CURSO DE METODOS NUMERICOS

QUINTA PARTE

SOLUCIONES NUMERICAS DE SISTEMAS NO LINEALES

CAPITULO XXI. PUNTOS FIJOS PARA FUNCIONES DE VARIAS VARIABLES

1. PRELIMINARES

La forma general de un sistema de ecuaciones no lineales es:

$$\begin{aligned} f_1(x_1,x_2,\ldots,x_n) &= 0 \ , \\ f_2(x_1,x_2,\ldots,x_n) &= 0 \ , \\ \ldots \quad \ldots & \quad \ldots \\ f_n(x_1,x_2,\ldots,x_n) &= 0 \ , \end{aligned} \tag{XXI.1}$$

donde cada función f_i puede tomarse como una aplicación de un vector $\mathbf{x}$ del espacio $n-$dimensional $\mathcal{R}^n$, $\mathbf{x} = (x_1,\ x_2,\ \ldots,\ x_n)^t$, en la recta real $\mathcal{R}$. Alternativamente, el sistema puede representarse definiendo una función $\mathbf{F}$, de $\mathcal{R}^n$ en $\mathcal{R}^n$ por

$$\mathbf{F}(x_1,x_2,\ldots,x_n) = (f_1(x_1,\ldots,x_n), f_2(x_1,\ldots,x_n),\ldots,f_n(x_1,\ldots,x_n))^t = \mathbf{0} \ .$$

Usando notación vectorial para representar las variables x_i, el sistema $(XXI.1)$ asume la forma:

$$\mathbf{F}(\mathbf{x}) = \mathbf{0} \ . \tag{XXI.2}$$

Las funciones $f_1, f_2, \ldots, f_n$ se llaman **funciones coordenadas de F**.

Antes de discutir la solución de un sistema dado en las formas $(XXI.1)$ ó $(XXI.2)$, necesitamos considerar algunos resultados concernientes a la continuidad y a la diferenciabilidad de funciones de $\mathcal{R}^n$ en $\mathcal{R}^n$.

Definición. Sea f una función definida en un conjunto $D \subset \mathcal{R}^n$ y con valores en $\mathcal{R}$. Se dice que la función f tiene **límite** L en $\mathbf{x}_0$, denotado por $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = L$, si, dado cualquier $\varepsilon > 0$, existe un número $\delta > 0$ con la propiedad de que $|f(\mathbf{x}) - L| < \varepsilon$ siempre que $x \in D$ y $0 < ||\mathbf{x} - \mathbf{x}_0|| < \delta$.

Debe hacerse notar que la existencia de un límite es independiente de la norma vectorial particular usada debido a la equivalencia de las normas vectoriales en $\mathcal{R}^n$.

Definición. Sea f una función del conjunto $D \subset \mathcal{R}^n$ en $\mathcal{R}$. Se dice que la función f es **continua** en $\mathbf{x}_0 \in D$ siempre y cuando el $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x})$ exista y $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0)$. Se dice, además que f es **continua** en un conjunto D si f es continua en cada punto de D. Este concepto se expresa escribiendo $f \in \mathcal{C}(D)$.

Definición. Sea $\mathbf{F}$ una función del conjunto $D \subset \mathcal{R}^n$ en $\mathcal{R}^n$ y supongamos que $\mathbf{F}$ tiene la representación $\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}),\ldots,f_n(\mathbf{x}))^t$, donde f_i para cada i es una aplicación de $\mathcal{R}^n$ en $\mathcal{R}$. Definimos $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} \mathbf{F}(\mathbf{x}) = \mathbf{L} = (L_1, L_2, \ldots, L_n)^t$ si y sólo si $\lim\limits_{\mathbf{x}\to\mathbf{x}_0} f_i(\mathbf{x}) = L_i$ para cada $i = 1, 2, \ldots, n$.

Definición. Sea $\mathbf{F}$ una función del conjunto $D \subset \mathcal{R}^n$ en $\mathcal{R}^n$ con la representación $\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}),\ldots,f_n(\mathbf{x}))^t$. Se dice que la función $\mathbf{F}$ es **continua** en $\mathbf{x}_0 \in D$ si

$\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{F}(\mathbf{x})$ existe y $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{F}(\mathbf{x}) = \mathbf{F}(\mathbf{x}_0)$. Se dice que $\mathbf{F}$ es **continua** en el conjunto D si $\mathbf{F}$ es continua en cada $\mathbf{x}$ de D. Este concepto se expresa escribiendo $\mathbf{F} \in \mathcal{C}(D)$.

Enunciaremos un Teorema que relaciona la continuidad de una función de n variables en un punto con las derivadas parciales de la función en ese punto.

Teorema XXI.1

Sea f una función del conjunto $D \subset \mathcal{R}^n$ en $\mathcal{R}^n$ y $\mathbf{x}_0 \in D$. Si existen constantes $\delta > 0$ y $K > 0$ con

$$\Big|\frac{\partial f(\mathbf{x})}{\partial x_j}\Big| \leq K \qquad \text{para cada} \;\; j = 1, 2, \ldots, n \;,$$

siempre que $0 < ||\mathbf{x} - \mathbf{x}_0|| < \delta$ y $\mathbf{x} \in D$, entonces f es continua en $\mathbf{x}_0$.

2. METODO DE ITERACION Y EJEMPLOS

En el capítulo XVI, se desarrolló un proceso iterativo para resolver la ecuación $f(x) = 0$ transformando primero esta ecuación en una ecuación de la forma $x = g(x)$. La función g tiene sus puntos fijos precisamente en las soluciones de la ecuación original. Aquí se investigará un procedimiento similar para funciones de $\mathcal{R}^n$ en $\mathcal{R}^n$.

Sea dado un sistema de ecuaciones no lineales de un tipo especial:

$$\begin{aligned} x_1 &= g_1(x_1, x_2, \ldots, x_n) \;, \\ x_2 &= g_2(x_1, x_2, \ldots, x_n) \;, \\ \ldots & \quad \ldots \quad \ldots \quad \ldots \\ x_n &= g_n(x_1, x_2, \ldots, x_n) \;, \end{aligned} \tag{XXI.3}$$

donde las funciones $g_1, \; g_2, \ldots, g_n$ son reales, definidas y continuas en un conjunto D, vecindad de una solución separada $(x_1^*, x_2^*, \ldots, x_n^*)$ del sistema $(XXI.3)$. De forma más compacta el sistema $(XXI.3)$ se puede escribir como:

$$\mathbf{x} = \mathbf{G}(\mathbf{x}) \;. \tag{XXI.4}$$

Definición. Se dice que una función $\mathbf{G} : D \subset \mathcal{R}^n \to \mathcal{R}^n$ tiene un **punto fijo** en $\mathbf{p} \in D$ si $\mathbf{G}(\mathbf{p}) = \mathbf{p}$.

Para hallar la raíz vectorial $\mathbf{p} = (p_1, p_2, \ldots, p_n)$ de la ecuación $(XXI.4)$, frecuentemente resulta conveniente utilizar el **método de iteración**

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}) \;, \tag{XXI.5}$$

$k = 0, 1, 2, \ldots$, donde la aproximación inicial $\mathbf{x}^{(0)} \approx \mathbf{p}$. Obsérvese que si el proceso de iteración $(XXI.5)$ converge, entonces el valor límite $\xi = \lim_{k\to\infty} \mathbf{x}^{(k)}$ es definitivamente una raíz de la ecuación $(XXI.4)$, y pasando al límite en $(XXI.5)$ para $k \to \infty$ tendremos, en virtud de la continuidad de la función $\mathbf{G}(\mathbf{x})$,

$$\xi = \lim_{k\to\infty} \mathbf{x}^{(k+1)} = \mathbf{G}(\lim_{k\to\infty} \mathbf{x}^{(k)}) = \mathbf{G}(\xi) \;.$$

De este modo, ξ es una raíz de la ecuación vectorial $(XXI.4)$. Si, además, todas las aproximaciones $\mathbf{x}^{(k)}$ $(k = 0, 1, 2, \ldots)$ pertenecen al dominio D y $\mathbf{p}$ es una raíz única del sistema en D, entonces $\xi = \mathbf{p}$.

