

雙邊 Hardy 不等式及其幾何應用



演講者：陳木法教授(北京師範大學教授、
中國科學院院士)

時間：民國 101 年 10 月 22 日

地點：天文數學館 202室

整理：陳麗伍

今天要談的是雙邊的 Hardy 不等式及其在幾何中的應用。雙邊的 Hardy 不等式可能很多人不是特別熟悉，我會慢慢介紹。事實上，Hardy-type inequality 是調和分析中很大的一個主題。

下面分三個小標題，首先是背景介紹，第二個是今天的主題 Hardy-type inequality (Hardy-type 不等式)，第三個是幾何應用，介紹的主要是這兩年的工作。

1. 背景介紹

大概在上世紀，七十年代 (1970s)，也就是四、五十年前，數學重新回歸自然。一個代表性的例子就是概率論 (機率論) 跟統計力學的交叉形成的新的研究領域。當年，我們就是走這個被稱為交互作用的粒子系統 (interacting particle system) 的研究方向。統計力學研究的中心問題當然是相變 (Phase Transition) 現象。相變現象的數學是無窮維的數學。無窮維數學的工具很少，所以要四處找工具。因為這個原因，我尋訪過計算數學 (獲知其中的三對角線矩陣算法)，也尋訪過泛函分析、幾何代數、數學物理方法等等，發現幾何做的比較精彩。幾何裡面做了許多估計的工作，所以我們就去學幾何。隨後不久，發現概率方法也能證明幾何結果，這大概是在 1992 年，我們開始做一點幾何的估計。然後大約在 1999 年，距今十多年了，有個很大的發

感謝陳木法教授同意本刊轉載他擔任醉月湖講座 (由中央研究院數學研究所與臺灣大學數學系聯合主辦) 的演講紀錄。本文下方註解為本刊所加，非原稿所有，希望能方便讀者閱讀。

現，發現調和分析，也就是 Hardy 不等式¹ 這個工具對我們很有用。當時的十個判別準則全是在知道 Hardy 不等式以後才展開的、才完成的。十多年過去了，大概就這兩年，我們才想到用概率方法也能做 Hardy-type 不等式，由此迎來了新的篇章。

概率 (機率) 的一個特點就是跟其他領域比較起來相對年輕，所以比較開放。一方面，機率理論是在數學好多分支的撫育下成長起來的；另一方面，最近這些年來，機率的一些思想也滲透到很多其它的數學分支。今天講的就是一個例子，是機率應用於其它數學分支的一種比較深入的結果。同時，也希望讓年輕的朋友體會一下數學的功夫。我覺得做數學很大的程度要靠功夫，靠硬功夫。

2. Hardy-type 不等式

談到 Hardy-type 不等式，當然要從老 Hardy 不等式開始。老 Hardy 不等式是九十年前，也就是 1920 年提出的一個不等式。這個不等式非常的簡練。

$$\int_0^\infty \left(\frac{1}{x} \int_0^x f \right)^p dx \leq \left(\frac{p}{p-1} \right)^p \int_0^\infty f^p, \quad f \geq 0, p > 1$$

這就是 Hardy 在 1920 年提出的不等式。這裏及隨後，Lebesgue 測度 dx 常略去不寫。還有一種寫法把 0 到 x 局部的積分，寫成 $Hf(x)$ 。這個 H 是為了紀念 Hardy 而通行的記號，稱為 Hardy 算子。利用算子 H ，上式可改寫成

$$\int_0^\infty \left(\frac{1}{x} Hf(x) \right)^p dx \leq \left(\frac{p}{p-1} \right)^p \int_0^\infty f^p. \quad \boxed{\text{常數 } \frac{p}{p-1} \text{ 精確}}$$

還有一個更方便的解釋：把左邊改成離散的求和，0 到 x 的部分和的平均就是算術平均，左邊就是算術平均的 p 階矩 (moment)，不等式成為 f 的算術平均的 p 階矩小於等於 f 本身的 p 階矩乘以常數 $[p/(p-1)]^p$ ，這樣就特別容易記住：

$$\|Af\|_{L^p(dx)} \leq \frac{p}{p-1} \|f\|_{L^p(dx)}.$$

Hardy 為什麼要證明這個不等式呢？他是為了尋找 Hilbert² 的一個定理的初等證明。Hilbert 定理講的是離散的級數，證明一個 2 重級數 (Series) 的收斂性。Hilbert 的證明是用傅立葉分析 (Fourier Analysis)，Hardy 質疑一個初等的級數的收斂性為什麼要用傅立葉分析，嫌它太“高級”，認為應當有一個初等的證明。所以 Hardy 就引進這麼一個不等式，讓這個 2 重級數的和用一個單重級數的和來控制，這是引出這個不等式的原因。Hardy 這篇文章主要討論的是離散級數的情況，很奇怪的是他同時也把連續 (continuous) 的情況寫出來，但是

¹Godfrey Harold Hardy (1877~1947)，英國數學家，以在數論和數學分析方面的研究最爲人所知。

²David Hilbert (1862~1943)，德國數學家，十九世紀和二十世紀最有影響力的數學家之一。

沒有給出證明。這是 1920 年的文章。如果 $p = 2$, Hardy 不等式就成爲 Poincaré³ 不等式, 這個不等式我們做得更久, 也更老一點, 一八八幾年時就提出來了, 所以 Poincaré 不等式要比 Hardy 不等式更早一點。這個非常簡單的不等式, 實際上牽扯到那個時代的三位頂尖的數學家, 我認爲當時最前面的三位數學家就是這三位。所以很好玩, 這麼簡單的數學, 牽扯了這麼多一流的數學家。

看到這麼簡單的不等式, 大家都會想要自己證它一下。我想對學數學的人來說這個很要緊, 因爲數學是做出來的, 不是讀出來的, 更不像看戲那樣就能看會的。所以遇到甚麼問題, 就想自己證明它一下。在這裏, 首先容易聯想到 $L^p(\mu)$ 的性質, 這裏 μ 是概率測度 (measure)。不等式的左方的被積函數是關於概率測度 $\frac{1}{x} \int_0^x$ 的一階矩的 p 次方, 它受控於關於此測度的 p 階矩。 L^p 就是 p 階矩, 這個 p 階矩所對應的範數 (norm) 有一個單調上升 (monotone increasing) 的性質, 當 p 大於零時, $L^p(\mu)$ 範數關於 p 單調上升。有了這一點概念, 就可以給出證明。

$$\begin{aligned} \int_0^\infty \left(\frac{1}{x} \int_0^x f(y) dy \right)^p dx &\leq \int_0^\infty \left(\int_0^x f(y)^p dy \right) \frac{dx}{x} \\ &= \int_0^\infty f(y)^p dy \int_y^\infty \frac{dx}{x} \quad (\text{由 Fubini 定理}) \\ &= \infty \end{aligned}$$

倒數第二行中的 $1/x$ 積出來是無窮大, 所以這個結論沒有用。我們的第一輪證明失敗了。

我們不必爲此感到灰心喪氣, 實際上這個不等式的證明也不是平凡的。之前提到過, Hardy 在 1920 年寫出這個不等式但是沒有給證明, 這個不等式的證明是五年 (1925年) 之後才發表的。所以它的證明並不是平凡的, 不能一下子就看出來。讓我們回頭看看在哪裡摔的跤? 這個證明問題就在 $1/x$ 是個發散積分 (divergent integral)。把 $1/x$ 修改一下, 改成 $\frac{1}{x^{1+\delta}}$, 就會變成一個收斂積分 (convergent integral)。這樣, 我們可以將上面的論證重寫如下:

$$\begin{aligned} \int_0^\infty \left(\frac{1}{x} \int_0^x f(y) dy \right)^p \frac{dx}{x^\delta} &\leq \int_0^\infty \left(\int_0^x f(y)^p dy \right) \frac{dx}{x^{1+\delta}} \\ &= \int_0^\infty f(y)^p dy \int_y^\infty \frac{dx}{x^{1+\delta}} \quad (\text{由 Fubini 定理}) \\ &= \frac{1}{\delta} \int_0^\infty f(y)^p \frac{dy}{y^\delta}, \quad \boxed{\forall \delta > 0, f \geq 0}. \end{aligned}$$

這個改變很多時候可能就不平凡, 這就是絕地逢生。別人看沒路可走, 我們卻走出一條路來。可惜所得到的結果並不是我們所要的 Hardy 不等式。該怎麼回到 Hardy 不等式呢? 也就是需

