

# Numerical Optimization 07: 2nd order methods

Qiang Zhu

University of Nevada Las Vegas

May 20, 2020

# Overview

- 1 Newton's method
- 2 Extension to multivariate optimization
- 3 Secant Method
- 4 Summary

## Newton's method

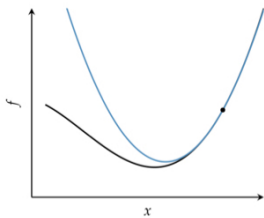
In optimization, knowing the first-order information can help determine the direction to travel, but does not help to determine how far to step to the local minimum. A better way is to use the second-order information. In univariable optimization, the quadratic approximation about a point  $(x^k)$  come from

$$q(x) = f(x^k) + (x - x^k)f'(x^k) + \frac{(x - x^k)^2}{2}f''(x^k)$$

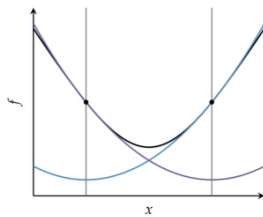
Setting the derivative to zero,

$$\begin{aligned}\frac{\partial q(x)}{\partial x} &= f'(x^k) + (x - x^k)f''(x^k) = 0 \\ x^{k+1} &= x^k - \frac{f'(x^k)}{f''(x^k)}\end{aligned}$$

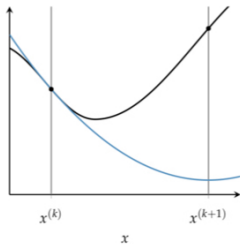
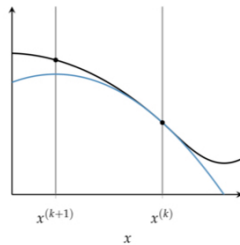
# Various cases



Oscillation



Overshoot

Negative  $f''$ 

## Extension to multivariate optimization

If  $f$  is a multivariate function

$$f(\mathbf{x}) = f(\mathbf{x}^k) + (\mathbf{g}^k)^T(\mathbf{x} - \mathbf{x}^k) + \frac{1}{2}(\mathbf{x} - \mathbf{x}^k)^T \mathbf{H}^k(\mathbf{x} - \mathbf{x}^k)$$

Setting the gradient to be zero,

$$\nabla q(\mathbf{x}^k) = \mathbf{g}^k + \mathbf{H}^k(\mathbf{x} - \mathbf{x}^k)$$

### Quiz

The Booth's function is

$$f(\mathbf{x}) = (x_1 + 2x_2 - 7)^2 + (2x_1 + x_2 - 5)^2$$

use the Newton's method to find the minimum when  $\mathbf{x} = [9, 8]$

# Newton's method with line search

Newtons method can also be used to supply a descent direction to line search or can be modified to use a step factor. Smaller steps toward the minimum or line searches along the descent direction can increase the methods robustness. The descent direction is:

$$\mathbf{d}^k = -(\mathbf{H}^k)^{-1} \mathbf{g}^k$$

# Secant Method

Newton's method for **univariate** function minimization needs to know the first and second derivatives. However, the second derivative is not easy to compute for some cases. The secant method use estimates of  $H$  as follows

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

The secant method requires **an additional initial design point**. It suffers from the same problems as Newton's method when quadratic function is not a good approximation.

# Summary

- Incorporating second-order information in descent methods often speeds convergence.
- Newton's method is a root-finding method that leverages second-order information to quickly descend to a local minimum.
- The secant method approximates Newton's method when the second-order information is not directly available.