CONVEX OPTIMIZATION

ABSTRACT. This course has a pretty good slides. But week 2 by Eric is pretty recordable.

CONTENTS

1. 1/9: determining convexity; generalized inequalities 1
1.1. Ways to determine convexity 1
1.2. Generalized inequalities 3
2. 1/11: Dual cones; Convex functions 5
2.1. Dual Cones and Generalized inequalities 5
2.2. Convex functions 6
Appendix A. a 7
Appendix B. b 7
Appendix C. c 7

1. 1/9: DETERMINING CONVEXITY; GENERALIZED INEQUALITIES

1.1. Ways to determine convexity.

So we start with the question of how might we show that a given set is convex. Let C be the considered set.

Method 1: Apply definition.

Method 2: Show that C is obtained from simple convex sets (hyperplane, half spaces, norm balls, etc) by operation that preserves convexity.

A few operations that preserves convexity are:

Operation (i): $S_1 \cap S_2$ or $\cap S_\alpha$.

Example 1.1. *Examples for this operations.*

(1) Any polyhedron can be written as an intersection of half-spaces and hyperplanes, which means that all polyhedra are convex.

(2) Define $p(t) := \sum_{k=1}^{m} x_k \cos(kt)$ for $x \in \mathbb{R}^m$. For instance, if $m = 2$ then $p(t) = x\cos(t) + y\cos(2t)$.

Then, the following set is convex:

$$\begin{aligned} S &:= \left\{ x \in \mathbb{R}^m \middle| |p(t) < 1|, |t| < \frac{\pi}{3} \right\} \\ &= \bigcap_{|t| \le \pi/3} \left\{ x \middle| (\cos(t), \dots, \cos(nt)) \cdot x \le 1 \right\} \end{aligned}$$

where each of the set in the intersection is convex since it is a slab set, for example when $m = 2$ and $t = -\frac{\pi}{3} = \left\{ x \middle| -1 \le \left(-\frac{1}{2}, \frac{1}{2}\right) \cdot x \le 1 \right\}$, so it's like a slanted slab.

Operation (ii): Affine functions. Suppose $f : \mathbb{R}^n \to \mathbb{R}^n$ is affine ($f(x) = A + b$), then $S \subset \mathbb{R}^n$ is convex means that $\{f(x) : x \in S\}$ is convex. In other words, the affine image of a convex set is convex. Similarly, the preimages of convex sets under f are also convex.

Example 1.2.

(1) Scaling: $f(x) = ax$.
(2) Translation: $a \in \mathbb{R}^n$, $f(x) = x + a$.
(3) Projection: e.g. $\pi_1(x, y) = x \in \mathbb{R}$ is convex. This means that $S \subset \mathbb{R}^m \times \mathbb{R}^n$ is convex implies that $\{x \in \mathbb{R}^m | (x, y) \in S\}$ is also convex.
(4) Suppose $A_1, \dots, A_n$ are symmetric m by m matrices ($\in S^m$). Then for $B \in S^m$ we have that

$$\{x \in \mathbb{R}^n : x_1 A_1 + \cdots + x_n A_n \le B\}$$

is convex since it's the preimage of a halfspace under the function $f(x) = B - (x_1 A_1 + \cdots + x_n A_n)$.
(5) Hyperbolic cone: the set

$$C = \{x \in \mathbb{R}^n \big| x^T P x \le (c^T x)^2, c^T x \ge 0\}$$

with $P \in S_+^n = \{P \in S^n | P \ge 0\}$. Then C is convex because it is the preimage under

$$f(x) = (P^{1/2} x, c^T x)$$

of the second order cone

$$\{(z, t) : z^T z \le t^2, t \ge 0\}.$$

Operation (ii): perspective function and linear fractional functions. So we need to define them first.

Def 1.1. *The* <u>*perspective function*</u> $p : \mathbb{R}^{n+1} \to \mathbb{R}^n$ *is given by* $p(x, t) = \frac{x}{t}$.

The idea is to rescale (x, t) such that the last component is 1, then drop it. This is called a perspective function because it mimics the way how photos are taken.

