15.093: Optimization Methods

Lecture 14: Lagrangean Methods

1 Outline

SLIDE 1

- The Lagrangean dual
- The strength of the Lagrangean dual
- Solution of the Lagrangean dual

2 The Lagrangean dual

SLIDE 2

- Consider

$$\begin{aligned} Z_{\mathrm{IP}} = \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{D}\boldsymbol{x} \geq \boldsymbol{d} \\ & \boldsymbol{x} \text{ integer} \end{aligned}$$

- $X = \{\boldsymbol{x} \text{ integer} \mid \boldsymbol{D}\boldsymbol{x} \geq \boldsymbol{d}\}$
- Optimizing over X can be done efficiently

2.1 Formulation

SLIDE 3

- Consider

$$\begin{aligned} Z(\boldsymbol{\lambda}) = \min \quad & \boldsymbol{c}'\boldsymbol{x} + \boldsymbol{\lambda}'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in X \end{aligned} \quad (D)$$

- For fixed $\boldsymbol{\lambda}$, problem can be solved efficiently
- $Z(\boldsymbol{\lambda}) = \min_{i=1,\dots,m} \left(\boldsymbol{c}'\boldsymbol{x}^i + \boldsymbol{\lambda}'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^i)\right)$.
- $Z(\boldsymbol{\lambda})$ is concave and piecewise linear

2.2 Weak Duality

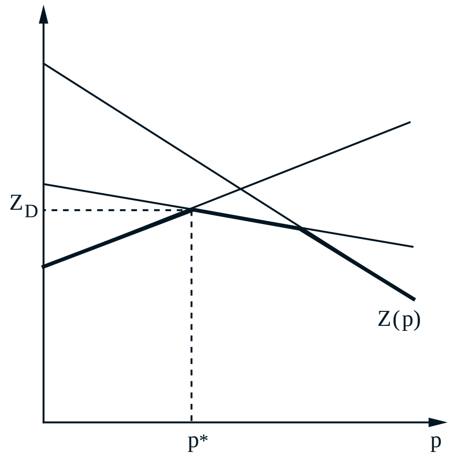
SLIDE 4

If problem (D) has an optimal solution and if $\boldsymbol{\lambda} \geq \boldsymbol{0}$, then $Z(\boldsymbol{\lambda}) \leq Z_{\mathrm{IP}}$

- **Proof:** $\boldsymbol{x}^*$ an optimal solution to (D).
- Then $\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^* \leq \boldsymbol{0}$ and, therefore,

$$\boldsymbol{c}'\boldsymbol{x}^* + \boldsymbol{\lambda}'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^*) \leq \boldsymbol{c}'\boldsymbol{x}^* = Z_{\mathrm{IP}}$$

- Since $\boldsymbol{x}^* \in X$, $Z(\boldsymbol{\lambda}) \leq \boldsymbol{c}'\boldsymbol{x}^* + \boldsymbol{\lambda}'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^*)$, and thus, $Z(\boldsymbol{\lambda}) \leq Z_{\mathrm{IP}}$



2.3 Key problem

SLIDE 5

- Consider the *Lagrangean dual*:

$$\begin{aligned} Z_{\mathrm{D}} = \max \quad & Z(\boldsymbol{\lambda}) \\ \text{s.t.} \quad & \boldsymbol{\lambda} \geq \mathbf{0} \end{aligned}$$

- $Z_{\mathrm{D}} \leq Z_{\mathrm{IP}}$
- We need to maximize a piecewise linear concave function

SLIDE 6

3 Strength of LD

3.1 Main Theorem

SLIDE 7

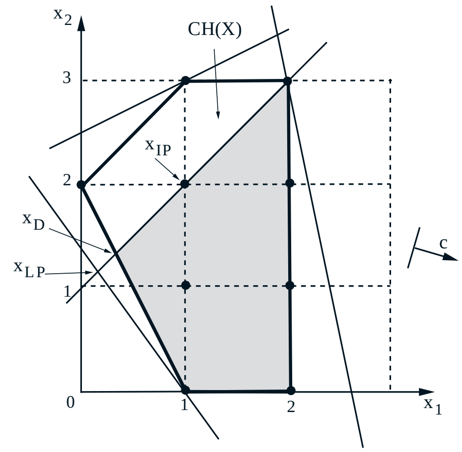
$X = \{\boldsymbol{x} \text{ integer} \mid \boldsymbol{D}\boldsymbol{x} \geq \boldsymbol{d}\}$. Note that CH$(X)$ is a polyhedron. Then

$$\begin{aligned} Z_{\mathrm{D}} = \min \quad & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \in \mathrm{CH}(X) \end{aligned}$$

3.2 Example

SLIDE 8

$$\begin{array}{llcrcr} \min & 3x_1 & - & x_2 & & \\ \text{s.t.} & x_1 & - & x_2 & \geq & -1 \\ & -x_1 & + & 2x_2 & \leq & 5 \\ & 3x_1 & + & 2x_2 & \geq & 3 \\ & 6x_1 & + & x_2 & \leq & 15 \\ & \multicolumn{3}{r}{x_1, x_2 \geq 0} & & \text{integer} \end{array}$$



Relax $x_1 - x_2 \geq -1$, X involves the remaining constraints

$$X = \big\{(1,0),(2,0),(1,1),(2,1),(0,2),$$
$$(1,2),(2,2),(1,3),(2,3)\big\}.$$

SLIDE 9

For $p \geq 0$, we have SLIDE 10

$$Z(p) = \min_{(x_1,x_2)\in X} \big(3x_1 - x_2 + p(-1 - x_1 + x_2)\big)$$

$$Z(p) = \begin{cases} -2 + p, & 0 \leq p \leq 5/3, \\ 3 - 2p, & 5/3 \leq p \leq 3, \\ 6 - 3p, & p \geq 3. \end{cases}$$

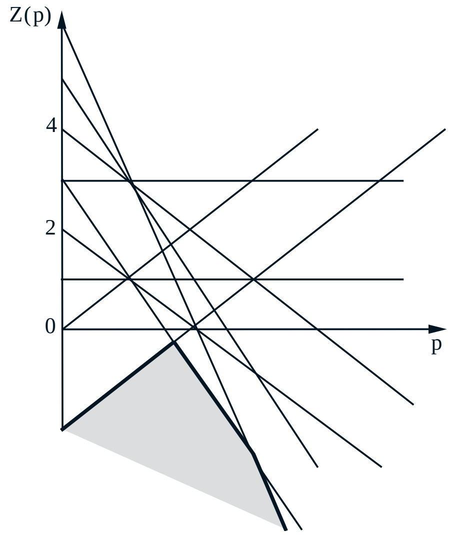
$p^* = 5/3$, and $Z_{\mathrm{D}} = Z(5/3) = -1/3$ SLIDE 11

SLIDE 12

- $\boldsymbol{x}_{\mathrm{D}} = \big(1/3, 4/3\big)$, $Z_{\mathrm{D}} = -1/3$
- $\boldsymbol{x}_{\mathrm{LP}} = \big(1/5, 6/5\big)$, $Z_{\mathrm{LP}} = -9/5$
- $\boldsymbol{x}_{\mathrm{IP}} = (1,2)$, $Z_{\mathrm{IP}} = 1$
- $Z_{\mathrm{LP}} < Z_{\mathrm{D}} < Z_{\mathrm{IP}}$

SLIDE 13

- In general, $Z_{\mathrm{LP}} \leq Z_{\mathrm{D}} \leq Z_{\mathrm{IP}}$
- For $\boldsymbol{c}'\boldsymbol{x} = 3x_1 - x_2$, we have $Z_{\mathrm{LP}} < Z_{\mathrm{D}} < Z_{\mathrm{IP}}$.
- For $\boldsymbol{c}'\boldsymbol{x} = -x_1 + x_2$, we have $Z_{\mathrm{LP}} < Z_{\mathrm{D}} = Z_{\mathrm{IP}}$.
- For $\boldsymbol{c}'\boldsymbol{x} = -x_1 - x_2$, we have $Z_{\mathrm{LP}} = Z_{\mathrm{D}} = Z_{\mathrm{IP}}$.
- It is also possible: $Z_{\mathrm{LP}} = Z_{\mathrm{D}} < Z_{\mathrm{IP}}$ but not on this example.



3.3 LP and LD

SLIDE 14

- $Z_{\mathrm{IP}} = Z_{\mathrm{D}}$ for all cost vectors $\boldsymbol{c}$, if and only if

$$\mathrm{CH}\Big(X \cap \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\}\Big) = \mathrm{CH}(X) \cap \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b}\}$$

- We have $Z_{\mathrm{LP}} = Z_{\mathrm{D}}$ for all cost vectors $\boldsymbol{c}$, if

$$\mathrm{CH}(X) = \{\boldsymbol{x} \mid \boldsymbol{D}\boldsymbol{x} \geq \boldsymbol{d}\}$$

- If $\{\boldsymbol{x} \mid \boldsymbol{D}\boldsymbol{x} \geq \boldsymbol{d}\}$, has integer extreme points, then $\mathrm{CH}(X) = \{\boldsymbol{x} \mid \boldsymbol{D}\boldsymbol{x} \geq \boldsymbol{d}\}$, and therefore $Z_{\mathrm{D}} = Z_{\mathrm{LP}}$

