代数方程式を解く平野法[1] (Advanced Topic)

Newton 法は局所的収束性に関しては理論上も実用上も十分であるが，必ずしも大域的収束性をもたないという難点がある．しかし，対象を代数方程式に限れば，Newton 法を少し修正することで，局所的収束速度を損なうことなく大域的収束性を保証することができる．1960 年代後半に平野菅保[2] と B. Dejon–K. Nickel[3] によって独立に提案された算法 (我が国では**平野法**と呼ばれる) は，素直なアイデアに基づいて Newton 法を改良した分かりやすい実用的算法である．

平野法の導出

複素数を係数とする n 次多項式 $p(z)$ が

$$p(z) = \sum_{k=0}^{n} a_{n-k} z^k \tag{1}$$

の係数 a_k $(k = 0, 1, \ldots, n)$ によって与えられているとする．平野法は，任意の初期値 $z^{(0)}$ から出発して，

$$z^{(\nu+1)} = z^{(\nu)} + \Delta z^{(\nu)} \qquad (\nu = 0, 1, \ldots)$$

の形の更新をおこないながら $p(z)$ のどれかの根に収束する近似解列 $\{z^{(\nu)}\}$ を生成する．

修正量 $\Delta z^{(\nu)}$ はつぎのように考えて定める．$p(z)$ を $z = z^{(\nu)}$ の周りで

$$p(z^{(\nu)} + \zeta) = \sum_{k=0}^{n} c_{n-k} \zeta^k \tag{2}$$

のように展開する ($c_n = p(z^{(\nu)})$ に注意)．通常の Newton 法では，$z^{(\nu)}$ が根に十分近いことを前提にして式 (2) の右辺から $k = 0$ と $k = 1$ に対応する項だけをとり出して

$$p(z^{(\nu)} + \zeta) \simeq c_{n-1}\zeta + c_n \tag{3}$$

と近似し，これを 0 にする値 $\zeta = -c_n/c_{n-1}$ を修正量としている．これに対し，平野法では，$k = 1, \ldots, n$ について

$$p(z^{(\nu)} + \zeta) \simeq c_{n-k}\zeta^k + c_n \tag{4}$$

という 2 項近似を考え，これから定まる

$$\zeta_k = (-c_n/c_{n-k})^{1/k} \tag{5}$$

[1]本資料は，杉原正顯, 室田一雄：数値計算法の数理, §5.1, 岩波書店, 1994 を編集したものである．

[2]平野菅保: 代数方程式の解法および誤差，第 8 回プログラミングシンポジウム報告集，情報処理学会, 1967.

[3]B. Dejon and K. Nickel: A never failing fast convergent root-finding algorithm, *Constructive Aspects of the Fundamental Theorems of Algebra* (B. Dejon and P. Henrici, eds.), John Wiley, 1969, pp.1–35.

($c_{n-k}=0$ のときは $\zeta_k=\infty$ とおく) を修正量の候補とし，それらの中で絶対値の最小のもの (これを ζ_m と記す) を修正量に採用する．最小の修正量を与える項が最も影響力のある項であって，この修正により関数値は減少するだろうと考えてのことである．

実際，次の補題のように，$|\zeta_m|$ が余裕をもって最小ならばこれが成り立つ．なお，$z^{(\nu)}$ が根 α の十分よい近似根ならば $m=1$ となり，通常の Newton 法に一致する．

補題 1 $0<\beta<1$ とする．

$$|\zeta_m| \le \frac{\beta}{1+\beta}|\zeta_k| \qquad (1\le k\le n, k\ne m) \tag{6}$$

ならば

$$|p(z^{(\nu)}+\zeta_m)| \le \beta|p(z^{(\nu)})| \tag{7}$$

である．

[証明] $1\le k\le n, k\ne m$，$c_{n-k}\ne 0$ であるような k に関する和を $\sum_k{}^*$ とかくと，$c_{n-m}{\zeta_m}^m+c_n=0$ だから

$$\begin{aligned}|p(z^{(\nu)}+\zeta_m)| &\le \sum_k{}^*|c_{n-k}||\zeta_m|^k\\ &\le \sum_k{}^*|c_{n-k}||\zeta_k|^k[\beta/(1+\beta)]^k = \sum_k{}^*|c_n|[\beta/(1+\beta)]^k \le \beta|c_n|. \qquad \square\end{aligned}$$

