SUBGRADIENT ALGORITHM

Author
Vinitra Muralikrishnan
11th March, 2024

Contents

1 Overview

Definition 1.1. This is an algorithm to minimize a non-differentiable function

It applies on:

- non differentiable f
- step lengths or sizes are not searched by backtracking line search. They are fixed beforehand
- Unlike ordinary gradient descent methods, subgradient method is not a descent method. The function value can increase or decrease

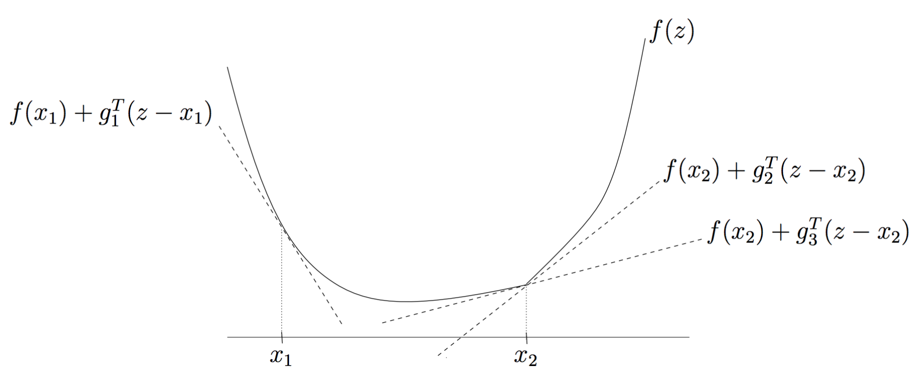
Definition 1.2. The subgradient of a convex function f at x is any $g \in \mathbb{R}^n$ s.t.

$$f(y) \geq f(x) + g^T(y - x), \ \forall y$$

The above definition stands for:

- always exists for convex f
- for f differentiable at x, then $g = \nabla f(x)$
- difficut to determine all subgradients at a point
- Also applicable for non-convex f

First order global underestimator:



Remark. max of 2 functions f_1, f_2 is non-convex whether or not they are convex

Definition 1.3. The subdifferential is the set of all subgrdients of a convex f at x

$$\partial f(x) = \{g \in \mathbb{R}^n : g \text{ is a subgradient of } f \text{ at } x\}$$

The above definition implies:

- non empty convex f
- $\partial f(x)$ is closed and convex, even for non convex f
- f is differentiable at $x \Leftrightarrow \partial f(x) = \{\nabla f(x)\}$

Remark. Subdifferential is a set while subgradient of a point is a vector.

Remark. Subgradient algorithm is not a descent method. It can move in any direction and keeps track of the best value so far.

2 References

- Stanford Lecture Notes