Chapter 1

Introduction

Mixed notes from the book Convex Optimization [1]. Some parts will be supplemented by the book of Dattoro [2]. The book of Dattoro is extremely useful for concretely illustrating most of the concepts.

Quotes

- "We study convex geometry because it is the easiest of geometries. For that reason, much of a practitioner's energy is expended seeking invertible transformation of problematic sets to convex ones". Dattoro [2].

Chapter 2

Convex Sets

2.1 Affine and convex sets and cones.

- **Affine set.** A set $C \subseteq \mathbf{R}^n$ is *affine* if the line through any two points in C lines in C, *i.e.* if $x_1, x_2 \in C$ and $\theta \in \mathbf{R}$ implies $\theta x_1 + (1-\theta)x_2 \in C$. More generally, an affine set is a set that contains all *affine combination* (see def. below) of two or more of its points.

    - * Any affine set is convex [2].
        * The intersection of an arbitrary collection of affine sets remains affine [2].
        * Any affine set is open in the sense that it contains no boundary, *e.g.* the empty set $\emptyset$, point, line, plane, hyperplane, subspace etc [2]. Converse not necessarily true (*e.g.* see point just below about subspace.)
        * If C is an affine set and $x_0 \in C$, then the set $V = C - x_o = \{x - x_0 | x \in C\}$ is a subspace.
    - *Affine combination.* A combination of points $\sum_{i=1}^{k} \theta_i x_i$ where $\sum_{i=1}^{k} \theta_i = 1$ is an affine combination.
    - *Ambient space.* The space where a given set lives in, *e.g.* a plane can live in $\mathbf{R}^2, \mathbf{R}^3$. The choice of ambient space has implications on, for example, the interior of a set ([2], p34)
    - *Affine hull*, denoted $\mathbf{aff}\, C$ is the smallest set that makes C affine.
    - *Affine dimension* of a set C is the dimension of $\mathbf{aff}\, C$. In fact dimension of a set is synonymous with affine dimension [2].
    - *Relative interior.* The interior of, for example, a plane in $\mathbf{R}^3$ is empty. To "fix" this issue, we define the relative interior of C as: $\mathbf{relint}\, C = \{x \in C | B(x,r) \cap \mathbf{aff}\, C \subseteq C \text{ for some } r > 0\}$

- **Convex sets**. A set C is convex if the line segment between any two points in C lies in C, *i.e.* $\theta x_1 + (1-\theta)x_2 \in C$ for any $x_1, x_2 \in C$ and $0 \le \theta \le 1$.

    - *Convex combination.* A combination of points $\sum_{i=1}^{k} \theta_i x_i$ where $\sum_{i=1}^{k} \theta_i = 1$ and $\theta_i \ge 0$ is a convex combination.
    - *Convex hull* $\mathbf{conv}\, C$ of a set C is the smallest set that makes C convex.

- **Cones**. A set C is called a cone if for every $x \in C$ and $\theta \ge 0$ we have $\theta x \in C$. Cones can have very unintuitive shapes, see Fig. 35-41 in [2].

    - *Convex cone* is a set that is cone and also convex, *i.e.* $\theta_1 x_1 + \theta_2 x_2 \in C$ for any $x_1, x_2 \in C$ and for $\theta_1, \theta_2 \ge 0$. Some differences between a convex set and a convex cone: (i) A convex set doesn't have to include the origin, a convex cone does; (ii) a convex set can be bounded but a convex cone cannot.

Some important examples and notes ([1] p27):

- Any subspace is affine and a convex cone

- A line segment is convex but not affine

- A ray (*i.e.* $\{\theta v + x_0 : x \ge 0\}$) is convex but not affine. It is convex cone if its base x_0 is 0.

- Any line is affine.

- The empty set, any single point and the whole space are affine (hence convex) subsets of $\mathbf{R}^n$

- Halfspaces (see below) are convex but not affine.

2.2 Hyperplanes, halfspaces, balls and polyhedra

- **Hyperplane** is a set of the form

$$\{x | a^T x = b\}$$

This set has several intuitive interpretations.