平野法の算法

平野法の修正量 $\Delta z^{(\nu)}$ は，さらに減速の手法も取り入れて，つぎのように定める．$0<\beta<1$, $\lambda>1$ は，減速の仕方をきめるパラメタである．例えば，$\beta=3/4$, $\lambda=2$ とする．

平野法の第 ν 段 (基本形)

1°: c_k $(k=0,1,\ldots,n)$ を組立除法 (Horner 法) で計算する; $\mu:=1$.

2°: $\zeta_k:=(-\mu c_n/c_{n-k})^{1/k}$ $\quad(k=1,\ldots,n)$.

3°: $|\zeta_k|$ が最小である k $(1\le k\le n)$ を m とする[4].

4°: $|p(z^{(\nu)}+\zeta_m)|\le(1-(1-\beta)\mu)|c_n|$ $\quad$ ならば，$\Delta z^{(\nu)}:=\zeta_m$ として第 ν 段を終了し，そうでなければ，$\mu:=\mu/\lambda$ として 2° に戻る． $\square$

上記 2° における k 乗根の分枝は，$|z^{(\nu)}+\zeta_k|$ が最も小さくなるようなものを選ぶ．具体的には，

$$2\pi\varphi=\arg z^{(\nu)}, \qquad 2\pi\psi_k=\arg(-c_n/c_{n-k}) \quad (0\le\varphi,\psi_k<1) \tag{8}$$

として，$k(0.5-\varphi)-\psi_k$ に最も近い整数 j_k を求め，

$$\zeta_k:=|\mu c_n/c_{n-k}|^{1/k}\exp[2\pi\imath(\psi_k+j_k)/k] \tag{9}$$

とする．なお，初期値 $z^{(0)}$ は求めたい根の近似値に選ぶのがよいが，それが未知の場合には $z^{(0)}=0$ とする．

[4] $|\zeta_k|$ を最小にする k が複数個ある場合には，k の値の最小のものを m とする．

表 1: 平野法 (基本形) による代数方程式の解法 (例 1: $p(z) = z^3 - 3z + 3$)
$\alpha_1 = -2.1038034\cdots$, $\alpha_2 = \overline{\alpha_3} = 1.0519017\cdots - \mathrm{i}\ 0.56523585\cdots$

ν	$z^{(\nu)}$		$p(z^{(\nu)})$		$\lvert p(z^{(\nu)})\rvert$	$m^{(\nu)}$	$\mu^{(\nu)}$
	実部	虚部	実部	虚部			
0	2.50000	0.0	11.1250	0.0	11.1	1	1
1	1.79365	0.0	3.38955	0.0	3.4	1	1
2	1.28406	0.0	1.26500	0.0	1.3	2	1
3	1.28406	−0.573048	0.0	−0.92723	0.93	1	1
4	1.08355	−0.529389	0.11052	−0.12810	0.17	1	1
5	1.04959	−0.564645	0.360E−2	0.785E−2	0.86E−2	1	1
6	1.05191	−0.565232	0.114E−4	−0.153E−4	0.19E−4	1	1
7	1.05190	−0.565236	0.0	0.9E−6	0.9E−6	−	−

例 1 3 次式 $p(z) = z^3 - 3z + 3$ に $z^{(0)} = 2.5$ として平野法を適用する (これは普通の Newton 法ではうまく解けない問題である). $z^{(1)}$ と $z^{(2)}$ は通常の Newton 法と同じ値であるが, $\nu = 2$ で $m^{(2)} = 2$ となって $z^{(3)}$ は複素数となり, $z^{(\nu)}$ は複素根

$$\alpha_2 = 1.0519017\cdots - \mathrm{i}\ 0.56523585\cdots$$

に 2 次収束する (→表 1). $|p(z^{(\nu)})|$ が単調に減少しているのが観察される. $z^{(7)}$ の共役複素数として α_2 の共役根 α_3 が求められ, 根 (解) と係数の関係 $\alpha_1 = -3/(\alpha_2\alpha_3) = -3/|\alpha_2|^2$ から α_1 の近似値 -2.10381 が求められる. □

