

# STEM 時代的數學觸角 (上)

## 從 18 項個案看數學在產業、科技、商業、醫學各方向的廣大應用

作者：美國工業與應用數學會 (SIAM) 譯者：中華民國數學會與臺灣工業與應用數學會

美國的 Society for Industrial and Applied Mathematics (SIAM) 成立於 1951 年，學會宗旨在結合數學、科技、工業各領域，並將焦點集中在應用數學、計算數學、工業數學，會員來源廣泛，目前人數已接近 15,000 人，並在世界各地設有分會。

在本文中，我們給出一個廣泛但不盡全面的，關於數學商業應用的綜覽。對於想要認識數學怎麼在「真實世界」應用的學生，我們希望以下 18 項個案研究可以提供一些答案。多數的個案都是實地訪查後才得知的，輔以已出版的文件作參考。

### 商業分析

軟體產業正在進行一個龐大的賭注... 他們認為資料導向決策是未來的趨勢。協助企業在資料中找尋有意義的資訊，則成為在軟體與服務業中所謂「商業智慧」或「商業分析」的一個快速成長的產業。科技大廠如 IBM、Oracle、SAP 與 Microsoft 總共花費了超過 250 億美金併購此領域的專業公司。[1]

「商業分析」(Business Analytics) 已成為無所不包的新詞，包含許多應用數學中生根已久的領域，像是作業研究與管理科學等等。但這個詞彙同時也有新的含義：對企業主管而言，巨大資料庫的應用愈來愈可行。

至少從 1950 年代開始，數學就一直應用在物流、倉儲與設施位置選擇上。美國空軍以及美國海軍研究辦公室的後勤規劃採用了丹齊 (George Dantzig)、卡魯席 (William Karush)、庫恩 (Harold Kuhn) 與杜克 (Albert Tucker) 早年針對最佳化的研究結果。這些最佳化技術諸如線性規劃及其變體，直到今日仍對業界至關重要。

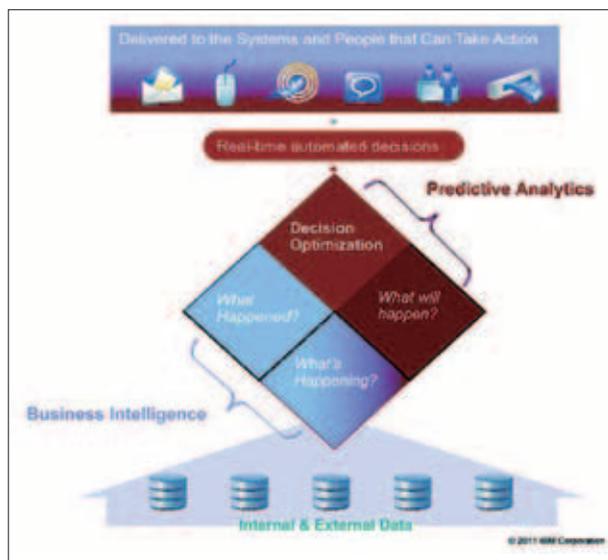
對於想要進入產業的企業及學生而言，機會就在於發展低成本的演算法與相關技術，使其能處理大量結構化或非結構化的資料。因為企業正在各個層面採納其商業情報 (即資料) 與分析結果 (即量化

方法)，包括行銷、人力資源、金融、供應鏈管理、設施位置選擇、風險管理和產品與流程設計等等。

### 個案 1：預測分析學

IBM 在 2009 與 2010 年協助紐約州稅務廳 (New York State Division of Taxation and Finance, DTF) 安裝一組全新的預測分析系統，該系統源於 IBM 極成功的西洋棋軟體「深藍」，以及電視節目《危險邊緣！》<sup>①</sup> 裡使用的遊戲引擎「華生」(Watson)。

這套系統稱為「稅捐稽徵最佳化解決方案」(Tax Collection Optimization Solution, TACOS)，可用來蒐集各種資料，包含國稅局的各種動作 (如撥打電話、探訪、追繳欠稅、徵稅、資產扣押) 及納



