P-2

半正定値計画緩和を用いた最急降下法の解析

Analysis of Gradient Descent Using Semidefinite Programming Relaxation

○酒井一輝 [1]
*Kazuki Sakei[1]

Abstract: This paper analyzes gradient descent by formulating worst-case complexity as a semidefinite program and incorporating an iteration-dependent sequence of smoothness constants. The approach removes the need for a fixed precomputed Lipschitz constant and enables concise theoretical analysis, with future work focusing on numerical validation and further refinement of complexity bounds.

1. はじめに

Performance Estimation Problem（PEP）は, 与えられた問題クラス上で一次最適化アルゴリズムの最悪性能を, 補間不等式と双対化を通じて半正定値計画として評価する枠組みである．本研究では凸関数 f の最小化問題に対して, ステップ係数 $\mathbf{h} := \{h_{i,k}\}$ をもつ fixed-step first-order methods(FSFOM)

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \frac{1}{L}\sum_{k=0}^{i} h_{i+1,k}\nabla f(\mathbf{x}_k), \quad i = 0, \dots, N-1$$

を解析する. 従来の PEP では勾配の Lipschitz 定数 L を固定値として事前に見積もり, アルゴリズムに入力する必要があるが, 解析上の制約でもあった. 本稿ではこの前提を緩め, 反復 i に依存して変化する平滑性定数列 L_i を PEP に直接組み込む定式化を提案する. 具体的には, 補間不等式を反復依存の形へ拡張し, グラム行列化と双対化により計算可能な半正定値計画へ落とし込むことで, 定数 L を前提としない評価を可能にする. 定数 L による従来解析との比較を通じて, 得られる上界の精度や計算効率に与える影響を検証する.

2. PEP の定義と基本的な変換手順

定義 1 (問題クラス) $L > 0$ に対し，$\mathcal{F}_L$ を *L-Lipschitz* 連続な勾配をもつ $\mathbb{R}^d$ 上の凸関数の集合とする.

仮定 1 凸関数 f は最小点をもち，初期点 $\mathbf{x}_0$ は $f(\mathbf{x}_0) - f^* \le \frac{1}{2}LR^2$ を満たすものとする. ただし $R > 0$ は定数である.

定義 2 (原問題 (P)) ステップ係数 $\mathbf{h} := \{h_{i,k}\}$ をもつ *FSFOM,* 反復回数 N, 問題の次元 d, 初期点 $\mathbf{x}_0$ および定数 $R(>0), L$ が与えられた場合, 仮定 1 の下で最悪ケース勾配ノルムの上限は

$$\begin{aligned}
\text{最大化:}\ & \frac{1}{L^2R^2}\left\|\nabla f(\mathbf{x}_N)\right\|^2 \\
\text{制約条件:}\ & \mathbf{x}_{i+1} = \mathbf{x}_i - \frac{1}{L}\sum_{k=0}^{i} h_{i+1,k}\nabla f(\mathbf{x}_k), \\
& \quad (i = 0, \dots, N-1), \\
& f \in \mathcal{F}_L, \mathbf{x}_0, \dots, \mathbf{x}_N \in \mathbb{R}^d \\
& f(\mathbf{x}_0) - f(\mathbf{x}^\star) \le \frac{1}{2}LR^2
\end{aligned} \tag{P}$$

によって定義される.

2.1　有限次元緩和

この問題 (P) は無限次元の関数制約を含むため計算不可能であるが,Kim と Fessler[2] の緩和手法を適用すれば,

$$\begin{aligned}
\text{最大化:}\ & \mathrm{Tr}(\mathbf{u}_N\mathbf{u}_N^\top Z) \\
\text{制約条件:}\ & \mathrm{Tr}(A_{i-1,i}(\mathbf{h})Z) \le \delta_{i-1} - \delta_i, \quad i = 1, \dots, N \\
& \mathrm{Tr}(B_{N,i}(\mathbf{h})Z) \le \delta_N - \delta_i, \quad i = 0, \dots, N-1 \\
& \mathrm{Tr}(C_N(\mathbf{h})Z) \le \delta_N, \\
& \delta_0 \le \frac{1}{2}, \\
& Z \succeq 0, \delta \in \mathbb{R}^{N+1}
\end{aligned} \tag{P1}$$

という有限次元の半正定値計画に緩和される．ただし, $\mathbf{u}_i := [0, \dots, 0, \underbrace{1}_{i+1}, 0, \dots, 0]^\top \in \mathbb{R}^{N+1}$ であり, $Z \succeq 0$ という表現は Z が半正定値行列であるということを表す.

1: 日大理工・院 (前)・数学

また $A_{i-1,i}(\mathbf{h}), B_{N,i}(\mathbf{h}), C_N$を次のように定義する.

$$\begin{aligned}A_{i-1,i}(\mathbf{h}) &:= \frac{1}{2}(\mathbf{u}_{i-1}-\mathbf{u}_i)^\top(\mathbf{u}_{i-1}-\mathbf{u}_i)\\ &\quad+\frac{1}{2}\sum_{k=0}^{i-1}h_{i,k}(\mathbf{u}_i\mathbf{u}_k^\top+\mathbf{u}_k\mathbf{u}_k^\top)\\ B_{N,i}(\mathbf{h}) &:= \frac{1}{2}(\mathbf{u}_N-\mathbf{u}_i)(\mathbf{u}_N-\mathbf{u}_i)^\top\\ &\quad+\frac{1}{2}\sum_{l=i+1}^{N}\sum_{k=0}^{l-1}h_{l,k}\big(\mathbf{u}_i\mathbf{u}_k^\top+\mathbf{u}_k\mathbf{u}_i^\top\big)\\ C_N &:= \frac{1}{2}\mathbf{u}_N\mathbf{u}_N^\top\end{aligned}$$

