Solution of Nonlinear Equations

1 Introduction

Finding the solutions of the nonlinear equations occurs often in the scientific computing. For example, let us consider the problem of finding the parameter λ in the curve $y = \lambda \cosh(x/\lambda)$ such that the length of the arc between $x = 0$ and $x = 5$ is 10. Now

$$10 = \int_0^5 \frac{ds}{dx} dx = \int_0^5 \cosh(x/\lambda)\, dx = \lambda \sinh(5/\lambda)$$

Hence, finding the appropriate curve needs the value of λ which can be found from the solution of the transcendal equation

$$\lambda \sinh(5/\lambda) = 10$$

In this chapter, we discuss few methods along with their convergence properties. Let α is a solution of $f(x) = 0$ and c_n is a approximation of the root. Here suffix n denotes an iteration index that will be introduced later. Now the error at the n-th stage is $e_n = \alpha - c_n$. If

$$|e_{n+1}| = A|e_n|^p,$$

then p is the order of the convergence and A is called the asymptotic rate constant. Clearly for $p = 1$, we need $A < 1$ for convergence.

2 Bisection method

Suppose that f is a continuous function on the interval $[a, b]$ and $f(a)f(b) < 0$. By intermediate value theorem, f has at least one zero in the interval $[a, b]$. We next calculate $c = (a + b)/2$ and test fc). If $f(c) = 0$, then c is the root and we are done. If not, then either $f(a)f(c) < 0$ or $f(b)f(c) < 0$. In the former case, a root lies in $[a, c]$ and we rename c as b and do the same process. In the latter case, we rename c as a and continue the same process. The root now lies in a interval whose length is half of the length of the original interval. The process is repeated and we stop the iteration when $f(c)$ is very nearly zero or length of the interval $[a, b]$ is very small.

2.1 Convergence

Let $a_0 = a$ and $b_0 = b$ and $[a_n, b_n]$ $(n \geq 0)$ are the successive intervals in the bisection process. Clearly

$$a_0 \leq a_1 \leq a_2 \leq \cdots \leq b_0 = b$$

and

$$b_0 \geq b_1 \geq b_2 \geq \cdots \geq a_0 = a$$

Now the sequence $\{a_n\}$ is monotonic increasing and bounded above and the sequence $\{b_n\}$ is monotonic decreasing and bounded below. Hence both the sequences converge. Further,

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} = \cdots = \frac{b - a}{2^n}$$

Hence $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \alpha$. Further, taking limit in $f(a_n)f(b_n) \leq 0$, we get $[f(r)]^2 \leq 0$ and that implies $f(r) = 0$. Hence, both a_n and b_n converges to a root of $f(x) = 0$.

Let us apply the bisection method to the interval $[a_n, b_n]$ and calculate midpoint $c_n = (a_n + b_n)/2$. Then the root lies either in $[a_n, c_n]$ or $[c_n, b_n]$. In either case

$$|\alpha - c_n| \leq \frac{b_n - a_n}{2} = \frac{b-a}{2^{n+1}}$$

Hence, $c_n \to \alpha$ as $n \to \infty$.

In this method, we can calculate the number of iteration n that need to be done to achieve a specified accuracy. Suppose we want relative accuracy ϵ of the root. Hence we want

$$\frac{|\alpha - c_n|}{|\alpha|} \leq \epsilon$$

Suppose that the root lies in $[a, b]$ where $b > a > 0$. Clearly $|\alpha| > a$ and hence the above relation is true if

$$\frac{|\alpha - c_n|}{a} \leq \epsilon$$

which is true if

$$\frac{b-a}{2^{n+1}a} \leq \epsilon$$

Solving this we can find minimum number of iteration needed to obtain the desired accuracy.

