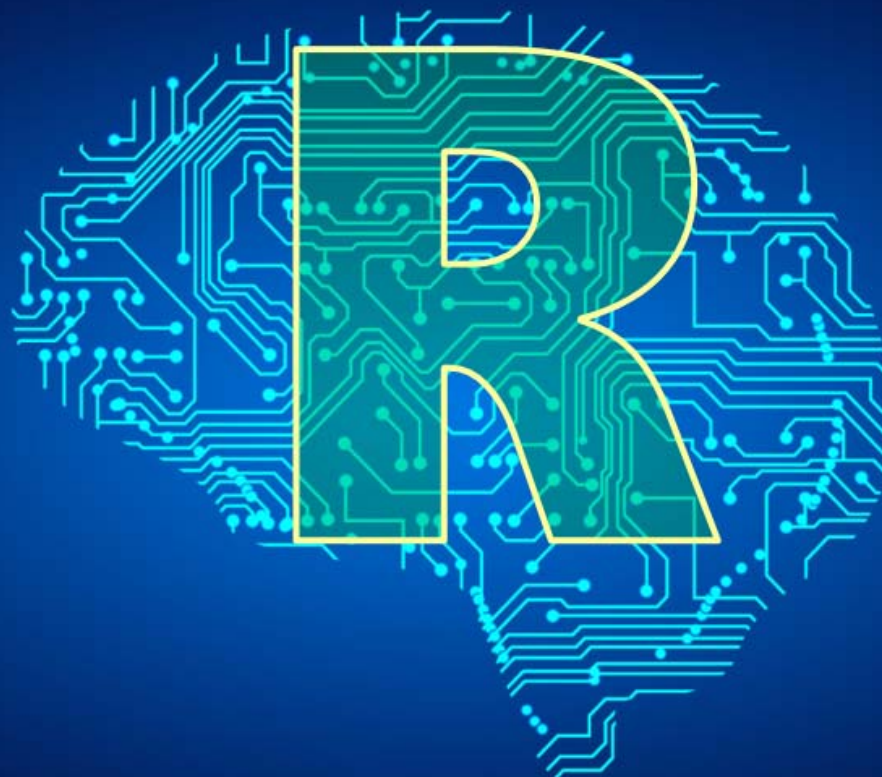


# 機器學習(ML) 深度學習(DL) 人工智慧(AI) 的 數學基礎



吳漢銘

國立臺北大學 統計學系

<http://www.hmwu.idv.tw>

# "Artificial Intelligence is All About Math" 2/144

科學月刊  
SCIENCE MONTHLY

關於科月 | 購買科月 | 訂閱科月 | 投稿須知 | 各期

封面故事 News Focus 專訪 專欄 科技報導 評論 精選文章與其他 活動訊息

2018年3月27日

## 人工智慧浪潮下的數學教育

魏澤人／任教於國立東華大學，創立花蓮-py社群及實做數學粉專。

我們曾學過的數學，究竟對人生有什麼幫助？

第一個尷尬點是，數學的實用性變得十分明顯，明顯到令人尷尬。學數學的人，常會聽到人問：「學數學有什麼用？」、「我高中數學都忘光了，還不是活得好好的？」、「工作上好像完全沒用到。」、「我寫程式這麼久了，也沒用到什麼數學。」

當然，內行人都知道數學在科學、技術、工程中，應用十分廣泛。特別在網路時代，網路加密、電腦運算、影像壓縮，甚至網頁和應用程式的自動排版都得用到數學。即使連照片編修這種屬於藝術文化的活動，其中的圖層操作就包含了向量概念，調色盤會看到16進位

及顏色轉換的線。電腦工程師們重新拿起統計、微積分及線性代數課本，想要了解現代的人工智慧在玩什麼把戲。現代人工智慧的領軍人物之一勒丘恩（Yann LeCun）說「人工智慧就是數學了傅立葉轉換和（artificial intelligence is all about math）」，他給想從深入人工智慧領域大學生的建議是：「如果在『iOS程式設計』及『量子力學』中要選一門課來修的話，選量子力學，且一定要選修微積分一、微積分二、微積分三……、線性代數、機率與統計，和盡可能的多選物理課程。即便如此，最重要的還是要會寫程式。」



