

**Análisis Numérico Matricial**  
**Universidad de Murcia**  
**curso 2019-2020**

*Sistemas de Ecuaciones no lineales*

Antonio José Pallarés Ruiz



# Índice general

<b>1. Introducción y complementos de análisis matricial.</b>	<b>5</b>
1.1. Origen de los problemas del análisis numérico matricial . . . . .	6
1.2. Repaso de álgebra matricial . . . . .	9
1.2.1. Sistemas de Ecuaciones . . . . .	9
1.2.2. Matrices . . . . .	10
1.2.3. Espacios vectoriales. Aplicaciones lineales . . . . .	12
1.2.4. Reducción de matrices . . . . .	18
1.2.5. Cocientes de Rayleigh. Matrices simétricas y hermitianas . . . . .	20
1.3. Normas matriciales . . . . .	22
1.3.1. Convergencia de matrices . . . . .	25
1.4. Análisis del error. Condicionamiento . . . . .	26
1.4.1. Condicionamiento en la búsqueda de valores y vectores propios . . . . .	30
1.5. Complementos del capítulo . . . . .	31
<b>2. Métodos Directos para Ecuaciones Lineales</b>	<b>35</b>
2.1. Sistemas fáciles de resolver . . . . .	36
2.1.1. Sistemas diagonales . . . . .	36
2.1.2. Sistemas triangulares superiores. Método ascendente . . . . .	37
2.1.3. Sistemas triangulares inferiores. Método descendente . . . . .	38
2.2. Factorización LU . . . . .	39
2.2.1. Algoritmos de factorización . . . . .	40
2.2.2. Complejidad de las factorizaciones LU . . . . .	41
2.2.3. Factorizaciones LDU . . . . .	43
2.2.4. Transformaciones de Gauss sin permutar filas . . . . .	44
2.3. Sistemas con matrices especiales . . . . .	46
2.3.1. Matrices con diagonal estrictamente dominante . . . . .	46
2.3.2. Matrices definidas positivas. . . . .	48
2.3.3. Matrices simétricas definidas positivas (SPD). . . . .	50
2.3.4. Matrices tridiagonales . . . . .	52
2.4. Método de Gauss . . . . .	54
2.4.1. Factorización LUP . . . . .	54

2.4.2. Factorización LUPQ (pivote total)	56
2.5. Factorización QR	59
2.5.1. Transformaciones de Householder	59
2.5.2. Factorización QR usando las transformaciones de Householder	62
2.6. Problemas de mínimos cuadrados	64
2.6.1. Modelo General de los problemas de aproximación	64
2.6.2. Aproximación por mínimos cuadrados	66
2.6.3. Sistemas sobredeterminados	69
2.6.4. Aplicaciones y Ejemplos	71
2.7. Actividades complementarias del capítulo	75
<b>3. Métodos iterativos de resolución de sistemas de ecuaciones</b>	<b>77</b>
3.1. Métodos iterativos. Criterios de Convergencia	78
3.1.1. Criterios de Convergencia	79
3.1.2. Construcción de Métodos iterativos	80
3.2. Método de Jacobi	81
3.3. Método de Gauss-Seidel	83
3.3.1. Convergencia de los Métodos de Jacobi y Gauss-Seidel	85
3.4. Método de Relajación	87
3.5. Método del gradiente conjugado	93
3.6. Actividades complementarias del capítulo	101
<b>4. Valores y vectores propios</b>	<b>103</b>
4.1. El problema de aproximar valores y vectores propios.	104
4.2. El método de la potencia	106
4.3. El método de Jacobi	111
4.4. El método QR	123
4.5. Actividades complementarias del capítulo	124
<b>5. Sistemas de Ecuaciones no lineales.</b>	<b>125</b>
5.1. Iteración de punto fijo.	127
5.2. Método de Newton	131
5.3. Método de Broyden	135
5.4. Método del descenso rápido: "Steepest Descent"	139
5.5. Ejercicios	140
<b>Bibliografía</b>	<b>143</b>

# Sistemas de Ecuaciones no lineales.

## Interrogantes centrales del capítulo

- Aproximar soluciones de sistemas de ecuaciones no lineales.
- Métodos numéricos para aproximación de soluciones:
  - Iteración de punto fijo.
  - Método de Newton.
  - Método del descenso rápido.

## Destrezas a adquirir en el capítulo

- Ser capaz de describir los algoritmos correspondientes a los distintos métodos expuestos en este capítulo.
- Implementar en el ordenador los programas de los métodos estudiados.
- Aplicar los programas construidos de forma efectiva en la búsqueda de aproximaciones de soluciones de sistemas de ecuaciones no lineales.

**Desarrollo de los contenidos fundamentales**

En este capítulo abordamos el problema de resolver sistemas de ecuaciones no lineales.

Un sistema de  $m$  ecuaciones no lineales con  $m$  incógnitas

$$\begin{cases} f_1(x_1, x_2, \dots, x_m) == 0 \\ f_2(x_1, x_2, \dots, x_m) == 0 \\ \dots\dots\dots \\ f_m(x_1, x_2, \dots, x_m) == 0 \end{cases}$$

definido por  $m$  funciones  $f_k : \mathbb{R}^m \rightarrow \mathbb{R}$ , se puede representar con la ecuación

$$f(\vec{x}) == 0, \text{ donde } f = (f_1, f_2, \dots, f_m) \text{ es una función } f : \mathbb{R}^m \rightarrow \mathbb{R}^m.$$

Resolver el sistema de ecuaciones se traduce en encontrar los ceros de  $f$ .

A veces, interesa escribir el sistema de ecuaciones en la forma

$$\begin{cases} g_1(x_1, x_2, \dots, x_m) == x_1 \\ g_2(x_1, x_2, \dots, x_m) == x_2 \\ \dots\dots\dots \\ g_m(x_1, x_2, \dots, x_m) == x_n \end{cases}$$

$$g(\vec{x}) == x, \text{ donde } g = (g_1, g_2, \dots, g_m) \text{ es una función } g : \mathbb{R}^m \rightarrow \mathbb{R}^m,$$

cuyas soluciones son los puntos fijos de  $g = (g_1, g_2, \dots, g_m)$ , es decir los vectores  $\vec{x}$  con  $g(\vec{x}) = \vec{x}$ .

Vamos a centrar nuestro estudio en las versiones multidimensionales del método de iteración de punto fijo y del método de Newton.

El método de iteración de punto fijo no requiere de herramientas adicionales para su aplicación, mientras que el de Newton necesita de métodos de resolución de sistemas lineales o de inversión de matrices.

Como alternativa al Método de Newton, se propone el estudio del método de Broyden, uno de los métodos denominados quasi-Newton en el que se renuncia a la convergencia cuadrática a cambio de no tener que evaluar el Jacobiano de una función en cada etapa.

Los ceros de funciones vectoriales se pueden interpretar como puntos donde alcanzan valores mínimos absolutos funciones de variable real, de forma análoga a como introducimos el método del gradiente para resolver sistemas de ecuaciones lineales. La búsqueda de mínimos puede hacerse con el método del gradiente o de descenso rápido, siguiendo las direcciones de máximo descenso en términos del gradiente de la función a minimizar.

Así, se puede acudir al método del descenso rápido que tiene buenas propiedades de convergencia aunque también es lento, para obtener la buena aproximación inicial de la solución que necesita el método de Newton para su convergencia.

## 5.1. Iteración de punto fijo.

La prueba del teorema de punto fijo para funciones contractivas de la asignatura de Cálculo Numérico en 1 variable real, se puede realizar en espacios métricos completos y son aplicables a funciones definidas en subconjuntos cerrados de  $\mathbb{R}^n$ .

