Alpha Theory to Certify Roots

Most applications require verification of computational results [5]. Examining the convergence of Newton's method [1], we explain approximate roots. In [6], polynomials with approximate coefficients are considered.

1 Point Estimates for Approximate Zeroes

Given a function f in one variable x, Newton's method defines an iteration

$$x_{k+1} = N_f(x_k), \quad k = 0, 1, \ldots \text{ and } N_f(x) = x - \frac{f(x)}{f'(x)}, \tag{1}$$

provided the derivative f' does not vanish at any of the x_k's.
We define z to be *an approximate zero for a root* ζ *of* f $(f(\zeta) = 0)$ if

1. $x_{k+1} = N_f(x_k)$ is defined for all $k = 0, 1, \ldots$ starting with $x_0 = z$; and
2. the convergence is quadratic:

$$|x_k - \zeta| \leq \left(\frac{1}{2}\right)^{2^k - 1} |z - \zeta|. \tag{2}$$

Observe that an approximate zero z is a very economical representation for a root ζ. With one step of Newton's method we double the accuracy of the approximation. We want to find a criterion to determine whether a given number is an approximate zero or not. In the affirmative case, this criterion then becomes a certificate for the computed approximation.

To formulate a criterion to determine if a given number is an approximate zero, we need an auxiliary function (the general definition for any f is in [1, page 156], here we take $f \in \mathbb{C}[x]$):

$$\begin{array}{rcl} \gamma & : & \mathbb{C}[x] \times \mathbb{C} \rightarrow \mathbb{R} \\ & & (f, x) \mapsto \gamma(f, x) = \displaystyle\max_{k=2}^{\deg(f)} \left| \frac{f^{(k)}(x)}{k!\ f'(x)} \right|^{1/(k-1)} \end{array} \tag{3}$$

where $f^{(k)}$ is the kth derivative of f. Note that $\gamma(f, x)$ is undefined for $f'(x) = 0$.

For polynomials in one variable, evaluating γ is straightforward with a computer algebra system. Consider for example: $\gamma(x^3 - 1, x) = \max \left(\left| \frac{6x}{2!\ 3x^2} \right|, \left| \frac{6}{3!\ 3x^2} \right|^{1/2} \right)$, defined for all $x \neq 0$. If we know the exact root ζ, then we can already formulate a criterion for an approximate zero:

Theorem 1.1 *Assume $f(\zeta) = 0$ and $f'(\zeta) \neq 0$. If $|x - \zeta| \leq \frac{3 - \sqrt{7}}{2\gamma(f, \zeta)}$, then x is an approximate zero for ζ.*

The number $\frac{3 - \sqrt{7}}{2\gamma(f, \zeta)}$ defines the radius (in the complex plane) of a disk centered at ζ for which Newton's method converges quadratically. For a criterion computation at a point x itself, we need to define two more auxiliary functions (also here we restrict to polynomials f):

$$\begin{array}{rcl} \beta & : & \mathbb{C}[x] \times \mathbb{C} \rightarrow \mathbb{R} \\ & & (f, x) \mapsto \beta(f, x) = |x - N_f(x)| = \left| \frac{f(x)}{f'(x)} \right| \end{array} \tag{4}$$

and

$$\begin{array}{rcl} \alpha & : & \mathbb{C}[x] \times \mathbb{C} \rightarrow \mathbb{R} \\ & & (f, x) \mapsto \alpha(f, x) = \beta(f, x)\gamma(f, x). \end{array} \tag{5}$$

Theorem 1.2 *There is a universal constant α_0: if $\alpha(f, x) < \alpha_0$ then x is an approximate zero for a root ζ of f and $|x - \zeta| \leq 2\beta(f, x)$.*

In [1], the value of this constant is reported as $\alpha_0 = \frac{1}{4}(13 - 3\sqrt{17}) \approx 0.157671$. An improvement by X. Wang and D. Han to $3 - 2\sqrt{2} \approx 0.171573$ is reported in [4].

