

Indice degli argomenti

Prefazione	13
Software e materiale di corredo	19
1 Introduzione all'utilizzo di MATLAB e alla gestione dei dati	23
1.1 L'interfaccia di MATLAB	23
1.2 L'esecuzione del codice	26
1.3 Classi di dati	28
1.4 Gli array	31
1.4.1 Creazione dei diversi tipi di array	31
1.4.2 Estrazione dei dati da un array	34
1.5 Le tabelle	36
1.5.1 Creazione di una tabella	37
1.5.2 Importazione di una tabella	38
1.5.3 Estrazione dei dati da una tabella	40
1.6 Gestione dei dati	43
1.6.1 Salvataggio dei risultati ottenuti	43
1.6.2 Estrazione dei dati in base a dei criteri	45
1.6.3 Intersezione di due table con le stesse variabili	46
1.6.4 Unione e intersezione di due table	47
1.7 Introduzione alla generazione di numeri casuali	52
1.8 Elementi di base di programmazione	53
1.8.1 Le istruzioni if e i cicli	53
1.8.2 L'istruzione assert	57
1.8.3 Alcune costanti speciali	57
2 Algebra lineare di base	61
2.1 Operazioni elementari con le matrici	61
2.2 Le matrici diagonali	64
2.3 Alcune matrici particolari	67
2.4 Le matrici idempotenti	68
2.5 Le matrici ortogonali	69
2.6 Moltiplicazione di matrici trasposte	70
2.7 Moltiplicazione di matrici inverse	71
2.8 La trasposta dell'inversa	72
2.9 Sistemi di equazioni lineari	72
2.10 La traccia	74
2.11 L'espansione implicita	76

2.12	Matrice degli scostamenti standardizzati tramite espressioni matriciali	78
3	La matrice dei dati e le analisi univariate	83
3.1	Analisi univariate di variabili categoriche	85
3.2	Analisi univariate di variabili quantitative	88
3.3	Grafici univariati per dati categorici o quantitativi	90
3.4	Le distribuzioni di frequenze di variabili quantitative	97
3.5	Analisi univariate di variabili quantitative in presenza di sottogruppi	100
3.6	Intervalli di confidenza e riepiloghi avanzati	104
4	Variabili casuali: densità e distribuzioni	111
4.1	La variabile Gaussiana o normale	111
4.2	La variabile aleatoria Uniforme	122
4.3	La variabile aleatoria chi quadrato	125
4.4	La variabile aleatoria T di Student	130
4.5	La distribuzione normale bivariata	133
4.6	La distribuzione uniforme discreta	140
5	I trattamenti preliminari dei dati	143
5.1	I dati mancanti e strategie per il loro trattamento	143
5.2	I valori anomali e le strategie per il loro trattamento	146
5.3	Analisi automatica dei missing e dei valori anomali univariati	151
5.4	Operazioni avanzate con i missing values	154
6	La relazione tra le variabili quantitative: correlazione e cograduazione	159
6.1	La matrice di covarianze e la matrice di correlazione	162
6.2	La significatività della correlazione	167
6.3	L'indice di cograduazione di Spearman	174
6.3.1	La cograduazione in presenza di gradi ripetuti	174
6.4	La relazione tra gli indici di correlazione e cograduazione	179
6.5	La correlazione e la cograduazione in presenza di missing values	181
7	L'associazione	185
7.1	Introduzione	185
7.2	Indici di associazione per le tabelle 2×2	185
7.2.1	Misure basate sulla statistica di Pearson	189
7.2.2	Misure basate sul rapporto dei prodotti incrociati	193

