

Indice degli argomenti

Prefazione	15
Prefazione alla seconda edizione	23
Software e materiale di corredo	25
1 Introduzione all'utilizzo di MATLAB e alla gestione dei dati	29
1.1 L'interfaccia di MATLAB	29
1.2 L'esecuzione del codice	32
1.3 Classi di dati	34
1.4 Gli array	37
1.4.1 Creazione dei diversi tipi di array	37
1.4.2 Estrazione dei dati da un array	40
1.5 Le tabelle	42
1.5.1 Creazione di una tabella	43
1.5.2 Importazione di una tabella	44
1.5.3 Introduzione ai task	46
1.5.4 Estrazione dei dati da una tabella	49
1.6 Gestione dei dati	52
1.6.1 Salvataggio dei risultati ottenuti	52
1.6.2 Estrazione dei dati in base a criteri	54
1.7 Introduzione alla generazione di numeri casuali	55
1.8 Elementi di base di programmazione	56
1.8.1 Le istruzioni <code>if</code> e i cicli	56
1.8.2 L'istruzione <code>assert</code>	60
1.8.3 Alcune costanti e funzioni speciali	61
1.8.4 Il sistema di help di MATLAB	63
1.8.5 Il comando <code>plot</code>	64
Esercizi di riepilogo	64

2	Algebra lineare di base	67
2.1	Operazioni elementari con le matrici	67
2.2	Le matrici diagonali	70
2.3	Alcune matrici particolari	73
2.4	Le matrici idempotenti	75
2.5	Le matrici ortogonali	75
2.6	Moltiplicazione di matrici trasposte	76
2.7	Moltiplicazione di matrici inverse	78
2.8	La trasposta dell'inversa	78
2.9	Sistemi di equazioni lineari	79
2.10	Le operazioni matriciali elemento per elemento	81
2.11	La traccia	82
2.12	L'espansione implicita	84
	Esercizi di riepilogo	89
3	Analisi esplorative dei dati e tabelle pivot	93
3.1	Analisi univariate di variabili categoriche	95
3.2	Analisi univariate di variabili quantitative	96
3.3	Grafici univariati per dati categorici o quantitativi	101
3.4	Le distribuzioni di frequenze di variabili quantitative	107
3.5	Analisi univariate di variabili quantitative in presenza di sottogruppi	111
3.6	Intervalli di confidenza e riepiloghi avanzati	116
3.7	Tabelle pivot tramite la creazione guidata	121
3.8	Tabelle pivot avanzate	122
3.9	Boxplot in base a più variabili di raggruppamento	129
	Esercizi di riepilogo	132
4	Importazione dei dati dal mondo web (in tempo reale)	135
4.1	Importazione dei dati da una pagina web	135
4.2	Importazione dei dati da una pagina di GitHub	136
4.2.1	Gestione dei file scaricati tramite GitHub	140
4.3	Importazione di dataset di grandi dimensioni	142
4.4	Introduzione alle timetable	145
4.4.1	Estrarre record dalle timetable	146
4.4.2	Creare sequenze di date	150
4.5	Importazione serie storiche ISTAT	151
4.6	Importazione personalizzata	154
4.6.1	Cambiamento della periodicità delle serie storica	156

4.7	Importazione in tempo reale di serie storiche finanziarie o dai social media	157
	Esercizi di riepilogo	160
5	Variabili casuali: densità, distribuzioni e stime dei parametri	165
5.1	La variabile Gaussiana o normale	165
5.2	La variabile aleatoria Uniforme	175
5.3	La variabile aleatoria chi quadrato	181
5.4	La variabile aleatoria T di Student	186
5.5	Stima dei parametri delle distribuzioni	189
5.6	La distribuzione normale bivariata	190
5.7	La distribuzione uniforme discreta	197
	Esercizi di riepilogo	199
6	I trattamenti preliminari dei dati	205
6.1	I dati mancanti e strategie per il loro trattamento	205
6.2	I valori anomali e le strategie per il loro trattamento	208
6.3	Analisi automatica dei missing e dei valori anomali univariati	214
6.4	Operazioni avanzate con i missing values	216
6.5	Verifica dell'ipotesi di normalità	219
	Esercizi di riepilogo	221
7	La relazione tra le variabili quantitative: correlazione e cograduazione	223
7.1	La matrice di covarianze e la matrice di correlazione	226
7.2	La significatività della correlazione	231
7.3	L'indice di cograduazione di Spearman	239
7.3.1	La cograduazione in presenza di gradi ripetuti	239
7.4	La relazione tra gli indici di correlazione e cograduazione . . .	243
7.5	La correlazione e la cograduazione in presenza di missing values	244
	Esercizi di riepilogo	247
8	L'associazione	253
8.1	Introduzione	253
8.2	Indici di associazione per le tabelle 2×2	253
8.2.1	Misure basate sulla statistica di Pearson	257
8.2.2	Misure basate sul rapporto dei prodotti incrociati	261
8.3	Indici di associazione per le tabelle $I \times J$	264
8.4	Indici di associazione basati sulla riduzione dell'errore	268

