Optimering med bivillkor

9.1. Optimering med bivillkor

Låt $f(\boldsymbol{x})$ vara en funktion av $\boldsymbol{x} \in \mathbb{R}$. Vi vill optimera funktionen f under *bivillkoret* $g(\boldsymbol{x}) = C$ (eller bivllkoren $g_1(\boldsymbol{x}) = C_1, \ldots, g_k(\boldsymbol{x}) = C_k$). Även olikheter $h_1(\boldsymbol{x}) \leq A_1, \ldots, h_m(\boldsymbol{x}) \leq A_m$ kan förekomma

Terminologi:

- f kallas en **målfunktion**;
- ekvationen $g_k(\boldsymbol{x}) = C_k$ kallas ett **bivillkor**.

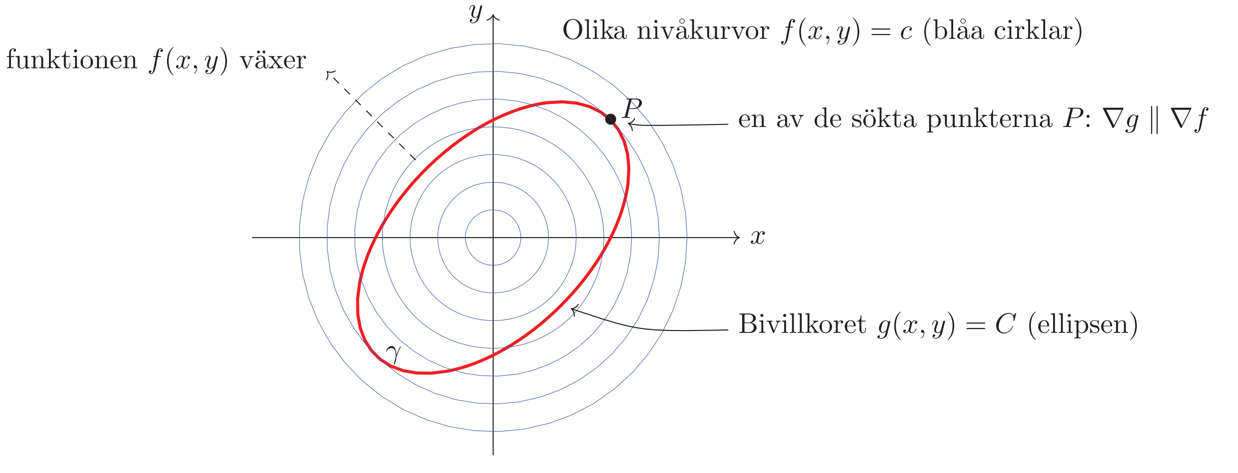
Det påminner randundersökningen i optimering på kompakta mängder men vi vill studera hur man kan lösa problem utan att parametrisera randen.

Exempel 9.1. Bestäm de punkter på ellipsen $x^2 - xy + y^2 = 4$ som ligger på den strsta avståndet från origo.

Lösning. Här

$$f(x, y) = \text{avståndet mellan } (x, y) \text{ och origo } = \sqrt{x^2 + y^2}$$

och bivillkoret ges av $g(x, y) := x^2 - xy + y^2 = 4$. Eftersom mängden $\gamma = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 1\}$ är kompakt (*varför?*), konstaterar vi att den kontinuerliga funktionen f antar sitt största värde på γ.



Alltså: $\nabla f \parallel \nabla g$ vilket är ekvivalent till att determinanten

$$\begin{vmatrix} f'_x & f'_y \\ g'_x & g'_y \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ 2x-y & 2y-x \end{vmatrix} = 0 \quad \Leftrightarrow \quad x(2y-x)-y(2x-y)=0 \quad \Leftrightarrow \quad (x-y)(x+y)=0$$

Alltså ska 'intressanta' punkter på uppfylla $y = \pm x$ tillsammans med bivillkoret $x^2 - xy + y^2 = 4$ vilket ger

$$\begin{cases} y &= x \\ x^2 &= 4 \end{cases} \qquad \text{eller} \qquad \begin{cases} y &= -x \\ 3x^2 &= 4 \end{cases}$$

Det första systemet ger två punkter $(2,2)$ och $(-2,2)$ med resp. funktionsvärdena $f(2,2) = f(-2,-2) = 2\sqrt{2}$, och det andra systemet ger $(\frac{2}{\sqrt{3}}, -\frac{2}{\sqrt{3}})$ och $(-\frac{2}{\sqrt{3}}, \frac{2}{\sqrt{3}})$ med resp. funktionsvärdena $f(\mp\frac{2}{\sqrt{3}}, \pm\frac{2}{\sqrt{3}}) = \frac{2\sqrt{2}}{\sqrt{3}}$.