### **Teorema (Teorema del punto fijo de Banach)**

Sea  $(M, d)$  un espacio métrico completo y  $g : M \rightarrow M$  una aplicación contractiva, es decir, tal que existe  $0 < c < 1$  de manera que

$$d(g(x), g(y)) \leq c d(x, y) \quad \text{para cada } x, y \in M.$$

Entonces  $g$  tiene un único punto fijo en  $\vec{x} \in M$ ,  $g(\vec{x}) = \vec{x}$ . Además,  $\vec{x}$  es el límite de la sucesión de iteradas  $\vec{x}_{n+1} = g(\vec{x}_n)$ , iniciada en cualquier punto  $\vec{x}_0 \in M$  y se cumple:

$$d(\vec{x}, \vec{x}_n) \leq \frac{c^n}{1-c} d(\vec{x}_1, \vec{x}_0).$$

En  $\mathbb{R}^m$  todas las normas definen distancias equivalentes a la distancia euclídea y con estas distancias  $\mathbb{R}^m$  y todos sus subconjuntos cerrados son espacios métricos completos. Reescribiendo el teorema anterior se tiene

**Corolario 5.1.1** Sea  $\|\cdot\|$  una norma en  $\mathbb{R}^m$ ,  $M \subset \mathbb{R}^m$  un subconjunto cerrado y  $g : M \rightarrow M$  una aplicación contractiva, es decir, tal que existe  $0 < c < 1$  de manera que

$$\|g(\vec{x}) - g(\vec{y})\| \leq c \|\vec{x} - \vec{y}\| \quad \text{para cada } \vec{x}, \vec{y} \in M.$$

Entonces  $g$  tiene un único punto fijo en  $\vec{x} \in M$ ,  $g(\vec{x}) = \vec{x}$ . Además,  $\vec{x}$  es el límite de la sucesión de iteradas  $\vec{x}_{n+1} = g(\vec{x}_n)$ , iniciada en cualquier punto  $\vec{x}_0 \in M$  y se cumple:

$$\|\vec{x} - \vec{x}_n\| \leq \frac{c^n}{1-c} \|\vec{x}_1 - \vec{x}_0\|.$$

Para aplicar el teorema de punto fijo, la cuestión fundamental ahora, es determinar condiciones que permitan comprobar fácilmente la función  $g$  es contractiva con respecto a una determinada norma.

Al igual que en el caso de funciones unidimensionales, también podemos recurrir al siguiente teorema del incremento finito para funciones diferenciables multidimensionales:

**Teorema 5.1.2** Sea  $A$  un abierto convexo en  $\mathbb{R}^m$  y  $g : A \rightarrow \mathbb{R}^m$  una función diferenciable en  $A$ . Sea  $\|\cdot\|$  una norma en  $\mathbb{R}^m$  y denotemos también por  $\|\cdot\|$  a la norma subordinada en el espacio de operadores lineales  $\mathcal{L}(\mathbb{R}^m, \mathbb{R}^m)$  definida por

$$\|T\| = \sup\{\|T(\vec{x})\| : \|\vec{x}\| \leq 1\} = \sup\left\{\frac{\|T(\vec{y})\|}{\|\vec{y}\|} : \vec{y} \neq 0\right\}.$$

Si  $c > 0$  es una constante tal que  $\|dg(\vec{x})\| \leq c$  para cada  $\vec{x} \in A$ , entonces

$$\|g(\vec{x}) - g(\vec{y})\| \leq c \|\vec{x} - \vec{y}\| \quad \text{para cada } \vec{x}, \vec{y} \in A.$$

Si las normas de las diferenciales están acotadas en todos los puntos de  $A$  por una constante  $c < 1$ , y  $M \subset A$  es un conjunto cerrado con  $g(M) \subset M$ , entonces se verifica el teorema de punto fijo anterior.

Localmente, los discos cerrados forman una base de entornos convexos. Así, es posible enunciar el siguiente resultado:

**Proposición 5.1.3** Sea  $g : A \rightarrow \mathbb{R}^m$  una función diferenciable, con derivadas parciales continuas, definida en el abierto  $A \subset \mathbb{R}^m$ , y sea  $\vec{x} \in A$  un punto fijo de  $g$  ( $g(\vec{x}) = \vec{x}$ ) con  $\|dg(\vec{x})\| < 1$ . Entonces existe  $r > 0$  tal que

$$\overline{B(\vec{x}, r)} = \{\vec{y} : \|\vec{x} - \vec{y}\| \leq r\} \subset A, \quad g(\overline{B(\vec{x}, r)}) \subset \overline{B(\vec{x}, r)},$$

y la sucesión de iteradas  $\vec{x}_{n+1} = g(\vec{x}_n)$  converge hacia  $\vec{x}$  para cualquier  $\vec{x}_0 \in \overline{B(\vec{x}, r)}$ .

La proposición se puede aplicar en el siguiente ejemplo:

#### Ejemplo 5.1.4

Sean

$$g_1(x, y) = A \cos x - B \sin y,$$

$$g_2(x, y) = A \sin x + B \cos y$$

las componentes de la función  $g(x, y) = (g_1(x, y), g_2(x, y))$ .  $g$  es una función diferenciable en todo  $(x, y)$ , con derivadas parciales

$$\frac{\partial g_1(x, y)}{\partial x} = A \cos x; \quad \frac{\partial g_1(x, y)}{\partial y} = -B \sin y;$$

$$\frac{\partial g_2(x, y)}{\partial x} = A \sin x; \quad \frac{\partial g_2(x, y)}{\partial y} = -B \cos y;$$

La diferencial de  $g$  actuando sobre un vector  $(h_1, h_2)$  es

$$dg(x, y)(h_1, h_2) = (A \cos(x)h_1 - B \sin(y)h_2, A \sin(x)h_1 + B \cos(y)h_2).$$

Si en  $\mathbb{R}^2$  consideramos la norma euclídea  $\|(x, y)\| = \sqrt{x^2 + y^2}$ , se tiene que

$$\begin{aligned} \|dg(x, y)(h_1, h_2)\|_2 &= \sqrt{A^2 h_1^2 + B^2 h_2^2 - 2AB \sin(y-x)h_1 h_2} \\ &\leq \sqrt{A^2 h_1^2 + B^2 h_2^2 + 2AB|h_1||h_2|} = |A||h_1| + |B||h_2| \\ &\leq \sqrt{A^2 + B^2} \|(h_1, h_2)\|_2 \end{aligned}$$

De donde resulta que  $\|dg(x, y)\|_2 \leq \sqrt{A^2 + B^2}$ .

Si consideramos el caso particular en el que  $A = 0.6$  y  $B = 0.2$ , resulta que  $\|dg\|_2 \leq \sqrt{0.4} < 1$ ,  $g$  es contractiva en todo  $\mathbb{R}^2$  y por lo tanto posee un único punto fijo  $\vec{x}$ , que es el límite de cualquier sucesión de iteradas de  $g$ .

