

Durand-Kerner算法的全局化*

王德人¹ 赵风光²

(汤任基推荐, 1995年10月16日收到, 1997年6月23日收到修改稿)

摘 要

本文利用对称多项式与一元多项式之间的关系, 结合连续同伦思想, 构造了一条概率为1的正则同伦曲线. 然后, 对这条同伦路径, 进行离散化跟踪, 导出了一类带有步长参数的 Durand-Kerner算法, 我们证明了这类算法的整体收敛性, 从而在理论上解决了人们关于 Durand-Kerner算法具有整体性的推测. 本文还深入讨论了步长参数的选择问题. 最后, 我们以足够的数值例子, 检验了理论的正确性.

关键词 Durand-Kerner算法 连续同伦 路径跟踪 整体收敛性 点估计

一、引 言

对于复 n 次多项式

$$p(z) = z^n + a_1 z^{n-1} + \dots + a_{n-1} z + a_n \quad (1.1)$$

全部零点寻求, 当且仅有单零点时, 是有许多很好的算法, 如 Durand-Kerner 算法、Aberth 算法、平方根法以及各种变形^[8,11,14], 所有这些算法的特点, 是具有高阶敛速和并行计算功能, 其中最为实用的, 应该是 Durand-Kerner 程序, 即

$$z_i^{k+1} = z_i^k - \frac{p(z_i^k)}{\prod_{j=1}^n (z_i^k - z_j^k)} \quad (i=1, 2, \dots, n; k=0, 1, \dots) \quad (1.2)$$

许多作者认为程序(1.2)可能会有大范围收敛性质, 但至今并未见到在任何文献上进行理论上的严格论证, 而讨论得较多的倒是它的局部收敛或半局部收敛理论^[1, 2, 3]. 其实, 从数值分析的角度看, 对于低次多项式, 程序(1.2)确实具有较好的收敛性质, 初值的要求并不苛刻, 但是随着多项式的次数增加(譬如10次以上的情形), 算法对于初值的不同选择, 就显得十分敏感, 其敛散行为显得十分复杂, 远非人们猜测的那样乐观.

我们对方程 $z^{10} - 1 = 0$, 采用随机数产生的不同初值, 按算法(1.2)计算, 所得数值结果如下图

* 国家自然科学基金资助项目.

1 上海大学数学系, 上海 201800.

2 复旦大学, 上海 200433.

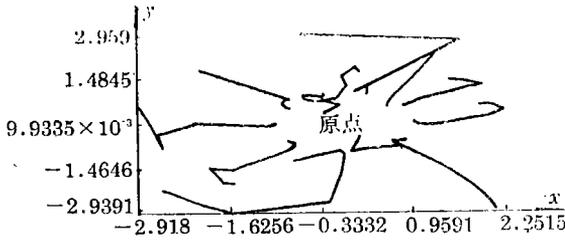


图1 按Durand-Kerner方法(1.2)计算迭代12步得 $z^{10}-1=0$ 的全部零点

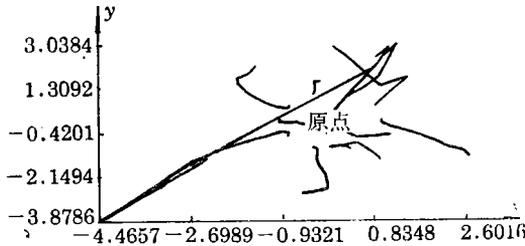


图2 按Durand-Kerner方法(1.2)计算迭代16步得 $z^{10}-1=0$ 的全部零点

由此看出，算法(1.2)对初值十分灵感，特别是图2，我们还计算过方程 $z^{20}-i=0$ ，此时，已找不到合适的初值，使算法能正常地计算下去，因而得不到任何数值结果。

为了改善 Durand-Kerner 程序的收敛性质，从理论上证实其整体收敛性，我们提出了连续化Durand-Kerner程序的思想，构造一条概率为1的正则同伦曲线，然后对这条同伦路径，进行离散化跟踪，导出了一类带有步长参数的修正算法。我们证明了这类算法的整体收敛性，从而在理论上解决了人们关于Durand-Kerner算法具有整体性的推测。同时，我们还深入分析了步长参数的选择，通过对方程 $z^{10}-1=0$ ， $z^{20}-1=0$ ， $z^{80}-1=0$ 的实际计算，显示了步长参数的不同取法，对于算法敛散行为的重要性。

二、同伦路径的建立

我们以 ${}^{\prime}C^n$ 表示 n 维复空间^①，元素

$$Z = [z_1, \dots, z_n]^T \in {}^{\prime}C^n$$

今引进关于 z_1, z_2, \dots, z_n 的基本对称多项式

$$\varphi_i(z_1, \dots, z_n) = \sum_{1 \leq j_1 < \dots < j_i \leq n} z_{j_1} \dots z_{j_i} \quad (i=1, 2, \dots, n)$$

并令 $p_i(z_1, \dots, z_n) = (-1)^i \varphi_i(z_1, \dots, z_n) - a_i \quad (i=1, 2, \dots, n)$

于是，我们可定义映射 F

$$F: {}^{\prime}C^n \rightarrow {}^{\prime}C^n \tag{2.1}$$

$$F(Z) = \begin{bmatrix} p_1(z_1, \dots, z_n) \\ \dots \\ p_n(z_1, \dots, z_n) \end{bmatrix} \in {}^{\prime}C^n, Z \in {}^{\prime}C^n$$

① ${}^{\prime}C, {}^{\prime}B, {}^{\prime}A$ 等分别表示 C, B, A 的双线性体。

且由一元多项式与 n 元对称多项式之间的关系, 我们容易得到, F 的任一零点恰好对应着 p 的所有零点.

