Math 701: Secant Method

The secant method approximates solutions to $f(x) = 0$ using an iterative scheme similar to Newton's method in which the derivative has been replace by

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

This results in the two-term recurrence

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

which needs a base of two different approximations x_0 and x_1 of the solution to get started.

Let p be the exact solution such that $f(p) = 0$ and suppose $f'(p) \neq 0$. Before proving that x_0 and x_1 sufficiently close to p implies $x_n \to p$ with order $\alpha = (1 + \sqrt{5})/2$, we first derive this rate of convergence heuristically.

Intuitive Derivation of the Rate of Convergence

Define $e_n = x_n - p$ and assume $x_n \to p$ as $n \to \infty$ and further that $|e_n| \sim M|e_{n-1}|^\alpha$ for some constants $M > 0$ and $\alpha \geq 1$.

Let $\epsilon > 0$ be arbitrary. By definition

$$\begin{aligned}
e_{n+1} = x_{n+1} - p = x_n - \frac{f(x_n)(x_n - x_{n-1})}{f(x_n) - f(x_{n-1})} - p = e_n - \frac{f(x_n)(e_n - e_{n-1})}{f(x_n) - f(x_{n-1})} \\
= e_n - \frac{f(x_n)(e_n - e_{n-1})}{f(x_n) - f(x_{n-1})} = \frac{e_n\big(f(x_n) - f(x_{n-1})\big) - f(x_n)(e_n - e_{n-1})}{f(x_n) - f(x_{n-1})} \\
= \frac{e_n\big(f(x_n) - f(x_{n-1})\big) - f(x_n)(e_n - e_{n-1})}{f(x_n) - f(x_{n-1})} = \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})} \\
= \left(\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}}\right)\left(\frac{e_n e_{n-1}}{f(x_n) - f(x_{n-1})}\right) \\
= \left\{\frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}}\right\}\left(\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})}\right)e_n e_{n-1}
\end{aligned}$$

Claim that

$$\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \approx \frac{1}{f'(p)} \qquad \text{and} \qquad \frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}} \approx \frac{f''(p)}{2}.$$

As this is a heuristic derivation of α there is no need to prove the above claims rigorously, but only to justify them from an intuitive point of view.

For the first part of the claim note the mean value theorem implies there is ξ_n between x_n and x_{n-1} such that

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

It follows that $f'(\xi_n) \to f'(p)$ since $\xi_n \to p$ as $n \to \infty$. Therefore

$$\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \to \frac{1}{f'(p)} \qquad \text{as} \qquad n \to \infty.$$

Consequently, if n is large enough then

$$\frac{x_n - x_{n-1}}{f(x_n) - f(x_{n-1})} \approx \frac{1}{f'(p)}$$

and the first part of the claim has been justified.

An intuitive justification of the second part of the claim is similar, but slightly more involved. By Taylor's theorem there is η_n between x_n and p such that

$$f(x_n) = f(p) + (x_n - p)f'(p) + \frac{1}{2}(x_n - p)^2 f''(p) + \frac{1}{3!}(x_n - p)^3 f'''(\eta_n).$$

Therefore

$$f(x_n) \approx e_n f'(p) + \frac{1}{2} e_n^2 f''(p).$$

This suggests that

$$\frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{x_n - x_{n-1}} \approx \frac{f'(p) - f'(p)}{x_n - x_{n-1}} + \frac{1}{2}\frac{e_n f''(p) - e_{n-1} f''(p)}{x_n - x_{n-1}} = \frac{f''(p)}{2}.$$

We are now ready to infer a plausible value for α the order of convergence of the secant method. Combine the results of the claim with the expression for e_{n-1} to obtain

$$|e_{n+1}| \approx C|e_n||e_{n-1}| \qquad \text{where} \qquad C = \left|\frac{f''(p)}{2f'(p)}\right|.$$

Substituting the relation $|e_n| \sim M|e_{n-1}|^\alpha$ yields

$$M|e_n|^\alpha \approx M^{1+\alpha}|e_{n-1}|^{\alpha^2} \approx CM|e_{n-1}|^\alpha |e_{n-1}|.$$

Solving for M and α from the relations

$$M^{1+\alpha} = CM \qquad \text{and} \qquad \alpha^2 = \alpha + 1$$

obtains

$$M = C^{1/\alpha} = \left|\frac{f''(p)}{2f'(p)}\right|^{1/\alpha} \qquad \text{and} \qquad \alpha = \frac{1+\sqrt{5}}{2}.$$

This finishes our heuristic derivation of the rate of convergence of the secant method.

Rigorous Proof of the Rate of Convergence

We assume f is twice continuously differentiable and that p is such that $f(p) = 0$ and $f'(p) \neq 0$. The secant method is

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

First claim if x_0 and x_1 with $x_0 \neq x_1$ are chosen close enough to p that $x_n \to p$ as $n \to \infty$. That is, the secant method converges. Let

$$\kappa(\alpha, \beta) = 1 - \frac{f'(\alpha)}{f'(\beta)}.$$

Since $f'(p) \neq 0$ it follows that the limit supremum of $|\kappa(\alpha, \beta)|$ is zero as $\alpha \to p$ and $\beta \to p$. Therefore, there is $\delta > 0$ so small such that

$$|\kappa(\alpha, \beta)| \leq \gamma < 1 \quad \text{whenever} \quad |\alpha - p| < \delta \quad \text{and} \quad |\beta - p| < \delta.$$

Choose x_0 and x_1 such that $|x_0 - p| < \delta$ and $|x_1 - p| < \delta$. By the mean-value theorem

$$\frac{f(x_n) - f(p)}{x_n - p} = f'(a_n) \qquad \text{and} \qquad \frac{f(x_n) - f(x_{n-1})}{x_n - x_{n-1}} = f'(b_n)$$

for some a_n between x_n and p and for some b_n between x_n and x_{n-1}. For induction assume $|x_n - p| < \delta$ and $|x_{n-1} - p| < \delta$, in which case $|a_n - p| < \delta$ and $|b_n - p| < \delta$. Denoting $e_n = x_n - p$ we obtain

$$\begin{aligned} e_{n+1} &= e_n - f(x_n)\frac{e_n - e_{n-1}}{f(x_n) - f(x_{n-1})} \\ &= e_n - \frac{(f(x_n) - f(p))(e_n - e_{n-1})}{f(x_n) - f(x_{n-1})} \\ &= e_n - \frac{f'(a_n)e_n(e_n - e_{n-1})}{f'(b_n)(e_n - e_{n-1})} = \Big(1 - \frac{f'(a_n)}{f'(b_n)}\Big)e_n. \end{aligned}$$

Therefore, $|e_{n+1}| \leq \gamma|e_n|$ and by induction $|e_{n+1}| \leq \gamma^n|e_1|$. Since $\gamma < 1$, it follows that $x_n \to p$ as $n \to \infty$ and the secant method converges.

Claim there exists C such that $|e_{n+1}|/|e_n e_{n-1}| \to C$ as $n \to \infty$. First note that

$$\begin{aligned} e_{n+1} &= e_n - f(x_n)\frac{e_n - e_{n-1}}{f(x_n) - f(x_{n-1})} \\ &= e_n - \frac{f(x_n)e_n - f(x_{n-1})e_n + f(x_{n-1})e_n - f(x_n)e_{n-1}}{f(x_n) - f(x_{n-1})} \\ &= \frac{f(x_n)e_{n-1} - f(x_{n-1})e_n}{f(x_n) - f(x_{n-1})} = \frac{f(x_n)/e_n - f(x_{n-1})/e_{n-1}}{f'(b_n)(x_n - x_{n-1})}e_n e_{n-1}. \end{aligned}$$

Define

$$A = \max\{\, |f''(x)| : |x - p| \le \delta \,\}.$$

Since f'' is continuous, it follows that $A < \infty$. By Taylor's theorem we have

$$f(x_n) = f(p) + f'(p)e_n + \int_p^{x_n} f''(t)(x_n - t)dt$$

Now

$$\begin{aligned}
\frac{f(x_n)}{e_n} - \frac{f(x_{n-1})}{e_{n-1}} &= \frac{1}{e_n}\int_p^{x_n} f''(t)(x_n - t)dt - \frac{1}{e_{n-1}}\int_p^{x_{n-1}} f''(t)(x_{n-1} - t)dt \\
&= J_1 + J_2
\end{aligned}$$

where

$$J_1 = \Big(\frac{1}{e_n} - \frac{1}{e_{n-1}}\Big)\int_p^{x_n} f''(t)(x_n - t)dt$$

and

$$\begin{aligned}
J_2 &= \frac{1}{e_{n-1}}\Big(\int_p^{x_n} f''(t)(x_n - t)dt - \int_p^{x_{n-1}} f''(t)(x_{n-1} - t)dt\Big) \\
&= J_3 + J_4
\end{aligned}$$

Here

$$J_3 = \frac{x_n - x_{n-1}}{e_{n-1}}\int_p^{x_n} f''(t)dt \qquad \text{and} \qquad J_4 = \frac{1}{e_{n-1}}\int_{x_{n-1}}^{x_n} f''(t)(x_{n-1} - t)dt.$$

Estimate

$$\begin{aligned}
|J_1| &\le A\Big|\frac{1}{e_n} - \frac{1}{e_{n-1}}\Big|\Big|\int_p^{x_n}(x_n - t)dt\Big| \\
&\le A\frac{|x_n - x_{n-1}|}{|e_n e_{n-1}|}\frac{|e_n|^2}{2} \le A\Big|\frac{e_n}{e_{n-1}}\Big|\frac{|x_n - x_{n-1}|}{2}
\end{aligned}$$