Now

$$|e_{n+1}| = |\alpha - c_{n+1}| \leq \frac{1}{2}(b_{n+1} - a_{n+1}) = \frac{1}{2}\frac{b_n - a_n}{2}$$

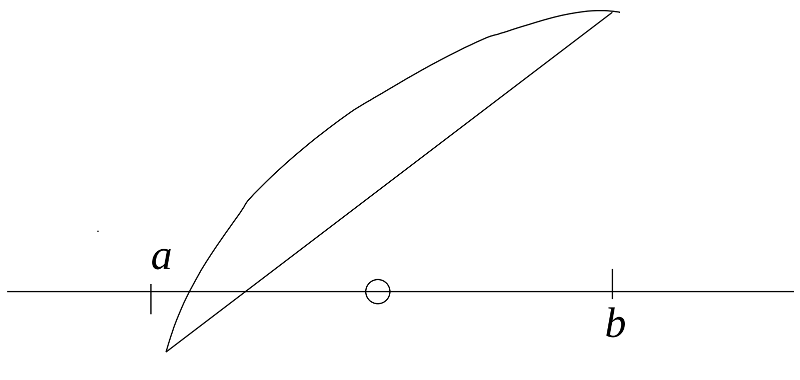
and

$$|e_n| = |\alpha - c_n| \leq \frac{1}{2}(b_n - a_n)$$

Thus we find

$$|e_{n+1}| \sim \frac{1}{2}|e_n|$$

Hence the bisection method converges linearly.



3 Regula falsi

Consider the figure in which the root lies between a and b. In the first iteration of bisection method, the approximation lies at the small circle. However, since $|f(a)|$ is small, we expect the root to lie near a. This can be achieved if we joint the coordinates $(a, f(a))$ and $(b, f(b))$

and take the intersection c of the line with the x axis as the first approximation. Hence the new approximation c satisfies

$$\frac{c-a}{b-a} = \frac{0-f(a)}{f(b)-f(a)} \implies c = \frac{af(b)-bf(a)}{f(b)-f(a)}$$

Since $f(a)$ and $f(b)$ are of opposite sign, the method is well defined. If $f(c) = 0$, then c is the exact root α. Otherwise we take $b = c$ if $f(a)f(c) < 0$ and $a = c$ if $f(c)f(b) < 0$. This process is then repeated. Of course, method does not work satisfactorily in all cases and certain modification can be made.

Let $[a_n, b_n]$ be the successive intervals of the regula falsi method. In this case $b_n - a_n$ may not go to zero as $n \to \infty$. However, the method still converges to the root.

To show this, we consider the worst case. We assume that $f''(x)$ exists and for some i, $f''(x) \geq 0$ in $[a_i, b_i]$. The case of $f''(x) \leq 0$ can be treated similarly. Also, suppose that $f(a_i) < 0$ and $f(b_i) > 0$. Let c_i be the new approximation which is nothing but the intersection of the straight line through $(a_i, f(a_i))$ and $(b_i, f(b_i))$ with the x-axis. We claim that $a_i < a_{i+1} = c_i < b_{i+1} = b_i$. To show that note that the straight line is nothing but the degree one polynomial $p(x)$ with $p(c_i) = 0$. Obviously $a_i < c_i < b_i$. For $x \in [a_i, b_i]$

$$f(x) - p(x) = (x-a_i)(x-b_i)f''(\eta)/2, \qquad \eta \in (a_i, b_i)$$

Putting $x = c_i$ and using the given conditions, we get

$$f(c_i) \leq 0$$

If $f(c_i) \neq 0$, then $a_{i+1} = c_i > a_i$ and $b_{i+1} = b_i$. Hence, if the condition holds, then $b_i = b_f$ (say) for $i \geq i_0$. Now a_i is monotonically increasing and bounded by b_f. Hence $\lim_{n\to\infty} a_i$ exists and is equal to ζ (say). Since f is continuous, $f(\zeta) = \lim_{n\to\infty} f(a_i) \leq 0$ and $f(b_f) > 0$ and hence $\zeta \neq b_f$. Taking limit in

$$c_i = \frac{a_i f(b_i) - b_i f(a_i)}{f(b_i) - f(a_i)}$$

we find

$$\zeta = \frac{\zeta f(b_f) - b_f f(\zeta)}{f(b_f) - f(\zeta)}$$

Hence we find

$$(\zeta - b_f)f(\zeta) = 0$$

Since $\zeta \neq b_f$, we must have $f(\zeta) = 0$ and a_i converges to ζ.

