

ランダムノイズを含む Vandermonde 行列に関する統計モデルとその一致推定量

相島 健助¹¹ 法政大学情報科学部

e-mail : aishima@hosei.ac.jp

1 はじめに

Vandermonde 行列に関する数値データに対する数値計算は、応用上よく現れる重要なものである。ここで考えるのは数値データに確率的な摂動が加わる統計モデルであり、このような場合は最小二乗法の応用によりパラメータ推定を行うことが多い。本稿では、信号の到来方向を推定する手法 [1] からの着想により、Vandermonde 行列に関する統計モデルへの一致推定量を構成する。

2 全最小二乗法と行列の特異値分解

ここでは、線型方程式において、係数行列や右辺ベクトルが正確に与えられるとは限らず、

$$\bar{A}X = \bar{B}, \quad A = \bar{A} + E_A, \quad B = \bar{B} + E_B \quad (1)$$

と定式化される問題を考える。線型方程式 $\bar{A}X = \bar{B}$ における \bar{A}, \bar{B} の代わりに摂動が加わった A, B が観測され、その情報を基に X を推定したい。そのために、下記の全最小二乗問題

$$\min_{\hat{A}, \hat{B}, \hat{X}} \|A - \hat{A}\|_F^2 + \|B - \hat{B}\|_F^2 \quad \text{s. t.} \quad \hat{A}\hat{X} = \hat{B} \quad (2)$$

を考える。上記の \hat{X} を求めるために、行列 A, B を部分行列とする $C = [A, B]$ に対して、 C の大きい方から n 個の特異値に対応する n 本の右特異ベクトルからなる $(n + \ell) \times n$ 行列を Y とおく。このとき、全最小二乗問題の解 \hat{X} は、 Y の部分行列を基に以下のように得られる。

$$[Y_1, Y_2] := Y^\top, \quad Y_1 \in \mathbf{R}^{n \times n}, \quad Y_2 \in \mathbf{R}^{n \times \ell}, \quad \hat{X} = Y_1^{-1}Y_2. \quad (3)$$

上記の推定値 \hat{X} ともとの X の関係を解析する上で、 $E = [E_A, E_B]$ とし、下記を仮定する。

仮定 1 行列 E の各行をなす行ベクトルは、平均が零ベクトル、各変数同士の共分散を表す共分散行列が $\sigma^2 I_n$ であり、すべての行ベクトルは独立同一分布に従う。

上記の仮定の下で、統計的な意味での一致性を示す下記の収束定理が得られる。

定理 2 $\lim_{m \rightarrow \infty} m^{-1} \bar{A}^\top \bar{A}$ が存在し、この極限の行列のランクは n とする。このとき仮定 1 の下で、 $m \rightarrow \infty$ として \hat{X} は X に確率 1 で収束する。

この定理と拡張は [3, 4] に詳しい。以下、この類推から Vandermonde 行列に関する一致性を示す。

3 Vandermonde 行列に関する統計モデルとその一致推定量

本稿では、Vandermonde 行列に関する統計モデルを考える。Vandermonde 行列とは

$$V = \begin{bmatrix} 1 & v_1 & v_1^2 & \cdots & v_1^{n-1} \\ 1 & v_2 & v_2^2 & \cdots & v_2^{n-1} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & v_\ell & v_\ell^2 & \cdots & v_\ell^{n-1} \end{bmatrix} \quad (4)$$

である．ここで考える問題設定では， V は直接観測できず，観測できるのは下記の M のみである．

$$L \in \mathbf{R}^{m \times \ell}, \quad \bar{M} = LV, \quad M = \bar{M} + E. \quad (5)$$

M の観測を基に， V をなす v_1, v_2, \dots, v_ℓ を推定する．この問題設定は，電波到来方向推定のための ESPRIT 法 [1] や，時系列データに対する動的モード分解の統計解析 [4] を意識したものである．

具体的な推定手法について論じる前に， V と固有値問題の関係を述べる．そのために

$$V_1 = \begin{bmatrix} 1 & v_1 & \cdots & v_1^{n-2} \\ 1 & v_2 & \cdots & v_2^{n-2} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & v_\ell & \cdots & v_\ell^{n-2} \end{bmatrix}, \quad V_2 = \begin{bmatrix} v_1 & v_1^2 & \cdots & v_1^{n-1} \\ v_2 & v_2^2 & \cdots & v_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ v_\ell & v_\ell^2 & \cdots & v_\ell^{n-1} \end{bmatrix}$$