为保证 F 零点的唯一性, 我们引进映射 F 在商空间 $'C^n/S^n$ 上的限制映射 F_S

$$F_S: 'C^n/S^n \rightarrow S^n \quad (2.2)$$

其中 S^n 为 n 个元素的置换群.

因此, 由映射 F 关于 n 个变元的对称性, 即知, 若有 $F(Z)=0$, 则对任何 $\sigma \in S^n$, 亦有 $F(Z_\sigma)=0$, 其中

$$Z_\sigma = [z_{\sigma(1)}, z_{\sigma(2)}, \dots, z_{\sigma(n)}]^T \quad (\sigma \in S^n)$$

于是, 对于限制映射 F_S (这种限制是合理), 可知 F_S 的零点是唯一的, 而且恰好对应着 p 的所有零点, 事实上, 我们只需对 F_S 的零点, 只注意其各个分量, 而不必考察各分量的顺序.

今对给定的 $Y \in 'C^n$, 考虑如下带参数 t 的方程组

$$F_S(Z) = e^{-t}Y \quad (t \geq 0) \quad (2.3)$$

为方便计, 我们把 $t = +\infty$ 亦看成一点, 使区间 $[0, +\infty)$ 紧化为区间 $[0, +\infty]$. 于是(2.3)当 $t = +\infty$ 时, 将导致方程组

$$F_S(Z) = 0 \quad (2.4)$$

其解恰好对应着 p 的所有零点.

由此, 如果我们记

$$F_S^{-1}(Y) = \{Z \in 'C^n \mid F_S(Z) = e^{-t}Y, t \geq 0\} \quad (2.5)$$

则我们将证明, $F_S^{-1}(Y)$ 形成一条可表示为参数 t 的连续曲线 $Z(t)$, $t \in [0, +\infty]$. 因此, 如果我们知道 $Z(0)$, 那么, 从 $Z(0)$ 开始跟踪这条曲线, 必将导致 p 的所有零点.

下面, 我们将讨论四个问题:

- (i) 当 $t \rightarrow +\infty$ 时, 曲线 $Z(t)$ 的极限形式是什么?
- (ii) 曲线 $Z(t)$ 是否为 t 的光滑函数?
- (iii) 在 $'C^1 \times [0, +\infty]$ 上, 曲线 $z_1(t), z_2(t), \dots, z_n(t)$ 的走向如何? 相交否?
- (iv) 初始 $Z(0)$ 将如何获得?

首先, 为回答问题(i), 我们讨论 $F_S^{-1}(Y)$ 的连续性问题. 为此, 我们在商空间 $'C^n/S^n$ 中引进拓扑. 取 $\{U_n\}$ 为 $'C^n$ 的一个拓扑基, 定义

$$V_n = \bigcup_{\sigma \in S^n} \{Z_\sigma \mid Z \in U_n\}$$

则 $\{V_n\}$ 形成了 $'C^n/S^n$ 中的一个拓扑基, 从而 $'C^n/S^n$ 构成拓扑流形, 且在此拓扑意义下, F_S 是光滑的.

定理2.1 F_S 是 $'C^n/S^n$ 到 $'C^n$ 的一个同胚映射.

证明 先证 F_S 为满射映射, 对 $\forall Y \in 'C^n$, 则 $F_S(Z) = Y$ 可展开为如下形式的多项式方程

$$p_Y(z) = z^n + (a_1 - y_1)z^{n-1} + \dots + (a_n - y_n) = 0 \quad (2.6)$$

$$Y = [y_1, y_2, \dots, y_n]^T \in 'C^n$$

由代数基本定理, p_Y 恰有 n 个不同的零点 $\alpha_1, \alpha_2, \dots, \alpha_n$. 因此, 对 $\forall \sigma \in S^n$, 亦有 $F_S(\alpha_\sigma) = Y$, 其中 $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_n]^T$, $\alpha_\sigma = [\alpha_{\sigma(1)}, \alpha_{\sigma(2)}, \dots, \alpha_{\sigma(n)}]^T$. 由此即知 $\alpha \in 'C^n/S^n$ 是 Y 的原象.