楊立昆

電腦科學家

楊立昆，法國籍計算機科學家，他在機器學習、計算機視覺、移動機器人和計算神經科學等領域都有很多貢獻。他最著名的工作是在光學字符識別和計算機視覺上使用卷積神經網絡，他也被稱為卷積網絡之父。他同Léon Bottou和Patrick Haffner等人創建了DjVu圖像壓縮技術。維基百科

出生資訊：1960年7月8日（60歲），法國蘇瓦西蘇蒙特莫朗西  
獲頒獎項：圖靈獎（2018）；AAAI Fellow（2019）；法國榮譽軍團勳章（2020）

論文：Modèles connexionnistes de l'apprentissage (1987)

模範生：沃伊切赫·薩倫巴

獲獎記錄：圖靈獎

學歷：巴黎第六大學（1983年-1987年），巴黎高等電子工程師學校，索邦大學



勒丘恩：「人工智慧完全是數學。」（Wikipedia）

## 每日頭條

[首頁](#)[健康](#)[娛樂](#)[時尚](#)[遊戲](#)[3C](#)[親子](#)[文化](#)[歷史](#)[動](#)

## AI（人工智慧）就是統計學？

2018-10-26 由 秦思生 產規劃 發表于 科技

諾貝爾經濟學獎托馬斯·薩金特在《財經》世界科技創新論壇上的演講中說過一句話：人工智慧首先是一些很華麗的辭藻。人工智慧其實就是統計學，只不過用了一個很華麗的辭藻，其實就是統計學。



<https://kknews.cc/zh-tw/tech/5z36z28.html>

<http://www.hmwu.idv.tw>

## AI 不過是統計學

Thomas J. Sargent：人工智慧只是統計學的延伸

2011年諾貝爾經濟學獎取得者Thomas J. Sargent在題為“共享全球智慧 引領未來科技”的世界科技創新論壇上表示：

計算機是非常擅長計算，它們可以非常快速地完成計算人算不了的東西，但最終必須由人來組織和分析這些計算。你可以看一些非常成功的人工智慧應用，它不僅是機器在「思考」，也是科學家在思考。像 AlphaGo 的演算法看上去是第一次出現，但其實有很多非常聰明的數學，並且是由人設定教學內容。人工智慧是由機器和人分飾兩角的，非常有趣。

AI時代的中層支柱：統計學: <https://www.mdeditor.tw/pl/2nBY/zh-tw>

但是為什麼不說搞統計學呢？很簡單，因為不如人工智慧說法高大上，為什麼要高大上？因為高大上有人投錢。人工智慧代表了最新科技、最熱行業，能吸引投資，你要說是做統計學的誰投錢？誰買單？實際上這些工作很早就有人研究，只是那時候都歸類於統計學領域。

就像顯示系統縱橫位置指示器就是滑鼠；人體表皮污垢學就是搓灰；人體表皮死細胞分離器就是搓澡巾；智能高端數字通訊設備表面高分子化合物線性處理就是手機貼膜……

所以現在大家都學聰明了，那怕是老生常談也得包裝一個好聽的名稱，你說搞機器學習、深度學習就有人投錢，有人出大價錢挖你，你要說搞統計學大家立馬就不感興趣了，其實做的還是一回事。當然資本也是知道這些道理的，那為什麼還要投錢，因為資本是逐利的，投錢是為了掙更多錢，高大上的外衣就是掙錢的保障之一。

人工智慧和統計學不能完全劃等號

人工智慧和統計學存在莫大的關係，或者說統計學是人工智慧的最重要的理論基礎，但統計學和人工智慧依然有著很大不同，更不是一回事。



## 郭董指數學很重要 台師大教授：數學人才出路好的時代來了



2019-03-01 21:28 聯合報 記者陳智華／即時報導 讚 7,878 分享

在美國，數學家的出路非常好，台灣這樣的時代也要來了嗎？

台師大電機工師系助理教授、數學專欄作家賴以威今在臉書貼文指出，鴻海集團董事長郭台銘說：「數學很重要！」他因之前受邀鴻海跟高階主管演講時，親耳聽郭董對總部3、4百位高階主管、十幾個遠端連線的各地分公司主管員工這麼說。

