

MA385 Part 1: Solving nonlinear equations

1.3: The Secant Method

Dr Niall Madden

September 2025

0. Outline

- 1 Motivation
- 2 Order of Convergence
- 3 Side quest: The Mean Value Theorem
- 4 Analysis of the Secant Method
- 5 Exercises

For more details, see Section 1.5 (The secant method) of [Süli and Mayers, *An Introduction to Numerical Analysis*](#)

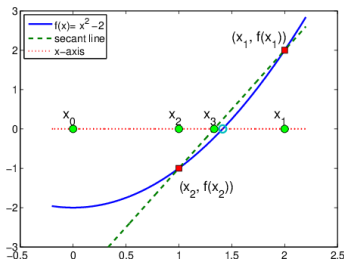
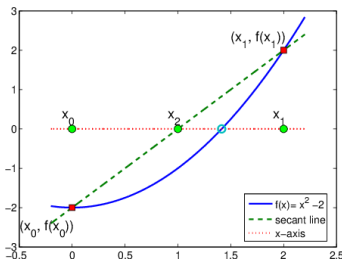
1. Motivation

We'll start with considering, heuristically, how we could improve upon bisection:

1. Motivation

Idea:

- ▶ Choose two points, x_0 and x_1 .
- ▶ Take x_2 to be the zero of the line joining $(x_0, f(x_0))$ to $(x_1, f(x_1))$.
- ▶ Take x_3 to be the zero of the line joining $(x_1, f(x_1))$ to $(x_2, f(x_2))$.
- ▶ Etc.



1. Motivation

The Secant Method

Choose x_0 and x_1 so that there is a solution in $[x_0, x_1]$. Then define

$$x_{k+1} = x_k - f(x_k) \frac{x_k - x_{k-1}}{f(x_k) - f(x_{k-1})}. \quad (1)$$

1. Motivation

Example 1.3.1

Use the Secant Method to solve $x^2 - 2 = 0$ in $[0, 2]$. Results are shown below. We see that, not only does the method appear to converge to the true solution, it seem to do so *much* more efficiently than Bisection. We'll return to why this is later.

	Secant		Bisection	
k	x_k	$ x_k - \tau $	x_k	$ x_k - \tau $
0	0.000000	1.41	0.000000	1.41
1	2.000000	5.86e-01	2.000000	5.86e-01
2	1.000000	4.14e-01	1.000000	4.14e-01
3	1.333333	8.09e-02	1.500000	8.58e-02
4	1.428571	1.44e-02	1.250000	1.64e-01
5	1.413793	4.20e-04	1.375000	3.92e-02
6	1.414211	2.12e-06	1.437500	2.33e-02
7	1.414214	3.16e-10	1.406250	7.96e-03

2. Order of Convergence

To compare different methods, we need the following concept.

Definition 1.3.1 (Linear Convergence)

Suppose that $\tau = \lim_{k \rightarrow \infty} x_k$. We say that the sequence $\{x_k\}_{k=0}^{\infty}$ converges to τ at least linearly if there is a sequence of positive numbers $\{\varepsilon_k\}_{k=0}^{\infty}$, and $\mu \in (0, 1)$, such that

$$\lim_{k \rightarrow \infty} \varepsilon_k = 0, \quad (2a)$$

and

$$|\tau - x_k| \leq \varepsilon_k \quad \text{for } k = 0, 1, 2, \dots \quad (2b)$$

and

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_{k+1}}{\varepsilon_k} = \mu. \quad (2c)$$

2. Order of Convergence

For Example 6, the bisection method converges at least linearly. As we have seen, the Secant Method converge more quickly than bisection. Now we'll give a more precise description of what “*more quickly*” means.

Definition 1.3.2 (Order of Convergence)

Let $\tau = \lim_{k \rightarrow \infty} x_k$. Suppose there exists $\mu > 0$ and a sequence of positive numbers $\{\varepsilon_k\}_{k=0}^{\infty}$ such that (2a) and (2b) both hold. Then we say that the sequence $\{x_k\}_{k=0}^{\infty}$ converges with at least order q if

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_{k+1}}{(\varepsilon_k)^q} = \mu.$$

2. Order of Convergence

$$\lim_{k \rightarrow \infty} \frac{\varepsilon_{k+1}}{(\varepsilon_k)^q} = \mu.$$

Two particular values of q are important to us:

- (i) If $q = 1$, and we have that $0 < \mu < 1$, then the rate is **linear**.
- (ii) If $q = 2$, the rate is **quadratic** for any $\mu > 0$.

3. Side quest: The Mean Value Theorem

The **Mean Value Theorem** (aka “MVT”) is simple, and incredibly useful. The modern version is due to Cauchy (1823), but there is a version called Rolle’s Theorem which applies to polynomials only from 1691, and one for the **sin** function to Parameshvara or Kerala (1380–1460).

Theorem 1.3.1 (Mean Value Theorem (MVT))

If $f : [a, b] \rightarrow R$ is continuous and differentiable, then there is a point $c \in (a, b)$, such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Informally: *there is a point between a and b where the tangent to f is parallel to the line joining $(a, f(a))$ at $(b, f(b))$.*

3. Side quest: The Mean Value Theorem

A visual proof:

3. Side quest: The Mean Value Theorem

From a numerical analysis point of view, the MVT tells us “how different $f(b)$ might be from $f(a)$ ”:

$$f(b) = f(a) + f'(c)(b - a)$$

gives that

$$|f(b) - f(a)| \leq \max_{a \leq x \leq b} |f'(x)|(b - a).$$

That is, in a computation, if we replace $f(a)$ with $f(b)$, the error we introduce depends on $b - a$ and also the magnitude of f' .

