

CPSC 542F WINTER 2017

Convex Analysis and Optimization

Lecture Notes

SUMMER RESEARCH INTERNSHIP, UNIVERSITY OF WESTERN ONTARIO

GITHUB.COM/LAURETHTEX/CLUSTERING

This research was done under the supervision of Dr. Pauline Barmby with the financial support of the MITACS Globalink Research Internship Award within a total of 12 weeks, from June 16th to September 5th of 2014.

First release, August 2014



Contents

1	Convex Sets	5
1.1	Convexity	5
1.1.1	Cone	5
1.2	Convex Functions	6
1.2.1	Epigraph	6
1.3	Support Function	8
1.4	Operations Preserve Convexity of Functions	9

1. Convex Sets

1.1 Convexity

1.1.1 Cone

Definition 1.1.1 — Cone. A set $K \in \mathbb{R}^n$, when $x \in K$ implies $\alpha x \in K$.

A non convex cone can be hyper-plane.

For convex cone $x + y \in K, \forall x, y \in K$.

Cone don't need to be "pointed". e.g.

Direct sums of cones $C_1 + C_2 = \{x = x_1 + x_2 | x_1 \in C_1, x_2 \in C_2\}$.

■ **Example 1.1** $S_1^n \{X | X = X^n, \lambda(x) \geq 0\}$

A matrix with positive eigenvalues.

Operations preserving convexity

Intersection $C \cap_{i \in \mathbb{I}} C_i$

Linear map Let $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear map. If $C \in \mathbb{R}^n$ is convex, so is $A(C) = \{Ax | x \in C\}$

Inverse image $A^{-1}(D) = \{x \in \mathbb{R}^n | Ax \in D\}$

Operations that induce convexity

Convex hull on $S = \cap \{C | S \in C, C \text{ is convex}\}$

■ **Example 1.2** $Co\{x_1, x_2, \dots, x_m\} = \{\sum_{i=1}^m \alpha_i x_i | \alpha \in \Delta_m\}$

For a convex set $x \in C \Rightarrow x = \sum \alpha_i x_i$.

Theorem 1.1.1 — Carathéodory's theorem. If a point $x \in \mathbb{R}^d$ lies in the convex hull of a set P , there is a subset P' of P consisting of $d+1$ or fewer points such that x lies in the convex hull of P' . Equivalently, x lies in an r -simplex with vertices in P .

1.2 Convex Functions

Definition 1.2.1 — Convex function. Let $C \in \mathbb{R}^n$ be convex, $f : C \rightarrow \mathbb{R}$ is convex on f if $x, y \in C \times C$. $\forall \alpha \in (0, 1)$, $f(\alpha x + (1 - \alpha)y) \leq f(\alpha x) + f((1 - \alpha)y)$

Definition 1.2.2 — Strictly Convex function. Let $C \in \mathbb{R}^n$ be convex, $f : C \rightarrow \mathbb{R}$ is strictly convex on f if $x, y \in C \times C$. $\forall \alpha \in (0, 1)$, $f(\alpha x + (1 - \alpha)y) < f(\alpha x) + f((1 - \alpha)y)$

Definition 1.2.3 — Strongly convex. $f : C \rightarrow \mathbb{R}$ is strongly convex with modulus $u \geq 0$ if $f - \frac{1}{2}u\|\cdot\|^2$ is convex.

Interpretation: There is a convex quadratic $\frac{1}{2}u\|\cdot\|^2$ that lower bounds f .

■ **Example 1.3** $\min_{x \in C} f(x) \leftrightarrow \min \bar{f}(x)$ Useful to turn this into an unconstrained problem.

$$\bar{f}(x) = \begin{cases} f(x) & \text{if } x \in C \\ \infty & \text{elsewhere} \end{cases}$$

Definition 1.2.4 A function $f : \mathbb{R}^n \rightarrow \mathbb{R} \cup \infty$ is convex if $x, y \in \mathbb{R}^n \times \mathbb{R}^n$, $\forall x, y, \bar{f}(\alpha x + (1 - \alpha)y) \leq f(\alpha x) + f((1 - \alpha)y)$

Definition 1 is equivalent to definition 2 if $f(x) = \infty$.

■ **Example 1.4** $f(x) = \sup_{j \in J} f_j(x)$

1.2.1 Epigraph

Definition 1.2.5 — Epigraph. For $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$, its epigraph $epi(f) \in \mathbb{R}^{n+1}$ is the set $epi(f) = \{(x, \alpha) | f(x) \leq \alpha\}$

Next: a function is convex i.f.f. its epigraph is convex.

Definition 1.2.6 A function $f : C \rightarrow \mathbb{R}$, $C \in \mathbb{R}^n$ is convex if $\forall x, y \in C$, $f(ax + (1 - a)x) \leq af(x) + (1 - a)f(y) \quad \forall a \in (0, 1)$.

Strict convex: $x \neq y \Rightarrow f(ax + (1 - a)x) < af(x) + (1 - a)f(y)$

(R) f is convex $\Rightarrow -f$ is concave.

