

TESIS DOCTORAL

El método de Newton
en espacios de Banach

José Manuel Gutiérrez Jiménez



UNIVERSIDAD DE LA RIOJA

TESIS DOCTORAL

El método de Newton
en espacios de Banach

José Manuel Gutiérrez Jiménez

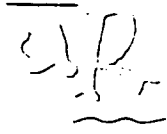
Universidad de La Rioja
Servicio de Publicaciones
2006

Esta tesis doctoral, dirigida por el doctor D. Miguel Ángel Hernández Verón, fue leída el 26 de mayo de 1995, y obtuvo la calificación de Apto Cum Laude Unanimidad.

© José Manuel Gutiérrez Jiménez

Edita: Universidad de La Rioja
Servicio de Publicaciones

ISBN 84-689-9234-8



Universidad de La Rioja
Dpto. Matemáticas y Computación

EL MÉTODO DE NEWTON
EN
ESPACIOS DE BANACH

por

José Manuel Gutiérrez Jiménez

Logroño, Mayo de 1995

EL MÉTODO DE NEWTON
EN
ESPACIOS DE BANACH

por

José Manuel Gutiérrez Jiménez

Memoria presentada para optar al grado de Doctor en Ciencias Matemáticas. Realizada bajo la dirección del Dr. D. Miguel A. Hernández Verón.

Agradecimientos

Quiero agradecer a los doctores Miguel Angel Hernández Verón y M. Amparo Salanova Martínez su colaboración en la realización de este trabajo. Sin su ayuda, este proyecto no se hubiera hecho realidad.

Asimismo, agradezco a los doctores José J. Guadalupe Hernández y Juan Luis Varona Malumbres las múltiples observaciones y correcciones que me han hecho, las cuales han permitido mejorar esta memoria en muchos aspectos. Hago extensivo este agradecimiento al resto de compañeros del Departamento de Matemáticas y Computación de la Universidad de La Rioja, donde nunca me ha faltado el apoyo para realizar este trabajo.

Por último, no me quiero olvidar de los profesores del área de Análisis Matemático de la Universidad de Zaragoza, donde inicié mi investigación.

a mis padres...

Introducción

Uno de los problemas que se presentan con mayor frecuencia a los matemáticos es el de resolver una ecuación. Con una misma notación, $F(x) = 0$, nos podemos referir a situaciones tan distintas como encontrar las raíces de un polinomio, resolver un sistema de ecuaciones o hallar la solución de una ecuación diferencial. Pero este problema no siempre resulta sencillo de resolver. En los casos en que no somos capaces de encontrar la solución exacta, nos conformamos con obtener una aproximación de dicha solución. Es aquí cuando adquieren interés los procesos iterativos para resolver una ecuación. A partir de un valor inicial x_0 más o menos cercano a la solución x^* de $F(x) = 0$, obtenemos un nuevo valor x_1 que se espera sea mejor aproximación de la solución que x_0 . A partir de x_1 , se consigue otra mejor aproximación, x_2 , y así sucesivamente construimos una sucesión $x_0, x_1, x_2, \dots, x_n, \dots$ que cumple que cada uno de los términos es mejor aproximación de la solución que el anterior.

Uno de los métodos iterativos más utilizados para resolver ecuaciones no lineales es el conocido como método de Newton o método de la tangente. En el caso de considerar ecuaciones de la forma $f(t) = 0$, con f una función real de variable real, el método de Newton tiene una clara interpretación geométrica. Dado un valor inicial t_0 , la siguiente aproximación se obtiene calculando la intersección de la recta tangente a f en el punto $(t_0, f(t_0))$ con el eje de abscisas. Se obtiene así

$$t_1 = t_0 - \frac{f(t_0)}{f'(t_0)}.$$

siempre que $f'(t_0) \neq 0$, y en general, reiterando el proceso,

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)}, \quad n \geq 0.$$

lo cual se puede hacer siempre que $f'(t_n) \neq 0$.

La función f y el punto de salida t_0 deben verificar una serie de condiciones para que la sucesión obtenida sea convergente a una solución de $f(t) = 0$, como se

puede consultar en cualquiera de los numerosos textos que sobre este tema hay en la bibliografía matemática, y entre los que citamos los de J.F. Traub [103], A.M. Ostrowski [83] y J. Baranger [17].

El propio Ostrowski, en [83], generalizó este método a sistemas de ecuaciones, y L.V. Kantorovich [64], [65], [66], a ecuaciones dadas a partir de un operador definido en espacios de Banach, obteniendo así una mayor generalidad a la hora de aplicar sus resultados. El caso en que realmente tiene interés aplicar el método de Newton es cuando el operador F es no lineal, ya que para operadores lineales existe una teoría más simple para solucionar el problema [92], [105]. Apoyándose en la teoría desarrollada para operadores lineales, L.B. Rall [92] introduce el método de Newton en espacios de Banach. Observar que lo que se hizo en el caso real para aproximar la solución de una ecuación no lineal $f(t) = 0$ fue resolver una ecuación lineal (obtenida igualando la recta tangente a cero) que en cierto sentido aproxima a la ecuación no lineal. En espacios de Banach haremos básicamente lo mismo.

Así, para operadores F diferenciables, por el desarrollo de Taylor de F en torno a un punto, tenemos que

$$F(x) = F(x_0) + F'(x_0)(x - x_0) + \eta(x, x_0),$$

con $\|\eta(x, x_0)\| \rightarrow 0$ cuando $x \rightarrow x_0$.

Entonces, si x^* es solución de $F(x) = 0$, se cumple

$$F(x^*) = F(x_0) + F'(x_0)(x^* - x_0) + \eta(x^*, x_0).$$

Si x^* está cerca de x_0 , despreciando la cantidad $\eta(x^*, x_0)$, obtenemos una ecuación lineal

$$F(x_0) + F'(x_0)(x - x_0) = 0.$$

cuya solución

$$x_1 = x_0 - [F'(x_0)]^{-1}F(x_0).$$

esperamos que sea una mejor aproximación a la solución que x_0 .

Repitiendo este proceso, se obtiene una sucesión $\{x_n\}$ conocida como sucesión de Newton.

El problema que se plantea a continuación es el de garantizar la convergencia de dicha sucesión a una solución de la ecuación $F(x) = 0$. Citamos aquí tres técnicas diferentes de abordar el problema.

La primera de ellas, debida a Kantorovich [64], [65], [66], establece, mediante relaciones de recurrencia y el empleo de sucesiones mayorizantes, condiciones para

la convergencia de la sucesión de Newton a una solución de $F(x) = 0$: además se garantiza la existencia y unicidad de dicha solución en un determinado dominio. Posteriormente, otros autores (Dennis [34], Miel [73], Ortega [80], Ostrowski [82], Tapia [102]) han dado versiones modificadas del resultado de Kantorovich.

W.C. Rheinboldt en [94] extrae la idea esencial de la teoría mayorizante introducida por Kantorovich y la aplica a las ecuaciones en diferencias para obtener resultados generales, que incluyen como casos particulares los de Kantorovich y el Teorema del Punto Fijo.

La tercera de las técnicas para estudiar la convergencia del método de Newton se debe a F.A. Potra y V. Pták [85], que utilizan la inducción no discreta (método desarrollado por el propio Pták [86]) para obtener como consecuencia resultados de convergencia y estimaciones del error para el método de Newton. Este método es utilizado posteriormente por otros autores para estudiar métodos parecidos al de Newton (Moret [75]) o el método de Newton con condiciones más débiles que las exigidas por Potra y Pták (Argyros [10]).

Otro de los puntos destacados en el estudio del método de Newton es el análisis del error cometido en cada iteración. La mayoría de los autores citados anteriormente dan sus propias estimaciones del error, obtenidas con las distintas hipótesis que manejan. Habitualmente, las estimaciones del error para la sucesión en el espacio de Banach se obtienen a partir de una sucesión real mayorizante. Destacamos aquí el trabajo de W.B. Gragg y R.A. Tapia [49], en el que se establecen estimaciones intermedias entre la sucesión en el espacio de Banach y su mayorizante. También es destacable, por el estudio comparativo de numerosas cotas del error obtenidas por distintos autores, el trabajo de T. Yamamoto [106].

En esta memoria desarrollaremos fundamentalmente la técnica de Kantorovich, pero con un nuevo enfoque. El trabajo realizado es mayoritariamente teórico, aunque algunos resultados aparecen ilustrados con ejemplos. Las tablas numéricas que aparecen en algunos de estos ejemplos han sido calculadas con ayuda del programa *Mathematica* (versión 2.2). Para una mayor información sobre el método de Newton desde el punto de vista computacional pueden consultarse los textos de O. López [71], M. Gasca [45] o R.E. Scraton [99], entre otros.

La presente memoria está dividida en cinco capítulos. Se ha intentado conseguir una memoria autocontenida. Así los resultados que se utilizan son citados con anterioridad, aunque en alguno de los casos se haya omitido la demostración, citando, eso sí, otras fuentes donde consultarlos. El primer capítulo es básicamente introductorio. Se comentan algunas de las propiedades de los espacios de Banach.

destacando los resultados que más nos van a interesar, tales como los relativos a inversión de operadores, diferenciabilidad o integrabilidad. Para funciones reales, se define el concepto de función logarítmicamente convexa y el de grado de convexidad logarítmico, que mide este tipo de convexidad. La extensión formal del grado de convexidad logarítmico de una función a espacios de Banach es un operador lineal, que denotaremos L_F . A este operador y, por analogía con el caso real, lo seguiremos llamando grado de convexidad logarítmico. No obstante, ha de quedar muy claro que en esta situación no se trata de una medida de la convexidad, sino que es simplemente un operador lineal cuyas propiedades nos van a permitir estudiar el método de Newton y otros procesos iterativos desde un nuevo punto de vista.

A continuación se analiza el método de Newton para resolver ecuaciones escalares. Se da una idea geométrica de la construcción del método y se comentan los principales resultados sobre la convergencia y el orden del mismo. También se pone de manifiesto la relación entre el grado de convexidad logarítmico de una función y la velocidad de convergencia de dicho proceso iterativo (ver Hernández [53]).

Finalmente, se estudia el método de Newton en espacios de Banach: se introduce el concepto de sucesión mayorizante y se destaca su importancia como herramienta para estudiar sucesiones definidas en dichos espacios. A continuación nos centramos en el método de Newton y, en concreto, en uno de los resultados clásicos sobre dicho método: El Teorema de Kantorovich. Este resultado proporciona condiciones para la convergencia de la sucesión de Newton a la raíz de la ecuación considerada, así como delimita regiones donde se puede garantizar la existencia y unicidad de solución, dando también estimaciones del error cometido en cada iteración. Como ya hemos comentado, el Teorema de Kantorovich ha sido tema de estudio por parte de diversos autores. En esta memoria enunciamos el resultado tal y como el propio autor escribe en [66], aunque también se citan versiones más recientes. Terminamos el capítulo mostrando la aplicación del método de Newton para resolver ecuaciones diferenciales e integrales.

En el Capítulo II se obtiene la derivada del operador

$$G(x) = x - [F'(x)]^{-1}F(x)$$

definido entre espacios de Banach, y se comprueba que esta derivada coincide con el operador lineal $L_F(x)$ que extiende el concepto de grado de convexidad logarítmico introducido en el Capítulo I. Posteriormente se utiliza este hecho para, a partir de técnicas similares a las del Teorema del Punto Fijo, dar nuevas condiciones de convergencia, en términos del operador L_F . Estas condiciones nos permiten

obtener unos dominios de puntos de salida, para los cuales la sucesión de Newton es convergente, más amplios que los permitidos por las condiciones de Kantorovich. Ilustramos este resultado mediante unos ejemplos al final del capítulo.

El interés del Capítulo III está en que a partir de un operador F que satisface las condiciones de Kantorovich, se construye, utilizando el grado de convexidad logarítmico y sus propiedades, la función real que nos proporciona la mejor estimación del error. Es decir, dado el operador F , construimos la sucesión de Newton $\{x_n\}$ para resolver la ecuación $F(x) = 0$. Entonces, hay infinitas funciones f que "mayorizan" a F , esto es, la correspondiente sucesión de Newton para resolver $f(t) = 0$ es una sucesión mayorizante de $\{x_n\}$. Cada una de estas funciones nos proporciona una estimación del error distinta; de entre todas ellas, obtenemos la que nos proporciona la mejor estimación. Esta función resulta ser un polinomio de segundo grado, y coincide con el que Kantorovich y otros autores emplean para mayorizar la sucesión de Newton.

La última sección de este capítulo se dedica a estudiar un resultado sorprendente acerca de estimaciones del error, debido a Gragg y Tapia [49]. Estos autores lograron obtener unas estimaciones intermedias entre la sucesión de Newton en un espacio de Banach y su mayorizante, así como acotaciones inferiores para la primera. Posteriormente, con la ayuda de un resultado de Miel [72], nosotros obtenemos nuevas cotas superiores que, aunque peores que las de Gragg y Tapia, son más sencillas de calcular desde el punto de vista operacional.

En el Capítulo IV se trabaja en dos direcciones bien diferenciadas. En la primera de ellas se utiliza el hecho de que si, dados un operador F y una función f , se verifica

$$\|L_F(x)\| \leq L_f(t), \quad \|x - x_0\| \leq t - t_0,$$

entonces, la sucesión obtenida aplicando el método de Newton a f es mayorizante de la correspondiente sucesión obtenida para F . Se encuentran así resultados de convergencia en condiciones distintas a las de Kantorovich, en función de las acotaciones que se tengan para el operador L_F . Por ejemplo, se estudia el caso de que

$$\|L_F(x)\| \leq M < 1,$$

y se encuentran funciones reales con grado de convexidad logarítmico constante. Se establecen así nuevos resultados de convergencia que tienen el inconveniente de proporcionar una estimación del error muy poco fina, ya que la función real obtenida tiene raíces múltiples, y es conocido que la convergencia del método de Newton en este caso es muy lenta.

Por otra parte, se estudia la convergencia del método de Newton en condiciones más fuertes que las de Kantorovich, por ejemplo, condiciones tipo Lipschitz sobre la derivada segunda del operador F , es decir

$$\|F''(x) - F''(y)\| \leq k\|x - y\|,$$

condiciones de continuidad tipo Hölder

$$\|F''(x) - F''(y)\| \leq k\|x - y\|^p, \quad 0 < p < 1,$$

o condiciones que engloban a las anteriores

$$\|F''(x) - F''(y)\| \leq k\|x - y\|^p, \quad 0 \leq p.$$

En todos estos casos, y a partir del operador L_F , se construye la función que nos dará la mejor estimación del error, dándose expresiones cómodas de éste en función de las raíces de dicha función, de forma similar a la que emplea Ostrowski [82] para polinomios de segundo grado.

Daremos ejemplos en los que la convergencia de la sucesión de Newton no queda garantizada con las condiciones clásicas de Kantorovich y, sin embargo, sí se garantiza con estas nuevas condiciones [62]. También se analizan las estimaciones del error obtenidas en las nuevas situaciones, comparándolas con las ya conocidas. Por último, se aplican estos resultados a algunos tipos de ecuaciones diferenciales e integrales, obteniendo para ellos resultados acerca de existencia y unicidad de solución.

En el Capítulo V, teniendo en cuenta la influencia de la convexidad en la velocidad de convergencia del método de Newton conocida para el caso real, se construye una aceleración para dicho método y se define un nuevo proceso iterativo de tercer orden, que llamaremos aceleración convexa del método de Newton. Se estudia este método en el caso escalar, dándose condiciones para la convergencia, expresiones del error en función de las raíces y algunas propiedades interesantes.

A continuación se extiende este método a espacios de Banach y se dan resultados de convergencia en distintas condiciones. Se concluye el capítulo con unos ejemplos donde se compara este método con otros de tercer orden conocidos como son el método de Chebyshev [58] y el de Halley [56], [109], comprobando experimentalmente que, siendo un método con el mismo orden, tiene mayor velocidad de convergencia que aquéllos. Este resultado fue comprobado para funciones escalares por Hernández y Salanova [55].

No queremos terminar esta introducción sin destacar una de las cualidades más interesantes de este trabajo a nuestro entender, como es el hecho de abrir nuevas

vías en el estudio de procesos iterativos en espacios de Banach mediante el operador grado de convexidad logarítmico. Este operador, como ya pusieron de manifiesto Ehrmann [42] y Gander [43] juega un papel fundamental en los procesos iterativos escalares de tercer orden, ya que cualquier método de tercer orden puede escribirse en función de dicho operador. En espacios de Banach, este operador simplifica bastante el trabajo de obtención de condiciones de convergencia, como se pone de manifiesto en el Capítulo V. Así, se abre la posibilidad de un nuevo estudio de los procesos iterativos de orden tres conocidos (Chebyshev, Halley), de nuevos procesos iterativos o, incluso, familias de procesos.

Otra vía que queda abierta es la de analizar la posible influencia del grado de convexidad logarítmico en la velocidad de convergencia de métodos definidos en espacios de Banach. Para ello consideramos espacios dotados de una estructura de orden parcial, donde poder definir la noción de operador convexo, [40], [63], [81], [104]. Así, se puede pensar en extender a estos espacios los resultados obtenidos por Hernández [53] para el caso real.

Los ejemplos sobre ecuaciones diferenciales e integrales que ilustran algunos resultados de esta memoria nos hacen pensar en que podemos obtener resultados más generales sobre existencia y unicidad de soluciones empleando estas nuevas técnicas: resultados que se pueden extender a otros tipos de ecuaciones, como ecuaciones en derivadas parciales o problemas de cálculo de valores y vectores propios en espacios de Banach [66].

En esta memoria se estudia el método de Newton empleando fundamentalmente la técnica desarrollada por Kantorovich, mediante el empleo de sucesiones mayorizantes. Como ya se ha comentado con anterioridad, no es esta la única manera de abordar el problema. Sería interesante analizar las técnicas utilizadas por Rheinboldt [94] (ecuaciones en diferencias) y Potra y Pták [85] (inducción no discreta) con la ayuda del grado de convexidad logarítmico. Por último y, aunque no ha sido apenas mencionada en esta memoria, la aplicación numérica de estos métodos también puede tener su interés. Para el problema general de resolver una ecuación en espacios de Banach por distintos métodos, hay que ser consciente de que el método de Newton es preferible a otros procesos iterativos de orden superior, ya que conlleva la inversión de un sólo operador en cada caso (incluso se consideran modificaciones de dicho método para las que no se tenga que hacer una inversión en cada paso). No obstante, en algunos casos, la aceleración de la convergencia puede justificar el coste operacional realizado para la inversión de operadores, si esta es fácilmente computable.

Índice

1	Preliminares	1
1.1	Operadores en espacios de Banach	1
1.2	El grado de convexidad logarítmico y el operador L_F	8
1.3	La convexidad y el método de Newton	12
1.4	El método de Newton en espacios de Banach. Teoremas de Kantorovich	16
1.5	Ejemplo de aplicación del método de Newton	22
2	Dominios de puntos de salida para el método de Newton	27
2.1	Introducción	27
2.2	El operador L_F y el método de Newton	28
2.3	Condiciones de convergencia para el método de Newton	32
2.4	Ampliación del dominio de puntos de salida	36
3	Estudio del error en el método de Newton	47
3.1	Introducción	47
3.2	Estudio del error en las condiciones de Kantorovich	48
3.3	Técnica de Gragg y Tapia	61
4	Nuevas sucesiones mayorizantes	75
4.1	Introducción	75
4.2	Sucesiones mayorizantes y convexidad	76
4.3	Sucesiones mayorizantes bajo condiciones de Kantorovich fuertes	84
4.4	Aplicación a las ecuaciones diferenciales e integrales	102
5	La aceleración convexa del método de Newton	107
5.1	Introducción	107
5.2	La aceleración convexa en espacios de Banach	115
5.3	Otro estudio de la convergencia para la aceleración convexa	131
5.4	Ejemplos	136

Capítulo 1

Preliminares

1.1 Operadores en espacios de Banach

Sea $F : X \rightarrow Y$ un operador definido entre dos espacios de Banach X e Y . El objetivo de esta memoria es estudiar la ecuación $F(x) = 0$, bajo ciertas condiciones sobre el operador F .

Comentaremos algunas nociones que serán utilizadas en el desarrollo posterior de esta memoria. Para un estudio más detallado de los espacios de Banach se puede consultar cualquiera de los numerosos tratados sobre el tema que hay en la bibliografía matemática, entre los que citamos los de Beauzamy [19], Berberian [20], Curtain-Pritchard [30], Day [33], Dunford-Schwarz [41], Kreyszig [67], Lindenstrauss-Tzafriri [69], [70] y Rudin [97]. Unos textos muy recomendables, ya que introducen los espacios de Banach con un objetivo similar al nuestro son los de Kantorovich [66] y Rall [92].

Al conjunto de los operadores lineales y acotados entre los espacios X e Y lo denotaremos a partir de ahora $\mathcal{L}(X, Y)$. Si $X = Y$ denotaremos simplemente $\mathcal{L}(X)$. Sea $L \in \mathcal{L}(X, Y)$. Denotamos $\mathcal{D}(L)$ a su dominio y $\mathcal{R}(L)$ a su rango. Si existe un operador L^{-1} que aplica $\mathcal{R}(L)$ en $\mathcal{D}(L)$, de manera que

$$L^{-1}Lx = x, \quad \text{para todo } x \in \mathcal{D}(L),$$

$$LL^{-1}y = y, \quad \text{para todo } y \in \mathcal{R}(L),$$

diremos que L es *invertible* y que L^{-1} es el *inverso* de L .

Si L^{-1} existe, entonces también es un operador lineal.

A continuación citamos unos resultados conocidos sobre inversión de operadores. Observar que estos resultados están dados para operadores de un espacio X en sí mismo. Para ver otros resultados más generales sobre inversión de operadores, pueden consultarse los textos de Nashed [76], [77], [78].

Lema 1.1. (Banach) Sea $L \in \mathcal{L}(X)$ verificando que

$$\|L\| \leq k < 1.$$

entonces el operador $I - L$ tiene inverso continuo y además

$$\|(I - L)^{-1}\| \leq \frac{1}{1 - k}.$$

Unas ligeras variantes del lema anterior son los siguientes resultados.

Lema 1.2. Sea $L \in \mathcal{L}(X)$. Existe L^{-1} si y sólo si existe un operador invertible $M \in \mathcal{L}(X)$ tal que

$$\|I - ML\| < 1.$$

En este caso,

$$L^{-1} = \sum_{n=0}^{\infty} (I - ML)^n M$$

y

$$\|L^{-1}\| \leq \frac{\|M\|}{1 - \|I - ML\|}.$$

Lema 1.3. Sea $L \in \mathcal{L}(X)$. Existe L^{-1} si y sólo si existe un operador invertible $M \in \mathcal{L}(X)$ tal que

$$\|M - L\| < \frac{1}{\|M^{-1}\|}.$$

En este caso,

$$L^{-1} = \sum_{n=0}^{\infty} (I - M^{-1}L)^n M^{-1}$$

y

$$\|L^{-1}\| \leq \frac{\|M^{-1}\|}{1 - \|I - M^{-1}L\|} \leq \frac{\|M^{-1}\|}{1 - \|M^{-1}\|\|M - L\|}.$$

Presentamos ahora algunos conceptos y resultados básicos del cálculo diferencial e integral en espacios de Banach.

Definición 1.4. Dado $x_0 \in X$, si existe un operador $L \in \mathcal{L}(X, Y)$ de manera que, para todo $x \in X$,

$$\lim_{h \rightarrow 0} \frac{F(x_0 + hx) - F(x_0)}{h} = L(x), \quad (1.1)$$

entonces F se dice *diferenciable Gâteaux* (o diferenciable débilmente) en x_0 . En esta situación, el operador lineal L es la *derivada de Gâteaux* (o la *diferencial de Gâteaux*) de F en x_0 , y se denota $L = dF(x_0)$.

Definición 1.5. Si el límite de la ecuación (1.1) es uniforme en el conjunto $\{x \in X: \|x\| = 1\}$, entonces F se dice *diferenciable Fréchet*, o simplemente diferenciable en x_0 . En este caso, el operador lineal L se llama *derivada de Fréchet* (o *diferencial de Fréchet*) de F en x_0 , y se denota $L = F'(x_0)$.

Equivalentemente, el concepto de diferenciabilidad se puede expresar de la siguiente forma. Si dado $x_0 \in X$, existe un operador lineal y continuo $L = F'(x_0)$ de manera que

$$\lim_{\|v\| \rightarrow 0} \frac{\|F(x_0 + v) - F(x_0) - Lv\|}{\|v\|} = 0,$$

entonces F es diferenciable Fréchet en x_0 .

Observar que si F es un operador diferenciable Fréchet en x_0 , entonces también es diferenciable Gâteaux en x_0 . Además $F'(x_0) = dF(x_0)$. Salvo que se indique lo contrario, cuando nos referimos a un operador diferenciable, lo entenderemos en el sentido de Fréchet.

Es bien conocido que las propiedades de la diferencial en espacios de Banach son análogas a las de la derivada en el caso escalar, salvo el Teorema del Valor Medio. El siguiente resultado generaliza el Teorema del Valor Medio conocido para funciones escalares.

Teorema 1.6. (del Valor Medio) Sean X e Y dos espacios de Banach. Sea $F : X \rightarrow Y$ un operador diferenciable en un conjunto convexo $\Omega \subseteq X$. Entonces, si $x_0, x_1 \in \Omega$, se tiene

$$\|F(x_1) - F(x_0)\| \leq \sup_{0 < \theta < 1} \|F'(x_0 + \theta(x_1 - x_0))\| \|x_1 - x_0\|.$$

Corolario 1.7. Con la notación anterior se tiene que

$$\begin{aligned} & \|F(x_1) - F(x_0) - F'(x_0)(x_1 - x_0)\| \\ & \leq \sup_{0 < \theta < 1} \|F'(x_0 + \theta(x_1 - x_0)) - F'(x_0)\| \|x_1 - x_0\|. \end{aligned}$$

A continuación comentamos algunos aspectos sobre el cálculo integral en espacios de Banach.

Definición 1.8. Sea F definida en un intervalo real $[a, b]$ y con valores en un espacio de Banach Y . Entonces podemos definir su integral como el límite de las siguientes sumas

$$\sum_{k=0}^{n-1} F(\tau_k)(t_{k+1} - t_k)$$

donde $a = t_0 < t_1 < \dots < t_n = b$ y $\tau_k \in [t_k, t_{k+1}]$, cuando $\max\{t_{k+1} - t_k\}$ tiende a cero. Si el límite dirigido anterior existe lo llamaremos *integral* de F y lo denotaremos

$$\int_a^b F(t) dt.$$

Evidentemente, si la integral existe, es un elemento de Y . Una condición suficiente para que exista dicha integral es que la función F sea continua en el intervalo $[a, b]$.

Las propiedades de esta integral se deducen de las conocidas para la integral de Riemann en el caso real. De entre ellas, destacamos tres por su posterior utilidad.

(i) Si consideramos un operador lineal y acotado $L \in \mathcal{L}(Y, Z)$, entonces

$$\int_a^b L(F(t)) dt = L\left(\int_a^b F(t) dt\right).$$

(ii) Si $F(t) = \phi(t)y_0$, donde y_0 es un elemento fijo de Y y ϕ es una función real integrable, entonces

$$\int_a^b F(t) dt = y_0 \int_a^b \phi(t) dt.$$

(iii) $\left\| \int_a^b F(t) dt \right\| \leq \int_a^b \|F(t)\| dt.$

Definición 1.9. Supongamos ahora que T es un operador definido en un segmento $[x_0, x_1] \subseteq X$ y con valores en el espacio $\mathcal{L}(X, Y)$. En este caso, se define su integral

$$\int_{x_0}^{x_1} T(x) dx = \int_0^1 T(x_0 + t(x_1 - x_0))(x_1 - x_0) dt.$$

Notar que, en este caso, la función $T(x_0 + [\cdot](x_1 - x_0))(x_1 - x_0)$ está en las condiciones de la Definición 1.8, con $[a, b] = [0, 1]$.

Los siguientes resultados sobre cálculo integral en espacios de Banach pueden consultarse en los libros de Kantorovich [66] o Rall [92].

Teorema 1.10. Sea F un operador de X en Y que tiene derivada continua en el segmento $[x_0, x_1] \subseteq X$. Entonces

$$\int_{x_0}^{x_1} F'(x) dx = F(x_1) - F(x_0).$$

Lema 1.11. Si se verifica

$$\|F(x_0 + \tau(x_1 - x_0))\| \leq o(t_0 + \tau(t_1 - t_0)), \quad \tau \in [0, 1]$$

y

$$\|x_1 - x_0\| \leq t_1 - t_0.$$

entonces

$$\left\| \int_{x_0}^{x_1} F(x) dx \right\| \leq \int_{t_0}^{t_1} o(t) dt.$$

Como variante del lema anterior, tenemos el siguiente resultado.

Corolario 1.12. Si se cumple la desigualdad

$$\|F(x)\| \leq o(t)$$

para x y t tales que

$$\|x - x_0\| \leq t - t_0.$$

entonces

$$\left\| \int_{x_0}^{x_1} F(x) dx \right\| \leq \int_{t_0}^{t_1} o(t) dt.$$

donde x_1 es un elemento que cumple $\|x_1 - x_0\| \leq t_1 - t_0$.

Si $F : X \rightarrow Y$ es un operador dos veces diferenciable en un punto $x_0 \in X$, entonces $F''(x_0) \in \mathcal{B}(X^2, Y)$, donde $\mathcal{B}(X^2, Y)$ es el conjunto de los operadores bilineales acotados de $X^2 = X \times X$ en Y , con la norma

$$\|B\| = \sup_{\|x_1\|, \|x_2\| \leq 1} \|B(x_1, x_2)\|, \quad B \in \mathcal{B}(X^2, Y).$$

Es conocido, [28], que existe una isometría entre el espacio $\mathcal{B}(X^2, Y)$ y $\mathcal{L}(X, \mathcal{L}(X, Y))$. A partir de ahora, y teniendo en cuenta esta isometría, identificaremos los elementos de ambos espacios. También se sabe, [92], que $F''(x_0)$ es un operador bilineal simétrico, es decir,

$$F''(x_0)x_1x_2 = F''(x_0)x_2x_1 \quad \text{para todo } x_1, x_2 \in X.$$

Para un operador bilineal de $\mathbb{R}^2 \times \mathbb{R}^2$ en \mathbb{R}^2 empleamos la notación dada por una matriz de bloques [12]

$$B = \begin{pmatrix} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22} \end{pmatrix}. \quad (1.2)$$

Entonces, si $x = (x_1, x_2)$, $y = (y_1, y_2)$, se tiene que

$$\begin{aligned} B(x, y) &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} b_1^{11} & b_1^{12} \\ b_1^{21} & b_1^{22} \\ b_2^{11} & b_2^{12} \\ b_2^{21} & b_2^{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \begin{pmatrix} b_1^{11}x_1 + b_1^{21}x_2 & b_1^{12}x_1 + b_1^{22}x_2 \\ b_2^{11}x_1 + b_2^{21}x_2 & b_2^{12}x_1 + b_2^{22}x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \begin{pmatrix} b_1^{11}x_1y_1 + b_1^{21}x_2y_1 + b_1^{12}x_1y_2 + b_1^{22}x_2y_2 \\ b_2^{11}x_1y_1 + b_2^{21}x_2y_1 + b_2^{12}x_1y_2 + b_2^{22}x_2y_2 \end{pmatrix}. \end{aligned}$$

En \mathbb{R}^2 consideramos la norma del máximo

$$\|(x_1, x_2)\| = \max\{|x_1|, |x_2|\},$$

y en $\mathcal{L}(\mathbb{R}^2, \mathbb{R}^2)$ la norma dada por

$$\|L\| = \max\{|a_{11}| + |a_{12}|, |a_{21}| + |a_{22}|\}.$$

con

$$L = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Entonces se comprueba fácilmente [90] que la norma del operador bilineal B definido por (1.2) viene dada por

$$\|B\| = \sup_{\|x\|=1} \max_i \sum_{j=1}^2 \left| \sum_{k=1}^2 b_i^{jk} x_k \right|.$$

Además se tiene la siguiente estimación de la norma anterior

$$\|B\| \leq \max \{ |b_1^{11}| + |b_1^{12}| + |b_1^{21}| + |b_1^{22}|, |b_2^{11}| + |b_2^{12}| + |b_2^{21}| + |b_2^{22}| \}.$$

El resultado de componer una aplicación bilineal B y una lineal L es una aplicación bilineal LB . Con la notación anterior, tenemos que la matriz asociada a LB es

$$LB = \begin{pmatrix} a_{11}b_1^{11} + a_{12}b_2^{11} & a_{11}b_1^{12} + a_{12}b_2^{12} \\ a_{11}b_1^{21} + a_{12}b_2^{21} & a_{11}b_1^{22} + a_{12}b_2^{22} \\ a_{21}b_1^{11} + a_{22}b_2^{11} & a_{21}b_1^{12} + a_{22}b_2^{12} \\ a_{21}b_1^{21} + a_{22}b_2^{21} & a_{21}b_1^{22} + a_{22}b_2^{22} \end{pmatrix}.$$

Si F es un operador de \mathbb{R}^2 en \mathbb{R}^2 que a un par $(x, y) \in \mathbb{R}^2$ le hace corresponder $(f(x, y), g(x, y)) \in \mathbb{R}^2$, tenemos que la derivada segunda de F en un punto (x_0, y_0) es el operador de $\mathbb{R}^2 \times \mathbb{R}^2$ en \mathbb{R}^2 dado por la matriz de la forma (1.2)

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \\ g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{pmatrix}.$$

donde las derivadas parciales anteriores están evaluadas en el punto (x_0, y_0) .

Finalizamos esta pequeña introducción a los espacios de Banach con un resultado que será de gran utilidad para nosotros, como es el Teorema de Taylor.

Teorema 1.13. (Taylor) *Supongamos que F es un operador n -veces diferenciable en la bola $B(x_0, r)$, $r > 0$, y que $F^{(n)}$ es integrable en el segmento $[x_0, x_1]$ con $x_1 \in B(x_0, r)$. Entonces*

$$F(x_1) = F(x_0) + \sum_{k=1}^{n-1} \frac{1}{k!} F^{(k)}(x_0)(x_1 - x_0)^k + R_n(x_0, x_1)$$

donde

$$\begin{aligned} R_n(x_0, x_1) &= \frac{1}{(n-1)!} \int_{x_0}^{x_1} F^{(n)}(x)(x_1 - x)^{n-1} dx \\ &= \frac{1}{(n-1)!} \int_0^1 F^{(n)}(x_0 + t(x_1 - x_0))(x_1 - x_0)^n (1-t)^{n-1} dt. \end{aligned}$$

1.2 El grado de convexidad logarítmico y el operador L_F

En [44], J. Garay y M.A. Hernández introducen el grado de convexidad logarítmico como una medida puntual de la convexidad de una función y lo utilizan para un problema de localización de neuronas. Esta forma de medir la convexidad surge del Teorema de Bohr-Mollerup [27], para la definición de la función gamma. En este resultado aparece el concepto de función logarítmicamente convexa, es decir, una función cuyo logaritmo es una función convexa. El grado de convexidad logarítmico es una medida puntual de la convexidad, dada por la resistencia de una función a ser "concavizada" por el operador logarítmico, es decir, el número de veces que hay que aplicar el operador logarítmico a una función convexa para obtener como resultado una función cóncava (ver [54], [55] o [95] para más detalles). En [53] se relaciona dicho grado de convexidad con la velocidad de convergencia del método de Newton. A continuación daremos una idea de la construcción de dicha medida de la convexidad.

Sea $f : U \subseteq \mathbb{R} \rightarrow \mathbb{R}$, siendo $U \neq \emptyset$ y no reducido a un punto, una función perteneciente a la clase $C^2(U)$ y convexa en $x_0 \in U$. Además, por simplificar los cálculos, suponemos que $f(x_0) = 1$. Por otra parte, denotamos

$$C_{(x_0, y_0)}^2(U) = \{g \in C^2(U); g(x_0) = y_0\}$$

y consideramos el operador normalización a uno, dado por

$$N : C_{(x_0, 0)}^2(U) \rightarrow C_{(x_0, 1)}^2(U)$$

tal que $N[g](x) = g(x) + 1$.

Entonces definimos las siguientes sucesiones de funciones:

$$\begin{aligned}
F_1(x) &= \log f(x), & G_1(x) &= N[F_1](x), \\
F_2(x) &= \log G_1(x), & G_2(x) &= N[F_2](x), \\
&\dots \\
F_n(x) &= \log G_{n-1}(x), & G_n(x) &= N[F_n](x).
\end{aligned}$$

El estudio de la convexidad en el punto x_0 de las funciones que forman la sucesión $\{F_n\}$ nos va a permitir medir la resistencia de la función f a ser "concajada" mediante sucesivas aplicaciones del operador logaritmo. Teniendo en cuenta que

$$F_n''(x_0) = f''(x_0) - n f'(x_0)^2.$$

resulta sencillo observar que la concavidad o convexidad de las funciones F_n en x_0 queda perfectamente determinada analizando el valor de la expresión

$$\frac{f''(x_0)}{f'(x_0)^2}.$$

Este es el origen de la definición del grado de convexidad logarítmico.

Definición 2.1. El grado de convexidad logarítmico de una función $f \in C_{(x_0,1)}^{(2)}$ en el punto x_0 es el número real dado por la expresión

$$\dot{L}_f(x_0) = \frac{f''(x_0)}{f'(x_0)^2}$$

cuando x_0 no es un mínimo de f , en cuyo caso $\dot{L}_f(x_0) = \infty$.

Notemos que para $f \in C_{(x_0,1)}^{(2)}$, $\dot{L}_f(x_0) \in \mathbb{R}$, y si consideramos funciones convexas, $\dot{L}_f(x_0) \in [0, \infty]$.

En la definición anterior, hemos definido el grado de convexidad para funciones que en el punto x_0 toman el valor uno. Si f toma un valor positivo arbitrario en el punto x_0 , consideraremos la función

$$f^* = \frac{f}{f(x_0)} \in C_{(x_0,1)}^{(2)}$$

dado que $\log f$ y $\log f^*$ tienen la "misma convexidad" en el punto x_0 , ya que se diferencian en una constante.

Definición 2.2. Con la notación anterior, si f es una función perteneciente a la clase $C^2(U)$, siendo U un intervalo que contiene al punto x_0 , definimos el grado de convexidad logarítmico de f en x_0 como el número real dado por la expresión

$$L_f(x_0) = \dot{L}_{f^*}(x_0) = \frac{f(x_0) f''(x_0)}{f'(x_0)^2}$$

si x_0 no es un mínimo de f , en cuyo caso $L_f(x_0) = \infty$.

Nota. Si consideramos una función positiva en un entorno del punto x_0 , es claro que el grado de convexidad logarítmico de una función convexa es un número real positivo, siendo "más convexa" respecto al logaritmo a medida que mayor sea dicho número. Por otra parte, aunque perdemos el significado geométrico del concepto, podemos extender formalmente la definición al espacio $C^2(U)$.

Observar también que el grado de convexidad logarítmico es una medida invariante por dilataciones pero no por traslaciones.

Si bien el concepto de grado de convexidad logarítmico puede considerarse como reciente, y su aplicación al estudio de la velocidad de convergencia de procesos iterativos tenga como origen [54], podemos encontrar algún antecedente en la bibliografía matemática. Así, por ejemplo, B. Neta[79] da la expresión de diversos métodos iterativos de tercer orden en función del grado de convexidad logarítmico, y en el trabajo de M. Altman [4] nos encontramos con la expresión

$$\frac{f(t)f''(t)}{f'(t)^2}.$$

que el autor utiliza para probar la convergencia de un proceso iterativo de orden tres. Además, en dicho trabajo se da una propiedad sobre el grado de convexidad logarítmico que destacamos por su utilidad en capítulos posteriores.

Lema 2.3. (Altman) *Sea f una función que satisface las condiciones $f(t) \geq 0$, $f'(t) \neq 0$ y $f'''(t) \geq 0$ en $[0, t^*]$ y $f(t^*) = 0$. Entonces*

$$L_f(t) < \frac{1}{2}, \quad t \in [0, t^*].$$

Nota. Si exigimos que $f'(t) \neq 0$ en $[0, t^*)$, entonces se tiene que

$$L_f(t) \leq \frac{1}{2}, \quad t \in [0, t^*].$$

La extensión del grado de convexidad logarítmico a funciones reales definidas en espacios de Banach, y en particular a funcionales lineales continuos sigue teniendo sentido geométrico. No ocurre lo mismo entre operadores definidos entre espacios

de Banach en general. Sin embargo, podemos entender el grado de convexidad logarítmico como un operador definido en espacios de Banach, dejando muy claro que ya no se trata de una medida de la convexidad, sino de un operador cuyas propiedades nos servirán para estudiar el método de Newton y otros procesos iterativos.

Dado un operador F entre dos espacios de Banach X e Y , y un punto $x \in X$, tenemos que

$$X \xrightarrow{F} Y \xrightarrow{\Gamma(x)} X$$

siendo $\Gamma(x)$ el operador inverso $\Gamma(x) = [F'(x)]^{-1} \in \mathcal{L}(Y, X)$. Por tanto, $\Gamma(x)F(x)$ es un elemento de X . Teniendo en cuenta la isometría entre operadores bilineales y operadores de $\mathcal{L}(X, \mathcal{L}(X, Y))$, tenemos que $F''(x)\Gamma(x)F(x) \in \mathcal{L}(X, Y)$, y por tanto,

$$L_F(x) : X \xrightarrow{F''(x)\Gamma(x)F(x)} Y \xrightarrow{\Gamma(x)} X.$$

Esto nos lleva a dar la siguiente definición.

Definición 2.4. Sean X e Y dos espacios de Banach y $F : X \rightarrow Y$ un operador con derivada segunda de Fréchet. Sea $x_0 \in X$ de manera que exista $\Gamma(x_0) = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$, entonces definimos el operador $L_F : X \rightarrow \mathcal{L}(X, X)$ que a $x_0 \in X$ le hace corresponder el operador lineal $L_F(x_0) : X \rightarrow X$, que actúa de la siguiente manera:

$$L_F(x_0)(x) = \Gamma(x_0)F''(x_0)\Gamma(x_0)F(x_0)x, \quad x \in X.$$

Observar que este operador es la generalización del grado de convexidad logarítmico definido para funciones reales. Por este motivo, lo seguiremos denominando igual.

Nos hemos dedicado principalmente a estudiar el concepto de función logarítmicamente convexa, así como a dar una medida de dicha convexidad. No se ha pretendido realizar un análisis minucioso del concepto de convexidad, ni de las propiedades que de él se deducen. Para un estudio detallado de la convexidad, pueden consultarse los trabajos de Borweim y Noll [22], Ciesielski [29], Giles [47], Hörmander [60] o Phelps [84], entre otros.

1.3 La convexidad y el método de Newton

El método de Newton, por su simplicidad y efectividad, es uno de los métodos iterativos más conocidos para resolver una ecuación no lineal. La idea original del método se debe a Newton, pero Raphson fue el primero en expresarlo como una fórmula de recurrencia. Por ello, también se denomina método de Newton-Raphson.

Dada una función $f : \mathbb{R} \rightarrow \mathbb{R}$, que satisface unas ciertas condiciones de derivabilidad, el proceso iterativo

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)} \quad (3.1)$$

es conocido como método de Newton para resolver la ecuación $f(t) = 0$.

Ostrowsky [82] generalizó el método anterior a sistemas de ecuaciones no lineales, y posteriormente Kantorovich [66] a ecuaciones en espacios de Banach en general.

Si estudiamos la ecuación

$$F(x) = 0, \quad (3.2)$$

con F un operador entre espacios de Banach, una de las maneras de abordar el problema es la conocida como linealización de la ecuación, esto es, encontrar un operador lineal L de forma que la solución de la ecuación $L(x) = y$ aproxime a la solución x^* de (3.2).

Para encontrar este operador lineal L podemos usar el desarrollo de Taylor de F entorno a un punto. Así, si F es diferenciable podemos escribir

$$F(x) = F(x_0) + F'(x_0)(x - x_0) + \eta(x, x_0)$$

donde

$$\lim_{x \rightarrow x_0} \frac{\|\eta(x, x_0)\|}{\|x - x_0\|} = 0.$$

Como $F(x^*) = 0$, entonces $F(x_0) + F'(x_0)(x^* - x_0) = -\eta(x^*, x_0)$. Si x^* está próximo a x_0 , despreciando la cantidad $\eta(x^*, x_0)$, obtenemos la siguiente ecuación lineal

$$F(x_0) + F'(x_0)(x - x_0) = 0. \quad (3.3)$$

Si esta ecuación tiene solución única en un entorno de x_0 , podemos pensar, como sucede en \mathbb{R} , que dicha solución x_1 será una aproximación de x^* mejor que x_0 . Repitiendo el proceso con x_1 en lugar de x_0 , obtenemos una nueva aproximación, x_2 , que se supone mejor que x_1 . Así obtenemos una sucesión $x_0, x_1, x_2, x_3, \dots$ que bajo ciertas condiciones converge a la solución de (3.2).

Notemos que para resolver la ecuación lineal (3.3) es suficiente que exista el operador $[F'(x_0)]^{-1}$. En este caso (3.3) tiene solución única y viene dada por

$$x_1 = x_0 - [F'(x_0)]^{-1}F(x_0).$$

En general, si existe $[F'(x_n)]^{-1}$ para $n = 0, 1, 2, \dots$, podemos definir la sucesión de Newton en espacios de Banach

$$x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n).$$

En la siguiente sección estudiaremos las condiciones en las que la sucesión anterior es convergente.

Volviendo al método de Newton en el caso escalar, hay varios resultados en los que se dan condiciones para su convergencia. Para no extendernos mucho en este sentido sólo daremos uno en el que probamos la convergencia global del método. Pueden consultarse otras condiciones de convergencia en [24], [32], [46], [51], [59], [82], [83], [93] o [101].

Teorema 3.1. *Sea $f : [a, b] \rightarrow \mathbb{R}$ una función que verifica*

- a) $f(a)f(b) < 0$.
- b) $f'(t) \neq 0$ si $t \in [a, b]$.
- c) $f''(t)$ no cambia de signo en $[a, b]$.

Entonces existe una única raíz t^ de $f(t) = 0$ en $[a, b]$, a la cual converge la sucesión $\{t_n\}$ definida por (3.1), con $t_0 \in [a, b]$ tal que $f(t_0)f''(t_0) \geq 0$.*

Las hipótesis de este teorema reciben el nombre de condiciones de Fourier, bajo las cuales podemos asegurar la convergencia global del método de Newton.

Definimos ahora uno de los conceptos más importantes en el estudio de los procesos iterativos, como es el de orden de convergencia.

Consideramos un proceso iterativo $x_{n+1} = o(x_n)$ y suponemos que la sucesión $\{x_n\}$ es convergente a un límite x^* . Para Schröder [98] el proceso anterior es de orden p si

$$o(x^*) = x^*: \quad o^{(j)}(x^*) = 0, \quad j = 1, 2, \dots, p-1: \quad o^{(p)}(x^*) \neq 0.$$

Esta definición solamente es válida para procesos iterativos definidos por funciones de una variable con p derivadas continuas. Una generalización de esta definición para sistemas de ecuaciones puede verse en [103].

La mayoría de los autores utilizan la siguiente definición, más reciente, de orden de convergencia.

Definición 3.2. Sea $\{x_n\}$ una sucesión convergente a un límite x^* . Si existe un número real p y una constante no nula C tal que

$$\lim_{n \rightarrow \infty} \frac{\|x^* - x_{n+1}\|}{\|x^* - x_n\|^p} = C, \quad (3.4)$$

diremos que $\{x_n\}$ es *convergente de orden p* y llamaremos a C *constante de error asintótico*.