とおく．ここで $\ell \leq n-1$ であれば， $W = V_1^\dagger V_2$ の非零固有値が v_1, \dots, v_ℓ となる．ただし， V_1^\dagger は V_1 の Moore–Penrose の一般化逆行列とする．この原理を基に，以下に具体的な推定法を述べる．観測できる M の大きい方から ℓ 個の特異値に対応する ℓ 本の右特異ベクトルからなる $n \times \ell$ 行列を N とおく．このとき， N^\top の第 1 列から第 $n-1$ 列からなる $\ell \times (n-1)$ 行列を N_1 とし，さらに， N^\top の第 2 列から第 n 列からなる $\ell \times (n-1)$ 行列を N_2 とおくと，ある $\ell \times \ell$ 行列 S により $N_1 \approx SV_1$ ， $N_2 \approx SV_2$ となることに着目する．この性質に基づき，下記の全最小二乗問題

$$\min_{\hat{N}_1, \hat{N}_2, \hat{W}} \|N_1 - \hat{N}_1\|_F^2 + \|N_2 - \hat{N}_2\|_F^2 \quad \text{s. t.} \quad \hat{N}_1 \hat{W} = \hat{N}_2 \quad (6)$$

を解き， \hat{W} の非零固有値を v_1, \dots, v_ℓ の推定値とすると，下記の一致性が成立する．

定理 3 観測される行列 M をなす式 (5) の L に対して， $\lim_{m \rightarrow \infty} m^{-1} \bar{L}^\top \bar{L}$ が存在し，この極限の行列のランクは ℓ とする．さらに， $\ell \leq n-1$ とする．このとき仮定 1 の下で， $m \rightarrow \infty$ として \hat{W} の非零固有値は，式 (4) の v_1, \dots, v_ℓ に確率 1 で収束する．

上記の定理は，[1] における提案手法である ESPRIT 法の一致性の理論保証を含む．この主定理の既存研究における位置づけ，応用や数値的な性質については，講演時に報告したい．

謝辞 本研究は JSPS 科研費 JP22K03422 の助成を受けた．

参考文献

- [1] A. Paulraj, R. Roy, T. Kailath, Estimation of signal parameters via rotational invariance techniques- Esprit, Proc. 19th Asolimar Conf. on Circuits Systems and Computers, 1985.
- [2] K. Aishima, Consistent estimation with the use of orthogonal projections for a linear regression model with errors in the variables, Linear Algebra Appl., Vol. 684 (2024), 101–126.
- [3] K. Aishima, Strong consistency of an estimator by the truncated singular value decomposition for an errors-in-variables regression model with collinearity, Linear Algebra Appl., in press.
- [4] K. Aishima, Strong consistency of the projected total least squares dynamic mode decomposition for datasets with random noise, Jpn. J. Ind. Appl. Math., Vol. 40 (2023), 691–707.

Taylor 級数法による特異解析と佐藤超関数の応用可能性について

和田 侑佳¹, 瀧 雄也², 石井 良夫³^{1,2}創価大学院理工学研究科情報システム工学専攻, ³創価大学理工学部情報システム工学科¹e23m5308@soka-u.jp, ²e24d5301@soka-u.jp, ³chaos@soka.ac.jp

1.はじめに

1.1 特異点の数値解析における問題点

非線形システムには特異点が存在し, この点は一般解として求められない特異解を意味している. システムの非線形性の要因の一つとして考えられている特異解を数値計算することは非常に困難であるが, 今日では Cauchy の主値積分による特異点を回避した積分路による数値計算を行うことで近似的な結果が得られている[1]. ここで厳密な数値解を得ることが困難な要因の一つとして, 特異点を含む項に零除算が発生することであるが, この問題は超関数を用いることで解析可能性が期待される.

1.2 超関数における記述の有用性

数値解析において, 特異点を含んだ関数を数学的に記述できる形式の一つとして後述する超関数が挙げられる[2]. 超関数を用いることで特異点を含んだ関数を解析的に扱うことが可能であるため, 特異点の数値解析において, 解析精度の向上といった大きな有用性が期待されている.

1.3 研究目的

Cauchy の主値積分の効率的な計算手法として今日では, 平山弘氏らによる Taylor 級数法を用いた解析法が考案されている[3]. これは, 被積分関数を一次の極を持つ部分と見かけ上の特異点を持つ正則な部分に分解したのちに, 正則部分に Taylor 級数法を適用することで高精度かつ効率的な解析を可能としている.

