

# Indice degli argomenti

<b>Prefazione</b>	<b>15</b>
<b>Prefazione alla seconda edizione</b>	<b>23</b>
<b>Software e materiale di corredo</b>	<b>25</b>
<b>1 Introduzione all'utilizzo di MATLAB e alla gestione dei dati</b>	<b>29</b>
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

<b>2</b>	<b>Algebra lineare di base</b>	<b>67</b>
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
<b>3</b>	<b>Analisi esplorative dei dati e tabelle pivot</b>	<b>93</b>
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
<b>4</b>	<b>Importazione dei dati dal mondo web (in tempo reale)</b>	<b>135</b>
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
<b>5</b>	<b>Variabili casuali: densità, distribuzioni e stime dei parametri</b>	<b>165</b>
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
<b>6</b>	<b>I trattamenti preliminari dei dati</b>	<b>205</b>
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
<b>7</b>	<b>La relazione tra le variabili quantitative: correlazione e cograduazione</b>	<b>223</b>
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
<b>8</b>	<b>L'associazione</b>	<b>253</b>
8.1	Introduzione . . . . .	253
8.2	Indici di associazione per le tabelle $2 \times 2$ . . . . .	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
<b>9</b>	<b>Le rappresentazioni grafiche</b>	<b>283</b>
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 . . . . .	300
9.10	Grafici a tre dimensioni . . . . .	303
9.11	Il balloonplot . . . . .	308
9.12	I diagrammi a bolle (bubblecharts) . . . . .	310
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 . . . . .	322
9.15	Le heatmap . . . . .	324
9.16	I grafici in coordinate parallele . . . . .	328
9.16.1	I grafici in coordinate parallele con variabile di rag- gruppamento . . . . .	330
9.17	Le rappresentazioni iconiche . . . . .	333
9.17.1	Le stelle . . . . .	333
9.17.2	Le facce di Chernoff . . . . .	335
	Esercizi di riepilogo . . . . .	336
<b>10</b>	<b>Algebra lineare avanzata</b>	<b>343</b>
10.1	La norma di un vettore . . . . .	343
10.2	Il prodotto scalare . . . . .	346
10.3	Le forme quadratiche . . . . .	350
10.3.1	Estrazione degli elementi tramite forme quadratiche	356
10.4	Spazi vettoriali . . . . .	357
10.4.1	Dipendenza, indipendenza lineare e base di uno spazio vettoriale . . . . .	358
10.5	Il rango . . . . .	362

Indice degli argomenti	11
10.6 Autovalori e autovettori . . . . .	365
10.7 Polinomio caratteristico . . . . .	366
10.8 Routine per il calcolo degli autovalori e degli autovettori . .	368
10.9 Scomposizione spettrale . . . . .	369
10.9.1 La scomposizione spettrale attraverso il calcolo sim- bolico . . . . .	370
10.10 Introduzione ai poligoni . . . . .	372
10.11 Proiezioni ortogonali . . . . .	378
Esercizi di riepilogo . . . . .	384
<b>11 Le distanze e gli indici di similarità</b>	<b>389</b>
11.1 Definizione di distanze . . . . .	389
11.2 Alcuni tipi di distanza . . . . .	390
11.3 Gli indici di distanza e gli indici di dissimilarità . . . . .	398
11.4 Lo spazio euclideo ponderato . . . . .	399
11.5 La distanza di Mahalanobis . . . . .	401
11.5.1 Proprietà della distanza di Mahalanobis . . . . .	407
11.6 La scala di misura delle distanze . . . . .	408
11.7 Gli indici di similarità . . . . .	409
11.7.1 Indici di similarità per fenomeni dicotomici . . . . .	410
11.7.2 Indici di similarità in presenza di fenomeni misti . .	416
Esercizi di riepilogo . . . . .	418
<b>12 La riduzione delle dimensioni</b>	<b>421</b>
12.1 Analisi in componenti principali (PC): introduzione . . . . .	421
12.2 La prima PC come combinazione lineare delle variabili originarie	422
12.3 Le prime $k$ PC come combinazioni lineari delle variabili origi- narie . . . . .	424
12.3.1 Relazione tra autovalori traccia e determinante . . .	425
12.4 La scomposizione in valori singolari (svd) . . . . .	427
12.5 Le prime $k$ PC come migliore rappresentazione di rango $k$ .	429
12.6 PC come proiezione ortogonale dei punti in un sottospazio di dimensione ridotta . . . . .	431
12.6.1 Retta di regressione e retta associata alla prima com- ponente principale . . . . .	432
12.6.2 Ricostruzione della matrice originaria con una matrice di rango ridotto . . . . .	437
12.6.3 Componenti principali come rotazione degli assi car- tesiani . . . . .	439
12.7 L'analisi in componenti principali in pratica . . . . .	449

