

多變量分析

第二版

Multivariate
Analysis

An Introduction
Second Edition

黃俊英著

中華民國七十五年元月

多變量分析

第二版

**Multivariate
Analysis**

An Introduction
Second Edition

黃俊英著

中華民國七十五年元月

環球經濟雜誌社

台北市信義路四段199巷2號9樓之1
電話：7075908

多變量分析

中華民國七十四年二月初版
中華民國七十五年元月二版

著者：黃俊英

出版者：中國經濟企業研究所

台北市忠孝東路四段233號10之17
電話：7525463

總經銷：華泰圖書文物公司

台北市金華街181-1號
電話：3936633・3416633

登記證：局版台業字第1201號

印刷者：雨利美術印刷有限公司

(定價：350元)

二版序

多變量分析是描述、分析和處理多變量資料的統計方法。由於管理科學、社會科學和行為科學的問題常涉及多變量的資料，多變量分析技術的應用是非常重要的。近些年來電腦軟體程式的發展與普及，已使多變量分析技術的繁複演算過程不再構成研究人員使用此種技術的阻力，多變量分析已逐漸成為研究人員的「工具箱」中一套不可或缺的工具。

本書是應用導向的（ applications-oriented ），也是使用者導向的（ users-oriented ）。對一個應用多變量分析的研究人員而言，最重要的是要認識各種多變量分析技術的優點和缺點，了解各種不同技術的適用場合，能夠針對研究問題的性質和資料的特徵選擇適當的技術，並能對分析結果提出適切的解釋。本書對各種分析技術的介紹和討論，都是針對使用者的需要，強調各種技術的基本觀念與應用，儘可能避免繁瑣的數學演算，任何一位研究人員，只要修讀過應用統計的課程，並具備最基本的矩陣代數知識，就能夠輕易地了解本書的內容。

多變量分析已被廣泛地應用到社會科學、行為科學及管理科學的領域。由於筆者係一修讀企業管理的學生，故本書所舉的應用實例皆屬於企業管理的領域。為了方便讀者對本書內容做進一步探討研究的需要，凡書中引用及提及的有關著作及研究論文，均在各章之後的附註中註明其出處，以便查考。在本書之後，並列有中文索引，以便讀者查閱。

筆者於民國63年自美返國後，即開始在政治大學企管研究所開授「多變量分析」課程；亦曾在台灣大學及淡江大學開授相關課程。爲應教學需要，並供企業界人士之參考，本書初版於去年2月出版。爲使本書內容能更切合教學、研究以及企業應用的需要，乃根據初版架構充實內容，增訂再版。同時，鑒於近幾年來，國內管理研究所的博士和碩士論文應用多變量分析技術進行研究者日益增多，因此，本版中增列若干國內管理方面的應用實例，以反映多變量分析在國內管理學術研究之應用日益普及的事實。本書之出版，吾師楊必立教授的鼓勵以及在出版方面的協助，筆者衷心感謝。

本書是筆者教學和研究的一點心得，希望能有助於多變量分析在國內的推廣。多變量分析的發展甚爲快速，書中謬誤或遺漏之處恐將難免，尙請各界先進及讀者不吝指正，俾能於三版時補正。

黃俊英 謹識

民國 75 年 1 月

政治大學企業管理研究所

目 錄

第一章 基本概念	1
壹、資料的類型.....	2
貳、多變量分析技術的分類.....	4
第二章 變異數與共變數分析	11
壹、變異數分析.....	11
貳、二因素變異數分析.....	18
叁、共變數分析.....	26
第三章 多變項變異數及共變數分析	33
壹、多變項變異數分析.....	34
貳、多變項共變數分析.....	43
第四章 複迴歸	47
壹、基本概念.....	47
貳、簡單迴歸模式.....	48
叁、複迴歸模式.....	52
肆、判定係數與相關係數.....	54
伍、直線迴歸模式的基本假設.....	59
陸、複迴歸模式的檢定.....	62
柒、自相關與複共線性.....	65

