行列関数計算のためのクリロフ部分空間法

数理第 2 研究室 佐藤 暁史　指導教員 室田 一雄 教授

2009 年 2 月 2 日

1 はじめに

本研究では，行列 $A \in \mathbb{C}^{n\times n}$ とベクトル $\boldsymbol{b} \in \mathbb{C}^n$，そして $f(A)$ が定義されるような関数 $f : \mathbb{C} \subset D \to \mathbb{C}$ が与えられているときに，

$$f(A)\boldsymbol{b} \tag{1}$$

を近似的に評価する方法を扱う．この計算の応用においては，多くの場合行列 A は大規模な疎行列であり，なんらかの構造を有している．この場合，最初に $f(A)$ の計算をすることによって式 (1) を評価するのは不可能であることが多く，よって残りの計算をするために利用可能な様々なアルゴリズムを使うことができない．$f(A)$ の計算を行わず，式 (1) を直接的に得るための標準的なアプローチは，初期ベクトル $\boldsymbol{b}$ による A のクリロフ部分空間に基づいた方法である．この方法による長所は，行列とベクトルの掛け算にのみ A が必要となるということと，f が滑らかな関数ならば超一次収束をするということである．しかし，クリロフ部分空間による近似には一つの欠点がある．それは，次元数だけのベクトルをすべて保存する必要があるということである．ゆえに，しばしばメモリの制限によって，解ける問題のサイズが限定されてしまう．クリロフ部分空間近似による非エルミートの線形方程式系の一般的な解法は，クリロフ空間がある最大の次元に達するたびにアルゴリズムを再出発して保存領域を制限することである．Eiermann ら [3] は，有限個のベクトルを記憶するだけで計算できる再出発クリロフ部分空間近似を提案した．

Eiermann らのアルゴリズムのうちの特殊な場合は Afanasjew ら [1] によって最急降下法の一般化になっていることが示されている．Afanasjew らの解析が赤池 [2] による解析の一般化になっていることも述べる．

最後に，以上に述べたアルゴリズムにおいて，赤池のアルゴリズムを実際に実装し，解析結果を確かめる．

2 行列関数

$\Lambda(A) = \{\lambda_1, \lambda_2, \ldots, \lambda_k\}$ を，行列 $A \in \mathbb{C}^{n\times n}$ の相違なる k 個の固有値とする．複素関数 f が与えられているとき，行列 $f(A)$ は，$f^{(r)}(\lambda_j)$ が $r = 0, 1, \ldots, n_j - 1; j = 1, 2, \ldots, k$ において存在する場合に定義される．この場合，$q_{f,A} \in \mathcal{P}_{K-1}$ がたかだか $K-1$ 次の一意に定まる多項式であるとして $f(A) := q_{f,A}(A)$ と表すことができる．ここで，

$q_{f,A}(A) \in P_{K-1}$ は，$q^{(r)}_{f,A}(\lambda_j) = f^{(r)}(\lambda_j), \quad r = 0, 1, \ldots, n_j - 1,\ j = 1, 2, , \ldots, k$ を満たす一意に定まる多項式とする．

3 再出発クリロフ部分空間近似

$A \in \mathbb{C}^{n\times n}$ と $\mathbf{0} \neq \boldsymbol{b} \in \mathbb{C}^n$ で与えられるクリロフ部分空間を $\mathcal{K}_m(A, \boldsymbol{b}) := \mathrm{span}\{\boldsymbol{b}, A\boldsymbol{b}, \ldots, A^{m-1}\boldsymbol{b}\}$ と定義する．$f(A)\boldsymbol{b}$ の計算のための再出発クリロフ部分空間法は以下のようになる．

Given: $A, \boldsymbol{b}, f$
$f^{(0)} := f, \boldsymbol{f}^{(0)} := \mathbf{0}, \boldsymbol{b}^{(0)} := \boldsymbol{b}, \gamma^{(0)} := \|\boldsymbol{b}\|$
for $k = 1, 2, \ldots$ until convergence **do**
　$\mathcal{K}_m(A, \boldsymbol{b}^{(k-1)})$ の Arnoldi 分解 $AV^{(k)}_m = V^{(k)}_m H^{(k)}_m + \mu^{(k)}_{m+1,m}\boldsymbol{v}^{(k)}_{m+1}\boldsymbol{e}^{\top}_m$ を計算する．
　近似 $\boldsymbol{f}^{(k)}$ を更新する．
end for

4 再出発長 1 の場合とその性質

A を正則な $n \times n$ 実数行列，x, b を n 次元 実数ベクトル，x を変数とする連立一次方程式 $Ax = b$ の解を x^* とする．$Ax = b$ を解く最急降下法を以下の手順により定義する：

Step 0 x_0 を適当な n 次元ベクトルとし，$k=0$ とする．

Step 1 $A\epsilon_k = Ax_k - b$ とする．

Step 2 $\zeta_k = A^\top A\epsilon_k$ とする．

Step 3 $\gamma_k = \frac{\epsilon_k^\top A^\top A\zeta_k}{\zeta_k^\top A^\top A\zeta_k}$ とする．

Step 4 $x_{k+1} = x_k - \gamma_k\zeta_k$ とする．

Step 5 終了条件を満たせば終了．

Step 6 k の値を 1 増やして Step 1 へと戻る．

これに関して，赤池によって，以下の定理が示されている．

定理 4.1.

