Diskreta kaj reela optimumado

Prof. Christer Kiselman

1. Enkonduko

Optimumado signifas 'serĉadon de la plej bona'. (Latine *optimus* 'la plej bona'.) *Sed kio estas plej bona?* Oni devas unue decidi pri tio. Kio estas optimuma kontribuo al Universala Esperanto-Asocio? (Ĉu la plej alta aŭ la plej malalta?) Kia estas optimuma organizado de riparo de strato? (Ĉu tiu kiu ĝenas la trafikon malplej, aŭ tiu kiu ĝenas plej?)

Pli bona el kiuj? Tiu ĉi estas la dua demando. Kiuj estas la eblaj valoroj aŭ statoj kiuj rajtas aŭ povas ekzisti? Ekzemple oni kutime ne povas aĉeti duonan pomon, aŭton aŭ domon: oni devas aĉeti neniun aŭ unu aŭ du ...; nur entjeraj valoroj eblas. Male, optimuma longo povas esti 1,3 metroj aŭ $\pi \approx 3{,}14159265$ metroj; en tia situacio ĉiu reela nombro eblas.

Ni prezentos kelkajn gravajn diferencojn inter entjera kaj reela optimumado, kaj montros ke oni povas difini taŭgajn klasojn de funkcioj kun du entjeraj variabloj por kiuj validas dezirindaj rezultoj.

La reĝo Salomono komprenis la diferencon inter reela kaj diskreta optimumado antaŭ pli ol 2 900 jaroj:

La verdikto de la reĝo Salomono: "Kaj la reĝo diris: Dishaku la vivantan infanon en du partojn, kaj donu duonon al unu kaj duonon al la alia." (Biblio, Reĝoj, Libro Unua 3:25; bildo 1.)

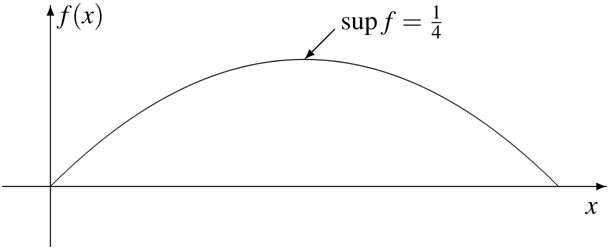
Li tiel decidis kiam du virinoj disputis pri infano: ambaŭ asertis esti ties patrino. Kaj la verdikto maksimumigis la feliĉon de ambaŭ — almenaŭ oni povas facile difini mezuron de feliĉo tia ke ĝia maksimumo okazas kiam la infano estas duonigita (bildo 2). Sed ĉu estas permesite dishaki infanon?

Pierre de Fermat (naskita eble en 1601, pli kredeble en 1607 aŭ 1608; morta en 1665) priskribis metodon por trovi optimuman valoron: *Trovu la nulejojn de la derivaĵo!*

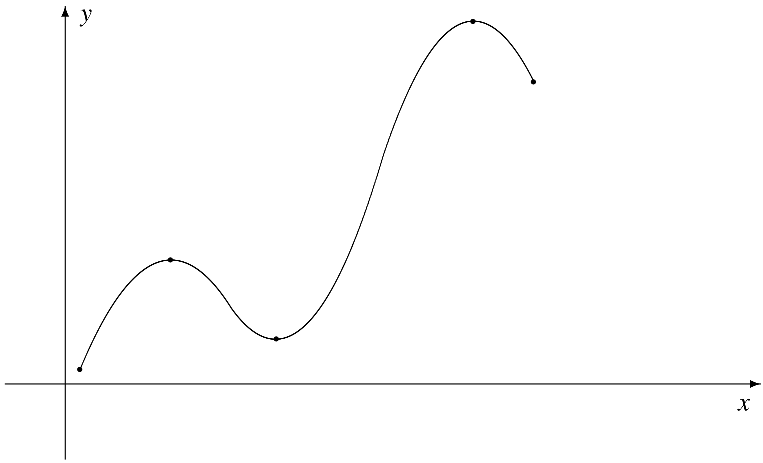
La regulo de Fermat: *Se funkcio f havas lokan maksimumon aŭ lokan minimumon en interna punkto a de intervalo kie ĝi estas difinita, kaj se f estas derivebla en a, tiam ĝia derivaĵo f' egalas al nulo en a (bildo 3).*

Eble Fermat ne bone komprenis la nocion de derivaĵo, sed li tamen komprenis kion signifas la fakto ke la derivaĵo nulas!

