1. Optimization Introduction

Motivation for Optimization and Convexity

Sun Woo Lim

Jan 6, 2021

What is an Optimization Problem?

$$p^* = \min_{x \in \mathbb{R}^n} \quad f_0(x)$$

$$\text{subject to:} \quad f_i(x) \le 0,\ i = 1, \ldots, m,$$

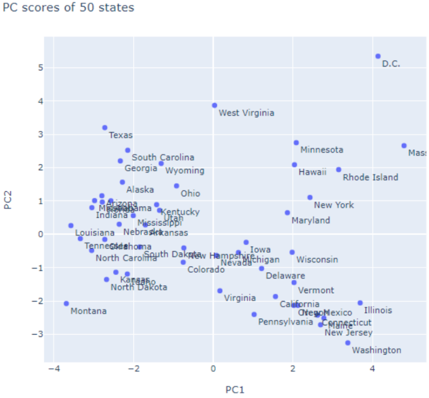
$$h_i(x) = 0,\ i = 1, \ldots, q$$

✓ Objective function you want to minimize / maximize
✓ m inequality constraints
✓ q equality constraints
✓ Why Study Optimization? Beautiful Theory along with zillions of real life problems!

Optimization Examples

1. Data Science and Statistics

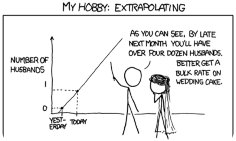
1) PCA on the US election data : "maximize variance"



Optimization Examples

1. Data Science and Statistics

2) Ordinary Least Squares



3) Supervised Learning

- RIDGE, LASSO, Logistic Regression, SVM ⊂ Supervised Learning

Figure 9.21 Comparison of original boat image (a), wavelet compression with $\lambda = 10$ (b), and wavelet compression with $\lambda = 30$ (c).

"Efficient Image Compression Using LASSO (⊂QP)"

Optimization Examples

2. Economics

1. Zero Sum Game in Game Theory

- You pick a row, opponent picks a column.
- Element : your payoff (that your opponent gives)
- Would you play first?

7	-8	-7	-8	3	5
9	-5	10	-2	-10	5
-8	1	10	9	7	-2
9	10	0	6	9	3
3	10	6	10	4	-7

Rational agents : if I go for the second, the opponent knows that I will choose the max element of the selected column. Then, the opponent gauges several colmax values: 9,10,10,10,9,5. So, the opponent has to choose the last column. This is "min-max" : minimize the maximum.

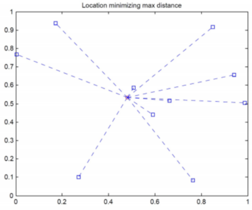
2. Markowitz (UCSD Prof., 1990 Nobel Prize in Economics) Portfolio

- $W \in \mathbb{R}^n$: vector of **returns**
- $p = E(W), \Sigma = E[(W-p)(W-p)^T]$.
- minimize$_x (-p'x + \gamma x'\Sigma x)$ s.t. $1'x = 1$ and $x_i \geq 0, \forall i$
- x : fraction of wealth inverstigated in different stocks.
- γ : risk sensitivity parameter.

Optimization Examples

3. Engineering

1. **Optimal location of Power supply station (minimax, minimize average distance)**



2. **GPS localization**

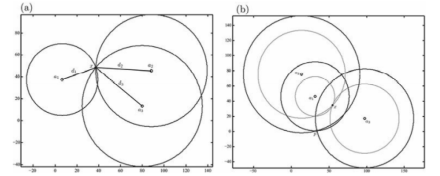


Figure 10.4 (a) Correct sync and localization; (b) wrong sync (light gray circles): p is the true point position, x is the estimated one; in this case, the correct sync (black circles) can be found by maximizing δ, instead of minimizing it.

Optimization Examples

4. Skier Path Optimization

Exercise 9.2 (A slalom problem) A two-dimensional skier must slalom down a slope, by going through n parallel gates of known position (x_i, y_i), and of width c_i, $i = 1, \ldots, n$. The initial position (x_0, y_0) is given, as well as the final one, (x_{n+1}, y_{n+1}). Here, the x-axis represents the direction down the slope, from left to right, see Figure 9.24.

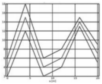
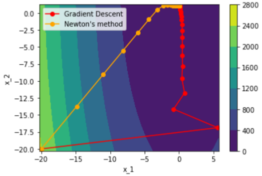


Figure 9.24 Slalom problem with $n = 5$ obstacles. "Uphill" (resp. "downhill") is on the left (resp. right) side. The middle path is dashed, initial and final positions are not shown.

