UVG

CURSO DE MÉTODOS NUMÉRICOS

Solución de ecuaciones de una variable II

Guatemala, 13 de julio de 2009

Método de Punto Fijo

<u>Definición</u>:

Sea F una función definida en el intervalo [a,b]. Decimos que $\alpha \in$ [a,b] es <u>punto fijo</u> de F, si $F(\alpha)=\alpha$.

<u>Definición (Recordatorio)</u>:

Un <u>espacio métrico</u> es un par (M,d), donde M es un conjunto no vacío y $d: M \times M \to \Re$, una función (llamada *métrica*) que cumple, para $x,y,z \in M$,

i) $d(x,y) \geq 0$; $d(x,y)=0$ ssi $x=y$;

ii) $d(x,y)=d(y,x)$;

iii) $d(x,y) \leq d(x,z)+d(z,y)$.

Método de Punto Fijo

Definición:

Una sucesión $\{x_n\}$ de elementos de un espacio métrico (M,d) es de *Cauchy* si, $d(x_n,x_m)\to 0$, cuando $n,m\to\infty$.

Definición:

Un espacio métrico en el que cada sucesión de Cauchy converge es *completo*.

Ejemplo:

El conjunto de números reales $\Re$, con la distancia usual es un espacio métrico completo.

Método de Punto Fijo

Definición:

Sea (M,d) un espacio métrico. Una función $F:M\to M$ es una *contracción*, si existe λ, $0<\lambda<1$, tal que, para cada x & y elementos de M,

$$d(F(x),F(y)) \leq \lambda d(x,y).$$

Teorema (del punto fijo de Banach)

Sean (M,d) un espacio métrico completo, y $F:M\to M$ una contracción. Entonces, existe un único punto fijo $\alpha \in M$ para F.

Nota: Cualquier función contractante sobre el conjunto de números reales (dotado de la métrica usual), tiene un punto fijo único.

Método de Punto Fijo

Método Iterativo:

Sea (M,d) un espacio métrico completo, y sea $F:M \rightarrow M$ es una contracción, entonces existe una única solución $\alpha \in M$, tal que $F(\alpha)= \alpha$, la cual es el límite de la sucesión $\{x_n\}$, donde $x_n=F(x_{n-1})$, n=1,2,3...,y donde x_0 es un punto arbitrario de M.

Al número x_n se le llama n-ésima iteración del método de punto fijo.

Nota:

- **Un error en el cálculo de la k-ésima iteración no afecta la convergencia del método.**
- **La convergencia del método es más rápida, mientras se seleccione el x_0 "cerca" del punto fijo.**

Método de Punto Fijo

Restricción a M=[a,b]:

En el caso que se considere la restricción del conjunto M a un intervalo cerrado de los reales, digamos [a,b], debemos resolver ahora el problema: ¿Cómo comprobamos que una función F es una contracción? (i.e., ¿Cómo se comprueba que F tiene un punto fijo en el intervalo considerado?

Recuerde: **El conjunto de números reales es completo, por lo que, para demostrar la convergencia del método a un punto fijo, basta probar que la función es una contracción.**

Método de Punto Fijo

<u>Teorema (5★)</u>:

Sea M=[a,b] y F:M→M, y suponga que existe la derivada F' en (a,b). Entonces, si se cumple que

$$\left|\frac{d}{dx}F(x)\right| \leq k < 1 \qquad \textbf{(*)}$$

para algún k, entonces F es una contracción.

<u>Nota:</u>

1. **A la condición (*) se le llama usualmente *Condición de Lipschitz.***
2. **Nótese que en el teorema anterior se supone que R es un espacio métrico, utilizando como métrica a usual de ese conjunto, es decir d(x,y)=|x-y |**

Método de Punto Fijo

Ejercicios:

1. **Resuelva la ecuación $tan x=x$, en el intervalo [4,5].**

2. **Resuelva la ecuación $x^3-x-1=0$, en el intervalo [1,2].**

3. **¿Puede proponer un estimado para el error de aproximación en cada iteración del método de punto fijo?**

4. **¿Cuál es la interpretación geométrica del método?**

Método de la Secante

<u>Problema:</u>

- **En cada aproximación x_n de la solución de F(x)=0, el método de NR evalúa dos funciones, F y F', en el punto x_{n-1}.**
- **Históricamente, el cálculo de esa derivada involucraba muchas dificultades. En la actualidad, esas dificultades se minimizan con los paquetes computacionales.**
- **Sin embargo, muchas son funciones no elementales, por lo que es deseable contar con métodos que, se asemejen a NR, pero que no incluyan la derivada en su ecuación de recurrencia.**