賴以威聽到郭董跟員工這樣說時，才知道郭董找他去演講是為了推廣數學，推廣成長型數學思維。

郭董指出，工業物聯網、人工智慧、資料分析，這些鴻海現今著重的科技領域，背後都需要數學。因此，讓集團意識到數學的重要性，知道如何學好數學的心態是非常重要的。

他表示，演講後的隔天，郭董跟他聊到，鴻海很歡迎電機系和數學系的人才加入，與相關的產學合作。他這兩天認識許多裡面的同事、主管也是相關科系。

賴以威說，美國就業網站 CareerCast 曾經統計過，在美國數學人才有非常好的工作機會。在台灣，看來這樣的時代也要來臨了。

但賴以威強調，數學很重要，這指的不是程序性的計算，而是懂得活用的數感。

賴以威說：「郭台銘董事長都這樣說了，你不覺得嗎？」

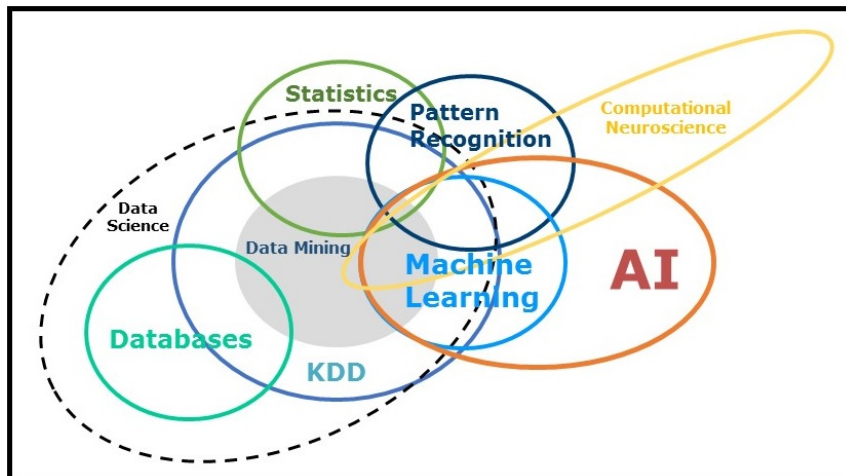


<https://udn.com/news/story/7266/3672385>

# Statistics, Data Mining, Machine Learning, Deep Learning, and AI

5/144

- **Statistics (STAT):** Statistics is the discipline that concerns the collection, organization, analysis, interpretation and presentation of data.
- **Data Mining (DM):** Data mining is a process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.
- **Machine Learning (ML):** Machine learning (ML) is the study of computer algorithms that improve automatically through experience. Machine learning algorithms build a mathematical model based on the training data, in order to make predictions or decisions.
- **Deep learning (DL):** Deep learning is part of a broader family of machine learning methods based on artificial neural networks with representation learning.
- **Artificial intelligence (AL):** the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings.

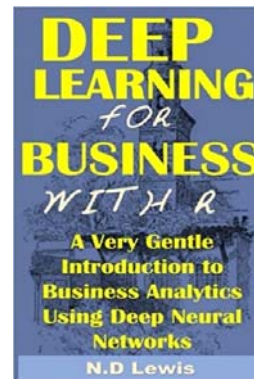
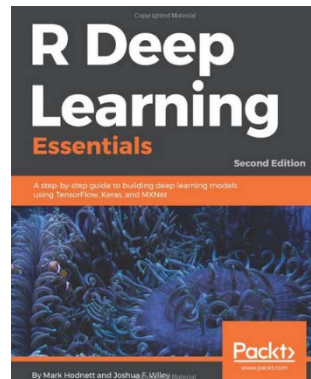
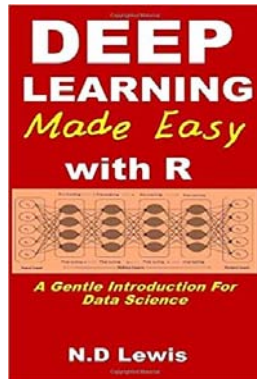


