

UNA PROPUESTA PARA ANALIZAR LA EFECTIVIDAD REAL DE LOS MÉTODOS ITERATIVOS PARA RESOLVER ECUACIONES NO LINEALES

Pedro José Hurtado Camargo



Trabajo especial de grado: Modalidad Seminario-Monografía

Tutor: Giovanni Calderón

Mérida, Julio 2008

Departamento de Matemáticas

Facultad de Ciencias, Universidad de Los Andes

**UNA PROPUESTA PARA ANALIZAR LA EFECTIVIDAD REAL
DE LOS MÉTODOS ITERATIVOS PARA RESOLVER
ECUACIONES NO LINEALES¹**

Requisito Especial de Grado, presentado por:

PEDRO JOSÉ HURTADO CAMARGO

en la modalidad de Seminario-Monografía para optar al título de:

Licenciado en Matemáticas

Tutor: Dr. Giovanni Calderón²

¹Este trabajo fue parcialmente financiado por Consejo de Desarrollo Científico Humanístico y Tecnológico, CDCHT-ULA, bajo el proyecto: C-1525-07-05-F

²Universidad de Los Andes, Departamento de Matemáticas, Grupo Ciencias de la Computación, Edificio Teórico de la Facultad de Ciencias, La Hechicera, Mérida 5101, Mérida - Venezuela.

A mis hermanos, Yonathan Yari, Dexis Hurtado y Samuel Hurtado y a mis padres
Rosa M. Camargo Mendoza y Pedro Samuel Hurtado.

RESUMEN

Encontrar una o más raíces de una ecuación no lineal de la forma, $f(x) = 0$, donde f es una función real de variable real, es uno de los problemas más comunes que ocurren en las matemáticas aplicadas.

En el presente existen numerosos métodos iterativos para resolver la ecuación no lineal, $f(x) = 0$; entre los más populares se pueden citar a: el método Bisección, Regula Falsi, Newton-Raphson, Secante y Müller, entre otros (ver referencia [1]-[3]). En los últimos años se han definido métodos iterativos que mejoran, en cierta forma, la precisión de los métodos clásicos. En la mayoría de los casos, la eficiencia de estos nuevos métodos a sido justificada sólo mediante su orden de convergencia y en otros casos en función del número de iteraciones realizadas al ser aplicados en algunos ejemplos. Este tipo de análisis puede llegar a sesgar la conclusión sobre la superioridad o no de un método específico.

Por tal motivo, el objetivo principal de este trabajo radica en definir una estrategia para analizar la efectividad de los métodos iterativos tomando en cuenta el tiempo de CPU usado por los mismos. El tiempos de CPU es usado para estimar la diferencia de dos medias poblacionales (tiempo de CPU) para establecer diferencias entre cada par de métodos y poder realizar una clasificación. Una vez hecho este estudio se concluye sobre la aplicabilidad de cada método y su costo computacional.

En este trabajo se expondrán los métodos iterativos propuestos en las referencias [4]-[11]. Se presenta en la mayoría de los métodos, su construcción y su análisis de convergencia.

AGRADECIMIENTO

Deseo expresar aquí mi gratitud a todas aquellas personas que con su valiosa colaboración contribuyeron, de una u otra manera, en el desarrollo de esta investigación.

Por sus recomendaciones, sugerencias, y paciente labor de corrección en este trabajo, agradezco encarecidamente al profesor Giovanni Calderón, mi tutor, quien siempre depositó su confianza en mí.

Por sus tazas de café, paciencia y apoyo incondicional, a mi hermano Samuel Hurtado, y a Dexis Hurtado mi querida hermana, por su paciente ayuda.

Finalmente, a quienes desde La Luz siempre me apoyaron, a Dominga Mendoza y Juan Camargo, mis abuelos; a Rosa M. Camargo Mendoza y Pedro Samuel Hurtado, mis padres; a José Caicea, mi padre adoptivo; y a Yonathan Yari, mi apreciado hermanito. A todos gracias!!

Índice general

Introducción	III
1. Preliminares	1
1.1. Conceptos básicos	1
1.2. Métodos clásicos	3
1.3. Teoría general de los métodos iterativos	7
1.4. Criterios de paradas	14
1.5. Distribuciones fundamentales de muestreo	15
1.5.1. Algunas estadísticas importantes	17
1.5.2. Distribuciones muestrales.	23
1.5.3. Distribuciones muestrales de medias.	23
1.6. Problemas de estimación de una y dos muestras.	25
2. Métodos iterativos para resolver ecuaciones no lineales	31
2.1. Método iterativo JRS	31
2.2. Método iterativo MKN1	34
2.3. Método iterativo MKN2	39
2.4. Método iterativo NSPP	40
2.5. Métodos iterativos CCHUN1 y CCHUN2	44
2.6. Métodos iterativos CCHUN3 y CCHUN4	48
2.7. Método iterativo SG	52
2.8. Método iterativo BM	54
3. Estrategia de clasificación de los métodos y experimentación numérica	57
3.1. Estrategia de clasificación.	59
3.1.1. Construcción del intervalo $[a, b]$	61
3.2. Experimentación numérica.	65
3.2.1. Test estadístico.	67

3.3. Aportes, conclusiones y comentarios finales	70
A. Base de funciones	73
Referencias	77

Introducción

Encontrar una o más raíces de una ecuación no lineal de la forma, $f(x) = 0$, donde f es una función real de variable real ($f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$), es uno de los problemas más comunes que ocurren tanto en las matemáticas como en las distintas ramas de las ciencias aplicadas. Dicho problema constituye el punto de partida de este trabajo.

Un caso más sencillo del problema viene dado cuando la función f es un polinomio. Los polinomios de primero y segundo grado fueron tratados con éxitos en tiempos muy remotos, los babilonios eran capaces de resolver ecuaciones cuadráticas. La historia del descubrimiento de la solución algebraica de la cúbica enfrentó a dos grandes rivales italianos: Cardano¹ y Tartaglia² hacia 1540. Subsiguientemente, Ferrari³ alumno y secretario de Cardano resolvió en 1545 la ecuación de cuarto grado. Posteriormente fueron muchos los matemáticos eminentes que trataron de resolver las ecuaciones de grado superior a cuatro, aunque en vano puesto que el matemático noruego Abel⁴ probó que es imposible resolver por radicales la ecuación general de grado mayor que cuatro. En consecuencia, para calcular las raíces de polinomios de grado mayor que cuatro es imprescindible usar técnicas numéricas. A partir de este momento, y a lo largo de la historia, la construcción de métodos numéricos para resolver ecuaciones no lineales ha atraído la atención de matemáticos puros y aplicados.

Estos métodos numéricos consisten en hallar, mediante un proceso iterativo, un valor aproximado de x que satisfaga bajo ciertos criterios de error la ecuación no lineal $f(x) = 0$, tanto para el caso en que x representa una raíz (cero) simple como para raíces múltiples. En la actualidad existen numerosos métodos iterativos para resolver la ecuación no lineal, $f(x) = 0$; entre los más populares y por lo cual se les suele llamar clásicos, se pueden citar a: método de Bisección, Regula Falsi, Newton-Raphson, Secante y Müller, entre otros, ver referencia [1]-[3]. No obstante, en el presente, se sigue investigando en el tema y progresivamente surgen nuevos métodos iterativos que “mejoran” las propiedades de convergencia de los métodos clásicos. Pero, en línea general, estas mejoras quedan justificadas sólo mediante su orden de convergencia o, en casos peores, sobre la base de resultados numéricos (número de iteraciones) de algunos pocos ejemplos.

Desde un punto de vista práctico, en la comparación de dos o más métodos, además de sus propiedades teóricas (orden de convergencia, constante asintótica del error, selec-

ción de las aproximaciones iniciales), se debería tener en cuenta su costo computacional (tiempo de CPU usado). Pues, en la mayoría de los casos, las condiciones exigidas para obtener la convergencia de un método son restrictivas y en la práctica suele ser muy difícil que se cumplan en todo el proceso iterativo. Por tal motivo, existe la necesidad de hacer una clasificación de los nuevos métodos iterativos a partir del tiempo de CPU usado.

Con todo lo dicho, el objetivo principal de este trabajo radica en definir e implementar una estrategia para analizar la efectividad de los métodos iterativos tomando en cuenta el tiempo de CPU usado por los mismos. Una vez hecho este estudio se podrá concluir sobre la aplicabilidad de cada método y su costo computacional. Ahora bien, debido a la gran cantidad de referencias⁵ recientes en el tema (nuevos métodos iterativos), se han seleccionado 8 artículos para ser analizados (ver referencias [4]-[11]) a lo largo del trabajo. Los métodos propuestos en estas referencias están definidos para tratar el problema de ceros simples de funciones reales con variable real ($f : D \subseteq \mathbb{R} \rightarrow \mathbb{R}$). Para llevar a cabo la implementación y aplicación de cada uno de los métodos, se trabaja sobre una base de funciones no lineales y se eligen las condiciones iniciales de tal manera que cumplan con las propiedades de convergencia de cada uno de los métodos iterativos.

Para lograr el objetivo general antes dicho se deben alcanzar los siguientes objetivos:

- Definir un índice de clasificación para los métodos iterativos a partir de los tiempos de CPU obtenidos al usar distintas aproximaciones iniciales (número de muestras). Para este fin, se usará estimación de la diferencia de dos medias poblacionales para establecer diferencias entre cada par de métodos y poder realizar una clasificación.
- Como consecuencia del objetivo anterior resulta necesario la implementación de los distintos métodos iterativos. Las implementaciones numéricas de los métodos que surgen de [4]-[11] es llevada a cabo en MATLAB. No obstante, los códigos no son dados explícitamente en el trabajo.
- Comparar los resultados obtenidos de la experimentación numérica con los que surgen de resultados teórico, tales como: orden de convergencia y efectividad computacional. A partir de este estudio podremos concluir sobre la aplicabilidad de cada método y su costo computacional.
- De igual forma, para lograr el objetivo anterior, se construyen las pruebas de convergencia de los métodos iterativos dados en [4]-[11]. Sin embargo, vale la pena resaltar que algunas de estas pruebas son contribuciones originales de este trabajo. Pues, las mismas no se presentan por los autores en las referencias respectivas.

Los objetivos planteados se cubren en totalidad a lo largo del trabajo, el trabajo queda organizado en tres capítulos: en el Capítulo 1 se desarrollan los preliminares teóricos necesarios para abordar adecuadamente el estudio de la convergencia de los

métodos iterativos que se han introducidos en [4]-[11]. En el Capítulo 2 se expondrán los métodos iterativos propuestos en las referencias [4]-[11]. Se presenta en la mayoría de los métodos la construcción de cada uno de estos métodos y se estudia, también en la mayoría de los métodos, el análisis de convergencia. En el último Capítulo se define e implementa la estrategia para la clasificación de los métodos iterativos. Se presentan los resultados numéricos y las conclusiones generales del trabajo. En el Apéndice A se expone la base de funciones no lineales usadas en la experimentación numérica.

¹**Georolamo Cardano** nació en Pavia, Italia, el 24 de septiembre de 1501 y murió en Roma el 21 de septiembre de 1576. Cardano era médico de profesión, pero su celebridad es alcanzada por sus trabajos en las matemáticas (álgebra), llegando a ser el principal miembro de la escuela de Bolonia, que se dedicaba principalmente al estudio del álgebra. Es conocido por la resolución algebraica de polinomios de tercer grado. Sin embargo, a lo largo de la historia, se conoció que Cardano plagió, copió y publicó como propio el método de resolución de ecuaciones de tercer grado de Niccolò Fontana (más conocido como Tartaglia). Cardano hizo importantes contribuciones al Álgebra, Probabilidad, Hidrodinámica, Mecánica y Geología y publicó dos enciclopedias de Ciencias Naturales. Además, fue un jugador y apostador empedernido (especialmente en los dados y ajedrez).

²**Tartaglia**, sobrenombre de Niccolò Fontana (c. 1500-1557), matemático italiano nacido en Brescia, uno de los descubridores de la solución de la ecuación de tercer grado. Se le conoce como Tartaglia (el tartamudo) por su defecto en el habla, debido a las heridas que le causó de niño un soldado francés durante la invasión de su ciudad natal. Tartaglia enseñó matemáticas en varias universidades antes de instalarse en Florencia en 1542. Escribió sobre artillería y tradujo los Elementos de Euclides. Reveló su método de resolución de ecuaciones de tercer grado a otro famoso matemático renacentista, Gerolamo Cardano, y éste lo publicó en 1545, por lo que se conoce como "fórmula de Cardano".

³**Lodovico Ferrari** nació en Bolonia, Italia, el 2 de febrero de 1522 y murió en la misma ciudad envenenado por su hermana el 5 de octubre de 1565. Llegó a ser uno de los mayores representantes de la escuela de Bolonia (discípulo de Cardano). Uno de sus mayores aportes esta dado por la resolución algebraica de la ecuación general de cuarto orden.

⁴**Niels Henrik Abel** nació en Findö, Noruega, el 5 de agosto de 1802 y murió en Froland, Noruega, el 6 de abril de 1829. Fue un célebre matemático; se dedico también a las funciones analíticas, ámbito en el que desarrolló un método general para la construcción de funciones periódicas recíproca de la integral elíptica.

⁵Ver, entre otras, las contribuciones dadas en las revistas: Applied Mathematics And Computation, Applied Numerical Mathematics, Mathematics Letters, Computer & Mathematics with Applications, Journal of Computational and Applied Mathematics, Mathematics and Computer in Simulation.

Preliminares

El objetivo principal de este capítulo es introducir algunos conceptos básicos sobre la teoría de métodos iterativos para funciones no lineales. En particular: referentes a orden de convergencia de una sucesión, multiplicidad de una raíz o cero de una función. Del mismo modo, se realiza una introducción de los métodos iterativos más conocidos en la literatura, comunmente llamados métodos clásicos y algunos resultados de la Teoría General de los Métodos Iterativos necesarios para el estudio a realizarse en los capítulos posteriores. Los resultados teóricos a introducir en las siguientes secciones se pueden encontrar en textos de Análisis Numérico y como referencia pueden consultar [1],[2] y [3].

1.1. Conceptos Básicos

Definición 1.1 Sea $\{x_n\}_{n=0}^{\infty}$ una sucesión de números reales que converge a α y $\varepsilon_n = x_n - \alpha$, $n \geq 0$ el error en la iteración n . Si existe un número $p \geq 1$ y una constante $C > 0$ tal que

$$\lim_{n \rightarrow \infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|^p} = C,$$

entonces p se llama el **orden de convergencia** de la sucesión $\{x_n\}_{n=0}^{\infty}$

Para $p = 1, 2, 3$, la convergencia es lineal, cuadrática y cúbica, respectivamente. Si $1 < p < 2$ la convergencia se llama superlineal. En el caso lineal, la constante C se llama la *tasa de convergencia* (o constante asintótica) y necesariamente $0 < C < 1$.

También se puede definir al orden de convergencia de la sucesión $\{x_n\}_{n=0}^{\infty}$ como el número $p \geq 1$ tal que

$$|\varepsilon_{n+1}| \leq C|\varepsilon_n|^p \text{ para algún } C > 0, \text{ con } 0 < C < 1 \text{ para } p = 1.$$

Observación 1.1 *En general, una sucesión con mayor orden de convergencia se aproxima a la raíz más rápidamente que una de orden inferior. La constante asintótica influye en la rapidez de convergencia, pero no es tan importante como el orden.*

Proposición 1.1 *Sea $\{x_n\}_{n=0}^{\infty}$ una sucesión que converge a α , y $\varepsilon_n = x_n - \alpha$, $p \geq 1$, $k_i \in \mathbb{R}$, $k_1 \neq 0$, $i = 1, 2, \dots$. Si $\varepsilon_{n+1} = k_1\varepsilon_n^p + k_2\varepsilon_n^{p+1} + \dots$ ó $\varepsilon_{n+1} = k_1\varepsilon_n^p + O(\varepsilon_n^{p+1})$, entonces*

$$\lim_{n \rightarrow \infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|^p} = |k_1|.$$

Demostración:

Sea $\{x_n\}$ una sucesión que converge a α . $\varepsilon_n = x_n - \alpha$, $n \geq 0$, $p \geq 1$, $k_i \in \mathbb{R}$, $k_1 \neq 0$, $i = 1, 2, \dots$ Por hipótesis,

$$\varepsilon_{n+1} = k_1\varepsilon_n^p + k_2\varepsilon_n^{p+1} + k_3\varepsilon_n^{p+2} + \dots$$

Tomando ε_n^p como factor común, se obtiene

$$\varepsilon_{n+1} = \varepsilon_n^p(k_1 + k_2\varepsilon_n + k_3\varepsilon_n^2 + \dots),$$

Así,

$$\frac{\varepsilon_{n+1}}{\varepsilon_n^p} = k_1 + k_2\varepsilon_n + k_3\varepsilon_n^2 + \dots$$

Como $\varepsilon_n \rightarrow 0$ cuando $n \rightarrow \infty$, se obtiene,

$$\lim_{n \rightarrow \infty} \frac{|\varepsilon_{n+1}|}{|\varepsilon_n|^p} = |k_1|.$$

□

Definición 1.2 *Decimos que α es una raíz de multiplicidad m de una función f si existe una función h continua en α tal que*

$$f(x) = (x - \alpha)^m h(x) \quad y \quad h(\alpha) \neq 0.$$

Si $m = 1$, se dice que α es una raíz simple; si $m > 1$, se dice que α es una raíz no simple o raíz múltiple.

Teorema 1.1 *$f \in C^m[a, b]$ tiene un cero α de multiplicidad m en (a, b) sí y sólo si*

$$f^{(j)}(\alpha) = 0, \quad j = 0 : m - 1 \quad y \quad f^{(m)}(\alpha) \neq 0.$$

Demostración: Por definición

$$f(x) = (x - \alpha)^m h(x).$$

Aplicando la fórmula de Leibnitz para la derivada de un producto:

$$\begin{aligned} f^{(j)}(x) &= \sum_{k=0}^j \binom{j}{k} [(x - \alpha)^m]^{(k)} h(x)^{(j-k)} \\ &= \sum_{k=0}^j \binom{j}{k} m(m-1) \cdots (m-k+1) (x - \alpha)^{(m-k)} h(x)^{(j-k)}. \end{aligned}$$

Resulta entonces que si $j < m$ $f^{(j)}(\alpha) = 0$ y, si $j = m$, $f^{(m)}(\alpha) = m!$.
Recíprocamente,

$$\begin{aligned} f(x) &= f(\alpha) + f'(\alpha)(x - \alpha) + \frac{1}{2} f''(\alpha)(x - \alpha)^2 + \cdots \\ &= \cdots + \frac{1}{(p-1)!} f^{(p-1)}(\alpha)(x - \alpha)^{p-1} + \frac{f^{(p)}(\xi)}{p!} (x - \alpha)^p, \quad x < \xi < \alpha. \end{aligned}$$

Aplicando la hipótesis se tiene que

$$f(x) = \frac{f^{(p)}(\xi)}{p!} (x - \alpha)^p.$$

Siendo $f^{(p)}(x)$ continua en α , $f(x) = (x - \alpha)^p h(x)$ con $h(x) = \frac{f^{(p)}(\xi)}{p!}$, $\xi = \xi(x)$. \square

Ejemplo 1.1 La función descrita por $f(x) = e^x - x - 1$ tiene un cero de multiplicidad dos en $\alpha = 0$, porque $f(0) = e^0 - 0 - 1 = 0$ y $f'(0) = e^0 - 1 = 0$, pero $f''(0) = e^0 = 1$.

1.2. Métodos Clásicos

En esta sección se presentan de manera breve los métodos de Bisección, Secante, Newton y Steffensen. Se muestra en que consiste cada uno de los métodos mencionados y se enuncian algunos resultados de convergencia.

Método de Bisección

Supongamos que f es una función continua definida en el intervalo $[a, b]$ con $f(a)$ y $f(b)$ de signos diferentes. De acuerdo con el Teorema del Valor Intermedio, existe un número α en (a, b) tal que $f(\alpha) = 0$. Si bien el procedimiento se aplica aunque exista más de una raíz en el intervalo (a, b) , por razones de simplicidad se supone que la raíz de este

intervalo es única. El método requiere dividir varias veces a la mitad los subintervalos de $[a, b]$ y, en cada paso, localizar la mitad que contenga a α .

Para empezar, supongamos que $a_1 = a$ y $b_1 = b$, y sea x_1 el punto medio de $[a, b]$; es decir,

$$x_1 = a_1 + \frac{b_1 - a_1}{2} = \frac{a_1 + b_1}{2}.$$

Si $f(x_1) = 0$, entonces $\alpha = x_1$; de no ser así, entonces $f(x_1)$ tiene el mismo signo que $f(a_1)$ o $f(b_1)$. Si $f(x_1)$ y $f(a_1)$ tienen el mismo signo, entonces $\alpha \in (x_1, b_1)$ y se definen $a_2 = x_1$ y $b_2 = b_1$. Si $f(x_1)$ y $f(a_1)$ tienen signos opuestos, entonces $\alpha \in (a_1, x_1)$ y tomamos $a_2 = a_1$ y $b_2 = x_1$. Después se vuelve a aplicar el proceso al intervalo $[a_2, b_2]$ y así sucesivamente (ver Figura 1.1).

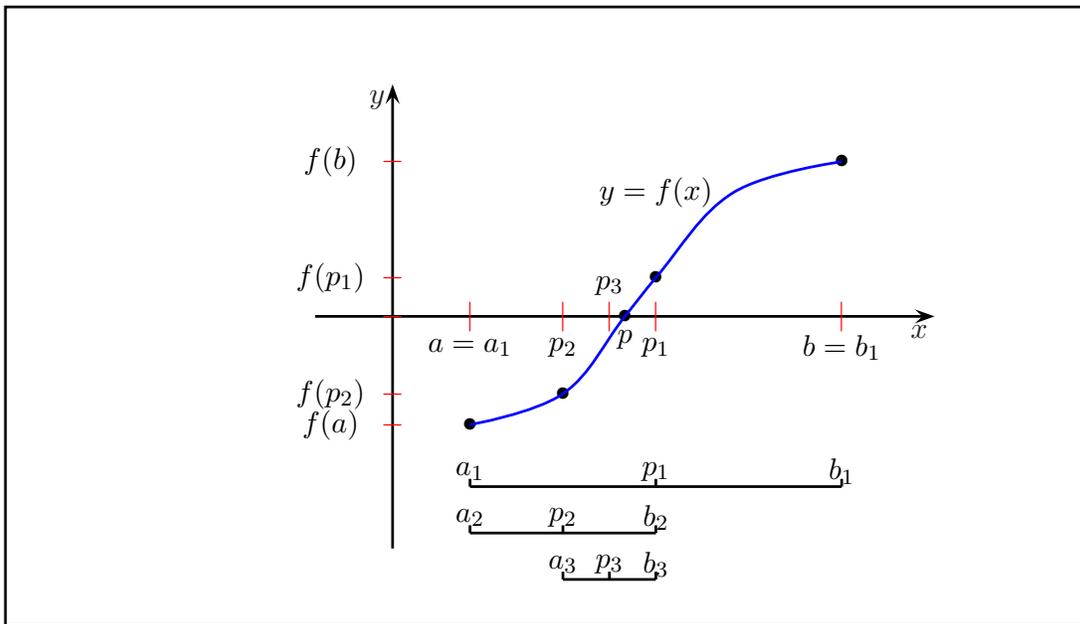


Figura 1.1: Método de Bisección: Secuencia de los intervalos y puntos medios.