業務過程中的自動化決策。(IBM 提供)

① 譯註：Jeopardy! 機智問答節目。

稅人的回應（如付清稅款、申訴、宣告破產）。在實務上這些動作有時受到限制，例如人力有限或預算不足以架設電話客服中心等。TACOS 中的模型涵蓋這些活動間的相依性，能預測各種收稅策略可能造成的結果，像是通話與探訪的時間點等。模型採用的數學方法稱作馬可夫決策過程（Markov decision process），該流程會將每個納稅人賦予一個狀態，並預測在此狀態之下進行某個動作能獲得什麼報酬。依據這個過程得出的計畫或策略可最大化來自整個納稅族群預期收益，而非僅是單一納稅人的最大收益。

TACOS 系統使 DTF 在 2009 與 2010 年的收益增加了 8300 萬美金（增加 8%），但支出卻沒有增加。其中，每封追繳欠稅令（或稅務扣押權）所追回的美金增加了 22%，每一份扣押信函追回的美金增加了 11%，案件分派到各地辦公室所需的時間也減少了 9.3% [2]。縱使細節或許不同，但其他企業顯然也可以使用類似的方法處理收款與帳款。

### 個案 2：影像分析與資料探勘

SAIC 這間公司發展軍事用途的情報、監視、偵查（intelligence, surveillance and reconnaissance，ISR）系統。美國軍方在阿富汗戰爭裡大量布署這些自動系統。2009 年，無人航空載具（unmanned aerial vehicle，UAV）拍攝了足以放映 24 年份量的全動態影像，而且預計這些載具在 2011 年，將會再拍攝 30 倍的影像。

這產生一個大問題：影像要如何整理才能變成有用的資料？軍方當然不可能讓數千名士兵守在電腦螢幕前觀看數十年份的影像。即使真這麼做，人工

不僅容易犯錯也容易疲勞。連續監看數小時的影像可能會遺漏真正發生問題的關鍵片段，比如，曾涉嫌運送炸彈的車子正停在某棟屋子前面。

SAIC 研發的「後設資料」系統稱為 AIMES。這系統會向人們指出資料堆中的疑點。首先，AIMES 處理影像，補償並修正 UAV 因移動而造成的影像偏差，光是這點就是有趣的數學挑戰。接著，AIMES 會搜尋影像視野中的物體，並存入可搜尋的資料庫。AIMES 也會「融合」影像資料與其他類型資料。例如，若 UAV 的操作者說：「把畫面拉近那輛卡車！」。程式就會知道視野中的物體是一輛卡車，且該物體可能十分重要。再者，AIMES 方便攜帶布署在戰場上，只需要一台伺服器與二或三台螢幕即可使用（參閱 [3]）。

即便位居美國本土的產業不需如此顧慮恐怖分子或路邊炸彈，但影音監視仍是保護廠房或其他建築安全不可或缺的設備。相機加麥克風的組合還能用於其他用途：比如，麥克風能比操作者先偵知機器的異常，監視設備能讓首批抵達現場的人員找到火警或意外中的受害者。參閱 [4]。

### 個案 3：作業研究

美國第七大的混凝土公司維吉尼亞混凝土公司（Virginia Concrete），自 2002 年起運用最佳化軟體，為卡車駕駛安排運送順序。該公司擁有 120 輛卡車，供十座預拌廠使用。一個緊要的限制是混凝土在卡車裡大約兩小時就會乾掉，因此得在兩小時之內運送。而且，灌漿時程非常不可預測，常常一天之內就有 95% 的訂單會臨時調整。

