Optimization
SC 607 IITB
by
Ankur Kulkarni

https://somphene.github.io/notes/

Notes
by
Som S. Phene

Spring 2020

A visitor remarked that the library at the newly constructed IHES was lacking.
Grothendieck replied "We don't read books, we write them."

This is an **incomplete draft**. Please send corrections, comments, pictures of bad drawings, etc. to somphene1@gmail.com.

Last updated January 25, 2020.

Contents

1 Lecture Jan 10

§1.1 Introduction

Question 1.1.1. What is Optimization?

(Answer). Choosing the best alternative out of all the given alternatives for a specific job.

§1.1.i Principle

For finite or infinite lists of alternatives in which the alternatives can be run through in finite or reasonable time, optimization is easy: can be done by direct checking of the alternatives in the list. It is the other case where things get interesting. The problems we encounter in daily life are primarily of the second case. Hence formulating techniques better than just going through every alternative to solve such complex problems are needed. This is what we plan to cover in our course.

Basic elements of any Optimization problem are :

- Quantification of alternatives using a real valued function which we call f. Convention: We would like to find the least value of f.
- Set of alternatives is called **feasible region**, which we denote by S.

In this course we mainly deal with optimization problems for which S is a subset of Euclidean Space ($S \subseteq \mathbb{R}^n; f\colon \mathbb{R}^n \to \mathbb{R}$)

Question 1.1.2. (Optimization Problem) Find an $x^* \in S$ such that

$$f(x^*) \leq f(x) \quad \forall x \in S \tag{1.1}$$

Or sometimes we don't necessarily want x^* but the value of $f(x^*)$ such that eq. (1.1) holds, Such an x^* is called **Optimal Solution** and $f(x^*)$ is called **Optimal Value**.

Example 1.1.3 (Google Maps)

Optimization Problem of finding the optimal path to take from IIT Bombay main gate to IIT Delhi main gate. Here optimal may mean distance, time or some other cost function, denoted by f, associated with the path. Google maps solves for the time optimal path constrained to given traffic rules within seconds.

Example 1.1.4 (Classroom Assignment to Courses)

With hundreds of courses running in IIT Bombay, which of the hundreds of classrooms to assign to the course with constraints of size, computers, labs, preferences of instructor and students, so as to maximize closeness to the ideal scenario for everyone becomes an optimization problem.

§1.1.ii Classical Approach

> **Question 1.1.5.** **Queen Dido's problem**, also known as **Isoperimetric problem**: Given a rope of constant length l, find the shape of maximum area that can be enclosed by the rope.

Solution: We shall present a specific way of solving the problem so as to highlight some features that will be needed later, We outline the steps as follows:

1. The optimal shape can not have a dent protruding inwards (as shown in fig. 1.1), ie. it must be convex.

   Definition 1.1.6. **(Convex Set)** A set S is said to be **convex** if

   $$\forall\, x, y \in S; \quad \forall\, \lambda \in [0,1]; \quad \lambda x + (1-\lambda) y \in S$$

