Gradient Descent

$x, y \in \mathbb{R}^n$

f: convex, ρ lipschitz $\longrightarrow$ $|f(x) - f(y)| \leq \rho \|x - y\|$

$x^{t+1} = x^t - \eta \nabla^T f(x^t)$ output $\bar{x} = \frac{1}{T} \sum x^t$

$x^* = \operatorname{argmin} f(x)$

Subgradients

Recall: For convex C^1 f, $f(x) \geq f(y) + \nabla^T f(y)(x - y)$ $\forall x, y \in \mathbb{R}^n$

Def^n: For $x \in \mathbb{R}^n$, $\partial f(x) := \{v \in \mathbb{R}^n : f(y) \geq f(x) + v^T(y - x) \ \forall y \in \mathbb{R}^n\}$

called differential set. Any $v \in \partial f(x)$ called subgradient @ x

Subgradient Bound

Lemma 1: f convex. f ρ-Lipschitz $\iff$ $\forall x \in \mathbb{R}^n$ $\forall v \in \partial f(x)$ $\|v\| \leq \rho$

Pf $\Leftarrow$ $v \in \partial f(x)$, $\|v\| \leq \rho$

$$f(x) - f(y) \leq v^T(x - y) \overset{CS}{\leq} \|v\| \|x - y\| \leq \rho \|x - y\|$$

$\Rightarrow$ f is ρ-lipschitz

Let $x \in \mathbb{R}^n$, $v \in \partial f(x)$

Let $y = x + \epsilon \frac{v}{\|v\|}$

$$\rho\epsilon = \rho\|y - x\| \geq f(y) - f(x) \geq v^T(y - x) = \epsilon\|v\|$$

$\Rightarrow$ $\|v\| \leq \rho$