Método de la Secante

<u>Teorema (Método de la secante):</u>

Suponga que F es una función $C^2[a,b]$ y que existe un número $p\in[a,b]$, tal que $F(p)=0$. Entonces, si $F'(p)\neq 0$, entonces la sucesión $\{x_n\}$ definida por la iteración

$$x_{n+1} = x_n - \frac{F(x_n)\cdot(x_n - x_{n-1})}{F(x_n) - F(x_{n-1})}$$

converge al número p, para ciertas aproximaciones iniciales $x_0, x_1 \in[a,b]$.

Método de la Secante

Nota:

- **El método de la secante converge más rápido que el método de bisección, pero menos rápido que el método NR.**

- **El método de la secante considera las dos aproximaciones más recientes para la generación de la siguiente aproximación (¡¡aunque no estemos encerrando a la raíz de la ecuación!!!). ¿Qué pasaría si conserváramos un punto que está produciendo el encajonamiento de la raíz? (Note que esta es la misma idea que el método de bisección, con la única diferencia que en lugar de seleccionar el punto medio de cada intervalo como aproximación, se aplica en dicho subintervalo el método de la secante)**

Método de la Secante

<u>Nota:</u>

- **El método de la secante converge más rápido que el método de bisección, pero menos rápido que el método NR.**

- **El método de la secante considera las dos aproximaciones más recientes para la generación de la siguiente aproximación (¡¡aunque no estemos encerrando a la raíz de la ecuación!!!). ¿Qué pasaría si conserváramos un punto que está produciendo el encajonamiento de la raíz? (Note que esta es la misma idea que el método de bisección, con la única diferencia que en lugar de seleccionar el punto medio de cada intervalo como aproximación, se aplica en dicho subintervalo el método de la secante).**

Ceros de Polinomios

<u>Definición</u>: Un *polinomio de grado n* es una expresión de la forma $P(x)=a_n x^n+...+a_1 x+a_0$, $a_n \neq 0$, donde las a_i son constantes denominadas *coeficientes de* P, y donde *x* es una indeterminada.

Caso especial lo constituye el polinomio constante P(x)=0, al cual no se le asigna grado.

<u>Teorema</u> (Fundamental del Álgebra (D'Alambert 1746-Gauss 1799))

V1 El polinomio P(x) con grado $n \geq 1$ y de coeficientes reales tiene, a lo más, n raíces reales.

V2 El polinomio P(x) con grado $n \geq 1$ y de coeficientes complejos tiene exactamente n raíces complejas.

V3 Cada polinomio no constante con coeficientes complejos tiene una raíz

V4 Si P(x) es un polinomio de grado $n \geq 1$ con coeficientes reales o complejos, entonces P(x)=0 tiene al menos una raíz (posiblemente compleja) (Burden & Faires).

Paréntesis Histórico

- **Los estudios de Al-Juarizmi (Siglo IX) se enfocaron únicamente en números positivos, por lo que el TFA no era relevante.**
- **En el siglo XVI, Cardano fue quien propuso el estudio de las raíces de los polinomios con cantidades "mas generales que los números reales". Este descubrimiento fue derivado de su estudio de la ecuación x^3=15x+4. Estas cantidades se llamaron "números complejos".**
- **Bombelli (en *Algebra*, publicada en 1572), produjo un conjunto de reglas para manipular los "números complejos".**
- **En 1637, Descartes dijo que "uno puede imaginar que cada ecuación de grado n tiene n raíces, pero esas raíces imaginarias no corresponden a ninguna cantidad real".**
- **Viète mostró ecuaciones de grado n con n raíces.**
- **La primera conjetura respecto a que *todo* polinomio de grado n tiene n raíces, fue enunciada por el matemático flamenco Albert Girard, en 1629.**

Paréntesis Histórico

•Durante muchos años se consideró que el enunciado de Girard era autoevidente (¡postulado!).

• El matemático inglés Harriot propuso que un polinomio que se anula en *c*, tiene un factor *x-c*. Sin embargo, este resultado fue formalizado por Descartes en 1637 (*La géométrie)*.

•En 1702, Leibniz "probó" que el polinomio x^4+c^4 no puede escribirse como el producto de dos polinomios reales y cuadráticos. Este constituía un contraejemplo a la conjetura del ahora TFA. Claramente su argumentación era limitada.

•En 1742, Euler, mediante correspondencia con N. Bernoulli y con Goldbach, demostró que el contraejemplo de Leibniz era falso.