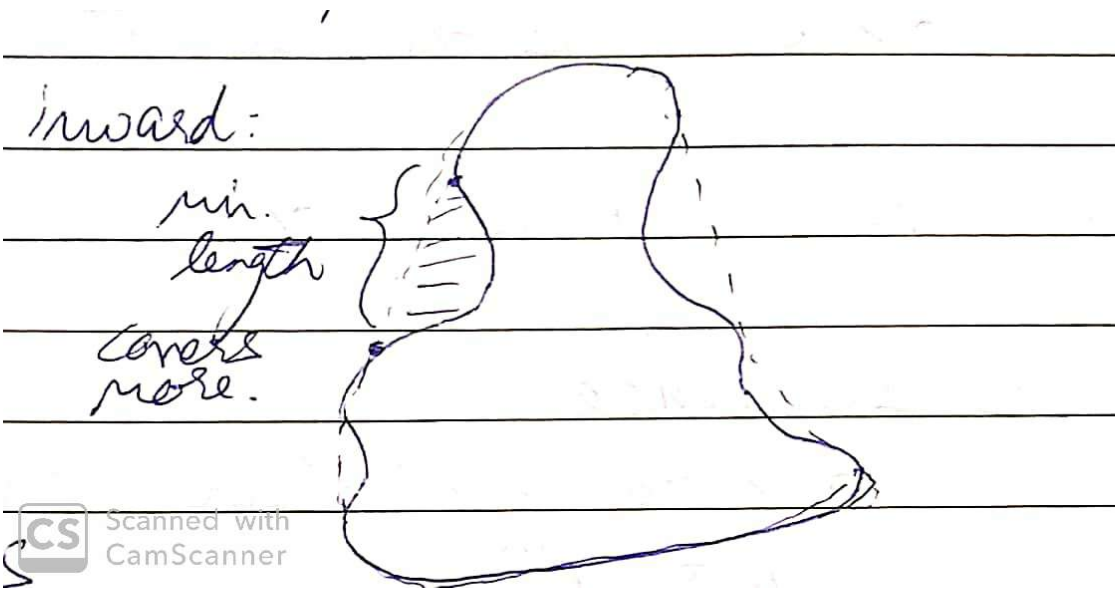

Figure 1.1: A non-convex set is not optimal.

2. **Symmetry argument**: Now that we have reduced the admissible set to convex shapes, choose a starting point on the rope. Travel a distance of $l/2$ along the rope, stop and mark the point. Draw a line segment L joining the starting point and the final point. This line L divides the enclosed shape into two parts. If the area enclosed by the rope on either of the two sides of the segment is not equal then simply reflect the side with larger area about L to increase the total area while maintaining the total length of the rope to be l (as shown in fig. 1.2). Hence the shape must be symmetric about L.

3. **Geometry of semicircle**: By the previous symmetry argument, it suffices to find the optimal shape for one side of the shape enclosed by the rope of length $l/2$ and the line segment L. Suppose that the optimal curve is traced by the rope of length $l/2$. Then fix a point on the rope and imagine it to be a hinge. The lengths of this point from the starting and final points on the rope are fixed, say a and b. Keeping the rope's $l/2$ length constant, area enclosed between the rope and line segment L can be increased by increasing the area of the inscribed triangle as shown in

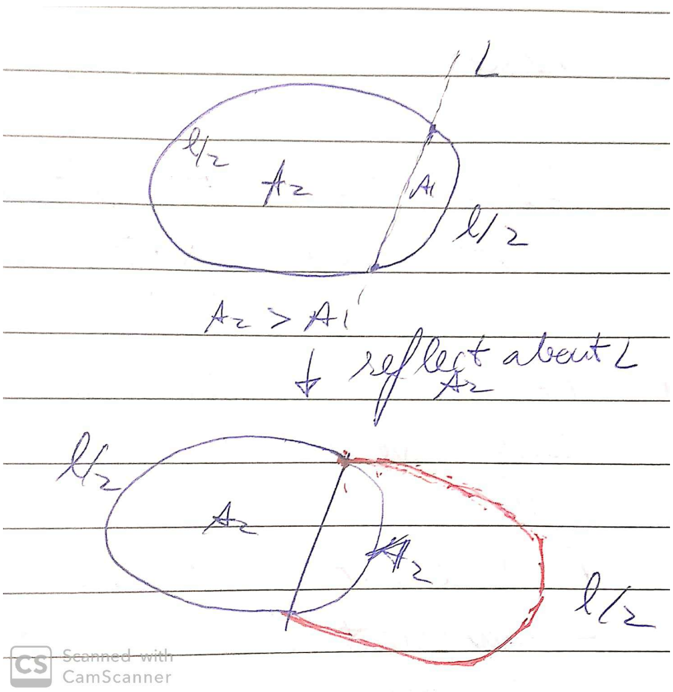

Figure 1.2: An asymmetric shape about L is not optimal.

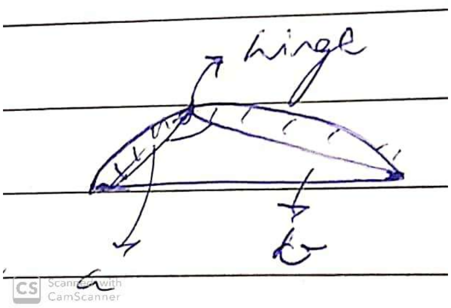

Figure 1.3: Any shape that is not the semicircle on L is not optimal.

fig. 1.3 (the shaded region has fixed area, only the inscribed triangle are can be varied). Fixing any one of a or b as the base, one may obtain the greatest area for the inscribed triangle by making the other side perpendicular to the base so as to get the maximum height (Area of $\Delta = 1/2 \times$ base $\times$ height). This achieves the maximum area of $\frac{ab}{2}$ for the inscribed triangle. Note that the angle subtended by the line segment L at the fixed hinge point is $\pi/2$ for the optimal solution. However the hinge point was arbitrarily chosen to be on the rope. Hence the line segment subtends an angle of $\pi/2$ at every point on the rope hat traces the optimal solution. This geometric property uniquely identifies that the optimal solution cannot be anything that is not a semicircle on L.

