Newton-Raphson Method for Convex Optimization

Conner DiPaolo, Jeffrey Rutledge, Colin Adams

Harvey Mudd College

February, 2016

Finding Roots of Functions

Given a differentiable function $f : \mathbb{R} \mapsto \mathbb{R}$ we want to find the instances when $f(x) = 0$ (not generally solvable in closed form).

Newton's Method:

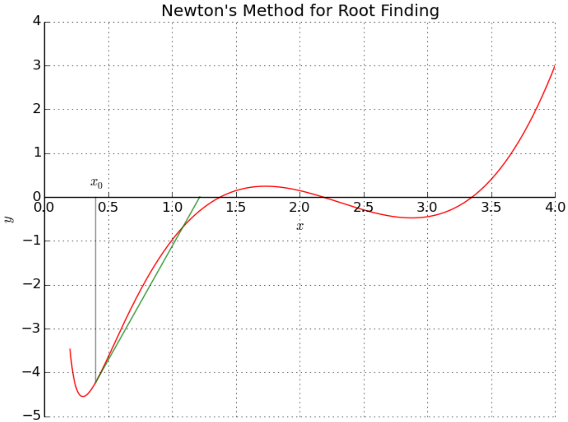
1. Take a starting position x_0
2. Find where the tangent line $y = f(x_n) + (x_{n+1} - x_n)f'(x_n)$ is 0 and iterate:

$$y = f(x_n) + (x_{n+1} - x_n)f'(x_n) = 0$$

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

Finding Roots of Functions

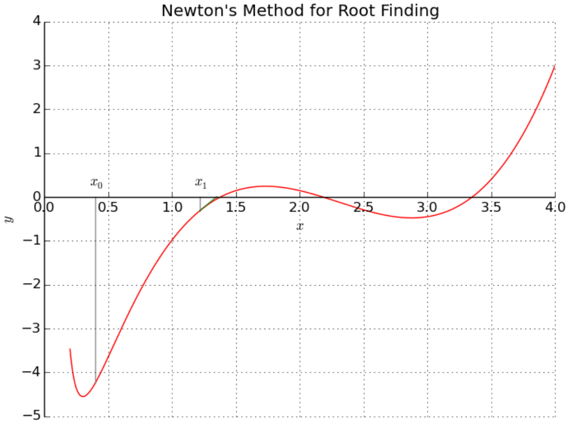
Initial $x_0 = 0.4,\ f(x) = (x-1)(x-3)^2 + e^{\frac{1}{3x}} - \cos\left(\frac{x}{2}\right) - 1.5$



$x_0 = 0.4 \quad \Delta = 0.9805$

Finding Roots of Functions

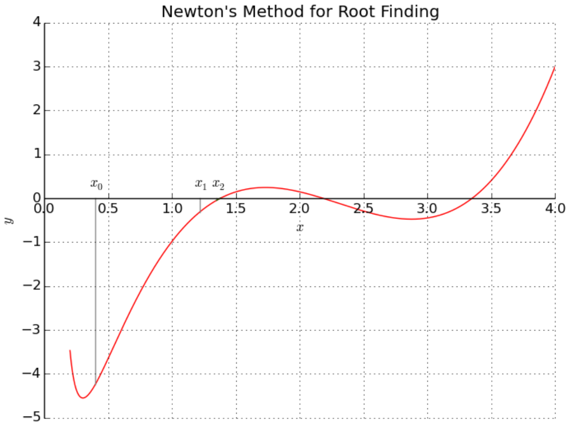
Initial $x_0 = 0.4$, $f(x) = (x-1)(x-3)^2 + e^{\frac{1}{3x}} - \cos\left(\frac{x}{2}\right) - 1.5$



$x_1 = 1.2167 \quad \Delta = 0.1638$

Finding Roots of Functions

Initial $x_0 = 0.4,\ f(x) = (x-1)(x-3)^2 + e^{\frac{1}{3x}} - \cos\left(\frac{x}{2}\right) - 1.5$



$$x_2 = 1.3487 \quad \Delta = 0.0319$$

Converges in 6 iterations.

Optimizing Function Using Newton's Method

Instead of finding where $f(x) = 0$, how can we find where $f'(x) = 0$?

Optimizing Function Using Newton's Method

Instead of finding where $f(x) = 0$, how can we find where $f'(x) = 0$?

Instead of:

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

Use derivatives:

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

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

How do we find the derivatives of functions where $f : \mathbb{R}^n \mapsto \mathbb{R}$?

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

How do we find the derivatives of functions where $f : \mathbb{R}^n \mapsto \mathbb{R}$?

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

becomes,

$$x_{n+1} = x_n - H_f^{-1}(x_n)\nabla f(x_n)$$

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

How do we find the derivatives of functions where $f : \mathbb{R}^n \mapsto \mathbb{R}$?

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

becomes,

$$x_{n+1} = x_n - H_f^{-1}(x_n)\nabla f(x_n)$$

Def: The gradient of $f : \mathbb{R}^n \mapsto \mathbb{R}$ is

$$\nabla f = \begin{bmatrix} \frac{\partial}{\partial x_1} f & \frac{\partial}{\partial x_2} f & \dots & \frac{\partial}{\partial x_n} f \end{bmatrix}$$

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

How do we find the derivatives of functions where $f : \mathbb{R}^n \mapsto \mathbb{R}$?

