A Universal Constant for the Convergence of the Newton Method

Jean-Claude Yakoubsohn
Université Paul Sabatier, Toulouse

February 28, 1994

[summary by Xavier Gourdon]

Abstract
A new theorem is given concerning the convergence of the Newton method. In this result appears the constant $h_0 = 0.162434 \ldots$ which plays a fundamental part in the localization of “good” initial points.

1. Introduction

Given an algebraic equation over $\mathbb{C}$, $P(z) = 0$, it is well-known that the Newton iteration

$$z_0 \in \mathbb{C}, \quad z_{n+1} = z_n - \frac{P(z_n)}{P'(z_n)} \quad (1)$$

converges to a solution $z^*$ provided the initial value $z_0$ is sufficiently close to $z^*$. This iteration is generally used as a refining step in a root finding algorithm to increase the accuracy of the solutions, for example in the exclusion algorithm described in [1]. The problem of giving sufficient conditions on $z_0$ for (1) to converge is classical. For example, the Newton-Kantorovitch theorem [2, p. 263] states that under the condition

$$2 \left| \frac{P(z_0)}{P'(z_0)^2} \right| \sup_{|z-z_0|<\epsilon} |P''(z)| < 1, \quad (2)$$

with some $\epsilon > 0$, the Newton iteration is well defined and converges to the unique solution in $|z-z_0| < 2|P(z_0)/P'(z_0)|$ of the equation $P(z) = 0$. This result presents two disadvantages in practice: condition (2) is not expressed at only one point $z_0$, and the discs of unicity of a solution are generally small. The first result concerning the convergence of Newton method with a punctual criterion is given by Smale in [3]. A new result of this type is given in the following.

Theorem 1. Let $P$ be a univariate complex polynomial of degree $d$. Let $h_0 \approx 0.162434 \ldots$ be the first positive root of the polynomial $4h^3 - 12h^2 + 8h - 1$. Let $z_0 \in \mathbb{C}$ and $h \in [0, h_0]$ such that

$$\left| \frac{P^{(k)}(z_0)}{P'(z_0)^k} \right| \leq h^{k-1}, \quad 2 \leq k \leq d. \quad (3)$$

Then (convergence) the Newton iteration (1) converges to a simple solution $z^*$ of the algebraic equation $P(z) = 0$; (complexity) the convergence is super-quadratic, that is

$$|z_{n+1} - z_n| \leq a^n |z_1 - z_0| \left( \frac{h}{a^2} \right)^{2n-1},$$

49
where \( a = 2h_0^2 - 4h_0 + 1 \approx 0.404488 \ldots \) and \( h_0/a^2 \approx 0.990156 \ldots \); (set of unicity) for \( z \in \mathbb{C} \), define the polynomials in \( t \)

\[
L(z, t) = 1 - \sum_{k=1}^{d-1} \frac{|P^{(k)}(z)|}{k!} t^{k-1} \quad \text{and} \quad \overline{L}(z, t) = tL(z, t) - |P(z)|.
\]

Denote by \( \ell(z^*) \) the positive root of \( L(z^*, t) \). The polynomial \( \overline{L}(z, t) \) is concave over \( \mathbb{R} \) and admits either no real roots or two positive roots \( \ell^-(z) \leq \ell^+(z) \). Then each set of form \( |z - z_n| < \ell^+(z_n) \) for the indices \( n \) such that \( \ell(z^*) \geq \ell^-(z_n) \) (this happens for \( n \) large) contains only one solution of \( P(z) = 0 \) which is \( z^* \).

This result generalizes well for algebraic systems [4].

2. Proof of convergence

It is interesting to give a general idea of the proof to understand the origin of the universal constant \( h_0 \). Suppose \( z_0 \) satisfies conditions (3). A first inequality on \( P(z_1) \) is easily derived:

\[
|P(z_1)| = \left| P \left( z_0 - \frac{P(z_0)}{P'(z_0)} \right) \right| \leq \sum_{k=2}^{d} \frac{h^{k-1}}{k!} |P(z_0)| \leq \frac{h}{1 - h} |P(z_0)|.
\]

Next, we would like \( z_1 \) to satisfy conditions (3). Expanding, it is easy to obtain the inequalities

\[
\left| \frac{P^{(k)}(z_1)P(z_1)^{d-k}}{P'(z_1)^k} \right| \leq h^{k-1} \left( \frac{h}{1 - h} \right)^{k-1} S_{k, d}(h), \quad 2 \leq k \leq d
\]

where

\[
S_{k, d}(h) = \sum_{i=0}^{d-k} \binom{k + i}{i} h^i \quad \text{and} \quad T_{d}(h) = 1 - \sum_{i=1}^{d-1} (i + 1) h^i.
\]

Thus, we need \( Y_{k, d}(h) = h^{k-1} S_{k, d}(h) - (1 - h)^{k-1} T_{d}(h) \) to be negative. It is technical but feasible to show that the polynomials \( Y_{k, d} \) have only one positive root \( y_{k, d} \), and that they satisfy \( Y_{k, d}(h) < 0 \) for \( 0 \leq h \leq y_{2, d} \). The sequence \( y_{k, d} \) is strictly decreasing and tends to the smallest root \( h_0 \approx 0.162434 \ldots \) of the polynomial \( 4h^3 - 12h^2 + 8h - 1 \) (therefore it is possible to replace \( h_0 \) by \( y_{2, d} \) in the theorem). Now, by induction, inequality (4) leads to \( |P(z_n)| \leq (\frac{h}{1 - h})^n |P(z_0)| \), showing that \( P(z_n) \to 0 \) and by continuity, \( (z_n) \) converges to a solution \( z^* \) of \( P(z) = 0 \).

3. Conclusion

This result gives a good refining algorithm that fits well with the exclusion method [1]. The result of stability in the theorem also provides good bounds for a classical homotopy method: starting from the roots of a polynomial \( Q(z) \), we find the roots of \( P(z) \) by finding those of the polynomials \( H_t(z) = tP(z) + (1 - t)Q(z) \) for successive values of \( t \) between 0 and 1.

Bibliography