本研究では, Taylor 級数を用いた数値解析法をもとに超関数による数値計算法の構築を試みるとともに, 複素平面での数値積分に対する有用性について検討した.

2 超関数

超関数とは, 一つの関数に対して一つの実数値を対応させる一般的な関数の概念を拡張したもので, 特異性を有する関数や不連続関数を便宜的に扱いやすくした広義の関数である. 1940 年代に数学者の L.シュワルツが超関数(distribution)の理論としてまとめ, 超関数の数学的な定義が厳密に作られたという経緯がある[4].

2.1 δ 関数

超関数の代表例として δ 関数が挙げられる. δ 関数の定義は次式のとおりである.

$$\delta(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \neq 0 \end{cases} \quad (1)$$

物理学者の P.A.M.ディラックが自著「量子力学」に導入したこの関数は, テスト関数 $f(x)$ との畳み込みによって特異点の解析的な記述を可能にしている[5].

$$\int f(x)\delta(x)dx = f(0) \quad (2)$$

2.2 佐藤超関数

1958 年に佐藤幹夫氏によって, 複素平面上に定義された正則関数の組が佐藤超関数である[6]. 佐藤超関数は, 解析関数の組を実軸近傍に収束させた境界値を超関数 $f(x)$ とみなして扱うことから, その定義式は次式として表される.

$$f(x) = \lim_{\varepsilon \rightarrow 0} \{F_+(x + i\varepsilon) - F_-(x - i\varepsilon)\} \quad (3)$$

佐藤超関数は, 設定する解析関数の組に応じて定義できることから, 関数としての表現の

自由度が高く,特異性のある関数を多種多様な関数形式で扱える利点を有している.

3 特異点の数値解析

3.1 Taylor 級数法

Cauchy 主値積分は一次の極をもつ項と見かけ上の特異点を含む項に,すなわち次式として分割できる[7].

$$P.V. \int_a^b \frac{f(x)}{x-\lambda} dx = P.V. \int_a^b \frac{f(\lambda)}{x-\lambda} dx + \int_a^b \frac{f(x)-f(\lambda)}{x-\lambda} dx \quad (4)$$

なお,右辺第一項は対数項とみなして近似する.ここで,右辺第二項は $x = \lambda$ のときに零除算の発生が懸念されるが,これは見かけ上の特異点のため除去が可能である.具体的には $f(x)$ を Taylor 展開したものから $f(\lambda)$ を引いたのちに分母を付加することで零除算の問題が解決可能である.以上の手順から,平山氏は数値計算の高精度かつ高効率化が可能であることを示している.

3.2 佐藤超関数による特異解析の試み

前述した Cauchy の主値積分の分割と Taylor 級数法をもとに超関数を用いた数値計算法の構築及び複素平面での拡張を試みた.まず,一次の極をもつ部分の解析において, Cauchy の主値積分と δ 関数の間で成立する次式の定理をもとに考察する.

$$\frac{1}{x+i\varepsilon} = P.V. \int_a^b \frac{1}{x} dx - i\pi\delta(x) \quad (\varepsilon \rightarrow 0) \quad (5)$$

式(5)から,式(4)の右辺第一項は次のように書き換えることができる.

$$P.V. \int_a^b \frac{f(x)}{x-\lambda} dx = f(x) \left\{ \frac{1}{x-\lambda+i\varepsilon} + i\pi\delta(x-\lambda) \right\} \quad (6)$$

以上のように式展開を行うことで,超関数を用いて一次の極をもつ部分の表現が可能であるといえる.次に,見かけ上の特異点であるが,これは Taylor 級数展開を複素平面に拡張した次式を考察した.

$$f_{\pm}(x) \approx \sum_{k=0}^n \frac{f^{(k)}(\lambda)}{k!} (x-\lambda \pm i\varepsilon)^k \quad (7)$$

式(6)の超関数に佐藤超関数の定義を導入し,式(7)と組み合わせることで,複素平面上に拡張しつつ高精度な主値積分の数値計算が可能ではないかと考えられる.

4 おわりに

本研究では,平山氏の Taylor 級数法を超関数を利用した展開を試みることで,超関数における特異解析の有用性について考察した. Taylor 級数法と佐藤超関数を掛け合わせた展開は,特異解析の有用性だけでなく,複素特異点の解析における応用可能性も期待できるのではないかと考えられる. 今後は,具体的な数値計算を通して定量的に超関数の特異解析における有効性を検証していきたい.