El método de iteración puede aplicarse también al sistema general $(XXI.1)$ ó $(XXI.2)$, donde $\mathbf{F}(\mathbf{x})$ es una función vectorial, definida y continua en la vecindad D de una raíz separada $\mathbf{p}$. Por ejemplo, escribamos nuevamente este sistema de la forma

$$\mathbf{x} = \mathbf{x} + \Lambda\ \mathbf{F}(\mathbf{x}) = \mathbf{G}(\mathbf{x}) \ , \qquad (XXI.6)$$

donde Λ es una matriz no singular. El método ordinario de iteración $(XXI.5)$ es directamente aplicable a esta ecuación. Si la función $\mathbf{F}(\mathbf{x})$ tiene una derivada continua $\mathbf{F}'(\mathbf{x})$ en D, se sigue que

$$\mathbf{G}'(\mathbf{x}) = I + \Lambda\ \mathbf{F}'(\mathbf{x}) = I + \Lambda\ J(\mathbf{x}) \ ,$$

donde la matriz $J(\mathbf{x}) = \mathbf{F}'(\mathbf{x})$ es la **matriz Jacobiana**, definida como $J(\mathbf{x}) = J_{ij} = (\frac{\partial f_i(\mathbf{x})}{\partial x_j})$. Entonces elegiremos (debido a lo que se demostrará más adelante) la matriz Λ de forma que

$$\mathbf{G}'(\mathbf{x}^{(0)}) = I + \Lambda\ J(\mathbf{x}^{(0)}) = 0 \ ,$$

de donde, si la matriz $J(\mathbf{x}^{(0)}) = \mathbf{F}'(\mathbf{x}^{(0)})$ es no singular, tendremos

$$\Lambda = -[J(\mathbf{x}^{(0)})]^{-1} \ .$$

Si $\det[J(\mathbf{x}^{(0)})] = 0$, debe eligirse una aproximación inicial $\mathbf{x}^{(0)}$ diferente.

Es decir, es esencialmente el proceso de Newton modificado (que veremos en un próximo capítulo) aplicado a la ecuación $(XXI.2)$.

Ejemplo. Utilícese el método de iteración para dar una solución aproximada del sistema

$$\begin{aligned} x_1^2 + x_2^2 &= 1 \ , \\ x_1^3 - x_2 &= 0 \ . \end{aligned}$$

A partir de una construcción gráfica, puede verse que el sistema dado tiene dos soluciones que difieren únicamente en el signo. Nos limitaremos a la solución positiva. Se puede tomar $\mathbf{x}^{(0)} = (0.9, 0.5)^t$ para la aproximación inicial de la solución positiva. Estableciendo

$$\mathbf{F}(\mathbf{x}) = \begin{pmatrix} x_1^2 + x_2^2 - 1 \\ x_1^3 - x_2 \end{pmatrix}$$

tenemos

$$J(\mathbf{x}) = \mathbf{F}'(\mathbf{x}) = \begin{pmatrix} 2\ x_1 & 2\ x_2 \\ 3\ x_1^2 & -1 \end{pmatrix}$$

de donde

$$J(\mathbf{x}^{(0)}) = \begin{pmatrix} 1.8 & 1 \\ 2.43 & -1 \end{pmatrix} \ , \qquad \text{y} \qquad \det[J(\mathbf{x}^{(0)})] = -1.8 - 2.43 = -4.23 \ .$$

Como la matriz $J(\mathbf{x}^{(0)})$ es no singular, existe una matriz inversa, y de este modo

$$\Lambda = -[J(\mathbf{x}^{(0)})]^{-1} = \frac{1}{4.23}\begin{pmatrix} -1 & -1 \\ -2.43 & 1.8 \end{pmatrix} .$$

Ahora hacemos

$$\begin{aligned} \mathbf{G}(\mathbf{x}) &= \mathbf{x} + \Lambda\ \mathbf{F}(\mathbf{x}) = \\ &= \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - \frac{1}{4.23}\begin{pmatrix} 1 & 1 \\ 2.43 & -1.8 \end{pmatrix}\begin{pmatrix} x_1^2 + x_2^2 - 1 \\ x_1^3 - x_2 \end{pmatrix} . \end{aligned}$$

Usando esta última fórmula hallamos las aproximaciones sucesivas siguientes para la solución del sistema dado (con aritmética de cuatro dígitos significativos):

$$\begin{aligned} \mathbf{x}^{(1)} &= \begin{pmatrix} x_1^{(0)} \\ x_2^{(0)} \end{pmatrix} - \frac{1}{4.23}\begin{pmatrix} 1 & 1 \\ 2.43 & -1.8 \end{pmatrix}\begin{pmatrix} {x_1^{(0)}}^2 + {x_2^{(0)}}^2 - 1 \\ {x_1^{(0)}}^3 - x_2^{(0)} \end{pmatrix} \\ &= \begin{pmatrix} 0.9 \\ 0.5 \end{pmatrix} - \begin{pmatrix} 0.06832 \\ -0.06298 \end{pmatrix} = \begin{pmatrix} 0.8317 \\ 0.5630 \end{pmatrix} \\ \mathbf{x}^{(2)} &= \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} - \frac{1}{4.23}\begin{pmatrix} 1 & 1 \\ 2.43 & -1.8 \end{pmatrix}\begin{pmatrix} {x_1^{(1)}}^2 + {x_2^{(1)}}^2 - 1 \\ {x_1^{(1)}}^3 - x_2^{(1)} \end{pmatrix} \\ &= \begin{pmatrix} 0.8317 \\ 0.5630 \end{pmatrix} - \begin{pmatrix} 0.005035 \\ -0.0002700 \end{pmatrix} = \begin{pmatrix} 0.8267 \\ 0.5633 \end{pmatrix} \\ \mathbf{x}^{(3)} &= \begin{pmatrix} x_1^{(2)} \\ x_2^{(2)} \end{pmatrix} - \frac{1}{4.23}\begin{pmatrix} 1 & 1 \\ 2.43 & -1.8 \end{pmatrix}\begin{pmatrix} {x_1^{(2)}}^2 + {x_2^{(2)}}^2 - 1 \\ {x_1^{(2)}}^3 - x_2^{(2)} \end{pmatrix} \\ &= \begin{pmatrix} 0.8267 \\ 0.5633 \end{pmatrix} - \begin{pmatrix} 0.0006383 \\ -0.0001489 \end{pmatrix} = \begin{pmatrix} 0.8261 \\ 0.5634 \end{pmatrix} \\ \mathbf{x}^{(4)} &= \begin{pmatrix} x_1^{(3)} \\ x_2^{(3)} \end{pmatrix} - \frac{1}{4.23}\begin{pmatrix} 1 & 1 \\ 2.43 & -1.8 \end{pmatrix}\begin{pmatrix} {x_1^{(3)}}^2 + {x_2^{(3)}}^2 - 1 \\ {x_1^{(3)}}^3 - x_2^{(3)} \end{pmatrix} \\ &= \begin{pmatrix} 0.8261 \\ 0.5634 \end{pmatrix} - \begin{pmatrix} 0.00002364 \\ -0.0002426 \end{pmatrix} = \begin{pmatrix} 0.8261 \\ 0.5636 \end{pmatrix} \end{aligned}$$

y así sucesivamente. Finalizando con la cuarta aproximación, tenemos las raíces

$$x_1 = 0.8261, \qquad x_2 = 0.5636$$

y $\mathbf{F}(\mathbf{x}) = (0.0000, 0.0001)^t$, con una exactitud de 10^{-4}.

3. CONDICIONES PARA LA CONVERGENCIA DEL PROCESO DE ITERACION

Definición. La aplicación $\mathbf{y} = \mathbf{G}(\mathbf{x})$ se denomina **aplicación sobreyectiva** o **de contracción** en el dominio D si existe una fracción propia q, $(0 \le q < 1)$, tal que para dos puntos cualesquiera $\mathbf{x}_1,\ \mathbf{x}_2 \in D$ sus imágenes $\mathbf{y}_1 = \mathbf{G}(\mathbf{x}_1)$ y $\mathbf{y}_2 = \mathbf{G}(\mathbf{x}_2)$ satisfagan la condición

$$||\mathbf{y}_1 - \mathbf{y}_2|| \le q\ ||\mathbf{x}_1 - \mathbf{x}_2|| \ ,$$

para cualquier norma canónica.

