

Sistemas basados en lógica y en aritmética difusas

Oscar G. Duarte V.

19 de abril de 2025

Índice general

I Lógica difusa	1
1. Elementos de la teoría de conjuntos difusos	3
1.1. Motivación	3
1.2. Conjuntos difusos	5
1.2.1. Pertenencia a conjuntos convencionales	6
1.2.2. Los conjuntos difusos	8
1.2.3. Notación sumatoria (de Zadeh)	9
1.2.4. Conjuntos difusos sobre universos de discurso continuos	10
1.2.5. Para destacar	10
1.3. Algunas definiciones	13
1.3.1. Soporte de un conjunto difuso	13
1.3.2. Núcleo de un conjunto difuso	14
1.3.3. α -corte de un conjunto difuso	16
1.3.4. Producto cartesiano	22
1.3.5. Extensión cilíndrica	24
1.3.6. Proyección cilíndrica	27
2. Operaciones entre conjuntos difusos	31
2.1. Operaciones básicas	31
2.1.1. Operaciones sobre conjuntos concretos	32
2.1.2. Operaciones Difusas	35
2.2. \mathcal{T} -Normas y \mathcal{S} -Normas	36
2.2.1. Operadores de agregación	38
2.2.2. \mathcal{T} -normas	38
2.2.3. \mathcal{S} -normas	41
2.3. Relaciones entre conjuntos	44
2.3.1. Relaciones concretas entre conjuntos concretos	44
2.3.2. Relaciones difusas entre conjuntos concretos	46
2.3.3. Relación entre conjuntos difusos	50

2.3.4. Composición de relaciones difusas	52
3. Principios de lógica difusa	57
3.1. Variables lingüísticas	57
3.1.1. Formas usuales de conjuntos difusos	60
3.1.2. Formas usuales de las variables lingüísticas	64
3.2. Lógica difusa	67
3.2.1. Lógica proposicional concreta	67
3.2.2. Un isomorfismo útil	69
3.2.3. Lógica proposicional difusa	70
4. Sistemas basados de reglas	75
4.1. Estructura de los sistemas basados en reglas	75
4.2. Base de reglas	82
4.2.1. Representación del conocimiento de una regla	84
4.2.2. Propiedades de una base de reglas	88
4.3. Máquina de inferencia	90
4.3.1. Inferencia	92
4.3.2. Inferencia numérica	93
4.3.3. Inferencia difusa	98
4.4. Difusor	104
4.5. Con cresor	104
4.6. Sistemas de múltiples entradas y múltiples salidas	112
4.6.1. Sistemas tipo MISO	114
4.6.2. Sistemas tipo MIMO	119
4.6.3. La maldición de la dimensionalidad	119
5. Anotaciones sobre los SBR	121
5.1. Modificadores lingüísticos	121
5.2. Sobre las implicaciones difusas	124
5.2.1. Implicación concreta	124
5.2.2. Implicaciones difusas	125
5.2.3. Sistemas con inferencias de implicación	126
5.3. Control difuso	129
5.3.1. Rudimentos de control	129
5.3.2. Controlador difuso	132
5.3.3. Unas anotaciones sobre el control difuso	136
5.4. Sistemas tipo Takagi-Sugeno	136
6. Aprendizaje de máquina	139
6.1. Introducción	139

6.2.	Aproximación de funciones con SBR	142
6.2.1.	Aproximadores universales de funciones	143
6.2.2.	SBR como aproximadores universales de funciones	145
6.3.	Construcción de la función semántica a partir de encuestas	146
6.4.	Agrupamiento difuso	151
6.4.1.	El algoritmo <i>Fuzzy c-means</i>	152
6.4.2.	Diseño de la función semántica usando <i>Fuzzy c-means</i>	160
6.5.	Identificación de reglas de SBR	165
6.5.1.	Entrenamiento de universos fijos	167
6.5.2.	Entrenamiento de universos variables	176
6.6.	Optimización mediante heurísticas	180
6.6.1.	Algoritmos genéticos	182
6.6.2.	Optimización de SBR con algoritmos genéticos	187
II	Aritmética difusa	205
7.	Números difusos	207
7.1.	Números difusos	207
7.1.1.	Números trapezoidales y triangulares	209
7.1.2.	α -cortes de número difusos	212
7.2.	Valor representativo y nivel de optimismo	215
7.3.	Ambigüedad	223
7.4.	Distancia entre números difusos	226
7.5.	Tendencias de los parámetros de un número difuso	229
7.6.	Caracterización de números trapezoidales	231
7.7.	Aproximaciones trapezoidales de números difusos	235
7.7.1.	Conjunto de criterios 0	238
7.7.2.	Conjunto de criterios 1	241
7.7.3.	Conjunto de criterios 2	243
7.7.4.	Conjunto de criterios 3	246
7.7.5.	Conjunto de criterios 4	250
8.	Aritmética intervalar	255
8.1.	Definiciones básicas	255
8.2.	Extensión de funciones de una variable	257
8.3.	Extensión de funciones de una variable continuas y monótonas	257
8.4.	Extensión de funciones multivariable	263
8.5.	Monotonía de funciones multivariable	265
8.6.	Extensión de funciones multivariable continuas y monótonas	268
8.7.	Operaciones aritméticas sobre intervalos	272

8.7.1. Suma y resta de intervalos	272
8.7.2. Producto de intervalos	272
8.7.3. Potencia de un intervalo	275
8.8. Algunas precauciones	277
9. El principio de extensión	279
9.1. Extensión de funciones sobre \mathbb{R}	281
9.2. Extensión de funciones sobre \mathbb{R}^n	286
9.3. Extensión mediante α -cortes	289
10. Aritmética difusa	295
10.1. Suma de números difusos	296
10.2. Resta de números difusos	297
10.3. Producto de números difusos	299
10.4. División de números difusos	302
10.5. Aritmética de trapecios positivos	306
11. Algoritmos de extensión de funciones monótonas	311
11.1. Representación discreta de un número difuso	313
11.2. Extensión directa	314
11.3. Extensión inversa posible	319
11.4. Extensión inversa necesaria	328
11.5. Extensiones inversas intermedias	339
11.6. Medida de la existencia de la extensión inversa de una función	344
12. Sistemas basados en aritmética	349
12.1. Estructura	349
12.2. Máquina de inferencia	354
12.3. Función de razonamiento aproximado	355
12.3.1. Galería de fra	356
12.3.2. Funciones <i>ad hoc</i>	368
12.4. Interpretación lingüística	372
12.4.1. Escalamiento	376
12.5. Aproximación lingüística	379
12.5.1. Consistencia entre dos conjuntos difusos	380
12.5.2. Escalamiento	381
12.5.3. Construcción de la salida	382
13. Anotaciones sobre los SBA	389
13.1. Sistemas de razonamiento inverso	389
13.2. Redes de SBA	393

A. Obtención de los pesos de un operador OWA	403
A.1. Orness y entropía	403
A.2. Una familia de operadores OWA	403
A.2.1. Operadores del tipo AND ($0 \leq \nu \leq 0.5$)	404
A.2.2. Selección de $f(\cdot)$	407
A.2.3. Operadores tipo OR ($0.5 \leq \alpha \leq 1$)	408
B. Introducción a los algoritmos genéticos	409
B.1. El algoritmo básico	409
B.1.1. Operadores de Probabilidad	411
B.1.2. Operadores de Selección	412
B.1.3. Operadores de Asignación de Parejas	412
B.1.4. Operadores de Reproducción	412
B.2. Codificación del problema	413
B.2.1. Codificación Binaria	413
B.2.2. Codificación Real	414
B.3. Codificación Híbrida	417
C. Herramientas de Software	419
C.1. UNFuzzy	419
C.2. FuzzyNet	422
C.3. UNGenético	424
C.4. scilab	427
C.5. python	427

Índice de figuras

1.1.	Ejemplo de conjunto convencional	7
1.2.	Ejemplo de conjunto difuso	9
1.3.	Ejemplo de conjuntos difusos sobre la recta real	11
1.4.	Tres conjuntos difusos continuos, y sus soportes	15
1.5.	Tres conjuntos difusos continuos, y sus núcleos	17
1.6.	α -cortes del ejemplo 1.6.	20
1.7.	α -cortes del ejemplo 1.7.	21
1.8.	α -cortes del ejemplo 1.8.	22
1.9.	Producto cartesiano del ejemplo 1.10	23
1.10.	Extensión cilíndrica en universos discretos	25
1.11.	Extensión cilíndrica en universos continuos	26
1.12.	Proyección cilíndrica sobre el universo discreto V	28
1.13.	Proyección cilíndrica sobre el universo discreto U	29
1.14.	Proyección cilíndrica sobre universos continuos	30
2.1.	Operaciones entre conjuntos concretos. Diagramas de Venn.	33
2.2.	Operaciones entre conjuntos difusos.	37
2.3.	Operadores máx y mín en el cubo unitario	39
2.4.	Operador mín	39
2.5.	Ejemplos de \mathcal{T} -Normas	40
2.6.	Ejemplos de \mathcal{T} -normas paramétricas	42
2.7.	Operador máx	43
2.8.	Ejemplos de \mathcal{S} -Normas	44
2.9.	Ejemplos de \mathcal{S} -normas paramétricas	45
2.10.	Relación concreta entre conjuntos concretos	47
2.11.	Relación entre conjuntos difusos	51
2.12.	Composición de relaciones	53
2.13.	Composición de relaciones. Un ejemplo.	54
2.14.	Composición de relaciones. Un cálculo.	55

3.1. Función semántica en el ejemplo 3.1	59
3.2. Tipos de conjuntos difusos definidos como trapecios	62
3.3. Tipos de conjuntos difusos curvos	62
3.4. Transiciones rectas y curvas para funciones de pertenencia	65
3.5. Efecto de los parámetros γ y λ en la curvatura.	65
3.6. Cuatro formas usuales de variables lingüísticas.	66
3.7. Ejemplos de variables lingüísticas que no son particiones	68
3.8. Operaciones básicas equivalentes	70
3.9. Función semántica en el ejemplo 3.2	71
3.10. Grado de verdad difuso	72
3.11. Operación de disyunción en lógica difusa	72
3.12. Operación de conjunción en lógica difusa	73
4.1. Sistemas de cómputo	76
4.2. Tipos de sistemas numéricos según el número de entradas y salidas	77
4.3. Estructura de un Sistema Basado en Reglas	77
4.4. Bloques de un Sistema Basado en Reglas	79
4.5. Variables lingüísticas del ejemplo 4.1	80
4.6. Función de entrada-salida para el ejemplo 4.1	81
4.7. Varias funciones de entrada-salida para el ejemplo 4.1	81
4.8. Efecto de las reglas sobre la relación entrada-salida en el ejemplo 4.3 . .	85
4.9. Variables lingüísticas del ejemplo 4.4	86
4.10. Representación difusa del conocimiento de la base de reglas en el ejemplo 4.4	87
4.11. Cubrimiento de cada una de las reglas	89
4.12. Completez y coherencia de la base de reglas para el ejemplo 4.6	91
4.13. Inferencia numérica con conocimiento sin incertidumbre	96
4.14. Inferencia numérica con conocimiento sin incertidumbre	97
4.15. Inferencia difusa con un hecho sin incertidumbre	101
4.16. Inferencia difusa con un hecho con incertidumbre	102
4.17. Tipos de difusor	105
4.18. Esquema de agregación-concreción	107
4.19. Tipos de funciones de concreción.	109
4.20. Esquema de concreción-agregación. Concesor tipo Altura.	110
4.21. Relación Entrada-Salida del ejemplo 4.9	111
4.22. Cálculo paso a paso. Implicación tipo Mamdani	113
4.23. Sistema de múltiples entradas y una salida	114
4.24. SBR del ejemplo 4.10	116
4.25. Relación Entrada-Salida del ejemplo 4.10	117
4.26. Relación Entrada-Salida del ejemplo 4.11	117
4.27. Difusor para múltiples entradas	118

4.28. SBR tipo MIMO con sistemas MISO en paralelo	119
4.29. SBR tipo MIMO con inferencia común	120
5.1. Efecto de los modificadores lingüísticos	122
5.2. Efecto de los modificadores lingüísticos en el ejemplo 5.1. En azul sin modificadores y en rojo con modificadores.	123
5.3. Relación Entrada-Salida del ejemplo 5.2	128
5.4. Cálculo paso a paso. Implicación tipo Lukasiewicz	130
5.5. Planta a ser controlada	131
5.6. Retroalimentación de una planta	132
5.7. Controlador difuso	133
5.8. Estructuras usuales de controladores difusos	133
5.9. Planta y simulación para el ejemplo 5.3	134
5.10. Controlador Difuso del ejemplo 5.3	135
5.11. Planta y simulación para el ejemplo 5.3	135
5.12. Inferencia en el ejemplo 5.4	138
6.1. Esquemas del aprendizaje de máquina	141
6.2. Sistema basado en reglas	142
6.3. Aproximador de una función	143
6.4. Probabilidad acumulada en el ejemplo 6.2	147
6.5. Conjunto difuso asociado al término ‘Adulto mayor’ en el ejemplo 6.2 . .	148
6.6. Curva $R(\alpha)$ del término ‘Adulto’ en el ejemplo 6.2	149
6.7. Función semántica del ejemplo 6.2	149
6.8. Función semántica curva del ejemplo 6.2	150
6.9. El proceso de agrupamiento	152
6.10. Método de Picard aplicado el algoritmo <i>Fuzzy c-means</i>	155
6.11. Diagrama de flujo del algoritmo <i>Fuzzy c-means</i>	156
6.12. Grupos obtenidos en el Ejemplo 6.4	156
6.13. Evolución del algoritmo <i>Fuzzy c-means</i>	157
6.14. Efecto del parámetro m del algoritmo <i>Fuzzy c-means</i>	158
6.15. Puntos y grupos obtenidos en el ejemplo 6.7	158
6.16. Grupos individuales obtenidos en el ejemplo 6.7	159
6.17. Puntos en el ejemplo 6.8	159
6.18. Grupos obtenidos en el ejemplo 6.8	160
6.19. Curvas de nivel de los grupos obtenidos en el ejemplo 6.8	160
6.20.	162
6.21. Datos para el ejemplo 6.9	163
6.22. Variables simétricas en el ejemplo 6.9	163
6.23. Agrupamiento <i>Fuzzy c-means</i> en el ejemplo 6.9	164
6.24. Estrategia FCM en el ejemplo 6.9	164

6.25. Estrategia del conjunto central en el ejemplo 6.9	164
6.26. Modelos en el ejemplo 6.9	165
6.27. Algoritmos de identificación de reglas	166
6.28. Entrenamiento de universos fijos	169
6.29. Función de prueba para entrenamiento de SBR	171
6.30. Variables lingüísticas en el ejemplo 6.10	171
6.31. Aproximación obtenida en el ejemplo 6.10	172
6.32. Desempeño de los SBR del Ejemplo 6.11	173
6.33. Aproximaciones en el ejemplo 6.12 con 3 conjuntos	174
6.34. Aproximaciones en el ejemplo 6.12 con 5 conjuntos	175
6.35. Aproximaciones en el ejemplo 6.12 con 7 conjuntos	175
6.36. Aproximaciones en el ejemplo 6.12 con 9 conjuntos	175
6.37. Aproximaciones en el ejemplo 6.12 con 11 conjuntos	175
6.38. Aproximaciones en el ejemplo 6.12 con 13 conjuntos	176
6.39. Aproximaciones en el ejemplo 6.12 con 15 conjuntos	176
6.40. Entrenamiento de universos variables	178
6.41. Aproximaciones en el ejemplo 6.13 con ancho 0.1	179
6.42. Aproximaciones en el ejemplo 6.13 con ancho 0.2	180
6.43. Aproximaciones en el ejemplo 6.13 con ancho 0.3	180
6.44. Heurísticas	181
6.45. Algoritmo genético básico	183
6.46. Función a optimizar en el ejemplo 6.14	184
6.47. Poblaciones de algunas generaciones en el ejemplo 6.14. En rojo el punto solución.	185
6.48. Evolución de la optimización en el ejemplo 6.14	186
6.49. Histograma del ejemplo 6.15	187
6.50. Heurísticas aplicadas al diseño de SBR	189
6.51. Optimización y función objetivo en el ejemplo 6.16	191
6.52. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 3 conjuntos	192
6.53. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 5 conjuntos	192
6.54. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 7 conjuntos	193
6.55. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 9 conjuntos	193
6.56. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 11 conjuntos	193
6.57. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 13 conjuntos	194
6.58. Optimizaciones y funciones objetivo en el ejemplo 6.17 con 15 conjuntos	194
6.59. variable lingüística ‘Estatura’	195
6.60. Variables sin sentido lingüístico	196
6.61. Partición parametrizada	197
6.62. Optimizaciones y funciones objetivo en el ejemplo 6.18 con 3 conjuntos	197
6.63. Optimizaciones y funciones objetivo en el ejemplo 6.18 con 5 conjuntos	198
6.64. Optimizaciones y funciones objetivo en el ejemplo 6.18 con 7 conjuntos	198

6.65. Optimizaciones y funciones objetivo en el ejemplo 6.18 con 9 conjuntos	198
6.66. Optimizaciones y funciones objetivo en el ejemplo 6.18 con 11 conjuntos	199
6.67. Optimizaciones y funciones objetivo en el ejemplo 6.18 con 13 conjuntos	199
6.68. Funciones semánticas en el ejemplo 6.18	200
6.69. Optimizaciones y funciones objetivo en el ejemplo 6.19 con 3 conjuntos	201
6.70. Optimizaciones y funciones objetivo en el ejemplo 6.19 con 5 conjuntos	201
6.71. Optimizaciones y funciones objetivo en el ejemplo 6.19 con 7 conjuntos	201
6.72. Optimizaciones y funciones objetivo en el ejemplo 6.19 con 9 conjuntos	202
6.73. Optimizaciones y funciones objetivo en el ejemplo 6.19 con 11 conjuntos	202
6.74. Optimizaciones y funciones objetivo en el ejemplo 6.19 con 13 conjuntos	202
6.75. Variables lingüísticas en el ejemplo 6.19	203
7.1. Condiciones de un número difuso	210
7.2. Número difusos trapezoidales y triangulares	211
7.3. Números difusos para ordenar	215
7.4. Ejemplo de funciones de ponderación $h(\alpha) = \alpha^\gamma$	216
7.5. Valor representativo del ejemplo 7.1	218
7.6. Números y valor representativo del ejemplo 7.2	220
7.7. Número difuso del ejemplo 7.3	222
7.8. Número difuso del ejemplo 7.3	222
7.9. Número difuso del ejemplo 7.3	225
7.10. Valor representativo en el ejemplo 7.8	234
7.11. Ambigüedad en el ejemplo 7.8	235
7.12. Aproximaciones trapezoidales del ejemplo 7.9	253
8.1. Extensión de funciones de una variable a intervalos	258
8.2. Monotonía de una función	259
8.3. Función del ejemplo 8.1	261
8.4. Función del ejemplo 8.2	262
8.5. Función del ejemplo 8.3	263
8.6. Función del ejemplo 8.4	264
8.7. Función del ejemplo 8.5	265
8.8. Producto intervalar. Visualización de los casos de la Tabla 8.1	274
9.1. Función del ejemplo 9.1	280
9.2. Problema del ejemplo 9.2	282
9.3. Grado de pertenencia para $y = 0.75$ en el ejemplo 9.2	283
9.4. Respuesta del ejemplo 9.2	283
9.5. Problema del ejemplo 9.3	284
9.6. Grado de pertenencia para $y = 1$ en el ejemplo 9.3	285
9.7. Respuesta del ejemplo 9.3	285

9.8. Problema del ejemplo 9.4	287
9.9. Parejas de puntos que cumplen $x_1 + x_2 = 4$	287
9.10. Funciones de pertenencia parametrizadas en el ejemplo 9.4	288
9.11. Mínimo de funciones de pertenencia parametrizadas en el ejemplo 9.4	288
9.12. Resultado del ejemplo 9.4	289
9.13. Solución del ejemplo 9.5	291
9.14. Solución del ejemplo 9.5	292
9.15. Solución del ejemplo 9.7	293
10.1. Suma de números difusos	297
10.2. Resta de números difusos	299
10.3. Producto de números difusos	301
10.4. Producto de números difusos positivos	302
10.5. División de números difusos	304
10.6. División de números difusos positivos	305
10.7. Comparación de resultados en el ejemplo 10.7	308
10.8. Comparación de resultados en el ejemplo 10.8	309
11.1. Números del ejemplo 11.3	316
11.2. Función del ejemplo 11.4	317
11.3. Números del ejemplo 11.4	318
11.4. Función del ejemplo 11.5	318
11.5. Números del ejemplo 11.4	319
11.6. Números del ejemplo 11.7	324
11.7. Solución del ejemplo 11.8	326
11.8. Solución del ejemplo 11.9	327
11.9. Números del ejemplo 11.12	335
11.10. Solución del ejemplo 11.13	337
11.11. Solución del ejemplo 11.14	338
11.12. Números del ejemplo 11.15	342
11.13. Números del ejemplo 11.16	343
11.14. Números del ejemplo 11.17	343
11.15. Flujo de dinero en el ejemplo 11.21	347
11.16. TIR posible en el ejemplo 11.21	347
11.17. TIR necesaria en el ejemplo 11.21	348
11.18. VPN necesaria en el ejemplo 11.21	348
12.1. Estructura interna de un SBA	350
12.2. Bloques de un Sistema Basado en Aritmética	351
12.3. Sistema MIMO del ejemplo 12.1	352
12.4. Variables lingüísticas del ejemplo 12.1	353

12.5. Respuesta del sistema del ejemplo 12.1	353
12.6. Función logística parametrizada	361
12.7. Galería de fra (1/3)	365
12.8. Galería de fra (2/3)	366
12.9. Galería de fra (3/3)	367
12.10 Interpretación lingüística en el ejemplo 12.5	379
12.11 Consistencia entre conjuntos difusos	381
12.12 Aproximación lingüística en el ejemplo 12.7	386
12.13 Sensibilidad a β en el ejemplo 12.7	387
13.1. Estructura de un SRI	391
13.2. Variables en el ejemplo 13.1	392
13.4. Red de 3 SBA	393
13.3. Variables en el ejemplo 13.1	393
13.5. Red general de SBA	394
13.6. Red para la estimación del impacto ambiental de vertederos	395
13.7. Red para la estimación de la probabilidad de contaminación a la atmósfera en vertederos	396
13.8. Red para la estimación de la probabilidad de contaminación al suelo en vertederos	396
A.1. Pesos para 3 OWA con $p = 5$	404
A.2. Función $f(\nu) = 1 - (2\nu - 1)^\beta$ y sus restricciones	408
C.1. Galería de capturas de pantalla de UNFuzzy	421
C.2. Galería de capturas de pantalla de FuzzyNet	425
C.3. Uso de UNGenético	426
C.4. Herencia múltiple	427

Índice de tablas

1.1.	Comparación entre los conjuntos concretos y los conjuntos difusos	11
2.1.	Operaciones entre conjuntos concretos. Grados de pertenencia.	34
2.2.	Tabla de pertenencia a las tres funciones básicas	35
2.3.	Producto de grados de pertenencia	35
2.4.	Definiciones de algunas \mathcal{T} -normas y \mathcal{S} -normas	41
2.5.	Relación concreta entre conjuntos concretos	46
2.6.	Relación difusa entre conjuntos concretos	48
2.7.	Relación difusa entre conjuntos concretos	49
3.1.	Tipos de conjuntos difusos rectos	61
3.2.	Vértices de conjuntos igualmente espaciados. δ se calcula según la ecuación 3.11	66
3.3.	Vértices de conjuntos igualmente espaciados. δ se calcula según la ecuación 3.12	66
3.4.	Isomorfismo entre tres campos de la matemática	69
4.1.	Base de reglas en el ejemplo 4.2.	83
4.2.	Tautología del <i>modus ponens</i>	94
4.3.	Funciones de concreción	108
4.4.	Base de reglas del ejemplo 4.10	116
4.5.	Base de reglas del ejemplo 4.11	117
4.6.	Explosión de la base de reglas	120
5.1.	Base de reglas en el ejemplo 5.1	123
5.2.	Función de verdad del operador de implicación	124
5.3.	Primera tautología del operador de implicación	125
5.4.	Segunda tautología del operador de implicación	125
5.5.	Algunos operadores de implicación difusos	126
5.6.	Base de reglas del ejemplo 5.3	135
5.7.	Base de reglas del ejemplo 5.4	138

6.1. Respuestas en el ejemplo 6.2	147
6.2. Probabilidad acumulada en el ejemplo 6.2	147
6.3. Errores en el ejemplo 6.9	165
6.4. Reglas obtenidas en el ejemplo 6.10	172
6.5. Errores en el ejemplo 6.12	174
6.6. Errores en el ejemplo 6.13	179
6.7. Mejor individuo en el ejemplo 6.16	191
7.1. α -cortes de los números difusos más usuales	214
7.2. Casos de la aproximación trapezoidal que minimiza la distancia	236
7.3. Discriminantes de aproximación trapezoidal que minimiza la distancia	237
7.4. Matrices para la aproximación a números trapezoidales	239
7.5. Conjuntos de criterios de semejanza geométrica entre X_k y A	240
7.6. Vértices de las aproximaciones trapezoidales del ejemplo 7.9	253
7.7. Propiedades de las aproximaciones trapezoidales del ejemplo 7.9	254
8.1. Producto intervalar. Casos basados en los límites	273
8.2. División intervalar. Casos basados en los límites	275
8.3. Potenciación intervalar. Casos basados en los límites	276
10.1. Aritmética de trapecios positivos	306
11.1. Números del ejemplo 11.3	317
11.2. Sentido de las monotonías de funciones inversas	322
11.3. Números del ejemplo 11.7	325
11.4. Números del ejemplo 11.12	336
11.5. Diferencia en el cálculo de las extensiones inversas posible y necesaria	339
12.1. Comparación de fra	363
12.2. Entradas válidas al bloque de interpretación lingüística y su salida	377
12.3. Salidas disponibles del bloque de aproximación lingüística	385
12.4. Salidas disponibles del ejemplo 12.7	386
B.1. Operadores de los AG	411

Listado de algoritmos

1.	Diseño de una función semántica a partir de <i>Fuzzy c-means</i>	162
2.	Universos Fijos	170
3.	Universos Variables	177
4.	Trapecio equivalente por valores y ambigüedades	244
5.	Trapecio equivalente por valores extremos	247
6.	Trapecio equivalente por tendencias del Valor central y la Ambigüedad .	250
7.	Trapecio equivalente por rampas laterales	252
8.	Extensión directa	315
9.	Extensión inversa posible	323
10.	Extensión inversa necesaria	333
11.	Extensión inversa intermedia	340
12.	Medida de la existencia de la extensión inversa	345
13.	Pesos de un operador OWA	373

Listado de ejemplos

Ejemplo 1.1. Soporte de un conjunto difuso discreto	13
Ejemplo 1.2. Soporte de tres conjuntos difusos continuos	13
Ejemplo 1.3. Nucleo de un conjunto difuso discreto	14
Ejemplo 1.4. Núcleo de tres conjuntos difusos continuos	16
Ejemplo 1.5. α -cortes de un conjunto discreto	16
Ejemplo 1.6. α -cortes de un conjunto continuo	18
Ejemplo 1.7. α -corte de un conjunto no convexo	19
Ejemplo 1.8. α -corte de un conjunto discontinuo	19
Ejemplo 1.9. Producto cartesiano de dos conjuntos concretos discretos	22
Ejemplo 1.10. Producto cartesiano de dos conjuntos concretos continuos	23
Ejemplo 1.11. Extensión cilíndrica sobre universos discretos	24
Ejemplo 1.12. Extensión cilíndrica sobre universos continuos	26
Ejemplo 1.13. Proyección cilíndrica sobre universos discretos	27
Ejemplo 1.14. Proyección cilíndrica sobre universos continuos	27
Ejemplo 3.1. Variable lingüística	58
Ejemplo 3.2. Grado de verdad difuso	71
Ejemplo 3.3. Operación de disyunción en lógica difusa	72
Ejemplo 3.4. Operación de conjunción en lógica difusa	73
Ejemplo 4.1. Sistema de lógica difusa tipo SISO	80
Ejemplo 4.2. Base de reglas	82

Ejemplo 4.3. Efecto de la base de reglas	83
Ejemplo 4.4. El conocimiento en las reglas	86
Ejemplo 4.5. Cubrimiento de una regla	88
Ejemplo 4.6. Completez y coherencia de una base de reglas	90
Ejemplo 4.7. <i>Modus Ponens</i>	92
Ejemplo 4.8. <i>Modus Ponens</i> generalizado	98
Ejemplo 4.9. Sistema tipo Mamdani	111
Ejemplo 4.10. SBR tipo MISO	115
Ejemplo 4.11. SBR tipo MISO	117
Ejemplo 5.1. Modificadores lingüísticos	123
Ejemplo 5.2. Sistema con implicación de Lukasiewicz	127
Ejemplo 5.3. Controlador difuso	134
Ejemplo 5.4. Sistema Takagi-Sugeno de 1 entrada	137
Ejemplo 6.1. Teorema de Weierstrass	144
Ejemplo 6.2. Construcción de un conjunto difusa a partir de encuestas	146
Ejemplo 6.3. Variación sobre el ejemplo 6.2	150
Ejemplo 6.4. <i>Fuzzy c-means</i> aplicado a una distribución homogénea en \mathbb{R}	156
Ejemplo 6.5. Convergencia del algoritmo <i>Fuzzy c-means</i>	157
Ejemplo 6.6. Efecto del parámetro m del algoritmo <i>Fuzzy c-means</i>	157
Ejemplo 6.7. <i>Fuzzy c-means</i> aplicado a una distribución homogénea en \mathbb{R}^2	158
Ejemplo 6.8. <i>Fuzzy c-means</i> aplicado a una distribución no homogénea en \mathbb{R}^2	159
Ejemplo 6.9. Diseño de variables lingüísticas con <i>Fuzzy c-means</i>	161
Ejemplo 6.10. Entrenamiento de Universos Fijos	171
Ejemplo 6.11. Entrenamiento de Universos Fijos. Número de etiquetas	173
Ejemplo 6.12. Entrenamiento de Universos Fijos. Tipo de conjuntos	173
Ejemplo 6.13. Entrenamiento de Universos Variables	179
Ejemplo 6.14. Búsqueda mediante algoritmos genéticos	184

Ejemplo 6.15. Aleatoriedad de los algoritmos genéticos	186
Ejemplo 6.16. Optimización de modificadores lingüísticos mediante AG	190
Ejemplo 6.17. Optimizaciones de modificadores lingüísticos mediante AG	192
Ejemplo 6.18. Optimización de la variable de entrada mediante AG	197
Ejemplo 6.19. Optimización de dos variables mediante AG	200
Ejemplo 7.1. Valor representativo de números difusos	217
Ejemplo 7.2. Interpretación del nivel de optimismo	219
Ejemplo 7.3. Valor representativo de un número difuso	221
Ejemplo 7.4. Ambigüedad de números difusos	224
Ejemplo 7.5. Ambigüedad de un número difuso	224
Ejemplo 7.6. Error en la aritmética de números positivos	226
Ejemplo 7.7. Distancia entre dos números difusos	227
Ejemplo 7.8. Valor representativo y ambigüedad de números trapezoidales	234
Ejemplo 7.9. Aproximaciones trapezoidales de un número difuso	252
Ejemplo 8.1. Extensión intervalar de e^x	260
Ejemplo 8.2. Extensión intervalar de e^{-x}	262
Ejemplo 8.3. Extensión intervalar de $\sin(x)$	263
Ejemplo 8.4. Extensión de funciones multivariable. Suma	264
Ejemplo 8.5. Extensión de funciones multivariable. Producto	264
Ejemplo 8.6. Monotonía de un hiperplano	267
Ejemplo 8.7. Monotonía del producto	267
Ejemplo 8.8. Extensión de funciones multivariable monótonas. Suma	269
Ejemplo 8.9. Extensión de funciones multivariable monótonas. Resta	270
Ejemplo 8.10. Extensión de funciones multivariable monótonas. Hiperplano	271
Ejemplo 9.1. Extensión de una función sobre un universo discreto	280
Ejemplo 9.2. Principio de extensión de una función sobre \mathbb{R} uno a uno	282
Ejemplo 9.3. Principio de extensión de una función sobre \mathbb{R} no uno a uno	284

Ejemplo 9.4. Principio de extensión de una función sobre \mathbb{R}^2	286
Ejemplo 9.5. Extensión mediante α -cortes de una función uno a uno	290
Ejemplo 9.6. Extensión mediante α -cortes de una función sobre \mathbb{R}	291
Ejemplo 9.7. Extensión mediante α -cortes de una función sobre \mathbb{R}^2	292
Ejemplo 10.1. Suma de números difusos	296
Ejemplo 10.2. Resta de números difusos	298
Ejemplo 10.3. Producto de números difusos	300
Ejemplo 10.4. Producto de números difusos positivos	301
Ejemplo 10.5. División de números difusos	303
Ejemplo 10.6. División de números difusos positivos	305
Ejemplo 10.7. Aritmética de números trapezoidales positivos. Producto	307
Ejemplo 10.8. Aritmética de números trapezoidales positivos. División	308
Ejemplo 11.1. Funciones a extender	312
Ejemplo 11.2. Representación discreta de un número difuso	314
Ejemplo 11.3. Extensión directa de un hiperplano	316
Ejemplo 11.4. Extensión directa de una función positiva	317
Ejemplo 11.5. Extensión directa de una función acotada	318
Ejemplo 11.6. Extensión inversa posible en un intervalo	320
Ejemplo 11.7. Extensión posible de un hiperplano	324
Ejemplo 11.8. Extensión posible de una función positiva	325
Ejemplo 11.9. Extensión posible de una función acotada	326
Ejemplo 11.10Extensión inversa necesaria en un intervalo	328
Ejemplo 11.11Extensión inversa necesaria en un intervalo sin solución	328
Ejemplo 11.12Extensión necesaria de un hiperplano	334
Ejemplo 11.13Extensión necesaria de una función positiva	336
Ejemplo 11.14Extensión necesaria de una función acotada	337
Ejemplo 11.15Extensiones inversas intermedias de un hiperplano	341

Ejemplo 11.16 Extensiones inversas intermedias de una función positiva	342
Ejemplo 11.17 Extensiones inversas intermedias de una función acotada	343
Ejemplo 11.18 Existencia de la extensión inversa de un hiperplano	344
Ejemplo 11.19 Existencia de la extensión inversa de una función positiva	345
Ejemplo 11.20 Existencia de la extensión inversa de una función acotada	346
Ejemplo 11.21 Tasa interna de retorno difusa	346
Ejemplo 12.1. Análisis de riesgo de daño causado por rayos	352
Ejemplo 12.2. $\text{fra}(\cdot)$ para el análisis de riesgo de daño causado por rayos	364
Ejemplo 12.3. $\text{fra}(\cdot)$ <i>ad hoc.</i> Valor ambiental del suelo.	369
Ejemplo 12.4. Operador OWA	371
Ejemplo 12.5. Interpretación lingüística	378
Ejemplo 12.6. Consistencia entre conjuntos difusos	380
Ejemplo 12.7. Aproximación lingüística	384
Ejemplo 13.1. Razonamiento inverso en 2D	391
Ejemplo 13.2. Red de SBA	394

Prólogo

La necesidad de tender puentes entre números y palabras ha llevado a la construcción de modelos matemáticos del lenguaje y del razonamiento. Los modelos extensos de lenguaje o LLM (del inglés *Large Language Models*) son tan sólo el más reciente intento, uno muy exitoso, de sistematizar la generación del lenguaje escrito. Sin embargo, podemos encontrar ejemplos relativamente recientes en la historia como ELIZA¹ o remontarnos mucho más atrás, hasta la lógica aristotélica, para argumentar que la formalización del lenguaje ha sido un área de estudio fructífera y abundante.

Los sistemas difusos pueden enmarcarse en esa tradición. Como se explica en este texto, toman algo de aritmética básica y algo de lógica elemental para ofrecer soluciones prácticas a problemas específicos que involucran simultáneamente números y palabras. Quiero insistir en lo anterior: el propósito de los sistemas difusos es la solución práctica de problemas. Eso no es de extrañar, ya que nacen en la ingeniería y para la ingeniería, aunque luego su formalización y aplicación se han extendido a otra áreas del conocimiento.

La gran diferencia entre el enfoque de los sistemas difusos y los LLM es el lugar en que se ubican en la conocida dicotomía interpretabilidad vs precisión. Los sistemas difusos son altamente interpretables porque su estructura interna es simple; podemos explicar porqué generan los resultados obtenidos. Por el contrario, los LLM son sofisticados sistemas con millones de parámetros ajustables capaces de generar magníficos resultados, pero no podemos explicar porqué se obtiene un resultado específico.

Cada herramienta tiene su nicho específico de aplicación. Hay que evitar la tentación de emplear sólo una de ellas como si fuese la solución universal. Por ello vale la pena insistir en el poder de las soluciones que ofrecen simplicidad y elegancia en los procedimientos, aún a costa de la precisión de sus resultados.

Este texto se ha diseñado teniendo en mente dos tipos de sistemas difusos: aquellos

¹ELIZA es un programa desarrollado en la década de los 60, con capacidades de procesamiento de lenguaje natural. Capacidades limitadas para los estándares de hoy en día, pero que todavía nos sorprenden por la simplicidad y efectividad de su enfoque.

basados en lógica difusa y aquellos basados en aritmética difusa. Se trata de un texto con vocación didáctica. Por ello, abarca desde los principios básicos hasta las aplicaciones sencillas. Esa misma orientación didáctica ha guiado el diseño del texto y de cada uno de sus capítulos. El resultado ha sido un documento con 244 figuras, 53 tablas y 116 ejemplos, es decir, es un texto en el que no se han escatimado esfuerzos para hacer que las explicaciones sean claras y simples. Por ello, se ha utilizado un estilo de redacción que combina el rigor matemático con claridad en la presentación de definiciones, conceptos, procedimientos y aplicaciones.

De los dos tipos de sistemas, indudablemente los basados en lógica son más conocidos que los basados en aritmética. Existen varios textos disponibles dedicados a los sistemas basados en lógica difusa la mayoría de los cuales están orientados a su aplicación en el control de procesos.

Los sistemas basados en aritmética son menos populares. Buena parte del contenido de este texto que se dedica a ellos es la recopilación de mis trabajos de investigación en ese tema. En ese sentido, la segunda parte de este texto marca una gran diferencia con otras obras didácticas sobre teoría y aplicación de lógica difusa.

El texto se ha organizado en 13 capítulos y 3 apéndices. Los capítulos están agrupados en dos partes de tamaño semejante. La primera parte se ocupa de los sistemas basados en lógica difusa y la segunda de los basados en aritmética difusa.

La primera parte consta de 6 capítulos. En el capítulo 1 se presentan los principios de la teoría de conjuntos difusos y algunas definiciones básicas. El capítulo 2 se dedica a las operaciones básicas entre conjuntos difusos. La lógica difusa hace su aparición en el capítulo 3, que desemboca en la presentación de los sistemas difusos basados que hacen uso de ella en el capítulo 4. Algunas observaciones importantes sobre este tipo de sistemas se efectúan en el capítulo 5. La primera parte termina con una explicación sobre el uso de aprendizaje de máquina para el diseño y optimización de sistemas difusos en el capítulo 6.

La segunda parte contiene 7 capítulos. Para abordar el tema de aritmética difusa es necesario comenzar por la definición de número difuso. Eso se hace en el capítulo 7, que se aprovecha para presentar propiedades y aproximaciones de los números difusos. Debido a que los números difusos pueden entenderse como generalizaciones de la noción de intervalo, es conveniente estudiar la aritmética de intervalos antes de dar el paso a la aritmética difusa, y a ello se destina el capítulo 8. La aritmética difusa también puede entenderse como un conjunto de funciones que operan sobre números difusos y una forma de definir tales funciones es a través del principio de extensión, que no es otra cosa que un mecanismo para extender la aplicación de funciones definidas sobre conjuntos convencionales a conjuntos difusos; el capítulo 9 se dedica a dicho principio. A la aritmética difusa se llega en el capítulo 10 en el que se presentan las cuatro operaciones

básicas de suma, resta, multiplicación y división sobre números difusos. La aplicación de los principios anteriores se puede sistematizar a través de unos algoritmos que se desarrollan en el capítulo 11. Esos algoritmos están en el corazón de los sistemas basados en aritmética difusa que se presentan en los capítulos 12 y 13.

La redacción de este texto ha sido para mi un viaje a través de mi historia personal. Cada párrafo y cada ecuación me ha recordado algún amigo de tantos que me ha dejado la academia. Espero que disfruten tanto su lectura como yo lo he hecho con su escritura.

Oscar Duarte

Parte I

Lógica difusa

Capítulo 1

Elementos de la teoría de conjuntos difusos

1.1. Motivación

Las matemáticas nos ha permitido construir representaciones de la realidad, así como nuevas realidades. No en vano suele decirse que las son el lenguaje de la ciencia. El lenguaje matemático nos ayuda a expresar el conocimiento científico y, quizás más importante, a establecer los límites de ese conocimiento, su rango de validez.

Contrario a lo que puede pensarse, la universalidad del conocimiento científico está acotada. La ley de gravitación universal de Newton, por ejemplo, fue *universal* hasta que la evidencia demostró las ideas de Einstein. Hoy en día, esa ley sigue siendo válida y universal en lo que podríamos llamar ‘su propio universo’, que son todos aquellos fenómenos que estudiamos y analizamos en los que las predicciones newtonianas son suficientemente adecuadas para nuestros propósitos.

Aunque sabemos bien que el modelo relativista de la gravedad es más completo y preciso, renunciamos a él para estudiar la órbita de un satélite artificial. Es más, renunciamos también a la formulación newtoniana cuando estudiamos la caída de un cuerpo como si la aceleración de la gravedad fuese constante.

Lo hacemos, por supuesto, por simplicidad. Pero para hacerlo debemos analizar el efecto de esa simplificación en la precisión de nuestras estimaciones y predicciones. Sabemos también que todo experimento científico, y por tanto toda conclusión científica, tiene una precisión limitada a las capacidades tecnológicas de los instrumentos de precisión¹.

¹Además, si se trata de fenómenos cuánticos, la precisión también está limitada por el principio de incertidumbre de Heisenberg.

Por esas razones, no es temerario afirmar que la ciencia es el estudio de la incertidumbre de nuestro conocimiento. La cuantificación de la incertidumbre es la clave del conocimiento científico.

Es por ello que la estadística y la probabilidad son fundamentales en el desarrollo de la ciencia. Gracias a ellas tenemos herramientas matemáticas para entender y cuantificar la incertidumbre.

Sin embargo, hay muchas situaciones en las que el enfoque probabilístico puede no ser el más adecuado. La información que tenemos a nuestra disposición no es perfecta y por ello hay muchas causas potenciales de errores en nuestro modelos. Destaquemos aquí algunos tipos de imperfecciones en la información :

- Información cuantitativa (numérica) tan escasa que no es posible construir una distribución de probabilidad confiable.
- Información expresada en forma lingüística, usualmente con cuantificadores tales como 'poco' o 'mucho'.
- Información que resulta de aplicar modelos que generan un gran margen de error a partir de la información disponible.

Zadeh expresó de forma elegante estas ideas en 1973 bajo el nombre del *principio de incompatibilidad* (véase [57]):

Principio de incompatibilidad: Al incrementarse la complejidad de un problema, nuestra capacidad para producir afirmaciones categóricas y precisas, disminuye. Y el sentido y significado de tales afirmaciones también disminuye.

Para explicar el mensaje del principio de incompatibilidad de Zadeh, consideremos una obra de arte. La *Monna Lisa* de Leonardo da Vinci, por ejemplo. Podemos hacer afirmaciones precisas y categóricas sobre algunas de las propiedades de ese cuadro. Por ejemplo, sobre sus dimensiones. Podemos medir y establecer cuántos centímetros tiene de largo y de ancho y además podemos determinar la incertidumbre de esa medición. Eso es preciso. Podemos hacerlo porque las dimensiones, largo y ancho, son conceptos bastante sencillos.

Ahora, pensemos en el concepto de belleza. ¿Qué tan bello es el cuadro de la *Monna Lisa*? Nos enfrentamos a un fenómeno muchísimo más complejo que el de las meras dimensiones. Pretender hacer una afirmación categórica sobre la belleza de ese cuadro es mucho más complicado y carece de sentido. Supongamos, por ejemplo, que le asignamos un número de belleza. 'De 0 a 10, la *Monna Lisa* para el crítico de arte número tres es 8.35 bella'. Esa es una afirmación categórica y precisa pero carente de sentido. Ese número no significa nada ¿por qué? Porque la belleza es un fenómeno muy complejo.

Conforme crece la complejidad del fenómeno, disminuye nuestra capacidad de hacer afirmaciones categóricas. Y si las hacemos, tales afirmaciones carecen de sentido.

El principio de incompatibilidad de Zadeh tiene consecuencias importantes en la toma de decisiones sobre sistemas complejos. A manera de ejemplo, consideremos la evaluación de impacto ambiental. No conocemos cómo son todas las relaciones entre los fenómenos naturales en un ecosistema. Eso es tremadamente complejo. Y tenemos información parcial y limitada, tanto del estado actual de un ecosistema como de las actuaciones futuras que haremos en un determinado sistema. Por eso, tomar decisiones basados en evaluaciones de impacto ambiental implica hacer aseveraciones categóricas sobre un sistema complejo.

El principio de incompatibilidad nace de nuestras limitaciones para modelar la imperfección de la información y la forma en que esta se propaga cuando crece la complejidad de un problema. Esas dificultades se hacen evidentes a la hora de tomar decisiones. Enumeremos algunas de esas dificultades:

1. La imprecisión de las estimaciones numéricas.
 2. La vaguedad de la información lingüística .
 3. La subjetividad en la interpretación de la información.
 4. La necesidad de combinar información cuantitativa y cualitativa.
 5. La actitud hacia el riesgo en la toma de decisiones.
- .

Este es un texto dedicado a los sistemas difusos. Vamos a mostrar que los sistemas difusos nos permiten modelar problemas presentes a la hora de tomar decisiones. Nos referimos a los problemas de incertidumbre, de ambigüedad, de subjetividad. Utilizaremos estas técnicas para lograr modelar esas dificultades e incorporarlas a nuestro proceso de toma de decisiones. La base de los Sistemas Difusos es la Teoría de Conjuntos Difusos, originalmente propuesta por Zadeh en 1965 (Véase [56]). Es una idea simple, pero poderosa, que se presenta en las secciones 1.2 y 1.3.

1.2. Conjuntos difusos

El concepto de *Conjunto difuso* es la base de los sistemas difusos así como de una gran colección de conceptos que llevan el adjetivo ‘difuso’ (aritmética difusa, lógica difusa, geometría difusa, por ejemplo). En esta sección se introduce la definición de conjunto difuso como el resultado de relajar una condición de los conjuntos convencionales.

En ese sentido, podría pensarse que los conjuntos difusos son una extensión de los conjuntos convencionales o, lo que es igual, que los conjuntos convencionales son un tipo especial de conjunto difuso. No obstante, como se verá en la sección 1.2.5, la situación es justamente al revés: los conjuntos difusos son un tipo especial de conjuntos convencionales.

Para abordar esta aparente paradoja, antes de presentar la definición de conjunto difuso (sección 1.2.2), conviene recordar algunos conceptos sencillos de la teoría de conjuntos convencional (sección 1.2.1).

1.2.1. Pertenencia a conjuntos convencionales

Supongamos que tenemos la tarea de construir tres conjuntos: a) el conjunto de los países de América, b) el conjunto de los dígitos y c) el conjunto de las vocales en castellano. Es clara la tarea que tenemos. No hay dudas a la hora de intentar construir esa colección de elementos que son las vocales en castellano, por ejemplo.

Pero hay otras situaciones en donde no es tan obvio decidir cómo está conformado el conjunto que hay que crear. Consideremos dos conjuntos difíciles de construir:

1. El conjunto de los adolescentes. ¿Cuándo se es adolescente y cuándo no? ¿Hay un momento mágico de entrada a ese conjunto y un momento de salida? ¿Hay alguna condición completamente definida que nos permita decir que una persona es o no es, adolescente?
2. El conjunto de los latinos. ¿Cuáles son las condiciones socioculturales para hacer esa definición? ¿Tiene que ver sólo con el lugar de nacimiento? Probablemente, no. Hay muchas otras condiciones culturales, que hacen que sea muy difícil establecer si una persona pertenece o no directamente a ese conjunto.

¿Qué es un conjunto, según la Real Academia de la Lengua Española? La acepción que nos sirve desde el punto de vista matemático es la que define un conjunto como “una colección de elementos que tienen una propiedad en común”. Sin embargo, hay situaciones en las que se puede tener esa propiedad *parcialmente*.

¿Cómo se aborda la definición matemática de conjuntos? Lo primero que se debe hacer es definir un *universo de discurso*, es decir, una colección de todas las cosas que nos puede llegar a interesar. Dentro de ese universo de discurso, se seleccionan unos ciertos elementos (una colección) para construir cada conjunto.

Veamos un ejemplo sencillo apoyados en la figura 1.1.

Supongamos que el universo de discurso está representado por los nueve íconos que están en U . Supongamos que estamos hablando sobre los nueve deportes representados

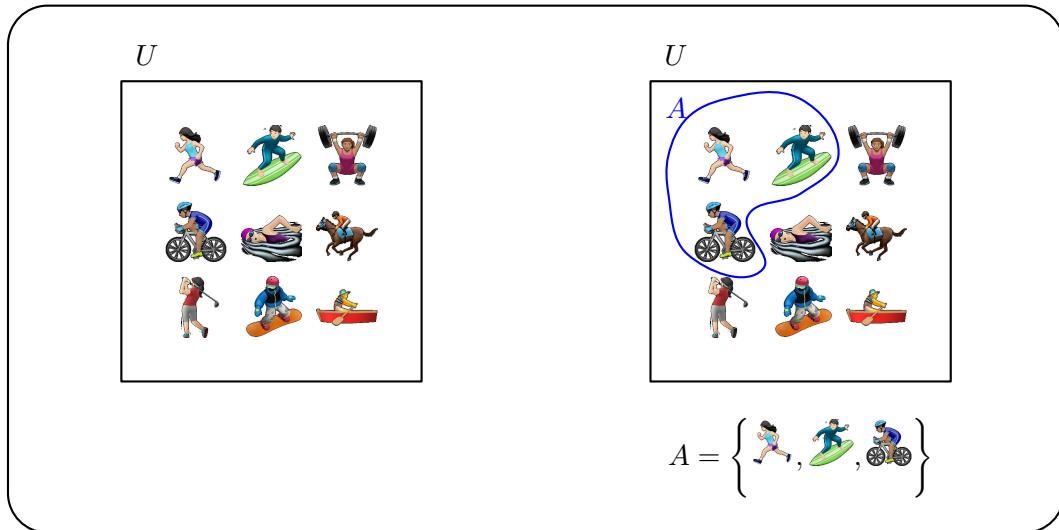


Figura 1.1: Ejemplo de conjunto convencional

por esos íconos y que esos deportes son lo único que nos interesa. Ese es todo nuestro universo de discurso.

En el momento en que se seleccionan algunos de ellos, por ejemplo los tres íconos encerrados por la curva azul, se construye un conjunto. El conjunto A , entonces, está constituido por tres elementos, por los deportes atletismo, surf y ciclismo.

Al hacer eso, se ha dividido el universo en dos partes, lo que está adentro y lo que está afuera del conjunto. En otras palabras, al definir ese conjunto, cada uno de los elementos del universo de discurso sólo puede tener una de dos opciones: o está adentro o está afuera del conjunto. Cada elemento puede *pertenecer* o *no pertenecer* al conjunto A . Utilizamos un símbolo para representar ese hecho: \in y \notin :

- Si x pertenece al conjunto A se simboliza por: $x \in A$
- Si x no pertenece al conjunto A se simboliza por : $x \notin A$

Para ilustrarlo, expresíamos que el atletismo pertenece al conjunto A y que la equitación no pertenece al conjunto A así:

$$\text{atletismo} \in A, \quad \text{equitación} \notin A$$

Ahora bien, hay otra forma de decir exactamente lo mismo: con el *grado de pertenencia*.

Podemos asignar a cada elemento del universo de discurso un grado de pertenencia al conjunto A . El grado de pertenencia es un número que puede ser 0 o 1. Será 0 si el elemento no pertenece al conjunto A y será 1 en caso de que sí pertenezca al conjunto A .

Utilizaremos la letra μ para representar el grado de pertenencia. $\mu_A(x)$ será el grado de pertenencia al conjunto A del elemento x . Por ejemplo, el grado de pertenencia del atletismo al conjunto A es 1 y el grado de pertenencia de la equitación al conjunto A , es 0. Es decir:

$$\mu_A \left(\begin{array}{c} \text{Atleta} \\ \text{(corriendo)} \end{array} \right) = 1 \quad \mu_A \left(\begin{array}{c} \text{Equitación} \\ \text{(caballo)} \end{array} \right) = 0$$

1.2.2. Los conjuntos difusos

Presentamos ahora los conjuntos difusos. La característica especial de los conjuntos difusos es que **los elementos pueden pertenecer parcialmente al conjunto**.

A diferencia del ejemplo de la figura 1.1, en donde cada elemento sólo puede estar o adentro o afuera del conjunto (es decir, solo puede pertenecer completamente o no pertenecer completamente al conjunto), en los conjuntos difusos se permite una pertenencia parcial.

Para representar esta situación, relajamos una condición del grado de pertenencia: el cambio consiste en que en los conjuntos difusos el grado de pertenencia no está restringido a los valores 0 o 1, sino que puede tomar cualquier valor entre 0 y 1, es decir, puede ser cualquier número en el intervalo cerrado $[0, 1]$. Por ejemplo, podrá haber un grado de pertenencia 0.3 que represente una pertenencia parcial a un conjunto.

Veamos un nuevo ejemplo ilustrado en la figura 1.2. Construyamos ahora un conjunto difuso con el mismo universo de discurso U de los nueve deportes y démosle el nombre \tilde{B} .

Para construir el conjunto \tilde{B} debemos asignarle a cada elemento del universo de discurso un grado de pertenencia. A manera de ejemplo, hemos asignado arbitrariamente los valores que se ven en la figura 1.2. Al atletismo se le asignó 0.0, lo que significa que está completamente por fuera del conjunto. Al levantamiento de pesas, 1.0, lo que significa que están completamente adentro del conjunto. Al *surf*, 0.5 lo que significa que está parcialmente adentro y parcialmente afuera del conjunto.

En la misma figura 1.2 hemos intentado dibujar el borde de este conjunto en el diagrama de Venn con un trazo de color rojo. Al intentarlo hemos encontrado problemas:

Claramente podríamos dejar por fuera al atletismo que tiene grado de pertenencia 0 y adentro al levantamiento de pesas que tiene grado de pertenencia 1. Pero ¿qué haríamos con el ciclismo? El ciclismo tiene grado de pertenencia 0.5; no está ni completamente

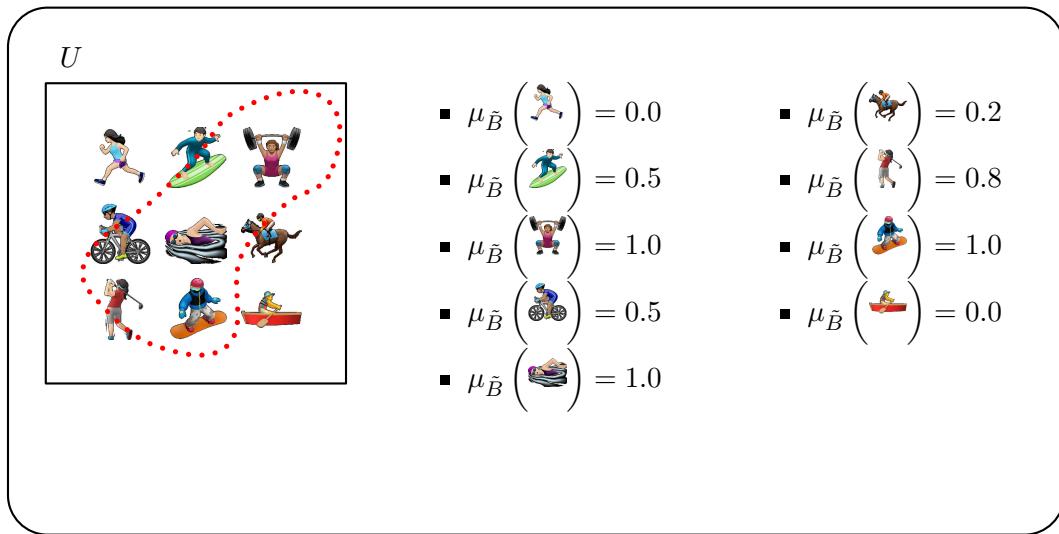


Figura 1.2: Ejemplo de conjunto difuso

adentro ni completamente afuera. Está parcialmente adentro y parcialmente afuera. Por eso hemos representado con líneas punteadas ese supuesto borde de nuestro conjunto. Realmente no podemos trazar el borde de manera nítida. La frontera no es nítida. La frontera es *difusa* y eso es lo que da origen al nombre de los *conjuntos difusos*.

1.2.3. Notación sumatoria (de Zadeh)

Hay otra forma de representar los conjuntos difusos, denominada notación sumatoria o notación de Zadeh. Al usar esa notación, el conjunto \tilde{B} de la figura 1.2 se describe así:

$$\tilde{B} = \left\{ \frac{0.0}{\text{Running}} + \frac{0.5}{\text{Surfing}} + \frac{1.0}{\text{Weightlifting}} + \frac{0.5}{\text{Cycling}} + \frac{1.0}{\text{Snowboarding}} + \frac{0.2}{\text{Horseback Riding}} + \frac{0.8}{\text{Swimming}} + \frac{0.5}{\text{Golf}} + \frac{0.0}{\text{Rowing}} \right\}$$

Aun cuando hemos usado el símbolo $+$, no se trata de una suma. Es sólo una forma de describir un conjunto difuso. Al usar esta notación, se expresa cada elemento del universo de discurso y su grado de pertenencia representados en una fracción; pero tampoco se trata de una fracción, es sólo una forma de escritura.

De manera un poco más compacta, suelen omitirse aquellos elementos cuyo grado de pertenencia es 0. De manera tal que el conjunto difuso \tilde{B} puede representarse así:

$$\tilde{B} = \left\{ \frac{0.5}{\text{Surfing}} + \frac{1.0}{\text{Weightlifting}} + \frac{0.5}{\text{Cycling}} + \frac{1.0}{\text{Snowboarding}} + \frac{0.2}{\text{Horseback Riding}} + \frac{0.8}{\text{Swimming}} + \frac{0.5}{\text{Golf}} + \right\}$$

Generalizando, diremos que un conjunto difuso \tilde{B} puede representarse mediante la notación

$$\tilde{B} = \left\{ \frac{\mu_{\tilde{B}}(x_1)}{x_1} + \frac{\mu_{\tilde{B}}(x_2)}{x_2} + \dots \right\}$$

$$\tilde{B} = \left\{ \sum_i \frac{\mu_{\tilde{B}}(x_i)}{x_i} \right\}$$

Ahora bien, y esto es importante, el ejemplo de los nueve deportes es un ejemplo en donde podemos enumerar cada uno de los elementos del universo. Decimos entonces que se trata de un universo de discurso *discreto*.

Sin embargo, existen también universos de discurso en los que no podemos enumerar los elementos porque son *continuos*. La recta real o un intervalo de la recta real son conjuntos continuos en los que, además, hay infinitos elementos.

Cuando tenemos un Universo de discurso continuo cambiamos la notación de sumatoria por la de una integral.

$$\tilde{B} = \left\{ \int \frac{\mu_{\tilde{B}}(x)}{x} \right\}$$

1.2.4. Conjuntos difusos sobre universos de discurso continuos

Veamos un ejemplo con universos de discurso continuos. Utilicemos como universo de discurso la recta real.

En la figura 1.3 se presentan dos conjuntos difusos, \tilde{C} y \tilde{D} , cada uno con un color diferente. ¿Qué debo hacer para construir un conjunto difuso? Debo asignarle a cada valor de la recta real un grado de pertenencia a ese conjunto.

En estas condiciones, a veces se utiliza el término *función de pertenencia*, en lugar de *grado de pertenencia*, porque la función asigna el valor entre cero y uno del grado de pertenencia.

Las formas que resultan pueden ser muy variadas y en la figura 1.3 simplemente se muestran dos funciones cualesquiera. Será usual utilizar conjuntos difusos con formas sencillas, como por ejemplo, el trapecio del conjunto difuso \tilde{D} .

1.2.5. Para destacar

Recordemos. El principal aporte de la definición de los conjuntos difusos es que ahora un elemento puede pertenecer parcialmente al conjunto y, a la vez, al complemento del conjunto.

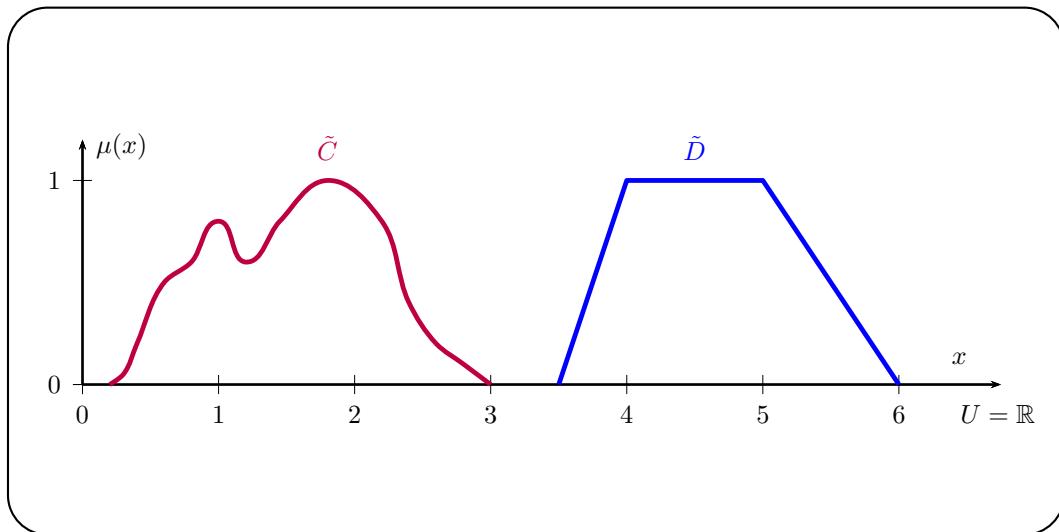


Figura 1.3: Ejemplo de conjuntos difusos sobre la recta real

Los conjuntos convencionales vamos a denominarlos, a partir de este momento, *conjuntos concretos*². En inglés los conjuntos difusos se denominan *fuzzy sets* y los concretos *crisp sets*.

En la tabla 1.1 se hace una comparación entre los conjuntos concretos y los conjuntos difusos. Tenemos por lo menos tres cosas que los diferencian:

1. La primera comparación tiene que ver con los grados de pertenencia. La pertenencia a un conjunto concreto solo puede ser cero o uno, mientras que la pertenencia

²Hace unos años se utilizaba en español también el término *conjunto borroso*, en particular, de la literatura proveniente de España y México. Hoy en día, ese término ha caído en desuso.

Tabla 1.1: Comparación entre los conjuntos concretos y los conjuntos difusos

Conjuntos concretos (<i>crisp sets</i>)	Conjuntos difusos (<i>fuzzy sets</i>)
Pertenencia 0 ó 1 $\mu_x \in \{0, 1\}$	Pertenencia entre 0 y 1 $\mu_x \in [0, 1]$
Principio del tercer excluido	Pertenencia simultánea a C y a C'
Fronteras nítidas	Fronteras difusas

a un conjunto difuso puede ser cualquier valor entre cero y uno. Eso deriva en el siguiente punto de comparación.

2. Los conjuntos concretos satisfacen el principio del tercer excluido. ¿Qué significa eso? Que un elemento solo tiene dos opciones, pertenecer o no pertenecer al conjunto. Cualquier tercera opción está excluida. Ese principio no se satisface en los conjuntos difusos. De hecho, se puede pertenecer al conjunto y a su complemento, simultáneamente.
3. Como consecuencia de lo anterior, las fronteras de los concretos, son nítidas, son concretas. Por su parte, las fronteras de los conjuntos difusos son, precisamente, difusas.

Quizás algún lector pueda estar en este momento preguntándose si hemos removido todo el andamiaje de la matemática. Al fin de cuentas, a la hora de formalizar la matemática, la teoría de conjuntos es la base para muchas ramas de la matemática: para la lógica, para el álgebra, para la aritmética, etcétera. Y ahora hemos ‘cambiado’ la teoría de conjuntos. ¿Es este, entonces, un cisma matemático?

No, realmente no. La teoría de conjuntos clásica, de conjuntos concretos, nos permite definir dentro de ella a los conjuntos difusos.

Un conjunto difuso es un conjunto de parejas: parejas elemento y grado de pertenencia. Eso se ve muy claro en la notación sumatoria. En otras palabras, podemos emplear los conjuntos concretos para definir los conjuntos difusos así:

$$SD = \{(x, \mu(x)) \mid x \in U, \mu(x) \in [0, 1]\}$$

Por eso, deberíamos, en rigor, hablar de una *teoría de subconjuntos difusos*, porque están definidos al interior de la teoría de conjuntos concretos.

Lo que sí tenemos es una gran oportunidad. Con los conjuntos difusos podemos, eso sí, construir nuevas versiones de las áreas clásicas de la matemática, como la lógica, la aritmética, la geometría, etcétera. En particular, en este curso, abordaremos aplicaciones basadas en lógica difusa y en aritmética difusa.

Para concluir esta sección, invito al lector a reflexionar sobre lo siguiente: ¿qué tipo de conjunto, concreto o difuso, representaría de mejor forma los siguientes casos?

- Las personas altas.
- Las personas vivas.
- Las poblaciones vulnerables.
- Los proyectos de impacto social alto.

- Los clientes satisfechos.

1.3. Algunas definiciones

1.3.1. Soporte de un conjunto difuso

El soporte de un conjunto difuso es un subconjunto del universo de discurso. Lo forman aquellos elementos del universo de discurso que tienen grado de pertenencia al conjunto mayor que 0, es decir, un grado de pertenencia que no es nulo y por tanto pertenecen, así sea con un valor muy pequeño, al conjunto difuso. En términos matemáticos:

$$\begin{aligned} \text{soporte}(A) &\subset U \\ \text{soporte}(A) &= \{x \in U \mid \mu_A(x) > 0\} \end{aligned} \quad (1.1)$$

Ejemplo 1.1 (Soporte de un conjunto difuso discreto) Ilustremos el concepto con un ejemplo. Tomamos como universo de discurso los deportes y definimos el conjunto difuso A , utilizando la notación sumatoria:

$$A = \left\{ \frac{0.0}{\text{fútbol}} + \frac{0.0}{\text{atletismo}} + \frac{0.5}{\text{surf}} + \frac{0.0}{\text{natación}} + \frac{0.0}{\text{equitación}} + \frac{0.7}{\text{natación}} + \frac{0.3}{\text{equitación}} + \frac{0.0}{\text{atletismo}} + \frac{1.0}{\text{surf}} + \frac{0.0}{\text{snowboard}} \right\}$$

¿Cuál es el soporte? El soporte son esos elementos del universo de discurso que tienen grado de pertenencia mayor que 0. En este caso, son cuatro elementos, el *surf* que tiene grado de pertenencia 0.5, la natación que tiene grado de pertenencia 0.7, la equitación que tiene 0.3 y el *snowboard* que tiene grado de pertenencia 1.0. Es decir:

$$\text{soporte}(A) = \left\{ \text{surf}, \text{natación}, \text{equitación}, \text{snowboard} \right\}$$

Ejemplo 1.2 (Soporte de tres conjuntos difusos continuos) Vamos a presentar unos ejemplos sobre un universo de discurso continuo: la recta real. En la figura 1.4 se muestran tres conjuntos difusos definidos sobre \mathbb{R} , y sus respectivos soportes. Cada uno de los conjuntos se define a través de la función de pertenencia graficada.

El soporte son esos elementos del universo de discurso que tienen grado de pertenencia mayor que 0. En el ejemplo de la figura 1.4a corresponde al intervalo que va desde 1 hasta 9:

$$\text{soporte}(A) = [1, 9]$$

El ejemplo de la figura 1.4b tiene la particularidad de que son dos trapecios. Si buscamos el soporte, resulta ser entonces la unión de dos intervalos:

$$\text{soporte}(B) = [1, 4] \cup [2, 8]$$

El ejemplo de la figura 1.4c tiene una forma especial que utilizaremos en algunas ocasiones. Se trata de un conjunto tipo *singleton*. Los *singleton* son conjuntos que tienen grado de pertenencia 0 en todo el universo de discurso, salvo en un elemento. Esa es la singularidad del *singleton*, y en esa singularidad tiene un grado de pertenencia 1. Entonces, el soporte de un conjunto difuso *singleton*, es justamente su singularidad:

$$\text{soporte}(C) = 4.5$$

1.3.2. Núcleo de un conjunto difuso

El núcleo de un conjunto difuso es un subconjunto del universo de discurso. Lo forman aquellos elementos del universo de discurso que tienen grado de pertenencia al conjunto igual a 1. En términos matemáticos:

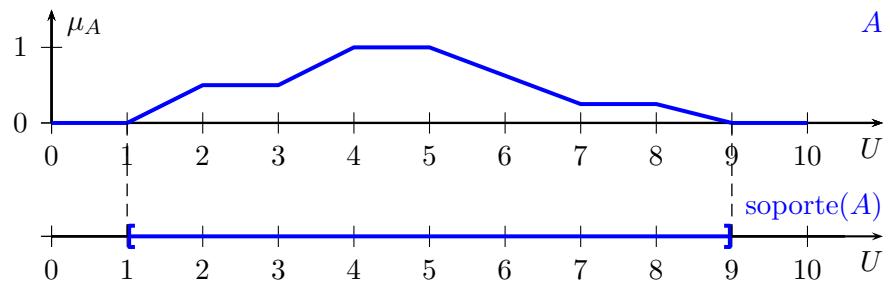
$$\begin{aligned} \text{núcleo}(A) &\subset U \\ \text{núcleo}(A) &= \{x \in U \mid \mu_A(x) = 1\} \end{aligned} \tag{1.2}$$

Nótese que esta definición implica que si un elemento pertenece al núcleo, también pertenece al soporte, o lo que es igual:

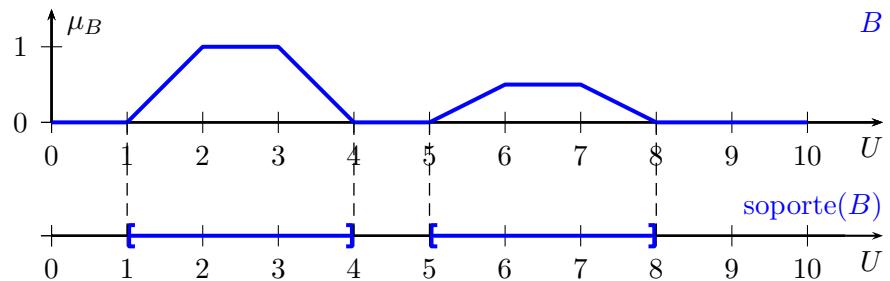
$$\text{núcleo}(A) \subset \text{soporte}(A) \tag{1.3}$$

Ejemplo 1.3 (Núcleo de un conjunto difuso discreto) Si consideramos el conjunto A del ejemplo 1.1, encontramos que solamente hay un elemento que tiene grado de pertenencia igual a 1, el *surf*. Por esta razón:

$$\text{núcleo}(A) = \left\{ \text{surf} \right\}$$



(a) Un conjunto convexo



(b) Un conjunto no convexo

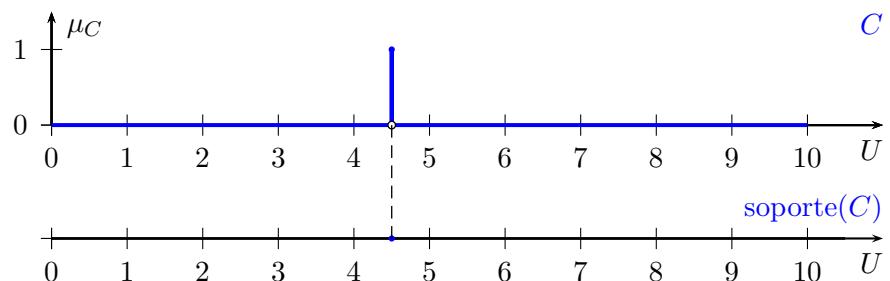
(c) Un conjunto tipo *singleton*

Figura 1.4: Tres conjuntos difusos continuos, y sus soportes

Ejemplo 1.4 (Núcleo de tres conjuntos difusos continuos) Consideremos ahora los tres conjuntos del ejemplo 1.2 que se han graficado en la figura 1.5.

El núcleo son esos elementos del universo de discurso que tienen grado de pertenencia igual a 1, que resultan ser:

$$\begin{aligned}\text{núcleo}(A) &= [4, 5] \\ \text{núcleo}(B) &= [2, 3] \\ \text{núcleo}(C) &= 4.5\end{aligned}$$

1.3.3. α -corte de un conjunto difuso

El α -corte de un conjunto difuso también es un subconjunto del universo de discurso. α es un número entre 0 y 1 y el α -corte está formado por los elementos del universo de discurso que tienen un grado de pertenencia mayor o igual a α . Si seleccionamos, por ejemplo, $\alpha = 0.3$, entonces deberíamos buscar los elementos del universo de discurso cuyo grado de pertenencia es mayor o igual a 0.3. En esta definición se hace una salvedad para el caso en que $\alpha = 0$:

Sea A un conjunto difuso definido sobre U , y $\alpha \in [0, 1]$. Se define:

$$A_\alpha = \begin{cases} x \in U \mid \mu_A(x) \geq \alpha & \text{si } \alpha > 0 \\ x \in U \mid \mu_A(x) > \alpha & \text{si } \alpha = 0 \end{cases} \quad (1.4)$$

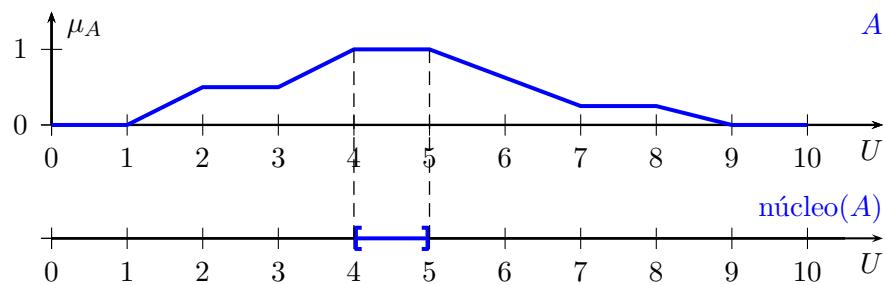
o lo que es igual:

$$A_\alpha = \begin{cases} x \in U \mid \mu_A(x) \geq \alpha & \text{si } \alpha > 0 \\ \text{soporte}(A) & \text{si } \alpha = 0 \end{cases} \quad (1.5)$$

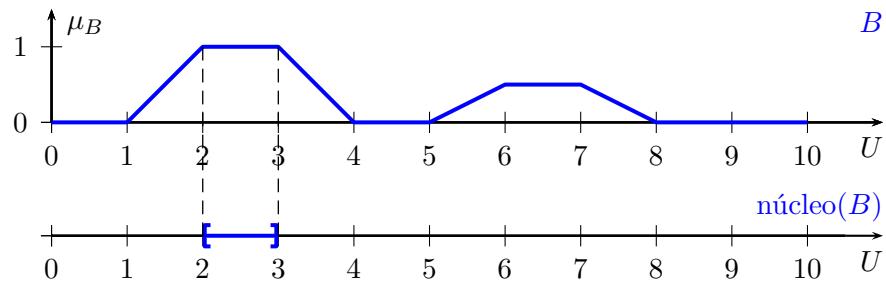
La excepción para $\alpha = 0$ se debe a que el conjunto de los elementos con grado de pertenencia mayor o igual a 0 es todo el universo de discurso. Esta excepción será particularmente útil en las implementaciones en software de las operaciones de aritmética difusa.

Ejemplo 1.5 (α -cortes de un conjunto discreto) Usando como universo de discurso el conjunto de los deportes, definimos el conjunto difuso A :

$$A = \left\{ \frac{0.0}{\text{Fútbol}} + \frac{0.0}{\text{Atletismo}} + \frac{0.5}{\text{Natación}} + \frac{0.0}{\text{Voleibol}} + \frac{0.0}{\text{Pádel}} + \frac{0.7}{\text{Tenis}} + \frac{0.3}{\text{Bádminton}} + \frac{0.0}{\text{Gimnasia rítmica}} + \frac{1.0}{\text{Boxeo}} + \frac{0.0}{\text{Ajedrez}} \right\}$$



(a) Un conjunto convexo



(b) Un conjunto no convexo

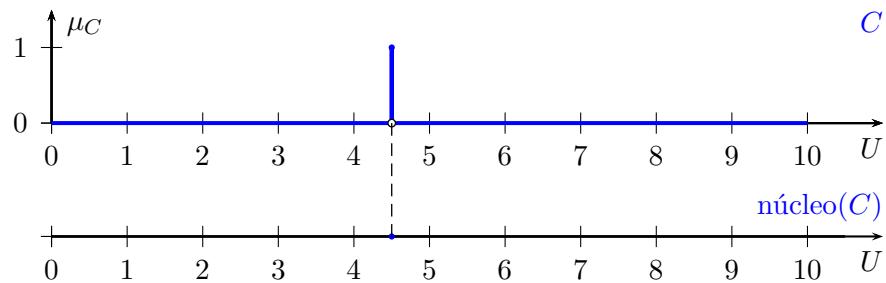
(c) Un conjunto tipo *singleton*

Figura 1.5: Tres conjuntos difusos continuos, y sus núcleos

A continuación se muestran los α -cortes de A para diferentes valores de α : 1.0, 0.8, 0.7, 0.4 y 0.0:

$$\begin{aligned} A_{1.0} &= \left\{ \text{snowboard} \right\} \\ A_{0.8} &= \left\{ \text{surf}, \text{snowboard} \right\} \\ A_{0.7} &= \left\{ \text{surf}, \text{natación}, \text{snowboard} \right\} \\ A_{0.4} &= \left\{ \text{surf}, \text{natación}, \text{equitación}, \text{snowboard} \right\} \\ A_{0.0} &= \left\{ \text{surf}, \text{natación}, \text{equitación}, \text{snowboard} \right\} \end{aligned}$$

¿Cuál es el α -corte de A cuando $\alpha = 1$? Será el conjunto formado por todos aquellos elementos del universo de discurso cuyo grado de pertenencia es mayor o igual que 1. En este caso, únicamente el *snowboard*.

Si bajamos el valor de α hasta 0.8 para encontrar el α -corte debemos buscar los elementos del universo de discurso cuyo grado de pertenencia es mayor o igual que 0.8. En nuestro ejemplo sigue siendo solo el *snowboard*.

Pero al bajar hasta 0.7 entra dentro del α -corte un elemento más, la natación, que tiene grado de pertenencia 0.7. Al disminuir α hasta 0.4 se incorpora un nuevo deporte, el *surf*, que tiene grado de pertenencia 0.5.

Y si bajamos hasta 0, debemos reemplazar el signo \geq por el signo $>$, o lo que es igual, debemos tomar el soporte de ese conjunto: los cuatro deportes *surf*, natación, equitación y *snowboard*.

Nótese que al ir disminuyendo el valor de α desde 1 hacia 0 los conjuntos van creciendo, se van incorporando más elementos. Lo que decimos, entonces, es que esos *alpha*-cortes están anidados uno dentro del otro. El más pequeño será el correspondiente a $\alpha = 1$, y este crece conforme se disminuye el valor de α .

Ejemplo 1.6 (α -cortes de un conjunto continuo) La figura 1.6 muestra un conjunto A continuo, definido sobre la recta real. En esa misma figura se han dibujado los α -cortes para diferentes valores de α : 1.0, 0.75, 0.50, 0.25 y 0.0.

Para encontrar el α -corte cuando $\alpha = 1.0$, trazamos una línea en horizontal 1 y vemos cuáles son los elementos del universo de discurso que quedaron intersectados por esa línea

o que quedan por encima de ella. En este caso, el intervalo $[4.0, 5.0]$. Es decir:

$$A_{1.0} = [4.0, 5.0]$$

Para los demás valores de α se obtiene:

$$\begin{aligned} A_{0.75} &= [3.5, 5.66] \\ A_{0.50} &= [2.0, 6.33] \\ A_{0.25} &= [1.5, 8.0] \\ A_{0.00} &= [1.0, 9.0] \end{aligned}$$

Ejemplo 1.7 (α -corte de un conjunto no convexo) La figura 1.7 muestra un conjunto A que no es convexo, definido sobre la recta real. En esa misma figura se han dibujado los α -cortes para diferentes valores de α : 1.0, 0.75, 0.50, 0.25 y 0.0.

En este ejemplo se observa que un α -corte de un conjunto continuo no siempre es un intervalo cerrado: para $\alpha = 1.0$ el corte resultante es la unión de un intervalos y un punto, mientras que para $\alpha = 0.75$ el corte resultante es la unión de dos intervalos:

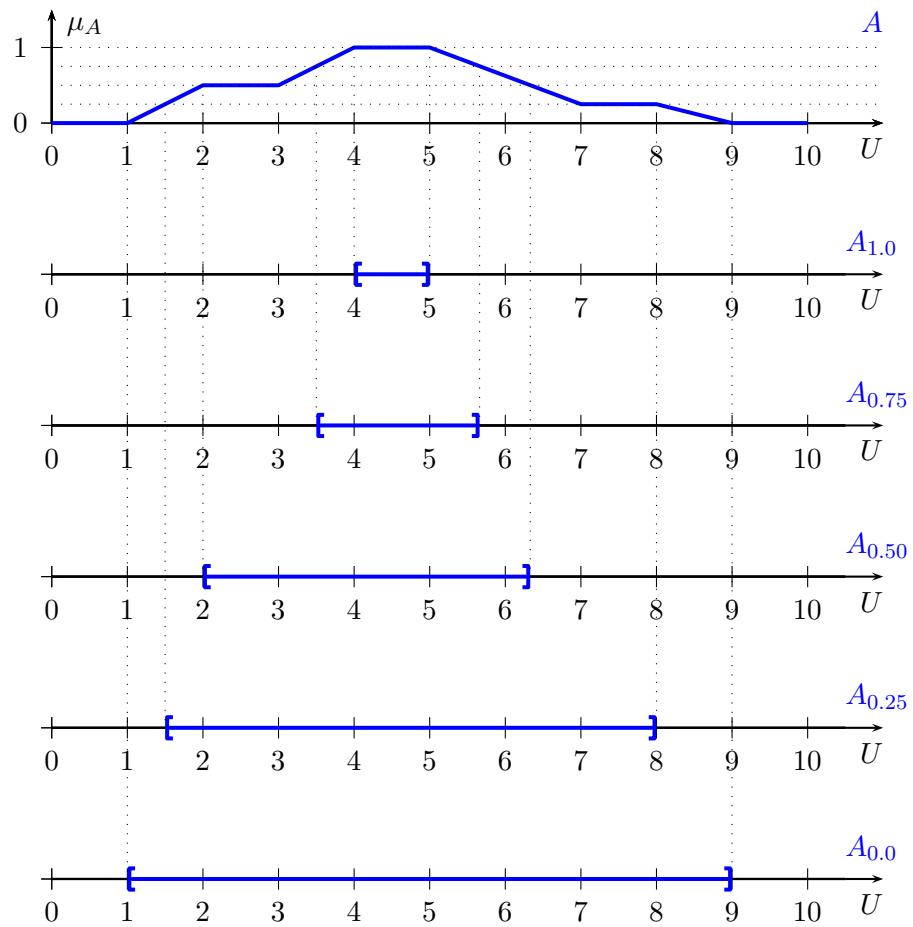
$$\begin{aligned} A_{1.00} &= [2.00, 3.00] \cup \{7\} \\ A_{0.75} &= [1.75, 3.50] \cup [6.50, 7.25] \\ A_{0.50} &= [1.50, 7.50] \\ A_{0.25} &= [1.25, 7.75] \\ A_{0.00} &= [1.00, 8.00] \end{aligned}$$

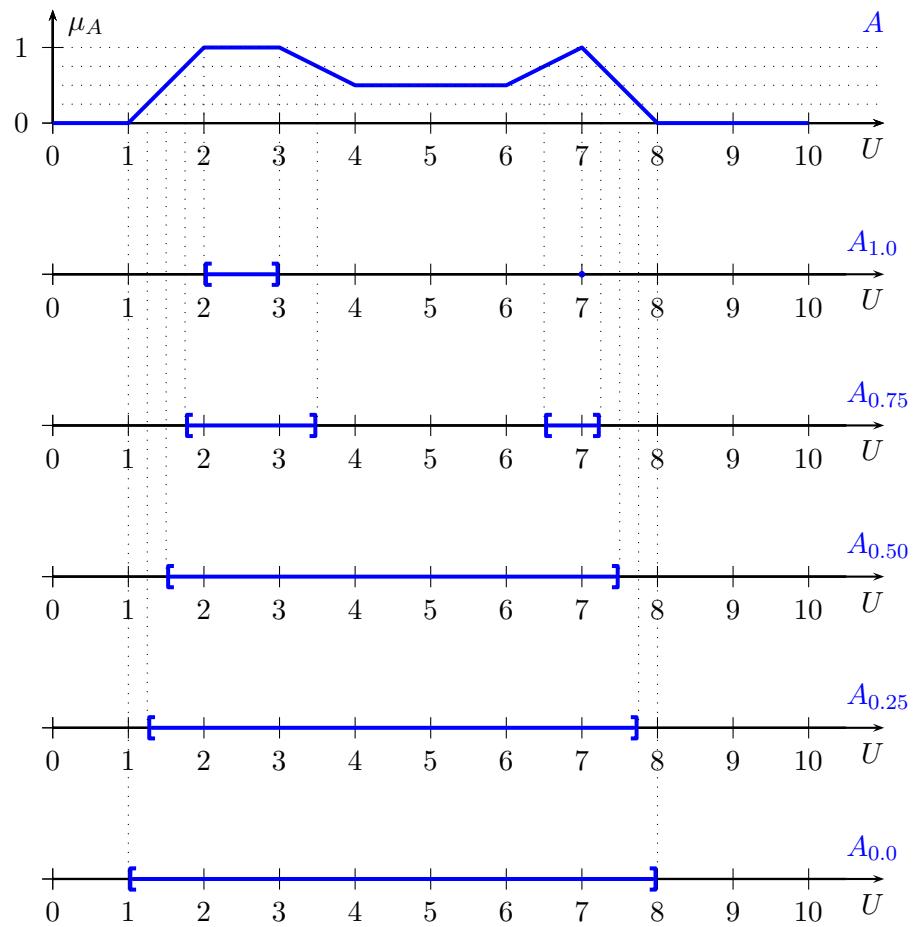
Ejemplo 1.8 (α -corte de un conjunto discontinuo) La figura 1.8 muestra un conjunto A discontinuo, definido sobre la recta real. En esa misma figura se han dibujado un α -corte para $\alpha = 0.5$. Existen dos discontinuidades, una en $x = 3$ y otra en $x = 7$.

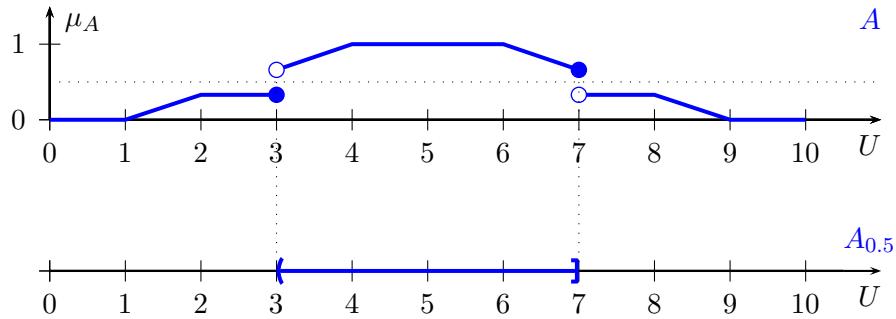
¿Qué valor toma la función en las discontinuidades? Para resolver la ambigüedad se ha utilizado la siguiente convención para las fronteras de la discontinuidad: el punto relleno indica el valor de la función. De acuerdo a lo anterior se tiene que:

$$\mu_A(3) = 0.33 \quad \mu_A(7) = 0.66$$

En esas condiciones, el α -corte correspondiente a $\alpha = 0.5$ tiene la particularidad ser un

Figura 1.6: α -cortes del ejemplo 1.6.

Figura 1.7: α -cortes del ejemplo 1.7.

Figura 1.8: α -cortes del ejemplo 1.8.

intervalo con uno de sus extremos abiertos y el otro cerrado:

$$A_{0.5} = (3, 7]$$

1.3.4. Producto cartesiano

El producto cartesiano de dos conjuntos es el conjunto de todas las parejas posibles que se puedan obtener tomando un elemento del primer conjunto y un elemento del segundo conjunto. Suele utilizarse el símbolo \times para denotar la operación de producto cartesiano.

De acuerdo a lo anterior, para dos conjuntos concretos A y B su producto cartesiano se expresa de la siguiente forma:

$$A \times B = \{(x, y) \mid x \in A, y \in B\} \quad (1.6)$$

Ejemplo 1.9 (Producto cartesiano de dos conjuntos concretos discretos)

Definamos los conjuntos A y B como:

$$A = \left\{ \begin{array}{c} \text{🏃} \\ \text{🏄} \end{array} \right\} \qquad B = \left\{ \begin{array}{c} \text{🦉} \\ \text{🐬} \\ \text{🦢} \end{array} \right\}$$

El producto cartesiano estará formado por todas las posibles parejas de elementos que tienen

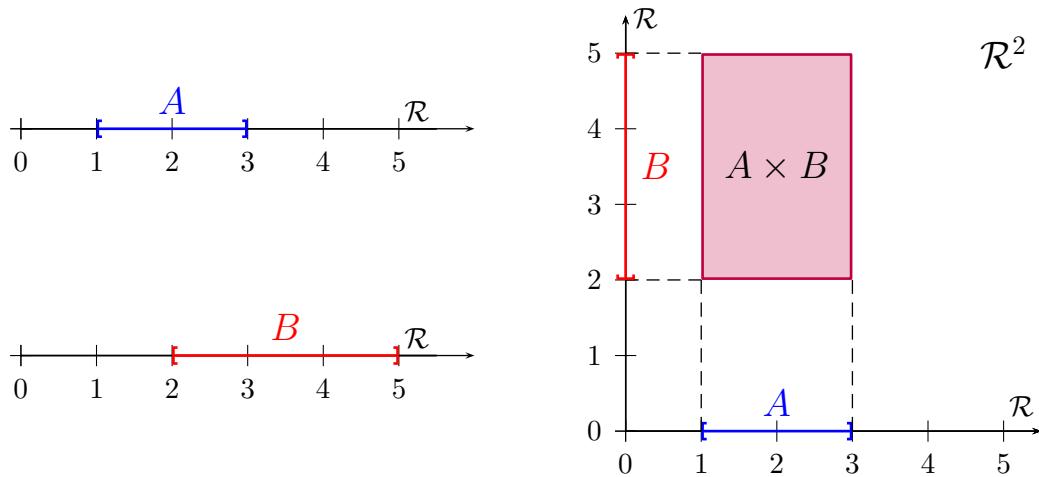


Figura 1.9: Producto cartesiano del ejemplo 1.10

un elemento de A y un elemento de B :

$$A \times B = \left\{ \begin{array}{l} \left(\begin{array}{c} \text{runner} \\ \text{surfer} \end{array}, \begin{array}{c} \text{owl} \\ \text{shark} \end{array} \right), \quad \left(\begin{array}{c} \text{runner} \\ \text{surfer} \end{array}, \begin{array}{c} \text{shark} \\ \text{flamingo} \end{array} \right), \quad \left(\begin{array}{c} \text{runner} \\ \text{surfer} \end{array}, \begin{array}{c} \text{flamingo} \\ \text{flamingo} \end{array} \right), \\ \left(\begin{array}{c} \text{owl} \\ \text{shark} \end{array}, \begin{array}{c} \text{owl} \\ \text{shark} \end{array} \right), \quad \left(\begin{array}{c} \text{owl} \\ \text{shark} \end{array}, \begin{array}{c} \text{shark} \\ \text{flamingo} \end{array} \right), \quad \left(\begin{array}{c} \text{owl} \\ \text{shark} \end{array}, \begin{array}{c} \text{flamingo} \\ \text{flamingo} \end{array} \right) \end{array} \right\}$$

Ejemplo 1.10 (Producto cartesiano de dos conjuntos concretos continuos)

La figura 1.9 muestra dos conjuntos continuos definidos sobre la recta real y su producto cartesiano.

El conjunto A corresponde al intervalo $[2, 3]$ y el conjunto B al intervalo $[2, 5]$. Como el producto cartesiano es un conjunto de parejas de los intervalos, debemos representarlo en \mathbb{R}^2 . Los puntos forman un rectángulo cuyos lados corresponden a los intervalos de los conjuntos A y B .

1.3.5. Extensión cilíndrica

Supongamos que el conjunto difuso A está definido sobre el universo de discurso U . Es posible extender ese conjunto hacia el universo de discurso $U \times V$, en donde V es otro universo. Si denotamos por x a los elementos de U , y por y a los elementos de V , el conjunto original tendrá una función de pertenencia $\mu_A(x)$ mientras que la función de pertenencia del conjunto extendido será $\mu_{\tilde{A}}(x, y)$ que se calcula así:

$$\mu_{\tilde{A}}(x, y) = \mu_A(x) \quad (1.7)$$

La operación anterior recibe el nombre de *extensión cilíndrica* o *extensión ortogonal* (es decir, perpendicular)³. Al analizar esta operación, se observa que la extensión cilíndrica es una operación que lleva un conjunto difuso definido en un universo de discurso, hacia otro universo de discurso, de una dimensión mayor, que contiene al universo original.

Ejemplo 1.11 (Extensión cilíndrica sobre universos discretos) Utilizando el universo U que se muestra en la figura 1.1, el de los deportes, definimos el conjunto C :

$$C = \left\{ \frac{0.4}{\text{levantamiento pesas}}, \frac{0.6}{\text{ciclismo}}, \frac{1.0}{\text{natación}}, \frac{0.5}{\text{caballo}} \right\}$$

Tomemos ahora el universo de discurso V formado por los animales:

$$V = \left\{ \text{águila}, \text{pato}, \text{dragón}, \text{pez}, \text{ búho}, \text{flamingo}, \text{gato}, \text{mono}, \text{zorro} \right\}$$

Para extender el conjunto C al universo de discurso $U \times V$ construimos la tabla que se muestra en la figura 1.10a. En esta tabla, a partir de la tercera columna cada casilla es una pareja de deporte y animal. Es decir, esta tabla representa el producto cartesiano de esos dos universos de discurso. Se ha empleado la segunda columna para mostrar los grados de pertenencia de C sobre U .

En la figura 1.10b se muestra la dirección en la que se quiere extender C para construir el conjunto C_{ext} sobre $U \times V$. Para ello, los valores de la segunda columna se copian en las demás columnas, extendiendo así el conjunto, de la misma forma en que un círculo de \mathcal{R}^2 se extendería para formar un cilindro en \mathcal{R}^3 . El resultado se muestra en la figura 1.10c.

En este ejemplo se han omitido los grados de pertenencia nulos, para facilitar la explicación. La figura 1.10d muestra la tabla con todos los grados de pertenencia, incluidos aquellos que son nulos.

³Es factible realizar una extensión en una dirección no ortogonal; este tipo de extensiones se denominan *oblicuas*, pero no se incluyen en este texto.

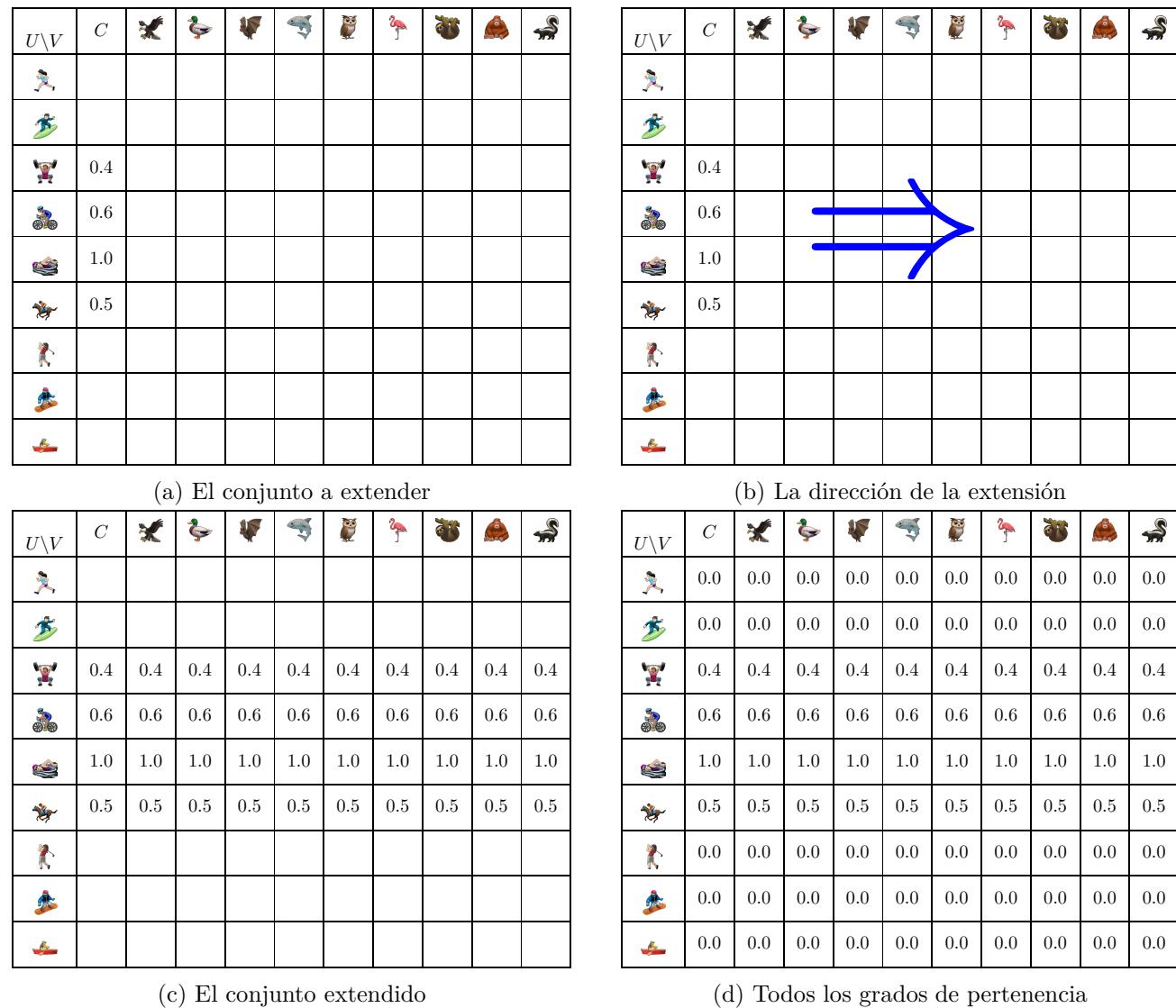


Figura 1.10: Extensión cilíndrica en universos discretos

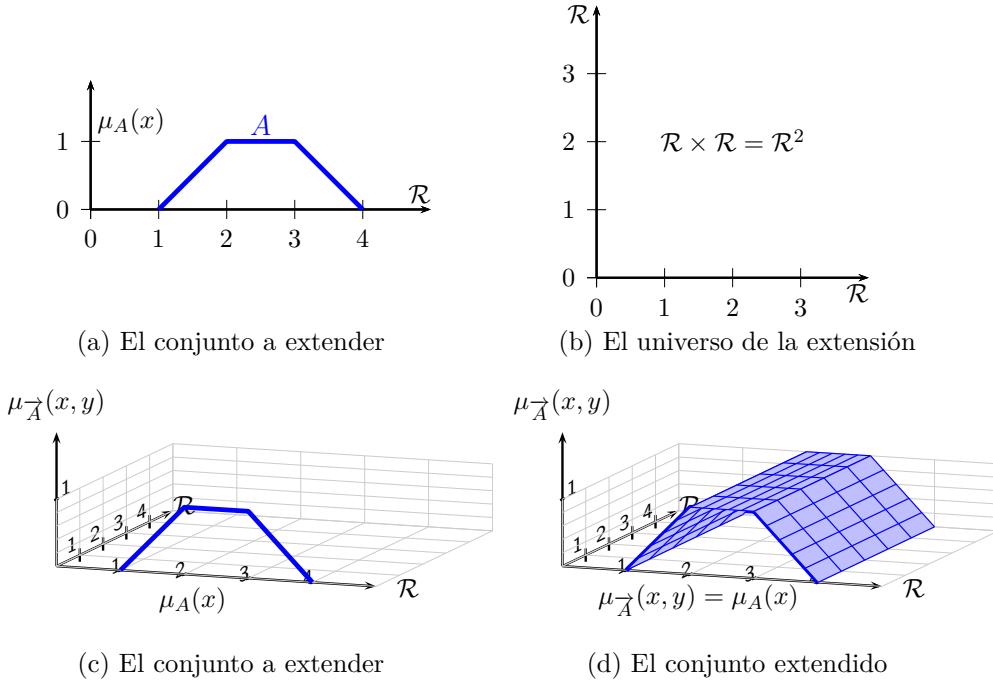


Figura 1.11: Extensión cilíndrica en universos continuos

Ejemplo 1.12 (Extensión cilíndrica sobre universos continuos) Consideremos el conjunto difuso A que se muestra en la figura 1.11a. A está definido sobre el universo de discurso \mathcal{R} , que es la recta real. Supongamos ahora que queremos extender A sobre \mathcal{R}^2 , tal como se muestra en la figura 1.11b.

La función de pertenencia del conjunto extendido, $\mu_{\vec{A}}$, debe visualizarse en \mathcal{R}^3 . La figura 1.11c muestra $\mu_A(x)$ en \mathcal{R}^3 , y su extensión ortogonal se visualiza en 1.11d.

Como $\mu_{\vec{A}}(x, y) = \mu_A(x)$, la forma del conjunto \vec{A} se puede visualizar como si se hubiese 'estirado' el conjunto A en la dirección de la nueva dimensión, es decir, perpendicularmente (ortogonalmente).

1.3.6. Proyección cilíndrica

La proyección cilíndrica (también denominada *proyección ortogonal*), puede asimilarse al proceso inverso de la extensión cilíndrica⁴. Dado un conjunto difuso A definido sobre un universo de discurso compuesto por el producto cartesiano $U \times V$, la proyección nos permite llevar ese conjunto a uno de los dos universos U ó V , mediante las operaciones:

$$\begin{aligned} \text{Proyección sobre } U : \mu_A^{\leftarrow}(x) &= \max_V (\mu_A(x, y)) \\ \text{Proyección sobre } V : \mu_A^{\leftarrow}(y) &= \max_U (\mu_A(x, y)) \end{aligned} \quad (1.8)$$

Ejemplo 1.13 (Proyección cilíndrica sobre universos discretos) Consideremos ahora el conjunto A que se muestra en la figura 1.12a. Se trata de un conjunto definido sobre un universo de discurso $U \times V$, en donde tanto U como V son discretos. U está formado por algunos deportes, mientras que V está formado por algunos animales. En la figura 1.12b se han borrado los valores nulos de $\mu_A(x, y)$ para facilitar la explicación.

Para realizar la proyección de A sobre U , aplicando la ecuación 1.8, se debe tomar el mayor valor de $\mu_A(x, y)$ en la dirección que indica la figura 1.12c. El resultado se muestra en la figura 1.12d, y resulta ser:

$$\overleftarrow{A}(x) = \left\{ \frac{0.7}{\text{levantamiento de pesas}} + \frac{1.0}{\text{ciclismo}} + \frac{0.8}{\text{natación}} + \frac{0.5}{\text{caballo}} \right\}$$

Ahora bien, si se desea realizar la proyección de A sobre U se debe tomar el mayor valor de $\mu_A(x, y)$, pero en la dirección que indica la figura 1.13c. El resultado se muestra en la figura 1.13d, y resulta ser:

$$\overleftarrow{A}(x) = \left\{ \frac{0.1}{\text{pato}} + \frac{0.5}{\text{oso}} + \frac{0.8}{\text{pez}} + \frac{1.0}{\text{ búho}} + \frac{0.9}{\text{flamingo}} \right\}$$

Ejemplo 1.14 (Proyección cilíndrica sobre universos continuos) Consideremos ahora el conjunto A de la figura 1.14a. Está definido sobre $U \times V$, en donde tanto U como V son continuos; de hecho, son rectas reales.

⁴También es posible realizar proyecciones que no son perpendiculares, es decir, proyecciones oblicuas. Una presentación elegante de las herramientas de álgebra lineal que permiten realizar estas proyecciones se encuentra en [50].

$U \setminus V$	🦅	🦆	🦇	🦈	🦉	🦩	🦍	🐿️	🐿️
🏃	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
🏄	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
🏋️	0.0	0.0	0.3	0.5	0.7	0.0	0.0	0.0	0.0
🏍️	0.0	0.0	0.5	0.8	1.0	0.9	0.0	0.0	0.0
🎂	0.0	0.1	0.4	0.8	0.7	0.0	0.0	0.0	0.0
🐎	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0
🤸	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
🏂	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

(a) El conjunto a proyectar

$U \setminus V$	🦅	🦆	🦇	🦈	🦉	🦩	🦍	🐿️	🐿️
🏃									
🏄									
🏋️									
🏍️									
🎂									
🐎									
🤸									
🏂									

(b) El conjunto a proyectar

$U \setminus V$		🦅	🦆	🦇	🦈	🦉	🦩	🦍	🐿️	🐿️
🏃										
🏄										
🏋️										
🏍️										
🎂										
🐎										
🤸										
🏂										

(c) Dirección de la Proyección sobre U

$U \setminus V$	pry	🦅	🦆	🦇	🦈	🦉	🦩	🦍	🐿️	🐿️
🏃										
🏄										
🏋️	0.7									
🏍️	1.0									
🎂	0.8									
🐎	0.5									
🤸										
🏂										

(d) Conjunto proyectado sobre U Figura 1.12: Proyección cilíndrica sobre el universo discreto V

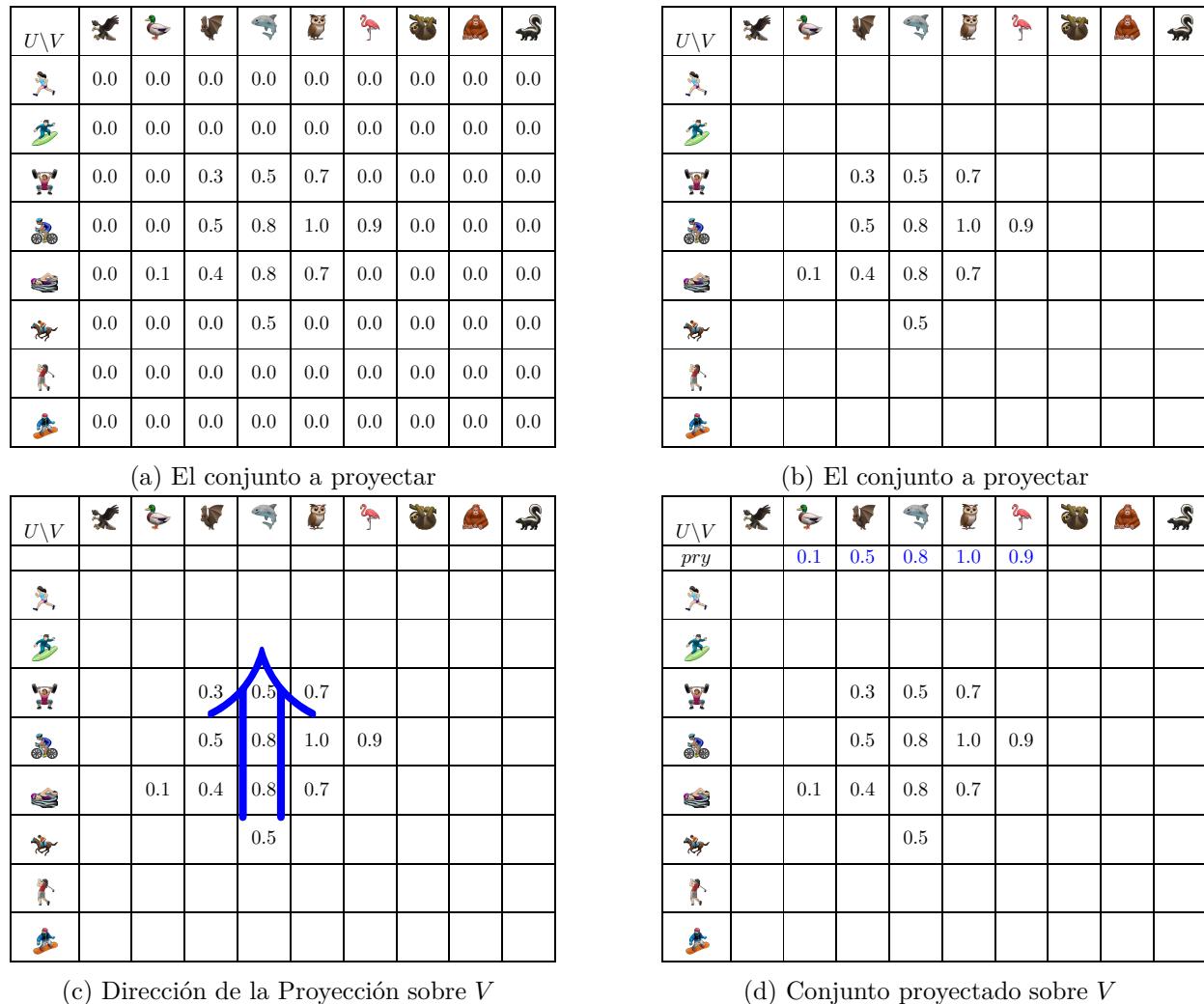


Figura 1.13: Proyección cilíndrica sobre el universo discreto U

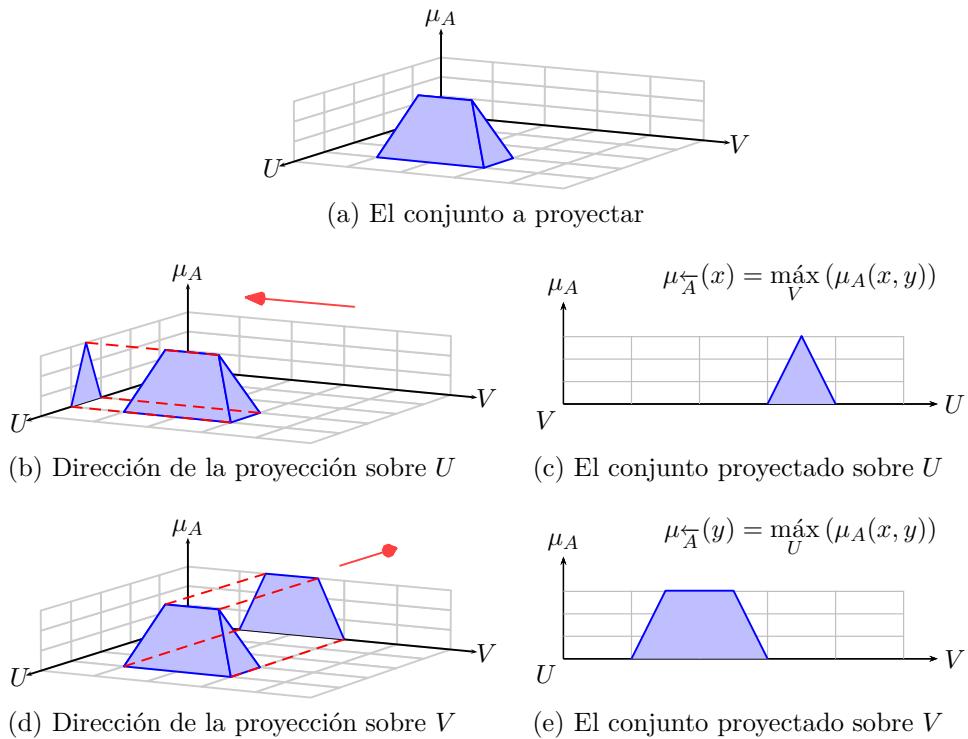


Figura 1.14: Proyección cilíndrica sobre universos continuos

Para proyectar A sobre U aplicamos la ecuación 1.8. El proceso se puede visualizar como si se aplicara una fuente de luz plana en la dirección que muestra la figura 1.14b, y la proyección fuese la sombra obtenida, que se muestra en la figura 1.14c.

Por su parte, las figuras 1.14d y 1.14e ilustran la proyección sobre V .

Capítulo 2

Operaciones entre conjuntos difusos

En este capítulo se presentan las definiciones básicas que permiten hacer operaciones entre conjuntos difusos. Así como la noción de conjunto difuso puede entenderse como una extensión de la noción de conjunto concreto, las operaciones entre conjuntos difusos pueden entenderse como una extensión de las operaciones entre conjuntos concretos. Por esa razón, la estrategia que utilizaremos para presentarlas consiste en, primero, recordar cómo son las operaciones entre conjuntos concretos para después efectuar la extensión a los conjuntos difusos.

En la sección 2.1 se exploran las operaciones básicas de Unión, Intersección y Complemento. Allí encontraremos que hay más de una forma de extender esas operaciones a los conjuntos difusos, lo que nos lleva a la definición de los operadores denominados \mathcal{T} -Normas y \mathcal{S} -Normas, que se presentan en la sección 2.2. Posteriormente, en la sección 2.3, se estudian las relaciones entre conjuntos; este concepto resulta clave en la construcción de la teoría de lógica difusa, específicamente en la inferencia de proposiciones (véase la sección 4.3.3).

2.1. Operaciones básicas

Las teoría de conjuntos define tres operaciones básicas: Unión, Intersección y Complemento. A partir de ellas se derivan operaciones más complejas y conceptos más sofisticados. Algo semejante ocurre con la teoría de conjuntos difusos, tal como se verá más adelante. Las operaciones básicas entre conjutnos difusos se presentan en la sección 2.1.2, pero antes de ello, en la sección 2.1.1, se presenta un resumen de las operaciones entre conjuntos concretos.

2.1.1. Operaciones sobre conjuntos concretos

En esta sección presentamos, a manera de recordatorio, las operaciones básicas entre conjuntos concretos. Para ello, utilizaremos el universo de discurso U de la figura 2.1a que consta de algunos deportes. Sobre ese universo, definimos los conjuntos A y B :

$$A = \left\{ \begin{array}{c} \text{Fútbol} \\ \text{Atletismo} \\ \text{Ciclismo} \\ \text{Natación} \\ \text{Snowboard} \end{array} \right\}$$

$$B = \left\{ \begin{array}{c} \text{Fútbol} \\ \text{Atletismo} \\ \text{Surf} \\ \text{Ajedrez} \end{array} \right\}$$

En la figura 2.1 se ilustra, mediante diagramas de Venn, las tres operaciones básicas:

Unión: la unión de dos conjuntos, denotada por el operador \cup , es un nuevo conjunto cuyos elementos están contenidos en uno cualquiera de los dos conjuntos originales.

Intersección: la intersección de dos conjuntos, denotada por el operador \cap , es un nuevo conjunto cuyos elementos están contenidos simultáneamente en los dos conjuntos originales.

Complemento: el complemento de un conjunto, denotado por el operador $'$, es un nuevo conjunto cuyos elementos no están contenidos en el conjunto original.

Es posible representar también esas operaciones mediante los grados de pertenencia. La tabla 2.1 muestra los grados de pertenencia de cada elemento de U a los conjuntos A y B , y al resultado de aplicar sobre ellos las operaciones básicas. Para ilustrar el contenido de ese tabla, consideremos el elemento ciclismo. Dicho elemento pertenece a A pero no pertenece a B , es decir:

$$\mu_A \left(\begin{array}{c} \text{Ciclista} \\ \text{Bicicleta} \end{array} \right) = 1 \quad \mu_B \left(\begin{array}{c} \text{Ciclista} \\ \text{Bicicleta} \end{array} \right) = 0$$

En consecuencia, el ciclismo sí pertenece a la unión de los dos conjuntos, pero no pertenece a su intersección:

$$\mu_{A \cup B} \left(\begin{array}{c} \text{Ciclista} \\ \text{Bicicleta} \end{array} \right) = 1 \quad \mu_{A \cap B} \left(\begin{array}{c} \text{Ciclista} \\ \text{Bicicleta} \end{array} \right) = 0$$

Además, también podemos afirmar que el ciclismo no pertenece al complemento de A , pero sí al complemento de B :

$$\mu_{A'} \left(\begin{array}{c} \text{Ciclista} \\ \text{Bicicleta} \end{array} \right) = 0 \quad \mu_{B'} \left(\begin{array}{c} \text{Ciclista} \\ \text{Bicicleta} \end{array} \right) = 1$$

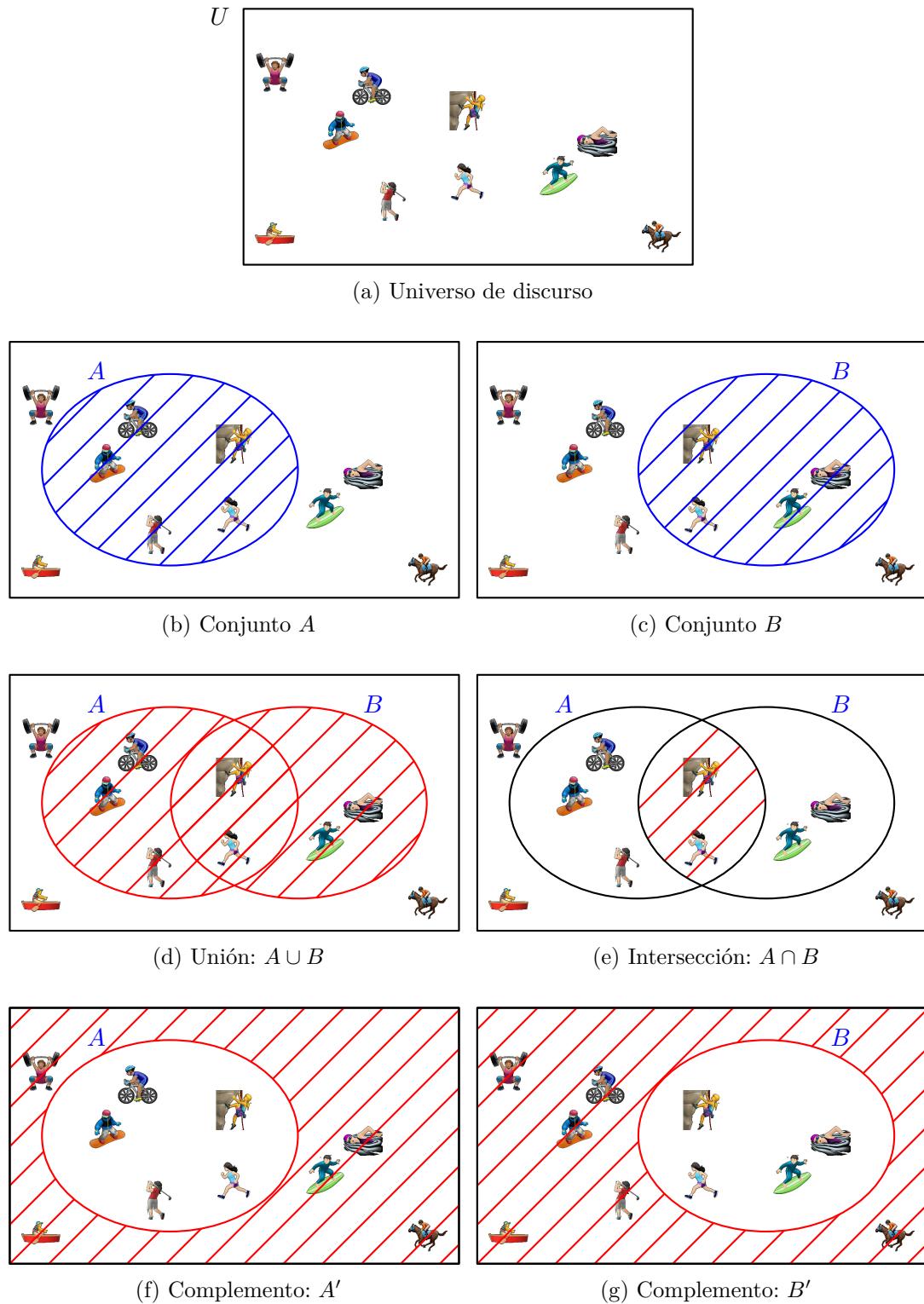


Figura 2.1: Operaciones entre conjuntos concretos. Diagramas de Venn.

Tabla 2.1: Operaciones entre conjuntos concretos. Grados de pertenencia.

x	$\mu_A(x)$	$\mu_B(x)$	$\mu_{A \cup B}(x)$	$\mu_{A \cap B}(x)$	$\mu_{A'}(x)$	$\mu_{B'}(x)$
	1	1	1	1	0	0
	1	1	1	1	0	0
	0	1	1	0	1	0
	0	0	0	0	1	1
	1	0	1	0	0	1
	0	1	1	0	1	0
	0	0	0	0	1	1
	1	0	1	0	0	1
	1	0	1	0	0	1
	0	0	0	0	1	1

La tabla 2.1 nos permite plantear un interrogante: ¿Será posible efectuar las operaciones básicas entre conjuntos mediante operaciones aritméticas entre los grados de pertenencia? Dicho de otra forma, ¿Será posible encontrar unas funciones matemáticas $\mathcal{U}, \mathcal{I}, \mathcal{C}$ que satisfagan las siguientes relaciones?:

$$\left\{ \begin{array}{l} \mu_{A \cup B}(x) = \mathcal{U}(\mu_A(x), \mu_B(x)) \\ \mu_{A \cap B}(x) = \mathcal{I}(\mu_A(x), \mu_B(x)) \\ \mu_{A'}(x) = \mathcal{C}(\mu_A(x)) \end{array} \right. \quad (2.1)$$

La respuesta es afirmativa. Para ilustrarlo, consideremos la tabla 2.2, en la que se muestran todas las posibles combinaciones (4 en total) de valores de $\mu_A(x)$ y $\mu_B(x)$, junto con los grados de pertenencia de x a las tres operaciones básicas. Debajo de esa tabla se ha construido otra que muestra los resultados de tres operaciones aritméticas simples, $\max(x, y)$, $\min(x, y)$, $1 - x$.

Se observa claramente que esas tres operaciones aritméticas reproducen los resultados de las operaciones de unión, intersección y complemento. Por lo tanto, sí es posible efectuar las operaciones básicas entre conjuntos mediante operaciones aritméticas entre los grados de pertenencia así:

Tabla 2.2: Tabla de pertenencia a las tres funciones básicas

$\mu_A(x)$	$\mu_B(x)$	$\mu_{A \cup B}(x)$	$\mu_{A \cap B}(x)$	$\mu_{A'}(x)$
0	0	0	0	1
0	1	1	0	1
1	0	1	0	0
1	1	1	1	0

x	y	$\max(x, y)$	$\min(x, y)$	$1 - x$
0	0	0	0	1
0	1	1	0	1
1	0	1	0	0
1	1	1	1	0

Tabla 2.3: Producto de grados de pertenencia

x	y	xy
0	0	0
0	1	1
1	0	1
1	1	1

$$\begin{cases} \mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)) \\ \mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \\ \mu_{A'}(x) = 1 - \mu_A(x) \end{cases} \quad (2.2)$$

Estas no son las únicas tres funciones que pueden emplearse para calcular la unión, la intersección y el complemento. Este hecho se explora en la sección 2.2. Por lo pronto, basta con mostrar en la tabla 2.3 que el producto también puede usarse para calcular la intersección.

2.1.2. Operaciones Difusas

Para efectuar las tres operaciones básicas (unión, intersección y complemento) entre conjuntos difusos, podemos emplear las expresiones de la ecuación 2.2 sobre los grados de pertenencia que definen a los conjuntos.

Para ilustrarlo, vamos a utilizar como universo de discurso al conjunto de los reales positivos \mathcal{R}^+ . Definimos los conjuntos A y B que se muestran en la figura 2.2a. Para

mayor claridad, los dos conjuntos se han dibujado también por separado en las figuras 2.2b y 2.2c.

Para construir la unión de A y B podemos aplicar 2.2 sobre sus funciones de pertenencia, es decir:

$$\mu_{A \cup B}(x) = \max(\mu_A(x), \mu_B(x)) \quad (2.3)$$

Por otra parte, para construir la intersección de A y B podemos también aplicar 2.2 sobre sus funciones de pertenencia, es decir:

$$\mu_{A \cap B}(x) = \min(\mu_A(x), \mu_B(x)) \quad (2.4)$$

Los complementos de A y B se pueden construir aplicando 2.2 así:

$$\begin{cases} \mu_{A'}(x) &= 1 - \mu_A(x) \\ \mu_{B'}(x) &= 1 - \mu_B(x) \end{cases} \quad (2.5)$$

Los resultados de las operaciones anteriores se muestran en la figura 2.2. Nótese que esta representación gráfica guarda una gran semejanza con los diagramas de Venn que se muestran en la figura 2.1d, en especial para las operaciones de Unión e Intersección.

2.2. \mathcal{T} -Normas y \mathcal{S} -Normas

La Tabla 2.2 muestra que los operadores \max y \min pueden utilizarse para efectuar las operaciones de Unión e Intersección de conjuntos. En esta sección se exploran las propiedades de esos dos operadores, con el propósito de identificar otros operadores que puedan utilizarse para efectuar esas mismas operaciones.

Encontraremos que al generalizar las propiedades del operador \min podemos definir un conjunto versátil de operadores que también pueden usarse para efectuar la intersección. A esos operadores los denominaremos \mathcal{T} -normas (sección 2.2.2). Por su parte, al generalizar las propiedades del operador \max definiremos otro conjunto de operadores que también pueden usarse para efectuar la unión, a los que denominaremos \mathcal{S} -normas (sección 2.2.3).

Antes de hacerlo, presentaremos el concepto de operadores de agregación (sección 2.2.1), que sirve como marco de referencia para entender el papel que cumplen los dos tipos de normas.

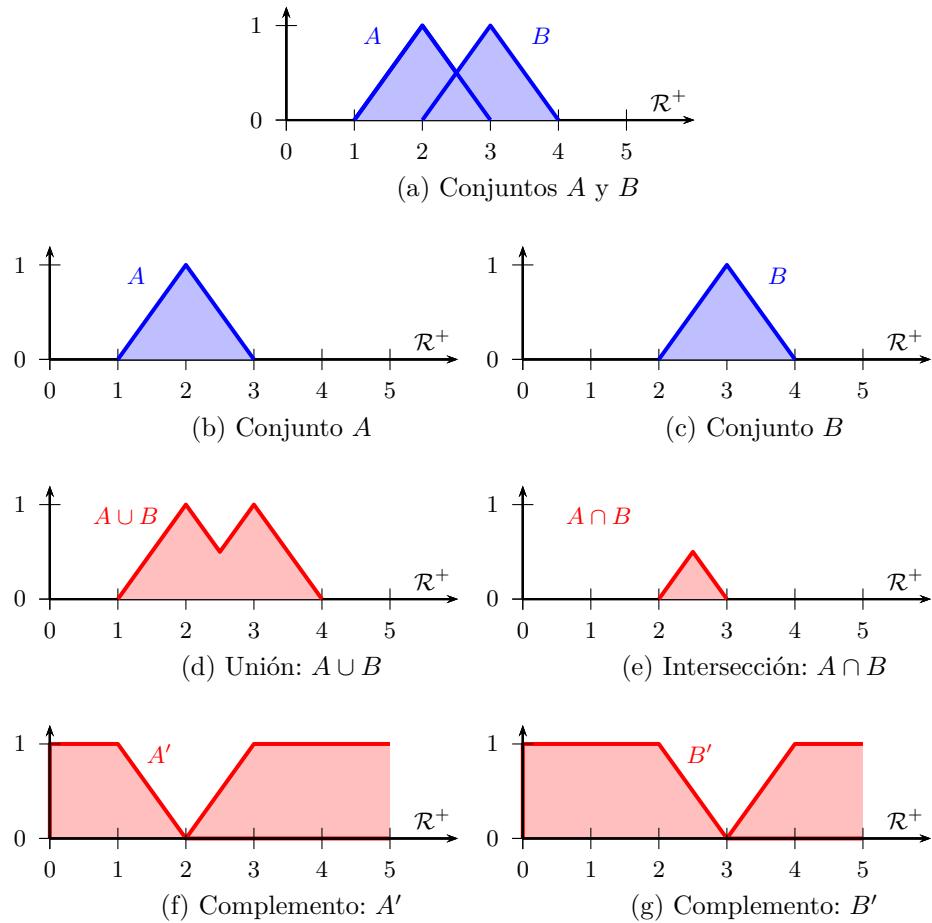


Figura 2.2: Operaciones entre conjuntos difusos.

2.2.1. Operadores de agregación

En el contexto del procesamiento de información, es adecuado enmarcar los operadores máx y mín como *operadores de agregación* (véase [17, 1]). Supongamos que tenemos un conjunto de n números reales x_1, x_2, \dots, x_n y que el valor de cada uno de ellos tiene un significado, es decir, que cada uno de ellos representa una *información*. Un operador de agregación es una función que asigna un número real a ese conjunto de datos, con la intención de generar nueva información:

$$y = Agr(x_1, x_2, \dots, x_n) \quad (2.6)$$

Para que una cierta función sea considerada un operadores de agregación debe satisfacer algunas propiedades:

$$\left\{ \begin{array}{ll} \textbf{Rango y recorrido unitarios:} & Agr : [0, 1]^n \rightarrow [0, 1] \\ \textbf{Identidad:} & Agr(x) = x \\ \textbf{Condición de frontera inferior:} & Agr(0, \dots, 0) = 0 \\ \textbf{Condición de frontera superior:} & Agr(1, \dots, 1) = 1 \\ \textbf{Monotonía no decreciente:} & (x_1, \dots, x_n) \leq (z_1, \dots, z_n) \implies \\ & Agr(x_1, \dots, x_n) \leq Agr(z_1, \dots, z_n) \end{array} \right. \quad (2.7)$$

En la figura 2.3 se muestra el resultado de aplicar los operadores máx y mín sobre dos variables x, y que toman valores en el intervalo $[0, 1]$. Esta figura permite verificar que los dos operadores satisfacen las condiciones anteriores y por tanto son operadores de agregación.

La misma figura permite observar que las dos funciones tienen un gran parecido. En ambos casos la superficie resultante la forman caras con forma de triángulos rectángulos; en ambos casos los triángulos han rotado sobre ejes paralelos a los ejes x y y . La diferencia es que en el operador mín se han rotado desde la parte de abajo de cubo unitario, mientras que en el operador máx lo han hecho desde la parte de arriba.

2.2.2. \mathcal{T} -normas

Estudiemos ahora el operador mín. La figura 2.4 muestra el operador mín en el cubo unitario, y destaca los cuatro puntos que reproducen la operación de intersección entre conjuntos concretos (ver Tabla 2.2). Nótese, además, que existe una simetría respecto al plano $x = y$ o, lo que es igual, que el operador es conmutativo ($\min(x, y) = \min(y, x)$).

Al combinar las propiedades de la ecuación 2.7 con la necesidad de reproducir los cuatro puntos de la figura 2.4 y la condición de conmutatividad se obtiene un tipo especial de operador denominado \mathcal{T} -norma.

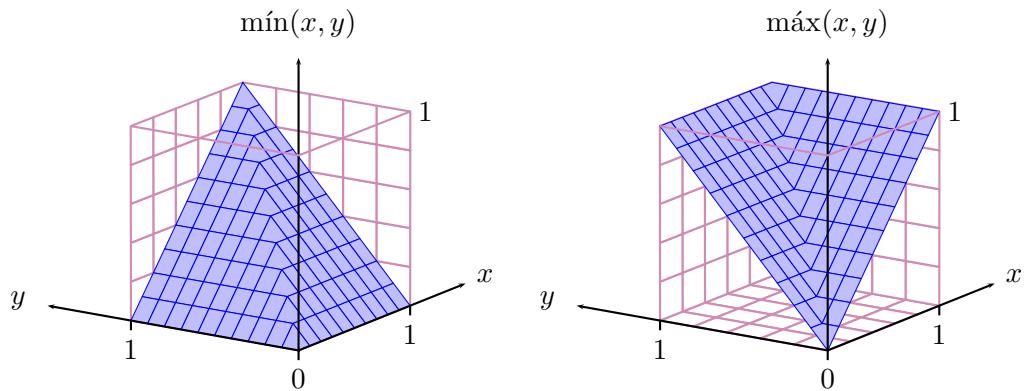


Figura 2.3: Operadores máx y mín en el cubo unitario

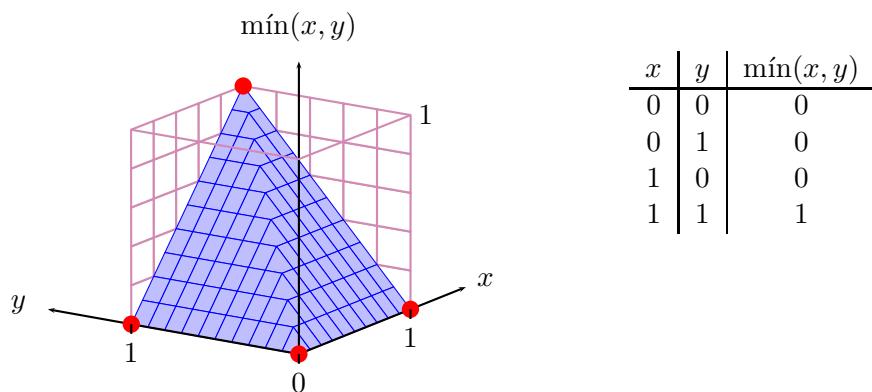
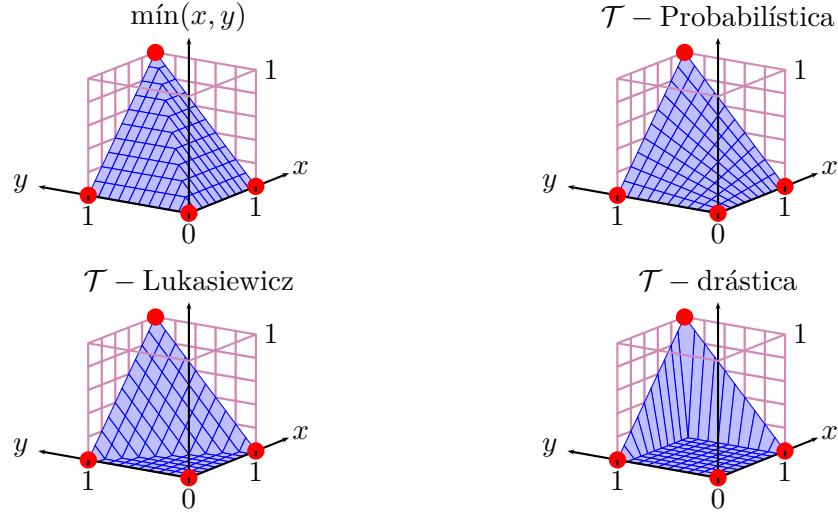


Figura 2.4: Operador mín

Figura 2.5: Ejemplos de \mathcal{T} -Normas

Las \mathcal{T} -normas son operadores de agregación que operan sobre dos variables, es decir:

$$\mathcal{T} : [0, 1] \times [0, 1] \rightarrow [0, 1]$$

y que satisfacen las siguientes propiedades ([43]):

$$\left\{ \begin{array}{ll} \mathcal{T}(x, y) = \mathcal{T}(y, x) & (T1) \text{ Comutativa} \\ \mathcal{T}(x, y) \leq \mathcal{T}(u, v), \text{ si } x \leq u, y \leq v & (T2) \text{ Monotonía no decreciente} \\ \mathcal{T}(x, \mathcal{T}(y, z)) = \mathcal{T}(\mathcal{T}(x, y), z) & (T3) \text{ Asociativa} \\ \mathcal{T}(x, 1) = x & (T3) \text{ Modulativa} \end{array} \right. \quad (2.8)$$

La figura 2.5 muestra 4 ejemplos de \mathcal{T} -normas, cuyas definiciones se pueden ver en la Tabla 2.4. En esa figura se han destacado los cuatro puntos necesarios para calcular la intersección de conjuntos concretos para resaltar que todas las \mathcal{T} -normas cruzan por esos puntos.

Las \mathcal{T} -normas ‘mínimo’ y ‘drástica’ son esenciales. Se puede demostrar que todas las \mathcal{T} -normas son menores o iguales a la del mínimo y mayores o iguales que la de la drástica. Es decir, estas dos normas son casos extremos: todas las otras \mathcal{T} -normas están entre esos dos casos extremos.

Este hecho sugiere una idea: si las \mathcal{T} -normas están comprendidas entre la drástica y la mínima, ¿Podrá diseñarse una \mathcal{T} -norma que dependa de un parámetro, de tal manera

Tabla 2.4: Definiciones de algunas \mathcal{T} -normas y \mathcal{S} -normas

Nombre	\mathcal{T} -norma	\mathcal{S} -norma
min-max	$\min(x, y)$	$\max(x, y)$
Probabilística	xy	$x + y - xy$
Lukasiewicz	$\max(x + y, 1)$	$\min(x + y, 1)$
Drástica	$\begin{cases} x & \text{si } y = 1 \\ y & \text{si } x = 1 \\ 0 & \text{otro caso} \end{cases}$	$\begin{cases} x & \text{si } y = 0 \\ y & \text{si } x = 0 \\ 1 & \text{otro caso} \end{cases}$
Sugeno	$\max\left(\frac{x+y-1+\lambda_S \cdot x \cdot y}{1+\lambda_S}, 0\right)$	$\max(x + y - 1 + \lambda_S \cdot x \cdot y, 1)$
Hamacher	$\frac{x \cdot y}{\gamma + (1-\gamma)(x+y-x \cdot y)}$	$\frac{x+y-x \cdot y-(1-\gamma) \cdot x \cdot y}{1-(1-\gamma) \cdot x \cdot y}$

que al modificar este parámetro cambie su forma, acercándola o alejándola de los dos casos extremos?

