

区分的微分可能関数と一般化自動微分

Piecewise Smooth Function and Generalized Algorithmic Differentiation

情報工学専攻 15N8100018J 吉田 北斗

概要: 区分的微分可能関数と呼ばれる微分不可能点を含む関数に対して、自動微分を用いた最適化を行う。区分的微分可能関数はその構成要素や特徴ごとにレベル付けすることができ、本研究ではその中でも絶対値関数により表現されるものを対象とする。対象とする関数は、微分不可能点の周辺で A. Griewank らにより提唱されている Abs-Normal Form により一般化勾配を導出することができる。これを利用して従来手法の拡張を行い、その収束の安定性を評価をする。

キーワード: 区分的微分可能, 区分線形近似, 一般化勾配, 自動微分

1 はじめに

非線形最適化問題を解く種々の反復解法は、その目的関数や制約式の微分情報に依存していることが多い。しかし、工学問題ではしばしば区分的微分可能 (Piecewise Smooth, Piecewise Differentiable) なモデルを用いることがあり、そういったモデルを表す目的関数は至るところで微分不可能となっているため、従来の手法では解軌道がジグザグになるなど収束が不規則になる、または収束しないといったことが起こる。このような挙動を回避する、区分的微分可能なモデルに対する安定した手法が必要である。

本研究では、区分的微分可能なモデルの中でも、絶対値関数によって表現されるものに対する微分情報を用いた最適化手法の提案とその収束の安定性の評価を行う。

2 自動微分 [1]

自動微分 (Algorithmic Differentiation) とは、関数値を計算するある数値アルゴリズム α_f が与えられたとき、その偏導関数値を計算する数値アルゴリズム α_g を導出する算法である。自動微分による微分係数の導出は合成関数の微分則である連鎖律 (chain rule) に従う。

2.1 基本演算と要素的偏導関数

基本演算 (Basic Operation, Elementary Function) φ とは、プログラム中で扱うことのできる以下のような演算 (関数) のことである。

- 四則演算 (+, -, *, /)
- べき乗 (pow)
- 開平 (sqrt)
- 初等超越関数 (log, exp, sin, ...)

2.2 計算過程と中間変数

入力を定めてアルゴリズムを実行すると、アルゴリズム中の基本演算と引数の関係を表す計算の履歴が定められる。この履歴を計算過程 (Computational Process) と呼ぶ。

中間変数 (Intermediate Variables) とは、基本演算と 1 対 1 に対応 (単一代入) し、その基本演算の実行結果を格納する変数のことである。 $\mathbf{y} = f(\mathbf{x})$ ($f: \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$) を自動微分により実行したときの計算過程は表 1 のようになる。ここで、 \mathbf{x} は入力変数 (Input Variables), \mathbf{y} は出力変数 (Output Variables) と呼ばれる。

表 1 計算過程

v_{i-n}	$=$	x_i	$i = 1, \dots, n$
v_i	$=$	$\varphi_i(v_j)_{j < i}$	$i = 1, \dots, l$
y_{m-i}	$=$	v_{l-i}	$i = 0, \dots, m-1$

基本演算 φ_i に対応する中間変数が v_i である。 φ_i の添字の i は i 番目に行われた基本演算であることを表している。

計算過程に現れる任意の二つの変数 u, v の間に “ u の値を計算するには、 v の値が計算されていることが必要” という関係があるとき、二つの変数には半順序関係があり、このとき $u \prec v$ と表記する。

3 区分的微分可能関数

3.1 区分的微分可能関数 [2]

開集合 $\mathcal{D} \subset \mathbb{R}^n$ が与えられたとき、 $f: \mathcal{D} \rightarrow \mathbb{R}^m$ が \mathcal{N} 上で連続する点 \mathbf{x} の開近傍 $\mathcal{N} \subset \mathcal{D}$ と、 C^1 級関数 $f_i: \mathcal{N} \rightarrow \mathbb{R}^m$ ($i = 1, \dots, k$) が存在するとき、 f_i を f の点 \mathbf{x} 周辺での selection function と呼ぶ。