*) Ett undantagsfall: $\nabla g(x,y) = \mathbf{0}$ ger $2x - y = 2y - x = 0$, alltså $(x,y) = (0,0)$ som *inte* ligger på γ.

Sammanfattningsvis är $2\sqrt{2}$ det strsta avståndet.

Allmänt gäller följande satsen. Vi anatar att både f och g är av klass $\mathscr{C}^1$.

Sats 14. *Antag att punkten (a,b) är en lokal extrempunkt till funktionen $f(x,y)$ under bivillkoret $g(x,y) = C$. Antag vidare att (a,b) är en inre punkt till definitionsmängder D_f och D_g. Då gäller det att $\nabla f(a,b) \parallel \nabla g(a,b)$. Allmänt, antag att $\boldsymbol{a} \in \mathbb{R}^n$ är en lokal extrempunkt till funktionen $f(\boldsymbol{x})$ under bivillkoren $g_1(\boldsymbol{x}) = C_1, \ldots, g_k(\boldsymbol{x}) = C_k$ och antag att $\boldsymbol{a}$ är en inre punkt till $D_f \cap D_{g_1} \cap \ldots \cap D_{g_k}$. Då gäller det att $\nabla f(\boldsymbol{a})$, $\nabla g_1(\boldsymbol{a})$, $\ldots$, $\nabla g_k(\boldsymbol{a})$ är linjärt beroende.*

Bevis för $n = 2$ och $k = 1$. Fallet $\nabla g(a,b) = \mathbf{0}$ är triviallt. Antag att $\nabla g(a,b) \neq \mathbf{0}$. Enligt implicita funktionenssatsen kan vi parametrisera $\gamma = \{(x,y) \in \mathbb{R}^2 : g(x,y) = C\}$ i en omgivning av (a,b): $\boldsymbol{x}(t) = (x(t), y(t))$ så att $\boldsymbol{x}(0) = (a,b)$ och $g(\boldsymbol{x}(t)) = C$. Den nya funktionen av en variabel $h(t) = f(\boldsymbol{x}(t))$ har ett (lokalt) extremvärde i $t = 0$, så det gäller att $h'(0) = 0$. Enligt kedjeregeln:

$$0 = h'(0) = \frac{\partial f}{\partial x}(a,b) \cdot \frac{dx}{dt}(0) + \frac{\partial f}{\partial y}(a,b) \cdot \frac{dy}{dt}(0) \quad \Leftrightarrow \quad \nabla f(a,b) \cdot \dot{\boldsymbol{x}}(0) = 0 \quad \Leftrightarrow \quad \nabla f(a,b) \perp \dot{\boldsymbol{x}}(0),$$

d.v.s. $\nabla f(a,b)$ och tangentvektorn till nivåkurvan γ är ortogonala. Det betyder att $\nabla f(a,b)$ är en normalvektor till γ, alltså den är parallell med gradienten $\nabla g(a,b)$, v.s.b. □

Algoritmen:

(i) Visa att ett optimalt värde existerar (ofta *ad hoc*)

(ii) Bestäm 'kandidater' genom att lösa sambandet $\nabla f \parallel \nabla g$ *tillsammans* med bivillkoret $g(x,y) = C$.

(iii) Optimera: beräkna funktionesvrdena i kandidaterna och bestäm det extremala värdet .

Anmärkning. I (ii) kan man bestämma intressanta punkter genom at skriva

$$\nabla f \parallel \nabla g \quad \Leftrightarrow \quad \nabla f = \lambda \nabla g, \quad \lambda \in \mathbb{R}. \tag{iia}$$

λ kallas för **Lagranges multiplikatorn**. Alternativt kan man använda **determinanten**

$$\begin{vmatrix} f'_x & f'_y \\ g'_x & g'_y \end{vmatrix} = 0 \tag{iib}$$

Alternativet (iib) förefaller ofta enklare än det första (Lagranges) alternativet (iia), men (iib) fungerar bara för två variabler.

Exempel 9.2. Bestäm det största värde av $f(x, y, z) = xyz$ på enhetsfären i $\mathbb{R}^3$.