捌、虛變數複迴歸模式·····	69
玖、選取預測變數的方法·····	72
拾、應用實例·····	77
附錄 4 — 1：逐步迴歸的運算過程·····	84

第五章 自動互動檢視法····· 93

壹、基本概念·····	93
貳、應用實例·····	100
叁、優點與限制·····	110
肆、與其他方法之結合·····	112

第六章 區別分析····· 121

壹、基本概念·····	121
貳、幾何圖形的描述·····	122
叁、二群體區別函數·····	125
肆、複區別分析·····	140
伍、顯著性檢定·····	143
陸、預測變數的相對重要性·····	146
柒、應用實例·····	148

第七章 聯合分析····· 157

壹、基本概念·····	158
貳、偏好模式的選擇·····	160
叁、資料收集的方法·····	163
肆、受測體的描述·····	168
伍、優點與限制·····	169

陸、應用實例	171
--------	-----

附錄 7 - 1：對稱直交排列法	182
------------------	-----

第八章 規則相關分析 187

壹、基本概念	187
--------	-----

貳、規則相關係數	188
----------	-----

叁、顯著性檢定	192
---------	-----

肆、重疊指數	195
--------	-----

伍、應用實例	198
--------	-----

第九章 因素分析 207

壹、基本概念	207
--------	-----

貳、巴氏的球形檢定	215
-----------	-----

叁、共同性的估計	217
----------	-----

肆、共同因素之抽取	219
-----------	-----

伍、因素的轉軸	227
---------	-----

陸、結果之解釋	234
---------	-----

柒、應用實例	236
--------	-----

附錄 9 - 1：重心法之運算過程	244
-------------------	-----

第十章 集群分析 253

壹、基本概念	253
--------	-----

貳、相似性的衡量	254
----------	-----

叁、建立集群的方法	259
-----------	-----

肆、層次集群方法的演算	261
-------------	-----

伍、非層次集群方法的演算	271
--------------	-----

陸、集群方法的比較.....	274
柒、二階段集群分析.....	275
捌、集群之描述及顯著性檢定.....	277
玖、應用實例.....	281

第十一章 多元尺度法..... 291

壹、多元尺度法的分類.....	291
貳、基本運算步驟.....	294
叁、計量多元尺度法.....	296
肆、非計量多元尺度法.....	299
伍、相似及偏好資料的收集.....	304
陸、單調迴歸.....	309

附 表 313

1 常態曲綫面積.....	313
2 t 值表.....	314
3 卡方 (χ^2) 分配表.....	315
4 F 分配表.....	316
5 杜賓—瓦特生 (DW) 檢定表.....	318

索 引 319

第一章 基本概念

多變量分析 (Multivariate Analysis) 又稱多變量統計分析 (Multivariate Statistical Analysis) ,是分析多變量資料的統計方法。多變量分析所處理的資料常包含許多變數的數值,演算的過程繁雜,通常非借助於電子計算機不可。由於電腦技術的發展與電腦知識的普及,加上電腦軟體程式的不斷開發,已使多變量分析技術在社會科學、行為科學和管理科學等領域中的應用日益普遍,多變量分析已逐漸成為從事社會科學、行為科學和管理科學研究工作人員必須通曉的一套統計技術。

多變量分析是協助研究人員從事科學調查研究的一種工具。科學調查研究的目的包括：

- 1 資料縮減或結構簡化：將所要研究的現象儘可能用簡單的結構來表達,而不致於犧牲有價值的資訊,以利解釋。

- 2 分類和分群：根據某些特性,將「相似的」事物 (objects) 集合成一群體。

- 3 調查變數間的相依關係：所有的變數都是相互獨立或是具有相依關係？

- 4 預測 (prediction)：建立變數間的關係,俾能根據某些變數的觀察值來預測其他變數的數值。

- 5 假設的建立和檢定：檢定特定的統計假設。(註 1)