³Jules Henri Poincaré (1854~1912), 法國數學家、理論物理學家、工程師及科學哲學家, 對數學及其應用具有全面知識。

要將最後一行中的分母 y^δ 去除, 就有希望了。移除分母最簡單的方法就是變量替換 (variable substitution), 將不等式兩邊的 f 改寫成 g :

$$\text{設 } g(y) = f(y^{1+\gamma})y^\gamma, \quad \gamma = \frac{\delta}{p-1}, \quad \delta = 1$$

這個一看就知道是變量替換公式。這一步可說是神來之筆, 用了這一個步驟後, 答案就呼之欲出, 回到 Hardy 不等式了。實際上, 這個 δ 是最小化的解, 等於 1, 如果要知道答案, 把 δ 代進去就完了, 但這當中包含了很深刻的道理。這樣, 我們就完成了 Hardy 不等式的證明。

這個證明我自己覺得相當神奇。實際上 Hardy 1925 年的證明之後, 經過了 40 年, 這個證明才由 E. K. Godunova⁴ 找到。儘管這篇文章在 1970 年, 也就是五年後翻譯成英文, 在中研院數學所的圖書館可以找到這篇文章的資料, 但是實際上沒有人注意到這個工作。之後又過了差不多 40 年的時間, 在 2002 年 Stein Kaijser⁵, Lars-Erik Persson⁶ 和 Anders Öberg⁷ 發表了一篇文章⁸, 重新發表這個證明, 這時候大家才注意到有一個很簡單的方法來證明這個不等式, Hardy 始終不知道有這樣一個證明。

這是一個滿漂亮的證明, 可以體現我第一個想說的話, 體現出數學的功夫。有功夫能看出來, 沒有功夫還是不行。當然, 這個 Hardy 不等式太簡單了。說老實話, 我在做這個方向的早期就知道, 但是它對我來說沒有什麼用處, 因為太簡單了。如果熟悉隨機過程, 看看其對應的算子就知道這太簡單了, 只是一種非常特別的情況, 沒有太多用處, 所以就要往前走。接下來的故事就滿長的, 因為要一步一步發展。

剛才的 Hardy 算子寫法是這個樣子,

$$\int_0^\infty \left(\frac{1}{x} Hf(x) \right)^p dx \leq \left(\frac{p}{p-1} \right)^p \int_0^\infty f^p$$

或等價地

$$\|Hf\|_{L^p(x^{-p}dx)} \leq \frac{p}{p-1} \|f\|_{L^p(dx)}$$

現在用另一個觀點來看, 左方老 Hardy 處理的是 Hf 的 L^p 範數 (norm), 針對的是測度 $x^{-p}dx$ 。當然, 上式左右的兩個測度都很簡單, 因為簡單, 所以下一個任務就要推廣一下。把兩邊的測度換成一般的、抽象的 Borel 測度 $=(x^{-p}dx, dx) \rightarrow (\mu, \nu)$, 就變成非常一般的形式:

$$\|Hf\|_{L^p(\mu)} \leq A \|f\|_{L^p(\nu)}$$

⁴E. K. Godunova, Russian mathematician,

⁵Sten Kaijser, 任教於瑞典烏普薩拉大學 (Uppsala University)。

⁶Lars-Erik Persson, 任教於瑞典呂勒奧科技大學 (Lulea University of Technology)。

⁷Anders Öberg, 任教於瑞典耶夫勒大學 (University College of Gävle)。

⁸Kaijser, Sten; Persson, Lars-Erik; Öberg, Anders, On Carleman and Knopp's inequalities. J. Approx. Theory 117 (2002), no. 1, 140-151.

這就是今天所要講的 Hardy-type 不等式。從現在開始，為節省記號，不等式中的常數 A 均指最佳常數。這個 Hardy-type 不等式經歷了大約七十年，約在 1970 年前後，這個不等式才研究得比較完整，有多篇文章，G. Talenti (1969), G. Tomaselli (1969), R.S. Chisholm & W.N. Everitt (1970-71) 等，其中 B. Muckenhoupt 在 1972 年發表的文章⁹ 被引用得多一點，時間上也晚一點。其他的文章，例如 M. Artola (1968-69) 和 D.W. Boyd & J.A. Erdős (1972) 根本沒有發表，因為知道別人已經寫出來了，所以即使投稿了的稿件也自己撤回。總之，這裡已經牽扯到六篇文章。

下一步想想能不能再推廣到更一般的 p 和 q ：

$$\|Hf\|_{L^q(\mu)} \leq A\|f\|_{L^p(\nu)},$$

這個時候兩邊幾乎對等了。但一邊是 L^p ，另一邊是 L^q ，難度又增加了。差不多又做了 20 年，經過 P. Gurka (1984), E.N. Batuev & V.D. Stepanov (1989)，直到 1990 年 B. Opic & A. Kufner 才做得比較完整。應當說，這種 (p, q) 情形也很早就注意到了，Hardy 本人在 1930 年就開始做一些特殊的情況 (G.H. Hardy & J.E. Littlewood, 1930)，不過這篇也沒有發表。最後指出，《Hardy-type Inequalities》¹⁰ 是一本專著，總結了之前的故事，是關於 Hardy 不等式比較經典的著作。

再往下說之前，還要提一下，總體的目標是用測度 μ 和 ν 來算出 A 的一個較好的估計 (當然，如能 sharp 就更好)。回憶 Hf 是 f 從 0 到 x 的積分。如果將 Hf 改寫成 f ，原本的 f 就成為 f' 。這樣改寫後，左邊就成為 f 的 q 階範數，右邊就變成 f' 的 p 階範數，再將區間改成更廣一些的從 $-M$ 到 N ($M, N \leq \infty$) 的區間，我們得到如下形式的 Hardy-type 不等式：

$$\left(\int_{-M}^N |f|^q d\mu \right)^{\frac{1}{q}} \leq A \left(\int_{-M}^N |f'|^p d\nu \right)^{\frac{1}{p}}.$$

當然，這個不等式也可以寫成

$$\|f\|_{\mu, q} \leq A\|f'\|_{\nu, p}.$$

這樣一寫，就要小心。原來的 Hardy 算子初值為 0，是 0 到 x 的積分。所以新的函數 f 在左端點要等於 0，邊界條件要寫上： $f(-M) = 0$ 。事實上，這時候可以寫兩個邊界條件： $f(-M) = 0$ 和 $f(N) = 0$ 。兩個邊界條件就是今天的主題，雙邊的 Hardy-type 不等式。在雙邊的時候，不能寫成單邊的 Hardy 算子的形式，尤其你由一個 0 初值出去，只要非負、非零，積分後就不可能

⁹Muckenhoupt, Benjamin, Weighted norm inequalities for the Hardy maximal function. Trans. Amer. Math. Soc. 165 (1972), 207-226.

¹⁰B. Opic and A. Kufner, Hardy-type Inequalities, Pitman Research Notes in Mathematics Series, Longman, New York, 1990.

再回到 0, 所以不可能寫成 Hardy 算子的形式。這樣, 前面的單邊情形變換 $(Hf, f) \rightarrow (f, f')$ 前後沒有差別, 但是後面的雙邊情形就有不同。後者是今天討論的主要情況。

因為這個題目很多人不熟悉, 所以我要強調一下。這是相當時髦的題目, 就我所知, 總共已有五本書討論這個題目。

- [1] Opic, B. and Kufner, A. *Hardy-type Inequalities*. Longman, New York, 1990.
- [2] Kufner, A. and Persson, L.E. *Weighted Inequalities of Hardy-type*. World Scientific, 2003.
- [3] Kufner, A., Maligranda, L. and Persson, L.E. *The Hardy Inequality: About its History and Some Related Results*. Vydavatel'sky Servis, 2007.
- [4] Kokilashvili, V., Meshki, A. and Persson, L.E. *Weighted Norm Inequalities for Integral Transforms with Product Weights*. Nova Sci. Publ., New York, 2010.
- [5] Maz'ya, V. *Sobolev Spaces with Applications to Elliptic Partial Differential Equations* (2nd Ed.). Springer, 2011.