Teorema XXI.2

Sea D un dominio cerrado y la aplicación $\mathbf{y} = \mathbf{G}(\mathbf{x})$ una aplicación sobreyectiva. En tal caso, si para el proceso de iteración

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}) \ , \qquad (XXI.5)$$

todas las aproximaciones sucesivas $\mathbf{x}^{(k)}$ están en D $(k = 0, 1, 2, \ldots)$, quiere decir que:

(a) independientemente de la elección de la aproximación inicial $\mathbf{x}^{(0)}$ el proceso de iteración $(XXI.5)$ converge, es decir existe el límite $\mathbf{p} = \lim_{k\to\infty} \mathbf{x}^{(k)}$;

(b) el vector límite $\mathbf{p}$ es la única solución de la ecuación $\mathbf{x} = \mathbf{G}(\mathbf{x})$ en el dominio D;

(c) se cumple la estimación

$$||\mathbf{p} - \mathbf{x}^{(k)}|| \le \frac{q^k}{1-q} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| \ . \qquad (XXI.7)$$

Demostración: (a) para probar la convergencia de la secuencia de aproximaciones $\mathbf{x}^{(k)}$ $(k = 0, 1, 2, \ldots)$, apliquemos el criterio de Cauchy para obtener

$$\begin{aligned} ||\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}|| &= ||(\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}) + (\mathbf{x}^{(k+2)} - \mathbf{x}^{(k+1)}) + \ldots + (\mathbf{x}^{(k+p)} - \mathbf{x}^{(k+p-1)})|| \\ &\le ||\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}|| + ||\mathbf{x}^{(k+2)} - \mathbf{x}^{(k+1)}|| + \ldots + ||\mathbf{x}^{(k+p)} - \mathbf{x}^{(k+p-1)}|| \end{aligned}$$

Utilizando la relación de iteración y la condición de contracción, obtendremos sucesivamente:

$$\begin{aligned} ||\mathbf{x}^{(s+1)} - \mathbf{x}^{(s)}|| &= ||\mathbf{G}(\mathbf{x}^{(s)}) - \mathbf{G}(\mathbf{x}^{(s-1)})|| \\ &\le q \ ||\mathbf{x}^{(s)} - \mathbf{x}^{(s-1)}|| \le q^2 \ ||\mathbf{x}^{(s-1)} - \mathbf{x}^{(s-2)}|| \le \ldots \\ &\le q^s \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| \ , \end{aligned}$$

donde $s \ge 0$. Por consiguiente, hallamos

$$||\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}|| \le q^k \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| + q^{k+1} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| + \ldots + q^{k+p-1} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| \ ,$$

ó, utilizando la fórmula de la suma de los términos de una progresión geométrica, hallaremos

$$||\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}|| \le \frac{q^k - q^{k+p}}{1-q} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| \le \frac{q^k}{1-q} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}|| \ .$$

Como $0 \le q < 1$ y, como $q^k \to 0$ para $k \to \infty$, se deduce que para cualquier $\varepsilon > 0$ existe $N = N(\varepsilon)$ tal que para $k > N(\varepsilon)$ y $p > 0$ se cumple la desigualdad

$$||\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}|| < \varepsilon$$

esto es, el criterio de Cauchy es válido para la secuencia $\mathbf{x}^{(k)}$ $(k = 0, 1, 2, \ldots)$. Por consiguiente, $\mathbf{p} = \lim_{k\to\infty} \mathbf{x}^{(k)}$ y $\mathbf{p} \in D$ ya que el dominio D es cerrado.

(b) El vector $\mathbf{p}$ es una solución de la ecuación $\mathbf{x} = \mathbf{G}(\mathbf{x})$ ya que, pasando al límite en la relación de iteración $(XXI.5)$ para $k \to \infty$ y teniendo en cuenta la continuidad en D de la función vectorial $\mathbf{G}(\mathbf{x})$, tenemos

$$\mathbf{p} = \lim_{k\to\infty} \mathbf{x}^{(k)} = \lim_{k\to\infty} \mathbf{G}(\mathbf{x}^{(k-1)}) = \mathbf{G}(\mathbf{p}) \ .$$

Esta solución es única en D. En efecto, sea $\mathbf{p}^{'}$ otra solución, es decir $\mathbf{p}^{'} = \mathbf{G}(\mathbf{p}^{'})$. Restando tenemos:

$$\mathbf{p} - \mathbf{p}^{'} = \mathbf{G}(\mathbf{p}) - \mathbf{G}(\mathbf{p}^{'})$$

de donde

$$||\mathbf{p} - \mathbf{p}^{'}|| = ||\mathbf{G}(\mathbf{p}) - \mathbf{G}(\mathbf{p}^{'})|| \leq q\ ||\mathbf{p} - \mathbf{p}^{'}||$$

ó

$$(1 - q)\ ||\mathbf{p} - \mathbf{p}^{'}|| \leq 0 \ .$$

Como $(1-q) > 0$, la desigualdad obtenida puede mantenerse únicamente si $||\mathbf{p}-\mathbf{p}^{'}|| = 0$, esto es cuando $\mathbf{p} = \mathbf{p}^{'}$. Por lo tanto no puede haber otra solución de la ecuación $\mathbf{x} = \mathbf{G}(\mathbf{x})$ en el dominio D.

(c) Pasando al límite en la desigualdad

$$||\mathbf{x}^{(k+p)} - \mathbf{x}^{(k)}|| \leq \frac{q^k}{1-q}\ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}||$$

para $p \to \infty$ obtenemos la estimación $(XXI.7)$. Esto completa la demostración del Teorema. **c.q.d.**

En el capítulo XIII hemos definido las normas l_1, l_2 y l_∞ para el vector $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ como

$$||\mathbf{x}||_1 = \sum_{i=1}^{n} |x_i| \ , \qquad ||\mathbf{x}||_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \quad \text{y} \quad ||\mathbf{x}||_\infty = \max_{1\leq i\leq n} |x_i| \ .$$

Consecuentemente, las normas matriciales asociadas tienen las formas

$$||A||_1 = \max_{||\mathbf{x}||_1=1} ||A\mathbf{x}||_1 = \max_{1\leq j\leq n} \sum_{i=1}^{n} |a_{ij}| \ , \qquad \text{norma } \ l_1 \ ,$$

$$||A||_2 = \max_{||\mathbf{x}||_2=1} ||A\mathbf{x}||_2 \ , \qquad \text{norma } \ l_2 \ ,$$

y

$$||A||_\infty = \max_{||\mathbf{x}||_\infty=1} ||A\mathbf{x}||_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{ij}| \ , \qquad \text{norma } \ l_\infty \ .$$

Además con respecto al dominio D definiremos las normas:

$$||\mathbf{G}'(\mathbf{x})||_I = \max_{\mathbf{x}\in D} ||\mathbf{G}'(\mathbf{x})||_1 \qquad \text{y} \qquad ||\mathbf{G}'(\mathbf{x})||_{II} = \max_{\mathbf{x}\in D} ||\mathbf{G}'(\mathbf{x})||_\infty \ ,$$

donde, obviamente,

$$||\mathbf{G}'(\mathbf{x})||_1 = \max_{1\le j\le n} \sum_{i=1}^{n} \Big|\frac{\partial g_i(\mathbf{x})}{\partial x_j}\Big| \qquad \text{y} \qquad ||\mathbf{G}'(\mathbf{x})||_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} \Big|\frac{\partial g_i(\mathbf{x})}{\partial x_j}\Big| \ .$$

Lema XXI.3

Si $\mathbf{F}(\mathbf{x}) = (f_{ij}(\mathbf{x}))$ es una función matricial $n\times r$, donde las $f_{ij}(\mathbf{x})$ son continuas juntamente con sus derivadas parciales de primer orden en un dominio convexo que contenga los punto $\mathbf{x}$ y $\mathbf{x}+\mathbf{\Delta x}$, entonces,

$$||\mathbf{F}(\mathbf{x}+\mathbf{\Delta x}) - \mathbf{F}(\mathbf{x})||_\infty \le r\ ||\Delta x||_\infty\ ||\mathbf{F}'(\xi)||_\infty$$

donde $\xi = \mathbf{x} + \theta\ \mathbf{\Delta x}$, $0<\theta<1$.

Teorema XXI.4

Sean las funciones $\mathbf{G}(\mathbf{x})$ y $\mathbf{G}'(\mathbf{x})$ continuas en el dominio D, cumpliéndose en D la desigualdad

$$||\mathbf{G}'(\mathbf{x})||_{II} \le q < 1\ , \qquad (XXI.8)$$

donde q es una constante. Si las aproximaciones sucesivas

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}) \qquad (XXI.5)$$

$(k=0,1,2,\ldots)$ están contenidas en D, el proceso de iteración $(XXI.5)$ converge y el vector límite

$$\mathbf{p} = \lim_{k\to\infty} \mathbf{x}^{(k)}$$

es la única solución del sistema $\mathbf{y}=\mathbf{G}(\mathbf{x})$ en el dominio D.

