認識極值問題的對偶現象

王進賢

一、引言

談起對偶(duality),它可是個很叫座的名詞,許多地方都有它的踪跡,例如,布氏代數(Boolean algebra)、投影幾何、電路、博戲論(Game theory)、經濟學、圖形學(Graph theory)等等(註一)。而本文所要談的是極值問題的對偶現象,我們將討論“在某些限制下求 n 元實值函數的極小值”這類問題的對偶問題。通常在數學計劃(Mathematical programming)這一類的書中總是先“定義”出原問題的對偶問題,然後再加以證明兩個問題等值(註二),這樣的做法讓人有“對偶問題是從天上掉下來的”的感覺。本文中我們將仿照參考資料〔1〕的做法,透過Lagrange乘數法則(multiplier rule)把對偶問題“推導”出來。

二、逗留值問題的變化

我們先從求二元實函數的逗留值(stationary value)問題出發。我們說函數 $f(x,y)$ 在點 (x_0,y_0) 逗留若且唯若 $df(x_0,y_0)=0$,也就是 $f_x(x_0,y_0)=f_y(x_0,y_0)=0$。為簡單方便起見,底下所談函數均有良好的連續和可微分等現象(註三)。我們要討論的問題是:

(Ⅰ) 考慮函數 $f(x,y)$ 在 (x,y) 要滿足限制 $g(x,y)=0$ 的情況下的逗留現象。

這個問題和下述問題

(Ⅱ) 考慮三元函數 $F(x,y,\lambda)=f(x,y)+\lambda g(x,y)$ 的逗留現象。

等值(亦即有相同的逗留點 (x_0,y_0))。因為使問題(Ⅱ)有逗留點的情況也就是 $dF(x,y,\lambda)=0$ 有解,而這個微分式相當於 $F_x=0$,$F_y=0$ 和 $F_\lambda=0$,那也就是 $f_x(x,y)+\lambda g_x(x,y)=0$,$f_y(x,y)+\lambda g_y(x,y)=0$ 和 $g(x,y)=0$。根據一個顯而易明的原則:“增加一些原問題所必須滿足的條件到原問題上去所新形成的問題和原問題有相同的逗留現象。”,因此問題(Ⅱ)實際上是和“求在 $g(x,y)=0$ 的限制下,$F(x,y,\lambda)$ 的逗留現象”這問題等值,而這個問題又可看成問題(Ⅰ)。假若我們附加 $F(x,y,\lambda)$ 逗留的另兩個條件,我們就得到了下述等值問題。

(Ⅲ) 考慮 $F(x,y,\lambda)=f+\lambda g$ 在限制 $f_x+\lambda g_x=0$ 和 $f_y+\lambda g_y=0$ 成立的情況下的逗留現象。

假若在逗留點附近,從兩個限制我們能把 x、y 表為 λ 的函數(註四),那麼 $F(x,y,\lambda)$ 就換成 λ 的函數 $\psi(\lambda)$,而我們也就進入問題。

(Ⅵ) 考慮 $\psi(\lambda)$ 的逗留現象。

三、極小問題的對偶

上面的結果告訴我們考慮(Ⅰ)的逗留現象與考慮(Ⅱ)或(Ⅲ)或(Ⅳ)的逗留現象一樣。現在我們要問，如果考慮的是極大或極小問題，仿照上面這些推演結果會有何種變化？考慮問題(Ⅰ)的極小和考慮問題(Ⅳ)的極小還是極大有關？我們的答案是後者。為了考慮問題(Ⅰ)的極小現象(稱它為問題(Ⅰ′))，我們引進問題

(Ⅱ′)考慮三元函數 $F(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ 的極小現象。

假設(Ⅱ′)在點 $(\bar{x}, \bar{y}, \bar{\lambda})$ 處有相對極小值(以下均簡稱極小值) $d = F(\bar{x}, \bar{y}, \bar{\lambda})$ (註五)，則問題(Ⅰ′)在點 $(\bar{x}, \bar{y})$ 處有極小值 d。現在讓我們來轉一轉思路，假設我們讓 $\bar{\lambda}$ 做小幅度變動，就稱它為 λ，對於每一個 λ (暫時固定)設 $F(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ 的極小值為 d_λ，則 $d_\lambda \leq d$ (因為 d 相當於由目前固定 λ 的 $F(x, y, \lambda)$ 函數加上 $g(x, y) = 0$ 的限制下所得的極小，能參與比較的 (x, y) 較少，當然其所得極小值不小於沒加限制的)。另外，因為 $d = d_{\bar{\lambda}}$，因此 $d = \max(d_\lambda)$。上述過程中，給一個 λ 對應一個 d_λ 值可定為函數 $\phi(\lambda)$，也就是 $\phi(\lambda) = \min F(x, y, \lambda)$，它就是我們通常所稱〔問題(Ⅰ′)〕的對偶函數(dual function)(見參考資料〔2〕)。$\phi(\lambda)$ 能用明式表達出來嗎？因為 $\phi(\lambda) = \min F(x, y, \lambda) = \min[f(x, y) + \lambda g(x, y)]$ 而對於給一個固定 λ 找 $f(x, y) + \lambda g(x, y)$ 的極小值這件事是和問題