Teorema 1.2 (Convergencia del Método de Bisección) Sea $f \in C[a, b]$ y $f(a)f(b) < 0$. El método de bisección genera una sucesión $\{x_n\}_{n=1}^{\infty}$ que aproxima a un cero α de f tal que

$$|x_n - \alpha| \leq \frac{b - a}{2^n}, \quad \text{donde } n \geq 1.$$

Demostración: Ver Capítulo 2: Teorema 2.1 de [2]. □

Método de la Secante

La gráfica de $y = f(x)$ se aproxima por una recta secante en una vecindad de la raíz, pero esta secante se determina en dos iterados sucesivos, independientemente del cambio de signo de f . Tampoco la raíz queda encerrada entre los iterados. El método se define de la siguiente manera

$$\begin{cases} x_0, x_1 & \text{aproximaciones iniciales a } \alpha \quad (f(\alpha) = 0) \\ x_{n+1} = & x_n - f(x_n) \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}. \end{cases}$$

Teorema 1.3 (Convergencia del Método de la Secante) *Sea f dos veces continuamente diferenciable en un intervalo que contine a α y supongamos que $f'(\alpha) \neq 0$. Si x_0 y x_1 están suficientemente próximos a α , entonces la sucesión $\{x_n\}$ de iterados converge a α . El orden de convergencia es*

$$p = \frac{1 + \sqrt{5}}{2} \approx 1,62.$$

Demostración: Ver Capítulo 2 de [3]. □

Método de Newton

El método de Newton (o método de Newton-Raphson) es una de las técnicas numéricas para resolver el problema de búsqueda de raíces $f(x) = 0$ más poderosa y conocida. Hay muchas formas de introducirlos. La más común consiste en considerarlo gráficamente. Otra posibilidad consiste en derivarlo como una técnica que permite lograr una convergencia más rápida que la que ofrecen otros tipos de iteración funcional. Una tercera forma de introducir el método de Newton, es considerando los polinomios de Taylor. La tercera forma de introducirlo es la que se va a usar en este trabajo. Supongamos que $f \in C^2[a, b]$ y $x_0 \in [a, b]$ una aproximación de α tal que $f'(x_0) \neq 0$ y $|\alpha - x_0|$ es "pequeño". Consideremos el primer polinomio de Taylor para $f(x)$ expandido alrededor de x_0 ,

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \frac{(x - x_0)^2}{2}f''(\xi(x)),$$

donde $\xi(x)$ está entre x y x_0 . Dado que $f(\alpha) = 0$ entonces de la ecuación anterior para $x = \alpha$, se tiene

$$0 = f(x_0) + (\alpha - x_0)f'(x_0) + \frac{(\alpha - x_0)^2}{2}f''(\xi(\alpha)).$$

Se deriva el método de Newton, suponiendo que, como $|\alpha - x_0|$ es tan pequeño, el término que contiene $(\alpha - x_0)^2$ es mucho menor y que

$$0 \approx f(x_0) + (\alpha - x_0)f'(x_0).$$

Despejando α de esta ecuación se obtiene

$$\alpha \approx x_0 - \frac{f(x_0)}{f'(x_0)}.$$

De esta manera se introduce el método de Newton, el cual comienza con una aproximación inicial x_0 y genera la sucesión $\{x_n\}_{n=0}^{\infty}$ definida por

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}, \quad \text{para } n \geq 1.$$

En la Figura 1.2 se muestra gráficamente cómo se obtienen las aproximaciones usando tangentes sucesivas. Comenzando con la aproximación inicial x_0 , la aproximación x_1 es la intersección con el eje x de la línea tangente a la gráfica de f en $(x_0, f(x_0))$. La aproximación x_2 es la intersección con el eje x de la línea tangente a la gráfica de f en $(x_1, f(x_1))$, y así sucesivamente.

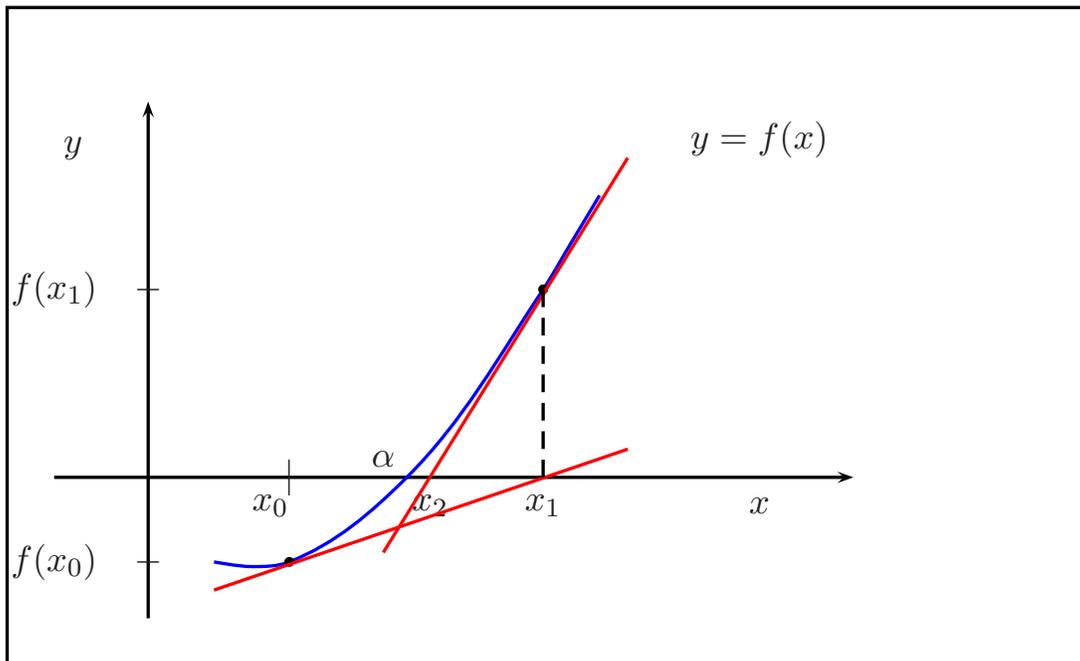


Figura 1.2: Método de Newton

El siguiente teorema de convergencia para el método de Newton muestra la impor-

tancia teórica de la elección de x_0 .

Teorema 1.4 (Convergencia del Método de Newton) *Sea $f \in C^2[a, b]$. Si $\alpha \in [a, b]$ es tal que $f(\alpha) = 0$ y $f'(\alpha) \neq 0$, entonces existe $\delta > 0$ tal que el método de Newton genera una sucesión $\{x_n\}_{n=1}^{\infty}$ que converge a α para cualquier aproximación inicial $x_0 \in [\alpha - \delta, \alpha + \delta]$ y además la convergencia es cuadrática.*

La demostración de este teorema se puede realizar considerando los resultados de la Teoría General de los Métodos Iterativos que se presentan en la siguiente sección. Para mayores detalles de la demostración del teorema anterior pueden consultar el Capítulo 2 de [2].

Método de Steffensen

Es una modificación del método de Newton, análogo al de la Secante pero usando una aproximación diferente de la derivada de f . Dicha aproximación es la siguiente

$$f'(x_n) \approx \frac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}.$$

El método se define como

$$\begin{cases} x_0 & \text{una aproximación inicial a } \alpha \quad (f(\alpha) = 0) \\ x_{n+1} & = x_n - \frac{f(x_n)}{\frac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}}. \end{cases}$$

En este método se hacen dos evaluaciones de funciones en cada paso.

1.3. Teoría General de los Métodos Iterativos

Los métodos iterativos de Newton, Secante y Steffensen pueden considerarse casos particulares del siguiente método iterativo más general: sea x_{n+1} determinado por evaluaciones de la función y/o de las derivadas en los puntos $x_n, x_{n-1}, \dots, x_{n-m+1}$ y pongamos

$$x_{n+1} = g(x_n, x_{n-1}, \dots, x_{n-m+1}), \quad n = m, m+1, \dots$$

La función g se llama *función de iteración*, y $\{x_n\}_{n=0}^{\infty}$ la *sucesión de iterados*.

En los métodos presentados en la sección anterior se tiene que

$$g(x) = x - \frac{f(x)}{f'(x)}, \quad m = 1, \quad \text{para Newton.}$$

$$g(x, y) = x - f(x) \frac{x - y}{f(x) - f(y)}, \quad m = 2, \quad \text{para la Secante.}$$

$$g(x) = x - \frac{f(x)}{\frac{f(x + f(x)) - f(x)}{f(x)}}, \quad m = 1, \quad \text{para Steffensen.}$$

La teoría general de los métodos iterativos es simple cuando $m = 1$, es decir,

$$x_{n+1} = g(x_n), \quad n = 0, 1, 2, \dots \quad (1.1)$$

Sea $\{x_n\}_{n=0}^{\infty}$ una sucesión generada por (1.1) para un x_0 inicial dado. Supongamos que $x_n \rightarrow \alpha$ y que g es continua. Tomando límites en (1.1):

$$\alpha = g(\alpha). \quad (1.2)$$

Si (1.2) es cierto decimos que α es un *punto fijo* de g . Para resolver el problema $f(x) = 0$, se puede construir una función g tal que un punto fijo de g sea una raíz de f .

Lema 1.1 Si $g \in C[a, b]$ y $g : [a, b] \rightarrow [a, b]$, entonces g tiene un punto fijo en $[a, b]$.

Demostración: Si $g(a) = a$ o si $g(b) = b$, entonces g tendrá un punto fijo en un extremo. Supongamos que no es así; entonces deberá ser cierto que $g(a) > a$ y que $g(b) < b$. La función $h(x) = g(x) - x$ es continua en $[a, b]$ y tenemos

$$h(a) = g(a) - a > 0 \quad \text{y} \quad h(b) = g(b) - b < 0.$$

Por el Teorema del Valor Intermedio se tiene que existe $\alpha \in (a, b)$ para el cual $h(\alpha) = 0$. Este número α es un punto fijo de g .

$$0 = h(\alpha) = g(\alpha) - \alpha \quad \text{implica que} \quad g(\alpha) = \alpha.$$

□

Lema 1.2 Sea $g : [a, b] \rightarrow [a, b]$ y $g \in C[a, b]$. Si además g' existe en $[a, b]$ y existe una constante positiva $k < 1$ con

$$|g'(x)| \leq k, \quad \text{para toda } x \in (a, b),$$

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

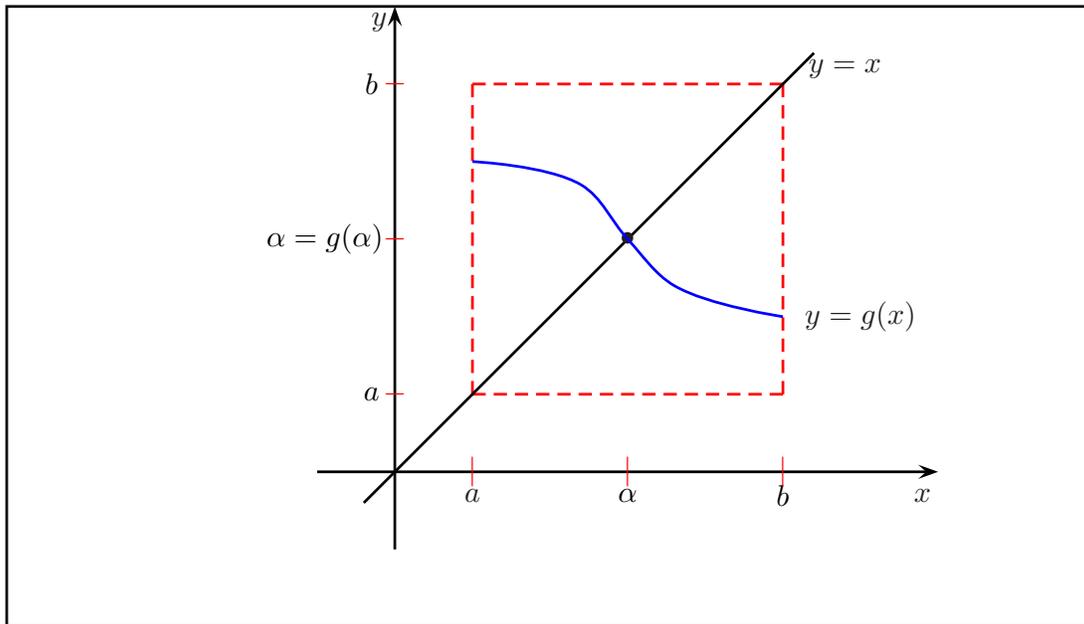


Figura 1.3: Lema 1.2

Demostración: Supongamos que $|g'(x)| \leq k \leq 1$ y que α y β son puntos fijos en $[a, b]$ tal que $\alpha \neq \beta$. Según el teorema del valor medio, existe un número ξ entre α y β y, por tanto, en $[a, b]$ tal que

$$\frac{g(\alpha) - g(\beta)}{\alpha - \beta} = g'(\xi).$$

Por tanto,

$$|\alpha - \beta| = |g(\alpha) - g(\beta)| = |g'(\xi)||\alpha - \beta| \leq k|\alpha - \beta| < |\alpha - \beta|,$$

lo cual es una contradicción. Esta contradicción se debe solamente a la suposición, $\alpha \neq \beta$. Es decir, $\alpha = \beta$ y el punto fijo en $[a, b]$ es único. \square

Teorema 1.5 (Teorema de punto fijo) Sea $g \in C[a, b]$ tal que $g(x) \in [a, b]$ para toda $x \in [a, b]$. Además, supongamos que existe g' en (a, b) y una constante positiva $0 < k < 1$ tales que

$$|g'(x)| \leq k, \quad \text{para toda } x \in (a, b).$$

Entonces, para cualquier número x_0 en $[a, b]$, la sucesión definida por

$$x_n = g(x_{n-1}), \quad n \geq 1,$$

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

Demostración: Los Lemas (1.1) y (1.2) implican que existe un punto fijo único en $[a, b]$. Puesto que g transforma $[a, b]$ en sí mismo, la sucesión $\{x_n\}_{n=0}^{\infty}$ se define para todo $n \geq 0$ y $x_n \in [a, b]$ para toda n . Aplicando el hecho de que $|g'(x)| \leq k$ y el Teorema del Valor Medio, tenemos, para cada n ,

$$|x_n - \alpha| = |g(x_n) - g(\alpha)| = |g'(\xi_n)||x_{n-1} - \alpha| \leq k|x_{n-1} - \alpha|,$$

donde $\xi_n \in (a, b)$. Al aplicar esta desigualdad inductivamente se obtiene

$$|x_n - \alpha| \leq k|x_{n-1} - \alpha| \leq k^2|x_{n-2} - \alpha| \leq \dots \leq k^n|x_0 - \alpha|.$$

Puesto que $0 < k < 1$, se tiene $\lim_{n \rightarrow \infty} k^n = 0$ y

$$\lim_{n \rightarrow \infty} |x_n - \alpha| \leq \lim_{n \rightarrow \infty} k^n|x_0 - \alpha| = 0.$$

Por lo tanto, $\{x_n\}_{n=0}^{\infty}$ converge a α . □

Corolario 1.1 *Si g satisface las hipótesis del Teorema (1.5), las cotas del error que supone utilizar x_n para aproximar α están dadas por las siguientes desigualdades*

$$|x_n - \alpha| \leq k^n \max\{x_0 - a, b - x_0\}$$

y

$$|x_n - \alpha| \leq \frac{k^n}{1 - k}|x_1 - x_0|, \quad \text{para toda } n \geq 1.$$

Demostración: ver [2, pp. 62]. □

Teorema 1.6 *Sea g continuamente diferenciable en algún intervalo (c, d) que contenga el punto fijo α de g . Si $|g'(\alpha)| < 1$, entonces existe una $\delta > 0$ tal que la iteración de punto fijo converge para cualquier aproximación x_0 siempre que $|x_0 - \alpha| \leq \delta$.*

Demostración: Por ser g' continua y $|g'(\alpha)| < 1$, entonces existe $\delta > 0$ tal que $|g'(x)| < 1$ para todo $x \in [\alpha - \delta, \alpha + \delta]$. Para continuar la prueba es necesario demostrar que g envía $[\alpha - \delta, \alpha + \delta]$ en $[\alpha - \delta, \alpha + \delta]$. Si $x \in [\alpha - \delta, \alpha + \delta]$, el Teorema del Valor Medio implica que, para algún número ξ entre x y α , $|g(x) - g(\alpha)| = |g'(\xi)||x - \alpha|$. Por tanto,

$$|g(x) - \alpha| = |g(x) - g(\alpha)| = |g'(\xi)||x - \alpha| < |x - \alpha|.$$

Puesto que $x \in [\alpha - \delta, \alpha + \delta]$, se tiene que $|x - \alpha| \leq \delta$ y que $|g(x) - \alpha| < \delta$.

Este resultado implica que g manda $[\alpha - \delta, \alpha + \delta]$ en $[\alpha - \delta, \alpha + \delta]$

Ahora, por el Teorema (1.5), la iteración definida por $x_n = g(x_{n-1})$, $n \geq 1$ converge para cualquier aproximación $x_0 \in [\alpha - \delta, \alpha + \delta]$, con lo cual queda demostrado el Teorema (1.6). □

Teorema 1.7 *Sea g una función de iteración tal que $g^{(p)}$ es continua en un entorno de α , $p \geq 2$ y $\varepsilon_n = x_n - \alpha$. Supongamos además que*

$$g(\alpha) = \alpha; \quad g^{(j)}(\alpha) = 0, \quad j = 1, 2, \dots, p-1; \quad g^{(p)}(\alpha) \neq 0.$$

Si x_0 se elige suficientemente próximo a α , la iteración

$$x_{n+1} = g(x_n), \quad n \geq 0;$$

tendrá un orden de convergencia igual a p y

$$\lim_{n \rightarrow \infty} \frac{x_{n+1} - \alpha}{(x_n - \alpha)^p} = \lim_{n \rightarrow \infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^p} = \frac{g^{(p)}(\alpha)}{p!}.$$

Demostración: g es continuamente diferenciable en un entorno de α y $g'(\alpha) = 0 < 1$. Según el Teorema (1.6), se obtiene que $x_n \rightarrow \alpha$. Siendo

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

con ξ_n entre α y x_n y $g^{(j)}(\alpha) = 0 \quad j = 1, \dots, (p-1)$, resulta que

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

Así,

$$\lim_{n \rightarrow \infty} \frac{\varepsilon_{n+1}}{\varepsilon_n^p} = \frac{g^{(p)}(\alpha)}{p!}.$$

y de la definición (1.1), la iteración

$$x_{n+1} = g(x_n), \quad n \geq 0;$$

es de orden p . □

Definición 1.3 (Eficiencia Computacional) *Sea $f(x) = 0$ una ecuación no lineal y Ψ la función de iteración que converge al cero α con un orden de convergencia igual a r . Se dice que el número $r^{1/d}$ es la **eficiencia computacional** de Ψ con $d := \sum_{j=0}^n h_{f^{(j)}}$, en donde*

1. n es el número máximo de derivadas de f realizadas en Ψ , y
2. $h_{f^{(j)}}$ es la cantidad de evaluaciones de $f^{(j)}$ realizada en Ψ

Ejemplo 1.2 (Eficiencia Computacional) Sea Ψ la función de iteración de Newton dada por

$$\Phi(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$

Entonces del Teorema 1.4, tenemos que Φ tiene segundo orden de convergencia ($r = 2$). Es fácil ver que Φ requiere de una evaluación de f y f' en cada iteración, es decir $d = 2$. Así, la eficiencia computacional de Φ es:

$$r^{1/d} = 2^{1/2} \approx 1,42$$

□

Los siguientes resultados teóricos serán de utilidad en el análisis de el orden de los métodos iterativos a estudiar en el Capítulo 2.

Teorema 1.8 Sea $f \in C^{(3)}$, $h \in C^{(2)}$ en un entorno de una raíz simple de f . Si $h(\alpha) = 1$ y $h'(\alpha) = \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)}$, entonces existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $\Phi(x_n) = x_n - h(x_n) \frac{f(x_n)}{f'(x_n)}$ converge a α y además es de al menos orden tres.

Demostración: Para demostrar que Φ es convergente en un entorno de α tenemos que verificar que $\Phi(\alpha) = \alpha$ y $|\Phi'(\alpha)| < 1$

$$\Phi(\alpha) = \alpha - h(\alpha) \frac{f(\alpha)}{f'(\alpha)} = \alpha, \quad (1.3)$$

$$\Phi'(x) = 1 - h'(x) \frac{f(x)}{f'(x)} - h(x) \left[\frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} \right], \quad (1.4)$$

$$\begin{aligned} \Phi'(\alpha) &= 1 - h'(\alpha) \frac{f(\alpha)}{f'(\alpha)} - h(\alpha) \left[\frac{f'(\alpha)^2 - f(\alpha)f''(\alpha)}{f'(\alpha)^2} \right], \\ &= 1 - \frac{f'(\alpha)^2}{f'(\alpha)^2} = 1 - 1 = 0 < 1. \end{aligned} \quad (1.5)$$

De (1.3) y (1.5) y el Teorema (1.6), existe $\delta > 0$ tal que Φ converge a α para cualquier aproximación inicial $x_0 \in [\alpha - \delta, \alpha + \delta]$.

Derivando (1.4), se obtiene

$$\begin{aligned}
\Phi''(x) &= - \left[h''(x) \frac{f(x)}{f'(x)} + 2h'(x) \left(\frac{f'(x)^2 - f(x)f''(x)}{f'(x)^2} \right) \right. \\
&\quad + h(x) \left(\frac{[2f'(x)f''(x) - (f'(x)f''(x) + f(x)f'''(x))]f'(x)^2}{f'(x)^4} \right. \\
&\quad \left. \left. - \frac{(f'(x)^2 - f(x)f''(x))2f'(x)f''(x)}{f'(x)^4} \right) \right] \\
\Phi''(\alpha) &= - \left[2h'(\alpha) + h(\alpha) \left(\frac{f'(\alpha)^3 f''(\alpha) - 2f'(\alpha)^3 f''(\alpha)}{f'(\alpha)^4} \right) \right] \\
&= - \left[2h'(\alpha) - \frac{f'(\alpha)^3 f''(\alpha)}{f'(\alpha)^4} \right] \\
&= - \left[2 \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)} - \frac{f''(\alpha)}{f'(\alpha)} \right] = 0
\end{aligned} \tag{1.6}$$

De (1.4), (1.5), (1.6) y el Teorema (1.7) se deduce que Φ es de al menos orden tres. \square

Corolario 1.2 Sea α una raíz simple de $f \in C^3(I)$, con I un intervalo abierto que contiene a α . Supongamos que $h \in C^2(I)$. Si

$$F(x) = x e^{-\frac{f(x)}{x f'(x)}}, \quad \phi(\alpha) = \alpha, \quad \phi'(\alpha) = 0, \quad \phi''(\alpha) \neq 0, \quad \alpha \neq 0;$$

y $h(x) = 1 + \frac{1}{2} F'(x) + \ln \left(\sqrt{\left| \frac{\phi(x)}{x} \right|} \right)$, entonces existe $\delta > 0$ tal que para todo $x_0 \in$

$[\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $\Phi(x_n) = x_n - h(x_n) \frac{f(x_n)}{f'(x_n)}$ converge a α y tiene por lo menos un tercer orden de convergencia.

Demostración: La idea es verificar que $h(\alpha) = 1$ y $h'(\alpha) = \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)}$. Realizando los cálculos de F' y F'' evaluadas en α se llega,

$$F'(\alpha) = 0, \tag{1.7}$$

$$F''(\alpha) = \frac{1}{\alpha} + \frac{f''(\alpha)}{f'(\alpha)}, \tag{1.8}$$

y al evaluar en α la derivada de $\ln \left(\sqrt{\left| \frac{\phi(x)}{x} \right|} \right)$ se obtiene $-\frac{1}{2\alpha}$.

Así,

$$h(\alpha) = 1 + \frac{1}{2}F'(\alpha) + \ln \left(\sqrt{\left| \frac{\phi(\alpha)}{\alpha} \right|} \right) = 1,$$

$$h'(\alpha) = \frac{1}{2} \left(\frac{1}{\alpha} + \frac{f''(\alpha)}{f'(\alpha)} \right) - \frac{1}{2\alpha} = \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)}$$

y del Teorema 1.8, existe $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $\Phi(x_n) = x_n - h(x_n) \frac{f(x_n)}{f'(x_n)}$ converge a α y tiene por lo menos un tercer orden de convergencia. □

1.4. Criterios de paradas

Resulta común que un proceso iterativo se detenga ya sea porque alcanzó un máximo número de iteraciones o porque logró alcanzar una tolerancia del error (prescrita por el usuario) entre la solución aproximada y la solución exacta del problema. Para este último propósito existen criterios de parada que dependen tanto de la multiplicidad de la raíz como del orden de convergencia de la función de iteración. Debido a que los métodos a ser analizados están propuestos para ceros simples y sus ordenes de convergencia son mayores a uno, dos criterios de parada son implementados: para una tolerancia ε del error se pide que

$$|f(x_n)| \leq \varepsilon \quad \text{o} \quad |x_{n+1} - x_n| \leq \varepsilon.$$

A continuación se analizan los criterios.