4 Solution of LD

SLIDE 15

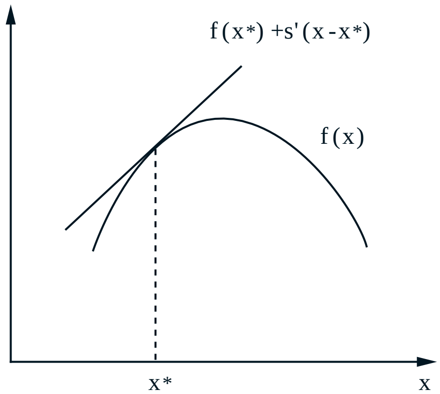
- $Z(\boldsymbol{\lambda}) = \min_{i=1,\ldots,m} \ \big(\boldsymbol{c}'\boldsymbol{x}^i + \boldsymbol{\lambda}'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^i)\big)$, i.e.,

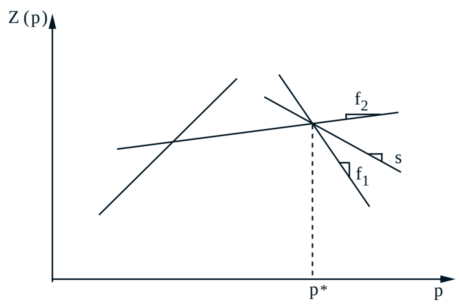
$$Z(\boldsymbol{\lambda}) = \min_{i=1,\ldots,m} \big(h_i + \boldsymbol{f}_i'\boldsymbol{\lambda}\big).$$

- Motivation: classical steepest ascent method for maximizing $Z(\boldsymbol{\lambda})$

$$\boldsymbol{\lambda}^{t+1} = \boldsymbol{\lambda}^t + \theta_t \nabla Z(\boldsymbol{\lambda}^t), \quad t = 1, 2, \ldots$$

- Problem: $Z(\boldsymbol{\lambda})$ is not differentiable





4.1 Subgradients

Slide 16

- A function $f : \Re^n \mapsto \Re$ is concave if and only if for any $\boldsymbol{x}^* \in \Re^n$, there exists a vector $\boldsymbol{s} \in \Re^n$ such that

$$f(\boldsymbol{x}) \leq f(\boldsymbol{x}^*) + \boldsymbol{s}'(\boldsymbol{x} - \boldsymbol{x}^*),$$

  for all $\boldsymbol{x} \in \Re^n$.

- Let f be a concave function. A vector $\boldsymbol{s}$ such that

$$f(\boldsymbol{x}) \leq f(\boldsymbol{x}^*) + \boldsymbol{s}'(\boldsymbol{x} - \boldsymbol{x}^*),$$

  for all $\boldsymbol{x} \in \Re^n$, is called a **subgradient** of f at $\boldsymbol{x}^*$.

Slide 17

Slide 18

4.2 Subgradient Algorithm

Slide 19

1. Choose a starting point $\boldsymbol{\lambda}^1$; let $t = 1$.
2. Given $\boldsymbol{\lambda}^t$, choose a subgradient $\boldsymbol{s}^t$ of the function $Z(\cdot)$ at $\boldsymbol{\lambda}^t$. If $\boldsymbol{s}^t = \boldsymbol{0}$, then $\boldsymbol{\lambda}^t$ is optimal and the algorithm terminates. Else, continue.

t	p^t	s^t	$Z(p^t)$
1	5.00	-3	-9.00
2	2.60	-3	-1.80
3	0.68	1	-1.32
4	1.19	1	-0.81
5	1.60	1	-0.40
6	1.92	-2	-0.84
7	1.40	1	-0.60
8	1.61	1	-0.39
9	1.78	-2	-0.56
10	1.51	1	-0.49

3. Let $\boldsymbol{\lambda}^{t+1} = \boldsymbol{\lambda}^t + \theta_t \boldsymbol{s}^t$, where θ_t is a positive stepsize parameter. Increment t and go to Step 2.

3a If $\boldsymbol{\lambda} \geq \mathbf{0}$, $p_j^{t+1} = \max\left\{p_j^t + \theta_t s_j^t, 0\right\}, \qquad \forall\ j.$

4.2.1 Step sizes

SLIDE 20

- $Z(\boldsymbol{p}^t)$ converges to the unconstrained maximum of $Z(\cdot)$, for any stepsize sequence θ_t such that

$$\sum_{t=1}^{\infty} \theta_t = \infty, \qquad \text{and} \qquad \lim_{t\to\infty} \theta_t = 0.$$

- Examples $\theta_t = 1/t$
- $\theta_t = \theta_0 \alpha^t, \qquad t = 1, 2, \ldots,$
- $\theta_t = \frac{\hat{Z}_{\mathrm{D}} - Z(\boldsymbol{p}^t)}{||\boldsymbol{s}^t||^2}\alpha^t$, where α satisfies $0 < \alpha < 1$, and $\hat{Z}_{\mathrm{D}}$ is an estimate of the optimal value Z_{D}.

4.3 Example

SLIDE 21

Recall $p^* = 5/3 = 1.66$ and $Z_{\mathrm{D}} = -1/3 = -0.33$. Apply subgradient optimization:

MIT OpenCourseWare
http://ocw.mit.edu

15.093J / 6.255J Optimization Methods
Fall 2009

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.