•En 1746, D'Alambert hizo el primer intento de demostrar formalmente (¡utilizando un método iterativo!), el TFA. La prueba es considerada incorrecta por utilizar argumentos circulares.

Paréntesis Histórico

•**Luego Euler demostró el TFA para polinomios con coeficientes reales y de grado a lo más 6. En 1749 presentó una demostración para el caso general de polinomios con coeficientes reales.**

• **En 1772, Lagrange presentó objeciones a la prueba de Euler.**

•**En 1795, Laplace intentó probar el TFA utilizando las propiedades del discriminante del polinomio. Su prueba tiene la dificultad que asume la existencia de las raíces.**

•**K.F. Gauss es acreditado como el primero en dar una demostración formal y general de FTA, en su tesis doctoral de 1799. Sin embargo, esta prueba carecía del rigor que en la actualidad se requiere para aceptarla.**

•**Desarrollos posteriores (destacando el suizo Argand, Cauchy y de nuevo Gauss (quien introdujo el concepto de número complejo en 1831), condujeron a mejoras en la calidad de las pruebas.**

•**En 1849 (¡50 años después!), Gauss presenta su prueba para el caso de coeficientes complejos).**

Ceros de Polinomios

<u>Corolario</u>:

Si P(x) es un polinomio de grado $n \geq 1$ con coeficientes reales o complejos, entonces existen constantes (reales o complejas) únicas $x_1,...,x_k$, $k \leq n$, y enteros positivos $m_1,...,m_k$, tales que $m_1+...+m_k=1$, entonces

$$P(x) = a_n (x - x_1)^{m_1} (x - x_2)^{m_2} \cdots (x - x_k)^{m_k}$$

<u>Corolario</u>:

Sean P(x) y Q(x) polinomios de grado a lo más n. Si $x_1,...,x_k$, $k>n$, son números distintos con $P(x_i)=Q(x_i)$, $i=1,2,...,k$, entonces P(x)=Q(x), para todos los valores de x.

Método de Horner y Deflación

- **El método de Horner es el conocido popularmente como *división sintética*.**
- **La idea es aplicar inicialmente el método NR a un polinomio P(x).**
- **Al encontrar una aproximación adecuada x_0 de la raíz del polinomio, podemos reescribir, por el método de Horner, a P(x) de la forma siguiente:**

$$P(x)=(x-x_0)Q(x)+b_0$$

- **Nótese que $P(x_0)=b_0$. Además, $P'(x)=Q(x)+(x-x_0)Q'(x)$, por lo que $P'(x_0)=Q(x_0)$.**

Método de Horner y Deflación

- **Nótese que al avanzar en la aproximación de la raíz, se tiene que, digamos, y_1 es uno de los ceros que buscamos. Entonces,**

$$P(x) \approx (x-y_1)Q_1(x)$$

- **Podemos aplicar el mismo procedimiento a $Q_1(x)$, y obtener una aproximación de la raíz y_2 de Q_1. Entonces,**

$$P(x) \approx (x-y_1)(x-y_2)Q_2(x)$$

- **En general, se deflata el grado del polinomio, en el procedimiento**

$$P(x) \approx (x-y_1)(x-y_2)\ldots(x-y_k)Q_k(x)$$

Método de Müller

- **La idea del método es análoga a la del método de la secante.**
- **El método de Müller (MMü) utiliza tres puntos iniciales (digamos x_0,x_1,x_2),y encuentra la parábola que pasa por ellos.**
- **Entonces, encuentra la aproximación de la raíz de la función, mediante la intersección x_3 de la parábola que se construyó, con el eje X.**
- **Se repite el procedimiento, encontrando x_4, a partir de x_1,x_2 y x_3.**

Método de Müller

- **La idea del método es análoga a la del método de la secante.**
- **El método de Müller (MMü) utiliza tres puntos iniciales (digamos x_0,x_1,x_2),y encuentra la parábola que pasa por ellos.**
- **Entonces, encuentra la aproximación de la raíz de la función, mediante la intersección x_3 de la parábola que se construyó, con el eje X.**
- **Se repite el procedimiento, encontrando x_4, a partir de x_1,x_2 y x_3.**

Método de Müller

<u>Deducción del método</u>

Consideremos los puntos

$$(x_0,f(x_0)),\ (x_1,f(x_1)),\ (x_2,f(x_2))$$

y encontremos la parábola que contenga a estos tres puntos.

