KTH Matematik

5B1815 Tillämpad linjär optimering

Föreläsning 11

Subgradientmetoder.

Subgradient och subgradientmetod

Betrakta problemet

$$\begin{array}{ll} \text{max} & \varphi(u) \\ \text{då} & u \in I\!R^m, \end{array}$$

där φ är en konkav funktion på $I\!R^m$.

Definition. *En vektor $s \in I\!R^m$ är en* subgradient *till φ i en punkt u om $\varphi(v) \le \varphi(u) + s^T(v-u)$ för alla $v \in I\!R^m$.*

Om φ är differentierbar i u ges den unika subgradienten av $\nabla\varphi(u)$.

Idén bakom en subgradientmetod är att iterativt maximera $\varphi(u)$ genom att beräkna subgradienter och ta steg i subgradientens riktning.

Några egenskaper hos subgradienten

Påstående. *Antag att u inte är optimal till (D). Låt s vara en subgradient till φ i u och låt u^* vara optimal till (D). Då gäller att $s^T(u^* - u) > 0$.*

Bevis. Idé: $0 < \varphi(u^*) - \varphi(u) \leq s^T(u^* - u)$. $\square$

OBS! Det gäller *inte* säkert att $\varphi(u + \theta s) > \varphi(u)$ för θ positiv nära noll.

Påstående. *Antag att u inte är optimal till (D). Låt s vara en subgradient till φ i u och låt u^* vara optimal till (D). Då gäller att*

$$\|u + \theta s - u^*\|_2^2 < \|u - u^*\|_2^2 \quad \textit{för} \quad \theta \in \left(0, \frac{2s^T(u^* - u)}{\|s\|_2^2}\right).$$

Bevis. Idé: $\|u + \theta s - u^*\|_2^2 = \|u - u^*\|_2^2 - 2\theta s^T(u^* - u) + \theta^2 \|s\|_2^2$. $\square$

Ett litet steg i subgradientens riktning tar oss närmare optimallösningen.

Några egenskaper hos subgradienten, forts.

Påstående. *En punkt $u^* \in I\!R^m$ är optimal till (D) om och endast om nollvektorn är en subgradient till φ i u^*.*

Bevis. Idé: $0^T(u - u^*) \geq \varphi(u) - \varphi(u^*)$ för alla u om och endast om $\varphi(u) \leq \varphi(u^*)$ för alla u. $\square$

OBS! Detta gäller då (D) inte har några bivillkor.

Påstående. *Antag att s^1 och s^2 är subgradienter till φ i u. Då är $(1-\alpha)s^1 + \alpha s^2$ också en subgradient till φ i u för $\alpha \in [0, 1]$.*

Bevis. Idé:

$$\left.\begin{aligned} \varphi(v) &\leq \varphi(u) + (s^1)^T(v-u) \\ \varphi(v) &\leq \varphi(u) + (s^2)^T(v-u) \end{aligned}\right. \Rightarrow \varphi(v) \leq \varphi(u) + ((1-\alpha)s^1 + \alpha s^2)^T(v-u).$$

Subgradienter till lagrangeduala problem

Påstående. *Antag att* $\varphi(u) = \min_{x \in X}\{f(x) - u^T g(x)\}$. *Då är* $-g(x(u))$ *en subgradient till* φ *i* u, *där* $x(u)$ *är en optimallösning till* $\min_{x \in X}\{f(x) - u^T g(x)\}$.

Bevis. Idé: Låt $x(v)$ vara optimal till $\min_{x \in X}\{f(x) - v^T g(x)\}$. Då får vi

$$\begin{aligned}
\varphi(v) = f(x(v)) - v^T g(x(v)) &\leq f(x(u)) - v^T g(x(u)) \\
&= f(x(u)) - u^T g(x(u)) - g(x(u))^T (v - u) \\
&= \varphi(u) - g(x(u))^T (v - u). \quad \square
\end{aligned}$$

Genom att lösa det lagrangerelaxerade problemet $\min_{x \in X}\{f(x) - u^T g(x)\}$ får vi ur optimallösningen $x(u)$ en subgradient $-g(x(u))$.

