逼近零点与计算复杂性理论

王则柯

徐森林

(中山大学数学系, 广州) (中国科学技术大学数学系, 合肥)

摘 要

本文比较 Smale 对 Newton 方法所作的成本估计和作者对 Kuhn 算法所作的成本估计. 结果表明, 无论是算零点还是算所谓逼近零点, 后者都比前者好得多, 其比率分别为 n^9/μ^7 对 $n^2 \log(n/\epsilon)$ 和 n^9/μ^7 对 $n^3 \log(n/\mu)$, 这里 n 是多项式的阶数, $\epsilon > 0$ 是计算零点的精度要求, μ 是允许相应的论断失败的概率, $0 < \mu < 1$.

一、引 言

计算复系数多项式的零点, 有两种值得注意的方法: Newton 方法和 Kuhn 算法.

Newton 方法是一种广泛应用的方法. 除了历史的原因外, Newton 方法公式之简单也是人们乐意首先考虑用它的重要因素. Newton 方法及其变种, 本质上是基于一个迭代过程或迭代公式. 公式固然简单, 但用起来却很不容易, 因为总是有一个初值选择的问题要解决. Newton 法对于退化情况是失败的, 或者收敛速度变得很慢. 为此, Smale 对 Newton 方法的效率提出了一种概率的讨论^[1]. Smale 引进逼近零点的概念, 并得到一些有意义的结果.

Kuhn 算法则是基于一种拓扑结构^[2]. 虽然算法的叙述稍费篇幅, 但一旦编成程序, 使用却极其方便, 唯一要做的就是输入多项式的系数组及计算的精度要求. 然后, 算法总是可以把多项式的全部零点找到, 根本不存在初值选择的问题, 不会发生算法失败的情况.

谈到收敛速度, Smale 证明 Newton 方法找多项式

$$f(z) = z^n + a_1 z^{n-1} + \cdots + a_n, |a_i| < 1,$$

的一个“逼近零点”所需的迭代次数顶多是 $[100(n+2)]^9/\mu^7$, μ 是允许论断失败的概率, $0 < \mu < 1$. 本文证明, Kuhn 算法算出该多项式全部零点所需的函数计值次数顶多是

$$150(n+1)^3 \log(n/\epsilon),$$

$\epsilon > 0$ 是计算的精度要求. 大家知道, 一次 Newton 迭代需要两次函数计值.

本文的目的是对两种算法的上述成本估计进行比较. 基于代数几何某些定理的一种概率分析, 是这一比较的主要方法. 此外, 与 Beiberbach 猜想有关的一些结果在这一比较中起着重要作用.

二、Kuhn 算法的拓扑结构

考虑计算复系数多项式 $f(z) = z^n + a_1 z^{n-1} + \cdots + a_n$ 的全部零点的问题. Kuhn 算法

(详见文献 [2]) 可以概述如下: 对半空间 $C \times [-1, +\infty)$ 进行单纯剖分, 使得顶点都在类如 $C_d = C \times \{d\}$, 整数 $d \geq -1$ 的平面上, 平面 C_d 被剖分为直角边长为 $s(d)$ 的等腰直角三角形, $s(-1) = 1$, 而 $s(d) = 2^{-d}$, $d \geq 0$. 复数幅角限制在 $(-\pi, \pi]$, C_d , $d \geq 0$, 上顶点的标号为:

$$l(z) = \begin{cases} 1, & \text{若 } f(z) = 0, \text{ 或 } -\pi/3 \leq \arg f(z) \leq \pi/3, \\ 2, & \text{若 } \pi/3 < \arg f(z) \leq \pi, \\ 3, & \text{若 } -\pi < \arg f(z) < -\pi/3; \end{cases}$$

而对 C_{-1} 上顶点的标号, 则在上式中用 z^n 代替 $f(z)$. 设 Q 是 C_{-1} 上由 $y = \pm m$ 和 $x = \pm m$ 界定的正方形, 这里 $z = x + iy$, 而 $m = \lceil 3(1 + \sqrt{2})n/4\pi \rceil^2$. 三角形称作是完全标号的, 若它的顶点正好具有标号 1, 2, 3. 从 Q 的边界 ∂Q 上的 n 个标号为 (1, 2) 的棱出发, 可以找到 Q 内 n 个全标三角形. 据剖分, 每个全标三角形 (z^1, z^2, z^3) 确定唯一的新顶点 z' . 若 $l(z') = l$, 令 z' 为新的 z^l . 按此轮迴规则, 从上述 Q 内 n 个全标三角形开始的计算, 正好收敛到多项式的 n 个零点.