Bildo 1. La verdikto de la reĝo Salomon. Gravuraĵo fare de Gustave Doré (1832–1883).

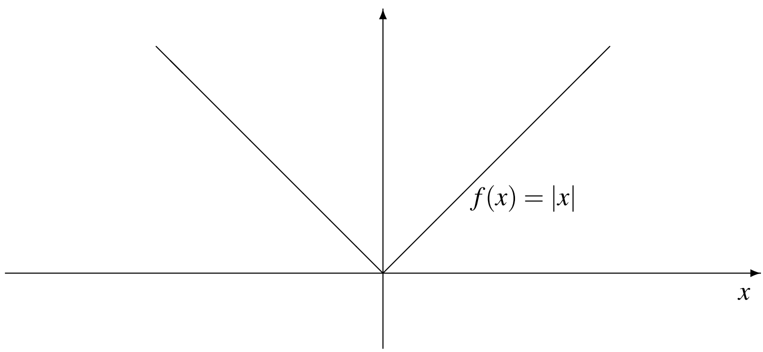


Bildo 2. La objekta funkcio $f(x) = x(1-x)$, kies maksimumo estas atingita ĉe $x = \frac{1}{2}$.



Bildo 3. La regulo de Fermat. La derivaĵo nulas ĉe tri punktoj. Ekzistas du finpunktoj. La serĉado estas limigita al kvin punktoj.

Sed kion fari se la derivaĵo ne ekzistas? Ekzemple se oni volas trovi la plej malgrandan valoron de la funkcio kiu por la nombro x donas ĝian absolutan valoron, do $x \mapsto |x|$? Vidu bildon 4.



Bildo 4. La neglata funkcio $f(x) = |x|$, kiu havas minimumon ĉe $x = 0$.

Ĝi ja havas la optimumon egalan al la minimumo en la origino, sed tie la derivaĵo ne ekzistas. Tiaj demandoj kondukis sciencistojn al nova esplorbranĉo: *neglata analitiko*. Mi menciu du esploristojn kiuj estas konataj pro ilia laboro en tiu ĉi fako: Francis Clarke kaj Boris Mordukovich.

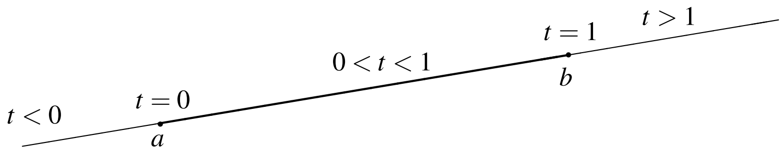
2. Konveksaj aroj kaj funkcioj kun reelaj variabloj

Kio estas konveksa aro? Por difini konveksan aron ni unue bezonas la nocion de streko. La ***streko*** inter du punktoj a kaj b estas la aro

$$[a,b] = \{(1-t)a + tb;\ t \in \mathbf{R},\ 0 \leqslant t \leqslant 1\}.$$

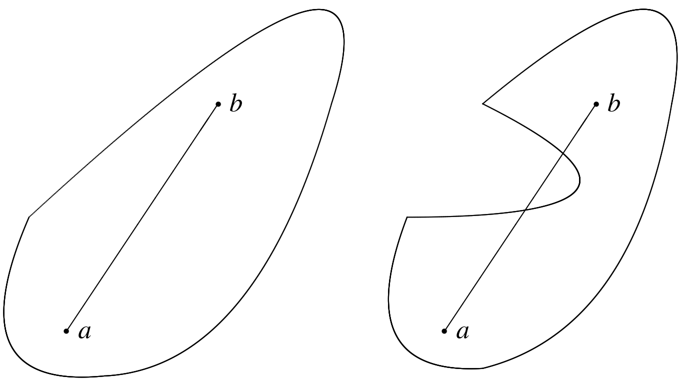
Ĉi tie $\mathbf{R}$ estas la aro de ĉiuj reelaj nombroj.

