

MATLAB

機率與數理統計

張德豐 等編著

溫坤禮 校訂



本書特色

- 著重機率與數理統計分析的思考方法
從具體實例引出相關內容
- 輕鬆掌握機率與數理統計內容
同時學習使用MATLAB進行資料處理的
基本方法和技巧

MATLAB

機率與數理統計

張德豐 等編著

溫坤禮 校訂



本書特色

- 著重機率與數理統計分析的思考方法
從具體實例引出相關內容
- 輕鬆掌握機率與數理統計內容
同時學習使用MATLAB進行資料處理的
基本方法和技巧

五南圖書出版公司 印行

國家圖書館出版品預行編目資料

MATLAB 機率與數理統計 / 張德豐等編著. — 初

版. — 臺北市：五南， 2013.06

面：公分.

I S B N: 978-957-11-7074-9 (平裝)

1. Matlab (電腦程式) 2. 機率 3. 數理統計

312.49M384

102005859



5DG3

MATLAB 機率與數理統計

作 著 — 張德豐 等

校 訂 — 溫坤禮

發 行 人 — 楊榮川

總 編 輯 — 王翠華

主 編 — 穆文娟

責任編輯 — 王者香

圖文編輯 — 林秋芬

封面設計 — 小小設計有限公司

出 版 者 — 五南圖書出版股份有限公司

地 址：106 台北市大安區和平東路二段 339 號 4 樓

電 話：(02)2705-5066 傳 真：(02)2706-6100

網 址：<http://www.wunan.com.tw>

電子郵件：wunan@wunan.com.tw

劃撥帳號：01068953

戶 名：五南圖書出版股份有限公司

台中市駐區辦公室 / 台中市區中山路 6 號

電 話：(04)2223-0891 傳 真：(04)2223-3549

高雄市駐區辦公室 / 高雄市新興區中山一路 290 號

電 話：(07)2358-702 傳 真：(07)2350-236

法律顧問 元貞聯合法律事務所 張澤平律師

出版日期 2013 年 6 月初版一刷

定 價 新臺幣 580 元

版權聲明

※本書為中國機械工業出版社獨家授權的中文繁體字版本，僅限於臺灣地區出版發行。本書專有版權屬中國機械工業出版社所有，該專有版權受法律保護。未經許可，不得擅自使用。

※版權所有，欲利用本書內容，必須徵求本公司同意※

MATLAB 是一種主要用於工程計算的高級電腦語言。美國的 MathWorks 公司自 1984 年推出 MATLAB 的 DIS 版本後，又推出了它的 Windows 版本，並且不斷推出更新的版本，使得 MATLAB 的應用領域越來越廣。到目前為止，MATLAB 已經包括模擬工具 Simulink、自動控制、信號處理、圖像處理、神經網路、模式識別、小波（Wavelet）分析、數理統計、生物資訊等 30 多個工具箱（Toolbox）。由於其靈活的編程方法和極高的編程效率，加上其在使用者介面和功能上的不斷擴展，自推出以來，日益受到廣大大專院校師生和研發人員的青睞。

MATLAB R2009a 是 MATLAB 的新版本，它對以往版本中的產品模組進行了一些調整。例如，MATLAB Builder for COM 的功能被整合到了 MATLAB Builder for .NET 中，Financial Time Series Toolbox 的功能被整合到了 Financial Toolbox 中。MATLAB 將高性能的數值計算和視覺化整合在一起，並提供了大量的內建函數，從而被廣泛地應用於科學計算、控制系統、資訊處理等領域的分析、模擬和設計工作。利用 MATLAB 產品的開放式結構，可以非常容易地對其功能進行擴充，從而不斷深化對工程問題的認識。

MATLAB 開放的產品體系使其成為了諸多領域開發的首選軟體。MATLAB 還具有 500 多家第三方合作夥伴，分佈在科學計算、機械動力、化工、電腦通信、汽車和金融等領域。介面方式包括了聯合建模、資料共用和開發流程銜接等。

為了更好地適應大專院校培養高等技術應用型人才的需要，解決大專院校「機率與統計」理論課與實驗課相結合的問題，並根據應用數學與專業相融、基礎數學為專業服務的基本要求和以應用為目的，以及必須與夠用為準的基本原則，作者在多年從事高等教育教學實驗的基礎上，編寫了本書。

