

ERRORES, OPTIMIZACIÓN Y RESOLUCIÓN NUMÉRICA DE SISTEMAS

Valentín Gregori Gregori
Bernardino Roig Sala

Enero 2015

Presentación

La Universidad Española ha emprendido una etapa inédita con el denominado Plan Bolonia. En el nuevo plan el tiempo del que dispone el profesorado para la impartición de la docencia matemática se ha reducido drásticamente. De esta manera la clásica clase magistral del siglo anterior se vuelve, en ocasiones, menos expositiva y más orientada hacia la búsqueda de conocimientos en los que el universitario deberá involucrarse de una manera más activa.

Este texto encuentra un hueco entre los textos matemáticos de investigación numérica en los que se estudia la teoría con detalle y se realizan algunas aplicaciones, y aquellos otros más centrados en la ingeniería que sin apenas justificación alguna se dedican a enfatizar las aplicaciones.

El presente libro es un texto sobre la resolución numérica de problemas esenciales para los alumnos de ingeniería que se graduarán en estos nuevos planes. Básicamente el contenido corresponde a un curso que los autores han impartido en la Escuela Politécnica Superior de Gandia en anteriores años académicos. El poco tiempo de que se dispone para su impartición queda patente, en cierta manera, en la ausencia de demostraciones que sólo aparecen esporádicamente en letra pequeña si éstas permiten entender mejor el capítulo. Ello, sin embargo, permite una lectura más fluida del texto.

Este libro sirve de base para primeros cursos de grado enfatizando la parte algorítmica y computacional y para cursos superiores en donde ya es posible incidir con detalle en todo el contenido del libro.

Para la comprensión del texto sólo se requieren nociones del análisis matemático y matricial, así como una base de algorítmica y programación. Todos los requisitos teóricos utilizados se han explicitado previamente de forma breve. Los algoritmos o comandos presentados se han escrito de la forma que se ha creído más sencilla y compacta en la versión actual de Matlab. Obviamente dichos algoritmos se pueden escribir de forma más general y abarcando más casuísticas, ahora bien, ello desvía el enfoque de este libro que incide más en la comprensión de los métodos y conceptos tratados. En <http://personales.upv.es/~broig/> se puede encontrar un fichero con todos los comandos Matlab incluidos en el texto. Se han recuadrado las ecuaciones más esenciales desde la perspectiva numérica sobre todo para facilitar su uso a estudiantes de primeros cursos.

Los autores agradecerán cualquier sugerencia tendente a mejorar el presente texto en ediciones sucesivas.

Los autores.

NOTACIÓN:

En este texto se ha evitado un lenguaje excesivamente simbólico. No obstante, el lector debe conocer la siguiente terminología básica que se usa en matemáticas y ciencias tecnológicas:

$\forall$	Cuantificador universal. Se lee “para todo” o “para cada”
$\exists$	Cuantificador existencial. Se lee “existe”
$\iff$	Equivalencia proposicional. Se lee “si y sólo si”
sii	Abreviatura de “si y sólo si”
$\Rightarrow$	Implicación proposicional. La proposición de la izquierda implica la de la derecha. Se lee “implica”
	Se lee “tal (tales) que”
:	Se lee “tal (tales) que”
i.e.	En latín <i>id est</i> y se lee “es decir”
$\in$	Símbolo de pertenencia
$\subset$	Símbolo de inclusión
$\cup$	Símbolo de unión
$\cap$	Símbolo de intersección
$\mathbb{N}$	Conjunto de los números naturales (incluye al cero)
$\mathbb{N}^*$	El conjunto $\mathbb{N}$ sin el cero
$\mathbb{Z}$	El anillo de los números enteros
$\mathbb{Q}$	El cuerpo de los números racionales
$\mathbb{R}$	El cuerpo de los números reales
$\mathbb{C}$	El cuerpo de los números complejos

Sumario

1 REPRESENTACIÓN NUMÉRICA Y TEORÍA DE ERRORES	9
1.1 INTRODUCCIÓN	9
1.2 REPRESENTACIÓN DE NÚMEROS AL ORDENADOR	10
1.2.1 El sistema decimal y el binario	11
1.2.2 Números naturales	12
1.2.3 Números enteros	13
1.2.4 Números reales	14
1.2.5 Ejemplo con 4 decimales en binario	17
1.2.6 El formato en punto flotante IEEE-754	18
1.2.7 El formato punto flotante de precisión doble	19
1.3 CÁLCULO APROXIMADO DE FUNCIONES	22
1.3.1 Teorema de Rolle	22
1.3.2 Teorema del valor medio (de Lagrange)	22
1.3.3 Teorema de Taylor (fórmula de Taylor)	22
1.3.4 Teorema de Taylor en funciones de varias variables	28
1.3.5 Aproximaciones de primer y segundo orden de una función de una variable	31
1.3.6 Aproximaciones de primer y segundo orden de una función de varias variables	32
1.3.7 Aproximación de primer orden de un campo vectorial	37
1.4 ERRORES	38
1.4.1 Definiciones de errores	39
1.4.2 Observación sobre el error máximo de truncamiento o redondeo	40
1.4.3 Definiciones de precisión	41
1.4.4 Cifras exactas de un decimal	41
1.4.5 Cotas del error	42
1.5 PROPAGACIÓN DE ERRORES	42
1.5.1 Definiciones	43
1.5.2 Propagación del error en operaciones aritméticas	44
1.5.3 Problemas por cancelaciones	45