Ni havas $[a,a] = \{a\}$ kaj $[a,b] = [b,a]$; la streko ne havas direkton.



Bildo 5. La streko $[a,b]$ inter du punktoj a kaj b.

Difino 2.1. Aro K en $\mathbf{R}^n$ estas nomata ***konveksa*** se por ĉiuj punktoj $a,b \in \mathbf{R}^n$ validas $\{a,b\} \subset K \;\Rightarrow\; [a,b] \subset K$. □



Bildo 6. Konveksa aro; nekonveksa aro.

Difino 2.2. Estu A ajna aro en la spaco $\mathbf{R}^n$. Ĝia ***konveksa tegaĵo***, $\mathbf{teg}(A)$, estas la plej malgranda konveksa aro entenanta la aron A.

Difino 2.3. Funkcio $F\colon \mathbf{R}^n \to [-\infty, +\infty]$ estas nomata ***konveksa*** se la aro

$$\{(x,t) \in \mathbf{R}^n \times \mathbf{R};\ F(x) \leqslant t\} \subset \mathbf{R}^{n+1}$$

estas konveksa. □

Funkcio $F\colon \mathbf{R}^2 \to [-\infty, +\infty]$ estas konveksa se kaj nur se, por ĉiuj $a, b \in \mathbf{R}^2$, la funkcio kun unu variablo

$$\mathbf{R} \ni t \mapsto F(a + tb)$$

estas konveksa.

Difino 2.4. Estu f ajna funkcio difinita en la spaco $\mathbf{R}^n$ aŭ nur en subaro de ĝi. Ĝia ***konveksa envelopo***, $\mathbf{env}(f)$, estas la plej granda konveksa malpliedanto de f. □

Plianto de donita nombro estas nombro kiu estas pli granda ol la donita nombro; pliedanto estas nombro kiu ne malsuperas la donitan nombron. Simile oni parolas pri pliedanto kaj malpliedanto de funkcio.[1]

Oni do difinas konveksajn funkciojn per konveksaj aroj. Tiu estas la plej facila vojo. La aro

$$\mathbf{epi}^{\mathrm{F}}(F) = \{(x,t) \in \mathbf{R}^n \times \mathbf{R};\ F(x) \leqslant t\} \subset \mathbf{R}^{n+1}$$

nomiĝas la ***finia epigrafeo*** de la funkcio. La aro

$$\mathbf{epi}^{\mathrm{F}}_{\mathrm{s}}(F) = \{(x,t) \in \mathbf{R}^n \times \mathbf{R};\ F(x) < t\} \subset \mathbf{epi}^{\mathrm{F}}(F)$$

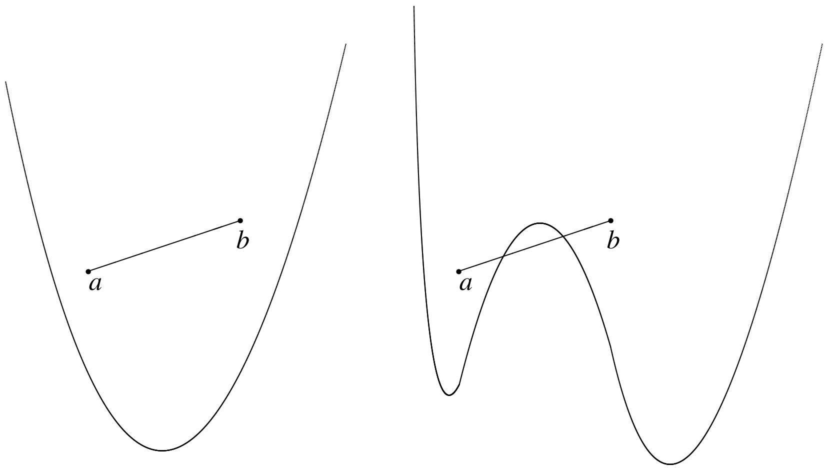
nomiĝas la ***strikta finia epigrafeo*** de la funkcio.