Combining the above arguments, we see that any shape that is not a **circle** cannot be the optimal solution. This approach is due to Jakob Steiner. The **logic** underlying all arguments was that there was a method by which area could be increased while keeping the perimeter lesser than or equal to the given length l.

Claimed Theorem: Of all positive integers, 1 is the largest.

Claimed Proof: Based on the same logic that was stated above, given any positive integer that is not 1, we can increase its value by some method (say by squaring). Hence proved.

Flaw in logic: We may not have an optimal solution in the first place. The above example shows that we cannot take the existence of optimal solution for granted. We must prove the existence of an optimal solution. This naturally needs a review of concepts from Analysis, here in particular that of Real Analysis, which forms the topic of our next section 1.2.

§1.1.iii Another Approach

Here, I independently give an alternative proof to the isoperimetric problem, which is only based on reflection symmetrization and hence simpler to follow.

Proof. I shall use two lemmas.

Lemma 1.1.7 (Reflection symmetrization for non-convex to convex sets)
Any S that is not convex, can be transformed to a convex set with increase in the area enclosed while keeping its perimeter constant.

Proof of Lemma by Contradiction. Denote the region enclosed in the rope by S. Suppose that S is not convex. Then there must exist at least one pair of points $x, y \in S$ such that the line segment L in not entirely inside S, ie. $\lambda x + (1-\lambda)y \notin S$, $\forall\, \lambda \in [0, 1]$. In case the set of points on the segment L not in S are isolated, they have measure zero and the area enclosed is unaffected by their addition, hence these points can be added to make S convex. In case the set of points on $L \notin S$ are not isolated, the measure (one-dimensional) may be positive and hence we cannot simply add them as the perimeter could increase. In this case, take the minimal union of intervals on L which cover these points to give subsegments of L. For each such subsegment, reflect the points on the rope about that subsegment as shown in fig. 1.4. Reflection in the plane being an isometry, keeps the length of the points on the rope and hence that of the entire rope fixed, but the reflected part of the rope encloses more area.

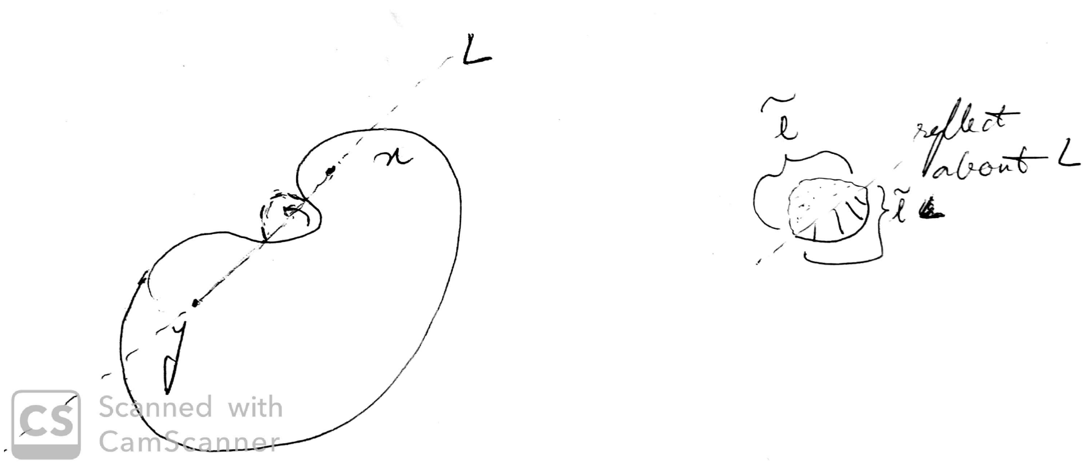

Figure 1.4: Any shape that is not the semicircle on L is not optimal.

□

Lemma 1.1.8 (Reflection Symmetrization of Area)

Given any two points on the rope, dividing it into arcs of equal length, the line segment joining them must also divide the optimal S into two parts of equal area.

Proof. Suppose not, then reflect part of greater area in the line segment to increase the bounded area. This keeps the perimeter constant as reflection in the plane is an isometry. Clearly then, the S we started with was not optimal. □

Suppose we assume that the optimal solution exists, lemma 1.1.7 gives us that the optimal S must be convex, in particular it is connected and lemma 1.1.8 gives us that the optimal S must have a reflection axis of symmetry in every direction. This uniquely identifies the optimal S as the region in the plane enclosed by a circle. □

Exercise 1.1.9. In the above proof, a subtle result was used: Show that the segments found by joining two points on the rope that divide it into arcs of equal length, are in bijective correspondence with directions in the plane.