Level set: $S_\alpha f = \{x | f(x) \leq \alpha\}$.

$S_\alpha f$ is convex $\Leftrightarrow f$ is convex.

Definition 1.2.7 — Strongly convex. $f : C \rightarrow \mathbb{R}$ is strongly convex with modulus μ if $\forall x, y \in C$, $\forall \alpha \in (0, 1)$, $f(ax + (1 - \alpha)y) \leq af(x) + (1 - \alpha)f(y) - \frac{1}{2\mu}\alpha(1 - \alpha)\|x - y\|^2$.

(R)

- f is 2nd-differentiable, f is convex $\Leftrightarrow \nabla^2 f(x) \succ 0$.
- f is strongly convex $\Leftrightarrow \nabla^2 f(x) \succ \mu I \Leftrightarrow x \geq \mu$

Definition 1.2.8 — 2. $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is convex if $x, y \in \mathbb{R}, \alpha \in (0, 1), f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$.

The effective domain of f is $\text{dom } f = \{x | f(x) < +\infty\}$

■ **Example 1.5 — Indicator function.** $\delta_C(x) = \begin{cases} 0 & x \in C \\ +\infty & \text{elsewhere} \end{cases}$.
 $\text{dom } \delta_C(x) = C$

Definition 1.2.9 — Epigraph. The epigraph of f is $\text{epif} = \{(x, \alpha) | f(x) \leq \alpha\}$

The graph of epif is $\{(x, f(x)) | x \in \text{dom } f\}$.

Definition 1.2.10 — III. A function $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is

Theorem 1.2.1 $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ is convex $\iff \forall x, y \in \mathbb{R}^n, \alpha \in (0, 1), f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y)$.

Proof. \Rightarrow take $x, y \in \text{dom } f, (x, f(x)) \in \text{epif}, (y, f(y)) \in \text{epif}$. ■

■ **Example 1.6 — Distance.** Distance to a convex set $d_C(x) = \inf\{\|z - x\| | z \in C\}$. Take any two sequences $\{y_k\}$ and $\{\bar{y}_k\} \subset C$ s.t. $\|y_k - x\| \rightarrow d_C(x)$, $\|\bar{y}_k - \bar{x}\| \rightarrow d_C(\bar{x})$. $z_k = \alpha y_k + (1 - \alpha)\bar{y}_k$.

$$\begin{aligned} d_C(\alpha x + (1 - \alpha)\bar{x}) &\leq \|z_k - \alpha x - (1 - \alpha)\bar{x}\| \\ &= \|\alpha(y_k - x) + (1 - \alpha)(\bar{y}_k - \bar{x})\| \\ &\leq \alpha\|y_k - x\| + (1 - \alpha)\|\bar{y}_k - \bar{x}\| \end{aligned}$$

Take $k \rightarrow \infty, d_C(\alpha x + (1 - \alpha)\bar{x}) \leq \alpha d(x) + (1 - \alpha)d(\bar{x})$

■ **Example 1.7 — Eigenvalues.** Let $X \in S^n := \{n \times \text{nsymmetric matrix}\}$. $\lambda_1(x) \geq \lambda_2(x) \geq \dots \geq \lambda_n(x)$.

$$f_k(x) = \sum_i^n \lambda_i(x).$$

Equivalent characterization

$$\begin{aligned} f_k(x) &= \max\{\sum_i v_i^T X v_i | v_i \perp v_j, i \neq j\} \\ &= \max\{\text{tr}(V^T X V) | V^T V = I_k\} \end{aligned}$$

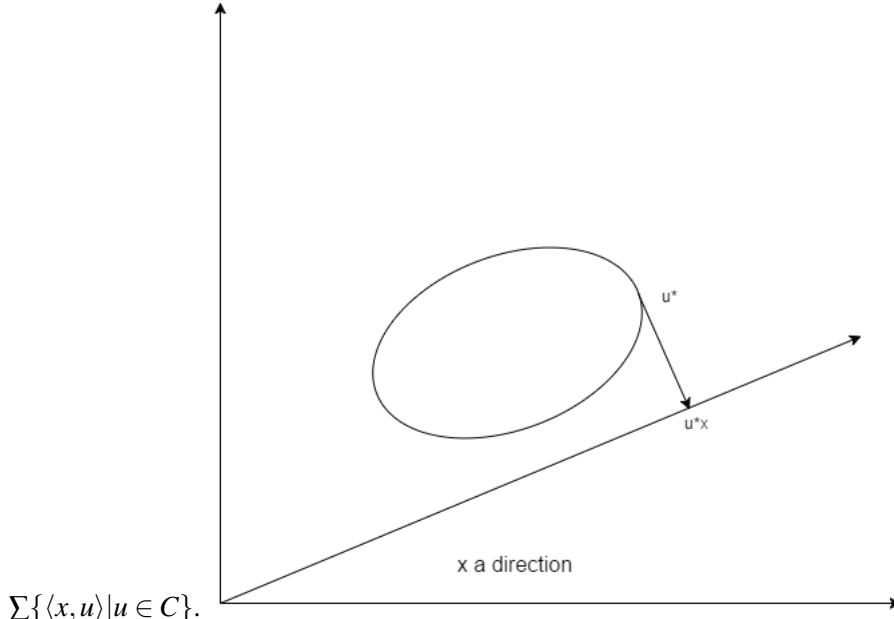
$\max\{\text{tr}(V^T X V)\}$ by circularity