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

becomes,

$$x_{n+1} = x_n - H_f^{-1}(x_n)\nabla f(x_n)$$

Def: The gradient of $f : \mathbb{R}^n \mapsto \mathbb{R}$ is

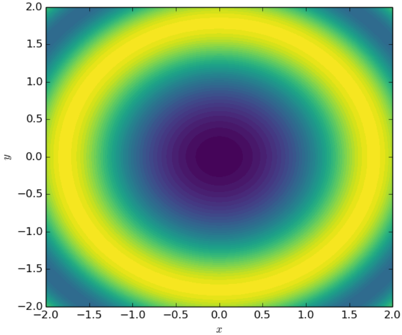
$$\nabla f = \begin{bmatrix} \frac{\partial}{\partial x_1} f & \frac{\partial}{\partial x_2} f & \dots & \frac{\partial}{\partial x_n} f \end{bmatrix}$$

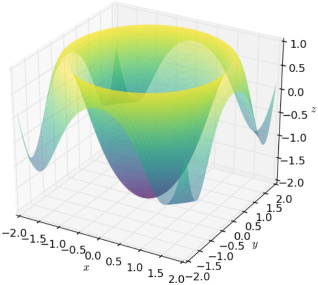
Def: The Hessian of $f : \mathbb{R}^n \mapsto \mathbb{R}$ is

$$H_f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \ddots & & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{bmatrix}$$

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

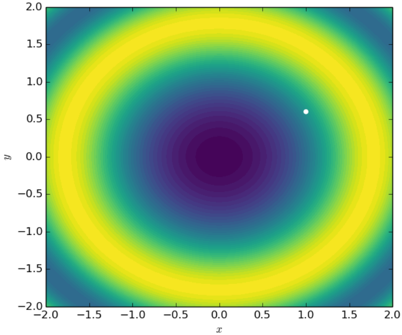
Initial $x_0 = \begin{bmatrix} 1 & 0.6 \end{bmatrix}$, $f(x) = -\cos(x^2 + y^2) - e^{-(x^2+y^2)}$

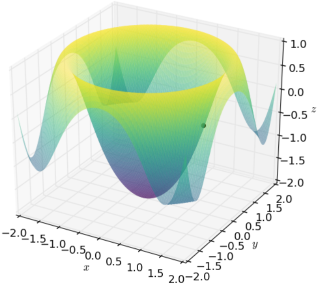




Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

Initial $x_0 = \begin{bmatrix} 1 & 0.6 \end{bmatrix}$, $f(x) = -\cos(x^2 + y^2) - e^{-(x^2+y^2)}$

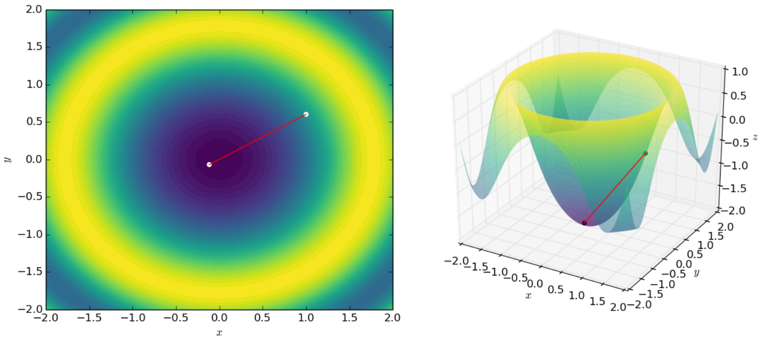




$x_0 = \begin{bmatrix} 1 & 0.6 \end{bmatrix}$ $\Delta = 1.1661$

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

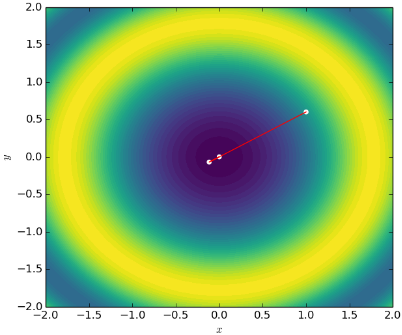
Initial $x_0 = \begin{bmatrix} 1 & 0.6 \end{bmatrix}$, $f(x) = -\cos(x^2 + y^2) - e^{-(x^2+y^2)}$

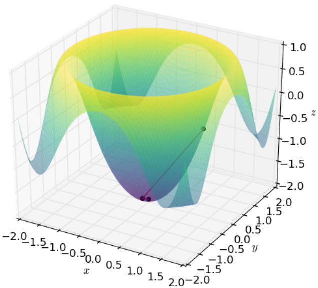


$x_1 = \begin{bmatrix} -0.1167 & -0.0700 \end{bmatrix}$ $\Delta = 0.1361$

Optimizing $f : \mathbb{R}^n \mapsto \mathbb{R}$

Initial $x_0 = \begin{bmatrix} 1 & 0.6 \end{bmatrix}$, $f(x) = -\cos(x^2 + y^2) - e^{-(x^2+y^2)}$





$$x_2 = \begin{bmatrix} -7.8 \times 10^{-5} & -4.7 \times 10^{-5} \end{bmatrix} \qquad \Delta = 9.15 \times 10^{-5}$$

Converges in 4 iterations.