Validas la ekvivalentoj

$$F \text{ estas konveksa } \Leftrightarrow \mathbf{epi}^{\mathrm{F}}(F) \text{ estas konveksa } \Leftrightarrow \mathbf{epi}^{\mathrm{F}}_{\mathrm{s}}(F) \text{ estas konveksa,}$$

kie la unua ekvivalento estas difino, kaj la dua facile pruvebla teoremo.

[1] Ĉi tie *ed* estas malfortiga sufikso, kiu signalas ke egaleco estas permesata, do *pozitiveda* egalas al 'pozitiva aŭ nula', alivorte al 'nenegativa'. Simile pri *pliedanto*.



Bildo 7. Konveksa epigrafeo; nekonveksa epigrafeo.

Se funkcio kies minimumon oni serĉas estas konveksa, tio estas granda helpo: loka minimumo estas ĉiam tutspaca minimumo. Konveksaj funkcioj do estas speciale agrablaj kiam temas pri serĉado de minimumo.

Alia bona eco estas ke la ombro de konveksa korpo estas konveksa.

3. Konveksaj funkcioj kun entjeraj variabloj

Se ni havas ne reelajn variablojn sed ekzemple entjerajn variablojn, tiam tute ne estas klare kiel la bonaj ecoj menciitaj en la antaŭa sekcio povas travivi.

Ĉi tie mi provos montri la malfacilaĵojn kiuj estiĝas kiam oni klopodas optimumigi per entjeraj variabloj $\ldots,-2,-1,0,1,2,3,\ldots$.

3.1. Loka minimumo ne nepre estas tutspaca

La jam menciita eco de konveksaj funkcioj kun reelaj variabloj, nome ke loka minimumo estas aŭtomate tutspaca, estas tre utila en la reela opimumado. Alivorte, se F estas konveksa en la reela ebeno, tiam por scii ke $F(x,y) \geqslant F(a,b)$ por ĉiuj (x,y) en la tuta ebeno sufiĉas kontroli ke $F(x,y) \geqslant F(a,b)$ por ĉiuj punktoj (x,y) sufiĉe proksimaj al (a,b).

Sed rigardu la funkcion

$$f(x,y) = x + |x-2my|, \qquad (x,y) \in \mathbf{Z}^2,$$

kie m estas pozitiva entjero, kiom ajn granda. Ĉi tie $\mathbf{Z}$ estas la aro de ĉiuj entjeraj nombroj, kaj $\mathbf{Z}^2$ la aro de ĉiuj paroj de entjeroj.

Tiam $f(x,y) \geqslant f(0,0) = 0$ kiam $x \geqslant -m$, do kiam (x,y) apartenas al granda aro kiu enhavas la originon, sed

$$f(-m-1,-1) = -2 < f(0,0).$$

Tamen f estas la malvastigaĵo al $\mathbf{Z}^2$ de konveksa funkcio sur $\mathbf{R}^2$:

$$F(x,y) = x + |x - 2my|, \qquad (x,y) \in \mathbf{R}^2.$$

Kion fari? Ekzistas tri ebloj.
• Ni povas rifuzi nomi f konveksa.
• Ni povas klopodi uzi grandegajn ĉirkaŭaĵojn.
• Ni simple rezignu la rezulton.
Temas pri esplorado! Ni devas elekti.
La solvo kiun mi proponos estas ke ni fortigu la kondiĉojn sur f.

3.2. Sovaĝaj marĝenaj funkcioj

Estas geometrie evidente ke la ombro ĵetita sur $\mathbf{R}^2$ de konveksa aro en $\mathbf{R}^3$ estas konveksa. Oni kutime esprimas tion parolante pri marĝenaj funkcioj.

Difino 3.1. Se estas donita funkcio $F\colon \mathbf{R}^2 \to [-\infty, +\infty]$, tiam ĝia ***marĝena funkcio*** estas la funkcio

$$H(x) = \inf_{y \in \mathbf{R}} F(x,y), \qquad x \in \mathbf{R}. \qquad \square$$

Oni vidas ke la strikta epigrafeo de H tutsimple estas la projekciaĵo (la ombro) de la strikta epigrafeo de F per la projekcio $(x,y,t) \mapsto (x,t)$, la projekcio kiu forgesas la variablon y. Ni ja havas $H(x) < t$ se kaj nur se ekzistas y tia ke $F(x,y) < t$.

