Iteration Functions

1 Taylor Polynomial

Given a polynomial $p(z)$ of degree n and a complex number a we have

$$p(z) = p(a) + p'(a)(z-a) + \frac{p''(a)}{2!}(z-a)^2 + \cdots + \frac{p^{(n)}(a)}{n!}(z-a)^n. \tag{1}$$

Suppose θ is a root of $p(z)$, i.e. $p(\theta) = 0$. Let $z = \theta$, $a = z$ in the above we get

$$0 = p(\theta) = p(z) + p'(z)(\theta - z) + \frac{p''(z)}{2!}(\theta - z)^2 + \cdots + \frac{p^{(n)}(z)}{n!}(\theta - z)^n. \tag{2}$$

The above holds for any root of θ and for any z. Adding z to both sides of the above, define

$$B_1(z) \equiv z - p(z) = z + p'(z)(\theta - z) + \frac{p''(z)}{2!}(\theta - z)^2 + \cdots + \frac{p^{(n)}(z)}{n!}(\theta - z)^n. \tag{3}$$

Equivalently,

$$B_1(z) = z - p(z) = \theta + (p'(z) - 1)(\theta - z) + \frac{p''(z)}{2!}(\theta - z)^2 + \cdots + \frac{p^{(n)}(z)}{n!}(\theta - z)^n. \tag{4}$$

Note that $B_1(\theta) = \theta$. So θ is a fixed point of $B_1(z)$.
Given z_0, define the fixed point iteration

$$z_{k+1} = B_1(z_k), \quad k \geq 0. \tag{5}$$

What can we say about the fixed point iteration? Will it converge when z_0 is close to θ?
From (4) we can write

$$B_1(z) - \theta = (p'(z) - 1)(\theta - z) + \frac{p''(z)}{2!}(\theta - z)^2 + \cdots + \frac{p^{(n)}(z)}{n!}(\theta - z)^n. \tag{6}$$

Or,

$$B_1(z) - \theta = (\theta - z)\left((p'(z) - 1) + \frac{p''(z)}{2!}(\theta - z) + \cdots + \frac{p^{(n)}(z)}{n!}(\theta - z)^{n-1}\right). \tag{7}$$

Notice that we can write the above after factoring $(\theta - z)$ as

$$B_1(z) - \theta = (\theta - z)\Big((p'(z) - 1) + (\theta - z)G(z)\Big), \tag{8}$$

where $G(z)$ is a sum of terms. What is important is that when $(\theta - z)$ is small, $(\theta - z)G(z)$ is small. So if $|p'(\theta) - 1| < 1$. Then there is a neighborhood of the root θ so that for any z_0 in this neighborhood fixed point iteration converges to θ. A neighborhood of θ means the disc of some radius $r > 0$ centered at θ:

$$D_r(\theta) = \{z : |z - \theta| < r\}. \tag{9}$$

More formally, using the triangle inequality we can write,

$$|(p'(\theta)-1)+(\theta-z)G(z)| \leq |(p'(\theta)-1)|+|(\theta-z)G(z)|. \tag{10}$$

So for example if $|(p'(\theta)-1)| < .9$, there will be a neighborhood where $|(p'(z)-1)| < .95$, and $|(\theta - z)G(z)| < .1$ for every z in this neighborhood.

2 Newton Method

From (2) we also get

$$zp'(z)-p(z)=\theta p'(z)+\frac{p''(z)}{2!}(\theta-z)^2+\cdots+\frac{p^{(n)}(z)}{n!}(\theta-z)^n. \tag{11}$$

Dividing by $p'(z)$ we get Newton's iteration function

$$B_2(z) \equiv z-\frac{p(z)}{p'(z)}=\theta+\frac{p''(z)}{p'(z)2!}(\theta-z)^2+\cdots+\frac{p^{(n)}(z)}{p'(z)n!}(\theta-z)^n. \tag{12}$$

This means

$$B_2(z) \equiv z-\frac{p(z)}{p'(z)}-\theta \approx \frac{p''(z)}{p'(z)2!}(\theta-z)^2. \tag{13}$$

If $p'(\theta) \neq 0$, i.e. θ is a simple roots of $p(z)$, then $B_2(\theta)=\theta$, i.e. θ is a fixed point of $B_2(z)$. In fact multiple roots are also fixed points. And any fixed point of $B_2(z)$ is a root of $p(z)$.