现证 F_S 是单射的. 设存在 $Z_1, Z_2 \in 'C^n/S^n$, 使成立

$$F_S(Z_1) = Y, F_S(Z_2) = Y$$

则此两方程组的展开, 同样具有(2.6)的形式, 由代数基本定理, 可知 Z_1 和 Z_2 的分量分别对应着方程(2.6)的所有 n 个不同零点, 故总存在 $\sigma \in S^n$, 使 $Z_1 = Z_2$ (或 $Z_{1\sigma} = Z_{2\sigma}$), 此即 $Z_1 = Z_2$.

于是, 由区域不变性定理 (Schwartz^[10]), 即知 F_S 为同胚映射. Q.E.D

利用定理2.1, 我们可把 $F_S^{-1}(Y)$ 表示为关于 t 的连续函数, 即

$$Z(t) = F_S^{-1}(e^{-t}Y) \quad (t \geq 0) \tag{2.7}$$

因此, 当 $t \rightarrow \infty$ 时, 由(2.7)即得极限式

$$\lim_{t \rightarrow \infty} Z(t) = F_S^{-1}(0)$$

此即我们的需要解决的问题(i).

其次, 为回答问题(iii), 我们把 $Z(t)$ 表示成分量形式, 即 $Z(t) = [z_1(t), z_2(t), \dots, z_n(t)]^T$. 显然所有的 $z_i(t) (i=1, 2, \dots, n)$, 表示由 $[0, +\infty]$ 到 $'C^1$ 的 n 条连续曲线, 因为由定理2.1保证, 它们不可能间断, 为清楚起见, 我们将它们画在 $'C^1 \times [0, +\infty]$ 上, 如图3. 它们可能有三种情况, 首先, 由定理2.1保证, 它们不可能发散到无穷, 即情况A不可能出现, 只可能象情况B那样, 从一端光滑地到达另一端, 或者象情况E, F那样在D点相交后, 又各自分开到达另一端. 事实上, 最后这种情况我们亦是不愿意遇到的, 所以我们在下面证明了这样一个事实, 就是通过 $'C^n$ 中随机选取 Y 的办法, 使遇到这种相交情形的概率为0.

为此, 我们将(2.3)改写为

$$h_{Y,t}(z) = z^n + (a_1 - e^{-t}y_1)z^{n-1} + \dots + (a_n - e^{-t}y_n) = 0$$

$$Y = [y_1, y_2, \dots, y_n]^T$$

同时, 我们利用多项式结式概念, 设 p, q 为两个多项式, 以符号 $\text{Res}(p, q)$ 表示结式, 它是相应于 p 和 q 的系数的一个行列式^[17].

引理2.2 多项式 p 有重根的充分必要条件是

$$\text{dis}(p) = \text{Res}(p, p') = 0$$

证明 见[17]

现在, 我们把 $h_{Y,t}(z)$ 看成系数为 $a_1 - e^{-t}y_1, \dots, a_n - e^{-t}y_n$ 的多项式, 并定义

$$'B = \{Y \in 'C^n \mid \exists t_0 > 0, \text{使} \text{dis}(h_{Y,t}(z)) = \text{Res}(h_{Y,t}(z), h'_{Y,t}(z)) = 0\} \tag{2.8}$$

显然, 由(2.8)的定义, 当 $Y \in 'C^n - 'B$ 时, 则对任意固定的 $t > 0$, $h_{Y,t}$ 没有重零点, 下面我们证明由(2.8)定义的集合 $'B$ 是一个零测集.

定理2.3 由(2.8)定义的 $'B$ 是 $'C^n$ 中的零测集.

证明 由于 $\text{dis}(h_{Y,t}(z))$ 可表示为一个关于系数

$$a_1 - e^{-t}y_1, a_2 - e^{-t}y_2, \dots, a_n - e^{-t}y_n$$

的行列式, 因此, 对固定的 $t > 0$, 我们可以得到一个关于 y_1, y_2, \dots, y_n 的多项式方程

$$P_t(y_1, \dots, y_n) = 0$$

这个多项式不恒为0, 因为我们可以通过选不同的 $Y \in 'C^n$, 而使 $h_{Y,t}$ 无重零点. 于是, 集合

$$P_t^{-1}(0) = \{Y \in 'C^n \mid P_t(Y) = 0\}$$

是 $'C^n$ 中余维为1的流形, 亦即是 R^{2n} 中余维为2的流形, 故集合 $'B = \bigcup_{t>0} P_t^{-1}(0)$ 是 $'C^n$ 中的零测集.

Q.E.D

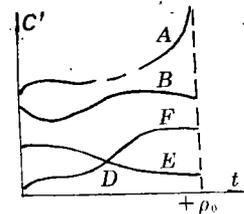


图 3

由此, 只要我们所选的 $Y \in 'C^n$ 不会落在 $'B$ 中, 那么上面所图示的相交情形就永远不会发生, 除非 p 本身有重零点, 此时将会导致在 $+\infty$ 处相交. 然而这种情形已不是我们在跟踪过程中所需考虑的问题.

