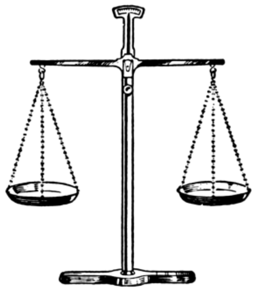


Univariate Non Linear Programming: Newton method

Univariate non linear optimization

Algorithm can be used for a ***numerical*** solution of the problem

These optimization algorithms generate a sequence of points converging to the optimal solution

There are two types of optimization algorithms:

- Dicotonomous (find the root of the equation of the derivative equal to zero and at each iteration reduce the search interval).
- We have seen the bisection method, another algorithm of this type is the secant method (explained in the exercise lecture).

- Approximation: these algorithms use local approximations of the function to be optimised

Termination criteria

- When to stop:
  - The solution is accurate with a given level of tolerance
  - Very small improvement from one iteration to the next:

$$|x_k - x_{k-1}| \le \varepsilon_x, \quad |f(x_k) - f(x_{k-1})| \le \varepsilon_f$$

  - The maximum number of iterations has been reached
  - The solution diverges
  - The solutions are in a loop

Newton's method

Fit a quadratic approximation to $f(x)$ using both gradient and curvature information at x.

- Use Taylor approximation:

$$f(x+h) = f(x) + f'(x)*h + \tfrac{1}{2} f''(x)\, h^2$$

- Root finding of $f\,'$

$$f'(x+h) = f'(x) + f''(x)h + o(h^2)$$

$$\Rightarrow h = -\frac{f'(x)}{f''(x)}$$

Newton's method

Fit a quadratic approximation to $f(x)$ using both gradient and curvature information at x.

- Use Taylor approximation:

$$f(x+h) = f(x) + f'(x)*h + \tfrac{1}{2} f''(x)\, h^2$$

- Root finding of f'

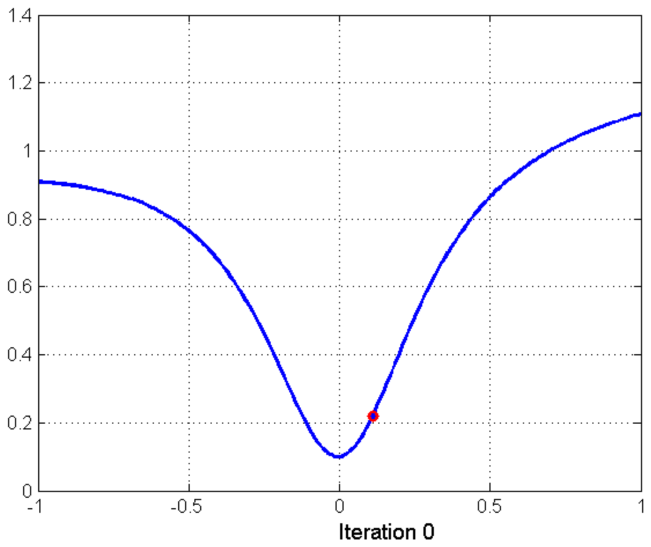
$$f'(x+h) = f'(x) + f''(x)h + \cancel{o(h^2)}$$