1.5.4	La diferencial de una función	48
1.5.5	Propagación del error en funciones	49
1.5.6	Propagación del error en funciones de varias variables	53
1.5.7	Propagación del error en campos vectoriales	57
1.6	ALGUNAS INDICACIONES PARA LA COMPUTACIÓN	61
1.7	EJERCICIOS	61
1.8	MAPA CONCEPTUAL	65
2	RESOLUCIÓN DE ECUACIONES	67
2.1	ELEMENTOS DE ANÁLISIS	68
2.1.1	Teorema de Bolzano	68
2.1.2	Minimización de funciones	68
2.1.3	Convergencia de una sucesión	69
2.1.4	Orden de convergencia	70
2.1.5	Teorema de completitud de $\mathbb{R}$	71
2.2	RESOLUCIÓN DE ECUACIONES NO LINEALES	72
2.2.1	Criterios de finalización	73
2.2.2	Método de bisección	74
2.2.3	Método de regula-falsi	77
2.3	NUEVOS ELEMENTOS DEL ANÁLISIS	81
2.3.1	Funciones contractivas	81
2.3.2	Teorema. Condición suficiente de contractividad	82
2.3.3	Teorema del punto fijo	83
2.3.4	Cotas del error en el método iterativo del punto fijo	83
2.3.5	Teorema sobre el orden del método iterativo	85
2.4	MÉTODOS ITERATIVOS DE RESOLUCIÓN DE ECUACIONES	85
2.4.1	Método iterativo del punto fijo	85
2.4.2	Método de Newton	88
2.4.3	Método de la secante	92
2.4.4	Estudio del caso de raíces múltiples en el método de Newton	96
2.4.5	Aceleración de la convergencia	98
2.4.6	Notas	100
2.4.7	Extensión a $\mathbb{C}$	103
2.4.8	El comando fzero de Matlab	103
2.5	EJERCICIOS	104
2.6	MAPA CONCEPTUAL	112
3	RESOLUCIÓN DE SISTEMAS DE ECUACIONES LINEALES	113
3.1	ELEMENTOS DE ÁLGEBRA	114
3.1.1	Producto escalar	114
3.1.2	Normas vectoriales	117

3.1.3	Ángulo entre dos vectores	118
3.1.4	Aproximaciones del error en un algoritmo iterativo	119
3.1.5	Norma matricial, número de condición y cotas del error	119
3.1.6	Estudio del error en sistemas de ecuaciones lineales	122
3.1.7	Ejemplo de propagación del error en sistemas	126
3.1.8	Precondicionado en sistemas lineales	128
3.1.9	Observación sobre los números complejos	130
3.2	MÉTODOS DIRECTOS DE RESOLUCIÓN DE SISTEMAS LINEALES . .	131
3.2.1	Resolución de sistemas simples	131
3.2.2	Método de Gauss y descomposición LU	132
3.2.3	Método de Gauss y descomposición LU con pivotaje	135
3.3	MÉTODOS ITERATIVOS DE RESOLUCIÓN DE SISTEMAS LINEALES .	140
3.3.1	Convergencia de métodos iterativos matriciales	141
3.3.2	Criterios de finalización	143
3.3.3	Método Jacobi	144
3.3.4	Método Gauss-Seidel	145
3.3.5	Métodos de sobrerelajación	147
3.3.6	Observaciones	149
3.4	EJERCICIOS	155
3.5	MAPA CONCEPTUAL	164
4	OPTIMIZACIÓN Y RESOLUCIÓN DE SISTEMAS DE ECUACIONES NO LINEALES	165
4.1	ELEMENTOS DEL ÁLGEBRA Y DEL ANÁLISIS	166
4.1.1	Direcciones conjugadas. Ortogonalidad	166
4.1.2	Curvas de nivel y dirección de máximo descenso	168
4.1.3	Clasificación de puntos críticos en un campo escalar	169
4.1.4	Campo vectorial contractivo y teorema del punto fijo	170
4.2	RESOLUCIÓN DE SISTEMAS POR MÍNIMOS CUADRADOS	171
4.2.1	Ejemplo de ajuste polinómico	172
4.2.2	Extensión a los complejos	173
4.3	RESOLUCIÓN DE SISTEMAS DE ECUACIONES NO LINEALES	174
4.3.1	Criterios de finalización de un esquema iterativo	174
4.3.2	Método iterativo en varias variables	175
4.3.3	Método de Newton multivariable	177
4.3.4	Métodos cuasi-Newton	181
4.3.5	Observaciones	184
4.4	MÉTODOS DE OPTIMIZACIÓN	185
4.4.1	Criterios de finalización de un esquema iterativo de minimización . .	185
4.4.2	Método del máximo descenso	186
4.4.3	Método del gradiente conjugado	191

