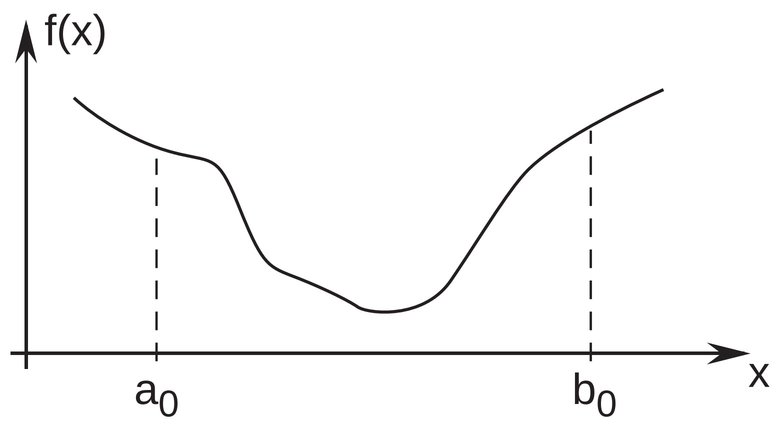
One-dimensional search methods

- Given $f : \mathbb{R} \to \mathbb{R}$. Want to find minimizer of f.
- Use iterative algorithm.
- Special case of the multivariable problem.
- Applicable to line search used in general multivariable algorithms.

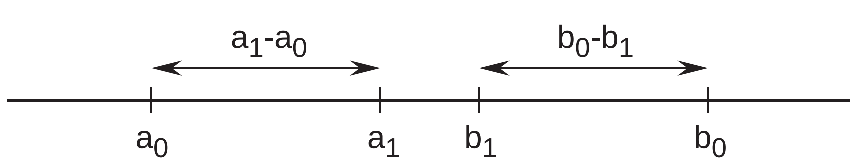
- Iterative algorithm: next point $x^{(k+1)}$ depends on $x^{(k)}$ and f.
- The algorithm may use only f, and perhaps f' and even f''.
- We study three algorithms:
  - Golden section method [uses only f]
  - Secant method [uses only f']
  - Newton's method [uses f' and f'']

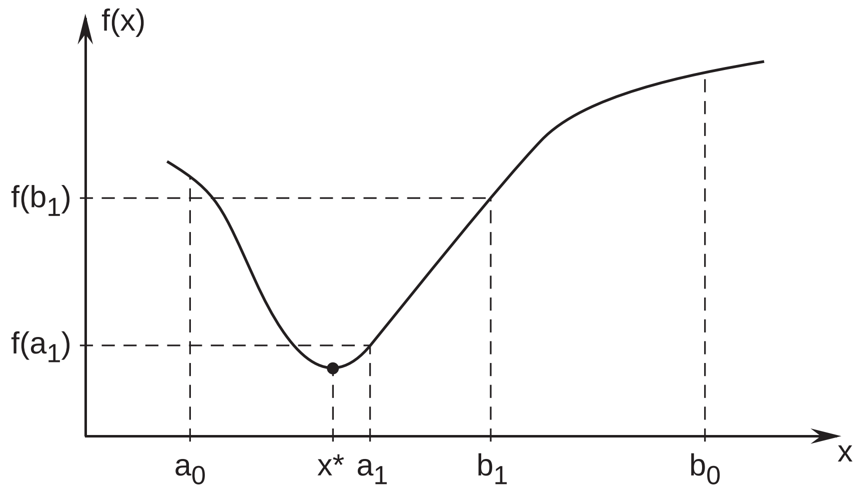
Golden section search (§7.1)

- Given $f : [a_0, b_0] \to \mathbb{R}$, unimodal.

