Introduction to Numerical Analysis I
Handout 5

1 Root finding

We consider equation $f(x) = 0$.

1.2 Fixed Point Iteration

Consider a sequence $\{x_n\}_{n=0}^{\infty}$ given by

$$x_{n+1} = g(x_n).$$

Each member of the sequence x_n is denoted an **iteration** and the function g is the **iteration function**. The main idea is that if $x_n \to \bar{x}$ and g is continuous then

$$\bar{x} = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} g(x_{n+1}) = g(\bar{x}).$$

Definition 1.1. A point $x \in X$ is called a fixed point of a function $g : X \to X$ if and only if $x = g(x)$.

Example 1.2. How to manipulate f to get fixed point iteration

1. The first option is to rewrite the formula such that x is be on the right hand side, for example $f(x) = x^2 - 2x - 3$ can be rewritten as $x = (x^2 - 3)/2 = g(x)$ and then

   $$x_{n+1} = (x_n^2 - 3)/2 = g(x).$$

2. The other option $g(f(x), x) = f(x) + x = x$ and then

   $$x_{n+1} = f(x_n) + x_n.$$

3. And even $g(f(x), x) = f(x)h(x) + x = x$ for $h(x) \neq 0$ and then

   $$x_{n+1} = f(x_n)h(x_n) + x_n.$$

   For example the Newton method suggests $h(x) = -1/f'(x)$.

Theorem 1.3 (Fixed Point Theorem). Let $g(x)$ be differentiable function in a closed interval $I = [a, b]$. If

- $g(x) \subset I$ for each $x \in I$ and
- $|g'(x)| < 1$ for each $x \in I$

then g is fixed point iteration function and there is unique $\bar{x} \in I$ such that $g(\bar{x}) = \bar{x}$. In this case $x_{n+1} = g(x_n) \to \bar{x}$ for any initial guess $x_0 \in I$. Furthermore, there is exists a number $0 < L < 1$ such that $e_n = |x_{n+1} - x| \leq L|x_n - x|$.

Proof: Existence: If $g(a) = a$ or $g(b) = b$ then we done. Otherwise, let $h(x) = g(x) - x$. Since since $a < g(x) < b$ it comes that $h(a) = g(a) - a > 0$ and $h(b) = g(b) - b < 0$ and therefore there is $\bar{x} \in I$ such that $0 = h(\bar{x}) = g(\bar{x}) - \bar{x}$. Thus $\bar{x} = g(\bar{x})$.

Uniqueness: Assume g has two different fixed points $\bar{x}_1 \neq \bar{x}_2$ in I. In this situation $h(\bar{x}_1) = 0 = h(\bar{x}_2)$ and therefore, by Roll's theorem, there is $c \in I$ such that $h'(c) = 0$. However $h'(c) = g'(c) - 1 = 0$ means $g'(c) = 1$ which contradicts the assumption that $|g'(x)| < 1$ for each $x \in I$.

Convergence:

$$\frac{|x_{n+1} - x|}{|x_n - x|} = \frac{|g(x_n) - g(x)|}{|x_n - x|} \to L = |g'(x_n)| < 1$$

Thus $e_n = |x_{n+1} - x| \leq L|x_n - x|$

Note: Actually, the requirement that g' is differentiable at the root is too strong. One may require instead that g is Lipschitz continues, that is there is a Lipschitz constant $0 < L < 1$ such that $|g(x) - g(y)| \leq L|x - y|$ for any $x, y \in I$. Any differentiable function is necessarily satisfy this condition with $L = \max_x g'(x)$.

Theorem 1.4 (Another Useful Theorem). Let $g \in \mathcal{C}^2$ be twice continuously differentiable fixed point iteration and let $\bar{x} = g(\bar{x})$ be iteration point. If $g'(\bar{x}) < 1$ then there is a neighborhood $N(\bar{x})$, s.t. $x_n \to \bar{x}$ for any initial guess $x_0 = N(\bar{x})$.

1.2.1 Convergence Rate

Consider linear approximation of the error: $e_{n+1} = x_{n+1} - r = -r + g(x_n) = -r + g(r + e_n) \approx \cancel{-r + g(r)} + e_n g'(r) = e_n g'(r)$. Thus we learn that $|e_{n+1}| \approx |e_n| \cdot |g'(r)|$ and that the smaller $|g'(r)|$ the faster the convergence. Furthermore, consider the Taylor expansion of the error

$$e_{n+1} \approx \cancel{-r + g(r)} + \sum_{k=1}^{N} \frac{e_n^k}{k!} g^{(n)}(r) + \frac{e_n^{N+1}}{(N+1)!} g^{(N+1)}(c),$$

for $c \in (r, r + e_{n+1}$. If we assume that the first N derivatives of g vanishes at r then we get that

$$\frac{|e_{n+1}|}{|e_n|^{N+1}} = g^{(N+1)}(c),$$

that is, the order of convergence is $p = N + 1$.