Demostración: en virtud del Teorema XXI.2, es suficiente demostrar que la aplicación $\mathbf{y}=\mathbf{G}(\mathbf{x})$ es, dada la condición $(XXI.8)$, una aplicación sobreyectiva en D para la norma l_∞.

Sean $\mathbf{x}_1,\mathbf{x}_2\in D$ e $\mathbf{y}_i=\mathbf{G}(\mathbf{x}_i)$ $(i=1,2)$. Por el Lema XXI.3 tenemos:

$$\begin{aligned}||\mathbf{y}_1-\mathbf{y}_2||_\infty &= ||\mathbf{G}(\mathbf{x}_1)-\mathbf{G}(\mathbf{x}_2)||_\infty \le \\ &\le ||\mathbf{x}_1-\mathbf{x}_2||_\infty\ ||\mathbf{G}'(\xi)||_\infty \le ||\mathbf{x}_1-\mathbf{x}_2||_\infty\ ||\mathbf{G}'(\mathbf{x})||_{II}\ .\end{aligned}$$

De donde

$$||\mathbf{y}_1-\mathbf{y}_2||_\infty \le q\ ||\mathbf{x}_1-\mathbf{x}_2||_\infty$$

con $0\le q<1$, lo que completa la demostración. **c.q.d.**

Corolario XXI.5

El proceso de iteración

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}) \qquad (XXI.5)$$

$(k = 0, 1, 2, \ldots)$ converge si

$$\sum_{j=1}^{n} \Big| \frac{\partial g_i(\mathbf{x})}{\partial x_j} \Big| \leq q_i < 1 \qquad (XXI.9)$$

$(i = 1, 2, \ldots, n)$ cuando $\mathbf{x} \in D$.

Demostración: Evidentemente, del sistema de desigualdades $(XXI.9)$ sigue la condición $(XXI.8)$ del Teorema XXI.4. **c.q.d.**

Notése que en razón del Teorema XXI.2 obtenemos el siguiente estimado para la aproximación $\mathbf{x}^{(k)}$:

$$||\mathbf{p} - \mathbf{x}^{(k)}||_\infty \leq \frac{q^k}{1-q} \; ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}||_\infty \ ,$$

$(k = 0, 1, 2, \ldots)$.

Lema XXI.6

Si la función vectorial $\mathbf{F}(\mathbf{x}) = (f_i(\mathbf{x}))$ es continua juntamente con su derivada $\mathbf{F}'(\mathbf{x})$, en un dominio convexo que contenga los puntos $\mathbf{x}$ y $\mathbf{x} + \mathbf{\Delta x}$, entonces,

$$||\mathbf{F}(\mathbf{x} + \mathbf{\Delta x}) - \mathbf{F}(\mathbf{x})||_1 \leq ||\Delta x||_1 \; ||\mathbf{F}'(\xi)||_1$$

donde $\xi = \mathbf{x} + \theta \; \mathbf{\Delta x}$, $0 < \theta < 1$.

Teorema XXI.7

Sean las funciones $\mathbf{G}(\mathbf{x})$ y $\mathbf{G}'(\mathbf{x})$ continuas en el dominio D cerrado, convexo y acotado, cumpliéndose en D la desigualdad

$$||\mathbf{G}'(\mathbf{x})||_I \leq q < 1 \ , \qquad (XXI.10)$$

donde q es una constante. Si $\mathbf{x}^{(0)} \in D$ y todas las aproximaciones sucesivas

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}) \qquad (XXI.5)$$

$(k = 0, 1, 2, \ldots)$ pertenecen a D, el proceso de iteración $(XXI.5)$ converge y el vector límite

$$\mathbf{p} = \lim_{k \to \infty} \mathbf{x}^{(k)}$$

es la única solución del sistema $\mathbf{y} = \mathbf{G}(\mathbf{x})$ en el dominio D.

Demostración: demostraremos que $\mathbf{y} = \mathbf{G}(\mathbf{x})$ es aplicación sobreyectiva en D para la norma l_1.

Supongamos $\mathbf{x}_1, \mathbf{x}_2 \in D$ y $\mathbf{y}_i = \mathbf{G}(\mathbf{x}_i)$ $(i = 1, 2)$. Por el Lema XXI.6 tenemos:

$$||\mathbf{y}_1 - \mathbf{y}_2||_1 = ||\mathbf{G}(\mathbf{x}_1) - \mathbf{G}(\mathbf{x}_2)||_1 \leq ||\mathbf{x}_1 - \mathbf{x}_2||_1 \; ||\mathbf{G}'(\xi)||_1$$

donde $\xi \in D$. Como

$$||\mathbf{G}'(\xi)||_1 \leq \max_{\mathbf{x} \in D} ||\mathbf{G}'(\mathbf{x})||_1 = ||\mathbf{G}'(\mathbf{x})||_{II} \leq q$$

se deduce que

$$||\mathbf{y}_1 - \mathbf{y}_2||_1 \leq q \ ||\mathbf{x}_1 - \mathbf{x}_2||_1$$

donde $0 \leq q < 1$, lo que completa la demostración. **c.q.d.**

Corolario XXI.8

El proceso de iteración

$$\mathbf{x}^{(k+1)} = \mathbf{G}(\mathbf{x}^{(k)}) \qquad (XXI.5)$$

$(k = 0, 1, 2, \ldots)$ converge si

$$\sum_{j=1}^{n} \Big| \frac{\partial g_j(\mathbf{x})}{\partial x_i} \Big| \leq q_i < 1 \qquad (XXI.11)$$

$(i = 1, 2, \ldots, n)$ cuando $\mathbf{x} \in D$.

Demostración: evidentemente, del sistema de desigualdades $(XXI.11)$ se sigue la condición $(XXI.10)$ del Teorema XXI.7. **c.q.d.**

Nótese que por el Teorema XXI.2 obtenemos la siguiente estimación para la aproximación $\mathbf{x}^{(k)}$:

$$||\mathbf{p} - \mathbf{x}^{(k)}||_1 \leq \frac{q^k}{1-q} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}||_1 \ ,$$

$(k = 0, 1, 2, \ldots)$.

El siguiente Teorema generaliza los Teoremas de punto fijo VII.1 y VII.2 al caso $n-$dimensional. Este Teorema es un caso especial del bien conocido **Teorema de la Aplicación Contractiva**.

Teorema XXI.9

Sea $D = \{(x_1, x_2, \ldots, x_n)^t \mid a_i \leq x_i \leq b_i \ \text{ para cada } \ i = 1, 2, \ldots, n\}$ para alguna colección de constantes $a_1, a_2, \ldots, a_n$ y $b_1, b_2, \ldots, b_n$. Supongamos que $\mathbf{G}$ es una función continua con primeras derivadas parciales continuas de $D \subset \mathcal{R}^n$ a $\mathcal{R}^n$ con la propiedad de que $\mathbf{G}(\mathbf{x}) \in D$ para cada $\mathbf{x} \in D$. Entonces, $\mathbf{G}$ tiene un punto fijo en D. Además, supóngase que existe una constante $K < 1$ con

$$\Big| \frac{\partial g_i(\mathbf{x})}{\partial x_j} \Big| \leq \frac{K}{n} \qquad \text{siempre que } \ \mathbf{x} \in D \ ,$$

para cada $j = 1, 2, \ldots, n$ y cada función componente g_i. Entonces la sucesión $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$ definida por un $\mathbf{x}_0$ en D seleccionada arbitrariamente y generada por

$$\mathbf{x}^{(k)} = \mathbf{G}(\mathbf{x}^{(k-1)}) \qquad \text{para cada } \ k \geq 1$$

converge al punto fijo único $\mathbf{p} \in D$ y

$$||\mathbf{x}^{(k)} - \mathbf{p}||_\infty \leq \frac{K^k}{1-K} \ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}||_\infty \ . \qquad (XXI.12)$$

Ejemplo. Considérese el sistema no lineal dado por

$$\begin{aligned} 3\ x_1 - \cos(x_2 x_3) - \frac{1}{2} &= 0\ , \\ x_1^2 - 81\ (x_2 + 0.1)^2 + \sin x_3 + 1.06 &= 0\ , \\ e^{-x_1 x_2} + 20\ x_3 + \frac{10\pi - 3}{3} &= 0\ . \end{aligned}$$