7.3	Indici di associazione per le tabelle $I \times J$	196
7.4	Indici di associazione basati sulla riduzione dell'errore	200
7.5	Indici di associazione basati sulla riduzione dell'eterogeneità	203
7.6	Indici di associazione per variabili ordinali	207
8	Le rappresentazioni grafiche	213
8.1	Rappresentazioni grafiche per serie storiche univariate	213
8.2	I grafici a cascata	215
8.3	Rappresentazioni grafiche per serie storiche bivariate	217
8.4	I grafici ad imbuto	219
8.5	Grafici per la stima della densità univariata	220
8.6	Grafici a dispersione personalizzati	223
8.7	Grafici a dispersione con istogrammi o boxplot ai margini	224
8.8	Grafici con istogrammi bivariati	228
8.9	Grafici esplorativi per l'analisi di regressione	229
8.10	Grafici a tre dimensioni	234
8.11	Il balloonplot	237
8.12	La matrice dei diagrammi a dispersione	239
8.12.1	La matrice dei diagrammi a dispersione con variabile di raggruppamento	243
8.13	Le heatmap	245
8.14	I grafici in coordinate parallele	250
8.14.1	I grafici in coordinate parallele con variabile di raggruppamento	251
8.15	Le rappresentazioni iconiche	254
8.15.1	Le stelle	255
8.15.2	Le facce di Chernoff	256
9	Algebra lineare avanzata	259
9.1	La norma di un vettore	259
9.2	Il prodotto scalare	262
9.3	Le forme quadratiche	265
9.3.1	Estrazione degli elementi tramite forme quadratiche	272
9.4	Lo spazio vettoriale	273
9.4.1	Dipendenza, indipendenza lineare e base di uno spazio vettoriale	273
9.5	Il rango	277
9.6	Autovalori e autovettori	280
9.7	Polinomio caratteristico	280
9.8	Routines per il calcolo degli autovalori e degli autovettori	283

9.9	Scomposizione spettrale	283
9.9.1	La scomposizione spettrale attraverso il calcolo simboli- co	284
9.10	Introduzione ai poligoni	286
9.11	Proiezioni ortogonali	293
10	Le distanze e gli indici di similarità	297
10.1	Definizione di distanze	297
10.2	Alcuni tipi di distanza	298
10.3	Gli indici di distanza e gli indici di dissimilarità	306
10.4	Lo spazio euclideo ponderato	307
10.5	La distanza di Mahalanobis	309
10.5.1	Proprietà della distanza di Mahalanobis	315
10.6	La scala di misura delle distanze	316
10.7	Gli indici di similarità	317
10.7.1	Indici di similarità per fenomeni dicotomici	318
10.7.2	Indici di similarità in presenza di fenomeni misti	324
11	La riduzione delle dimensioni	329
11.1	Analisi in componenti principali (PC): introduzione	329
11.2	La prima PC come combinazione lineare delle variabili origina- rie	330
11.3	Le prime k PC come combinazioni lineari delle variabili origi- narie	332
11.3.1	Relazione tra autovalori traccia e determinante	333
11.4	La scomposizione in valori singolari (svd)	335
11.5	Le prime k PC come migliore rappresentazione di rango k delle variabili originarie	337
11.6	PC come proiezione ortogonale dei punti in un sottospazio di dimensione ridotta	340
11.6.1	Retta di regressione e retta associata alla prima compo- nente principale	341
11.6.2	Ricostruzione della matrice originaria con una matrice di rango ridotto	346
11.6.3	Componenti principali come rotazione degli assi carte- siani	348
11.7	L'analisi in componenti principali in pratica	356
11.8	Il biplot	367
11.8.1	Prima rappresentazione dei punti riga e dei punti colonna	371

11.8.2 Seconda rappresentazione dei punti riga e dei punti colonna	372
11.9 La funzione pcaFS	376
11.10 Componenti principali su \tilde{X} oppure su Z	384
12 L'analisi delle corrispondenze	389
12.1 Notazione	392
12.2 Giudizi sulla bontà dell'analisi e punteggi	403
12.3 Contributi all'inerzia del punto o all'inerzia della dimensione latente	409
12.4 La funzione CorAna	411
13 Cluster analysis	419
13.1 Metodi di cluster gerarchico	421
13.2 Il dendrogramma	426
13.2.1 Concetti alla base della strategia agglomerativa	426
13.2.2 Definizione di distanza tra due gruppi e metodo di raggruppamento	428
13.3 Caratteristiche dei metodi gerarchici	430
13.4 Le scelte nell'analisi dei gruppi	431
13.5 Metodi di clustering non gerarchici	434
13.5.1 Il metodo delle k -medie	437
13.6 La scelta del numero ottimo di gruppi	442
13.7 Il confronto tra partizioni	444
13.8 Classificazioni basati su modelli di misture finite	448