平野法の算法において, m は通常小さな値になることが経験的に知られている. そこで, すべての c_k $(k = 0, 1, \ldots, n)$ を最初から計算するのではなく, (組立除法 (Horner 法) のための 1 次元配列を保持しながら) c_k を $k = n, n-1, \ldots$ の順に必要に応じて計算していく方が実際的には効率がよい.

平野法の第 ν 段 (実用的変種)

0°: c_n, c_{n-1} を組立除法 (Horner 法) で計算する; $\zeta_1 := -c_n/c_{n-1}$;
$|p(z^{(\nu)} + \zeta_1)| \le \beta|c_n|$ ならば, $\Delta z^{(\nu)} := \zeta_1$ として第 ν 段を終了し, そうでなければ, $\mu := 1/\lambda$, $K := 2$ として 1° に進む.

1°: $K \le n$ ならば, 組立除法 (Horner 法) を続けて c_{n-K} を計算し, $\zeta_K := (-\mu c_n/c_{n-K})^{1/K}$ とする (分枝のとり方は (9) による).

2°: $\zeta_k := \zeta_k/\lambda^{1/k}$ $\quad (k = 1, \ldots, K-1)$; $K > n$ ならば $K := n$ とする.

3°: $|\zeta_k|$ が最小である k $(1 \le k \le K)$ を m とする.

4°: $|p(z^{(\nu)} + \zeta_m)| \le (1 - (1-\beta)\mu)|c_n|$ ならば, $\Delta z^{(\nu)} := \zeta_m$ として第 ν 段を終了し, そうでなければ, $\mu := \mu/\lambda$, $K := K + 1$ として 1° に戻る. □

n 次多項式 $p(z)$ の一つの根 α がわかれば，$p(z)$ の他の根を求める問題は，$n-1$ 次多項式 $q(z)=p(z)/(z-\alpha)$ の根を求める問題に帰着される．これを**減次**と呼ぶ．Horner 法

$$b_0 := a_0; \quad b_k := b_{k-1}\alpha + a_k \quad (k=1,\ldots,n) \tag{10}$$

(ここで，$b_n = p(\alpha) = 0$) によって $q(z)=\sum_{k=0}^{n-1} b_{n-1-k}z^k$ の係数 b_k が計算できる．その際，$|\alpha|$ が小さいほうが b_k に含まれる丸め誤差が小さい．

平野法の収束性

平野法の収束性は，局所的には ($m=1$ となるので)Newton 法と同じである．すなわち，単根に対して 2 次収束，重根に対して 1 次収束である．第 ν 段が $\mu=\mu^{(\nu)}$ で終了すれば，

$$|p(z^{(\nu+1)})| \le (1-(1-\beta)\mu^{(\nu)})|p(z^{(\nu)})| \tag{11}$$

が成り立ち，関数値が減少する．さらに，次の定理のように厳密な意味の**大域的収束性**が理論的に保証される[5]．

定理 1 任意の初期値 $z^{(0)}$ に対し，平野法 (基本形) の近似解列 $\{z^{(\nu)}\}$ は

$$|p(z^{(\nu+1)})| \le [1-(1-\beta)\theta]\,|p(z^{(\nu)})| \tag{12}$$

を満たす．ここで，

$$\theta = \left(\frac{\beta}{1+\beta}\right)^{2n^3} \lambda^{-n}. \tag{13}$$

したがって，$\{z^{(\nu)}\}$ は $p(z)$ のある根に収束する．

[証明] 第 ν 段だけを考えるので，上添字 $^{(\nu)}$ は省略する．$c_n \ne 0$ として，

$$\zeta_k(\mu) = (-\mu c_n/c_{n-k})^{1/k}, \quad \zeta_k = \zeta_k(1) = (-c_n/c_{n-k})^{1/k}$$

