(Version: 2021/3/30)

区間演算による大域最適化

柏木 雅英

1　はじめに

関数 $f:\mathbb{R}^n \to \mathbb{R}$ と、区間ベクトル $I_{\mathrm{init}} \subset \mathbb{R}^n$ に対して、関数 f の最小値を与える点の包含や、最小値の包含を区間演算で求めることを考える。

例えば f に凸性があるなどの条件があれば、局所最適解がそのまま大域最適解となるので効率的なアルゴリズムが考えられる。が、ここでは局所最適解がいくつもあるような場合の、大域最適化を考える。

f は区間演算さえ出来れば本質的には他に仮定は要らないが、ここでは f は微分可能で f' に対する区間演算も可能であることを仮定している。

2　アルゴリズム

f の区間 $I_{\mathrm{init}} \subset \mathbb{R}^n$ における最小値を達成し得るような区間のリストを得ることを考える。区間 I の第 i 成分を I_i 、区間 I における f の傾き $f'(I)$ の第 i 成分を $f'(I)_i$ と書く。

擬似コードで書くと、次の通り。

function globalopt(f, I_{init}, th) {

　　$L = \{I_{\mathrm{init}}\}$　# 調査対象の区間のリスト

　　$S = \{\}$　# 最小値を含む可能性のある区間のリスト

　　$\delta = \infty$　# 最小値の上限

　　while ($L \neq \emptyset$) {

　　　　$I = L.\mathrm{pop}()$　# L からひとつの要素を取り出す

　　　　if ($f(I) > \delta$) continue　# I に最小値は存在しない

　　　　$c = \mathrm{mid}(I)$

　　　　if ($f(c) + f'(I)(I-c) > \delta$) continue　# I に最小値は存在しない

　　　　($\star$1)

　　　　$\delta = \min(\delta, \overline{f(c)})$　# δ を更新

　　　　$b_i = \begin{cases} \underline{I_i} & (f'(I)_i < 0) \\ \overline{I_i} & (f'(I)_i > 0) \\ \mathrm{mid}(I_i) & (other) \end{cases}$　# b は値が小さくなりそうな点

　　　　$\delta = \min(\delta, \overline{f(b)})$　# δ を更新

(⋆2)
if (width(I) $< th$) {
    S.append(I)   # 解のリストに追加
    continue
}
$I_1, I_2 =$ divide(I)   # I を 2 分割
L.append(I_1)
L.append(I_2)
}
return S
}

3 傾き情報を利用した最小値の非存在判定

更に、アルゴリズムの (⋆1) の部分で次のような非存在判定を行っている。
すべての成分 $1 \leq i \leq n$ について、

- $f'(I)_i > 0$ かつ $\underline{I_i} \neq \underline{I_{\mathrm{init}}}$ ならば、I に最小値は存在しない。
- $f'(I)_i < 0$ かつ $\overline{I_i} \neq \overline{I_{\mathrm{init}}}$ ならば、I に最小値は存在しない。

すなわち、ある区間において関数がある成分について左下がりで区間の左端が I_{init} の左端でないならば、(左隣の区間内で関数値がより小さくなるので) 区間内に最小値は存在しないということを利用している。右隣についても同様。

4 区間縮小

アルゴリズムの (⋆2) の部分に次のような区間の縮小操作を挿入することもできる。
f の平均値形式を考えると、$c = \mathrm{mid}(I)$ として

$$f(c) + \sum_{k=1}^{n} f'(I)_k (x_k - c_k)$$

となるが、この x_k をある x_j だけ残し残りを区間 I_k で置き換えると、

$$f(c) + \sum_{j=1, j\neq k}^{n} f'(I)_k (I_k - c_k) + f'(I)_j (x_j - c_j)$$

となる。

$$J_j = f(c) + \sum_{j=1, j\neq k}^{n} f'(I)_k (I_k - c_k)$$

とおくと、

$$J_j + f'(I)_j (x_j - c_j)$$

と書ける。この値が δ より小さい領域にしか最小値は存在しないので、

$$J_j + f'(I)_j (x_j - c_j) = [-\infty, \delta]$$

という x_j に関する方程式を考え、x_j について解く。この区間解と I の共通部分にしか最小値は存在しないので、その領域に I を絞り込むことが出来る。

が、あまりいい性能が出ないようなので、現状この縮小法は採用していない。