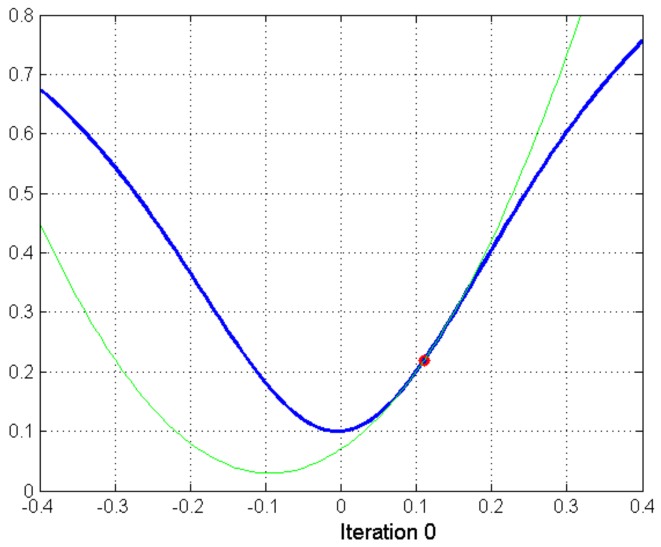
Linear approximation

$$\Rightarrow h = -\frac{f'(x)}{f''(x)}$$

- New guess: $x_{k+1} = x_k + h_k = x_k - \dfrac{f'(x_k)}{f''(x_k)}$

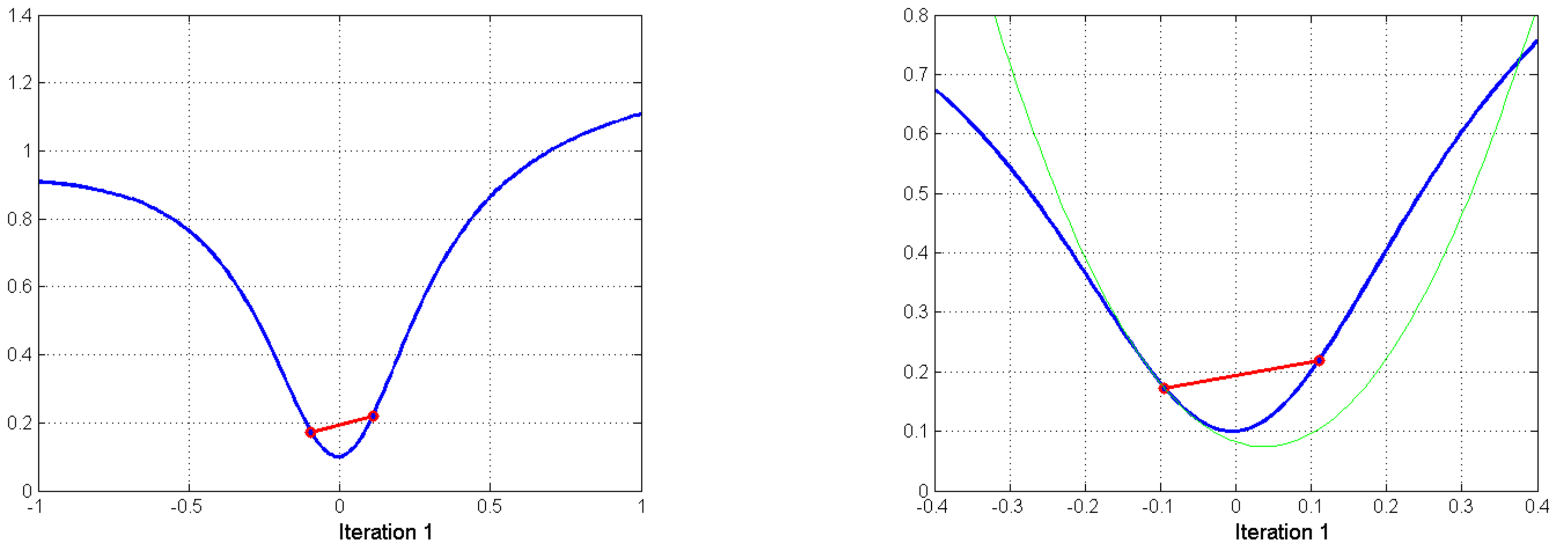
Newton method





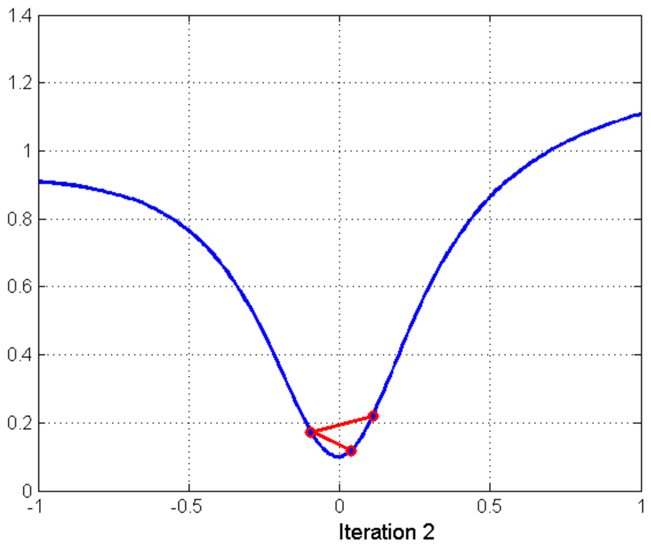