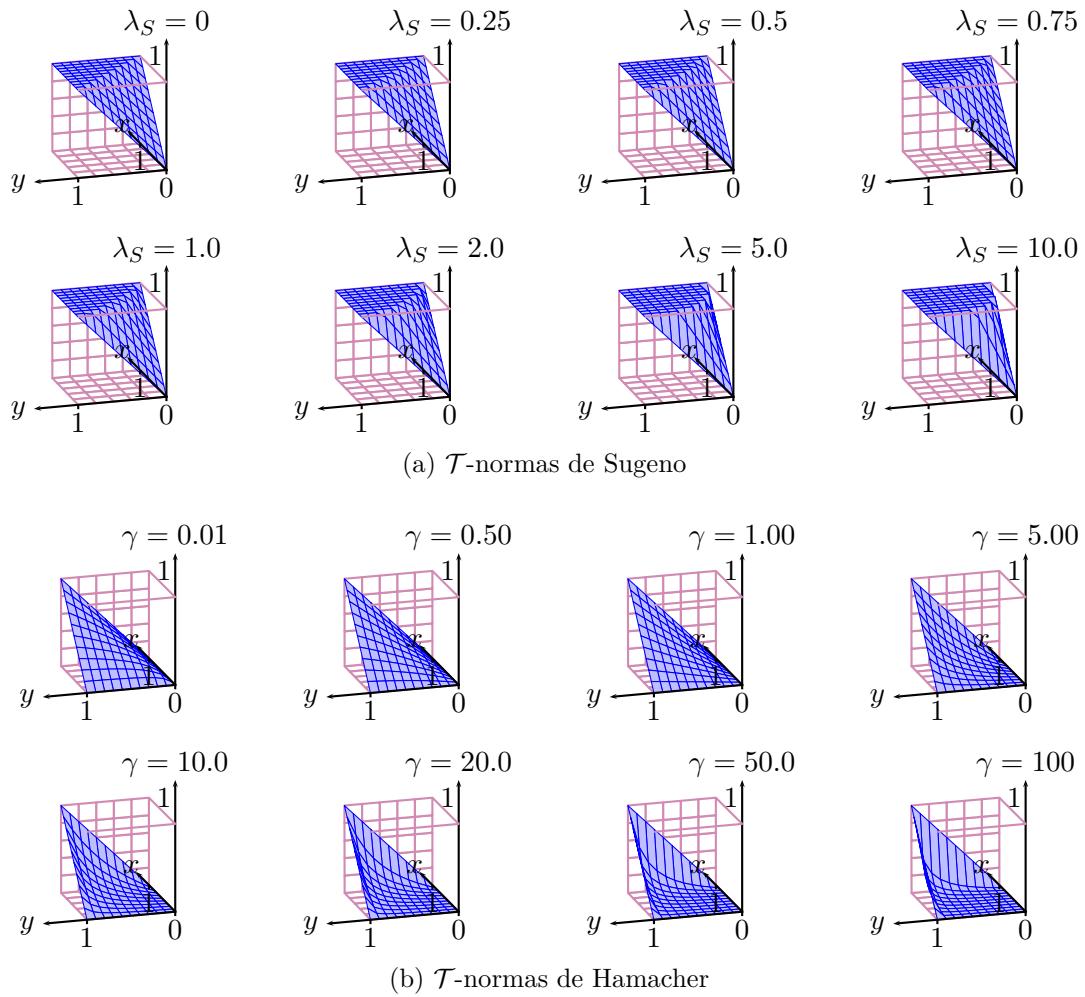
Hay una gran colección de \mathcal{T} -normas parametrizadas, familias de \mathcal{T} -normas, que cumplen tales requisitos. La figura 2.6 muestra dos de esas familias, la de Sugeno y la de Hamacher, cuyas definiciones se encuentran en la Tabla 2.4.

La familia Sugeno depende del parámetro λ_S y la familia Hamacher del parámetro γ . Nótese que al variar esos parámetros se consigue que la norma resultante se acerque o se aleje de los casos extremos (mínimo y drástica). De esta forma, es posible controlar mediante un único parámetro el comportamiento de una \mathcal{T} -norma. Otras definiciones de \mathcal{T} pueden encontrarse en [17].

2.2.3. \mathcal{S} -normas

Estudiemos ahora el operador \max . La figura 2.7 muestra el operador \max en el cubo unitario, y destaca los cuatro puntos que reproducen la operación de unión entre conjuntos concretos (ver Tabla 2.2). Nótese, además, que existe una simetría respecto al plano $x = y$ o, lo que es igual, que el operador es conmutativo ($\max(x, y) = \max(y, x)$).

Al combinar las propiedades de la ecuación 2.7 con la necesidad de reproducir los cuatro puntos de la figura 2.7 y la condición de conmutatividad se obtiene un tipo especial de operador denominado \mathcal{S} -norma.

Figura 2.6: Ejemplos de \mathcal{T} -normas paramétricas

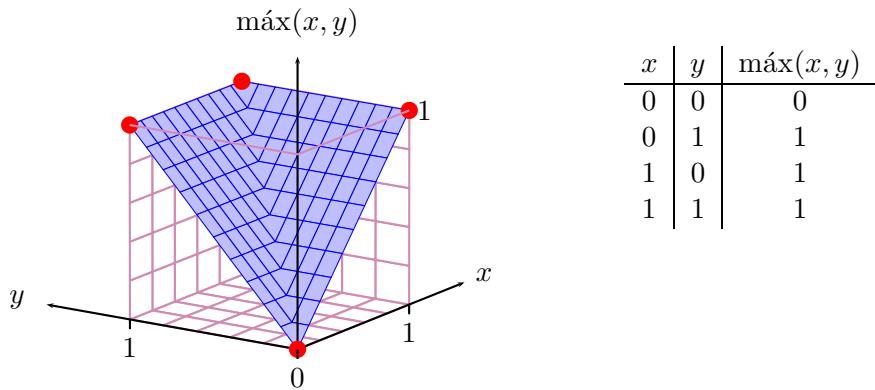


Figura 2.7: Operador máx

Las \mathcal{S} -normas son operadores de agregación que operan sobre dos variables, es decir:

$$\mathcal{S} : [0, 1] \times [0, 1] \rightarrow [0, 1]$$

y que satisfacen las siguientes propiedades:

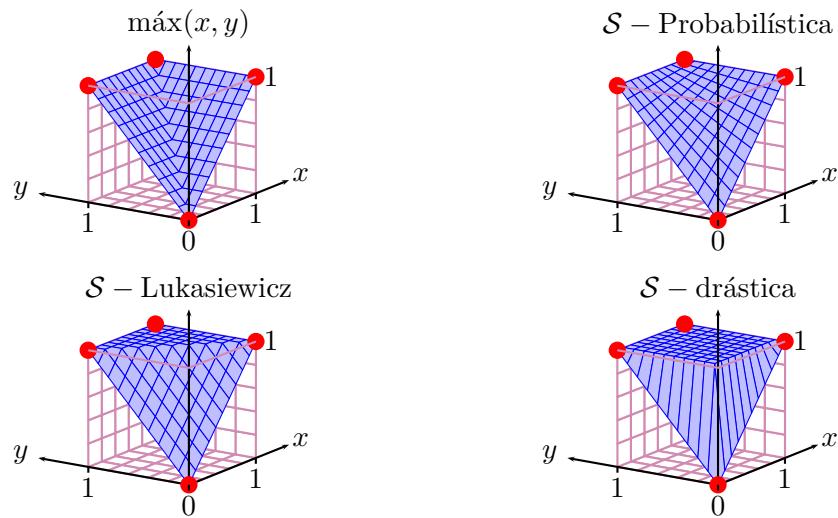
$$\left\{ \begin{array}{ll} \mathcal{S}(x, y) = \mathcal{S}(y, x) & (S1) \text{ Comutativa} \\ \mathcal{S}(x, y) \leq \mathcal{S}(u, v), \text{ si } x \leq u, y \leq v & (S2) \text{ Monotonía no decreciente} \\ \mathcal{S}(x, \mathcal{S}(y, z)) = \mathcal{S}(\mathcal{S}(x, y), z) & (S3) \text{ Asociativa} \\ \mathcal{S}(x, 0) = x & (S3) \text{ Modulativa} \end{array} \right. \quad (2.9)$$

Nótese la similitud entre las ecuaciones 2.8 y 2.9. La única diferencia se encuentra en el módulo: el módulo de las \mathcal{T} -normas es 1, mientras que el de las \mathcal{S} -normas es 0.

Como consecuencia de esa semejanza, es posible construir una \mathcal{S} -norma a partir de un \mathcal{T} -norma y viceversa. Tales normas se denominan *duales*. Esta dualidad justifica que las \mathcal{S} -normas también sean denominadas \mathcal{T} -conormas. Las expresiones que relacionan dos normas T y S duales son las siguientes:

$$\begin{aligned} S(x, y) &= 1 - T(1 - x, 1 - y) \\ T(x, y) &= 1 - S(1 - x, 1 - y) \end{aligned} \quad (2.10)$$

La figura 2.8 muestra 4 ejemplos de \mathcal{S} -normas, cuyas definiciones se pueden ver en la Tabla 2.4. En esa figura se han destacado los cuatro puntos necesarios para calcular la unión de conjuntos concretos para resaltar que todas las \mathcal{S} -normas cruzan por esos puntos.

Figura 2.8: Ejemplos de S -Normas

De nuevo, es posible identificar dos normas que son casos extremos: la máxima y la drástica. todas las otras S -normas están entre esos dos casos extremos. También es posible definir S -normas paramétricas como las que se muestran en la figura 2.9, cuyas definiciones se encuentran en la tabla 2.4

2.3. Relaciones entre conjuntos

Las relaciones entre conjuntos pueden entenderse como conjuntos de relaciones entre sus elementos. Si este nuevo conjunto (el conjunto de las relaciones) es difuso, tendremos una *relación difusa*.

Para abordar este concepto, en la sección 2.3.1 se hace un repaso sobre la noción convencional entre conjuntos, es decir, la relación concreta entre conjuntos concretos; posteriormente, en la sección 2.3.2 se explica cómo se pueden formular relaciones difusas entre conjuntos concretos; en la sección 2.3.3 se da un paso más para construir entre conjuntos difusos. El proceso de composición de relaciones, que permite encadenar relaciones entre más de dos conjuntos, se explica en la sección 2.3.4

2.3.1. Relaciones concretas entre conjuntos concretos

Consideremos ahora los conjuntos concretos A y B que se muestran en la figura 2.10: A está definido en el universo de discurso U (los deportes) mientras que B lo está en V (los animales).

En la misma figura 2.10 se muestra la relación R entre los dos conjuntos. Esa relación es,

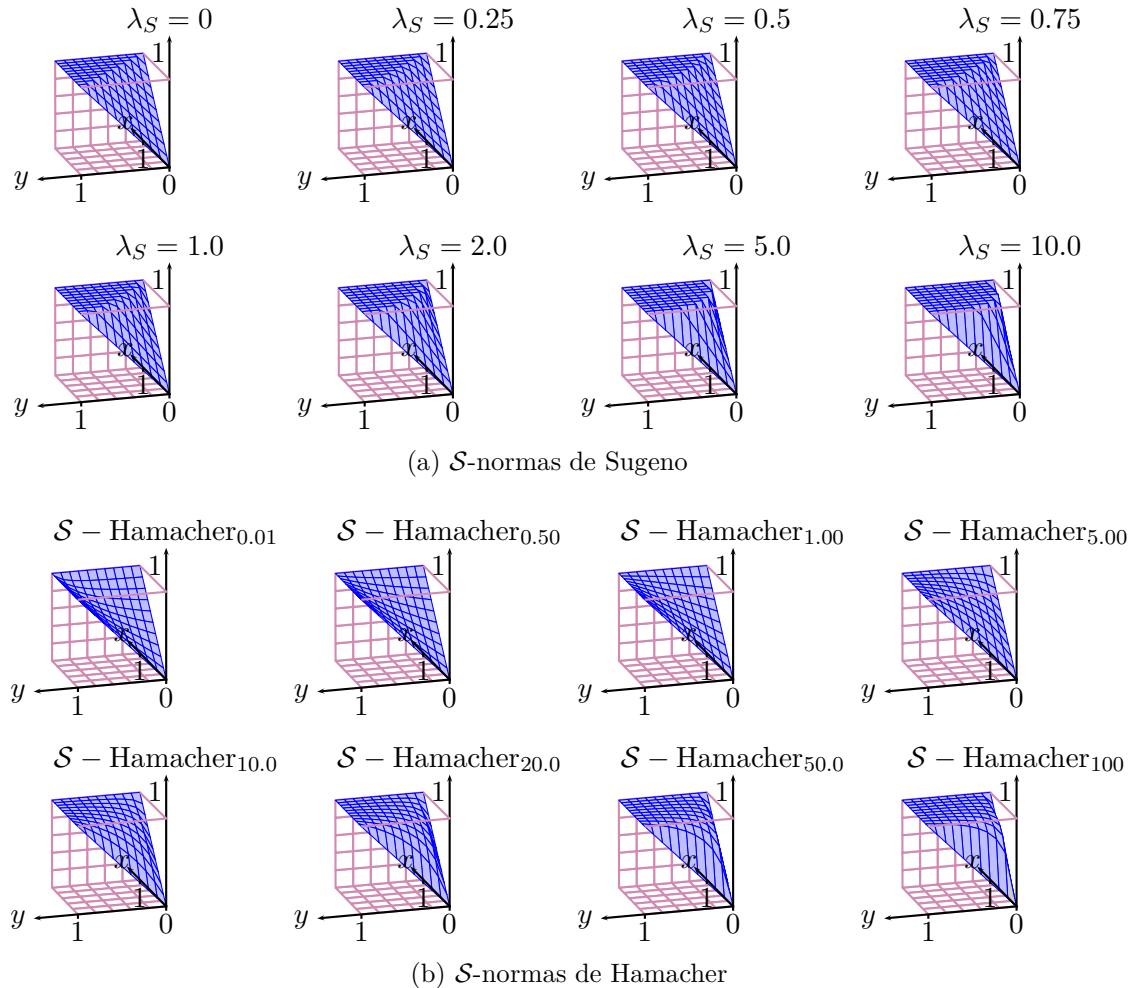
Figura 2.9: Ejemplos de \mathcal{S} -normas paramétricas

Tabla 2.5: Relación concreta entre conjuntos concretos

A \ B		
A		
	1	1
	1	0
	0	0

en realidad, un conjunto de relaciones entre los elementos. Cada relación entre elementos se ha representado por una flecha en el diagrama sagital de la parte inferior. Podemos afirmar, por ejemplo, que el ciclista no se relaciona con ninguno de los elementos del conjunto B, mientras que la atleta se relaciona tanto con el búho como con el flamingo. Es posible que esa relación tenga algún significado, pero desde el punto de vista matemático eso es irrelevante.

Lo que sí es relevante es que cada una de esas relaciones (cada flecha en el diagrama) da origen a una pareja de elementos que está formada por un elemento de A y el correspondiente elemento de B con el cual se relaciona. La relación R es el conjunto de todas las relaciones entre sus elementos, y por tanto puede representarse por el conjunto de esas parejas:

$$R = \left\{ \left(\begin{array}{c} \text{runner} \\ \text{atleta} \end{array}, \begin{array}{c} \text{owl} \\ \text{flamingo} \end{array} \right), \left(\begin{array}{c} \text{runner} \\ \text{atleta} \end{array}, \begin{array}{c} \text{flamingo} \\ \text{flamingo} \end{array} \right), \left(\begin{array}{c} \text{surfer} \\ \text{clown} \end{array}, \begin{array}{c} \text{owl} \\ \text{owl} \end{array} \right) \right\} \quad (2.11)$$

La Tabla 2.5 muestra la misma información pero de otra forma. Cada celda muestra el grado de pertenencia de la pareja correspondiente a la relación R , es decir, al conjunto de sus parejas.

2.3.2. Relaciones difusas entre conjuntos concretos

Dado que las relaciones entre conjuntos concretos pueden representarse mediante conjuntos de parejas, es posible definir relaciones difusas haciendo que esos conjuntos de parejas sean difusos. Para ello, debe relajarse la condición de pertenencia a la relación, permitiendo que sea parcial.

La Tabla 2.6 muestra una relación difusa entre los mismos conjuntos A y B de la figura 2.10. En cada celda se ha consignado el grado de pertenencia a la relación, que es un

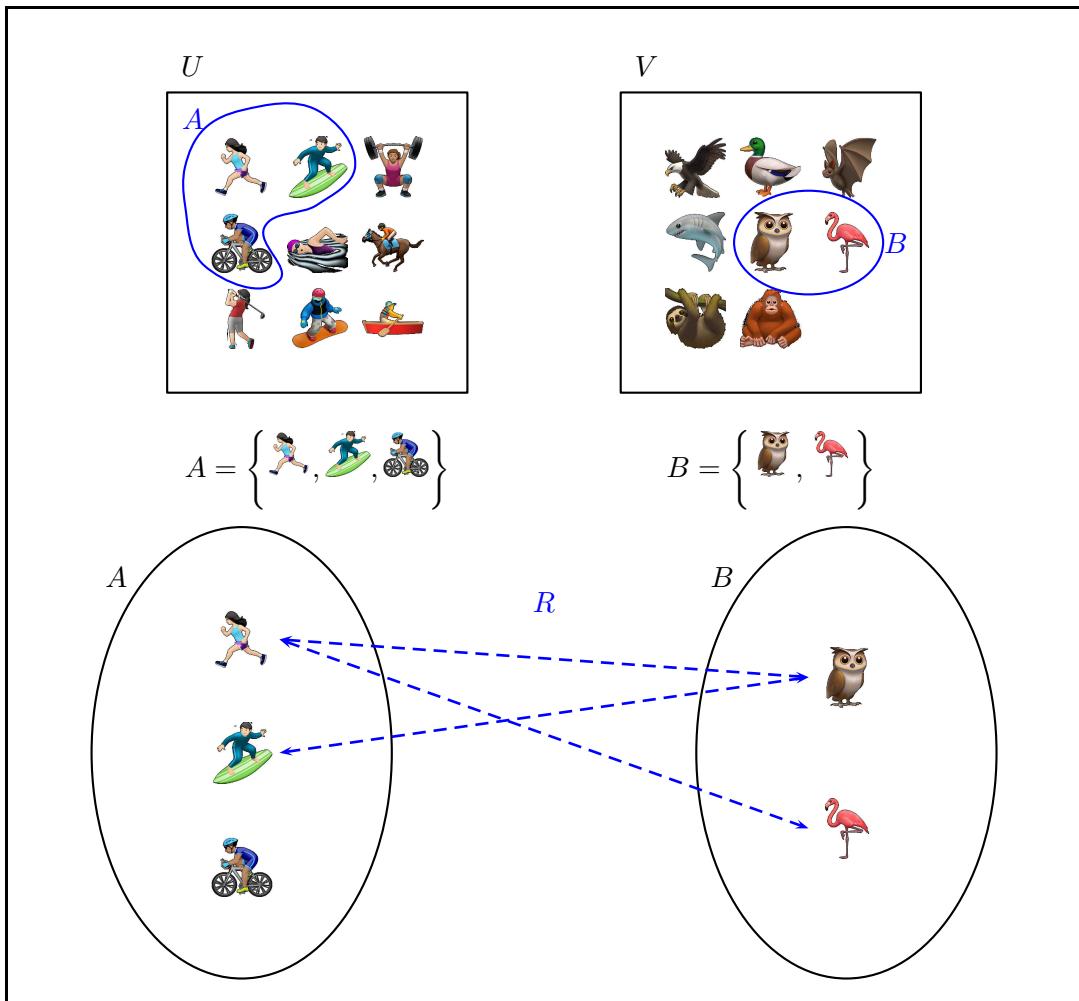


Figura 2.10: Relacion concreta entre conjuntos concretos

Tabla 2.6: Relación difusa entre conjuntos concretos

A \ B		
	0.8	0.6
	1.0	0.7
	0.5	0.2

número en el intervalo $[0, 1]$. Esta relación \tilde{R} se definiría formalmente como el siguiente conjunto:

$$\tilde{R} = \frac{0.8}{\begin{pmatrix} \text{runner}, \text{owl} \end{pmatrix}} + \frac{0.6}{\begin{pmatrix} \text{runner}, \text{flamingo} \end{pmatrix}} + \frac{1.0}{\begin{pmatrix} \text{surfer}, \text{owl} \end{pmatrix}} + \frac{0.7}{\begin{pmatrix} \text{surfer}, \text{flamingo} \end{pmatrix}} + \frac{0.5}{\begin{pmatrix} \text{clown}, \text{owl} \end{pmatrix}} + \frac{0.2}{\begin{pmatrix} \text{clown}, \text{flamingo} \end{pmatrix}} \quad (2.12)$$

Si se consideran la totalidad de los elementos de los universo de discurso U y V , la relación se podría representar tal como se muestra en la Tabla 2.7. Ahora bien, para conjuntos discretos como los que estamos empleando, es posible asignar un número entero a cada elemento del conjunto (el elemento 1, el elemento 2, etc.). Gracias a ello, podemos representar la relación como una matriz. La celda i,j de la matriz tendrá el grado de pertenencia de la pareja formada por los elementos i y j a la relación. En nuestro ejemplo esta matriz resulta ser:

$$\tilde{R} = \begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 & \mathbf{0.8} & \mathbf{0.6} & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & \mathbf{1.0} & \mathbf{0.7} & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & \mathbf{0.5} & \mathbf{0.2} & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix} \quad (2.13)$$

El grado de pertenencia a una relación puede interpretarse como una medida de la fortaleza de esa relación. Un grado de pertenencia cercano a 1 puede interpretarse como

Tabla 2.7: Relación difusa entre conjuntos concretos

una relación ‘fuerte’ entre los elementos.

2.3.3. Relación entre conjuntos difusos

En esta sección exploramos el significado de la afirmación ‘el conjunto \tilde{C} está relacionado con el conjunto \tilde{D} ’. Esta exploración es importante, porque cada conjunto difuso puede representar un concepto, en cuyo caso dicha afirmación sería equivalente a decir que ‘los conceptos \tilde{C} y \tilde{D} están relacionados’. Para hacerlo, consideramos que tanto \tilde{C} como \tilde{D} son conjuntos difusos, como, por ejemplo:

$$\begin{aligned}\tilde{C} &= \left\{ \frac{0.2}{\text{surf}} + \frac{0.6}{\text{murciélagos}} + \frac{1.0}{\text{bicicleta}} \right\} \\ \tilde{D} &= \left\{ \frac{0.3}{\text{murciélagos}} + \frac{0.7}{\text{surf}} + \frac{1.0}{\text{ búho}} + \frac{0.5}{\text{flamingo}} \right\}\end{aligned}\tag{2.14}$$

La afirmación ‘el conjunto \tilde{C} está relacionado con el conjunto \tilde{D} ’ se interpreta como el producto cartesiano de los conjuntos \tilde{C} y \tilde{D} . Es decir, es un nuevo conjunto difuso \tilde{R} de parejas de elementos de \tilde{C} y de \tilde{D} .

$$\begin{aligned}\tilde{R} &= \tilde{C} \times \tilde{D} \\ \tilde{R} &\subset U \times V\end{aligned}\tag{2.15}$$

El grado de pertenencia de cada pareja a la relación se calcula aplicando el operador \min a los frados de pertenencia de cada elemento a los conjuntos o, de forma más general, aplicando cualquier \mathcal{T} -norma

$$\begin{aligned}\mu_{\tilde{R}}(x, y) &= \min(\mu_{\tilde{C}}(x), \mu_{\tilde{D}}(y)) \\ \mu_{\tilde{R}}(x, y) &= \mathcal{T}(\mu_{\tilde{C}}(x), \mu_{\tilde{D}}(y))\end{aligned}\tag{2.16}$$

La figura 2.11 ilustra la situación. El conjunto \tilde{C} está definido sobre el universo U (los deportes) mientras que el conjunto \tilde{D} lo está en V (los animales). Los grados de pertenencia se muestran en la misma figura. Tomemos, por ejemplo, la pareja formada por el *surf* y el murciélagos. El grado de pertenencia a \tilde{C} del elemento *surf* es 0.2 y el del murciélagos al \tilde{D} es 0.3. Tomamos el valor más pequeño (0.2) para representar la relación entre el *surf* y el murciélagos.

Para conjuntos discretos como los que estamos empleando, también podemos representar la relación como una matriz, semejante a la que se muestra en la ecuación 2.13. En nuestro ejemplo esta matriz resulta ser:

\tilde{R}	V								
U	$\begin{array}{c} \tilde{D} \\ \diagdown \\ \tilde{C} \end{array}$	0.0	0.0	0.3	0.7	1.0	0.5	0.0	0.0
	0.0								
	0.2			0.2	0.2	0.2	0.2		
	0.6			0.3	0.6	0.6	0.5		
	1.0			0.3	0.7	1.0	0.5		
	0.0								
	0.0								
	0.0								
	0.0								
	0.0								

Figura 2.11: Relación entre conjuntos difusos

$$\tilde{R} = \begin{bmatrix} 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & \mathbf{0.2} & \mathbf{0.2} & \mathbf{0.2} & \mathbf{0.2} & 0.0 & 0.0 \\ 0.0 & 0.0 & \mathbf{0.3} & \mathbf{0.6} & \mathbf{0.6} & \mathbf{0.5} & 0.0 & 0.0 \\ 0.0 & 0.0 & \mathbf{0.3} & \mathbf{0.7} & \mathbf{1.0} & \mathbf{0.5} & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 & 0.0 \end{bmatrix} \quad (2.17)$$

2.3.4. Composición de relaciones difusas

La composición de relaciones es la concatenación de relaciones. La situación se ilustra en la figura 2.12. Existe una relación R entre los conjuntos A y B ; también existe una relación \tilde{S} entre los conjuntos B y C . Lo que queremos es inferir cuál es la relación entre los conjuntos A y C a partir de las relaciones \tilde{R} y \tilde{S} .

La nueva relación se denota por $R \circ S$. Formalmente tenemos que:

$$\begin{array}{c} A \xrightarrow{\tilde{R}} B \xrightarrow{\tilde{S}} C \\ A \xrightarrow{\tilde{R} \circ \tilde{S}} C \end{array} \quad (2.18)$$

En la misma figura 2.12 se muestran los valores numéricos de las relaciones \tilde{R} y \tilde{S} para todas las parejas posibles. Utilizaremos esos valores para ilustrar cómo se obtiene la relación compuesta entre un elemento de A y uno de C .

Consideremos el caso en que tomamos el elemento atletismo de A y el elemento cruasán de C . Este caso se ilustra en la figura 2.13. En esa figura se han trazado las relaciones que llevarían desde el atletismo hasta el cruasán. Hay 2 caminos distintos: uno a través del búho y otro a través del flamenco.

El procedimiento consiste en encontrar cuál de los dos caminos es mejor, es decir, cuál de los dos caminos permite establecer una relación más fuerte. Para cada camino, la relación total será la del eslabón más débil.

Por ejemplo, el camino que pasa por el búho, tiene grados de pertenencia 0.2 y 0.3; por tanto, la relación total de ese camino tiene una fortaleza de 0.2. Por otra parte, el camino que pasa por el flamenco tiene un camino con grados de pertenencia 0.7 y 0.4 y por tanto la fortaleza de esa relación es de 0.4. ¿Qué camino tomamos? El mejor, el que pasa por el flamenco. La relación entre los 2 extremos, el atletismo y el cruasán es de 0.4.

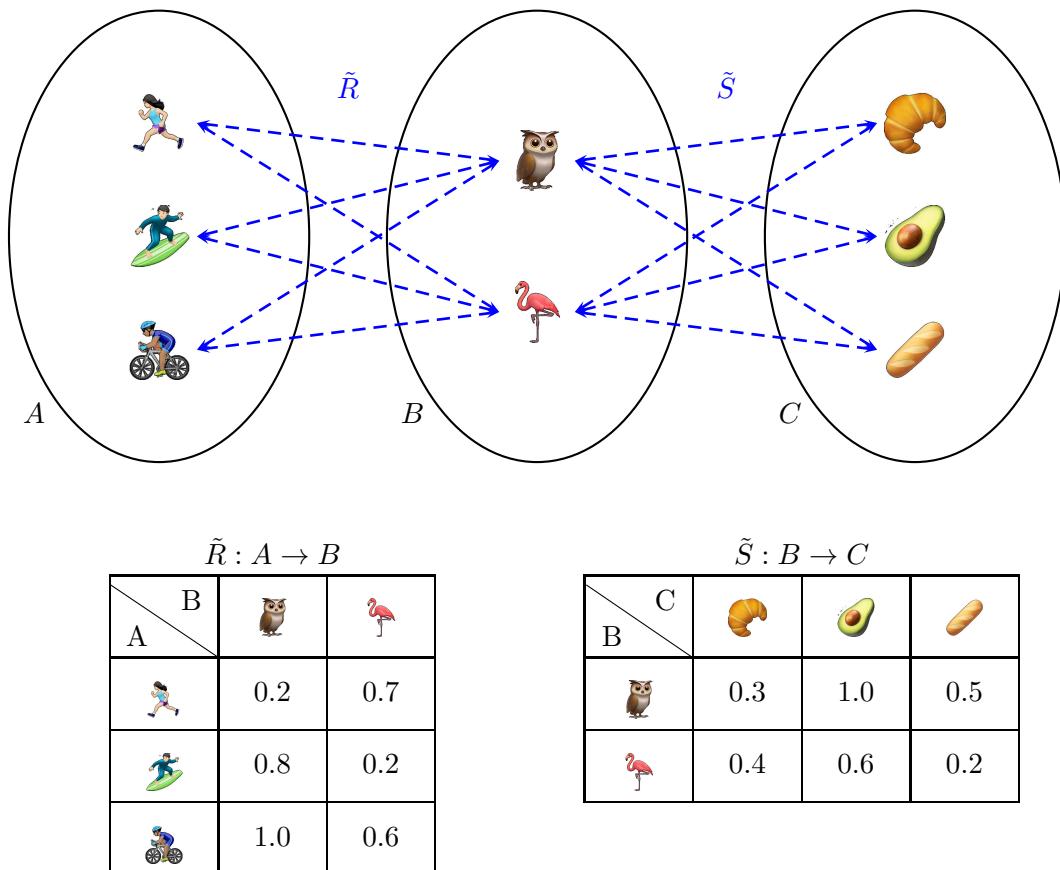


Figura 2.12: Composición de relaciones

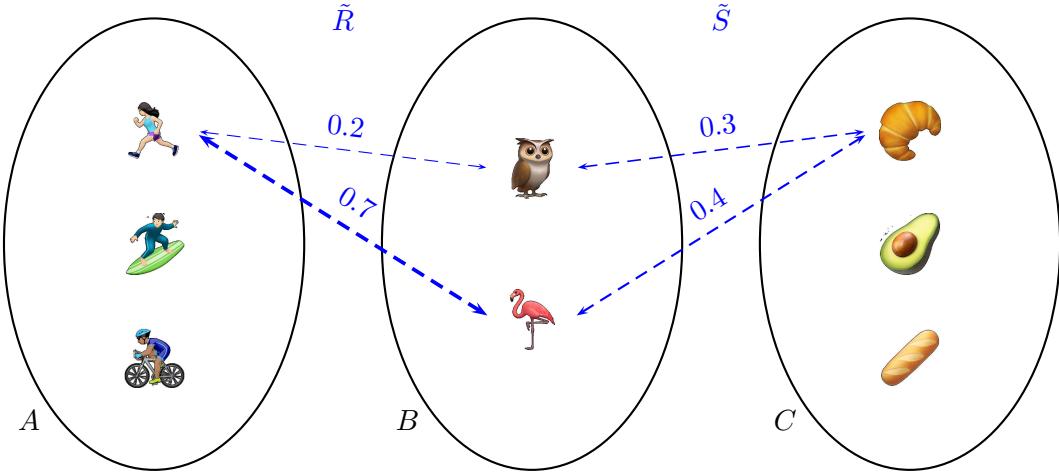


Figura 2.13: Composición de relaciones. Un ejemplo.

Nótese que la operación que hemos realizado es una combinación de los operadores máx y mín que puede formalizarse así:

$$(\tilde{R} \circ \tilde{S})_{i,j} = \max_{k \in B} (\min(\tilde{R}_{i,k}, \tilde{S}_{k,j})) \quad (2.19)$$

Esta expresión puede generalizarse empleando cualquier combinación de \mathcal{T} -normas y \mathcal{S} -normas:

$$(\tilde{R} \circ \tilde{S})_{i,j} = \mathcal{S}_{k \in B} (\mathcal{T}(\tilde{R}_{i,k}, \tilde{S}_{k,j})) \quad (2.20)$$

Las expresiones de las ecuaciones 2.19 y 2.20 pueden sistematizarse al emplear matrices para representar las relaciones \tilde{R} y \tilde{S} . Estas matrices pueden representar relaciones difusas de conjuntos concretos, como en la ecuación 2.13 o relaciones entre conjuntos difusos, como en la ecuación 2.17.

Para ilustrar este procedimiento, consideremos de nuevo el ejemplo de las figuras 2.12 y 2.13. La matriz R representa la primera de las relaciones \tilde{R} , la matriz S representa la relación \tilde{S} y la matriz $R \circ S$ representa la relación compuesta $\tilde{R} \circ \tilde{S}$.

Comencemos por efectuar un análisis de las dimensiones de esas matrices. La matriz R es 3×2 , porque el conjunto A tiene 3 elementos y el conjunto B tiene 2 elementos. La matriz S es 2×3 , porque va desde B , que tiene 2 elementos, hasta C , que tiene 3. La matriz compuesta es una matriz 3×3 , porque va desde el A hasta C y ambos conjuntos tienen 3 elementos.

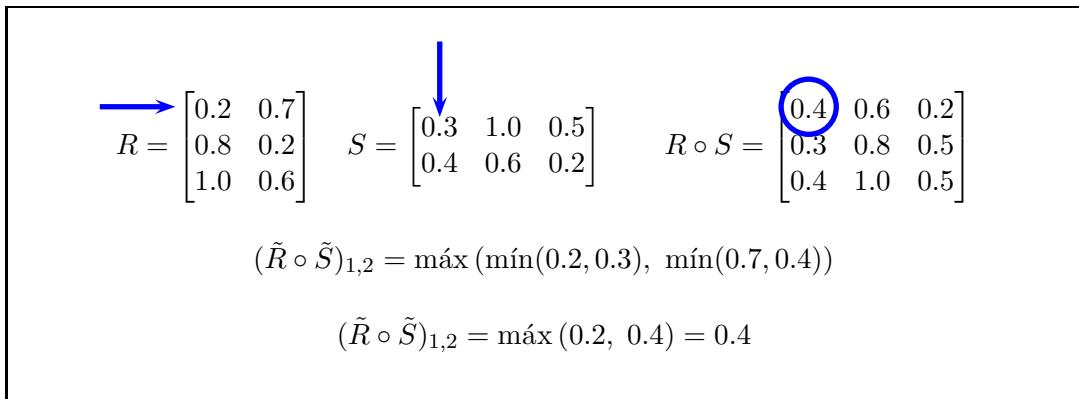


Figura 2.14: Composición de relaciones. Un cálculo.

En la figura 2.14 se indica cómo debe procederse para calcular el término (1,1) de la matriz resultante: es necesario tomar la primera fila de R y la primera columna de S y efectuar el mínimo de cada pareja para luego tomar el máximo de esos resultados.

Este procedimiento recuerda a la multiplicación de matrices: para obtener el elemento (i,j) de la multiplicación de dos matrices es necesario tomar la fila i de la primera matriz y la columna j de la segunda, para efectuar el producto de cada pareja para luego la suma de esos resultados. Si se remplaza ‘producto’ por ‘mínimo’ (o por ‘ \mathcal{T} -norma’) y ‘suma’ por ‘máximo’ (o por ‘ \mathcal{S} -norma’), se obtiene el procedimiento general de composición de relaciones utilizando matrices.

Capítulo 3

Principios de lógica difusa

En un sentido amplio, la *Lógica* puede definirse como el estudio de la forma en que razonamos y argumentamos. Es un área del conocimiento que puede abordarse desde varios marcos conceptuales como los que brindan la filosofía, la matemática y la informática. La matemática, por ejemplo, emplea un lenguaje basado en símbolos que representan operaciones y conceptos cuidadosamente definidas para construir su marco conceptual formalizado.

Este capítulo tiene por propósito presentar los principios de la lógica difusa, que puede entenderse como una formalización de la manera en que argumentamos sobre conceptos que son representados por conjuntos difusos.

Inicialmente se presenta el concepto de Variable Lingüística (sección 3.1), que resulta vital, porque permite interpretar el significado de los conceptos que se representan mediante conjuntos difusos. Posteriormente, en la sección 3.2, se presentan las nociones de lógica difusa propiamente dicha, que en rigor debería denominarse lógica proposicional difusa. Las operaciones que allí se presentan brindan el soporte matemático para los Sistemas Basados en Reglas que se estudian en el capítulo 4.

3.1. Variables lingüísticas

En términos coloquiales, *una variable lingüística es una variable cuyos valores se pueden expresar por palabras*. A manera de ejemplo, consideremos el conjunto de todos los libros que hay en una cierta biblioteca. Si queremos analizar el contenido de los libros podemos definir una variable denominada ‘Tipo de contenido’; los valores que puede tomar esa variable corresponden a los tipos de contenido que puede tener un libro, por ejemplo ‘novela’ o ‘libro de texto’.

En un sentido más restrictivo, en el contexto de los sistemas basados en lógica difusa *una variable lingüística es una variable cuyos valores se pueden expresar por palabras o por números*. Considerando el mismo conjunto de los libros, para referirnos al tamaño de los libros podríamos definir la variable ‘extensión del libro’; dicha variable podría tener dos tipos de valores:

- Palabras que describan la extensión, tales como ‘corto’, ‘largo’, ‘muy largo’, etc.
- Números que indiquen el tamaño, como por ejemplo el número de páginas, el número de palabras, el número de caracteres, etc.

Cualquiera de los dos caminos que utilicemos nos puede brindar información sobre la variable que estamos analizando, la extensión del libro. Dependiendo del contexto del problema que se esté analizando es posible que uno de los dos caminos sea más conveniente que el otro. También es posible que en ocasiones sea necesario combinar información de ambos tipos.

La siguiente es una definición formal del concepto: una variable lingüística \mathcal{V} es un conjunto formado por cuatro elementos:

$$\mathcal{V} : \langle \mathcal{X}, \mathcal{L}\mathcal{X}, \mathcal{U}, \mathcal{M}_{\mathcal{U}} \rangle \quad (3.1)$$

cuyo significado es el siguiente:

- \mathcal{X} es un **Nombre**: una expresión con la que se identifica la variable.
- $\mathcal{L}\mathcal{X}$ es **Conjunto de términos lingüísticos**: una colección de expresiones con las que se pueda valorar la variable.
- \mathcal{U} es un **Universo de discurso**: el conjunto de valores factibles que puede tomar la variable en un cierto dominio. Si el dominio se expresa como un número (lo cuál es muy usual), el universo de discurso resulta ser un intervalo sobre la recta real.
- $\mathcal{M}_{\mathcal{U}}$ es una **Función semántica**: es una relación entre $\mathcal{L}\mathcal{X}$ y \mathcal{U} que permite interpretar los términos de forma numérica y los valores numéricos de forma lingüística. Una función semántica puede definirse a través de conjuntos difusos, como se ilustra en el ejemplo 3.1.

Ejemplo 3.1 (Variable lingüística) Consideremos el conjunto de todos los libros que están guardados en una cierta biblioteca y que estamos interesados en definir una variable que nos ayude a valorar qué tan grande es cada libro. Un **Nombre** adecuado para esa variable es ‘Extensión’, es decir:

$$\mathcal{X} = \text{'Extensión'} \quad (3.2)$$

Para definir el **Conjunto de términos** seleccionamos una colección de términos que nos permiten calificar la extensión de cada libro. Por ejemplo:

$$\mathcal{L}\mathcal{X} = \{\text{'Corto'} , \text{'Intermedio'} , \text{'Largo'} , \text{'Muy largo'} \} \quad (3.3)$$

Para definir el **Universo de discurso** seleccionamos una propiedad que caracterice la extensión de un libro y determinamos los valores mínimo y máximo que puede tener. Supongamos que decidimos que esa propiedad sea el *número de páginas*, y después de hacer una revisión de todos los libros disponibles encontramos que no hay libros con menos de 10 páginas ni más de 450 páginas. El universo de discurso debe abarcar todos esos valores, por ejemplo:

$$\mathcal{U} = [0, 500] \quad (3.4)$$

Para construir la **función semántica** $\mathcal{M}_{\mathcal{U}}$ asignamos a cada término de $\mathcal{L}\mathcal{X}$ un conjunto difuso definido sobre \mathcal{U} . En la figura 3.1 se han trazado las gráficas de las funciones de pertenencia de los conjuntos difusos asociados a cada uno de los cuatro términos. Se ha usado un color diferente para cada conjunto y se han escrito las iniciales del término encima de cada uno de ellos.

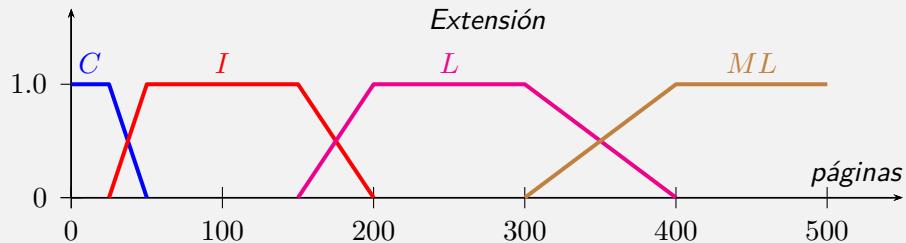


Figura 3.1: Función semántica en el ejemplo 3.1

Consideremos el término ‘Corto’, cuyo conjunto difuso asociado se ha trazado en azul. A través de la función de pertenencia de ese conjunto podemos interpretar qué tan acertado es asignar la etiqueta ‘Corto’ a un cierto libro según su número de páginas.

La función semántica establece un puente entre los términos lingüísticos de $\mathcal{L}\mathcal{X}$ y los valores numéricos de \mathcal{U} . El puente se construye a través del significado. Gracias a la función semántica, tiene sentido el número. Pero además, la función semántica permite

dar un sentido numérico a los términos y por tanto nos permitirá efectuar operaciones matemáticas sobre las palabras.

Si bien es cierto que una gráfica como la de la figura 3.1 corresponde a una función semántica, es usual decir que esa gráfica *es* la variable lingüística. No se trata de una sinécdoca, sino que realmente esa gráfica contiene los cuatro componentes de la variable lingüística. Veamos:

- \mathcal{X} , el nombre de la variable, parece como el título de la gráfica: ‘Extensión’.
- $\mathcal{L}\mathcal{X}$, el conjunto de términos, se infiere de las iniciales que aparecen sobre cada conjunto: ‘Corto’, ‘Intermedio’, ‘Largo’ y ‘Muy largo’.
- \mathcal{U} , el universo de discurso, se descubre al ver el intervalo sobre el que se definen los conjuntos difusos: $[0, 500]$.
- $\mathcal{M}_{\mathcal{U}}$, a función semántica, se visualiza con la forma de las funciones de pertenencia de cada conjunto difuso.

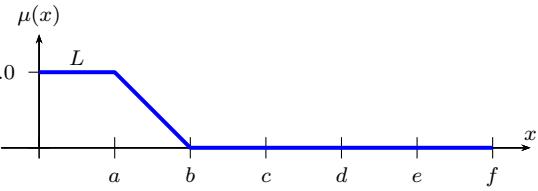
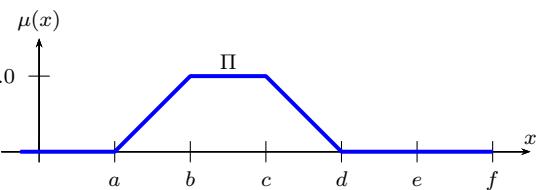
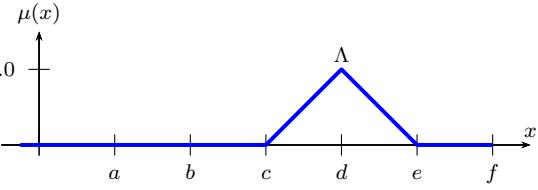
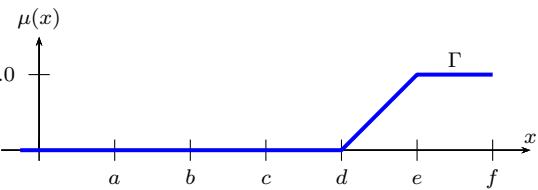
3.1.1. Formas usuales de conjuntos difusos

En la tabla 3.1 se muestra una colección de cuatro tipos de conjuntos difusos que permiten construir de forma práctica funciones semánticas como la que se muestra en la figura 3.1. En todos los casos se ha supuesto un universo de discurso \mathcal{U} definido sobre \mathcal{R} , que corresponde al intervalo $[0, f]$. En la misma tabla se muestra la expresión matemática de las respectivas funciones de pertenencia. Los cuatro tipo de conjuntos son:

- **Conjuntos tipo L:** los valores bajos de x pertenecen totalmente al conjunto mientras que los valores altos están totalmente excluidos del conjunto. Hay una transición gradual mediante una línea recta.
- **Conjuntos tipo II:** los valores bajos y altos de x están excluidos del conjunto. hay un conjunto de valores centrales que pertenecen totalmente al conjunto y dos transiciones graduales mediante líneas rectas.
- **Conjuntos tipo Λ :** es semejante al conjunto tipo II, con la diferencia de que sólo existe un valor de x que pertenece totalmente al conjunto.
- **Conjuntos tipo Γ :** los valores bajos de x están totalmente excluidos del conjunto mientras que los valores altos pertenecen totalmente al conjunto. Hay una transición gradual mediante una línea recta.

Debido a su forma, los conjuntos tipo II también se denominan *Conjuntos trapezoidales*. La función de pertenencia de un conjunto trapezoidal está completamente definida por

Tabla 3.1: Tipos de conjuntos difusos rectos

	$\mu_L(x) = \begin{cases} 1 & x < a \\ \frac{b-x}{b-a} & a \leq x < b \\ 0 & b \leq x \end{cases}$
	$\mu_{\text{II}}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & b \leq x < c \\ \frac{x-d}{c-d} & c \leq x < d \\ 0 & d \leq x \end{cases}$
	$\mu_{\Lambda}(x) = \begin{cases} 0 & x < c \\ \frac{x-c}{d-c} & c \leq x < d \\ \frac{x-e}{d-e} & d \leq x < e \\ 0 & e \leq x \end{cases}$
	$\mu_{\Gamma}(x) = \begin{cases} 0 & x < d \\ \frac{x-d}{e-d} & d \leq x < e \\ 1 & e \leq x \end{cases}$

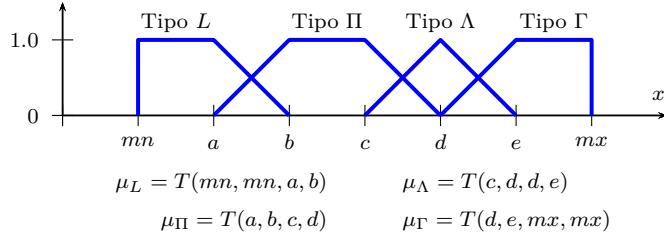


Figura 3.2: Tipos de conjuntos difusos definidos como trapecios

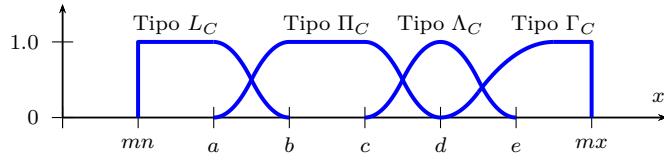


Figura 3.3: Tipos de conjuntos difusos curvos

los cuatro vértices del trapecio. Esto permite definir la siguiente nomenclatura para simplificar la expresión de su función de pertenencia:

$$T(a, b, c, d) = \mu_{\Pi}(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & b \leq x < c \\ \frac{x-d}{c-d} & c \leq x < d \\ 0 & d \leq x \end{cases} \quad (3.5)$$

Nótese que los cuatro casos de la Tabla 3.1 pueden visualizarse como conjuntos trapezoidales. La Figura 3.2 muestra la forma de expresar los cuatro tipos de conjuntos como trapecios cuando el universo de discurso \mathcal{U} es un intervalo genérico $[mn, mx]$.

También es posible construir versiones suavizadas de los tipos de conjuntos trapezoidales. Para ello basta con definir curvas suaves que remplacen las rectas mediante las que se realizan las transiciones. Hay varias formas de definir tales curvas suaves, como por ejemplo a partir de exponenciales, de la función arcotangente, de la campana gaussiana, etc. La figura 3.3 muestra la versión suavizada de los tipos de conjunto usuales.

En la figura 3.4 se muestra un tipo de suavización que se basa en dos tramos curvos.

Consideremos la curva ascendente que va desde $(a, 0)$ hasta $(c, 1)$ definida por:

$$\mu_c(x) = \begin{cases} \lambda \left(\frac{x-a}{\underline{\chi}-a} \right)^\gamma & a \leq x \leq \underline{\chi} \\ 1 - (1-\lambda) \left(\frac{x-b}{\underline{\chi}-b} \right)^\gamma & \underline{\chi} \leq x \leq b \end{cases} \quad (3.6)$$

$\mu_c(x)$ depende de los parámetros λ, γ . Los valores que pueden tomar estos parámetros están dados por:

$$\begin{aligned} \lambda &\in [0, 1] \\ \gamma &\in [1, \infty) \end{aligned} \quad (3.7)$$

El efecto de dichos parámetros sobre la forma de la curva se observa en la figura 3.5 y se explica a continuación:

- λ es un parámetro que controla el valor de $\underline{\chi}$, que es el punto en donde termina la primera curva y comienza la segunda:

$$\underline{\chi} = a + \lambda(b - a) \quad (3.8)$$

Si $\lambda = 0.5$ el cambio sucede justo en la mitad del intervalo $[a, b]$; para los casos en que $\lambda < 0.5$ el cambio sucede antes de la mitad y para los casos en que $\lambda > 0.5$ el cambio sucede después de la mitad. Al aumentar λ desde 0 a 1 el cambio se aleja de a y se acerca a b .

- λ es también el valor que toma la función en el momento del cambio. En efecto, al calcular $\mu_c(\underline{\chi})$ con cualquiera de los dos tramos de 3.6 se obtiene:

$$\mu_c(\underline{\chi}) = \lambda \quad (3.9)$$

- γ es un parámetro que controla la curvatura. Si $\gamma = 1$ las dos curvas resultan ser rectas con la misma pendiente, pero al incrementar γ la curva se aleja cada vez más de la recta.
- γ también controla qué tan empinadas son las curvas alrededor del punto en donde cambian. La derivada de ambas curvas en $x = \underline{\chi}$ resulta ser:

$$\left. \frac{d\mu_c(x)}{dx} \right|_{x=\underline{\chi}} = \frac{\gamma}{b-a} \quad (3.10)$$

Lo que hemos dicho aquí para la curva ascendente que se muestra en la figura 3.4 se aplica también para la curva descendente que se muestra en la misma figura. La demostración la omitimos, pero la simetría de las curvas debería bastar para ello. En

las curvas descendentes, para los casos en que $\lambda > 0.5$ el cambio sucede antes de la mitad. Al aumentar λ desde 0 a 1 el cambio se aleja de b y se acerca a a .

Es muy importante destacar que la derivada de $\mu_c(x)$ es cero tanto en a como en c . Esto implica si construimos las funciones de transferencia de los conjuntos difusos que se muestran en la figura 3.3 usando $\mu_c(x)$, tanto las funciones de transferencia como sus derivadas serán continuas.

3.1.2. Formas usuales de las variables lingüísticas

Al momento de construir una variable lingüística, una de las primeras preocupaciones es la definición del número de términos lingüísticos. En la mayoría de las aplicaciones se requiere que los términos sean interpretables (comprendibles) por personas. En esos casos, lo usual es que el número de términos sea de 3 a 5, y en todo caso nunca de más de 7. No obstante, para aquellas aplicaciones en las que no se requiere que la variable lingüística sea interpretable por personas el número de términos puede ser mucho mayor.

Otro de los aspectos a decidir es la forma de los conjuntos difusos asociados a cada término. La figura 3.6 muestra cuatro formas usuales con las que se diseñan las variables lingüísticas. En los extremos se emplean conjuntos tipo L y Γ , mientras que en la región central se utilizan conjuntos tipo Π o Λ , bien sea rectos o curvos. El uso de trapecios o triángulos incide en el desempeño numérico de los sistemas en donde se use la variable, lo que determina la selección.

La figura 3.6 puede dar la impresión errónea de que los conjuntos difusos deben estar homogéneamente espaciados en el universo de discurso. Esto no necesariamente es así, como puede verse en el ejemplo 3.1. Sin embargo, si efectivamente se desea construir unos conjuntos homogéneamente espaciados, se puede seguir el procedimiento que se explica a continuación.

Supongamos que se desean construir n conjuntos sobre un universo de discurso \mathcal{U} que sea el intervalo $[\mathcal{U}_L, \mathcal{U}_R]$ y que estos conjuntos tengan la forma de los que se muestran en la Figura 3.6a o en la Figura 3.6b. En ese caso, los conjuntos serán trapecios cuyos vértices a, b, c, d deben calcularse como se muestra en la Tabla 3.2, en donde i es el número del conjunto y δ está dado por

$$\delta = \frac{\mathcal{U}_R - \mathcal{U}_L}{2n - 1} \quad (3.11)$$

Si, por el contrario, se desea que los conjuntos tengan la forma que se muestra en la Figura 3.6c o en la Figura 3.6d entonces los vértices deben calcularse como se muestra en la Tabla 3.3, y δ está dado por

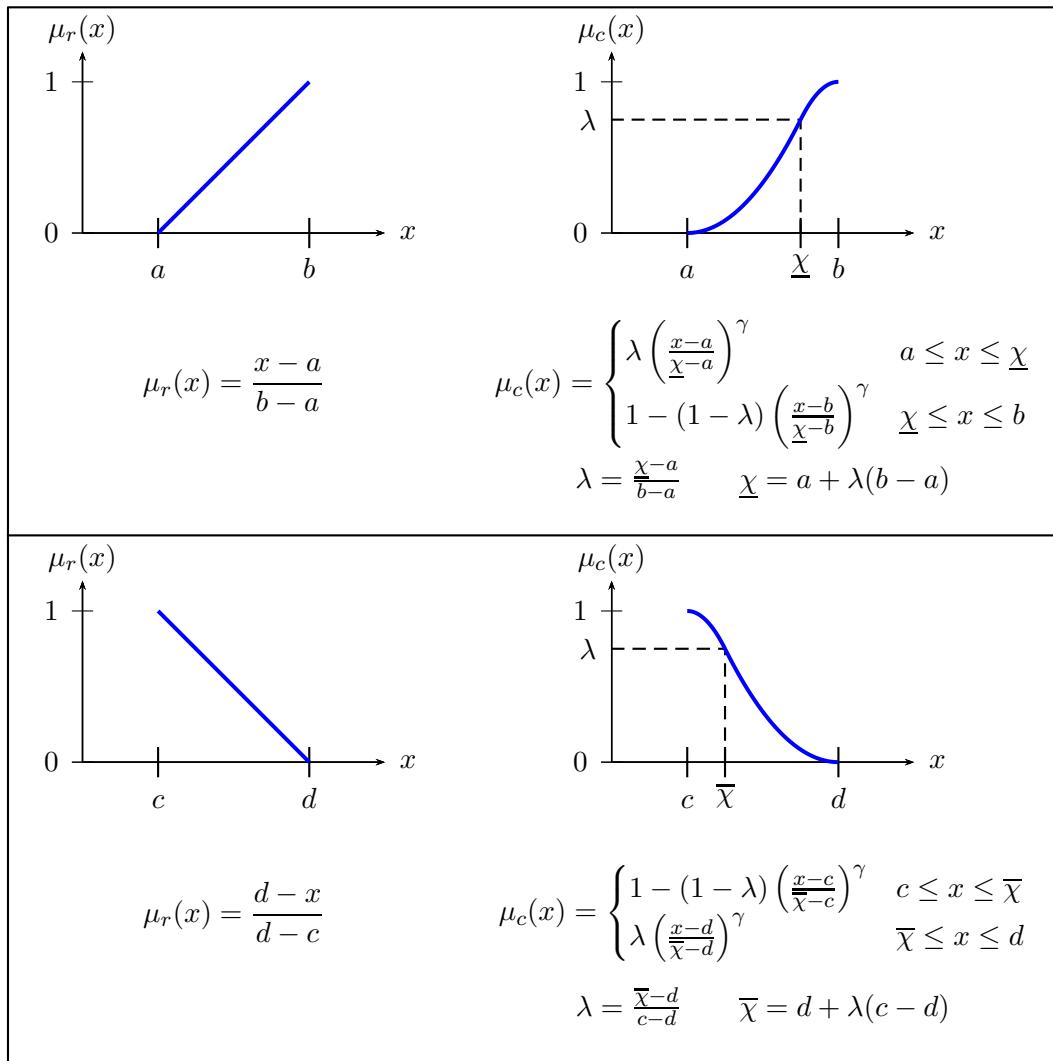
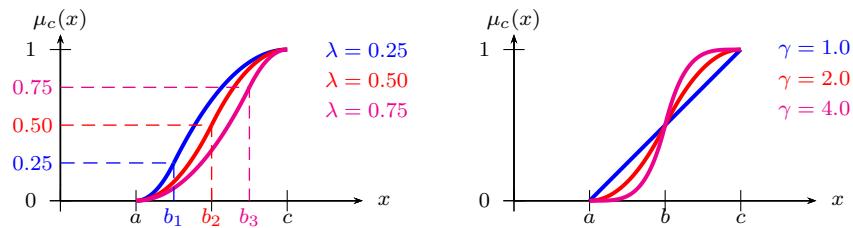


Figura 3.4: Transiciones rectas y curvas para funciones de pertenencia

Figura 3.5: Efecto de los parámetros γ y λ en la curvatura.

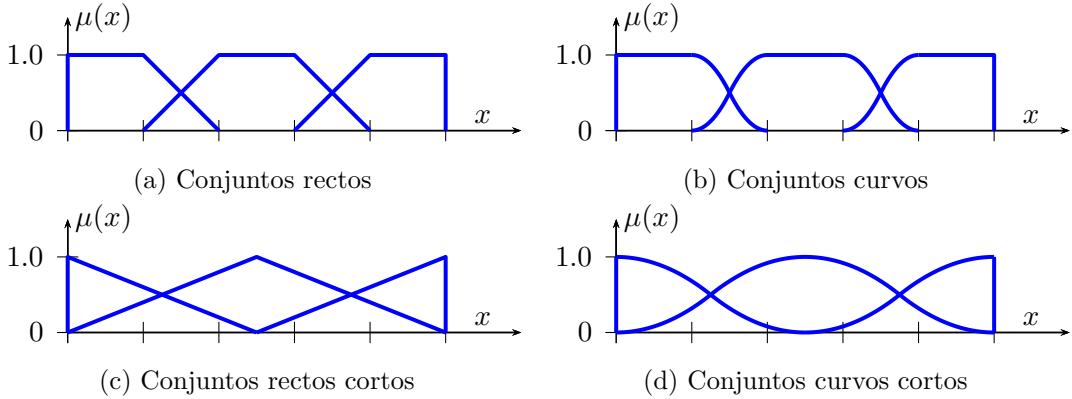


Figura 3.6: Cuatro formas usuales de variables lingüísticas.

Tabla 3.2: Vértices de conjuntos igualmente espaciados. δ se calcula según la ecuación 3.11

i	a	b	c	d
1	\mathcal{U}_L	\mathcal{U}_L	$\mathcal{U}_L + \delta$	$\mathcal{U}_L + 2\delta$
$2, \dots, n-1$	$\mathcal{U}_L + (2i-1)\delta$	$\mathcal{U}_L + (2i)\delta$	$\mathcal{U}_L + (2i+1)\delta$	$\mathcal{U}_L + (2i+2)\delta$
n	$\mathcal{U}_R - 2\delta$	$\mathcal{U}_R - \delta$	\mathcal{U}_R	\mathcal{U}_R

$$\delta = \frac{\mathcal{U}_R - \mathcal{U}_L}{n-1} \quad (3.12)$$

En general, el diseño de cada variable lingüística depende de la aplicación en donde se utilice. Si se requiere que la variable sea interpretable por personas, hay varias recomendaciones que suelen aplicarse:

1. La función semántica debe establecer una *partición* del universo de discurso \mathcal{U} . Un conjunto de conjuntos difusos con funciones de pertenencia $\mu_1(x), \mu_2(x), \dots, \mu_N(x)$

Tabla 3.3: Vértices de conjuntos igualmente espaciados. δ se calcula según la ecuación 3.12

i	a	b	c	d
1	\mathcal{U}_L	\mathcal{U}_L	\mathcal{U}_L	$\mathcal{U}_L + \delta$
$2, \dots, n-1$	$\mathcal{U}_L + (i-1)\delta$	$\mathcal{U}_L + (i)\delta$	$\mathcal{U}_L + (i)\delta$	$\mathcal{U}_L + (i+1)\delta$
n	$\mathcal{U}_R - \delta$	\mathcal{U}_R	\mathcal{U}_R	\mathcal{U}_R

establecen una partición del universo de discurso \mathcal{U} si cumplen dos condiciones:

$$\begin{cases} \sum_{i=1,\dots,N} \mu_i(x) = 1.0 & \forall x \in U \\ \max_{\mathcal{U}} (\mu_i(x)) > 0 & i = 1, 2, \dots, N \end{cases} \quad (3.13)$$

La primera condición en 3.13 establece una condición para todos los elementos del universo de discurso, que consiste en que al sumar su grado de pertenencia a todos los conjuntos el resultado debe ser 1. La segunda condición establece que no debe haber conjuntos vacíos.

2. Todos los conjuntos difusos deben ser *normales*. Un conjunto difuso es normal si al menos existe un elemento del universo de discurso en el que la función de pertenencia vale 1. Por ello, esta recomendación puede expresarse así:

$$\max_{\mathcal{U}} (\mu_i(x)) = 1 \quad i = 1, 2, \dots, N \quad (3.14)$$

Esta recomendación es más restrictiva que la segunda condición en 3.13

3. Los conjuntos difusos asociados a términos no adyacentes deben ser *disyuntos*. Dos o más conjuntos difusos son disyuntos si su intersección es vacía. Para poder establecer si dos términos son adyacentes es necesario establecer un orden que responda al significado del término¹.

La figura 3.7 muestra tres ejemplos de variables lingüísticas que no siguen las recomendaciones arriba listadas.

3.2. Lógica difusa

El término *lógica difusa* se utiliza para referirse al estudio de las proposiciones que contienen conceptos representados por conjuntos difusos. Por esa razón, en rigor debería denominarse *lógica proposicional difusa*, pero por antonomasia no es así. Llegaremos a ese concepto en la sección 3.2.3, pero para ello primero resumiremos los aspectos de la lógica proposicional concreta, que es binaria, en la sección 3.2.1 y luego presentaremos el argumento matemático que nos permite conectar la teoría de conjuntos difusos con la lógica difusa, en la sección 3.2.2.

3.2.1. Lógica proposicional concreta

El campo de la lógica matemática que estudia la forma en que expresamos enunciados y la forma en que creamos nuevos enunciados se denomina *Lógica proposicional* o *Cálculo*

¹Por ejemplo, los términos ‘Bajo’-‘Medio’-‘Alto’ están correctamente ordenados según su significado, mientras que los términos ‘Bajo’-‘Alto’-‘Medio’ no lo están.

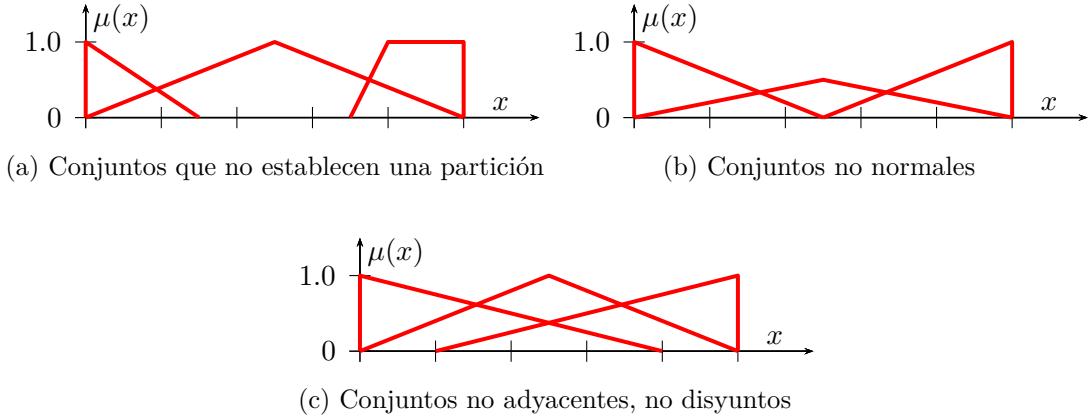


Figura 3.7: Ejemplos de variables lingüísticas que no son particiones

proposicional. En este campo, una *proposición* es un enunciado, es decir, una afirmación o una idea sobre la que se puede discutir si es *verdadera* o *falsa*. Una proposición simple tiene la forma

$$X \text{ es } P \quad (3.15)$$

La lógica proposicional estudia cómo determinar el grado de verdad de proposiciones compuestas que se construyen como combinaciones de proposiciones simples que tienen la forma de 3.15. Más exactamente, estudia cómo se propaga la verdad (y por tanto, la falsedad/falsedad) de una o más proposiciones cuando éstas se combinan para crear nuevas proposiciones.

La combinación de proposiciones se realiza a través de *conectores lógicos*, también denominados *conectivas*. La forma en que se propaga la verdad a través de cada tipo de conector se denomina *función de verdad* de ese conector.

Existen tres conectores básicos a partir de los cuales se pueden desarrollar otros más complejos: 1) la *disyunción* 2) la *conjunción* y 3) la *negación*. La figura 3.8a muestra las funciones de verdad de estos tres operadores. Cada una de esas funciones se ha descrito con una tabla, que se denomina la *tabla de verdad* del conector.

Para ilustrar el uso de las tablas de verdad, consideremos la proposición

$$p : \text{'Hoy es viernes'} \quad (3.16)$$

y una segunda proposición q que se obtiene a partir de la negación de p :

$$q = \neg p : \text{'Hoy no es viernes'} \quad (3.17)$$

Tabla 3.4: Isomorfismo entre tres campos de la matemática

Lógica Proposicional	Teoría de Conjuntos	Álgebra Booleana
Disyunción $p \vee q$	Unión $A \cup B$	OR $x \oplus y$
Conjunción $p \wedge q$	Intersección $A \cap B$	AND $x \otimes y$
Negación $\neg p$	Complemento A'	NOT $\ominus x$

La tabla de verdad de la negación (Figura 3.8a) establece que si la proposición p es verdadera, su negación es falsa. Y viceversa: si la proposición p es falsa, su negación es verdadera. Nótese que la tabla de verdad no nos permite saber, per se, si la afirmación ‘Hoy no es viernes’ es cierta o falsa; lo que nos permite es saber cómo se propaga la verdad de la afirmación ‘Hoy es viernes’ cuando se construye una nueva afirmación que la niega.

3.2.2. Un isomorfismo útil

La matemática demuestra que existe un *isomorfismo* entre tres campos del conocimiento: 1) la Teoría de conjuntos. 2) la Lógica proposicional y 3) el Álgebra booleana.

Eso significa que los tres campos del conocimiento tienen una estructura equivalente. Por tanto, un problema definido en uno de esos campos tiene un problema equivalente definido en los otros campos. Y si se encuentra la solución en uno de ellos, se puede hallar la solución equivalente en los otros dos.

Una forma de explorar el isomorfismo existente entre la teoría de conjuntos y la lógica proposicional es convertir enunciados en relaciones entre conjuntos. Por ejemplo, el enunciado ‘Buenos Aires es una ciudad’ puede interpretarse como ‘el elemento Buenos Aires pertenece al conjunto de las Ciudades’. Otro ejemplo conocido es el siguiente: el enunciado ‘Todos los animales son mortales’ puede interpretarse como ‘el conjunto de los animales está contenido dentro del conjunto de los mortales’.

De forma más específica, la Tabla 3.4 identifica las operaciones básicas de cada uno de los tres campos que son equivalentes. Por ejemplo, si tomamos la primera línea de esa tabla encontramos que la *unión de conjuntos* es equivalente a la *disyunción de proposiciones* y a la *operación OR de variables booleanas*.

Esta equivalencia se pone de manifiesto cuando se comparan las tablas de las tres operaciones. Estas se muestran en la figura 3.8. Nótese que los ‘0’ y ‘1’ en la tabla de

Disyunción <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>p</th> <th>q</th> <th>$p \vee q$</th> </tr> </thead> <tbody> <tr><td>F</td><td>F</td><td>F</td></tr> <tr><td>F</td><td>V</td><td>V</td></tr> <tr><td>V</td><td>F</td><td>V</td></tr> <tr><td>V</td><td>V</td><td>V</td></tr> </tbody> </table>	p	q	$p \vee q$	F	F	F	F	V	V	V	F	V	V	V	V	Conjunción <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>p</th> <th>q</th> <th>$p \wedge q$</th> </tr> </thead> <tbody> <tr><td>F</td><td>F</td><td>F</td></tr> <tr><td>F</td><td>V</td><td>F</td></tr> <tr><td>V</td><td>F</td><td>F</td></tr> <tr><td>V</td><td>V</td><td>V</td></tr> </tbody> </table>	p	q	$p \wedge q$	F	F	F	F	V	F	V	F	F	V	V	V	Negación <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>p</th> <th>$\neg p$</th> </tr> </thead> <tbody> <tr><td>F</td><td>V</td></tr> <tr><td>V</td><td>F</td></tr> </tbody> </table>	p	$\neg p$	F	V	V	F
p	q	$p \vee q$																																				
F	F	F																																				
F	V	V																																				
V	F	V																																				
V	V	V																																				
p	q	$p \wedge q$																																				
F	F	F																																				
F	V	F																																				
V	F	F																																				
V	V	V																																				
p	$\neg p$																																					
F	V																																					
V	F																																					
(a) Lógica proposicional																																						
Unión <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>μ_A</th> <th>μ_B</th> <th>$\mu_{A \cup B}$</th> </tr> </thead> <tbody> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </tbody> </table>	μ_A	μ_B	$\mu_{A \cup B}$	0	0	0	0	1	1	1	0	1	1	1	1	Intersección <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>μ_A</th> <th>μ_B</th> <th>$\mu_{A \cap B}$</th> </tr> </thead> <tbody> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </tbody> </table>	μ_A	μ_B	$\mu_{A \cap B}$	0	0	0	0	1	0	1	0	0	1	1	1	Complemento <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>μ_A</th> <th>$\mu_{A'}$</th> </tr> </thead> <tbody> <tr><td>0</td><td>1</td></tr> <tr><td>1</td><td>0</td></tr> </tbody> </table>	μ_A	$\mu_{A'}$	0	1	1	0
μ_A	μ_B	$\mu_{A \cup B}$																																				
0	0	0																																				
0	1	1																																				
1	0	1																																				
1	1	1																																				
μ_A	μ_B	$\mu_{A \cap B}$																																				
0	0	0																																				
0	1	0																																				
1	0	0																																				
1	1	1																																				
μ_A	$\mu_{A'}$																																					
0	1																																					
1	0																																					
(b) Teoría de conjuntos																																						
OR <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>x</th> <th>y</th> <th>$x \oplus y$</th> </tr> </thead> <tbody> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>1</td></tr> <tr><td>1</td><td>0</td><td>1</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </tbody> </table>	x	y	$x \oplus y$	0	0	0	0	1	1	1	0	1	1	1	1	AND <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>x</th> <th>y</th> <th>$x \otimes y$</th> </tr> </thead> <tbody> <tr><td>0</td><td>0</td><td>0</td></tr> <tr><td>0</td><td>1</td><td>0</td></tr> <tr><td>1</td><td>0</td><td>0</td></tr> <tr><td>1</td><td>1</td><td>1</td></tr> </tbody> </table>	x	y	$x \otimes y$	0	0	0	0	1	0	1	0	0	1	1	1	NOT <table border="1" style="border-collapse: collapse; width: 100%;"> <thead> <tr> <th>x</th> <th>$\ominus x$</th> </tr> </thead> <tbody> <tr><td>0</td><td>1</td></tr> <tr><td>1</td><td>0</td></tr> </tbody> </table>	x	$\ominus x$	0	1	1	0
x	y	$x \oplus y$																																				
0	0	0																																				
0	1	1																																				
1	0	1																																				
1	1	1																																				
x	y	$x \otimes y$																																				
0	0	0																																				
0	1	0																																				
1	0	0																																				
1	1	1																																				
x	$\ominus x$																																					
0	1																																					
1	0																																					
(c) Álgebra booleana																																						

Figura 3.8: Operaciones básicas equivalentes

la Unión ocupan las mismas posiciones que las ‘F’ y ‘V’ en la tabla de la disyunción y que los ‘0’ y ‘1’ en la tabla del operador OR. Exactamente lo mismo puede decirse de las equivalencias que se muestran en las demás tablas.

3.2.3. Lógica proposicional difusa

En la lógica proposicional concreta la verdad de una proposición puede tomar dos valores: Verdadero o Falso. Por esa razón es posible construir tablas de verdad como las que se muestran en la figura 3.8a.

Por su parte, en la *lógica proposicional difusa* la verdad de una proposición puede tomar infinitos valores. El *grado de verdad* de una proposición es un número real que está en el intervalo $[0, 1]$. El 0 representa la falsedad absoluta y el 1 la verdad absoluta. Un valor intermedio se interpreta como un verdad parcial. Por esta razón, se dice que la lógica proposicional difusa es un tipo de lógica *multivaluada*.

La lógica proposicional difusa es, sin duda, la rama de la lógica difusa de más amplia aplicación. Por esta razón, por antonomasia utilizaremos en este texto la expresión *lógica difusa* para referirnos a ella.

El camino más directo para presentar las nociones básicas de la lógica difusa hace uso de dos herramientas: 1) la definición de variable lingüística presentada en la sección 3.1 y 2) el isomorfismo presentado en la sección 3.2.2.

Por una parte, a través de la función semántica de la variable lingüística podremos asimilar el grado de verdad de una proposición al grado de pertenencia de un conjunto difuso. Y por otra, gracias al isomorfismo podremos construir nuevas proposiciones difusas empleando los operadores disponibles para operar sobre conjuntos difusos. Los ejemplos 3.2, 3.3 y 3.4 ilustran estos conceptos.

Ejemplo 3.2 (Grado de verdad difuso) Consideremos la proposición “*65 km/h es una velocidad alta*”. El grado de verdad de esa afirmación depende del contexto en el que se emplee. No es lo mismo referirse a la velocidad de un auto de competencia que a la de un ciclista, por ejemplo. Para poder establecer el significado de la expresión ‘velocidad alta’ empleamos una variable lingüística. En este ejemplo suponemos que el contexto en el que se emplea la proposición es tal que nos permite construir la variable lingüística ‘velocidad’ que se muestra en la figura 3.9.

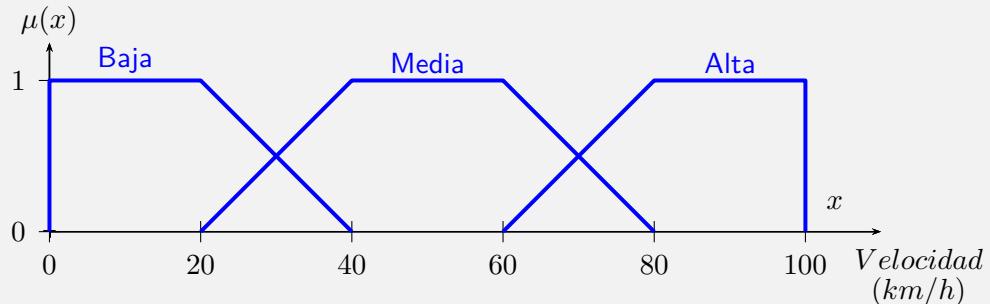


Figura 3.9: Función semántica en el ejemplo 3.2

La variable lingüística se ha construido con tres etiquetas: Baja, Media y Alta.

El grado de verdad de la proposición “*65 km/h es una velocidad alta*” es el grado de pertenencia de $x = 65 \text{ km/h}$ al conjunto difuso asociado a la etiqueta Alta. Este valor resulta ser 0.25, tal como se muestra en la figura 3.10.

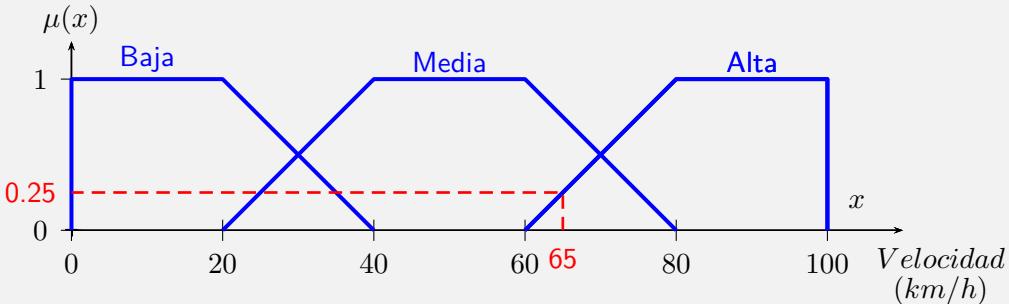


Figura 3.10: Grado de verdad difuso

Ejemplo 3.3 (Operación de disyunción en lógica difusa) Consideremos la proposición “65 km/h es una velocidad media o alta” en el mismo contexto del ejemplo 3.2.

Para asignar un grado de verdad a esa proposición debemos darle sentido a la expresión ‘velocidad media o alta’. El conector ‘o’ implica que esta proposición se ha construido como la disyunción de dos proposiciones:

“65 km/h es una velocidad media” **o** “65 km/h es una velocidad alta”

Haciendo uso del isomorfismo que se resume en la tabla 3.4 identificamos que el grado de verdad de esa disyunción se puede asimilar al grado de pertenencia a la unión de los conjuntos asociados a las etiquetas Media y Alta de la variable lingüística Velocidad (Figura 3.9).

En la figura 3.11 se ha trazado en color azul la función de pertenencia de esa unión. Para calcular la unión se ha empleado el operador máx. En esa misma figura se muestra en rojo que el grado de pertenencia de $x = 65 \text{ km/h}$ al conjunto resultante es 0.75. Por tanto, el grado de verdad asignado a la proposición “65 km/h es una velocidad media o alta” es 0.75.

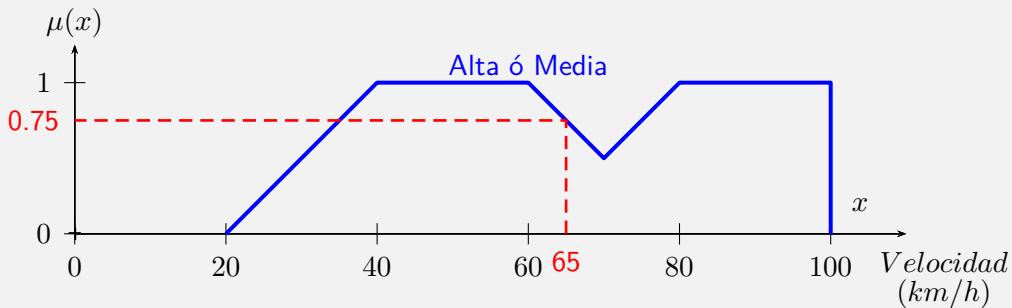


Figura 3.11: Operación de disyunción en lógica difusa

Nótese que la operación de unión podría haberse realizado con cualquier otra \mathcal{S} -Norma diferente al operador máx. En tal caso, el resultado numérico podría haber sido otro.

Ejemplo 3.4 (Operación de conjunción en lógica difusa) Consideremos la proposición “65 km/h es una velocidad alta y 30 km/h es una velocidad baja” en el mismo contexto del ejemplo 3.2.

El conector ‘y’ implica que esta proposición se ha construido como la conjunción de dos proposiciones:

“65 km/h es una velocidad alta” y “30 km/h es una velocidad baja”

El grado de verdad de cada una de esas dos expresiones se puede obtener por separado siguiendo el mismo procedimiento que se muestra en el ejemplo 3.2. La figura 3.12 muestra que esos valores son 0.5 y 0.25, respectivamente.

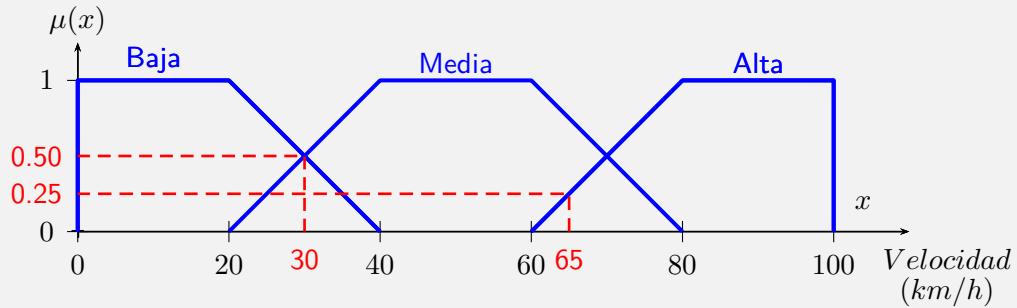


Figura 3.12: Operación de conjunción en lógica difusa

Haciendo uso del isomorfismo que se resume en la tabla 3.4 identificamos que el grado de verdad de la conjunción de las dos proposiciones se puede asimilar a la intersección de conjuntos. En consecuencia, podemos utilizar el operador \min para obtener el grado de verdad de la conjunción:

$$\min(0.25, 0.50) = 0.25$$

Nótese que la operación de intersección podría haberse realizado con cualquier otra \mathcal{T} -Norma diferente al operador \min . En tal caso, el resultado numérico podría haber sido otro.

Capítulo 4

Sistemas basados de reglas

Sin duda alguna, la aplicación más popular de los conjuntos difusos son los sistemas que utilizan la lógica difusa para efectuar cálculos matemáticos. Estos sistemas se suelen denominar *Sistemas de Lógica Difusa*, pero en este texto los llamamos *Sistemas Basados en Reglas* (SBR) para enfatizar que el conocimiento se plasma en reglas de comportamiento. Además, esta denominación nos permitirá diferenciarlos claramente de los *Sistemas Basados en Aritmética* que se presentan en el capítulo 12.

La estructura de este capítulo es la siguiente: en la sección 4.1 se muestra la estructura de un SBR, que consta de cuatro bloques. Esos bloques se explican en las secciones 4.2, 4.3, 4.4 y 4.5, para el caso de sistemas de un única entrada y un única salida. En la sección 4.6 se explica cómo operan los sistemas de múltiples entadas y salidas.

4.1. Estructura de los sistemas basados en reglas

En términos simples, un sistema de cómputo es un método para procesar información. Tal como se ilustra en la Figura 4.1a, el sistema se alimenta con cierta información (las Entradas) y genera nueva información (las Salidas).

Cuando las entradas y las salidas son números (Figura 4.1b), el papel del sistema de cómputo puede asimilarse al cálculo de una función matemática:

$$\mathbf{Y} = \mathbf{F}(\mathbf{X}) \tag{4.1}$$

en donde \mathbf{X} es un vector que contiene las variables de entrada y \mathbf{Y} es otro vector que contiene las variables de salida.

Dependiendo del número de variables de entrada y salida, los sistemas se clasifican en SISO (una entrada y una salida), MISO (varias entradas y una salida) ó MIMO (varias

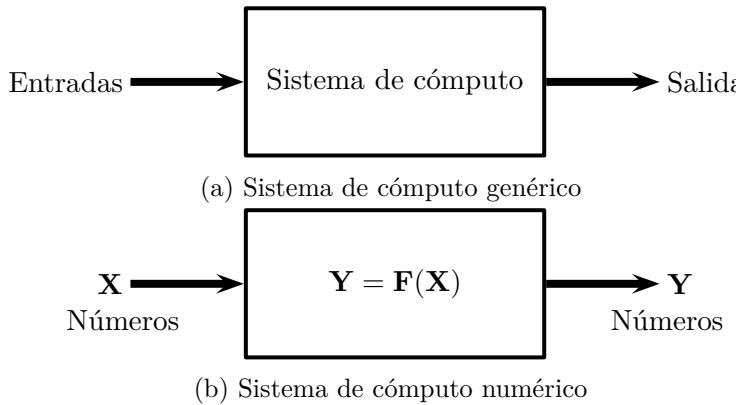


Figura 4.1: Sistemas de cómputo

entradas y varias salidas). La figura 4.2 ilustra esta clasificación y muestra el significado de cada una de esas tres siglas en inglés.

Los *sistemas basados en lógica difusa* que presentamos en este texto son sistemas de cómputo numérico. Específicamente presentaremos el modelo MIMO formulado por Mamdani en [40]. Estos sistemas también suelen denominarse *sistemas basados en reglas-SBR*, *sistemas tipo Mamdani* o *sistemas de lógica difusa*.

Aunque los SBR son, en general, sistemas MIMO, para facilitar la presentación se supondrá inicialmente un sistema tipo SISO. En la sección 4.6 se extiende esta explicación al caso MIMO.

La estructura interna de un SBR se muestra en la figura 4.3. Consta de cuatro bloques que se presentan en detalle en las secciones 4.2, 4.3, 4.4 y 4.5. El papel que cumplen cada uno de esos bloques se ilustra en la figura 4.4 y se explica a continuación:

Base de reglas : la base de reglas es una colección de reglas de la forma ‘*si la entrada es X entonces la salida es Y* ’. En su versión más usual, X y Y son etiquetas de variables lingüísticas asociadas a las variables de entrada y de salida, respectivamente.

Mediante estas reglas se describe el comportamiento que debe tener el sistema. Es una descripción lingüística, es decir, con palabras, de cómo deben obtenerse las salidas a partir de las entradas. Por esa razón, la base de reglas contiene el conocimiento sobre el comportamiento deseado del sistema.

Máquina de inferencia : la máquina de inferencia es el bloque encargado de aplicar el conocimiento contenido en la base de reglas a una situación en particular. La máquina de inferencia recibe la información en forma de conjuntos difusos y en-

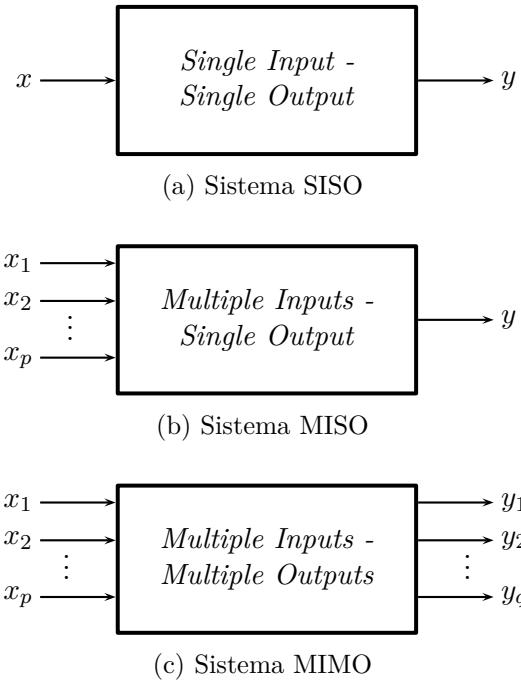


Figura 4.2: Tipos de sistemas numéricos según el número de entradas y salidas

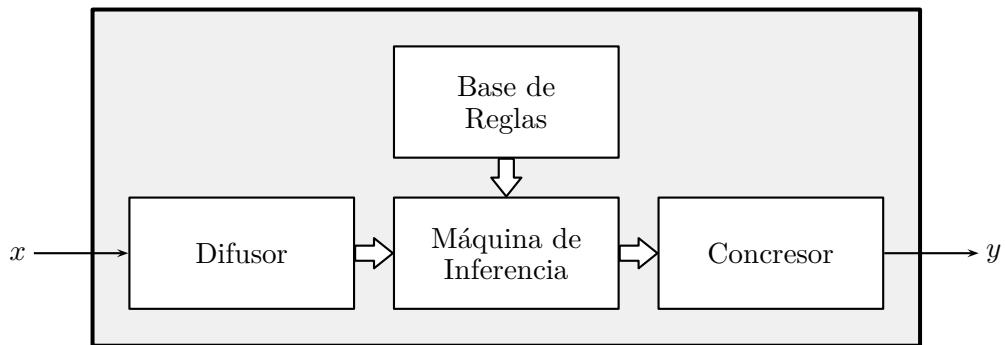


Figura 4.3: Estructura de un Sistema Basado en Reglas

trega su inferencia también en forma de conjuntos difusos. La máquina genera un conjunto difuso por cada una de las reglas almacenadas en la base de reglas.

Difusor : el difusor es el bloque encargado de recibir la entrada como un número y convertirla en un conjunto difuso. Este conjunto difuso es el que procesará el bloque de la máquina de inferencia. Este bloque también se conoce con los nombres *fuzzificador*, *emborronador* y *fuzzifier*.

Concresor : el concresor es el bloque encargado de calcular un valor numérico a partir de los conjuntos difusos que ha generado la máquina de inferencia. Este bloque también se conoce con los nombres *defuzzificador*, *desemborronador* y *defuzzyfier*.

De acuerdo a lo anterior, para construir un SBR es necesario realizar las siguientes tareas que se ilustran en el ejemplo 4.1:

1. Diseñar la Variable Lingüística para la entrada.
2. Diseñar la Variable Lingüística para la salida.
3. Construir la Base de Reglas.
4. Ajustar los operadores matemáticos.

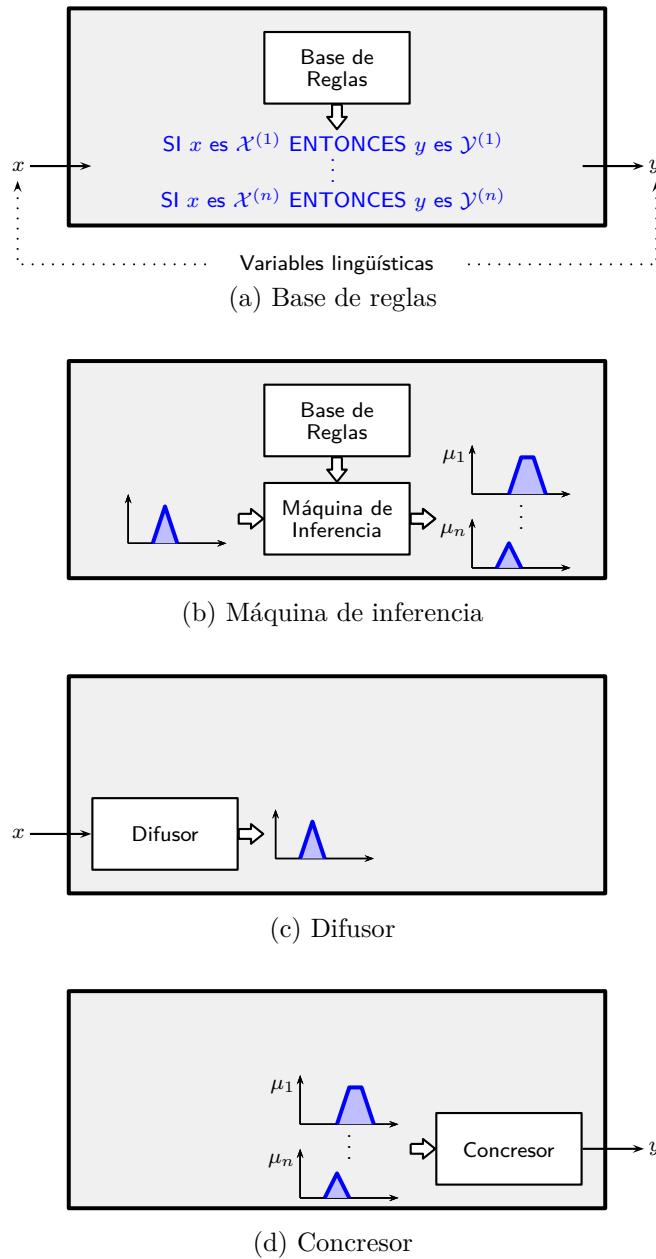


Figura 4.4: Bloques de un Sistema Basado en Reglas

Ejemplo 4.1 (Sistema de lógica difusa tipo SISO) En este ejemplo se diseña un SBR de una entrada y una salida. Iniciamos por la construcción de las variables lingüísticas. Se ha decidido que cada una de las dos variables tenga tres etiquetas: Baja, Media y Alta. También se ha decidido que la forma de los conjuntos difusos asociados sea la que se muestra en la figura 4.5

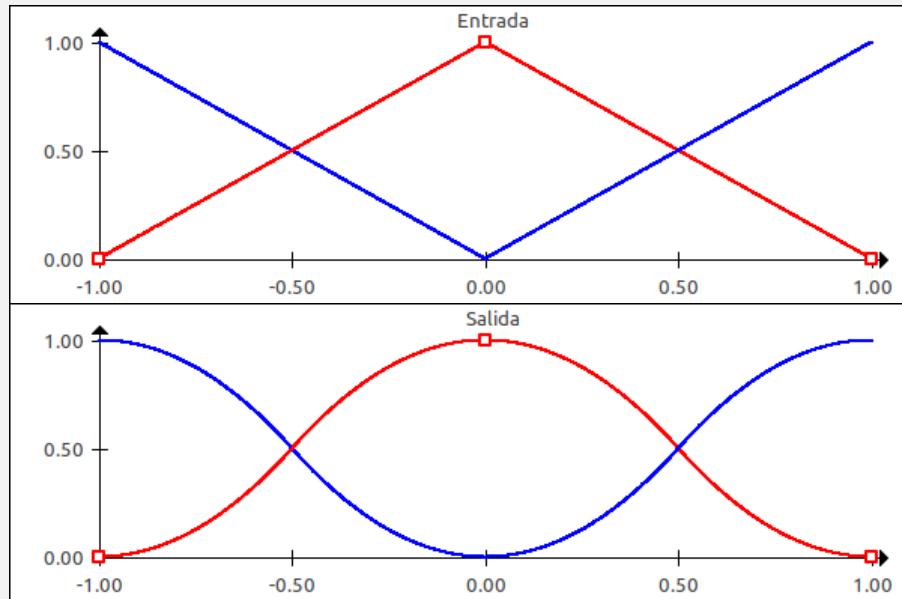


Figura 4.5: Variables lingüísticas del ejemplo 4.1

En este ejemplo se desea que conforme la entrada sea más grande, la salida sea más pequeña. Por esa razón se ha construido una base de reglas que tiene tres reglas de comportamiento:

- Regla 1: Si la entrada es Baja Entonces la salida es Alta
- Regla 2: Si la entrada es Media Entonces la salida es Media
- Regla 3: Si la entrada es Alta Entonces la salida es Baja

Haciendo uso de una herramienta de software adecuada (sección C.1) se puede visualizar ya la relación entre la entrada y la salida, tal como se muestra en la figura 4.6. Nótese cómo el sistema tiene el comportamiento deseado: conforme la entrada se hace más grande, la salida se hace más pequeña.

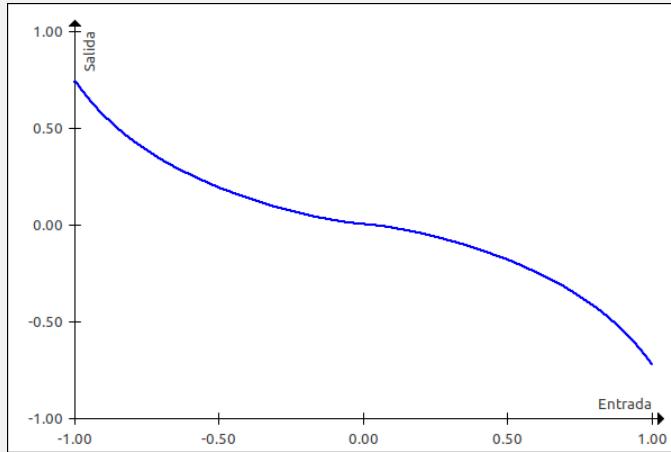


Figura 4.6: Función de entrada-salida para el ejemplo 4.1

Ahora bien, es posible realizar ajustes de varios operadores matemáticos que están dentro de cada uno de los cuatro bloques principales para tener comportamientos numéricamente distintos manteniendo ese comportamiento deseado. La figura 4.7 muestra cuatro funciones de entrada salida diferentes para el mismo sistema, que se han obtenido modificando algunos de esos parámetros. Nótese que en todos ellos la tendencia es decreciente, que corresponde al comportamiento especificado en la base de reglas.

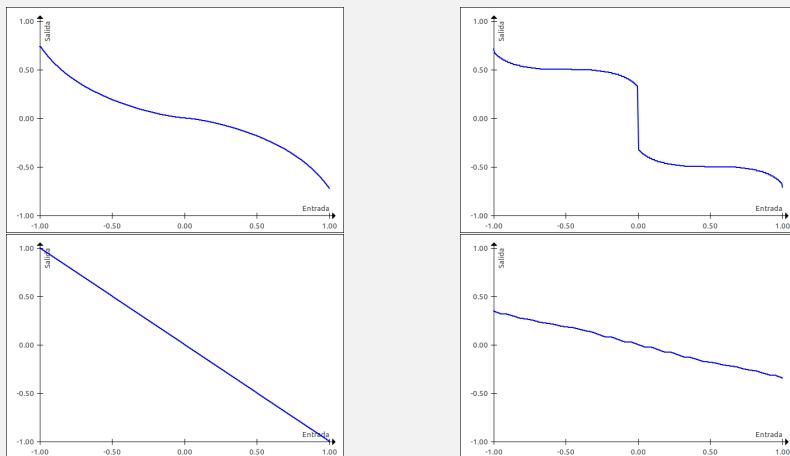


Figura 4.7: Varias funciones de entrada-salida para el ejemplo 4.1

4.2. Base de reglas

Supongamos un SBR de una entrada y una salida denotadas por x , y , respectivamente (Figura 4.3). La base de reglas es un conjunto de n reglas:

- Regla 1: **Si** x es $\mathcal{X}^{(1)}$ **entonces** y es $\mathcal{Y}^{(1)}$
- Regla 2: **Si** x es $\mathcal{X}^{(2)}$ **entonces** y es $\mathcal{Y}^{(2)}$
- ⋮
- Regla n : **Si** x es $\mathcal{X}^{(n)}$ **entonces** y es $\mathcal{Y}^{(n)}$

Podemos utilizar el operador \Rightarrow para definir de forma resumida la base de reglas \mathcal{BR} así:

$$\mathcal{BR} = \left\{ \mathcal{X}^{(k)} \Rightarrow \mathcal{Y}^{(k)} \quad , k = 1, 2, \dots, n \right\} \quad (4.2)$$

En donde $\mathcal{X}^{(k)}$ es alguna de las etiquetas lingüísticas de la variable x , mientras que $\mathcal{Y}^{(k)}$ es alguna de las etiquetas lingüísticas de la variable y . Los dos componentes de cada regla, $\mathcal{X}^{(k)}$, $\mathcal{Y}^{(k)}$ se denominan el *antecedente* y el *consecuente* de la regla, respectivamente.

En el contexto de los SBR, una regla del tipo ‘si antecedente entonces consecuente’ debe interpretarse como una relación de causalidad: ‘si a la entrada sucede el antecedente, entonces a la salida debe darse el consecuente’. Por ello, la base de reglas contiene la información sobre cómo debe ser el comportamiento del SBR. Es una información representada de forma lingüística y por ello se dice que un SBR es un *sistema de computación con palabras*.

Ejemplo 4.2 (Base de reglas) Supongamos que se desea diseñar un SBR que ayude a calcular el riesgo de daño al que está expuesto un cierto aparato debido a que está expuesto a una fuente de calor. El riesgo es mayor cuanto mayor sea el calor emanado por esa fuente.

Para ello debemos construir un SBR con una entrada y una salida. La variable de entrada es el **calor** emanado y la variable de salida es el **riesgo** de daño. Utilicemos las siguientes etiquetas:

- Para la variable **calor**: *poco*, *medio* y *mucho*.
- Para la variable **riesgo**: *aceptable*, *neutro* e *inaceptable*.

El conocimiento que poseemos (El riesgo es mayor cuanto mayor sea el calor emanado por esa fuente) puede expresarse mediante reglas que relacionan las etiquetas lingüísticas:

- Regla 1: Si el calor es poco entonces el riesgo es aceptable.
- Regla 2: Si el calor es medio entonces el riesgo es neutro.
- Regla 3: Si el calor es mucho entonces el riesgo es inaceptable.

La base de reglas será el conjunto de esas tres reglas. La base puede resumirse en la Tabla 4.1 que contiene los antecedentes y consecuentes de cada regla:

Tabla 4.1: Base de reglas en el ejemplo 4.2.

Calor	Riesgo
Poco	Aceptable
Medio	Neutro
Mucho	Inaceptable

Ejemplo 4.3 (Efecto de la base de reglas) En este ejemplo se ilustra el efecto de la base de reglas sobre el comportamiento de un SBR. Para ello, suponemos un sistema SISO, cuyas variables de entrada y salida se denominan, respectivamente *Entrada* y *Salida*. Ninguna sorpresa.

Además, suponemos que cada una de esas dos variables tiene tres etiquetas denominadas: *Bajo*, *Medio* y *Alto*. Las variables lingüísticas se han definido como conjuntos rectos cortos (Figura 3.6c).

En esas condiciones, se han construido cinco bases de reglas diferentes y se ha utilizado una herramienta de software adecuada (sección C.1) para visualizar la relación entrada-salida resultante en cada caso. Los resultados se muestran en la figura 4.8.

- En la base de reglas del primer caso (Figura 4.8a) al avanzar en las etiquetas del antecedente (Bajo, Medio, Alto) también se avanza en las etiquetas del consecuente (Bajo, Medio, Alto). El efecto final es que la relación entrada-salida es creciente, es decir, al incrementar el valor numérico de la entrada se incrementa el valor numérico de la salida.
- En la base de reglas del segundo caso (Figura 4.8b) al avanzar en las etiquetas del antecedente (Bajo, Medio, Alto) se retrocede en las etiquetas del consecuente (Alto, Medio, Bajo). El efecto final es que la relación entrada-salida es decreciente, es decir, al incrementar el valor numérico de la entrada se disminuye el valor numérico de la salida.
- En la base de reglas del tercer caso (Figura 4.8c) el consecuente siempre es igual

(Medio). El efecto final es que la relación entrada-salida es constante el valor numérico de la salida no cambia.

- En la base de reglas del cuarto caso (Figura 4.8d), al avanzar en las etiquetas del antecedente (Bajo, Medio, Alto) inicialmente se retrocede en las etiquetas del consecuente y luego se avanza (Alto, Medio, Alto). El efecto final es que la relación entrada-salida es concava hacia arriba, semejante a una U.
- En la base de reglas del quinto caso (Figura 4.8e), al avanzar en las etiquetas del antecedente (Bajo, Medio, Alto) inicialmente se avanza en las etiquetas del consecuente y luego se retrocede (Bajo, Medio, Bajo). El efecto final es que la relación entrada-salida es concava hacia abajo, semejante a una U invertida.'

4.2.1. Representación del conocimiento de una regla

Las curvas que se muestran en las figuras 4.6, 4.7 y 4.8 son ejemplos de relaciones entrada-salida. Estas relaciones están construidas sobre el espacio que se obtiene al efectuar el producto cartesiano de los universos de discurso de las variables de entrada y de salida. Para los SBR estos universos son intervalos en \mathcal{R} y por eso las curvas se dibujan en \mathcal{R}^2 .

Ahora bien, cada una de las reglas contiene una parte del conocimiento que se utiliza para construir esa relación. En otras palabras, cada una de las reglas aporta conocimiento para construir una parte de la relación de entrada-salida. ¿Cómo se visualiza ese conocimiento?

Una regla de la forma

$$\text{Si } x \text{ es } \mathcal{X}^{(k)} \text{ entonces } y \text{ es } \mathcal{Y}^{(k)} \quad (4.3)$$

establece una relación entre las etiquetas $\mathcal{X}^{(k)}, \mathcal{Y}^{(k)}$. Recordemos que cada etiqueta tiene un conjunto difuso asociado en la variable lingüística correspondiente, que podemos denominar $x^{(k)}, y^{(k)}$, respectivamente. La relación entre las etiquetas puede expresarse como una relación entre sus conjuntos difusos asociados.

Denotemos por $\mu_x^{(k)}(x)$ la función de pertenencia del conjunto $x^{(k)}$ y por $\mu_y^{(k)}(y)$ la del conjunto $y^{(k)}$. La relación R entre esos dos conjuntos puede analizarse con los principios que se presentan en la sección 2.3.3: la relación entre $x^{(k)}, y^{(k)}$ es el producto cartesiano de los dos conjuntos, cuya función de pertenencia se calcula aplicando una \mathcal{T} -norma:

$$\mu_R(x, y) = \mathcal{T}(\mu_A(x), \mu_B(y)) \quad (4.4)$$

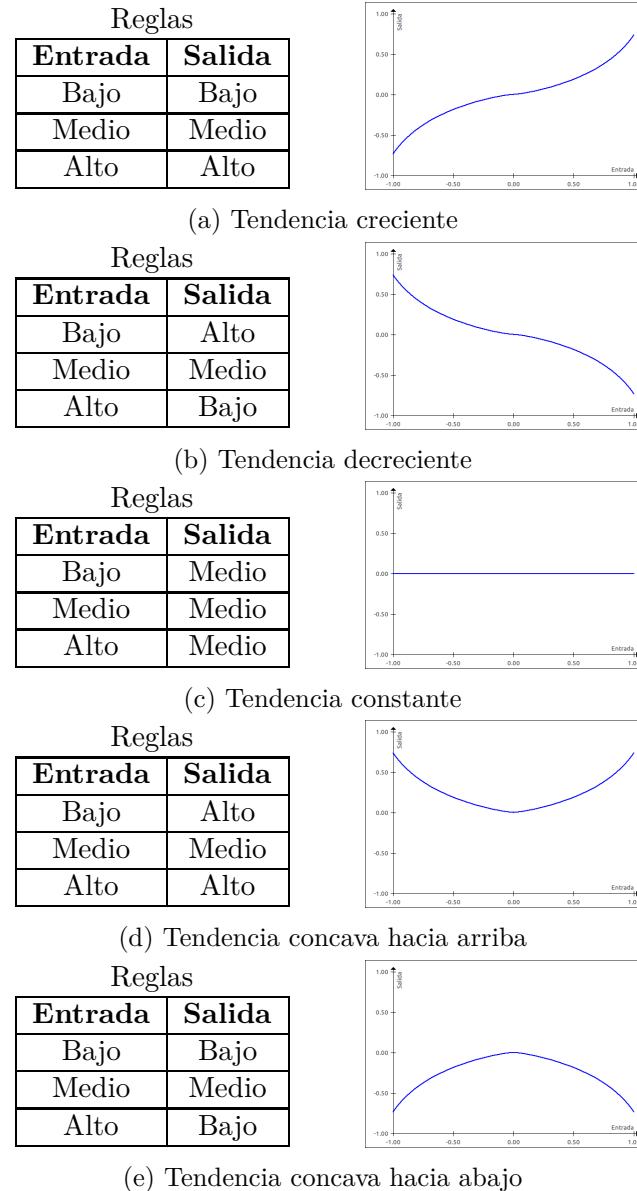


Figura 4.8: Efecto de las reglas sobre la relación entrada-salida en el ejemplo 4.3

Ejemplo 4.4 (El conocimiento en las reglas) Retomemos el ejemplo 4.2. En dicho ejemplo se formula un SBR de una entrada (calor) y una salida (riesgo); cada variable tiene tres etiquetas. La base de reglas se consigna en la Tabla 4.1 y se reproduce aquí:

Calor	Riesgo
Poco	Aceptable
Medio	Neutro
Mucho	Inaceptable

Suponiendo que las dos variables se han definido como conjuntos rectos cortos (Figura 4.4) ¿Cómo se puede visualizar el conocimiento de cada una de las reglas?

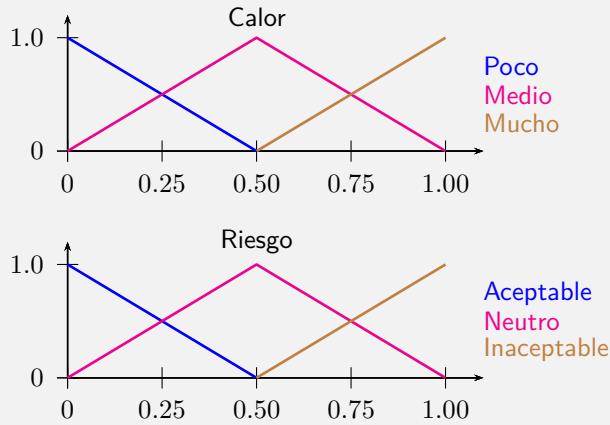
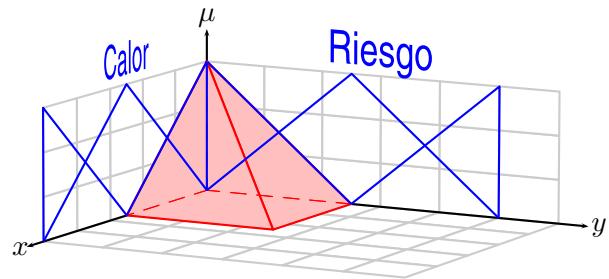


Figura 4.9: Variables lingüísticas del ejemplo 4.4

La figura 4.10a muestra el caso de la primera regla (si el calor es poco entonces el riesgo es aceptable). La gráfica es en 3 dimensiones porque se trata de un conjunto difuso sobre \mathcal{R}^2 : un eje corresponde a la variable x , otro a la variable y y el tercero (el vertical) al grado de pertenencia del conjunto.

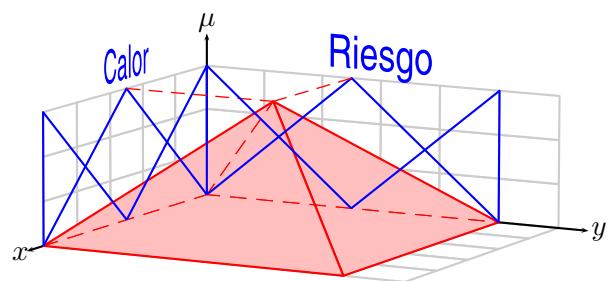
En esa figura se han dibujado las dos variables lingüísticas (las mismas que aparecen en la figura 4.9). Las proposiciones 'calor es poco' y 'riesgo es aceptable' corresponden a los dos conjuntos difusos triángulares más cercanos al origen. Para calcular la relación entre esos dos conjuntos se extienden sobre \mathcal{R}^2 y se calcula la intersección. EDn este ejemplo hemos usado el operador \min para ello. El resultado es la pirámide que se observa en color rojo.

El mismo procedimiento se ha empleado para las otras dos reglas. En cada caso la pareja de conjuntos que hay que relacionar es diferente y, por tanto, la pirámide que se obtiene es diferente. Ello se puede ver en las figuras 4.10b y 4.10c.



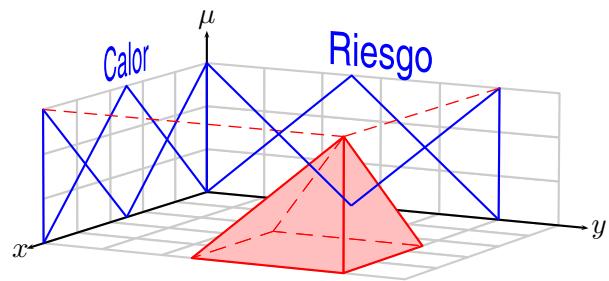
Calor es Poco \rightarrow Riesgo es Aceptable

(a) Regla 01



Calor es Medio \rightarrow Riesgo es Neutro

(b) Regla 02



Calor es Mucho \rightarrow Riesgo es Inaceptable

(c) Regla 03

Figura 4.10: Representación difusa del conocimiento de la base de reglas en el ejemplo 4.4

Ejemplo 4.5 (Cubrimiento de una regla) Hemos mencionado que cada regla aporta aporta conocimiento para construir una parte de la relación de entrada-salida. Continuando con el ejemplo 4.4, en este ejemplo se muestra en qué trozo de la relación entrada-salida incide cada una de las reglas.

Consideremos de nuevo la primera regla (si el calor es poco entonces el riesgo es aceptable), ahora visualizada en la Figura 4.11a. Allí se ha trazado la relación emtrada salida (es la curva azul); también se han dibujado las dos variables lingüísticas, una bajo el eje horizontal y la otra a la izquierda del eje vertical.

En esas dos variables se ha dibujado en rojo el conjunto difuso asociado al antecedente y al consecuente de la regla('calor es poco' y 'riesgo es aceptable'). También se ha dibujado un rectángulo que proviene de cruzar las proyecciones en \mathbb{R}^2 de esos conjuntos (más exactamente, de los soportes de esos conjuntos).

El trozo de la curva que está contenido en ese rectángulo es el trozo en donde la regla incide, de alguna manera ¿de qué manera? Eso se explica en las secciones 4.3 y 4.5.

Algo semejante se ha realizado con las reglas 2 y 3 (figuras 4.11b y 4.11c, respectivamente). Nótese que la regla 2 logra incidir en toda la curva, porque los soportes de los conjuntos 'calor es medio' y 'riesgo es neutro' abarcan los dos universos de discurso.

4.2.2. Propiedades de una base de reglas

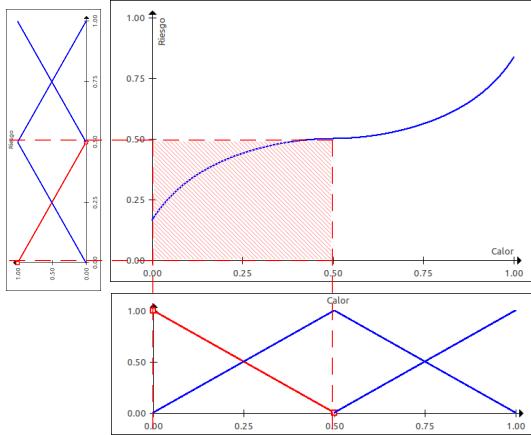
Una buena base de reglas debe satisfacer dos propiedades:

Completez : Una base de reglas es *completa* si existe al menos una regla para cualquier antecedente posible. Como los antecedentes dependen de las etiquetas de las variables lingüísticas de entrada, lo anterior es equivalente a decir que una base de reglas es completa si existe al menos una regla para cualquier combinación de etiquetas de entrada.

Supongamos que un SBR tiene dos variables de entrada; supongamos también que la primera variable tiene tres etiquetas (Bajo, Medio, Alto) y la segunda tiene cinco (Helado, Frio, Tibio, Caliente, Muy caliente). De esta forma, un posible antecedente para este sistema sería 'si la entrada 1 es Alto y la entrada 2 es Frio'. En tal sistema, hay $3 \times 5 = 15$ posibles combinaciones de etiquetas. Por tanto, una base de reglas debería tener al menos 15 reglas para considerarse completa.

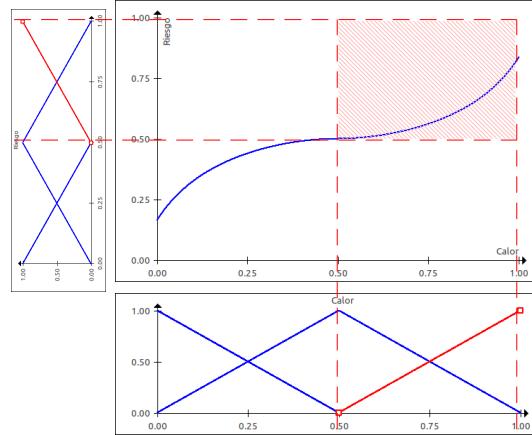
Generalizando, para un sistema de m entradas cuyas variables tienen e_1, e_2, \dots, e_m etiquetas, el número mínimo de reglas que debe tener la base de reglas para ser

Si el calor es poco entonces
el riesgo es aceptable



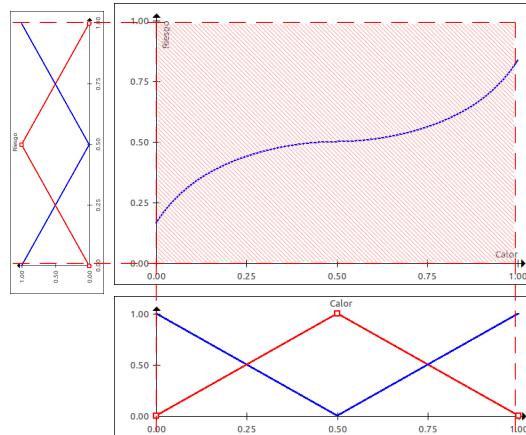
(a) Regla 01.

Si el calor es mucho entonces
el riesgo es inaceptable



(b) Regla 03.

Si el calor es medio entonces
el riesgo es neutro



(c) Regla 02.

Figura 4.11: Cubrimiento de cada una de las reglas

completo está dado por:

$$n_{min} = \prod_{k=1}^m e_k \quad (4.5)$$

Coherencia : Una base de reglas es *coherente* si no existen reglas que tengan el mismo antecedente y diferente consecuente.

Ejemplo 4.6 (Completez y coherencia de una base de reglas) Consideremos un sistema SISO, con cinco etiquetas a la entrada y cinco a la salida (Muy bajo, Bajo, Medio, Alto y Muy alto). Se ha diseñado de forma tal que su relación entrada salida sea una línea recta. En la figura 4.12a se muestra la base de reglas y la relación entrada salida; también se destaca la región que cubre la regla:

Si la entrada es Alto es poco entonces la salida es Alto.

La figura 4.12b muestra el efecto de eliminar esa regla. Al hacerlo, la región que anteriormente era cubierta por la regla queda mal definida. Es lógico: la máquina de inferencia no encontrará el conocimiento necesario para obtener una conclusión adecuada.

Por otra parte, en la figura 4.12c se muestra el efecto de incorporar dos reglas que tengan el mismo antecedente (Alto) y consecuentes distintos (Alto y Muy bajo). La máquina de inferencia obtendrá dos conclusiones contradictorias y por tanto la relación entrada-salida se verá afectada.

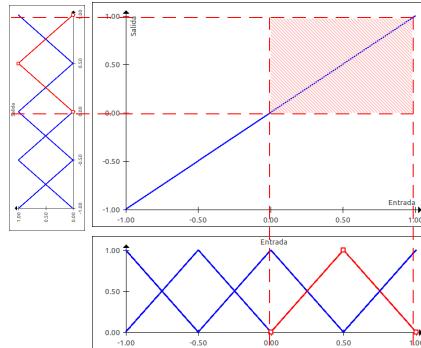
4.3. Máquina de inferencia

Para comprender en detalle el funcionamiento de un SBR conviene comenzar por entender cómo opera la máquina de inferencia, que es su bloque más importante. Ese es el propósito de esta sección.

Tal como se ilustra en la Figura 4.4b, la máquina de inferencia debe aprovechar el conocimiento almacenado en la base de reglas para inferir cuál debe ser la salida del sistema ante una cierta entrada. La máquina de inferencia realiza la misma operación con cada una de las reglas presentes en la base. Por ello, basta con comprender cómo se realiza la inferencia con una sola regla.

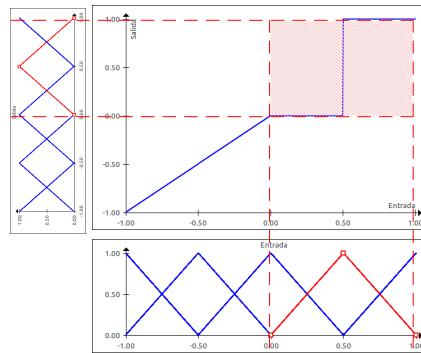
Por simplicidad, en esta sección se supone que el SBR es del tipo SISO, es decir, que tiene una sola entrada y una sola salida. La explicación está organizada de la siguiente forma: en la sección 4.3.1 se precisa en qué consiste la inferencia; posteriormente, en la sección 4.3.2 se ilustra cómo solemos hacer inferencia con números, aunque no sea usual

Reglas	
Entrada	Salida
Muy bajo	Muy bajo
Bajo	Bajo
Medio	Medio
Alto	Alto
Muy alto	Muy alto



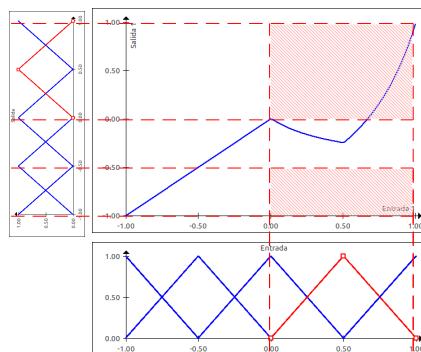
(a) Conjunto completo de reglas

Reglas	
Entrada	Salida
Muy bajo	Muy bajo
Bajo	Bajo
Medio	Medio
Muy alto	Muy alto



(b) Conjunto incompleto de reglas

Reglas	
Entrada	Salida
Muy bajo	Muy bajo
Bajo	Bajo
Medio	Medio
Alto	Alto
Alto	Muy bajo
Muy alto	Muy alto



(c) Conjunto incoherente de reglas

Figura 4.12: Completez y coherencia de la base de reglas para el ejemplo 4.6

que llamemos a ese proceso ‘inferencia’; en la sección 4.3.3 se extiende el concepto de la inferencia numérica al caso de los conjuntos difusos, lo que permite efectuar inferencias empleando las reglas del SBR.

4.3.1. Inferencia

En lógica, la inferencia es la obtención de nuevas proposiciones (conclusiones) a partir de unas proposiciones iniciales (premisas). Específicamente, la lógica se preocupa por determinar cómo se puede asegurar que las conclusiones preservan la validez de las premisas, es decir, cómo asegurar que si las premisas son verdaderas la conclusión también lo sea.

Ello se logra a través del uso de unas *reglas de inferencia*. Una regla de inferencia toma unas premisas, analiza su sintaxis, y devuelve una conclusión o conclusiones. El uso adecuado de esas reglas asegura la preservación de la validez de las premisas.

Los SBR utilizan una regla de inferencia específica denominada *modus ponens*¹ y que puede expresarse de la siguiente forma:

$$\text{si } A \text{ implica } B; \text{ y si } A \text{ es verdad; entonces } B \text{ también es verdad} \quad (4.6)$$

Para comprender el significado de esa regla, conviene desglosarla así:

1. *A implica B* es un conocimiento que se tiene y que relaciona a *A* y a *B* a través de una implicación.
2. *A es verdad* es un hecho.
3. *A también es verdad* es la conclusión que se obtiene.

Lo anterior nos permite escribir la regla del *modus ponens* así:

$$\frac{\begin{array}{c} A \Rightarrow B \\ A \end{array}}{B} \quad \begin{array}{l} \text{Conocimiento} \\ \text{Hecho} \\ \hline \text{Conclusión} \end{array} \quad (4.7)$$

Ejemplo 4.7 (Modus Ponens) Este ejemplo ilustra de forma simple la naturaleza de la regla de inferencia del *modus ponens*. Sabemos que el color de los tomates está relacionado con su grado de madurez. Tenemos ese conocimiento. En particular, sabemos que los tomates rojos están maduros.

¹También conocida como *modus ponendo ponens* que es una expresión en latín que puede traducirse como: ‘el modo que, al afirmar, afirma’.

Supongamos ahora que tenemos en nuestras manos un cierto tomate cuyo color es rojo. ¿Qué podemos inferir de su grado de madurez? obviamente, que está maduro.

Para formalizar lo anterior, construimos dos premisas:

$$\begin{aligned} A &: \text{el tomate está rojo} \\ B &: \text{el tomate está maduro} \end{aligned}$$

y con ellas formulamos la siguiente regla que contiene nuestro conocimiento:

$$A \Rightarrow B : \text{Si el tomate está rojo Entonces el tomate está maduro}$$

La aplicación de la fórmula del *modus ponens* (ecuación 4.7) resulta ser:

$$\frac{\begin{array}{c} \text{El tomate está rojo} \\ \text{El tomate está rojo} \end{array}}{\text{El tomate está maduro}} \quad \Rightarrow \quad \text{el tomate está maduro} \quad (4.8)$$

La razón de ser de la regla del *modus ponens* se revela al emplear las tablas de verdad para estudiar cómo se propaga la validez de las premisas. Para ello, primero reescribimos la ecuación 4.7 como

$$((A \Rightarrow B) \wedge A) \Rightarrow B \quad (4.9)$$

El paréntesis $(A \Rightarrow B)$ contiene el conocimiento. Al conectar ese conocimiento con el hecho A a través del conector \wedge , es decir, a través de la conjunción, estamos afirmando que tanto el conocimiento como el hecho suceden. La implicación $\Rightarrow B$ representa la inferencia, es decir, si se dan el conocimiento y el hecho, entonces se obtiene la conclusión, que es B .

La tabla 4.2 analiza la expresión 4.9. En la última columna se puede observar que todas las posibles combinaciones de verdad/falsedad para A y para B resultan verdaderas, es decir, que 4.9 es una *tautología*. En otras palabras, que la regla del *modus ponens* es correcta.

4.3.2. Inferencia numérica

En esta sección empleamos la expresión *inferencia numérica* para explorar cómo se realiza la inferencia cuando el conocimiento, los hechos y las conclusiones son numéricos. No

Tabla 4.2: Tautología del *modus ponens*

A	B	$A \Rightarrow B$	$((A \Rightarrow B) \wedge A)$	$((A \Rightarrow B) \wedge A) \Rightarrow B$
F	F	V	F	V
F	V	V	F	V
V	F	F	F	V
V	V	V	V	V

es una expresión de uso generalizado, pero se emplea aquí para facilitar la presentación, en la sección 4.3.3, del concepto de *inferencia difusa*.

Supongamos que existen dos variables numéricas x, y que pueden tomar valores reales. Si existe una relación entre esas dos variables, puede representarse en el plano cartesiano. En la figura 4.13a se ha supuesto que esa relación corresponde a la función $y = f(x)$ y se ha dibujado en color azul. Si, ahora, sabemos que el valor de la variable x es x_o , podemos conocer el valor de y aplicando $y = f(x_o)$, que es sencillamente *leer el resultado en el plano cartesiano* tal como se muestra en la misma figura.

Nótese que El conocimiento que tenemos sobre la relación que existe entre las dos variables reside en el plano \mathcal{R}^2 y está representado por el conjunto de puntos que forman la curva $y = f(x)$. Por su parte, cada una de las dos variables reside en \mathcal{R} , pero cada una de ellas tiene su propio universo, representado por su propio eje real.

¿Pero qué es *leer el resultado en el plano cartesiano*? El proceso puede desglosarse así:

1. Se ubica el hecho $x = x_o$ en su propio eje real.
2. Se traza una línea perpendicular al eje de la variable x sobre el plano cartesiano.
3. Se identifica el cruce de esa línea con la curva de la función $y = f(x)$.
4. Se traza una línea perpendicular al eje de la variable y que pasa por el cruce identificado en el paso anterior.
5. Se lee en el eje de la variable y el valor que cruza la línea trazada en el punto anterior.

Este procedimiento trivial puede resumirse de la siguiente forma empleando los conceptos de extensión cilíndrica y proyección cilíndrica (secciones 1.3.5 y 1.3.6):

1. Efectuar la extensión cilíndrica de x_o sobre \mathcal{R}^2
2. Obtener la intersección del conjunto obtenido en el paso 1 con el conjunto formado por todos los puntos de $y = f(x)$.
3. Obtener la proyección cilíndrica del conjunto obtenido en el paso 2 sobre el eje y

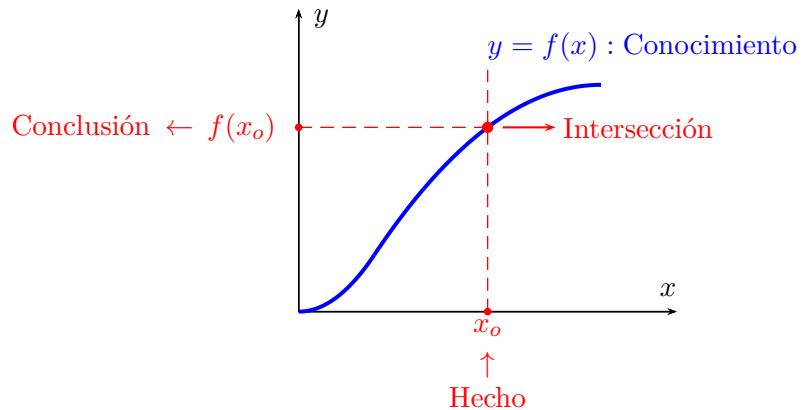
La figura 4.13b muestra cómo se aplica ese procedimiento cuando el hecho tiene incertidumbre. Se ha supuesto ahora que el hecho consiste en que la variable x corresponde al intervalo A . La extensión cilíndrica de este nuevo hecho sobre \mathcal{R}^s es ahora una franja y por tanto la intersección con la curva deja de ser un punto para convertirse ahora en un trozo de la curva. Al proyectar este trozo de recta sobre el eje y se obtiene un intervalo $f(A)$.

En ocasiones el conocimiento mismo puede contener incertidumbre. La figura 4.14 ilustra esa situación. Allí, los puntos que relacionan las variables x y y no forman una curva sino una franja en \mathcal{R}^2 , que se representa como una región de color azul tenue.

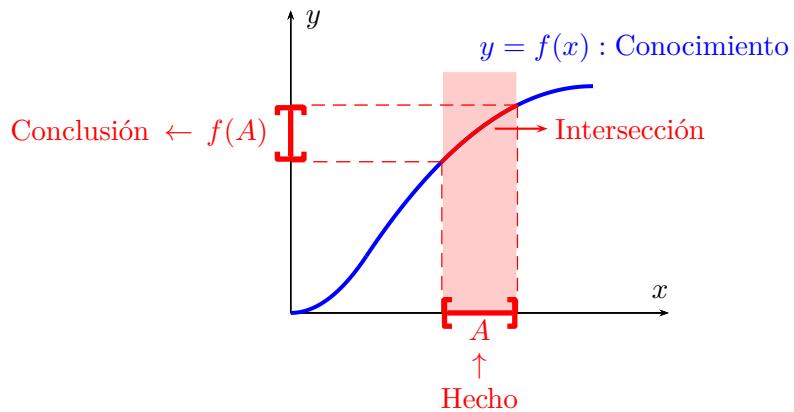
Si el hecho no contiene incertidumbre (figura 4.14a) al efectuar el procedimiento Extensión-Intersección-Proyección se obtiene un intervalo, debido a que la intersección entre la franja del conocimiento y la línea vertical de la extensión no es un único punto sino todo un intervalo.

Si además de la incertidumbre del conocimiento también hay incertidumbre en el hecho (figura 4.14b) la intersección entre la franja del conocimiento y la franja de la extensión resulta ser toda una región (marcada en rojo) y su proyección, un intervalo.

En resumen, cuando tanto el conocimiento como el hecho se expresan de forma numérica es posible efectuar una inferencia a partir del procedimiento de Extensión-Intersección-Proyección. La conclusión que así se obtiene puede ser un único valor o un conjunto de valores, dependiendo de la presencia o no de incertidumbre en el conocimiento y en el hecho.

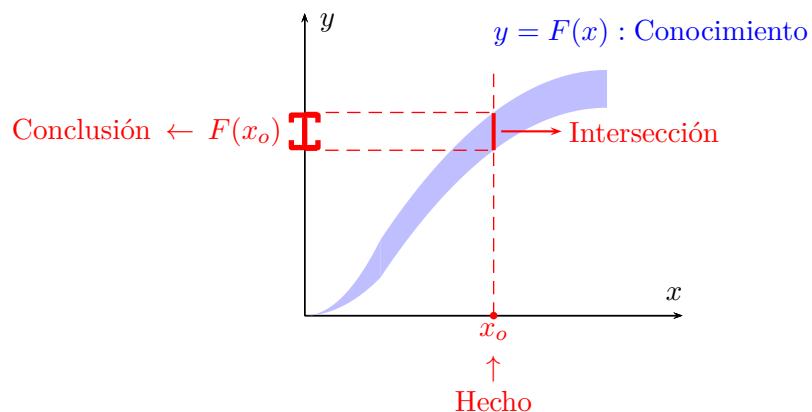


(a) Inferencia con un hecho sin incertidumbre

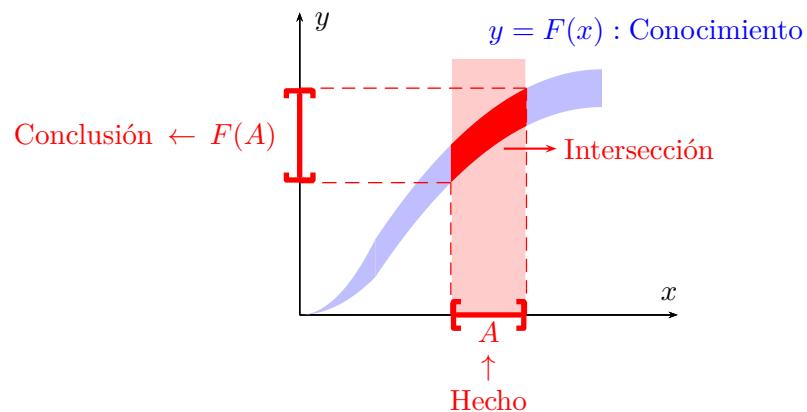


(b) Inferencia con un hecho con incertidumbre

Figura 4.13: Inferencia numérica con conocimiento sin incertidumbre



(a) Inferencia con un hecho sin incertidumbre



(b) Inferencia con un hecho con incertidumbre

Figura 4.14: Inferencia numérica con conocimiento sin incertidumbre

4.3.3. Inferencia difusa

La inferencia difusa lleva un paso más allá la inferencia numérica presentada en la sección 4.3.2. Tal como se ilustra en la figura 4.4, la inferencia difusa utiliza un conocimiento expresado como una regla de la forma:

$$\text{Si } x \text{ es } \mathcal{X}^{(k)} \text{ entonces } y \text{ es } \mathcal{Y}^{(k)} \quad (4.10)$$

En la sección 4.2 se explica que, $\mathcal{X}^{(k)}$ y $\mathcal{Y}^{(k)}$ son etiquetas de las variables lingüísticas asociadas a x y a y , respectivamente. Para facilitar la presentación del proceso de inferencia difusa vamos a suponer que una de las etiquetas de x es A y que una de las etiquetas de y es B y que existe una regla que es:

$$\text{Si } x \text{ es } A \text{ entonces } y \text{ es } B \quad (4.11)$$

El propósito es poder aplicar esa regla, es decir, poder usar ese conocimiento, aún cuando el hecho no sea *exactamente A* sino *aproximadamente A*. En otras palabras, lo que se busca es extender la regla convencional del *modus ponens* (ecuación 4.7) de la siguiente forma:

$$\frac{\begin{array}{c} A \Rightarrow B \\ \approx A \end{array}}{\approx B} \begin{array}{l} \text{Conocimiento} \\ \text{Hecho} \\ \hline \text{Conclusión} \end{array} \quad (4.12)$$

En donde $\approx A$ y $\approx B$ pueden interpretarse como *aproximadamente A* y *aproximadamente B*, respectivamente.

Ejemplo 4.8 (Modus Ponens generalizado) Para ilustrar el propósito de la inferencia difusa, retomemos el ejemplo 4.7. En ese ejemplo el conocimiento está expresado en la regla:

$$A \Rightarrow B : \text{Si el tomate está rojo Entonces el tomate está maduro}$$

Supongamos ahora que tenemos en nuestras manos un cierto tomate cuyo color no es completamente rojo, es *casi* rojo. ¿Podemos inferir algo sobre su grado de madurez?. Si usáramos el *modus ponens* convencional no podríamos inferir nada, porque para usar esa regla el hecho debe ser exactamente igual al antecedente de la regla y en nuestro caso el hecho (el tomate está *casi* rojo) no es igual al antecedente (el tomate está rojo).

El propósito de la inferencia difusa consiste en construir un tipo de *modus ponens* difuso que nos permita obtener una conclusión como la siguiente:

$$\begin{array}{l} \text{Si el tomate está rojo} \\ \text{El tomate está casi rojo} \\ \hline \text{El tomate está casi maduro} \end{array} \Rightarrow \text{está maduro} \quad (4.13)$$

4.3.3.1. Inferencia difusa con un hecho sin incertidumbre

La figura 4.15 ayuda a ilustrar el proceso de inferencia difusa cuando se tiene la regla ‘Si x es A entonces y es B ’ y además se sabe que $x = x_o$

El proceso puede desglosarse así:

- Como A y B son etiquetas de las variables lingüísticas asociadas a x y y , respectivamente, cada una de ellas tiene un conjunto difuso asociado. Estos conjuntos están definidos sobre su propio eje real, tal como se muestran en la figura 4.15a.
- La regla ‘Si x es A entonces y es B ’ establece una relación entre los conjuntos A y B . Como se explica en la sección 2.3.3, la relación entre conjuntos difusos es un nuevo conjunto difuso definido sobre el producto cartesiano de los dos universos de discurso.

En esa misma sección se explica que el grado de pertenencia de la relación se puede obtener haciendo uso de cualquier \mathcal{T} -norma (ecuación 2.16). Si denotamos por R a esa relación obtenemos el conjunto difuso cuya función de pertenencia está dada por

$$\mu_{\tilde{R}}(x, y) = \mathcal{T}_C(\mu_A(x), \mu_B(y)) \quad (4.14)$$

Si decidimos que la \mathcal{T} -norma sea el operador \min obtenemos el conjunto difuso que se muestra en la figura 4.15b, cuya función de pertenencia está dada por

$$\mu_{\tilde{R}}(x, y) = \min(\mu_A(x), \mu_B(y)) \quad (4.15)$$

El conjunto difuso que hemos obtenido tiene forma de pirámide. Este conjunto es la versión numérica del conocimiento expresado en la regla ‘Si x es A entonces y es B ’

- El hecho que queremos usar para obtener una conclusión es $x = x_o$. Para representarlo construimos un conjunto difuso tipo singleton centrado en $x = x_o$, tal como se muestra en la figura 4.15c.

