Robust Newton Method for Polynomials

Bahman Kalantari

1 Introduction

Consider a complex polynomial

$$p(z) = a_n z^n + \cdots + a_1 z + a_0, \tag{1}$$

with coefficients $a_j \in \mathbb{C}$, $z = x + iy$, $i = \sqrt{-1}$, and $x, y \in \mathbb{R}$. The Newton iterations are defined recursively by the formula

$$z_{j+1} = z_j - \frac{p(z_j)}{p'(z_j)}, \quad j = 0, 1, ..., \tag{2}$$

where $z_0 \in \mathbb{C}$ is the starting point, or *seed*, and $p'(z)$ is the derivative of $p(z)$. The sequence $\{z_j\}_{j=0}^{\infty}$ is called the *orbit* of z_0. The *basin of attraction* of a root θ of $p(z)$ is the set of all seeds z_0 whose orbit converges to θ. The *immediate* basin of attraction of a root is the *maximal connected component* of the basin of attraction containing the root. An interpretation of Newton iterations is the application of *fixed point* iterations to the rational function $N_p(z) = z - p(z)/p'(z)$. Notably, the iterate z_{j+1} is undefined if z_j is a *critical point* of $p(z)$, *i.e.,* if $p'(z_j) = 0$. When the iterate z_{j+1} is defined, it can be interpreted as the root of linearized approximation to $p(z)$.

The *modulus* of a complex number $z = x + iy$ is $|z| = \sqrt{x^2 + y^2}$. Equivalently, $|z| = \sqrt{z\overline{z}}$, where $\overline{z} = x - iy$ is the *conjugate* of z. In general, Newton iterates do not necessarily monotonically decrease the modulus of the polynomial after each iteration: at an arbitrary step j, we may obtain $p(z_{j+1})$ such that $|p(z_{j+1})| > |p(z_j)|$. For example, if $p(z) = z^2 - 1$, then $|p(z_{j+1})| >> |p(z_j)|$ for small $|z_j|$. However, near *simple* roots θ (i.e. $p'(\theta) \neq 0$), the rate of convergence is quadratic, thus requiring very few iterations to get highly accurate approximations. Another drawback of the Newton Method is that its orbits need not even converge; some cycle. For instance, in the case of $p(z) = z^3 - 2z + 2$, the Newton iterate at $z_0 = 0$ is $z_1 = 1$ and the iterate at z_1 is z_0, resulting in a cycle between 0 and 1. Finally, an orbit may converge yet be complex to the point of practical unwieldiness, even when $p(z)$ is a cubic polynomial.

In this forgoing approach we consider minimization of the modulus of a complex polynomial $p(z)$. It is more convenient to consider square of the modulus,

$$F(z) = |p(z)|^2 = p(z)\overline{p(z)}. \tag{3}$$

Clearly, θ is a root of $p(z)$ if and only if $F(\theta) = 0$.

2 The Robust Newton Method

Given $z_0 \in \mathbb{C}$ with $p(z_0) \neq 0$, define the following quantities:

$$
\begin{array}{l|l|l}
k = \min\{j \in \{1,\dots,n\} : p^{(j)}(z_0) \neq 0\} & & \\
u_k = \dfrac{1}{k!} p(z_0)\overline{p^{(k)}(z_0)} & \theta = \begin{cases} 0, & \text{if } c_k = |\gamma|, \gamma < 0 \\ \pi/k, & \text{if } c_k = |\gamma|, \gamma > 0 \\ \pi/(2k), & \text{if } c_k = |\delta|, \delta < 0 \\ 3\pi/(2k), & \text{if } c_k = |\delta|, \delta > 0 \end{cases} & A = \max_{j \geq 0}\{\frac{|p^{(j)}(z_0)|}{j!}\}. \\
\gamma = \quad 2 \cdot Re(u_k^{k-1}) & & \\
\delta = -2 \cdot Im(u_k^{k-1}) & & \\
c_k = \max\{|\gamma|, |\delta|\} & &
\end{array} \tag{4}
$$

In particular, note that A is equal to the maximum of the modulus taken over the coefficients of the Taylor expansion of $p(z)$ at z_0.