Subgradienter i heltalsprogrammeringsproblem

Vi kan tillämpa detta på heltalsprogrammeringsproblemet

$$(IP) \qquad \begin{array}{ll} \text{min} & c^T x \\ \text{då} & Ax \geq b, \\ & x \in X, \end{array}$$

där $X = \{x \in I\!R^n : Cx \geq d,\ x \geq 0,\ x \text{ heltalig}\}$.

För ett givet $u \in I\!R^m$, $u \geq 0$, blir den duala målfunktionen

$$\varphi(u) = \min_{x \in X} \{c^T x - u^T(Ax - b)\}.$$

Om $x(u)$ är en optimallösning till $\min_{x \in X}\{c^T x - u^T(Ax - b)\}$ blir $b - Ax(u)$ en subgradient till φ i u.

Exemplproblem

$$\begin{array}{ll} \text{min} & -x_1 - x_2 - x_3 \\ \text{då} & x_1 - x_2 - x_3 \geq 0, \\ & 2x_2 - 2x_3 \geq 0, \\ & x_1 + x_2 + x_3 = 1, \\ & x_j \in \{0,1\}, \quad j = 1,2,3. \end{array}$$

Låt $g(x) = \begin{pmatrix} x_1 - x_2 - x_3 \\ 2x_2 - 2x_3 \end{pmatrix}$, $\mathcal{I} = \{1,2\}$, $\mathcal{E} = \emptyset$,
$X = \{x \in I\!R^3 : x_1 + x_2 + x_3 = 1,\ x_j \in \{0,1\}, j = 1,2,3\}$.

Lagrangerelaxering av bivillkor 1 och 2 i exempelproblemet

Antag att bivillkoren $x_1 - x_2 - x_3 \ge 0$ och $2x_2 - 2x_3 \ge 0$ lagrangerelaxeras med motsvarande multiplikatorer u_1 och u_2.

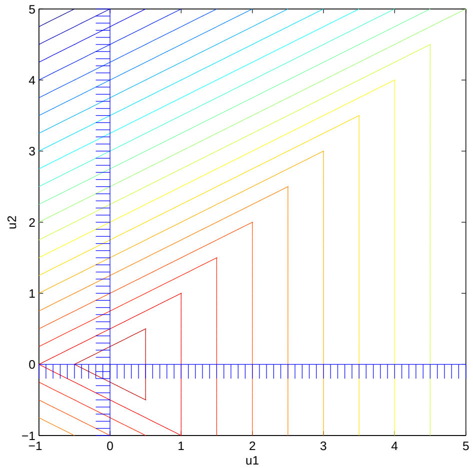
Det duala problemet ges av

$$(D) \quad \begin{array}{ll} \text{max} & \varphi(u) \\ \text{då} & u \ge 0, \end{array} \quad \text{där}$$

$$\begin{aligned} \varphi(u) = \min\ & -x_1 - x_2 - x_3 - u_1(x_1 - x_2 - x_3) - u_2(2x_2 - 2x_3) \\ \text{då}\quad & x_1 + x_2 + x_3 = 1,\ x_j \in \{0,1\},\ j = 1,2,3, \\ = \min & (-u_1 - 1)x_1 + (u_1 - 2u_2 - 1)x_2 + (u_1 + 2u_2 - 1)x_3 \\ \text{då}\quad & x \in \{(1\ 0\ 0)^T, (0\ 1\ 0)^T, (0\ 0\ 1)^T\}, \\ = \min & \{-u_1 - 1, u_1 - 2u_2 - 1, u_1 + 2u_2 - 1\}. \end{aligned}$$

Geometrisk illustration av duala problemet till exempelproblemet

Det tvådimensionella duala problemet (D) kan illustreras grafiskt:

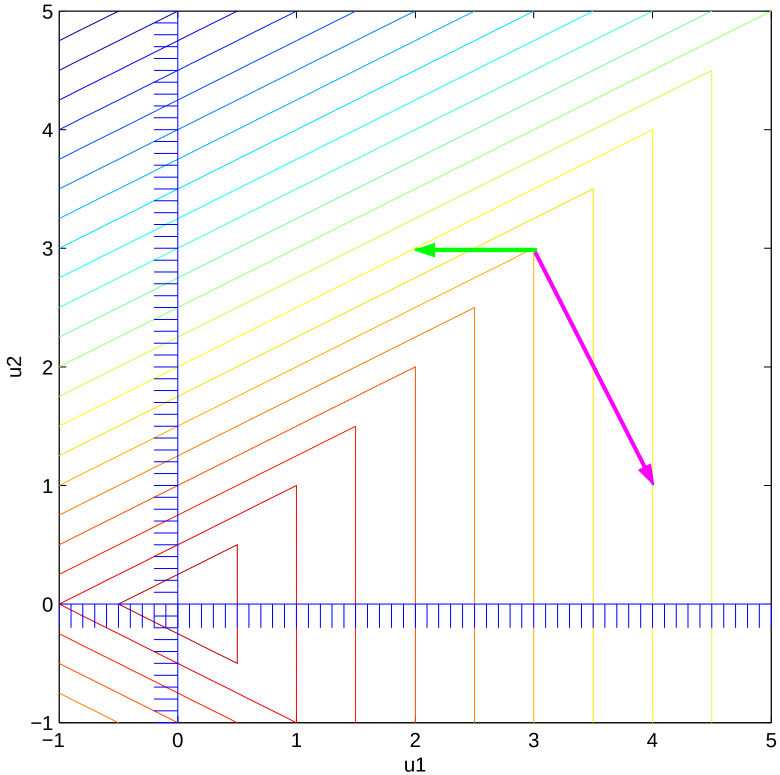


$$u_2 \geq u_1 \geq 0 \Rightarrow x(u) = (0\ 1\ 0)^T,\ \varphi(u) = u_1 - 2u_2 - 1,\ g(x(u)) = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$$

$$u_1 \geq u_2 \geq 0 \Rightarrow x(u) = (1\ 0\ 0)^T,\ \varphi(u) = -u_1 - 1,\ g(x(u)) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Geometrisk illustration av subgradienter till det duala problemet

Det tvådimensionella duala problemet (D) kan illustreras grafiskt:



Då $u_1 = u_2$ är subgradienten inte unik. Exempelvis $u_1 = u_2 = 3$. Notera att φ avtar i de markerade subgradienternas riktningar.

Subgradientmetod

Iteration k i en subgradientmetod tar följande form givet $u^k \in I\!R^m$ där $u_i^k \geq 0$, $i \in \mathcal{I}$:

$$x^k \leftarrow \text{optimallösning till } \min_{x \in X}\{c^T x - (u^k)^T(Ax - b)\},$$

$$s^k \leftarrow b - Ax^k,$$

$$\theta^k \leftarrow \text{steglängd},$$

$$u^{k+1} \leftarrow u^k + \theta^k s^k, \qquad u_i^{k+1} \leftarrow \max\{u_i^{k+1}, 0\},\ i \in \mathcal{I}.$$

Några kommentarer:

- Punkten $u^k + \theta^k s^k$ projiceras på tillåtna området $\{u : u_i \geq 0, i \in \mathcal{I}\}$.
- Möjligt val av steglängd är $\theta^k = 1/k$. (Bättre val finns.)
- Ofta långsam konvergens "nära" u^*.
- Viktigt att generera primalt tillåtna lösningar efterhand.

Grov sammanfattning av kursen

Kursen kan grovt summeras enligt följande:

- Modeller och metoder för linjära optimeringsproblem där variablerna eventuellt varit heltaliga.
- Osäkerhet i problemdata har modellerats med stokastisk programmering.
- Metoder för LP-problem: simplexmetoden och inrepunktsmetoder.
- Dekomposition och kolumngenerering.
- Trädsökning, LP-relaxering och lagrangerelaxering för IP-problem.
- Subgradientmetoder för duala problem.
- Projekten avklarade. Tentamen återstår. Lycka till!