

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

$$\begin{aligned} Z_{\text{IP}} = \min \quad & \mathbf{c}'\mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} \geq \mathbf{b} \\ & \mathbf{Dx} \geq \mathbf{d} \\ & \mathbf{x} \text{ integer} \end{aligned}$$

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

2.1 Formulation

SLIDE 3

- Consider

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

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

2.2 Weak Duality

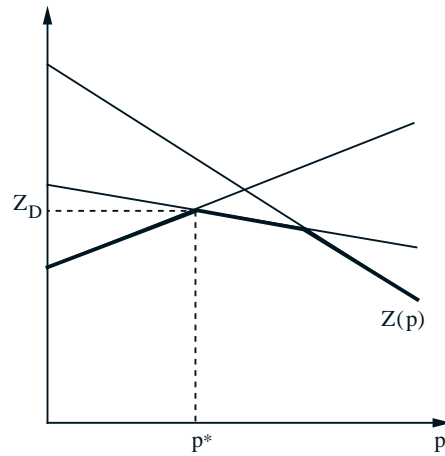
SLIDE 4

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

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

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

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



2.3 Key problem

SLIDE 5

- Consider the *Lagrangian dual*:

$$Z_D = \max_{\lambda \geq 0} Z(\lambda)$$

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

SLIDE 6

3 Strength of LD

3.1 Main Theorem

SLIDE 7

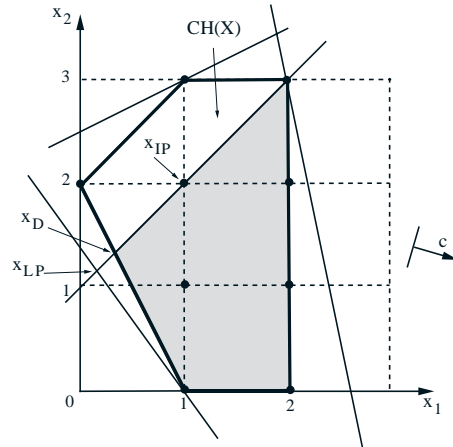
$X = \{x \text{ integer} \mid Dx \geq d\}$. Note that $\text{CH}(X)$ is a polyhedron. Then

$$Z_D = \min_{\substack{c'x \\ Ax \geq b \\ x \in \text{CH}(X)}}$$

3.2 Example

SLIDE 8

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



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

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

For $p \geq 0$, we have

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

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

$p^* = 5/3$, and $Z_D = Z(5/3) = -1/3$

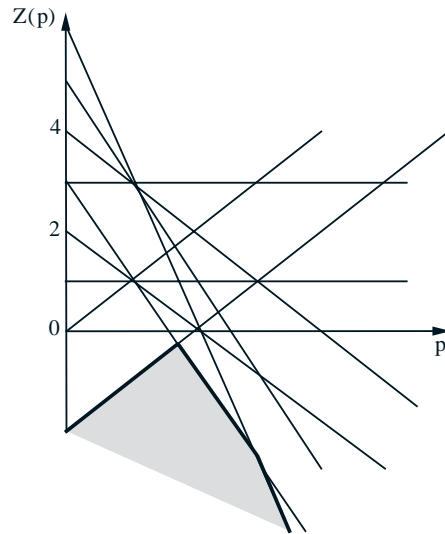
- $\mathbf{x}_D = (1/3, 4/3)$, $Z_D = -1/3$
- $\mathbf{x}_{LP} = (1/5, 6/5)$, $Z_{LP} = -9/5$
- $\mathbf{x}_{IP} = (1, 2)$, $Z_{IP} = 1$
- $Z_{LP} < Z_D < Z_{IP}$

SLIDE 9
SLIDE 10

SLIDE 11
SLIDE 12

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

SLIDE 13



3.3 LP and LD

SLIDE 14

- $Z_{IP} = Z_D$ for all cost vectors c , if and only if

$$\text{CH}(X \cap \{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}\}) = \text{CH}(X) \cap \{\mathbf{x} \mid \mathbf{Ax} \geq \mathbf{b}\}$$

- We have $Z_{LP} = Z_D$ for all cost vectors c , if

$$\text{CH}(X) = \{\mathbf{x} \mid \mathbf{Dx} \geq \mathbf{d}\}$$

- If $\{\mathbf{x} \mid \mathbf{Dx} \geq \mathbf{d}\}$, has integer extreme points, then $\text{CH}(X) = \{\mathbf{x} \mid \mathbf{Dx} \geq \mathbf{d}\}$, and therefore $Z_D = Z_{LP}$

4 Solution of LD

SLIDE 15

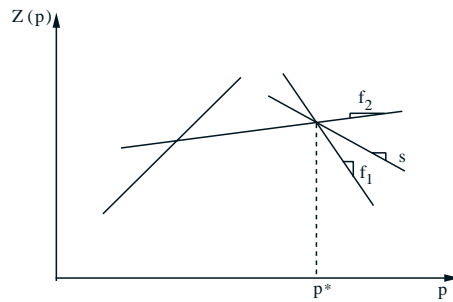
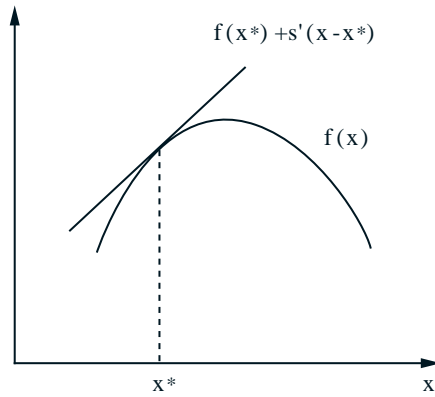
- $Z(\boldsymbol{\lambda}) = \min_{i=1, \dots, m} (c^i \mathbf{x}^i + \boldsymbol{\lambda}'(\mathbf{b} - \mathbf{Ax}^i))$, i.e.,

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

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

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

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



4.1 Subgradients

SLIDE 16

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

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

for all $\mathbf{x} \in \mathfrak{R}^n$.

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

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

for all $\mathbf{x} \in \mathfrak{R}^n$, is called a **subgradient** of f at \mathbf{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 \mathbf{s}^t of the function $Z(\cdot)$ at $\boldsymbol{\lambda}^t$. If $\mathbf{s}^t = \mathbf{0}$, 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 $\lambda^{t+1} = \lambda^t + \theta_t s^t$, where θ_t is a positive stepsize parameter. Increment t and go to Step 2.

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

4.2.1 Step sizes

SLIDE 20

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

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

- Examples $\theta_t = 1/t$
- $\theta_t = \theta_0 \alpha^t$, $t = 1, 2, \dots$,
- $\theta_t = \frac{\hat{Z}_D - Z(p^t)}{\|s^t\|^2} \alpha^t$, where α satisfies $0 < \alpha < 1$, and \hat{Z}_D is an estimate of the optimal value Z_D .

4.3 Example

SLIDE 21

Recall $p^* = 5/3 = 1.66$ and $Z_D = -1/3 = -0.33$. Apply subgradient optimization: