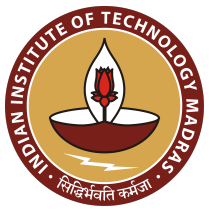
Constrained Optimization

Saurav Samantaray

Department of Mathematics

Indian Institute of Technology Madras

October 21, 2026



Motivation

Manufacturing

- Suppose we have m different materials; for each material i we have s_i units of it available in stock.
- k different products can be manufactures; product j gives profit p_j and while using up c_{ij} amount of material i to make.
- To maximize profits, we can solve the following optimization problem for the total amount x_j we should manufacture of each item j:

$$\max_{x \in \mathbb{R}^n} \sum_{j=1}^{k} p_j x_j$$

$$\text{such that } x_j \geq 0 \ \forall \ j \in \{1, 2, \ldots, k\} \tag{1}$$

$$\sum_{j=1}^{k} c_{ij} x_j \leq s_i, \ \forall \ i \in \{1, 2, \ldots, m\}$$

- The first constraint ensures that we do not make negative numbers of any product,
- and the second ensures that we do not use more than our stock of each material.

Constrained Problem

A general formulation of these problems is:

$$\min_{x \in \mathbb{R}^n} f(x) \qquad \text{subject to} \qquad \begin{cases} c_i(x) & = 0, \quad i \in \mathscr{E} \\ c_j(x) & \geq 0, \quad j \in \mathscr{I} \end{cases} \tag{2}$$

f and c_i are scalar valued functions of the vector of unknowns x and $\mathscr{E}$ and $\mathscr{I}$ are set of indices.

- x is a vector of variables, also called unknown or parameters;
- f is the objective function, a function of x that we want to optimise (minimise or maximise);
- c is the vector function of constraints that must be satisfied by the unknowns x.
- c_i, $i \in \mathscr{E}$ are the equality constraints.
- c_i, $i \in \mathscr{I}$ are the inequality constraints.

Compact form of Constrained Problem

Definition

Define the feasible set Ω to be the set of points x that satisfy the constraints; that is,

$$\Omega = \{x \mid c_i(x) = 0, \quad i \in \mathscr{E}; \quad c_i(x) \geq 0, \quad i \in \mathscr{I}\}, \tag{3}$$

Now (2) can be rewritten more compactly as:

Constrained Problem

$$\min_{x \in \Omega} f(x). \tag{4}$$

Characterizations of the Solutions

- For the unconstrained optimization problems the solution point x^* was characterised in the following way:
- Necessary conditions: Local minima of unconstrained problems have

$$\nabla f(x^*) = 0$$

and,

$$\nabla^2 f(x^*) \quad \text{is positive semidefinite}$$

- Sufficient conditions: Any point x^* at which $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite is a strong local minimiser of f.

LOCAL AND GLOBAL SOLUTIONS

- We have seen already that global solutions are difficult to find even when there are no constraints.
- The situation may improve when we add constraints.
- The feasible set might exclude many of the local minima.
- It might be comparatively easy to pick the global minimum from those that remain.

LOCAL AND GLOBAL SOLUTIONS

- Consider the problem

$$\min_{x \in \mathbb{R}^n} ||x||_2^2, \quad \text{subject to } ||x||_2^2 \geq 1. \tag{5}$$

- Without the constraint, this is a convex quadratic problem with unique minimiser $x = 0$.
- When the constraint is added, any vector x with $||x|| = 1$ solves the problem.
- There are infinitely many such vectors (hence, infinitely many local minima) whenever $n \geq 2$

LOCAL AND GLOBAL SOLUTIONS

- Addition of a constraint produces a large number of local solutions that do not form a connected set.
- Consider

$$\min_{x \in \mathbb{R}^2} (x_2 + 100)^2 + 0.01 x_1^2, \quad \text{subject to } x_2 - \cos x_1 \geq 0, \quad (6)$$

- Without the constraint, the problem has the unique solution $(-100, 0)$.
- With the constraint there are local solutions near the points

