15.083J/6.859J Integer Optimization

Lecture 9: Duality II

1 Outline

Slide 1

- Solution of Lagrangean dual
- Geometry and strength of the Lagrangean dual

2 The TSP

Slide 2

$$\begin{aligned}
\sum_{e\in\delta(\{i\})} x_e &= 2, && i \in V,\\
\sum_{e\in E(S)} x_e &\le |S|-1, && S \subset V,\ S \ne \emptyset, V,\\
x_e &\in \{0,1\}.
\end{aligned}$$

$$\begin{aligned}
\min \quad & \sum_{e\in E} c_e x_e \\
\text{s.t.} \quad & \sum_{e\in\delta(\{i\})} x_e = 2, && i \in V\setminus\{1\},\\
& \sum_{e\in\delta(\{1\})} x_e = 2,\\
& \sum_{e\in E(S)} x_e \le |S|-1, && S \subset V\setminus\{1\},\ S \ne \emptyset, V\setminus\{1\},\\
& \sum_{e\in E(V\setminus\{1\})} x_e = |V|-2,\\
& x_e \in \{0,1\}.
\end{aligned}$$

Dualize $\sum_{e\in\delta(\{i\})} x_e = 2, \quad i \in V\setminus\{1\}$.
What is the relation of Z_{D} and Z_{LP}?

3 Solution

Slide 3

- $Z(\boldsymbol{\lambda}) = \min_{k\in K} \left(\boldsymbol{c}'\boldsymbol{x}^k + \boldsymbol{\lambda}'(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x}^k)\right)$, $\boldsymbol{x}^k,\ k\in K$ are extreme points of $\mathrm{conv}(X)$.
- $\boldsymbol{f}_k = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}^k$ and $h_k = \boldsymbol{c}'\boldsymbol{x}^k$.
- $Z(\boldsymbol{\lambda}) = \min_{k\in K}\left(h_k + \boldsymbol{f}_k'\boldsymbol{\lambda}\right)$, piecewise linear and concave.
- Recall $\boldsymbol{\lambda}^{t+1} = \boldsymbol{\lambda}^t + \theta_t \nabla Z(\boldsymbol{\lambda}^t)$

3.1 Subgradients

Slide 4

- Prop: $f : \Re^n \mapsto \Re$ is concave if and only if for any $\boldsymbol{x}^* \in \Re^n$, there exists a vector $\boldsymbol{s} \in \Re^n$ such that
$$f(\boldsymbol{x}) \le f(\boldsymbol{x}^*) + \boldsymbol{s}'(\boldsymbol{x}-\boldsymbol{x}^*).$$

- Def: f concave. A vector $\boldsymbol{s}$ such that for all $\boldsymbol{x} \in \Re^n$:

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

  is called a **subgradient** of f at $\boldsymbol{x}^*$. The set of all subgradients of f at $\boldsymbol{x}^*$ is denoted by $\partial f(\boldsymbol{x}^*)$ and is called the **subdifferential** of f at $\boldsymbol{x}^*$.
- Prop: $f : \Re^n \mapsto \Re$ be concave. A vector $\boldsymbol{x}^*$ maximizes f over $\Re^n$ if and only if $\boldsymbol{0} \in \partial f(\boldsymbol{x}^*)$.

Slide 5

$$\begin{aligned} Z(\boldsymbol{\lambda}) &= \min_{k \in K} \left(h_k + \boldsymbol{f}'_k \boldsymbol{\lambda}\right), \\ E(\boldsymbol{\lambda}) &= \left\{k \in K \mid Z(\boldsymbol{\lambda}) = h_k + \boldsymbol{f}'_k \boldsymbol{\lambda}\right\}. \end{aligned}$$

Then, for every $\boldsymbol{\lambda}^* \geq \boldsymbol{0}$ the following relations hold:

- For every $k \in E(\boldsymbol{\lambda}^*)$, $\boldsymbol{f}_k$ is a subgradient of the function $Z(\cdot)$ at $\boldsymbol{\lambda}^*$.
- $\partial Z(\boldsymbol{\lambda}^*) = \text{conv}\big(\{\boldsymbol{f}_k \mid k \in E(\boldsymbol{\lambda}^*)\}\big)$, i.e., a vector $\boldsymbol{s}$ is a subgradient of the function $Z(\cdot)$ at $\boldsymbol{\lambda}^*$ if and only if $Z(\boldsymbol{\lambda}^*)$ is a convex combination of the vectors $\boldsymbol{f}_k,\ k \in E(\boldsymbol{\lambda}^*)$.

3.2 The subgradient algorithm

Slide 6

Input: A nondifferentiable concave function $Z(\boldsymbol{\lambda})$.

Output: A maximizer of $Z(\boldsymbol{\lambda})$ subject to $\boldsymbol{\lambda} \geq \boldsymbol{0}$.