本書介紹機率與數理統計的基本原理、典型應用，以及使用 MATLAB 進行實際工程分析的基本方法。全書共分 9 章。第 1 章介紹 MATLAB 的資料基礎；第 2 章介紹機率與數理統計基本概念，包括隨機事件及其機率、事件及運算、條件機率與事件的獨立性等內容；第 3 章介紹多維隨機變數，包括二維隨機變數、隨機向量的分佈、隨機向量函數的分佈、二維隨機向量的數字特徵等內容；第 4 章介紹統計估計及統計特徵，包括統計圖的繪製、變數分佈估計、參數的點估計、區間估計及機率分佈的統

計特徵等內容；第 5 章介紹統計檢驗方法——假設檢驗，包括假設檢驗概述、單常態總體的假設檢驗、兩常態總體參數的假設檢驗、非常態總體參數的假設檢驗及變數分佈形態的檢驗等內容；第 6 章介紹方差分析及曲線擬合，包括方差分析的相關概念、單因素方差分析、雙因素方差分析、資料曲線擬合及二次響應曲面模型等內容；第 7 章介紹迴歸分析，包括一元線性迴歸分析、多元線性迴歸分析、偏最小二乘迴歸分析等內容；第 8 章介紹多元統計分析，包括因素分析、聚類分析、正交實驗設計分析、多元方差分析、判別分析及實驗設計分析等內容；第 9 章介紹隱藏式馬可夫模型及統計工具箱的示範程式，包括隱藏式馬可夫模型、示範程式等內容。

本書具有如下特點：

- 一、注意機率與數理統計的思想方法介紹。在闡述某一統計概念方法時，一般是從具體實例開始引出相關內容，或是以具體實例結束相關內容。
- 二、本書在重視公式和定理推導的同時，也重視 MATLAB 應用於機率與數理統計方法時的簡單性、實用性和可操作性。實際中，機率與統計幾乎遍及各個領域，成為解決實際問題的重要工具。
- 三、凸顯了知識的技能化和應用意識的養成。

經由對本書的學習，讀者不僅可以掌握機率與數理統計的內容，同時也能初步掌握使用 MATLAB 進行資料處理的基本方法和技巧。

作者

目 錄

前言

1

MATLAB 的資料基礎 1

- 1.1 MATLAB 的主要功能 2
 - 1.1.1 MATLAB 簡介 2
 - 1.1.2 MATLAB 的資料及數值分析 4
 - 1.1.3 MATLAB 矩陣的建立及基本操作 19
 - 1.1.4 符號運算 23
 - 1.1.5 MATLAB 的繪圖功能 25
 - 1.1.6 MATLAB 資料類型及輸出輸入 38
- 1.2 MATLAB 的程式 42
 - 1.2.1 關係及邏輯運算 42
 - 1.2.2 M 函數檔案 43
 - 1.2.3 M 檔案 45
 - 1.2.4 程式控制語句 45
 - 1.2.5 編寫程式要點 50

2

機率與數理統計基本概念 51

- 2.1 隨機事件及其機率 52
 - 2.1.1 隨機事件 52
 - 2.1.2 機率 53
 - 2.1.3 排列與組合 58
 - 2.1.4 古典機率 60

2.2 事件及運算	63
2.3 條件機率與事件的獨立性	70
2.3.1 條件機率	70
2.3.2 乘法公式	71
2.3.3 獨立性	72
2.4 機率空間	76
2.4.1 基本概念	76
2.4.2 機率空間	77
2.5 總體樣本	83
2.5.1 總體與樣本的基礎	83
2.5.2 分佈定理	85
2.6 統計量與抽樣分佈	86
2.6.1 統計量	86
2.6.2 經驗分佈函數	87
2.6.3 χ^2 分佈	90
2.6.4 t 分佈	94
2.6.5 F 分佈	94
2.6.6 超幾何分佈	95
2.6.7 常態分佈	96
2.6.8 常態總體的樣本均值與樣本方差的分佈	99
2.6.9 機率密度函數對比——直方圖估計法	105
2.7 統計檢驗	108
2.7.1 統計檢驗的基本原理	108
2.7.2 異常值檢驗	109
2.7.3 方差檢驗	111
2.7.4 分佈擬合檢驗	112