維吉尼亞混凝土公司從喬治梅森大學（Goerge

Mason University) 與決定性分析公司 (Decisive Analytics Corporation) 引進一群數學家研發可自動派車的工具。數學家發現在舊的模型裡，每一輛卡車會被指派至一座「主要」預拌廠。若捨棄此模型，便能大幅減少混凝土運送時間。數學家因此建議卡車應該前往最近的預拌廠。至於跨夜排程，納入虛擬卡車的想法很有幫助，虛擬卡車代表可能被取消的訂單，而沒被取消的訂單則可重新指派至實體卡車。

為了測試，維吉尼亞混凝土公司用此軟體做所有的排程決策，然而派車人員仍然擁有變更電腦決策的權限。結果該系統讓該公司每一位駕駛所運送的混凝土量增加了 26%。[5]

### 金融數學

……市場可能越來越不重視非標準型金融衍生商品 (exotic derivative)，更多的交易將發生在匯兌上。未來的模型都應呈現符合觀察結果的真實動態，執行成本控管也至關重要。要順利控管，深入了解市場的微結構與交易資料是不可或缺的。[摘自訪談]

2007 與 2008 年的信用危機重挫了人們對金融量化方法的信心，多數專家將此解釋為量化模型無法正確詮釋市場資料的依存關係。風險模型假設兩地 (比如邁阿密和洛杉磯) 的房地產違約互不相關，或相關性很低；但實際上當恐慌發生時，所有事情之間的相關係數都變成 1 了<sup>2</sup>。

然而，經此危機震撼以及隨後的經濟衰退，金融經理人也上了很有價值的一課。他們學到了數學模型不可以隨便套用，必須嚴肅檢視模型背後的假

設。若干簡化模型的失敗並不代表數學模型都是糟糕的，而是模型必須更貼近現實。最重要的是，學生要了解金融產業不需捨棄量化研究。數學家與應用數學家的需求仍高居不下。隨著量化模型越來越複雜，以及經理人想知道這些模型的極限，數學家的專業只會越來越受到重視。只是，相較於以往，學生的數學技能應該要建立在更廣博的金融產業知識之上。

### 個案 4：算法交易 (自動化交易)

2009 年，霍夫 (Christian Hauff) 與艾姆葛連 (Robert Almgren) 離開美國銀行 (Bank of America) 這家世界頂尖的股票與衍生物算法交易公司，創立了一家叫做「量化經紀人」 (Quantitative Brokers) 的公司。他們當時發現一個難得的良機，打算將高頻交易的相同原則應用在尚未高度自動化的一類資產：利率期貨。

自動化交易在選擇權市場早已司空見慣。部份原因是金融數學迫切需要自動化交易這項工具。大銀行想要以風險中立的方式擁有資產：無論市場怎麼變化，都能讓銀行賺錢 (或至少避免虧損)。1970 年代早期，休斯 (Myron Scholes) 與布萊克 (Fischer Black) 發現了動態避險 (dynamic hedging) 策略以達成風險中立。此策略需要持續進行小規模交易。

布萊克與休斯考量的重點是選擇權定價。又過了 20 年，金融工程師才開始將執行交易的過程納入

<sup>2</sup> 譯註：意即完全正相關。



考量。不貿然進行全盤交易有許多原因：你也許想等等看是否有人願意出比較好的價錢，或者等到市場價格接近你的目標價之後再出手。如果你交易的資產在市場的當日交易佔了重要比例，你或許會審慎進行，以免過度影響市場價格。

「量化經紀人」這間公司的主要業務是執行交易。該公司運用電腦演算法，為客戶規劃策略路徑，從交易一開始的狀況導向客戶所期望的形勢（例如以低於 Y 價格購入 X 股的歐元期貨）。每一位客戶都有一定程度的風險規避，因此，客戶的效用函數是期望獲利與期望風險的線性組合。而「量化經紀人」的 STROBE 演算法能找出一條路線以最佳化客戶的效用函數。該演算法並能生成最佳值附近的一個包絡以概括出可接受的偏差範圍。其中所用的數學工具包括微分方程以及變分學。參閱 [6]。