- avoids the need to bracket the root
- quadratic convergence (decimal accuracy doubles at every iteration)

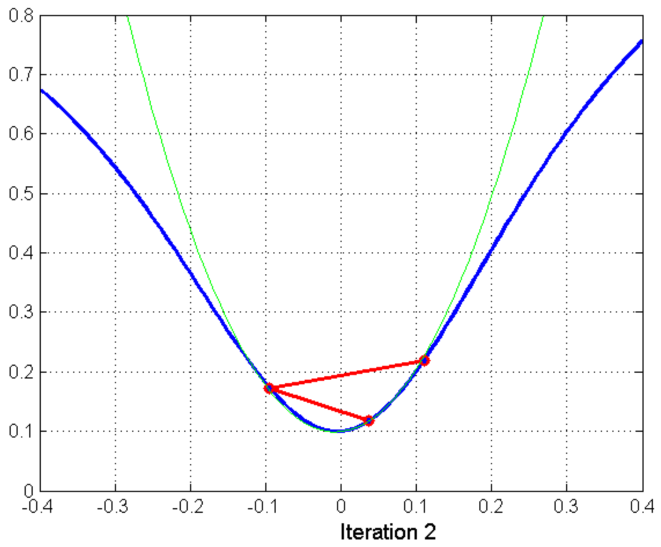
Newton method



- avoids the need to bracket the root
- quadratic convergence (decimal accuracy doubles at every iteration)