  1. It is the hyperplane with a normal vector a and an offset b from the origin.
  2. Let b be $a^T x = b$. Then, $\{x | a^T x = b\} = \{x | a^T(x - x_0)\} = x_0 + a^\perp$ where $a^\perp$ is the orthogonal complement of a, and x_0 is any point in the hyperplane.
  3. More interpretations on p27-28.

- **Halfspace**. Each hyperplane divides $\mathbf{R}^n$ into two halfspaces. A (closed) halfspace is of the form

$$\{x | a^T x \leq b\},$$

  where $a \neq 0$.

- **Norm ball**. A norm ball is the set of the form $B(x_c, r) = \{x : ||x_c - x|| \leq r, x \in \mathbf{R}^n\}$, where $|| \cdot ||$ is a given norm. Another common representation of the ball is $B(x_c, r) = \{x_c + ru : ||u|| \leq 1, u \in \mathbf{R}^n\}$.

  Norm ball is convex (p30).

- **Norm cone** is the set $C(x, t) = \{(x, t) : ||x|| \leq x \in \mathbf{R}^{n+1}, t \in \mathbf{R}\} \subseteq \mathbf{R}^{n+1}$. It is a convex cone.

- **Proper cone** is a cone $K \subseteq \mathbf{R}^n$ that satisfies the following:

  - is convex
  - is closed
  - is solid, *i.e.* it has nonempty interior
  - is pointed, which means that it contains no line (or, equivalently, $x \in K, -x \in K \implies x = 0$)

  The concept of Proper Cone will be central in defining *generalized inequalities*.

- **Polyhedra**. A polyhedron is the solution set of a finite number of equalities and inequalities:

$$\mathcal{P} = \{x : a_j^T x \leq b_j, j = 1, \ldots, m, c_i^T x = d_i, i = 1, \ldots, m\}.$$

  A simpler notation is $\mathcal{P} = \{x | Ax \preceq b, Cx = d\}$, where the symbol $\preceq$ is *vector* or *componentwise* inequality (p32).

- **Simplexes.** Simplexes are a family of polyhedra; they are also a generalization of the triangle (and its interior); *i.e.* a 1D simplex is a line segment, a 2D simplex is the triangle and its interior, a 3D simplex is tetrahedron.

  - *Affine independence* means that for $v_0, \ldots, v_k \in \mathbf{R}^n$, the points $v_1 - v_0, \ldots, v_k - v_0$ are linearly independent.
  - A simplex can be defined in terms affinely independent points: $C = \mathbf{conv}\{v_0, \ldots, v_k\}$.

2.3 Operations that preserve convexity

- **Intersection**: Convexity is preserved under intersection; the intersection of even infinite convex sets is convex.

- **Affine functions.** Let f be an affine function, *i.e.* $f(x) = Ax + b$. Then the image of S under f,

$$f(S) = \{f(x) \mid x \in S\}$$

  and the inverse image of S under f,

$$f^{-1}(S) = \{x \mid f(x) \in S\}$$

  are both convex if S is convex.

  - *Cartesian product.* Define $S := S_1 \times S_2$ for two convex sets S_1, S_2. Then, S is convex.
  - *Sum* The sum S of two convex sets S_1, S_2, $S = \{x_1 + x_2 : x_1 \in S_1, S_2 \in x_2\}$ is convex.

- **Linear-fractional and perspective functions.**

  - *Perspective function* is the $R^{n+1} \to \mathbf{R}^n$ function $P(x, t) = x/t$ with domain $\mathbf{dom}\, P = \mathbf{R}^{n+1} \times \mathbf{R}_{++}$. That is, the perspective function normalizes the input vector so the last element is one, and then drops this last element. If a set $C \subseteq \mathbf{dom}\, P$ is convex, then its image under P is also convex.
  - *Linear-fractional function* is the composition $P \circ g$ of a perspective function P with an affine function g. It is easy to show that linear-fraction functions preserve convexity: If S is convex, then its image $g(S)$ under g will be convex, then its image under perspective will also be convex.

