Mathematical optimization in ML & AI

Introduction

Optimization is the process of **finding the best solution** from all feasible solutions. It is a central element in Machine Learning as optimization plays a central role in training models and **has an immediate impact on predictions or decisions** of the final model

In practice, this means to **maximize or minimize some function** by systematically choosing input values from within an allowed set and computing the value of the function.

Types of Optimization Problems

Optimization problems can be broadly classified based on the nature of the objective function, the constraints, and the variables involved. Main types include:

Linear Optimization

- Both the objective function and the constraints are linear functions of the decision variables
- In ML: E.g. Linear Regression parameter fitting, Linear Support Vector Machines (SVM)

Nonlinear Optimization

- The objective function or at least one of the constraints is a nonlinear function of the decision variables
- In ML:
    - Convex problems (global minimum, efficient to find): Kernelized Support Vector Machines, Lasso and Ridge Regression (e.g. Feature Selection)
    - Non-convex problems (multiple local minima, challenging): Neural Net training, Generative Adversarial Networks Optimization
- Of great relevance for practical ML

Integer Optimization

- All or some of the decision variables are restricted to integer values
- In ML: Decision Tree depth limitation, Feature selection
- Often NP-hard

Combinatorial Optimization

- The optimization of an objective function whose domain is a discrete but large configuration space
- In ML: Hyperparameter tuning, Neural architecture search (NAS)
- Often resorting to heuristics to find suboptimal solutions

Approach

The approach to solving an optimization problem largely depends on the type of problem. There's two main camps:

Analytical Methods

- Involves solving the problem using mathematical techniques
- Mostly useful for simpler problems or problems that can be reduced to a simpler form
- Closed-form solutions often exists, as e.g. in linear regression

Numerical Methods

- Typically involve iterative methods to approximate the solution
- Fallback when analytical solutions are not feasible