

# Uri, 我和最優分解

黃光明

上一篇 (我和數學研究的第一次邂逅) 寫的是我和數學糾結的起始, 這一篇可看成是終結篇。其間當然還有四十年的橫隔, 點綴著人生的一些小喜小悲, 則將隨緣或記或忘。

Uri Rothblum 和我本是天南地北八竿子打不著的兩個人。講先天, 他是亞洲最西的以色列人, 我是亞洲最東的中國人。我對猶太人本無偏見, 也不願爲了崇高但一知半解的國際政治任意開口得罪朋友, 但有時被他直接面問以色列和中東諸國的糾紛時, 我只有說:「中國在二戰時, 被日本任意攻擊、轟炸, 死了多少人, 我的同情是永遠在被侵略的一方的」, 從此我們不談國事。講後天, 我進入數學界是有些偶然的事 (見「我和數學研究的第一次邂逅」一文), 而他是正規數學系出身, 史丹佛的博士, 然後在耶魯教了八年書, 是大家一致看好的青年才俊。初看, 我們應該是起步就不在一條跑道上的人。講性格, 我是急性的人, 對守時幾乎抬高到了守法的規格; 而他恰是一個把守時視爲糞土的人, 十二點吃午餐的約會加上半小時開車的距離, 他能在電話上東扯西扯到十二點差一分才走出辦公室, 然後就看運氣了, 只要在辦公室到停車場的路上有任何一個人和他打招呼, 他就能停下脚步聊個五分、十分。可憐我這個有糖尿病經不起餓的人在辦公室一下發狠不等了, 一下又延五分鐘, 不知循環幾次。光爲這事, 我就能和他斷交一百次, 我當然是在理的, 但我能知道別人多少次包容了我的缺點付之一笑嗎?

我和他相識在 1990 年代中期, 那時我在 Bell Labs 數學中心, 而他是該中心某人請的顧問, 雖常在走廊上擦身而過, 但未正式會面。其實, 我們在研究上早已是比鄰了。原來在 1981 年我發表了一篇 Optimal Partitions 的文章 [4], 討論如何把有限的資源分配給衆多用戶且達到分配最優化的效果。我這篇文章可算是第一次把這一個分配的問題作爲一個數學問題提出, 換句話說, 以前也有文章在個別特殊應用上提出一些最優化的結果, 我則是把這些結果看出其普遍的規律, 使其適用於一大類問題。Uri 在 1982 年也和合作者發表了類似的結果。雖然我似乎先了一步, 但這可以是反映兩個雜誌處理投稿的快慢, 我們其實是各自獨立做出了結果。之後, 我們兩人都在這領域繼續發表文章, 彼此引用對方文章。一日, 不知他怎麼心血來潮的跨入了我的辦公室, 相互一聲 Hi, 就開始了近二十年的合作, 他也變成了我聘請的顧問。從此我們就合力來推廣這個新的數學問題, 建立基礎結構, 拓展應用領域, 合寫了二十幾篇論文直至 2012

年他猝逝為止。當時還剩下很多還沒有成文的構想，一些尚待思考的課題，但大賢已去，而我也已退休了五年，沒有心力來把它們一一完成，不過至少完成了一個十幾年前就開始的他的心願。原來他雖博學，但因太認真，不敢遽然寫書，看我一本本的出版，就慢慢覺得不一定是那麼一件難事了，於是我們開始寫 *Optimal Partitions* 一書。果然他是持非常謹慎的態度，將文字千錘百鍊，還不敢讓自己的嬰兒入世去撞天下。後來在出版商的壓力下，終於同意把已寫好的三百五十頁作為上冊 [6] 在 2012 年出版。我本以為這樣也算是個交待了，其餘的他愛寫不寫，我聽天由命。但他的離世則使我覺得已責無旁貸非完成這套書以慰在天之靈不可，乃在陳宏賓博士的幫助下一年內完成了下冊 [7]，在 2013 年四月出版了，為 Uri 關上了門，也為將來做應用數學的年輕學者，開了一扇門。