- Para realizar la inferencia emplearemos el procedimiento explicado en la sección [4.3.2](#), que consta de tres pasos: 1) Extensión cilíndrica del hecho 2) Intersección de la extensión con el conocimiento y 3) Proyección cilíndrica de la intersección.
- La figura [4.15d](#) muestra la extensión cilíndrica del hecho sobre \mathcal{R}^2 y cómo se intersecta con el conocimiento. Como el hecho es un singleton, la extensión resulta ser un rectángulo que corta la pirámide. Ese corte es justamente la intersección, y tiene forma de trapecio.
- Las figuras [4.15e](#) muestra el resultado de la intersección y cuál es el resultado de efectuar la proyección cilíndrica sobre el universo de discurso de y .
- La conclusión que se ha obtenido resulta ser el conjunto difuso que se muestra en la figura [4.15f](#).

4.3.3.2. Inferencia difusa con un hecho con incertidumbre

Consideremos ahora el caso en que el hecho tiene incertidumbre y por tanto no podemos representarlo por un conjunto tipo singleton. Para ilustrarlo, vamos a suponer que el hecho y su incertidumbre su representan por un conjunto difuso de forma triangular. El proceso de inferencia se ha graficado en la figura [4.16](#).

El hecho se muestra en la figura [4.16c](#). Nótese como la extensión cilíndrica de ese hecho ahora es una pirámide que al cortar la pirámide de la regla genera una tercera pirámide (figuras [4.16d](#) y [4.16e](#)). Al realizar la proyección cilíndrica se obtiene un trapecio definido sobre el universo de discurso de la variable y . Este trapecio, al igual que el que se obtuvo en la figura [4.15f](#), resulta ser un recorte del conjunto B . No es de extrañar que ese tipo de resultado se suela denominar el *recortado* de la inferencia.

4.3.3.3. Ecuación de la inferencia difusa

El proceso de inferencia consiste, como se ha explicado ya, en obtener una conclusión a partir del uso de un conocimiento previo y un hecho que sucede. Para efectuar la inferencia difusa ha sido necesario construir un conjunto difuso que representa el conocimiento. Para ello hemos empleado la ecuación [2.16](#). Denotemos ese conjunto por C ; su función de pertenencia será entonces:

$$\begin{aligned}\mu_C(x, y) &= \mathcal{T}_C(\mu_A(x), \mu_B(y)) \\ \mu_C(x, y) &= \min(\mu_A(x), \mu_B(y))\end{aligned}\tag{4.16}$$

También hemos utilizado un conjunto difuso para representar el hecho. Su extensión cilíndrica puede obtenerse empleando la ecuación [1.7](#). Si denotamos por H el conjunto del hecho y por H_{ext} el conjunto extendido, tenemos:

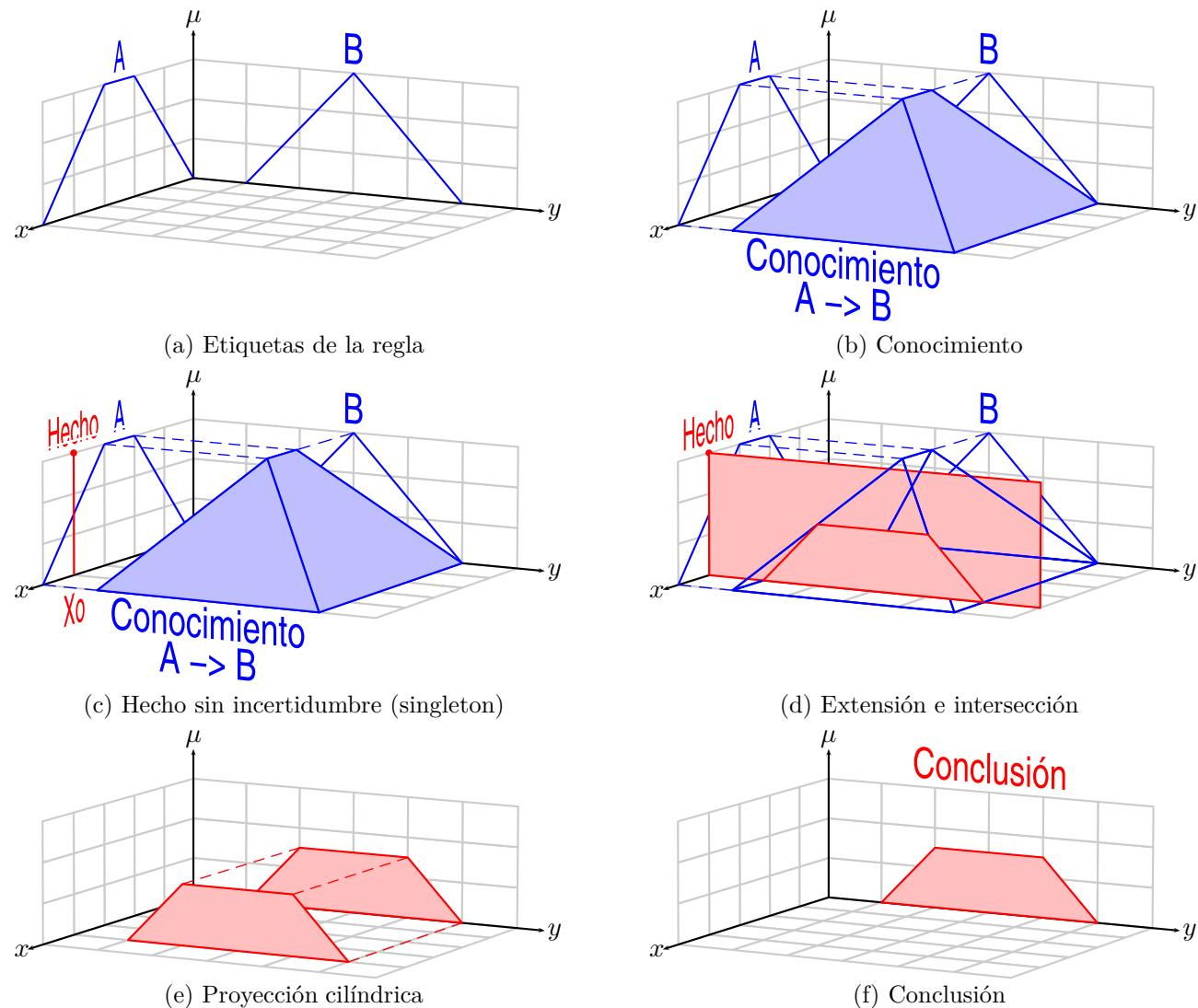


Figura 4.15: Inferencia difusa con un hecho sin incertidumbre

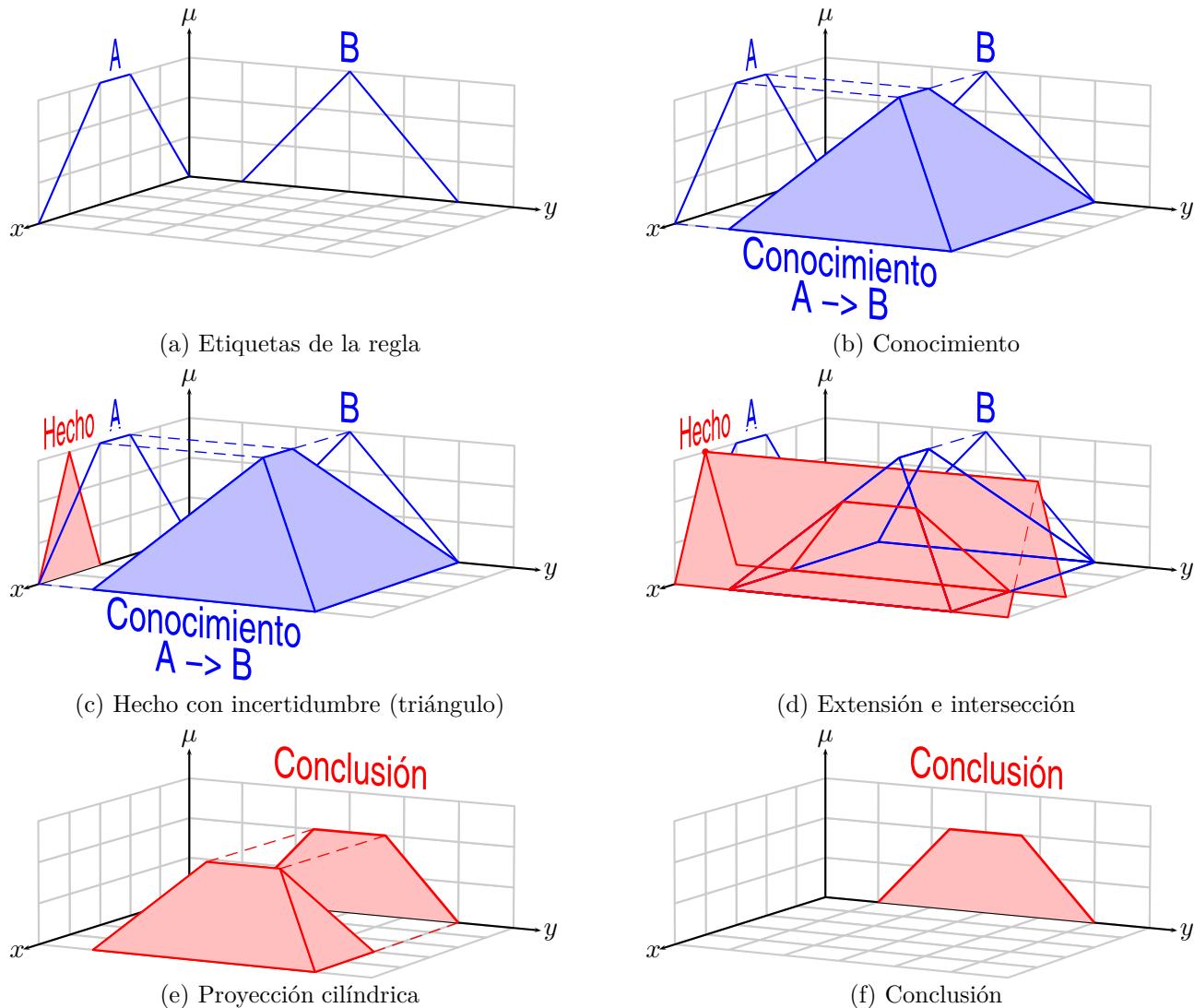


Figura 4.16: Inferencia difusa con un hecho con incertidumbre

$$\mu_{H_{ext}}(x, y) = \mu_H(x) \quad (4.17)$$

La intersección se ha realizado con el operador mín. Si denominamos a ese conjunto \cap , su función de pertenencia se calcula como

$$\begin{aligned} \mu_{\cap}(x, y) &= \min(\mu_{H_{ext}}(x, y), \mu_C(x, y)) \\ \mu_{\cap}(x, y) &= \min(\mu_H(x), \mu_C(x, y)) \end{aligned} \quad (4.18)$$

La proyección cilíndrica de la intersección puede obtenerse empleando la ecuación 1.8. Si denotamos por D esa proyección, es decir, la conclusión, obtenemos:

$$\begin{aligned} \mu_D(y) &= \max_X(\mu_{\cap}(x, y)) \\ \mu_D(y) &= \max_X(\min(\mu_H(x), \mu_C(x, y))) \end{aligned} \quad (4.19)$$

La ecuación 4.19 se denomina la *composición* de la inferencia. Nótese que la composición utiliza dos operadores, el máx y el mín, y por ello se denomina composición máx – mín.

Ahora bien, los operadores máx y mín pueden remplazarse por otros operadores que cumplan el mismo rol, es decir, pueden remplazarse por cualquier S -norma y por cualquier T -norma, respectivamente. Por ello, podemos obtener una versión generalizada de 4.19 así:

$$\mu_D(y) = \max_X(S_c(T_c(\mu_H(x), \mu_C(x, y))) \quad (4.20)$$

La ecuación 4.20 tiene la misma forma de la ecuación que se emplea para efectuar la composición de relaciones (ecuación 2.20). Por esta razón, puede afirmarse que la inferencia difusa es la composición de un hecho y un conocimiento:

$$\text{Conclusión} = \text{Hecho} \circ \text{Conocimiento} \quad (4.21)$$

Por otra parte, el operador mín que se utiliza en 4.16 para construir $\mu_C(x, y)$ también puede remplazarse por otro tipo de operador. En este caso, el remplazo se realiza por un operador que cumpla la función de representar la relación entre los conjuntos A y B . Tales operadores se denominan *operadores de implicación*, y se exploran en la sección 5.2. Si denotamos por I al operador de implicación, obtenemos la expresión más general de la inferencia difusa:

$$\mu_D(y) = \max_X(S_c(T_c(\mu_H(x), I(\mu_A(x), \mu_B(y)))) \quad (4.22)$$

4.4. Difusor

En la sección 4.1 se explica que el papel del difusor es el de recibir la entrada al sistema en forma de número real y convertir ese número en un conjunto difuso que será utilizado por la máquina de inferencia (figura 4.4c). Específicamente, la salida del difusor será el hecho que, en combinación con las reglas, permitirá obtener conclusiones.

En esencia, si la entrada es $x = x_o$, el difusor produce un conjunto difuso centrado en x_o . Existen diferentes tipos de difusor, según sea la forma del conjunto difuso que producen.

La figura 4.17 muestra 5 tipos de difusores. El más popular, sin duda alguna, es el difusor tipo singleton debido a que la inferencia resulta más fácil de implementar, tal como se puede verificar al comparar las figuras 4.15 y 4.16.

Sin embargo, en ocasiones se quiere modelar la incertidumbre presente en la información que entra al SBR. En esos casos el difusor tipo singleton deja de ser conveniente y puede emplearse otra de las opciones que se muestran en la figura 4.17. Estos tipos de difusor producen conjuntos difusos cuyas funciones de pertenencia se explican en la sección 3.1.1.

Cada uno de esos tipos de difusor utiliza algunos parámetros que permiten controlar la forma de los conjuntos producidos. Estos parámetros se muestran también en la figura 4.17 y su significado es el siguiente:

- A es el ancho del soporte del conjunto difuso producido.
- B es el ancho del α -corte del conjunto difuso producido, correspondiente a $\alpha = 1$.
- λ es el parámetro que se usa en la ecuación 3.6 y que controla el punto de inflexión de las curvas.
- γ es el parámetro que se usa en la ecuación 3.6 y que controla la curvatura.

4.5. Concresor

El papel del concresor (sección 4.1 y figura 4.4d) consiste en producir un número real a partir de las conclusiones generadas por la máquina de inferencia.

En la sección 4.3 se explica que la máquina de inferencia produce una conclusión al combinar un hecho con un conocimiento; el hecho proviene del difusor y el conocimiento está formalizado en reglas. En esa misma sección se muestra que cada conclusión resulta ser un conjunto difuso (figuras 4.15f y 4.16f).

Si en la base de reglas hay n reglas, la máquina de inferencia deberá obtener n conclusiones, una por cada una de las reglas en la base. Por tanto, el concresor recibirá n

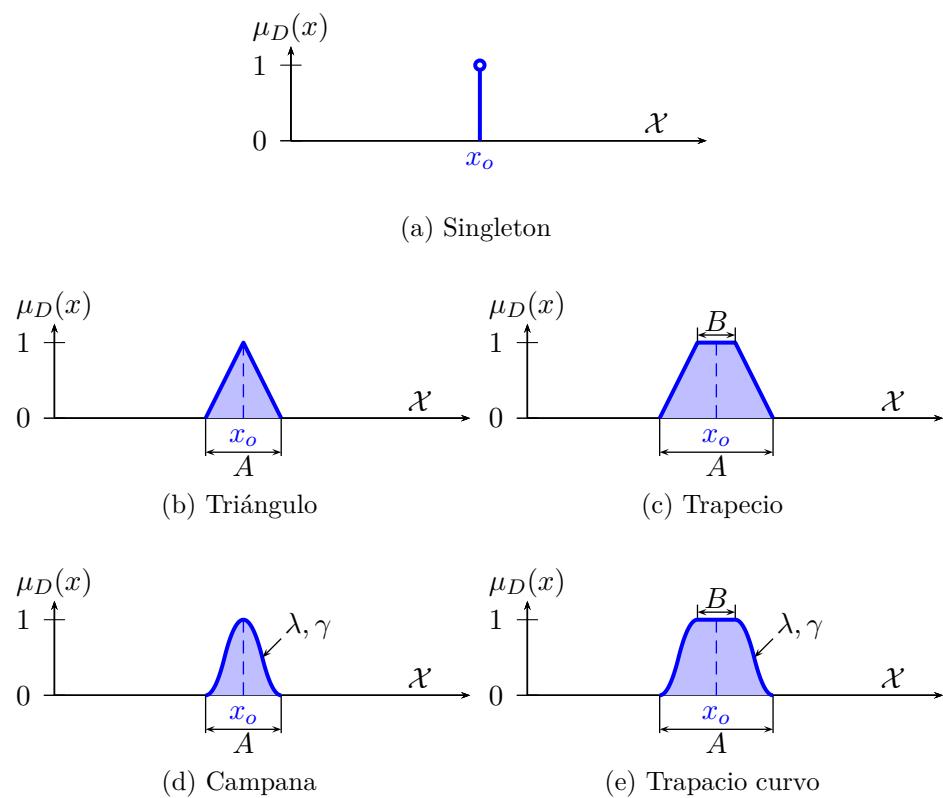


Figura 4.17: Tipos de difusor

conjuntos difusos y deberá producir un número real.

Existen dos estrategias para efectuar la tarea del congresor:

Agregación-concreción : Consiste en agregar los n conjuntos difusos en uno sólo y después obtener un número concreto que represente a ese conjunto difuso agregado.

La figura 4.18 ilustra el proceso: el congresor recibe n conjuntos difusos, uno por cada regla de la base, cuyas funciones de pertenencia son $\mu^{(1)}(y)$, $\mu^{(2)}(y)$, \dots , $\mu^{(n)}(y)$. Utilizando un *operador de agregación* \mathcal{A} se obtiene un único conjunto difuso cuya función de pertenencia es:

$$\mu_{Ag}(y) = \mathcal{A} \left\{ \mu^{(1)}(y), \dots, \mu^{(n)}(y) \right\} \quad (4.23)$$

Posteriormente se utiliza un *operador de concreción* \mathcal{C} que calcula un valor concreto y_o a partir del conjunto difuso agregado.

El operador \mathcal{A} usualmente es el operador máx o cualquier \mathcal{S} -norma, pero para ciertos tipos de operador de implicación puede necesitarse una \mathcal{T} -norma. Este aspecto se aclara en la sección 5.2.

Por otra parte, la tabla 4.3 muestra los operadores de concreción \mathcal{C} más utilizados. La forma en que operan se ilustra en la figura 4.19.

Concreción-agregación : Consiste en obtener un número concreto por cada uno de los n conjuntos difusos y después agregar esa información para obtener un único número concreto.

El operador más representativo de este tipo de congresores es el denominado *congresor de altura*, cuyo funcionamiento se ilustra en la figura 4.20: cada uno de los conjuntos difusos de la variable lingüística de salida se caracteriza por un *valor representativo* $y^{(k)}$.

Como el resultado de la inferencia resulta ser el conjunto difuso de la etiqueta del consecuente de la regla pero *recortado*, de cada conclusión pueden extraerse dos valores concretos: $y^{(k)}$, h^k ; es decir, el valor representativo y la altura a la que se efectúa el recorte.

El congresor de altura agrega la información resultante mediante un promedio ponderado de los $y^{(k)}$, en donde el coeficiente de ponderación es la altura de la conclusión. Es decir:

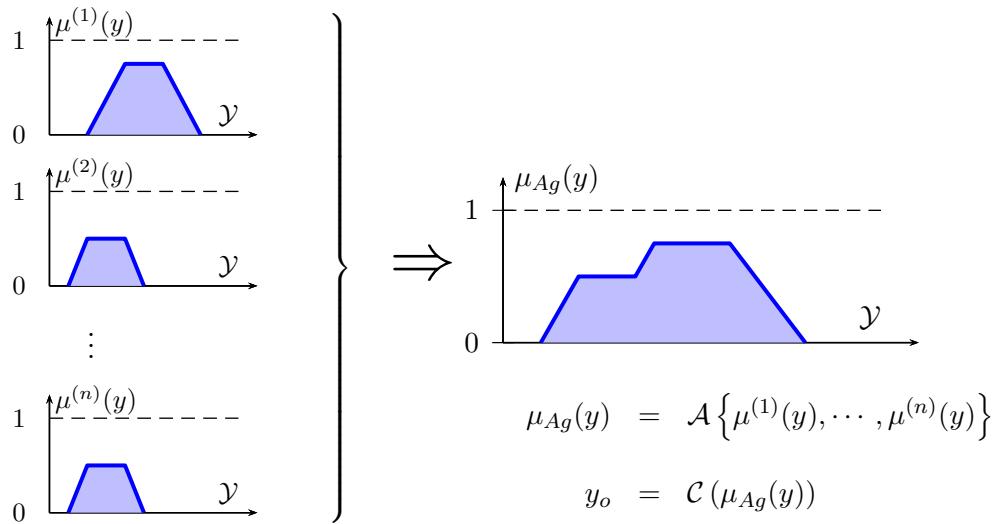


Figura 4.18: Esquema de agregación-concreción

$$y_o = \frac{\sum_{k=1}^n y^{(k)} h^k}{\sum_{k=1}^n h^k} \quad (4.24)$$

Tabla 4.3: Funciones de concreción

Tipo	Explicación	Función de concreción $\mathcal{C}(\mu_{Ag}(y))$
Primer máximo	Toma el valor de y más pequeño entre aquellos que tienen el grado de pertenencia más alto	$\inf \left\{ \bar{y} \in \mathcal{Y} \mid \mu_{Ag}(\bar{y}) = \max (\mu_{Ag}(y)) \right\}$
Último máximo	Toma el valor de y más alto entre aquellos que tienen el grado de pertenencia más alto	$\sup \left\{ \bar{y} \in \mathcal{Y} \mid \mu_{Ag}(\bar{y}) = \max (\mu_{Ag}(y)) \right\}$
Media de máximos	Toma el promedio de los valores de y entre aquellos que tienen el grado de pertenencia más alto	media $\left\{ \bar{y} \in \mathcal{Y} \mid \mu_{Ag}(\bar{y}) = \max (\mu_{Ag}(y)) \right\}$
Centro de gravedad	Toma el promedio de y ponderado por el grado de pertenencia	$\frac{\int_{\mathcal{Y}} y \mu_{Ag}(y) dy}{\int_{\mathcal{Y}} \mu_{Ag}(y) dy}$

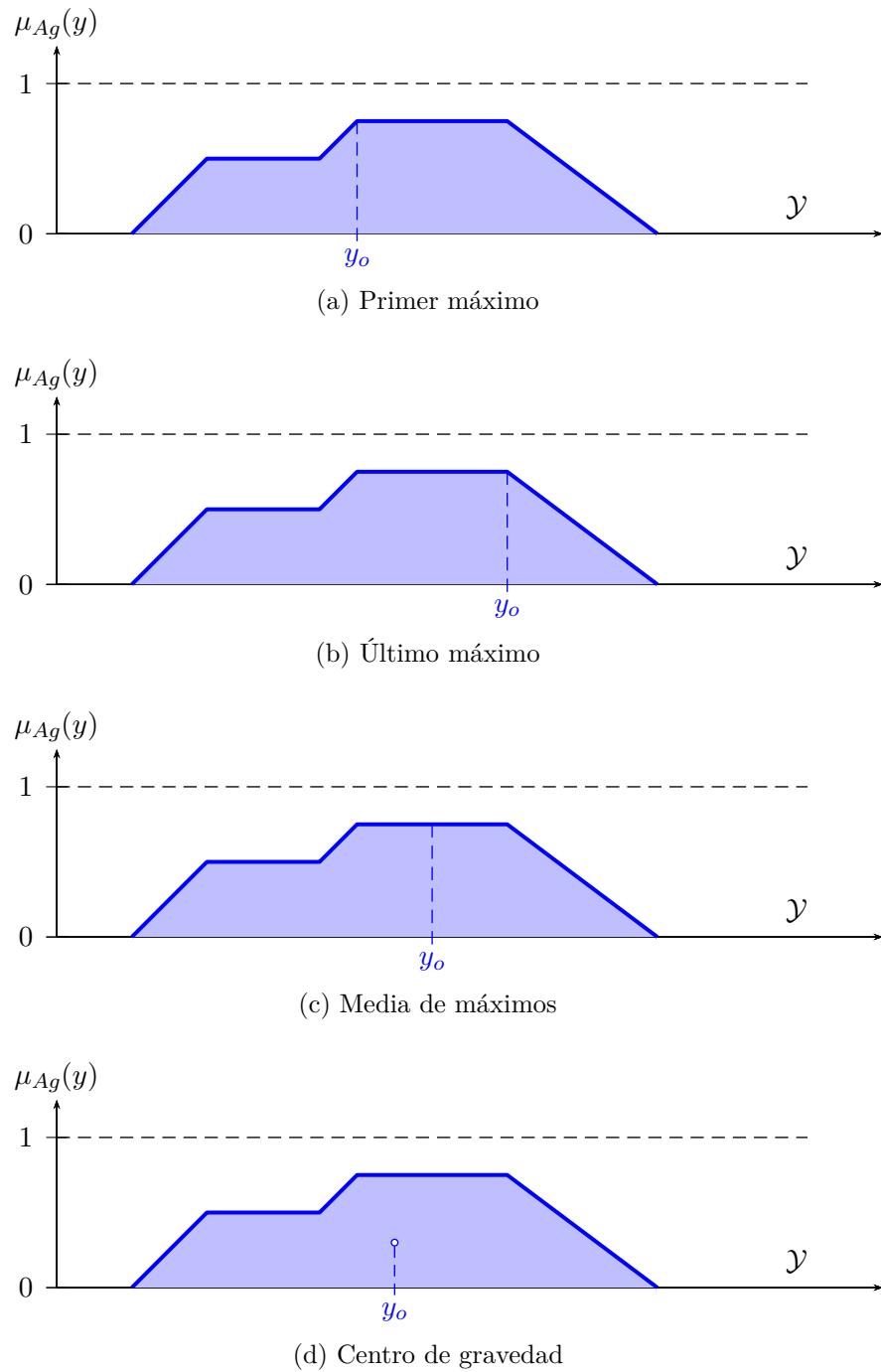


Figura 4.19: Tipos de funciones de concreción.

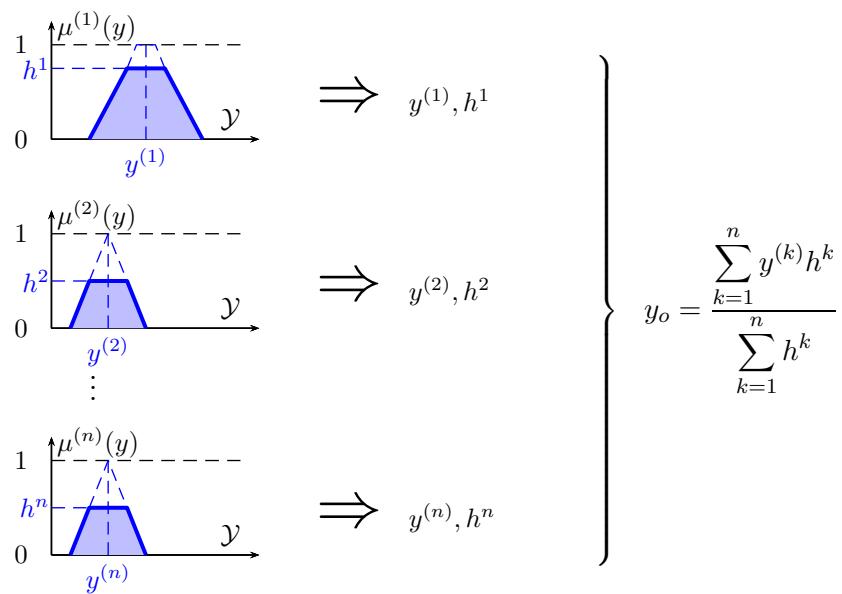


Figura 4.20: Esquema de concreción-agregación. Concresor tipo Altura.

Ejemplo 4.9 (Sistema tipo Mamdani) En este ejemplo se muestra el proceso de agregación de reglas. Para hacerlo, utilizaremos un sistema SISO con tres etiquetas a la entrada y a la salida (Bajo, Medio, Alto). El universo de discurso de las dos variables es el intervalo $[-1, 1]$ y sus variables lingüísticas son del tipo conjuntos cortos rectos. También se ha definido una base de reglas creciente ($\text{Si Bajo} \Rightarrow \text{Bajo}$, $\text{Si Medio} \Rightarrow \text{Medio}$, $\text{Si Alto} \Rightarrow \text{Alto}$).

Otras características importantes del sistema son las siguientes:

-
- El difusor es del tipo singleton.
- La representación del conocimiento (ecuación 4.16) emplea el operador \min .
- La inferencia emplea (ecuación 4.19) los operadores \max y \min .
- La agregación de conclusiones (ecuación 4.23) emplea el operador \max .
- La concreción se realiza con el operador Centro de Gravedad (Tabla 4.3).

Un SBR con los operadores anteriormente listados corresponde a un *Sistema tipo mamdani*. La figura 4.21 muestra la relación Entrada-Salida de nuestro sistema.

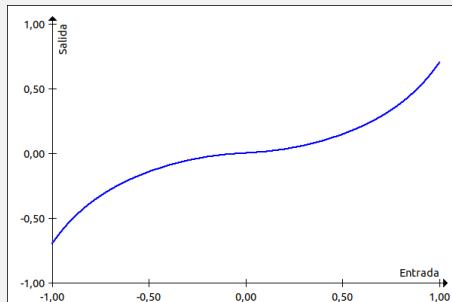


Figura 4.21: Relación Entrada-Salida del ejemplo 4.9

Se ha elaborado la figura 4.22 para visualizar el proceso interno de inferencia y concreción. Para ello se han considerado 9 entradas diferentes ($x = -1.0, x = -0.75, \dots, x = 0.65, x = 1.0$). Para cada entrada se han trazado 5 gráficas:

- La primera gráfica muestra en azul la variable lingüística de entrada y en rojo la salida del difusor.
- Las siguientes tres gráficas muestran las conclusiones obtenidas en cada una de las tres reglas. Nótese que las conclusiones corresponden a los conjuntos recortados de los consecuentes.

- La quinta gráfica muestra en azul el resultado de agregar las conclusiones de las reglas y en rojo la salida del congresor.

En la primera columna se observa que al aumentar x desde -1 hasta 1 el singleton se desplaza hacia la derecha y por tanto la intersección con los tres conjuntos de la variable lingüística va cambiando.

Consideremos la columna 2, correspondiente a la primera regla (*Si Bajo \Rightarrow Bajo*). El conjunto Bajo es un triángulo cuyos vértices están en -1 , -1 y 0 . Cuando $x = -1$ la conclusión es el mismo conjunto Medio. Conforme vamos aumentando el valor de x desde -1 hasta 0 , la altura a la que se efectúa el recorte del conjunto va disminuyendo. A partir de $x = 0$ la regla produce un conjunto vacío, es decir, no genera ninguna conclusión porque la regla no se activa.

Veamos ahora la columna 3, correspondiente a la segunda regla (*Si Medio \Rightarrow Medio*). El conjunto Medio es un triángulo cuyos vértices están en -1 , 0 y 1 . Conforme vamos aumentando el valor de x desde -1 hasta 0 , la altura a la que se efectúa el recorte del conjunto va aumentando; cuando $x = 0$ la conclusión es el conjunto completo. Si seguimos aumentando x desde 0 hasta 1 la altura a la que se efectúa el recorte del conjunto va disminuyendo.

Ya hora estudiemos la columna 4, correspondiente a la segunda regla (*Si Alto \Rightarrow Alto*). El conjunto Alto es un triángulo cuyos vértices están en 0 , 1 y 1 . La regla no se activa para valores de x menores que 0 . Conforme vamos aumentando el valor de x desde 0 hasta 1 , la altura a la que se efectúa el recorte del conjunto va aumentando; cuando $x = 1$ la conclusión es el conjunto completo.

Veamos ahora la columna 4 que contiene la agregación de las conclusiones y la concreción. Recordemos que la agregación se efectúa con el operador máx, es decir, se efectúa la unión de las conclusiones de las reglas. Es fácil observar que el conjunto resultante se va desplazando paulatinamente hacia la derecha; por esa razón su centro de gravedad se desplaza también paulatinamente hacia la derecha, es decir, la salida del congresor aumenta paulatinamente. Por ello, la relación Entrada-Salida resulta ser creciente (Figura 4.21).

4.6. Sistemas de múltiples entradas y múltiples salidas

La explicación de los bloques de un SBR que se consigna en las secciones 4.2 a 4.5 corresponde a sistemas de una entrada y una salida. No obstante, es posible construir SBR de múltiples entradas y múltiples salidas (ver Figura 4.2). En esta sección se explica qué modificaciones hay que realizar para lograrlo. Inicialmente, en la sección 4.6.1 se aborda la construcción de sistemas tipo MISO (*Multiple Inputs - Single Output*) y

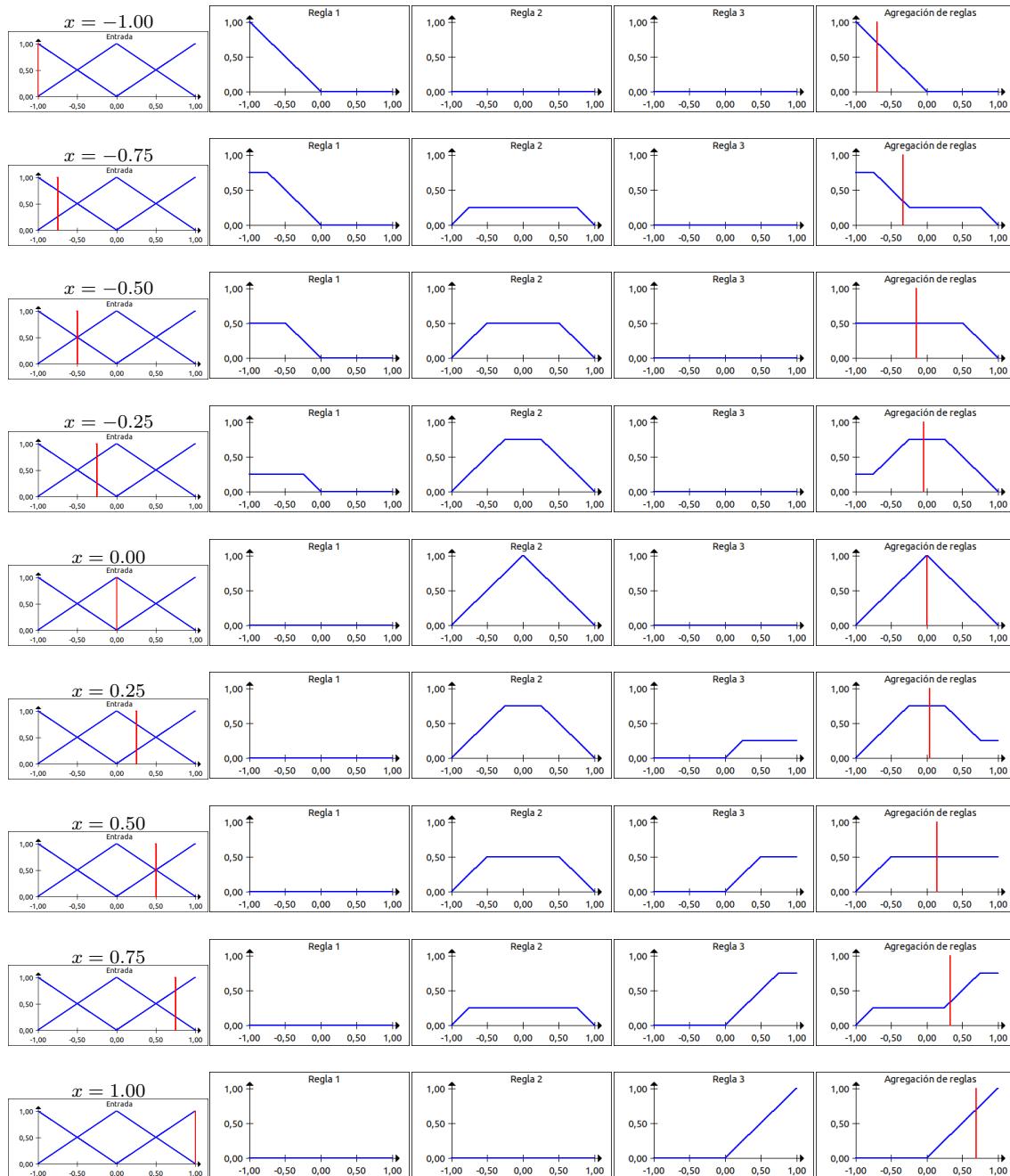


Figura 4.22: Cálculo paso a paso. Implicación tipo Mamdani

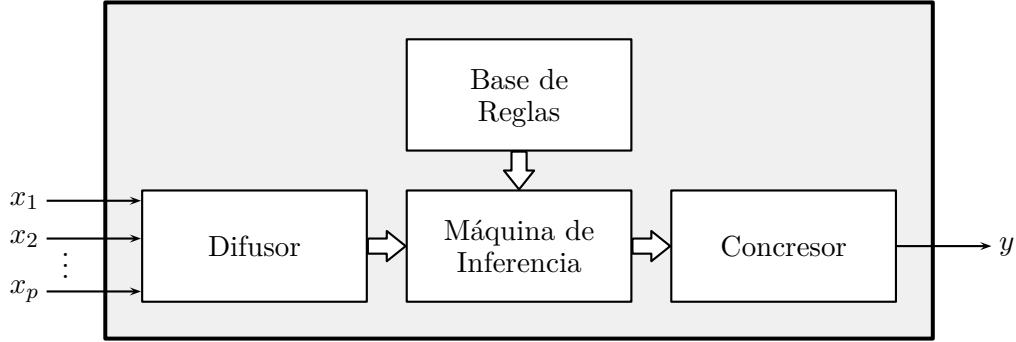


Figura 4.23: Sistema de múltiples entradas y una salida

posteriormente, en la sección 4.6.2 se aborda la construcción de sistemas tipo MIMO (*Multiple Inputs - Multiple Output*).

Para facilitar la explicación se utilizará la siguiente notación:

- p es el número de entradas del sistema.
- q es el número de salidas del sistema.
- x_1, x_2, \dots, x_p son las variables de entrada del sistema.
- y_1, y_2, \dots, y_q son las variables de salida del sistema.
- \mathbf{x} es un vector columna que contiene las variables de entrada del sistema.
- \mathbf{y} es un vector columna que contiene las variables de salida del sistema.

En resumen, las entradas y salidas al sistema se denotarán por:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_q \end{bmatrix} \quad (4.25)$$

4.6.1. Sistemas tipo MISO

Consideremos ahora un sistema MISO como el que se muestra en la Figura 4.23. Cada una de las p entradas deberá tener una variable lingüística asociada y, a su vez, cada una de esas variables tendrá un conjunto de etiquetas.

Una regla de comportamiento debe especificar la relación entre las entradas y la salida. La forma general de la regla número k será ahora:

$$\text{Si } x_1 \text{ es } \mathcal{X}_1^{(k)} \text{ Y } x_2 \text{ es } \mathcal{X}_2^{(k)} \text{ Y } \dots \text{ Y } x_p \text{ es } \mathcal{X}_p^{(k)} \text{ entonces } y \text{ es } \mathcal{Y}^{(k)} \quad (4.26)$$

En donde $\mathcal{X}_j^{(k)}$ es una de las etiquetas de x_j y $\mathcal{Y}^{(k)}$ es una de las etiquetas de la variable de salida y . Es decir, el antecedente de cada regla será una proposición compuesta que se construye conectando varias proposiciones simples con el conector lógico **Y**, es decir, con una conjunción.

Para poder efectuar una inferencia a partir de ese tipo de reglas, es necesario adecuar el difusor y la inferencia, como se muestra en las secciones 4.6.1.1 y 4.6.1.2. El bloque concresor no requiere cambios, porque la máquina de inferencia genera una conclusión por cada regla, tal como en el caso SISO.

Ejemplo 4.10 (SBR tipo MISO) En este ejemplo se muestra el comportamiento de un sistema de dos entradas y una salida. La Figura 4.24 ayuda a presentar algunas de las características de ese sistema:

- Los nombres de las entradas son x_1 , x_2 .
- El nombre de la salida es y .
- La variable lingüística de x_1 tiene tres etiquetas: B, M, A. Estos son nombres abreviados para Bajo, Medio y Alto.
- La variable lingüística de x_2 tiene cinco etiquetas: MB, B, M, A, MA. Estos son nombres abreviados para Muy Bajo, Bajo, Medio, Alto y Muy Alto.
- La variable lingüística de y tiene siete etiquetas: MMB, MB, B, M, A, MA, MMA. Estos son nombres abreviados para Muy-Muy Bajo, Muy Bajo, Bajo, Medio, Alto, Muy Alto y Muy-Muy Alto.
- Las tres variables lingüísticas tienen conjuntos curvos cortos.

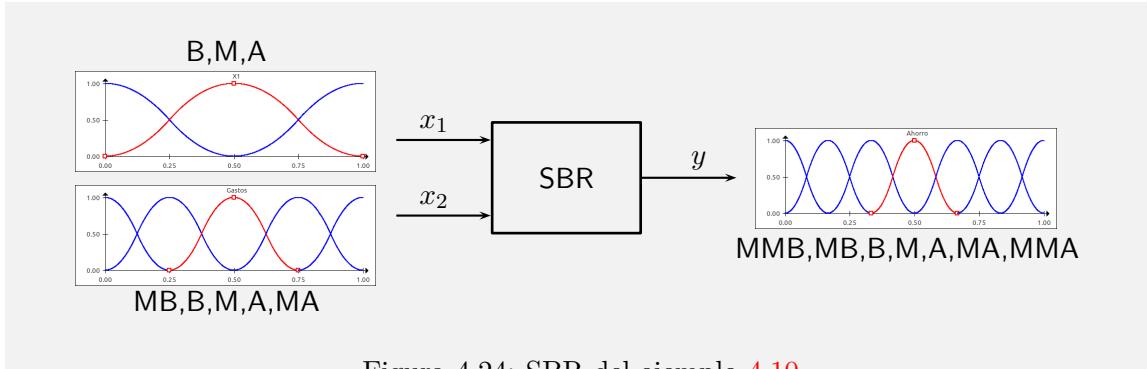


Figura 4.24: SBR del ejemplo 4.10

Consideremos ahora la Base de Reglas del sistema. En este ejemplo queremos construir un sistema cuya salida crezca conforme x_1 y x_2 van creciendo. En otras palabras, cuando las entradas sean pequeñas la salida debe ser pequeña y cuando las entradas sean grandes la salida debe ser grande. Este conocimiento podría reflejarse en dos reglas:

- Si x_1 es Bajo Y x_2 es Muy Bajo entonces y es Muy-Muy Bajo.
- Si x_1 es Alto Y x_2 es Muy Alto entonces y es Muy-Muy Alto.

Tal como esperábamos, el antecedente de cada una de esas reglas combina dos proposiciones. Debido a que x_1 tiene tres etiquetas y a que x_2 tiene cinco etiquetas, existen $3 \times 5 = 15$ posibles combinaciones diferentes de esas proposiciones para construir un antecedente. Eso significa que necesitamos 15 reglas para tener una base completa.

En lugar de escribir las 15 reglas, elaboramos la Tabla 4.4. En ella, cada fila corresponde a una etiqueta de x_1 , cada columna a una etiqueta de x_2 y en cada celda aparece una etiqueta de y .

Cada celda corresponde a una regla. La fila y la columna de esa celda corresponden a las etiquetas del antecedente y el contenido de la celda corresponde al consecuente. A manera de ejemplo, se ha resaltado en azul la regla

- Si x_1 es Bajo Y x_2 es Alto entonces y es Medio.

Tabla 4.4: Base de reglas del ejemplo 4.10

$x_2 \backslash x_1$	MB	B	M	A	MA
B	MMB	MB	B	M	A
M	MB	B	M	A	MA
A	B	M	A	MA	MMA

Nótese que al recorrer la tabla hacia la izquierda o hacia abajo las etiquetas del consecuente

van siendo mayores. De esta forma se logra el efecto creciente deseado, tal como puede verificarse al trazar la relación Entrada-Salida. Esta relación se muestra en la Figura 4.25. Es una superficie en \mathbb{R}^3 debido a que hay dos entradas y una salida.

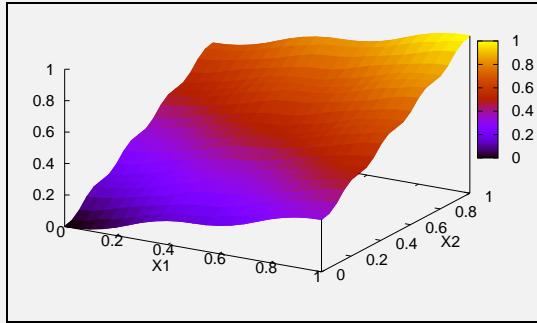


Figura 4.25: Relación Entrada-Salida del ejemplo 4.10

Ejemplo 4.11 (SBR tipo MISO) Retomemos el ejemplo 4.11 y modifiquemos su base de reglas. La nueva base se muestra en la Tabla 4.5.

Tabla 4.5: Base de reglas del ejemplo 4.11

$x_2 \backslash x_1$	MB	B	M	A	MA
B	MMB	MB	B	MB	MMB
M	MB	B	M	B	MB
A	MMB	MB	B	MB	MMB

Nótese que al recorrer la tabla hacia la izquierda o hacia abajo las etiquetas del consecuente primero van siendo mayores, pero a partir de la mitad de la tabla van siendo menores. La nueva relación Entrada-Salida se muestra en la Figura 4.26.

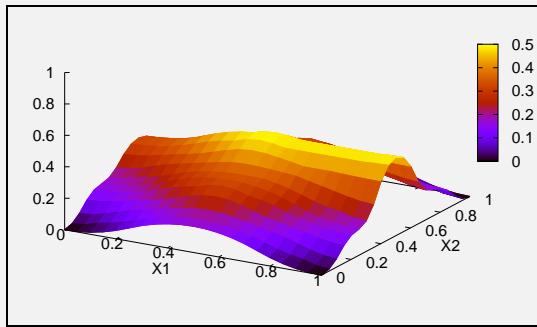


Figura 4.26: Relación Entrada-Salida del ejemplo 4.11

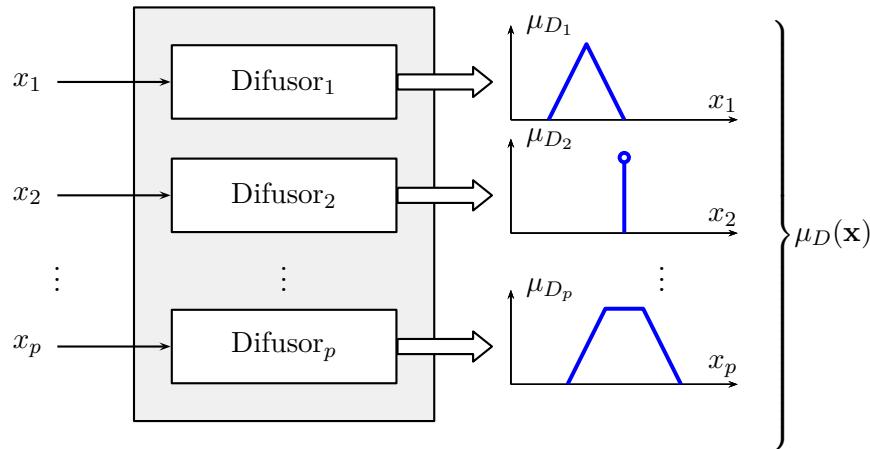


Figura 4.27: Difusor para múltiples entradas

4.6.1.1. Difusor para múltiples entradas

La Figura 4.27 ilustra cómo se ha concebido el difusor para múltiples entradas: para cada una de las entradas se diseña un difusor de una sola entrada, y posteriormente se combinan sus salidas para obtener $\mu_D(\mathbf{x})$.

La forma en que se combinan las salidas de los difusores debe reflejar la forma en que se combinan las proposiciones simples en los antecedentes de las reglas. Como esta combinación se efectúa con el operador **Y** (ecuación 4.26), esta combinación se realiza a través de una \mathcal{T} -norma:

$$\begin{cases} \mu_D(\mathbf{x}) = \mu_{D_1}(x_1) \mathbf{Y} \mu_{D_2}(x_2) \mathbf{Y} \cdots \mathbf{Y} \mu_{D_p}(x_p) \\ \mu_D(\mathbf{x}) = \mathcal{T}_{AND} \{\mu_{D_1}(x_1), \mu_{D_2}(x_2), \dots, \mu_{D_p}(x_p)\} \end{cases} \quad (4.27)$$

4.6.1.2. Inferencia para múltiples entradas

De acuerdo con la ecuación 5.10, en los sistemas SISO el conocimiento de una regla de la forma ‘si x es $A \Rightarrow y$ es B ’ se representa por

$$\mu_C(x, y) = \mathcal{I}(\mu_A(x), \mu_B(y)) \quad (4.28)$$

En donde \mathcal{I} es un operador de implicación (ecuación 5.11). Para representar el conocimiento de una regla como la de la ecuación 4.26 debemos ahora considerar que el antecedente está formado por varias proposiciones conectadas con el operador **Y**.

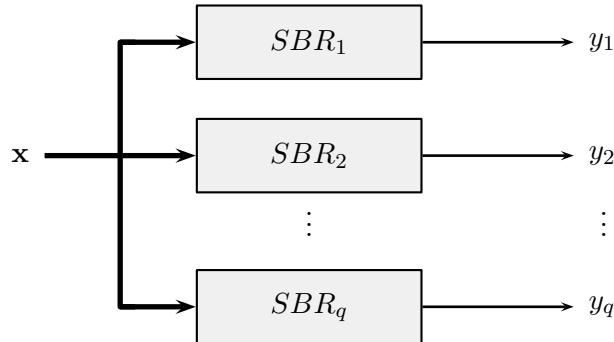


Figura 4.28: SBR tipo MIMO con sistemas MISO en paralelo

Si consideremos la k -ésima regla, la grado de pertenencia debe entonces calcularse así:

$$\mu_C^{(k)}(\mathbf{x}, y) = \mathcal{I}\left(\mu_x^{(k)}(\mathbf{x}), \mu_y^{(k)}(y)\right) \quad (4.29)$$

Por su parte, la función de pertenencia del antecedente de múltiples entradas debe calcularse con una \mathcal{T} -norma, para reflejar la naturaleza del conector \mathbf{Y} . Es decir:

$$\begin{cases} \mu_x^{(k)}(\mathbf{x}) = \mu_{x_1}^{(k)}(x_1) \text{ AND } \mu_{x_2}^{(k)}(x_2) \cdots \text{ AND } \mu_{x_p}^{(k)}(x_p) \\ \mu_x^{(k)}(\mathbf{x}) = \mathcal{T}_{AND} \left\{ \mu_{x_1}^{(k)}(x_1), \mu_{x_2}^{(k)}(x_2), \dots, \mu_{x_p}^{(k)}(x_p) \right\} \end{cases} \quad (4.30)$$

4.6.2. Sistemas tipo MIMO

De forma general, un sistema tipo MIMO con p entradas y q salidas se puede construir siguiendo una de dos posibles estrategias:

- Construyendo q sistemas tipo MISO que reciben la misma entrada. Esta estrategia se ilustra en la figura 4.28.
- Construyendo un único sistema que internamente tiene q concresores independientes. Esta estrategia se ilustra en la figura 4.29

En cualquier caso, los sistemas tipo MIMO no son muy empleados en la práctica, porque es preferible considerarlos como sistemas independientes.

4.6.3. La maldición de la dimensionalidad

Los ejemplos 4.10 y 4.11 sirven para mostrar que el número de reglas necesario para tener una base completa depende del número de entradas y del número de etiquetas de

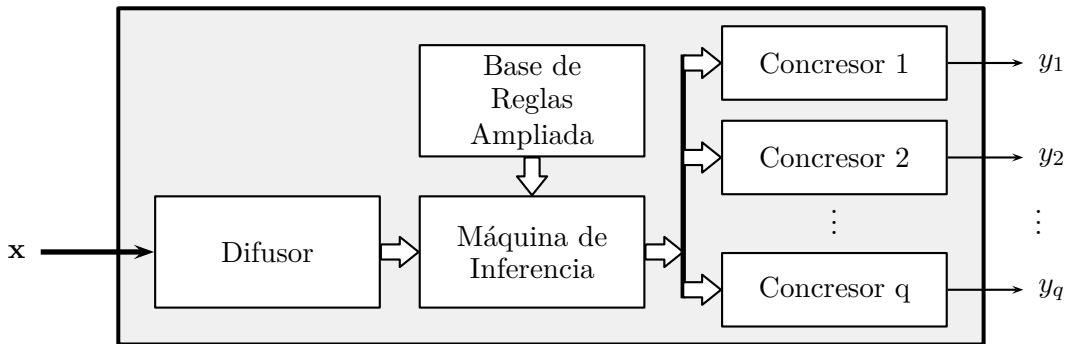


Figura 4.29: SBR tipo MIMO con inferencia común

Tabla 4.6: Explosión de la base de reglas

$\begin{matrix} L \\ \diagdown \\ p \end{matrix}$	2	3	4	5
1	2	3	4	5
2	4	9	16	25
3	8	27	64	125
4	16	81	256	625
5	32	243	1024	3125
10	1024	59049	1048576	9765625

las variables lingüísticas de entrada. Lo anterior se debe a que se necesita que exista una regla para cada posible combinación de etiquetas de las entradas.

Específicamente, si un sistema tiene p entradas y el número de etiquetas lingüísticas de cada una de ellas es L_1, L_2, \dots, L_q entonces el número R de reglas necesario está dado por

$$R = L_1 \times L_2 \times \dots \times L_p = \prod_{i=1}^p L_i \quad (4.31)$$

Una de las consecuencias de la ecuación 4.31 es que el número de reglas necesario crece muy rápidamente conforme crece el número de entradas ó de etiquetas. Este hecho se pone de evidencia en la Tabla 4.6 en la que se ha calculado R para distintos escenarios. Este fenómeno se conoce como la *explosión de la base de reglas* y no es otra cosa que la manifestación de que los sistemas basados en reglas están afectados por la *maldición de la dimensionalidad*.

Capítulo 5

Anotaciones sobre los SBR

Hay una literatura extensa sobre los SBR en la que se formulan variantes, aplicaciones y se estudian sus propiedades. En este capítulo presentamos someramente dos aspectos que destacamos por su relevancia: 1) En la sección 5.1 se presenta una variación del proceso de inferencia que permite el ajuste del desempeño de un SBR. 2) En la sección 5.2 se presentan los operadores de implicación difusa, que también pueden usarse como base para la inferencia en reglas de la forma ‘si antecedente entonces consecuente’. 3) En la sección 5.3 se presenta una de las aplicaciones más conocidas de los SBR, consistente en el cálculo de las reglas de control para sistemas dinámicos.

5.1. Modificadores lingüísticos

Un modificador lingüístico es un operador matemático que cambia el contenido semántico de un conjunto difuso. En esta sección se presenta un tipo de modificador que consiste en elevar el grado de pertenencia a algún exponente positivo, es decir:

$$\mu_{mod}(x) = (\mu(x))^\varsigma \quad (5.1)$$

La Figura 5.1 muestra el efecto de este tipo de modificadores, para algunos valores de ς . Se ha empleado una función de pertenencia triangular para ello, que corresponde a la gráfica de la mitad, en la que $\varsigma = 1$.

Nótese que valores de $\varsigma > 1$ angostan la función, mientras que valores de $\varsigma < 1$ la ensanchan. En particular, se pueden definir dos operadores basados en los modificadores lingüísticos:

- Operador de concentración: emplea $\varsigma = 2.0$

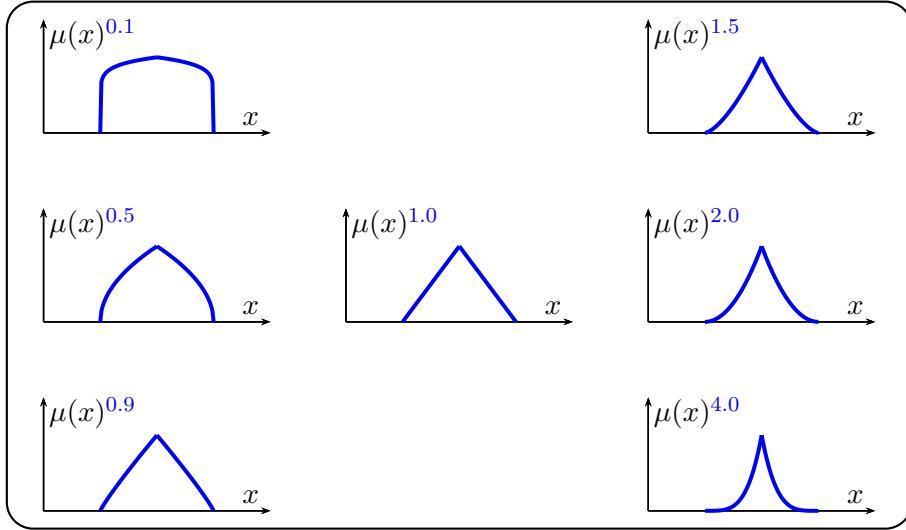


Figura 5.1: Efecto de los modificadores lingüísticos

- Operador de dilatación: emplea $\varsigma = 0.5$

Algunos autores que interpretan la concentración y la dilatación en términos lingüísticos. La concentración la interpretan como una definición más restrictiva del contenido semántico y la dilatación como una definición menos restrictiva; a partir de allí, sugieren utilizar los adverbios ‘muy’ y ‘poco’ para expresar en palabras el efecto del modificador lingüístico. Sin embargo, al menos en castellano el uso de esos adverbios no resulta adecuado.

Veamos ahora cómo se pueden incorporar los modificadores lingüísticos en un SBR. En la sección 4.6.1.2 se explica el proceso de inferencia para sistemas tipo MISO. Allí se explica que el conocimiento de la k -ésima regla como la de la ecuación 4.26 se representa por

$$\mu_C^{(k)}(\mathbf{x}, y) = \mathcal{I}\left(\mu_x^{(k)}(\mathbf{x}), \mu_y^{(k)}(y)\right) \quad (5.2)$$

en donde \mathcal{I} es un operador de implicación que opera sobre el antecedente y el consecuente. También se explica que la función de pertenencia del antecedente debe calcularse como:

$$\mu_x^{(k)}(\mathbf{x}) = \mathcal{T}_{AND} \left\{ \mu_{x_1}^{(k)}(x_1), \mu_{x_2}^{(k)}(x_2), \dots, \mu_{x_p}^{(k)}(x_p) \right\} \quad (5.3)$$

Podemos definir para cada regla un conjunto de p modificadores lingüísticos para aplicarlos a cada una de las funciones de pertenencia del antecedente en 5.3. Si denotamos por ς a esos modificadores podemos escribir:

$$\varsigma^{(k)} = [\varsigma_1^{(k)} \quad \varsigma_2^{(k)} \quad \dots \quad \varsigma_p^{(k)}] \quad (5.4)$$

Y calcular ahora la función de pertenencia del antecedente así:

$$\mu_x^{(k)}(\mathbf{x}) = \mathcal{T}_{AND} \left\{ \left(\mu_{x_1}^{(k)}(x_1) \right)^{\varsigma_1^{(k)}}, \left(\mu_{x_2}^{(k)}(x_2) \right)^{\varsigma_2^{(k)}}, \dots, \left(\mu_{x_p}^{(k)}(x_p) \right)^{\varsigma_p^{(k)}} \right\} \quad (5.5)$$

Ejemplo 5.1 (Modificadores lingüísticos) Para ilustrar el efecto de los modificadores lingüísticos se ha construido un SBR tipo SISO como el de la figura 4.8a. Ese sistema tiene tres etiquetas a la entrada, tres etiquetas a la salida y una base de tres reglas que se muestra en la Tabla 5.1

Tabla 5.1: Base de reglas en el ejemplo 5.1

Regla	Antecedente	Consecuente
R1	1	1
R2	2	2
R3	3	3

El antecedente de cada regla es una única etiqueta. Por esa razón, podemos definir hasta 3 modificadores lingüísticos. Por ejemplo, podemos usar:

$$\varsigma_1^{(1)} = 2.0 \quad \varsigma_1^{(2)} = 1.0 \quad \varsigma_1^{(3)} = 0.5 \quad (5.6)$$

La Figura 5.2 muestra el efecto del uso de esos modificadores. Allí se ha dibujado la relación entrada - salida en dos condiciones: sin usar modificadores (en azul) y usando los modificadores (en rojo). Es evidente que los modificadores alteran la relación.

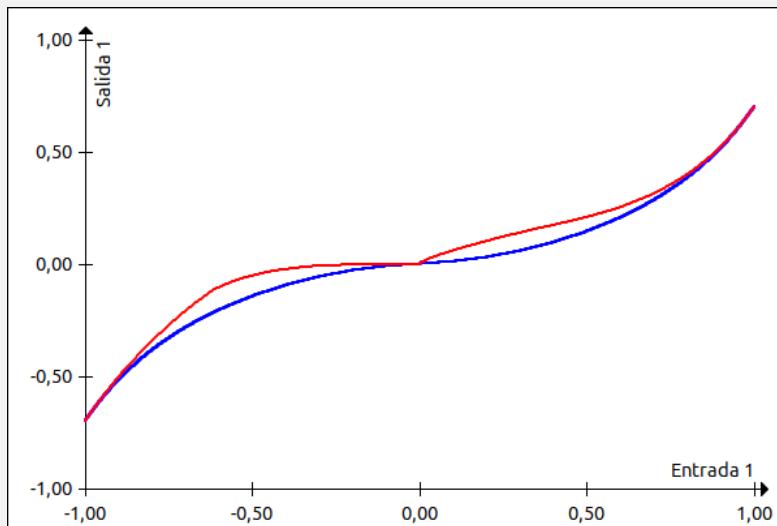


Figura 5.2: Efecto de los modificadores lingüísticos en el ejemplo 5.1. En azul sin modificadores y en rojo con modificadores.

Tabla 5.2: Función de verdad del operador de implicación

Implicación		
p	q	$p \Rightarrow q$
F	F	V
F	V	V
V	F	F
V	V	V

5.2. Sobre las implicaciones difusas

En la sección 4.3.3 se explica que una regla de la forma ‘si antecedente entonces consecuente’ debe interpretarse como una relación de causalidad entre el antecedente y el consecuente: el hecho de que a la entrada suceda el antecedente tiene como consecuencia que a la salida suceda el consecuente. Esta interpretación es la que justifica que el conocimiento se pueda representar usando una \mathcal{T} -norma en la ecuación 4.14. Sin embargo, existe otra forma de interpretar ese tipo de reglas empleando el conector lógico de implicación.

5.2.1. Implicación concreta

En la lógica proposicional concreta \Rightarrow es el *operador de implicación*. La expresión $A \Rightarrow B$ se lee ‘A implica B’. A es el antecedente y B el consecuente.

La función de verdad de la implicación está dada por la Tabla 5.2. El significado de esa tabla puede expresarse con la máxima según la cual *la verdad nunca implica falsedad*, ya que:

- Un antecedente falso puede implicar un consecuente falso.
- Un antecedente falso puede implicar un consecuente verdadero.
- Un antecedente verdadero *no* puede implicar un consecuente falso.
- Un antecedente verdadero puede implicar un consecuente verdadero.

El operador de implicación puede derivarse de diferentes formas a partir de los operadores básicos de conjunción (\wedge), disyunción (\vee) y negación (\neg). Por ejemplo, las siguientes dos tautologías, que se demuestran en las Tablas 5.3 y 5.4, permiten hacerlo:

Tabla 5.3: Primera tautología del operador de implicación

p	q	$\neg p$	$\neg p \vee q$	\Leftrightarrow	$p \Rightarrow q$
F	F	V	V	V	V
F	V	V	V	V	V
V	F	F	F	V	F
V	V	F	V	V	V

Tabla 5.4: Segunda tautología del operador de implicación

p	q	$\neg p$	$p \wedge q$	$\neg p \vee (p \wedge q)$	\Leftrightarrow	$p \Rightarrow q$
F	F	V	F	V	V	V
F	V	V	F	V	V	V
V	F	F	F	F	V	F
V	V	F	V	V	V	V

$$\begin{aligned} p \Rightarrow q &\Leftrightarrow \neg p \vee q \\ p \Rightarrow q &\Leftrightarrow \neg p \vee (p \wedge q) \end{aligned} \tag{5.7}$$

5.2.2. Implicaciones difusas

Así como los operadores de conjunción y disyunción pueden extenderse a la lógica difusa empleando \mathcal{T} -normas y \mathcal{S} -normas, cabe preguntarse si el operador \Rightarrow puede también ser extendido.

La respuesta es afirmativa. De hecho, existe una muy rica literatura especializada en ese tópico (Véase [3]) que aún hoy en día sigue explorando las propiedades, aplicaciones y formas de construcción de operadores difusos de implicación.

Una implicación es un operador $I : [0, 1]^2 \rightarrow [0, 1]$ que permite calcular la función de verdad de la Tabla 5.2. Por ello, debe cumplir que:

$$\left\{ \begin{array}{lcl} I(0, 0) & = & 1 \\ I(0, 1) & = & 1 \\ I(1, 0) & = & 0 \\ I(1, 1) & = & 1 \end{array} \right. \tag{5.8}$$

Adicionalmente, suele exigirse que el operador cumpla también algunas condiciones

Tabla 5.5: Algunos operadores de implicación difusos

Nombre	$I(x, y)$
Kleene-Dienes	$\max(1 - x, y)$
Lukasiewicz	$\min(1 - x + y, 1)$
Zadeh	$\max(1 - x, \min(x, y))$
Estocástica	$\max(xy, 1 - x)$
Goguen	$\begin{cases} 1 & \text{si } x \leq y \\ y/x & \text{si } x > y \end{cases}$
Gödel	$\begin{cases} 1 & \text{si } x \leq y \\ y & \text{si } x > y \end{cases}$
Aguda	$\begin{cases} 1 & \text{si } x < 1 \text{ ó } y = 1 \\ 0 & \text{si } x = 1 \text{ y } y < 1 \end{cases}$

adicionales como¹: Para todos los posibles valores de x, x_1, x_2, y, y_1, y_2

$$\begin{cases} x_1 < x_2 \Rightarrow I(x_1, y) \geq I(x_2, y) \\ y_1 < y_2 \Rightarrow I(x, y_1) \leq I(x, y_2) \end{cases} \quad (5.9)$$

La Tabla 5.5 muestra algunos operadores de implicación difusos. Lo que es importante destacar es que todos ellos buscan calcular la validez de una conclusión, que puede ser cierta o falsa.

5.2.3. Sistemas con inferencias de implicación

Es posible diseñar un SBR en donde la representación del conocimiento emplee una implicación difusa en lugar de una \mathcal{T} -norma. Para ello, definimos una versión más general de la ecuación 4.16 es la siguiente:

$$\mu_C(x, y) = \mathcal{I}(\mu_A(x), \mu_B(y)) \quad (5.10)$$

¹En la literatura técnica se encuentran varias definiciones de operadores de implicación que no necesariamente coinciden en todas las condiciones impuestas. En este texto hemos querido dejar la versión más elemental posible.

En donde $\mathcal{I}(x, y)$ es el *operador de implicación* que puede ser de dos tipos:

$$\mathcal{I}(x, y) : \begin{cases} \mathcal{T}_C(x, y) & \text{Una } \mathcal{T}\text{-norma} \\ \mathcal{I}(x, y) & \text{Una implicación difusa} \end{cases} \quad \text{ó} \quad (5.11)$$

Sin embargo, estas dos opciones no son equivalentes. Cada una de ellas tiene un significado diferente:

- Al usar una \mathcal{T} -norma el conocimiento expresado en la regla ‘Si x es $A \Rightarrow y$ es B ’ establece una relación entre de causalidad entre A y B . Es especialmente importante destacar que B **no puede suceder** si ‘ x es A ’ no es cierto.
- Al usar una implicación difusa el conocimiento expresado en la regla ‘Si x es $A \Rightarrow y$ es B ’ establece la forma en que se propaga la verdad o la falsedad de la proposición ‘ x es A ’ a través de la regla. Es especialmente importante destacar que B **si puede suceder** aunque ‘ x es A ’ no sea cierto.

Lo anterior tiene una consecuencia muy importante para la función de agregación de las conclusiones que se obtienen a partir de todas las reglas. Cada regla aporta una conclusión parcial; la agregación debe fusionar esas conclusiones parciales y producir una conclusión general. Debido a que cada uno de las dos opciones interpreta de forma diferente lo que puede o no suceder si ‘ x es A ’ no es cierto, cada una de las dos opciones requiere un tipo de fusión diferente:

- Las conclusiones obtenidas por una \mathcal{T} -norma deben *unirse*, porque cada una de ellas sólo es verdadera cuando el antecedente es verdadero.
- Las conclusiones obtenidas por una implicación difusa deben *intersectarse*, porque la conclusión general debe identificar las conclusiones que son verdaderas en todas las reglas.

El ejemplo 5.2 ayuda a entender mejor la afirmación anterior.

Ejemplo 5.2 (Sistema con implicación de Lukasiewicz) En este ejemplo se diseña un sistema muy semejante al sistema tipo Mamdani del ejemplo 4.9. Los cambios que se realizan son los siguientes:

- En lugar de representar el conocimiento con el operador \min , se realiza con la implicación de Lukasiewicz (Tabla 5.5).
- En lugar de efectuar la agregación con el operador \max se realiza con el operador \min .

La figura 5.3 muestra la relación Entrada-Salida del sistema (en azul) y se compara con la del ejemplo 4.9 (en rojo). Es notoria la semejanza de las dos relaciones. Sin embargo, internamente suceden dos procesos muy diferentes.

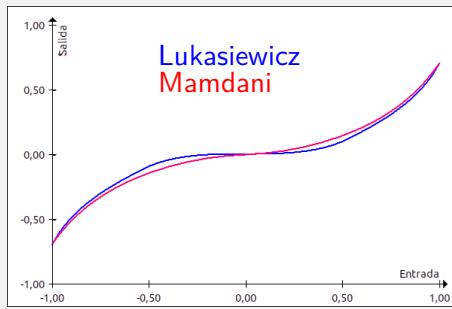


Figura 5.3: Relación Entrada-Salida del ejemplo 5.2

Para visualizar el proceso interno de inferencia y concreción se ha preparado la Figura 5.4, que puede compararse con la Figura 4.22 del ejemplo 4.9. En la construcción de la figura se han considerado 9 entradas diferentes ($x = -1.0, x = -0.75, \dots, x = 0.65, x = 1.0$). Para cada entrada se han trazado 5 gráficas:

- La primera gráfica muestra en azul la variable lingüística de entrada y en rojo la salida del difusor.
- Las siguientes tres gráficas muestran las conclusiones obtenidas en cada una de las tres reglas.
- La quinta gráfica muestra en azul el resultado de agregar las conclusiones de las reglas y en rojo la salida del congresor.

Al analizar la primera columna se observa que no hay ninguna diferencia con el sistema tipo Mamdani, lo cual no es ninguna sorpresa porque ambos sistemas tienen el mismo difusor.

La primera gran diferencia se observa en las columnas 2, 3 y 4. Nótese que las conclusiones ya no son versiones recortadas de los conjuntos de los consecuentes. Es más, ahora, cuando una regla no se activa se obtiene una conclusión que es un conjunto difuso con grado de pertenencia 1. Por ejemplo, cuando $x = -1$ (primera línea) la regla 3 no se activa (Si Alto \Rightarrow Alto) porque el difusor no intersecta con el conjunto Alto; sin embargo, la conclusión de la regla 3 no es vacía, como en el caso Mamdani, sino todo lo contrario: un conjunto cuya función de pertenencia siempre es 1.

Lo anterior sucede porque la función de implicación asigna un grado de verdad alto a cualquier entrada falsa. En otras palabras, como la intersección es 0, la implicación es 1. Esto es coherente con el grado de verdad de las dos primeras líneas de la Tabla 5.2.

Es por eso que la función de agregación no puede ser una unión. Si lo fuera, la conclusión siempre sería un conjunto difuso de grado de pertenencia 1. En cambio, al usar una intersección (el operador \min) se obtiene un conjunto que refleja aquellas conclusiones con alto grado de verdad *en todas las reglas*.

En otras palabras, las reglas que sí se activan ayudan a 'limpiar' la conclusión final. Esto puede verse claramente cuando $x = 0$ (línea 5). Las reglas 1 y 3 producen conjuntos con grado de pertenencia 1, mientras que la regla 2 ($\text{Si Medio} \Rightarrow \text{Medio}$) produce un triángulo que es justamente el conjunto Medio. La intersección de las tres conclusiones se observa en la columna 5, y corresponde exactamente al triángulo de la regla 2, que es la única que se activa.

5.3. Control difuso

Una de las aplicaciones más conocidas de los SBR es en el control de sistemas dinámicos. De hecho, esta aplicación es tan popular que, por antonomasia, algunos autores denominan a los SBR *controladores difusos*. En esta sección se hace una breve introducción al uso de SBR como controladores. Para ello, inicialmente se formula un problema clásico de control, consistente en la estabilización de una planta de una entrada y una salida para explicar la estrategia fundamental del control, que es la retroalimentación. En la sección 5.3.2 se muestra la forma convencional de utilizar un SBR para efectuar la tarea de control.

5.3.1. Rudimentos de control

¿Qué es el control? En términos generales, las técnicas de control se utilizan para lograr que un determinado sistema se comporte *de una forma deseada* en lugar de como lo haría naturalmente. Está presente en prácticamente todos los dispositivos que usamos diariamente, desde los más simples hasta los más sofisticados. Gracias al control logramos que un calentador para el agua tenga la temperatura que ajustamos, que una cinta para caminar avance a la velocidad establecida, que el tanque de un inodoro no se desborde por ejemplo; y también gracias al control logramos que un robot construya autónomamente una pieza, que un satélite mantenga su trayectoria y que un vehículo se parquee sin intervención del conductor.

El sistema que se desea controlar se denomina la *planta*. Para controlar una planta, lo primero que se debe hacer es identificar sus entradas y salidas ¿Cuáles son los estímulos que recibe y cuáles sus respuestas? La figura 5.5 muestra una situación típica: la planta recibe un estímulo $x(t)$ y reacciona con una respuesta $y(t)$. Cuando se aplica un estímulo constante la respuesta de la planta crece sin detenerse.

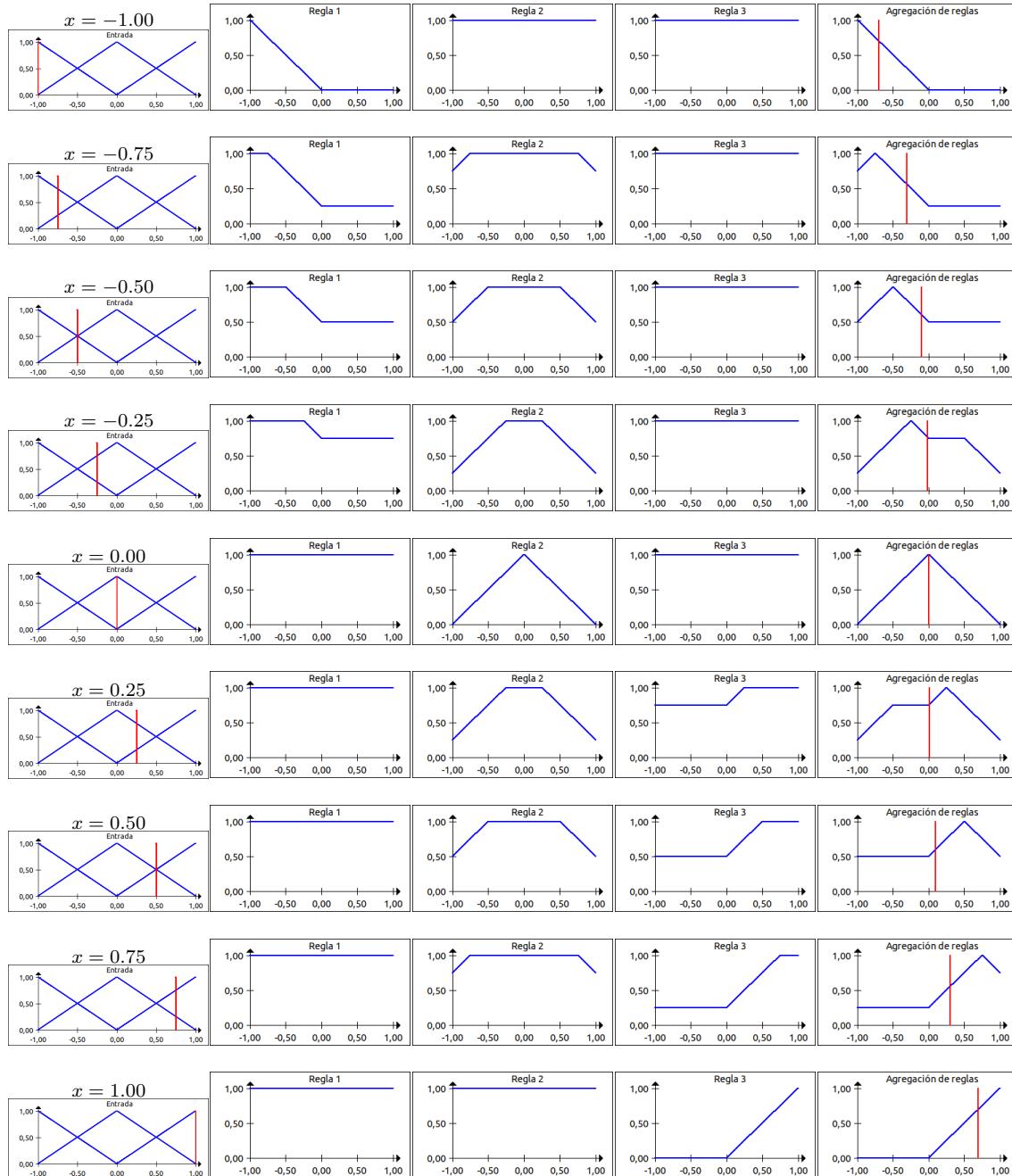


Figura 5.4: Cálculo paso a paso. Implicación tipo Lukasiewicz

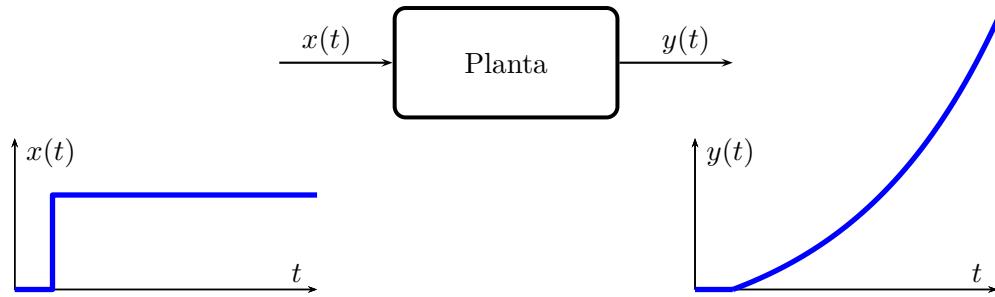


Figura 5.5: Planta a ser controlada

Veamos algunos ejemplos de esa situación: supongamos que la planta es un tanque de agua, $x(t)$ es la posición de una válvula que se abre o se cierra para permitir que entre más o menos agua y $y(t)$ es el nivel del tanque. Al abrir la válvula empezará a entrar agua y el nivel del tanque crecerá sin parar; si no hacemos algo, el tanque se desbordará.

Otro ejemplo: supongamos que la planta es un vehículo eléctrico, $x(t)$ es la posición de una perilla que regula la tensión de alimentación al motor y $y(t)$ es la posición del carro. Al girar la perilla el carro empezará a moverse y por tanto su posición cambiará. Si mantenemos la posición de la perilla el carro se moverá sin detenerse y la posición aumentará.

Los dos ejemplos hipotéticos anteriores son, desde el punto de vista matemático, semejantes: ante una entrada de tipo escalón la salida tiende a infinito. Se trata de dos ejemplos de plantas *inestables*, en el sentido en que ante una entrada que es acotada, la salida no es acotada². Es muy posible que queramos modificar ese comportamiento *natural* de la planta para, por ejemplo, evitar que se desborde el tanque o para lograr que nuestro vehículo llegue a una posición deseada.

La estrategia fundamental para controlar un sistema es la realimentación, que se ilustra en la figura 5.6. La planta ya no está sola. Hemos adicionado otros elementos. El *sensor* es un elemento que mide el valor de la salida $y(t)$. La señal $r(t)$ indica cuál es el *comportamiento deseado* para nuestra planta (cuál es el nivel de agua que queremos en el tanque o a qué posición queremos que llegue nuestro vehículo). Se efectúa una comparación entre el comportamiento deseado y el que hemos medido, restando $r(t) - y_m(t)$ nos permite identificar un *error*, que se representa por $e(t)$.

El *controlador* recibe el error y decide qué estímulo $x(t)$ debe darse a la planta, con

²Existen varias definiciones de estabilidad. Aquí estamos empleando la estabilidad BIBO (*Bounded Input - Bounded Output*).

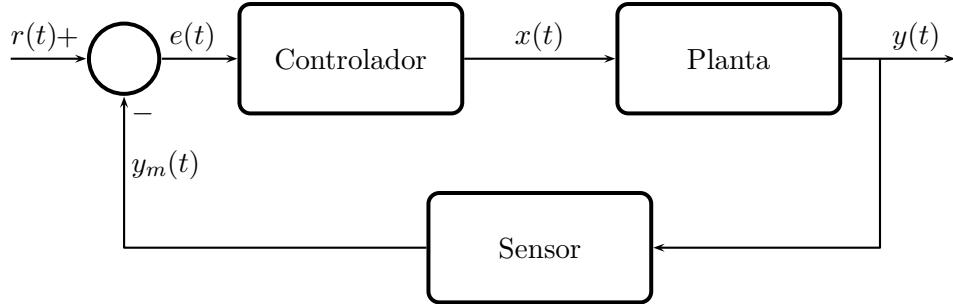


Figura 5.6: Retroalimentación de una planta

el propósito de disminuir ese error. El controlador es el cerebro del sistema que decide cómo debe actuarse para lograr el comportamiento deseado.

5.3.2. Controlador difuso

Un controlador difuso es un dispositivo que cumple la función del controlador de la Figura 5.6 y cuyo funcionamiento se basa en la teoría de conjuntos difusos. Una de las estructuras más usadas se muestra en la Figura 5.7, en la que se utiliza un SBR como elemento de cálculo. El universo de discursos tanto de la entrada como de la salida es el intervalo $[-1, 1]$. Los bloques *Normalizador* y *Denormalizador* son cambios de escala que permiten conectar el SBR en aplicaciones disímiles.

En ocasiones, la naturaleza de la planta o el tipo de problema a resolver obligan a sofisticar el controlador. Por ejemplo, es posible que además de utilizar el valor del error para definir la acción de control, sea útil utilizar la derivada del error o su derivada. Un controlador PID es un controlador que calcula la acción del error como una combinación lineal Proporcional al error, a su Integral y a su Derivada.

$$u_{PID}(t) = K_P e(t) + K_I \int e(t) dt + K_D \frac{de(t)}{dt} \quad (5.12)$$

Las constantes K_P , K_I y K_D se ajustan para obtener el efecto deseado. Las versiones P , PI y PD de esta estrategia de control, utilizan el error, el error y su integral ó el error y su derivada.

Es posible extender esta estrategia de control a los controladores difusos. La Figura 5.8 muestra las entradas que tendrían los controladores difusos tipo P , PI , PD y PID . El controlador tipo P necesita un SBR SISO, mientras que los otros tres necesitan controladores MISO de 2 o tres entradas.

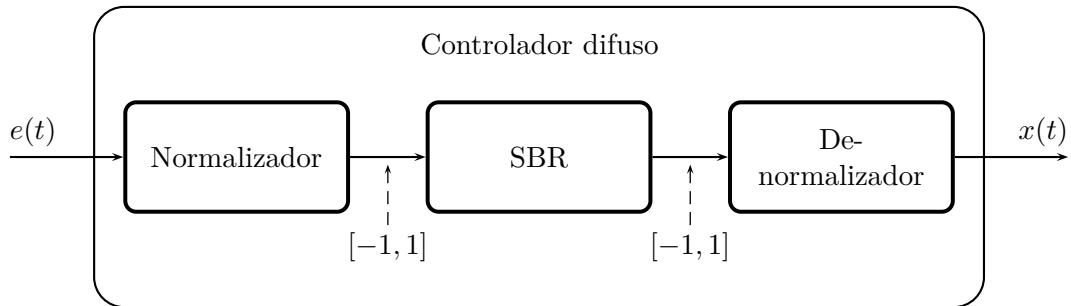


Figura 5.7: Controlador difuso

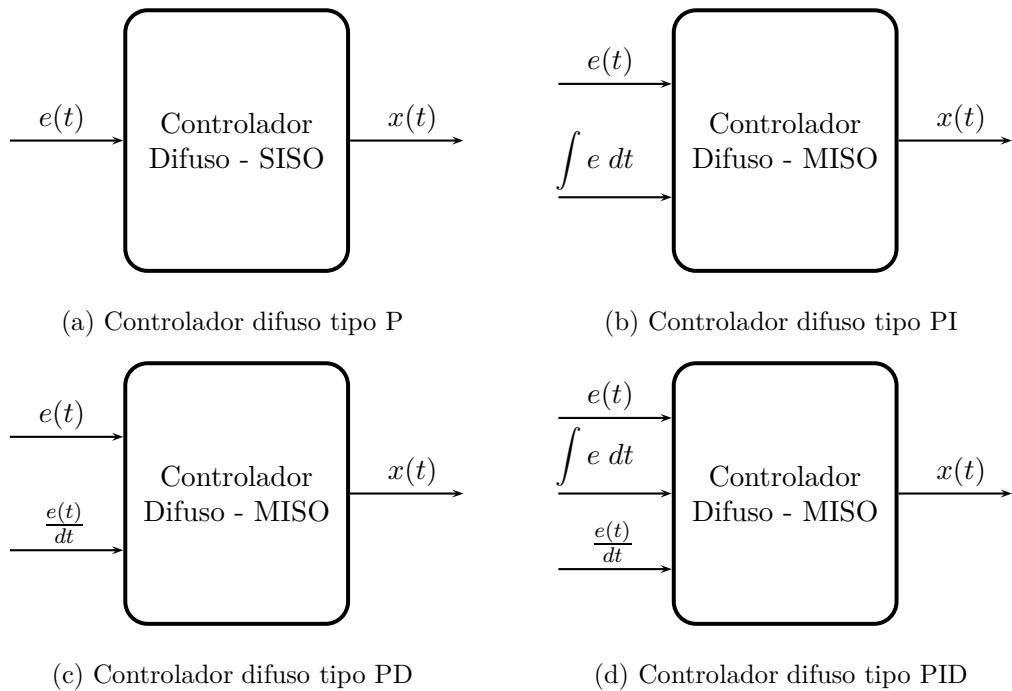


Figura 5.8: Estructuras usuales de controladores difusos

Ejemplo 5.3 (Controlador difuso) Supongamos que la planta tiene un comportamiento quer se describe por la ecuación diferencial

$$\frac{d^2y(t)}{dt^2} + \frac{dy(t)}{dt} = x(t) \quad (5.13)$$

Podemos utilizar la Transformada de Laplace para analizar la ecuación: 1

$$s^2\mathbf{Y}(s) + s\mathbf{Y}(s) = \mathbf{X}(s) \quad (5.14)$$

Y obtener la función de transferencia de la planta:

$$\mathbf{F}(s) = \frac{\mathbf{Y}(s)}{\mathbf{X}(s)} = \frac{1}{s^2 + s} \quad (5.15)$$

Utilizando un software de simulación (scilab, sección C.4) podemos evaluar cuál sería la salida de la planta ante una entrada tipo escalón. La Figura 5.9 muestra el esquema de simulación y el comportamiento de las variables $x(t)$, $y(t)$. La tendencia creciente de $y(t)$ demuestra que la planta es inestable.

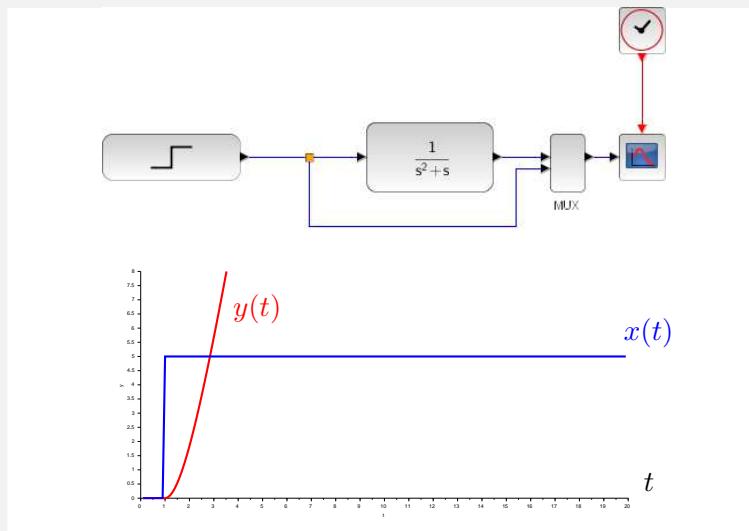


Figura 5.9: Planta y simulación para el ejemplo 5.3

Con el propósito de estabilizar la planta se propone el uso de un controlador difuso. Se decide probar un controlador tipo P . Las variables de entrada y de salida se definen con tres conjuntos cortos rectos y la base de reglas consta de tres reglas que se consignan en la Tabla 5.6.

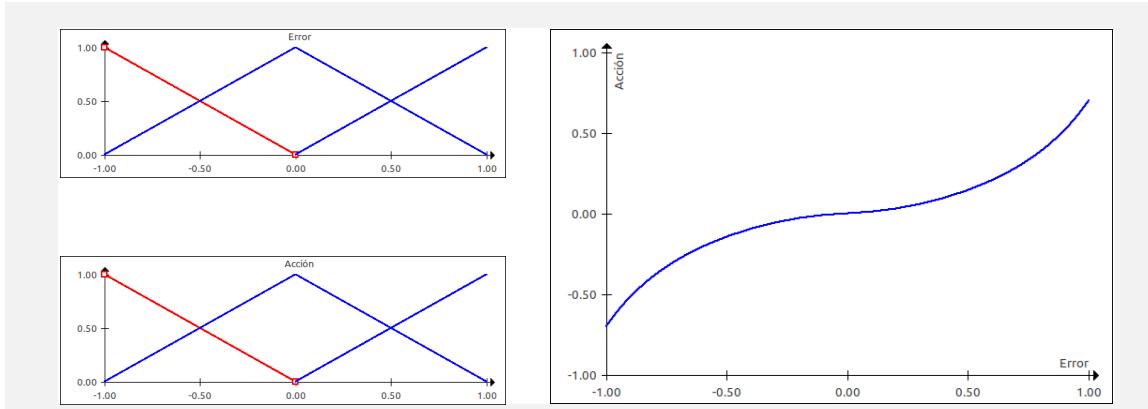


Figura 5.10: Controlador Difuso del ejemplo 5.3

Tabla 5.6: Base de reglas del ejemplo 5.3

Error	Acción
Negativo	Baja
Cero	Media
Positivo	Alta

Haciendo uso del mismo software de simulación se construye el sistema retroalimentado y se simula la respuesta ante una entrada escalón. La Figura 5.11 muestra el esquema empleado y el resultado de la simulación. Nótese que la variable y ahora no crece indefinidamente, sino que su valor se estabiliza.

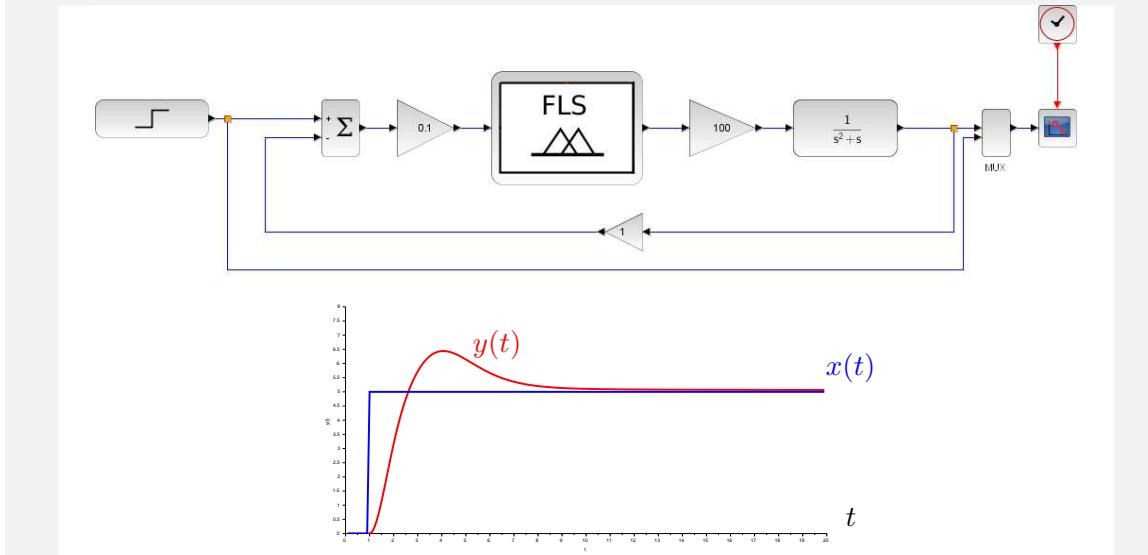


Figura 5.11: Planta y simulación para el ejemplo 5.3

5.3.3. Unas anotaciones sobre el control difuso

La teoría de control de sistemas dinámicos es amplia y profunda. El control difuso es solo una de muchas estrategias disponibles. Para abordar una tarea de control es necesario un bagaje sobre análisis de sistemas dinámicos que excede en mucho el alcance de este libro.

Los controladores difusos tienen una ventaja comparativa con otras técnicas de control: permiten el diseño a partir de una descripción lingüística. Sin embargo, esa ventaja es la raíz de su principal desventaja: debido a esa descripción lingüística es fácil utilizar todas las herramientas que brinda la teoría de control para analizar las propiedades del sistema realimentado.

Por ejemplo, no es para nada sencillo determinar la estabilidad de un sistema realimentado con un controlador difuso. No es posible construir una función de Liapunov, por ejemplo. En consecuencia, no hay herramientas para valorar la robustez del sistema realimentado. Y eso es grave.

La excepción se da cuando las plantas tienen un ‘buen comportamiento’, como en el caso de las plantas pasivas. En tales casos es la realimentación, más que el controlador mismo, quien aporta la solución.

Las primeras aplicaciones de los sistemas difusos se conocieron en el área de control. De hecho, el nacimiento de la lógica difusa estuvo inspirada en problemas de control. El control difuso le dio un impulso muy grande a los sistemas difusos, primero en la década de los 70s y luego en la de los 90s y a comienzos del siglo XXI. Sin embargo, en la actualidad la aplicabilidad del control difuso está revaluada a la baja.

Por otra parte, existen otros tipos de sistemas basados en reglas diferentes a los que aquí se han considerado, que también pueden usarse como controladores difusos. Entre estos controladores destacan los tipo Takagi-Sugeno a los que se dedica la sección 5.4.

5.4. Sistemas tipo Takagi-Sugeno

Los sistemas tipo Takagi-Sugeno (TS) también son Sistemas Basados en Reglas, pero con un tipo diferente de reglas. Específicamente, el consecuente de las reglas es diferente. En lugar de ser expresiones de la forma ‘ y es $\mathcal{Y}^{(k)}$ ’, el consecuente es una función matemática de las entradas.

A manera de ejemplo, supóngamos un sistema TS de dos entradas x_1 y x_2 y una salida y . Supongamos también que cada una de las entradas tiene una variable lingüística

asociada con las etiquetas Bajo, Medio y Alto. En esas condiciones, la base de reglas del sistema será de la forma:

- Regla 1: **Si** x_1 es Bajo **Y** x_2 es Bajo **entonces** $y = f_1(x_1, x_2)$
- Regla 2: **Si** x_1 es Bajo **Y** x_2 es Medio **entonces** $y = f_2(x_1, x_2)$
- ⋮
- Regla 9: **Si** x_1 es Alto **Y** x_2 es Alto **entonces** $y = f_9(x_1, x_2)$

Ese tipo de reglas no pueden interpretarse con una máquina de inferencia como la presentada en la sección 4.3 y por tanto tampoco puede emplearse un congresor como los de la sección 4.5. En su lugar, la salida del sistema se calcula como el promedio ponderado de los consecuentes de todas las reglas. La ponderación se hace con la función de pertenencia a los antecedentes.

Es decir, que para un sistema de p entradas y n reglas, la salida y del sistema está dada por:

$$y = \frac{\sum_{k=1}^p \mu_D(\mathbf{x}) f_k(\mathbf{x})}{\sum_{k=1}^p \mu_D(\mathbf{x})} \quad (5.16)$$

en donde $\mu_D(\mathbf{x})$ es la salida del difusor, que se calcula de la misma forma que para un sistema tipo Mamdani (ecuación 4.27).

En principio, las funciones $f_k(\mathbf{x})$ que van en los consecuentes pueden ser de cualquier tipo. Sin embargo, lo usual es utilizar constantes, combinaciones lineales, funciones afines y polinomios. La simplicidad de estas funciones, combinada con la versatilidad que brinda una base de reglas, hace que los sistemas TS sean muy atractivos en el diseño de sistemas de control.

Ejemplo 5.4 (Sistema Takagi-Sugeno de 1 entrada) c Retomemos los ejemplos 4.2 y 4.4. Podemos formular un sistema Takagi Sugeno para modelar la relación entre la entrada 'Calor' y la salida 'Riesgo'.

La base de reglas que se propone se muestra en la Tabla 5.7. En la figura 5.12.a. se han trazado las tres rectas que corresponden a cada una de las reglas, destacando en rojo la tercera de ellas.

Se ha trazado también la figura 5.12.b. en la que se muestra la curva de la relación entre entrada y salida. Esa curva resulta ser una interpolación de las tres rectas. Los pesos del promedio ponderado de la ecuación 5.16 dependen del valor de la entrada y por tanto, como

se observa en esa figura, la curva se acerca a la tercera recta conforme la entrada es cada vez mayor, es decir, conforme la entrada tiene un grado de pertenencia mayor a la etiqueta 'Mucho'.

Tabla 5.7: Base de reglas del ejemplo 5.4

Calor	Riesgo
Poco	$y = m_1x + b_1$
Medio	$y = m_2x + b_2$
Mucho	$y = m_3x + b_3$

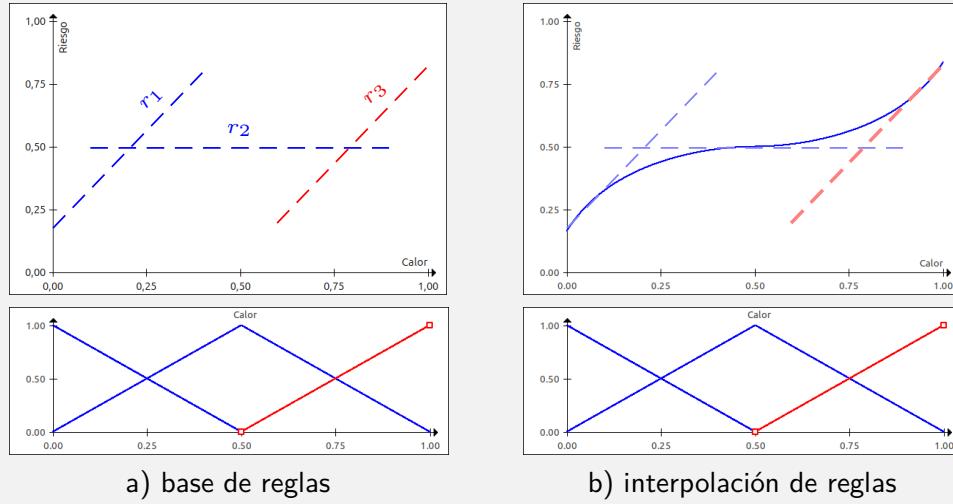


Figura 5.12: Inferencia en el ejemplo 5.4

Capítulo 6

Aprendizaje de máquina

6.1. Introducción

El término ‘aprendizaje de máquina’ comenzó a utilizarse en la década de los 50s y desde ese entonces su significado ha venido adaptándose a los desarrollos tecnológicos. Hoy en día, es usual definir el ‘aprendizaje de máquina’ como una de las ramas de la inteligencia artificial que se especializa en el desarrollo de programas de cómputo que puedan aprender a realizar ciertas tareas sin necesidad de que hayan sido explícitamente programados para ello.

La figura 6.1a puede ayudar a aclarar esta definición. Allí se observa un ‘Sistema de cómputo’ que no es otra cosa que un *software* que procesa una información de entrada para producir una información de salida. El esquema básico del aprendizaje de máquina implica que hay otro *software*, que en la figura se denomina ‘Sistema de diseño u optimización’, que ayuda a mejorar el desempeño del sistema principal. Para hacerlo, este segundo software utiliza como fuente de información las entradas y salidas del sistema principal, así como información proveniente del entorno.

Lo anterior nos permite formular la siguiente definición genérica del aprendizaje de máquina:

Algoritmos que ayudan a que otros algoritmos aprendan (es decir, que mejoren su desempeño) a partir del análisis de datos relacionados con el entorno y con su propio desempeño.

La definición anterior abarca muchas estrategias de aprendizaje. Algunas de ellas son específicas para un cierto tipo de problemas y otras son más genéricas. Estas estrategias suelen clasificarse en tres grandes tipos de aprendizaje, según la información disponible para que la máquina aprenda:

Aprendizaje supervisado: la característica principal del aprendizaje supervisado consiste en que existe información sobre cómo debería desempeñarse el sistema principal, al menos en ciertas ocasiones. Dicho de otra forma, existe una colección de ejemplos que muestran cómo deben ser las salidas ante unas ciertas entradas. El aprendizaje consiste, entonces, en lograr que la máquina reproduzca esos comportamientos.

La figura 6.1b ilustra esta situación: el sistema de diseño recibe las entradas y salidas directamente del sistema principal y del entorno recibe la información de cómo debería ser la salida. El sistema puede comparar la salida deseada con la salida obtenida y deducir qué cambios deben hacerse en el diseño del sistema para que esas dos salidas sean lo más parecidas que se pueda.

Aprendizaje no supervisado: en el aprendizaje no supervisado no existe la colección de ejemplos y por tanto no es posible comparar la salida del sistema con una salida deseada. El aprendizaje consiste en encontrar patrones de semejanza entre los datos disponibles.

La figura 6.1c ilustra esta situación: el sistema de diseño recibe las entradas directamente del sistema principal. A partir de esa información debe decidir qué ajustes realizar en el diseño del sistema principal.

Aprendizaje por refuerzo: en el aprendizaje por refuerzo tampoco existe una colección de ejemplos de entradas y salidas. Pero sí se tiene acceso a una información que proviene del entorno y que se refiere a qué tan bien (o mal) desempeñó su tarea el sistema principal.

La figura 6.1d ilustra esta situación: a partir de la medición de las entradas y de la información sobre el desempeño del sistema principal, el sistema de decide qué aspectos del diseño del sistema principal deben modificarse.

La Figura 6.2 muestra la estructura de un Sistema Basado en Reglas. El diseño de uno de esos sistemas implica:

- El diseño de las variables lingüísticas, que a la vez implica, para cada variable, la selección de un nombre, el diseño del universo de discurso, la selección de los términos lingüísticos y la construcción de la función semántica:
- El diseño de la base de reglas.
- La selección y ajuste de los siguientes operadores matemáticos: T_{AND} , \mathcal{I} , \mathcal{S}_C , \mathcal{T}_C , \mathcal{A}
- La selección del tipo de difusor y sus parámetros asociados.
- La selección del tipo de congresor.

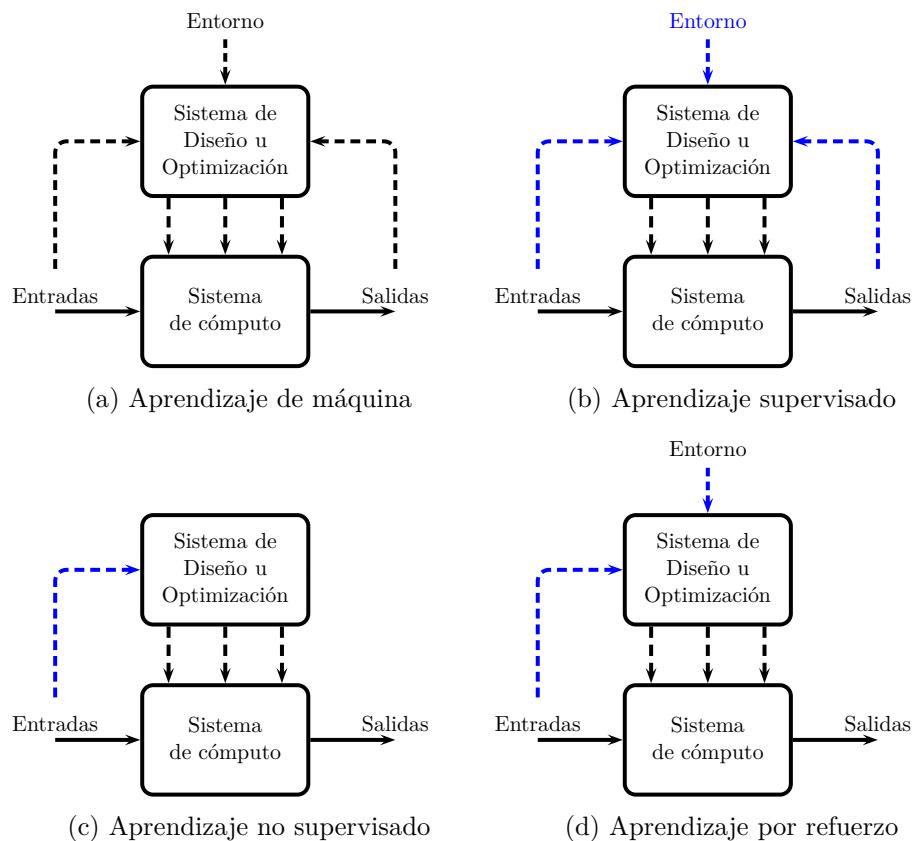


Figura 6.1: Esquemas del aprendizaje de máquina

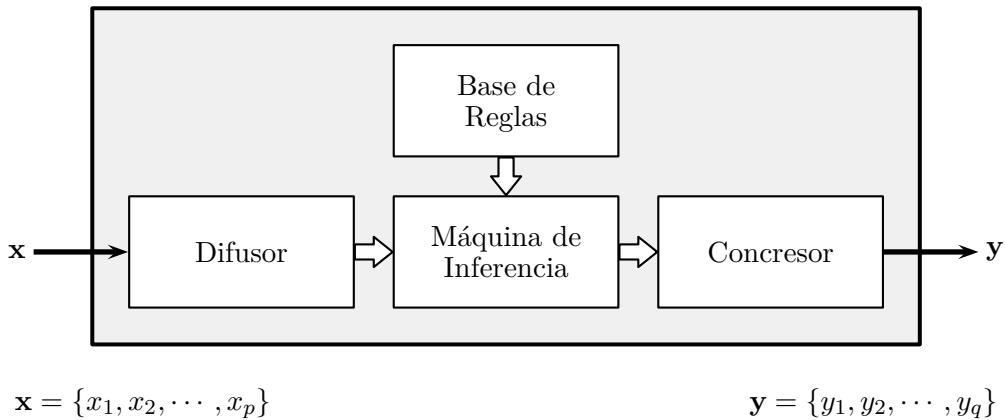


Figura 6.2: Sistema basado en reglas

En este texto se muestran algunas estrategias de aprendizaje de máquina aplicadas al diseño de sistemas difusos basados en reglas. Es decir, se presentarán algunos algoritmos que pueden emplearse para diseñar u optimizar el diseño de algunos de los componentes de un Sistema basado en reglas. Es una presentación somera cuyo propósito principal es entregar el siguiente mensaje: los sistemas difusos son susceptibles de ser diseñados y optimizados mediante aprendizaje de máquina.

En la sección 6.2 se explica qué es la aproximación de funciones y cómo realizarse mediante SBR. En la sección 6.3 se presenta una estrategia para diseñar la función semántica de una variable lingüística a partir de información proveniente de una encuesta. Posteriormente, en la sección 6.4, se presenta una estrategia muy popular para construir conjuntos difusos a partir de información numérica. La sección 6.5 muestra dos estrategias muy prácticas para encontrar las reglas de un SBR, mientras que la sección 6.6 ilustra el uso de heurísticas para la optimización del diseño de SBR,

6.2. Aproximación de funciones con SBR

El paradigma del aprendizaje supervisado (Figura 6.1b) busca automatizar el diseño de un algoritmo para que imite el desempeño de un sistema conocido. Si pensamos en aplicar ese paradigma al diseño de un SBR, entonces, es porque necesitamos que nuestro SBR imite el desempeño de un sistema conocido.

¿Por qué podemos estar interesados en diseñar un SBR que imite el desempeño de

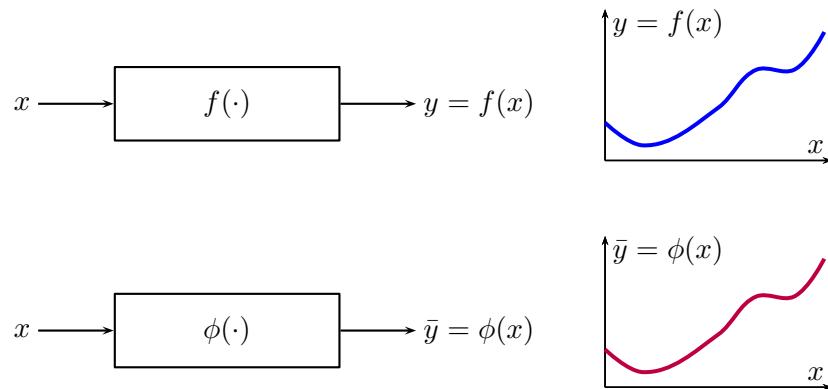


Figura 6.3: Aproximador de una función

un sistema conocido? ¿Por qué no usar directamente el sistema conocido? Podemos imaginar, al menos, dos situaciones que explican nuestro interés:

- Es posible que no tengamos acceso al sistema conocido, sino únicamente a la información sobre cómo se desempeña. El SBR que queremos construir nos puede ayudar a simular el comportamiento del sistema conocido.
- Es posible que tengamos a la mano un modelo muy complicado del sistema conocido y que necesitemos uno más sencillo.

Antes de presentar cómo podríamos diseñar el SBR imitador, en esta sección abordamos una pregunta apremiante: *¿Tiene sentido pretender que un SBR sea capaz de imitar el desempeño de cualquier sistema?* o, dicho de otra manera *¿podemos estar seguros de que existe algún SBR como el que necesitamos o estaremos buscando algo inexistente?*

Para dar responder esa pregunta vamos a presentar la noción de un *aproximador universal de funciones* y luego vamos a explorar si los SBR son o no, aproximadores universales de funciones. Espóiler: los SBR que nos interesan en este documento sí lo son.

6.2.1. Aproximadores universales de funciones

Un *aproximador de una función* es otra función que tiene un comportamiento muy parecido a la primera. La Figura 6.3 ilustra la situación: $f(\cdot)$ es una cierta función que cuando recibe la entrada x produce la salida $y = f(x)$; por su parte, $\phi(\cdot)$ es otra función diferente que cuando recibe la misma entrada x produce la salida $\bar{y} = \phi(x)$. En esas condiciones, se dice que $\phi(\cdot)$ es un *aproximador* de $f(\cdot)$ si las dos salidas son parecidas, es decir si $|y - \bar{y}|$ es pequeño.

Para hablar de *aproximadores universales de funciones*, debemos introducir antes el concepto de *familia de funciones*, que no es más que un conjunto de funciones que

comparten la misma estrategia con la que calculan las salidas, pero se diferencian en los parámetros específicos de cada miembro. Usaremos la notación $\Phi(\mathbf{x}; \theta) : \mathbb{R}^d \rightarrow \mathbb{R}$ para referirnos a una familia de funciones. En esa notación, $\mathbf{x} \in \mathbb{R}^d$ es el vector de entradas de las funciones y $\theta \in \mathbb{R}^q$ es un vector que contiene sus parámetros.

Por ejemplo, para la familia de las rectas $y = mx + b$ se tiene que:

$$y = \Phi(x, \theta) = mx + b \quad x \in \mathbb{R} \quad \theta = \begin{bmatrix} m \\ b \end{bmatrix} \quad (6.1)$$

Todas las rectas usan la misma estrategia general para calcular la salida: multiplicar por m la entrada y sumarle b . Lo que hace que una recta sea diferente de la otra es el valor de sus parámetros (m, b) .

En términos generales, una familia de funciones $\Phi(\mathbf{x}; \theta) : \mathbb{R}^d \rightarrow \mathbb{R}$ es un *aproximador universal de funciones* si es lo suficientemente versátil para aproximar una función arbitraria.

Podemos precisar esta noción en términos matemáticos mediante la siguiente definición:

Una familia de funciones $\Phi(\mathbf{x}; \theta) : \mathbb{R}^d \rightarrow \mathbb{R}$ es un *aproximador universal* del conjunto de funciones $\mathcal{F}(\mathbf{x}) : \mathbb{R}^d \rightarrow \mathbb{R}$ si para cualquier subconjunto cerrado y acotado C en \mathbb{R}^d , cualquier función $f(\mathbf{x}) \in \mathcal{F}$ y cualquier $\varepsilon > 0$, existe al menos un miembro ϕ de la familia tal que

$$\sup_C |f(\mathbf{x}) - \phi(\mathbf{x})| < \varepsilon \quad (6.2)$$

La definición anterior es suficiente para los propósitos de este documento. Sin embargo, debemos anotar que hay definiciones más sofisticadas que son válidas para sistemas de múltiples salidas y para espacios compactos¹.

Ejemplo 6.1 (Teorema de Weierstrass) Karl Weierstrass demostró en 1885 que la familia de polinomios en x es un *aproximador universal* de las funciones $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ continuas. El teorema puede enunciarse así

Supóngase que f es una función real continua definida sobre el intervalo real $[a, b]$. Para todo $\varepsilon > 0$, existe un polinomio p tal que

$$|f(x) - p(x)| < \varepsilon \quad \forall x \in [a, b]$$

¹El término ‘espacio compacto’ hace referencia a un concepto desarrollado en la topología matemática, que busca generalizar la idea de conjunto finito

6.2.2. SBR como aproximadores universales de funciones

Al inicio de la sección 6.1 formulamos la pregunta: ¿Tiene sentido pretender que un SBR sea capaz de imitar el desempeño de cualquier sistema?. Tenemos ahora los elementos necesarios para reformularla en términos más precisos: *¿Es la familia de SBR un aproximador universal de algún tipo de funciones?*

A lo largo de la historia se han encontrado varias respuestas parciales a esa pregunta. Las siguientes son las más relevantes:

- En 1992 Buckley demostró en [9] que una familia de SBR parecida a los sistemas Takagi-Sugeno (sección 5.4) es un aproximador universal de funciones continuas. La familia se caracteriza por tener polinomios en los consecuentes de cada regla y usar como concretor una expresión diferente a la ecuación 5.16.
- También en 1992, Wang demostró en [51] que una familia de SBR tipo Mamdani es un aproximador universal de funciones continuas. Esa familia se caracteriza por tener variables lingüísticas con conjuntos gausianos, y usar como agregador el producto. Esta última característica es muy curiosa, porque lo usual es que el agregador sea una \mathcal{S} -norma y el producto es una \mathcal{S} -norma.
- En 1995 Castro [12] demostró que una familia de SBR tipo Mamdani es un aproximador universal de funciones continuas. En 1996, junto a Delgado extendieron ese resultado en [13] a una familia SBR tipo Takagi Sugeno. Ambas familias son muy amplias, es decir, cubren una gran cantidad de SBR. Para formar parte de esa familia, un SBR debe cumplir con las siguientes características:
 1. Los soportes de los conjuntos difusos de las variables lingüísticas deben ser intervalos.
 2. La agregación debe hacerse con una \mathcal{S} -norma.
 3. La implicación debe hacerse con una \mathcal{T} -norma o con una implicación tal que $I(x, y) = 0$ implica que tanto x como y son 0.
 4. Si el sistema es de tipo Mamdani, el concretor debe ser tal que su salida siempre esté incluida en el soporte del conjunto difuso que recibe.
 5. Si el sistema es de tipo Sugeno, las funciones de los consecuentes deben entregar siempre un valor numérico.

La importancia del resultado de Castro es muy grande. Demuestra que ‘prácticamente’ todos los SBR son aproximadores universales de funciones continuas. De todos los tipos de SBR presentados en este documento, por ejemplo, los únicos que no entran en esa familia son los sistemas que utilizan inferencias de implicación (sección 5.2.3) porque esos hacen la agregación con una \mathcal{T} -norma, lo que viola la segunda condición.

La conclusión es clara: si se quiere usar aprendizaje supervisado para diseñar un SBR, éste no debe usar inferencias de implicación, porque ni siquiera tenemos certeza de que exista el SBR que estamos buscando.

Ahora bien, lo que hemos encontrado es que si queremos diseñar un SBR que imite el desempeño de algún sistema conocido, tiene sentido intentarlo con un sistema tipo Mamdani o tipo Takagi-Sugeno convencionales, porque tenemos certeza de que al menos existe uno que satisface nuestros propósitos. Pero ¿cómo encontrarlo? es allí donde entran las estrategias de aprendizaje de máquina. En las siguientes secciones presentamos algunas alternativas para abordar esa búsqueda.

6.3. Construcción de la función semántica a partir de encuestas

La función semántica es uno de los elementos que definen una variable lingüística (Sección 3.1, ecuación 3.1). En términos prácticos, construir la función semántica equivale a construir los conjuntos difusos que acompañan a cada uno de los términos de la variable lingüística (ejemplo 3.1).

En esta sección presentamos una estrategia para construir conjuntos difusos a partir de encuestas de opinión realizadas a un grupo de personas. Para hacerlo, interpretamos la función de pertenencia del conjunto asociado a término \mathcal{E} así:

La función de pertenencia $\mu_{\mathcal{E}}(x)$ representa la probabilidad de que una persona asigne la etiqueta \mathcal{E} al valor x

A partir de una encuesta podemos estimar la probabilidad de que una persona de un determinado grupo asigne la etiqueta \mathcal{E} al valor x y a partir de allí construir la función de pertenencia. Adicionalmente, es factible simplificar la función de pertenencia mediante un ajuste de línea recta para obtener conjuntos trapezoidales. Esta estrategia se ilustra en el Ejemplo 6.2.

Ejemplo 6.2 (Construcción de un conjunto difusa a partir de encuestas)

Tomemos como ejemplo la variable lingüística ‘edad’. Supongamos que seleccionamos 5 términos lingüísticos: ‘niño’, ‘adolescente’, ‘joven’, ‘adulto’ y ‘adulto mayor’. Formalmente, definimos la variable lingüística \mathcal{V} (ecuación 3.1) así:

$$\begin{aligned} \mathcal{V} &= \langle \mathcal{X}, \mathcal{L}_{\mathcal{X}}, \mathcal{U}, \mathcal{M}_{\mathcal{U}} \rangle \\ \mathcal{X} &= \text{‘Edad’} \\ \mathcal{L}_{\mathcal{X}} &= \{ \text{‘Niño’}, \text{‘Adolescente’}, \text{‘Joven’}, \text{‘Adulto’}, \text{‘Adulto mayor’} \} \\ \mathcal{U} &= [0, 100] \end{aligned} \tag{6.3}$$

En la definición anterior queda pendiente la construcción de la función semántica $\mathcal{M}_{\mathcal{U}}$. Es decir, aún falta construir un conjunto difuso para cada uno de los cinco términos lingüísticos.

Comencemos por el término ‘Adulto mayor’. Se decide aplicar una encuesta a un grupo de 10 personas con una única pregunta: *¿cuál es la edad (en años) a la que se pasa de ‘Adulto’ a ‘Adulto mayor’?* y se obtienen las respuestas que aparecen en la tabla 6.1

Tabla 6.1: Respuestas en el ejemplo 6.2

70	62	67	66	65	67	69	67	64	74
----	----	----	----	----	----	----	----	----	----

A partir de esas respuestas se construye la Tabla 6.2 en la que x es la edad, $n(x)$ es el número de respuestas que dicen que el cambio de adulto a adulto mayor se da a una edad menor a x y $p(x)$ es la fracción de las respuestas que dicen que el cambio de adulto a adulto mayor se da a una edad menor a x . Como el total de respuestas es 10, se tiene que

$$p(x) = \frac{n(x)}{10} \quad (6.4)$$

Tabla 6.2: Probabilidad acumulada en el ejemplo 6.2

x	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
$n(x)$	0	0	1	1	2	3	4	7	7	8	9	9	9	9	10
$p(x)$	0.0	0.0	0.1	0.1	0.2	0.3	0.4	0.7	0.7	0.8	0.9	0.9	0.9	0.9	1.0

La Figura 6.4 muestra el comportamiento de $p(x)$. Por la forma en que fue construida, $p(x)$ es la estimación de la probabilidad acumulada de que una persona asigne la etiqueta ‘Adulto mayor’ a alguien que tenga x años de edad.

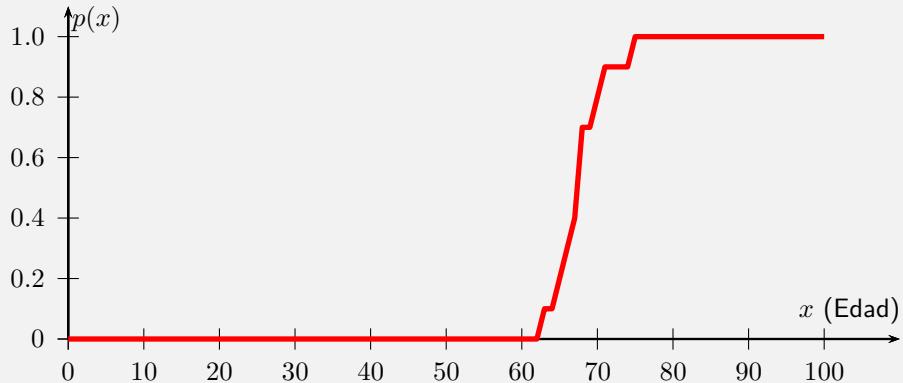


Figura 6.4: Probabilidad acumulada en el ejemplo 6.2

La interpretación que hemos hecho de la función de pertenencia nos permite utilizar $p(x)$ como la curva $L(\alpha)$ del conjunto difuso asociado a la etiqueta ‘Adulto mayor’. Sin embargo, quizás la forma de la curva no nos satisface y busquemos una función más simple.

Por ejemplo, si queremos que la función semántica este formada por conjuntos trapezoidales, como en la Figura 3.6a, podemos buscar la mejor aproximación con alguno de los criterios que se explican en la sección 7.7. En particular, el conjunto de criterios 4 (sección 7.7.5) resulta ser muy práctico.

En efecto, a partir de los datos de la tabla 6.2 podemos encontrar la tendencia de la curva $L(\alpha)$ del conjunto. Para ello efectuamos un ajuste de línea recta con aquellas parejas $(p(x), x)$ para las cuales $0 < p(x) < 1$. La ecuación que se obtiene es:

$$x = m_L \times p(x) + b_L \quad m_L = 10.4 \quad b_L = 62.5 \quad (6.5)$$

Utilizando la ecuación 7.129 se obtienen los vértices a y b del trapecio:

$$\begin{aligned} a &= b_L &= 62.50 \\ b &= m_L + b_L &= 72.94 \end{aligned} \quad (6.6)$$

Debido a que la etiqueta ‘Adulto mayor’ es la que está más a la derecha de todas, el conjunto difuso asociado será del tipo Γ y, por tanto, los vértices c y d son el límite superior del universo de discurso, es decir $c = d = 100$. En conclusión, el conjunto difuso trapezoidal que debemos asociar al término ‘Adulto mayor’ resulta ser $T(62.50, 72.94, 100, 100)$, que se muestra en la figura 6.5

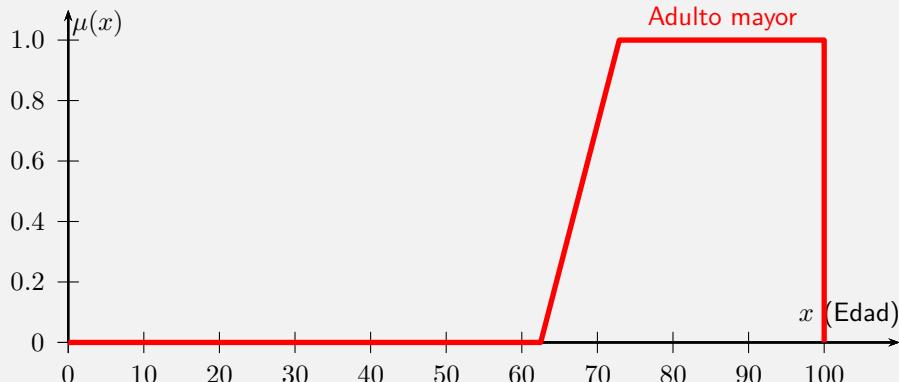
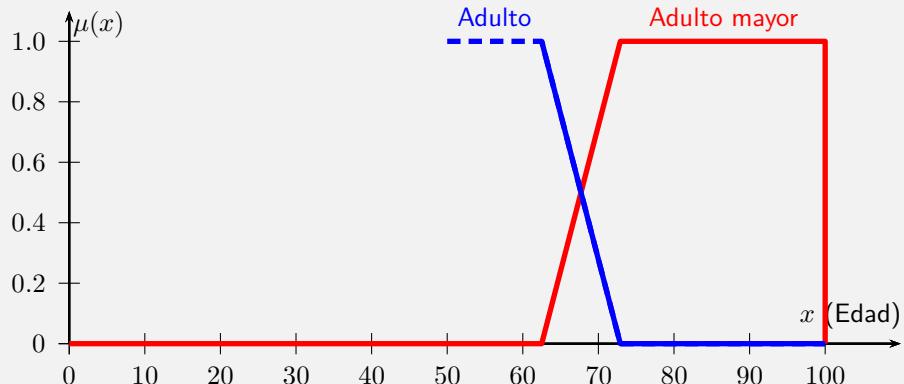


Figura 6.5: Conjunto difuso asociado al término ‘Adulto mayor’ en el ejemplo 6.2

A partir de la curva $L(\alpha)$ del conjunto asociado al término ‘Adulto mayor’, es posible obtener la curva $R(\alpha)$ del conjunto asociado al término ‘Adulto’. Para ello, nótese que una función semántica como la de la Figura 3.6a es una partición difusa (ecuación 3.13). Eso significa que la transición que realiza la función de transferencia del ‘Adulto mayor’ desde 62.5 hasta 72.9 debe complementarse con la transición del ‘Adulto’.

Es decir, el conjunto difuso asociado al término ‘Adulto’ debe ser $T(a, b, 62.50, 72.94)$, como se muestra en la Figura 6.6

Figura 6.6: Curva $R(\alpha)$ del término ‘Adulto’ en el ejemplo 6.2

Si se realizan otras encuestas para preguntar las edades de transición entre cada par de términos lingüísticos consecutivos, se puede construir una función semántica completa semejante a la que se muestra en la Figura 6.7

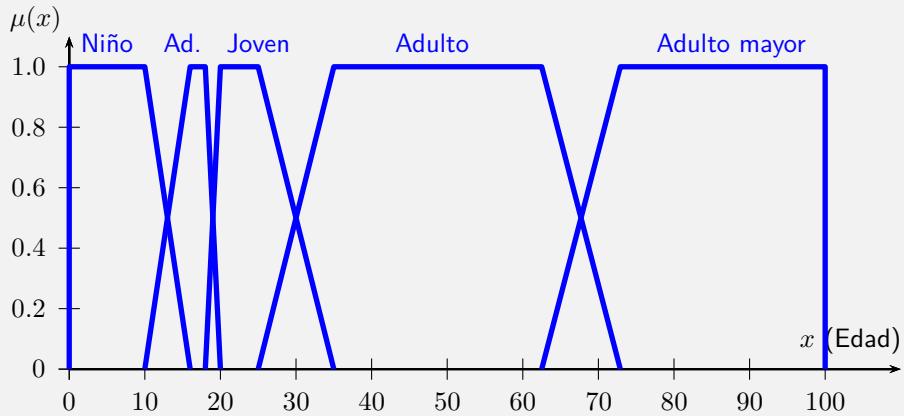


Figura 6.7: Función semántica del ejemplo 6.2

Por supuesto, también es posible construir una versión de la función semántica con conjuntos suavizados (Figura 3.6b). Para ello podemos emplear los mismos vértices y aplicar las funciones de la Figura 3.4 con lo que obtendríamos la función semántica que se muestra en la Figura 6.8

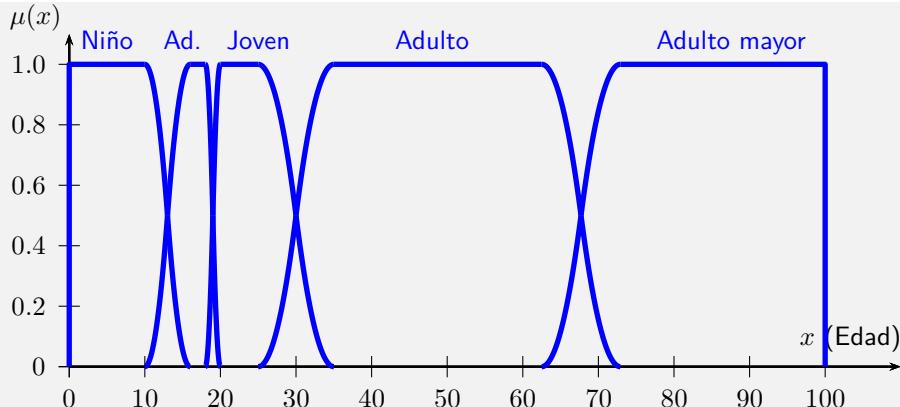


Figura 6.8: Función semántica curva del ejemplo 6.2

Ejemplo 6.3 (Variación sobre el ejemplo 6.2) El elemento más importante del ejemplo 6.2 es el ajuste de línea recta (ecuación 6.5). A partir de los datos de la Tabla 6.2 se tomas las parejas $(p(x), x)$ para las cuales $0 < p(x) < 1$ y a partir de allí se efectúa el ajuste.

En este ejemplo hacemos una variación sobre ese procedimiento. En lugar de tomar las parejas $(p(x), x)$, tomamos las parejas $(x, p(x))$ para efectuar el ajuste. El resultado del ajuste es ahora:

$$p(x) = m \times x + b \quad m = 0.0877622 \quad b = -5.4367133 \quad (6.7)$$

Al interpretar que la función de pertenencia es $p(x)$, obtenemos

$$\mu(x) = 0.0877622 \times x - 5.4367133 \quad (6.8)$$

Para un trapecio, los vértices a y b satisfacen $\mu(a) = 0$ y $\mu(b) = 1$. Por lo tanto

$$\begin{aligned} 0 &= 0.0877622 \times a - 5.4367133 & 1 &= 0.0877622 \times b - 5.4367133 \\ a &= \frac{5.4367133}{0.0877622} & 1 &= \frac{1 + 5.4367133}{0.0877622} \\ a &= 61.95 & b &= 73.34 \end{aligned} \quad (6.9)$$

Los vértices así obtenidos son cercanos, pero diferentes, a los obtenidos en el ejemplo 6.2 ($a = 62.50$ $b = 72.94$ ecuación 6.6). Esto se debe a que para un conjunto de parejas de datos

(x, y) el ajuste de línea recta convencional minimiza el error en la predicción de la variable y y este resultado es distinto al que se obtendría al minimizar el error en la predicción de la variable x .

¿Cuál de las dos opciones es preferible, tomar las parejas $(p(x), x)$ o las parejas $(x, p(x))$? No hay duda. Deben tomarse las parejas $(p(x), x)$ porque corresponden al criterio 4 de aproximación trapezoidal, y por tanto minimiza la distancia entre el conjunto original y el trapecio obtenido.

El ejemplo 6.2 muestra que es posible usar los datos provenientes de una encuesta para construir una transición suave los conjuntos que representan dos conceptos. Al repetir ese procedimiento en las distintas transiciones, se construye la función semántica.

Vale la pena destacar que estamos aprovechando los datos para dar un sentido a los conceptos. Es, por tanto un ejemplo de aprendizaje de máquina. Corresponde a una estrategia de aprendizaje supervisado, porque existe información sobre cómo debe desempeñarse la función semántica: debe emular la opinión de los encuestados.

Por ello, la función semántica que se obtiene representa la subjetividad de un colectivo de personas, ya no de un individuo. Es, además, una subjetividad que se puede representar gráficamente (Figura 6.7). Por ello, este procedimiento puede ser una herramienta muy útil para ayudar a construir consensos entre personas.

6.4. Agrupamiento difuso

El objetivo del agrupamiento es la obtención de conjuntos (de grupos), buscando que se satisfagan dos condiciones: 1) que los elementos dentro de un mismo conjunto sean parecidos entre sí y 2) que los elementos que estén en conjuntos distintos sean diferentes entre sí.

Precisemos la idea del párrafo anterior con la ayuda de la Figura 6.9. Supongamos un conjunto U de n individuos, que será nuestro universo de discurso. Supongamos también que podemos conocer ciertas *características numéricas* de cada uno de esos individuos. El número de características que conocemos de cada individuo es d y, por tanto, podemos construir una tabla de datos con n filas y d columnas. Los algoritmos de agrupamiento buscan construir grupos, es decir, conjuntos, en el universo de discurso U , a partir de las características numéricas de los individuos.

Para determinar qué tan parecidos entre sí son dos individuos, los algoritmos de agrupamiento se apoyan en *medidas de distancia* que comparan sus características numéricas. Una medida de distancia para \mathbb{R}^d es una función $d(\mathbf{x}_1, \mathbf{x}_2) : \mathbb{R}^d \rightarrow \mathbb{R}$ que satisfaga las

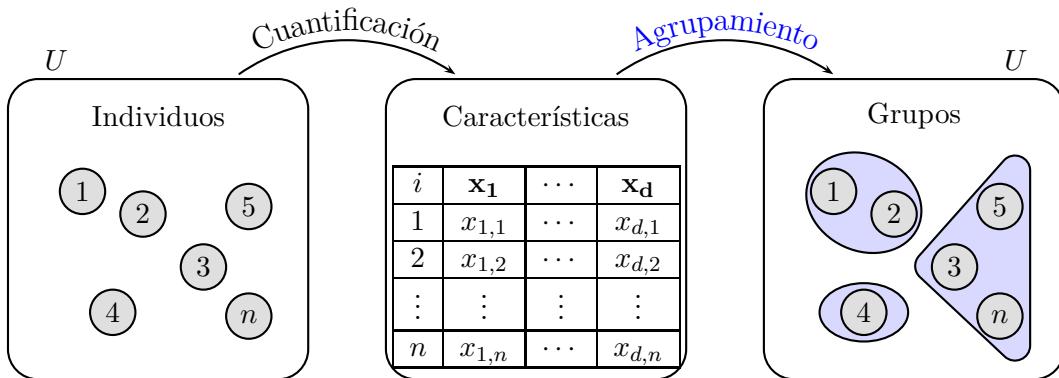


Figura 6.9: El proceso de agrupamiento

siguientes propiedades:

$$\begin{cases} d(\mathbf{x}, \mathbf{x}) = 0 & \forall \mathbf{x} \\ d(\mathbf{x}_1, \mathbf{x}_2) > 0 & \forall \mathbf{x}_1, \mathbf{x}_2 \mid \mathbf{x}_1 \neq \mathbf{x}_2 \\ d(\mathbf{x}_1, \mathbf{x}_2) = d(\mathbf{x}_2, \mathbf{x}_1) & \forall \mathbf{x}_1, \mathbf{x}_2 \\ d(\mathbf{x}_1, \mathbf{x}_3) < d(\mathbf{x}_1, \mathbf{x}_2) + d(\mathbf{x}_2, \mathbf{x}_3) & \forall \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \end{cases} \quad (6.10)$$

Existe una gran colección de algoritmos de agrupamiento, que aún hoy en día sigue creciendo. Algunas de las propiedades que se emplean para comparar distintos algoritmos son²: 1. el tipo de partición resultante 2. la estrategia de agrupamiento 3. la medida de distancia empleada 4. la complejidad numérica

El estudio detallado de las técnicas de agrupamiento excede el alcance de este documento. Debemos, eso sí, destacar que todas ellas responden al esquema del aprendizaje no supervisado (Figura 6.1c) porque no cuentan con información directa sobre cuál debería ser el agrupamiento correcto. Además, sólo unas pocas de esas técnicas están enfocadas en generar conjuntos difusos, que es nuestro tema central.

6.4.1. El algoritmo *Fuzzy c-means*

El algoritmo *Fuzzy c-means* es, sin duda, el algoritmo de agrupamiento para obtener conjuntos difusos más conocido. También es la base de una gran colección de algoritmos. Fue propuesto originalmente por Bezdek [6], como una generalización del conocido algoritmo *k-means*.

²Para tener un panorama general de los métodos de agrupamiento véase [54], por ejemplo. Aunque ese artículo documento fue publicado hace más de 20 años, y desde ese entonces han aparecido nuevos algoritmos la clasificación que allí se realiza y el análisis general siguen siendo válidos.