To find the order of convergence, let us consider the worst case discussed just above. Now writing x_i instead of a_i and $b_i = b$, the iteration can be written as

$$x_{i+1} = \frac{bf(x_i) - x_i f(b)}{f(x_i) - f(b)} = \phi(x_i)$$

where

$$\phi(x) = \frac{bf(x) - xf(b)}{f(x) - f(b)}$$

Hence after $i \geq i_0$, the regula falsi method usually becomes a fixed point iteration method. This method will converge if $|\phi'(\zeta)| < 1$. Now we can write

$$\phi'(\zeta) = 1 - \frac{\zeta - b}{f(\zeta) - f(b)} f'(\zeta) = 1 - \frac{f'(\zeta)}{f'(\eta_1)}, \qquad \zeta < \eta_1 < b.$$

By mean-value theorem, there exists $\eta_2 \in (x_i, \zeta)$ such that

$$f(x_i) - f(\zeta) = (x_i - \zeta) f'(\eta_2)$$

Since $f''(x) \geq 0$ in $[x_i, b]$, $f'(x)$ is monotonically increasing in $[x_i, b]$ and $f'(\eta_2) = f(x_i)/(x_i - \eta) > 0$. This implies

$$0 < f'(\eta_2) \leq f'(\zeta) \leq f'(\eta_1) \implies 0 < \frac{f'(\zeta)}{f'(\eta_1)} \leq 1$$

Hence

$$0 \leq 1 - \frac{f'(\zeta)}{f'(\eta_1)} < 1 \implies 0 \leq \phi'(\zeta) < 1.$$

Hence the fixed point iteration converges to ζ i.e. $\zeta = \phi(\zeta)$. Now

$$e_{n+1} = \zeta - x_{n+1} = \phi(\zeta) - \phi(x_n) = e_n \phi'(\eta) = K e_n,$$

where $0 \leq K < 1$. Hence the convergence is linear.

4 Secant method

Here we don't insist on bracketing of roots. Given two initial guess. Given two approximation x_{n-1}, x_n, we take the next approximation x_{n+1} as the intersection of line joining $(x_{n-1}, f(x_{n-1}))$ and $(x_n, f(x_n))$ with the x-axis. Thus x_{n+1} need not lie in the interval $[x_{n-1}, x_n]$. If the root is α and α is a simple zero, then it can be proved that the method converges for initial guess in sufficiently small neighbourhood of α. Now x_{n+1} is given by

$$x_{n+1} = \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})}$$

4.1 Convergence

Now

$$\begin{aligned}
e_{n+1} = \alpha - x_{n+1} &= \alpha - \frac{x_{n-1} f(x_n) - x_n f(x_{n-1})}{f(x_n) - f(x_{n-1})} \\
&= \alpha - x_n + \frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} f(x_n) \\
&= \alpha - x_i + \frac{f(x_n) - f(\alpha)}{f[x_n, x_{n-1}]} \\
&= -(\alpha - x_n)(\alpha - x_{n-1}) \frac{f[\alpha, x_n, x_{n-1}]}{f[x_n, x_{n-1}]} \\
&= -e_n e_{n-1} \frac{-f''(\eta_1)}{2f'(\eta_2)},
\end{aligned}$$

where $\eta_1 \in I(x_n, x_{n-1}, x_n)$ and $\eta_2 \in I(x_n, x_{n-1})$. Here, $I(a, b, c)$ denotes the interior of the interval formed by a, b and c. Since α is a simple zero, $f'(\alpha) \neq 0$. Consider the interval $J = \{x : |x - \alpha| \leq \delta\}$ such that

$$\left| \frac{-f''(\eta_1)}{2f'(\eta_2)} \right| \leq M, \qquad \eta_1, \eta_2 \in J$$

Now we have

$$|e_{n+1}| \le M|e_n||e_{n-1}|$$

Let $\varepsilon_n = M|e_n|$. Then $\varepsilon_{n+1} \le \varepsilon_n \varepsilon_{n-1}$. Now choose initial guess x_0 and x_1 such that

$$|x_i - \alpha| < \min\{1/M, \delta\}, \quad i = 0, 1$$

This implies $\varepsilon_i = M|x_i - \alpha| < \min\{1, M\delta\}$ for $i = 0, 1$. Now choose $0 < D < \min\{1, M\delta\}$ and thus $0 < D < 1$ and $\varepsilon_0, \varepsilon_1 \le D < 1$. Now

$$\varepsilon_2 \le \varepsilon_1 \varepsilon_0 \le D^2$$

Also, $\varepsilon_3 \le \varepsilon_2 \varepsilon_1 \le D^3$ etc. By induction, we can show that $\varepsilon_n \le D^{\lambda_n}$ where $\lambda_0 = \lambda_1 = 1$ and $\lambda_n = \lambda_{n-1} + \lambda_{n-2}$ for $n \ge 2$. Using $\lambda_n \propto r^n$, we find

$$\lambda_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{n+1}\right] \sim \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1}, \qquad \text{as} \quad n \to \infty$$

Since $D < 1$ and $\lambda_n \to \infty$, we get $\varepsilon_n \to 0$ as $n \to \infty$ and thus $x_n \to \alpha$.