- **人工智慧**是讓機器具有如同人類甚至更多的思辨能力。
- **機器學習**則是能夠達成人工智慧的方法，透過與人類相似的學習方法，訓練機器進行資料分類、處理與預測。
- **深度學習**代表實現機器學習的一種技術。

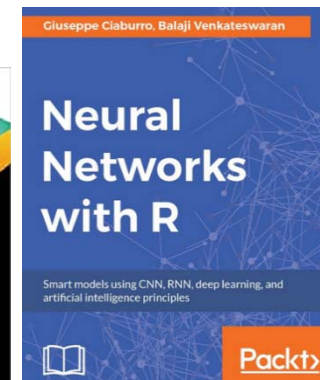
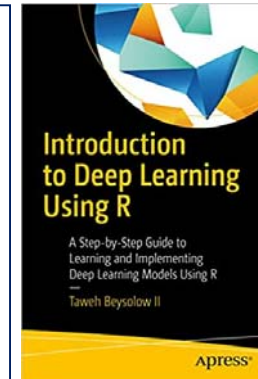
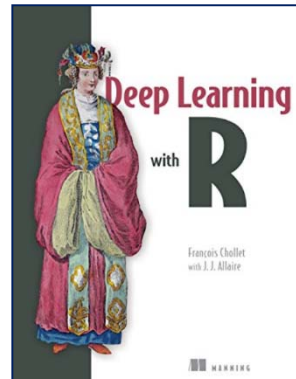
<https://jamesmccaffrey.wordpress.com/2016/09/29/machine-learning-data-science-and-statistics/>

# Deep Learning with R

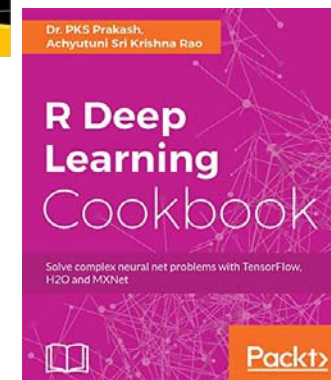
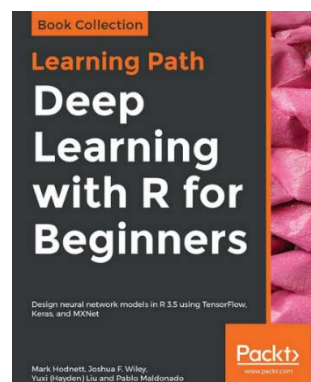
2016



2017



2018



2020



# 書籍：深度學習的數學

7/144



[日] 涌井良幸、涌井貞美  
**深度學習的數學**  
出版商：人民郵電  
出版日期：2019-05-01  
語言：簡體中文



## 深度學習的數學

作者：涌井良幸, 涌井貞美  
譯者：楊瑞龍  
出版社：博碩  
出版日期：2020/05/04

### 第1章 神經網路的思想

- 1-1 神經網路和深度學習 2
- 1-2 神經元工作的數學表示 6
- 1-3 啟動函數：將神經元的工作一般化 12
- 1-4 什麼是神經網路 18
- 1-5 用惡魔來講解神經網路的結構 23
- 1-6 將惡魔的工作翻譯為神經網路的語言 31
- 1-7 網路自學習的神經網路 36

### 第2章 神經網路的數學基礎

- 2-1 神經網路所需的函數 40
- 2-2 有助於理解神經網路的數列和遞推關係式 46
- 2-3 神經網路中經常用到的 $\Sigma$ 符號 51
- 2-4 有助於理解神經網路的向量基礎 53
- 2-5 有助於理解神經網路的矩陣基礎 61
- 2-6 神經網路的導數基礎 65
- 2-7 神經網路的偏導數基礎 72
- 2-8 誤差反向傳播法必需的鏈式法則 76
- 2-9 梯度下降法的基礎：多變數函數的近似公式 80
- 2-10 梯度下降法的含義與公式 83
- 2-11 用Excel 體驗梯度下降法 91
- 2-12 最優化問題和回歸分析 94

