

オンライン制約下での半正定値対称行列の近似手法

数理工学専攻 48-186205 合田 理貴
指導教員 岩田 覚 教授

1 はじめに

半正定値対称行列の近似手法とは行列 $\mathbf{A} \in \mathbb{R}^{n \times d}$ ($n \gg d$) を入力として

$$(1 - \epsilon) \mathbf{A}^\top \mathbf{A} \preceq \tilde{\mathbf{A}}^\top \tilde{\mathbf{A}} \preceq (1 + \epsilon) \mathbf{A}^\top \mathbf{A}$$

を満たすような行列 $\tilde{\mathbf{A}} \in \mathbb{R}^{m \times d}$ を構築する手法のことをいう。ここで関係 \preceq を対称行列 \mathbf{X}, \mathbf{Y} について ” $\mathbf{X} \preceq \mathbf{Y} \stackrel{\text{def}}{\iff} \mathbf{Y} - \mathbf{X}$ が半正定値対称行列” と定義する。このとき $\tilde{\mathbf{A}}$ を \mathbf{A} の $(1 \pm \epsilon)$ -スペクトル近似と呼び、その行数 m が小さいほど望ましいとされる。本研究ではスペクトル近似を行列 \mathbf{A} の各行の重み付けサンプリングによって構成することを考える。つまり行列 $\tilde{\mathbf{A}}$ は非負値を対角成分にもつ対角行列 $\mathbf{D} \in \mathbb{R}^{m \times m}$ 、各行に非零成分としてちょうど1つの1をもつ行列 $\mathbf{S} \in \mathbb{R}^{m \times n}$ を用いて $\tilde{\mathbf{A}} = \mathbf{D}\mathbf{S}\mathbf{A}$ の形で表されるとする。

2 既存研究および準備

初めに制約を課さないもとでスペクトル近似を求める手法について説明する。以降の議論で行列 \mathbf{A} は d 次元列ベクトル \mathbf{a}_i を用いて $\mathbf{A} = (\mathbf{a}_1, \dots, \mathbf{a}_n)^\top$ のように成分表示されるとする。行列 \mathbf{A} のスペクトル近似は定義 2.1 で定められるレバレッジスコアと呼ばれる値を用いて構築できることが知られている (定理 2.2)。

定義 2.1 (レバレッジスコア). 行列 $\mathbf{A} \in \mathbb{R}^{n \times d}$ について第 i 行 \mathbf{a}_i^\top のレバレッジスコア $\tau(\mathbf{A}, \mathbf{a}_i)$ を

$$\tau(\mathbf{A}, \mathbf{a}_i) \stackrel{\text{def}}{=} \mathbf{a}_i^\top (\mathbf{A}^\top \mathbf{A})^+ \mathbf{a}_i$$

と定義する。ただし $(\mathbf{A}^\top \mathbf{A})^+$ は行列 $\mathbf{A}^\top \mathbf{A}$ のムーアペンローズ逆行列を表すとする。

定理 2.2 (Li, Miller, Peng [3]). $\mathbf{A} \in \mathbb{R}^{n \times d}$ をランク r の行列, $\epsilon \in (0, 1)$ を実数とする。確率変数 $\mathbf{S} \in \mathbb{R}^{n \times n}$ を対角成分 S_{ii} が確率 $p_i \stackrel{\text{def}}{=} \min(\tau(\mathbf{A}, \mathbf{a}_i) \epsilon^{-2} \log d, 1)$ で $1/\sqrt{p_i}$, 確率 $1 - p_i$ で 0 をとる対角行列とする。このとき高い確率で $\mathbf{S}\mathbf{A}$ は \mathbf{A} の $(1 \pm \epsilon)$ -スペクトル近似となり, $\mathbf{S}\mathbf{A}$ の非零な行の数は $O(r\epsilon^{-2} \log d)$ となる。

次にオンライン制約下での近似手法について考える。ここでオンライン制約とは簡単に ”入力として行列 \mathbf{A} の各行が逐次的に与えられること”, ”各行をサンプリ

ングするかどうかの決定をその行が与えられた際に行い, 以降その決定を変更できないこと” の 2 つを課した制約を指す。オンライン制約下では全体の行列 \mathbf{A} が未知であるため各行のレバレッジスコアを直接計算することはできない。そこで代わりに現在までにサンプリングした行列を用いてレバレッジスコアのようなものを計算し, それをもとにサンプリング確率を定めることを考える。以下ではその値を相対レバレッジスコアとして導入する。定義における \mathbf{B} がその時点までにサンプリングした行からなる行列に対応する。

