

交互近接勾配法とラッソ回帰への応用

小野寺優希也，松下慎也，徐粒

秋田県立大学

2021 年 9 月 17 日

発表の流れ

- ① はじめに
- ② 準備
- ③ アルゴリズム
- ④ 数値実験
- ⑤ 結論
- ⑥ 参考文献

目次

① はじめに

② 準備

③ アルゴリズム

④ 数値実験

⑤ 結論

⑥ 参考文献

はじめに

本研究では以下の最適化問題を考える：

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1 \quad (1)$$

- $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $\|\cdot\|_1$ は ℓ_1 ノルム, $\|\cdot\|_2$ は ℓ_2 ノルム, $\lambda > 0$
- 式 (1) の最適化問題はラッソ回帰と呼ばれている
- ラッソ回帰の応用例
 - 画像信号処理 [1]
 - 曲線フィッティング [2] [3]

[1] A. Chambolle and T. Pock: A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging Vision, **40**-1, 120/145 (2011)

[2] M. Nagahara: Sparsity methods for systems and control, 31-53, Now Publishers (2020)

[3] 鈴木讓: スパース推定 100 問 with Python, 2-22, 共立出版 (2021)

例：曲線フィッティング

データを近似する曲線を見つける問題を考える。

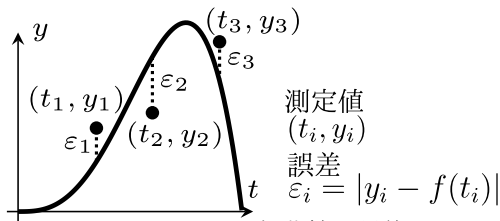


図 a：近似曲線と誤差

$f(t) = \beta_4 t^4 + \beta_3 t^3 + \beta_2 t^2 + \beta_1 t^1 + \beta_0$, $x := [\beta_4 \ \beta_3 \ \beta_2 \ \beta_1 \ \beta_0]$ として、求めたい問題の解が疎 (**スパース**) のとき、以下のように定式化できる。

$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & \varepsilon_1^2 + \varepsilon_2^2 + \varepsilon_3^2 \approx 0 \end{array} \Leftrightarrow \text{minimize} \quad \|Ax - b\|_2^2 + \lambda \|x\|_1$$

はじめに

ラッソ回帰

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1$$

- ラッソ回帰の解法：近接勾配法
 - 近接勾配法の性質

$$\|x_{k+1} - x^*\|_2 \leq \|x_k - x^*\|_2 \quad (k = 1, 2, \dots) \quad (2)$$

ここで、 x^* は問題 (1) の最適解である。以後、 x^* は問題 (1) の最適解とする。

- 近接勾配法の加速化：FISTA [4]
 - 式 (2) を満たさない

[4] A. Beck and M. Teboulle: A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci., 2-1, 183/202 (2009)

はじめに

式 (2)

$$\|x_{k+1} - x^*\|_2 \leq \|x_k - x^*\|_2 \quad (k = 1, 2, \dots)$$

- 式 (2) を満たす場合
 - 一度解に到達すると解から離れない
- 式 (2) を満たさない場合
 - 一度解に到達しても次の繰り返しでは解から離れてしまう

本研究の目的

ラッソ回帰において式 (2) を満たす近接勾配法の改良について検討する。

目次

- ① はじめに
- ② 準備
- ③ アルゴリズム
- ④ 数値実験
- ⑤ 結論
- ⑥ 参考文献

近接写像の例

Definition 2 (近接写像)

$\gamma > 0$, $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \{+\infty\}$ を凸関数とする. このとき, f の近接写像 $\text{prox}_{\gamma f}$ を以下で定義する.

$$\text{prox}_{\gamma f}(x) := \arg \min_{y \in \mathbb{R}^n} \left(f(y) + \frac{1}{2\gamma} \|x - y\|_2^2 \right) \quad (4)$$

アルゴリズムにおいて ℓ_1 ノルム

$$\|x\|_1 := \sqrt{|x_1| + |x_2| + \cdots + |x_n|}$$

の近接写像を計算する必要がある.

近接写像の例

Example 1 (ℓ_1 ノルムの近接写像)

ℓ_1 ノルムの近接写像は $i = 1, 2, \dots, n$ に対して以下で与えられる.

$$[\text{prox}_{\gamma \|\cdot\|_1}(x)]_i = \text{sgn}(x_i) \max\{|x_i| - \gamma, 0\} \quad (5)$$

ただし, sgn は符号を表す.