Supongamos que tenemos un universo de discurso de n individuos, y que cada uno de ellos tiene d características que se han cuantificado. Si denotamos por $x_{l,k}$ a la característica número l del individuo número k podemos definir:

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & & \ddots & \\ x_{d,1} & x_{d,2} & \cdots & x_{d,n} \end{bmatrix}_{d \times n} \quad \mathbf{X} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \cdots \ \mathbf{x}_n] \quad \mathbf{x}_k = \begin{bmatrix} x_{1,k} \\ \vdots \\ x_{l,k} \\ \vdots \\ x_{d,k} \end{bmatrix} \quad (6.11)$$

Nótese que \mathbf{x}_k es un vector columna que contiene las características numéricas del k -ésimo individuo, mientras que \mathbf{X} es la matriz que contiene todas las características de todos los individuos.

El objetivo del algoritmo *Fuzzy c-means* es el de obtener c conjuntos difusos. Es necesario asignar a cada individuo un grado de pertenencia a cada conjunto. Si denotamos por $u_{i,k}$ es el grado de pertenencia del individuo k al grupo i podemos definir:

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & u_{1,2} & \cdots & u_{1,n} \\ u_{2,1} & u_{2,2} & \cdots & u_{2,n} \\ \vdots & & \ddots & \\ u_{c,1} & u_{c,2} & \cdots & u_{c,n} \end{bmatrix}_{c \times n} \quad (6.12)$$

Los grados de pertenencia no pueden ser arbitrarios. Por una parte, sabemos $u_{i,k} \in [0, 1]$. Además, para que los grupos constituyan una partición del universo de discurso deben cumplirse las siguientes condiciones:

$$\sum_{i=1}^c u_{i,k} = 1 \quad \forall k \in \{1, 2, \dots, n\} \quad (6.13a)$$

$$\sum_{k=1}^n u_{i,k} > 0 \quad \forall i \in \{1, 2, \dots, c\} \quad (6.13b)$$

La condición 6.13a significa que la pertenencia de un individuo se distribuye en su totalidad entre los c grupos, en tanto que la condición 6.13b significa que no hay conjuntos vacíos.

Existen infinitas matrices \mathbf{U} que satisfacen las condiciones 6.13. ¿Cómo determinar cuál de ellas es la mejor?. Bezdek comienza por asignar a cada uno de los conjuntos difusos

un *prototipo*, que podemos interpretar como un valor central del conjunto. Si denotamos por $v_{l,i}$ a la característica número l del prototipo del conjunto número i podemos definir:

$$\mathbf{V} = \begin{bmatrix} v_{1,1} & v_{1,2} & \cdots & v_{1,c} \\ v_{2,1} & v_{2,2} & \cdots & v_{2,c} \\ \vdots & & \ddots & \\ v_{d,1} & v_{d,2} & \cdots & v_{d,c} \end{bmatrix}_{d \times c} \quad \mathbf{V} = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_c] \quad \mathbf{v}_i = \begin{bmatrix} v_{1,i} \\ \vdots \\ v_{l,i} \\ \vdots \\ v_{d,i} \end{bmatrix} \quad (6.14)$$

Nótese que \mathbf{v}_i es un vector columna que contiene las características numéricas del i -ésimo prototipo, mientras que \mathbf{V} es la matriz que contiene todas las características de todos los prototipos.

Bezdek utiliza un indicador para determinar qué tan bien hecha está una partición. Ese indicador es el siguiente:

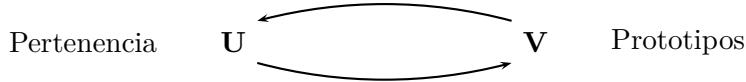
$$J_m(\mathbf{U}, \mathbf{V}) = \sum_{k=1}^n \sum_{i=1}^c (u_{i,k})^m \| \mathbf{x}_k - \mathbf{v}_i \|^2 \quad (6.15)$$

en donde $m \in \mathbb{R}$, $m > 1$, $\|\cdot\|$ es una norma diferenciable. Para ayudar a interpretar el significado de $J_m(\mathbf{U}, \mathbf{V})$, nótese que estamos haciendo el siguiente uso de los subíndices:

$$\left\{ \begin{array}{ll} i, j : & 1, 2, \dots, c \quad \text{Número de grupo} \\ k : & 1, 2, \dots, n \quad \text{Número de individuo} \end{array} \right. \quad (6.16)$$

¿Qué significa $J_m(\mathbf{U}, \mathbf{V})$? Veamos:

- $\| \mathbf{x}_k - \mathbf{v}_i \|^2$ es la distancia al cuadrado entre el individuo k y el prototipo i .
- $(u_{i,k})^m \| \mathbf{x}_k - \mathbf{v}_i \|^2$ es el error que se comete al asignar el individuo k al conjunto i , ponderado por (una potencia) del grado de pertenencia de ese individuo a ese conjunto.
- $\sum_{i=1}^c (u_{i,k})^m \| \mathbf{x}_k - \mathbf{v}_i \|^2$ es la suma de todos los errores cometidos en la asignación del individuo k a los c grupos, ponderado por (una potencia) sus grado de pertenencia.
- $\sum_{k=1}^n \sum_{i=1}^c (u_{i,k})^m \| \mathbf{x}_k - \mathbf{v}_i \|^2$ es la suma global de todos los errores de asignación de los n individuos a los c grupos, ponderado por (una potencia) sus grado de pertenencia.

Figura 6.10: Método de Picard aplicado el algoritmo *Fuzzy c-means*

Con esa interpretación, es claro que la mejor partición es aquella que minimice $J_m(\mathbf{U}, \mathbf{V})$.

Para minimizar $\mathbf{J}_m(\mathbf{U}, \mathbf{V})$ hay que encontrar las matrices \mathbf{U} , \mathbf{V} . En otras palabras, es necesario encontrar todos los grados de pertenencia $u_{i,k}$ y todas las componentes $v_{l,i}$ de los prototipos. En total son $n \times c^2 \times d$ parámetros. Bezdek demostró³ que las funciones de pertenencia $u_{i,k}^*$ y los vectores prototipo \mathbf{v}_i^* que minimizan \mathbf{J}_m deben cumplir las siguientes condiciones:

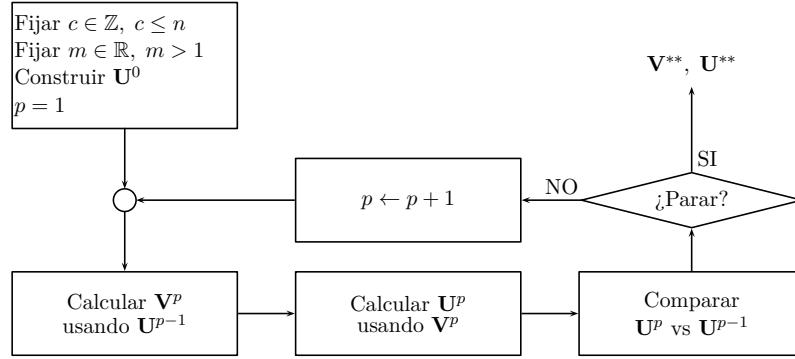
$$u_{i,k}^* = \left(\sum_{j=1}^c \left(\frac{\|\mathbf{x}_k - \mathbf{v}_i^*\|}{\|\mathbf{x}_k - \mathbf{v}_j^*\|} \right)^{\frac{2}{m-1}} \right)^{-1} \quad \forall i, k \quad (6.17a)$$

$$\mathbf{v}_i^* = \frac{\sum_{k=1}^n (u_{i,k}^*)^m \mathbf{x}_k}{\sum_{k=1}^n (u_{i,k}^*)^m} \quad \forall i \quad (6.17b)$$

Al analizar la ecuación 6.17 encontramos una tropiezo: para calcular los grados de pertenencia $u_{i,k}$ necesitamos los prototipos \mathbf{v}_i y para calcular los prototipos \mathbf{v}_i necesitamos los grados de pertenencia $u_{i,k}$. ¿Qué fue primero, el huevo o la gallina?.

Para solucionar este *impasse*, Bezdek propone el uso del método de Picard, que se ilustra en la Figura 6.10. La idea consiste en encontrar una matriz \mathbf{U} válida que nos permita encontrar una matriz \mathbf{V} , para luego recalcular \mathbf{U} y con esta nueva \mathbf{U} obtener una nueva \mathbf{V} ... y así entrar en un ciclo de mejoramiento de las dos matrices. El resultado es el algoritmo *Fuzzy c-means*, cuyo diagrama de flujo se muestra en la Figura 6.11. Los ejemplos 6.4 a 6.8 ilustran el funcionamiento del algoritmo. En todos ellos se ha utilizado la implementación disponible en **scilab** (sección C.4).

³La demostración original se publicó en [6]. Sin embargo, este documento no es fácil de obtener, por lo que se recomienda consultar alguna publicación posterior como, por ejemplo, [7]. La demostración de 6.17a emplea multiplicadores de Lagrange para incorporar la condición 6.13a, mientras que la de 6.17b es una optimización convencional sin restricciones.

Figura 6.11: Diagrama de flujo del algoritmo *Fuzzy c-means*

Ejemplo 6.4 (Fuzzy c -means aplicado a una distribución homogénea en \mathbb{R})
 Se ha creado un conjunto de 100 números aleatorios, cada uno de los cuales está en el intervalo $[0, 1]$. El conjunto se ha creado siguiendo una distribución de probabilidad homogénea. Se ha aplicado el algoritmo *Fuzzy c-means* para agrupar esos números. Se ha utilizado el algoritmo 4 veces, variando el número de conjuntos deseados, es decir, varioando c . Los valores que se han utiliozado son $c = 2, 3, 5, 10$. En todos los casos se ha utilizado $m = 2$.

La Figura 6.12 muestra los grupos obtenidos. Nótese que los conjuntos resultantes no son números difusos, porque para valores pequeños de α , los α -cortes no son intervalos cerrados. Ese pequeño 'rebote' de las funciones de transferencia es característico del algoritmo *Fuzzy c-means*.

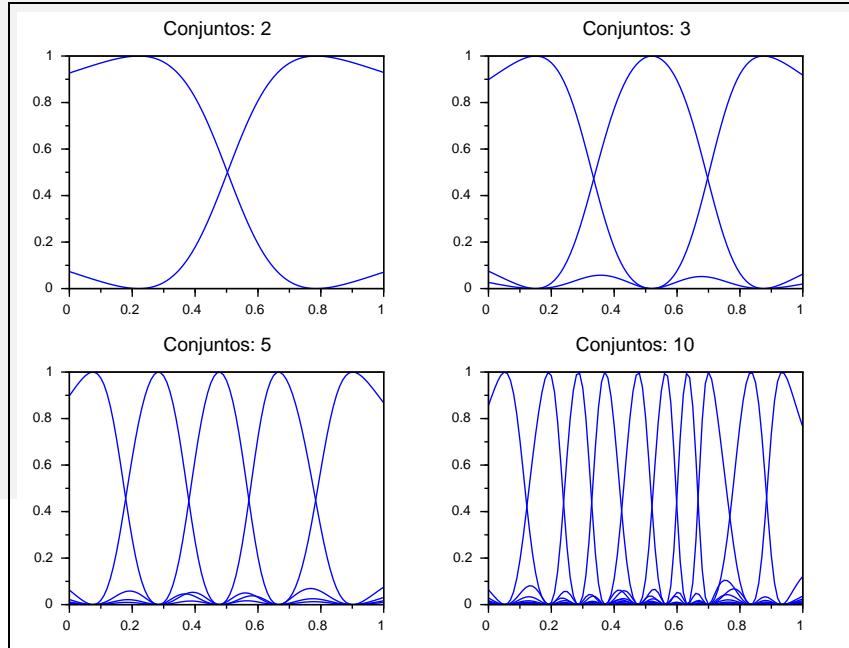


Figura 6.12: Grupos obtenidos en el Ejemplo 6.4

Ejemplo 6.5 (Convergencia del algoritmo *Fuzzy c-means*) Para ilustrar la forma en que evoluciona el algoritmo al aplicar el método de Picard, se ha tomado el caso en que $c = 3$ en el ejemplo 6.4. La Figura 6.13 muestra las funciones de pertenencia de los tres conjuntos después de 1, 3, 5 y 10 iteraciones. Para este ejemplo, los prototipos de la iteración inicial fueron obtenidos de forma aleatoria.

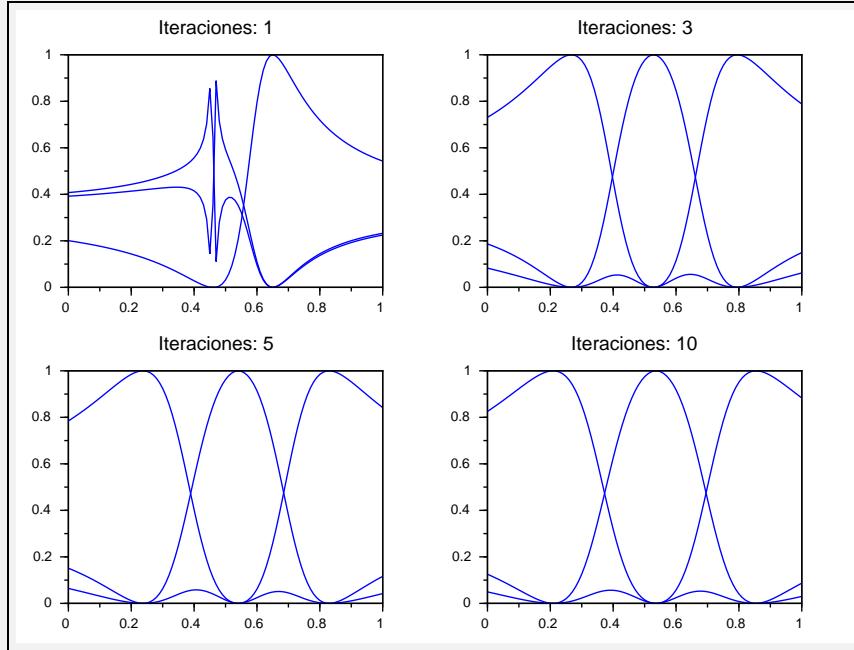


Figura 6.13: Evolución del algoritmo *Fuzzy c-means*

Ejemplo 6.6 (Efecto del parámetro m del algoritmo *Fuzzy c-means*) Para ilustrar el efecto del parámetro m sobre el agrupamiento resultante, se ha tomado el caso en que $c = 3$ en el ejemplo 6.4. La Figura 6.14 muestra las funciones de pertenencia de los tres conjuntos para cuatro valores diferentes de m . Los valores utilizados han sido $m = 1.1, 2.0, 3.0, 5.0$. Nótese que valores bajos de m llevan el algoritmo a una definición casi *crisp* de los conjuntos. De hecho, el valor $m = 1$ convierte el algoritmo *Fuzzy c-means* en el algoritmo concreto *k – means*. Por otra parte, valores altos de m alteran la forma haciendo más anchos los α -cortes bajos y angostos los α -cortes altos.

Este ejemplo ilustra la razón por la que es muy poco frecuente utilizar valores de m diferentes a 2.

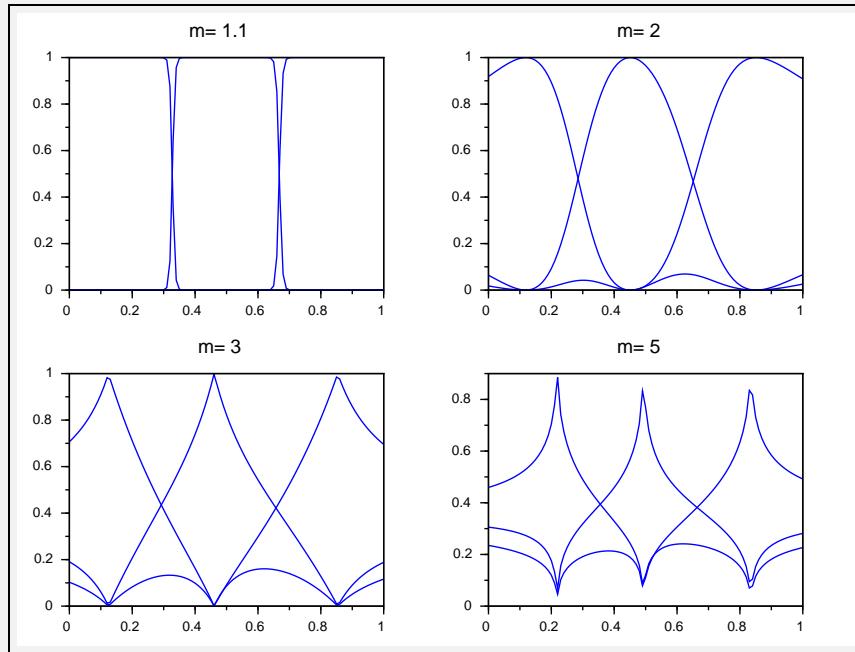


Figura 6.14: Efecto del parámetro m del algoritmo *Fuzzy c-means*

Ejemplo 6.7 (*Fuzzy c-means* aplicado a una distribución homogénea en \mathbb{R}^2)

Se ha creado un conjunto de 100 puntos aleatorios, cada uno de los cuales está en el cuadrado unitario $[0, 1] \times [0, 1]$. El conjunto de datos se ha creado siguiendo dos distribuciones de probabilidad homogéneas y puede verse en la Figura 6.15.a.

Se ha aplicado el algoritmo *Fuzzy c-means* para obtener 4 grupos. Los grupos obtenidos se muestran en la Figura 6.15.b.

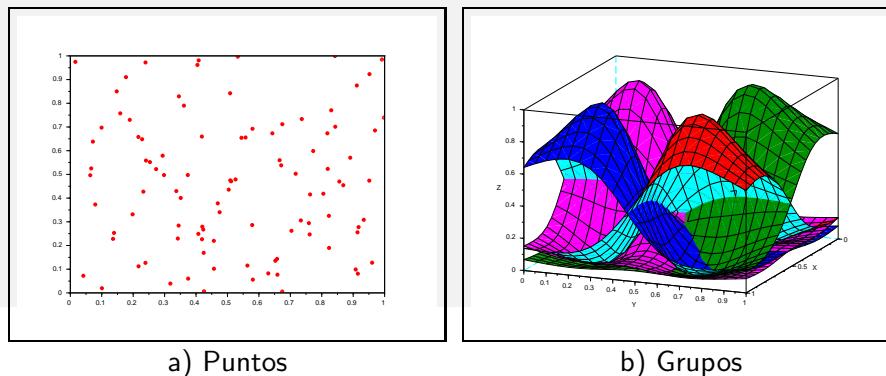


Figura 6.15: Puntos y grupos obtenidos en el ejemplo 6.7

Para ayudar a visualizar la forma de los grupos obtenidos, se ha elaborado la Figura 6.16 en la que se muestran de forma individual las superficies 3D y las curvas de nivel de cada uno de los cuatro conjuntos.

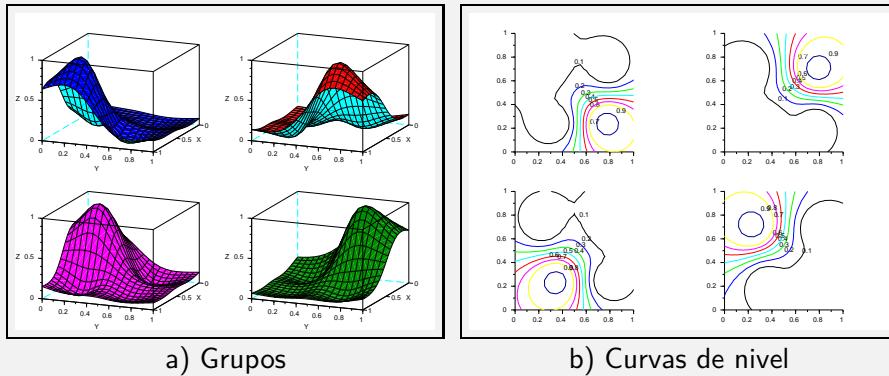


Figura 6.16: Grupos individuales obtenidos en el ejemplo 6.7

Ejemplo 6.8 (*Fuzzy c-means* aplicado a una distribución no homogénea en \mathbb{R}^2)
 Se ha creado el conjunto de 100 puntos en el cuadrado unitario $[0, 1] \times [0, 1]$ que se muestra en la Figura 6.17. Aunque allí se ve que los puntos podrían agruparse adecuadamente en dos grupos, se ha utilizado el algoritmo *Fuzzy c-means* para crear 4 grupos. Lo hemos hecho para destacar, precisamente, que el algoritmo por si sólo no ayuda a determinar cuál es el número más adecuado de grupos que se deberían obtener, sino que es un parámetro decidido *a priori* por el usuario.

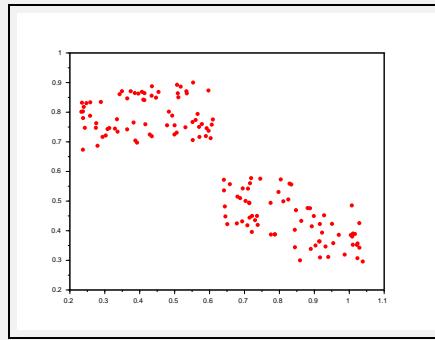


Figura 6.17: Puntos en el ejemplo 6.8

La Figura 6.18 muestra las superficies en 3D de los cuatro grupos obtenidos, primero juntos y luego cada uno por separado.

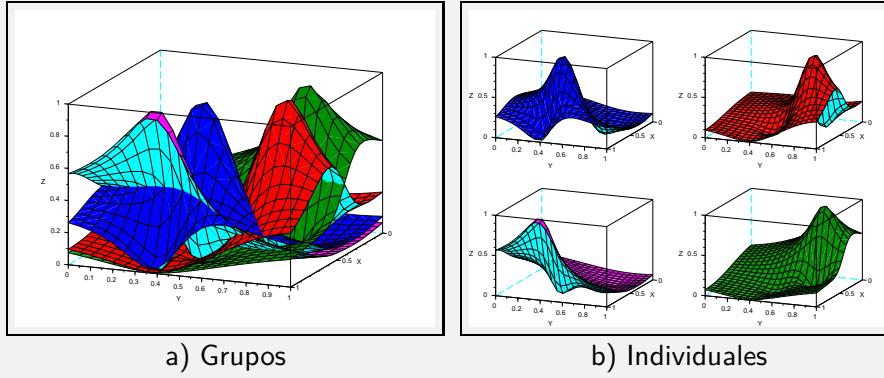


Figura 6.18: Grupos obtenidos en el ejemplo 6.8

Para ayudar a visualizar la forma de los grupos obtenidos, se ha elaborado también la Figura 6.19a. en la que se muestran de forma individual las curvas de nivel de cada uno de los cuatro conjuntos. También se ha elaborado también la Figura 6.19b. en la que además de las curvas de nivel se han marcado los puntos que se han agrupado.

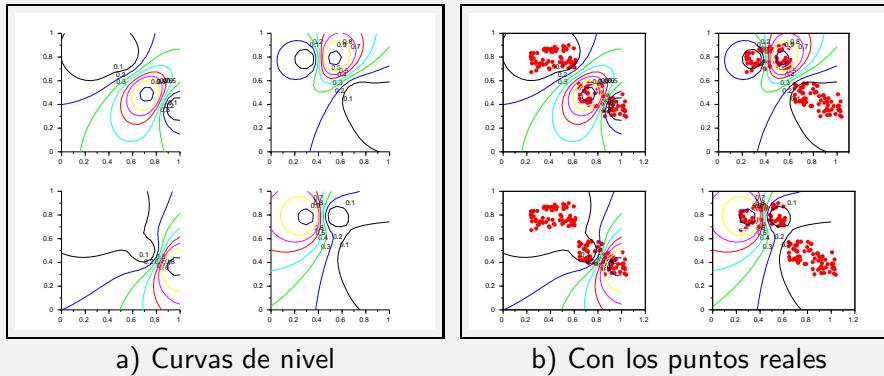


Figura 6.19: Curvas de nivel de los grupos obtenidos en el ejemplo 6.8

6.4.2. Diseño de la función semántica usando *Fuzzy c-means*

En esta sección exploramos una alternativa para diseñar variables lingüísticas utilizando *Fuzzy c-means*. Específicamente, nos referimos aquí a la obtención de la función semántica de la variable, es decir, a la obtención de los conjuntos difusos que acompañan a cada uno de los términos de la variable lingüística.

La situación que imaginamos es la siguiente: tenemos un conjunto de n puntos, cada uno de ellos en \mathbb{R}^d ; punto representa a un individuo y cada una de las d dimensiones

representa una característica de los individuos; la tarea consiste en construir una variable lingüística para cada una de esas características.

La estrategia que aquí se propone consiste en trabajar cada una de las dimensiones por separado. Con las coordenadas correspondientes se efectúa un agrupamiento difuso, y a partir de ese agrupamiento se obtiene la función de pertenencia de cada conjunto. Veamos en más detalle el procedimiento.

Supongamos que los n puntos están contenidos en una matriz \mathbf{X} como en la ecuación 6.11. En esa matriz la fila número l contiene las coordenadas de cada punto en la dimensión número l . Tomamos esa fila y le aplicamos el algoritmo *Fuzzy c-means*. En la Figura 6.20, a la izquierda, se ilustra el resultado para unos datos hipotéticos, en el que se ha usado $c = 3$. Lo que queremos es obtener una función semántica como la que se muestra a la derecha en esa misma figura.

La propuesta consiste en encontrar la mejor aproximación a número trapezoidal de cada uno de los grupos obtenidos mediante *Fuzzy c-means*, aplicando para ello las ideas de la sección 7.7.5. No obstante, para hacerlo tenemos que tener en cuenta lo siguiente:

- El orden de los grupos que entrega el algoritmo *Fuzzy c-means* no está predeterminado.
- El algoritmo *Fuzzy c-means* entrega como resultado sólo las matrices \mathbf{U} y \mathbf{V} . Es decir, no contamos con expresiones explícitas de los α -cortes de los grupos.
- Los grupos producidos por el algoritmo *Fuzzy c-means* no son números difusos.

Para obviar estas dificultades, se propone el algoritmo 1, cuya lógica es la siguiente: se construyen los conjuntos de izquierda a derecha; cada conjunto se obtiene utilizando el algoritmo 7, que necesita emplea como insumo el ajuste de línea recta de las rampas L y R . La rampa L del primer conjunto y la rampa R del último conjunto son líneas verticales. Para otros casos se procede así: La rampa R de un conjunto se obtiene con las condiciones de optimalidad (ecuación 6.17) aplicadas a un conjunto de puntos entre prototipo de ese conjunto y el del siguiente. La rampa L de un conjunto se obtiene como el complemento de la rampa L del anterior.

Ejemplo 6.9 (Diseño de variables lingüísticas con *Fuzzy c-means*) Este ejemplo ilustra el procedimiento propuesto para diseñar las variables lingüísticas en de un Sistema Basado en Reglas (SBR), empleando *Fuzzy c-means*. Para ello, hemos construido el conjunto de puntos que se muestra en la Figura 6.21a. El propósito es el de obtener un SBR que modele esos puntos, es decir, que si se aplica como entrada la coordenada x de uno de esos puntos, el SBR entregue como salida la coordenada y de ese punto.

Se ha construido manualmente un SBR cuyo desempeño se observa en la Figura 6.21b. en

Algoritmo 1 Diseño de una función semántica a partir de *Fuzzy c-means***Entradas:**

- \mathbf{U}, \mathbf{V} las matrices que arroja *Fuzzy c-means*
 c el número de grupos obtenidos
 l las variable cuya función semántica se quiere obtener

```

1: procedure FUNCIONSEMANTICA( $\mathbf{U}, \mathbf{V}, l$ )
2:    $p_1, \dots, p_c \leftarrow \text{sort}(v_{l,1}, v_{l,2}, \dots, v_{l,c})$             $\triangleright$  Ordenamiento de los prototipos
3:   for  $i = 1, 2, \dots, c$  do
4:     if  $i = 1$  then                                          $\triangleright$  Primer conjunto
5:        $m_{L,i} = 0$ 
6:        $b_{L,i} = \min_x$                                           $\triangleright$  Rampa vertical
7:     else
8:        $M \leftarrow 101$                                           $\triangleright$  Un valor arbitrario
9:        $\delta_x \leftarrow (v_i - v_{i-1})/M$                           $\triangleright$  Paso en  $x$ 
10:       $X[] \leftarrow \text{array}(M)$                                  $\triangleright$  Un arreglo de tamaño  $M$ 
11:       $A[] \leftarrow \text{array}(M)$                                  $\triangleright$  Un arreglo de tamaño  $M$ 
12:      for  $k = 1, 2 \dots M$  do
13:         $X[k] \leftarrow v_{i-1} + \delta_x \times k$                        $\triangleright$  Valor de  $x$ 
14:         $A[k] \leftarrow \left( \sum_{j=1}^c \left( \frac{x_k - p_i}{x_k - p_j} \right)^{\frac{2}{m-1}} \right)^{-1}$   $\triangleright$  Grado de pertenencia, es decir,  $\alpha$ 
15:         $m_{L,i}, b_{L,i} \leftarrow \text{Regression}(A, X)$                    $\triangleright$  Ajuste de linea recta
16:      if  $i = c$  then                                          $\triangleright$  Último conjunto
17:         $m_{R,i} = 0$ 
18:         $b_{R,i} = \max_x$                                           $\triangleright$  Rampa vertical
19:      else
20:         $m_{R,i} = -m_{L,i-1}$ 
21:         $b_{R,i} = b_{L,i}$ 
22:       $T_i \leftarrow \text{trapecio}(m_{L,i}, b_{L,i}, m_{R,i}, b_{R,i})$            $\triangleright$  Algoritmo 7, ecuación 7.130

```

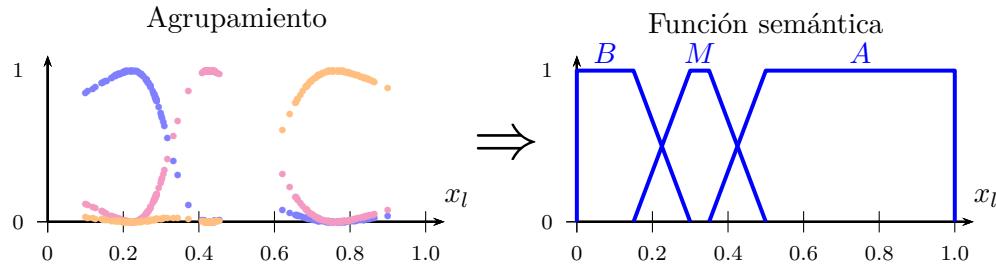


Figura 6.20

color azul. En esa misma figura se han marcado los puntos originales en color gris, para visualizar qué tan adecuado es el SBR como modelo de los datos. Hemos denominado a ese SBR ‘simétrico’, porque las variables lingüísticas de x y y se han construido manualmente con particiones difusas simétricas del intervalo $[0, 1]$, tal como se observa en la Figura 6.22.

Otras características del SBR son las siguientes:

- Difusor: singleton.
- Etiquetas de x : Bajo, Medio y Alto.
- Etiquetas de y : Bajo, Medio y Alto.
- Base de reglas: tres reglas ascendentes.
- Concresor: Altura.

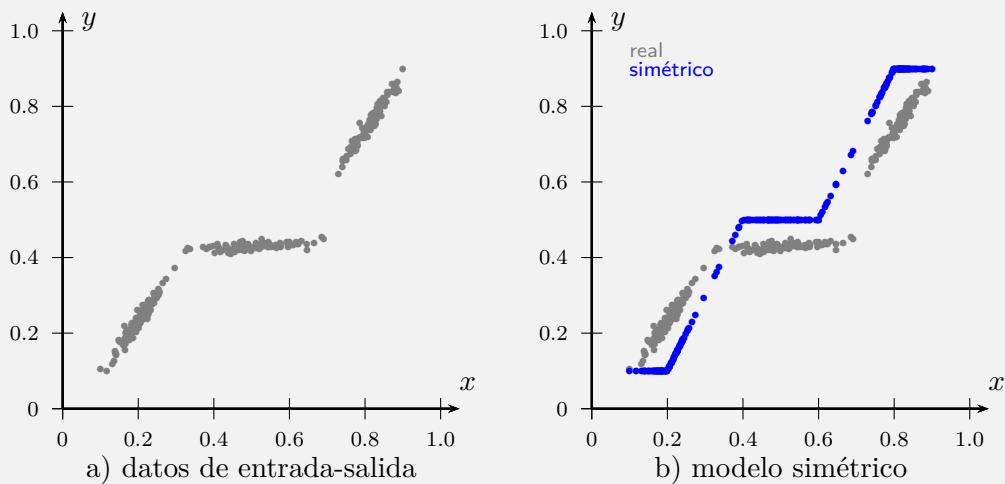


Figura 6.21: Datos para el ejemplo 6.9

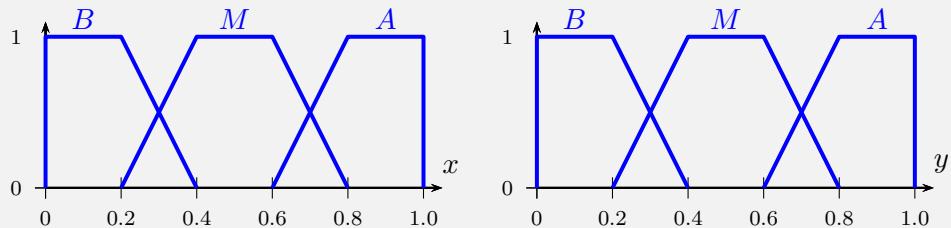


Figura 6.22: Variables simétricas en el ejemplo 6.9

Ahora formulamos la alternativa de diseñar las variables lingüísticas apoyándonos en el algoritmo Fuzzy c-means. Para ello, aplicamos el algoritmo con $c = 3$ al conjunto de coordenadas

x y y por separado. El resultado se muestra en la Figura 6.23.

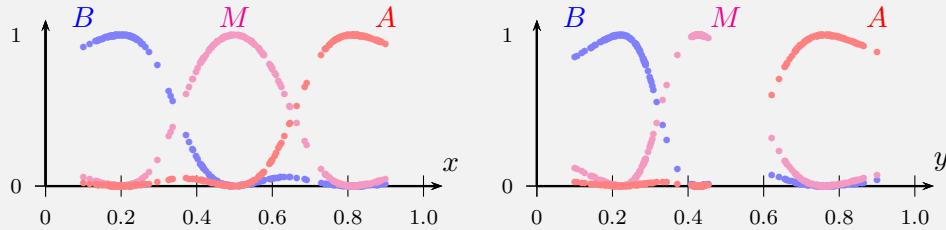


Figura 6.23: Agrupamiento *Fuzzy c-means* en el ejemplo 6.9

Para obtener las variables lingüísticas a partir de los resultados del agrupamiento se han utilizado las siguientes estrategias:

- FCM - rectos: se han obtenido ajustes de línea recta de las rampas de cada uno de los grupos y se ha obtenido la mejor aproximación trapezoidal correspondiente (ver sección 7.7). Las variables obtenidas se muestran en la Figura 6.24.
- FCM - curvos: utilizando los vértices de los trapecios anteriores, se han construido conjuntos curvos (ver Figura 3.3).
- Conjunto central - rectos: se ha tomado el núcleo del conjunto Medio de cada variable para construir una partición difusa con triángulos en los conjuntos Bajo y Alto. Las variables obtenidas se muestran en la Figura 6.25.
- Conjunto central - curvos: utilizando los vértices de los trapecios anteriores, se han construido conjuntos curvos (ver Figura 3.3).

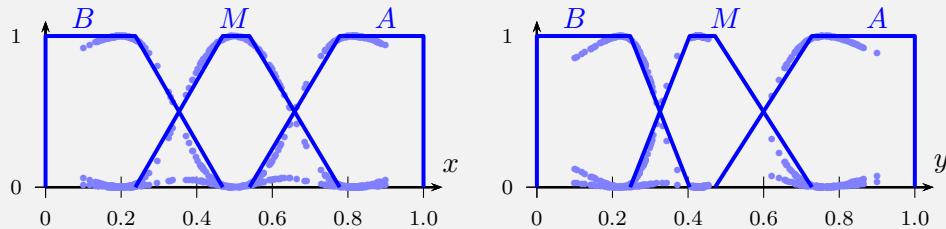


Figura 6.24: Estrategia FCM en el ejemplo 6.9

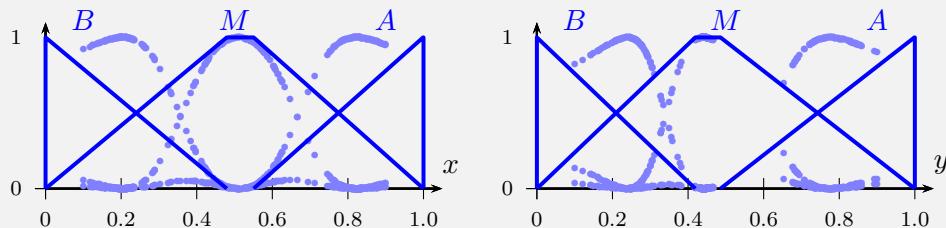


Figura 6.25: Estrategia del conjunto central en el ejemplo 6.9

El desempeño de los cuatro SBR así obtenidos se muestra en la Figura 6.26. También se han dibujado allí los datos originares y el desempeño del sistema simétrico, para comparar visualmente los cambios en los modelos correspondientes.

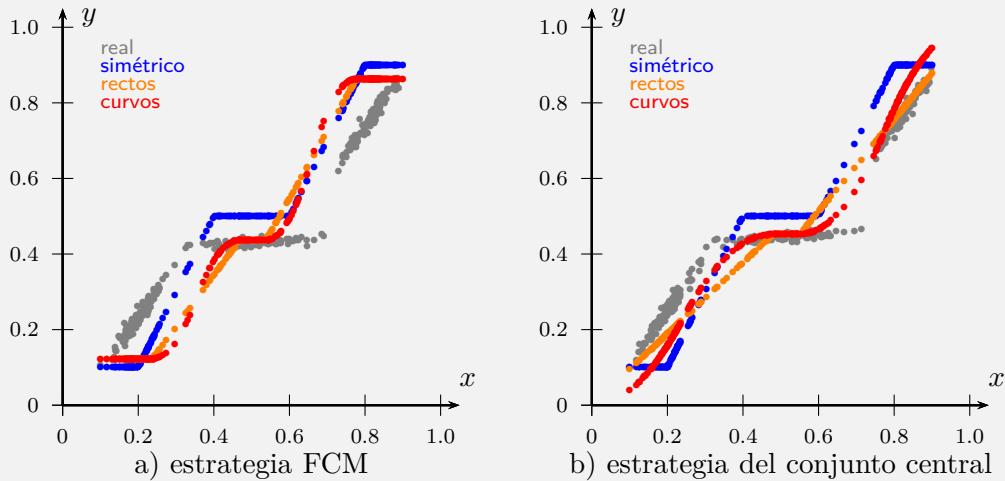


Figura 6.26: Modelos en el ejemplo 6.9

La Tabla 6.3 permite hacer una comparación cuantitativa de los desempeños. La columna ‘error’ muestra el error relativo promedio de cada modelo en relación con los datos reales. Por su parte, la columna ‘mejoría’ muestra qué tanto disminuyó ese error relativo al emplear el diseño basado en el agrupamiento, en relación con el diseño simétrico.

Tabla 6.3: Errores en el ejemplo 6.9

Estrategia	Forma	Error	Mejoría (%)
Simétrico	Rectos	0.006209	
FCM	Rectos	0.005931	4.470131
FCM	Curvos	0.006022	3.012082
Conjunto central	Rectos	0.002934	53.145731
Conjunto central	Curvos	0.003671	41.376293

6.5. Identificación de reglas de SBR

En esta sección se presentan dos algoritmos que permiten construir la Base de Reglas a partir del desempeño esperado del SBR. En la sección 4.2 se muestra que en la Base de

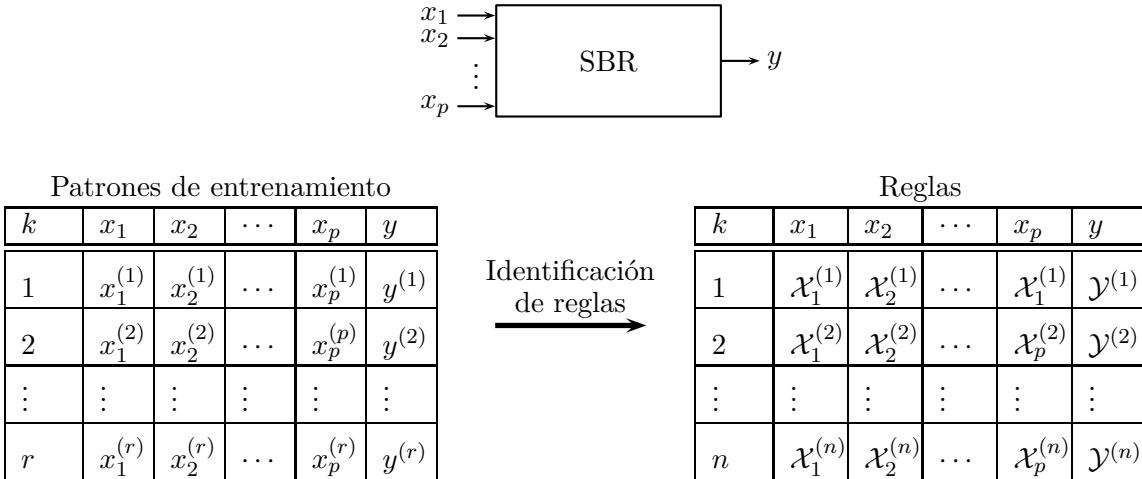


Figura 6.27: Algoritmos de identificación de reglas

Reglas se aloja el conocimiento sobre cómo debe desempeñarse un SBR. Por esa razón, se dice que estos algoritmos ayudan a *descubrir conocimiento*.

La situación se ilustra en la Figura 6.27 para un SBR de p entradas y una salida: A partir de un conjunto de r *patrones de entrenamiento* el algoritmo debe identificar un conjunto de n reglas.

Cada patrón de entrenamiento es una conjunto de $p + 1$ datos numéricos (uno por cada una de las p entradas y uno más por la salida). El significado de cada uno de esos patrones es el siguiente: si el SBR recibe las entradas de ese patrón, debe producir la salida de ese patrón. Por eso, cada uno de los patrones de entrenamiento es un comportamiento individual deseado. En conjunto, toda la tabla debe describir el comportamiento deseado global del SBR.

Por su parte, cada una de las reglas que se deben identificar son del tipo “**si** antecedente **entonces** consecuente”. Eso significa que los elementos \mathcal{X} y \mathcal{Y} que aparecen en la base de reglas son términos lingüísticos de las variables de entrada y salida.

Debido a que estos algoritmos utilizan el desempeño deseado como fuente de información, se trata de algoritmos de aprendizaje supervisado (Figura 6.1b). Hay un gran marco conceptual alrededor del aprendizaje supervisado que no abordaremos en este texto⁴. Ese marco permite evaluar aspectos como la complejidad del conjunto de patrones de entrenamiento y la validez de la base de reglas obtenida. En este texto, sin embar-

⁴Véase, por ejemplo, [5]

go, nos limitaremos a presentar en las secciones 6.5.1 y 6.5.2 cómo funcionan los dos algoritmos más sencillos que se conocen en la literatura⁵.

6.5.1. Entrenamiento de universos fijos

Este algoritmo fue propuesto por Wang en [52]. El nombre de *universos fijos* se lo hemos dado nosotros para indicar que los universos de discurso de entrada y salida no se modifican al ejecutar el algoritmo⁶.

La lógica del algoritmo es la siguiente: cada patrón de entrenamiento puede ayudarnos a descubrir una regla específica. Debido a que varios patrones pueden referirse a la misma regla, o a dos reglas parecidas, se necesita un criterio para discernir a cuál patrón ‘creerle más’.

La Figura 6.28 ayuda a explicar cómo funciona el algoritmo de entrenamiento de universos fijos. Se ha supuesto un sistema de dos entradas (x_1, x_2) y una salida (y). También se ha supuesto que cada una de esas variables tiene una variable lingüística que ya ha sido definida, con las siguientes etiquetas:

- x_1 : Bajo, Medio, Alto, representadas por las letras B, M, A en la Figura 6.28.b.
- x_2 : Negativo, Cero, Positivo, representadas por las letras N, C, P en la Figura 6.28.b.
- y : Poco, Algo, Mucho, representadas por las letras P, A, M en la Figura 6.28.b.

El algoritmo funciona recorriendo la tabla de patrones de entrenamiento (Figura 6.27) línea por línea. Al llegar a la línea número k tenemos el k -ésimo patrón de entrenamiento que se muestra en la Figura 6.28.b.: la variable x_1 tiene el valor numérico $x_1^{(k)}$, la variable x_2 tiene el valor numérico $x_2^{(k)}$ y la variable y tiene el valor numérico $y^{(k)}$.

Para cada variable, se calcula el grado de pertenencia del valor numérico del patrón a todas sus etiquetas y se selecciona aquella que tenga el mayor grado de pertenencia. En la Figura 6.28 se muestra esto así:

- Para x_1 : El grado de pertenencia de $x_1^{(k)}$ al término Bajo es 0.75, al término Medio es 0.25 y al término Alto es 0.0. Se selecciona la etiqueta Bajo.
- Para x_2 : El grado de pertenencia de $x_2^{(k)}$ al término Negativo es 0.0, al término Cero es 1.0 y al término Positivo es 0.0. Se selecciona la etiqueta Cero.

⁵Ambos algoritmos están implementados en UNFuzzy, que es la herramienta empleada en los ejemplos de esta sección (véase C.1)

⁶El nombre asignado por Wang es el de “Diseño de sistemas difusos mediante un esquema de búsqueda en tablas”.

- Para y : El grado de pertenencia de $y^{(k)}$ al término Poco es 0.0, al término Algo es 0.25 y al término Mucho es 0.75. Se selecciona la etiqueta Mucho.

Con las etiquetas seleccionadas se construye una regla candidata (Figura 6.28.c.) y se calcula su grado de *certeza*. La certeza de la regla candidata es el producto de los grados de pertenencia de las entradas a las etiquetas seleccionadas.

Antes de incorporar la regla a la base de reglas se verifica si ya existe en la base alguna regla que tenga el mismo antecedente. En tal caso, es necesario decidir cuál de las dos reglas debe permanecer en la base para asegurar su coherencia (Ejemplo 4.6). La regla ganadora es la que tenga la mayor certeza.

El algoritmo 2 formaliza la explicación anterior.

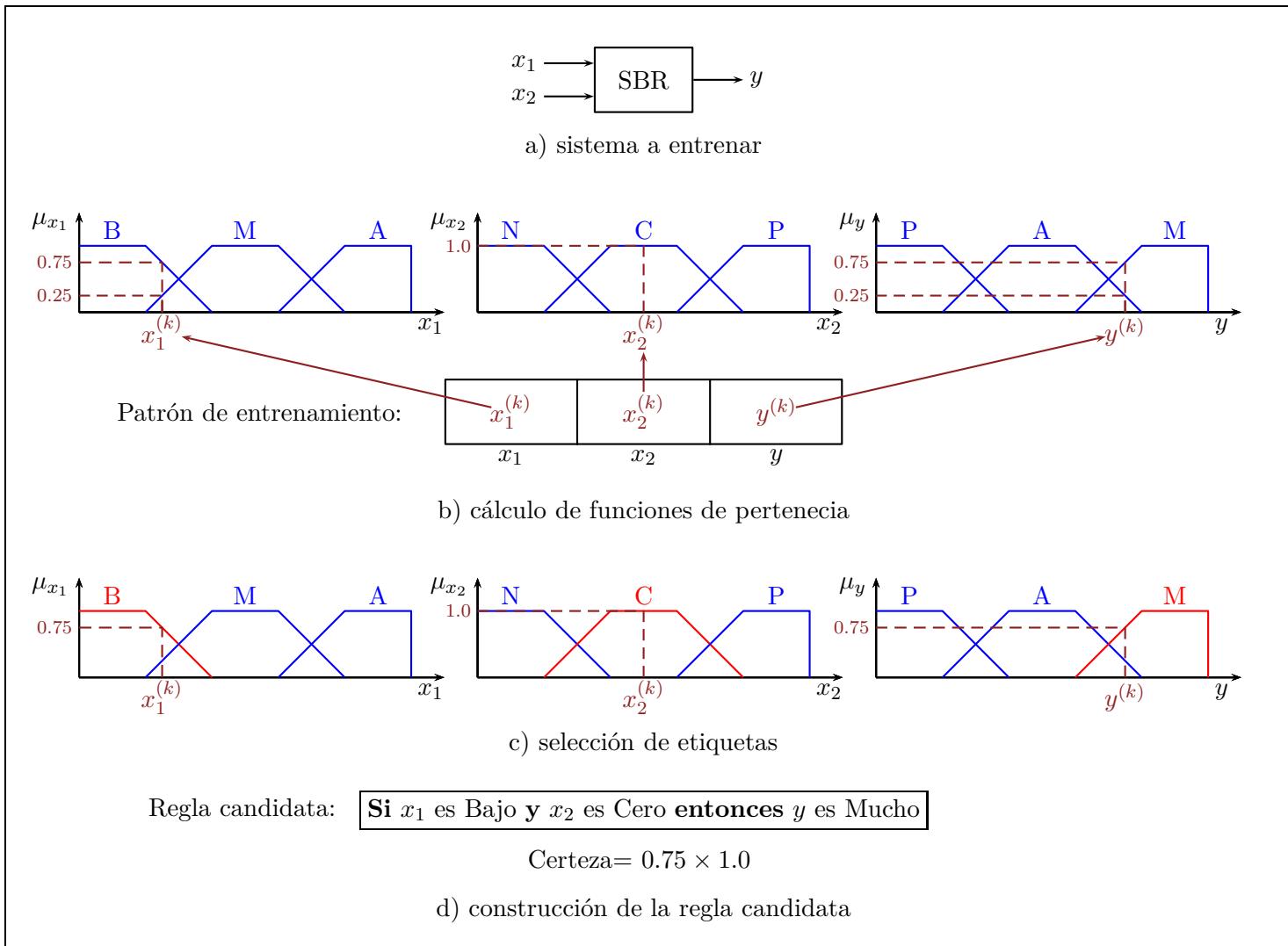


Figura 6.28: Entrenamiento de universos fijos

Algoritmo 2 Universos Fijos**Entradas:** S un SBR tipo Mamdani N matriz con los patrones de entrenamiento

```

1: procedure UNIVERSOSFIJOS( $S, N$ )
2:    $p \leftarrow \text{entradas}(S)$                                  $\triangleright$  Número de entradas de  $S$ 
3:    $q \leftarrow \text{salidas}(S)$                                  $\triangleright$  Número de salidas de  $S$ 
4:    $r \leftarrow \text{size}(N)$                                      $\triangleright$  Número de patrones
5:   for  $k = 1, 2, \dots, r$  do                                 $\triangleright$  Iterar sobre cada patrón de entrenamiento
6:      $A \leftarrow$  un vector vacío de tamaño  $p$                    $\triangleright$  Antecedentes de una regla
7:      $C \leftarrow$  un vector vacío de tamaño  $q$                    $\triangleright$  Consecuentes de una regla
8:      $\sigma \leftarrow 1$                                           $\triangleright$  Certeza
9:     for  $i = 1, 2, \dots p$  do                                 $\triangleright$  Iterar sobre cada entrada
10:     $N_{k,i} \leftarrow N[k][i]$                                  $\triangleright$  Entrada  $i$  del patrón número  $k$ 
11:     $t_i \leftarrow$  Número de términos de la entrada  $i$ 
12:     $m \leftarrow$  un vector vacío de tamaño  $t_i$ 
13:    for  $j = 1, 2, \dots t_i$  do                                 $\triangleright$  Pertenencia a cada término de la entrada
14:       $m[j] \leftarrow \mu_{E_{i,j}}(N_{k,i})$                    $\triangleright$  Pertenencia a cada término de la entrada
15:       $A[i] \leftarrow i \mid m[i] = \max(m)$                    $\triangleright$  Término de máxima pertenencia
16:       $\sigma \leftarrow \sigma \times A[i]$                              $\triangleright$  Actualización de la certeza
17:    for  $i = 1, 2, \dots q$  do                                 $\triangleright$  Iterar sobre cada salida
18:       $N_{k,i} \leftarrow N[k][p+i]$                                  $\triangleright$  Salida  $i$  del patrón número  $k$ 
19:       $t_i \leftarrow$  Número de términos de la salida  $i$ 
20:       $m \leftarrow$  un vector vacío de tamaño  $t_i$ 
21:      for  $j = 1, 2, \dots t_i$  do                                 $\triangleright$  Pertenencia a cada término de la salida
22:         $m[j] \leftarrow \mu_{S_{i,j}}(N_{k,i})$                    $\triangleright$  Pertenencia a cada término de la salida
23:         $C[i] \leftarrow i \mid m[i] = \max(m)$                    $\triangleright$  Término de máxima pertenencia
24:         $Rc \leftarrow \text{Regla}(A, C, \sigma)$                    $\triangleright$  Construir la regla candidata
25:         $r \leftarrow \text{BuscarRegla}(S, A)$   $\triangleright$  Buscar en la base una regla con igual antecedente
26:        if  $r == \text{NULL}$  then                                 $\triangleright$  No se encontró la regla
27:          Adicionar  $Rc$  a la base de reglas
28:        else
29:           $\sigma_r \leftarrow$  certeza de la regla  $r$ 
30:          if  $\sigma > \sigma_r$  then
31:            Sustituir  $r$  por  $Rc$  en la base de reglas
32:        return  $S$                                           $\triangleright$  el sistema entrenado

```

Ejemplo 6.10 (Entrenamiento de Universos Fijos) Para ilustrar el algoritmo de entrenamiento de universos fijos vamos a suponer que queremos construir un SBR que imite el comportamiento de un sistema de una entrada x y una salida y que se rige por la ecuación 6.18, y cuyo comportamiento se muestra en la Figura 6.29

$$y(x) = 0.7 [0.6 \sin(\pi x) + 0.3 \sin(3\pi x) + 0.1 \sin(5\pi x)] \quad (6.18)$$

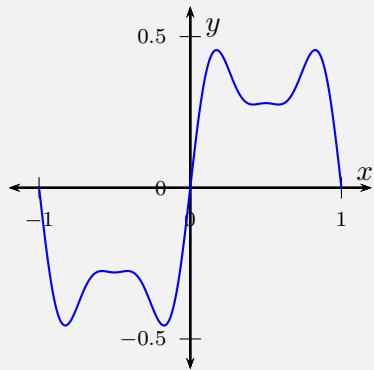


Figura 6.29: Función de prueba para entrenamiento de SBR

También vamos a suponer que los patrones de entrenamiento han sido obtenidos al aplicar a la entrada una señal sinusoidal descrita por la ecuación 6.19 y que las mediciones de la entrada y la salida no tienen errores ni incertidumbre. En total, tenemos una tabla con 250 patrones de entrenamiento y cada uno de ellos consiste en una pareja de datos $(x^{(k)}, y^{(k)})$.

$$x(k) = \sin(0.008\pi k) \quad k = 0, 1, \dots, 250 \quad (6.19)$$

Para construir el SBR se ha decidido que tanto x como y tengan una variable lingüística con cinco etiquetas (muy bajo, bajo, neutro, alto y muy alto) y que sus conjuntos difusos sean los que se muestran en la Figura 6.30.

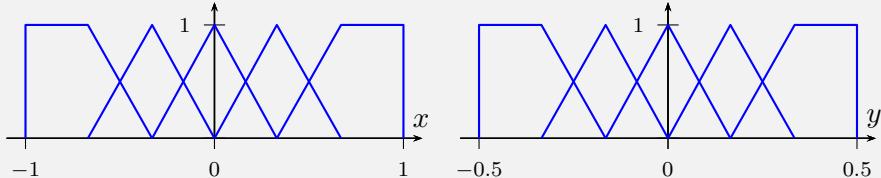


Figura 6.30: Variables lingüísticas en el ejemplo 6.10

Tabla 6.4: Reglas obtenidas en el ejemplo 6.10

Regla	x_1	y_1
R 1	4	5
R 2	5	5
R 3	3	3
R 4	2	1
R 5	1	1

La Tabla 6.4 muestra el resultado de aplicar el algoritmo de universos fijos. Se han identificado 5 reglas. Los números en esa tabla indican el número de la etiqueta que debe ir en la regla. El número 1 indica la etiqueta 'muy bajo', el número 2 la etiqueta 'bajo' y así sucesivamente. Podemos, entonces, redactar las siguientes reglas:

- Regla 1: **Si** x es alto **entonces** y es muy alto.
- Regla 2: **Si** x es muy alto **entonces** y es muy alto.
- Regla 3: **Si** x es neutro **entonces** y es neutro.
- Regla 4: **Si** x es bajo **entonces** y es muy bajo.
- Regla 5: **Si** x es muy bajo **entonces** y es muy bajo.

Al utilizar un SBR tipo Mamdani, con congresor de centro de gravedad se obtiene un desempeño que se muestra en color rojo en la Figura 6.31. Puede verse que el sistema logra modelar (aproximar) el desempeño del sistema original, aunque no de forma perfecta. En los ejemplos posteriores se mejorará esa aproximación.

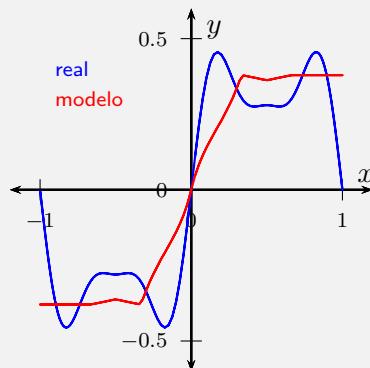


Figura 6.31: Aproximación obtenida en el ejemplo 6.10

Ejemplo 6.11 (Entrenamiento de Universos Fijos. Número de etiquetas)

Retomemos el Ejemplo 6.10. En ese ejemplo se utilizó un SBR en el que cada variable tenía cinco etiquetas. Ahora vamos a explorar el efecto de utilizar un número diferente de etiquetas.

La Figura 6.32 muestra el desempeño del SBR cuando se utilizan 5, 7, 9, 11, 13 y 15 etiquetas en cada variable lingüística. Es evidente que el número de etiquetas incide en la precisión de la aproximación de la función.

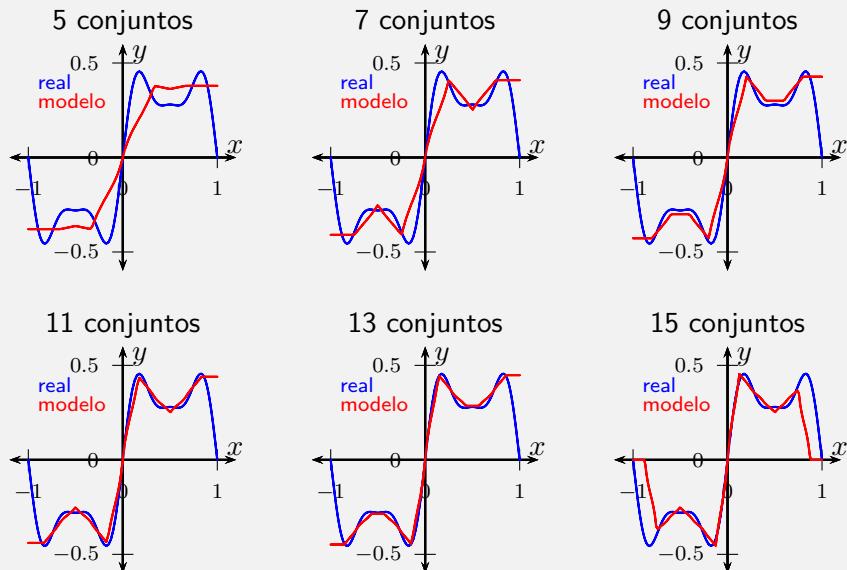


Figura 6.32: Desempeño de los SBR del Ejemplo 6.11

Ejemplo 6.12 (Entrenamiento de Universos Fijos. Tipo de conjuntos)

Retomemos los ejemplos 6.10 y 6.11. Vamos ahora a explorar el efecto de modificar la forma de los conjuntos difusos. Para ello, consideramos cuatro casos diferentes:

- Conjuntos rectos: definidos de forma semejante al ejemplo 6.10.
- Conjuntos rectos cortos: definidos como en la Figura 3.6c.
- Conjuntos curvos: definidos suavizando los conjuntos rectos como se ilustra en la Figura 3.4.

- Conjuntos curvos cortos: definidos suavizando los conjuntos rectos cortos como se ilustra en la Figura 3.4.

Para cada uno de los casos se probó el desempeño con 3, 5, 7, 9, 11, 13 y 15 etiquetas. Las Figuras 6.33 a 6.39 muestran los desempeños de cada SBR.

Para entender los resultados se ha preparado la Tabla 6.5. En ella se han consignado los errores cuadráticos promedio de cada SBR. Puede observarse allí que el menor se obtiene con el sistema de 15 conjuntos rectos cortos.

También es interesante observar que los sistemas con conjuntos cortos (rectos y curvos) tienen un mejor desempeño que los que no usan conjuntos cortos. Esto, sin duda, se debe al comportamiento de los sistemas en los extremos de la curva, ya que al usar conjuntos no cortos, hay un tramo en que la salida es constante.

Tabla 6.5: Errores en el ejemplo 6.12

Número de conjuntos	Tipo de conjuntos			
	Rectos	Rectos cortos	Curvos	Curvos cortos
3	0.132	0.283	0.135	0.283
5	0.135	0.183	0.138	0.186
7	0.116	0.128	0.119	0.133
9	0.109	0.084	0.109	0.091
11	0.105	0.062	0.104	0.064
13	0.105	0.043	0.105	0.047
15	0.099	0.030	0.102	0.033

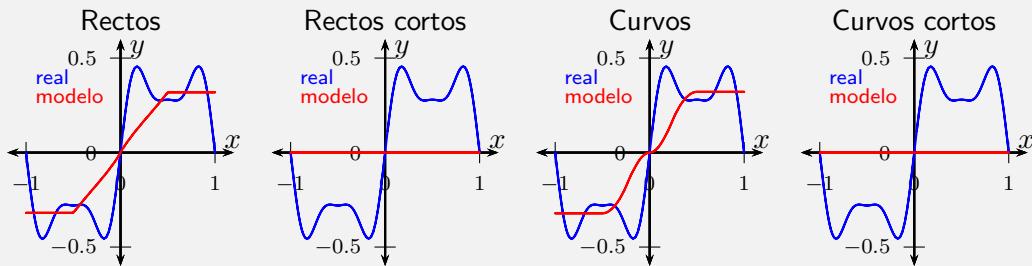


Figura 6.33: Aproximaciones en el ejemplo 6.12 con 3 conjuntos

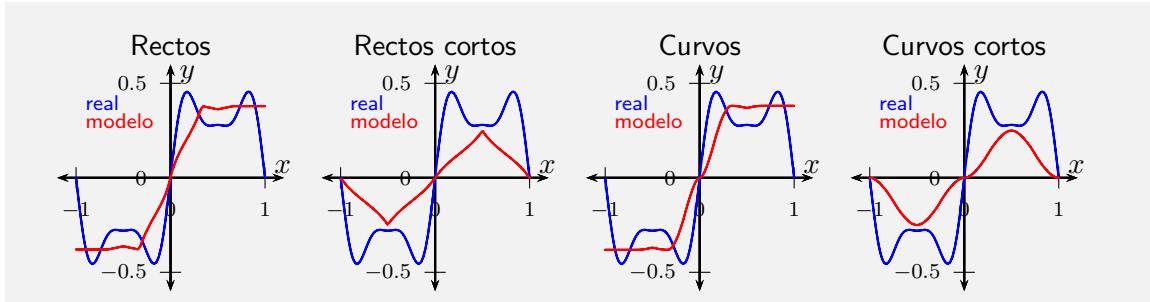


Figura 6.34: Aproximaciones en el ejemplo 6.12 con 5 conjuntos

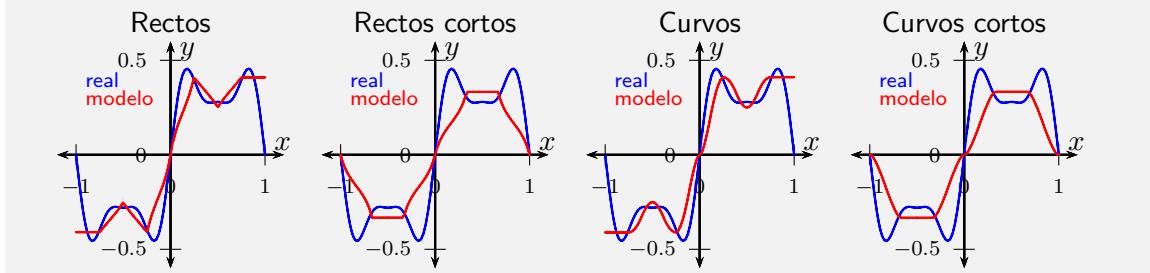


Figura 6.35: Aproximaciones en el ejemplo 6.12 con 7 conjuntos

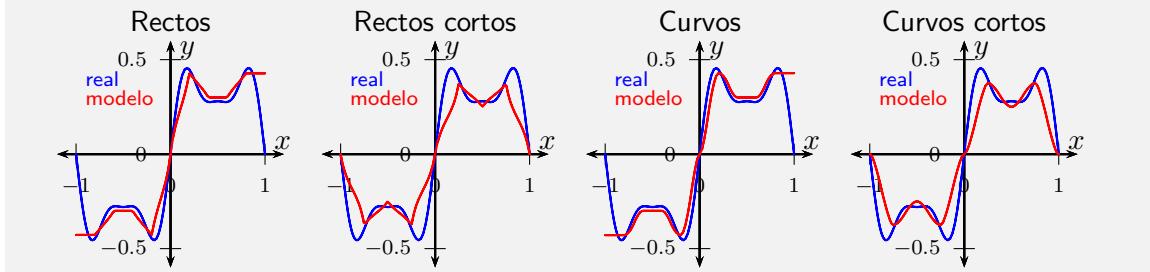


Figura 6.36: Aproximaciones en el ejemplo 6.12 con 9 conjuntos

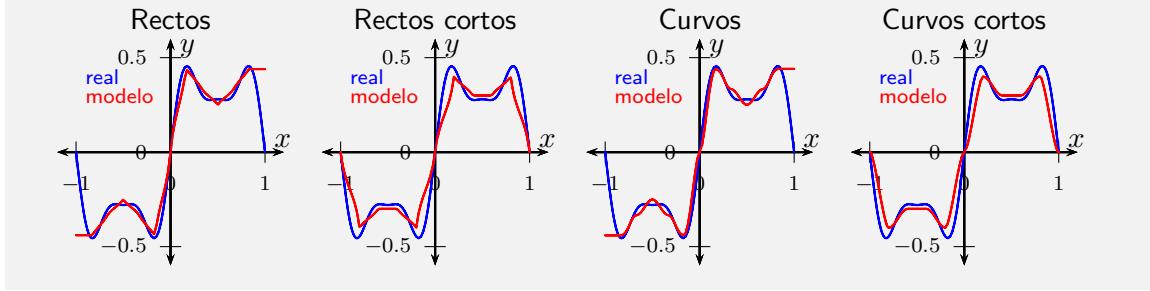


Figura 6.37: Aproximaciones en el ejemplo 6.12 con 11 conjuntos

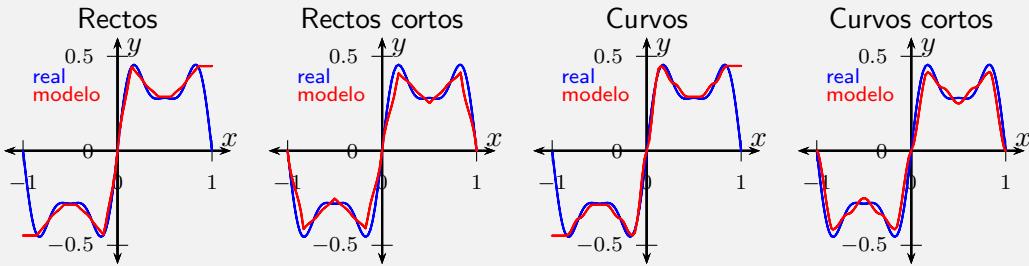


Figura 6.38: Aproximaciones en el ejemplo 6.12 con 13 conjuntos

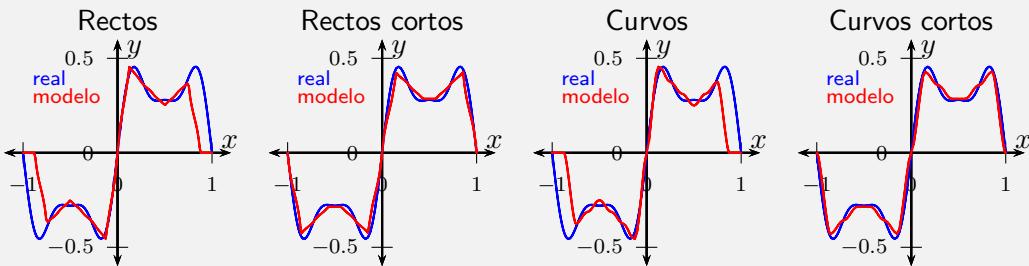


Figura 6.39: Aproximaciones en el ejemplo 6.12 con 15 conjuntos

6.5.2. Entrenamiento de universos variables

Este algoritmo es una versión muy simplificada de otro propuesto por Wang en [53, 42]. El nombre de *universos variables* se lo hemos dado nosotros para indicar que los universos de discurso de entrada y salida son modificados al ejecutar el algoritmo⁷.

La Figura 6.40 ayuda a explicar cómo funciona el algoritmo de entrenamiento de universos fijos. Se ha supuesto un sistema de dos entradas (x_1, x_2) y una salida (y). Ninguna de estas variables tiene definida su función semántica, es decir, no se han seleccionado sus etiquetas ni los conjuntos difusos asociados a ellas.

El algoritmo funciona recorriendo la tabla de patrones de entrenamiento (Figura 6.27) línea por línea. Al llegar a la línea número k tenemos el k -ésimo patron de entrenamiento que se muestra en la Figura 6.28.b.: la variable x_1 tiene el valor numérico $x_1^{(k)}$, la variable x_2 tiene el valor numérico $x_2^{(k)}$ y la variable y tiene el valor numérico $y^{(k)}$.

⁷El algoritmo propuesto por Wang identifica Funciones Difusas Base para poder efectuar un ajuste de mínimos cuadrados.

Para cada variable, se construye un conjunto difuso centrado en su valor numérico. La forma del conjunto debe haber sido predeterminada. Por ejemplo, en la Figura 6.40 se han usado conjuntos triangulares para x_1 y para y y un conjunto trapezoidal para x_2 . A cada uno de esos conjuntos se le asigna una etiqueta arbitraria (En la Figura 6.40 hemos usado ‘Etiqueta k’ para nombrar los tres conjuntos) y con esas etiquetas se construye una regla que se adiciona directamente a la base.

El algoritmo 3 formaliza la explicación anterior.

Algoritmo 3 Universos Variables

Entradas:

S un SBR tipo Mamdani
 N matriz con los patrones de entrenamiento
 Q_e Conjuntos prototipo de entrada
 Q_s Conjuntos prototipo de salida

```

1: procedure UNIVERSOSVARIABLES( $S, P, Q_e, Q_s$ )
2:    $p \leftarrow \text{entradas}(S)$                                  $\triangleright$  Número de entradas de  $S$ 
3:    $q \leftarrow \text{salidas}(S)$                                  $\triangleright$  Número de salidas de  $S$ 
4:    $r \leftarrow \text{size}(N)$                                      $\triangleright$  Número de patrones
5:   for  $k = 1, 2, \dots, r$  do                                $\triangleright$  Iterar sobre cada patrón de entrenamiento
6:      $A \leftarrow$  un vector vacío de tamaño  $p$                    $\triangleright$  Antecedentes de una regla
7:      $C \leftarrow$  un vector vacío de tamaño  $q$                    $\triangleright$  Consecuentes de una regla
8:     for  $i = 1, 2, \dots, p$  do                             $\triangleright$  Iterar sobre cada entrada
9:        $N_{k,i} \leftarrow N[k][i]$                                  $\triangleright$  Entrada  $i$  del patrón número  $k$ 
10:       $c \leftarrow Q_e[i]$                                       $\triangleright$  un nuevo conjunto difuso
11:      desplazar el centro de  $c$  hasta  $N_{k,i}$ 
12:      adicionar  $c$  a la variable lingüística de la entrada  $i$ 
13:       $m \leftarrow$  posición de  $c$  en la variable
14:       $A[i] = m$ 
15:      for  $i = 1, 2, \dots, q$  do                             $\triangleright$  Iterar sobre cada salida
16:         $N_{k,i} \leftarrow N[k][p + i]$                        $\triangleright$  Salida  $i$  del patrón número  $k$ 
17:         $c \leftarrow Q_s[i]$                                       $\triangleright$  un nuevo conjunto difuso
18:        desplazar el centro de  $c$  hasta  $N_{k,i}$ 
19:        adicionar  $c$  a la variable lingüística de la salida  $i$ 
20:         $m \leftarrow$  posición de  $c$  en la variable
21:         $C[i] = m$ 
22:       $\sigma \leftarrow 1$                                           $\triangleright$  Certeza
23:       $Rc \leftarrow \text{Regla}(A, C, \sigma)$                      $\triangleright$  Construir la regla candidata
24:      Adicionar  $Rc$  a la base de reglas
25:    return  $S$                                           $\triangleright$  el sistema entrenado

```

Es importante destacar que el algoritmo de universos variables crea una nueva etiqueta

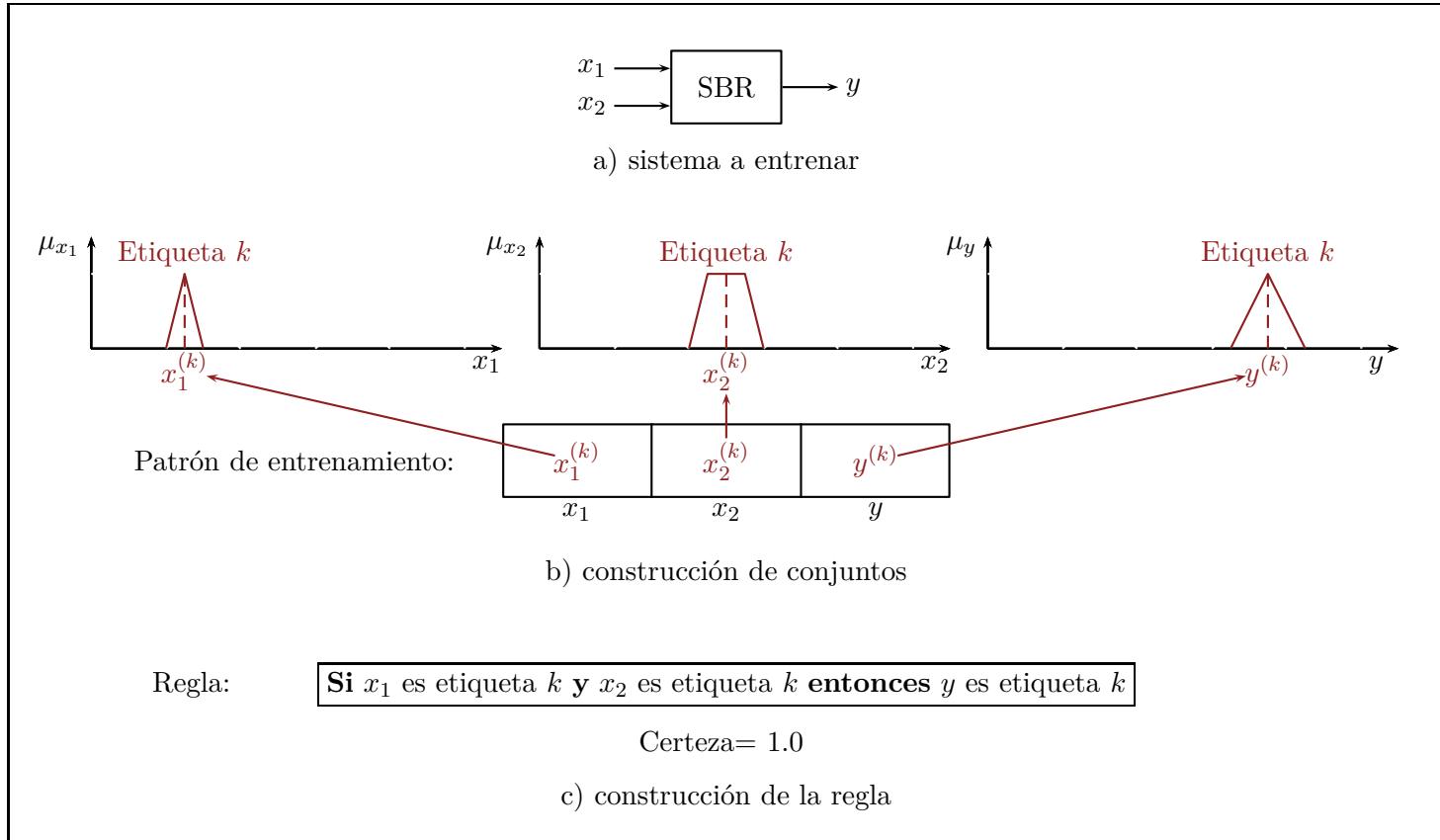


Figura 6.40: Entrenamiento de universos variables

en cada variable por cada patrón. Así, un conjunto de entrenamiento de 200 patrones producirá variables lingüísticas con 200 etiquetas. Es imposible asignar un significado lingüístico a semejante cantidad de etiquetas. De hecho, es difícil asignar un significado lingüístico a más de 9 etiquetas. Es por eso que las etiquetas que se asignan son arbitrarias, porque no se pretende que tengan un sentido lingüístico.

Este hecho hace que la Base de Reglas obtenida no sea *interpretable* por un humano. Nótese que el entrenamiento por Universos Fijos sí puede generar un SBR interpretable (Ejemplo 6.10). ¿Cuál método es mejor? Si se requiere interpretar las reglas obtenidas, no hay duda: el entrenamiento por Universos Fijos es mejor. Pero si no se requiere que sea interpretable, es muy probable que el entrenamiento por universos variable arroje aproximaciones más exactas.

Ejemplo 6.13 (Entrenamiento de Universos Variables) Retomemos el ejemplo 6.10 y apliquemos ahora el algoritmo de universos variables a los mismos patrones de entrenamiento.

Para aplicar el algoritmo es necesario definir la forma del conjunto difuso que se va a adicionar con cada patrón de entrenamiento. En este ejemplo hemos probado cuatro formas diferentes (Triángulo, Pi, Campana y Pi-campana), con anchos (soportes) de 0.1, 0.3 y 0.5. Los desempeños de los SBR resultantes se muestran en las Figuras 6.41 a 6.43.

También se ha preparado la Tabla 6.6 con los errores cuadráticos promedio, para comparar los desempeños. Estos desempeños también pueden compararse con los del ejemplo 6.12 (Tabla 6.5). Es evidente que los SBR con ancho 0.1 tienen un mejor desempeño que los SBR entrenados con universos fijos. Este desempeño superior es a costa de la interpretabilidad de la Base de reglas obtenida.

Tabla 6.6: Errores en el ejemplo 6.13

Ancho	Tipo de conjuntos			
	Triángulo	Pi	Campana	PiCampana
0.1	0.013	0.014	0.011	0.014
0.2	0.037	0.041	0.030	0.038
0.3	0.065	0.072	0.055	0.070

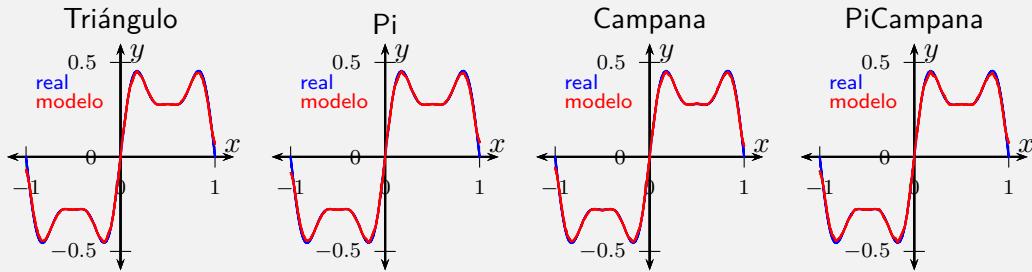
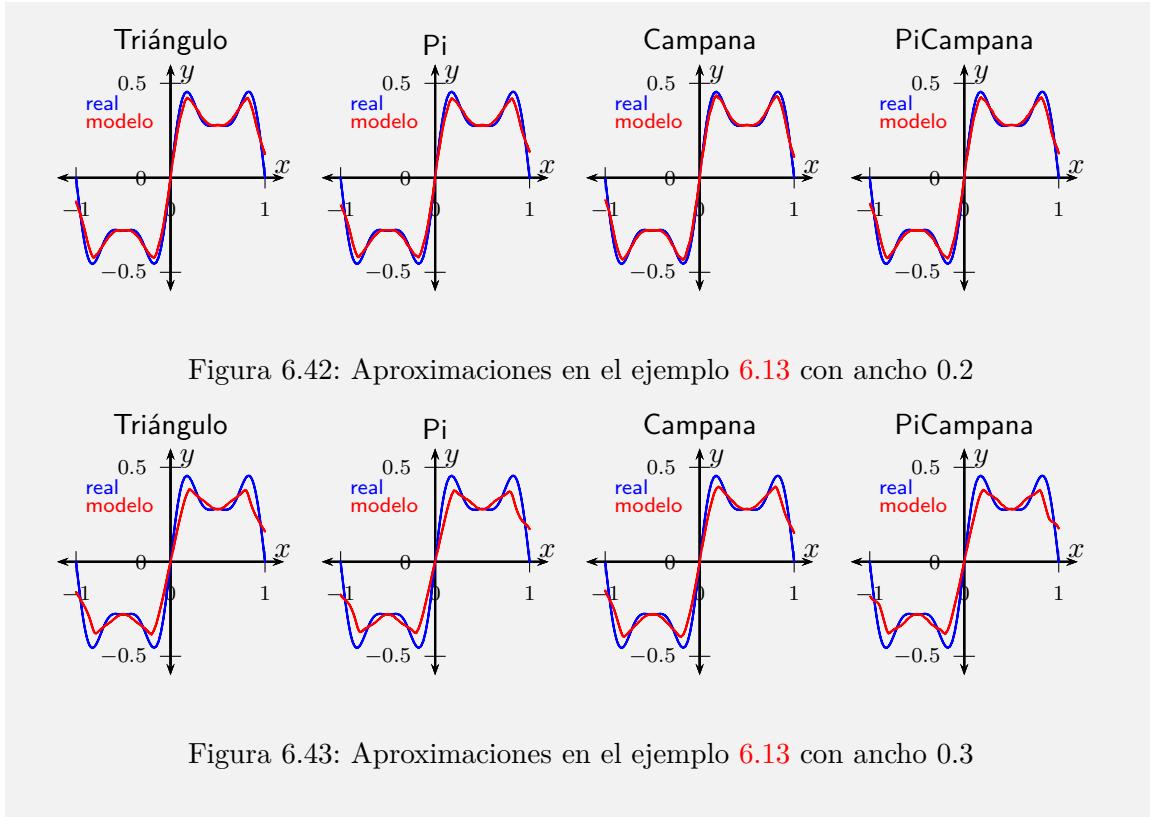


Figura 6.41: Aproximaciones en el ejemplo 6.13 con ancho 0.1



6.6. Optimización mediante heurísticas

A la hora de buscar la solución a un problema dado, en ocasiones contamos con procedimientos explícitos para encontrar esa solución. Sabemos, por ejemplo, cómo encontrar las raíces de un polinomio de segundo orden. No obstante, existen problemas en los que no es viable obtener una solución exacta, bien porque no se conoce método alguno que provea esa solución, o bien porque se conoce pero resulta ser muy complejo de implementar o muy lento en su ejecución.

¡Qué bueno sería contar con un algoritmo capaz de encontrar la mejor solución a cualquier problema! Algo así como lo que se ilustra en la Figura 6.44a. Ese algoritmo, sin embargo, no existe o al menos no lo conocemos.

En informática, una *heurística* es un algoritmo de propósito general que ayuda a buscar buenas soluciones a problemas dados (véase [18]). La idea fundamental de una heurística es la de *buscar* “buenas” soluciones al problema, aunque no se pueda asegurar que

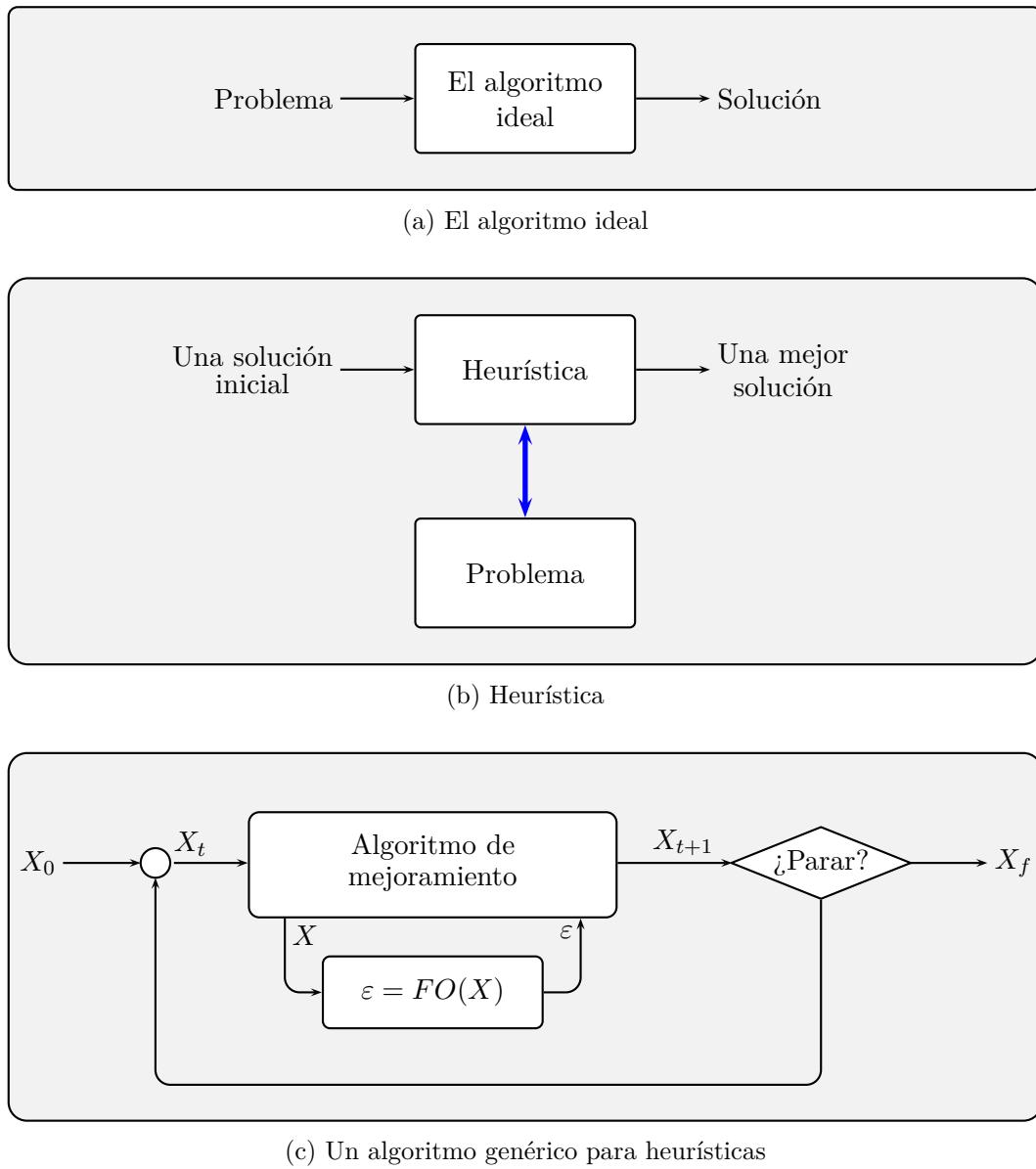


Figura 6.44: Heurísticas

sean soluciones “exactas”⁸. La Figura 6.44b ilustra esta situación: Las heurísticas son algoritmos que pueden interactuar con un problema dado, y que a partir de una solución inicial pueden entregar una mejor solución.

La interacción entre la heurística y el problema, representada por la flecha en azul de la Figura 6.44b requiere una explicación. Las heurísticas deben poder probar diferentes soluciones al problema y deben poder determinar entre dos soluciones cuál es mejor y cuál es peor. Para hacerlo, el problema debe formularse matemáticamente de la siguiente forma: *Encontrar el valor óptimo de X , denotado por X_{opt} tal que*

$$\begin{aligned} X_{opt} &= X | FO(X) = \min_{X \in \mathcal{X}} (FO(X)) \\ FO(X) &: \mathcal{X} \times \mathbb{R} \end{aligned} \quad (6.20)$$

En donde \mathcal{X} es el *espacio de búsqueda*, es decir, es el conjunto que contiene todas las soluciones posibles. $FO(X)$ es una función que calcula un número real, denominada la *Función Objetivo*. Nótese que el problema se ha formulado como la búsqueda del valor de X que hace que $FO(X)$ sea mínimo; por ello Las heurísticas son *métodos de optimización*.

La Figura 6.44c muestra la forma en que funcionan muchas heurísticas. Son algoritmos iterativos que parten de una solución inicial X_0 y en cada iteración emplean un algoritmo que busca una mejor solución X_{t+1} a partir de la mejor solución que se ha encontrado hasta ese momento, X_t . Para hacerlo, necesitan interactuar con una pieza de software que calcule el valor de $\varepsilon = FO(X)$.

Hay una gran colección de heurísticas, con nombres curiosos: recocido simulado, escalada simple, búsqueda tabú, algoritmos genéticos, enjambres de partículas, colonias de abejas, colonias de hormigas, migración de aves, etc. Los nombres suelen estar asociados a procesos naturales que inspiran el algoritmo.

Es posible emplear heurísticas para optimizar el diseño de SBR. A continuación se ilustra ese proceso con una heurística específica, denominada algoritmos genéticos (AG). Para ello, en la sección 6.6.1 se hace una presentación somera de dicha heurística, y en la sección 6.6.2 se muestran dos estrategias para aplicarlos a la optimización de SBR.

6.6.1. Algoritmos genéticos

La razón por la cual los Algoritmos Genéticos tienen ese nombre, es porque se trata de estrategias que fueron concebidas como una imitación simplificada de la mecánica que ha gobernado los procesos genéticos de selección natural que se dan en la naturaleza.

⁸Curiosamente, aunque estas técnicas se encargan de *buscar*, el término *heurístico* proviene del griego *heuriskien*, que significa *encontrar*.

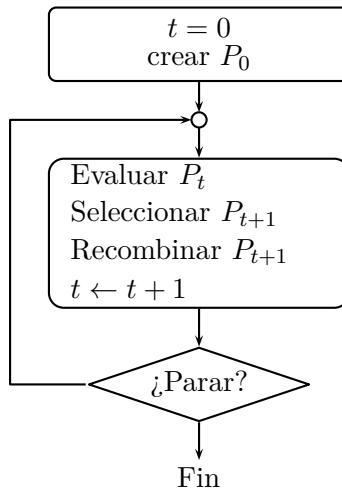


Figura 6.45: Algoritmo genético básico

Por tanto, ha tomado algunos términos propios de las ciencias biológicas para describir los componentes del algoritmo (se habla de *genes*, *mutación*, etc.). No hay que perder de vista, sin embargo, que los AG no pretenden modelar el proceso biológico, tan sólo se trata de una estrategia heurística de búsqueda de soluciones inspirada en dicho proceso.

La estructura básica de un AG se muestra en la Figura 6.45. Cada iteración del AG se denomina una *generación*. En cada generación hay una colección de posibles soluciones; cada posible solución es un *individuo*, y todos los individuos de una misma generación son una *población*. El AG crea una nueva población para cada generación, buscando que haya mejores individuos. Primero selecciona un subconjunto de la población, buscando mantener aquellos que tienen una mejor función objetivo. Despues, obtiene nuevos individuos recombinándolos. Para ello, a cada individuo se le asigna un *genoma* que es una pieza de información que contiene los aspectos que hacen que un individuo sea distinto a otro. El AG opera sobre el genoma para efectuar la recombinación, imitando los procesos de selección natural. En todas las etapas se utilizan procesos aleatorios.

En resumen, en un AG:

- P_t es una población de individuos.
- Cada individuo es una posible solución.
- Cada individuo tiene un genoma.
- Se seleccionan los mejores individuos, según su función objetivo.

- Se recombinan los genomas para crear nuevos individuos.
- En cada etapa hay procesos aleatorios.

Un aspecto crítico del diseño de un AG es la definición del genoma. Esta tarea consiste en codificar y decodificar la información de cada individuo de forma tal que pueda usarla el AG. En el apéndice B se hace una presentación detallada de cada etapa del AG y de sus operadores internos para distintos tipos de codificación. Los ejemplos 6.14 y 6.15 ilustran cómo se pueden usar para optimizar una función matemática elemental⁹.

Ejemplo 6.14 (Búsqueda mediante algoritmos genéticos) Consideremos ahora la siguiente función de dos variables, cuyo comportamiento se muestra en la Figura 6.46

$$f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2 \quad (6.21)$$

El valor mínimo de esa función sucede en el punto $(3, 2)$ y es cero. El propósito de este ejemplo es ilustrar la forma en que un algoritmo genético busca una solución. Por eso, vamos a suponer que no sabemos en qué punto sucede el mínimo de la función ni cuánto vale ese mínimo.

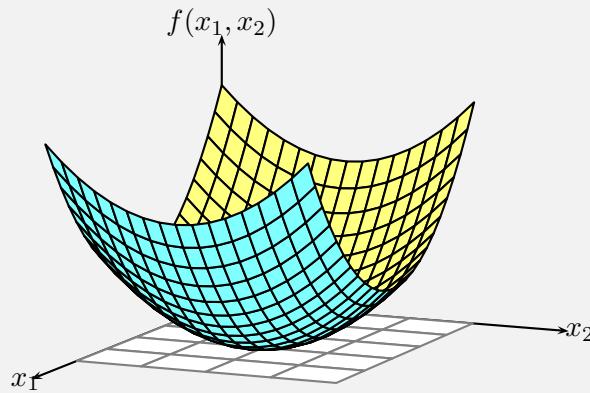


Figura 6.46: Función a optimizar en el ejemplo 6.14

Lo primero que debemos hacer es formular el problema como la búsqueda del valor más pequeño de una función objetivo. En este caso esta tarea es trivial, porque el problema consiste, precisamente, en encontrar el mínimo valor de $f(x_1, x_2)$. Por lo tanto, la función objetivo es

$$FO(x_1, x_2) = f(x_1, x_2) = (x_1 - 3)^2 + (x_2 - 2)^2 \quad (6.22)$$

⁹para desarrollar los ejemplos de esta sección se ha empleado la librería UNGenético (sección C.3). Además, los ejemplos en los que se optimiza un SBR emplean el código fuente generado por UNFuzzy (sección C.1)

Para aplicar el algoritmo de genético básico (Figura 6.45), es necesario seleccionar algunos valores y operadores internos del algoritmo. Los más importantes son los siguientes:

- Espacio de búsqueda: $[0, 6] \times [0, 6]$
- Número máximo de generaciones: 20
- Número de individuos en cada generación: 11
- Tipo de codificación genética: codificación real

La Figura 6.47 muestra las poblaciones de algunas de las generaciones que se han obtenido al correr el algoritmo. La población inicial (Generación 0) se ha construido de forma aleatoria. Son 11 puntos (x_1, x_2) aleatorios en el rectángulo $[0, 0] \times [6, 6]$, que se han dibujado con color azul.

Al ejecutar el algoritmo, se obtiene una nueva población para cada generación. En cada generación habrá 11 individuos nuevos. En cada una de las gráficas de la Figura 6.47 se ha marcado en rojo el punto $(3, 2)$ que es la solución del problema. Nótese que al avanzar en la iteración algunos individuos de la población (es decir, algunos puntos azules) se acercan al punto solución.

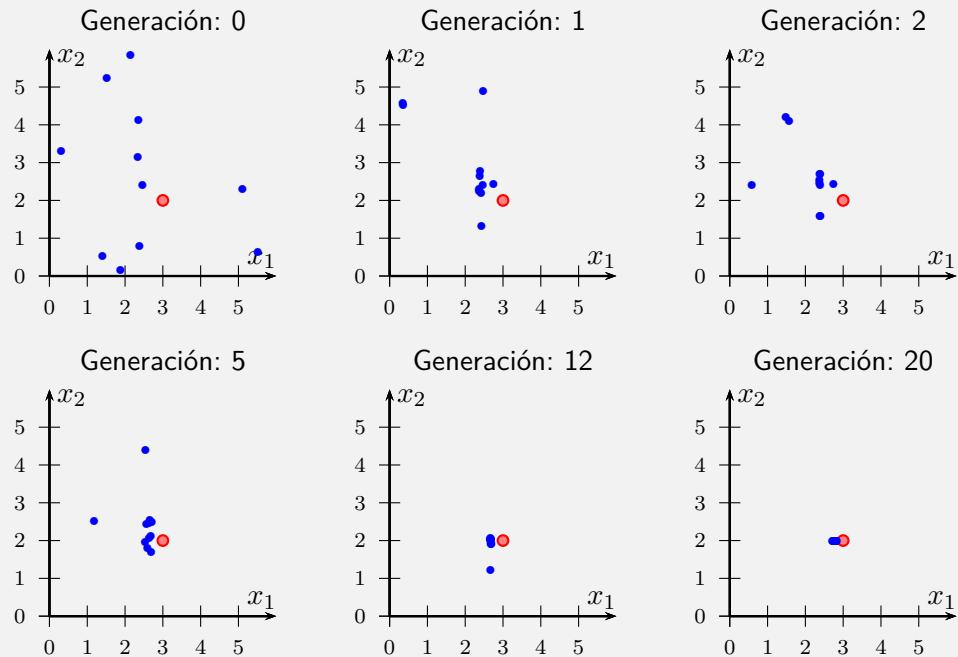


Figura 6.47: Poblaciones de algunas generaciones en el ejemplo 6.14. En rojo el punto solución.

Para cada uno de los individuos de cada una de las generaciones es posible calcular su función objetivo empleando 6.22. La Figura 6.48 muestra el mejor valor de la función objetivo (es decir, el más pequeño) para cada una de las 20 generaciones que se han obtenido al correr el algoritmo. Puede verse que conforme avanza el algoritmo se va encontrando una mejor solución.

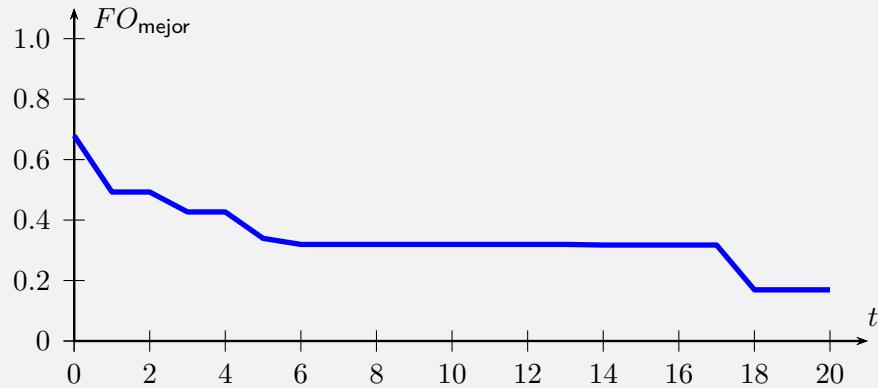


Figura 6.48: Evolución de la optimización en el ejemplo 6.14

Ejemplo 6.15 (Aleatoriedad de los algoritmos genéticos) Los algoritmos genéticos utilizan decisiones aleatorias en varias etapas del proceso. Por esa razón, dos ejecuciones del mismo algoritmo para resolver el mismo problema pueden arrojar soluciones diferentes. Para ilustrar este hecho, se ha corrido 100 veces un algoritmo genético para resolver el mismo problema del ejemplo 6.14, con las siguientes condiciones:

- Espacio de búsqueda: $[0, 6] \times [0, 6]$
- Número máximo de generaciones: 100
- Número de individuos en cada generación: 31
- Tipo de codificación genética: codificación real

Pese a que se han empleado las mismas condiciones en las 100 ejecuciones del algoritmo, la mejor solución obtenida en cada una de ellas ha sido diferente. Con las 100 soluciones obtenidas se ha trazado construido el histograma que se muestra en la Figura 6.49

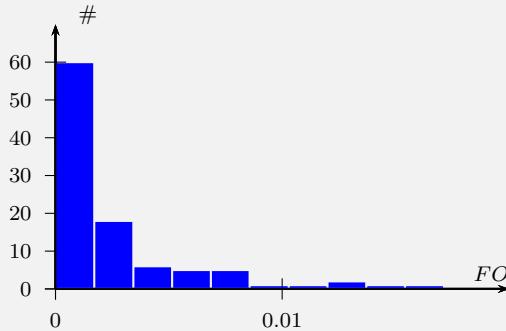


Figura 6.49: Histograma del ejemplo 6.15

6.6.2. Optimización de SBR con algoritmos genéticos

El diseño de un SBR implica el diseño de sus componentes, cada uno de los cuales tiene varios parámetros. Por ejemplo, es necesario definir para cada variable lingüística el número de etiquetas, el tipo de conjunto difuso para cada etiqueta y los parámetros que definen la forma de cada conjunto; también es necesario definir el número de reglas y para cada regla las etiquetas en los antecedentes y consecuentes así como los posibles modificadores lingüísticos; también es necesario definir el tipo de concesor, el tipo de difusor y los parámetros que definen su forma; no debemos olvidar que cada SBR tiene operadores AND, de Implicación, de Composición y de Agregación de Reglas, cada uno de los cuales debe seleccionarse y, si resultan ser paramétricos, escoger dicho parámetro.

La consecuencia de lo anterior es que aún un sistema pequeño como puede ser uno SISO de tres etiquetas puede tener más de 40 parámetros internos. A diseñar un SBR es usual utilizar la definición por defecto de muchos de ellos, lo que simplifica enormemente la tarea. Esta decisión, sin embargo, deja abierta una pregunta: ¿cuál es la mejor selección de parámetros para una aplicación dada?

Supongamos que necesitamos que el SBR aproxime una función y que para ello tenemos una colección de ejemplos de cómo se comporta esa función. Es una situación semejante a la descrita en la sección 6.5 e ilustrada en la Figura 6.27: tenemos una colección de r patrones de entrenamiento y necesitamos construir un SBR que reproduzca ese comportamiento.

Supongamos que tenemos un SBR tipo MISO con p entradas y denotemos por $SBR(\mathbf{x})$ la salida que se obtiene de ese SBR cuando se aplica el vector \mathbf{x} a las entradas. Denotemos

ahora el i -ésimo patrón de entrenamiento por $\mathbf{x}^{(i)}$, es decir:

$$\mathbf{x}^{(i)} = \begin{bmatrix} x_1^{(i)} & x_2^{(i)} & \cdots & x_p^{(i)} \end{bmatrix} \quad (6.23)$$

Si se aplica el i -ésimo patrón de entrenamiento a la entrada del SBR la salida será $SBR(\mathbf{x}^{(i)})$. Como queremos que el SBR reproduzca los patrones de entrenamiento, queremos que esa salida sea igual a $y^{(i)}$. Por eso, podemos definir el error cometido por el SBR al reproducir el i -ésimo patrón de entrenamiento como

$$\varepsilon^{(i)} = SBR(\mathbf{x}^{(i)}) - y^{(i)} \quad (6.24)$$

Para considerar los errores cometidos al reproducir todos los patrones, calculamos el error promedio total así:

$$\varepsilon = \frac{1}{r} \sqrt{\sum_{i=1}^r (\varepsilon^{(i)})^2} \quad (6.25)$$

es decir

$$\varepsilon = \frac{1}{r} \sqrt{\sum_{i=1}^r (SBR(\mathbf{x}^{(i)}) - y^{(i)})^2} \quad (6.26)$$

Ahora bien, hemos dicho antes que el diseño de un SBR implica seleccionar muchos parámetros. Supongamos que hemos decidido dejar fijos algunos de esos parámetros y otros no. Definimos θ como el vector de los parámetros del SBR que podemos cambiar:

$$\theta = [\theta_1 \ \theta_2 \ \cdots \ \theta_z] \quad (6.27)$$

Claramente, la salida del SBR dependerá de los valores de los parámetros θ . Para indicar eso, modificamos la notación: $SBR(\mathbf{x}; \theta)$ la salida que se obtiene del SBR con parámetros θ cuando se aplica el vector \mathbf{x} a las entradas.

Con esa nueva notación, el error total al reproducir los patrones de comportamiento resulta ser

$$\varepsilon(\theta) = \frac{1}{r} \sqrt{\sum_{i=1}^r (SBR(\mathbf{x}^{(i)}; \theta) - y^{(i)})^2} \quad (6.28)$$

El mejor SBR será aquél que minimice el error calculado usando la ecuación 6.28. Podemos usar una heurística para buscar los mejores valores de θ , es decir, aquellos que minimicen $\varepsilon(\theta)$. La Figura 6.50 ilustra esta estrategia: el algoritmo de mejoramiento de la heurística buscará minimizar la función objetivo $FO(\theta)$ que es, precisamente, el error

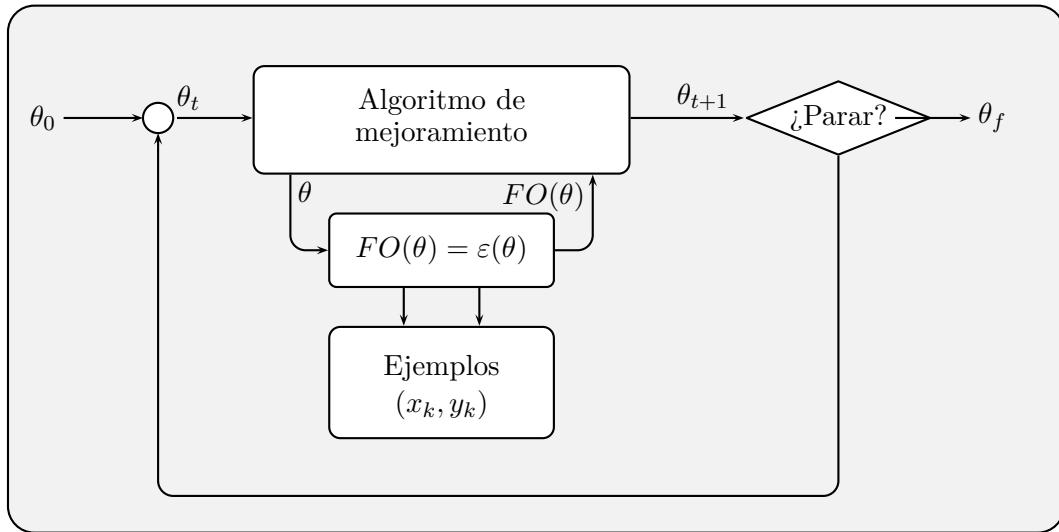


Figura 6.50: Heurísticas aplicadas al diseño de SBR

total. Para calcular $FO(\theta)$ es necesario utilizar la información de todos los patrones de entrenamiento.

La heurística que vamos a usar en las secciones 6.6.2.1 y 6.6.2.2 es la de algoritmos genéticos. La diferencia en las dos secciones es el conjunto de parámetros θ que se modifican.

6.6.2.1. Optimización de modificadores lingüísticos

El comportamiento de un SBR puede modificarse a través de modificadores lingüísticos (Véase la sección 5.1 y el ejemplo 5.1). Un modificador lingüístico es un exponente positivo que se aplica a la función de pertenencia de cada antecedente de cada regla. Un SBR con p entradas y n reglas puede tener hasta np modificadores.

Denotemos por $\varsigma_i^{(k)}$ al modificador lingüístico del i -ésimo antecedente de la k -ésima regla. Para encontrar los mejores valores de todos los $\varsigma_i^{(k)}$ definimos el vector θ como el vector que contiene los np modificadores:

$$\theta = [\varsigma_1^{(1)} \ \dots \ \varsigma_p^{(1)} \ \varsigma_1^{(2)} \ \dots \ \varsigma_p^{(2)} \ \dots \ \varsigma_1^{(n)} \ \dots \ \varsigma_p^{(n)}] \quad (6.29)$$

Para optimizar el diseño de un SBR podemos aplicar una heurística que encuentre el valor de este vector θ que minimice la función objetivo 6.28. Los ejemplos 6.16 y 6.17 ilustran el uso de AG con este propósito.

Ejemplo 6.16 (Optimización de modificadores lingüísticos mediante AG)

Retomemos el ejemplo 6.10. En ese ejemplo se utiliza el entrenamiento de universos fijos para obtener un SBR que aproxima la función de la ecuación 6.18. El entrenamiento logra identificar un conjunto de 5 reglas que se muestran en la Tabla 6.4. Además, en la Figura 6.31 se compara la función real con el modelo obtenido, lo que permite valorar visualmente qué tan buena es la aproximación.

Nos proponemos ahora optimizar el SBR obtenido, es decir, lograr que aproxime mejor la función original. Para ello, se decide aplicar un modificador lingüístico a los antecedentes de cada una de las reglas. Además, para determinar el valor que debe tener cada uno de esos modificadores se decide utilizar un algoritmo genético.

La ecuación 6.30 muestra la función objetivo que debe minimizar el algoritmo genético. Esta función mide el error promedio que se comete al utilizar $SBR(x_i; \theta)$ para aproximar los r patrones de entrenamiento.

$$FO(\theta) = \frac{1}{r} \sqrt{\sum_{i=1}^r (SBR(\mathbf{x}^{(i)}; \theta) - y^{(i)})^2} \quad (6.30)$$

Por su parte, $SBR(x_i; \theta)$ es el SBR que se desea optimizar. θ es el vector que contiene los parámetros del SBR que se van a ajustar, es decir, los modificadores lingüísticos. Como la base de reglas tiene 5 antecedentes, se utilizarán 5 modificadores lingüísticos y por tanto θ es un vector de 5 números reales. Observando la Figura 5.1 se decide que cada modificador debe estar en el intervalo $[0.1, 4.0]$, es decir:

$$\theta = [\varsigma_1^{(1)} \quad \varsigma_1^{(2)} \quad \varsigma_1^{(3)} \quad \varsigma_1^{(4)} \quad \varsigma_1^{(5)}] \in [0.1, 4.0]^5 \quad (6.31)$$

Con la información anterior, se diseña un algoritmo genético con las siguientes características:

- Función a optimizar: ecuación 6.30
- Espacio de búsqueda: $[0.1, 4.0]^5$
- Número máximo de generaciones: 50
- Número de individuos en cada generación: 11
- Tipo de codificación genética: codificación real

La Figura 6.51 muestra los resultados de la ejecución del algoritmo. La gráfica de la izquierda muestra en color azul la función que se desea aproximar, en color negro la aproximación del

SBR antes de ejecutar el algoritmo y en color rojo la mejor aproximación obtenida gracias al algoritmo.

Puede observarse que en algunos trozos de la curva la aproximación roja se acerca más que la aproximación negra a la función azul. Es decir, se observa que el algoritmo genético mejoró la aproximación.

La gráfica de la derecha, por su parte, muestra la evolución de la mejor función objetivo en cada generación a lo largo de la ejecución del algoritmo. El comportamiento descendente de esa curva es la evidencia cuantitativa de que el algoritmo logra mejorar el desempeño del SBR original.

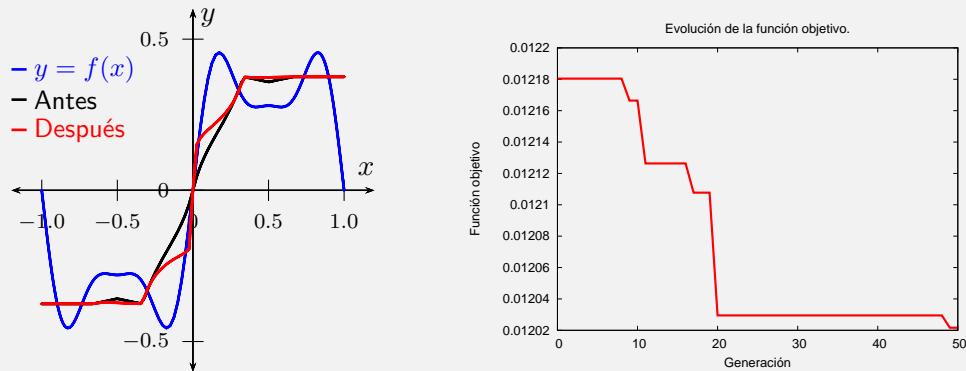


Figura 6.51: Optimización y función objetivo en el ejemplo 6.16

La Tabla 6.7 muestra el valor de los genes del mejor individuo obtenido, es decir, muestra los cinco modificadores lingüísticos que hay que aplicar a los antecedentes de la base de reglas. Estos valores, no obstante, corresponden a una única ejecución del algoritmo; debido a la naturaleza aleatoria del mismo, otra ejecución seguramente arrojará otros modificadores.

Tabla 6.7: Mejor individuo en el ejemplo 6.16

Gen	Valor
1	0.910531
2	0.353713
3	0.091039
4	0.673560
5	0.159540

Ejemplo 6.17 (Optimizaciones de modificadores lingüísticos mediante AG)

Las mismas ideas del ejemplo 6.16 pueden aplicarse a SBR construidos con otro número de reglas. En este ejemplo mostramos el resultado de hacerlo sobre los SBR del ejemplo 6.11 que emplean 3, 5, 5, 9, 11, 13 y 15 reglas, y cuyo desempeño se puede ver en la Figura 6.32.

Las Figuras 6.52 a 6.58 muestran, para cada caso, las aproximaciones y la evolución de la función objetivo.

Nótese que al aumentar el número de reglas las curvas negra y roja son cada vez más parecidas, es decir, la modificación que logra hacer el AG es más pequeña. Esto se debe a que al usar más reglas el SBR original aproxima mejor la función azul y por tanto, cada vez es más difícil encontrar un mejor desempeño. No obstante, las gráficas de la derecha son todas descendentes, lo que demuestra que el AG logra alguna mejora.

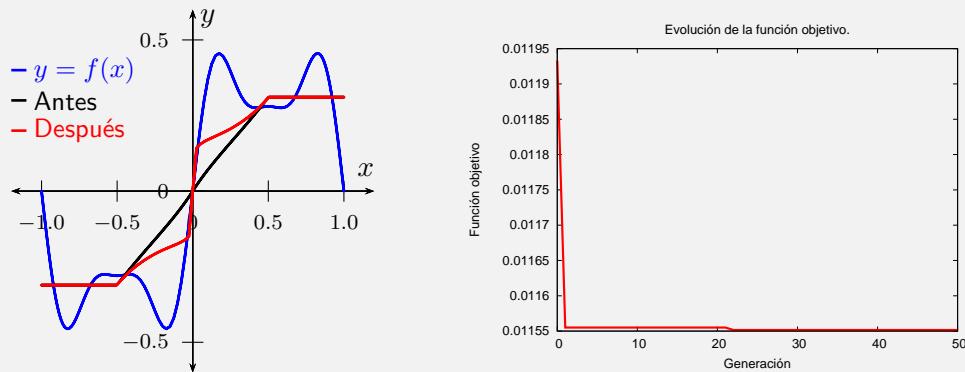


Figura 6.52: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 3 conjuntos

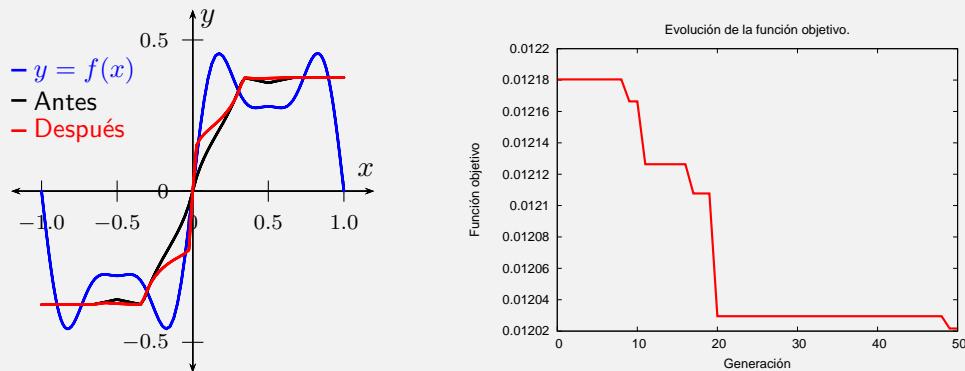


Figura 6.53: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 5 conjuntos

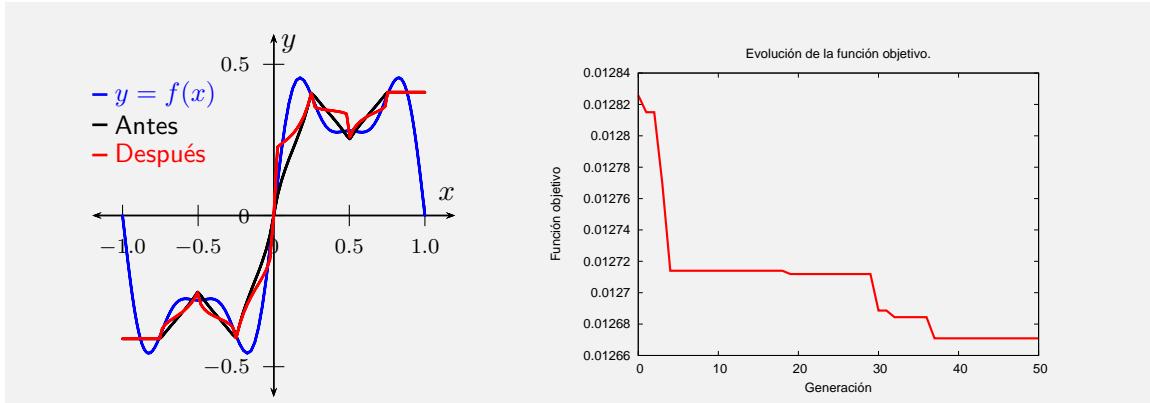


Figura 6.54: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 7 conjuntos

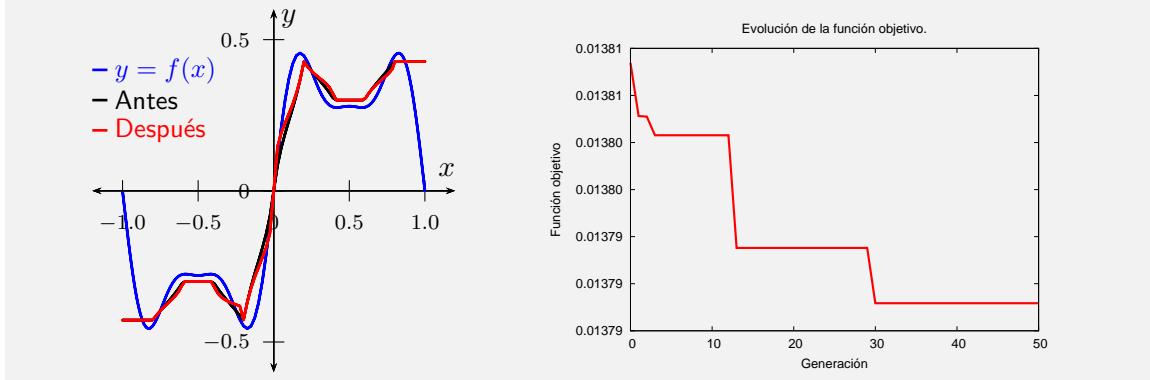


Figura 6.55: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 9 conjuntos

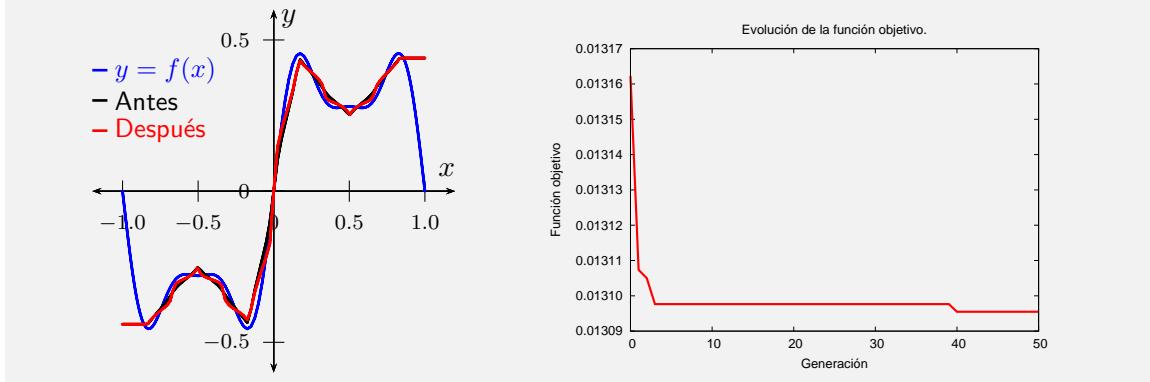


Figura 6.56: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 11 conjuntos

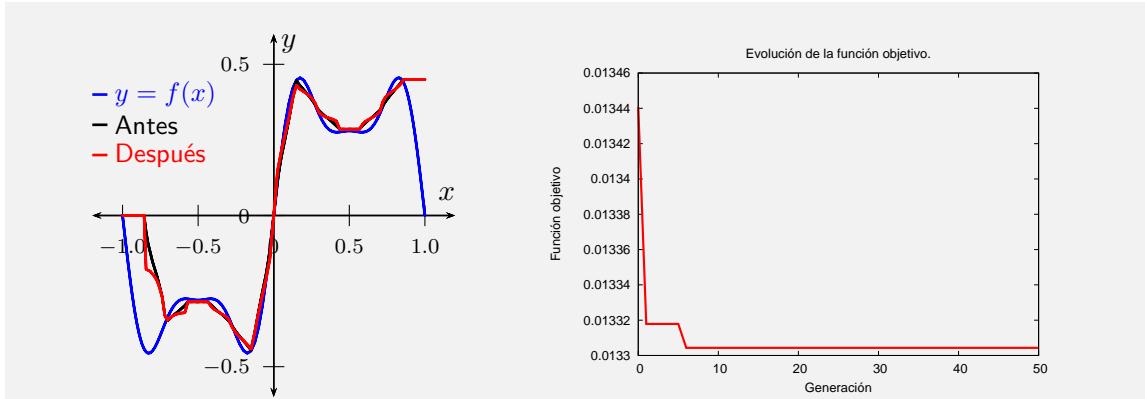


Figura 6.57: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 13 conjuntos

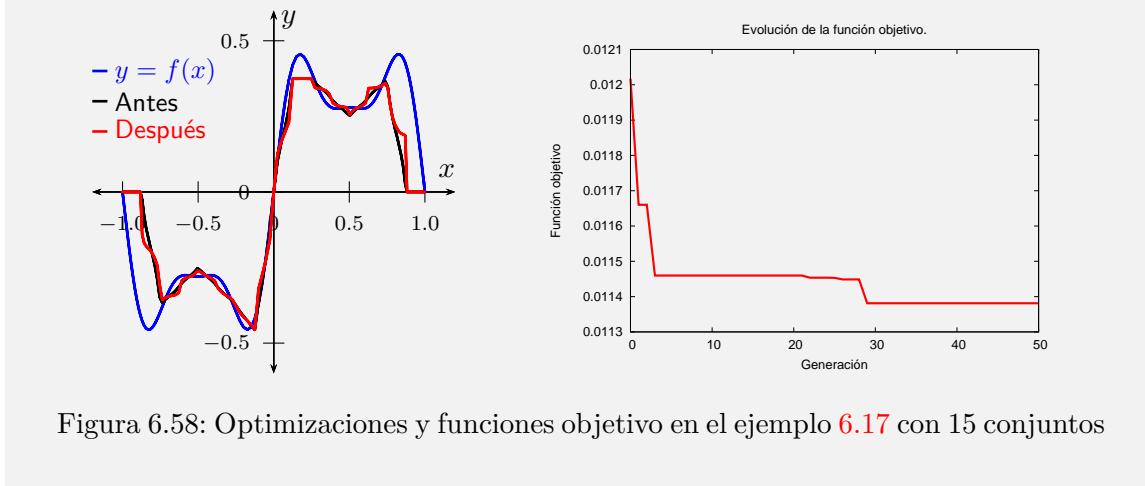


Figura 6.58: Optimizaciones y funciones objetivo en el ejemplo 6.17 con 15 conjuntos

6.6.2.2. Optimización de las variables lingüísticas

El funcionamiento de un SBR depende fuertemente de la función semántica de las variables lingüísticas, es decir, depende fuertemente de la forma de los conjuntos difusos de cada una de las variables de entrada y de salida. Por ello, resulta lógico intentar optimizar un SBR modificando la forma de esos conjuntos.

Esta modificación, sin embargo, no puede hacerse de cualquier forma, porque es posible llegar a definiciones absurdas o inconvenientes de la función semántica.

Para explicar este hecho, supóngase que se desea optimizar la definición de los conjuntos de la variable lingüística *Estatura* que aparece en la figura 6.59. Al modificar la forma de estos conjuntos pueden darse los siguientes casos no deseados:

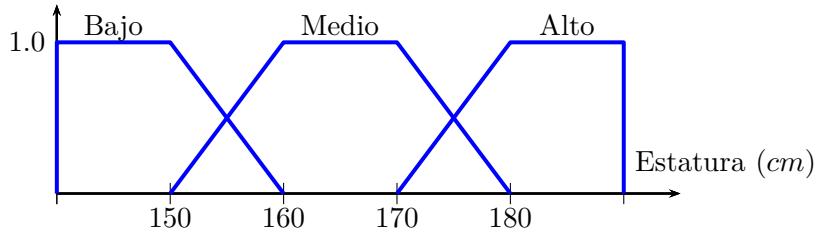


Figura 6.59: variable ligüística ‘Estatura’

Zonas indefinidas: si una región del espacio de entrada no tiene ninguna etiqueta que la califique, no estará cubierta por ninguna regla y por lo tanto el modelo no podrá simular el comportamiento del sistema en esta región (Ver Figura 6.60a).

Conjuntos subsumidos: si un conjunto queda contenido dentro de otro, está región quedará doblemente etiquetada, y existirán reglas igualmente válidas (y posiblemente contradictorias) para la región (Ver Figura 6.60b).

Etiquetas ilógicas: es posible que la Base de Reglas pierda todo su significado si las etiquetas lingüísticas cambian de orden, por ejemplo *Medio* mayor que *Alto* (Ver Figura 6.60c).

Por lo anterior es necesario restringir la forma que los conjuntos difusos puedan tomar y asegurar que la función semántica sea una partición del universo de discurso. Una posibilidad para lograrlo consiste en definir la forma de los conjuntos a partir de unos parámetros, según se explica en el siguiente procedimiento que se refiere a la figura 6.61:

Sea E el número de conjuntos a construir. El procedimiento para definir tales conjuntos consiste en:

1. Dividir el universo de discurso en $N = 2E - 1$ intervalos de igual tamaño. Cada intervalo tendrá un ancho Δ , y estarán limitado por los puntos p_1, p_2, \dots, p_{N-2} .
2. Definir la forma de los conjuntos difusos como trapecios en forma tal que los vértices c, d de un conjunto coincidan con los vértices a, b del conjunto siguiente.
3. Denominar los puntos en donde aparecen los vértices de los trapecios como v_1, v_2, \dots, v_{N-2} .
4. Definir $\delta_i = v_i - p_i$
5. Restringir los valores posibles de δ_i al intervalo $[-\frac{\Delta}{2}, \frac{\Delta}{2}]$

Esta estrategia no sólo asegura la existencia de una partición correcta, sino que además parametriza la forma de los conjuntos en función de un número pequeño de parámetros $\delta_0, \delta_1, \dots, \delta_{N-2}$, que pueden ser optimizados. Para ello, definimos θ como

$$\theta = [\delta_1 \quad \delta_2 \quad \dots \quad \delta_{N-2}] \quad (6.32)$$

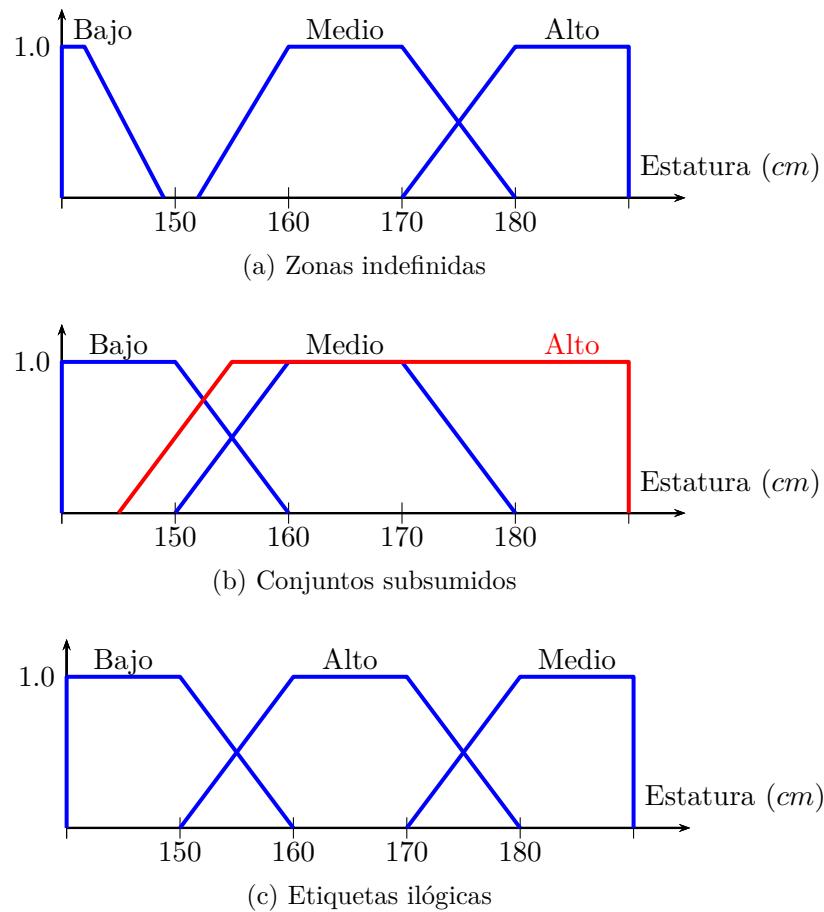


Figura 6.60: Variables sin sentido lingüístico

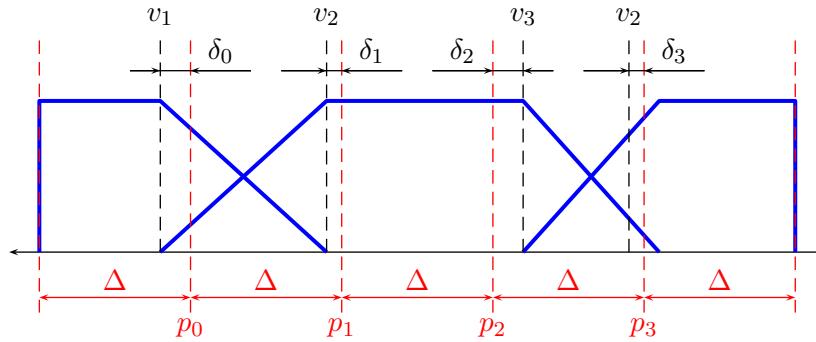


Figura 6.61: Partición parametrizada

Podemos ahora aplicar una heurística que encuentre el valor de este vector θ que minimice la función objetivo 6.28. Los ejemplos 6.18 y 6.19 ilustran el uso de AG con este propósito.

Ejemplo 6.18 (Optimización de la variable de entrada mediante AG)

Retomemos el ejemplo 6.10 en el que se utiliza el entrenamiento de universos fijos para obtener un SBR que aproxima la función de la ecuación 6.18. En este ejemplo se utiliza un AG para optimizar la función semántica de la variable de entrada.

El AG se ha utilizado para optimizar un SBR de 3, 5, 7, 9, 11 y 13 etiquetas. En cada caso se ha empleado el entrenamiento de universos fijos para identificar la base de reglas y posteriormente se ha optimizado la función semántica de entrada.

Las Figuras 6.62 a 6.67 muestran el desempeño del SBR antes y después de la optimización y la forma en que evolucionó la función objetivo.

Por su parte, la Figura 6.68 muestra las funciones semánticas antes y después de la optimización.