### 系統生物學

製藥研究者嘗試了許多新作法與科技以應對研發藥品逐漸攀升的費用。生物標記、適應性試驗設計、建模、試驗模擬、預測性新陳代謝、資料探勘、疾病模型等等都重塑了研發的方式。量化製藥學提升了原來以模型為基礎的方法，運作於文化與科技的層次，整合了資料領域與科學領域……[7]

在 2000 年完成的人類基因組計劃應該引領我們進入一個發展個人化醫療與標靶藥物的新紀元。但

是我們最終發現，只有少數不尋常的疾病或其變種直接源於獨立的幾個基因突變。多數常見的疾病，像是糖尿病以及癌症（藥品研究的頭號目標），都肇因於複雜基因網絡的機能失常。修復單一基因就能治好這類疾病的想法變得有點天真，就像以為更換一顆螺絲就能修好引擎一樣。醫生實際上需要以特定的劑量、在特定的時間、於基因網絡的特定地點實施一連串的介入治療，才有可能治療這些疾病。隨著基因網絡的複雜度愈加明顯，分析複雜網絡的數學方法就變得益發重要。

若干生物科技的研究重點已經從基因體學（genomics）轉到「體學」（omics），像是蛋白質體學（proteomics），主要研究藥物所標靶的蛋白質的形狀與摺疊方式。分子動力學模擬則從最根本的層級著手，利用量子力學的原理進行研究。近幾年來，演算法、軟體、硬體的發展，在一毫秒內就能模擬幾千萬個原子組成的分子，而一毫秒正是許多重要生物過程發生的時間尺度。

其他數學模型則反其道而行，分析整個有機體。譬如說，這樣的模型可用來預測病人族群（其中的病人生理機能各自不同）在不同公衛手段介入下的反應。

最終，整體病人模型可能與基因組資料結合，為病人提供真正的個人化醫療服務。這些模型背後的數學與計算技術，包括網絡科學、決定性與隨機微分方程、貝氏網絡與隱式馬可夫模型、最佳化、統計、控制論、模擬論，以及不確定性的量化分析。

### 個案 5：分子動力學

蕭（David Shaw）是位電腦科學家。先前他是

某避險基金的執行長，一向以電腦系統執行交易策略。2001年，他建立一間全新的私人研究實驗室 D. E. Shaw Research (DESRES)，專注於研究蛋白質摺疊問題。蕭的實驗室量身訂作了一台名為 Anton 的超級電腦，配備了 512 顆晶片，用來加速計算原子間的交互作用。然而，即使有如此強大的硬體，如果使用蠻力 (brute-force) 模擬計算，仍然無法在合理時間內完成對一個蛋白質分子的模擬。因此另一個關鍵，是蕭實驗室裡的分子動力學軟體 Desmond，它使用明慎逼近法 (judicious approximation) 簡化力場的計算過程，也運用新平行演算法來減少 Anton 處理器之間的通訊量。

一開始沒人知道 Anton 是否比其他方法有效。比如別的演算法會將計算問題分割，分配至不同電腦進行 (這是史丹佛大學 Folding@Home 計畫採取的方法)。然而在 2010 年，DESRES 宣布他們已經完成了週期為 100 微秒的 FiP35 蛋白質摺疊與展開模擬。FiP35 有 13,564 個原子。與以前最好的程式結果相比，他們的模擬所涵蓋的時間長了十倍，程式大約花了三個星期進行模擬計算。

選擇 FiP35 是因為以往的實驗已經對該蛋白質摺疊和展開時的結構有相當的理解。即使如此，此模擬仍然帶來嶄新的科學洞見：每一次從摺疊到展開的途徑基本上是相同的。《科學》(Science) 雜誌將蕭的模擬成果列為橫跨所有科學領域的年度十大科學突破。相關模擬可能在未來開啟新一波的藥物與蛋白質交互作用的研究。由於這種交互作用發生得太快，科學家無法在傳統實驗室裡進行研究。參閱 [8]。