式 (5) は

$$[\text{prox}_{\gamma \|\cdot\|_1}(x)]_i = \begin{cases} x_i - \gamma & (x_i > \gamma) \\ 0 & (-\gamma \leq x_i \leq \gamma) \\ x_i + \gamma & (x_i < -\gamma) \end{cases}$$

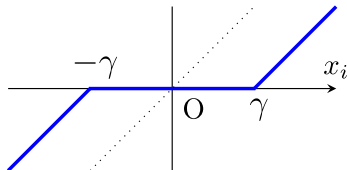


図 b : $[\text{prox}_{\gamma \|\cdot\|_1}(x)]_i$

が成り立つ.

目次

- ① はじめに
- ② 準備
- ③ アルゴリズム
- ④ 数値実験
- ⑤ 結論
- ⑥ 参考文献

近接勾配法

Algorithm 1 近接勾配法

Input: $x_0 \in \mathbb{R}^n$, $\gamma \in (0, 2/\|A^T A\|)$, $\lambda > 0$

- 1: **while** 停止条件を満たすまで **do**
- 2: $x_{k+1} = \text{prox}_{\gamma\lambda\|\cdot\|_1}(x_k - \gamma A^T(Ax_k - b))$
- 3: **end while**

Output: x_k

- Algorithm 1 によって生成された点列 $\{x_k\}$ は問題 (1) の最適解に収束する [7].

[7] H. H. Bauschke and P. L. Combettes: Convex analysis and monotone operator theory in Hilbert spaces, Springer (2017)

FISTA [4]

Algorithm 2 FISTA [4]

Input: $x_0, x_1 \in \mathbb{R}^n$, $\gamma \in (0, 1/\|A^T A\|]$, $t_1 = 1$, $\lambda > 0$

- 1: **while** 停止条件を満たすまで **do**
- 2: $t_{k+1} = (1 + \sqrt{1 + 4t_k^2})/2$
- 3: $\alpha_k = (t_k - 1)/t_{k+1}$
- 4: $y_k = x_k + \alpha_k(x_k - x_{k-1})$
- 5: $x_{k+1} = \text{prox}_{\gamma\lambda\|\cdot\|_1}(y_k - \gamma A^T(Ay_k - b))$
- 6: **end while**

Output: x_k

- 近接勾配法との違い：2～4 の処理を追加

[4] A. Beck and M. Teboulle: A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci., 2-1, 183/202 (2009)

近接勾配法と FISTA の比較

- 数値実験の詳細は「4. 数値実験」を参照

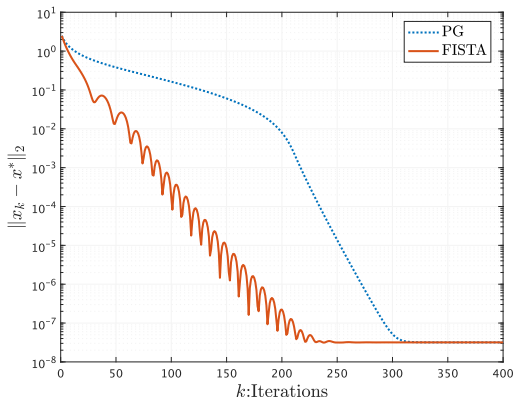


Fig. 1: ラッソ回帰に対して近接勾配法 (PG) と FISTA を適用して点列 $\{x_k\}$ と最適解 x^* の差 $\|x_k - x^*\|_2$ を比較した結果

FISTA の問題点

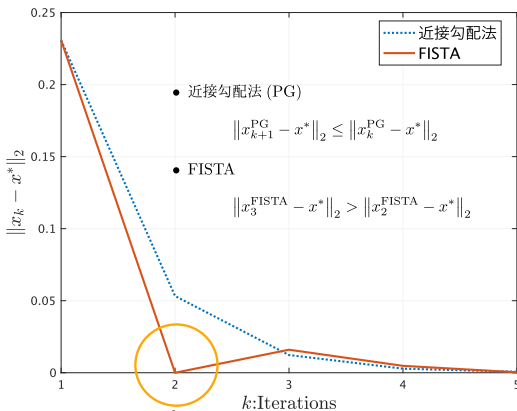


図 c : $\min \frac{x^2}{2}$ に対して近接勾配法と FISTA を適用した結果

- Z. Mu and Y. Peng: A note on the inertial proximal point method, Stat. Optim. Inf. Comput., 3-3, 241/248 (2015)