為什麼這本書寫了十幾年（我其它書大概都是一年內完成的），有三個原因。第一是 Uri 的外務太多，隨著他在學界聲望的隆起，他做了講座教授，系主任，院長，乃至副校長 (provost)，行政任務非常繁重。又先後擔任一些著名期刊的編務（去世時猶是 *Linear Algebra and Applications* 的主編），加上纏身的教學任務，有時一整年都無法和我討論寫書事宜。第二是他勤而不快的作風，你看到他工作時不眠不休的專心，就知道他表面的慢完全是精工出細貨的緣故。我們合作的方式是經過提出一個問題、討論及獲得一些成果後，我很快的寫完初稿，通常是相當粗糙且不周延的。然後就給他把關精讀，他會找出不少毛病，並加強結果及寫作上的嚴謹。雖然頁數會增加一倍，時間也拖得很久，但是都是值得的，常常給我清水變雞湯的感覺。第三是最優分解是一個新的研究課題，因此很多工具和附件都不具備，需要自己去建構，為此就中斷了寫書的進程而去做工具了，做成後，這工具當然也是要公諸於世的，所以又去寫文章，花了不少額外時間。同時一本書拖的時間愈長就愈難收尾，因為在這一段過程裡又有很多新的文獻發表，你要把它們收入書中，這是一件聽來容易其實頭痛的事情，因為不但要找到合適的位置，還要照顧到對前後文的影響，正所謂牽一髮而動全局。

第一次見面，Uri 就對我提出一個問題，要敘述這個問題，必須先介紹 1992 年 Barnes, Hoffman 和 Rothblum 發表的一篇經典作 [1]（以後簡稱 BHR 文）提出了一種新的觀點來作最優分解問題。我們先舉一個例來讓讀者明白什麼是一個分解問題。統計學的目的就是把大量的數據簡化成幾個特徵數（如平均值）及討論這樣的簡化會失去多少資訊。一項重要的工具就是把原數據分成數群，使得相似的數據都分在同一群，這樣就把原數據簡化成幾個群了（這個步驟叫做「群聚」，英文是 clustering）。譬如一個暑期營要把參加的青少年按體能、技術分組，則可能選出身高、體重、年齡、年級等等數項來做分組的依據，這時每一個青少年就被這幾項數據代表了。要形成一個分解問題，還缺一個目標函數，就是把我們希望分解達成的目的用數學函數表示出來。這時候合理的選擇可能不止一個，選擇的時候，除了合理性外，也要顧慮到對應的數學問題是否易解。請注意「群聚」雖是很重要的一類分解問題，但還有很多別類的（目標函數不同），這裡只是舉一例。

1958 年 Fisher [3] 考慮了上述的群聚問題, 他假設每一元素(前例中青少年) 被一向量代表(前例中身高、體重、年齡、年級四數) 而一共要分為  $p$  群, 目標函數是把每一群算出下述一個量: 取群內每一對向量之差再平方後之和。注意當每一群的向量都接近時, 向量差及其平方都小, 諸群之和也小, 所以當最優分解定為達到最小目標函數值時的分解, 則的確和我們分組的目的之一致。他用一般分析方法得到了當向量只含一個數時的最優分解, 它有這樣的一個特徵: 如果向量只是身高時, 則最高的一批形成一群, 次高的一批也形成一群, 如此類推, 直到最矮的一批形成一群。這個特徵很強, 滿足它的分解數顯然是多項式型的。他並且在文中說出一個願望, 希望有人能把他的發現推廣到含任意多項的向量模式上。這個願望沉澱了四十年後才由 Boros 和 Hwang [2] 作出回應, 他們用更精密的分析得到了向量多於一項時最優分解的特徵。包含一群點的最小多面體稱為它的外包(convex hull)。則特徵是: 任意兩群的外包或不接觸, 或其一的外包被另一外包包含, 但不含後者的任何點。這樣的敘述似乎是不知所云, 我們回歸到一維向量 ( $d = 1$ ) 看看它在說什麼。在一維時, 特徵保持了每一群都是連續元素的性質, 但容許一群可以整體的插入另一群兩個連續元素間(因此比 Fisher 的結論弱)。譬如三個青少年依身高次序排為 1, 2, 3 號, 設群  $A$  含(1, 3), 群  $B$  含(2), 則這分解滿足外包  $B$  不含群  $A$  的點的條件, 卻不滿足最高的兩個在一群的條件。