第一本書是剛才提到的 1990 年的名著。之後的四本都是最近十年出現的著作。前三本的書名中都有 Hardy 不等式或是 Hardy-type 不等式。第四本書名中沒有提到 Hardy 不等式, 但是 weighted norm inequalities 完全是 Hardy 不等式, 只不過做的是高維的 Hardy 不等式。第三本書談的是歷史, 內容是好玩的。一個分支的研究狀況, 如果東西少, 就沒有歷史可談; 這裏邊講了許多歷史, 說明這個題目已經有了充分的發展。第五本書標題也沒有 Hardy 不等式, 因為這本書有 894 頁, Hardy 不等式只是其中一部分。

一開始的時候, 我提過想要用譜值論、特徵值理論來研究相變現象。我做了整整十年之後, 還不知道有 Hardy-type 不等式, 我只知道有古典的 Hardy 不等式, 而不知道有這麼一般的 Hardy-type 不等式。這個一般的不等式對我來說是個非常大的 favor。知道這個就不得了, 好多東西都做出來了, 對我們幫助非常大。比如說大家可能聽說過的 Log-Sobolev 不等式。Log-Sobolev 不等式我們也做了幾十年, 但是其準則只是因為有了 Hardy-type 不等式後才找到的。很多很要緊的結果就是因為用了 Hardy-type 不等式, 所以我們十年中不斷地用。我有一個列表, 一維情形有十個準則。在當時被稱為十大準則。Hardy-type 不等式為什麼在調和分析中有如此地位就是因為它是很要緊的、核心的一個問題。然後大約在二年前 (2010年), 我們找到反向的、原來的 Hardy 理論中沒有的東西, 引導我們往這個方向發展。

在講我的結果之前, 先解釋一下記號。第一個是測度 $\hat{\nu}$, 它是由測度 ν 的絕對連續部分 (absolutely continuous part) 導出的。接下來所有的事情都要通過測度 μ 和 $\hat{\nu}$ 表示出來, 所以 $\hat{\nu}$ 是很要緊的。

$$\hat{\nu}(dx) = \hat{\nu}_p(dx) = \left(\frac{d\nu^\#}{dx} \right)^{\frac{-1}{p-1}} dx \quad \nu^\# : \nu \text{ 的絕對連續部分,}$$

此處需要假定 $\nu^\#$ 的密度非零, 否則可用極限過渡。接著需要常數 $k_{q,p}$:

$$k_{q,p} = \left(1 + \frac{q}{p'}\right)^{\frac{1}{q}} \left(1 + \frac{p'}{q}\right)^{\frac{1}{p'}} \leq 2 \text{ 若 } q \geq p,$$

其中 p' 是 p 的共軛指數: $1/p + 1/p' = 1$ 。後面要用到的這個 $k_{q,p}$, 當 $q = p$ 時是精確的; 當 $q \neq p$ 時, 可換成更小一些的精確常數, 此處略去。這個式子不需要記住, 只要知道 $k_{q,p}$ 永遠小於等於 2。所以粗糙地講, 把 $k_{q,p}$ 當成 2 好了。有了這兩個記號後, 就可以進入今天的第一個主要結果。

定理 2.1 (陳, Acta Math. Sin. Eng. Ser. 2013). 在 *Hardy-type* 不等式中的最佳常數 A 滿足

(1) $A \leq k_{q,p} B^*$ 如 $1 < p \leq q < \infty$ 且 $\mu_{pp} = 0$,

(2) $A \geq B_*$ 如 $1 < p, q < \infty$, 其中

$$B^* = \sup_{x \leq y} \frac{\mu[x, y]^{\frac{1}{q}}}{\left[\hat{\nu}[-M, x]^{\frac{q(1-p)}{p}} + \hat{\nu}[y, N]^{\frac{q(1-p)}{p}}\right]^{\frac{1}{q}}},$$

$$B_* = \sup_{x \leq y} \frac{\mu[x, y]^{\frac{1}{q}}}{\left[\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}\right]^{\frac{1}{q}}},$$

此外, 當 $q \geq p$ 時, 我們有 $B_* \leq B^* \leq 2^{\frac{1}{p} - \frac{1}{q}} B_*$ 。

換言之, 最佳常數 A 有上、下控制, 分別由 (1) 和 (2) 給出。 B^* 和 B_* 這兩個表達式由測度 μ 和 $\hat{\nu}$ 完全刻劃。只是對於上界, 要一些條件: q 要大於等於 p , 測度 μ 要連續。熟知測度 μ 的分解有三部分: 第一部分絕對連續, 第二部分奇異連續 (singularly continuous), 第三部分是純點的 (discrete), 這裏假定純點是 0, 也就是 $\mu_{pp} = 0$, 沒有純跳的部分。在這裡我們要強調: 上界估計是最重要、最有用的東西。如果把 (1) 中的條件都去掉, 會是什麼情況, 目前還沒有答案。所以這裡是一個開放的地盤, 還有很多的問題。

講一下這個定理有那些漂亮的地方。首先, 第一個漂亮的地方是上、下界相差的常數是 universal, 是一個普適常數, 跟測度沒有關係。第二個優點是這裡有兩個邊界, 一個是 $-M$, 另一個是 N 。 M 和 N 都可以是無窮大。那麼邊界條件 $F(-M) = 0$, 如果 M 是無窮大, 表示趨於負無窮大時 F 趨於 0。兩個邊界, 一個 $-M$, 一個 N , 這兩個邊界在這兩個常數中是對稱的, 這是一個漂亮的地方。無論是 B^* 或是 B_* , 這兩個邊界都是對稱的, 這是第二個優點。第三個優點, 如果不喜歡一個上、一個下, 只利用一個也是可以。因為 B^* 可以由 B_* 統起來, 相差一個常數倍而已。將這個常數倍放大, 寫到第一行裡, 這個新常數仍然是小於等於 2。如果 $q = p$,

兩個常數就變成一個,

$$B^* = B_* = \sup_{x \leq y} \frac{\mu[x, y]^{\frac{1}{q}}}{[\hat{\nu}[-M, x]^{1-p} + \hat{\nu}[y, N]^{1-p}]^{\frac{1}{q}}},$$

這又是另一個優點。最後一個優點, 我做的這個雙邊的, 可以回到單邊。如果把第二加項 $\hat{\nu}[y, N]$ 去掉, 那麼 y 不起作用, 可取為 N , 因此 N 的邊界條件就不起作用, y 就可以移到 N 那邊, 就變成單邊。所以右邊的邊界條件不起作用, 就變成單個邊界條件, 這個表達式就變得非常簡單。因為那項去掉後, 分母可以倒過來, 就很簡單。如果把第一項 $\hat{\nu}[-M, x)$ 去掉, 因為是對稱的, 邊界條件就變成在右邊。也就是

$$\text{去掉 } \hat{\nu}[y, N], \text{ 得到 } \boxed{f(-M) = 0}, B^- = \sup_x \mu[x, N]^{\frac{1}{q}} \hat{\nu}[-M, x)^{\frac{p-1}{p}}.$$

$$\text{去掉 } \hat{\nu}[-M, x), \text{ 得到 } \boxed{f(N) = 0}, B^+ = \sup_y \mu[-M, y]^{\frac{1}{q}} \hat{\nu}[y, N]^{\frac{p-1}{p}}.$$

總結起來, 我們所研究的 Hardy-type 不等式是

$$\text{在 } (-M, N) \text{ 上, } \|f\|_{\mu, q} \leq A \|f'\|_{\nu, p},$$

其中 μ 和 ν 是 Borel 測度。對於單邊情形, 如 $f(-M) = 0$, 即左邊 = 0, 最佳常數寫成 A^- ; 如果右邊 = 0, 最佳常數寫成 A^+ 。區分開來, 這時候就有了以下這個定理:

定理 2.2. 令 $q \geq p$, 則 $B^\pm \leq A^\pm \leq k_{q,p} B^\pm$, 其中

$$B^- = \sup_x \mu[x, N]^{\frac{1}{q}} \hat{\nu}[-M, x)^{\frac{p-1}{p}},$$

$$B^+ = \sup_y \mu[-M, y]^{\frac{1}{q}} \hat{\nu}[y, N]^{\frac{p-1}{p}}.$$

這就回答前面說的, 從 1920 年到 1990 年間 70 年研究的結果, 就是這個結果。容易看出, 從後一定理到前一個, 有相當的跨度。

現在的問題是, 該如何從單邊走向雙邊, 這是我的核心問題, 也是困難的地方。原本 $p = q = 2$ 時, 我們會證明。最早的發現, 也就是從這個地方發現的。只是那是在用了三大工具, 經過五個步驟後才做出來的, 所以很辛苦。但是我還是不甘心, 覺得這樣簡單的問題, 為什麼需要那麼大的工具, 特別是需要容度 (capacity) 等等。經過長時間的摸索, 才找到新的辦法。下面要講的就是這個新的證明。相信許多讀者, 不論是經驗老到的或是初學者, 都會在此處停下來, 想想能否找到自己的證明。

