Pure Adaptive Search In Global Optimization

Z.Zabinsky & R.Smith

Michael Yee
Kwong-Meng Teo

This presentation is based on: Zabinsky, Zelda B., and Robert L. Smith. *Pure Adaptive Search in Global Optimization. Mathematical Programming* 55, 1992, pp. 323-338.

1 Outline

SLIDE 1

- Extension to Discrete Optimization
  Pure Adaptive Search for Finite Global Optimization
  Z.Zabinsky et al., Math. Programming 69 (1995)

- Summary of other results leading from PAS
- Further comments

2 Pure Adaptive Search

2.1 Discrete Case

SLIDE 2

Consider:

$$\begin{array}{ll} min_x & f(x) \\ s.t. & x \in S \end{array}$$

where $f(x) \in R$ and S is a finite set

- Strong PAS: domain with strictly improving cost: $S_k = \{x : x \in S, f(x) < f(x_k)\}$

- Weak PAS: domain with equal or improving cost: $S_k = \{x : x \in S, f(x) \leq f(x_k)\}$

SLIDE 3

Define:

- $y_* = y_1 < y_2 < \cdots < y_K = y^*$ are all possible distinct objective values attained by $x \in S$
- $\pi_j = P(f(x) = y_j)$, x is random sample from S
- $p_j = \sum_{i=1}^{j} \pi_i$

Given $x_m, f(x_m) = y_k$, assume:

SLIDE 4

- Strong PAS:
  $P(f(x_{m+1}) = y_j) = \begin{cases} \pi_j / p_{k-1}, & j < k; \\ 0, & \text{o.w.} \end{cases}$
- Weak PAS:
  $P(f(x_{m+1}) = y_j) = \begin{cases} \pi_j / p_k, & j \leq k; \\ 0, & \text{o.w.} \end{cases}$

SLIDE 5

Theorem: The expected number of iterations to solve the finite optimization problem is:
(i) $1 + \sum_{j=2}^{K} \pi_j / p_j$ for strong PAS and
(ii) $1 + \sum_{j=2}^{K} \pi_j / p_{j-1}$ for weak PAS

Proof: Model stochastic process $\{W_m = f(x_m) | m = 0, 1, \ldots\}$ as a Markov chain with states $y_1, \ldots, y_K$, and π_i, p_i define the transition probabilities. Given initial probability distribution of $W_0 = \underline{\pi}$, derive expected number of transitions to converge to absorption state y_1.

SLIDE 6

Corollary: The expected number of strong PAS iterations to solve the finite optimization problem is bounded above by $1 + log(\frac{1}{\pi_1})$

Proof:
$0 < x < 1 \;\Rightarrow\; x < -log(1 - x)$
Therefore, $\pi_j/p_j < -log(1 - \pi_j/p_j) = log(p_j/p_{j-1})$
$\forall j = 2, \ldots, K$ SLIDE 7
Corollary: The expected number of iterations for finite global optimization, given a uniform distribution on the objective function values, is
(i) $\sum_{j=1}^{K} \frac{1}{j}$, bounded above by $1 + logK$ for strong PAS and
(ii) $1 + \sum_{j=1}^{K-1} \frac{1}{j}$, bounded above by $2 + log(K - 1)$ for weak PAS

Comparison to continuous case:
Consider $S = \{$vertices of n-dim lattice $\{1, \ldots, k\}^n\}$, each with unique objective functions.
Expected number of iterations is bounded by $2 + log(K) = 2 + nlog(k)$ since $K = k^n$.
Linear in n.

3 Summary of PAS results

SLIDE 8

- Polynomial time implementation of PAS for LP
  Linear Optimization in Random Polynomial Time
  A.Gademann, PhD Thesis, (1993)

- That there exists a polynomial time implementation for the PAS algorithm for most convex programming problems
  Implementing PAS for Global Optimization using Markov Chain Sampling
  D.Reaume, H.Romeijn & R.Smith, Journal of Global Optimization 20, (2001)

4 Further Comments

SLIDE 9

Iteration k:
Step 1 : Start with x_k
Step 2 : Obtain sample of $x_{k+1} \sim U(S_k)$
Step 3 : If stop criterion met, stop, else start $k + 1$

where

$S_k = \{x : x \in S$ and $f(x) < x_k\}$

5 Further Comments

5.1 Generalization

SLIDE 10

Iteration k:
Step 1 : Start with x_k
Step 2 : Obtain x'_{k+1} $s.t$ $E[x'_{k+1}] \leq E[x_{k+1}]$
Step 3 : Let $x_{k+1} := x'_{k+1}$
Step 4 : If stop criterion met, stop, else start $k+1$

$x_k \sim U(S_k), \; S_k = \{x : x \in S \text{ and } f(x) < x_k\}$

6 Further Comments

6.1 wrt Interior Point Algo.

SLIDE 11

Iteration k:
Step 1 : Start with x_k
Step 2 : Obtain x'_{k+1} $s.t$ $E[x'_{k+1}] \leq E[x_{k+1}]$
Step 3 : Let $x_{k+1} := x'_{k+1}$
Step 4 : If stop criterion met, stop, else start $k+1$

$x_k \sim U(S_k), \; S_k = \{x : x \in S \text{ and } f(x) < x_k\}$

Possible approach to explain empirical observations of number of iterations required by Interior Point Method ?

7 Further Comments

SLIDE 12

Iteration k:
Step 1 : Start with x_k
Step 2 : Obtain sample of $x_{k+1} \sim U(S_k)$
Step 3 : If stop criterion met, stop, else start $k+1$

If we use information from x_k to find x_{k+1}, will we be limited to finding local optimum only ?

- upper bound of minimum cost
- feasible direction and starting point
- moments at x_k

8 Final Comment !

SLIDE 13

Iteration k:
Step 1 : Start with x_k
Step 2 : Obtain sample of $x_{k+1} \sim U(S_k)$
Step 3 : If stop criterion met, stop, else start $k+1$

If we can find x_{k+1} without local information from x_k, is it equivalent to finding a feasible point, if possible, of arbitrary objective function value ?

Consider finding $x_{k+1} \sim U(S'_k)$ where
$S'_k = \{x : x \in S,\ f(x) > x_k - \epsilon,\ f(x) < x_k + \epsilon\}$