近似GCDでのNewtonSLRAアルゴリズムの効果的な利用に向けて

Toward finding better approximate GCD with NewtonSLRA

神戸大学 大学院 人間発達環境学研究科　長坂 耕作 *1

Kosaku Nagasaka

Graduate School of Human Development and Environment, Kobe University

Abstract

In this talk, we briefly review the NewtonSLRA/1 algorithm for the well-known structured low rank approximation, and consider its variants for finding better approximate polynomial GCD.

1　はじめに

次のように定義される近似 GCD には様々な算法が知られているが，本報告では，Sylvester 行列（部分終結式行列）のゼロ空間から余因子の係数ベクトルを抽出し，そこから近似 GCD を求める方法に着目する。この方法は様々な最適化問題として解かれるが，本報告では，SLRA (Structured Low Rank Approximation) に帰着する方法についてのみ取り扱う。

定義 1 (近似 GCD – 次数指定型 –)
多項式 $f(x), g(x) \in \mathbb{R}[x]$ と次数 $k \in \mathbb{N}$ に対し，次式を満たす $d(x) \in \mathbb{R}[x]$ で摂動最小のものを，次数 k の近似 *GCD* という。なお，摂動は $\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2$ とする。

$$\exists \Delta_f(x), \Delta_g(x), f_1(x), g_1(x) \in \mathbb{R}[x],\ \deg(\Delta_f) \le \deg(f),\ \deg(\Delta_g) \le \deg(g),\ \deg(d) = k,$$
$$f(x) + \Delta_f(x) = f_1(x)d(x),\ g(x) + \Delta_g(x) = g_1(x)d(x)$$

ここで，SLRA (Structured Low Rank Approximation) とは，与えられた行列の構造を保ちつつ，その階数をより小さくした行列を求める問題である。本報告では，Schost と Spaenlehauer による論文 [SS16] に基づき，解くべき SLRA を以下のように定義しておく。なお，以下では，$m \times n$ の実行列全体を $\mathcal{M}_{m,n}(\mathbb{R})$ により表し，$\mathcal{M}_{m,n}(\mathbb{R})$ の内積として $\langle M_1, M_2 \rangle := \mathrm{trace}(M_1 M_2^T)$ を導入する（この内積で導かれるノルムは，Frobenius ノルム）。加えて，階数が r となる行列の集合を $\mathcal{D}_r \subset \mathcal{M}_{m,n}(\mathbb{R})$ を用いて表し，$E \subset \mathcal{M}_{m,n}(\mathbb{R})$ を $\mathcal{M}_{m,n}(\mathbb{R})$ の affine 部分空間とし，その正規直交基底（内積は前述のもの）を $E_1, \ldots, E_d$ とする。

定義 2 (SLRA)
行列 $M \in E$ と正整数 r に対し，$\|M - M^*\|$ が小さくなるような行列 $M^* \in E \cap \mathcal{D}_r$ を求めよ。

*1 E-mail: nagasaka@main.h.kobe-u.ac.jp

SLRA の解法としては，Cadzow[Cad88] による Lift-and-Project アルゴリズム（または，Cadzow アルゴリズム）がシンプルで使いやすい。このアルゴリズムは，構造を持たない行列として目的の階数を持つ最近行列を特異値分解により求め（$\mathcal{D}_r$ への Lifting），それを正射影により本来の構造を持つ行列に変形（E への Projection）することを繰り返す反復法である（収束は線形）。一方，Schost と Spaenlehauer[SS16] による NewtonSLRA/1 アルゴリズム（問題条件により NewtonSLRA/2 も提案されている）では，E への Projection を正射影ではなく，Newton 法の考え方を踏襲し，局所的な二次収束を実現している。

2 NewtonSLRA/1 の詳細

Schost と Spaenlehauer[SS16] による NewtonSLRA/1 アルゴリズムの詳細な手順を述べる。前述の問題（定義 2）に基づき，入力の行列 $M \in E$ と正整数 r に対し，$M_0 = M$ として，M_i から M_{i+1} を求める反復法であり，その収束先（便宜上 M_∞ とおく）が最終的に $M_\infty \in E \cap \mathcal{D}_r$ を満たすように構成されている。