12.8	Il biplot . . . . .	461
12.8.1	Prima rappresentazione dei punti riga e dei punti colonna . . . . .	463
12.8.2	Seconda rappresentazione dei punti riga e dei punti colonna . . . . .	465
12.9	Qualità della proiezione per ogni punto . . . . .	470
12.10	La funzione <code>pcaFS</code> . . . . .	475
12.11	Componenti principali su $\tilde{X}$ oppure su $Z$ . . . . .	483
	Esercizi di riepilogo . . . . .	488
<b>13</b>	<b>L'analisi delle corrispondenze</b>	<b>493</b>
13.1	Notazione . . . . .	496
13.2	Giudizi sulla bontà dell'analisi e punteggi . . . . .	509
13.3	Contributi all'inerzia del punto o all'inerzia della dimensione latente . . . . .	513
13.4	La funzione <code>CorAna</code> . . . . .	515
13.5	Strumenti avanzati di analisi delle corrispondenze . . . . .	522
13.5.1	Il moonplot . . . . .	522
13.5.2	Personalizzazione dei grafici di analisi delle corrispondenze . . . . .	523
	Esercizi di riepilogo . . . . .	526
<b>14</b>	<b>Cluster analysis</b>	<b>531</b>
14.1	Metodi di cluster gerarchico . . . . .	533
14.2	Il dendrogramma . . . . .	538
14.2.1	Concetti alla base della strategia agglomerativa . . . . .	538
14.2.2	Definizione di distanza tra due gruppi e metodo di raggruppamento . . . . .	540
14.3	Caratteristiche dei metodi gerarchici . . . . .	542
14.4	Le scelte nell'analisi dei gruppi . . . . .	543
14.5	Metodi di clustering non gerarchici . . . . .	546
14.5.1	Il metodo delle $k$ -medie . . . . .	549
14.6	La scelta del numero ottimo di gruppi . . . . .	554
14.7	Il confronto tra partizioni . . . . .	556
14.8	Classificazioni basati su modelli di misture finite . . . . .	560
	Esercizi di riepilogo . . . . .	567
<b>15</b>	<b>Analisi delle serie storiche</b>	<b>571</b>
15.1	Introduzione . . . . .	571
15.2	Le trasformazioni della serie e le loro caratteristiche descrittive	572

15.2.1	Confronto nel tempo e misura delle variazioni . . . .	573
15.2.2	Aggregazione contemporanea e comparazione di aggregati nel tempo . . . . .	577
15.2.3	Dipendenza seriale: autocovarianza e autocorrelazione	580
15.2.4	Identificazione dei cicli: il periodogramma . . . . .	585
15.3	Modelli e metodi per serie temporali . . . . .	587
15.4	Il modello classico (deterministico) di scomposizione di una serie temporale . . . . .	589
15.4.1	Trend . . . . .	590
15.4.2	Ciclo . . . . .	591
15.4.3	Stagionalità . . . . .	591
15.4.4	Irregolarità del calendario, valori anomali e cambiamenti strutturali . . . . .	598
15.4.5	Stima e previsione . . . . .	600
15.5	Regressione polinomiale locale e medie mobili . . . . .	608
15.5.1	Polinomi locali e stima kernel . . . . .	609
15.5.2	Il trattamento delle estremità delle serie. La stima in tempo reale . . . . .	614
15.5.3	Inferenza . . . . .	616
15.5.4	Stima di $h$ per convalida incrociata . . . . .	618
15.5.5	Stima di $\sigma^2$ e stima intervallare di $\mu_t$ . . . . .	619
15.5.6	Le medie mobili . . . . .	619
15.5.7	Dipendenza seriale spuria, effetto ampiezza e fase . .	621
15.6	Trend locali e livellamento esponenziale . . . . .	623
15.6.1	Previsione mediante livellamento esponenziale . . . .	623
15.6.2	Il livellamento esponenziale applicato alla misura del rischio di mercato . . . . .	627
15.6.3	Previsione mediante il metodo di Holt-Winters . . . .	628
15.6.4	Metodo di Holt-Winters stagionale . . . . .	630
15.7	La modellazione del breve periodo . . . . .	631
15.7.1	Stazionarietà . . . . .	632
15.7.2	Previsione lineare ottimale . . . . .	634
15.7.3	Funzione di autocorrelazione parziale . . . . .	635
15.7.4	Stima della media e della funzione di autocovarianza	636
15.7.5	La verifica dell'ipotesi di incorrelazione . . . . .	637
15.7.6	Il teorema di Wold . . . . .	638
15.7.7	Processi Autoregressivi . . . . .	639
15.7.8	Processi Media Mobile . . . . .	643
15.7.9	Processi misti . . . . .	645

15.7.10 Inferenza . . . . .	647
15.7.11 Verifica del modello stimato . . . . .	651
15.7.12 Nonstazionarietà: trend deterministici o stocastici? . . . . .	652
15.7.13 Modelli ARIMA . . . . .	654
15.7.14 Modelli ARIMA stagionali . . . . .	655
15.7.15 Modelli ARIMA con variabili esogene e modelli regA- RIMA . . . . .	656
15.7.16 Previsione . . . . .	659
Appendice: Stima del modello di regressione multipla . . . . .	667
Esercizi di riepilogo . . . . .	673
<b>Bibliografia</b>	<b>677</b>