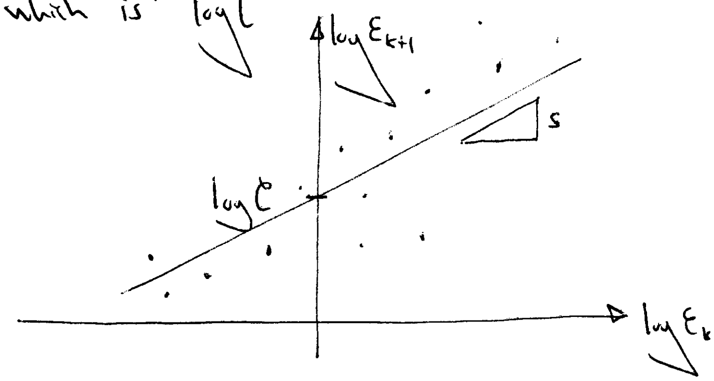
Rate of Convergence of a Sequence.

A sequence x_k is said to converge to α with order s iff $\frac{\varepsilon_{k+1}}{\varepsilon_k^s} \to \mathcal{C}$ where $\mathcal{C}$ is a nonzero constant, (and of course $\varepsilon_k = x_k - \alpha$). Most theoretical problems in Math128a involve showing that $s=1$ (called linear convergence) or that $s=2$ (called quadratic convergence). Note that

$$\log \varepsilon_{k+1} = s \cdot \log \varepsilon_k + \log \mathcal{C} \qquad (1)$$

and so a practical method of finding s and $\mathcal{C}$ is to run code on a computer, generating error values ε_k, and to then plot $\log \varepsilon_{k+1}$ versus $\log \varepsilon_k$. The data points should fall roughly on a line, the slope of which is s, and the y-intercept of which is $\log \mathcal{C}$

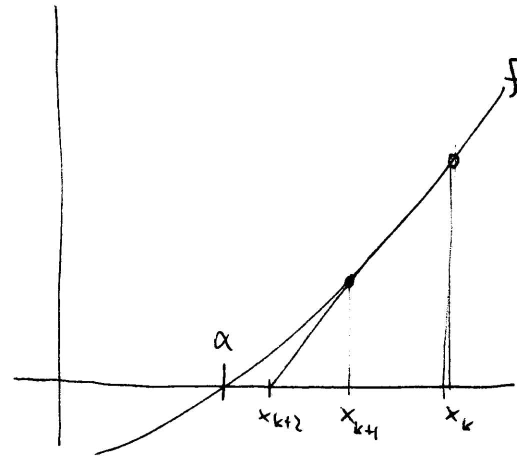


It is worth noting that in practice, a sequence undergoing quadratic convergence gets to within numerical precision ($\sim 10^{-16}$) of its limit after only a handful of iterations.

Although for theoretical problems s is normally 1 or 2, we can show that for the secant method, $s = \frac{1+\sqrt{5}}{2} \approx 1.61803\ldots$

$\mathcal{P}$ turn

Recall that the Secant Method generates a new iterate x_{k+2} given two previous iterates x_k and x_{k+1} as illustrated below



Observe the similar triangles:

$$\frac{f(x_k)}{x_k - x_{k+2}} = \frac{f(x_{k+1})}{x_{k+1} - x_{k+2}} \qquad (2)$$

Giving

$$x_{k+2}\Big(f(x_k) - f(x_{k+1})\Big) = x_{k+1} f(x_k) - x_k \cdot f(x_{k+1}) \qquad (3)$$

which we can use to find x_{k+2}.

The geometry of the method and thus the rate of convergence is unaffected by our choice of coordinates; we make the following computations easier by setting the origin at the root α. Taylor series expanding about the (new) origin we obtain

$$f(x_k) = f'(0)\cdot x_k + f''(0)\cdot \frac{x_k^2}{2} + O(x_k^3)$$

$$f(x_{k+1}) = f'(0)\cdot x_{k+1} + f''(0)\cdot \frac{x_{k+1}^2}{2} + O(x_{k+1}^3)$$

Note $0 < |x_{k+1}| < |x_k|$.

Thus

$$x_{k+1}\cdot f(x_k) - x_k \cdot f(x_{k+1}) = \frac{f''(0)}{2}\cdot\left(x_{k+1}x_k^2 - x_k x_{k+1}^2\right) + O(x_k^4)$$

$$= \frac{f''(0)}{2} x_{k+1} x_k \cdot \left(x_k - x_{k+1}\right) + O(x_k^4)$$

Also,

$$f(x_k) - f(x_{k+1}) = f'(0)\cdot(x_k - x_{k+1}) + O(x_k^2)$$

Thus (3) can be written as

$$x_{k+2}\cdot \cancel{f'(0)}\cdot\cancel{(x_k - x_{k+1})}\cdot\Big(1 + O(x_k)\Big) = \cancel{f'(0)}\cdot\cancel{(x_k - x_{k+1})}\left(\frac{1}{2}\frac{f''(0)}{f'(0)} x_{k+1}x_k + O(x_k^3)\right)$$

Recall from a problem set that $\frac{1}{1+x} = 1 - x + x^2 \cdots$ and so

$$x_{k+2} = \frac{1}{2}\frac{f''(0)}{f'(0)} x_{k+1} x_k + O(x_k^3) \qquad (4)$$

when the x's are close enough to the root (0) we have

$$x_{k+2} = M x_{k+1} x_k$$

where $M = \frac{1}{2}\frac{f''(0)}{f'(0)}$ is assumed nonzero (or we'd have to look at even higher order terms!).

Define $y_k = M \cdot x_k$ and note that

$$x_{k+2} = M \cdot x_{k+1} \cdot x_k$$

$$\frac{y_{k+2}}{\cancel{M}} = \cancel{M} \cdot \frac{y_{k+1}}{\cancel{M}} \cdot \frac{y_k}{\cancel{M}}$$

And so $y_{k+2} = y_{k+1} \cdot y_k \longrightarrow z_{k+2} = z_{k+1} + z_k$ where $z_k = \log y_k$.

Notice that

$$\underbrace{\begin{bmatrix} z_{k+1} \\ z_{k+2} \end{bmatrix}}_{\underline{v}_{k+1}} = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix}}_{\underline{A}} \underbrace{\begin{bmatrix} z_k \\ z_{k+1} \end{bmatrix}}_{\underline{v}_k}$$

repeated matrix multiplication.

which gives the nice expression $\underline{v}_k = \underline{A}^k \underline{v}_0$ for any $\underline{v}_k$. A short computation reveals that $\underline{A}$ has eigenvalues $\lambda_1 = \frac{1+\sqrt{5}}{2}$ and $\lambda_2 = \frac{1-\sqrt{5}}{2}$, and corresponding eigenvectors $\underline{u}_1 = \begin{bmatrix} \lambda_1 - 1 \\ 1 \end{bmatrix}$ and $\underline{u}_2 = \begin{bmatrix} \lambda_2 - 1 \\ 1 \end{bmatrix}$, which obviously span $\mathbb{R}^2$. Writing $\underline{v}_0 = \alpha \underline{u}_1 + \beta \underline{u}_2$ we obtain $\underline{v}_k = \underline{A}^k \underline{v}_0 = \alpha \underline{A}^k \underline{u}_1 + \beta \underline{A}^k \underline{u}_2 = \alpha \lambda_1^k \underline{u}_1 + \beta \lambda_2^k \underline{u}_2$. Note that $|\lambda_2| < 1$ and so $\lambda_2^k \longrightarrow 0$ as $k \to \infty$, thus for large k, $\underline{v}_k \approx \alpha \lambda_1^k \underline{u}_1$. Taking the 2^{nd} component we obtain $z_{k+2} \approx \alpha \lambda_1^k$. A sufficient condition for $z_k \to -\infty$ is $\alpha < 0$; For convenience we take $\alpha = -1$, (it is easy to see that α doesn't affect the following argument).

$$z_{k+2} = \log \frac{x_{k+2}}{M} \longrightarrow x_{k+2} = M e^{-\lambda_1^k}$$

turn

Remember that our original task was to find s so that $\frac{\varepsilon_{k+1}}{\varepsilon_k^s} = R \longrightarrow$ a non zero constant.

$$R = \frac{\varepsilon_{k+1}}{\varepsilon_k^s} = \frac{x_{k+1}}{x_k^s} = \frac{Me^{-\lambda_1^{k+1}}}{Me^{-s\cdot\lambda_1^k}} = e^{s\lambda_1^k - \lambda_1^{k+1}} = e^{\lambda_1^k\cdot(s-\lambda_1)}$$

If $s \neq \lambda_1$ then clearly R goes to either 0 or infinity. It must be that

$$s = \lambda_1 = \frac{1+\sqrt{5}}{2}$$