Example 1.5. Let $r > 0$. Consider we looking for an iteration to approximate $x = \sqrt[m]{r}$. Define

$$g(x) = \sum_{k=0}^{N} a_k x^{1-mk} r^k$$

. In order to find the coefficients a_k for the maximal OOC, one solves the following linear system:

- $g\left(r^{1/m}\right) = \sum_{k=0}^{N} a_k \left(r^{1/m}\right)^{1-mk} r^k = \sum_{k=0}^{N} a_k r^{1/m} = r^{1/m}$
- $g^{(k)}\left(r^{1/m}\right) = 0$ for all $k = 1, ...N$.

1.2.2 Series acceleration

Aitken δ^2 method Consider a convergent sequence $x_n \to r$. As we already saw

$$\frac{x_{n+1} - r}{x_n - r} = \frac{e_{n+1}}{e_n} \approx g'(r) \approx \frac{e_{n+2}}{e_{n+1}} \approx \frac{x_{n+2} - r}{x_{n+1} - r}$$

Solve, $\frac{x_{n+1}-r}{x_n-r} \approx \frac{x_{n+2}-r}{x_{n+1}-r}$ for r to get

$$r \approx \frac{x_{n+2}x_n - x_{n+1}^2}{x_{n+2} - 2x_{n+1} + x_n} = x_n - \frac{(x_n - x_{n+1})^2}{x_{n+2} - 2x_{n+1} + x_n}$$

One use it to create accelerated sequence

$$\tilde{x}_n = x_n - \frac{(x_n - x_{n+1})^2}{x_{n+2} - 2x_{n+1} + x_n} = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n},$$

where $\Delta x_n = x_n - x_{n+1}$ and therefore $\Delta^2 x_n = \Delta\Delta x_n = \Delta(x_n - x_{n+1}) = \Delta x_n - \Delta x_{n+1} = (x_n - x_{n+1}) - (x_{n+1} - x_{n+2})$.

The $\tilde{x}_n$ converges faster then x_n in sense of

$$\frac{\tilde{x}_n - r}{x_n - r} \to 0$$

Aitken Algorithm
For an iteration function $g(n)$ and initial guess x^0
1) calculate $x_1 = g(x_0)$ and $x_2 = g(x_1)$,
2) then calculate $\tilde{x}$
3) if $|\tilde{x} - x| > \textit{tollerance}$ continue to 1 with $x_0 = x_1$

There is no reason to use x_1, since $\tilde{x}$ is better approximation to r, thus the Aitken Algorithm can be improved by Steffensen.

Steffensen Algorithm
For an iteration function $g(n)$ and initial guess x^0
1) calculate $x_1 = g(x_0)$ and $x_2 = g(x_1)$,
2) then calculate $\tilde{x}$
3) if $|\tilde{x} - x| > \textit{tollerance}$ continue to 1 with $x_0 = \tilde{x}$

The algorithm have similar computational complexity problem as Newton's method: each step we have to compute 2 functions. The half of the consolation is that we don't have to know the derivative. Still, if function computation is costly it may become a disadvantage. The convergence rate is at best quadratic.

Steffenson Method Let $g(x)$ be fixed point iteration with a fixed point $r = g(r)$. If $f(r) = 0$ then $h(x) = g(x) - f(x)$ is also an iteration function with $h(r) = r$. Define $g(x) = x + f(x)$ and use Steffensen Algorithm:

Given x_n, one get $x_{n+1} = g(x_n) = x_n + f(x_n)$, then $\Delta x_n = x_n - x_{n+1} = -f(x_n)$ and $\Delta^2 x_n = -\Delta f(x_n) = f(x_{n+1}) - f(x_n) = f(x_n + f(x_n)) - f(x_n)$. Finally define

$$x_{n+1} = \tilde{x}_n = x_n - \frac{(f(x_n))^2}{f(x_n + f(x_n)) - f(x_n)}$$

The method has quadratic convergence rate. This is sort of Secant with $f'(x) \approx \frac{f(x_n + f(x_n)) - f(x_n)}{f(x_n)}$. Similarly, Applying $g(x) = x + f(x)$ to Secant would give

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

Finally change x_{n-1} with x_n to get the same result.

1.3 Durand–Kerner method

DK used for simultaneous finding of roots of polynomial. Here is example for polynom of degree 3. Consider a monic polynomial (coefficient of highest degree =1). $P_3(x) = x^3 + a_2x^2 + a_1x + a_0$ Let r_1, r_2, r_3 be roots of P_3, i.e. $P_3(x) = (x - r_1)(x - r_2)(x - r_3)$.

A Newton like iteration for r_1 is

$$r_1^{n+1} = r_1^n - \frac{P_n(r_1^n)}{(r_1^n - r_2)(r_1^n - r_3)}$$

the denominator is simply the derivative P'_m. Now lets find all roots, starting with initial guess $\vec{r}_0 = (r_1^0, r_2^0, r_3^0)$

while $\sqrt{|r_1^n|^2 + |r_2^n|^2 + |r_3^n|^2} > \epsilon$

$$r_1^{n+1} = r_1^n - \frac{P_n(r_1^n)}{(r_1^n - r_2)(r_1^n - r_3)}$$

$$r_2^{n+1} = r_1^n - \frac{P_n(r_2^n)}{(r_2^n - r_1^{n+1})(r_2^n - r_3^n)}$$

$$r_3^{n+1} = r_1^n - \frac{P_n(r_3^n)}{(r_3^n - r_1^{n+1})(r_2^n - r_2^{n+1})}$$