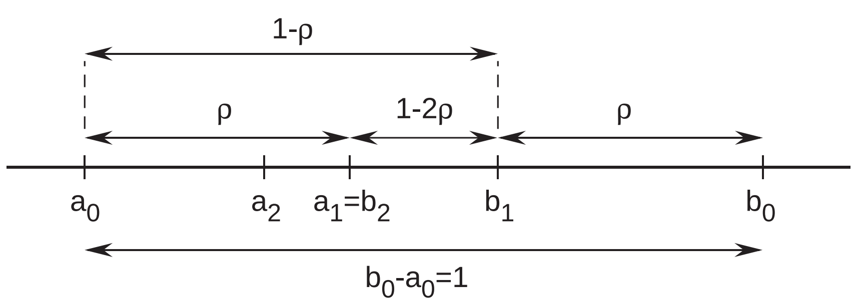


- Evaluate f at two intermediate points a_1, b_1, with $a_1 - a_0 = \rho(b_0 - a_0)$.





- Suppose $f(a_1) < f(b_1)$.
- We now have a reduced uncertainty interval $[a_0, b_1]$.
- Repeat procedure.
- Problem: function evaluations are "expensive"!
- To reduce the number of function evaluations, want to choose position of intermediate points in such a way that a previously evaluated point is one of the intermediate points.



- Need to have $\rho(1-\rho) = 1 - 2\rho$.
- Solving, we get $\rho = (3-\sqrt{5})/2 \approx 0.382$.
- At each step the reduction in the range of the uncertainty interval is a factor of $1-\rho$. Hence, after N steps, we have $(1-\rho)^N \approx 0.618^N$.
- If we desire a final uncertainty range of ≤ 0.1, say, we need N steps, where

$$(\text{Initial range}) \times 0.618^N \leq 0.1.$$

- See Example 7.1.
- Read about Fibonacci method.

Newton's method (§7.3)

- Start with $x^{(0)}$. Approximate f by a quadratic:

$$q_0(x) = f(x^{(0)}) + f'(x^{(0)})(x - x^{(0)}) + \frac{1}{2}f''(x^{(0)})(x - x^{(0)})^2.$$

  We match the first and second derivatives.

- Use the minimizer of q_0 as our next iterate, $x^{(1)}$.

- By FONC, we need

$$q_0'(x^{(1)}) = f'(x^{(0)}) + f''(x^{(0)})(x^{(1)} - x^{(0)}) = 0.$$

- We get

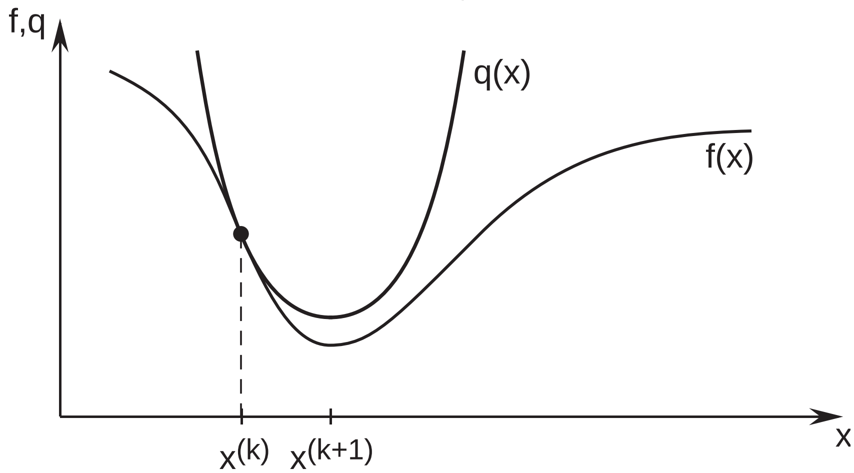
$$x^{(1)} = x^{(0)} - \frac{f'(x^{(0)})}{f''(x^{(0)})}.$$

- General Newton's algorithm:

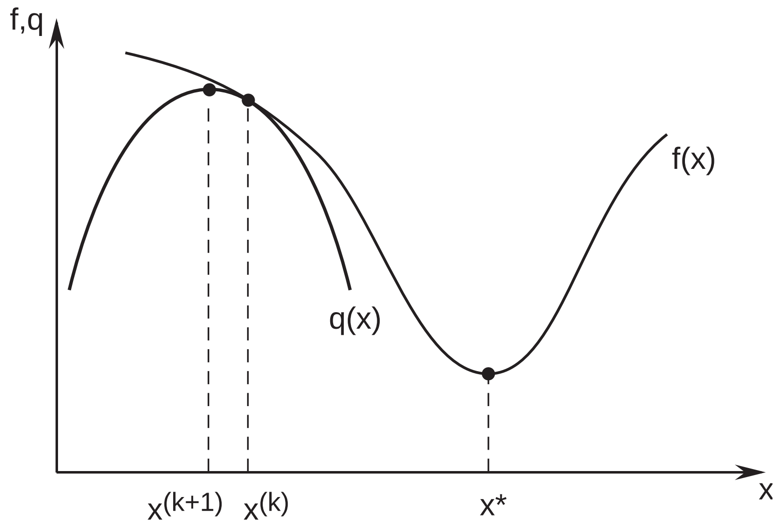
$$x^{(k+1)} = x^{(k)} - \frac{f'(x^{(k)})}{f''(x^{(k)})}.$$

- See Example 7.3.

Newton's method works well if we start close enough.



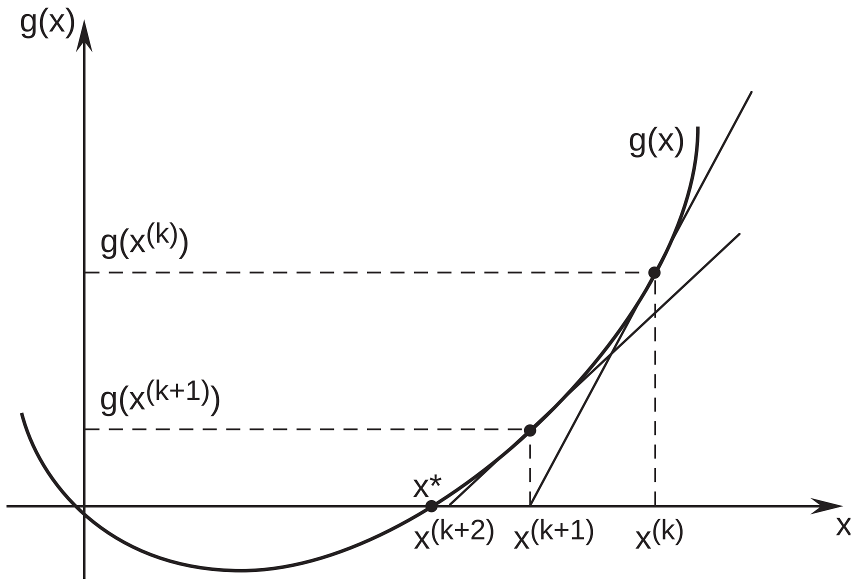
Newton's method may fail if we don't start close enough.



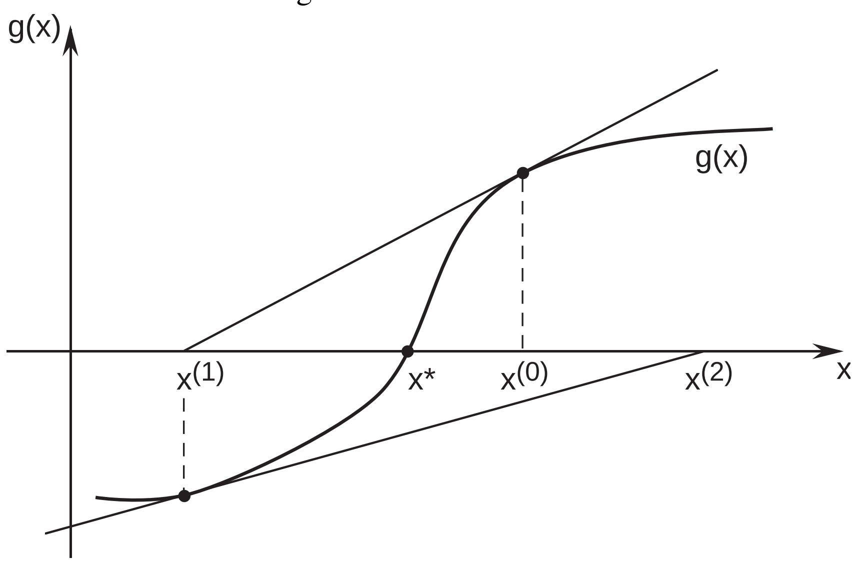
Newton's method of tangents

- Write $g(x) = f'(x)$.
- Newton's method becomes
$$x^{(k+1)} = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}.$$
Drives $g(x^{(k)})$ to zero.
- Can apply the same algorithm to find the zero of a function $g : \mathbb{R} \to \mathbb{R}$.
- Name: *Newton's method of tangents*.
- See Example 7.4.
- Need to start "close enough" to minimizer.

