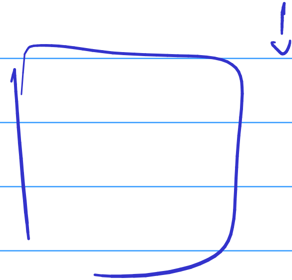
Overview

$f(x)=0$

Solving in 1D for a scalar



$\vec{f}(\vec{x}) = \vec{0}$

n components

$\min f(x) \rightsquigarrow f'(x)=0$ solve

Solving Nonlinear Equations

- What is the goal here?

$$f(x, y) = 0 \quad \rightsquigarrow \quad x^2 + y = 1$$

$$g(x, y) = 0 \quad \rightsquigarrow \quad y^2 - 3xy - y = 18$$

Newton's method

- What does Newton's method look like in n dimensions?

1D: $f(x) \rightsquigarrow f(x_k+s) = f(x_k) + f'(x_k)s$

$$f_1(x,y,z),\quad f_2(x,y,z),\quad f_3(x,y,z) \qquad \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \vdots & \vdots & \vdots \\ \frac{\partial f_3}{\partial x} & \cdots & \frac{\partial f_3}{\partial z} \end{pmatrix} = J_f(x,y,z)$$

nD: $\vec{f}(\vec{x}) \rightsquigarrow \vec{f}(\vec{x}_k+\vec{s}) = \vec{f}(\vec{x}_k) + J_f(\vec{x}_k)\vec{s} = \vec{0}$

$\Leftrightarrow \quad J_f(\vec{x}_k)\vec{s} = -\vec{f}(\vec{x}_k)$

$\Leftrightarrow \quad \vec{s} = -J_f^{-1}(\vec{x}_k)\vec{f}(\vec{x}_k)$

$\vec{x}_{k+1} = \vec{x}_k + \vec{s} = \vec{x}_k - J_f^{-1}(\vec{x}_k)\vec{f}(\vec{x}_k)$

1D: $x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$

Newton: Example

- Set up Newton's method to find a root of

$$f(x,y)=\begin{pmatrix} x+2y-2 \\ x^2+4y^2-4 \end{pmatrix}\begin{matrix} f_1 \\ f_2 \end{matrix} = 0$$

Demo: Newton's method in n dimensions

$$J_f(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}\ \boxed{1} & \frac{\partial f_1}{\partial y}\ \boxed{2} \\ \frac{\partial f_2}{\partial x}\ \boxed{2x} & \frac{\partial f_2}{\partial y}\ \boxed{8y} \end{pmatrix}$$

17 Finding the Best: Optimization in 1D

Optimization

- State the problem.

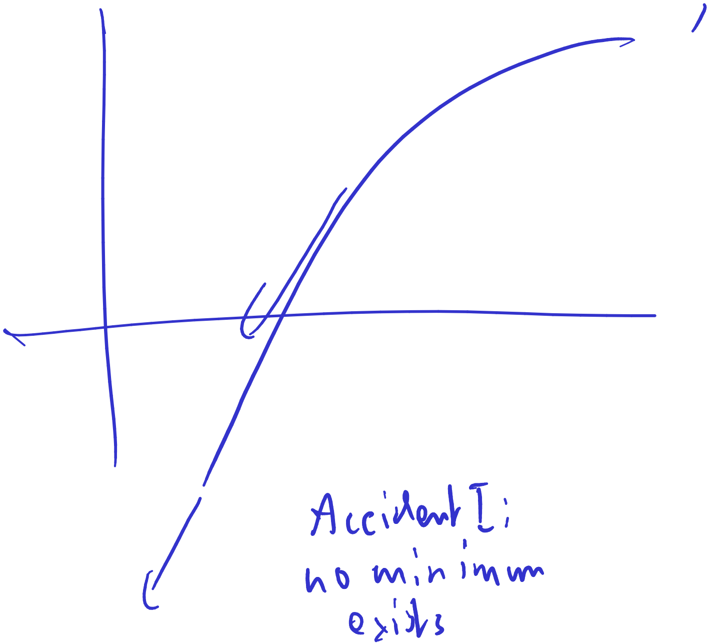
$f: \mathbb{R} \to \mathbb{R}$ ← objective function

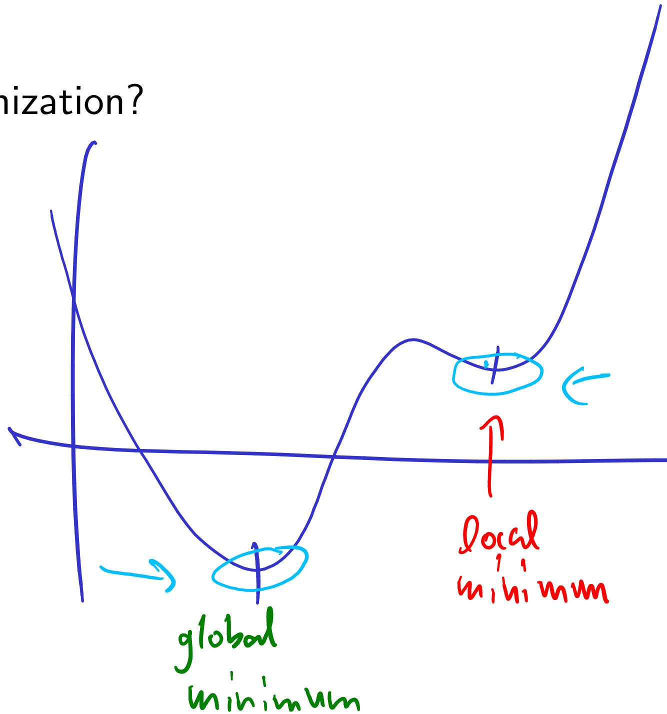
Find x^* so that $f(x^*) = \min_x f(x)$.