4. Analysis of the Secant Method

Theorem 1.3.2

Suppose that f and f' are real-valued functions, continuous and defined in an interval $I = [\tau - h, \tau + h]$ for some $h > 0$. If $f(\tau) = 0$ and $f'(\tau) \neq 0$, then the sequence (1) converges at least linearly to τ .

4. Analysis of the Secant Method

- ▶ We wish to show that $|\tau - x_{k+1}| < |\tau - x_k|$.
- ▶ From the MVT, there is a point $w_k \in [x_{k-1}, x_k]$ such that

$$\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} = f'(w_k). \quad (3)$$

- ▶ Also by the MVT, there is a point $z_k \in [x_k, \tau]$ such that

$$\frac{f(x_k) - f(\tau)}{x_k - \tau} = \frac{f(x_k)}{x_k - \tau} = f'(z_k). \quad (4)$$

Therefore $f(x_k) = (x_k - \tau)f'(z_k)$.

4. Analysis of the Secant Method

► Using (3) and (4), we can show that

$$\tau - x_{k+1} = (\tau - x_k) \left(1 - f'(z_k)/f'(w_k) \right).$$

Therefore

$$\frac{|\tau - x_{k+1}|}{|\tau - x_k|} = \left| 1 - \frac{f'(z_k)}{f'(w_k)} \right|.$$

4. Analysis of the Secant Method

- Suppose that $f'(\tau) > 0$. (If $f'(\tau) < 0$ just tweak the arguments accordingly). Saying that f' is *continuous in the region* $[\tau - h, \tau + h]$ means that, for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$|f'(x) - f'(\tau)| < \varepsilon \text{ for any } x \in [\tau - \delta, \tau + \delta].$$

Take $\varepsilon = f'(\tau)/4$. Then $|f'(x) - f'(\tau)| < f'(\tau)/4$. Thus

$$\frac{3}{4}f'(\tau) \leq f'(x) \leq \frac{5}{4}f'(\tau) \quad \text{for any } x \in [\tau - \delta, \tau + \delta].$$

Then, so long as w_k and z_k are both in $[\tau - \delta, \tau + \delta]$

$$\frac{f'(z_k)}{f'(w_k)} \leq \frac{5}{3}.$$

4. Analysis of the Secant Method

Given enough time and effort we *could* show that the Secant Method converges faster than linearly. In particular, that the order of convergence is

$$q = (1 + \sqrt{5})/2 \approx 1.618.$$

This number arises as the only positive root of $q^2 - q - 1$. It is called the **Golden Mean**, and arises in many areas of Mathematics, including finding an explicit expression for the Fibonacci Sequence:

$$f_0 = 1,$$

$$f_1 = 1,$$

$$f_{k+1} = f_k + f_{k-1} \text{ for } k = 2, 3, \dots$$

That gives, $f_0 = 1$, $f_1 = 1$, $f_2 = 2$, $f_3 = 3$, $f_4 = 5$, $f_5 = 8$, $f_6 = 13$,
...

4. Analysis of the Secant Method

The connection here is that it turns out that $\varepsilon_{k+1} \leq C\varepsilon_k\varepsilon_{k-1}$. Repeatedly using this we get:

- ▶ Let $r = |x_1 - x_0|$ so that $\varepsilon_0 \leq r$ and $\varepsilon_1 \leq r$,
- ▶ Then $\varepsilon_2 \leq C\varepsilon_1\varepsilon_0 \leq Cr^2$
- ▶ Then $\varepsilon_3 \leq C\varepsilon_2\varepsilon_1 \leq C(Cr^2)r = C^2r^3$.
- ▶ Then $\varepsilon_4 \leq C\varepsilon_3\varepsilon_2 \leq C(C^2r^3)(Cr^2) = C^4r^5$.
- ▶ Then $\varepsilon_5 \leq C\varepsilon_4\varepsilon_3 \leq C(C^4r^5)(C^2r^3) = C^7r^8$.
- ▶ And in general, $\varepsilon_k = C^{f_k-1}r^{f_k}$.

5. Exercises

Exercise 1.3.1

Suppose we define the Secant Method as follows.

Choose any two points x_0 and x_1 .

For $k = 1, 2, \dots$, set x_{k+1} to be the point where the line through $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$ that intersects the x -axis.

Show how to derive the formula for the secant method.

5. Exercises

Exercise 1.3.2

Another interpretation of the Secant Method, is that it computes x_{k+1} as a weighted average:

$$x_{k+1} = (1 - \sigma_k)x_k + \sigma_k x_{k-1},$$

where σ_k chosen to obtain fast convergence to the true solution. (This is something called a “relaxation method”).

Derive a formula for σ_k in terms of $f(x_{k-1})$ and $f(x_k)$. Can you give a physical interpretation to σ_k and $1 - \sigma_k$ that might justify the claim that the method should be more efficient than bisection?

5. Exercises

Exercise 1.3.3

- (i) Is it possible to construct a problem for which the bisection method will work, but the secant method will fail? If so, give an example.
- (ii) Is it possible to construct a problem for which the secant method will work, but bisection will fail? If so, give an example.