とおく．$c_{n-k}=0$ ならば $\zeta_k=\infty$ であるが，$c_0 \ne 0$ だから ζ_n は有限値である．

μ $(0<\mu<1)$ が

$$B_m = \{\mu \mid |\zeta_m(\mu)| \le \frac{\beta}{1+\beta}|\zeta_k(\mu)| \quad (1 \le k \le n, k \ne m)\}$$

に属するならば，

$$p(z+\zeta_m(\mu)) = \left[\sum_{k=1}^{n} c_{n-k}\zeta_m(\mu)^k + \mu c_n\right] + (1-\mu)c_n$$

の $[\cdots]$ の部分は補題 1 により $|\,[\cdots]\,| \le \beta\mu|c_n|$ と抑えられるので，

$$|p(z+\zeta_m(\mu))| \le \beta\mu|c_n| + (1-\mu)|c_n| = (1-(1-\beta)\mu)|c_n|$$

となり，算法の 4° の判定条件が成り立って，$\Delta z = \zeta_m(\mu)$ で終了する．したがって，算法の $2^\circ \sim 4^\circ$ を繰り返して $\mu = 1, \lambda^{-1}, \lambda^{-2}, \ldots$ と減速していくとき，$\mu \in \bigcup_{m=1}^{n} B_m$ となれば修正量 Δz が定まって終了する．

後で示すように B_m は (空でなければ) 区間になるが，もし

[5] 室田一雄: 平野の変形 Newton 法の大域的収束性，情報処理学会論文誌，Vol. 21 (1980), pp. 469–474.

区間の族 $\{B_m \cap [\theta, 1]\}_{m=1}^n$ の中に "対数的長さ"(=上限と下限の比の対数) が
$\log\lambda$ 以上の区間がある $\cdots$ (#)

ならば，μ の列 $1, \lambda^{-1}, \lambda^{-2}, \ldots, \lambda^{-N}$ $(\geq \theta > \lambda^{-N-1})$ はそのような "長い" 区間を飛び越すことはできないので，少なくとも一つはそのような区間に含まれる．すなわち，(#) が成り立つならば，ある $\mu \geq \theta$ に対して終了条件が満たされ，定理の式 (12) が成り立つ．

以下，(#) を示そう．B_m の定義式より，$\zeta_m = \infty$ のときは $B_m = \emptyset$ である．$\zeta_m \neq \infty$ ならば，$\zeta_k(\mu) = \mu^{1/k}\zeta_k$ を代入して計算すると，

$$B_m = \{\mu \mid \underline{b}_m \leq \mu \leq \overline{b}_m\}$$

となる．ただし，$\gamma = \beta/(1+\beta)$ とおいて

$$\underline{b}_m = \max_{1\leq j\leq m-1}\left(\frac{1}{\gamma}\left|\frac{\zeta_m}{\zeta_j}\right|\right)^{mj/(m-j)}, \quad \overline{b}_m = \min_{m+1\leq k\leq n}\left(\gamma\left|\frac{\zeta_k}{\zeta_m}\right|\right)^{km/(k-m)} \tag{14}$$

である．ここで，$\underline{b}_m > \overline{b}_m$ なら $B_m = \emptyset$ であり，$c_{n-k} \neq 0$ である最小の $k \geq 1$ を k_1 として，$\underline{b}_{k_1} = 0, \overline{b}_n = +\infty$ と約束する．$1 \leq j \neq k \leq n$ に対して記号

$$\Delta(j,k) = \frac{\log|\zeta_j| - \log|\zeta_k|}{1/j - 1/k} \tag{15}$$

を導入すると，$\log\gamma < 0$ に注意して，

$$\begin{aligned}
\log\underline{b}_m &= \max_{1\leq j\leq m-1}\left(-\Delta(j,m) - \frac{\log\gamma}{1/j-1/m}\right) \\
&\leq -\min_{1\leq j\leq m-1}\Delta(j,m) - n^2\log\gamma, \\
\log\overline{b}_m &= \min_{m+1\leq k\leq n}\left(-\Delta(m,k) + \frac{\log\gamma}{1/m-1/k}\right) \\
&\geq -\max_{m+1\leq k\leq n}\Delta(m,k) + n^2\log\gamma
\end{aligned}$$