Cuando $p = 2$, se dice que la convergencia es cuadrática, y cuando $p = 3$, cúbica.

A menudo, para resolver una ecuación $f(x) = 0$, se utiliza un proceso iterativo $x_{n+1} = \phi(x_n)$, donde ϕ es un funcional que depende de la función f . Si queremos asociar el concepto de orden con el proceso iterativo que genera la sucesión, podemos escribir (3.4) como sigue

$$\lim_{x \rightarrow x^*} \frac{\|x^* - \phi(x)\|}{\|x^* - x\|^p} = C.$$

Un estudio más detallado del orden de convergencia puede verse en [S1], donde se pone de manifiesto la relación entre el orden y la velocidad de convergencia de la sucesión, así como la independencia de la norma elegida en la definición de orden para el caso de operadores definidos en los espacios \mathbb{R}^n . En concreto, consideramos dos procesos iterativos ϕ_1 y ϕ_2 con el mismo límite x^* y órdenes p_1 y p_2 respectivamente. Si $p_1 < p_2$ la sucesión generada por ϕ_2 converge a x^* más rápidamente que la generada por ϕ_1 . Además esto ocurre con cualquier norma. Si $p_1 = p_2$, entonces el proceso con menor constante de error asintótico es el más rápido, aunque en este caso no se puede hablar de independencia de la norma elegida.

En un proceso iterativo $x_{n+1} = \phi(x_n)$, el funcional ϕ depende de una función f . Por tanto, el orden del proceso puede variar para distintas clases de funciones f . Así, para hablar con precisión deberíamos decir que un proceso es de al menos un cierto orden. Esta ambigüedad queda suprimida en otras definiciones de orden, como las que dan Ortega y Rheinboldt [S1].

La multiplicidad de una raíz puede alterar el orden de un proceso iterativo. Así, cuando decimos que un proceso iterativo es de un cierto orden p , se entiende que esto es cierto para los ceros de una misma multiplicidad. Mientras no se diga lo contrario, supondremos que todos los ceros de las funciones que tratamos son de multiplicidad uno. Algunas consideraciones acerca de procesos iterativos asociados a funciones con raíces múltiples pueden verse en [21] o [103].

Es bien conocido. [51], [81], [82], que el método de Newton tiene convergencia cuadrática.

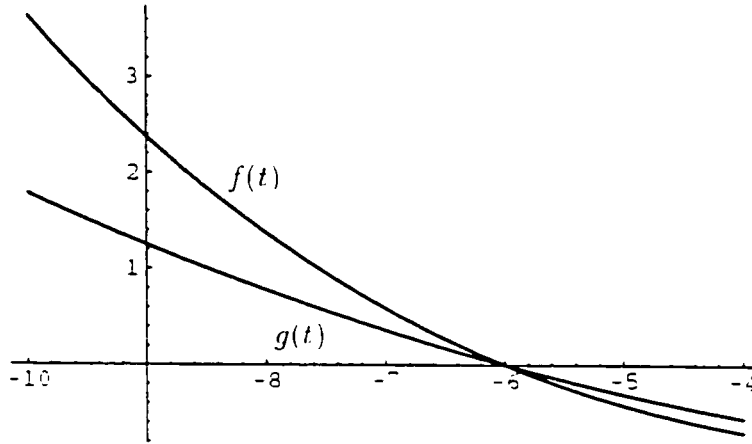


Figura 1.1: Gráficas de las funciones $f(t) = \frac{-t^3}{216} - 1$ y $g(t) = \frac{t^2}{36} - 1$

Si empleamos un método iterativo para construir una sucesión convergente a la solución x^* de una ecuación $F(x) = 0$, el inconveniente que tiene la anterior definición de orden es que en su expresión aparece x^* , que en principio, es desconocido. Recientemente, Dong Chen [36] dió una nueva definición de orden de convergencia para procesos iterativos en espacios de Banach, sin necesidad de conocer previamente x^* . No obstante, esta nueva definición no tiene la generalidad anterior y sólo se puede aplicar a procesos iterativos de una determinada forma.

Pasamos a estudiar la influencia de la convexidad en el método de Newton. Consideramos, sin pérdida de generalidad, una función f decreciente y convexa en un intervalo $[a, b]$ tal que $f(a)f(b) < 0$. En estas condiciones, tomando $t_0 \in [a, b]$ tal que $f(t_0) > 0$, el Teorema 3.1 garantiza la convergencia de la sucesión $\{t_n\}$ dada por (3.1) a t^* , la única raíz de $f(t) = 0$ en $[a, b]$. Además, la sucesión $\{t_n\}$ es creciente a t^* .

Sea g una función en las mismas condiciones que f en $[a, b]$, tal que $g(t^*) = 0$. Consideramos la sucesión

$$s_{n+1} = s_n - \frac{g(s_n)}{g'(s_n)}, \quad s_0 = t_0. \quad (3.5)$$

Hernández y Salanova, [53], [54], utilizan el grado de convexidad logarítmico dado en la Definición 2.2 para comparar las sucesiones $\{t_n\}$ y $\{s_n\}$ definidas por (3.1) y

(3.5) respectivamente. En concreto, prueban el siguiente resultado.

Teorema 3.3. *En las condiciones anteriores para f y g , si además $L_g(t) < L_f(t)$ para $t \in [t_0, t^*)$, entonces la sucesión $\{s_n\}$ converge a t^* más rápidamente que $\{t_n\}$. Es más, se tiene que $t_n < s_n$, $n \geq 1$.*

Ejemplo 3.4. Sean las funciones

$$f(t) = \frac{-t^3}{216} - 1, \quad g(t) = \frac{t^2}{36} - 1.$$

Es inmediato comprobar que f y g están en las condiciones de Fourier en el intervalo $[-10, -6]$, siendo ambas decrecientes y convexas. Además, se tiene que $L_g(t) < L_f(t)$ si $t^3 + 108t + 864 > 0$, y esto ocurre en dicho intervalo $[-10, -6)$. Por tanto, considerando $t_0 = s_0 = -10$, la sucesión $\{s_n\}$ construida para g converge a la raíz más rápidamente que la sucesión $\{t_n\}$ construida para f , como se puede ver en la Tabla 1.1.

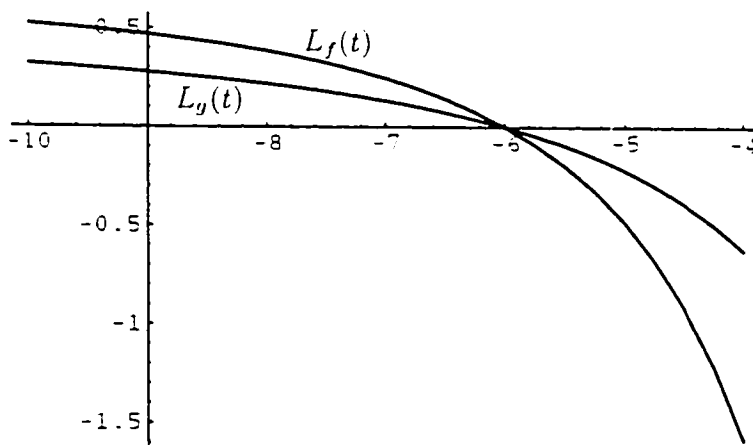


Figura 1.2: Gráficas de $L_f(x)$ y $L_g(x)$

1.4 El método de Newton en espacios de Banach. Teoremas de Kantorovich

Recogemos en esta sección alguno de los principales resultados conocidos sobre la convergencia del método de Newton en espacios de Banach.

iteración	t_n	s_n
0	-10.000000000000000	-10.000000000000000
1	-7.3866666666666700	-6.800000000000000
2	-6.244023743014800	-6.047058823529500
3	-6.009412497423800	-6.000183108262900
4	-6.000014735026500	-6.000000002793900
5	-6.000000000036300	-6.000000000000000
6	-6.000000000000100	-6.000000000000000
7	-6.000000000000100	-6.000000000000000
8	-6.000000000000100	-6.000000000000000
9	-6.000000000000100	-6.000000000000000

Tabla 1.1: Influencia de la convexidad

En 1948, Kantorovich [64] estableció, utilizando relaciones de recurrencia, un resultado conocido actualmente como Teorema de Kantorovich y que resume los resultados básicos referentes a convergencia del método, estimación del error y unicidad de soluciones (ver también [66], [81] y [82]). Un año más tarde, utilizando sucesiones mayorizantes, el mismo Kantorovich [65], dió otra demostración del mismo teorema.

La técnica de Kantorovich no es la única forma de abordar el estudio del método de Newton en espacios de Banach. Posteriormente, otros autores han estudiado dicho método en distintos contextos. Así, Rheinboldt [94] obtuvo resultados acerca de la convergencia del método de Newton como caso particular de unos resultados sobre ecuaciones en diferencias. Destacamos también el estudio realizado por Pták [88], [89] y en colaboración con Potra [85], a partir del método llamado inducción no discreta o inducción continua. La aplicación de este método a los procesos iterativos [86] abrió nuevas vías de estudio, sobre todo en lo relativo al análisis del error [110], [75], [10], [16]. Por último, citamos también los trabajos de Alefeld-Herzberger [2] y Moore [74], en los que se aplican las herramientas del análisis de intervalos para probar la convergencia de procesos iterativos (método de Newton y similares) para resolver sistemas de ecuaciones no lineales.

En esta sección nos ocuparemos del Teorema de Kantorovich. Para ello, empezamos introduciendo el concepto de sucesión mayorizante y viendo cómo se emplea para probar la convergencia de sucesiones en espacios de Banach.

Definición 4.1. Sea $\{x_n\}$ una sucesión en un espacio de Banach X y $\{t_n\}$ una sucesión de números reales. Diremos que la sucesión $\{t_n\}$ *mayoriza* a $\{x_n\}$ si se verifica la condición

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \quad n = 0, 1, 2, \dots$$

Observar que del hecho de ser mayorizante se sigue que la sucesión $\{t_n\}$ debe ser creciente. El interés de las sucesiones mayorizantes es que de su convergencia se puede deducir la convergencia de la sucesión en el espacio de Banach. En efecto, si $\{t_n\}$ converge a $t^* < \infty$, entonces existe $x^* \in X$ de manera que la sucesión $\{x_n\}$ converge a x^* y además

$$\|x^* - x_n\| \leq t^* - t_n, \quad n = 0, 1, 2, \dots$$

La desigualdad anterior nos permite también obtener cotas del error para la sucesión definida en espacios de Banach en términos de su correspondiente mayorizante.

En este resultado, y a partir de ahora, en toda esta memoria, denotaremos

$$B(x_0, r) = \{x \in X: \|x - x_0\| < r\}$$

$$\overline{B(x_0, r)} = \{x \in X: \|x - x_0\| \leq r\}.$$

Llamaremos a estos conjuntos, respectivamente, bola abierta y bola cerrada de centro x_0 y radio r .

Teorema 4.2. Sea $G : X \rightarrow X$ un operador definido en espacios de Banach con derivada continua en la bola cerrada $\Omega_0 = \overline{B(x_0, r_0)} \subseteq X$ y sea g una función real diferenciable en el intervalo $[t_0, t']$, $t' - t_0 \leq r_0$. Supongamos que

(i) $\|G(x_0) - x_0\| \leq g(t_0) - t_0$,

(ii) $\|G'(x)\| \leq g'(t)$ si $\|x - x_0\| \leq t - t_0$,

y que $t - g(t)$ tiene una raíz en el intervalo $[t_0, t']$. Entonces la sucesión $t_{n+1} = g(t_n)$ es convergente a \tilde{t} , la menor raíz de

$$t - g(t) \tag{4.1}$$

en el intervalo $[t_0, t']$, y además es mayorizante de la sucesión $x_{n+1} = G(x_n)$. En consecuencia $\{x_n\}$ está contenida en Ω_0 y es convergente a x^* , solución de

$$G(x) = x. \tag{4.2}$$

La ecuación (4.2) puede tener otras soluciones, además de x^* , en la bola Ω_0 , incluso si (4.1) tiene una única raíz en $[t_0, t']$. En el siguiente resultado se establecen condiciones acerca de la unicidad de solución de (4.2).

Teorema 4.3. *Supongamos que, además de las condiciones del teorema anterior, se verifica que*

$$g(t') \leq t'$$

y que (4.1) tiene una raíz única en $[t_0, t']$. Entonces, (4.2) tiene una solución única en Ω_0 . Además la sucesión $\{x_n\}$ converge a dicha solución empezando en cualquier punto de Ω_0 .

Una vez introducida la idea de sucesión mayorizante, pasamos a exponer el Teorema de Kantorovich. En [66], Kantorovich y Akilov establecen la convergencia del método de Newton exigiendo condiciones al punto de salida de la sucesión y acotando la derivada segunda del operador. Sea F una aplicación entre dos espacios de Banach X e Y de clase C^2 en la bola cerrada $\Omega_0 = \overline{B(x_0, r_0)} \subseteq X$. Dichos autores estudian la solución de la ecuación

$$F(x) = 0 \tag{4.3}$$

mediante el método de Newton

$$x_{n+1} = x_n - \Gamma_n F(x_n), \quad \Gamma_n = [F'(x_n)]^{-1}. \tag{4.4}$$

Para ello, consideran una función real auxiliar f de clase C^2 en el intervalo $[t_0, t']$ con $t' - t_0 \leq r_0$, y prueban, bajo las hipótesis descritas a continuación (que llamamos condiciones generales de Kantorovich), que la convergencia de la sucesión de Newton está garantizada.

Teorema 4.4. *Supongamos que se cumplen las siguientes condiciones:*

- (i) *Existe el operador inverso $\Gamma_0 = [F'(x_0)]^{-1}$.*
- (ii) *$f'(t_0) < 0$.*
- (iii) *$\|\Gamma_0 F(x_0)\| \leq -\frac{f(t_0)}{f'(t_0)}$.*

(iv) $\|\Gamma_0 F''(x)\| \leq -\frac{f''(t)}{f'(t_0)}$ si $\|x - x_0\| \leq t - t_0 \leq r_0$.

(v) La ecuación $f(t) = 0$ tiene una raíz en $[t_0, t']$.

Entonces el método de Newton (4.4) para resolver (4.3), empezando en x_0 , nos proporciona una sucesión convergente a una raíz x^* de $F(x) = 0$. Además

$$\|x^* - x_n\| \leq t^* - t_n, \quad n = 0, 1, \dots$$

donde t^* es la menor raíz de

$$f(t) = 0 \tag{4.5}$$

en el intervalo $[t_0, t']$.

Por último, si $f(t') \leq 0$ y (4.5) tiene una sola raíz en $[t_0, t']$, x^* es la única raíz de (4.3) en Ω_0 .

En la práctica, la aplicación del teorema anterior puede resultar complicada, ya que en él aparece una función f desconocida. Por ello el siguiente resultado tiene especial interés, ya que si el operador F cumple ciertas condiciones, podemos tomar como f en el Teorema 4.4 el polinomio $p(t) = \frac{k}{2}t^2 - t + a$.

Teorema 4.5. Sea $F : X \rightarrow Y$ un operador en las condiciones anteriores, es decir, con derivada segunda continua en la bola cerrada $\Omega_0 = \overline{B(x_0, r)}$. Supongamos además que

(i) Existe el operador inverso $\Gamma_0 = [F'(x_0)]^{-1}$.

(ii) $\|\Gamma_0 F(x_0)\| \leq a$.

(iii) $\|\Gamma_0 F''(x)\| \leq k$ si $x \in \Omega_0$.

Entonces, si se cumple

$$h = ka \leq \frac{1}{2}$$

y

$$r \geq r_0 = \frac{1 - \sqrt{1 - 2h}}{h} a,$$

la ecuación (4.3) tiene solución x^* y la sucesión de Newton dada por (4.4) es convergente a esta solución. Además,

$$\|x^* - x_0\| \leq r_0.$$

es decir, $x^* \in \overline{B(x_0, r_0)}$. Por otra parte, si $2h < 1$ y

$$r < r_1 = \frac{1 + \sqrt{1 - 2h}}{h} a$$

o bien $2h = 1$ y $r \leq r_1$, la solución x^* es única en Ω_0 . Por último, se tiene la siguiente estimación del error

$$\|x^* - x_n\| \leq 2^{1-n} (2h)^{2^n - 1} a, \quad n = 0, 1, \dots$$

La demostración de los dos teoremas anteriores puede verse en [66]. A las condiciones del Teorema 4.5 las llamaremos a partir de ahora condiciones de Kantorovich (no confundir con las condiciones generales de Kantorovich, que son las del Teorema 4.4).

A continuación, presentamos una versión más reciente [107] del resultado anterior, con condiciones más débiles que las dadas por Kantorovich. En lugar de acotar la derivada segunda del operador, se exige una condición de tipo Lipschitz sobre la derivada primera y que los conjuntos donde el operador F está definido sean conjuntos convexos en lugar de bolas. Se prueba la convergencia de la sucesión (4.4) a una solución de (4.3), se estudian las regiones de existencia y unicidad de dicha solución y se da una estimación del error cometido en cada paso. Estas condiciones, más débiles que las del Teorema 4.4, suelen denominarse condiciones de tipo Kantorovich.

Teorema 4.6. Sea $F : D \subseteq X \rightarrow Y$ un operador diferenciable en un conjunto abierto y convexo $D_0 \subseteq D$ y tal que para algún $x_0 \in D_0$ exista el operador $\Gamma_0 = [F'(x_0)]^{-1} \in \mathcal{L}(Y, X)$. Podemos suponer, sin pérdida de generalidad, que $F(x_0) \neq 0$. Supongamos que F' satisface la condición de Lipschitz

$$\|\Gamma_0(F'(x) - F'(y))\| \leq k\|x - y\|, \quad x, y \in D_0.$$

Denotamos $a = \|\Gamma_0 F(x_0)\|$, $h = ka$ y $t^* = \frac{2a}{1 + \sqrt{1 - 2h}}$. Suponemos también que $h \leq 1/2$ y que

$$\overline{B} = \overline{B(x_1, t^* - a)} \subseteq D_0.$$

Entonces

- (i) La sucesión dada por (4.4) está bien definida, es convergente a una solución $x^* \in \overline{B}$ de la ecuación (4.3) y además, para $n \geq 1$, los términos de la sucesión $\{x_n\}$ están en la bola abierta $B = B(x_1, t^* - a)$.

(ii) La solución x^* es única en $B(x_0, t^{**}) \cap D_0$ con

$$t^{**} = \frac{1 + \sqrt{1 - 2h}}{k}.$$

si $2h < 1$, o en $\overline{B(x_0, t^{**})} \cap D_0$ si $2h = 1$.

(iii) Una estimación del error viene dada por

$$\|x^* - x_n\| \leq t^* - t_n = \frac{2a_n}{1 + \sqrt{1 - 2h_n}} \leq 2^{1-n}(2h)^{2^n-1}a.$$

donde $\{t_n\}$ es la sucesión definida por

$$t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}, \quad t_0 = 0$$

con

$$p(t) = \frac{k}{2}t^2 - t + a.$$

y $\{a_n\}$, $\{h_n\}$ son las sucesiones definidas por las siguientes relaciones de recurrencia

$$b_0 = 1, \quad a_0 = a, \quad h_0 = h = ka.$$

$$b_{n+1} = \frac{b_n}{1 - h_n}, \quad a_{n+1} = \frac{h_n a_n}{2(1 - h_n)}, \quad h_n = kb_n a_n, \quad n \geq 0.$$

1.5 Ejemplo de aplicación del método de Newton

A continuación vamos a utilizar el método de Newton para resolver algunas ecuaciones diferenciales e integrales. La técnica general para resolver este tipo de problemas puede verse en [92]. La clave de estos problemas está en encontrar los inversos de ciertos operadores diferenciales. Esto se consigue teniendo en cuenta que, en algunos casos, resolver una ecuación diferencial es equivalente a resolver una ecuación integral. En efecto, resolver un problema de valores en la frontera

$$\begin{cases} \frac{d^2 y}{dx^2} = \phi(x, y(x)) \\ y(0) = y_0, \quad y(1) = y_1. \end{cases} \quad (5.1)$$

con ϕ una función continua para $x \in [0, 1]$ e $y \in C^2[0, 1]$, es equivalente, ver [87], a resolver la ecuación integral de Fredholm de segunda clase

$$y(x) = \int_0^1 G(x, t)\phi(t, y(t)) dt + (y_1 - y_0)x + y_0, \quad 0 \leq x \leq 1,$$

donde $G(x, t)$ es el núcleo de Green

$$G(x, t) = \begin{cases} (x-1)t, & 0 \leq t \leq x, \\ (t-1)x, & x \leq t \leq 1. \end{cases}$$

Desde otro punto de vista, el operador inverso de

$$P : C^2[0, 1] \rightarrow C[0, 1]$$

definido por

$$P(y) = \frac{d^2 y}{dx^2}$$

es el operador $Q : C[0, 1] \rightarrow C^2[0, 1]$, que a $\phi \in C[0, 1]$ le hace corresponder $Q(\phi) \in C^2[0, 1]$, definido de la siguiente manera

$$Q(\phi)(x) = \int_0^1 G(x, t)\phi(t, y(t)) dt + (y_1 - y_0)x + y_0. \quad (5.2)$$

Por ejemplo, consideramos el problema de valores en la frontera siguiente

$$\begin{cases} \frac{d^2 y}{dx^2} = xy^2 - 1 \\ y(0) = y(1) = 0. \end{cases}$$

Si denotamos F al operador de $C^2[0, 1]$ en $C[0, 1]$ que a una función $y \in C^2[0, 1]$ le hace corresponder

$$F(y) = \frac{d^2 y}{dx^2} - xy^2 + 1,$$

el problema de valores en la frontera anterior consiste en resolver la ecuación

$$F(y) = 0 \quad (5.3)$$

con las condiciones

$$y(0) = y(1) = 0. \quad (5.4)$$

La derivada del operador F en $y = y(x)$ satisface

$$F'(y)(z) = \frac{d^2 z}{dx^2} - 2xy(x)z(x). \quad (5.5)$$

con $z \in C^2[0,1]$. Podemos escribir el método de Newton para resolver (5.3) de la siguiente forma

$$F'(y_n)(y_{n+1} - y_n) = -F(y_n).$$

Denotando $u_n = y_{n+1} - y_n$ y teniendo en cuenta (5.5), nos queda

$$\frac{d^2 u_n}{dx^2} = 2xy_n u_n - F(y_n). \quad (5.6)$$

Teniendo en cuenta la equivalencia entre resolver (5.1) y (5.2), en nuestro caso particular, y considerando las condiciones en la frontera (5.4), podemos “despejar” u_n de (5.6) para obtener

$$u_n(x) = 2 \int_0^1 G(x,t) t y_n(t) u_n(t) dt + \int_0^1 G(x,t) F(y_n(t)) dt. \quad (5.7)$$

Así, por ejemplo, si tomamos como punto de salida la función

$$y_0(x) = \frac{x}{2}(1-x), \quad 0 \leq x \leq 1,$$

tenemos, sustituyendo en (5.7),

$$\begin{aligned} u_0(x) &= \int_0^1 G(x,t) t^2 (1-t) u_0(t) dt + \frac{1}{4} \int_0^1 G(x,t) t^3 (1-t)^2 dt \\ &= \int_0^1 G(x,t) t^2 (1-t) u_0(t) dt + \frac{x^7}{42} - \frac{x^6}{15} + \frac{x^5}{20} - \frac{x}{140}. \end{aligned}$$

Si denotamos T al operador integral de núcleo

$$K(x,t) = G(x,t) t^2 (1-t),$$

podemos escribir la ecuación anterior de la forma

$$(I - T)(u_0) = f(x). \quad (5.8)$$

con

$$f(x) = \frac{x^7}{42} - \frac{x^6}{15} + \frac{x^5}{20} - \frac{x}{140}.$$

En $C[0,1]$ consideramos la norma

$$\|y\|_\infty = \max_{x \in [0,1]} |y(x)|$$

y en el espacio $C^2[0,1]$ la norma

$$\|y\| = \max\{\|y\|_\infty, \|y'\|_\infty, \|y''\|_\infty\}.$$

Con estas normas, obtenemos para el operador integral

$$Ty(x) = \int_0^1 G(x, t)t^2(1-t)y(t) dt$$

la estimación

$$\|Ty\| = \max\{\|Ty\|_\infty, \|(Ty)'\|_\infty, \|(Ty)''\|_\infty\} \leq \frac{4}{27}\|y\|_\infty.$$

Luego $\|T\| \leq 4/27$. Por el Lema de Banach, existe $[I - T]^{-1}$ y por tanto la ecuación integral (5.8) tiene solución. Resolviendo esta ecuación integral hallamos $u_0(x)$, y en consecuencia la función

$$y_1(x) = y_0(x) + u_0(x)$$

es el siguiente paso en la aplicación del método de Newton.

Para ver que la sucesión de Newton es convergente a una solución de (5.3), veremos que se cumplen las condiciones del Teorema 4.5 de Kantorovich.

Notar que, por (5.8) y por el Lema de Banach,

$$\begin{aligned} \|[F'(y_0)]^{-1}F(y_0)\| &= \|y_1 - y_0\| = \|u_0\| \leq \|[I - T]^{-1}\| \|f(x)\|_\infty \\ &\leq \frac{27}{23} \|f(x)\|_\infty \leq \frac{27}{23} \frac{31}{420} \approx 0.087. \end{aligned}$$

Por tanto, $a = 0.087$.

Por otra parte, como $[F'(y_0)]^{-1} = [I - T]^{-1}G$, se sigue que

$$\|[F'(y_0)]^{-1}\| \leq \|[I - K]^{-1}\| \|G\| \leq \frac{\|G\|}{1 - \|K\|},$$

y como $\|G\| \leq 1$,

$$\|[F'(y_0)]^{-1}\| \leq \frac{1}{1 - \|K\|} = \frac{27}{23}.$$

Por último, $F''(y) = -2xI_2$, donde I_2 es el operador bilineal definido por la multiplicación.

$$I_2xy(t) = x(t)y(t).$$

Observar que $F''(y)$ es un operador bilineal constante y que

$$\|F''(y)\| \leq 2.$$

En consecuencia,

$$\|[F'(y_0)]^{-1}F''(y)\| \leq 2\frac{27}{23} = k.$$

y $h = ak = 2(0.087)(27/23) \simeq 0.20426 < 1/2$.

Por tanto, el Teorema 4.5 garantiza la existencia de solución de (5.3) con las condiciones en la frontera (5.4), así como la convergencia a dicha solución de la sucesión de Newton $\{y_n(x)\}$ empezando en $y_0(x)$.

Existen ejemplos de ecuaciones en los que no se cumplen las condiciones de Kantorovich, ni tampoco las condiciones más débiles del Teorema 4.6 (ver [8]). Pensar por ejemplo en la ecuación escalar definida en $[0, 1]$ por

$$h(t) = 0, \quad \text{con } h(t) = t^{1+p} - \frac{t}{4}, \quad 0 < p < 1,$$

o en la ecuación diferencial

$$y'' + y^{1+p} = 0, \quad 0 < p < 1$$

con las condiciones en la frontera $y(0) = y(1) = 0$.

En ambos casos, las derivadas segundas de los operadores pueden no estar definidas, con lo que no se puede utilizar el Teorema 4.5.

Las condiciones del Teorema 4.6 tampoco se verifican, ya que para la diferencia de derivadas se obtienen estimaciones del tipo

$$\|F'(x) - F'(y)\| \leq k\|x - y\|^p, \quad 0 < p < 1, \quad (5.9)$$

y no se puede garantizar que $\|x - y\|^p \leq \|x - y\|$.

Sin embargo, algunos autores, [96], [8], [13], [14], han logrado probar la convergencia del método de Newton con condiciones del tipo (5.9), que se denominan condiciones tipo Hölder.

La idea de probar la convergencia del método de Newton con condiciones más generales que las de Kantorovich ha atraído a muchos autores y se ha convertido en un reciente e interesante tema de investigación. Así, a parte de las condiciones tipo Hölder comentadas anteriormente, se ha estudiado el caso de ecuaciones definidas por un operador no diferenciable [15] o el caso de que la derivada del operador no sea inversible en el sentido clásico [78].

Capítulo 2

Dominios de puntos de salida para el método de Newton

2.1 Introducción

Como ya se ha estudiado en el Capítulo 1, entre las condiciones para la convergencia del método de Newton aparece siempre alguna sobre el punto de salida x_0 para que la sucesión $\{x_n\}$ sea convergente a la raíz de la ecuación considerada. En este capítulo trataremos de suavizar este tipo de condiciones para el punto de salida, de manera que ampliemos el dominio de puntos que pueden serlo.

La caracterización de dominios de puntos de salida para que el método de Newton converja a una solución de una ecuación no lineal es un interesante y difícil problema, incluso en el caso escalar [48]. Si una ecuación tiene varias raíces, podemos pensar en determinar los dominios de puntos de salida desde los cuales el método de Newton converge a cada una de las raíces. No es sencillo este problema, ya que puntos de salida próximos pueden originar sucesiones convergentes a distintas raíces o, incluso, una sucesión convergente y otra no [92]. Nos adentramos así en otro interesante campo como es el del caos y los fractales y su relación con el método de Newton. Puede verse una introducción a estos temas en el libro de Becker y Dörfler [18].

En esta sección realizamos un estudio local del problema, en concreto, dada una ecuación no lineal, daremos condiciones para que en un cierto dominio exista una raíz de la ecuación considerada y delimitaremos un conjunto de puntos de salida para los que se obtiene convergencia a dicha raíz.

Notemos que el problema de resolver una ecuación de la forma $F(x) = 0$ se puede transformar en un problema de localización de un punto fijo del operador $G(x) = x - F(x)$, y recíprocamente. Así, resultan dos problemas equivalentes y por tanto las técnicas empleadas en una de las dos situaciones son perfectamente extensibles a la otra.

En el Teorema del Punto Fijo de Banach aparecen dos condiciones. En primer lugar, debemos tener un operador que transforma un espacio de Banach X en sí mismo y además, en segundo lugar, este operador debe ser una contracción. Observar que el método de Newton para la resolución de la ecuación $F(x) = 0$ consiste en construir la sucesión dada por

$$x_{n+1} = G(x_n), \text{ con } G(x) = x - [F'(x)]^{-1}F(x).$$

Por tanto, resulta evidente que esta sucesión es la misma que la obtenida aplicando el método de las aproximaciones sucesivas al operador G , y que en las condiciones anteriores converge a un punto fijo de G .

Es conocido [52], [57], que para funciones reales se verifica que $G'(x) = L_F(x)$, siendo L_F el grado de convexidad logarítmico de la función real F . En este capítulo vamos a probar, a partir de la definición del operador $L_F(x)$ vista en el capítulo anterior, que el resultado citado también se verifica para operadores definidos en espacios de Banach. Así, conocida la derivada del operador que define el método de Newton, podemos aplicar las técnicas del Teorema del Punto Fijo para estudiar la convergencia de este proceso iterativo. Esta idea, que a priori puede parecer un poco artificiosa y que nunca ha sido utilizada, nos permitirá dar condiciones para la convergencia del método de Newton y analizar en profundidad los dominios de puntos de salida que garanticen la convergencia de dicho método. En lo relativo a este último aspecto, se ha conseguido probar la convergencia del método de Newton a partir de puntos de salida donde no se verificaban las condiciones clásicas de Kantorovich [66], como se pondrá de manifiesto en los ejemplos que aparecen en este capítulo.

2.2 El operador L_F y el método de Newton

Como ya vimos en el Capítulo 1, dada una función $f : [a, b] \rightarrow \mathbb{R}$, podemos interpretar el grado de convexidad logarítmico de f como la función $L_f : [a, b] \rightarrow \mathbb{R}$ definida por

$$L_f(t) = \frac{f(t)f''(t)}{f'(t)^2}.$$

También vimos como extender este concepto a espacios de Banach. Así, dado un operador $F : X \rightarrow Y$, el operador lineal

$$L_F(x) = \Gamma(x)F''(x)[\Gamma(x)F(x)] \in \mathcal{L}(X, X).$$

con $\Gamma(x) = [F'(x)]^{-1} \in \mathcal{L}(Y, X)$ generaliza el grado de convexidad logarítmico definido para funciones reales. Hay que dejar claro que en este caso, este operador no es una medida de la convexidad de F , y por tanto, se pierde el sentido geométrico que se tenía en el caso escalar. No obstante, el estudio de este operador nos permitirá analizar el método de Newton y otros de orden tres con un enfoque distinto al dado por otros autores.

Como ya indicamos en la introducción, dada una ecuación escalar $f(t) = 0$, la sucesión de Newton se obtiene aplicando el método de las aproximaciones sucesivas $t_{n+1} = g(t_n)$ a la función

$$g(t) = t - \frac{f(t)}{f'(t)}$$

comenzando en un t_0 determinado. En este caso es una comprobación inmediata que $g'(t) = L_f(t)$. Veamos a continuación la generalización de este resultado a espacios de Banach. Para ello necesitaremos el siguiente lema.

Lema 2.1. Sean X e Y espacios de Banach. Consideramos dos operadores

$$Q : X \rightarrow Y, \quad P : X \rightarrow \mathcal{L}(Y, X)$$

diferenciables en un punto x_0 . Si definimos el operador H de X en X mediante

$$H(x) = [P(x)Q](x), \quad X \xrightarrow{Q} Y \xrightarrow{P(x)} X.$$

entonces H es diferenciable en x_0 y su derivada $H'(x_0) : X \rightarrow X$ es la aplicación lineal que a $z \in X$ le hace corresponder

$$H'(x_0)[z] = (P'(x_0)[z]Q)(x_0) + (P(x_0)Q'(x_0))[z].$$

Demostración. Utilizando la definición de derivada tenemos, para $z \in X$,

$$\begin{aligned} H'(x_0)(z) &= \lim_{s \rightarrow 0} \frac{H(x_0 + sz) - H(x_0)}{s} \\ &= \lim_{s \rightarrow 0} \frac{[P(x_0 + sz)Q](x_0 + sz) - [P(x_0)Q](x_0)}{s} \end{aligned}$$

$$\begin{aligned}
&= \lim_{s \rightarrow 0} \frac{[P(x_0 + sz) - P(x_0)]Q(x_0 + sz) + P(x_0)[Q(x_0 + sz) - Q(x_0)]}{s} \\
&= (P'(x_0)(z)Q)(x_0) + (P(x_0)Q'(x_0))(z). \quad \blacksquare
\end{aligned}$$

La sucesión de Newton en espacios de Banach se obtiene aplicando el método de las aproximaciones sucesivas al operador

$$G(x) = x - [\Gamma(x)F](x), \quad \text{con } \Gamma(x) = [F'(x)]^{-1}.$$

En el teorema siguiente probamos que la derivada del operador G en un punto x_0 es precisamente el operador grado de convexidad logarítmico $L_F(x_0)$, que se ha introducido en la Definición I.2.4.

Teorema 2.2. *Con la notación anterior se tiene que*

$$G'(x_0) = L_F(x_0).$$

Demostración. Denotamos $H(x) = [\Gamma(x)F](x)$, siendo Γ el operador de X en $\mathcal{L}(Y, X)$ que a un punto $x \in X$ le hace corresponder $\Gamma(x) = [F'(x)]^{-1} \in \mathcal{L}(Y, X)$. Entonces $G'(x_0) = I - H'(x_0)$. Para derivar el operador H , notemos que es de la forma de los operadores que aparecen en el Lema 2.1. En consecuencia, para $z \in X$ tenemos que

$$H'(x_0)(z) = [\Gamma'(x_0)(z)F](x_0) + [\Gamma(x_0)F'(x_0)](z) = [\Gamma'(x_0)(z)F](x_0) + z.$$

Por tanto, $G'(x_0)(z) = -[\Gamma'(x_0)(z)F](x_0)$.

Ahora bien, podemos expresar Γ como la siguiente composición de operadores

$$\Gamma = TF'$$

siendo $T : \mathcal{L}(X, Y) \rightarrow \mathcal{L}(Y, X)$ el operador inversión $T(F) = F^{-1}$. Como

$$\Gamma : X \xrightarrow{F'} \mathcal{L}(X, Y) \xrightarrow{T} \mathcal{L}(Y, X),$$

por la regla de la cadena para derivar una composición de operadores, sabemos que $\Gamma'(x_0) \in \mathcal{L}(X, \mathcal{L}(Y, X))$ y, además, dado $x \in X$,

$$\Gamma'(x_0)[x] = T'(F'(x_0))F''(x_0)(x).$$

Ahora bien, teniendo en cuenta la expresión de la derivada del operador inversión [92], se sigue que, para $z \in X$,

$$\Gamma'(x_0)(z) = T'(F(x_0))(F''(x_0)(z)) = -\Gamma(x_0)F''(x_0)(z)\Gamma(x_0) \in \mathcal{L}(Y, X).$$

Además, como $F''(x_0)$ es simétrica, se obtiene, para $z \in X$,

$$\begin{aligned} G'(x_0)(z) &= \Gamma(x_0)F''(x_0)(z)\Gamma(x_0)F(x_0) \\ &= \Gamma(x_0)F''(x_0)\Gamma(x_0)F(x_0)(z) = L_F(x_0)(z). \quad \blacksquare \end{aligned}$$

La condición de existencia del operador L_F en un cierto conjunto parece muy fuerte, ya que conlleva la existencia de los operadores inversos $\Gamma(x) = [F'(x)]^{-1}$ en todos los puntos de dicho conjunto. Sin embargo, exigiendo unas ciertas condiciones a la derivada segunda de F , se garantiza la existencia de L_F en una cierta bola. Como se ha visto en el Capítulo 1, estas condiciones son similares a las que necesita Kantorovich en el Teorema I.4.5 para probar la convergencia del método de Newton.

Teorema 2.3. *Sea $F : X \rightarrow Y$ un operador entre espacios de Banach, con derivada segunda continua en un conjunto convexo $\Omega_0 \subseteq X$. Sea $x_0 \in \Omega_0$ de manera que existe $\Gamma_0 = \Gamma(x_0) = [F'(x_0)]^{-1}$. Supongamos además que $\|\Gamma_0 F''(x)\| \leq k$ para $x \in \Omega_0$. Entonces existe $L_F(x)$ para todo $x \in B(x_0, 1/k) \cap \Omega_0$ y además*

$$\|L_F(x)\| \leq \frac{k}{(1 - k\|x - x_0\|)^2} \|\Gamma_0 F(x)\|.$$

Demostración. Tener en cuenta que para $x \in \Omega_0$

$$\int_{x_0}^x \Gamma_0 F''(z) dz = \Gamma_0 F'(x) - I,$$

y por tanto,

$$\|\Gamma_0 F'(x) - I\| = \left\| \int_{x_0}^x \Gamma_0 F''(z) dz \right\| \leq k\|x - x_0\| < 1.$$

Entonces, por el Lema de Banach para la inversión de operadores se sigue que, si $x \in B(x_0, 1/k)$, existe $[\Gamma_0 F'(x)]^{-1}$ y además

$$\|[\Gamma_0 F'(x)]^{-1}\| \leq \frac{1}{1 - k\|x - x_0\|}.$$

Utilizando que

$$L_F(x) = [\Gamma_0 F'(x)]^{-1} \Gamma_0 F''(x) [\Gamma_0 F'(x)]^{-1} \Gamma_0 F(x).$$

se obtiene el resultado. ■

Notas. • La condición de que F'' sea continua puede reemplazarse por la de que F'' sea integrable y el resultado sigue siendo válido.

• En las condiciones generales de Kantorovich (Teorema I.4.4), se sigue la existencia del operador $L_F(x)$ para todo $x \in B(x_0, -f'(t_0)/M)$, siendo $M = \max\{f''(t); t \in [t_0, t']\}$, sin más que aplicar este resultado para $\Omega_0 = \overline{B(x_0, r_0)}$.

2.3 Condiciones de convergencia para el método de Newton

Nos ocuparemos ahora de resolver una ecuación de la forma $F(x) = 0$ siendo F un operador entre dos espacios de Banach X e Y . Para ello construimos formalmente la sucesión de Newton dada por

$$x_{n+1} = x_n - [F'(x_n)]^{-1} F(x_n). \quad (3.1)$$

y comenzando en x_0 .

Denotamos G al operador de X en X definido por

$$G(x) = x - \Gamma(x)F(x), \text{ con } \Gamma(x) = [F'(x)]^{-1}.$$

Observar que la sucesión (3.1) se obtiene aplicando el método de las aproximaciones sucesivas al operador G empezando en x_0 . Como ya hemos comentado anteriormente, el problema de encontrar una solución de $F(x) = 0$ es equivalente al de encontrar un punto fijo de G .

Un resultado clásico sobre puntos fijos, que por su generalidad utilizaremos en distintos contextos, afirma que si X es un espacio de Banach, $G : X \rightarrow X$ un operador continuo y la sucesión $x_{n+1} = G(x_n)$ comenzando en $x_0 \in X$ converge a algún $x^* \in X$, entonces x^* es un punto fijo de G , es decir, $G(x^*) = x^*$. La demostración de este resultado se sigue sin más que tener en cuenta la continuidad de G .

Ya estamos en condiciones de obtener un primer resultado de convergencia para el método de Newton utilizando el operador grado de convexidad logarítmico y argumentos similares a los del Teorema del Punto Fijo.

Teorema 3.1. *Con la notación anterior, supongamos que L_F está definido en un cierto conjunto cerrado y convexo Ω y que $\|L_F(x)\| \leq M < 1$, para todo $x \in \Omega$. Así mismo, sea $x_0 \in \Omega$ y*

$$r_0 = \frac{\|x_0 - G(x_0)\|}{1 - M} = \frac{\|[F'(x_0)]^{-1}F(x_0)\|}{1 - M}. \quad (3.2)$$

En estas condiciones, si $\overline{B(x_0, r_0)} \subseteq \Omega$, se tiene que

- (1) Existe un punto fijo x^* de G en $\overline{B(x_0, r_0)}$.
- (2) x^* es el único punto fijo de G en Ω .
- (3) La sucesión $x_{n+1} = G(x_n)$, comenzando en x_0 , converge a x^* .

Además se verifica

$$\|x^* - x_n\| \leq M^n r_0.$$

Demostración. En primer lugar, veamos que la sucesión $\{x_n\}$ está contenida en $\overline{B(x_0, r_0)}$. Procederemos por inducción. Para $n = 1$ tenemos, por (3.2),

$$\|x_1 - x_0\| = \|G(x_0) - x_0\| = (1 - M)r_0 < r_0.$$

En consecuencia, $x_1 \in B(x_0, r_0)$.

Supongamos que $x_1, x_2, \dots, x_n \in B(x_0, r_0)$ y que

$$\|x_n - x_0\| \leq (1 - M^n)r_0 < r_0.$$

Entonces, utilizando que $G'(x) = L_F(x)$, como ya probamos en el Teorema 2.2, y teniendo en cuenta el Teorema del Valor Medio,

$$\begin{aligned} \|x_{n+1} - x_n\| &= \|G(x_n) - G(x_{n-1})\| \\ &\leq \sup_{z \in [x_{n-1}, x_n]} \|G'(z)\| \|x_n - x_{n-1}\| \leq M \|x_n - x_{n-1}\|. \end{aligned} \quad (3.3)$$

En consecuencia, reiterando el proceso,

$$\|x_{n+1} - x_n\| \leq M \|x_n - x_{n-1}\| \leq \dots$$

$$\dots \leq M^n \|x_1 - x_0\| = M^n(1 - M)r_0. \quad (3.4)$$

y así

$$\begin{aligned} \|x_{n+1} - x_0\| &\leq \|x_{n+1} - x_n\| + \|x_n - x_0\| \\ &\leq M^n(1 - M)r_0 + (1 - M^n)r_0 = (1 - M^{n+1})r_0 < r_0. \end{aligned}$$

luego $x_{n+1} \in B(x_0, r_0)$.

Tenemos por tanto que $x_n \in B(x_0, r_0)$ para todo $n \geq 1$ y además

$$\|x_n - x_0\| \leq (1 - M^n)r_0.$$

Veamos ahora que $\{x_n\}$ es una sucesión de Cauchy. Para ello consideramos un entero $p \geq 1$: entonces, aplicando sucesivamente la desigualdad dada por (3.3),

$$\begin{aligned} \|x_{n+p} - x_n\| &\leq \|x_{n+p} - x_{n+p-1}\| + \dots + \|x_{n+1} - x_n\| \\ &\leq M\|x_{n+p-1} - x_{n+p-2}\| + M\|x_{n+p-2} - x_{n+p-3}\| + \dots + \|x_{n+1} - x_n\| \leq \dots \\ &\dots \leq (M^{p-1} + M^{p-2} + \dots + M^2 + M + 1)\|x_{n+1} - x_n\| = \frac{1 - M^p}{1 - M}\|x_{n+1} - x_n\|. \end{aligned}$$

Teniendo en cuenta (3.4) se sigue que

$$\|x_{n+p} - x_n\| \leq \frac{1 - M^p}{1 - M} M^n(1 - M)r_0 = (1 - M^p)M^n r_0. \quad (3.5)$$

De aquí se deduce que $\{x_n\}$ es de Cauchy y, por tanto, convergente.

Sea $x^* = \lim_{n \rightarrow \infty} x_n$. Como G es un operador diferenciable, luego continuo, sabemos que x^* es un punto fijo de G , y en consecuencia una solución de $F(x) = 0$.

A continuación veamos que x^* es el único punto fijo de G en Ω . Supongamos que x^{**} es otro punto fijo de G en Ω . Entonces, aplicando el Teorema 2.2, el Teorema del Valor Medio, y teniendo en cuenta que L_F está acotado en Ω , si suponemos que $x^{**} \neq x^*$ obtenemos

$$\begin{aligned} \|x^* - x^{**}\| &= \|G(x^*) - G(x^{**})\| \\ &\leq \sup_{z \in [x^*, x^{**}]} \|L_F(z)\| \|x^* - x^{**}\| \leq M \|x^* - x^{**}\| < \|x^* - x^{**}\|, \end{aligned}$$

que es una contradicción.

Por último, para dar una estimación del error, basta observar que, por (3.5),

$$\|x_{n+p} - x_n\| \leq (1 - M^p)M^n r_0.$$

y haciendo $p \rightarrow \infty$, se sigue que

$$\|x^* - x_n\| \leq M^n r_0. \quad (3.6)$$

con lo que se prueba el teorema. ■

Notemos que el único punto fijo del operador G resulta ser la única raíz de la ecuación $F(x) = 0$, ya que de

$$x^* = G(x^*) = x^* - [F'(x^*)]^{-1}F(x^*)$$

se deduce que $[F'(x^*)]^{-1}F(x^*) = 0$. Aplicando el operador lineal $F'(x^*)$ a la igualdad anterior, se sigue que $F(x^*) = 0$.

En el teorema anterior, hemos probado la convergencia de la sucesión dada por el método de Newton (3.1) fijando un punto de salida. A continuación vamos a obtener un resultado que garantiza la convergencia de dicha sucesión para un conjunto más amplio de puntos de salida. En concreto, ya hemos probado la convergencia de la sucesión de Newton (3.1) comenzando en el centro de una bola. Podemos demostrar que, además, comenzando en cualquier otro punto de dicha bola también obtenemos una sucesión convergente a la solución de $F(x) = 0$.

