Overview

We consider problem of the form,
\[ \text{minimize } f(x), \quad \text{subject to } x \in \Omega \subseteq \mathbb{Z}^n. \quad (P) \]
\(\Omega\) bounded, \(f\) convex on \(\Omega\), \(f\) unvarational if \(x \notin \mathbb{Z}^n\).
- Evaluation of \(f\) is expensive, derivatives are unavailable
- Applications: optimal design of concentrating solar power plants, performance tuning codes for HPC etc.
- Any surrogate model of continuous relaxation can not be solved, prohibiting branch-and-bound.
- IP techniques like Bender’s decomposition, Outer Approximation etc. can not be used due to unavailability of \(\nabla f\).
- DFO: global optimal not guaranteed (even for convex \(f\)).

We provide a method for solving \(P\) to global optimality under the stated assumptions on \(f\) and \(\Omega\).

The underestimation lemma

Given a set of points \(X\) satisfying \(|X| \geq n+1\) and their function values, we define a \textit{secant linear map}, using any \(n+1\) affinely independent points, \(x^i = (i_1, \ldots, i_n)\), by solving
\[
\begin{aligned}
\begin{bmatrix}
\nu^i
\end{bmatrix} =
\begin{bmatrix}
x^i
\end{bmatrix}
x^i = \begin{bmatrix}
x^i
\end{bmatrix} + 
\begin{bmatrix}
u^i
\end{bmatrix}
\end{aligned}
\]
where
\[
\begin{aligned}
X^i \triangleq \begin{bmatrix}
x^i
\end{bmatrix}^T,
\nu^i \triangleq \begin{bmatrix}
1
\end{bmatrix},
\text{ and } f^i \triangleq f^{\nu^i}. 
\end{aligned}
\]

\textbf{Lemma 1}: Let \(f\) be convex and \(X\) be poised. The unique linear map \(m^i(x) \triangleq \langle \nu^i \rangle x + b^i\) satisfying \(m^i(x^i) = f^i\), \(\nu^i \in \mathbb{R}^n\), \(f(x) \geq m^i(x)\), \(\forall x \in \mathbb{R}^n\) is bounded, then
\[ f(x) \geq f^i = \min_{x \in \{x^i\}} f(x). \]

\textbf{Corollary 1}: The linear mapping \(m^i(x)\) satisfies
\[ f(x) \geq f^i \quad \text{for } x \text{ such that } f(x) \geq f^i. \]

Pictorial demonstration

An MILP formulation for \((PLM)\)

- Define \(\eta = \max_{i} \text{ maximum of piecewise linear secant functions:} \)
  \[ \eta \geq \langle \nu^i \rangle x + b^i - M_0(1 - \sum_{j=1}^{n} \chi_j), \quad \forall i \in W(X), \quad (PLM) \]
- Define \(\nu^i = 1 \quad \text{if } x \in \text{ cone}(X \setminus -x^i), \quad \forall i \in W(X), \forall j \in i. \)

- For a given poised set \(X\), the \(n+1\) cones are disjoint:
  \[ \sum_{j=1}^{n} \chi_j \leq 1, \quad \forall i \in W(X), \quad (4) \]

- A lower bound on \(\lambda\) variables:
  \[ \chi_j \geq \lambda_j \geq -M_0(1 - \chi_j), \quad \forall i \in W(X), \forall j, i, j \neq l \]

- Define \(w^i = 0 \quad \text{if } \chi_j < 0 \quad \text{for } j \in i \quad \text{and } j \neq l \)

- At least one of the \(w^i\) be \(0 \quad \text{if corresponding } \chi_j = 0 \quad \text{for } j \in i \quad \text{and } j \neq l \)

Finally, the full MILP model is given as:
\[ \text{minimize } \eta \quad \text{subject to } (2) - (7) \]
\[ w^i, z^i \in \{0, 1\}, \forall j, i, j \in \{1, \ldots, n+1\}, \forall j \neq l, \forall i \quad \text{and } x \in \mathbb{Z}^n. \]

Derivation of model parameters

- If \(f\) is a valid lower bound on \(f\) over \(\Omega\) :
  \[ M_0 = \max_{i \in W(X)} \langle \nu^i \rangle x + b^i \]
- Error bound:
  \[ f(x) - f(x^i) \leq \eta \]

Ongoing interesting stuff

- Selecting “useful” subsets of \(W(X) \times \{x^i\} [\text{bottleneck}]
- Possible extension of \(\text{“cone”} \text{underestimation lemma}
- Algorithmic refinements: surrogate models, trust region etc.
- Understanding “(minimal cardinality)” optimal sets

References