

INDICE

Prefacio	VII
1. Introducción	
1.1. El análisis de datos multivariantes	1
1.2. Estructura del libro	4
1.3. Datos y programas de ordenador	7
1.4. Un poco de historia	8
1.5. Lecturas complementarias	11
2. Algebra matricial	
2.1. Introducción	13
2.2. Vectores	15
2.2.1. Definiciones básicas	16
2.2.2. Dependencia lineal	20
2.3. Matrices	21
2.3.1. Definiciones básicas	22
2.3.2. Productos ente matrices	23
2.3.3. Rango de una matriz	25
2.3.4. Matrices cuadradas	26
2.3.5. Matrices particionadas	36
2.4. Vectores y valores propios	
2.4.1. Definición	37
2.4.2. Valores y vectores propios de matrices simétricas	40
2.4.3. Diagonalización de matrices simétricas	41
2.4.4. Raíz cuadrada de una matriz semidefinida positiva	43
2.4.5. Descomposición en valores singulares	
2.4.6. (*) Diagonalización de matrices generales	46
2.4.7. (*) Inversas generalizadas	47
2.5. (*) Proyección ortogonal	
2.5.1. Matrices idempotentes	48
2.5.2. Proyección ortogonal	50
2.6. (*) Derivadas matriciales	54
3. Descripción de datos multivariantes	
3.1. Introducción	61
3.2. Datos multivariantes	
3.2.1. Tipos de variables	62
3.2.2. La matriz de datos	63
3.2.3. Análisis univariante	65
3.3. Medidas de centralización: el vector de medias	67
3.4. La matriz de varianzas y covarianzas	
3.4.1. Calculo a partir de la matriz de datos centrados	70
3.4.2. Propiedades	73
3.4.3. Variables redundantes: el caso con matriz S singular	74
3.5. Medidas globales de variabilidad	77
3.5.1. La varianza total y la varianza media	
3.5.2. La varianza generalizada	78
3.5.3. La varianza efectiva	80
3.6. Variabilidad y distancias	
3.6.1. El concepto de distancia	81

3.6.2. La distancia de Mahalanobis	83
3.6.3. Distancias promedio	84
3.7. Medidas de dependencia lineal	
3.7.1. Dependencia por pares: la matriz de correlación	86
3.7.2. Dependencia de cada variable y le resto: regresión múltiple	87
3.7.3. Dependencia directa entre pares: correlaciones parciales	90
3.7.4. El coeficiente de dependencia efectiva	92
3.8. La matriz de precisión	94
3.9. Coeficientes de asimetría y kurtosis	95
Apéndice 3.1. La estructura de la matriz de preescisión	97
Apéndice 3.2. Los determinantes de S y R	99
Apéndice 3.3. Correlaciones parciales	100
4. Análisis grafico y datos atípicos	
4.1. Introducción	103
4.2. Representaciones graficas	
4.2.1. Histogramas y diagramas de dispersión	104
4.2.2. Representación mediante figuras	109
4.2.3. (*)Representación de proyecciones	111
4.3. Transformaciones lineales	
4.3.1. Consecuencias	113
4.3.2. Estandarización univariante	
4.3.3. (*)Estandarización multivariante	114
4.4. Trasformaciones no lineales	
4.4.1. Simplicidad en las distribuciones	117
4.4.2. Simplicidad en las relaciones	118
4.5. Datos atípicos	
4.5.1. Definición	120
4.5.2. Los efectos de los atípicos	121
4.5.3. (*) Identificaron de grupos de atípicos	123
4.6. Lecturas complementarias	
Apéndice 4.1. Calculo de las direcciones de máxima Kurtosis	131
5. Componentes principales	
5.1. Introducción	133
5.2. Planteamiento del problema	134
5.3. Calculo de los componentes	
5.3.1. Calculo del primer componente	137
5.3.2. Calculo del segundo componente	140
5.3.3. Generalización	143
5.4. Propiedades de los componentes	145
5.5. Análisis normado o con correlaciones	148
5.6. Interpretación de los componentes	152
5.6.1. Selección del numero de componentes	155
5.6.2. Representación grafica	156
5.6.3. Datos atípicos	
5.6.4. Distribuciones de los componentes	159
5.7. Generalizaciones	165
5.8. Lecturas complementarias	166
Apéndices 5.1. Distancias entre puntos y proyecciones	167
Apéndices 5.2. Los componentes como predictores óptimos	168