Si de la $i-$ésima ecuación se despeja x_i, el sistema puede cambiarse a un problema de punto fijo

$$\begin{aligned} x_1 &= \frac{1}{3}\ \cos(x_2 x_3) + \frac{1}{6}\ , \\ x_2 &= \frac{1}{9}\ \sqrt{x_1^2 + \sin x_3 + 1.06} - 0.1\ , \\ x_3 &= -\ \frac{1}{20}\ e^{-x_1 x_2} - \frac{10\pi - 3}{60}\ . \end{aligned}$$

Sea $\mathbf{G} : \mathcal{R}^3 \to \mathcal{R}^3$ definida por $\mathbf{G}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x}), g_3(\mathbf{x}))^t$ donde

$$\begin{aligned} g_1(x_1, x_2, x_3) &= \frac{1}{3}\ \cos(x_2 x_3) + \frac{1}{6}\ , \\ g_2(x_1, x_2, x_3) &= \frac{1}{9}\ \sqrt{x_1^2 + \sin x_3 + 1.06} - 0.1\ , \\ g_3(x_1, x_2, x_3) &= -\ \frac{1}{20}\ e^{-x_1 x_2} - \frac{10\pi - 3}{60}\ . \end{aligned}$$

Se usarán los Teoremas XXI.1 y XXI.9 para demostrar que $\mathbf{G}$ tiene un único punto fijo en

$$D = \{(x_1, x_2, x_3)^t \mid -1 \le x_i \le 1 \ \text{ para cada } \ i = 1, 2, 3\}\ .$$

Para $\mathbf{x} = (x_1, x_2, x_3)^t \in D$,

$$\begin{aligned} |g_1(x_1, x_2, x_3)| &\le \frac{1}{3}\ |\cos(x_2 x_3)| + \frac{1}{6} \le 0.50\ , \\ |g_2(x_1, x_2, x_3)| &= |\frac{1}{9}\ \sqrt{x_1^2 + \sin x_3 + 1.06} - 0.1| \le \\ &\le \frac{1}{9}\ \sqrt{1 + \sin 1 + 1.06} - 0.1 < 0.09\ , \\ |g_3(x_1, x_2, x_3)| &= \frac{1}{20}\ e^{-x_1 x_2} + \frac{10\pi - 3}{60} \le \\ &\le \frac{1}{20}\ e + \frac{10\pi - 3}{60} < 0.61\ ; \end{aligned}$$

así que $-1 \le g_i(x_1, x_2, x_3) \le 1$, para cada $i = 1, 2, 3$. Por lo tanto, $\mathbf{G}(\mathbf{x}) \in D$ siempre que $\mathbf{x} \in D$.

Para las cotas de las derivadas parciales en D se obtiene lo siguiente:

$$\left|\frac{\partial g_1}{\partial x_1}\right| = 0\ , \qquad \left|\frac{\partial g_2}{\partial x_2}\right| = 0\ , \qquad \text{y} \qquad \left|\frac{\partial g_3}{\partial x_3}\right| = 0\ ,$$

mientras que

$$\begin{aligned}
\left|\frac{\partial g_1}{\partial x_2}\right| &\leq \frac{1}{3}\ |x_3|\ |\sin x_2 x_3| \leq \frac{1}{3}\ \sin 1 < 0.281\ ,\\
\left|\frac{\partial g_1}{\partial x_3}\right| &\leq \frac{1}{3}\ |x_2|\ |\sin x_2 x_3| \leq \frac{1}{3}\ \sin 1 < 0.281\ ,\\
\left|\frac{\partial g_2}{\partial x_1}\right| &\leq \frac{|x_1|}{9\ \sqrt{x_1^2+\sin x_3+1.06}} \leq \frac{1}{9\ \sqrt{0.218}} < 0.238\ ,\\
\left|\frac{\partial g_2}{\partial x_3}\right| &\leq \frac{|\cos x_3|}{18\ \sqrt{x_1^2+\sin x_3+1.06}} \leq \frac{1}{18\ \sqrt{0.218}} < 0.119\ ,\\
\left|\frac{\partial g_3}{\partial x_1}\right| &\leq \frac{|x_2|}{20}\ e^{-x_1 x_2} \leq \frac{1}{20}\ e < 0.14\ ,\\
\left|\frac{\partial g_3}{\partial x_2}\right| &\leq \frac{|x_1|}{20}\ e^{-x_1 x_2} \leq \frac{1}{20}\ e < 0.14\ .
\end{aligned}$$

Como las derivadas parciales de g_1, g_2 y g_3 están acotadas en D, el Teorema XXI.1 implica que estas funciones son continuas en D. Consecuentemente, $\mathbf{G}$ es continua en D. Además, para todo $\mathbf{x} \in D$

$$\left|\frac{\partial g_i(\mathbf{x})}{\partial x_j}\right| \leq 0.281 \qquad \text{para cada } \ i=1,2,3 \quad \text{y} \quad j=1,2,3\ ,$$

y de aquí la condición de la segunda parte del Teorema XXI.9 se satisface con $K = 0.843 < 1$.

De la misma manera puede demostrarse que $\frac{\partial g_i}{\partial x_j}$ es continua en D para cada $i = 1,2,3$ y $j = 1,2,3$. Consecuentemente, $\mathbf{G}$ tiene un punto fijo único en D y el sistema no lineal tiene una solución en D.

Es necesario hacer notar que, el que $\mathbf{G}$ tenga una solución única en D no implica que la solución al sistema original sea única en este dominio. En nuestro caso, la solución para x_2 implica la elección de la raíz cuadrada principal (positiva).

Para aproximar el punto fijo $\mathbf{p}$ escogemos $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$. La sucesión de vectores generada por

$$\begin{aligned}
x_1^{(k)} &= \frac{1}{3}\ \cos(x_2^{(k-1)} x_3^{(k-1)}) + \frac{1}{6}\ ,\\
x_2^{(k)} &= \frac{1}{9}\ \sqrt{(x_1^{(k-1)})^2 + \sin x_3^{(k-1)} + 1.06} - 0.1\ ,\\
x_3^{(k)} &= -\ \frac{1}{20}\ e^{-x_1^{(k-1)} x_2^{(k-1)}} - \frac{10\pi - 3}{60}\ ,
\end{aligned}$$

convergerá a la solución única del problema dado en forma de punto fijo. En este ejemplo se ha generado la sucesión hasta que k fuera tal que

$$||\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}||_\infty < 10^{-5}\ .$$

Tabla 1

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|\|_\infty$
0	0.100000000	0.100000000	−0.100000000	—
1	0.499983333	0.009441150	−0.523101267	0.4231
2	0.499995935	0.000025568	−0.523363316	9.4×10^{-3}
3	0.500000000	0.000012337	−0.523598136	2.3×10^{-4}
4	0.500000000	0.000000034	−0.523598467	1.2×10^{-5}
5	0.500000000	0.000000016	−0.523598775	3.1×10^{-7}

Usando la cota de error $(XXI.12)$ con $K = 0.843$ da

$$||\mathbf{x}^{(5)} - \mathbf{p}||_\infty \leq \frac{K^5}{1-K}\ ||\mathbf{x}^{(1)} - \mathbf{x}^{(0)}||_\infty \leq \frac{(0.843)^5}{1-0.8431}\ (0.423) < 1.15\ ,$$

lo cual no indica la precisión real de $\mathbf{x}^{(5)}$, debido a la inexacta aproximación inicial. La solución real es

$$\mathbf{p} = \big(0.5, 0, -\frac{\pi}{6}\big)^t \approx (0.5, 0, -0.5235987756)^t\ ,$$

así que el error real es

$$||\mathbf{x}^{(5)} - \mathbf{p}||_\infty \leq 2 \times 10^{-8}\ .$$

Una manera general de acelerar la convergencia de la iteración de punto fijo en la solución de sistemas no lineales es usar las últimas estimaciones $x_1^{(k)}, \ldots, x_{i-1}^{(k)}$, en lugar de $x_1^{(k-1)}, \ldots, x_{i-1}^{(k-1)}$, para calcular $x_i^{(k)}$, como se hizo en el método de Gauss-Seidel para sistemas lineales. Las ecuaciones son entonces:

$$\begin{aligned} x_1^{(k)} &= \frac{1}{3}\ \cos(x_2^{(k-1)} x_3^{(k-1)}) + \frac{1}{6}\ , \\ x_2^{(k)} &= \frac{1}{9}\ \sqrt{(x_1^{(k)})^2 + \sin x_3^{(k-1)} + 1.06} - 0.1\ , \\ x_3^{(k)} &= -\ \frac{1}{20}\ e^{-x_1^{(k)} x_2^{(k)}} - \frac{10\pi - 3}{60}\ . \end{aligned}$$

Con $\mathbf{x}^{(0)} = (0.1, 0.1, -0.1)^t$, los resultados de estos cálculos se muestran en la tabla 2.