定義 2.3 (相対レバレッジスコア). ベクトル \mathbf{a} , 行列 \mathbf{B} について \mathbf{a} の \mathbf{B} に対する相対レバレッジスコア $\tau^{\mathbf{B}}(\mathbf{a})$ を以下のように定義する。

$$\tau^{\mathbf{B}}(\mathbf{a}) \stackrel{\text{def}}{=} \mathbf{a}^\top \left(\begin{pmatrix} \mathbf{B} \\ \mathbf{a}^\top \end{pmatrix} \begin{pmatrix} \mathbf{B} \\ \mathbf{a}^\top \end{pmatrix} \right)^+ \mathbf{a}.$$

先行研究として Cohen, Musco, Pachocki [2] によってオンライン制約下で

$$(1 - \epsilon) \mathbf{A}^\top \mathbf{A} - \delta \mathbf{I} \preceq \tilde{\mathbf{A}}^\top \tilde{\mathbf{A}} \preceq (1 + \epsilon) \mathbf{A}^\top \mathbf{A} + \delta \mathbf{I}$$

の近似保証を達成する (ϵ, δ) -スペクトル近似を構築するアルゴリズムが提案されている。本研究では相対レバレッジスコアの値を用いて ”近似保証” および ”スペクトル近似の行数” についてその改善を行った。

3 オンライン制約下での近似手法

Algorithm 1 では各行 \mathbf{a}_i^\top について $i - 1$ 行目までにサンプリングした行からなる行列 $\tilde{\mathbf{A}}_{i-1}$ に対する相対レバレッジスコア $\tau^{\tilde{\mathbf{A}}_{i-1}}(\mathbf{a}_i)$ を求める。それを元にサンプリング確率を設定することで高い確率でスペクトル近似を構築することが可能となる (定理 3.1)。

定理 3.1. ランク r の行列 $\mathbf{A} \in \mathbb{R}^{n \times d}$ および実数 $\epsilon \in (0, 1/2]$ を入力としたとき **Algorithm 1** は高い確率で行数 $O(r\epsilon^{-2} \log \mu(\mathbf{A}) \log d)$ の $(1 \pm \epsilon)$ -スペクトル近似を返す。ただし $\mu(\mathbf{A})$ は以下のように定める。

$$\mu(\mathbf{A}) \stackrel{\text{def}}{=} \frac{\sigma_{\max}(\mathbf{A})}{\min_{1 \leq i \leq n} \sigma_{\min}(\mathbf{A}_i)}.$$

$\sigma_{\max}(\mathbf{X}), \sigma_{\min}(\mathbf{X})$ はそれぞれ行列 \mathbf{X} の最大特異値, 最小非零特異値を表し, 行列 $\mathbf{A}_i \in \mathbb{R}^{i \times d}$ は行列 \mathbf{A}

の第 1 行目から第 i 行目までからなる部分行列とする.

Algorithm 1 OnlineRowSampling (\mathbf{A}, ϵ)

- 1: **入力:** 行列 $\mathbf{A} \in \mathbb{R}^{n \times d}$, 実数 $\epsilon \in (0, 1/2]$.
 - 2: **出力:** 行列 $\tilde{\mathbf{A}}_n$.
 - 3: $c = 5\epsilon^{-2} \log d$.
 - 4: $\tilde{\mathbf{A}}_0 \leftarrow \mathbf{O}$.
 - 5: **for** $i = 1, \dots, n$ **do**
 - 6: $p_i \leftarrow \min \left(c\tau^{\tilde{\mathbf{A}}_{i-1}}(\mathbf{a}_i), 1 \right)$.
 - 7: $\tilde{\mathbf{A}}_i \leftarrow \begin{cases} \begin{pmatrix} \tilde{\mathbf{A}}_{i-1} \\ \mathbf{a}_i^\top / \sqrt{p_i} \end{pmatrix} & \text{確率 } p_i, \\ \tilde{\mathbf{A}}_{i-1} & \text{その他の場合.} \end{cases}$
 - 8: **end for**
 - 9: **return** $\tilde{\mathbf{A}}_n$.
-

4 オンラインランダム順列制約下での近似手法

Algorithm 1 では相対レバレッジスコアの計算のために毎回擬似逆行列の計算を行う必要があり, またスペクトル近似の行数も最悪の場合 $\Theta(n)$ になるという問題点が存在する. そのため本研究では入力にランダム順列の仮定を課したオンラインランダム順列制約を新たに導入し, より効率的なアルゴリズムの設計を行った.