2.4 Generalized inequalities, minimum and minimal elements

- **Generalized inequality**. A proper cone K (see above) can be used to define a *generalized inequality* as follows:

$$x \preceq_K y \iff y - x \in K$$

A strict generalized inequality $x \prec y$ is defined as

$$x \prec_K y \iff y - x \in \textbf{int } K$$

Example. The nonnegative orthant $\mathbf{R}_+^n$ is a proper cone and for $K = \mathbf{R}_+^n$ the associated inequality $\preceq_K$ corresponds to componentwise inequality $x \prec y$.

Properties of generalized inequalities:

1. $\preceq_K$ is preserved under addition: If $x \preceq_K y$ and $u \preceq_K v$, then $x + u \preceq_K y + v$
2. $\preceq_K$ is transitive: If $x \preceq_K y$ and $y \preceq_K z$, then $x \preceq_K z$.
3. $\preceq_K$ is preserved under nonnegative scaling: If $x \preceq_K y$ and $\alpha > 0$, then $\alpha x \preceq_K \alpha y$.
4. $\preceq_K$ is reflexive: If $x \preceq_K x$.
5. $\preceq_K$ is antisymmetric: If $x \preceq_K y$ and $y \preceq_K x$, then $x = y$.
6. $\preceq_K$ is preserved under limits: If $x_i \preceq_K y_i$ for $i = 1, 2, \ldots$ and $x_i \to x$ and $y_i \to i$, then $x \preceq_K y$.

Properties 1,2,3 are shared by the strict inequality $x \prec_K y$ too, property 4 is strictly *not* shared by it. Also, and additional property for strict generalized inequalities (probably shared by non-strict too):

- if $x \prec_K y$, then for small enough u, v we have $x + u \prec_K y + v$.

- **Minimum and minimal elements**. An essential difference between a regular inequality and a generalized one is that *not all points are comparable*; that is, one of the two inequalities $x \le y$ or $x \le y$ has to hold. This is not the case for generalized inequality.

Example. Consider the proper cone $K = \mathbf{R}_+^n$, and points $x = (3, 3)$, $y = (5, 5)$ and $z = (4, 2)$. Clearly, x and y are comparable and $x \preceq_K y$. Similarly, y and z are comparable and $z \preceq_K y$. However, x and z are not comparable.

- *Minimum element.* We say that $x \in S$ is the minimum element of S (w.r.t. $\preceq_K$) if for every $y \in S$ we have $x \preceq_K y$, which happens if and only if

$$S \subseteq x + K$$

where $x + K$ is the set of all the points that are (i) comparable to x and (ii) greater than or equal to x (confer Fig. 2.17 of [1] or Fig. 43 of [2]).

There can be at most *one* minimum point.

- *Minimal element.* First of all, a minimum point is also a minimal point. But a minimal point can exist even if there is no minimum. There can be more than one minimal points.

We say that $x \in S$ is the minimal point of S (w.r.t. $\preceq_K$) if for any $y \in S$, $y \preceq_K$ holds only if $y = x$. Or, equivalently,

$$(x - K) \cap S = \{x\}$$

where $x - K$ denotes the set of all points that are comparable to x and are less then or equal to x w.r.t. $\preceq_K$.

Confer Fig. 2.17 of [1] or Fig. 43 of [2]. Note that in Fig. 43b it's impossible to draw the cone K (centered on any of the minimal points) that would contain the entire $\mathcal{C}_2$, therefore $\mathcal{C}_2$ has no minimum.

2.5 Separating and supporting hyperplanes

- **Separating hyperplane** is a hyperplane that separates two convex sets.

Separating hyperplane theorem. Let C and D be two convex sets that do not intersect. Then, there exists $a \ne 0$ and b such that $a^T x \le b$ for all $x \in C$ and $a^T x \ge b$ for all $x \in D$.

- *Strict separation* is defined similarly when $\ge$ and $\le$ are replaced by $>$ and $<$.

Converse of separating hyperplane theorem is not in general true, but one can obtain it by adding additional constraints. One variant of converse theorem would be (see p50): Any two convex sets at least one of which is open are disjoint if and only if there exists a separating hyperplane.