Criterio 1: $|f(x_n)| \leq \varepsilon$. Para analizar este criterio, supongamos que f es diferenciable en algún intervalo $I \subseteq \mathbb{R}$ que contenga a α , considere que $x_n \in I$ para cada $n \in \mathbb{N}$. Sea x_n la n -ésima aproximación a α , por el Teorema del Valor Intermedio se tiene

$$f(x_n) = f(x_n) - f(\alpha) = f'(\xi_n)(x_n - \alpha), \quad \text{con} \quad \xi_n \in \text{int}(x_n, \alpha),$$

donde $\text{int}(x_n, \alpha)$ representa el interior entre los valores de x_n y α . Por lo tanto,

$$x_n - \alpha = \frac{f(x_n)}{f'(\xi_n)}. \tag{1.9}$$

De (1.9) se tiene 3 casos:

1. Si $|f'(\alpha)| \approx 1$, entonces $|x_n - \alpha| \approx |f(x_n)| \leq \varepsilon$.
2. Si $|f'(\alpha)| \lll 1$, entonces $|x_n - \alpha| > \varepsilon$.

3. Si $|f'(\alpha)| \gg 1$, entonces $|x_n - \alpha| \ll \varepsilon$.

Consecuentemente, este criterio de parada no funciona para raíces múltiples, pues si $f'(\alpha) = 0$ entonces se cumple el caso (2), lo cual produciría una parada del proceso iterativo sin haber alcanzado la tolerancia prescrita. En cambio, para ceros simples, el criterio funciona si se cumple el caso (1) o el caso (3); aunque, para este último se estaría haciendo iteraciones innecesarias.

Criterio 2: $|x_{n+1} - x_n| \leq \varepsilon$. Sea $x_{n+1} = g(x_n)$ y $x_n \approx \alpha$, aplicando el Teorema del Valor Intermedio se tiene que existe ξ_n entre x_n y α tal que

$$x_{n+1} - \alpha = g(x_n) - g(\alpha) = g'(\xi_n)(x_n - \alpha) \approx g'(\alpha)(x_n - \alpha). \quad (1.10)$$

De (1.10) resulta,

$$\begin{aligned} x_{n+1} - x_n &= (x_{n+1} - \alpha) - (x_n - \alpha) \\ &\approx g'(\alpha)(x_n - \alpha) - (x_n - \alpha) = (x_n - \alpha)(g'(\alpha) - 1). \end{aligned}$$

Por lo tanto,

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

Luego, si $g'(\alpha) \approx 1$, entonces $|x_n - \alpha| \geq |x_{n+1} - x_n|$, para este caso podría no ser un buen criterio para aquellos métodos que presentan una convergencia lineal, pues $g'(\alpha) \neq 0$. Si $-1 < g'(\alpha) < 0$, entonces $|x_n - \alpha| \leq |x_{n+1} - x_n|$, por lo tanto, el criterio es bueno.

El caso del método de Newton,

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}, \quad \text{para cada } n \geq 0.$$

Por lo tanto,

$$x_{n+1} - x_n = -\frac{f(x_n)}{f'(x_n)} \approx -\frac{f(x_n)}{f'(\xi_n)} = -(x_n - \alpha) \quad \text{con } \xi_n \in \text{int}(x_n, \alpha),$$

donde $\text{int}(x_n, \alpha)$ denota el interior formado por x_n y α ($x_n < \alpha$ o $\alpha < x_n$). Si $|x_{n+1} - x_n| \leq \varepsilon$, entonces $|x_n - \alpha| \leq \varepsilon$. Así, este criterio funciona para este método.

1.5. Distribuciones fundamentales de muestreo

En estadística la **población**¹, también llamada universo o colectivo es el conjunto de elementos de referencia sobre el que se realizan las observaciones. El número de elementos

¹En el pasado la palabra **población** se refería a observaciones que se obtenían de estudios estadísticos con personas. En la actualidad, el estadístico utiliza el término para referirse a observaciones relevantes de cualquier cosa de interés, ya sea de grupos de personas, animales o todos los resultados posibles de algún sistema biológico o de ingeniería complicado.

o sujetos que componen una población estadística es igual o mayor que el número de elementos que se obtienen de ella en una muestra. En otras palabras, una **muestra** es un subconjunto de una población.

Si en la escuela hay 600 estudiantes que se quieren clasificar según su grupo sanguíneo, se dice que se tiene una población de tamaño 600. Algunas poblaciones finitas son tan grandes que en teoría se suponen infinitas.

Cada observación en una población es un valor de una variable aleatoria X que tiene alguna distribución de probabilidad $f(x)$. Si se inspeccionan artículos que salen de una línea de montaje en busca de defectos, entonces cada observación en la población podría ser un valor 0 ó 1 de la variable aleatoria binomial X con distribución de probabilidad

$$b(x; 1, p) = p^x q^{1-x}, \quad x = 0, 1$$

donde 0 indica un artículo no defectuoso y 1 indica uno defectuoso. Por supuesto, se supone que p , la probabilidad de que cualquier artículo esté defectuoso, permanece constante de una prueba a otra. En el experimento de tipo sanguíneo la variable aleatoria X representa el tipo de sangre al tomar un valor del 1 al 8. A cada estudiante se le asigna uno de los valores de la variable aleatoria discreta. Ahora, por ejemplo, las duraciones de la baterías de almacenamiento son valores que toma una variable aleatoria continua que quizá tiene una distribución normal.

En el campo de la inferencia estadística el estadístico se interesa en llegar a conclusiones con respecto a la población cuando es imposible o poco práctico observar todo el conjunto de observaciones que constituyen la población. Por ejemplo, al intentar determinar la longitud promedio de la vida de cierta marca de foco, sería imposible probar todos los focos si tenemos que venderlos. Los costos exorbitantes también pueden ser un factor prohibitivo para estudiar toda la población. Por tanto, debemos depender de un subconjunto de observaciones de la población para ayudarnos a hacer inferencias con respecto a la misma población. Esto nos lleva a considerar la noción de muestreo.

Si nuestras inferencias a partir de la muestra para la población han de ser válidas, debemos obtener muestras que sean representativas de la población. Con mucha frecuencia estamos tentados a elegir una muestra mediante la selección de los miembros más convenientes de la población. Tal procedimiento puede conducir a inferencias erróneas con respecto a la población. Cualquier procedimiento de muestreo que produzca inferencias que sobreestimen o subestimen de forma consistente alguna característica de la población se dice que está **sesgado**. Para eliminar cualquier posibilidad de sesgo en el procedimiento de muestreo, es deseable elegir una **muestra aleatoria** en el sentido de que las observaciones se realizan de forma independiente y al azar.

Para seleccionar una muestra aleatoria de tamaño n de una población $f(x)$, definamos la variable aleatoria X_i , $i = 1, 2, \dots, n$, que represente la i -ésima medición o valor de la muestra que observemos. Las variables aleatorias X_1, X_2, \dots, X_n constituirán entonces una muestra aleatoria de la población $f(x)$ con valores numéricos x_1, x_2, \dots, x_n

si las mediciones se obtienen al repetir el experimento n veces independientes bajo esencialmente las mismas condiciones. Debido a las condiciones idénticas bajo las que se seleccionan los elementos de la muestra, es razonable suponer que las n variables aleatorias X_1, X_2, \dots, X_n son independientes y que cada una tiene la misma distribución de probabilidad $f(x)$. Es decir, las distribuciones de probabilidad de X_1, X_2, \dots, X_n son, respectivamente, $f(x_1), f(x_2), \dots, f(x_n)$ y su distribución de probabilidad conjunta es

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2) \cdots f(x_n).$$

Definición 1.4 Sean X_1, X_2, \dots, X_n variables aleatorias independientes, cada una con la misma distribución de probabilidad $f(x)$. Definimos entonces a X_1, X_2, \dots, X_n como una **muestra aleatoria** de tamaño n de la población $f(x)$ y escribimos su distribución de probabilidad conjunta como

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2) \cdots f(x_n).$$

Si se realiza una selección aleatoria de $n = 8$ baterías de almacenamiento de un proceso de fabricación, que mantiene las mismas especificaciones, y registramos la duración de cada batería con la primera medición x_1 como un valor de X_1 , la segunda medición x_2 como un valor de X_2 , etcétera, entonces x_1, x_2, \dots, x_8 son los valores de la muestra aleatoria X_1, X_2, \dots, X_8 . Si suponemos que la población de duraciones de las baterías es normal, los valores posibles de cualquier $X_i, i = 1, 2, \dots, 8$, serán precisamente los mismos que los de la población original, y por ello X_i tiene la misma distribución normal idéntica que X .

1.5.1. Algunas estadísticas importantes

El propósito fundamental al seleccionar muestras aleatorias es obtener información acerca de los parámetros desconocidos de la población. Suponga, por ejemplo, que deseamos llegar a una conclusión con respecto a la proporción de personas bebedoras de café en Venezuela que prefieren cierta marca de café. Sería imposible preguntar a cada bebedor de café venezolano a fin de calcular el valor del parámetro p que representa la proporción de la población. En su lugar, se selecciona una muestra aleatoria grande y se calcula la proporción \hat{p} de personas en esta muestra que prefieren la marca de café en cuestión. El valor \hat{p} se utiliza ahora para hacer una inferencia con respecto a la proporción p verdadera.

Ahora bien, \hat{p} es una función de los valores observados en la muestra aleatoria; como son posible muchas muestras aleatorias a partir de la misma población, esperaríamos que \hat{p} variara algo de una muestra a otra. es decir, \hat{p} es un valor de una variable aleatoria que representamos con \hat{P} . Tal variable aleatoria se llama **estadística**.

Definición 1.5 Cualquier función de las variables aleatorias que forman una muestra aleatoria se llama *estadística*.

Tendencia central en la muestra.

Las estadísticas que por lo general se utilizan más para medir el centro de un conjunto de datos, acomodados en orden de magnitud, son la **media**, **mediana** y **moda**.

Definición 1.6 Si X_1, X_2, \dots, X_n representan una muestra aleatoria de tamaño n , entonces la media de la muestra se define mediante la estadística

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}.$$

Nótese que la estadística \bar{X} toma el valor $\bar{x} = \sum_{i=1}^n x_i/n$ cuando X_1 toma el valor x_1 , X_2 toma el valor x_2 y así sucesivamente. En la práctica al valor de una estadística por lo general se le da el mismo nombre de la estadística. Por ejemplo, el término, **media de la muestra** se aplica tanto a la estadística \bar{X} como a su valor calculado \bar{x} .

Ejemplo 1.3 Un inspector de alimentos examina una muestra aleatoria de siete latas de cierta marca de atún para determinar el porcentaje de impurezas externas. Se registran los siguientes datos: 1.8, 2.1, 1.6, 0.9, 2.7 y 1.8. Calcule la media de la muestra.

Solución. El valor observado \bar{x} de la estadística \bar{X} es

$$\bar{x} = \frac{1,8 + 2,1 + 1,7 + 1,6 + 0,9 + 2,7 + 1,8}{7} = 1,8 \%$$

La segunda estadística más útil para medir el centro de un conjunto de datos es la mediana.

Definición 1.7 Si X_1, X_2, \dots, X_n representa una muestra aleatoria de tamaño n , acomodada en orden creciente de magnitud, entonces la **mediana de la muestra** se define mediante la estadística

$$\tilde{X} = \begin{cases} X_{(n+1)/2} & \text{si } n \text{ es impar} \\ \frac{X_{n/2} + X_{(n/2)+1}}{2} & \text{si } n \text{ es par.} \end{cases}$$

Ejemplo 1.4 El número de barcos extranjeros que llegan a un puerto de la costa este en siete días seleccionados al azar son 8, 3, 9, 5, 6, 8 y 5. Encuentre la mediana de la muestra.

Solución. Al acomodar las observaciones en orden creciente de magnitud, obtenemos 3,5,5,6,8,8,9 y de aquí $\tilde{x} = 6$.

Ejemplo 1.5 *Se encuentra que el contenido de nicotina para una muestra aleatoria de seis cigarrillos de cierta marca son 2.3, 2.7, 2.5, 2.9, 3.1, 3.1 y 1.9 miligramos. Encuentre la mediana.*

Solución. Si acomodamos estos contenidos de nicotina en orden creciente de magnitud, obtenemos 1.9,2.3,2.5,2.7,2.9,3.1 y la mediana es entonces la media de 2.5 y 2.7. por tanto,

$$\tilde{x} = \frac{2,5 + 2,7}{2} = 2,6 \text{ miligramos.}$$

La tercera y última estadística para medir el centro de una muestra aleatoria que presentaremos es la moda, que se designa con la estadística M .

Definición 1.8 *Si X_1, X_2, \dots, X_n , no necesariamente diferentes, representan una muestra aleatoria de tamaño n , entonces la **moda** M es aquel valor de la muestra que ocurre más a menudo o con la mayor frecuencia. La moda puede no existir, y cuando existe no necesariamente es única.*

La moda no siempre existe. Esto es realmente cierto cuando todas las observaciones ocurren con la misma frecuencia. Para algunos conjuntos de datos habrá varios valores que ocurren con la mayor frecuencia, en cuyo caso tenemos más de una moda.

Ejemplo 1.6 *Si las donaciones de una muestra aleatoria de residentes de Fairway Forest para la Asociación del Pulmón de Virginia se registraron como 9, 10, 5, 9, 9, 7, 8, 6, 10 y 11 dólares, entonces la moda es $m = \$9$, el valor ocurre con la mayor frecuencia.*

Ejemplo 1.7 *El número de películas vistas el mes pasado por una muestra aleatoria de 12 estudiantes de preparatoria se registraron como sigue: 2, 0, 3, 1, 2, 4, 2, 5, 4, 0, 1 y 4. En este caso hay dos modas, 2 y 4, pues ambas ocurren con la mayor frecuencia. Se dice que la distribución es binomial.*

Ejemplo 1.8 *No existe moda para el contenido de nicotina del ejemplo 1.5, puesto que cada medición ocurre sólo una vez.*

En resumen, consideremos los méritos relativos de la media, mediana y moda. La media es la medida de localización central más comúnmente utilizada en estadística. Emplea toda la información disponible. Las distribuciones de medias que se obtienen en muestreos repetidos de una población son bien conocidos, y en consecuencia los métodos que se utilizan en la inferencia estadística para estimar μ se basa en la media de la

muestra. La única desventaja real de la media es que puede resultar afectada de manera adversa por valores extremos.

La mediana tiene la ventaja de ser fácil de calcular si el número de observaciones es relativamente pequeño. No resulta influida por valores extremos. Si intentamos estimar el centro de una población con base a un valor de la muestra, la media es más estable que la mediana. por tanto una media de la muestra probablemente esté más cercana que la mediana de la muestra a la media de la población.

La moda es la medición que se utiliza menos de las tres. Para conjuntos pequeños de datos su valor casi no tiene utilidad si es que, de hecho, existe. Sólo en el caso de una gran cantidad de datos tiene un sentido significativo. Sus dos ventajas principales son que 1) no requiere cálculo y 2) se puede usar para datos cualitativos así como cuantitativos. Entonces, si correr es la forma de ejercicio que la mayoría de las personas expresa como preferida, decimos que correr es la **selección modal**.

Variabilidad de la muestra.

Las tres medidas de localización central que se definieron antes no dan por sí mismas una descripción adecuada de nuestros datos. Necesitamos conocer como se dispersan las observaciones del promedio. Se remite al lector al Capítulo 1 de [21] para una discusión más amplia. Es muy posible tener dos conjuntos de observaciones con la misma media o mediana que difieren de manera considerable en la variabilidad de sus mediciones alrededor del promedio.

Considere las siguientes mediciones, en litros, para dos muestras de jugo de naranja embotellado por las compañías A y B:

Muestra A	0.97	1.00	0.94	1.03	1.11
Muestra B	1.06	1.01	0.88	0.91	1.14

Ambas muestras tienen la misma media, 1.00 litros. Es bastante obvio que la compañía A embotella el jugo de naranja con un contenido más uniforme que la B. Decimos que la variabilidad o la dispersión de las observaciones del promedio es menor para la muestra A que para la muestra B. Por tanto, al comprar jugo de naranja, tendremos más confianza de que la botella seleccionada esté más cerca del promedio anunciado si la compramos de la compañía A.

Las estadísticas más importantes para medir la variabilidad de una muestra aleatoria son el **rango** y la **varianza**. La más simple de calcular es el rango:

Definición 1.9 *El **rango** de una muestra aleatoria X_1, X_2, \dots, X_n se define con la estadística $X_{(n)} - X_{(1)}$, donde $X_{(n)}$ y $X_{(1)}$ son, respectivamente, las observaciones más grande y más pequeña en la muestra.*

En el caso de las compañías que embotellan jugo de naranja, el rango para la compañía A es 0.17 litros comparado con un rango de 0.26 litros para la B, lo que indica una mayor dispersión de los valores para la compañía B.

El rango puede ser una medida pobre de la variabilidad, en particular si el tamaño de la muestra o población es grande. Considera sólo los valores extremos y no nos dice nada acerca de la distribución de los valores intermedios. Considere, por ejemplo, los dos conjuntos de datos siguientes, ambos con un rango de 12:

	Datos Observados								
Conjunto A	3	4	5	6	8	9	10	12	15
Conjunto B	3	7	7	7	8	8	8	9	15

En el primer conjunto la media y la mediana son ocho, pero los números varían a lo largo de todo el intervalo de 3 a 15. En el segundo conjunto la media y la mediana también son ocho, pero la mayor parte de los valores están más cerca del centro de los datos. Aunque el rango falla al medir esta variabilidad entre la observación superior y la inferior, tiene algunas aplicaciones útiles. En la industria el rango se puede predeterminar al especificar por adelantado que una medición particular de los artículos que salen de una línea de producción deba caer dentro de cierto intervalo. En tanto que todas las mediciones caigan dentro del intervalo especificado, se dice que el proceso está bajo control.

Para compensar la desventaja del rango, consideraremos una medida de variabilidad, a saber, la **varianza de la muestra**, que considera la posición de cada observación en relación con la media de la muestra.

Definición 1.10 Si X_1, X_2, \dots, X_n representa una muestra aleatoria de tamaño n , entonces la varianza de la muestra se define con la estadística

$$S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}.$$

El valor calculado de S^2 para una muestra dada se denota con s^2 .

Ejemplo 1.9 Una comparación de precios de café en cuatro tiendas de abarrotes seleccionadas al azar en San Diego muestra aumentos en comparación con el mes anterior de 12, 15, 17 y 20 centavos para una bolsa de una libra. Encuentre la varianza de esta muestra aleatoria de aumentos de precio.

Solución. Al calcular la media de la muestra, obtenemos

$$\bar{x} = \frac{12 + 15 + 17 + 20}{4} = 16 \text{ centavos.}$$

Por tanto,

$$\begin{aligned}
 s^2 &= \frac{\sum_{i=1}^4 (x_i - 16)^2}{4} \\
 &= \frac{(12 - 16)^2 + (15 - 16)^2 + (17 - 16)^2 + (20 - 16)^2}{4} \\
 &= \frac{(-4)^2 + (-1)^2 + (1)^2 + (4)^2}{4} = \frac{34}{4}.
 \end{aligned}$$

Si \bar{x} es un número decimal que se redondea, acumulamos un error grande cuando utilizamos la fórmula de la varianza de la muestra de la Definición 1.10. Para evitar esto, derivemos una fórmula muy útil en el siguiente teorema.

Teorema 1.9 *Si S^2 es la varianza de una muestra aleatoria de tamaño n , podemos escribir*

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}{n^2}.$$

Demostración. Por definición,

$$\begin{aligned}
 S^2 &= \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n} = \frac{\sum_{i=1}^n (X_i^2 - 2\bar{X}X_i + \bar{X}^2)}{n} \\
 &= \frac{\sum_{i=1}^n X_i^2 - 2\bar{X} \sum_{i=1}^n X_i + n\bar{X}^2}{n}.
 \end{aligned}$$

Al reemplazar \bar{X} por $\sum_{i=1}^n X_i/n$, se obtiene la fórmula de cálculo más útil

$$S^2 = \frac{n \sum_{i=1}^n X_i^2 - \left(\sum_{i=1}^n X_i \right)^2}{n^2}$$

□.

Definición 1.11 *La desviación estándar de la muestra, que se denota con S , es la raíz cuadrada positiva de la varianza de la muestra.*

Ejemplo 1.10 *Encuentre la varianza de los datos 3, 4, 5, 6, 6 y 7, que representan el número de truchas atrapadas por una muestra aleatoria de seis pescadores el 19 de junio de 1996, en el lago Muskoka.*

Solución. Encontramos que $\sum_{i=1}^6 x_i^2 = 171$, $\sum_{i=1}^6 x_i = 31$, $n = 6$. De aquí,

$$s^2 = \frac{(6)(171) - (31)^2}{36} = \frac{65}{36}.$$

1.5.2. Distribuciones muestrales.

Definición 1.12 La distribución de probabilidad de una estadística se llama **distribución muestral**.

La distribución de probabilidad de \bar{X} se llama **distribución muestral de la media**. La distribución muestral de una estadística depende del tamaño de la población, el tamaño de las muestras y el método de elección de las muestras.

La distribución muestral de \bar{X} con tamaño muestral n es la distribución que resulta cuando un experimento se lleva a cabo una y otra vez (siempre con tamaño muestral n) y resultan los diferentes valores de \bar{X} . Esta distribución muestral, entonces describe la variabilidad de los promedios muestrales alrededor de la media de la población μ . Se aplica el mismo principio en el caso de la distribución de S^2 . La distribución muestral produce información acerca de la variabilidad de los valores de s^2 alrededor de σ^2 en experimentos que se repiten.

1.5.3. Distribuciones muestrales de medias.

Suponga que una muestra aleatoria de n observaciones se toma de una población normal con media μ y varianza σ^2 . Cada observación X_i , $i = 1, 2, \dots, n$, de la muestra aleatoria tendrá entonces la misma distribución normal que la población que se muestrea. De aquí, por la propiedad reproductiva de la distribución normal que se establece en el Teorema 7.11 de [21, C.7, pág. 194], concluimos que

$$\bar{X} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

tiene distribución normal con media

$$\mu_{\bar{X}} = \frac{\mu + \mu + \dots + \mu}{n} = \mu$$

y varianza

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2 + \sigma^2 + \dots + \sigma^2}{n^2} = \frac{\sigma^2}{n}.$$

Si tomamos muestras de una población con distribución desconocida, finita o infinita, la distribución muestral de \bar{X} aún será aproximadamente normal con media μ y varianza σ^2/n siempre que el tamaño de la muestra sea grande. Este asombroso resultado es una consecuencia inmediata del siguiente teorema, que se llama teorema del límite central.

Teorema 1.10 (Teorema del límite central) *Si \bar{X} es la media de una muestra aleatoria de tamaño n tomada de una población con media μ y varianza finita σ^2 , entonces la forma límite de la distribución de*

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}},$$

conforme $n \rightarrow \infty$, es la distribución normal estándar $n(z; 0, 1)$.

La aproximación normal para \bar{X} por lo general será buena si $n \geq 30$ sin importar la forma de la población. Si $n < 30$, la aproximación es buena sólo si la población no es muy diferente de una distribución normal y, como se estableció antes, si se sabe que la población es normal, la distribución muestral de \bar{X} seguirá una distribución normal exacta, no importa qué tan pequeño sea el tamaño de las muestras.

Suponga que tenemos dos poblaciones, la primera con media μ_1 y varianza σ_1^2 , y la segunda con media μ_2 y varianza σ_2^2 . Representemos con la estadística \bar{X}_1 la media de una muestra aleatoria de tamaño n_1 seleccionada de la primera población, y con la estadística \bar{X}_2 la media de una muestra aleatoria seleccionada de la segunda población, independiente de la muestra de la primera población. ¿Qué podemos decir acerca de la distribución de muestreo de la diferencia $\bar{X}_1 - \bar{X}_2$ para muestras repetidas de tamaño n_1 y n_2 ? De acuerdo con el Teorema 1.10, las variables \bar{X}_1 y \bar{X}_2 están distribuidas aproximadamente de forma normal con medias μ_1 y μ_2 y varianzas σ_1^2/n_1 y σ_2^2/n_2 , respectivamente. Esta aproximación mejora conforme n_1 y n_2 aumentan. Al elegir muestras independientes de las dos poblaciones las variables \bar{X}_1 y \bar{X}_2 serán independientes, y entonces con el uso del Teorema 7.11 de [21, C.7, pág. 194], con $a_1 = 1$ y $a_2 = -1$, podemos concluir que $\bar{X}_1 - \bar{X}_2$ está distribuida aproximadamente de forma normal con media

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

y varianza

$$\sigma_{\bar{X}_1 - \bar{X}_2}^2 = \sigma_{\bar{X}_1}^2 + \sigma_{\bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

Teorema 1.11 *Si se extraen al azar muestras independientes de tamaño n_1 y n_2 de dos poblaciones, discretas o continuas, con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente, entonces la distribución muestral de las diferencia de las medias, $\bar{X}_1 - \bar{X}_2$, está distribuida aproximadamente de forma normal con media y varianza dadas por*

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2 \quad y \quad \sigma_{\bar{X}_1 - \bar{X}_2}^2 = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}.$$

De aquí,

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}}.$$

es aproximadamente una variable normal estándar.