我们将把 (z^1, z^2, z^3) 确定新顶点 z' , 若 $l(z') = l$, 则令 z' 为新的 z^l 这样一步轮迴直观地称作是计算经过四面体 (z^1, z^2, z^3, z') . 三角形作为计算经过四面体的一个门的先决条件是它必须是该四面体的面, 并且是全标三角形. 更准确地说, 一个门称作是一个进口, 若从四面体内部看去它的顶点标号依逆时针方向正好是 (1, 2, 3); 若正好是 (3, 2, 1), 则称作是一个出口. 事实上, 在每步轮迴中, 旧全标三角形是该四面体的进口, 新全标三角形是出口.

命题 2.1. 每个四面体顶多允许计算经过一次.

证. 记半空间的剖分为 T , 以 $V(T)$ 记 T 的顶点集. 标号法实际上是一个对应

$$l: V(T) \rightarrow \{1, 2, 3\}.$$

所以, 对任何标号四面体, 或者它没有门, 或者它正好有一对门, 一个进口一个出口. 命题 2.1 证完.

以 T_j 记上述 Q 内第 j 个全标单纯形开始的计算所经过的所有开单纯形的点集并, $1 \leq j \leq n$. 我们显然有下述

命题 2.2. $T_i \cap T_j = \emptyset$, 当 $i \neq j$.

下面我们提出 Kuhn 算法误差估计的一个改进. 首先叙述一个易证, 然而有用的命题.

命题 2.3. $|w| < 1$ 蕴含 $|\arg(1+w)| \leq |w|\pi/2$.

定理 2.1. 若 $\{z^1, z^2, z^3\}$ 是一个全标三角形的顶点集, 标号由 f 给出,

$$\text{diam}\{z^1, z^2, z^3\} = \delta.$$

那末 f 的零点集与 $\{z^1, z^2, z^3\}$ 的距离不超过 $3n\delta/4$.

证. $n = 1$ 的情况是平凡的, 这时 f 只是一个平移. 下面考虑 $n > 1$ 的情况.

改写 $f(z)$ 为 $f(z) = \prod_{j=1}^n (z - \xi_j)$. 若定理不真, 则对所有 j , $1 \leq j \leq n$, $|z^1 - \xi_j| > 3n\delta/4$. 于是

$$|(z^2 - z^1)/(z^1 - \xi_i)| < \delta/(3n\delta/4) = 4/3n < 1.$$

1) $\lceil x \rceil$ 表示不小于实数 x 的最小整数.

根据命题 2.3,

$$\begin{aligned} \left| \arg \frac{f(z^2)}{f(z^1)} \right| &= \left| \arg \frac{(z^2 - \xi_1) \cdots (z^2 - \xi_n)}{(z^1 - \xi_1) \cdots (z^1 - \xi_n)} \right| \leq \sum_{j=1}^n \left| \arg \frac{z^2 - \xi_j}{z^1 - \xi_j} \right| \\ &= \sum_{j=1}^n \left| \arg \left(1 + \frac{z^2 - z^1}{z^1 - \xi_j} \right) \right| \leq \sum_{j=1}^n \frac{\pi}{2} \left| \frac{z^2 - z^1}{z^1 - \xi_j} \right| < n \cdot \frac{\pi}{2} \cdot \frac{4}{3n} = \frac{2\pi}{3}. \end{aligned} \quad (2.5)$$

同样,

$$\left| \arg f(z^1)/f(z^3) \right| < 2\pi/3, \quad (2.6)$$

$$\left| \arg f(z^2)/f(z^3) \right| < 2\pi/3. \quad (2.7)$$

根据标号法, $\arg f(z^2) - \arg f(z^1) > 4\pi/3$, $\arg f(z^1) - \arg f(z^3) > 4\pi/3$ 和 $\arg f(z^2) - \arg f(z^3) < 2\pi/3$ 均不可能, 故 (2.5)–(2.7) 分别等价于

$$0 < \arg f(z^2) - \arg f(z^1) < 2\pi/3, \quad (2.8)$$

$$0 < \arg f(z^1) - \arg f(z^3) < 2\pi/3, \quad (2.9)$$

和

$$\arg f(z^2) - \arg f(z^3) > 4\pi/3. \quad (2.10)$$

从(2.10)式减去(2.8)式, 得

$$\arg f(z^1) - \arg f(z^3) > 2\pi/3,$$

与(2.9)式矛盾. 定理 2.1 得证.

