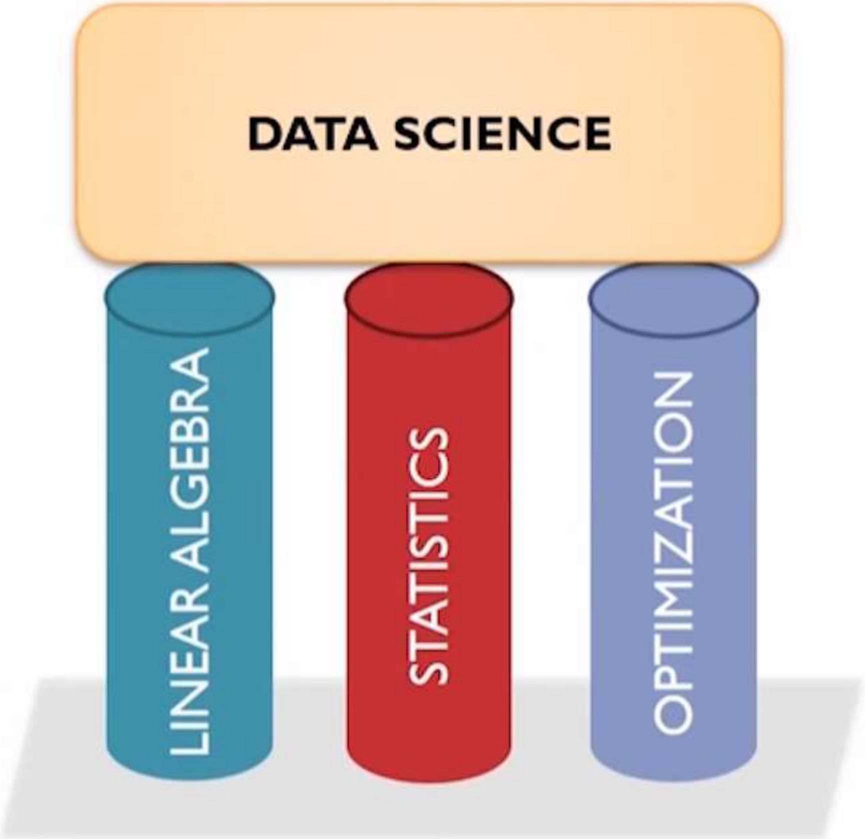
Optimization for Data Science

Unconstrained nonlinear optimization

Constrained nonlinear optimization

Connections to data science

Three pillars of data science



Fundamentals of optimization

What is optimization ?

"An optimization problem consists of maximizing or minimizing a real function by systematically choosing input values from within an allowed set and computing the value of the function."*

*"http://en.wikipedia.org/wiki/Mathematical_optimization"

What is optimization?

- … the use of specific methods to determine the "best" solution to a problem
  - Find the best functional representation for data
  - Find the best hyperplane to classify data

Minimize / maximize $f(X)$

X = Decision variables acting on X

minimize error $\sum_{i=1}^{m} e_i^2$

a_0, a_1

$e_1 = y_1 - a_0 - a_1 x_1$

$e_n = y_n - a_0 - a_1 x_n$

$y_1 = a_0 + a_1 x_1$

$y_n = a_0 + a_1 x_n$

$\sum_{i=1}^{n} e_i^2$

y	x
y_1	x_1
y_2	x_2
y_n	x_n

$y = a_0 + a_1 x$

Why optimization for machine learning

- (Almost) All machine learning (ML) algorithms can be viewed as solutions to optimization problems
  - Even in cases where, the original machine learning technique has a basis derived from other fields
- A basic understanding of optimization approaches help in
  - More deeply understand the working of the ML algorithm
  - Rationalize the workings of the algorithm
  - And (may be !!!), develop new algorithms ourselves

Components of an optimization problem

- Objective function
  - We look at minimization problem
- Decision variables
- Constraints

f

f

$-f$

$\min f(x)$

$x \in \Omega$

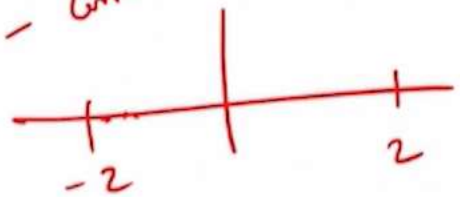
Types of optimization problems

- Depending on the type of objective function, constraints and decision variables
  - Linear programming problem ✓
  - Nonlinear programming problem ✓
    - Convex vs Non-convex
  - Integer programming problem (linear and nonlinear)
  - Mixed integer linear programming problem
  - Mixed integer nonlinear programming problem