Si n_1 y n_2 son mayores que o iguales a 30, la aproximación normal para la distribución $\bar{X}_1 - \bar{X}_2$ es muy buena sin importar las formas de las dos poblaciones. Sin embargo, aun cuando n_1 y n_2 sean menores que 30, la aproximación normal es razonablemente buena excepto cuando las poblaciones no son definitivamente normales. Por supuesto, si ambas poblaciones son normales, entonces $\bar{X}_1 - \bar{X}_2$ tiene una distribución normal sin importar cuáles son los tamaños de n_1 y n_2 .

1.6. Problemas de estimación de una y dos muestras.

El propósito de estas presentaciones es la construcción de las bases que permitan extraer conclusiones acerca de los parámetros de la población a partir de los datos experimentales. Por ejemplo, el teorema del límite central proporciona información sobre la distribución de la media muestral \bar{X} . La distribución involucra la media de la población μ . Así, cualesquiera conclusiones que se extraigan con respecto a μ a partir de un promedio muestral observado debe depender del conocimiento de su distribución muestral.

En esta sección se comienza por señalar el propósito de la inferencia estadística. Se sigue con la presentación del problema de la **estimación de los parámetros de la población**.

Inferencia estadística.

La teoría de la **inferencia estadística** consiste en aquellos métodos por los que se realizan inferencias o generalizaciones acerca de una población.

La inferencia estadística se puede dividir en dos áreas principales: **estimación** y **pruebas de hipótesis**. En éste trabajo nos ocuparemos de los métodos de estimación.

Métodos clásicos de estimación.

Una **estimación puntual** de algún parámetro de la población θ es un solo valor $\hat{\theta}$ de una estadística $\hat{\Theta}$. Por ejemplo, el valor \bar{x} de la estadística \bar{X} , que se calcula a partir de la muestra de tamaño n , es una estimación puntual del parámetro poblacional μ .

No se espera que un estimador realice la estimación del parámetro poblacional sin error. No esperamos que \bar{X} estime μ exactamente, sino que en realidad esperamos que no esté muy alejado.

Estimador insesgado.

¿Cuáles son las propiedades deseables de una “buena” función de decisión que influirán sobre nosotros para elegir un estimador en lugar de otro? Sea $\hat{\Theta}$ un estimador

cuyo valor $\hat{\theta}$ es una estimación puntual de algún parámetro poblacional desconocido θ . Ciertamente, desearíamos que la distribución muestral de $\hat{\Theta}$ tuviera una media igual al parámetro estimado. Se dice que un estimador que posee esta propiedad es **insesgado**.

Varianza de un estimador puntual.

Si $\hat{\Theta}_1$ y $\hat{\Theta}_2$ son dos estimadores insesgados del mismo parámetro poblacional θ , elegiríamos el estimador cuya distribución muestral tuviera la menor varianza. De aquí, si $\sigma_{\hat{\Theta}_1}^2 < \sigma_{\hat{\Theta}_2}^2$, decimos que $\hat{\Theta}_1$ es un **estimador más eficiente** de θ que $\hat{\Theta}_2$.

Para poblaciones normales se puede mostrar que \bar{X} y \tilde{X} son estimadores insesgados de la media de la población μ , pero la varianza de \bar{X} es más pequeña que la varianza de \tilde{X} . De esta manera las estimaciones \bar{x} y \tilde{x} serán, en promedio, iguales a la media poblacional μ , pero \bar{x} es probable que esté más cerca de μ para una muestra dada, y por ello \bar{X} es más eficiente que \tilde{X} .

Es improbable que incluso el estimador insesgado más eficiente estime el parámetro poblacional con exactitud. Es cierto que nuestra precisión aumenta con muestras grandes, pero no hay razón por la cual debemos esperar que una estimación **puntual** de una muestra dada sea exactamente igual al parámetro poblacional que se supone estima. Hay muchas situaciones en las que es preferible determinar un intervalo dentro del cual esperaríamos encontrar el valor del parámetro. Tal intervalo se llama **intervalo de estimación**.

Estimación por intervalo.

Una estimación por intervalo de un parámetro poblacional θ es un intervalo de la forma $\hat{\theta}_L < \theta < \hat{\theta}_U$, donde $\hat{\theta}_L$ y $\hat{\theta}_U$ dependen del valor de la estadística $\hat{\Theta}$ para una muestra particular y también de la distribución de muestreo de $\hat{\Theta}$. A medida que el tamaño de la muestra aumenta, sabemos que $\sigma_{\bar{X}}^2 = \sigma^2/n$ disminuye, y en consecuencia es probable que nuestra estimación esté cercana al parámetro μ , lo que tiene como resultado un intervalo más pequeño. De esta manera el intervalo estimado indica, por su longitud, la precisión de la estimación puntual.

Como muestras distintas por lo general darán valores diferentes de $\hat{\Theta}$ y, por tanto, valores diferentes de $\hat{\theta}_L$ y $\hat{\theta}_U$. De la distribución muestral de $\hat{\Theta}$ seremos capaces de determinar $\hat{\theta}_L$ y $\hat{\theta}_U$ de modo que $P(\hat{\theta}_L < \theta < \hat{\theta}_U)$ sea igual a algún valor fraccional positivo que queramos especificar. Si, por ejemplo, encontramos $\hat{\theta}_L$ y $\hat{\theta}_U$ tales que

$$P(\hat{\theta}_L < \theta < \hat{\theta}_U) = 1 - \alpha,$$

para $0 < \alpha < 1$, tenemos entonces una probabilidad de $1 - \alpha$ de seleccionar una variable aleatoria que produzca un intervalo que contenga θ . El intervalo $\hat{\theta}_L < \theta < \hat{\theta}_U$, que se calcula a partir de la muestra seleccionada, se llama entonces **intervalo de confianza**

de $(1 - \alpha)100\%$, la fracción $1 - \alpha$ se llama **coeficiente de confianza** o **grado de confianza**, y los extremos, $\hat{\theta}_L$ y $\hat{\theta}_U$, se denominan **límites de confianza** inferior y superior. Así, cuando $\alpha = 0.05$, tenemos un intervalo de confianza de 95% , y cuando $\alpha = 0.01$ obtenemos un intervalo de confianza más amplio de 99% . Entre más amplio sea el intervalo de confianza podemos tener más confianza de que el intervalo dado contenga el parámetro desconocido. Por supuesto, es mejor tener una confianza de 95% de que la vida promedio de cierto transistor de televisor está entre seis y siete años que tener una confianza de 99% de que está entre tres y 10 años. Idealmente, preferimos un intervalo corto con un alto grado de confianza. Algunas veces las restricciones en el tamaño de nuestra muestra nos impiden tener intervalos cortos sin sacrificar algo de nuestro grado de confianza.

Para el caso de intervalos de confianza para una sola muestra (estimación de la media), se remite al lector a [21, Cap. 9, pág. 243].

Intervalos de confianza para la diferencia de dos medias poblacionales.

Si tenemos dos poblaciones con medias μ_1 y μ_2 y varianzas σ_1^2 y σ_2^2 , respectivamente, un estimador puntual de la diferencia entre μ_1 y μ_2 está dado por la estadística $\bar{X}_1 - \bar{X}_2$. Por tanto, para obtener una estimación puntual de $\mu_1 - \mu_2$, seleccionamos dos muestras aleatorias independientes, una de cada población, de tamaño n_1 y n_2 , y calculamos la diferencia $\bar{x}_1 - \bar{x}_2$, de las medias muestrales. Claramente, debemos considerar las distribuciones muestrales de $\bar{X}_1 - \bar{X}_2$.

De acuerdo con el Teorema 1.11, podemos esperar que la distribución muestral de $\bar{X}_1 - \bar{X}_2$ esté distribuida de forma aproximadamente normal con media $\mu_{\bar{X}_1 - \bar{X}_2} = \mu_1 - \mu_2$ y desviación estándar $\sigma_{\bar{X}_1 - \bar{X}_2} = \sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}$. Por tanto, podemos asegurar con una probabilidad de $1 - \alpha$ que la variable normal estándar

$$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}},$$

caerá entre $-z_{\alpha/2}$ y $z_{\alpha/2}$. Con referencia, a la Figura 1.4, escribimos

$$P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha.$$

Al sustituir para Z , establecemos de manera equivalente que

$$P \left[-z_{\alpha/2} < \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)}} < z_{\alpha/2} \right] = 1 - \alpha,$$

que conduce al siguiente intervalo de confianza de $(1 - \alpha)100\%$ para $\mu_1 - \mu_2$.

De nuevo, se invoca el teorema del límite central y de esta forma el intervalo de confianza que sigue es aproximado a menos que las dos poblaciones sean normales.

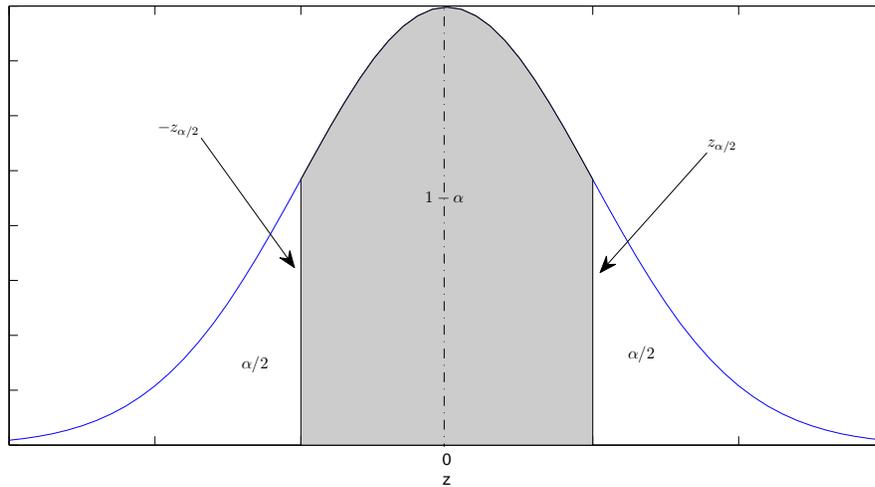


Figura 1.4: $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$

Intervalo de confianza para $\mu_1 - \mu_2$; con σ_1^2 y σ_2^2 conocida. Si \bar{x}_1 y \bar{x}_2 son las medias de muestras aleatorias independientes de tamaño n_1 y n_2 de poblaciones con varianzas conocidas σ_1^2 y σ_2^2 , respectivamente, un intervalo de confianza de $(1 - \alpha)100\%$ para $\mu_1 - \mu_2$ está dado por

$$(\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}},$$

donde $z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha.

Ejemplo 1.11 *Se lleva a cabo un experimento en que se comparan dos tipos de motores, A y B. Se mide el rendimiento en millas por galón de gasolina. Se realizan 50 experimentos con el motor tipo A y 75 con el motor tipo B. La gasolina que se utiliza y las demás condiciones se mantienen constantes. El rendimiento promedio de gasolina para el motor A es de 36 millas por galón y el promedio para el motor B es 42 millas por galón. Encuentre un intervalo de confianza de 96% sobre $\mu_B - \mu_A$, donde μ_B y μ_A son el rendimiento de gasolina medio poblacional para los motores B y A, respectivamente. Suponga que las desviaciones estándar poblacionales son seis y ocho para los motores A y B, respectivamente.*

Solución. La estimación puntual de $\mu_B - \mu_A$ es $\bar{x}_B - \bar{x}_A = 42 - 36 = 6$. Con el uso de $\alpha = 0.04$, encontramos $z_{0,02} = 2.05$ de la tabla A.3 de [21, pág. 681]. De aquí con la

sustitución en la fórmula anterior, el intervalo de confianza de 96 % es

$$6 - 2,05\sqrt{\frac{64}{75} + \frac{36}{50}} < \mu_1 - \mu_2 < 6 + 2,05\sqrt{\frac{64}{75} + \frac{36}{50}},$$

o simplemente $3.43 < \mu_B - \mu_A < 8.57$. \square

Si no se está dispuesto a suponer normalidad, muestras grandes (digamos mayores que 30) permitirán el uso de s_1 y s_2 en lugar de σ_1 y σ_2 , respectivamente, con la explicación de que $s_1 \simeq \sigma_1$ y $s_2 \simeq \sigma_2$. De nuevo, por supuesto, el intervalo de confianza es aproximado.

Intervalo de confianza para $\mu_1 - \mu_2$ con distribución y varianzas desconocidas.

Si \bar{x}_1 y \bar{x}_2 son las medias de muestras aleatorias independientes de tamaño n_1 y $n_2 \geq 30$ de poblaciones con distribución y varianzas desconocidas, se puede construir un intervalo de confianza de $(1 - \alpha)100\%$ para $\mu_1 - \mu_2$ de la siguiente manera:

$$\mathbf{LI} < \mu_1 - \mu_2 < \mathbf{LS}.$$

$$\mathbf{LI}: \quad = \quad (\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}},$$

$$\mathbf{LS}: \quad = \quad (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}.$$

donde $z_{\alpha/2}$ es el valor z que deja un área de $\alpha/2$ a la derecha y s_1 y s_2 son las desviaciones estándar de cada una de las muestras.

Observación 1.2 Si $0 < \mathbf{LI} < \mu_1 - \mu_2 < \mathbf{LS}$ o $\mathbf{LI} < \mu_1 - \mu_2 < \mathbf{LS} < 0$ se puede concluir con un $(1 - \alpha)100\%$ de confianza que $\mu_2 < \mu_1$ o $\mu_1 < \mu_2$, es decir μ_1 y μ_2 son diferentes.

Esta observación se debe a lo siguiente:

- Si $0 < \mathbf{LI} < \mu_1 - \mu_2 < \mathbf{LS}$, es decir, $0 < \mu_1 - \mu_2$, entonces $\mu_1 > \mu_2$ con $(1 - \alpha)100\%$ de confianza.
- Si $\mathbf{LI} < \mu_1 - \mu_2 < \mathbf{LS} < 0$, es decir, $\mu_1 - \mu_2 < 0$, entonces $\mu_1 < \mu_2$ con $(1 - \alpha)100\%$ de confianza.
- En el caso de $\mathbf{LI} < \mu_1 - \mu_2 < \mathbf{LS}$, con $\mathbf{LI} < 0$ y $\mathbf{LS} > 0$ no se puede asegurar que $\mu_1 > \mu_2$ o $\mu_1 < \mu_2$.

Capítulo 2

Métodos Iterativos para resolver ecuaciones no lineales

Encontrar una o más raíces de una ecuación no lineal de la forma, $f(x) = 0$, donde $f : I \rightarrow \mathbb{R}$ es una función real de variable real, es uno de los problemas más comunes que ocurren en las matemáticas aplicadas. Los métodos analíticos para resolver tales problemas son difíciles de encontrar o en línea general no existen. Por lo tanto, para resolver este problema resulta necesario aplicar técnicas numéricas basadas en procedimientos iterativos como por ejemplo, Newton-Raphson. En los últimos años se han definido métodos iterativos que mejoran, en cierta forma, la precisión de los métodos clásicos; ver, por ejemplo, [4]-[11].

En este capítulo estudiaremos los métodos iterativos presentados en las referencias [4]-[11]. Se presenta la construcción de cada método y se estudia, en la mayoría de los métodos, el análisis de convergencia.

A lo largo de este trabajo se considera la ecuación no lineal

$$f(x) = 0. \tag{2.1}$$

donde f es una función real de variable real y se asume que α es una raíz simple de (2.1).

2.1. Método iterativo JRS

El método es introducido por J. R. Sharma en [4]. El método es un miembro de la familia de métodos iterativos (2.7) en la que $a_n = 1$.

Construcción del método. El proceso de aproximación se lleva a cabo considerando la ecuación cuadrática en x y y en la siguiente forma:

$$x^2 + ay^2 + bx + cy + d = 0. \quad (2.2)$$

Geoméricamente, la ecuación (2.2) representa las siguientes curvas cuadráticas(ver [16, p.730])

1. Un círculo, si $a=1$,
2. Una parábola, si $a=0$,
3. Una elipse, si a es una cantidad positiva distinta de 1 y
4. Una hipérbola, si a es una cantidad negativa.

La idea es construir una familia de métodos de tercer orden por derivación geométrica con la ayuda de todas las posibles curvas cuadráticas dependiendo del parámetro a en la ecuación (2.2).

Por conveniencia y tomando en cuenta que la ecuación cuadrática (2.2) pasa por el punto $(x_n, y(x_n))$, esta puede ser escrita en la forma equivalente

$$Q(x, y) = (x - x_n)^2 + a_n(y - y(x_n))^2 + b_n(x - x_n) + c_n(y - y(x_n)) + d_n = 0. \quad (2.3)$$

Los valores de b_n , c_n y d_n son determinados en términos de a_n por las condiciones de que f y y , f' y y' , f'' y y'' coinciden en el punto $(x_n, y(x_n))$. Imponiendo las condiciones tangenciales

$$y(x_n) = f(x_n), \quad y'(x_n) = f'(x_n) \quad y \quad y''(x_n) = f''(x_n); \quad (2.4)$$

y resolviendo las ecuaciones

$$Q(x_n, y(x_n)) = 0, \quad Q'(x_n, y(x_n)) = 0 \quad y \quad Q''(x_n, y(x_n)) = 0, \quad (2.5)$$

se obtiene

$$\begin{aligned} b_n &= \frac{2(1 + a_n f'(x_n)^2) f'(x_n)}{f''(x_n)}, \\ c_n &= -\frac{2(1 + a_n f'(x_n)^2)}{f''(x_n)}, \\ d_n &= 0. \end{aligned} \quad (2.6)$$

Note que el sistema de ecuaciones (2.5) tiene una única solución, si y sólo si, $f''(x_n) \neq 0$. Sustituyendo en (2.3) los valores de b_n , c_n y d_n obtenidos en (2.6) y tomando el

próximo iterado x_{n+1} como el punto de intersección de $Q(x, y) = 0$ con el eje x , eso significa resolver $Q(x_{n+1}, 0) = 0$, se obtiene

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \frac{L_1(x_n)}{L_2(x_n) \pm \sqrt{(L_2(x_n))^2 - L(x_n)(L_1(x_n))}}, \quad (2.7)$$

donde

$$\begin{aligned} L(x_n) &= \frac{f(x_n)f''(x_n)}{f'(x_n)^2}, \\ L_1(x_n) &= 2 + a_n f'(x_n)^2 (2 + L(x_n)), \\ L_2(x_n) &= 1 + a_n f'(x_n)^2 \end{aligned}$$

y a_n es un parámetro dependiendo de n .

El signo del denominador se elige de la siguiente manera:

- el signo $+$ si $1 + a_n f'(x_n)^2 > 0$
- el signo $-$ si $1 + a_n f'(x_n)^2 < 0$
- cualquiera de los dos signos si $1 + a_n f'(x_n)^2 = 0$

Análisis de convergencia. El método JRS definido en (2.7) requiere de una evaluación de f , f' y f'' en cada iteración. El método presenta al menos un tercer orden de convergencia. La demostración de la convergencia es realizada en el siguiente teorema. La eficiencia computacional del método es: $r^{1/d} = 3^{1/3} \approx 1,442$.

Teorema 2.1 *Sea α una raíz simple de $f \in C^5(I)$, con I un entorno de α . Supongamos que $f'(\alpha) \neq 0$ y $a_n \neq -1/f'(x_n)^2$. Entonces existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$ la familia de métodos iterativos definidos por (2.7) convergen a α y además tienen un orden de convergencia de al menos tres.*

Demostración: Para realizar la demostración, se usa la función auxiliar h definida por las siguientes expresiones:

$$h(x) = \frac{B(x)}{C(x)},$$

con

$$\begin{aligned} B(x) &= 2A(x) + a_n f'(x)^2 L(x), \\ A(x) &= 1 + a_n f'(x)^2, \\ L(x) &= \frac{f(x)f''(x)}{f'(x)^2}, \\ C(x) &= A(x) \pm \sqrt{A(x)^2 - L(x)B(x)}. \end{aligned}$$

Así, el proceso iterativo (2.7) se puede reescribir como

$$g(x_n) = x_n - h(x_n) \frac{f(x_n)}{f'(x_n)} \quad (2.8)$$

Evaluando A , L , C y B y simplificando se obtiene

$$\begin{aligned} A(\alpha) &= 1 + a_n f'(\alpha)^2, \\ L(\alpha) &= 0, \\ C(\alpha) &= 2A(\alpha), \end{aligned} \quad (2.9)$$

$$B(\alpha) = 2A(\alpha). \quad (2.10)$$

Al evaluar h en α se obtiene

$$h(\alpha) = \frac{B(\alpha)}{C(\alpha)} = 1. \quad (2.11)$$

Si se deriva h se obtiene

$$h'(x) = \frac{B'(x)C(x) - B(x)C'(x)}{C(x)^2}. \quad (2.12)$$

Al evaluar B' y C' en α se obtiene

$$B'(\alpha) = 5a_n f'(\alpha) f''(\alpha), \quad (2.13)$$

$$C'(\alpha) = 2A'(\alpha) - L'(\alpha), \quad (2.14)$$

con

$$A'(\alpha) = 2a_n f'(\alpha) f''(\alpha),$$

$$L'(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}.$$

Sustituyendo (2.9), (2.10), (2.13), (2.14) y $C(\alpha)^2$ en (2.12) y simplificando los cálculos resulta

$$h'(\alpha) = \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)}. \quad (2.15)$$

Los valores (2.11) y (2.15) implican que h satisface las hipótesis del Teorema 1.8 el cual garantiza que existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$ el proceso iterativo definido por (2.8) converge a α y además es de al menos orden tres. \square

2.2. Método iterativo MKN1

El método MKN1 es introducido por Muhammad Noor y Khalida Noor en [5]. Los autores usan la técnica de descomposición de Noor[17] para sugerir una familia de métodos iterativos de un paso, dos paso y tres paso para resolver ecuaciones no lineales.

Construcción del método Se asume que γ es una condición inicial suficientemente cercana a α . Se puede reescribir la ecuación no lineal (2.1) usando las series de Taylor como:

$$\begin{aligned} f(x) &= f(\gamma) + (x - \gamma)f'(\gamma) + \frac{(x - \gamma)^2 f''(\gamma)}{2} + \dots, \\ f(\gamma) + (x - \gamma)f'(\gamma) + \frac{(x - \gamma)^2 f''(\gamma)}{2} &= 0, \end{aligned} \quad (2.16)$$

donde γ es una aproximación inicial a un cero de (2.1).

Podemos reescribir la ecuación (2.16) en la siguiente forma:

$$x = c + N(x), \quad (2.17)$$

donde

$$c = \gamma - \frac{f(\gamma)}{f'(\gamma)}; \quad (2.18)$$

y

$$N(x) = -\frac{(x - \gamma)^2}{2f'(\gamma)} f''(\gamma). \quad (2.19)$$

Notemos que el operador $N(x)$ es no lineal. Ahora construimos una sucesión de métodos iterativos de orden superior usando el siguiente método de descomposición, que es esencialmente debido a Noor[17] y Noor y Noor[18]. Esta descomposición del operador no lineal $N(x)$ es bastante diferente que la descomposición de Adomian. La idea principal de esta técnica es buscar una solución de la ecuación (2.17) mediante una serie de la forma:

$$x = \sum_{i=0}^{\infty} x_i. \quad (2.20)$$

El operador no lineal N puede ser descompuesto como

$$N\left(\sum_{i=0}^{\infty} x_i\right) = N(x_0) + \sum_{i=1}^{\infty} \left\{ N\left(\sum_{j=0}^i x_j\right) \right\}. \quad (2.21)$$

Combinando (2.17),(2.20) y (2.21), obtenemos

$$x = \sum_{i=0}^{\infty} x_i = c + N(x_0) + \sum_{i=1}^{\infty} \left\{ N\left(\sum_{j=0}^i x_j\right) \right\}.$$

Así se obtiene el siguiente esquema iterativo:

$$\begin{aligned}
 x_0 &= c, \\
 x_1 &= N(x_0), \\
 x_2 &= N(x_0 + x_1), \\
 &\vdots \\
 x_{n+1} &= N(x_0 + x_1 + \cdots + x_n), \quad n = 1, 2, \dots;
 \end{aligned}
 \tag{2.22}$$

Entonces

$$\begin{aligned}
 x_1 + x_2 + \cdots + x_{m+1} &= N(x_0) + N(x_0 + x_1) + \cdots \\
 &\quad + N(x_0 + x_1 + \cdots + x_m), \quad m = 1, 2, \dots.
 \end{aligned}$$

y

$$x = c + \sum_{i=1}^{\infty} x_i.
 \tag{2.23}$$

Se sigue de (2.18),(2.19),(2.20) y (2.22), que

$$x_0 = c = \gamma - \frac{f(\gamma)}{f'(\gamma)}
 \tag{2.24}$$

y

$$x_1 = N(x_0) = -\frac{(x_0 - \gamma)^2}{2f'(\gamma)} f''(\gamma).
 \tag{2.25}$$

De (2.22),(2.23) y (2.24), obtenemos

$$x = c = x_0 = \gamma - \frac{f(\gamma)}{f'(\gamma)}.$$

Esto nos permite sugerir el siguiente método iterativo de un paso para resolver la ecuación no lineal (2.1).