$$(x_1, x_2) = (k\pi, -1), \quad \text{for } k = \pm 1, \pm 3, \pm 5, \ldots$$

LOCAL AND GLOBAL SOLUTIONS

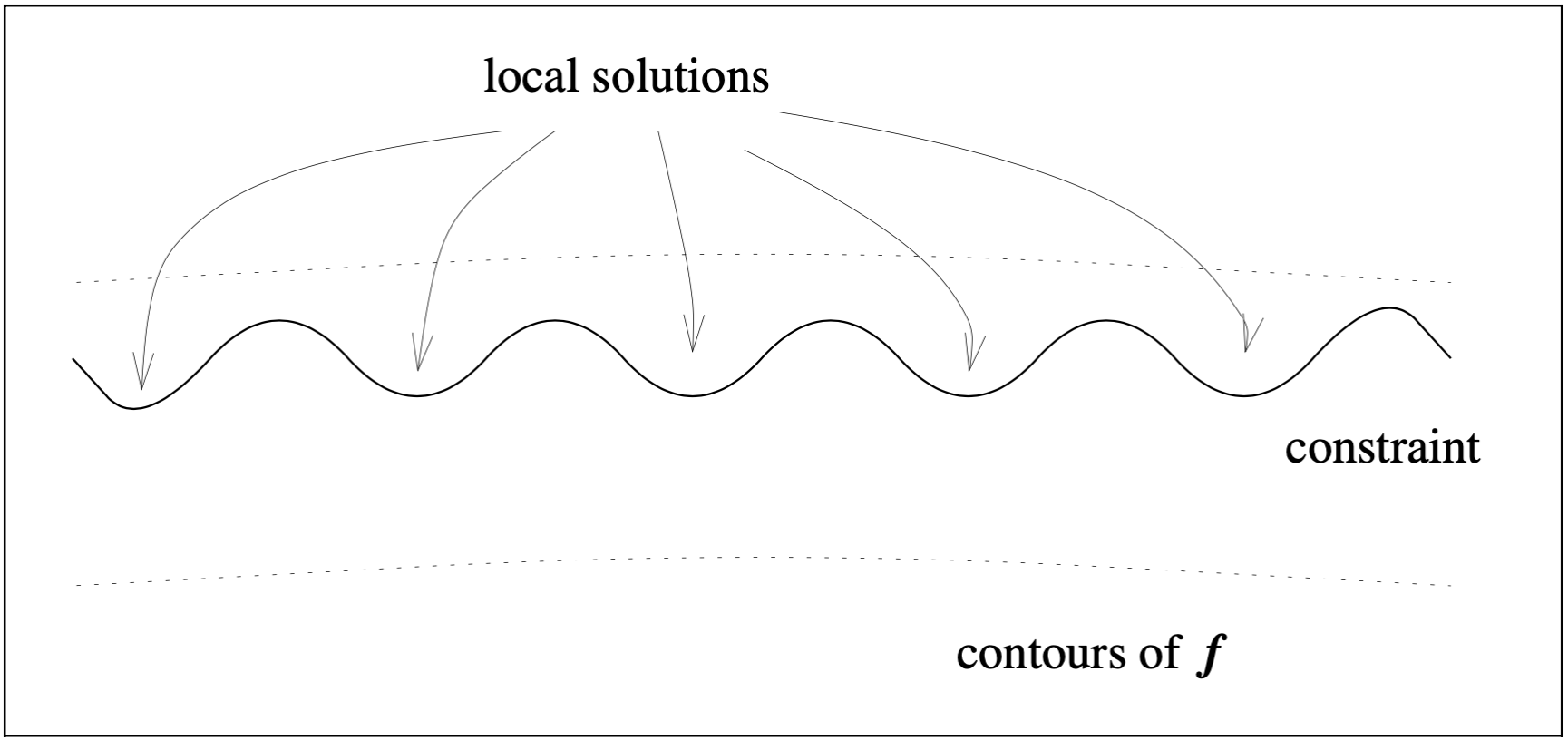


Figure Constrained problem with many isolated local solutions.

LOCAL AND GLOBAL SOLUTIONS

- Local and global solutions are defined in a very similar fashion as they were for the unconstrained case.
- The new caveat that comes into action in the definitions for the constrained case is the inclusion of constraints leading to a restriction imposed via a feasible set (space).

Definition

A vector x^* is a local solution of the constrained minimisation problem (4) if $x^* \in \Omega$ and there exists a neighbourhood $\mathcal{N}$ of x^* such that

$$f(x^*) \leq f(x) \quad \text{for all } x \in \Omega \cap \mathcal{N}$$

LOCAL AND GLOBAL SOLUTIONS

Definition

A vector x^* is called a strict local solution (also called a strong local solution) if $x^* \in \Omega$ and there is a neighbourhood $\mathcal{N}$ of x^* such that

$$f(x^*) < f(x) \quad \text{for all } x \in \mathcal{N} \cap \Omega \quad \text{with } x \neq x^*$$

Definition

A point x^* is an isolated local solution if $x^* \in \Omega$ and there is a neighbourhood $\mathcal{N}$ of x^* such that x^* is the only local minimiser in $N \cap \Omega$.

Smoothness

- Smoothness of objective functions and constraints is an important issue in characterizing solutions.
- Just as in the unconstrained case, it ensures that the objective function and the constraints all behave in a reasonably predictable way.
- Allows algorithms to make good choices for search directions.
- Non-smooth functions contain “kinks” or “jumps” where the smoothness breaks down.
- The feasible region for any given constrained optimization problem usually contains many kinks and sharp edges.

Smoothness

- Does this mean that the constraint functions that describe these regions are non-smooth?