- **Supporting hyperplane**. Suppose that $C \subseteq \mathbf{R}^n$ and x_0 is a point in its boundary, *i.e.* $x_o \in \mathbf{bd}\, C = \mathbf{cl}\, C \setminus \mathbf{int}\, C$. If $a \neq 0$ satisfies $a^T x \leq a^T x_0$ for all $x \in C$ (*i.e.* the entire set C lies on one side of the hyperplane), then the hyperplane $\{x : a^T x = a^T x_0\}$ is called a *supporting hyperplane*.

  Convexity and supporting hyperplanes are intimately connected:

  – *Supporting hyperplane theorem*: If C is convex, then there exists a supporting hyperplane for any $x_0 \in \mathbf{bd}\, C$.

  – Partial converse of the theorem: If a set is closed, has nonempty interior and has a supporting hyperplane at every point on its boundary, then it is convex.

2.6   Dual cones and generalized inequalities

- **Dual cone**. Let K be a cone. Then, the set $K^* = \{y : x^T y \geq 0 \text{ for all } x \in K\}$ is called the dual cone of K. Some properties of dual cone K^* (p51 and p53):

  – A dual cone K^* is a ... cone.

  – K^* is always *convex*.

  – K^* is always closed.

  – $K_1 \subseteq K_2 \implies K_2^* \subseteq K_1^*$.

  – If K has nonempty interior, then K^* is pointed.

  – If the closure of K is pointed, than K^* has nonempty interior.

  – K^{**} is the closure of the convex hull of K (hence, if K is convex and closed, $K^{**} = K$).

  – If K is a proper cone, then K^* is also a proper cone.

  An intuitive way of grasping the meaning a dual cone is shown in Fig. 369 [2]: Draw the two hyperplanes that are orthogonal to the two vertices that define the original cone $\mathcal{K}$ in the figure, and then their intersection (i.e., the purple area) is the dual cone (the intersection is necessary because the definition of the dual cone says that any vector in the dual cone should form an acute angle with all points of the original cone).

  – The dual cone of $K = \mathbf{R}^n_+$ is itself.

  – The dual cone of a line in space is its orthogonal complement.

  – More generally, the dual cone of a subspace $V \subseteq \mathbf{R}^n$ is its orthogonal complement $\{y : y^T v = 0 \text{ for all } v \in V\}$

- **Dual generalized inequalities**. Like any proper cone, the dual of a proper cone, K^*, induces a generalized inequality $\preceq_{K^*}$. The following relationships relating the generalized inequality of proper cone and its dual seem to be fundamental:

  – $x \preceq_K y$ iff $\lambda^T x \leq \lambda^T y$ for all $\lambda \succeq_{K^*} 0$.

  – $x \prec_K y$ iff $\lambda^T x < \lambda^T y$ for all $\lambda \succ_{K^*} 0$, $\lambda \neq 0$.

  Those are central properties that allow us to characterize minimum and minimal points w.r.t. a cone K in terms of its dual generalized inequalities, $\preceq_{K^*}$.

  – *Minimum point.* $x \in S$ is the minimum point in S w.r.t. generalized inequality $\preceq_K$ iff for all $\lambda \succ_{K^*} 0$, x is the unique minimizer of $\lambda^T z$ over $z \in S$. Geometrically, this means that for any $\lambda \succ_{K^*}$, the hyperplane

  $$\{z : \lambda^T(z - x) = 0\}$$

  is a strict supporting hyperplane (strict means that the hyperplane intersects S only at one point, *e.g.* it's not a trivial hyperplane such a line being supporting hyperplane of a line).

  – *Minimal point.* There is a gap between necessary and sufficient conditions. A note to illustrate why: $x \in S$ may be minimal of S but there may be no λ for which x minimizes $\lambda^T z$ over $z \in S$ (see Figure 2.25, 2.26, p57).

Bibliography

[1] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[2] J. Dattorro, *Convex optimization & Euclidean distance geometry*, 2010.