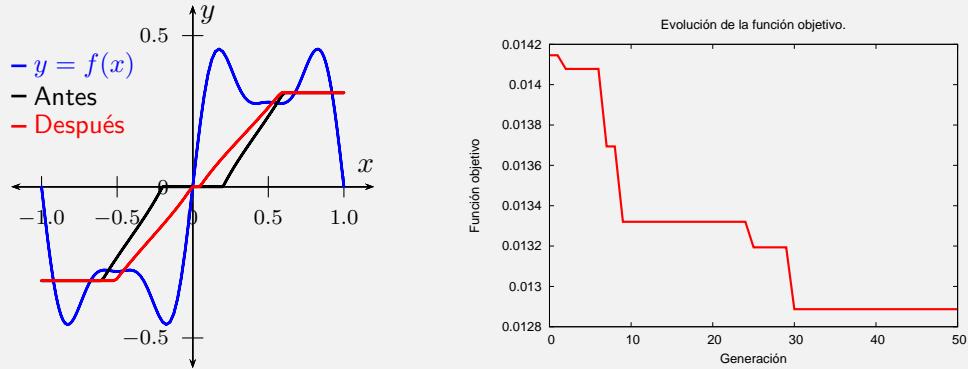


Figura 6.62: Optimizaciones y funciones objetivo en el ejemplo 6.18 con 3 conjuntos

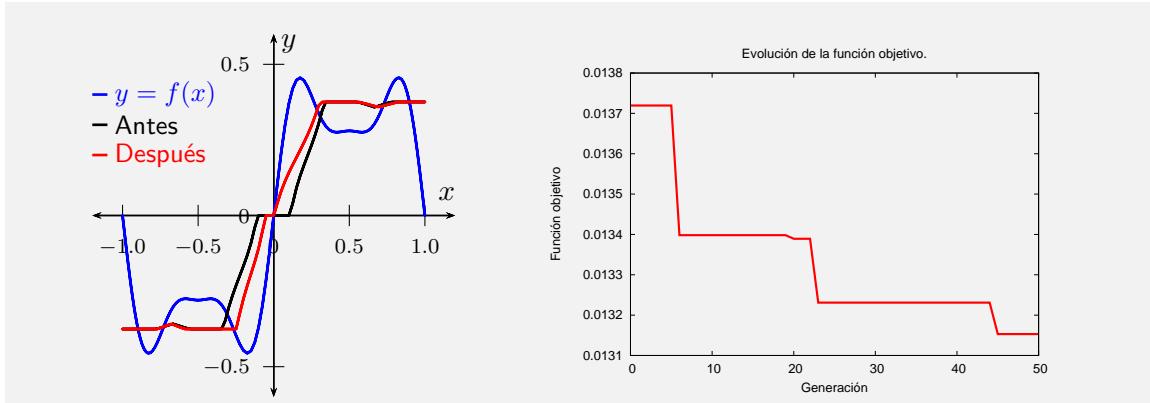


Figura 6.63: Optimizaciones y funciones objetivo en el ejemplo 6.18 con 5 conjuntos

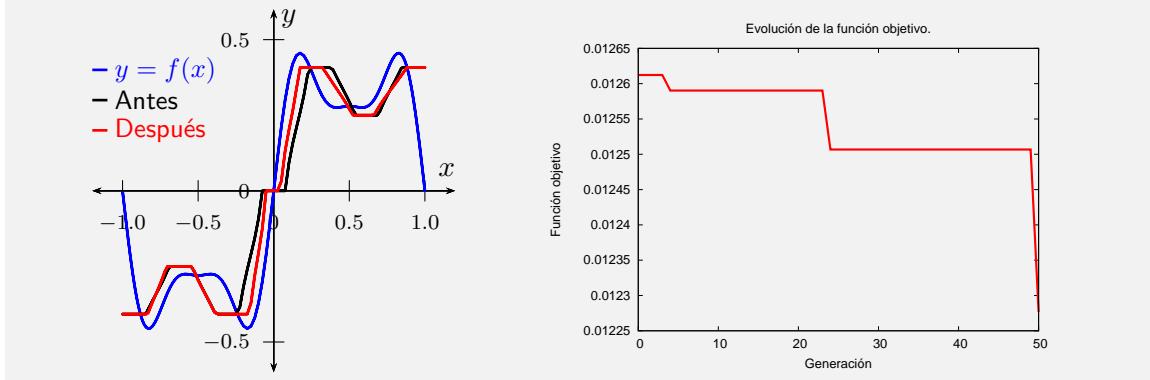


Figura 6.64: Optimizaciones y funciones objetivo en el ejemplo 6.18 con 7 conjuntos

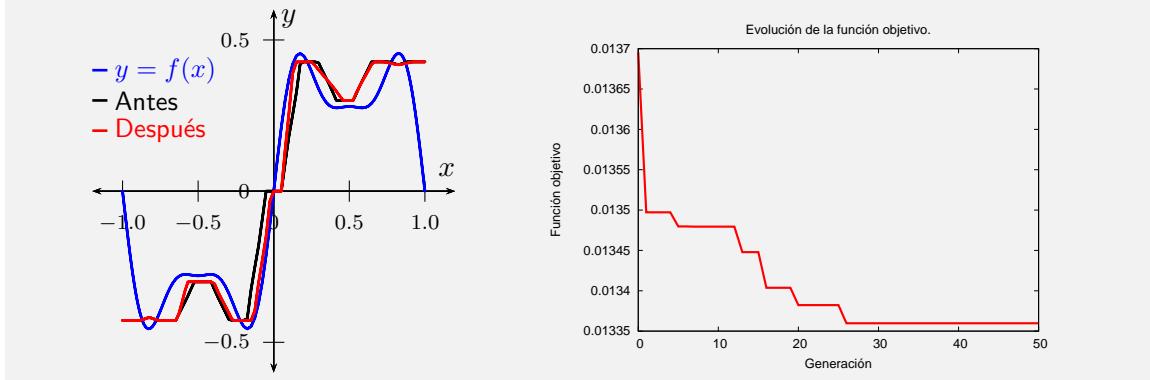


Figura 6.65: Optimizaciones y funciones objetivo en el ejemplo 6.18 con 9 conjuntos

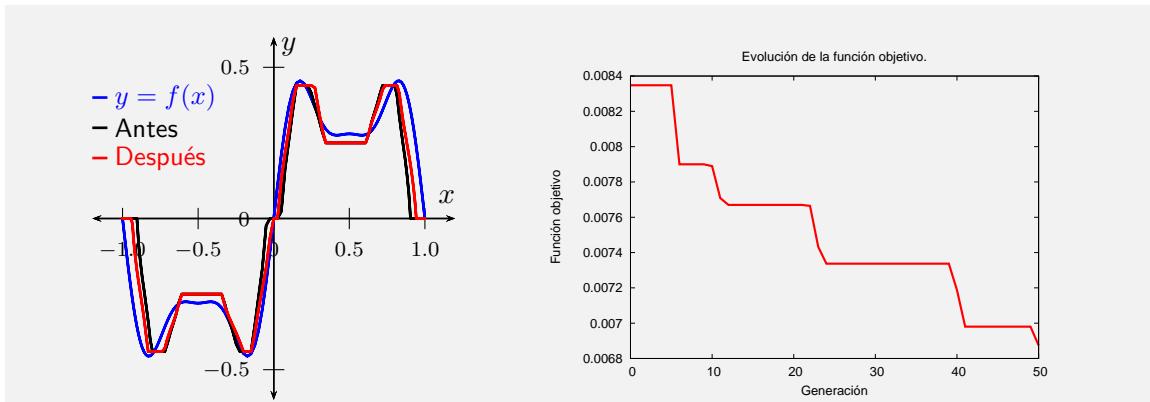


Figura 6.66: Optimizaciones y funciones objetivo en el ejemplo 6.18 con 11 conjuntos

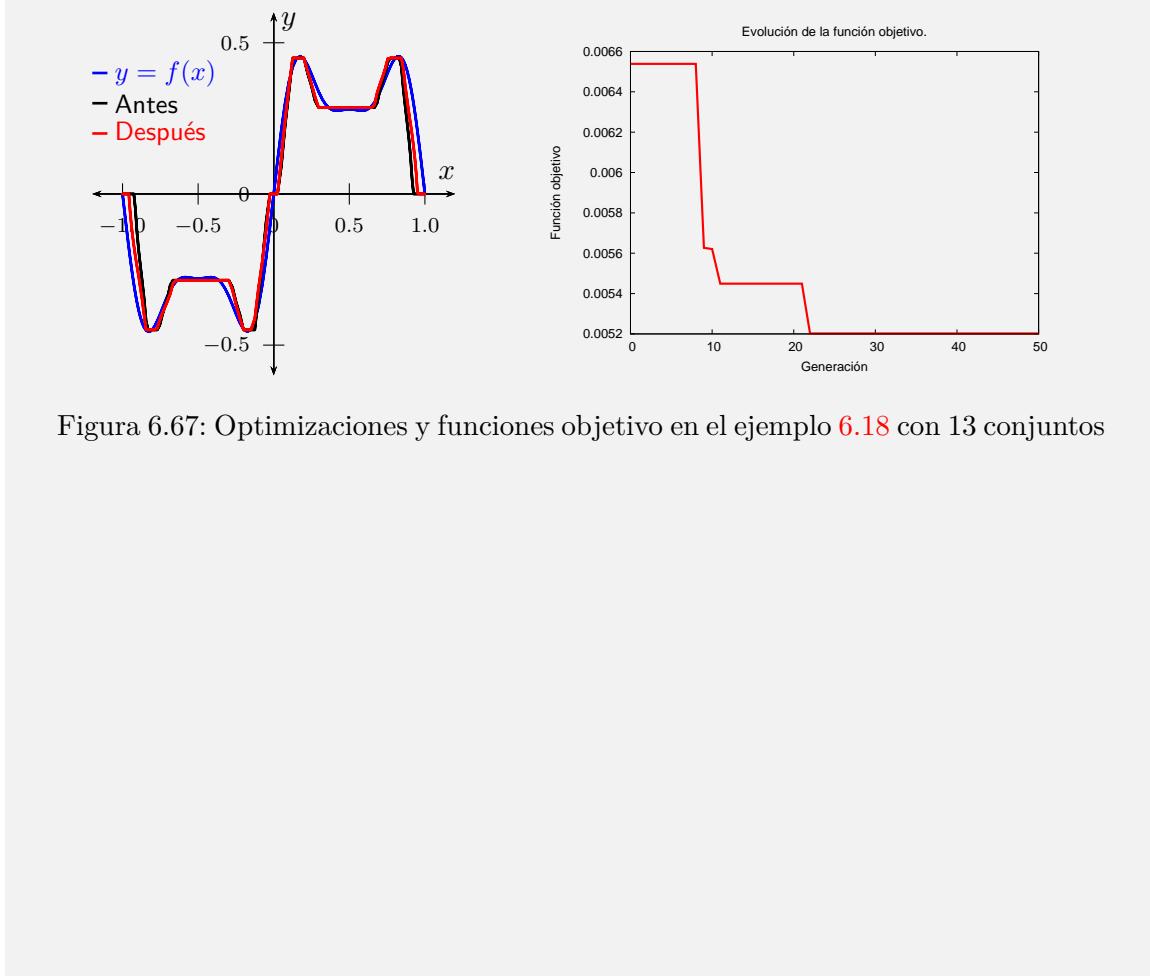


Figura 6.67: Optimizaciones y funciones objetivo en el ejemplo 6.18 con 13 conjuntos

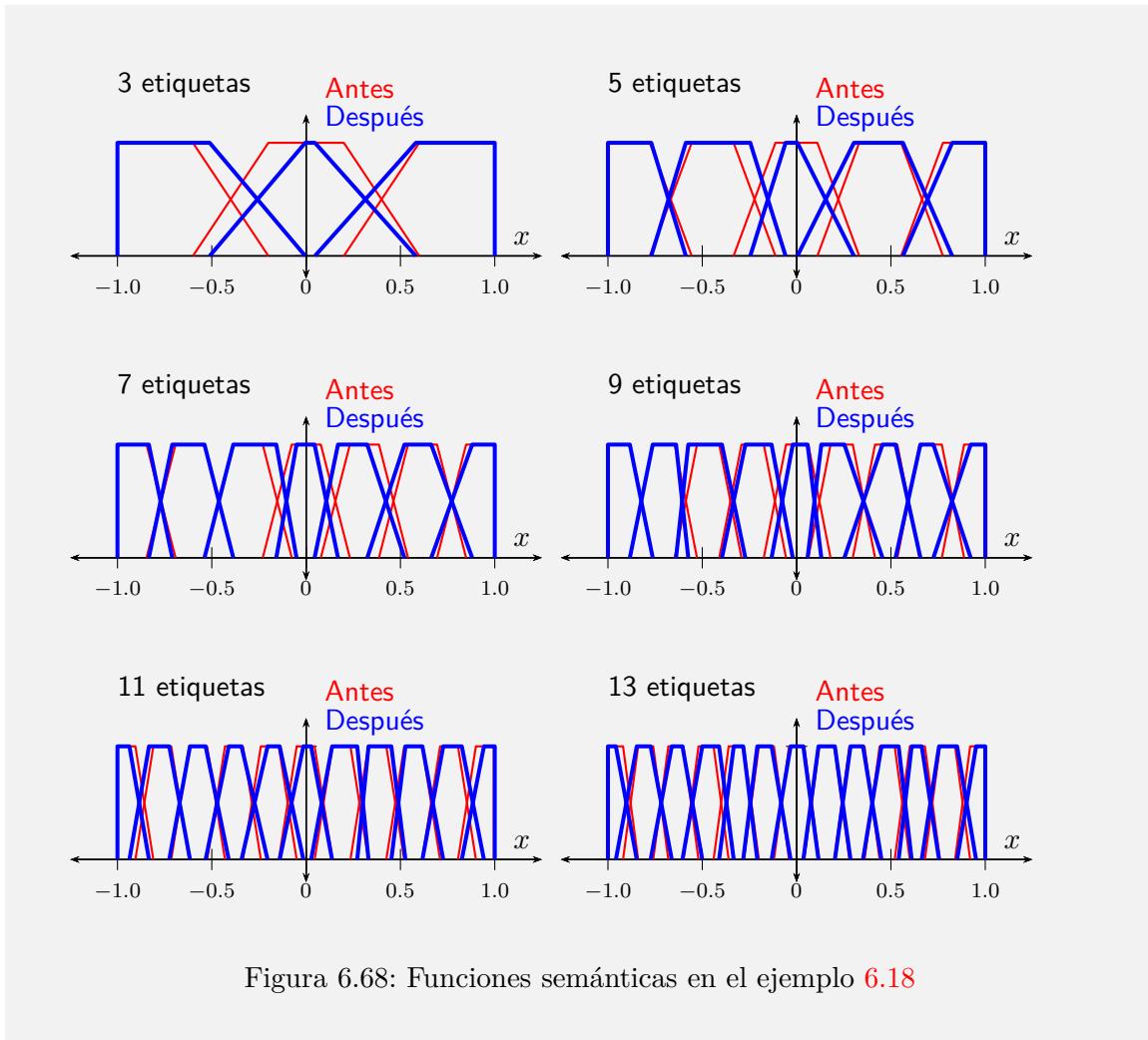


Figura 6.68: Funciones semánticas en el ejemplo 6.18

Ejemplo 6.19 (Optimización de dos variables mediante AG) Continuando con el ejemplo 6.18, en este ejemplo utilizamos un AG para optimizar simultáneamente la función semántica de las variables de entrada y de salida. En este ejemplo también se ha empleado el AG para un SBR de 3, 5, 7, 9, 11 y 13 etiquetas; al igual que en el referido ejemplo, en este también se ha empleado en cada caso el entrenamiento de universos fijos para identificar la base de reglas y posteriormente se han optimizado las funciones semánticas de entrada y de salida.

Las Figuras 6.69 a 6.74 muestran el desempeño del SBR antes y después de la optimización y la forma en que evolucionó la función objetivo.

Por su parte, la Figura 6.75 muestra las funciones semánticas antes y después de la optimización.

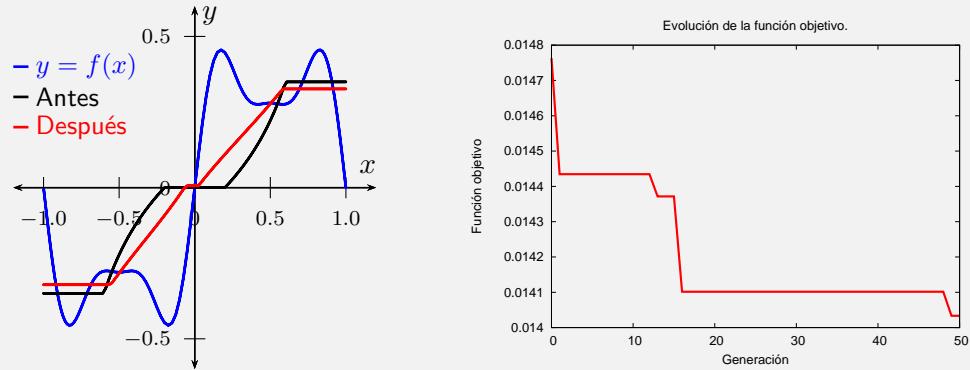


Figura 6.69: Optimizaciones y funciones objetivo en el ejemplo 6.19 con 3 conjuntos

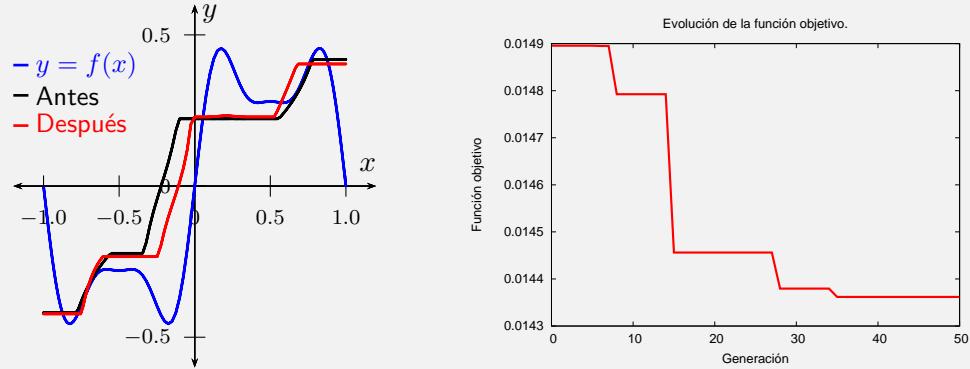


Figura 6.70: Optimizaciones y funciones objetivo en el ejemplo 6.19 con 5 conjuntos

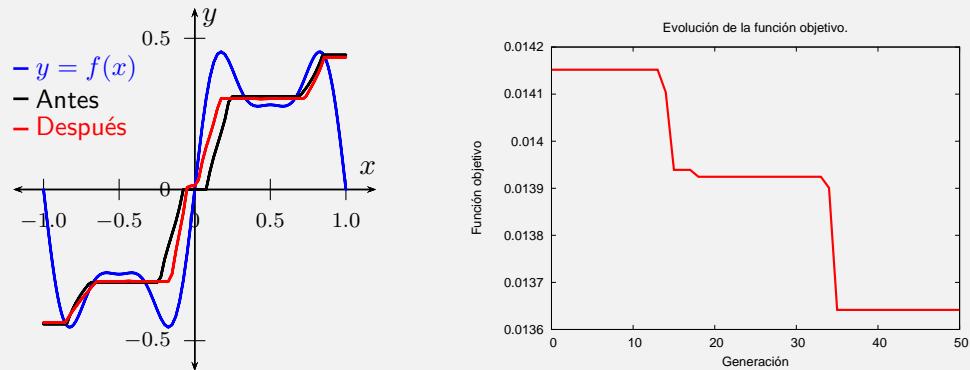


Figura 6.71: Optimizaciones y funciones objetivo en el ejemplo 6.19 con 7 conjuntos

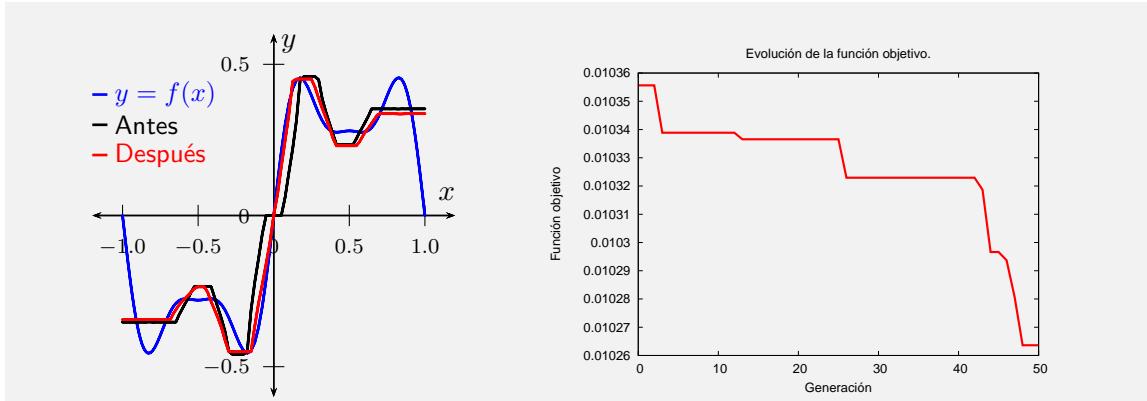


Figura 6.72: Optimizaciones y funciones objetivo en el ejemplo 6.19 con 9 conjuntos

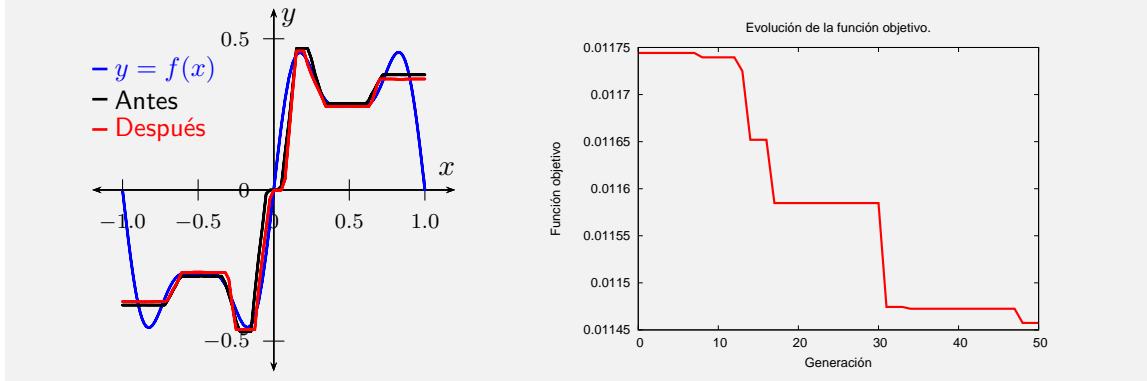


Figura 6.73: Optimizaciones y funciones objetivo en el ejemplo 6.19 con 11 conjuntos

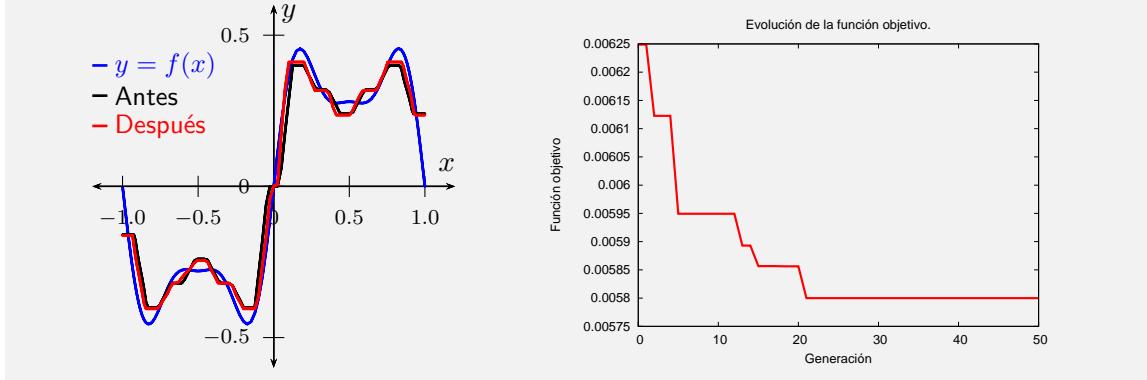


Figura 6.74: Optimizaciones y funciones objetivo en el ejemplo 6.19 con 13 conjuntos

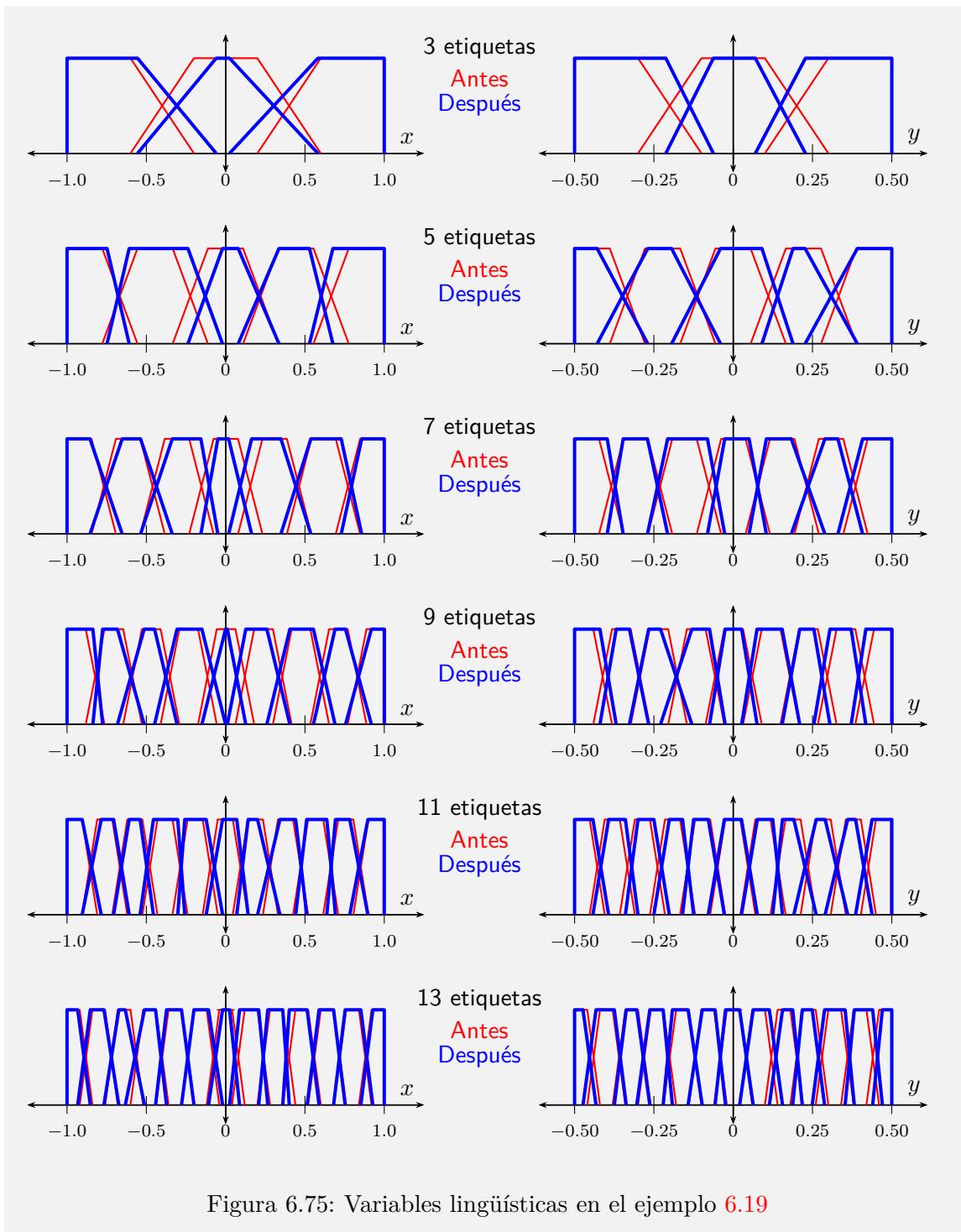
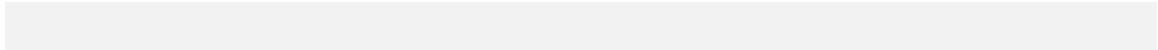


Figura 6.75: Variables lingüísticas en el ejemplo 6.19



Parte II

Aritmética difusa

Capítulo 7

Números difusos

El estudio de la *aritmética difusa* debe iniciar con el concepto de *número difuso*, y ese es el propósito de este capítulo. La definición básica y los elementos mínimos para construir números difusos se presentan en la sección 7.1.

En las secciones 7.2 y 7.3 se presentan dos propiedades numéricas que sirven para caracterizar los números difusos: el valor representativo y la ambigüedad. Una de las estrategia para determinar qué tan parecidos son dos números difusos es la de determinar la distancia entre ello, concepto que se presenta en la sección 7.4.

Un tipo especial de números difusos son los trapezoidales, denominados así porque su función de pertenencia tiene forma de trapecio. En la sección 7.6 se estudian sus propiedades, mientras que en la sección 7.7 se presentan varias estrategias para aproximar un número que no es trapezoidal a uno que sí lo es.

7.1. Números difusos

En esta sección se presenta el concepto de *número difuso*. Se trata de un tipo de conjunto difuso que permite representar cantidades que tienen incertidumbre o que están definidas de forma vaga y de esta forma dar sentido matemático a expresiones tales como ‘llegaré cerca de las 10 pm’, ‘cuesta más o menos 100 dólares’ o ‘mide aproximadamente 2 hectáreas’. A partir de el concepto de número difuso se puede definir también un conjunto de operaciones aritméticas básicas (capítulo 10).

Hay dos formas equivalentes de definir los números difusos ([8]):

- Un número difuso es un conjunto difuso sobre \mathbb{R} , normal, convexo y semicontínuo superiormente.

- Un número difuso es un conjunto difuso sobre \mathbb{R} cuyos α -cortes son intervalos cerrados.

La definición de α -corte se encuentra en la sección 1.3.3, específicamente en las ecuaciones 1.4 y 1.5. Además, los ejemplos 1.6, 1.7 y 1.8 ayudan a ilustrar la explicación del concepto de número difuso que sigue a continuación.

La primera de las definiciones establece cuatro condiciones para que un conjunto difuso sea un número difuso:

1. Debe estar definido sobre \mathbb{R} . La Figura 7.1a muestra un ejemplo de número difuso en el que puede verse que el universo de discurso es la recta real. En las implementaciones prácticas de software, el universo de discurso suele limitarse a un intervalo en la recta; esto debe interpretarse asumiendo que por fuera de ese intervalo el grado de pertenencia del conjunto es cero.

La misma figura sirve para ilustrar la segunda definición: se ha construido allí un α -corte que resulta ser un intervalo cerrado. Al variar el valor de α entre 0 y 1 también se obtiene un intervalo cerrado.

2. Debe ser *normal*, es decir, debe existir al menos un valor de x en el que el grado de pertenencia sea 1:

$$\max_U(\mu(x)) = 1 \quad (7.1)$$

En la Figura 7.1b se muestra un conjunto difuso que no satisface la condición de normalidad y por tanto no es un número difuso. Para construir el α -corte se ha trazado la recta constante de altura α y el corte resultante es el conjunto vacío. En otras palabras, no se ha obtenido un intervalo cerrado.

Algunos autores establecen una diferencia entre el concepto de *número difuso* y el de *intervalo difuso* así: un número difuso tiene un único valor de x en el que $\mu(x) = 1$, mientras que un intervalo difuso tiene un conjunto de valores $x \in [x_a, x_b]$ para los cuales $\mu(x) = 1$. Sin embargo, en este texto hemos optado por una definición de número difuso que abarca ambos casos.

3. Debe ser *convexo*. Un conjunto difuso sobre \mathbb{R} es convexo ([49]) si para todos los valores de a, b tales que $\mu(a) > 0, \mu(b) > 0$, y para todos los valores de $\lambda \in [0, 1]$ se satisface:

$$\mu(\lambda a + (1 - \lambda)b) \geq \lambda\mu(a) + (1 - \lambda)\mu(b) \quad (7.2)$$

La ecuación 7.2 es la definición convencional de convexidad de curvas aplicada a la función de pertenencia $\mu(x)$. Nótese que $(\lambda a + (1 - \lambda)b)$ permite recorrer los valores de x desde a hasta b al variar λ desde 0 hasta 1 y por tanto el término a la izquierda de la desigualdad no es más que la versión parametrizada de la función

de pertencia entre a y b . Por otra parte, el término de la derecha de la desigualdad es el promedio ponderado de los grados de pertenencia o, lo que es igual, la versión parametrizada de una línea recta que pasa por los puntos $(a, \mu(a))$ y $(b, \mu(b))$.

Por lo anterior, podemos interpretar la condición de convexidad de la siguiente forma: $\mu(x)$ es convexa si para cualquier par de puntos a, b tales que $\mu(a) > 0$, $\mu(b) > 0$ la línea recta que une los puntos $(a, \mu(a))$ y $(b, \mu(b))$ no corta la curva $\mu(x)$ entre a y b (puede tocar la curva, pero no cortarla).

La Figura 7.1c muestra un conjunto sobre \mathbb{R} que no satisface la condición de convexidad y por tanto no es un número difuso. Al trazar el α -corte se ha cortado la curva $\mu(x)$ y como consecuencia el α -corte resultante es la unión de dos intervalos cerrados, es decir, no es un intervalo cerrado.

4. Debe ser *semincontínuo superiormente*. Una función $\mu(x)$ es semicontínua superiormente en el punto x si para cualquier $\epsilon > 0$ existe un $\delta > 0$ tal que

$$\mu(y) < \mu(x) + \epsilon \quad (7.3)$$

para todos los $y \in \mathbb{R}$ que $\|y - x\| < \delta$. Para que el conjunto difuso sea semicontínuo superiormente, $\mu(x)$ debe serlo en todo $x \in \mathbb{R}$.

Esta condición es más débil que la condición de continuidad¹. Es decir, algunas funciones no continuas pueden ser semicontínuas superiormente. Las Figuras 7.1d y 7.1e muestran dos funciones de pertenencia que no son continuas. La primera no satisface la condición de semicontinuidad superior y la segunda sí lo hace. Es decir, la primera no es un número difuso y la segunda sí lo es.

Ambas funciones son discontinuas en los mismos puntos. La diferencia entre las dos funciones está en cómo se ha definido esa discontinuidad: en la primera figura la curva superior está abierta en sus extremos (representados por los círculos de fondo blanco) y las curvas inferiores están cerradas en sus extremos (representados por los círculos azules), mientras que en la segunda función es justo al contrario. Por esa razón, el α -corte de la primera función es un intervalo abierto, mientras que en la segunda es un intervalo cerrado.

7.1.1. Números trapezoidales y triangulares

Existen dos tipos de números difusos muy utilizados, los trapezoidales y los triangulares, además de sus versiones suavizadas. La Figura 7.2 muestra la forma de sus funciones

¹De hecho, una de las primeras definiciones de número difuso se encuentra en [25]. Allí se exige como condición la continuidad de $\mu(x)$. No obstante, en [8] se demuestra que la condición de continuidad se puede relajar por la de semicontinuidad superior y por ello ha sido ampliamente adoptada desde ese entonces.

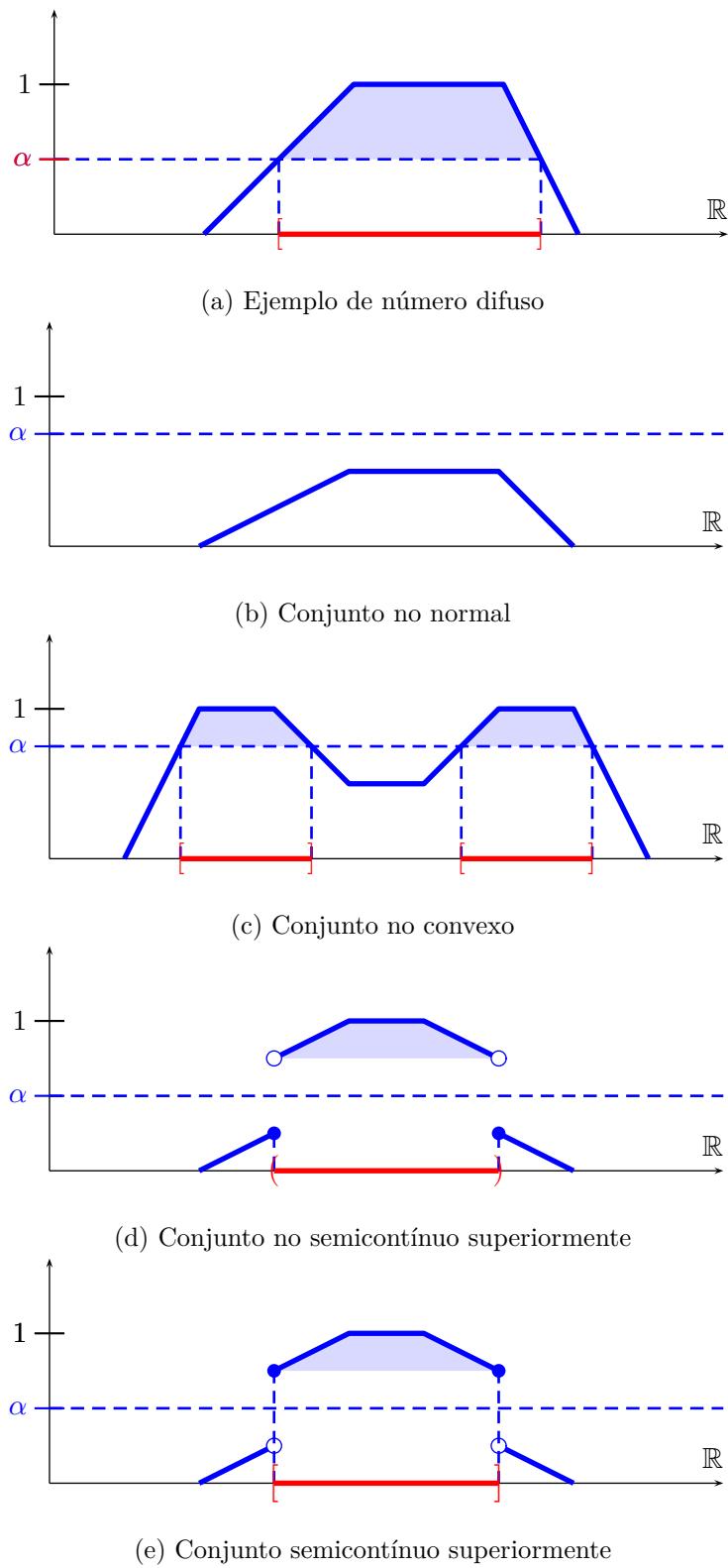


Figura 7.1: Condiciones de un número difuso

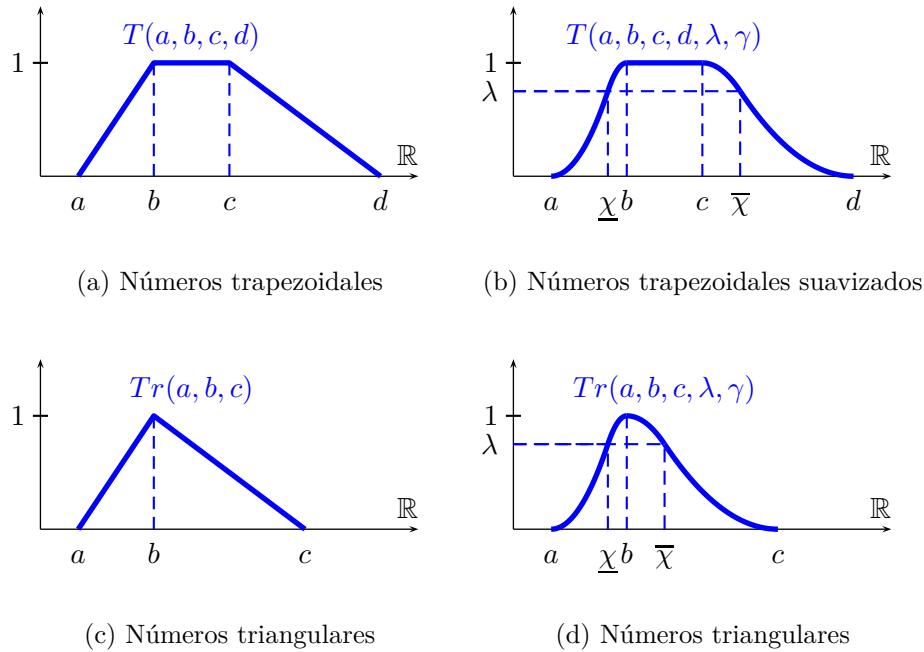


Figura 7.2: Número difusos trapezoidales y triangulares

de pertenencia. En particular, las versiones no suavizadas son muy fáciles de operar, debido a que se construyen con trozos de líneas rectas.

En la figura 7.2a puede verse que para especificar un número difuso trapezoidal basta con especificar sus cuatro vértices. La ecuación 7.4 muestra la expresión matemática de la función de pertenencia.

$$T(a, b, c, d) : \quad \mu(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ 1 & b \leq x < c \\ \frac{x-d}{c-d} & c \leq x < d \\ 0 & d \leq x \end{cases} \quad (7.4)$$

De forma similar, en la figura 7.2c puede verse que para especificar un número difuso triangular basta con especificar sus tres vértices. La ecuación 7.5 muestra la expresión matemática de la función de pertenencia.

$$Tr(a, b, c) : \mu(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \leq x < b \\ \frac{x-c}{b-c} & b \leq x < c \\ 0 & c \leq x \end{cases} \quad (7.5)$$

Utilizando el método de suavización que se muestra en la Figura 3.4 se puede definir un *números trapezoidal suavizados*, cuya función de pertenencia resulta ser:

$$T_c(a, b, c, d, \lambda, \gamma) : \mu(x) = \begin{cases} 0 & x < a \\ \lambda \left(\frac{x-a}{\underline{\chi}-a} \right)^\gamma & a \leq x \leq \underline{\chi} \\ 1 - (1-\lambda) \left(\frac{x-b}{\underline{\chi}-b} \right)^\gamma & \underline{\chi} \leq x \leq b \\ 1 & b \leq x < c \\ 1 - (1-\lambda) \left(\frac{x-c}{\bar{\chi}-c} \right)^\gamma & c \leq x \leq \bar{\chi} \\ \lambda \left(\frac{x-d}{\bar{\chi}-d} \right)^\gamma & \bar{\chi} \leq x \leq d \\ 0 & d \leq x \end{cases} \quad (7.6)$$

donde $\underline{\chi} = a + \lambda(b - a)$ $\bar{\chi} = d + \lambda(c - d)$

A partir de la definición anterior se puede construir un *número triangular suavizado* $Tr_c(a, b, c, \lambda, \gamma) = T_c(a, b, b, c, \lambda, \gamma)$, cuya función de pertenencia resulta ser:

$$Tr_c(a, b, c, \lambda, \gamma) : \mu(x) = \begin{cases} 0 & x < a \\ \lambda \left(\frac{x-a}{\underline{\chi}-a} \right)^\gamma & a \leq x \leq \underline{\chi} \\ 1 - (1-\lambda) \left(\frac{x-b}{\underline{\chi}-b} \right)^\gamma & \underline{\chi} \leq x \leq b \\ 1 - (1-\lambda) \left(\frac{x-b}{\bar{\chi}-b} \right)^\gamma & b \leq x \leq \bar{\chi} \\ \lambda \left(\frac{x-c}{\bar{\chi}-c} \right)^\gamma & \bar{\chi} \leq x \leq c \\ 0 & c \leq x \end{cases} \quad (7.7)$$

donde $\underline{\chi} = a + \lambda(b - a)$ $\bar{\chi} = c + \lambda(b - c)$

7.1.2. α -cortes de número difusos

Supongamos un número difuso A . Debido a que los α -cortes de un número difuso son intervalos cerrados, podemos definir los α -cortes de A de esta forma:

$$A_\alpha = [L_A(\alpha), R_A(\alpha)] \quad (7.8)$$

En donde $L_A(\alpha)$, $R_A(\alpha)$ son unas funciones que calculan los extremos izquierdos y derechos del intervalo en función de α , respectivamente:

$$\begin{aligned} L_A(\alpha) & : [0, 1] \rightarrow \mathbb{R} \\ R_A(\alpha) & : [0, 1] \rightarrow \mathbb{R} \end{aligned} \quad (7.9)$$

Para que A_α sea un intervalo cerrado, estas funciones deben satisfacer ciertas condiciones:

$$\begin{cases} L_A(\alpha) \leq R_A(\alpha) & \forall \alpha \in [0, 1] \\ L_A(\alpha_1) \leq L_A(\alpha_2) & \forall \alpha_1, \alpha_2 \in [0, 1] \quad | \alpha_1 < \alpha_2 \\ R_A(\alpha_1) \geq R_A(\alpha_2) & \forall \alpha_1, \alpha_2 \in [0, 1] \quad | \alpha_1 < \alpha_2 \end{cases} \quad (7.10)$$

La primera condición se explica fácilmente, porque si no se satisface para algún α , el extremo izquierdo sería mayor que el derecho y por tanto ese α -corte no sería un intervalo cerrado. Las otras dos condiciones son necesarias también porque si $\alpha_1 < \alpha_2$ se cumple que $A_{\alpha_2} \subset A_{\alpha_1}$.

Consideremos ahora el número trapezoidal $T(a, b, c, d)$ de la Figura 7.2. Este número está construido con trozos de líneas rectas, cuyas funciones de pertenencia están definidas en la ecuación 7.4. El trozo ascendente corresponde justamente a la función $L(\alpha)$ y el descendente a $R(\alpha)$. Es decir, se satisface que:

$$\begin{cases} \alpha = \frac{L(\alpha)-a}{b-a} \therefore L(\alpha) = a + \alpha(b-a) \\ \alpha = \frac{R(\alpha)-d}{c-d} \therefore R(\alpha) = d - \alpha(d-c) \end{cases} \quad (7.11)$$

Se puede entonces especificar un número trapezoidal $A = T(a, b, c, d)$ a partir de sus α -corte así :

$$T(a, b, c, d) : A_\alpha = [a + \alpha(b-a), d - \alpha(d-c)] \quad (7.12)$$

En la Tabla 7.1 se muestran las funciones $L(\alpha)$ y $R(\alpha)$ para los números trapezoidales, triangulares y sus versiones suavizadas. El procedimiento para obtenerlos es semejante al desarrollado para obtener la ecuación 7.12: a partir de las ecuaciones 7.4, 7.5, 7.6 y 7.7 se han identificado los trozos de líneas ascendentes y descendentes, que corresponden a $L(\alpha)$ y $R(\alpha)$ para de allí despejar sus expresiones. En las versiones suavizadas, además se ha empleado el hecho que en los cambios de curvatura la función de pertenencia vale λ .

Tabla 7.1: α -cortes de los números difusos más usuales

Tipo de conjunto	Ecuación para $\mu(x)$	$L(\alpha)$	$R(\alpha)$
Trapezoidal	7.4	$a + \alpha(b - a)$	$d - \alpha(d - c)$
Triangular	7.5	$a + \alpha(b - a)$	$c - \alpha(c - b)$
Trapezoidal suavizado	7.6	$\begin{cases} a + (\underline{\chi} - a) \left(\frac{\alpha}{\lambda} \right)^{1/\gamma} & \text{si } 0 \leq \alpha \leq \lambda \\ b + (\underline{\chi} - b) \left(\frac{\alpha-1}{\lambda-1} \right)^{1/\gamma} & \text{si } \lambda \leq \alpha \leq 1 \end{cases}$ $\underline{\chi} = a + \lambda(b - a)$	$\begin{cases} d + (\bar{\chi} - d) \left(\frac{\alpha}{\lambda} \right)^{1/\gamma} & \text{si } 0 \leq \alpha \leq \lambda \\ b + (\bar{\chi} - c) \left(\frac{\alpha-1}{\lambda-1} \right)^{1/\gamma} & \text{si } \lambda \leq \alpha \leq 1 \end{cases}$ $\bar{\chi} = d + \lambda(c - d)$
Triangular suavizado	7.7	$\begin{cases} a + (\underline{\chi} - a) \left(\frac{\alpha}{\lambda} \right)^{1/\gamma} & \text{si } 0 \leq \alpha \leq \lambda \\ c + (\underline{\chi} - b) \left(\frac{\alpha-1}{\lambda-1} \right)^{1/\gamma} & \text{si } \lambda \leq \alpha \leq 1 \end{cases}$ $\underline{\chi} = a + \lambda(b - a)$	$\begin{cases} c + (\bar{\chi} - c) \left(\frac{\alpha}{\lambda} \right)^{1/\gamma} & \text{si } 0 \leq \alpha \leq \lambda \\ b + (\bar{\chi} - b) \left(\frac{\alpha-1}{\lambda-1} \right)^{1/\gamma} & \text{si } \lambda \leq \alpha \leq 1 \end{cases}$ $\bar{\chi} = c + \lambda(b - c)$

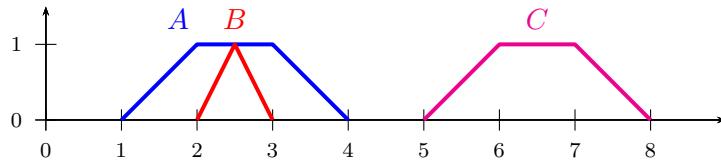


Figura 7.3: Números difusos para ordenar

7.2. Valor representativo y nivel de optimismo

El concepto de valor representativo permite establecer un ordenamiento entre los números difusos, es decir, permite comparar dos números y decidir si uno es mayor que otro. Consideremos los números difusos de la Figura 7.3:

$$\begin{aligned} A &= T(1.0, 2.0, 2.0, 3.0) \\ B &= T(2.0, 2.5, 2.5, 3.0) \\ C &= T(5.0, 6.0, 7.0, 8.0) \end{aligned} \quad (7.13)$$

Claramente el número C está a la derecha de A y de B , y por tanto podemos decir que $C > A$ y que $C > B$. Sin embargo, ¿se puede afirmar algo similar sobre A y B ?

Para responder esa pregunta vamos a utilizar un enfoque inspirado en [15] al que le hemos agregado la noción de *optimismo*. La idea es asignar a cada número difuso un *valor representativo* que ayude a ubicar el número en la recta real. El enfoque consiste en asignar un *valor representativo* a cada α -corte y luego efectuar un promedio ponderado sobre todos ellos. Formalicemos la idea:

Sea A un número difuso. Definimos $v_A(\alpha)$ como un valor real contenido en el α -corte:

$$v_A(\alpha) \in [L_A(\alpha), R_A(\alpha)] \quad (7.14)$$

¿Cuál de todos los valores que están en ese intervalo debemos tomar? Nuestra propuesta consiste en emplear una variable auxiliar que nos ayude a tomar esa decisión. Podemos definir β como un número entre 0 y 1 que nos permita recorrer todo el intervalo. Si $\beta = 0$ tomaremos el extremo inferior; al aumentar β tomaremos un valor cada vez mayor, hasta que al llegar a $\beta = 1$ tomaremos el límite superior:

$$v_A(\alpha, \beta) = (1 - \beta)L_A(\alpha) + \beta R_A(\alpha) \quad \beta \in [0, 1] \quad (7.15)$$

El nombre que le deremos a β es el de *nivel de optimismo*. En efecto, la decisión de qué valor del intervalo tomar para que represente al α -corte es subjetiva. Si suponemos que un valor alto es mejor que uno alto², entonces seleccionar un valor cercano al extremo

²Por ejemplo, si el número representa una ganancia.

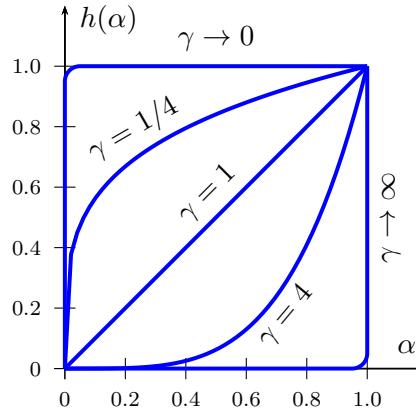


Figura 7.4: Ejemplo de funciones de ponderación $h(\alpha) = \alpha^\gamma$

inferior puede calificarse como una actitud *pesimista* y, por el contrario, seleccionar un valor cercano al extremo superior puede calificarse como una actitud *optimista*³.

Ahora bien, para ponderar el aporte de cada α -corte en la obtención del valor representativo del número, empleamos una función $h(\alpha)$, monótonamente no decreciente tal que $h(0) = 0$ y $h(1) = 1$. El propósito de esta ponderación consiste en darle un mayor peso a la información a la que le corresponde una mayor función de pertenencia. Una de las funciones que se pueden usar para efectuar esta ponderación es:

$$\begin{aligned} h(\alpha) & : [0, 1] \rightarrow [0, 1] \\ h(\alpha) & = \alpha^\gamma \quad \text{con } \gamma \geq 0 \end{aligned} \tag{7.16}$$

En la Figura 7.4 se ha graficado el comportamiento de α^γ para diferentes valores de γ . Si tomamos como referencia $\gamma = 0$ se tendría $h(\alpha) = 1$ y por tanto todos los α -cortes estarían ponderados por el mismo valor. Al incrementar el valor de γ la ponderación cambia: los α -cortes de mayor altura⁴ empiezan a ser ponderados con pesos más grandes que aquellos de menor altura.

Al utilizar $\gamma = 1$ se tiene $h(\alpha) = \alpha$ lo que significa que la ponderación es proporcional a la altura del α -corte. Valores de α mayores que 1 ponderan de manera aún más generosa los α -cortes de mayor altura, en detrimento de los de menor. En el límite ($\alpha \rightarrow \infty$) la situación es equivalente a trabajar solamente con el α -corte superior ($\alpha = 1$).

Ahora, definimos el valor representativo del número difuso A , denotado por V_A como

³Por su puesto, si es mejor un valor bajo que uno alto, las etiquetas de *pesimista* y *optimista* se intercambiarían. Esto puede suceder si el número representa una pérdida o el riesgo de que suceda una pérdida, por ejemplo.

⁴Recuérdese que la altura de cada α -corte es, justamente, α

el promedio de los valores representativos de todos los α -cortes de A , ponderado por la función $h(\alpha)$.

$$\mathbf{V}_A(\beta) = \frac{\int_0^1 h(\alpha)v_A(\alpha, \beta)d\alpha}{\int_0^1 h(\alpha)d\alpha} \quad (7.17)$$

Y si utilizamos la función de ponderación $h(\alpha) = \alpha^\gamma$ escribiremos:

$$\mathbf{V}_A(\gamma, \beta) = \frac{\int_0^1 \alpha^\gamma \mathbf{v}_A(\alpha, \beta)d\alpha}{\int_0^1 \alpha^\gamma d\alpha} \quad (7.18)$$

$$\mathbf{V}_A(\gamma, \beta) = (\gamma + 1) \int_0^1 \alpha^\gamma \mathbf{v}_A(\alpha, \beta)d\alpha$$

Ejemplo 7.1 (Valor representativo de números difusos) En este ejemplo analizaremos el valor representativo de los números difusos de la Figura 7.3:

$$\begin{aligned} A &= T(1.0, 2.0, 3.0, 4.0) \\ B &= T(2.0, 2.5, 2.5, 3.0) \\ C &= T(5.0, 6.0, 7.0, 8.0) \end{aligned} \quad (7.19)$$

Las funciones $L(\alpha)$ y $R(\alpha)$ de los tres números son:

$$\begin{aligned} L_A(\alpha) &= 1 + \alpha & R_A(\alpha) &= 4 - \alpha \\ L_B(\alpha) &= 2 + \alpha/2 & R_A(\alpha) &= 3 - \alpha/2 \\ L_C(\alpha) &= 5 + \alpha & R_A(\alpha) &= 8 - \alpha \end{aligned} \quad (7.20)$$

Al aplicar 7.15 resulta

$$\begin{aligned} \mathbf{v}_A(\alpha, \beta) &= (1 - \beta)(1 + \alpha) + \beta(4 - \alpha) \\ \mathbf{v}_B(\alpha, \beta) &= (1 - \beta)(2 + \alpha/2) + \beta(3 - \alpha/2) \\ \mathbf{v}_C(\alpha, \beta) &= (1 - \beta)(5 + \alpha) + \beta(8 - \alpha) \end{aligned} \quad (7.21)$$

Al expandir y reorganizar se puede escribir:

$$\begin{aligned} \mathbf{v}_A(\alpha, \beta) &= (1 + 3\beta) + \alpha(1 - 2\beta) \\ \mathbf{v}_B(\alpha, \beta) &= (2 + \beta) + \alpha(1/2 - \beta) \\ \mathbf{v}_C(\alpha, \beta) &= (5 + 3\beta) + \alpha(1 - 2\beta) \end{aligned} \quad (7.22)$$

Si se utiliza la función ponderado $h(\alpha) = \alpha^\gamma$ se puede aplicar 7.18 para obtener

$$\begin{aligned}\mathbf{V}_A(\gamma, \beta) &= (1 + 3\beta) + (1 - 2\beta)\frac{\gamma + 1}{\gamma + 2} \\ \mathbf{V}_B(\gamma, \beta) &= (2 + \beta) + (\frac{1}{2} - \beta)\frac{\gamma + 1}{\gamma + 2} \\ \mathbf{V}_C(\gamma, \beta) &= (5 + 3\beta) + (1 - 2\beta)\frac{\gamma + 1}{\gamma + 2}\end{aligned}\quad (7.23)$$

Si seleccionamos $\gamma = 1$ en 7.16 se obtiene.

$$\begin{aligned}\mathbf{V}_A(1, \beta) &= \frac{5}{3} + \frac{5}{3}\beta \\ \mathbf{V}_B(1, \beta) &= \frac{7}{3} + \frac{1}{3}\beta \\ \mathbf{V}_C(1, \beta) &= \frac{17}{3} + \frac{5}{3}\beta\end{aligned}\quad (7.24)$$

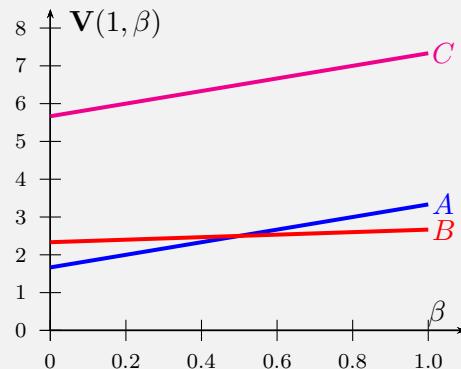


Figura 7.5: Valor representativo del ejemplo 7.1

La Figura 7.5 nos permite analizar mejor la relación que hay entre los números dibujados en la Figura 7.3. C está a la derecha de A y de B y por tanto no es de extrañar que el valor representativo de C sea mayor que el de A y el de B .

Ahora bien, la Figura 7.3 muestra que el valor representativo de B es mayor que el de A para $\beta < 0.5$, mientras que es menor para $\beta > 0.5$. Esto significa que la pregunta ¿es B mayor que A ? puede responderse así: depende del nivel de optimismo; para un nivel bajo (es decir, pesimista) B es mayor que A , mientras que para un nivel alto (es decir, optimista) B es menor que A .

Ejemplo 7.2 (Interpretación del nivel de optimismo) Este ejemplo busca ilustrar la utilidad de la variable β y la razón por la que la denominamos *nivel de optimismo*. Para hacerlo, formulamos una aplicación del uso de números difusos en el campo de la toma de decisiones.

Imaginemos que un inversionista tiene la oportunidad de hacer una inversión en una de dos alternativas disponibles, la A ó la B . Él (o ella) debe tomar la decisión basándose en la utilidad esperada de la inversión.

Imaginemos también un primer escenario en donde la información disponible es la siguiente:

- La utilidad esperada con la alternativa A es de \$2 por cada \$10 invertidos.
- La utilidad esperada con la alternativa B es de \$1 por cada \$10 invertidos.

Con ese escenario la decisión es obvia: debe seleccionarse la alternativa A porque tiene una utilidad esperada mayor que la alternativa B .

Imaginemos ahora un segundo escenario en donde la información disponible tiene incertidumbre:

- La utilidad esperada con la alternativa A está comprendida entre $-\$2$ y $\$2$ por cada \$10 invertidos. En otras palabras, puede perderse hasta \$2 o ganarse hasta \$2 por cada \$10 invertidos.
- La utilidad esperada con la alternativa B está comprendida entre $-\$1$ y $\$1$ por cada \$10 invertidos. En otras palabras, puede perderse hasta \$1 o ganarse hasta \$1 por cada \$10 invertidos.

Así las cosas ¿Qué decisión debe tomarse? La decisión ha dejado de ser obvia porque si bien es cierto que con la alternativa A puede llegar a ganarse más dinero que con la B , es igualmente cierto que puede llegar a perderse más. No es obvio cómo establecer cuál de las dos alternativas es mejor que la otra.

Esta es una situación muy conocida y estudiada en el campo de *Teoría de la decisión* (Véase, por ejemplo, [35]). En uno de los enfoques empleados, se analiza cuál es la *actitud hacia el riesgo* de la persona que debe tomar la decisión.

Una persona con actitud *optimista* tenderá a privilegiar la información favorable en detrimento de la desfavorable el momento de tomar la decisión. Lo opuesto será lo que haga una persona con una actitud *pesimista*.

La variable β que hemos introducido en el cálculo del valor representativo de los números difusos permite modelar la actitud hacia el riesgo de una persona. Por eso la hemos denominado *nivel de optimismo*.

En nuestro segundo escenario, podemos representar la utilidad esperada en cada alternativa mediante los siguientes números difusos

$$U_A = T(-3, -1, 1, 3) \quad U_B = T(-2, 0, 0, 2) \quad (7.25)$$

Si calculamos el valor representativo de U_A y U_B para $\gamma = 0$ (siguiendo un procedimiento similar al del ejemplo 7.1) se obtiene:

$$\mathbf{V}_{U_A}(0, \beta) = -2 + 4\alpha \quad \mathbf{V}_{U_B}(0, \beta) = -1 + 2\alpha \quad (7.26)$$

Al calcular los valores representativos para los valores extremos de β se obtiene:

$$\begin{cases} \mathbf{V}_{U_A}(0, 0) = -2 \\ \mathbf{V}_{U_A}(0, 1) = 2 \end{cases} \quad \begin{cases} \mathbf{V}_{U_B}(0, 0) = -1 \\ \mathbf{V}_{U_B}(0, 1) = 1 \end{cases} \quad (7.27)$$

Lo que demuestra que U_A y U_B representan adecuadamente la información del segundo escenario.

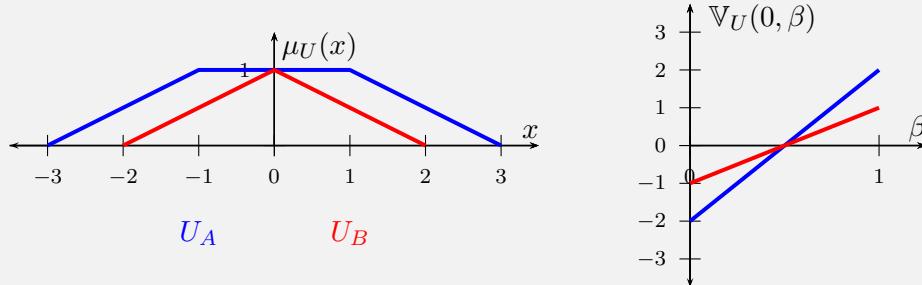


Figura 7.6: Números y valor representativo del ejemplo 7.2

En la Figura 7.6 se han graficado las funciones de pertenencia y los valores representativos de los números U_A y U_B . Analicemos ahora el efecto de la variable β :

Cuando β toma valores bajos (es decir, cercanos a 0) el valor representativo de cada número está más hacia la izquierda que cuando toma valores altos (es decir, cercanos a 1). Debido a

que los números representan las utilidades esperadas, cuando β toma valores bajos le da más peso a la información correspondiente a las peores utilidades posibles de obtener que a las mejores. En otras palabras, privilegia las peores posibilidades en detrimento de las mejores. Cuando un decisor hace eso, se dice que tiene una *actitud hacia el riesgo pesimista*. Por ello, valores bajos de β sirven para modelar la interpretación esperada que haría un decisor pesimista sobre la información disponible.

Lo contrario sucede para valores altos de β : se le da más peso a la información correspondiente a las mejores utilidades posibles de obtener que a las peores. En consecuencia, valores altos de β sirven para modelar la interpretación esperada que haría un decisor *optimista* sobre la información disponible.

Por lo anterior, la gráfica de la derecha en la Figura 7.6 permite analizar el efecto de la actitud hacia el riesgo de un decisor sobre la decisión adoptada. Una actitud pesimista (β cerca a 0) conlleva a seleccionar la alternativa B , porque con ella se espera perder menos, mientras que una actitud optimista (β cerca a 1) conlleva a seleccionar la alternativa A , porque con ella se espera ganar más.

Ejemplo 7.3 (Valor representativo de un número difuso) Consideremos el número difuso N definido por

$$\begin{aligned} L_N(\alpha) &= p + (q - p)\alpha \\ R_N(\alpha) &= s - (s - r)\sqrt{\alpha} \\ p \leq q \leq r \leq s \end{aligned} \quad (7.28)$$

El valor representativo de cada α -corte estará dado por la ecuación 7.15

$$v_N(\alpha, \beta) = (1 - \beta)(p + (q - p)\alpha) + \beta(s - (s - r)\sqrt{\alpha}) \quad (7.29)$$

Que puede expandirse y reorganizarse así:

$$v_N(\alpha, \beta) = (p + \beta(s - p)) + \beta(r - s)\sqrt{\alpha} + (1 - \beta)(q - p)\alpha \quad (7.30)$$

Al seleccionar $h(\alpha) = \alpha^\gamma$ puede emplearse 7.18 para calcular el valor representativo del número:

$$\begin{aligned} V_N(\gamma, \beta) &= (\gamma + 1) \left[\int_0^1 (p + \beta(s - p))\alpha^\gamma d\alpha \right. \\ &\quad \left. + \int_0^1 (\beta(r - s))\alpha^{\gamma+1/2} d\alpha \right. \\ &\quad \left. + \int_0^1 ((1 - \beta)(q - p))\alpha^{\gamma+1} d\alpha \right] \end{aligned} \quad (7.31)$$

Lo que deriva en

$$\mathbf{V}_N(\gamma, \beta) = (p + \beta(s - p)) + \frac{\gamma + 1}{\gamma + 3/2}(\beta(r - s)) + \frac{\gamma + 1}{\gamma + 2}((1 - \beta)(q - p)) \quad (7.32)$$

Por ejemplo, si $p = -2, q = 0, r = 0, s = 1$ se obtiene el número N que se muestra en la figura 7.7.

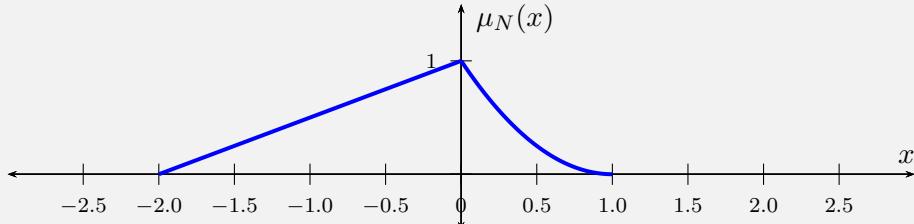


Figura 7.7: Número difuso del ejemplo 7.3

En la Figura 7.8 se muestra cómo varía el valor representativo de N con γ y con β . En la gráfica de la izquierda se puede ver que para cualquier valor fijo de γ el valor representativo cambia con β siguiendo una línea recta. Esto está acorde con la ecuación 7.15, y es así aún cuando la función de pertenencia del número no sea construida con trozos de líneas recta.

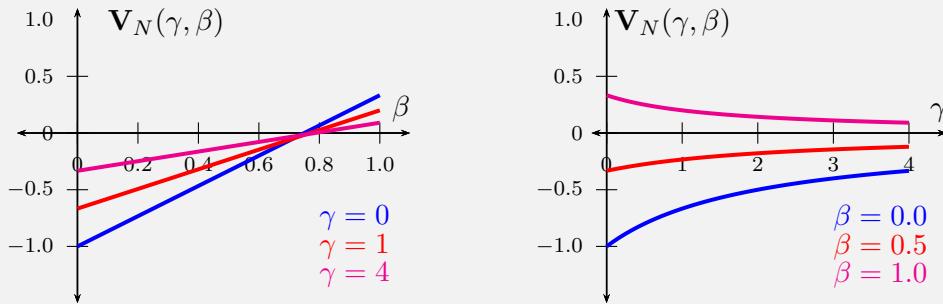


Figura 7.8: Número difuso del ejemplo 7.3

En cambio, la figura de la derecha muestra que el efecto de γ sobre el valor representativo es diferente. Incrementar el valor γ implica darle un mayor peso a los α -cortes superiores. En el límite ($\gamma \rightarrow \infty$) la ecuación 7.32 se convierte en

$$\begin{aligned} \lim_{\gamma \rightarrow \infty} \mathbf{V}_N(\gamma, \beta) &= (p + \beta(s - p)) + (\beta(r - s)) + ((1 - \beta)(q - p)) \\ &= q + \beta(r - q) \end{aligned} \quad (7.33)$$

Al remplazar por los valores del número N se obtiene

$$\lim_{\gamma \rightarrow \infty} \mathbf{V}_N(\gamma, \beta) = 0 + \beta(0 - 0) = 0 \quad (7.34)$$

Es decir, en el límite el único α -corte que pesa es $\alpha = 1$, y como $\mathbf{v}_N(1, \beta) = 0$ entonces el valor representativo del número se hace cero.

7.3. Ambigüedad

La ambigüedad de un número difuso es una medida de la incertidumbre asociada al valor que representa. Si graficamos la función de pertenencia del número, la ambigüedad es una medida de qué tan ancho es el número.

En este documento utilizamos la definición de ambigüedad formulada en [15]. Se define primero la ambigüedad de un α -corte del número difuso A como el ancho de ese intervalo:

$$\mathbf{a}_A(\alpha) = R_A(\alpha) - L_A(\alpha) \quad (7.35)$$

Mientras que la ambigüedad del número difuso A se define como el promedio de las ambigüedades de sus α -cortes ponderados por una función $h(\alpha)$ como la de la ecuación 7.16

$$\mathbf{A}_A = \frac{\int_0^1 h(\alpha) \mathbf{a}_A(\alpha) d\alpha}{\int_0^1 h(\alpha) d\alpha} \quad (7.36)$$

Y si utilizamos la función de ponderación $h(\alpha) = \alpha^\gamma$ escribiremos:

$$\mathbf{A}_A(\gamma) = \frac{\int_0^1 \alpha^\gamma \mathbf{a}_A(\alpha) d\alpha}{\int_0^1 \alpha^\gamma d\alpha} \quad (7.37)$$

$$\mathbf{A}_A(\gamma) = (\gamma + 1) \int_0^1 \alpha^\gamma \mathbf{a}_A(\alpha) d\alpha$$

Ejemplo 7.4 (Ambigüedad de números difusos) En este ejemplo continuamos con el ejemplo 7.1, para analizar la ambigüedad de los números difusos de la Figura 7.3.

Podemos aplicar directamente la ecuación 7.35 aprovechando que conocemos las funciones L y R (ecuación 7.20). Se obtiene:

$$\begin{aligned}\mathbf{a}_A(\alpha) &= (4 - \alpha) - (1 + \alpha) = 3 - 2\alpha \\ \mathbf{a}_B(\alpha) &= (3 - \alpha/2) - (2 + \alpha/2) = 1 - \alpha \\ \mathbf{a}_C(\alpha) &= (8 - \alpha) - (5 + \alpha) = 3 - 2\alpha\end{aligned}\quad (7.38)$$

Si se utiliza la función ponderado $h(\alpha) = \alpha^\gamma$ se puede aplicar 7.37 para obtener

$$\begin{aligned}\mathbf{A}_A(\gamma) &= 3 - 2 \frac{\gamma + 1}{\gamma + 2} \\ \mathbf{A}_B(\gamma) &= 1 - 1 \frac{\gamma + 1}{\gamma + 2} \\ \mathbf{A}_C(\gamma) &= 3 - 3 \frac{\gamma + 1}{\gamma + 2}\end{aligned}\quad (7.39)$$

Si seleccionamos $\gamma = 1$ en 7.16 se obtiene

$$\mathbf{A}_A(1) = 2 \quad \mathbf{A}_B(1) = 1/2 \quad \mathbf{A}_C(1) = 2 \quad (7.40)$$

Las ambigüedades de los números A y C son iguales, porque tienen la misma forma, aunque C está desplazado a la derecha respecto a A . Por otra parte, B tiene menos ambigüedad que A y que C , porque es un número más angosto (Ver Figura 7.3).

Ejemplo 7.5 (Ambigüedad de un número difuso) En este ejemplo retomamos el tipo de números difusos estudiados en el ejemplo 7.3, para determinar su ambigüedad. Se trata de números definidos por

$$\begin{aligned}L_N(\alpha) &= p + (q - p)\alpha \\ R_N(\alpha) &= s - (s - r)\sqrt{\alpha} \\ p \leq q \leq r \leq s\end{aligned}\quad (7.41)$$

La ambigüedad de cada α -corte estará dado por la ecuación 7.35

$$\mathbf{a}_N(\alpha) = (s - (s - r)\sqrt{\alpha}) - (p + (q - p)\alpha) \quad (7.42)$$

Que puede reorganizarse así:

$$\mathbf{a}_N(\alpha, \beta) = (s - p) + +(r - s)\sqrt{\alpha} + (p - q)\alpha \quad (7.43)$$

Al seleccionar $h(\alpha) = \alpha^\gamma$ puede emplearse 7.37 para calcular el valor representativo del número:

$$\begin{aligned} \mathbf{A}_N(\gamma, \beta) &= (\gamma + 1) \left[\int_0^1 (s - p)\alpha^\gamma d\alpha \right. \\ &\quad + \int_0^1 (r - s)\alpha^{\gamma+1/2} d\alpha \\ &\quad \left. + \int_0^1 (p - q)\alpha^{\gamma+1} d\alpha \right] \end{aligned} \quad (7.44)$$

Lo que deriva en

$$\mathbf{A}_N(\gamma) = (s - p) + \frac{\gamma + 1}{\gamma + 3/2}(r - s) + \frac{\gamma + 1}{\gamma + 2}(p - q) \quad (7.45)$$

Si utilizamos $p = -2, q = 0, r = 0, s = 1$ se obtiene el número N que se muestra en la figura 7.7. En la Figura 7.9 se muestra cómo varía la ambigüedad de N con γ .

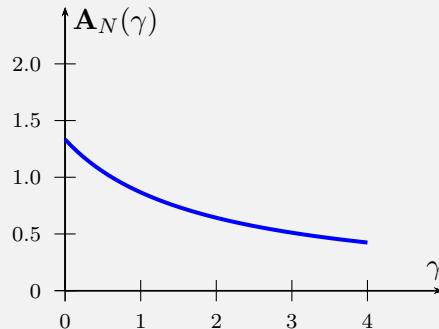


Figura 7.9: Número difuso del ejemplo 7.3

Incrementar el valor γ implica darla un mayor peso a los α -cortes superiores. En el límite ($\gamma \rightarrow \infty$) la ecuación 7.45 se convierte en

$$\lim_{\gamma \rightarrow \infty} \mathbf{A}_N(\gamma) = (s - p) + (r - s) + (p - q) = r - q \quad (7.46)$$

Al remplazar por los valores del número N se obtiene

$$\lim_{\gamma \rightarrow \infty} \mathbf{A}_N(\gamma) = 0 - 0 = 0 \quad (7.47)$$

Es decir, en el límite el único α -corte que pesa es $\alpha = 1$, y como $a_N(1) = 0$ entonces el valor representativo del número se hace cero.

7.4. Distancia entre números difusos

Consideremos ahora dos números difusos A y B . Para determinar qué tan parecidos (o diferentes) son se utiliza alguna medida de la *distancia* entre ellos, como por ejemplo (Véase [32]):

$$\begin{aligned}\mathbf{d}(A, B) &= \left[\int_0^1 (L_A(\alpha) - L_B(\alpha))^2 d\alpha + \int_0^1 (R_A(\alpha) - R_B(\alpha))^2 d\alpha \right]^{\frac{1}{2}} \\ \mathbf{d}^2(A, B) &= \int_0^1 (L_A(\alpha) - L_B(\alpha))^2 + (R_A(\alpha) - R_B(\alpha))^2 d\alpha\end{aligned}\quad (7.48)$$

Esa medida de distancia es una versión particular (la más usada) de una distancia generalizada que emplea un parámetro $q \geq 1$

$$\mathbf{d}_q(A, B) = \left[\int_0^1 (L_A(\alpha) - L_B(\alpha))^q d\alpha + \int_0^1 (R_A(\alpha) - R_B(\alpha))^q d\alpha \right]^{\frac{1}{q}} \quad (7.49)$$

Ejemplo 7.6 (Error en la aritmética de números positivos) Consideremos ahora el ejemplo 10.7 que ilustra el uso de la aritmética de números trapezoidales positivos. En ese ejemplo se calcula el producto de dos números; en la Figura 10.7 se compara el resultado exacto con el obtenido mediante la aritmética de números trapezoidales positivos. Para determinar qué tan parecidos son los dos resultados, podemos calcular la distancia entre los dos.

Los números del ejemplo son $A = T(0, 1, 1, 2)$ y $B = T(1, 2, 3, 4)$ por lo tanto

$$\begin{cases} L_A(\alpha) = \alpha \\ R_A(\alpha) = 2 - \alpha \end{cases} \quad \begin{cases} L_B(\alpha) = 1 + \alpha \\ R_B(\alpha) = 4 - \alpha \end{cases} \quad (7.50)$$

Y por tanto

$$L_{AB}(\alpha) = \alpha + \alpha^2 \quad R_{AB}(\alpha) = 8 - 6\alpha + \alpha^2 \quad (7.51)$$

El resultado obtenido mediante la aritmética de números trapezoidales positivos es $C = T(0, 2, 3, 8)$, es decir

$$L_C(\alpha) = 2\alpha \quad R_C(\alpha) = 8 - 5\alpha \quad (7.52)$$

Podemos ahora calcular la distancia entre AB y C empleando 7.48:

$$\begin{aligned} d^2(AB, C) &= \int_0^1 ((\alpha + \alpha^2) - (2\alpha))^2 + ((8 - 6\alpha + \alpha^2) - (8 - 5\alpha))^2 d\alpha \\ d^2(AB, C) &= \int_0^1 2\alpha^4 - 4\alpha^3 + 2\alpha^2 d\alpha \\ d^2(AB, C) &= \left. \frac{2}{5}\alpha^5 - \alpha^4 + \frac{2}{3}\alpha^3 \right|_0^1 = \frac{2}{5} - 4 + \frac{2}{3} = \frac{1}{15} \end{aligned} \quad (7.53)$$

Es decir, el error que se ha cometido al emplear la aritmética de números trapezoidales positivos ha sido

$$d(AB, C) = \frac{1}{\sqrt{15}} \quad (7.54)$$

Ejemplo 7.7 (Distancia entre dos números difusos) En este ejemplo se estudia la distancia entre un número trapezoidal $M = T(a, b, c, d)$ y un número difuso N como los estudiados en los ejemplos 7.3 y 7.5, definido por:

$$\begin{aligned} L_N(\alpha) &= p + (q - p)\alpha \\ R_N(\alpha) &= s - (s - r)\sqrt{\alpha} \\ p \leq q \leq r \leq s \end{aligned} \quad (7.55)$$

La aplicación de la ecuación 7.48 resulta en:

$$\begin{aligned} d^2(M, N) &= \int_0^1 ((a + \alpha(b - a)) - (p + (q - p)\alpha))^2 d\alpha + \\ &\quad \int_0^1 ((d - \alpha(d - c)) - (s - (s - r)\sqrt{\alpha}))^2 d\alpha \end{aligned} \quad (7.56)$$

Consideremos primero los integrandos, a los que denominaremos I_L e I_R :

$$\begin{aligned} I_L &= ((a + \alpha(b - a)) - (p + (q - p)\alpha))^2 \\ &= ((a - p) + (b - a + p - q)\alpha)^2 \\ &= (A + B\alpha)^2 \\ &= A^2 + 2AB\alpha + B^2\alpha^2 \\ I_R &= ((d - \alpha(d - c)) - (s - (s - r)\sqrt{\alpha}))^2 \\ &= ((d - s) + (s - r)\sqrt{\alpha} + (c - d)\alpha)^2 \\ &= (C + D\sqrt{\alpha} + E\alpha)^2 \\ &= C^2 + 2CD\alpha^{1/2} + (D^2 + 2CE)\alpha + 2DE\alpha^{3/2} + E^2\alpha^2 \end{aligned} \quad (7.57)$$

En donde hemos hecho:

$$\begin{aligned} A &= a - b & B &= b - a + p - q \\ C &= d - s & D &= s - r & E &= c - d \end{aligned} \quad (7.58)$$

La suma de los integrandos es, entonces:

$$I_L + I_R = (A^2 + C^2) + 2CD\alpha^{1/2} + (D^2 + 2CE + 2AB)\alpha + 2DE\alpha^{3/2} + (B^2 + E^2)\alpha^2 \quad (7.59)$$

Dado que $\int_0^1 \alpha^\gamma d\alpha = 1/(\gamma + 1)$, podemos calcular la integral de la ecuación 7.56 directamente:

$$\mathbf{d}^2(M, N) = (A^2 + C^2) + \frac{2CD}{3/2} + \frac{D^2 + 2CE + 2AB}{2} + \frac{2DE}{5/2} + \frac{B^2 + E^2}{3} \quad (7.60)$$

Consideremos ahora el caso en que el número N es el mismo empleado en los ejemplos 7.3 y 7.5, es decir, $p = -2, q = 0, r = 0, s = 1$, y calculemos su distancia al número trapezoidal $M = T(-2, 0, 0, 0)$.

$$A = 0 \quad B = 0 \quad C = 0 \quad D = 1 \quad E = -1 \quad (7.61)$$

$$\mathbf{d}(M, N) = \sqrt{\frac{1}{2} + \frac{4}{5} + \frac{1}{3}} = \sqrt{\frac{1}{30}} = 0.1826 \quad (7.62)$$

7.5. Tendencias de los parámetros de un número difuso

Para un número difuso A , definimos la tendencia de uno de sus parámetros como la recta que mejor representa la variación de ese parámetro respecto a α . Además, adoptamos el criterio de mínimos cuadrados convencional para determinar las pendientes y términos independientes de dicha recta.

Para aplicar la anterior definición, supongamos un parámetro $f(\alpha)$ y una recta cuya ecuación es $m + b\alpha$. Los parámetros m, b de la tendencia de $f(\alpha)$ son los que minimizan:

$$E(m, b) = \int_0^1 (f(\alpha) - (m\alpha + b))^2 d\alpha \quad (7.63)$$

Para determinar m, b calculamos las derivadas parciales:

$$\begin{aligned} \frac{\partial E}{\partial m} &= \int_0^1 \frac{(f(\alpha) - (m\alpha + b))^2}{\partial m} d\alpha \\ \frac{\partial E}{\partial b} &= \int_0^1 \frac{(f(\alpha) - (m\alpha + b))^2}{\partial b} d\alpha \end{aligned} \quad (7.64)$$

Utilizando la regla de la cadena:

$$\begin{aligned} \frac{\partial E}{\partial m} &= \int_0^1 2(f(\alpha) - m\alpha - b)(-\alpha) d\alpha \\ \frac{\partial E}{\partial b} &= \int_0^1 2(f(\alpha) - m\alpha - b)(-1) d\alpha \end{aligned} \quad (7.65)$$

Lo que resulta en:

$$\begin{aligned} \frac{\partial E}{\partial m} &= -2 \int_0^1 \alpha f(\alpha) d\alpha + 2m \int_0^1 \alpha^2 d\alpha + 2b \int_0^1 \alpha d\alpha \\ \frac{\partial E}{\partial b} &= -2 \int_0^1 f(\alpha) d\alpha + 2m \int_0^1 \alpha d\alpha + 2b \int_0^1 d\alpha \end{aligned} \quad (7.66)$$

Calculando las integrales:

$$\begin{aligned} \frac{\partial E}{\partial m} &= -2 \int_0^1 \alpha f(\alpha) d\alpha + \frac{2}{3}m + b \\ \frac{\partial E}{\partial b} &= -2 \int_0^1 f(\alpha) d\alpha + m + 2b \end{aligned} \quad (7.67)$$

Para encontrar los valores de m, b que minimizan F se igualan las derivadas parciales a cero, con lo que se obtiene un sistema de 2 ecuaciones lineales:

$$\begin{cases} \frac{1}{3}m + \frac{1}{2}b = \int_0^1 \alpha f(\alpha) d\alpha \\ \frac{1}{2}m + b = \int_0^1 f(\alpha) d\alpha \end{cases} \quad (7.68)$$

Usando la definición de $\mathbf{F}(\gamma)$ podemos escribir el sistema de ecuaciones de forma matricial

$$\begin{bmatrix} 1/3 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} \mathbf{F}(1) \\ \mathbf{F}(0) \end{bmatrix} \quad (7.69)$$

Lo que nos permite obtener m, b despejando:

$$\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{F}(1) \\ \mathbf{F}(0) \end{bmatrix} \quad (7.70)$$

Lo anterior lo podemos consignar en el Teorema 7.1

Teorema 7.1 *Sea $f(\alpha)$ la función que define un cierto parámetro geométrico de un número difuso y $\mathbf{F}(\gamma) = \int_0^1 \alpha^\gamma f(\alpha) d\alpha$. En esas condiciones la recta $\hat{f}(\alpha) = m\alpha + b$ que minimiza el error cuadrático está determinada por*

$$\begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{F}(1) \\ \mathbf{F}(0) \end{bmatrix} \quad (7.71)$$

Además, recordemos que $L_A(\alpha) = \mathbf{v}(\alpha, 0)$, $R_A(\alpha) = \mathbf{v}(\alpha, 1)$, y que el denominador en las ecuaciones 7.18 y 7.37 es $1/(\gamma + 1)$ se obtiene:

$$\begin{aligned} \begin{bmatrix} m_L \\ b_L \end{bmatrix} &= \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{V}(1, 0)/2 \\ \mathbf{V}(0, 0)/1 \end{bmatrix} & \begin{bmatrix} m_R \\ b_R \end{bmatrix} &= \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{V}(1, 1)/2 \\ \mathbf{V}(0, 1)/1 \end{bmatrix} \\ \begin{bmatrix} m_v \\ b_v \end{bmatrix} &= \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{V}(1, 0.5)/2 \\ \mathbf{V}(0, 0.5)/1 \end{bmatrix} & \begin{bmatrix} m_a \\ b_a \end{bmatrix} &= \begin{bmatrix} 12 & -6 \\ -6 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{A}(1)/2 \\ \mathbf{A}(0)/1 \end{bmatrix} \end{aligned} \quad (7.72)$$

Por lo tanto, las tendencias definidas por:

$$\begin{aligned} \text{Tendencia de } L_A(\alpha) &: m_L \times \alpha + b_L \\ \text{Tendencia de } R_A(\alpha) &: m_R \times \alpha + b_R \\ \text{Tendencia de } \mathbf{v}_A(\alpha, 0.5) &: m_v \times \alpha + b_v \\ \text{Tendencia de } \mathbf{a}_A(\alpha) &: m_a \times \alpha + b_a \end{aligned} \quad (7.73)$$

Se pueden calcular empleando

$$\begin{aligned} \begin{bmatrix} m_L \\ b_L \end{bmatrix} &= \begin{bmatrix} 6 & -6 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{V}(1,0) \\ \mathbf{V}(0,0) \end{bmatrix} & \begin{bmatrix} m_R \\ b_R \end{bmatrix} &= \begin{bmatrix} 6 & -6 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{V}(1,1) \\ \mathbf{V}(0,1) \end{bmatrix} \\ \begin{bmatrix} m_v \\ b_v \end{bmatrix} &= \begin{bmatrix} 6 & -6 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{V}(1,0.5) \\ \mathbf{V}(0,0.5) \end{bmatrix} & \begin{bmatrix} m_a \\ b_a \end{bmatrix} &= \begin{bmatrix} 6 & -6 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} \mathbf{A}(1) \\ \mathbf{A}(0) \end{bmatrix} \end{aligned} \quad (7.74)$$

Por otra parte, de la ecuación 7.74 se tiene que

$$\begin{aligned} m_L &= 6(\mathbf{V}_A(1,0) - \mathbf{V}_A(0,0)) \\ m_R &= 6(\mathbf{V}_A(1,1) - \mathbf{V}_A(0,1)) \\ m_a &= 6(\mathbf{A}_A(1) - \mathbf{A}_A(0)) \end{aligned} \quad (7.75)$$

Por otra parte, como $L(\alpha)$ es monótonamente no decreciente entonces $m_L \geq 0$. De forma análoga, como $R(\alpha)$ es monótonamente no creciente entonces $m_R \leq 0$. Además, esas mismas propiedades implican que $\mathbf{a}(\alpha)$ es monótonamente no creciente y por tanto $m_a \leq 0$. De allí se desprende que

$$\begin{aligned} \mathbf{V}_A(1,0) - \mathbf{V}_A(0,0) &\geq 0 \\ \mathbf{V}_A(1,1) - \mathbf{V}_A(0,1) &\leq 0 \\ \mathbf{A}_A(1) - \mathbf{A}_A(0) &\leq 0 \end{aligned} \quad (7.76)$$

O, lo que es igual, al comparar los valores representativos de cualquier número difuso A para $\gamma = 1$ y $\gamma = 0$ se tienen las siguientes desigualdades

$$\begin{aligned} \mathbf{V}_A(1,0) &\geq \mathbf{V}_A(0,0) \\ \mathbf{V}_A(1,1) &\leq \mathbf{V}_A(0,1) \\ \mathbf{A}_A(1) &\leq \mathbf{A}_A(0) \end{aligned} \quad (7.77)$$

7.6. Caracterización de números trapezoidales

Consideremos el número difuso trapezoidal $A = T(a, b, c, d)$ para el cual

$$\begin{aligned} L_A(\alpha) &= a + \alpha(b - a) \\ R_A(\alpha) &= d - \alpha(d - c) \end{aligned} \quad (7.78)$$

El valor representativo y la ambigüedad de cada uno de sus α -cortes se obtiene aplicando 7.15 y 7.35:

$$\begin{aligned} \mathbf{v}_A(\alpha, \beta) &= (1 - \beta)(a + \alpha(b - a)) + \beta(d - \alpha(d - c)) \\ \mathbf{a}_A(\alpha) &= d - \alpha(d - c) - a - \alpha(b - a) \end{aligned} \quad (7.79)$$

Teniendo en cuenta que

$$\begin{cases} \mathbf{v}_A(0, \beta) = (1 - \beta)a + \beta d \\ \mathbf{a}_A(0) = d - a \end{cases} \quad \begin{cases} \mathbf{v}_A(1, \beta) = (1 - \beta)b + \beta c \\ \mathbf{a}_A(1) = c - b \end{cases} \quad (7.80)$$

La ecuación 7.79 puede reescribirse:

$$\begin{aligned} \mathbf{v}_A(\alpha, \beta) &= \mathbf{v}_A(0, \beta) + \alpha (\mathbf{v}_A(1, \beta) - \mathbf{v}_A(0, \beta)) \\ \mathbf{a}_A(\alpha) &= \mathbf{a}_A(0) + \alpha (\mathbf{a}_A(1) - \mathbf{a}_A(0)) \end{aligned} \quad (7.81)$$

La ecuación 7.81 significa que para los α -cortes de los números trapezoidales su valor representativo y su ambigüedad varían linealmente desde los valores correspondientes a $\alpha = 0$ hasta los de $\alpha = 1$.

Al emplear la función de ponderación $h(\alpha) = \alpha^\gamma$ el denominador de las ecuaciones 7.17 y 7.36 resulta ser

$$\int_0^1 \alpha^\gamma d\alpha = \frac{\alpha^{\gamma+1}}{\gamma+1} \Big|_0^1 = \frac{1}{\gamma+1} \quad (7.82)$$

Y los numeradores:

$$\begin{aligned} \int_0^1 \alpha^\gamma \mathbf{v}_A(\alpha, \beta) d\alpha &= \int_0^1 \alpha^\gamma (\mathbf{v}_A(0, \beta) + \alpha (\mathbf{v}_A(1, \beta) - \mathbf{v}_A(0, \beta))) d\alpha \\ &= \mathbf{v}_A(0, \beta) \int_0^1 \alpha^\gamma d\alpha + (\mathbf{v}_A(1, \beta) - \mathbf{v}_A(0, \beta)) \int_0^1 \alpha^{\gamma+1} d\alpha \quad (7.83) \\ &= \mathbf{v}_A(0, \beta) \frac{1}{\gamma+1} + (\mathbf{v}_A(1, \beta) - \mathbf{v}_A(0, \beta)) \frac{1}{\gamma+2} \end{aligned}$$

$$\begin{aligned} \int_0^1 \alpha^\gamma \mathbf{a}_A(\alpha) d\alpha &= \int_0^1 \alpha^\gamma (\mathbf{a}_A(0) + \alpha (\mathbf{a}_A(1) - \mathbf{a}_A(0))) d\alpha \\ &= \mathbf{a}_A(0) \int_0^1 \alpha^\gamma d\alpha + (\mathbf{a}_A(1) - \mathbf{a}_A(0)) \int_0^1 \alpha^{\gamma+1} d\alpha \quad (7.84) \\ &= \mathbf{a}_A(0) \frac{1}{\gamma+1} + (\mathbf{a}_A(1) - \mathbf{a}_A(0)) \frac{1}{\gamma+2} \end{aligned}$$

Por lo tanto se tiene que

$$\left\{ \begin{array}{l} \mathbf{V}_A(\gamma, \beta) = \frac{\mathbf{v}_A(0, \beta) \frac{1}{\gamma+1} + (\mathbf{v}_A(1, \beta) - \mathbf{v}_A(0, \beta)) \frac{1}{\gamma+2}}{\frac{1}{\gamma+1}} \\ \mathbf{A}_A(\gamma) = \frac{\mathbf{a}_A(0) \frac{1}{\gamma+1} + (\mathbf{a}_A(1) - \mathbf{a}_A(0)) \frac{1}{\gamma+2}}{\frac{1}{\gamma+1}} \end{array} \right. \quad (7.85)$$

Simplificando se llega a

$$\left\{ \begin{array}{l} V_A(\gamma, \beta) = \mathbf{v}_A(0, \beta) + \frac{\gamma+1}{\gamma+2} (\mathbf{v}_A(1, \beta) - \mathbf{v}_A(0, \beta)) \\ A_A(\gamma) = \mathbf{a}_A(0) + \frac{\gamma+1}{\gamma+2} (\mathbf{a}_A(1) - \mathbf{a}_A(0)) \end{array} \right. \quad (7.86)$$

Comparando las ecuaciones 7.81 y 7.86 se observa que el valor representativo y la ambigüedad de un número trapezoidal son iguales a las de uno de sus α -corte. Es decir:

$$\mathbf{V}_A(\gamma, \beta) = \mathbf{v}_A(\tilde{\alpha}, \beta) \quad \mathbf{A}_A(\gamma) = \mathbf{a}_A(\tilde{\alpha}) \quad \tilde{\alpha} = \frac{\gamma+1}{\gamma+2} \quad (7.87)$$

Lo que deriva en

$$\left\{ \begin{array}{l} \mathbf{V}_A(\gamma, \beta) = (a + \beta(d-a)) + \tilde{\alpha}((b + \beta(c-b)) - (a + \beta(d-a))) \\ \mathbf{A}_A(\gamma) = (d-a) + \tilde{\alpha}((c-b) - (d-a)) \end{array} \right. \quad (7.88)$$

Podemos también escribir 7.88 en forma matricial así:

$$\begin{bmatrix} \mathbf{V}_A(\gamma, \beta) \\ \mathbf{A}_A(\gamma) \end{bmatrix} = \begin{bmatrix} (1-\tilde{\alpha})(1-\beta) & \tilde{\alpha}(1-\beta) & \tilde{\alpha}\beta & (1-\tilde{\alpha})\beta \\ -(1-\tilde{\alpha}) & -\tilde{\alpha} & \tilde{\alpha} & (1-\tilde{\alpha}) \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (7.89)$$

Ejemplo 7.8 (Valor representativo y ambigüedad de números trapezoidales)

Consideremos el número difuso $A = T(0, 2, 2, 4)$, y estudiamos primero su valor representativo. Empleando la ecuación 7.89 se han preparado las gráficas de la Figura 7.10.

La gráfica de la izquierda ayuda a entender el efecto de γ sobre el valor representativo. Para hacerlo, se ha trazado $V_A(\gamma, \beta)$ para tres valores fijos de β . Incrementar el valor de γ significa darle más peso a los α -cortes superiores. Por esa razón, las gráficas se acercan a un cierto valor, que corresponde al valor representativo del α -corte más alto, $v(1, \beta)$. En este ejemplo ese valor es 2, para todos los valores de β .

Es necesario hacer una observación especial sobre el rango de valores válidos de γ . De acuerdo con la ecuación 7.16, γ debe ser mayor o igual a 0, lo que corresponde a las curvas trazadas con una línea sólida. Sin embargo, en este ejemplo hemos explorado el efecto de utilizar valores negativos de γ , en concreto, hemos explorado $\gamma \in (-1, 0)$. Ese comportamiento se ha trazado con líneas punteadas.

Valores negativos de γ hacen que la función de ponderación $h(\alpha) = \alpha^\gamma$ asigne un peso mayor a los α -cortes inferiores. Una ponderación así carece de sentido para el cálculo de un valor representativo, porque un mayor grado de pertenencia se interpreta como una mayor posibilidad de ocurrencia.

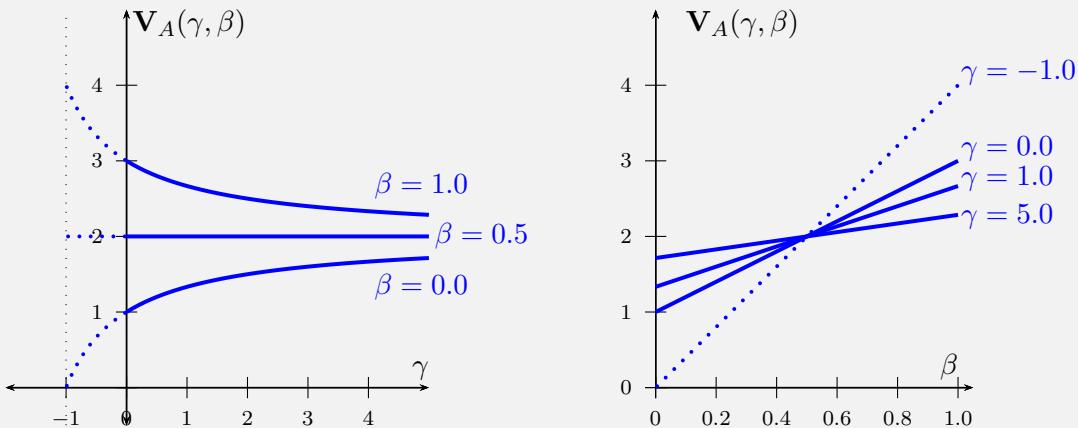


Figura 7.10: Valor representativo en el ejemplo 7.8

También se ha calculado la ambigüedad empleando la ecuación 7.89 para varios valores de γ , lo que nos ha permitido dibujar la Figura 7.11. De nuevo, se ha trazado en línea punteada el comportamiento para $\gamma \in (-1, 0)$. Es lógico que la curva sea descendente, porque al

aumentar γ se ponderan con más peso los α -cortes superiores, que son más angostos, es decir, que tiene menor ambigüedad.

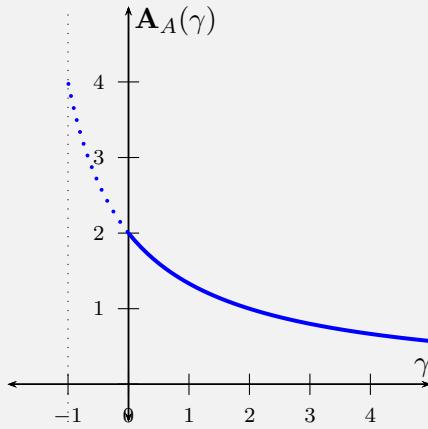


Figura 7.11: Ambigüedad en el ejemplo 7.8

7.7. Aproximaciones trapezoidales de números difusos

Los números trapezoidales son sencillos de representar y sus operaciones son fáciles de implementar y analizar. Además, para una persona es relativamente fácil interpretar la información que representa un número difuso que cuantifique una determinada magnitud. Por ello, no es de extrañar que sean muy utilizados para representar el conocimiento de los expertos o el resultado de mediciones directas o de estimaciones de variables.

Sin embargo, pese a que la información inicial se represente mediante números trapezoidales, al efectuar operaciones entre números trapezoidales el resultado puede ser otro tipo de número. Operaciones simples como el producto y la división tiene ese efecto (secciones 10.3 y 10.4). También pueden aparecer números no trapezoidales al calcular la inversa de funciones no lineales (Sección 11.3), situación que sucede en varias aplicaciones prácticas.

Las herramientas computacionales permiten calcular y manipular números de formas completamente arbitrarias. Sin embargo, es usual que esos resultados sean presentados a una persona que debe interpretarlos, y esa tarea puede ser difícil de realizar si la forma es extraña. Por esa razón, puede ser preferible presentar los resultados con números trapezoidales.

Lo anterior plantea un problema concreto: Dado un cierto número difuso ¿Cuál es el

Tabla 7.2: Casos de la aproximación trapezoidal que minimiza la distancia

Caso 1	Caso 2
$a = \int_0^1 (4 - 6\alpha)L_A(\alpha)d\alpha$ $b = \int_0^1 (6\alpha - 2)L_A(\alpha)d\alpha$ $c = \int_0^1 (6\alpha - 2)R_A(\alpha)d\alpha$ $d = \int_0^1 (4 - 6\alpha)R_A(\alpha)d\alpha$	$a = 6 \int_0^1 \alpha L_A(\alpha)d\alpha - 4 \int_0^1 \alpha R_A(\alpha)d\alpha$ $b = 2 \int_0^1 \alpha R_A(\alpha)d\alpha$ $c = b$ $d = b$
Caso 3	Caso 4
$a = 2 \int_0^1 \alpha L_A(\alpha)d\alpha$ $b = a$ $c = a$ $d = 6 \int_0^1 \alpha R_A(\alpha)d\alpha - 4 \int_0^1 \alpha L_A(\alpha)d\alpha$	$a = 2 \int_0^1 L_A(\alpha)d\alpha + \int_0^1 (6\alpha - 2)R_A(\alpha)d\alpha$ $b = \int_0^1 (3\alpha - 1)L_A(\alpha)d\alpha + \int_0^1 (3\alpha - 1)R_A(\alpha)d\alpha$ $c = b$ $d = \int_0^1 (2 - 6\alpha)L_A(\alpha)d\alpha + 2 \int_0^1 R_A(\alpha)d\alpha$

número trapezoidal que más se le parece?

Para abordar ese problema primero hay que precisar qué significa que dos números difusos sean parecidos. En [4], por ejemplo, se busca minimizar la distancia entre el número y su aproximación. Para ello, primero se definen el valor y la ambigüedad de un número difuso A así:

$$\text{Val}(A) = \int_0^1 \alpha(R_A(\alpha) + L_A(\alpha))d\alpha \quad \text{Amb}(A) = \int_0^1 \alpha(R_A(\alpha) - L_A(\alpha))d\alpha \quad (7.90)$$

En [4] se demuestra que el número $T(a, b, c, d)$ cuya distancia a un número difuso A es mínima y que preserva su valor y ambigüedad se obtiene como uno los casos que se muestran en la Tabla 7.2. También se demuestra que para determinar cuál de los casos es el correcto se pueden utilizar las desigualdades de la Tabla 7.3 así:

- Si se cumple 1, entonces el caso es 1.
- Si no se cumple 1 y se cumple 2, entonces el caso es 2.
- Si no se cumple 1 y se cumple 3, entonces el caso es 3.
- Si no se cumplen ni 1, ni 2 ni 3, entonces el caso es 4.

En este texto empleamos un enfoque diferente [19]. Una forma de interpretar el significado de ‘parecido’ es el de imponer dos condiciones: 1) que tengan el mismo valor

Tabla 7.3: Discriminantes de aproximación trapezoidal que minimiza la distancia

1 :	$\int_0^1 (3\alpha - 1)L_A(\alpha)d\alpha - \int_0^1 (3\alpha - 1)R_A(\alpha)d\alpha \leq 0$
2 :	$\int_0^1 (3\alpha - 1)L_A(\alpha)d\alpha + \int_0^1 (\alpha - 1)R_A(\alpha)d\alpha > 0$
3 :	$\int_0^1 (\alpha - 1)L_A(\alpha)d\alpha + \int_0^1 (3\alpha - 1)R_A(\alpha)d\alpha > 0$

representativo y 2) que tengan la misma ambigüedad. Sin embargo, debido a que un número difuso se define por cuatro variables (la posición de sus vértices), imponer sólo dos condiciones no es suficiente para resolver el problema, y es necesario refinar más el significado de ‘parecido’. Dado que el valor representativo y la ambigüedad dependen de la selección de los parámetros γ y β , se pueden construir más condiciones que permitan obtener de forma explícita la mejor aproximación trapezoidal de cualquier número difuso, bajo otros criterios de semejanza.

Precisemos el problema. Dado un número difuso A , se desea obtener un número difuso trapezoidal $X = T(a, b, c, d)$ que satisfaga algunos criterios específicos seleccionados para que la información que representa X sea semejante a la que representa A .

En las secciones 7.7.1 a 7.7.5 se presentan 5 alternativas. Cada una de ellas emplea un conjunto de 4 parámetros que describen alguna propiedad geométrica de A . Supongamos que uno de esos conjuntos es $\{y_1(A), y_2(A), y_3(A), y_4(A)\}$. Supongamos también que para el caso de los números trapezoidales es posible calcular esos 4 parámetros como una combinación lineal de sus vértices utilizando una cierta matriz invertible P_k . Si esto es así, entonces podemos escribir:

$$Y_k(A) = \begin{bmatrix} y_1(A) \\ y_2(A) \\ y_3(A) \\ y_4(A) \end{bmatrix} = P_k \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (7.91)$$

El subíndice k se usa para indicar que hay más de un conjunto de cuatro parámetros que puede usarse. Como P_k es invertible, para cada uno de esos conjuntos de parámetros podemos escribir:

$$\begin{bmatrix} a & b & c & d \end{bmatrix}^T = M_k \times Y_k(A) \quad (7.92)$$

que permite encontrar un número $X_k = T(a, b, c, d)$ a partir de un vector $Y(A)$ que contiene ciertas propiedades del número A que es necesario calcular. La matriz M_k es una matriz constante y puede estar previamente calculada.

Sin embargo, para asegurar que el resultado sea efectivamente un número difuso se encuentra que es necesario imponer ciertas restricciones a $Y_k(A)$, es decir, a las propiedades de A . Por ello, se han preparado también unos algoritmos que verifican si se cumplen o no tales restricciones y que, en caso de no cumplirse, relajan las condiciones del problema, modificando $Y_k(A)$.

La Tabla 7.4 consigna el vector $Y_k(A)$, las matrices P_k , M_k y las restricciones para cada uno de los 5 conjuntos de parámetros establecidos. El contenido de esa tabla se demuestra en las secciones 7.7.1 a 7.7.5.

Realmente, cada uno de los conjuntos de parámetros de la Tabla 7.4 corresponde a un conjunto de criterios geométrico de semejanza. La Tabla 7.5 contiene el criterio asociado a cada uno de ellos.

7.7.1. Conjunto de criterios 0

El primer criterio que presentamos se especifica mediante dos condiciones. Dado un número difuso A , su representación trapezoidal X debe ser tal que:

- el soporte de X sea igual al de A
- el núcleo de X sea igual al de A

Le hemos dado el número 0 a este criterio, porque realmente es un enfoque que no considera ni el valor representativo de A ni su ambigüedad. En ese sentido, es una aproximación ingenua que utilizaremos solo con fines de contrastación con los otros criterios que son más elaborados.

El soporte y el núcleo de X están determinados por sus vértices, mientras que los de A corresponden a los α -cortes inferior y superior. Se puede utilizar 7.15 para escribir

$$\begin{cases} \text{soporte}(X) = [a, d] = [\mathbf{v}_A(0, 0), \mathbf{v}_A(0, 1)] = \text{soporte}(A) \\ \text{núcleo}(X) = [b, c] = [\mathbf{v}_A(1, 0), \mathbf{v}_A(1, 1)] = \text{núcleo}(A) \end{cases} \quad (7.93)$$

De tal manera que

$$\begin{cases} a = \mathbf{v}_A(0, 0) & b = \mathbf{v}_A(1, 0) \\ c = \mathbf{v}_A(1, 1) & d = \mathbf{v}_A(0, 1) \end{cases} \quad (7.94)$$

Que en forma matricial se convierte en una expresión de la forma que tiene la ecuación 7.92:

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{v}_A(0, 1) \\ \mathbf{v}_A(1, 0) \\ \mathbf{v}_A(1, 1) \\ \mathbf{v}_A(0, 1) \end{bmatrix} \quad (7.95)$$

Tabla 7.4: Matrices para la aproximación a números trapezoidales

k	Y_k	P_k	M_k	Restricción
0	$\begin{bmatrix} \mathbf{v}(0,0) \\ \mathbf{v}(1,0) \\ \mathbf{v}(1,1) \\ \mathbf{v}(0,1) \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	
1	$\begin{bmatrix} \mathbf{V}(0,0.5) \\ \mathbf{A}(0) \\ \mathbf{V}(1,0.5) \\ \mathbf{A}(1) \end{bmatrix}$	$\begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -1/2 & -1/2 & 1/2 & 1/2 \\ 1/6 & 1/3 & 1/3 & 1/6 \\ -1/3 & -2/3 & 2/3 & 1/3 \end{bmatrix}$	$\begin{bmatrix} 4 & -2 & -3 & 3/2 \\ -2 & 1 & 3 & -3/2 \\ -2 & -1 & 3 & 3/2 \\ 4 & 2 & -3 & -3/2 \end{bmatrix}$	$\frac{\mathbf{A}(1)}{\mathbf{A}(0)} \geq \frac{2}{3}$
2	$\begin{bmatrix} \mathbf{V}(0,0) \\ \mathbf{V}(0,1) \\ \mathbf{V}(1,0) \\ \mathbf{V}(1,1) \end{bmatrix}$	$\begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 2/3 & 1/3 \end{bmatrix}$	$\begin{bmatrix} 4 & 0 & -3 & 0 \\ -2 & 0 & 3 & 0 \\ 0 & -2 & 0 & 3 \\ 0 & 4 & 0 & -3 \end{bmatrix}$	$\frac{\mathbf{V}(1,1) - \mathbf{V}(1,0)}{\mathbf{V}(0,1) - \mathbf{V}(0,0)} \geq \frac{2}{3}$
3	$\begin{bmatrix} m_v \\ b_v \\ m_a \\ b_a \end{bmatrix}$	$\begin{bmatrix} -1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 1 & -1 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & -1/2 \\ 1 & 1 & -1/2 & -1/2 \\ 1 & 1 & 1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 \end{bmatrix}$	$m_a + b_a \geq 0$
4	$\begin{bmatrix} m_L \\ b_L \\ m_R \\ b_R \end{bmatrix}$	$\begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$m_R + b_R \geq m_L + b_L$

Tabla 7.5: Conjuntos de criterios de semejanza geométrica entre X_k y A .

k	Nombre y Criterio
0	Núcleo y soporte : <ul style="list-style-type: none"> ■ T debe preservar el núcleo de A ■ T debe preservar el soporte de A
1	Valor central y ambigüedad : <ul style="list-style-type: none"> ■ T debe preservar el valor representativo central de A al menos para dos valores de γ diferentes ■ T debe preservar la ambigüedad de A al menos para dos valores de γ diferentes
2	Niveles de optimismo extremos : <ul style="list-style-type: none"> ■ T debe preservar el valor representativo pesimista de A al menos para dos valores de γ diferentes ■ T debe preservar el valor representativo optimista de A al menos para dos valores de γ diferentes
3	Tendencias de valor y ambigüedad : <ul style="list-style-type: none"> ■ T debe preservar la tendencia del valor representativo central de A. ■ T debe preservar la tendencia de la ambigüedad de A.
4	Tendencias L y R : <ul style="list-style-type: none"> ■ T debe preservar la tendencia de la curva $L(\alpha)$ de A. ■ T debe preservar la tendencia de la curva $R(\alpha)$ de A.

Para verificar que los valores de a, b, c, d permiten construir un número trapezoidal debe analizarse si se satisface la condición

$$a \leq b \leq c \leq d \quad (7.96)$$

Dado que A es un número difuso, se sabe que sus vértices están ordenados, y por tanto se puede asegurar que la condición 7.96 siempre se satisface.

7.7.2. Conjunto de criterios 1

En esta sección exploramos el siguiente criterio de semejanza: Dado un número difuso A , su representación trapezoidal X debe ser tal que:

- el valor representativo central (optimismo neutro) X sea igual al de A , al menos para dos valores de γ diferentes
- la ambigüedad de X sea igual a la de A , al menos para dos valores de γ diferentes

La condición de optimismo neutro implica $\beta = 0.5$. Para aplicar este criterio se puede emplear la ecuación 7.89 con dos valores diferentes de γ . Por ejemplo, usando $\gamma = 1$ y $\gamma = 0$ se obtiene:

$$\begin{bmatrix} \mathbf{V}_A(0, 0.5) \\ \mathbf{A}_A(0) \\ \mathbf{V}_A(1, 0.5) \\ \mathbf{A}_A(1); \end{bmatrix} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ -1/2 & -1/2 & 1/2 & 1/2 \\ 1/6 & 1/3 & 1/3 & 1/6 \\ -1/3 & -2/3 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (7.97)$$

Despejando se consigue una expresión de la forma que tiene la ecuación 7.92:

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 4 & -2 & -3 & 3/2 \\ -2 & 1 & 3 & -3/2 \\ -2 & -1 & 3 & 3/2 \\ 4 & 2 & -3 & -3/2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_A(0, 0.5) \\ \mathbf{A}_A(0) \\ \mathbf{V}_A(1, 0.5) \\ \mathbf{A}_A(1); \end{bmatrix} \quad (7.98)$$

Además, se sabe que:

$$\begin{cases} \mathbf{A}_A(0) \geq 0 & \text{la ambigüedad es positiva} \\ \mathbf{A}_A(1) \geq 0 & \text{la ambigüedad es positiva} \\ \mathbf{A}_A(0) \geq \mathbf{A}_A(1) & \text{Por convexidad} \end{cases} \quad (7.99)$$

Para que sea un número difuso se necesita verificar la condición de la ecuación 7.96. Para ello escribimos las siguientes desigualdades:

$$\begin{cases} b - a &= 6(\mathbf{V}_A(1, 0.5) - \mathbf{V}_A(0, 0.5)) - 3(\mathbf{A}_A(1) - \mathbf{A}_A(0)) \geq 0 \\ c - b &= -2\mathbf{A}_A(0) + 3\mathbf{A}_A(1) \geq 0 \\ d - c &= -6(\mathbf{V}_A(1, 0.5) - \mathbf{V}_A(0, 0.5)) - 3(\mathbf{A}_A(1) - \mathbf{A}_A(0)) \geq 0 \end{cases} \quad (7.100)$$

La primera y la tercera desigualdad se pueden formular así:

$$\begin{cases} 6\left(\frac{\mathbf{V}_A(1, 0) + \mathbf{V}_A(1, 1)}{2} - \frac{\mathbf{V}_A(0, 0) + \mathbf{V}_A(0, 1)}{2}\right) \\ -3((\mathbf{V}_A(1, 1) - \mathbf{V}_A(1, 0)) - (\mathbf{V}_A(0, 1) - \mathbf{V}_A(0, 0))) \geq 0 \\ -6\left(\frac{\mathbf{V}_A(1, 0) + \mathbf{V}_A(1, 1)}{2} - \frac{\mathbf{V}_A(0, 0) + \mathbf{V}_A(0, 1)}{2}\right) \\ -3((\mathbf{V}_A(1, 1) - \mathbf{V}_A(1, 0)) - (\mathbf{V}_A(0, 1) - \mathbf{V}_A(0, 0))) \geq 0 \end{cases} \quad (7.101)$$

Lo que conduce a

$$\begin{cases} \mathbf{V}_A(1, 0) - \mathbf{V}_A(0, 0) \geq 0 \\ -\mathbf{V}_A(1, 1) + \mathbf{V}_A(0, 1) \geq 0 \end{cases} \quad (7.102)$$

Las dos desigualdades de 7.102 las satisface todo número difuso (Ecuación 7.77). Además, como las ambigüedades son positivas, la única restricción resultante se puede reducir a:

$$\frac{\mathbf{A}_A(1)}{\mathbf{A}_A(0)} \geq \frac{2}{3} \quad (7.103)$$

Si no se satisface la desigualdad 7.103 el resultado de aplicar la ecuación 7.98 son cuatro números que no cumplen el requisito $a \leq b \leq c \leq d$ y por tanto con ellos no se puede construir un número difuso trapezoidal. En ese caso, es necesario relajar las condiciones del criterio para encontrar unas nuevas condiciones tales que se pueda satisfacer 7.103. Para hacerlo, es necesario tener en cuenta que la primera y tercera desigualdad conducen a

$$2|\mathbf{V}_A(1, 0.5) - \mathbf{V}_A(0, 0.5)| \geq |\mathbf{A}_A(1) - \mathbf{A}_A(0)| \quad (7.104)$$

Y, por tanto, si se relaja alguna de las condiciones de ambigüedad también debe verificarse la desigualdad 7.104. Con eso en mente, al menos se pueden encontrar las siguientes opciones para relajar las condiciones del criterio:

- **Opción 1:** mantener la ambigüedad para $\gamma = 1$ y modificar la de $\gamma = 0$ así:

$$A_A(0) \leftarrow \frac{3}{2}A_A(1) \quad (7.105)$$

Si no se cumple 7.104 hacer:

$$\begin{aligned}\Delta_V &\leftarrow \operatorname{sgn}(V_A(0, 0.5) - V_A(1, 0.5)) \\ V_A(0, 0.5) &\leftarrow V_A(1, 0.5) + \Delta_V \frac{|\mathbf{A}_A(1) - \mathbf{A}_A(0)|}{2}\end{aligned}\quad (7.106)$$

- **Opción 2:** mantener la ambigüedad para $\gamma = 0$ y modificar la de $\gamma = 1$ así:

$$A_A(1) \leftarrow \frac{2}{3}A_A(0) \quad (7.107)$$

Si no se cumple 7.104 hacer:

$$\begin{aligned}\Delta_V &\leftarrow \operatorname{sgn}(V_A(0, 0.5) - V_A(1, 0.5)) \\ V_A(1, 0.5) &\leftarrow V_A(0, 0.5) - \Delta_V \frac{|\mathbf{A}_A(1) - \mathbf{A}_A(0)|}{2}\end{aligned}\quad (7.108)$$

El algoritmo 4 verifica si se cumplen las condiciones de la ecuación 7.103. Si no se cumplen prueba las dos opciones anteriores y selecciona aquella con la que se obtiene la menor distancia entre A y X .

7.7.3. Conjunto de criterios 2

En esta sección exploramos el siguiente criterio de semejanza: Dado un número difuso A , su representación trapezoidal X debe ser tal que:

- el menor valor representativo de X sea igual al de A , al menos para dos valores de γ diferentes
- el mayor valor representativo de X sea igual al de A , al menos para dos valores de γ diferentes

El menor y el mayor valor representativo corresponden a el menor y el mayor nivel de optimismo, respectivamente. En otras palabras, corresponden a $\beta = 0$ y $\beta = 1$. Por ello, la primera condición implica $V_X(\gamma, 0) = V_A(\gamma, 0)$ en tanto que la segunda implica $V_X(\gamma, 1) = V_A(\gamma, 1)$. Al aplicar 7.89 para $\beta \in \{0, 1\}$ y $\gamma \in \{0, 1\}$ se obtiene:

$$\begin{bmatrix} \mathbf{V}_A(0, 0) \\ \mathbf{V}_A(0, 1) \\ \mathbf{V}_A(1, 0) \\ \mathbf{V}_A(1, 1) \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 \\ 1/3 & 2/3 & 0 & 0 \\ 0 & 0 & 2/3 & 1/3 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (7.109)$$

Algoritmo 4 Trapecio equivalente por valores y ambigüedades

Entradas: N un número difuso

```

1: procedure TRAPECIOEQUIVALENTEVALORAMBIGÜEDAD( $N$ )
2:    $V_0 \leftarrow \mathbf{V}_N(0, 0.5)$                                  $\triangleright$  Ecuación 7.18 con  $\gamma = 0$  y  $\beta = 0.5$ 
3:    $A_0 \leftarrow \mathbf{A}_N(0)$                                  $\triangleright$  Ecuación 7.37 con  $\gamma = 0$ 
4:    $V_1 \leftarrow \mathbf{V}_N(1, 0.5)$                                  $\triangleright$  Ecuación 7.18 con  $\gamma = 1$  y  $\beta = 0.5$ 
5:    $A_1 \leftarrow \mathbf{A}_N(1)$                                  $\triangleright$  Ecuación 7.37 con  $\gamma = 1$ 
6:   Construir la matriz  $M$ 
7:    $Y \leftarrow [V_0, A_0, V_1, A_1]^T$                              $\triangleright$  Ver ecuación 7.98
8:    $\Delta_V = V_1 - V_0$ 
9:    $\Delta_A = A_1 - A_0$ 
10:  if  $A_1/A_0 < 2/3$  then
11:     $Y_1 \leftarrow [V_0, \tilde{A}_0, V_1, A_1]^T$                    $\triangleright$  Ecuación 7.105 y 7.106
12:     $Y_2 \leftarrow [V_0, A_0, V_1, \tilde{A}_1]^T$                    $\triangleright$  Ecuación 7.107 y 7.108
13:     $X_1 \leftarrow M \times Y_1$ 
14:     $X_2 \leftarrow M \times Y_2$ 
15:     $d_1 \leftarrow \mathbf{d}(A, X_1)$                                  $\triangleright$  Ecuación 7.48
16:     $d_2 \leftarrow \mathbf{d}(A, X_2)$                                  $\triangleright$  Ecuación 7.48
17:    if  $d_1 \leq d_2$  then
18:       $Y \leftarrow Y_1$ 
19:    else
20:       $Y \leftarrow Y_2$ 
21:     $[a, b, c, d]^T \leftarrow M \times Y$                              $\triangleright$  Ecuación 7.98
22:  return  $a, b, c, d$                                           $\triangleright$  Vértices del trapecio equivalente

```

Despejando se consigue una expresión de la forma que tiene la ecuación 7.92:

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 4 & 0 & -3 & 0 \\ -2 & 0 & 3 & 0 \\ 0 & -2 & 0 & 3 \\ 0 & 4 & 0 & -3 \end{bmatrix} \begin{bmatrix} \mathbf{V}_A(0,0) \\ \mathbf{V}_A(0,1) \\ \mathbf{V}_A(1,0) \\ \mathbf{V}_A(1,1) \end{bmatrix} \quad (7.110)$$

Además, se sabe que:

$$\begin{cases} \mathbf{V}_A(0,0) \leq \mathbf{V}_A(0,1) & \text{Extremo inferior menor o igual que el superior} \\ \mathbf{V}_A(1,0) \leq \mathbf{V}_A(1,1) & \text{Extremo inferior menor o igual que el superior} \\ \mathbf{V}_A(0,0) \leq \mathbf{V}_A(1,0) & L_A(\alpha) \text{ es no decreciente} \\ \mathbf{V}_A(1,1) \leq \mathbf{V}_A(0,1) & R_A(\alpha) \text{ es no creciente} \end{cases} \quad (7.111)$$

Para verificar la condición de la ecuación 7.96 escribimos las siguientes desigualdades:

$$\begin{cases} b - a = -6\mathbf{V}_A(0,0) + 6\mathbf{V}_A(1,0) \geq 0 \\ c - b = -2(\mathbf{V}_A(0,1) - \mathbf{V}_A(0,0)) + 3(\mathbf{V}_A(1,1) - \mathbf{V}_A(1,0)) \geq 0 \\ d - c = 6\mathbf{V}_A(0,1) - 6\mathbf{V}_A(1,1) \geq 0 \end{cases} \quad (7.112)$$

La primera y la tercera condición se satisfacen para todo número difuso. La segunda condición puede reescribirse así

$$\frac{\mathbf{V}_A(1,1) - \mathbf{V}_A(1,0)}{\mathbf{V}_A(0,1) - \mathbf{V}_A(0,0)} \geq \frac{2}{3} \quad (7.113)$$

El numerador y el denominador de la ecuación 7.113 son las ambigüedades de A para $\gamma = 1$ y $\gamma = 0$, respectivamente. Eso nos permite reescribir las condiciones como cualquiera de las siguientes expresiones, que son equivalentes:

$$\begin{cases} \frac{\mathbf{A}_A(1)}{\mathbf{A}_A(0)} \geq \frac{2}{3} \\ \mathbf{A}_A(0) - \mathbf{A}_A(1) \leq \frac{\mathbf{A}_A(0)}{3} \\ \mathbf{A}_A(0) - \mathbf{A}_A(1) \leq \frac{\mathbf{A}_A(1)}{2} \end{cases} \quad (7.114)$$

Si no se satisface la condición que hemos encontrado el resultado de aplicar la ecuación 7.110 son cuatro números que no cumplen el requisito $a \leq b \leq c \leq d$ y por tanto con ellos

no se puede construir un número difuso trapezoidal. En ese caso, es necesario relajar las condiciones del criterio para encontrar unas nuevas condiciones tales que se pueda satisfacer 7.113. Al menos se pueden encontrar las siguientes opciones para hacerlo:

- **Opción 1:** mantener las condiciones para $\gamma = 1$ y modificar las de $\gamma = 0$ que se requieran así:

$$\begin{aligned}\Delta_A &\leftarrow \frac{\mathbf{V}_A(1,1) - \mathbf{V}_A(1,0)}{2} \\ \Delta_L &\leftarrow \mathbf{V}_A(1,0) - \mathbf{V}_A(0,0) \\ \Delta_R &\leftarrow \mathbf{V}_A(0,1) - \mathbf{V}_A(1,1) \\ \mathbf{V}_A(0,0) &\leftarrow \mathbf{V}_A(1,0) - \Delta_A \frac{\Delta_L}{\Delta_L + \Delta_R} \\ \mathbf{V}_A(0,1) &\leftarrow \mathbf{V}_A(1,1) + \Delta_A \frac{\Delta_R}{\Delta_L + \Delta_R}\end{aligned}\tag{7.115}$$

- **Opción 2:** mantener las condiciones para $\gamma = 0$ y modificar las de $\gamma = 1$ que se requieran así:

$$\begin{aligned}\Delta_A &\leftarrow \frac{\mathbf{V}_A(0,1) - \mathbf{V}_A(0,0)}{3} \\ \Delta_L &\leftarrow \mathbf{V}_A(1,0) - \mathbf{V}_A(0,0) \\ \Delta_R &\leftarrow \mathbf{V}_A(0,1) - \mathbf{V}_A(1,1) \\ \mathbf{V}_A(1,0) &\leftarrow \mathbf{V}_A(0,0) + \Delta_A \frac{\Delta_L}{\Delta_L + \Delta_R} \\ \mathbf{V}_A(1,1) &\leftarrow \mathbf{V}_A(0,1) - \Delta_A \frac{\Delta_R}{\Delta_L + \Delta_R}\end{aligned}\tag{7.116}$$

En ambas opciones Δ_A es el cambio más grande que se puede tener en la ambigüedad al cambiar γ , de acuerdo con las condiciones de la ecuación 7.114. Este cambio máximo se ajusta moviendo los extremos izquierdo y derecho de los valores extremos del número A para el γ seleccionado. La cantidad que se desplazan se distribuye proporcionalmente a Δ_L y Δ_R , que son las cantidades previstas en el problema originalmente planteado.

El algoritmo 5 verifica si se cumple la condición de la ecuación 7.113. Si no se cumplen prueba las dos opciones anteriores y selecciona aquella con la que se obtiene la menor distancia entre A y X .

7.7.4. Conjunto de criterios 3

En esta sección exploramos el siguiente criterio de semejanza: Dado un número difuso A , su representación trapezoidal X debe ser tal que:

Algoritmo 5 Trapecio equivalente por valores extremos

Entradas: N un número difuso

```

1: procedure TRAPECIOEQUIVALENTEVALORESEXTREMOS( $N$ )
2:    $V_{0,0} \leftarrow \mathbf{V}_N(0, 0)$                                  $\triangleright$  Ecuación 7.18 con  $\gamma = 0$  y  $\beta = 0$ 
3:    $V_{0,1} \leftarrow \mathbf{V}_N(0, 1)$                                  $\triangleright$  Ecuación 7.18 con  $\gamma = 0$  y  $\beta = 1$ 
4:    $V_{1,0} \leftarrow \mathbf{V}_N(1, 0)$                                  $\triangleright$  Ecuación 7.18 con  $\gamma = 1$  y  $\beta = 0$ 
5:    $V_{1,1} \leftarrow \mathbf{V}_N(1, 1)$                                  $\triangleright$  Ecuación 7.18 con  $\gamma = 1$  y  $\beta = 1$ 
6:   Construir la matriz  $M$                                           $\triangleright$  Ver ecuación 7.110
7:    $Y \leftarrow [V_{0,0}, V_{0,1}, V_{1,0}, V_{1,1}]^T$ 
8:   if  $V_{1,1} - V_{1,0} < 2(V_{0,1} - V_{0,0})/3$  then
9:      $Y_1 \leftarrow [V_{0,0}, \tilde{V}_{0,1}, V_{1,0}, V_{1,1}]^T$            $\triangleright$  Ecuación 7.115
10:     $Y_2 \leftarrow [V_{0,0}, V_{0,1}, \tilde{V}_{1,0}, \tilde{V}_{1,1}]^T$          $\triangleright$  Ecuación 7.116
11:     $X_1 \leftarrow M \times Y_1$ 
12:     $X_2 \leftarrow M \times Y_2$ 
13:     $d_1 \leftarrow \mathbf{d}(A, X_1)$                                       $\triangleright$  Ecuación 7.48
14:     $d_2 \leftarrow \mathbf{d}(A, X_2)$                                       $\triangleright$  Ecuación 7.48
15:    if  $d_1 \leq d_2$  then
16:       $Y \leftarrow Y_1$ 
17:    else
18:       $Y \leftarrow Y_2$ 
19:     $[a, b, c, d]^T \leftarrow M \times Y$                                 $\triangleright$  Ecuación 7.110
20:    return  $a, b, c, d$                                           $\triangleright$  Vértices del trapecio equivalente

```

- La tendencia del valor representativo central de X debe ser igual a la de A
- La tendencia de la ambigüedad de X debe ser igual a la de A

Para estudiar los valores representativos centrales de todos los α -cortes del trapecio X tomamos $\beta = 0.5$ y aplicamos la ecuación 7.81 que se transforma en

$$\begin{cases} \mathbf{v}_A(\alpha, 0.5) = \frac{a+d}{2} + \alpha \left(\frac{b+c}{2} - \frac{a+d}{2} \right) = m_v \alpha + b_v \\ \mathbf{a}_A(\alpha) = (d-a) + \alpha ((c-b) - (d-a)) = m_a \alpha + b_a \end{cases} \quad (7.117)$$

En donde se observa que tanto el valor representativo como la ambigüedad dependen de α siguiendo un patrón de líneas rectas cuyas pendientes y términos independientes son:

$$\begin{cases} m_v = \frac{b+c}{2} - \frac{a+d}{2} \\ b_v = \frac{a+d}{2} \\ m_a = (c-b) - (d-a) \\ b_a = d-a \end{cases} \quad (7.118)$$

Que en forma matricial se escribe así:

$$\begin{bmatrix} m_v \\ b_v \\ m_a \\ b_a \end{bmatrix} = \begin{bmatrix} -1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & 0 & 0 & 1/2 \\ 1 & -1 & 1 & -1 \\ -1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \quad (7.119)$$

Despejando se consigue una expresión de la forma que tiene la ecuación 7.92:

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & -1/2 \\ 1 & 1 & -1/2 & -1/2 \\ 1 & 1 & 1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 \end{bmatrix} \begin{bmatrix} m_v \\ b_v \\ m_a \\ b_a \end{bmatrix} \quad (7.120)$$

Para aplicar este criterio, los valores m_v, b_v, m_a, b_a se estiman mediante ajustes de línea recta sobre los valores representativos y las ambigüedades de los α -cortes, lo que minimiza el error cuadrático. Con esta estimación se puede aplicar 7.120 directamente para obtener los vértices del trapecio X .

Para verificar la condición de la ecuación 7.96 escribimos las siguientes desigualdades:

$$\begin{cases} b - a = m_v - m_a/2 \geq 0 \\ c - b = m_a + b_a \geq 0 \\ d - c = -m_v - m_a/2 \geq 0 \end{cases} \quad (7.121)$$

Utilizando la ecuación 7.74 se puede escribir

$$\begin{cases} m_v = 6\mathbf{V}_A(1, 0.5) - 6\mathbf{V}_A(0, 0.5) \\ m_a = 6\mathbf{A}_A(1) - 6\mathbf{A}_A(0) \end{cases} \quad (7.122)$$

y de allí

$$\begin{cases} m_v = 6\frac{\mathbf{V}_A(1, 0) + \mathbf{V}(1, 1)}{2} - 6\frac{\mathbf{V}_A(0, 0) + \mathbf{V}(0, 1)}{2} \\ m_a = 6(\mathbf{V}_A(1, 1) - \mathbf{V}_A(1, 0)) - 6(\mathbf{V}_A(0, 1) - \mathbf{V}_A(0, 0)) \end{cases} \quad (7.123)$$

Y por tanto la primera y la tercera desigualdad de 7.121 se pueden escribir así:

$$\begin{cases} m_v - \frac{m_a}{2} = 6\mathbf{V}_A(1, 0) - 6\mathbf{V}_A(0, 0) \geq 0 \\ -m_v - \frac{m_a}{2} = 6\mathbf{V}_A(0, 1) - 6\mathbf{V}_A(1, 1) \geq 0 \end{cases} \quad (7.124)$$

Estas condiciones se cumplen para todo número difuso (ecuación 7.77). Por lo tanto, la única restricción restante es

$$m_a + b_a \geq 0 \quad (7.125)$$

Si no se satisface la desigualdad 7.125 no se satisface la condición 7.96. En ese caso, es necesario relajar las condiciones del criterio. Al menos se pueden encontrar las siguientes opciones para hacerlo:

- **Opción 1:** mantener la pendiente y variar el término independiente:

$$b_a \leftarrow -m_a \quad (7.126)$$

- **Opción 2:** mantener el término independiente y variar la pendiente:

$$m_a \leftarrow -b_a \quad (7.127)$$

El algoritmo 6 verifica si se cumple la condición de la ecuación 7.113. Si no se cumplen prueba las dos opciones anteriores y selecciona aquella con la que se obtiene la menor distancia entre A y X .

Algoritmo 6 Trapecio equivalente por tendencias del Valor central y la Ambigüedad

Entradas:

N un número difuso

```

1: procedure TRAPECIOEQUIVALENTEAJUSTESVALORAMBIGÜEDAD( $N$ )
2:    $m_V, b_V \leftarrow$  Regresión( $\mathbf{v}_N(\alpha, 0.5)$ )            $\triangleright$  Regresión lineal con  $\beta = 0.5$ 
3:    $m_A, b_A \leftarrow$  Regresión( $\mathbf{a}_N(\alpha)$ )            $\triangleright$  Regresión lineal
4:   Construir la matriz  $M$             $\triangleright$  Ver ecuación 7.120
5:    $Y \leftarrow [m_v, b_v, m_a, b_a]^T$ 
6:   if  $|m_V| > |m_A/2|$  OR  $b_a < -m_a$  then
7:      $Y_1 \leftarrow [m_v, b_v, m_a, \tilde{b}_a]^T$             $\triangleright$  Ecuación 7.126
8:      $Y_2 \leftarrow [m_v, b_v, \tilde{m}_a, b_a]^T$             $\triangleright$  Ecuación 7.127
9:      $X_1 \leftarrow M \times Y_1$ 
10:     $X_2 \leftarrow M \times Y_2$ 
11:     $d_1 \leftarrow \mathbf{d}(A, X_1)$             $\triangleright$  Ecuación 7.48
12:     $d_2 \leftarrow \mathbf{d}(A, X_2)$             $\triangleright$  Ecuación 7.48
13:    if  $d_1 \leq d_2$  then
14:       $Y \leftarrow Y_1$ 
15:    else
16:       $Y \leftarrow Y_2$ 
17:     $[a, b, c, d]^T \leftarrow M \times Y$             $\triangleright$  Ecuación 7.120
18:    return  $a, b, c, d$             $\triangleright$  Vértices del trapecio equivalente

```

7.7.5. Conjunto de criterios 4

En esta sección exploramos el siguiente criterio de semejanza: Dado un número difuso A , su representación trapezoidal X debe ser tal que:

- La tendencia de su curva $L(\alpha)$ sea igual a la de A .
- La tendencia de su curva $R(\alpha)$ sea igual a la de A .

El enfoque de este criterio consiste en construir las rampas izquierda y derecha del trapecio mediante dos ajustes de línea recta: para la rampa izquierda se ajusta la curva $L(\alpha)$ y para la derecha se ajusta la curva $R(\alpha)$:

$$\begin{cases} L_A(\alpha) = \mathbf{a}_A(\alpha, 0) = m_L\alpha + b_L \\ R_A(\alpha) = \mathbf{a}_A(\alpha, 1) = m_R\alpha + b_R \end{cases} \quad (7.128)$$

La obtención de los vértices del trapecio es directa, porque corresponde a los valores 0 y 1 de α , es decir:

$$\begin{cases} a = b_L \\ b = m_L + b_L \\ c = m_R + b_R \\ d = b_R \end{cases} \quad (7.129)$$

Lo que permite obtener una expresión de la forma que tiene la ecuación 7.92

$$\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} m_L \\ b_L \\ m_R \\ b_R \end{bmatrix} \quad (7.130)$$

Para verificar la condición de la ecuación 7.96 escribimos las siguientes desigualdades:

$$\begin{cases} b - a = m_L \geq 0 \\ c - b = (m_R + b_R) - (m_L + b_L) \geq 0 \\ d - c = -m_R \geq 0 \end{cases} \quad (7.131)$$

En el ajuste se obtienen pendientes $m_L \geq 0$ y $m_R \leq 0$, debido a que las curvas $L(\alpha)$ y $R(\alpha)$ son no decrecientes y no crecientes, respectivamente. Por lo anterior se deduce que la única condición que hay que confirmar para que sea un número difuso es:

$$m_R + b_R \geq m_L + b_L \quad (7.132)$$

Si no se satisface la condición de la ecuación 7.132 el resultado de aplicar la ecuación 7.130 son cuatro números que no cumplen el requisito $a \leq b \leq c \leq d$ y por tanto con ellos no se puede construir un número difuso trapezoidal. En ese caso, es necesario relajar las condiciones del criterio para encontrar unas nuevas condiciones tales que sé se pueda satisfacer 7.132. Al menos se pueden encontrar las siguientes opciones para hacerlo:

- **Opción 1:** mantener las pendientes y ajustar los términos independientes.

$$\begin{aligned} \Delta_b &\leftarrow (m_L - m_R) - (b_R - b_L) \\ b_L &\leftarrow b_L - \Delta_b/2 \\ b_R &\leftarrow b_L + \Delta_b/2 \end{aligned} \quad (7.133)$$

- **Opción 2:** mantener los términos independientes y ajustar las pendientes.

$$\begin{aligned} \Delta_m &\leftarrow (b_L - b_R) - (m_R - m_L) \\ m_L &\leftarrow m_L - \Delta_m/2 \\ m_R &\leftarrow m_L + \Delta_m/2 \end{aligned} \quad (7.134)$$

El algoritmo 7 verifica si se cumplen las condiciones de la ecuación 7.132. Si no se cumplen prueba las dos opciones anteriores y selecciona aquella con la que se obtiene la menor distancia entre A y X .

Algoritmo 7 Trapecio equivalente por rampas laterales

Entradas:

N un número difuso

```

1: procedure TRAPECIOEQUIVALENTEVALORRAMPASLATERALES( $N$ )
2:    $m_L, b_L \leftarrow$  Regresión( $v_N(\alpha, 0)$ )            $\triangleright$  Regresión lineal con  $\beta = 0, h(\alpha) = 1$ 
3:    $m_R, b_R \leftarrow$  Regresión( $A_N(\alpha)$ )            $\triangleright$  Regresión lineal con  $h(\alpha) = 1$ 
4:   Construir la matriz  $M$                             $\triangleright$  Ver ecuación 7.130
5:    $Y \leftarrow [m_L, b_L, m_R, b_R]^T$ 
6:   if  $(m_R + b_R) < (m_L + b_L)$  then
7:      $Y_1 \leftarrow [m_L, \tilde{b}_L, m_R, \tilde{b}_R]^T$             $\triangleright$  Ecuación 7.126
8:      $Y_2 \leftarrow [\tilde{m}_L, b_L, \tilde{m}_R, b_R]^T$             $\triangleright$  Ecuación 7.127
9:      $X_1 \leftarrow M \times Y_1$ 
10:     $X_2 \leftarrow M \times Y_2$ 
11:     $d_1 \leftarrow \mathbf{d}(A, X_1)$                           $\triangleright$  Ecuación 7.48
12:     $d_2 \leftarrow \mathbf{d}(A, X_2)$                           $\triangleright$  Ecuación 7.48
13:    if  $d_1 \leq d_2$  then
14:       $Y \leftarrow Y_1$ 
15:    else
16:       $Y \leftarrow Y_2$ 
17:     $[a, b, c, d]^T \leftarrow M \times Y$             $\triangleright$  Ecuación 7.130
18:    return  $a, b, c, d$                           $\triangleright$  Vértices del trapecio equivalente

```

Ejemplo 7.9 (Aproximaciones trapezoidales de un número difuso) En este ejemplo se obtienen diferentes aproximaciones trapezoidales del número difuso de la Figura 7.7, que ya ha sido estudiado en los ejemplos 7.3, 7.5 y 7.7.

Se trata de un número difuso N definido por

$$\begin{aligned}
L_N(\alpha) &= p + (q - p)\alpha \\
R_N(\alpha) &= s - (s - r)\sqrt{\alpha} \\
p &\leq q \leq r \leq s \\
p &= -2, q = 0, r = 0, s = 1
\end{aligned} \tag{7.135}$$

Al aplicar los cinco criterios 0, 1, 2, 3 y 4 presentados en las secciones 7.7.1 a 7.7.5 se obtienen cinco números trapezoidales X_0 , X_1 , X_2 , X_3 y X_4 .

Ese número difuso también ha sido estudiado en [4]; allí se obtiene una aproximación trapezoidal utilizando un método que minimiza la distancia mientras que mantiene un valor representativo y la ambigüedad. La aproximación que allí se consigue es el número difuso:

$$X_d = T(-29/15, -1/30, -1/30, 2/3) \quad (7.136)$$

Las seis aproximaciones trapezoidales se han graficado en la Figura 7.12. En cada caso los resultados (en azul) se han graficado sobre el número original A (en rojo), lo que permite visualizar la bondad de la aproximación. Por su parte, la Tabla 7.6 muestra los vértices de sus funciones de pertenencia; allí se puede constatar que los métodos producen resultados diferentes, aunque cercanos.