## 個案 6：全病人模型 (Whole-patient model)

在 2020 年之前，虛擬細胞、器官與動物都會廣泛應用在製藥研究中。[9]

Entelos 與 Archimedes 這兩家位於舊金山灣區的公司是全身電腦建模領域的先驅。雖然兩家公司離完成人體生理的完整模擬還很遠，但至少已經成功模擬了主要的子系統，像是心血管系統以及與糖尿病相關的新陳代謝網絡。

Entelos 的 PhysioLab 模型以及 Archimedes 模型能預測臨床藥物的不良反應與結果。顯而易見的，藥廠可以因此在投入臨床試驗之前先篩掉無效或有害的化合物，從而省下大量的時間和金錢。除此之外，模擬能探索多重藥物療法的效果，而這些效果極難透過臨床實驗測得。20 種不同的藥物組合可能需要 20 個不同的臨床試驗，但模擬實驗可以快速鎖定出最可能有效的特定組合。

舉例來說，Archimedes 曾受某大型健康維護組織 (health maintenance organization, HMO) 委託，評估一種稱為 A-L-L，針對糖尿病與心臟病患者新預防性療法的效果。根據 Archimedes 模型的預測，該組合療法應能降低目標族群 71% 的心臟病發作與中風機率。此模型隨後在臨床研究獲得證實：心臟病發作與中風機率降低了 60%。從而該 HMO 建議組織中的醫師，只要病人符合標準，都可採取這種新療法。

Entelos 與 Archimedes 運用到的數學與電腦科學領域包括非線性動力學、控制理論、微分方程，以及物件導向程式設計。參閱 [10] 和 [11]。

## 石油探勘與開採

對石油開採業來說，這是一個充滿風險與機會的時代。即使全世界都關注氣候變遷，並且被迫要減少碳足跡，但至少可在預見的將來，我們仍得依賴石油與天然氣。對「石油頂峰」(peak oil)的憂慮至今未減<sup>⑥</sup>，但主因可能是我們低估了石油工業創新和研發非傳統石油來源的能力。

諸如把二氧化碳注入地面之類的進階生產技術，能從現有油井取得更多石油，也能隔離原先會洩漏到空氣中的碳。以往我們認為重油堆積物的萃取開發太過昂貴(如加拿大亞伯達省的瀝青沙、美國科羅拉多州與懷俄明州的油頁岩)，但隨著石油價格飆升，相關的研發變得越來越有吸引力。另外深海鑽井亦已受到重視，但也存在許多新風險，像是英國石油公司 2010 年在墨西哥灣的漏油事件。

隨著石油探勘的難度與開採成本越來越高，數學演算法與模擬對企業就變得更加重要。以地震數據逆推(用地震紀錄定位地下岩層)一直以來都是石油探勘的重要工具。演算法與電腦軟硬體的演進讓三維、甚至四維的模擬不再遙不可及。大尺度的盆地模型(basin model)可幫助公司決定某段岩層適不適合鑽井。小尺度的油藏模型則用於已開採區塊，預測該區塊的石油流動、研擬最佳開採率的策略，以及預測像是地質斷層因為油藏岩塊的壓力改變而重新活動等問題。

在採油設施正式開採之前，動態模擬能讓石油公司分析並將意外風險降至最低。但英國石油公司的漏油事件讓大家了解更有效的風險分析與建模是必要的。顯然，我們需要更快的、可針對即時資訊分



析的模型，當突發狀況發生時，能快速監控油井與評估意外損失。

### 個案 7：盆地建模

地質條件必須非常特別才能孕育出石油，必須要有油源岩(含有機質的沉積岩)、讓石油流注的儲油岩(通常不會是油源岩)、能困住石油不讓其溢出地表的圈閉岩(不透水岩層)，以及迫使油源岩深入地底的覆蓋岩層，使高溫高壓得以「煮熟」有機質並生成石油。但即使這四種岩層都齊備了，仍然可能沒有石油，因為時機也很重要。如果圈閉岩層太晚形成，石油可能早就漏光了。