$f(x)$

S.t constraints → linear

x – continuous variables



$f(x_1, x_2)$

$x_1 \in [0,1,2,3]$

x_2 is continuous

$\min f(x_1, x_2)$

$x_1 \in [0,1,2,3]$

$x_2 \in [0,1,2,3]$

constraints on x_1 and x_2

$x_1 \in [0,1]$

$x_2 \in [0,1]$

Nonlinear Optimization

UNCONSTRAINED CASE

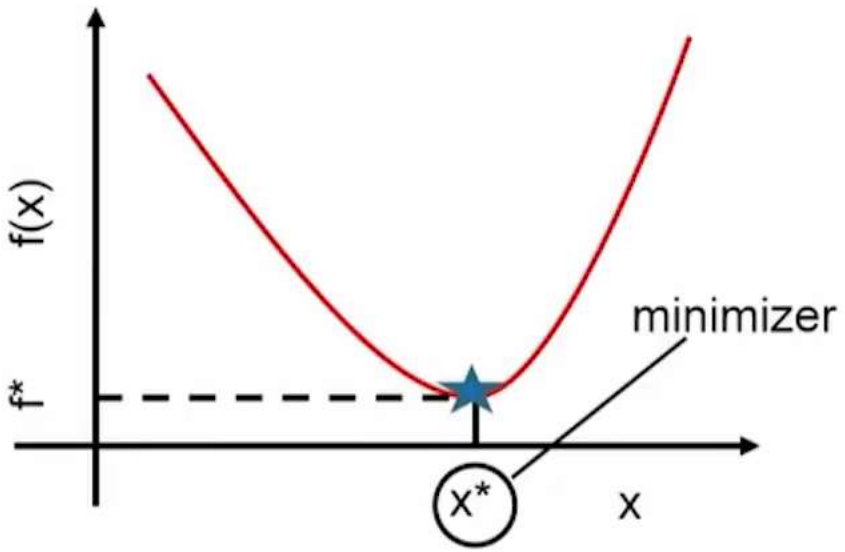
Univariate Optimization – Local and Global Optimum