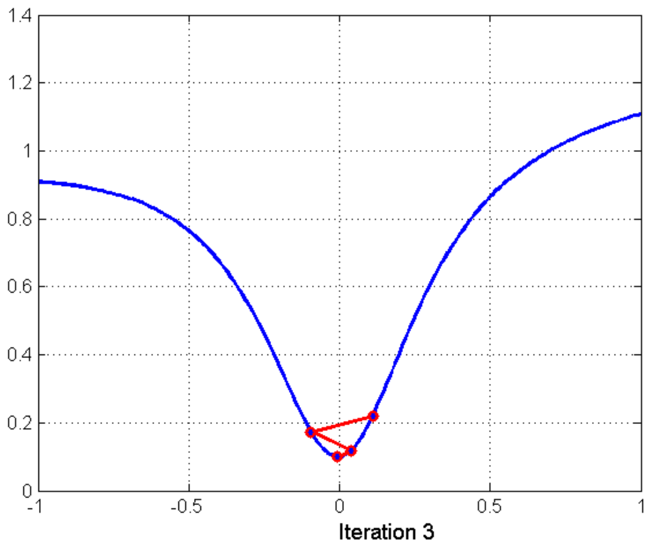
Newton method

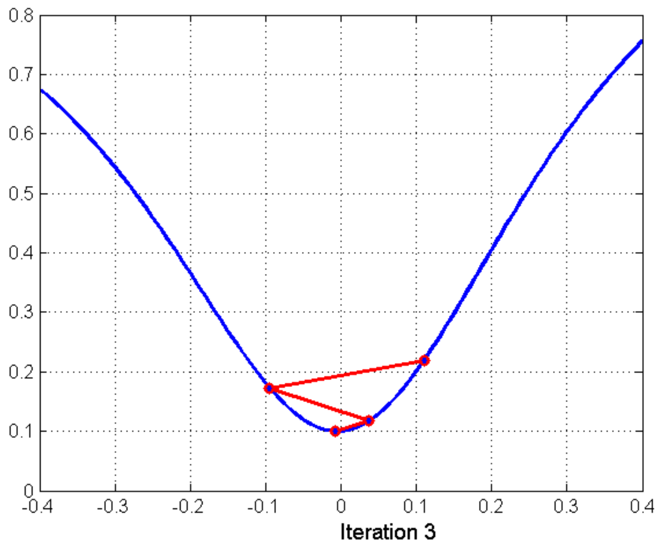




- avoids the need to bracket the root
- quadratic convergence (decimal accuracy doubles at every iteration)

Newton method

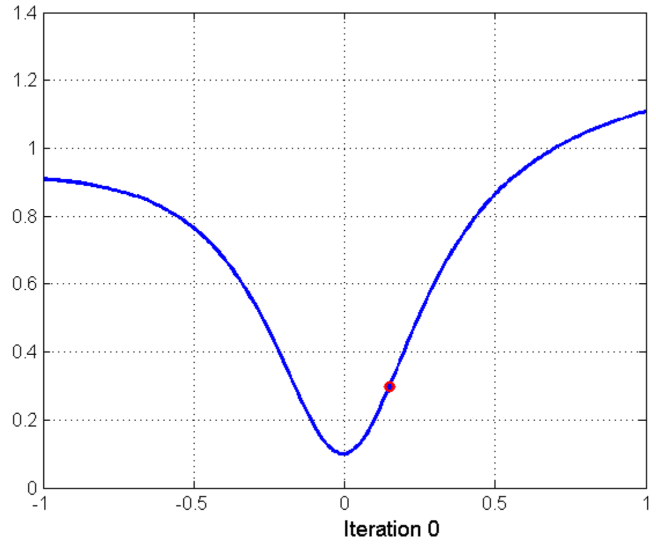


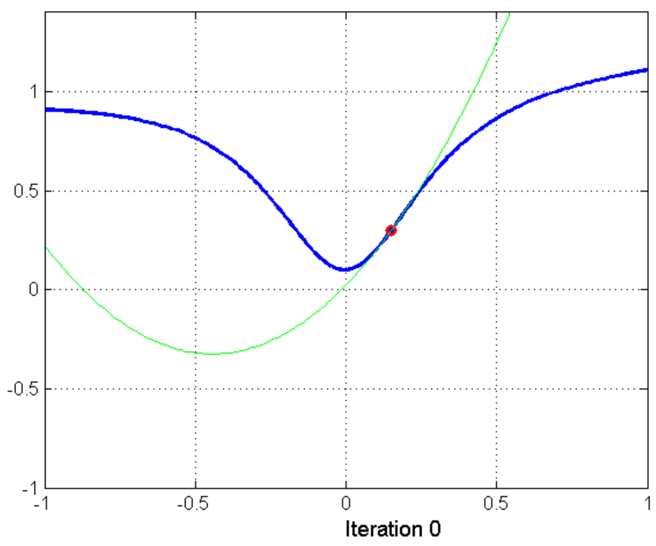


- avoids the need to bracket the root
- quadratic convergence (decimal accuracy doubles at every iteration)

