21-690: Methods of Optimization

Krish Matta

Contents

1 Introduction

Below are my notes for the course 21-690: Methods of Optimization taught in the Spring 2025 semester by Professor Nicholas Boffi at Carnegie Mellon University.

2 Affine Sets

Definition (Affine). A set $C \subseteq \mathbb{R}^n$ is said to be *affine* if for all $x, y \in C$, we have that $\theta x + (1-\theta)y \in C$ for all $\theta \in \mathbb{R}$.

Geometrically, affine sets are sets in which the line formed by any two points in the set is entirely contained in the set as well.

Definition (Affine Combination). An *affine combination* of $\{x_i\}_{i=1}^k \subseteq \mathbb{R}^n$ is a linear combination

$$\sum_{i=1}^{k} \theta_i x_i$$

where $\theta_i \in \mathbb{R}$ for all $i \in [k]$ and $\sum_{i=1}^k \theta_i = 1$.

Proposition. Let $C \subseteq \mathbb{R}^n$ be an affine set. Then, any affine combination of $\{x_i\}_{i=1}^k \subseteq C$ is contained in C.

Proof. By induction on k.

Definition (Affine Hull). The *affine hull* of a set $C \subseteq \mathbb{R}^n$ is the set

$$\operatorname{aff}(C) = \{\theta x + (1-\theta)y : x, y \in C, \theta \in \mathbb{R}\}.$$

Exercise. Prove that the affine hull of $C \subseteq \mathbb{R}^n$ is the smallest affine set containing C.

3 Convex Sets

Affine sets are certainly unbounded, as lines are unbounded. We can consider bounded sets by considering only line segments rather than entire lines. Thus lies the idea behind convexity.

Definition (Convex). A set $C \subseteq \mathbb{R}^n$ is said to be *convex* if for all $x, y \in C$, we have that $\theta x + (1-\theta)y \in C$ for all $\theta \in [0, 1]$.

Geometrically, convex sets are sets in which the line segment formed by any two points in the set is entirely contained in the set as well. Note that we only consider line segments by restricting our coefficients to $[0, 1]$.

We can similarly extend the idea of affine combinations and affine hulls to convexity.

Definition (Convex Combination). A *convex combination* of $\{x_i\}_{i=1}^k \subseteq \mathbb{R}^n$ is a linear combination

$$\sum_{i=1}^{k} \theta_i x_i$$

where $\theta_i \in [0,1]$ for all $i \in [k]$ and $\sum_{i=1}^k \theta_i = 1$.

Proposition. Let $C \subseteq \mathbb{R}^n$ be a convex set. Then, any convex combination of $\{x_i\}_{i=1}^k \subseteq C$ is contained in C.

Proof. By induction on k.

Definition (Convex Hull). The *convex hull* of a set $C \subseteq \mathbb{R}^n$ is the set

$$\mathrm{conv}(C) = \{\theta x + (1-\theta)y : x, y \in C, \theta \in [0,1]\}.$$

Exercise. Prove that the convex hull of $C \subseteq \mathbb{R}^n$ is the smallest convex set containing C.

3.1 Examples

3.1.1 Cones

Definition (Cone). A set C is a *cone* if for all $x \in C$, $\lambda x \in C$ for all $\lambda > 0$.

The traditional image of a "cone" is itself a cone, if extended to infinity.

Not all cones are convex: the union of two different lines is a cone, but not convex.

3.1.2 Hyperplanes and Halfspaces

Definition (Hyperplane). Let $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. The set

$$\{x \in \mathbb{R}^n : a^T x = b\}$$

is a *hyperplane*.

Geometrically, hyperplanes are lines in $\mathbb{R}^2$ and planes in $\mathbb{R}^3$.

Proposition. All hyperplanes are affine.

Proof. Consider the hyperplane

$$S = \{x \in \mathbb{R}^n : a^T x = b\}.$$

Consider $x, y \in S$ and $\theta \in \mathbb{R}$. Then,

$$a^T(\theta x + (1-\theta)y) = \theta(a^T x) + (1-\theta)(a^T y) = \theta b + (1-\theta)b = b.$$

Hence, $\theta x + (1-\theta)y \in S$ and so S is affine.

Definition (Halfspace). Let $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$. The set

$$\{x \in \mathbb{R}^n : a^T x \leq b\}$$

is a *halfspace*.

Geometrically, halfspaces are one side of a hyperplane.