### 第3章 神經網路的最優化

- 3-1 神經網路的參數和變數 102
- 3-2 神經網路的變數的關係式 111
- 3-3 學習數據和正解 114
- 3-4 神經網路的代價函數 119
- 3-5 用Excel體驗神經網路 127

### 第4章 神經網路和誤差反向傳播法

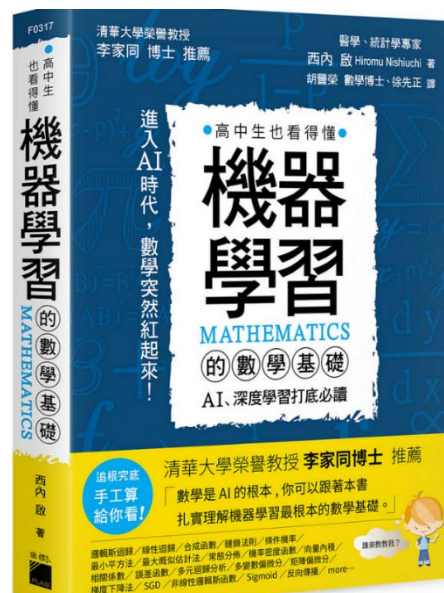
- 4-1 梯度下降法的回顧 134
- 4-2 神經單元誤差 141
- 4-3 神經網路和誤差反向傳播法 146
- 4-4 用Excel體驗神經網路的誤差反向傳播法 153

### 第5章 深度學習和卷積神經網路

- 5-1 小惡魔來講解卷積神經網路的結構 168
- 5-2 將小惡魔的工作翻譯為卷積神經網路的語言 174
- 5-3 卷積神經網路的變數關係式 180
- 5-4 用Excel體驗卷積神經網路 193
- 5-5 卷積神經網路和誤差反向傳播法 200
- 5-6 用Excel體驗卷積神經網路的誤差反向傳播法 212

# 書籍：機器學習的數學基礎

8/144



**機器學習的數學基礎：**  
**AI、深度學習打底必讀**  
醫學統計學專家 西內啟 著  
胡豐榮博士, 徐先正 合譯，  
出版商: 旗標科技出版  
(2020-01-31)

## 目錄：

序篇 AI、機器學習需要什麼樣的數學能力  
單元01 21世紀每個人都需要具備數學能力  
單元02 數學金字塔

## 第1篇 機器學習的數學基礎

單元03 將事物用數字來表現  
單元04 將數字用字母符號代替  
單元05 減法是負數的加法, 除法是倒數的乘法  
單元06 機率先修班：集合  
單元07 機率先修班：命題的邏輯推理  
單元08 機率、條件機率與貝氏定理

## 第2篇 機器學習需要的一次函數與二次函數

單元09 座標圖與函數  
單元10 聯立方程式求解與找出直線的斜率與截距  
單元11 用聯立不等式做線性規劃  
單元12 從線性函數進入二次函數  
單元13 利用二次函數標準式求出最大值與最小值  
單元14 找出二次函數最適當的解  
單元15 用最小平方方法找出誤差最小的直線

## 第3篇 機械學習需要的二項式定理、對數、三角函數

單元16 二項式定理與二項式係數  
單元17 利用二項分布計算重複事件發生的機率  
單元18 指數運算規則與指數函數圖形  
單元19 用對數的觀念處理大數字  
單元20 對數的性質與運算規則  
單元21 尤拉數 e 與邏輯斯迴歸  
單元22 畢氏定理計算兩點距離  
單元23 三角函數的基本觀念  
單元24 三角函數的弧度制與單位圓

## 第4篇 機械學習需要的 $\Sigma$ 、向量、矩陣

單元25 整合大量數據的  $\Sigma$  運算規則  
單元26 向量基本運算規則  
單元27 向量的內積  
單元28 向量內積在計算相關係數的應用  
單元29 向量、矩陣與多元線性迴歸  
單元30 矩陣的運算規則  
單元31 轉置矩陣求解迴歸係數

