{x} - \mathbf{f}(\mathbf{x})$ para realizar la transformación.

También podría despejarse (total o parcialmente) de la primera ecuación x_1, de la segunda x_2, y así sucesivamente hasta despejar x_n de la n-ésima ecuación.

$$\mathbf{f}(\mathbf{x}) = \begin{cases} f_1(x_1, \ldots, x_n) = 0 \\ f_2(x_1, \ldots, x_n) = 0 \\ \vdots \\ f_n(x_1, \ldots, x_n) = 0 \end{cases} \Longleftrightarrow \mathbf{g}(\mathbf{x}) = \begin{cases} g_1(x_1, \ldots, x_n) = x_1 \\ g_2(x_1, \ldots, x_n) = x_2 \\ \qquad \vdots \\ g_n(x_1, \ldots, x_n) = x_n \end{cases}$$

Una vez escrito el sistema de ecuaciones en la forma $\mathbf{g}(\mathbf{x}) = \mathbf{x}$, el método de las aproximaciones sucesivas consiste en, a partir de un vector inicial $\mathbf{x}^{(\mathbf{0})}$, realizar el siguiente esquema iterativo

$$\left\{ \begin{array}{cl} \mathbf{x^{(0)}} & \text{dado} \\ \mathbf{x^{(n+1)}} = \mathbf{g}(\mathbf{x^{(n)}}) & n = 0, 1, 2, \dots \end{array} \right. \tag{3.1}$$

de forma que se genera una sucesión de vectores $\left\{\mathbf{x^{(n+1)}}\right\}_{n=0}^{\infty}$. Según el teorema del punto fijo se tiene el siguiente resultado:

Teorema 3.1 *Sea D cerrado de $\mathbb{R}^n$ y sea $\mathbf{g} : D \longrightarrow D$ una aplicación. Si $\mathbf{g}$ es una contracción sobre D, entonces la sucesión $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ dada por*

$$\mathbf{x^{(n+1)}} = \mathbf{g}(\mathbf{x^{(n)}}),$$

empezando en cualquier $\mathbf{x^{(0)}} \in D$, converge a una solución $\mathbf{x}^*$ en D del sistema de ecuaciones $\mathbf{x} = \mathbf{g}(\mathbf{x})$. Además, la solución $\mathbf{x}^*$ es única en D.

Demostración.
Por ser D un dominio cerrado de $\mathbb{R}^n$ será completo. Como $\mathbf{g}(\mathbf{x})$ es una contracción para la norma $\|\cdot\|$, por el teorema del punto fijo se verificará que:

i) Existe un único punto $\mathbf{x}^* \in D$ tal que es un punto fijo de $\mathbf{g}(\mathbf{x})$, esto es, $\mathbf{x}^* = \mathbf{g}(\mathbf{x}^*)$.

ii) La sucesión $\left\{\mathbf{x^{(n+1)}}\right\}_{n=0}^{\infty}$, dada por (3.1), converge hacia $\mathbf{x}^*$.

□

Puesto que el sistema de ecuaciones $\mathbf{x} = \mathbf{g}(\mathbf{x})$ es equivalente al sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, en las condiciones del teorema anterior el método de aproximaciones sucesivas converge hacia una solución $\mathbf{x}^*$ del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, considerando para ello un punto de salida $\mathbf{x^{(0)}}$ cualquiera de D. Es decir, hemos obtenido un resultado de convergencia global.

Demostrar que una aplicación $\mathbf{g}(\mathbf{x})$ es una contracción mediante la determinación de su constante de Lipschitz puede ser laborioso en ciertas ocasiones, resultando el teorema anterior de poca utilidad en la práctica. Ante esto, pueden considerarse algunas versiones del mismo que, pese a ser más restrictivas en sus condiciones, son más fáciles de aplicar en la práctica.

Para ello necesitaremos utilizar la matriz Jacobiana de la aplicación $\mathbf{g}(\mathbf{x})$, que denotaremos por $J_{\mathbf{g}}(\mathbf{x})$. Recordemos que, asumiendo que todas las componentes de $\mathbf{g}(\mathbf{x})$ son derivables en todos los puntos del dominio de trabajo, se define como la siguiente matriz de dimensión $n \times n$

$$J_{\mathbf{g}}(\mathbf{x}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\mathbf{x}) & \frac{\partial g_1}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial g_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial g_2}{\partial x_1}(\mathbf{x}) & \frac{\partial g_2}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial g_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_n}{\partial x_1}(\mathbf{x}) & \frac{\partial g_n}{\partial x_2}(\mathbf{x}) & \cdots & \frac{\partial g_n}{\partial x_n}(\mathbf{x}) \end{pmatrix}$$

Teorema 3.2 *Sea D cerrado de $\mathbb{R}^n$ y $\mathbf{g} : D \longrightarrow D$ una aplicación de clase $C^{(1}(D)$. Si existe $L < 1$ tal que*

$$\|J_{\mathbf{g}}(\mathbf{x})\| \leqslant L < 1 \qquad \forall\, \mathbf{x} \in D$$

entonces la sucesión $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ dada por

$$\mathbf{x^{(n+1)}} = \mathbf{g}(\mathbf{x^{(n)}}),$$

empezando en cualquier punto $\mathbf{x^{(0)}} \in D$, converge a la única solución $\mathbf{x}^*$ en D del sistema de ecuaciones $\mathbf{x} = \mathbf{g}(\mathbf{x})$.

Demostración.
Por extensión del teorema del valor medio se verifica que

$$\forall\, \mathbf{x}, \mathbf{y} \in D \quad \exists\, \mathbf{z} \in \overset{\circ}{D} \text{ tal que } \|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leqslant \|J_{\mathbf{g}}(\mathbf{z})(\mathbf{x} - \mathbf{y})\|,$$

donde $\overset{\circ}{D}$ es el interior de D. Como, por hipótesis, el valor de $\|J_{\mathbf{g}}(\mathbf{x})\|$ está acotado por $L < 1$, trabajando con la norma vectorial a la que está subordinada la anterior obtenemos

$$\begin{gathered} \forall\, \mathbf{x}, \mathbf{y} \in D \quad \|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\mathbf{y})\| \leqslant \|J_{\mathbf{g}}(\mathbf{z})(\mathbf{x} - \mathbf{y})\| \leqslant \\ \leqslant \|J_{\mathbf{g}}(\mathbf{z})\| \|\mathbf{x} - \mathbf{y}\| \leqslant L\, \|\mathbf{x} - \mathbf{y}\| < \|\mathbf{x} - \mathbf{y}\| \end{gathered}$$

por lo que, teniendo en cuenta que $\mathbf{g} : D \longrightarrow D$, resulta que $\mathbf{g}(\mathbf{x})$ es una contracción para la norma $\|\cdot\|$. Aplicando el Teorema 3.1 a $\mathbf{g}(\mathbf{x})$ se concluye la demostración.
□

Los dos teoremas anteriores establecen condiciones suficientes de convergencia global del método sobre un dominio cerrado D. Esto es, aseguran la convergencia independientemente del punto $\mathbf{x^{(0)}} \in D$ elegido para iniciar el proceso iterativo.

Cuando podemos afirmar la existencia de una solución $\mathbf{x}^*$ y se conoce un entorno de la solución, pueden establecerse resultados de convergencia local (esto es, para puntos $\mathbf{x^{(0)}}$ suficientemente próximos a la solución). Así por ejemplo, se tiene el siguiente teorema:

Teorema 3.3 *Sea D cerrado de $\mathbb{R}^n$ y sea $\mathbf{g} : D \longrightarrow D$ una aplicación de clase $C^{(1}(D)$. Supongamos que $\mathbf{x}^* \in D$ es una solución del sistema de ecuaciones $\mathbf{x} = \mathbf{g}(\mathbf{x})$ tal que $\|J_{\mathbf{g}}(\mathbf{x}^*)\| < 1$. Entonces existe un valor $\delta > 0$ tal que si $\|\mathbf{x}^* - \mathbf{x^{(0)}}\| < \delta$, la sucesión $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ dada por*

$$\mathbf{x^{(n+1)}} = \mathbf{g}(\mathbf{x^{(n)}})$$

verifica que:

i) $\|\mathbf{x}^* - \mathbf{x^{(n)}}\| < \delta$ *para cualquier* $\mathbf{x^{(n)}}$,

ii) $\lim_{n\to\infty} \mathbf{x^{(n)}} = \mathbf{x}^*$.

Demostración.
Por ser $\frac{\partial g_i}{\partial x_j}(\mathbf{x})$ $(i, j = 1, \ldots, n)$ continuas en todo $\mathbf{x} \in D$ existirá una bola abierta de centro $\mathbf{x}^*$ y radio δ', $B(\mathbf{x}^*, \delta')$, tal que en ella se verifique

$$\|J_{\mathbf{g}}(\mathbf{x})\| \leqslant L < 1 \quad \forall \mathbf{x} \in B(\mathbf{x}^*, \delta').$$

Considerando un valor $\delta < \delta'$ se tendrá por tanto que

$$\|J_{\mathbf{g}}(\mathbf{x})\| \leqslant L < 1 \quad \forall \mathbf{x} \in \overline{B(\mathbf{x}^*, \delta)},$$

donde $\overline{B(\mathbf{x}^*, \delta)}$ es una bola cerrada de centro $\mathbf{x}^*$ y radio δ.

Además,

$$\begin{aligned}\|\mathbf{x^{(n)}} - \mathbf{x}^*\| = \|\mathbf{g}(\mathbf{x^{(n-1)}}) - \mathbf{g}(\mathbf{x}^*)\| < L\|\mathbf{x^{(n-1)}} - \mathbf{x}^*\| <\\ < L^2\|\mathbf{x^{(n-2)}} - \mathbf{x}^*\| < \cdots < L^n\|\mathbf{x^{(0)}} - \mathbf{x}^*\| < L^n\delta < \delta,\end{aligned}$$

consecuentemente $\mathbf{g}(\mathbf{x})$ es una contracción en $\overline{B(\mathbf{x}^*, \delta)}$.

Por otra parte, al ser $L < 1$ bastará con escoger $N \in \mathbb{R}$ suficientemente grande para que todos los elementos de la sucesión con índice $n > N$ sean tan cercanos a $\mathbf{x}^*$ como se desee. En otros términos, $\mathbf{x}^* = \lim_{n\to\infty} \mathbf{x^{(n)}}$.

□

Notemos que los resultados de convergencia local, como el que acabamos de estudiar, tienen el inconveniente de que necesitamos conocer a priori la existencia de una solución, además de exigir una condición sobre ella para asegurar la convergencia.

3.2. Método de Newton

Considérese nuevamente el sistema de n ecuaciones no lineales con n incógnitas representado por

$$\mathbf{f}(\mathbf{x}) = \mathbf{0} \Leftrightarrow \left\{\begin{array}{c} f_1(x_1, x_2, \ldots, x_n) = 0 \\ f_2(x_1, x_2, \ldots, x_n) = 0 \\ \vdots \\ f_n(x_1, x_2, \ldots, x_n) = 0 \end{array}\right.$$

Supongamos que en un dominio cerrado $D \subset \mathbb{R}^n$, $\mathbf{f}(\mathbf{x})$ es una función de clase $C^{(2}(D)$ y supongamos, además, que la ecuación anterior admite una solución $\mathbf{x}^*$ en el dominio D. Para cualquier otro vector $\mathbf{x^{(0)}} \in D$, denotando por $\delta\mathbf{x}$ al vector tal que $\mathbf{x}^* = \mathbf{x^{(0)}} + \delta\mathbf{x}$, la expresión del desarrollo en serie de Taylor [3] nos permitiría afirmar, para cada una de las ecuaciones del sistema, que existen los valores $\theta_j \in [0, 1]$, $(j = 1, 2, \ldots, n)$, tales que:

$$\begin{aligned}0 = f_j(\mathbf{x}^*) = f_j(\mathbf{x^{(0)}} + \delta\mathbf{x}) =\\ = f_j(\mathbf{x^{(0)}}) + \left\{\nabla f_j(\mathbf{x^{(0)}})\right\}^T \delta\mathbf{x} + \frac{1}{2}\left\{\delta\mathbf{x}\right\}^T \left(H_{f_j}(\mathbf{x^{(0)}} + \theta_j\delta\mathbf{x})\right)\delta\mathbf{x},\end{aligned}$$

donde $\nabla f_j(\mathbf{x})$ es el gradiende de la función $f_j(\mathbf{x})$ y $H_{f_j}(\mathbf{x})$ es la matriz Hessiana de la función $f_j(\mathbf{x})$, la cual se define como

$$H_{f_j}(\mathbf{x}) = \begin{pmatrix} \frac{\partial^2 f_j}{\partial x_1^2} & \frac{\partial^2 f_j}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f_j}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f_j}{\partial x_2 \partial x_1} & \frac{\partial^2 f_j}{\partial x_2^2} & \cdots & \frac{\partial^2 f_j}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial^2 f_j}{\partial x_n \partial x_1} & \frac{\partial^2 f_j}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f_j}{\partial x_n^2} \end{pmatrix}.$$

Si conocido $\mathbf{x^{(0)}}$ se fuese capaz de determinar $\delta\mathbf{x}$ resolviendo el sistema formado para $j = 1, 2, \ldots, n$ por las ecuaciones

$$f_j(\mathbf{x^{(0)}} + \delta\mathbf{x}) = f_j(\mathbf{x^{(0)}}) + \left\{\nabla f_j(\mathbf{x^{(0)}})\right\}^T \delta\mathbf{x} + \frac{1}{2}\left\{\delta\mathbf{x}\right\}^T \left(H_{f_j}(\mathbf{x^{(0)}} + \theta_j \delta\mathbf{x})\right)\delta\mathbf{x} = 0$$

podría determinarse $\mathbf{x}^*$ como $\mathbf{x}^* = \mathbf{x^{(0)}} + \delta\mathbf{x}$. Pero para resolver este sistema primero deberíamos conocer los valores de θ_j (lo cual no es obvio) y, una vez conocidos, resolver un sistema, en general no lineal ya que $\delta\mathbf{x}$ interviene en la expresión de las matrices Hessianas $\left(H_{f_j}(\mathbf{x^{(0)}} + \theta_j \cdot \delta\mathbf{x})\right)$. Por tanto, salvo en situaciones muy particulares, no se ganaría gran cosa reemplazando el problema de resolver $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ por el de resolver el sistema anterior.