Proposition. All halfspaces are convex.

3.1.3 Balls and Ellipsoids

Definition (Ball). A ball is a set

$$B(x_c, r) = \{x \in \mathbb{R}^n : ||x - x_c|| \leq r\},$$

where $r > 0$.

Proposition. All balls are convex.

Proof. Consider some ball $B(x_c, r)$. Take $x, y \in B(x_c, r)$ and $\theta \in [0,1]$. Let $z = \theta x + (1-\theta)y$. We wish to show that $z \in B(x_c, r)$. Observe that by triangle inequality,

$$||z - x_c|| = ||\theta x + (1-\theta)y - x_c|| \leq ||\theta x - \theta x_c|| + ||(1-\theta)y - (1-\theta x_c)|| \leq \theta r + (1-\theta)r = r,$$

hence $z \in B(x_c, r)$.

Definition (Ellipsoid). An ellipsoid is a set

$$\{x \in \mathbb{R}^n : (x - x_c)^T P^{-1}(x - x_c) \leq 1\}$$

where $x_c \in \mathbb{R}^n$ and $P \in S^n_{++}$.

Proposition. All ellipsoids are convex.

Proof. Consider some ellipsoid

$$C = \{x \in \mathbb{R}^n : (x - x_c)^T P^{-1}(x - x_c) \leq 1\}.$$

Take $x, y \in C$, and $\theta \in [0,1]$. Let $z = \theta x + (1-\theta)y$. We wish to show that $z \in C$. Observe that

$$
\begin{aligned}
&(z - x_c)^T P^{-1}(z - x_c) \\
&= (\theta x + (1-\theta)y - x_c)^T P^{-1}(\theta x + (1-\theta)y - x_c) \\
&= (\theta x + (1-\theta)y - \theta x_c - (1-\theta)x_c)^T P^{-1}(\theta x + (1-\theta)y - \theta x_c - (1-\theta)x_c) \\
&= (\theta(x - x_c) + (1-\theta)(y - x_c))^T P^{-1}(\theta(x - x_c) + (1-\theta)(y - x_c)) \\
&= \theta^2 (x - x_c)^T P^{-1}(x - x_c) + \theta(1-\theta)((x - x_c)^T P^{-1}(y - x_c) + (y - x_c)^T P^{-1}(x - x_c)) \\
&+ (1-\theta)^2 (y - x_c)^T P^{-1}(y - x_c) \\
&= \theta^2 (x - x_c)^T P^{-1}(x - x_c) + 2\theta(1-\theta)(x - x_c)^T P^{-1}(y - x_c) + (1-\theta)^2 (y - x_c)^T P^{-1}(y - x_c) \\
&\leq \theta^2 (x - x_c)^T P^{-1}(x - x_c) + 2\theta(1-\theta)\sqrt{(x - x_c)^T P^{-1}(x - x_c)}\sqrt{(y - x_c)^T P^{-1}(y - x_c)} \\
&+ (1-\theta)^2 (y - x_c)^T P^{-1}(y - x_c) \\
&= \left(\theta\sqrt{(x - x_c)^T P^{-1}(x - x_c)} + (1-\theta)\sqrt{(y - x_c)^T P^{-1}(y - x_c)}\right)^2 \\
&\leq (\theta + (1-\theta))^2 \\
&= 1.
\end{aligned}
$$

Hence, $z \in C$, and so all ellipsoids are convex.

3.1.4 Polyhedra

Definition (Polyhedra). A polyhedra is a set

$$P = \{x \in \mathbb{R}^n : \big(\forall i \in [m], a_i^T x \leq b_i\big) \wedge \big(\forall i \in [p], c_i^T x = d_i\big)\}.$$

Polyhedra are typically presented with the notation

$$P = \{x \in \mathbb{R}^n : Ax \preceq b, Cx = d\}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, C \in \mathbb{R}^{p \times n}, d \in \mathbb{R}^p$.

Geometrically, polyhedra are the intersection of a finite number of hyperplanes and halfspaces.

3.2 Operations that Preserve Convexity

3.2.1 Intersection

Proposition. Intersection preserves convexity.

Proof. Let I be an index set such that for all i, $C_i \subseteq \mathbb{R}^n$ is a convex set. Let $C = \bigcap_{i \in I} C_i$. We claim that C is convex.