参考文献

- [1] Longman, I. M., “On the Numerical Evaluation of Cauchy principal values of Integrals”, MTAC, 12 (1958), pp. 205–207.
- [2] 今井功, “応用超関数論 I”, サイエンス社, 1981.
- [3] 平山弘, 小宮聖司, “Taylor 級数法による Cauchy の主値積分の数値積分法”, 日本応用数理学会論文誌, Vol. 30, No. 4, 2020, pp. 375–392
- [4] L. Schwartz, “Theorie des distributions”, HERMANN, 1997.
- [5] P.A.M. Dirac, “The Principles of Quantum Mechanics”, Oxford Univ. Press, 1958.
- [6] 佐藤幹夫, “超関数の理論”, 数学 10, 1958, pp. 1–27.
- [7] Kumar, S., “A Note on Quadrature Formulae for Cauchy Principal Value Integral”, J. Inst. Math. Its Appl., 26 (1980), pp. 447–451.

テイラー級数法と超関数によるヒルベルト変換の考察

三谷 大河¹, 瀧 雄也², 石井良夫³

^{1,2}創価大学理工学研究科情報システム工学専攻, ³創価大学理工学部情報システム工学科
e-mail: ¹e24m5317@soka-u.jp, ²e24d5301@soka-u.jp, ³chaos@soka.ac.jp

1 はじめに

様々な信号の解析において,フーリエ変換やウェーブレット変換などが有用とされており,さらにヒルベルト変換は信号の瞬時振幅と瞬時位相の解析において重要な変換の一つである.しかし,これらの変換を用いても信号の不連続点もしくは特異点の解析に関しては,厳密な数値解の導出が困難であると考えられている.

本研究では,テイラー級数法によるコーシーの主値積分の展開をヒルベルト変換への導入を試みることで,特異性解析における数値計算の高精度化について考察した.また,超関数をヒルベルト変換に適用することの応用可能性についても検討した.

2 佐藤超関数について

超関数とは実数値同士を対応させている従来の関数の定義に則っていないにもかかわらず,従来の関数の計算規則に対応することを可能とした関数である[1].

佐藤超関数とは,上半平面と下半平面にある正則関数の実軸近傍における境界値での差であり.佐藤幹夫によって導入された概念および超関数の一つである.これを用いることによってより広範囲の関数を定義することができ,微分方程式や複素解析などに優れている.正則関数 $F(z)$ において佐藤超関数 $f(x)$ は以下のように定義される[2][3].

$$f(x) = F(x + i0) - F(x - i0)$$

3 ヒルベルト変換

ヒルベルト変換 \mathcal{H} とは,実関数を複素解析関数に変換するものであり,次式で表される[4].

$$\mathcal{H}f(x) := \frac{1}{\pi} P.V. \int_{-\infty}^{\infty} \frac{f(\tau)}{x - \tau} d\tau$$

ここで,P.V. はコーシーの主値を意味する.このヒルベルト変換により,信号解析では信号の瞬時振幅と瞬時位相の解析において有用性を示しているが,信号の不連続点もしくは特異点の解析に関しては厳密な数値解の導出が困難である.

4 テイラー級数法を用いたヒルベルト変換

本研究では,ヒルベルト変換について課題となる不連続点や特異点の解析に着目した.具体的にはテイラー級数法と超関数を用いた解析の高精度化について考察した.

4-1 テイラー級数法

コーシーの主値積分は次式で展開されることがわかっている[5].

$$A = P.V. \int_a^b \frac{f(x)}{x - \lambda} dx = P.V. \int_a^b \frac{f(\lambda)}{x - \lambda} dx + \int_a^b \frac{f(x) - f(\lambda)}{x - \lambda} dx$$

ここで式の右辺第1項は,一次の極を持つが,右辺の第二項に関してはテイラー級数法を用いることで見かけ上の特異点の除去が可能になっている[6].以下の手順で展開できる.

まず, $f(x)$ を $x = \lambda$ で n 次まで以下のようにテイラー展開を行う.

$$f(x) \doteq f_0 + f_1(x - \lambda) + f_2(x - \lambda)^2 + f_3(x - \lambda)^3 + \cdots + f_n(x - \lambda)^n$$

ここで $f(\lambda) = f_0$ なので非積分関数をテイラー展開すると次式になる.

$$\frac{f(x) - f(\lambda)}{x - \lambda} \doteq f_1 + f_2(x - \lambda)^2 + f_3(x - \lambda)^3 + \cdots + f_n(x - \lambda)^{n-1}$$

このようにして見かけ上の特異点を除去した解析が可能となる.