定理 3 (Kim and Fessler[2]) $f \in \mathcal{F}_L$ と仮定 1 の下で, $\mathbf{x}_0\cdots\mathbf{x}_N \in \mathbb{R}^d$ は最急降下法によって生成されたとする. つまり,*FSFOM* によるステップ係数 $h_{i+1,k}$ が $k=i$ のとき 1 で, それ以外のとき 0 とする. このとき $N \geq 1$ に対して

$$\|\nabla f(\mathbf{x}_N)\|^2 \leq \frac{L^2R^2}{2N+1} \tag{1}$$

証明方針. Kim と Fessler[2] は (P1) のラグランジュ双対の実行可能解を与え, それが最適解であることを示した.

3. PEP の一般化

原問題 (P) は L-Lipschitz 連続な勾配をもつ凸関数上で最悪ケースの勾配ノルムを最大化する無限次元最適化である. 本研究では, これをより柔軟な有限次元の半正定値計画 (P1) に緩和するため，対 $(i-1,i)$, (N,i), $(N,*)$ に対して Kim と Fessler[2] が課していた補間不等式に反復依存の平滑性定数 L_i と更新係数 M_i を導入した.

3.1　補間不等式と反復式

$(i-1,i)$,(N,i) のみ平滑性定数を反復依存した L_{i-1} とし

$$\begin{aligned}&\frac{1}{2L_{i-1}}\|\nabla f(\mathbf{x}_i)-\nabla f(\mathbf{x}_{i-1})\|^2\\ &\qquad\leq f(\mathbf{x}_i)-f(\mathbf{x}_{i-1})-\langle\nabla f(\mathbf{x}_{i-1}),\mathbf{x}_i-\mathbf{x}_{i-1}\rangle\end{aligned}$$

を課す. $(N,*)$ では従来通り定数 L を用いる. 更新式における L と区別するため

$$\mathbf{x}_{i+1}=\mathbf{x}_i-\frac{1}{M_i}\sum_{k=0}^{i}h_{i+1,k}\,\nabla f(\mathbf{x}_k),\qquad i=0,\ldots,N-1$$

とする.

3.2　正規化と有限次元化

従来と同様に

$$\begin{aligned}\mathbf{g}_i &= \frac{1}{LR}\nabla f(\mathbf{x}_i),\quad G := [\,\mathbf{g}_0,\ldots,\mathbf{g}_N\,]^\top,\\ \delta_i &= \frac{1}{LR^2}\big(f(\mathbf{x}_i)-f^\star\big), i=0,\ldots,N,\end{aligned}$$

と正規化し，グラム行列 $Z=GG^\top$ を導入する. 補間不等式は Z と標準基底 $\mathbf{u}_i$ を用いて

$$\mathrm{Tr}\big(\widetilde{A}_{i-1,i}(\mathbf{h})Z\big) \leq \frac{L_{i-1}}{L}(\delta_{i-1}-\delta_i)$$

$$\mathrm{Tr}\big(\widetilde{B}_{N,i}(\mathbf{h})Z\big) \leq \frac{L_N}{L}(\delta_N-\delta_i)$$

などの線形制約にまとめられる.

3.3　半正定値計画と双対問題

以上より (P) は

$$\max_{Z\succeq 0,\,\delta}\ \mathrm{Tr}(\mathbf{u}_N\mathbf{u}_N^\top Z)$$

という有限次元の半正定値計画 ($\widetilde{P1}$) に緩和され，(P1) の係数行列が $\widetilde{A}_{i-1,i}(\mathbf{h})$，$\widetilde{B}_{N,i}(\mathbf{h})$，$C_N$ に置き換えられる. ラグランジュ双対化により

$$\min_{\mathbf{a},\mathbf{b},c,e\geq 0}\ \frac{1}{2}e$$

を目的とする双対問題 (D) が得られ，行列

$$\sum_{i=1}^{N}a_i\widetilde{A}_{i-1,i}(\mathbf{h})+\sum_{i=0}^{N-1}b_i\widetilde{B}_{N,i}(\mathbf{h})+cC_N-\mathbf{u}_N\mathbf{u}_N^\top$$

が半正定値となること，および δ に関する線形等式制約が付随する.

4. 最悪勾配上界に関する課題

定理 3 では定数 L の下で

$$\|\nabla f(\mathbf{x}_N)\|^2 \leq \frac{L^2R^2}{2N+1}$$

が示された. 今後は, 本研究で導入した反復依存の平滑性定数 L_{i-1} および更新係数 M_i をもつ最急降下法に対しても, 定理 3 を一般化したより柔軟な不等式が成り立つかどうか検証していく.

5. 参考文献

[1] Y. Drori and M. Teboulle, Performance of first-order methods for smooth convex minimization: a novel approach, *Mathematical Programming*, **145**, pp. 451–82, 2014.

[2] D. Kim and J. A. Fessler, Optimizing the efficiency of first-order methods for decreasing the gradient of smooth convex functions, *Journal of Optimization Theory and Applications*, **188**, pp.192–219, 2021.