Optimization for Machine Learning

Lionel Tondji

African Master's in Machine Intelligence

July 15, 2024

Short Presentation

- Name: Lionel Tondji
- Email: tngoupeyou@aimsammi.org
- Post Doctoral researcher since November 2023
- Research : Optimization, Inverse Problems, Machine Learning

Core Info and Course Outline

- Course : 15.07 - 02.08
  - From Monday to Thursday : 2 PM - 4 PM
- Lab projects: students send their lab python projects to: tngoupeyou@aimsammi.org
- Exercises : Monday to Thursday 2.30PM - 6.30PM
- Quiz every Friday : 2 - 4PM
- Lab every Friday : 2.30PM - 6.30PM

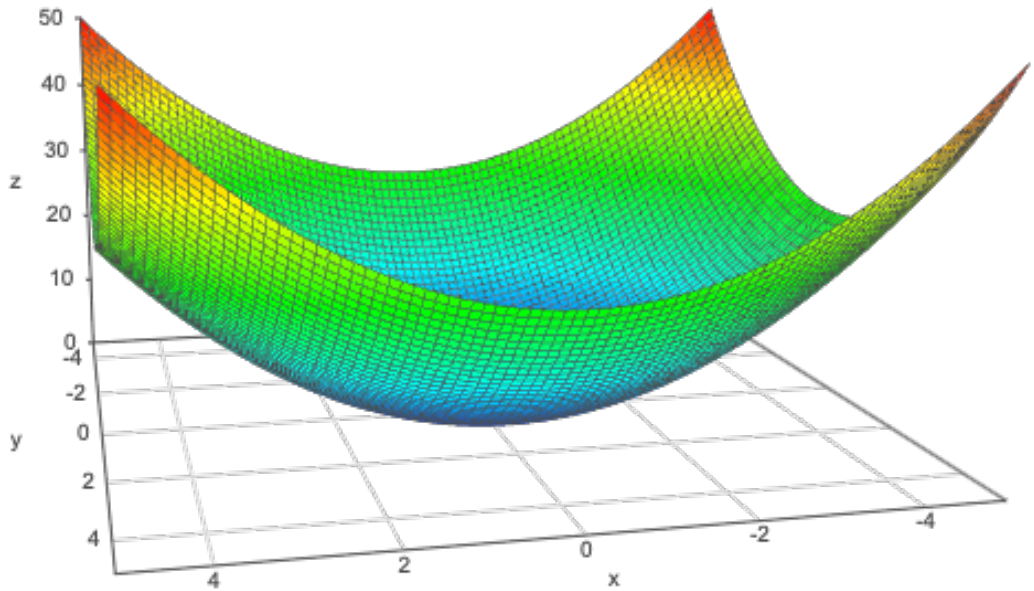
Introduction

Optimization:

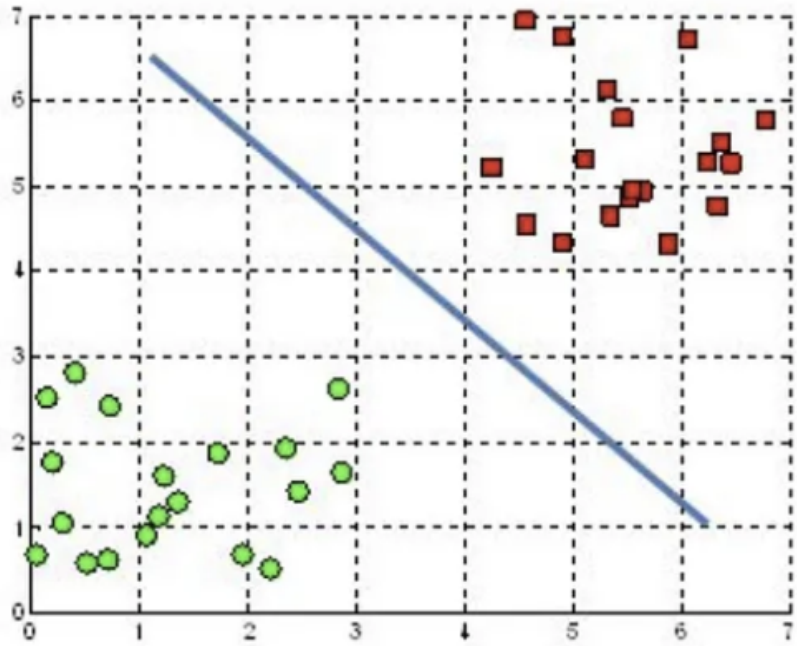
- Trying to find settings of parameters in order to optimize something.
- Optimize a cost function
- Optimizing is going to be associated with the words cost.
- Can also think about: maximizing profit, maximizing benefit, rewards.

Example : Minimizing a function in 3D

Trying to find the lowest point.