Now as $x_n \to \alpha$, $\eta_1, \eta_2 \to \alpha$. This implies

$$|e_{n+1}| \sim C|e_n||e_{n-1}|, \qquad C = \left|\frac{f''(\alpha)}{2f'(\alpha)}\right|$$

Let

$$|e_{n+1}| \sim C^{\beta}|e_n|^p \implies |e_n| \sim C^{\beta}|e_{n-1}|^p \implies |e_{n-1}| = simC^{-\beta/p}|e_n|^{1/\beta}$$

Now from $|e_{n+1}| \sim C|e_n||e_{n-1}|$, we get

$$C|e_n|^p \sim C|e_n||e_n|^{1/p}C^{-\beta/p},$$

which is true provided

$$p = 1 + 1/p, \quad \beta = 1 - \beta/p \implies \beta = p/(1+p) = p - 1$$

Taking the positive value of p, we find $p = (1+\sqrt{5})/2 = r$ (golden ratio) and $\beta = r - 1$. Hence

$$|e_{n+1}| \sim C^{r-1}|e_n|^r$$

The order of convergence is non-integer which is greater than one. Hence, the convergence is superlinear.

5 Newton-Raphson method

Let x_0 be an initial guess to the root α of $f(x) = 0$. Let h is the correction i.e. $\alpha = x_0 + h$. Then $f(\alpha) = 0$ implies $f(x_0 + h) = 0$. Now assuming h small and f twice continuously differentiable, we find

$$f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(\eta) = 0, \qquad \eta \in I(x_0, x_0 + h]$$

Neglecting quadratic and higher order term and assuming that α is a simple root, we find

$$h \approx -f(x_0)/f'(x_0) \implies x_1 = x_0 - f(x_0)/f'(x_0)$$

might be a better approximation to α than x_0. We can continue this process with x_1. Hence, method is given by

$$x_{n+1} = x_n - f(x_n)/f'(x_n), \qquad n = 0, 1, 2, \cdots$$

Geometrically, x_{n+1} is the intersection of tangent with the $x-$axis that passes through the point $(x_n, f(x_n))$. This method can also be derived from the secant method if x_{n-1} approach x_n. The method may or may not converge if the initial guess is too far from the root.

5.1 Convergence

If α is simple root, the $f'(\alpha) \neq 0$ and hence $f'(x) \neq 0$ in a sufficiently small neighbourhood of α. Consider the interval $J = \{x : |x - \alpha| \leq \delta\}$ in which $f'(x) \neq 0$ and

$$\left| \frac{f''(\xi_n)}{2f'(x_n)} \right| \leq M, \qquad x_n, \xi_n \in J$$

If $e_n = \alpha - x_n$, then from $f(x_n + e_n) = 0$, we find

$$f(x_n) + e_n f'(x_n) + e_n^2 f''(\xi_n)/2f'(x_n) = 0$$

Now we write $e_n = \alpha - x_n$, divide both side by $f'(x_n)$ and use the iteration formula for Newton-Raphson method to arrive at

$$e_{n+1} = \left(-\frac{f''(\xi_n)}{2f'(x_n)} \right) e_n^2 \implies |e_{n+1}| \leq M|e_n|^2$$

Let $\varepsilon_n = M|e_n|$. Then $\varepsilon_{n+1} \leq \varepsilon_n^2$. Now choose initial guess x_0 such that

$$|x_0 - \alpha| < \min\{1/M, \delta\}$$

This implies $\varepsilon_0 = M|x_0 - \alpha| < \min\{1, M\delta\}$. Now choose $0 < D < \min\{1, M\delta\}$ and thus $0 < D < 1$ and $\varepsilon_0 \leq D < 1$. Now

$$\varepsilon_n \leq \varepsilon_{n-1}^2 \leq \varepsilon_{n-2}^{2^2} \leq \cdots \leq (\varepsilon_0)^{2^n} = D^{2^n}$$

Since $D < 1$, $\varepsilon_n \to 0$ as $n \to \infty$ and thus $x_n \to \alpha$ as $n \to \infty$.