式 (3.1) を満たすとき、 f は点 $\mathbf{x} \in \mathcal{D}$ で区分的微分可能であるという。

$$f(\mathbf{x}) \in \{f_1(\mathbf{x}), \dots, f_k(\mathbf{x})\} \quad \text{for } \mathbf{x} \in \mathcal{N} \quad (3.1)$$

さらに、このとき selection function f_i がすべて線形であるならば、 f は区分的線形関数と呼ばれる。

3.2 局所リプシッツ連続 [2, 3, 4]

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ が点 $\mathbf{x} \in \mathbb{R}^n$ で局所リプシッツ連続 (Locally Lipschitz Continuous) であるとは、ある $K > 0, \epsilon > 0$ ($K, \epsilon \in \mathbb{R}$) が存在して、 \mathbf{x} の ϵ 近傍 $B(\mathbf{x}; \epsilon)$ に対して式 (3.2) が成り立つことをいう。

$$|f(\mathbf{x}'') - f(\mathbf{x}')| \leq K|\mathbf{x}'' - \mathbf{x}'| \quad \text{for all } \mathbf{x}'', \mathbf{x}' \in B(\mathbf{x}; \epsilon) \quad (3.2)$$

$f: \mathbb{R}^n \rightarrow \mathbb{R}$ が $\mathcal{D} \subseteq \mathbb{R}^n$ 上で局所リプシッツ連続であるとは、 \mathcal{D} に属するすべての点で式 (3.2) が成り立つことをいう。

区分的微分可能関数は局所リプシッツ連続であることが知られている。

3.3 区分的微分可能関数の構成要素

区分的微分可能関数とその構成要素からいくつかのレベルに分けることにする。

レベル 0 (Smooth)

定義域内に微分不可能点を含まない滑らかな関数。

レベル 1 (Piecewise Smooth)

局所リプシッツ連続であるが、定義域内に微分不可能点を含む関数。その微分不可能点は abs, min, max 関数によって表される。

レベル 2 (Piecewise Semismooth)

局所リプシッツ連続であるが、定義域内に微分不可能点を含む関数。その微分不可能点はレベル 1 にユークリッドノルムを加えたものによって表される。

レベル 3 (Discontinuous Piecewise Smooth)

非連続な関数。その微分不可能点はレベル 2 に sign 関数と条件分岐を加えたものによって表される。

このレベルの順に、その関数の微分不可能周辺での扱いは困難にある。本研究では、レベル 1 の区分的微分可能関数を対象としている。

3.4 区分線形関数の min/max による書き換え

ある区分的線形関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ と、 f と区分的に一致する線形関数 $f_1(\mathbf{x}) = \mathbf{a}_1^T \mathbf{x} + b_1, \dots, f_k(\mathbf{x}) = \mathbf{a}_k^T \mathbf{x} + b_k$ に対して、以下を満たす有限な添字集合 $M_1, \dots, M_l \subseteq \{1, \dots, k\}$ が存在することが知られている [2].

$$f(\mathbf{x}) = \max_{1 \leq i \leq l} \min_{j \in M_i} \mathbf{a}_j^T \mathbf{x} + b_j \quad (3.3)$$

すなわち、あらゆる区分的線形関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ は min/max で表現することができる。

3.5 min/max の絶対値関数による書き換え

\mathbb{R} 上で、min/max は絶対値関数を用いて以下のように書き換えられることが知られている。

$$\min(u, w) = \frac{1}{2}(u + w - |u - w|), \quad (3.4)$$

$$\max(u, w) = \frac{1}{2}(u + w + |u - w|) \quad (3.5)$$

4 Abs-Normal Form [5, 6, 7]

Abs-Normal Form とは A. Griewank らにより提唱されている区分的微分可能関数を自動微分で扱うための手法の一つである。

4.1 Increment Function

各基本演算の増分の近似を以下のように考える。

$$\begin{aligned} \Delta v_i &= \Delta v_j \pm \Delta v_k & \text{when } v_i &= v_j \pm v_k \\ \Delta v_i &= v_j * \Delta v_k + v_k * \Delta v_j & \text{when } v_i &= v_j * v_k \\ \Delta v_i &= \varphi'_i(v_j)_{j < i} * \Delta(v_j)_{j < i} & \text{when } v_i &= \varphi_i(v_j)_{j < i} \neq |v_j| \\ \Delta v_i &= |v_j + \Delta v_j| - v_i & \text{when } v_i &= |v_j| \end{aligned} \quad (4.1)$$