Given z_0, define the fixed point iteration

$$z_{k+1}=B_2(z_k), \quad k \geq 0. \tag{14}$$

We have,

$$B_2(z)-\theta=\frac{p''(z)}{p'(z)2!}(\theta-z)^2+\cdots+\frac{p^{(n)}(z)}{p'(z)n!}(\theta-z)^n. \tag{15}$$

This means if $p'(\theta) \neq 0$, there is a neighborhood around θ so that starting with any z_0 in this neighborhood the fixed point iteration converges. Furthermore,

$$\lim_{k\to\infty}\frac{z_{k+1}-\theta}{(\theta-z_k)^2}=\frac{p''(\theta)}{p'(\theta)2!}. \tag{16}$$

We say the rate of convergence is quadratic. This, roughly speaking, says when z_k is close enough to a root θ the error in each iteration doubles. Thus if it is 10^{-1}, you expect the next errors to be approximately 10^{-2}, 10^{-4}, 10^{-8}, etc.

3 Newton Method for Multiple Roots

Given a polynomial $p(z)$ and a root θ we say it is a root of multiplicity m if

$$p(z)=(z-\theta)^m q(z),$$

where $m \geq 1$ and $q(\theta) \neq 0$. When $m=1$ we say θ is a simple root.

It is easy to show that in Newton's method,

$$B_2'(\theta)=1-\frac{1}{m}.$$

This implies

$$\lim_{k\to\infty}\frac{z_{k+1}-\theta}{(z_k-\theta)}=\frac{m-1}{m}. \tag{17}$$

4 Some Facts on Dynamics of Newton's Method

The basin of attraction of a root θ of a polynomial is the set of all input z_0 so that orbit of z_0, denoted by $O(z_0) = \{z_1, z_2, \ldots,\}$, converges to θ. The basin of attraction is denoted by $A(\theta)$.

What kind of a set it it?

A subset O of the Euclidean plane is said to be open if given any point z_0 in O, there is an open disk $D_r(z_0) = \{z : |z - z_0| < r\}$ that is contained in O.

We claim $A(\theta)$ is an open set. First, there is a neighborhood, open disk at θ, say $D_r(\theta)$ for which any point in it converges to θ.

Fact: Newton's iteration function is continuous at θ.

Fact: Under continuity, inverse image of an open set is an open set.

$B_2^{-1}(D_r(\theta)) = \{z : B_2(z) \in D_r(\theta)\}$.

Now if $O(z_0)$ lies in $D_r(\theta)$, that is if the orbit at z_0 converges to θ, we claim that there is an open neighborhood of z_0, say $D_t(z_0)$, for some $t > 0$, such that any point z' in $D_t(z_0)$, the orbit of z' converges to θ. To prove the existence of $D_t(z_0)$, we use the fact that since z_0 converge to θ this means after so many Newton iterations, say N iterations, all the subsequent iterates stay in $D_r(\theta)$. Now take the inverse of $D_r(\theta)$. This is an open set, say O_1. Take the inverse image of O_1, say O_2. This is an open set. So after N inverse images we get an open set O_N which contains z_0 and every point in O_N converges to θ.

Immediate basin of attraction: The largest connected component of $A(\theta)$ that is contains θ.

5 Halley Method

We start again with the equation

$$0 = p(z) + p'(z)(\theta - z) + \frac{p''(z)}{2!}(\theta - z)^2 + \cdots + \frac{p^{(n)}(z)}{n!}(\theta - z)^n. \tag{18}$$

Using the above and a mixture of $B_1(z)$ and $B_2(z)$ we would like to make a new iteration function that its order of convergence is cubit.