Example : 2D classification problem

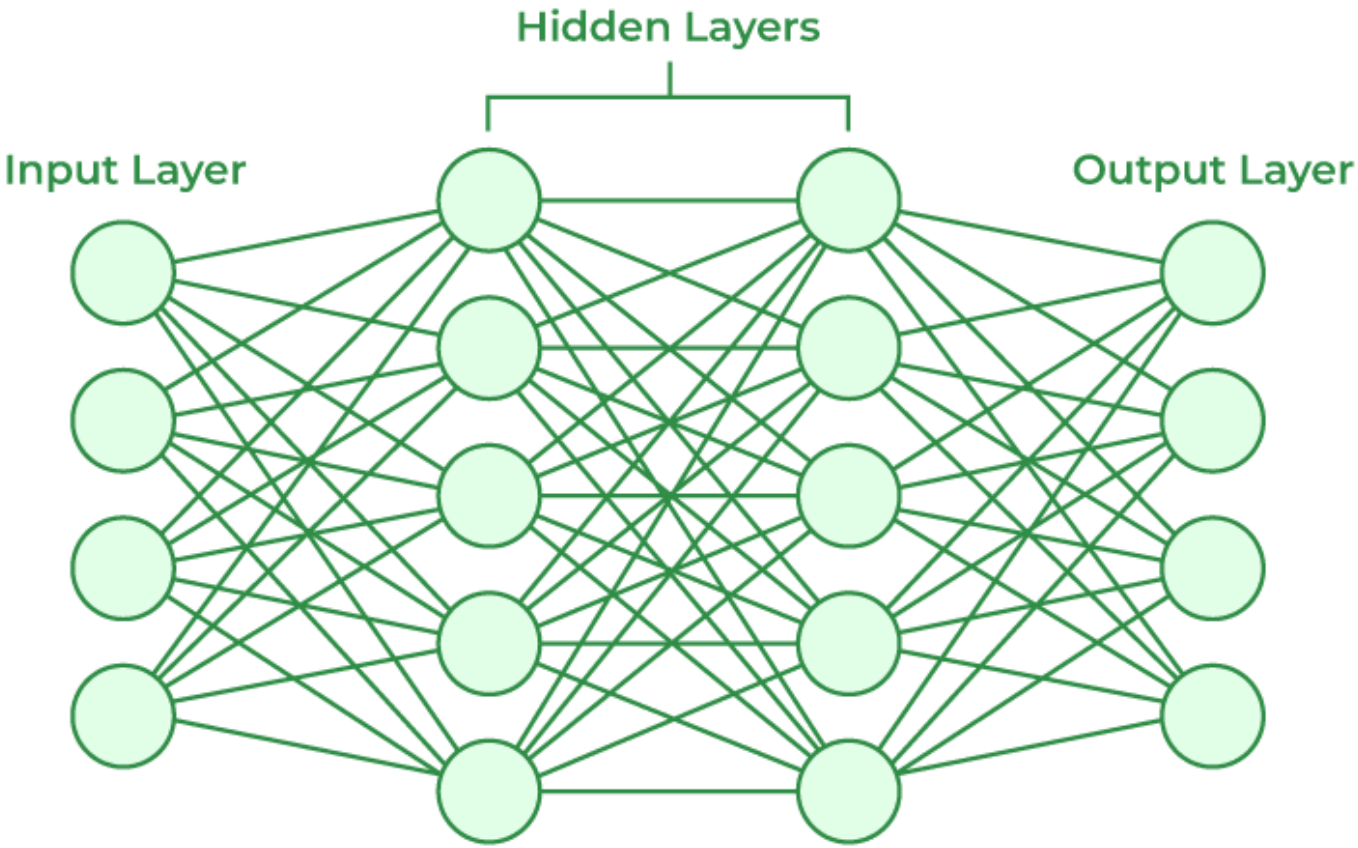


Example : Denoisy Problem

Remove all the noise to the best possible of our abilities

Example : Neural Networks

We want to minimize the error.



This class is going to focus on

- Not so much on these particular applications but rather the methodology
- Optimization as a powerful and general purpose tool to replace the impossible (brute force) by something that can be implemented and executed reliably and efficiently.

What these lectures do and do not cover

Optimization:

$$\begin{aligned} &\text{minimize: } f(x) \\ &\text{subject to : } x \in \mathcal{X} \subseteq \mathbb{R}^n. \end{aligned}$$

- **What can this model?**
  1. Classification: find parameters that give the minimum error
  2. Image denoising: Cost = distance from the original image + "how noisy image looks"
  3. Image In-painting
     - Parameters: values of missing pixels
     - Objective: "$f(x)$" measure of many sharp boundaries I have
  4. matrix completion and collaborative Filtering
  5. Neural Networks:
     - $f(x)$ = Classification error and x : all the weights of the NN

What these lectures do and do not cover

Optimization:

$$\begin{aligned} &\text{minimize: } f(x) \\ &\text{subject to : } x \in \mathcal{X} \subseteq \mathbb{R}^n. \end{aligned}$$

- **When can we solve it**? We need f and $\mathcal{X}$ to be convex.
- what this course will not cover?
  1. How do we model interesting problems using convex Optimization
  2. What are specialized solvers we can then call to solve these problems?
- Will cover
  1. What are the basic algorithms for solving these convex problems?
  2. How do we develop new algorithms? How fast do they run, and what computing resources do they need.

Example of problems this course will cover

- Least-square Regression
- Ridge (L2-regularized) Regression
- Lasso (L1-regularized) Regression
- Logistic Regression
- General convex functions for which we can compute derivatives.

In Summary

This class

1. Fundamental (broad) problem classes
2. Algorithms tailored for these classes of functions
3. Analysis of convergence rates

Acknowledgements

Thanks for your attention!