

「非凸最適化アルゴリズムとその計算量解析」演習問題

丸茂 直貴

2024/8/24–26

■記法について

- $\mathbb{N} := \{0, 1, \dots\}$ とします.
- ベクトル $x, y \in \mathbb{R}^d$ に対し, 内積を $\langle x, y \rangle := \sum_{i=1}^d x_i y_i$, ノルムを $\|x\| := \sqrt{\langle x, x \rangle}$ と定めます.
- 行列 $A \in \mathbb{R}^{m \times n}$ に対し, $\|A\|$ をその作用素ノルムとします. $\|A\| := \sup_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|}$ です.
- 正定値とは限らない対称行列 $A \in \mathbb{R}^{d \times d}$ とベクトル $x \in \mathbb{R}^d$ に対し, $\langle Ax, x \rangle$ を $\|x\|_A^2$ と表します. ($\|x\|_A^2 < 0$ になることもあります, 記号の濫用をします.)

問題 1. 基礎的問題

関数 $f: \mathbb{R}^d \rightarrow \mathbb{R}$ は 2 階微分可能とする.

- 任意の $x \in \mathbb{R}^d$ に対し $\|\nabla^2 f(x)\| \leq L$ ならば, ∇f は L -Lipschitz 連続であることを示せ.
- ∇f が L -Lipschitz 連続ならば, 任意の $x \in \mathbb{R}^d$ に対し $\|\nabla^2 f(x)\| \leq L$ であることを示せ.
- ∇f の Lipschitz 定数は大きい, $\nabla^2 f$ の Lipschitz 定数は小さい関数 f の例を挙げよ.
- ∇f の Lipschitz 定数は小さい, $\nabla^2 f$ の Lipschitz 定数は大きい関数 f の例を挙げよ.

問題 2. 正則化無しの Newton 法が収束しない強凸平滑関数の例

関数 $g: \mathbb{R} \rightarrow \mathbb{R}$ を

$$g(x) := \sqrt{1 + x^2}$$

と定める.

- g が凸関数であること, また g' が 1-Lipschitz 連続であることを示せ.
- g'' が Lipschitz 連続であることを示せ.
- $\mu > 0$ に対し, 関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ を

$$f(x) := g(x) + \frac{\mu}{2} x^2$$

と定め, 関数 f に対する Newton 法

$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)} \quad (k = 0, 1, \dots)$$

を考える. μ が十分小さいとき, ある $\alpha \in \mathbb{R}$ が存在し, 初期点 $x_0 = \alpha$ から始めた Newton 法が収束しないことを示せ.

問題 3. ステップサイズを大きくすることも許す場合の最急降下法の計算量

関数 $f: \mathbb{R}^d \rightarrow \mathbb{R}$ について, ∇f は L -Lipschitz 連続であると仮定する. $x_0 \in \mathbb{R}^d$, $\ell_{\text{init}} > 0$, $\alpha > 1 \geq \beta > 0$ を任意にとり, $\Delta := f(x_0) - \inf_{x \in \mathbb{R}^d} f(x)$, $\ell_{\text{max}} := \max\{\ell_{\text{init}}, \alpha L\}$ とおく. 以下のアルゴリズムは, $\|\nabla f(x_k)\| \leq \varepsilon$ なる点 x_k を

$$\left\lceil \frac{2\ell_{\text{max}}\Delta}{\varepsilon^2} \right\rceil \log_{\alpha} \left(\frac{\alpha}{\beta} \right) + \log_{\alpha} \left(\frac{\ell_{\text{max}}}{\ell_{\text{init}}} \right)$$

回以下の反復で見つけることを示せ. (k が増加しない反復もあることに注意.)

```

1:  $\ell \leftarrow \ell_{\text{init}}, k \leftarrow 0$ 
2: repeat
3:    $x \leftarrow x_k - \frac{1}{\ell} \nabla f(x_k)$ 
4:   if  $f(x) - f(x_k) \leq \langle \nabla f(x_k), x - x_k \rangle + \frac{\ell}{2} \|x - x_k\|^2$  :
5:      $x_{k+1} \leftarrow x, k \leftarrow k + 1, \ell \leftarrow \beta \ell$ 
6:   else
7:      $\ell \leftarrow \alpha \ell$ 

```

問題 4. 最急降下法の収束が遅くなる関数の例

(a) $r_0 > r_1 > 0$ を定数とする. 微分可能な関数 $g: \mathbb{R} \rightarrow \mathbb{R}$ のうち, 条件

$$g(0) = 0, \quad g'(0) = -r_0, \quad g(r_0) \geq -r_0^2, \quad g'(r_0) = -r_1,$$

と

$$|g'(x) - g'(y)| \leq |x - y|, \quad \forall x, y \in [0, r_0]$$

を満たすものを一つ求めよ.

(b) $0 < \delta < 1$ を定数とする. 微分可能な関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ と実数 $L, \Delta > 0$ の組のうち, 以下の条件

(i)–(iii) を全て満たすものを一つ構成せよ:

(i) f' は L -Lipschitz 連続である.