Newton's method of tangents ...



... may fail if we don't start close enough.



Secant method (§7.4)

- Newton's method requires f''.
- If f'' not available, may approximate it using finite differences.
- Replace $f''(x^{(k)})$ in Newton's method by

$$\frac{f'(x^{(k)}) - f'(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}.$$

- This is like fitting a quadratic by matching slopes at two points.

- For a quadratic, the secant approximation is exact:

$$f''(x^{(k)}) = \frac{f'(x^{(k)}) - f'(x^{(k-1)})}{x^{(k)} - x^{(k-1)}}.$$

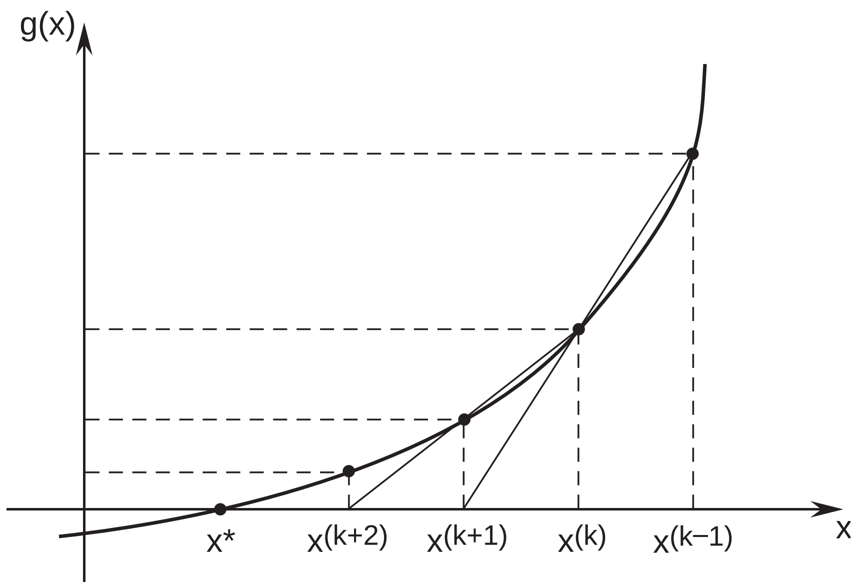
- Secant algorithm:

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{f'(x^{(k)}) - f'(x^{(k-1)})} f'(x^{(k)}).$$

- Requires two initial points: $x^{(-1)}$ and $x^{(0)}$.
- Can use for zero finding.
- Secant algorithm to find zero of g:

$$x^{(k+1)} = x^{(k)} - \frac{x^{(k)} - x^{(k-1)}}{g(x^{(k)}) - g(x^{(k-1)})} g(x^{(k)}).$$

- See Example 7.5.



Remarks on line search (§7.5)

- Recall that one dimensional methods are used for line search:

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$$

- Choose α_k to minimize

$$\phi_k(\alpha) = f(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}).$$

- Derivative of ϕ_k is
$$\phi_k'(\alpha) = \boldsymbol{d}^{(k)T}\nabla f(\boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}^{(k)}).$$

- Practical line search algorithms are usually complicated.

- Often not necessary to do accurate line searches.

- To still guarantee convergence of overall algorithm, need to use appropriate termination criteria for line search.

Next to be discussed ...

- Various algorithms for multivariable optimization.

- All have the basic form:
$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{d}^{(k)}$$

- Desired properties:
  - Descent property
  - Convergence

- Direction $\boldsymbol{d}^{(k)}$ usually depends on $\nabla f(\boldsymbol{x}^{(k)})$, and possibly $\boldsymbol{F}(\boldsymbol{x}^{(k)})$ or $\boldsymbol{d}^{(k-1)}$.

- Background: Need to recall limits/convergence.

