Bundle Methods

(Primal and Dual interpretations)

Ashutosh Mahajan

Department of Industrial and Systems Engineering
Lehigh University

COR@L Seminar Series, Spring 2006

References:

- *Applying Bundle Methods to the Optimization of Polyhedral Functions: An Applications-Oriented Development*, C.Paolo, F.Antonio, N.Maddalena, Technical Report, TR-96-17, December 01, 1997, University of Pisa
- *Nonlinear Programming: Theory and Applications*, M.S.Bazaraa, H.D. Sherali, C.M.Shetty, 2nd Ed., 1993, John Wiley and Sons..

Review

Consider a polyhedral function:

$$
\begin{aligned}
\varphi_{CP}(\mathbf{y}) &= \min_i \{\varphi(\mathbf{y}_i) + \mathbf{g}(\mathbf{y}_i).(\mathbf{y} - \mathbf{y}_i) | i \in \beta\} \\
&= \min_i \{\varphi(\mathbf{y}_i) + \varphi(\bar{\mathbf{y}}) - \varphi(\bar{\mathbf{y}}) + \mathbf{g}(\mathbf{y}_i)(\bar{\mathbf{y}} - \mathbf{y}_i) + \\
&\quad \mathbf{g}(\mathbf{y}_i)(\mathbf{y} - \bar{\mathbf{y}}) | i \in \beta\} \\
&= \min_i \{\varphi(\mathbf{y}_i) - \varphi(\bar{\mathbf{y}}) + \mathbf{g}(\mathbf{y}_i)(\bar{\mathbf{y}} - \mathbf{y}_i) + \\
&\quad \mathbf{g}(\mathbf{y}_i)(\mathbf{y} - \bar{\mathbf{y}}) | i \in \beta\} + \varphi(\bar{\mathbf{y}}) \\
\varphi_{CP}(\mathbf{d}) &= \min_i \{\alpha_i + \mathbf{g}_i \mathbf{d} | i \in \beta\} + \varphi(\bar{\mathbf{y}})
\end{aligned}
$$

where,

$$\mathbf{g}_i = \mathbf{g}(\mathbf{y}_i), \alpha_i = \varphi(\mathbf{y}_i) - \varphi(\bar{\mathbf{y}}) + \mathbf{g}(\mathbf{y}_i)(\bar{\mathbf{y}} - \mathbf{y}_i) \geq 0$$

Review(2)

$$(D_\beta) : \max_{d}\{\varphi_d(\mathbf{d})\} \quad = \quad \max_{v,d}\{v | v \leq \alpha_i + \mathbf{g}_i\mathbf{d}, i \in \beta\}$$

Stabilizing, we get:

$$(\Pi_{\beta t}) : \max_{v,d}\{v - \frac{1}{2t}||d||^2 | v \leq \alpha_i + \mathbf{g}_i\mathbf{d}, i \in \beta\}$$

Dual:

$$(\Delta_{\beta t}) : \min_{\theta}\{\frac{1}{2}t||\sum_i \mathbf{g}_i\theta_i||^2 + \sum_i \alpha_i\theta_i | \sum_i \theta_i = 1, \theta_i \geq 0\}$$

Review(3)

$$(\Delta_{\beta t}) : t. \min_{\theta} \{ \frac{1}{2} || \sum_i \mathbf{g}_i \theta_i ||^2 + \frac{1}{t} \sum_i \alpha_i \theta_i | \sum_i \theta_i = 1, \theta_i \geq 0 \}$$

Same as the problem:

$$(\Delta_{\beta}^{\epsilon}) : \min_{\theta} \frac{1}{2} || \sum_i \mathbf{g}_i \theta_i ||^2$$

$$\text{s.t. } \sum_i \alpha_i \theta_i \leq \epsilon$$

$$\sum_i \theta_i = 1, \theta_i \geq 0$$

which is a smoothing of:

$$\min_{g} \{ ||\mathbf{g}|| : \mathbf{g} \in conv(\{ \mathbf{g}_i : \alpha_i \leq \epsilon \}) \}$$

Primal Interpretation

Consider the problem:

$$(P) : min_x \{\mathbf{cx} | \mathbf{x} \in X, \mathbf{Ax} = \mathbf{b}\}$$

- Let, $X' = \{\mathbf{x_i}\}$, be the set of all extreme points of X.
- Let X be bounded.
- Suppose its easy to solve (P) over X alone.