Also, $|e_{n+1}| \sim C|e_n|^2$ which implies quadratic convergence and the asymptotic rate constant is $C = |f''(\alpha)/2f'(\alpha)|$.

Also, note that

$$-f(x_n) = f(\alpha) - f(x_n) = (\alpha - x_n)f'(c_n) \sim (\alpha - x_n)f'(x_n)$$

This implies

$$e_n = \alpha - x_n \sim -f(x_n)/f'(x_n) = x_{n+1} - x_n$$

Hence, the error is approximately the difference between the two successive iteration values. Thus the difference between the successive iteration values can be used for stopping criterion.

6 Fixed point iteration

In this method, one writes $f(x) = 0$ in the form $x = g(x)$ so that any solution of $x = g(x)$ (which is also called fixed point) is a solution of $f(x) = 0$. This can be accomplished in many ways. For example, with $f(x) = x^2 - 5$, we can write $g(x) = (x + 5/x)/2$ or $g(x) = x + 5 - x^2$ or $g(x) = x - (5 - x^2)/2$ etc.

The function $g(x)$ is also called an iteration function. Once an $g(x)$ is chosen, then we carry out the iteration (starting from initial guess x_0)

$$x_{n+1} = g(x_n), \qquad n = 0, 1, 2, \cdots$$

Theorem: Let g be defined in an interval $I = [a, b]$ such that $g(x) \in I$ i.e. $g(x)$ maps I into itself. Further, suppose that g is differentiable in I and there exists a nonnegative constant $K < 1$ such that $g'(x) \le K$ for all $x \in I$. Then there exists a unique fixed point ξ in I and $x_n \to \xi$ as $n \to \infty$.

Proof: If $g(a) = a$ or $g(b) = b$, then obviously g have a fixed point. Hence suppose that $g(a) \ne a$ and $g(b) \ne b$. Since g maps I into I, we must have $g(a) > a$ and $g(b) < b$. Now consider the function $h(x) = x - g(x)$ and we must have $h(a) < 0$ and $h(b) > 0$. By intermediate value theorem, there exists ξ such that $h(\xi) = 0$ and hence existence of fixed point is proved. To prove uniqueness, suppose that ξ and η are distinct fixed point. Then

$$|\xi - \eta| = |g(\xi) - g(\eta)| = |g'(\zeta)||\xi - \eta| \le K|\xi - \eta| < |\xi - \eta|, \quad \zeta \in I(\xi, \eta)$$

which is a contradiction. Hence fixed point is unique.

To prove convergence, consider $e_n = \xi - x_n$. Then

$$e_n = g(\xi) - g(x_{n-1}) = g'(c_n)e_{n-1}, \qquad c_n \in I(\xi, x_n)$$

Hence

$$|e_n| \le K|e_{n-1} \le K^2|e_{n-2}| \le \cdots \le K^n|e_0|$$

Since $0 \le K < 1$, we have $K^n \to 0$ as $n \to \infty$ and hence $e_n \to 0$ as $n \to \infty$.

Also, assuming that g is twice differentiable, we have

$$e_{n+1} = \xi - x_{n+1} = g(\xi) - g(x_n) = g(\xi) - g(\xi - e_n) = e_n g'(\xi) - \frac{e_n^2}{2} g''(c_n), \qquad c_n \in I(\xi, x_n)$$

If $g'(\xi) \ne 0$, then

$$|e_{n+1}| \sim A|e_n|, \qquad A \approx |g'(\xi)|$$

showing that the convergence is first order. On the other hand, when $g'(\xi) = 0$, then

$$|e_{n+1}| \sim C|e_n|^2, \qquad C \approx |g''(\xi)|/2$$

showing that the convergence is 2nd order. For example, the Newton-Raphson is a special case of fixed point iteration in which $g(x) = x - f(x)/f'(x)$. If ξ is a simple root of f, then the convergence of Newton-Raphson method is 2nd order.

It is often very difficult to verify the assumptions of the previous theorem. Hence many times, we check the following condition: If g is continuously differentiable in some open interval J containing ξ and if $|g'(\xi)| < 1$ in J, then there exists an $\delta > 0$ such that the fixed point iteration converges whenever we start with x_0 that satisfies $|x_0 - \xi| \le \delta$.