Comenzando a iterar en  $\vec{x}_0 = (0, 0)$ ,  $\vec{x}_1 = (-0.2, 0.6)$  y  $n = 72$  es el primer entero para el que se verifica la segunda desigualdad en:

$$\|\vec{x} - \vec{x}_n\| \leq \frac{\sqrt{0.4}^n}{1 - \sqrt{0.4}} \|\vec{x}_1 - \vec{x}_0\| < 10^{-14}.$$



En la tabla aparecen algunas iteradas  $\vec{x}_n$  de  $g$ , comenzando en  $\vec{x}_0 = (0, 0)$ ,

$n$	$\vec{x}_n = (x_{n+1}, y_{n+1})$	$\ \vec{x}_n - \vec{x}_{n-1}\ $
1	$(-0.2, 0.6)$	0.6324555320336759
2	$(-0.2842687214589724, 0.700968441383752)$	0.13151366306127302
3	$(-0.32111694605433183, 0.7049118182407534)$	0.03705862756317865
11	$(-0.3730792873028048, 0.6855454643808043)$	$7.616517317889067E - 4$
12	$(-0.37350553632157, 0.6853447127117006)$	$4.711575730540949E - 4$
13	$(-0.3737690885750729, 0.685220363357465)$	$2.9141474263543366E - 4$
31	$(-0.37419600023647853, 0.6850185752549108)$	$5.093329333095609E - 8$
32	$(-0.3741960287050984, 0.6850185617837148)$	$3.149500657466283E - 8$
33	$(-0.37419604630889514, 0.6850185534536928)$	$1.9475187463943547E - 8$
70	$(-0.37419607483155515, 0.6850185399569221)$	$3.510833468576701E - 16$
71	$(-0.37419607483155537, 0.685018539956922)$	$2.482534153247273E - 16$
72	$(-0.3741960748315555, 0.685018539956922)$	$1.1102230246251565E - 16$

El vector  $\vec{x}_{72} = (-0.37419607483155, 0.68501853995692)$  ya aproxima al punto fijo de  $g$  con precisión de  $10^{-14}$ .

En el ejemplo hemos comprobado que la función era contractiva utilizando la norma euclídea. En la siguiente proposición se establecen condiciones suficientes para controlar la norma de operadores asociada a la norma del supremo  $\|(x_1, \dots, x_m)\|_\infty = \max\{|x_1|, \dots, |x_m|\}$ , de la diferencial de una función.

**Proposición 5.1.5** Sea  $R = \{(x_1, \dots, x_m) : x_i \in [a_i, b_i] \ i = 1, \dots, m\}$  un intervalo  $m$ -dimensional y  $g : R \rightarrow R$  una función diferenciable para la que existe una constante  $0 < K < 1$  verificando:

$$\left| \frac{\partial g_i(\vec{x})}{\partial x_j} \right| \leq \frac{K}{m} \quad \text{para } \vec{x} \in R \text{ e } i, j = 1, \dots, m.$$

Entonces  $\|dg\|_\infty \leq K$  y  $g$  es contractiva para la norma del supremo

$$\|g(\vec{x}) - g(\vec{y})\|_\infty \leq K \|\vec{x} - \vec{y}\|_\infty \quad \text{para cada par } \vec{x}, \vec{y} \in R.$$

En consecuencia,  $g$  tiene un único punto fijo  $\vec{x}$  en  $R$  que es el límite de cualquier sucesión de iteradas  $\vec{x}_{n+1} = g(\vec{x}_n)$  con  $\vec{x}_0 \in R$ .

**Ejemplo 5.1.6** Consideremos el sistema de ecuaciones no lineales

$$\begin{cases} x_1 - 0.3 \cos(x_2 x_3) - 0.15 = 0 \\ x_1^2 - 100(x_2 + 0.2)^2 + \sin(x_3) + 1.1 = 0 \\ e^{-x_1 x_2 - 3} + x_3 + 0.5 = 0 \end{cases}$$

que podemos interpretar como un problema de punto fijo despejando  $x_i$  en cada ecuación:

$$\begin{cases} g_1(x_1, x_2, x_3) = 0.3 \cos(x_2 x_3) + 0.15 = x_1 \\ g_2(x_1, x_2, x_3) = 0.1 \sqrt{x_1^2 + \sin(x_3) + 1.1} - 0.2 = x_2 \\ g_3(x_1, x_2, x_3) = -e^{-x_1 x_2 - 3} - 0.5 = x_3 \end{cases}$$

Comenzamos observando en la primera ecuación que  $g_1(x_1, x_2, x_3) \in [-0.15, 0.45]$  para cualquier valor de  $x_2$  y  $x_3$ .

Suponiendo que  $x_1 \in [-0.15, 0.45]$ , se tiene que  $g_2(x_1, x_2, x_3) \in [-0.2, 0]$ . Y suponiendo también que  $x_2 \in [-0.2, 0]$ ,  $-x_1 x_2 - 3 \in [-3.3, -2.97]$ , y  $g_3(x_1, x_2, x_3) \in [-0.55, -0.53]$ .

Así, tenemos que  $g(M) \subset M$  para

$$M = [-0.9, 0.9] \times [-0.9, 0.9] \times [-0.9, -0.9] = \{(x_1, x_2, x_3) : -0.9 \leq x_1, x_2, x_3 \leq 0.9\}.$$

Vamos a usar la proposición 5.1.5 para comprobar que  $g$  es contractiva en el abierto convexo de  $\mathbb{R}^3$  que contiene a  $M$ ,  $R = (-1, 1) \times (-1, 1) \times (-1, 1)$ . Para ello calculamos y acotamos las derivadas parciales en  $R$ :

$$\begin{array}{lll} \frac{\partial g_1(\vec{x}_n)}{\partial x_1} = 0 & \left| \frac{\partial g_1(\vec{x}_n)}{\partial x_2} \right| = |0.3x_3 \sin(x_2x_3)| < 0.3 & \left| \frac{\partial g_1(\vec{x}_n)}{\partial x_3} \right| = |0.3x_2 \sin(x_2x_3)| < 0.3 \\ \left| \frac{\partial g_2(\vec{x}_n)}{\partial x_1} \right| = \left| \frac{0.1x_1}{\sqrt{x_1^2 + \sin(x_3) + 1.1}} \right| < \frac{0.1}{\sqrt{0.1}} < 0.32 & \frac{\partial g_2(\vec{x}_n)}{\partial x_2} = 0 & \left| \frac{\partial g_2(\vec{x}_n)}{\partial x_3} \right| = \left| \frac{0.1 \cos(x_3)}{2\sqrt{x_1^2 + \sin(x_3) + 1.1}} \right| < 0.16 \\ \left| \frac{\partial g_3(\vec{x}_n)}{\partial x_1} \right| = |x_2 e^{-x_1x_2-3}| < e^{-2} < 0.3 & \left| \frac{\partial g_3(\vec{x}_n)}{\partial x_2} \right| = |x_1 e^{-x_1x_2-3}| < 0.3 & \frac{\partial g_3(\vec{x}_n)}{\partial x_3} = 0 \end{array}$$

Con estas acotaciones, se tiene que  $\|dg(\vec{x})\|_\infty < 3 \times 0.32 = 0.96 < 1$  y por lo tanto,  $g$  es contractiva. Si realizamos las iteraciones  $\vec{x}_{n+1} = g(\vec{x}_n)$  comenzando en  $\vec{x}_0 = (0.1, 0.1, -0.1)$  obtenemos los siguientes valores que proporcionan una aproximación al único punto fijo de  $g$ :

$n$	$\vec{x}_n = (x_{n1}, x_{n2}, x_{n3})$	$\ \vec{x}_n - \vec{x}_{n-1}\ $
0	(0.1, 0.1, -0.1)	
1	(0.44998500012499953, -0.09949295629891547, -0.5492916787604621)	0.6034483843120066
2	(0.44955210741231366, -0.11165956397528992, -0.5520667014574826)	0.012486572364764038
3	(0.4494301918036675, -0.11181558685374286, -0.5523500070129856)	3.456422890955236E - 4
11	(0.4494277988953497, -0.11183576613150253, -0.552353426099008)	2.4279387088187486E - 15
12	(0.4494277988953497, -0.11183576613150269, -0.552353426099008)	1.5265566588595902E - 16
13	(0.4494277988953497, -0.11183576613150269, -0.552353426099008)	0.0

## Gauss-Seidel

Una forma de acelerar la convergencia en las iteraciones de punto fijo para funciones de varias variables consiste en utilizar en cada etapa del algoritmo las aproximaciones más recientes del punto fijo.

