Några kommentarer om optimering under bivillkor

Thomas Andrén

Ett vanligt optimeringsproblem bygger på att man vill finna de variabelvärden som gör att en funktion tar ett så stort eller litet värde som möjligt (inom sin värdemängd). Denna funktion kallas ofta för ***målfunktion*** och problemet brukar tecknas på följande sätt:

$$\underset{x,y}{\text{Max}}\, z = f(x,y)$$

Här utgör $f(x, y)$ vår målfunktion, och problemet består i att finna x och y som gör att z blir så stor som möjligt. Detta är ett exempel på ***optimering utan bivillkor***. Låt oss nu säga att vi vill begränsa definitionsmängden för f genom att införa ett villkor för vilka värden x och y kan ta. I beskrivningen ovan tänkte vi oss att x och y kunde ta vilka värden som helst. Vi skulle till exempel kunna tänka oss att följande villkor skulle gälla: $x+y=10$. Om vi optimerar vår funktion f samtidigt som vi inför denna begränsning för x och y så har vi utfört en ***optimering under bivillkor***. Den typ av bivillkor som behandlas under kursen avser likhetsvillkor av typen $g(x,y)=c$. Dvs. bivillkoret som måste gälla avser en funktion beroende av två eller fler variabler som är lika med en viss konstant c. För att lösa denna typ av problem kan vi använda oss av ***Lagranges multiplikator metod***. Denna lösningsmetod består av följande steg:

Steg 1: Formulera Lagranges funktion
Steg 2: Bestäm stationära punkter för denna funktion
Steg 3: Klassificera de stationära punkterna

Ett optimeringsproblem under bivillkor formuleras på följande sätt:

$$\underset{x,y}{\text{Max}}\, z = f(x,y) \qquad \text{(Målfunktion)}$$
$$u.b. \quad g(x,y) = c \qquad \text{(Bivillkor,u.b=under bivillkor)}$$

Lagranges funktion sammanför målfunktionen och bivillkoret i en gemensam funktion. Denna gemensamma funktion optimeras på vanligt sätt. Lagranges funktion kan formuleras på följande två sätt:

$$L(x,y,\lambda) = f(x,y) + \lambda(c - g(x,y)) \tag{1}$$

eller

$$L(x,y,\lambda) = f(x,y) - \lambda(g(x,y) - c) \tag{2}$$

Vilken formulering man använder är godtyckligt, bara man inte blandar ihop dem. I denna framställning använder vi uteslutande formulering (1). Det speciella med Lagranges funktion är att den introducerar ytterligare en variable λ, som brukar kallas för ***Lagranges multiplikator***. Tolkningen av Lagranges multiplikator kan i vissa ekonomiska sammanhang vara viktig. Vid en stationär punkt tolkas optimalt λ som marginalförändringen av optimalt z när c ökar med en enhet. Till exempel, om vår målfunktion utgör en nyttofunktion för en individ och vårt bivillkor är ett budget villkor som innebär att man inte kan konsumerar för mer än sin inkomst, och där inkomsten kommer att representeras av c. Lagranges multiplikator kommer då att vara ett mått på hur mycket nyttan förändras om inkomsten ökar med 1 enhet.

Lagranges funktion optimeras på vanligt sätt genom att bestämma partialderivatan till samtliga variabler. Detta skulle kunna se ut på följande sätt:

$$\frac{\partial L}{\partial x} = \frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \tag{3}$$

$$\frac{\partial L}{\partial y} = \frac{\partial f}{\partial y} - \lambda \frac{\partial g}{\partial y} = 0 \tag{4}$$

$$\frac{\partial L}{\partial \lambda} = c - g(x, y) = 0 \tag{5}$$

Här har vi ett simultant ekvationssystem som man måste lösa med avseende på x, y och λ. Det vanligaste förfaringssättet är dock att ur ekvationerna (3) och (4) härleda en relation mellan x och y som sedan substitueras in i (5) som då blir en ekvation med en okänd variabel som lätt kan bestämmas. Lösningen till ekvationssystemet kommer att utgöras av de eftersökta stationära punkterna (eller punkten).

