共役勾配法 (CG 法) の収束性*

梅谷　信行

平成 22 年 11 月 2 日

目 次

1 チェビシェフ多項式による収束性の予測

ここで、$\|\mathbf{x}\|_\mathbf{A}$ は次のように定義される作用素ノルムである。

$$\|\mathbf{x}\|_\mathbf{A}^2 = (\mathbf{x}, \mathbf{A}\mathbf{x}) \tag{1}$$

CG 法は Krylov 部分空間 $\mathcal{K}_\Updownarrow$ の中で誤差の $\mathbf{A}$ ノルムを最小化する手法であった。つまり、方程式の解を $\mathbf{x}^*$、m 回反復後の近似解を $\mathbf{x}_m$ とすると、

*これは忘れっぽい著者が昔勉強したことを書き留めておく備忘録です．きっと間違えを多く含んでいます．すいません．ご指摘ご意見などありましたら教えてもらえると有り難いです．

$\mathbf{x}_m - \mathbf{x}_0$ が Krylov 部分空間に含まれることから、

$$\mathbf{x}_m = \mathbf{x}_0 + q_m(\mathbf{A})r_0 \tag{2}$$

とかける、但し、多項式 q_m は $m-1$ 次である。m 回反復目の CG 法の誤差 $\mathbf{e}_m$ の $\mathbf{A}$ ノルム $\|\mathbf{e}_m\|_\mathbf{A}$ は、

$$\begin{aligned}
\|\mathbf{e}_m\|_\mathbf{A} &= \|\mathbf{x}^* - \mathbf{x}_m\|_\mathbf{A} && (3)\\
&= \|\mathbf{x}^* - (\mathbf{x}_0 + q_m(\mathbf{A})r_0)\|_\mathbf{A} && (4)\\
&= \|\mathbf{e}_0 - q_m(\mathbf{A})\mathbf{A}\mathbf{e}_0\|_\mathbf{A} && (5)\\
&= \|(\mathbf{I} - \mathbf{A}q_m(\mathbf{A}))\mathbf{e}_0\|_\mathbf{A} && (6)
\end{aligned}$$

のように初期誤差 $\mathbf{e}_0$ を用いて表すことができる。ここで、$\mathbf{e}_0 = \mathbf{A}\mathbf{e}_0$ を用いた。q_m はこれを $m-1$ 次の多項式全体 P_{m-1} の中で最小化するので

$$\begin{aligned}
\|\mathbf{e}_m\|_\mathbf{A} &= \|(\mathbf{I} - \mathbf{A}q_m(\mathbf{A}))\mathbf{e}_0\|_\mathbf{A} && (7)\\
&= \min_{q\in P_{m-1}} \|(\mathbf{I} - \mathbf{A}q(\mathbf{A}))\mathbf{e}_0\|_\mathbf{A} && (8)
\end{aligned}$$

となる。ここで簡単のため $\mathbf{I} - \mathbf{A}q(\mathbf{A}) = r(\mathbf{A})$ とおく。但し、r は $r(0) = 1$ を満たす m 次の多項式である。これを用いると上の式は

$$\|\mathbf{e}_m\|_\mathbf{A} = \min_{r\in P_m \quad r(0)=1} \|r(\mathbf{A})\mathbf{e}_0\|_\mathbf{A} \tag{9}$$

と表すことができる。

1.1　固有値による誤差の A ノルムの評価

ここで、$\lambda_1, \ldots, \lambda_n$ は $\mathbf{A}$ の固有値であるとし、$\mathbf{x}_1, \ldots, \mathbf{x}_n$ は固有ベクトルであるとする。$\xi_1, \ldots, \xi_n$ は $\mathbf{e}_0$ を固有ベクトルの線形結合で表した時の係数であるとする

$$\begin{aligned}
\|r(\mathbf{A})\mathbf{e}_0\|_\mathbf{A} &= \|\sum_i^n r(\lambda_i)\xi_i\mathbf{x}_i\|_\mathbf{A} && (10)\\
&\leq |\sum_i^n r(\lambda_i)| \quad \|\sum_i^n \xi_i\mathbf{x}_i\|_\mathbf{A} && (11)
\end{aligned}$$

$$\leq \max_{i} \quad |r(\lambda_i)| \quad \|\mathbf{e}_0\|_{\mathbf{A}} \tag{12}$$

$$\leq \max_{\lambda \in [\lambda_{min}, \lambda_{max}]} \quad |r(\lambda)| \quad \|\mathbf{e}_0\|_{\mathbf{A}} \tag{13}$$