Algoritmo 2.2.1 Para un x_0 dado, calcular la solución aproximada x_{n+1} por el esquema iterativo

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

El método que se sugiere en el Algoritmo 2.2.1 es el Método de Newton que tiene segundo orden de convergencia.

De nuevo usando (2.22) y (2.23)-(2.25), concluimos que

$$x = c + x_1 = x_0 + N(x_0) = \gamma - \frac{f(\gamma)}{f'(\gamma)} - \frac{(x - \gamma)^2}{2f'(\gamma)} f''(\gamma).$$

Usando esta relación, podemos sugerir el siguiente método iterativo de dos pasos para resolver la ecuación (2.1).

Algoritmo 2.2.2 Para un x_0 dado, calcular la solución aproximada x_{n+1} por el esquema iterativo:

Predictor:

$$y_n = x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0.$$

Corrector:

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

Podemos reescribir el Algoritmo 2.2.2 en la siguiente forma: para un x_0 dado, encontrar la solución aproximada x_{n+1} por el esquema iterativo

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

que es conocido como la iteración de Householder, ver [19].

Nuevamente usando (2.22) y (2.19), podemos calcular

$$x_2 = N(x_0 + x_1) = -\frac{(x_0 + x_1 - \gamma)^2}{2f'(\gamma)} f''(\gamma). \quad (2.26)$$

De (2.22) y (2.23)-(2.26), concluimos que

$$\begin{aligned} x &= c + x_1 + x_2 \\ &= x_0 + N(x_0) + N(x_0 + x_1) \\ &= \gamma - \frac{f(\gamma)}{f'(\gamma)} - \frac{(x_0 - \gamma)^2}{2f'(\gamma)} f''(\gamma) - \frac{(x_0 + x_1 - \gamma)^2}{2f'(\gamma)} f''(\gamma). \end{aligned} \quad (2.27)$$

Usando (2.27), podemos sugerir y analizar el siguiente método iterativo de tres pasos para resolver la ecuación no lineal (2.1).

Algoritmo 2.2.3 (MKN1) Para un x_0 dado, calcular la solución aproximada x_{n+1} por el esquema iterativo

Predictor:

$$\begin{aligned} y_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0, \\ z_n &= -\frac{(y_n - x_n)^2}{2f'(x_n)} f''(x_n). \end{aligned}$$

Corrector:

$$x_{n+1} = g(x_n) = y_n + z_n - \frac{(y_n + z_n - x_n)^2}{2f'(x_n)} f''(x_n), \quad n = 0, 1, 2, \dots \quad (2.28)$$

Análisis de convergencia El método MKN1 requiere de una evaluación de f , f' y f'' en cada iteración. El método tiene un orden de convergencia igual a dos, orden igual al método de Newton con la diferencia de una evaluación de la segunda derivada para el método MKN1. La eficiencia computacional del método es: $r^{1/d} = 2^{1/3} \approx 1,259$.

Teorema 2.2 *Sea α una raíz simple de $f \in C^4(I)$, con I un intervalo abierto que contiene a α . Entonces existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq I$ el método iterativo definido por (2.28) converge a α y tiene un orden de convergencia igual a dos.*

Demostración: Denotemos por T al último término de la ecuación (2.28). Primero se tiene que demostrar que α es un punto fijo de g . Sea x_n una sucesión en I que converge a α . Esto implica, que

$$y(\alpha) = \alpha, \quad z(\alpha) = 0.$$

Así,

$$g(\alpha) = \alpha. \quad (2.29)$$

Calculando la primera y segunda derivada de y, z y T , simplificando los cálculos y manteniendo la hipótesis de que $x_n \rightarrow \alpha$, se obtiene lo siguiente

$$y'(\alpha) = 0, \quad y''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}, \quad (2.30)$$

$$z(\alpha) = z'(\alpha) = 0, \quad z''(\alpha) = -\frac{f''(\alpha)}{f'(\alpha)}, \quad (2.31)$$

$$T(\alpha) = T'(\alpha) = 0, \quad T''(\alpha) = -\frac{f''(\alpha)}{f'(\alpha)}. \quad (2.32)$$

De (2.30),(2.31) y (2.32) se obtiene,

$$g'(\alpha) = y'(\alpha) + z'(\alpha) + T'(\alpha) = 0 \quad (2.33)$$

y

$$g''(\alpha) = y''(\alpha) + z''(\alpha) + T''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)} - \frac{f''(\alpha)}{f'(\alpha)} - \frac{f''(\alpha)}{f'(\alpha)} = -\frac{f''(\alpha)}{f'(\alpha)}. \quad (2.34)$$

De (2.33),(2.29) y el Teorema 1.6, existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq I$ el método iterativo definido por (2.28) converge a α .

De (2.29),(2.33),(2.34) y el Teorema 1.7, el método iterativo (2.28) es de segundo orden.

□

2.3. Método iterativo MKN2

El método es introducido por Muhammad Noor y Khalida Noor en [6], es un método de dos pasos o método tipo predictor-corrector.

Construcción del método. Se asume que α es una raíz simple de (2.1) y γ es una condición inicial suficientemente cercana a α . Usando las series de Taylor, obtenemos

$$f(x) = f(\gamma) + (x - \gamma)f'(\gamma) + \frac{(x - \gamma)^2}{2}f''(\gamma) + R_n, \quad (2.35)$$

donde γ es la aproximación inicial para un cero de (2.1) y R_n denota los términos de orden superior.

De (2.1) y (2.35), obtenemos

$$f(\gamma) + (x - \gamma)f'(\gamma) + \frac{(x - \gamma)^2}{2}f''(\gamma) + R_n = 0. \quad (2.36)$$

Asumiendo que la aproximación inicial γ esta muy cercana a la solución exacta x , la ecuación (2.36) puede ser aproximada como

$$f(\gamma) + (x - \gamma)f'(\gamma) + \frac{(x - \gamma)^2}{2}f''(\gamma) \approx 0.$$

Usando la ecuación anterior, podemos sugerir el siguiente método iterativo para resolver la ecuación no lineal (2.1).

Algoritmo 2.3.1 (MKN2) Para un x_0 dado, calcular la solución aproximada x_{n+1} por el esquema iterativo

Predictor:

$$y(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}, \quad f'(x_n) \neq 0.$$

Corrector:

$$x_{n+1} = g(x_n) = x_n + \frac{-f'(x_n) \pm \sqrt{f'^2(x_n) - 2f(x_n)f''(y(x_n))}}{f''(y(x_n))}. \quad (2.37)$$

El signo del numerador del segundo término se elige de la siguiente manera:

- signo +, si $f'(\alpha) > 0$,
- signo -, si $f'(\alpha) < 0$.

Análisis de convergencia. El método requiere de una evaluación de f , f' y f'' en cada iteración. El método tiene un orden de convergencia de al menos tres. La eficiencia computacional del método es: $r^{1/d} = 3^{1/3} \approx 1,422$.

Teorema 2.3 *Sea α una raíz simple de $f \in C^4(I)$, con I un intervalo abierto que contiene a α . Entonces existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq I$ el método iterativo definido por (2.37) converge a α y tiene un orden de convergencia de al menos tres.*

Demostración: Sea x_n una sucesión en I que converge a α . Esto implica,

$$y(\alpha) = \alpha, \quad y'(\alpha) = 0, \quad y''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}.$$

Al realizar el calculo de g evaluada en α , se obtiene

$$g(\alpha) = \alpha. \quad (2.38)$$

La ecuación (2.38) implica que g tiene un punto fijo en I .

Al realizar el calculo de la primera y segunda derivada de g y realizando simplificaciones, se obtiene

$$g'(\alpha) = 0, \quad y \quad g''(\alpha) = 0 \quad (2.39)$$

De (2.38), (2.39) y el Teorema 1.6, existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq I$ el método iterativo definido por (2.37) converge a α . Ahora, de (2.38),(2.39) y el Teorema 1.7, el método iterativo (2.37) es de al menos orden tres. □

2.4. Método iterativo NSPP

El método es introducido por Kou Jishen, Li Yitian, Wang Xiuhua en [7]. Los autores plantean un método iterativo para encontrar una raíz simple de la ecuación no lineal $f(x) = 0$, donde $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ para un intervalo abierto D .

Construcción del método. En [12], es propuesto el método de Potra-Pták también conocido como método de dos pasos,

$$x_{n+1} = x_n - \frac{f(x_n) + f(x_n - f(x_n)/f'(x_n))}{f'(x_n)}. \quad (2.40)$$

Este método es de tercer orden y es obtenido a partir de modificaciones del método de Newton.

Por la composición del método de Newton y el método de Steffensen[13, 14], Sharma[15] presenta un método de tercer orden, llamado método de Newton-Steffensen, dado por

$$x_{n+1} = x_n - \frac{f(x_n)^2}{f'(x_n)(f(x_n) - f(x_n - f(x_n)/f'(x_n)))}. \quad (2.41)$$

Las fórmulas (2.40) y (2.41) no requieren la evaluación de la segunda derivada. A partir de las formulas (2.40) y (2.41) resulta un nuevo método de cuarto orden basado en la composición del método de Newton-Steffensen [13] y el método de Potra-Pták [12].

Ahora se considera la combinación lineal de las formulas (2.40) y (2.41) que produce una clase de métodos iterativos de orden superior

$$x_{n+1} = x_n - \theta \frac{f(x_n) + f(y_n)}{f'(x_n)} - (1 - \theta) \frac{f(x_n)^2}{f'(x_n)(f(x_n) - f(y_n))}, \quad (2.42)$$

donde $\theta \in \mathbb{R}$ y $y_n = x_n - f(x_n)/f'(x_n)$ es la iteración de Newton. Obviamente, cuando $\theta = 1$, en la formula (2.42) se obtiene (2.40) y si $\theta = 0$ en (2.42) se obtiene (2.41).

Análisis de convergencia. El método requiere de dos evaluaciones de f y una evaluación de f' en cada iteración. El método presenta un orden de convergencia de al menos tres para $\theta \in \mathbb{R}$ y orden de convergencia cuatro si $\theta = -1$. La eficiencia computacional del método es: $r^{1/d} = 4^{1/3} \approx 1,587$.

Necesitamos verificar que el esquema iterativo (2.42), denotado por NSPP, es convergente para cualquier x_0 perteneciente a un entorno de la raíz simple α de f .

Teorema 2.4 *Sea $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$, donde D es un intervalo abierto y $f \in C^3(D)$ y supongamos que f tiene una raíz simple $\alpha \in D$. Entonces, existe un $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta]$, el esquema iterativo (2.42) converge a α*

Demostración: Sea $g : D \subset \mathbb{R} \rightarrow \mathbb{R}$ definida de la siguiente manera,

$$g(x) = \begin{cases} x - \theta \frac{f(x) + f(y)}{f'(x)} - (1 - \theta) \frac{f(x)^2}{f'(x)(f(x) - f(y))}, & x \neq \alpha \\ \alpha & x = \alpha \end{cases}$$

donde $y = x - f(x)/f'(x)$ y $\theta \in \mathbb{R}$.

Considérese los siguientes límites:

$$\lim_{x \rightarrow \alpha} \left(x - \theta \frac{f(x) + f(y)}{f'(x)} \right) = \alpha, \quad (2.43)$$

$$\lim_{x \rightarrow \alpha} \frac{f(x)^2}{f'(x)(f(x) - f(y))} = \lim_{x \rightarrow \alpha} \frac{2f(x)f'(x)}{f''(x)(f(x) - f(y)) + f'(x)(f'(x) - f'(y)y')} = 0. \quad (2.44)$$

De (2.43) y (2.44), se obtiene $\lim_{x \rightarrow \alpha} g(x) = \alpha = g(\alpha)$, es decir, g es continua en α .

Considérese la derivada de g en α .

$$\begin{aligned}
g'(\alpha) &= \lim_{x \rightarrow \alpha} \frac{g(x) - g(\alpha)}{x - \alpha} \\
&= \lim_{x \rightarrow \alpha} \frac{x - \alpha - \theta \frac{f(x) + f(y)}{f'(x)} - (1 - \theta) \frac{f(x)^2}{f'(x)(f(x) - f(y))}}{x - \alpha} \\
&= \lim_{x \rightarrow \alpha} \left(1 - \theta \left[\frac{(f'(x) + f'(y)y')f'(x) - (f(x) + f(y))f''(x))}{f'(x)^2} \right] \right) \\
&\quad - \lim_{x \rightarrow \alpha} (1 - \theta) \left[\frac{2f(x)f'(x)^2(f(x) - f(y)) - f(x)^2[f''(x)(f(x) - f(y)) + \frac{f'(x)(f'(x) - f'(y)y')}{(f'(x)(f(x) - f(y))^2)}]}{(f'(x)(f(x) - f(y)))^2} \right] \\
&= (1 - \theta) - (1 - \theta) \\
&= 0.
\end{aligned} \tag{2.45}$$

De (2.45), $|g'(\alpha)| < 1$. Como g es continua en α y $|g'(\alpha)| < 1$ por el Teorema (1.6), existe $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq D$ el esquema iterativo (2.42) converge a α . \square

Teorema 2.5 *Sea $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$, donde D es un intervalo abierto. Supongamos que f tiene una raíz simple $\alpha \in D$. Si f es suficientemente diferenciable en un entorno de la raíz α , entonces el orden de convergencia del método definido por (2.42) es de al menos tres; Además, si $\theta = -1$, entonces el orden del método es cuatro.*

Demostración: Sea $e_n = x_n - \alpha$ y $d_n = y_n - \alpha$, donde $y_n = x_n - f(x_n)/f'(x_n)$. Usando expansión de Taylor y teniendo en cuenta que $f(\alpha) = 0$, se tiene

$$f(x_n) = f'(\alpha)[e_n + c_2e_n^2 + c_3e_n^3 + c_4e_n^4 + O(e_n^5)], \tag{2.46}$$

donde $c_k = (1/k!)f^{(k)}(\alpha)/f'(\alpha)$, $k = 2, 3, \dots$. Además, se tiene

$$f'(x_n) = f'(\alpha)[1 + 2c_2e_n + 3c_3e_n^2 + 4c_4e_n^3 + O(e_n^4)]. \tag{2.47}$$

Dividiendo (2.46) por (2.47) se obtiene

$$\frac{f(x_n)}{f'(x_n)} = e_n - c_2e_n^2 + 2(c_2^2 - c_3)e_n^3 + (7c_2c_3 - 4c_2^3 - 3c_4)e_n^4 + O(e_n^5), \tag{2.48}$$

$$d_n = y_n - \alpha = e_n - \frac{f(x_n)}{f'(x_n)} = c_2e_n^2 - 2(c_2^2 - c_3)e_n^3 - (7c_2c_3 - 4c_2^3 - 3c_4)e_n^4 + O(e_n^5). \tag{2.49}$$

Expandiendo de nuevo $f(y_n)$ sobre α , se tiene

$$f(y_n) = f'(\alpha)[d_n + c_2 d_n^2 + c_3 d_n^3 + c_4 d_n^4 + O(d_n^5)],$$

y de (2.49), se obtiene

$$f(y_n) = f'(\alpha)[c_2 e_n^2 - 2(c_2^2 - c_3)e_n^3 - (7c_2 c_3 - 5c_2^3 - 3c_4)e_n^4 + O(e_n^5)] \quad (2.50)$$

Así, de (2.46) y (2.50) se obtiene

$$f(x_n) + f(y_n) = f'(\alpha)[e_n + 2c_2 e_n^2 - (2c_2^2 - 3c_3)e_n^3 - (7c_2 c_3 - 5c_2^3 - 4c_4)e_n^4 + O(e_n^5)]. \quad (2.51)$$

Dividiendo (2.51) por (2.47) resulta

$$\frac{f(x_n) + f(y_n)}{f'(x_n)} = e_n - 2c_2^2 e_n^3 - (7c_2 c_3 - 9c_2^3)e_n^4 + O(e_n^5). \quad (2.52)$$

Por otro lado, de (2.46) y (2.50)

$$f(x_n) - f(y_n) = f'(\alpha)[e_n + (2c_2^2 - c_3)e_n^3 + (7c_2 c_3 - 5c_2^3 - 2c_4)e_n^4 + O(e_n^5)],$$

obteniéndose

$$\frac{f(x_n)}{f(x_n) - f(y_n)} = 1 + c_2 e_n + 2(c_3 - c_2^2)e_n^2 + 3(c_2^3 + c_4 - 2c_2 c_3)e_n^3 + O(e_n^4). \quad (2.53)$$

De (2.48) y (2.53) se tiene,

$$\frac{f(x_n)^2}{f'(x_n)(f(x_n) - f(y_n))} = e_n - c_2^2 e_n^3 - 3(c_2 c_3 - c_2^3)e_n^4 + O(e_n^5). \quad (2.54)$$

De (2.42)

$$e_{n+1} = e_n - \theta \frac{f(x_n) + f(y_n)}{f'(x_n)} - (1 - \theta) \frac{f(x_n)^2}{f'(x_n)(f(x_n) - f(y_n))}, \quad (2.55)$$

Sustituyendo (2.52) y (2.54) en (2.55) se obtiene

$$\begin{aligned} e_{n+1} &= e_n - \theta[e_n - 2c_2^2 e_n^3 - (7c_2 c_3 - 9c_2^3)e_n^4] \\ &\quad - (1 - \theta)[e_n - c_2^2 e_n^3 - 3(c_2 c_3 - c_2^3)e_n^4] + O(e_n^5) \\ &= (1 + \theta)c_2^2 e_n^3 + [(4\theta + 3)c_2 c_3 - 3(2\theta + 1)c_2^3]e_n^4 + O(e_n^5). \end{aligned} \quad (2.56)$$

Esto significa que el método definido por (2.42) es al menos cúbicamente convergente para toda $\theta \in \mathbb{R}$. Además, cuando $\theta = -1$, el orden de convergencia del método es cuatro y de (2.56) se tiene la ecuación de error

$$e_{n+1} = (3c_2^3 - c_2 c_3)e_n^4 + O(e_n^5).$$

□

2.5. Métodos iterativos CCHUN1 y CCHUN2

Los métodos CCHUN1 y CCHUN2 son introducidos por Changbum Chun en [8]. Antes de presentarse la construcción de cada uno de estos métodos, se plantea y desarrolla la idea general.

Se considera una función de iteración de la forma

$$\Phi(x) = x - h(x) \frac{f(x)}{f'(x)}, \quad (2.57)$$

donde h denota una función peso. Para desarrollar los métodos iterativos con tercer orden de convergencia, se construye la función peso h haciendo uso de las fórmulas iterativas de segundo orden.

Se puede mostrar que después de simples manipulaciones, el método iterativo definido por $x_{n+1} = \Phi(x_n)$ tiene la ecuación de error

$$e_{n+1} = (1 - h(\alpha))e_n + \frac{1}{2} \frac{h(\alpha)f''(\alpha) - 2h'(\alpha)f'(\alpha)}{f'(\alpha)} e_n^2 + O(e_n^3), \quad (2.58)$$

donde $e_n = x_n - \alpha$. Así, para que la función de iteración (2.57) sea de al menos orden tres, se debe satisfacer

$$h(\alpha) = 1, \quad (2.59)$$

$$h'(\alpha) = \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)}. \quad (2.60)$$

Antes de construir los dos métodos a estudiar, se va a construir un método para ilustrar la manera de construir la función peso h . Considérese la función de iteración de Newton, ϕ , definida por

$$\phi(x) = x - \frac{f(x)}{f'(x)}.$$

En vista de que $\phi''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$, es natural considerar h en la forma

$$h(x) = \frac{1}{2} \phi'(x) + C, \quad (2.61)$$

donde C es una constante arbitraria, así h satisface la condición (2.60). Como $\phi'(\alpha) = 0$, de (2.61) y (2.59) se obtiene que $C = 1$. Así la función h definida por

$$h(x) = \frac{1}{2} \phi'(x) + 1, \quad (2.62)$$

satisface las condiciones (2.59) y (2.60). Incorporando (2.62) en (2.57) se obtiene la función de iteración de tercer orden

$$\Phi(x) = x - \left[1 + \frac{1}{2}\phi'(x) \right] \frac{f(x)}{f'(x)}.$$

Por consiguiente, se obtiene el método iterativo con convergencia cúbica

$$x_{n+1} = x_n - \left[1 + \frac{1}{2}\phi'(x_n) \right] \frac{f(x_n)}{f'(x_n)}, \quad (2.63)$$

o

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

conocido como la iteración de Householder.

Método CCHUN1. El método se obtiene al considerar la función de iteración

$$\phi_1(x) = x - \frac{f(x)f'(x)}{f(x)^2 + f'(x)^2}.$$

En este caso, $\phi_1''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$. Así, de (2.63), se obtiene el nuevo método iterativo con convergencia cúbica

$$x_{n+1} = x_n - \left[1 + \frac{1}{2} \left[1 - \left(\frac{[f(x_n)^2 - f'(x_n)^2][f''(x_n)f(x_n) - f'(x_n)^2]}{[f(x_n)^2 + f'(x_n)^2]^2} \right) \right] \right] \quad (2.64)$$

En (2.64), se reemplaza $f''(x_n)$ por la diferencia $\frac{f'(x_n + f(x_n)) - f'(x_n)}{f(x_n)}$. De esta manera se define

$$\phi'(x) = 1 - \left[\frac{[f(x)^2 - f'(x)^2][f'(x + f(x)) - f'(x) - f'(x)^2]}{[f(x)^2 + f'(x)^2]^2} \right] \quad (2.65)$$

Así, el nuevo método iterativo CCHUN1, es dado por

$$x_{n+1} = x_n - L_f(x_n) \frac{f(x_n)}{f'(x_n)}, \quad (2.66)$$

donde

$$L_f(x_n) = 1 + \frac{1}{2}\phi'(x_n).$$

Análisis de convergencia del método CCHUN1. El método requiere de una evaluación de f y dos evaluaciones de f' en cada iteración. En el siguiente teorema se demuestra que el método iterativo definido en (2.66) es convergente y que es de al menos un tercer orden de convergencia. La eficiencia computacional del método es: $r^{1/d} = 3^{1/3} \approx 1,442$.