4-2 ヒルベルト変換を用いたテイラー級数法

ヒルベルト変換はテイラー級数法によって次式に書き換えることが可能である.

$$\mathcal{H}f(x) = \frac{1}{\pi} \left(P.V. \int_a^b \frac{f(\tau)}{t - \tau} d\tau + \int_a^b \frac{f(t) - f(\tau)}{t - \tau} d\tau \right)$$

従来では第一項を対数項とみなして解析に取り組んでいるが,本研究では第一項に佐藤超関数を用いることによって新たな解析について取り組んだ.具体的には,対数項を佐藤超関数で定義される δ 関数を用いることで,解析の高精度化を試みた.佐藤超関数の δ 関数は次式で表される.

$$f(x) = -\frac{1}{2\pi i} \left(\frac{1}{x + i0} + \frac{1}{x - i0} \right)$$

ここでコーシーの主値と δ 関数の間で成立する定理に,佐藤超関数の δ 関数を適用する.以上の展開から佐藤超関数による対数項の高精度近似を試みた.佐藤超関数は上下両半平面における母関数が正則性を持つ限り,多種多様な関数形式で定義することが可能である.これは δ 関数にとどまらず,導出する超関数によって対数項の更なる高精度近似が可能であると考えられる.

5 おわりに

本研究では,テイラー級数法と超関数を組み合わせた解析法を考察した.これによってヒルベルト変換における信号の特異性について,解析実現性が期待できるのではないかと考えられた.特に佐藤超関数を適用することによって,複素特異点の解析に関しても有用性があると期待できる.今後の展望として,実際に数値計算を行うことで特異性を持った信号の包絡線の検出などにおいて超関数が有用であるか検討を試みる.

参考文献

- [1] L. Schwartz, “Theorie des distributions”, HERMANN, 1997.
- [2] 佐藤幹夫, “超関数の理論”, 数学 10, 1958, pp. 1-27.
- [3] 今井功, “応用超関数論 I”, サイエンス社, 1981.
- [4] 今井功, “応用超関数論 II”, サイエンス社, 1981.
- [5] Kumar, S., “A Note on Quadrature Formulae for Cauchy Principal Value Integral”, J. Inst. Math. Its Appl., 26 (1980), pp. 447-451.
- [6] 平山弘, 小宮聖司, “Taylor 級数法による Cauchy の主値積分の数値積分法”, 日本応用数理学会論文誌, Vol. 30, No. 4, 2020, pp. 375-392

On application of Krylov subspace methods to linear systems $AB\mathbf{x} = \mathbf{b}$ with matrix product structure

Zhixuan XU, Tomohiro SOGABE, Tomoya KEMMOCHI, Shao-Liang ZHANG
 Department of Applied Physics, Graduate School of Engineering, Nagoya University
 e-mail : z-xu@na.nuap.nagoya-u.ac.jp

1 Introduction

Consider linear systems whose coefficient matrix can be written as the product of two matrices:

$$AB\mathbf{x} = \mathbf{b}. \quad (1)$$

Linear systems of the form (1) are being considered in the field of statistics, see [1, 2]. Here to reduce storage, only $A \in \mathbb{R}^{m \times l}$ and $B \in \mathbb{R}^{l \times n}$ can be accessed instead of the full system $C \in \mathbb{R}^{m \times n}$ ($C = AB$). In this article, we assume $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times n}$ to be nonsingular large and sparse matrices, and $\mathbf{b} \in \mathbb{R}^n$ is a right hand side vector. For the above problem, there are three equivalent systems to consider. We present four approaches to solving (1) based on these equivalent systems using Krylov subspace methods.

To calculate the Krylov subspace of the **original system** (1), in each iteration, several matrix-vector products in the form $AB\mathbf{x}$ are required. A common approach computes $AB\mathbf{x}$ in the order $A(B\mathbf{x})$, which is later written as ORI. Note that in this article, the approach $(AB)\mathbf{x}$ is also considered for comparison, written as ORI-AB.

For the **subsystems**

$$\begin{cases} A\mathbf{y} = \mathbf{b}, \\ B\mathbf{x} = \mathbf{y}, \end{cases} \quad (2)$$

we apply KS methods first to the subsystem $A\mathbf{y} = \mathbf{b}$, then to the subsystem $B\mathbf{x} = \mathbf{y}$. This approach is written as SUB.