证明中包含一个以后经常要用到的事实, 兹叙述如下:

引理 2.1. 若对某三点组中任两点 z' 和 z'' 均成立不等式 $|\arg f(z')/f(z'')| < 2\pi/3$, 则该三点组不是由 f 完全标号的.

三角形的三个顶点正好构成一个三点组.

三、零点计算的成本估计

这一节对算法可能区域中位于表示精度要求的高度以下的四面体总数给出一个上界, 这样, 计算多项式全部零点所需要的 f -计值次数的上界也就有了, 这是因为每一个轮迴顶多包含一次多项式计值. 主要结果是定理 3.1 及推论 3.1.

定理 3.1. 设 $\epsilon > 0$ 为计算的精度要求. 用 Kuhn 算法算出多项式 f 的全部零点所需的 f 计值次数不超过

$$\pi[5M^2 + 28n(1 + 0.75n)^2 \lceil \log_2(3\sqrt{2}n/4\epsilon) \rceil],$$

这里

$$M = \max \left\{ 3\sqrt{2} (2 + \pi)n/4\pi, \frac{n}{n-1} \max |a_k| + 1 \right\} + \sqrt{2}.$$

通常, 成本是以函数计值次数来衡量的, 所以有以下推论

推论 3.1. 算出一个零点的平均成本不超过

$$\pi[5M^2/n + 28(1 + 0.75n)^2 \lceil \log_2(3\sqrt{2}n/4\epsilon) \rceil].$$

定理的证明由下述命题及引理给出.

命题 3.1. 设 Σ_d 是剖分中位于 C_d 和 C_{d+1} 之间的一个基本方体, B_d 是以 $\{0\} \times [d, d +$

1) 为轴的圆柱体. 记 σ_d 为 Σ_d 的整个包含在 B_d 内的四面体的数目, 有

$$\sigma_d \leq \begin{cases} 5 \cdot \text{vol}(\Sigma_d \cap B_d), & \text{若 } d = -1, \\ 14 \cdot \text{vol}(\Sigma_d \cap B_d) \cdot 2^{2d}, & \text{若 } d \geq 0. \end{cases}$$

这里 $\text{vol}(A)$ 表示 $C \times [-1, +\infty)$ 的子集 A 在三维欧氏空间中的通常体积.

注意命题 3.1 对基本方体的位置及圆柱体的半径均无限制.

证. 为方便起见, 记 $V_d = \text{vol}(\Sigma_d \cap B_d)$. 对 $d = -1$, 总有 $0 \leq V_d \leq 1$. 显然, 若 $V_d \leq 1/2$, 则 $\sigma_d = 0$; 若 $V_d = 1$, 则 $\Sigma_d \subset B_d$, 知 $\sigma_d = 5$. 若 $1/2 < V_d < 1$, 则 $\sigma_d \leq 1$, 这是因为 Σ_d 的每一垂直棱都与 Σ_d 的四个四面体接触. 在 $d = -1$ 的上述所有情形下, 命题 3.1 成立.