4.4.4	Método de minimización de Newton	194
4.4.5	Métodos de minimización cuasi-Newton	199
4.4.6	La función fminunc	204
4.4.7	Resolución de sistemas mediante métodos de minimización	205
4.5	EJERCICIOS	206
4.6	MAPA CONCEPTUAL	211
	BIBLIOGRAFÍA	213

Índice

- ángulo entre vectores, 118
- aceleración de la convergencia, 98
- ajuste
 - polinómico, 172
- ancho de banda, 114
- aproximación
 - cuadrática, 31–33
 - de primer orden, 31
 - de segundo orden, 31
 - del error absoluto, 71, 119
 - del error relativo, 71, 119
 - del orden de convergencia, 72
 - lineal, 31–33, 37
- backtracking, 187
- base canónica, 115
- bien condicionado, 43, 50, 123, 134
- campo vectorial, 37, 57
 - contractivo, 170
- cancelación numérica, 45
- catenaria, 107
- Cholesky, 138
- cifras
 - exactas, 41
 - significativas, 41
- coma flotante, 15
- complejos, 103, 130, 173
- comportamiento
 - del error, 101
 - lineal, 101
 - potencial, 101
- composición de funciones, 82
- condición de convergencia, 141
 - suficiente, 141, 144, 146
 - suficiente y necesaria, 141
- constante
 - de convergencia, 70
 - de error asintótico, 70
- convergencia
 - cúbica, 70
 - cuadrática, 70, 90
 - de primer orden, 70
 - de segundo orden, 70
 - de tercer orden, 70
 - lineal, 70, 71
 - superlineal, 70
- coste
 - computacional, 134
 - operativo, 103
- cota
 - del error, 11, 73, 74, 143
 - absoluto, 42
 - absoluto en el método iterativo, 84
 - relativo, 42
 - del número de iteraciones, 84, 143, 171
 - del residuo de Taylor, 23, 29, 30
- criterio
 - de Sylvester, 115
 - para finalizar un algoritmo, 73, 143, 174, 185
- cuádriga tangente, 33
- curva de nivel, 169
- dígitos significativos, 41
- decimales
 - correctos, 41
 - exactos, 41
- definida
 - negativa, 115
 - positiva, 115
- derivada
 - direccional, 168
 - logarítmica, 49
- descomposición LU, 133
- diagonal principal, 114
- diferencia
 - finita, 161
 - progresiva, 99

- diferencia finita, 179
- diferencial, 48
 - logarítmica, 49
- dirección conjugada, 118, 166
- eficiencia, 103
- eliminación gaussiana, 132
- epsilon, 20
- error
 - absoluto, 11, 39, 59
 - de una iteración, 70
 - en una iteración, 142
 - cuadrático, 171
 - de discretización, 39
 - de incertidumbre, 38
 - de modelización, 38
 - de redondeo, 38, 39
 - de representación, 38, 43
 - de truncamiento, 38, 39
 - experimental, 38
 - máximo
 - de redondeo, 40
 - de truncamiento, 40
 - relativo, 11, 39, 59
 - de una iteración, 70
 - residual, 124
 - vectorial, 141
- esquema iterativo, 140
- estable, 43
- fórmula de Sherman-Morrison, 182
- factorización de Cholesky, 138
- fminunc, 204
- función
 - cóncava, 68
 - compleja, 103
 - componente, 37, 57
 - contractiva, 82
 - convexa, 68
 - objetivo, 185
- gradiente, 168
- incremento
 - de $\mathbf{x}$, 32, 33
 - de x , 161
- inestable, 43
- linealización, 31, 33
- LU, 133
- máxima pendiente, 168
- máximo, 68, 120, 170
 - ascenso, 168
 - descenso, 169
- método
 - Δ^2 de Aitken, 99
 - convergente, 69
 - cuasi-Newton, 181
 - de bisección, 74, 76
 - de Broyden, 182
 - de Cholesky, 138
 - de descenso, 185
 - de eliminación de Gauss, 132
 - de Gauss
 - con pivotaje completo, 137
 - con pivotaje parcial, 135
 - sin pivotaje, 136
 - de Gauss-Seidel, 145, 146
 - de interpolación cuadrática inversa, 103
 - de Jacobi, 144, 145
 - de la cuerda de Whitaker, 91
 - de la posición falsa, 77
 - de la secante, 92, 94
 - de las diferencias finitas, 161
 - de minimización, 185
 - de Newton, 88, 91, 177
 - de Newton modificado, 178
 - de optimización, 185
 - de regula-falsi, 77, 79
 - de regula-falsi modificado, 78, 81
 - de relajación, 148
 - de Schröder, 96, 97
 - de sobrerrelajación, 147–149
 - de sobrerrelajación sucesiva, 147
 - de Steffensen, 94, 95, 99
 - de subrelajación, 147
 - del máximo descenso, 186
 - directo, 131
 - iterativo, 72, 140
 - del punto fijo, 83, 85, 87
 - multivariable, 175
 - por intervalos, 73
 - puntual, 73
 - SOR, 147, 148
- mínima pendiente, 169
- mínimo, 68, 169, 185
- mal condicionado, 43, 50, 123, 134
- matriz
 - ampliada, 133
 - banda, 114
 - de Hilbert, 155