2 The Telescope System

The following system is inspired on an example in [2, page 219]:

$$\begin{cases} x_1 - \alpha = 0 \\ x_2 - x_1^2 = 0 \\ x_3 - x_2^2 = 0 \\ x_4 - x_3^2 = 0 \end{cases} \tag{6}$$

where we consider the constant α as a parameter. Obviously, there is only one solution: $(\alpha, \alpha^2, \alpha^4, \alpha^8)$. Although the system (6) consists only of quadrics, its is a telescope for its solution, depending whether $\alpha < 1$ or $\alpha > 1$, the solution either decreases or increases. This telescoping effect poses challenges for numerical methods. The difficulties that arise when dealing with high degree polynomials using limited precision do not disappear when introducing new variables as in the telescope.

3 Rewriting Taylor Series Expansions

In the first lemma below we see that the form of the function $\gamma(f, \zeta)$ function comes from a Taylor series development of the polynomial f about the root ζ.

Lemma 3.1 *For a polynomial f in one variable x of degree d and with root ζ, $f'(\zeta) \neq 0$, we have:*

$$|N_f(x) - \zeta| \leq \left|\frac{f'(\zeta)}{f(x)}\right| \left[\sum_{k=2}^{d} (k-1) \left(\gamma(f,\zeta)|x-\zeta|\right)^{k-1}\right] |x - \zeta|. \tag{7}$$

Proof. We develop f as a Taylor series about ζ, using $f(\zeta) = 0$:

$$f(x) = \sum_{k=1}^{d} \frac{f^{(k)}(\zeta)}{k!}(x-\zeta)^k \quad \text{and} \quad f'(x) = \sum_{k=1}^{d} \frac{f^{(k)}(\zeta)}{(k-1)!}(x-\zeta)^{k-1}. \tag{8}$$

As $\frac{1}{(k-1)!} - \frac{1}{k!} = \frac{k-1}{k!}$:

$$(x-\zeta)f'(x) - f(x) = \sum_{k=1}^{d} (k-1)\frac{f^{(k)}(\zeta)}{k!}(x-\zeta)^k. \tag{9}$$

So we have:

$$N_f(x) - \zeta = \frac{1}{f'(x)} \left[\sum_{k=2}^{d} (k-1)\frac{f^{(k)}(\zeta)}{k!}(x-\zeta)^{k-1}\right] (x-\zeta) \tag{10}$$

$$= \frac{f'(\zeta)}{f'(x)} \left[\sum_{k=2}^{d} (k-1)\frac{f^{(k)}(\zeta)}{k! f'(\zeta)}(x-\zeta)^{k-1}\right] (x-\zeta). \tag{11}$$

Then by repeated application of $|a+b| \leq |a| + |b|$ we obtain

$$|N_f(x) - \zeta| \leq \left|\frac{f'(\zeta)}{f'(x)}\right| \left[\sum_{k=2}^{d} (k-1) \left(\left|\frac{f^{(k)}(\zeta)}{k! f'(\zeta)}\right|^{1/(k-1)} |x-\zeta| \right)^{k-1}\right] |x-\zeta|. \tag{12}$$

The statement of the lemma in (7) follows from the definition of $\gamma(f, \zeta)$ in (3). $\square$

The first lemma already gives a relation between the original error $|x - \zeta|$ and the error after one Newton iteration: $|N_f(x) - \zeta|$. In the second lemma we simplify this relationship.

Lemma 3.2 *For a polynomial f in one variable x of degree d and with root ζ, $f'(\zeta) \neq 0$, we have:*

$$|N_f(x) - \zeta| \leq \left| \frac{f'(\zeta)}{f(x)} \right| \frac{u}{(1-u)^2} |x - \zeta|, \quad \textit{for } u = \gamma(f,\zeta)|x-\zeta| < 1. \tag{13}$$