これを用いると m 回反復時における誤差の $\mathbf{A}$ ノルム $\|\mathbf{e}_m\|_{\mathbf{A}}$ は次のように評価される。

$$\|\mathbf{e}_m\|_{\mathbf{A}} \leq \min_{q \in P_m \quad r(0)=1} \quad \max_{\lambda \in [\lambda_{min}, \lambda_{max}]} \quad |r(\lambda)| \quad \|\mathbf{e}_0\|_{\mathbf{A}} \tag{14}$$

つまり、$r(0) = 1$ を満たす m 次の多項式 r の中で、区間 $\lambda \in [\lambda_{min}, \lambda_{max}]$ での最大の絶対値を最小値によって、誤差の $\mathbf{A}$ ノルムは抑えられる

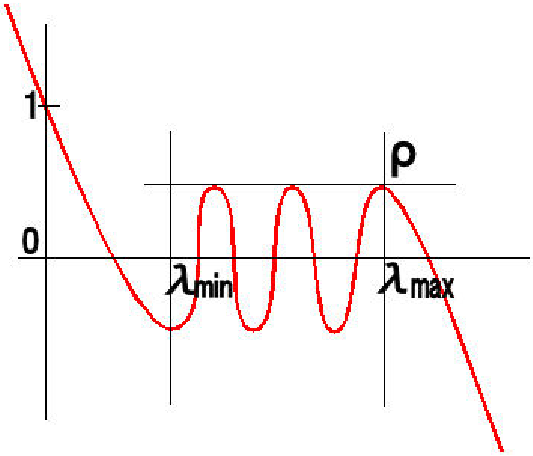

図 1: 誤差の評価関数 (イメージ).λ_{min} から λ_{max} の間で関数の絶対値の最大値が最小となるその時の最小値が ρ

1.2 チェビシェフ多項式

チェビシェフ多項式はある区間で絶対値の最大値を最小化するという性質をもった多項式である。具体的には n 次のチェビシェフ多項式 $C_n(x)$ は最高次の係数が 1 であるような任意の n 次の多項式 $p_n(x)$ の中で

$$\max_{|x|<1} |p_n(x)| \tag{15}$$

が最小になるような多項式を 2^{n-1} 倍したものと一致する。チェビシェフ多項式は次のように定義される

$$C_k(t) = \begin{cases} \cos(k\cos^{-1} t) & |t| \le 1 \\ \cosh(k\cosh^{-1} t) & |t| > 1 \end{cases} \tag{16}$$

一見複雑な超越関数に見えるが実はこれは単純な多項式である。チェビシェフ多項式同士には次のような関係がある。$C_{k+1}(t) = 2tC_k(t) - C_{k-1}(t)$ これを使えば $C_1(t) = t$、$C_2(t) = 2t^2 - 1$、$C_3(t) = 4t^3 - 3t$、$C_4(t) = 8t^4 - 8t^2 + 1$、$\cdots$ となることがわかる。チェビシェフ多項式は次のようにも表すことができる。

$$C_k(t) = \frac{1}{2}\left[(t + \sqrt{t^2 - 1})^k + (t + \sqrt{t^2 - 1})^{-k}\right] \tag{17}$$

チェビシェフ多項式の絶対値は区間 [-1,1] で最大値 1 をとることが知られている。

1.3 チェビシェフ多項式の変数変換とスケーリング

k 次のチェビシェフ多項式 C_k は区間 [-1,1] の間の絶対値の最大値を最小化する最高次数の係数が 2^{k-1} の k 次の多項式であった。チェビシェフ多項式の最適化区間を $[-1, 1]$ から $[\alpha, \beta]$ に変換するために、$t = 1 + 2\frac{t'-\beta}{\beta-\alpha}$ という変数変換を用いたものを $\bar{C}_k$ とすると

$$\bar{C}_k(t) = C_k\left(1 + 2\frac{t - \beta}{\beta - \alpha}\right) \tag{18}$$