现在, 我们再来考虑问题(ii), 即曲线 $Z(t)$ 的光滑性. 由于 F_s 并不是光滑微分同胚, 因此, $Z(t)$ 的光滑性是需要考虑的. 为此, 我们先考虑下述引理.

引理2.4 $\det(F'(Z)) \neq 0$ 的充分必要条件为

$$z_i \neq z_j \quad (i \neq j, i, j = 1, 2, \dots, n)$$

其中 $Z = [z_1, z_2, \dots, z_n]^T \in 'C^n$.

证明 首先, 我们利用下述关系式

$$U(Z)F'(Z) = V(Z) \quad (2.9)$$

其中

$$U(Z) = \begin{bmatrix} z_1^{n-1} & \dots & z_1 & 1 \\ z_2^{n-1} & \dots & z_2 & 1 \\ \dots & \dots & \dots & \dots \\ z_n^{n-1} & \dots & z_n & 1 \end{bmatrix}$$

$$V(Z) = \text{diag}(-\prod_{j \neq 1} (z_1 - z_j), \dots, -\prod_{j \neq n} (z_n - z_j))$$

今假定 Z 中只有 r ($r \leq n$) 个不同分量, 且分别有重复数, m_i ($i = 1, 2, \dots, r$), 则有

$$\text{rank}(U(Z)) = n - \sum_{m_i > 1} (m_i - 1)$$

$$\text{rank}(V(Z)) = n - \sum_{m_i > 1} m_i$$

于是, 我们容易验证

$$\text{rank}(F'(Z)) = n \iff m_i = 1 \quad (i = 1, 2, \dots, n)$$

此表明 Z 中不可能有重复分量, 即 $z_i \neq z_j$ ($i \neq j, i, j = 1, 2, \dots, n$) Q.E.D.

定理2.5 若 $Y \in 'C^n - 'B$, 则 $F_s^{-1}(Y)$ 可唯一表示为 t 的光滑函数 $Z(t)$, 且满足 $Z(0) = F_s^{-1}(Y)$, 同时有

$$\lim_{t \rightarrow \infty} Z(t) = F_s^{-1}(0) \quad (2.10)$$

证明 由于 $Y \in 'B$, 按集合 $'B$ 的定义, 对任意 $t_0 > 0$, 方程组(2.3)或多项式 $h_{Y,t}(z)$ 无重根, 即若用 $Z(t_0)$ 表示(2.3)的解, 则 $z_i(t_0) \neq z_j(t_0)$ ($i \neq j, i, j = 1, 2, \dots, n$) 成立. 于是由引理2.4得 $\det F'(Z(t_0)) \neq 0$. 从而由隐函数定理, 保证了在 $(Z(t_0), t_0)$ 附近, Z 可表示为 t 的光滑函数 $Z(t)$. 再利用对称性, 这种表示在空间 $'C^n/S^n$ 中是唯一的. 于是, 我们可以从 $t=0$ 出发, 逐步延拓. 如果存在某个 $t_0 > 0$, 使延拓无法进行, 则将会出现矛盾, 故定理结论成立.

关于极限式(2.10), 则是由于 F_s 是同胚映射 (定理2.1), 故(2.10)成立. Q.E.D

最后, 我们讨论初始 $Z(0)$ 的选取. 本文的目的是寻求多项式(1.1)的全部零点, 即 $F_s^{-1}(0)$. 然而, 在以上的论证中却遇到了同样难求的 $F_s^{-1}(Y)$. 一个自然的想法是, 随机选取点 $Z^0 \in 'C^n$, 并以 $F(Z^0)$ 作为(2.3)中的 Y , 这样选取的好处是可以避免选到 $z_i = z_j$ (对某 i, j) 的点. 下面我们需要研究这种选取的合理性. 首先运用文[15]中的引理:

引理2.6 设 $D \subset R^n$ 是一个开子集, $F: D \rightarrow R^n$ 是 C^1 映射. 如 $V \subset D$ 是零测集, 则 $F(V)$ 也是 R^n 中的零测集

今定义

$$\mathcal{X}^n = \{Z \in C^n \mid z_i \neq z_j, i \neq j, i, j = 1, 2, \dots, n\}$$

则有

$$'A = F(C^n - \mathcal{X}^n) = F(C^n/S^n - \mathcal{X}^n/S^n) \quad (2.11)$$

由此, 由于 F 的光滑性以及 $C^n - \mathcal{X}^n$ 是 C^n 中的零测集, 故由引理2.6知, $'A$ 亦是零测集.

引理2.7 映射 F_s 是 \mathcal{X}^n/S^n 到 $C^n - 'A$ 的光滑微分同胚.

证明 由(2.11)及定理2.1知, F_s 是 \mathcal{X}^n/S^n 到 $C^n - 'A$ 的同胚映射, 又由于对 $\forall Z \in \mathcal{X}^n/S^n$, 有 $\text{rank}(F_s(Z)) = \text{rank}(F(Z)) = n$, 且 F_s 又是光滑映射, 因此, F_s 是局部微分同胚, 从而得证 F_s 是由 \mathcal{X}^n/S^n 到 $C^n - 'A$ 上的光滑微分同胚. Q. E. D

今令

$$F_{S, Z^0}^{-1}(0) = \{Z \in C^n \mid F(Z) = tF(Z^0), t > 0\}$$

于是, 我们可得下述定理

定理2.8 对几乎所有的 $Z^0 \in C^n$, $F_{S, Z^0}^{-1}(0)$ 是连接 Z^0 到 $F_S^{-1}(0)$ 的光滑无穷道路.