Lösning. Vi använder Lagranges multiplikatormetod under bivillkoret $g := x^2 + y^2 + z^2 = 1$, alltså

$$\begin{cases} \nabla f = \lambda \nabla g \\ g = C \end{cases} \quad \Leftrightarrow \quad \begin{cases} yz = 2\lambda x \\ xz = 2\lambda y \\ xy = 2\lambda z \\ x^2 + y^2 + z^2 = 1 \end{cases}$$

Om $\lambda = 0$ så så får vi att det finns exakt två koordinate är lika med noll (varför?). Det ger 6 punkter $(\pm 1, 0, 0)$, $(0, \pm 1, 0)$ och $(0, 0, \pm 1)$ för vilka $f = 0$. Antag nu att $\lambda \neq 0$. Ett enkelt resonemang visar att $xyz \neq 0$ (ty, om t.ex. $x = 0$ så $yz = \lambda \cdot 0 = 0$, alltså $y = 0$ eller $z = 0$ vilket medför att $\lambda = 0$). Genom att dividera ekvationer får vi $x^2 = y^2 = z^2 = 1/3$. Detta ger oss 8 punkter $(\pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}}, \pm\frac{1}{\sqrt{3}})$ där alla teckenkombinationer kan förekomma. Funktionensvärde är antigen $f = \frac{1}{3\sqrt{3}}$ (sökta största värde) eller $f = -\frac{1}{3\sqrt{3}}$ (sökta minsta värde).

En **alternativ lösning via determinanter**. Notera att två vektorer i $\mathbb{R}^3$ är parallella om och endast om

$$\nabla f \parallel \nabla g \quad \Leftrightarrow \quad \nabla f \times \nabla g = 0 \quad \Leftrightarrow \quad \begin{vmatrix} f'_x & f'_y \\ g'_x & g'_y \end{vmatrix} = \begin{vmatrix} f'_x & f'_z \\ g'_x & g'_z \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} yz & xz \\ 2x & 2y \end{vmatrix} = \begin{vmatrix} yz & xy \\ 2x & 2z \end{vmatrix} = 0$$

vilket ger i vårt fall:

$$\begin{cases} z(y^2 - x^2) = 0 \\ y(z^2 - x^2) = 0 \\ x^2 + y^2 + z^2 = 1 \end{cases} \quad \Rightarrow$$

Vi har alltså två alternativen: antigen $x^2 = y^2 = z^2 = 1/3$ eller två koordinater är lika med noll, och sedan som ovan.

9.2. Optimering med två bivillkor

Vi skall nu studera problemet att optimera en funktion $f(x, y, z)$ under två bivillkor

$$g(x, y, z) = C_1, \qquad h(x, y, z) = C_2.$$

Enligt satsen, gradienterna är linjärt beroende, vilket ger systemet

$$\begin{cases} \begin{vmatrix} f'_x & f'_y & f'_z \\ g'_x & g'_y & g'_z \\ h'_x & h'_y & h'_z \end{vmatrix} = 0 \\ g = C_1 \\ h = C_2 \end{cases}$$

Exempel 9.3. Planet $x + y + z = 2$ skär cylindern $y^2 + z^2 = 4$ längs en ellips. Bestäm den punkt på ellipsen som ligger närmast origo.

Lösning. Vi använder Lagranges multiplikatorsmetod för $f(x, y, z) = x^2 + y^2 + z^2$ (avståndet i kvadrat) under bivillkoren $g := x + y + z = 2$ och $h := y^2 + z^2 = 4$, alltså

$$\begin{vmatrix} 2x & 2y & 2z \\ 1 & 1 & 1 \\ 0 & 2y & 2z \end{vmatrix} = 0 \quad \Leftrightarrow \quad \begin{vmatrix} x & 0 & 0 \\ 1 & 1 & 1 \\ 0 & y & z \end{vmatrix} = 0 \quad \Leftrightarrow \quad x \cdot \begin{vmatrix} 1 & 1 \\ y & z \end{vmatrix} = 0 \quad \Leftrightarrow \quad x(y - z) = 0$$

Tillsammans med bivillkoren ger det systemet

$$\begin{cases} x(y-z)=0 \\ x+y+z=2 \\ y^2+z^2=4 \end{cases}$$

a) I fallet $x=0$ får vi $y+z=2$ och $y^2+z^2=4$ vilket ger två punkter: $(0,2,0)$ och $(0,0,2)$.

b) I fallet $x \neq 0$ har vi systemet

$$\begin{cases} y=z \\ x+y+z=2 \\ y^2+z^2=4 \end{cases} \Leftrightarrow \begin{cases} y=z \\ x+2z=2 \\ z^2=2 \end{cases} \Leftrightarrow \begin{cases} y=z \\ x+2z=2 \\ z^2=2 \end{cases}$$

vilket ger $z=\pm\sqrt{2}$ och alltså ytterliggare två punkter:

$$(2-2\sqrt{2},\sqrt{2},\sqrt{2}) \quad \text{och} \quad (2+2\sqrt{2},-\sqrt{2},-\sqrt{2}).$$

Vi beräknar funktionsvärdena i dessa punkter:

$$f(0,2,0)=f(0,0,2)=4, \quad f(2-2\sqrt{2},\sqrt{2},\sqrt{2})=16-8\sqrt{2}, \quad f(2+2\sqrt{2},-\sqrt{2},-\sqrt{2})=16+8\sqrt{2},$$

vilket medför att minsta avståndet är $\sqrt{4}=2$ och det uppnås i punkterna $(0,2,0)$ och $(0,0,2)$.