El método de Newton se sustenta en simplificar las expresiones anteriores linealizándolas. Para ello considera que si se está suficientemente cerca de la solución (es decir, si $\|\delta\mathbf{x}\|$ es suficientemente pequeño) los términos

$$\frac{1}{2}\left\{\delta\mathbf{x}\right\}^T \left(H_{f_j}(\mathbf{x^{(0)}} + \theta_j \delta\mathbf{x})\right)\delta\mathbf{x}$$

podrán despreciarse frente a otros términos de cada ecuación del sistema. Por ello en este método se resuelve el sistema lineal

$$\mathbf{f}(\mathbf{x^{(0)}}) + J_{\mathbf{f}}(\mathbf{x^{(0)}})\Delta\mathbf{x^{(0)}} = 0,$$

donde $\Delta\mathbf{x^{(0)}}$ representa la variación o cambio de valor de $\mathbf{x^{(0)}}$, del que se obtiene que

$$\Delta\mathbf{x^{(0)}} = J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\mathbf{x^{(0)}}).$$

Obviamente, al ser diferente el sistema linealizado que el proporcionado por el desarrollo de Taylor, se tendrá que $\Delta\mathbf{x^{(0)}} \neq \delta\mathbf{x}$ y por tanto

$$\mathbf{x}^* = \mathbf{x^{(0)}} + \delta\mathbf{x} \neq \mathbf{x^{(1)}} = \mathbf{x^{(0)}} + \Delta\mathbf{x^{(0)}}.$$

De una forma intuitiva puede pensarse que aunque $\mathbf{x^{(1)}}$ sea diferente de $\mathbf{x}^*$, será un vector más próximo a $\mathbf{x}^*$ que $\mathbf{x^{(0)}}$ pues lo hemos obtenido aproximando el valor $\delta\mathbf{x}$ que nos llevaba de $\mathbf{x^{(0)}}$ a $\mathbf{x}^*$. De esta forma el método de Newton propone repetir este proceso de forma recursiva hasta estar lo suficientemente cercanos a la solución buscada. Más concretamente el método de Newton consiste en:

Dado un vector $\mathbf{x^{(0)}}$, generar la sucesión

$$\left\{\mathbf{x^{(n+1)}} = \mathbf{x^{(n)}} - J_{\mathbf{f}}\left(\mathbf{x^{(n)}}\right)^{-1}\mathbf{f}\left(\mathbf{x^{(n)}}\right)\right\}_{n=0}^{\infty}$$

Sobre este método, en primer lugar, puede observarse que si denotamos por

$$\mathbf{g}(\mathbf{x}) = \mathbf{x} - J_{\mathbf{f}}(\mathbf{x})^{-1}\mathbf{f}(\mathbf{x})$$

estamos en presencia de un caso particular del método de aproximaciones sucesivas. En otros términos, se tiene la siguiente propiedad:

Proposición 3.4 *Sea D cerrado de $\mathbb{R}^n$ y sea $\mathbf{f} : D \longrightarrow D$ una aplicación de clase $C^{(1}(D)$. Si la función $\mathbf{g}(\mathbf{x}) = \mathbf{x} - J_{\mathbf{f}}(\mathbf{x})^{-1}\mathbf{f}(\mathbf{x})$ es una contracción definida en D, la sucesión $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ dada por*

$$\mathbf{x^{(n+1)}} = \mathbf{x^{(n)}} - J_{\mathbf{f}}(\mathbf{x^{(n)}})^{-1}\mathbf{f}(\mathbf{x^{(n)}}),$$

obtenida a partir de cualquier $\mathbf{x^{(0)}} \in D$, converge hacia la única solución $\mathbf{x}^*$ de la ecuación $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ en D.

Demostración.
Es un caso particular del Teorema 3.1 de convergencia del método de aproximaciones sucesivas. $\square$

Lema 3.5 *Sean D cerrado y convexo de $\mathbb{R}^n$ y $\mathbf{f} : D \longrightarrow D$ una aplicación de clase $C^{(1}(D)$. Si existe una constante $L > 0$ tal que se verifica*

$$\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\| \quad \forall\, \mathbf{x}, \mathbf{y} \in D$$

entonces se verifica también que

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) - J_{\mathbf{f}}(\mathbf{y})(\mathbf{x} - \mathbf{y})\| \leqslant \frac{L}{2}\|\mathbf{x} - \mathbf{y}\|^2 \quad \forall\, \mathbf{x}, \mathbf{y} \in D.$$

Demostración.
Siendo $\mathbf{x}$ e $\mathbf{y}$ dos vectores genéricos de D denotemos por $\mathbf{q}(t)$ a la función vectorial dependiente de un único parámetro real definida por

$$\mathbf{q}(t) = \mathbf{f}(\mathbf{y} + t(\mathbf{x} - \mathbf{y})).$$

Esta función, habida cuenta de las hipótesis realizadas sobre $\mathbf{f}$ es derivable $\forall\, t \in [0, 1]$. Así, denotando por $\mathbf{z} = \mathbf{y} + t(\mathbf{x} - \mathbf{y})$ se tiene que

$$\mathbf{q}'(t) = J_{\mathbf{f}}(\mathbf{z})(\mathbf{x} - \mathbf{y}),$$

de donde

$$\|\mathbf{q}'(t) - \mathbf{q}'(0)\| = \|J_{\mathbf{f}}(\mathbf{z})(\mathbf{x} - \mathbf{y}) - J_{\mathbf{f}}(\mathbf{y})(\mathbf{x} - \mathbf{y})\| \leqslant \|J_{\mathbf{f}}(\mathbf{z}) - J_{\mathbf{f}}(\mathbf{y})\|\|\mathbf{x} - \mathbf{y}\| =$$

$$= \|J_{\mathbf{f}}(\mathbf{y}+t(\mathbf{x}-\mathbf{y})) - J_{\mathbf{f}}(\mathbf{y})\|\|\mathbf{x}-\mathbf{y}\| \leqslant Lt\|\mathbf{x}-\mathbf{y}\|^2.$$

Esta desigualdad, a su vez, puede utilizarse en el proceso siguiente

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y}) - J_{\mathbf{f}}(\mathbf{y})(\mathbf{x}-\mathbf{y})\| = \|\mathbf{q}(1) - \mathbf{q}(0) - \mathbf{q}'(0)\| =$$

$$= \left\|\int_0^1 (\mathbf{q}'(t) - \mathbf{q}'(0))\, dt\right\| \leqslant \int_0^1 \|\mathbf{q}'(t) - \mathbf{q}'(0)\|\, dt \leqslant \int_0^1 Lt\|\mathbf{x}-\mathbf{y}\|^2\, dt = \frac{L}{2}\|\mathbf{x}-\mathbf{y}\|^2.$$

□

Observación 3.1 *El que se verifique la hipótesis del lema precedente*

$$\exists\, L \in \mathbb{R}^+ \quad \textit{tal que} \quad \|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x}-\mathbf{y}\| \quad \forall\, \mathbf{x}, \mathbf{y} \in D$$

se expresa diciendo que ***la matriz Jacobiana es Lipschitziana de razón*** L.

Con ayuda de este lema puede procederse a presentar y demostrar el siguiente teorema:

Teorema 3.6 *Sean D cerrado y convexo de $\mathbb{R}^n$ y sea $\mathbf{f} : D \longrightarrow D$ una aplicación de clase $C^{(1}(D)$ para la que se verifican las dos hipótesis siguientes:*

i) Existe $\beta \in \mathbb{R}^+$ tal que $\|J_{\mathbf{f}}(\mathbf{x})^{-1}\| < \beta$, para $\mathbf{x} \in D$.

ii) Existe $L \in \mathbb{R}^+$ tal que $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x}-\mathbf{y}\|$ $\forall\, \mathbf{x}, \mathbf{y}, \in D$.

Entonces la sucesión $\left\{\mathbf{x}^{(\mathbf{n})}\right\}_{n=0}^{\infty}$ dada por

$$\mathbf{x}^{(\mathbf{n+1})} = \mathbf{x}^{(\mathbf{n})} - J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{n})})^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{n})})$$

obtenida a partir de cualquier $\mathbf{x}^{(\mathbf{0})} \in D$ verifica que

$$\|\mathbf{x}^{(\mathbf{n+1})} - \mathbf{x}^{(\mathbf{n})}\| \leqslant \frac{L\beta}{2}\|\mathbf{x}^{(\mathbf{n})} - \mathbf{x}^{(\mathbf{n-1})}\|^2.$$

Demostración.
Se tiene que

$$\|\mathbf{x}^{(\mathbf{n+1})} - \mathbf{x}^{(\mathbf{n})}\| = \|-J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{n})})^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{n})})\| \leqslant \|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{n})})^{-1}\|\|\mathbf{f}(\mathbf{x}^{(\mathbf{n})})\| \leqslant \beta\|\mathbf{f}(\mathbf{x}^{(\mathbf{n})})\|$$

y como de

$$\mathbf{x}^{(\mathbf{n+1})} = \mathbf{x}^{(\mathbf{n})} - J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{n})})^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{n})})$$

se tiene que

$$\mathbf{f}(\mathbf{x}^{(\mathbf{n})}) = -J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{n})})(\mathbf{x}^{(\mathbf{n+1})} - \mathbf{x}^{(\mathbf{n})}),$$

luego, utilizando el lema precedente, se sigue

$$\|\mathbf{x}^{(\mathbf{n+1})} - \mathbf{x}^{(\mathbf{n})}\| \leqslant \beta\|\mathbf{f}(\mathbf{x}^{(\mathbf{n})})\| =$$

$$= \beta\|\mathbf{f}(\mathbf{x}^{(\mathbf{n})}) - \mathbf{f}(\mathbf{x}^{(\mathbf{n})}) - J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{n})})(\mathbf{x}^{(\mathbf{n+1})} - \mathbf{x}^{(\mathbf{n})})\| \leqslant$$

$$\leqslant \frac{L\beta}{2}\|\mathbf{x^{(n)}} - \mathbf{x^{(n-1)}}\|^2.$$

□

El teorema anterior nos muestra que la relación entre la norma del vector diferencia entre las aproximaciones halladas en las iteraciones $(n+1)$ y n es proporcional (con factor $C = \frac{L\beta}{2}$) al cuadrado de la norma del vector diferencia entre las aproximaciones n y $(n+1)$. Pero por sí solo este teorema no nos justifica que el método converja. Para acabar de obtener un resultado que garantice la convergencia del método es necesario imponer más condiciones, las cuales analizamos en el siguiente capítulo.

3.2.1. Programación del método de Newton

El método de Newton

$$\mathbf{x^{(n+1)}} = \mathbf{x^{(n)}} - J_{\mathbf{f}}(\mathbf{x^{(n)}})^{-1}\mathbf{f}(\mathbf{x^{(n)}})$$

exige calcular la matriz inversa de $J_{\mathbf{f}}(\mathbf{x^{(n)}})$ en cada iteración. Ello, en el caso de que el número de ecuaciones sea elevado, requiere un gran esfuerzo computacional. Por este motivo, a la hora de programar el método, resulta conveniente calcular,en su lugar, en cada paso la expresión equivalente

$$J_{\mathbf{f}}(\mathbf{x^{(n)}})(\mathbf{x^{(n)}} - \mathbf{x^{(n+1)}}) = \mathbf{f}\left(\mathbf{x^{(n)}}\right).$$

De esta forma, en cada iteración el valor del vector $\mathbf{x^{(n+1)}}$ se obtiene mediante la resolución de un sistema de ecuaciones lineal, reduciendo significativamente el coste computacional.

A continuación se muestra un ejemplo de programación del método de Newton realizado con *Mathematica*. En concreto, se implementa la función *Newton*$(f, x, x0, tol)$ donde las variables de entrada son:

- $f \to$ vector cuyas componentes son las funciones componentes $f_i(\mathbf{x})$, $i = 1, \ldots n$,
- $x \to$ vector cuyas componentes son las incógnitas x_i, $i = 1, \ldots, n$,
- $x0 \to$ vector inicial $\mathbf{x^{(0)}}$,
- $tol \to$ valor numérico que indica el criterio de parada deseado: $\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| < tol$.

```
Newton[f_, x_, x0_, tol_] :=
  Module[{u = x, n, Jac, dif = 2 tol, k = 0, A, B, a, b, res = x0},

   n = Length[f];
   Jac = Table[D[f[[i]], u[[j]]], {i, n}, {j, n}];

   While[dif > tol,
     A = Jac /. Table[u[[i]] → res[[i]], {i, n}];
     If[Det[A] == 0,
      Print["La matriz Jacobiana asociada al sistema es singular."],
      B = Table[f[[i]] /. Table[u[[j]] → res[[j]], {j, n}], {i, n}];
      a = NSolve[A.Table[res[[i]] - u[[i]], {i, n}] == B, Table[u[[i]], {i, n}],
         Reals][[1]];
      b = Table[u[[i]], {i, n}] /. a;
      dif = Norm[res - b, Infinity];
      res = b;
      k++;]]

    Print[res]; Print["Número de iteraciones: ", k]];
```

Capítulo 4

Convergencia semilocal del método de Newton

4.1. Teorema de Kantorovich

En este capítulo presentamos, a partir de los resultados desarrollados por Kantorovich, un resultado de convergencia semilocal para el método de Newton [5]. La característica más importante de estos resultados es que se exigen condiciones al punto de partida pero no a la solución, de forma que no sólo se ocupan de garantizar la convergencia del método, sino que también aportan resultados sobre la existencia de soluciones de la ecuación. A diferencia de la convergencia global, obtendremos condiciones para que un punto determinado $\mathbf{x}^{(\mathbf{0})}$ pueda asegurar la convergencia de la sucesión dada por el método de Newton.