8.5	Indici di associazione basati sulla riduzione dell'eterogeneità	271
8.6	Indici di associazione per variabili ordinali	275
	Esercizi di riepilogo	280
9	Le rappresentazioni grafiche	283
9.1	Rappresentazioni grafiche per serie storiche univariate	283
9.2	I grafici a cascata	285
9.3	Rappresentazioni grafiche per serie storiche bivariate	287
9.4	I grafici ad imbuto	289
9.5	Grafici per la stima della densità univariata	290
9.6	Grafici a dispersione personalizzati	293
9.7	Grafici a dispersione con istogrammi o boxplot ai margini . .	296
9.8	Grafici con istogrammi bivariati	299
9.9	Grafici esplorativi per l'analisi di regressione	299
9.10	Grafici a tre dimensioni	303
9.11	Il balloonplot	307
9.12	I diagrammi a bolle (bubblecharts)	308
	9.12.1 I grafici a mappa: bubblechart geografico	312
9.13	I grafici a sciame (swarmchart)	315
9.14	La matrice dei diagrammi a dispersione	317
	9.14.1 La matrice dei diagrammi a dispersione con variabile di raggruppamento	321
9.15	Le heatmap	323
9.16	I grafici in coordinate parallele	327
	9.16.1 I grafici in coordinate parallele con variabile di rag- gruppamento	329
9.17	Le rappresentazioni iconiche	332
	9.17.1 Le stelle	332
	9.17.2 Le facce di Chernoff	334
	Esercizi di riepilogo	335
10	Algebra lineare avanzata	341
10.1	La norma di un vettore	341
10.2	Il prodotto scalare	344
10.3	Le forme quadratiche	348
	10.3.1 Estrazione degli elementi tramite forme quadratiche	354
10.4	Spazi vettoriali	355
	10.4.1 Dipendenza, indipendenza lineare e base di uno spazio vettoriale	356
10.5	Il rango	360

10.6	Autovalori e autovettori	363
10.7	Polinomio caratteristico	364
10.8	Routine per il calcolo degli autovalori e degli autovettori	366
10.9	Scomposizione spettrale	367
10.9.1	La scomposizione spettrale attraverso il calcolo simbolico	368
10.10	Introduzione ai poligoni	370
10.11	Proiezioni ortogonali	376
	Esercizi di riepilogo	382
11	Le distanze e gli indici di similarità	387
11.1	Definizione di distanze	387
11.2	Alcuni tipi di distanza	388
11.3	Gli indici di distanza e gli indici di dissimilarità	396
11.4	Lo spazio euclideo ponderato	397
11.5	La distanza di Mahalanobis	399
11.5.1	Proprietà della distanza di Mahalanobis	405
11.6	La scala di misura delle distanze	406
11.7	Gli indici di similarità	407
11.7.1	Indici di similarità per fenomeni dicotomici	408
11.7.2	Indici di similarità in presenza di fenomeni misti	414
	Esercizi di riepilogo	416
12	La riduzione delle dimensioni	419
12.1	Analisi in componenti principali (PC): introduzione	419
12.2	La prima PC come combinazione lineare delle variabili originarie	420
12.3	Le prime k PC come combinazioni lineari delle variabili originarie	422
12.3.1	Relazione tra autovalori traccia e determinante	423
12.4	La scomposizione in valori singolari (svd)	425
12.5	Le prime k PC come migliore rappresentazione di rango k	427
12.6	PC come proiezione ortogonale dei punti in un sottospazio di dimensione ridotta	429
12.6.1	Retta di regressione e retta associata alla prima componente principale	430
12.6.2	Ricostruzione della matrice originaria con una matrice di rango ridotto	435
12.6.3	Componenti principali come rotazione degli assi cartesiani	437
12.7	L'analisi in componenti principali in pratica	447