## 第5篇 機器學習需要的微分與積分

單元32 函數微分找出極大值或極小值的位置  
單元33 n 次函數的微分  
單元34 積分基礎—從幾何學角度瞭解連續型機率密度函數  
單元35 積分基礎—用積分計算機率密度函數  
單元36 合成函數微分、鏈鎖法則與代換積分  
單元37 指數函數、對數函數的微分積分  
單元38 概似函數與最大概似估計法  
單元39 常態分佈的機率密度函數  
單元40 多變數積分—雙重積分算機率密度函數係數

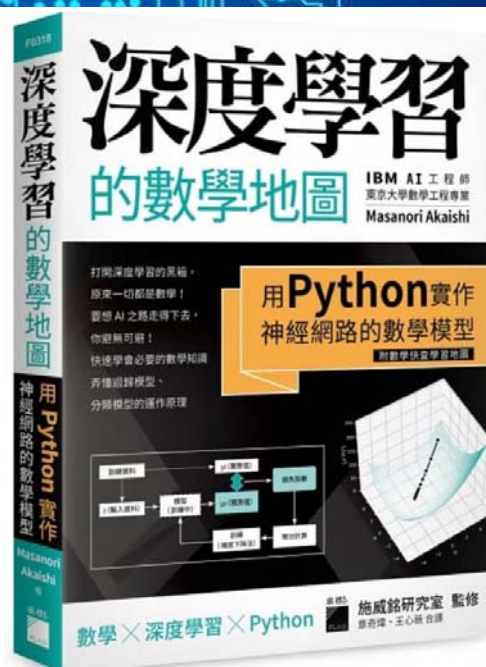
## 第6篇 深度學習需要的數學能力

單元41 多變數的偏微分—對誤差平方和的參數做偏微分  
單元42 矩陣型式的偏微分運算  
單元43 多元迴歸分析的最大概似估計法與梯度下降  
單元44 由線性迴歸瞭解深度學習的多層關係  
單元45 多變數邏輯斯迴歸與梯度下降法  
單元46 神經網路的基礎—用非線性邏輯斯函數組合出近似函數  
單元47 神經網路的數學表示法  
單元48 反向傳播—利用隨機梯度下降法與偏微分鏈鎖法則



# 書籍：深度學習的數學地圖

9/144



深度學習的數學地圖：  
用 Python 實作神經網路的數學模型  
作者：Masanori Akaishi  
譯者：章奇煒, 王心薇  
出版社：旗標  
出版日期：2020/05/28

## 第6章 機率、統計

- 6.1 隨機變數與機率分佈
- 6.2 機率密度函數與累積分佈函數  
專欄 Sigmoid 函數的機率密度函數
- 6.3 概似函數與最大概似估計法  
專欄 為何概似函數的極值是求最大值，而不是最小值？

## 目錄

### 【導入篇 機器學習快速指引】

#### 第1章 機器學習入門

- 1.1 何謂機器學習
  - 1.1.1 何謂機器學習模型
  - 1.1.2 機器學習的訓練方法
  - 1.1.3 監督式學習的迴歸、分類模型
  - 1.1.4 訓練階段與預測階段
  - 1.1.5 損失函數與梯度下降法
- 1.2 第一個機器學習模型：簡單線性迴歸模型
- 1.3 本書討論的機器學習模型
- 1.4 數學是深度學習的核心
- 1.5 本書架構

#### 第4章 多變數函數的微分

- 4.1 多變數函數
- 4.2 偏微分
- 4.3 全微分
- 4.4 全微分與合成函數
- 4.5 梯度下降法 (GD)  
專欄 梯度下降法與局部最佳解

### 【實踐篇 機器學習、深度學習實作】

#### 第7章 線性迴歸模型（迴歸）

#### 第8章 邏輯斯迴歸模型（二元分類）

#### 第9章 邏輯斯迴歸模型（多類別分類）

#### 第10章 深度學習

### 【發展篇 實務上的解決方法】

#### 第11章 以實用的深度學習為目標

### 【理論篇 數學速學課程】

#### 第2章 微分、積分

- 2.1 函數
  - 2.1.1 函數運作行為
  - 2.1.2 函數的圖形
- 2.2 合成函數與反函數
  - 2.2.1 合成函數  
專欄 合成函數的表示法
  - 2.2.2 反函數
- 2.3 微分與極限
  - 2.3.1 微分的定義
  - 2.3.2 函數值增量與微分的關係
  - 2.3.3 切線方程式  
專欄 切線方程式與訓練階段、預測階段的關係
- 2.4 極大值與極小值
- 2.5 多項式的微分
  - 2.5.1  $x^n$  的微分 ( $n$  是正整數)
  - 2.5.2 微分計算的線性關係與多項式的微分
  - 2.5.3  $x^r$  的微分 ( $r$  是實數)