Sed rigardu la funkcion

$$f(x,y) = |x - 2my|, \qquad (x,y) \in \mathbf{Z}^2,$$

kie m estas pozitiva entjero. Ĝia marĝena funkcio estas $h(x) = |x|$ por $-m \leqslant x \leqslant m$, kaj estas donita por aliaj valoroj de x per tio ke ĝi estas perioda kun periodo $2m$. Do h estas funkcio kun segildentoj kiom ajn grandaj.

Tamen estas tente nomi f konveksa: ĝi ja estas la malvastigaĵo al $\mathbf{Z}^2$ de konveksa funkcio difinita sur $\mathbf{R}^2$:

$$F(x,y) = |x - 2my|, \qquad (x,y) \in \mathbf{R}^2.$$

Ankaŭ ĉi tie ni devas elekti. Ĉu ni volas malaprobi la funkcion f kiel konveksan funkcion? Ĉu ni volas nomi la marĝenan funkcion konveksa? Ĉu ni volas rezigni la rezulton?

4. Difinoj por funkcioj kun entjeraj variabloj

Ni rigardu unue funkciojn kun unu entjera variablo.

Difino 4.1. Funkcio $f\colon \mathbf{Z} \to \mathbf{R}$ nomiĝas ***konvekse vastigebla***, se ekzistas konveksa funkcio $F\colon \mathbf{R} \to \mathbf{R}$ tia ke $F(x) = f(x)$ por ĉiu $x \in \mathbf{Z}$. □

Oni povas facile montri ke funkcio $f\colon \mathbf{Z} \to \mathbf{R}$ estas konveksa se kaj nur se

$$f(x) \leqslant \tfrac{1}{2}f(x-1) + \tfrac{1}{2}f(x+1), \qquad x \in \mathbf{Z}.$$

Tio signifas ke la valoro de la funkcio en iu punkto nenie superas la mezan valoron de la valoroj en la du najbaraj punktoj.

Se oni donas la pezon (gravon) $+1$ al la punktoj $x-1$ kaj $x+1$ kaj la pezon -2 al la punkto x, tiam la sumo de la valoroj estas pozitiveda:[2]

$$(+1)f(x-1) + (-2)f(x) + (+1)f(x+1) = f(x-1) - 2f(x) + f(x+1) \geqslant 0.$$

Oni povas skribi tion:

$$f(a) - f(b) - f(c) + f(d) \geqslant 0,$$

kie a estas ajna punkto, $b = c = a+1$ kaj $d = a+2$.

Ni nun konsideru funkciojn kun du entjeraj variabloj. Ni povas fari similajn difinojn.

Difino 4.2. Funkcio $f\colon \mathbf{Z}^2 \to \mathbf{R}$ kun du entjeraj variabloj nomiĝas ***konvekse vastigebla***, se ekzistas konveksa funkcio $F\colon \mathbf{R}^2 \to \mathbf{R}$ kun du reelaj variabloj tia ke $F(x) = f(x)$ por ĉiu $x \in \mathbf{Z}^2$. (F estas konveksa vastigaĵo de f al la tuta reela ebeno $\mathbf{R}^2$; f estas malvastigaĵo de F al $\mathbf{Z}^2$.) □

Difino 4.3. Funkcio $f\colon \mathbf{Z}^2 \to \mathbf{R}$ kun du entjeraj variabloj nomiĝas ***romboide konveksa***, se

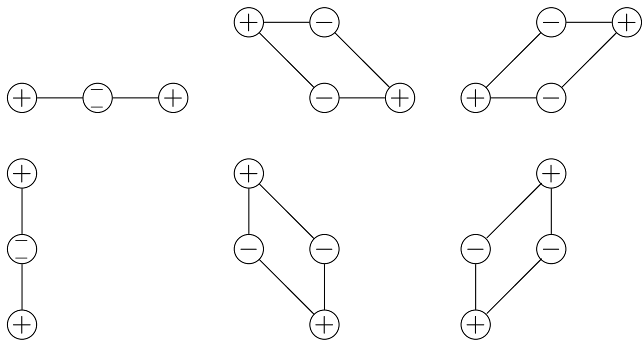
$$f(a) - f(b) - f(c) + f(d) \geqslant 0$$

kiam a estas ajna punkto kaj b, c, d estas la jenaj:

[2] Pri la sufikso *ed* vidu piednoton 1 sur paĝo 76.