证明 由

$$\begin{aligned} C^n - F^{-1}('B) &\supseteq C^n - F^{-1}('A \cup 'B) \\ &= C^n - F^{-1}('A) - F^{-1}('B - 'A) = \mathcal{X}^n - F^{-1}('B - 'A) \end{aligned}$$

可得

$$\begin{aligned} F^{-1}('B) &\subseteq (C^n - \mathcal{X}^n) \cup F^{-1}('B - 'A) \\ &= (C^n - \mathcal{X}^n) \cup_{\sigma \in S^n} (F_S^{-1}('B - 'A))_c \end{aligned} \quad (2.12)$$

其中 $'A, 'B$ 分别由(2.8)和(2.11)定义, 而 $(W)_c = \{w_\sigma \mid w \in W\}$. 又因为 $C^n - \mathcal{X}^n$ 是 C^n 中的零测集, $'B - 'A$ 是 $C^n - 'A$ 中的零测集, F_S^{-1} 限制在 $C^n - 'A$ 上光滑, 从而 $F_S^{-1}('B - 'A)$ 是零测集. 于是, 由(2.12)推出 $F^{-1}('B)$ 是 C^n 中的零测集. Q. E. D

这样我们通过随机选取初始 Z^0 , 使 $Z^0 \in F^{-1}('B)$ 的可能性为0, 亦就是说几乎对所有的 $Z^0 \in C^n$, $Y = F(Z^0) \in 'B$. 从而由定理保证了 $F_{S, Z^0}^{-1}(0)$ 光滑地连接 Z^0 到 $F_S^{-1}(0)$.

到以, 我们已把上面所提出的问题, 完善地得到了解决, 使以后的数值跟踪过程, 得以顺利完成.

三、数值跟踪算法的建立

由第二节中的分析, 我们总有很大的可能性选到不属于 $F^{-1}('B)$ 的初始点 Z^0 . 于是, 由定理2.5可知(2.3)中的解曲线 $Z(t)$ 是关于 t 的光滑函数, 因此, 我们对(2.3)两边关于 t 求导得

$$F'(Z(t))\dot{Z}(t) = -e^{-t}Y \quad (3.1)$$

将(2.3)代入(3.1)得

$$\left. \begin{aligned} F'(Z(t))\dot{Z}(t) &= -F(Z(t)) \\ Z(0) &= Z^0 \quad (t > 0) \end{aligned} \right\} \quad (3.2)$$

又由于沿曲线 $Z(t)$, 其各分量互不相同, 从而 $F'(Z(t))$ 可逆, 于是(3.2)可改写为

$$\left. \begin{aligned} \dot{Z}(t) &= -F'(Z(t))^{-1}F(Z(t)) \\ Z(0) &= Z^0 \end{aligned} \right\} \quad (3.3)$$

对此初值问题, 我们利用Euler折线法跟踪, 于是由(3.3)可得如下迭代关系

$$Z^{k+1} = Z^k - h_k F'(Z^k)^{-1} F(Z^k) \quad (k=0, 1, 2, \dots) \quad (3.4)$$

其中 Z^0 为初值, h_k 为参数. 又由映射 F 的定义及 $F'(Z^k)^{-1}$ 的具体形式^[1], 我们可得(3.4)的分量形式

$$z_i^{k+1} = z_i^k - \frac{h_k p(z_i^k)}{\prod_{j \neq i} (z_i^k - z_j^k)} \quad (i=1, 2, \dots, n; k=0, 1, \dots) \quad (3.5)$$

(3.5)就是Durand-Kerner的修正形式, 亦就是多项式(1.1)的 n 个零点的 n 条跟踪曲线的离散形式, 它具有很好的并行计算性能和整体收敛性能.

在第二节中讨论的理论基础上, 我们对于(3.5)可得到下述整体收敛定理

定理3.1 对几乎所有的 $Z^0 \in 'C^n$ 及 $\varepsilon > 0$, 都存在常数 $h > 0$, $M > 0$, 使从 Z^0 出发的迭代(3.5), 当 $h_k = h$, $k \leq M$ 时, 算法(3.5)有意义, 且对每个 i , 存在 p 的某个零点 α_i , 满足

$$|z_i^M - \alpha_i| \leq \varepsilon$$

证明 由于 $F^{-1}('B)$ 是零测集. 因此, 随机选取 Z^0 不落在 $F^{-1}('B)$ 中的概率为1. 不妨设 $Z^0 \in F^{-1}('B)$, 则由定理2.8知, $F_{S, Z^0}^{-1}(0)$ 是一条光滑道路 $Z(t)$, 且当 $t \rightarrow \infty$ 时, 有