Tabla 2

k	$x_1^{(k)}$	$x_2^{(k)}$	$x_3^{(k)}$	$\|\|\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}\|\|_\infty$
0	0.100000000	0.100000000	−0.100000000	—
1	0.499983333	0.022229794	−0.523046126	0.4230
2	0.499977468	0.000028154	−0.523598072	2.2×10^{-2}
3	0.500000000	0.000000038	−0.523598775	2.8×10^{-5}
4	0.500000000	0.000000000	−0.523598776	3.8×10^{-8}

La iteración $\mathbf{x}^{(4)}$ es exacta dentro de 10^{-7} en la norma l_∞; usando el método de Seidel se ha acelerado realmente la convergencia para este problema. Se debe observar, sin embargo, que el método de Seidel no *siempre* acelera la convergencia.

CAPITULO XXII. METODO DE NEWTON

1. INTRODUCCION Y METODO

Aún cuando el problema presentado en el ejemplo del Capítulo XXI puede transformarse fácilmente en un formato convergente de punto fijo, esto no sucede frecuentemente. En este capítulo consideraremos un procedimiento algorítmico que puede usarse para realizar la transformación para un problema general.

Así como en el Capítulo IX hemos introducido el método de Newton-Raphson para la resolución del problema de la búsqueda de raíces de la ecuación $f(x) = 0$ usando un enfoque intuitivo basado en el polinomio de Taylor, aquí también usaremos este procedimiento para introducir el método.

La forma general de un sistema de ecuaciones no lineales es:

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= 0 \ , \\ f_2(x_1, x_2, \ldots, x_n) &= 0 \ , \\ \ldots \quad \ldots & \quad \ldots \\ f_n(x_1, x_2, \ldots, x_n) &= 0 \ . \end{aligned} \qquad (XXII.1)$$

Usando notación vectorial, el sistema $(XXII.1)$ asume la forma:

$$\mathbf{F}(\mathbf{x}) = \mathbf{0} \ . \qquad (XXII.2)$$

Resolveremos el sistema $(XXII.2)$ por el método de aproximaciones sucesivas. Supongamos que hemos hallado la aproximación k-ésima, $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)})^t$, de una de las raíces separadas $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ de la ecuación vectorial $(XXII.2)$. La raíz exacta de $(XXII.2)$ puede representarse entonces como

$$\mathbf{x} = \mathbf{x}^{(k)} + \varepsilon^{(k)} \ , \qquad (XXII.3)$$

donde $\varepsilon^{(k)} = (\varepsilon_1^{(k)}, \varepsilon_2^{(k)}, \ldots, \varepsilon_n^{(k)})^t$ es la corrección (**error de la raíz**).

Sustituyendo $(XXII.3)$ en $(XXII.2)$, tenemos

$$\mathbf{F}(\mathbf{x}^{(k)} + \varepsilon^{(k)}) = \mathbf{0} \ . \qquad (XXII.4)$$

Suponiendo que la función $\mathbf{F}(\mathbf{x})$ es continuamente diferenciable en un cierto dominio convexo que contiene a $\mathbf{x}$ y $\mathbf{x}^{(k)}$, desarrollemos el primer miembro de la ecuación $(XXII.4)$ en potencias del pequeño vector $\varepsilon^{(k)}$ limitándonos a los términos lineales,

$$\mathbf{F}(\mathbf{x}^{(k)} + \varepsilon^{(k)}) = \mathbf{F}(\mathbf{x}^{(k)}) + \mathbf{F}'(\mathbf{x}^{(k)}) \ \varepsilon^{(k)} = \mathbf{0} \ . \qquad (XXII.5)$$

De la fórmula $(XXII.5)$ se deduce que la derivada $\mathbf{F}'(\mathbf{x})$ ha de ser considerada como la **matriz Jacobiana**, $J(\mathbf{x})$, del conjunto de funciones $f_1, f_2, \ldots, f_n$ con respecto a las variables $x_1, x_2, \ldots, x_n$; esto es,

$$J(\mathbf{x}) = \mathbf{F}'(\mathbf{x}) = \Big(\frac{\partial f_i}{\partial x_j}\Big) = \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \ldots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \ldots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial f_n(\mathbf{x})}{\partial x_1} & \frac{\partial f_n(\mathbf{x})}{\partial x_2} & \ldots & \frac{\partial f_n(\mathbf{x})}{\partial x_n} \end{pmatrix} \qquad (XXII.6)$$

$(i, j = 1, 2, \ldots, n)$.

El sistema $(XXII.5)$ es un sistema lineal en las correcciones $\varepsilon_i^{(k)}$ $(i = 1, 2, \ldots, n)$ con la matriz $J(\mathbf{x}^{(k)})$, y de aquí que la fórmula $(XXII.5)$ pueda escribirse como:

$$\mathbf{F}(\mathbf{x}^{(k)}) + J(\mathbf{x}^{(k)})\ \varepsilon^{(k)} = \mathbf{0}\ ,$$

de donde, dando por supuesto que la matriz $J(\mathbf{x}^{(k)})$ es no singular, tenemos

$$\varepsilon^{(k)} = -J^{-1}(\mathbf{x}^{(k)})\ \mathbf{F}(\mathbf{x}^{(k)})\ .$$

En consecuencia

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} - J^{-1}(\mathbf{x}^{(k)})\ \mathbf{F}(\mathbf{x}^{(k)})\ , \qquad (XXII.7)$$

$(k = 0, 1, 2, \ldots)$. Para la aproximación de orden cero $\mathbf{x}^{(0)}$ podemos tomar un valor aproximado de la raíz deseada.

Además del enfoque intuitivo basado en el polinomio de Taylor, para construir el algoritmo que nos llevó a un método apropiado de punto fijo en el caso unidimensional, en el Capítulo X, tratamos de encontrar una función ϕ con la propiedad de que

$$g(x) = x - \phi(x)\ f(x)\ , \qquad (XXII.8)$$

dando convergencia cuadrática al punto fijo p de g. De esta condición surgió el método de Newton, escogiendo $\phi(x) = 1/f'(x)$.

Usando un enfoque similar para el caso $n-$dimensional, es necesaria una matriz

$$A(\mathbf{x}) = \begin{pmatrix} a_{11}(\mathbf{x}) & a_{12}(\mathbf{x}) & \ldots & a_{1n}(\mathbf{x}) \\ a_{21}(\mathbf{x}) & a_{22}(\mathbf{x}) & \ldots & a_{2n}(\mathbf{x}) \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1}(\mathbf{x}) & a_{n2}(\mathbf{x}) & \ldots & a_{nn}(\mathbf{x}) \end{pmatrix} \qquad (XXII.9)$$

donde cada una de las componentes $a_{ij}(\mathbf{x})$ sea una función de $\mathcal{R}^n$ a $\mathcal{R}$. El procedimiento requiere encontrar una matriz $A(\mathbf{x})$ tal que

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} - A(\mathbf{x})^{-1}\ \mathbf{F}(\mathbf{x})\ , \qquad (XXII.10)$$

dé convergencia cuadrática a la solución de $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, siempre que, desde luego, $A(\mathbf{x})$ sea no singular en el punto fijo de $\mathbf{G}$. El siguiente Teorema verifica que este enfoque puede usarse para justificar la elección de A.