Newton method

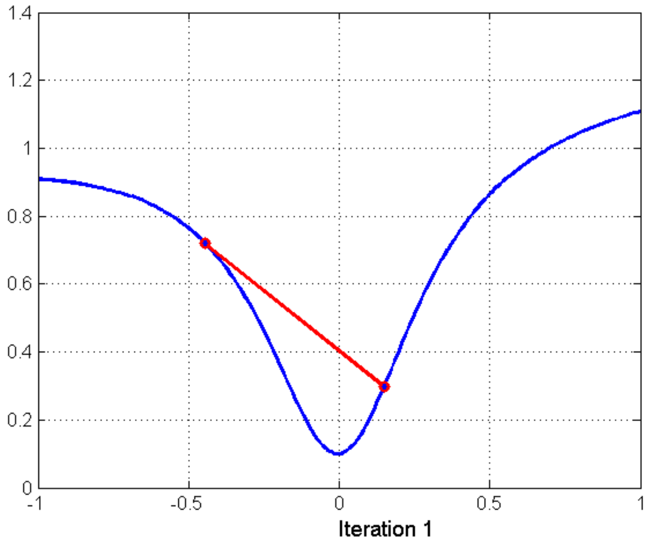
- Global convergence of Newton's method is poor.
- Often fails if the starting point is too far from the minimum.

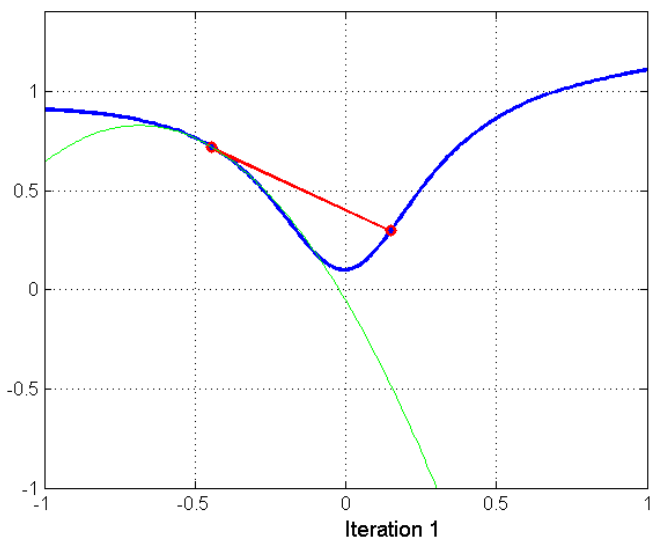




Newton method

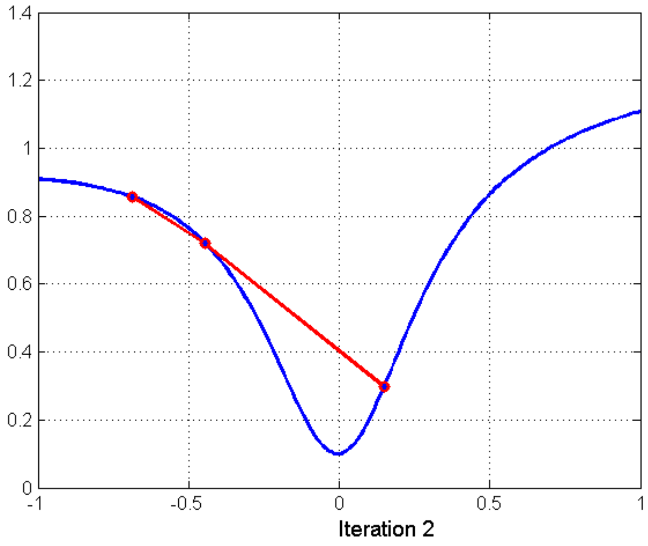
- Global convergence of Newton's method is poor.
- Often fails if the starting point is too far from the minimum.