となる。これにさらに適当なスケーリングを行って、引数が γ の時に値 1 を取るようにした関数 $\hat{C}_k$ とすると

$$\hat{C}_k(t) = \frac{\bar{C}_k(t)}{\bar{C}_k(\gamma)} = \frac{C_k(1 + 2\frac{t-\beta}{\beta-\alpha})}{C_k(1 + 2\frac{\gamma-\beta}{\beta-\alpha})} \tag{19}$$

$\hat{C}_k$ は区間 $[\alpha, \beta]$ の間の絶対値の最大値を最小化する引数 γ において 1 をとる k 次の多項式である。$C_k(1 + 2\frac{t-\beta}{\beta-\alpha})$ の絶対値は区間 $[\alpha, \beta]$ で最大値 1 をとることから $\hat{C}_k(t)$ の絶対値は区間 $[\alpha, \beta]$ で最大値 $\frac{1}{|C_k(1+2\frac{\gamma-\beta}{\beta-\alpha})|}$ をとる。つまり、

$$\min_{p \in P_k, p(\gamma)=1} \max_{t \in [\alpha, \beta]} |p(t)| = \frac{1}{|C_k(1 + 2\frac{\gamma-\beta}{\beta-\alpha})|} \tag{20}$$

1.4　CG法の収束性の評価

引数0のとき1をとるm次の多項式$r(\lambda)$の中で、区間$[\lambda_{min}, \lambda_{max}]$において絶対値の最大値の最小値は、上の式の$\alpha = \lambda_{max}$、$\beta = \lambda_{min}$、$\gamma = 0$とおくことで

$$\min_{r \in P_k, r(0)=1} \max_{\lambda \in [\lambda_{min}, \lambda_{max}]} |r(\lambda)| = \frac{1}{|C_m(1 + 2\frac{-\lambda_{max}}{\lambda_{max}-\lambda_{min}})|} \tag{21}$$

$$= \frac{1}{|C_m(1 + 2\frac{\lambda_{min}}{\lambda_{max}-\lambda_{min}})|} \tag{22}$$

$$= \frac{1}{C_m(1 + 2\frac{\lambda_{min}}{\lambda_{max}-\lambda_{min}})} \tag{23}$$

$$= \frac{1}{C_m(1 + 2\eta)} \tag{24}$$

ここで、$\eta = \frac{\lambda_{min}}{\lambda_{max}-\lambda_{min}}$とおいた。また、$C_m(t) > 1 \quad (\quad for \quad t > 1 \quad)$を用いた。ここから、$C_m(1 + 2\eta)$を評価する

$$C_m(t) = \frac{1}{2}\left[(t + \sqrt{t^2 - 1})^m + (t + \sqrt{t^2 - 1})^{-m}\right] \geq \frac{1}{2}(t + \sqrt{t^2 - 1})^m \tag{25}$$

を用いる。

$$C_m(1 + 2\eta) \geq \frac{1}{2}(1 + 2\eta + \sqrt{(1 + 2\eta)^2 - 1})^m \tag{26}$$

$$= \frac{1}{2}(1 + 2\eta + 2\sqrt{\eta(\eta + 1)})^m \tag{27}$$

$$= \frac{1}{2}\{(\sqrt{\eta} + \sqrt{1 + \eta})^2\}^m \tag{28}$$

$$(\sqrt{\eta} + \sqrt{1 + \eta})^2 = \frac{(\sqrt{\lambda_{min}} + \sqrt{\lambda_{max}})^2}{\lambda_{max} - \lambda_{min}} \tag{29}$$

$$= \frac{\sqrt{\lambda_{min}} + \sqrt{\lambda_{max}}}{\sqrt{\lambda_{max}} - \sqrt{\lambda_{min}}} \tag{30}$$

$$= \frac{\sqrt{\kappa} + 1}{\sqrt{\kappa} - 1} \tag{31}$$

を用いる。但し、$\kappa = \frac{\lambda_{max}}{\lambda_{min}}$とおいた。$\kappa$は正定値対称行列のスペクトル条件数(Spectorum Condition Number)である。

$$C_m(1+2\eta) \geq \frac{1}{2}\left(\frac{\sqrt{\kappa}+1}{\sqrt{\kappa}-1}\right)^m \tag{32}$$

$$\min_{r\in P_k, r(0)=1} \max_{\lambda\in[\lambda_{min},\lambda_{max}]} |r(\lambda)| \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^m \tag{33}$$