(1) $b = c = a + (1,0)$, $d = a + (2,0)$;
(2) $b = a + (1,0)$, $c = a + (1,-1)$, $d = a + (2,-1)$;
(3) $b = a + (1,0)$, $c = a + (1,1)$, $d = a + (2,1)$;
(4) $b = c = a + (0,1)$, $d = a + (0,2)$;
(5) $b = a + (0,1)$, $c = a + (-1,1)$, $d = a + (-1,2)$.
(6) $b = a + (0,1)$, $c = a + (1,1)$, $d = a + (1,2)$; □

La kondiĉoj similas al tiuj por unu variablo, sed temas nun pri ses kondiĉoj por ĉiu punkto a. La kondiĉojn (1) kaj (4) ni nomas unu-variablaj; la ceterajn romboidaj: (2) kaj (3) horizontalaj kaj (5) kaj (6) vertikalaj.



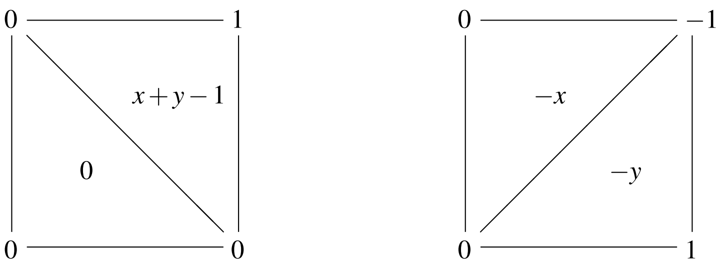
Bildo 8. Maldekstre la du unu-variablaj kondiĉoj; dekstre la kvar romboidaj kondiĉoj.

Ekzemplo 4.4. Rigardu la funkcion kiu egalas al xy en $\{0,1\}^2$. Kiel vastigi ĝin al konveksa funkcio difinita en la tuta kvadrato $[0,1]^2$? Evidente $F(x,y) = xy$ ne taŭgas, ĉar ĝi ne estas konveksa. Sed ni povas preni $F(x,y) = (x+y-1)^+ = \max(x+y-1,0)$.

Kion pri $g(x,y) = -xy$? Ni ne povas uzi $G(x,y) = -xy$. Sed $G(x,y) = \max(-x,-y)$ bone taŭgas. Vidu bildon 9. □

Difino 4.5. Al donita funkcio $f\colon \mathbf{Z}^2 \to \mathbf{R}$ ni difinas ĝian ***kanona***n ***vastigaĵo***n, notitan $\mathbf{kan}(f)\colon \mathbf{R}^2 \to \mathbf{R}$, tiel: en ĉiu kvadrato $p + [0,1]^2$, kie $p \in \mathbf{Z}^2$, $\mathbf{kan}(f)$ egalas al $\mathbf{env}\left(f\big|_{p+\{0,1\}^2}\right)$. □

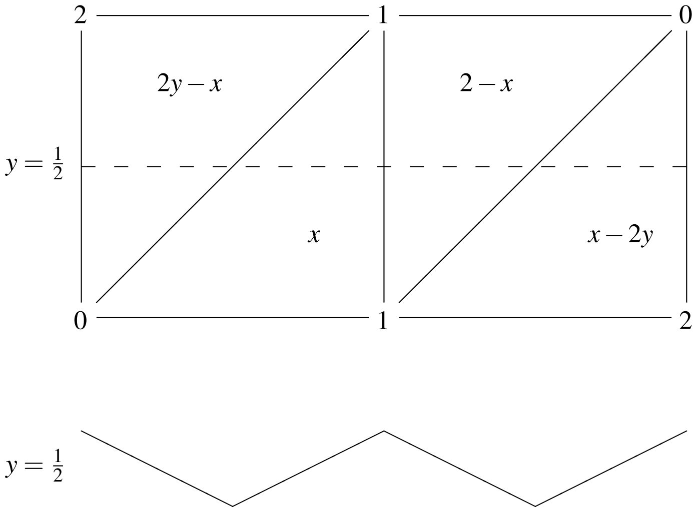
Povas okazi ke punkto x apartenas al du kvadratoj, $p + [0,1]^2$ kaj $q + [0,1]^2$. Tiam la difino donas la saman valoron sendepende de la kvadrato uzata.