6. Escalado multidimensional	
6.1. Introducción	171
6.2. Escalados métricos: coordenadas principales	173
6.3. Matrices compatibles con métricas euclídeas	176
6.4. Construcción de las coordenadas principales	178
6.5. Relación entre coordenadas y componentes principales	182
6.6. Biplots	183
6.7. Escalado no métrico	187
6.8. Lecturas complementarias	
Apéndice 6.1. Maximización del stress	190
7. Análisis de correspondencia	
7.1. Introducción	193
7.2. Búsqueda de la mejor proyección	
7.2.1. Proyección de las filas	195
7.2.2. Proyección de las columnas	203
7.2.3. Análisis conjunto	204
7.3. La distancia ji-cuadrado	206
7.4. Asignación de puntuaciones	212
5.8. Lecturas complementarias	217
8. Análisis de conglomerados	
8.1. Fundamentos	219
8.2. Métodos clásicos de partición	
8.2.1. Fundamentos del algoritmo de k-medias	220
8.2.2. Implementación del algoritmo	221
8.2.3. Numero de grupos	223
8.3. Métodos jerárquicos	
8.3.1. Distancias y similitudes	230
8.3.2. Algoritmos jerárquicos	233
8.3.3. Métodos aglomerativos	234
8.4. Conglomerados por variables	239
8.5. Lecturas complementarias	242
Apéndice 8.1. Calculo del estadístico Ji-Cuadrado en tablas 2X2	243
9. Distribuciones multivariantes	
9.1. Conceptos básicos	245
9.1.1. Variables aleatorias vectoriales	
9.1.2. Distribución conjunta	246
9.1.3. Distribuciones marginales y condicionadas	247
9.1.4. Independencia	250
9.1.5. La maldición de la dimensión	251
9.2. Propiedades de variables vectoriales	
9.2.1. Vector de medias	252
9.2. Esperanza de una función	
9.2.3. Matriz de varianzas y covarianzas	253
9.2.4. Transformaciones de vectores aleatorios	254
9.2.5. Esperanzas y varianzas de transformaciones lineales	255
9.3. Dependencia entre variables aleatorias	
9.3.1. Esperanzas condicionadas	256
9.3.2. Varianzas condicionadas	257
9.3.3. Matriz de correlación	258

9.3.4. Correlaciones múltiples	
9.3.5. Correlaciones parciales	259
9.4. La distribución multinomial	260
9.5. La distribución de Dirichlet	263
9.6. La normal k-dimensional	
9.6.1. Propiedades básicas	264
9.6.2. Distribuciones condicionales	267
9.7. Distribuciones elípticas	270
9.7.1. Distribuciones esféricas	271
9.7.2. Función de densidad	272
9.8. (*) La distribución de Wishart	
9.8.1. Concepto	273
9.8.2. Propiedades de la distribución	275
9.9. La T2 de Hotelling	276
9.10. Distribuciones mezcladas	277
9.11. Lecturas complementarias	280
Apéndice 9.1. La distribución Wishart invertida	283
10. Inferencia con datos multivariantes	
10.1. Introducción	285
10.2. Fundamentos de la estimación máximo verosímil	286
10.3. Estimación de los parámetros de variables p-dimensionales	288
10.4. El método de la razón de verosimilitudes	290
10.5. Contraste sobre la media de una población normal	292
10.6. Contraste sobre la matriz de varianzas de una población normal	294
10.6.1. Contraste de un valor particular	
10.6.2. Contraste de independencia	295
10.6.3. Contraste de esfericidad	296
10.6.3. (*) Contraste de esfericidad parcial	
10.6.5. Ajustes en a distribución	297
10.7. Contraste de igualdad de varias medias: el análisis de la varianza multivariante	298
10.8. Contraste de datos atípicos	303
10.9. Contraste de normalidad	
10.9.1. Contraste básicos	304
10.9.2. Transformaciones	305
10.10. Lecturas complementarias	307
Apéndice 10.1. Inadmisibilidad de la media muestral para $p > 3$	308
Apéndice 10.2. Razón de verosimilitud y la T de Hotelling	309
Apéndice 10.3. Contraste de valores atípicos	
Apéndice 10.4. La distribución de Wilks	310
11. Métodos de inferencia avanzada multivariante	
11.1. Introducción	311
11.2. Estimación MV con datos faltantes	312
11.2.1. Estimación MV con el algoritmo EM	314
11.2.2. Estimación MV de mezclas	316
11.2.3. Estimación de poblaciones normales con datos ausentes	322
11.3. Estimación robusta	325
11.4. Estimación bayesiana	
11.4.1. Concepto	328