1. ϵ_k は $A^\top A$ の 2 つの固有ベクトルの線形結合に近づいていく．
2. ϵ_k は漸近的に 2 つの特定の方向へ近づく．

この性質をみると，ζ_k と ζ_{k-2} がほとんど同じ方向を向いているとき，$x_{k-2} - x_k$ の方向を用いて x_{k+1} を計算すると良いことがわかる．すなわち，$\hat{\gamma}_k = \frac{\epsilon_k^\top A^\top A(x_{k-2}-x_k)}{(x_{k-2}-x_k)^\top A^\top A(x_{k-2}-x_k)}$ とし，$x_{k+1} = x_k - \hat{\gamma}_k(x_{k-2} - x_k)$ によって，x_{k+1} を計算する．これを加速手続きとする．

再出発長 1 の場合，再出発クリロフ部分空間法は

Given: $A, \boldsymbol{b}$
$\sigma_1 := \|\boldsymbol{b}\|, v_1 := \boldsymbol{b}/\sigma_1$
for $k = 1, 2, \dots$ **do**
  $\boldsymbol{w} := A\boldsymbol{v}_k;\ \rho_k := \boldsymbol{v}_k^H\boldsymbol{w}$
  $\boldsymbol{w} := \boldsymbol{w} - \rho_k\boldsymbol{v}_k;\ \sigma_{k+1} := \|\boldsymbol{w}\|$
  $\boldsymbol{v}_{k+1} := \boldsymbol{w}/\sigma_{k+1}$
end for

となる．これに関して，次の定理が示された．

定理 4.2. A：をエルミート行列，$\lambda_{\min}$ を最小固有値，$\lambda_{\max}$ を最大固有値とし，$\boldsymbol{b}$ は $\lambda_{\max}$ や $\lambda_{\min}$ に対応する固有ベクトルを含むとする．さらに $\theta \in (0,1)$ であるとする．この時以下が成立する．

$$\begin{aligned}\lim_{k\to\infty} \rho_{2k-1} &= \theta\lambda_{\min} + (1-\theta)\lambda_{\max} =: \rho_1^*,\\ \lim_{k\to\infty} \rho_{2k} &= (1-\theta)\lambda_{\min} + \theta\lambda_{\max} =: \rho_2^*,\\ \lim_{k\to\infty} \sigma_{k+1} &= \sqrt{\theta(1-\theta)}(\lambda_{\max} - \lambda_{\min}) =: \sigma^*.\end{aligned}$$

これは赤池の結果の一般化である．

5 数値計算

以下の値を用い，線型方程式を解くための最急降下法について，赤池の性質を確かめる数値計算を行った．

- $A = \begin{pmatrix} \sqrt{0.00268704} & & & & & \\ & \sqrt{0.01581310} & & & 0 & \\ & & \sqrt{0.08234830} & & & \\ & & & \sqrt{0.17590130} & & \\ & 0 & & & \sqrt{0.25946632} & \\ & & & & & \sqrt{0.49823436} \end{pmatrix}$
- $b = 0$, $x_0 = (1,1,1,1,1,1)^\top$

結果の一部を図 1 に示す．

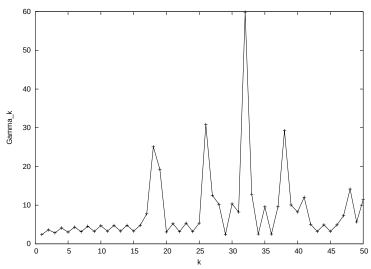

図 1 γ_k の値

参考文献

[1] M. Afanasjew, M. Eiermann, O. G. Ernst, and S. Güttel. A generalization of the steepest descent method for matrix functions. *Electronic Transactions on Numerical Analysis*, Vol. 28, pp. 206–222, 2008.

[2] H. Akaike. On a successive transformation of probability distribution and its application to the analysis of the optimum gradient method. *Annals of the Institute of Statistical Mathematics*, Vol. 11, pp. 1–17, 1959.

[3] M. Eiermann and O. G. Ernst. A restarted krylov subspace method for the evaluation of matrix functions. *SIAM J. Numer. Anal.*, Vol. 44, No. 6, pp. 2481–2504, 2006.