在介紹我們的證明之前, 讓我們回顧一下, 單邊的情形結論已經有了。現在, 我們把 $-M$ 跟 N 的區間中間加了一個點 θ , 將區間分成兩部分。



令 A_θ^- 為在 $(-M, \theta)$ 上, 當 $f(-M) = 0$ 時的最佳常數,
 令 A_θ^+ 為在 (θ, N) 上, 當 $f(N) = 0$ 時的最佳常數。

加了這個點 θ 後, 變成兩個單邊。在左邊這個半區間裡關心的是左端點 $-M$, 在另一個半邊關心的是右端點的 N , 然後將剛才的結果重寫一次, 得到下面的結果。這個跟上面的差別, 從左邊半區間看, 無非是用 θ 代表先前的 N ; 從右邊半區間看, 無非是用 θ 代表先前的 $-M$; 也就是換一下記號而已。這是下面我們要用到的已知結果。

令 $\boxed{q \geq p}$. 則我們有 $B_\theta^\pm \leq A_\theta^\pm \leq k_{q,p} B_\theta^\pm$, 其中

$$B_\theta^- = \sup_{x < \theta} \mu[x, \theta]^{\frac{1}{q}} \hat{\nu}[-M, x]^{\frac{p-1}{p}},$$

$$B_\theta^+ = \sup_{y > \theta} \mu[\theta, y]^{\frac{1}{q}} \hat{\nu}[y, N]^{\frac{p-1}{p}}.$$

現在, 我開始證明我們的上界估計: $\boxed{k_{q,p} B^* \geq A}$ 。首先, 改寫 B^* :

$$\begin{aligned} B^* &= \sup_{x \leq y} \frac{\mu[x, y]^{\frac{1}{q}}}{\left[\hat{\nu}[-M, x]^{\frac{q(1-p)}{p}} + \hat{\nu}[y, N]^{\frac{q(1-p)}{p}} \right]^{\frac{1}{q}}} \\ &= \left\{ \sup_{x \leq y} \frac{\mu[x, y]}{\hat{\nu}[-M, x]^{\frac{q(1-p)}{p}} + \hat{\nu}[y, N]^{\frac{q(1-p)}{p}}} \right\}^{\frac{1}{q}} \\ &=: \left\{ \sup_{x \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} \right\}^{\frac{1}{q}} \end{aligned}$$

右方的分數可分拆, 然後使用合分比定理 (proportional property) 得出

$$\begin{aligned} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} &= \frac{\mu[x, \theta] + \mu(\theta, y)}{\varphi(x) + \psi(y)} \quad \boxed{\theta \in [x, y]} \\ &\geq \frac{\mu[x, \theta]}{\varphi(x)} \wedge \frac{\mu(\theta, y)}{\psi(y)}, \end{aligned}$$

此處 $x \wedge y = \min\{x, y\}$, 類似地, $x \vee y = \max\{x, y\}$ 。兩邊對 θ 取上確界, 得

$$\frac{\mu[x, y]}{\varphi(x) + \psi(y)} \geq \sup_{\theta \in [x, y]} \left\{ \frac{\mu[x, \theta]}{\varphi(x)} \wedge \frac{\mu(\theta, y)}{\psi(y)} \right\}.$$

再對 $x \leq y$ 取上確界, 注意兩個 \sup 可交換, 得出

$$\begin{aligned} \sup_{x \leq y} \frac{\mu[x, y]}{\varphi(x) + \psi(y)} &\geq \sup_{x \leq y} \sup_{\theta \in [x, y]} \left\{ \frac{\mu[x, \theta]}{\varphi(x)} \wedge \frac{\mu(\theta, y]}{\psi(y)} \right\} \\ &= \sup_{\theta} \sup_{\theta \in [x, y]} \{ \cdots \} \\ &= \sup_{\theta} \left\{ \left[\sup_{x \leq \theta} \frac{\mu[x, \theta]}{\varphi(x)} \right] \wedge \left[\sup_{y \geq \theta} \frac{\mu(\theta, y]}{\psi(y)} \right] \right\}. \end{aligned}$$

兩邊取 $1/q$ 次方, 完成了證明的第一步。

引理 2.3. $B^* \geq \sup_{\theta} (B_{\theta}^- \wedge B_{\theta}^+)$.

這裡我們做了什麼? 我們把兩個單邊的情況跟一個整體的雙邊的情況聯繫起來。然而, 右方只是最小值, 那是不夠的, 我們需要的是最大值。由測度 μ 的連續性假定 $\mu_{pp} = 0$, 存在 $\bar{\theta}$ 使得 $B_{\bar{\theta}}^- = B_{\bar{\theta}}^+$, 當然更有

$$B^* \geq \sup_{\theta} (B_{\theta}^- \wedge B_{\theta}^+) \geq B_{\bar{\theta}}^-.$$

於是

$$\begin{aligned} k_{q,p} B^* &\geq (k_{q,p} B_{\bar{\theta}}^-) \vee (k_{q,p} B_{\bar{\theta}}^+) \\ &\geq A_{\bar{\theta}}^- \vee A_{\bar{\theta}}^+ \quad (\text{上面所述的已知定理}) \\ &\geq \inf_{\theta} (A_{\theta}^- \vee A_{\theta}^+) \\ &\geq A \quad (\text{分裂技術}). \quad \square \end{aligned}$$

最後一步是兩個半區間上的最佳常數與整個區間上的最佳常數的比較, 是典型的分裂技術, 其證明也不長, 我們十多年前就會了, 此處不再解釋。我們指出: 最後這一步及開頭的定理 2.2, 都用到條件 $q \geq p$ 。

這個證明總共有三個要點, 第一個是分裂技術 (splitting technique); 第二個是 $\bar{\theta}$ 的使用, 做了一個過河拆橋的好事, 不需要將 $\bar{\theta}$ 解出, 過了河就拆了橋; 最後一個技巧就是使用合分比, 將上界估計證出來。這個證明做出來的時候, 真會讓人高興得跳起來, 因為是這麼漂亮又這麼簡單。所以這是一個讓我們很得意的證明。實際上, 當我說那個複雜的表達式很漂亮, 可能很多人心裡並不同意, 因為還是有點複雜。我說它漂亮, 是因為我喜歡, 因為證明是這麼漂亮, 所以當然會喜歡。

我已經證明了雙邊的 Hardy 的情況, 用概率 (機率) 的語言講, 這是 Transient 情形。還有 Ergodic 的情況。

$$\|f\|_{\mu, q} \leq A \|f'\|_{\nu, p} \quad \text{Transient,}$$

$$\|f - \mu(f)\|_{\mu,q} \leq A \|f'\|_{\nu,p} \quad \text{Ergodic.}$$

如果 μ 是概率測度，自然考慮 Ergodic 情況。若只是將兩邊等於 0 的邊界條件都去掉，那麼所得不等式是平凡的，因為恒等於 1 的常數 f 給出 $A = \infty$ 。所以中間要減去一個概率測度的平均值。這是概率論最重要的情形。這種情形讓我吃了許多苦頭，之前提到的三大工具、五個步驟就是證明這種情形的平行的結果。然而，剛才我們所用證明的方法完全可以套過來，所以我們那個方法在發表的論文中至少用了三次。它不僅能處理這種情況，還能處理 Log-Sobolev 不等式以及 Nash 不等式。所以有好多的故事。

3. 幾何應用

這最後十幾分鐘的時間我想跟大家講應用，這個 Hardy-type 不等式有甚麼用處。現在給大家一個很簡單情況的應用。如果 $q = p$ ，這些估計都能改進。如果 $q \neq p$ ，應該也可以改進，但是還沒做出來，這是一個開放的問題（校稿註：現已基本解決）。我們考慮如下的 $q = p = 2$ 的特殊例子：

$$\|f\|_{\mu,2} \leq A \|f'\|_{\nu,2} \quad \text{on } (-M, N), \quad f(N) = 0.$$