Teorema 2.6 *Sea α una raíz simple de $f \in C^3(I)$, con I un intervalo abierto que contiene a α . Supongamos que $L_f \in C^2(I)$. Si $L_f(x) = \frac{1}{2}\phi'(x) + 1$, entonces existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $x_{n+1} = x_n - L_f(x_n)\frac{f(x_n)}{f'(x_n)}$ converge a α y tiene por lo menos un tercer orden de convergencia.*

Demostración: Se necesita verificar que $L_f(\alpha) = 1$ y $L'_f(\alpha) = \frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)}$, es decir, $\phi'(\alpha) = 0$ y $\phi''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)}$ y después aplicar el Teorema 1.8 donde h es reemplazada por L_f . De (2.65),

$$\phi'(\alpha) = 1 - \left[\frac{(-f'(\alpha)^2)(-f'(\alpha)^2)}{f'(\alpha)^4} \right] = 1 - 1 = 0. \quad (2.67)$$

Derivando ϕ' y evaluándola en α , se obtiene

$$\begin{aligned} \phi''(\alpha) &= - \left[\frac{[2f'(\alpha)f''(\alpha) - f''(\alpha)[1 + f'(\alpha)] + f''(\alpha) + 2f'(\alpha)f''(\alpha)] - 4f'(\alpha)f''(\alpha)}{f'(\alpha)^2} \right] \\ &= - \left[-\frac{f''(\alpha)f'(\alpha)}{f'(\alpha)^2} \right] \\ &= \frac{f''(\alpha)}{f'(\alpha)}. \end{aligned} \quad (2.68)$$

De (2.67) y (2.68) se obtiene $L_f(\alpha) = 1$ y $L'_f(\alpha) = \frac{1}{2}\frac{f''(\alpha)}{f'(\alpha)}$, es decir $h = L_f$ satisface las hipótesis del Teorema 1.8. Por lo tanto, existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $x_{n+1} = x_n - L_f(x_n)\frac{f(x_n)}{f'(x_n)}$ converge a α y tiene por lo menos un tercer orden de convergencia. □

Método CCHUN2. Consideremos la función de iteración F definida por

$$F(x) = xe^{-\frac{f(x)}{xf'(x)}} \quad (x \neq 0). \quad (2.69)$$

Es fácil ver que $F(\alpha) = \alpha$, y de las ecuaciones (1.7) y (1.8), se tiene que $F'(\alpha) = 0$, $F''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)} + \frac{1}{\alpha}$, $\alpha \neq 0$. Se requiere que $h'(\alpha) = \frac{1}{2} \frac{f''(\alpha)}{f'(\alpha)}$, lo cual sugiere que se tome a h como

$$h(x) = \frac{1}{2}(F'(x) - \ln|x|) + C, \quad (2.70)$$

donde C es una constante arbitraria. Se puede determinar C haciendo $h(\alpha) = 1$ y teniendo en cuenta que $F'(\alpha) = 0$. Así, $C = 1 + \frac{1}{2} \ln|\alpha|$. De lo anterior, la función peso h queda definida por

$$h(x) = 1 + \frac{1}{2}F'(x) + \ln \sqrt{\left|\frac{\alpha}{x}\right|}, \quad (2.71)$$

la cual satisface las condiciones (2.59) y (2.60). Notemos que (2.71) contiene el cero α , que es generalmente desconocido, haciendo esta función inapropiada para la función de iteración (2.57). Para superar esta dificultad reemplazamos α con cualquier función de iteración $\phi(x)$ de orden dos, esto es, con $\phi(\alpha) = \alpha$, $\phi'(\alpha) = 0$, $\phi''(\alpha) \neq 0$. Esto produce una nueva función

$$h(x) = 1 + \frac{1}{2}F'(x) + \ln \sqrt{\left|\frac{\phi(x)}{x}\right|}. \quad (2.72)$$

Incorporando (2.72) en (2.57), obtenemos una familia de métodos iterativos de tercer orden:

$$x_{n+1} = x_n - W_f(x_n) \frac{f(x_n)}{f'(x_n)}, \quad (2.73)$$

donde

$$W_f(x_n) = 1 + \frac{1}{2}F'(x_n) + \ln \sqrt{\left|\frac{\phi(x_n)}{x_n}\right|}. \quad (2.74)$$

Sustituyendo $\phi(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$ en (2.73), obtenemos el método iterativo CCHUN2

$$x_{n+1} = x_n - W_f(x_n) \frac{f(x_n)}{f'(x_n)}, \quad (2.75)$$

donde

$$\begin{aligned} W_f(x_n) &= 1 + \frac{1}{2}F'(x_n) + \ln \sqrt{|z_{n+1}|}, \\ z_{n+1} &= 1 - \frac{f(x_n)}{x_n f'(x_n)}. \end{aligned}$$

Análisis de convergencia. El método CCHUN2 requiere de una evaluación de f , f' y f'' en cada iteración. Además el método es de al menos un tercer orden de convergencia. La eficiencia computacional del método es: $r^{1/d} = 3^{1/3} \approx 1,442$.

Teorema 2.7 *Sea α una raíz simple de $f \in C^3(I)$, con I un intervalo abierto que contiene a α . Supongamos que $W_f \in C^2(I)$. Si $W_f(x_n) = 1 + \frac{1}{2}F'(x_n) + \ln \sqrt{\left| \frac{\phi(x_n)}{x_n} \right|}$ y $\phi(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$, entonces existe $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $x_{n+1} = x_n - W_f(x_n) \frac{f(x_n)}{f'(x_n)}$ converge a α y tiene por lo menos un tercer orden de convergencia.*

Demostración: Se tiene que,

$$F(x) = xe^{-\frac{f(x)}{xf'(x)}}, \quad \phi(\alpha) = \alpha, \quad \phi'(\alpha) = 0, \quad \phi''(\alpha) = \frac{f''(\alpha)}{f'(\alpha)} \neq 0$$

y del Corolario 1.2, con $h = W_f$, se tiene que existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta]$ el método iterativo definido por $x_{n+1} = x_n - W_f(x_n) \frac{f(x_n)}{f'(x_n)}$ converge a α y tiene por lo menos un tercer orden de convergencia. \square

2.6. Métodos iterativos CCHUN3 y CCHUN4

Los métodos CCHUN3 y CCHUN4 son introducidos por Changbum Chun en [9]. La forma de presentar los dos métodos es parecida a la utilizada en la sección anterior.

Consideremos el método de Newton de dos pasos dados por

$$z_n = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (2.76)$$

$$x_{n+1} = z_n - \frac{f(z_n)}{f'(z_n)}. \quad (2.77)$$

La idea es encontrar una corrección para el segundo término del paso (2.77) que produzca un método de tercer orden. Para hacer esto, primeros consideramos adaptar la función f alrededor del punto $(x_n, f(x_n))$ con el polinomio de tercer grado

$$h(x) = ax^3 + bx^2 + cx + d. \quad (2.78)$$

Imponiendo sobre (2.78) las condiciones tangenciales en la n -ésima iteración x_n

$$h'(x_n) = f'(x_n), \quad (2.79)$$

tenemos

$$c = f'(x_n) - 3ax_n^2 - 2bx_n, \quad (2.80)$$

de este modo se obtiene la aproximación polinomial de la primera derivada

$$h'(x) = 3ax^2 + 2bx + f'(x_n) - 3ax_n^2 - 2bx_n. \quad (2.81)$$

Ahora, cuando z_n es definido por (2.76), aproximamos $f'(z_n)$ como

$$\begin{aligned} f'(z_n) &\approx h'(z_n) = 3az_n^2 + 2bz_n + f'(x_n) - 3ax_n^2 - 2bx_n \\ &= 3a \left(x_n - \frac{f(x_n)}{f'(x_n)} \right)^2 + 2b \left(x_n - \frac{f(x_n)}{f'(x_n)} \right) + f'(x_n) - 3ax_n^2 - 2bx_n \\ &= -6ax_n \frac{f(x_n)}{f'(x_n)} + 3a \frac{f(x_n)^2}{f'(x_n)^2} - 2b \frac{f(x_n)}{f'(x_n)} + f'(x_n) \\ &= \frac{-6ax_n f(x_n) f'(x_n) + 3a f(x_n)^2 - 2b f(x_n) f'(x_n) + f'(x_n)^3}{f'(x_n)^2} \\ &= \frac{f'(x_n)^3 - (6ax_n + 2b) f(x_n) f'(x_n) + 3a f(x_n)^2}{f'(x_n)^2}. \end{aligned} \quad (2.82)$$

El autor del artículo [9], al aproximar $f'(z_n)$ por $h'(z_n)$, obtiene que

$$f'(z_n) \approx h'(z_n) = \frac{f'(x_n)^2 - (6ax_n + 2b) f(x_n) + 3a f(x_n)^2}{f'(x_n)},$$

y al reemplazarla en la ecuación (2.77) obtiene una familia de métodos de dos parámetros

$$z_n = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (2.83)$$

$$x_{n+1} = g(x_n) = z_n - \frac{f(z_n) f'(x_n)}{f'(x_n)^2 + (\lambda - 2\mu x_n) f(x_n) + \mu f(x_n)^2} \quad (2.84)$$

donde

$$\lambda = -2b \in \mathbb{R},$$

$$\mu = 3a \in \mathbb{R}.$$

Si se usa la aproximación de $f'(z_n)$ obtenida en (2.82) y se reemplaza en (2.77) obtenemos la nueva familia de métodos de dos parámetros

$$z_n = x_n - \frac{f(x_n)}{f'(x_n)}, \quad (2.85)$$

$$x_{n+1} = g(x_n) = z_n - \frac{f(z_n) f'(x_n)^2}{f'(x_n)^3 + (\lambda - 2\mu x_n) f(x_n) f'(x_n) + \mu f(x_n)^2}, \quad (2.86)$$

donde

$$\lambda = -2b \in \mathbb{R},$$

$$\mu = 3a \in \mathbb{R}.$$

Análisis de convergencia. Cada uno de los métodos iterativos de la familia (2.84) requieren de dos evaluaciones de f y una evaluación de f' en cada iteración. Además el orden de convergencia es tres. La eficiencia computacional de cada uno de estos métodos es: $r^{1/d} = 3^{1/3} \approx 1,442$.

Teorema 2.8 *Sea α una raíz simple de $f \in C^2(I)$, con I un intervalo abierto que contiene a α . Entonces existe $\delta > 0$ tal que $\forall x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq I$ la familia de métodos iterativos definidos por (2.84) converge a α .*

Demostración: Como $f \in C^2(I)$, entonces $g \in C^1(I)$. Sea x_n una sucesión en I que converge a α . Esto implica,

$$z_n \rightarrow \alpha \quad \text{y} \quad z'_n \rightarrow 0.$$

Como α es una raíz simple de f ,

$$f(z_n) \rightarrow 0, \quad f'(x_n) \rightarrow f'(\alpha) \neq 0.$$

Así,

$$g(x_n) \rightarrow \alpha,$$

es decir, α es un punto fijo de g . Derivando g y calculando $\lim_{x_n \rightarrow \alpha} g'(x_n)$ obtenemos

$$g'(x_n) \rightarrow 0.$$

Por lo tanto, el Teorema 1.6, garantiza que existe $\delta > 0$ tal que para todo $x_0 \in [\alpha - \delta, \alpha + \delta] \subseteq I$ la familia de métodos iterativos definidos por (2.84) converge a α . \square

Teorema 2.9 *Sea $\alpha \in I$ una raíz simple de una función $f : I \rightarrow \mathbb{R}$, para un intervalo I abierto y f es suficientemente diferenciable. Si x_0 esta suficientemente cercano a α , entonces el método definido por (2.84) tiene tercer orden de convergencia.*

Demostración:

$$f(x_n) = f'(\alpha)[e_n + c_2 e_n^2 + O(e_n^3)], \quad (2.87)$$

$$f'(x_n) = f'(\alpha)[1 + 2c_2 e_n + O(e_n^2)], \quad (2.88)$$

donde $e_n = x_n - \alpha$, $c_k = \frac{f^{(k)}(\alpha)}{k!f'(\alpha)}$ y $\lambda, \mu \in \mathbb{R}$, $k = 1, 2, \dots$. Por simples cálculos, obtenemos

$$\frac{f(x_n)}{f'(x_n)} = e_n - c_2 e_n^2 + 2(c_2^2 - c_3)e_n^3 + O(e_n^4), \quad (2.89)$$

de manera que

$$z_n = x_n - \frac{f(x_n)}{f'(x_n)} = \alpha + c_2 e_n^2 - 2(c_2^2 - c_3)e_n^3 + O(e_n^4), \quad (2.90)$$

de donde

$$f(z_n) = f'(\alpha)[c_2 e_n^2 - 2(c_2^2 - c_3)e_n^3 + O(e_n^4)]. \quad (2.91)$$

De (2.87),(2.88) y (2.91), tenemos

$$f(z_n)f'(x_n) = f'(\alpha)^2[c_2 e_n^2 + 2c_3 e_n^3 + O(e_n^4)], \quad (2.92)$$

y

$$f'(x_n)^2 + (\lambda - 2\mu x_n)f(x_n) + \mu f(x_n)^2 = f'(\alpha)[f'(\alpha) + (4c_2 f'(\alpha) + \lambda - 2\mu\alpha)e_n + O(e_n^2)], \quad (2.93)$$

de donde

$$\frac{f(z_n)f'(x_n)}{f'(x_n)^2 + (\lambda - 2\mu x_n)f(x_n) + \mu f(x_n)^2} = c_2 e_n^2 - \left(4c_2^2 + \frac{\lambda - 2\mu\alpha}{f'(\alpha)}c_2 - 2c_3\right) e_n^3 + O(e_n^4). \quad (2.94)$$

De (2.90) y (2.94)

$$\begin{aligned} x_{n+1} &= z_n - \frac{f(z_n)f'(x_n)}{f'(x_n)^2 + (\lambda - 2\mu x_n)f(x_n) + \mu f(x_n)^2} \\ &= \alpha + c_2 \left(2c_2 + \frac{\lambda - 2\mu\alpha}{f'(\alpha)}\right) e_n^3 + O(e_n^4). \end{aligned} \quad (2.95)$$

Puesto que $e_{n+1} = x_{n+1} - \alpha$, esto muestra que el método iterativo definido por (2.84) tiene tercer orden de convergencia. Esto completa la demostración. \square

Método CCHUN3. El método se obtiene al considerar $\lambda = 1$ y $\mu = 0$ en (2.84):

$$x_{n+1} = z_n - \frac{f(z_n)f'(x_n)}{f'(x_n)^2 + f(x_n)},$$

con

$$z_n = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Método CCHUN4. El método se obtiene al considerar $\lambda = 0$ y $\mu = 1$ en (2.84):

$$x_{n+1} = z_n - \frac{f(z_n)f'(x_n)}{f'(x_n)^2 - x_n f(x_n) + f(x_n)^2},$$

con

$$z_n = x_n - \frac{f(x_n)}{f'(x_n)}.$$

2.7. Método iterativo SG

El método es introducido por J. R. Sharma y R. K. Guha en [10]. Los autores desarrollan un esquema que mejora el orden de convergencia del método de Ostrowski. Así, ellos empiezan con el siguiente esquema de iteración multipunto,

$$\begin{aligned} z_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - \frac{f(z_n)}{f'(x_n)} \frac{f(x_n)}{f(x_n) - 2f(z_n)}, \\ \tilde{x}_{n+1} &= x_{n+1} - \frac{f(x_{n+1})}{f'(x_n)} \frac{f(x_n) + af(z_n)}{f(x_n) + bf(z_n)}. \end{aligned} \quad (2.96)$$

donde a y b son parámetros que pueden ser determinados del siguiente teorema de convergencia.

Análisis de convergencia. El método SG requiere de tres evaluaciones de f y una evaluación de f' en cada iteración. El método presenta un sexto orden de convergencia. Por lo cual la eficiencia computacional del método es: $r^{1/d} = 6^{1/4} \approx 1,565$

Teorema 2.10 *Sea $f : I \rightarrow \mathbb{R}$ una función de valor real definida sobre I , donde I es entorno de una raíz simple α de f . Supongamos que $f \in C^4(I)$. Entonces el esquema de iteración (2.96) define una familia de un parámetro, de sexto orden de convergencia si $b = a - 2$.*

Demostración: Usando expansión de Taylor de $f(x_n)$ sobre α y teniendo en cuenta que $f'(\alpha) \neq 0$, se tiene

$$f(x_n) = f'(\alpha)[e_n + c_2 e_n^2 + c_3 e_n^3 + c_4 e_n^4 + O(e_n^5)], \quad (2.97)$$

donde $e_n = x_n - \alpha$ y $c_k = (1/k!)f^{(k)}(\alpha)/f'(\alpha)$, $k = 2, 3, \dots$. Además,

$$f'(x_n) = f'(\alpha)[1 + 2c_2 e_n + 3c_3 e_n^2 + 4c_4 e_n^3 + O(e_n^4)], \quad (2.98)$$

y

$$\frac{f(x_n)}{f'(x_n)} = e_n - c_2 e_n^2 + 2(c_2^2 - c_3)e_n^3 + (7c_2 c_3 - 4c_2^3 - 3c_4)e_n^4 + O(e_n^5) \quad (2.99)$$

Sustituyendo (2.99) en $z_n = x_n - f(x_n)/f'(x_n)$ resulta

$$z_n - \alpha = c_2 e_n^2 - 2(c_2^2 - c_3)e_n^3 - (7c_2 c_3 - 4c_2^3 - 3c_4)e_n^4 + O(e_n^5), \quad (2.100)$$

Expandiendo $f(z_n)$ sobre α y usando (2.100), se tiene

$$f(z_n) = f'(\alpha)[c_2 e_n^2 - 2(c_2^2 - c_3)e_n^3 - (7c_2 c_3 - 5c_2^3 - 3c_4)e_n^4 + O(e_n^5)], \quad (2.101)$$

De (2.97) y (2.101), se obtiene

$$f(x_n) - 2f(z_n) = f'(\alpha)[e_n - c_2 e_n^2 + (4c_2^2 - 3c_3)e_n^3 + (14c_2c_3 - 10c_2^3 - 5c_4)e_n^4 + O(e_n^5)]. \quad (2.102)$$

Dividiendo (2.97) por (2.102) y simplificando, se obtiene

$$\frac{f(x_n)}{f(x_n) - 2f(z_n)} = 1 + 2c_2 e_n + 2(-c_2^2 + 2c_3)e_n^2 + 2(-2c_2c_3 + 3c_4)e_n^3 + O(e_n^4). \quad (2.103)$$

También

$$\frac{f(z_n)}{f'(x_n)} = c_2 e_n^2 + 2(-2c_2^2 + c_3)e_n^3 + (-14c_2c_3 + 13c_2^3 + 3c_4)e_n^4 + O(e_n^5). \quad (2.104)$$

Multiplicando (2.103) por (2.104) resulta

$$\frac{f(z_n)}{f'(x_n)} \frac{f(x_n)}{f(x_n) - 2f(z_n)} = c_2 e_n^2 + 2(-c_2^2 + c_3)e_n^3 + 3(-2c_2c_3 + c_2^3 + c_4)e_n^4 + O(e_n^5). \quad (2.105)$$

Así,

$$x_{n+1} = z_n - \frac{f(z_n)}{f'(x_n)} \frac{f(x_n)}{f(x_n) - 2f(z_n)} = \alpha + c_2(c_2^2 - c_3)e_n^4 + O(e_n^5), \quad (2.106)$$

que es

$$x_{n+1} - \alpha = c_2(c_2^2 - c_3)e_n^4 + O(e_n^5). \quad (2.107)$$

De (2.97) y (2.101), se obtiene

$$f(x_n) + af(z_n) = f'(\alpha)[e_n + (1+a)c_2 e_n^2 + \{-2ac_2^2 + (1+2a)c_3\}e_n^3 + O(e_n^4)], \quad (2.108)$$

y

$$f(x_n) + bf(z_n) = f'(\alpha)[e_n + (1+b)c_2 e_n^2 + \{-2bc_2^2 + (1+2b)c_3\}e_n^3 + O(e_n^4)]. \quad (2.109)$$

Al dividir (2.108) por (2.109) y simplificado, se obtiene

$$\frac{f(x_n) + af(z_n)}{f(x_n) + bf(z_n)} = 1 + (a-b)c_2 e_n + (a-b)\{- (b+3)c_2^2 + 2c_3\}e_n^2 + O(e_n^3). \quad (2.110)$$

Ahora la expansión de Taylor de $f(x_{n+1})$ sobre α es dada por

$$f(x_{n+1}) = f'(\alpha)[(x_{n+1} - \alpha) + O\{(x_{n+1} - \alpha)^2\}]. \quad (2.111)$$

De (2.98) y (2.111) se encuentra

$$\frac{f(x_{n+1})}{f'(x_n)} = [1 - 2c_2 e_n + (4c_2^2 - 3c_3)e_n^2](x_{n+1} - \alpha) + O(e_n^7). \quad (2.112)$$

De la multiplicación de (2.110) y (2.112) resulta

$$\frac{f(x_{n+1})f(x_n) + af(z_n)}{f'(x_n)f(x_n) + bf(z_n)} = [1 + (a - b - 2)c_2e_n + \{(4 - (a - b)(b + 5))c_2^2 - (3 - 2(a - b))c_3\}e_n^2](x_{n+1} - \alpha) + O(e_n^7). \quad (2.113)$$

Sustituyendo (2.107) y (2.113) en el tercer paso de (2.96), se obtiene la ecuación de error

$$\begin{aligned} \tilde{e}_{n+1} &= [(2 - a + b)c_2e_n - \{(4 - (a - b)(b + 5))c_2^2 - (3 - 2(a - b))c_3\}e_n^2](x_{n+1} - \alpha) + O(e_n^7) \\ &= c_2(c_2^2 - c_3)[(2 - a + b)c_2 - \{(4 - (a - b)(b + 5))c_2^2 - (3 - 2(a - b))c_3\}e_n]e_n^5 + O(e_n^7). \end{aligned} \quad (2.114)$$

Para que el esquema iterativo (2.96) sea de sexto orden, el coeficiente de e_n^5 debe desaparecer. De (2.114) se encuentran las siguientes condiciones sobre los parámetros a y b , con

$$b = a - 2.$$

Con este valor de b , la ecuación de error (2.114), queda

$$\tilde{e}_{n+1} = c_2(c_2^2 - c_3)[2(a + 1)c_2^2 - c_3]e_n^6 + O(e_n^7). \quad (2.115)$$

Así, la ecuación (2.115) establece el máximo orden de convergencia igual a seis para el esquema iterativo (2.96). Esto completa la prueba. \square

Método iterativo SG. El método se obtiene al considerar $a = 1$ en (2.96):

$$\begin{aligned} z_n &= x_n - \frac{f(x_n)}{f'(x_n)}, \\ x_{n+1} &= z_n - \frac{f(z_n)}{f'(x_n)} \frac{f(x_n)}{f(x_n) - 2f(z_n)}, \\ \tilde{x}_{n+1} &= x_{n+1} - \frac{f(x_{n+1})f(x_n) + f(z_n)}{f'(x_n)f(x_n) - f(z_n)}. \end{aligned}$$

2.8. Método iterativo BM

El Método de BM es introducido por Giovanni Calderón en [11]. El método es derivado por el autor al combinar el método de Bisección con Müller.

Sean $a, b \in \mathbb{R}$ y $f \in C[a, b]$ tal que $f(a)f(b) < 0$, por el Teorema del Valor Intermedio se tiene que existe $\alpha \in (a, b)$ tal que $f(\alpha) = 0$. El método consiste en construir el polinomio de segundo grado (polinomio de Müller) que pasa por los puntos $(a, f(a))$,

$((a+b)/2, f((a+b)/2))$ y $(b, f(b))$. La raíz del polinomio de Müller define una aproximación al cero α . Luego, para la siguiente iteración se define el intervalo $[a, b]$ como el menor segmento que contiene a α : es decir $(a, (a+b)/2)$ o $((a+b)/2, b)$. El algoritmo del método es el siguiente.

Algoritmo de BM

Entrada: $f \in [a, b]$ tal que $f(a)f(b) < 0$; ε tolerancia; máx número máximo de iteraciones.

Salida: α solución aproximada a $f(x) = 0$.

1. $n = 0$.
2. Mientras $n \leq$ máx hacer los pasos 3 hasta 7.
3. Calcular, $c = \frac{a+b}{2}$; $a_0 = \frac{(c-b)[f(a) - f(b)] - (a-b)[f(c) - f(b)]}{(a-b)(c-b)(a-c)}$,
 $a_1 = \frac{f(b) - f(c)}{b-c} + (b-c)a_0$, $a_2 = f(b)$ raíz = $b - \frac{2a_2}{a_1 - \sqrt{a_1^2 - 4a_0a_2}}$,
 a_0, a_1 y a_2 son los coeficientes del polinomio de segundo grado
 $P(x) = a_0(x - x_2)^2 + a_1(x - x_2) + a_2$ (Müller)

4. Si $a < \text{raíz} < b$, entonces $x_{n+1} = \text{raíz}$.

En caso contrario: $x_{n+1} = b - \frac{2a_2}{a_1 + \sqrt{a_1^2 - 4a_0a_2}}$.

5. Si $|x_{n+1} - x_n| < \varepsilon$ o $|f(x_{n+1})| < \varepsilon$, entonces $\alpha = \text{raíz}$, detener.

6. Si $f(a)f(x_{n+1}) < 0$, entonces $b = \text{raíz}$.
 Si $f(a)f(c) > 0$, entonces $a = c$.

En caso contrario: $a = \text{raíz}$.
 Si $f(b)f(c) > 0$, entonces $b = c$

7. $n = n + 1$, ir al paso 2.