First note

$$B_1(z) - B_2(z) = -p(z) + \frac{p(z)}{p'(z)} = (p'(z) - 1)(\theta - z) + \sum_{i=2}^{n} \frac{(p'(z) - 1)p^{(i)}(z)}{i!p'(z)}(\theta - z)^i. \tag{19}$$

Multiply the above by $p(z)$ and (18) by $-(p'(z) - 1)(\theta - z)$ and adding, we get

$$p(z)(B_1(z) - B_2(z)) = p^2(z)\frac{(1 - p'(z))}{p'(z)} = \sum_{i=2}^{n} u_i(z)(\theta - z)^i, \tag{20}$$

where

$$u_i(z) = (p'(z) - 1)\left(\frac{p(z)p^{(i)}(z)}{i!p'(z)} - \frac{p^{(i-1)}(z)}{(i-1)!}\right). \tag{21}$$

Multiplying (20) by

$$\frac{-p''(z)}{2p'(z)u_2(z)} \tag{22}$$

and adding it to the expansion of $B_2(z)$ we get

$$B_3(z) \equiv z - p(z)\frac{p'(z)}{p'(z)^2 - p(z)p''(z)/2} = \theta + \sum_{i=3}^{n} v_i(z)(\theta - z)^i, \tag{23}$$

where

$$v_i(z) = \left(\frac{p^{(i)}(z)}{i!p'(z)} - \frac{p''(z)}{2p'(z)}\frac{u_i(z)}{u_2(z)}\right). \tag{24}$$

The above iteration function is called Halley's method and has cubic order of convergence. This roughly means near a simple root if the current error is 10^{-1}, the next one roughly 10^{-3}, 10^{-6}, etc.

6 Horner's Method

Consider a polynomial

$$p_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0.$$

To efficiently evaluate the polynomial at a point z_0 we use nested multiplication:

$$p(z_0) = (\cdots((a_n z_0 + a_{n-1})z_0 + a_{n-2})z_0 + \cdots + a_1)z_0 + a_0.$$

Let $b_n = a_n$ and recursively define

$$b_{n-m} = b_{n-m+1} z_0 + a_{n-m}, \quad m = 1, \ldots, n.$$

Performing n multiplications and n additions we evaluate $p(z)$ at z_0 to get

$$p(z_0) = b_0.$$

Let

$$p_1(z) = b_n z^{n-1} + b_{n-2} z^{n-2} + \cdots + b_2 z + b_1.$$

We have

$$p_1(z)(z - z_0) + b_0 = p(z).$$

From the above, differentiating we get

$$p'(z) = p_1'(z)(z - z_0) + p_1(z). \tag{25}$$

Thus

$$p'(z_0) = p_1(z_0).$$

So from Horner's recursion we get $p(z_0)$ and $p'(z_0)$. By repeating this process we can compute all the normalized derivatives of $p(z)$ at z_0. By induction from (25) we get

$$p^{(i)}(z) = p_1^{(i)}(z)(z - z_0) + i p_1^{(i-1)}(z). \tag{26}$$

Substituting $z = z_0$ gives

$$p^{(i)}(z_0) = i p_1^{(i-1)}(z_0), \quad i = 1, \ldots, n. \tag{27}$$

To summarize, given a polynomial

$$p_n(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0,$$

and a particular z_0, we can compute all the normalized derivatives of $p(z)$ at z_0:

$$b_i^{(i)} = \frac{p^i(z_0)}{i!}, \quad i = 0, \ldots, n,$$

Set

$$b_m^{(-1)} = a_m, \quad m = 0, \ldots, n.$$

For $i = 0, \ldots, n$, do

$$b_n^{(i)} = b_n^{(i-1)}.$$

For $m = 1, \ldots, n - i$, do

$$b_{n-m}^{(i)} = z_0 b_{n-m+1}^{(i)} + b_{n-m}^{(i-1)}.$$