Bundle Methods
(Leave no Primals)

Ashutosh Mahajan

Department of Industrial and Systems Engineering
Lehigh University

COR@L Seminar Series, Spring 2006
References:


Subgradient Optimization

\[
(D) \quad \max_{y} \{ \varphi(y) : y \in Y \}
\]

\( \varphi() : \mathbb{R}^m \to \mathbb{R} \cup [-\infty], \)
continuous, non-differentiable concave function

\( Y \subseteq \mathbb{R}^m \)

(Draw figure)
Subgradient Optimization

- Gradient $\nabla \varphi(\bar{y})$
- Subdifferential $\partial \varphi(\bar{y})$.
- Subgradient $g \in \partial \varphi(\bar{y})$.
- $y^* \in \{y : \varphi(y) \geq \varphi(\bar{y})\} \subseteq$ half space pointed by $g$.
- Directional derivative $\varphi'(\bar{y}, d)$

(Draw figure)
Subgradient Optimization: Basic Results

- \( \varphi() \) differentiable \( \Rightarrow \varphi'(\bar{y}, d) = \nabla\varphi(\bar{y}).d. \)
- \( \varphi'(\bar{y}, d) = \min_v \{ v.d : v \in \partial\varphi(\bar{y}) \}. \)
- \( \nabla\varphi(\bar{y}) = \arg\max_d \{ \varphi'(\bar{y}, d) : ||d|| = 1 \}. \)

\( \arg\min_g \{ ||g|| : g \in \partial\varphi(\bar{y}) \} = \arg\max_d \{ \varphi'(\bar{y}, d) : ||d|| = 1 \}. \)

Optimality results follow.

Where is the problem?
Subgradient Optimization: $\varepsilon$-optimality

- $\varepsilon$-subgradient: $\varphi(y) \leq \varphi(\bar{y}) + g.(\bar{y} - y) + \varepsilon$.
- $\varepsilon$-ascent direction
- $d$ is of $\varepsilon$-ascent iif $\varphi'(\bar{y}, d) > 0$.

Exactly same results follow.
How does this help?
Subgradient Optimization: $\epsilon$-optimality

\[ \mathbf{g} \in \partial \varphi(\bar{\mathbf{y}}) \Rightarrow \mathbf{g} \in \partial \varphi_{\alpha}(\bar{\mathbf{y}}) \]

For all $\mathbf{y}$ and

\[ \alpha \geq \varphi(\bar{\mathbf{y}}) + \mathbf{g} \cdot (\mathbf{y} - \bar{\mathbf{y}}) - \varphi(\mathbf{y}) + \epsilon \]
Put up in a place where it’s easy to see the cryptic admonishment
T. T. T.
When you feel how depressingly slowly you climb, it’s well to remember that Things Take Time.
– Piet Hein
Bundle Methods: motivation

- Cutting Plane Method.
- $\varphi_{CP}(y) = \min_i \{ \varphi(y_i) + g(y_i) \cdot (y - y_i : i \in \beta) \}$
- In terms of direction $d$.
- Stabilization.
- Dual and the QP.
- $\epsilon$ revisited.
- *Smoothing.*