$$Z(t) \rightarrow F_S^{-1}(0)$$

从而存在一个 $T > 0$, 使 $t \geq T$ 时, 有

$$\|Z(t) - F_S^{-1}(0)\|_\infty < \varepsilon/2 \quad (3.6)$$

而在区间 $[0, T]$ 上, (3.3)的右端, 在 $\{(Z, t) \in 'C^n \times [0, T] \mid F(Z) = e^{-t} Y\}$ 的一个管形区域内是光滑的. 因此, 必存在一个步长 h , 及剖分数 M , $hM = T$, 满足

$$\|Z(T) - Z^M\|_\infty < \varepsilon/2 \quad (3.7)$$

于是, 结合不等式(3.6)、(3.7)即得定理结论. Q. E. D

上述定理是对 h_k 为常数的情况下证明的, 对于变步长的情况, 同样可以证明, 只是叙述上要稍作变动.

四、步长 h_k 的选择

第三节主要证明了算法(3.5)几乎对初值的选取没有任何要求, 但是对步长参数 h_k 的选择, 并未作什么讨论, 事实上, 要选择适当的步长参数 h_k 亦非易事.

从理论上讲, 只要 h_k 选得尽可能小, 是会使(3.5)连续跟踪(2.3)中的道路, 从而也不会引起溢出, 因为, 沿此道路函数值 $F(Z)$ 是不断在下降. 但是, 过小的步长 h_k 无疑会增加不必要的计算工作量. 因此, 只有在不会引起计算过程中所可能出现的麻烦的情况下, 尽可能使 h_k 增大直至为1, 这才会使跟踪算法更有效.

为此, 我们设 $z_1(t), \dots, z_n(t)$ 为 $F_S^{-1}(Y)$ 中的 n 条道路并定义

$$d(t) = \min_{i \neq j} |z_i(t) - z_j(t)| \quad (4.1)$$

如果有 $Z^0 \in F^{-1}('B)$, 则对 $\forall t > 0$, 都是 $d(t) > 0$. 另外, 我们再引进

$$d^k = \min_{i \neq j} |z_i^k - z_j^k| \quad (4.2)$$

其中 $z_i^k (i=1, 2, \dots, n; k=0, 1, \dots)$, 由算法(3.5)产生, 显然, 要使算法(3.5)有效地跟踪 $F_s^{-1}(Y)$, 就必需通过选取 h_k , 使 $d^k > 0$, 此时可使迭代不中断. 基于这种想法, 我们假定已经得到了这样的 h_k 和序列 $\{z_i^k\}$, ($i=1, 2, \dots, n$), 且有 $d^k > 0$, 对于 d^{k+1} , 我们有

$$\begin{aligned} d^{k+1} &= \min_{i \neq l} |z_i^{k+1} - z_l^{k+1}| \\ &= \min_{i \neq l} \left| z_i^k - z_l^k - h_k \frac{p(z_i^k)}{\prod_{j \neq i} (z_i^k - z_j^k)} + h_k \frac{p(z_l^k)}{\prod_{j \neq l} (z_l^k - z_j^k)} \right| \end{aligned}$$

由此看出, 我们只要取

$$\left. \begin{aligned} h_k &< \min_{i \neq j} \left| \frac{z_i^k - z_j^k}{w_i^k - w_j^k} \right| \\ w_i^k &= p(z_i^k) / \prod_{j \neq i} (z_i^k - z_j^k) \end{aligned} \right\} \quad (i=1, 2, \dots, n) \quad (4.3)$$

此时就可保证 $d^{k+1} > 0$. 另一方面, 为保证 Durand-Kerner 算法的二次收敛性, h_k 不能大于 1. 因此, h_k 应满足

$$h_k = \min \left\{ 1, \beta \min_{i \neq j} \left| \frac{z_i^k - z_j^k}{w_i^k - w_j^k} \right| \right\} \quad (\beta < 1) \quad (4.4)$$

这种 h_k 的选择, 并没有带来额外函数值的计算量, 只是利用了已经算得的结果, 因此可以取得更好的效率.

由(2.4), 我们可对步长 h_k 作如此解释, 如果相应于 z_i^k 和 z_j^k 的修正量 w_i^k 和 w_j^k 的差值增长过快, h_k 将变小, 这时 h_k 起收缩作用. 相反, 如果迭代在零点附近, w_i^k 将逼近于 0, 这时, h_k 就会变大, 起加速作用, 直到 h_k 增大到 1 时, 算法(3.5)就变成了快速收敛的 Durand-Kerner 方法(1.2), 步长 h_k 与 d^k 的关系是, 当 d^k 很小时, h_k 需取小步长, 而 d^k 较大时, 则 h_k 相应地调大.