Proof. Recall the sum of the geometric series $\displaystyle\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$ and its derivative $\displaystyle\sum_{k=1}^{\infty} k r^{k-1} = \frac{1}{(1-r)^2}$, which holds for all $r \geq 0$ and $r < 1$.
Assuming $u < 1$, We apply these sums to the summation in the right hand side of (7):

$$\sum_{k=2}^{d} (k-1) \left(\gamma(f,\zeta)|x-\zeta|\right)^{k-1} \leq \sum_{k=1}^{\infty} (k-1) u^{k-1} = \sum_{k=1}^{\infty} k u^{k-1} - \sum_{k=1}^{\infty} u^{k-1} \tag{14}$$

$$= \frac{1}{(1-u)^2} - \frac{1}{1-u} = \frac{u}{(1-u)^2}. \tag{15}$$

This proves the lemma. □

Observe that Lemma 3.2 imposes a restriction on $u = \gamma(f,\zeta)|x-\zeta| < 1$, but because Theorem 1.1 leads to the bound $u = \gamma(f,\zeta)|x-\zeta| \leq \frac{3-\sqrt{7}}{2} < 1$, this restriction on u is satisfied for approximate zeroes. Also the next lemma imposes an upper bound on u which is also satisfied for approximate zeroes.

Lemma 3.3 *For a polynomial f in one variable x of degree d and with root ζ, $f'(\zeta) \neq 0$, we have:*

$$|N_f(x) - \zeta| \leq \left(\frac{u}{1 - 4u + 2u^2} \right) |x - \zeta|, \quad \textit{for } u = \gamma(f,\zeta)|x-\zeta| < 1 - \frac{\sqrt{2}}{2}. \tag{16}$$

Proof. Comparing (13) with (16), we need to show $\displaystyle\left| \frac{f'(\zeta)}{f(x)} \right| \leq \frac{(1-u)^2}{1-4u+2u^2}$.

We first look at $\displaystyle\frac{f'(x)}{f'(\zeta)}$:

$$\frac{f'(x)}{f'(\zeta)} = \frac{1}{f'(\zeta)} \left[\sum_{k=1}^{d} \frac{f^{(k)}(\zeta)}{(k-1)!} (x-\zeta)^{k-1} \right] \tag{17}$$

$$= \frac{1}{f'(\zeta)} \left[f'(\zeta) + \sum_{k=2}^{d} \frac{f^{(k)}(\zeta)}{(k-1)!} (x-\zeta)^{k-1} \right] \tag{18}$$

$$= 1 + \sum_{k=2}^{d} \frac{f^{(k)}(\zeta)}{(k-1)! f'(\zeta)} (x-\zeta)^{k-1} \tag{19}$$

$$= 1 + \sum_{k=2}^{d} k \left(\left(\frac{f^{(k)}(\zeta)}{k! f'(\zeta)} \right)^{1/(k-1)} (x-\zeta) \right)^{k-1} \tag{20}$$

$$= 1 + B, \text{ where } B = \sum_{k=2}^{d} k \left(\left(\frac{f^{(k)}(\zeta)}{k! f'(\zeta)} \right)^{1/(k-1)} (x-\zeta) \right)^{k-1}. \tag{21}$$

Now we can estimate $\displaystyle\frac{f'(\zeta)}{f'(x)}$:

$$\left| \frac{f'(\zeta)}{f'(x)} \right| = \frac{1}{|1+B|} \leq \sum_{k=0} |B|^k = \frac{1}{1-|B|}, \tag{22}$$

provided $|B| < 1$. We verify that $u < 1 - \frac{\sqrt{2}}{2}$ leads to $|B| < 1$:

$$|B| \leq \sum_{k=2}^{d} k u^{k-1} = \sum_{k=1}^{d} k u^{k-1} - 1 \leq \sum_{k=1}^{\infty} k u^{k-1} - 1 = \frac{1}{(1-u)^2} - 1. \tag{23}$$