其中測度 μ 和 ν 來自微分算子：

$$L = a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx}, \quad C(x) = \int_{\theta}^x \frac{b}{a},$$

$$\boxed{\mu(dx)} = \frac{e^C}{a}(x)dx, \quad \boxed{\nu(dx)} = e^{C(x)}dx,$$

$\theta \in (-M, N)$ 是一個參考點。有趣的是，最佳常數可寫成

$$\boxed{A^2 = \lambda_0^{-1}},$$

其中 λ_0 是算子 L 的一個主特徵值：

$$Lf = -\lambda_0 f, \quad f \neq 0, \quad f(N) = 0.$$

此處，當 $M < \infty$ 時，還需補充邊界條件 $f'(-M) = 0$ 。這就建立了不等式與特徵值之間的關係。一般地講，不等式是特徵值的變分形式；而特徵值所呈現的特徵方程是使不等式成為等式所導出的方程。所以剛才講的所有關於不等式最佳常數的估計都是特徵值的估計。在當前情況下，有 μ 也有 ν ，我們真正用的不是 ν 而是 $\hat{\nu}$ ， $\hat{\nu}$ 在這裡特別簡單， $\hat{\nu}(dx) = e^{-C(x)}dx$ 。

我們有以下的一般結果：

定理 3.1. 對於給定的算子 L ，存在系列 $\{\underline{k}_n\}_{n \geq 1}$ 和 $\{\bar{k}_n\}_{n \geq 1}$ ，使得當 $n \uparrow$ 時，

$$\uparrow \underline{k}_n^{-1} \leq \lambda_0 = A^{-2} \leq \bar{k}_n^{-1} \downarrow.$$

換言之，我們可以構造出兩串數列，一串升，一串降，慢慢靠近 λ_0 。實際上，我相信它們都收斂於 λ_0 ，但是還沒有證出來。特別地，我們寫出當 $n = 1$ 時的下界和上界如下：

推論 3.2. 對於一般的二階微分算子 L ，我們有 $\underline{k}^{-1} \leq \lambda_0 \leq \bar{k}^{-1}$ ，其中

$$\begin{aligned} \underline{k} &= \sup_{x \in (-M, N)} \frac{1}{\sqrt{\psi(x)}} \int_{-M}^N \psi(\cdot \vee x) \sqrt{\psi} \, d\mu \\ \bar{k} &= \sup_{x \in (-M, N)} \frac{1}{\psi(x)} \int_{-M}^N \psi(\cdot \vee x) \psi \, d\mu \\ \psi(x) &= \hat{\nu}(x, N) = \int_x^N e^{-C} \, , \quad x \vee y = \max\{x, y\}. \end{aligned}$$

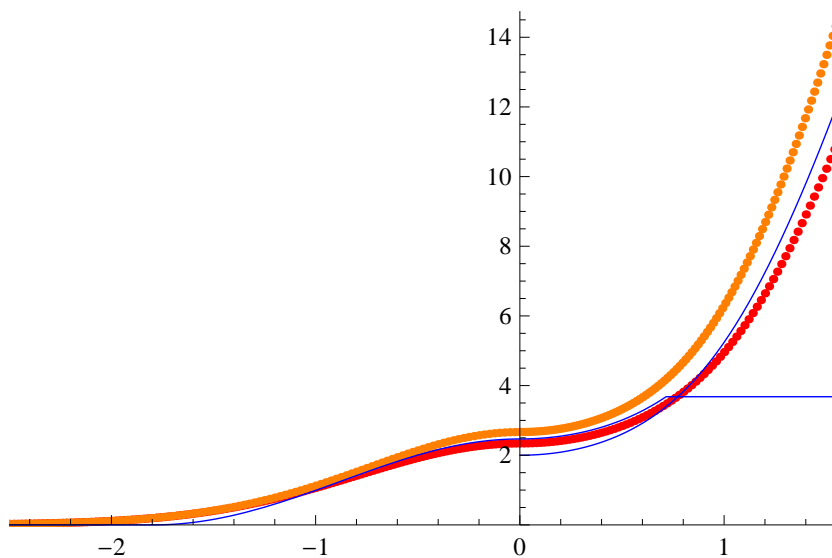
當然，這裏的上下界也是通過測度 μ 和 $\hat{\nu}$ 表出的，適用於一般的算子。下面考慮更特殊的特例。

例 3.3. 令 $(-M, N) = (0, 1)$,

$$a(x) \equiv 1 \text{ 和 } b(x) = -(d-1)\alpha \tanh(\alpha x), \quad x \in (0, 1),$$

其中 d 和 α 為給定的常數參數。

將常數 d 和 α 代入 \bar{k} 和 \underline{k} ，產生如下的圖形。



畫圖時，將常數 d 固定為 5，看上下界隨 α 的變動情況。因為 α 可能是純虛數（隨後給出表達式），所以這裏我們做了變換： $\alpha = \sqrt{-\text{sgn}(x)|x|}$ ， $x \in (-2.5, \pi/2)$ 。圖中有兩條粗一點的曲線，顏色深一點的是下界估計，顏色淺一點的（位於上方）是上界估計。從圖可以看出，兩條線

相差最大的地方，其倍數小於等於 2。左邊幾乎可以說是重疊，這是畫了圖後才發現的。一般情形在做很多例子的時候，上下界之比都小於等於 2，比 4 要改進一倍，但是沒想到會如此貼近。這個貼近是很不尋常的發現，因為負的時候更難，允許無限，在 -1 的地方就已經幾乎相疊，更別說更負的情形。

在這個圖中除了兩條粗一點的曲線外，還有兩條細一點的曲線。關於這兩條細曲線的故事就很長了。上圖的意義來自幾何。考慮緊的黎曼流形 (Riemann Manifold)，假設無邊，因為是緊的，所以譜值離散 (discrete spectrum)，考慮它的第一個非零的特徵值 λ_1 。要點是用三個幾何量做估計：流形維數 d ，流形直徑 D 和流形曲率的下界 K 。一個標準例子是 d 維單位球面：

$$\mathbb{S}^d : D = \pi, Ric = d - 1, \lambda_1 = d \quad \forall d \geq 1.$$

為方便，我們定義

$$\alpha = \frac{D}{2} \sqrt{\frac{-K}{d-1}} \quad \text{如 } K \neq 0, \quad \alpha = 0 \quad \text{如 } K = 0.$$

著名的 Myers 定理斷言：當 $K > 0$ 時， $|\alpha| \leq \pi/2$ 。

如開頭所說，在幾何中關於第一特徵值有相當完整的研究。當 $K \geq 0$ 時，代表性成果如下表。前三個結果對於單位高維曲面達到最優，其中的第三個最好；第四個是在單位圓、零曲率時達到最優，被認為是幾何中最艱深的結果。上一張圖中的兩條細曲線，在右方高一些的那一條就是下面的第三個結果。此處提及，這裏我們只列出最優估計而不可能談及其它結果，我們直奔目標而不顧細節。詳細文獻可在文末的論文中找到。

作者	下界估計: $K \geq 0$ 情形
A. Lichnerowicz (1958)	$\frac{d}{d-1} K$
P.H.Bérard, G.Besson & S. Gallot (1985)	$d \left\{ \frac{\int_0^{\pi/2} \cos^{d-1} t dt}{\int_0^{D/2} \cos^{d-1} t dt} \right\}^{2/d}, K = d - 1$
Chen & F.Y. Wang (1997)	$\frac{dK}{(d-1)(1 - \cos^d \alpha)}$
J.Q. Zhong & H.C. Yang (1984)	$\frac{\pi^2}{D^2}$

對於負曲率的情況，有三個最優估計。第二個優於第一個，它們都在 $\alpha = 0$ 處達到最優。第三

個在曲率很負時優於第二個, 但它們自身不能比較。

作者	下界估計: $K \leq 0$ 情形
H.C. Yang (1989) F. Jia (1991) $d \geq 5$ Chen & F.Y. Wang (1994)	$\frac{\pi^2}{D^2} e^{-(d-1)\alpha}$
Chen & F.Y. Wang (1997)	$\frac{1}{D^2} \sqrt{\pi^4 + 8(d-1)\alpha^2} \cosh^{1-d} \alpha$
Chen (1994)	$\frac{1}{D^2} ((d-1)\alpha \tanh \alpha \operatorname{sech} \theta)^2$