在本書以後各章所要介紹的各種多變量分析技術將可幫助研究人員達成上列的五項目的。

壹、資料的類型

資料矩陣 (data matrix) 是多變量分析的核心，資料矩陣如表 1 所示，是將一群數字排列成長方形的形狀。資料矩陣 (表 1) 包括 n 個事物 ($i = 1, 2, 3, \dots, n$)， m 個變數 ($j = 1, 2, 3, \dots, m$)， X_{ij} 代表第 i 個事物在第 j 個變數上的數值。為了便於了解所有原始數字的意義，我們通常都先利用某些彙總性的統計值，如平均數 (means)、變異數 (variances) 和共變數 (covariances) 等，來描述原始的資料。

表 1 資料矩陣

事物 (i)	變 數 (j)				
	1	2	3		m
1	X_{11}	X_{12}	X_{13}		X_{1m}
2	X_{21}	X_{22}	X_{23}		X_{2m}
3	X_{31}	X_{32}	X_{33}		X_{3m}
4	X_{41}	X_{42}	X_{43}		X_{4m}
⋮	⋮	⋮	⋮		⋮
⋮	⋮	⋮	⋮		⋮
⋮	⋮	⋮	⋮		⋮
⋮	⋮	⋮	⋮		⋮
n	X_{n1}	X_{n2}	X_{n3}		X_{nm}

X_{ij} 可能是名目尺度 (nominal scale, 亦稱名目量表)、順序尺度 (ordinal scale, 亦稱順序量表)、區間尺度 (interval scale, 亦稱區間量表) 或比率尺度 (ratio scale, 亦稱比率量表) 的資料。

一、名目尺度

名目尺度係爲了標示目的而指定的數字，譬如國民身份證的總號碼、學生的學號、或以 0 代表男性、以 1 代表女性等等。名目尺度彼此間無順序關係，其加減乘除的運算也毫無意義。

二、順序尺度

順序尺度可表示各事物間的等級或順序，但不表示不同順序或等級間的差異程度，換言之，它只能指出等級或順序，但不能衡量不同等級間的距離。譬如，可依據受訪者的某些特性，如對某品牌的喜愛程度，將受訪者按順序排列，也可依照消費者的偏好程度，將產品、品牌或商店排列。

三、區間尺度

區間尺度具有一個相同的衡量單位，不僅可表示順序或等級，還可表示不同等級間的距離。在區間尺度上的第 n 個點與第 $n + 1$ 個點間的距離，相等於第 $n - 1$ 個點與第 n 個點間的距離。區間尺度的零點因係任意選定，因此，我們不能說第 $2n$ 個點的強度比第 n 個點高（或低）一倍。譬如華氏或攝氏溫度即爲區間尺度，溫度從攝氏 10° 上升到攝氏 20° ，並不表示溫度上升了一

4 多變量分析

倍。

四、比率尺度

比率尺度與區間尺度的區別在於前者有一個獨一無二的零點，後者的零點則係自行選定。長度和重量都是比率尺度，因所有距離都是從一個絕對的零點算起，因此 2 公尺就是比 1 公尺長一倍，4 公斤就是比 2 公斤重一倍。

表 2 列舉各種尺度的基本比較、典型例子、平均數和顯著性檢定等重要特性（註 2）。

貳、多變量分析技術的分類

多變量分析技術，誠如 R. Johnson 和 D. Wichern 所言，是一個「混合袋」（mixed bag）（註 3），要將多變量分析技術做適當的分類是不容易的。不過已有若干學者根據不同的基準，提出一些不同的分類架構和分類方法。譬如，Paul Green 曾根據（1）研究的重點是事物或變數，（2）資料矩陣是否劃分為準則變數（criterion）和預測變數（predictor）兩部分以及各部分的變數數目，（3）資料是名目、順序或區間尺度的數值等三項基準，將多變量分析技術區分成四類，即（1）單準則變數和複預測變數結合的技術（single criterion, multiple predictor association），（2）複準則變數和複預測變數結合的技術（multiple criterion, multiple predictor association），（3）變數互依分析（analysis of variable interdependence）和（4）事物間相似分析（

