Bundle Methods
(Leave no Primals)

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..

Subgradient Optimization

$$(D) \qquad \max_{y}\{\varphi(\mathbf{y}) : \mathbf{y} \in Y\}$$

$\varphi() : \mathbb{R}^m \to \mathbb{R} \cup [-\infty]$,

continuous, non-differentiable concave function

$Y \subseteq \mathbb{R}^m$

(Draw figure)

Subgradient Optimization

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

(Draw figure)

Subgradient Optimization: Basic Results

- $\varphi()$ differentiable $\Rightarrow \varphi'(\bar{\mathbf{y}}, \mathbf{d}) = \nabla\varphi(\bar{\mathbf{y}}).\mathbf{d}$.
- $\varphi'(\bar{\mathbf{y}}, \mathbf{d}) = \min_v\{\mathbf{v}.\mathbf{d} : \mathbf{v} \in \partial\varphi(\bar{\mathbf{y}})\}$.
- $\nabla\varphi(\bar{\mathbf{y}}) = \text{argmax}_d\{\varphi'(\bar{\mathbf{y}}, \mathbf{d}) : ||d|| = 1\}$.

$\text{argmin}_g\{||\mathbf{g}|| : \mathbf{g} \in \partial\varphi(\bar{\mathbf{y}})\} = \text{argmax}_d\{\varphi'(\bar{\mathbf{y}}, \mathbf{d}) : ||d|| = 1\}$.
Optimality results follow.
Where is the problem?

Subgradient Optimization: ϵ-optimality

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

Exactly same results follow.
How does this help?

Subgradient Optimization: ϵ-optimality

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

$\forall \mathbf{y}$ and

$\alpha \geq \varphi(\bar{\mathbf{y}}) + \mathbf{g}.(\mathbf{y} - \bar{\mathbf{y}}) - \varphi(\mathbf{y}) + \epsilon$

T. T. T.

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.
- $\varphi_{CP}(\mathbf{y}) = \min_i \{\varphi(\mathbf{y}_i) + \mathbf{g}(\mathbf{y}_i).(\mathbf{y} - \mathbf{y}_i : i \in \beta\}$
- In terms of direction **d**.
- Stabilization.
- Dual and the QP.
- ϵ revisited.
- *Smoothing*.