Bildo 9. Kanona vastigaĵo de du funkcioj.

La kanona vastigaĵo estas difinita en $\mathbf{R}^2$ kaj konveksa en ĉiu kvadrato $p + [0,1]^2$, $p \in \mathbf{Z}^2$, sed ne nepre konveksa en la tuta ebeno $\mathbf{R}^2$ (bildo 10).

Ni ĉiam havas $\mathbf{env}(f) \leqslant \mathbf{kan}(f)$; egaleco validas se kaj nur se f estas romboide konveksa.



Bildo 10. La kanona vastigaĵo $\mathbf{kan}(f)$ de la funkcio $f(x,y) = x + 2y - 2xy$, $(x,y) \in \{0,1,2\} \times \{0,1\}$, ne estas konveksa. La konveksa envelopo de ĝi estas $\mathbf{env}(f)(x,y) = |x - 2y|$, $(x,y) \in [0,2] \times [0,1]$. Validas $1 = \mathbf{kan}(f)(1, \frac{1}{2}) > \mathbf{env}(f)(1, \frac{1}{2}) = 0$.

Ekzemplo 4.6. Ĉiu afina funkcio estas romboide konveksa. La maksimumo de du afinaj funkcioj $f(x,y) = \max(ax+by+c, Ax+By+C)$ estas ĉiam konvekse vastigebla; ĝi estas romboide konveksa se kaj nur se

$$(a-A)(b-B)(a-A-b+B)(a-A+b-B) = 0. \qquad \square$$

Ekzemplo 4.7. La polinomo $p(x,y) = ax^2 + 2bxy + cy^2$, $(x,y) \in \mathbf{Z}^2$, de grado du estas konvekse vastigebla se kaj nur se $a, c \geqslant 0$ kaj $ac \geqslant b^2$; kaj romboide konveksa se kaj nur se $a, c \geqslant |b|$. $\square$

Ekzemplo 4.8. La polinomo $p(x,y) = ax^4 + 6bx^2y^2 + cy^4$, $(x,y) \in \mathbf{Z}^2$, de grado kvar estas konvekse vastigebla se kaj nur se $a, b, c \geqslant 0$ kaj $ac \geqslant b^2$; ĝi estas romboide konveksa se kaj nur se $a, c \geqslant b \geqslant 0$. $\square$

5. Rezulto por marĝenaj funkcioj

Teoremo 5.1. *Se $f\colon \mathbf{Z}^2 \to \mathbf{R}$ estas romboide konveksa kaj malsupren barita, tiam ĝia marĝena funkcio*

$$h(x) = \inf_{y \in \mathbf{Z}} f(x,y), \qquad x \in \mathbf{Z},$$

estas konvekse vastigebla. $\square$

Romboida konvekseco estas esence necesa por ke la konkludo validu.

6. Rezulto por lokaj minimumoj

Por studi lokajn minimumojn estas interese trovi kiel eble plej malgrandajn ĉirkaŭaĵojn. Por optimuma rezulto ni lasu la ĉirkaŭaĵon dependi de la studata funkcio.