Análisis de Convergencia El método en la primera iteración realiza cuatro evaluaciones de función y de la segunda iteración realiza solo dos evaluaciones de función. En el siguiente teorema se establece que el método tiene convergencia cúbica cuya demostración es realizada en [20]. La eficiencia computacional del método es $r^{1/d} = 3^{1/2} \approx 1,732$.

Teorema 2.11 *Sea $f \in C[a, b]$ con a y b tales que $f(a)f(b) < 0$. Si α es un cero simple de f en $[a, b]$, entonces la sucesión x_n producida por el método de BM converge a α con a lo sumo un tercer orden de convergencia.*

Demostración: Ver Teorema 2.8 en [20].

□

Estrategia de clasificación de los métodos y experimentación numérica

Desde un punto de vista práctico, en la comparación de dos o más métodos, además de sus propiedades teóricas (orden de convergencia, constante asintótica del error, selección de las aproximaciones iniciales), se debería tener en cuenta su costo computacional (tiempo de CPU usado). Pues, en la mayoría de los casos, las condiciones exigidas para obtener la convergencia de un método son restrictivas y en la práctica suele ser muy difícil que se cumplan en todo el proceso iterativo. Por tal motivo, existe la necesidad de hacer una clasificación de los nuevos métodos iterativos a partir del tiempo de CPU usado.

Con lo ya mencionado, el objetivo de este capítulo consiste en definir e implementar una estrategia para analizar la efectividad de los métodos iterativos tomando en cuenta el tiempo de CPU usado por los mismos.

Como hay una gran cantidad de referencias recientes en el tema (nuevos métodos iterativos), se han seleccionado 8 artículos para ser analizados (ver referencias [4]-[11]) a lo largo del trabajo. Además de los métodos que surgen de las referencias [4]-[11], se anexaran para ser analizados según la estrategia que se va a definir, los métodos NU, NRF y RFNM cuyos análisis de convergencia fueron realizados en [20].

A partir de la información contenida en el Cuadro 3.1, se construye el Cuadro 3.2, en el cual se da la clasificación de los métodos a partir de la eficiencia computacional observándose un intercambio de posiciones de los métodos BM y SG (ver Cuadros 3.1 y 3.2). Sin embargo, la eficiencia computacional no toma en cuenta que tipo de función se

Métodos	Orden	Nf	Nf'	Nf''
SG	6	3	1	0
NSPP	4	2	1	0
BM	3	2	0	0
RFNM	3	2	1	0
NRF	3	2	1	0
CCHUN3	3	2	1	0
CCHUN4	3	2	1	0
CCHUN1	3	1	2	0
CCHUN2	3	1	1	1
MKN2	3	1	1	1
JRS	3	1	1	1
NEWTON	2	1	1	0
NU	2	2	1	0
MKN1	2	1	1	1

Cuadro 3.1: Clasificación por orden de convergencia. Los términos Nf, Nf' y Nf'', representan el número de evaluaciones de f , f' y f'' respectivamente (por iteración).

Métodos	Eficiencia Computacional ($r^{1/d}$)
BM	1.7320
NSPP	1.5874
SG	1.5650
RFNM	1.4422
NRF	1.4422
CCHUN3	1.4422
CCHUN4	1.4422
CCHUN1	1.4422
CCHUN2	1.4422
MKN2	1.4422
JRS	1.4422
NEWTON	1.4142
NU	1.2599
MKN1	1.2599

Cuadro 3.2: Métodos ordenados por eficiencia computacional

esta evaluando; esto es, si es f o alguna de sus derivadas. Por ejemplo, en el Cuadro 3.2, se le da la misma eficiencia computacional al método CCHUN3 y al método MKN2, el primero con dos evaluaciones de f y una de f' , mientras el segundo tiene una evaluación de f , una evaluación de f' y una evaluación de f'' . Se puede observar en el Cuadro 3.2 que los métodos RFNM, NRF, CCHUN3, CCHUN4, CCHUN1, CCHUN2, MKN2 y JRS tienen igual eficiencia computacional, ¿será que al estimar los tiempos de CPU de cada uno de estos métodos, los tiempos serán iguales? Además, en general, los tiempos de cálculo de f' y f'' son más grandes que el tiempo de cálculo de f . Por lo cual, surge la pregunta: ¿Influirá la primera derivada o la segunda derivada en los tiempos

de CPU de los métodos que utilizan una o ambas derivadas? Las preguntas formuladas anteriormente se espera que sean respondidas después que se muestre y analice la data que se obtiene al aplicar la estrategia de clasificación a definir.

3.1. Estrategia de clasificación.

Se supone que se tienen m métodos iterativos a clasificar y n funciones. Cada función tiene un intervalo $[a, b]$ en el cual tienen una raíz simple. En cada uno de los intervalos $[a, b]$ se colocan k puntos, es decir, $a = x_1 < x_2 < \dots < x_{k-1} < x_k = b$ y se procede a aplicar cada método de la siguiente manera: se toma el primer punto, $a = x_1$, de cada uno de los intervalos $[a, b]$ de las respectivas funciones f_1, f_2, \dots, f_n . Se aplica el método M_1 a las funciones f_1, f_2, \dots, f_n , obteniéndose los respectivos tiempos $t_{11}^1, t_{12}^1, \dots, t_{1n}^1$, cuyo promedio es T_{11} (ver Cuadro 3.3 para el método M_1). A continuación, se aplica el método

Tiempo de los métodos para la base de funciones					
Métodos	Ejemplos				
	f_1	f_2	\dots	f_n	
M_1	t_{11}^1	t_{12}^1	\dots	t_{1n}^1	$= \frac{\sum_{i=1}^n t_{1i}^1}{n} = T_{11}$
M_2	t_{21}^1	t_{22}^1	\dots	t_{2n}^1	$= \frac{\sum_{i=1}^n t_{2i}^1}{n} = T_{12}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
M_m	t_{m1}^1	t_{m2}^1	\dots	t_{mn}^1	$= \frac{\sum_{i=1}^n t_{mi}^1}{n} = T_{1m}$

Cuadro 3.3: Esquema para obtener el tiempo de cada método para el valor de $c = 1$. El total de métodos es $m = 14$ y el total de ejemplos que pertenecen a la base de funciones es $n = 47$. Por ejemplo, el valor t_{12}^1 es el tiempo que tarda el método 1 en hallar la raíz, α , de la función 2.

M_2 a las funciones f_1, f_2, \dots, f_n , obteniéndose las respectivos tiempos $t_{21}^1, t_{22}^1, \dots, t_{2n}^1$, cuyo promedio es T_{12} (ver Cuadro 3.3 para el método M_2). Este procedimiento se repite hasta aplicar el método M_m a las funciones f_1, f_2, \dots, f_n , obteniéndose los respectivos tiempos $t_{m1}^1, t_{m2}^1, \dots, t_{mn}^1$, cuyo promedio es T_{1m} (ver Cuadro 3.3 para el método M_m).

Seguidamente se toma el segundo punto, x_2 , de cada uno de los intervalos $[a, b]$ de las respectivas funciones f_1, f_2, \dots, f_n . Para este punto x_2 , se aplica el procedimiento

usado para el punto x_1 , obteniéndose los tiempos promedios, $T_{21}, T_{22}, \dots, T_{2m}$, de los respectivos métodos M_1, M_2, \dots, M_m (ver los tiempos para el valor de $c = 2$ en el Cuadro 3.4).

Tiempo de los métodos para k puntos en $[a, b]$				
c	Métodos			
	M_1	M_2	\dots	M_m
1	T_{11}	T_{12}	\dots	T_{1m}
2	T_{21}	T_{22}	\dots	T_{2m}
\vdots	\vdots	\vdots	\ddots	\vdots
k	T_{k1}	T_{k2}	\dots	T_{km}

Cuadro 3.4: Esquema para obtener el tiempo de cada método para un total de $k = 1500$ puntos sobre cada intervalo. El total de métodos es $m = 14$.

```

Algoritmo A

 $f_i, [a_i, b_i] \quad x_c^i = a_i + ((b_i - a_i)/(k - 1))(c - 1);$ 
 $i = 1 : n \quad c = 1 : k.$ 
 $M_j \quad j = 1 : m.$ 
Para  $c = 1 : k$  hacer
  Para  $i = 1 : n$  hacer
    Para  $j = 1 : m$  hacer
      Si  $f_i(x_c^i)f_i(b_i) < 0$  entonces
         $a_i = x_c^i;$ 
      sino
         $b_i = x_c^i;$ 
      Fin Si.
       $t_{ji}^c = \text{TiempoMetodo}(M_j, f_i, a_i, b_i, x_c^i);$ 
    Fin Para
  Fin Para
  Para  $p = 1 : m$  hacer
     $T_{cp} = \text{Promedio}(t_{p(1:n)}^c);$ 
  Fin Para
Fin Para

```

De esta manera se procede para los restantes puntos hasta llegar al punto $x_k = b$ de cada uno de los intervalos $[a, b]$ de las respectivas funciones f_1, f_2, \dots, f_n . Para este

punto $x_k = b$, se obtienen los tiempos promedios, $T_{k1}, T_{k2}, \dots, T_{km}$, de los respectivos métodos M_1, M_2, \dots, M_m (ver los tiempos para el valor de $c = k$ en el Cuadro 3.4).

El procedimiento anterior se resume en el algoritmo A. Ahora a partir de los tiempos obtenidos para cada método en los respectivos k puntos sobre $[a, b]$, se clasifican los métodos M_1, M_2, \dots, M_m , en función de la media de los tiempos obtenidos por dichos métodos.

3.1.1. Construcción del intervalo $[a, b]$

Para que una función de iteración, g , tenga un punto fijo en un intervalo $[a, b]$, según el Lema 1.1, se necesita que dicha función sea continua en $[a, b]$ y además que g mande $[a, b]$ en $[a, b]$. Como se tienen más de dos métodos a clasificar, se necesita que todos esos métodos sean convergente sobre el mismo intervalo $[a, b]$.

Considerando lo dicho anteriormente, se aplica el procedimiento siguiente con el cual se construyen cada uno de los intervalos de las respectivas funciones $f_1, f_2 \dots, f_n$

- Para cada una de las funciones se tomo un intervalo $[a, b]$ tal que $f(a)f(b) < 0$. Esta parte es suficiente para los métodos que necesitan un intervalo como por ejemplo, BM, NRF, RFNM.
- Se busca un intervalo $[a_1, b_1] \subseteq [a, b]$ tal que $f'(x) \neq 0$, para toda $x \in [a_1, b_1]$. Esta parte se debe a que la mayoría de los métodos que se analizaron en el Capítulo 2, usan a NEWTON como predictor.
- Se grafican las funciones de iteración de cada uno de los métodos iterativos sobre el intervalo $[a_1, b_1]$. Se hace un análisis gráfico para ver si los métodos iterativos convergen a la raíz, α de f . En caso de que exista un punto en el cual uno de los métodos iterativos no pueda converger, entonces se busca un intervalo $[a_2, b_2] \subseteq [a_1, b_1]$ de tal manera que el método sea convergente para cada $x_0 \in [a_2, b_2]$. Hasta este punto se logra definir la base de funciones como la expuesta en el Apéndice A de este trabajo.

Para ilustrar los puntos anteriores, se va a construir el intervalo $[a, b]$ para la función 22 del Apéndice A.

Se puede observar en la Figura 3.1, que la función $f = e^{\sin(x)} - x - 1$ tiene una raíz en el intervalo $[1, 3]$. Además se puede observar que existe $x \in [1, 1.2]$ tal que $f'(x) = 0$. Por tal motivo, nos quedamos con el intervalo $[1.2, 3]$.

En la Figura 3.2 se muestra el gráfico de la función de iteración del método de NEWTON, en la cual hay puntos entre $[1.2, 1.5]$ tal que las imágenes de dichos puntos no caen en $[1.2, 3]$. Por lo tanto, nos quedamos con el intervalo $[1.3, 3]$.

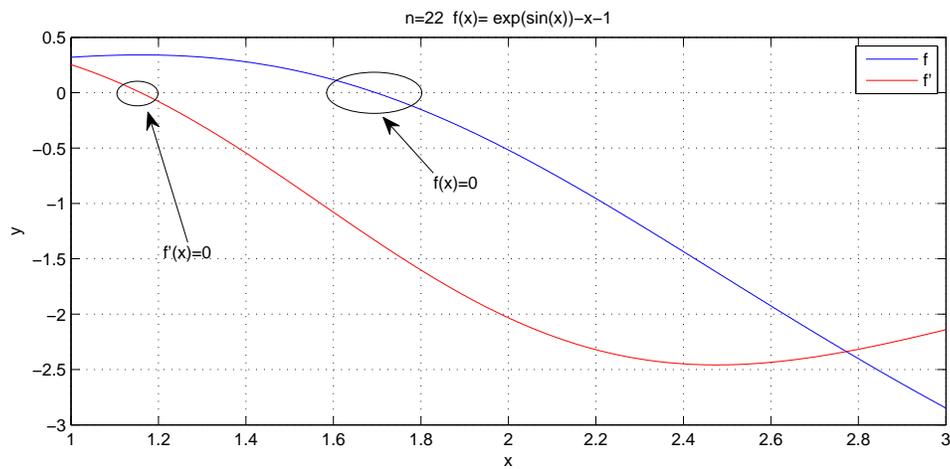


Figura 3.1: f y f' sobre $[a,b]=[1,3]$. En esta figura y en las figuras 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8 y 3.9 se colocaran en cada una de las leyendas de las figuras antes mencionadas, los nombres de los métodos en minúscula.

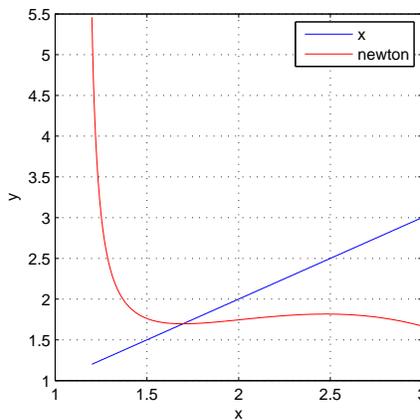


Figura 3.2: Función de iteración del método de NEWTON

En la Figura 3.3 se muestra el gráfico de la función de iteración del método NSPP en la cual hay puntos entre $[1.3, 1.4]$ tal que la imagen de dichos puntos no caen en $[1.3, 3]$. De esta manera nos quedamos con el intervalo $[1.4, 3]$.

En la Figura 3.4 se muestra el gráfico de la función de iteración del método CCHUN1 en la cual hay puntos entre $[2.4, 3]$ tal que la imagen de dichos puntos no caen en $[1.4, 3]$. Por esta razón nos quedamos con el intervalo $[1.4, 2.4]$.

En la Figura 3.5 se muestra el gráfico de la función de iteración del método CCHUN2

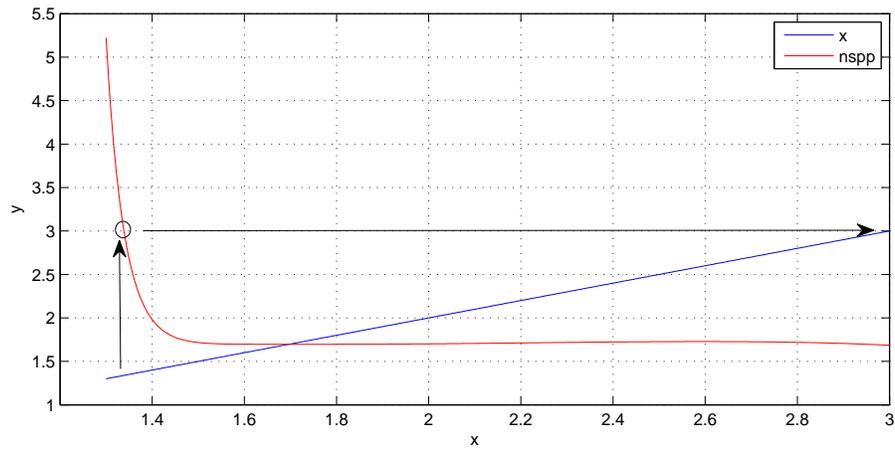


Figura 3.3: Función de iteración del método NSPP

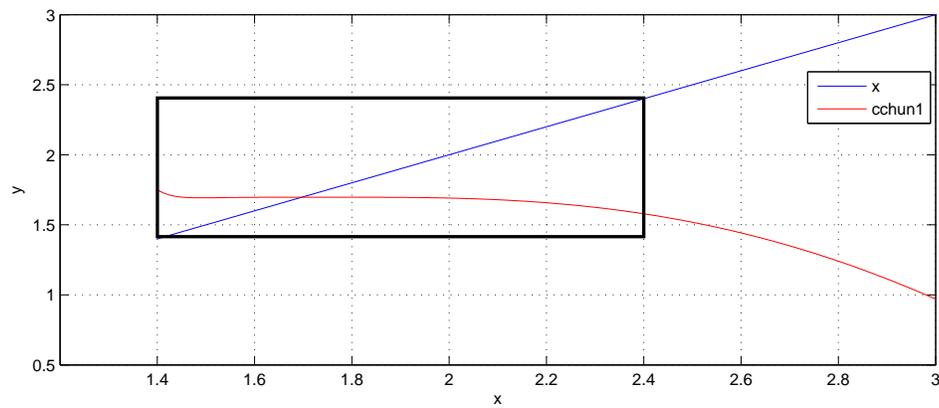


Figura 3.4: Función de iteración del método CCHUN1

en la cual hay puntos entre $[1.4, 1.5]$ tal que la imagen de dichos puntos no caen en $[1.4, 2.4]$. Por lo tanto nos quedamos con el intervalo $[1.5, 2.4]$.

En la Figura 3.6 se muestra el gráfico de la función de iteración del método CCHUN4 en la cual hay puntos entre $[1.5, 1.6]$ tal que la imagen de dichos puntos no caen en $[1.5, 2.4]$. Por lo tanto se escoge el intervalo $[1.6, 2.4]$.

En la Figura 3.7 se muestra de nuevo el gráfico de la función de iteración del método CCHUN1 en la cual hay puntos entre $[2.25, 2.4]$ tal que la imagen de dichos puntos no caen en $[1.6, 2.4]$, lo cual sugiere un nuevo intervalo, $[1.6, 2.25]$.

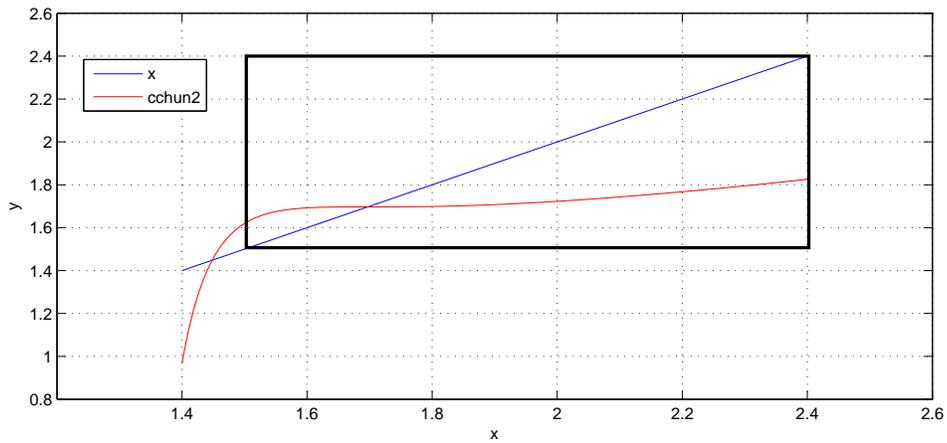


Figura 3.5: Función de iteración del método CCHUN2

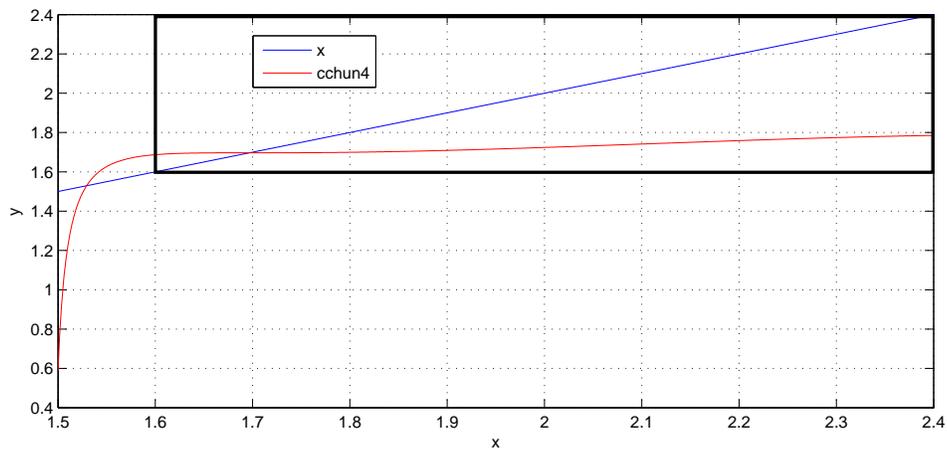


Figura 3.6: Función de iteración del método CCHUN4

Para terminar esta parte, se grafican las funciones de iteración de los métodos de NEWTON, JRS, MKN1, MKN2, NSPP, CCHUN1, CCHUN2, CCHUN3, CCHUN4, SG (ver Figura 3.8) y se escoge el intervalo $[a,b]=[1.6, 2.1]$ en el cual la imagen de cada punto de dicho intervalo mediante cada una de las funciones de iteración es un punto del mismo intervalo. El intervalo $[a,b]=[1.6, 2.1]$ es el que corresponde con el que está en el Apéndice A para la función 22.

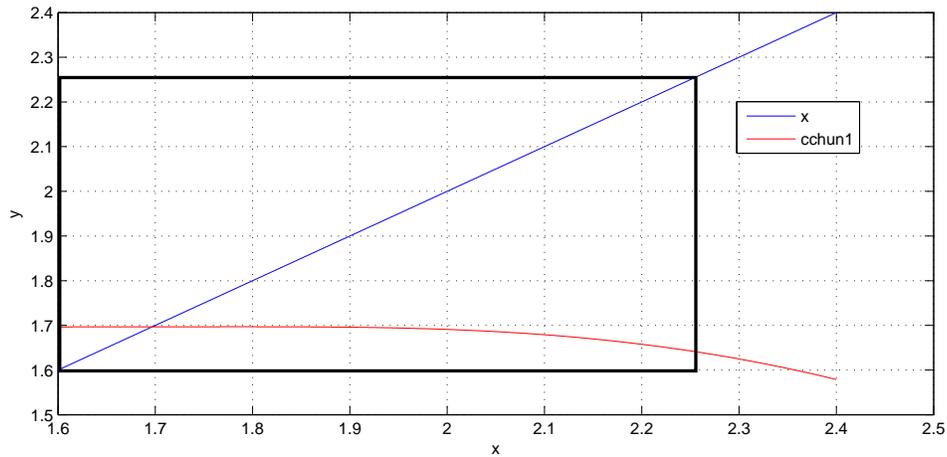


Figura 3.7: Función de iteración del método CCHUN1

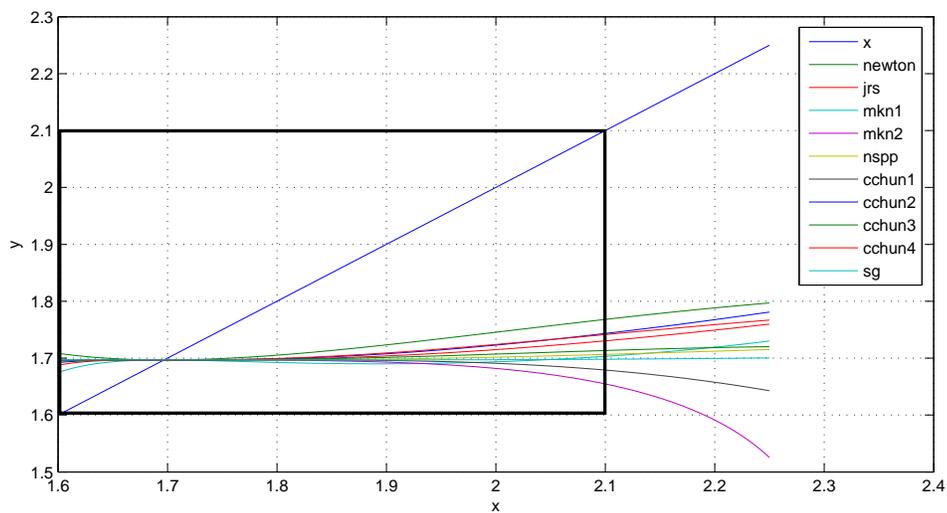


Figura 3.8: Funciones de iteración de los métodos iterativos: NEWTON, JRS, MKN1, MKN2, NSPP, CCHUN1, CCHUN2, CCHUN3, CCHUN4, SG .