BHR 文中所提出的方法是用於下列一類問題的(稱為「和分解問題」): 如把含  $d$  數的向量看成一個  $d$  維點, 則在和分解問題中, 一個含  $p$  群的分解, 可以表示為一個  $d$  列  $p$  行的方陣(因此可看成是  $dp$  維空間的一個點), 第  $j$  行的向量是第  $j$  群點(每一點都是一個  $d$  維向量)的和。這樣以和來代表一群點的合理性端看目標函數是否只依賴於每群的和, 如果是, 則以群中點的和來代表該群並無不可。當然, 和分解問題只是分解問題的一類, 但從迄今的文獻看來, 它是最重要的一類, 也是在現實中出現最多的一類(雖然前段所舉之例並不是)。本文的討論將只限於和分解問題。

一個做最優化問題的人都熟知的定理說當目標函數是凹向上函數(或稱凸向下函數(convex function)), 以下簡稱凹函數)時, 最優點會落在一個頂點上, 這樣我們就只要比較頂點而可以忽略所有的內點了。我解釋一下何以如此。這裡只說平面上的點, 則凹函數是一個 U 字形的曲線, 所以任何一線段上的點其值落在 U 形曲線的一段時, 最大值必落在兩頂點之一上, 現在再考慮一個多面體裡所有的點, 如果最優點不是多面體的頂點, 則必能取一直線段通過該點而交於多面體表面。根據前論這條線段上有一端會是最優點, 這新的最優點已在多面體的表面上, 如還不是多面體頂點, 則再做如上的移動, 直到移到頂點為止(讀者可以假設多面體是在平面上的多角形而畫出如何從形內一點經過兩次直線變換而移到邊界, 再移到頂點)。估計頂點數目也是離散數學中一個重要的課題, 有人做了一些, 我們也補充了一些, 但常常數目還是偏高或根本不會算, 這就需要找別的辦法了。大家應該都聽到過「線性規劃」, 據說是二十世紀應用數學上最了不起的發現, 也是當今應用最廣的數學。它以一組所有解答必須遵守的不等式把解答的空

間切割出一個多面體，目標函數是線性時（線性函數也是凹函數，故最優解答落在頂點上），線性規劃發展出一些效率很高的算法來找到一個最優頂點。但我們的最優分解問題不能直接用上線性規劃，因為還不知道分解問題所建構的多面體是不是線性規劃切出的多面體，這就是 Uri 提出討論的問題了。我們是找了外援才解決的，當時賓州州大的李文卿教授也在訪問 Bell Labs 做顧問，她是從小學一年級起就終身第一名的名數學家，另有一大陸清華高材生高彪當時在明尼蘇達唸博士而暑假來跟我做研究，群賢畢至，星光閃耀，難題又何懼？經幾次討論各自深思後，終由最年輕的高彪，看出關鍵的一步而宣告破案。因為敘述較繁，此處略過。