Difino 6.1. Se funkcio $f\colon \mathbf{Z}^2 \to \mathbf{R}$ kaj punkto $p \in \mathbf{Z}^2$ estas donitaj, ni diros ke punkto $q \in \mathbf{Z}^2$ estas ***f-najbaro de** p*, se $\|q-p\|_\infty = 1$ kaj se la streko $[p,q]$ estas knikstreko de $\mathbf{kan}(f)$, la kanona vastigaĵo de f. Ni notu per $N_f(p)$ la aron de ĉiuj f-najbaroj de p; ĝi estas la ***f-ĉirkaŭaĵo*** de la punkto p. $\square$

Ĉi tie la normo $\|x\|_\infty$ estas la l^∞-normo, $\|x\|_\infty = \max(|x_1|, |x_2|)$.

Knikstreko de konveksa funkcio estas streko $[a,b]$ tia ke f ne estas afina en iu ĉirkaŭaĵo de $]a,b[= [a,b] \smallsetminus \{a,b\}$.

Punkto povas havi $0, 2, 3, \ldots, 8$ najbarojn. Laŭdifine ili ĉiuj estas elementoj de la sfero

$$\begin{aligned} S(p) &= \{x \in \mathbf{Z}^2;\ \|x-p\|_\infty = 1\} \\ &= p + \{(1,0), (1,1), (0,1), (-1,1), (-1,0), (-1,-1), (0,-1), (1,-1)\}. \end{aligned}$$

Ĝi povas ĉiam servi kiel f-ĉirkaŭaĵon por ĉiu romboide konveksa funkcio f, sed kelkfoje pli malgranda aro taŭgas.

Teoremo 6.2. *Estu $f\colon \mathbf{Z}^2 \to \mathbf{R}$ romboide konveksa funkcio. Ni supozu ke la konveksa tegaĵo de la aro $\{x \in S(a);\ f(x) \geqslant f(a)\}$ enhavas la punkton a en sia interno. Tiam $f(a)$ estas la tutebena minimumo de f, t.e.*

$$f(a) = \inf_{y \in \mathbf{Z}^2} f(y).$$

Speciale, se la konveksa tegaĵo de $N_f(a)$ enhavas a en sia interno, tiam sufiĉas supozi ke $f(b) \geqslant f(a)$ por ĉiuj $b \in N_f(a)$. □

Ni vidas ke la nova klaso solvas la du problemojn: la dezirataj rezultoj validas en la nova klaso de romboide konveksaj funkcioj.

Referencoj

Favati, Paola; Tardella, Fabio (1990). Convexity in nonlinear programming. *Ricerca operativa*, **53**, 3–44.

Kiselman, Christer O. (2007). *Trois problèmes en convexité digitale : minima locaux, fonctions marginales et hyperplans séparants.* Lekcio ĉe GeoNet VI, 2007-06-12. Kaptebla ĉe `www2.math.uu.se/~kiselman`

Kiselman, Christer O. (2008). Minima locaux, fonctions marginales et hyperplans séparants dans l'optimisation discrète. *C. R. Acad. Sci. Paris Sér. I* **346**, 49–52.

Kiselman, Christer O. (2010). Local minima, marginal functions, and separating hyperplanes in discrete optimization. **En:** Rajendra Bhatia (red.), *Abstracts: Short communications; Posters*. International Congress of Mathematicians, Hyderabad, August 19–27, 2010, pp. 572–573.

Murota, Kazuo (2003). *Discrete Convex Analysis.* Philadelphia: Society for Industrial and Applied Mathematics.

ДИСКРЕТНА И ЦЕЛОЧИСЛЕНА ОПТИМИЗАЦИЯ

Проф. Кристер Киселман

Оптимизацията е важен клон от математиката. Думата означава търсене на най-доброто, от лат. *optimum* ‘най-добро, оптимално’. За да се намери най-доброто, първо да се реши какво е добро: за продавача най-добра е най-високата цена, за купувача – най-ниската. На второ място трябва да се реши кои са приемливите състояния и стойности. Например, обикновено не може да се купи половин ябълка, кола или къща: трябва да се купи нито едно или едно, две, . . . ; възможни са само цели положителни числа. От друга страна, оптималната дължина може да е 1,3 метра или $\pi \approx 3{,}14159265$ метра: в такава ситуация всички естествени числа са възможни.

Представят се няколко главни разлики между целочислена и естествена оптимизация и се показва как може да се дефинират подходящите класове функции с целочислени променливи, за които са валидни желателните резултати.