Esta parábola se puede construir de diferentes formas, siendo la que nos conviene, en este caso,

$$P(x)=a(x-x_2)^2+b(x-x_2)+c.$$

Encontremos, ahora, las constantes a,b y c adecuadas.

Método de Müller

Para ello, evaluemos el polinomio

$$P(x)=a(x-x_2)^2+b(x-x_2)+c.$$

en cada uno de los puntos dados. Es decir,

$$f(x_0)=a(x_0-x_2)^2+b(x_0-x_2)+c$$
$$f(x_1)=a(x_1-x_2)^2+b(x_1-x_2)+c$$
$$f(x_2)=c$$

Entonces, dado que $c=f(x_2)$, debemos resolver el sistema de ecuaciones lineales:

Método de Müller

$$f(x_0) - f(x_2) = a(x_0 - x_2)^2 + b(x_0 - x_2)$$

$$f(x_1) - f(x_2) = a(x_1 - x_2)^2 + b(x_1 - x_2)$$

$$\Leftrightarrow$$

$$\begin{bmatrix} (x_0 - x_2)^2 & (x_0 - x_2) \\ (x_1 - x_2)^2 & (x_1 - x_2) \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} f(x_0) - f(x_2) \\ f(x_1) - f(x_2) \end{bmatrix}$$

Entonces,

$$\det \begin{bmatrix} (x_0 - x_2)^2 & (x_0 - x_2) \\ (x_1 - x_2)^2 & (x_1 - x_2) \end{bmatrix} = (x_0 - x_2)^2 (x_1 - x_2) - (x_1 - x_2)^2 (x_0 - x_2)$$

$$= (x_0 - x_2)(x_1 - x_2)(x_0 - x_1)$$

Método de Müller

Resolviendo por el método de Cramer, obtenemos:

$$a = \frac{\det\begin{bmatrix} f(x_0)-f(x_2) & x_0-x_2 \\ f(x_1)-f(x_2) & x_1-x_2 \end{bmatrix}}{(x_0-x_2)(x_1-x_2)(x_0-x_1)} =$$

$$= \frac{[f(x_0)-f(x_2)](x_1-x_2)-[f(x_1)-f(x_2)](x_0-x_2)}{(x_0-x_2)(x_1-x_2)(x_0-x_1)}$$

$$b = \frac{\det\begin{bmatrix} (x_0-x_2)^2 & f(x_0)-f(x_2) \\ (x_1-x_2)^2 & f(x_1)-f(x_2) \end{bmatrix}}{(x_0-x_2)(x_1-x_2)(x_0-x_1)} =$$

$$= \frac{[f(x_1)-f(x_2)](x_0-x_2)^2-[f(x_0)-f(x_2)](x_1-x_2)^2}{(x_0-x_2)(x_1-x_2)(x_0-x_1)}$$

Método de Müller

Conocidos los parámetros a,b y c, entonces encontramos la aproximación para la raíz de f(x), haciendo P(x)=0. Para ello aplicamos la fórmula de Viéte a la parábola:

$$P(x) = a(x - x_2)^2 + b(x - x_2) + c = 0,$$

$$\Rightarrow x - x_2 = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Entonces, haciendo $x = x_3$, se tiene

$$x_3 = x_2 - \frac{2c}{b \pm \sqrt{b^2 - 4ac}}$$

Método de Müller

Nótese que en esta última expresión se muestra que existen dos posibilidades para x_3, dependiendo del signo que se tome frente al radical.

Dado que el objetivo es que las aproximaciones de las raíces formen una sucesión de Cauchy, entonces requerimos que en cada iteración el denominador del cociente sea cada vez mayor...de esa cuenta, nos conviene tomar el signo del radical igual al signo de b.

$$\Rightarrow x_3 = x_2 - \frac{2c}{b + \text{signo}(b)\sqrt{b^2 - 4ac}}$$

Método de Müller

Obtenida la aproximación, se procede de forma análoga, tomando ahora x_1,x_2 y x_3, y encontrando mediante la construcción de la parábola que pasa por los puntos correspondientes y de esa forma calcular la nueva aproximación x_4.

Debe notarse que el discriminante de la parábola podría ser negativo, por lo que el método tendería a aproximar raíces complejas del polinomio.

Método de Aitken

<u>Objetivo:</u>

Acelerar la convergencia de una sucesión que converge linealmente, independientemente del punto inicial y de la regla de aplicación.

<u>Idea del método</u>: Supóngase que $\{y_n\}$ es una sucesión que converge linealmente a un límite c. Entonces, el método consiste en producir, a partir de los x_n, otra sucesión $\{\hat{y}_n\}$ que converge más rápidamente a c.