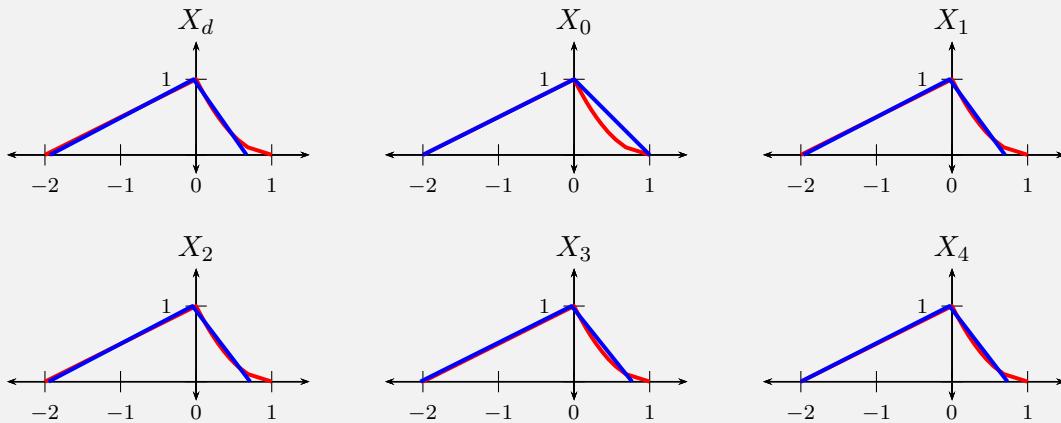


Figura 7.12: Aproximaciones trapezoidales del ejemplo 7.9

Tabla 7.6: Vértices de las aproximaciones trapezoidales del ejemplo 7.9

Número	a	b	c	d
X_d	-1.933	-0.033	-0.033	0.667
X_0	-2.000	0.000	0.000	1.000
X_1	-1.967	-0.033	-0.033	0.700
X_2	-1.952	-0.048	-0.048	0.714
X_3	-2.033	-0.033	-0.033	0.767
X_4	-2.000	-0.033	-0.033	0.734

Para comparar los seis resultados obtenidos se ha preparado la Tabla 7.7. En ella se compara el valor representativo central \bar{V} y la ambigüedad A para dos valores de γ , así como la distancia d al número original N .

En esa tabla puede verse que la peor aproximación (la de la mayor distancia) es X_0 , correspondiente al criterio 0. Esto no sorprende porque, tal como se menciona en la sección 7.7.1, es una aproximación ingenua que no considera ni el valor representativo ni la ambigüedad.

Lo que sí sorprende es que las aproximaciones X_1 a X_4 resulten mejores o iguales que X_d . La sorpresa radica en que X_d es una aproximación que minimiza la distancia. ¿Cómo es posible esto? La respuesta está en las restricciones. Mientras que el procedimiento para obtener X_d asegura que el valor representativo y la ambigüedad de la aproximación sean iguales a la del número original (usando $\gamma = 1$), los otros procedimientos utilizan otras condiciones y en ocasiones las relajan.

Tabla 7.7: Propiedades de las aproximaciones trapezoidales del ejemplo 7.9

Número	$\gamma = 0$		$\gamma = 1$		d
	V	A	V	A	
N	-0.333	1.333	-0.233	0.867	0.000
X_d	-0.333	1.300	-0.233	0.867	0.067
X_0	-0.250	1.500	-0.167	1.000	0.183
X_1	-0.333	1.333	-0.233	0.889	0.055
X_2	-0.333	1.333	-0.238	0.889	0.056
X_3	-0.333	1.400	-0.233	0.933	0.067
X_4	-0.333	1.367	-0.233	0.911	0.055

Capítulo 8

Aritmética intervalar

La aritmética intervalar es el área que estudia cómo realizar operaciones aritméticas entre intervalos. Más exactamente, se preocupa por determinar un intervalo en el que, con total certeza, estará el resultado de una cierta operación aritmética, aun cuando esta se haya realizado con incertidumbre.

La aritmética intervalar nace en los años 60 con el propósito de estudiar la incertidumbre de los cálculos realizados por computadores, es decir, por máquinas con una precisión limitada [44]. Esta idea ha evolucionado a lo largo del tiempo hasta tal punto que hoy contamos inclusive con un estándar sobre cómo implementar en software la aritmética intervalar (véase [36, 48]).

Mediante la aritmética intervalar se pueden abordar el estudio de los problemas derivados del redondeo, el truncamiento y la incertidumbre en los datos (véase [33]). También se encuentran aplicaciones en campos tan variados como el control, el análisis de sistemas de potencia, la investigación de operaciones, el planeamiento, etc.

En este capítulo introducimos los rudimentos de la aritmética intervalar, como un paso necesario para entender la aritmética difusa, ya que esta se basa en la aritmética de sus α -cortes y, como se muestra en la sección 7.1, los α -cortes de los números difusos son intervalos cerrados. Una presentación más amplia puede encontrarse en [45, 47].

8.1. Definiciones básicas

Partimos de la definición convencional de *intervalo cerrado* como un subconjunto de \mathbb{R} definido por dos números reales $L \leq R$ denominados los *límites inferior* y *superior* (o *límites izquierdo* y *derecho*), respectivamente, así:

$$A = [L_A, R_A] = \{x \in \mathbb{R} \mid L_A \leq x \leq R_A\} \quad (8.1)$$

Una observación importante: la aritmética intervalar se ha construido para determinar un intervalo en el que, con total certeza, estará el resultado de una cierta operación aritmética. Por esta razón trabaja sobre intervalos cerrados. En esta sección, aunque se omita el adjetivo *cerrado*, se sobreentiende que el intervalo lo es.

En el caso especial en que $L = R$ se dice que el intervalo es *degenerado*. Un intervalo degenerado contiene un único elemento x , lo que permite utilizar la siguiente notación:

$$x = [x, x] \quad (8.2)$$

Dos intervalos A, B son iguales si y solo si son el mismo conjunto. Para que esto suceda sus límites deben ser iguales:

$$A = B \Leftrightarrow L_A = L_B \text{ y } R_A = R_B \quad (8.3)$$

La *intersección* de dos intervalos A, B puede ser vacía. Sus límites permiten determinar si esto sucede:

$$A \cap B = \emptyset \Leftrightarrow R_A < L_B \text{ ó } R_B < L_A \quad (8.4)$$

En caso de que no sea vacía, el resultado

$$\begin{aligned} A \cap B &= \{x \mid x \in A \text{ y } x \in B\} \\ A \cap B &= [\max\{L_A, L_B\}, \min\{R_A, R_B\}] \end{aligned} \quad (8.5)$$

Si la intersección de dos intervalos no es vacía, entonces su *unión* también es un intervalo:

$$\begin{aligned} A \cup B &= \{x \mid x \in A \text{ ó } x \in B\} \\ A \cup B &= [\min\{L_A, L_B\}, \max\{R_A, R_B\}] \quad \text{si } A \cap B \neq \emptyset \end{aligned} \quad (8.6)$$

Sin embargo, la unión de dos intervalos puede no ser un intervalo. Es decir, la unión no es clausurativa en el conjunto de los intervalos.

Para subsanar la no clausuratividad de la unión, se define el *intervalo hull*¹ que sí es lo es y puede usarse en cálculos de intervalos:

$$A \underline{\cup} B = [\min\{L_A, L_B\}, \max\{R_A, R_B\}] \quad (8.7)$$

Para establecer una relación de ordenamiento, se define que un intervalo es menor que otro si todos los elementos del primero son menores que los del segundo, es decir:

$$\begin{aligned} A < B &\Leftrightarrow \{x < y \mid \forall x \in A \text{ y } \forall y \in B\} \\ A < B &\Leftrightarrow R_A < L_B \end{aligned} \quad (8.8)$$

¹Una posible traducción del término en inglés *interval hull* es ‘intervalo envoltorio’

Y a partir de allí se define la relación mayor que:

$$A < B \Leftrightarrow B > A \quad (8.9)$$

Estas relaciones de ordenamiento resultan ser transitiva:

$$\begin{aligned} A < B \quad \text{y} \quad B < C &\Rightarrow A < C \\ A > B \quad \text{y} \quad B > C &\Rightarrow A > C \end{aligned} \quad (8.10)$$

Utilizando la notación de la ecuación 8.2 sabemos que $0 = [0, 0]$, lo que nos permite decir que un intervalo A es *positivo* si $A > 0$ y que es *negativo* si $A < 0$.

8.2. Extensión de funciones de una variable

Dada una función $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ el resultado de aplicar $f(\cdot)$ a un intervalo A se define como el conjunto de los resultados de aplicar $f(\cdot)$ a todos los elementos de A :

$$f(A) = \{f(x) \mid x \in A\} \quad (8.11)$$

Esta notación permite escribir

$$f(x) = f([x, x]) = \{f(x) \mid x \in [x, x]\} = f(x) \quad (8.12)$$

Es decir, la definición de la ecuación 8.11 implica que hay una relación uno a uno entre la función $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ y su aplicación en intervalos degenerados. En ese sentido, se extiende el uso de $f(\cdot)$ de los reales a los intervalos.

El sentido de la ecuación 8.11 se ilustra en la Figura 8.1a. La función $f(x)$ se caracteriza por la curva que se ha trazado; el trozo dibujado en azul corresponde a los puntos $(x, f(x))$ para todos los valores de x en el intervalo A . La proyección de estos puntos sobre el eje vertical resulta ser $f(A)$.

En la Figura 8.1b se muestra que el resultado de aplicar $f(\cdot)$ a un intervalo A no necesariamente es un intervalo. En esa figura $f(x)$ presenta una discontinuidad y como consecuencia $f(A)$ es la unión de dos intervalos.

8.3. Extensión de funciones de una variable continuas y monótonas

La definición de continuidad de una función implica que si $f(x)$ es continua en el intervalo cerrado A entonces el resultado de aplicar $f(\cdot)$ en todos los puntos de A es también

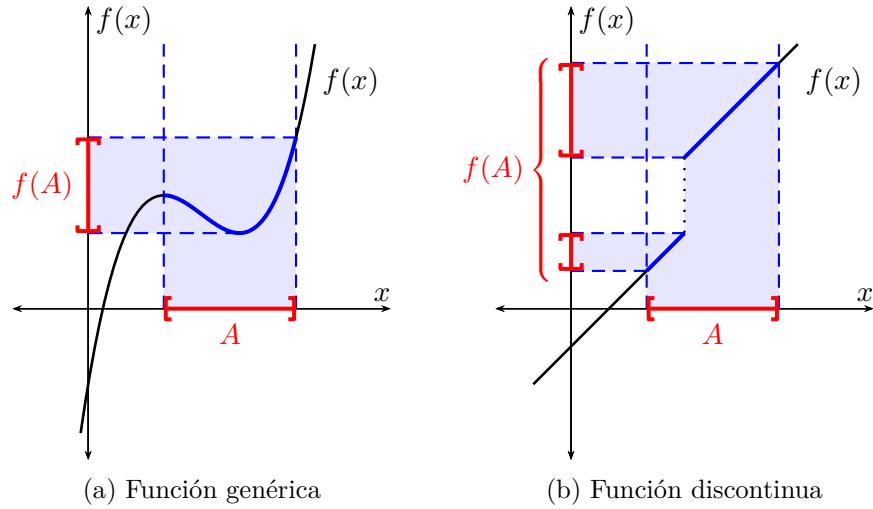


Figura 8.1: Extensión de funciones de una variable a intervalos

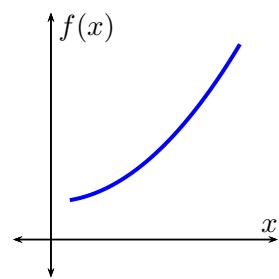
un intervalo cerrado. Si, además, la función satisface alguna condición de monotonía, entonces es particularmente sencillo calcular los límites de dicho intervalo.

Supongamos una función $f(x) : \mathbb{R} \rightarrow \mathbb{R}$. Se tienen las siguientes definiciones de monotonía: $f(x)$ es

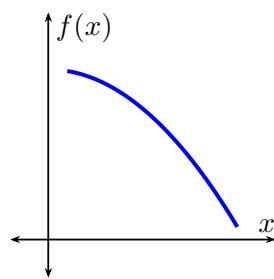
$$\begin{cases} \text{monótonamente creciente} & \text{si } f(x_1) > f(x_2) \quad \forall x_1 > x_2 \\ \text{monótonamente decreciente} & \text{si } f(x_1) < f(x_2) \quad \forall x_1 > x_2 \\ \text{monótonamente no-decreciente} & \text{si } f(x_1) \geq f(x_2) \quad \forall x_1 > x_2 \\ \text{monótonamente no-creciente} & \text{si } f(x_1) \leq f(x_2) \quad \forall x_1 > x_2 \end{cases} \quad (8.13)$$

Estas definiciones se ilustran en la Figura 8.2. La monotonía creciente (Figura 8.2a) implica que al incrementar el valor de x también se incrementa el valor de $f(x)$. La monotonía no-decreciente (Figura 8.2c) es menos exigente, porque permite que al incrementar el valor de x permanezca constante el valor de $f(x)$, lo importante es que no disminuya. Algo semejante puede decirse de las funciones decrecientes y no-crecientes (Figuras 8.2b y 8.2d). Como comparación, la función de la Figura 8.1a no es monótona porque en ciertos tramos crece y en otros decrece.

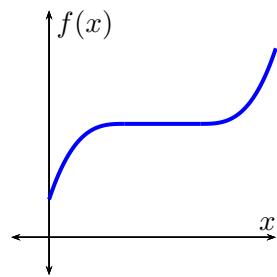
En caso de que $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ sea continua y derivable, la definición de la ecuación 8.13 puede reformularse en términos de las derivadas. $f(x)$ es



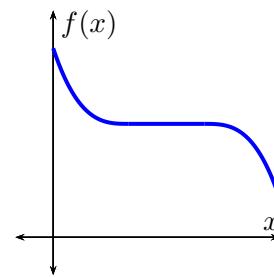
(a) Monotonía creciente



(b) Monotonía decreciente



(c) Monotonía no-decreciente



(d) Monotonía no-creciente

Figura 8.2: Monotonía de una función

$$\left\{ \begin{array}{ll} \text{monótonamente creciente} & \text{si } \frac{df}{dx} > 0 \quad \forall x \\ \text{monótonamente decreciente} & \text{si } \frac{df}{dx} < 0 \quad \forall x \\ \text{monótonamente no-decreciente} & \text{si } \frac{df}{dx} \geq 0 \quad \forall x \\ \text{monótonamente no-creciente} & \text{si } \frac{df}{dx} \leq 0 \quad \forall x \end{array} \right. \quad (8.14)$$

Recapitulando, tenemos dos hechos importantes: y 1) Si $f(x)$ es continua, sabemos que al aplicarla en un intervalo el resultado será otro intervalo. 2) Si $f(x)$ es monótona, podemos saber qué valores de x generan los valores más pequeños y más grandes de $f(x)$.

Estos dos hechos se combinan en el Teorema 8.1:

Teorema 8.1 Extensión de funciones continuas y monótonas

1. *Sea $A = [L_A, R_A]$ un intervalo cerrado y $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ una función continua monótonamente no-decreciente en A . En esas condiciones:*

$$f(A) = [f(L_A), f(R_A)] \quad (8.15)$$

2. *Sea $A = [L_A, R_A]$ un intervalo cerrado y $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ una función continua monótonamente no-creciente en A . En esas condiciones:*

$$f(A) = [f(R_A), f(L_A)] \quad (8.16)$$

Ejemplo 8.1 (Extensión intervalar de e^x) Consideremos el cálculo de $f(x) = e^x$ en el intervalo $[-1, 1]$.

La Figura 8.3 muestra la gráfica de la función $f(x)$. Nótese que es una función continua y creciente, por lo que satisface las condiciones del Teorema 8.1 y podemos aplicar la ecuación 8.15:

$$\begin{aligned} f(A) &= [f(L_A), f(R_A)] \\ &= [f(-1), f(1)] \\ &= [e^{-1}, e^1] \\ &= [0.368, 2.718] \end{aligned} \quad (8.17)$$

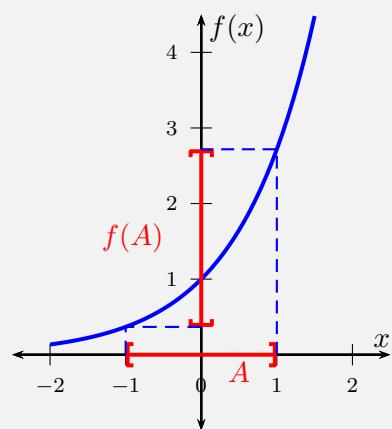


Figura 8.3: Función del ejemplo 8.1

La Figura 8.3 también muestra que para obtener $f(A)$ basta con calcular el valor de $f(x)$ en los extremos del A .

Ejemplo 8.2 (Extensión intervalar de e^{-x}) Consideremos el cálculo de $f(x) = e^{-x}$ en el intervalo $[-1, 1]$.

La Figura 8.4 muestra la gráfica de la función $f(x)$. Nótese que es una función continua y decreciente, por lo que satisface las condiciones del Teorema 8.1 y podemos aplicar la ecuación 8.16:

$$\begin{aligned} f(A) &= [f(R_A), f(L_A)] \\ &= [f(1), f(-1)] \\ &= [e^{-1}, e^1] \\ &= [0.368, 2.718] \end{aligned} \tag{8.18}$$

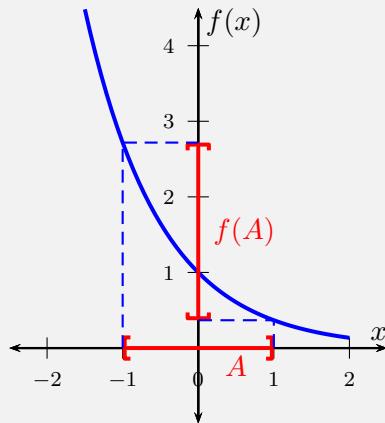


Figura 8.4: Función del ejemplo 8.2

La Figura 8.4 también muestra que para obtener $f(A)$ basta con calcular el valor de $f(x)$ en los extremos del A . Al compararla con la Figura 8.3 puede observarse que ha sido necesario utilizar los extremos contrarios, debido a que la función es decreciente.

Ejemplo 8.3 (Extensión intervalar de $\sin(x)$) Consideremos el cálculo de $f(x) = \sin(x)$ en el intervalo $[-1, 1]$.

La Figura 8.5 muestra la gráfica de la función $f(x)$. Nótese que es una función continua pero no es monótona. Aparentemente no cumpliría las condiciones del Teorema 8.1. No obstante ese teorema establece que la función debe ser monótona en el *intervalo de evaluación* que en este caso es $[-1, 1]$. La función $\sin(x)$ sí es monótonamente creciente en el intervalo $[-\pi/2, \pi/2]$ y por tanto es creciente en $[-1, 1]$. Eso nos permite aplicar la ecuación 8.15:

$$\begin{aligned} f(A) &= [f(L_A), f(R_A)] \\ &= [f(-1), f(1)] \\ &= [\sin(-1), \sin(1)] \\ &= [-0.841, 0.841] \end{aligned} \tag{8.19}$$

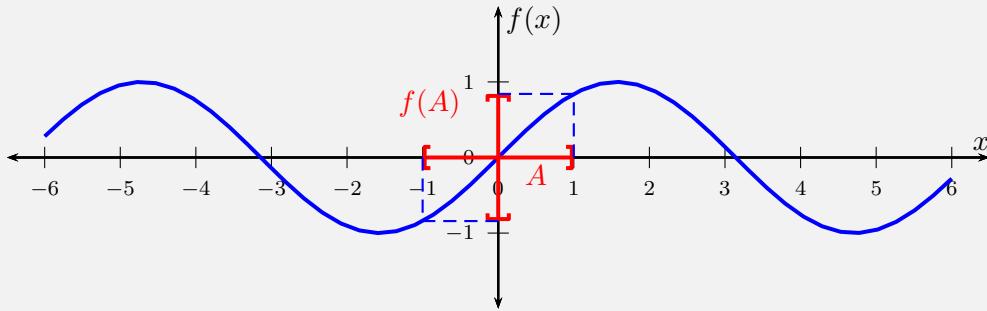


Figura 8.5: Función del ejemplo 8.3

8.4. Extensión de funciones multivariable

Consideremos ahora una función de múltiples variables $f(x_1, x_2, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$. Nos interesa aplicar $f(\cdot)$ a un conjunto de n intervalos A_1, A_2, \dots, A_n . El es será el conjunto de los resultados de aplicar $f(\cdot)$ en todas las posibles combinaciones de valores de los argumentos:

$$f(A_1, A_2, \dots, A_n) = \{f(x_1, x_2, \dots, x_n) \mid x_1 \in A_1, x_2 \in A_2, \dots, x_n \in A_n\} \tag{8.20}$$

Que también puede escribirse de forma más compacta utilizando una notación vectorial:

$$\begin{aligned}\mathbf{x} &= \{x_1, x_2, \dots, x_n\} \\ \mathbf{A} &= A_1 \times A_2 \times \dots \times A_n \\ f(\mathbf{A}) &= \{f(\mathbf{x}) \mid \mathbf{x} \in \mathbf{A}\}\end{aligned}\tag{8.21}$$

Ejemplo 8.4 (Extensión de funciones multivariable. Suma) Consideremos la función que suma dos números:

$$\begin{aligned}z &= f(x_1, x_2) : \mathbb{R}^2 \rightarrow \mathbb{R} \\ z &= f(x_1, x_2) = x_1 + x_2\end{aligned}\tag{8.22}$$

Supongamos que se desea aplicar esa función a los intervalos $A_1 = [1, 3]$ y $A_2 = [2, 3]$. En la Figura 8.6 se muestra en color azul la superficie de la función, y en color magenta el resultado de aplicar $f(\cdot)$ en todas las parejas de puntos definidas por los intervalos.

Como la superficie es continua, todos los valores están entre el más pequeño y el más grande, que resultan ser 3 y 6, respectivamente. Para ilustrarlo se han trazado los planos $z = 3$ y $z = 6$.

En conclusión,

$$f(A_1, A_2) = [3, 6]\tag{8.23}$$

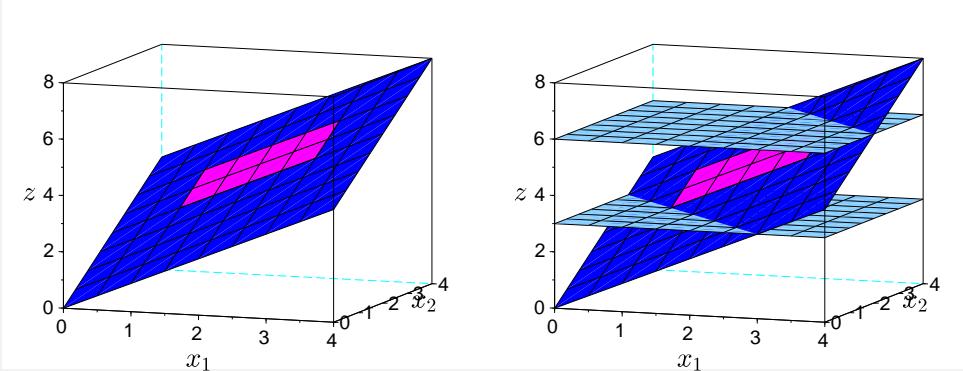


Figura 8.6: Función del ejemplo 8.4

Ejemplo 8.5 (Extensión de funciones multivariable. Producto) Consideremos la función que multiplica dos números:

$$\begin{aligned}z &= f(x_1, x_2) : \mathbb{R}^2 \rightarrow \mathbb{R} \\ z &= f(x_1, x_2) = x_1 x_2\end{aligned}\tag{8.24}$$

Supongamos que se desea aplicar esa función a los intervalos $A_1 = [-1, 2]$ y $A_2 = [-1, 2]$. En la Figura 8.7 se muestra en color azul la superficie de la función, y en color magenta el resultado de aplicar $f(\cdot)$ en todas las parejas de puntos definidas por los intervalos.

Como la superficie es continua, todos los valores están entre el más pequeño y el más grande, que resultan ser -2 y 4 , respectivamente. Para ilustrarlo se han trazado los planos $z = -2$ y $z = 4$.

En conclusión,

$$f(A_1, A_2) = [-2, 4] \quad (8.25)$$

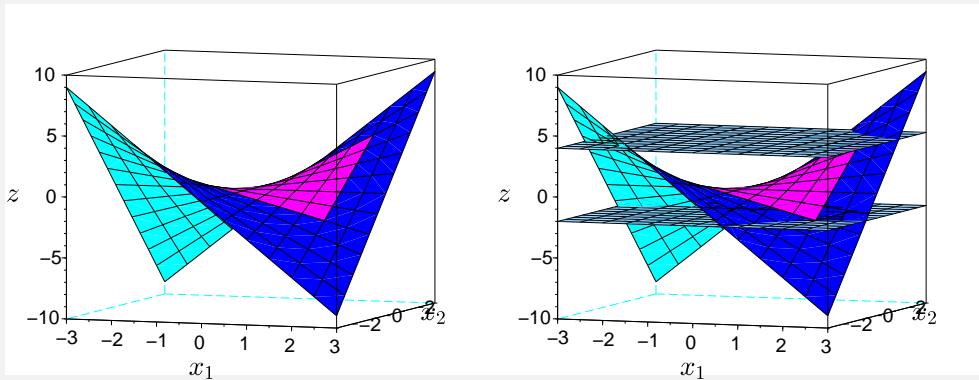


Figura 8.7: Función del ejemplo 8.5

8.5. Monotonía de funciones multivariable

Para analizar la monotonía de $f(\cdot)$, consideremos primero la k -ésima variable. Definamos un vector \mathbf{x}_k^o en el que todas las variables, salvo la k -ésima, son fijas:

$$\mathbf{x}_k^o = \{x_1^o, x_2^o, \dots, x_{k-1}^o, x_k, x_{k+1}^o, \dots, x_n^o\}$$

Podemos definir una función $f_k(\mathbf{x}_k^o) : \mathbb{R} \rightarrow \mathbb{R}$ para cada posible combinación de los $n-1$ valores fijos. Estas funciones sólo varían en la dirección de la variable x_k y por lo tanto podemos estudiar su monotonía siguiendo las definiciones de la ecuación 8.13: $f(\cdot)$ es monótonamente:

$$\left\{ \begin{array}{lll} \text{creciente en } x_k & \text{si } f(\mathbf{x}_k^o) \text{ es monótonamente creciente} & \forall \mathbf{x}_k^o \\ \text{decreciente en } x_k & \text{si } f(\mathbf{x}_k^o) \text{ es monótonamente decreciente} & \forall \mathbf{x}_k^o \\ \text{no-decreciente en } x_k & \text{si } f(\mathbf{x}_k^o) \text{ es monótonamente no-decreciente} & \forall \mathbf{x}_k^o \\ \text{no-creciente en } x_k & \text{si } f(\mathbf{x}_k^o) \text{ es monótonamente no-creciente} & \forall \mathbf{x}_k^o \end{array} \right. \quad (8.26)$$

En caso de que $f(\cdot)$ sea continua y derivable, la monotonía en una variable también puede definirse a partir de las derivadas parciales.

$$\left\{ \begin{array}{lll} \text{creciente en } x_k & \text{si } \frac{\partial f}{\partial x_k} > 0 & \forall \mathbf{x} \\ \text{decreciente en } x_k & \text{si } \frac{\partial f}{\partial x_k} < 0 & \forall \mathbf{x} \\ \text{no-decreciente en } x_k & \text{si } \frac{\partial f}{\partial x_k} \geq 0 & \forall \mathbf{x} \\ \text{no-creciente en } x_k & \text{si } \frac{\partial f}{\partial x_k} \leq 0 & \forall \mathbf{x} \end{array} \right. \quad (8.27)$$

La monotonía en una variable implica que para todo el espacio en que está definida $f(\cdot)$, al efectuar un cambio en la variable k se puede saber, con certeza, el sentido del cambio de $f(\cdot)$: aumentará, disminuirá, no disminuirá o no aumentará, dependiendo el tipo de monotonía.

El siguiente paso es establecer una definición de monotonía en todas las variables:

Sea $f(x_1, x_2, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$ una función. $f(\cdot)$ es una función *monótona* si para cada una de sus n variables satisface alguna de las cuatro condiciones de monotonía de la ecuación la ecuación 8.26.

En tal caso, tiene sentido definir la variable auxiliar $d_f(k)$ que indica el sentido de la monotonía de cada variable como:

$$d_f(k) = \begin{cases} 0 & \text{si } f \text{ es monótonamente creciente o no-decreciente en } x_k \\ 1 & \text{si } f \text{ es monótonamente decreciente o no-creciente en } x_k \end{cases} \quad (8.28)$$

Podemos ser más restrictivos y es tablecere que $f(\cdot)$ es una función *estrictamente monótona* si para cada una de sus n variables satisface alguna de las dos primeras condiciones de monotonía de la ecuación la ecuación 8.26.

Ejemplo 8.6 (Monotonía de un hiperplano) Consideremos la función que define un hiperplano en \mathbf{R}^4 :

$$f(x_1, x_2, \dots, x_n) = 2 + 3x_1 - 2x_2 + 5x_3 \quad (8.29)$$

Las derivadas parciales son fáciles de calcular y corresponden a los coeficientes de cada variable:

$$\frac{\partial f}{\partial x_1} = 3 \quad \frac{\partial f}{\partial x_2} = -2 \quad \frac{\partial f}{\partial x_3} = 5 \quad (8.30)$$

Al revisar la ecuación 8.27 se comprueba que:

- $f(\cdot)$ es monótonamente creciente en x_1
- $f(\cdot)$ es monótonamente decreciente en x_2
- $f(\cdot)$ es monótonamente creciente en x_3

Y por tanto podemos afirmar que $f(\cdot)$ es monótona. Nótese que no es necesario que la monotonía sea igual en todas las variables. De hecho, se tiene que:

$$d_f(1) = 0 \quad d_f(2) = 1 \quad d_f(3) = 0 \quad (8.31)$$

Ejemplo 8.7 (Monotonía del producto) Consideremos ahora la función que calcula el producto de dos números cuya gráfica se muestra en la Figura 8.7:

$$\begin{aligned} f(x_1, x_2) &: \mathbb{R}^2 \rightarrow \mathbb{R} \\ f(x_1, x_2) &= x_1 x_2 \end{aligned} \quad (8.32)$$

Las derivadas parciales resultan ser:

$$\frac{\partial f}{\partial x_1} = x_2 \quad \frac{\partial f}{\partial x_2} = x_1 \quad (8.33)$$

Es decir, los signos las derivadas parciales pueden ser positivas, negativas o cero. Por lo tanto, no se satisfacen las condiciones de la ecuación 8.27 en ninguna de las dos variables y por tanto la función no es monótona.

Sin embargo, si se restringe la función a alguno de los cuatro cuadrantes la situación cambia. Por ejemplo, si se restringe al primer cuadrante ($x_1 \geq 0$ y $x_2 \geq 0$) la nueva función es

$$\begin{aligned} f(x_1, x_2) &: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R} \\ f(x_1, x_2) &= x_1 x_2 \end{aligned} \quad (8.34)$$

Las derivadas parciales se calculan de la misma forma, pero ahora podemos asegurar que no son negativas. En consecuencia podemos afirmar que en el primer cuadrante

$$\begin{aligned} f(\cdot) &\text{ es monótonamente no-decreciente en } x_1 \\ f(\cdot) &\text{ es monótonamente no-decreciente en } x_2 \end{aligned}$$

Y por tanto $f(\cdot)$ es monótona en el primer cuadrante con $d_f(1) = d_f(2) = 0$.

8.6. Extensión de funciones multivariable continuas y monótonas

Sea $f(x_1, x_2, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$ una función de varias variables. Consideremos ahora unos intervalos A_1, A_2, \dots, A_n , y su producto cartesiano $\mathbf{A} = A_1 \times A_2 \times \dots \times A_n$.

La definición de continuidad de una función implica que si $f(\cdot)$ es continua en \mathbf{A} entonces el resultado de aplicar $f(\cdot)$ en todos los puntos de \mathbf{A} es también un intervalo cerrado. Si, además, la función satisface alguna condición de monotonía, entonces es particularmente sencillo calcular los límites de dicho intervalo, como se explica a continuación.

Cada uno de los n intervalos se caracteriza por sus límites inferior y superior:

$$A_k = [L_{A_k}, R_{A_k}] \quad (8.35)$$

Vamos a construir dos conjuntos de n valores. El primero debe contener los valores de las n variables con los que se obtiene el menor valor de $f(\mathbf{A})$ y el segundo debe contener los los valores de las n variables con los que se obtiene el mayor valor de $f(\mathbf{A})$. Si denotamos esos conjuntos por \mathcal{L}_{A_k} y \mathcal{R}_{A_k} respectivamente, lo que estamos diciendo es que

$$f(\mathbf{A}) = [f(\mathcal{L}_A), f(\mathcal{R}_A)] \quad (8.36)$$

Llamemos \mathcal{L}_{A_k} el valor de la k -ésima variable que debe estar en \mathcal{L}_A . De forma análoga, llamemos \mathcal{R}_{A_k} el valor de la k -ésima variable que debe estar en \mathcal{R}_A . Es decir:

$$\begin{aligned} \mathcal{L}_A &= \{\mathcal{L}_{A_1}, \mathcal{L}_{A_2}, \dots, \mathcal{L}_{A_n}\} \\ \mathcal{R}_A &= \{\mathcal{R}_{A_1}, \mathcal{R}_{A_2}, \dots, \mathcal{R}_{A_n}\} \end{aligned} \quad (8.37)$$

¿Cuáles son los valores de \mathcal{L}_{A_k} y de \mathcal{R}_{A_k} ? Depende de la monotonía de la variable x_k . Si x_k es no-decreciente, \mathcal{L}_{A_k} debe ser el límite inferior y \mathcal{R}_{A_k} debe ser el límite superior.

Pero si Si x_k es no-creciente los papeles se invierten es decir, \mathcal{L}_{A_k} debe ser el límite superior y \mathcal{R}_{A_k} el límite inferior.

El razonamiento anterior nos permite formular el Teorema 8.2.

Teorema 8.2 Extensión de funciones multivariable continuas y monótonas
Sea $f(x_1, x_2, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}$ una función continua y monótona. Sean A_1, A_2, \dots, A_n unos intervalos. El resultado de aplicar $f(\cdot)$ en dichos intervalos es otro intervalo:

$$f(A_1, A_2, \dots, A_n) = [f(\mathcal{L}_A), f(\mathcal{R}_A)] \quad (8.38)$$

En donde

$$\begin{aligned} \mathcal{L}_A &= \{\mathcal{L}_{A_1}, \mathcal{L}_{A_2}, \dots, \mathcal{L}_{A_n}\} \\ \mathcal{R}_A &= \{\mathcal{R}_{A_1}, \mathcal{R}_{A_2}, \dots, \mathcal{R}_{A_n}\} \end{aligned} \quad (8.39)$$

Los elementos de \mathcal{L}_A y \mathcal{R}_A se obtienen a partir de los límites de cada intervalo $A_k = [L_{A_k}, R_{A_k}]$ y de la función auxiliar $d_f(k)$ que indica la monotonía de la variable x_k así:

$$\mathcal{L}_{A_k} = \begin{cases} L_{A_k} & \text{si } d_f(k) = 0 \\ R_{A_k} & \text{si } d_f(k) = 1 \end{cases} \quad \mathcal{R}_{A_k} = \begin{cases} R_{A_k} & \text{si } d_f(k) = 0 \\ L_{A_k} & \text{si } d_f(k) = 1 \end{cases} \quad (8.40)$$

ó, lo que es igual:

$$\begin{aligned} \mathcal{L}_{A_k} &= (1 - d_f(k)) L_{A_k} + d_f(k) R_{A_k} \\ \mathcal{R}_{A_k} &= (1 - d_f(k)) R_{A_k} + d_f(k) L_{A_k} \end{aligned} \quad (8.41)$$

Ejemplo 8.8 (Extensión de funciones multivariable monótonas. Suma)

Retomemos el ejemplo 8.4, en el que se aplica una función que suma dos números a los intervalos $A_1 = [1, 3]$ y $A_2 = [2, 3]$. La función es

$$\begin{aligned} f(x_1, x_2) &: \mathbb{R}^2 \rightarrow \mathbb{R} \\ f(x_1, x_2) &= x_1 + x_2 \end{aligned} \quad (8.42)$$

Claramente la función es monótonamente creciente en las dos variables. Esto implica que:

$$d_f(1) = d_f(2) = 0 \quad (8.43)$$

Y por tanto

$$\begin{aligned} \mathcal{L}_{A_1} &= L_{A_1} = 1 & \mathcal{R}_{A_1} &= R_{A_1} = 3 \\ \mathcal{L}_{A_2} &= L_{A_2} = 2 & \mathcal{R}_{A_2} &= R_{A_2} = 3 \end{aligned} \quad (8.44)$$

Podemos construir los conjuntos de los valores con los que se obtienen los menores y mayores valores de la función:

$$\mathcal{L}_A = \{1, 2\} \quad \mathcal{R}_A = \{3, 3\} \quad (8.45)$$

Y por lo tanto,

$$\begin{aligned} f(A_1, A_2) &= [f(1, 2), f(3, 3)] \\ f(A_1, A_2) &= [(1+2), (3+3)] \\ f(A_1, A_2) &= [3, 6] \end{aligned} \quad (8.46)$$

Ejemplo 8.9 (Extensión de funciones multivariable monótonas. Resta) En este ejemplo se aplica una función que resta dos números a los intervalos $A_1 = [1, 3]$ y $A_2 = [2, 3]$. La función es

$$\begin{aligned} f(x_1, x_2) &: \mathbb{R}^2 \rightarrow \mathbb{R} \\ f(x_1, x_2) &= x_1 - x_2 \end{aligned} \quad (8.47)$$

Claramente la función es monótonamente creciente la primera variable y decreciente en la segunda. Es decir:

$$d_f(1) = 0 \quad d_f(2) = 1 \quad (8.48)$$

Y por tanto

$$\begin{aligned} \mathcal{L}_{A_1} &= L_{A_1} = 1 & \mathcal{R}_{A_1} &= R_{A_1} = 3 \\ \mathcal{L}_{A_2} &= R_{A_2} = 3 & \mathcal{R}_{A_2} &= L_{A_2} = 2 \end{aligned} \quad (8.49)$$

Podemos construir los conjuntos de los valores con los que se obtienen los menores y mayores valores de la función:

$$\mathcal{L}_A = \{1, 3\} \quad \mathcal{R}_A = \{3, 2\} \quad (8.50)$$

Y por lo tanto,

$$\begin{aligned} f(A_1, A_2) &= [f(1, 3), f(3, 2)] \\ f(A_1, A_2) &= [(1-3), (3-2)] \\ f(A_1, A_2) &= [-2, 1] \end{aligned} \quad (8.51)$$

Ejemplo 8.10 (Extensión de funciones multivariable monótonas. Hiperplano)
Consideremos ahora la función que define un hiperplano

$$\begin{aligned} f(x_1, x_2) & : \mathbb{R}^4 \rightarrow \mathbb{R} \\ f(x_1, x_2) & = 2x_1 - 3x_2 + x_3 - 2x_4 + 1 \end{aligned} \quad (8.52)$$

Que se desea aplicar a los siguientes intervalos:

$$A_1 = [0, 2] \quad A_2 = [1, 2] \quad A_3 = [3, 5] \quad A_4 = [4, 6] \quad (8.53)$$

Tal como se ilustra en el ejemplo 8.6, la monotonía depende del signo de los coeficientes de cada variable, es decir:

$$d_f(1) = 0 \quad d_f(2) = 1 \quad d_f(3) = 0 \quad d_f(4) = 1 \quad (8.54)$$

Y por tanto

$$\begin{array}{llll} \mathcal{L}_{A_1} = L_{A_1} = 0 & \mathcal{R}_{A_1} = R_{A_1} = 2 \\ \mathcal{L}_{A_2} = R_{A_2} = 2 & \mathcal{R}_{A_2} = L_{A_2} = 1 \\ \mathcal{L}_{A_3} = L_{A_3} = 3 & \mathcal{R}_{A_3} = R_{A_3} = 5 \\ \mathcal{L}_{A_4} = R_{A_4} = 6 & \mathcal{R}_{A_4} = L_{A_4} = 4 \end{array} \quad (8.55)$$

Podemos construir los conjuntos de los valores con los que se obtienen los menores y mayores valores de la función:

$$\mathcal{L}_A = \{0, 2, 3, 6\} \quad \mathcal{R}_A = \{2, 1, 5, 4\} \quad (8.56)$$

Y por lo tanto,

$$\begin{aligned} f(\mathbf{A}) & = [f(0, 2, 3, 6), f(2, 1, 5, 4)] \\ f(\mathbf{A}) & = [(2 \cdot 0) - (3 \cdot 2) + 3 - (2 \cdot 6) + 1, (2 \cdot 2) - (3 \cdot 1) + 5 - (2 \cdot 4) + 1] \\ f(\mathbf{A}) & = [-14, -1] \end{aligned} \quad (8.57)$$

Es decir, $f(\mathbf{A}) = [-14, -1]$

8.7. Operaciones aritméticas sobre intervalos

En esta sección exploramos en más detalle la extensión de las operaciones básicas de suma, resta, multiplicación, división y potencia a los intervalos. Para ello, se aplicarán los principios y resultados de las secciones 8.2 a 8.6 y se analizarán los resultados.

8.7.1. Suma y resta de intervalos

Sean A y B dos intervalos. Definimos los operadores de suma $+$ y resta $-$ como la extensión de las operaciones convencionales $f(x_1, x_2) = x_1 + x_2$ y $g(x_1 - x_2) = x_1 - x_2$. Estas operaciones son monótonas y el sentido de la monotonía está determinado por el signo de cada variable. Por lo tanto, al aplicar el teorema 8.2 se obtiene

$$\begin{aligned} A + B &= [L_A + L_B, R_A + R_B] \\ A - B &= [L_A - R_B, R_A - L_B] \end{aligned} \quad (8.58)$$

Nótese que la ecuación $A - A = 0$ es incorrecta. En efecto, para que el resultado sea $0 = [0, 0]$ se necesita que $L_A = R_A$, es decir, que el intervalo sea degenerado. Por tanto, para cualquier intervalo no degenerado se tiene que $A - A \neq 0$.

8.7.2. Producto de intervalos

Sean A y B dos intervalos. Definimos el operador de producto \cdot como la extensión de la operación convencional $f(x_1, x_2) = x_1 x_2$. En el dominio de \mathbf{R}^2 esta operación no es monótona (Véase el ejemplo 8.5) y por lo tanto no es posible aplicar el teorema 8.2. Sin embargo, dado que la función sí es continua, podemos asegurar que el resultado es un intervalo definido por

$$\begin{aligned} A \cdot B &= [\min\{S\}, \max\{S\}] \\ S &= \{L_A L_B, L_A R_B, R_A L_B, R_A R_B\} \end{aligned} \quad (8.59)$$

Es decir, para definir los límites del resultado basta con calcular los cuatro productos de los límites de los intervalos y determinar los valores mínimo y máximo.

Sin embargo, también es posible calcular $A \cdot B$ revisando los signos de los límites de los intervalos y efectuando sólo dos productos. La Tabla 8.1 muestra cuáles deben ser las parejas de valores que deben multiplicarse para obtener los límites inferior y superior del producto. Cada uno de los 9 casos puede visualizarse en la Figura 8.8 como un rectángulo, que corresponde al producto cartesiano $A \times B$.

El caso 1 corresponde a intervalos positivos, es decir, $A \geq 0$ y $B \geq 0$. En otras palabras, el caso 1 corresponde a la situación en la que $A \times B$ reside en el primer cuadrante de

Tabla 8.1: Producto intervalar. Casos basados en los límites

	Caso		L	R	Ejemplo	
1)	$0 \leq L_A$	y	$0 \leq L_B$	$L_A L_B$	$R_A R_B$	$[1, 2][3, 4] = [3, 8]$
2)	$R_A \leq 0$	y	$0 \leq L_B$	$L_A R_B$	$R_A L_B$	$[-2, -1][3, 4] = [-8, -3]$
3)	$R_A \leq 0$	y	$R_B \leq 0$	$R_A R_B$	$L_A L_B$	$[-2, -1][-4, -3] = [3, 8]$
4)	$0 \leq L_A$	y	$R_B \leq 0$	$R_A L_B$	$L_A R_B$	$[1, 2][-4, -3] = [-8, -3]$
5)	$L_A \leq 0 \leq R_A$	y	$0 \leq L_B$	$L_A R_B$	$R_A R_B$	$[-1, 2][3, 4] = [-4, 8]$
6)	$L_A \leq 0 \leq R_A$	y	$R_B \leq 0$	$R_A L_B$	$L_A L_B$	$[-1, 2][-4, -3] = [-8, 3]$
7)	$0 \leq L_A$	y	$L_B \leq 0 \leq R_B$	$R_A L_B$	$R_A R_B$	$[1, 2][-3, 4] = [-6, 8]$
8)	$R_A \leq 0$	y	$L_B \leq 0 \leq R_B$	$L_A R_B$	$L_A L_B$	$[-2, -1][-3, 4] = [-8, 6]$
9)	$L_A \leq 0 \leq R_A$	y	$L_B \leq 0 \leq R_B$	$\min S_L$	$\max S_R$	$[-1, 2][-3, 4] = [-6, 8]$
				$S_L = \{L_A R_B, R_A L_B\}$		
				$S_R = \{L_A L_B, R_A R_B\}$		

\mathbb{R}^2 . En el ejemplo 8.5 se ilustra que en esta situación la función sí es monótona y por tanto puede aplicarse el Teorema 8.2.

De forma análoga, los casos 2, 3 y 4 corresponden a situaciones en las que $A \times B$ reside en el segundo, tercer y cuarto cuadrante de \mathbb{R}^2 , respectivamente. En todos estos casos la función es monótona porque las derivadas parciales no cambian de signo. No obstante, el sentido de la monotonía de cada variable varía según el cuadrante.

En los casos 5 a 8 el rectángulo respectivo cruza uno de los dos ejes y por tanto se intersecta con dos cuadrantes, mientras que en el caso 9 se intersecta con los cuatro cuadrantes. Estas intersecciones hacen que las derivadas parciales cambien de signo y por tanto que la función no sea monótona.

El producto nos permite definir el operador de cambio de signo – como:

$$-A = -1 \cdot A = [-1, -1] \cdot A \quad (8.60)$$

Que en la Tabla 8.1 corresponde al caso 2, 3 ú 8. En cualquiera de esos casos se obtiene

$$-A = [-R_A, -L_A] \quad (8.61)$$

Con el operador – podemos hacer una definición alternativa de la resta como $A - B = A + (-B)$. En efecto:

$$\begin{aligned} A - B &= A + (-B) \\ A - B &= [L_A, R_A] + (-[L_B, R_B]) \\ A - B &= [L_A, R_A] + [-R_B, -L_B] \\ A - B &= [L_A - R_B, R_A - L_B] \end{aligned} \quad (8.62)$$

Que coincide con la definición de la ecuación 8.58.

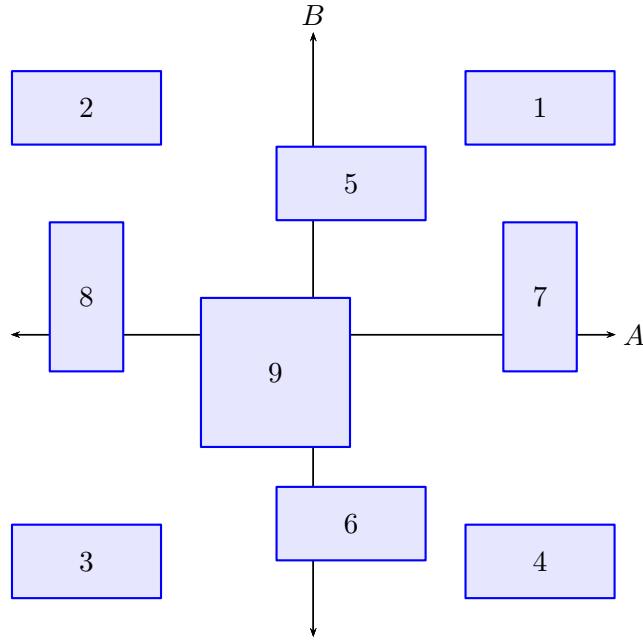


Figura 8.8: Producto intervalar. Visualización de los casos de la Tabla 8.1

8.7.2.1. División de intervalos

Para definir la división, primero definimos la operación $1/B$ como la extensión de la función $f(x) : \mathbb{R}^{-0} \rightarrow \mathbb{R}$ tal que $f(x) = 1/x$.

$$\frac{1}{B} = \left\{ \frac{1}{x} \mid \forall x \in B \right\} \quad (8.63)$$

Nótese que $f(x)$ está definida para los reales, excluyendo el cero. Por lo tanto, la operación $1/B$ sólo está definida para los intervalos que no contienen el 0, es decir para aquellos que satisfacen:

$$B < 0 \quad \text{ó} \quad B > 0 \quad (8.64)$$

Debido a que $1/x$ es una función decreciente en \mathbb{R}^{-0} , podemos aplicar el Teorema 8.1 para obtener

$$\frac{1}{B} = \left[\frac{1}{R_B}, \frac{1}{L_B} \right] \quad (8.65)$$

Podemos ahora definir la división de intervalos haciendo uso de la multiplicación así:

$$\frac{A}{B} = A \cdot \frac{1}{B} \quad (8.66)$$

Tabla 8.2: División intervalar. Casos basados en los límites

	Caso	L	R	Ejemplo
1)	$0 \leq L_A$ y $0 \leq L_B$	L_A/R_B	R_A/L_B	$[1, 2]/[3, 4] = [1/3, 2/4]$
2)	$R_A \leq 0$ y $0 \leq L_B$	L_A/L_B	R_A/R_B	$[-2, -1]/[3, 4] = [-2/3, -1/4]$
3)	$R_A \leq 0$ y $R_B \leq 0$	R_A/L_B	L_A/R_B	$[-2, -1]/[-4, -3] = [1/4, 2/3]$
4)	$0 \leq L_A$ y $R_B \leq 0$	R_A/R_B	L_A/L_B	$[1, 2]/[-4, -3] = [-2/3, -1/4]$
5)	$L_A \leq 0 \leq R_A$ y $0 \leq L_B$	L_A/L_B	R_A/L_B	$[-1, 2]/[3, 4] = [-1/3, 2/3]$
6)	$L_A \leq 0 \leq R_A$ y $R_B \leq 0$	R_A/R_B	L_A/R_B	$[-1, 2]/[-4, -3] = [-2/3, -1/3]$

Gracias a esta definición podemos construir la Tabla 8.2, en la que se muestra cuáles deben ser las parejas de valores que deben dividirse para obtener los límites inferior y superior de la división. Al comparar esta tabla con la del producto (Tabla 8.1) se observa que hay menos casos posibles: no aparecen los casos 7, 8 y 9, porque corresponden a intervalos que no satisfacen la ecuación 8.64 y por tanto para ellos no está definida la división.

Nótese que la ecuación $A/A = 1$ es incorrecta. En efecto, como dividendo y divisor son iguales, solo es posible que se de el caso 1 o el caso 3 de la Tabla 8.2, es decir

$$\frac{A}{A} = \begin{cases} \left[\frac{L_A}{R_A}, \frac{R_A}{L_A} \right] & \text{si es el caso 1} \\ \left[\frac{R_A}{L_A}, \frac{L_A}{R_A} \right] & \text{si es el caso 2} \end{cases} \quad (8.67)$$

En cualquiera de los dos casos, para que el resultado sea $1 = [1, 1]$ se necesita que $L_A = R_A$, es decir, que el intervalo sea degenerado. Por tanto, para cualquier intervalo no degenerado se tiene que $A/A \neq 1$.

8.7.3. Potencia de un intervalo

Consideramos ahora la función de potenciación. Dado un número natural n , es decir $n \in \mathbb{N}$, la función $f(x) : \mathbb{R} \rightarrow \mathbb{R}$ de potenciación se define como

$$f(x) = x^n = \underbrace{x \times x \times \cdots \times x}_{n \text{ veces}} \quad (8.68)$$

Esta definición se generaliza para el caso en que el exponente es un entero negativo

$$x^{-n} = \frac{1}{x^n} \quad n \in \mathbb{N} \quad (8.69)$$

o un número racional

$$x^{m/n} = \sqrt[n]{x^m} \quad m, n \in \mathbb{N} \quad (8.70)$$

Tabla 8.3: Potenciación intervalar. Casos basados en los límites

Caso	L	R	Ejemplo
1) $0 \leq L_A$	L_A^2	R_A^2	$[1, 2]^2 = [1^2, 2^2] = [1, 4]$
2) $R_A \leq 0$	R_A^2	L_A^2	$[-2, -1]^2 = [(-1)^2, (-2)^2] = [1, 4]$
3) $L_A \leq 0 \leq R_A$	0 $S_R = \{L_A^2, R_A^2\}$	$\max S_R$	$[-2, 1]^2 = [0, (-2)^2] = [0, 4]$

Sin embargo, esta última generalización *sólo es válida para bases positivas*, es decir $x > 0$

Una generalización válida para un exponente real y una base real positiva da origen a la función de *exponenciación*:

$$f(x) = x^y = e^{y \cdot \ln(x)} \quad x \in \mathbb{R}^+ \quad y \in \mathbb{R} \quad (8.71)$$

En esta sección estudiamos la extensión de la función de potenciación más sencilla, $f(x) = x^2$. El propósito específico de elegir esta función es el de demostrar que si A es un intervalo, entonces, en general²

$$A^2 \neq A \cdot A \quad (8.72)$$

La función $f(x) = x^2$ no es monótona porque su derivada es $2x$, lo que implica que su signo no es constante. La función corresponde a una parábola cuyo mínimo valor es $f(0) = 0$. La función es monótonamente decreciente para $x < 0$ y monótonamente creciente para $x > 0$.

La Tabla 8.3 muestra la forma de calcular $f(A) = A^2$ en función de los límites de A . El caso 1 corresponde a un intervalo positivo, y allí podemos aplicar la ecuación 8.15 porque la función es creciente. El caso 2 corresponde a un intervalo negativo, y allí podemos aplicar la ecuación 8.16 porque la función es decreciente.

En el caso 3 hay algunos valores negativos y otros positivos, lo que implica que el $0 \in A$ y por tanto $0 \in f(A)$. Como $f(A) \geq 0$, se demuestra que el límite inferior de $f(A)$ es 0. El límite superior del caso 3 viene dado por las monotonías parciales de la función para $x > 0$ y para $x < 0$.

Consideremos ahora las funciones:

$$\begin{cases} f_1(x) = \mathbb{R} \rightarrow \mathbb{R} \\ f_1(x) = x^2 \end{cases} \quad \begin{cases} f_2(x, y) = \mathbb{R}^2 \rightarrow \mathbb{R} \\ f_2(x, y) = x \cdot y \end{cases} \quad (8.73)$$

²'en general' significa aquí que para algunos intervalos sí se satisface que $x^2 = x \cdot x$, pero que es un error aseverar que se cumple para todos.

y el intervalo $A = [-1, 1]$. La aplicación de las dos funciones al intervalo A da como resultado:

$$\begin{aligned} f_1(A) &= A^2 &= \{x^2 \mid x \in A\} &= [0, 1] \\ f_2(A, A) &= A \cdot A &= \{x \cdot y \mid x \in A, y \in A\} &= [-1, 1] \end{aligned} \quad (8.74)$$

Con lo que se demuestra que $A^2 \neq A \cdot A$.

8.8. Algunas precauciones

Resumimos aquí, de forma compacta, algunos resultados que han sido previamente presentados. Lo hacemos con el propósito de destacar que algunos resultados obvios en los números reales no son ciertos al extender las funciones a los complejos. Esto implica que debemos tratar con precaución la aritmética intervalar, especialmente a la hora de despejar ecuaciones, aún las más simples³.

$$\begin{aligned} A + B - B &\neq A \\ A - A &\neq 0 \\ A/A &\neq 1 \\ A \cdot A &\neq A^2 \end{aligned} \quad (8.75)$$

³El significado del signo \neq es ‘en general no es igual’, es decir, es posible que para algunos intervalos específicos sí se cumpla la igualdad. Por ejemplo $[1, 1] - [1, 1] = [0, 0] = 0$

Capítulo 9

El principio de extensión

En el capítulo 8 se muestra cómo aplicar a los intervalos una función $f(\cdot)$ que ha sido originalmente definida para ser aplicada sobre números reales. Además, hemos denominado a ese proceso la *extensión de la función*.

Sin embargo, en este capítulo, y a partir de aquí en el resto del texto, al utilizar esa expresión nos referiremos a la aplicación a conjuntos difusos de funciones que han sido definidas para hacerlo sobre conjuntos concretos. La estrategia natural que utilizaremos es el *principio de extensión*, un principio establecido por Zadeh desde la definición primigenia de los conjuntos difusos [56].

Para presentar el principio, supongamos dos universos de discurso U , V y una función que a cada elemento de U asigna un elemento de V

$$y = f(x) : U \rightarrow V \quad (9.1)$$

Supongamos ahora que hemos definido un conjunto difuso A sobre U y que queremos aplicar la función f a ese conjunto. ¿Cómo calcular $f(A)$?

De acuerdo con el principio de extensión, el resultado de aplicar $f(A)$ será un conjunto difuso sobre V . Llamemos a ese conjunto B , es decir:

$$B = f(A) \quad (9.2)$$

Para definir B debemos conocer el grado de pertenencia de todos los elementos de V a ese conjunto. En otras palabras, debemos poder calcular $\mu_B(y)$. El principio de extensión determina que

$$\mu_B(y) = \max_{f(x)=y} \{\mu_A(x)\} \quad (9.3)$$

Es decir, para calcular $\mu_B(y)$ debemos identificar todos los valores de x tales que $f(x) = y$; después, debemos identificar cuál es el grado de pertenencia de todos esos valores al conjunto A y tomar el máximo grado. Ese será el grado de pertenencia de y al conjunto $B = f(A)$

Ejemplo 9.1 (Extensión de una función sobre un universo discreto)

Supongamos que el universo de discurso U es el de los deportes y el universo de discurso V es el de los animales. Supongamos también que hemos definido una función $y = f(x)$ como la que se muestra en la figura 9.1.

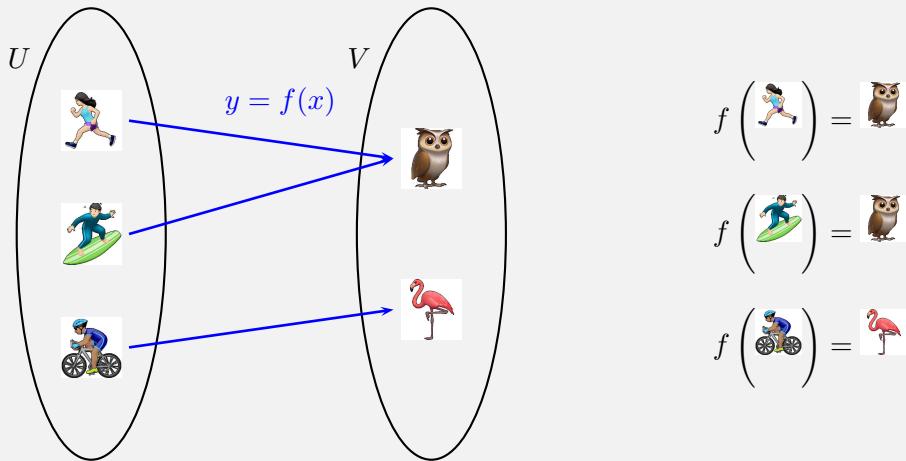


Figura 9.1: Función del ejemplo 9.1

f asigna un animal a cada deporte. En particular, asigna el buho al atletismo y al surf y asigna el flamenco al ciclismo.

Supongamos, además, que hemos construido el siguiente conjunto difuso A sobre U

$$A = \left\{ \frac{0.3}{\text{runner}} + \frac{0.5}{\text{surfer}} + \frac{0.2}{\text{cyclist}} \right\} \quad (9.4)$$

Lo que queremos ahora es calcular $B = f(A)$, mediante el principio de extensión. Necesitamos calcular μ_B para todos los elementos de V , es decir, para el buho y para el flamenco.

Consideraremos primero el buho. Los elementos de A a los que se le asigna el buho son el atletismo y el surf. Los grados de pertenencia de estos elementos a A son, respectivamente, 0.3 y 0.5. Por lo tanto

$$\mu_B \left(\begin{array}{c} \text{owl} \\ \text{flamingo} \end{array} \right) = \max \left\{ \mu_A \left(\begin{array}{c} \text{runner} \\ \text{cyclist} \end{array} \right), \mu_A \left(\begin{array}{c} \text{surfer} \\ \text{flamingo} \end{array} \right) \right\} = \max \{0.3, 0.5\} = 0.5 \quad (9.5)$$

Veamos ahora el flamenco. Solo hay un elemento de A al que se le asigna en flamenco. Es el ciclismo, y su grado de pertenencia a A es 0.2. Por lo tanto

$$\mu_B \left(\begin{array}{c} \text{flamingo} \end{array} \right) = \max \left\{ \mu_A \left(\begin{array}{c} \text{cyclist} \end{array} \right) \right\} = \max \{0.2\} = 0.2 \quad (9.6)$$

Combinando los dos resultados anteriores podemos escribir el conjunto B como

$$B = \left\{ \frac{0.5}{\text{owl}} + \frac{0.2}{\text{flamingo}} \right\} \quad (9.7)$$

9.1. Extensión de funciones sobre \mathbb{R}

Supongamos que ahora $f(x)$ opera sobre los reales, de tal manera que

$$y = f(x) : \mathbb{R} \rightarrow \mathbb{R} \quad (9.8)$$

Al aplicar el principio de extensión a funciones sobre conjuntos continuos como los reales es necesario efectuar una sutil modificación a la ecuación 9.3. La ecuación que debe aplicarse es:

$$\mu_B(y) = \sup_{f(x)=y} \{\mu_A(x)\} \quad (9.9)$$

Al comparar las dos ecuaciones es evidente que se ha sustituido el operador \max por el operador \sup . ¿Por qué? La respuesta la brinda el análisis real.

Para un conjunto S , el valor *supremo* (\sup), si existe, el valor más pequeña que sirve de cota superior al conjunto. Este valor puede o no formar parte del conjunto; si forma parte del conjunto se denomina *máximo* (\max).

Para ilustrar la diferencia consideremos los intervalos $I_1 = [0, 1]$ e $I_2 = [0, 1)$. Para ambos casos el valor 1 es el valor más pequeño que sirve de cota superior. Este valor si

está contenido en I_1 pero no lo está en I_2 . Por eso:

$$\left\{ \begin{array}{lcl} \sup \{[0, 1]\} & = & 1 \\ \max \{[0, 1]\} & = & 1 \\ \sup \{[0, 1)\} & = & 1 \\ \max \{[0, 1)\} & : & \text{no existe} \end{array} \right. \quad (9.10)$$

Ejemplo 9.2 (Principio de extensión de una función sobre \mathbb{R} uno a uno)

Consideremos la función $f(x)$ y el conjunto difuso A que se muestran en la Figura 9.2, que corresponden a:

$$f(x) = \frac{x^3}{6} \quad (9.11)$$

$$A : T(-0.5, 0.5, 1.0, 2.0)$$

¿Cuál es el resultado de aplicar $f(x)$ al conjunto A ? En otras palabras ¿Cuál es el resultado de $f(A)$?

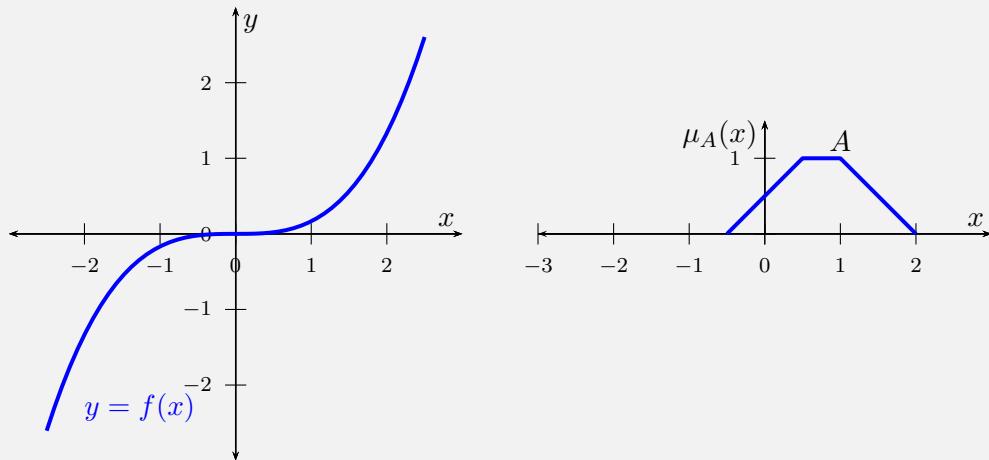


Figura 9.2: Problema del ejemplo 9.2

Sabemos que el resultado de $f(A)$ será un conjunto difuso. Llamemos a ese conjunto B y a su función de pertenencia $\mu_B(y)$. Para ilustrar la aplicación del principio de extensión vamos a calcular el grado de pertenencia en un solo valor de y . Hemos seleccionado $y = 0.75$ y el proceso se muestra en la Figura 9.3:

Debemos determinar todos los valores de x tales que $f(x) = 0.75$. Al despejar en 9.11 se obtiene un único valor \bar{x} (ver la línea azul a trozos en la Figura 9.3):

$$\bar{x} = \sqrt[3]{6 * 0.75} = 1.651 \quad (9.12)$$

Para calcular $\mu_A(1.651)$ aplicamos 7.4 al trapecio $T(-0.5, 0.5, 1.0, 2.0)$, resultando en

$$\mu_A(1.651) = \frac{2.0 - 1.651}{2.0 - 1 - 0} = 0.349 \quad (9.13)$$

Ese valor corresponde a $\mu_B(0.75)$ (ver la línea roja a trozos en la Figura 9.3)

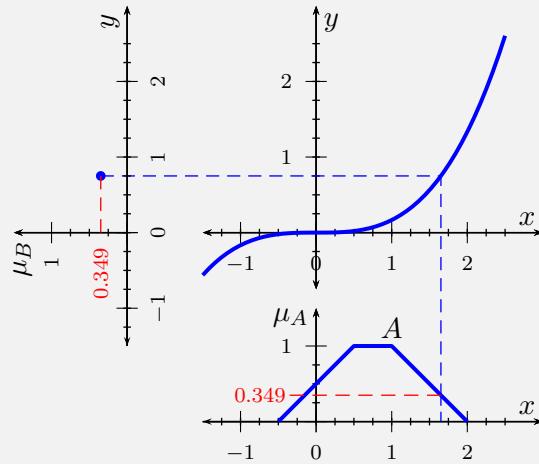


Figura 9.3: Grado de pertenencia para $y = 0.75$ en el ejemplo 9.2

Ahora bien, el procedimiento anterior hay que repetirlo para todos los valores de y . En la Figura 9.4 se muestra el proceso para unos cuantos valores de y y el resultado final que es el conjunto $B = f(A)$

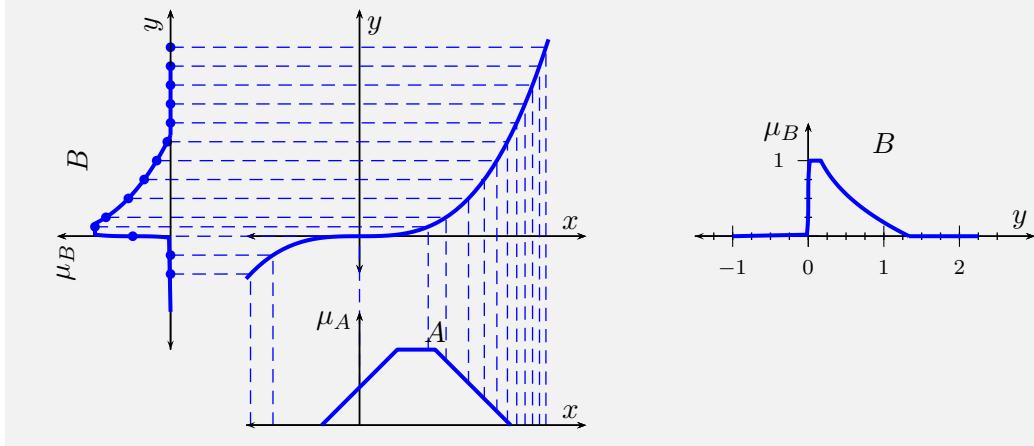


Figura 9.4: Respuesta del ejemplo 9.2

Ejemplo 9.3 (Principio de extensión de una función sobre \mathbb{R} no uno a uno)

Consideremos la función $f(x)$ y el conjunto difuso A que se muestran en la Figura 9.5, que corresponden a:

$$f(x) = \begin{cases} x + 0.5 & \text{si } x < 0.5 \\ 1 & \text{si } 0.5 \leq x < 2 \\ 3x - 5 & \text{si } x \geq 2 \end{cases} \quad (9.14)$$

$$A : T(-0.5, 0.5, 1.0, 2.0)$$

¿Cuál es el resultado de aplicar $f(x)$ al conjunto A ? En otras palabras ¿Cuál es el resultado de $f(A)$?

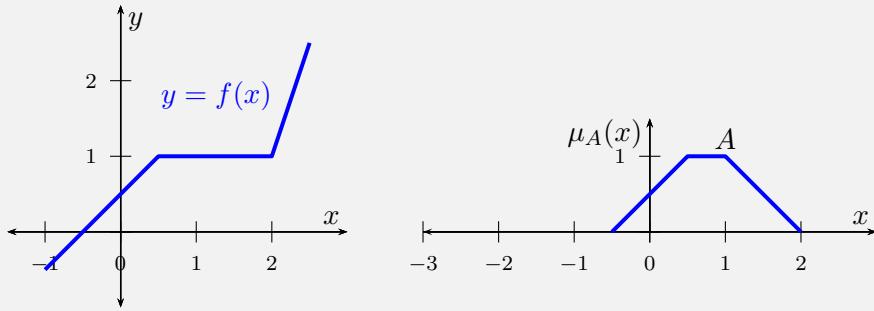
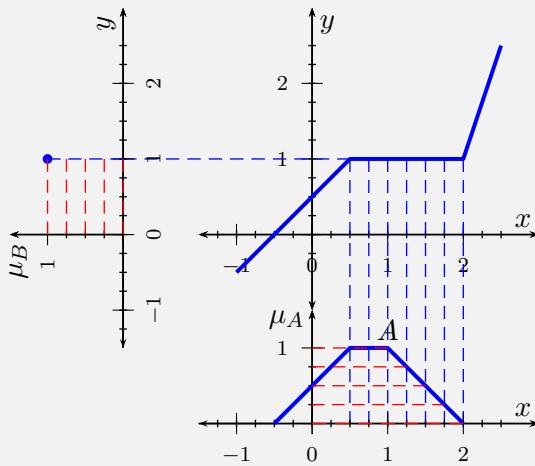


Figura 9.5: Problema del ejemplo 9.3

Existe una diferencia importante entre $f(x)$ en este ejemplo y en el ejemplo 9.2: el trozo de línea recta horizontal de la función hace que exista más de un valor de x para el cual $f(x) = 1$; de hecho para todo x en $[0.5, 2.0]$ la función vale 1. En otras palabras, la función no es uno a uno.

Para calcular $\mu_B(1)$ debemos determinar todos los valores de x tales que $f(x) = 1$. Como ya hemos dicho, estos valores forman el intervalo $[0.5, 2.0]$. Debemos calcular $\mu_A(x)$ para todos los valores y tomar el sup de esos valores. En la Figura 9.6 se muestra el proceso para algunos valores, suficientes para mostrar que $\mu_B(1.0) = 1.0$ (ver la línea roja a trozos).

Figura 9.6: Grado de pertenencia para $y = 1$ en el ejemplo 9.3

El procedimiento anterior hay que repetirlo para todos los valores de y . En la Figura 9.7 se muestra el proceso para unos cuantos valores de y y el resultado final que es el conjunto $B = f(A)$

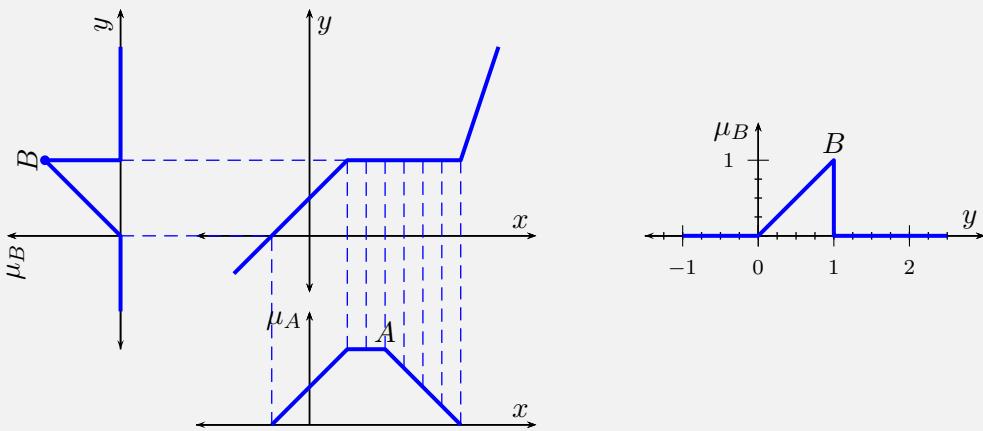


Figura 9.7: Respuesta del ejemplo 9.3

9.2. Extensión de funciones sobre \mathbb{R}^n

Supongamos que ahora una función de múltiples argumentos reales $f(x_1, x_2, \dots, x_n)$. Para simplificar la notación construimos el vector $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$ de tal manera que

$$\begin{aligned} y = f(\mathbf{x}) & : \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R} \rightarrow \mathbb{R} \\ y = f(\mathbf{x}) & : \mathbb{R}^n \rightarrow \mathbb{R} \end{aligned} \quad (9.15)$$

También suponemos que existe un conjunto difuso A_k en cada una de las n variables y queremos determinar

$$B = f(A_1, A_2, \dots, A_n) \quad (9.16)$$

Para aplicar el principio de extensión en estas condiciones (ecuación 9.9) es necesario identificar cuál es el grado de pertenencia del argumento de la función. Utilizamos para ello la \mathcal{T} -norma estándar, es decir el operador \min y obtenemos

$$\mu_B(y) = \sup_{f(\mathbf{x})=y} \{\min\{\mu_{A_1}(x_1), \mu_{A_2}(x_2), \dots, \mu_{A_n}(x_n)\}\} \quad (9.17)$$

Ejemplo 9.4 (Principio de extensión de una función sobre \mathbb{R}^2) Para ilustrar el significado de la ecuación 9.17 utilizaremos una función simple definida en \mathbb{R}^2 así:

$$\begin{aligned} f(x_1, x_2) & : \mathbb{R}^2 \rightarrow \mathbb{R} \\ f(x_1, x_2) & = x_1 + x_2 \end{aligned} \quad (9.18)$$

Definimos los conjuntos difusos triangulares $A_1 = Tr(1, 2, 3)$ y $A_2 = Tr(2, 3, 4)$ con el propósito de calcular $B = f(A_1, A_2)$, es decir, con el propósito de calcular $B = A_1 + A_2$.

La Figura 9.8 ilustra el problema: la función se visualiza en \mathbb{R}^3 como un plano cuya ecuación es $y = x_1 + x_2$; en la parte inferior se muestran las funciones de pertenencia de A_1 y de A_2 , mientras que el problema consiste en determinar $\mu_B(y)$.

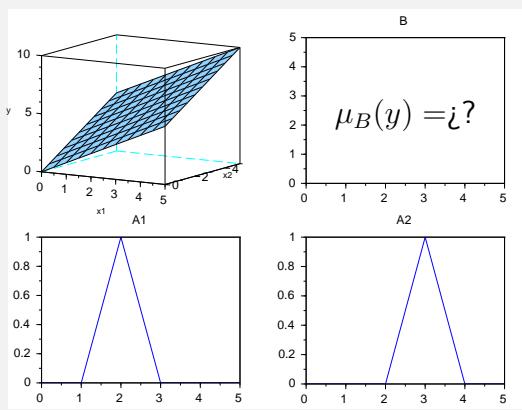
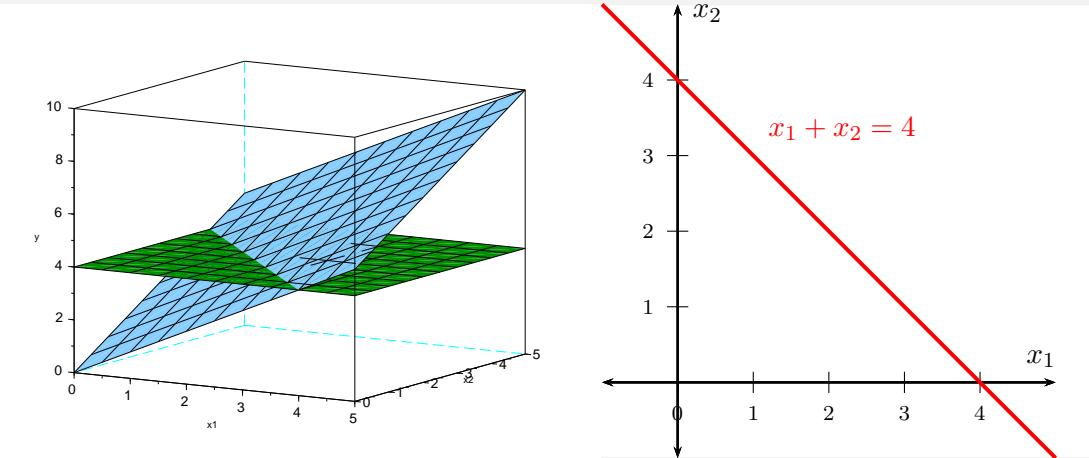


Figura 9.8: Problema del ejemplo 9.4

Para ilustrar el proceso, vamos a calcular $\mu_B(4)$. Lo primero que se necesita es determinar cuáles son las parejas (x_1, x_2) tales que $f(x_1, x_2) = 4$; en otras palabras, cuales son las parejas de puntos tales que $x_1 + x_2 = 4$. La Figura 9.9 muestra tales parejas: en \mathbb{R}^3 se visualizan como la intersección de los planos definidos por las ecuaciones

$$\begin{aligned} x_1 + x_2 &= 4 \\ y &= 2 \end{aligned} \tag{9.19}$$

La proyección en \mathbb{R}^2 de la intersección es la recta de color rojo en esa misma figura.

Figura 9.9: Parejas de puntos que cumplen $x_1 + x_2 = 4$

La recta está formada por infinitos parejas de puntos (x_1, x_2) . Debemos encontrar en cada caso $\mu_{A_1}(x_1)$, $\mu_{A_2}(x_2)$ y comparar sus valores. Para hacerlo, vamos a recorrer la recta utilizando un parámetro t , o lo que es igual, vamos a escribir primero la ecuación de la recta en forma paramétrica.

Seleccionamos de forma arbitraria dos puntos de la recta $p_0 = (0, 4)$, $p_1 = (4, 0)$ y decidimos que p_0 corresponda a $t = 0$, mientras que p_1 corresponda a $t = 1$. La ecuación de cada una de las coordenadas será:

$$\begin{aligned} x_1 &= 4t \\ x_2 &= 4 - 4t \end{aligned} \quad (9.20)$$

Con lo que la ecuación de la recta resulta ser:

$$R : (4t, 4 - 4t) \quad t \in (-\infty, \infty) \quad (9.21)$$

Ahora podemos reescibir la ecuación 7.5 para $A_1 = Tr(1, 2, 3)$ y $A_2 = Tr(2, 3, 4)$ en función del parámetro t . El resultado se ha graficado en la Figura 9.10 y corresponde a las siguientes expresiones:

$$\mu_{A_1}(t) = \begin{cases} 0 & \text{si } t < 1/4 \\ 4t - 1 & \text{si } 1/4 \leq t < 1/2 \\ 3 - 4t & \text{si } 1/2 \leq t < 3/4 \\ 0 & \text{si } 3/4 \leq t \end{cases} \quad \mu_{A_2}(t) = \begin{cases} 0 & \text{si } t \geq 1/2 \\ 2 - 4t & \text{si } 1/2 > t \geq 1/4 \\ 4t & \text{si } 1/4 > t \geq 0 \\ 0 & \text{si } 0 > t \end{cases} \quad (9.22)$$

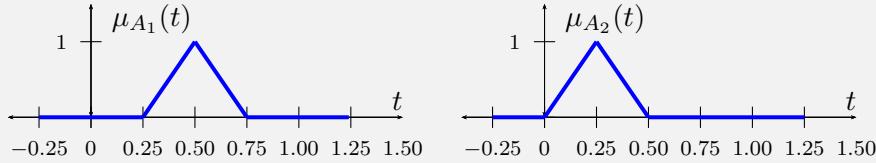


Figura 9.10: Funciones de pertenencia parametrizadas en el ejemplo 9.4

Ahora podemos encontrar $\min \{\mu_{A_1}(t), \mu_{A_2}(t)\}$, tal como se muestra en la Figura 9.11. Allí también se observa que:

$$\sup \{\min \{\mu_{A_1}(t), \mu_{A_2}(t)\}\} = 0.5 \quad (9.23)$$

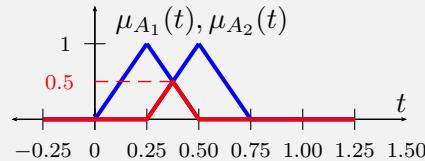


Figura 9.11: Mínimo de funciones de pertenencia parametrizadas en el ejemplo 9.4

La ecuación 9.23 nos permite entonces afirmar que $\mu_B(4) = 0.5$. Para obtener $\mu_B(y)$ debemos repetir el procedimiento para todos los valores de y . El resultado se muestra en la Figura 9.12, en la que se destaca en rojo el caso estudiado en detalle ($y = 4$).

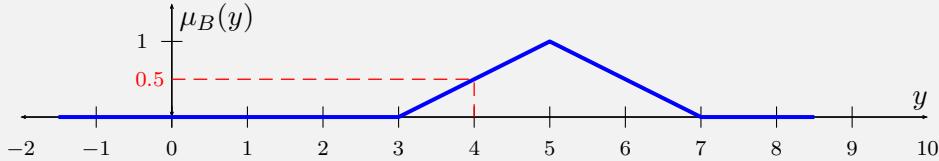


Figura 9.12: Resultado del ejemplo 9.4

9.3. Extensión mediante α -cortes

Los ejemplos 9.2, 9.2 y 9.20 ilustran que la aplicación del principio de extensión puede ser una tarea engorrosa, sobre todo si se quiere extender una operación definida sobre \mathbb{R}^n . En esta sección se demuestra que existe un procedimiento equivalente mucho más simple.

El procedimiento extiende las operaciones trabajando con los α -cortes de los conjuntos (sección 1.3.3). Específicamente, se apoya en el siguiente teorema:

Teorema 9.1 *Sea función $f : U \rightarrow V$ y A un conjunto difuso $A \in U$ y $B = f(A)$, entonces cada α -corte de B se puede obtener así:*

$$B_\alpha = [f(A)]_\alpha = f(A_\alpha) \quad (9.24)$$

La interpretación de 9.24 es la siguiente: B_α es un conjunto concreto definido sobre V , mientras que A_α es un conjunto concreto definido sobre U . B_α puede obtenerse calculando f en todos los elementos de A_α .

Para demostrar 9.24 consideraremos primero un valor de $\alpha > 0$ y un elemento \bar{y} que forma parte de B_α . Por definición se tiene que

$$\bar{y} \in B_\alpha \Leftrightarrow \mu_B(\bar{y}) \geq \alpha \quad (9.25)$$

Al aplicar la ecuación 9.9 se tiene que

$$\bar{y} \in B_\alpha \Leftrightarrow \sup_{f(x)=\bar{y}} \{\mu_A(x)\} \geq \alpha \quad (9.26)$$

Para que se cumpla la desigualdad que está a la derecha de la doble implicación se necesita que exista al menos un valor \bar{x} que cumpla dos condiciones:

- \bar{x} debe cumplir la restricción, es decir $f(\bar{x}) = \bar{y}$
- El grado de pertenencia a A debe ser al menos α , es decir $\mu_A(\bar{x}) \geq \alpha$

Por tanto, podemos escribir

$$\bar{y} \in B_\alpha \Leftrightarrow \exists \bar{x} \mid \bar{y} = f(\bar{x}) \quad \mathbf{Y} \quad \mu_A(\bar{x}) \geq \alpha \quad (9.27)$$

ó, lo que es igual

$$\bar{y} \in B_\alpha \Leftrightarrow \exists \bar{x} \mid \bar{y} = f(\bar{x}) \quad \mathbf{Y} \quad \bar{x} \in A_\alpha \quad (9.28)$$

El término de la derecha de la doble implicación significa que \bar{y} forma parte de la imagen de f cuando se calcula sobre A_α ,

$$\bar{y} \in B_\alpha \Leftrightarrow \bar{y} \in f(A_\alpha) \quad (9.29)$$

La doble implicación en la ecuación 9.29 significa que hemos concluido la demostración de 9.24 para $\alpha > 0$. ¿Y para $\alpha = 0$? La demostración sigue los mismos pasos, remplazando el símbolo \geq por $>$ en las ecuaciones 9.25 a 9.29.

Ejemplo 9.5 (Extensión mediante α -cortes de una función uno a uno)

Retomemos ahora el ejemplo 9.2, para solucionarlo mediante la aplicación del principio de extensión sobre los α -cortes. Lo que se desea es calcular $f(A)$ sabiendo que $f(x) = \frac{x^3}{6}$ y A es el número trapezoidal $A : T(-0.5, 0.5, 1.0, 2.0)$ (Figura 9.2).

Usando la ecuación 7.12 podemos calcular los α -cortes de A

$$A_\alpha = [(-0.5 + \alpha), (2.0 - \alpha)] \quad (9.30)$$

Para aplicar 9.24 deberíamos calcular $f(x)$ en todos los puntos del intervalo A_α . Sin embargo esto no es necesario, porque $f(x)$ es continua y monótonamente creciente, por lo que podemos aplicar el Teorema 8.1 y obtener :

$$\begin{aligned} B_\alpha &= [L_B(\alpha), R_B(\alpha)] \\ B_\alpha &= [f(L_A(\alpha)), f(R_A(\alpha))] \\ L_B(\alpha) &= f(L_A(\alpha)) &= \frac{(-0.5+\alpha)^3}{6} \\ R_B(\alpha) &= f(R_A(\alpha)) &= \frac{(2.0-\alpha)^3}{6} \end{aligned} \quad (9.31)$$

En la Figura 9.13 se han graficado las funciones $L_B(\alpha)$ y $R_B(\alpha)$. La curva en azul corresponde a $L_B(\alpha)$ y la que está en rojo a $R_B(\alpha)$. Para facilitar la visualización de las curvas

estas se han trazado en punteado, y sobre ellas se ha dibujado en una línea sólida más gruesa el trozo que corresponde al intervalo $\alpha \in [0, 1]$.

También se ha graficado $\mu_B(y)$, que se obtiene fácilmente al intercambiar los ejes. Este resultado se puede comparar con el que se muestra en la Figura 9.4 para verificar que los dos métodos llevan a la misma solución. Recuérdese que por fuera del soporte el grado de pertenencia es cualquier conjunto difuso es 0.

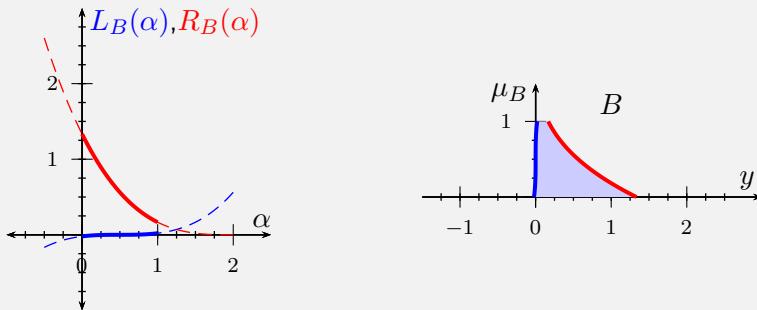


Figura 9.13: Solución del ejemplo 9.5

Ejemplo 9.6 (Extensión mediante α -cortes de una función sobre \mathbb{R})

Retomemos ahora el ejemplo 9.3, para solucionarlo mediante la aplicación del principio de extensión sobre los α -cortes. Lo que se desea es calcular $f(A)$ sabiendo que A es el número trapezoidal $A : T(-0.5, 0.5, 1.0, 2.0)$ y que $f(x)$ es la función que se muestra en la Figura 9.5, es decir:

$$f(x) = \begin{cases} x + 0.5 & \text{si } x < 0.5 \\ 1 & \text{si } 0.5 \leq x < 2 \\ 3x - 5 & \text{si } x \geq 2 \end{cases} \quad (9.32)$$

Debido a que esta función también es continua y monótonamente no decreciente también podemos hacer uso del Teorema 8.1 y calcular los α -cortes de $B = f(A)$ como:

$$\begin{aligned} B_\alpha &= [f(A_\alpha), f(A_\alpha)] \\ B_\alpha &= [f(-0.5 + \alpha), f(2.0 - \alpha)] \\ B_\alpha &= [\alpha, 1.0] \end{aligned} \quad (9.33)$$

El resultado se muestra en la Figura 9.14 y puede compararse con el de conjunto obtenido en la Figura 9.7

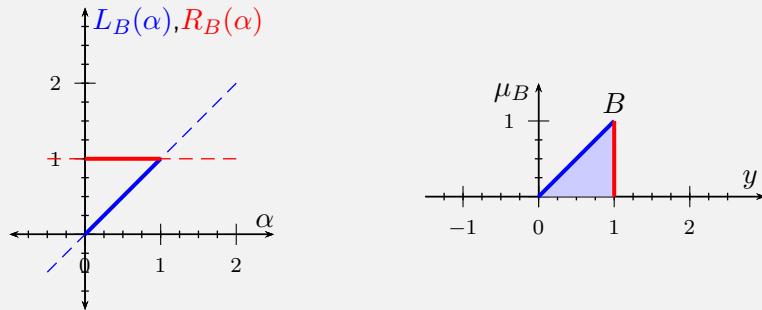


Figura 9.14: Solución del ejemplo 9.5

Ejemplo 9.7 (Extensión mediante α -cortes de una función sobre \mathbb{R}^2)

Retomemos ahora el ejemplo 9.4, para solucionarlo mediante la aplicación del principio de extensión sobre los α -cortes. El propósito es calcular $f(x_1, x_2) = x_1 + x_2$ cuando $A_1 = Tr(1, 2, 3)$ y $A_2 = Tr(2, 3, 4)$.

Podemos obtener los α -cortes de A_1, A_2 usando la tabla 7.1:

$$\begin{aligned} A_{1_\alpha} &= [(1 + \alpha), (3 - \alpha)] \\ A_{2_\alpha} &= [(2 + \alpha), (4 - \alpha)] \end{aligned} \quad (9.34)$$

Para obtener $B = f(A_1, A_2)$ podemos aplicar la ecuación 9.24. Para ello, es necesario tener en cuenta que el A_α está ahora definido sobre \mathbb{R}^2 y por tanto será el producto cartesiano de los α -cortes de A_1 y A_2 :

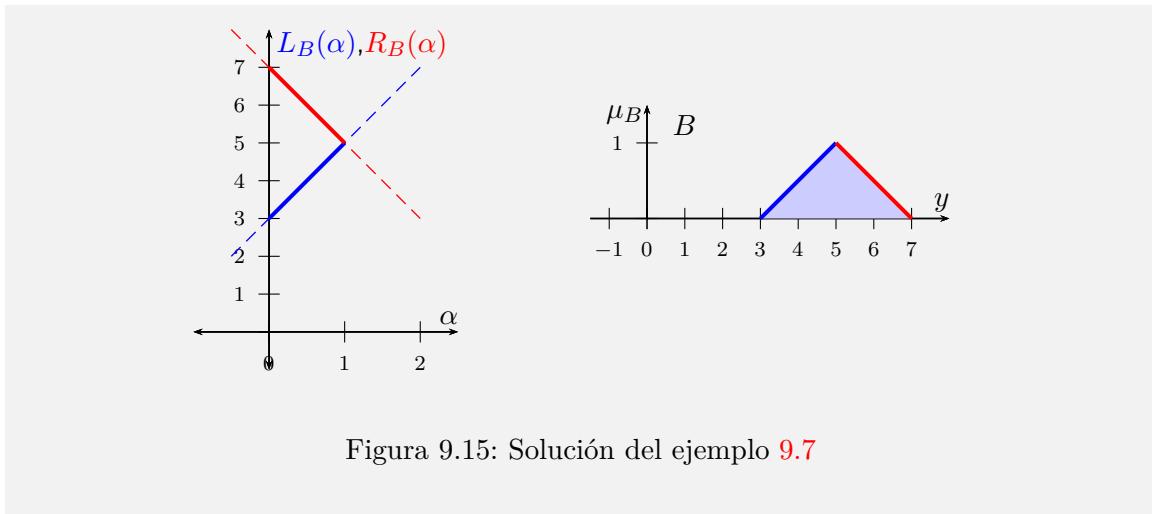
$$A_\alpha = A_{1_\alpha} \times A_{2_\alpha} \quad (9.35)$$

Tanto A_{1_α} como A_{2_α} son intervalos, por tanto A_α es un rectángulo. Deberíamos calcular $f(x_1, x_2)$ en los infinitos puntos de ese rectángulo para obtener B_α . Sin embargo, esto no es necesario por las razones que se explican a continuación.

$f(x_1 + x_2)$ es una suma, lo implica que es una función monótona creciente en sus dos variables. Podemos entonces aplicar el Teorema 8.2 directamente:

$$\begin{aligned} \mathcal{L}(\alpha) &= \{L_{A_1}(\alpha), L_{A_2}(\alpha)\} \\ \mathcal{R}(\alpha) &= \{R_{A_1}(\alpha), R_{A_2}(\alpha)\} \\ L_B(\alpha) &= f(L_{A_1}(\alpha), L_{A_2}(\alpha)) = (1 + \alpha) + (2 + \alpha) = 3 + 2\alpha \\ R_B(\alpha) &= f(R_{A_1}(\alpha), R_{A_2}(\alpha)) = (3 + \alpha) + (4 + \alpha) = 7 - 2\alpha \end{aligned} \quad (9.36)$$

El resultado se muestra en la Figura 9.15 y puede compararse con el de conjunto obtenido en la Figura 9.12



Capítulo 10

Aritmética difusa

En capítulos previos hemos presentado y explorado tres hechos que ahora conectaremos:

- Los α -cortes de los números difusos son intervalos cerrados (Sección 7.1).
- Existen formas muy concisas de efectuar las operaciones aritméticas básicas sobre intervalos (Sección 8.7).
- Es posible extender una función a números difusos utilizando los α -cortes (Sección 9.3).

El propósito de este capítulo consiste en ilustrar cómo se pueden efectuar operaciones aritméticas sobre números difusos¹. Lo haremos empleando los tres hechos arriba mencionados, conectados de la siguiente forma:

Las operaciones aritméticas entre números difusos se realizan mediante los α -cortes de los números difusos, que son intervalos cerrados, cuyas operaciones aritméticas se pueden realizar de forma muy concisa.

En este capítulo centraremos nuestra atención en las cuatro operaciones aritméticas básicas, mientras que en el capítulo 11 se estudiará la implementación en software de otras funciones más generales. Si denotamos por \odot a un operador genérico para las cuatro operaciones, es decir $\odot \in \{+, -, \cdot, /\}$ el párrafo anterior se expresa como sigue:

$$(A \odot B)_{\alpha} = [A_{\alpha} \odot B_{\alpha}] \quad (10.1)$$

¹En realidad, en el ejemplo 9.7 ya se ilustró cómo se conectan estos tres hechos con la operación de suma.

10.1. Suma de números difusos

Para aplicar 10.1 al operador de suma, basta con recordar que la suma de intervalos se rige por la ecuación 8.58, lo que resulta en:

$$\begin{aligned}(A + B)_\alpha &= [A_\alpha + B_\alpha] \\ (A + B)_\alpha &= [L_A(\alpha) + L_B(\alpha), R_A(\alpha) + R_B(\alpha)]\end{aligned}\quad (10.2)$$

Como caso especial, consideremos la suma de dos números trapezoidales:

$$\begin{aligned}A &= T(a, b, c, d) = [a + \alpha(b - a), d - \alpha(d - c)] \\ B &= T(e, f, g, h) = [e + \alpha(f - e), h - \alpha(h - g)]\end{aligned}\quad (10.3)$$

Cuya suma resulta ser

$$(A + B)_\alpha = [(a + \alpha(b - a)) + (e + \alpha(f - e)), (d - \alpha(d - c)) + (h - \alpha(h - g))] \quad (10.4)$$

Es decir,

$$\begin{aligned}L_{A+B}(\alpha) &= (a + e) + \alpha((b + f) - (a + e)) \\ R_{A+B}(\alpha) &= (d + h) - \alpha((d + h) - (c + g))\end{aligned}\quad (10.5)$$

Que corresponden a los límites del número trapezoidal $T(a + e, b + f, c + g, d + h)$. En otras palabras, la suma de dos números trapezoidales resulta en otro número trapezoidal, cuyos vértices se obtienen sumando los respectivos vértices:

$$T(a, b, c, d) + T(e, f, g, h) = T(a + e, b + f, c + g, d + h) \quad (10.6)$$

Ejemplo 10.1 (Suma de números difusos) Consideremos los números trapezoidales

$$\begin{aligned}A &= T(1, 2, 3, 4) \\ B &= T(2, 3, 4, 5)\end{aligned}\quad (10.7)$$

La Figura 10.1 muestra los números A y B . De acuerdo con la ecuación 7.12 sus respectivos α -cortes están definidos por

$$\begin{aligned}A_\alpha &= [1 + \alpha(2 - 1), 4 - \alpha(4 - 3)] = [1 + \alpha, 4 - \alpha] \\ B_\alpha &= [2 + \alpha(3 - 2), 5 - \alpha(5 - 4)] = [2 + \alpha, 5 - \alpha]\end{aligned}\quad (10.8)$$

Aplicando la ecuación 10.2 se obtienen los α -cortes de $C = A + B$

$$\begin{aligned}
 C_\alpha &= [L_A(\alpha) + L_B(\alpha), R_A(\alpha) + R_B(\alpha)] \\
 C_\alpha &= [(1+\alpha) + (2+\alpha), (4-\alpha) + (5-\alpha)] \\
 C_\alpha &= [3+2\alpha, 9-2\alpha]
 \end{aligned} \tag{10.9}$$

Que corresponde al número trapezoidal $T(3, 5, 7, 9)$ de la Figura 10.1. Este resultado se podría haber obtenido aplicar directamente la ecuación 10.6:

$$\begin{aligned}
 A + B &= T(1+2, 2+3, 3+4, 4+5) \\
 A + B &= T(3, 5, 7, 9)
 \end{aligned} \tag{10.10}$$

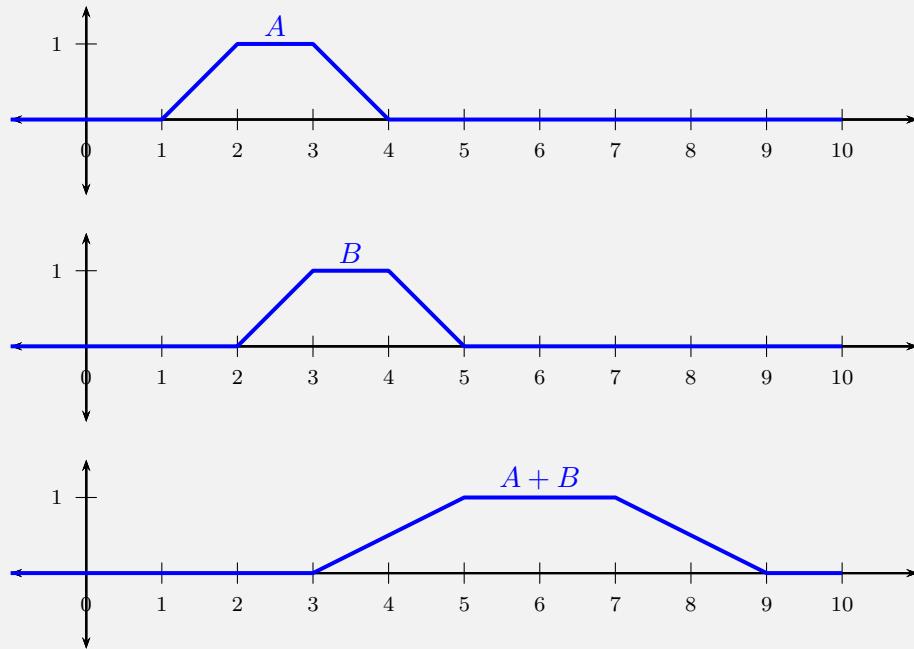


Figura 10.1: Suma de números difusos

10.2. Resta de números difusos

Para aplicar 10.1 al operador de resta, basta con recordar que la resta de intervalos se rige por la ecuación 8.58, lo que resulta en:

$$\begin{aligned}(A - B)_\alpha &= [A_\alpha - B_\alpha] \\ (A - B)_\alpha &= [L_A(\alpha) - R_B(\alpha), R_A(\alpha) - L_B(\alpha)]\end{aligned}\quad (10.11)$$

Como caso especial, consideremos la resta de dos números trapezoidales:

$$\begin{aligned}A &= T(a, b, c, d) = [a + \alpha(b - a), d - \alpha(d - c)] \\ B &= T(e, f, g, h) = [e + \alpha(f - e), h - \alpha(h - g)]\end{aligned}\quad (10.12)$$

Cuya resta resulta ser

$$(A - B)_\alpha = [(a + \alpha(b - a)) - (h - \alpha(h - g)), (d - \alpha(d - c)) - (e + \alpha(f - e))] \quad (10.13)$$

Es decir,

$$\begin{aligned}L_{A-B}(\alpha) &= (a - h) + \alpha((b - g) - (a - h)) \\ R_{A-B}(\alpha) &= (d - e) - \alpha((d - e) - (c - f))\end{aligned}\quad (10.14)$$

Que corresponden a los límites del número trapezoidal $T(a - h, b - g, c - f, d - e)$. En otras palabras, la suma de dos números trapezoidales resulta en otro número trapezoidal, cuyos vértices se obtienen restando algunos de sus vértices:

$$T(a, b, c, d) - T(e, f, g, h) = T(a - h, b - g, c - f, d - e) \quad (10.15)$$

Ejemplo 10.2 (Resta de números difusos) Consideremos los números trapezoidales

$$\begin{aligned}A &= T(5, 6, 7, 8) \\ B &= T(1, 2, 3, 4)\end{aligned}\quad (10.16)$$

La Figura 10.2 muestra los números A y B . De acuerdo con la ecuación 7.12 sus respectivos α -cortes están definidos por

$$\begin{aligned}A_\alpha &= [5 + \alpha(6 - 5), 8 - \alpha(8 - 7)] = [5 + \alpha, 8 - \alpha] \\ B_\alpha &= [1 + \alpha(2 - 1), 4 - \alpha(4 - 3)] = [1 + \alpha, 4 - \alpha]\end{aligned}\quad (10.17)$$

Aplicando la ecuación 10.11 se obtienen los α -cortes de $C = A - B$

$$\begin{aligned}C_\alpha &= [L_A(\alpha) - R_B(\alpha), R_A(\alpha) - L_B(\alpha)] \\ C_\alpha &= [(5 + \alpha) - (4 - \alpha), (8 - \alpha) - (1 + \alpha)] \\ C_\alpha &= [1 + 2\alpha, 7 - 2\alpha]\end{aligned}\quad (10.18)$$

Que corresponde al número trapezoidal $T(1, 3, 5, 7)$ de la Figura 10.2. Este resultado se podría haber obtenido aplicar directamente la ecuación 10.15:

$$\begin{aligned} A - B &= T(5 - 4, 6 - 3, 7 - 2, 8 - 1) \\ A - B &= T(1, 3, 5, 7) \end{aligned} \quad (10.19)$$

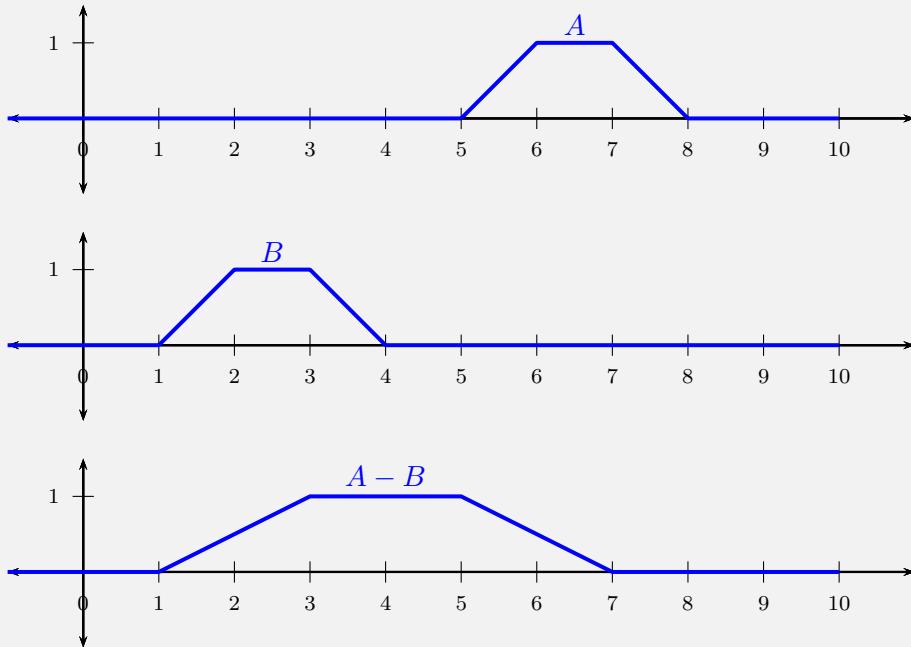


Figura 10.2: Resta de números difusos

10.3. Producto de números difusos

La aplicación de la ecuación 10.1 al operador de producto resulta en:

$$(A \cdot B)_\alpha = [A_\alpha \cdot B_\alpha] \quad (10.20)$$

A diferencia de los casos de suma y resta, para el producto no hay una expresión tan compacta como la que desemboca en las ecuaciones 10.2 y 10.11. De hecho, en la sección 8.7.2 se demuestra que el producto de dos intervalos se rige por la ecuación 8.59 o, lo que es igual, por las expresiones de la Tabla 8.1.

Como caso especial, consideremos el producto de dos números trapezoidales positivos:

$$\begin{aligned} A &= T(a, b, c, d) = [a + \alpha(b - a), d - \alpha(d - c)] & 0 < a \\ B &= T(e, f, g, h) = [e + \alpha(f - e), h - \alpha(h - g)] & 0 < e \end{aligned} \quad (10.21)$$

Cuya producto está dado por el caso 1 de la Tabla 8.1 y resulta ser

$$(A \cdot B)_\alpha = [(a + \alpha(b - a))(e + \alpha(f - e)), (d - \alpha(d - c))(h - \alpha(h - g))] \quad (10.22)$$

Es decir,

$$\begin{aligned} L_{A \cdot B}(\alpha) &= ae + \alpha(be + af - 2ae) + \alpha^2(bf + ae - be - af) \\ R_{A \cdot B}(\alpha) &= dh - \alpha(2dh - ch - dg) + \alpha^2(dh + cg - dg - ch) \end{aligned} \quad (10.23)$$

Es decir, el resultado no corresponde a un número trapezoidal. En efecto, la presencia de los términos que dependen de α^2 significa que las curvas $L(\alpha)$ y $R(\alpha)$ no son líneas rectas sino expresiones cuadráticas.

Ejemplo 10.3 (Producto de números difusos) Consideremos los números trapezoidales

$$\begin{aligned} A &= T(-0.5, 0.5, 1.5, 2.5) \\ B &= T(0, 1, 1, 2) \end{aligned} \quad (10.24)$$

La Figura 10.3 muestra los números A y B , así como el resultado del producto $A \cdot B$. De acuerdo con la ecuación 7.12 sus respectivos α -cortes están definidos por

$$\begin{aligned} A_\alpha &= [-0.5 + \alpha(1.5 - 0.5), 2.5 - \alpha(2.5 - 1.5)] = [0.5 + \alpha, 2.5 - \alpha] \\ B_\alpha &= [0.0 + \alpha(1.0 - 0.0), 2.0 - \alpha(2.0 - 1.0)] = [0.0 + \alpha, 2.0 - \alpha] \end{aligned} \quad (10.25)$$

Nótese que el número A no es positivo ni negativo. A continuación mostramos algunos de los α -cortes de los dos números:

$$\begin{array}{lll} A_{0.00} &= [-0.50, 2.50] & B_{0.00} = [0.00, 2.00] \\ A_{0.25} &= [-0.25, 2.00] & B_{0.25} = [0.25, 1.75] \\ A_{0.50} &= [0.00, 2.00] & B_{0.50} = [0.50, 1.50] \\ A_{0.75} &= [0.75, 2.00] & B_{0.75} = [0.75, 1.25] \\ A_{1.00} &= [0.50, 1.50] & B_{1.00} = [1.00, 1.00] \end{array} \quad (10.26)$$

Es necesario destacar que los α -cortes correspondientes a $0 \leq \alpha < 0.5$ caen en el caso 5 de la Tabla 8.1, mientras que los casos correspondientes a $0.5 \leq \alpha \leq 1.0$ caen en el caso 1.

En esa Tabla se verifica que los valores de $L_{A \cdot B}$ son diferentes en los dos casos, mientras que son iguales para $R_{A \cdot B}$:

$$L_{A \cdot B}(\alpha) = \begin{cases} (0.5 + \alpha)(2 - \alpha) & \text{si } 0.0 \leq \alpha < 0.5 \\ (0.5 + \alpha)\alpha & \text{si } 0.5 \leq \alpha < 1.0 \end{cases} \quad (10.27)$$

$$R_{A \cdot B}(\alpha) = (2.5 - \alpha)(2 - \alpha)$$

Este resultado puede visualizarse en la Figura 10.3. En efecto, allí se puede observar que la curva ascendente de $A \cdot B$ tiene un quiebre justo en $x = 0$, que corresponde a $\alpha = 0.5$, mientras que la curva descendente no presenta ese quiebre.

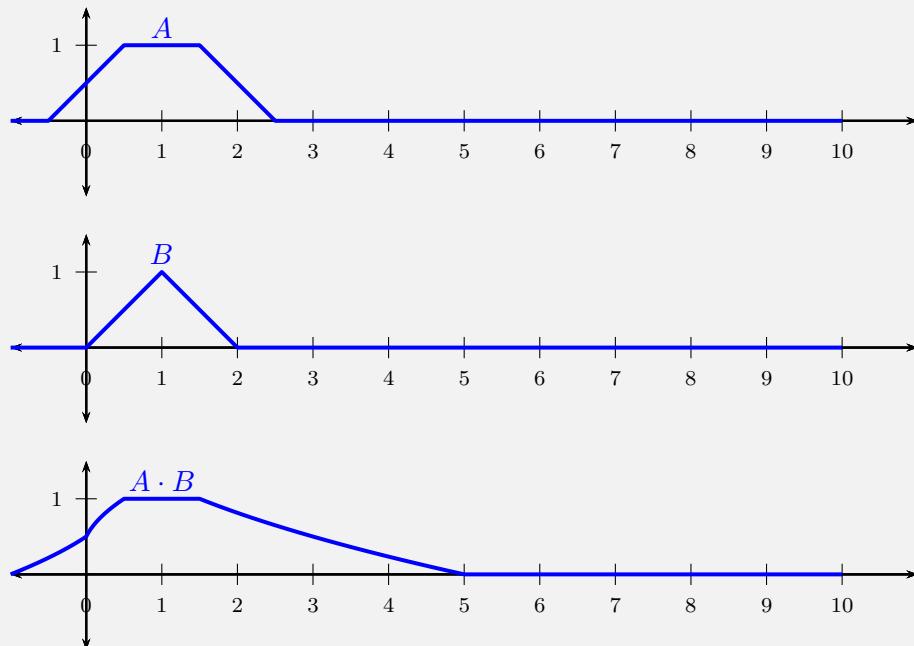


Figura 10.3: Producto de números difusos

Ejemplo 10.4 (Producto de números difusos positivos) Consideremos los números trapezoidales positivos

$$\begin{aligned} A &= T(0, 1, 1, 2) \\ B &= T(1, 2, 3, 4) \end{aligned} \quad (10.28)$$

La Figura 10.3 muestra los números A y B , así como el resultado del producto $A \cdot B$. Al aplicar la ecuación 10.23 se obtiene

$$\begin{aligned} L_{A \cdot B}(\alpha) &= 0 \cdot 1 + \alpha(1 \cdot 1 + 0 \cdot 2 - 2 \cdot 0 \cdot 1) + \alpha^2(1 \cdot 2 + 0 \cdot 1 - 1 \cdot 1 - 0 \cdot 2) \\ R_{A \cdot B}(\alpha) &= 2 \cdot 4 - \alpha(2 \cdot 2 \cdot 4 - 1 \cdot 4 - 2 \cdot 3) + \alpha^2(2 \cdot 4 + 1 \cdot 3 - 2 \cdot 3 - 1 \cdot 4) \end{aligned} \quad (10.29)$$

Lo que resulta en

$$\begin{aligned} L_{A \cdot B}(\alpha) &= \alpha + \alpha^2 \\ R_{A \cdot B}(\alpha) &= 8 - 6\alpha + \alpha^2 \end{aligned} \quad (10.30)$$

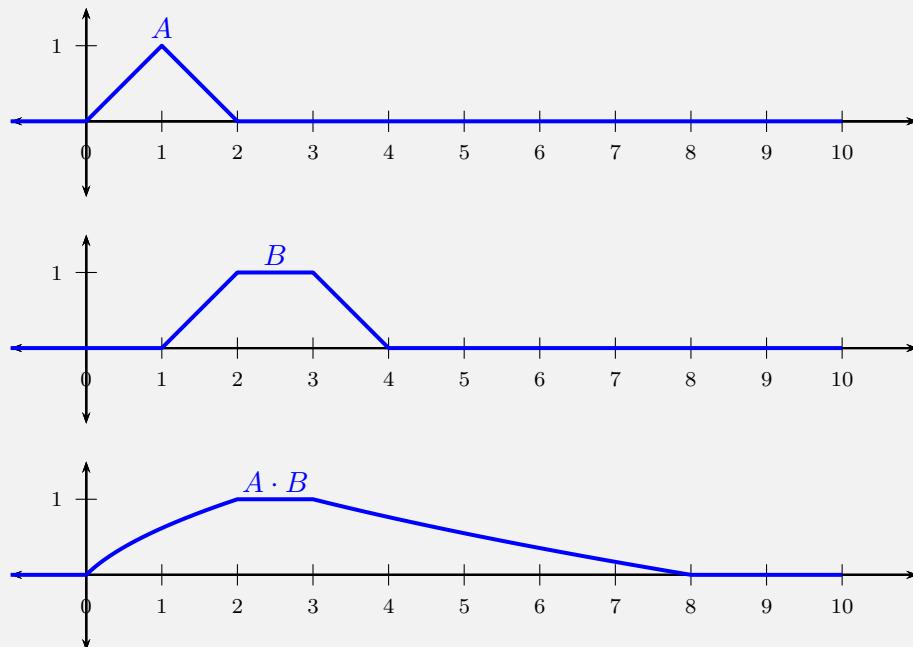


Figura 10.4: Producto de números difusos positivos

10.4. División de números difusos

La aplicación de la ecuación 10.1 al operador de producto resulta en:

$$(A/B)_\alpha = [A_\alpha/B_\alpha] \quad (10.31)$$

Al igual que en producto de números difusos, para la división no hay una expresión tan compacta como la que desemboca en las ecuaciones de suma y resta, 10.2 y 10.11. De hecho, en la sección 8.7.2.1 se demuestra que la división de dos intervalos se rige por la ecuación 8.66 o, lo que es igual, por las expresiones de la Tabla 8.2.

Como caso especial, consideremos el producto de dos números trapezoidales positivos:

$$\begin{aligned} A &= T(a, b, c, d) = [a + \alpha(b - a), d - \alpha(d - c)] & 0 < a \\ B &= T(e, f, g, h) = [e + \alpha(f - e), h - \alpha(h - g)] & 0 < e \end{aligned} \quad (10.32)$$

Cuya división está dada por el caso 1 de la Tabla 8.2 y resulta ser

$$(A/B)_\alpha = \left[\frac{a + \alpha(b - a)}{h - \alpha(h - g)}, \frac{d - \alpha(d - c)}{e + \alpha(f - e)} \right] \quad (10.33)$$

Es decir,

$$\begin{aligned} L_{A/B}(\alpha) &= \frac{a + \alpha(b - a)}{h - \alpha(h - g)} \\ R_{A/B}(\alpha) &= \frac{d - \alpha(d - c)}{e + \alpha(f - e)} \end{aligned} \quad (10.34)$$

Que, claramente, son expresiones que no corresponden a líneas rectas y por tanto el resultado no es un número trapezoidal.

Ejemplo 10.5 (División de números difusos) Consideremos los números trapezoidales

$$\begin{aligned} A &= T(5, 6, 7, 8) \\ B &= T(1, 2, 2, 3) \end{aligned} \quad (10.35)$$

La Figura 10.5 muestra los números A y B , así como el resultado de la división A/B . De acuerdo con la ecuación 7.12 sus respectivos α -cortes están definidos por

$$\begin{aligned} A_\alpha &= [-2 + \alpha(2 - (-2)), 6 - \alpha(6 - 4)] = [-2 + 4\alpha, 6 - 2\alpha] \\ B_\alpha &= [1 + \alpha(2 - 2), 3 - \alpha(3 - 1)] = [1 + \alpha, 3 - \alpha] \end{aligned} \quad (10.36)$$

Nótese que el número A no es positivo ni negativo. A continuación mostramos algunos de los α -cortes de los dos números:

$$\begin{aligned} A_{0.00} &= [-2.00, 6.00] & B_{0.00} &= [1.00, 3.00] \\ A_{0.25} &= [-1.00, 5.50] & B_{0.25} &= [1.25, 2.75] \\ A_{0.50} &= [0.00, 5.00] & B_{0.50} &= [1.50, 2.50] \\ A_{0.75} &= [1.00, 4.50] & B_{0.75} &= [1.75, 2.25] \\ A_{1.00} &= [2.00, 4.00] & B_{1.00} &= [2.00, 2.00] \end{aligned} \quad (10.37)$$

Es necesario destacar que los α -cortes correspondientes a $0 \leq \alpha < 0.5$ caen en el caso 5 de la Tabla 8.2, mientras que los casos correspondientes a $0.5 \leq \alpha \leq 1.0$ caen en el caso 1. En esa Tabla se verifica que los valores de $L_{A \cdot B}$ son diferentes en los dos casos, mientras que son iguales para $R_{A \cdot B}$:

$$L_{A/B}(\alpha) = \begin{cases} \frac{-2 + 4\alpha}{1 + \alpha} & \text{si } 0.0 \leq \alpha < 0.5 \\ \frac{-2 + 4\alpha}{3 - \alpha} & \text{si } 0.5 \leq \alpha < 1.0 \end{cases} \quad (10.38)$$

$$R_{A/B}(\alpha) = \frac{6 - 2\alpha}{1 + \alpha}$$

Este resultado puede visualizarse en la Figura 10.5. En efecto, allí se puede observar que la curva ascendente de $A \cdot B$ tiene un quiebre justo en $x = 0$, que corresponde a $\alpha = 0.5$, mientras que la curva descendente no presenta ese quiebre.

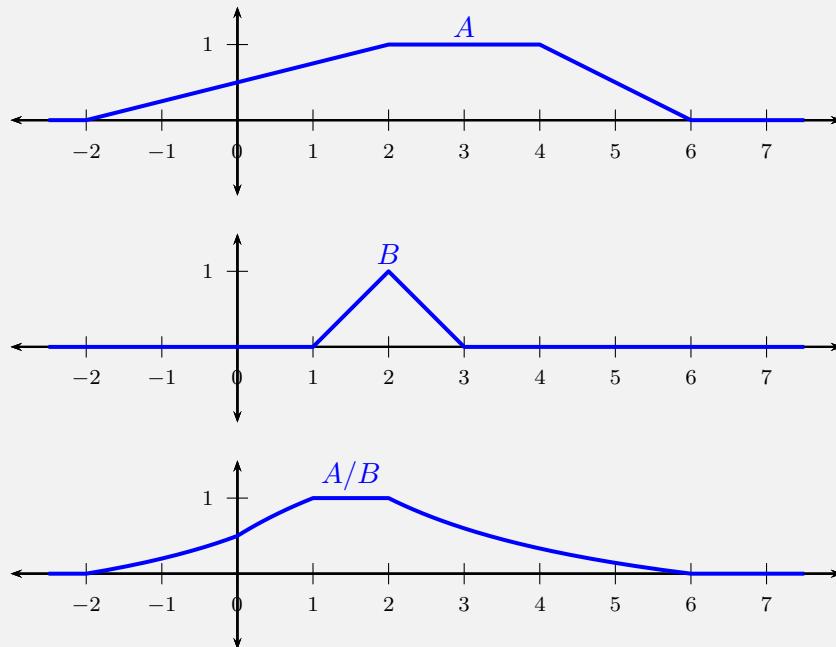


Figura 10.5: División de números difusos

Ejemplo 10.6 (División de números difusos positivos) Consideremos los números trapezoidales positivos

$$\begin{aligned} A &= T(5, 6, 7, 8) \\ B &= T(1, 2, 2, 3) \end{aligned} \quad (10.39)$$

La Figura 10.6 muestra los números A y B , así como el resultado de la división A/B . Al aplicar la ecuación 10.34 se obtiene

$$\begin{aligned} L_{A/B}(\alpha) &= \frac{5 + \alpha(6 - 5)}{3 - \alpha(3 - 2)} = \frac{5 + \alpha}{3 - \alpha} \\ R_{A/B}(\alpha) &= \frac{8 - \alpha(8 - 7)}{1 + \alpha(2 - 1)} = \frac{8 - \alpha}{1 + \alpha} \end{aligned} \quad (10.40)$$

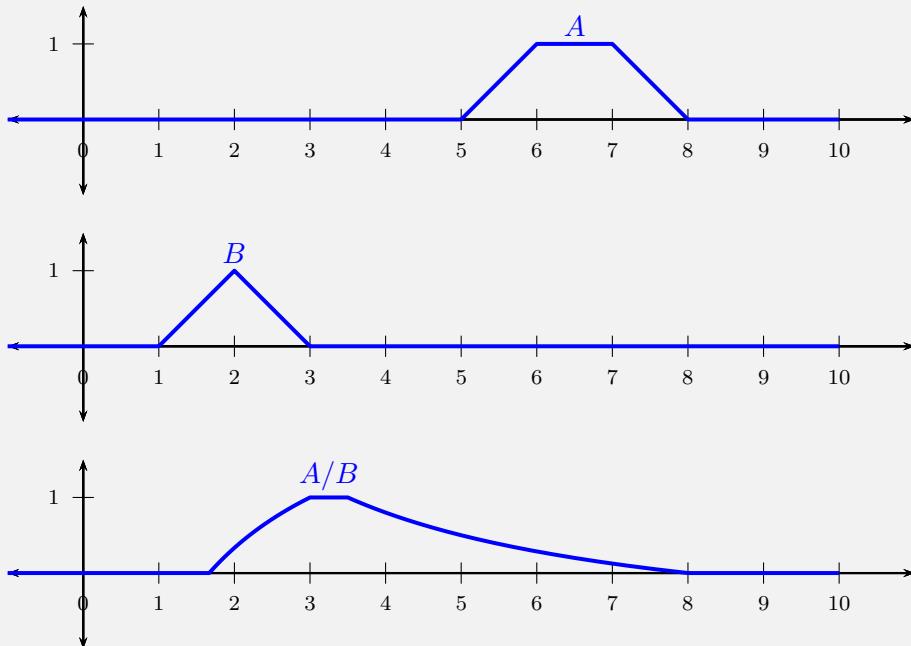


Figura 10.6: División de números difusos positivos

Tabla 10.1: Aritmética de trapecios positivos

	Definición
A	$T(a, b, c, d) > 0$
B	$T(e, f, g, h) > 0$
$A + B$	$T(a + e, b + f, c + g, d + h)$
$A - B$	$T(a - h, b - g, c - f, d - e)$
$A \cdot B$	$T(ae, bf, cg, dh)$
A/B	$T(a/h, b/g, c/f, d/e)$

10.5. Aritmética de trapecios positivos

En esta sección vamos a presentar una versión muy simplificada de la aritmética de números difusos. Se trata de una versión que considera que los únicos números difusos que existen son los trapezoidales positivos.

Comencemos por definir el conjunto \mathbb{T} de todos los posibles números trapezoidales positivos:

$$\mathbb{T} = \{T(a, b, c, d) \mid 0 < a \leq b \leq c \leq d < \infty\} \quad (10.41)$$

Y lo dotamos de cuatro operaciones a las que les damos los nombres y símbolos usuales (suma, resta, producto y división; $+$, $-$, \cdot , $/$). La Tabla 10.1 muestra la definición que hacemos de esas operaciones, aplicadas a dos números $A = T(a, b, c, d)$ y $B = T(e, f, g, h)$.

Con las definiciones de la Tabla 10.1 se tiene que:

- El operador de **suma** coincide con la expresión general de la suma de números trapezoidales de la ecuación 10.6
- El operador de **resta** coincide con la expresión general de la resta de números trapezoidales de la ecuación 10.15
- El operador de **producto** coincide con la expresión general del producto de números trapezoidales positivos de la ecuación 10.23 para los valores de $\alpha = 0$ y $\alpha=1$. En efecto:

$$\begin{aligned} L_{A \cdot B}(0) &= ae + 0(be + af - 2ae) + 0^2(bf + ae - be - af) = ae \\ L_{A \cdot B}(1) &= ae + 1(be + af - 2ae) + 1^2(bf + ae - be - af) = bf \\ R_{A \cdot B}(1) &= dh - 1(2dh - ch - dg) + 1^2(dh + cg - dg - ch) = cg \\ R_{A \cdot B}(0) &= dh - 0(2dh - ch - dg) + 0^2(dh + cg - dg - ch) = dh \end{aligned} \quad (10.42)$$

- El operador de **division** coincide con la expresión general de la división de números

trapezoidales positivos de la ecuación 10.34 para los valores de $\alpha = 0$ y $\alpha = 1$. En efecto:

$$\begin{aligned}
 L_{A/B}(0) &= \frac{a + 0(b - a)}{h - 0(h - g)} = \frac{a}{h} \\
 L_{A/B}(1) &= \frac{a + 1(b - a)}{h - 1(h - g)} = \frac{b}{g} \\
 R_{A/B}(1) &= \frac{d - 1(d - c)}{e + 1(f - e)} = \frac{c}{f} \\
 R_{A/B}(0) &= \frac{d - 0(d - c)}{e + 0(f - e)} = \frac{d}{e}
 \end{aligned} \tag{10.43}$$

Pese a las coincidencias, *la aritmética de trapecios positivos es diferente a la aritmética de números difusos aplicada a los trapecios positivos*. La primera considera únicamente dos valores de α , que son 0 y 1, mientras que la segunda considera todos los valores de α en el intervalo $[0, 1]$.

Si la aritmética de trapecios positivos sólo contempla los valores de α cero y uno ¿Qué pasa con los demás valores? No se necesitan para efectuar los cálculos. Sin embargo, si se llegan a requerir se puede efectuar una interpolación lineal que reconstruya el trapecio completo.

Los ejemplos 10.7 y 10.8 muestran las diferencias entre los dos tipos de aritmética, e ilustran lo parecidos que pueden llegar a ser los resultados obtenidos con ambas.

Ejemplo 10.7 (Aritmética de números trapezoidales positivos. Producto)

Retomemos el ejemplo 10.4 para analizarlo como números trapezoidales positivos.

El ejemplo consiste en obtener $A \cdot B$ con

$$\begin{aligned}
 A &= T(0, 1, 1, 2) \\
 B &= T(1, 2, 3, 4)
 \end{aligned} \tag{10.44}$$

Si trabajamos en el marco de los trapecios positivos, entonces:

$$A \in \mathbb{T} \quad B \in \mathbb{T} \quad A \cdot B \in \mathbb{T} \tag{10.45}$$

Y podemos calcular $A \cdot B$ usando la definición de la Tabla 10.1, que resulta en

$$A \cdot B = T(0, 2, 3, 8) \tag{10.46}$$

La Figura 10.7 compara los resultados obtenidos con los dos enfoques. En azul se ha trazado el resultado obtenido en el ejemplo 10.4 que han calculado los infinitos α -cortes, con $\alpha \in [0, 1]$. En rojo se ha trazado el resultado de trabajar en \mathbb{T} , en el que se han calculado los resultados de sólo dos α -cortes, con $\alpha \in \{0, 1\}$ y se ha efectuado una interpolación lineal para los demás valores.

Los resultados no son iguales, pero sí son muy parecidos. Dependiendo de la aplicación que se vaya a efectuar, es posible que la pérdida de información debida al simplificar el cálculo no sea significativa.

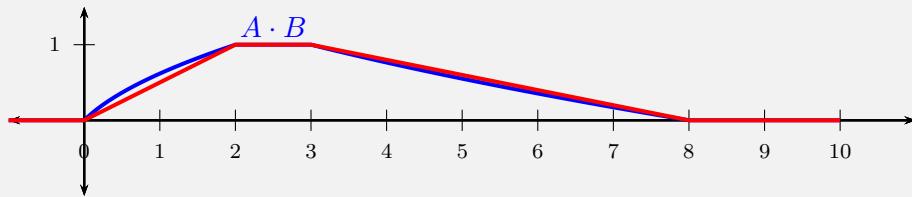


Figura 10.7: Comparación de resultados en el ejemplo 10.7

Ejemplo 10.8 (Aritmética de números trapezoidales positivos. División)

Retomemos el ejemplo 10.6 para analizarlo como números trapezoidales positivos.

El ejemplo consiste en obtener A/B con

$$\begin{aligned} A &= T(5, 6, 7, 8) \\ B &= T(1, 2, 2, 3) \end{aligned} \tag{10.47}$$

Si trabajamos en el marco de los trapecios positivos, entonces:

$$A \in \mathbb{T} \quad B \in \mathbb{T} \quad A/B \in \mathbb{T} \tag{10.48}$$

Y podemos calcular A/B usando la definición de la Tabla 10.1, que resulta en

$$A/B = T(5/3, 3, 7/2, 8) \tag{10.49}$$

La Figura 10.8 compara los resultados obtenidos con los dos enfoques. En azul se ha trazado el resultado obtenido en el ejemplo 10.6 que han calculado los infinitos α -cortes, con $\alpha \in [0, 1]$. En rojo se ha trazado el resultado de trabajar en \mathbb{T} , en el que se han calculado los resultados de sólo dos α -cortes, con $\alpha \in \{0, 1\}$ y se ha efectuado una interpolación lineal para los demás valores.

Los resultados no son iguales, pero sí son muy parecidos. Dependiendo de la aplicación que se vaya a efectuar, es posible que la pérdida de información debida al simplificar el cálculo no sea significativa.

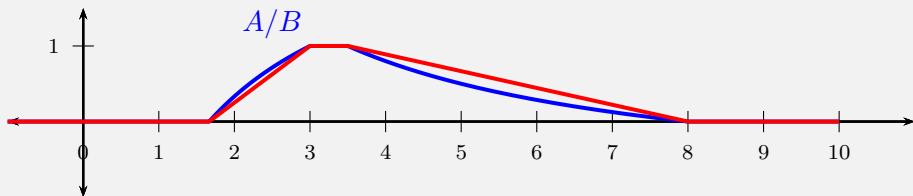


Figura 10.8: Comparación de resultados en el ejemplo 10.8

Capítulo 11

Algoritmos de extensión de funciones monótonas

En este capítulo se estudia cómo implementar en software algunos algoritmos que permitan aplicar a números difusos funciones definidas originalmente para hacerlo sobre números concretos [23, 24]). Específicamente, consideraremos el caso de una función multivariante estrictamente monótona y continua:

$$f(x_1, x_2, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R} \quad \text{estrictamente monótona y continua} \quad (11.1)$$

Y sus funciones inversas:

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2, x_3, \dots, x_n) \\ x_2 &= f_2^{-1}(y, x_1, x_3, \dots, x_n) \\ &\vdots \\ x_n &= f_n^{-1}(y, x_1, x_2, \dots, x_{n-1}) \end{aligned} \quad (11.2)$$

La pregunta general que se aborda en el capítulo es ¿cómo implementar algoritmos que permitan aplicar las funciones que se muestran en 11.1 y 11.2 a números difusos?

Se presupone que es posible implementar en software tanto $f(\cdot)$ como sus inversas y que, por tanto, lo que se necesita desarrollar son unos algoritmos que permitan *extender* el uso de esas implementaciones a números difusos. Específicamente, se presentan en este capítulo los siguientes algoritmos:

- **Extensión directa:** este algoritmo permite aplicar la función $f(\cdot)$ a números difusos. Este algoritmo se presenta en la sección 11.2.

- **Extensión inversa posible:** este algoritmo permite aplicar las funciones inversas de $f(\cdot)$ a números difusos, bajo la siguiente consideración: se tiene información, representada por números difusos, sobre el resultado de aplicar la función $f(\cdot)$ y de todos los valores de x , salvo de x_k ; el algoritmo permite encontrar un número difuso coherente con la información disponible. Este algoritmo se presenta en la sección 11.3.
- **Extensión inversa necesaria:** este algoritmo permite aplicar las funciones inversas de $f(\cdot)$ a números difusos, bajo la siguiente consideración: se tiene información, representada por números difusos, sobre todos los valores de x , salvo de x_k ; además, se desea que al aplicar la función $f(\cdot)$ se obtenga un resultado deseado. El algoritmo permite encontrar el número difuso faltante con el que se obtiene el resultado deseado, o al menos uno parecido. Este algoritmo se presenta en la sección 11.4.
- **Extensión inversa intermedia:** las extensiones inversas posible y necesaria pueden visualizarse como dos casos extremos de una misma forma de extender las funciones inversas. Este algoritmo, que se presenta en la sección 11.5 permite encontrar las soluciones intermedias entre esos dos extremos.

Adicionalmente, en la sección 11.6 se presenta un concepto teórico, y su respectivo algoritmo, que permite evaluar la existencia de la extensión de funciones inversas. Este concepto es relevante, por que si bien es cierto que la extensión inversa posible siempre existe, la extensión inversa necesaria puede no existir (Cuando esto último sucede, el algoritmo propuesto modifica las condiciones del problema para obtener una buena solución).

Ejemplo 11.1 (Funciones a extender) Para ilustrar el objetivo de este capítulo, consideremos el caso de un hiperplano:

$$y = f(x_1, x_2, x_3) = \sigma_0 + \sigma_1 x_1 + \sigma_2 x_2 + \sigma_3 x_3 \quad (11.3)$$

$\sigma_k \neq 0 \quad k = 1, 2, \dots, n$

La restricción $\sigma_k \neq 0$ se introduce para asegurar que $f(\cdot)$ sea estrictamente monótona. En esas condiciones es posible despejar cada una de las n variables que forman parte del argumento de la función y y obtener las funciones inversas, que resultan ser

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2, x_3) = \frac{y - \sigma_0 - \sigma_2 x_2 - \sigma_3 x_3}{\sigma_1} \\ x_2 &= f_2^{-1}(y, x_1, x_2) = \frac{y - \sigma_0 - \sigma_1 x_1 - \sigma_3 x_3}{\sigma_2} \\ x_3 &= f_3^{-1}(y, x_1, x_2) = \frac{y - \sigma_0 - \sigma_1 x_1 - \sigma_2 x_2}{\sigma_3} \end{aligned} \quad (11.4)$$

Los algoritmos que se muestran en este capítulo permiten aplicar $f(\cdot)$ y sus tres funciones inversas a números difusos. El algoritmo de extensión directa, por ejemplo, permiten calcular $f(A_1, A_2, A_3)$ cuando A_k son unos números trapezoidales

$$\begin{aligned} A_1 &= T(a_1, b_1, c_1, d_1) \\ A_2 &= T(a_2, b_2, c_2, d_2) \\ A_3 &= T(a_3, b_3, c_3, d_3) \end{aligned} \quad (11.5)$$

11.1. Representación discreta de un número difuso

Los algoritmos propuestos hacen uso del Teorema 9.1, que establece que la extensión de funciones a los conjuntos difusos puede resolverse para cada α -corte de forma independiente. Por esa razón, los algoritmos operan sobre números difusos especificados a partir de sus α -cortes. Más exactamente, operan sobre números especificados a partir de un conjunto finito de sus α -cortes. Igualmente, los resultados de los algoritmos son un conjunto finito de α -cortes de los números obtenidos.

Definimos el conjunto Υ como al conjunto finito de α -cortes de interés

$$\begin{aligned} \Upsilon &= \{\alpha_1, \alpha_2, \dots, \alpha_p\} \\ 0 = \alpha_1 < \alpha_2 < \dots < \alpha_{p-1} < \alpha_p = 1 \end{aligned} \quad (11.6)$$

Si, además, los α -cortes están igualmente espaciados se obtiene

$$\begin{aligned} \Upsilon &= \{0.0, d\alpha, 2d\alpha, \dots, (p-2)d\alpha, 1.0\} \\ d\alpha &= 1/(p-1) \end{aligned} \quad (11.7)$$

De acuerdo con lo anterior, la representación discreta de un número difuso N consiste en una colección de α -cortes:

$$\{[L_N(\alpha_1), R_N(\alpha_1)], [L_N(\alpha_2), R_N(\alpha_2)], \dots, [L_N(\alpha_p), R_N(\alpha_p)]\} \quad (11.8)$$

La representación discreta empleada en los algoritmos que se muestran en este capítulo se realiza utilizando tres arreglos de número reales, que contienen toda la información que se requiere para construir esa colección de α -cortes:

$$N \cong \begin{cases} \Upsilon &= [\alpha_1, \alpha_2, \dots, \alpha_p] \\ L_N &= [L_N(\alpha_1), L_N(\alpha_2), \dots, L_N(\alpha_p)] \\ R_N &= [R_N(\alpha_1), R_N(\alpha_2), \dots, R_N(\alpha_p)] \end{cases} \quad (11.9)$$

En donde hemos utilizado el símbolo \cong para indicar que el número N es representado por la expresión de la derecha.