Teorema 3.2. *En las condiciones del teorema anterior, si construimos la sucesión $\tilde{x}_{n+1} = G(\tilde{x}_n)$ comenzando en un punto cualquiera \tilde{x}_0 de la bola $\overline{B(x_0, r_0)}$, con r_0 dado por (3.2), tenemos que $\{\tilde{x}_n\}$ es convergente al único punto fijo del operador G en $\overline{B(x_0, r_0)}$. Además*

$$\|x^* - \tilde{x}_n\| \leq \frac{M^n}{1-M} \|\tilde{x}_1 - \tilde{x}_0\|$$

y

$$\|x^* - \tilde{x}_n\| \leq M^n \|x^* - \tilde{x}_0\|.$$

Demostración. Sea x un punto cualquiera de $\overline{B(x_0, r_0)}$. Entonces

$$\|x_1 - G(x)\| = \|G(x_0) - G(x)\| \leq M\|x_0 - x\| \leq Mr_0.$$

Así, G aplica la bola $\overline{B(x_0, r_0)}$ en la bola $\overline{B(x_1, Mr_0)}$.

Además tenemos que $\overline{B(x_1, Mr_0)} \subseteq \overline{B(x_0, r_0)}$. En efecto, si x está en $\overline{B(x_1, Mr_0)}$, y teniendo en cuenta (3.2), se sigue que

$$\|x - x_0\| = \|x - x_1\| + \|x_1 - x_0\| \leq Mr_0 + (1-M)r_0 = r_0.$$

Por tanto, $G : \overline{B(x_0, r_0)} \rightarrow \overline{B(x_0, r_0)}$ y es una contracción con constante de Lipschitz $M < 1$. Entonces, por el Teorema del Punto Fijo se sigue la tesis. ■

A continuación, vamos a considerar el caso particular en que el conjunto convexo Ω donde está definido y acotado el operador L_F sea una bola cerrada $\Omega = \overline{B(w_0, R_0)}$.

Veamos que se sigue verificando el Teorema 3.2 en lo referente a los puntos de salida que podemos escoger en esta nueva situación.

Corolario 3.3. *En las condiciones habituales, sean $x_0 \in \Omega = \overline{B(w_0, R_0)}$ y r_0 dado por (3.2). Si $\|w_0 - x_0\| \leq R_0 - r_0$, entonces la sucesión de Newton (3.1) es convergente a la única raíz de la ecuación $F(x) = 0$, tomando como punto de salida cualquier punto de la bola $\overline{B(x_0, r_0)}$.*

Demostración. Observar que dado $z \in \overline{B(x_0, r_0)}$ se verifica que

$$\|z - w_0\| \leq \|z_0 - x_0\| + \|x_0 - w_0\| \leq R_0.$$

Por tanto, $\overline{B(x_0, r_0)} \subseteq \Omega = \overline{B(w_0, R_0)}$ y el resultado se sigue sin más que aplicar el Teorema 3.2. ■

2.4 Ampliación del dominio de puntos de salida

En este apartado vamos a centrar nuestro estudio en la determinación de dominios de puntos de salida, lo más amplios posibles, para el método de Newton

$$x_{n+1} = x_n - [F'(x_n)]^{-1}F(x_n). \quad (4.1)$$

En el Teorema 3.2 vimos que si el punto de salida x_0 de la sucesión de Newton (4.1) cumplía una determinada condición, no sólo tenemos garantizada la convergencia desde dicho punto, sino también desde todos los puntos de un cierto entorno suyo.

En el Corolario 3.3, para el caso en que el conjunto Ω , donde el operador L_F está acotado, sea una bola cerrada $\Omega = \overline{B(w_0, R_0)}$, ya vimos cómo encontrar dicho entorno de puntos de salida. Continuando en esta situación, es decir, con Ω una bola cerrada, vamos a suponer que w_0 , el centro de dicha bola, es un punto de salida, esto es,

$$\varphi(w_0) \leq R_0(1 - M),$$

donde

$$\varphi(w_0) = \|[F'(w_0)]^{-1}F(w_0)\|. \quad (4.2)$$

En estas condiciones, el Corolario 3.3 nos asegura que (4.1) es convergente empezando en cualquier punto de la bola $\overline{B(w_0, r_0)}$ con

$$r_0 = \frac{\varphi(w_0)}{1 - M}.$$

A continuación, y en función de la cantidad $\varphi(w_0)$, vamos a ampliar el dominio de salida dado por el Corolario 3.3.

Teorema 4.1. Sea $\Omega = \overline{B(w_0, R_0)}$ una bola donde existe $L_F(x)$ y donde además $\|L_F(x)\| \leq M < 1$. Sea $\varphi(w_0)$ dada por (4.2). Entonces:

- (i) Si $\varphi(w_0) \leq (1 - M)^2 R_0$, la sucesión de Newton (4.1) es convergente empezando en cualquier punto de Ω .
- (ii) Si $(1 - M)^2 R_0 < \varphi(w_0) \leq \frac{1-M}{1+M} R_0$, la sucesión de Newton (4.1) es convergente empezando en cualquier punto de la bola

$$\overline{B\left(w_0, \frac{R_0}{M} - \frac{\varphi(w_0)}{M(1-M)}\right)}.$$

- (iii) Si $\frac{1-M}{1+M} R_0 < \varphi(w_0) \leq (1 - M) R_0$, la sucesión de Newton (4.1) es convergente empezando en cualquier punto de la bola

$$\overline{B\left(w_0, \frac{\varphi(w_0)}{1-M}\right)}.$$

Demostración. Supongamos que, empezando en un punto x_0 de Ω , el método (4.1) nos proporciona una sucesión que está contenida en Ω . Se prueba fácilmente que en este caso dicha sucesión es de Cauchy, y por tanto convergente. En efecto,

$$\begin{aligned} \|x_{n+1} - x_n\| &= \|G(x_n) - G(x_{n-1})\| \\ &\leq \sup_{x \in [x_{n-1}, x_n]} \|L_F(x)\| \|x_n - x_{n-1}\| \leq M \|x_n - x_{n-1}\|. \end{aligned}$$

luego,

$$\|x_{n+1} - x_n\| \leq M^n \|x_1 - x_0\|.$$

Así, la sucesión $\{x_n\}$ converge a algún $x^* \in \overline{B(w_0, R_0)}$. Como G es continua, $G(x^*) = x^*$ y por tanto x^* es solución de $F(x) = 0$.

En consecuencia, sólo tenemos que comprobar qué condición debe cumplir un punto x_0 para que la sucesión dada por (4.1) comenzando en dicho punto esté contenida en Ω . Observar que si denotamos por $\Gamma(w_0)$ a $[F'(w_0)]^{-1}$ tenemos que

$$\|x_1 - w_0\| \leq \|G(x_0) - G(w_0) - \Gamma(w_0)F(w_0)\| \leq M\|x_0 - w_0\| + \varphi(w_0).$$

$$\|x_2 - w_0\| \leq \|G(x_1) - G(w_0) - \Gamma(w_0)F(w_0)\| \leq M\|x_1 - w_0\| + \varphi(w_0)$$

$$\leq M^2 \|x_0 - w_0\| + \varphi(w_0)(1 + M)$$

y, en general,

$$\begin{aligned} \|x_n - w_0\| &\leq M^n \|x_0 - w_0 - \Gamma(w_0)F(w_0)\| (1 + M + \dots + M^{n-1}) \\ &< M \|x_0 - w_0\| + \frac{1}{1 - M} \varphi(w_0). \end{aligned}$$

De aquí obtenemos una condición suficiente para que x_n pertenezca a la bola $\overline{B}(w_0, R_0)$. Esta consiste en que

$$M \|x_0 - w_0\| + \frac{1}{1 - M} \varphi(w_0) \leq R_0,$$

lo que ocurre si

$$\|x_0 - w_0\| \leq \frac{R_0}{M} - \frac{1}{M(1 - M)} \varphi(w_0). \quad (4.3)$$

Observar que el segundo miembro de la desigualdad anterior es una cantidad positiva, ya que, al ser w_0 un punto de salida de (4.1),

$$\frac{\varphi(w_0)}{1 - M} \leq R_0.$$

Como $\|x_0 - w_0\| \leq R_0$, una condición suficiente para que se cumpla (4.3) es que

$$R_0 \leq \frac{R_0}{M} - \frac{\varphi(w_0)}{M(1 - M)}$$

o, equivalentemente,

$$\frac{1}{M(1 - M)} \varphi(w_0) \leq R_0 \left(\frac{1}{M} - 1 \right),$$

y esto ocurre si $\varphi(w_0) \leq R_0(1 - M)^2$. Por tanto, en este caso, todos los puntos de $\Omega = \overline{B}(w_0, R_0)$ tomados como puntos de salida de (4.1) nos darían una sucesión convergente. De aquí se sigue (i).

Por otra parte, si

$$R_0(1 - M)^2 < \varphi(w_0) \leq \frac{1 - M}{1 + M} R_0,$$

entonces

$$\frac{R_0}{M} - \frac{1}{M(1 - M)} \varphi(w_0) \geq \frac{\varphi(w_0)}{1 - M}.$$

Por tanto, en este caso, si $x \in B\left(w_0, \frac{R_0}{M} - \frac{\|[F'(w_0)]^{-1}F(w_0)\|}{M(1 - M)}\right)$, la condición (4.3) es cierta y, en consecuencia, la sucesión (4.1) es convergente empezando en cualquier punto de dicha bola, que es mayor que la dada por el Corolario 3.3.

Por último, si

$$\frac{1-M}{1+M}R_0 < \|\Gamma(w_0)F(w_0)\| \leq (1-M)R_0,$$

entonces no conseguimos mejorar la bola de puntos de salida dada por el Corolario 3.3. ■

Observación. Las bolas que aparecen en los apartados (i), (ii) y (iii) verifican las siguientes relaciones de contenido

$$B\left(w_0, \frac{\varphi(w_0)}{(1-M)}\right) \subseteq B\left(w_0, \frac{R_0}{M} - \frac{\varphi(w_0)}{M(1-M)}\right) \subseteq \Omega.$$

Notemos que en este último teorema no hemos expresado las estimaciones del error, aunque se seguirán verificando las indicadas en los resultados de la sección anterior.

Como aplicación de estos resultados, veremos que hay ecuaciones en las que podemos considerar puntos de salida que no verifican las condiciones de Kantorovich (ver Teorema 1.4.5) y, por tanto, no está garantizada la convergencia de la sucesión (4.1) comenzando en dichos puntos y aplicando los resultados clásicos de convergencia estudiados en el Capítulo 1; sin embargo, si cumplen nuestras hipótesis y, por tanto, sí podemos asegurar la convergencia de la sucesión de Newton. Dada una ecuación $F(x) = 0$, hemos conseguido ampliar el dominio de puntos de salida para que la sucesión (4.1) sea convergente a la solución de dicha ecuación. No obstante, la velocidad de convergencia de la sucesión formada empezando en un punto donde se cumplen nuestras condiciones y no las de Kantorovich puede ser más lenta, ya que ésta depende de la rapidez con que la cantidad M^n se acerque a cero, que es menor cuando M está cerca de uno.

Veamos unos ejemplos en los que se consideran puntos de salida para la sucesión de Newton que hacen de esta una sucesión convergente, aunque no se verifican las condiciones de Kantorovich.

Ejemplo 4.2. Consideramos la ecuación escalar $p(t) = 0$, siendo p el polinomio $p(t) = (t-1)(t-3) = t^2 - 4t + 3$. En este caso se tiene

$$L_p(t) = \frac{p(t)p''(t)}{(p'(t))^2} = \frac{t^2 - 4t + 3}{2(t-2)^2}.$$

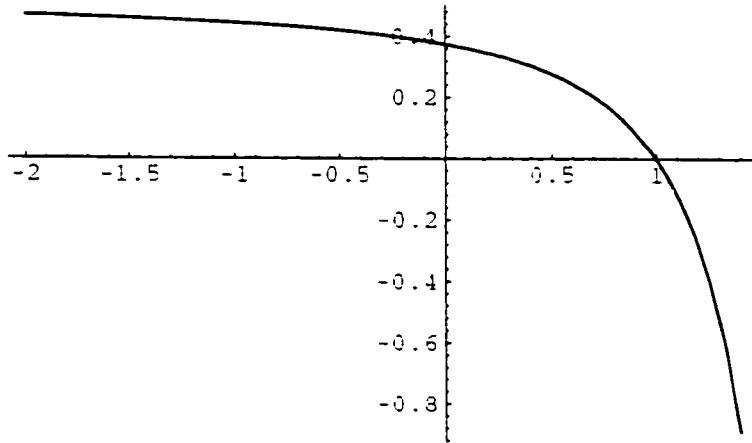


Figura 2.1: $L_p(t)$ con $p(t) = t^2 - 4t + 3$

Se prueba fácilmente que $|L_p(t)| < 1$ si $t \in (-\infty, 2 - \frac{\sqrt{3}}{3})$.

Sea $\Omega = [-100, 1.36754]$. Entonces

$$M = \sup_{t \in \Omega} |L_p(t)| = |L_p(1.36754)| = 3/4.$$

Tomamos $t_0 = 0.88$ y denotamos

$$r_0 = \frac{|p(t_0)/p'(t_0)|}{1 - M} = -\frac{4p(t_0)}{p'(t_0)} = 0.45429.$$

Como $\overline{B(t_0, r_0)} = [0.42571, 1.33429] \subseteq \Omega$, sabemos, por el Teorema 3.2, que la sucesión de Newton dada por (4.1) es convergente comenzando desde cualquier punto del intervalo $[0.42571, 1.33429]$. En particular, en el extremo superior del intervalo, $\tilde{t}_0 = 1.33429$, la sucesión (4.1) es convergente y sin embargo no se dan las condiciones de convergencia de Kantorovich. En efecto, sean

$$a = \left| \frac{p(\tilde{t}_0)}{p'(\tilde{t}_0)} \right|, \quad k = \left| \frac{p''(\tilde{t}_0)}{p'(\tilde{t}_0)} \right|;$$

entonces, como $p''(t)$ es constante, $ak = |L_p(\tilde{t}_0)| = 0.668221 > 1/2$ y comenzando en \tilde{t}_0 no se garantiza la convergencia de la sucesión (4.1) con las condiciones de Kantorovich. Sin embargo, con nuestras hipótesis sí se prueba la convergencia. ■

Ejemplo 4.3. En este ejemplo consideramos el sistema de ecuaciones no lineales que resulta de hacer $F(x, y) = 0$ si $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ viene dada por

$$F(x, y) = (xy - 1, xy + x - 2y).$$

Tenemos entonces

$$F'(x, y) = \begin{pmatrix} y & x \\ y + 1 & x - 2 \end{pmatrix}.$$

Por tanto, si (x, y) no está en la recta $x + 2y = 0$, existe $[F'(x, y)]^{-1}$ y este operador lineal viene dado por

$$[F'(x, y)]^{-1} = \frac{1}{x + 2y} \begin{pmatrix} 2 - x & x \\ y + 1 & -y \end{pmatrix}.$$

Teniendo en cuenta la expresión conocida para la derivada segunda de un operador de \mathbb{R}^2 en \mathbb{R}^2 , se tiene que

$$F''(x, y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

En este caso el operador lineal $L_{F(x, y)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ actúa de la manera siguiente. Dado $(s, t) \in \mathbb{R}^2$,

$$\begin{aligned} L_{F(x, y)}(s, t) &= [F'(x, y)]^{-1} F''(x, y) ([F'(x, y)]^{-1} F(x, y), (s, t)) \\ &= \frac{\lambda}{(x + 2y)^2} \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \text{con } \lambda = t(x^2 + x - 2) + s(2y^2 - y - 1). \end{aligned}$$

Si en \mathbb{R}^2 trabajamos con la norma del máximo

$$\|(x, y)\| = \max\{|x|, |y|\},$$

tenemos que

$$\begin{aligned} \|L_{F(x, y)}(s, t)\| &= \max \left\{ \frac{2|\lambda|}{(x + 2y)^2}, \frac{|\lambda|}{(x + 2y)^2} \right\} = \frac{2|\lambda|}{(x + 2y)^2} \\ &\leq 2 \frac{|x^2 + x - 2| + |2y^2 - y - 1|}{(x + 2y)^2} \|(s, t)\|. \end{aligned}$$

En consecuencia,

$$\|L_{F(x, y)}\| \leq 2 \frac{|x^2 + x - 2| + |2y^2 - y - 1|}{(x + 2y)^2} \equiv g(x, y).$$

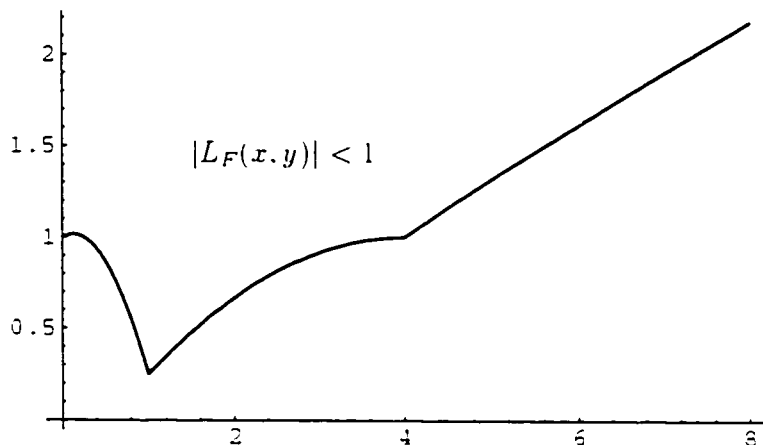


Figura 2.2: Región donde $|L_F(x, y)| < 1$

Se tiene que $\|L_F(x, y)\| \leq 1$ en la región de la Figura 2.2. Considerando $w_0 = (1.13, 1.13)$ y $R_0 = 0.43$ se tiene que

$$\Omega = \overline{B(w_0, 0.43)} = [0.7, 1.56] \times [0.7, 1.56].$$

Entonces

$$M = \sup_{(x,y) \in \Omega} \|L_F(x, y)\| \leq g(0.7, 0.7) = 0.7.$$

Por otra parte, como

$$(1 - M)^2 R_0 < \frac{1 - M}{1 + M} r_0 < \|[F'(x_0)]^{-1} F(x_0)\| < (1 - M) R_0,$$

nos encontramos en la situación (iii) del Teorema 4.1, y por tanto en la del Corolario 3.3. Así calculamos

$$\frac{\|[F'(w_0)]^{-1} F(w_0)\|}{1 - M} = 0.4083815,$$

donde la norma considerada es la norma habitual para matrices definida en el primer capítulo.

Entonces sabemos que la sucesión (4.1) es convergente desde cualquier punto de la bola $\overline{B(w_0, 0.4083815)}$. En esta bola, sin embargo, hay puntos donde no se garantizan las condiciones de Kantorovich. Por ejemplo, en $\tilde{w}_0 = (x_0, y_0) = (0.73, 0.73)$ tenemos que

$$\alpha = \|[F'(\tilde{w}_0)]^{-1} F(\tilde{w}_0)\| = 0.336753.$$

$$b = \|[F'(\tilde{w}_0)]^{-1}F''(X)\| = 1.826484.$$

y por tanto $ab = 0.615 > 1/2$ y no estamos en las condiciones del Teorema I.4.5. En resumen, empezando en $\tilde{w}_0 = (0.73, 0.73)$ no se garantiza la convergencia de la sucesión (4.1) con las condiciones de Kantorovich y sí con las nuestras (ver Tabla 2.1). ■

iteración	x_n	y_n
0	0.730000000000	0.730000000000
1	1.06657534247	1.03328767123
2	1.00141463870	1.00070731935
3	1.00000066644	1.00000033322
4	1.00000000000	1.00000000000
5	1.00000000000	1.00000000000

Tabla 2.1: Sucesión de Newton en el Ejemplo 3.4

En el Teorema 4.1, considerando que el conjunto donde el operador L_F está acotado es una bola cerrada $\Omega = \overline{B(w_0, R_0)}$, encontramos, según los valores de $\|[F'(w_0)]^{-1}F(w_0)\|$, unas bolas de puntos de salida que aseguraban la convergencia de la sucesión (4.1). A continuación, utilizando otra técnica en la demostración, mejoramos considerablemente este resultado ya que probamos la convergencia de la sucesión de Newton (4.1) empezando desde cualquier punto de Ω sin más que exigir que el centro de la bola, tomado como punto de salida de (4.1), proporcione una sucesión convergente, condición exigida en el Teorema 3.1.

Teorema 4.4. *Sea $\Omega = \overline{B(w_0, R_0)}$ una bola tal que $\|L_F(x)\| \leq M < 1$, $x \in \Omega$. Si $\|[F'(w_0)]^{-1}F(w_0)\| \leq (1 - M)R_0$, entonces la sucesión dada por (4.1) empezando en cualquier punto de Ω es convergente a x^* , única solución de $F(x) = 0$ en Ω .*

Demostración. Observar que la sucesión (4.1) se obtiene aplicando el método de las aproximaciones sucesivas al operador G definido por

$$G(x) = x - [F'(x)]^{-1}F(x).$$

Teniendo en cuenta que la derivada de G en un punto x_0 es $L_F(x_0)$, se sigue que G es una contracción. Además, G envía la bola Ω en sí misma. En efecto, sea x un

punto cualquiera de Ω . Entonces

$$\|x_1 - G(x)\| = \|G(x_0) - G(x)\| \leq M\|x_0 - x\| \leq MR_0.$$

Así, G aplica la bola $\overline{B(x_0, R_0)}$ en la bola $\overline{B(x_1, MR_0)}$.

Además $\overline{B(x_1, MR_0)} \subseteq \overline{B(x_0, R_0)}$. En efecto, si $x \in \overline{B(x_1, MR_0)}$, como

$$r_0 = \frac{\|[F'(w_0)]^{-1}F(w_0)\|}{1-M} \leq R_0,$$

de (3.4) se sigue que

$$\begin{aligned} \|x - x_0\| &= \|x - x_1\| + \|x_1 - x_0\| \\ &\leq MR_0 + (1-M)r_0 \leq MR_0 + (1-M)R_0 = R_0. \end{aligned}$$

Por tanto, utilizando el Teorema del Punto Fijo se concluye el resultado. ■

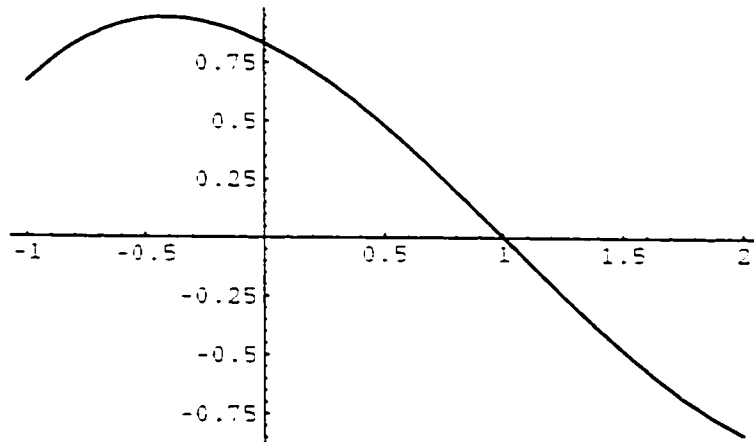


Figura 2.3: Gráfica de $p(t) = \frac{(t-1)^2}{6} - (t-1)$

Veamos a continuación si con este resultado podemos ampliar el dominio de puntos de salida para el método de Newton obtenido en el Ejemplo 4.3. Allí probamos que la sucesión (4.1) era convergente considerando como punto de salida cualquiera de la bola $\overline{B(w_0, 0.4083815)}$ con $w_0 = (1.13, 1.13)$. A partir de este resultado que acabamos de probar, como

$$\frac{\|[F'(w_0)]^{-1}F(w_0)\|}{1-M} = 0.4083815 \leq R_0 = 0.43.$$

la convergencia queda garantizada comenzando la sucesión de Newton en cualquier punto de la bola $\overline{B}(w_0, 0.43)$.

Para acabar esta sección, haremos un comentario acerca del método de Newton en el caso escalar. Para resolver una ecuación escalar $f(t) = 0$ usando el método de Newton, la mayoría de los autores exigen que la función f esté en las condiciones de Fourier, esto es, $f' \neq 0$, que f'' no cambie de signo en un intervalo $[a, b]$ y que además $f(a)f(b) < 0$. Una función que tenga una raíz en un punto de inflexión no está en las condiciones de Fourier. Sin embargo, con nuestros resultados de convergencia, podemos demostrar que el método de Newton converge a la solución de $f(x) = 0$ también en esta situación.

iteración	t_n
0	0.6000000000000000
1	1.0231884057971016
2	0.9999995842730325
3	1.0000000000000000
4	1.0000000000000000

Tabla 2.2: Sucesión de Newton en el Ejemplo 4.5

Ejemplo 4.5. El polinomio $p(t) = \frac{(t-1)^3}{6} - (t-1)$ tiene una raíz en el punto $t = 1$, que además es un punto de inflexión de dicho polinomio. Con la ayuda del Teorema 4.1 vamos a encontrar un intervalo de puntos de salida que nos garanticen la convergencia de la sucesión dada por (4.1) para $F = p$.

Consideramos $\overline{B}(0.8, 0.35) = [0.45, 1.15]$. Entonces

$$M = \sup_{t \in \Omega} |L_f(t)| = |L_f(0.45)| = \left| \frac{f(0.45)f''(0.45)}{f'(0.45)^2} \right| = 0.39875.$$

Como

$$\begin{aligned} \frac{1-M}{1+M} R_0 &= 0.172 < \left| \frac{f(w_0)}{f'(w_0)} \right| = \left| \frac{f(0.8)}{f'(0.8)} \right| = 0.2027 \\ &< (1-M)R_0 = (1-0.39875)0.4 = 0.2405 \end{aligned}$$

sabemos que el método de Newton es convergente empezando en cualquier punto de la bola

$$\overline{B}\left(w_0, \frac{\| [f'(w_0)]^{-1} f(w_0) \|}{(1-M)}\right) = \overline{B}(0.8, 0.2027). \quad \blacksquare$$

Capítulo 3

Estudio del error en el método de Newton

3.1 Introducción

Resulta una comprobación inmediata que las estimaciones del error que se obtienen a partir del Teorema del Punto Fijo en el Capítulo 2 son peores que las obtenidas utilizando sucesiones mayorizantes en el Capítulo 1. Dentro de estas últimas, las más estudiadas han sido las obtenidas en las condiciones del Teorema I.4.5, que hemos denominado condiciones de Kantorovich (ver también [66] o [68]). Es por ello que nuestro primer estudio relativo a estimaciones del error en el método de Newton hará referencia a esta situación. Así, resulta evidente que existen infinitas funciones que verifican este tipo de condiciones de Kantorovich y que por tanto nos proporcionan infinitas sucesiones mayorizantes. Cada una de ellas nos da una estimación del error diferente, dependiendo ésta de la situación de las raíces y, por supuesto, del grado de convexidad logarítmico de dichas funciones, ya que estamos aplicando el método de Newton a una ecuación escalar y ya vimos en la Sección I.3 la influencia que tiene la convexidad en la velocidad de convergencia de este método.

Veremos que entre todas las funciones mayorizantes que verifican las condiciones de Kantorovich, la mejor estimación del error nos la da un polinomio de segundo grado. Tanto en la construcción de este polinomio como en la demostración de que nos proporciona la mejor sucesión mayorizante, interviene de forma fundamental el grado de convexidad logarítmico. Para ello utilizaremos el hecho de que sabemos

comparar, en función de dicho grado, la velocidad de convergencia de dos sucesiones reales obtenidas aplicando el método de Newton a dos funciones determinadas.

El siguiente estudio a realizar consistirá en considerar directamente las condiciones del Teorema I.4.4, las condiciones generales de Kantorovich. Veremos que, en este caso, las estimaciones del error son mejores que en las condiciones anteriores. No obstante, su aplicación práctica puede resultar complicada.

Por último, terminamos este capítulo aplicando las técnicas de Gragg-Tapia [49] para mejorar la estimación del error obtenida a partir de sucesiones mayorizantes. Para ello, a partir de un resultado de Miel [72], obtenemos unas cotas intermedias entre la sucesión en el espacio de Banach y la sucesión mayorizante que, si bien son peores que las obtenidas por Gragg y Tapia, son más sencillas de calcular desde el punto de vista práctico.

3.2 Estudio del error en las condiciones de Kantorovich

El Teorema I.4.1 garantizaba la convergencia de la sucesión de Newton en espacios de Banach para resolver la ecuación $F(x) = 0$, dada por

$$x_{n+1} = x_n - \Gamma_n F(x_n), \quad \Gamma_n = [F'(x_n)]^{-1}, \quad (2.1)$$

en términos de una función real f que mayoriza, en cierto sentido, al operador F y a sus derivadas.

En el Teorema I.4.5, se parte de las denominadas condiciones de Kantorovich, esto es

$$\|\Gamma_0 F(x_0)\| \leq a, \quad (2.2)$$

$$\|\Gamma_0 F''(x)\| \leq k \quad \text{si } x \in \Omega_0 = \overline{B(x_0, r_0)}, \quad (2.3)$$

y $ak \leq 1/2$.

Bajo estas hipótesis, la función real que los autores prefijan para probar la convergencia de la sucesión (2.1) y que proporciona la estimación del error cometido es el polinomio

$$p(t) = \frac{k}{2}t^2 - t + a. \quad (2.4)$$

A continuación, veremos una nueva demostración de la convergencia de la sucesión de Newton para un operador F que verifica (2.2) y (2.3), utilizando el operador L_F definido en la Sección I.2. En esta demostración, construiremos el polinomio (2.4), que otros autores prefijan, y estableceremos las bases para probar la convergencia de (2.1) en condiciones distintas a (2.2) y (2.3), construyendo en estos casos las funciones correspondientes que sustituyan al polinomio (2.4).

En el Teorema II.2.3 veíamos una estimación del grado de convexidad logarítmico de un operador F que cumple unas ciertas condiciones. En concreto, sean $F : X \rightarrow Y$ un operador entre espacios de Banach y $x_0 \in X$ un punto tal que existe $\Gamma_0 = [F'(x_0)]^{-1}$. Supongamos además que $\Gamma_0 F''$ es continua y $\|\Gamma_0 F''(x)\| \leq k$ para $x \in \Omega_0$. Entonces, existe $L_F(x)$ para todo $x \in B(x_0, 1/k) \cap \Omega_0$ y además

$$\|L_F(x)\| \leq \frac{k}{(1 - k\|x - x_0\|)^2} \|\Gamma_0 F(x)\|. \quad (2.5)$$

Utilizando la expresión (2.5) podemos dar otro método para encontrar sucesiones mayorizantes. Para ello nos será de utilidad el siguiente resultado que, como consecuencia de los Teoremas I.4.2 y II.2.2, nos da una condición para obtener sucesiones mayorizantes en términos del operador grado de convexidad logarítmico.

Teorema 2.1. *Sea F un operador definido entre dos espacios de Banach X e Y . Consideramos el método de Newton para resolver la ecuación $F(x) = 0$ definido por $x_{n+1} = G(x_n)$, siendo*

$$G(x) = x - [F'(x)]^{-1} F(x).$$

Entonces, si existe una función real f de manera que

$$(i) \quad \|\Gamma_0 F(x_0)\| \leq -\frac{f(t_0)}{f'(t_0)},$$

$$(ii) \quad \|L_F(x)\| \leq L_f(t) \text{ si } \|x - x_0\| \leq t - t_0,$$

la sucesión de Newton construida para resolver la ecuación $f(t) = 0$ empezando en t_0 es mayorizante de la sucesión de Newton $\{x_n\}$ definida en el espacio de Banach X .

Demostración. Basta aplicar el Teorema I.4.2 al operador G y a la función $g(t) = t - \frac{f(t)}{f'(t)}$ y tener en cuenta que $G'(x) = L_F(x)$, $g'(t) = L_f(t)$, como quedó probado en el Teorema II.2.2. ■

Lo que hacemos a continuación es ver que, en las condiciones de Kantorovich y empezando en x_0 , podemos definir el siguiente término de la sucesión de Newton, x_1 .

Además, para todo x perteneciente al segmento que une x_0 y x_1 se prueba la existencia del operador lineal $L_F(x)$ y su acotación por el grado de convexidad logarítmico de una función real. Esta función real es precisamente el polinomio definido en (2.4). La estimación dada es la mejor posible ya que todas las desigualdades se transformarían en igualdades en el caso de trabajar con funciones reales. Los resultados obtenidos para $x \in [x_0, x_1]$ se generalizan a $x \in [x_1, x_2]$ y, en general, a $x \in [x_n, x_{n+1}]$, siendo $\{x_n\}$ la sucesión de Newton dada por (2.1).

En consecuencia, y aplicando el Teorema 2.1, podemos construir una sucesión mayorizante de $\{x_n\}$ sin más que emplear el método de Newton para resolver la ecuación escalar $p(t) = 0$, con p definido por (2.4). Además, a partir de las observaciones anteriores y de la relación entre el grado de convexidad logarítmico y la velocidad de convergencia del método de Newton (Sección I.3), probaremos que esta sucesión mayorizante es la que nos proporciona una estimación óptima del error.

Teorema 2.2. *Consideramos la ecuación $F(x) = 0$, siendo F un operador entre dos espacios de Banach, X e Y , dos veces diferenciable Fréchet, con F'' continua en un conjunto convexo Ω_0 . Sea x_0 un punto de Ω_0 en el que existe $\Gamma_0 = [F'(x_0)]^{-1}$. Supongamos que $\|\Gamma_0 F''(x)\| \leq k$ si $x \in \Omega_0$ y que además $ak \leq 1/2$, siendo $\|\Gamma_0 F(x_0)\| \leq a$. Supongamos también que $B(x_0, 1/k) \subseteq \Omega_0$. Entonces la sucesión $\{t_n\}$ definida por*

$$t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}, \quad t_0 = 0,$$

siendo p el polinomio dado por (2.4), es una sucesión mayorizante de $\{x_n\}$ definida por (2.1).

Demostración. En las condiciones anteriores, y comenzando en x_0 , escribimos la sucesión de Newton (2.1) de la siguiente forma

$$x_{n+1} = x_n - \Gamma_n F(x_n) = G(x_n), \quad \text{donde } \Gamma_n = [F'(x_n)]^{-1}.$$

Con nuestras condiciones tenemos que x_1 está definido. Para poder definir x_2 necesitamos que exista Γ_1 . Para ello basta con que x_1 esté en la bola $B(x_0, 1/k)$ (ver Teorema II.2.3) o, lo que es igual,

$$\|x_1 - x_0\| = \|\Gamma_0 F(x_0)\| \leq a < 1/k,$$

es decir, $ak < 1$, lo cual es cierto por hipótesis.

En este caso, y teniendo en cuenta el Teorema II.2.2, se tiene que

$$\int_{x_0}^{x_1} L_F(z) dz = G(x_1) - G(x_0) = x_2 - x_1. \quad (2.6)$$

Además si $x \in [x_0, x_1]$, $x = x_0 + s(x_1 - x_0) = x_0 + s\Gamma_0 F(x_0)$ con $0 \leq s \leq 1$. Por tanto, si desarrollamos en serie de Taylor

$$\begin{aligned} \Gamma_0 F(x) &= \Gamma_0 F(x_0) + (x - x_0) + \int_{x_0}^x \Gamma_0 F''(z)(x - z) dz \\ &= \Gamma_0 F(x_0)(1 - s) + \int_{x_0}^x \Gamma_0 F''(z)(x - z) dz. \end{aligned}$$

Luego, teniendo en cuenta (2.2) y (2.3),

$$\begin{aligned} \|\Gamma_0 F(x)\| &\leq (1 - s)\|\Gamma_0 F(x_0)\| + \left\| \int_{x_0}^x \Gamma_0 F''(z)(x - z) dz \right\| \\ &\leq (1 - s)a + \frac{k}{2}\|x - x_0\|^2 = (1 - s)a + \frac{k}{2}s^2a^2. \end{aligned}$$

Si hacemos $t = \|x - x_0\| = s\|x_1 - x_0\| = sa$, tenemos que

$$\|\Gamma_0 F(x)\| \leq a - t + \frac{k}{2}t^2.$$

En particular, para $x = x_1$, tenemos

$$\|\Gamma_0 F(x_1)\| \leq -\frac{p(t_1)}{p'(t_0)}. \quad (2.7)$$

Así, para $x \in [x_0, x_1]$ y $t = \|x - x_0\|$, se sigue que

$$\|L_F(x)\| \leq \frac{k(a - t + \frac{k}{2}t^2)}{(1 - kt)^2}.$$

La expresión del segundo miembro es el grado de convexidad logarítmico del polinomio (2.4). Por tanto, si $x \in [x_0, x_1]$, $t = \|x - x_0\|$ y $p(t)$ es el polinomio dado por (2.4)

$$\|L_F(x)\| \leq L_p(t). \quad (2.8)$$

Observar que este polinomio tiene raíces reales si y sólo si $1 - 2ak \geq 0$, es decir $ak \leq 1/2$. Con esta condición se garantiza la existencia de Γ_1 que nos apareció anteriormente.

Si utilizamos el método de Newton para aproximar la menor de las raíces del polinomio $p(t)$ dado por (2.4), empezando en $t_0 = 0$, obtenemos que

$$\|x_1 - x_0\| \leq a = t_1 = t_0 - \frac{p(t_0)}{p'(t_0)}.$$

luego $\|x_1 - x_0\| \leq t_1 - t_0$.

¿Qué ocurre con los demás términos? De (2.6) y (2.8) se sigue que

$$\|x_2 - x_1\| \leq \left\| \int_{x_0}^{x_1} L_F(z) dz \right\| \leq \int_{t_0}^{t_1} L_p(t) dt = t_2 - t_1.$$

Además $\|x_2 - x_0\| \leq \|x_2 - x_1\| + \|x_1 - x_0\| \leq (t_2 - t_1) + (t_1 - t_0)$

$$= t_2 < r_1 = \frac{1 - \sqrt{1 - 2ak}}{k} < \frac{1}{k}.$$

Por tanto, $x_2 \in B(x_0, 1/k)$, existe Γ_2 y podemos definir x_3 . Por otra parte,

$$x_3 - x_2 = \int_{x_1}^{x_2} L_F(z) dz$$

y, para que $\|x_3 - x_2\| \leq t_3 - t_2$, basta con que se cumpla $\|L_F(x)\| \leq L_p(t)$ si $x \in [x_1, x_2]$ y $\|x - x_1\| = t$.

Para probar esto, veamos que podemos reemplazar x_0 por x_1 en el paso anterior.

En primer lugar, notemos que se verifica

$$\|\Gamma_1 F''(x)\| \leq \|\Gamma_1 F'(x_0)\| \|\Gamma_0 F''(x)\|.$$

Por el Lema de Banach,

$$\|\Gamma_1 F'(x_0)\| = \|[\Gamma_0 F'(x_1)]^{-1}\| \leq \frac{1}{1 - k\|x_1 - x_0\|} = \frac{1}{1 - kt_1} = \frac{p'(t_0)}{p'(t_1)}. \quad (2.9)$$

y en consecuencia se sigue que

$$\|\Gamma_1 F''(x)\| \leq \frac{k}{1 - kt_1}.$$

Además, como

$$\int_{x_1}^x \Gamma_1 F''(z) dz = \Gamma_1 F'(x) - I$$

obtenemos que

$$\|\Gamma_1 F'(x) - I\| \leq \frac{k}{1 - kt_1} \|x - x_1\| < 1$$

ya que, si $x \in [x_1, x_2]$, $x = x_1 + s(x_2 - x_1)$, con $0 \leq s \leq 1$ y entonces

$$\|x - x_1\| = s\|x_2 - x_1\| \leq \|x_2 - x_1\| \leq t_2 - t_1 < \frac{1}{k} - t_1 = \frac{1 - kt_1}{k}.$$

Luego, aplicando de nuevo el Lema de Banach obtenemos que existe $[\Gamma_1 F'(x)]^{-1}$ y

$$\|[\Gamma_1 F'(x)]^{-1}\| \leq \frac{1}{1 - \frac{k}{1 - kt_1} \|x - x_1\|} = \frac{1 - kt_1}{1 - k(t_1 + \|x - x_1\|)}.$$

Por último, si $x \in [x_1, x_2]$, desarrollando en serie de Taylor tenemos que

$$\Gamma_1 F(x) = \Gamma_1 F(x_1) + (x - x_1) + \int_{x_1}^x \Gamma_1 F''(z)(x - z) dz.$$

Como $x - x_1 = s(x_2 - x_1) = -s\Gamma_1 F(x_1)$, se sigue que

$$\Gamma_1 F(x) = (1 - s)\Gamma_1 F(x_1) + \int_{x_1}^x \Gamma_1 F''(z)(x - z) dz$$

y entonces

$$\begin{aligned} \|\Gamma_1 F(x)\| &\leq (1 - s)\|\Gamma_1 F(x_1)\| + \left\| \int_{x_1}^x \Gamma_1 F''(z)(x - z) dz \right\| \\ &\leq (1 - s)\|\Gamma_1 F(x_1)\| + \frac{k}{1 - kt_1} \frac{\|x - x_1\|^2}{2} \\ &= (1 - s)\|\Gamma_1 F(x_1)\| + \frac{k}{1 - kt_1} \frac{s^2 \|\Gamma_1 F(x_1)\|^2}{2}. \end{aligned}$$

Si denotamos $t = \|x - x_1\| = s\|\Gamma_1 F(x_1)\|$, con

$$t \in [0, \|\Gamma_1 F(x_1)\|] = [0, \|x - x_1\|] \subseteq [0, t_2 - t_1],$$

tenemos que

$$\|\Gamma_1 F(x)\| \leq \|\Gamma_1 F(x_1)\| - t + \frac{k}{2(1 - kt_1)} t^2.$$

Por otra parte, de (2.7), obtenemos

$$\begin{aligned} \|\Gamma_1 F(x_1)\| &\leq \|\Gamma_1 F'(x_0)\| \|\Gamma_0 F(x_1)\| = \|[\Gamma_0 F'(x_1)]^{-1}\| \|\Gamma_0 F(x_1)\| \\ &\leq \frac{1}{1 - kt_1} p(t_1) = -\frac{p(t_1)}{p'(t_1)}. \end{aligned} \tag{2.10}$$

y por tanto

$$\begin{aligned} \|\Gamma_1 F(x)\| &\leq \frac{1}{1 - kt_1} \left[p(t_1) - t(1 - kt_1) + \frac{k}{2} t^2 \right] \\ &= \frac{1}{1 - kt_1} \left[a - t_1 + \frac{k}{2} t_1^2 - t + kt_1 t + \frac{k}{2} t^2 \right] = \frac{1}{1 - kt_1} p(t_1 + t). \end{aligned}$$

Entonces, si $x \in [x_1, x_2]$ y $\|x - x_1\| = t$,

$$\begin{aligned} \|L_F(x)\| &\leq \|[\Gamma_1 F'(x)]^{-1}\| \|\Gamma_1 F''(x)\| \|[\Gamma_1 F'(x)]^{-1}\| \|\Gamma_1 F(x)\| \\ &\leq \frac{1 - kt_1}{1 - k(t_1 + t)} \frac{k}{1 - kt_1} \frac{1 - kt_1}{1 - k(t_1 + t)} \frac{p(t_1 + t)}{1 - kt_1} = L_p(t_1 + t) \end{aligned}$$

$$= \frac{kp(t_1 + t)}{(1 - k(t_1 + t))^2} = \frac{kp(t_1 + t)}{(p'(t_1 + t))^2} = L_p(t_1 + t).$$

De aquí, si $x \in [x_1, x_2]$,

$$\|x - x_1\| \leq s\|x_2 - x_1\| \leq \|x_2 - x_1\| \leq t_2 - t_1,$$

y por tanto

$$\begin{aligned} \|x_3 - x_2\| &= \left\| \int_{x_1}^{x_2} L_F(x) dx \right\| \leq \int_0^{t_2 - t_1} L_p(t_1 + t) dt \\ &= \int_{t_1}^{t_2} L_p(z) dz = t_3 - t_2. \end{aligned}$$

Se puede probar sin dificultad que podemos reemplazar x_2 por x_1 en el paso anterior, y en general x_{n+1} por x_n , obteniendo así que si $x \in [x_{n-1}, x_n]$ y $t = \|x - x_{n-1}\|$ entonces $\|L_F(x)\| \leq L_p(t_{n-1} + t)$ y, en consecuencia,

$$\begin{aligned} \|x_{n+1} - x_n\| &= \left\| \int_{x_{n-1}}^{x_n} L_F(x) dx \right\| \leq \int_0^{t_n - t_{n-1}} L_p(t_{n-1} + t) dt \\ &= \int_{t_{n-1}}^{t_n} L_p(z) dz = t_{n+1} - t_n. \quad \blacksquare \end{aligned}$$

Observación. Como $\{t_n\}$ es una sucesión convergente, hemos probado, con nuevos argumentos, la convergencia de la sucesión $\{x_n\}$ para F y x_0 en las condiciones de Kantorovich. Nos encontramos ahora en condiciones de enunciar el siguiente resultado.

Teorema 2.3. Con la misma notación que el teorema anterior, la sucesión de Newton $\{x_n\}$ converge a una solución x^* de la ecuación $F(x) = 0$. Además, si denotamos t^* y t^{**} a las raíces del polinomio (2.4), ($t^* \leq t^{**}$), entonces

$$x^* \in \overline{B(x_0, t^*)}$$

y además es la única solución en

$$B(x_0, t^{**}) \quad \text{si } t^* < t^{**}$$

o en

$$\overline{B(x_0, t^{**})} \quad \text{si } t^* = t^{**}.$$

Demostración. Del teorema anterior se sigue que

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n.$$

Por tanto, $\{x_n\}$ es una sucesión convergente a un límite $x^* \in \overline{B(x_0, t^*)} \subseteq B(x_0, 1/k)$. Como se puso de manifiesto en la demostración del Teorema 2.2.

$$\|\Gamma_n F(x_n)\| \leq -\frac{p(t_n)}{p'(t_n)}.$$

(ver (2.10)). Por tanto, la convergencia de $\{t_n\}$ a una solución de $p(t) = 0$ y el hecho de que $F'(x^*)$ es inversible (ver Teorema II.2.3) implican que $F(x^*) = 0$.

Para comprobar la unicidad, supongamos que $t^* < t^{**}$ y que existe otra solución \tilde{x} de $F(x) = 0$ en la bola $B(x_0, t^{**})$. Entonces.

$$\|\tilde{x} - x_0\| \leq \rho(t^{**} - t_0), \quad \text{con } 0 < \rho < 1.$$

Por inducción vamos a probar que

$$\|\tilde{x} - x_n\| \leq \rho^{2^n} (t^{**} - t_n), \quad n \geq 0.$$

Para ello observar que

$$\begin{aligned} \tilde{x} - x_{n+1} &= \tilde{x} - x_n + \Gamma_n F(x_n) \\ &= -\Gamma_n [F(\tilde{x}) - F(x_n) - F'(x_n)(\tilde{x} - x_n)] \\ &= -\Gamma_n \int_{x_n}^{\tilde{x}} F''(x)(\tilde{x} - x) dx \\ &= [-\Gamma_0 F'(x_n)]^{-1} \int_0^1 \Gamma_0 F''(x_n + t(\tilde{x} - x_n))(1-t) dt (\tilde{x} - x_n)^2. \end{aligned}$$

Así, teniendo en cuenta que $\|\Gamma_0 F''(x)\| \leq k$ y que, por (2.9),

$$\|[\Gamma_0 F'(x_n)]^{-1}\| \leq \frac{1}{1 - kt_n} = \frac{-1}{p'(t_n)}.$$

deducimos que

$$\begin{aligned} \|\tilde{x} - x_{n+1}\| &\leq \frac{-1}{p'(t_n)} \int_0^1 k(1-t) dt \|\tilde{x} - x_n\|^2 \\ &= \frac{-k}{2p'(t_n)} \|\tilde{x} - x_n\|^2. \end{aligned}$$

Reiterando el mismo procedimiento para el polinomio p tenemos que

$$t^{**} - t_{n+1} = \frac{-k}{2p'(t_n)} (t^{**} - t_n)^2.$$

De estas dos igualdades, y utilizando la hipótesis de inducción, se sigue que

$$\|\tilde{x} - x_{n+1}\| \leq \frac{-k(t^{**} - t_n)^2}{2p'(t_n)} \left(\frac{\|\tilde{x} - x_n\|}{t^{**} - t_n} \right)^2 \leq (t^{**} - t_{n+1}) \rho^{2^{n+1}}.$$

Hemos probado así que $\|\tilde{x} - x_n\| \leq (t^{**} - t_n)\rho^{2^n}$, si $t^* < t^{**}$.