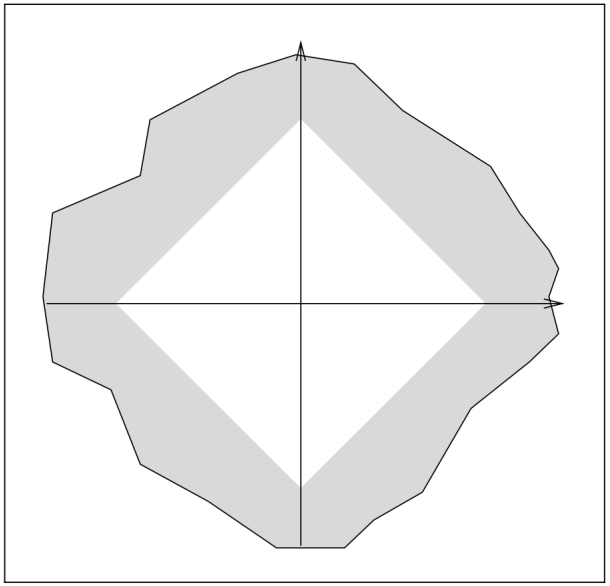


Figure: A feasible region with a non-smooth boundary can be described by smooth constraints.

- The answer is often no, because the non-smooth boundaries can often be described by a collection of smooth constraint functions.

Smoothness

- The figure above shows a diamond-shaped feasible region in $\mathbb{R}^2$.
- It could be described by the single non-smooth constraint

$$||x||_1 = |x_1| + |x_2| \leq 1.$$

- Or, it could also be brought out as an intersection of four smooth (in fact, linear) constraints:

$$x_1 + x_2 \leq 1, \quad x_1 - x_2 \leq 1, \quad -x_1 + x_2 \leq 1, \quad -x_1 - x_2 \leq 1.$$

- Each of the four constraints represents one edge of the feasible polytope.
- The constraint functions are chosen so that each one represents a smooth piece of the boundary of Ω.

Smoothness

- In general, the constraint functions are chosen so that each one represents a smooth piece of the boundary of Ω.
- Non-smooth, unconstrained optimization problems can sometimes be reformulated as smooth constrained problems.
- Consider the unconstrained scalar problem of minimizing a non-smooth function $f(x)$ defined by,

$$f(x) = \max(x^2, x)$$

- It has kinks at $x = 0$ and $x = 1$.
- The solution at $x^* = 0$.
- A smooth, constrained formulation of this problem can be obtained by adding an artificial variable t and writing,

Smoothness

$$\min\ t, \quad \text{s.t}, \quad t \geq x, \quad t \geq x^2.$$

- In the examples above we expressed inequality constraints in a slightly different way from the form $c_i(x) \geq 0$.
- However, any collection of inequality constraints with $\geq$ or $\leq$ and nonzero right-hand-sides can be expressed in the form $c_i(x) \geq 0$ by simple rearrangement of the inequality.
- $$t - x \geq 0, \quad t - x^2 \geq 0.$$

Convexity

Convex Function

A function $f : \Omega \subset \mathrm{R}^n \to \mathrm{R}$ is convex if

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y), \tag{7}$$

for all $x, y \in \Omega$ and Ω is a domain and $\alpha \in [0, 1]$.

Necessary and Sufficient Condition

For a twice-differentiable function, a sufficient and necessary condition for convexity on a convex domain is

$$\nabla^2 f(x) \geq 0 \tag{8}$$

That is, the Hessian is positive semidefinite everywhere on the domain.

Convex Set

A set is convex if the line segment joining any two points remains in the set.

The penalty method

Penalty Method for Equality Constraint

Consider the constrained optimization problem

$$\begin{aligned} &\min_{x \in \mathbb{R}^n} f(x) \\ &\text{subject to} \quad h(x) = 0 \end{aligned} \tag{9}$$

- we cannot freely choose x: the solution must lie on the constraint set $h(x) = 0$
- replaces this constrained problem with an unconstrained one by modifying the objective function:

$$F_\mu(x) = f(x) + \frac{\mu}{2} h(x)^2, \quad \mu > 0 \tag{10}$$

- $h(x)^2$ measures the violation of the equality constraint.
- μ controls how strongly constraint violations are penalized.

The penalty method

- The new objective has two competing goals:
  1. Minimize the original objective $f(x)$.
  2. Keep the constraint violation $h(x)^2$ small.
- When $h(x) = 0$, the penalty term vanishes.
- When $h(x) \neq 0$, the objective increases.
- Instead of forcing the solution to satisfy the constraint exactly, we make violating the constraint increasingly expensive.

Example

Consider

$$\begin{aligned} \min_{x,y} \quad & f(x,y) = (x-2)^2 + (y-1)^2, \\ \text{subject to} \quad & x + y = 1. \end{aligned}$$

The penalty formulation is

$$\boxed{F_\mu(x,y) = (x-2)^2 + (y-1)^2 + \frac{\mu}{2}(x+y-1)^2.}$$

- Small penalty parameter:
  - The algorithm prioritizes minimizing the original objective.
  - It may accept a substantial constraint violation to remain close to the unconstrained minimizer.
- Large penalty parameter
  - Violating the constraint becomes expensive.
  - The minimizer moves closer to the line $x + y = 1$, even if this increases the original objective.