Ejemplo 11.2 (Representación discreta de un número difuso) Consideremos el número trapezoidal $N = T(1, 2, 3, 4)$. Para obtener una representación discreta de 6 α -cortes igualmente espaciados podemos aplicar la información de la Tabla 7.1 y obtener:

$$N \cong \begin{cases} \Upsilon &= [0.0, 0.2, 0.4, 0.6, 0.8, 1.0] \\ L_N &= [1.0, 1.2, 1.4, 1.6, 1.8, 2.0] \\ R_N &= [4.0, 3.8, 3.6, 3.4, 3.2, 3.0] \end{cases} \quad (11.10)$$

11.2. Extensión directa

Supongamos que queremos aplicar la función $f(\cdot)$ definida en 11.1 a un conjunto de números difusos $A = \{A_1, A_2, \dots, A_n\}$. Para ello, lo primero que haremos será definir un conjunto de p α -cortes discretos Υ (ecuación 11.6).

Al utilizar la representación discreta, el cálculo de $B = f(A)$ se traduce en el cálculo de sus arreglos L_B, R_B . Es decir,

$$B \cong \begin{cases} \Upsilon &= [\alpha_1, \alpha_2, \dots, \alpha_p] \\ L_B &= [L_B(\alpha_1), L_B(\alpha_2), \dots, L_B(\alpha_p)] \\ R_B &= [R_B(\alpha_1), R_B(\alpha_2), \dots, R_B(\alpha_p)] \end{cases} \quad (11.11)$$

Debido a que hemos impuesto a $f(\cdot)$ las condiciones de continuidad y monotonía podemos aplicar el Teorema 8.2 para calcular cada uno de los α -cortes de B . Para ello, definimos

$$D_{A_k}(\alpha, d) = \begin{cases} L_{A_k}(\alpha) & \text{si } d = 0 \\ R_{A_k}(\alpha) & \text{si } d = 1 \end{cases} \quad (11.12)$$

La función $D_{A_k}(\alpha, d)$ devuelve uno de los dos extremos del α -corte de la k -ésima variable. El extremo se selecciona con la variable booleana d .

Por otra parte, como $f(\cdot)$ es estrictamente monótona, la ecuación 8.28 se convierte en

$$d_f(k) = \begin{cases} 0 & \text{si } f \text{ es monótonamente creciente en } x_k \\ 1 & \text{si } f \text{ es monótonamente decreciente en } x_k \end{cases} \quad (11.13)$$

Y construimos el arreglo d_f que contiene los valores booleanos que determinan el sentido de la monotonía de $f(\cdot)$ en cada una de sus n variables:

$$d_f = \{d_f(1), d_f(2), \dots, d_f(n)\} \quad (11.14)$$

Asumimos la variable $d_f(k)$ como booleana, lo que nos permite emplear el operador de negación¹, denotado por \neg , para definir los siguientes arreglos:

$$\begin{aligned} \mathcal{D}_A(\alpha, d_f) &= \{D_{A_1}(\alpha, d_f(1)), D_{A_2}(\alpha, d_f(2)), \dots, D_{A_n}(\alpha, d_f(n))\} \\ \mathcal{D}_A(\alpha, \neg d_f) &= \{D_{A_1}(\alpha, \neg d_f(1)), D_{A_2}(\alpha, \neg d_f(2)), \dots, D_{A_n}(\alpha, \neg d_f(n))\} \end{aligned} \quad (11.15)$$

La aplicación de las ecuaciones 8.36 a 8.40 del Teorema 8.2 resulta en una forma práctica de calcular los límites de cada α -corte de B :

$$\begin{aligned} B(\alpha) &= f(A_k)(\alpha) = [f(\mathcal{L}_A(\alpha)), f(\mathcal{R}_A(\alpha))] \\ \mathcal{L}_A(\alpha) &= \mathcal{D}_A(\alpha, d_f) \\ \mathcal{R}_A(\alpha) &= \mathcal{D}_A(\alpha, \neg d_f) \end{aligned} \quad (11.16)$$

Lo anterior nos permite formular el algoritmo 8 que permite obtener el resultado de aplicar $f(\cdot)$ a una colección de números difusos.

Algoritmo 8 Extensión directa

Entradas:

- f una función (11.1)
 - d_f un vector de booleanos (11.14)
 - A un vector n de números difusos con igual Υ
-

```

1: procedure DIRECTA( $f, d_f, A$ )
2:    $p \leftarrow \text{size}(\Upsilon)$                                  $\triangleright$  Número de  $\alpha$ -cortes
3:    $L_B \leftarrow \text{zeros}(p)$                              $\triangleright$  Vector vacío
4:    $R_B \leftarrow \text{zeros}(p)$                              $\triangleright$  Vector vacío
5:    $j \leftarrow 1$ 
6:   while  $j < p$  do                                 $\triangleright$  Calcular para cada  $\alpha$ -corte
7:      $\alpha \leftarrow \Upsilon[j]$ 
8:      $L_B[j] \leftarrow f(\mathcal{L}_A(\alpha))$                    $\triangleright$  Ecuaciones 11.15 y 11.16
9:      $R_B[j] \leftarrow f(\mathcal{R}_A(\alpha))$                    $\triangleright$  Ecuaciones 11.15 y 11.16
10:     $j \leftarrow j + 1$ 
11:   return  $L_B, R_B$                                  $\triangleright$  Vectores con los límites de  $B = f(A)$  para cada  $\alpha$ -corte

```

¹ $\neg 0 = 1$ y $\neg 1 = 0$

Ejemplo 11.3 (Extensión directa de un hiperplano) Consideremos la función $f(x_1, x_2, x_3) : \mathbb{R}^3 \rightarrow \mathbb{R}$ que define un hiperplano en \mathbb{R}^4 :

$$y = f(x_1, x_2, x_3) = 1 + x_1 - 2x_2 - x_3 \quad (11.17)$$

Se desea calcular $B = f(A_1, A_2, A_3)$, para los siguientes números difusos:

$$\begin{aligned} A_1 &= T(0, 1, 2, 3) \\ A_2 &= T(2, 3, 4, 5) \\ A_3 &= T(-3, -2, -1, 0) \end{aligned} \quad (11.18)$$

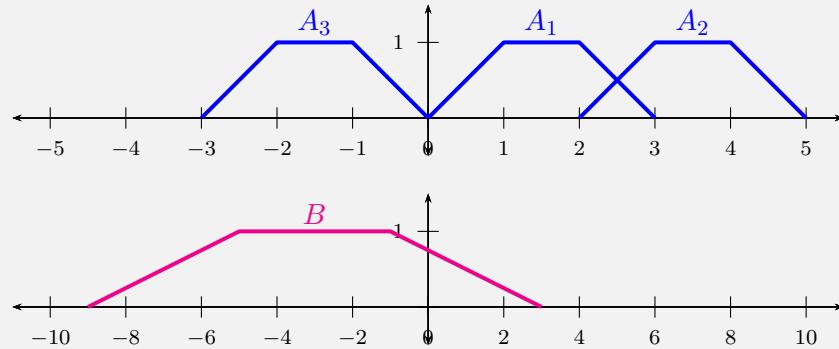


Figura 11.1: Números del ejemplo 11.3

La Figura 11.1 muestra los números difusos de entrada A_1, A_2, A_3 y el resultado B de aplicar el algoritmo de extensión directa.

Exploraremos un poco cómo ha operado este algoritmo. Iniciemos por destacar que $f(\cdot)$ efectivamente es una función continua y monótona (Véase el ejemplo 8.6) y que el sentido de las monotones está dado por el signo de los coeficientes que acompañan a cada variable, por lo tanto:

$$d_f(1) = 0 \quad d_f(2) = 1 \quad d_f(3) = 1 \quad (11.19)$$

Por otra parte, se ha utilizado una representación discreta que emplea 5 α -cortes igualmente espaciados. La Tabla 11.1 muestra los límites inferior (L) y superior (R) para cada uno de los números y cada uno de los α -cortes. Además, en esa tabla se han utilizado dos colores para explicar el funcionamiento del algoritmo, así:

Para un mismo α -corte, los valores del mismo color están relacionados por la ecuación 11.17. Por ejemplo, para $\alpha = 0$ se tiene que

$$\begin{aligned} -9.0 &= 1.0 + 0.0 - 2 \times 5.0 - 0.0 \\ 3.0 &= 1.0 + 3.0 - 2 \times 2.0 - (-3.0) \end{aligned} \quad (11.20)$$

Tabla 11.1: Números del ejemplo 11.3

α	A_1		A_2		A_3		B	
	L	R	L	R	L	R	L	R
0.000	0.000	3.000	2.000	5.000	-3.000	0.000	-9.000	3.000
0.250	0.250	2.750	2.250	4.750	-2.750	-0.250	-8.000	2.000
0.500	0.500	2.500	2.500	4.500	-2.500	-0.500	-7.000	1.000
0.750	0.750	2.250	2.750	4.250	-2.250	-0.750	-6.000	0.000
1.000	1.000	2.000	3.000	4.000	-2.000	-1.000	-5.000	-1.000

El resultado obtenido es coherente al que resulta de aplicar los principios de aritmética difusa:

$$B = T(-9, -7, -1, 3) \quad (11.21)$$

Ejemplo 11.4 (Extensión directa de una función positiva) Consideremos la función $f(x_1, x_2) : \mathbb{R}^{+2} \rightarrow \mathbb{R}^{+}$ definida por:

$$y = f(x_1, x_2) = \frac{x_1^2}{x_2} \quad (11.22)$$

Se desea calcular $B = f(A_1, A_2)$, para los siguientes números difusos:

$$\begin{aligned} A_1 &= T(1.0, 1.8, 2.2, 3.0) \\ A_2 &= T(0.5, 0.9, 1.1, 1.5) \end{aligned} \quad (11.23)$$

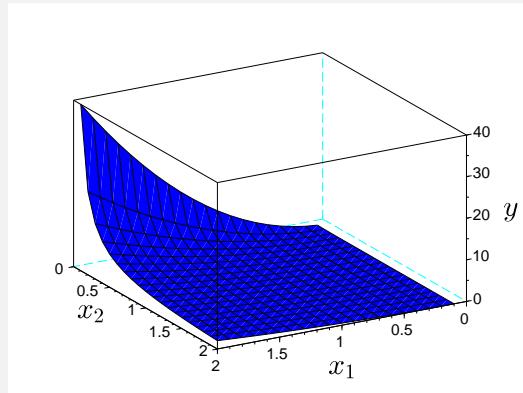


Figura 11.2: Función del ejemplo 11.4

Destaquemos que $f(\cdot)$ es continua y monótona en el universo de discurso que se ha definido, los reales positivos. La Figura 11.2 muestra la superficie de $f(\cdot)$ en \mathbb{R}^3 . Una implementación

en software de esa función debe establecer un valor mínimo de x_2 , para evitar la división por 0. En este ejemplo se ha utilizado 0.1. El sentido de la monotonía implica que

$$d_f(1) = 0 \quad d_f(2) = 1 \quad (11.24)$$

La Figura 11.3 muestra los números difusos de entrada A_1, A_2 y el resultado B de aplicar el algoritmo de extensión directa.

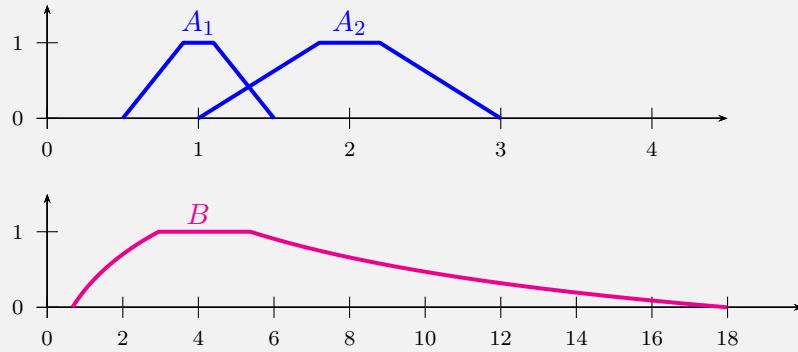


Figura 11.3: Números del ejemplo 11.4

Ejemplo 11.5 (Extensión directa de una función acotada) Consideremos la función $f(x_1, x_2) : (0, 1]^2 \rightarrow (0, 1]$ definida por:

$$y = f(x_1, x_2) = \cos(x_1\pi/2) \sin(x_2\pi/2) \quad (11.25)$$

Se desea calcular $B = f(A_1, A_2)$, para los siguientes números difusos:

$$\begin{aligned} A_1 &= T(0.0, 0.1, 0.2, 0.3) \\ A_2 &= T(0.2, 0.3, 0.4, 0.5) \end{aligned} \quad (11.26)$$

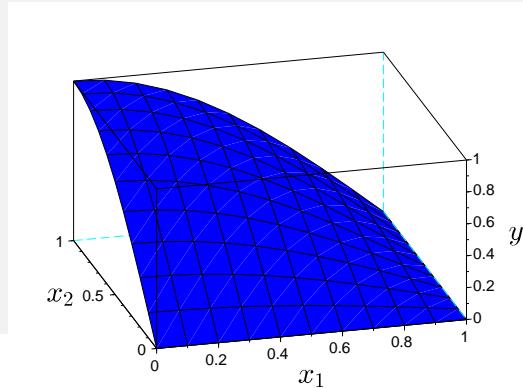


Figura 11.4: Función del ejemplo 11.5

Destaquemos que $f(\cdot)$ es continua y monótona en el universo de discurso que se ha definido, los intervalos unitarios. La Figura 11.4 muestra la superficie de $f(\cdot)$ en \mathbb{R}^3 . Una implementación en software de esa función debe establecer un valor mínimo de x_1 y de x_2 , porque en esos valores la función no es creciente ni decreciente. Esa es la razón por la que la función se define sobre intervalos abiertos en 0. En este ejemplo se ha utilizado como límite 0.01. El sentido de la monotonía implica que

$$d_f(1) = 1 \quad d_f(2) = 0 \quad (11.27)$$

La Figura 11.5 muestra los números difusos de entrada A_1, A_2 y el resultado B de aplicar el algoritmo de extensión directa.

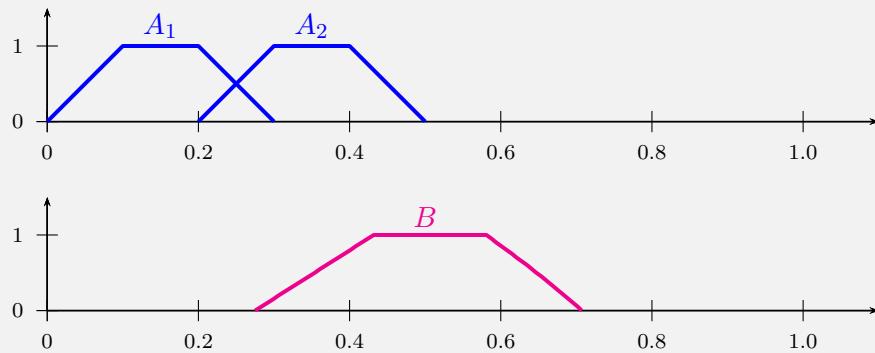


Figura 11.5: Números del ejemplo 11.4

Pese a que la función no es lineal, el resultado es muy cercano a un número trapezoidal. Lo anterior, debido a que la superficie de $f(\cdot)$ puede ser aproximada sin mucho error a un plano tangente en cercanías de cualquiera de sus puntos.

11.3. Extensión inversa posible

Supongamos que queremos aplicar una de las funciones inversas 11.2:

$$x_k = f_k^{-1}(y, x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \quad (11.28)$$

En las siguientes condiciones: la función $f(\cdot)$ se está utilizando para modelar un cierto fenómeno que relaciona las variables de entrada x con la variable de salida y ; además, se tiene información sobre el valor de la salida y y del de todas las variables de entrada x , salvo del de x_k . La información disponible, sin embargo, contiene incertidumbre y por ello se ha optado por representarla mediante números difusos.

En esas condiciones, se desea determinar que valores *puede* tener x_k que sean compatibles con la información disponible.

Ejemplo 11.6 (Extensión inversa posible en un intervalo) Un breve ejemplo haciendo uso de la aritmética intervalar puede ayudar a aclarar el problema que se ha formulado. Supongamos que $y = x_1 - x_2$ y que se sabe que $y \in [1, 3]$ y $x_2 \in [1, 2]$. Lo que se desea es encontrar todos los valores de x_1 tales que es posible encontrar una terna $\{x_1, x_2, y\}$ que cumpla simultáneamente las tres condiciones:

$$\begin{cases} y &= x_1 - x_2 \\ x_2 &\in [1, 2] \\ y &\in [1, 3] \end{cases} \quad (11.29)$$

Por ejemplo, el número $x_1 = 5$ es uno de los valores buscados, porque $x_1 = 5, x_2 = 2, y = 3$ forman una terna que cumple con las tres condiciones. De hecho, para todos los valores de x_1 que caen en el intervalo $[2, 5]$ es posible encontrar una terna adecuada $\{x_1, x_2, y\}$, y por ello podemos decir que

$$x_1^{pos} = [2, 5] \quad (11.30)$$

Por la forma en que está formulado el problema, es equivalente a aplicar el algoritmo de extensión directa a la función $f_k^{-1}(\cdot)$ de la ecuación 11.28.

Comencemos por simplificar la notación definiendo X_k como el arreglo que contiene todas las entradas x , salvo la k -ésima:

$$X_k = \{x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_n\} \quad (11.31)$$

Y aclarando que queremos obtener los valores *posibles* de x_k , es decir, los que son compatibles con la información disponible.

$$x_k^{pos} = f_k^{-1}(y, X_k) = g(y, X_k) \quad (11.32)$$

Para aplicar el algoritmo de extensión directa a 11.32 necesitamos conocer el sentido de la monotonía de $g(\cdot)$ en sus variables y, X_k que, a su vez, dependerá de las derivadas parciales. Para ello, consideraremos el hiperplano tangente a $f(\cdot)$ que se obtiene al tomar su polinomio de Taylor de primer orden.

$$y - p_0 = a_1(x_1 - p_1) + a_2(x_2 - p_2) + \dots + a_n(x_n - p_n) \quad (11.33)$$

En donde $P = \{p_1, p_2, \dots, p_n\}$ es un punto en \mathbb{R}^n y además

$$\begin{aligned} p_0 &= f(P) \\ a_j &= \left. \frac{\partial f(X)}{\partial x_j} \right|_{X=P} \quad j = 1, 2, \dots, n \end{aligned} \tag{11.34}$$

Lo que nos permite despejar el término $(x_k - p_k)$:

$$\begin{aligned} x_k - p_k &= \frac{1}{a_k}(y - p_0) - \frac{a_1}{a_k}(x_1 - p_1) \cdots \\ &\quad - \frac{a_{k-1}}{a_k}(x_{k-1} - p_{k-1}) - \frac{a_{k+1}}{a_k}(x_{k+1} - p_{k+1}) \cdots \\ &\quad - \frac{a_n}{a_k}(x_n - p_n) \end{aligned} \tag{11.35}$$

La ecuación 11.35 corresponde al hiperplano que resulta de efectuar una aproximación de primer orden del polinomio de Taylor de $g(y, X_k)$. Por lo tanto, podemos afirmar que

$$\begin{aligned} \frac{\partial g(y, X_k)}{\partial y} &= \frac{1}{a_k} = \frac{1}{\frac{\partial f(X)}{\partial x_k}} \\ \frac{\partial g(y, X_k)}{\partial x_j} &= -\frac{a_j}{a_k} = \frac{\frac{\partial f(X)}{\partial x_j}}{\frac{\partial f(X)}{\partial x_k}} \end{aligned} \tag{11.36}$$

Ahora bien, como los sentidos de las monotonías dependen de los signos de las derivadas parciales, y recordando que $g(y, X_k) = f_k^{-1}(y, X_k)$, los resultados de la ecuación 11.36 pueden escribirse así:

- El sentido de la monotonía de $f_k^{-1}(y, X_k)$ respecto a y es el mismo sentido de la monotonía de $f(X)$ respecto a x_k .
- El sentido de la monotonía de $f_k^{-1}(y, X_k)$ respecto a x_j está determinado por el sentido de la monotonía de $f(X)$ respecto a x_j y a x_k así:
 - Si los dos sentidos de la monotonía de $f(X)$ respecto a x_j y a x_k son iguales, la monotonía de $f_k^{-1}(y, X_k)$ respecto a x_j es decreciente.
 - Si los dos sentidos de la monotonía de $f(X)$ respecto a x_j y a x_k son diferentes, la monotonía de $f_k^{-1}(y, X_k)$ respecto a x_j creciente.

Las mismas conclusiones se muestran de forma compacta en la Tabla 11.2. Allí se ha utilizado la definición de $d_f(\cdot)$ que se encuentra en la ecuación 11.13.

Si, además, definimos

Tabla 11.2: Sentido de las monotonías de funciones inversas

	Respecto a x_j		Respecto a y
	$d_f(j) : \nearrow$	$d_f(j) : \searrow$	
$d_f(k) : \nearrow$	\searrow	\nearrow	\nearrow
$d_f(k) : \searrow$	\nearrow	\searrow	\searrow

$$\begin{aligned}
d_{f_y}(k) &= \begin{cases} 0 & \text{si } f_k^{-1}(y, Xk) \text{ es monótonamente creciente en } y \\ 1 & \text{si } f_k^{-1}(y, Xk) \text{ es monótonamente decreciente en } y \end{cases} \\
d_f(j, k) &= \begin{cases} 0 & \text{si } f_k^{-1}(y, Xk) \text{ es monótonamente creciente en } x_j \\ 1 & \text{si } f_k^{-1}(y, Xk) \text{ es monótonamente decreciente en } x_y \end{cases}
\end{aligned} \tag{11.37}$$

Entonces podemos hacer uso del operador XOR, denotado por \odot para escribir

$$\begin{aligned}
d_{f_y}(k) &= d_f(k) \\
d_f(j, k) &= \neg d_f(k) \odot d_f(j)
\end{aligned} \tag{11.38}$$

La ecuación 11.38 nos permite calcular el sentido de las monotonías de las funciones inversas 11.2. Podemos ahora aplicar el algoritmo de extensión directa para obtener la extensión inversa posible así. Para ello, definimos

$$A_k = \{A_1, A_2, \dots, A_{k-1}, A_{k+1}, \dots, A_n\} \tag{11.39}$$

Un arreglo de números difusos que representan la información sobre las variables de entrada X_k , y un número difuso Y que representa la información sobre la variable de salida y . El resultado de la extensión inversa posible de $f_k^{-1}(y, Xk)$ será un número difuso x_k^{pos} cuyos α -cortes serán

$$x_k^{pos}(\alpha) = [L_{x_k^{pos}}(\alpha), R_{x_k^{pos}}(\alpha)] \tag{11.40}$$

Para obtener los límites de esos α -cortes se debe aplicar la función inversa en un conjunto adecuado de valores

$$x_k^{pos}(\alpha) = [f_k^{-1}(\mathcal{L}_Y^{pos}(\alpha), \mathcal{L}_{A_k}^{pos}(\alpha)), f_k^{-1}(\mathcal{R}_Y^{pos}(\alpha), \mathcal{R}_{A_k}^{pos}(\alpha))] \tag{11.41}$$

En donde los valores de Y que hay que utilizar están determinados por

$$\begin{aligned}\mathcal{L}_Y^{pos}(\alpha) &= \mathcal{D}_Y(\alpha, d_f(k)) \\ \mathcal{R}_Y^{pos}(\alpha) &= \mathcal{D}_Y(\alpha, \neg d_f(k))\end{aligned}\quad (11.42)$$

Y los de las variables X_k están determinados por

$$\begin{aligned}\mathcal{L}_{A_k}^{pos}(\alpha) &= \left\{ \mathcal{L}_{A_1}^{pos}(\alpha), \mathcal{L}_{A_2}^{pos}(\alpha), \dots, \mathcal{L}_{A_{k-1}}^{pos}(\alpha), \mathcal{L}_{A_{k+1}}^{pos}(\alpha), \dots, \mathcal{L}_{A_n}^{pos}(\alpha) \right\} \\ \mathcal{R}_{A_k}^{pos}(\alpha) &= \left\{ \mathcal{R}_{A_1}^{pos}(\alpha), \mathcal{R}_{A_2}^{pos}(\alpha), \dots, \mathcal{R}_{A_{k-1}}^{pos}(\alpha), \mathcal{R}_{A_{k+1}}^{pos}(\alpha), \dots, \mathcal{R}_{A_n}^{pos}(\alpha) \right\} \\ \mathcal{L}_{A_j}^{pos}(\alpha) &= \mathcal{D}(\alpha, \neg d_f(k) \odot d_f(j)) \\ \mathcal{R}_{A_j}^{pos}(\alpha) &= \mathcal{D}(\alpha, d_f(k) \odot d_f(j))\end{aligned}\quad (11.43)$$

Podemos ahora establecer el algoritmo 9, que resume el procedimiento para efectuar la extensión inversa posible de una función de multiples variables, aplicada a números difusos.

Algoritmo 9 Extensión inversa posible

Entradas:

- f_k^{-1} una función (11.2)
 - d_f un vector de booleanos (11.14)
 - A_k un vector $n - 1$ de números difusos con igual Υ
 - Y un número difuso con igual Υ que A_k
-

```

1: procedure INVERSAPOSIBLE( $f_k^{-1}, d_f, Y, A_k$ )
2:    $p \leftarrow \text{size}(\Upsilon)$                                  $\triangleright$  Número de  $\alpha$ -cortes
3:    $L_{A_k} \leftarrow \text{zeros}(p)$                            $\triangleright$  Vector vacío
4:    $R_{A_k} \leftarrow \text{zeros}(p)$                            $\triangleright$  Vector vacío
5:    $j \leftarrow 1$ 
6:   while  $j < p$  do                                      $\triangleright$  Calcular para cada  $\alpha$ -corte
7:      $\alpha \leftarrow \Upsilon[j]$ 
8:      $L_{A_k}[j] \leftarrow f_k^{-1} \left( \mathcal{L}_Y^{pos}(\alpha), \mathcal{L}_{A_k}^{pos}(\alpha) \right)$        $\triangleright$  Ecuaciones 11.41, 11.42 y 11.43
9:      $R_{A_k}[j] \leftarrow f_k^{-1} \left( \mathcal{R}_Y^{pos}(\alpha), \mathcal{R}_{A_k}^{pos}(\alpha) \right)$        $\triangleright$  Ecuaciones 11.41, 11.42 y 11.43
10:     $j \leftarrow j + 1$ 
11:   return  $L_{A_k}, R_{A_k}$                                  $\triangleright$  Vectores con los límites de  $A_k = f_k^{-1}(Y, A_k)$  para cada  $\alpha$ -corte

```

Ejemplo 11.7 (Extensión posible de un hiperplano) En este ejemplo consideramos la misma función $f(\cdot)$ del ejemplo 11.3:

$$y = f(x_1, x_2, x_3) = 1 + x_1 - 2x_2 - x_3 \quad (11.44)$$

Cuyas funciones inversas son

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2, x_3) = y - 1 + 2x_2 + x_3 \\ x_2 &= f_2^{-1}(y, x_1, x_3) = (y - 1 - x_1 + x_3) / -2 \\ x_3 &= f_3^{-1}(y, x_1, x_2) = y - 1 - x_1 + 2x_2 \end{aligned} \quad (11.45)$$

En este ejemplo se busca obtener la extensión inversa posible en la segunda variable, $A_2^{pos} = f_2^{-1}(B, A_1, A_3)$, con los siguientes números difusos:

$$\begin{aligned} B &= T(-4, -3, -2, -1) \\ A_1 &= T(0, 1, 2, 3) \\ A_3 &= T(-3, -2, -1, 0) \end{aligned} \quad (11.46)$$

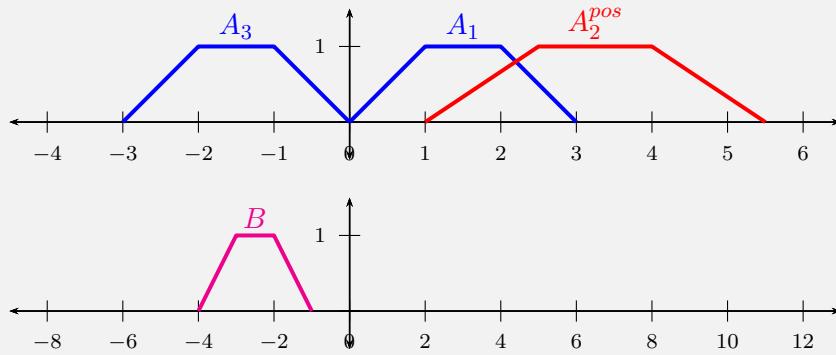


Figura 11.6: Números del ejemplo 11.7

La Figura 11.6 muestra los números difusos B, A_1, A_3 y el resultado A_2^{pos} de aplicar el algoritmo de extensión inversa posible. Recordemos que el sentido de la monotonía de $f(\cdot)$ determina que:

$$d_f(1) = 0 \quad d_f(2) = 1 \quad d_f(3) = 1 \quad (11.47)$$

Como se quiere aplicar la inversa a la segunda variable, entonces $k = 2$ y por tanto $d_f(k) = 1$.

Definamos D_X como una variable auxiliar que puede valer 0 ó 1. Valdrá 0 si se necesita el límite inferior para obtener A_2^{pos} en cada α -corte y 1 en caso contrario. Aplicando las ecuaciones 11.41, 11.42 y 11.43 se obtiene

$$\begin{aligned} D_{A_1} &= \neg d_f(k) \odot d_f(1) = 0 \odot 0 = 0 \\ D_{A_2} &= \neg d_f(k) \odot d_f(3) = 0 \odot 1 = 0 \\ D_{A_3} &= \neg d_f(k) \odot d_f(3) = 0 \odot 1 = 1 \\ D_B &= d_f(k) = 1 \end{aligned} \quad (11.48)$$

La Tabla 11.3 muestra los límites inferior (L) y superior (R) para cada uno de los números y cada uno de los α -cortes. Además, en esa tabla se han utilizado dos colores para explicar el funcionamiento del algoritmo, así: Para un mismo α -corte, los valores del mismo color están relacionados por la ecuación $x_2 = (y - 1 - x_1 + x_3) / -2$. Eso implica que los colores de las columnas L y R de cada variable se pueden seleccionar usando lo calculado en la ecuación 11.48

Por ejemplo, para $\alpha = 0$ se tiene que

$$\begin{aligned} 1.0 &= (-1.0 - 1 - (0.0) + (0.0)) / -2 \\ 5.5 &= (-4.0 - 1 - (3.0) + (-3.0)) / -2 \end{aligned} \quad (11.49)$$

Tabla 11.3: Números del ejemplo 11.7

α	A_1		A_2		A_3		B	
	L	R	L	R	L	R	L	R
0.000	0.000	3.000	1.000	5.500	-3.000	0.000	-4.000	-1.000
0.250	0.250	2.750	1.375	5.125	-2.750	-0.250	-3.750	-1.250
0.500	0.500	2.500	1.750	4.750	-2.500	-0.500	-3.500	-1.500
0.750	0.750	2.250	2.125	4.375	-2.250	-0.750	-3.250	-1.750
1.000	1.000	2.000	2.500	4.000	-2.000	-1.000	-3.000	-2.000

Ejemplo 11.8 (Extensión posible de una función positiva) En este ejemplo consideramos la misma función $f(\cdot)$ del ejemplo 11.4:

$$y = f(x_1, x_2) = \frac{x_1^2}{x_2} \quad (11.50)$$

Cuyas funciones inversas son

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2) = \sqrt{yx_2} \\ x_2 &= f_2^{-1}(y, x_1) = \frac{x_1^2}{y} \end{aligned} \quad (11.51)$$

En este ejemplo se busca obtener la extensión inversa posible en cada una de las dos variables:

$$\left\{ \begin{array}{l} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1^{pos} = ? \\ A_2 = T(0.5, 0.9, 1.1, 1.5) \end{array} \right. \quad \left\{ \begin{array}{l} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1 = T(1.0, 1.8, 2.2, 3.0) \\ A_2^{pos} = ? \end{array} \right. \quad (11.52)$$

La Figura 11.7 muestra los resultados de la aplicación del algoritmo de extensión inversa posible a cada uno de los dos casos. Debe destacarse que dada la naturaleza no lineal de las funciones, se obtienen números que no son trapezoidales. En este ejemplo se ha utilizado una representación discreta de 21 α -cortes, suficiente para que las funciones de pertenencia resultantes sean suaves.

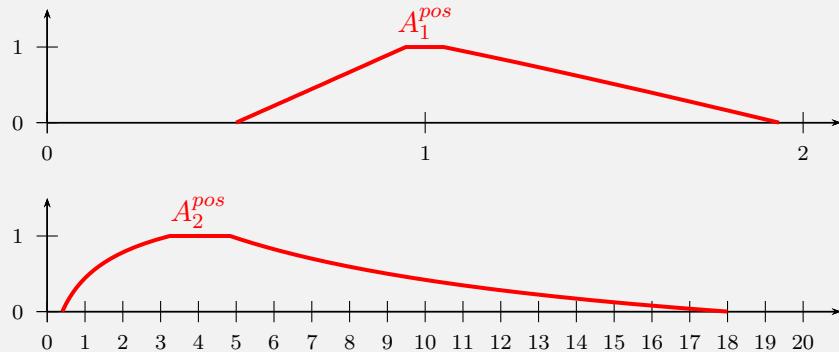


Figura 11.7: Solución del ejemplo 11.8

Ejemplo 11.9 (Extensión posible de una función acotada) En este ejemplo consideramos la misma función $f(\cdot)$ del ejemplo 11.5:

$$y = f(x_1, x_2) = \cos(x_1\pi/2) \sin(x_2\pi/2) \quad (11.53)$$

Cuyas funciones inversas son

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2) = \frac{2}{\pi} \cos^{-1} \left(\frac{y}{\sin(x_2\pi/2)} \right) \\ x_2 &= f_2^{-1}(y, x_1) = \frac{2}{\pi} \sin^{-1} \left(\frac{y}{\cos(x_1\pi/2)} \right) \end{aligned} \quad (11.54)$$

En este ejemplo se busca obtener la extensión inversa posible en cada una de las dos variables:

$$\begin{cases} B = T(0.7, 0.8, 0.9, 1.0) \\ A_1^{pos} = ? \\ A_2 = T(0.2, 0.3, 0.4, 0.5) \end{cases} \quad \begin{cases} B = T(0.7, 0.8, 0.9, 1.0) \\ A_1 = T(0.0, 0.1, 0.2, 0.3) \\ A_2^{pos} = ? \end{cases} \quad (11.55)$$

Antes de proceder a la aplicación del algoritmo es importante destacar que las funciones inversas no están definidas para todas las combinaciones de sus argumentos. Consideremos $f_1^{-1}(\cdot)$: Hay dos condiciones para las cuales no está definida:

- Para $\sin(x_2) = 0$, es decir, para $x_2 = 0$. Esta situación no sucederá, porque $f(x_1, x_2) : (0, 1]^2 \rightarrow (0, 1]$, es decir, $x_2 = 0$ está fuera del universo de discurso.
- Para $\frac{y}{\sin(x_2)} > 1$, es decir para $y > \sin(x_2)$

Podemos ajustar la función para asegurar que esté definida en toda combinación de valores así:

$$\begin{aligned} x_1 = f_1^{-1}(y, x_2) &= \begin{cases} \cos^{-1}\left(\frac{y}{\sin(x_2)}\right) & \text{si } y \leq \sin(x_2) \\ 0 & \text{si } y > \sin(x_2) \end{cases} \\ x_2 = f_2^{-1}(y, x_1) &= \begin{cases} \sin^{-1}\left(\frac{y}{\cos(x_2)}\right) & \text{si } y \leq \cos(x_2) \\ 1 & \text{si } y > \cos(x_2) \end{cases} \end{aligned} \quad (11.56)$$

La Figura 11.8 muestra los resultados de la aplicación del algoritmo de extensión inversa posible a cada uno de los dos casos. Nótese que en ambos casos los ajustes efectuados sobre las funciones inversas aseguran que el resultado existe en el universo de discurso.

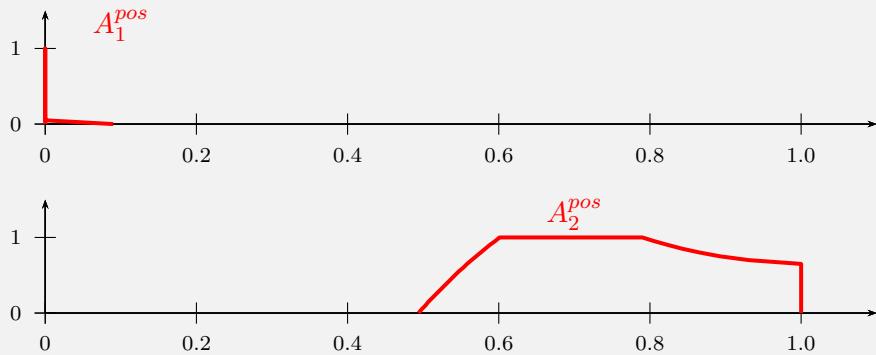


Figura 11.8: Solución del ejemplo 11.9

11.4. Extensión inversa necesaria

Supongamos ahora que queremos aplicar una de las funciones inversas 11.2:

$$x_k = f_k^{-1}(y, x_1, x_2, \dots, x_{k-1}, x_{k+1}, \dots, x_n) \quad (11.57)$$

En las siguientes condiciones: la función $f(\cdot)$ se está utilizando para modelar un cierto fenómeno que relaciona las variables de entrada x con la variable de salida y ; además, se tiene información sobre el valor de todas las variables de entrada x , salvo del de x_k y se necesita asegurar que y cumpla unas ciertas condiciones. La información disponible y la condición deseada para y contienen incertidumbre y por ello se ha optado por representarlas mediante números difusos.

En esas condiciones, se desea determinar que valores *debe* tener x_k que sean compatibles con la información disponible y que se satisfaga la condición impuesta para y .

Ejemplo 11.10 (Extensión inversa necesaria en un intervalo) Un breve ejemplo haciendo uso de la aritmética intervalar puede ayudar a aclarar el problema que se ha formulado. Supongamos que $y = x_1 - x_2$ y que se sabe que $x_2 \in [1, 2]$ y se necesita que $y \in [1, 3]$. Lo que se desea es encontrar todos los valores de x_1 tales que sea imposible encontrar una terna $\{x_1, x_2, y\}$ que cumpla simultáneamente las tres condiciones:

$$\begin{cases} y &= x_1 - x_2 \\ x_2 &\in [1, 2] \\ y &\notin [1, 3] \end{cases} \quad (11.58)$$

Por ejemplo, el número $x_1 = 5$ no puede ser uno de los valores buscados, porque $x_1 = 5, x_2 = 1$ hace que $y = 4 \notin [1, 3]$ y por tanto sí se pueden cumplir las tres condiciones.

En cambio, el número $x_1 = 3$ sí es uno de los valores buscados, porque

$$3 - [1, 2] = [3, 3] - [1, 2] = [1, 2] \subset [1, 3] \quad (11.59)$$

De hecho, para todos los valores de x_1 que caen en el intervalo $[3, 4]$ se satisface que $x_1 - [1, 2] \subset [1, 3]$, ya que $[3, 4] - [1, 2] = [1, 3]$ y por ello podemos decir que

$$x_1^{nec} = [3, 4] \quad (11.60)$$

Ejemplo 11.11 (Extensión inversa necesaria en un intervalo sin solución)

Modifiquemos ahora el ejemplo 11.10 para mostrar que en ocasiones la extensión necesaria

puede no tener solución. Consideramos la misma función $y = x_1 - x_2$ y que se sabe que $x_2 \in [1, 2]$, pero ahora se necesita que $y = 2$, es decir, $y \in [2, 2]$. Lo que se desea es encontrar todos los valores de x_1 tales que sea imposible encontrar una terna $\{x_1, x_2, y\}$ que cumpla simultáneamente las tres condiciones:

$$\begin{cases} y &= x_1 - x_2 \\ x_2 &\in [1, 2] \\ y &= 2 \end{cases} \quad (11.61)$$

Consideremos el valor de $x_1 = 3.5$. Es posible encontrar un valor de $x_2 \in [1, 2]$ tal que $2 = x_1 - x_2$. Dicho valor es $x_2 = 1.5$. Sin embargo, *también es posible encontrar otros valores de $x_2 \in [1, 2]$ tales que $2 \neq x_1 - x_2$* , por ejemplo $x_2 = 1$, y por esa razón $x_1 = 3.5$ no puede ser uno de los valores buscados.

De hecho, no existe ningún valor de x_1 con el que podamos asegurar que

$$[2, 2] = [x_1, x_1] - [1, 2] \quad (11.62)$$

Ya que eso implicaría que

$$[2, 2] = [x_1 - 2, x_1 - 1] \quad \begin{cases} x_1 - 2 &= 2 \\ x_1 - 1 &= 2 \end{cases} \quad (11.63)$$

Lo cual llevaría a la contradicción $3 = 4$. Por lo anterior podemos afirmar que

$$x_1^{nec} = \emptyset \quad (11.64)$$

Al comparar los ejemplos 11.6 y 11.10 se observa que la extensión necesaria es más restrictiva que la posible. Además, el ejemplo 11.11 muestra que puede suceder que la extensión necesaria no tenga solución.

Formalmente, la extensión inversa necesaria es la búsqueda de un número difuso A_k^{nec} tal que si construimos el arreglo

$$A^{nec} = \{A_1, A_2, \dots, A_{k-1}, A_k^{nec}, A_{k+1}, \dots, A_n\} \quad (11.65)$$

Se obtiene

$$f(A^{nec}) = B \quad (11.66)$$

En donde B , $\{A_1, A_2, \dots, A_{k-1}, A_{k+1}, \dots, A_n\}$ son unos números difusos conocidos. Podemos aplicar 11.16 para analizar la ecuación 11.66, lo que resulta en:

$$\begin{aligned} B(\alpha) &= f(A^{nec})(\alpha) = [f(\mathcal{L}_{A^{nec}}(\alpha)), f(\mathcal{R}_{A^{nec}}(\alpha))] \\ \mathcal{L}_{A^{nec}}(\alpha) &= \mathcal{D}_{A^{nec}}(\alpha, d_f) \\ \mathcal{R}_{A^{nec}}(\alpha) &= \mathcal{D}_{A^{nec}}(\alpha, \neg d_f) \end{aligned} \quad (11.67)$$

De la ecuación 11.15 se tiene que

$$\begin{aligned} \mathcal{D}_{A^{nec}}(\alpha, d_f) &= \{D_{A_1}(\alpha, d_f(1)), \dots, D_{A_{k-1}}(\alpha, d_f(k-1)), \\ &\quad D_{A_k^{nec}}(\alpha, d_f(k)), \\ &\quad D_{A_{k+1}}(\alpha, d_f(k+1)), \dots, D_{A_n}(\alpha, d_f(n))\} \\ \mathcal{D}_{A^{nec}}(\alpha, \neg d_f) &= \{D_{A_1}(\alpha, \neg d_f(1)), \dots, D_{A_{k-1}}(\alpha, \neg d_f(k-1)), \\ &\quad D_{A_k^{nec}}(\alpha, \neg d_f(k)), \\ &\quad D_{A_{k+1}}(\alpha, \neg d_f(k+1)), \dots, D_{A_n}(\alpha, \neg d_f(n))\} \end{aligned} \quad (11.68)$$

Lo que nos permite despejar, usando las funciones inversas 11.2:

$$\begin{aligned} D_{A_k^{nec}}(\alpha, d_f(k)) &= f_k^{-1} \left(\begin{array}{l} D_B(\alpha, 0), D_{A_1}(\alpha, d_f(1)), \dots, \\ D_{A_{k-1}}(\alpha, d_f(k-1)), D_{A_{k+1}}(\alpha, d_f(k+1)), \\ \dots, D_{A_n}(\alpha, d_f(n)) \end{array} \right) \\ D_{A_k^{nec}}(\alpha, \neg d_f(k)) &= f_k^{-1} \left(\begin{array}{l} D_B(\alpha, 1), D_{A_1}(\alpha, \neg d_f(1)), \dots, \\ D_{A_{k-1}}(\alpha, \neg d_f(k-1)), D_{A_{k+1}}(\alpha, \neg d_f(k+1)), \\ \dots, D_{A_n}(\alpha, \neg d_f(n)) \end{array} \right) \end{aligned} \quad (11.69)$$

El contenido de la ecuación 11.69 puede reescribirse de forma más adecuada. Para ello, recordemos que:

$$\begin{aligned} D_{A_k^{nec}}(\alpha, 0) &= L_{A_k^{nec}}(\alpha) \\ D_{A_k^{nec}}(\alpha, 1) &= R_{A_k^{nec}}(\alpha) \end{aligned} \quad (11.70)$$

El resultado de la extensión inversa necesaria de $f_k^{-1}(y, Xk)$ será un número difuso x_k^{nec} cuyos α -cortes serán

$$x_k^{nec}(\alpha) = [L_{x_k^{nec}}(\alpha), R_{x_k^{nec}}(\alpha)] \quad (11.71)$$

Para obtener los límites de esos α -cortes se debe aplicar la función inversa en un conjunto adecuado de valores

$$x_k^{nec}(\alpha) = \left[f_k^{-1} \left(\mathcal{L}_Y^{nec}(\alpha), \mathcal{L}_{A_k}^{nec}(\alpha) \right), f_k^{-1} \left(\mathcal{R}_Y^{nec}(\alpha), \mathcal{R}_{A_k}^{nec}(\alpha) \right) \right] \quad (11.72)$$

En donde los valores de Y que hay que utilizar están determinados por

$$\begin{aligned} \mathcal{L}_Y^{nec}(\alpha) &= \mathcal{D}_Y(\alpha, d_f(k)) \\ \mathcal{R}_Y^{nec}(\alpha) &= \mathcal{D}_Y(\alpha, \neg d_f(k)) \end{aligned} \quad (11.73)$$

Y los de las variables X_k están determinados por

$$\begin{aligned} \mathcal{L}_{A_k}^{nec}(\alpha) &= \left\{ \mathcal{L}_{A_1}^{nec}(\alpha), \mathcal{L}_{A_2}^{nec}(\alpha), \dots, \mathcal{L}_{A_{k-1}}^{nec}(\alpha), \mathcal{L}_{A_{k+1}}^{nec}(\alpha) \dots, \mathcal{L}_{A_n}^{nec}(\alpha) \right\} \\ \mathcal{R}_{A_k}^{nec}(\alpha) &= \left\{ \mathcal{R}_{A_1}^{nec}(\alpha), \mathcal{R}_{A_2}^{nec}(\alpha), \dots, \mathcal{R}_{A_{k-1}}^{nec}(\alpha), \mathcal{R}_{A_{k+1}}^{nec}(\alpha) \dots, \mathcal{R}_{A_n}^{nec}(\alpha) \right\} \\ \mathcal{L}_{A_j}^{nec}(\alpha) &= \mathcal{D}(\alpha, d_f(k) \odot d_f(j)) \\ \mathcal{R}_{A_j}^{nec}(\alpha) &= \mathcal{D}(\alpha, \neg d_f(k) \odot d_f(j)) \end{aligned} \quad (11.74)$$

Las ecuaciones 11.71, 11.72, 11.73 y 11.74 nos permiten calcular $L_{x_k^{nec}}(\alpha)$ y $R_{x_k^{nec}}(\alpha)$ que son los límites del α -corte de x_k^{nec} . Sin embargo *estas expresiones no aseguran que $L_{x_k^{nec}}(\alpha) \leq R_{x_k^{nec}}(\alpha)$ y por tanto, no podemos asegurar que el resultado obtenido sea un intervalo..*

Recordemos que hemos supuesto que tanto $f(\cdot)$ como sus inversas están bien definidas en el universo de discurso sobre el que trabajamos. En esas condiciones ¿por qué razón puede darse la situación en la que la extensión inversa necesaria no tiene solución?

La respuesta está relacionada con la situación del ejemplo 11.11: en la extensión necesaria se impone una condición: la $y = f(X)$ no puede tener valores por incompatibles con B , es decir $f(X) \subset B$. Si esta condición es muy restrictiva, entonces el problema no tiene solución.

El párrafo anterior implica que si se relajan las condiciones impuestas sobre y , entonces sí se puede encontrar una solución a la extensión inversa necesaria. Justamente esa es la inspiración del algoritmo 10. El algoritmo recibe un valor deseado de y^{des} y devuelve, además de x_k^{nec} , otro valor y^{nec} para el que existe una solución. Si las condiciones impuestas originalmente permitían encontrar la solución, entonces $y^{nec} = y^{des}$.

Recordemos que la ecuación 7.10 establece las condiciones que deben cumplir las funciones $L_A(\alpha)$ y $R_A(\alpha)$ para que A sea un número difuso. Al utilizar una representación discreta con p α -cortes, estas condiciones se convierten en:

$$\begin{cases} L(\alpha_p) \leq R(\alpha_p) \\ L(\alpha_j) \leq L_{\alpha_{j+1}} & j = 1, 2, \dots, p-1 \\ R(\alpha_j) \geq R_{\alpha_{j+1}} & j = 1, 2, \dots, p-1 \end{cases} \quad (11.75)$$

El algoritmo propuesto tiene los siguientes pasos:

1. Se resuelve el problema para el α -corte superior, es decir, para $\alpha = 1$, de la siguiente forma:
 - a) Se calculan $L_{x_k^{nec}(\alpha_p)}$, $R_{x_k^{nec}(\alpha_p)}$ con las ecuaciones 11.71, 11.72, 11.73 y 11.74 usando $y^{nec}(\alpha_p)$.
 - b) Se verifica si el resultado es un intervalo válido. En caso de no serlo, se modifica $y^{nec}(\alpha_p)$ hasta que lo sea. La modificación consiste en *ensanchar* el α -corte de $y^{nec}(\alpha_p)$ para relajar las condiciones impuestas sobre y .
2. Se resuelve el problema sucesivamente para los demás α -cortes, disminuyendo paulatinamente el valor de $j = p-1, p-2, \dots, 1$, de la siguiente forma:
 - a) Se verifica que el α -corte de y^{nec} esté contenido en el α -corte inmediatamente superior, y en caso de no serlo se modifica:

$$L_{y_j^{nec}} \leftarrow \min(L_{y_{j+1}^{nec}}, L_{y_j^{nec}}) \quad R_{y_j^{nec}} \leftarrow \max(R_{y_{j+1}^{nec}}, R_{y_j^{nec}}) \quad (11.76)$$

- b) Se calcula el límite inferior de $x_k^{nec}(\alpha_j)$ así:

- 1) Se calculan $L_{x_k^{nec}(\alpha_j)}$ con las ecuaciones 11.71, 11.72, 11.73 y 11.74 usando $y^{nec}(\alpha_j)$
- 2) Se verifica si $L_{x_k^{nec}(\alpha_j)} \leq L_{x_k^{nec}(\alpha_{j+1})}$. En caso de no cumplir la condición, se busca el valor de y^* que sí lo cumple haciendo:

$$\begin{aligned} x_k^{nec}(\alpha_j) &= x_k^{nec}(\alpha_{j+1}) \\ A^{nec}(\alpha_j) &= \{A_1(\alpha_j), \dots, x_k^{nec}(\alpha_j), \dots, A_n^{nec}(\alpha_j)\} \\ Y^* &= f(A^{nec}(\alpha_j)) \\ y^* &= LY^* \end{aligned} \quad (11.77)$$

y se ensancha adecuadamente el α -corte de y^{nec} así:

$$\begin{cases} \text{si } d_f(k) = 0 \text{ entonces } L_{y^{nec}}(\alpha_j) \leftarrow y^* \\ \text{si } d_f(k) = 1 \text{ entonces } R_{y^{nec}}(\alpha_j) \leftarrow y^* \end{cases} \quad (11.78)$$

c) Se calcula el límite superior de $x_k^{nec}(\alpha_j)$ así:

1) Se calculan $R_{x_k^{nec}(\alpha_j)}$ con las ecuaciones 11.71, 11.72, 11.73 y 11.74 usando $y^{nec}(\alpha_j)$

2) Se verifica si $R_{x_k^{nec}(\alpha_j)} \geq R_{x_k^{nec}(\alpha_{j+1})}$. En caso de no cumplir la condición, se busca el valor de y^* que sí lo cumple haciendo:

$$\begin{aligned} x_k^{nec}(\alpha_j) &= x_k^{nec}(\alpha_{j+1}) \\ A^{nec}(\alpha_j) &= \{A_1(\alpha_j), \dots, x_k^{nec}(\alpha_j), \dots, A_n^{nec}(\alpha_j)\} \\ Y^* &= f(A^{nec}(\alpha_j)) \\ y^* &= R_{Y^*} \end{aligned} \quad (11.79)$$

y se ensancha adecuadamente el α -corte de y^{nec} así:

$$\begin{cases} \text{si } d_f(k) = 0 \text{ entonces } R_{y^{nec}}(\alpha_j) \leftarrow y^* \\ \text{si } d_f(k) = 1 \text{ entonces } L_{y^{nec}}(\alpha_j) \leftarrow y^* \end{cases} \quad (11.80)$$

Podemos ahora establecer el algoritmo 10, que resume el procedimiento para efectuar la extensión inversa posible de una función de múltiples variables, aplicada a números difusos.

Algoritmo 10 Extensión inversa necesaria

Entradas:

- f_k^{-1} una función (11.2)
- d_f un vector de booleanos (11.14)
- A_k un vector $n - 1$ de números difusos con igual Υ
- Y un número difuso con igual Υ que A_k

```

1: procedure INVERSANECESARIA( $f_k^{-1}, d_f, Y, A_k$ )
2:    $p \leftarrow \text{size}(\Upsilon)$                                  $\triangleright$  Número de  $\alpha$ -cortes
3:    $L_{A_k} \leftarrow \text{zeros}(p)$                            $\triangleright$  Vector vacío
4:    $R_{A_k} \leftarrow \text{zeros}(p)$                            $\triangleright$  Vector vacío
5:    $L_Y^{nec} \leftarrow L_Y$                                  $\triangleright$  Copiar el vector  $Y$ 
6:    $R_Y^{nec} \leftarrow R_Y$                                  $\triangleright$  Copiar el vector  $Y$ 
7:    $Flag \leftarrow 0$                                       $\triangleright$  Indicador de modificación de  $Y$ 
8:    $j \leftarrow p$                                           $\triangleright$  Paso 1.  $\alpha = 1$ 
9:    $\alpha \leftarrow \Upsilon[j]$ 
10:   $L_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{nec}(\alpha), \mathcal{L}_{A_k}^{nec}(\alpha))$      $\triangleright$  Ecuaciones 11.71, 11.72, 11.73 y 11.74
11:   $R_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{R}_Y^{nec}(\alpha), \mathcal{R}_{A_k}^{nec}(\alpha))$      $\triangleright$  Ecuaciones 11.71, 11.72, 11.73 y 11.74
12:  while  $L_{A_k}[j] > R_{A_k}[j]$  do
13:     $dY \leftarrow \epsilon + \eta(R_Y(j) - L_Y(j))$ 

```

```

14:    $L_{Y^{nec}}[j] \leftarrow L_{Y^{nec}}[j] - dY$ 
15:    $R_{Y^{nec}}[j] \leftarrow R_{Y^{nec}}[j] + dY$ 
16:    $L_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{nec}(\alpha), \mathcal{L}_{A_k}^{nec}(\alpha))$   $\triangleright$  Ecuaciones 11.71, 11.72, 11.73 y 11.74
17:    $R_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{R}_Y^{nec}(\alpha), \mathcal{R}_{A_k}^{nec}(\alpha))$   $\triangleright$  Ecuaciones 11.71, 11.72, 11.73 y 11.74
18:    $Flag \leftarrow 1$ 
19:   for all  $j = \{p-1, p-2, \dots, 1\}$  do  $\triangleright$  Paso 2. Ciclo para los demás  $\alpha$ -cortes
20:      $\alpha \leftarrow \Upsilon[j]$ 
21:      $L_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{nec}(\alpha), \mathcal{L}_{A_k}^{nec}(\alpha))$   $\triangleright$  Ecuaciones 11.71, 11.72, 11.73 y 11.74
22:     if  $L_{A_k}[j] > L_{A_k}[j+1]$  then
23:        $L_{A_k}[j] = L_{A_k}[j+1]$ 
24:        $y^* \leftarrow f(A^{nec}(\alpha_j))$   $\triangleright$  Ecuación 11.77
25:       if  $d_f(k) = 0$  then
26:          $L_{Y^{nec}}[j] \leftarrow y^*$ 
27:       else
28:          $R_{Y^{nec}}[j] \leftarrow y^*$ 
29:        $Flag \leftarrow 1$ 
30:      $R_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{nec}(\alpha), \mathcal{L}_{A_k}^{nec}(\alpha))$   $\triangleright$  Ecuaciones 11.71, 11.72, 11.73 y 11.74
31:     if  $R_{A_k}[j] < R_{A_k}[j+1]$  then
32:        $R_{A_k}[j] = R_{A_k}[j+1]$ 
33:        $y^* \leftarrow f(A^{nec}(\alpha_j))$   $\triangleright$  Ecuación 11.79
34:       if  $d_f(k) = 0$  then
35:          $R_{Y^{nec}}[j] \leftarrow y^*$ 
36:       else
37:          $L_{Y^{nec}}[j] \leftarrow y^*$ 
38:        $Flag \leftarrow 1$   $\triangleright$  Vectores con los límites de  $A_k$  y
39:     return  $L_{A_k}, R_{A_k}, L_{Y^{nec}}, R_{Y^{nec}}, Flag$   $\triangleright$  de  $Y^{nec} = f(A)$  para cada  $\alpha$ -corte.  

 $Flag$  indica si se modificó o no  $Y$ 

```

Ejemplo 11.12 (Extensión necesaria de un hiperplano) En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.3 y 11.7:

$$y = f(x_1, x_2, x_3) = 1 + x_1 - 2x_2 - x_3 \quad (11.81)$$

Cuyas funciones inversas son

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2, x_3) = y - 1 + 2x_2 + x_3 \\ x_2 &= f_2^{-1}(y, x_1, x_3) = (y - 1 - x_1 + x_3) / -2 \\ x_3 &= f_3^{-1}(y, x_1, x_2) = y - 1 - x_1 + 2x_2 \end{aligned} \quad (11.82)$$

En este ejemplo se busca obtener la extensión inversa posible en la segunda variable, $A_2^{nec} = f_2^{-1}(B, A_1, A_3)$, con los siguientes números difusos:

$$\begin{aligned} B &= T(-4, -3, -2, -1) \\ A_1 &= T(0, 1, 2, 3) \\ A_3 &= T(-3, -2, -1, 0) \end{aligned} \quad (11.83)$$

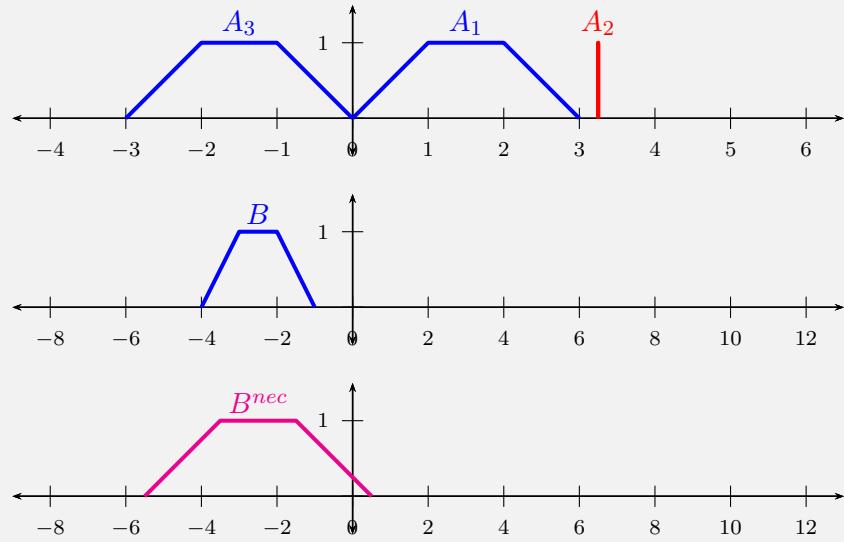


Figura 11.9: Números del ejemplo 11.12

La Figura 11.9 muestra los números difusos B, A_1, A_3 y el resultado A_2^{nec} de aplicar el algoritmo de extensión inversa necesaria. Recordemos que el sentido de la monotonía de $f(\cdot)$ determina que:

$$d_f(1) = 0 \quad d_f(2) = 1 \quad d_f(3) = 1 \quad (11.84)$$

Como se quiere aplicar la inversa a la segunda variable, entonces $k = 2$ y por tanto $d_f(k) = 1$.

Definamos D_X como una variable auxiliar que puede valer 0 ó 1. Valdrá 0 si se necesita el límite inferior para obtener A_2^{nec} en cada α -corte y 1 en caso contrario. Aplicando las ecuaciones 11.71, 11.72, 11.73 y 11.74 se obtiene

$$\begin{aligned} D_{A_1} &= d_f(k) \odot d_f(1) = 1 \odot 0 = 1 \\ D_{A_2} &= d_f(k) \odot d_f(3) = 1 \odot 1 = 0 \\ D_{A_3} &= d_f(k) \odot d_f(3) = 1 \odot 1 = 0 \\ D_B &= d_f(k) = 1 \end{aligned} \quad (11.85)$$

La Tabla 11.4 muestra los límites inferior (L) y superior (R) para cada uno de los números y cada uno de los α -cortes. Además, en esa tabla se han utilizado dos colores para explicar el funcionamiento del algoritmo, así: Para un mismo α -corte, los valores del mismo color están relacionados por la ecuación $y = 1 + x_1 - 2x_2 - x_3$. Eso implica que los colores de las columnas L y R de cada variable se pueden seleccionar usando lo calculado en la ecuación 11.85

Por ejemplo, para $\alpha = 0$ se tiene que

$$\begin{aligned} -5.501 &= 1 + 0.0 - 2 \times 3.251 - 0.0 \\ 0.506 &= 1 + 3.0 - 2 \times 3.247 - (-3.0) \end{aligned} \quad (11.86)$$

Tabla 11.4: Números del ejemplo 11.12

α	A_1		A_2		A_3		B		B^{nec}	
	L	R	L	R	L	R	L	R	L	R
0.000	0.000	3.000	3.247	3.251	-3.000	0.000	-4.000	-1.000	-5.501	0.506
0.250	0.250	2.750	3.247	3.251	-2.750	-0.250	-3.750	-1.250	-5.001	0.006
0.500	0.500	2.500	3.247	3.251	-2.500	-0.500	-3.500	-1.500	-4.501	0.494
0.750	0.750	2.250	3.247	3.251	-2.250	-0.750	-3.250	-1.750	-4.001	0.994
1.000	1.000	2.000	3.247	3.251	-2.000	-1.000	-3.000	-2.000	-3.501	1.494

Es importante destacar que el algoritmo ha modificado el número B y lo ha convertido en el número B^{nec} . Este nuevo número es más ancho, es decir, tiene más incertidumbre. Lo que ha sucedido es que el algoritmo ha detectado que con las condiciones del problema no era posible encontrar una solución; el algoritmo ha modificado el valor B , agregándole incertidumbre y en esas condiciones ha logrado encontrar un número A_2^{nec} que, a efectos prácticos, es un singleton, es decir, no tiene incertidumbre. Realmente el algoritmo ha encontrado un conjunto rectangular muy angosto, tan angosto que prácticamente es un singleton.

Ejemplo 11.13 (Extensión necesaria de una función positiva) En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.4 y 11.8:

$$y = f(x_1, x_2) = \frac{x_1^2}{x_2} \quad (11.87)$$

Cuyas funciones inversas son

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2) = \sqrt{yx_2} \\ x_2 &= f_2^{-1}(y, x_1) = \frac{x_1^2}{y} \end{aligned} \quad (11.88)$$

En este ejemplo se busca obtener la extensión inversa necesaria en cada una de las dos variables:

$$\begin{cases} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1^{nec} = ? \\ A_2 = T(0.5, 0.9, 1.1, 1.5) \end{cases} \quad \begin{cases} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1 = T(1.0, 1.8, 2.2, 3.0) \\ A_2^{nec} = ? \end{cases} \quad (11.89)$$

La Figura 11.10 muestra los resultados de la aplicación del algoritmo de extensión inversa necesaria a cada uno de los dos casos. En ambos casos fue necesario modificar la restricción impuesta a B para encontrar una solución, ensanchando sus α -cortes. Nótese que el ensanchamiento se hace de forma separada para cada uno de los límites; eso se observa claramente en B_1^{nec} , en donde se aprecia que las curvas L y R presentan unos quiebres a alturas diferentes.

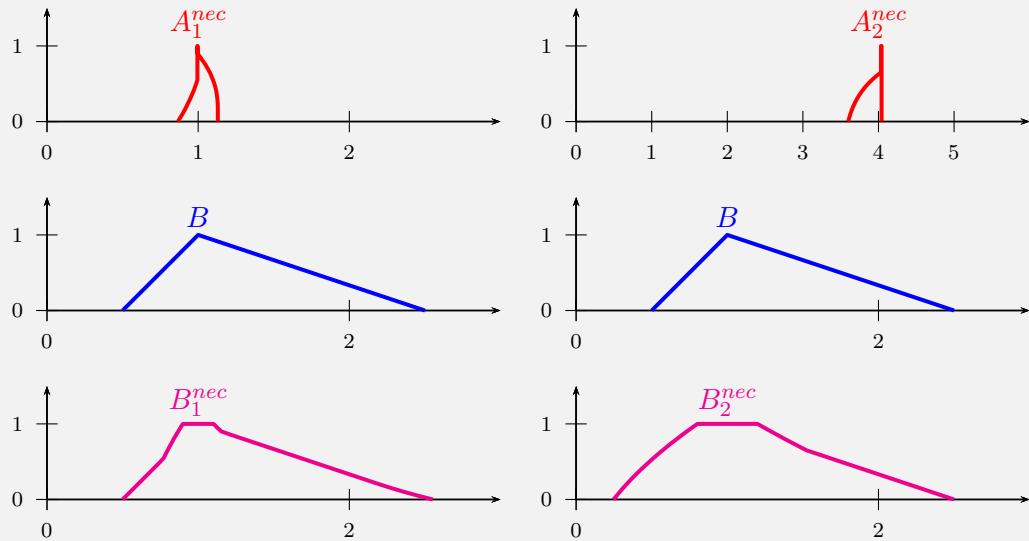


Figura 11.10: Solución del ejemplo 11.13

Ejemplo 11.14 (Extensión necesaria de una función acotada) En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.5 y 11.9:

$$y = f(x_1, x_2) = \cos(x_1\pi/2) \sin(x_2\pi/2) \quad (11.90)$$

Cuyas funciones inversas son, en principio

$$\begin{aligned} x_1 &= f_1^{-1}(y, x_2) = \frac{2}{\pi} \cos^{-1} \left(\frac{y}{\sin(x_2 \pi/2)} \right) \\ x_2 &= f_2^{-1}(y, x_1) = \frac{2}{\pi} \sin^{-1} \left(\frac{y}{\cos(x_1 \pi/2)} \right) \end{aligned} \quad (11.91)$$

Pero que se han ajustado a la ecuación 11.56, tal como se explica en el ejemplo 11.9. En este ejemplo se busca obtener la extensión inversa necesaria en cada una de las dos variables:

$$\begin{cases} B &= T(0.7, 0.8, 0.9, 1.0) \\ A_1^{nec} &= ? \\ A_2 &= T(0.2, 0.3, 0.4, 0.5) \end{cases} \quad \begin{cases} B &= T(0.7, 0.8, 0.9, 1.0) \\ A_1 &= T(0.0, 0.1, 0.2, 0.3) \\ A_2^{nec} &= ? \end{cases} \quad (11.92)$$

La Figura 11.11 muestra los resultados de la aplicación del algoritmo de extensión inversa necesaria a cada uno de los dos casos. En ninguno de los dos casos fue necesario modificar la restricción impuesta a B .

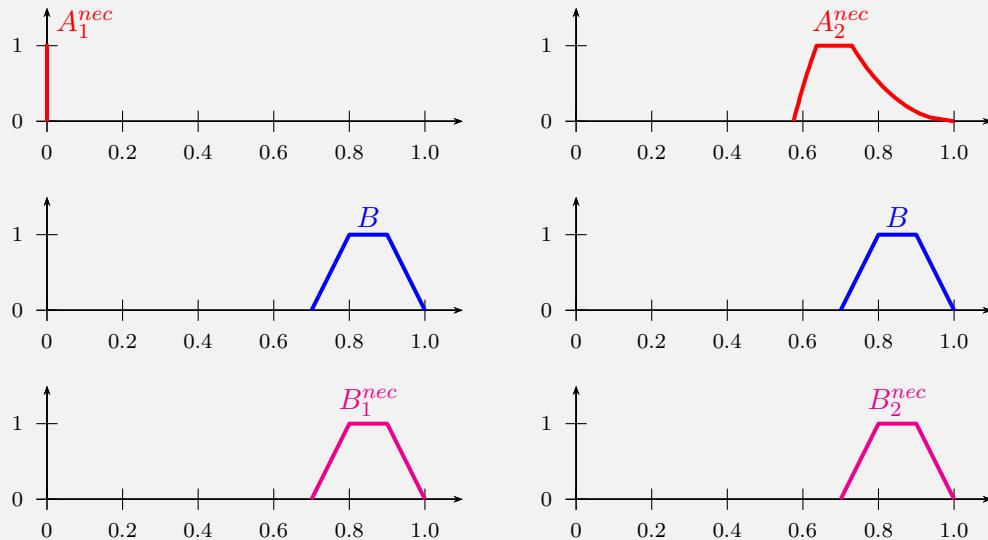


Figura 11.11: Solución del ejemplo 11.14

Tabla 11.5: Diferencia en el cálculo de las extensiones inversas posible y necesaria

Cálculo	Possible	Necesaria
$\mathcal{L}_{A_j}(\alpha)$	$\mathcal{D}(\alpha, \neg d_f(k) \odot d_f(j))$	$\mathcal{D}(\alpha, d_f(k) \odot d_f(j))$
$\mathcal{R}_{A_j}(\alpha)$	$\mathcal{D}(\alpha, d_f(k) \odot d_f(j))$	$\mathcal{D}(\alpha, \neg d_f(k) \odot d_f(j))$

11.5. Extensiones inversas intermedias

Al comparar los algoritmos de extensión inversa posible y necesaria se encuentra que la principal diferencia en los cálculos radica en las ecuaciones 11.43, y 11.74. Esta diferencia se ponen de manifiesto en la Tabla 11.5 y consiste en el papel de $d_f(k)$ en la selección de los límites de los α -cortes de los números A que hay que emplear a la hora de usar las funciones inversas 11.2.

Recordemos que $d_f(k)$ indica el sentido de la monotonía de $f(\cdot)$ en la variable k . Esta información hace que si en el cálculo de la extensión posible se usa el límite inferior de un cierto α -corte, en el cálculo de la extensión necesaria se use el límite superior; y viceversa.

Este hecho sugiere la definición de otro tipo de extensión inversa que no utilice ninguno de los valores límites de los intervalos, sino valores intermedios. Para ello, defnimos r , una variable auxiliar

$$r \in [0, 1] \quad (11.93)$$

r permite parametrizar esta familia de extensiones inversas intermedias. Si $r = 0$ la extensión será la posible, si $r = 1$ la extensión será la necesaria y al variar r desde 0 hasta 1 se obtendrán diferentes extensiones que varían suavemente.

Para obtener estas extensiones intermedias basta con sustituir el papel de las ecuaciones 11.71, 11.72, 11.73 y 11.74 por la ecuación 11.94 en el algoritmo de la extensión inversa necesaria.

$$\left\{
\begin{aligned}
x_k^{int}(\alpha, r) &= [L_{x_k^{int}}(\alpha, r), R_{x_k^{int}}(\alpha, r)] \\
x_k^{int}(\alpha, r) &= [f_k^{-1}(\mathcal{L}_Y^{int}(\alpha), \mathcal{L}_{A_k}^{int}(\alpha, r)), f_k^{-1}(\mathcal{R}_Y^{int}(\alpha), \mathcal{R}_{A_k}^{int}(\alpha, r))] \\
\mathcal{L}_y^{int}(\alpha) &= \mathcal{L}_y^{pos}(\alpha) = \mathcal{L}_y^{nec}(\alpha) \\
\mathcal{R}_y^{int}(\alpha) &= \mathcal{R}_y^{pos}(\alpha) = \mathcal{R}_y^{nec}(\alpha) \\
\mathcal{L}_{A_k}^{int}(\alpha, r) &= \{\mathcal{L}_{A_1}^{int}(\alpha, r), \mathcal{L}_{A_2}^{int}(\alpha, r), \dots, \mathcal{L}_{A_{k-1}}^{int}(\alpha, r), \mathcal{L}_{A_{k+1}}^{int}(\alpha, r) \dots, \mathcal{L}_{A_n}^{int}(\alpha, r)\} \\
\mathcal{R}_{A_k}^{int}(\alpha, r) &= \{\mathcal{R}_{A_1}^{int}(\alpha, r), \mathcal{R}_{A_2}^{int}(\alpha, r), \dots, \mathcal{R}_{A_{k-1}}^{int}(\alpha, r), \mathcal{R}_{A_{k+1}}^{int}(\alpha, r) \dots, \mathcal{R}_{A_n}^{int}(\alpha, r)\} \\
\mathcal{L}_{A_j}^{int}(\alpha, r) &= \mathcal{L}_{A_j}^{pos}(\alpha) + r(\mathcal{L}_{A_j}^{nec}(\alpha) - \mathcal{L}_{A_j}^{pos}(\alpha)) \\
\mathcal{R}_{A_j}^{int}(\alpha, r) &= \mathcal{R}_{A_j}^{pos}(\alpha) + r(\mathcal{R}_{A_j}^{nec}(\alpha) - \mathcal{R}_{A_j}^{pos}(\alpha))
\end{aligned}
\right. \tag{11.94}$$

Nótes el papel de la variable r en las últimas dos expresiones de 11.94: permite llevar una transición suave desde la extensión posible a la extensión necesaria.

Algoritmo 11 Extensión inversa intermedia

Entradas:

- f_k^{-1} una función (11.2)
- d_f un vector de booleanos (11.14)
- A_k un vector $n - 1$ de números difusos con igual Υ
- Y un número difuso con igual Υ que A_k
- r un número en el intervalo $[0, 1]$

```

1: procedure INVERSAINTERMEDIA( $f_k^{-1}, d_f, Y, A_k$ )
2:    $p \leftarrow \text{size}(\Upsilon)$                                  $\triangleright$  Número de  $\alpha$ -cortes
3:    $L_{A_k} \leftarrow \text{zeros}(p)$                            $\triangleright$  Vector vacío
4:    $R_{A_k} \leftarrow \text{zeros}(p)$                            $\triangleright$  Vector vacío
5:    $L_Y^{int} \leftarrow L_Y$                                   $\triangleright$  Copiar el vector  $Y$ 
6:    $R_Y^{int} \leftarrow R_Y$                                   $\triangleright$  Copiar el vector  $Y$ 
7:    $Flag \leftarrow 0$                                       $\triangleright$  Indicador de modificación de  $Y$ 
8:    $j \leftarrow p$                                           $\triangleright$  Paso 1.  $\alpha = 1$ 
9:    $\alpha \leftarrow \Upsilon[j]$ 

```

```

10:    $L_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{int}(\alpha), \mathcal{L}_{A_k}^{int}(\alpha))$                                  $\triangleright$  Ecuación 11.94
11:    $R_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{R}_Y^{int}(\alpha), \mathcal{R}_{A_k}^{int}(\alpha))$                                  $\triangleright$  Ecuación 11.94
12:   while  $L_{A_k}[j] > R_{A_k}[j]$  do
13:      $dY \leftarrow \epsilon + \eta(R_Y(j) - L_Y(j))$ 
14:      $L_{Y^{int}}[j] \leftarrow L_{Y^{int}}[j] - dY$ 
15:      $R_{Y^{int}}[j] \leftarrow R_{Y^{int}}[j] + dY$ 
16:      $L_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{int}(\alpha), \mathcal{L}_{A_k}^{int}(\alpha))$                                  $\triangleright$  Ecuación 11.94
17:      $R_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{R}_Y^{int}(\alpha), \mathcal{R}_{A_k}^{int}(\alpha))$                                  $\triangleright$  Ecuación 11.94
18:      $Flag \leftarrow 1$ 
19:   for all  $j = \{p-1, p-2, \dots, 1\}$  do       $\triangleright$  Paso 2. Ciclo para los demás  $\alpha$ -cortes
20:      $\alpha \leftarrow \Upsilon[j]$ 
21:      $L_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{int}(\alpha), \mathcal{L}_{A_k}^{int}(\alpha))$                                  $\triangleright$  Ecuación 11.94
22:     if  $L_{A_k}[j] > L_{A_k}[j+1]$  then
23:        $L_{A_k}[j] = L_{A_k}[j+1]$ 
24:        $y^* \leftarrow f(A^{int}(\alpha_j))$                                  $\triangleright$  Ecuación 11.77
25:       if  $d_f(k) = 0$  then
26:          $L_{Y^{int}}[j] \leftarrow y^*$ 
27:       else
28:          $R_{Y^{int}}[j] \leftarrow y^*$ 
29:        $Flag \leftarrow 1$ 
30:      $R_{A_k}[j] \leftarrow f_k^{-1}(\mathcal{L}_Y^{int}(\alpha), \mathcal{L}_{A_k}^{int}(\alpha))$                                  $\triangleright$  Ecuación 11.94
31:     if  $R_{A_k}[j] < R_{A_k}[j+1]$  then
32:        $R_{A_k}[j] = R_{A_k}[j+1]$ 
33:        $y^* \leftarrow f(A^{int}(\alpha_j))$                                  $\triangleright$  Ecuación 11.79
34:       if  $d_f(k) = 0$  then
35:          $R_{Y^{int}}[j] \leftarrow y^*$ 
36:       else
37:          $L_{Y^{int}}[j] \leftarrow y^*$ 
38:        $Flag \leftarrow 1$           Vectores con los límites de  $A_k$  y
39:     return  $L_{A_k}, R_{A_k}, L_{Y^{int}}, R_{Y^{int}}, Flag$            $\triangleright$  de  $Y^{int} = f(A)$  para cada  $\alpha$ -corte.  

                                          $Flag$  indica si se modificó o no  $Y$ 

```

Ejemplo 11.15 (Extensiones inversas intermedias de un hiperplano) En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.3, 11.7 y 11.12.

En este ejemplo se busca obtener la familia de extensiones inversas intermedias en la segunda variable, $A_2^{int}(r) = f_2^{-1}(B, A_1, A_3)$, con los siguientes números difusos:

$$\begin{aligned} B &= T(-4, -3, -2, -1) \\ A_1 &= T(0, 1, 2, 3) \\ A_3 &= T(-3, -2, -1, 0) \end{aligned} \quad (11.95)$$

La Figura 11.12 muestra una colección de las extensiones obtenidas. Se ha graficado tanto el resultado para $A_2^{int}(r)$ como el número B , que pudo haber sido modificado por el algoritmo.

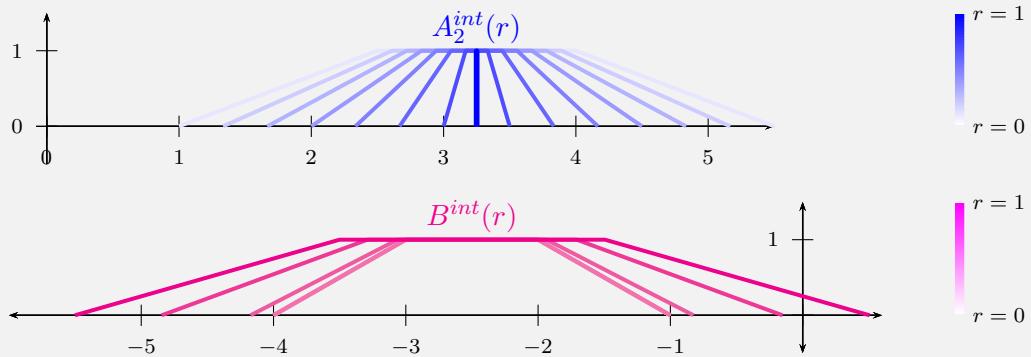


Figura 11.12: Números del ejemplo 11.15

Los resultados para $r = 0$ coinciden con los de la extensión inversa posible (Ejemplo 11.7) y los de $r = 1$ con los de la necesaria (Ejemplo 11.12).

Ejemplo 11.16 (Extensiones inversas intermedias de una función positiva)

En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.4, 11.8 y 11.13.

En este ejemplo se busca obtener la familia de extensiones inversas intermedias en cada una de las dos variables:

$$\left\{ \begin{array}{l} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1^{int}(r) = ? \\ A_2 = T(0.5, 0.9, 1.1, 1.5) \end{array} \right. \quad \left\{ \begin{array}{l} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1 = T(1.0, 1.8, 2.2, 3.0) \\ A_2^{int}(r) = ? \end{array} \right. \quad (11.96)$$

La Figura 11.13 muestra una colección de las extensiones obtenidas. Se ha graficado tanto el resultado para $A_1^{int}(r), A_2^{int}(r)$ como el número B , que pudo haber sido modificado por el algoritmo.

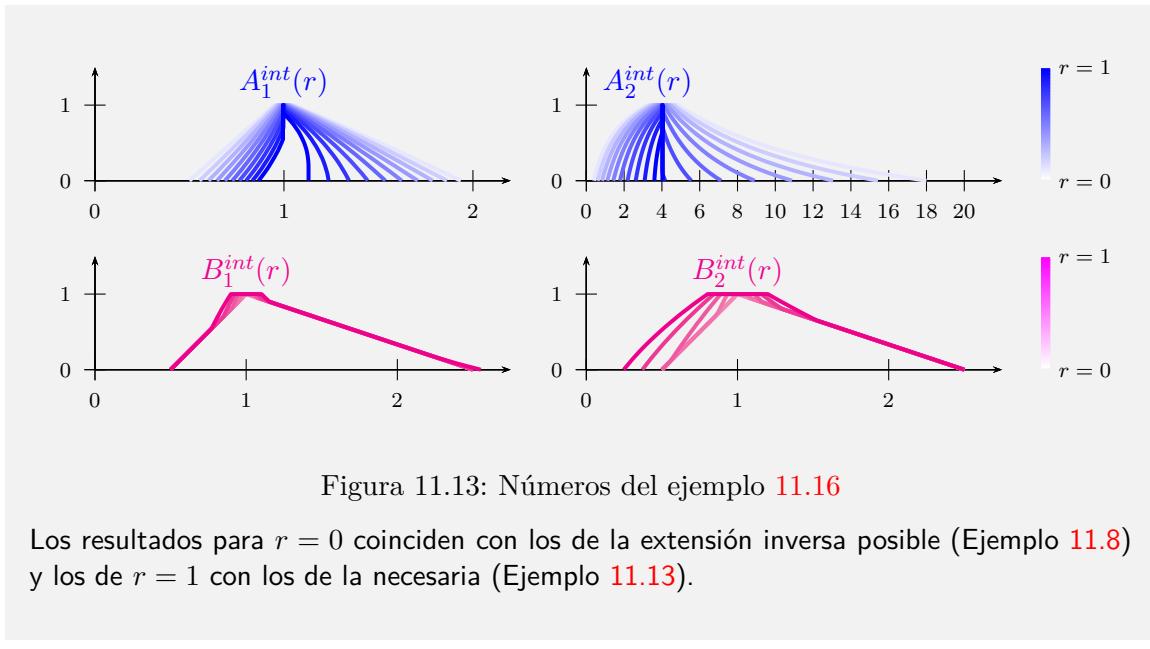


Figura 11.13: Números del ejemplo 11.16

Los resultados para $r = 0$ coinciden con los de la extensión inversa posible (Ejemplo 11.8) y los de $r = 1$ con los de la necesaria (Ejemplo 11.13).

Ejemplo 11.17 (Extensiones inversas intermedias de una función acotada)

En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.5, 11.9 y 11.14.

En este ejemplo se busca obtener la familia de extensiones inversas intermedias en cada una de las dos variables:

$$\left\{ \begin{array}{l} B = T(0.7, 0.8, 0.9, 1.0) \\ A_1^{\text{int}}(r) = ? \\ A_2 = T(0.2, 0.3, 0.4, 0.5) \end{array} \right. \quad \left\{ \begin{array}{l} B = T(0.7, 0.8, 0.9, 1.0) \\ A_1 = T(0.0, 0.1, 0.2, 0.3) \\ A_2^{\text{int}}(r) = ? \end{array} \right. \quad (11.97)$$

La Figura 11.14 muestra una colección de las extensiones obtenidas. Se ha graficado tanto el resultado para $A_1^{\text{int}}(r)$, $A_2^{\text{int}}(r)$ como el número B , que no fue modificado por el algoritmo para ningún valor de r .

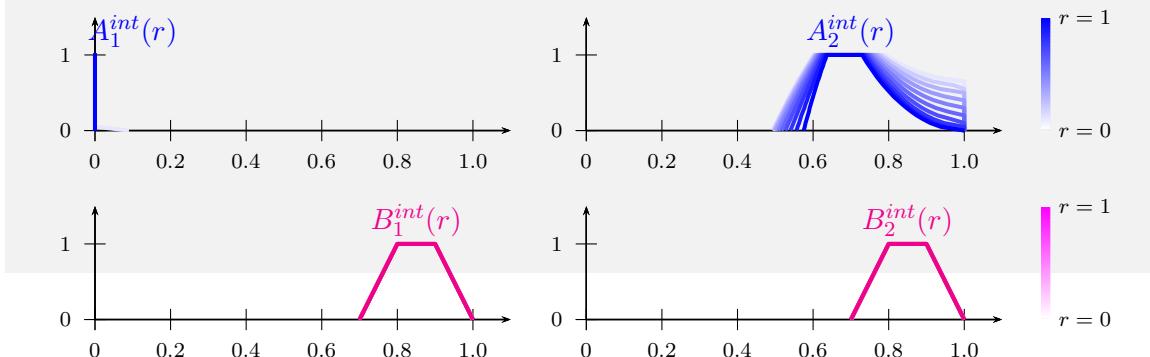


Figura 11.14: Números del ejemplo 11.17

Los resultados para $r = 0$ coinciden con los de la extensión inversa posible (Ejemplo 11.9) y los de $r = 1$ con los de la necesaria (Ejemplo 11.14).

11.6. Medida de la existencia de la extensión inversa de una función

La extensión inversa posible siempre existe, mientras que la necesaria puede no existir. Supongamos un caso en el que la necesaria no existe. Debido a que la familia de extensiones intermedias transita desde la posible hasta la necesaria al llevar r desde 0 hasta 1 debe haber un valor r_o a partir del cual la extensión intermedia deja de existir.

Lo anterior nos permite definir una *medida de la existencia de la extensión inversa de una función*, r_o de la siguiente forma:

Sean:

- Una función y sus funciones inversas tales como 11.1 y 11.2.
- B y $A_k = \{A_1, \dots, A_{k-1}, A_{k+1}, \dots, A_n\}$ unos números difusos.
- $x_k^{int}(r)$ el número difuso que se obtiene al aplicar el algoritmo de extensión inversas intermedia de $f_k^{-1}(B, A_k)$ con el parámetro r .
- $B^{nec}(r)$ el número difuso que representa a B y que es (posiblemente) una modificación de B efectuada por el algoritmo aplicado en el paso anterior.

La medida de la existencia de $f_k^{-1}(B, A_k)$ es el número r_o tal que

$$r_o = \max_r \mid B^{nec}(r) = B \mid \quad (11.98)$$

Es decir, es el máximo valor de r para el cual el algoritmo de extensión intermedia no modifica la salida B .

Ejemplo 11.18 (Existencia de la extensión inversa de un hiperplano) En este ejemplo consideraremos la misma función $f(\cdot)$ de los ejemplos 11.3, 11.7 y 11.12 y 11.15.

El propósito es el de determinar la medida de la existencia de la extensión inversa en la segunda variable, $A_2 = f_2^{-1}(B, A_1, A_3)$, con los siguientes números difusos:

Algoritmo 12 Medida de la existencia de la extensión inversa**Entradas:**

- f_k^{-1} una función (11.2)
- d_f un vector de booleanos (11.14)
- A_k un vector $n - 1$ de números difusos con igual Υ
- Y un número difuso con igual Υ que A_k

```

1: procedure MEDIDAEXISTENCIA( $f_k^{-1}, d_f, Y, A_k$ )
2:    $r_o \leftarrow 0$ 
3:    $r \leftarrow 0$ 
4:    $Flag \leftarrow False$ 
5:    $dr \leftarrow 0.01$                                  $\triangleright$  Incremento
6:   while  $r \leq 1.0$  do                       $\triangleright$  Ciclo para distintos valores de  $r$ 
7:     Obtener la extensión intermedia  $A_k^{int}(r)$            $\triangleright$  Algoritmo 11
8:      $Flag \leftarrow$  bandera de la ejecución del algoritmo
9:     if  $Flag = True$  then
10:        $r_o = r$ 
11:        $r = 1.5$ 
12:     else
13:        $r \leftarrow r + dr$ 
14:   return  $r_o$ 

```

$$\begin{aligned}
B &= T(-4, -3, -2, -1) \\
A_1 &= T(0, 1, 2, 3) \\
A_3 &= T(-3, -2, -1, 0)
\end{aligned} \tag{11.99}$$

En el ejemplo 11.15 se muestra que para obtener algunas de las extensiones intermedias el algoritmo modifica el número B , por esta razón sabemos que $r < 1.0$. Al aplicar el algoritmo 12 se obtiene $r_0 = 0.75$.

Ejemplo 11.19 (Existencia de la extensión inversa de una función positiva)
En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.4, 11.8 y 11.13 y 11.16.

El propósito es el de determinar la medida de la existencia de la extensión inversa en cada una de las dos variables, con los siguientes números difusos:

$$\left\{ \begin{array}{l} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1^{int}(r) = ? \\ A_2 = T(0.5, 0.9, 1.1, 1.5) \end{array} \right. \quad \left\{ \begin{array}{l} B = T(0.5, 1.0, 1.0, 2.5) \\ A_1 = T(1.0, 1.8, 2.2, 3.0) \\ A_2^{int}(r) = ? \end{array} \right. \quad (11.100)$$

En el ejemplo 11.16 se muestra que para obtener algunas de las extensiones intermedias el algoritmo modifica el número B , por esta razón sabemos que $r < 1.0$. Al aplicar el algoritmo 12 se obtiene, para ambas variables, $r_0 = 0.5$.

Ejemplo 11.20 (Existencia de la extensión inversa de una función acotada)

En este ejemplo consideramos la misma función $f(\cdot)$ de los ejemplos 11.5, 11.9 y 11.14 y 11.17.

El propósito es el de determinar la medida de la existencia de la extensión inversa en cada una de las dos variables, con los siguientes números difusos:

$$\left\{ \begin{array}{l} B = T(0.7, 0.8, 0.9, 1.0) \\ A_1^{int}(r) = ? \\ A_2 = T(0.2, 0.3, 0.4, 0.5) \end{array} \right. \quad \left\{ \begin{array}{l} B = T(0.7, 0.8, 0.9, 1.0) \\ A_1 = T(0.0, 0.1, 0.2, 0.3) \\ A_2^{int}(r) = ? \end{array} \right. \quad (11.101)$$

En el ejemplo 11.17 se muestra que para obtener las de las extensiones intermedias el algoritmo no tuvo que modificar el número B , por esta razón sabemos que $r = 1.0$. Al aplicar el algoritmo 12 se obtiene, efectivamente, para ambas variables, $r_0 = 1.0$.

Ejemplo 11.21 (Tasa interna de retorno difusa) Las matemáticas financieras proveen algunas herramientas básicas para el análisis del flujo esperado de dinero de un determinado proyecto. Una de ellas es el Valor Presente Neto (VPN) de ese flujo, que es la suma de cada ingreso y egreso de dinero traído al presente. Suponiendo que esos flujos suceden en los instantes de tiempo $k = 0, 1, \dots, n$, denotándolos por d_0, d_1, \dots, d_n y asumiendo una tasa constante r , el VPN se calcula así:

$$\text{VPN}(d_0, d_1, \dots, d_n, r) = d_0 + \frac{d_1}{(1+r)} + \frac{d_2}{(1+r)^2} + \dots + \frac{d_n}{(1+r)^n} \quad (11.102)$$

$$\text{VPN}(d_0, d_1, \dots, d_n, r) = \sum_{k=0}^n \frac{d_k}{(1+r)^k}$$

Otra de las herramientas de análisis es la Tasa Interna de Retorno (TIR), que se define como el valor de r que hace que el VPN sea cero:

$$\text{TIR} = r \left| \sum_{k=0}^n \frac{d_k}{(1+r)^k} = 0 \right. \quad (11.103)$$

Para aquellos casos en que d_0 es negativo y los demás flujos son positivos, la ecuación TIR está bien definida (tiene una única solución). Esta situación es usual en proyecto que requieren una única inversión inicial, representada por d_0 y que generan ingresos en los demás tiempos.

En este ejemplo analizamos un proyecto estudiado en [10, 11]. El proyecto tiene un flujo de dinero representado por los cinco números difusos que se muestran en la Figura 11.15 y se analiza a través de su TIR.

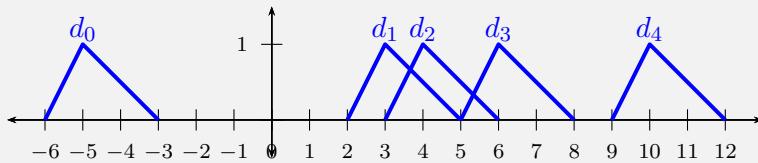


Figura 11.15: Flujo de dinero en el ejemplo 11.21

El enfoque con el que analizaremos el problema es diferente al original. Lo que haremos será considerar que la TIR es una de las funciones inversas del VPN y aplicar los algoritmos de extensión de funciones inversas.

Nótese que es aunque sí posible despejar cada uno de los flujos de dinero d_k en la ecuación 11.102 no es posible hacerlo con r . Sin embargo, sí podemos hallar numéricamente su valor aplicando, por ejemplo, el método de Newton-Raphson.

Si buscamos la extensión inversa posible de la TIR se obtiene el número que se muestra en la Figura 11.16. Este resultado coincide con el de los autores en [10, 11], porque allí se asocia la distribución de posibilidad de la TIR a la distribución de la posibilidad de que $\text{VPN} = 0$.

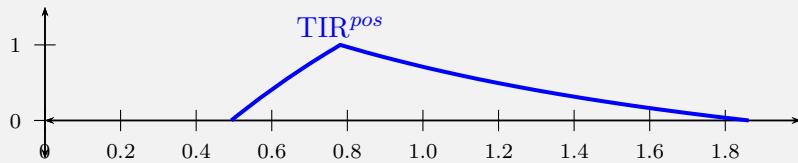


Figura 11.16: TIR posible en el ejemplo 11.21

Sin embargo, otro enfoque es posible y a juicio de este autor, más acertado: dado que la TIR es el valor de r que satisface $\text{VPN} = 0$, para encontrar la TIR asociada a un flujo de dinero debe usarse la extensión necesaria, usando como salida deseada $\text{VPN} = T(0, 0, 0, 0)$.

El resultado de este enfoque se muestra en la figura 11.17. El algoritmo también detecta que debido a la incertidumbre de los flujos de dinero es imposible obtener $\text{VPN} = T(0, 0, 0, 0)$ y encuentra el VPN que se muestra en la Figura 11.18.

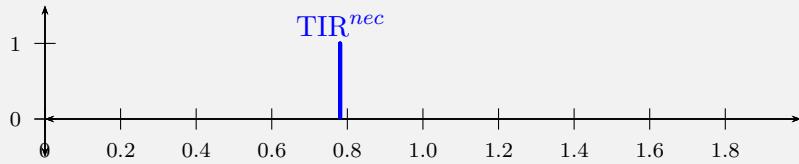


Figura 11.17: TIR necesaria en el ejemplo 11.21

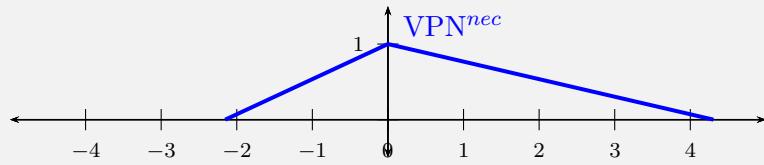


Figura 11.18: VPN necesario en el ejemplo 11.21

Capítulo 12

Sistemas basados en aritmética

Los sistemas basados en aritmética difusa, SBA, nacen como una respuesta al problema de la explosión de la base de reglas, propios de los sistemas basados en reglas, SBR (Sección 4.6.3). Este problema impide aprovechar el potencial lingüístico de la lógica difusa en sistemas con unas pocas variables y etiquetas lingüísticas (Véase [23, 14]).

Al igual que los SBR, los SBA también utilizan Variables Lingüísticas, lo que les permite combinar información cualitativa y cuantitativa. Sin embargo, no utilizan lógica difusa para calcular las salidas, sino aritmética difusa y, por tanto, no requieren de una base de reglas. Por ello los SBA no adolecen del problema de la explosión de la base de reglas y es posible utilizarlos en casos en los que haya un número elevado de variables y etiquetas.

En este capítulo se explica la estructura de los SBA y se ilustra su potencial con algunos ejemplos. Para el diseño y la implementación de SBA y de redes de SBA se ha desarrollado **FuzzyNet** (sección C.2), una herramienta de software libre que ha servido de apoyo para la preparación de los ejemplos de este capítulo y los de los capítulos 11 y 13 .

12.1. Estructura

Los SBA son sistemas tipo MISO (Figura 4.2). Su estructura interna se muestra en la Figura 12.1. Esta estructura recuerda a la de los SBR (Figura 4.3), porque los bloques correspondientes cumplen funciones análogas. Sin embargo, cada uno de ellos es radicalmente diferente al de su contraparte SBR. Las entradas se han marcado como **E**, en lugar de **x**, con el propósito de destacar que pueden ser de varios tipos, y no únicamente números reales como en los SBR. Por su parte, la salida se ha marcado como **s**, en lugar de **y**, para destacar que es una sola salida, pero que puede ser de diferentes tipos y no

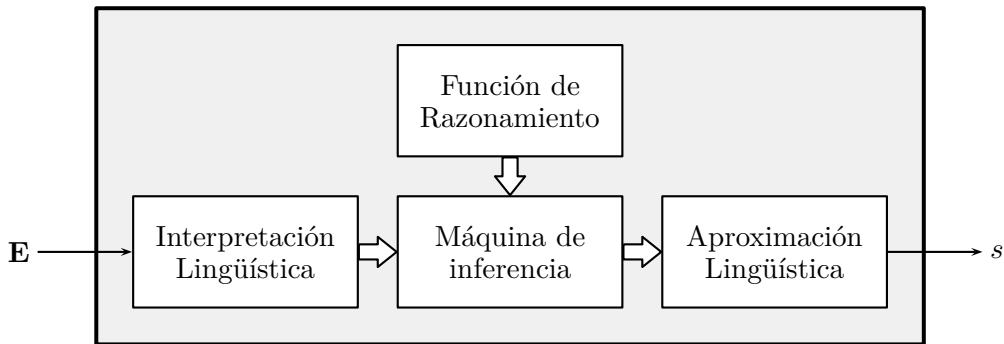


Figura 12.1: Estructura interna de un SBA

únicamente números. En particular, tanto entradas como salidas pueden ser palabras, lo que significa que permiten calcular palabras a partir de palabras; como el cálculo se realiza con aritmética difusa, decimos que los SBA son *sistemas de computación con palabras basados en aritmética difusa* (*Arithmetic Based Computing With Words Systems*, ABCWS).

Un SBA consta de cuatro bloques que se presentan en detalle en las secciones 12.2, 12.3, 12.4 y 12.5. El papel que cumplen cada uno de esos bloques se ilustra en la Figura 12.2 y se explica a continuación:

Función de razonamiento aproximado : la función de razonamiento aproximado, $fra(\cdot)$ es una función $\mathbb{R}^p \rightarrow \mathbb{R}$, en donde p es el número de entradas. Esta función controla la relación entre las entradas y la salida del sistema. En otras palabras, la fra captura el conocimiento sobre el comportamiento del sistema.

Razonamiento aproximado : la $fra(\cdot)$ se aplica sobre números difusos y produce un número difuso. La forma de hacerlo es mediante la aplicación del principio de extensión. Más específicamente, mediante el algoritmo de extensión directa presentado en la sección 11.2. Debido a que este algoritmo solo puede usarse con funciones monótonas, una $fra(\cdot)$ debe ser, precisamente, monótona.

Interpretación lingüística : para aplicar el algoritmo de extensión directa se requieren números difusos. El bloque de interpretación lingüística recibe las entradas del sistema y genera un número difuso por cada una de ellas. Es sobre estos números que se aplicará la $fra(\cdot)$. Este bloque acepta entradas de varios tipos, como por ejemplo palabras. Para operar, el bloque necesita que cada entrada esté asociada a una variable lingüística.

Aproximación lingüística : el resultado de aplicar la $fra(\cdot)$ es un número difuso.

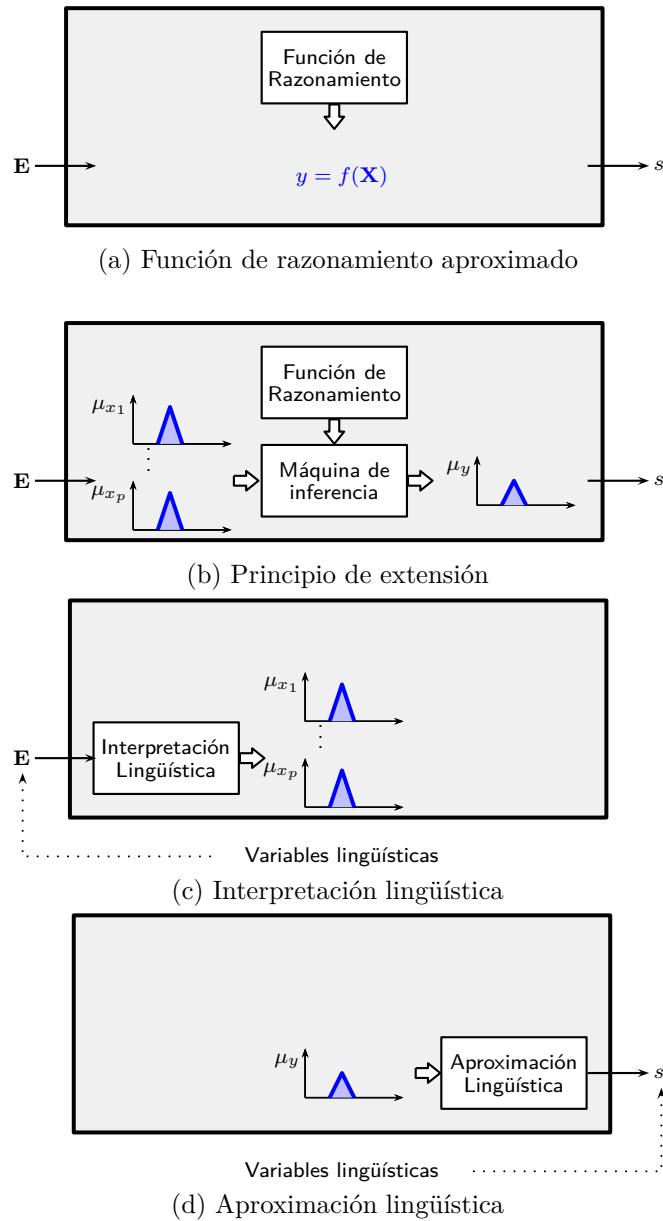


Figura 12.2: Bloques de un Sistema Basado en Aritmética

El bloque de aproximación lingüística transforma este resultado en la salida del sistema. Hay distintos tipos de salidas que se pueden obtener, como por ejemplo palabras. Para operar, el bloque necesita que la salida esté asociada a una variable lingüística.

De acuerdo a lo anterior, para construir un SBA es necesario realizar las siguientes tareas:

- Diseñar una Variable Lingüística para cada una de las entradas.
- Diseñar la Variable Lingüística para la salida.
- Construir la función de razonamiento aproximado.

Ejemplo 12.1 (Análisis de riesgo de daño causado por rayos) En este ejemplo consideremos el problema de estimar el riesgo del daño causado por las descargas eléctricas atmosféricas (los rayos) que puedan impactar una determinada instalación. Este problema ha sido estudiado a fondo en [30, 29]. Lo que aquí presentamos es una versión extremadamente simplificada que tiene propósitos meramente ilustrativos.

Supongamos que hemos consultado a un experto en el tema, y que él ha identificado que el riesgo depende de cuatro variables:

- I : la *Intensidad* de la descarga, que es el valor máximo de la corriente del rayo. Es un valor que se mide en kiloamperios (kA).
- D : la *Densidad* de descargas a tierra, que es el número de rayos por unidad de área en un periodo de un año. Es un valor que se mide en $1/km^2$.
- P : el nivel de *Protección* del equipo y las personas, que depende del sistema y equipos de protección. Se evalúa en una escala artificial sin unidades.
- M : la *Importancia* del equipo protegido, que depende de la función que cumplen y la dificultad para reponerlo en caso de daño. También se evalúa en una escala artificial sin unidades.

Con esa información, se desea construir un SBA como el de la Figura 12.3 que ayude a calcular R , el *Riesgo* del daño causado por el impacto de una descarga eléctrica atmosférica.

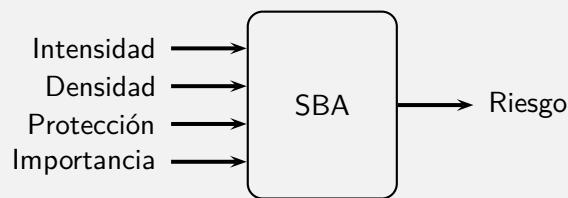


Figura 12.3: Sistema MIMO del ejemplo 12.1

A partir del conocimiento del experto se han construido las 5 variables lingüísticas que requiere el SBA, 4 para las entradas y 1 para la salida. Estas pueden verse en la Figura 12.4.

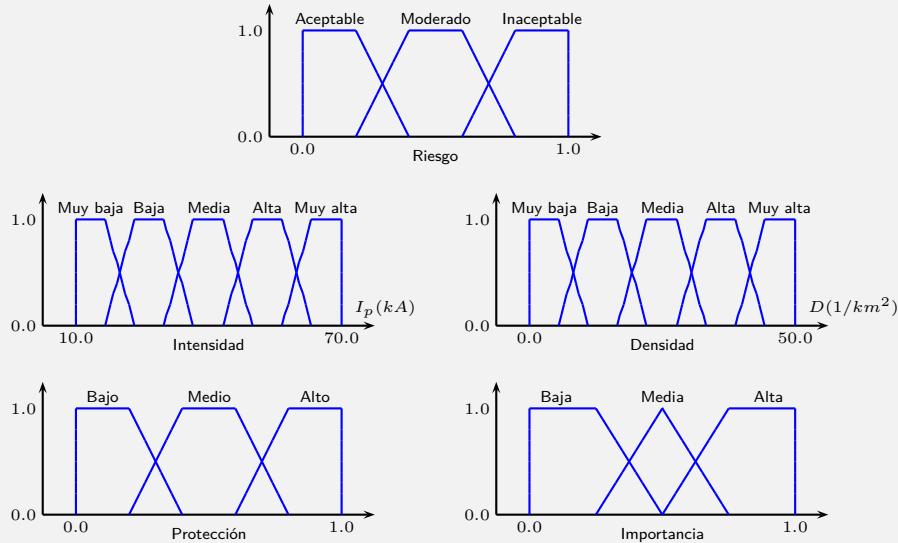


Figura 12.4: Variables lingüísticas del ejemplo 12.1

Para construir la función de razonamiento aproximado se han empleado las técnicas que se presentan en la sección 12.3. El conocimiento del experto se condensa en una función matemática simple:

$$\text{fra}(I, D, P, M) = 0.4I + 0.2D + 0.2(1 - P) + 0.2M \quad (12.1)$$

Una vez construido el SBA, se ha utilizado para calcular el riesgo de una instalación específica. La información disponible estaba en forma de etiquetas lingüísticas. La Figura 12.5 muestra la valoración de cada variable de entrada, así como la valoración del riesgo, que resulta ser 'moderado'.

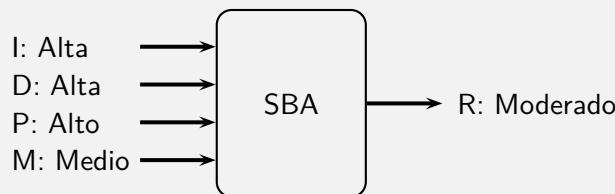


Figura 12.5: Respuesta del sistema del ejemplo 12.1

Para destacar, nótese que un SBR equivalente requeriría de $5 \times 5 \times 3 \times 3 = 225$ reglas. Semejante cantidad de reglas ha sido remplazada por una única función aritmética en este

SBA.

12.2. Máquina de inferencia

Un SBA como el que se muestra en la Figura 12.1 tiene p entradas E_1, E_2, \dots, E_p y una salida s . Cada una de esas variables tiene asociada una variable lingüística (ecuación 3.1):

$$\begin{aligned} \mathcal{V}_{E_1} & : \langle \mathcal{X}_1, \mathcal{L}\mathcal{X}_1, \mathcal{U}_1, \mathcal{M}_{\mathcal{U}_1} \rangle \\ \mathcal{V}_{E_2} & : \langle \mathcal{X}_2, \mathcal{L}\mathcal{X}_2, \mathcal{U}_2, \mathcal{M}_{\mathcal{U}_2} \rangle \\ & \dots \\ \mathcal{V}_{E_p} & : \langle \mathcal{X}_p, \mathcal{L}\mathcal{X}_p, \mathcal{U}_p, \mathcal{M}_{\mathcal{U}_p} \rangle \\ \mathcal{V}_s & : \langle \mathcal{Y}, \mathcal{L}\mathcal{Y}, \mathcal{W}, \mathcal{M}_{\mathcal{W}} \rangle \end{aligned} \tag{12.2}$$

En donde \mathcal{U}_k es el universo de discurso de la entrada k , y \mathcal{W} el de la salida s . Denotemos por \mathcal{F}_T el conjunto de todos los conjuntos difusos que pueden definirse sobre el universo de discurso T . En esas condiciones, una máquina de inferencia, bien sea de un SBR o de un SBA, puede verse como una función:

$$\begin{aligned} MI & : \mathcal{F}_U \rightarrow \mathcal{F}_W \\ \mathcal{F}_U & = \mathcal{F}_{U_1} \times \mathcal{F}_{U_1} \cdots \mathcal{F}_{U_p} \end{aligned} \tag{12.3}$$

Sin embargo, en el contexto de los SBR o de los SBA no cualquier conjunto difuso es una entrada válida para una máquina de inferencia. Tanto el difusor de un SBR como el bloque de interpretación lingüística de un SBA producen números difusos, que son las entradas para la máquina de inferencia.

Denotemos por \mathcal{N}_T el conjunto de todos los números difusos que pueden definirse sobre el universo de discurso T que, para ello, es un intervalo cerrado en \mathbb{R} . Claramente $\mathcal{N}_T \subset \mathcal{F}_T$. En esas condiciones, una máquina de inferencia, bien sea de un SBR o de un SBA, puede verse como una función:

$$\begin{aligned} MI & : \mathcal{N}_U \rightarrow \mathcal{F}_W \\ \mathcal{N}_U & = \mathcal{N}_{U_1} \times \mathcal{N}_{U_1} \cdots \mathcal{N}_{U_p} \end{aligned} \tag{12.4}$$

La máquina de inferencia de un SBR produce un resultado que no es, en general, un número difuso, sino que es la agregación (usualmente el máximo) de los conjuntos difusos convexos resultantes de cada una de las reglas. En contraposición, la máquina

de inferencia de un SBA sí genera un número difuso. Por esta razón, puede describirse como una función *FRA* que opera sobre números difusos y produce números difusos:

$$FRA(\cdot) : \mathcal{N}_U \rightarrow \mathcal{N}_W \quad (12.5)$$

Más específicamente, *FRA* es la extensión a números difusos de una función *fra* definida para los reales:

$$fra(\cdot) : \mathbb{R}^p \rightarrow \mathbb{R} \quad (12.6)$$

fra(·) recibe el nombre de *función de razonamiento aproximado* y se explora en detalle en la sección 12.3. Para aplicar *fra*(·) a números difusos se utiliza, por supuesto, el principio de extensión (capítulo 9).

12.3. Función de razonamiento aproximado

El elemento diferenciador de los SBA, respecto a los SBR, es el uso de una función matemática para establecer el funcionamiento del sistema, en lugar de una base de reglas. La motivación para hacerlo es la siguiente: en muchas aplicaciones prácticas, el conocimiento que se tiene sobre cómo debería funcionar el sistema es simple y puede, por tanto, representarse por funciones matemáticas simples.

A manera de ejemplo, supóngase un sistema de dos entradas en las que cada una de ellas afecta positivamente la salida. Es decir, que cuando una de esas entradas crece la salida también lo hace. Supóngase también que se sabe que el efecto de cada una de las dos entradas es igual de importante sobre la salida. Ese comportamiento puede modelarse por un promedio simple sobre una escala adecuada; algo de la forma:

$$y = \frac{x_1 + x_2}{2} \quad (12.7)$$

No son pocas las aplicaciones en las que sucede algo similar. En esta sección presentamos una estrategia para modelar con funciones sencillas, situaciones sencillas. Ahora bien, al combinar varios sistemas de este tipo se pueden modelar situaciones más complejas, tal como se muestra en la sección 13.2 pero, por lo pronto, consideraremos un único sistema.

Supongamos un SBA con p entradas. La función de razonamiento aproximado *fra*(·) de ese sistema será, en general, una función $fra(\cdot) : \mathbb{R}^p \rightarrow \mathbb{R}$. Sin embargo, para efectuar una implementación práctica del principio de extensión es preferible imponer algunas condiciones:

- fra debe ser una función monótona, de tal forma que se pueda emplear el algoritmo de extensión directa (sección 11.2).
- Debido a que las operaciones de producto y división de números difusos no son monótonas en todos los cuadrantes, es necesario restringir los valores de las variables a números positivos.
- Para facilitar la interpretación del significado de diferentes $\text{fra}(\cdot)$ es preferible que sean comparables entre sí.

Todas estas condiciones se consiguen fácilmente si las $\text{fra}(\cdot)$ satisfacen las siguientes condiciones:

$$\begin{cases} \text{fra}(\cdot) : [0, 1]^p \rightarrow [0, 1] \\ \text{fra}(\cdot) \text{ es monótona en todas sus variables} \end{cases} \quad (12.8)$$

En adelante, si no se establece algo diferente daremos por sentado que una $\text{fra}(\cdot)$ satisface esas condiciones. Además, usaremos x_i, y para referirnos a las variables de entrada y salida, de tal forma que podremos escribir

$$y = \text{fra}(x_1, x_2, \dots, x_p) \quad (12.9)$$

12.3.1. Galería de fra

En esta sección presentamos una colección de $\text{fra}(\cdot)$ que satisfacen la condición de la ecuación 12.8. Estas funciones hacen uso de unas variables auxiliares, s_i, w_i definidas de la siguiente forma:

$$s_i = \begin{cases} 1 & \text{si } \text{fra}(\cdot) \text{ es monótonamente creciente en } x_i \\ 0 & \text{si } \text{fra}(\cdot) \text{ es monótonamente decreciente en } x_i \end{cases} \quad (12.10)$$

$$0 \leq w_i \leq 1 \quad \sum_{i=1}^p w_i = 1 \quad (12.11)$$

s_i es una variable discreta que ayuda a discernir cuál es el sentido de la monotonía de la variable x_i . Por su parte, w_i es una variable real que sirve para ponderar la importancia de la variable x_i en el cálculo de y . Podemos también calcular la suma de los pesos de las variables crecientes decrecientes, w^+ y w^- , respectivamente :

$$w^+ = \sum_{i=1}^p s_i w_i \quad w^- = \sum_{i=1}^p (1 - s_i) w_i \quad (12.12)$$

Las funciones que se presentan en esta sección están diseñadas para que sea fácil diseñar una $\text{fra}(\cdot)$ cuando se parte del siguiente conocimiento:

- Se sabe que la salida varía monótonamente con las entradas, y se conoce el sentido de la monotonía. Gracias a este conocimiento se puede establecer el valor de cada s_i .
- Se sabe cuáles variables inciden en mayor medida sobre la salida. Gracias a este conocimiento se pueden asignar los pesos relativos w_i .

La colección está compuesta por 4 opciones principales, cada una de las cuales tiene 3 variantes. Las 12 funciones resultantes se han graficado en las Figuras 12.7, 12.8 y 12.9 con diferentes ajustes, para facilitar su comparación.

12.3.1.1. Opciones

Opción 1, variante 1. Combinación lineal. La opción más elemental es una combinación lineal de las entradas. Para asegurar que se respete la monotonía de cada variable la combinación se hace sobre x_i en aquellas variables crecientes y sobre $(1 - x_i)$ en las decrecientes:

$$y = \text{fra}_1(x_1, \dots, x_p) = \sum_{i=1}^p s_i w_i x_i + \sum_{i=1}^p (1 - s_i) w_i (1 - x_i) \quad (12.13)$$

o, lo que es igual:

$$y = \text{fra}_1(x_1, \dots, x_p) = w^- + \sum_{i=1}^p (-1)^{s_i+1} w_i x_i \quad (12.14)$$

Estas son algunas de las propiedades¹ de $\text{fra}_1(\cdot)$:

$$\left\{ \begin{array}{lcl} \max_x \{\text{fra}_1(\cdot)\} & = & 1 \\ \min_x \{\text{fra}_1(\cdot)\} & = & 0 \\ \text{rango}\{\text{fra}_1(\cdot)\} & = & 1 \\ \text{fra}_1(0, 0, \dots, 0) & = & w^- \\ \text{fra}_1(1, 1, \dots, 1) & = & w^+ \\ \text{fra}_1(0.5, 0.5, \dots, 0.5) & = & 0.5 \end{array} \right. \quad (12.15)$$

$\text{fra}_1(\cdot)$ genera un hiperplano. En las figuras 12.7a, 12.8a y 12.9a se pueden observar unos ejemplos en \mathbb{R}^3 .

¹La demostración es directa al notar que $\max_x \{\text{fra}_1(\cdot)\} = \text{fra}_1(s_1, s_2, \dots, s_p)$, así como $\min_x \{\text{fra}_2(\cdot)\} = \text{fra}_1((1-s_1), (1-s_2), \dots, (1-s_p))$ y, por supuesto, que $\text{rango}\{\text{fra}_1(\cdot)\} = \max_x \{\text{fra}_1(\cdot)\} - \min_x \{\text{fra}_1(\cdot)\}$

Opción 2, variante 1. Combinación lineal centrada en cero. Esta segunda opción también es una combinación lineal de las entradas. Se diferencia de la opción 1 en que está diseñada para que la salida sea 0.5 cuando las entradas sean cero, es decir:

$$\text{fra}_2(0, 0, \dots, 0) = 0.5 \quad (12.16)$$

Las dos opciones son coincidentes cuando $w^+ = w^-$, porque $\text{fra}_1(0, 0, \dots, 0) = w^-$ y si $w^+ = w^-$ entonces $w^+ = w^- = 0.5$.

Para encontrar la función que satisface 12.16, la formulamos como una transformación afín de $\text{fra}_1(\cdot)$

$$\text{fra}_2(\cdot) = \phi + \kappa \times \text{fra}_1(\cdot) \quad (12.17)$$

$$\begin{aligned} \text{fra}_2(0, \dots, 0) &= \phi + \kappa \times \text{fra}_1(\cdot) \\ \text{fra}_2(0, \dots, 0) &= \phi + \kappa \times w^- \\ 0.5 &= \phi + \kappa \times w^- \\ \phi &= 0.5 - \kappa w^- \end{aligned} \quad (12.18)$$

Por otra parte, el valor de κ está condicionado por el hecho de que $\text{fra}_2(\cdot) \in [0, 1]$.

$$\begin{aligned} \min_x\{\text{fra}_2(\cdot)\} &= \phi + \kappa \min_x\{\text{fra}_1(\cdot)\} \geq 0 \\ \min_x\{\text{fra}_2(\cdot)\} &= \phi \geq 0 \\ \max_x\{\text{fra}_2(\cdot)\} &= \phi + \kappa \max_x\{\text{fra}_1(\cdot)\} \leq 1 \\ \max_x\{\text{fra}_2(\cdot)\} &= \phi + \kappa \leq 1 \end{aligned} \quad (12.19)$$

Combinando 12.18 y 12.19 se obtiene

$$\left\{ \begin{array}{l} \kappa \leq 1 - \phi \\ \kappa \leq 1 - (0.5 - \kappa w^-) \\ \kappa \leq 0.5 + \kappa w^- \\ \kappa(1 - w^-) \leq 0.5 \\ \kappa w^+ \leq 0.5 \\ \kappa \leq \frac{0.5}{w^+} \end{array} \right. \quad \left\{ \begin{array}{l} \phi \geq 0 \\ 0.5 - \kappa w^- \geq 0 \\ 0.5 \geq \kappa w^- \\ \frac{0.5}{w^-} \geq \kappa \end{array} \right. \quad (12.20)$$

Las dos condiciones para κ se obtienen haciendo

$$\kappa = \frac{0.5}{\max\{w^+, w^-\}} \quad (12.21)$$

En resumen, $\text{fra}_2(\cdot)$ se define así:

$$\left\{ \begin{array}{l} \kappa = \frac{0.5}{\max\{w^+, w^-\}} \\ \phi = 0.5 - \kappa w^- \\ y = \text{fra}_2(x_1, \dots, x_p) = \phi + \kappa \left[\sum_{i=1}^p s_i w_i x_i + \sum_{i=1}^p (1-s_i) w_i (1-x_i) \right] \end{array} \right. \quad (12.22)$$

o lo que es igual:

$$y = \text{fra}_2(x_1, \dots, x_p) = 0.5 + \kappa \sum_{i=1}^p (-1)^{s_i+1} w_i x_i \quad (12.23)$$

Estas son algunas de las propiedades de $\text{fra}_2(\cdot)$:

$$\left\{ \begin{array}{l} \max_x \{\text{fra}_2(\cdot)\} = 0.5 + \kappa w^+ \\ \min_x \{\text{fra}_2(\cdot)\} = 0.5 - \kappa w^- \\ \text{rango}\{\text{fra}_2(\cdot)\} = \kappa \\ \text{fra}_2(0, 0, \dots, 0) = 0.5 \\ \text{fra}_2(1, 1, \dots, 1) = 0.5 + \kappa(w^+ - w^-) \\ \text{fra}_2(0.5, 0.5, \dots, 0.5) = 0.5(1 + \kappa(w^+ - w^-)) \end{array} \right. \quad (12.24)$$

$\text{fra}_2(\cdot)$ genera un hiperplano. En las figuras 12.7b, 12.8b y 12.9b se pueden observar unos ejemplos en \mathbb{R}^3 . En las figuras 12.8 las opciones 1 y 2 generan el mismo plano, porque en ese ejemplo $w^+ = w^- = 0.5$. En las figuras 12.7 y 12.9 esta condición no se da y opr eso los dos planos son diferentes.

Opción 3, variante 1. Producto. Esta opción emplea el producto de las entradas para calcular la salida. Para asegurar que se respete la monotonía de cada variable el producto se hace sobre x_i en aquellas variables crecientes y sobre $(1-x_i)$ en las decrecientes. Es importante recordar que $x_i, w_i, s_i, (1-x_i), (1-s_i)$ son no negativas, por lo que la monotonía del producto no se ve afectada por los signos de los multiplicandos.

La función que resulta es:

$$y = \text{fra}_3(x_1, \dots, x_p) = \frac{\prod_{i=1}^p (s_i w_i x_i + (1-s_i) w_i (1-x_i))}{\prod_{i=1}^p w_i} \quad (12.25)$$

Y estas son algunas de sus propiedades:

$$\left\{ \begin{array}{l} \max_x \{\text{fra}_3(\cdot)\} = 1 \\ \min_x \{\text{fra}_3(\cdot)\} = 0 \\ \text{rango}\{\text{fra}_3(\cdot)\} = 1 \\ \\ \text{fra}_3(0, 0, \dots, 0) = \begin{cases} 1 & \text{si } s_1 = \dots = s_p = 0 \\ 0 & \text{otro caso} \end{cases} \\ \\ \text{fra}_3(1, 1, \dots, 1) = \begin{cases} 1 & \text{si } s_1 = \dots = s_p = 1 \\ 0 & \text{otro caso} \end{cases} \\ \\ \text{fra}_3(0.5, 0.5, \dots, 0.5) = 0.5^p \end{array} \right. \quad (12.26)$$

$\text{fra}_3(\cdot)$ genera una hipersuperficie. En las figuras 12.7c, 12.8c y 12.9c se pueden observar unos ejemplos en \mathbb{R}^3 .

Opción 4, variante 1. Función logística La función logística

$$y = \frac{1}{1 + e^{-x}} \quad (12.27)$$

Es una función sigmoide que tiene el punto de inflexión en 0 (Figura 12.6). Esta función puede generalizarse usando dos parámetros $\kappa > 0$ y $0 < \phi < 1$ (por ejemplo $\kappa = 10$, $\phi = 0.5$) así:

$$y = \frac{1}{1 + e^{-\kappa(x_i - \phi)}} \quad (12.28)$$

Como puede verse en la Figura 12.6, el efecto de κ es el de hacer más o menos abrupta la sigmoide, mientras que el de ϕ es el de modificar el punto de inflexión.

Esta opción está concebida como la aplicación de la ecuación 12.28 sobre la combinación lineal de las entradas. Para asegurar que se respete la monotonía de cada variable la combinación se hace sobre x_i en aquellas variables crecientes y sobre $(1 - x_i)$ en las decrecientes. El resultado se normaliza para asegurar que esté en el intervalo $[0, 1]$.

Obtenemos primero los valores mínimo y máximo de la función así;

$$\begin{aligned} Y_{max} &= \frac{1}{1 + e^{-\kappa(1-\phi)}} \\ Y_{min} &= \frac{1}{1 + e^{-\kappa(0-\phi)}} \end{aligned} \quad (12.29)$$

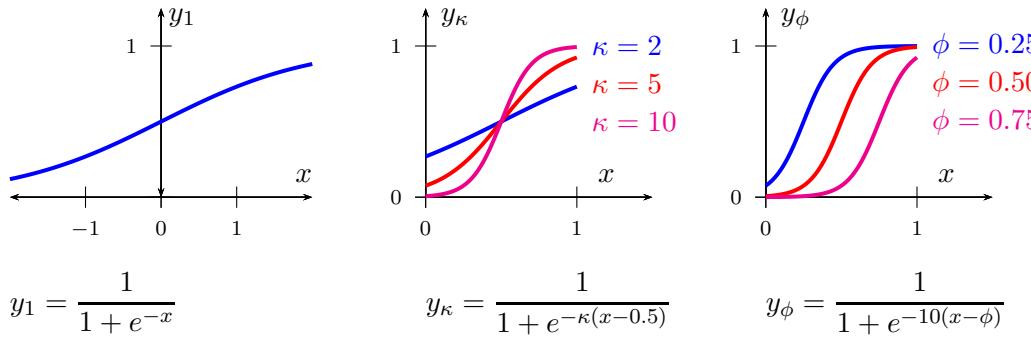


Figura 12.6: Función logística parametrizada

Y a partir de allí calculamos la función y la normalizamos:

$$\begin{aligned}
 X &= \sum_{i=1}^p s_i w_i x_i + \sum_{i=1}^p (1 - s_i) w_i (1 - x_i) \\
 Y &= \frac{1}{1 + e^{-A(X-b)}} \\
 y = \text{fra}_4(x_1, \dots, x_p) &= \frac{Y - Y_{min}}{Y_{max} - Y_{min}}
 \end{aligned} \tag{12.30}$$

Estas son algunas de sus propiedades de la función obtenida:

$$\left\{
 \begin{array}{lcl}
 \max_x \{\text{fra}_4(\cdot)\} &=& 1 \\
 \min_x \{\text{fra}_4(\cdot)\} &=& 0 \\
 \text{rango}\{\text{fra}_4(\cdot)\} &=& 1 \\
 \text{fra}_4(0, 0, \dots, 0) &=& \frac{w^- - Y_{min}}{Y_{max} - Y_{min}} \\
 \text{fra}_4(1, 1, \dots, 1) &=& \frac{w^+ - Y_{min}}{Y_{max} - Y_{min}} \\
 \text{fra}_4(0.5, 0.5, \dots, 0.5) &=& \frac{0.5 - Y_{min}}{Y_{max} - Y_{min}}
 \end{array}
 \right. \tag{12.31}$$

$\text{fra}_4(\cdot)$ genera una hipersuperficie. En las figuras 12.7d, 12.8d y 12.9d se pueden observar unos ejemplos en \mathbb{R}^3 .

La Tabla 12.1 permite comparar las propiedades de las 4 opciones presentadas. Está construida con los resultados de las ecuaciones 12.15, 12.24, 12.26 y 12.31. Se ha introducido una notación que simplifica la escritura (por cuestiones de espacio): el símbolo $\tilde{}$ significa que todas las variables tiene el mismo valor. Por ejemplo $\tilde{0}$ significa $x_1 = x_2 = \dots = x_p = 0$.

12.3.1.2. Variantes

A partir de estas 4 opciones se pueden desarrollar variantes². En esta sección presentamos dos de ellas que consisten en preprocesar cada x_i por una función $g_i(x_i) : [0, 1] \rightarrow [0, 1]$ monótonamente crecientes y tales que $g(0) = 0$ y $g(1) = 1$. Mediante ese preprocesamiento se logra modificar la forma de la función $\text{fra}(\cdot)$.

Variante 2. Monomios

$$\tilde{x}_i = g_2(x_i) = x_i^{r_i} \quad 0 \leq r_i \quad (12.32)$$

Consideremos dos situaciones diferentes en las que x_i toma los valores $3/4$ y $1/4$, respectivamente. Es decir, el primer valor de x_i es el 3 veces el segundo. Ahora bien, esta relación se convierte en 9 si $r_i = 2$ y en 1.73 si $r_i = 1/2$. Este ejemplo sirve páreja ilustrar que, en general, $r_i > 1$ sirve para subvalorar los valores bajos de x_i en relación de los valores altos, mientras que $r_i < 1$ sirve para sobrevalorarlos.

Variante 3. Función logística La función logística (ecuación 12.28) también sirve para preprocesar x_i .

$$\tilde{x}_i = g_3(x_i) = \frac{1}{1 + e^{-\kappa_g(x_i - \phi_g)}} \quad (12.33)$$

Esta variante tiene el efecto de concentrar el efecto de la variación de las entradas alrededor de el punto de inflexión, disminuyéndolo cerca de los extremos.

Estas variantes preservan las propiedades de la Tabla 12.1, salvo la correspondiente a $\text{fra}(\tilde{0.5})$. Eso significa que el papel de las variantes puede entenderse como una deformación en la manera en que la hipersuperficie conecta los puntos extremos.

En las Figuras 12.7, 12.8 y 12.9 se han dibujado las cuatro opciones con sus 3 variantes, en tres escenarios distintos. Para poder dibujarlas ha sido necesario tomar sólo dos variables de entrada ($p = 2$). Los tres escenarios seleccionados han sido:

- $s_1 = 1, s_2 = 1, w_1 = 0.5, w_2 = 0.5$, es decir, ambas entradas son crecientes y tienen igual peso.

²Las hemos denominado variantes 2 y 3, reservando el número 1 a la versión presentada en la sección 12.3.1.1

Tabla 12.1: Comparación de fra

	$\text{fra}_1(\cdot)$	$\text{fra}_2(\cdot)$	$\text{fra}_3(\cdot)$	$\text{fra}_4(\cdot)$
$\max\{\text{fra}(\cdot)\}$	1	$0.5 + \kappa w^+$	1	1
$\min\{\text{fra}(\cdot)\}$	0	$0.5 - \kappa w^-$	0	0
rango(fra)	1	κ	1	1
$\text{fra}(\tilde{0})$	w^-	0.5	$\begin{cases} 1 & \text{si } \tilde{s} = 0 \\ 0 & \text{otro caso} \end{cases}$	$\frac{w^- - Y_{min}}{Y_{max} - Y_{min}}$
$\text{fra}(\tilde{1})$	w^+	$0.5 + \kappa(w_+ - w^-)$	$\begin{cases} 1 & \text{si } \tilde{s} = 1 \\ 0 & \text{otro caso} \end{cases}$	$\frac{w^+ - Y_{min}}{Y_{max} - Y_{min}}$
$\text{fra}(0.5)$	0.5	$0.5 + \kappa(1 + w_+ - w^-)$	0.5^p	$\frac{0.5 - Y_{min}}{Y_{max} - Y_{min}}$

- $s_1 = 1, s_2 = 0, w_1 = 0.5, w_2 = 0.5$, es decir, ambas entradas tienen igual peso y tienen monotonías contrarias; la entrada 1 es creciente y la 2 es decreciente.
- $s_1 = 1, s_2 = 1, w_1 = 0.9, w_2 = 0.1$, es decir, ambas entradas son crecientes y tienen pesos muy diferentes, siendo mayor el de la primera entrada.

Al comparar el escenario 1 con el escenario 2 (Figuras 12.7 vs 12.8), se puede observar que el cambio de monotonía puede visualizarse como un simple cambio en el sentido de los ejes de las entradas, pero que en esencia preserva la forma de la función.

Por su parte, al comparar el escenario 1 con el escenario 3 (Figuras 12.7 vs 12.9), puede verse que el efecto de tener pesos diferentes en las entradas consiste en que la salida es mucho más sensible a las variaciones de las entradas con mayor peso. En efecto, las gráficas de la Figura 12.9 son mucho más horizontales en la dirección de la segunda entrada (x_2) que las de la Figura 12.7.

En las tres figuras también puede verse que las tres variantes comparten los mismos puntos extremos. Por ejemplo, al comparar las Figuras 12.7a, 12.7e y 12.7i se observa que las tres superficies guardan la misma monotonía y además toman los mismos valores en las aristas verticales del cubo. Lo que las diferencia es la forma en que conectan esos extremos. Al tomar otra terna de superficies que estén en la misma columna de las Figuras 12.7, 12.8 y 12.9 se llega a la misma conclusión.

Ejemplo 12.2 (fra(·) para el análisis de riesgo de daño causado por rayos)

Retomemos ahora el ejemplo 12.1 en el que se presenta un SBA que permite estimar riesgo de daño causado por rayos. Para construir una $\text{fra}(·)$ para ese sistema, comenzamos por asignar los siguientes significados a las variables x_i :

$$x_1 : I \quad x_2 : D \quad x_3 : P \quad x_4 : M \quad (12.34)$$

Analicemos ahora el sentido de la monotonía de las variables involucradas:

- I : la *Intensidad* de la descarga incrementa el riesgo, porque está asociada a la energía del rayo que es, en últimas, la que puede causar daños.
- D : la *Densidad* de descargas a tierra incrementa el riesgo, porque incrementa la posibilidad de que un rayo impacte la instalación.
- P : el nivel de *Protección* disminuye el riesgo de daño, porque un mayor nivel de protección logra aislar rayos más peligrosos de los equipos y las personas.
- M : la *Importancia* del equipo protegido incrementa el riesgo porque cuanto más importante sea el equipo dañado, mayor es su efecto sobre la operación de la instalación.

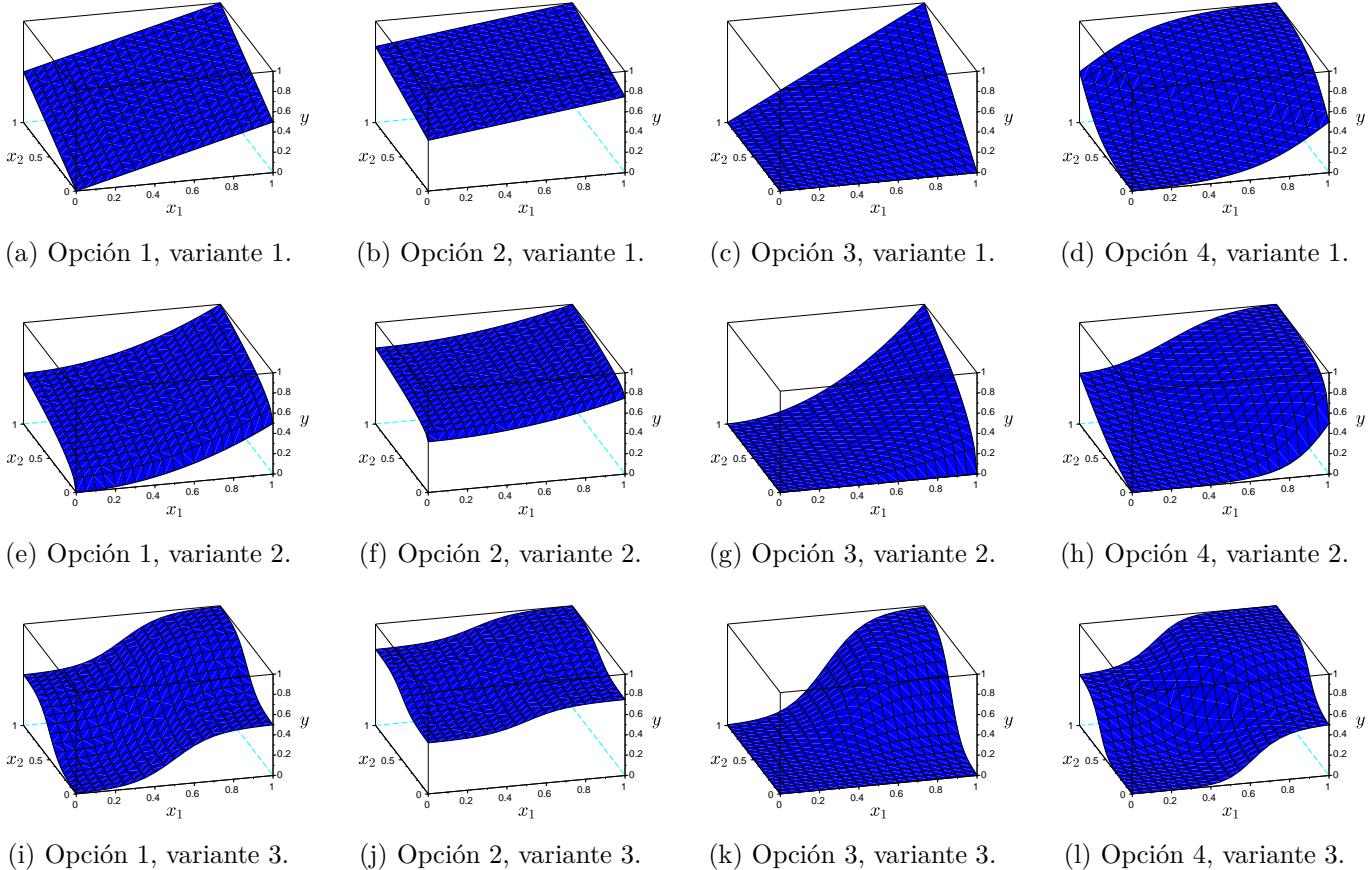


Figura 12.7: Galería de fra (1/3). $p = 2$, $s_1 = 1, s_2 = 1$, $w_1 = 0.5$, $w_2 = 0.5$. Para la opción 4 se ha usado $\kappa = 10$, $\phi = 0.5$. Para la variante 2 se ha usado $r_1 = 2.0$, $r_2 = 0.5$. Para la variante 3 se ha usado $\kappa_g = 10$, $\phi_g = 0.5$.