其中的 θ 可由如下方式得到: 定義

$$\begin{aligned} \theta_1 &= 2^{-1}(d-1)\alpha \tanh \alpha, \\ \theta_n &= \theta_1 \tanh \theta_{n-1}, \quad n \geq 2, \quad \text{則 } \theta_n \downarrow \theta. \end{aligned}$$

在零曲率的情況, 最優估計如下表。其中第二個優於第一個。第三個在零附近優於第二個, 誤差可以達到 10^{-7} , 非常精確。

作者	下界估計: $K \approx 0$ 情形
11 authors	$\frac{\pi^2}{D^2} + \frac{K}{2}, \quad K \in \mathbb{R}$
Y.M. Shi & H.C. Zhang (2007)	$\sup_{s \in (0,1)} s \left[4(1-s) \frac{\pi^2}{D^2} + K \right] \quad (*)$
Chen & E. Scacciatelli & L. Yao (2002)	$\frac{\pi^2}{D^2} + \frac{K}{2} + (10 - \pi^2) \frac{K^2 D^2}{16}, \quad K \leq \frac{4}{D^2}$

其中的第二個結果有顯式:

$$(*) = \begin{cases} \left(\frac{\pi}{D} + \frac{KD}{4\pi} \right)^2, & -4\pi^2 \leq KD^2 \leq 4\pi^2 \\ K, & KD^2 \in (4\pi^2, (d-1)\pi^2] \\ 0, & KD^2 < -4\pi^2 \end{cases}$$

總而言之, 從 1958 年到現在的這五、六十年間所得到的 10 個最優估計中, 經過比較, 以下 5 個

結果領先。

作者	下界估計
Chen & F.Y. Wang (1997)	$\frac{dK}{(d-1)(1-\cos^d \alpha)}, K \geq 0$
Chen & F.Y. Wang (1997)	$\frac{1}{D^2} \sqrt{\pi^4 + 8(d-1)\alpha^2} \cosh^{1-d}\alpha$
Chen (1994)	$\frac{1}{D^2} ((d-1)\alpha \tanh \alpha \operatorname{sech} \theta)^2$
Y. Shi & H.C. Zhang (2007)	$\sup_{s \in (0,1)} s \left[4(1-s) \frac{\pi^2}{D^2} + K \right]$
Chen & E. Scacciatelli & L. Yao (2002)	$\frac{\pi^2}{D^2} + \frac{K}{2} + (10-\pi^2) \frac{K^2 D^2}{16}, K \leq \frac{4}{D^2}$

現在，我們可以給出上圖的完整的解釋。我們給出了新的統一的下界估計，它通常優於此表中的第二、第三個估計。在右端點近旁，可用此表中的第一個估計作補充；在零點近旁，可用此表中第四個估計的改進形式作補充（它即是圖中的另一細曲線）。如前所述，表中的第五個估計很精確，但它只適用於零點的一個小鄰域，此處不再討論。總之，此圖已顯示出我們對於幾何所獲得的進步：一個統一的下界加上兩個補充。圖中的上界估計最初的用意是作為判斷下界估計優劣的參照物。然而，圖中顯示出上、下界的曲線十分相似，這就啟發我們去做凸平均 (convex mean)。挑選凸平均的係數時，自然選一個過零點，選另一個過右端點，因為這兩點都是最優點。於是我們得到兩條新曲線 $\underline{\eta}$ 和 $\bar{\eta}$ ，依賴於 α 和 d ，滿足 $\underline{\eta} \leq \lambda_0 \leq \bar{\eta}$ 。它們都是 \bar{k}^{-1} 和 \underline{k}^{-1} 的凸平均：

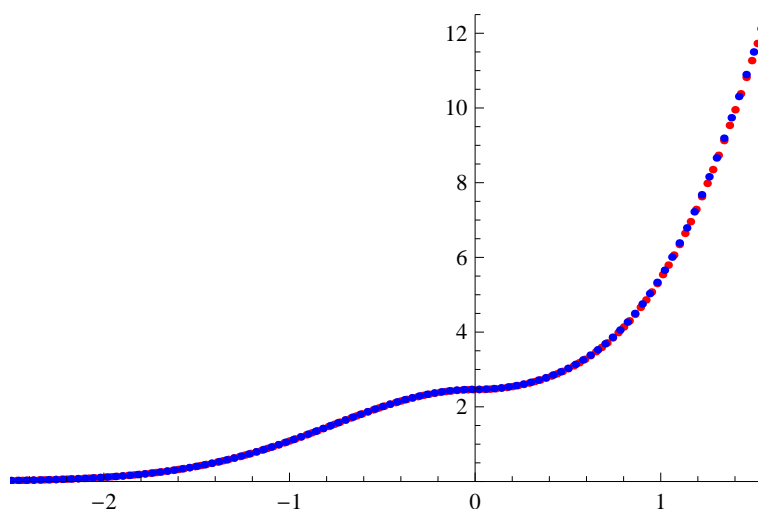
$$\eta = \gamma \bar{k}^{-1} + (1 - \gamma) \underline{k}^{-1}$$

$$\bar{\gamma} = \frac{5^{3/2} - 5\pi^2/16}{5^{3/2} - 10/3} \approx 0.39 \quad \boxed{\text{曲線經過 } |\alpha| = 0} \quad (\text{與 } \alpha, d \text{ 無關})$$

$$\underline{\gamma} = \frac{d\pi^2/4 - \underline{k}^{-1}}{\bar{k}^{-1} - \underline{k}^{-1}} \Big|_{|\alpha|=\pi/2} \quad \boxed{\text{曲線經過 } |\alpha| = \frac{\pi}{2}} \quad (\text{只與 } d \text{ 有關})$$

$$\approx 0.367 \text{ if } d = 5$$

這兩個凸平均所畫出的圖如下：



這兩個上、下界，在圖形上幾乎看不出差別，相差非常小。如果查閱本文的電子版，你會發現這裏實際上有紅和藍兩條曲線。換句話說，我們差不多把馬上要講到的 $\bar{\lambda} = \lambda_0$ 找出來了。使用三個幾何量 d , D 和 K 的估計，差不多也就這樣了。取得這樣的結果是預先難以想像的，回想五、六十年來的那麼多人的努力，可見來之不易。在這個過程中，只有最後一步做凸平均時用到特定的幾何行為，其它結論都適用於遠為一般的情形。今天所講的所有內容都是一維的，至於怎麼樣從高維化爲一維，下面的比較定理完全是用機率方法做出來，是非常機率的（參見附錄）。

定理 3.4 (比較定理)。我們有

$$\lambda_1 \geq \frac{4}{D^2} \bar{\lambda} \quad (\text{陳和王鳳雨 1994: 耦合方法}),$$

此處 $\bar{\lambda}$ 是算子 \bar{L} 的主特徵值：

$$\bar{L} = \frac{d^2}{dx^2} + (d-1)\alpha \tanh(\alpha x) \frac{d}{dx}, \quad f(0) = 0, \quad f'(1) = 0.$$

其次，我們有

$$\bar{\lambda} = \lambda_0 \quad (\text{陳 2011: 對偶方法}),$$

其中 λ_0 是算子 L 的主特徵值：

$$L = \frac{d^2}{dx^2} - (d-1)\alpha \tanh(\alpha x) \frac{d}{dx}, \quad f'(0) = 0, \quad f(1) = 0.$$

由以上應用也許可以看出，機率這門學科現在已經充分地長大，如果一個數學系裡面沒有機率，是非常危險的。這話我不只在這裡說，早年對大陸的名校的朋友也說過。以上的所有這些方法，對幾何都有用。我講的是緊流形，沒有邊界，就算是凸邊界 (convex boundary) 也是一樣。事實上，就算是非緊流形，也類似，只是需要添加一個向量場。我看到最新的是一個 Finsler-Laplacian 估計，結構也是一樣，只不過要用不同記號而已。謝謝大家。