Our usage here is to say that if $C \subset \text{domain}(P)$ is convex, then so is $P(C)$. The key fact in proving that is nothing but to note that P maps line segments into line segments.

Similarly, the preimage of a convex set under such transformation is also convex.

Def 1.2. *For $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n, d \in \mathbb{R}$, the linear fractional function $f : \mathbb{R}^n \to \mathbb{R}^m$ is given by*

$$f(x) = \frac{Ax + b}{c^T x + d}.$$

One remark here is that this is the composition of an affine function

$$g(x) = \begin{pmatrix} A \\ c^T \end{pmatrix} x + \begin{pmatrix} b \\ d \end{pmatrix}$$

with P. And that when $c = 0, d > 0$, the function f is affine.

An example is

$$f(x) = \frac{1}{x_1 + x_2 + 1} x$$

defined on the halfspace where the denominator is positive.

1.2. **Generalized inequalities.**

Def 1.3. *A convex cone $K \subset \mathbb{R}^n$ is a proper cone if K is*

- *closed;*
- *solid: with non empty interior;*
- *and pointed: contains no line.*

Few examples are

Example 1.3.

(1) Non-negative orthant $K = \mathbb{R}^n_+ := \{x \in \mathbb{R}^n | x_i \geq 0, \forall i\}$.
(2) Positive semidefinite cone

$$K = S^n_+ = \{A \in S^n | A \geq 0\}.$$

(3) Non-negative polynomials on $[0, 1]$, i.e.

$$K := \{x \in \mathbb{R}^n | x_1 + x_2 t + \cdots + x^n t^{n-1} \geq 0, t \in [0, 1]\}.$$

And we'll see that the notion of proper cones is not out of blue. In fact it is crucial to the works below.

Def 1.4. *Given a proper cone K, the* <u>*generalized inequality*</u> *defined by K is the relation $\preceq_K$ given by*

$$x \preceq_k y \iff y - x \in K.$$

We also use the notation $\prec_K$ in the following way:

$$x \prec_K y \iff y - x \in \text{interior}(K).$$

Example 1.4.

(1) For $K = \mathbb{R}^n_+$, $\preceq_K$ is the component wise inequality.
(2) For $K = S^n_+$, $\preceq_K$ is the matrix inequality, i.e.

$$X \preceq_K Y \iff Y - X \text{ is positive semi definite.}$$

Note that we may drop the subscript K on both cases above since they are so common.

Another remark is that $\preceq_K$ forms a order, though not a total order.

Def 1.5. *Given a set S, then $x \in S$ is a* <u>*minimum element*</u> *of S with respect to $\preceq_K$ if $y \in S$ imples $x \preceq_K y$.*

Remark: If a set S has a minimum element, then it's unique. To prove this, assume $x, z \in S$ such that $\forall y \in S$, $x \preceq_K y$ and $z \preceq_K y$, then $x - z \in K$ and $z - x \in K$, yet that means a line is in the proper cone, contradiction.

Def 1.6. *Given a set S, $x \in S$ is a* <u>*minimal element*</u> *of S with respect to $\preceq_K$ if $\forall y \in S$, $y \preceq_K x \Rightarrow x = y$.*

Note that minimal elements are not necessarily unique.

Theorem 1.1. *(Separating Hyperplane Theorem): For C, D non-empty and disjoint convex sets contained in $\mathbb{R}^n$, we have that $\exists a \in \mathbb{R}^n \backslash \{0\}$ such that*

(i) $a^T x \leq b$, $\forall x \in C$;
(ii) $a^T x \geq b$, $\forall x \in D$;

i.e. the hyper plane $\{x : a^T x = b\}$ separates C and D.

Note that the equal sign is in the inequality. And indeed for a strict separation condition additional assumptions are required.

The idea of the proof is the following: Suppose $d(C, D) = \inf_{u \in C, v \in D} ||u - v||_2$ and $\exists c \in C, d \in D$ such that $d(C, D) = ||c - d||_2$. Note that this is satisfied only when both C, D are closed and D bounded.