実際の反復計算（$M_i \to M_{i+1}$）は，次の Lifting と Projection に分かれる。

Lifting（$\tilde{M}_i \in \mathcal{D}_r$ なる M_i に最近の行列を求める）

- $U\Sigma V^T = M_i$: 特異値分解（Σ は特異値が降順に対角に並ぶ）
- $\tilde{M}_i := U\tilde{\Sigma}V^T$: $\tilde{\Sigma}$ は $r+1$ 番目からの特異値を 0 に置き換え

Project （$M_{i+1} \in E$ なる行列を $\tilde{M}_i$ から求める）

- $\mathcal{D}_r$ の $\tilde{M}_i$ での normal space の基底 $N_k := \tilde{u}_s \tilde{v}_t^T$ （U, V の $r+1$ 列目以降の組み合わせ）
- $A = (a_{st}) \in \mathcal{M}_{(m-r)(n-r),d}(\mathbb{R})$, $\vec{b} = (b_s) \in \mathcal{M}_{(m-r)(n-r),1}(\mathbb{R})$, $a_{st} = \langle N_s, E_t \rangle$, $b_s = \langle N_s, \tilde{M}_i - M_i \rangle$ として，$M_{i+1} := M_i + (E_1 \ \cdots \ E_d)(A^\dagger \vec{b})$ なる更新を行う（$\cdot^\dagger$ は疑似逆行列）

3 改善の試みと実験結果

Cadzow アルゴリズムは，正射影を行っていることから，E の正規直交基底 $\{E_1, \ldots, E_d\}$ の選び方に非依存であるが，NewtonSLRA アルゴリズムは，$\mathcal{D}_r$ の $\tilde{M}_i$ での normal space 上での最小二乗解を用いているため基底により結果が変化する。特に近似 GCD では，SLRA で落とすべき階数は常に 1 であり，反復の最終段階の最小二乗解（疑似逆行列の部分）の部分は制約式が近似 GCD の次数 $k+1$ しかない過少決定系となり，正規直交基底の選び方に依存する部分は比較的大きいと言える。また，正規直交基底を入力の多項式の各係数それぞれに対応させる形で構成した場合でも，Frobenius ノルムで正規化する関係から，最終段階の最小二乗解は多項式の係数ベクトルの観点からは最小ノルム解にはならない。しかしながら，重み行列を $A^\dagger \vec{b}$ なる最小二乗解の求解時の前後に適用することで，近づけることは可能である。さらに，NewtonSLRA は必ずしも最近の（もしくはそれに近い）M_∞ に収束するとは限らず，近似 GCD が求める摂動の最小化を著しく達成できない可能性がある。そのため，反復初期に大きな修正項が発生しないよう小さなステップサイズに抑えることも有効かもしれない（減衰付き Newton 法とは逆）。

以上の修正の試みについて，多項式ペア $\{f(x) = \tilde{f}(x) \times d(x) + \Delta_f(x),\ g(x) = \tilde{g}(x) \times d(x) + \Delta_g(x)\}$ を次のように生成し有効性の検証を行った。なお，各係数は積を構成する前の段階で，$[-1, 1] \subset \mathbb{R}$ からランダムに，摂動多項式 $\Delta_f(x)$ と $\Delta_g(x)$ の各係数は $[-err, err] \subset \mathbb{R}$ からランダムに生成した。

small perturbation（各 100 ペア）

- default: $\deg(\tilde{f}) = 30$, $\deg(\tilde{g}) = 28$, $\deg(d) = 5$, $err = 1.0\text{e-}8$

- small GCD: $\deg(\tilde{f}) = 30$, $\deg(\tilde{g}) = 28$, $\deg(d) = 2$, $err = 1.0$e-8
- large GCD: $\deg(\tilde{f}) = 20$, $\deg(\tilde{g}) = 18$, $\deg(d) = 10$, $err = 1.0$e-8
- assymmetry: $\deg(\tilde{f}) = 30$, $\deg(\tilde{g}) = 12$, $\deg(d) = 5$, $err = 1.0$e-8

large perturbation（各 100 ペア）

- default: $\deg(\tilde{f}) = 30$, $\deg(\tilde{g}) = 28$, $\deg(d) = 5$, $err = 1.0$e-2
- small GCD: $\deg(\tilde{f}) = 30$, $\deg(\tilde{g}) = 28$, $\deg(d) = 2$, $err = 1.0$e-2
- large GCD: $\deg(\tilde{f}) = 20$, $\deg(\tilde{g}) = 18$, $\deg(d) = 10$, $err = 1.0$e-2
- assymmetry: $\deg(\tilde{f}) = 30$, $\deg(\tilde{g}) = 12$, $\deg(d) = 5$, $err = 1.0$e-2