(ii) $f(0) - \inf_{x \in \mathbb{R}} f(x) \leq \Delta$.

(iii) 数列 $(x_k)_{k \in \mathbb{N}}$ を, 初期点 $x_0 = 0$, ステップサイズ $\frac{1}{L}$ の最急降下法

$$x_{k+1} = x_k - \frac{1}{L} f'(x_k) \quad (k = 0, 1, \dots)$$

により定める. このとき, 任意の $\varepsilon > 0$ に対し, $|f'(x_k)| \leq \varepsilon$ を満たす最小の k は

$$\frac{\delta L \Delta}{2\varepsilon^{2-\delta}} - 1$$

以上である.

ヒント: 不等式 $\sum_{k=1}^{\infty} k^{-\frac{2}{2-\delta}} \leq \frac{2}{\delta}$ を証明無しに用いてよい.

問題 5. 3 次正則化 Newton 法の子問題の性質

定数 $M > 0$, ベクトル $g \in \mathbb{R}^d$, (半正定値とは限らない) 対称行列 $H \in \mathbb{R}^{d \times d}$ に対し, 最適化問題

$$\min_{u \in \mathbb{R}^d} \left\{ f(u) := \langle g, u \rangle + \frac{1}{2} \|u\|_H^2 + \frac{M}{6} \|u\|^3 \right\}$$

を考える. この問題の任意の最適解を $u^* \in \mathbb{R}^d$ とおき, さらに $\sigma := \frac{M}{2} \|u^*\|$ とおく.

(a) 等式 $(H + \sigma I)u^* = -g$ が成り立つことを示せ.

(b) 行列

$$H + \sigma I + \sigma \left(\frac{u^*}{\|u^*\|} \right) \left(\frac{u^*}{\|u^*\|} \right)^\top$$

は半正定値であることを示せ.

(c) $\langle u^*, v \rangle \neq 0$ なる $v \in \mathbb{R}^d$ を任意に固定する. 等式

$$f(u^* + \alpha v) - f(u^*) = \frac{\alpha^2}{2} \|v\|_{H+\sigma I}^2$$

を満たす $\alpha \neq 0$ を一つ求めよ.

(d) 行列 $H + \sigma I$ は半正定値であることを示せ.

問題 6. Hesse 行列の Lipschitz 連続性から従う Hessian-free 不等式

関数 $f: \mathbb{R}^d \rightarrow \mathbb{R}$ について, $\nabla^2 f$ は M -Lipschitz 連続であると仮定する.

(a) $x_1, \dots, x_n \in \mathbb{R}^d$ と, $\sum_{i=1}^n \lambda_i = 1$ なる $\lambda_1, \dots, \lambda_n \geq 0$ を任意にとり, $\bar{x} := \sum_{i=1}^n \lambda_i x_i$ とおく. 以下の式を示せ:

$$\left\| \nabla f(\bar{x}) - \sum_{i=1}^n \lambda_i \nabla f(x_i) \right\| \leq \frac{M}{2} \sum_{i=1}^n \lambda_i \|x_i - \bar{x}\|^2 = \frac{M}{2} \sum_{1 \leq i < j \leq n} \lambda_i \lambda_j \|x_i - x_j\|^2.$$

(b) 任意の $x, y \in \mathbb{R}^d$ に対し, 以下の不等式を示せ:

$$f(x) - f(y) \leq \frac{1}{2} \langle \nabla f(x) + \nabla f(y), x - y \rangle + \frac{M}{12} \|x - y\|^3.$$

ヒント: (a) を用いる.

問題 7. Heavy-ball 法における関数値減少量の評価

関数 $f: \mathbb{R}^d \rightarrow \mathbb{R}$ について, ∇f は L -Lipschitz 連続, $\nabla^2 f$ は M -Lipschitz 連続であると仮定する. 点列 $(x_k)_{k \in \mathbb{N}}$ を, 慣性係数が 1 の Heavy-ball 法

$$x_{k+1} = 2x_k - x_{k-1} - \frac{1}{L} \nabla f(x_k) \quad (k = 0, 1, \dots)$$

で定める. ただし $x_{-1} := x_0$ とする. 表記の簡略化のため, $k \geq -1$ に対し $u_k := x_{k+1} - x_k$ とおき, $k \geq 0$ に対し $S_k := \sum_{i=0}^{k-1} \|u_i\|^2$ とおく.

(a) $k \geq 0$ に対し, 以下の不等式を示せ:

$$f(x_{k+1}) - f(x_k) \leq \frac{L}{2} \langle u_{k-1}, u_k \rangle - \frac{L}{2} \langle u_k, u_{k+1} \rangle + \frac{M}{12} \|u_k\|^3.$$

(b) $k \geq 0$ に対し、以下の不等式を示せ:

$$\frac{f(x_{k+1}) + f(x_k)}{2} - f(x_0) \leq -\frac{L}{4}\|u_k\|^2 + \frac{M}{12} \sum_{i=0}^{k-1} \|u_i\|^3.$$

ただし $\sum_{i=0}^{-1} \|u_i\|^3 = 0$ と約束する.