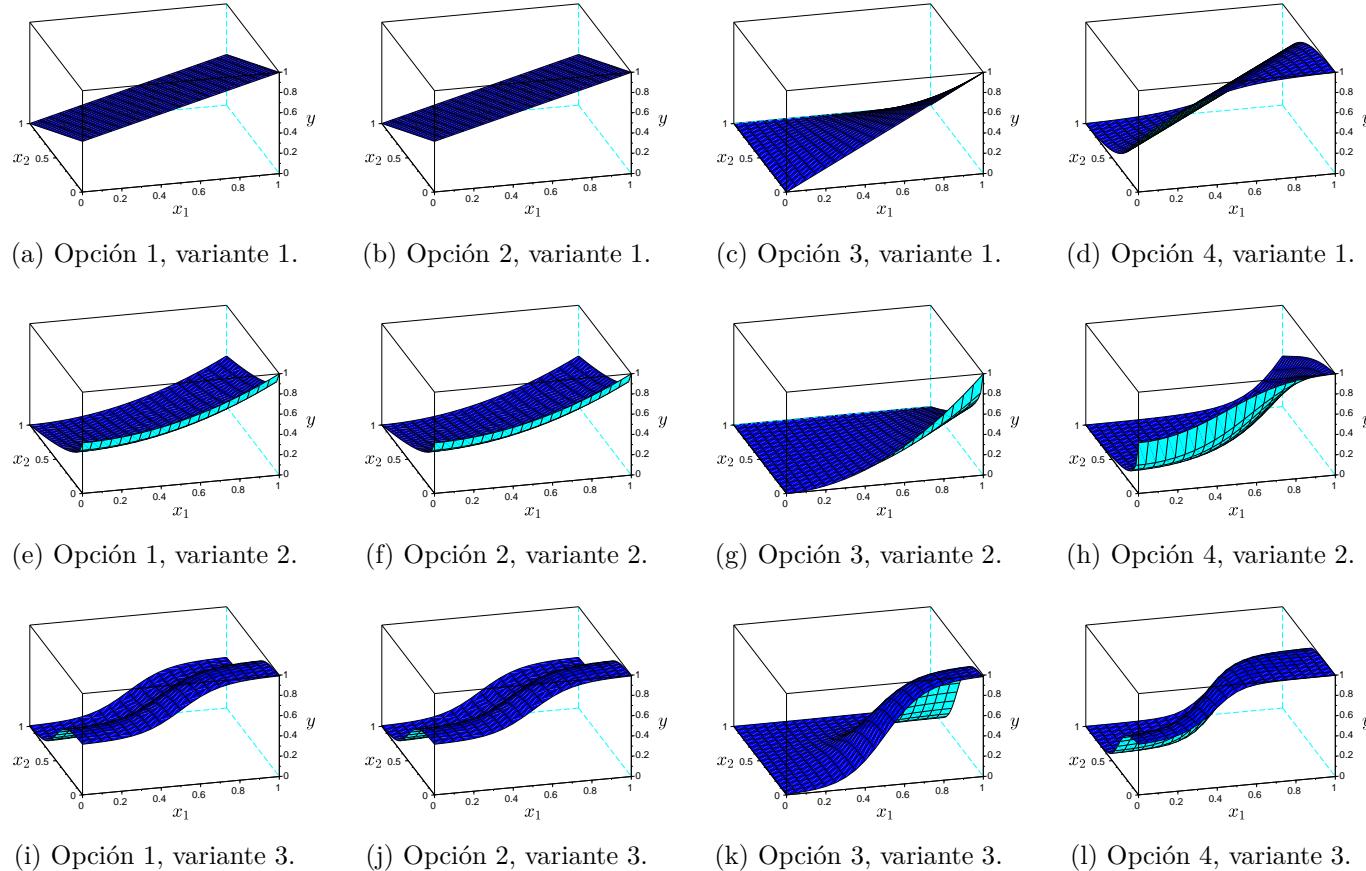


Figura 12.8: Galería de fra (2/3). $p = 2, s_1 = 1, s_2 = 0, w_1 = 0.5, w_2 = 0.5$. Para la opción 4 se ha usado $\kappa = 10$, $\phi = 0.5$. Para la variante 2 se ha usado $r_1 = 2.0, r_2 = 0.5$. Para la variante 3 se ha usado $\kappa_g = 10, \phi_g = 0.5$.

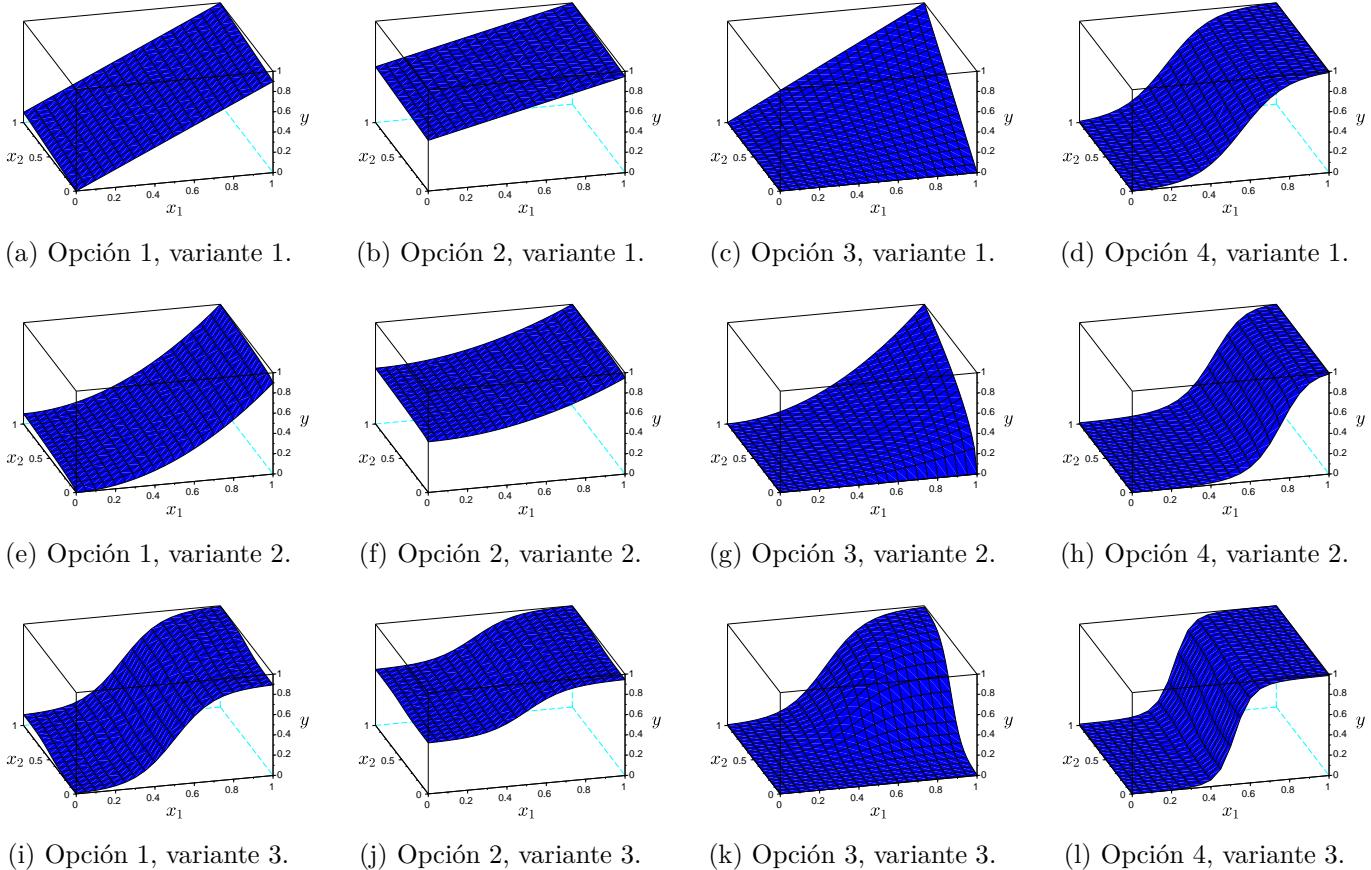


Figura 12.9: Galería de fra (3/3). $p = 2$, $s_1 = 1, s_2 = 1$, $w_1 = 0.9$, $w_2 = 0.1$. Para la opción 4 se ha usado $\kappa = 10$, $\phi = 0.5$. Para la variante 2 se ha usado $r_1 = 2.0$, $r_2 = 0.5$. Para la variante 3 se ha usado $\kappa_g = 10$, $\phi_g = 0.5$.

De lo anterior se deduce que

$$s_1 = 1 \quad s_2 = 1 \quad s_3 = 0 \quad s_4 = 1 \quad (12.35)$$

Supongamos ahora que los expertos en el tema han llegado a la conclusión que la variable I es la más importante de las cuatro en la estimación del riesgo, y que consideran que su importancia duplica a las de las otras tres. Lo anterior se traduce en

$$w_1 = 0.4 \quad w_2 = 0.2 \quad w_3 = 0.2 \quad w_4 = 0.2 \quad (12.36)$$

Si se emplea la opción 1 y la variante 1 para construir la $\text{fra}(\cdot)$ se obtiene

$$\text{fra}(x_1, x_2, x_3, x_4) = 0.4x_1 + 0.2x_2 + 0.2(1 - x_3) + 0.2x_4 \quad (12.37)$$

Que coincide con la ecuación 12.1 que se presentó en el ejemplo 12.1

12.3.2. Funciones *ad hoc*

Las opciones presentadas en la sección 12.3.1.1 son eso: opciones. No es obligatorio que la $\text{fra}(\cdot)$ sea una de ellas. De igual forma, tampoco es obligatorio el uso del algoritmo de extensión directa y por tanto las condiciones que se establecen en la ecuación 12.8 tampoco son obligatorias. Lo que sí es obligatorio es que la máquina de inferencia sea capaz de calcular 12.5.

Ahora bien, en algunas aplicaciones se cuenta con un conocimiento que relaciona las entradas con la salida, y este conocimiento está expresado en forma matemática como una función

$$y = \Phi(x_1, x_2, \dots, x_p) \quad (12.38)$$

y se quiere emplear un SBA porque se necesita aplicar ese conocimiento a variables que se expresan de forma lingüística.

En esas condiciones, $FRA(\cdot)$ en la ecuación 12.5 debe ser la extensión a números difusos de 12.38. Esto significa que, para algunas aplicaciones, puede ser necesario construir una máquina de inferencia ‘a la medida’ de las necesidades.

No obstante, también es usual que si se dispone de una ecuación como 12.38, ésta resulte ser monótona. Si eso es así, entonces una forma práctica de extender $\Phi(\cdot)$ es mediante

el algoritmo de extensión directa. En ese caso, la $\text{fra}(\cdot)$ debe ser $\Phi(\cdot)$ o una adaptación de ella.

Estas $\text{fra}(\cdot)$ son, entonces, funciones *ad-hoc*. Son funciones diseñadas explícitamente para una aplicación concreta. Son funciones monótonas, pero que no necesariamente están definidas en los intervalos unitarios. Por ello, debe prestarse especial atención en que en su implementación no se violen las condiciones de monotonía, por ejemplo, con la multiplicación de valores cuyos signos no se conozcan con antelación.

Para ilustrar estas ideas, el ejemplo 12.3 muestra el uso de una función *ad hoc* en una aplicación real en el área de la evaluación de impacto ambiental. Además, en la sección 12.3.2.1 se introducen los operadores OWA, que son un tipo de operadores muy versátiles para la construcción de funciones a partir del conocimiento de los expertos.

Ejemplo 12.3 ($\text{fra}(\cdot)$ *ad hoc*. Valor ambiental del suelo.) La Evaluación de Impacto Ambiental es un procedimiento que busca identificar, predecir e interpretar los impactos ambientales que un proyecto produciría, si es ejecutado [26]. El procedimiento suele incluir un estudio por separado de cada uno de los impactos identificados, y luego una estimación del impacto total del proyecto. Existen varias metodologías de uso general, es decir, que se pueden aplicar a una amplia variedad de proyectos, y otras diseñadas para propósitos específicos.

Dentro de este último caso se encuentra EVIAVE (Evaluación de Impacto Ambiental de Vertederos), una metodología diseñada en la Universidad de Granada específicamente para diagnosticar ambientalmente los vertederos de residuos sólidos urbanos y establecer su grado de interacción y afectación al medio ambiente [31, 58, 2, 16, 20].

La metodología es compleja y en este ejemplo solamente tomamos, con propósitos ilustrativos, el cálculo de una de las muchas variables que intervienen, el ‘valor ambiental del suelo’. Aprovechamos, eso sí, para mencionar que EVIAVE se implementa como una red de muchos SBA interconectados entre sí (sección 13.2, ejemplo 13.2).

Para el cálculo del valor ambiental del suelo, EVIAVE utiliza tres variables:

- D_1 : uso del suelo.
- D_2 : tipo de vegetación.
- D_3 : cobertura vegetal.

Cada una de esas tres variables tiene asociada una variable lingüística cuyo universo de discurso es el intervalo $[0, 5]$. La variable V_s (valor ambiental del suelo) también tiene asociada su propia variable lingüística, con el mismo universo de discurso. La metodología EVIAVE establece la forma de estimar V_s así:

$$V_s = \frac{D_1 + D_2 + D_3}{3} \quad (12.39)$$

En EVIAVE la variable V_s (y todas las demás) se calcula con un SBA. El conocimiento para construir la $\text{fra}(\cdot)$ está explícito en la ecuación 12.39: V_s es el promedio simple de las variables D_1 , D_2 y D_3 . Por ello, podemos utilizar una combinación lineal (opción 1, variante 1) con los siguientes parámetros:

$$\begin{aligned} x_1 &: D_1 & x_2 &: D_2 & x_3 &: D_3 \\ s_1 &= 1 & s_2 &= 1 & s_3 &= 1 \\ w_1 &= \frac{1}{3} & w_2 &= \frac{1}{3} & w_3 &= \frac{1}{3} \end{aligned} \quad (12.40)$$

12.3.2.1. Operadores OWA

En la sección 2.2.1 se presentaron someramente los operadores de agregación. De ellos, los más conocidos son:

1. Mínimo: $y = \min(x_1, x_2, \dots, x_p)$
2. Máximo: $y = \max(x_1, x_2, \dots, x_p)$
3. Promedio simple: $y = (x_1 + x_2 + \dots + x_p)/p$
4. Promedio ponderado $y = \sum_{i=1}^p w_i x_i$ con $0 \leq w_i \leq 1$ y $\sum_{i=1}^p w_i = 1$

Un grupo de operadores de agregación especial corresponde a los Promedios Ponderados Ordenados, que se conocen como los OWA, por su sigla en inglés (*Ordered Weighted Average*). Un OWA es un promedio ponderado, que se aplica sobre las entradas previamente ordenadas de mayor a menor, y de allí su nombre. En términos matemáticos:

$$\begin{aligned} \text{owa}(\cdot) &: \mathcal{R}^p \rightarrow \mathcal{R} \\ y &= \text{owa}(x_1, x_2, \dots, x_p) \\ y &= \sum_{i=1}^p w_i \bar{x}_i \quad \sum_{i=1}^p w_i = 1 \quad 0 \leq w_i \leq 1 \quad \forall i \end{aligned} \quad (12.41)$$

Donde \bar{x}_i es el i -ésimo elemento de $\{x_1, x_2, \dots, x_p\}$ previamente ordenado de mayor a menor.

Ejemplo 12.4 (Operador OWA) Supongamos un operador $\text{owa}(\cdot)$ definido para tres entradas con el siguiente vector de pesos:

$$\mathbf{w} = \{0.5, 0.3, 0.2\} \quad (12.42)$$

Nótese que todos los pesos son positivos y su suma es 1, por lo que definen un operador $\text{owa}(\cdot)$ válido.

Supongamos ahora que las entradas que debe agregar el operador son las siguientes:

$$x_1 = 2 \quad x_2 = 5 \quad x_3 = 3 \quad (12.43)$$

Para aplicar el operador, primero se ordenan las entradas de mayor a menor, lo que resulta en

$$\bar{x}_1 = 5 \quad \bar{x}_2 = 3 \quad \bar{x}_3 = 2 \quad (12.44)$$

Y después se calcula el promedio ponderado

$$y = \text{owa}(x_1, x_2, x_3) = 0.5 \times 5 + 0.3 \times 3 + 0.2 \times 2 = 3.8 \quad (12.45)$$

Si, en cambio, las entradas resultan ser

$$x_1 = 3 \quad x_2 = 1 \quad x_3 = 2 \quad (12.46)$$

Para aplicar el operador, primero se ordenan las entradas de mayor a menor, lo que resulta en

$$\bar{x}_1 = 3 \quad \bar{x}_2 = 2 \quad \bar{x}_3 = 1 \quad (12.47)$$

Y después se calcula el promedio ponderado

$$y = \text{owa}(x_1, x_2, x_3) = 0.5 \times 3 + 0.3 \times 2 + 0.2 \times 1 = 2.2 \quad (12.48)$$

Este ejemplo permite ilustrar que los pesos no están asignados a una entrada en específico, sino a una posición resultante del ordenamiento de las entradas.

Una de las características de los OWA que los hace muy interesantes, es que la selección del vector de pesos $\mathbf{w} = \{w_1, \dots, w_n\}$ establece diferentes tipos de agregaciones; por ejemplo:

1. Mínimo: con $\mathbf{w} = \{0, 0, \dots, 0, 1\}$

2. Máximo: con $\mathbf{w} = \{1, 0, 0, \dots, 0\}$
3. Promedio simple: con $\mathbf{w} = \{1/n, 1/n, \dots, 1/n\}$

En la literatura hay una gran colección de propuestas sobre cómo construir un vector de pesos adecuado para un operador owa(\cdot) [39, 27, 28, 38]. Nosotros sugerimos el uso de una familia de operadores propuesta en [21]. Lo que tiene de interesante esta familia es que facilita la selección del vector de pesos, en función del tipo de agregación que se quiera realizar. Mediante un parámetro $\Phi \in [0, 1]$ se controla el tipo de agregación del operador OWA así:

- Cuando $\Phi = 0.0$ el operador es equivalente al operador mínimo.
- Cuando $\Phi = 0.5$ el operador es equivalente al operador promedio simple.
- Cuando $\Phi = 1.0$ el operador es equivalente al operador máximo.
- Al aumentar Φ la agregación varía suavemente.

En otras palabras, las agregaciones tipo OR y tipo AND pueden realizarse con un operador OWA al que se le ajusta el parámetro Φ : un valor de Φ cercano a 1 permitirá una agregación tipo OR, y un valor de Φ cercano a 0 permitirá una agregación tipo AND. Además, se pueden realizar agregaciones intermedias. El nombre que recibe Φ es el de *orness* del operador owa(\cdot), que podríamos traducir como ‘qué tan cercano al operador OR resulta ser’ el operador owa(\cdot).

La forma de obtener el vector de pesos se plasma en el algoritmo 13. La demostración se encuentra en [21], y se ha reproducido en el apéndice A.

Ahora bien, para que un operador owa(\cdot) se pueda utilizar como función de razonamiento aproximado deben cumplirse las condiciones de monotonía. Esto establece la siguiente restricción importante: *Un operador owa(\cdot) es una fra(\cdot) válida sólo en aquellos casos en que todas las entradas tiene la misma monotonía*. La monotonía puede ser creciente o decreciente; lo importante es que sea igual para todas las entradas.

12.4. Interpretación lingüística

En la sección 12.1 se explica que el papel del bloque de interpretación lingüística es el de construir los números difusos sobre los que operará la máquina de inferencia. El bloque de interpretación lingüística produce un número difuso por cada una de las entradas (Figura 12.2c). En la explicación que viene a continuación se asume que hay una única entrada, lo que simplifica la notación. Para varias entradas, el proceso se realiza de forma similar e independiente en cada una de ellas.

Algoritmo 13 Pesos de un operador OWA

Entradas:

p , el número de entradas del operador owa(\cdot)
 Φ , el *orness* deseado del operador owa(\cdot). Es un número en el intervalo $[0, 1]$

```

1: procedure PESOSOWA( $p, \Phi$ )
2:   Defina  $f(x) = 1 - (1 - 2x)^\beta$  con  $\beta \in [1, 1.5]$ .            $\triangleright$  por defecto  $\beta = 1.5$ 
3:   if  $\Phi < 0.5$  then
4:      $\nu \leftarrow \Phi$ 
5:   else
6:      $\nu \leftarrow 1 - \Phi$ 
7:    $m = p - 1$ 
8:    $\Delta = f(\nu)m/p$ 
9:    $K \leftarrow 6(f(\nu) - 2\nu)/(m^2 - 1)$ 
10:   $b \leftarrow f(\nu)/p - Kp/2$ 
11:  if  $\Phi < 0.5$  then
12:     $w_p \leftarrow 1 - \Delta$ 
13:    for  $i = 1$  to  $m$  do
14:       $w_i \leftarrow Ki + b$ 
15:    else
16:       $w_1 \leftarrow 1 - \Delta$ 
17:      for  $i = 1$  to  $m$  do
18:         $w_{i+1} \leftarrow K(p - i) + b$ 
19:  return  $w$             $\triangleright$  Vector con los pesos del operador owa( $\cdot$ )

```

Supongamos que la variable de entrada tiene asociada la variable lingüística \mathcal{V}_e definida por (ecuación 3.1):

$$\mathcal{V}_e : \langle \mathcal{X}, \mathcal{L}\mathcal{X}, \mathcal{U}, \mathcal{M}_{\mathcal{U}} \rangle \quad (12.49)$$

En donde, además:

- El universo de discurso \mathcal{U} es el intervalo

$$\mathcal{U} = [\mathcal{U}_L, \mathcal{U}_R] \quad (12.50)$$

- El conjunto de términos lingüísticos (etiquetas) es:

$$\mathcal{L}\mathcal{X} = \{\Gamma_1, \Gamma_2, \dots, \Gamma_t\} \quad (12.51)$$

- La k -ésima etiqueta tiene un conjunto difuso asociado \mathcal{E}_k que es un número difuso. La función de pertenencia de ese número se denota por $\mu_{\mathcal{E}_k}(x)$ y sus α -cortes son

$$\mathcal{E}_{k_\alpha} = [L_{\mathcal{E}_k}(\alpha), R_{\mathcal{E}_k}(\alpha)] \quad (12.52)$$

El bloque de interpretación lingüística es capaz de construir el número difuso que alimenta la máquina de inferencia a partir de entradas de diferentes tipos. La Tabla 12.2 muestra qué tipo de entradas son válidas y un ejemplo de cuál sería el número difuso que se produciría en cada caso. Esos ejemplos suponen que la variable lingüística tiene 5 términos (Muy Bajo, Bajo, Medio, Alto y Muy Alto) cuyos conjuntos difusos asociados se han dibujado en línea discontinua azul. La salida del bloque de interpretación lingüística es el número difuso x , que se puede caracterizar por sus α -cortes, x_α . En la Tabla, x se ha dibujado en color rojo.

A continuación se explica qué tipo de entradas son válidas, y cómo se construye x a partir de ellas:

Entrada del tipo ‘etiqueta’ : la entrada puede ser una de las etiquetas de la variable lingüística, es decir, uno de sus términos lingüísticos. Por ejemplo, la entrada puede ser la palabra ‘Medio’.

Como cada etiqueta tiene asociado un conjunto difuso, que es un número difuso, la salida obvia del bloque de interpretación lingüística es precisamente ese número. Si la entrada es la k -ésima etiqueta, entonces los α -cortes de x serán los mismos de \mathcal{E}_k :

$$x_\alpha = [L_{\mathcal{E}_k}(\alpha), R_{\mathcal{E}_k}(\alpha)] \quad (12.53)$$

Entrada del tipo ‘etiqueta modificada’ : la entrada puede ser una etiqueta con un modificador lingüístico. Los modificadores válidos son ‘**por lo menos**’ y ‘**a lo sumo**’:

- El modificador ‘**por lo menos**’ se interpreta como ‘mayor o igual que’, de tal forma que una entrada ‘por lo menos Medio’ se interpreta como ‘mayor o igual que Medio’. Para construir el número difuso x que representa esa entrada, se mantiene la curva $L(\alpha)$ de la etiqueta, mientras que la curva $R(\alpha)$ se remplaza por el valor más pequeño posible, es decir, el límite superior del universo de discurso. En conclusión,

$$x_\alpha = [L_{\mathcal{E}_k}(\alpha), \mathcal{U}_R,] \quad (12.54)$$

- El modificador ‘**a lo sumo**’ se interpreta como ‘menor o igual que’, de tal forma que una entrada ‘a lo sumo Medio’ se interpreta como ‘menor o igual que Medio’. Para construir el número difuso x que representa esa entrada, se mantiene la curva $R(\alpha)$ de la etiqueta, mientras que la curva $L(\alpha)$ se remplaza por el valor más pequeño posible, es decir, el límite inferior del universo de discurso. En conclusión,

$$x_\alpha = [\mathcal{U}_L, R_{\mathcal{E}_k}(\alpha)] \quad (12.55)$$

Entrada del tipo ‘combinación de etiquetas’ : la entrada puede ser la combinación de dos etiquetas con el modificador ‘entre’, por ejemplo ‘entre Bajo y Alto’. Esta entrada se interpreta como un ‘mayor o igual que Bajo y menor o igual que Alto’. Si los números difusos de las dos etiquetas son \mathcal{E}_{k_1} y \mathcal{E}_{k_2} , el número difuso x se construye tomando los valores menores y mayores de sus α -cortes:

$$x_\alpha = [\min\{L_{\mathcal{E}_{k_1}}(\alpha), L_{\mathcal{E}_{k_2}}(\alpha)\}, \max\{R_{\mathcal{E}_{k_1}}(\alpha), R_{\mathcal{E}_{k_2}}(\alpha)\}] \quad (12.56)$$

Si, como es de suponer, los términos lingüísticos están ordenados con $k_1 < k_2$ y sus conjuntos difusos establecen una partición difusa del universo de discurso, entonces la ecuación 12.56 se convierte en:

$$x_\alpha = [L_{\mathcal{E}_{k_1}}(\alpha), R_{\mathcal{E}_{k_2}}(\alpha)] \quad (12.57)$$

Entrada del tipo ‘expresión’ : la entrada puede ser una de las siguientes expresiones: ‘**cualquier cosa**’, ‘**lo mínimo**’ ó ‘**lo máximo**’:

- La expresión ‘**cualquier cosa**’ se interpreta como ignorancia total sobre el valor de la entrada. Por tanto, x es un número difuso que cubre todo el universo de discurso.

$$x_\alpha = [\mathcal{U}_L, \mathcal{U}_R,] \quad (12.58)$$

- La expresión ‘**mínimo**’ se interpreta como un conocimiento preciso sobre el valor de la entrada que corresponde al menor valor posible. Por tanto, x es un singleton en el límite inferior del universo de discurso.

$$x_\alpha = [\mathcal{U}_L, \mathcal{U}_L] \quad (12.59)$$

- La expresión ‘**máximo**’ se interpreta como un conocimiento preciso sobre el valor de la entrada que corresponde al mayor valor posible. Por tanto, x es un singleton en el límite superior del universo de discurso.

$$x_\alpha = [\mathcal{U}_R, \mathcal{U}_R] \quad (12.60)$$

Entrada del tipo ‘número concreto’ : la entrada puede ser un número concreto.

Este tipo de entrada es un conocimiento preciso sobre su valor . Por tanto, x es un singleton centrado en ese valor. Si el número concreto es c , entonces x está definido por

$$x_\alpha = [c, c] \quad (12.61)$$

Entrada del tipo ‘intervalo’ : la entrada puede ser un intervalo. Este tipo de entrada es un conocimiento preciso sobre los límites de su valor . Por tanto, x es el número difuso que representa ese intyervalo, es decir, un número rectangular definido por sus límites. Si el intervalo es $[a, b]$, entonces x está definido por

$$x_\alpha = [a, b] \quad (12.62)$$

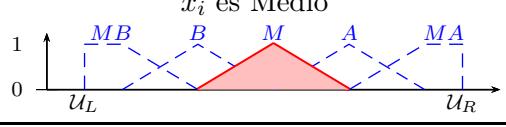
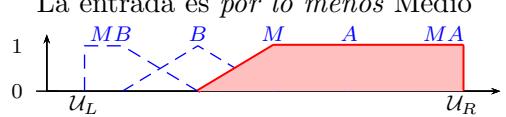
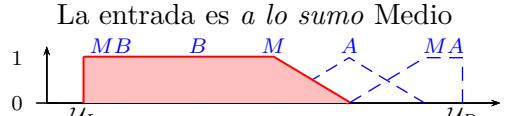
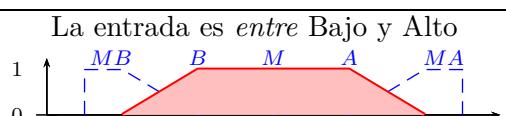
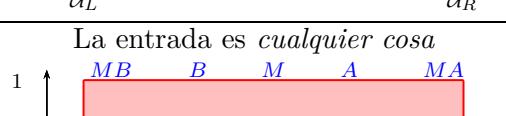
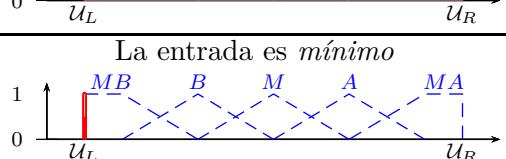
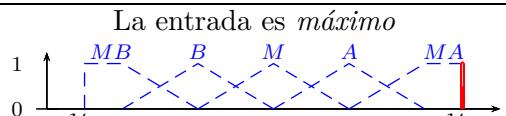
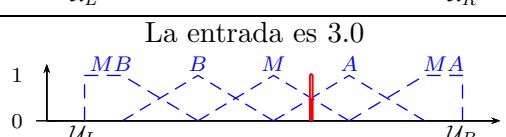
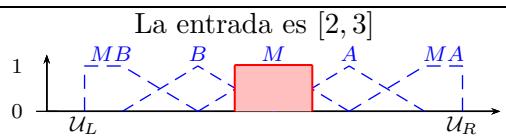
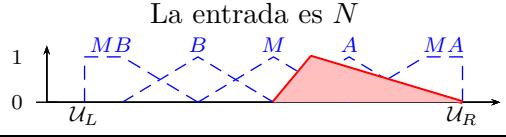
Entrada del tipo ‘número difuso’ : la entrada puede ser un número difuso N , en cuyo caso x debe ser ese mismo número difuso, es decir:

$$x_\alpha = N_\alpha \quad (12.63)$$

12.4.1. Escalamiento

En ocasiones, el rango en que se ha definido el universo de discurso de una variable lingüística de entrada no coincide con el rango en que se ha definido la fra(.). Esto es usual si, por ejemplo, la fra(.) se ha definido sobre los intervalos unitarios, mientras que la variable lingüística se ha definido sobre un universo que tiene un sentido físico.

Tabla 12.2: Entradas válidas al bloque de interpretación lingüística y su salida

Tipo de entrada		Ejemplo
1	Etiqueta $x_\alpha = [L_{\mathcal{E}_k}(\alpha), R_{\mathcal{E}_k}(\alpha)]$	
2	Etiqueta modificada $x_\alpha = [L_{\mathcal{E}_k}(\alpha), U_R]$	La entrada es <i>por lo menos</i> Medio 
3	Etiqueta modificada $x_\alpha = [U_L, R_{\mathcal{E}_k}(\alpha)]$	La entrada es <i>a lo sumo</i> Medio 
4	Combinación de etiquetas $x_\alpha = [L_{\mathcal{E}_{k_1}}(\alpha), R_{\mathcal{E}_{k_2}}(\alpha)]$	La entrada es <i>entre</i> Bajo y Alto 
5	Expresión $x_\alpha = [U_L, U_R]$	La entrada es <i>cualquier cosa</i> 
6	Expresión $x_\alpha = [U_L, U_L]$	La entrada es <i>mínimo</i> 
7	Expresión $x_\alpha = [U_R, U_R]$	La entrada es <i>máximo</i> 
8	Número concreto $x_\alpha = [c, c]$	La entrada es 3.0 
9	Intervalo $x_\alpha = [a, b]$	La entrada es [2, 3] 
10	Número difuso $x_\alpha = N_\alpha$	La entrada es <i>N</i> 

Cuando esto sucede, es necesario efectuar un escalamiento del número difuso antes de usarlo como entrada de la $\text{fra}(\cdot)$. Supongamos que la $\text{fra}(\cdot)$ fue construida considerando que la entrada está definida en el rango $[E_L, E_R]$. El número escalado x_e estará definido por sus α -cortes así:

$$\begin{cases} L_{x_e}(\alpha) = E_L + (L_x(\alpha) - \mathcal{U}_L) \frac{E_R - E_L}{\mathcal{U}_R - \mathcal{U}_L} \\ R_{x_e}(\alpha) = E_L + (R_x(\alpha) - \mathcal{U}_L) \frac{E_R - E_L}{\mathcal{U}_R - \mathcal{U}_L} \end{cases} \quad (12.64)$$

Ahora bien, lo más usual es que se necesite escalar al intervalo $[0, 1]$, en cuyo caso $E_L = 0$ y $E_R = 1$ lo que lleva a

$$\begin{cases} L_{x_e}(\alpha) = \frac{L_x(\alpha) - \mathcal{U}_L}{\mathcal{U}_R - \mathcal{U}_L} \\ R_{x_e}(\alpha) = \frac{R_x(\alpha) - \mathcal{U}_L}{\mathcal{U}_R - \mathcal{U}_L} \end{cases} \quad (12.65)$$

Ejemplo 12.5 (Interpretación lingüística) Retomemos ahora los ejemplo 12.1 y 12.2 en los que se esboza un SBA que permite estimar el riesgo de daño causado por descargas eléctricas atmosféricas. En dicho ejemplo los valores de las entradas entradas al sistema son (Figura 12.5):

- I : la Intensidad es Alta.
- D : la Densidad es Alta.
- P : el Nivel de protección es Alto.
- M : la importancia es Media

Nótese que los cuatro valores corresponden a etiquetas de las variables lingüísticas (Figura 12.4). Las cuatro variables se han construido espaciando homogéneamente los conjuntos. Por ello, podemos establecer que los valores de las cuatro entradas son los siguientes trapecios

$$\begin{cases} I = T(390/9, 450/9, 510/9, 570/9) \\ D = T(250/9, 300/9, 350/9, 400/9) \\ P = T(3/5, 4/5, 5/5, 5/5) \\ M = T(1/4, 2/4, 2/4, 3/4) \end{cases} \quad (12.66)$$

Ahora bien, como la fra(\cdot) del sistema estpá definida sobre los intervalos unitarios (Ecuación 12.37 del ejemplo 12.2), es necesario escalar los números difusos de las variables I , D . El resultado se muestra en la Figura 12.10 y corresponde a

$$\begin{cases} x_1 = T(5/9, 6/9, 7/9, 8/9) \\ x_2 = T(5/9, 6/9, 7/9, 8/9) \\ x_3 = T(3/5, 4/5, 5/5, 5/5) \\ x_4 = T(1/4, 2/4, 2/4, 3/4) \end{cases} \quad (12.67)$$

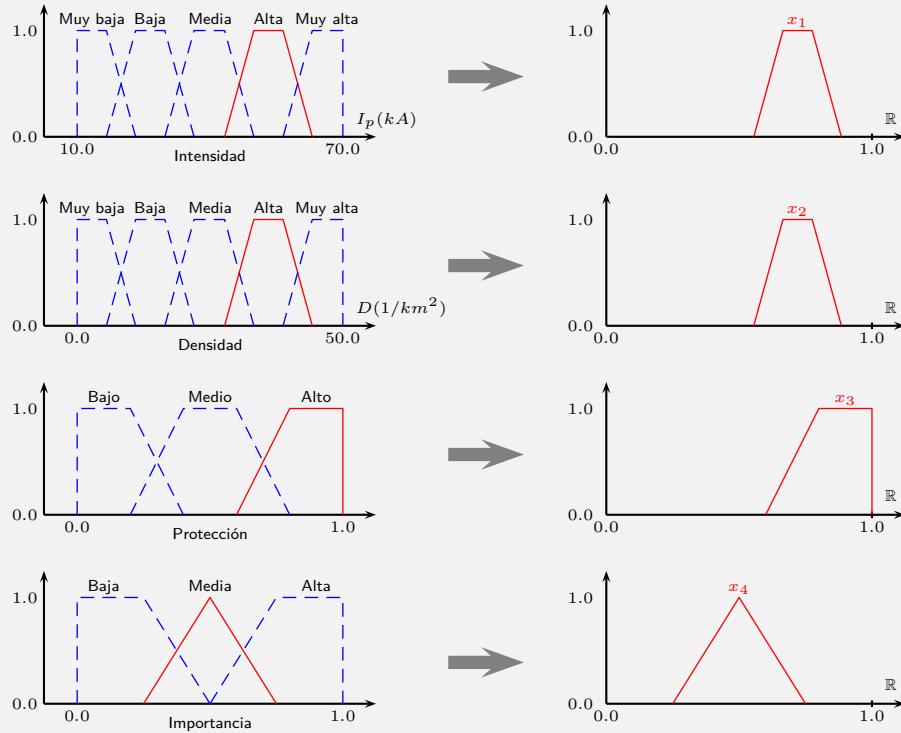


Figura 12.10: Interpretación lingüística en el ejemplo 12.5

12.5. Aproximación lingüística

En la sección 12.1 se explica que el papel del bloque de aproximación lingüística es el de transformar la salida de la máquina de inferencia, que es un número difuso y , en la salida del sistema (Figura 12.2d).

Como los SBA son sistemas tipo MISO, sólo hay una salida \mathbf{s} , que tiene una variable lingüística asociada \mathcal{V}_s :

$$\mathcal{V}_s : \langle \mathcal{X}, \mathcal{L}\mathcal{X}, \mathcal{U}, \mathcal{M}_{\mathcal{U}} \rangle \quad (12.68)$$

Si el rango de salida de la $\text{fra}(\cdot)$ es diferente al universo de discurso $\mathcal{U} = [\mathcal{U}_L, \mathcal{U}_R]$, es necesario efectuar primero un escalamiento de y antes de realizar la aproximación lingüística propiamente dicha. Este procedimiento se presenta en la sección 12.5.2. Antes, sin embargo, presentamos en la sección 12.5.1 el concepto de consistencia entre números difusos, que será fundamental en el proceso de aproximación lingüística.

12.5.1. Consistencia entre dos conjuntos difusos

Supongamos dos conjuntos definidos sobre el mismo universo de discurso \mathcal{U} . La consistencia entre dichos conjuntos es una medida de qué tanto pueden llegar a representar un mismo significado. Ese significado compartido se determina, meramente, mediante su intersección, de allí que la consistencia sea definida como el máximo grado de pertenencia de la intersección de los dos conjuntos:

$$\text{cons}(A, B) = \max_{x \in \mathcal{U}} \{\mu_{A \cap B}(x)\} \quad (12.69)$$

Si utilizamos la \mathcal{T} -norma del operador \min para realizar la intersección, obtenemos

$$\text{cons}(A, B) = \max_{x \in \mathcal{U}} \{\min(\mu_A(x), \mu_B(x))\} \quad (12.70)$$

Ejemplo 12.6 (Consistencia entre conjuntos difusos) Para ilustrar el concepto de consistencia, consideremos la variable lingüística \mathcal{V} que se muestra en la Figura 12.11a y el conjunto difuso y de la Figura 12.11b.

Las Figuras 12.11c, 12.11d y 12.11e muestran el procedimiento para calcular la consistencia entre cada uno de los conjuntos de \mathcal{V} y el conjunto y . En cada caso se ha obtenido la intersección entre los conjuntos respectivos, y se ha buscado cuál es el máximo grado de pertenencia de esa intersección.

Los resultados obtenidos son:

$$\left\{ \begin{array}{lcl} \text{cons(Bajo, } y) & = & 0.60 \\ \text{cons(Medio, } y) & = & 0.83 \\ \text{cons(Alto, } y) & = & 0.17 \end{array} \right. \quad (12.71)$$

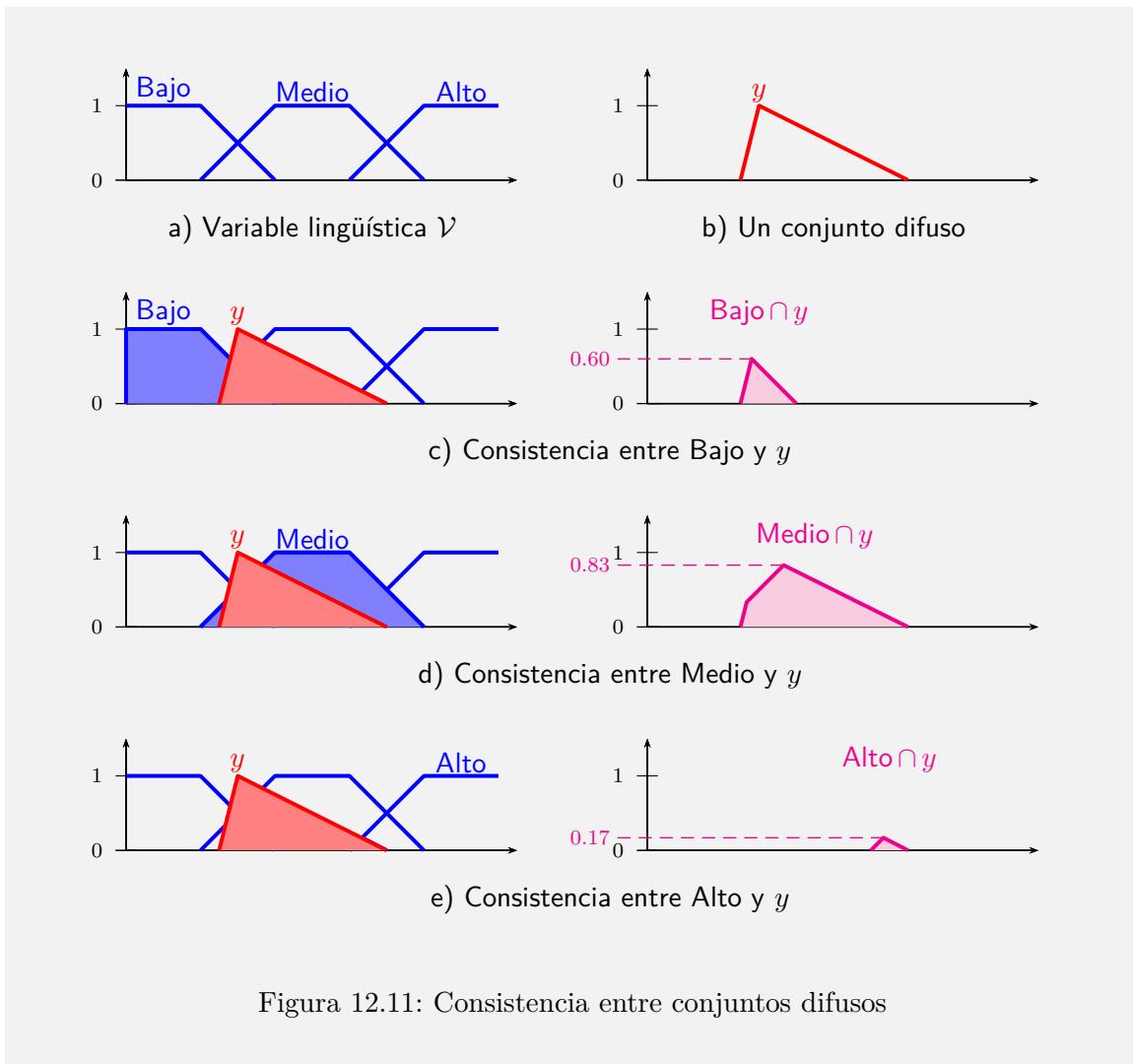


Figura 12.11: Consistencia entre conjuntos difusos

12.5.2. Escalamiento

Supongamos que la $\text{fra}(\cdot)$ del SBA fue construida considerando que la entrada está definida en el rango $[E_L, E_R]$. El número y_e escalado al universo de discurso $\mathcal{U} = [\mathcal{U}_L, \mathcal{U}_R]$ estará definido por sus α -cortes así:

$$\begin{cases} L_{y_e}(\alpha) = \mathcal{U}_L + (L_y(\alpha) - E_L) \frac{\mathcal{U}_R - \mathcal{U}_L}{E_R - E_L} \\ R_{y_e}(\alpha) = \mathcal{U}_L + (R_y(\alpha) - E_L) \frac{\mathcal{U}_R - \mathcal{U}_L}{E_R - E_L} \end{cases} \quad (12.72)$$

Si, como es usual, se necesita escalar desde el intervalo $[0, 1]$, se tiene $E_L = 0$ y $E_R = 1$, en cuyo caso

$$\begin{cases} L_{y_e}(\alpha) = \mathcal{U}_L + L_y(\alpha)(\mathcal{U}_R - \mathcal{U}_L) \\ R_{y_e}(\alpha) = \mathcal{U}_L + R_y(\alpha)(\mathcal{U}_R - \mathcal{U}_L) \end{cases} \quad (12.73)$$

12.5.3. Construcción de la salida

Supongamos que la variable de salida tiene asociada la variable lingüística \mathcal{V}_s definida por (ecuación 3.1):

$$\mathcal{V}_s : \langle \mathcal{X}, \mathcal{L}\mathcal{X}, \mathcal{U}, \mathcal{M}_{\mathcal{U}} \rangle \quad (12.74)$$

En donde, además:

- El universo de discurso \mathcal{U} es el intervalo

$$\mathcal{U} = [\mathcal{U}_L, \mathcal{U}_R] \quad (12.75)$$

- El conjunto de términos lingüísticos (etiquetas) es:

$$\mathcal{L}\mathcal{X} = \{\Gamma_1, \Gamma_2, \dots, \Gamma_t\} \quad (12.76)$$

- La k -ésima etiqueta tiene un conjunto difuso asociado \mathcal{E}_k que es un número difuso.

El bloque de aproximación lingüística recibe un número difuso que proviene de la máquina de inferencia y es capaz de generar salidas de diferentes tipos. La Tabla 12.3 muestra qué tipo de salidas disponibles y un ejemplo de cuál sería el número difuso que se produciría en cada caso. Estos ejemplos, a su vez, utilizan el caso estudiado en el Ejemplo 12.6.

A continuación se explica qué tipo de salidas están disponibles, y cómo se construye la salida s a partir de ellas, cuando el resultado de la inferencia es el número $y = \text{fra}(\cdot)$:

Salida del tipo ‘etiqueta’ : la salida puede ser una de las etiquetas de la variable lingüística de salida. Para determinar cuál de ellas, se calcula la consistencia entre su respectivo conjunto difuso asociado y el resultado de la inferencia, para tomar aquella etiqueta que tenga la mayor consistencia.

$$\mathbf{s} = \Gamma_k \mid \text{cons}(\mathcal{E}_k, y) = \max_i \{\text{cons}(\mathcal{E}_i, y)\} \quad (12.77)$$

En caso de empate, la aplicación específica en que se esté usando el SBA determinará cuál es el mejor criterio de desempate.

Salida del tipo ‘etiqueta subjetiva’ : la salida puede ser una de las etiquetas de la variable lingüística de salida que dependa del nivel de optimismo. Para determinar cuál de ellas, se calcula el valor representativo del resultado de la inferencia en función del optimismo y se construye un número tipo singleton centrado en ese valor. Posteriormente se calcula la consistencia entre el singleton y los conjuntos difusos asociados a las etiquetas, para tomar aquella etiqueta que tenga la mayor consistencia.

Nótese que si B es un singleton centrado en c , entonces para cualquier otro conjunto A se tiene que

$$\text{cons}(A, B) = \mu_A(c) \quad (12.78)$$

Por tanto,

$$\mathbf{s}(\beta) = \Gamma_k \mid \mu_{\mathcal{E}_k}(\mathbf{V}_y(\beta)) = \max_i \{\mu_{\mathcal{E}_i}(\mathbf{V}_y(\beta))\} \quad (12.79)$$

Salida del tipo ‘frase’ : la salida puede ser una frase descriptiva que tenga información sobre la consistencia de y con cada una de las etiquetas. Para ello, su utilizan las expresiones ‘posiblemente’ y ‘muy posiblemente’ como adverbios calificadores que pueden acompañar a cada etiqueta.

Para cada etiqueta Γ se construye una expresión que puede ser ‘muy posiblemente Γ ’, ‘posiblemente Γ ’ o estar vacía. El criterio para decidir cuál de ellas se basa en consistencia entre el conjunto difuso asociado a la etiqueta y y y . Si denotamos por μ dicha consistencia, podemos usar un criterio basado en umbrales así:

$$\text{pos}(\mu, \Gamma) = \begin{cases} \text{‘muy posiblemente } \Gamma\text{’} & \text{si } 2/3 \leq \mu \leq 1 \\ \text{‘posiblemente } \Gamma\text{’} & \text{si } 1/3 \leq \mu < 2/3 \\ ‘ & \text{si } 0 \leq \mu < 1/3 \end{cases} \quad (12.80)$$

La salida \mathbf{s} será la concatenación de las expresiones que resultan de cada etiqueta. Si usamos el símbolo de sumatoria para referirnos a esa concatenación podremos escribir

$$\mathbf{s} = \sum_{k=1}^t \text{pos}(\text{cons}(\mathcal{E}_k, y), \Gamma_k) \quad (12.81)$$

Salida del tipo ‘número concreto’ : la salida puede ser un número concreto. El número corresponde al valor representativo (ecuación 7.17) del número y

$$\mathbf{s}(\beta) = \mathbf{V}_y(\beta) \quad (12.82)$$

El número concreto que resulta de esta salida depende el nivel de optimismo β , lo que permitirá valorar la subjetividad de un decisor. Por defecto, se toma $\beta = 0.5$

Salida del tipo ‘pareja de números concreto’ : la salida puede ser una pareja de números concretos. El primer número es el valor representativo de y y el segundo su ambigüedad (ecuación 7.36):

$$\mathbf{s}(\beta) = (\mathbf{V}_y(\beta), \mathbf{A}_y) \quad (12.83)$$

Salida del tipo ‘número difuso’ : la salida puede ser el mismo número difuso y . Esta opción es equivalente a ignorar por completo el bloque de aproximación lingüística.

$$\mathbf{s} = y \quad (12.84)$$

Ejemplo 12.7 (Aproximación lingüística) Retomemos ahora los ejemplos 12.1, 12.2 y 12.5 en los que se utiliza un SBA que permite estimar el riesgo de daño causado por descargas eléctricas atmosféricas.

Para las entradas del ejemplo, el bloque de interpretación lingüística produce 4 números trapezoidales (ecuación 12.67):

$$\begin{cases} x_1 &= T(5/9, 6/9, 7/9, 8/9) \\ x_2 &= T(5/9, 6/9, 7/9, 8/9) \\ x_3 &= T(3/5, 4/5, 5/5, 5/5) \\ x_4 &= T(1/4, 2/4, 2/4, 3/4) \end{cases} \quad (12.85)$$

La máquina de inferencia debe aplicar la $\text{fra}(\cdot)$ del sistema (ecuación 12.37):

Tabla 12.3: Salidas disponibles del bloque de aproximación lingüística

Tipo de salida		Ejemplo
1	Etiqueta $s = \Gamma_k \mid \text{cons}(\mathcal{E}_k, y) = \max_i \{\text{cons}(\mathcal{E}_i, y)\}$	La salida es Medio
1	Etiqueta subjetiva $s(\beta) = \Gamma_k \mid \mu_{\mathcal{E}_k}(\mathbf{V}_y(\beta)) = \max_i \{\mu_{\mathcal{E}_i}(\mathbf{V}_y(\beta))\}$	Para $\beta = 0.5$ la salida es Medio
3	Frase $s = \sum_{k=1}^t \text{pos}(\text{cons}(\mathcal{E}_k, y), \Gamma_k)$	La salida es muy posiblemente Medio y posiblemente Bajo
4	Número concreto $s(\beta) = \mathbf{V}_y(\beta)$	Para $\beta = 0.5$ la salida es 3.58
5	Pareja de números concretos $s(\beta) = (\mathbf{V}_y(\beta), \mathbf{A}_y)$	Para $\beta = 0.5$ la salida es (3.58, 1.5)
6	Número difuso $s = y$	La salida es y

$$y = \text{fra}(x_1, x_2, x_3, x_4) = 0.4x_1 + 0.2x_2 + 0.2(1 - x_3) + 0.2x_4 \quad (12.86)$$

Si bien es cierto que para obtener y podemos usar el algoritmo de extensión directa, también es cierto que el cálculo puede realizarse de forma más sencilla por tratarse de operaciones aritméticas elementales entre trapezios. En cualquier caso el resultado es:

$$y = T(0.463, 0.540, 0.566, 0.683) \quad (12.87)$$

El resultado se muestra en color rojo en la Figura 12.12, junto con la variable lingüística de la salida, Riesgo.

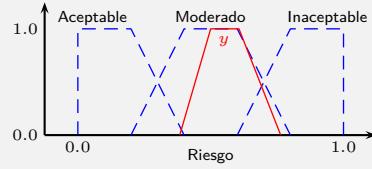


Figura 12.12: Aproximación lingüística en el ejemplo 12.7

El valor representativo y y la ambigüedad de y resultan ser

$$\mathbf{V}_y(0.5) = 0.56 \quad \mathbf{A}_y = 0.20 \quad (12.88)$$

y su consistencia con los conjuntos asociados a las etiquetas ‘Aceptable’, ‘Moderado’ e ‘Inaceptable’ son, respectivamente:

$$\begin{aligned} \text{cons}(\mathcal{E}_1, y) &= 0.06 \\ \text{cons}(\mathcal{E}_2, y) &= 1.00 \\ \text{cons}(\mathcal{E}_3, y) &= 0.46 \end{aligned} \quad (12.89)$$

Tabla 12.4: Salidas disponibles del ejemplo 12.7

Tipo de salida	Salida
Etiqueta	El riesgo es moderado
Etiqueta subjetiva	Para $\beta = 0.5$ el riesgo es moderado
Frase	El riesgo es muy posiblemente Moderado y posiblemente Inaceptable
Número concreto	Para $\beta = 0.5$ el riesgo es 0.56
Pareja de números concretos	Para $\beta = 0.5$ el riesgo es (0.56, 0.20)
Número difuso	El riesgo es $T(0.463, 0.540, 0.566, 0.683)$

También podemos explorar la incidencia del nivel de optimismo en la salida. En la Figura 12.13 se ha graficado la variación del valor representativo de y con los cambios de β de dos formas. A la izquierda destacando la variación numérica y a la derecha en relación con la variable lingüística. El valor varía desde 0.4611 hasta 0.6589 al aumentar β . En ese intervalo la etiqueta que tiene el máximo grado de pertenencia siempre es 'Moderado', y por tanto la salida del tipo Etiqueta subjetiva será esa para todos los valores de β .

Nótese que en esta aplicación, es peor un valor de salida alto que uno bajo: es peor un riesgo alto que uno bajo. Por ello β debería interpretarse así: β bajo es una actitud hacia el riesgo optimista, mientras que β alto es una actitud hacia el riesgo pesimista.

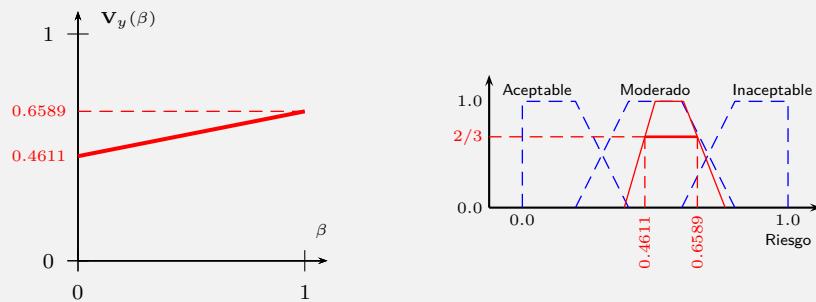


Figura 12.13: Sensibilidad a β en el ejemplo 12.7

Capítulo 13

Anotaciones sobre los SBA

En este capítulo destacamos dos características muy interesantes de los SBA que los diferencian de los SBR. Por una parte, la posibilidad de descubrir el valor de una entrada a partir del valor de la salida y de las demás entradas: esta propiedad la denominamos *razonamiento inverso* y se presenta en la sección 13.1. Por otra parte, la facilidad de conectar SBA entre sí para construir redes de sistemas de computación con palabras; esta propiedad, que se presenta en la sección 13.2 permite construir sistemas de una alta complejidad, con numerosas variables de entrada.

13.1. Sistemas de razonamiento inverso

El elemento central de los SBA es la aplicación del principio de extensión que se facilita mediante el uso del algoritmo de extensión directa. Con esto en mente, formulamos ahora un tipo de sistema complementario cuyo elemento central utiliza el algoritmo que permite obtener las extensiones inversas intermedias (sección 11.5). Los sistemas de este tipo los denominamos **Sistemas de Razonamiento Inverso - SRI**.

Recordemos que el algoritmo en cuestión permite descubrir el valor de una entrada cuando se conocen las demás entradas y la salida. Además, dicho algoritmo permite obtener la extensión inversa posible, la extensión inversa necesaria y una familia de extensiones inversas que son intermedias a esas dos.

La estructura de un SRI se muestra en la Figura 13.1. Para explicarla, comencemos por destacar cuáles son las entradas y cuáles las salidas de un SRI:

Entradas :

- k : es el número de la variable de entrada que se desea descubrir.

- \mathbf{E}_k : es el vector que contiene las todas las entradas, salvo la k -ésima.
- \mathbf{s} : es el valor de la salida, que se supone conocida y por ello se convierte en entrada al SRI
- r : es el parámetro que controla qué tipo de extensión inversa se debe obtener (ecuaciones 11.93 y 11.94).

Salidas :

- \mathbf{e}_k : es la k -ésima entrada, descubierta por el SRI
- \mathbf{s}^{int} : es el valor de la salida que es coherentge con \mathbf{e}_k . No siempre es igual a la entrada \mathbf{s} , porque el algoritmo puede modificarla.

Veamos ahora los bloques que conforman un SRI:

Interpretación lingüística : Cada una de las entradas \mathbf{E}_k , así como la salida \mathbf{s} debe tener una variable lingüística asociada. La interpretación lingüística utiliza estas variables lingüísticas y se realiza tal como se hace para los SBA (sección 12.4).

Aproximación lingüística : La entrada que se calcula \mathbf{e}_k también debe tener una variable lingüística asociada y se asume que a \mathbf{s}^{int} se le asocia la misma variable lingüística que a \mathbf{s} . La aproximación lingüística utiliza estas variables lingüísticas y se realiza tal como se hace para los SBA (sección 12.5).

Máquina de inferencia inversa : este bloque obtiene la extensión inversa intermedia (sección 11.5) mediante la aplicación del algoritmo 11. El algoritmo encuentra la entrada \mathbf{e}_k y construye la salida \mathbf{s}^{int} que es compatible con ella y las demás entradas \mathbf{E}_k .

Funciones inversas : para construir un SRI se supone que existe una $\text{fra}(\cdot)$ monótona, quizás utilizada en un SRA de la forma:

$$\begin{aligned} \text{fra}(\cdot) &: \mathbb{R}^p \rightarrow \mathbb{R} \\ y &= \text{fra}(x_1, \dots, x_p) \end{aligned} \tag{13.1}$$

y que, además, se cuenta con las funciones inversas que despejan cada una de las p entradas:

$$\begin{aligned} \text{fra}_k^{-1}(\cdot) &: \mathbb{R}^p \rightarrow \mathbb{R} \\ x_k &= \text{fra}(x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_p, y) \end{aligned} \tag{13.2}$$

Es importante que estas funciones satizfagan dos condiciones:

- Cada $\text{fra}_k^{-1}(\cdot)$ debe estar definida para todas las posibles combinaciones de sus argumentos.

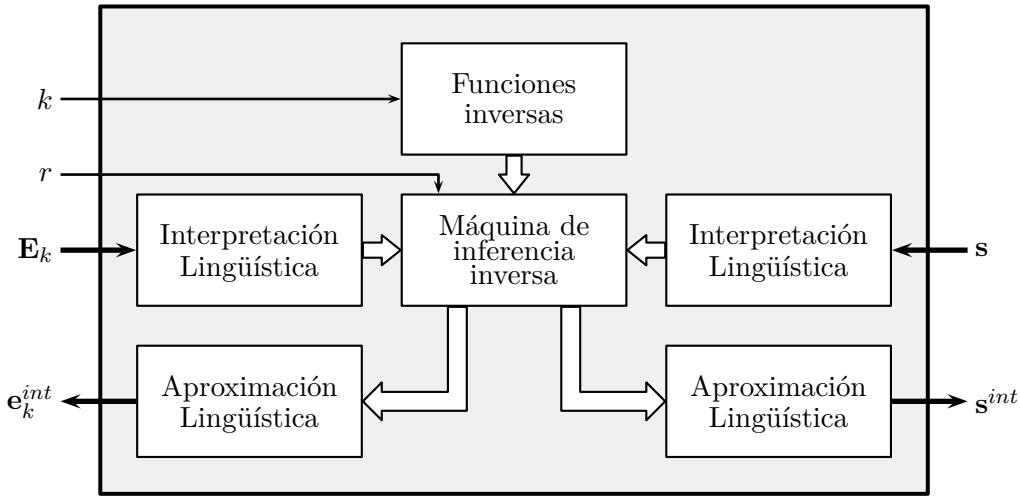


Figura 13.1: Estructura de un SRI

- El rango de cada una de las $\text{fra}_k^{-1}(\cdot)$ debe coincidir con el universo de discurso de la variable lingüística, considerando el escalamiento.

No siempre es evidente si una determinada fra_k^{-1} satisface estas dos condiciones y en ocasiones es necesario forzar el escalamiento cuando se excede el rango. Nótese que esta verificación no la realiza el algoritmo 11 sino que debe implementarse por aparte, bien sea como parte del escalamiento o como parte de la máquina de inferencia inversa.

Esta implementación es un proceso de dos pasos. Supongamos que se han obtenido tanto \mathbf{e}_k como \mathbf{s}^{int} y que el universo de discurso de \mathcal{E}_k es el intervalo $\mathcal{U}_k = [U_{L_k}, U_{R_k}]$. Los dos pasos son:

1. Verificar si $\mathbf{e}_k \in \mathcal{U}_k$. Si esa condición no se cumple, debe modificarse \mathbf{e}_k así:

$$\begin{cases} L_x(\alpha) \leftarrow \max \left(\min \left(L_x(\alpha), U_R \right), U_L \right) \\ R_x(\alpha) \leftarrow \min \left(\max \left(R_x(\alpha), U_L \right), U_R \right) \end{cases} \quad (13.3)$$

2. Recalcular \mathbf{s}^{int} utilizando la $\text{fra}(\cdot)$ y el algoritmo de extensión directa.

Ejemplo 13.1 (Razonamiento inverso en 2D) Consideremos un SBA de dos entradas, x_1 y x_2 . Supongamos que el universo de discurso de cada una de esas variables es el

intervalo $[0, 1]$ y que los pesos y las monotonías están determinadas por:

$$w_1 = 0.8 \quad w_2 = 0.2 \quad s_1 = 1 \quad s_2 = 0 \quad (13.4)$$

Supongamos que se decide implementar la $\text{fra}(\cdot)$ con una combinación lineal (ecuación 12.13), lo que resulta en

$$\begin{aligned} \text{fra}(\cdot) &: [0, 1]^2 \rightarrow [0, 1] \\ \text{fra}(x_1, x_2) &= 0.2 + 0.8x_1 - 0.2x_2 \end{aligned} \quad (13.5)$$

Las funciones inversas son, entonces:

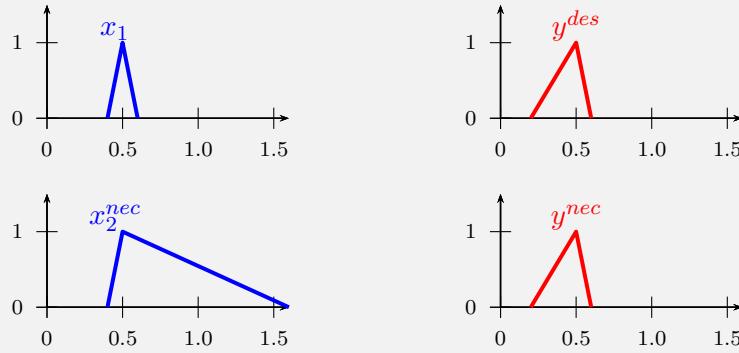
$$\begin{aligned} \text{fra}_1^{-1}(x_2, y) &= \frac{y - 0.2 + 0.2x_2}{0.8} \\ \text{fra}_2^{-1}(x_1, y) &= \frac{y - 0.2 - 0.8x_1}{-0.2} \end{aligned} \quad (13.6)$$

Pese a que la función directa está definida en los intervalos unitarios, las funciones inversas no lo están. Por ejemplo

$$\text{fra}_2^{-1}(0.4, 0.6) = \frac{0.6 - 0.2 - 0.8 \times 0.4}{-0.2} = \frac{0.08}{-0.2} = -0.4 \notin [0, 1] \quad (13.7)$$

Este hecho puede tener por consecuencia que el algoritmo entregue como resultado números difusos que estén por fuera del universo de discurso de su variable. Para ilustrar este hecho tomemos $x_1 = T(0.4, 0.5, 0.5, 0.6)$ y supongamos que se desea descubrir el valor de x_2 con el cual se obtiene una salida deseada que es $y^{des} = T(0.2, 0.5, 0.5, 0.6)$.

En la parte superior de a Figura 13.2 se muestran los números x_1 y y^{des} , mientras que en la parte inferior se muestran los resultados de la aplicación del algoritmo de extensión necesaria: por una parte se observa el valor de x_2^{nec} y el valor de la variable de salida y^{nec} , que coincide con y^{des} .



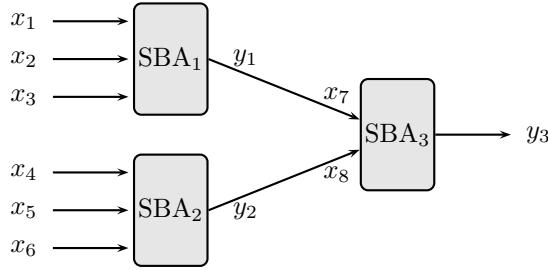


Figura 13.4: Red de 3 SBA

Nótese que el soporte de x_2^{nec} no está contenido en el intervalo $[0, 1]$, que es el universo de discurso de x_2 . Para corregir esta situación utilizamos la ecuación 13.3 sobre x_2^{nec} y posteriormente aplicamos la extensión directa para obtener una nueva versión de y^{nec} . Los resultados se muestran en la Figura 13.3.

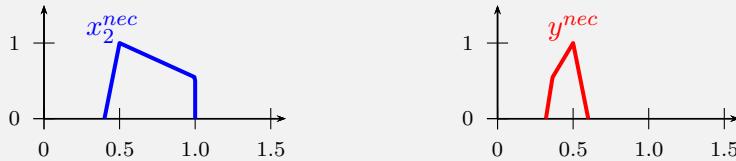


Figura 13.3: Variables en el ejemplo 13.1

13.2. Redes de SBA

Consideremos ahora el caso que se representa en la Figura 13.4. El sistema SBA₃ recibe dos entradas (x_7, x_8) y genera una salida (y_3). Lo interesante es que sus entradas provienen, a su vez, de otros dos SBA: x_7 proviene de la salida de SBA₁ mientras que x_8 proviene del SBA₂.

¿Es eso posible? ¿Podemos interconectar directamente distintos Sistemas Basados en Aritmética difusa. En otras palabras y usando el ejemplo de la Figura 13.4, ¿es correcto escribir $x_7 = y_1$ y $x_8 = y_2$, como se insinúa allí?

La respuesta la tenemos en las tablas 12.2 y 12.3, que resumen el tipo de entradas válidas y de salidas disponibles en estos sistemas. Nótese que en ambas tablas aparece el tipo ‘número difuso’ y, por tanto, la respuesta es afirmativa.

Ahora bien, recordemos que en un SBA tanto las entradas como la salida tienen una

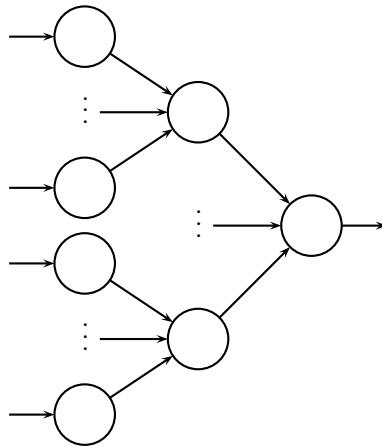


Figura 13.5: Red general de SBA

variable lingüística asociada. Por esa razón, podemos establecer las dos condiciones necesarias para conectar la salida y_a de un cierto SBA a la entrada x_b de otro:

- Las dos variables lingüísticas deben ser iguales:

$$\mathcal{V}_a = \mathcal{V}_b \quad (13.8)$$

- y_a , es decir, el resultado de la aproximación lingüística empleada en el primer sistema, debe ser del tipo ‘número difuso’.

La segunda condición no obsta para que se pueda realizar otro tipo de aproximaciones lingüísticas en el primer sistema, siempre y cuando no se utilicen como entradas a otros sistemas. Por ejemplo, pueden usarse para brindar información lingüística a un usuario interesado en conocer los resultados intermedios.

Podemos generalizar las ideas anteriores. Una *Red de Sistemas de Computación con Palabras basadas en aritmética difusa* es una red como la que se muestra en la Figura 13.5, en la que cada nodo es un SBA y las conexiones entre nodos se efectúan mediante números difusos que comparten las mismas variables lingüísticas. El número de capas y conexiones es arbitrario, siempre y cuando no se cierren lazos, es decir, siempre y cuando la red sea del tipo ‘hacia adelante’.

Ejemplo 13.2 (Red de SBA) En el Ejemplo 12.3 se menciona la metodología EVIAVE (Evaluación de Impacto Ambiental de Vertederos), diseñada específicamente para diagnos-

ticar ambientalmente los vertederos de residuos sólidos urbanos y establecer su grado de interacción y afectación al medio ambiente [31, 58, 2, 16, 20].

La metodología evalúa 36 variables que describen la situación del vertedero. La evaluación se realiza considerando cinco elementos ambientales: 1) Aguas superficiales 2) Aguas subterráneas 3) Atmósfera 4) Suelo y 5) Salud.

Para cada uno de esos elementos, la metodología estima su valor ambiental (*VA*) y la probabilidad de contaminación (*PBC*). Con esta información calcula luego el Índice de riesgo de afectación ambiental (*IRA*) de ese elemento.

Finalmente, se calcula el Índice de Interacción Medio-Vertedero *IMV* a partir de los *IRA* de los cinco elementos. Este índice busca “cuantificar de manera conjunta las diferentes afecciones a cada elemento del medio, con el objetivo de ser representativo del estado ambiental global existente entre la interacción del punto de vertido y el entorno ambiental.”

La Figura 13.6 es la representación gráfica de una red de SBA que se ha construido para implementar EVIAVE. El nodo superior (*IMV*) recibe los resultados de los cinco nodos inferiores, que corresponden a los *IRA* de cada uno de los cinco elementos ambientales.

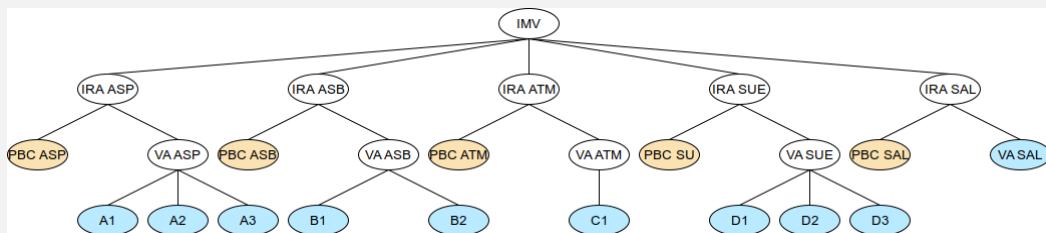
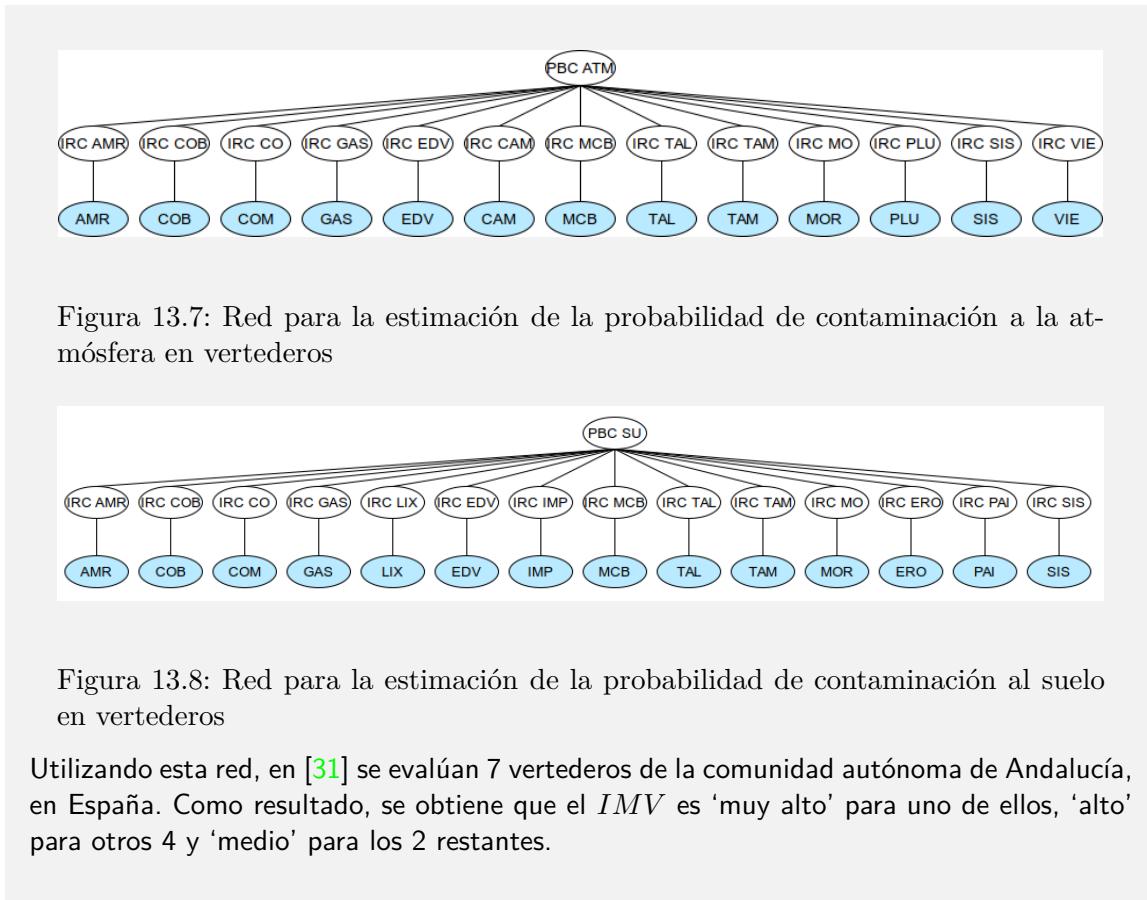


Figura 13.6: Red para la estimación del impacto ambiental de vertederos

Cada uno de estos, a su vez, reciben información de 2 nodos, uno correspondiente a su *PBC* y otro a su *VA*. Los nodos destinados al cálculo de los *VA* reciben como entradas unos nodos en color azul, que son entradas a toda la red.

Por su parte, los nodos destinados el cálculo de las *PBC* (en color sepia) son en realidad el resultado de otras redes de SBA. Las Figuras 13.7 y 13.8 muestran dos de ellas, las correspondientes a los elementos atmósfera y suelo. Nótese que estas dos redes comparten varios nodos de entrada (en color azul), como por ejemplo AMR (Condición para la variable Asentamiento de la Masa de Residuos) y COB (Condición para la variable Cobertura vegetal)



Bibliografía

- [1] Muhammad Akram and Ayesha Bashir. Complex fuzzy ordered weighted quadratic averaging operators. *Granular Computing*, 6(3):523–538, 2021.
- [2] Gabriela Arrieta, Ignacio Requena, Javier Toro, and Montserrat Zamorano. Adaptation of eviave methodology for monitoring and follow-up when evaluating the environmental impact of landfills. *Environmental Impact Assessment Review*, 56:168–179, 2016.
- [3] Michał Baczyński, Gleb Beliakov, Humberto Sola, and Ana Pradera. *Advances in Fuzzy Implication Functions*, volume 300. 01 2013.
- [4] A. Ban, A. Brândăş, L. Coroianu, C. Negruiu, and O. Nica. Approximations of fuzzy numbers by trapezoidal fuzzy numbers preserving the ambiguity and value. *Computers & Mathematics with Applications*, 61(5):1379–1401, 2011.
- [5] Michael R. Berthold, Christian Borgelt, Frank Höppner, Frank Klawonn, and Rossaria Silipo. *Guide to Intelligent Data Science. How to Intelligently Make Use of Real Data*. Springer, Cham, 2020.
- [6] James C. Bezdek. *Fuzzy mathematics in pattern classification*. PhD thesis, University of Cornell, 1973.
- [7] James C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Pres, 1981.
- [8] Bernadette Bouchon-Meunier, Olga Kosheleva, Vladik Kreinovich, and Hung T. Nguyen. Fuzzy numbers are the only fuzzy sets that keep invertible operations invertible. *Fuzzy Sets and Systems*, 91(2):155–163, 1997. Fuzzy Arithmetic.
- [9] J.J. Buckley. Universal fuzzy controllers. *Automatica*, 28(6):1245–1248, 1992.
- [10] Christer Carlsson and Robert Fuller. On fuzzy internal rate of return. Technical report, 1998.

- [11] Christer Carlsson and Robert Fuller. Capital budgeting problems with fuzzy cash flows. *Mathware and Soft Computing*, 6(1):81–89, 1999.
- [12] J.L. Castro. Fuzzy logic controllers are universal approximators. *IEEE Transactions on Systems, Man, and Cybernetics*, 25(4):629–635, 1995.
- [13] J.L. Castro and M. Delgado. Fuzzy systems with defuzzification are universal approximators. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(1):149–152, 1996.
- [14] M. Delgado, O. Duarte, and I. Requena. An arithmetic approach for the computing with words paradigm. *International Journal of Intelligent Systems*, 21(2):121–142, 2006.
- [15] M. Delgado, M.A. Vila, and W. Voxman. On a canonical representation of fuzzy numbers. *Fuzzy Sets and Systems*, 93(1):125–135, 1998.
- [16] Miguel Delgado, Oscar G. Duarte, and Ignacio Requena. A fuzzy methodology for the environmental impact evaluation of garbage dumps. In *ISFA-2005 World Congress Beijing*, 2005.
- [17] M. Detyniecki. *Fundamentals on Aggregation Operators*. Computer Science Division, University of California, Berkeley, 2001.
- [18] Adenso Diaz and Fan T. Tseng. Introducción a las técnicas heurísticas. In Adenso Diaz, editor, *Optimización Heurística Y Redes Neuronales*, chapter 1, pages 19–36. editorial paraninfo, Madrid, 1996.
- [19] Oscar Duarte. New approach to obtain trapezoidal approximations of fuzzy numbers. *Fuzzy sets ans systems*, submitted, 2025.
- [20] Oscar Duarte, Ignacio Requena, and Monserrat Zamorano. Environmental diagnosis of municipal waste landfill using intelligent. In *Proceedings of SECOTOX Conference and the Internacional Conference on Environment Engineering, Planning and Economics*, Skiathos (Grecia), 2007. ISBN 978-960-89818-1-2.
- [21] Oscar Duarte and Sandra Téllez. A family of OWA operators based on Faulhaber's formulas. <https://arxiv.org/abs/1801.10545>, 2018.
- [22] Oscar G. Duarte. Unfuzzy-software para el análisis, diseño, simulación e implementación de sistemas de lógica difusa. Master's thesis, Universidad Nacional de Colombia. Maestría en Automatización Industrial., 1997.
- [23] Oscar G. Duarte. *Técnicas difusas en la evaluación de impacto ambiental*. PhD thesis, Universidad de Granada, 2000.

- [24] Oscar G. Duarte, Miguel Delgado, and Ignacio Requena. Algorithms to extend crisp functions and their inverse functions to fuzzy numbers. *International journal of intelligent systems*, 18(8):855–876, 2003.
- [25] Didier Dubois and Henri Prade. Operations on fuzzy numbers. *International Journal of Systems Science*, 9:613–626, 1978.
- [26] Conesa Fdez-Vitoria. *Guía metodologica para la evaluación del impacto ambiental*. Ediciones mundi-prensa,, Madrid:, 3 ed. edition, 1996.
- [27] Dimitar Filev and Ronald R. Yager. On the issue of obtaining owa operator weights. *Fuzzy Sets and System*, 94:157–169, 1998.
- [28] Robert Fullér and Péter Majlender. An analytic approach for obtaining maximal entropy owa operator weights. Technical report, Turku Centre for Computer Science, 2000.
- [29] Luis E Gallego, Oscar Duarte, Horacio Torres, Mauricio Vargas, Johny Montaña, Ernesto Pérez, Javier Herrera, and Camilo Younes. Lightning risk assessment using fuzzy logic. *Journal of Electrostatics*, 60(2):233–239, 2004. Selected Papers from the 26th International Conference on Lightning Protection.
- [30] Luis Eduardo Gallego. Metodología para el análisis de riesgo de daño por rayos en colombia utilizando técnicas difusas. Master's thesis, Universidad Nacional de Colombia, 2002.
- [31] María Encarnación Garrido Vegara. *Metodología de diagnóstico ambiental de vertederos, adaptación para su informatización utilizando técnicas difusas y su aplicación en vertederos de Andalucía*. PhD thesis, Universidad de Granada, 2008.
- [32] Przemysław Grzegorzewski. Metrics and orders in space of fuzzy numbers. *Fuzzy Sets and Systems*, 97(1):83–94, 1998.
- [33] Y. Hijazi, H. Hagen, C. D. Hansen, and K. I. Joy. Why interval arithmetic is so useful. Visualization of large and unstructured data sets, 2008.
- [34] J.H. Holland. *Adaptation in Natural and Artificial Systems*. The MIT Press, London, 1975.
- [35] Ronald Howard and Ali E. Abbas. *Foundations of Decision Analysis*. Pearson Education, 2016.
- [36] IEEE. Ieee standard for interval arithmetic. *IEEE Std 1788-2015*, pages 1–97, 2015.

- [37] Manuel Laguna and Pablo Moscato. Moscatto, pablo. In Adenso Diaz, editor, *Optimización Heurística Y Redes Neuronales*, pages 67–104. Editorial Paraninfo, Madrid, 1996.
- [38] María Teresa Lamata and Elio Cables Perez. Owa weights determination by means of linear functions. *Mathware and soft computing*, 16(2), 2009.
- [39] Xinwang Liu. *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice*, volume 265 of *Studies in Fuzziness and Soft Computing*, chapter A Review of the OWA Determination Methods: Classification and Some Extensions, pages 60–90. Springer, 2011.
- [40] Ebrahim H Mamdani. Application of fuzzy algorithms for control of simple dynamic plant. In *Proceedings of the institution of electrical engineers*, volume 121, pages 1585–1588. IET, 1974.
- [41] José.J Martínez and Sergio A.. Rojas. *Introducción a la Informática Evolutiva. Un Nuevo Enfoque Para Resolver Problemas de Ingeniería*. Facultad de Ingeniería. Universidad Nacional de Colombia, 1999.
- [42] J.M. Mendel. Fuzzy logic systems for engineering: a tutorial. *Proceedings of the IEEE*, 83(3):345–377, 1995.
- [43] K Menger. Statistical metrics. *Proceedings of the National Academy of Sciences of the United States of America*, 28(12):535–537, 1942.
- [44] Ramon E. Moore. *Interval Analysis*. Prentice Hall, Englewood Cliffs, NJ, 1966.
- [45] Ramon E. Moore, R. Baker Kearfott, and Michael J. Cloud. *Introduction to Interval Analysis*. Society for Industrial and Applied Mathematics, 2009.
- [46] Erika S. Olaya, Carlos J. Cortés, and Oscar G. Duarte. Despliegue de función de calidad (qfd) apoyado mediante técnicas difusas: Caso protésis mioeléctrica de mano. *Ingeniería e Investigación*, 25(2):4–14, 2005.
- [47] Navid Razmjooy. *Introduction to Interval Analysis and Solving the Problems with Interval Uncertainties*, pages 29–90. 2024.
- [48] Nathalie Revol. Introduction to the ieee 1788-2015 standard for interval arithmetic. In Alessandro Abate and Sylvie Boldo, editors, *Numerical Software Verification*, pages 14–21, Cham, 2017. Springer International Publishing.
- [49] Yu-Ru Syau, E.S. Lee, and Lixing Jia. Convexity and upper semicontinuity of fuzzy sets. *Computers & Mathematics with Applications*, 48(1):117–129, 2004.
- [50] Peter van Overschee and B.L. de Moor. *Subspace Identification for Linear Systems*. Springer, 1996.

- [51] L.-X. Wang. Fuzzy systems are universal approximators. In *[1992 Proceedings IEEE International Conference on Fuzzy Systems*, pages 1163–1170, 1992.
- [52] L.-X. Wang and J.M. Mendel. Generating fuzzy rules by learning from examples. In *Proceedings of the 1991 IEEE International Symposium on Intelligent Control*, pages 263–268, 1991.
- [53] Li-Xin Wang. *Adaptative Fuzzy Systems and Control. Design and Stability Analysis*. Prentice Hall, New Jersey, 1995.
- [54] Rui Xu and D. Wunsch. Survey of clustering algorithms. *IEEE Transactions on Neural Networks*, 16(3):645–678, 2005.
- [55] Ronald R. Yager. On ordered weighted averaging aggregation operators in multicriteria decision making. *IEEE Transactions on Systems Man and Cybernetics*, 18(1):183–190, 1988.
- [56] Lotfi A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.
- [57] Lotfi A. Zadeh. Outline of a new approach to the analysis of complex systems and decision processes. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(1):28–44, 1973.
- [58] Montserrat Zamorano, Emilio Molero, Álvaro Hurtado, Alejandro Grindlay, and Ángel Ramos. Evaluation of a municipal landfill site in southern spain with gis-aided methodology. *Journal of Hazardous Materials*, 160(2):473–481, 2008.

Apéndice A

Obtención de los pesos de un operador OWA

A.1. Orness y entropía

Supongamos un operador $\text{owa}(\cdot)$ con p entradas y un vector de pesos \mathbf{w} . Yager define el orness $\text{orness}(\mathbf{w})$ y la dispersión o entropía de ese operador así [55]:

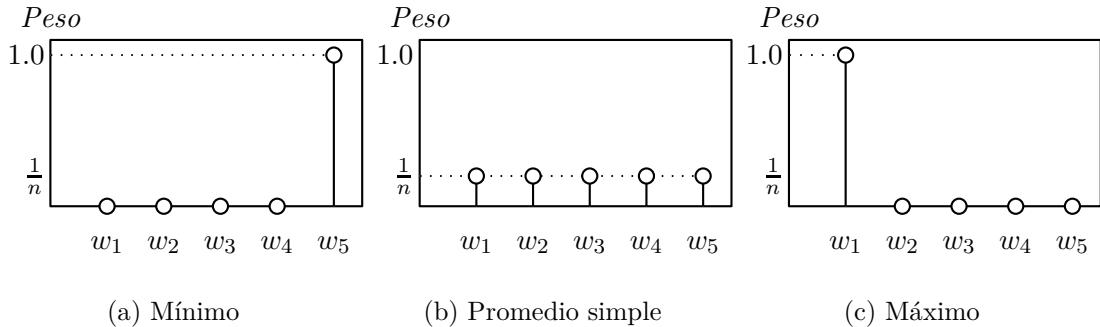
$$\text{orness}(\mathbf{w}) = \frac{1}{p-1} \sum_{i=1}^p (p-i)w_i \quad (\text{A.1})$$

$$\text{disp}(\mathbf{w}) = - \sum_{i=1}^p w_i \ln w_i \quad (\text{A.2})$$

El *orness* se refiere al grado de semejanza de la agregación efectuada por el operador, respecto a la que haría un operador del tipo *OR*. Por su parte, la dispersion el grado en el que \mathbf{w} considera toda la información de la agregación.

A.2. Una familia de operadores OWA

Consideremos ahora el problema de encontrar los pesos de un operador OWA con un *orness* deseado. El operador debe agregar la información de p entradas. Definimos $\nu \in [0, 1]$ como una variable que es igual al *orness* deseado. Sin embargo, en una perspectiva más general, $\nu = g(\text{orness}) \in [0, 1]$ en donde $g(\cdot)$ es una función monótonamente creciente con $g(0) = 0$ and $g(1) = 1$.

Figura A.1: Pesos para 3 OWA con $p = 5$

Imponemos además las siguientes condiciones con el propósito de construir una familia de operadores cuyo comportamiento sea fácil de entender:

1. Cuando $\nu = 0.0$ el operador es equivalente al operador mínimo.
2. Cuando $\nu = 0.5$ el operador es equivalente al operador promedio simple.
3. Cuando $\nu = 1.0$ el operador es equivalente al operador máximo.
4. Al aumentar ν la agregación varía suavemente.
5. El comportamiento del operador debe ser simétrico respecto a $\nu = 0.5$. En otras palabras, la forma en que el operador evoluciona desde el mínimo hasta el promedio simple al variar ν desde 0.0 hasta 0.5 debe ser la misma forma en que evoluciona desde el operador máximo hasta el promedio simple al variar ν desde 1.0 hasta 0.5.

La última condición se cumple fácilmente si se construye una familia para $0 \leq \nu \leq 0.5$ y luego se utiliza la simetría para definir el comportamiento para $0.5 < \nu \leq 1.0$.

A.2.1. Operadores del tipo AND ($0 \leq \nu \leq 0.5$)

Sea ν el *orness* deseado, y consideremos primero el caso en que $0 \leq \nu \leq 0.5$. La lógica del método que aquí se describe es la siguiente: cuando ν varía desde 0 hasta 0.5 los pesos deben variar desde aquellos que se muestran en la Figura A.1a hasta aquellos que se muestran en la Figura A.1b; para lograrlo, primero variaremos w_p de forma monótona desde 1.0 hasta $1/p$. La cantidad en que se ha disminuido w_p se distribuye entre los demás pesos para mantener la suma total de pesos constante. Esta distribución se realiza usando una línea recta cuyos parámetros son fáciles de calcular, tal como se muestra a continuación.

Definimos $f(x) : [0, 0.5] \rightarrow [0, 1]$ una función monótonamente creciente con $f(0) = 0$

y $f(0.5) = 1$. Más adelante se encuentran otras condiciones que debe cumplir $f(x)$. El valor de w_p debe variar de 1 hasta $1/p$ cuando ν varía desde 0 hasta 0.5. Usamos para ello $f(\cdot)$:

$$w_p = 1 - \Delta \quad \Delta = f(\nu)(1 - 1/p) = f(\nu)(p - 1)/p \quad (\text{A.3})$$

Bajo estas condiciones, la suma de los otros pesos debe ser

$$\sum_{i=1}^{p-1} w_i = \sum_{i=1}^m w_i = \Delta \quad m = p - 1 \quad (\text{A.4})$$

De acuerdo con la ecuación A.1 el *orness* será

$$\text{orness}(\mathbf{w}) = \frac{1}{p-1} \sum_{i=1}^p (p-i)w_i = \frac{1}{m} \sum_{i=1}^m (p-i)w_i = \nu \quad (\text{A.5})$$

La distribución de Δ se realiza siguiendo la ecuación de una línea recta:

$$w_i = Ki + b \quad i = 1, 2, \dots, m \quad (\text{A.6})$$

Las ecuaciones A.4 y A.5 se convierten en

$$\begin{cases} \sum_{i=1}^m (Ki + b) = \Delta \\ \frac{1}{m} \sum_{i=1}^m (p-i)(Ki + b) = \nu \end{cases} \quad (\text{A.7})$$

Que pueden reorganizarse como

$$\begin{cases} K \sum_{i=1}^m i + \sum_{i=1}^m b = \Delta \\ \frac{K}{m} \left(p \sum_{i=1}^m i - \sum_{i=1}^m i^2 \right) + \frac{b}{m} \left(\sum_{i=1}^m p - \sum_{i=1}^m i \right) = \nu \end{cases} \quad (\text{A.8})$$

Las sumas se pueden calcular usando las fórmulas de Faulhaber:

$$\sum_{i=1}^m i = \frac{m(m+1)}{2} \quad \sum_{i=1}^m i^2 = \frac{m(m+1)(2m+1)}{6} \quad (\text{A.9})$$

Para obtener

$$\begin{cases} K \frac{m(m+1)}{2} + mb = \Delta \\ \frac{K}{m} \left(\frac{m(m+1)^2}{2} - \frac{m(m+1)(2m+1)}{6} \right) + \frac{b}{m} \left(m(m+1) - \frac{m(m+1)}{2} \right) = \nu \end{cases} \quad (\text{A.10})$$

Reorganizamos los términos para mostrar de forma explícita que hemos obtenido un sistema de dos ecuaciones lineales con dos incógnitas:

$$\begin{cases} K \frac{m(m+1)}{2} + bm = \Delta \\ K \frac{(m+1)(m+2)}{6} + b \frac{m+1}{2} = \nu \end{cases} \quad (\text{A.11})$$

A partir de A.11 se obtiene expresiones directas para el cálculo de K y b

$$\begin{cases} K = 6 \frac{\Delta n/m - 2\alpha}{m^2 - 1} \\ b = \frac{\Delta}{m} - K \frac{(m+1)}{2} \end{cases} \quad (\text{A.12})$$

Usando la ecuación A.3 se pueden calcular K y b a partir de ν y m (ó de p)

$$\begin{cases} K = 6 \frac{f(\nu) - 2\nu}{m^2 - 1} = 6 \frac{f(\nu) - 2\nu}{p(p-2)} \\ b = \frac{f(\nu)}{m+1} - K \frac{(m+1)}{2} = \frac{f(\nu)}{p} - K \frac{p}{2} \end{cases} \quad (\text{A.13})$$

A.2.2. Selección de $f(\cdot)$

Como todos los pesos deben ser no-negativos, estudiemos ahora esa condición para el primer peso $w_1 = K \times 1 + b$

$$\begin{aligned}
K + b &\geq 0 \\
K + \frac{\Delta}{m} - K^{\frac{m+1}{2}} &\geq 0 \\
K^{\frac{1-m}{2}} + \frac{\Delta}{m} &\geq 0 \\
K^{\frac{1-m}{2}} &\geq -\frac{\Delta}{m} \\
3(1-m) \frac{\Delta p/m - 2\nu}{m^2-1} &\geq -\frac{\Delta}{m} \\
-\frac{3}{p}(\Delta \frac{p}{m} - 2\nu) &\geq -\frac{\Delta}{m} \\
2m\nu - p\Delta &\geq -p\Delta/3 \\
2m\nu &\geq 2p\Delta/3 \\
3\nu &\geq \Delta p/m \\
3\nu &\geq f(\nu)
\end{aligned} \tag{A.14}$$

Para asegurar que los demás pesos sean no negativos, imponemos la condición $K \geq 0$. Al considerar la ecuación A.13 se obtiene directamente que $f(\nu) \geq 2\nu$.

Hemos encontrado dos restricciones para $f(\nu)$:

$$2\nu \leq f(\nu) \leq 3\nu \tag{A.15}$$

Si se escoge $f(\nu) = 1 - (1 - 2\nu)^\beta$ para algún $0 \leq \beta \leq 1$ (Figura A.2), las desigualdades

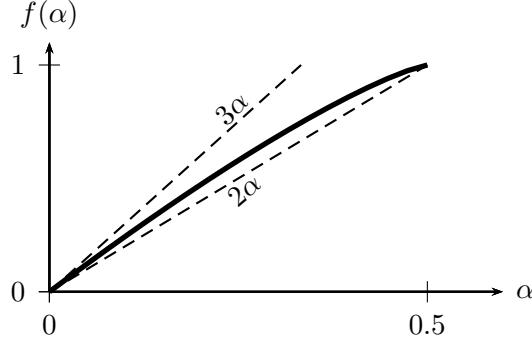


Figura A.2: Función $f(\nu) = 1 - (2\nu - 1)^\beta$ y sus restricciones

de A.15 se satisfacen si la pendiente en 0 está en el intervalo $[2, 3]$

$$\begin{aligned}
 2 &\leq \frac{df(\nu)}{d\nu} \Big|_{\nu=0} \leq 3 \\
 2 &\leq 2\beta(1 - 2\nu)^{\beta-1} \Big|_{\nu=0} \leq 3 \\
 2 &\leq 2\beta \leq 3 \\
 1.0 &\leq \beta \leq 1.5
 \end{aligned} \tag{A.16}$$

Nótese que en el caso límite en que $\beta = 1.0$ se obtiene $f(\nu) = 2\nu$. Este caso lleva a $K = 0$ and $b = 1/p$ lo que significa que Δ se distribuye homogáneamente entre los pesos w_1, \dots, w_{p-1} .

A.2.3. Operadores tipo OR ($0.5 \leq \alpha \leq 1$)

Para asegurar el comportamiento simétrico de la familia los resultados anteriores deben adaptarse así:

- Calcular los pesos para $\bar{\nu} = 1 - \nu$
- Invertir el orden de los pesos obtenidos.

Las ecuaciones A.3, A.4 y A.13, junto con esta adaptación son las que se implementan en el algoritmo 13.

Apéndice B

Introducción a los algoritmos genéticos

Este apéndice resume las ideas básicas de los AG en la sección B.1; una presentación más extensa puede encontrarse en [34], [37] y [41]. En la sección B.2 se destaca la importancia crucial que tiene para los AG el problema de codificación de la información, mientras que en la sección B.3 se propone una estrategia de codificación.

B.1. El algoritmo básico

El objetivo de un Algoritmo Genético es la optimización de una *Función de Evaluación* (también denominada *fitness* o *Función objetivo*), sujeta a ciertas restricciones. Para ello, las variables de la función se codifican en *genes*; una colección de genes, que corresponde a un punto en el espacio de entrada de la función de evaluación, es un *individuo*. En un AG se produce una colección de individuos (una *población*) que a través de un cierto ciclo de iteraciones se va mejorando. Cada iteración corresponde a una nueva *generación*.

La estrategia básica de los AG se resume en el siguiente algoritmo ($P(t)$ es la población en la generación t):

1. $t = 0$
2. Crear $P(t)$
3. Hacer hasta que se cumpla la condición de parada
 - **Evaluuar** la $P(t)$
 - **Seleccionar** $P(t + 1)$ a partir de $P(t)$

- Recombinar $P(t + 1)$
- $t = t + 1$
- 4. Fin del ciclo

La creación de $P(0)$ generalmente se hace de forma aleatoria. La **Evaluación** de $P(t)$ consiste en calcular el valor de la Función de Evaluación para cada individuo de la población.

El proceso de **Selección** consta de dos etapas:

1. Asignar una probabilidad de supervivencia a cada individuo basada en la función de evaluación, para ello se emplea un **Operador de Probabilidad**
2. Ejecutar algún algoritmo de muestreo (de selección) para decidir cuántas copias de cada individuo se crean en la siguiente generación; se emplea un **Operador de Selección**

El proceso de **Recombinación** consta de dos etapas:

1. El proceso de **Cruce**: Es la generación de nuevos individuos (hijos) a partir de individuos de la actual generación (padres). Para efectuar este proceso se necesita:
 - Asignar parejas con un **Operador de Asignación de Parejas**.
 - Decidir si cada pareja se cruza y cuántos individuos genera con un **Operador de Reproducción**.
 - Cruzar los individuos con un **Operador de Cruce**
2. El proceso de Mutación: Es la alteración aleatoria de alguna de las características (genes) de un individuo mediante un **Operador de Mutación**.

Adicionalmente al ciclo básico, existe la posibilidad de adaptar el algoritmo genético modificando alguna de sus características conforme avanza el algoritmo. Para esto se requiere algún **Operador de adaptación**. Es usual seguir la estrategia de **Elitismo**, según la cual el mejor individuo de una generación siempre pasa a la siguiente.

De acuerdo a lo anterior, un algoritmo genético requiere los operadores que se resumen en la tabla B.1. Para cada tipo de operador existe una amplia gama de posibilidades en la literatura, lo cual da origen a múltiples versiones del algoritmo básico. Algunos operadores dependen del tipo de codificación genética empleada (ver sección B.2); a continuación se muestran algunos de los operadores que no dependen del tipo de codificación.

El desempeño de un AG suele analizarse siguiendo la función de evaluación del mejor de los individuos de cada generación; sin embargo, también pueden seguirse la del peor

Tabla B.1: Operadores de los AG

Tipo de Operador	observaciones
Operador de probabilidad	no depende de la codificación
Operador de selección	no depende de la codificación
Operador de asignación de parejas	no depende de la codificación
Operador de reproducción	no depende de la codificación
Operador de cruce	si depende de la codificación
Operador de mutación	si depende de la codificación
Operador de adaptación	es opcional y depende de los otros operadores seleccionados

individuo, el promedio de las funciones de evaluación de todos los individuos y/o su desviación estándar. Existen también otras medidas de desempeño como las siguientes:

Medida Online: Promedio de las funciones de evaluación de todos los individuos que han aparecido hasta ese momento en todas las generaciones.

Medida Offline: Promedio de las funciones de evaluación de los mejores individuos de cada generación hasta ese momento.

B.1.1. Operadores de Probabilidad

El operador de probabilidad asigna una *probabilidad de supervivencia* $p(i)$ a los N individuos de una cierta generación. Un individuo con una buena función de evaluación $f(i)$ deberá tener una mayor probabilidad de supervivencia que otro con una peor función de evaluación. Estos son algunos de los operadores más usuales:

Ranking Lineal: Se ordena la población de mejor individuo a peor. La probabilidad de supervivencia depende de su posición i según la siguiente ecuación (n_{min} es un parámetro seleccionable por el usuario):

$$n_{max} = 2 - n_{min}$$

$$p(i) = (n_{max} - (n_{max} - n_{min}) * ((i)/(N - 1)))/N$$

Proporcional: La probabilidad de supervivencia depende de la función de evaluación (mejor función significa mayor probabilidad).

$$\text{Para maximizar: } p(i) = \frac{f(i)}{\sum_{j=1}^N f(j)}$$

$$\text{Para minimizar: } p(i) = 1.0 - \frac{f(i)}{\sum_{j=1}^N f(j)}$$

Homogénea: A todos los individuos se les asigna la misma probabilidad de supervivencia $\frac{1}{N}$

B.1.2. Operadores de Selección

El Operador de Selección establece cuántas copias se crearán de cada individuo en la siguiente generación. Suele establecerse que el número de individuos en todas las generaciones se mantenga constante. Estos son algunos de los operadores más usuales:

Estocástica con Remplazo: se crea un ‘ruleta’ con tantas casillas como individuos tenga la población; el ángulo de cada casilla es proporcional a la probabilidad de supervivencia. Se juega a la ruleta tantas veces como individuos tenga la población y se crea una copia de cada individuo ganador.

Estocástica Universal: se crea una ruleta como en el caso anterior, pero sólo se juega una vez; las casillas ganadoras están igualmente espaciadas, y hay tantas como individuos tenga la población.

B.1.3. Operadores de Asignación de Parejas

El Operador de Asignación de Parejas decide qué individuos se cruzan entre sí para generar nuevos individuos. Estos son algunos de los operadores más usuales:

Aleatoria: se buscan aleatoriamente las parejas.

Siguiente: la pareja de cada individuo es la siguiente en el arreglo de individuos que forma la población.

Extremos: al primer individuo le corresponde el último, al segundo el penúltimo, etc.

B.1.4. Operadores de Reproducción

El Operador de Reproducción establece una estrategia general de reproducción entre parejas que se cruzan. Estos son algunos de los operadores más usuales:

Dos padres dos hijos: por cada pareja se crean dos hijos que remplazan a sus padres.

Mejor padre-mejor hijo: por cada pareja se crean dos hijos, pero sólo el mejor hijo remplaza al peor parente

Mejores entre padres e hijos: por cada pareja se crean dos hijos; de los cuatro se mantienen los dos mejores.

B.2. Codificación del problema

Tal como se comentó en la sección B.1, para optimizar una función mediante AG es necesario codificar las variables de dicha función en *genes*. No existe una única forma de codificar dichas variables; en esta sección se muestran dos de las estrategias más comunes para ese proceso, la *codificación binaria* y la *codificación real*, si bien en la literatura se encuentran aplicaciones con codificación entera, en listas y arreglos, etc.

La elección de uno u otro tipo de codificación puede afectar el desempeño de un AG, principalmente por que: 1) los AG manipulan una representación codificada del problema y 2) la representación puede limitar en gran medida la visión que tiene un sistema sobre su entorno.

Por otra parte, ciertas codificaciones pueden acomodarse más fácilmente para representar ciertos tipos de variables, y por tanto pueden resultar en una implementación más rápida y efectiva del AG. Es necesario resaltar que los operadores de cruce y mutación dependen del tipo de codificación empleada.

B.2.1. Codificación Binaria

Históricamente, esta fue la primera estrategia de representación empleada. Se busca representar el espacio de búsqueda $E = E_1 \times E_2 \times \dots \times E_n$ mediante un alfabeto binario. Para ello, se codifica cada variable del espacio de búsqueda mediante una función $cod_i : E_i \rightarrow \{0, 1\}^{L_i}$ con $L_i \in N$ y $i = 1, 2, \dots, n$; es decir, cada variable del espacio de búsqueda, E_i se representa mediante una cadena de unos y ceros de longitud L_i .

La representación de un elemento cualquiera $x = (x_1, x_2, \dots, x_n) \in E$, tal que $x_i \in E_i$ se lleva a cabo mediante la concatenación de las n cadenas $cod(x) = cod_1(x_1)cod_2(x_2)\dots cod_n(x_n)$. Esta nueva cadena de L casillas se denomina el *genoma*¹, y cada una de las casillas contiene un *gen*, que será un uno o un cero.

B.2.1.1. Cruce Binario

El operador de cruce es el operador que permite que dos individuos compartan su información genética, para producir nuevos individuos. Estos son algunos de los operadores de cruce con codificación binaria más usuales:

Cruce simple: Dados dos individuos (*padres*) con genomas $P_1 = (c_1^1, \dots, c_L^1)$ y $P_2 = (c_1^2, \dots, c_L^2)$, respectivamente, se generan dos nuevos individuos (*hijos*) con genomas $H_1 = (c_1^1, \dots, c_i^1, c_{i+1}^2, \dots, c_L^2)$ y $H_2 = (c_1^2, \dots, c_i^2, c_{i+1}^1, \dots, c_L^1)$, donde i (denominado el *punto de cruce*) es un número entero aleatorio del intervalo $[1, L - 1]$.

¹algunos autores emplean el término *cromosoma*

Cruce doble: Similar al anterior, pero con dos puntos de cruce aleatorios i, j . Los genomas de los hijos serán $H_1 = (c_1^1, \dots, c_i^1, c_{i+1}^2, \dots, c_j^2, c_{j+1}^1, \dots, c_L^1)$ y $H_2 = (c_1^2, \dots, c_i^2, c_{i+1}^1, \dots, c_j^1, c_{j+1}^2, \dots, c_L^2)$.

B.2.1.2. Mutación Binaria

El operador de mutación efectúa alteraciones aleatorias en la información genética. Con esta función se busca explorar nuevas zonas del espacio de búsqueda, y así intentar escapar de posibles óptimos locales en los que pueda quedar atrapado el algoritmo. Estos son algunos de los operadores de mutación con codificación binaria más usuales:

Uniforme: Para cada se genera un número aleatorio $m \in [0, 1]$. Si $m \leq p_m$ se altera el valor del gen (si es un 1 cambia a 0, y viceversa). p_m es un número del intervalo $[0, 1]$ denominado *probabilidad de mutación*, seleccionado por el usuario y constante en toda la ejecución del algoritmo.

No Uniforme: Similar al anterior, pero la probabilidad de mutación disminuye conforme avanza la ejecución del algoritmo, usualmente siguiendo una curva exponencial $p_m = p_{max} e^{-b\frac{t}{T}}$, donde t es el número de la generación, T el número máximo de generaciones, y p_{max}, b dos parámetros seleccionados por el usuario ($p_{max} \in [0, 1]$ y $b > 0$).

B.2.2. Codificación Real

En la codificación real se busca representar el espacio de búsqueda $E = E_1 \times E_2 \times \dots \times E_n$ mediante un arreglo de números reales, usando una función $cod : E \rightarrow R^n$. Es usual emplear n funciones de codificación, una para cada una de las variables del espacio de búsqueda: $cod_i : E_i \rightarrow R$; en estos casos la representación de un elemento cualquiera $x = (x_1, x_2, \dots, x_n) \in E$, tal que $x_i \in E_i$ se lleva a cabo mediante la concatenación de los n reales $cod(x) = [cod_1(x_1)cod_2(x_2) \dots cod_n(x_n)]$. Este arreglo de n elementos es el *genoma*, y cada uno de los elementos del arreglo es un *gen*, que será un número real.

B.2.2.1. Cruce Real

La variedad de operadores de cruce para codificación real es mucho mayor que para la codificación binaria. Algunos de los operadores más empleados se han listado a continuación. En todos los casos se ha supuesto que los dos individuos $C_1 = \{c_1^1, c_2^1, \dots, c_n^1\}$ y $C_2 = \{c_1^2, c_2^2, \dots, c_n^2\}$ se desean cruzar para generar m hijos $H_k = \{h_1^k, h_2^k, \dots, h_n^k\}$ con $k = 1, 2, \dots, m$. Cada gen h_i^k debe pertenecer al intervalo $[a_i, b_i]$:

Plano: Se genera un hijo ($m = 1$). h_i^k se genera aleatoriamente en el intervalo $[c_i^1, c_i^2]$.

Cruce simple: Una posición j se escoge aleatoriamente ($j < n - 1$). Se generan dos

hijos así:

$$\begin{aligned} H_1 &= \{c_1^1, c_2^1, \dots, c_j^1, c_{j+1}^2, \dots, c_n^2\} \\ H_2 &= \{c_1^2, c_2^1, \dots, c_j^2, c_{j+1}^1, \dots, c_n^1\} \end{aligned}$$

Cruce Aritmético: Se generan dos hijos ($\lambda \in [0, 1]$ es un parámetro seleccionable por el usuario); cada gen se obtiene así:

$$\begin{aligned} h_i^1 &= \lambda c_i^1 + (1 - \lambda) c_i^2 \\ h_i^2 &= \lambda c_i^2 + (1 - \lambda) c_i^1 \end{aligned}$$

Cruce $BLX - \alpha$: Se generan dos hijos ($\alpha \in [0, 1]$ es un parámetro seleccionable por el usuario); cada gen se obtiene aleatoriamente del intervalo $[c_{min} - I\alpha, c_{max} + I\alpha]$ en donde $c_{min} = \min(c_i^1, c_i^2)$, $c_{max} = \max(c_i^1, c_i^2)$, $I = c_{max} - c_{min}$. El cruce correspondiente a $\alpha = 0$ es igual al cruce plano.

Cruce Lineal: Se generan tres hijos, de los cuales se seleccionan sólo los dos con mejor función de evaluación. Cada gen de los tres hijos se genera así:

$$\begin{aligned} h_i^1 &= \frac{1}{2} c_i^1 + \frac{1}{2} c_i^2 \\ h_i^2 &= \frac{3}{2} c_i^1 - \frac{1}{2} c_i^2 \\ h_i^3 &= -\frac{1}{2} c_i^1 + \frac{3}{2} c_i^2 \end{aligned}$$

Cruce Discreto: Cada gen h_i se escoge aleatoriamente entre c_i^1 y c_i^2 .

Cruce Línea extendida: $h_i = c_i^1 + \alpha(c_i^2 - c_i^1)$. α se escoge aleatoriamente en el intervalo $[-0.25, 1.25]$.

Cruce Extendido intermedio: $h_i = c_i^1 + \alpha_i(c_i^2 - c_i^1)$. α_i se escoge aleatoriamente en el intervalo $[-0.25, 1.25]$. Equivale al cruce $BLX - 0.25$.

Cruce Heurístico de Wrigth: Supóngase que C_1 es el padre con la mejor función de evaluación. En esas condiciones $h_i = r(c_i^1 - c_i^2) + c_i^1$ con r un número aleatorio perteneciente al intervalo $[0, 1]$.

Cruce BGA Lineal: En las mismas condiciones del cruce de Wrigth, se obtiene cada

gen así:

$$\begin{aligned} h_i &= c_i^1 + rango_i \gamma \Delta \\ \delta &= \frac{c_i^2 - c_i^1}{\|C_1 - C_2\|} \\ \gamma &= \sum_{k=0}^{15} \alpha_k 2^{-k} \end{aligned}$$

α_k puede ser 0 ó 1, con probabilidad de ser 1: $p(\alpha_k = 1) = \frac{1}{16}$. Usualmente, $rango_i = 0.5(b_i - a_i)$

B.2.2.2. Mutación Real

Al igual que con los operadores de cruce, existe una mayor variedad de operadores de mutación con codificación real que con codificación binaria. Algunos de los más usados son los siguientes:

Uniforme: Cada gen tiene una probabilidad de mutar. La probabilidad es seleccionable por el usuario. Si un individuo está formado por los genes c_1, c_2, \dots, c_n , y el gen c_i debe pertenecer al intervalo $[a_i, b_i]$, entonces el nuevo gen c'_i es un número aleatorio perteneciente al intervalo $[a_i, b_i]$

No Uniforme: al aplicar este operador en la generación t , con el número máximo de generaciones es g_{max} , el nuevo gen c'_i es

$$c'_i = \begin{cases} c_i + \Delta(t, b_i - c_i) & \text{si } \tau > 0 \\ c_i - \Delta(t, b_i - c_i) & \text{si } \tau \leq 0 \end{cases}$$

$$\Delta(t, y) = y \left(1 - r^{(1 - \frac{t}{g_{max}})^b} \right)$$

b es un parámetro seleccionable por el usuario; r es un número aleatorio en el intervalo $[0, 1]$; τ es un número aleatorio que puede valer 0 ó 1.

Mutación de Muhlenbein:

$$\begin{aligned} c'_i &= c_i \pm \gamma rango_i \\ \gamma &= \sum_{k=0}^{15} \alpha_k 2^{-k} \end{aligned}$$

$rango_i$ define el rango de la mutación, y usualmente se emplea como $0.1(b_i - a_i)$. El signo + ó - se escoge aleatoriamente con igual probabilidad, y α_k puede ser 0 ó 1, con probabilidad de ser 1: $p(\alpha_k = 1) = \frac{1}{16}$.

B.3. Codificación Híbrida

Al analizar los distintos operadores de las codificaciones real y binaria, puede verse que quizás la primera de ellas tiene algunas propiedades en las que supera a la segunda; por ejemplo la posibilidad de controlar más fácilmente la relación *Exploración-Explotación* del algoritmo, gracias a la existencia de operadores parametrizables como el operador de cruce $BLX - \alpha$.

No obstante, las comparaciones que se hacen sobre las distintas estrategias de codificación demuestran que no puede afirmarse que una de ellas sea superior a las otras en todos los casos. Por el contrario, lo único que puede concluirse es que la elección de la estrategia de codificación más adecuada depende del tipo de problema a optimizar. Tómese por ejemplo el caso en el que una de las variables del espacio de búsqueda sea una variable booleana; no parece adecuado representar esta variable mediante un número real. También es sabido que la codificación binaria suele presentar inconvenientes cuando los espacios de búsqueda son continuos.

El problema se complica un poco más si consideramos que en un gran número de aplicaciones reales el espacio de búsqueda está compuesto por variables heterogéneas: algunas serán reales, otras binarias, otras enteras, etc. A título de ejemplo, supóngase que se desea optimizar la función $f(x_1, x_2, x_3)$

$$f(x_1, x_2, x_3) : R \times \{0, 1\} \times Z \rightarrow R$$

es decir, una función en la que la primera variable es real, la segunda es booleana y la tercera es entera. ¿cuál puede ser la estrategia de codificación más adecuada en este caso?

Justamente para enfrentar ese tipo de problemas, se propone aquí emplear una *Codificación Híbrida*, en la que en un mismo genoma coexisten genes de distinto tipo. Retomando el ejemplo anterior, se propone emplear un genoma de tres genes: el primero con codificación real, el segundo con codificación binaria y el tercero con codificación entera.

Es claro que el problema anterior puede intentar solucionarse empleando un único tipo de codificación (por ejemplo codificación binaria) para todas las variables. Si esto es así, ¿cuál es la ganancia de este enfoque? La respuesta está en la posibilidad de tener una *representación más natural* del problema, y así obtener una implementación más sencilla y eficiente del algoritmo.

Debido a que los operadores de cruce y mutación dependen del tipo de codificación empleada, al tener una codificación híbrida será necesario definir distintos operadores de cruce y mutación para cada uno de los distintos tipos de genes presentes en el genoma.

Apéndice C

Herramientas de Software

En este apéndice se hace una somera presentación de varias herramientas de software que ayudan a la implementación de sistemas difusos. Tres de ellas han sido diseñadas por el autor de este texto (**UNFuzzy**, **FuzzyNet** y **UNGenético**) y dos son librerías que forman parte de herramientas de propósito general (**scilab** y **python**). Todas ellas son de software libre.

Este apéndice no es un manual de usuario, sino una explicación de alto nivel sobre el propósito y alcance de cada una de esas herramientas. También se explican algunos detalles de su implementación.

C.1. UNFuzzy

UNFuzzy es una herramienta para el diseño, análisis e implementación de Sistemas de Lógica Difusa basados en reglas - SBR. La primera versión data de 1997 [22] y la versión más reciente de 2022. Es una herramienta de software libre¹ desarrollada en C++ con el apoyo de la librería gráfica **wxWidgets**². Sus principales funcionalidades son las siguientes:

Diseño: a través de la interfaz gráfica el usuario puede especificar características tales como:

- El número de variables de entrada y de salida, así como la función semántica de cada una de ellas. Hay 9 formas disponibles para definir la función de pertenencia de cada conjunto, así como algoritmos diseño y etiquetado rápidos.

¹Disponible en [○](#)

²Disponible en [○](#)

- El tipo y ancho del difusor para cada variable de entrada. Hay 5 opciones disponibles.
- El tipo de concretor para cada variable de salida. Hay 5 opciones disponibles.
- La base de reglas. A cada regla es posible asociarle un nivel de certeza y a cada antecedente un operador lingüístico. Adicionalmente hay algoritmos de definición rápida de la base de reglas de forma manual.
- La selección de cinco operadores internos: operador de implicación (9 opciones), operador mínimo de composición (8 opciones), operador máximo de composición (4 opciones), operador And (9 opciones) y operador de agregación de reglas (13 opciones).

Adicionalmente, están disponibles dos algoritmos de identificación de reglas: entrenamiento de universos fijos y entrenamiento de universos variables. Estos algoritmos pueden utilizarse con patrones de entrenamiento almacenados en archivos `csv` o especificados manualmente.

Análisis: a través de la interfaz gráfica el usuario puede:

- Calcular las salidas del SBR para unas entradas específicas.
- Explorar el comportamiento interno del SBR mediante la visualización de las salidas de cada uno de los bloques funcionales del sistema.
- Graficar en 2D la relación entrada - salida del SBR.
- Exportar en formato plano una tabla con la relación entrada - salida. El formato permite la visualización en 3D de la relación empleando el programa `gnuplot`.

Implementación: a través de la interfaz gráfica el usuario puede generar un archivo con un código fuente en lenguaje `C++` que puede ser compilado externamente. Este archivo contiene todas las funcionalidades del SBR.

Gracias a la generación de código fuente, un programador puede construir piezas de software con SBR en su interior sin preocuparse por los detalles de implementación. Puede también, por ejemplo, usar las estrategias de definición rápida de variables, los algoritmos de entrenamiento y cambiar dinámicamente el diseño del sistema.

Junto con `UNFuzzy` se distribuye `UNFuzzyNetwork`. Este último permite el diseño de redes multicapa de SBR, es decir, SBR interconectados de tal forma que las salidas de unos de ellos llegan directamente a las entradas de otros, en una arquitectura de varias capas. Mediante `UNFuzzyNetwork` también se puede generar automáticamente el código fuente en lenguaje `C++`, para obtener implementaciones externas.



Figura C.1: Galería de capturas de pantalla de UNFUZZY

La Figura C.1 muestra algunas capturas de pantalla de UNFuzzy y UNFuzzyNetwork. Para la preparación de este texto, se ha usado UNFuzzy directamente en los siguientes ejemplos: 4.1, 4.10, 4.11, 4.3, 4.5, 4.6, 4.9, 5.1, 5.2, 5.4 . Además, el código fuente en C++ generado por UNFuzzy se ha empleado en los ejemplos 6.10, 6.11, 6.12, 6.13, 6.16, 6.17, 6.18, 6.19, 6.9 .

C.2. FuzzyNet

FuzzyNet es una herramienta para el diseño e implementación de Redes de Sistemas de Computación con Palabras Basados en Aritmética Difusa - SBA. La primera versión data del 2005 (Véase, por ejemplo, una aplicación en [46]) y la versión más reciente de 2022. Es una herramienta de software libre³ desarrollada en C++ con el apoyo de la librería gráfica wxWidgets⁴. Si bien es cierto que el propósito de **FuzzyNet** es el diseño de redes de SBA, también permite el diseño de un SBA individual.

Para comprender el funcionamiento y alcance de **FuzzyNet** es necesario conocer el uso que allí se le da a los siguientes términos:

- **Nodo de Cálculo:** un nodo de cálculo es, en esencia, un SBA. Cada variable de entrada y de salida del SBA tiene una variable lingüística asociada.
- **Estrategia:** una estrategia es una red de SBA. Cada estrategia debe tener un nombre único que la distingue. Las redes de cálculo son del tipo ‘hacia adelante’, es decir, no se admiten lazos en la red.
- **Caso:** un caso es el conjunto de la información necesaria para calcular algo con una estrategia. Esencialmente es un conjunto de parejas (V, X) en donde V es el nombre de una variable y X el valor que toma esa variable. Cada caso debe tener un nombre único que lo distingue.
- **Proyecto:** un proyecto es un conjunto de estrategias aplicadas a un conjunto de casos.

Las principales funcionalidades de **FuzzyNet** son las siguientes

Diseño: el tipo de red de SBA que se puede diseñar está definido por las siguientes características del programa:

- Cada variable lingüística puede definirse manualmente o a través de algoritmos de diseño rápido.
- El número de α -cortes de cada número difuso es configurable.

³Disponible en [○](#)

⁴Disponible en [○](#)

- Los tipos de entrada válidos para cada nodo de cálculo son los siguientes: etiquetas lingüísticas, etiquetas lingüísticas con modificadores, intervalos, números reales, la salida de otro nodo de cálculo.
- Los tipos de salidas válidos para cada nodo son los siguientes: el valor representativo, la pareja valor representativo y ambigüedad, una etiqueta lingüística, una frase descriptiva.
- Es posible modificar globalmente el nivel de optimismo. Esta variable incide en el cálculo del valor representativo de cada número difuso.
- Hay un conjunto de 9 funciones de razonamiento aproximado disponibles para cada nodo.

Visualización: a través de la interfaz gráfica se puede visualizar un proyecto de diferentes formas:

- **Estrategia como red:** se visualiza la estrategia seleccionada como una red de SBA. Esta opción permite la edición de la red y de cada uno de sus componentes.
- **Caso como red:** se visualiza la información del caso seleccionado, para la estrategia seleccionada, como una red de SBA. Cada nodo toma un color característico según el tipo de información que fluye a través de él. Al seleccionar un nodo puede editarse la información relativa a él, para el caso seleccionado. Hay tres alternativas de visualización: 1. La etiqueta cuya consistencia con el número difuso del nodo es máxima, 2. el valor representativo del número difuso del nodo y 3. la ambigüedad del número difuso del nodo.
- **Caso como lista:** se visualiza la información del caso seleccionado, para la estrategia seleccionada como una tabla de datos.
- **Variables del caso actual:** se visualizan los datos de entrada del caso seleccionado, bien sea para la estrategia seleccionada, o para todo el proyecto.
- **Matriz:** se visualiza la información de todos los casos y todos los nodos de todas las estrategias como una tabla de datos.

Reportes: Es posible generar reportes altamente configurables en los siguientes formatos 1. L^AT_EX, 2. HTML y 3. Texto. Los reportes pueden ser creados para todo el proyecto, para la estrategia seleccionada, para todas las estrategias, para el caso seleccionado, para todos los casos o como análisis de resultados del proyecto.

Librerías: una librería es un archivo que guarda información parcial de un proyecto. Específicamente puede guardar la definición de estrategias, casos, variables o una combinación de ellas.

Implementación: el programa puede intercambiar información de forma estática con otros programas a través de la importación y exportación de archivos de texto. Adicionalmente, el programa **FuzzyNetConsole**⁵ permite usar desde la línea de comandos un archivo prediseñado en FuzzyNet.

La Figura C.2 muestra algunas capturas de pantalla de FuzzyNet Fuzzynet se ha usado directamente en este texto en los siguientes ejemplos: 12.7, 13.2 .

C.3. UNGenético

UNGenético es una librería en lenguaje C++ para la implementación de algoritmos genéticos⁶. La primera versión data de 2001 y la versión más reciente de 2025. Incorpora todas las definiciones que se explican en el apéndice B.

La Figura C.3 ilustra el uso de la librería. Esencialmente, el usuario tiene dos tareas:

1. Definir una clase heredera de la clase **Individuo** en la que se debe representar el problema a optimizar.
2. Ejecutar el algoritmo siguiendo mediante una corta secuencia de instrucciones.

La ejecución del algoritmo puede hacerse de dos formas: de forma continua ó paso a paso. La primera alternativa ejecuta al algoritmo con una única instrucción, mientras que la segunda permite definir un ciclo para cada iteración (para cada generación), lo que permite tener acceso a los resultados intermedios del algoritmo.

En la Figura C.3 también se observa que la clase heredera de **Individuo** que debe diseñar el usuario tiene 5 funciones. Para comprender el papel de estas funciones debe tenerse en cuenta que:

- Cada individuo posee un arreglo de genes denominado **Genoma**. Allí pueden contenerse genes de diferentes tipos, es decir, objetos de distintas clases herederas de **Gen**.
- Las características que definen a cada Individuo son propiedades diferentes a **Genoma** mediante las función **codificar()** esas propiedades se almacenan en el **Genoma** siguiendo alguna estrategia de codificación, mientras que con la función **decodificar()** se lee la información genética de **Genoma**, y se actualizan las propiedades del Individuo.
- La clase **Individuo** define la función virtual pura **codificar()**. Toda clase heredera debe redefinir esta función; se encarga de codificar las propiedades del Individuo en el **Genoma**

⁵Disponible en [○](#)

⁶Disponible en [○](#)

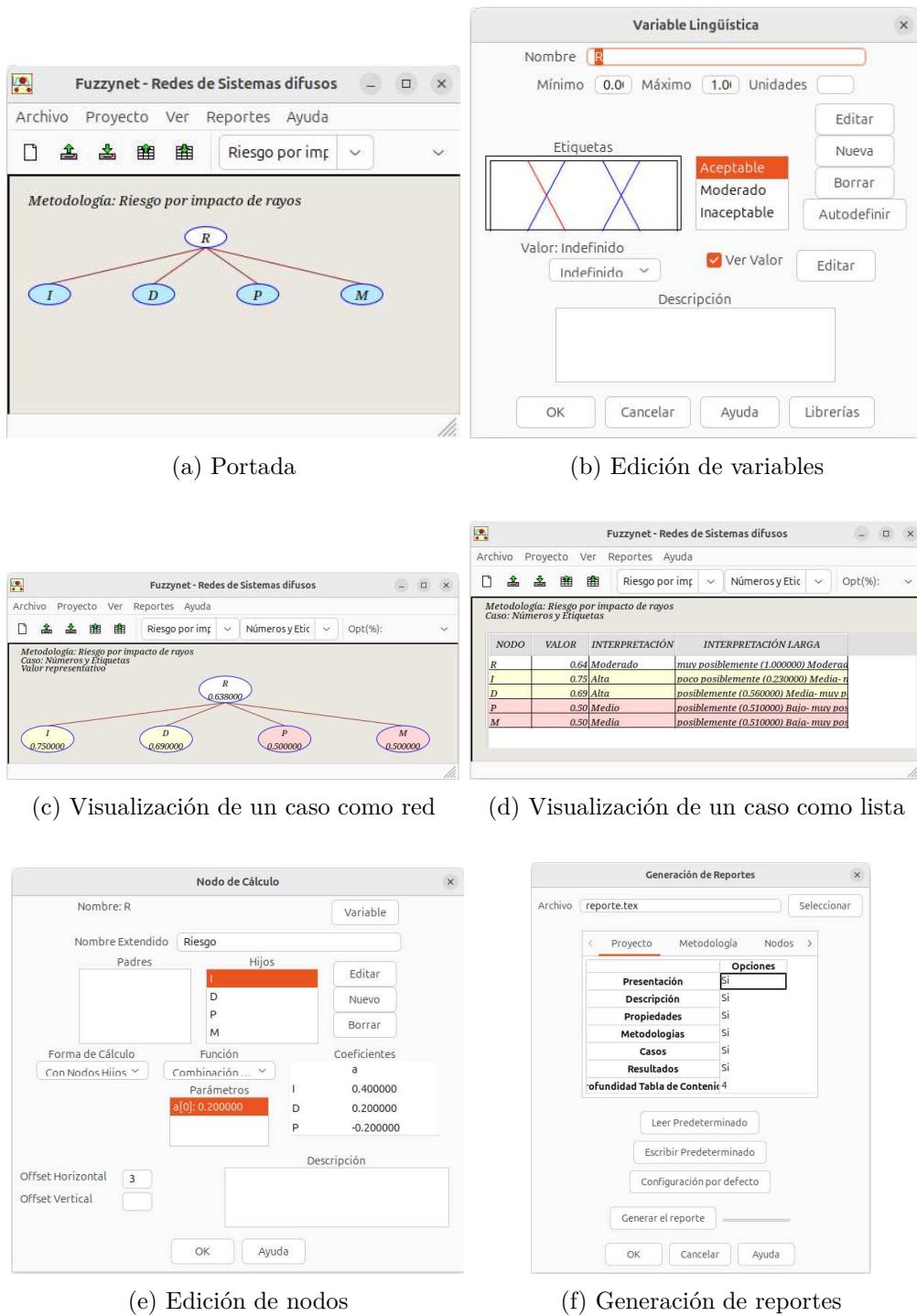


Figura C.2: Galería de capturas de pantalla de FuzzyNet

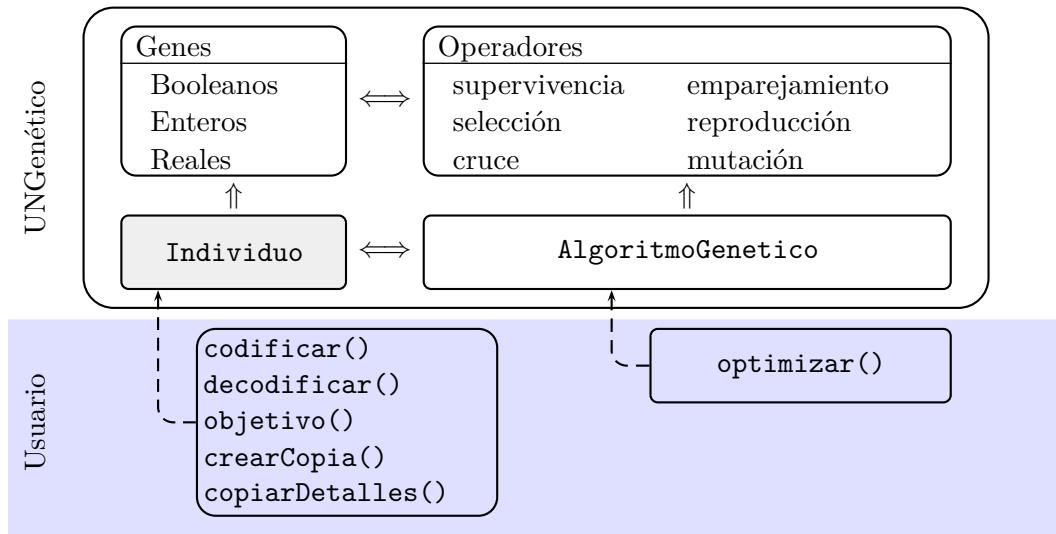


Figura C.3: Uso de UNGenético

- La clase `Individuo` define la función virtual pura `decodificar()`. Toda clase heredera debe redefinir esta función; se encarga de decodificar las propiedades del `Genoma` y actualizar los valores de las propiedades del `Individuo`.
- La clase `Individuo` define la función virtual pura `Individuo* crearCopia()`. Toda clase heredera debe redefinir esta función; en ella debe crearse un objeto de la misma clase, y retornar el apuntador a la clase padre `Individuo`
- La clase `Individuo` define la función virtual pura `copiarDetalles(Individuo *other)`. Toda clase heredera debe redefinir esta función; en esta función se copian las propiedades de `other` y se codifican
- La clase `Individuo` define la función virtual pura `double objetivo()`. Toda clase heredera debe redefinir esta función; aquí se calcula el valor de la función objetivo del `Individuo`.

UNGénetico se ha usado directamente en este texto en los siguientes ejemplos: [6.14](#), [6.15](#), [6.16](#), [6.17](#), [6.18](#), [6.19](#) .

En particular, aquellos ejemplos en los que se optimiza un SBR se han construido empleando tanto UNFuzzy como UNGénetico, mediante herencia múltiple. La idea se ilustra en la Figura C.4: se ha implementado la clase SLG_AG que hereda las definiciones de un SBR de la clase SLD de UNFuzzy y las definiciones de la clase `Individuo` de UNGénetico.

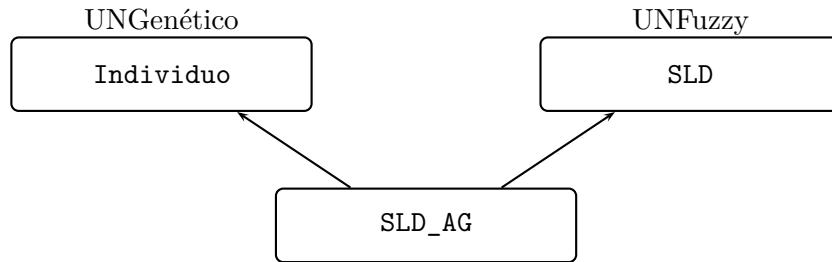


Figura C.4: Herencia múltiple

C.4. scilab

scilab es un conocido programa de cálculo numérico⁷. Fue desarrollado originalmente en los años 80 por el *Institut National de Recherche en Informatique et en Automatique* - INRIA (Instituto Nacional de Investigación en Informática y Automática) y hoy en día mantenido y publicado por Dassault Systèmes.

scilab es una poderosa herramienta de propósito matemático general que dispone de una amplia colección de librerías especializadas. Estas librerías se conocen como *toolboxes* o módulos independientes. En relación con los sistemas difusos, destacamos que:

- Existe un módulo que implementa varias estrategias de agrupamiento denominado *Clustering Toolbox*. Dicho módulo incluye el algoritmo Fuzzy-Cmeans.
- Existe un módulo específico para implementar SBR denominado *Fuzzy Logic Toolbox*. Desafortunadamente este módulo no es compatible con todas las versiones actuales de **scilab**. Funciona con las versiones **scilab 6.x** pero no con las versiones **scilab 202X**.
- La herramienta **scicos**, que forma parte de **scilab**, es particularmente útil para las aplicaciones de SBR como controladores.

scilab se ha usado directamente en este texto en los siguientes ejemplos: [5.3](#), [6.2](#), [6.3](#), [6.4](#), [6.5](#), [6.6](#), [6.7](#), [6.8](#), [6.9](#), [7.9](#), [8.4](#), [8.5](#), [9.4](#).

C.5. python

python es un lenguaje de programación de alto nivel y de propósito general. Fue desarrollado a finales de los 80 y hoy en día es uno de los lenguajes más populares para el

⁷La página web del proyecto se encuentra en [○](#)

desarrollo de aplicaciones científicas en el ámbito académico.

En particular, la librería **SciKit**⁸ provee una gran cantidad de funciones para el aprendizaje de máquina y en general para la analítica de datos. El módulo **skfuzzy** de esa librería está dedicado a sistemas difusos y aritmética difusa. Desafortunadamente, la implementación de aritmética difusa no es robusta, y puede arrojar resultados erróneos.

⁸La página web de **SciKit** está disponible en [○](#)