To verify whether $|B| < 1$ we solve $\frac{1}{(1-u)^2} - 1 = 1$ and find two positive roots. The smallest positive root is $1 - \frac{\sqrt{2}}{2}$, so $|B| < 1$ is implied by $u < 1 - \frac{\sqrt{2}}{2}$. Putting it all together leads to

$$\left| \frac{f'(\zeta)}{f'(x)} \right| \leq \frac{1}{1 - \left(\frac{1}{(1-u)^2} - 1 \right)} = \frac{(1-u)^2}{1 - 4u + 2u^2} \tag{24}$$

and the lemma is proven. □

Proof of Theorem 1.1. For x to be an approximate zero of f for the root ζ we must have $|N_f(x) - \zeta| \leq \frac{1}{2}|x - \zeta|$. Applying (16) from Lemma 3.3, we consider $\frac{u}{1-4u+2u^2} = \frac{1}{2}$. This equation has two positive solutions: $\frac{3 \pm \sqrt{7}}{2}$. We take the smallest solution as a bound for $u = \gamma(f,\zeta)|x - \zeta| \leq \frac{3-\sqrt{7}}{2}$, which leads to the criterion $|x - \zeta| \leq \frac{3-\sqrt{7}}{2\gamma(f,\zeta)}$. □

4 Contraction Maps

In this section we outline some of the ideas of [1, §8.2] in the proof of Theorem 1.2.

We denote by $B(r,z)$ a closed disk of radius r, centered at z, or formally: $B(r,z) = \{ x \in \mathbb{C} \mid |x - z| \leq r \}$.

Lemma 4.1 *If $g : B(r,z) \rightarrow B(r,z)$ is continuously differentially with $|g'(x)| \leq c$ for all $x \in B(r,z)$, then $|g(y_1) - g(y_2)| \leq c|y_1 - y_2|$ for all $y_1, y_2 \in B(r,z)$.*

If the constant c in the Lemma above is strictly less than 1, we call g a contraction and with contraction constant c. Note that at a root ζ is a fixed point of the Newton operator: $N_f(\zeta) = \zeta$. So we want $|N_f(z)| \leq 1$. The proof of the next proposition relates $\alpha(f,z)$ with $\beta(f,z)$ and $\gamma(f,z)$.

Proposition 4.1 *For all z: $|N_f(z)| \leq 2\alpha(f,z)$.*

Proof. We take the derivative of $N_f(z) = z - f(z)/f'(z)$:

$$\begin{aligned} N_f'(z) &= 1 - \frac{f'(z)f'(z) - f(z)f''(z)}{[f'(z)]^2} &(25)\\ &= \frac{f(z)f''(z)}{[f'(z)]^2} &(26)\\ &= \frac{f''(z)}{f'(z)} \frac{f(z)}{f'(z)} &(27)\\ &= \frac{f''(z)}{f'(z)} \beta(f,z). &(28) \end{aligned}$$

By definition of $\gamma(f,z)$, we have $|f''(z)/(2f'(z))| \leq \gamma(f,z)$, so $|N_f'(z)| \leq 2\gamma(f,z)\beta(f,z) = 2\alpha(f,z)$. □

For the Newton operator to be a contraction map, we thus must have that $2\alpha(f,z) < 1$ so 0.5 is a strict upper bound on the universal constant α_0.

The following theorem considers the application of the Newton operator to a closed disk.

Theorem 4.1 *If* $r < \dfrac{1-\sqrt{2}/2}{\gamma(f,z)}$ *then for all* $x \in B(r,z)$*:*

$$N_f'(x) \leq \frac{2(\alpha(f,z)+u)}{[1-4u+2u^2]^2}, \quad u = r\gamma(f,z) \tag{29}$$

and

$$N_f(B(r,z)) \subset B\left(\frac{2(\alpha(f,z)+u)}{[1-4u+2u^2]^2}, N_f(z)\right). \tag{30}$$

As a corollary to this theorem, for $u < 1-\sqrt{2}/2$, N_f is a contraction map for the disk $B(u/\gamma(f,z),z)$ onto itself with contraction constant $2(\alpha(f,z)+u)/[1-4u+2u^2]^2 < 1$.