3

多維隨機變數

117

3.1	二維隨機變數	118
3.1.1	二維隨機變數的定義	118
3.1.2	離散型隨機向量	118
3.1.3	連續型隨機向量	120
3.1.4	隨機向量的均勻分佈	121
3.2	隨機向量的分佈	124
3.2.1	邊緣分佈	124
3.2.2	條件分佈	130
3.2.3	二維常態分佈	134
3.3	隨機向量函數的分佈	135
3.3.1	二維隨機向量函數的概念	135
3.3.2	函數分佈	136
3.4	二維隨機向量的數字特徵	141
3.4.1	數學期望值	141
3.4.2	邊緣分佈的期望與方差	142
3.4.3	協方差	143
3.4.4	相關係數	145
3.4.5	矩與協方差矩陣	147
3.5	大數定律與中心極限定理	152
3.5.1	切比雪夫 (Chebyshev) 不等式	152
3.5.2	大數定律	153
3.5.3	中心極限定理	160

4

統計估計及統計特徵

167

4.1	統計圖的繪製	168
-----	--------	-----



5

4.1.1	盒狀圖	168
4.1.2	分佈圖	170
4.1.3	散度圖	177
4.2	變數分佈估計	179
4.2.1	頻率分佈表與頻率直方圖	179
4.2.2	五數概括與盒狀圖	185
4.3	參數的點估計	189
4.3.1	矩估計法	189
4.3.2	極大似然估計法	190
4.3.3	估計量的性能分析	197
4.4	區間估計	201
4.4.1	區間估計的概念	201
4.4.2	單常態總體參數的區間估計	205
4.4.3	單側置信區間	208
4.5	機率分佈的統計特徵	210
4.5.1	機率密度和累積分佈密度	210
4.5.2	機率分佈的均值和方差	211
	統計檢驗方法——假設檢驗	215
5.1	假設檢驗概述	216
5.1.1	假設檢驗的邏輯	216
5.1.2	假設檢驗的步驟	218
5.1.3	檢驗的 p 值	220
5.1.4	假設檢驗錯誤與勢函數	223
5.1.5	假設檢驗與區間估計的關係	225
5.2	單常態總體的假設檢驗	227
5.2.1	總體均值的檢驗	227



5.2.2	總體 $N(\mu, \sigma^2)$ 方差 σ^2 的檢驗	235
5.3	兩常態總體參數的假設檢驗	238
5.3.1	方差未知但相等時兩個常態總體均值的檢驗	238
5.3.2	兩個常態總體方差齊性（相等）的檢驗	242
5.4	非常態總體參數的假設檢驗	244
5.5	變數分佈形態的檢驗	246
5.5.1	χ^2 擬合優度檢驗	247
5.5.2	$K_{LJIMORPOB}-C_{MHPHOB}$ 檢驗	257
5.5.3	常態性檢驗	261
5.5.4	符號檢驗法	268
5.5.5	秩和檢驗法	269

6

方差分析及曲線擬合 271

6.1	方差分析的相關概念	272
6.1.1	基本概念	272
6.1.2	方差分析的必要性	272
6.1.3	方差分析的基本想法	274
6.2	單因素方差分析	275
6.2.1	單因素統計模型及檢驗方法	275
6.2.2	效應與誤差方差的估計	283
6.2.3	重複數相同的方差分析	285
6.2.4	多重比較	289
6.2.5	方差齊性檢驗	292
6.3	雙因素方差分析	295
6.3.1	雙因素無重複實驗的方差分析	295
6.3.2	雙因素重複實驗的方差分析	298
6.3.3	多因素方差分析	301



6.4 資料曲線擬合	305
6.4.1 多項式擬合	305
6.4.2 連分式展開及連分式的有理近似	307
6.4.3 有理式擬合	311
6.4.4 函數線性組合的曲線擬合方法	314
6.4.5 最小二乘曲線擬合	317
6.5 二次響應曲面模型	321

7

迴歸分析 325

7.1 一元線性迴歸分析	326
7.1.1 一元線性迴歸分析的基本定義	326
7.1.2 未知參數估計	326
7.1.3 迴歸方程的顯著性檢驗	329
7.1.4 利用迴歸方程進行預測	335
7.1.5 一元非線性迴歸模型	337
7.2 多元線性迴歸分析	341
7.2.1 多元線性迴歸分析的基本定義	341
7.2.2 矩陣表示法	342
7.2.3 未知參數估計	343
7.2.4 誤差方差 σ^2 的估計	344
7.2.5 有關的統計推斷	345
7.3 偏最小二乘迴歸分析	361
7.3.1 偏最小二乘迴歸方法的資料結構與建模思想	361
7.3.2 偏最小二乘迴歸方法的演算法步驟	362
7.3.3 偏最小二乘迴歸方法的輔助分析	365