Si $t^* = t^{**}$, se prueba de forma similar que $\|\tilde{x} - x_n\| \leq t^* - t_n$.

En ambas situaciones se tiene que $\|\tilde{x} - x_n\| \rightarrow 0$ cuando $n \rightarrow \infty$. En consecuencia, $\tilde{x} = x^*$. ■

A continuación, vamos a dar una estimación del error para la sucesión $\{x_n\}$ utilizando la estimación del error para polinomios de segundo grado debida a Ostrowski [83].

Teorema 2.4. *En las condiciones de los dos teoremas anteriores, obtenemos la siguiente estimación del error*

$$\|x^* - x_n\| \leq \|t^* - t_n\| = \frac{t^{**} - t^*}{1 - \theta^{2^n}} \theta^{2^n}, \quad \text{con } \theta = \frac{t^*}{t^{**}}, \quad \text{si } ak < 1/2.$$

$$\|x^* - x_n\| \leq t^* - t_n = \frac{t^*}{2^n} \quad \text{si } ak = 1/2.$$

Demostración. En primer lugar, tener en cuenta que $\{t_n\}$ es mayorizante de $\{x_n\}$, entonces, si $p \in \mathbb{N}$,

$$\begin{aligned} \|x_{n+p} - x_n\| &\leq \|x_{n+p} - x_{n+p-1}\| + \cdots + \|x_{n+1} - x_n\| \\ &\leq (t_{n+p} - t_{n+p-1}) + \cdots + (t_{n+1} - t_n) = t_{n+p} - t_n. \end{aligned}$$

Si hacemos $p \rightarrow \infty$, de la convergencia de $\{x_n\}$ y $\{t_n\}$ se sigue que

$$\|x^* - x_n\| \leq t^* - t_n.$$

Vamos a dar una expresión que nos permita obtener la estimación del error $t^* - t_n$ sin necesidad de calcular los términos de la sucesión $\{t_n\}$. Escribimos $p(t) = (k/2)(t^* - t)(t^{**} - t)$ y denotamos $a_n = t^* - t_n$, $b_n = t^{**} - t_n$ para $n \geq 0$. Se tiene entonces que $p(t_n) = (k/2)a_nb_n$ y $p'(t_n) = -(k/2)(a_n + b_n)$.

De la expresión del método de Newton se sigue que

$$a_{n+1} = t^* - t_{n+1} = t^* - t_n + \frac{p(t_n)}{p'(t_n)} = a_n - \frac{a_nb_n}{a_n + b_n} = \frac{a_n^2}{a_n + b_n}, \quad n \geq 0.$$

y análogamente

$$b_{n+1} = \frac{b_n^2}{a_n + b_n}, \quad n \geq 0.$$

Utilizando estas relaciones tenemos

$$\frac{a_n}{b_n} = \left[\frac{a_{n-1}}{b_{n-1}} \right]^2 = \left[\frac{a_{n-2}}{b_{n-2}} \right]^{2^2} = \cdots = \left[\frac{a_0}{b_0} \right]^{2^n} = \left[\frac{t^*}{t^{**}} \right]^{2^n} = \theta^{2^n}, \quad n \geq 0.$$

Si $ak < 1/2$, entonces $t^* < t^{**}$. En este caso existe $d > 0$ de manera que $t^{**} = t^* + d$. En consecuencia,

$$t^* - t_n = (t^{**} - t_n)\theta^{2^n} = (t^* - t_n)\theta^{2^n} + d\theta^{2^n}.$$

de donde

$$t^* - t_n = \frac{d\theta^{2^n}}{1 - \theta^{2^n}}, \quad n \geq 0.$$

En definitiva, hemos obtenido que

$$\|x^* - x_n\| \leq t^* - t_n = \frac{d\theta^{2^n}}{1 - \theta^{2^n}}, \quad n \geq 0.$$

Si $ak = 1/2$ entonces $t^* = t^{**} = 1/k$ y, en consecuencia, $a_n = b_n$, $a_{n+1} = \frac{a_n}{2}$ y, por tanto, $a_n = \frac{t^*}{2^n}$. ■

La siguiente cuestión es preguntarnos si este polinomio (2.4) considerado habitualmente nos proporciona la mejor estimación del error para el método de Newton en las condiciones de Kantorovich.

En concreto, sea f una función que verifica

$$(i') \quad \|\Gamma_0 F(x_0)\| = -\frac{p(t_0)}{p'(t_0)} < -\frac{f(t_0)}{f'(t_0)}$$

$$(ii') \quad \sup_{x \in \Omega_0} \|\Gamma_0 F''(x)\| = -\frac{p''(t)}{p'(t_0)} < -\frac{f''(t)}{f'(t_0)}, \quad t \in [0, t^*].$$

donde t^* es la menor raíz de (2.4). El Teorema 1.4.4 nos asegura que la sucesión de Newton construida para resolver la ecuación $f(t) = 0$ es mayorizante de (2.1). Nos preguntamos ahora qué función nos proporcionará mejor estimación del error, p ó f . Para ello tendremos en cuenta la relación existente entre la convexidad y la velocidad de convergencia del método de Newton vista en la Sección 1.3. Compararemos así las estimaciones del error obtenidas por dichas funciones. Para ello, veremos el siguiente resultado, fundamental en nuestro estudio.

Lema 2.5. Sean f y g dos funciones decrecientes y convexas en el intervalo $[0, t^*]$ tales que, para $t \in (0, t^*)$, se verifica

$$a) \quad -\frac{g(0)}{g'(0)} < -\frac{f(0)}{f'(0)}.$$

$$\text{b) } -\frac{g''(t)}{g'(0)} < -\frac{f''(t)}{f'(0)}.$$

Entonces $L_f(t) > L_g(t)$ en $[0, t^*]$.

Demostración. Consideramos $t \in (0, t^*)$. Entonces, por b) tenemos

$$0 < \int_0^t \left(-\frac{f''(s)}{f'(0)} + \frac{g''(s)}{g'(0)} \right) ds = -\frac{f'(t)}{f'(0)} + \frac{g'(t)}{g'(0)},$$

de donde

$$-\frac{f'(t)}{f'(0)} > -\frac{g'(t)}{g'(0)}, \quad t \in (0, t^*). \quad (2.11)$$

Teniendo en cuenta (2.11) podemos escribir

$$0 < \int_0^t \left(-\frac{f'(s)}{f'(0)} + \frac{g'(s)}{g'(0)} \right) ds = \left(-\frac{f(t)}{f'(0)} + \frac{g(t)}{g'(0)} \right) - \left(-\frac{f(0)}{f'(0)} + \frac{g(0)}{g'(0)} \right),$$

y aplicando a) se sigue que

$$-\frac{f(t)}{f'(0)} > -\frac{g(t)}{g'(0)}, \quad t \in (0, t^*). \quad (2.12)$$

Ahora bien, por b) tenemos

$$f''(t) > \frac{f'(0)}{g'(0)} g''(t), \quad t \in (0, t^*). \quad (2.13)$$

y de (2.11) se deduce que

$$f'(t) > \frac{f'(0)}{g'(0)} g'(t), \quad t \in (0, t^*),$$

y como $f''(t) \leq 0$ en $(0, t^*)$, obtenemos

$$\frac{1}{f'(t)^2} > \frac{1}{\left(\frac{f'(0)}{g'(0)}\right)^2 g'(t)^2}. \quad t \in (0, t^*). \quad (2.14)$$

En consecuencia, por (2.12), (2.13) y (2.14),

$$L_f(t) = \frac{f(t)f''(t)}{f'(t)^2} > \frac{g(t)g''(t)}{g'(t)^2} = L_g(t), \quad t \in (0, t^*). \quad \blacksquare$$

Nota. El resultado anterior sigue siendo válido si se da la igualdad en una de las dos condiciones a) y b) anteriores.

Por el Lema 2.5. sabemos que en las condiciones (i') y (ii') se tiene que $L_p(t) < L_f(t)$ para $t \in [0, t^*]$. En consecuencia, por el Teorema I.3.3. es conocido que el método de Newton aplicado a la función f es más lento que si se lo aplicamos a p , y así, la estimación del error obtenida para el polinomio (2.4) será óptima en las condiciones (i') y (ii'). A partir de estas consideraciones, podemos enunciar el siguiente resultado.

Teorema 2.6. *El polinomio (2.4) proporciona la mejor estimación del error para la sucesión de Newton (2.1) en las condiciones de Kantorovich, empleando sucesiones mayorizantes.*

Demostración. En los comentarios previos al enunciado. ■

Nota. En [50] se encontró, usando también argumentos de convexidad, el polinomio de segundo grado que proporcionaba una estimación del error óptima para el método de Newton en las siguientes condiciones, un poco más fuertes que las de Kantorovich:

$$(I) \quad \|\Gamma_0\| \leq \frac{-1}{p'(t_0)}.$$

$$(II) \quad \|F(x_0)\| \leq p(t_0).$$

$$(III) \quad \|F''(x)\| \leq p''(t) \text{ para } \|x - x_0\| \leq t - t_0.$$

Este polinomio resultó ser, como cabía esperar, el dado por (2.4).

Ahora bien, notemos que (i') y (ii') son unas condiciones más restrictivas que las condiciones generales de Kantorovich dadas en el Teorema I.4.4. En efecto, si a viene dado por (2.2) y

$$J(t) = \|\Gamma_0 F''(x)\| \quad \text{si } \|x - x_0\| = t - t_0,$$

tenemos que el polinomio (2.4) surge de considerar la cota más elemental que puede conseguirse de la curva $J(t)$, que es la recta

$$y = \sup_{x \in \Omega_0} \|\Gamma_0 F''(x)\|.$$

Ahora bien, si la función J es "suficientemente buena" veremos que puede mejorarse esta acotación y, por el Lema 2.5, obtener una mejor estimación del error en el método de Newton.

Así, sea \mathcal{J} una función continua en el intervalo $[t_0, t^*]$ tal que la ecuación diferencial

$$y''(t) = r\mathcal{J}(t) \quad (2.15)$$

con las condiciones iniciales $y'(t_0) = -r$, $y(t_0) = a$, siendo $r = \|\Gamma_0\|^{-1}$ y a dado por (2.2), tenga única solución. Entonces, la solución de esta ecuación nos proporcionará la estimación óptima del error mediante sucesiones mayorizantes en el método de Newton.

Encontrar esta función \mathcal{J} en la práctica puede resultar muy complicado. En el siguiente ejemplo damos un sistema de ecuaciones para el cual se puede determinar la función \mathcal{J} y comparamos las estimaciones del error obtenidas a partir del polinomio (2.4) y de la solución de la ecuación diferencial (2.15).

Ejemplo 2.7. Consideramos el sistema de ecuaciones no lineales resultante de hacer

$$F(x, y) = 0, \quad \text{con } F(x, y) = \left(x^3 - 2y + \frac{1}{3}y^3 - 4x + \frac{2}{3} \right). \quad (2.16)$$

Estudiamos la posible utilización del método de Newton, y estimamos el error en el caso de que haya convergencia. Teniendo en cuenta la expresión matricial de las derivadas primera y segunda del operador F vista en el Capítulo 1, así como la de sus normas, tenemos que

$$F'(x, y) = \begin{pmatrix} 3x^2 & -2 \\ -4 & 3y^2 \end{pmatrix}, \quad F''(x, y) = \begin{pmatrix} 6x & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 6y \end{pmatrix}.$$

Si partimos de $(x_0, y_0) = (0, 0)$, entonces

$$[F'(x, y)]^{-1} = \frac{1}{\Delta} \begin{pmatrix} 3y^2 & 2 \\ 4 & 3x^2 \end{pmatrix} \quad \text{con } \Delta = 9x^2y^2 - 8$$

y

$$\Gamma_0 = [F'(x_0, y_0)]^{-1} = \begin{pmatrix} 0 & -1/4 \\ -1/2 & 0 \end{pmatrix}.$$

$$\Gamma_0 F''(x, y) = \begin{pmatrix} 0 & 0 \\ -3x & 0 \\ 0 & -3y/2 \\ 0 & 0 \end{pmatrix}.$$

De aquí, $a = \|\Gamma_0 F(x_0, y_0)\| = 1/6$ y $\|\Gamma_0 F''(x, y)\| = 3\|(x, y)\|$.

Estudiamos la solución de la ecuación (2.16) en Ω_0 , la bola de centro (x_0, y_0) y radio 0.97 (para asegurar la existencia de $[F'(x, y)]^{-1}$ en dicha bola). Tomando como de costumbre $t_0 = 0$, se sigue que si $\|(x, y)\| = t$ entonces,

$$\|\Gamma_0 F''(x, y)\| = 3t.$$

Si utilizamos el Teorema I.4.5 de Kantorovich, tenemos que

$$\|\Gamma_0 F''(x, y)\| = 3t \leq 3(0.97) = 2.91 = k,$$

por consiguiente, el polinomio de la forma (2.4) correspondiente es $p(t) = \frac{2.91}{2}t^2 - t + \frac{1}{6}$ (notar que $ak = \frac{2.91}{6} < \frac{1}{2}$).

Sin embargo, si resolvemos la ecuación diferencial (2.15), en el caso de que $J(t) = 3t$, $r = 1/2$, $a = 1/3$, obtenemos la función $q(t) = t^3 - 2t + \frac{1}{3}$, que nos da una estimación del error sensiblemente mejor que la dada por el polinomio anterior, como se puede ver en la Tabla 3.1.

Denotamos t^* a la menor raíz del polinomio p y s^* a la menor raíz positiva de q . Entonces las sucesiones $\{t_n\}$ y $\{s_n\}$ obtenidas aplicando el método de Newton a p y q respectivamente nos dan las siguientes estimaciones para el error cometido en cada iteración de (2.1)

$$E_n = t^* - t_n, \quad E'_n = s^* - s_n.$$

En la tabla se dan las primeras estimaciones del error cometido por ambas sucesiones y se ve claramente que son mucho mejor las obtenidas con la sucesión $\{s_n\}$.

3.3 Técnica de Gragg y Tapia

En la sección anterior estudiamos la ecuación $F(x) = 0$ cuando F es un operador dos veces diferenciable y cuya derivada segunda verifica la condición $\|\Gamma_0 F''(x)\| \leq k$ en un conjunto convexo Ω_0 (condiciones de Kantorovich).

En concreto, vimos que si $x_0 \in \Omega_0$ es un punto en el que existe el operador inverso $\Gamma_0 = [F'(x_0)]^{-1}$ y se cumple la condición $ak \leq 1/2$, siendo $\|\Gamma_0 F(x_0)\| \leq a$, entonces la sucesión

$$x_{n+1} = x_n - \Gamma_n F(x_n), \quad \Gamma_n = [F'(x_n)]^{-1} \tag{3.1}$$

iteración	E_n	E'_n
0	0.284121965376	0.169083656929
1	0.117455298709	0.002416990262
2	0.038976334308	0.000001531325
3	0.007711681245	0.000000000000
4	0.000442272563	0.000000000000
5	0.000001631050	0.000000000000
6	0.000000000023	0.000000000000
7	0.000000000000	0.000000000000

Tabla 3.1: Estimaciones del error

está bien definida y converge a x^* solución de $F(x) = 0$.

Probamos que la sucesión de Newton

$$t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}, \quad t_0 = 0. \quad (3.2)$$

construida para resolver la ecuación $p(t) = 0$, siendo p el polinomio

$$p(t) = \frac{k}{2}t^2 - t + a, \quad (3.3)$$

es mayorizante de $\{x_n\}$, esto es,

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \quad n \geq 0. \quad (3.4)$$

y

$$\|x^* - x_n\| \leq t^* - t_n, \quad n \geq 0. \quad (3.5)$$

siendo t^* la menor raíz positiva de $p(t) = 0$.

Esto nos permite dar una estimación del error para la sucesión $\{x_n\}$ a partir de la conocida estimación del error para polinomios de segundo grado debida a Ostrowsky [83], como pusimos de manifiesto en el Teorema 2.4. Teniendo en cuenta la relación entre la velocidad de convergencia de la sucesión de Newton y el grado de convexidad logarítmico, probamos (Teorema 2.6) que dicha estimación es la mejor que se puede obtener mediante el empleo de sucesiones mayorizantes.

Nuestro siguiente objetivo será encontrar estimaciones intermedias entre la sucesión en el espacio de Banach y su mayorizante, es decir, unas ciertas cantidades σ_n que verifiquen

$$\|x^* - x_n\| \leq \sigma_n \leq t^* - t_n.$$

La técnica que seguiremos se debe a Gragg y Tapia [49]. Posteriormente, obtendremos nuevas estimaciones intermedias con la ayuda de unos resultados que pueden encontrarse en el trabajo de Miel [72], y que extienden las obtenidas por Gragg y Tapia.

En primer lugar probaremos un lema que necesitaremos más adelante [109].

Lema 3.1. *Con la notación anterior, se tiene que*

$$\frac{\|x^* - x_{n+1}\|}{t^* - t_{n+1}} \leq \left(\frac{\|x^* - x_n\|}{t^* - t_n} \right)^2.$$

Demostración. Teniendo en cuenta (3.1) y que x^* es solución de $F(x) = 0$, podemos escribir

$$\begin{aligned} x^* - x_{n+1} &= x^* - x_n + \Gamma_n F(x_n) \\ &= -\Gamma_n (F(x^*) - F(x_n) - F'(x_n)(x^* - x_n)) = -\Gamma_n \left(\int_{x_n}^{x^*} [F'(x) - F'(x_n)] dx \right) \\ &= -\Gamma_n \left(\int_0^1 [F'(x_n + t(x^* - x_n)) - F'(x_n)] (x^* - x_n) dt \right). \end{aligned}$$

Observar que a partir de la acotación $\|\Gamma_0 F''(x)\| \leq k$, y por el Teorema del Valor Medio, deducimos

$$\|\Gamma_0 [F'(x) - F'(y)]\| \leq \sup_{z \in \{x, y\}} \|\Gamma_0 F''(z)\| \|x - y\| \leq k \|x - y\|.$$

Utilizando esta desigualdad, se tiene

$$\begin{aligned} \|x^* - x_{n+1}\| &\leq \|\Gamma_n F'(x_0)\| \left\| \int_{x_n}^{x^*} \Gamma_0 [F'(x) - F'(x_n)] dx \right\| \\ &\leq \frac{-1}{p'(t_n)} \int_0^1 kt dt \|x^* - x_n\|^2 = \frac{-1}{p'(t_n)} \int_0^1 kt(t^* - t_n)^2 dt \left[\frac{\|x^* - x_n\|}{t^* - t_n} \right]^2. \end{aligned}$$

y haciendo el cambio de variable $z = t_n + t(t^* - t_n)$, se sigue

$$\begin{aligned} \|x^* - x_{n+1}\| &\leq \frac{-1}{p'(t_n)} \int_{t_n}^{t^*} k(z - t_n) dz \left[\frac{\|x^* - x_n\|}{t^* - t_n} \right]^2 \\ &= \frac{-1}{p'(t_n)} \int_{t_n}^{t^*} [p'(z) - p'(t_n)] dz \left[\frac{\|x^* - x_n\|}{t^* - t_n} \right]^2 \\ &= \int_{t_n}^{t^*} \left[1 - \frac{p'(z)}{p'(t_n)} \right] dz \left[\frac{\|x^* - x_n\|}{t^* - t_n} \right]^2 = \left(t^* - t_n + \frac{p(t_n)}{p'(t_n)} \right) \left[\frac{\|x^* - x_n\|}{t^* - t_n} \right]^2 \end{aligned}$$

$$= (t^* - t_{n+1}) \left[\frac{\|x^* - x_n\|}{t^* - t_n} \right]^2. \quad \blacksquare$$

En el resultado siguiente [49] vamos a mejorar la estimación para la cantidad $\|x^* - x_n\|$ dada por (3.5), obteniendo además una estimación inferior de dicha cantidad.

Teorema 3.2. *Sea la ecuación $F(x) = 0$, donde F es un operador entre dos espacios de Banach X e Y dos veces diferenciable en un conjunto convexo $\Omega_0 \subseteq X$. Sea x_0 un punto de Ω_0 en el que existe $\Gamma_0 = [F'(x_0)]^{-1}$. Supongamos además que se verifica $\|\Gamma_0 F''(x)\| \leq k$ para $x \in \Omega_0$ y que $ak \leq 1/2$, siendo $\|\Gamma_0 F(x_0)\| \leq a$. Llamamos*

$$k_n = \frac{t^* - t_{n+1}}{(t^* - t_n)^2} \quad (3.6)$$

y $\delta_n = \|x_{n+1} - x_n\|$. Sean τ_n y σ_n la única raíz positiva y la menor raíz positiva de los polinomios

$$o_n(t) = k_n t^2 - t + \delta_n \quad (3.7)$$

y

$$v_n(t) = k_n t^2 + t - \delta_n.$$

respectivamente. Entonces

$$\tau_n \leq \|x^* - x_n\| \leq \sigma_n \leq t^* - t_n.$$

Demostración. Por el lema anterior,

$$\begin{aligned} k_n \|x^* - x_n\|^2 &\geq \|x^* - x_{n+1}\| \geq \left| \|x^* - x_n\| - \|x_{n+1} - x_n\| \right| \\ &\geq \|x^* - x_n\| - \|x_{n+1} - x_n\|. \end{aligned}$$

y por tanto, $o_n(\|x^* - x_n\|) \geq 0$.

Observar además que

$$o_n(t^* - t_n) = k_n (t^* - t_n)^2 - (t^* - t_n) + \delta_n = \delta_n - (t_{n+1} - t_n) < 0$$

ya que la sucesión $\{t_n\}$ mayoriza a $\{x_n\}$, y se cumple (3.4).

Como $0 < \|x^* - x_n\| \leq t^* - t_n$ tenemos, (ver Figura 3.1).

$$\|x^* - x_n\| \leq \sigma_n \leq t^* - t_n.$$

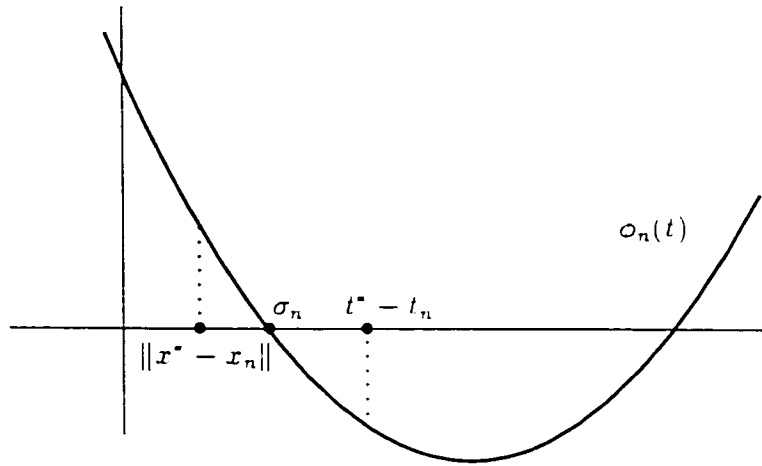


Figura 3.1: Gráfica del polinomio $\phi_n(t)$

Por otra parte, para obtener cotas inferiores, notar que también se cumple

$$\begin{aligned} k_n \|x^* - x_n\|^2 &\geq \|x^* - x_{n+1}\| \geq \left| \|x^* - x_n\| - \|x_{n+1} - x_n\| \right| \\ &\geq \|x_{n+1} - x_n\| - \|x^* - x_n\|. \end{aligned}$$

Luego $v_n(\|x^* - x_n\|) \geq 0$. (ver Figura 3.2).

Como $v_n(0) < 0$, deducimos que

$$\|x^* - x_n\| \geq \tau_n.$$

de donde se sigue el resultado. ■

Corolario 3.3. En la situación del teorema anterior se tiene

$$\frac{1}{5}\delta_n < (-2 + \sqrt{8})\delta_n \leq \|x^* - x_n\| \leq 2\delta_n.$$

Demostración. El polinomio ϕ_n tiene un mínimo en $t = \frac{1}{2k_n}$. Además

$$\phi_n\left(\frac{1}{2k_n}\right) = \delta_n - \frac{1}{4k_n} < 0.$$

luego $k_n\delta_n < 1/4$. Como $\phi_n(2\delta_n) = \delta_n(4k_n\delta_n - 1) < 0$, se sigue que $\sigma_n < 2\delta_n$ y, por tanto,

$$\|x^* - x_n\| \leq 2\delta_n.$$

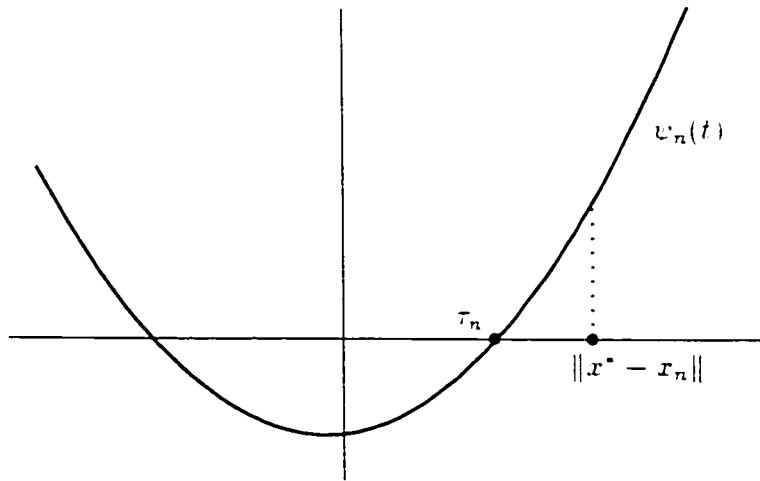


Figura 3.2: Gráfica del polinomio $w_n(t)$

Por otra parte, sea

$$\overline{w}_n(t) = \frac{1}{4\delta_n}t^2 + t - \delta_n.$$

Observar que $w_n(t) < \overline{w}_n(t)$ para $t > 0$. Entonces, denotando $a = -2 + \sqrt{8}$, se tiene

$$\overline{w}_n(a\delta_n) = 4\delta_n(a^2 + 4a - 4) = 0,$$

y así, $a\delta_n < \tau_n \leq \|x^* - x_n\|$. ■

Vamos a ilustrar estos resultados con un ejemplo en el que se pondrá de manifiesto la mejora en la estimación del error obtenida con las cantidades σ_n y τ_n .

Ejemplo 3.4. Consideramos el polinomio de tercer grado $F(x) = x^3 - 9x + 3$, y lo estudiamos en el intervalo $[0, 5/4]$, donde sabemos que tiene una única raíz. Si $x_0 = 0$, tenemos que

$$a = -\frac{F(0)}{F'(0)} = \frac{1}{3}.$$

$$b = \sup_{x \in [0, 5/4]} -\frac{F''(x)}{F'(0)} = \frac{5}{6}.$$

Como $ab = 5/18 < 1/2$, se cumplen las condiciones de Kantorovich y en consecuencia el polinomio de segundo grado

$$p(t) = \frac{5}{12}t^2 - t + \frac{1}{3}$$

mayoriza a F , esto es, la sucesión de Newton $\{t_n\}$ construida para resolver la ecuación $p(t) = 0$ es mayorizante de $\{x_n\}$, la sucesión de Newton para resolver $F(x) = 0$.

Los primeros términos de estas sucesiones son

$$t_0 = 0, \quad t_1 = \frac{1}{3}, \quad t_2 = \frac{31}{78};$$

$$x_0 = 0, \quad x_1 = \frac{1}{3}, \quad x_2 = \frac{79}{234}.$$

Teniendo en cuenta que la menor raíz de $p(t) = 0$ es $t^* = 2/5$, vamos a obtener, para $n = 0, 1$, las expresiones de k_n y δ_n que aparecen en los coeficientes de los polinomios ϕ_n y ψ_n definidos en el teorema anterior, y que en este caso son

$$k_0 = \frac{t^* - t_1}{(t^* - t_0)^2} = \frac{5}{12},$$

$$k_1 = \frac{t^* - t_2}{(t^* - t_1)^2} = \frac{15}{26},$$

$$\delta_0 = \|x_1 - x_0\| = \frac{1}{3},$$

$$\delta_1 = \|x_2 - x_1\| = \frac{1}{234}.$$

En consecuencia,

$$\phi_0(t) = k_0 t^2 - t + \delta_0 = \frac{5}{12} t^2 - t + \frac{1}{3}$$

y

$$\psi_0(t) = k_0 t^2 + t - \delta_0 = \frac{5}{12} t^2 + t - \frac{1}{3}.$$

Igualando a cero estos polinomios hallamos σ_0 y τ_0 , que son

$$\sigma_0 = \frac{2}{5} = 0.4, \quad \tau_0 = \frac{-6 + 2\sqrt{14}}{5} \simeq 0.2966629.$$

Obtenemos así las siguientes estimaciones para la cantidad $\|x^* - x_0\|$, con x^* la raíz de $F(x) = 0$ en el intervalo $[0, 5/4]$,

$$0.2966629 \leq \|x^* - x_0\| \leq \frac{2}{5} = 0.4 = t^* - t_0.$$

Observar que, en este caso, el polinomio ϕ_0 coincide con p y, en consecuencia, la acotación superior que obtenemos no mejora la que ya conocíamos y que viene dada por $t^* - t_0$.

Para $n = 1$ tenemos

$$o_1(t) = k_1 t^2 - t + \delta_1 = \frac{1}{26} t^2 - t + \frac{1}{234}.$$

$$v_1(t) = k_1 t^2 + t - \delta_1 = \frac{1}{26} t^2 + t - \frac{1}{234}.$$

Calculando las correspondientes raíces de estos polinomios, obtenemos

$$\sigma_1 = \frac{13 - \sqrt{\frac{502}{3}}}{15} \simeq 0.0042840867.$$

$$\tau_1 = \frac{-13 + \sqrt{\frac{512}{3}}}{15} \simeq 0.0042630196.$$

En consecuencia, se deducen las siguientes cotas del error para $\|x^* - x_1\|$

$$0.0042630196 \leq \|x^* - x_1\| \leq 0.0042840867$$

$$< t^* - t_1 = \frac{1}{15} = 0.066666\dots$$

Observar que las cotas del error obtenidas con las cantidades τ_1 y σ_1 mejoran sensiblemente la acotación superior $t^* - t_1$. No obstante, desde el punto de vista práctico, se plantea el inconveniente de que para estimar $\|x^* - x_n\|$ es necesario conocer x_{n+1} . ■

A continuación, vamos a obtener nuevas estimaciones intermedias para (3.5). Para ello necesitaremos los siguientes lemas, que pueden encontrarse en el trabajo de Miel [72].

Lema 3.5. Sean $\{x_n\}$ y $\{t_n\}$ las sucesiones definidas por (3.1) y (3.2) respectivamente. Entonces

$$\frac{\|x_{n+1} - x_n\|}{t_{n+1} - t_n} \leq \left[\frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} \right]^2. \quad (3.8)$$

Demostración. Notar que de (3.1) se sigue que

$$\|x_{n+1} - x_n\| = \|[F'(x_n)]^{-1} F(x_n)\| \leq \|[\Gamma_0 F'(x_n)]^{-1}\| \|\Gamma_0 F(x_n)\|.$$

Por otra parte, se tiene

$$\|[\Gamma_0 F'(x_n)]^{-1}\| \leq \frac{1}{1 - k\|x_n - x_0\|} \leq \frac{1}{1 - kt_n}.$$

Además,

$$\begin{aligned}\Gamma_0 F(x_n) &= \Gamma_0 [F(x_n) - F(x_{n-1}) - F'(x_{n-1})(x_n - x_{n-1})] \\ &= \int_{x_{n-1}}^{x_n} \Gamma_0 [F'(z) - F'(x_{n-1})] dz \\ &= \int_0^1 \Gamma_0 [F'(x_{n-1} + s(x_n - x_{n-1})) - F'(x_{n-1})] (x_n - x_{n-1}) ds.\end{aligned}$$

Como

$$\Gamma_0 [F'(x_{n-1} + s(x_n - x_{n-1})) - F'(x_{n-1})] \leq ks \|x_n - x_{n-1}\|.$$

se sigue que

$$\|\Gamma_0 F(x_n)\| \leq k \|x_n - x_{n-1}\|^2 \int_0^1 s ds = \frac{k}{2} \|x_n - x_{n-1}\|^2.$$

y en consecuencia,

$$\|x_{n+1} - x_n\| \leq \frac{k \|x_n - x_{n-1}\|^2}{2(1 - kt_n)}.$$

Por último, tenemos que probar que la igualdad

$$\frac{k}{2} \frac{1}{1 - kt_n} = \frac{t_{n+1} - t_n}{[t_n - t_{n-1}]^2}$$

es cierta. En efecto, si repetimos el proceso anterior, tomando como F el polinomio p definido en (3.3), tenemos que

$$\begin{aligned}t_{n+1} - t_n &= -\frac{p(t_n)}{p'(t_n)} = \frac{-1}{p'(t_n)} [p(t_n) - p(t_{n-1}) - p'(t_{n-1})(t_n - t_{n-1})] \\ &= \frac{1}{1 - kt_n} [p(t_n) - p(t_{n-1}) - p'(t_{n-1})(t_n - t_{n-1})].\end{aligned}$$

Como

$$\begin{aligned}p(t_n) - p(t_{n-1}) - p'(t_{n-1})(t_n - t_{n-1}) &= \int_{t_{n-1}}^{t_n} [p'(t) - p'(t_{n-1})] dt \\ &= \int_{t_{n-1}}^{t_n} k(t - t_{n-1}) dt = \frac{k}{2} (t_n - t_{n-1})^2.\end{aligned}$$

se sigue que el resultado es cierto. ■

Lema 3.6. Para $\{x_n\}$ y $\{t_n\}$ definidas por (3.1) y (3.2) respectivamente, las siguientes desigualdades son ciertas:

$$\begin{aligned} \|x_{n+1} - x_n\| &\stackrel{(1)}{\leq} \frac{t_{n+1} - t_n}{[t_n - t_{n-1}]^2} \|x_n - x_{n-1}\|^2 \stackrel{(2)}{\leq} \frac{t_{n+1} - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\| \\ &\stackrel{(3)}{\leq} \frac{t_{n+1} - t_n}{t_1 - t_0} \|x_1 - x_0\| \stackrel{(4)}{\leq} t_{n+1} - t_n. \end{aligned} \quad (3.9)$$

Demostración. La desigualdad (1) se sigue del Lema 3.5. Para la segunda desigualdad, notar que, por (3.4),

$$\frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} \leq 1,$$

por tanto,

$$\left[\frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} \right]^2 \leq \frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}}.$$

Observar además que, si $1 \leq \lambda \leq 2$, también es cierto que

$$\left[\frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} \right]^2 \leq \left[\frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} \right]^\lambda \leq \frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}}.$$

Reiterando la desigualdad dada por

$$\|x_{n+1} - x_n\| \leq \frac{t_{n+1} - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\|$$

se tiene que

$$\begin{aligned} (t_{n+1} - t_n) \frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} &\leq \frac{t_{n+1} - t_n}{t_n - t_{n-1}} \frac{t_n - t_{n-1}}{t_{n-1} - t_{n-2}} \|x_{n-1} - x_{n-2}\| \leq \dots \\ &\dots \leq \frac{t_{n+1} - t_n}{t_n - t_{n-1}} \frac{t_n - t_{n-1}}{t_{n-1} - t_{n-2}} \dots \frac{t_2 - t_1}{t_1 - t_0} \|x_1 - x_0\|, \end{aligned} \quad (3.10)$$

de donde se sigue la desigualdad (3).

Por último, sin más que usar que $\|x_1 - x_0\| \leq t_1 - t_0$, se deduce (4). ■

Cambiando el término independiente del polinomio ϕ_n definido en (3.7) por alguna de las cantidades intermedias dadas en el Lema 3.6, obtenemos nuevas cotas superiores para $\|x^* - x_n\|$, como veremos en el siguiente resultado.

Teorema 3.7. Sean $\{x_n\}$ y $\{t_n\}$ las sucesiones definidas por (3.1) y (3.2). Sean σ_n la menor raíz positiva del polinomio (3.7). $\sigma_n^{(1)}$, $\sigma_n^{(2)}$ y $\sigma_n^{(3)}$ las menores raíces positivas de los polinomios $\phi_n^{(i)}(t) = k_n t^2 - t + \delta_n^{(i)}$, $i = 1, 2, 3$, $n \geq 1$, donde k_n viene dado por (3.6) y

$$\delta_n^{(1)} = \frac{t_{n+1} - t_n}{(t_n - t_{n-1})^2} \|x_n - x_{n-1}\|^2, \quad \delta_n^{(2)} = \frac{t_{n+1} - t_n}{t_n - t_{n-1}} \|x_n - x_{n-1}\|^2,$$

$$\delta_n^{(3)} = \frac{t_{n+1} - t_n}{t_1} \|x_1 - x_0\|.$$

Entonces

$$\|x^* - x_n\| \leq \sigma_n \leq \sigma_n^{(1)} \leq \sigma_n^{(2)} \leq \sigma_n^{(3)} \leq t^* - t_n.$$

Demostración. Observar entonces que, por el Lema 3.4.

$$\phi_n^{(i)}(\|x^* - x_n\|) \geq \phi_n(\|x^* - x_n\|) \geq 0$$

y

$$\phi_n^{(i)}(t^* - t_n) = \delta_n^{(i)} - (t_{n+1} - t_n) \leq 0.$$

En consecuencia, teniendo en cuenta la forma de los polinomios $\phi_n^{(i)}$ y que $\|x^* - x_n\| \leq t^* - t_n$, si llamamos $\sigma_n^{(i)}$ a la menor raíz positiva de $\phi_n^{(i)}$, se sigue (ver Figura 3.3) que

$$\|x^* - x_n\| \leq \sigma_n^{(i)} \leq t^* - t_n.$$

Además, para cada t se tiene que

$$\phi_n(t) \leq \phi_n^{(1)}(t) \leq \phi_n^{(2)}(t) \leq \phi_n^{(3)}(t).$$

luego

$$\|x^* - x_n\| \leq \sigma_n \leq \sigma_n^{(1)} \leq \sigma_n^{(2)} \leq \sigma_n^{(3)} \leq t^* - t_n. \quad \blacksquare$$

Aunque, evidentemente, las estimaciones del error $\sigma_n^{(1)}$, $\sigma_n^{(2)}$ y $\sigma_n^{(3)}$ son peores que σ_n dada por Gragg y Tapia, notar el interés que, desde el punto de vista práctico, tienen las primeras, ya que en los tres casos, para estimar $\|x^* - x_n\|$, no es necesario conocer x_{n+1} , como ocurre con σ_n . Destacar en este sentido la importancia de la estimación $\sigma_n^{(3)}$, en la que simplemente conociendo los dos primeros términos de la sucesión en el espacio de Banach, podemos obtener, a partir de la sucesión mayorizante, las sucesivas estimaciones de la cantidad $\|x^* - x_n\|$.

Notar además que la cantidad

$$\frac{\|x_1 - x_0\|}{t_1 - t_0}$$

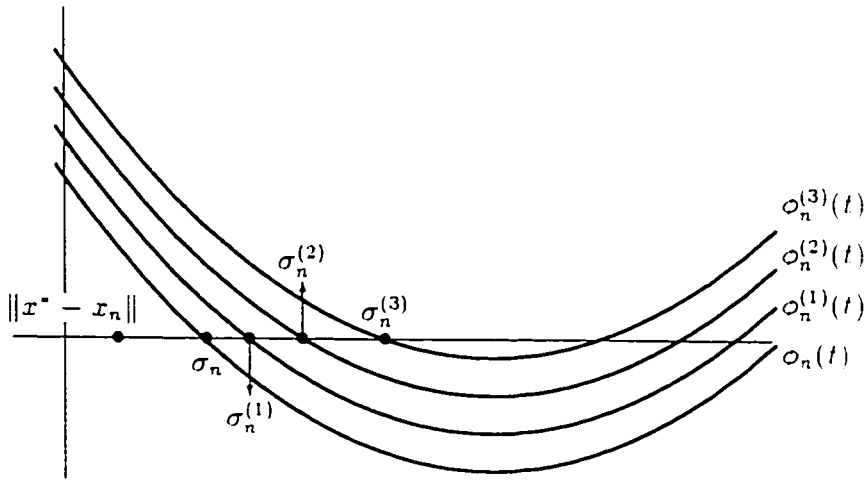


Figura 3.3: Gráficas de los polinomios $\phi_n(t)$ y $\phi_n^{(j)}(t)$, $j = 1, 2, 3$

puede ser sustituida en la desigualdad (3) del Lema 3.6 por

$$\frac{\|x_2 - x_1\|}{t_2 - t_1} \quad \text{o por} \quad \frac{\|x_{k+1} - x_k\|}{t_{k+1} - t_k}$$

para algún $k \leq n$ determinado (ver 3.10), obtenido en este caso

$$\begin{aligned} (t_{n+1} - t_n) \frac{\|x_n - x_{n-1}\|}{t_n - t_{n-1}} &\leq \frac{t_{n+1} - t_n}{t_k - t_{k-1}} \|x_k - x_{k-1}\| \\ &\leq \frac{t_{n+1} - t_n}{t_2 - t_1} \|x_2 - x_1\| \leq \frac{t_{n+1} - t_n}{t_1} \|x_1 - x_0\|. \end{aligned}$$

De la misma forma, reiterando la desigualdad (1) del Lema 3.6, podemos obtener otra serie de términos independientes de forma similar a $\delta_n^{(1)}$. Vamos a ilustrar estos comentarios con un ejemplo en el que hallamos las estimaciones $\sigma_j^{(1)}$, $\sigma_j^{(2)}$ y $\sigma_j^{(3)}$ para $j = 1, 2$.

Ejemplo 3.8. Con la misma notación que en el Ejemplo 3.4 tenemos que

$$\delta_1^{(1)} = \frac{t_2 - t_1}{(t_1 - t_0)^2} \|x_1 - x_0\|^2 = \frac{5}{78} = \frac{t_2 - t_1}{t_1 - t_0} \|x_1 - x_0\| = \delta_1^{(2)},$$

$$\delta_1^{(3)} = t_2 - t_1.$$

Observar que, en nuestro ejemplo, la desigualdad (4) del Lema 3.6 se transforma en igualdad, ya que $\|x_1 - x_0\| = t_1 - t_0$. En esta situación es cuando tiene pleno interés

el comentario previo a este ejemplo en el que se ofrecía la posibilidad de sustituir la expresión

$$\frac{\|x_1 - x_0\|}{t_1 - t_0},$$

que en este caso vale uno, por otra menor que uno, por ejemplo

$$\frac{\|x_2 - x_1\|}{t_2 - t_1}.$$

Las estimaciones $\sigma_1^{(1)} = \sigma_1^{(2)} = 1/15$, obtenidas igualando a cero los polinomios

$$o_1^{(1)}(t) = o_1^{(2)}(t) = k_1 t^2 - t + \delta_1^{(1)}$$

no mejoran la ya conocida $\|x^* - x_1\| \leq r_1 - t_1 = 1/15$.

Sin embargo, al acotar $\|x^* - x_2\|$ por $\sigma_2^{(1)}$ y $\sigma_2^{(2)}$ obtendremos unas notables mejoras. Además, destacamos el hecho de que para dar estas estimaciones no necesitamos conocer x_3 , como ocurre si queremos calcular σ_2 .

Efectuando los cálculos oportunos obtenemos que

$$k_2 = \frac{195}{313}, \quad \delta_2^{(1)} = \frac{5}{439452}, \quad \delta_2^{(2)} = \frac{25}{146484}.$$

En consecuencia, igualando a cero el polinomio

$$o_2^{(1)}(t) = k_2 t^2 - t + \delta_2^{(1)}$$

obtenemos $\sigma_2^{(1)} \simeq 0.0000113779$.

Análogamente igualando a cero el polinomio

$$o_2^{(2)}(t) = k_2 t^2 - t + \delta_2^{(2)}$$

hallamos $\sigma_2^{(2)} \simeq 0.000170685$.

Concluimos por tanto que

$$\|x^* - x_2\| \leq \sigma_2^{(1)} \leq \sigma_2^{(2)} \leq r_1 - t_2 = \frac{1}{390} = 0.0025641. \quad \blacksquare$$

Capítulo 4

Nuevas sucesiones mayorizantes

4.1 Introducción

En este capítulo nos vamos a dedicar a encontrar funciones que nos proporcionen sucesiones mayorizantes para el método de Newton en condiciones distintas a las de Kantorovich.

Para ello avanzaremos en dos direcciones. En primer lugar intentaremos debilitar las condiciones de Kantorovich y, mediante un estudio geométrico del grado de convexidad logarítmico, encontraremos funciones que proporcionan sucesiones mayorizantes en condiciones menos restrictivas que las de Kantorovich.

En segundo lugar, estudiaremos lo que ocurre en condiciones más fuertes que las de Kantorovich. La mayoría de los autores [66], [82], [92], [107], para resolver una ecuación $F(x) = 0$ por el método de Newton, utilizan las denominadas condiciones de Kantorovich o condiciones similares a éstas. En este tipo de condiciones se suelen dar acotaciones sobre la derivada segunda del operador F , o bien una condición tipo Lipschitz sobre la derivada primera de F . Nosotros consideraremos condiciones en principio más restrictivas que las anteriores, como condiciones tipo Lipschitz sobre la derivada segunda de F . Encontraremos, utilizando el grado de convexidad logarítmico, la función que nos proporciona una sucesión mayorizante óptima en este caso, y que nos permitirá dar un resultado de convergencia y análisis del error para la sucesión de Newton.

Veremos además que hay situaciones en las que las condiciones de Kantorovich no garantizan la convergencia del método de Newton y sí que lo hacen estas nuevas condiciones.

4.2 Sucesiones mayorizantes y convexidad

Partimos como de costumbre de una ecuación $F(x) = 0$, con F un operador entre dos espacios de Banach X e Y , dos veces diferenciable Fréchet en un conjunto convexo $\Omega_0 \subseteq X$. Si F cumple las condiciones de Kantorovich, esto es, existe un punto $x_0 \in \Omega_0$ tal que está definido el operador inverso $\Gamma_0 = \Gamma(x_0) = [F'(x_0)]^{-1}$, además $\|\Gamma_0 F''(x)\| \leq k$ si $x \in \Omega_0$ y $ak \leq 1/2$, siendo $a = \|\Gamma_0 F(x_0)\|$, entonces sabemos que la sucesión de Newton

$$x_{n+1} = x_n - \Gamma_n F(x_n), \quad \Gamma_n = [F'(x_n)]^{-1} \quad (2.1)$$

converge a x^* solución de $F(x) = 0$ y que el polinomio

$$p(t) = \frac{k}{2}t^2 - t + a. \quad (2.2)$$

proporciona la sucesión mayorizante óptima, en el sentido de que la estimación del error que nos da es la mejor entre todas las que se pueden conseguir mediante sucesiones mayorizantes.