Algorithm:

1. Choose a starting point $\boldsymbol{\lambda}^1 \geq \boldsymbol{0}$; let $t = 1$.

2. Given $\boldsymbol{\lambda}^t$, check whether $\boldsymbol{0} \in \partial Z(\boldsymbol{\lambda}^t)$. If so, then $\boldsymbol{\lambda}^t$ is optimal and the algorithm terminates. Else, choose a subgradient $\boldsymbol{s}^t$ of the function $Z(\boldsymbol{\lambda}^t)$.

3. Let $\lambda_j^{t+1} = \max\left\{\lambda_j^t + \theta_t s_j^t, 0\right\}$, where θ_t is a positive stepsize parameter. Increment t and go to Step 2.

3.2.1 Step length

Slide 7

- $\sum_{t=1}^{\infty} \theta_t = \infty, \qquad$ and $\qquad \lim_{t\to\infty} \theta_t = 0.$
- Example: $\theta_t = 1/t$.
- Example: $\theta_t = \theta_0 \alpha^t, \qquad t = 1, 2, \ldots, 0 < \alpha < 1.$
- $\theta_t = f \frac{\hat{Z}_D - Z(\boldsymbol{\lambda}^t)}{||\boldsymbol{s}^t||^2}$, where f satisfies $0 < f < 2$, and $\hat{Z}_D$ is an estimate of the optimal value Z_D.
- The stopping criterion $\boldsymbol{0} \in \partial Z(\boldsymbol{\lambda}^t)$ is rarely met. Typically, the algorithm is stopped after a fixed number of iterations.

3.3 Example

Slide 8

- $Z(\lambda) = \min\left\{3 - 2\lambda,\ 6 - 3\lambda,\ 2 - \lambda,\ 5 - 2\lambda,\ -2 + \lambda,\ 1,\ 4 - \lambda,\ \lambda,\ 3\right\}$,
- $\theta_t = 0.8^t$.

-

λ^t	s^t	$Z(\lambda^t)$
1.5.00	-3	-9.00
2.2.60	-2	-2.20
3.1.32	-1	-0.68
4.1.83	2	-0.66
5.1.01	1	-0.99
6.1.34	1	-0.66
7.1.60	1	-0.40
8.1.81	-2	-0.62
9.1.48	1	-0.52
10.1.61	1	-0.39

4 Nonlinear problems

SLIDE 9

- $$\begin{aligned} Z_{\rm P} = \min \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0}, \\ & \boldsymbol{x} \in X. \end{aligned}$$
- $Z(\boldsymbol{\lambda}) = \min_{\boldsymbol{x}\in X} \left\{ f(\boldsymbol{x}) + \boldsymbol{\lambda}'\boldsymbol{g}(\boldsymbol{x}) \right\}$.
- $Z_{\rm D} = \max_{\boldsymbol{\lambda}\ge\boldsymbol{0}} Z(\boldsymbol{\lambda})$.
- $Y = \{(y, \boldsymbol{z}) \mid y \ge f(\boldsymbol{x}), \boldsymbol{z} \ge \boldsymbol{g}(\boldsymbol{x}),\ \text{for all } \boldsymbol{x} \in X\}$.
- $\begin{aligned} Z_{\rm P} = \min \quad & y \\ \text{s.t.} \quad & (y, \boldsymbol{0}) \in Y. \end{aligned}$
- $Z(\boldsymbol{\lambda}) \le f(\boldsymbol{x}) + \boldsymbol{\lambda}'\boldsymbol{g}(\boldsymbol{x}) \le y + \boldsymbol{\lambda}'\boldsymbol{z}, \qquad \forall (y, \boldsymbol{z}) \in Y.$
- Geometrically, the hyperplane $Z(\boldsymbol{\lambda}) = y + \boldsymbol{\lambda}'\boldsymbol{z}$ lies below the set Y.
- Theorem:
$$\begin{aligned} Z_{\rm D} = \min \quad & y \\ \text{s.t.} \quad & (y, \boldsymbol{0}) \in \text{conv}(Y). \end{aligned}$$

