Lecture Notes for Math 5630/6630 **Fall 2024**

Note 5: Iterative Methods

Tags: `math.na` *Date: 09/10/2024*

Disclaimer: *This lecture note is for math 5630/6630 class only.*

In theory, the bracket methods are also iterative. However, one important phenomenon of bracket methods is that the latest iterations may not generate the current iteration. For instance, the false position method will have a stalled endpoint once the iterations are near the root.

In the following, we will introduce some iterative methods, which generate new iterations based on the latest iterations.

1 Fixed Point Iteration

Starting from any number, if you repeatedly press the $\cos(x)$ button on a calculator, the answer will finally stay at a specific number. That is quite related to the topic of "fixed point".

Definition 1.1. *If $x^* = g(x^*)$, then x^* is called an "fixed point" of $g(x)$. This means x^* will not change after applying g.*

> Geometrically, if graphs of $y = g(x)$ and the straight line $y = x$ have an intersection point, that point corresponds to a fixed point.

The following two theorems constitute the Fixed Point Theorem (existence and uniqueness).

Theorem 1.2. *If $g \in C[a, b]$ satisfies that $a \le g(x) \le b$ for all $x \in [a, b]$, then there is a fixed point x^* in the interval $[a, b]$.*

> Let $h(x) := x - g(x)$. It satisfies that $h(a) \ge 0$ and $h(b) \le 0$, then by Intermediate Value Theorem, a root of $h(x)$ (a fixed point of g) exists on $[a, b]$.

Theorem 1.3. *If in addition that g' exists and $|g'(x)| \le \rho$ for a constant $\rho < 1$, then the fixed point x^* is unique.*

If there are two fixed points x^* and y^*, then

$$x^* = g(x^*), \quad y^* = g(y^*)$$

It implies that there exists $\zeta \in [a, b]$ that

$$|x^* - y^*| = |g(x^*) - g(y^*)| = |g'(\zeta)||x^* - y^*| \le \rho |x^* - y^*|$$

Therefore, $|x^* - y^*| = 0$.

Computationally, the fixed point iteration is written as

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

If x_n **converges** to a finite number, we find a fixed point for $g(x)$.

Theorem 1.4. *Suppose $g \in C[a, b]$ and $a \le g(x) \le b$ with $|g'| \le \rho < 1$, then the fixed point iteration converges the unique fixed point.*

Since $|x_{n+1} - x^*| = |g(x_n) - g(x^*)| \le \rho |x_n - x^*| \le \cdots \le \rho^n |x_0 - x^*| \to 0$. The convergence is **linear** and the rate is bounded by ρ. The sequence is called a **contraction** by factor ρ.

Suppose $g \in C^2[a, b]$. Using Taylor expansion, we notice that

$$\begin{aligned} x_{n+1} = g(x_n) &= g(x^*) + g'(x^*)(x_n - x^*) + \frac{g''(\zeta)}{2}(x_n - x^*)^2 \\ &= x^* + g'(x^*)(x_n - x^*) + \frac{g''(\zeta)}{2}(x_n - x^*)^2 \end{aligned}$$

for certain $\zeta \in [x_n, x^*]$. It implies that

$$|x_{n+1} - x^*| = |(x_n - x^*)||g'(x^*) + \frac{g''(\zeta)}{2}(x_n - x^*)|.$$

If x_n is near the fixed point x^*, $\frac{g''(\zeta)}{2}(x_n - x^*) \approx 0$, then we can see

$$|x_{n+1} - x^*| \approx |(x_n - x^*)||g'(x^*)|.$$

It means *locally* the fixed point iteration converges if $|g'(x^*)| < 1$, and diverges if $|g'(x^*)| > 1$.

Example 1.5. *For instance, let $g(x) = \frac{x^2}{4} + \frac{3}{4}$, it has two fixed points $x_1^* = 1$ and $x_2^* = -3$. We can find $g'(x_1^*) = \frac{1}{2}$, $g'(x_2^*) = -\frac{3}{2}$. Then, the fixed point iteration converges near x_1^* but diverges near x_2^*.*

Example 1.6. *For instance, let* $g(x) = \frac{x^2}{4} - 3$*, it has two fixed points* $x_1^* = -2$ *and* $x_2^* = 6$*. We can find* $g'(x_1^*) = -1$*,* $g'(x_2^*) = 3$*. Then, the fixed point iteration diverges near* x_2^**. Near* x_1^**, it is not straightforward to tell whether locally the iteration converges since* $|g'(x_1^*)| = 1$*.*