The **augmented system**

$$\mathcal{A}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}, \quad (3)$$

where

$$\mathcal{A} = \begin{bmatrix} I & -B \\ A & O \end{bmatrix}, \tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix}, \tilde{\mathbf{b}} = \begin{bmatrix} \mathbf{0} \\ \mathbf{b} \end{bmatrix}$$

can also be used for computations. This approach is noted as AUG.

Our purpose is to clarify the characteristics of KS methods on these equivalent systems.

2 Numerical Experiment

Here we provide the iterations and true relative residual of different approaches in Figs. 1-2. In total, 19 pairs of matrices selected from [3] are tested. In this experiment, preconditioned GPBi-CG method [4] with ILUTP preconditioner was used, and the threshold was set to 1×10^{-5} .

From the results shown in Fig. 1, we discuss the speed of convergence. ORI, SUB and AUG approaches converge really well with appropriate preconditioning, and the iteration is less than 10 for almost cases. ORI-AB fails for five problems, indicating its inferiority.

We now consider the true relative residual shown in Fig. 2. True residual of ORI approach is less than the precision $\varepsilon = 10^{-8}$ for 15 problems, while those of SUB, AUG and ORI-AB approaches are less than ε for 11 problems. ORI approach is the most accurate under our testing conditions.

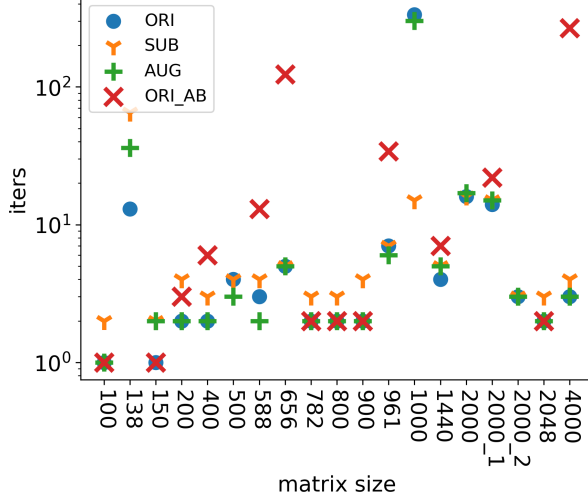


Fig.1. Iterations of different approaches

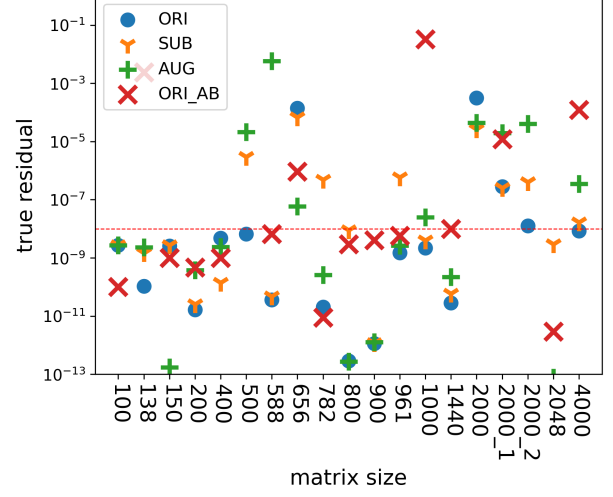


Fig.2. True residual of different approaches

3 Discussions

The above experiment prepares ground for the following discussion of characteristics of different systems. If AB possesses well-condition, ORI is worth considering in its convergence rate. SUB is a good choice if AB is hard to solve. However, this method can face accuracy issues. In our problems, SUB loses three digits for the problem of A (olm500, $n = 500$) and B (tomography, $n = 500$), while ORI causes no digit loss. This is because the accuracy of two subsystems doesn't necessarily means the accuracy of (1). The dimension of \mathcal{A} is twice as much as that of A and B , which will probably cause slower convergence for AUG. However, the structured block form can be exploited by preconditioning, which speeds up the convergence.

4 Conclusion

In this article, we show the characteristics of the 4 approaches presented. This will provide useful information for future algorithm development.

Acknowledgment This work was supported by JSPS KAKENHI Grant Numbers JP20H00581 and JP20K20397.

References

- [1] Anna Ma et al., SIAM J. Matrix Anal. & Appl. 39 (2018) 104–122.
- [2] Kui Du, Appl. Math. Comput. 466 (2024), 128468, 13 pp.
- [3] Timothy Davis et al., ACM Trans.Math. Softw. 38 (2011) 1:1-1:25.
- [4] Shao-Liang Zhang, SIAM J. Sci. Comput. 18 (1997) 537-551.