Definition 1. The *robust Newton iterate* at z_0 is

$$
\widehat{N}_p(z_0) = z_0 + \frac{C_k}{3}\frac{u_k}{|u_k|}e^{i\theta}, \quad C_k = \frac{c_k |u_k|^{2-k}}{6A^2}. \tag{5}
$$

We call $(u_k/|u_k|)e^{i\theta}$ the *normalized robust Newton direction* at z_0. We call $C_k/3$ the *step-size*. In particular, when $k = 1$, we have $c_1 = 2$ and $\theta = \pi$. Thus $e^{i\theta} = e^{i\pi} = -1$, and $C_1 = |u_1|/3A^2$ so that

$$
\widehat{N}_p(z_0) = z_0 - \frac{p(z_0)\overline{p'(z_0)}}{9A^2} = z_0 - \frac{|p'(z_0)|^2}{9A^2}\left(\frac{p(z_0)}{p'(z_0)}\right). \tag{6}
$$

Definition 1 defines the Robust Newton Method everywhere, including at critical points. In particular, when $k = 1$ the normalized robust Newton direction is simply a positive scalar multiple of the standard Newton direction. Also, by the definition of A, $|p'(z_0)|^2/9A^2 \leq 1/9$. Thus the robust Newton iterate always lies on the line segment between z_0 and the standard Newton iterate, $z_0 - p(z_0)/p'(z_0)$. This seemingly simple modification when $k = 1$, together with the ability to define the iterates when $k > 1$, will guarantee that the polynomial modulus at the new point, $z_1 = \widehat{N}_p(z_0)$, will necessarily decrease by a computable estimate.

3 A Generic Robust Newton Method

Algorithm 1 Robust Newton Method

```
Pick z_0 ∈ ℂ
t ← 0
while |p(z_t)p'(z_t)| ≠ 0,  do
    z_{t+1} ← N̂_p(z_t), t ← t + 1
end while
```

Except at most $(n-1)$ critical points, the index k equals 1 so that the iterate is defined according to (6). This simple modification of Newton Method assures global convergence to a root or a critical point of $p(z)$, while reducing $|p(z)|$ at each iteration.

The orbit of z_0 will necessarily converge to a root of $p(z)$ if $|p(z_0)|$ is below a critical threshold. This threshold is the minimum of $|p(z)|$ taken over critical points that are not roots of $p(z)$.

Now let $\varepsilon \in (0, 1)$ be a selected tolerance. We wish to iterate the algorithm until z_t satisfies $|p(z_t)| \leq \varepsilon$. However, the algorithm may produce instead a point with $|p(z_t)p'(z_t)| \leq \varepsilon$. To turn the generic algorithm into a practical one we consider a modification.

It can be shown that as long as $|p(z_t)p'(z_t)| \geq \varepsilon$, each iteration of the Robust Newton Method decreases $F(z)$ by at least $\varepsilon^2/9A^2$. When $|p(z_t)| > \varepsilon$, but $|p(z_t)p'(z_t)| < \varepsilon$, the decrement could be small. In such a case $p'(z_t)$ is small, giving an indication that the subsequent iterates may be converging to a critical point. To avoid this, we treat z_t as if it were a critical point and redefine its index as the smallest k such that $|p(z_t)p^{(k)}(z_t)|/k! \geq \varepsilon$. Then we proceed to define the next iterate as if z_t were a critical point with index k. Such an index is well defined for any ε less than $|p^{(n)}(z)|/n!$, the modulus of the coefficient of z^n in $p(z)$. Since we have adjusted the next iterate, z_{t+1}, the inequality $|p(z_{t+1})| < |p(z_t)|$ may not hold. If the inequality holds, we have succeeded to reduce the modulus and proceed as usual. However, if $|p(z_{t+1})| \geq |p(z_t)|$, we return to z_t and proceed to compute the next robust Newton iterate, repeating this process. Eventually, using this scheme, either we avoid convergence to a critical point while monotonically reducing $|p(z)|$, or the sequence of iterates will near a critical point. However, by continuity and the formula for robust Newton iterate at a critical point, we can be assured that the scheme explained here will escape the critical point and from that point on we proceed as usual.