We notice the sequence $\{x_{2n}\}_{n\geq 1}$ forms a fixed point iteration for the function

$$g(g(x)) = \frac{1}{4}\left(\frac{x^2}{4} - 3\right)^2 - 3 = \frac{x^4}{64} - \frac{3}{8}x^2 - \frac{3}{4}.$$

And we find that $(g(g(x)))' = \frac{1}{16}x^3 - \frac{3}{4}x$ attains a local maximum value 1 at $x_1^* = -2$. In other words, the fixed point iteration with $g(g(x))$ converges near x_1^*. Therefore the original fixed point iteration converges locally at x_1^*. However, the convergence can be sublinear.

Theorem 1.7. *Suppose* x^* *is a fixed point of* $g(x)$ *and* $|g'(x^*)| = 1$*. If* $|\frac{d}{dx}g(g(x^*))| \leq 1$ *for all* $x \in (x^* - \varepsilon, x^* + \varepsilon)$ *and the equal sign holds* ***only*** *at* $x = x^*$*, then locally the fixed point iteration converges to* x^**.*

Remark 1.8. *Let* I *be an interval and* $x^* \in I$ *satisfies* $f(x^*) = 0$*. Suppose the function* $g(x) = x + f(x)h(x)$ *for certain function* $h(x)$ *satisfies* $g(I) \subset I$*. Then* $x^* \in I$ *is also a fixed point of* $g(x) = x + f(x)h(x)$*. The corresponding fixed point iteration is*

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

If x_n *is close to a root* x^**, then*

$$\begin{aligned} x_{n+1} - x^* &= x_n - x^* + f(x_n)h(x_n) \\ &= x_n - x^* + \big(f(x^*)h(x_n^*) + (f'(x^*)h(x_n^*) + f(x_n)h'(x_n^*))(x_n - x^*) + O(|x_n - x^*|^2)\big) \end{aligned}$$

which is

$$x_{n+1} - x^* \approx (x_n - x^*)(1 + (f'(x^*)h(x^*) + f(x^*)h'(x^*))) = (x_n - x^*)(1 + f'(x^*)h(x^*)).$$

For the local convergence, we need the contraction property $|1 + f'(x^*)h(x^*)| < 1$*.*

2 Newton-Raphson Method

Let $f \in C^2$, if the current iteration x_n is close to a root, then the Newton-Raphson method computes the next iteration x_{n+1} by the **fixed point iteration**

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

Geometrically, x_{n+1} is the x-intercept of the tangent line of f at $(x_n, f(x_n))$, see Fig 5.1. Asymptotically, the Newton-Raphson method has a quadratic convergence rate if the root is simple.

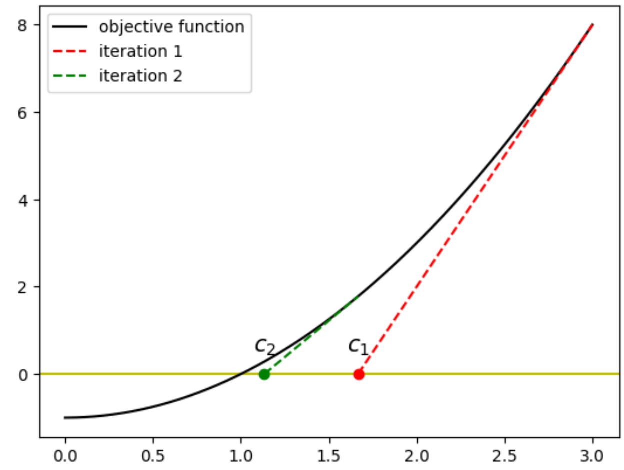

Figure 5.1: Newton's Method

The Newton's method uses $h(x) = -\frac{1}{f'(x)}$ in Remark 1.8, which minimizes the factor $1 + f'(x^*)h(x^*) = 0$ and leads to

$$x_{n+1} - x^* = \mathcal{O}(|x_n - x^*|^2).$$

Theorem 2.1. *Suppose $f \in C^2$ and the root is simple, then the order of convergence for the Newton-Raphson method is 2.*

The Taylor expansion of f at x_n is

$$f(x) = f(x_n) + f'(x_n)(x - x_n) + \frac{f''(\zeta)}{2}(x - x_n)^2$$

where $\zeta \in (x, x_n)$. The root is x^*, then $f(x^*) = 0$, thus

$$0 = f(x_n) + f'(x_n)(x^* - x_n) + \frac{f''(\zeta)}{2}(x^* - x_n)^2.$$

Therefore,

$$-\frac{f''(\zeta)}{f'(x_n)}(x^* - x_n)^2 = x^* - x_n + \frac{f(x_n)}{f'(x_n)} = x^* - x_{n+1}.$$

If we denote the error $\epsilon_n = x_n - x^*$, then $\epsilon_{n+1} = \frac{f''(\zeta)}{f'(x_n)}\epsilon_n^2$, which implies the order of convergence is 2.