3.2. Experimentación numérica.

Al aplicarse el Algoritmo A a la base de funciones se obtiene para cada método una muestra de 1500 valores de los tiempos de cómputos. A partir de la muestra de cada método se calcula la media y la desviación estándar y se construye un tabla en forma

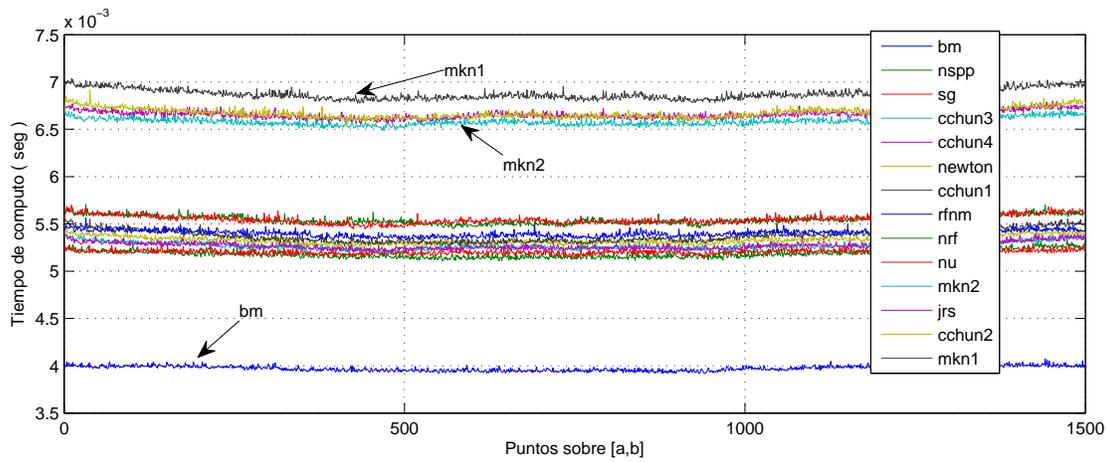


Figura 3.9: Tiempos de computo para cada uno de los métodos iterativos, sobre la base de funciones

ascendente a partir de las medias de la muestra de cada método (ver Cuadro 3.5). En la Figura 3.9 puede observar el comportamiento de los tiempos de cada método para cada uno de los 1500 puntos sobre el intervalo $[a,b]$.

M_i	Métodos	$\bar{\mu} (\times 10^{-3} \text{seg})$	Desviación Estándar ($\times 10^{-3} \text{seg}$)
1	BM	3.974248	0.027862
2	NSPP	5.184684	0.039859
3	SG	5.205262	0.022764
4	CCHUN3	5.271237	0.040149
5	CCHUN4	5.273448	0.037473
6	NEWTON	5.333006	0.040973
7	CCHUN1	5.376351	0.064658
8	RFNM	5.403892	0.033782
9	NRF	5.548643	0.044952
10	NU	5.551426	0.043933
11	MKN2	6.584139	0.039242
12	JRS	6.659552	0.040757
13	CCHUN2	6.677363	0.053237
14	MKN1	6.877562	0.055150

Cuadro 3.5: Métodos ordenados en función de la media o promedio, $\bar{\mu}$. M_i , $i = 1 : 14$, representa al método i . Por ejemplo, M_6 , representa al método 6 que es NEWTON.

3.2.1. Test estadístico.

Cada uno de los métodos tiene una media real, la cual proviene de promediar todos los tiempos de CPU, usado por cada método, que surgen al usar cada punto del intervalo $[a, b]$ (infinitos puntos). Estos tiempos no se pueden obtener, por tal razón se calculan los tiempos para k puntos en $[a, b]$ (muestra finita). A partir de los k tiempos obtenidos se obtiene una media muestral, la cual es una estimación de la media real.

En el caso de tener dos métodos M_1 y M_2 cuyas medias muestrales $\bar{\mu}_1$ y $\bar{\mu}_2$ estén muy cercanas, puede suceder que las medias muestrales sean estimaciones de una misma media real, μ , de los métodos M_1 y M_2 , es decir, que las medias reales de M_1 y M_2 sean $\mu_1 = \mu_2 = \mu$. También puede suceder que las medias muestrales $\bar{\mu}_1$ y $\bar{\mu}_2$ estén muy cercanas pero sean estimaciones de medias diferentes.

En [21, Cap. 9, pág. 243], al realizar estimación de la media real usando intervalo a partir de la media muestral, se asegura con cierta probabilidad, que la media real esta contenida en el intervalo que se construye. Dichos intervalos son los llamados intervalos de confianza.

Para el caso de una sola media, el intervalo de confianza simplemente produce límites de error sobre la media. Los valores contenidos en el intervalo se deben ver como valores razonables dados los datos experimentales.

En el caso de una diferencia entre dos medias, la interpretación se puede extender a una de comparación de las dos medias. Por ejemplo, si se tiene gran confianza de que una diferencia $\mu_1 - \mu_2$ es positiva, realmente se infiere que $\mu_1 > \mu_2$ con poco riesgo de caer en un error.

A partir de la información del Cuadro 3.5 surge la necesidad de saber si las medias muestrales de cada uno de los métodos que se presentan en dicho cuadro, son estimaciones de medias diferentes. Para ello, vamos a construir intervalos de confianza para la diferencia de dos medias (promedio) poblacionales (ver Sección 1.6) y usar la Observación 1.2 de la Sección 1.6 para concluir que pares de medias son diferentes.

Se van a construir intervalos de confianza¹ de 99% para $\mu_{i+1} - \mu_i$, $i = 1, 2, \dots, 13$,

¹Para mayores detalles de los intervalos de confianza, se remite al lector a la Subsección 1.6, pág:26

de la siguiente manera:

$$\mathbf{LI} < \mu_{i+1} - \mu_i < \mathbf{LS}$$

$$\mathbf{LI:} \quad = \quad (\bar{\mu}_{i+1} - \bar{\mu}_i) - 2,58 \sqrt{\frac{s_{i+1}^2}{n_{i+1}} + \frac{s_i^2}{n_i}}$$

$$\mathbf{LS:} \quad = \quad (\bar{\mu}_{i+1} - \bar{\mu}_i) + 2,58 \sqrt{\frac{s_{i+1}^2}{n_{i+1}} + \frac{s_i^2}{n_i}}$$

donde s_{i+1} y s_i son las desviaciones estándar de cada una de las muestras de tamaños n_{i+1}, n_i respectivamente.

M_{i+1}	M_i	$\mathbf{LI} (\times 10^{-3}\text{seg})$	$<$	$\mu_{i+1} - \mu_i$	$<$	$\mathbf{LS} (\times 10^{-3}\text{seg})$
2	1	1.207196	$<$	$\mu_2 - \mu_1$	$<$	1.213675
3	2	0.017519	$<$	$\mu_3 - \mu_2$	$<$	0.023635
4	3	0.062900	$<$	$\mu_4 - \mu_3$	$<$	0.069049
5	4	-0.001447	$<$	$\mu_5 - \mu_4$	$<$	0.005869
6	5	0.055859	$<$	$\mu_6 - \mu_5$	$<$	0.063257
7	6	0.038245	$<$	$\mu_7 - \mu_6$	$<$	0.048444
8	7	0.022680	$<$	$\mu_8 - \mu_7$	$<$	0.032400
9	8	0.141005	$<$	$\mu_9 - \mu_8$	$<$	0.148496
10	9	-0.001403	$<$	$\mu_{10} - \mu_9$	$<$	0.006970
11	10	1.028788	$<$	$\mu_{11} - \mu_{10}$	$<$	1.036636
12	11	0.071644	$<$	$\mu_{12} - \mu_{11}$	$<$	0.079182
13	12	0.013343	$<$	$\mu_{13} - \mu_{12}$	$<$	0.022276
14	13	0.195093	$<$	$\mu_{14} - \mu_{13}$	$<$	0.205305

Cuadro 3.6: Intervalos de confianza para la diferencia de las medias μ_i y μ_{i+1} , de los respectivos métodos M_i, M_{i+1} , $i = 1 : 13$. **LI** y **LS** son los límites inferior y superior del intervalo de confianza para la diferencia de las medias.

En el Cuadro 3.6 se muestra la diferencia de dos medias próximas entre sí, el límite inferior, **LI** y el límite superior, **LS** necesarios para construir los intervalos de confianza para cada una de las diferencias de las medias con un nivel de confianza del 99%. Por ejemplo, en la tercera fila se muestra lo siguiente:

$$\mathbf{4 \quad 3 \quad 0.062900 \quad < \quad \mu_4 - \mu_3 \quad < \quad 0.069049,}$$

y significa que se tiene un nivel de confianza del 99% de que la diferencia de las medias de los métodos **4** y **3** es un punto del intervalo $(0.062900 \times 10^{-3}\text{seg}, 0.069049 \times 10^{-3}\text{seg})$. Además, de la Observación 1.2 se puede asegurar con un nivel de confianza del 99%, que $\mu_3 < \mu_4$.

También se observa en el Cuadro 3.6, que

$$-0.001447 < \mu_5 - \mu_4 < 0.005869,$$

y significa que se tiene un nivel de confianza del 99% de que la diferencia de las medias de los métodos **5** y **4** es un punto del intervalo $(-0.001447 \times 10^{-3} \text{seg}, 0.005869 \times 10^{-3} \text{seg})$. Se aclara que no se puede establecer diferencias entre las medias μ_5 y μ_4 , es decir, no se puede concluir que $\mu_5 < \mu_4$ o $\mu_5 > \mu_4$ por el hecho de que $\mu_5 - \mu_4$ pertenece al intervalo

$$(-0.001447 \times 10^{-3} \text{seg}, 0.005869 \times 10^{-3} \text{seg}),$$

dicha diferencia puede ser un valor del intervalo $(-0.001447 \times 10^{-3} \text{seg}, 0 \times 10^{-3} \text{seg})$ o del intervalo $(0 \times 10^{-3} \text{seg}, 0.005869 \times 10^{-3} \text{seg})$ o ser **0**. Esta aclaratoria también se aplica a la diferencia $\mu_{10} - \mu_9$, dada en el Cuadro 3.6.

Del Cuadro 3.6 y de la Observación 1.2 se puede concluir con un 99% de confianza que $\mu_1 < \mu_2$, $\mu_2 < \mu_3$, $\mu_3 < \mu_4$, $\mu_5 < \mu_6$, $\mu_6 < \mu_7$, $\mu_7 < \mu_8$, $\mu_8 < \mu_9$, $\mu_{10} < \mu_{11}$, $\mu_{11} < \mu_{12}$, $\mu_{12} < \mu_{13}$ y $\mu_{13} < \mu_{14}$. Para el caso de $\mu_5 - \mu_4$ y $\mu_{10} - \mu_9$, como ya se ha dicho, no se puede establecer diferencias entre las respectivas medias μ_4, μ_5 y μ_9, μ_{10} . A partir de la manera como están ordenadas las medias reales de los respectivos métodos estudiados, se construye el Cuadro 3.7, en el cual los métodos están ordenados en función de la media muestral, con sus respectivos ordenes de convergencia, evaluaciones de funciones y eficiencia computacional.

Métodos	$\bar{\mu}$ ($\times 10^{-3} \text{seg}$)	r	Nf	Nf'	Nf''	$r^{1/d}$
BM	3.974248	3	2	0	0	1.7320
NSPP	5.184684	4	2	1	0	1.5874
SG	5.205262	6	3	1	0	1.5650
CCHUN3	5.271237	3	2	1	0	1.4422
CCHUN4	5.273448	3	2	1	0	1.4422
NEWTON	5.333006	2	1	1	0	1.4142
CCHUN1	5.376351	3	1	2	0	1.4422
RFNM	5.403892	3	2	1	0	1.4422
NRF	5.548643	3	2	1	0	1.4422
NU	5.551426	2	2	1	0	1.2599
MKN2	6.584139	3	1	1	1	1.4422
JRS	6.659552	3	1	1	1	1.4422
CCHUN2	6.677363	3	1	1	1	1.4422
MKN1	6.877562	2	1	1	1	1.2599

Cuadro 3.7: Métodos ordenados según su media o promedio $\bar{\mu}$. r es el orden de convergencia y las abreviaturas **Nf**, **Nf'** y **Nf''** representan el número de evaluaciones de **f**, **f'** y **f''** respectivamente. **d** es el número de evaluaciones de funciones y $r^{1/d}$ es la eficiencia computacional.

3.3. Aportes, conclusiones y comentarios finales

Como aportes de este trabajo se tienen los siguientes:

1. Se definió una estrategia para analizar la efectividad de los métodos iterativos tomando en cuenta el tiempo de CPU usado por los mismos.
2. Se realizaron demostraciones de Teoremas de convergencia para los siguientes métodos iterativos:

1 Método iterativo JRS. Para este método se enuncia y demuestra el Teorema 2.1, referente a la convergencia del método y orden de convergencia. Para demostrar el Teorema 2.1 se usa el Teorema 1.8 que también es un aporte de este trabajo.

2 Método iterativo MKN1. Para este método se demuestra el Teorema 2.2 referente a la convergencia del método y orden de convergencia. Para demostrar que el método es convergente se usó el Teorema 1.6 que también es un aporte de este trabajo. En el caso de orden de convergencia se usó la teoría de métodos iterativos.

3 Método iterativo MKN2. Para este método se demuestra el Teorema 2.3 referente a la convergencia del método y orden de convergencia.

4 Método iterativo NSPP. Para este método se demuestra el Teorema 2.4 referente a la convergencia del método. Dicha demostración es realizada a partir del uso del Teorema 1.6.

5 Métodos iterativos CCHUN1 y CCHUN2. Para estos métodos se demuestran los siguientes teoremas:

- a). Teorema 2.6, referente a convergencia y orden de convergencia del método CCHUN1.
- b). Teorema 2.7, referente a convergencia y orden de convergencia del método CCHUN2.

6 Métodos iterativos CCHUN3 y CCHUN4. Para estos métodos se demuestra el Teorema 2.8, referente a la convergencia de los métodos, en el cual se usa el Teorema 1.6.

Eficiencia de los métodos. En este trabajo se obtuvo una clasificación de los métodos iterativos según el tiempo de CPU usado, la cual se muestra en el Cuadro 3.7. A partir de la clasificación dada en el Cuadro 3.7, se pueden responder las preguntas planteadas al inicio de este capítulo.

En los resultados mostrados en el Cuadro 3.2, los métodos RFNM, NRF, CCHUN3, CCHUN4, CCHUN1, CCHUN2, MKN2 y JRS tienen igual eficiencia computacional. Sin embargo, a partir de los resultados de la clasificación (Cuadro 3.7) se logran clasificar estos métodos, pues los tiempos de CPU de cada uno de estos métodos resultan diferentes. No obstante, dentro de este grupo de métodos están los métodos CCHUN3 y CCHUN4 para los cuales no se puede establecer diferencias entre los tiempos de CPU. Para responder a la segunda pregunta realizada al inicio del capítulo, ¿influirá la primera derivada o la segunda derivada en los tiempos de CPU de los métodos que utilizan una o ambas derivadas?, se puede decir según la clasificación dada en el Cuadro 3.7 que sí influye, ya que los métodos MKN2, JRS, CCHUN2 y MKN1 que utilizan ambas derivadas son los que tienen mayor tiempo de CPU.

Los métodos NEWTON y NU, a pesar de tener eficiencias computacionales bajas, remontan posiciones al ser ordenados según el tiempo de CPU.

Los métodos BM, NSPP y SG con eficiencias computacionales mayores que los demás métodos, son los que tienen menor tiempo de CPU. De esta manera nos permite sugerir los métodos BM, NSPP y SG si se tiene los siguientes casos:

- Si se conoce el intervalo $[a, b]$ en el cual la función f tiene una raíz, se sugiere aplicar el método BM.
- Si no se conoce el intervalo $[a, b]$, se sugiere aplicar, en primer lugar el método NSPP y en segundo lugar el método SG.

Por último, como continuación de este trabajo, se espera realizar un estudio, según la estrategia definida, a métodos que sean tipo predictor-corrector y que dicho predictor sea NEWTON. Además se está trabajando en otra estrategia de clasificación en función del tiempo de CPU usado por cada uno de los métodos. La estrategia en la cual se está trabajando, se puede ver como una modificación de la estrategia ya definida.

Base de Funciones

A continuación se da una lista de funciones que se aplicaran a cada método iterativo estudiado en este trabajo, con una tolerancia de error de $0,5 * 10^{-14}$.

Función	Intervalo
▪ $f_1(x) = (\sin(x))^2 - x^2 + 1$	[1.2 , 2]
▪ $f_2(x) = x^2 - e^{(x)} - 3x + 2$	[0.2 , 0.4]
▪ $f_3(x) = \cos(x) - x$	[0.5 , 1.3]
▪ $f_4(x) = (x - 1)^3 - 1$	[1.8 , 4]
▪ $f_5(x) = x^3 - 10$	[2 , 2.5]
▪ $f_6(x) = xe^{x^2} - \sin(x)^2 + 3\cos(x) + 5$	[-1.25 , -1.1]
▪ $f_7(x) = e^{x^2+7x-30} - 1$	[2.976 , 3.015]
▪ $f_8(x) = x + \sin(x) - 2$	[0.6 , 1.4]
▪ $f_9(x) = x^6 - 36x^5 + 450x^4 - 2400x^3 + 5400x^2 - 4320x + 720$	[1.186 , 1.2]
▪ $f_{10}(x) = x^3 + 4x^2 - 10$	[1 , 2]
▪ $f_{11}(x) = \sin(x) - 0,5x$	[-2 , -1.75]
▪ $f_{12}(x) = (x - 1)^2 - 1$	[1.7 , 2.5]
▪ $f_{13}(x) = (x - 1)^3 - 2$	[2 2.5]
▪ $f_{14}(x) = 10x\exp(-x^2) - 1$	[0.06 , 0.3]

▪ $f_{15}(x) = x^{10} - 1$	[0.96 , 1.025]
▪ $f_{16}(x) = \sin(x)$	[-3.4 , -3]
▪ $f_{17}(x) = x^2 - 4$	[-3 , -1.4]
▪ $f_{18}(x) = e^x - 1 - \cos(\pi x)$	[-0.75 , -0.5]
▪ $f_{19}(x) = x^3 - e^{-x}$	[0.51 , 2]
▪ $f_{20}(x) = \sin(1/x) - x$	[0.8 , 1]
▪ $f_{21}(x) = e^x - 3x^2$	[-1 , -0.3]
▪ $f_{22}(x) = e^{\sin(x)} - x - 1$	[1.6 , 2.1]
▪ $f_{23}(x) = 1 - 11(x^{11})$	[0.781 , 0.85]
▪ $f_{24}(x) = xe^{-x} - 0,1$	[0.06 , 0.26]
▪ $f_{25}(x) = x^3 + 1$	[-1.2 , -0.8]
▪ $f_{26}(x) = 2xe^{-5} + 1 - 2e^{-5x}$	[0.1 , 0.2]
▪ $f_{27}(x) = 2xe^{-10} + 1 - 2e^{-10x}$	[0.06 , 0.1]
▪ $f_{28}(x) = 2xe^{-20} + 1 - 2e^{-20x}$	[0.025 , 0.06]
▪ $f_{29}(x) = 257x - (1 - 5x)^4$	[0.995 , 1.01]
▪ $f_{30}(x) = 6562x - (1 - 10x)^4$	[0.9996 , 1.001]
▪ $f_{31}(x) = 130322x - (1 - 20x)^4$	[0.9999 , 1.0001]
▪ $f_{32}(x) = x^2 + \sin(x/5) - 1/4$	[-1 , -0.5]
▪ $f_{33}(x) = (5x - 1)/(4 * x)$	[0.18 , 0.25]
▪ $f_{34}(x) = x - 3\log(x)$	[1 , 2.1]
▪ $f_{35}(x) = e^x - 4(x^2)$	[-1 , -0.35]
▪ $f_{36}(x) = e^{-x} + \cos(x)$	[1.3 , 2]
▪ $f_{37}(x) = 1000000e^x + 435000/x(e^x - 1) - 1564000$	[0.1009 , 0.101005]
▪ $f_{38}(x) = x^3 - 2x^2 - 5$	[2.4 , 4]
▪ $f_{39}(x) = 1/x - 1$	[0.5 , 1.2]

- $f_{40}(x) = e^{(1-x)} - 1$ [0.4 , 1.35]
- $f_{41}(x) = 1/x - \sin(x) + 1$ [-0.84 , -0.5]
- $f_{42}(x) = x - 2 - e^{-x}$ [1.9 , 2.3]
- $f_{43}(x) = x^2 - (1 - x)^5$ [0.2 , 0.5]
- $f_{44}(x) = x^3 + 4x^2 + 8x + 8$ [-2.5 , -1.5]
- $f_{45}(x) = \sin(x) + e^x + \log_{10}(x^2 + 1)$ [-2.6 , -2]
- $f_{46}(x) = (x - 4)^5 + 4x + 4$ [2.3 , 2.42]
- $f_{47}(x) = x^3 - 3x^2 - 9x + 9$ [4.2 , 5.5]

Referencias

- [1] J. F. Traub, *Iterative Methods for the Solution of Equations*, Prentice-Hall, New Jersey, 1964.
- [2] Richard L. Burden, J. Douglas Faires, *Análisis Numérico*, 7th ed., 2002.
- [3] Trevisan M. C., *Notas de Análisis Numérico*, Notas mimeografiadas, ULA, 2001.
- [4] J. R. Sharma, *A family of third-order methods to solve nonlinear equations by quadratic curves approximation*, Appl. Math. Comput. 184 (2007) 210-215.
- [5] Muhammad Aslam Noor and Khalida Inayat Noor, *Improved iterative methods for solving nonlinear equations*, Appl. Math. Comput. 184 (2007) 270-275.
- [6] Muhammad Aslam Noor and Khalida Inayat Noor, *Modified iterative methods with cubic convergence for solving nonlinear equations*, Appl. Math. Comput. 184 (2007) 322-325.
- [7] Kou Jisheng, Li Yitian and Wang Xiuhua, *A composite fourth-order iterative method for solving non-linear equations*, Appl. Math. Comput. 184 (2007) 471-475.
- [8] Changbum Chun, *Construction of third-order modifications of Newton's methods*, Appl. Math. Comput. 189 (2007) 662-668.
- [9] Changbum Chun, *A two-parameter third-order family of methods for solving non-linear equations*, Appl. Math. Comput. 189 (2007) 1822-1827.
- [10] J. R. Sharma and R. K. Guha, *A family of modified Ostrowski methods with accelerated sixth order convergence*, Appl. Math. Comput. 190 (2007) 111-115.
- [11] Giovanni Calderón, *Predictor-corrector type methods to solving nonlinear equations*, Sometido a revisión en Div. Mat. 2007.

-
- [12] F. A. Potra, V. Pták, *Nondiscrete induction and iterative processes*, Research Notes in Mathematics, vol. 103, Pitman, Boston, 1984.
- [13] L. W. Johnson, R. D. Riess, *Numerical Analysis*, Addison-Wesley, Reading, MA, 1977.
- [14] I. F. Steffensen, *Remarks on iteration*, Skand. Aktuarietidskr. 16(1933)64-72.
- [15] J.R. Sharma, *A composite third order Newton-Steffensen method for solving nonlinear equations*, Appl. Math. Comput. 169 (2005) 242-246.
- [16] G.B. Thomas Jr., R.L. Finney, *Calculus and Analytic Geometric*, Addison-Wesley, Publishing Company, Inc., MA 1998.
- [17] M. Aslam Noor, *Numerical Analysis and Optimization*, Lecture Notes, COMSATS Institute of Information Technology, Islamabad, Pakistan, 2006.
- [18] M. Aslam Noor, K. Inayat Noor, *Some iterative schemes for nonlinear equations*, Appl. Math. Comp., in press, doi:1016/j.amc.2006.05.084
- [19] A. S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*, McGraw-Hill, New York, 1970.
- [20] José Prieto. *Métodos iterativos para resolver ecuaciones no lineales*. Trabajo Especial de Grado. Universidad de Los Andes. Febrero 2008.
- [21] W. Ronal E., M. Raymond H., M. Sharon L., *Probabilidad y estadística para ingenieros*, 6a. ed., Prentice-Hall Hispanoamericana, S.A., México, 1999.