#### 第5章 指數函數、對數函數

- 5.1 指數函數
  - 5.1.1 連乘的定義與公式
  - 5.1.2 連乘觀念的推廣
  - 5.1.3 將連乘寫成指數函數形式
- 5.2 對數函數  
專欄 對數函數的意義
- 5.3 對數函數的微分  
專欄 用 Python 來計算尤拉數  $e$
- 5.4 指數函數的微分  
專欄 以  $e$  為底的指數函數也可用  $\exp$  表示
- 5.5 Sigmoid 函數
- 5.6 Softmax 函數  
專欄 Sigmoid 和 Softmax 函數的關係

#### 第3章 向量、矩陣

- 3.1 向量入門
  - 3.1.1 何謂向量
  - 3.1.2 向量的標記法
  - 3.1.3 向量的分量
  - 3.1.4 往多維擴展
  - 3.1.5 分量的符號
- 3.2 向量和、向量差、純量乘積
  - 3.2.1 向量和
  - 3.2.2 向量差
  - 3.2.3 向量與純量的乘積
- 3.3 向量的長度（絕對值）與距離
  - 3.3.1 向量的長度（絕對值）
  - 3.3.2  $\Sigma$  可整合冗長的加法算式
  - 3.3.3 向量間的距離
- 3.4 三角函數
  - 3.4.1 三角比：三角函數的基本定義
  - 3.4.2 單位圓上的座標
  - 3.4.3 三角函數的圖形
  - 3.4.4 用三角函數表示直角三角形的邊長
- 3.5 向量內積
  - 3.5.1 向量內積的幾何定義
  - 3.5.2 用分量來表示內積公式
- 3.6 餘弦相似性
  - 3.6.1 兩個二維向量的夾角
  - 3.6.2  $n$  維向量的餘弦相似性  
專欄 餘弦相似性的應用範例
- 3.7 矩陣運算
  - 3.7.1 一個輸出節點的內積表示法
  - 3.7.2 三個輸出節點的矩陣相乘