对 $d \geq 0$, $0 \leq V_d \leq 2^{-2d}$. 我们区分七种情况. 利用剖分的性质, 通过简单的体积分析即可得到相应的结论:

$$1) 0 \leq V_d \leq \frac{1}{4} (2^{-d})^2 = \frac{1}{4} 2^{-2d}, \sigma_d = 0;$$

$$2) \frac{1}{4} 2^{-2d} < V_d \leq \frac{3}{8} 2^{-2d}, \sigma_d \leq 2;$$

$$3) \frac{3}{8} 2^{-2d} < V_d \leq \frac{1}{2} 2^{-2d}, \sigma_d \leq 4;$$

$$4) \frac{1}{2} 2^{-2d} < V_d \leq \frac{6}{8} 2^{-2d}, \sigma_d \leq 7;$$

$$5) \frac{6}{8} 2^{-2d} < V_d \leq \frac{7}{8} 2^{-2d}, \sigma_d \leq 8;$$

$$6) \frac{7}{8} 2^{-2d} < V_d < 2^{-2d}, \sigma_d \leq 9;$$

$$7) V_d = 2^{-2d}, \sigma_d = 14.$$

在所有七种情况, $\sigma_d \leq 14 \cdot \text{vol}(\Sigma_d \cap B_d) / 2^{-2d}$. 命题 3.1 得证.

下面是关于计算过程的三个引理.

引理 3.1. 整个计算在以原点为中心的半径

$$M = \max \left\{ 3\sqrt{2} (2 + \pi)n/4\pi, \frac{n}{n-1} \max |a_k| + 1 \right\} + \sqrt{2}$$

的圆柱内进行.

证. 三角形的直径顶多是 $\sqrt{2}$, 故若一个三角形有一顶点在此大圆柱外, 则它的三条边都位于同轴的半径 $r = M - \sqrt{2}$ 的圆柱外. 由引理 2.1, 我们只需证明对于剖分的位于半径 r 的圆柱外的每一条棱 (z', z'') , 有 $|\arg F(z')/F(z'')| < 2\pi/3$ 即可, 这里若 $z \in C_{-1}$, $F(z) = z^n$, 否则 $F(z) = f(z)$. 这样一来, 任何全标三角形必须和半径 r 的圆柱体接触, 也就一定要整个包含在半径 M 的圆柱体内, 引理即可得出.

改写 $f(z) = z^n(1 + a_1/z + \cdots + a_n/z^n) = z^n(1 + g(z))$. 若 z' 及 z'' 均不在 C_{-1} , 我们有

$$|g(z'')| \leq |a_1|/r + \cdots + |a_n|/r^n < \max |a_k|/(r-1) \leq (n-1)/n,$$

$$|g(z') - g(z'')| \leq |a_1| \cdot \left| \frac{1}{z'} - \frac{1}{z''} \right| + \cdots + |a_n| \cdot \left| \frac{1}{z'^n} - \frac{1}{z''^n} \right|$$

$$\begin{aligned} &\leq \max |a_k| \cdot |z' - z''| \left(\frac{1}{r^2} + \frac{2}{r^3} + \cdots + \frac{n}{r^{n+1}} \right) < \sqrt{2} \max |a_k| / (r-1)^2 \\ &\leq \frac{\sqrt{2}}{r-1} \cdot \frac{n-1}{n} \leq \frac{n-1}{n} \sqrt{2} / \left(\frac{3\sqrt{2}(2+\pi)n}{4\pi} - 1 \right) \\ &< \frac{(n-1)\sqrt{2}}{n} / \frac{3\sqrt{2}(2+\pi)(n-1)}{4\pi} = \frac{4\pi}{3(2+\pi)n}, \\ &\left| \frac{g(z') - g(z'')}{1 + g(z'')} \right| \leq \frac{4\pi}{3(2+\pi)n} / \left(1 - \frac{n-1}{n} \right) = \frac{4\pi}{3(2+\pi)} < 1, \end{aligned}$$

所以

$$\begin{aligned} &|\arg F(z')/F(z'')| = |\arg f(z')/f(z'')| \\ &\leq n |\arg z' - \arg z''| + \left| \arg \left(1 + \frac{g(z') - g(z'')}{1 + g(z'')} \right) \right| \\ &\leq n \cdot \frac{\sqrt{2}}{3\sqrt{2}(2+\pi)n/4\pi} + \frac{\pi}{2} \left| \frac{g(z') - g(z'')}{1 + g(z'')} \right| \\ &< \frac{4\pi}{3(2+\pi)} + \frac{\pi}{2} \frac{4\pi}{3(2+\pi)} = \frac{2\pi}{3}. \end{aligned}$$

z' , z'' 之一或两者均在 C_{-1} 时的证明是类似的,事实上更简单一些,从略. 引理 3.1 证完.