To show this we take $q = (1 + |g'(\xi))/2 < 1$ and take $\epsilon = (1 - |g'(\xi))/2 > 0$. Since g' is continuous, there exists a $\delta > 0$ such that

$$|g'(x) - g'(\xi)| \leq \epsilon \qquad |x - \xi| \leq \delta.$$

Now

$$|g'(x)| \leq |g'(x) - g'(\xi)| + |g'(\xi)| \leq \epsilon + |g'(\xi) = q$$

Consider $I = [\xi - \delta, \xi + \delta]$ and we show that g maps I to itself. To show this, we note that for $x \in I$

$$|g(x) - \xi| = |g(x) - g(\xi)| = g'(\eta)(x - \xi)$$

where $\eta \in I(x, \xi)$ and hence $\eta \in I = [\xi - \delta, \xi + \delta]$. Thus

$$|g(x) - \xi| \leq q|x - \xi| < \epsilon$$

showing that g maps I to I.

7 Roots of polynomial

Finding roots of polynomial also deals with complex roots. Also, sometimes we are interested in finding all the roots of a polynomial. We know that a polynomial of degree n has n roots (counting multiplicity) in the complex field. We first deal with some localization theorem.
Theorem: All roots of the polynomial

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

lie in the open disk whose centre is at the origin of the complex plane and whose radius is

$$\rho = 1 + |a_n|^{-1} \max_{0 \leq i \leq n} |a_i|$$

Proof: Let $c = \max_{0 \leq i \leq n} |a_i|$. If $c = 0$, then nothing to prove. Hence assume that $c > 0$ and then $\rho > 1$. Now we show that $p(z)$ does not vanish in the region $|z| \geq \rho$. To show this, we find (noting that $|z| \geq \rho > 1$ and $c|a_n|^{-1} = \rho - 1$)

$$\begin{aligned}
|p(z)| &\geq |a_n z^n| - |a_{n-1} z^{n-1} + a_{n-2} z^{n-2} + \cdots + a_0| \\
&\geq |a_n z^n| - c \sum_{i=0}^{n-1} |z|^n \\
&> |a_n z^n| - c|z|^n (|z| - 1)^{-1} \\
&= |a_n z^n| [1 - c|a_n|^{-1} (|z| - 1)^{-1}] \\
&\geq |a_n z^n| [1 - c|a_n|^{-1} (\rho - 1)^{-1}] = 0
\end{aligned}$$

Here we have used $|z| \geq \rho \implies |z| - 1 \geq \rho - 1$.
Corollary: Note that if we consider $s(z) = z^n p(1/z)$ then

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

Note that $p(z_0) = 0$ implies $s(1/z_0) = 0$. Hence, if all the roots of s lies inside the disk $|z| \leq \rho$, then all the nonzero roots of p are outside the disk $|z| < 1/\rho$.

8 Horner's algorithm

This is also known as nested multiplication and as synthetic division. For a polynomial $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ and a number z_0, we can write $p(z) = (x - z_0)q(z) + p(z_0)$ where

$$q(z) = b_{n-1} z^{n-1} + b_{n-2} z^{n-2} + \cdots + b_0$$

is polynomial of degree one less than that of p. Substituting $q(z)$ and equating like powers we find

$$b_{n-1} = a_n, b_{n-2} = a_{n-1} + b_{n-1} z_0, \cdots, b_0 = a_1 + b_1 z_0, p(z_0) = a_0 + b_0 z_0$$

Thus we can use Horner's algorithm to find value of a polynomial at any point z_0. This can also be used to deflate a polynomial if we know that z_0 is a root. This method also can be used to find Taylor expansion of a polynomial about any point. Suppose

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 = b_n (z - z_0)^n + b_{n-1} (z - z_0)^{n-1} + \cdots + b_0$$

Clearly $b_k = p^{(k)}(z_0)/k!$. We can use Horner's algorithm to find c_k efficiently. Since $c_0 = p(z_0)$ which is obtain by applying nested multiplication to $p(z)$. The method also gives

$$q(z) = (p(z) - p(z_0))/(z - z_0) = b_n (z - z_0)^{n-1} + b_{n-1} (z - z_0)^{n-2} + \cdots + b_1$$

Hence we can obtain c_1 by applying nested multiplication to $q(z)$. This process can be repeated.