## 機器學習的數學： 用數學引領你走進AI的神秘世界

作者：孫博

譯者：博碩文化

出版社：博碩

出版日期：2020/09/09

### 目錄

第 1 章 向量和它的朋友們

第 2 章 矩陣的威力

第 3 章 距離

第 4 章 導數

第 5 章 微分與積分

第 6 章 弧長與曲面

第 7 章 偏導

第 8 章 多重積分

第 9 章 參數方程式

第 10 章 超越直角座標系

第 11 章 梯度下降

第 12 章 誤差與近似

第 13 章 牛頓法

第 14 章 無解之解

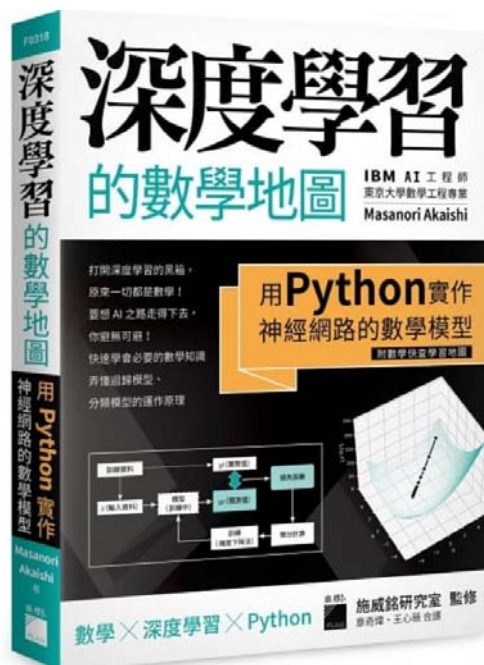
第 15 章 極大與極小

第 16 章 尋找最佳解

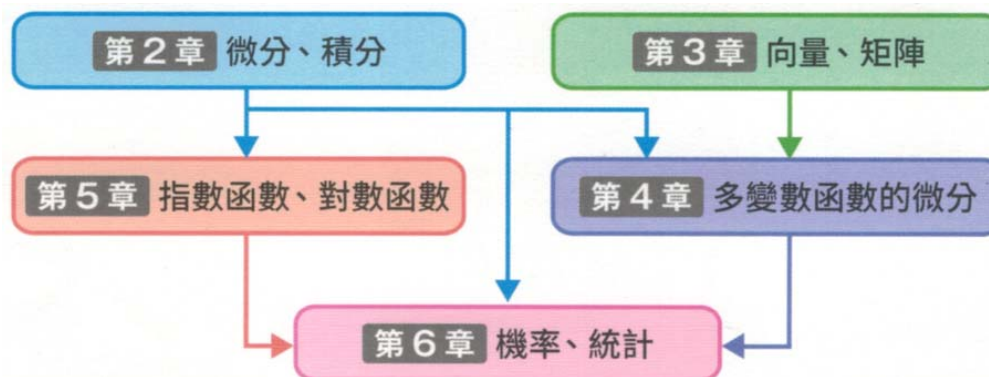
第 17 章 最佳形態

第 18 章 硬幣與骰子

第 19 章 機率分佈



深度學習的數學地圖：  
用 Python 實作神經網路的數學模型  
作者：Masanori Akaishi  
譯者：章奇煒, 王心薇  
出版社：旗標  
出版日期：2020/05/28



## 實踐篇 機器學習、深度學習實作

重點	第1章	第7章	第8章	第9章	第10章
實現深度學習所需概念	迴歸 1	迴歸 2	二元分類	多類別分類	深度學習
1 損失函數	○	○	○	○	○
3.7 矩陣運算				○	○
4.5 梯度下降法		○	○	○	○
5.5 Sigmoid 函數			○		○
5.6 Softmax 函數				○	○
6.3 概似函數與最大概似估計法			○	○	○
10 反向傳播					○

\* 迴歸 1 是指簡單線性迴歸，迴歸 2 是指多元線性迴歸



# 數學理論學習地圖

12/144

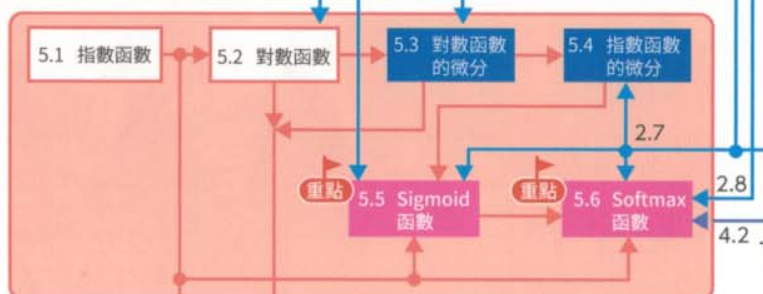
## 第2章 微分、積分



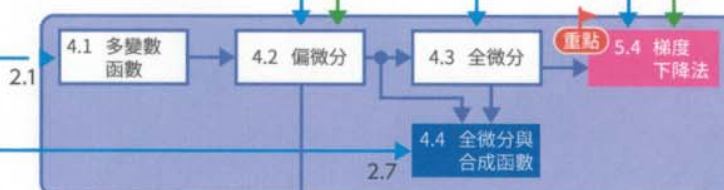
## 第3章 向量、矩陣



## 第5章 指數函數、對數函數



## 第4章 多變數函數的微分



## 第6章 機率、統計



實踐篇  
機器學習  
深度學習  
實作

## 第7章 線性迴歸模型(迴歸)

## 第8章 邏輯斯迴歸模型(二元分類)

## 第9章 邏輯斯迴歸模型(多類別分類)

## 第10章 深度學習



## 發展篇 實務上的解決方法 第11章 以實用的深度學習為目標