Exercise 1.1.10. Show that there exists an optimal solution to the isoperimetric problem.

Remark 1.1.11 (Method of Moving Planes) — The technique I used in the above proof can be generalized to give what is called the **Method of Moving planes** [Ca].

§1.2 Review of Real Analysis

Definition 1.2.1. (**Sequences in $\mathbb{R}^n$**) Function $f\colon \mathbb{N} \to \mathbb{R}^n$ defines a collection of points in $\mathbb{R}^n$ which is called a sequence and denoted by $\{x_k\}_{k\in\mathbb{N}} \subseteq \mathbb{R}^n$.

Definition 1.2.2. (**Bounded Sequences**) We say $\{x_k\}_{k\in\mathbb{N}}$ is bounded if

$$\exists\, M > 0 \quad \text{such that} \quad \|x_k\| \leq M \quad \forall\, k \in \mathbb{N}$$

Example 1.2.3 (Unbounded sequence)

$x_k = k^2$ is not bounded.

Example 1.2.4 (Bounded but not convergent sequence)

$x_k = (-1)^k$ is bounded. Note that even though it is bounded, it is oscillating. There is no single value which the sequence eventually attains or settles towards.

Definition 1.2.5 (Convergence)**.** We say that $\{x_k\}_{k\in\mathbb{N}}$ **converges to x^*** if

$$\forall\, \epsilon > 0; \quad \exists\, n_0 \in \mathbb{N}; \quad \text{such that } \|x_k - x*\| < \epsilon \quad \forall\, k > n_0$$

x^* is called the **limit** of the sequence.

Proposition 1.2.6

$x_k = \frac{1}{k}$ converges to 0.

Proof. Given any $\epsilon > 0$, take $n_0 = \lceil \frac{1}{\epsilon} \rceil + 1$ so that $x_{n_0} - 0 < \epsilon$. $\{x_k\}_{k\in\mathbb{N}}$ being a monotone decreasing sequence, satisfies the criterion for convergence to 0. □

Proposition 1.2.7

If the limit of a sequence exists, it is unique.

Proof. Suppose $\{x_k\}_{k\in\mathbb{N}}$ has two distinct limits l_1 and l_2. Straightforward application of Triangle Inequality gives

$$\|l_1 - l_2\| \leq \|l_1 - x_k\| + \|x_k - l_2\|$$

Since $\{x_k\}_{k\in\mathbb{N}}$ converges to both l_1 and l_2, both the terms on the right hand side are arbitrarily close to 0 for sufficiently large k, ie. For any small $\epsilon > 0$

$$\|l_1 - l_2\| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad \text{for} \quad k > n_0$$

$$\implies \|l_1 - l_2\| = 0$$

By definition of metric, $l_1 = l_2$ □

Definition 1.2.8. A sequence $\{x_k\}_{k\in\mathbb{N}}$ **does not converge** if

$$\forall\, x^* \in \mathbb{R}^n,\, \exists\, \epsilon > 0; \quad \forall\, n_0 \in \mathbb{N}; \quad \|x_k - x*\| \geq \epsilon \quad \text{for some} \quad k > n_0$$

The above definition can be used to generate a point $x_{k_{n_0}}$ by taking for each n_0, the smallest k for which the condition in the definition holds. This is itself a sequence $\{x_{k_{n_0}}\}_{n_0\in\mathbb{N}}$. This gives rise to the notion of a subsequence.

Definition 1.2.9. A **subsequence** $\{x_{k_i}\}_{i\in\mathbb{N}}$ of $\{x_k\}_{k\in\mathbb{N}}$ is a sub-collection of points x_{k_i} from the original sequence such that $k_{i+1} > k_i$ is satisfied.

For more details of the notions defined in the above review of Real Analysis, there are many standard references such as [Ru+64] or [Pu02].

Bibliography

[**Ca**] Xavier Cabré. "Elliptic PDE's in probability and geometry: Symmetry and regularity of solutions". In: *Discrete & Continuous Dynamical Systems-A* 20.3 (), p. 425.

[**Pu02**] C. C. Pugh. *Real mathematical analysis*. New York: Springer, 2002. ISBN: 978-0387952970.

[**Ru+64**] Walter Rudin et al. *Principles of mathematical analysis*. Vol. 3. McGraw-hill New York, 1964.