Pattes pysselhörna, del 2: Svar

Patte

November 15, 1995

Introduktion

Problemet denna gång gällde huruvida den välkända formeln för skuggpris i LP,

$$y_i^* = \frac{\partial z^*}{\partial b_i} = \frac{\partial(c_B^{\mathrm{T}} B^{-1} b)}{\partial b_i} = (c_B^{\mathrm{T}} B^{-1})_i, \tag{1}$$

alltid är giltig, om värdet av skuggpriset *alltid* återfinns i en optimaltablå (undantaget de skuggpris som motsvarar likhetsvillkor förstås), *varför* formeln är giltig, och, om den inte alltid är giltig, *när* den är det.

Frågorna

Fråga 1. LP-problemet:

$$\begin{array}{lrcrcrcrcl}
\min & z = -3x_1 & - & x_2 & & & & & & \\
\text{då} & x_1 & + & x_2 & + & x_3 & & & = & 1, \quad (*) \\
& 2x_1 & + & x_2 & & & + & x_4 & = & 2, \\
& x_1, & & x_2, & & x_3, & & x_4, & \geq & 0.
\end{array}$$

Optimaltablåer:

bas	$-z$	x_1	x_2	x_3	x_4	$\bar{b}$
$-z$	1	0	2	3	0	3
x_1	0	1	1	1	0	1
x_4	0	0	-1	-2	1	0

bas	$-z$	x_1	x_2	x_3	x_4	$\bar{b}$
$-z$	1	0	1/2	0	3/2	3
x_3	0	0	1/2	1	$-1/2$	0
x_1	0	1	1/2	0	1/2	1

Table 1: Två optimaltablåer

Samtliga optimala duallösningar:

$$y^* = t\begin{pmatrix}-3\\0\end{pmatrix} + [1-t]\begin{pmatrix}0\\-3/2\end{pmatrix}, \qquad 0 \leq t \leq 1.$$

Fråga 2. Skuggpriset för bivillkor $(*)$ [dvs. när vi talar om en *ökning* av b_1] är 0; den *negativa* marginella förändringen är -3. Vi kan alltså inte tala om *ett* skuggpris, för det fordrar att det två värdena är desamma. Vi ser att vi kan återfinnna dessa värden i de två tablåerna; men med tillgång till *en* optimaltablå (vilket de flesta programpaket tillhandahåller) kan vi inte besvara frågan. Formeln (1) tycks vara giltig för högerled 1 när dess optimala duala lösning är unik, men inte annars. Detta ska få sin matematiska förklaring i svaren på frågorna 3 och 4.

Fråga 3. Betrakta funktionen

$$f(\delta) = \min\{\, c^{\mathrm{T}} x \mid Ax = b + \delta; \quad x \geq 0 \,\}, \qquad \delta \in \Re^m.$$

(Vi noterar parentetiskt att $f(\delta) = c_B^{\mathrm{T}} B^{-1}(b+\delta)$ för varje optimal bas B i det LP-problem som motsvarar högerledet $b+\delta$.) Det som söks är *riktningsderivatan* av f i riktning d respektive $-d$, beräknad i $\delta = 0$, dvs.

$$f'(0; d) = \lim_{t \downarrow 0} \frac{1}{t}[f(td) - f(0)]$$

respektive

$$-f'(0; -d) = \lim_{t \uparrow 0} \frac{1}{t}[f(td) - f(0)].$$

Några egenskaper hos f. För att lättare studera funktionen f:s egenskaper skriver vi om f med hjälp av "tricket" LP-dualitet; med index-mängden $\mathcal{K}$ definierad som mängden av extrempunkter i det duala tillåtna området fås

$$\begin{aligned} f(\delta) &= \min\{\, c^{\mathrm{T}} x \mid Ax = b + \delta; \quad x \geq 0 \,\} \\ &= \max\{\, (b+\delta)^{\mathrm{T}} y \mid A^{\mathrm{T}} y \leq c \,\} \\ &= \max\{\, (b+\delta)^{\mathrm{T}} y^k \mid k \in \mathcal{K} \,\}. \end{aligned}$$