Maximize? Just use $-f$.

Optimization: What could go wrong?

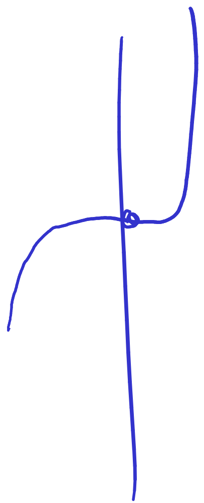
- What are some potential problems in optimization?





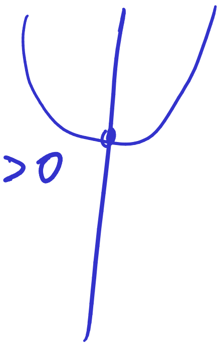
Optimization: What is a solution?

- How can we tell that we have a (at least local) minimum? (Remember calculus!)



Necessary cond: $f'(x) = 0$

Sufficient cond: $f'(x) = 0$ and $f''(x) > 0$



Newton's Method

- Let's steal the idea from Newton's method for equation solving: Build a simple version of f and minimize that.

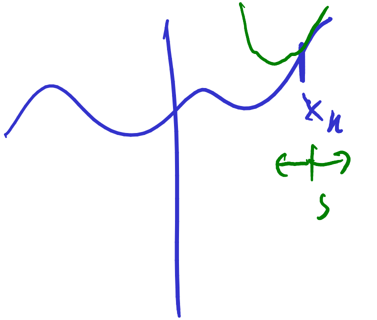
To minimize $f(x)$

$$\tilde{f}(x_k+s) = f(x_k) + f'(x_k)\cdot s + \frac{f''(x_k)}{2}s^2$$

take deriv. w.r.t. s

$$\tilde{f}'(x+s) = 0 + f'(x_k) + f''(x_k)\cdot s = 0$$

necessary for min



$$f'(x)=0$$

$$f''(x_k)\,s = -f'(x_k)$$

$$s = -\frac{f'(x_k)}{f''(x_k)}$$

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

Demo: Newton's method in 1D
In-class activity: Optimization Methods

18 Optimization in n Dimensions

Optimization in *n* dimensions: What is a solution?

- How can we tell that we have a (at least local) minimum? (Remember calculus!)

$f(\vec{x}) \leftarrow$ objective function $f: \mathbb{R}^n \to \mathbb{R}$

nec: $f'(x) = 0$

suf: $f''(x) > 0$

$$\frac{\partial f}{\partial x} = 0 \quad \frac{\partial f}{\partial y} = \quad \frac{\partial f}{\partial z} = 0$$

$\nabla f(x) = 0$ / $f(x, y, z) = 15$

$$H_f(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial z \partial x} & \cdots & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}$$

$\frac{\partial^2 f}{\partial x \partial z} = \frac{\partial^2 f}{\partial z \partial x}$

$H_f(x)$ symmetric

H_f positive definite

Steepest Descent

- Given a scalar function $f\colon \mathbb{R}^n \to \mathbb{R}$ at a point $\boldsymbol{x}$, which way is down?

Demo: Steepest Descent

Newton's method (nD)

- What does Newton's method look like in n dimensions?

$f(\vec{x})$

$$\tilde{f}(\vec{x}_k + \vec{s}) = f(\vec{x}_k) + \nabla f(\vec{x}_k) \cdot \vec{s} + \frac{1}{2} \vec{s}^T H_f(x_k) \vec{s}$$

$$= \nabla_{\vec{s}} \tilde{f}(\vec{x}_k) = 0 + \nabla f(\vec{x}_k) + H_f(\vec{x}_k) \vec{s}$$

$$\vec{s} = - H_f(\vec{x}_k)^{-1} \nabla f(\vec{x}_k)$$

$$\vec{x}_{k+1} = \vec{x}_k - H_f^{-1}(\vec{x}_k) \cdot \nabla f(\vec{x}_k)$$

Demo: Newton's method in n dimensions

Demo: Nelder-Mead Method

$$x_{k+1} = x_k - \frac{f(x)}{\text{slope}}$$

$$\text{slope} = f'(x)$$

$$\text{slope} = \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}$$

Nonlinear Least Squares/Gauss-Newton

- What if the f to be minimized is actually a 2-norm?

$$f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_2, \qquad \boldsymbol{r}(\boldsymbol{x}) = \boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x})$$

Demo: Gauss-Newton