また，参考文献 [UM17] の 7.2.2 章の実験に実際に使用された多項式（各 20 ペア）をその著者から入手し，同様に実験を行った（$f(x) = (x^3 + x^2 + x + 1)u_d$, $g(x) = (-x^3 + x^2 - x + 1)u_d$ を摂動させたもの）。

3.1 実験結果の概要

表 1 から表 10 に実験結果（平均値）を掲載している。LRA は部分終結式行列そのままから近似 GCD を抽出したもので，UVGCD は参考文献 [Zen11] のアルゴリズム，SSGCD が参考文献 [SS16] の NewtonSLRA/1 アルゴリズム，SSGCD-W は最終段階の最小二乗解を係数ベクトルの最小化に重み行列で寄せたもの，SSGCD-D は反復初期のステップサイズを抑制したもの，SSGCD-DW はそれらの組み合わせ，SSGCD-BW は正規直交基底を変えたもの（表から読み取れるように効果がないため，詳細は割愛する）を表す。なお，これらの実験は Mathematica 12 上に各アルゴリズムを実装して行った（計算時間は今回使用した実装に依存するため，あくまでも相対的な目安に過ぎないことに留意されたい）。

実験結果から，正規直交基底の選び方や重み行列などにより収束先が変わることは確認できるが，改善されたとは言い難く，抜本的な改善へは程遠いと思われる。しかしながら，NewtonSLRA/1 自体が UVGCD の不得手な問題に対して効果的であることも実験結果（表 9 と表 10）は示唆している。残念ながら改善には至らなかったが，より摂動の小さな近似 GCD を見つけうることはわかったため，機会があればさらなる改善の検討を行っていきたい。

謝　辞

This work was supported by JSPS KAKENHI Grant Number 19K11827.

参 考 文 献

[Cad88] James A. Cadzow. Signal enhancement-a composite property mapping algorithm. *IEEE Trans. Acoust. Speech Signal Process.*, 36(1):49–62, 1988.

[SS16] Éric Schost and Pierre-Jean Spaenlehauer. A quadratically convergent algorithm for structured low-rank approximation. *Found. Comput. Math.*, 16(2):457–492, 2016.

[UM17] Konstantin Usevich and Ivan Markovsky. Variable projection methods for approximate (greatest) common divisor computations. *Theoret. Comput. Sci.*, 681:176–198, 2017.

[Zen11] Zhonggang Zeng. The numerical greatest common divisor of univariate polynomials. In *Randomization, relaxation, and complexity in polynomial equation solving*, volume 556 of *Contemp. Math.*, pages 187–217. Amer. Math. Soc., Providence, RI, 2011.

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0616	0.00	7.6707e-8
UVGCD	0.0585	2.23	6.9460e-9
SSGCD	4.2880	2.51	6.9461e-9
SSGCD-W	4.3625	2.65	6.9460e-9
SSGCD-D	4.7276	3.34	6.9461e-9
SSGCD-DW	4.7658	3.37	6.9460e-9
SSGCD-BW	10.0179	2.40	6.9461e-9

表 1: default: $\deg(\tilde{f}) = 30, \deg(\tilde{g}) = 28, \deg(d) = 5, err = 1.0\mathtt{e-}8$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0548	0.00	4.7895e-7
UVGCD	0.0489	2.21	3.9970e-9
SSGCD	3.2112	2.10	3.9995e-9
SSGCD-W	3.2034	2.10	3.9970e-9
SSGCD-D	3.4384	3.03	3.9995e-9
SSGCD-DW	3.4541	3.05	3.9970e-9
SSGCD-BW	7.5230	2.09	3.9995e-9

表 2: small GCD: $\deg(\tilde{f}) = 30, \deg(\tilde{g}) = 28, \deg(d) = 2, err = 1.0\mathtt{e-}8$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0675	0.00	1.9092e-7
UVGCD	0.0542	2.24	1.0241e-8
SSGCD	4.1942	3.04	1.0248e-8
SSGCD-W	3.9856	2.79	1.0241e-8
SSGCD-D	4.6064	3.94	1.0248e-8
SSGCD-DW	4.6158	3.93	1.0241e-8
SSGCD-BW	9.5100	3.04	1.0248e-8