En concreto, vimos (Teorema III.2.2) que si F estaba en las condiciones de Kantorovich, entonces

$$\|L_F(x)\| \leq L_p(t)$$

si $x \in [x_n, x_{n+1}]$ y t es su correspondiente punto del intervalo $[t_n, t_{n+1}]$. En esta situación, el Teorema III.2.1 nos asegura que la sucesión $\{t_n\}$ obtenida aplicando el método de Newton a la ecuación $p(t) = 0$, con $t_0 = 0$ es una sucesión mayorizante de $\{x_n\}$.

Si analizamos el grado de convexidad logarítmico de este polinomio vemos que es de la forma de la Figura 4.1.

Sigamos ahora el proceso inverso, es decir, partamos de la función $w(t) = L_p(t)$. Hemos visto que, si $\|L_F(x)\| \leq w(t)$, entonces la función que tiene como grado de convexidad logarítmico $w(t)$ nos proporciona una sucesión mayorizante de (2.1). Esta función es precisamente el polinomio p definido por (2.2) y las condiciones que debe cumplir F para que exista convergencia son las de Kantorovich. ¿Qué ocurre si consideramos como $w(t)$ otra función distinta? ¿Obtendremos convergencia en condiciones distintas a las de Kantorovich?

Sea r_1 la menor raíz de la ecuación $p(t) = 0$ con p dado por (2.2). Consideramos ahora otra función g de manera que $g(r_1) = 0$ y $g(t) \neq 0$ para $t \in [0, r_1)$. Si la gráfica

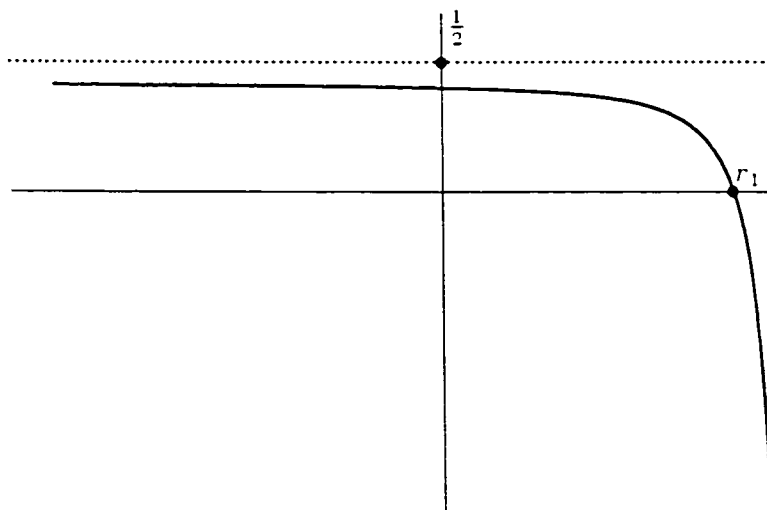


Figura 4.1: Gráfica de $L_p(t)$, con p dado por (2.2)

del grado de convexidad logarítmico de esta función está por encima de la de L_p en el intervalo $[0, r_1]$, es decir, $L_p(t) \leq L_g(t)$ si $t \in [0, r_1]$, entonces el Teorema I.3.3 nos asegura que $t_n \leq s_n$ para $n \geq 0$, donde $\{s_n\}$ es la sucesión obtenida aplicando el método de Newton a la ecuación $g(t) = 0$.

Así, la sucesión $\{s_n\}$ también es mayorizante de $\{x_n\}$, pero, por supuesto, nos proporciona una peor estimación del error que $\{t_n\}$. No obstante, cabe preguntarse si la sucesión mayorizante dada por g permite debilitar las condiciones de Kantorovich.

Para responder a estas cuestiones comenzamos analizando desde este nuevo punto de vista una situación familiar: Los resultados obtenidos en el Capítulo 2 utilizando el Teorema del Punto Fijo. Allí, vimos que si el grado de convexidad logarítmico de F estaba acotado por una constante menor que uno, entonces podíamos probar la convergencia de la sucesión (2.1) empezando en puntos de salida donde no se garantizaban las condiciones de Kantorovich (ver Ejemplos II.4.2 y II.4.3).

Desde nuestro punto de vista geométrico, y empleando sucesiones mayorizantes, tenemos que buscar funciones g cuyo grado de convexidad logarítmico sea constante. Observar que si exigimos que estas funciones tengan como raíz r_1 , del hecho de que $L_g(t)$ sea constante en el intervalo $[0, r_1]$ se deduce que r_1 debe ser también raíz de g' . Como g tiene raíces múltiples, es conocido [103] que el método de Newton tiene convergencia lineal, por tanto, las estimaciones del error obtenidas en este caso son sensiblemente peores que las obtenidas por el polinomio (2.2).

Veamos ahora qué funciones tienen grado de convexidad logarítmico constante.

Sea g una función tal que $L_{g'}(t) = M$. Sabemos que $L_g(t) = h'(t)$, siendo h la función dada por

$$h(t) = t - \frac{g(t)}{g'(t)}.$$

Integrando la expresión

$$\left(t - \frac{g(t)}{g'(t)}\right)' = M$$

se obtiene

$$t - \frac{g(t)}{g'(t)} = Mt + C.$$

Si suponemos que r_1 es una raíz de g , encontramos el valor que debe tomar la constante C , que es $C = r_1(1 - M)$.

De aquí se sigue que

$$\frac{g'(t)}{g(t)} = \frac{1}{(M-1)(r_1-t)}.$$

y volviendo a integrar,

$$\log g(t) = \log C_1(r_1 - t)^{\frac{1}{1-M}}.$$

Teniendo en cuenta que el grado de convexidad logarítmico de una función no cambia si multiplicamos a ésta por una constante, se obtiene la función

$$g(t) = (r_1 - t)^{\frac{1}{1-M}}.$$

Observar que estas funciones están definidas en el intervalo $[0, r_1]$ y además son decrecientes y convexas en dicho intervalo. En estas condiciones la convergencia del método de Newton a la raíz r_1 está garantizada empezando en un punto cualquiera del intervalo $[0, r_1]$.

A continuación vamos a estudiar las condiciones que una función de la forma anterior debe cumplir para que nos proporcione una sucesión mayorizante de (2.1). Para ello nos serán de utilidad los Teoremas I.4.2 y III.2.1 que nos dicen cuando una función f mayoriza a un operador F , esto es, la sucesión real obtenida aplicando el método de Newton a la ecuación $f(t) = 0$ es mayorizante de (2.1), la sucesión de Newton para resolver $F(x) = 0$. Consideraremos funciones de la forma

$$f(t) = (r - t)^{\frac{1}{1-M}}. \quad (2.3)$$

Supongamos que $\|L_F(x)\| \leq M < 1$ si $x \in \overline{B(x_0, R_0)}$. En esta situación, las funciones de la forma (2.3) nos aseguran de antemano la segunda condición del Teorema III.2.1.

Vamos a dar condiciones sobre r para que se cumpla la primera condición de dicho teorema. Tiene que ocurrir que

$$a \leq -\frac{f(t_0)}{f'(t_0)} = (1-M) \frac{r^{\frac{1}{1-M}}}{r^{\frac{1}{1-M}-1}} = r(1-M),$$

y por tanto, $r \geq \frac{a}{1-M}$, donde $a = \|\Gamma_0 F(x_0)\|$.

Por último, como también debe cumplirse que la longitud del intervalo donde está definida la sucesión real, que en este caso es el intervalo $[0, r]$, no exceda del radio de la bola donde el grado de convexidad logarítmico L_F está acotado, podemos tomar r tal que

$$\frac{a}{1-M} \leq r \leq R_0.$$

Para que este intervalo esté definido, tiene que ocurrir que

$$\frac{a}{1-M} \leq R_0,$$

que es la misma condición que se obtenía para probar la convergencia en el Corolario II.3.3.

En definitiva, la sucesión

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)} \quad (2.4)$$

construida aplicando el método de Newton a la función f dada por (2.3) con

$$\frac{a}{1-M} \leq r \leq R_0 \quad (2.5)$$

y comenzando en $t_0 = 0$ nos proporciona una sucesión mayorizante de la sucesión $\{x_n\}$ dada por (2.1). Con esto hemos probado el siguiente teorema.

Teorema 2.1. *Sea la ecuación $F(x) = 0$, con F un operador entre espacios de Banach y sea $\overline{B(x_0, R_0)}$ una bola en la que $\|L_F(x)\| \leq M < 1$. Denotamos $a = \|[F'(x_0)]^{-1}F(x_0)\|$. Si $\frac{a}{1-M} \leq R_0$, entonces la sucesión de Newton (2.4) para resolver la ecuación escalar $f(t) = 0$, con f dada por (2.3) y tomando como punto de salida $t_0 = 0$, es una sucesión mayorizante de $\{x_n\}$, la sucesión de Newton para resolver $F(x) = 0$ y empezando en x_0 .*

Además, en estas condiciones, $\{t_n\}$ es convergente, luego $\{x_n\}$ también lo es.

Que la situación del resultado anterior es más general que la de Kantorovich es una comprobación inmediata, sin más que considerar por ejemplo la función real

$$F(t) = (1-t)^4.$$

Esta función está en las condiciones del teorema anterior ($L_f(t) = 3/4$ para todo t) y sin embargo no cumple las condiciones de Kantorovich en el intervalo $[0, 1]$. En efecto

$$a = -\frac{F(0)}{F'(0)} = \frac{1}{4},$$

$$k = \sup_{t \in [0,1]} \left\{ -\frac{F''(t)}{F'(0)} \right\} = 3.$$

y $ak = 3/4 > 1/2$.

Este ejemplo nos conduce de forma inmediata al sistema de ecuaciones de la forma $F(x, y) = ((1-x)^4, (1-y)^4)$, para el cual se tiene que $\|L_F(x, y)\| = 3/4 < 1$ para todo $(x, y) \in \mathbb{R}^2$ y, sin embargo, si tomamos $(x_0, y_0) = (0, 0)$, tenemos

$$a = \|\Gamma_0 F(x_0, y_0)\| = \frac{1}{4},$$

$$k \geq \|\Gamma_0 F''(x_0, y_0)\| = 3$$

y, por tanto, $ak \geq 3/4 > 1/2$.

A continuación vamos a analizar desde este nuevo punto de vista el Ejemplo II.4.3. Como paso previo, haremos la siguiente observación.

Observación. Por el Teorema II.3.2 sabemos que, en las condiciones del teorema anterior, la sucesión $\{x_n\}$ definida por (2.1) es convergente a la solución de $F(x) = 0$, comenzando en cualquier punto de la bola $\overline{B(x_0, R_0)}$.

En consecuencia, si $\tilde{x}_0 \in \overline{B(x_0, R_0)}$ la sucesión $\{\tilde{x}_n\}$ definida por (2.1) es convergente a x^* , solución de $F(x) = 0$.

En este caso, para poder encontrar una función de la forma (2.3) que nos proporcione una sucesión mayorizante de $\{\tilde{x}_n\}$ tendremos que exigir que

$$\|[F'(\tilde{x}_0)]^{-1}F(\tilde{x}_0)\| \leq -\frac{f(t_0)}{f'(t_0)} = r(1 - M),$$

y por tanto

$$r \geq \frac{\|[F'(\tilde{x}_0)]^{-1}F(\tilde{x}_0)\|}{1 - M}. \quad (2.6)$$

Ejemplo 2.2. Consideramos el operador $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ del Ejemplo II.4.3, que viene dado por $F(x, y) = (xy - 1, xy + x - 2y)$.

Como se vió en ese ejemplo, si tomamos

$$\Omega = \overline{B(x_0, 0.43)} = [0.7, 1.56] \times [0.7, 1.56],$$

con $x_0 = (1.13, 1.13)$ se tiene que

$$\sup_{(x,y) \in \Omega} \|L_F(x,y)\| \leq 0.7 = M.$$

Entonces la sucesión $\{x_n\}$ dada por (2.1) y empezando en $x_0 = (1.13, 1.13)$ es convergente a la solución de $F(x,y) = 0$. Podemos encontrar una función de la forma (2.3) que nos proporcione una sucesión mayorizante de $\{x_n\}$ si tomamos $r = 0.43$, el radio de la bola Ω . En este caso nos queda

$$f(t) = [0.43 - t]^{10/3}.$$

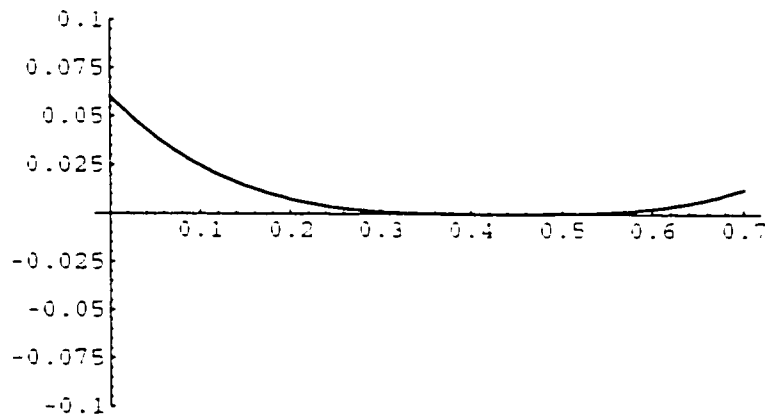


Figura 4.2: Gráfica de $f(t) = [0.43 - t]^{10/3}$

Otra función que también nos daría sucesiones mayorizantes es la función de la forma (2.3) con

$$r = \frac{\|[F'(x_0)]^{-1}F(x_0)\|}{1 - M} = 0.4083815.$$

y en general, las funciones de la forma

$$f(t) = [r - t]^{10/3} \text{ con } 0.4083815 \leq r \leq 0.43. \quad \blacksquare$$

Por último, comentar que al aplicar el método de Newton (2.1) a una función de la forma (2.3), nos queda un proceso bastante sencillo, ya que

$$-\frac{f(t)}{f'(t)} = (1 - M)(r - t).$$

y la sucesión que se obtiene es

$$t_{n+1} = t_n + (1 - M)(r - t_n). \quad (2.7)$$

Podemos dar ahora una estimación del error cometido aproximando r por t_n . Si llamamos $E_n = r - t_n$, sustituyendo (2.7) en E_{n+1} tenemos

$$E_{n+1} = r - t_{n+1} = r - t_n - (1 - M)(r - t_n) = M E_n.$$

Reiterando el proceso, se sigue que

$$E_{n+1} = M^{n+1} E_0 = M^{n+1} r.$$

Observar que cuanto menor sea r tendremos una mejor estimación del error. Por tanto, en (2.5) y (2.6) nos interesa tomar

$$r = \frac{\| [F'(x_0)]^{-1} F(x_0) \|}{1 - M} \quad \text{y} \quad r = \frac{\| [F'(\tilde{x}_0)]^{-1} F(\tilde{x}_0) \|}{1 - M}$$

respectivamente.

En el Teorema 2.1 estudiamos el caso de que el grado de convexidad logarítmico del operador F estuviese acotado por una constante menor que uno. Nos podemos plantear ahora el estudio en otras situaciones. Por supuesto, la situación ideal es encontrar una función w que verifique

$$\|L_F(x)\| = w(t) \quad \text{si} \quad \|x - x_0\| \leq t - t_0.$$

Repetiendo el proceso de integración visto con anterioridad, encontraríamos una función f cuyo grado de convexidad logarítmico sea la función w . Esta función f nos proporcionaría la sucesión mayorizante que nos da la mejor estimación del error. Pero esta situación es muy difícil de comprobar en la práctica, por lo que puede resultar interesante analizar otros casos.

Una vez estudiado el caso de que el grado de convexidad logarítmico esté acotado por una recta paralela al eje OX , podemos pensar en estudiar el caso de que el grado de convexidad logarítmico esté acotado por una recta que no sea paralela al eje OX . Evitaremos así el problema de tener raíces múltiples que teníamos cuando $L_f(x) = M$, con M constante. Además, para este caso, la expresión del método de Newton es muy sencilla, ya que no hay necesidad de efectuar divisiones. Buscaremos una función f cuyo grado de convexidad logarítmico sea de la forma

$$L_f(t) = \alpha(r - t), \quad \alpha, r > 0.$$

Teniendo en cuenta que $L_f(t)$ es la derivada de $t - f(t)/f'(t)$, integrando la igualdad anterior, y exigiendo que r sea una raíz de f , tenemos

$$\frac{f(t)}{f'(t)} = (r - t) \left(\frac{\alpha}{2}(r - t) - 1 \right).$$

de donde se sigue que

$$f(t) = \frac{r - t}{|t - r + (2/\alpha)|}.$$

Para evitar problemas de continuidad en el intervalo $[0, r]$, suponemos $r - (2/\alpha) < t$ para todo $t \in [0, r]$ o, lo que es lo mismo,

$$\alpha r < 2.$$

En este caso, denotando $\beta = (2/\alpha) - r$, nos quedan unas funciones de la forma

$$f(t) = \frac{r - t}{\beta + t}, \quad \beta, r > 0.$$

Con esto, tendríamos probado el siguiente resultado.

Teorema 2.3. *Sea F un operador entre dos espacios de Banach X e Y , dos veces diferenciable Fréchet en un conjunto convexo $\Omega_0 \subseteq X$. Sea x_0 un punto en el que existe $\Gamma_0 = [F'(x_0)]^{-1}$, y sea $a = \|\Gamma_0 F(x_0)\|$. Si*

$$\|L_F(x)\| \leq \alpha(r - t) \quad \text{para } \|x - x_0\| \leq t - t_0,$$

con $\alpha r < 2$ y $a \leq \frac{r}{2}(2 - \alpha r)$, entonces la sucesión de Newton (2.1) converge a x^* , solución de $F(x) = \bar{0}$. Además se tiene que

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n \quad \text{y} \quad \|x^* - x_n\| \leq r - t_n,$$

donde $\{t_n\}$ es la sucesión de Newton para resolver la ecuación $f(t) = 0$ con

$$f(t) = \frac{r - t}{\beta + t}, \quad \beta = \frac{2}{\alpha} - r.$$

Demostración. Por los comentarios previos al enunciado del teorema, sabemos que $\|L_F(x)\| \leq L_f(t)$ si $\|x - x_0\| \leq t - t_0$. Además

$$\|\Gamma_0 F(x_0)\| = a \leq -\frac{f(t_0)}{f'(t_0)} = \frac{r\beta}{r + \beta} = \frac{r}{2}(2 - \alpha r).$$

De la convergencia de la sucesión $\{t_n\}$, y sin más que aplicar el Teorema III.2.1, se sigue el resultado. ■

En lugar de considerar rectas, como en los Teoremas 2.1 y 2.2, podemos considerar otro tipo de funciones, obteniendo así, en función del grado de convexidad logarítmico, nuevas condiciones para probar la convergencia de la sucesión de Newton (2.1). No obstante, estos resultados son interesantes desde el punto de vista teórico pero, en la práctica, es difícil encontrar acotaciones cómodas para el operador L_F .

4.3 Sucesiones mayorizantes bajo condiciones de Kantorovich fuertes

En el Teorema III.2.2 encontramos, por medio del grado de convexidad logarítmico, una sucesión mayorizante que nos permite resolver la ecuación $F(x) = 0$ en el caso de que la derivada segunda de F esté acotada. Esta sucesión se construía aplicando el método de Newton a la ecuación $p(t) = 0$, con p un polinomio de segundo grado. A continuación vamos a generalizar este resultado al caso de que la derivada segunda verifique una condición tipo Lipschitz

$$\|\Gamma_0[F''(x) - F''(x_0)]\| \leq k\|x - x_0\|. \quad (3.1)$$

Este tipo de condiciones nos permitirá garantizar la convergencia de la sucesión de Newton en situaciones en las que las condiciones de Kantorovich no lo hace, como en el siguiente ejemplo [62].

Ejemplo 3.1. Sean $X = Y = [-1, 1]$, $f : X \rightarrow Y$ el polinomio

$$f(x) = \frac{1}{6}x^3 + \frac{1}{6}x^2 - \frac{5}{6}x + \frac{1}{3}$$

y $x_0 = 0$.

Sea M la menor cota que verifica

$$\|[f'(x_0)]^{-1}[f'(x) - f'(y)]\| \leq M\|x - y\|, \quad x, y \in X.$$

En este caso $M = 8/5$. Denotamos $a = \|[f'(x_0)]^{-1}f(x_0)\| = 2/5$.

Observar que $aM = 16/25 > 1/2$, luego no se cumplen las condiciones de Kantorovich y no se puede garantizar la convergencia de la sucesión de Newton empezando en x_0 . Más adelante probaremos la convergencia usando nuevos argumentos. ■

Construimos la sucesión de Newton

$$x_{n+1} = x_n - \Gamma_n F(x_n), \quad \Gamma_n = [F'(x_n)]^{-1}. \quad (3.2)$$

El siguiente resultado estudia la convergencia de la sucesión bajo la condición (3.1).

Teorema 3.2. *Sea la ecuación $F(x) = 0$. con F un operador entre dos espacios de Banach X e Y dos veces diferenciable Fréchet en un conjunto convexo $\Omega_0 \subseteq X$. Sea x_0 un punto de Ω_0 en el que existe el operador $\Gamma_0 = [F'(x_0)]^{-1}$. Supongamos $\|\Gamma_0 F(x_0)\| \leq a$ y $\|\Gamma_0 F''(x_0)\| \leq b$ y que F satisface la condición (3.1) en Ω_0 . Sea p el polinomio dado por*

$$p(t) = a - t + \frac{b}{2}t^2 + \frac{k}{6}t^3. \quad (3.3)$$

Si se verifica

$$p(m) \leq 0. \quad (3.4)$$

con

$$m = \frac{2}{b + \sqrt{b^2 + 2k}}. \quad (3.5)$$

entonces el polinomio p tiene dos raíces positivas, r_1 y r_2 ($r_1 \leq r_2$). Si además $B(x_0, m) \subseteq \Omega_0$, la sucesión $\{x_n\}$ definida por (3.2) es convergente a x^* , solución de $F(x) = 0$ en $\overline{B(x_0, r_1)}$, y única solución en $B(x_0, r_2)$ si $r_1 < r_2$ o en $\overline{B(x_0, r_1)}$ si $r_1 = r_2$. Por último, se tiene también que

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n,$$

siendo $\{t_n\}$ la sucesión obtenida aplicando el método de Newton al polinomio $p(t)$, a partir de $t_0 = 0$.

Antes de demostrar el teorema, veamos un lema y algunas observaciones sobre el polinomio (3.3).

Lema 3.3. *Supongamos que F es un operador en las condiciones del teorema anterior. Si $x \in B(x_0, m)$, con m dado por (3.5), el operador grado de convexidad logarítmico está definido y además*

$$\|L_F(x)\| \leq \frac{(b + k\|x - x_0\|)\|\Gamma_0 F(x)\|}{\left[1 - \frac{k}{2}\|x - x_0\|^2 - b\|x - x_0\|\right]^2}.$$

Demostración. Observar que

$$\int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)] dy = \Gamma_0 F'(x) - I - \Gamma_0 F''(x_0)(x - x_0).$$

Por tanto, teniendo en cuenta (3.1).

$$\begin{aligned} \|\Gamma_0 F'(x) - I - \Gamma_0 F''(x_0)(x - x_0)\| &= \left\| \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)] dy \right\| \\ &= \left\| \int_0^1 \Gamma_0 [F''(x_0 + t(x - x_0)) - F''(x_0)] (x - x_0) dt \right\| \\ &\leq k \|x - x_0\|^2 \int_0^1 t dt = \frac{k \|x - x_0\|^2}{2}. \end{aligned}$$

De aquí se sigue

$$\|\Gamma_0 F'(x) - I\| \leq \frac{k}{2} \|x - x_0\|^2 + b \|x - x_0\|.$$

Analizando el signo del polinomio $\frac{k}{2}t^2 + bt - 1$, deducimos una condición suficiente para que el segundo miembro de la desigualdad anterior sea menor que uno.

$$\|x - x_0\| < \frac{2}{b + \sqrt{b^2 + 2k}}.$$

En este caso, por el Lema de Banach, se sigue que existe el operador $[\Gamma_0 F'(x)]^{-1}$ y además

$$\|[\Gamma_0 F'(x)]^{-1}\| \leq \frac{1}{1 - \left(\frac{k}{2}\|x - x_0\|^2 + b\|x - x_0\|\right)}.$$

Con esto, teniendo en cuenta la expresión del operador L_F y que de (3.1) se deduce que

$$\|\Gamma_0 F''(x)\| \leq b + k\|x - x_0\|,$$

se prueba el resultado. ■

Observación. Vamos a analizar un poco el polinomio dado por (3.3). Este polinomio es de la forma de la Figura 4.3 con un máximo en

$$t = M = -\frac{b + \sqrt{b^2 + 2k}}{k} < 0$$

y un mínimo en $t = m > 0$, con m definido por (3.5). Para que p tenga raíces positivas tiene que ocurrir que $p(m) \leq 0$, que no es otra cosa que la condición (3.4).

En este caso, si denotamos r_1 a la menor de las raíces positivas de $p(t)$, se sigue que $r_1 \leq m$ (la igualdad se obtiene si en (3.4) también se da la igualdad).

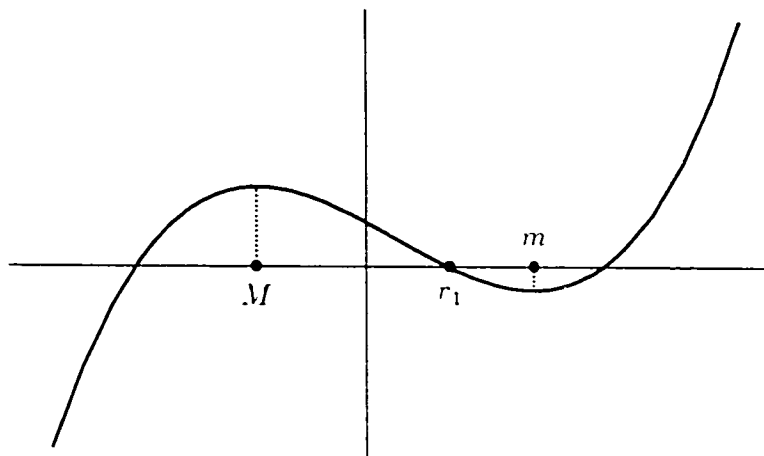


Figura 4.3: Ejemplo de polinomio de la forma (3.3)

El polinomio p tiene un punto de inflexión en $t = -b/k < 0$. Por lo tanto, en el intervalo $[0, r_1]$, p es decreciente y convexo. Es conocido que en esta situación la sucesión $\{t_n\}$ definida por

$$t_{n+1} = t_n - \frac{p(t_n)}{p'(t_n)}, \quad \text{con } t_0 = 0,$$

y p dado por (3.3), es creciente y convergente a r_1 .

Por último, comentar que la condición (3.4) puede reemplazarse por otra condición más sencilla de probar. De hecho, para que p tenga alguna raíz positiva, basta con que $p(r) \leq 0$ para algún $r > 0$. Un valor muy cómodo con el que probar es $t = 1/b$. En concreto, podemos sustituir (3.4) por la condición

$$a - \frac{1}{2b} + \frac{k}{6b^3} \leq 0. \quad (3.6)$$

más sencilla de probar, aunque más fuerte (notemos que (3.6) \Rightarrow (3.4)).

Z. Huang [62] dio un par de condiciones equivalentes a (3.4) y más cómodas de probar, ya que dependen únicamente de los coeficientes del polinomio p . Estas condiciones son

$$6ab^3 + 9a^2k^2 + 18abk \leq 3b^2 + 8k$$

y

$$3ak^2 + 3bk + b^3 \leq [b^2 + 2k]^{3/2}. \quad (3.7)$$

Demostración del Teorema 3.2. Seguiremos un razonamiento análogo al de la demostración del Teorema III.2.2. En nuestras condiciones tenemos que x_1 está definido. Además,

$$\|x_1 - x_0\| = \|\Gamma_0 F(x_0)\| \leq a = t_1 - t_0.$$

Observar que, además, $x_1 \in B(x_0, m)$, con m definido en (3.5). En efecto, teniendo en cuenta la observación anterior,

$$\|x_1 - x_0\| \leq a = t_1 < r_1 \leq m,$$

y por el Lema 3.3 existe $L_F(x_1)$ y, por tanto, podemos definir x_2 . Además,

$$x_2 - x_1 = \int_{x_0}^{x_1} L_F(x) dx.$$

Si $x \in [x_0, x_1]$ podemos escribir $x = x_0 + s(x_1 - x_0)$, con $0 \leq s \leq 1$. Entonces, por el Teorema de Taylor

$$\begin{aligned} \Gamma_0 F(x) &= \Gamma_0 F(x_0) + (x - x_0) + \frac{1}{2} \Gamma_0 F''(x_0)(x - x_0)^2 \\ &\quad + \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)](x - y) dy. \end{aligned}$$

Como $x - x_0 = s(x_1 - x_0) = -s\Gamma_0 F(x_0)$,

$$\begin{aligned} \|\Gamma_0 F(x)\| &\leq (1 - s)\|\Gamma_0 F(x_0)\| + \frac{1}{2}\|\Gamma_0 F''(x_0)\|\|(x - x_0)^2\| \\ &\quad + \left\| \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)](x - y) dy \right\|. \end{aligned}$$

Ahora bien, teniendo en cuenta (3.1),

$$\begin{aligned} &\left\| \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)](x - y) dy \right\| \\ &= \left\| \int_0^1 \Gamma_0 [F''(x_0 + s(x - x_0)) - F''(x_0)](x - x_0)^2(1 - s) ds \right\| \\ &\leq k\|x - x_0\|^3 \int_0^1 s(1 - s) ds = \frac{k}{6}\|x - x_0\|^3. \end{aligned}$$

En consecuencia, como $\|x - x_0\| = s\|\Gamma_0 F(x_0)\| \leq sa$,

$$\|\Gamma_0 F(x)\| \leq (1 - s)a + \frac{1}{2}bs^2a^2 + \frac{k}{6}a^3s^3.$$

Si hacemos $t = sa$, entonces

$$\|\Gamma_0 F(x)\| \leq a - t + \frac{1}{2}bt^2 + \frac{k}{6}t^3$$

y, por tanto, del Lema 3.3 se sigue que

$$\|L_F(x)\| \leq \frac{(b+kt)(a-t+\frac{1}{2}bt^2+\frac{k}{6}t^3)}{[1-\frac{k}{2}t^2-bt]^2} = L_p(t).$$

Así,

$$\|x_2 - x_1\| = \left\| \int_{x_0}^{x_1} L_F(x) dx \right\| \leq \int_{t_0}^{t_1} L_p(t) dt = t_2 - t_1.$$

Siguiendo la demostración del Teorema III.2.2 podemos sustituir en el razonamiento anterior x_2 por x_1 y en general, x_{n+1} por x_n .

Para demostrar la unicidad, se puede emplear la misma técnica que en la demostración del Teorema III.2.3. Con esto, queda probado el resultado. ■

El análisis del método de Newton aplicado a polinomios de segundo grado realizado por Ostrowski [82] es utilizado para estimar el error de la sucesión (3.2) cuando el operador F cumple las condiciones de Kantorovich. Recordar que en este caso se obtenía una sucesión mayorizante de (3.2) aplicando el método de Newton a la ecuación $p(t) = 0$, siendo p un polinomio de segundo grado.

En las condiciones del Teorema 3.2 la sucesión mayorizante de (3.2) se obtiene a partir de un polinomio de tercer grado. Estudiaremos a continuación el método de Newton aplicado a $p(t) = 0$, siendo p un polinomio de tercer grado, en concreto, de la forma (3.3).

Teorema 3.4. *Sea p un polinomio de la forma (3.3), con dos raíces positivas y una negativa. Entonces la sucesión de Newton $\{t_n\}$ construida para resolver la ecuación $p(t) = 0$, empezando en $t_0 = 0$, es convergente a r_1 , la menor raíz positiva de p . Además*

$$(r_2 - r_1) \frac{\alpha^{2^n}}{r - \alpha^{2^n}} \leq r_1 - t_n \leq (r_2 - r_1) \frac{\theta^{2^n}}{R - \theta^{2^n}},$$

donde

$$r = \frac{r_0 - r_2}{r_0 - r_1}, \quad R = 1 - \frac{r_2 - r_1}{r_0 + r_1},$$

$$\alpha = r \frac{r_1}{r_2} < 1, \quad \theta = R \frac{r_1}{r_2} < 1.$$

$-r_0$ es la raíz negativa de p y r_2 la otra raíz positiva.

Demostración. La convergencia de la sucesión $\{t_n\}$ es conocida ya que el polinomio (3.3) es una función decreciente y convexa en el intervalo $[0, r_1]$.

Para ver la estimación del error, podemos escribir el polinomio (3.3) de la siguiente manera

$$p(t) = \frac{k}{6}(r_1 - t)(r_2 - t)(r_0 + t). \quad 0 < r_1 \leq r_2. \quad 0 < r_0.$$

Entonces

$$p'(t) = -\frac{k}{6}[(r_2 - t)(r_0 + t) + (r_1 - t)(r_0 + t) - (r_1 - t)(r_2 - t)].$$

Si llamamos $a_n = r_1 - t_n$, $b_n = r_2 - t_n$ y $c_n = r_0 + t_n$, tenemos

$$\frac{p(t_n)}{p'(t_n)} = -\frac{a_n b_n c_n}{b_n c_n + a_n c_n - a_n b_n}.$$

Como consecuencia

$$\begin{aligned} a_{n+1} &= r_1 - t_{n+1} = r_1 - t_n + \frac{p(t_n)}{p'(t_n)} \\ &= a_n - \frac{a_n b_n c_n}{b_n c_n + a_n c_n - a_n b_n} = \frac{a_n^2 (c_n - b_n)}{b_n c_n + a_n c_n - a_n b_n}. \end{aligned}$$

Análogamente

$$b_{n+1} = \frac{b_n^2 (c_n - a_n)}{b_n c_n + a_n c_n - a_n b_n}.$$

Por lo tanto,

$$\frac{a_{n+1}}{b_{n+1}} = \left(\frac{a_n}{b_n}\right)^2 \frac{c_n - b_n}{c_n - a_n}.$$

Observar que

$$\frac{c_n - b_n}{c_n - a_n} = \frac{(r_0 - r_2) + 2t_n}{(r_0 - r_1) + 2t_n}$$

y que

$$r = \frac{r_0 - r_2}{r_0 - r_1} \leq \frac{(r_0 - r_2) + 2t_n}{(r_0 - r_1) + 2t_n} \leq \frac{r_0 - r_2 + 2r_1}{r_0 + r_1} = R. \quad (3.8)$$

De aquí se sigue que

$$\begin{aligned} \frac{a_{n+1}}{b_{n+1}} &\leq R \left(\frac{a_n}{b_n}\right)^2 \leq R^3 \left(\frac{a_{n-1}}{b_{n-1}}\right)^4 \leq \dots \\ &\dots \leq R^{2^{n+1}-1} \left(\frac{a_0}{b_0}\right)^{2^{n+1}} = \frac{1}{R} \left(R \frac{r_1}{r_2}\right)^{2^{n+1}}. \end{aligned}$$

Como $\theta = R \frac{r_1}{r_2}$, y teniendo en cuenta que $b_n = r_2 - r_1 + a_n$, se deduce que

$$a_n \left(1 - \frac{\theta^{2^n}}{R}\right) \leq \frac{r_2 - r_1}{R} \theta^{2^n}.$$

y de aquí

$$r_1 - t_n \leq (r_2 - r_1) \frac{\theta^{2^n}}{R - \theta^{2^n}}.$$

Para ver la estimación inferior, notar que, a partir de (3.8),

$$\begin{aligned} \frac{a_{n+1}}{b_{n+1}} &\geq r \left(\frac{a_n}{b_n} \right)^2 \geq r^3 \left(\frac{a_{n-1}}{b_{n-1}} \right)^4 \geq \dots \\ \dots &\geq r^{2^{n+1}-1} \left(\frac{a_0}{b_0} \right)^{2^{n+1}} = \frac{1}{r} \left(r \frac{r_1}{r_2} \right)^{2^{n+1}} = \frac{1}{r} \alpha^{2^{n+1}}. \end{aligned}$$

Por tanto,

$$a_n \left(1 - \frac{\alpha^{2^n}}{r - \alpha^{2^n}} \right) \geq \frac{r_2 - r_1}{r} \alpha^{2^n}.$$

de donde se deduce

$$(r_2 - r_1) \frac{\alpha^{2^n}}{r - \alpha^{2^n}} \leq r_1 - t_n. \quad \blacksquare$$

Observaciones. • Para un polinomio de la forma (3.3) se puede demostrar que $r_1 \leq r_2 \leq r_0$. En efecto, con la misma notación que en la demostración del teorema anterior, se tiene que

$$0 \leq \frac{a_1}{b_1} = \frac{a_0^2 r_0 - r_2}{b_0^2 r_0 - r_1}.$$

Por otra parte, analizando las raíces de p' , esto es, m (mínimo relativo de p) y M (máximo relativo de p), tenemos que

$$r_1 \leq m < -M < r_0.$$

De las desigualdades anteriores se deduce la relación entre las raíces del polinomio (3.3).

• A partir de dicha relación se prueba sin dificultad que r y R definidos en el enunciado del Teorema 3.4 son ambos positivos y menores que uno.

• A la vista de la relación anterior y de la estimación del error dada en el Teorema 3.4, en la que aparece el cociente r_1/r_2 , podemos pensar en comparar las expresiones de a_{n+1} y c_{n+1} en lugar de a_{n+1} y b_{n+1} . De esta forma la expresión del error vendría dada en términos del cociente $r_1/r_0 < r_1/r_2$.

Procediendo como en la demostración del teorema anterior, se sigue que

$$\frac{a_{n+1}}{c_{n+1}} \leq \frac{a_n^2 r_0 + r_1}{c_n^2 r_2 - r_1}.$$

Desafortunadamente, la cantidad $\frac{r_0 + r_1}{r_2 - r_1}$ por la que estamos acotando es mayor que uno, y si seguimos el razonamiento de la demostración del Teorema 3.4 obtenemos una estimación del error que depende de la cantidad

$$\frac{r_0 + r_1 r_1}{r_2 - r_1 r_0}$$

que no tiene por qué ser menor que uno, como prueba el siguiente ejemplo. Consideramos el polinomio de la forma (3.3) dado por

$$p(t) = \frac{1}{6}t^3 + \frac{1}{6}t^3 - t + \frac{2}{3}.$$

En este caso tenemos $r_1 = 1$, $r_2 = -1 + \sqrt{5}$ y $r_0 = 1 + \sqrt{5}$ y

$$\frac{r_0 + r_1 r_1}{r_2 - r_1 r_0} = \frac{2 + \sqrt{5}}{-2 + \sqrt{5}} \frac{1}{1 + \sqrt{5}} = \frac{11 + 5\sqrt{5}}{4} > 1.$$

Este ejemplo nos hace abandonar la idea de obtener una expresión del error en términos del cociente r_1/r_0 .

A continuación veremos un ejemplo en el que podemos aplicar el Teorema 3.2 y en el que no se cumplen las condiciones de Kantorovich.

Ejemplo 3.5. Volviendo a la situación planteada en el Ejemplo 3.1, y manteniendo la misma notación empleada en esta sección, tenemos que $a = 2/5$, $b = 2/5$ y $k = 6/5$. Entonces

$$3ak^2 + 3bk + b^3 = 3\frac{36}{25}\frac{2}{5} + 3\frac{6}{5}\frac{2}{5} + \frac{8}{125} = \frac{404}{125}.$$

$$(b^2 + 2k)^{3/2} = \left(\frac{4}{25} + 2\frac{6}{5}\right)^{3/2} = \left(\frac{8}{5}\right)^3 = \frac{512}{125}.$$

Por tanto, se cumple la condición (3.7) y en consecuencia, se puede aplicar el Teorema 3.2 para probar la convergencia de la sucesión de Newton.

Sin embargo, como se puso de manifiesto en el Ejemplo 3.1, las condiciones de Kantorovich no bastaban para garantizar la convergencia de dicha sucesión. ■

En algunos casos, la convergencia de una sucesión se puede establecer indistintamente por el Teorema de Kantorovich o por el Teorema 3.2. Entonces, podemos comparar las estimaciones del error que se obtienen en cada caso. Así,

supongamos que un operador F cumple las condiciones del Teorema 3.2 y también las del Teorema de Kantorovich, esto es.

$$\| [F'(x_0)]^{-1}(F'(x) - F'(y)) \| \leq c \|x - y\|, \quad x, y \in \Omega_0. \quad (3.9)$$

con $ac \leq \frac{1}{2}$. En estas condiciones, las sucesiones mayorizantes se obtienen aplicando el método de Newton a los polinomios $p(t)$ dado por (3.3) y

$$q(t) = a - t + \frac{c}{2}t^2, \quad (3.10)$$

respectivamente.

Denotamos \hat{s}_1 y \hat{s}_2 a las raíces de q . $\{s_n\}$ a la sucesión de Newton para resolver $q(t) = 0$ con $s_0 = 0$. Son conocidas (Teorema III.2.4) las siguientes expresiones del error para $\{s_n\}$:

$$\hat{s}_1 - s_n = (\hat{s}_2 - \hat{s}_1) \frac{\lambda^{2^n}}{1 - \lambda^{2^n}}, \quad \lambda = \frac{\hat{s}_1}{\hat{s}_2} < 1. \quad (3.11)$$

El lema siguiente nos da una condición suficiente para obtener mejores estimaciones del error a partir del polinomio de tercer grado (3.3) que a partir del polinomio de segundo grado (3.10).

Lema 3.6. Sean p y q los polinomios definidos en (3.3) y (3.10). Si

$$k(1 + \sqrt{1 - 2ac}) < 3c(c - b), \quad (3.12)$$

entonces $r_1 - t_n < \hat{s}_1 - s_n$, $n \geq 0$.

Demostración. Notar que

$$p(\hat{s}_2) = q(\hat{s}_2) + \frac{\hat{s}_2^2}{2} \left(\frac{k}{3}\hat{s}_2 + b - c \right) < 0.$$

Además

$$p(\hat{s}_1) = q(\hat{s}_1) + \frac{\hat{s}_1^2}{2} \left(\frac{k}{3}\hat{s}_1 + b - c \right) < \frac{\hat{s}_1^2}{2} \left(\frac{k}{3}\hat{s}_2 + b - c \right) < 0.$$

Por tanto, se tiene (ver Figura 4.4), $r_1 < \hat{s}_1 < \hat{s}_2 < r_2$ y $\frac{r_1}{r_2} < \frac{\hat{s}_1}{\hat{s}_2}$.

A partir del Teorema 3.4 y de (3.11), se sigue

$$r_1 - t_n \leq (r_2 - r_1) \frac{\theta^{2^n}}{R - \theta^{2^n}} < (\hat{s}_2 - \hat{s}_1) \frac{\lambda^{2^n}}{1 - \lambda^{2^n}} = \hat{s}_1 - s_n, \quad n \geq 0. \quad (3.13)$$

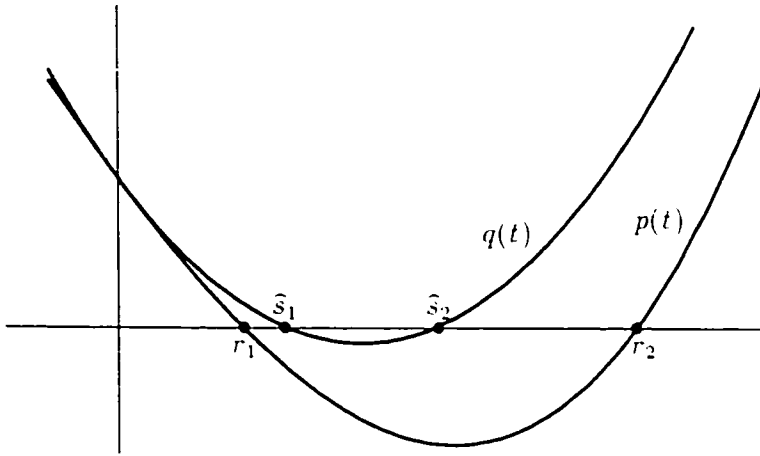


Figura 4.4: Polinomios definidos en (3.3) y (3.10)

En efecto, para $n = 0$ se tiene $r_1 < \hat{s}_1$. Supongamos que (3.13) se cumple para $k \leq n$. Entonces,

$$\begin{aligned} (r_2 - r_1) \frac{\theta^{2^{n+1}}}{R - \theta^{2^{n+1}}} &= (r_2 - r_1) \frac{\theta^{2^n}}{R - \theta^{2^n}} \frac{\theta^{2^n} (R - \theta^{2^n})}{R - \theta^{2^{n+1}}} \\ &< (\hat{s}_2 - \hat{s}_1) \frac{\lambda^{2^n}}{1 - \lambda^{2^n}} \frac{\theta^{2^n} (R - \theta^{2^n})}{R - \theta^{2^{n+1}}} < (\hat{s}_2 - \hat{s}_1) \frac{\lambda^{2^n}}{1 - \lambda^{2^n}} \frac{\lambda^{2^n}}{1 + \lambda^{2^n}} \\ &= (\hat{s}_2 - \hat{s}_1) \frac{\lambda^{2^{n+1}}}{1 - \lambda^{2^{n+1}}} = \hat{s}_1 - s_n. \quad \blacksquare \end{aligned}$$

Ejemplo 3.7. Consideramos los espacios de Banach $X = [0, 1] \times [0, 1]$, $Y = \mathbb{R}^2$ y el operador $F : X \rightarrow Y$ dado por

$$F(x, y) = \left(\frac{x^3}{24} + \frac{y^2}{4} - x + \frac{1}{3}, \frac{y^3}{8} + \frac{3x^2}{4} - 3y + 1 \right). \quad (3.14)$$

Para resolver la ecuación $F(x, y) = 0$ por el método de Newton partimos de $(x_0, y_0) = (0, 0)$. Entonces tenemos

$$\begin{aligned} F'(x, y) &= \begin{pmatrix} \frac{x^2}{8} - 1 & \frac{y}{2} \\ \frac{3x}{2} & \frac{3y^2}{8} - 3 \end{pmatrix}, \\ \Delta = \det[F'(x, y)] &= \frac{3x^2 y^2}{64} - \frac{3x^2}{8} - \frac{3y^2}{2} + 3 - \frac{3xy}{4}. \end{aligned}$$

y por tanto,

$$[F'(x, y)]^{-1} = \frac{1}{\Delta} \begin{pmatrix} \frac{3y^2}{8} - 3 & -\frac{y}{2} \\ -\frac{3x}{2} & \frac{x^2}{8} - 1 \end{pmatrix}.$$

De aquí se sigue que

$$\Gamma_0 = [F'(0, 0)]^{-1} = \begin{pmatrix} -1 & 0 \\ 0 & -1/3 \end{pmatrix}.$$

$$\Gamma_0 F(0, 0) = \begin{pmatrix} -1 & 0 \\ 0 & -1/3 \end{pmatrix} \begin{pmatrix} 1/3 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/3 \\ -1/3 \end{pmatrix}$$

y, por tanto, $a = \|\Gamma_0 F(0, 0)\| = 1/3$.

La derivada segunda de F viene dada por la siguiente expresión matricial

$$F''(x, y) = \begin{pmatrix} x/4 & 0 \\ 0 & 1/2 \\ 3/2 & 0 \\ 0 & 3y/4 \end{pmatrix}.$$

Entonces,

$$\Gamma_0 F''(0, 0) = \begin{pmatrix} -1 & 0 \\ 0 & -1/3 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1/2 \\ 3/2 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & -1/2 \\ -1/2 & 0 \\ 0 & 0 \end{pmatrix}$$

y por tanto, $\|\Gamma_0 F''(0, 0)\| = 1/2 = b$.