これを用いると,$\|\mathbf{e}_m\|_\mathbf{A} \leq 2\left(\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right)^m \|\mathbf{e}_0\|_\mathbf{A}$ となり、誤差の $\mathbf{A}$ ノルムに対して上のように評価できることがわかる。

このことから **条件数が小さいほど収束が早い** ということがわかる。

2　反復数と条件数

上の式から、予測される収束率 ρ_e は次のようにかける。

$$\rho_e = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} \tag{34}$$

これは κ が十分大きい場合、次のように変形される。

$$\rho_e = \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1} = \frac{(\sqrt{\kappa}+1)-2}{\sqrt{\kappa}+1} = 1-\frac{2}{\sqrt{\kappa}+1} \simeq 1-\frac{2}{\sqrt{\kappa}} \tag{35}$$

例えば κ が a^2 倍されたとしたときに収束率 ρ_e' がもとの収束率 ρ_e からどのように近似的に求められるかを考える。

$$\rho_e' = 1-\frac{2}{\sqrt{a^2\kappa}} = 1-\frac{1}{a}\frac{2}{\sqrt{\kappa}} \simeq (1-\frac{2}{\sqrt{\kappa}})^{\frac{1}{a}} = \rho_e{}^{\frac{1}{a}} \tag{36}$$

ここで κ が十分大きいという近似を用いた。κ が a^2 倍されると収束率がだいたい $\frac{1}{a}$ 乗されることがわかる。これは同じ収束を得るまでに a 倍の反復数を要するということを意味する。

2.1　メッシュサイズとの関係

ラプラス演算子を有限要素法離散化した場合、メッシュサイズと条件数 κ はだいたい次のような関係にあることが知られている。

$$\kappa \propto \frac{1}{h^2} \tag{37}$$

このことから例えばメッシュサイズが a 倍されるとだいたい次のことがいえる。
h が a 倍　$\rightarrow$ 条件数 κ が $1/a^2$ 倍　$\rightarrow$ 収束率が a 乗　$\rightarrow$ 反復数が $1/a$ 倍

つまり、**メッシュサイズが半分になると反復数がだいたい２倍になる**ということがわかる。

3　収束の超線形性

CG 法の収束は超線形性を示す。収束が超線形であるとは、対数グラフで残差の変化を表したときにグラフが直線的ではなく徐々に加速して始めより急な収束を示すことをいう。

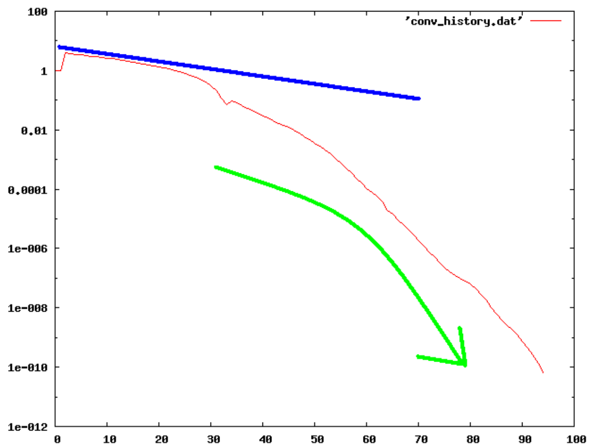

図 2: 典型的な CG 法の残差収束の様子

図 3 に典型的な CG 法の残差収束の様子を示す．青い線は始めの収束の様子を表している。緑の線のようにあきらかに収束が加速しているのがわかる。このような超線形の収束は Jacobi 法や Gauss-Sidel 法や古典的などの古典的な反復法ではみられず、Krylov 部分空間法ならではのものである。上で行ったチェビシェフ多項式による収束の予測では線形の収束しか予測できないので、実際のところ上の式で得られる収束よりも現実の方がずっと早い収束が得られることが多い。

4　参考にしたもの

[1]

参考文献

[1] Saad, Y.: *Iterative Methods for Sparse Linear Systems, Second Edition*, Society for Industrial and Applied Mathematics, 2 edition (2003).