Optimering under osäkerhet

Föreläsning 5

Ann-Brith Strömberg

9 september 2003

Värdet av stokastisk optimering

- När tjänar vi på att använda stokastisk optimering?
- Kan vi använda en enklare modell?
- Vad förlorar vi i så fall?
- Kan vi beräkna övre och undre gränser för kostnaden?

First stage feasibility set
I ett tvåstegsproblem är "first stage feasibility set" de värden på steg 1-variablerna som medför att steg 2 har minst en tillåten lösning. Om vi förenklar problemet riskerar vi att välja en steg 1-lösning för vilken tillåtna steg 2-lösningar saknas.

Hädanefter antar vi att modellen har "relatively complete recourse", d.v.s. steg 2 *kan inte bli otillåtet* p.g.a. ett "dåligt" val i steg 1.

2-stegsproblem

Vi betraktar följande problem

$$\begin{aligned} \min \quad & c^{\mathrm{T}}x + q^{\mathrm{T}}y, \\ \text{då} \quad & Ax = b, \\ & T(\xi)x + Wy = h(\xi). \end{aligned}$$

Som tidigare kan vi skriva

$$\begin{aligned} Q(x,\xi) = \min \quad & q^{\mathrm{T}}y, \\ \text{då} \quad & Wy = h(\xi) - T(\xi)x. \end{aligned}$$

Vi definierar även

$$\begin{aligned} g(x,\xi) &= c^{\mathrm{T}}x + Q(x,\xi) \\ X_0 &= \{x | Ax = b\} \end{aligned}$$

2-stegsproblem

Vi har nu det optimala värdet för recourseproblemet (TS=Two-stage)

$$\mathrm{TS} = \min_{x \in X_0} \mathsf{E}_{\xi}[g(x,\xi)]$$

med optimal lösning x^*.

Om de stokastiska variablerna har diskreta fördelningar fås istället

$$\mathrm{TS} = \min_{x \in X_0} \left\{ \sum_{i=1}^{n} p_i g(x, \xi^i) \right\}$$

där $p_i = P(\xi = \xi^i)$.

Expected value solution

Vad händer om vi sätter in väntevärden i stället för stokastiska variabler? Antag diskreta fördelningar och låt

$$\bar{T} = \sum_{i=1}^{n} p_i T(\xi^i) \quad \text{och} \quad \bar{h} = \sum_{i=1}^{n} p_i h(\xi^i).$$

Vi formulerar problemet (EV = Expected Value)

$$\text{EV} = \min_{x \in X_0} g(x, \mathsf{E}[\xi]),$$

där

$$\begin{aligned} g(x, \mathsf{E}[\xi]) = c^{\mathrm{T}}x + \min \quad & q^{\mathrm{T}}y, \\ \text{då} \quad & Wy = \bar{h} - \bar{T}x, \end{aligned}$$

med optimal lösning $\bar{x}$.

Mått på osäkerhet

Ett mått på hur mycket osäkerheten påverkar oss kan vi nu få genom att beräkna EEV (Expected Value of using the Expected solution)

$$\text{EEV} = \mathsf{E}_\xi[g(\bar{x}, \xi)].$$

Vi får nu värdet av den stokastiska lösningen (VSS) som

$$\text{VSS} = \text{EEV} - \text{TS}.$$

Vi vet att TS $\leq$ EEV ty $\bar{x}$ är tillåten men ej nödvändigtvis optimal i orginalproblemet.

Värdet av perfekt information

Vad skulle vi vara beredda att betala för en kristallkula som förutsäger framtiden?

Detta avgörs av värdet WS (Wait and See solution).

Om vi hela tiden vet det framtida utfallet kan vi för varje enskilt ξ lösa problemet

$$\min_{x \in X_0} \quad g(x, \xi)$$

med optimal lösning $x_\xi \in X_0 \Rightarrow$

$$\text{WS} = \mathsf{E}_\xi[g(x_\xi, \xi)].$$

Förväntat värde av perfekt information

Vi kan nu få fram “Expected Value of Perfect Information” (EVPI) som TS – WS (för minproblem).

Detta är samtidigt det högsta pris vi är beredda att betala för att exakt kunna förutsäga framtiden.