附錄：耦合方法與概率度量

這次演講集中於這兩年的新進展，幾乎未涉及概率方法，後者以前訪問臺灣時曾多次講過。這裏略作介紹，以饗讀者。

A.1. 測度的耦合

回顧在學微積分學的重積分的時候，總是利用累次積分，化成比較容易處理的單重積分來計算。但也有特別的例外，把單重積分提升為多重積分來計算。最典型的例子是計算正態分布

$$\int_{\mathbb{R}} e^{-x^2/2} dx = \sqrt{2\pi},$$

為計算左方的單重積分。我們先將它提升為二重積分

$$\int_{\mathbb{R}} \int_{\mathbb{R}} e^{-(x^2+y^2)/2} dx dy,$$

此時使用極坐標很容易算出結果。這個巧妙的算法屬於 Simón-Denis Poisson (1781~1840)。換言之，我們把單個測度 dx 提升為乘積測度 $dx \times dy$ 。也可以說，我們把單個概率測度（正態分布） $\mu(dx)$ 提升為乘積空間 $\mathbb{R} \times \mathbb{R}$ 上的乘積概率測度 $\mu(dx) \times \mu(dy)$ 。這已經是一種特殊的耦合，稱為獨立耦合。

想想為什麼這種耦合的應用很有限，原因就在於“獨立性”的要求太強了。只要脫離獨立性，容許相關性，那就是一片廣闊新天地。給定兩個可測空間 (E_k, \mathcal{E}_k) 上的概率測度 $\mu_k, k = 1, 2$ ，稱乘積空間 $(E_1 \times E_2, \mathcal{E}_1 \times \mathcal{E}_2)$ 上的概率測度 $\tilde{\mu}$ 為 μ_1 和 μ_2 的耦合，如果下述邊緣性條件成立：

$$\begin{aligned} \tilde{\mu}(A_1 \times E_2) &= \mu_1(A_1), & A_1 \in \mathcal{E}_1, \\ \tilde{\mu}(E_1 \times A_2) &= \mu_2(A_2), & A_2 \in \mathcal{E}_2. \end{aligned}$$

近乎平凡的耦合是獨立耦合： $\tilde{\mu}_0 = \mu_1 \times \mu_2$ 。它卻有非平凡的應用。我們斷言直線上的每一個概率測度 μ ，總滿足如下的 **FKG** 不等式：

$$\int_{\mathbb{R}} fg d\mu \geq \int_{\mathbb{R}} f d\mu \int_{\mathbb{R}} g d\mu, \quad f, g \in \mathcal{M},$$

其中 \mathcal{M} 是 \mathbb{R} 上有界單調增函數的全體。使用獨立耦合 $\tilde{\mu}_0 = \mu \times \mu$ 和直線的全序性，證明只需一行：

$$\iint_{\mathbb{R} \times \mathbb{R}} \tilde{\mu}_0(dx, dy)[f(x) - f(y)][g(x) - g(y)] \geq 0, \quad f, g \in \mathcal{M}.$$

展開左方的積分，即得所求。

下面是一種很有用的基本耦合 $\tilde{\mu}_b$ 。設 $E_k = E$ (Hausdorff 空間), $k = 1, 2$ 。記 Δ 為 E 的對角線集: $\Delta = \{(x, x) : x \in E\}$ 。命

$$\tilde{\mu}_b(dx_1, dx_2) = (\mu_1 \wedge \mu_2)(dx_1)I_\Delta + \frac{(\mu_1 - \mu_2)^+(dx_1)(\mu_1 - \mu_2)^-(dx_2)}{(\mu_1 - \mu_2)^+(E)}I_{\Delta^c},$$

其中 ν^\pm 為符號測度 ν 的 Jordan-Hahn 分解, 而 $\nu_1 \wedge \nu_2 = \nu_1 - (\nu_1 - \nu_2)^+$ 。特別地, 如 μ_k 關於 Lebesgue 測度有密度 h_k , 則 $(\mu_1 - \mu_2)^\pm$ 和 $\mu_1 \wedge \mu_2$ 亦然: 分別有密度 $(h_1 - h_2)^\pm$ (函數的正、負部) 和 $h_1 \wedge h_2 := \min\{h_1, h_2\}$ 。每一種耦合都有它的基本特徵。這裏是第一項。有了它之後, 由邊緣性自然導出第二項。因為 $(\mu_1 - \mu_2)^+$ 表示 μ_1 比 μ_2 多出的部分, $\mu_1 \wedge \mu_2$ 是最大的共有部分。此耦合的第一項就是把這個共有部分都放到對角線上。特別地, 如果 $\mu_1 = \mu_2$, 則第二部分消失。基本耦合的重要性在於它刻畫了概率測度的全變差距離。詳言之, 暫設 ρ 為離散距離:

$$\rho(x, y) = \begin{cases} 1 & x \neq y, \\ 0 & x = y. \end{cases}$$

則我們有如下的 Dobrushin 定理:

$$\int_{E \times E} \rho d\tilde{\mu}_b =: \tilde{\mu}_b(\rho) = \inf_{\tilde{\mu} \in \mathcal{C}(\mu_1, \mu_2)} \tilde{\mu}(\rho) = \frac{1}{2} \|\mu_1 - \mu_2\|_{\text{Var}} \quad (= (\mu_1 - \mu_2)^+),$$

其中 $\mathcal{C}(\mu_1, \mu_2) := \{\tilde{\mu} : \tilde{\mu} \text{ 是 } \mu_1 \text{ 和 } \mu_2 \text{ 的耦合}\}$ 。對於一般的距離函數 ρ , 如下的定義

$$W(\mu_1, \mu_2) = \inf_{\tilde{\mu} \in \mathcal{C}(\mu_1, \mu_2)} \tilde{\mu}(\rho)$$

稱為 μ_1 和 μ_2 的 **Wasserstein** 距離。這是 1969 年在研究隨機場 (概率論與統計物理的交叉學科) 時引進的, 隨後逐步成為研究諸多數學課題 (特別是無窮維) 的基本工具。

進一步, 將上述距離函數 ρ 換成更一般的費用函數 c (非負):

$$\inf_{\tilde{\mu} \in \mathcal{C}(\mu_1, \mu_2)} \int_{E \times E} c(x, y) \tilde{\mu}(dx, dy),$$

並考察輸運函數 $T : E \rightarrow E$, 使得 μ_2 是在 μ_1 之下、由 T 導出的分布, 而且 $\int_E c(x, T(x)) \mu_1 dx$ 達到上式的下確界, 這就構成了當今 PDE 中很熱鬧的研究方向: 最優輸運 (optimal transport)。

至此為止, 我們講的是靜態情形 (即不含時間 t)。下面轉入動態情形。對於給定的兩個馬氏過程, 依照上述方法, 我們可以定義它們對於固定的不同出發點、在固定的同一時間 t 的分布 (概率測度) 的耦合。然而, 這種耦合未必是馬氏的。雖然非馬氏耦合有它的用處, 但為簡單計, 此處我們只考慮馬氏耦合 (即給定邊緣過程是馬氏的, 要求耦合過程也是馬氏的)。

A.2. 馬氏過程的耦合

給定 (E_k, \mathcal{E}_k) 的馬氏半群 $P_k(t)$ 或轉移概率 $P_k(t, x_k, \cdot)$, $k = 1, 2$, 稱乘積空間上的馬氏半群 $\tilde{P}(t)$ 或轉移概率 $\tilde{P}(t; x_1, x_2; \cdot)$ 是 $P_k(t)$ 或 $P_k(t, x_k, \cdot)$ ($k = 1, 2$) 的耦合, 如果下述 (關於過程的) 邊緣性成立:

$$\begin{aligned}\tilde{P}(t; x_1, x_2; A_1 \times E_2) &= P_1(t, x_1, A_1), \\ \tilde{P}(t; x_1, x_2; E_1 \times A_2) &= P_2(t, x_2, A_2), \quad t \geq 0, x_k \in E_k, A_k \in \mathcal{E}_k, k = 1, 2.\end{aligned}\tag{MP}$$

等價地,

$$\begin{aligned}\tilde{P}(t)f(x_1, x_2) &= P_1(t)f(x_1), \quad f \in {}_b\mathcal{E}_1 \\ \tilde{P}(t)f(x_1, x_2) &= P_2(t)f(x_2), \quad f \in {}_b\mathcal{E}_2, t \geq 0, x_k \in E_k, k = 1, 2,\end{aligned}\tag{MP}$$

此處 ${}_b\mathcal{E}$ 是有界 \mathcal{E} 可測函數的全體; 在等式的左邊, 原為單變量的函數 f 都被視為雙變量函數。

上述定義“(MP)”實際上沒有多少用處, 因為 $P_k(t)$ 或 $P_k(t, x_k, \cdot)$ ($k = 1, 2$) 都是未知的。然而, 從它們出發, 自然導出如下關於算子的耦合。