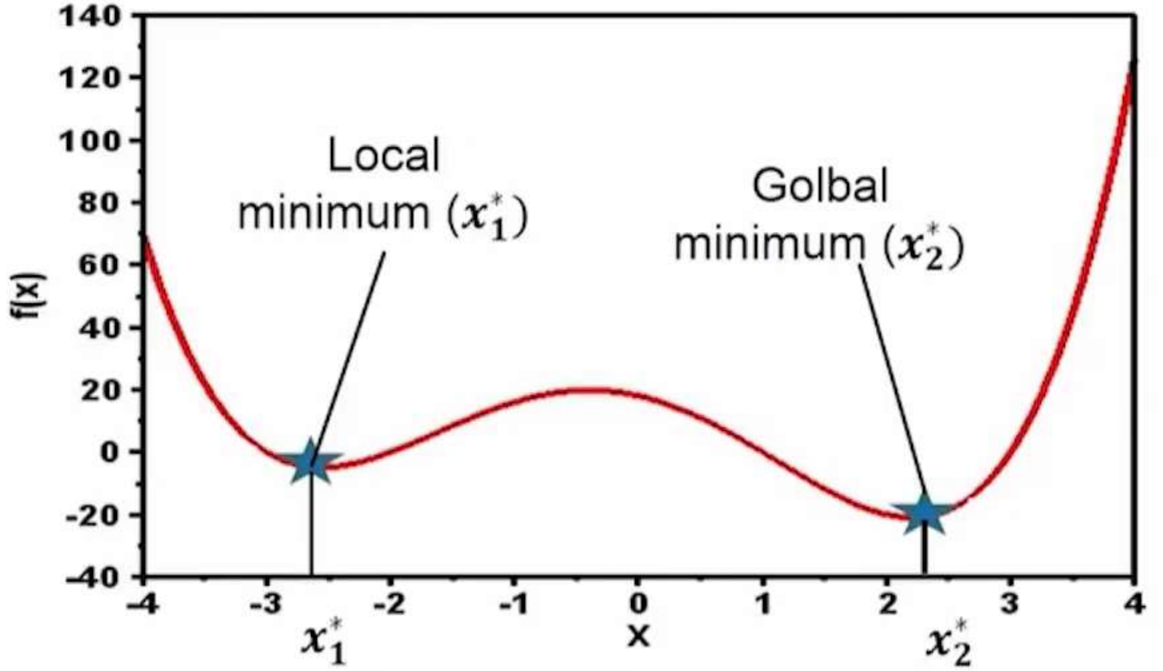
Univariate optimization

$$\boxed{\begin{array}{c} \min_{x} \; f(x) \\ x \in R \end{array}}$$

Decision variable

Objective function

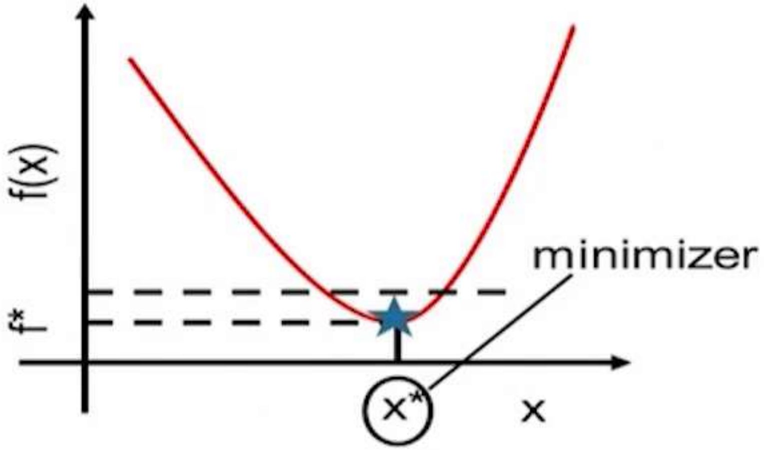




Univariate Optimization – Conditions for Local Optimum

Univariate optimization

$$\boxed{\begin{array}{c}\min\limits_{x} \quad f(x) \\ x \in R\end{array}}$$



Approximate $f(x)$ as a quadratic function using Taylor series at a point x^k

$$f(x) \approx f(x^k) + \frac{1}{1!} f'(x^k)(x - x^k) + \frac{1}{2!} f''(x^k)(x - x^k)^2$$

When $x^k = x^*$,

$$f(x) \approx f(x^*) + \underbrace{\cancel{\frac{1}{1!} f'(x^*)(x - x^*)}}_{0} + \frac{1}{2!} f''(x^*)(x - x^*)^2$$

$$\underbrace{f(x) - f(x^*)}_{\text{Positive}} \approx \frac{1}{2!} \underbrace{f''(x^*)}_{\text{Has to be positive}} \underbrace{(x - x^*)^2}_{\text{Always positive}}$$

Univariate Optimization – Summary

Univariate optimization

$$\boxed{\begin{array}{c}\min_{x} \quad f(x) \\ x \in R\end{array}}$$

Necessary and sufficient conditions for x^* to be the minimizer of the function $f(x)$

First order necessary conditon: $f'(x^*) = 0$

Second order sufficiency condition: $f''(x^*) > 0$

Univariate Optimization – Numerical Example

$$\min_{x} \; f(x)$$

$$f(x) = 3x^4 - 4x^3 - 12x^2 + 3$$

First order condition

$$f'(x) = 12x^3 - 12x^2 - 24x = 0$$

$$= 12x(x^2 - x - 2x) = 0$$

$$= 12x(x+1)(x-2) = 0$$

$$x = 0, x = -1, x = 2$$

Second order condition

$$f''(x) = 36x^2 - 24x - 24$$

$$f''(x)\big|_{x=0} = -24$$

$$f''(x)\big|_{x=-1} = 36 > 0$$

$$f''(x)\big|_{x=2} = 72 > 0$$

$$f(-1) = -2$$

$$f(2) = -29$$

$x^* = -1$, is a local minimizer of $f(x)$

$x^* = 2$, is a global minimizer of $f(x)$