このとき、目的関数 f の点 \mathbf{x} での方向 $\Delta \mathbf{x}$ への増分の近似を

$$\Delta f(\mathbf{x}; \Delta \mathbf{x}) \quad (4.2)$$

と表し、increment function と呼ぶ。式 (4.2) が計算可能であるとすると、 f の区分線形近似を

$$f_{\text{PL}, \mathbf{x}}(\Delta \mathbf{x}) \equiv f(\mathbf{x}) + \Delta f(\mathbf{x}; \Delta \mathbf{x}) \quad (4.3)$$

と表すことができる。

4.2 Abs-Normal Form

絶対値関数を含む方程式 $y = f(\mathbf{x})$ ($f: \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$) を、自動微分によって実行したときの計算過程を表 2 のように、絶対値関数の実行結果である中間変数 w_i , w_i の引数となる中間変数 z_i に分けて考える。

表 2 絶対値関数を特殊化した計算過程

w_{i-n}	$= x_i$	$i = 1, \dots, n$
z_i	$= \varphi_i(w_j)_{j < i}$	
σ_i	$= \text{sign}(z_i)$	$i = 1, \dots, s$
w_i	$= \sigma_i z_i = z_i $	
y	$= \varphi_{s+1}(w_j)_{j < s+1}$	

表 2 中の関数 sign は

$$\text{sign}(z) = \begin{cases} -1 & (z < 0) \\ 0 & (z = 0) \\ 1 & (z > 0) \end{cases}$$

で定義される。また、 φ_i は絶対値関数以外の基本演算、 s は f に含まれる絶対値関数の個数である。

絶対値関数の引数となる中間変数 z_i を、対応する φ_i が実行された順番に並べたベクトル \mathbf{z} を switching vector と呼ぶ。

$$\mathbf{z} \equiv \mathbf{z}(\mathbf{x}) \in \mathbb{R}^s$$

switching vector \mathbf{z} が 0 を含むような \mathbf{x} の集合 Ω は関数 f の微分不可能点の集合である。

$$\Omega \equiv \{\mathbf{x} \mid 0 \in \mathbf{z}(\mathbf{x}), \mathbf{x} \in \mathcal{D}\}$$

また、switching vector の符号を表すベクトル $\boldsymbol{\sigma}$ を signature vector と呼ぶ。

$$\boldsymbol{\sigma} \equiv \boldsymbol{\sigma}(\mathbf{x}) \equiv \text{sign}(\mathbf{z}(\mathbf{x})) \in \{-1, 0, 1\}^s$$

中間変数 \mathbf{w} を入力変数と考えれば、表 2 より

$$\mathbf{z} \equiv \Phi(\mathbf{x}, \mathbf{w}) \equiv \Phi(\mathbf{x}, |\mathbf{z}|) \quad (\Phi: \mathbb{R}^{n+s} \rightarrow \mathbb{R}^s), \quad (4.4)$$

$$f(\mathbf{x}) \equiv \tilde{f}(\mathbf{x}, \mathbf{w}) \equiv \tilde{f}(\mathbf{x}, |\mathbf{z}|) \quad (\tilde{f}: \mathbb{R}^{n+s} \rightarrow \mathbb{R}) \quad (4.5)$$

となる \tilde{f} により

$$y \equiv \tilde{f}(\mathbf{x}, |\mathbf{z}|) \quad (4.6)$$

と定義することができる。このような方程式の再解釈を Abs-Normal Form と呼ぶ。

4.3 Piecewise Linearization

式 (4.4), 式 (4.6) より、 \mathbf{z} , y の増分を以下のように定義する。

$$\Delta \mathbf{z} = \frac{\partial \Phi}{\partial \mathbf{x}} \Delta \mathbf{x} + \frac{\partial \Phi}{\partial |\mathbf{z}|} \Delta |\mathbf{z}|,$$

$$\Delta y = \frac{\partial \tilde{f}}{\partial \mathbf{x}} \Delta \mathbf{x} + \frac{\partial \tilde{f}}{\partial |\mathbf{z}|} \Delta |\mathbf{z}|$$