步长 h_k 的另一种选择方法, 是来自于 Durand-Kerner 方法的点估计理论. 此时, 我们可以按照算法的点估计收敛条件来选取 h_k , 以保证(3.5)在迭代过程中总有 $d^k > 0^{[1]}$. 我们可以证明, 当

$$h_k = \min \left\{ 1, 0.204378d^k / \sum_i |w_i^k| \right\} \quad (4.5)$$

时, 足够保证 $d^k > 0$, 且同样具有收缩和加速作用. 但是, 这样给出的 h_k 往往过于精细, 从而会导致迭代步数的增多. 数值试验表明, 采用(4.5)的效率确实没有(4.4)来得好.

五、数值试验

为了检验迭代法(3.5)的效率与稳定性, 我们分别利用 Durand-Kerner 方法(1.2), 修正 Durand-Kerner 方法(3.5)(4.4)和(3.5)(4.5)计算了以下例子.

例1 在引言中, 我们已经对方程 $z^{10} - 1 = 0$, 采用随机选取不同初值的办法, 按方法(1.2)计算, 其数值结果已在(图1)(图2)中显示. 在此, 我们以同样选取初值的方法, 在固定 $h_k = h$ 的情况下, 用方法(3.5)计算, 所得结果, 如(图4)(图5).



图4 按修正算法(3.5), 在定步长 $h=0.2$ 计算经过24步迭代得 $z^{10}-1=0$ 的全部零点

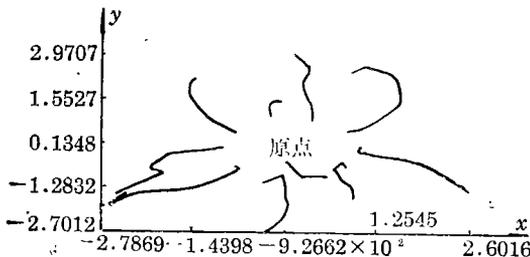


图5 按修正算法(3.5), 在定步长 $h=0.02$ 计算经过22步迭代得 $z^{10}-1=0$ 的全部零点

由此容易看出, 图4、图5中的跟踪路径比图1、图2中的路径要光滑得多。显然, 算法的收敛速度要慢一些, 但初值的选取几乎没有什么困难, 而且数值稳定性亦较好, 所不足的是追踪过程还有尖点与跳跃的情况存在, 曲线还显得不够光滑。

我们还试图用Durand-Kerner方法去求解方程 $z^{20}-1=0$, $z^{30}-1=0$ 等, 但是, 由于几乎对所有初值均出现溢出现象, 而使算法(1.2)无法继续进行下去, 宣告失败。下面的例子是对这些方程采用了算法(3.5)(4.4)和算法(3.5)(4.5)所得的数值结果, 此时, h_k 可按公式(4.4)或(4.5)变动, 我们获得了更理想的结果。

例2 求解方程 $z^{10}-1=0$, $z^{20}-1=0$, $z^{30}-1=0$, 我们利用算法(3.5)(4.4)进行计算, 其数值结果如图6、图7、图8。

由这些图形看, 不论方程的次数多么高, 其跟踪的路径都是相当光滑, 且不出现紊乱行为, 这是图1~图4所无法比拟的。更重要的是算法(3.5)、(4.4)和(3.5)、(4.5)对任何初值几乎都适用, 特别, 对高次多项式无溢出现象, 数值稳定性极好。这些算例和图形充分体现了算法(3.5)、(4.4)和(3.5)、(4.5)的效能。

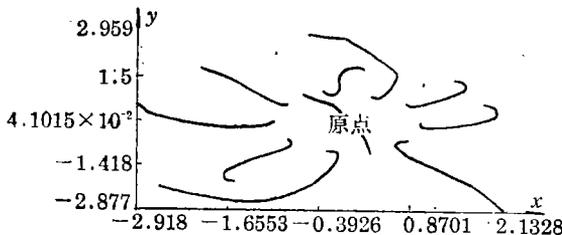


图6 按修正算法(3.5)(4.4)计算方程 $z^{10}-1=0$ 的全部零点 (迭代步数为22)



图7 按修正算法(3.5)(4.4)计算方程 $z^{20}-1=0$ 的全部零点(迭代步数为31)

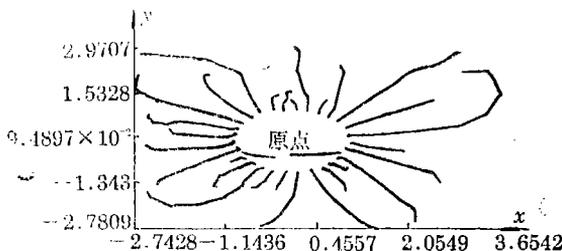


图8 按修正算法(3.5)(4.4)计算方程 $z^{30}-1=0$ 的全部零点(迭代步数为84)

最后, 我们想通过算法(1.2), 算法(3.5)、(4.4)和算法(3.5)、(4.5)时间的比较, 进一步阐明这些算法的优劣, 以作为本文的结束语。

例3 我们对方程 $z^{30}-1=0$ 的求解, 分别利用算法(1.2), 算法(3.5)、(4.4)和算法(3.5)、(4.5), 按随机办法选取不同的初值进行计算, 它们的敛散行为, 列表如下(精度相同)

