Preliminary – Convex optimization

Basic definitions

Convex set: A set $K \subseteq \mathbb{R}^n$ is said to be convex if for every pair of points $x, y \in K$ and for every $\lambda \in [0,1]$, we have

$$\lambda x + (1-\lambda) y \in K.$$

Convexity: A function $f: K \to \mathbb{R}$, defined over a convex set K, is convex if for all $x, y \in K$ and $\lambda \in [0,1]$, we have

$$f(\lambda x + (1-\lambda) y) \leq \lambda f(x) + (1-\lambda) f(y).$$

Properties of convex sets and convex functions

Necessary and sufficant conditions for convexity

First-order characterization of convexity

A differentiable function $f: K \to \mathbb{R}$ over a convex set K is convex if and only if

$$\forall x, y \in K: \ f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle.$$

Second-order notion of convexity

Suppose K is convex and open. If $f: K \to \mathbb{R}$ is twice continuously differentiable, then it is convex if and only if

$$\forall x \in K: \ \nabla^2 f(x) \succeq 0.$$

Subgradients

For a convex function f defined over a convex set K, a vector v is said to be a subgradient of f at a point $x \in K$ if for any $y \in K$

$$f(y) \geq f(x) + \langle v, y - x \rangle.$$

* Although a convex function can be non-differentiable, at least one subgradient must exist in all points in the interior of the function domain.

Separating and supporting hyperplanes

Suppose $K \subseteq \mathbb{R}^n$ is convex. For $h \in \mathbb{R}^n$ and $c \in \mathbb{R}$, the hyperplane

$$H = \{ x \in \mathbb{R}^n : \langle h, x \rangle = c \}$$

is said to separate $y \in \mathbb{R}^n$ from K if $\langle h, x \rangle \leq c$ for all $x \in K$, but $\langle h, y \rangle > c$.

Further, if $y \in \partial K$ and $\langle h, y\rangle = c$, then H is said to be a supporting hyperplane for K, where ∂K is the boundary of K.

Separating and supporting hyperplane theorem

Let $K \subseteq \mathbb{R}^n$ be a nonempty, closed, and convex set. Then, given $y \in \mathbb{R}^n \setminus K$, there is an $h \in \mathbb{R}^n \setminus \{0\}$ such that

$$\forall x \in K: \ \langle h, x\rangle < \langle h, y\rangle$$

if $y \in \partial K$, then there exists a supporting hyperplane containin

Algorithmic results for convex optimization

1st order oracle: A first order oracle for a function $f: \mathbb{R}^n \to \mathbb{R}$ is a primitive that given $x \in \mathbb{Q}^n$ outputs the value $f(x) \in \mathbb{Q}$ and a vector $g(x) \in \mathbb{Q}^n$ such that

$$\forall z \in \mathbb{R}^n: \ f(z) \geq f(x) + \langle z - x, g(x)\rangle$$

note that $g(x)$ can be any of the subgradient of f at the point x.

Separation oracle: For any non-empty, closed and convex set $K \subseteq \mathbb{R}^n$, a separation oracle is a primitive that given any point $x \in \mathbb{R}^n$, either asserting $x \in K$ or returning a hyperplane separating x and K.

* The separating hyperplane theorem ensures the existence of such a separation oracle.

Efficient algorithm for convex optimization

Given

1. a first-order orcale for a convex function $f: \mathbb{R}^n \to \mathbb{R}$
2. a separation orcale for a convex set $K \subseteq \mathbb{R}^n$
3. numbers $r > 0$ and $R > 0$ such that $K \subseteq B(0, R)$ where $B(0,R)$ is the Euclidean ball of radius R around point 0 and K contains a Euclidean ball of radius r.
4. bounds l and u such that for all $x \in K$, $l \leq f(x) \leq u$
5. $\varepsilon > 0$

Outputs a point $\hat{x} \in K$ such that

$$f(\hat{x}) \leq f(x^*) + \varepsilon$$

where x^* is any minimizer of f over K. The running time of the algorithm is

$$O\left((n^2 + T_K + T_f) \cdot n^2 \cdot \log\left(\frac{R}{r} \cdot \frac{u_0 - l_0}{\varepsilon}\right)^2\right)$$

where T_K and T_f are the running times for the separation oracle for

$\mathcal{K}$ and the first-order oracle for f, respectively.