Theorem 4.2 (Robust α Theorem) *There are positive real constants* u_0 *and* α_0 *such that:*
If $\alpha(f,z) < \alpha_0$*, then there is a root* ζ *of* f *such tat*

$$B\left(\frac{u_0}{\gamma(f,z)}, z\right) \subset B\left(\frac{3-\sqrt{7}}{2\gamma(f,\zeta)}, \zeta\right) \tag{31}$$

and N_f *maps* $B\left(\dfrac{u_0}{\gamma(f,z)}, z\right)$ *into* $B\left(\dfrac{u_0}{\gamma(f,\zeta)}, \zeta\right)$ *with contraction constant* $\leq 1/2$*.*

Theorem 1.2 follows from the proof of the Robust α Theorem.

5 Approximate Zeroes of Systems

For a system $f(\mathbf{x}) = \mathbf{0}$ of n equations in n variables $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, we consider the complex projective space $\mathbb{P}^n$ and use coordinates $[z_0 : z_1 : z_2 : \cdots : z_n]$. When using these homogeneous coordinates we have n equations in $n+1$ variables. For a unique answer, we will restrict Newton's method to the tangent space $T_{\mathbf{z}}$ at $\mathbf{z}$:

$$T_{\mathbf{z}} = \{ \ \mathbf{y} \in \mathbb{P}^{n+1} \mid z_0y_0 + z_1y_1 + z_2y_2 + \cdots + z_ny_n = 0 \ \}. \tag{32}$$

When we compute the inverse of the Jacobian matrix J_f of f we restrict to $T_{\mathbf{z}}$, denoted by $J_f(\mathbf{z})|_{T_{\mathbf{x}}}^{-1}$. Then the Newton operator N_f for systems in projective space is $N_f(\mathbf{z}) = \mathbf{z} - J_f(\mathbf{z})|_{T_{\mathbf{z}}}^{-1} f(\mathbf{x})$.

To illustrate the projective Newton method, consider the following example:

$$f(x_1,x_2) = \left\{ \begin{array}{r} x_1x_2 - 1 = 0 \\ x_1^2 - 0.002 = 0 \end{array} \right. \tag{33}$$

where the constant `0.002` could be viewed as a parameter. The smaller we take this constant, the larger the value for x_2 because of the vertical asymptote at $x_1 = 0$, so the need for projective coordinates is justified. Replacing x_1 by z_1/z_0 and x_2 by z_2/z_0 leads to the homogeneous polynomials $h_1(\mathbf{z}) = z_1z_2 - z_0^2$ and $h_2(\mathbf{z}) = z_1^2 - 0.002z_0^2$. The linear system we solve to compute the update $\Delta\mathbf{z}$ by Newton's method is

$$\left[\begin{array}{ccc} -2z_0 & z_2 & z_1 \\ -0.004z_0 & 2z_1 & 0 \\ z_0 & z_1 & z_2 \end{array} \right] \Delta\mathbf{z} = - \left[\begin{array}{c} h_1(\mathbf{z}) \\ h_2(\mathbf{z}) \\ 0 \end{array} \right]. \tag{34}$$

The coordinates for $\mathbf{z}$ can be computed for this small example as first evaluating `x1 = sqrt(0.002)` and $x_2 =$ `1/sqrt(0.002)` to some precision, then taking for $z_0 = 1/\max(x_1,x_2)$. The solution to the linear system above computed in a higher working precision will give $\mathbf{z} + \Delta\mathbf{z}$ as a more accurate approximation for the solution.

The criterion for an approximate zero of a system, uses a generalized γ function, as in [1, page 262]:

$$\begin{array}{rcl} \gamma \; : \; \mathbb{P}[x] \times \mathbb{P}^n & \rightarrow & \mathbb{R} \\ (f, x) & \mapsto & \gamma(f, x) = ||\mathbf{x}|| \max_{k=2}^{d} \left|\left| J_f(\mathbf{x})|_{T_\mathbf{x}}^{-1} \frac{D^k f(x)}{k!} \right|\right|^{1/(k-1)} \end{array} \quad (35)$$

where d is the largest degree of the polynomials in f, and $D^k f$ is the kth derivative of f, considered as a k-linear map, $D^1 f = J_f$. With $|| \cdot ||$ we denote vector and matrix norms. Evaluating γ becomes harder, but most polynomial systems in practice are of low degree and one can estimate $||D^k f(\mathbf{x})||$. An application to robotics is worked out in [3]. The same restrictions on the regularity of the root apply, i.e.: the Jacobian matrix J_f must be invertible.