<u>¿Cómo?</u>

$$\hat{y}_n = y_n - \frac{\left(y_{n+1} - y_n\right)^2}{y_{n+2} - 2y_{n+1} + y_n}$$

(Ver detalle en Burden & Faires, sección 2.5)

Método de Aitken

<u>Notación:</u>

•Dada la sucesión $\{y_n\}$, se define la <u>diferencia progresiva</u> $\Delta y_n := y_{n+1} - y_n,\ n \geq 0$.

•En el caso de potencias de la diferencia progresiva, se tiene que $\Delta^k y_n = \Delta(\Delta^{(k-1)} y_n)$.

•Lo anterior indica que $\Delta^2 y_n = \Delta(\Delta y_n) = \Delta(y_{n+1} - y_n)$, entonces por la aplicación recursiva, la segunda diferencia progresiva es

$$\Delta^2 y_n = (y_{n+2} - y_{n+1}) - (y_{n+1} - y_n) = y_{n+2} - 2y_{n+1} + y_n$$

•Entonces, el generador de la sucesión acelerada en el método de Aitken, se reescribe como,

$$\hat{y}_{n+1} = y_n - \frac{(\Delta y_n)^2}{\Delta^2 y_n}$$

Método de Aitken

¿Qué quiere decir que la sucesión $\{\hat{y}_n\}$ converja "más rápidamente?, i.e., ¿En qué sentido este método es un acelerador?

<u>Teorema</u>:

Suponga que $\{y_n\}$ es una sucesión que converge linealmente a c, y que para valores suficientemente grandes de n, se tiene que $(y_n-c)(y_{n+1}-c)>0$. Entonces, la sucesión $\{\hat{y}_n\}$ converge a c con mayor rapidez que $\{y_n\}$, en el sentido que

$$\operatorname*{Lim}_{n\to\infty} \frac{\hat{y}_n - c}{y_n - c} = 0$$

Ceros Múltiples

<u>Definición:</u>

El número c es un cero de multiplicidad m de la función F(x), si para $x \neq c$, se tiene la representación $F(x)=(x-c)^m q(x)$, donde

$$\lim_{x \to c} q(x) \neq 0$$

<u>Teorema A:</u> La función F en $C^1[a,b]$ tiene un cero simple en $c \in (a,b)$ ssi $F(c)=0$, pero $F'(c) \neq 0$.

<u>Teorema B:</u> La función F en $C^m[a,b]$ tiene un cero de multiplicidad m en $c \in (a,b)$ ssi se cumple que $0=F(c)=F'(c)=F''(c)=\ldots=F^{(m-1)}(c)$, pero $F^{(m)}(c) \neq 0$.

Ceros Múltiples

<u>Resultados:</u>

- **El teorema A indica que, alrededor de $x=c$ existe un intervalo (vecindad), tal que el método NR converge cuadráticamente a c, para cualquier aproximación inicial x_0, siempre que c sea un cero simple.**
- **En Burden & Faires (pág. 83) se presenta el ejemplo de una función ($F(x)=e^x-x-1$) que tiene un cero múltiple en $x=0$. Al aplicar NR a esta función, considerando $x_0=1$, el método converge a la raíz, pero pierde eficiencia (i.e., converge menos que cuadráticamente).**

<u>Problema</u>: ¿Es posible ajustar NR para que converja cuadráticamente a ceros múltiples?

Análisis de Error para Métodos Iterativos

<u>Definición:</u>

Sea $\{x_n\}$ una sucesión convergente a x, para $x_n \neq x$, para cada n. Si existen constantes positivas λ y α con:

$$\lim_{n \to \infty} \frac{|x_{n+1} - x|}{|x_n - x|^{\alpha}} = \lambda$$

entonces, $\{x_n\}$ <u>converge a x con orden α y constante de error asintótica λ .</u>

<u>Nota</u>: Se dice que un método iterativo $x_n = F(x_{n-1})$ es de orden α, si la sucesión $\{x_n\}$ converge a x con orden α.

Análisis de Error para Métodos Iterativos

Resultados:

1. **El método de punto fijo de Banach converge linealmente a su solución.**
2. **El método de NR converge cuadráticamente.**
3. **El método de la secante tiene un orden de convergencia α, tal que, $1< \alpha<2$.**

Problema: ¿Es posible acelerar la convergencia de los métodos estudiados? En particular, ¿Es posible acelerar la convergencia del método NR?