さらに、 \mathbf{x} を $\mathbf{z} = \mathbf{0}$ であるような基準点と考えると、式 (4.1) より

$$\Delta \mathbf{z}(\mathbf{x}; \Delta \mathbf{x}) = \Phi(\mathbf{x}, \mathbf{0}) + \frac{\partial \Phi}{\partial \mathbf{x}} \Delta \mathbf{x} + \frac{\partial \Phi}{\partial |\mathbf{z}|} (|\mathbf{z} + \Delta \mathbf{z}| - |\mathbf{z}|),$$

$$\Delta y(\mathbf{x}; \Delta \mathbf{x}) = \tilde{f}(\mathbf{x}, \mathbf{0}) + \frac{\partial \tilde{f}}{\partial \mathbf{x}} \Delta \mathbf{x} + \frac{\partial \tilde{f}}{\partial |\mathbf{z}|} (|\mathbf{z} + \Delta \mathbf{z}| - |\mathbf{z}|) \quad (4.7)$$

$\Sigma \equiv \text{diag}(\boldsymbol{\sigma}(\Delta \mathbf{x}))$ として、式 (4.7) を行列の形でまとめると以下のようなになる。

$$\begin{aligned} \begin{bmatrix} \Delta \mathbf{z} \\ \Delta y \end{bmatrix} &= \begin{bmatrix} \Phi(\mathbf{x}, \mathbf{0}) \\ \tilde{f}(\mathbf{x}, \mathbf{0}) \end{bmatrix} + \begin{bmatrix} \frac{\partial \Phi}{\partial \mathbf{x}} & \frac{\partial \Phi}{\partial |\mathbf{z}|} \\ \frac{\partial \tilde{f}}{\partial \mathbf{x}} & \frac{\partial \tilde{f}}{\partial |\mathbf{z}|} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ |\mathbf{z} + \Delta \mathbf{z}| - |\mathbf{z}| \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{c}_z \\ \mathbf{c}_y \end{bmatrix} + \begin{bmatrix} Z & L \\ \mathbf{a}^T & \mathbf{b}^T \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x} \\ \Sigma \Delta \mathbf{z} \end{bmatrix} \end{aligned} \quad (4.8)$$

$$\mathbf{c}_z \equiv \Phi(\mathbf{x}, \mathbf{0}), \quad \mathbf{c}_y \equiv \tilde{f}(\mathbf{x}, \mathbf{0}),$$

$$Z \in \mathbb{R}^{s \times n}, \quad L \in \mathbb{R}^{s \times s}, \quad \mathbf{a} \in \mathbb{R}^n, \quad \mathbf{b} \in \mathbb{R}^s$$

以降では、この式 (4.8) を指して Abs-Normal Form と呼ぶことにする。

4.4 Selection Function

式 (4.8) を $\Delta \mathbf{x}$ について整理する.

$$\Delta \mathbf{z} = (I - L\Sigma)^{-1}(\mathbf{c}_z + Z\Delta \mathbf{x}),$$

$$\Delta y = c_y + \mathbf{b}^T \Sigma (I - L\Sigma)^{-1} \mathbf{c}_z + (\mathbf{a}^T + \mathbf{b}^T \Sigma (I - L\Sigma)^{-1} Z) \Delta \mathbf{x}$$

区分線形近似 $f_{\text{PL}, \mathbf{x}}$ を以下で定義する.

$$f_{\text{PL}, \mathbf{x}}(\Delta \mathbf{x}) \equiv \Delta y$$

ある σ を定めたときの区分線形近似 $f_{\text{PL}, \mathbf{x}}$ を $f_{\sigma, \mathbf{x}}$ と表し, これを selection function と呼ぶ.

$$f_{\sigma, \mathbf{x}}(\Delta \mathbf{x}) \equiv \gamma_\sigma + \mathbf{g}_\sigma^T \Delta \mathbf{x} \quad (4.9)$$

