10-425/625: Introduction to Convex Optimization **(Fall 2023)**

Lecture 2: Overview of Opt. / Convex Sets

Instructor:[1] *Matt Gormley* *August 30, 2023*

2.1 Optimization of Optimization (continued)

2.1.1 Convex Optimization Problems – Standard Form

For now it is worth noting (and re-visiting once the definitions are in place), that the explicit constraints define a convex set, and their intersection with the domain $\mathcal{D}$ is also a convex set. If we denote this convex set $\mathcal{C}$ then our convex optimization problem can be equivalently, succinctly described as:

$$\min_{x \in \mathcal{C}} f_0(x),$$

i.e. a *convex optimization problem* is simply the problem of minimizing a *convex function* over a *convex set.*

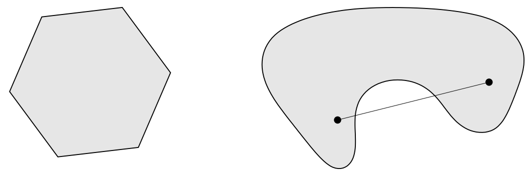
2.1.2 The Key Feature of Convex Optimization Problems

The most important structural feature of convex optimization problems is that *every local minima is a global minima.* This in turn makes local search algorithms effective for convex optimization.

We'll need to define some things in order to make sense of this claim. First, lets briefly define convex sets and functions:

Definition 2.1 (**Convex Set**). *A set C is convex, if for every $x_1, x_2 \in C$ and $0 \leq \theta \leq 1$ we have that, $\theta x_1 + (1 - \theta)x_2 \in C$.*

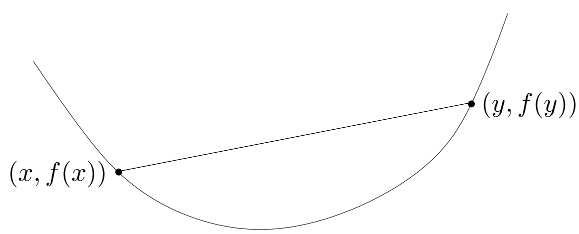
[1]These notes were originally written by Siva Balakrishnan for 10-725 Spring 2023 (original version: here) and were edited and adapted for 10-425/625.



Definition 2.2 (**Convex Function**). *A function* $f : \mathbb{R}^d \mapsto \mathbb{R}$ *is a convex function if,*

1. *dom*(f) *is a convex set,*
2. *for every* $x, y \in$ *dom*(f)*, and* $0 \leq \theta \leq 1$ *we have that,*

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



Definition 2.3 (**Nonconvex Function**). *A function* f *is said to be non-convex if it is* not *convex.*

Next, we'll need to understand what local optima are:

Definition 2.4 (Local & Global Optima). *A point* x *is a local optima, if* x *is feasible, and minimizes* f_0 *in a local neighborhood, i.e. for some* $\rho > 0$*,*

$$f_0(x) \leq f_0(y),$$

for all y *which are feasible, and* $\|x - y\|_2 \leq \rho$*. A point* x^* *is a global optima, if* x^* *is feasible and*

$$f_0(x) \leq f_0(y),$$

for all y *which are feasible.*

Theorem 2.5. *For a convex optimization problem any local optima is a global optima.*

Proof: Let x be a local optima. Suppose for contradiction of global optimality, that there is some x^* which is feasible, and has the property that,

$$f_0(x^*) < f_0(x).$$

Now, lets examine a new point,

$$x_0 = \left(1 - \frac{\rho}{\|x - x^*\|_2}\right) x + \frac{\rho}{\|x - x^*\|_2} x^*.$$

Notice that,

1. x_0 is feasible, since it is a convex combination of two feasible points x and x^*, and the set of feasible points is a convex set.
2. It is within a ρ-neighborhood of the local optima x, i.e.

$$\|x - x_0\|_2 = \frac{\rho}{\|x - x^*\|_2} \|x - x^*\|_2 = \rho.$$

3. Finally, observe that the objective value at x_0 by using the convexity of f_0 can be upper bounded as,

$$\begin{aligned} f_0(x_0) &\leq \left(1 - \frac{\rho}{\|x - x^*\|_2}\right) f_0(x) + \frac{\rho}{\|x - x^*\|_2} f_0(x^*) \\ &= f_0(x) + \frac{\rho}{\|x - x^*\|_2}(f_0(x^*) - f_0(x)) < f_0(x), \end{aligned}$$