1. Find the path that minimizes the total length of the path. Your answer should come in the form of an optimization problem.

5. Algorithm (Gradient Descent vs Newton's)

```
visualize_comparison(x_grad2, x_newton2, f2, watch_path=False)
```



Course Structures

1. Linear Algebra

: LA form a basis for this whole course + some optimization problems solely can be solved using *LA* knowledge
Keywords : Vector, Matrix, Norm, Symmetric Matrix, PSD, PD, EVD, SVD, Projection, Matrix Condition Number

2. Convex Optimization Problem

: Start discussing optimization problem, especially convex optimization problem. Learn how to form a dual of the original problem.
Keywords : Convex Set, Convex Function, Lagrangian, KKT, Slater's, Strong Dual, Weak Dual

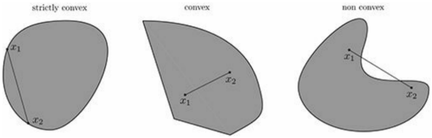
3. Types of Convex Optimization Problem

Although 'convex problem' is a broad subject, can be formulated as LP, QP, QCQP, SOCP, SDP. Full of real life examples
Keywords: LP, QP, QCQP, SOCP, SDP

4. Algorithms

Learn Algorithms to solve optimization problems *without constraints*. Not merely learning what Gradient Descent + Newton's method are, analyze how to adjust the 'step size' to make the algorithm converge.
Keywords: Descent Algorithm, Step Size, Gradient Descent, Newton's Method, Armijo Condition, Exact Line Search, Backtracking Line Search

1. Convex Set



$K \subseteq \mathbb{R}^n$ is a **convex set** if $\forall x_1, x_2 \in K$, $\forall \lambda \in [0, 1]$, $\lambda x_1 + (1 - \lambda)x_2 \in K$
"The Convex Combination also is in the set".

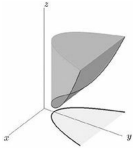
✓ Which of the following is / are convex?

1. **Empty Set** ϕ
2. **A set of single point** $\{x_0\}$
3. $\{z \in R^n : ||z - z_0||_2 \leq \epsilon\}$ **for some** $\epsilon > 0$
4. $\{z \in R^n : ||z - z_0||_2 = \epsilon\}$ **for some** $\epsilon > 0$
5. $[-2, -1] \cup [1, 2]$

Convexity

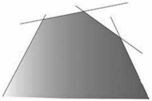
1. Convex Set - Operations that preserve convexity of a set

1. **Intersection of convex sets**
2. **Hyperplane** $\{x|a^T x - b = 0\}$ **and Half Spaces** $\{x|a^T x - b \leq 0\}$ **and** $\{x|a^T x - b > 0\}$
3. **Projection of a convex set onto a hyperplane**



4. **"Convex Hull of** A**"** $Co(A) := \{\sum_{i=1}^{m} \lambda_i x_i | x_i \in A, \lambda_i \geq 0, \forall i, \sum \lambda_i = 1\}$
5. **"Conic Hull of** A**" :=**$\{\sum_{i=1}^{m} \lambda_i x_i | x_i \in A, \lambda_i \geq 0, \forall i\}$
6. **"Affine Hull of** A**" :=**$\{\sum_{i=1}^{m} \lambda_i x_i | x_i \in A, \sum \lambda_i = 1\}$

Q) Why is a polyhedron convex?



Convexity

2. Convex Functions defined on the domain of a convex set

For $f : R^n \rightarrow R$ defined for $x \in dom(f)$: Convex set,
f is convex if $\lambda f(x_1) + (1-\lambda) f(x_2) \geq f(\lambda x_1 + (1-\lambda) x_2)$

✓ Properties of Convex Functions

1. **Pairwise Supremum of convex sets is a convex function**
2. **Nonnegative linear combination of Convex Functions is a convex function**

✓ Iff conditions for differentiable convex functions
For f which has an open domain and differentiable on $dom(f)$,

1. **First order (gradient) condition for convexity**
   f convex $\leftrightarrow f(y) \geq f(x) + \nabla f(x)^T (y-x), \forall x, y \in dom(f)$.

2. **Second order (Hessian) condition for convexity**
   f convex $\leftrightarrow \nabla^2 f(x) \succeq 0, \forall x \in dom(f)$. "Hessian is PSD".