Además, tenemos que

$$\begin{aligned} \Gamma_0 [F''(x, y) - F''(0, 0)] &= \begin{pmatrix} -1 & 0 \\ 0 & -1/3 \end{pmatrix} \begin{pmatrix} x/4 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 3y/4 \end{pmatrix} \\ &= \begin{pmatrix} -x/4 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & -y/4 \end{pmatrix} \end{aligned}$$

y así, $\|\Gamma_0 [F''(x, y) - F''(0, 0)]\| \leq \max \left\{ \frac{|x|}{4}, \frac{|y|}{4} \right\} = \frac{1}{4} \|(x, y)\|$, de donde se deduce que $k = 1/4$.

Construimos el polinomio dado por (3.3), que en este caso es

$$p(t) = \frac{1}{3} - t + \frac{1}{4}t^2 + \frac{1}{24}t^3.$$

Observar que se cumple la condición (3.6), ya que

$$a - \frac{1}{2b} + \frac{k}{6b^3} = -\frac{1}{3} < 0.$$

Por tanto, $p(t)$ tiene raíces positivas. Así la sucesión obtenida aplicando el método de Newton a la ecuación $p(t) = 0$ con $t_0 = 0$, es mayorizante de $\{x_n\}$, definida por (3.2), empezando en $(0,0)$ y con F dada por (3.14).

Observar que en este caso también se verifican las condiciones de Kantorovich, ya que

$$\Gamma_0 F''(x, y) = \begin{pmatrix} -x/4 & 0 \\ 0 & -1/2 \\ -1/2 & 0 \\ 0 & -y/4 \end{pmatrix}$$

y, en consecuencia,

$$\|\Gamma_0 F''(x, y)\| \leq \frac{3}{4} = c, \quad \forall (x, y) \in X = [0, 1] \times [0, 1].$$

En este caso se tiene

$$ac = \frac{13}{34} = \frac{1}{4} < \frac{1}{2}.$$

luego se cumplen las condiciones de Kantorovich. El polinomio (3.10) que nos proporciona la sucesión mayorizante correspondiente es

$$q(t) = \frac{c}{2}t^2 - t + a = \frac{3}{8}t^2 - t + \frac{1}{3}.$$

Como

$$k \left(1 + \sqrt{1 - 2ac}\right) = \frac{1 + \sqrt{\frac{1}{2}}}{4} < 3c(c - b) = \frac{9}{16},$$

se cumple la condición (3.12), y obtenemos mejores estimaciones del error a partir del polinomio (3.3).

Denotamos por x^* a la solución de $F(x, y) = 0$, con F dada por (3.14). $\{x_n\}$ a la sucesión de Newton para resolver dicha ecuación. r_1 y \widehat{s}_1 son las menores raíces positivas de $p(t) = 0$ y $q(t) = 0$ respectivamente y $\{t_n\}$, $\{s_n\}$ son las sucesiones de Newton para resolver las ecuaciones anteriores, con $t_0 = s_0 = 0$.

iteración	$r_1 - t_n$	$\widehat{s}_1 - s_n$
0	0.3695850618081907	0.3905242917512699
1	0.0362517284748574	0.1067190958417936
2	0.0004701842187366	0.0163540286238108
3	0.0000000820038027	0.0000014159342765
4	0.0000000000000025	0.000000000010632

Tabla 4.1: Comparación del error

Por la teoría mayorizante sabemos que

$$\|x^* - x_n\| \leq r_1 - t_n$$

y también

$$\|x^* - x_n\| \leq \widehat{s}_1 - s_n.$$

En la Tabla 4.1 comparamos las estimaciones del error obtenidas en ambos casos. Como se puede comprobar, son mucho mejores las obtenidas a partir del polinomio de tercer grado, en las condiciones del Teorema 3.2. ■

Como vimos, la condición de que F'' esté acotada, que aparecía en las condiciones de Kantorovich, se puede debilitar a una condición tipo Lipschitz sobre F' [107] o incluso a condiciones más suaves de diferenciabilidad sobre F' [96], [8], [13]. En nuestro caso, también podemos pensar en dar condiciones sobre la derivada primera, del tipo $\|F'(x) - F'(y)\| \leq k\|x - y\|^2$.

Al exigir condiciones como (3.1) obtenemos resultados distintos a los obtenidos por Kantorovich y otros autores ya que, si bien imponemos condiciones más fuertes al operador F obtenemos un dominio de puntos de salida más amplio, como se vio en los ejemplos 3.1 y 3.5.

A continuación extenderemos los resultados de los Teoremas III.2.2 y 3.1, dando un resultado que generaliza estas dos situaciones. Para ello supondremos que F'' verifica una condición del tipo

$$\|\Gamma_0[F''(x) - F''(x_0)]\| \leq k\|x - x_0\|^p.$$

Notar que, para $p = 0$, tenemos $\|\Gamma_0 F''(x)\| \leq k + \|\Gamma_0 F''(x_0)\| = k'$, luego estamos en las condiciones del Teorema III.2.2; y para $p = 1$, estamos en las condiciones

del Teorema 3.1. En general, para $p \in [0, 1]$, diremos que F'' es continua de tipo (k, p) -Hölder [8].

Teorema 3.8. *Sea la ecuación $F(x) = 0$, con F un operador entre dos espacios de Banach X e Y , dos veces diferenciable Fréchet en un conjunto convexo Ω_0 . Sea $x_0 \in \Omega_0$ un punto en el que existe el operador inverso $\Gamma_0 = [F'(x_0)]^{-1}$. Supongamos que F'' verifica la condición*

$$\|\Gamma_0(F''(x) - F''(x_0))\| \leq k\|x - x_0\|^p, \quad k > 0, \quad p \geq 0, \quad x \in \Omega_0. \quad (3.14)$$

y que $\|\Gamma_0 F'(x_0)\| \leq a$, $\|\Gamma_0 F''(x_0)\| \leq b$. Si la función

$$f(t) = a - t + \frac{b}{2}t^2 + \frac{k}{(p+1)(p+2)}t^{p+2}, \quad t \geq 0, \quad (3.15)$$

tiene alguna raíz real positiva y $B(x_0, m) \subseteq \Omega_0$ (m es la única solución positiva de $f'(t) = 0$), entonces $F(x) = 0$ tiene solución x^* y además esta solución es el límite de la sucesión $\{x_n\}$ obtenida aplicando el método de Newton al operador F empezando en x_0 .

Observación. Antes de demostrar el Teorema 3.8, vamos a estudiar la función dada por (3.15) que aparece en su enunciado. Analizando sus derivadas, vemos que f' es creciente si $t \geq 0$ y, como $f'(0) < 0$ y $f'(t) > 0$ para t suficientemente grande, la ecuación $f'(t) = 0$ tiene una sola solución positiva, que denotaremos m . En consecuencia, m es un mínimo de f . La condición necesaria y suficiente para que f tenga raíces positivas es que $f(m) < 0$.

Como paso previo a la demostración del teorema, veamos el siguiente lema.

Lema 3.9. *Supongamos que F es un operador que cumple las condiciones del teorema anterior. Si $x \in B(x_0, m)$, siendo m el mínimo de la función f dada por (3.15), entonces el operador grado de convexidad logarítmico está definido y además*

$$\|L_F(x)\| \leq \frac{(b + k\|x - x_0\|^p)\|\Gamma_0 F(x)\|}{\left[1 - \frac{k}{(p+1)}\|x - x_0\|^{p+1} - b\|x - x_0\|\right]^2}.$$

Demostración. Siguiendo las mismas técnicas que en resultados anteriores, observar que

$$\int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)] dy = \Gamma_0 F'(x) - I - \Gamma_0 F''(x_0)(x - x_0).$$

Por tanto, teniendo en cuenta (3.14),

$$\begin{aligned} \|\Gamma_0 F'(x) - I - \Gamma_0 F''(x_0)(x - x_0)\| &= \left\| \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)] dy \right\| \\ &= \left\| \int_0^1 \Gamma_0 [F''(x_0 + t(x - x_0)) - F''(x_0)](x - x_0) dt \right\| \\ &\leq k \|x - x_0\|^{p+1} \int_0^1 t^p dt = \frac{k \|x - x_0\|^{p+1}}{p+1}. \end{aligned}$$

Así,

$$\|\Gamma_0 F'(x) - I\| \leq \frac{k}{p+1} \|x - x_0\|^{p+1} + b \|x - x_0\|.$$

Usando ahora que, $\|x - x_0\| \leq m$ y que $f'(t) = \frac{k}{p+1} t^{p+1} + bt < 1$ si $0 \leq t < m$, se sigue, por el Lema de Banach, que existe $[\Gamma_0 F'(x)]^{-1}$ y además

$$\|[\Gamma_0 F'(x)]^{-1}\| \leq \frac{1}{1 - \frac{k}{p+1} \|x - x_0\|^2 - b \|x - x_0\|}.$$

Considerando la expresión del operador grado de convexidad logarítmico, la estimación anterior, y que de (3.14) se deduce que

$$\|\Gamma_0 F''(x)\| \leq b + k \|x - x_0\|^p.$$

se obtiene la acotación buscada para el operador L_F . ■

Demostración del Teorema 3.8. El razonamiento es análogo al de las demostraciones de los Teoremas III.2.2 y 3.1. En las condiciones del teorema se tiene que x_1 está definido. Además, si $x \in [x_0, x_1]$, podemos escribir $x = x_0 + s(x_1 - x_0)$. Entonces, desarrollando en serie de Taylor la función $\Gamma_0 F$ en torno a x_0 ,

$$\begin{aligned} \Gamma_0 F(x) &= \Gamma_0 F(x_0) + (x - x_0) + \frac{1}{2} \Gamma_0 F''(x_0)(x - x_0)^2 \\ &\quad + \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)](x - y) dy. \end{aligned}$$

Como $x - x_0 = s(x_1 - x_0) = -s\Gamma_0 F(x_0)$,

$$\begin{aligned} \|\Gamma_0 F(x)\| &\leq (1 - s)\|\Gamma_0 F(x_0)\| + \frac{1}{2}\|\Gamma_0 F''(x_0)\| \|s\Gamma_0 F(x_0)\|^2 \\ &\quad + \left\| \int_{x_0}^x \Gamma_0 [F''(y) - F''(x_0)](x - y) dy \right\|. \end{aligned}$$

Teniendo en cuenta que

$$\begin{aligned} \left\| \int_{x_0}^x \Gamma_0[F''(y) - F''(x_0)](x - y) dy \right\| &\leq k \|x - x_0\|^{2+p} \int_0^1 s^p(1 - s) ds \\ &= \frac{k \|x - x_0\|^{2+p}}{(p + 1)(p + 2)}. \end{aligned}$$

deducimos

$$\|\Gamma_0 F(x)\| \leq (1 - s)a + \frac{1}{2}bs^2a^2 + \frac{k}{(p + 1)(p + 2)}s^{2+p}a^{2+p}.$$

Si llamamos $t = sa$, concluimos que

$$\|\Gamma_0 F(x)\| \leq a - t + \frac{1}{2}bt^2 + \frac{k}{(p + 1)(p + 2)}t^{2+p}.$$

Entonces, el Lema 3.9 nos da la siguiente estimación para el operador grado de convexidad logarítmico. Si $x \in [x_0, x_1]$ y $t = \|x - x_0\|$, se sigue que

$$\|L_F(x)\| \leq \frac{\left(a - t + \frac{1}{2}bt^2 + \frac{k}{(p+1)(p+2)}t^{2+p}\right) (b + kt^p)}{\left(1 - bt - \frac{k}{p+1}t^{1+p}\right)^2} = L_f(t)$$

con f dada por (3.15).

Como $x_2 - x_1 = \int_{x_0}^{x_1} L_F(x) dx$, se deduce que x_2 está definido y además

$$\|x_2 - x_1\| = \left\| \int_{x_0}^{x_1} L_F(x) dx \right\| \leq \int_{t_0}^{t_1} L_f(t) dt = t_2 - t_1.$$

Podemos intercambiar x_1 por x_0 y x_2 por x_1 en el proceso anterior para obtener que x_3 está definido y

$$\|x_3 - x_2\| \leq t_3 - t_2.$$

En general, siguiendo un razonamiento inductivo, probamos que

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n.$$

En consecuencia, la sucesión $\{t_n\}$ mayoriza a $\{x_n\}$ y la convergencia de la primera implica la de la segunda. ■

Vamos a estudiar ahora el método de Newton aplicado a funciones f de la forma (3.15). De acuerdo con la observación que sigue al enunciado del Teorema 3.5.

sabemos que f tiene dos raíces reales positivas, que denotaremos r_1 y r_2 ($r_1 \leq r_2$). En consecuencia, podemos escribir

$$f(t) = (r_1 - t)(r_2 - t)g(t), \quad \text{con } g(r_1) \neq 0 \neq g(r_2). \quad (3.16)$$

Daremos a continuación un resultado que proporciona estimaciones del error cometido al emplear el método de Newton en este caso. Este resultado generaliza los Teoremas 3.3 y III.2.4.

Teorema 3.10. *En las condiciones anteriores, la sucesión de Newton $\{t_n\}$ para resolver $f(t) = 0$, con f dada por (3.15) y $t_0 = 0$, es convergente a r_1 , la menor raíz positiva de $f(t) = 0$. Además, para $r_1 < r_2$, se tienen las siguientes estimaciones del error*

$$\frac{(r_2 - r_1)}{m - \alpha^{2^n}} \alpha^{2^n} \leq r_1 - t_n \leq \frac{(r_2 - r_1)}{M - \theta^{2^n}} \theta^{2^n},$$

donde

$$\begin{aligned} \theta &= M \frac{r_1}{r_2}, & \alpha &= m \frac{r_1}{r_2}, \\ m &= \min_{t \in [0, r_1]} H(t), & M &= \max_{t \in [0, r_1]} H(t) \end{aligned}$$

y

$$H(t) = \frac{g(t) - (r_2 - t)g'(t)}{g(t) - (r_1 - t)g'(t)}.$$

Demostración. Como f es decreciente y convexa en $[0, r_1]$, tenemos garantizada la convergencia del método de Newton a la raíz r_1 tomando como punto de salida $t_0 = 0$.

Por otra parte, a partir de (3.16) deducimos que

$$f'(t) = -(r_2 - t)g(t) - (r_1 - t)g(t) + (r_1 - t)(r_2 - t)g'(t).$$

Entonces, denotando $a_n = r_1 - t_n$, $b_n = r_2 - t_n$ y por un proceso similar al seguido en la demostración del Teorema 3.4, tenemos que

$$a_{n+1} = a_n^2 \frac{-g(t_n) + b_n g'(t_n)}{f'(t_n)}, \quad b_{n+1} = b_n^2 \frac{-g(t_n) + a_n g'(t_n)}{f'(t_n)}.$$

En consecuencia,

$$m \left(\frac{a_n}{b_n} \right)^2 \leq \frac{a_{n+1}}{b_{n+1}} = \left(\frac{a_n}{b_n} \right)^2 \frac{g(t_n) - b_n g'(t_n)}{g(t_n) - a_n g'(t_n)} \leq M \left(\frac{a_n}{b_n} \right)^2.$$

De aquí se sigue que

$$\frac{1}{m} \left(m \frac{r_1}{r_2} \right)^{2^n} \leq \frac{a_n}{b_n} \leq \frac{1}{M} \left(M \frac{r_1}{r_2} \right)^{2^n}.$$

Sin más que tener en cuenta que $b_n = r_2 - r_1 + a_n$, se deduce el resultado. ■

En el caso en que $r_1 = r_2$, sabemos que la convergencia de la sucesión de Newton a la raíz doble r_1 es muy lenta. Siguiendo un razonamiento similar al utilizado en la demostración anterior, podemos ver que

$$a_{n+1} = r_1 - t_{n+1} = a_n \frac{g(t_n) - a_n g'(t_n)}{2g(t_n) - a_n g'(t_n)}.$$

Entonces, si

$$m = \min_{t \in [0, r_1]} H(t), \quad M = \max_{t \in [0, r_1]} H(t)$$

con

$$H(t) = \frac{g(t) - (r_1 - t)g'(t)}{2g(t) - (r_1 - t)g'(t)},$$

podemos dar las siguientes estimaciones del error

$$m^n r_1 \leq r_1 - t_n \leq M^n r_1.$$

4.4 Aplicación a las ecuaciones diferenciales e integrales

A continuación vamos a emplear los resultados de la sección anterior para estudiar algunas ecuaciones diferenciales e integrales, dando resultados sobre convergencia y unicidad de solución para ellas.

Así, por ejemplo, consideramos el espacio $X = \{y \in C^1[0, a]; y(0) = 0\}$ con la norma

$$\|y\| = \max_{t \in [0, a]} |y(t)| + \lambda \max_{t \in [0, a]} |y'(t)|,$$

con $\lambda > 0$ que se determinará más adelante.

Vamos a estudiar la ecuación diferencial

$$\begin{cases} y'(t) - o(t, y(t)) = 0 \\ y(0) = 0. \end{cases} \quad (4.1)$$

Suponemos que $o(t, u)$ es continua y tiene derivada segunda continua respecto a u en la región

$$0 \leq t \leq a, \quad |u - y_0(t)| \leq \delta, \quad y_0 \in X.$$

Sea $\Omega = \{y \in X: \|y - y_0\| \leq \delta\}$ y $F: \Omega \subseteq X \rightarrow C[0, a]$ el operador definido por

$$F(y)(t) = y'(t) - o(t, y(t)), \quad y \in \Omega.$$

El problema de resolver la ecuación diferencial (4.1) es equivalente al de encontrar una solución de $F(y) = 0$.

Para $y_0 \in \Omega$, tenemos

$$F'(y_0)y(t) = y'(t) - o'_u(t, y_0(t))y(t),$$

$$F''(y_0)xy(t) = -o''_{u^2}(t, y_0(t))x(t)y(t),$$

con $x, y \in \Omega$.

Sea $\Gamma_0 = F'(y_0)^{-1}$. Nuestro objetivo es encontrar una cota superior para $\|\Gamma_0\|$. Para ello sea $x = \Gamma_0(y)$. Entonces x es solución de la ecuación diferencial

$$x' = y(t) + o'_u(t, y_0(t))x,$$

esto es, [23], [100],

$$x(t) = v(t) \int_0^t \frac{y(s)}{v(s)} ds,$$

con $v(t) = \exp\left(\int_0^t o'_u(s, y_0(s)) ds\right)$.

Observar que

$$\max_{t \in [0, a]} |x(t)| \leq a \frac{\max_{t \in [0, a]} |v'(t)|}{\min_{t \in [0, a]} |v'(t)|} \|y\|$$

y que

$$\max_{t \in [0, a]} |x'(t)| \leq \max_{t \in [0, a]} |o'_u(t, y_0(t))| \max_{t \in [0, a]} |x(t)| + \|y\| \leq \theta \|y\|,$$

con

$$\theta = 1 + a \max_{t \in [0, a]} |o'_u(t, y_0(t))| \frac{\max_{t \in [0, a]} |v'(t)|}{\min_{t \in [0, a]} |v'(t)|}.$$

En consecuencia,

$$\|\Gamma_0\| \leq \lambda \theta + a \frac{\max_{t \in [0, a]} |v'(t)|}{\min_{t \in [0, a]} |v'(t)|}. \quad (4.2)$$

A partir de esta cota podemos dar resultados sobre existencia y unicidad de solución de la ecuación diferencial (4.1), aplicando los Teoremas 3.2 y 3.8.

Teorema 4.1. Para $y_0 \in \Omega$. supongamos que se cumplen las siguientes condiciones:

(i) $|y'_0(t) - \phi(t, y_0(t))| \leq a'$, $t \in [0, a]$.

(ii) $|\phi'_u(t, y_0(t))| \leq M_1$, $t \in [0, a]$.

(iii) $|\phi''_{u^2}(t, y_0(t))| \leq b'$, $t \in [0, a]$.

(iv) $|\phi''_{u^2}(t, u) - \phi''_{u^2}(t, v)| \leq k'|u - v|$, $t \in [0, a]$, $|u - y_0(t)| \leq \delta$, $|v - y_0(t)| \leq \delta$.

Sean $\theta_1 = ae^{2aM_1}$, $\theta_2 = 1 + M_1\theta_1$, $B = \theta_1 + \lambda\theta_2$, $a = a'B$, $b = b'B$ y $k = k'B$. Si el polinomio

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a$$

tiene dos raíces reales positivas $r_1 \leq r_2$ y se cumple $m \leq \delta$ (m es el mínimo de p), entonces la ecuación diferencial (4.1) tiene solución y^* definida en $[0, a]$. Además

$$|y^*(t) - y_0(t)| \leq r_1, \quad t \in [0, a]$$

y la solución es única en la región $|u - y_0(t)| \leq r_2$.

Demostración. Se sigue sin más que tener en cuenta que

$$\max_{t \in [0, a]} |v(t)| \leq e^{aM_1}, \quad \min_{t \in [0, a]} |v(t)| \geq e^{-aM_1},$$

y, por (4.2), $\|\Gamma_0\| \leq \theta_1 + \lambda\theta_2 = B$. Como

$$\|\Gamma_0 F(x_0)\| \leq \|\Gamma_0\| \|F(x_0)\| \leq Ba' = a,$$

y, análogamente

$$\|\Gamma_0 F''(x_0)\| \leq b, \quad \|\Gamma_0 [F''(y) - F''(y_0)]\| \leq k\|y - y_0\|,$$

el resultado es consecuencia inmediata del Teorema 3.2. ■

Notar que si tomamos $b = \theta_1$ y el polinomio p tiene dos raíces positivas distintas, podemos garantizar la existencia de solución tomando λ suficientemente pequeño.

Con un razonamiento similar, pero usando el Teorema 3.8, podemos enunciar el siguiente resultado general, que incluye como caso particular el Teorema 4.1.

Teorema 4.2. *En las condiciones del teorema anterior, pero cambiando (iv) por (iv') $|\phi''_{u^2}(t, u) - \phi''_{u^2}(t, v)| \leq k'|u - v|^p$, $p \geq 0$, para t, u y v satisfaciendo*

$$t \in [0, a], \quad |u - y_0(t)| \leq \delta, \quad |v - y_0(t)| \leq \delta.$$

Sean $\theta_1 = ne^{2aM_1}$, $\theta_2 = 1 + M_1\theta_1$, $B = \theta_1 + \lambda\theta_2$, $a = a'B$, $b = b'B$ y $k = k'B$. Si la función

$$f(t) = \frac{k}{(p+1)(p+2)}t^{2+p} + \frac{b}{2}t^2 - t + a$$

satisface las condiciones del Teorema 3.8, esto es, tiene dos raíces reales positivas $r_1 \leq r_2$ y se cumple $m \leq \delta$ (m es el mínimo de f), entonces la ecuación diferencial (4.1) tiene solución y^* definida en $[0, a]$. Además

$$|y^*(t) - y_0(t)| \leq r_1, \quad t \in [0, a]$$

y la solución es única en la región $|u - y_0(t)| \leq r_2$.

Nota. Si en lugar de exigir (iv) o (iv') se supone que $\phi''_{u^2}(t, u)$ está acotada para $t \in [0, a]$ y $|u - y_0(t)| \leq \delta$, se cumplen las condiciones clásicas de Kantorovich, y se obtienen resultados similares a partir de un polinomio de segundo grado [66].

Consideramos ahora la ecuación integral

$$x(s) = \int_0^1 K(s, t, x(t)) dt, \quad s \in [0, 1], \quad (4.3)$$

donde $K(s, t, u)$ es una función continua en todos sus argumentos y tiene derivadas continuas de cualquier orden. Definimos en un espacio de funciones X el operador F de manera que $F(x) = y$, con

$$y(s) = x(s) - \int_0^1 K(s, t, x(t)) dt.$$

Queremos resolver la ecuación $F(x) = 0$. Para ello, partimos de una aproximación inicial x_0 . Supongamos que X y K son lo suficientemente buenos para que $F'(x_0)$ se obtenga derivando bajo el signo integral, es decir, si $z = F'(x_0)x$, entonces

$$z(s) = x(s) - \int_0^1 K'_u(s, t, x_0(t))x(t) dt. \quad (4.4)$$

Es bien conocido [66], [87], que el operador $\Gamma_0 = F'(x_0)^{-1}$ tiene la forma $w = \Gamma_0(y)$ con

$$w(s) = y(s) + \int_0^1 G(s, t)y(t) dt.$$

y $G(s, t)$ resolvente de la ecuación (4.4) para el núcleo $K(s, t) = K'_u(s, t, x_0(t))$. Entonces, si

$$\int_0^1 |G(s, t)| dt \leq B, \quad s \in [0, 1],$$

se tiene que $\|\Gamma_0\| \leq 1 + B$.

A partir de esta cota, podemos establecer resultados de existencia y unicidad de solución para la ecuación (4.3) como consecuencia de los Teoremas 3.2 y 3.8.

Teorema 4.3. *En la situación anterior, supongamos que además se cumple*

- (i) $\left| \int_0^1 K(s, t, x_0(t)) dt - x_0(s) \right| \leq a', \quad s \in [0, 1].$
- (ii) $\left| \int_0^1 K''_{u^2}(s, t, x_0(t)) dt \right| \leq b', \quad s \in [0, 1].$
- (iii) $\left| \int_0^1 [K''_{u^2}(s, t, u) - K''_{u^2}(s, t, v)] dt \right| \leq k'|u - v|$, para s, u y v satisfaciendo $s \in [0, 1]$, $|u - x_0(s)| \leq \delta$, $|v - x_0(s)| \leq \delta$.

Sean $a = a'(1 + B)$, $b = b'(1 + B)$ y $k = k'(1 + B)$. Si el polinomio

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a$$

tiene dos raíces reales positivas $r_1 \leq r_2$ y se cumple $m \leq \delta$ (m es el mínimo de p), entonces la ecuación integral (4.3) tiene solución x^* definida en $[0, 1]$. Además

$$|x^*(t) - x_0(t)| \leq r_1, \quad t \in [0, 1]$$

y la solución es única en la región $|u - x_0(t)| \leq r_2$.

De forma similar, sustituyendo (iii) por

- (iii') $\left| \int_0^1 [K''_{u^2}(s, t, u) - K''_{u^2}(s, t, v)] dt \right| \leq k'|u - v|^p$, $p \geq 0$ para s, u y v satisfaciendo $s \in [0, 1]$, $|u - x_0(s)| \leq \delta$, $|v - x_0(s)| \leq \delta$.

se obtiene un resultado a partir del Teorema 3.8 y la función

$$f(t) = \frac{k}{(p+1)(p+2)}t^{2+p} + \frac{b}{2}t^2 - t + a.$$

Capítulo 5

La aceleración convexa del método de Newton

5.1 Introducción

Como es habitual en el estudio de procesos iterativos para la resolución de una ecuación escalar no lineal $f(t) = 0$, suponemos que f verifica las condiciones de Fourier [83]. En este caso, sin pérdida de generalidad, consideramos una función f estrictamente decreciente y convexa en un intervalo $[a, b]$, y que cumple $f(a) > 0 > f(b)$. En otro caso cambiamos $f(t)$ por $-f(t)$, $f(-t)$ o $-f(-t)$. En estas condiciones es conocido que la sucesión $\{t_n\}$ obtenida a partir del método de Newton-Raphson

$$t_{n+1} = t_n - \frac{f(t_n)}{f'(t_n)}, \quad t_0 \in [a, b], \quad f(t_0) > 0 \quad (1.1)$$

converge a r , única solución de $f(t) = 0$ en $[a, b]$. Además la sucesión obtenida es creciente.

Sea g una función en las mismas condiciones que f en $[a, b]$, con $g(r) = 0$, y $\{s_n\}$ la sucesión

$$s_{n+1} = s_n - \frac{g(s_n)}{g'(s_n)}, \quad s_0 = t_0.$$

Las dos sucesiones $\{t_n\}$ y $\{s_n\}$ convergen a r . Además, ambas son crecientes. Como vimos en el Capítulo 1, podemos comparar la velocidad con la que ambas sucesiones convergen a r mediante un estudio del grado de convexidad logarítmico de

las funciones f y g . Así, en el Teorema I.3.2 vimos que si $L_g(t) < L_f(t)$ para todo $t \in [a, r]$, entonces $\{s_n\}$ converge a r más rápidamente que $\{t_n\}$.

Utilizando este resultado intentaremos obtener una aceleración punto a punto para el método de Newton-Raphson. Es decir, dada $\{t_n\}$ por (1.1) tratamos de definir una aceleración $\{s_n\}$ tal que $s_{n+1} = G(t_n)$, a partir de un operador G . Este tipo de aceleraciones tienen gran interés, ya que una vez construido G nos permiten definir nuevos procesos iterativos

$$z_{n+1} = G(z_n)$$

independientemente de la sucesión $\{t_n\}$. Para esto, observar que el grado de convexidad logarítmico de una recta es cero y, por tanto, es el menor grado de convexidad logarítmico que puede tener una función convexa. Así, la recta $g(t) = f'(r)(t - r)$ cumple las condiciones exigidas para f y su grado de convexidad logarítmico es cero. Obviamente, la sucesión obtenida por el método de Newton-Raphson para esta recta converge más rápidamente que la sucesión obtenida para la función f .

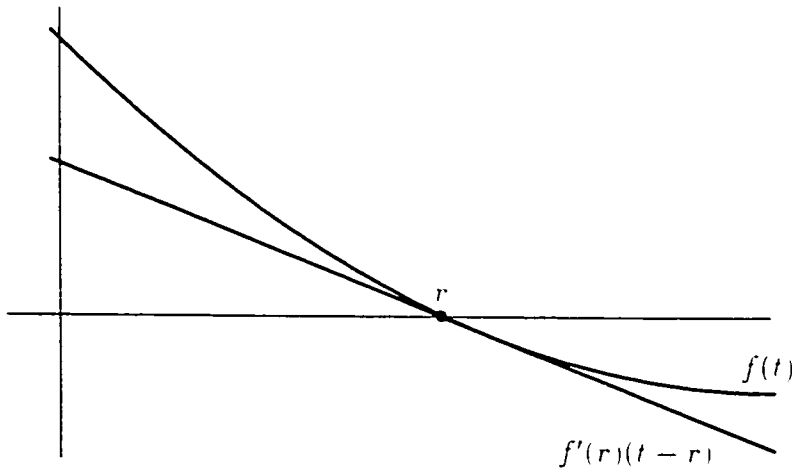


Figura 5.1: Aproximación de una función por su recta tangente

Teniendo en cuenta el desarrollo de Taylor para la función f , tomaremos la cantidad $f(t) - \frac{f''(r)}{2}(t-r)^2$ como aproximación de la recta $f'(r)(t-r)$.

El problema que surge es que r , en principio, es desconocido. Por ello hacemos las siguientes estimaciones.

Si consideramos $g(t) = f(t) - \frac{f''(r)}{2}(t-r)^2$ entonces podemos aproximar

$$g(t_n) \sim f(t_n) - \frac{f''(t_n)}{2}(t_n - t_{n+1})^2.$$

ya que $f''(t_n) \sim f''(r)$, la sucesión $\{t_n\}$ converge a r y $(t_n - t_{n+1})^2 \sim (t_n - r)^2$ debido a que

$$\lim_{n \rightarrow \infty} \frac{t_n - r}{t_n - t_{n+1}} = 1.$$

como se comprueba fácilmente.

Por otra parte, como $g'(t) = f'(t) - f''(r)(t-r)$ entonces

$$g'(t_n) \sim f'(t_n) - f''(t_n)(t_n - t_{n+1}).$$

Con estas aproximaciones tratamos de construir una aceleración para el método de Newton

$$\begin{aligned} s_{n+1} &= t_n - \frac{g(t_n)}{g'(t_n)} = t_n - \frac{f(t_n) - \frac{f''(t_n)}{2} \left(\frac{f(t_n)}{f'(t_n)} \right)^2}{f'(t_n) - f''(t_n) \frac{f(t_n)}{f'(t_n)}} \\ &= t_n - \frac{f(t_n)}{2f'(t_n)} \frac{2 - L_f(t_n)}{1 - L_f(t_n)}. \end{aligned} \quad (1.2)$$

Lema 1.1. Si llamamos

$$G(t) = t - \frac{f(t)}{2f'(t)} \frac{2 - L_f(t)}{1 - L_f(t)} = t - \frac{f(t)}{2f'(t)} \left[1 + \frac{1}{1 - L_f(t)} \right], \quad (1.3)$$

tenemos que

$$G'(t) = \frac{L_f^2(t)(L_f(t) - L_{f'}(t))}{2(1 - L_f(t))^2}.$$

Demostración. Es una comprobación inmediata. ■

Teorema 1.2. Si f es una función que satisface las condiciones anteriores, entonces la sucesión $\{s_n\}$ definida por (1.2) es una aceleración de la sucesión de Newton (1.1).

Demostración. Basta con que veamos que

$$\lim_{n \rightarrow \infty} \frac{r - s_{n+1}}{r - t_{n+1}} = \lim_{n \rightarrow \infty} \frac{r - G(t_n)}{r - F(t_n)} = 0.$$

donde G es la función definida por (1.3) y $F(t) = t - \frac{f(t)}{f'(t)}$.

Teniendo en cuenta el Lema 1.1 y que la derivada de F es L_f , se sigue que

$$\lim_{t \rightarrow r} \frac{r - G(t)}{r - F(t)} = \lim_{t \rightarrow r} \frac{G'(t)}{F'(t)} = \lim_{t \rightarrow r} \frac{L_f(t)(L_f(t) - L_{f'}(t))}{2(1 - L_f(t))^2} = 0. \quad \blacksquare$$

La aceleración anterior nos permite definir un nuevo proceso iterativo. Comenzando en $t_0 \in [a, b]$ tal que $f(t_0) > 0$ definimos la sucesión

$$t_{n+1} = G(t_n) \quad (1.4)$$

con G dada por (1.3).

En los siguientes resultados vamos a ver que esta nueva sucesión es convergente a r , solución de $f(t) = 0$ y que además esta convergencia es al menos de tercer orden, con lo que mejoramos el método de Newton.

Teorema 1.3. *Sea f una función decreciente y convexa en un intervalo $[a, b]$ verificando $f(a) > 0 > f(b)$. Sea $r \in (a, b)$ tal que $f(r) = 0$. Si $f'(t) \neq 0$ y*

$$L_{f'}(t) \leq L_f(t) < 1 \quad (1.5)$$

para todo $t \in [a, r]$, entonces la sucesión dada por (1.4) converge a r , solución de $f(t) = 0$. Además esta sucesión es creciente a r .

Demostración. Observar que la sucesión dada por (1.4) se puede escribir

$$t_{n+1} = \frac{1}{2} [F(t_n) + Q(t_n)]$$

con

$$F(t) = t - \frac{f(t)}{f'(t)},$$

$$Q(t) = t - \frac{q(t)}{q'(t)} \quad \text{siendo} \quad q(t) = \frac{f(t)}{f'(t)}.$$

Entonces, teniendo en cuenta que $F'(t) = L_f(t)$ y $Q'(t) = L_q(t)$,

$$\begin{aligned} F'(t) + Q'(t) &= L_f(t) + L_q(t) \\ &= L_f(t) - \frac{L_f(t) [1 - L_f(t)(2 - L_{f'}(t))]}{(1 - L_f(t))^2} \\ &= \frac{L_f^2(t)}{(1 - L_f(t))^2} (L_f(t) - L_{f'}(t)). \end{aligned}$$

De (1.5) se sigue que $F'(t) + Q'(t) \geq 0$ en $[a, r]$. Entonces,

$$t_1 - r = \frac{F+Q}{2}(t_0) - \frac{F+Q}{2}(r) \leq \left(\frac{F+Q}{2}\right)'(z_0)(t_0 - r)$$

para algún $z_0 \in (t_0, r)$. En consecuencia, $t_1 - r \leq 0$. Por un razonamiento inductivo se puede probar sin dificultad que $t_n \leq r \forall n \in \mathbb{N}$.

A continuación, veamos que $\{t_n\}$ es una sucesión creciente. Observar que

$$t_{n+1} - t_n = -\frac{f(t_n)}{f'(t_n)}H(L_f(t_n)),$$

donde

$$H(t) = \frac{2-t}{2(1-t)}.$$

Por las condiciones del teorema, $f(t) > 0$ y $f'(t) < 0$ en $[a, r]$. Además, como $L_f(t) < 1$ en $[a, r]$ y $H(t) > 0$ si $t < 1$, tenemos que $t_{n+1} - t_n \geq 0$.

Así, la sucesión $\{t_n\}$ es creciente y acotada, por tanto tiene límite. Sea

$$u = \lim_{n \rightarrow \infty} t_n \leq r.$$

Tomando límites en (1.2), y teniendo en cuenta que $H(L_f(t_n)) \neq 0$, obtenemos que $f(u) = 0$. Como r es la única solución de $f(t) = 0$ en $[a, b]$, se sigue que $u = r$. ■

Nota. Una condición suficiente para que se verifique la condición (1.5) es que $f'''(t) \geq 0$ en $[t_0, r]$. En efecto, en estas condiciones se cumple el Lema de Altman 1.2.3, por tanto, $L_f(t) < 1/2 < 1$. La otra desigualdad es inmediata, ya que $L_f(t) \geq 0 \geq L_{f'}(t)$ si $t \in [t_0, r]$.

Observar que esta condición la cumplen los polinomios de segundo grado y los de tercer grado decrecientes y convexos en $[t_0, r]$, con coeficiente director positivo.

Después de esta observación podemos enunciar la siguiente consecuencia.

Corolario 1.4. *Sea f una función decreciente y convexa en un intervalo $[a, b]$ verificando $f(a) > 0 > f(b)$. Sea $r \in (a, b)$ tal que $f(r) = 0$. Si $f'(t) \neq 0$ y $f'''(t) \geq 0$ para t en $[a, r]$, entonces la sucesión dada por (1.4) converge a r , solución de $f(t) = 0$, y además de forma creciente.*

Se pueden estudiar otros casos de convergencia para este nuevo proceso iterativo en los que la condición (1.5) no se cumpla. Algunas de estas situaciones son más complicadas, ya que la sucesión que obtenemos puede no ser creciente a r . En [53] pueden verse condiciones que garantizan la convergencia en otros casos.

Estudiaremos ahora el orden del proceso iterativo dado por (1.4). Como consecuencia de un resultado de W. Gander [43], veremos que, en general, el método tiene al menos orden tres.

Teorema 1.5. *Sea r un cero simple de f y H una función verificando $H(0) = 1$, $H'(0) = 1/2$, $H''(0) < \infty$. Entonces la iteración $t_{n+1} = F(t_n)$ con*

$$F(t) = t - \frac{f(t)}{f'(t)} H(L_f(t))$$

es al menos de tercer orden.

En nuestro caso, la función $H(t) = \frac{2-t}{2(1-t)}$ está en las condiciones de este resultado de Gander. Por tanto, el método obtenido a partir de (1.4) tiene al menos convergencia cúbica. Sin embargo, para una clase concreta de funciones (polinomios de segundo grado), se alcanza la convergencia de orden cuatro. Todas estas consideraciones acerca del orden de convergencia se entienden para funciones con raíces simples. Para raíces de multiplicidad mayor que uno, la convergencia es lineal, al igual que sucedía en el método de Newton [103].

Veamos ahora lo que ocurre si utilizamos el método de la aceleración convexa para resolver una ecuación

$$p(t) = 0.$$

con p un polinomio de segundo grado.

Teorema 1.6. *Sea p un polinomio de segundo grado y r una raíz simple de la ecuación $p(t) = 0$. En esta situación, la aceleración convexa del método de Newton es un método de cuarto orden.*

Demostración. Sea G la función dada por (1.3). Observar que $G(r) = r$. Aplicando el Lema 1.1 a un polinomio de segundo grado tenemos

$$G'(t) = \frac{L_p^3(t)}{2(1 - L_p(t))^2}$$

y, por tanto, $G'(r) = 0$. Derivando la expresión anterior se comprueba que $G''(r) = 0$ y también $G'''(r) = 0$. Sin embargo,

$$G^{(4)}(r) = \frac{3p''(r)^3}{p'(r)^3} \neq 0.$$

En consecuencia, el método definido por (1.4) es de orden cuatro si lo aplicamos a polinomios de segundo grado. ■

Nota. Observar además que la constante del error asintótico [61] en este caso viene dada por

$$\lim_{t \rightarrow r} \frac{G(t) - r}{t - r} = \lim_{t \rightarrow r} \frac{G^{(4)}(t)}{24} = \frac{p''(r)^3}{8p'(r)^3}.$$

En el teorema anterior hemos probado que la aceleración convexa del método de Newton para resolver la ecuación $p(t) = 0$, donde p es un polinomio de segundo grado, es un método de orden cuatro. Pero aún se puede probar más: en efecto, se tiene que, en este caso, al hacer dos iteraciones en el método de Newton se obtiene el mismo punto que al hacer una iteración en su aceleración convexa.

Teorema 1.7. *Sea p un polinomio de segundo grado con dos raíces positivas. Supongamos, sin pérdida de generalidad, que*

$$p(t) = bt^2 - t + a, \quad ab \leq \frac{1}{4}.$$

Denotamos $\{z_n\}$ la sucesión de Newton para resolver $p(t) = 0$, y $\{t_n\}$ su aceleración convexa (1.4), con $z_0 = t_0$. Entonces $z_{2n} = t_n$ para $n \geq 0$.

Demostración. Sea G la función definida en (1.3) con $f = p$ y

$$F(t) = t - \frac{p(t)}{p'(t)}.$$

Entonces

$$\begin{aligned} F(F(t)) &= \frac{b(bt^2 - a)^2 - a(2bt - 1)^2}{(2bt - 1)(2b(bt^2 - a) - (2bt - 1))} \\ &= \frac{b^3t^4 - 6ab^2t^2 + 4abt + a^2b - a}{(2bt - 1)(2b(bt^2 - a) - (2bt - 1))} = G(t). \end{aligned}$$

Por tanto,

$$z_2 = F(z_1) = F(F(z_0)) = G(z_0) = G(t_0) = t_1.$$

$$z_4 = F(z_3) = F(F(z_2)) = G(z_2) = G(t_1) = t_2.$$

y así sucesivamente. ■

Por último, hacemos algunas consideraciones en el caso de que la aceleración convexa del método de Newton se aplique a una función con raíces de multiplicidad

mayor que uno. Vemos que, en este caso, la convergencia obtenida es lineal. No obstante, la constante de error asintótico en esta situación es menor que la del método de Newton y, por tanto, siendo procesos con el mismo orden de convergencia, será más rápido el correspondiente a la aceleración convexa del método de Newton [81]. Partimos ahora de una ecuación

$$f(t) = 0,$$

donde $f(t) = (r - t)^\gamma g(t)$ con $\gamma > 1$ y $g(r) \neq 0$. En esta situación se tienen los siguientes resultados.

Lema 1.8. *Si f es diferenciable un número suficiente de veces entonces*

$$L_f(r) = \frac{\gamma - 1}{\gamma},$$

$$L_{f'}(r) = \frac{\gamma - 2}{\gamma - 1}$$

y, en general,

$$L_{f^{(n)}}(r) = \frac{\gamma - (n + 1)}{\gamma - n}.$$

Demostración. Las igualdades anteriores se obtienen sin más que sustituir, en el punto $x = r$, el valor de f y de sus derivadas en la expresión correspondiente del grado de convexidad logarítmico. ■

Teorema 1.9. *Sea r una raíz de multiplicidad $\gamma > 1$ de la función f . Entonces la aceleración convexa del método de Newton para resolver la ecuación $f(t) = 0$ tiene convergencia lineal. En esta situación, si denotamos C_N y C_A a las constantes de error asintótico del método de Newton y su aceleración convexa respectivamente, tenemos que*

$$C_A = \frac{\gamma - 1}{2\gamma} = \frac{1}{2}C_N.$$

Demostración. Sea G la función definida por (1.3). Observar que $G(r) = r$. Teniendo en cuenta la expresión para la derivada de G dada por el Lema 1.1 y las expresiones de $L_f(r)$ y $L_{f'}(r)$ cuando r es una raíz múltiple, vistas en el lema anterior, se sigue que

$$G''(r) = \frac{\gamma - 1}{2\gamma} \neq 0.$$

Si usamos que la derivada de

$$t - \frac{f(t)}{f'(t)}$$

es $L_f(t)$. el resto de la demostración se sigue inmediatamente de la igualdad anterior y del Lema 1.8. ■

5.2 La aceleración convexa en espacios de Banach

Vamos a generalizar el estudio del método anterior a espacios de Banach. Partimos de una ecuación no lineal $F(x) = 0$ siendo F un operador $F : \Omega \subseteq X \rightarrow Y$. con X e Y espacios de Banach y Ω un conjunto convexo. Formalmente, podemos escribir el método de Newton acelerado como sigue

$$x_{n+1} = x_n - \left[I + \frac{1}{2} L_F(x_n) \Delta_n^{-1} \right] \Gamma_n F(x_n) \quad (2.1)$$

siendo $\Delta_n = [I - L_F(x_n)]$ y $L_F(x_n)$ el operador lineal que aplica X en X de la siguiente manera

$$L_F(x_n)(x) = [F'(x_n)]^{-1} F''(x_n) [F'(x_n)]^{-1} F(x_n)(x), \quad x \in X$$

y que se introdujo en el Capítulo 1.

A continuación estudiamos las condiciones que tienen que cumplirse para que este método converja a una solución de $F(x) = 0$. Para este estudio empleamos sucesiones mayorizantes. en concreto, la sucesión obtenida aplicando la aceleración convexa del método de Newton a la ecuación $p(t) = 0$ con $p(t)$ una función real decreciente y convexa en un intervalo $[a, b]$, con $p(a) > 0 > p(b)$. Denotamos por r a la única solución de $p(t) = 0$ en $[a, b]$. Llamamos $\{t_n\}$ a esta sucesión, que podemos escribir de la forma

$$t_{n+1} = t_n - \frac{2 - L_p(t_n)}{2(1 - L_p(t_n))} \frac{p(t_n)}{p'(t_n)} = t_n - \left[1 + \frac{1}{2} \frac{L_p(t_n)}{1 - L_p(t_n)} \right] \frac{p(t_n)}{p'(t_n)}. \quad (2.2)$$

En primer lugar, damos un resultado general de convergencia, en términos de una función p que mayoriza al operador F . Este resultado es similar al Teorema I.4.4 dado por Kantorovich para probar la convergencia del método de Newton, o del dado por Altman [4] para probar la convergencia del método de Halley.

Teorema 2.1. Con la notación anterior, sea $x_0 \in \Omega$ un punto en el que existe $\Gamma_0 = [F'(x_0)]^{-1}$ y $t_0 \in [a, b]$ tal que $p(t_0) > 0$. Si se cumplen las condiciones

$$[C_1] \quad \|\Gamma_0 F(x_0)\| \leq -\frac{p(t_0)}{p'(t_0)}.$$

$$[C_2] \quad \|\Gamma_0 F''(x)\| \leq -\frac{p''(t)}{p'(t_0)} \text{ si } \|x - x_0\| \leq t - t_0 \leq r_1 - t_0.$$

$$[C_3] \quad \|\Gamma_0 F'''(x)\| \leq -\frac{p'''(t)}{p'(t_0)} \text{ si } \|x - x_0\| \leq t - t_0 \leq r - t_0.$$

entonces la sucesión dada por (2.1) está bien definida y converge a x^* , solución de $F(x) = 0$. Además $x^* \in B(x_0, r - t_0)$.

Demostración. En primer lugar, observar que en las condiciones del teorema, la sucesión real (2.2) es convergente a r , ya que, por $[C_3]$, $p'''(t) \geq 0$ y, en consecuencia, podemos aplicar el Corolario 1.4.