Por ejemplo la aceleración de Gauss-Seidel para el problema  $g(x_1, \dots, x_m) = (x_1, \dots, x_m)$  consiste en considerar, en lugar de  $\vec{x}_{n+1} = g(\vec{x}_n)$ , la sucesión

$$\vec{y}(n) = (y(n)_1, y(n)_2, \dots, y(n)_m)$$

definida por:

$$\begin{aligned} y(n+1)_1 &= g_1(y(n)_1, y(n)_2, \dots, y(n)_m) \\ y(n+1)_2 &= g_2(y(n+1)_1, y(n)_2, \dots, y(n)_m) \\ &\dots \\ y(n+1)_k &= g_k(y(n+1)_1, \dots, y(n+1)_{k-1}, y(n)_k, \dots) \\ &\dots \end{aligned} \tag{5.1}$$

En la definición de cada coordenada se utilizan las coordenadas recién calculadas.

Si se acelera la iterada de punto fijo del ejemplo anterior con este método de Gauss-Seidel se tienen los siguientes valores

$n$	$\vec{y}_n = (y_{n1}, y_{n2}, y_{n3})$	$\ \vec{y}_n - \vec{y}_{n-1}\ $
0	(0.1, 0.1, -0.1)	
1	(0.44998500012499953, -0.09033445920934564, -0.5518525664282725)	0.6024014014449389
2	(0.4496273043295007, -0.11180141591872639, -0.552353785613811)	0.021475786302636194
3	(0.44942814948717713, -0.1118357656165837, -0.5523534281396025)	$2.0209572175488585E - 4$
4	(0.4494277988963906, -0.11183576622996699, -0.5523534261013309)	$3.505972481043094E - 7$
5	(0.4494277988943376, -0.11183576613166642, -0.552353426099006)	$9.834949617314477E - 11$
6	(0.449427798895348, -0.11183576613150267, -0.552353426099008)	$1.0235978201637122E - 12$
7	(0.4494277988953497, -0.11183576613150269, -0.552353426099008)	$1.6653923600497446E - 15$
8	(0.4494277988953497, -0.11183576613150269, -0.552353426099008)	0.0

En este caso el método de Gauss-Seidel ha reducido el número de iteraciones. Aunque conviene resaltar que este método no siempre acelera la convergencia de punto fijo (como puede comprobar al repetir los cálculos del ejemplo 5.1.4)

## 5.2. Método de Newton

Recordemos que el método de Newton para resolver ecuaciones

$$f(x) = 0,$$

para funciones derivables una variable, consiste en sustituir en la ecuación la función  $f$  por el polinomio de Taylor de orden 1 en una aproximación  $x_n$  de la solución, después se resuelve la ecuación lineal

$$f(x_n) + f'(x_n)(x - x_n) = 0$$

(suponiendo que  $f'(x_n) \neq 0$ ) y se toma la solución de esta ecuación  $x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$ , como nueva aproximación de la solución buscada.

Al iterar este proceso, tenemos que si  $f$  es dos veces derivable con  $f''$  continua en un entorno de  $x_s$ , con  $f(x_s) = 0$  y  $f'(x) \neq 0$ , entonces  $x_n$  converge hacia  $x_s$  supuesto que  $x_0$  está próximo a  $x_s$ .

Vamos a estudiar cómo se puede trasladar fácilmente el método para aproximar ceros de funciones diferenciables de varias variables.

Consideramos un sistema de ecuaciones no lineales

$$\begin{cases} f_1(x_1, x_2, \dots, x_m) = 0 \\ f_2(x_1, x_2, \dots, x_m) = 0 \\ \dots\dots\dots \\ f_m(x_1, x_2, \dots, x_m) = 0 \end{cases}$$

definido por  $m$  funciones  $f_j : \mathbb{R}^m \rightarrow \mathbb{R}$ , representado por la ecuación vectorial

$$f(\vec{x}) = \vec{0}, \quad (5.2)$$

donde  $f = (f_1, f_2, \dots, f_m)$ .

Supongamos que  $f$  es diferenciable y que  $\vec{x}_k$  es una aproximación de una solución de (5.2). Aproximando en la ecuación,  $f(\vec{x})$  por la diferencial en  $\vec{x}_k$ , se tiene el sistema lineal:

$$f(\vec{x}_k) + df(\vec{x}_k)(\vec{x} - \vec{x}_k) = \vec{0}, \quad (5.3)$$

Que escrito en términos de las coordenadas de  $\vec{h}_k = \vec{x} - \vec{x}_k = (h_1, h_2, \dots, h_m)$  y las componentes de  $f$ , se escribe:

$$\begin{cases} \frac{\partial f_1(\vec{x}_k)}{\partial x_1} h_1 + \frac{\partial f_1(\vec{x}_k)}{\partial x_2} h_2 + \dots + \frac{\partial f_1(\vec{x}_k)}{\partial x_m} h_m = -f_1(\vec{x}_k) \\ \frac{\partial f_2(\vec{x}_k)}{\partial x_1} h_1 + \frac{\partial f_2(\vec{x}_k)}{\partial x_2} h_2 + \dots + \frac{\partial f_2(\vec{x}_k)}{\partial x_m} h_m = -f_2(\vec{x}_k) \\ \dots\dots\dots \\ \frac{\partial f_m(\vec{x}_k)}{\partial x_1} h_1 + \frac{\partial f_m(\vec{x}_k)}{\partial x_2} h_2 + \dots + \frac{\partial f_m(\vec{x}_k)}{\partial x_m} h_m = -f_m(\vec{x}_k) \end{cases}$$

Si la aplicación lineal  $df(x_k)$  de  $\mathbb{R}^n$  en  $\mathbb{R}^n$ , es biyectiva o, lo que es equivalente, si la matriz Jacobiana  $\left(\frac{\partial f_i(x_k)}{\partial x_j}\right)_{i,j}$  es no singular, el sistema lineal 5.3

$$(df(\vec{x}_k)) \vec{h}_k = -f(\vec{x}_k)$$

tiene una única solución  $\vec{h}_k$  que define una nueva aproximación de la solución:

$$\vec{x}_{k+1} = \vec{x}_k + \vec{h}_k.$$

Para uniformizar la notación con el caso de una variable podemos escribir

$$\vec{x}_{k+1} = \vec{x}_k - (df(\vec{x}_k))^{-1}(f(\vec{x}_k)),$$

aunque debemos resaltar que esta asignación no se utiliza en el algoritmo porque el cálculo de la inversa de una matriz requiere normalmente, muchas más operaciones que las necesarias para resolver el sistema lineal.

El algoritmo aparece en la página siguiente.

Razonando de forma análoga al caso de una variable se tiene el siguiente resultado de convergencia local

**Teorema 5.2.1** Sea  $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$  una función diferenciable de clase  $C^2(A)$ <sup>a</sup> Supongamos que  $\vec{x}_s \in A$  es un cero de  $f$ ,  $f(\vec{x}_s) = \vec{0}$ , y que  $df(\vec{x}_s)$  es biyectiva, o en términos de la matriz jacobiana, que ésta es no singular. Entonces, existe  $r > 0$  tal que la sucesión del método de Newton

$$\vec{x}_{n+1} = \vec{x}_n - (df(\vec{x}_n))^{-1}(f(\vec{x}_n)),$$

converge hacia  $\vec{x}_s$  para cualquier condición inicial  $x_0 \in A$  con  $\|x_0 - \vec{x}_s\| < r$ . Además la convergencia es de orden al menos 2. Es decir, existe una constante  $M > 0$  tal que

$$\|\vec{x}_{n+1} - \vec{x}\| \leq M\|x_n - \vec{x}\|^2$$

<sup>a</sup>  $f$  es de clase  $C^2(A)$ , si la función  $\vec{x} \rightarrow df(\vec{x}) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n) \equiv \mathbb{R}^{2m}$  es diferenciable con derivadas parciales continuas.