En la versión clásica del teorema de Kantorovich (adaptada al espacio $\mathbb{R}^n$), para asegurar la convergencia del método de Newton, empezando en $\mathbf{x}^{(\mathbf{0})} \in \mathbb{R}^n$, a una solución $\mathbf{x}^* \in \mathbb{R}^n$ de un sistema de ecuaciones no lineales $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, deben verificarse una serie de condiciones, siendo las dos primeras:

i) Existe $J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}$ y es tal que $\|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}\| \leqslant \beta$,

ii) $\|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}\mathbf{f}(\mathbf{x_0})\| \leqslant \eta$.

con $\beta, \eta \in \mathbb{R}^+$, que son las condiciones sobre el punto de salida para el método de Newton.

La tercera de ellas exige condiciones sobre la segunda derivada de $\mathbf{f}$, de forma que no sólo hace necesario que la función $\mathbf{f}$ sea de clase $C^{(2}$ en su dominio Ω sino que además, en la práctica, resulta difícil de comprobar. Notemos también que la matriz Hessiana no aparece en la expresión del método de Newton, por lo que su cálculo resulta de poca utilidad. Esta condición puede suavizarse, siendo sustituida por la siguiente:

$$\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\| \qquad \forall\, \mathbf{x}, \mathbf{y} \in \Omega \subset \mathbb{R}^n,$$

esto es, debe verificarse que la matriz Jacobiana sea Lipschitziana de razón L para la norma considerada.

Por último, añadiremos una cuarta condición que relacionará los parámetros introducidos, β, η y L para asegurar la convergencia.

De esta forma, obtenemos la siguiente versión del teorema de Kantorovich:

Teorema 4.1 *: Sea Ω abierto y convexo de $\mathbb{R}^n$ y sea $\mathbf{f} : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación de clase $C^{(1}(\Omega)$. Supongamos que $J_{\mathbf{f}}(\mathbf{x^0})^{-1}$ existe en algún $\mathbf{x^{(0)}} \in \Omega$ y que se verifican las condiciones:*

i) $\|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\| \leqslant \beta$,

ii) $\|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\mathbf{x_0})\| \leqslant \eta$,

iii) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad \forall\, \mathbf{x}, \mathbf{y} \in \Omega.$

Si $a^{(0)} = L\beta\eta \leqslant 1/2$ y $B(\mathbf{x^{(0)}}, R) \in \Omega$, donde $R = \frac{2\left(1-a^{(0)}\right)}{2-3a^{(0)}}\eta$, la sucesión de Newton $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ dada por

$$\mathbf{x^{(n+1)}} = \mathbf{x^{(n)}} - J_{\mathbf{f}}\left(\mathbf{x^{(n)}}\right)^{-1} \mathbf{f}\left(\mathbf{x^{(n)}}\right),$$

empezando en $\mathbf{x^{(0)}}$, converge a una solución $\mathbf{x}^*$ del sistema de ecuaciones $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, la solución x^* y las iteraciones $\mathbf{x^{(n)}}$ pertenecen a $\overline{B(\mathbf{x^{(0)}}, R)}$ y la solución $\mathbf{x}^*$ es única en la región $B(\mathbf{x^{(0)}}, \eta/a^{(0)}) \cap \Omega$.

La demostración de este resultado de convergencia semilocal se lleva a cabo mediante la construcción, a partir de $a^{(0)} = L\beta\eta > 0$, de la sucesión escalar

$$a^{(n+1)} = a^{(n)} h\left(a^{(n)}\right)^2 g\left(a^{(n)}\right), \qquad n \geqslant 0, \tag{4.1}$$

donde

$$h(x) = \frac{1}{1-x} \qquad \text{y} \qquad g(x) = \frac{x}{2}. \tag{4.2}$$

Como veremos más adelante, esta sucesión verifica un sistema de dos relaciones de recurrencia cuyo análisis garantiza la convergencia de la sucesión de Newton $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$. Previamente, vamos a presentar algunos resultados sobre la sucesión (4.1) que nos serán de utilidad para el desarrollo de la demostración del teorema.

Lema 4.2 *Sean h y g las funciones reales dadas en (4.2). Entonces*

a) h *es creciente y* $h(x) > 1$ *en* $(0, 1)$,

b) g *es creciente,*

c) para $\gamma \in (0, 1)$, *tenemos* $h(\gamma x) < h(x)$ *si* $x > 0$ *y* $g(\gamma x) = \gamma g(x)$.

Demostración.

a) La función h es creciente puesto que $h'(x) = \frac{1}{(1-x)^2} > 0 \ \forall x \in \mathbb{R}$. Además, $x \in (0,1) \Rightarrow h(x) = \frac{1}{1-x} \in (1, \infty)$.

b) Del mismo modo, la función g es creciente ya que $g'(x) = 1/2 > 0 \ \forall x \in \mathbb{R}$.

c) Si $x > 0$, se tiene que $h(\gamma x) = \frac{1}{1-\gamma x} < \frac{1}{1-x} = h(x)$, ya que $\gamma \in (0,1)$.
Por último, $g(\gamma x) = \frac{\gamma x}{2} = \gamma g(x)$.

$\square$

Lema 4.3 *Sean h y g las funciones reales dadas en (4.2).*
Si $a^{(0)} \in (0, 1/2)$, entonces

a) $h\left(a^{(0)}\right)^2 g\left(a^{(0)}\right) < 1$,

b) la sucesión $\left\{a^{(n)}\right\}$ es estrictamente decreciente y $a^{(n)} < 1$ para todo $n \geqslant 1$.

Si $a^{(0)} = 1/2$, entonces $a^{(n)} = a^{(0)} < 1$, $\forall n \geqslant 1$.

Demostración.
En primer lugar, consideramos el caso $a^{(0)} \in (1, 1/2)$. El apartado $a)$ se sigue de forma inmediata. Demostramos el caso b) por inducción sobre n. Como $h\left(a^{(0)}\right)^2 g\left(a^{(0)}\right) < 1$, es claro que $a^{(1)} < a^{(0)}$. Suponemos ahora que $a^{(j)} < a^{(j-i)}$ para $j = 1, 2, \dots n$. Entonces,

$$a^{(n+1)} = a^{(n)} h\left(a^{(n)}\right)^2 g\left(a^{(n)}\right) < a^{(n)} h\left(a^{(0)}\right)^2 g\left(a^{(0)}\right) < a^{(n)},$$

puesto que h y g son crecientes. Por tanto, la sucesión $\left\{a^{(n)}\right\}$ es estrictamente decreciente. Y, en consecuencia, $a^{(n)} < 1, \forall n \geqslant 1$.
Si, por otra parte, $a^{(0)} = 1/2$, entonces $h\left(a^{(0)}\right)^2 g\left(a^{(0)}\right) = 1$ y, por lo tanto, $a^{(n)} = a^{(0)} = 1/2 < 1$, $\forall n \geqslant 0$. $\square$

Lema 4.4 *Sean h y g las funciones reales dadas en (4.2).*
Si $a^{(0)} \in (0, 1/2)$, definimos $\gamma = \frac{a^{(1)}}{a^{(0)}}$, y entonces

a) $a^{(n)} < \gamma^{2^{n-1}} a^{(n-1)}$ *y* $a^{(n)} < \gamma^{2^n - 1} a^{(0)}$, $\forall n \geqslant 2$.

b) $h(a^{(n)}) g(a^{(n)}) < \gamma^{2^n - 1} h(a^{(0)}) g(a^{(0)}) = \gamma^{2^n} / h(a^{(0)})$, $\forall n \geqslant 1$.

Si $a^{(0)} = 1/2$, entonces $h(a^{(n)}) g(a^{(n)}) = h(a^{(0)}) g(a^{(0)}) = 1/h(a^{(0)})$, $\forall n \geqslant 1$.

Demostración.
Probamos a) mediante proceso de inducción. Si $n = 2$, aplicando el apartado b) del Lema 4.2, obtenemos

$$a^{(2)} = a^{(1)} h(a^{(1)})^2 g(a^{(1)}) = \gamma a^{(0)} h(a^{(0)})^2 g(\gamma a^{(0)}) < \gamma^2 a^{(1)} = \gamma^3 a^{(0)}.$$

Supongamos ahora que

$$a^{(n-1)} < \gamma^{2^{n-2}} a^{(n-2)} < \gamma^{2^{n-1}-1} a^{(0)}.$$

Entonces, por el mismo razonamiento,

$$a^{(n)} = a^{(n-1)} h(a^{(n-1)})^2 g(a^{(n-1)}) < \gamma^{2^{n-2}} a^{(n-2)} h\left(\gamma^{2^{n-2}} a^{(n-2)}\right)^2 g\left(\gamma^{2^{n-2} a^{(n-2)}}\right) <$$

$$< \gamma^{2^{n-1}} a^{(n-1)} < \gamma^{2^{n-1}} \gamma^{2^{n-2}} < \cdots < \gamma^{2^n-1} a^{(0)}.$$

Para probar b), observamos que

$$f(a^{(n)}) g(a^{(n)}) < h\left(\gamma^{2^n-1} a^{(0)}\right) g\left(\gamma^{2^n-1}\right) < \gamma^{2^n-1} h(a^{(0)}) g(a^{(0)}) = \gamma^{2^n}/h(a^{(0)}),\ n \geqslant 1,$$

con lo que se completa la demostración. □

Relaciones de recurrencia:

A continuación, definimos las dos relaciones de recurrencia siguientes:

(I) $\|\Gamma_n\| = \|J_{\mathbf{f}}(\mathbf{x^{(n)}})^{-1}\| \leqslant h(a^{(n-1)})\|\Gamma_{n-1}\|$,

(II) $\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \leqslant h(a^{(n-1)}) g(a^{(n-1)}) \|\mathbf{x^{(n)}} - \mathbf{x^{(n-1)}}\|$.

y probamos que se cumplen $\forall\, n \geqslant 0$ mediante inducción sobre n.

Para ello, suponemos que $\mathbf{x^{(n)}} \in \Omega\ \forall\ n \geqslant 0$ (lo probaremos más adelante). Entonces teniendo en cuenta que, por el Lema 4.3 (b), $a^{(n)} < 1\ \forall n \geqslant 0$,

$$\|\mathbf{I} - \Gamma_0 J_{\mathbf{f}}(x^{(1)})\| \leqslant \|\Gamma_0\| \|J_{\mathbf{f}}(\mathbf{x^{(0)}}) - J_{\mathbf{f}}(\mathbf{x^{(1)}})\| \leqslant L\beta \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| \leqslant L\beta\eta = a^{(0)} < 1.$$

Luego, aplicando el Lema 2.5, existe $\Gamma_1 = J_{\mathbf{f}}(\mathbf{x^{(1)}})^{-1}$ y

$$\|\Gamma_1\| \leqslant \frac{\|\Gamma_0\|}{1 - \|\mathbf{I} - \Gamma_0 J_{\mathbf{f}}(\mathbf{x^{(1)}})\|} \leqslant h(a^{(0)}) \|\Gamma_0\|.$$

A partir de la fórmula de Taylor [3], se obtiene

$$\mathbf{f}(\mathbf{x^{(1)}}) = \mathbf{f}(\mathbf{x^{(0)}}) + \int_0^1 J_{\mathbf{f}}(\mathbf{x^{(0)}} + t(\mathbf{x^{(1)}} - \mathbf{x^{(0)}}))(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})\, dt,$$

y, como por la fórmula de Newton

$$\mathbf{x^{(1)}} = \mathbf{x^{(0)}} - J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1} \mathbf{f}(\mathbf{x^{(0)}}) \Rightarrow \mathbf{f}(\mathbf{x^{(0)}}) = -J_{\mathbf{f}}(\mathbf{x^{(0)}})(\mathbf{x^{(1)}} - \mathbf{x^{(0)}}),$$

se tiene que

$$\mathbf{f}(\mathbf{x^{(1)}}) = -\int_0^1 J_{\mathbf{f}}(\mathbf{x^{(0)}})(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})\, dt + \int_0^1 J_{\mathbf{f}}(\mathbf{x^{(0)}} + t(\mathbf{x^{(1)}} - \mathbf{x^{(0)}}))(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})\, dt =$$

$$= \int_0^1 (J_{\mathbf{f}}(\mathbf{x^{(0)}} + t(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})) - J_{\mathbf{f}}(\mathbf{x^{(0)}}))(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})\, dt$$

Por lo tanto,

$$\|\mathbf{f}(\mathbf{x^{(1)}})\| \leqslant \int_0^1 \|(J_\mathbf{f}(\mathbf{x^{(0)}} + t(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})) - J_\mathbf{f}(\mathbf{x^{(0)}}))(\mathbf{x^{(1)}} - \mathbf{x^{(0)}})\| dt \leqslant$$

$$\leqslant \int_0^1 L\, t \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|^2 dt = \frac{L}{2}\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|^2 \leqslant \frac{L\eta}{2}\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|.$$

Luego

$$\|\mathbf{x^{(2)}} - \mathbf{x^{(1)}}\| = \|\Gamma_1 \mathbf{f}(\mathbf{x^{(1)}})\| \leqslant \|\Gamma_1\|\|\mathbf{f}(\mathbf{x^{(1)}})\| \leqslant h(a^{(0)})g(a^{(0)})\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|.$$

Si suponemos ahora que las relaciones (I) y (II) se verifican para un cierto $n \geqslant 1$, se sigue de forma análoga al caso $n = 1$ que (I) y (II) también se cumplen para $n + 1$. En consecuencia, queda probado por inducción que (I) y (II) se verifican para cualquier $n \in \mathbb{N}$.