Therefore

$$\frac{|J_1|}{|x_n - x_{n-1}|} \le \frac{A}{2}\Big|\frac{e_n}{e_{n-1}}\Big| = \frac{A}{2}\Big|1 - \frac{f'(a_n)}{f'(b_n)}\Big| \to 0 \qquad \text{as} \qquad n \to \infty.$$

Also

$$\frac{|J_3|}{|x_n - x_{n-1}|} \le A\Big|\frac{e_n}{e_{n-1}}\Big| \to 0 \qquad \text{as} \qquad n \to \infty.$$

The estimate of J_4 will be done more carefully. Consider two cases: If $x_{n-1} < x_n$ then the mean-value theorem for integrals implies

$$\frac{2}{(x_n - x_{n-1})^2} \int_{x_{n-1}}^{x_n} f''(t)(t - x_{n-1})dt = f''(\xi_n) \qquad \text{where} \qquad \xi_n \in [x_{n-1}, x_n].$$

If $x_n < x_{n-1}$ then

$$\frac{2}{(x_{n-1} - x_n)^2} \int_{x_n}^{x_{n-1}} f''(t)(x_{n-1} - t)dt = f''(\xi_n) \qquad \text{where} \qquad \xi_n \in [x_n, x_{n-1}].$$

In either case it holds that

$$\frac{J_4}{x_n - x_{n-1}} = f''(\xi_n)\frac{x_n - x_{n-1}}{2e_{n-1}} = \frac{f''(\xi_n)}{2}\Big(\frac{e_n}{e_{n-1}} + 1\Big) \to \frac{f''(p)}{2}$$

as $n \to \infty$. It follows that

$$J_2 \to \frac{f''(p)}{2} \qquad \text{as} \qquad n \to \infty.$$

Consequently

$$\left|\frac{e_{n+1}}{e_n e_{n-1}}\right| \to C \qquad \text{where} \qquad C = \left|\frac{f''(p)}{2f'(p)}\right| \qquad \text{as} \qquad n \to \infty.$$

Claim that the secant method converges with order α. Note that

$$\frac{1}{\alpha - 1} = \alpha, \qquad \alpha^2 - 1 = \alpha, \qquad \text{and} \qquad \frac{1}{\alpha} + \frac{1}{\alpha^2} = 1.$$

Define $K_n = |e_{n+1}|/|e_n e_{n-1}|$ and $M_n = |e_{n+1}|/|e_n|^\alpha$. Then

$$M_n^\alpha M_{n-1} = \Big(\frac{|e_{n+1}|^\alpha}{|e_n|^{\alpha^2}}\Big)\Big(\frac{|e_n|}{|e_{n-1}|^\alpha}\Big) = \Big(\frac{|e_{n+1}|}{|e_n e_{n-1}|}\Big)^\alpha = K_n^\alpha.$$

It follows that

$$M_n = \frac{K_n}{M_{n-1}^{1/\alpha}} \qquad \text{and similarly} \qquad M_{n+1} = \frac{K_{n+1}}{M_n^{1/\alpha}}.$$

Combining the above two inequalities implies

$$M_{n+1} = \frac{K_{n+1}}{K_n^{1/\alpha}} M_{n-1}^{1/\alpha^2}.$$

Since $K_n \to C$ as $n \to \infty$ then

$$\frac{K_{n+1}}{K_n^{1/\alpha}} \to C^{1-1/\alpha} = C^{2-\alpha} \qquad \text{as} \qquad n \to \infty.$$

Since $2-\alpha > 0$, the above limit makes sense even when $f''(p) = 0$. Let $L = 1 + C^{2-\alpha}$. By the definition of limit, there exists N large enough such that

$$M_{n+1} \le L M_{n-1}^{1/\alpha^2} \qquad \text{for all} \qquad n \ge N.$$

The above inequality and the fact that $1/\alpha^2 < 1$ implies that the sequence M_n is bounded. In particular, suppose $M_{2n-1} > L^\alpha$ where $2n \ge N$, then

$$M_{2n+1} \le L M_{2n-1}^{1/\alpha^2} < M_{2n-1}^{1/\alpha + 1/\alpha^2} = M_{2n-1}.$$

Therefore $M_{2(n+k)+1} \le M_{2n-1}$ for all $k \in \mathbf{N}$. Similarly if $M_{2n} > L^\alpha$ for some $2n \ge N$, then $M_{2(n+k)} \le M_{2n}$ for all $k \in \mathbf{N}$. Having consider both even and odd terms, we conclude in general that M_k is bounded. Consequently there exists M large enough such that $M_n \le M$ for every $n \in \mathbf{N}$. Thus

$$|e_{n+1}| \le M |e_n|^\alpha \qquad \text{for every} \qquad n \in \mathbf{N}$$

and so the secant method converges with order at least α.

The following references were consulted when preparing the above proof:

[1] Burden, Fairs and Burden, *Numerical Analysis, Tenth Edition*, hint given in for Problem 14 in Section 2.4.

[2] Dahlquist and Björck, *Numerical Methods in Scientific Computer, Volume I*, proof of Theorem 6.2.1.