表 1

初 值 \ 迭代次数	算法(1.2)	算法(3.5)(4.4)	算法(3.5)(4.5)
$(Z^0)_1$	∞	60	149
$(Z^0)_2$	∞	54	125
$(Z^0)_3$	∞	54	186
$(Z^0)_4$	∞	65	197
$(Z^0)_5$	∞	71	172
$(Z^0)_6$	∞	49	160
$(Z^0)_7$	∞	50	131

由此表看出, 算法(1.2)几乎对所有的初值均告失败, 而算法(3.5)、(4.4)和算法(3.5)、(4.4)均收敛, 但它们各自的迭代次数都有很大区别, 这说明由点估计方法选择的步长 h_k 过于精细, 因此影响了该算法的收敛速度. 所以在实用中, 我们还是推崇算法(3.5)、(4.4), 它会带来很好的计算效果。

参 考 文 献

- [1] 赵风光、王德人, Smale点估计理论与Durand-Kerner算法的收敛性, 计算数学, 2 (1993), 196-206.
- [2] 郑士明, 关于同时求解多项式所有根的Durand-Kerner算法的收敛性, 科学通报, 9 (1982),

515—517.

- [3] G. Alefeld and J. Herzberger, On the convergence speed of some algorithms for the simultaneous approximation of polynomial roots, *SIAM J. Numer. Anal.* **11** (1974), 237—243.
- [4] C. B. Garcia and W. I. Zangwill, Finding all solutions to polynomial systems and Other systems of equations, *Meth Prog.*, **16** (1979) 159—176.
- [5] T. Y. Li and T. Sauer, Homotopy method for generalized eigenvalue problems $Ax=aBx$, *LAA.*, **91** (1987), 65—74.
- [6] T. Y. Li, T. Sauer and J. Yorke, Numerical solution of a class of deficient polynomial systems, *SIAM J. Numer. Anal.*, **24** (1987), 435—451.
- [7] T. Y. Li, T. Sauer and J. Yorke, The Cheater's homotopy: An efficient procedure for solving systems of polynomial equations, *SIAM J. Numer. Anal.*, **26** (1989), 1241—1251.
- [8] T. Y. Li, Z. Zeng and L. Cong' Solving eigenvalue problems of real nonsymmetric matrices with real homotopies, *SIAM J. Numer. Anal.*, **29** (1992) 229—248.
- [9] Shui-Nee Chow, J. M. Paret and J. A. Yorke, Finding zeros of maps: Homotopy methods that are constructive with probability one, *Math. Comp.*, **32** (1978), 887—899.
- [10] S. T. Schwartz, *Nonlinear Functional Analysis Gordon and Breach*, New York (1969).
- [11] E. Durand, Solutions numériques des équations Algébriques, Tome I: *Equations du Type $F(x)=0$, Racines dun polynôme*, Masson, Paris (1960).
- [12] Wang Deren and Zhao Fengguang, Complexity analysis of a process for Simultaneously obtaining all zeros of polynomials, *Computing*, **43** (1989), 187—197.
- [13] Wang Deren and Wu Yujiang, Some modifications of the parallel Halley iteration method and their convergence, *Computing*, **38** (1987), 75—87.
- [14] Wang Deren and Zhao Fengguang, The theory of Smale's point estimation and its some application, *JCAM*, **60** (1993), 253—264.
- [15] 徐森林、王则柯, 《代数方程组和计算复杂性理论》, 科学出版社 (1989).
- [16] I. O. Kerner, Ein gesamtstschritlverfahren zur berechnung der nullstellen von polynomen, *Numer Math.*, **8** (1966), 290—294.
- [17] 莫宗坚等, 《代数学》(上), 北京大学出版社 (1987).
- [18] E. L. Allgower and K. Georg, *Numerical Continuation Methods, an Introduction*, Springer-Verlag, New York (1990).
- [19] S. Smale, Newton's method estimates from data at one point, in *The Morning of Disciplines: New Directions in Pure, Applied and Computational Mathematics* (R. Ewing, K. Gross, and G. Marfin Editors), Springer-Verlag, New York (1986), 185—196.

The Globalization of Durand-Kerner Algorithm

Wang Deren

(Shanghai University, Shanghai 201800, P. R. China)

Zhao Fengguang

(Fudan University, Shanghai 200433, P. R. China)

Abstract

Making use of the theory of continuous homotopy and the relation between symmetric polynomial and polynomial in one variable the authors devoted this article to constructing a regularly homotopic curve with probability one. Discrete tracing along this homotopic curve leads to a class of Durand-Kerner algorithm with step parameters. The convergence of this class of algorithms is given, which solves the conjecture about the global property of Durand-Kerner algorithm. The problem for steplength selection is thoroughly discussed. Finally, sufficient numerical examples are used to verify our theory.

Key words Durand-Kerner algorithm, continuous homotopy, path tracing, global convergence, point estimation