Representando la diferencial  $df(\vec{x})$  mediante la matriz jacobiana, sus componentes son las derivadas parciales de  $f$ ,  $\frac{\partial f(\vec{x})}{\partial x_k}$ . Así las derivadas parciales de la diferencial son las denominadas derivadas parciales segundas de  $f$ :

$$\frac{\partial^2 f(\vec{x})}{\partial x_j \partial x_k} = \frac{\partial \frac{\partial f(\vec{x})}{\partial x_k}}{\partial x_j}.$$

Estamos suponiendo que las parciales segundas de  $f$  son continuas.

**Algoritmo 5.1** Método de Newton (vectorial)**Datos de entrada:***F* %Función vectorial*dF* %Función diferencial de *F**X0* % aproximación inicial *X0**nmaxiteraciones* % número máximo de iteradas

global precision; % precisión para parada

**Variables:**

npasos=0;

*X*=*X0*;*FX*=*F*(*X*);*h*=*X*;**Fuio del programa:****while** npasos<nmaxiteraciones*J*=*dF*(*X*);**try***h*=solveGaussParcial(*J*, *FX*);**catch****error**('Diferencial singular');**end***X*=*X*-*h*;*FX*=*F*(*X*);**if** or(**norm**(*h*) < precision, **norm**(*FX*)<precision)**display**('cero de *F* en *X*')**break****endif**

npasos=npasos+1;

**endwhile****if** npasos==nmaxiteraciones

npasos

**error**('Newton no converge');**endif****Datos de salida:** [*X*,*FX*,npasos]. Raíz en *X* con *F*(*X*) = *FX* en npasos etapas, o mensaje de Error

## DEMOSTRACIÓN:

*Ideas que intervienen*

- Si se prueba que  $g(\vec{x}) = \vec{x} - (df(\vec{x}))^{-1}(f(\vec{x}))$  es una función contractiva en las proximidades de  $\vec{x}_s$  y este vector es su punto fijo. La sucesión  $\vec{x}_{n+1} = g(\vec{x}_n)$  converge hacia  $\vec{x}_s$ .
- Para probar que *g* es contractiva, bastará con probar que es diferenciable con derivadas parciales continuas y que  $dg(\vec{x}_s) = \vec{0}$ . En este caso *dg* es continua y existe  $r > 0$  tal que  $\|dg(\vec{x})\| < 0.5 < 1$  para todo  $\vec{x}$  con  $\|\vec{x} - \vec{x}_s\| < r$ .

- Para estimar el orden de convergencia se necesita una acotación del error al aproximar  $f(\vec{x}) - f(\vec{y})$  con la diferencial:

**Lema** Si la aplicación diferencial  $df : \vec{x} \rightarrow df(\vec{x})$  es Lipschitziana en un abierto convexo  $A$ , e.d. si existe  $K > 0$  tal que  $\|df(\vec{x}) - df(\vec{y})\| < K\|\vec{x} - \vec{y}\|$  para cada par  $\vec{x}, \vec{y}$  en  $A$ . Entonces,

$$\|f(\vec{x}) - f(\vec{y}) - df(\vec{y})(\vec{x} - \vec{y})\| \leq \frac{K}{2}\|\vec{x} - \vec{y}\|^2. \quad (5.4)$$

La prueba del lema se puede encontrar en el libro de Stoer y Burlisch [8, pág. 269]

La aplicación diferencial  $df$  es Lipschitziana en el supuesto de que  $f$  es de clase  $C^2(A)$  con  $\|d(df)(\vec{x})\| < K$  para cada  $\vec{x} \in A$ .

- Si se cumple (5.4), entonces

$$\begin{aligned} \|\vec{x}_{n+1} - \vec{x}_s\| &= \|(\vec{x}_n - \vec{x}_s) - (df(\vec{x}_n))^{-1}(f(\vec{x}_n))\| \\ &= \|(df(\vec{x}_n))^{-1}(df(\vec{x}_n)(\vec{x}_n - \vec{x}_s) - f(\vec{x}_n))\| \\ &\leq \|(df(\vec{x}_n))^{-1}\| \|df(\vec{x}_n)(\vec{x}_n - \vec{x}_s) - f(\vec{x}_n)\| \\ &= \|(df(\vec{x}_n))^{-1}\| \|f(\vec{x}_s) - f(\vec{x}_n) - df(\vec{x}_n)(\vec{x}_s - \vec{x}_n)\| \\ &\leq \|(df(\vec{x}_n))^{-1}\| \frac{K}{2} \|\vec{x}_n - \vec{x}_s\|^2 \end{aligned}$$

Tomando  $M > \sup\{\|(df(\vec{x}))^{-1}\| : \|\vec{x} - \vec{x}_s\| < r\} \frac{K}{2}$  se obtiene la convergencia de orden al menos 2, de  $\vec{x}_n$  hacia  $\vec{x}_s$ .

Vamos a ver  $g(\vec{x}) = \vec{x} - (df(\vec{x}))^{-1}(f(\vec{x}))$  es diferenciable con derivadas parciales continuas, en el caso particular de funciones de dos variables.

Supongamos  $f(x, y) = (f_1(x, y), f_2(x, y))$ , de clase  $C^2(A)$ .

$$df(\vec{x}) \equiv \begin{pmatrix} \frac{\partial f_1(\vec{x})}{\partial x} & \frac{\partial f_1(\vec{x})}{\partial y} \\ \frac{\partial f_2(\vec{x})}{\partial x} & \frac{\partial f_2(\vec{x})}{\partial y} \end{pmatrix}.$$

La diferencial  $df(\vec{x})$  es invertible si  $\det(df(\vec{x})) \neq 0$ . La función  $\det(df(\vec{x})) = \frac{\partial f_1(\vec{x})}{\partial x} \frac{\partial f_2(\vec{x})}{\partial y} - \frac{\partial f_1(\vec{x})}{\partial y} \frac{\partial f_2(\vec{x})}{\partial x}$  es suma de productos de funciones continuas diferenciables (aquí usamos que  $f$  es de clase 2), y en consecuencia, es continua y diferenciable en  $A$ . En particular, en un entorno de  $\vec{x}_s$  las diferenciales  $df(\vec{x})$  tienen inversas.

Las inversas de las diferenciales

$$df(\vec{x})^{-1} \equiv \begin{pmatrix} \frac{1}{\det(df(\vec{x}))} \frac{\partial f_2(\vec{x})}{\partial y} & -\frac{1}{\det(df(\vec{x}))} \frac{\partial f_1(\vec{x})}{\partial y} \\ -\frac{1}{\det(df(\vec{x}))} \frac{\partial f_2(\vec{x})}{\partial x} & \frac{1}{\det(df(\vec{x}))} \frac{\partial f_1(\vec{x})}{\partial x} \end{pmatrix} = \begin{pmatrix} a_{1,1}(\vec{x}) & a_{1,2}(\vec{x}) \\ a_{2,1}(\vec{x}) & a_{2,2}(\vec{x}) \end{pmatrix}.$$

tienen componentes diferenciables con derivadas parciales continuas (otra vez utilizamos que  $f$  es de clase 2).