Con los resultados obtenidos hasta ahora, nos encontramos en condiciones de demostrar el Teorema 4.1. Notemos que para ello bastará con probar que $\mathbf{x^{(n)}} \in \Omega$ y que la sucesión generada por el método de Newton es de Cauchy.

Demostración del Teorema 4.1.
Probamos en primer lugar que $\mathbf{x^{(n)}} \in B(\mathbf{x^{(0)}}, R)$. Observamos que

$$\|\mathbf{x^{(n)}} - \mathbf{x^{(0)}}\| \leqslant \sum_{i=0}^{n-1}\|\mathbf{x^{(i+1)}} - \mathbf{x^{(i)}}\| \overset{(II)}{\leqslant} \left(\sum_{i=0}^{n-2}\left(\prod_{j=0}^{i} h(a^{(j)})g(a^{(j)})\right)\right)\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| <$$

$$\overset{\text{Lema 4.4 (b)}}{<} \left[1 + \sum_{i=0}^{n-2}\left(\prod_{j=0}^{i} \gamma^{2^j-1} h(a^{(0)})g(a^{(0)})\right)\right]\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|$$

donde $\gamma = \frac{a^{(1)}}{a^{(0)}} < 1$ y $\Delta = h(a^{(0)})g(a^{(0)})/\gamma = 1/h(a^{(0)}) = 1 - a^{(0)} < 1$.

Ahora, por la desigualdad de Bernoulli[1], se sigue $\gamma^{2^{1+i}-1} = \gamma^{2(2^i-1)+1} \leqslant \gamma^{2i+1}$, y en consecuencia

$$\|\mathbf{x^{(n)}} - \mathbf{x^{(0)}}\| < \left(1 + \gamma\Delta\sum_{i=0}^{n-2}\gamma^{2i}\Delta^i\right)\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| < \left(1 + \gamma\Delta\frac{1-(\gamma^2\Delta)^{n-1}}{1-\gamma^2\Delta}\right)\eta <$$

$$< \frac{\eta}{1-\gamma\Delta} = \frac{2(1-a^{(0)})}{2-3a^{(0)}}\eta = R.$$

Por tanto, $\mathbf{x^{(n)}} \in B(\mathbf{x^{(0)}}, R)$ y, como $B(\mathbf{x^{(0)}}, R) \subset \Omega$, entonces $\mathbf{x^{(n)}} \in \Omega\ \forall\, n \geqslant 0$.

Probamos ahora que $\left\{\mathbf{x^{(n)}}\right\}_{n=0}^{\infty}$ es una sucesión de Cauchy, procediendo de manera análoga al apartado anterior. Así, para $m \geqslant 1$ y $n \geqslant 1$,

$$\|\mathbf{x^{(n+m)}} - \mathbf{x^{(n)}}\| \leqslant \sum_{i=n}^{n+m-1}\|\mathbf{x^{(i+1)}} - \mathbf{x^{(i)}}\| \overset{(II)}{\leqslant} \left(\sum_{i=n}^{n+m-2}\left(\prod_{j=n}^{i} h(a^{(j)})g(a^{(j)})\right)\right)\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \leqslant$$

[1] Desigualdad de Bernoulli: $(1+z)^n - 1 \geqslant nz$ si $z > -1$.

$$\overset{(II)}{\leqslant} \sum_{i=n-1}^{n+m-2} \left(\prod_{j=0}^{i} h(a^{(j)}) g(a^{(j)}) \right) \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| <$$

$$\overset{\text{Lema 4.4 (b)}}{<} \left(\sum_{i=n-1}^{n+m-2} \left(\prod_{j=0}^{i} \gamma^{2^j-1} h(a^{(0)}) g(a^{(0)}) \right) \right) \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| =$$

$$= \sum_{i=n-1}^{n+m-2} \left(\prod_{j=0}^{i} \left(\gamma^{2^j} \Delta \right) \right) \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| = \sum_{i=n-1}^{n+m-2} \left(\gamma^{2^{1+i}-1} \Delta^{1+i} \right) \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| =$$

$$= \sum_{i=0}^{m-1} \left(\gamma^{2^{n+i}-1} \Delta^{n+i} \right) \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|.$$

Ahora, de nuevo por la desigualdad de Bernoulli, $\gamma^{2^{n+i}-1} = \gamma^{2^n-1}\gamma^{2^n(2^i-1)} \leqslant \gamma^{2^n-1}\gamma^{2^n i}$, y por tanto

$$\|\mathbf{x^{(n+m)}} - \mathbf{x^{(n)}}\| < \left(\sum_{i=0}^{m-1} \left(\gamma^{2^n i} \Delta^i \right) \right) \gamma^{2^n-1} \Delta^n \eta < \frac{1 - \left(\gamma^{2^n} \Delta \right)^m}{1 - \gamma^{2^n} \Delta} \gamma^{2^n-1} \Delta^n \eta. \quad (4.3)$$

Luego la sucesión $\left\{ \mathbf{x^{(n)}} \right\}_{n=0}^{\infty}$ es de Cauchy.

Veamos ahora que el límite de la sucesión dada por el método de Newton es solución de $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.
Hemos probado que $\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\|$ tiende a cero cuando $n \to \infty$. Teniendo en cuenta que

$$\|\mathbf{f}(\mathbf{x^{(n)}})\| = \|-J_{\mathbf{f}}(\mathbf{x^{(n)}})(\mathbf{x^{(n-1)}} - \mathbf{x^{(n)}})\| \leqslant \|J_{\mathbf{f}}(\mathbf{x^{(n)}})\| \|\mathbf{x^{(n-1)}} - \mathbf{x^{(n)}}\|$$

y que la sucesión $\left\{ \|J_{\mathbf{f}}(\mathbf{x^{(n)}})\| \right\}$ está acotada, puesto que

$$\|J_{\mathbf{f}}(\mathbf{x^{(n)}})\| \leqslant \|J_{\mathbf{f}}(\mathbf{x^{(n)}}) - J_{\mathbf{f}}(\mathbf{x^{(0)}})\| + \|J_{\mathbf{f}}(\mathbf{x^{(0)}})\| \leqslant$$

$$\leqslant \|J_{\mathbf{f}}(\mathbf{x^{(0)}})\| + L\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| < \|J_{\mathbf{f}}(\mathbf{x^{(0)}})\| + LR,$$

se cumple que $\|\mathbf{f}(\mathbf{x^{(n)}})\| \to 0$ cuando $n \to \infty$. Por lo tanto, si $\mathbf{x}^* = \lim \mathbf{x^{(n)}}$, entonces $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$ por la continuidad de $\mathbf{f}$ en $B(\mathbf{x^{(0)}}, R)$.

Para probar la unicidad, suponemos que $\mathbf{z}^*$ es otra solución de $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ en la región $B(\mathbf{x^{(0)}}, \eta/a^{(0)}) \cap \Omega$. Podemos entonces realizar la aproximación

$$0 = \Gamma_0(\mathbf{f}(\mathbf{z}^*) - \mathbf{f}(\mathbf{x}^*)) = \left(\int_0^1 \Gamma_0 J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*)) dt \right) (\mathbf{z}^* - \mathbf{x}^*) = \mathbf{P}(\mathbf{z}^* - \mathbf{x}^*).$$

y probar que existe el inverso de $\mathbf{P} = \int_0^1 \Gamma_0 J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*)) dt$, para lo cual, por el Lema 2.4, bastará con verificar que $\|\mathbf{I} - \mathbf{P}\| < 1$. En efecto,

$$\|\mathbf{I} - \mathbf{P}\| = \|\mathbf{I} - \Gamma_0 \int_0^1 J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))\| \, dt =$$

$$= \|\Gamma_0 \left(J_{\mathbf{f}}(\mathbf{x^{(0)}}) - \int_0^1 J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))\, dt \right) \| \leqslant$$

$$\leqslant \|\Gamma_0\| \int_0^1 \|J_{\mathbf{f}}(\mathbf{x^{(0)}}) - J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))\|\, dt \leqslant \beta L \int_0^1 \|\mathbf{x^{(0)}} - \mathbf{x}^* + t(\mathbf{x}^* - \mathbf{z}^*)\|\, dt =$$

$$= \frac{\beta L}{2} \left(\|(\mathbf{x^{(0)}} - \mathbf{x}^*)\| + \|(\mathbf{x^{(0)}} - \mathbf{z}^*)\| \right) < \frac{2\beta L \eta}{2a^{(0)}} = \frac{\beta L \eta}{a^{(0)}} = \frac{a^{(0)}}{a^{(0)}} = 1.$$

Por lo tanto, existe el inverso de $\mathbf{P}$ y entonces $\mathbf{P}(\mathbf{z}^* - \mathbf{x}^*) = \mathbf{0} \Rightarrow \mathbf{z}^* - \mathbf{x}^* = 0 \Rightarrow \mathbf{z}^* = \mathbf{x}^*$. $\square$

Por otra parte, si $a^{(0)} = 1/2$, entonces $a^{(n)} = a^{(0)} = 1/2\ \forall\, n \geqslant 0$. Siguiendo un procedimiento totalmente análogo al del caso anterior en el que $a^{(0)} \in (0, 1/2)$, obtenemos los mismos resultados, teniendo en cuenta ahora que $\gamma = 1$ y $\Delta = h(a^{(0)})g(a^{(0)}) < 1$.

Ejemplo 4.1 ***Ecuación integral de Chandrasekhar.***

La ecuación de Chandrasekhar [10] es una ecuación integral que aparece en diversos problemas físicos. Siendo el espacio $X = C(0, 1)$ de las funciones continuas en $(0, 1)$, se trata de encontrar una función $x \in X$ que cumpla

$$x(s) = 1 + \lambda x(s) \int_0^1 \frac{s}{s+t} x(t)\, dt, \qquad s \in [0, 1]. \tag{4.4}$$

Con el fin de obtener un sistema de ecuaciones no lineales al que pueda aplicarse el Método de Newton, transformamos el problema continuo en uno discreto mediante una fórmula de integración numérica.

En concreto, utilizamos la fórmula de Gauss-Legendre con n nodos

$$\int_0^1 f(t)\, dt \simeq \sum_{j=1}^{n} p_j f(t_j),$$

donde t_j y p_j son respectivamente los nodos y los pesos de la fórmula de cuadratura en $[0, 1]$, que están tabulados para distintos valores de n mediante las fórmulas

$$t_j = \frac{y_j + 1}{2} \qquad y \qquad p_j = \frac{1}{\left(1 - y_j^2\right) \left(P_n'(y_j)\right)^2},$$

donde $P_n(y)$ denota el polinomio de Legendre de grado n e y_j, $j = (1, \ldots, n)$, son sus raíces.

Consideramos, por ejemplo, el caso particular $\lambda = 1/4$.

Denotando x_i a las aproximaciones de $x(t_i)$, para $i = 1, 2, \ldots, n$, se llega al siguiente sistema de ecuaciones no lineales

$$x_i = 1 + \frac{1}{4} x_i t_i \sum_{j=1}^{n} p_j \frac{x_j}{t_i + t_j}, \qquad i = 1, 2, \ldots, n, \tag{4.5}$$

o, equivalentemente,

$$\mathbf{f}(\mathbf{x}) = 1 - x_i + \frac{1}{4} x_i t_i \sum_{j=1}^{n} p_j \frac{x_j}{t_i + t_j} = 0, \qquad i = 1, 2, \ldots, n,$$

donde las incógnitas x_i, $i = 1, \ldots, n$, son las componentes del vector $\mathbf{x}$, que pertenece al espacio $\mathbb{R}^n$ dotado de la norma $\|\cdot\|_\infty$.

Tomamos un número de nodos $n = 8$. Como de (4.4) se deduce que $x(0) = 1$, parece razonable elegir como punto inicial $\mathbf{x^{(0)}} = (1, 1, \ldots, 1) \in \mathbb{R}^8$. Utilizando *Mathemática* como programa de cálculo, obtenemos:

i) $\|J_\mathbf{f}(\mathbf{x^{(0)}})^{-1}\|_\infty = 1,48062 = \beta$.

ii) $\|J_\mathbf{f}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\mathbf{x^{(0)}})\|_\infty = 0,255133 = \eta$.