Theorem 5.1 *Let $f(\mathbf{x}) = \mathbf{0}$ be a polynomial system with coordinates in projective space and $\zeta \in \mathbb{P}^n$ a root. If*

$$d_T(z, \zeta)\gamma(f, \zeta) \leq \frac{3 - \sqrt{7}}{2}, \quad d_T(z, \zeta) = \frac{||z - \zeta||}{||\zeta||} \quad (36)$$

then z is an approximate zero of f for ζ.

The function d_T is the so-called Riemannian distance function.

6 Exercises

1. Consider formula (2) for an approximate zero z close enough to a zero ζ of f. Assuming $|z - \zeta| \leq 0.1$ and given some $\epsilon > 0$, derive a bound on k, the number of iterations so that $|x_k - \zeta| \leq \epsilon$. Illustrate your bound with numerical examples.

2. Define a Maple function to evaluate γ for any polynomial f at any point x. Instead of Maple you may also use another computer algebra system.

3. Consider the polynomial $f(x) = x(x - r)$, where r is a parameter for the second root of f.

   (a) For r = 1, compute the radius $\frac{3-\sqrt{7}}{2\gamma(f,\zeta)}$ for all roots ζ.

   (b) Let r go to zero (take $r = 1/10^k$, for $k = 1, 2, \ldots$) and recompute $\frac{3-\sqrt{7}}{2\gamma(f,\zeta)}$ for all the roots. What do you observe as the second root gets closer to zero?

   (c) Let r grow larger (take $r = 10^k$, for $k = 1, 2, \ldots$) and recompute $\frac{3-\sqrt{7}}{2\gamma(f,\zeta)}$ for all the roots. What do you observe as the second root moves farther away?

4. Suppose that ζ_1 and ζ_2 are two regular roots of f, i.e.: $f(\zeta_1) = 0 = f(\zeta_2)$ and $f'(\zeta_1) \neq 0 \neq f'(\zeta_2)$. Show that ζ_1 and ζ_2 are separated as follows:

$$|\zeta_1 - \zeta_2| \geq \frac{5 - \sqrt{17}}{4\gamma(f, \zeta_2)}.$$

5. Use a computer algebra package to write a procedure to perform one step of the projective Newton method. On input are the polynomials in the system $f(\mathbf{x}) = \mathbf{0}$ with an approximate solution in the original affine coordinates. The procedure performs the transformation to projective coordinates $\mathbf{z}$, computes the extended Jacobian matrix and solves a linear system for $\Delta\mathbf{z}$. On return is the updated solution in projective and affine coordinates.

References

[1] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation.* Springer-Verlag, 1998.

[2] P. Bürgisser, M. Clausen, and M.A. Shokrollahi. *Algebraic Complexity Theory.* Springer–Verlag, 1997.

[3] M. Khan. Certifying approximate solutions of polynomial systems. Master's thesis, University of Illinois at Chicago, 2007.

[4] M. Petković. *Point Estimation of Root Finding Methods*, volume 1933 of *Lecture Notes in Mathematics.* Springer–Verlag, 2008.

[5] P.J. Roache. *Verification and Validation in Computational Science and Engineering.* Hermosa, 1998.

[6] V. Sharma, Z. Du, and C.K. Yap. Robust approximate zeros. In G. Støltìng Brodal and S. Leonardi, editors, *Algorithms - ESA 2005, 13th Annual European Symposium, Palma de Mallorca, Spain, October 3-6, 2005, Proceedings*, volume 3669 of *Lecture Notes in Computer Science*, pages 874–886. Springer-Verlag, 2005.