Con esta notación, si escribimos  $g(\vec{x}) = (g_1(x, y), g_2(x, y)) = \vec{x} - (df(\vec{x}))^{-1}(f(\vec{x}))$  se tiene

$$\begin{cases} g_1(x, y) &= x - a_{1,1}(x, y)f_1(x, y) - a_{1,2}(x, y)f_2(x, y) \\ g_2(x, y) &= y - a_{2,1}(x, y)f_1(x, y) - a_{2,2}(x, y)f_2(x, y) \end{cases}$$

Las derivadas parciales de  $g_1$  son:

$$\begin{aligned}\frac{\partial g_1(\vec{x})}{\partial x} &= 1 - \frac{\partial a_{1,1}(\vec{x})}{\partial x} f_1(\vec{x}) - a_{1,1}(\vec{x}) \frac{\partial f_1(\vec{x})}{\partial x} - \frac{\partial a_{1,2}(\vec{x})}{\partial x} f_2(\vec{x}) - a_{1,2}(\vec{x}) \frac{\partial f_2(\vec{x})}{\partial x} \\ \frac{\partial g_1(\vec{x})}{\partial y} &= -\frac{\partial a_{1,1}(\vec{x})}{\partial y} f_1(\vec{x}) - a_{1,1}(\vec{x}) \frac{\partial f_1(\vec{x})}{\partial y} - \frac{\partial a_{1,2}(\vec{x})}{\partial y} f_2(\vec{x}) - a_{1,2}(\vec{x}) \frac{\partial f_2(\vec{x})}{\partial y}\end{aligned}$$

Teniendo en cuenta que  $f(\vec{x}_s) = (f_1(\vec{x}_s), f_2(\vec{x}_s)) = (0, 0)$ , se tiene que

$$\begin{aligned}\frac{\partial g_1(\vec{x}_s)}{\partial x} &= 1 - a_{1,1}(\vec{x}_s) \frac{\partial f_1(\vec{x}_s)}{\partial x} - a_{1,2}(\vec{x}_s) \frac{\partial f_2(\vec{x}_s)}{\partial x} \\ &= 1 - \frac{1}{\det(df(\vec{x}))} \left( \frac{\partial f_2(\vec{x}_s)}{\partial y} \frac{\partial f_1(\vec{x}_s)}{\partial x} - \frac{\partial f_1(\vec{x}_s)}{\partial y} \frac{\partial f_2(\vec{x}_s)}{\partial x} \right) = 1 - 1 = 0 \\ \frac{\partial g_1(\vec{x}_s)}{\partial y} &= -a_{1,1}(\vec{x}_s) \frac{\partial f_1(\vec{x}_s)}{\partial y} - a_{1,2}(\vec{x}_s) \frac{\partial f_2(\vec{x}_s)}{\partial y} \\ &= -\frac{1}{\det(df(\vec{x}))} \left( \frac{\partial f_2(\vec{x}_s)}{\partial y} \frac{\partial f_1(\vec{x}_s)}{\partial y} - \frac{\partial f_1(\vec{x}_s)}{\partial y} \frac{\partial f_2(\vec{x}_s)}{\partial y} \right) = 0\end{aligned}$$

De forma análoga, haciendo las derivadas parciales de  $g_2$  se observa que también se anulan en  $\vec{x}_s$ , y por lo tanto  $dg(\vec{x}_s) = \vec{0}$ . Concluyendo la prueba en este caso bidimensional.

Para otras dimensiones el resultado sigue siendo cierto, y se puede comprobar utilizando la regla de la cadena para la diferencial.  $\square$

### 5.3. Método de Broyden

Una dificultad del método de Newton en sistemas de ecuaciones es la necesidad de calcular la matriz Jacobiana en cada aproximación y resolver un sistema lineal para determinar la aproximación siguiente. Para ver la complejidad de cálculo que esto significa intentaremos cuantificar las operaciones que conlleva cada etapa:

Supongamos que vamos a aplicar Newton para resolver una ecuación  $F(x) = 0$  donde  $F : \mathbb{R}^m \rightarrow \mathbb{R}^m$  en cada etapa tendremos que

- (I) Determinar y evaluar las  $m^2$  derivadas parciales que aparecen en el Jacobiano. Aquí, podríamos usar derivación numérica como aprendimos en la asignatura anterior lo que supondría un mínimo de  $m^2$  evaluaciones de las componentes de  $F$  para aproximar el Jacobiano.
- (II) Resolver un sistema lineal de  $m$  ecuaciones con  $m$  incógnitas, que tendrá un coste del orden de  $m^3$  operaciones,

**En total: la complejidad la podemos cuantificar con  $m^2$  evaluaciones de las componentes y un número de operaciones del orden de  $m^3$ .**

Cuando estudiamos ecuaciones no lineales en una variable, la dificultad que supone el cálculo de derivadas la abordamos comparando Newton con el método de la secante. Para funciones de varias variables vamos a ver una generalización de este método conocida como el **método de Broyden que requerirá sólo  $m$  evaluaciones y un número de operaciones del**

**orden de  $m^2$  en cada etapa.** Este es uno de los métodos denominados cuasi-Newton que consisten en sustituir en cada etapa la matriz Jacobiana de la diferencial por una matriz de aproximación a costa de perder algo en el orden de convergencia, pasando de convergencia cuadrática a convergencia **superlineal**:

$$\lim_n \frac{\|\bar{x} - x_{n+1}\|}{\|\bar{x} - x_n\|} = 0.$$



Descripción del método:

- (I) Se parte de una aproximación inicial  $x_0$  del sistema no lineal  $F(x) = 0$ . Se calcula la siguiente aproximación  $x_1$  como en el método de Newton determinando el Jacobiano de  $F$ ,  $A_0 = J_{x_0}$  y definiendo  $x_1 - x_0 = -A_0^{-1}F(x_0)$ .
- (II) Pensando en el caso de dimensión 1, para calcular  $x_2$  podemos apartarnos del método de Newton y usar la aproximación de la derivada en  $x_1$  que proporciona el cociente incremental (pendiente de la secante en los puntos de abscisas  $x_0$  y  $x_1$ )

$$f'(x_1) \approx \frac{f(x_1) - f(x_0)}{x_1 - x_0}.$$

Para funciones de varias variables no podemos hacer el cociente anterior pero podemos buscar una matriz  $A_1$  que aproxime al Jacobiano de  $F$  en  $x_1$  y que cumpla

$$A_1(x_1 - x_0) = F(x_1) - F(x_0).$$

Pensando en  $A_1$  como en el operador  $T_{A_1}(x) = A_1x$ , tenemos definido  $T_{A_1}(x_1 - x_0)$  por la expresión anterior y sólo necesitamos definirlo en el subespacio ortogonal al vector  $u = (x_1 - x_0)$  ya que todo vector de  $\mathbb{R}^m$  se expresa de forma única como suma de un múltiplo de  $u$  y un vector ortogonal a  $u$ .