Veamos ahora que la Jacobiana es Lipschitziana. Teniendo en cuenta que

$$\frac{\partial f_i(x)}{\partial x_i} = -1 + \frac{1}{4}\left(\frac{p_i x_i}{2} + t_i \sum_{j=1}^{8} \frac{p_j x_j}{t_i + t_j}\right), \qquad i = 1, \ldots, 8,$$

$$\frac{\partial f_i(x)}{\partial x_j} = \frac{1}{4}\frac{p_j t_i x_i}{t_i + t_j}, \qquad i, j = 1, \ldots, 8, i \neq j,$$

se tiene que $\mathbf{A} = J_\mathbf{f}(\mathbf{x}) - J_\mathbf{f}(\mathbf{y})$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, es una matriz de dimensión 8×8 formada por los elementos

$$a_{ii} = \frac{1}{4}(x_i - y_i) + t_i \sum_{j=1, j\neq i}^{8} \frac{p_j(x_j - y_j)}{t_i + t_j},$$

$$a_{ij} = \frac{1}{4}\frac{p_j t_i (x_i - y_i)}{t_i + t_j}, \qquad i \neq j.$$

Por lo tanto,

$$\|J_\mathbf{f}(\mathbf{x}) - J_\mathbf{f}(\mathbf{y})\|_\infty = \|\mathbf{A}\|_\infty =$$

$$= \frac{1}{4} \text{máx} \left\{ \left| p_i(x_i - y_i) + t_i \sum_{j=1, j\neq i}^{8} \frac{p_j(x_j - y_j)}{t_i + t_j} + \sum_{j=1, j\neq i}^{8} \frac{p_j t_i (x_i - y_i)}{t_i + t_j} \right|, i = 1, \ldots, 8 \right\} =$$

$$= \frac{1}{4} \text{máx} \left\{ \left| t_i \left(\left(\frac{p_i}{t_i} + \sum_{j=1, j\neq i}^{8} \frac{p_j}{t_i + t_j} \right)(x_i - y_i) + \sum_{j=1, j\neq i}^{8} \frac{p_j(x_j - y_j)}{t_i + t_j} \right) \right|, i = 1, \ldots, 8 \right\} =$$

$$= \frac{1}{4}\|\mathbf{B}(\mathbf{x} - \mathbf{y})\|_\infty,$$

donde $\mathbf{B}$ es una matriz 8×8 cuyos elementos son

$$b_{ii} = t_i \left(\frac{p_i}{t_i} + \sum_{j=1, j\neq i}^{8} \frac{p_j}{t_i + t_j} \right), \qquad i = 1, \ldots, 8,$$

$$b_{ij} = \frac{p_j t_i}{t_i + t_j}, \qquad i, j = 1, \ldots 8, i \neq j.$$

De este modo se concluye que

$$\|J_\mathbf{f}(\mathbf{x}) - J_\mathbf{f}(\mathbf{y})\|_\infty = \frac{1}{4}\|\mathbf{B}(\mathbf{x} - \mathbf{y})\|_\infty \leqslant \frac{1}{4}\|\mathbf{B}\|_\infty \|\mathbf{x} - \mathbf{y}\|_\infty = 0{,}344631\|\mathbf{x} - \mathbf{y}\|_\infty,$$

verificándose la condición *iii*) del Teorema 4.1 para una constante de Lipschitz $L = 0{,}344631$.

Como $a^{(0)} = L\beta\eta = 0{,}130186 < 1/2$, se cumplen todas las condiciones del Teorema de Kantovorich y se prueba la existencia de una solución $\mathbf{x}^*$ de la ecuación (4.4) con $\lambda = 1/4$, que está en la bola $\overline{B(\mathbf{x^{(0)}}, 1{,}08089)}$ y es única en $B(\mathbf{x^{(0)}}, 1{,}95976)$.

Aplicando el método de Newton para $n = 8$ se obtiene, en 3 iteraciones, la siguiente solución numérica de (4.5):

$$\mathbf{x}^* = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) =$$
$$= (1{,}02172,\ 1{,}07319,\ 1{,}12572,\ 1,16975,\ 1,20307,\ 1,22649,\ 1,24152,\ 1,24945).$$

```
P = LegendreP[8, z];
y = z /. {ToRules[NRoots[P == 0., z]]};
t = Table[(y[[i]] + 1) / 2, {i, 8}];
p[i_] := 1 / ((1 - (y[[i]]^2)) (D[P, z] /. z → y[[i]])^2);
f[x_, i_] := 1 - x[[i]] + 1 / 4 x[[i]] t[[i]] Sum[p[j] x[[j]] / (t[[i]] + t[[j]]), {j, 1, 8}];

x = {x1, x2, x3, x4, x5, x6, x7, x8};
x0 = Table[1, 8];
F = Table[f[x, i], {i, 8}];


Newton[F, x, x0, 10^(-3)]
{1.02172, 1.07319, 1.12572, 1.16975, 1.20307, 1.22649, 1.24152, 1.24945}
Número de iteraciones: 3
```

Finalmente, a partir de estos valores y mediante un proceso de interpolación, obtenemos una solución aproximada de (4.4), cuya gráfica se muestra a continuación.

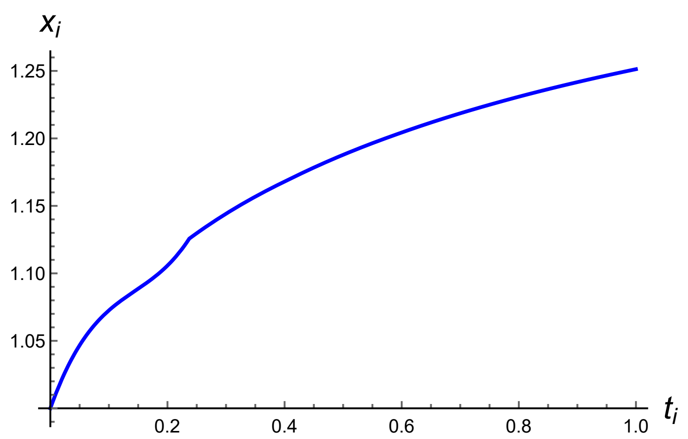

Figura 4.1: Solución aproximada de la ecuación (4.4).

Capítulo 5

Dominios de convergencia global restringida del método de Newton

5.1. Introducción

El Teorema 4.1 de convergencia semilocal para el método de Newton expuesto en el capítulo anterior, nos proporciona condiciones bajo las cuales se asegura la convergencia del método a partir del punto inicial $\mathbf{x}^{(\mathbf{0})}$. Como podemos observar, es este punto sobre el que se exigen las condiciones de los teoremas de convergencia semilocal, por lo que resulta clave realizar una elección adecuada del mismo para asegurar la convergencia del método. Esta necesidad de disponer de un punto inicial $\mathbf{x}^{(\mathbf{0})}$ suficientemente bueno se suprime en los teoremas de convergencia global, como vimos en el Capítulo 3. Así, en este capítulo mostramos un teorema de convergencia global para el método de Newton, que nos proporciona regiones de $\mathbb{R}^n$ en las que cualquier punto contenido en ellas puede utilizase para iniciar el método asegurando su convergencia. Esta convergencia global será restringida a dominios, generalmente bolas, de $\mathbb{R}^n$.

Este estudio se realizará a partir de la consideración de un punto denominado auxiliar $\tilde{\mathbf{x}}$. Entonces suponiendo las condiciones puntuales en este punto auxiliar seremos capaces de obtener un dominio de convergencia global para el método de Newton.

Definición 5.1 *Sean* $(E, \|\cdot\|)$ *y* $(V, \|\cdot\|)$ *dos espacios métricos normados,* $\mathbf{g} : E \longrightarrow V$ *una aplicación definida en* E *y con valores en* V *y* $\tilde{\mathbf{x}} \in E$*. Se dice que* $\mathbf{g}$ *es una* ***aplicación Lipschitz-centrada en*** $\tilde{\mathbf{x}}$ *cuando existe una constante real* $\tilde{L} > 0$ *tal que:*

$$\|\mathbf{g}(\mathbf{x}) - \mathbf{g}(\tilde{\mathbf{x}})\| \leqslant \tilde{L}\|\mathbf{x} - \tilde{\mathbf{x}}\| \qquad \forall\, \mathbf{x} \in E.$$

Resulta obvio que si se verifica la condición *iii*) del Teorema 4.1, es decir,

$$\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad \forall\, \mathbf{x}, \mathbf{y} \in \Omega, \tag{5.1}$$

entonces para cualquier punto $\tilde{\mathbf{x}} \in \Omega$ se verifica que

$$\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\tilde{\mathbf{x}})\| \leqslant \tilde{L}\|\mathbf{x} - \tilde{\mathbf{x}}\|, \tag{5.2}$$

con $\tilde{L} \leqslant L$. De este modo, una vez fijado $\tilde{\mathbf{x}} \in \Omega$, puede definirse $\mu = \frac{\tilde{L}}{L} \in (0, 1]$.

5.2. Resultado principal

Como hemos indicado, mediante la utilización de un punto auxiliar $\tilde{\mathbf{x}}$ ya estamos en disposición de obtener un resultado de convergencia global restringida.

Teorema 5.1 *Sea Ω abierto y convexo de $\mathbb{R}^n$ y sea $\mathbf{f} : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación de clase $C^{(1}(\Omega)$. Supongamos que $J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}$ existe en algún $\tilde{\mathbf{x}} \in \Omega$ y que se verifican las condiciones:*

i) $\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\| \leqslant \beta$,

ii) $\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\tilde{\mathbf{x}})\| \leqslant \eta$,

iii) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad L \in \mathbb{R}^+, \forall\, \mathbf{x}, \mathbf{y} \in \Omega.$

iv) Existe $R > 0$ tal que $B(\tilde{\mathbf{x}}, R) \subset \Omega$.

Si se satisface que

$$L\beta\eta \leqslant \frac{1}{2(1 + 2\mu)}, \tag{5.3}$$

siendo $\mu = \frac{\tilde{L}}{L}$ con $\tilde{L}$ la constante positiva que verifica (5.2) , la sucesión de Newton $\left\{\mathbf{x}^{(\mathbf{n})}\right\}_{n=0}^{\infty}$ dada por

$$\mathbf{x}^{(\mathbf{n+1})} = \mathbf{x}^{(\mathbf{n})} - J_{\mathbf{f}}\left(\mathbf{x}^{(\mathbf{n})}\right)^{-1}\mathbf{f}\left(\mathbf{x}^{(\mathbf{n})}\right),$$

empezando en cualquier $\mathbf{x}^{(\mathbf{0})} \in B(\tilde{\mathbf{x}}, R)$, converge a una solución $\mathbf{x}^*$ del sistema de ecuaciones $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, la solución $\mathbf{x}^*$ y las iteraciones $\mathbf{x}^{(\mathbf{n})}$ pertenecen a $\overline{B(\tilde{\mathbf{x}}, R)}$ y la solución $\mathbf{x}^*$ es única en la región $B(\tilde{\mathbf{x}}, r) \cap \Omega$, donde $r = \frac{2}{L\beta} - R$.

Antes de comenzar con la demostración del teorema, presentamos un Lema previo que será de utilidad para su desarrollo.

Lema 5.2 *Si $L\beta\eta \leqslant \frac{1}{2(1+2\mu)}$,*

i) se obtiene que $[R_1, R_2] \neq \emptyset$, donde R_1 y R_2 son las dos soluciones de la ecuación de segundo grado

$$(1 + 2\mu)L\beta t^2 - 2t + 2\eta = 0 \qquad y$$

ii) se verifica $\alpha = Lab < 2$, donde $a = \frac{\beta}{1-\mu L\beta R}$ y $b = \frac{\eta + R + \frac{1}{2}\mu L\beta R^2}{1-\mu L\beta R}$. Esta condición puede escribirse, equivalentemente, como

$$2(2 - L\beta\eta) - 2(1 + 4\mu)(L\beta R) + \mu(4\mu - 1)(L\beta R)^2 > 0.$$

La demostración de este lema se obtiene directamente a partir de sencillos desarrollos numéricos y puede verse en [6].

Teniendo en cuenta el lema anterior, nos disponemos a realizar la demostración del Teorema 5.1.

Demostración del Teorema 5.1.
Elegimos R tal que $\alpha < 2$. En primer lugar probamos, para todo $\mathbf{x} \in B(\tilde{\mathbf{x}}, R)$, que $J_{\mathbf{f}}(\mathbf{x})^{-1}$ existe y $\|J_{\mathbf{f}}(\mathbf{x})^{-1}\| \leqslant \frac{\beta}{1-\mu K\beta R}$.Para ello, vemos que

$$\|I - J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}J_{\mathbf{f}}(\mathbf{x})\| \leqslant \|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\|\|J_{\mathbf{f}}(\tilde{\mathbf{x}}) - J_{\mathbf{f}}(\mathbf{x})\| \leqslant \beta\tilde{L}\|\tilde{\mathbf{x}} - \mathbf{x}\| \leqslant$$

$$\leqslant \mu L\beta R \leqslant \frac{1}{2(1+2\mu)} < 1,$$

puesto que $\mu \in (0, 1]$.