問題のインスタンスを \mathbb{R}^d 上のベクトル n 個からなる集合 $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ とする. 任意の n 次の置換 σ について $n \times d$ 行列 $(\mathbf{x}_{\sigma(1)}, \dots, \mathbf{x}_{\sigma(n)})^\top$ を考え, そのような $n!$ 個の $n \times d$ 行列からなる集合を $\mathcal{S}(\mathbf{X})$ とする. 台集合を $\mathcal{S}(\mathbf{X})$ とする離散一様分布を $\mathcal{A}(\mathbf{X})$ としたときアルゴリズムの入力 \mathbf{A} を $\mathcal{A}(\mathbf{X})$ に従う確率変数と考える. **Algorithm 2** では入力を $O(\log n)$ 個のブロックに分け, i 番目のブロック内の各行 \mathbf{a}_j^\top の相対レバレッジスコアの計算を $i-1$ 番目のブロックまでにサンプリングされた行からなる行列 $\tilde{\mathbf{M}}_{i-1}$ を用いて計算することで擬似逆行列の計算回数を全体で $O(\log n)$ 回に抑えている. また構築されるスペクトル近似の行数についても良い上界が得られることを示した (定理 4.1).

定理 4.1. \mathbf{X} を \mathbb{R}^d 上の n 個のベクトルからなる集合, ϵ を $(0, 1/2]$ を満たす実数とする. このとき $n \times d$ 行列 $\mathbf{A} \sim \mathcal{A}(\mathbf{X})$ について高い確率で以下が成り立つ.

行列 \mathbf{A} および実数 ϵ を入力としたとき **Algorithm 2** は高い確率で $O(d\epsilon^{-2} \log n \log d)$ 行のスペクトル近似を出力する.

Algorithm 2 ScaledSampling (\mathbf{A}, ϵ)

- 1: **入力:** 行列 $\mathbf{A} \in \mathbb{R}^{n \times d}$, 実数 $\epsilon \in (0, 1/2]$.
 - 2: **出力:** 行列 $\tilde{\mathbf{A}}_n$.
 - 3: $K = d \log d$, $c = 9\epsilon^{-2} \log d$, $\alpha = \log_2(n/K)$.
 - 4: $\tilde{\mathbf{A}}_0 \leftarrow \mathbf{O}$.
 - 5: **for** $j = 1, \dots, K$ **do**
 - 6: $\tilde{\mathbf{A}}_j \leftarrow \begin{pmatrix} \tilde{\mathbf{A}}_{j-1} \\ \mathbf{a}_j^\top \end{pmatrix}$.
 - 7: **end for**
 - 8: **for** $i = 1, \dots, \alpha$ **do**
 - 9: $\tilde{\mathbf{M}}_{i-1} \leftarrow \tilde{\mathbf{A}}_{(2^i-1)K}$.
 - 10: **for** $j = (2^i-1)K + 1, \dots, (2^{i+1}-1)K$ **do**
 - 11: $p_j \leftarrow \min \left(c\tau^{\tilde{\mathbf{M}}_{i-1}}(\mathbf{a}_j), 1 \right)$.
 - 12: $\tilde{\mathbf{A}}_j \leftarrow \begin{cases} \begin{pmatrix} \tilde{\mathbf{A}}_{j-1} \\ \mathbf{a}_j^\top / \sqrt{p_j} \end{pmatrix} & \text{確率 } p_j, \\ \tilde{\mathbf{A}}_{j-1} & \text{その他の場合.} \end{cases}$
 - 13: **end for**
 - 14: **end for**
 - 15: **return** $\tilde{\mathbf{A}}_n$.
-

5 スペクトル近似のサイズについての下界

最後にスペクトル近似のサイズ (行数) の下界に関する結果として, オンラインランダム順列制約下で構築されるスペクトル近似のサイズは $\Omega(d\epsilon^{-2} \log n)$ 行となることを示した. 参考として制約を課さないもとのスペクトル近似のサイズの下界は $\Omega(d\epsilon^{-2})$ 行であることが知られている [1].

定理 5.1. R をオンラインランダム順列制約下で $(1 \pm \epsilon)$ -スペクトル近似を構築するアルゴリズムとする. このとき次を満たす \mathbb{R}^d 上の n 個のベクトルからなる集合 \mathbf{X} が存在する.

行列 $\mathbf{A} \sim \mathcal{A}(\mathbf{X})$, 実数 $\epsilon \in (0, 1)$ を入力としたとき, R が出力する $(1 \pm \epsilon)$ -スペクトル近似のサイズは高い確率で $\Omega(d\epsilon^{-2} \log n)$ となる.

参考文献

- [1] J. D. Batson, D. A. Spielman, and N. Srivastava. Twice-ramanujan sparsifiers. *CoRR*, abs/0808.0163, 2008.
- [2] M. B. Cohen, C. Musco, and J. W. Pachocki. Online row sampling. *CoRR*, abs/1604.05448, 2016.
- [3] G. L. Miller and R. Peng. Iterative approaches to row sampling. *CoRR*, abs/1211.2713, 2012.