Eftersom x^* (optimala lösningen för recourseproblemet) för varje värde på ξ är tillåten men ej nödvändigtvis optimal för problemet

$$\min_{x \in X_0} g(x, \xi),$$

har vi att EVPI ≥ 0 eller, om vi sammanfattar,

$$\text{WS} \leq \text{TS} \leq \text{EEV}.$$

Jordbruksexemplet

I jordbruksexemplet får vi

TS	-108 390
EEV	-107 240
WS	-115 406

vilket ger VSS $=108390 - 107240 = 1150$

samt EVPI $=115406 - 108390 = 7016$

Konvexitet

Ibland är $Q(x,\xi)$ konvex som funktion av x för varje fixt ξ.
Exempel:

$$\begin{aligned} Q(x,\xi) = \min \quad & q^{\mathrm{T}} y \\ \text{då} \quad & Wy = h(\xi) - T(\xi)x \end{aligned}$$

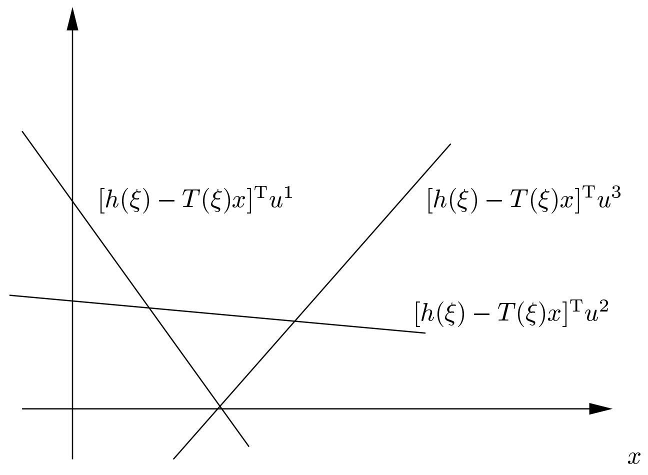
För att se detta tar vi fram dualen:

$$\begin{aligned} Q(x,\xi) = \max \quad & [h(\xi) - T(\xi)x]^{\mathrm{T}} u \\ \text{då} \quad & W^{\mathrm{T}} u \le q \end{aligned}$$

Den duala tillåtna polyedern är oberoende av x. För varje fixt (x,ξ) är en extrempunkt hos polyedern optimal. Om en given extrempunkt u^i är optimal för (x,ξ) gäller

$$Q(x,\xi) = [h(\xi) - T(\xi)x]^{\mathrm{T}} u^i.$$

Konvexitet



Konvexitet

Vi fixerar nu x till ett visst värde, säg $\widehat{x}$.

$$\begin{aligned} Q(\widehat{x},\xi) = \max \quad & [h(\xi) - T(\xi)\widehat{x}]^{\mathrm{T}} u \\ \text{då} \quad & W^{\mathrm{T}} u \le q \end{aligned}$$

Målfunktionen $h(\xi) - T(\xi)\widehat{x}$ är en linjär funktion av ξ.

Därmed är den konvex i både x och ξ var och en för sig (dock ej säkert om båda varieras samtidigt).

Men om vi även antar att T är konstant (d.v.s. $T(\xi) = T$) så är $g(x,\xi)$ konvex både i x och ξ samtidigt.

Jensens olikhet

För en konvex funktion f gäller konvexitetsolikheten:

$$f\left((1-\lambda)x+\lambda y\right) \le (1-\lambda)f(x)+\lambda f(y), \quad \forall \lambda \in [0,1], \forall x,y$$

Mera generellt gäller Jensens olikhet:

$$f\left(\textstyle\sum_i \lambda_i x^i\right) \le \sum_i \lambda_i f(x^i)$$

$$\text{då} \qquad \sum_i \lambda_i = 1 \quad \text{och} \quad \lambda_i \ge 0 \quad \forall i$$

Gäller även vid integration:

$$f\left(\textstyle\int_\Xi \xi dP(\xi)\right) \le \int_\Xi f(\xi) dP(\xi)$$

Jensens olikhet

För en funktion $g(x,\xi)$ *som är konvex som funktion av* ξ gäller därför att

$$g(x,\mathsf{E}[\xi]) = g\left(x, \textstyle\int_\Xi \xi dP(\xi)\right) \le \int_\Xi g(x,\xi)dP(\xi) = \mathsf{E}_\xi[g(x,\xi)]$$

Då denna olikhet gäller för alla x gäller den även för $\bar{x}$.