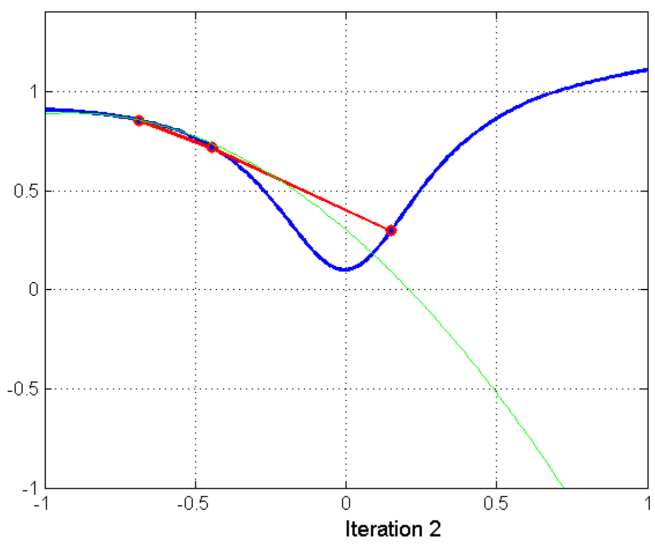




Newton method

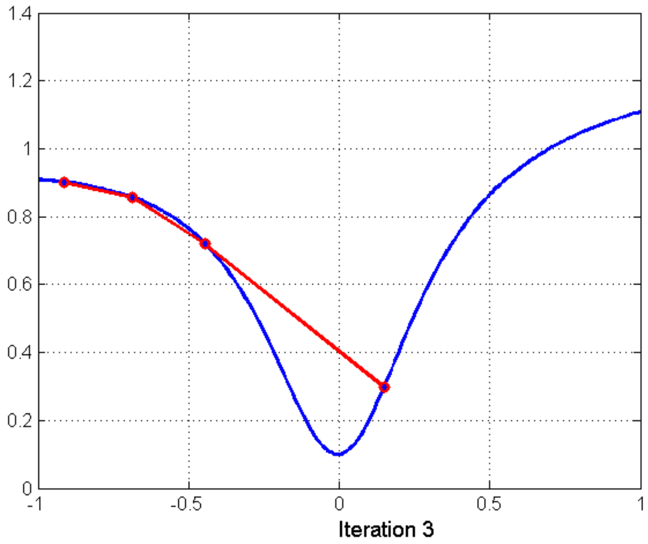
- Global convergence of Newton's method is poor.
- Often fails if the starting point is too far from the minimum.

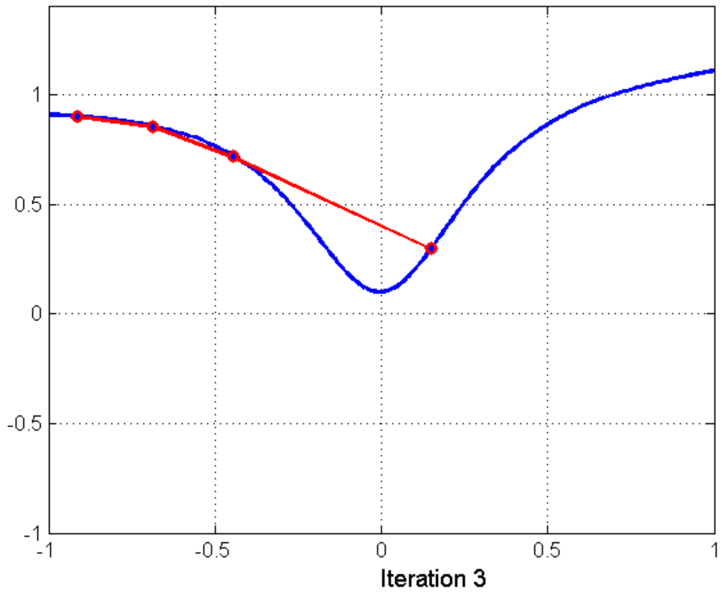




Newton method

- Global convergence of Newton's method is poor.
- Often fails if the starting point is too far from the minimum.





- in practice, must be used with a globalization strategy which reduces the step length until function decrease is assured

Extension to N (multivariate) dimensions

- How big N can be?
  - problem sizes can vary from a handful of parameters to many thousands