式 (4.9) の $\gamma_\sigma, \mathbf{g}_\sigma$ はそれぞれ

$$\gamma_\sigma = c_y + \mathbf{b}^T \Sigma (I - L\Sigma)^{-1} \mathbf{c}_z, \quad \mathbf{g}_\sigma^T = \mathbf{a}^T + \mathbf{b}^T \Sigma (I - L\Sigma)^{-1} Z$$

である. また, 定義より明らかに

$$\nabla f_{\sigma, \mathbf{x}} = \mathbf{g}_\sigma$$

である.

$f_{\sigma, \mathbf{x}}$ は $\sigma \equiv \sigma(\Delta \mathbf{x}) \in \{-1, 0, 1\}^s$ によって分割された $f_{\text{PL}, \mathbf{x}}$ のある多面体領域を表す.

式 (4.10) を例として, 点 (1, 1) を基準に分割された多面体領域の図を載せる.

$$f(x_1, x_2) = (x_2^2 - (x_1)_+)_+ \quad \text{with} \quad x_+ \equiv \max(x, 0) \quad (4.10)$$

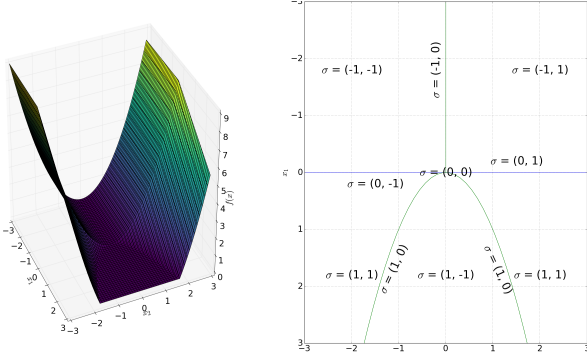


図1 式 (4.10) (左) と $\sigma(\mathbf{x})$ による領域分割 (右)

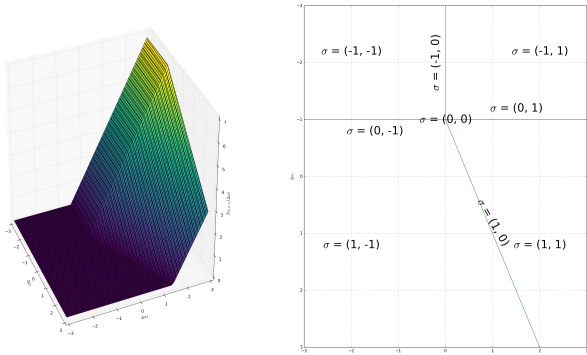


図2 $f_{\text{PL}, \mathbf{x}}(\Delta \mathbf{x})$ (左) と $\sigma(\Delta \mathbf{x})$ による領域分割 (右)

以下では, σ によって決まる多面体領域を

$$P_\sigma = \{\Delta \mathbf{x} : \sigma(\Delta \mathbf{x}) = \sigma\}$$

と表すことにする. σ には以下のような半順序関係を定義できる.

$$\sigma \preceq \tilde{\sigma} \iff \sigma_i^2 \leq \tilde{\sigma}_i \sigma_i \quad (i = 1, \dots, s) \quad (4.11)$$

式 (4.11) で等号が成り立つとき P_σ と $P_{\tilde{\sigma}}$ は同一であり, そ

他のとき P_σ と $P_{\tilde{\sigma}}$ は隣接している. また, この多面体領域の中でも P_σ ($\sigma \in \{-1, 1\}^s$) は開集合であり, その閉集合は

$$\bar{P}_\sigma = \text{cl } P_\sigma = P_\sigma \cup \{\sigma' \mid \sigma' \prec \sigma, \sigma' \in \{-1, 0, 1\}^s \setminus \{-1, 1\}^s\}$$

となる.