9.3. Optimering med godtyckliga bivillkor

Se även ett exempel på http://courses.mai.liu.se/GU/TATA43/Dokument/opt-kompl.pdf (L. Alexandersson).

Exempel 9.4. Optimera $f=xy+2z$ då $g=x^2+y^2+z^2 \leq 6$ och $h=x+y+z \geq 0$.

Lösning. Mängden D som ges av bivillkoren $g \leq 6$ och $h \geq 0$ är sluten och begränsad, alltså kompakt. Funktionen f är kontinuerlig, så f antar sitt största och minsta värde på D. Vi har följande uppdelningen:

$$D = B \cup S_1 \cup S_2 \cup \gamma, \quad \text{där}$$

(i) B är ett halvklot $\{g<6, h>0\}$;

(ii) S_1 är ett halvsfär $\{g=6, h>0\}$;

(iii) S_2 är en cirkelskiva $\{g<6, h=0\}$;

(iv) γ är en cirkel $\{g=6, h=0\}$.

Nu optimerar vi på var och en av dessa. (i) Optimering med bara stationära punkter:

$$\begin{cases} \nabla f = 0 \\ g<6 \\ h>0 \end{cases} \Leftrightarrow \begin{cases} f'_x = y = 0 \\ f'_y = x = 0 \\ f'_z = 2 = 0 \\ g<6 \\ h>0 \end{cases}$$

lösningen saknas, alltså finns inga stationära punkter i B.

(ii) Optimering med ett bivillkor ($g = 6$). Kandidater fås ur

$$\begin{cases} \nabla f \times \nabla g = 0 \\ g = 6 \\ h > 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} y^2 = x^2 \\ yz = 2x \\ xz = 2y \\ x^2 + y^2 + z^2 = 6 \\ x + y + z > 0 \end{cases}$$

Om $x = 0$ så $y = 0$, alltså $z = +\sqrt{6}$ ($z = -\sqrt{6}$ upfyller inte olikheten $h > 0$) vilket ger

$$\boxed{f(0, 0, \sqrt{6}) = 2\sqrt{6}}$$

I fall om $x \neq 0$ får vi antigen $x = y \neq 0$ med $z = 2$, alltså $(1, 1, 2)$ (den andra lösningen $(-1, -1, 2)$ uppfyller inte olikheten $h > 0$) vilket ger ytterligare en kandidat

$$\boxed{f(1, 1, 2) = 5}$$

eller $x = -y$ och $z = -2$. Här lösningarna saknas (uppfyller inte $h > 0$).

(iii) Optimering med ett bivillkor ($h = 0$). Kandidater fås ur

$$\begin{cases} \nabla f \times \nabla h = 0 \\ g < 6 \\ h = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = 2 \\ y = 2 \\ x^2 + y^2 + z^2 < 6 \\ x + y + z = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = 2 \\ y = 2 \\ x^2 + y^2 + z^2 < 6 \\ z = -4 \end{cases}$$

Eftersom $2^2 + 2^2 + (-4)^2 = 24 \not< 6$, lsningen saknas.

(iv) Optimering med två bivillkor ($h = 0$, $g = 6$). Kandidater fås ur

$$\begin{cases} \det(\nabla f, \nabla g, \nabla h) = 0 \\ g = 6 \\ h = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} \begin{vmatrix} y & x & 2 \\ 2x & 2y & 2z \\ 1 & 1 & 1 \end{vmatrix} = 0 \\ x^2 + y^2 + z^2 = 6 \\ x + y + z = 0 \end{cases} \quad \Leftrightarrow \quad \begin{cases} (y - x)(x + y - 2 - z) = 0 \\ x^2 + y^2 + z^2 = 6 \\ x + y + z = 0 \end{cases}$$

- Om $y = x$ så $z = -2x$, alltså $(1, 1, -2)$ eller $(-1, -1, 2)$:

$$\boxed{f(1, 1, -2) = -3, \qquad f(-1, -1, 2) = 5}$$

- Om $x + y = 2 + z$ så får vi från den tredje ekvationen att $z = -1$, alltså $(2, -1, -1)$ och $(-1, 2, -1)$:

$$\boxed{f(2, -1, -1) = f(-1, 2, -1) = -4}$$

Eftersom

$$-4 < -3 < 2\sqrt{6} = \sqrt{24} < 5 = \sqrt{25}$$

får vi att det största och minsta värdena är resp. 5 och -4.