Una manera sencilla de definir  $T_{A_1}$  en este espacio ortogonal es haciendo que actúe como el Jacobiano  $A_0$ , sin aportar información adicional sobre la variación de  $F$  en las direcciones perpendiculares a  $u$ . Así pondríamos

$$A_1v = A_0v \quad \text{si} \quad \langle v, (x_1 - x_0) \rangle = 0.$$

Obsérvese que la matriz  $A_1$  que cumple las dos condiciones impuestas viene determinada por la expresión:

$$A_1 = A_0 + \frac{1}{\|x_1 - x_0\|_2^2} [F(x_1) - F(x_0) - A_0(x_1 - x_0)](x_1 - x_0)^t = A_0 + \frac{1}{\|x_1 - x_0\|_2^2} F(x_1)(x_1 - x_0)^t$$

Ahora definimos una nueva aproximación a la solución de  $F(x) = 0$  haciendo

$$x_2 = x_1 - A_1^{-1}F(x_1).$$

Para evitar calcular la inversa de  $A_1$  o resolver el sistema lineal correspondiente lo que significaría mantener la complejidad de la evaluación de una etapa del método de Newton, podemos acudir a la **Fórmula de Sherman-Morrison** cuya verificación podéis ver en el enlace

[http://en.wikipedia.org/wiki/Sherman%E2%80%93Morrison\\_formula](http://en.wikipedia.org/wiki/Sherman%E2%80%93Morrison_formula)

$$(A + vy^t)^{-1} = A^{-1} - \frac{1}{1 + y^t A^{-1}v} A^{-1}vy^t A^{-1}$$

en el supuesto de que  $y^t A^{-1}x \neq -1$ .

Ahora podemos escribir en la segunda etapa del método (en la definición de  $A_1^{-1}$  usaremos  $y = (x_1 - x_0)$  en lugar de  $y = \frac{1}{\|x_1 - x_0\|_1^2}(x_1 - x_0)$  afectando al denominador de la fórmula de Sherman-Morrison como sigue):

- a)  $B = A_0^{-1}$
- b)  $y = (x_1 - x_0)$
- c)  $v = F(x_1)$ .
- d)  $A_1^{-1} = A - \frac{1}{y^t(y+Bv)}(Bv)(y^t B)$
- e)  $x_2 = x_1 - A_1^{-1}F(x_1)$ .

Observad que una vez guardado  $A_0^{-1}$  en la primera etapa (para calcular  $x_1$ ), en la segunda etapa el número de operaciones es el de multiplicar matrices por vectores fila o columna, multiplicar vectores o sumar vectores y matrices, que es del orden de  $m^2$ .

- (III) Supuesto que tenemos calculada la aproximación  $x_n$  y guardada la inversa  $B = A_{n-1}^{-1}$  podemos repetir el proceso de la segunda etapa para ir construyendo aproximaciones sucesivas a la solución del sistema  $F(x) = 0$ .

### Algoritmo 5.2 Método de Boyren (vectorial)

#### Datos de entrada:

$F$  %Función vectorial  
 $X0$  % aproximación inicial  $X0$   
 $dFX0$  % Diferencial de  $F$  en  $X0$ , Jacobiano  
 $nmaxiteraciones$  % número máximo de iteradas  
 $global\ precision$ ; % precisión para parada

#### Variables:

$npasos=0$ ;  
 $X=X0$ ;  
 $FX=fval(F,X)$ ;  
 $h=X$ ;  
 $invJa=inv(dFX0)$ ;

#### Fujo del programa:

```

while npasos<nmaxiteraciones
h=-invJa*FX;
X=X+h;
FX=F(X);
if or(norm(h)< precision , norm(FX)<precision)
display('cero de F en X')
break
endif
npasos=npasos+1;
u=invJa*FX;
iden=1/(dot(h+u,h));
w=(iden*h)'*invJa;
invJa=invJa-u*w;
endwhile
if npasos==nmaxiteraciones
npasos
error('quasinewton (Boyren) no converge');
endif

```

**Datos de salida:**  $[X,FX,npasos]$ . Raíz en  $X$  con  $F(X) = FX$  en  $npasos$  etapas, o mensaje de Error

En prácticas implementaremos el método y lo compararemos con el de Newton. Nuestra esperanza será que converja en condiciones análogas a las del método de Newton, cuando iniciemos el algoritmo con una buena aproximación de la solución, utilizando alguna etapa más que éste pero haciendo menos operaciones.

## 5.4. Método del descenso rápido: “Steepest Descent”

(Seguiremos la sección 10.3 del libro de Burden-Faires [4])

El método de Newton o el de Broyden para resolver sistemas es eficiente *si partimos de una aproximación a la solución adecuada*.

El método del gradiente o *descenso rápido* que veíamos para resolver sistemas lineales también puede usarse para localizar aproximaciones a soluciones de sistemas no lineales que sirvan para iniciar Newton o Boyren.

Consideremos el sistema de ecuaciones no lineales

$$\begin{cases} f_1(x_1, x_2, \dots, x_m) == 0 \\ f_2(x_1, x_2, \dots, x_m) == 0 \\ \dots\dots\dots \\ f_m(x_1, x_2, \dots, x_m) == 0 \end{cases}$$

Los ceros de este sistema son los vectores donde se alcanzan los *mínimos absolutos* de la siguiente función:

$$g((x_1, x_2, \dots, x_m)) = \sum_{k=1}^m (f_k(x_1, x_2, \dots, x_m))^2.$$

Esta función anterior recibe el nombre de “función objetivo del sistema”.

### Ideas del método

Método del gradiente o descenso rápido para localizar mínimos locales (relativos) de  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  una función diferenciable:

- (I) Se parte de una aproximación inicial  $x^{(0)} = (x_1^{(0)}, x_2^{(0)}, \dots, x_m^{(0)})$ .
- (II) Si se han realizado  $k$  iteraciones y se tiene una aproximación  $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_m^{(k)})$ :  
Se evalúa  $g(x^{(k)})$  (Para funciones objetivo ver si  $|g(x^{(k)})|$  es pequeño y parar)
- (III) Se evalúa el gradiente  $\nabla g(x^{(k)})$  que proporciona la dirección de máximo crecimiento y también la de máximo descenso (su opuesto).
- (IV) Nos movemos en esa dirección una distancia “adecuada” para fijar una nueva aproximación del mínimo  $x^{(k+1)} = (x_1^{(k+1)}, x_2^{(k+1)}, \dots, x_m^{(k+1)})$
- (V) Se repiten los pasos (II) a (IV) mientras  $\nabla g(x^{(k)})$  no sea nulo o sobrepasemos un número prefijado de pasos.

**Paso IV, movimiento a una distancia “adecuada”**

- (a) Se normaliza el gradiente  $u = \frac{1}{\|\nabla g(x^{(k)})\|} \nabla g(x^{(k)})$
- (b) Se hace  $\alpha = 1$
- (c) Mientras  $g(x^{(k)} - \alpha u) > g(x^{(k)})$  se hace la reducción  $\alpha = \frac{1}{2}\alpha$
- (d) Si  $|\alpha|$  es menor que la tolerancia admitida se termina avisando que no se puede mejorar la aproximación  $x_k$  a un mínimo. En otro caso
- (e) se toman los tres puntos  $(0, g(x^{(k)}))$ ,  $(\frac{\alpha}{2}, g(x^{(k)} + \frac{\alpha}{2}u))$  y  $(\alpha, g(x^{(k)} + \alpha u))$ , se construye el polinomio interpolador que representa una parábola, y se busca la abscisa  $\alpha_0$  del vértice de la misma.
  - $g_1 = g(x^{(k)}), g_2 = g(x^{(k)} + \frac{\alpha}{2}u), g_3 = g(x^{(k)} + \alpha u)$
  - $h_1 = 2(g_2 - g_1)/\alpha$  (diferencia dividida de Newton)
  - $h_2 = 2(g_3 - g_2)/\alpha$
  - $h_{1,2} = (h_2 - h_1)/\alpha$  (segunda diferencia dividida)
  - $\alpha_0 = \frac{1}{2}(\frac{\alpha}{2} - \frac{h_1}{h_{1,2}})$
- (f) Si  $g(x^{(k)} + \alpha_0 u) < g(x^{(k)} + \alpha u)$  (en el vértice hay un mínimo) se hace  $x^{(k+1)} = x^{(k)} + \alpha_0 u$ , en otro caso (en el vértice hay un máximo) se hace  $x^{(k+1)} = x^{(k)} + \alpha u$ .