表 2 四種尺度的比較

尺度	基本比較	典型例子	典型性的統計		
			平均數	顯著性	檢定
名目	本身	男一女 使用者—非使用者 職業	衆數	卡方 (χ^2) McNemar Cochran Q	
順序	次序	品牌偏好 社會階層 材料的硬度 木材的分類品質	中位數	Mann Whitney U Kruskal Wallis 等級相關	檢定 檢定
區間	區間的比較	溫度 平均成績 對品牌的態度 廣告的知曉程度	算術平均數	z 檢定 t 檢定 變異數分析 相關	
比率	絕對大小的比較	銷售量 買者人數 購買機率 重量	幾何平均數 調和平均數		

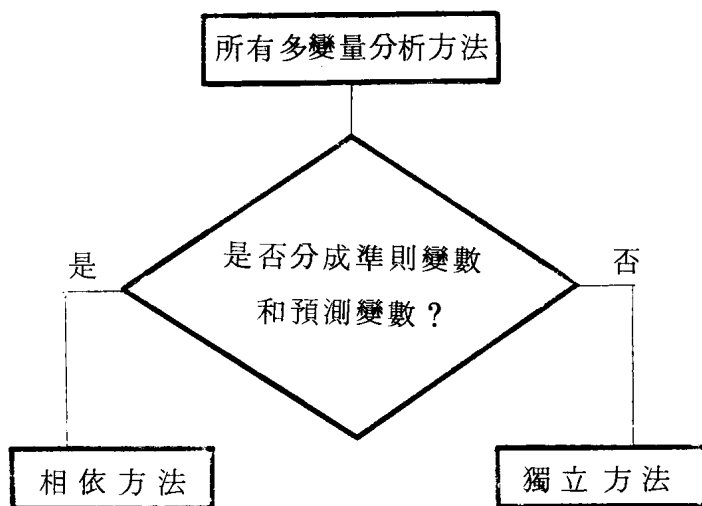
6 多變量分析

analysis of interobject similarity) (註 4) ; Jagdish Sheth曾以是否有某些變數相依於其他變數、相依變數的數目及投入資料是否為計量的資料等三個基準，提出一種分類的架構 (註 5) 。David Aaker 也曾根據是否有相依變數，而將所有多變量分析方法分為相依分析 (analysis of dependence) 和互依分析 (analysis of interdependence) 兩大類 (註 6) 。

前述準則變數常稱為相依變數 (dependent variable，亦稱因變數)，預測變數常稱為獨立變數 (independent variable，亦稱為自變數)。因相依變數 (或因變數) 與獨立變數 (或自變數) 有時會被一般人誤為具有因果關係，因此本書採用準則變數和預測變數這兩個名詞。

在分類方面，本書先根據是否將資料矩陣分為預測和準則變數這個基準，將所有多變量分析方法分為相依方法 (dependence methods) 和獨立方法 (independence methods) 兩大類，如圖 1 所示。

圖 1 多變數分析之分類



一、相依方法

關於相依方法的分類，T. Kinner和J. Taylor曾提出一種有用的分類方法。（註⁷）他們根據準則變數的數目、準則變數的尺度水準、預測變數的尺度水準等三項基準進行分類，如圖2所示。本書參採P. Green的分類精神，參照Kinner和Taylor的分類方法，將相依方法分成以下兩類：

（一）單準則變數和複預測變數結合的技術，包括複迴歸（multiple regression）、變異數和共變數分析（analysis of variance and covariance）、聯合分析（conjoint analysis或conjoint measurement）、二群體區別分析（two-group discriminant analysis）。

（二）複準則變數和複預測變數結合的技術，包括規則相關（canonical correlation）、多變項變異數和共變數分析（multivariate analysis of variance and covariance）、複區別分析（multiple discriminant analysis）。

二、獨立方法

關於獨立方法的分類，本書根據研究重點究係在資料矩陣的變數或事物這項基準，將之分爲：（見圖3）

（一）互依分析技術，包括因素分析（factor analysis）、多元尺度法（multidimensional scaling）、和其他減少構面的方法。

（二）相似分析技術，包括集群分析（cluster analysis）和其他將事物依相似性分群的方法。

圖 2 相依方法的分類