(Sista likheten följer av LP:s fundamentalsats.) Med den omskrivningen är det lätt att visa att f är en konvex funktion via konvexitetsdefinitionen. Det resultatet är i själva verket ett specialfall av det användbara resultatet att en funktion som är definerad som maximum av ett ändligt antal konvexa funktioner är en konvex funktion (nyttig övning i konvexitet). Värdet av funktionen bestäms av de duala extrempunkter som är *aktiva*, dvs. de extrempunkter y^k, $k \in \mathcal{K}(\delta) \subset \mathcal{K}$, för vilket

$$f(\delta) = (b+\delta)^{\mathrm{T}} y^k.$$

Värden av δ där mängden $\mathcal{K}(\delta)$ inte förändras motsvarar ett delområde där en dual baslösning är unik, och enligt formeln ovan är f linjär på varje sådant delområde. (Den är därför också differentierbar där.) De värden av δ där flera duala extrempunkter är optimala motsvarar (möjliga) icke-differentierbara punkter för f. På hela definitionsrummet för f (området av $\Re^m$ där f är ändlig—notera att för vissa värden av δ kan primala problemet sakna lösning) är alltså f en "styckvis" linjär funktion. (Här har jag valt att använda begreppet för det en-dimensionella fallet i brist på ett bättre.)

Riktningsderivatan. Vi börjar med att titta på vad subdifferentialen till f är, och utnyttjar att f är maximum av ett antal differentierbara, konvexa funktioner. För generalitetens skull tittar vi på den allmänna funktionen

$$f(x) = \max\{ f_i(x) \mid i \in \mathcal{I} \}, \qquad x \in \Re^n \tag{2}$$

där $\mathcal{I}$ är en ändlig mängd och samtliga f_i är konvexa och differentierbara funktioner. För denna funktion gäller att

$$\partial f(x) = \texttt{conv} \; \{ \nabla f_i(x) \mid i \in \mathcal{I}(x) \}, \qquad x \in \Re^n$$

där $\mathcal{I}(x) = \{ i \in \mathcal{I} \mid f_i(x) = f(x) \}$, dvs. det konvexa höljet av gradienterna till de aktiva funktionerna.

I vårt fall urartar detta till formeln

$$\partial f(\delta) = \texttt{conv} \; \{ y^k \mid k \in \mathcal{K}(\delta) \},$$

dvs. det konvexa höljet av de extremlösningar som är dualt optimala. Men eftersom detta är ett LP-problem är denna mängd lika med mängden av duala optimallösningar för det givna värdet av δ, och vi sammanfattar:

$$\partial f(\delta) = Y^*(\delta),$$

där $Y^*(\delta)$ betecknar mängden av optimala duallösningar för det givna värdet av δ.

Det vi är intresserade av är funktionen f:s uppträdande kring $\delta = 0$; vi vet nu att

$$\partial f(0) = Y^*(0),$$

dvs. att subdifferentialen av f i 0 sammanfaller med mängden av duala optimallösningar. Detta har vi nytta av när vi beräknar riktningsderivatan, genom att konvexiteten hos f ger att riktningsderivatan kan beräknas via formeln

$$f'(0; d) = \max\{ \xi_f^{\mathrm{T}} d \mid \xi_f \in \partial f(0) \}$$

dvs. den ges av det största värdet av skalärprodukten av d och en subgradient till f i $\delta = 0$ (se t.ex. Bazaraa m.fl.).

Formeln för den positiva marginella förändringen är därmed klar:

$$f'(0; d) = \max\{ {y^k}^{\mathrm{T}} d \mid k \in \mathcal{K}(0) \}. \tag{3}$$

(Anmärkning för den intresserade: vi kunde, inser jag nu, gått rakare på målet. Det finns ett resultat i minimax-teori av Danskin som säger att en funktion på formen (2), där varje funktion f_i är kontinuerligt differentierbar (de behöver inte vara konvexa), gäller att

$$f'(x;d) = \max\{ \nabla f_i(x)^{\mathrm{T}} d \mid i \in \mathcal{I}(x) \};$$

resultatet ovan är ett specialfall av detta. Det är lätt att vara efterklok, men det "krokigare" beviset gav lite mer insikt å andra sidan.)