Dantzig-Wolfe Decomposition

If $\mathbf{x} \in X$, then there exist $\theta_i \geq 0$ s.t.

$$\mathbf{x} = \sum_i \mathbf{x}_i \theta_i, \sum_i \theta_i = 1$$

Hence,

$$(P) : \min_x \{\mathbf{cx} | \mathbf{x} \in X, A\mathbf{x} = \mathbf{b}\}$$

is same as:

$$(M) : \min_\theta \mathbf{c}(\sum_i \mathbf{x}_i \theta_i)$$

$$\text{s.t. } A(\sum_i \mathbf{x}_i \theta_i) = \mathbf{b}$$

$$\sum_i \theta_i = 1, \theta_i \geq 0$$

Dantzig-Wolfe Decomposition(2)

$$(M_\beta): \min_\theta \mathbf{c}(\sum_{i\in\beta} \mathbf{x}_i\theta_i)$$

$$\text{s.t. } A(\sum_{i\in\beta} \mathbf{x}_i\theta_i) = \mathbf{b}, \sum_{i\in\beta} \theta_i = 1, \theta_i \geq 0$$

$$(L_\beta): \max_{y,v} \mathbf{yb} + v$$

$$\text{s.t. } v + \mathbf{y}A\mathbf{x}_i \leq \mathbf{cx}_i, i \in \beta$$

Call a new problem:

$$(P_{c-y*A}): \min_x \{(\mathbf{c} - \mathbf{y}*A)\mathbf{x} | \mathbf{x} \in X$$

If $v^* \leq (\mathbf{c} - \mathbf{y}*A)\mathbf{x}$, we are done, else add $\mathbf{x}^*$ to M_β.

Bundles

- Let, $\varphi(\mathbf{y}) = \mathbf{yb} + \min\{(\mathbf{c} - \mathbf{y}A)\mathbf{x} : \mathbf{x} \in X\}$
- Then, $\mathbf{g}(\bar{\mathbf{y}}) = \mathbf{b} - A\mathbf{x}(\bar{\mathbf{y}}) \in \partial\varphi(\bar{\mathbf{y}})$.
- Let $\mathbf{x}_i = \mathbf{x}(\mathbf{y}_i)$, $\mathbf{g}_i = \mathbf{g}(\mathbf{y}_i)$.
- Then, $\varphi(\mathbf{y}) = \mathbf{cx}(\mathbf{y}) + \mathbf{y}.\mathbf{g}(\mathbf{y})$.

$$\begin{aligned} \alpha_i &= \varphi(\mathbf{y}_i) - \varphi(\bar{\mathbf{y}}) + \mathbf{g}_i(\bar{\mathbf{y}} - \mathbf{y}_i) \\ &= (\mathbf{c} - \bar{\mathbf{y}}A)(\mathbf{x_i} - \mathbf{x}(\bar{\mathbf{y}})) \end{aligned}$$

For $\bar{\mathbf{y}} = 0, \varphi(\mathbf{0}) = \mathbf{cx}(\mathbf{0}), \alpha_i = \mathbf{c}(\mathbf{x_i} - \mathbf{x}(\mathbf{0}))$.

Bundles

- (L_β) : $\max_{y,v}\{v | v \leq (\mathbf{c} - \mathbf{y}A)\mathbf{x}_i + \mathbf{yb} - \mathbf{cx}(\mathbf{0}), i \in \beta\}$.
- (L_β) : $\max_{y,v}\{v | v \leq \mathbf{g}_i\mathbf{y} + \alpha_i, i \in \beta\} + \varphi(\mathbf{0})$.
- $L_\beta = D_\beta$.
- M_β : $\min_\theta\{\alpha\theta | \sum_i(\mathbf{b} - A\mathbf{x}_i)\theta_i = \mathbf{0}, \sum_i \theta_i = 1, \theta \geq 0\} + \varphi(\bar{\mathbf{y}})$.
- M_β^ϵ : $min_\theta\{\alpha\theta | \frac{1}{2}||G_\beta\theta||^2 \leq \epsilon\}$.