Exempel 1
Finn stationära punkter till följande problem

$$\begin{aligned} &z = 4xy \\ &u.b. \quad x^2 + y^2 = 1 \end{aligned}$$

Steg 1: Formulera Lagranges funktion:

$$L = 4xy + \lambda(1 - x^2 - y^2)$$

Steg 2: Bestäm stationära punkter:
Detta gör vi genom att bestämma första ordningens villkor för stationära punkter, dvs derivera partiellt med avseende på de inblandade variablerna:

$$\frac{\partial L}{\partial x} = 4y - 2\lambda x = 0 \qquad \Rightarrow \quad \lambda = 2\frac{y}{x} \tag{6}$$

$$\frac{\partial L}{\partial y} = 4x - 2\lambda y = 0 \qquad \Rightarrow \quad \lambda = 2\frac{x}{y} \tag{7}$$

$$\frac{\partial L}{\partial \lambda} = 1 - x^2 - y^2 = 0 \tag{8}$$

(6) och (7) kan vi sammanföra. Detta ger då $\frac{y}{x} = \frac{x}{y} \quad \Leftrightarrow \quad x^2 = y^2$ som utgör vår relation mellan *x* och *y*. Denna relation kan vi substituera in i bivillkoret (8). Detta ger

då att $2x^2 = 1 \quad \Leftrightarrow \quad x^2 = \frac{1}{2} \quad \Leftrightarrow \quad x = \pm\frac{1}{\sqrt{2}}$ som också innebär att $y = \pm\frac{1}{\sqrt{2}}$. Med hjälp av dessa värden kan vi bestämma motsvarande λ med hjälp av (6) eller (7).

Detta ger oss följande stationära punkter:

$$(x_0, y_0, \lambda) = \left(1/\sqrt{2}, 1/\sqrt{2}, 2\right), \left(-1/\sqrt{2}, 1/\sqrt{2}, -2\right), \left(1/\sqrt{2}, -1/\sqrt{2}, -2\right), \left(-1/\sqrt{2}, -1/\sqrt{2}, 2\right)$$

Steg 3 innebär att vi måste bestämma vilka punkter som är lokala max respektive min punkter. För denna uppgift behöver vi en besluts regel som tyvärr avviker något från den regel som gäller vid multivariat optimering utan bivillkor. Beslutsregeln ser ut på följande sätt:

Utgå ifrån följande Lagrangefunktion:

$$L(x, y, \lambda) = f(x, y) + \lambda(c - g(x, y))$$

Bestäm partialderivatan med avseende på x och y:

$$\frac{\partial L}{\partial x} = \frac{\partial f(x, y)}{\partial x} - \lambda \frac{\partial g(x, y)}{\partial x} =$$

$$\frac{\partial L}{\partial y} = \frac{\partial f(x, y)}{\partial y} - \lambda \frac{\partial g(x, y)}{\partial y}$$

Bestäm nu andra derivatan och korsderivatan till funktionerna ovan:

$$\frac{\partial^2 L}{\partial x^2} = \frac{\partial^2 f(x, y)}{\partial x^2} - \lambda \frac{\partial^2 g(x, y)}{\partial x^2}$$

$$\frac{\partial^2 L}{\partial y^2} = \frac{\partial^2 f(x, y)}{\partial y^2} - \lambda \frac{\partial^2 g(x, y)}{\partial y^2}$$

$$\frac{\partial^2 L}{\partial x \partial y} = \frac{\partial^2 f(x, y)}{\partial x \partial y} - \lambda \frac{\partial^2 g(x, y)}{\partial x \partial y}$$

$$\frac{\partial^2 L}{\partial y \partial x} = \frac{\partial^2 f(x, y)}{\partial y \partial x} - \lambda \frac{\partial^2 g(x, y)}{\partial y \partial x}$$

Slutligen behöver vi bestämma första derivatan till bivillkoret

$$\frac{\partial g(x, y)}{\partial x} \quad \text{och} \quad \frac{\partial g(x, y)}{\partial y}$$

Med hjälp av deriveringarna ovan kan vi formulera en Hessian som gäller vid optimering under bivillkor. Denna hessian kallas för ***The Bordered Hessian*** och ser ut på följande sätt:

$$\overline{H} = \begin{bmatrix} 0 & \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial^2 L}{\partial x^2} & \dfrac{\partial^2 L}{\partial x \partial y} \\ \dfrac{\partial g}{\partial y} & \dfrac{\partial^2 L}{\partial y \partial x} & \dfrac{\partial^2 L}{\partial y^2} \end{bmatrix}, \quad \Rightarrow \quad D(x,y) = -\begin{vmatrix} 0 & \dfrac{\partial g}{\partial x} & \dfrac{\partial g}{\partial y} \\ \dfrac{\partial g}{\partial x} & \dfrac{\partial^2 L}{\partial x^2} & \dfrac{\partial^2 L}{\partial x \partial y} \\ \dfrac{\partial g}{\partial y} & \dfrac{\partial^2 L}{\partial y \partial x} & \dfrac{\partial^2 L}{\partial y^2} \end{vmatrix}$$