Veamos además que estas condiciones nos permiten definir x_1 . Observar que, por el Lema de Altman,

$$\|L_F(x_0)\| \leq \|\Gamma_0 F''(x_0)\| \|\Gamma_0 F(x_0)\| \leq L_p(t_0) < 1/2 < 1.$$

Entonces, aplicando el Lema de Banach, existe $\Delta_0^{-1} = [I - L_F(x_0)]^{-1}$ y

$$\|\Delta_0^{-1}\| = \|[I - L_F(x_0)]^{-1}\| \leq \frac{1}{1 - L_p(t_0)}.$$

En consecuencia, x_1 está bien definido y además, por (2.2),

$$\begin{aligned} \|x_1 - x_0\| &\leq \left[1 + \frac{1}{2} \|\Gamma_0 F''(x_0)\| \|\Delta_0^{-1}\|\right] \|\Gamma_0 F(x_0)\| \\ &\leq -\left[1 + \frac{1}{2} \frac{L_p(t_0)}{1 - L_p(t_0)}\right] \frac{p(t_0)}{p'(t_0)} = t_1 - t_0. \end{aligned}$$

A continuación veamos que las condiciones $[C_1]$, $[C_2]$ y $[C_3]$ son ciertas si cambiamos x_1 por x_0 . En primer lugar, probemos la existencia de Γ_1 .

Observar que

$$\begin{aligned} \|I - \Gamma_0 F'(x_1)\| &= \left\| \int_{x_0}^{x_1} \Gamma_0 F''(x) dx \right\| \\ &\leq -\frac{1}{p'(t_0)} \int_{t_0}^{t_1} p''(t) dt = 1 - \frac{p'(t_1)}{p'(t_0)} < 1. \end{aligned}$$

Por tanto existe $[I - (I - \Gamma_0 F'(x_1))]^{-1} = [\Gamma_0 F'(x_1)]^{-1}$. En consecuencia existe Γ_1 y $\Gamma_1 = [\Gamma_0 F'(x_1)]^{-1} \Gamma_0$. Además

$$\|[\Gamma_0 F'(x_1)]^{-1}\| \leq \frac{1}{1 - \|I - \Gamma_0 F'(x_1)\|} \leq \frac{p'(t_0)}{p'(t_1)}.$$

Con esto, $[C_2]$ y $[C_3]$ se siguen fácilmente, ya que

$$\|\Gamma_1 F''(x)\| = \|[\Gamma_0 F'(x_1)]^{-1} \Gamma_0 F''(x)\| \leq -\frac{p'(t_0) p''(t)}{p'(t_1) p'(t_0)} = -\frac{p''(t)}{p'(t_1)},$$

y análogamente

$$\|\Gamma_1 F'''(x)\| = \|[\Gamma_0 F'(x_1)]^{-1} \Gamma_0 F'''(x)\| \leq -\frac{p'(t_0) p'''(t)}{p'(t_1) p'(t_0)} = -\frac{p'''(t)}{p'(t_1)}.$$

Para ver $[C_1]$ usamos el desarrollo de Taylor de F entorno a x_0 .

$$\begin{aligned} F(x_1) &= F(x_0) + F'(x_0)(x_1 - x_0) + \frac{1}{2} F''(x_0)(x_1 - x_0)^2 \\ &\quad + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx. \end{aligned}$$

Teniendo en cuenta (2.1),

$$F'(x_0)(x_1 - x_0) = -F(x_0) - \frac{1}{2} F''(x_0) \Gamma_0 F(x_0) \Delta_0^{-1} \Gamma_0 F(x_0)$$

y en consecuencia, de la bilinealidad de $F''(x_0)$, se obtiene

$$\begin{aligned} F(x_1) &= -\frac{1}{2} F''(x_0) \Gamma_0 F(x_0) \Delta_0^{-1} \Gamma_0 F(x_0) \\ &\quad + \frac{1}{8} F''(x_0) \left[(\Gamma_0 F(x_0))^2 + (\Delta_0^{-1} \Gamma_0 F(x_0))^2 + 2 \Gamma_0 F(x_0) \Delta_0^{-1} \Gamma_0 F(x_0) \right] \\ &\quad + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx = -\frac{1}{4} F''(x_0) \Gamma_0 F(x_0) \Delta_0^{-1} \Gamma_0 F(x_0) \\ &\quad + \frac{1}{8} F''(x_0) (\Gamma_0 F(x_0))^2 + \frac{1}{8} (\Delta_0^{-1} \Gamma_0 F(x_0))^2 + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx \\ &= \frac{1}{8} F''(x_0) \Gamma_0 F(x_0) [I - \Delta_0^{-1}] \Gamma_0 F(x_0) \\ &\quad + \frac{1}{8} F''(x_0) [\Delta_0^{-1} - I] \Gamma_0 F(x_0) \Delta_0^{-1} \Gamma_0 F(x_0) + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx \\ &= \frac{1}{8} F''(x_0) y_0 \Delta_0^{-1} \Gamma_0 F(x_0) - \frac{1}{8} F''(x_0) \Gamma_0 F(x_0) y_0 + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx \\ &= \frac{1}{8} F''(x_0) y_0 [\Delta_0^{-1} - I] \Gamma_0 F(x_0) + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx \\ &= \frac{1}{8} F''(x_0) y_0^2 + \frac{1}{2} \int_{x_0}^{x_1} F'''(x)(x_1 - x)^2 dx. \end{aligned}$$

donde $y_0 = L_F(x_0) \Delta_0^{-1} \Gamma_0 F(x_0)$.

Tenemos entonces

$$\begin{aligned}
\|\Gamma_0 F(x_1)\| &\leq \frac{1}{8} \|\Gamma_0 F''(x_0)\| \|y_0\|^2 \\
&+ \frac{1}{2} \left\| \int_0^1 \Gamma_0 F'''(x_0 + t(x_1 - x_0))(x_1 - x_0)^3 (1-t)^2 dt \right\| \\
&\leq -\frac{1}{8} \frac{p''(t_0)}{p'(t_0)} \left(\frac{L_p(t_0)}{1 - L_p(t_0)} \right)^2 \left(\frac{p(t_0)}{p'(t_0)} \right)^2 \\
&\quad - \frac{1}{2} \frac{\|x_1 - x_0\|^3}{p'(t_0)} \int_0^1 p'''(t_0 + t(t_1 - t_0))(1-t)^2 dt \\
&\leq -\frac{1}{8} \frac{p(t_0)}{p'(t_0)} \frac{\alpha^3}{(1-\alpha)^2} - \frac{1}{2} \frac{(t_1 - t_0)^3}{p'(t_0)} \int_{t_0}^{t_1} p'''(t) \frac{(t_1 - t)^2}{(t_1 - t_0)^3} dt
\end{aligned}$$

siendo $\alpha = L_p(t_0)$.

Por otra parte para la función $p(t)$ tenemos que

$$\begin{aligned}
p(t_1) &= p(t_0) + p'(t_0)(t_1 - t_0) + \frac{1}{2} p''(t_0)(t_1 - t_0)^2 \\
&\quad + \frac{1}{2} \int_{t_0}^{t_1} p'''(t)(t_1 - t)^2 dt.
\end{aligned}$$

Como $t_1 - t_0 = -\frac{p(t_0)}{p'(t_0)} \frac{2 - \alpha}{2 - 2\alpha}$.

$$\begin{aligned}
p(t_1) &= p(t_0) - \frac{1}{2} p'(t_0) \frac{p(t_0)}{p'(t_0)} \frac{2 - \alpha}{1 - \alpha} + \frac{1}{8} p''(t_0) \frac{p^2(t_0)}{(p'(t_0))^2} \left(\frac{2 - \alpha}{1 - \alpha} \right)^2 \\
&\quad + \frac{1}{2} \int_{t_0}^{t_1} p'''(t)(t_1 - t)^2 dt = \frac{1}{8} p(t_0) \frac{\alpha^3}{(1 - \alpha)^2} + \frac{1}{2} \int_{t_0}^{t_1} p'''(t)(t_1 - t)^2 dt.
\end{aligned}$$

De lo anterior se sigue que

$$\begin{aligned}
\|\Gamma_1 F(x_1)\| &= \|[\Gamma_0 F'(x_1)]^{-1} \Gamma_0 F(x_1)\| \\
&\leq -\frac{1}{8} \frac{p'(t_0)}{p'(t_1)} \frac{p(t_0)}{p'(t_0)} \frac{\alpha^3}{(1 - \alpha)^2} - \frac{1}{2} \frac{p'(t_0)}{p'(t_1)} \frac{(t_1 - t_0)^3}{p'(t_0)} \int_{t_0}^{t_1} p'''(t) \frac{(t_1 - t)^2}{(t_1 - t_0)^3} dt \\
&= -\frac{p(t_1)}{p'(t_1)}.
\end{aligned}$$

De aquí obtenemos, además, que $\|x_2 - x_1\| \leq t_2 - t_1$.

Por lo tanto, se puede probar por inducción que las condiciones $[C_1]$, $[C_2]$ y $[C_3]$ son ciertas si sustituimos x_0 por x_n para $n = 1, 2, \dots$ con lo que la sucesión $\{x_n\}$ está bien definida y además $\|x_{n+1} - x_n\| \leq t_{n+1} - t_n$, es decir, $\{t_n\}$ es mayorizante

de $\{x_n\}$. De aquí se sigue que, para $p \in \mathbb{N}$, $\|x_{n+p} - x_n\| \leq t_{n+p} - t_n$. Luego haciendo $p \rightarrow \infty$ y $n = 0$ tenemos que $\|x^* - x_0\| \leq r - t_0$. ■

Nota. La ecuación $F(x) = 0$ puede tener otras raíces en la bola $B(x_0, r - t_0)$ a pesar de que la ecuación mayorizante $p(t) = 0$ tenga una única solución en el intervalo $[a, b]$. Una condición suficiente para que x^* sea la única raíz de $F(x) = 0$ en la bola anterior es que $p(b) \leq b$. Este es un resultado general sobre funciones mayorizantes que puede verse en el libro de Kantorovich [66].

Al igual que vimos en el estudio del método de Newton, suele ser difícil el uso directo del Teorema 2.1, ya que contiene una función desconocida p . Sin embargo, si el operador F verifica unas condiciones del tipo de las de Kantorovich, podemos determinar esta función y utilizarla para probar la convergencia del método (2.1), dar estimaciones del error, etc. Al tratarse de un método de tercer orden tenemos que exigir condiciones sobre la derivada tercera del operador, como hicimos en el Teorema 2.1 y como hace Altman [4], o condiciones algo más débiles como que la derivada segunda verifique una condición tipo Lipschitz, situación analizada por otros autores [109], [38].

En lo que sigue supondremos que $F : \Omega \subseteq X \rightarrow Y$ es un operador dos veces diferenciable Fréchet en un conjunto convexo abierto $\Omega_0 \subseteq \Omega$ y que se cumplen además las siguientes condiciones:

[C₁] Existe un punto $x_0 \in \Omega_0$ en el que el operador inverso $\Gamma_0 = [F'(x_0)]^{-1}$ está definido.

[C₂] $\|\Gamma_0(F''(x) - F''(y))\| \leq k\|x - y\|$, $x, y \in \Omega_0$, $k \geq 0$.

$$a = \|\Gamma_0 F(x_0)\|, \quad b = \|\Gamma_0 F''(x_0)\|.$$

[C₃] La ecuación $p(t) = 0$ tiene una raíz negativa y dos positivas $r_1 \leq r_2$, donde p es el polinomio

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a.$$

Como ya vimos en la Sección IV.3, en esta situación se tiene que la sucesión obtenida usando el método de Newton para resolver la ecuación $p(t) = 0$ es mayorizante de la sucesión de Newton correspondiente a aplicar el método de Newton a $F(x) = 0$. Nos podemos preguntar ahora si ocurre lo mismo al utilizar la aceleración convexa del método de Newton para resolver las ecuaciones $p(t) = 0$ y $F(x) = 0$.

En estas condiciones sabemos que la sucesión $\{t_n\}$ construida por el método de la aceleración convexa aplicado al polinomio p , y empezando en $t_0 = 0$, es convergente a r_1 . La respuesta a estas cuestiones la encontraremos en el lema siguiente.

Lema 2.2. *En las condiciones anteriores, la sucesión $\{t_n\}$ es mayorizante de $\{x_n\}$, es decir,*

$$\|x_{n+1} - x_n\| \leq t_{n+1} - t_n, \quad n \geq 0.$$

Demostración. Para cada $n \geq 0$, vamos a probar lo siguiente

$$[\text{I}_n] \text{ Existe } \Gamma_n = [F'(x_n)]^{-1}.$$

$$[\text{II}_n] \|\Gamma_0 F''(x_n)\| \leq -\frac{p''(t_n)}{p'(t_0)}.$$

$$[\text{III}_n] \|\Gamma_n F'(x_0)\| \leq \frac{p'(t_0)}{p'(t_n)}.$$

$$[\text{IV}_n] \|\Gamma_0 F(x_n)\| \leq -\frac{p(t_n)}{p'(t_0)}.$$

Lo haremos por inducción. Para $n = 0$ se sigue inmediatamente de las hipótesis. Supongamos que $[\text{I}_n] \dots [\text{IV}_n]$ son ciertas. Probemos $[\text{I}_{n+1}] \dots [\text{IV}_{n+1}]$.

En primer lugar, observar que, en este caso, el operador $L_{F(x_n)}$ está definido y además

$$\begin{aligned} \|L_{F(x_n)}\| &\leq \|\Gamma_n F'(x_0)\| \|\Gamma_0 F''(x_n)\| \|\Gamma_n F'(x_0)\| \|\Gamma_0 F(x_n)\| \\ &\leq \frac{p''(t_n)p(t_n)}{p'(t_n)^2} = L_p(t_n) < \frac{1}{2}. \end{aligned}$$

por el Lema de Altman 1.2.3.

En consecuencia, por el Lema de Banach, $\Delta_n^{-1} = [I - L_{F(x_n)}]^{-1}$ también está definido y además

$$\|\Delta_n^{-1}\| \leq \frac{1}{1 - L_p(t_n)}.$$

Así, tenemos garantizada la existencia del siguiente término de la sucesión, x_{n+1} . Veamos que también está definido Γ_{n+1} .

Observar que

$$\begin{aligned} I - \Gamma_0 F'(x_{n+1}) &= -\int_{x_0}^{x_{n+1}} \Gamma_0 F''(x) dx \\ &= -\int_{x_0}^{x_{n+1}} \Gamma_0 [F''(x) - F''(x_0)] dx - \Gamma_0 F''(x_0)(x_{n+1} - x_0). \end{aligned}$$

Así, por [C₁], [C₂] y [C₃].

$$\begin{aligned} \|I - \Gamma_0 F'(x_{n+1})\| &\leq \left\| \int_0^1 \Gamma_0 [F''(x_0 + t(x_{n+1} - x_0)) - F''(x_0)](x_{n+1} - x_0) dt \right\| \\ &+ \|\Gamma_0 F''(x_0)\| \|x_{n+1} - x_0\| \leq \frac{k}{2} \|x_{n+1} - x_0\|^2 + b \|x_{n+1} - x_0\| \\ &\leq \frac{k}{2} t_{n+1}^2 + b t_{n+1} = p'(t_{n+1}) + 1 < 1. \end{aligned}$$

Entonces, por el Lema de Banach, existe $[\Gamma_0 F'(x_{n+1})]^{-1}$ y además

$$\|[\Gamma_0 F'(x_{n+1})]^{-1}\| \leq \frac{1}{1 - \|I - \Gamma_0 F'(x_{n+1})\|} \leq \frac{1}{p'(t_{n+1})}. \quad (2.3)$$

De aquí se sigue que existe $\Gamma_{n+1} = [\Gamma_0 F'(x_{n+1})]^{-1} \Gamma_0$.

Además, por (2.3),

$$\|\Gamma_{n+1} F'(x_0)\| = \|[\Gamma_0 F'(x_{n+1})]^{-1}\| \leq \frac{1}{p'(t_{n+1})} = \frac{p'(t_0)}{p'(t_{n+1})},$$

de donde se sigue [III_{n+1}].

Por otra parte,

$$\begin{aligned} \|\Gamma_0 F''(x_{n+1})\| &\leq \|\Gamma_0 [F''(x_{n+1}) - F''(x_0)]\| + \|\Gamma_0 F''(x_0)\| \\ &\leq k \|x_{n+1} - x_0\| + b \leq k t_{n+1} + b = p''(t_{n+1}) = -\frac{p''(t_{n+1})}{p'(t_0)}. \end{aligned}$$

con lo que probamos [II_{n+1}].

Por último, veamos [IV_{n+1}]. Utilizando la técnica de Altman y teniendo en cuenta el desarrollo de Taylor de F en torno a x_n ,

$$\begin{aligned} F(x_{n+1}) &= F(x_n) + F'(x_n)(x_{n+1} - x_n) + \frac{1}{2} F''(x_n)(x_{n+1} - x_n)^2 \\ &+ \int_{x_n}^{x_{n+1}} [F''(x) - F''(x_n)](x_{n+1} - x) dx. \end{aligned}$$

Como $x_{n+1} - x_n = -[I + \frac{1}{2} L_F(x_n) \Delta_n^{-1}] \Gamma_n F(x_n)$, obtenemos

$$\begin{aligned} F'(x_n)(x_{n+1} - x_n) &= -F(x_n) - \frac{1}{2} F'(x_n) L_F(x_n) \Delta_n^{-1} \Gamma_n F(x_n) \\ &= -F(x_n) - \frac{1}{2} F''(x_n) \Gamma_n F(x_n) \Delta_n^{-1} \Gamma_n F(x_n). \end{aligned}$$

Por tanto,

$$F(x_{n+1}) = -\frac{1}{2} F''(x_n) \Gamma_n F(x_n) \Delta_n^{-1} \Gamma_n F(x_n) + \frac{1}{2} F''(x_n)(x_{n+1} - x_n)^2$$

$$\begin{aligned}
& + \int_{x_n}^{x_{n+1}} [F''(x) - F''(x_n)](x_{n+1} - x) dx = -\frac{1}{2}F''(x_n)\Gamma_n F(x_n)\Delta_n^{-1}\Gamma_n F(x_n) \\
& \quad + \frac{1}{2}F''(x_n)[\Gamma_n F(x_n)]^2 + \frac{1}{8}F''(x_n)[L_F(x_n)\Delta_n^{-1}\Gamma_n F(x_n)]^2 \\
& \quad + \frac{1}{2}F''(x_n)\Gamma_n F(x_n)L_F(x_n)\Delta_n^{-1}\Gamma_n F(x_n) \\
& \quad + \int_{x_n}^{x_{n+1}} [F''(x) - F''(x_n)](x_{n+1} - x) dx.
\end{aligned}$$

Como $\Delta_n^{-1} = I + L_F(x_n)\Delta_n^{-1}$, llamando $y_n = L_F(x_n)\Delta_n^{-1}\Gamma_n F(x_n)$, se sigue que

$$\begin{aligned}
F(x_{n+1}) & = \frac{1}{8}F''(x_n)[L_F(x_n)\Delta_n^{-1}\Gamma_n F(x_n)]^2 \\
& \quad + \int_{x_n}^{x_{n+1}} [F''(x) - F''(x_n)](x_{n+1} - x) dx.
\end{aligned}$$

Así, usando que

$$\|y_n\| \leq \|L_F(x_n)\| \|\Delta_n^{-1}\| \|\Gamma_n F(x_n)\| \leq -\frac{a_n p(t_n)}{(1 - a_n)p'(t_n)},$$

donde hemos denotado $a_n = L_p(t_n)$, se tiene que

$$\begin{aligned}
\|\Gamma_0 F(x_{n+1})\| & \leq \frac{1}{8}p''(t_n) \frac{a_n^2 p(t_n)^2}{(1 - a_n)^2 p'(t_n)^2} \\
& \quad + \left\| \int_0^1 \Gamma_0 [F''(x_n + t(x_{n+1} - x_n)) - F''(x_n)](x_{n+1} - x_n)^2 (1 - t) dt \right\| \\
& \leq \frac{1}{8} \frac{a_n^3 p(t_n)}{(1 - a_n)^2} + \frac{k}{6} \|x_{n+1} - x_n\|^3 \leq \frac{1}{8} \frac{a_n^3 p(t_n)}{(1 - a_n)^2} + \frac{k}{6} (t_{n+1} - t_n)^3.
\end{aligned}$$

Procediendo de forma análoga con el polinomio p tenemos que

$$p(t_{n+1}) = \frac{1}{8} \frac{a_n^3 p(t_n)}{(1 - a_n)^2} + \frac{k}{6} (t_{n+1} - t_n)^3$$

y, en consecuencia, $\|\Gamma_0 F(x_{n+1})\| \leq p(t_{n+1})$, de donde se sigue $[\mathbf{IV}_{n+1}]$. Así, queda probada la inducción. ■

Lema 2.3. *En las condiciones anteriores se tiene que*

$$\begin{aligned}
\frac{\|x^* - x_{n+1}\|}{r_1 - t_{n+1}} & \leq \left(\frac{\|x^* - x_n\|}{r_1 - t_n} \right)^3, \quad n \geq 0, \quad k > 0; \\
\frac{\|x^* - x_{n+1}\|}{r_1 - t_{n+1}} & \leq \left(\frac{\|x^* - x_n\|}{r_1 - t_n} \right)^4, \quad n \geq 0, \quad k = 0.
\end{aligned}$$

Demostración. Como $\Delta_n^{-1} = I + L_F(x_n)\Delta_n^{-1}$, podemos escribir

$$I + \frac{1}{2}L_F(x_n)\Delta_n^{-1} = \frac{1}{2}(I + \Delta_n^{-1}) = \frac{1}{2}\Delta_n^{-1}(\Delta_n + I) = \Delta_n^{-1}\left(I - \frac{1}{2}L_F(x_n)\right).$$

Por tanto,

$$\begin{aligned} x^* - x_{n+1} &= x^* - x_n + \Delta_n^{-1}\left(I - \frac{1}{2}L_F(x_n)\right)\Gamma_n F(x_n) \\ &= -\Delta_n^{-1}\Gamma_n\left[F(x^*) - F(x_n) - F'(x_n)(x^* - x_n) - \frac{1}{2}F''(x_n)(x^* - x_n)^2\right] \\ &\quad + (x^* - x_n) - \frac{1}{2}\Delta_n^{-1}L_F(x_n)\Gamma_n F(x_n) \\ &\quad - \Delta_n^{-1}(x^* - x_n) - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)(x^* - x_n)^2 \\ &= -\Delta_n^{-1}\Gamma_n\int_{x_n}^{x^*}[F''(x) - F''(x_n)](x^* - x)dx + [I - \Delta_n^{-1}](x^* - x_n) \\ &\quad - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\left[(\Gamma_n F(x_n))^2 + (x^* - x_n)^2\right]. \end{aligned}$$

Además $I - \Delta_n^{-1} = -\Delta_n^{-1}L_F(x_n)$, y como $F''(x_n)$ es simétrica, se sigue que

$$\begin{aligned} &[I - \Delta_n^{-1}](x^* - x_n) - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\left[(\Gamma_n F(x_n))^2 + (x^* - x_n)^2\right] \\ &= -\Delta_n^{-1}\Gamma_n F''(x_n)\left[\Gamma_n F(x_n)(x^* - x_n) + \frac{1}{2}[(\Gamma_n F(x_n))^2 + (x^* - x_n)^2]\right] \\ &= \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\Gamma_n\int_{x_n}^{x^*}F''(x)(x^* - x)dx\left[\Gamma_n F(x_n) + (x^* - x_n)\right] \\ &= \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\left[\Gamma_n\int_{x_n}^{x^*}F''(x)(x^* - x)dx\right]^2. \end{aligned}$$

En definitiva,

$$\begin{aligned} x^* - x_{n+1} &= -\Delta_n^{-1}\Gamma_n\int_{x_n}^{x^*}[F''(x) - F''(x_n)](x^* - x)dx \\ &\quad - \frac{1}{2}\Delta_n^{-1}\Gamma_n F''(x_n)\left[\Gamma_n\int_{x_n}^{x^*}F''(x)(x^* - x)dx\right]^2. \end{aligned} \tag{2.4}$$

Entonces, teniendo en cuenta los resultados del Lema 2.2 y las condiciones iniciales,

$$\begin{aligned} \left\|\int_{x_n}^{x^*}\Gamma_0 F''(x)(x^* - x)dx\right\| &\leq \|x^* - x_n\|^2 \int_0^1 p''(t_n + t(r_1 - t_n))(1 - t)dt \\ &= \left[\frac{\|x^* - x_n\|}{r_1 - t_n}\right]^2 \int_0^1 p''(t_n + t(r_1 - t_n))(r_1 - t_n)^2(1 - t)dt \end{aligned}$$

$$= \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^2 \int_{t_n}^{r_1} p''(z)(r_1 - z) dz.$$

por lo que, tomando normas en (2.4), obtenemos

$$\begin{aligned} \|x^* - x_{n+1}\| &= \|\Delta_n^{-1}\| \|\Gamma_n F'(x_0)\| \left\| \int_{x_n}^{x^*} \Gamma_0 [F''(x) - F''(x_n)](x^* - x) dx \right\| \\ &+ \frac{1}{2} \|\Delta_n^{-1}\| \|\Gamma_n F'(x_0)\| \|\Gamma_0 F''(x_n)\| \left\| \Gamma_n \int_{x_n}^{x^*} F''(x)(x^* - x) dx \right\|^2 \\ &\leq -\frac{k(r_1 - t_n)^3}{6p'(t_n)(1 - L_p(t_n))} \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^3 \\ &- \frac{p''(t_n)}{2p'(t_n)^3(1 - L_p(t_n))} \left[\int_{t_n}^{r_1} p''(z)(r_1 - z) dz \right]^2 \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^4. \end{aligned} \quad (2.5)$$

Si $k > 0$, como $\|x^* - x_n\| \leq r_1 - t_n$.

$$\begin{aligned} \|x^* - x_{n+1}\| &\leq -\left[\frac{k(r_1 - t_n)^3}{6p'(t_n)(1 - L_p(t_n))} \right. \\ &+ \left. \frac{p''(t_n)}{2p'(t_n)^3(1 - L_p(t_n))} \left[\int_{t_n}^{r_1} p''(z)(r_1 - z) dz \right]^2 \right] \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^3 \\ &= (r_1 - t_{n+1}) \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^3 \end{aligned}$$

sin más que repetir para t_n las operaciones hechas con x_n .

Ahora bien, si $k = 0$, de (2.5) se deduce que

$$\begin{aligned} \|x^* - x_{n+1}\| &\leq -\frac{p''(t_n)}{2p'(t_n)^3(1 - L_p(t_n))} \left[\int_{t_n}^{r_1} p''(z)(r_1 - z) dz \right]^2 \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^4 \\ &= (r_1 - t_{n+1}) \left[\frac{\|x^* - x_n\|}{r_1 - t_n} \right]^4. \quad \blacksquare \end{aligned}$$

El hecho de que $\|x^* - x_n\| \leq r - t_n$ nos permite encontrar estimaciones del error para la sucesión $\{x_n\}$ en términos de la sucesión real $\{t_n\}$. Si $k = 0$, la sucesión $\{t_n\}$ se obtiene a partir de un polinomio de segundo grado, y mediante un procedimiento similar al seguido por Ostrowski [82] podemos dar la siguiente expresión del error para el método de la aceleración convexa aplicado a un polinomio de segundo grado.

Lema 2.4. Sea p el polinomio dado por

$$p(t) = \frac{b}{2}t^2 - t + a.$$

Si p tiene dos raíces positivas $r_1 \leq r_2$ y $\{t_n\}$ es la sucesión definida por (2.2), entonces

$$r_1 - t_n = \frac{(r_2 - r_1)\theta^{4^n}}{1 - \theta^{4^n}}, \quad n \geq 0, \quad \text{si } r_1 < r_2, \quad \theta = \frac{r_1}{r_2}.$$

$$r_1 - t_n = \frac{r_1}{4^n}, \quad n \geq 0, \quad \text{si } r_1 = r_2.$$

Demostración. Llamamos $a_n = r_1 - t_n$, $b_n = r_2 - t_n$, $n \geq 0$. Entonces

$$p(t_n) = \frac{b}{2}a_nb_n, \quad p'(t_n) = -\frac{b}{2}(a_n + b_n).$$

Tenemos, por (2.4),

$$\begin{aligned} a_{n+1} = r_1 - t_{n+1} &= a_n - \frac{1}{2} \left[1 + \frac{(a_n + b_n)^2}{a_n^2 + b_n^2} \right] \frac{a_nb_n}{a_n + b_n} \\ &= \frac{a_n^4}{(a_n + b_n)(a_n^2 + b_n^2)}. \end{aligned} \quad (2.6)$$

y análogamente

$$b_{n+1} = \frac{b_n^4}{(a_n + b_n)(a_n^2 + b_n^2)}.$$

Si $r_1 < r_2$, llamamos $\theta = r_1/r_2 < 1$ y obtenemos

$$\frac{a_n}{b_n} = \left[\frac{a_n}{b_n} \right]^4 = \dots = \left[\frac{a_0}{b_0} \right]^{4^n} = \theta^{4^n}.$$

Teniendo en cuenta que $b_n = r_2 - t_n = (r_2 - r_1) + a_n$, se sigue la primera parte.

Ahora bien, si $r_1 = r_2$, entonces $a_n = b_n$ y, por (2.6), la segunda parte también es cierta. ■

En el caso $k > 0$ la sucesión $\{t_n\}$ se obtiene a partir de un polinomio de tercer grado. Encontrar acotaciones cómodas del error por el método de Ostrowski resulta complicado. En el siguiente lema hemos desarrollado una nueva técnica que nos permitirá encontrar estimaciones del error cometido al aplicar el método de la aceleración convexa a un polinomio de tercer grado, en función de las raíces de dicho polinomio.

Lema 2.5. Sea p el polinomio dado por

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a, \quad k > 0.$$

supongamos que p tiene dos raíces positivas $r_1 \leq r_2$ y una raíz negativa. $-r_0$. Sea $\{t_n\}$ la aceleración convexa del método de Newton definida en (2.2). Si $r_1 < r_2$, entonces

$$r_1 - t_n \sim \frac{(r_2 - r_1)\theta^{3^n}}{\sqrt{\lambda} - \theta^{3^n}}, \quad n \geq 0.$$

donde

$$\lambda = \frac{r_2 - r_1}{r_0 + r_1} < 1, \quad \theta = \sqrt{\lambda} \frac{r_1}{r_2} < 1.$$

Si $r_1 = r_2$, entonces

$$\frac{r_1}{4^n} \leq r_1 - t_n \leq \frac{r_1}{3^n}.$$

Demostración. Podemos escribir el polinomio p de la siguiente forma

$$p(t) = \frac{k}{6}(r_1 - t)(r_2 - t)(r_0 + t).$$

Comparando los coeficientes de t^2 se deduce que $r_1 + r_2 \leq r_0$.

Denotamos $a_n = r_1 - t_n$, $b_n = r_2 - t_n$ y $c_n = r_0 + t_n$. Entonces, $p(t_n) = \frac{k}{6}a_nb_nc_n$, $p'(t_n) = \frac{k}{6}(a_nb_n - a_nc_n - b_nc_n)$ y $p''(t_n) = \frac{k}{6}(2c_n - 2a_n - 2b_n)$. De (2.2) se sigue que

$$\begin{aligned} a_{n+1} &= r_1 - t_{n+1} = r_1 - t_n + \left(1 + \frac{L_p(t_n)}{2(1 - L_p(t_n))}\right) \frac{p(t_n)}{p'(t_n)} \\ &= a_n \left[1 + \left(1 + \frac{L_p(t_n)}{2(1 - L_p(t_n))}\right) \frac{k}{6} \frac{b_nc_n}{p'(t_n)}\right] \\ &= a_n^2 \left[b_n - c_n + \frac{k}{6} \frac{b_n^2 c_n^2 p''(t_n)}{2p'(t_n)^2 (1 - L_p(t_n))}\right] \\ &= \left(\frac{k}{6}\right)^3 \frac{a_n^3}{p'(t_n)^3 (1 - L_p(t_n))} [a_n(b_n - c_n)(b_n^2 + c_n^2) - b_n^2 c_n^2]. \end{aligned}$$

De forma parecida, se obtiene

$$b_{n+1} = \left(\frac{k}{6}\right)^3 \frac{b_n^3}{p'(t_n)^3 (1 - L_p(t_n))} [b_n(a_n - c_n)(a_n^2 + c_n^2) - a_n^2 c_n^2].$$

En consecuencia,

$$\frac{a_{n+1}}{b_{n+1}} = \frac{a_n^3}{b_n^3} H(t_n).$$

donde

$$H(t_n) = \frac{b_n^2 c_n^2 + a_n(c_n - b_n)(b_n^2 + c_n^2)}{a_n^2 c_n^2 + b_n(c_n - a_n)(a_n^2 + c_n^2)}.$$

Notar que

$$H(t_n) = \frac{b_n^3 a_{n+1}}{a_n^3 b_{n+1}} = \frac{(r_1 - G(t_n))(r_2 - t_n)^3}{(r_2 - G(t_n))(r_1 - t_n)^3}.$$

con G definida en (1.3). Como $G(r_1) = r_1$, $G'(r_1) = G''(r_1) = 0$, se sigue, para t suficientemente cercano a r_1

$$H(t) \sim \frac{G'''(r_1)}{6}(r_2 - r_1)^2 = -\frac{f'''(r_1)}{6f'(r_1)}(r_2 - r_1)^2 = \frac{r_2 - r_1}{r_0 + r_1}.$$

Puesto que $t_n \rightarrow r_1$ cuando $n \rightarrow \infty$, se obtiene

$$\frac{a_n}{b_n} \sim \left(\frac{a_{n-1}}{b_{n-1}}\right)^3 \lambda \sim \dots \sim \left(\frac{a_0}{b_0}\right)^{3^n} \lambda^{(3^n - 1)/2} = \left(\sqrt{\lambda} \frac{r_1}{r_2}\right)^{3^n} \frac{1}{\sqrt{\lambda}}.$$

Luego,

$$r_1 - t_n \sim (r_2 - r_1 + r_1 - t_n) \frac{\theta^{3^n}}{\sqrt{\lambda}}.$$

de donde se sigue la primera parte.

Si $r_1 = r_2$, escribimos $p(t) = \frac{k}{5}(r_1 - t)^2(r_0 + t)$, donde $2r_1 \leq r_0$. Con la misma técnica y notaciones, obtenemos

$$a_{n+1} = a_n \frac{a_n^3 - c_n a_n^2 - c_n^3}{(a_n - 2c_n)(a_n^2 + 2c_n^2)} = a_n \frac{(a_n/c_n)^3 - (a_n/c_n)^2 - 1}{((a_n/c_n) - 2)((a_n/c_n)^2 + 2)}.$$

Teniendo en cuenta que $\frac{a_n}{c_n} \leq \frac{1}{2}$ y que la función

$$f(x) = \frac{1 + x^2 - x^3}{(x^2 + 2)(2 - x)}$$

es creciente en $[0, 1/2]$, deducimos

$$\frac{a_n}{4} \leq a_{n+1} \leq \frac{a_n}{3}$$

y, por tanto, se prueba el teorema. ■

Observar que las estimaciones del Lema 2.5 son igualdades cuando $n = 0$, esto es, para $t = t_0$, el valor más lejano a la raíz de los que forman la sucesión $\{t_n\}$, y donde cabía esperar la estimación fuese la peor de todas las dadas.

Con la ayuda de estos lemas, podemos demostrar el siguiente resultado que estudia la convergencia y unicidad de solución de la sucesión definida por (2.2).

Teorema 2.6. Supongamos que se cumplen $[C_1]$, $[C_2]$ y $[C_3]$ y que, además,

$$\overline{B} = \overline{B(x_1, r_1 - t_1)} \subseteq \Omega_0.$$

Entonces, la sucesión (2.2) está bien definida, está contenida en B y converge a x^* , solución de $F(x) = 0$. Si $r_1 < r_2$ tenemos un método de tercer orden para $k > 0$ y de cuarto orden para $k = 0$, donde k es la constante introducida en la condición $[C_2]$. Si $r_1 = r_2$, la convergencia del método es lineal en ambos casos. También se tiene que la solución es única en

$$\dot{B} = \begin{cases} B(x_0, r_2) \cap \Omega_0 & \text{si } r_1 < r_2, \\ \frac{B(x_0, r_2)}{B(x_0, r_2) \cap \Omega_0} & \text{si } r_1 = r_2. \end{cases}$$

Además, podemos dar las siguientes estimaciones del error en función de las raíces del polinomio p :

- Si $k = 0$,

$$\|x^* - x_n\| \leq r_1 - t_n = \frac{(r_2 - r_1)\theta^{4^n}}{1 - \theta^{4^n}}, \quad n \geq 0,$$

$$\text{si } r_1 < r_2, \quad \theta = \frac{r_1}{r_2} < 1,$$

$$\|x^* - x_n\| \leq r_1 - t_n = \frac{r_1}{4^n}, \quad n \geq 0, \quad \text{si } r_1 = r_2.$$

- Si $k > 0$,

$$\|x^* - x_n\| \leq r_1 - t_n \sim \frac{(r_2 - r_1)\theta^{3^n}}{\sqrt{\lambda} - \theta^{3^n}}, \quad n \geq 0,$$

$$\text{si } r_1 < r_2, \quad \lambda = \frac{r_2 - r_1}{r_0 + r_1} < 1, \quad \theta = \sqrt{\lambda} \frac{r_1}{r_2} < 1,$$

$$\|x^* - x_n\| \leq r_1 - t_n, \quad \text{con } \frac{r_1}{4^n} \leq r_1 - t_n \leq \frac{r_1}{3^n}, \quad \text{si } r_1 = r_2.$$

Demostración. La primera parte se sigue inmediatamente del Lema 2.2, de los resultados conocidos sobre sucesiones mayorizantes y del hecho de que la sucesión real $\{t_n\}$ es convergente. Que el método (2.1) tiene convergencia cúbica para $k > 0$ se deduce del Lema 2.3 y del Teorema 1.5.

Para probar que (2.1) es un método de orden cuatro para $k = 0$, usamos los Lemmas 2.2, 2.3 y el Teorema 1.6, y deducimos, para $1 \leq \gamma < 4$,

$$\frac{\|x^* - x_{n+1}\|}{\|x^* - x_n\|^\gamma} \leq \frac{r_1 - t_{n+1}}{(r_1 - t_n)^\gamma} = \frac{1}{(r_1 - r_2)^{\gamma-1}} \frac{(1 - \theta^{4^n})^\gamma}{1 - \theta^{4^{n+1}}} \theta^{(4-\gamma)4^n} \rightarrow 0.$$

Sin embargo.

$$\lim_{n \rightarrow \infty} \frac{\|x^* - x_{n+1}\|}{\|x^* - x_n\|^4} \leq \frac{1}{(r_1 - r_2)^3} < \infty.$$

Por otra parte, teniendo en cuenta el Teorema 1.6.

$$\lim_{n \rightarrow \infty} \frac{\|x^* - x_{n+1}\|}{\|x^* - x_n\|^4} > 0.$$

Para ver la unicidad, supongamos que x' es otra solución de $F(x) = 0$ en \tilde{B} . Entonces, reemplazando x^* y r_1 por x' y r_2 en la demostración del Lema 2.4 (lo único que se necesita es que sean raíces de $F(x) = 0$ y $p(t) = 0$ respectivamente), tenemos

$$\frac{\|x' - x_{n+1}\|}{r_2 - t_{n+1}} \leq \left[\frac{\|x' - x_n\|}{r_2 - t_n} \right]^3 \leq \dots \leq \left[\frac{\|x' - x_0\|}{r_2 - t_0} \right]^{3^{n+1}} = \rho^{3^{n+1}}.$$

con $\rho = \frac{\|x' - x_0\|}{r_2 - t_0}$.

Notar que, si $r_1 < r_2$, entonces $\rho < 1$. Por otra parte, si $r_1 = r_2$, entonces $\rho \leq 1$ y además $r_2 - t_{n+1} = r_1 - t_{n+1} \rightarrow 0$ cuando $n \rightarrow \infty$.

En ambos casos tenemos

$$\|x' - x_{n+1}\| \leq (r_2 - t_{n+1})\rho^{3^{n+1}} \rightarrow 0.$$

Así, $x' = \lim_{n \rightarrow \infty} x_n = x^*$, con lo que tenemos probada la unicidad.

La estimación el error dada se obtiene de los lemas 2.4 y 2.5, con lo que se concluye el resultado. ■

El siguiente estudio consistirá en obtener unas estimaciones intermedias entre la sucesión en el espacio de Banach y su mayorizante, empleando la técnica que Gragg y Tapia [49] usaron para el método de Newton, desarrollada en el Capítulo 3, y que también ha sido usada en métodos de tercer orden por otros autores [109]. En nuestro caso, la sucesión $\{t_n\}$, mayorizante de (2.1), se obtiene aplicando el método de la aceleración convexa a la ecuación $p(t) = 0$, con

$$p(t) = \frac{k}{6}t^3 + \frac{b}{2}t^2 - t + a. \quad (2.7)$$

Teorema 2.7. Sean τ_n^* y σ_n^* respectivamente la menor raíz positiva y la única raíz positiva de

$$\begin{aligned}\phi_n(t) &= k_n t^3 - t + \delta_n \\ \psi_n(t) &= k_n t^3 + t - \delta_n.\end{aligned}$$

donde

$$k_n = \frac{r_1 - t_{n+1}}{(r_1 - t_n)^3}, \quad \delta_n = \|x_{n+1} - x_n\| > 0.$$

Entonces se tienen las siguientes estimaciones a posteriori:

$$\begin{aligned}\sigma_n^* &\leq \|x^* - x_n\| \leq \tau_n^* \leq r_1 - t_n. \\ \|x^* - x_{n+1}\| &\leq \tau_n^* - \delta_n.\end{aligned}$$

Demostración. Observar que, por el Lema 2.3,

$$\begin{aligned}\|x^* - x_n\| - \|x_n - x_{n+1}\| &\leq \left| \|x^* - x_n\| - \|x_n - x_{n+1}\| \right| \\ &\leq \|x^* - x_{n+1}\| \leq k_n \|x^* - x_n\|^3.\end{aligned}$$

En consecuencia, $\phi_n(\|x^* - x_n\|) \geq 0$.

Por otra parte,

$$\phi_n(r_1 - t_n) = k_n(r_1 - t_n)^3 - (r_1 - t_n) + \delta_n = \delta_n - (t_{n+1} - t_n) < 0.$$

Como $\phi_n(0) > 0$, y $\|x^* - x_n\| \leq r_1 - t_n$, se sigue que

$$\|x^* - x_n\| \leq \tau_n^* \leq r_1 - t_n.$$

Para obtener las cotas inferiores, usando de nuevo el Lema 2.3, observar que

$$\delta_n - \|x^* - x_n\| \leq \|x^* - x_{n+1}\| \leq k_n \|x^* - x_n\|^3.$$

por tanto, $\psi_n(\|x^* - x_n\|) \geq 0$.

Como ψ_n es una función creciente y $\psi_n(0) < 0$, se sigue que ψ_n tiene una única raíz, que además es positiva: σ_n^* . De aquí,

$$\|x^* - x_n\| \geq \sigma_n^*.$$

Por último, teniendo en cuenta que $\phi_n(\tau_n^*) = k_n(\tau_n^*)^3 - \tau_n^* + \delta_n = 0$, se obtiene

$$\|x^* - x_{n+1}\| \leq k_n \|x^* - x_n\|^3 \leq k_n (\tau_n^*)^3 = \tau_n^* - \delta_n. \quad \blacksquare$$

Corolario 2.8. *En las condiciones del teorema anterior, se tiene*

$$0.89\delta_n \leq \|x^* - x_n\| \leq 1.5\delta_n.$$

$$\|x^* - x_{n+1}\| \leq 0.5\delta_n.$$

Demostración. Manteniendo la notación del teorema anterior, tenemos que ϕ_n tiene un mínimo relativo en

$$\tau_n = \frac{1}{\sqrt{3k_n}}.$$

Además

$$0 > \phi_n(\tau_n) = \delta_n - \frac{2}{3}\tau_n.$$

Por consiguiente,

$$k_n\delta_n^2 = \frac{\delta_n^2}{3\tau_n^2} < \frac{4}{27}.$$

y $\phi_n(3\delta_n/2) < 0$, de donde se deduce que

$$\|x^* - x_{n+1}\| \leq \tau_n^* < \frac{3\delta_n}{2}$$

y

$$\|x^* - x_{n+1}\| \leq \tau_n^* - \delta_n < \frac{\delta_n}{2}.$$

Por otra parte, sea $\bar{\sigma}_n^*$ la raíz positiva de

$$\bar{w}_n(t) = \frac{4}{27\delta_n^2}t^3 + t - \delta_n.$$

Como $w_n(t) < \bar{w}_n(t)$ para $t > 0$, $\bar{\sigma}_n^* < \sigma_n^*$.

Por último, basta con verificar que $\bar{w}_n(0.89\delta_n) < 0$. ■

5.3 Otro estudio de la convergencia para la aceleración convexa

La aceleración convexa del método de Newton para resolver una ecuación escalar $p(t) = 0$ viene dada por la expresión

$$t_{n+1} = t_n - \frac{1}{2} \left[1 + \frac{1}{1 - L_p(t_n)} \right] \frac{p(t_n)}{p'(t_n)}. \quad (3.1)$$

donde $L_p(t_n) = \frac{p(t_n)p''(t_n)}{p'(t_n)^2}$.

En la Sección 7.2 ya hemos estudiado bajo qué condiciones este método converge a t^* , solución de $p(t) = 0$, así como la estimación del error obtenido. También sabemos que este es un método de orden tres en general, pero que aplicado a ecuaciones polinómicas de segundo grado, la convergencia es de orden cuatro.

Como en otros métodos, la técnica empleada para estudiar la aceleración convexa del método de Newton en espacios de Banach consiste en mayorizar la sucesión en el espacio de Banach por una sucesión real. La sucesión real que se obtiene de manera natural en este caso se construye a partir de un polinomio de tercer grado. Como consecuencia de los comentarios anteriores, puede resultar interesante buscar condiciones para que la sucesión mayorizante se obtenga a partir de un polinomio de segundo grado.

Consideramos ahora la ecuación $F(x) = 0$, siendo F un operador entre dos espacios de Banach con derivadas de Fréchet de segundo orden continuas en un abierto convexo Ω . Siguiendo la técnica de dividir un proceso iterativo en dos, empleada por Dong Chen en [37] y [38], podemos escribir formalmente el método de la aceleración convexa para resolver dicha ecuación de la siguiente forma

$$\begin{cases} y_n &= x_n - \Gamma_n F(x_n) \\ x_{n+1} &= y_n + \frac{1}{2} L_F(x_n) H(x_n) (y_n - x_n), \end{cases} \quad (3.2)$$

donde $\Gamma_n = [F'(x_n)]^{-1}$, $L_F(x_n) = \Gamma_n F''(x_n) \Gamma_n F(x_n)$ y $H(x_n) = [I - L_F(x_n)]^{-1}$.

Supondremos a partir de ahora que se cumplen las siguientes condiciones:

(I) Existe un punto $x_0 \in \Omega$ donde el operador $\Gamma_0 = [F'(x_0)]^{-1}$ está definido. Además $\|\Gamma_0\| \leq \beta$.

(II) $\|F''(x)\| \leq M$ para $x \in \Omega$.

(III) $\|F''(x) - F''(y)\| \leq N\|x - y\|$ para $x, y \in \Omega$.

(IV) $\|y_0 - x_0\| \leq \eta$.

(V) El polinomio

$$p(t) = \frac{k}{2}t^2 - \frac{1}{3}t + \frac{\eta}{3}. \quad (3.3)$$

donde $M^2 + \frac{N}{2\beta} \leq k^2$, tiene dos raíces reales positivas.