盆地模型能用基本物理原理模擬石油形成的所有步驟。例如 Schlumberger 公司的 PetroMod 軟體會從岩層的年齡與性質的資料開始，然後基於地層年代來計算每一層的壓力與溫度，用來為岩石多孔隙性、密度與其他特性建立模型，然後將這些資訊送到化學模型，此模型能模擬原油生成，以及原油如何裂解成不同分子量的氣體與石油。流體流動模型除了能追蹤烴類的移動，也會考慮烴是液態或氣態、岩石的透水性、以及是否有斷層等資訊。模型的計算結果會與試驗鑽孔的測量結果比較，以驗證結果的正確性。在多數情況下，模擬會調整不同的參數進行多次試驗，才能確認資料中不確定的部分可能造成的結果。

總而言之，盆地模型橫跨了許多領域，包括流體流動、熱傳導、化學動力學、地質學、微分方程、

隨機分析，以及地球上最強大超級電腦的計算結果等等。盆地模型攸關數十億美金的收益或虧損。

在此我們提供兩個例子，說明盆地模型的優點，以及不用盆地模型時對石油開採的影響。首先，阿拉斯加普拉德霍灣（Prudhoe Bay）的油田附近有個稱為慕克路克（Mukluk）的可能油田。石油公司在 1980 年代早期砸下 15 億美金買下慕克路克的租賃權，被稱作「歷史上最昂貴的枯井」。雖然慕克路克的地質結構與普拉德霍灣非常相似，但可能是時間次序不對，也可能是圈閉岩沒產生效用，該處根本找不到石油。

第二個例子比較正面。Mobil 與 Unocal 向印尼買下了一塊深海區域的開發權，這塊區域叫做望加錫海峽（Makassar Straits）。依照以往的經驗，人們咸信這是一塊不適合開採的區域，因為該處的油源岩已經「過熟」了。但是 Mobil 的電腦模型卻指出該處的油源岩仍持續生成石油。1998 年的試掘井證明電腦模型是正確的，Unocal 也於 2003 年開始在該處生產石油。這是印尼第一座深海油田，尖峰產量約為一天 20,000 桶。參閱 [12]。（本文將於下期刊完）<sup>③</sup>

本文參考資料請見〈數理人文資料網頁〉

<http://yaucenter.nctu.edu.tw/periodical.php>

#### 本文出處

本文譯自 *Mathematics in Industry* (2012), SIAM (Society for Industrial and Applied Mathematics)，此為報告中的第二章“The Role of Mathematical Sciences in Industry: Trends and Case Studies”。本刊特別感謝中華民國數學會與臺灣工業與應用數學會提供中譯本。

#### 譯者簡介

中華民國數學會的前身「中國數學會」成立於 1935 年，並於 1967

年更名，延續至今。其主要活動為促進臺灣數學教育及數學研究，並謀求與國際數學界發生密切聯繫。

台灣工業與應用數學會成立於 2012 年 6 月，是結合應用數學界、工程界、工業界的學會。緣起來自美國的工業與應用數學會（SIAM）的啟發。

#### 延伸閱讀

►《產業中的數學》（2015），譯本，中華民國數學會與臺灣工業與應用數學會。

[http://www.taiwanmathsoc.org.tw/webpage/mii\\_report\\_cover.html](http://www.taiwanmathsoc.org.tw/webpage/mii_report_cover.html)

原文可見 SIAM 網頁版：

<https://www.siam.org/reports/mii/2012/report.php>

或其 pdf 版：

<https://www.siam.org/reports/mii/2012/report.pdf>

---

<sup>③</sup> 預計石油生產量將達到某最高點，此後將一直下降。本來是地質學家賀伯特（K. Hubbert）對美國石油產量的預言（因此也稱為「賀伯特頂點」），現在也用來描述世界石油產量的走向。