Por lo tanto, por el Lema 2.5, existe $J_{\mathbf{f}}(\mathbf{x})^{-1}$ y es tal que

$$\|J_{\mathbf{f}}(\mathbf{x})^{-1}\| \leqslant \frac{\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\|}{1 - \|I - J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}J_{\mathbf{f}}(\mathbf{x})\|} \leqslant \frac{\beta}{1 - \mu L\beta R} = a.$$

A partir de la fórmula de Taylor [3], se obtiene

$$\mathbf{f}(\mathbf{x^{(0)}}) = \mathbf{f}(\tilde{\mathbf{x}}) + \int_0^1 J_{\mathbf{f}}(\tilde{\mathbf{x}} + t(\mathbf{x^{(0)}} - \tilde{\mathbf{x}}))(\mathbf{x^{(0)}} - \tilde{\mathbf{x}})\, dt =$$

$$= \mathbf{f}(\tilde{\mathbf{x}}) + \left(J_{\mathbf{f}}(\tilde{\mathbf{x}}) + \int_0^1 J_{\mathbf{f}}(\tilde{\mathbf{x}} + t(\mathbf{x^{(0)}} - \tilde{\mathbf{x}})) - J_{\mathbf{f}}(\tilde{\mathbf{x}})\, dt\right)(\mathbf{x^{(0)}} - \tilde{\mathbf{x}}) \Rightarrow$$

$$\Rightarrow J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\mathbf{x^{(0)}}) = J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\tilde{\mathbf{x}}) + \mathbf{x^{(0)}} - \tilde{\mathbf{x}} +$$

$$+ J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\int_0^1 \big(J_{\mathbf{f}}(\tilde{\mathbf{x}} + t(\mathbf{x^{(0)}} - \tilde{\mathbf{x}})) - J_{\mathbf{f}}(\tilde{\mathbf{x}})\big)\, dt\, (\mathbf{x^{(0)}} - \tilde{\mathbf{x}}). \tag{5.4}$$

Por lo tanto

$$\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\mathbf{x^{(0)}})\| \leqslant$$

$$\leqslant \|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\tilde{\mathbf{x}})\| + \|\mathbf{x^{(0)}} - \tilde{\mathbf{x}}\| + \|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\|\int_0^1 \tilde{L}t\|\mathbf{x^{(0)}} - \tilde{\mathbf{x}}\|\, dt\, \|\mathbf{x^{(0)}} - \tilde{\mathbf{x}}\| \leqslant$$

$$\leqslant \eta + R + \frac{1}{2}\tilde{L}\beta R^2 = \eta + R + \frac{1}{2}\mu L\beta R^2.$$

Como $\mathbf{x^{(0)}} \in B(\tilde{\mathbf{x}}, R)$, existe $J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}$, $\mathbf{x^{(1)}}$ está bien definido y, como $\|I - J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}J_{\mathbf{f}}(\mathbf{x^{(0)}})\| = \mu L\beta R < 1$, por el Lema 2.5 existe $(J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}J_{\mathbf{f}}(\mathbf{x^{(0)}}))^{-1} = J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}J_{\mathbf{f}}(\tilde{\mathbf{x}})$ y $\|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}J_{\mathbf{f}}(\tilde{\mathbf{x}})\| \leqslant \frac{1}{1-\mu L\beta R}$.

Así pues, por la fórmula de Newton

$$\|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\| \leqslant \|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\mathbf{x^{(0)}})\| \leqslant \|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}J_{\mathbf{f}}(\tilde{\mathbf{x}})\|\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{(-1)}\mathbf{f}(\mathbf{x^{(0)}})\| \leqslant$$

$$\leqslant \frac{\eta + R + \frac{1}{2}\mu L\beta R^2}{1-\mu L\beta R} = b.$$

De manera análoga a (5.4), mediante la fórmula de Taylor se obtiene

$$J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\tilde{\mathbf{x}}) = J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\mathbf{x^{(0)}}) + \tilde{\mathbf{x}} - \mathbf{x^{(0)}} +$$

$$+J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\int_0^1 J_{\mathbf{f}}(\mathbf{x^{(0)}} + t(\tilde{\mathbf{x}} - \mathbf{x^{(0)}})) - J_{\mathbf{f}}(\mathbf{x^{(0)}})\, dt\, (\tilde{\mathbf{x}} - \mathbf{x^{(0)}})$$

y, como por la fórmula de Newton $\mathbf{x^{(1)}} - \tilde{\mathbf{x}} = \mathbf{x^{(0)}} - J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\mathbf{x^{(0)}}) - \tilde{\mathbf{x}}$, entonces

$$\mathbf{x^{(1)}} - \tilde{\mathbf{x}} = -J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\mathbf{f}(\tilde{\mathbf{x}}) + J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\int_0^1 \left(J_{\mathbf{f}}(\mathbf{x^{(0)}} + t(\tilde{\mathbf{x}} - \mathbf{x^{(0)}})) - J_{\mathbf{f}}(\mathbf{x^{(0)}})\right) dt\, (\tilde{\mathbf{x}} - \mathbf{x^{(0)}})$$

y por tanto

$$\|\mathbf{x^{(1)}} - \tilde{\mathbf{x}}\| \leqslant$$

$$\leqslant \|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}J_{\mathbf{f}}(\tilde{\mathbf{x}})\|\|\mathbf{f}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\tilde{\mathbf{x}})\| + \|J_{\mathbf{f}}(\mathbf{x^{(0)}})^{-1}\|\int_0^1 Lt\|\tilde{\mathbf{x}} - \mathbf{x^{(0)}}\|\, dt\, \|\tilde{\mathbf{x}} - \mathbf{x^{(0)}}\| \leqslant$$

$$\leqslant \frac{\eta}{1-\mu L\beta R} + \frac{L\beta R^2}{2(1-\mu L\beta R)} = \frac{2\eta + L\beta R^2}{2(1-\mu L\beta R)}.$$

Entonces $\mathbf{x^{(1)}} \in B(\tilde{\mathbf{x}}, R)$, ya que a partir de la condición (5.1) y teniendo en cuenta el Lema 5.2,

$$(1+2\mu)L\beta R^2 - 2R + 2\eta \leqslant 0 \Rightarrow \frac{2\eta + L\beta R^2}{2(1-\mu L\beta R)} \leqslant R.$$

Ahora aplicaremos inducción. Así, suponemos que se verifican

$$(i)\ \|\mathbf{f}(\mathbf{x^{(n+1)}})\| \leqslant \frac{L}{2}\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\|, \tag{5.5}$$

$$(ii)\ \|\mathbf{x^{(n+1)}} - \tilde{\mathbf{x}}\| < \frac{2\eta + L\beta R^2}{2(1-\mu L\beta R)} \leqslant R, \tag{5.6}$$

$$(iii)\ \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| < \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n)}} - \mathbf{x^{(n-1)}}\|, \tag{5.7}$$

y probamos estos items para $n+2$.

Como $\mathbf{x^{(n+1)}} \in B(\tilde{\mathbf{x}}, R)$, existe $J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}$ con $\|J_{\mathbf{f}}(\mathbf{x^{(n+1)}})\| < \frac{\beta}{1-\mu L\beta R}$. Por el Lema 2.5 tenemos que $\|J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}J_{\mathbf{f}}(\tilde{\mathbf{x}})\| \leqslant \frac{1}{1-\|I-J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}J_{\mathbf{f}}(\mathbf{x^{(n+1)}})\|} \leqslant \frac{1}{1-\mu L\beta R}$. Por lo que $\mathbf{x^{(n+2)}}$ está bien definido y se tiene que:

(i) Aplicando de nuevo la fórmula de Taylor se obtiene

$$\mathbf{f}(\mathbf{x^{(n+2)}}) = \mathbf{f}(\mathbf{x^{(n+1)}}) + \int_0^1 J_{\mathbf{f}}(\mathbf{x^{(n+1)}} + t(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}))(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}})\, dt$$

y, como por la fórmula de Newton

$$\mathbf{x^{(n+2)}} = \mathbf{x^{(n+1)}} - J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}\mathbf{f}(\mathbf{x^{(n+1)}}) \Rightarrow \mathbf{f}(\mathbf{x^{(n+1)}}) = -J_{\mathbf{f}}(\mathbf{x^{(n+1)}})(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}),$$

se tiene que

$$\mathbf{f}(\mathbf{x^{(n+2)}}) = -\int_0^1 J_{\mathbf{f}}(\mathbf{x^{(n+1)}})(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}})\,dt+$$

$$+\int_0^1 J_{\mathbf{f}}(\mathbf{x^{(n+1)}} + t(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}))(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}})\,dt =$$

$$= \int_0^1 (J_{\mathbf{f}}(\mathbf{x^{(n+1)}} + t(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}})) - J_{\mathbf{f}}(\mathbf{x^{(n+1)}}))(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}})\,dt$$

Por lo tanto

$$\|\mathbf{f}(\mathbf{x^{(n+2)}})\| = \int_0^1 \|(J_{\mathbf{f}}(\mathbf{x^{(n+1)}}+t(\mathbf{x^{(n+2)}}-\mathbf{x^{(+1)}}))-J_{\mathbf{f}}(\mathbf{x^{(n+1)}}))(\mathbf{x^{(n+2)}}-\mathbf{x^{(n+1)}})\|dt \leqslant$$

$$\leqslant \|\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}\| \int_0^1 L\|(\mathbf{x^{n+1}} + t(\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}) - \mathbf{x^{(n+1)}})\|dt \leqslant$$

$$\leqslant L\|\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}\|^2 \int_0^1 t\,dt = \frac{L}{2}\|\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}\|^2.$$

(ii)

$$\mathbf{x^{(n+2)}} - \tilde{\mathbf{x}} = \mathbf{x^{(n+1)}} - J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}\mathbf{f}(\mathbf{x^{(n+1)}}) - \tilde{\mathbf{x}} = -J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}\mathbf{f}(\tilde{\mathbf{x}})+$$

$$+J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}(\tilde{\mathbf{x}} - \mathbf{x^{(n+1)}}) \int_0^1 (J_{\mathbf{f}}(\mathbf{x^{(n+1)}} + t(\tilde{\mathbf{x}} - \mathbf{x^{(n+1)}})) - J_{\mathbf{f}}(\mathbf{x^{(n+1)}}))\,dt$$

y

$$\|\mathbf{x^{(n+2)}} - \tilde{\mathbf{x}}\| \leqslant \|J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}J_{\mathbf{f}}(\tilde{\mathbf{x}})\|\|J_{\mathbf{f}}(\tilde{\mathbf{x}})\mathbf{f}(\tilde{\mathbf{x}})\|+$$

$$+\frac{L}{2}\|J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}\|\|\tilde{\mathbf{x}} - \mathbf{x^{(n+1)}}\|^2 \leqslant \frac{2\eta + L\beta R^2}{2(1-\mu L\beta R} \leqslant R.$$

(iii)

$$\|\mathbf{x^{(n+2)}} - \mathbf{x^{(n+1)}}\| \leqslant \|J_{\mathbf{f}}(\mathbf{x^{(n+1)}})^{-1}\mathbf{f}(\mathbf{x^{(n-1)}})\| \leqslant \frac{La}{2}\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\|^2 \leqslant$$

$$\leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\|\|J_{\mathbf{f}}(\mathbf{x^{(n-1)}})^{-1}\mathbf{f}(\mathbf{x^{(n-1)}})\| \leqslant$$

$$\leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \left(\frac{La}{2}\right)^2 \|\mathbf{x^{(n-1)}} - \mathbf{x^{(n-2)}}\|^2 \leqslant$$

$$\leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \left(\frac{La}{2}\right)^2 \|J_{\mathbf{f}}(\mathbf{x^{(n-2)}})^{-1}\mathbf{f}(\mathbf{x^{(n-2)}})\|^2 \leqslant$$

$$\leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \left(\frac{La}{2}\right)^4 \|\mathbf{x^{(n-2)}} - \mathbf{x^{(n-3)}}\|^4 \leqslant$$

$$\leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \left(\frac{La}{2}\right)^8 \|\mathbf{x^{(n-3)}} - \mathbf{x^{(n-4)}}\|^8 \leqslant \cdots$$

$$\cdots \leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| \left(\frac{La}{2}\right)^{2^{n-1}} \|\mathbf{x^{(1)}} - \mathbf{x^{(0)}}\|^{2^{n-1}} \leqslant$$

$$\leqslant \left(\frac{\alpha}{2}\right)^{2^{n-1}} \left(\frac{Lab}{2}\right)^{2^{n-1}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| < \left(\frac{\alpha}{2}\right)^{2^{n}} \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\| < \|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\|.$$

Por lo tanto, por inducción matemática se verifican las ecuaciones (5.5), (5.6) y (5.7) para todo $n \in \mathbb{N}$.

En consecuencia, se obtiene que $\mathbf{x^{(n)}} \in B(\tilde{\mathbf{x}}, R)$ para todo $n \geqslant 0$ y, como $\alpha < 2$, la sucesión $\left\{\|\mathbf{x^{(n+1)}} - \mathbf{x^{(n)}}\|\right\}$ es estrictamente decreciente, lo cual implica que $\left\{\mathbf{x^{(n)}}\right\}$ converge: $\lim \mathbf{x^{(n)}} = \mathbf{x}^*$. Además, como $\|\mathbf{x^{(n)}} - \mathbf{x^{n-1}}\| \to 0$ cuando $n \to \infty$, se cumple que $\|\mathbf{f}(\mathbf{x^{(n)}})\| \to 0$ cuando $n \to \infty$. Por lo tanto, $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$ por la continuidad de $\mathbf{f}$ en $B(\mathbf{x^{(0)}}, R)$.

Por último, probamos que $\mathbf{x}^*$ es la única solución de $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ en $B(\tilde{\mathbf{x}}, r) \cap \Omega$, donde $r = \frac{2}{L\beta} - R$.
Supongamos que $\mathbf{z}^*$ es otra solución de $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ en $B(\tilde{\mathbf{x}}, r) \cap \Omega$ tal que $\mathbf{z}^* \neq \mathbf{x}^*$. Realizamos la aproximación

$$\mathbf{0} = J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}(\mathbf{f}(\mathbf{z}^*) - \mathbf{f}(\mathbf{x}^*)) =$$

$$= \left(\int_0^1 J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1} J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))dt\right)(\mathbf{z}^* - \mathbf{x}^*) = \mathbf{P}(\mathbf{z}^* - \mathbf{x}^*)$$

y probamos que existe el inverso de $\mathbf{P} = \int_0^1 J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1} J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))dt$, para lo cual bastará con verificar que $\|\mathbf{I} - \mathbf{P}\| < 1$. En efecto,

$$\|\mathbf{I} - \mathbf{P}\| = \|\mathbf{I} - J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1} \int_0^1 J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))dt\| =$$

$$= \|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\left(J_{\mathbf{f}}(\tilde{\mathbf{x}}) - \int_0^1 J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))dt\right)\| \leqslant$$

$$\leqslant \|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1})\| \int_0^1 \|J_{\mathbf{f}}(\tilde{\mathbf{x}}) - J_{\mathbf{f}}(\mathbf{x}^* + t(\mathbf{z}^* - \mathbf{x}^*))\|dt \leqslant$$

$$\leqslant \|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\| L \int_0^1 \|\tilde{\mathbf{x}} - \mathbf{x}^* + t(\mathbf{x}^* - \mathbf{z}^*)\|dt \leqslant$$

$$\leqslant \frac{\beta L}{2}(\|\tilde{\mathbf{x}} - \mathbf{x}^*\| + \|\tilde{\mathbf{x}} - \mathbf{z}^*\|) < \frac{\beta L}{2}(r + R) = 1$$

Por lo tanto, por el Lema (2.4), existe el inverso de $\mathbf{P}$ y entonces
$\mathbf{P}(\mathbf{z}^* - \mathbf{x}^*) = \mathbf{0} \Rightarrow \mathbf{z}^* - \mathbf{x}^* = \mathbf{0} \Rightarrow \mathbf{z}^* = \mathbf{x}^*$.

