Unconstrained Optimisation

Saurav Samantaray

Department of Mathematics

Indian Institute of Technology Madras

January 21, 2024

Unconstrained Optimisation

- Minimise an objective function that depends on real variables.
- No restriction on the values of these variables (no constraints).

Mathematical Formulation:

$$\min_{x} f(x) \tag{1}$$

$$\text{where, } x \in \mathbb{R}^n, n \geq 1.$$

$$f : \mathbb{R}^n \to \mathbb{R} \quad \text{is smooth}$$

In a real world scenario

- The objective function "f" might not be known globally everywhere.
- Ideally, may have finitely many values of "f" or some derivatives of "f".
- Any information for "f" at arbitrary points usually do-not come very cheaply.
- Therefore, one should prefer for algorithms which do-not demand the same, unnecessarily.

Example

- Suppose we are trying to find a curve that fits some experimental data.
- (t_i, y_i), y_i signal is measured at time t_i.
- Let's assume based on the knowledge of the phenomenon under study we have the understanding that the signal has exponential and oscillatory behaviour of certain types.

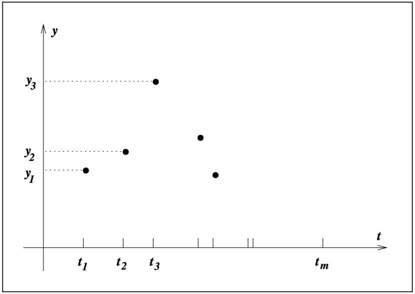

Figure: Least squares data fitting problem.

Example

- Choose the model function as

$$\phi(t,x) = x_1 + x_2 e^{-(x_3-t)^2/x_4} + x_5 \cos(x_6 t)$$

  where x_i's are the parameters of the model.
- What we want is the model should fit the observed data y_j, as closely as possible.
- Let $x = (x_1, x_2, x_3, x_4, x_5, x_6)$,
  We define the residual for each y_j as

$$r_j = y_j - \phi(t,x), \qquad j = 1, \ldots, m.$$

- We define the objective function as

$$\min_{x \in \mathbb{R}^6} f(x) = v_1^2(x) + \ldots + r_m^2(x)$$

  This is a non-linear least square problem, a special case of unconstrained optimisation.

Example

- Note that the equation of the objective function appears quite expensive even for small number of variables

$$n = 6$$

- say, if the no. of measurements i.e. $m = 10^5$, then the evaluation of f becomes quite a computational expense.

Lets Gain Some Perspective!!

- Suppose for a given set of data the optimal solution to the previous problem is approximately

$$x^* = (1.1, 0.01, 1.2, 1.5, 2.0, 1.5)$$

  and the corresponding function value is $f(x^*) = 0.34$.
- As at the optimal point the objective is non-zero there must be some discrepancy between the function values and the observations made.

Some Perspective

- i.e. y_j and $\phi(t_j, x^*)$ aren't the same for some or many $(y_j, t_j) \longleftrightarrow \phi(t_j, x^*)$
- The model hasn't produced al the data points correctly as

$$f(x^*) \neq 0$$

- Then how to know x^* is indeed a minimiser of f?
- In the sense that how to know which all points one should go close to or not?
- To answer this question, we need to define the term "solution and explain how to recognise solutions.

What is a solution?

A point x^* is a global minimiser of f if

$$f(x^*) \leq f(x) \qquad \forall\, x \in \mathbb{R}^n$$

or in the domain of interest.

- It would be the most ideal scenario if we could find a global minimiser.
- It might be difficult to get a global minimiser, owing to the limited (or local) knowledge of f.
- Most algorithms are only able to find a local minimiser.

What is a solution?

Local Minimiser

A point x^* is called a local minimiser, if there is a neighbourhood $\mathcal{N}$ of x^* such that

$$f(x^*) \leq f(x) \quad \forall\ x\mathcal{N}$$

- It's a points that achieves the smallest value of f in its neighbourhood.

Weak Local Minimiser	Strict (Strong) Local Minimiser
$f(x^*) \leq f(x) \qquad x \in \mathcal{N}$	$f(x^*) < f(x) \qquad x \in \mathcal{N},\ x \neq x^*$

Example

For a constant function $f(x) = 2$ every point is a weak local minimiser.
For $f(x) = (x-2)^4$, $x = 2$ is a strict local minimiser.

Isolated Local Minimiser

A point x^* is called an isolated local minimiser if there is a neighbourhood $\mathscr{N}$ of x^* such that x^* is the only local minimiser in $\mathscr{N}$.

Example

$$f(x) = x^4 \cos(1/x) + 2x^4 \quad f(0) = 0$$

- is twice continuously differentiable
- has a strict local minimiser at $x^* = 0$
- however, there are strict local minimisers at many nearby points x_j,
  and $x_j \to 0$ as $j \to \infty$

- Some strict local minimisers are not isolated
- All isolated local minimisers are strict
- It is often difficult to determine a global minimiser for an algorithm, as it often gets trapped in a locality (at a local minimiser)

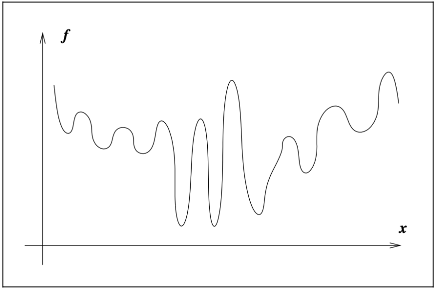

Figure: Showcases a function with many local minimisers.

How to detect minimisers?

- The simplest test being

$$f'(x^*) = 0$$

  is very insufficient to speak anything about the globality of the minimiser.
- These cases (having a lot of local minimisers) is quite standard for optimisation problems.
- Global knowledge about a function f may help identify global minima.
- For convex functions local minimiser is also a global minimiser.