近接勾配法と FISTA の特徴

Fig. 1 と図 c より

- 近接勾配法
 - 式 (2) を満たしている (単調に解に収束する)
- FISTA
 - 式 (2) を満たしてはいない (振動しながら解に収束する)
 - 近接勾配法よりも速く最適解に収束している

今回検討するアルゴリズム

- 近接勾配法よりも速く最適解に収束
- 式 (2) を満たしている

交互近接勾配法

Algorithm 3 交互近接勾配法

Input: $x_0, x_1 \in \mathbb{R}^n$, $\gamma \in (0, 2/\|A^T A\|)$, $\lambda > 0$, $\alpha_k \in [0, +\infty)$, ただし,
 $\alpha_k \leq \min\{1, 1/(\gamma\|A^T A\|)\} - 1/2$

```
1: while 停止条件を満たすまで do  
2:   if  $k$  is even then  
3:      $y_k = x_k$   
4:   else  
5:      $y_k = x_k + \alpha_k(x_k - x_{k-1})$   
6:   end if  
7:    $x_{k+1} = \text{prox}_{\gamma\lambda\|\cdot\|_1}(y_k - \gamma A^T(Ay_k - b))$   
8: end while
```

Output: x_k

- 近接勾配法との違い：2～6 の処理を追加

交互近接勾配法の収束定理 [6]

Theorem 2

$\{x_k\}$ を Algorithm 3 で生成された点列とすると、以下が成り立つ。

- ① $\{x_k\}$ は x^* に収束する。
- ② 以下が成り立つ、ただし、 k は偶数である。

$$\|x_{k+2} - x^*\|_2 \leq \|x_k - x^*\|_2. \quad (6)$$

本研究で取り組んだこと

- $\|x_k - x^*\|_2$ の評価は重要にも関わらず、先行研究では解析されていない。
- 本研究ではラッソ回帰に対する $\|x_k - x^*\|_2$ の評価を行なった。

[6] F. lutzeler and J. M. Hendrickx: On the proximal gradient algorithm with alternated inertia, J. Optim. Theory Appl., **176**-3, 688/710 (2018)

目次

- ① はじめに
- ② 準備
- ③ アルゴリズム
- ④ 数値実験
- ⑤ 結論
- ⑥ 参考文献

数値実験の概要

ラッソ回帰

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1$$

- ラッソ回帰において近接勾配法, FISTA, 交互近接勾配法を適用
- アルゴリズムの評価基準: $\|x_k - x^*\|_2$
- 最適解 x^* は CVX(凸最適化のライブラリ (<http://cvxr.com/>)) を用いて求める

Table 2: 数値実験で用いた計算環境

OS	macOS Big Sur
メモリ	8 GB 1600 MHz DDR3
CPU	1.8 GHz Intel Core i5
使用ソフト	MATLAB R2021a

数値実験の問題設定

ラッソ回帰

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1$$

- $A \in \mathbb{R}^{130 \times 80}$: 各成分が標準正規分布に従うように生成
- $b \in \mathbb{R}^{130}$: $b = Ax^* + e$
 - $x^* \in \mathbb{R}^{80}$: 標準正規分布から得られる 10 % スパースベクトル
 - $e \in \mathbb{R}^{130}$: 平均 0, 分散 0.001 の正規分布に従うように生成
- 初期点 : 原点
- 正則化パラメータ : $\lambda = 0.1$
- ステップサイズ : $\gamma = \frac{1}{2\|A^T A\|}$
- $\{\alpha_k\}$ (交互近接勾配法) : $\alpha_k = 0.5$

数値実験の結果

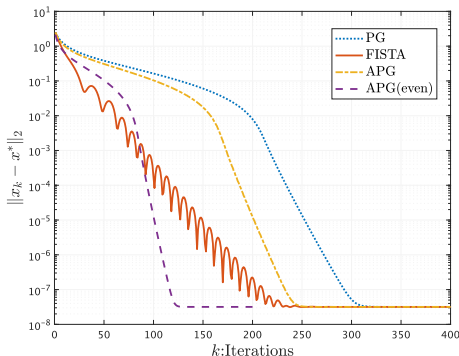


Fig. 2: ラッソ回帰に対して近接勾配法 (PG), FISTA, 交互近接勾配法 (APG) を適用して点列 $\{x_k\}$ と最適解 x^* の差 $\|x_k - x^*\|_2$ を比較した結果