På motsvarande sätt får vi att den negativa marginella förändringen är

$$-f'(0;-d) = \min\{ {y^k}^{\mathrm{T}} d \mid k \in \mathcal{K}(0) \}. \tag{4}$$

Dessa två värden är inte alltid lika; de är det under förutsättning att f är differentierbar, vilket alltså är detsamma som att dualen har en unik lösning! Om dualen har en unik optimallösning (vi brukar också säga att primala LP-lösningen är *icke-degenererad*) så är f differentierbar i $\delta = 0$ med $\nabla f(0) = y^*$, där y^* är den unika duallösningen, och formeln ovan ger

$$f'(0;d) = \nabla f(0)^{\mathrm{T}} d = {y^*}^{\mathrm{T}} d.$$

Vi ser nu att formeln (1) är giltig för alla bivillkor i precis då duallösningen är unik, och vi inser varför den inte är det annars; den partiella derivatan av z^* (alternativt f) som funktion av δ existerar ju inte annars!

Fråga 4. I fallet förändringar av endast *ett* högerled är förstås f en funktion i en variabel. Vi brukar betrakta enhetsförändringar, så vi inför beteckningen d_i^+ för en enhetsvektor i den i:te koordinaten och d_i^- för motsvarande negativa koordinatriktning. Formel (3) blir då:

$$f'(0;d_i^+) = \max\{ y_i^k \mid k \in \mathcal{K}(0) \},$$

dvs. det maximala värdet av den duala variabeln y_i bland alla optimala duallösningar. (Och motsvarande: det minsta värdet av y_i bland alla duala optimallösningar för en negativ förändring.) Vi får därför i exemplet:

$$f'(0;d_1^+) = 0 \qquad \text{och} \qquad -f'(0;d_1^-) = -3.$$

För att vi ska få ett entydigt värde (så att vi kan betrakta det som ett skuggpris) fordras att vänster- och högerderivatan av f är lika, dvs. att funktionen är deriverbar i $\delta = 0$, dvs. att duallösningen motsvarande högerledet i är unik. Formeln (3) sammanfaller då med den kända formeln (1), som sig bör. (Om vi bara är intresserade av att studera en-dimensionella förändringar så behöver alltså inte f vara differentierbar mer än i den aktuella koordiantriktningen.)

Fråga 5. Låt $d = (-.5, -.5)^{\mathrm{T}}$. Vi får:

$$f'(0;d) = 3/2 \qquad \text{respektive} \qquad -f'(0;-d) = 3/4.$$

[Notera att $-f'(0;-d) < f'(0;d)$.]

Fråga 6. Det framgår av formel (3) att man kan beräkna riktningsderivatan genom att jämföra värdet av skalärprodukten mellan d och samtliga duala extrema optimallösningar, så svaret på frågan är: JA. Men inte är det säkert praktiskt alla gånger; en rejält icke-degenererad optimallösning kan ju ha många optimala baser (det vet vi ju kan hända redan i fallet med två ursprungsvariabler och ett antal olikhetsvillkor; jämför uppgift i LP-exempelsamlingen), och speciellt nätverksproblem har *alltid* en degenererad optimallösning. För den som vill veta mera om hur man praktiskt kan gå tillväga hänvisar jag till Murtys bok om LP (1983, Sektion 8.15); där beskriver han hur man kan plocka fram marginella förändringen genom att [1] ta fram alla baser, dvs. utnyttja formeln (3) direkt, [2] genom att lösa ett nytt LP-problem över mängden av alla duala optimallösningar; detta LP-problem är detsamma som det som definerar $f'(0; d)$. Ett tredje alternativ är att använda sig av parametrisk LP.

Vi avslutar som vanligt med att tillsammans sjunga den glada signaturmelodin. Vinka med Era roliga hattar till tittarna! Vi ses nästa vecka. Hej då!