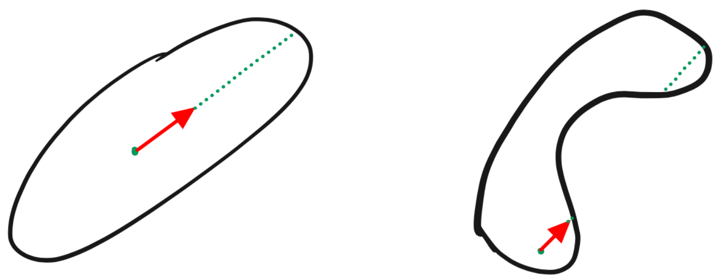
since $f_0(x^*) < f_0(x)$. However, since x_0 is in the ρ-neighborhood of x, this final claim contradicts the local optimality of x.

As a consequence we see that there cannot be any feasible x^* which satisfies $f_0(x^*) < f_0(x)$.

■

2.2 Convex Sets

We have already defined convex sets so let us briefly reflect on why they are so important in optimization. Here is picture you should have in your head, suppose we are optimizing some function over a set C and the function is simple (linear) and takes smaller values in the direction of the arrow. In case the domain is convex, we can follow the "good direction" and when we hit a "wall" declare that we're done. If it's not convex, we have a problem – there could be some "juicy" points (with much better objective value) somewhere "across the wall", and there is no easy way to optimize.



2.2.1 Examples of Convex Sets

Our next goal will be to describe some examples.

1. **Convex Hull:** For a given collection of points $x_1, \ldots, x_k \in \mathbb{R}^k$, a convex combination of the points is a linear combination,

$$\theta_1 x_1 + \ldots + \theta_k x_k,$$

with $\theta_i \geq 0$, and $\sum_{i=1}^k \theta_i = 1$. For a set C, the *convex hull* $\text{conv}(C)$ is the set of all convex combinations of elements of C. That is, $\text{conv}\{x_1, \ldots, x_k\} = \{\sum_{i=1}^k \theta_i x_i : 0 \leq \theta_i \leq 1, \sum_{i=1}^k \theta_i = 1\}$. This is always a convex set (and is the smallest convex set that contains C).

Many more examples (in each case, would be a good exercise to figure out how you would verify convexity):

2. Trivial ones: **empty set, point, line**
3. **Norm ball:** $\{x : \|x\| \leq r\}$, for any given norm $\|\cdot\|$ and radius $r \geq 0$.

Background: (Norms) When defined over real vector space, a norm is a function $g : \mathbb{R}^n \to \mathbb{R}$ that satisfies for all $x, y \in \mathbb{R}^n$ and $c \in \mathbb{R}$:

(a) The triangle inequality: $g(x+y) \leq g(x) + g(y)$

(b) Absolute homogeneity: $g(cx) = |c|g(x)$

(c) Positive definiteness: if $g(x) = 0$ then $x = 0$

Examples include: the absolute value function $|x|$ for $x \in \mathbb{R}$; the ℓ_1-norm $||x||_1 = \sum_{i=1}^n |x_i|$ for $x \in \mathbb{R}^n$; the ℓ_2-norm $||x||_2 = \sqrt{\sum_{i=1}^n x_i^2}$.

4. **Hyperplane:** $\{x : a^T x = b\}$ for a given a, b.

5. **Halfspace:** $\{x : a^T x \leq b\}$ for a given a, b. Note that halfspaces are fundamental convex sets. We will think about them in more detail when discussing the separating and supporting hyperplane theorems. They are also at the heart of convex duality.

6. **Affine space:** $\{x : Ax = b\}$, for given A, b.

Here is a slightly more interesting example.

7.

Theorem 2.6. *The set of optimal solutions X_{opt} to a convex optimization problem is a convex set.*

Proof: Suppose we consider, $x_1, x_2 \in X_{\text{opt}}$. Since they are both optimal we must have that $f_0(x_1) = f_0(x_2)$. Now, consider $x_0 = \theta x_1 + (1-\theta)x_2$, where $0 \leq \theta \leq 1$. x_0 is feasible, since the set of feasible solutions is convex. Further, by convexity of the objective we see that,

$$f_0(x_0) \leq \theta f_0(x_1) + (1-\theta) f_0(x_2) \leq f_0(x_1),$$

and so $x_0 \in X_{\text{opt}}$ also. ■

Segue... Next time we will consider a few more examples of convex sets and define operations that preserve convexity of a set, before turning to convex functions.