- de iteración, 140
- de permutación, 134
- de Vandermonde, 155
- diagonal, 114, 131
- diagonal dominante, 116
- dispersa, 114
- escasa, 114
- hermítica, 130
- hessiana, 32, 33, 169
- jacobiana, 37
- regular, 132
- triangular inferior, 114, 131
- triangular superior, 114, 131
- tridiagonal, 114
- multiplicadores, 132
- multiplicidad de una raíz, 96
- número de condición, 50, 55, 122
 - parcial, 54
- NaN, 20
- nodo, 161
- norma, 117
 - p , 117
 - del máximo, 117
 - euclídea, 117
 - infinito, 55, 117
 - matricial, 119, 122
 - asociada, 120
 - de Frobenius, 119
 - dos, 120
 - euclídea, 120
 - infinito, 59, 120
 - uno, 120
 - uno, 117
- numéricamente
 - estable, 42
 - inestable, 42
- orden
 - de aproximación, 23, 30
 - de convergencia, 70
 - de convergencia de un método, 70
 - del residuo de Taylor, 23, 30
- overflow, 20
- parábola tangente, 31
- pivotaje, 135
 - completo, 136
 - parcial, 135
- pivote, 132
- plano tangente, 32
- polinomio de Taylor, 23, 29
- precisión, 15
 - cuádruple, 18
 - doble, 18
 - simple, 18
- precondicionado, 128
- producto escalar, 114, 116
 - habitual, 116
- propagación del error
 - absoluto en funciones, 50, 54, 56
 - absoluto en sistemas lineales, 124
 - en operaciones aritméticas, 44
 - relativo en funciones, 50, 56
 - relativo en sistemas lineales, 124
- proyección ortogonal, 166
- punto
 - crítico, 68, 169
 - de inflexión, 68
 - fijo, 14, 82, 170
 - flotante, 15
 - silla, 170
- radio espectral, 121, 122
- rango de representación, 9, 14
- recta tangente, 31
- reescalado, 128
- residuo, 205
 - de Taylor, 23, 29, 30
- resolución
 - aproximada, 187
 - exacta, 187
 - inexacta, 187
- resto de Lagrange, 23
- semidefinida
 - negativa, 115
 - positiva, 115
- serie
 - de Mc Laurin, 23
 - de Taylor, 23
- sistema
 - bien condicionado, 123
 - binario, 11
 - decimal, 11
 - mal condicionado, 123
 - rígido, 113
- subespacio ortogonal, 166
- subespacio propio, 166
- sucesión
 - convergente, 69, 119
 - de Cauchy, 71

- supremo, 120
- sustitución
 - hacia adelante, 132
 - hacia atrás, 131
 - progresiva, 132
 - regresiva, 131
- teorema
 - condición suficiente de contractividad, 82
 - de Bolzano, 68
 - de Rolle, 22
 - de Taylor, 22, 28
 - del punto fijo, 83
 - del valor medio, 22
 - orden de un método iterativo, 85
- tipos de representación
 - de números enteros, 14
 - de números reales, 16
 - formato IEEE-754, 18
 - precisión doble, 19
- tolerancia, 73, 74, 143
- underflow, 20
- valor óptimo, 147
 - de w , 148, 149
- vector
 - diferencia, 32, 33, 166
 - direccional, 168
 - director, 168
 - error absoluto, 54, 58, 125
 - error relativo, 55, 58, 125
 - gradiente, 32, 33, 168
 - número de condición, 55
 - ortogonal, 118, 166
 - perpendicular, 118
 - unitario, 168
 - valor absoluto del gradiente, 54