Teorema XXII.1

Supóngase que $\mathbf{p}$ es una solución de $\mathbf{G}(\mathbf{x}) = \mathbf{x}$, para alguna función $\mathbf{G} = (g_1, g_2, \ldots, g_n)^t$, que manda a $\mathcal{R}^n$ en $\mathcal{R}^n$. Si existe un número $\delta > 0$ con la propiedad de que

i) $\partial g_i/\partial x_j$ es continua en $N_\delta = \{\mathbf{x} \mid ||\mathbf{x} - \mathbf{p}|| < \delta\}$ para cada $i = 1, 2, \ldots, n$ y $j = 1, 2, \ldots, n$,

ii) $\partial^2 g_i(\mathbf{x})/(\partial x_j \partial x_k)$ es continua y $|\partial^2 g_i(\mathbf{x})/(\partial x_j \partial x_k)| \leq M$ para alguna constante M, siempre que $\mathbf{x} \in N_\delta$ para cada $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$ y $k = 1, 2, \ldots, n$,

iii) $\partial g_i(\mathbf{p})/\partial x_j = 0$ para cada $i = 1, 2, \ldots, n$ y $j = 1, 2, \ldots, n$.

Entonces existe un número $\tilde{\delta} \leq \delta$ tal que la sucesión generada por $\mathbf{x}^{(k)} = \mathbf{G}(\mathbf{x}^{(k-1)})$ converge cuadráticamente a $\mathbf{p}$ para cualquier $\mathbf{x}^{(0)}$ siempre que $||\mathbf{x}^{(0)} - \mathbf{p}|| < \tilde{\delta}$. Además,

$$||\mathbf{x}^{(k)} - \mathbf{p}||_\infty \leq \frac{n^2 M}{2}\ ||\mathbf{x}^{(k-1)} - \mathbf{p}||_\infty^2 \qquad \text{para cada} \ \ k \geq 1 \ .$$

Este Teorema es extensión del Teorema X.1 y su demostración requiere que se pueda expresar $\mathbf{G}$ en términos de su serie de Taylor en n variables alrededor del punto $\mathbf{p}$.

Para utilizar el Teorema XXII.1, supongamos que $A(\mathbf{x})$ es una matriz $n \times n$ de funciones de $\mathcal{R}^n$ a $\mathcal{R}$ en la forma de la ecuación $(XXII.8)$, donde las componentes específicas se escogerán más adelante. Supongamos además, que $A(\mathbf{x})$ es no singular cerca de una solución $\mathbf{p}$ de $\mathbf{F}(\mathbf{x}) = \mathbf{0}$ y denotemos por $b_{ij}(\mathbf{x})$ a la componente de $A(\mathbf{x})^{-1}$ en el $i-$ésima fila y en la $j-$ésima columna. Como $\mathbf{G}(\mathbf{x}) = \mathbf{x} - A(\mathbf{x})^{-1}\ \mathbf{F}(\mathbf{x})$,

$$g_i(\mathbf{x}) = x_i - \sum_{j=1}^{n} b_{ij}(\mathbf{x})\ f_j(\mathbf{x}) \ ;$$

así que

$$\frac{\partial g_i(\mathbf{x})}{\partial x_k} = \begin{cases} 1 - \sum\limits_{j=1}^{n} b_{ij}(\mathbf{x})\ \frac{\partial f_j(\mathbf{x})}{\partial x_k} + \frac{\partial b_{ij}(\mathbf{x})}{\partial x_k}\ f_j(\mathbf{x}) \ , & \text{si } k = i \ ; \\ -\sum\limits_{j=1}^{n} b_{ij}(\mathbf{x})\ \frac{\partial f_j(\mathbf{x})}{\partial x_k} + \frac{\partial b_{ij}(\mathbf{x})}{\partial x_k}\ f_j(\mathbf{x}) \ , & \text{si } k \neq i \ . \end{cases}$$

El Teorema XXII.1 implica que necesitamos tener $\partial g_i(\mathbf{p})/\partial x_k = 0$ para cada $i = 1, 2, \ldots, n$ y $k = 1, 2, \ldots, n$. Esto significa que para $k = i$,

$$0 = 1 - \sum_{j=1}^{n} b_{ij}(\mathbf{p})\ \frac{\partial f_j(\mathbf{p})}{\partial x_i} \ ,$$

con lo que

$$\sum_{j=1}^{n} b_{ij}(\mathbf{p})\ \frac{\partial f_j(\mathbf{p})}{\partial x_i} = 1 \ , \qquad (XXII.11)$$

y cuando $k \neq i$,

$$0 = -\sum_{j=1}^{n} b_{ij}(\mathbf{p})\ \frac{\partial f_j(\mathbf{p})}{\partial x_k} \ ,$$

con lo que

$$\sum_{j=1}^{n} b_{ij}(\mathbf{p}) \; \frac{\partial f_j(\mathbf{p})}{\partial x_k} = 0 \; . \qquad (XXII.12)$$

Usando la matriz Jacobiana $J(\mathbf{x})$, vemos que las condiciones $(XXII.11)$ y $(XXII.12)$ requieren que

$$A(\mathbf{p})^{-1} \; J(\mathbf{p}) = I \; ,$$

y por lo tanto

$$A(\mathbf{p}) = J(\mathbf{p}) \; .$$

Una elección apropiada para $A(\mathbf{x})$ es consecuentemente $A(\mathbf{x}) = J(\mathbf{x})$, ya que la condición (iii) del Teorema XXII.1 se satisface con esta elección.

La función $\mathbf{G}$ se define como

$$\mathbf{G}(\mathbf{x}) = \mathbf{x} - J(\mathbf{x})^{-1} \; \mathbf{F}(\mathbf{x}) \; ,$$

y el procedimiento de iteración funcional surge de seleccionar $\mathbf{x}^{(0)}$ y de generar, para $k \geq 1$,

$$\mathbf{x}^{(k)} = \mathbf{G}(\mathbf{x}^{(k-1)}) = \mathbf{x}^{(k-1)} - J(\mathbf{x}^{(k-1)})^{-1} \; \mathbf{F}(\mathbf{x}^{(k-1)}) \; . \qquad (XXII.13)$$

Este método se llama **método de Newton para sistemas no lineales** y se espera que generalmente dé convergencia cuadrática siempre y cuando se conozca un valor inicial lo suficientemente exacto y que $J(\mathbf{p})^{-1}$ exista.

2. ALGORITMO Y EJEMPLOS

Una debilidad clara del procedimiento del método de Newton surge de la necesidad de invertir la matriz $J(\mathbf{x})$ en cada paso. En la práctica, el método se realiza generalmente en una forma de dos pasos. Primero, se encuentra un vector $\mathbf{y}$ que satisfaga

$$J(\mathbf{x}^{(k)}) \; \mathbf{y} = -\mathbf{F}(\mathbf{x}^{(k)}) \; .$$

Después de que se ha logrado esto, la nueva aproximación $\mathbf{x}^{(k+1)}$ se puede obtener sumando $\mathbf{y}$ a $\mathbf{x}^{(k)}$; es decir,

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{y} \; .$$

El siguiente algoritmo usa este procedimiento de dos pasos.

Algoritmo del método de Newton para sistemas no lineales.

==

Para aproximar una solución del sistema no lineal $\mathbf{F}(\mathbf{x}) = \mathbf{0}$, dada una aproximación inicial $\mathbf{x}$:

Entrada: número n de ecuaciones e incógnitas; aproximación inicial $\mathbf{x} = (x_1, \, x_2, \, \ldots, \, x_n)^t$; tolerancia TOL; número máximo de iteraciones N_0;

Salida: solución aproximada $\mathbf{x} = (x_1, \, x_2, \, \ldots, \, x_n)^t$ o mensaje de que el número de iteraciones fue excedido.

Paso 1: tomar $k = 1$;

Paso 2: mientras que $k \leq N_0$ seguir pasos 3–7;

Paso 3: calcular $\mathbf{F}(\mathbf{x})$ y $J(\mathbf{x})$, donde $(J(\mathbf{x}))_{ij} = (\partial f_i(\mathbf{x})/\partial x_j)$ para $1 \leq i, j \leq n$;

Paso 4: resolver el sistema lineal de $J(\mathbf{x})\ \mathbf{y} = -\mathbf{F}(\mathbf{x})$;

Paso 5: tomar $\mathbf{x} = \mathbf{x} + \mathbf{y}$;

Paso 6: si $||\mathbf{y}|| < TOL$ entonces SALIDA ($\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$); *(procedimiento completado satisfactoriamente)* PARAR;

Paso 7: tomar $k = k + 1$;

Paso 8: SALIDA ($'$*Número máximo* N_0 *de iteraciones excedido,* $N_0 =\ ', N_0$); PARAR.

==

Nótese que la convergencia del método de Newton se hace muy rápida una vez que estamos cerca de la solución $\mathbf{p}$. Esto ilustra la convergencia cuadrática del método cerca de la solución.