引理 3.2. 在 C_d , $d \geq 0$ 以上, 计算完全在以 $\xi_1, \cdots, \xi_n$ 为中心, 半径 $(1 + 0.75n)2^{1/2-d}$ 的 n 个垂直圆柱体内进行.

证. 定理 2.1 说全标三角形某顶点与某零点距离不超过 $3n\delta/4 = 0.75n\delta$, 所以整个三角形与此零点的距离不超过 $(1 + 0.75n)\delta$. 在 C_d 以上, $\delta \leq 2^{(1/2)-d}$, 并注意这里谈的是水平距离, 引理 3.2 即已得出.

引理 3.3. 令 $D = \lceil \log_2(3\sqrt{2}n/4\epsilon) \rceil$, 这里 $\epsilon > 0$ 是对计算零点的精度要求. 那末, C_D 以上的计算是不必要的.

证. 从 C_D 开始, 三角形直径不超过 $\sqrt{2} \cdot 2^{-D} \leq 4\epsilon/3n$, 全标三角形某顶点与零点的距离已不大于 ϵ , 所以不必再算下去. 证完

定理 3.1 的证明. 由引理 3.1 及 3.2, 计算只在 C_{-1} 和 C_0 之间一个半径 M 的底圆柱体和 C_0 以上 n 个圆柱体阶梯内进行. 底圆柱体体积为 πM^2 , 据命题 3.1, 它顶多包含 $5\pi M^2$ 个四面体. 每一节小圆柱的体积是 $\pi[(1 + 0.75n) \cdot 2^{0.5-d}]^2 = \pi(1 + 0.75n)^2 \cdot 2^{1-2d}$, 这里 C_d 是小圆柱下底所在的平面. 据命题 3.1, 每节小圆柱顶多包含 $28\pi(1 + 0.75n)^2$ 个四面体. 圆柱阶梯共有 n 个, C_D 以上不必再算, 所以在 C_D 以下计算的可能区域内四面体的总数不超过

$$5\pi M^2 + 28\pi n(1 + 0.75n)^2 \lceil \log_2(3\sqrt{2}n/4\epsilon) \rceil.$$

注意每个四面体顶多允许计算通过一次, 每步轮迴顶多需要一次多项式计值, 定理 3.1 因而得证.

Kuhn 算法允许方块 Q 的中心移动, 并允许剖分网格伸缩. 这在实际计算中是很有意义的, 但不会给上述讨论带来任何新问题.

四、逼近零点

下面的内容在于比较上述 Kuhn 算法的成本估计和 Smale 对 Newton 方法所作的成本

估计, 两种方法的结构大不相同, 但两种成本估计的对比($n^3 \log(n/\varepsilon)$ 与 n^9/μ^7)却是启发性的.

这一节主要讨论逼近零点, 下一节是基于代数几何理论的一个概率的比较. 虽然逼近零点概念只是 Smale 为讨论 Newton 方法的效率而提出来的, 但为比较, 也考虑用 Kuhn 方法算逼近零点.

Newton 方法可叙述为: 设 $z_0 \in C$, 归纳地定义 $z_k = z_{k-1} - f(z_{k-1})/f'(z_{k-1})$. 称 z_0 是 f 的一个逼近零点. 如果 z_k 对所有 $k \geq 1$ 定义, 当 $k \rightarrow \infty$ 时 z_n 收敛到 z^* , 这里 $f(z^*) = 0$, 并且对所有 $k \geq 1$, $|f(z_k)/f'(z_{k-1})| < 1/2$.

下面两个引理参看 Smale 的文献[1].

引理 4.1. 设 $g(z) = \sum_{i=1}^m b_i z^i$ 是一复多项式, $b_1 \neq 0$. 那末, 存在奇点 $\theta \in C$ (即 $f'(\theta) = 0$), 使对 $k = 2, 3, \dots$, 有

$$|b_k/b_1|^{1/(k-1)} |f(\theta)/b_1| \leq K.$$

引理 4.2. 设 $c \geq 1$, $\rho_f = \min_{\theta, f'(\theta)=0} |f(\theta)|$. 如果 $0 < |f(z)| < \rho_f/(cK + K + 1)$, 则

$$|f(z')/f(z)| < 1/c,$$

这里 $z' = z - f(z)/f'(z)$ (即 Newton 方法).