- 交互近接勾配法：式 (6) を満たしながら，近接勾配法よりも速く x^* に収束

参考：各アルゴリズムと CVX の実行時間

Table 3: ラッソ回帰において数値実験を 50 回行なった際の実行時間の平均値の結果 (近接勾配法：PG, 交互近接勾配法：APG)

Algorithm	CPU time [s]
PG	0.03652
FISTA	0.03814
APG	0.04281
CVX	2.1906

参考：交互近接勾配法の数値例

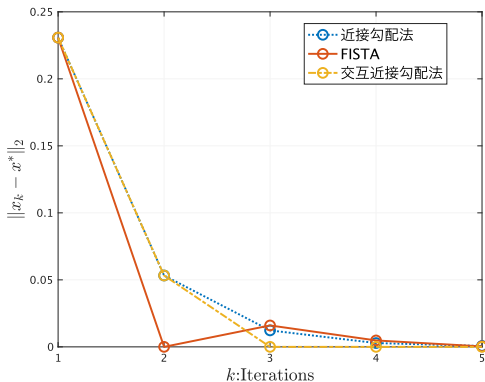


図 d : $\min \frac{x^2}{2}$ に対して近接勾配法, FISTA, 交互近接勾配法を適用した結果

- Z. Mu and Y. Peng: A note on the inertial proximal point method, Stat. Optim. Inf. Comput., 3-3, 241/248 (2015)

目次

- ① はじめに
- ② 準備
- ③ アルゴリズム
- ④ 数値実験
- ⑤ 結論
- ⑥ 参考文献

結論

本研究で取り組んだこと

- 先行研究：関数値による評価
 - $|F(x_k) - F(x^*)|$, ただし, $F(x) := \|Ax - b\|_2^2/2 + \lambda\|x\|_1$
 - 点列が最適解に収束しているか確認できない
- 本研究：点列と最適解の差による評価

結論

- 数値実験で点列 $\{x_k\}$ と最適解 x^* の差 $\|x_k - x^*\|_2$ を比較することにより交互近接勾配法の収束に関する有用性を検証することができた

今後の課題 (1)

画像信号処理

$$\min_{x \in \mathbb{R}^N} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|Kx\|_1$$

なお、 A は変換行列、 b は観測画像、 $K \in \mathbb{R}^{2N \times N}$ は差分変換行列である。

Observation



(a)

Result



(b)

図 e: 観測画像 (a) と観測画像に交互近接勾配法を適用した結果 (b)

- 図 e の場合は $N = 65536$ となる。
- 交互近接勾配法の点列の収束性とPSNR 値を評価

今後の課題 (2)

- 点列の収束性
 - 近接勾配法 (PG)

$$\|x_{k+1}^{\text{PG}} - x_k^{\text{PG}}\|_2 = o(1/\sqrt{k})$$

- FISTA

$$\|x_{k+1}^{\text{FISTA}} - x_k^{\text{FISTA}}\|_2 = o(1/k)$$

今後の課題 (2)

交互近接勾配法に対して点列の収束性の評価を理論的に保障する

参考文献

- [1] A. Chambolle and T. Pock: A first-order primal-dual algorithm for convex problems with applications to imaging, J. Math. Imaging Vision, **40**-1, 120/145 (2011)
- [2] M. Nagahara: Sparsity methods for systems and control, 31-53, Now Publishers (2020)
- [3] 鈴木讓: スパース推定 100 問 with Python, 2-22, 共立出版 (2021)
- [4] A. Beck and M. Teboulle: A fast iterative shrinkage-thresholding algorithm for linear inverse problems, SIAM J. Imag. Sci., **2**-1, 183/202 (2009)
- [5] F. Iutzeler and J. M. Hendrickx: A generic online acceleration scheme for optimization algorithms via relaxation and inertia, Optim. Methods Software, **34**-2, 383/405 (2019)
- [6] F. Iutzeler and J. M. Hendrickx: On the proximal gradient algorithm with alternated inertia, J. Optim. Theory Appl., **176**-3, 688/710 (2018)
- [7] H. H. Bauschke and P. L. Combettes: Convex analysis and monotone operator theory in Hilbert spaces, Springer (2017)