12.8	Il biplot	459
12.8.1	Prima rappresentazione dei punti riga e dei punti colonna	461
12.8.2	Seconda rappresentazione dei punti riga e dei punti colonna	463
12.9	Qualità della proiezione per ogni punto	468
12.10	La funzione <code>pcaFS</code>	473
12.11	Componenti principali su \tilde{X} oppure su Z	481
	Esercizi di riepilogo	486
13	L'analisi delle corrispondenze	491
13.1	Notazione	494
13.2	Giudizi sulla bontà dell'analisi e punteggi	507
13.3	Contributi all'inerzia del punto o all'inerzia della dimensione latente	511
13.4	La funzione <code>CorAna</code>	513
13.5	Strumenti avanzati di analisi delle corrispondenze	520
13.5.1	Il moonplot	520
13.5.2	Personalizzazione dei grafici di analisi delle corrispondenze	521
	Esercizi di riepilogo	524
14	Cluster analysis	529
14.1	Metodi di cluster gerarchico	531
14.2	Il dendrogramma	536
14.2.1	Concetti alla base della strategia agglomerativa	536
14.2.2	Definizione di distanza tra due gruppi e metodo di raggruppamento	538
14.3	Caratteristiche dei metodi gerarchici	540
14.4	Le scelte nell'analisi dei gruppi	541
14.5	Metodi di clustering non gerarchici	544
14.5.1	Il metodo delle k -medie	547
14.6	La scelta del numero ottimo di gruppi	552
14.7	Il confronto tra partizioni	554
14.8	Classificazioni basati su modelli di misture finite	558
	Esercizi di riepilogo	566
15	Analisi delle serie storiche	569
15.1	Introduzione	569
15.2	Le trasformazioni della serie e le loro caratteristiche descrittive	570

15.2.1	Confronto nel tempo e misura delle variazioni	571
15.2.2	Aggregazione contemporanea e comparazione di aggregati nel tempo	575
15.2.3	Dipendenza seriale: autocovarianza e autocorrelazione	578
15.2.4	Identificazione dei cicli: il periodogramma	583
15.3	Modelli e metodi per serie temporali	585
15.4	Il modello classico (deterministico) di scomposizione di una serie temporale	587
15.4.1	Trend	588
15.4.2	Ciclo	589
15.4.3	Stagionalità	589
15.4.4	Irregolarità del calendario, valori anomali e cambiamenti strutturali	596
15.4.5	Stima e previsione	598
15.5	Regressione polinomiale locale e medie mobili	606
15.5.1	Polinomi locali e stima kernel	607
15.5.2	Il trattamento delle estremità delle serie. La stima in tempo reale	612
15.5.3	Inferenza	614
15.5.4	Stima di h per convalida incrociata	616
15.5.5	Stima di σ^2 e stima intervallare di μ_t	617
15.5.6	Le medie mobili	617
15.5.7	Dipendenza seriale spuria, effetto ampiezza e fase . .	619
15.6	Trend locali e livellamento esponenziale	621
15.6.1	Previsione mediante livellamento esponenziale	621
15.6.2	Il livellamento esponenziale applicato alla misura del rischio di mercato	625
15.6.3	Previsione mediante il metodo di Holt-Winters	626
15.6.4	Metodo di Holt-Winters stagionale	628
15.7	La modellazione del breve periodo	629
15.7.1	Stazionarietà	630
15.7.2	Previsione lineare ottimale	632
15.7.3	Funzione di autocorrelazione parziale	633
15.7.4	Stima della media e della funzione di autocovarianza	634
15.7.5	La verifica dell'ipotesi di incorrelazione	635
15.7.6	Il teorema di Wold	636
15.7.7	Processi Autoregressivi	637
15.7.8	Processi Media Mobile	641
15.7.9	Processi misti	643

15.7.10	Inferenza	645
15.7.11	Verifica del modello stimato	649
15.7.12	Nonstazionarietà: trend deterministici o stocastici?	650
15.7.13	Modelli ARIMA	652
15.7.14	Modelli ARIMA stagionali	653
15.7.15	Modelli ARIMA con variabili esogene e modelli regA- RIMA	654
15.7.16	Previsione	657
	Appendice: Stima del modello di regressione multipla	665
	Esercizi di riepilogo	671
	Bibliografia	675