我們總結一下以多面體來做分解問題的精神，及尚需補充之處：

1. 當目標函數是凹函數時，最優分解（或最優點）只要在多面體的頂點裡去找就行了。這固然是大大的縮小了範圍，但我們仍需要知道頂點數是否仍太多，在計算學上就是問這數目是多項式型的，像  $n$  的  $k$  次方（ $k$  是一常數），或是冪數型的，像  $k$  的  $n$  次方。這兩個數目隨著  $n$  的變大而相差飛漲，譬如當  $k$  是 2 時， $n = 10$ ，兩數相差約是一千比一百， $n = 20$ ，則兩數相差就變成一百萬比四百，更不要說  $n = 100$  了。所以這兩型不只是量的區別，而是質的區別。如果是前一型，則我們可以列舉分解，逐一算出它的目標值，最大者即是最優分解；但如是冪數型的，那就不勝枚舉了。
2. 即使枚舉法可行，也是耗時耗力的事，所以如能證明分解問題中定義出來的多面體就是從線性規劃定義的那個多面體，則就可用發展成熟的線性規劃法很快的得到最優點。高、黃、李、Uri 四人雖然證明了這兩個多面體相等，那證明只是針對一類最簡單的多面體的，「簡單」主要是指下面兩事：
  - 向量只含一項。他們的證明用到了一維的向量可以比大小的性質，這在高維度時就辦不到（如  $(1, 2)$  和  $(2, 1)$  誰大？）。1999 年 Hwang, Onn 和 Rothblum [5] 開闢了一條新的路來銜接分解多面體和線性規劃多面體。這方法可形容為「賠了面子，贏了裡子」。第一步是把原題表現為一個等價但規模更大的問題，把一個分解寫成一個  $p \times n$  的方陣，每一行標示一個群，每一列標示一個向量（或元素）。令  $P$  是一個分解，則當向量  $i$  在  $P$  中屬於第  $j$  群時，方陣的第  $j$  行第  $i$  列就放進一個 1，否則放進 0。很顯明的，這個 0-1 方陣的確刻劃了  $P$ ，它沒有觸及到「和」的表示而把它留在目標函數裡。雖然這個方陣對應的多面體是  $pn$  維的，通常遠大於「和多面體」的  $dp$  維，但這大的多面體是「作業研究」學裡熟知的「廣義交通多面體」，它可以和線性規劃多面體銜接也是熟知的。
  - 他們的證明中假設了每一群含幾個元素是給定的，更一般的假設是規定第  $j$  群最少含幾個元素，最多幾個。但這些推廣並不算太困難，也都逐步完成了。

接著我們再考慮頂點的數目是多項式型還是冪數型，這裡又要引用 BHR 文的一個結果。

其實這結果本身就饒有趣味，因為是從一個代數的結果裡開出了幾何的花朵。該文探討的多面體是對應於多維向量分為  $p$  群、但每一群含多少元素有限制的分解題。他們證明了每一頂點所對應的分解有這樣的幾何性質。考慮每一群所含點的外包，則這些外包互不碰到（有這樣性質的分解稱為「可離分解」），注意前述 Boros-Hwang 特徵裡也提到過這一性質。為什麼這個結果重要且有趣呢？

- 它太出人意外，而這就是有趣。
- 從此我們對所有多面體都會期盼頂點所對應的分解，會有一些幾何性質，而這期盼也一直沒有令我們失望過。
- 要研究頂點任何性質，都多了一條道路，就是從它的幾何性質入手。

下面我們就利用這個性質把算頂點的問題換成算可離分解數的問題，後者是一道用在幾何上的組合數學問題。當時關心這問題的人頗多，且眾說紛紜，想不到答案卻很簡單，且出乎多人意料。1999 年 Hwang, Onn 和 Uri [5] 證明了可離分解的數目是多項式型，且項數不高，列舉不難。為了讀者可以在紙上操作，我們只給出點在平面上的證明。