8

多元統計分析 373

8.1 引言	374
--------	-----



8.2	因素分析	376
8.2.1	因素分析的理論介紹	376
8.2.2	因素分析的函數介紹	377
8.2.3	因素分析的應用示例分析	380
8.3	聚類分析	383
8.3.1	聚類分析的理論介紹	383
8.3.2	聚類分析的函數介紹	385
8.3.3	聚類分析的應用示例分析	391
8.4	正交實驗設計分析	394
8.4.1	正交表分析	395
8.4.2	不考慮交互作用正交實驗設計的基本程式分析	401
8.4.3	正交實驗設計分析的應用示例分析	414
8.5	多元方差分析	420
8.5.1	多元方差分析的理論介紹	420
8.5.2	多元方差分析的函數介紹	420
8.5.3	多元方差分析的應用示例分析	424
8.6	判別分析	425
8.6.1	判別分析概述	425
8.6.2	馬氏距離	428
8.6.3	多圖形平均法	431
8.7	實驗設計分析	433
8.7.1	實驗設計分析的理論介紹	433
8.7.2	實驗設計分析的函數介紹	434
8.7.3	實驗設計分析的應用示例分析	436

9

隱藏式馬可夫模型及統計工具箱的示範程式	443
---------------------	-----

9.1	隱藏式馬可夫模型	444
-----	----------	-----

9.1.1	基本理論	444
9.1.2	相關函數介紹	449
9.1.3	HMM 在語音識別中的應用	457
9.2	示範程式	462
9.2.1	aocool 函數	462
9.2.2	distool 函數	467
9.2.3	polytool 函數	468
9.2.4	randtool 函數	471
9.2.5	robustdemo 函數	472
9.2.6	rsmdemo 函數	473

附 錄 477

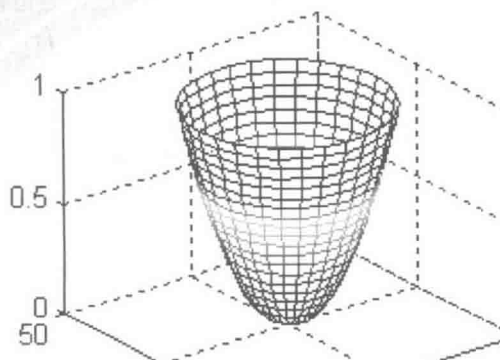
附錄 A	標準常態分佈函數表	478
附錄 B	χ^2 分佈上側分位點表	482
附錄 C	t 分佈上側分位點表	486

參考文獻 489

MATLAB 的資料基礎

學習重點

- 1.1 MATLAB 的主要功能
- 1.2 MATLAB 的程式





MATLAB 代表 Matrix Laboratory，是一個高性能的科學計算平台，整合了數值計算、矩陣計算和圖形繪製等眾多功能。在 MATLAB 中，問題的提出和解答只需要按照一般的數學方式表達和描述，不需要大量原始而傳統的編寫程式之過程，因此特別適用於研究、解決工程和數學問題。此外，MATLAB 還具有易擴展性，使用者都可以自行定義編寫函數或程式。