关于常数 K , 要略加说明. Smale 对 $K = 4$ 证明了引理 4.1, 他也证明了 $K \geq 1$. 如果 Beiberbach 猜想得到证实, K 将大大改善. Smale 本人的猜想是 $K = 1$. 但现在, 我们总可以用 $K = 4$.

定理 4.1. 设 $c > 1$, $|f(z_0)| < \rho_f/(cK + K + 1)$. 那末, Newton 方法收敛到 $f(z) = 0$ 的一个解 z^* .

证. 由引理 4.2 及 ρ_f 的定义, 易知 $z_k = z_{k-1} - f(z_{k-1})/f'(z_{k-1})$ 对所有 $k \geq 1$ 具有意义.

若对某个 l , $f(z_0) \neq 0, \dots, f(z_{l-1}) \neq 0$ 但 $f(z_l) = 0$, 则归纳有 $f'(z_k) \neq 0$ 及 $z_{k+1} = z_k$, $k = l, l+1, \dots$. 结论已得.

今若对所有 k , $f(z_k) \neq 0$. 首先由引理 4.2 及

$$|f(z_k)| = |f(z_k)/f'(z_{k-1})| \cdots |f(z_1)/f'(z_0)| \cdot |f(z_0)| < |f(z_0)|/c^k$$

可得

$$\lim_{k \rightarrow \infty} f(z_k) = 0.$$

现令 $L = \min_{f(z) \leq \rho_f/(cK+K+1)} |f'(z)|$, 显然 $L > 0$. 对任何自然数 k, m , 有

$$\begin{aligned} |z_{k+m} - z_k| &\leq \sum_{j=1}^m |z_{k+j} - z_{k+j-1}| = \sum_{j=1}^m |f(z_{k+j-1})/f'(z_{k+j-1})| \\ &\leq \frac{1}{L} \sum_{j=1}^m |f(z_{k+j-1})| \leq \frac{|f(z_0)|}{L} \sum_{j=1}^m \left(\frac{1}{c}\right)^{k+j-1} \\ &< |f(z_0)|/Lc^k \left(1 - \frac{1}{c}\right). \end{aligned}$$

所以序列 $\{z_k\}$ 收敛到某个 z^* , 且由 f 之连续性可知 $f(z^*) = \lim_{k \rightarrow \infty} f(z_k) = 0$. 定理 4.1 得证.

定理对于无重零点的多项式确立了逼近零点的存在性. 事实上我们有形式更明显的

推论 4.1. 若 $|f(z)| < \rho_f/(2K + K + 1) = \rho_f/(3K + 1)$, 则 z 是 f 的一个逼近零点.

五、计算逼近零点的成本

概率估计基于多项式空间中的体积分析. 令 P_n 记首项系数为 1 的 n 阶多项式的空间, 则 P_n 与 C^n 等同, 当 $f(z) = z^n + a_1 z^{n-1} + \cdots + a_n$ 时可写 $f \in C^n$. 令

$$P(R) = \{f \in P_n \mid |a_i| < R, i = 1, \cdots, n\},$$

则 $P(R)$ 的体积 $\text{vol } P(R)$ 是 $(\pi R^2)^n$, 这里对 P_n 用 $C^n = R^{2n}$ 中通常的体积.

令 $W_0 = \{f \in P_n \mid f(\theta) = 0, f'(\theta) = 0 \text{ 对某个 } \theta\}$. W_0 是有重零点的多项式的集合. 令

$$U_\rho(W_0) = \bigcup_{f_0 \in W_0} U_\rho(f_0),$$

这里 $U_\rho(f_0) = \{f \in P_n \mid |f(0) - f_0(0)| < \rho \text{ 且对所有 } z, f'(z) = f_0'(z)\}$. $U_\rho(W_0)$ 是容易引起麻烦的多项式的集合.

定理 5.1.

$$\frac{\text{vol}[U_\rho(W_0) \cap P(R)]}{\text{vol } P(R)} \leq \frac{(n-1)\rho^2}{R^2}.$$

引理 5.1. $f \in U_\rho(W_0)$ 当且仅当 $\rho_f < \rho$.