表3 P_σ が表す領域 (図2)

σ	P_σ	$\preceq \tilde{\sigma}$
(0, 0)	$\Delta x_1 = -1, \Delta x_2 = 0$	$\forall \{-1, 0, 1\}^2$
(0, 1)	$\Delta x_1 = -1, \Delta x_2 > 0$	(0, 0), (0, 1), ($\pm 1, 1$)
(1, 0)	$\Delta x_1 > -1, \Delta x_2 = \frac{\Delta x_1 + 1}{2}$	(0, 0), (1, 0), ($\pm 1, \pm 1$)
(0, -1)	$\Delta x_1 = -1, \Delta x_2 < 0$	(0, 0), (0, -1), ($\pm 1, -1$)
(-1, 0)	$\Delta x_1 < -1, \Delta x_2 = 0$	(0, 0), (-1, 0), (-1, ± 1)
(1, 1)	$\Delta x_1 > -1, \Delta x_2 > \frac{\Delta x_1 + 1}{2}$	(1, 1)
(1, -1)	$\Delta x_1 > -1, \Delta x_2 < \frac{\Delta x_1 + 1}{2}$	(1, -1)
(-1, 1)	$\Delta x_1 < -1, \Delta x_2 > 0$	(-1, 1)
(-1, -1)	$\Delta x_1 < -1, \Delta x_2 < 0$	(-1, -1)

5 制約なし連続最適化

対象とする区分線形な目的関数 $f: \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ は凸であるか, 局所最適解が存在することを仮定する.

Abs-Normal Form による区分的線形関数 f の区分線形近似とその selection function は [2] での selection function に一致する. Abs-Normal Form により, 各 selection function の関数と定義域

$$f_{\sigma, \mathbf{x}}(\Delta \mathbf{x}) \equiv \gamma_\sigma + \mathbf{g}_\sigma^T \Delta \mathbf{x}$$

$$\Sigma \Delta \mathbf{z} = \Sigma (I - L\Sigma)^{-1} (\mathbf{c}_z + Z\Delta \mathbf{x}) \leq 0$$

が得られているので, P_σ に対して線形計画問題 (以下 LP)

$$\begin{aligned} \min \quad & f_{\sigma, \mathbf{x}}(\Delta \mathbf{x}) \\ \text{s.t.} \quad & \Delta \mathbf{z} \leq 0 \\ & -\infty < \Delta x_j < \infty \quad (1 \leq j \leq n) \end{aligned} \quad (5.1)$$

を解くことで, 反復中の解 \mathbf{x}_k が属する P_σ の中で f_{σ, \mathbf{x}_k} を最小にする方向 $\Delta \mathbf{x}$ を求めることが可能であり, これを繰り返すことで局所最適解へと収束することができる.

区分的微分可能関数に対する微分不可能点での最適性条件は一般化勾配 $\partial^C f(\mathbf{x}) \equiv \text{conv} \{\partial^B f(\mathbf{x})\}$ ($\partial^B f(\mathbf{x})$: Bouligand Differential) [3] を用いて

$$0 \in \partial^C f(\mathbf{x})$$

で定義される. Abs-Normal Form による一般化勾配を

$$\partial f(\mathbf{x}) \equiv \{\nabla f_{\sigma, \mathbf{x}} \mid \sigma_0 \preceq \sigma\}$$

とすれば, 同様の条件が定義される.

また, σ と σ' の共通項の集合 $\sigma \cap \sigma' = \tilde{\sigma} \neq \emptyset$ について以下の LP を考える.

$$\begin{aligned} \min \quad & F(\Delta \mathbf{x}) \\ \text{s.t.} \quad & \tilde{\Sigma} \Delta \tilde{\mathbf{z}} \geq 0 \\ & -\infty < \Delta x_j < \infty \quad (1 \leq j \leq n) \end{aligned} \quad (5.2)$$

$F: \mathcal{D} \subset \mathbb{R}^n \rightarrow \mathbb{R}$ は任意でよい. $\tilde{\Sigma}, \Delta \tilde{\mathbf{z}}$ はそれぞれ $\tilde{\sigma}$ に対応する Σ の部分対角行列と $\Delta \mathbf{z}$ の部分 switching vector である.