(Ⅲ′)固定 λ，考慮 $F = f + \lambda g$ 在限制 $f_x + \lambda g_x = 0$ 和 $f_y + \lambda g_y = 0$ 的情況下的極小值。

這件事等值的。如果從 $f_x + \lambda g_x = 0$ 和 $f_y + \lambda g_y = 0$ 能唯一地將 x、y 表為 λ 的函數，則給一個 λ 定值只能找一組 (x, y)，因此 $\phi(\lambda)$ 就等於將表為 λ 函數的 x，y 代入 $F(x, y, \lambda)$ 即可得，而我們原來的問題(Ⅰ′)就和下述對偶問題

(Ⅳ′) $\max \phi(\lambda)$

等值。

四、兩個例子

例1：求在 $(x-3)^2 + y^2 = 5$ 的限制下，$f(x, y) = -xy$ 的極小。

我們知道極小點必須滿足

$$\begin{cases} f_x + \lambda g_x = -y + \lambda(2x - 6) = 0 & \cdots\cdots\cdots\cdots ① \\ f_y + \lambda g_y = -x + \lambda(2y) = 0 & \cdots\cdots\cdots\cdots ② \\ (x-3)^2 + y^2 = 5 \end{cases}$$

解之得 $x = 4$，$y = 2$，$\lambda = 1$ (另外從判別是否極小的充分條件得知該點實為極小點)，f 的極小值為 -8。

讓我們來考慮本問題的對偶函數，從①②得 $x = \dfrac{12\lambda^2}{4\lambda^2 - 1}$，$y = \dfrac{6\lambda}{4\lambda^2 - 1}$，故從(Ⅲ′)得知

$$\phi(\lambda) = \min\{-xy + \lambda[(x-3)^2 + y^2 - 5]\} = \frac{4\lambda + 4\lambda^3 - 80\lambda^5}{(4\lambda^2 - 1)^2}$$

(當 $\lambda > \frac{1}{2}$ 時成立，因為需要在 $\lambda = 1$ 的附近鄰域)

而 $\max \phi(\lambda)$ 也是在 $\lambda = 1$ 處得值 -8。

讓我們再舉個例子來看一看找對偶問題的幾何解釋。

例2：求在$g(x,y)=2x=0$的限制下，$f(x,y)=(x+1)^2+y^2$的極小點。

從幾何觀點來看，本問題是找拋物面$z=(x+1)^2+y^2$ 與平面$x=0$交集曲線的最低點（也就是一鉛直拋物線的頂點）。我們可以觀察得到$(x,y)=(0,0)$時有最小值1。再讓我們來看當λ固定時，拋物面$z=f+\lambda g=(x+\lambda+1)^2+y^2-(\lambda+1)^2+1$總是包含上說拋物線，而拋物面的頂點$(-\lambda-1,0,-(\lambda+1)^2+1)$總是比拋物線的頂點$(0,0,1)$低。讓$\lambda$變動，拋物面的頂點再怎麼昇高，也不會高過$(0,0,1)$。因此拋物線的頂點是爲所有拋物面$z=f+\lambda g$的最低點（頂點）中的最高點。

五、結　語

上面所討論二元函數在一個限制下的極值問題的對偶是很容易推廣到n元函數m個限制的情況。本文的主要目的是在喚起讀者的興趣，因此，我們不再討論下去，有興趣的讀者請參閱參考資料〔2〕，該書中另有一些變形情況的討論，如等式限制改爲不等式限制，部份對偶等等。又，本文中，討論極值問題的對偶這種處理方式亦可應用到處理 simplest variational problems 的對偶，有興趣的讀者請參閱參考資料〔1〕，同書中另有quadratic variational problems 等對偶現象的討論。

註一：布氏代數和投影幾何的對偶現象很容易從有關的書找到，請自行參閱。在G. Hodley的Linear Programming pp 226～227（Addison Wesley 1963）可找到一個在電路上的例子；同書第十一章另有線性計劃問題對偶現象在零總和二人式博戲論的解說；第十三章有在經濟學上的解說。筆者所寫“且談一筆劃和一線牢”（數學傳播二卷一期pp. 85～86）可做爲在圖形學上的一個例子。

註二：等值的意義是有相同的極值。另請參考資料〔2〕。

註三：在逗留值問題這裏，我們只要f、g均爲可微分連續卽可，在底下極值問題裏，我們將要求f，g的條件需滿足Lagrange乘數法則成立所需條件（請參看T.M. Apostol Mathematical Analysis pp. 152～156）。

註四：只要滿足隱函數定理條件卽可。

註五：我們是在極小値存在的情況下討論下去的，至於何種情況下極小值存在，請參閱參考資料〔2〕。

參考資料

〔1〕R. Courant Methods of Mathematical physics Vol. I pp. 231～242，252～257。

〔2〕D.G. Luenberger Introduction to Linear and Nonlinear Programming pp. 312～316。

（本文作者現為海軍官校數學系助教）