CURSO DE METODOS NUMERICOS

SEXTA PARTE

BIBLIOGRAFIA

BIBLIOGRAFIA BASICA

- R.L. Burden & J.D. Faires: *Análisis Numérico*, Grupo Editorial Iberoamérica, 1985; traducción de *Numerical Analysis*, Prindle, Weber and Schmidt, 1985.
- D. Kincaid & W. Cheney: *Análisis Numérico: las matemáticas del cálculo científico*, Addison-Wesley, 1994; traducción de *Numerical Mathematics and Computing*, Brooks/Cole Publ. Comp., 1991.
- B.P. Demidovich & I.A. Maron: *Cálculo Numérico Fundamental*, Paraninfo, 1988.
- J. Stoer & R. Bulirsch: *Introduction to Numerical Analysis*, Springer-Verlag, 1980.
- K.E. Atkinson: *An Introduction to Numerical Analysis*, John Wiley & Sons, 1978.
- P. Henrici: *Elementos de Análisis Numérico*, Editorial Trillas, 1972; traducción de *Elements of Numerical Analysis*, John Wiley & Sons, 1964.
- A. Ralston: *Introducción al Análisis Numérico*, Editorial Limusa-Wiles, 1970; traducción de *A first course in Numerical Analysis*, McGraw-Hill, 1965.

BIBLIOGRAFIA COMPLEMENTARIA

- A. Aubanell, A. Benseny & A. Delshams: *Útiles básicos de cálculo numérico*, Editorial Labor, 1993.
- A. Andronico, U. Bozzo, A. De Giorgio, I. Galligani, R. Marconi, E. Nervegna, G. Provenzano, M. Refice: *Principi di informatica*, Zanichelli, 1973.
- H.C. Bezner: *FORTRAN 77*, Prentice-Hall, 1989.
- P.M. Birns, P.B. Brown & J.C.C. Muster: *UNIX for People*, Prentice-Hall, 1985.
- E.K. Blum: *Numerical Analysis and Computation: Theory and Practise*, Addison-Wesley, 1972.
- M.P. Cherkasova: *Problems on Numerical Methods*, Wolters-Noordhoff Publ., 1972.
- C. Conde Lazaro & G. Winter Althaus: *Métodos y algoritmos básicos del álgebra numérica*, Editorial Reverté, 1990.
- G.B. Davis & T.R. Hoffmann: *FORTRAN: a structured disciplined style*, McGraw-Hill, 1978.
- G. Dahlquist & Å. Björck: *Numerical Methods*, Prentice-Hall, 1974.
- J.E. Dennis & R.B. Schnabel: *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall, 1983.
- D.K. Faddeev & V.M. Faddeeva: *Computational Methods of Linear Algebra*, W.H. Freeman & Comp., 1963.
- A.I. Forsythe, T.A. Keenan, E.I. Organick & W. Stenberg: *Lenguajes de diagramas de flujo*, Editorial Limusa-Wiley, 1973; traducción de *Computer Science: a first course*, John Wiley & Sons, 1969.
- G.E. Forsythe & C.B. Moler: *Computer solution of Linear Algebraic Systems*, Prentice-Hall, 1967.

- L. Fox: *An Introduction to Numerical Linear Algebra*, Clarendon Press, 1964.
- L. Fox & D.F. Mayers: *Computing methods for Scientists and Engineers*, Clarendon Press, 1968.
- C.E. Fröberg: *Introduction to Numerical Analysis*, Addison-Wesley, 1970.
- F. Garcia Merayo & A. Nevot Luna: *Análisis Numérico*, Editorial Paraninfo, 1992.
- C.F. Gerald & P.O. Wheatley: *Applied Numerical Analysis*, Addison-Wesley, 1984.
- P.E. Gill, W. Murray & M.H. Wright: *Numerical Linear Algebra and Optimization*, Addison-Wesley, 1991.
- H.H. Goldstine: *A Hystory of Numerical Analysis*, Springer-Verlag, 1977.
- G.H. Golub & C.F. Van Loan: *Matrix Computations*, North Oxford Academic, 1983.
- A.R. Gourlay & G.A. Watson: *Computational Methods for Matrix Eigenproblems*, John Wiley & Sons, 1973.
- W. Hackbusch: *Iterative Solution of Large Sparse systems of equations*, Springer-Verlag, Inc., 1994.
- V. Hernández, E. Ramos, R. Vélez & I. Yáñez: *Matemáticas Básicas*, Universidad Nacional de Educación a Distancia, 1992.
- J.D. Hoffman: *Numerical Methods for Engineers and Scientists*, McGraw-Hill, 1992.
- A.S. Householder: *The theory of Matrices in Numerical Analysis*, Blaisdell, 1965.
- A.S. Householder: *The Numerical Treatment of a single Nonlinear Equation*, McGraw-Hill, 1970.
- E. Isaacson & H.B. Keller: *Analysis of Numerical Methods*, John Wiley & Sons, 1966.
- J.A. Jacquez: *A First Course in Computing and Numerical Methods*, Addison-Wesley, 1970.
- I. Jacques & C. Judd: *Numerical Analysis*, Chapman and Hall Ltd., 1987.
- L.W. Johnson & R.D. Riess: *Numerical Analysis*, Addison-Wesley, 1982.
- E.B. Koffman & F.L. Friedman: *Problem solving and structured programming in Fortran 77*, Addison-Wesley, 1990.
- C. Lawson & R. Hanson: *Solving Least Squares Problems*, Prentice-Hall, 1974.
- S.J. Leon: *Linear Algebra with applications*, MacMillan, 1986.
- P. Lévy: *La invención del ordenador*, en *Historia de la Ciencias*, dirigida por M. Serres, Cátedra, 1991.
- M. El Lozy: *Editing in a Unix environment: the vi/ex Editor*, Prentice-Hall, 1985.
- M. Loukides: *UNIX for FORTRAN programmers*, O'Reilly & Associates, 1990.
- J.H. Mathews: *Numerical Methods for Computer Science, Engineering and Mathematics*, Prentice-Hall, 1987.
- J.M. McCormick & M.G. Salvadori: *Numerical Methods in Fortran*, Prentice-Hall, 1964.

- S.S. McNeary: *Introduction to computational methods for students of calculus*, Prentice-Hall, 1973.
- F. Michavila & L. Gavete: *Programación y cálculo numérico*, Edit. Reverté, 1992.
- D.G. Moursund & C.S. Duris: *Elementary Theory and Application of Numerical Analysis*, Dover Pubb., Inc., 1967.
- J.C. Nash: *Compact Numerical Methods for computers: linear algebra and function minimisation*, Adam Hilger, 1990.
- B. Noble & J.W. Daniel: *Applied Linear Algebra*, Prentice-Hall, 1977.
- T.R.F. Nonweiler: *Computational Mathematics: an introduction to numerical aproximations*, Ellis Horwood Lim. (John Wiley & Sons), 1984.
- J.P. Nougier: *Méthodes de calcul numérique*, Masson, 1993.
- J.M. Ortega: *Numerical Analysis*, Academic Press, 1972.
- J.M. Ortega & W.C. Rheinboldt: *Iterative solutions of Nonlinear Equations in several variables*, Academic Press, 1970.
- G.M. Phillips & P.J. Taylor: *Theory and Applications of Numerical Analysis*, Academic Press, 1973.
- G.W. Stewart: *Introduction to Matrix Computations*, Academic Press, 1973.
- E. Stiefel: *Introducción a la Matemática Numérica*, Editorial Labor, 1966.
- G. Strang: *Introduction to Applied Mathematics*, Wellesley-Cambridge Press, 1986.
- A.N. Tijonov & D.P. Kostomárov: *Algo acerca de la Matemática Aplicada*, Editorial MIR, 1979.
- J.C. Vaissiére & J.P. Nougier: *Programmes et exercices sur les méthodes numériques*, Masson, 1991.
- R. Wait: *The Numerical solutions of Algebraic Equations*, John Wiley & Sons, 1979.
- J.H. Wilkinson: *The Algebraic Eigenvalue Problem*, Clarendon Press, 1965.
- J.H. Wilkinson & C. Reinsch (Eds.): *Handbook for Automatic Computations. Vol.2: Linear Algebra*, 1971.
- D.M. Young & R.T. Gregory: *A survey of Numerical Mathematics*, Vol I–II, Addison-Wesley, 1972-1973.