この LP が解をもたない場合, $\tilde{\sigma} \subseteq \hat{\sigma}$ となるすべての多面体領域 $P_{\tilde{\sigma}}$ は降下方向をもたない. これを利用して, 分枝限定法を行うことで LP を解く回数を 2^s 回から減らすことが期待できる. これにより, 一般化勾配の列挙と最適性条件の判定を効

率化する。

以下に、本研究でのアルゴリズムを記す。

Algorithm 1 Piecewise Linear Minimization

```

procedure PLMIN( $f, \mathbf{x}$ )
   $\mathbf{x}_0 \leftarrow \mathbf{x}$ 
  for  $k \leftarrow 1, 2, \dots$  do
     $f_{\text{PL}, \mathbf{x}_k} \leftarrow \text{AbsNormalForm}(f, \mathbf{x})$ 
     $\boldsymbol{\sigma} \leftarrow \text{sign} \Delta \mathbf{z}(\mathbf{0})$ 
     $\Delta \mathbf{x} \leftarrow \underset{\Delta \mathbf{x}}{\text{argmin}} \{ f_{\boldsymbol{\sigma}, \mathbf{x}_k}(\Delta \mathbf{x}) \mid \Delta \mathbf{x} \in P_{\boldsymbol{\sigma}}, \boldsymbol{\sigma} \preceq \tilde{\boldsymbol{\sigma}} \}$ 
    if  $\Delta \mathbf{x} = \mathbf{0}$  or  $f(\mathbf{x}_k + \Delta \mathbf{x}) \geq f(\mathbf{x}_k)$  then
      STOP
     $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k + \Delta \mathbf{x}$ 
  return  $f(\mathbf{x}_k), \mathbf{x}_k$ 

```

5.1 数値実験・結果

以下に、[4]に掲載されているベンチマーク関数に対して行った実験を抜粋して載せる。

Hiriart-Urruty and Lemaréchal Example

$$f(x_1, x_2) = \max\{-100, 3x_1 - 2x_2, 3x_1 + 2x_2, 2x_1 - 5x_2, 2x_1 + 5x_2\} \quad (5.3)$$

Dimension	2
Optimum Point	$\mathbf{x}^* \in \{ \mathbf{x} \in \mathbb{R}^2 \mid x_1 \leq -50, x_2 \geq 0, 2x_1 \geq 5x_2 \}$ $\cup \{ \mathbf{x} \in \mathbb{R}^2 \mid x_1 \leq -50, x_2 \leq 0, 2x_1 \geq -5x_2 \}$
Optimum Value	$f(\mathbf{x}^*) = -100$
Starting Point	$\mathbf{x}_0 = (9, -3)$

表4 式(5.3)に対する実験結果

f^*	# f	# ∇f	#LP	反復数	時間 [sec]
-99.999999	2	2	3	2	0.00

Goffin

$$f(\mathbf{x}) = 50 \max_{1 \leq i \leq 50} x_i - \sum_{i=1}^{50} |x_i| \quad (5.4)$$

Dimension	50
Optimum Point	$\mathbf{x}^* = (0, \dots, 0)$
Optimum Value	$f(\mathbf{x}^*) = 0$
Starting Point	$\mathbf{x}_0 = (i - 25.5 \mid 1 \leq i \leq 50)$

表5 式(5.4)に対する実験結果

f^*	# f	# ∇f	#LP	反復数	時間 [sec]
1.2×10^{-5}	2	2	2	2	0.29

L1 hlib

$$f(\mathbf{x}) = \sum_{i=1}^n \left| \sum_{j=1}^n \frac{x_j}{i+j-1} \right| \quad (5.5)$$

Dimension	3, 4, 5, 6, 7, 8, 9, 10
Optimum Point	$\mathbf{x}^* = (0, \dots, 0)$
Optimum Value	$f(\mathbf{x}^*) = 0$
Starting Point	$\mathbf{x}_0 = (1, \dots, 1)$

表6 式(5.5)に対する実験結果

n	f^*	# f	# ∇f	#LP	反復数	時間 [sec]
3	3.0×10^{-12}	2	2	2	2	0.02
4	4.0×10^{-12}	2	2	2	2	0.02
5	5.0×10^{-12}	3	3	9	3	0.14
6	6.0×10^{-12}	5	5	103	5	1.16
7	7.0×10^{-12}	8	8	209	8	3.03
8	8.0×10^{-12}	7	7	425	7	8.48
9	9.0×10^{-12}	9	9	857	9	23.92
10	1.0×10^{-11}	5	7	769	5	27.80