Podemos escribir la aceleración convexa del método de Newton para resolver la ecuación $p(t) = 0$, con p dado por (3.3), de la siguiente forma

$$\begin{cases} s_n &= t_n - \frac{p(t_n)}{p'(t_n)} \\ t_{n+1} &= s_n + \frac{1}{2} \frac{L_p(t_n)}{1 - L_p(t_n)} (s_n - t_n). \end{cases} \quad (3.4)$$

Observar que en este caso, por el Corolario 1.4, el método (3.4) es convergente a r_1 , la menor raíz de $p(t) = 0$. Para estudiar la convergencia de (3.2) necesitaremos los siguientes lemas.

Lema 3.1. *En las condiciones anteriores podemos descomponer $F(x_{n+1})$ en una suma de la forma*

$$\begin{aligned} F(x_{n+1}) &= \int_{y_n}^{x_{n+1}} F''(x)(x_{n+1} - x) dx + \int_{x_n}^{y_n} F''(x)(x_{n+1} - y_n) dx \\ &+ \int_{x_n}^{y_n} F''(x)[1 - H(x_n)](y_n - x) dx + \int_{x_n}^{y_n} [F''(x) - F''(x_n)]H(x_n)(y_n - x_n) dx. \end{aligned}$$

Demostración. En primer lugar, notar que

$$\begin{aligned} F(x_{n+1}) &= F(x_{n+1}) - F(y_n) - F'(y_n)(x_{n+1} - y_n) + F(y_n) + F'(y_n)(x_{n+1} - y_n) \\ &= \int_{y_n}^{x_{n+1}} F''(x)(x_{n+1} - x) dx + F(y_n) + F'(y_n)(x_{n+1} - y_n). \end{aligned}$$

Por otra parte,

$$\begin{aligned} F(y_n) &= \int_{x_n}^{y_n} F''(x)(y_n - x) dx + F(x_n) + F'(x_n)(y_n - x_n) \\ &= \int_{x_n}^{y_n} F''(x)(y_n - x) dx. \end{aligned}$$

$$F'(y_n)(x_{n+1} - y_n) = \int_{x_n}^{y_n} F''(x)(x_{n+1} - y_n) dx + F'(x_n)(x_{n+1} - y_n).$$

y teniendo en cuenta la simetría del operador bilineal $F''(x_n)$,

$$F'(x_n)(x_{n+1} - y_n) = -\frac{1}{2} F''(x_n)H(x_n)(y_n - x_n)^2.$$

Entonces,

$$\int_{x_n}^{y_n} F''(x)(y_n - x) dx - \frac{1}{2} F''(x_n)H(x_n)(y_n - x_n)^2$$

$$\begin{aligned}
&= \int_0^1 F''(x_n + t(y_n - x_n))(y_n - x_n)^2(1-t) dt \\
&\quad - \int_0^1 F''(x_n)H(x_n)(y_n - x_n)^2(1-t) dt \\
&= \int_0^1 F''(x_n + t(y_n - x_n))[I - H(x_n)](y_n - x_n)^2(1-t) dt \\
&+ \int_0^1 [F''(x_n + t(y_n - x_n)) - F''(x_n)]H(x_n)(y_n - x_n)^2(1-t) dt \\
&\quad = \int_{x_n}^{y_n} F''(x)[I - H(x_n)](y_n - x) dx \\
&\quad + \int_{x_n}^{y_n} [F''(x) - F''(x_n)]H(x_n)(y_n - x_n) dx.
\end{aligned}$$

y en consecuencia, se sigue el resultado. ■

Lema 3.2. *Continuando con la misma notación se tiene, para $n \in \mathbb{N}$.*

$$(I_n) \quad \|\Gamma_n\| \leq -\frac{1}{p'(t_n)}.$$

$$(II_n) \quad \|H(x_n)\| \leq \frac{1}{1 - L_p(t_n)}.$$

$$(III_n) \quad \|I - H(x_n)\| \leq \frac{L_p(t_n)}{1 - L_p(t_n)}.$$

$$(IV_n) \quad \|y_n - x_n\| \leq s_n - t_n.$$

$$(V_n) \quad \|x_{n+1} - y_n\| \leq t_{n+1} - s_n.$$

$$(VI_n) \quad \|F(x_{n+1})\| \leq p(t_{n+1}).$$

Demostración. Para $n = 0$, (I_0) , (II_0) , (IV_0) y (V_0) se siguen inmediatamente de las condiciones (I)-(V). Para probar (III_0) , teniendo en cuenta la expresión del inverso de un operador ([92]), se sigue que

$$H(x_0) = [I - L_F(x_0)]^{-1} = \sum_{j=0}^{\infty} L_F(x_0)^j.$$

Así,

$$I - H(x_0) = -\sum_{n=1}^{\infty} L_F(x_0)^n = -L_F(x_0)H(x_0).$$

y por tanto,

$$\|I - H(x_0)\| \leq \|L_F(x_0)\| \|H(x_0)\| \leq \frac{L_p(t_0)}{1 - L_p(t_0)}.$$

Como suele ser habitual en estos casos, lo realmente difícil de probar es (VI_0) . Con la ayuda del Lema 3.1 podemos escribir

$$\begin{aligned}
F(x_1) &= \int_{y_0}^{x_1} F''(x)(x_1 - x) dx + \int_{x_0}^{y_0} F''(x)(x_1 - y_0) dx \\
&+ \int_{x_0}^{y_0} F''(x)[I - H(x_0)](y_0 - x) dx + \int_{x_0}^{y_0} [F''(x) - F''(x_0)]H(x_0)(y_0 - x_0) dx \\
&= \int_0^1 F''(y_0 + t(x_1 - y_0))(x_1 - y_0)^2(1 - t) dt \\
&+ \int_0^1 F''(x_0 + t(y_0 - x_0))(x_1 - y_0)(y_0 - x_0) dt \\
&+ \int_0^1 F''(x_0 + t(y_0 - x_0))[I - H(x_0)](y_0 - x_0)^2(1 - t) dt \\
&+ \int_0^1 [F''(x_0 + t(y_0 - x_0)) - F''(x_0)]H(x_0)(y_0 - x_0)^2 dt.
\end{aligned}$$

Así

$$\begin{aligned}
\|F(x_1)\| &\leq \frac{M}{2}\|x_1 - y_0\|^2 + M\|x_1 - y_0\|\|y_0 - x_0\| \\
&+ \frac{M}{2}\|I - H(x_0)\|\|y_0 - x_0\|^2 + \frac{N}{2}\|y_0 - x_0\|^3\|H(x_0)\| \\
&\leq \frac{M}{2}(t_1 - s_0)^2 + \frac{\|y_0 - x_0\|^2}{2(1 - L_p(t_0))} (2M\|L_F(x_0)\| + N\|y_0 - x_0\|) \\
&\leq \frac{M}{2}(t_1 - s_0)^2 + \frac{(s_0 - t_0)^3}{2(1 - L_p(t_0))} (2M^2\mathcal{J} + N).
\end{aligned}$$

Ahora bien, teniendo en cuenta la hipótesis (V), se sigue que

$$\begin{aligned}
\|F(x_1)\| &\leq \frac{M}{2}(t_1 - s_0)^2 + \frac{\mathcal{J}(s_0 - t_0)^3}{(1 - L_p(t_0))} \left(M^2 + \frac{N}{2\mathcal{J}} \right) \\
&\leq \frac{k}{2}(t_1 - s_0)^2 + \frac{k^2\mathcal{J}(s_0 - t_0)^3}{(1 - L_p(t_0))}.
\end{aligned}$$

Aplicando el Lema 3.1 al polinomio p dado por (3.3) se obtiene

$$p(t_1) = \frac{k}{2}(t_1 - s_0)^2 + \frac{k^2\mathcal{J}(s_0 - t_0)^3}{(1 - L_p(t_0))}.$$

y por tanto, $\|F(x_1)\| \leq p(t_1)$.

Empleando la misma técnica, y siguiendo un procedimiento inductivo, se puede probar que $(I_n) - (VI_n)$ se cumple para todo $n \in \mathbb{N}$. ■

Una vez vistos estos lemas podemos enunciar el siguiente resultado sobre la convergencia del método de la aceleración convexa en espacios de Banach.

Teorema 3.3. *Continuando con la misma notación anterior, consideramos el operador $F : \Omega \subseteq X \rightarrow Y$, con X e Y espacios de Banach y Ω un abierto convexo de X . Supongamos que F tiene derivadas de Fréchet de segundo orden continuas en Ω . Dado un punto $x_0 \in \Omega$, supongamos que se cumplen las condiciones (I)-(V) y que*

$$\overline{B(y_0, r_1 - \eta)} \subseteq \Omega.$$

Entonces el método de la aceleración convexa (3.2) es convergente. Además,

$$x_n, y_n \in \overline{B(y_0, r_1 - \eta)}, \quad n \in \mathbb{N}.$$

el límite $x^* = \lim_{n \rightarrow \infty} x_n$ es solución de la ecuación $F(x) = 0$ y se satisfacen las siguientes estimaciones del error

$$\|x^* - x_n\| \leq r_1 - t_n, \quad \|x^* - y_n\| \leq r_1 - s_n, \quad n \in \mathbb{N}.$$

Demostración. Se sigue inmediatamente a partir de los Lemas 3.1 y 3.2. ■

Notar que, por el Lema 2.4, podemos expresar $r_1 - t_n$ en función de las raíces del polinomio p definido en (3.3). Tenemos así

$$r_1 - t_n = \frac{(r_2 - r_1)\theta^{4^n}}{1 - \theta^{4^n}}, \quad n \geq 0, \quad \text{si } r_1 < r_2, \quad \theta = \frac{r_1}{r_2};$$

$$r_1 - t_n = \frac{r_1}{4^n}, \quad n \geq 0, \quad \text{si } r_1 = r_2.$$

5.4 Ejemplos

Concluimos este capítulo con unos cuantos ejemplos en los que se utiliza la aceleración convexa del método de Newton para resolver distintos tipos de ecuaciones. En primer lugar y, para funciones reales o de \mathbb{R}^2 en \mathbb{R}^2 , se compara la velocidad de convergencia a la raíz del método de Newton (orden dos), de su aceleración convexa (orden tres) y de otros procesos clásicos iterativos de tercer orden. A continuación, se estudian unas ecuaciones integrales analizando los resultados acerca de existencia y unicidad de solución que se obtienen.

En el primer ejemplo veremos que la aceleración convexa del método de Newton para resolver una ecuación polinómica de segundo grado mejora sensiblemente los resultados obtenidos por los métodos más conocidos de tercer orden, el método de Chebyshev [26], [37], [55], [58], y el método de Halley [1], [4], [25], [36], [38], [52], [109], y por supuesto, al método de Newton.

Ejemplo 4.1. Consideramos la ecuación

$$p(t) \equiv (t - 10)(t - 20) = t^2 - 30t + 200 = 0. \quad (4.1)$$

Denotamos $\{s_n\}$ la sucesión de Newton para resolver (4.1) y $\{t_n\}$, $\{u_n\}$ y $\{v_n\}$ las sucesiones de Chebyshev, Halley y la aceleración convexa del método de Newton respectivamente. Estas sucesiones vienen dadas por las siguientes expresiones:

$$\begin{aligned} s_{n+1} &= s_n - \frac{p(s_n)}{p'(s_n)} = \frac{s_n^2 - 200}{2s_n - 30}, \\ t_{n+1} &= t_n - \left[1 + \frac{L_p(t_n)}{2} \right] \frac{p(t_n)}{p'(t_n)} \\ &= t_n - \frac{(t_n^2 - 30t_n + 200)(5t_n^2 - 150t_n + 1100)}{8(t_n - 15)^3}, \\ u_{n+1} &= u_n - \left[\frac{2}{2 - L_p(u_n)} \right] \frac{p(u_n)}{p'(u_n)} = u_n - \frac{2(u_n^2 - 30u_n + 200)(u_n - 15)}{(3u_n^2 - 90u_n + 700)}, \\ v_{n+1} &= v_n - \frac{1}{2} \left[1 + \frac{1}{1 - L_p(v_n)} \right] \frac{p(v_n)}{p'(v_n)} \\ &= v_n - \frac{(3v_n^2 - 90v_n + 700)(v_n^2 - 30v_n + 200)}{4(v_n^2 - 30v_n + 250)(v_n - 15)}. \end{aligned}$$

Si empezamos en $s_0 = t_0 = u_0 = v_0 = 0$ obtenemos las sucesiones que se muestran en las Tablas 5.1 y 5.2. Como era de esperar, la sucesión $\{v_n\}$, obtenida por la aceleración convexa del método de Newton, converge a la solución de (4.1) más rápidamente que las otras sucesiones. ■

Ejemplo 4.2. Nos ocupamos ahora de resolver el sistema de ecuaciones dado por $F(x, y) = 0$, siendo

$$F(x, y) = (x^2 - y - 2, y^3 - x^2 + y + 1).$$

La sucesión de Newton para resolver dicha ecuación viene dada por

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) - \Gamma(x_n, y_n)F(x_n, y_n).$$

iteración	s_n	t_n
0	0.0000000000000000	0.0000000000000000
1	6.666666666666667	8.148148148148148
2	9.333333333333333	9.936934743362793
3	9.960784313372549	9.99995122930333
4	9.999847409781033	10.000000000000000
5	9.99999997671694	10.000000000000000
6	10.000000000000000	10.000000000000000

Tabla 5.1: Métodos de Newton y Chebyshev

iteración	u_n	v_n
0	0.0000000000000000	0.0000000000000000
1	8.571428571428571	9.333333333333333
2	9.980430528375734	9.999847409781033
3	9.99999925494193	10.000000000000000
4	10.000000000000000	10.000000000000000
5	10.000000000000000	10.000000000000000
6	10.000000000000000	10.000000000000000

Tabla 5.2: Métodos de Halley y de Newton acelerado

donde

$$\Gamma(x_n, y_n) = [F'(x_n, y_n)]^{-1} = \frac{1}{6x_n y_n^2} \begin{pmatrix} 3y^2 + 1 & 1 \\ 2x & 2x \end{pmatrix}.$$

En este caso se tiene la siguiente expresión para el método de Newton:

$$\begin{cases} x_{n+1} = \frac{3x^2 y^2 + 2y^3 + 6y^2 + 1}{6xy^2} \\ y_{n+1} = \frac{2y^3 + 1}{3y^2} \end{cases}$$

Para obtener la expresión de los métodos de Chebyshev, Halley y aceleración convexa del método de Newton nos será de utilidad conocer la matriz asociada al operador lineal $L_F(x, y)$, que en este caso es

$$L_F(x, y) = \frac{1}{6x^2 y^3} \begin{pmatrix} y(3x^2 y^2 - 2y^3 - 6y^2 - 1) & 2x(y^3 - 1) \\ 0 & 4x^2(y^3 - 1) \end{pmatrix}.$$

En función de este operador podemos definir los métodos iterativos anteriores como sigue. El método de Chebyshev es

$$(s_{n+1}, t_{n+1}) = (s_n, t_n) - \left[I + \frac{1}{2} L_F(s_n, t_n) \right] \Gamma(s_n, t_n) F(s_n, t_n).$$

es decir,

$$\begin{cases} s_{n+1} = s_n - \frac{(14s_n^2 t_n^2 - 2t_n^3 - 6t_n^2 - 1)(3s_n^2 t_n^2 - 2t_n^3 - 6t_n^2 - 1)}{72s_n^3 t_n^4} \\ \quad - \frac{(t_n^3 - 1)^2}{18s_n t_n^5} \\ t_{n+1} = t_n - \frac{(4t_n^3 - 1)(t_n^3 - 1)}{9t_n^5}. \end{cases}$$

Para el método de Halley podemos dar la siguiente expresión

$$(u_{n+1}, v_{n+1}) = (u_n, v_n) - \left[I + [2I - L_F(u_n, v_n)]^{-1} L_F(u_n, v_n) \right] \Gamma(u_n, v_n) F(u_n, v_n)$$

o, equivalentemente,

$$\begin{cases} u_{n+1} = u_n - \frac{6u_n(3u_n^2 v_n^2 - 2v_n^3 - 6v_n^2 - 1)}{9u_n^2 v_n^2 + 2v_n^3 + 6v_n^2 + 1} \\ \quad - \frac{2u_n(v_n^3 - 1)^2}{(2v_n^3 + 1)(9u_n^2 v_n^2 + 2v_n^3 + 6v_n^2 + 1)} \\ v_{n+1} = v_n - \frac{v_n(v_n^3 - 1)}{(2v_n^3 + 1)}. \end{cases}$$

iteración	x_n	y_n
0	6.000000000000000000	3.000000000000000000
1	3.336419753086419753	2.037037037037037037
2	2.183486059447010306	1.438355269870421386
3	1.806201995430546569	1.120022383336509062
4	1.737006139001477947	1.012402260381869450
5	1.732101431071650477	1.000151311360582637
6	1.732050814916382536	1.000000022890509658
7	1.732050807568877460	1.0000000000000000524
8	1.732050807568877294	1.000000000000000000
9	1.732050807568877294	1.000000000000000000

Tabla 5.3: Método de Newton

Por último, la aceleración convexa del método de Newton viene dada por

$$(w_{n+1}, z_{n+1}) = (w_n, z_n) - \frac{1}{2} \left[I + [I - L_F(w_n, z_n)]^{-1} \right] \Gamma(w_n, z_n) F(w_n, z_n).$$

Efectuando las operaciones correspondientes, llegamos a

$$\left\{ \begin{array}{l} w_{n+1} = w_n - \frac{(9w_n^2 z_n^2 + 2z_n^3 + 6z_n^2 + 1)(3w_n^2 z_n^2 - 2z_n^3 - 6z_n^2 - 1)}{12w_n z_n^2 (3w_n^2 z_n^2 + 2z_n^3 + 6z_n^2 + 1)} \\ \quad - \frac{(z_n^3 - 1)^2}{6w_n z_n^5} \\ z_{n+1} = z_n - \frac{(2z_n^3 + 1)(z_n^3 - 1)}{3z_n^2 (z_n^3 + 2)}. \end{array} \right.$$

Si comenzamos todas estas iteraciones en el mismo punto

$$(x_0, y_0) = (s_0, t_0) = (u_0, v_0) = (w_0, z_0) = (6, 3)$$

obtenemos las sucesiones dadas en las Tablas 5.3 a 5.6.

Como se puede observar en estas tablas, la aceleración convexa del método de Newton nos proporciona en cada iteración la mejor aproximación a la solución. Además, trabajando con esta precisión, llega a la solución exacta un paso antes que los métodos de Chebyshev y Halley y cuatro antes que el método de Newton.

■

iteración	s_n	t_n
0	6.000000000000000000	3.000000000000000000
1	2.719439840217446527	1.727937814357567444
2	1.830721648025601499	1.138820258896789643
3	1.733249320860063453	1.003038870277247239
4	1.732050823057711034	1.000000046348271498
5	1.732050807568877294	1.000000000000000000
6	1.732050807568877294	1.000000000000000000

Tabla 5.4: Método de Chebyshev

iteración	u_n	v_n
0	6.000000000000000000	3.000000000000000000
1	2.527771600300525920	1.581818181818181818
2	1.774166519412615274	1.057032561548124306
3	1.732099394264064235	1.000113501450570644
4	1.732050807569220348	1.000000000000974628
5	1.732050807568877294	1.000000000000000000
6	1.732050807568877294	1.000000000000000000

Tabla 5.5: Método de Halley

iteración	w_n	z_n
0	6.000000000000000000	3.000000000000000000
1	2.195930445526441461	1.173690932311621967
2	1.726757444904059338	0.998981852656109923
3	1.732050203990691682	1.00000000352888045
4	1.732050807568877294	1.000000000000000000
5	1.732050807568877294	1.000000000000000000

Tabla 5.6: Aceleración convexa del método de Newton

En el siguiente ejemplo, estudiamos una ecuación con raíces múltiples. Diremos que x^* es una raíz múltiple de la ecuación $F(x) = 0$ si el operador $F'(x^*)$ no es inversible. Esta situación ha sido menos tratada y, en general, es más complicada que el caso de raíces simples [91]. La convergencia a la solución suele ser muy lenta, incluso en el caso de ecuaciones escalares.

Ejemplo 4.3. Consideramos el sistema dado por $F(x, y) = 0$, donde

$$F(x, y) = ((x + y - 2)^2, (x - y)^2),$$

y cuya única solución es $(x^*, y^*) = (1, 1)$.

Tenemos en este caso que $L_F(x, y) = \frac{1}{2}I$, con I el operador identidad en \mathbb{R}^2 . En consecuencia, la aceleración convexa del método de Newton viene dada por la expresión

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) - \frac{3}{2}[F'(x_n, y_n)]^{-1}F(x_n, y_n) = \left(\frac{x_n + 3}{4}, \frac{y_n + 3}{4}\right).$$

Por otra parte, el método de Newton es

$$(x_{n+1}, y_{n+1}) = \left(\frac{x_n + 1}{2}, \frac{y_n + 1}{2}\right).$$

el de Halley

$$(x_{n+1}, y_{n+1}) = (x_n, y_n) - \frac{4}{3}[F'(x_n, y_n)]^{-1}F(x_n, y_n) = \left(\frac{x_n + 2}{3}, \frac{y_n + 2}{3}\right).$$

y el de Chebyshev coincide con la aceleración convexa del método de Newton en este caso.

Estudiando las sucesiones escalares que aparecen en estos métodos, y que denotamos

$$s_{n+1} = \frac{s_n + 1}{2}, \quad t_{n+1} = \frac{t_n + 2}{3}, \quad u_{n+1} = \frac{u_n + 3}{4}.$$

se puede ver sin dificultad que dichas sucesiones son convergentes a uno, empezando en cualquier punto de \mathbb{R} . Además la sucesión $\{u_n\}$ converge más rápidamente que $\{t_n\}$, y esta a su vez más que $\{s_n\}$. En la siguiente tabla se comparan estas sucesiones, tomando como punto de salida $s_0 = t_0 = u_0 = 0$. Como se puede ver (Tabla 5.7), la convergencia a la solución en todos los casos es muy lenta. ■

En el siguiente ejemplo, utilizamos la aceleración convexa del método de Newton para probar la existencia y unicidad de una ecuación integral (ver también [25], [26], [39]).

iteración	s_n	t_n	u_n
0	0.0000000000000000	0.0000000000000000	0.0000000000000000
1	0.5000000000000000	0.6666666666666667	0.7500000000000000
2	0.7500000000000000	0.8888888888888889	0.9375000000000000
3	0.8750000000000000	0.9629629629629639	0.9843750000000000
4	0.9375000000000000	0.9876543209876543	0.9960937500000000
5	0.9687500000000000	0.9958847736625514	0.9990234375000000
6	0.9843750000000000	0.9986282578875171	0.9997558593750000
7	0.9921875000000000	0.9995427526291724	0.9999389648437500
8	0.9960937500000000	0.9998475842097241	0.99998474121093753
9	0.9980468750000000	0.9999491947365747	0.99999618530273446
10	0.9990234375000000	0.9999830649121916	0.99999904632568363
11	0.9995117187500000	0.9999943549707305	0.99999976158142095
12	0.9997558593750000	0.9999981183235768	0.99999994039535527
13	0.9998779296875000	0.9999993727745256	0.99999998509883881
14	0.9999389648437500	0.9999997909248419	0.99999999627470970
15	0.9999694824218750	0.9999999303082806	0.9999999906867748

Tabla 5.7: Sucesiones de Newton, Halley y aceleración convexa del método de Newton

Ejemplo 4.4. En el espacio $X = C[0, 1]$ de las funciones continuas sobre el intervalo $[0, 1]$, con la norma del máximo,

$$\|x\| = \max_{s \in [0, 1]} |x(s)|.$$

consideramos la ecuación $F(x) = 0$, donde

$$F(x)(s) = x(s) - s + \frac{1}{2} \int_0^1 s \cos(x(t)) dt, \quad x \in C[0, 1], \quad s \in [0, 1].$$

Con la notación habitual, y tomando como punto de salida la función $x_0 = x_0(s) = s$, obtenemos, usando las definiciones de las derivadas de Fréchet de primer y segundo orden,

$$F'(z)x(s) = x(s) - \frac{s}{2} \int_0^1 x(t) \operatorname{sen}(z(t)) dt,$$

$$[F'(z)]^{-1}y(s) = y(s) + \frac{s}{2\Phi(t)} \int_0^1 y(t) \cos(z(t)) dt,$$

con $\Phi(t) = 1 - \frac{1}{2} \int_0^1 t \operatorname{sen}(z(t)) dt$,

$$F''(z)xy(s) = -\frac{s}{2} \int_0^1 x(t)y(t) \cos(z(t)) dt,$$

para $x, y, z \in C[0, 1]$ y $s \in [0, 1]$. A partir de aquí se deduce

$$a = b = \frac{\sin 1}{2 - \sin 1 + \cos 1}, \quad k = \frac{1}{2 - \sin 1 + \cos 1}.$$

Así, el polinomio que nos proporciona una sucesión mayorizante es

$$\begin{aligned} p(t) &= \frac{k}{6}t^3 - \frac{b}{2}t^2 - t + a \\ &= \frac{1}{6(2 - \sin 1 + \cos 1)} [t^3 + 3(\sin 1)t^2 - 6(2 - \sin 1 + \cos 1)t + 6 \sin 1]. \end{aligned}$$

Las raíces positivas de p son

$$r_1 = 0.6095694860276291, \quad r_2 = 1.70990829134757.$$

En consecuencia, el Teorema 2.6 nos asegura que $F(x) = 0$ tiene una raíz en $\overline{B(x_0, r_1)}$ y, además, esta es la única raíz en $B(x_0, r_2)$.

Denotamos $\{s_n\}$, $\{t_n\}$ y $\{u_n\}$ respectivamente a las sucesiones de Chebyshev, Halley y aceleración convexa para resolver la ecuación $p(t) = 0$. Cada una de ellas nos proporciona una cota del error para la cantidad

$$\|x^* - x_n\|.$$

n	$r_1 - s_n$	$r_1 - t_n$	$r_1 - u_n$
0	0.6095694860276291	0.6095694860276291	0.6095694860276291
1	0.0534834955243040	0.0495130055348865	0.0349873303274992
2	0.0001520166774545	0.0000984825547302	0.0000560164474543
3	0.0000000000042804	0.0000000000009129	0.0000000000001218
4	0.0000000000000000	0.0000000000000000	0.0000000000000000

Tabla 5.8: Cotas del error

siendo $\{x_n\}$ la correspondiente sucesión de funciones obtenida para cada método. En la Tabla 5.8 se muestran algunas cotas del error.

Observar que las mejores cotas se obtienen con la aceleración convexa del método de Newton. Para esta misma ecuación y , usando el método de Halley. Döring [39] obtuvo la cota

$$\|x^* - x_2\| \leq 0.000825.$$

Posteriormente, Candela and Marquina, [25], [26], dieron las cotas

$$\|x^* - x_2\| \leq 0.00037022683427694$$

and

$$\|x^* - x_2\| \leq 0.00014987029635502$$

empleando los métodos de Chebyshev y Halley respectivamente. Notar que las cotas que se muestran en la Tabla 5.8 mejoran sensiblemente las anteriores. ■

Como se ha puesto de manifiesto en la tercera sección de este capítulo, la aceleración convexa del método de Newton tiene un interés especial cuando se aplica a ecuaciones $F(x) = 0$, donde F es un operador que verifica que $F''(x)$ es constante. Aunque en general se trata de un proceso iterativo de tercer orden, en esta situación se alcanza orden cuatro y la convergencia es mucho más rápida que por el resto de métodos de tercer orden (ver Ejemplo 4.1). Para finalizar, vamos a estudiar una ecuación integral que surgió del trabajo de S. Chandrasekhar [31] relacionado con las teorías de la transferencia radiactiva, el transporte de neutrones y la teoría cinética de gases (ver también [92] y las referencias que allí se citan). Esta ecuación satisface que la derivada segunda del operador implicado es constante.

Ejemplo 4.5. En el espacio $C[0, 1]$ de las funciones continuas sobre el intervalo $[0, 1]$, dotado de la norma del máximo, se considera la ecuación $F(x) = 0$, donde

$$F(x)(s) = \alpha x(s) \int_0^1 \frac{s}{s+t} x(t) dt - x(s) + 1,$$

con $s \in [0, 1]$, $x \in C[0, 1]$ y $0 < \alpha \leq 2$.

A partir del Teorema 2.6 vamos a obtener información acerca de la existencia y unicidad de solución de dicha ecuación integral. Las expresiones de las derivadas del operador F son

$$F'(z)x(s) = \alpha z(s) \int_0^1 \frac{s}{s+t} x(t) dt + \alpha x(s) \int_0^1 \frac{s}{s+t} z(t) dt - x(s),$$

$$F''(z)xy(s) = \alpha x(s) \int_0^1 \frac{s}{s+t} y(t) dt + \alpha y(s) \int_0^1 \frac{s}{s+t} x(t) dt - x(s),$$

donde $x, y, z \in C[0, 1]$, $s \in [0, 1]$. Observar que $F''(z)$ es constante. Tomando como punto de partida la función $x_0 = x_0(s) = 1$, y para el caso particular de $\alpha = 1/4$, obtenemos (ver [37] para más detalles)

$$\|\Gamma_0\| = 1.53039421.$$

Así, manteniendo la notación empleada en el resto del capítulo, tenemos

$$\|\Gamma_0 F(x_0)\| \leq \|\Gamma_0\| \|F(x_0)\| = (1.53039421)\alpha \log 2 = 0.2651971 = a,$$

$$\|\Gamma_0 F''(x_0)\| \leq \|\Gamma_0\| \|F''(x_0)\| = 2a = b$$

y $k = 0$. Con estos datos, construimos el polinomio

$$p(t) = at^2 - t + a.$$

Las raíces de este polinomio son $r_1 = 0.2870485$ y $r_2 = 3.4837317$. Sabemos, por tanto que la ecuación integral anterior tiene una solución en la región $|s - 1| < r_1$, que además es única en $|s - 1| < r_2$. Por último y, usando de nuevo el Teorema 2.6, podemos dar unas expresiones del error en función de las raíces de p :

$$\|x^* - x_n\| \leq \frac{(r_2 - r_1)\theta^{4^n}}{1 - \theta^{4^n}}, \quad \theta = \frac{r_1}{r_2} = 0.08239685. \quad \blacksquare$$

Podemos mejorar los resultados obtenidos en los ejemplos de la Sección IV.4 sin más que utilizar, en lugar del método de Newton, su aceleración convexa.

Para finalizar, comentar que, como la aceleración convexa del método de Newton es el proceso iterativo de tercer orden que mejores resultados nos da cuando lo aplicamos a un operador con derivada segunda constante (ver ejemplos 4.1 y 4.5), sería interesante efectuar un estudio de este método sobre ecuaciones definidas por operadores cuadráticos generales [90], esto es, de la forma

$$Bx^2 + Lx + y = 0.$$

donde B es un operador bilineal simétrico entre dos espacios de Banach X e Y , L es un operador lineal de X en Y e y es un punto dado de Y . Un gran número de problemas son casos particulares de la ecuación anterior, como la propia ecuación de Chandrasekhar estudiada en el Ejemplo 4.5, y otros que aparecen en teoría de elasticidad, problemas de transporte de neutrones o doblamiento de vigas. Sobre este aspecto, hay una extensa bibliografía, de entre la que citamos los textos de Allgower y Prenter [3], Ames [5], Anselone y Moore [6] y Argyros [9], [11].

Bibliografía

- [1] ALEFELD. G.. On the convergence of Halley's method, *Amer. Math. Monthly*, **88**, (1981), 530–536.
- [2] ALEFELD. G. Y HERZBERGER. J.. "Introduction to interval computations", Academic Press, 1983.
- [3] ALLGOWER. E.L. Y PRENTER. P.M.. On the branching of solutions of quadratic differential equations, *Aequationes Mathematicae*, **10**, (1974), 81–96.
- [4] ALTMAN. M.. Concerning the method of tangent hyperbolas for operator equations, *Bull. Acad. Pol. Sci., Serie Sci. Math., Ast. et Phys.*, **9**, (1961), 633–637.
- [5] AMES. W.. "Ordinary differential equations in transport processes", Academic Press, 1968.
- [6] ANSELONE. P.M. Y MOORE. R.H.. An extension of the Newton-Kantorovich method for solving nonlinear equations with an application to elasticity, *J. Math. Anal. Appl.*, **13**, (1966), 476–501.
- [7] APOSTOL. T.M.. "Calculus". Vol. 1. Ed. Reverté, 1980.
- [8] ARGYROS. I.K.. Newton-like methods under mild differentiability conditions with error analysis, *Bull. Austral. Math. Soc.*, **37**, (1988), 131–147.
- [9] ARGYROS. I.K.. Quadratic equations and applications to Chandrasekhar's and related equations, *Bull. Austr. Math. Soc.*, **32**, (1988), 275–292.
- [10] ARGYROS. I.K.. On Newton's method and nondiscrete mathematical induction, *Bull. Austr. Math. Soc.*, **38**, (1988), 131–140.

- [11] ARGYROS, I.K.. On a class of nonlinear integral equations arising in Neutron transport. *Aequationes Mathematicae*, **36**, (1988), 99–111.
- [12] ARGYROS, I.K.. A Newton-like method for solving nonlinear equations in Banach space. *Acta. Math. Hung.*, **27**, (1992), 368–378.
- [13] ARGYROS, I.K.. Some methods for finding error bounds for Newton-like methods under mild differentiability conditions. *Studia Sci. Math. Hungar.*, **61(3–4)**, (1993), 183–194.
- [14] ARGYROS, I.K.. Remarks on the convergence of Newton's method under Hölder continuity conditions. *Tamkang J. Math.*, **23**, (1992), 269–277.
- [15] ARGYROS, I.K.. On the solution of equations with nondifferentiable operators. *Tamkang J. Math.*, **24**, (1993), 237–249.
- [16] ARGYROS, I.K.. Newton-like methods and non-discrete mathematical induction. *Studia Sci. Math. Hungar.*, **28**, (1993), 417–426.
- [17] BARANGER, J.. "Introduction à l'analyse numérique". Hermann, 1977.
- [18] BECKER, K.H. Y DÖRFLER, M.. "Dynamical systems and fractals". Cambridge, 1991.
- [19] BEAUZAMY, B.. "Introduction to Banach spaces and their geometry". North Holland, Math. Studies 68, 1985.
- [20] BERBERIAN, S.K.. "Lectures in functional analysis and operator theory". Springer-Verlag, 1974.
- [21] BODEWIG, E.. On types of convergence and on the behavior of approximations in the neighborhood of a multiple root of an equation. *Amer. Math. Monthly*, **7**, (1949), 325–333.
- [22] BORWEIM, J.M. Y NOLL, D.. Second order differentiability of convex functions in Banach spaces. *Trans. Amer. Math. Soc.*, **342**, (1994), 43–80.
- [23] BRAUN, M.. "Ecuaciones diferenciales y sus aplicaciones". Grupo Editorial Iberoamérica, 1990.
- [24] BURDEN, R.L. Y FAIRES, J.D.. "Análisis Numérico". Grupo Editorial Iberoamérica, 1985.

- [25] CANDELA, V. Y MARQUINA, A.. Recurrence relations for rational cubic methods I: The Halley method. *Computing*, **44**, (1990), 169–184.
- [26] CANDELA, V. Y MARQUINA, A.. Recurrence relations for rational cubic methods II: The Chebyshev method. *Computing*, **45**, (1990), 355–367.
- [27] CAMPBELL, R.. “Les intégrales eulériennes et leurs aplicaciones”. Dunod, 1966.
- [28] CARTAN, H.. “Calcul différentiel”. Hermann, 1971.
- [29] CIESIELSKI, Z.. Some properties of convex functions of higher orders. *Annales Polonici Mathematici*, **7**, (1959), 1–7.
- [30] CURTAIN, R.F. Y PRITCHARD, A.J.. “Functional analysis in modern applied mathematics”. Academic Press, 1977.
- [31] CHANDRASEKHAR, S.. “Radiative transfer”. Dover, 1960.
- [32] DAHLQUIST, G. Y BJÖRCK, A.. “Numerical Methods”. Prentice-Hall, 1974.
- [33] DAY, M.M.. “Normed linear spaces”. Springer-Verlag, 1958.
- [34] DENNIS, J.E.. On the Kantorovich hypothesis for Newton’s method. *SIAM J. Numer. Anal.*, **6**, (1969), 493–507.
- [35] DONG CHEN, On the convergence of the Halley’s method for nonlinear equation of one variable. *Tamkang. J. Math.*, **24**, (1993), 461–467.
- [36] DONG CHEN, Ostrowski-Kantorovich theorem and S -order of convergence of Halley method in Banach spaces. *Comment. Math. Univ. Carolin.*, **34**, (1993), 153–163.
- [37] DONG CHEN Y ARGYROS, I.K.. Results on the Chebyshev method in Banach spaces. *Proyecciones*, **12-2**, (1993), 119–128.
- [38] DONG CHEN, ARGYROS, I.K. Y QIAN, Q.S.. A note on the Halley’s method in Banach spaces. *Appl. Math. Comput.*, **58**, (1993), 215–224.
- [39] DÖRING, B.. Einige Sätze über das Verfahren der Tangierenden Hyperbeln in Banach-Räumen. *Aplikace Mat.*, **15**, (1970), 418–464.

- [40] DUMITRESCU, C. Y OLTEANU, O.. The inverse of a convex operator. Newton's method for convex operators. *Polytech. Inst. Bucharest Sci. Bull. Mech. Engrg.*, **53** (3-4), (1991), 23-29.
- [41] DUNFORD, N. Y SCHWARTZ, J.T.. "Linear operators. Part I. General theory". Interscience, 1958.
- [42] EHRMANN, H.. Konstruktion und durchführung von iterationsverfahren höherer ordnung. *Arch. Rat. Mech. Anal.*, **4**, (1959), 65-88.
- [43] GANDER, W.. On Halley's iteration method. *Amer. Math. Monthly*, **92**, (1985), 131-134.
- [44] GARAY, J. Y HERNÁNDEZ, M.A.. "Degree of logarithmic convexity". Publicaciones del Sem. Mat. García Galdeano, serie II, nº 26, Zaragoza, 1988.
- [45] GASCA, M.. "Cálculo Numérico: Resolución de ecuaciones y sistemas". Zaragoza, 1987.
- [46] GERALS, C.F. Y WHEATLEY, P.O.. "Applied Numerical Analysis". Addison-Wesley, 1994.
- [47] GILES, J.R.. "Convex analysis with application in differentiation of convex functions". Pitman, 1982.
- [48] GORN, S.. Maximal convergence intervals and a Gibbs type phenomenon for Newton's approximation procedure. *Ann. of Math.*, **59**, (1954), 463-476.
- [49] GRAGG, W.B. Y TAPIA, R.A.. Optimal error bounds for the Newton-Kantorovich theorem. *SIAM J. Numer. Anal.*, **11**, (1974), 10-13.
- [50] GUTIÉRREZ, J.M., HERNÁNDEZ, M.A. Y SALANOVA, M.A.. Optimización del error en el método de Newton mediante polinomios de segundo grado. *Actas III Jornadas Zaragoza-Pau de Mat. Apl.*, (1993).
- [51] HENRICI, P.. "Essentials of Numerical Annalysis". Wiley, 1982.
- [52] HERNÁNDEZ, M.A.. A note on Halley's method. *Numer. Math.*, **59-3**, (1991), 273-276.
- [53] HERNÁNDEZ, M.A.. Newton-Raphson's method and convexity. *Zb. Rad. Prirod.-Mat. Fak. Ser. Mat.*, **22-1**, (1992), 159-166.

- [54] HERNÁNDEZ. M.A. Y SALANOVA M.A.. "Grados de convexidad y concavidad de una curva. Su aplicación al estudio de procesos iterativos para la resolución de ecuaciones". Publicaciones del Sem. Mat. García Galdeano. serie II. nº 4. Zaragoza. 1991.
- [55] HERNÁNDEZ. M.A. Y SALANOVA M.A.. "La convexidad y concavidad en la convergencia de procesos iterativos para la resolución de ecuaciones". Publicaciones del Sem. Mat. García Galdeano. serie II. nº 5. Zaragoza. 1993.
- [56] HERNÁNDEZ. M.A. Y SALANOVA M.A.. A family of Chebyshev-Halley type methods. *J. Comput. Math.*, **47**. (1993). 59-63.
- [57] HERNÁNDEZ. M.A. Y SALANOVA M.A.. A family of Newton type iterative process. *Inter. J. Computer Math.*, **51**. (1994). 205-214.
- [58] HERNÁNDEZ. M.A. Y SALANOVA M.A.. On Chebyshev method. *Inter. J. Computer App. Math.*. por aparecer.
- [59] HILDEBRAND. F.B.. "Introduction to Numerical Analysis". Dover. 1987.
- [60] HÖRMANDER. L.. "Notions of convexity". Birkhäuser. 1994.
- [61] HOUSEHOLDER. A.S.. "The numerical treatment of a single nonlinear equation". McGraw-Hill. 1970.
- [62] HUANG. Z.. A note on the Kantorovich theorem for Newton iteration. *J. Comput. Appl. Math.*, **47**. (1993). 211-217.
- [63] KALABA. R.. On nonlinear differential equations. the maximum operation. and monotone convergence. *J. Math. Mech.*, **8**. (1959). 519-574.
- [64] KANTOROVICH. L.V.. On Newton's method for functional equations. *Dokl. Akad. Nauk S.S.S.R.*, **59**. (1948). 1237-1240.
- [65] KANTOROVICH. L.V.. On Newton's method. *Trudy Mat. Inst. Steklov.* **28**. (1949). 104-144.
- [66] KANTOROVICH. L.V. Y AKILOV. G.P.. "Functional Analysis". Pergamon Press. 1982.
- [67] KREYSZIG. E.. "Introductory Functional Analysis with applications". Wiley. 1978.

- [68] LANCASTER. P.. Error analysis for the Newton-Raphson method. *Numer. Math.*, **9**. (1966), 55-68.
- [69] LINDENSTRAUSS J. Y TZAFRIRI. L.. "Classical Banach spaces I". Springer-Verlag, 1977.
- [70] LINDENSTRAUSS J. Y TZAFRIRI. L.. "Classical Banach spaces II". Springer-Verlag, 1979.
- [71] LÓPEZ. O.. "Métodos iterativos de resolución de ecuaciones" Alhambra, 1986.
- [72] MIEL. G.J.. Majorizing sequences and error bounds for iterative methods. *Math. Comput.*, **34**. (1980), 185-202.
- [73] MIEL. G.J.. An updated version of the Kantorovich theorem for Newton's method. *Computing*, **27**. (1981), 237-244.
- [74] MOORE. R.E.. "Methods and applications of interval analysis". SIAM Studies in Applied Mathematics, 1979.
- [75] MORET. I.. A note on Newton type iterative methods. *Computing*, **33**. (1984), 65-73.
- [76] NASHED. M.Z.. "Generalized inverses and applications". Academic Press, 1976.
- [77] NASHED. M.Z.. Inner, outer and generalized inverses in Banach and Hilbert spaces. *Numer. Funct. Anal. Optimization*, **9**. (1987), 261-325.
- [78] NASHED. M.Z.. Convergence of Newton-like methods for singular operator equations using outer inverses. *Numer. Math.*, **66**. (1993), 235-257.
- [79] NETA. B.. Several new methods for solving equations. *Inter. J. Computer. Math.*, **23**. (1988), 265-282.
- [80] ORTEGA. J.M.. The Newton-Kantorovich theorem. *Amer. Math. Monthly*, **75**. (1968), 658-660.
- [81] ORTEGA. J.M. Y RHEINBOLDT. W.C.. "Iterative solution of nonlinear equations in several variables". Academic Press, 1970.
- [82] OSTROWSKI. A.M.. "Solution of equations in euclidean and Banach spaces". Academic Press, 1943.

- [83] OSTROWSKI, A.M.. "Solution of equations and systems of equations". Academic Press, 1966.
- [84] PHELPS, R.R.. "Convex functions, monotone operators and differentiability". Springer-Verlag, 1993.
- [85] POTRA, F.A. Y PTÁK, V.. Sharp error bounds for Newton process. *Numer. Math.*, **34**, (1980), 63-72.
- [86] POTRA, F.A. Y PTÁK, V.. "Nondiscrete induction and iterative processes". Pitman, 1984.
- [87] PORTER, D. Y STIRLING, D.S.G.. "Integral equations, a practical treatment, from spectral theory to applications". Cambridge, 1990.
- [88] PTÁK, V.. Concerning the rate of convergence of Newton's process. *Comm. Math. Univ. Carolinae*, **16**, (1975), 699-705.
- [89] PTÁK, V.. The rate of convergence of Newton's process. *Numer. Math.*, **25**, (1976), 279-285.
- [90] RALL, L.B.. Quadratic equations in Banach spaces. *Rend. Circ. Mat. Palermo*, **10**, (1961), 314-332.
- [91] RALL, L.B.. Convergence of the Newton process to multiple solutions. *Numer. Math.*, **9**, (1966), 23-37.
- [92] RALL, L.B.. "Computational solution of nonlinear operator equations". Robert E. Krieger Publishing Company, 1979.
- [93] RALSTON, A. Y RABINOWITZ, P.. "A first course in Numerical Analysis". McGraw-Hill, 1988.
- [94] RHEINOLDT, W.C.. A unified convergence theory for a class of iterative processes. *SIAM J. Numer. Anal.*, **5**, (1968), 42-63.
- [95] ROBERTS, A.W. Y VARBERG, D.E.. "Convex functions". Academic Press, 1973.
- [96] ROCKNE, J.. Newton's method under mild differentiability conditions with error analysis. *Numer. Math.*, **18**, (1972), 401-412.

- [97] RUDIN. W.. "Functional analysis". McGraw-Hill. 1974.
- [98] SCHRÖDER. E.. Über unendlich viele Algorithmen zur Auflösung der Gleichungen. *Math. Ann.* **2**, (1870), 317-365.
- [99] SCRATON. R.E.. "Métodos numéricos básicos". McGraw-Hill. 1987.
- [100] SIMMONS. "Ecuaciones diferenciales con aplicaciones y notas históricas". McGraw-Hill. 1977.
- [101] STOER. J. Y BULIRSCH. R. "Introduction to numerical analysis". Springer-Verlag, 1980.
- [102] TAPIA. R.A.. The Kantorovich theorem for Newton's method. *Amer. Math. Monthly*. **78**, (1971), 389-392.
- [103] TRAUB. J.F.. "Iterative methods for solution of equations". Prentice-Hall. 1964.
- [104] VANDERGRAFT. J.. Newton's method for convex operators in partially ordered spaces. *SIAM J. Numer. Anal.* **4** (3), (1967), 406-432.
- [105] WOODFORD. C.. "Solving linear and non-linear equations". Ellis Horwood. 1992.
- [106] YAMAMOTO. T.. A unified derivation of several error bounds for Newton's process. *J. Comput. Appl. Math.* **12 & 13**, (1985), 179-191.
- [107] YAMAMOTO. T.. A method for finding sharp error bounds for Newton's method under the Kantorovich assumption. *Numer. Math.* **49**, (1986), 203-220.
- [108] YAMAMOTO. T.. A convergence theorem for Newton-like methods in Banach spaces. *Numer. Math.* **51**, (1987), 545-557.
- [109] YAMAMOTO. T.. On the method of tangent hyperbolas in Banach spaces. *J. Comput. Appl. Math.* **21**, (1988), 75-86.
- [110] ZABREJKO. P.P. Y NGUEN. D.F.. The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates. *Numer. Funct. Anal. and Optimiz.* **9** (5 & 6), (1987), 671-684.