9 Newton-Raphson method

We use the iteration

$$z_{k+1} = z_k - p(z_k)/p'(z_k), \qquad k = 0, 1, 2, \cdots$$

Note that $p(z) = (x - z_k)q(x) + p(z_k)$ and $p'(z_k) = q(z_k)$. Hence, both numerator and denominator can be obtained by two steps of nested multiplication. To obtain complex roots, we need a complex initial guess. Alternatively, we also can use $z = \alpha + i\beta$ and substituting in $p(z) = 0$, we get two equation $F(\alpha, \beta) = 0, G(\alpha, \beta) = 0$ that can be solved together using Newton-Raphson technique with real arithmetic.

Suppose we obtain a root α_1 by Newton-Raphson method. Then by deflation method, $p(z) \approx (z - \alpha_1)q_1(z)$ where q_1 is a polynomial of degree one less than p. Now we apply Newton-Raphson to q_1 and find another root α_2. We can proceed this way and find all the roots $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that

$$p(z) \approx (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n)$$

Of course, the error in the roots increases from α_2 to α_n since the error gradually built up in the deflation method. One remedy is to take α_2 to α_n as initial guess and work with the full polynomial.

10 Müller's method

This is generalization of secant method. This method works well for simple and multiple roots. This method may converge to a complex root even if we start with a real initial guesses. It works for non polynomial too. Here we fit a polynomial of degree 2 through three interpolatory points x_{i-2}, x_{i-1}, x_i:

$$\begin{aligned} p(x) &= f[x_i] + f[x_i, x_{i-1}](x - x_i) + f[x_i, x_{i-1}, x_{i-2}](x - x_i)(x - x_{i-1}) \\ &= f[x_i] + f[x_i, x_{i-1}](x - x_i) + f[x_i, x_{i-1}, x_{i-2}][(x - x_i)^2 + h_i](x - x_i) \\ &= a_i(x - x_i)^2 + 2b_i(x - x_i) + c_i, \end{aligned}$$

where $a_i = f[x_i, x_{i-1}, x_{i-2}]$, $b_i = [h_i f[x_i, x_{i-1}, x_{i-2}] + f[x_i, x_{i-1}]/2$, $c_i = f[x_i]$ and $h_i = x_i - x_{i-1}$. Let δ_i be the root of smallest absolute value of the quadratic equation $a_i\delta^2 + 2b_i\delta + c_i = 0$. Then $x_{i+1} = x_i + \delta_i$ is the root of $p(x) = 0$ closest to x_i. Note that

$$\delta_i = \frac{-b_i \pm \sqrt{b_i^2 - a_i c_i}}{2a_i} = -\frac{c_i}{b_i \pm \sqrt{b_i^2 - a_i c_i}}$$

We need the sign that make the denominator largest in absolute value. Hence

$$\delta_i = -\frac{c_i}{b_i + \operatorname{sgn}(b_i)\sqrt{b_i^2 - a_i c_i}}$$

Complex arithmetic have to be used since $b_i^2 - a_i c_i$ might be negative. Once we get x_{i+1}, then we repeat the same procedure with interpolatory points x_{i-1}, x_i and x_{i+1}. It can be shown that

$$e_{i+1} \sim e_i e_{i-1} e_{i-2} \left(-\frac{f^{(3)}(\alpha)}{6f'(\alpha)} \right)$$

Further, if $|e_{i+1}| \sim |e_i|^p$, then $p \approx 1.84$ and hence the convergence is superlinear and better than secant method.

11 Sensitivity of polynomial roots

Consider $p(x) = x^2 - 2x + 1$ which has real roots $x_1 = x_2 = 1$. If we change the coefficient -2 to -1.9999, the roots will be complex. Hence, the character of the roots change from real to complex. Of course the roots are equal here.

However, the well separated roots of a polynomial will be sensitive too. For example, consider Wilkinson polynomial

$$p(x) = (x - 1)(x - 2) \cdots (x - 20)$$

whose roots are simple, real and well separated. A small perturbation to the coefficient of x^{19} from -210 to $-210 + 2^{-23}$ will change the roots significantly and some of the becomes complex too. Hence, care must be exercised while finding the roots of a polynomial.