1.1 MATLAB 的主要功能

1.1.1 MATLAB 簡介

啟動 MATLAB 後，系統將自動打開指令視窗，如圖 1-1 所示。

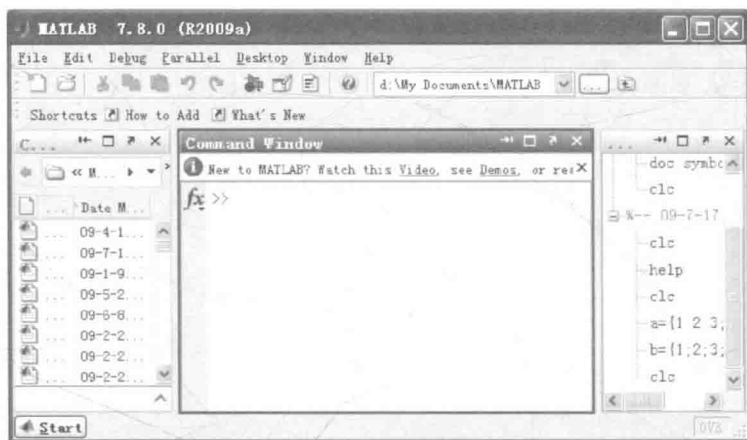


圖 1-1 MATLAB 指令視窗

MATLAB 按照設定的路徑啟動，可以經由 `pathtool` 函數改變路徑，亦即在指令視窗中進行「`>>pathtool`」即可以打開路徑設置對話方塊。圖中「`>>`」為 MATLAB 指令的提示符，顯示正在等待執行指令。此時，如果輸入相應的指令，MATLAB 就會執行，並得到結果。可以使用游標鍵 `<→`、`<↑`、`<↓` 或 `<→` 呼叫指令，以及在指令行中移動游標位置以修改指令。

點擊【File】功能表下【New】子功能表下的【M-file】選項，或者點擊工具欄中的  按鈕，則會彈出程式編輯視窗，如圖 1-2 所示。



MATLAB 的變數、注釋與標點、函數及 Script 檔案介紹如下。

1. 變數

MATLAB 變數的命名應該遵守一定的規則：變數以字母開頭（區分大小寫），之後可以是任意字母、數字或下畫線，但最長不能超過 36 個字元，也不能與 MATLAB 中的特殊變數（如 `ans`、`pi`、`eps`、`inf`、`NaN`、`i`、`j`、`nargin`、`nargout` 等）同名。



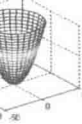
圖 1-2 MATLAB 程式編輯視窗

2. 注釋與標點

指令行中「%」符號後的所有文字為注釋，電腦不會執行。多條指令可以放在同一行，但是要使用逗號或者分號隔開。指令後的逗號表示顯示結果，分號則禁止顯示結果。符號「...」表示語句的餘下部分將出現在下一行，但是它不能出現在變數名或運算符之間。

3. 函數

MATLAB 本身的函數稱為內部函數。內部函數支援常用的數學函數，書寫方式和一般型式相同，例如三角函數 $\sin(x)$ 、常用對數 $\lg(x)$ 、自然對數 $\log(x)$ 、指數函數 $\exp(x)$ ，開方根 $\sqrt{x}$ 等。使用者還可以編寫各種自定義的函數，然後像 MATLAB 內部函數一樣，在工作環境下、Script 檔案和其他函數中呼叫。



4. Script 檔案

將 MATLAB 指令放在一個檔案中，然後告訴 MATLAB 打開檔案並順次執行其中的指令，這個檔案被稱為 Script 檔案。它可藉由單擊【File】功能表下【New】子功能表下的【M-file】選項建立。Script 檔案具有全局性，檔案中的所有變數將在整個工作環境中有效。

Script 檔案可直接在編輯或工作視窗中執行，也可被其他 M 檔案和函數呼叫。在工作視窗直接輸入 Script 檔案名便可執行，而在編輯視窗中執行 Script 檔案需單擊【Debug】功能表下的【Save File and Run】指令，然後切換到工作視窗查看執行結果。

1.1.2 MATLAB 的資料及數值分析

1. 資料分析

MATLAB 在作資料分析時，如果輸入的是向量，運算是對整個向量進行的；如果輸入的是陣列（矩陣），則運算按列進行。

利用 MATLAB 可以進行資料的基本統計計算，如下列各種函數。運算時，如果使用格式中有 dim，則指明運算按照指定維數進行。

- (1) $\max(x, \text{dim})$ ：求最大元素。
- (2) $\min(x, \text{dim})$ ：求最小元素。
- (3) $\text{median}(x, \text{dim})$ ：求中位值。
- (4) $\text{mean}(x, \text{dim})$ ：求平均值。
- (5) $\text{std}(x, \text{flag})$ ：求標準差，flag 指明標準差的不同計算方式。
- (6) $\text{prod}(x, \text{dim})$ ：求積。
- (7) $\text{sum}(x, \text{dim})$ ：求和。
- (8) $\text{cumsum}(x, \text{dim})$ ：求累計和。
- (9) $\text{cumprod}(x, \text{dim})$ ：求累計積。
- (10) $\text{cov}(x)$ ：求協方差陣。
- (11) $\text{cov}(x, y)$ ：求相關陣。
- (12) $\text{corrcoef}(x)$ ：求兩隨機變數的協方差。