Consider $x, y \in C$, and $\theta \in [0, 1]$. Set $z = \theta x + (1 - \theta) y$. It suffices to show that $z \in C$.

By definition of C, $x, y \in C_i$ for all $i \in I$. By convexity of C_i, we have that $z \in C_i$ for all $i \in I$. Hence, $z \in C$, as desired.

Exercise. The positive semidefinite cone, S_+^n, is convex.

Solution. Observe that

$$S_+^n = \bigcap_{z \in \mathbb{R}^n} \{X \in \mathbb{R}^{n \times n} : z^T X z \geq 0\}.$$

Every set on the right hand side is convex. The intersection of convex sets is convex, hence the positive semidefinite cone is convex.

3.2.2 Affine Functions

Definition (Affine Function). A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *affine* if it is of the form

$$f(x) = Ax + b.$$

Proposition. The image of a convex set over an affine function is convex.

Proof. Let $C \subseteq \mathbb{R}^n$ be a convex set and $f : \mathbb{R}^n \to \mathbb{R}^m$ an affine function via $f(x) = Ax + b$. We wish to show that $f[C]$ is convex.

Consider $x, y \in f[C]$, and $\theta \in [0, 1]$. Define $z = \theta x + (1 - \theta) y$. It suffices to show that $z \in f[C]$.

As $x, y \in f[C]$, there exists $s, t \in C$ such that $f(s) = x$ and $f(t) = y$.

Let $u = \theta s + (1 - \theta) t$. As C is convex, we have that $u \in C$.

Thus,

$$
\begin{aligned}
z &= \theta x + (1-\theta)y \\
&= \theta f(s) + (1-\theta)f(t) \\
&= \theta As + \theta b + (1-\theta)At + (1-\theta)b \\
&= A(\theta s + (1-\theta)t) + b \\
&= Au + b \\
&= f(u) \in f[C],
\end{aligned}
$$

as desired.

The above proposition implies that scaling, translation, and projection preserve convexity.

Proposition. The pre-image of a convex set over an affine function is convex.

Proof. Let $C \subseteq \mathbb{R}^m$ be a convex set and $f : \mathbb{R}^n \to \mathbb{R}^m$ an affine function via $f(x) = Ax + b$. We wish to show that $f^{-1}[C]$ is convex.

Consider $x, y \in f^{-1}[C]$, and $\theta \in [0, 1]$. It suffices to show that $z = \theta x + (1-\theta)y \in f^{-1}[C]$, which in turn we must show that $f(z) \in C$.

Observe that

$$
\begin{aligned}
f(z) &= f\left(\theta x + (1-\theta)y\right) \\
&= \theta Ax + \theta b + (1-\theta)Ay + (1-\theta)b \\
&= \theta f(x) + (1-\theta)f(y) \in C
\end{aligned}
$$

by convexity of C.

3.2.3 Perspective Function

Definition (Perspective Function). The perspective function $P : \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}^n$ is defined via

$$
P(s, t) = \frac{s}{t}.
$$

Proposition. The image of a convex set over the perspective function is convex.

Proof. Let $C \subseteq \mathbb{R}^n \times \mathbb{R}_{++}$ be a convex set. We wish to show that $P[C]$ is convex.

Consider $x, y \in P[C]$. We wish to show that for any $\theta \in [0,1]$, $\theta x + (1-\theta)y \in P[C]$. The proof is difficult if done directly, hence we will take a slightly different approach.

Fix $\theta \in [0,1]$. Since $x, y \in P[C]$, there exists $(a,s),(b,t) \in C$ such that $P(a,c) = x$ and $P(b,t) = y$.

By convexity of C, we have that

$$(\theta a + (1-\theta)b, \theta s + (1-\theta)t) = \theta(a,s) + (1-\theta)(b,t) \in C.$$

Hence,

$$\frac{\theta s P(x)}{\theta s + (1-\theta)t} + \frac{(1-\theta)tP(y)}{\theta s + (1-\theta)t} = \frac{\theta a + (1-\theta)b}{\theta s + (1-\theta)t} = P(\theta a + (1-\theta)b, \theta s + (1-\theta)t) \in P[C]$$

Let

$$\mu = \frac{\theta s}{\theta s + (1-\theta)t}.$$

Through substitution, we have that

$$\mu P(x) + (1-\mu)P(y) \in P[C].$$

As we vary $\theta \in [0,1]$, μ varies from 0 to 1, implying that $P[C]$ is convex.