である．$\underline{b}_m^* = \max(\underline{b}_m, \theta), \overline{b}_m^* = \min(\overline{b}_m, 1)$ とおくと

$$B_m \cap [\theta, 1] = \{\mu \mid \underline{b}_m^* \leq \mu \leq \overline{b}_m^*\}$$

であるが,

$$\Delta(0,m) = \Delta_0 \equiv -n^2\log\gamma - \log\theta, \quad \Delta(m,n+1) = \Delta_{n+1} \equiv n^2\log\gamma$$

と定義すると，上の不等式により,

$$\begin{aligned}
\log\underline{b}_m^* &\leq -\min_{0\leq j\leq m-1}\Delta(j,m) - n^2\log\gamma, \\
\log\overline{b}_m^* &\geq -\max_{m+1\leq k\leq n+1}\Delta(m,k) + n^2\log\gamma
\end{aligned}$$

とかける．したがって，$L_m = \log\overline{b}_m^* - \log\underline{b}_m^*$ とおくと,

$$L_m \geq \min_{0\leq j\leq m-1}\Delta(j,m) - \max_{m+1\leq k\leq n+1}\Delta(m,k) + 2n^2\log\gamma \tag{16}$$

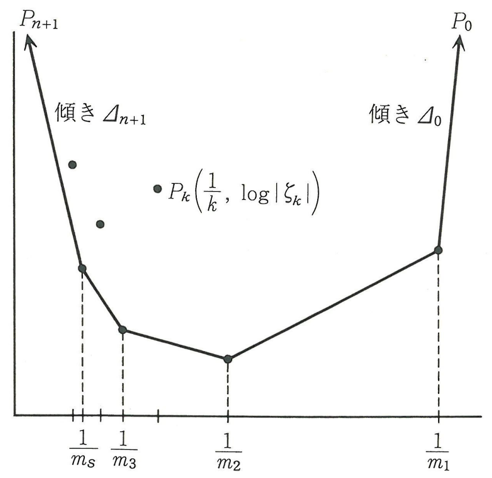

図 1: 定理 1 の証明に用いる凸図形

である．$L_m \geq 0$ ならば，L_m は区間 $B_m \cap [\theta, 1]$ の "対数的長さ" を表している．

以下，上式の右辺が $\log\lambda$ 以上であるような m の存在を示そう．図 1 のように，平面上に点

$$P_k = (1/k, \log|\zeta_k|) \qquad (k = 1, \ldots, n)$$

を描き，それらの点と二つの無限遠点

$$P_0 = +\infty(1, \Delta_0), \qquad P_{n+1} = +\infty(1, \Delta_{n+1})$$

との凸包をつくり，その頂点を

$$(1/m_i, \log|\zeta_{m_i}|) \qquad (1 \leq i \leq s; m_1 < m_2 < \cdots < m_s)$$

とし，$m_0 = 0, m_{s+1} = n+1$ とおく．$\Delta(j,k)$ が 2 点 P_j, P_k を通る直線の傾きを表すので，$m = m_i$ $(1 \leq i \leq s)$ に対する評価式 (16) は

$$L_{m_i} \geq \Delta(m_{i-1}, m_i) - \Delta(m_i, m_{i+1}) + 2n^2 \log\gamma$$

となる (すなわち，傾きの変化の大きい頂点は "長い" 区間に対応する)．これを $i = 1, \ldots, s$ について加え合わせると，

$$\sum_{i=1}^{s} L_{m_i} \geq \Delta_0 - \Delta_{n+1} + 2sn^2 \log\gamma \geq -\log\theta + 2(s-1)n^2 \log\gamma.$$

θ の値を代入すれば，この最右辺 $= n\log\lambda - 2(n-s+1)n^2\log\gamma \geq s\log\lambda$. ゆえに，ある i に対して，$L_{m_i} \geq \log\lambda$ となり，(#) が成り立つ．(証終)

以上 (2013-10-07)