El método del descenso rápido no hace honor a su nombre y no es un método muy rápido. Para usarlo en la búsqueda aproximaciones de ceros de sistemas de ecuaciones con la función objetivo, se consideran precisiones no muy finas y una vez obtenidas estas aproximaciones se usan con Newton o Broyden.

(Ver ejemplos de clase)

## 5.5. Ejercicios

**Ejercicio 5.1** Los puntos 2-periódicos de una función  $f$  son aquellos valores de  $x$  que cumplen  $x = f(f(x))$ . Observa que los puntos 2-periódicos satisfacen el sistema de ecuaciones

$$\begin{cases} x &= f(y) \\ y &= f(x) \end{cases}$$

Para su resolución, podemos usar un método iterativo de tipo Jacobi:

$$x_{n+1} = f(y_n), y_{n+1} = f(x_n),$$

o iteraciones de tipo Gauss-Seidel:

$$x_{n+1} = f(y_n), y_{n+1} = f(x_{n+1}).$$

Da condiciones suficientes de convergencia de los métodos citados suponiendo conocida la derivada de  $f$  en la solución.

Como aplicación, encuentra puntos 2-periódicos de  $f(x) = 1.4 \cos x$  a partir de la condición inicial  $x_0 = 1$ .

**Ejercicio 5.2** Sea  $C = \{(x, y, z) : a_1 \leq x \leq b_1, a_2 \leq y \leq b_2, a_3 \leq z \leq b_3\}$  un cubo en  $\mathbb{R}^3$ , y sea  $g : C \rightarrow C$  con componentes  $g_i$ ,  $g = (g_1, g_2, g_3)$ . Entonces  $g$  tiene un punto fijo. Supongamos que existe una constante  $K < 1$  tal que  $\left| \frac{\partial g_i(X)}{\partial x_j} \right| \leq \frac{K}{3}$  para todo  $X \in C$ . Prueba, usando la norma  $\|\cdot\|_\infty$ , que  $g$  es contractiva y que el punto fijo se puede obtener como límite de las iteradas funcionales  $X_{n+1} = g(X_n)$ .

**Ejercicio 5.3** El sistema no lineal

$$\begin{cases} x^2 - 10x + y^2 + 8 & == & 0 \\ xy^2 + x - 10y + 8 & == & 10 \end{cases}$$

puede transformarse en el problema de punto fijo

$$\begin{cases} g_1(x, y) = 0.1(x^2 + y^2 + 8) & == & x \\ g_2(x, y) = 0.1(xy^2 + x + 8) & == & y \end{cases}$$

Prueba que  $G = (g_1, g_2)$  tiene un punto fijo en  $R = [0, 1.5] \times [0, 1.5]$ .

Aplica la iteración funcional para aproximar la solución.

¿Acelera la convergencia el método de Gauss-Seidel?

**Ejercicio 5.4** Describe un método iterativo de punto fijo para resolver el sistema no lineal

$$\begin{cases} 3x - \cos(y + z) & = & \frac{1}{2} \\ x^2 - 81(y + 0.1)^2 + \sin z & = & -1.06 \\ e^{-xy} + 20z & = & \frac{10\pi - 3}{3} \end{cases}$$

**Ejercicio 5.5** El sistema no lineal

$$\begin{cases} 5x^2 - y^2 & = & 0 \\ y - \frac{1}{4}(\sin x + \cos y) & = & 0 \end{cases}$$

tiene una solución próxima a  $(\frac{1}{4}, \frac{1}{4})$ .

(I) Encuentra una función vectorial  $g$  y un conjunto  $D$  en el plano de forma que  $g$  tenga un único punto fijo en una solución al sistema anterior.

(II) Aplica la iteración de punto fijo para determinar esta solución con una precisión de  $10^{-3}$  para la norma  $\|\cdot\|_\infty$ .

¿Acelera la convergencia el método de Gauss-Seidel?

**Ejercicio 5.6** Utiliza la iteración funcional para obtener una solución aproximada al siguiente sistema no lineal, con una exactitud de  $10^{-3}$  para la norma  $\|\cdot\|_\infty$ .

$$\begin{cases} 3x^2 - y^2 & = & 0 \\ 3xy^2 - x^3 - 1 & = & 0 \end{cases}$$

¿Acelera la convergencia el método de Gauss-Seidel?

**Ejercicio 5.7** Elige una condición inicial y efectúa dos iteraciones del método de Newton en el sistema

$$\begin{cases} 4x^2 - y^2 = 0 \\ 4xy^2 - x = 1 \end{cases}$$

**Ejercicio 5.8** Comenzando en  $(0,0,1)$ , efectúa una iteración del método de Newton para el sistema de ecuaciones

$$\begin{cases} xy - z^2 = 1 \\ xyz - x^2 + y^2 = 2 \\ e^x - e^y + z = 3 \end{cases}$$

**Ejercicio 5.9** ¿A qué se reduce el método de Newton cuando se intenta aplicar en la resolución de una ecuación lineal  $A\vec{x} = \vec{b}$ ?

**Ejercicio 5.10** Utiliza el ordenador y los métodos de Newton y Boyren para encontrar una solución del sistema de ecuaciones no lineales:

$$\begin{cases} 4y^2 + 4y + 52x = 19 \\ 169x^2 + 3y^2 + 111x - 10y = 10 \end{cases}$$

**Ejercicio 5.11** Utiliza el método del descenso rápido junto con el de Newton o el de Boyren para aproximar soluciones de los sistemas o ecuaciones no lineales siguientes

(I)

$$\begin{cases} 4x^2 - 20x + 0.25y^2 + 8 = 0 \\ 14xy^2 + 2x - 5y + 8 = 0 \end{cases}$$

(II)

$$\begin{cases} x + \cos(xyz) - 1 = 0 \\ (1-x)^{0.25} + y + 0.05z^2 - 0.15z - 1 = 0 \\ -x^2 - 0.1y^2 + 0.01y + z - 1 = 0 \end{cases}$$

## Material adicional

Uso de Octave/Matlab para representar soluciones.

Práctica 8

# Bibliografía

- [1] A. Delshams A. Aubanell, A. Benseny, *Útiles básicos de cálculo numérico*, Ed. Labor - Univ. Aut. de Barcelona, Barcelona, 1993.
- [2] G. Allaire and S.M. Kaber, *Numerical linear algebra*, TAM 55, Springer,, New York, etc., 2008.
- [3] C. Brézinski, *Introduccion á la pratique du calcul numérique*, Dunod Université, Paris, 1988.
- [4] R.L. Burden and J.D. Faires, *Análisis numérico*, 7ª edición, Thomson-Learning, Mexico, 2002.
- [5] P.G. Ciarlet, *Introduction à l'analyse numérique matricielle et à l'optimisation*, Masson, Paris, 1990.
- [6] P.M. Cohn, *Algebra, vol. 1 y 2*, Ed. Hohn Wiley and Sons, London, 1974.
- [7] G. Hammerlin and K.H. Hoffmann, *Numerical mathematics*, Springer-Verlag, New York, 1991.
- [8] J.Stoer and R. Burlisch, *Introduction to numerical analysis*, Springer Verlag, New York, 1980.
- [9] D. Kincaid and W. Cheney, *Análisis numérico*, Ed. Addison-Wesley Iberoamericana, Reading, USA, 1994.
- [10] J. R. Shewchuk, *An introduction to the conjugate gradient method without the agonizing pain*, Tech. report, Carnegie Mellon University, USA,<http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient-figs.pdf>,doi=10.1.1.110.418, 1994.