□

Es claro que el resultado de convergencia es global restringido a una bola. Además, considerando $\tilde{\mathbf{x}} = \mathbf{x^{(0)}}$ o bien $\tilde{\mathbf{x}} = \mathbf{x}^*$ podemos obtener sendos resultados de convergencia semilocal y convergencia local respectivamente.

Ejemplo 5.1 ***Ecuación integral de tipo Hammerstein mixto.***

Consideramos la siguiente ecuación integral no linear de tipo Hammerstein mixto [11]: Siendo el espacio $X = C(0,1)$ de las funciones continuas en $(0,1)$ se trata de encontrar una función $x \in X$ que cumpla

$$x(s) = 1 + \lambda \int_0^1 G(s,t)\, x(t)^3 dt, \qquad (5.8)$$

donde el núcleo G es la función de Green en $[0,1] \times [0,1]$.

Consideramos el caso particular $\lambda = 3/4$. Procediendo de manera análoga al Ejemplo 4.1, usamos la fórmula de cuadratura de Gauss-Legendre con un número de pasos $n = 8$ para transformar esta ecuación en un sistema de ecuaciones no lineales. Si denotamos las aproximaciones de $x(t_i)$ como $x_i, i = 1, \ldots, n$, obtenemos

$$x_i = 1 + \frac{3}{4} \sum_{j=1}^{n} a_{ij} x_j^3, \qquad j = 1, 2, \ldots, n,$$

o, equivalentemente,

$$f_i(x) = 1 - x_i + \frac{3}{4} \sum_{j=1}^{n} a_{ij} x_j^3 = 0, \qquad j = 1, 2, \ldots, n, \qquad (5.9)$$

donde

$$a_{ij} = p_j G(t_i, t_j) = \begin{cases} p_j(1 - t_i)t_j, & j \leqslant i, \\ p_j(1 - t_j)t_i, & j > i \end{cases},$$

con $f_i : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n, i = 1, \ldots, n$, y siendo p_j y $t_i, j = 1, \ldots, n$ respectivamente los pesos y los nodos de la fórmula de cuadratura en $[0,1]$ tabulados para distintos valores de n.

Vamos a considerar una región $\Omega \in \mathbb{R}^n$ en la que se encuentre prelocalizada una solución numérica $\mathbf{x}^*$ de (5.8). Para ello, acotamos la norma de x^* solución de (5.8),

$$x^*(s) = 1 + \frac{3}{4} \int_0^1 G(s,t)\, x^*(t)^3\, dt.$$

Tomando normas en la expresión anterior se obtiene

$$\|x^*\| \leqslant 1 + \frac{3}{4} M \|x^*\|^3, \qquad \text{con } M = \max_s \int_0^1 |G(s,t)|\, dt,$$

esto es, $0 \leqslant 1 - \|x^*\| + \dfrac{3}{32}\|x^*\|^3 \Rightarrow \|x^*\| \leqslant 1,13826$.

Por lo tanto, consideramos la bola $\Omega = B(\mathbf{0}, 2)$, incluida en el espacio $\mathbb{R}^n$ dotado de la norma infinito. Nuestro objetivo es obtener una solución aproximada del sistema (5.9) aplicando el método de Newton y su correspondiente dominio de puntos de inicio.

Elegimos el punto $\tilde{\mathbf{x}} = (1,1;1,1;\ldots;1,1) \in \mathbb{R}^8$ y tomamos $R = 0{,}9$ obteniéndose que:

i) $\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\|_\infty = \beta = 1,47218$,

ii) $\|J_{\mathbf{f}}(\tilde{\mathbf{x}})^{-1}\mathbf{f}(\tilde{\mathbf{x}})\|_\infty = \eta = 0,093309$,

iii) $B(\tilde{\mathbf{x}}, R) \subset \Omega$.

Ahora, como

$$J_{\mathbf{f}}(\mathbf{x}) = \begin{pmatrix} -1+\frac{9}{4}a_{11}x_1^2 & \frac{9}{4}a_{12}x_2^2 & \cdots & \frac{9}{4}a_{18}x_8^2 \\ \frac{9}{4}a_{21}x_1^2 & -1+\frac{9}{4}a_{22}x_2^2 & \cdots & \frac{9}{4}a_{28}x_8^2 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{9}{4}a_{81}x_1^2 & \frac{9}{4}a_{82}x_2^2 & \cdots & -1+\frac{9}{4}a_{88}x_8^2 \end{pmatrix}$$

se tiene que

$$J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y}) = \frac{9}{4}\begin{pmatrix} a_{11}(x_1^2-y_1^2) & a_{12}(x_2^2-y_2^2) & \cdots & a_{18}(x_8^2-y_8^2) \\ a_{21}(x_1^2-y_1^2) & a_{22}(x_2^2-y_2^2) & \cdots & a_{28}(x_8^2-y_8^2) \\ \vdots & \vdots & \ddots & \vdots \\ a_{81}(x_1^2-y_1^2) & a_{82}(x_2^2-y_2^2) & \cdots & a_{88}(x_8^2-y_8^2) \end{pmatrix} \Rightarrow$$

$$\Rightarrow \|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\|_\infty = \frac{9}{4}\,\text{máx}\left\{\left|\sum_{j=1}^{8} a_{ij}(x_i^2 - y_i^2), i = 1,\ldots,8\right|\right\} = \|\mathbf{B}(\mathbf{x}^2 - \mathbf{y}^2)\|_\infty,$$

donde $\mathbf{B}$ es una matriz 8×8 cuyos elementos son $b_{ij} = \frac{9}{4}a_{ij}$.

Teniendo en cuenta que $\mathbf{x},\mathbf{y} \in B(\mathbf{0},2) \Rightarrow \|\mathbf{x}+\mathbf{y}\|_\infty \leqslant \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty \leqslant 4$, se tiene

$$\|\mathbf{B}(\mathbf{x}^2-\mathbf{y^2})\|_\infty \leqslant \|\mathbf{B}(\mathbf{x}+\mathbf{y})\|_\infty\|\mathbf{x}-\mathbf{y}\|_\infty \leqslant 4\|\mathbf{B}\|_\infty\|\mathbf{x}-\mathbf{y}\|_\infty = 1,11203\|\mathbf{x}-\mathbf{y}\|_\infty,$$

concluyéndose que la Jacobiana es Lipschitziana $\forall\, \mathbf{x},\mathbf{y} \in \mathbf{\Omega}$ con constante de Lipschitz $L = 1,11203$.

De manera análoga, como $\mathbf{x} \in B(\mathbf{0},2) \Rightarrow \|\mathbf{x}+\tilde{\mathbf{x}}\|_\infty \leqslant \|\mathbf{x}\|_\infty + \|\tilde{\mathbf{x}}\|_\infty \leqslant 3{,}1$, se tiene

$$\|J_{\mathbf{f}}(\mathbf{x})-J_{\mathbf{f}}(\tilde{\mathbf{x}})\|_\infty \leqslant \|\mathbf{B}(\mathbf{x}+\tilde{\mathbf{x}})\|_\infty\|\mathbf{x}-\tilde{\mathbf{x}}\|_\infty \leqslant 3{,}1\|\mathbf{B}\|_\infty\|\mathbf{x}-\tilde{\mathbf{x}}\|_\infty = 0,861824\|\mathbf{x}-\tilde{\mathbf{x}}\|_\infty,$$

por lo que la Jacobiana es Lipschitz-centrada en $\tilde{\mathbf{x}} = (1,1;1,1;\ldots;1,1)$ con constante de Lipschitz $\tilde{L} = 0{,}861824$ y entonces $\mu = \frac{\tilde{L}}{L} = 0,775$.

Como $L\beta\eta = 0,152757$ y $\frac{1}{2(1+2\mu)} = 0,196078$, se verifica la condición (5.3) y por el Teorema 5.1 puede asegurarse la convergencia del método de Newton a una solución $\mathbf{x}^* \in B(\tilde{\mathbf{x}},0,9)$ del sistema (5.9) tomando como punto inicial cualquier $\mathbf{x^{(0)}} \in B(\tilde{\mathbf{x}},0,9)$. Además, esta solución es única en la región $B(\tilde{\mathbf{x}},r)\cap\Omega$ donde

$r = \frac{2}{L\beta} - R = 0,321669.$

Aplicando el método de Newton con $\mathbf{x^{(0)}} = (1, 1, \ldots, 1) \in B(\tilde{\mathbf{x}}, 0, 9)$ se obtiene, en 4 iteraciones, la siguiente solución numérica de (5.8):

$$\mathbf{x}^* = (x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) =$$
$$= (1,00958,\ 1,04584,\ 1,09301,\ 1,12624,\ 1,12624,\ 1,09301,\ 1,04584,\ 1,00958).$$

```
P = LegendreP[8, z];
y = z /. {ToRules[NRoots[P == 0., z]]};
t = Table[(y[[i]] + 1) / 2, {i, 8}];
p[i_] := 1 / ((1 - (y[[i]]^2)) (D[P, z] /. z → y[[i]])^2);
a[i_, j_] := If[j > i, p[j] (1 - t[[j]]) t[[i]], p[j] (1 - t[[i]]) t[[j]]];
f[x_, i_] := 1 - x[[i]] + 3 / 4 Sum[a[i, j] x[[j]]^3, {j, 1, 8}];

x = {x1, x2, x3, x4, x5, x6, x7, x8};
x0 = Table[1, 8];
F = Table[f[x, i], {i, 8}];


Newton[F, x, x0, 10^(-6)]
{1.00958, 1.04584, 1.09301, 1.12624, 1.12624, 1.09301, 1.04584, 1.00958}
Número de iteraciones: 4
```

Por último, a partir de estos valores y mediante un proceso de interpolación, obtenemos una solución aproximada de (5.8), cuya gráfica se muestra a continuación.

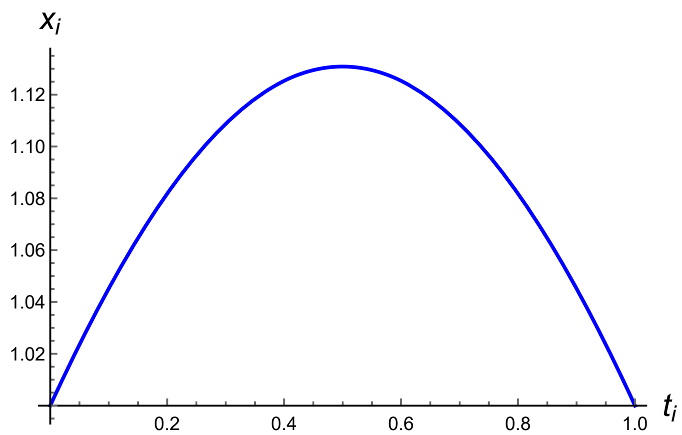

Figura 5.1: Solución aproximada de la ecuación (5.8).

5.3. Convergencia semilocal del método de Newton consecuencia del resultado principal

Si en el Teorema 5.1 de convergencia global tomamos como punto auxiliar un punto de inicio, $\tilde{\mathbf{x}} = \mathbf{x}^{(\mathbf{0})}$, obtenemos un teorema de convergencia semilocal para el método de Newton que, además de asegurar la convergencia, prueba la existencia de una solución $\mathbf{x}^*$ del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

Teorema 5.3 *Sea Ω abierto y convexo de $\mathbb{R}^n$ y sea $\mathbf{f} : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación de clase $C^{(1}(\Omega)$. Supongamos que $J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})$ existe en algún $\mathbf{x}^{(\mathbf{0})} \in \Omega$ y que se verifican las condiciones:*

i) $\|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}\| \leqslant \beta$,

ii) $\|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{0})})\| \leqslant \eta$,

iii) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad L \in \mathbb{R}^+, \forall\, \mathbf{x}, \mathbf{y} \in \Omega.$

iv) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})\| \leqslant \tilde{L}\|\mathbf{x} - \mathbf{x}^{(\mathbf{0})}\|, \qquad \tilde{L} \in \mathbb{R}^+, \forall\, \mathbf{x} \in \Omega.$

v) existe $R > 0$ tal que $B\left(\mathbf{x}^{(\mathbf{0})}, R\right) \subset \Omega$.

Si se satisface que

$$L\beta\eta \leqslant \frac{1}{2(1+2\mu)},$$

siendo $\mu = \frac{\tilde{L}}{L} \in (0,1]$, la sucesión de Newton $\left\{\mathbf{x}^{(\mathbf{n})}\right\}_{n=0}^{\infty}$ dada por

$$\mathbf{x}^{(\mathbf{n+1})} = \mathbf{x}^{(\mathbf{n})} - J_{\mathbf{f}}\left(\mathbf{x}^{(\mathbf{n})}\right)^{-1}\mathbf{f}\left(\mathbf{x}^{(\mathbf{n})}\right),$$

empezando en $\mathbf{x}^{(\mathbf{0})}$, converge a una solución $\mathbf{x}^*$ del sistema de ecuaciones $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, la solución $\mathbf{x}^*$ y las iteraciones $\mathbf{x}^{(\mathbf{n})}$ pertenecen a $B(\mathbf{x}^{(\mathbf{0})}, R)$ y la solución $\mathbf{x}^*$ es única en la región $B(\mathbf{x}^{(\mathbf{0})}, r) \cap \Omega$, donde $r = \frac{2}{L\beta} - R$.