11.4.2. Distribución a priori	330
11.4.3. Calculo de la posterior	332
11.4.4. (*) Estimación bayesiana de referencia en el modelo normal	333
11.4.5. (*) Estimación con información a priori	334
11.5. Contrastes bayesianos	
11.5.1. Conceptos básicos	336
11.5.2. Comparación entre los contrastes bayesianos y los clásicos	338
11.6. Selección de modelos	
11.6.1. El criterio de Akaike	339
11.6.2. El criterio BIC	340
11.6.3. Relación entre el BIC y el AIC	
11.7. Lecturas complementarias	
Apéndice 11.1. Convergencia del algoritmo EM	343
Apéndice 11. 2. Estimación bayesiana	344
12. Análisis factorial	
12.1. Introducción	347
12.2. El modelo factorial	
12.2.1. Hipótesis básicas	348
12.2.2. Propiedades	349
12.2.3. Unicidad del modelo	351
12.2.4. Normalización del modelo factorial	353
12.2.5. Numero máximo de factores	353
12.3. El método del factor principal	354
12.3.1. Estimación de las comunalidades	356
12.3.2. Generalizaciones	361
12.4. (*) Estimación máximo verosímil	
12.4.1. Estimación MV de los parámetros	363
12.4.2. Otros métodos de estimación	366
12.5. Determinación del numero de factores	
12.5.1 Contraste de verosimilitud	368
12.5.2. Criterios de selección	370
12.6. Rotaron de los factores	372
12.7. Estimación de los factores	
12.7.1. Los factores como parámetros	375
12.7.2. Los factores como variables aleatorias	
12.8. Diagnósis del modelo	377
12.9. Relación con componentes principales	380
12.10. Análisis factorial confirmatorio	
12.11. Lecturas complementarias	382
Apéndice 12.1. Estimación máximo-verosímil del modelo factorial	384
Apéndice 12.2. Contrastes sobre el rango de una matriz	386
Apéndice 12.3. Estimación de los factores	388
Apéndice 12.4. Interpretación bayesiana del estimador de los factores	389
13. Análisis discriminante	
13.1. Introducción	391
13.2. Clasificación entre dos poblaciones	
13.2.1. Planteamiento del problema	392
13.2.2. Poblaciones normales: función lineal discriminante	395
13.2.3. Interpretación geométrica	396

12.3.4. Cálculo de probabilidades de error	399
13.2.5. Probabilidades a posteriori	400
13.3. Generalización para varias poblaciones normales	
13.3.1. Planteamiento general	402
13.3.2. Procedimiento operativo	404
13.4. Poblaciones desconocidas. Caso general	
13.4.1. Regla estimada de clasificación	407
13.4.2. Cálculo de probabilidades de error	408
13.5. Variables canónicas discriminantes	
13.5.1. El caso de dos grupos	410
13.5.2. Varios grupos	411
13.5.3. Variables canónicas discriminantes	415
13.6. Discriminación cuadrática. Discriminación en poblaciones no normales	420
13.7. Discriminación bayesiana	422
13.8. Lecturas complementarias	424
Apéndice 13.1. El criterio minimizar la probabilidad del error	425
Apéndice 13.2. Discriminación y regresión	426
14. Discriminación logística y otros métodos de clasificación	
14.1. Introducción	429
14.2. Modelos de respuesta cualitativa	
14.2.1. Tipos de modelos	430
14.2.2. El modelo logístico	433
14.2.3. Interpretación del modelo logístico	
14.3. La estimación del modelo logit	435
14.3.1. Estimación MV	
14.3.2. Contrastes	439
14.3.3. Diagnóstico	442
14.4. El modelo multilogit	
14.5. Otros métodos de clasificación	443
14.5.1. Árboles de clasificación	
14.5.2. Redes neuronales	447
14.5.3. Métodos no paramétricos	450
14.5.4. Máquinas del vector soporte	452
14.6. Lecturas complementarias	454
15. Clasificación mediante mezcla de distribuciones	
15.1. Fundamentos	457
15.2. El método de k-medias para mezclas	
15.2.1. Criterios	458
15.2.2. Determinación del número de grupos	461
15.2.3. Estimación de mezclas de normales	
15.3.1. Las ecuaciones de máxima verosimilitud para la mezcla	465
15.3.2. Resolución mediante el algoritmo EM	467
15.3.3. Aplicación al análisis de conglomerados	469
15.4. Métodos bayesianos	475
15.5. Métodos de proyección	478
15.6. Conclusiones	
15.7. Lecturas complementarias	482
Apéndice 15.1. Comparación del criterio de la traza y el determinante	483

16. Dependencia entre conjunto de variables: correlación canónica y otros métodos	485
16.1. Introducción	
16.2. Construcción de variables canónicas	486
16.2.1. La primera variable canónica	487
16.2.2. Las r variables canónicas	489
16.3. Propiedades de las variables y correlaciones canónicas	491
16.4. Análisis muestral	492
16.5. Interpretación geométrica	496
16.6. Contrastes	498
16.7. Extensiones a más de dos grupos	500
16.8. Relación con otras técnicas estudiadas	501
16.9. Análisis canónico asimétrico	502
16.9.1. Coeficientes de redundancia	503
16.9.2. Análisis canónico asimétrico	504
16.10. Regresión multivariante y modelos estructurales	505
16.11. Modelos estructurales en variables latentes	507
16.11.1. Propiedades	508
16.11.2. Estimación	509
16.12. Lecturas complementarias	510
Apéndice 16.1. Estimación del modelo de regresión multivariante	511
Apéndice A	
1. EUROALI	513
2. EUROSEC	514
3. EPF	515
4. INVEST	517
5. MEDIFIS	518
6. MUNDODES	519
7. ACCIONES	522
Referencias	
Libros de análisis multivariante	523
Artículos y otras referencias	529
Índice	537