Med hjälp av bordered Hessian ($\overline{H}$), kan vi formulera uttrycket för en speciell determinant ($D(x,y)$). Det är denna determinant av "bordered Hessian" som leder oss till svaret om vår stationära punkt är ett lokalt maximum eller ett lokalt minimum. Observera minustecknet framför determinanten. Syftet med detta tecken är att vår beslutsregel skall likna beslutsregeln som gäller vid optimering utan bivillkor. Determinanten av "bordered Hessianen" ovan ser ut på följande sätt:

$$D(x,y) = -\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 L}{\partial x \partial y} - \frac{\partial g}{\partial y}\frac{\partial g}{\partial x}\frac{\partial^2 L}{\partial y \partial x} + \frac{\partial g}{\partial y}\frac{\partial g}{\partial y}\frac{\partial^2 L}{\partial x^2} + \frac{\partial g}{\partial x}\frac{\partial g}{\partial x}\frac{\partial^2 L}{\partial y^2}$$

som är lika med

$$D(x,y) = \left(\frac{\partial g}{\partial x}\right)^2 \frac{\partial^2 L}{\partial y^2} + \left(\frac{\partial g}{\partial y}\right)^2 \frac{\partial^2 L}{\partial x^2} - 2\frac{\partial g}{\partial x}\frac{\partial g}{\partial y}\frac{\partial^2 L}{\partial x \partial y}$$

Vår beslutsregel är baserad på vilket tecken denna determinant tar. Vi har följande 3 fall:

BESLUTSREGEL

Om (x_0, y_0) är en stationär punkt och

1) $D(x_0, y_0) < 0$	då är (x_0, y_0) ett lokalt maximum
2) $D(x_0, y_0) > 0$	då är (x_0, y_0) ett lokalt minimum
3) $D(x_0, y_0) = 0$	inget beslut

Låt oss nu återgå till vårt exempel och klassificera de stationära punkterna som vi fick. Vi börjar därför med att bestämma ingredienserna till ”bordered Hessian”.

$$\frac{\partial L}{\partial x^2} = -2\lambda \quad \frac{\partial L}{\partial y^2} = -2\lambda \quad \frac{\partial g}{\partial x} = 2x$$
$$\frac{\partial L}{\partial x \partial y} = 4 \quad \frac{\partial L}{\partial y \partial x} = 4 \quad \frac{\partial g}{\partial y} = 2y$$

Dessa komponenter ger en ”bordered Hessian” uttryckt på determinantformat med följande utseende:

$$D(x,y) = -\begin{vmatrix} 0 & 2x & 2y \\ 2x & -2\lambda & 4 \\ 2y & 4 & -2\lambda \end{vmatrix}$$

Om vi bestämmer determinanten till denna matris så får vi följande ekvation:

$$D(x,y) = -\left(16xy + 16xy + 8y^2\lambda + 8x^2\lambda\right) = -8\lambda(x^2 + y^2) - 32xy$$

Med hjälp av detta uttryck tar vi följande beslut:

$(x_0, y_0, \lambda_0) = (1/\sqrt{2}, 1/\sqrt{2}, 2) \quad \Rightarrow \quad D(1/\sqrt{2}, 1/\sqrt{2}, 2) = -32 < 0$ Lokalt maximum
$z_0 = 2$

$(x_0, y_0, \lambda_0) = (-1/\sqrt{2}, -1/\sqrt{2}, 2) \Rightarrow D(-1/\sqrt{2}, -1/\sqrt{2}, 2) = -32 < 0$ Lokalt maximum
$z_0 = 2$

$(x_0, y_0, \lambda_0) = (-1/\sqrt{2}, 1/\sqrt{2}, -2) \quad \Rightarrow \quad D(-1/\sqrt{2}, 1/\sqrt{2}, -2) = 32 > 0$ Lokalt minimum
$z_0 = -2$

$(x_0, y_0, \lambda_0) = (1/\sqrt{2}, -1/\sqrt{2}, -2) \quad \Rightarrow \quad D(1/\sqrt{2}, -1/\sqrt{2}, -2) = 32 > 0$ Lokalt minimum
$z_0 = -2$