其证明均见文献[1]. 但要注意, Smale 把定理错写成 $n\rho^2/R^2$.

Kuhn 算法可按任意精度得到零点, 其成本由定理 3.1 给出估计. 当充分接近一个单零点使得 $|f(z)| < \rho_f/(3K+1)$ 时, 对于 Newton 方法的一个逼近零点也就得到了. 由此我们提出下述定理

定理 5.2. 设 $0 < \mu < 1$ 和 n 给定. 令 $\sigma = (\mu/n)^{1/2}$, 则对任何 $f \in P(R)$, 下述为真的概率是 $1 - \mu$: 用 Kuhn 算法, 最多 s 次轮回就可给全部 n 个零点各找到一个使用 Newton 方法的逼近零点, 这里

$$s = \lceil 5\pi M^2 + 28n(1 + 0.75n)^2 \lceil \log_2(3\sqrt{2}n/4\epsilon) \rceil \rceil,$$

而

$$M = \max \left\{ 3\sqrt{2} (2 + \pi)n/4\pi, \frac{n}{n-1} \max |a_k| + 1 \right\} + \sqrt{2},$$

$$\epsilon = \sigma(1 - M)^2/13[1 - (n+1)M^n + nM^{n+1}].$$

证. 若 $\sigma R \leq \rho_f$, 且 z 对 f 的某零点 ξ 满足 $|z - \xi| < \epsilon$, 则

$$\begin{aligned} |f(z)| &= |f(z) - f(\xi)| = \left| \sum_{j=0}^n a_j(z^j - \xi^j) \right| \leq \epsilon R \sum_{j=1}^n jM^{j-1} \\ &\leq \epsilon R \left(\sum_{j=1}^n x^j \right)'_M = \epsilon R \left(\frac{x - x^{n+1}}{1-x} \right)'_M \\ &\leq \epsilon R \frac{1 - (n+1)M^n + nM^{n+1}}{(1-M)^2} \leq \frac{\sigma R}{13} \leq \frac{\rho_f}{13} \leq \frac{\rho_f}{3K+1}, \end{aligned}$$

知 z 是 f 的一个逼近零点.

据定理 3.1, 对满足 $\rho_f \geq \sigma R$ 的任何多项式均可在 s 步内对它的所有零点各找到一个逼近零点, 剩下只需证 $\text{vol}\{f \in P(R) \mid \rho_f < \sigma R\} / \text{vol } P(R) < \mu$, 但这由引理 5.1 和定理 5.1 即得. 定理 5.2 证完.

Smale 的结果是:

定理 5.3. 设 $0 < \mu < 1$ 和 n 给定, $R > \frac{1}{3}$. 令 $\sigma = \left(\frac{\mu}{150}\right)^{3/2} (n+2)^{-2}$. 那末, 对任何 $f \in P(R)$, 下述为真的概率是 $1 - \mu$: 对于 $z_0 = 0$ 及适当的 $0 < h < 1$, $z_k = z_{k-1} - h \cdot f(z_{k-1})/f'(z_{k-1})$ 对一切 $k \geq 1$ 均有定义, 且 z_s 是 f 的一个逼近零点, 这里

$$s = \left\lceil 4 \left\{ 3 + \frac{8R \log(15/\sigma)}{\sigma^2} \right\}^2 \right\rceil.$$

在 $R = 1$ 情形, 由此得出 $s = \lceil 100(n+2) \rceil^2 / \mu^7$.

比较定理 5.3 和 5.2, 通过细致的计算可知, 即使对于计算所谓逼近零点, 本文对 Kuhn 算法得到的结果亦远胜于 Smale 对 Newton 方法建立的结果. Smale 的处理看来是十分出色的, 所以我们倾向于认为 n^9/μ^7 与 $n^3 \log(n/\mu)$ 的对比反映着算法本身之间的某种对比.

作者对项武忠教授和 H. Kuhn 教授的建议深表感谢.

参 考 文 献

- [1] Smale, S., *Bulletin AMS*, 4(1981). 1—36.
 [2] Kuhn, H., in *Fixed Points: Algorithms and Applications*, Academic Press, 1977, 11—40.