A.3. 馬氏過程的耦合

分別以 L_k ($k = 1, 2$) 和 \tilde{L} 表邊緣半群 $P_k(t)$ ($k = 1, 2$) 和耦合半群 $\tilde{P}(t)$ 的形式無窮小生成元, 在 (MP) 兩邊關於 t 在 0 處取導數, 得出如下 (關於算子) 的邊緣性:

$$\begin{aligned}\tilde{L}f(x_1, x_2) &= L_1f(x_1), \quad f \in \mathcal{F}_1, \\ \tilde{L}f(x_1, x_2) &= L_2f(x_2), \quad f \in \mathcal{F}_2, x_k \in E_k, k = 1, 2;\end{aligned}\tag{MO}$$

其中 \mathcal{F}_k 是 ${}_b\mathcal{E}_2$ 的適當子集; 與前面的 (MP) 一樣, 等式左邊的函數 f 都被視為雙變量函數。

現在, 對於馬氏過程的馬氏耦合, 我們有了算子耦合的可行手段。例如, 對於 \mathbb{R}^d 上的擴散過程, 我們有二階微分算子

$$L = \sum_{i,j=1}^d a_{ij}(x) \frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^d b_i(x) \frac{\partial}{\partial x_i}.$$

為簡單計, 寫成 $L \sim (a(x), b(x))$ 。今給定邊緣算子 $L_k \sim (a_k(x), b_k(x))$, $k = 1, 2$, 則由耦合算子的邊緣性, 乘積空間 $\mathbb{R}^d \times \mathbb{R}^d$ 上的耦合算子 $\tilde{L} \sim (a(x, y), b(x, y))$ 應有如下形式:

$$a(x, y) = \begin{pmatrix} a_1(x) & c(x, y) \\ c(x, y)^* & a_2(y) \end{pmatrix}, \quad b(x, y) = \begin{pmatrix} b_1(x) \\ b_2(y) \end{pmatrix},$$

其中矩陣 $c(x, y)^*$ 是 $c(x, y)$ 的轉置, 條件是保證 $a(x, y)$ 非負定。更具體些, 設 $L_1 = L_2 \sim (a(x), b(x))$, $a(x) = \sigma(x)\sigma(x)^*$, $\det \sigma \neq 0$ 。則可取

$$c(x, y) = \sigma(x)[I - 2\bar{u}\bar{u}^*]\sigma(y)^*, \quad x \neq y,$$

其中 $\bar{u} = (x - y)/|x - y|$ 。矩陣 $I - 2\bar{u}\bar{u}^*$ 是反射矩陣 (即行列式為 -1 的正交矩陣)。這個耦合稱為反射耦合。

A.4. 應用於特徵值估計

回顧 $\{P_t\}_{t \geq 0}$ 與自身的耦合 $\{\tilde{P}_t\}_{t \geq 0}$ 滿足邊緣性: 對於一切 $f \in C_b^2(\mathbb{R}^d)$ 及 (x, y) ($x \neq y$), 有

$$\tilde{P}_t f(x, y) = P_t f(x) \quad (\text{相應地, } \tilde{P}_t f(x, y) = P_t f(y)), \quad (\text{A.1})$$

此處還是將 f 視為雙變量函數。我們由此出發來證明特徵值估計。

第一步. 設 g 是相應於 λ_1 的算子 $-L$ 的特徵函數: $-Lg = \lambda_1 g$ 。則由關於半群的標準的微分方程 (即 Kolmogorov 向前微分方程), 我們有

$$\frac{d}{dt} P_t g(x) = P_t Lg(x) = -\lambda_1 P_t g(x).$$

固定 g 和 x , 解這個關於函數 $t \rightarrow P_t g(x)$ 的常微分方程, 得出

$$P_t g(x) = g(x) e^{-\lambda_1 t}, \quad t \geq 0. \quad (\text{A.2})$$

這個恆等式很漂亮, 它將特徵值、特徵函數和半群三者統一到一個簡潔的公式裡。可惜此刻這個公式無用, 因為三者均未知。

第二步. 幸運的是: 耦合方法使得上述公式變得強有力。我們所需要的只是關於耦合算子的下述估計:

$$\tilde{L}\rho(x, y) \leq -\alpha\rho(x, y), \quad x \neq y, \quad (\text{A.3})$$

其中 $\alpha > 0$ 為常數。它等價於

$$\tilde{P}_t \rho(x, y) \leq \rho(x, y) e^{-\alpha t}, \quad t \geq 0. \quad (\text{A.4})$$

今考慮緊空間情形。此時 g 關於 ρ Lipschitz 連續。記其 Lipschitz 常數為 c_ρ 。命 $g_1(x, y) = g(x)$, $g_2(x, y) = g(y)$ 。那麼, 我們有

$$\begin{aligned} e^{-\lambda_1 t} |g(x) - g(y)| &= |P_t g(x) - P_t g(y)| \quad (\text{由 (A.2)}) \\ &= |\tilde{P}_t g_1(x, y) - \tilde{P}_t g_2(x, y)| \quad (\text{由 (A.1)}) \\ &= |\tilde{P}_t (g_1 - g_2)(x, y)| \quad (\text{因耦合在同一乘積空間上}) \\ &\leq \tilde{P}_t |g_1 - g_2|(x, y) \\ &\leq c_\rho \tilde{P}_t \rho(x, y) \quad (\text{Lipschitz 性}) \\ &\leq c_\rho \rho(x, y) e^{-\alpha t} \quad (\text{由 (A.4)}). \end{aligned}$$

因為 g 非常數, 存在 $x \neq y$ 使得 $g(x) \neq g(y)$ 。固定這樣一對 (x, y) , 然後令 $t \rightarrow \infty$, 必然得出 $\lambda_1 \geq \alpha$ 。經過這麼簡單的兩步, 我們就得到了所需的下界估計。

上面的緊性假定是可以避免的, 只需使用局部化程序。因此, 我們的方法適用於非常一般的情形。作為示例, 考慮 O.U. 過程:

$$L = \sum_{i=1}^d \left(\frac{\partial^2}{\partial x_i^2} - x_i \frac{\partial}{\partial x_i} \right).$$

取 $\rho(x, y) = |x - y|$ (通常的歐氏距離), 稍作計算, 便可看出條件 (A.3) 對於反射耦合及 $\alpha = 1$ 成立, 此時達到精確估計: $\lambda_1 = 1 = \alpha$ 。應當指出, 在一般情況下, 由條件 (A.3) 可見, 為得到好的下界 α , 不僅需要選擇好的耦合, 還需要選取好的距離。不難想象, 這裏有很多故事。

多年前, 我們曾經感到很奇怪, 為什麼文獻上處理的幾乎都是 Dirichlet 邊界情形, 後來才明白這種情形等價於極大值原理: 其特徵函數在內部為正、在邊界上達到最小值 0。對於 λ_1 , 特徵函數的零值曲面在區域內部, 極大值原理不適用。所以, 這裏的關鍵點是用耦合方法這一概率論工具代替了極大值原理。

致謝: 作者感謝中研院數學所的邀請 (hosts: 黃啓瑞、許順吉), 感謝黃、許、姜祖恕、周雲雄和他們夫人們的熱情款待。感謝陳隆奇、徐洪坤、羅夢娜、郭美惠、李育嘉及夫人、許元春和胡殿中等教授及其團隊的邀請和熱情款待。趁此機會, 也感謝數學所的其他同仁, 秘書組、圖書館等行政團隊所提供的幫助和溫暖如家的環境, 感謝陳麗伍助理為整理我的幾篇文章所付出的辛勞。同時, 作者感謝中國國家自然科學基金重點項目 (No. 11131003) 和教育部 973 項目的資助。

參考資料

1. Chen, Mu-Fa (2011). *General estimate of the first eigenvalue on manifolds*. *Front. Math. China* 6(6): 1025-1043.
2. Chen, Mu-Fa (2013). *Bilateral Hardy-type inequalities*. *Acta Math. Sin., Eng. Ser.* 29, no. 1, 1-32.

—本文演講者陳木法教授任教於北京師範大學數學科學學院, 整理者陳麗伍為中央研究院數學研究所助理—