Now we define the separation hyperplane by $a = d - c$, $b = \dfrac{||d||_2^2 - ||c||_2^2}{2}$ and claim that

$$f(x) := a^T x - b \begin{cases} \leq 0 & x \in C \\ \geq 0 & x \in D. \end{cases}$$

2. 1/11: DUAL CONES; CONVEX FUNCTIONS

An application is the supporting Hyperplane theorem is the following theorem.

Def 2.1. *For $C \subset \mathbb{R}^n$, $x_0 \in \partial C$, if $a \in \mathbb{R}^n$ with $a \neq 0$ satisfying $a^T x \leq a^T x_0$ for all $x \in C$, then the hyperplane*

$$\{x | a^T x = a^T x_0\}$$

*is a supporting hyperplane to **C**.*

*Equivalently, x_0 and **C** are separated by $\{x | a^T x = a^T x_0\}$.*

Theorem 2.1. *If $c \in \mathbb{R}^n$ is a non-empty convex set that there exists a supporting hyperplane for each $x \in \partial C$.*

The proof follows from the separation hyperplane theorem, for example, if $\int(C) \neq \emptyset$, then apply the separating hyperplane thoerem to $\mathrm{int}(C)$ and $\{x\}$.

2.1. **Dual Cones and Generalized inequalities.**

Def 2.2. *For any cone K, the dual cone of K is*

$$K^* := \{y : y^T x \geq 0, \forall x \in K\}.$$

Note that k^* is still a cone, and moreover K^* is always convex even when K is not.

The geometric idea of this is that

$$y \in K^* \iff -y \text{ is the normal of a hyper plane that supports } K \text{ at the origin.}$$

Example 2.1.

(1) $K = \mathbb{R}^n_+$, the first orthant, is it's own dual cone.
(2) $K = S^n_+$ is it's own dual cone.
(3) $K = \{(x,t) \Big| ||x||_2 \leq t\}$ is it's own dual cone.
(4) $K = \{(x,t) \Big| ||x||_1 \leq t\}$ has the dual cone $K^* = \{(x,t) \Big| ||x||_\infty \leq t\}$.

Remark 2.1.

- *Examples 1 to 3 above are self dual cones.*
- *Dual cones of proper cones are proper(!) and thus defines generalized inequalities. Moreover, $x \preceq_K y \iff \lambda^T x = \lambda^T y$ for all $\lambda \preceq_{K^*} 0$ and $x \prec_K y \iff \lambda^T x < \lambda^T y$ for all $\lambda \preceq_{K^*} 0$ with $\lambda \neq 0$.*

Now, what about minimal elements and minimum elements?

Proposition 2.2. *Given $S \subset \mathbb{R}^m$, $x \in S$ is the minimum element of S with respect to $\preceq_K$ iff $\forall \lambda \preceq_{K^*} 0$, x is the unique minimizer of $z \mapsto \lambda^T z$ over all S.*

Proposition 2.3. *If x minimizes $z \mapsto \lambda^T z$ over S for some $\lambda \prec_{K^*} 0$, then x is a minimal element of S with respect to $\preceq_K$. But the converse is false.*

One counter example is a bean shaped set with the point in middle. But when convexity is involved no such bad things could happen.

Proposition 2.4. *When S is convex, then if x is a minimal element of S with respect to $\preceq_K$, then $\exists \lambda \neq 0$ with $\lambda \preceq_{K^*} 0$ such that x minimizes $z \mapsto \lambda^T z$ over S.*

One application of this is the Pareto optimality, which is nothing but to find the minimum in the $\preceq_K$ sense for $K = \mathbb{R}^N_+$.

2.2. **Convex functions.**

Def 2.3. *$f : S \to \mathbb{R}$ for $S \subset \mathbb{R}^n$ is convex if S is convex and*

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta) f(y) \forall x, y \in S, \theta \in [0,1].$$

Remark that we say f is concave if $-f$ is convex.

Def 2.4. *$f : S \to \mathbb{R}$ is strictly convex if S is convex and*

$$f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta) f(y) \forall x, y \in S, \theta \in [0,1].$$

APPENDIX A. A

APPENDIX B. B

APPENDIX C. C

Acknowledgements.