(c) $k \geq 1$ に対し、以下の不等式を示せ:

$$\min_{1 \leq i \leq k} f(x_i) - f(x_0) \leq \frac{1}{2k-1} \left(-\frac{LS_k}{2} + \frac{M}{6}(k-1) \sum_{i=0}^{k-2} \|u_i\|^3 \right).$$

(d) $k \geq 1$ を任意に固定する. $\sqrt{(k-1)^5 S_{k-1}} \leq \frac{3L}{4M}$ であるとき、以下の不等式を示せ:

$$\min_{1 \leq i \leq k} f(x_i) - f(x_0) \leq -\frac{LS_k}{4k}.$$

ヒント: ℓ_p ノルムの単調性と、 $k \geq 2$ に対し $2(k-1)^{3/2} \geq k$ が成り立つことを使う.

問題 8. Heavy-ball 法における勾配ノルムの評価

本問題では前問と同じ設定を引き継ぐ. $k \geq 1$ に対し $\bar{x}_k := \frac{1}{k} \sum_{i=0}^{k-1} x_i$ とおく.

(a) $0 \leq i < j < k$ に対し、 $\|x_i - x_j\|^2 \leq kS_{k-1}$ を示せ.

ヒント: 三角不等式と Cauchy-Schwarz の不等式を用いる.

(b) $k \geq 1$ に対し、以下の不等式を示せ:

$$k\|\nabla f(\bar{x}_k)\| \leq L\|u_{k-1}\| + \frac{M}{4}k(k-1)S_{k-1}.$$

(c) $k \geq 1$ に対し、以下の不等式を示せ:

$$\sum_{i=1}^k i\|\nabla f(\bar{x}_i)\| \leq L\sqrt{kS_k} + \frac{M}{12}k(k^2-1)S_{k-1}.$$

(d) $k \geq 1$ を任意に固定する. $\sqrt{(k-1)^5 S_{k-1}} \leq \frac{3L}{4M}$ であるとき、以下の不等式を示せ:

$$\min_{1 \leq i \leq k} \|\nabla f(\bar{x}_i)\| \leq 3L\sqrt{\frac{S_k}{k^3}}.$$

問題 9. Levenberg-Marquardt 法の計算量 (難)

ベクトル値関数 $F: \mathbb{R}^d \rightarrow \mathbb{R}^n$ の Jacobi 行列関数 $J: \mathbb{R}^d \rightarrow \mathbb{R}^{n \times d}$ が L -Lipschitz 連続であると仮定する. 点列 $(x_k)_{k \in \mathbb{N}}$ を、最小二乗問題

$$\min_{x \in \mathbb{R}^d} \left\{ f(x) := \frac{1}{2} \|F(x)\|^2 \right\}$$

に対する Levenberg-Marquardt (LM) 法

$$x_{k+1} \in \operatorname{argmin}_{x \in \mathbb{R}^d} \left\{ \bar{f}_k(x) := \frac{1}{2} \|F(x_k) + J(x_k)(x - x_k)\|^2 + \frac{\lambda_k}{2} \|x - x_k\|^2 \right\} \quad (k = 0, 1, \dots)$$

で定める. ただしパラメータ λ_k は $\lambda_k := L\|F(x_k)\|$ と定める. 表記の簡略化のため $u_k := x_{k+1} - x_k$, $\Delta := f(x_0) - \inf_{x \in \mathbb{R}^d} f(x)$ とおく.

(a) $k \geq 0$ に対し, 以下の不等式を示せ:

$$\bar{f}_k(x_{k+1}) \leq f(x_k) - \frac{\lambda_k}{2} \|u_k\|^2.$$

(b) $k \geq 0$ に対し, 以下の不等式を示せ:

$$f(x_{k+1}) \leq \bar{f}_k(x_{k+1}).$$

ヒント: (a) を用いて $\|F(x_k) + J(x_k)u_k\| \leq \frac{\lambda_k}{L} - L\|u_k\|^2$ を示す.

以下では, $\sigma > 0$ を定数とし, 任意の $k \in \mathbb{N}$ に対し $\|J(x_k)\| \leq \sigma$ が成り立つと仮定する.

(c) $k \geq 0$ に対し, 以下の不等式を示せ:

$$\|\nabla f(x_{k+1})\| \leq 2\lambda_k \|u_k\| + \frac{L\sigma}{2} \|u_k\|^2.$$

(d) $k \geq 1$ に対し, 以下の不等式を示せ:

$$\sum_{i=0}^{k-1} \lambda_i \|u_i\| \leq \sqrt{2L\Delta \|F(x_0)\|} k.$$

ヒント: Cauchy–Schwarz の不等式を用いる.

(e) $k \geq 0$ に対し, 以下の不等式を示せ:

$$\|u_k\|^2 \leq \frac{2}{L^2} (\lambda_k - \lambda_{k+1}).$$

(f) $k \geq 1$ に対し, 以下の不等式を示せ:

$$\min_{1 \leq i \leq k} \|\nabla f(x_i)\| \leq \sqrt{\frac{8L\Delta \|F(x_0)\|}{k}} + \frac{\sigma \|F(x_0)\|}{k}.$$