表 3: large GCD: $\deg(\tilde{f}) = 20, \deg(\tilde{g}) = 18, \deg(d) = 10, err = 1.0\mathtt{e-}8$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0408	0.00	4.3615e-8
UVGCD	0.0385	2.12	7.0911e-9
SSGCD	3.0096	2.06	7.5052e-9
SSGCD-W	3.0271	2.15	7.0911e-9
SSGCD-D	3.2827	3.07	7.5052e-9
SSGCD-DW	3.3217	3.12	7.0911e-9
SSGCD-BW	5.1191	2.09	7.5052e-9

表 4: assymmetry: $\deg(\tilde{f}) = 30, \deg(\tilde{g}) = 12, \deg(d) = 5, err = 1.0\texttt{e-}8$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0744	0.00	5.8924e-2
UVGCD	0.2344	11.85	7.2848e-3
SSGCD	6.4976	5.55	7.1767e-3
SSGCD-W	6.2004	5.52	7.1734e-3
SSGCD-D	7.5990	7.17	1.9107e-2
SSGCD-DW	7.3477	7.12	1.9273e-2
SSGCD-BW	32.6104	5.67	7.1746e-3

表 5: default: $\deg(\tilde{f}) = 30, \deg(\tilde{g}) = 28, \deg(d) = 5, err = 1.0\texttt{e-}2$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0506	0.00	3.6073e-2
UVGCD	0.1624	10.73	3.9136e-3
SSGCD	4.0238	4.80	3.2046e-2
SSGCD-W	4.0194	4.79	1.4659e-2
SSGCD-D	4.4211	6.37	2.4332e-2
SSGCD-DW	4.4329	6.36	4.6206e-2
SSGCD-BW	22.5747	4.80	9.2822e-3

表 6: small GCD: $\deg(\tilde{f}) = 30, \deg(\tilde{g}) = 28, \deg(d) = 2, err = 1.0\texttt{e-}2$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0484	0.00	1.4409e-1
UVGCD	0.1498	11.92	1.0034e-2
SSGCD	5.9586	6.34	2.1549e-2
SSGCD-W	5.8510	6.08	1.2222e-2
SSGCD-D	6.5583	7.47	1.9450e-2
SSGCD-DW	6.5392	7.39	1.3590e-2
SSGCD-BW	21.2738	5.78	2.8131e-2

表 7: large GCD: $\deg(\tilde{f}) = 20, \deg(\tilde{g}) = 18, \deg(d) = 10, err = 1.0\texttt{e-}2$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.0503	0.00	4.3051e-2
UVGCD	0.1514	11.54	7.0656e-3
SSGCD	3.9488	5.03	7.5224e-3
SSGCD-W	3.9070	5.03	7.0891e-3
SSGCD-D	4.2638	6.47	7.4963e-3
SSGCD-DW	4.2644	6.40	7.0715e-3
SSGCD-BW	15.6781	5.00	7.5226e-3

表 8: assymmetry: $\deg(\tilde{f}) = 30, \deg(\tilde{g}) = 12, \deg(d) = 5, err = 1.0\text{e-}2$

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.2108	0.00	2.1110e-1
UVGCD	0.2270	10.50	1.0959e-2
SSGCD	16.9444	7.05	6.9406e-4
SSGCD-W	16.9606	7.05	6.9406e-4
SSGCD-D	19.3118	8.05	6.8487e-4
SSGCD-DW	19.3708	8.05	6.8487e-4
SSGCD-BW	26.3968	7.05	6.9702e-4

表 9: u_d の次数 50 の実験結果

アルゴリズム	計算時間	反復回数	$\sqrt{\|\Delta_f\|_2^2 + \|\Delta_g\|_2^2}$
LRA	0.3288	0.00	1.4099e-1
UVGCD	0.6398	11.40	1.1949e-1
SSGCD	30.3062	7.55	9.9627e-4
SSGCD-W	30.2838	7.55	9.9627e-4
SSGCD-D	34.4998	8.75	9.8903e-4
SSGCD-DW	34.5460	8.75	9.8903e-4
SSGCD-BW	110.8728	7.55	1.0067e-3

表 10: u_d の次数 100 の実験結果