3.1 Examples

Example 1. Consider the case where $p(z) = z^2 - 1$. As proven by Cayley, for any seed z_0 not on the y-axis, the orbit of z_0 under Newton iteration defined by $N_p(z) = z - (z^2-1)/2z$ converges to the root closest to z_0. No point on the y-axis converges. However, the orbits are different for the Robust Newton Method. We consider robust Newton iterations for $z_0 = 0$ and $z_0 = \varepsilon i$, $\varepsilon > 0$. Consider $z_0 = 0$. From Definition 1 and the values $p(0) = -1$, $p'(0) = 0$, and $p''(0)/2 = 1$, we get $A = 1$, $k = 2$, $u_2 = -1$, $u_2/|u_2| = -1$, $\gamma = -2$, $c_2 = 2$, $\theta = 0$, $e^{i\theta} = 1$ and $C_2 = 1/3$. It follows from (5) that the robust Newton iterate is $z_1 = -1/9$. The decrement is $F(z_1) - F(z_0) \approx -2/81$. Next, let $z_0 = \varepsilon i$, $\varepsilon > 0$. Then $p(z_0) = -(1+\varepsilon^2)$, $p'(z_0) = 2\varepsilon i$. Thus $k = 1$, $A = \max\{1+\varepsilon^2, 2\varepsilon, 1\} = 1 + \varepsilon^2$. Substituting these into (6) we get, $z_1 = \widehat{N}_p(z_0) = \varepsilon i - 2\varepsilon i/9(1+\varepsilon^2)$. We see that z_1 is closer to the origin than z_0 by a factor that improves iteratively. Thus, starting with any $\varepsilon \in (0, \infty)$, the sequence $z_{k+1} = \widehat{N}_p(z_k)$ monotonically converges to the origin, a critical point. By virtue of the fact the robust Newton iterate is defined at the origin, we adjust the iterates so as to avoid convergence to it. We treat a near-critical point as if it is critical point and compute the next iterate accordingly. Thus for ε small we treat $z_0 = \varepsilon i$ as if it is a critical point with index $k = 2$. We get $u_2 = p(z_0)p''(z_0)/2 = -(1+\varepsilon^2)$. Proceeding to define the robust Newton iterate with $k = 2$, $u_2/|u_2| = -1$, $\gamma = 2u_2 = -2(1+\varepsilon^2)$. Thus $c_2 = 2(1+\varepsilon^2)$ and $\theta = 0$ so that $e^{i\theta} = 1$ and $C_2 = 1/3(1+\varepsilon^2)$. Thus the robust Newton iterate becomes $z_1 = -1/9(1+\varepsilon^2) + \varepsilon i$. It is easy to see that for ε small enough $|p(z_1)| < 1 = |p(0)| < |p(z_0)|$. This together with the fact that in each iteration the Robust Newton Method decreases the current polynomial modulus implies the subsequent iterates will never get closer to the origin. In summary, by treating a near-critical point as a critical point and using the robust Newton iterates we have bypassed a critical point for good.

Example 2. Consider $p(z) = z^3 - 2z + 2$ at $z_0 = 0$ and $z_0 = 1$, the points in a cycle. If we pick $z_0 = \sqrt{2/3}$, a critical point, Newton's method is not defined. The only way to decrease the modulus here is to move into the complex plane; doing so is possible with the Robust Newton Method, and the orbit of z_0 will converge to a root as expected.

4 Programming Project

A number of programming projects can be described based on the Robust Newton Method. Some of them are described below.

1. Implement the Robust Newton Method and test it on some polynomial. You can also combine it with the usual Newton Method in this way:

Given an iterate z_t, generate Newton's iterate $z_{t+1} = z_t - p(z_t)/p'(z_t)$. If $|p(z_{t+1}| < |p(z_t)|$, pick z_{t+1} to be the next iterate. Otherwise, using z_t select z_{t+1} to be based on Robust Newton Method. Give a thorough description of the performance of these on some generated polynomials.

2. Produce a polynomiography of the performance of Robust Newton Method, and the combination of Robust Newton Method and Newton Method. This means pick a square, divide it into pixels and for each pixel iterate the Robust Newton Method and do color coding.

3. Use the Robust Newton Method to compute all roots of $p(z)$. This will be described later.