4.1 Figure

SLIDE 10

4.2 Example again

SLIDE 11

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

4.3 Subgradient algorithm

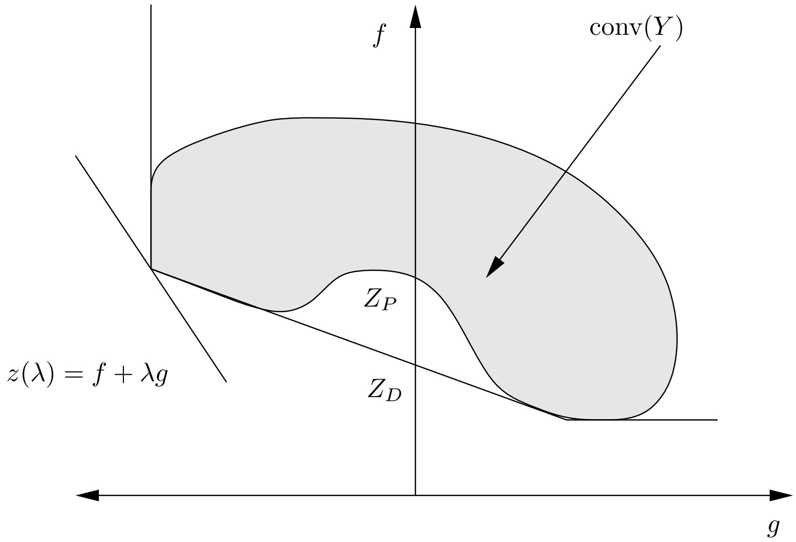
SLIDE 12

Input: Convex functions $f(\boldsymbol{x}), g_1(\boldsymbol{x}), \ldots, g_m(\boldsymbol{x})$ and a convex set X.

Output: An approximate minimizer.

Algorithm:

1. **(Initialization)** Select a vector $\overline{\boldsymbol{\lambda}}$ and solve $\min_{\boldsymbol{x}\in X} \left\{ f(\boldsymbol{x}) + \boldsymbol{\lambda}'\boldsymbol{g}(\boldsymbol{x}) \right\}$ to obtain the optimal value $\overline{Z}$ and an optimal solution $\overline{\boldsymbol{x}}$. Set $\boldsymbol{x}^0 = \overline{\boldsymbol{x}}$; $Z^0 = \overline{Z}$; $t = 1$.
2. **(Stopping criterion)** If $(|f(\overline{\boldsymbol{x}}) - \overline{Z}|/\overline{Z}) < \epsilon_1$ and $(\sum_{i=1}^m |\overline{\lambda_i}|/m) < \epsilon_2$ stop; Output $\overline{\boldsymbol{x}}$ and $\overline{Z}$ as the solution to the Lagrangean dual problem.



3. **(Subgradient computation)** Compute a subgradient $\boldsymbol{s}^t$; $\lambda_j^t = \max\{\overline{\lambda}_j + \theta_t s_j^t, 0\}$, where

$$\theta_t = g\frac{\hat{Z} - Z_{\mathrm{LP}}(\overline{\boldsymbol{\lambda}})}{||\boldsymbol{s}^t||^2}$$

with $\hat{Z}$ an upper bound on Z_{D}, and $0 < g < 2$. With $\boldsymbol{\lambda} = \boldsymbol{\lambda}^t$ solve $\min_{\boldsymbol{x}\in X} \left\{ f(\boldsymbol{x}) + \boldsymbol{\lambda}'\boldsymbol{g}(\boldsymbol{x}) \right\}$ to obtain the optimal value Z^t and an optimal solution $\boldsymbol{x}^t$.

4. **(Solution update)** Update

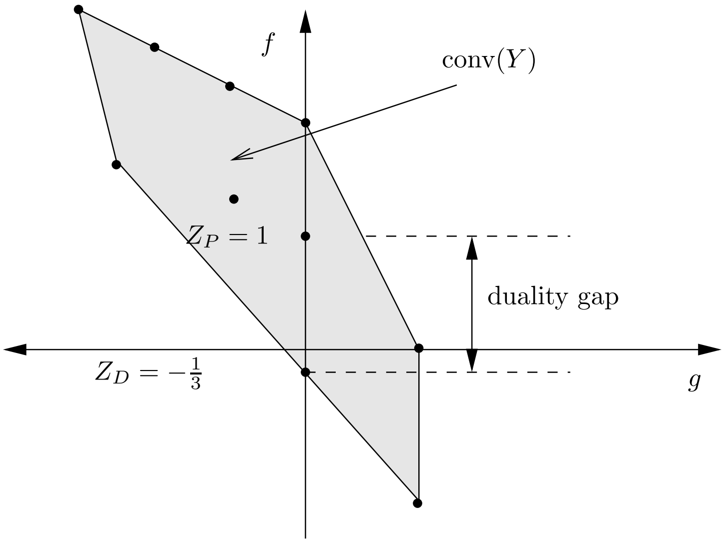
$$\overline{\boldsymbol{x}} \leftarrow \alpha \boldsymbol{x}^t + (1-\alpha)\overline{\boldsymbol{x}}$$

where $0 < \alpha < 1$.

5. **(Improving step)** If $Z^t > \overline{Z}$, then

$$\overline{\boldsymbol{\lambda}} \leftarrow \boldsymbol{\lambda}^t, \qquad \overline{Z} \leftarrow Z^t;$$

Let $t \leftarrow t + 1$ and go to Step 2.



f
conv(Y)
Z_P = 1
duality gap
Z_D = −1/3
g