表の # f は目的関数 f を評価した回数, # ∇f は目的関数 f の勾配を評価した回数, #LP は LP を解いた回数である。

6 考察

6.1 実験結果について

区分的線形関数に対するこの手法は、各反復で微分不可能点を表す kink [6] と呼ばれる超平面を伝って降下していく。このため、収束までの反復回数は理論上は $3^s - 2^s (= |\boldsymbol{\sigma} \in \{-1, 0, 1\}^s \setminus \{-1, 1\}^s|)$ 回以下となる。実験結果からもこれは正しいと考えられ、収束の様子は非常に安定している。

しかし、目的関数に含まれる絶対値関数の個数 s が増えることで、隣接する多面体領域が指数関数的に増えるため LP を解く回数が著しく増えてしまう。目的関数の変数の次元 n が上がることも s が増えることによる計算量の増加が激しいため、LP を解く回数を減らす、すなわち、効率的に一般化勾配を列挙する方法を考える必要がある。[8]では、Abs-Normal Form を表す行列のうち絶対値関数の入れ子構造を表す行列 L から最適性条件を考察することも行われている。今回は多面体領域を表す制約式の特徴から分枝限定法を行うことである程度現実的な処理時間で収束させることができている。

6.2 区分的非線形関数に対する手法と数値実験

また、今回のアルゴリズムを拡張し、隣接する多面体領域間での移動を踏まえた直線探索や、LP の目的関数に $\frac{\alpha}{2} \|\Delta \mathbf{x}\|^2$ の項を追加した QP (2次計画問題) を解くことで \mathbf{x}_k の近傍で目的関数を2次近似しながら降下方向を探索する手法を[4]の区分的非線形関数に対して実験を行ったが収束が安定しなかった。この拡張手法はいわば、信頼領域法をもとにする Bundle Method であったが、他にも準ニュートン法、非線形共役勾配法などをもとにした拡張も考えられる。しかし、各反復において多面体領域間で移動が発生した場合にはヘッセ行列や共役方向の近似がリセットされてしまうため、まだ考察と実験が必要である。

7 今後の課題

区分的微分可能関数に対する劣微分や一般化勾配の定義は他にも複数あり、今後はそれらの自動微分との関係や導出方法について考察していく。また、制約付き最適化問題や方程式系への応用を最終的な課題とする。

参考文献

- [1] 久保田 光一, 伊理 正夫, アルゴリズムの自動微分と応用. Automatic differentiation of algorithms and applications, 東京, Japan: コロナ社, 1998.7 1998.
- [2] S. Scholtes, *Piecewise Differentiable Functions*, pp. 91–111. New York, NY: Springer New York, 2012.
- [3] F. Clarke, *Optimization and Nonsmooth Analysis*. Society for Industrial and Applied Mathematics, 1990.
- [4] A. Bagirov, N. Karmitza, and M. M. Mkel, *Introduction to Nonsmooth Optimization: Theory, Practice and Software*. Springer Publishing Company, Incorporated, 2014.
- [5] A. Griewank, “On stable piecewise linearization and generalized algorithmic differentiation,” *Optimization Methods Software*, vol. 28, pp. 1139–1178, Dec. 2013.
- [6] A. Griewank, J.-U. Bernt, M. Radons, and T. Streubel, “Solving piecewise lineare equations in abs-normal form,” *Documenta Mathematica*, vol. Optimization Stories, 2015.
- [7] A. Griewank, A. Walther, S. Fiege, and T. Bosse, “On Lipschitz optimization based on gray-box piecewise linearization,” *Mathematical Programming*, vol. 158, no. 1, pp. 383–415, 2016.
- [8] A. Griewank and A. Walther, “First and second order optimality conditions for piecewise smooth objective functions,” *Optimization Methods and Software*, vol. 31, no. 5, pp. 904–930, 2016.