15.093: Optimization Methods

Lecture 14: Lagrangean Methods
1 Outline

• The Lagrangean dual
• The strength of the Lagrangean dual
• Solution of the Lagrangean dual

2 The Lagrangean dual

• Consider
  \[
  Z_{IP} = \min_{\text{s.t. } Ax \geq b, \quad Dx \geq d} c'x \quad \text{subject to } x \text{ integer}
  \]
  \[
  X = \{x \text{ integer } | Dx \geq d\}
  \]

• Optimizing over $X$ can be done efficiently

2.1 Formulation

• Consider
  \[
  Z(\lambda) = \min_{x \in X} c'x + \lambda'(b - Ax) \quad (D)
  \]

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

2.2 Weak Duality

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

• Proof: $x^*$ an optimal solution to $(D)$.

• Then $b - Ax^* \leq 0$ and, therefore,
  \[
  c'x^* + \lambda'(b - Ax^*) \leq c'x^* = Z_{IP}
  \]

• Since $x^* \in X$, $Z(\lambda) \leq c'x^* + \lambda'(b - Ax^*)$, and thus, $Z(\lambda) \leq Z_{IP}$
2.3 Key problem

- 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

3 Strength of LD

3.1 Main Theorem

\[ X = \{ x \text{ integer} \mid Dx \geq d \} \]

Note that \( \text{CH}(X) \) is a polyhedron. Then

\[ Z_D = \min_{c'x} \text{ s.t. } Ax \geq b, \quad x \in \text{CH}(X) \]

3.2 Example

\[
\begin{align*}
\text{min} & \quad 3x_1 - x_2 \\
\text{s.t.} & \quad x_1 - x_2 \geq -1 \\
& \quad -x_1 + 2x_2 \leq 5 \\
& \quad 3x_1 + 2x_2 \geq 3 \\
& \quad 6x_1 + x_2 \leq 15 \\
& \quad x_1, x_2 \geq 0 \quad \text{integer}
\end{align*}
\]
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 $Z_D = Z(5/3) = -1/3$

- $x_D = (1/3, 4/3)$, $Z_D = -1/3$
- $x_{LP} = (1/5, 6/5)$, $Z_{LP} = -9/5$
- $x_{IP} = (1, 2)$, $Z_{IP} = 1$
- $Z_{LP} < Z_D < Z_{IP}$

In general, $Z_{LP} \leq Z_D \leq Z_{IP}$

- For $c^T x = 3x_1 - x_2$, we have $Z_{LP} < Z_D < Z_{IP}$.
- For $c^T x = -x_1 + x_2$, we have $Z_{LP} < Z_D = Z_{IP}$.
- For $c^T x = -x_1 - x_2$, we have $Z_{LP} = Z_D = Z_{IP}$.
- It is also possible: $Z_{LP} = Z_D < Z_{IP}$ but not on this example.
3.3 LP and LD

- $Z_{LP} = Z_D$ for all cost vectors $c$, if and only if
  \[
  \text{CH}\left( X \cap \{ x \mid Ax \geq b \} \right) = \text{CH}(X) \cap \{ x \mid Ax \geq b \}
  \]

- We have $Z_{LP} = Z_D$ for all cost vectors $c$, if
  \[
  \text{CH}(X) = \{ x \mid Dx \geq d \}
  \]

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

4 Solution of LD

- $Z(\lambda) = \min_{i=1,..,m} (c^i x^i + \lambda^i (b - Ax^i))$, i.e.,
  \[
  Z(\lambda) = \min_{i=1,..,m} (h_i + f_i(\lambda)).
  \]

- Motivation: classical steepest ascent method for maximizing $Z(\lambda)$
  \[
  \lambda^{t+1} = \lambda^t + \theta_t \nabla Z(\lambda^t), \quad t = 1, 2, \ldots
  \]

- Problem: $Z(\lambda)$ is not differentiable
4.1 Subgradients

- A function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is concave if and only if for any \( x^* \in \mathbb{R}^n \), there exists a vector \( s \in \mathbb{R}^n \) such that

\[
f(x) \leq f(x^*) + s^T(x - x^*),
\]

for all \( x \in \mathbb{R}^n \).

- Let \( f \) be a concave function. A vector \( s \) such that

\[
f(x) \leq f(x^*) + s^T(x - x^*),
\]

for all \( x \in \mathbb{R}^n \), is called a subgradient of \( f \) at \( x^* \).

4.2 Subgradient Algorithm

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

- \( Z(p^t) \) converges to the unconstrained maximum of \( Z(\cdot) \), for any stepsize sequence \( \theta_t \) such that
  \[
  \sum_{t=1}^{\infty} \theta_t = \infty, \quad \text{and} \quad \lim_{t \to \infty} \theta_t = 0.
  \]
- Examples \( \theta_t = 1/t \)
- \( \theta_t = \theta_0 \alpha^t, \quad t = 1, 2, \ldots \)
- \( \theta_t = \frac{\tilde{Z}_t - Z(p^t)}{\|s^t\|} \alpha^t \), where \( \alpha \) satisfies \( 0 < \alpha < 1 \), and \( \tilde{Z}_D \) is an estimate of the optimal value \( Z_D \).

4.3 Example

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

\[
\begin{array}{|c|c|c|c|}
\hline
\text{t} & p^t & s^t & Z(p^t) \\
\hline
1 & 5.00 & -3 & -9.00 \\
2 & 2.60 & -3 & -1.80 \\
3 & 0.68 & 1 & -1.32 \\
4 & 1.19 & 1 & -0.81 \\
5 & 1.60 & 1 & -0.40 \\
6 & 1.92 & -2 & -0.84 \\
7 & 1.40 & 1 & -0.60 \\
8 & 1.61 & 1 & -0.30 \\
9 & 1.78 & -2 & -0.56 \\
10 & 1.51 & 1 & -0.49 \\
\hline
\end{array}
\]