Demostración.
Es un caso particular del Teorema (5.1) con $\tilde{\mathbf{x}} = \mathbf{x}^{(\mathbf{0})}$. □

Como puede observarse, este teorema presenta ciertas similitudes con el Teorema 4.1 de convergencia semilocal de Kantorovich expuesto en el Capítulo 4. En ambos se exige las condiciones

i) $\|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}\| \leqslant \beta$,

ii) $\|J_{\mathbf{f}}(\mathbf{x}^{(\mathbf{0})})^{-1}\mathbf{f}(\mathbf{x}^{(\mathbf{0})})\| \leqslant \eta$,

iii) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad L \in \mathbb{R}^+, \forall\, \mathbf{x}, \mathbf{y} \in \Omega,$

y, aunque de la condición *iii*) se sigue que necesariamente debe existir un $\tilde{L}$ tal que se verifica

iv) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{x}^{(0)})\| \leqslant \tilde{L}\|\mathbf{x} - \mathbf{x}^{(0)}\|, \qquad \tilde{L} \in \mathbb{R}^+, \forall\, \mathbf{x} \in \Omega,$

con $\tilde{L} \leqslant L$, esta condición no se exige en el Teorema 4.1.

Además, en el Teorema 5.3 debe verificarse $L\beta\eta \leq \dfrac{1}{2(1+2\mu)}$, con $\mu \in (0,1]$ dependiente de $\tilde{L}$, lo cual implica que $L\beta\eta \leqslant \alpha$ con $\alpha \in \left[\frac{1}{6}, \frac{1}{2}\right)$.

Notemos que, cuando se da el caso $\tilde{L} = L$, sus condiciones $i), ii), iii)$ y $iv)$ son idénticas a las del Teorema de Kantorovich 4.1. Sin embargo, al ser $\mu = 1$, para asegurar la convergencia del método deberá cumplirse que $L\beta\eta \leqslant \frac{1}{6}$, que es una condición más estricta que la exigida en el Teorema de Kantorovich $L\beta\eta \leqslant \frac{1}{2}$.

En cualquier caso, el Teorema de Kantorovich 4.1 resulta más conveniente que el Teorema 5.1. Este hecho puede apreciarse con claridad mediante un estudio de la accesibilidad del método de Newton a partir de las condiciones impuestas en ambos teoremas de convergencia semilocal [4]. El conjunto de los puntos de inicio $\mathbf{x}^{(0)}$ que hacen que el método de Newton converja a una solución permiten medir la accesibilidad del método. En este caso, la accesibilidad se observa de forma teórica mediante un dominio de parámetros, que establece gráficamente en un plano real la relación entre los parámetros L, β y η que se definen en las condiciones. Así, en la Figura (5.2) mostramos el dominio de parámetros del método de Newton asociado al Teorema de Kantorovich 4.1. Para representarlo gráficamente, se considera el plano xy, donde el eje de abcisas es $x = \eta$ y el eje de coordenadas es $y = L\beta$, y se colorean de rojo los valores de los parámetros que verifican la condición $L\beta\eta \leqslant 1/2$. De manera análoga, en la Figura (5.3) mostramos el dominio de parámetros del método de Newton asociado al Teorema 5.3, donde se colorean los valores de los parámetros que verifican $L\beta\eta \leqslant \alpha$, con $\alpha = \frac{1}{6}, \frac{1}{4}, \frac{1}{3}$ y $L\beta\eta < \frac{1}{2}$ (regiones verde,azul, amarilla y roja, repectivamente).

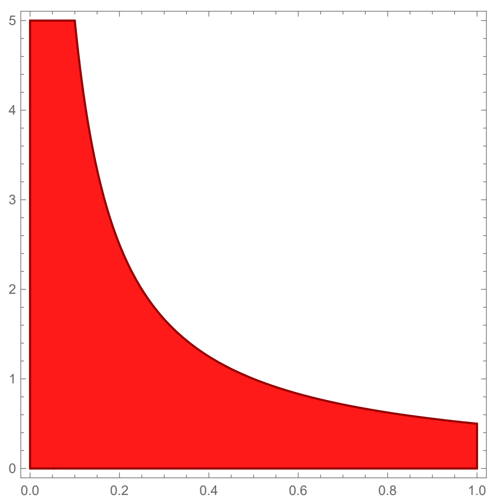

Figura 5.2: Dominio de parámetros del método de Newton asociado al Teorema de Kantorovich 4.1.

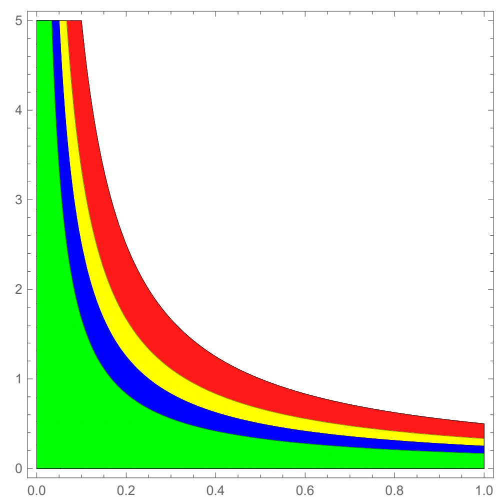

Figura 5.3: Dominio de parámetros del método de Newton asociado al Teorema 5.3.

Se observa que la accesibilidad del método de Newton asociada al Teorema de Kantorovich 4.1 es mejor que la asociada al Teorema 5.1, y que la accesibilidad del Teorema 5.1 mejora cuanto menor es el valor de $\mu = \frac{\tilde{L}}{L}$.

5.4. Convergencia local del método de Newton consecuencia del resultado principal

Si en el Teorema 5.1 de dominio de convergencia global tomamos el caso particular $\tilde{\mathbf{x}} = \mathbf{x}^*$, siendo $\mathbf{x}^*$ una solución del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, resulta que las condiciones del teorema se imponen sobre la función $\mathbf{f}$ y la solución $\mathbf{x}^*$, obteniéndose un resultado de convergencia local para el método de Newton.

Antes de enunciar el teorema, observemos que al ser $\mathbf{x}^*$ solución del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ se tiene que $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$. Por lo tanto, la condición *ii*) del Teorema 5.1 resulta

$$\|J_{\mathbf{f}}(\mathbf{x}^*)^{-1}\mathbf{f}(\mathbf{x}^*)\| = 0 = \eta,$$

lo que implica que la condición (5.3)

$$L\beta\eta = 0 \leqslant \frac{1}{2(1+2\mu)}$$

con $\mu = \frac{\tilde{L}}{L} \in (0, 1]$ siempre se verifica. Teniendo esto en cuenta, el teorema de convergencia local para el método de Newton se reduce al siguiente:

Teorema 5.4 *Sea Ω abierto y convexo de $\mathbb{R}^n$ y sea $\mathbf{f} : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ una aplicación de clase $C^{(1}(\Omega)$. Supongamos que $\mathbf{x}^*$ es una solución del sistema de ecuaciones $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ tal que existe $J_{\mathbf{f}}(\mathbf{x}^*)^{-1}$ y que se verifican las condiciones:*

i) $\|J_{\mathbf{f}}(\mathbf{x}^*)^{-1}\| \leqslant \beta$,

ii) $\|J_{\mathbf{f}}(\mathbf{x}) - J_{\mathbf{f}}(\mathbf{y})\| \leqslant L\|\mathbf{x} - \mathbf{y}\|, \qquad L \in \mathbb{R}^+, \forall\, \mathbf{x}, \mathbf{y} \in \Omega.$

iii) existe $R > 0$ tal que $B(\mathbf{x}^*, R) \subset \Omega$.

Entonces, la sucesión de Newton $\left\{\mathbf{x}^{(\mathbf{n})}\right\}_{n=0}^{\infty}$ dada por

$$\mathbf{x}^{(\mathbf{n+1})} = \mathbf{x}^{(\mathbf{n})} - J_{\mathbf{f}}\left(\mathbf{x}^{(\mathbf{n})}\right)^{-1}\mathbf{f}\left(\mathbf{x}^{(\mathbf{n})}\right),$$

empezando en cualquier $\mathbf{x}^{(\mathbf{0})} \in B(\mathbf{x}^*, R)$, converge a la solución $\mathbf{x}^*$ del sistema de ecuaciones $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, las iteraciones $\mathbf{x}^{(\mathbf{n})}$ pertenecen a $B(\mathbf{x}^*, R)$ y la solución $\mathbf{x}^*$ es única en la región $B(\mathbf{x}^*, r) \cap \Omega$, donde $r = \frac{2}{L\beta} - R$.

Demostración.
Es un caso particular del Teorema 5.1 con $\tilde{\mathbf{x}} = \mathbf{x}^*$. □

Notemos que, al igual que pasaba en el Teorema 5.1, este resultado no sólo asegura la convergencia del método de Newton si no que, además, proporciona un dominio de convergencia $B(\mathbf{x}^*, R)$ en el que cualquier $\mathbf{x}^{(\mathbf{0})}$ perteneciente a él puede ser tomado como punto de inicio para el método de Newton. Sin embargo, mientras el Teorema 5.1 prueba la existencia de una solución $\mathbf{x}^*$ del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, en el teorema que acabamos de mostrar, suponemos la existencia de una solución del sistema $\mathbf{f}(\mathbf{x}) = \mathbf{0}$.

Conclusiones

El método de Newton constituye una herramienta eficaz para la resolución de sistemas de ecuaciones no lineales. Destaca por su sencillez tanto en su aplicación como en su programación, siendo por ello frecuentemente utilizado. En este trabajo hemos podido ver que, además, existen teoremas de convergencia que, sin ser excesivamente difíciles de aplicar, son de utilidad no sólo para asegurar que el método va a converger a una solución del sistema, sino también para obtener otros resultados, como son la existencia de solución, determinar las regiones (generalmente bolas) en las que se encuentran las sucesivas iteraciones o en las que pueda asegurarse la unicidad de la solución o hallar dominios de convergencia. Así, en función del tipo de convergencia, se tiene que:

- Los teoremas de convergencia local exigen condiciones sobre la función $\mathbf{f}$ y la solución $\mathbf{x}^*$, lo cual es una desventaja al ser ésta, en general, desconocida. No obstante, es posible determinar un dominio de convergencia, como hemos visto en el Teorema 5.4.

- Los teoremas de convergencia semilocal exigen condiciones a la función $\mathbf{f}$ y al punto de partida $\mathbf{x}^{(0)}$ pero no a la solución $\mathbf{x}^*$, por lo que pueden ser vistos también como teoremas de existencia de solución.

- Los teoremas de convergencia global exigen condiciones a la función $\mathbf{f}$ y no exigen ninguna condición al punto de partida $\mathbf{x}^{(0)}$ o a la solución $\mathbf{x}^*$. Son resultados muy fuertes puesto que, además de probar la existencia de solución, nos indican que puede tomarse como punto inicial $\mathbf{x}^{(0)}$ cualquier punto del dominio de la función.

A lo largo del trabajo se han ido poniendo de manifiesto diferentes formas en las que se puede estudiar la convergencia de los métodos iterativos. En particular, se trata de encontrar resultados que se pueden aplicar a situaciones en los que los ya conocidos fracasan o que sean más generales. De ahí, el interés permanente por cambiar las condiciones de convergencia a la hora de estudiar la convergencia de los métodos iterativos.

El objetivo del TFG consiste en obtener un resultado de convergencia global, aunque sea restringida (particularmente a una bola), que además nos permita obtener resultados de convergencia local y semilocal, tal y como hemos visto en el último capítulo. Para ello, se utiliza una variación con respecto al Teorema de Kantorovich, la de sustituir que la derivada segunda de la función esté acotada por una condición Lipschitz-centrada en un punto auxiliar.

Bibliografía

[1] Ypma, Tjalling J., *Historical Development of the Newton-Raphson Method.* (1995). Mathematics. 93.

[2] Jarauta Bragulat, E., *Análisis matemático de una variable. Fundamentos y aplicaciones.* 45-47 (2000).

[3] Vera, G. *Lecciones de análisis matemático II.* (2011).

[4] Ezquerro, J.M., Hernández, M.A.,*Estrategias para mejorar la aplicabilidad de métodos iterativos que utilizan diferencias divididas.* Universidad de La Rioja, Servicio de Publicaciones (2014).

[5] Ezquerro, J.M., Gutiérrez, J.M., Hernández, M.A., Romero, N.,Rubio, M.J.,*El método de Newton: de Newton a Kantorovich.* La Gaceta de la RSME, Vol. 13, Núm 1, 53-76 (2010).

[6] Ezquerro, J.M., Hernández-Verón, M.A.,*Domains of global convergence for Newton's method from auxiliary points.* Applied Mathematics Letters 85, 48-56 (2018).

[7] Kantorovich, L.V.,*On Newton's method for functional equations.* Dokl Ajad, Nauk SSSR 59, 1237-1240 (1948).

[8] Kantorovich, L.V.,*The majoral principle and Newton's method.* Dokl Ajad, Nauk SSSR 76, 17-20 (1951).

[9] Kantorovich, L.V., Akilov, G.P.,*Functional Analysis.* Pergamon Press, Oxford (1982).

[10] Chandrasekar, S.,*Radiative transfer.* Dover, Nueva York (1960).

[11] Polyanin, A.D.,Manzhirov, A.V.,*Handbook of integral equations.* CRC Press, Boca Raton (1998).