先考慮  $p = 2$ 。令  $P$  為一個可離分解把  $n$  物分成兩群，且外包不碰到。則我們可以在兩個外包間畫一條直線  $L$  分開它們。所以任一條不含點的線都對應於一個可離分解；問題是有無限多的線對應於同一個可離分解，我們用下法找出每一個對應於  $P$  線-群的特徵代表：平行移動  $L$  直至它碰到一點，如果把線上點算成原群的話，則新線仍對應於  $P$ 。當碰到一點時，直線不能再如前平行移動（會跨過這一點而改變分解），但可以以碰到的點為中心旋轉移動直至碰到第二點。這時再不能移動此直線而仍保住原分解了。所以每一個可離分解都可轉換成一對點，而每一對點也只對應於一個可離分解。因此所有可離分解的數目就等於所有對的數目，在  $n$  點中對的數目是  $\frac{n(n-1)}{2}$ 。在討論多項式型和冪數型時，我們往往只關心最大的一項，因為它就決定了是那一型，所以其它的項都被略去。我們也略去系數，因為有它沒它，對電腦能否列舉沒有影響，我們以  $O(\cdot)$  來代表這種略去後的表示法，譬如  $\frac{n(n-1)}{2}$  就可寫成  $O(n^2)$ 。再考慮一般的  $p$ ，對任一對外包，我們都要求它們不碰，且知有  $O(n^2)$  的分解滿足這要求。現在有  $\frac{p(p-1)}{2}$  對外包，所以最多有  $O(n^2)$  的  $\frac{p(p-1)}{2}$  次方個分解是可離分解。為什麼說「最多」？因為滿足  $A, B$  兩外包不碰的分解不一定也滿足  $A, C$  兩外包不碰，但我們的計算把它們都算進去了。

Uri 和我都不是只做工不玩的人，我們每次討論完後，只要天氣許可，都會去網球場打四十分鐘球。有時時間緊迫，也會縮短討論時間。有時球打到一半，他忽然走到網前，又延續起前面的討論。時間到時總依依不捨，先說打最後三球，打完後又宣佈再打真正最後三球，打壞了的還要補打。他也迷看戲，住在曼哈頓時，買了不到一百元的年票，每年可看幾十場。請我們夫婦看了一場 off off Broadway，告訴我只要三元錢一張票。他喜歡旅行，藉著開會和訪問，到處公費旅行。應我邀請，來了台灣兩次（我也回報去訪問以色列一次），最喜歡拍公路上看到的廟宇，幾

乎一路上沒有漏拍的。有一次,我帶他到台北玉市,一塊小玉,小販要五元(大約是美金),我費盡九牛二虎之力還成三元,他卻以二元買到。因為他不通中文,只會豎起兩根手指,從頭到底就這一招,但對於小販的如簧巧舌也滴水不進,完全免疫,以靜制動,小販終於投降。2012年二月某晚,我太太接到一通 Uri 夫人 Naomi 的電話,說晚上上床後她在看書,Uri 在打電話,忽然有一陣沒聽到講話聲,再看,Uri 已不能動了,從此就不再醒來。兩個禮拜後,家族聽從了醫生的建議,拔掉了人工維持生命的器械,大賢已行!大賢已行!

誌謝: 本文承陳宏賓博士改寫成符合投稿規定的 pdf 形式,深表感謝。

## 參考資料

1. Barnes, E. R., Hoffman, A. J. and Rothblum, U. G., Optimal partitions having disjoint and conic hulls, *Math. Prog.* 53, 1992, 69-86.
2. Boros, E. and Hwang, F. K., Optimality of nested partitions and its application to cluster analysis, *SIAM J. Optim.* 6, 1996, 1153-1162.
3. Fisher, W. D., On grouping for maximal homogeneity, *J. Amer. Stat. Assoc.* 53, 1958, 789-798.
4. Hwang, F. K., Optimal partitions, *J. Optim. Thy and Appl.* 34, 1981, 1-10.
5. Hwang, F. K., Onn, S. and Rothblum, U. G., A polynomial time algorithm for shaped partition problems, *SIAM J. Optim.* 10, 1999, 70-80.
6. Hwang, F. K. and Rothblum, U. G., *Partitions: Optimality and Clustering, Vol. I*, World Scientific, 2012, Singapore.
7. Hwang, F. K., Rothblum, U. G. and Chen, H. B., *Partitions: Optimality and Clustering, Vol. II*, World Scientific, 2013, Singapore.

—本文作者為國立交通大學應用數學系講座教授(已退休)—