3 Secant Method

The secant method computes the next iteration x_{n+1} by

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

which computes the x-intercept of the secant line connecting $(x_{n-1}, f(x_{n-1}))$ and $(x_n, f(x_n))$. See Fig 5.2.

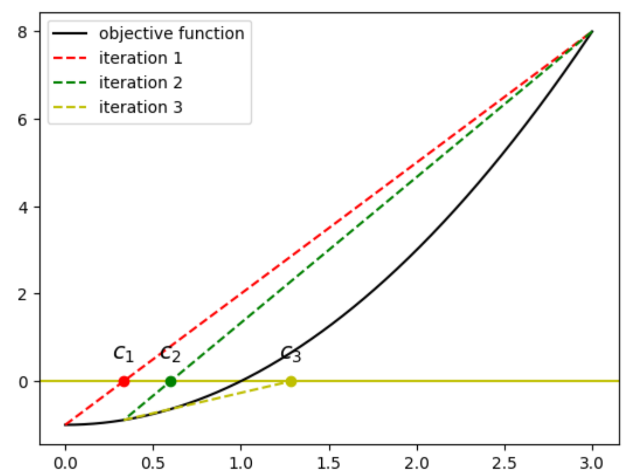

Figure 5.2: Secant Method

This method is similar to the false position method and Newton's method. The differences are:

1. The false position method uses the endpoints of a bracket, while the secant method always uses the latest two iterations to perform the next iteration.
2. Newton's method uses $f'(x_n)$ while the secant method replaces it with its approximation $\frac{f(x_n)-f(x_{n-1})}{x_n - x_{n-1}}$.

Theorem 3.1. *The mean order of convergence of the secant method is* $\frac{1+\sqrt{5}}{2} \approx 1.618$.

Let $\varepsilon_n = x_n - x^*$. Denote $\ell(x)$ the straight line connecting $(x_{n-1}, f(x_{n-1}))$ and $(x_n, f(x_n))$, and let $g(x) := f(x) - \ell(x)$, with $\ell' \equiv f'(\xi)$. There exists $\zeta \in (x_{n-1}, x_n)$ such that

$$\frac{1}{x_n - x_{n-1}} \left[\frac{g(x) - g(z)}{x - z}\Big|_{z=x_n} - \frac{g(x) - g(z)}{x - z}\Big|_{z=x_{n-1}} \right] = \frac{d}{dz}\left(\frac{g(x) - g(z)}{x - z} \right)\Big|_{z=\zeta}.$$

Since $g(x_n) = g(x_{n-1}) = 0$, the left-hand side equals to

$$\frac{1}{x_n - x_{n-1}} \left[\frac{g(x)}{x - x_n} - \frac{g(x)}{x - x_{n-1}} \right] = \frac{g(x)}{(x - x_n)(x - x_{n-1})}$$

The right-hand side equals to

$$\frac{-g'(z)(x - z) + (g(x) - g(z))}{(x - z)^2} = \frac{1}{2} f''(\zeta)$$

by the Taylor expansion $g(x) = g(z) + g'(z)(x - z) + \frac{1}{2} g''(\zeta)(x - z)^2$. Therefore,

$$g(x) = f(x) - \ell(x) = \frac{1}{2} f''(\zeta)(x - x_n)(x - x_{n-1}).$$

Put $x = x^*$, $f(x^*) = 0$ and $\ell(x^*) = f'(\xi)(x^* - x_{n+1})$, we obtain

$$\varepsilon_{n+1} = -\frac{f''(\zeta)}{2f'(\xi)} \varepsilon_n \varepsilon_{n-1}$$

On average, if we have the order p convergence $\varepsilon_{n+1} = \mathcal{O}(\varepsilon_n^p)$, then it must satisfy

$$p^2 = p + 1 \Rightarrow p = \frac{\sqrt{5} + 1}{2}.$$