Därmed gäller

$$g(\overline{x},\mathsf{E}[\xi]) = \mathrm{EV} \le \mathrm{EEV} = \mathsf{E}_\xi[g(\overline{x},\xi)]$$

Minimerar vi istället över $x \in X_0$ får vi

$$\mathrm{EV} = \min_{x\in X_0} g(x,\mathsf{E}[\xi]) \le \min_{x \in X_0} \mathsf{E}_\xi[g(x,\xi)] = \mathrm{TS}$$

Konvexitet—sammanfattning

Har vi konvexitet i ξ och x var och en för sig gäller

$$\mathrm{WS} \le \mathrm{TS} \le \mathrm{EEV}$$
$$\mathrm{EV} \le \mathrm{TS} \le \mathrm{EEV}$$

vilket medför

$$\mathrm{EVPI} = \mathrm{TS} - \mathrm{WS} \le \mathrm{EEV} - \mathrm{WS}$$
$$\mathrm{VSS} = \mathrm{EEV} - \mathrm{TS} \le \mathrm{EEV} - \max(\mathrm{EV},\mathrm{WS})$$

Edmundson-Madansky övre gräns

För en konvex funktion f gäller

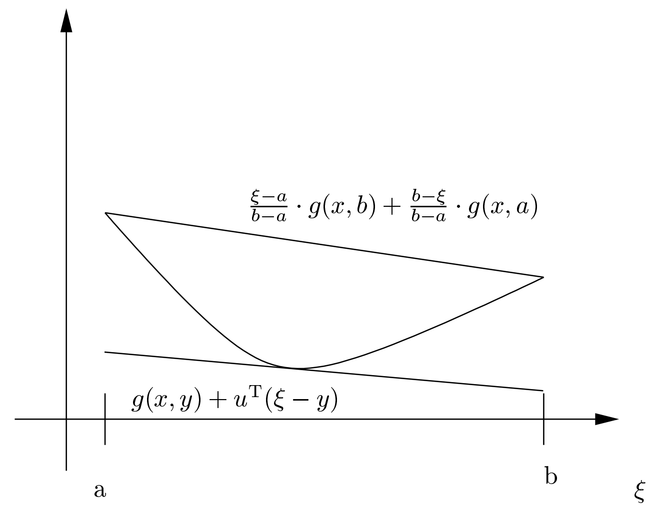
$$f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda)f(y).$$

Om vi har endast en stokastisk variabel ξ med utfallsrum intervallet $[a,b]$ har vi därför för alla x att

$$g(x,\xi) \le \frac{\xi - a}{b-a}\cdot g(x,b) + \frac{b-\xi}{b-a}\cdot g(x,a).$$

Integrerar vi detta får vi

$$\begin{aligned}\mathsf{E}_\xi[g(x,\xi)] &\le \mathsf{E}_\xi\left[\frac{\xi-a}{b-a}\cdot g(x,b) + \frac{b-\xi}{b-a}\cdot g(x,a)\right] \\ &= \frac{\mathsf{E}[\xi]-a}{b-a}\cdot g(x,b) + \frac{b-\mathsf{E}[\xi]}{b-a}\cdot g(x,a)\end{aligned}$$



Gränser

- Detta innebär att vi kan få fram en övre och undre gräns för TS, även om de slumpmässiga variablerna är kontinuerliga.
- Vi kan betrakta gränserna som givna av approximationer till funktionen.
- Vi kan lika gärna se dem som givna av approximationer av tähetsfunktionerna för de stokastiska variablerna.
- I fallet för den undre gränsen ersätter vi täthetsfunktionen med en diskret funktion med sannolikhet 1 i $\mathsf{E}[\xi]$, och i fall 2 ersätter vi den med en Bernoulli-fördelning med sannolikhet $\frac{\mathsf{E}[\xi]-a}{b-a}$ i b och $\frac{b-\mathsf{E}[\xi]}{b-a}$ i a.

Sammanfattning

En funktion $g(x,\xi)$ som tidigare, med $\xi \in [a,b]$. Antag att $g(x,\xi)$ är konvex i x och ξ var för sig. Vi söker värdet av

$$\min_{x \in X_0} \mathsf{E}_\xi[g(x,\xi)]$$

En undre gräns ges av

$$\min_{x \in X_0} g(x, \mathsf{E}[\xi]).$$

En övre gräns ges av

$$\min_{x \in X_0} \left(\frac{\mathsf{E}[\xi]-a}{b-a} g(x,b) + \frac{b-\mathsf{E}[\xi]}{b-a} g(x,a) \right).$$