Note $\langle A, B \rangle = \text{tr}(A, B)$ is true for symmetric matrix.

$$\langle A, A \rangle = |A|_F^2 = \sum_i A_{ii}^2$$

1.3 Support Function

Take a set $C \in \mathbb{R}^n$, not necessarily convex. The support function is $\sigma_C = \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$. $\sigma_C(x) =$



Fact 1.3.1 The support function binds the supporting hyper-plane.

Supporting functions are

- Positively homogeneous

$$\sigma_C(\alpha x) = \alpha \sigma_C(x) \forall \alpha > 0$$

$$\sigma_C(\alpha x) = \sup_{u \in C} \langle \alpha x, u \rangle = \alpha \sup_{u \in C} \langle x, u \rangle = \alpha \sigma_C(x)$$

- Sub-linear (a special case of convex, linear combination holds $\forall \alpha$).

$$\sigma_C(\alpha x + (1 - \alpha)y) = \sup_{u \in C} \langle \alpha x + (1 - \alpha)y, u \rangle \leq \alpha \sup_{u \in C} \langle x, u \rangle + (1 - \alpha) \sup_{u \in C} \langle y, u \rangle$$

■ **Example 1.8 — L2-norm.** $\|x\| = \sup_{u \in C} \{\langle x, u \rangle, u \in \mathbb{R}^n\}$.

$$\|x\|_p = \sup\{\langle x, u \rangle, u \in B_q\} \text{ where } \frac{1}{p} + \frac{1}{q} = 1. B_q = \{\|x\|_q \leq 1\}.$$

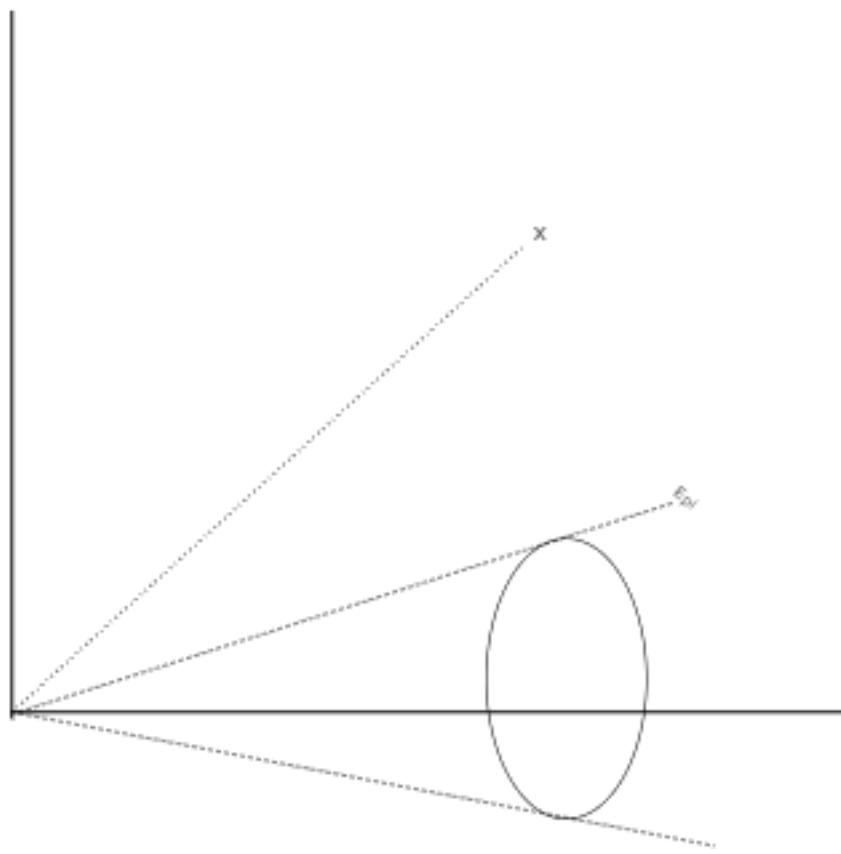
The norm is

- Positive homogeneous
- sub-linear
- If $0 \in C$, σ_C is non-negative.
- If C is central-symmetric, $\sigma_C(0) = 0$ and $\sigma_C(x) = \sigma_C(-x)$

■ **Fact 1.3.2 — Epigraph of a support function.** $epi\sigma_C = \{(x, t) | \sigma_C(x) \leq t\}$. Suppose $(x, t) \in epi\sigma_C$.

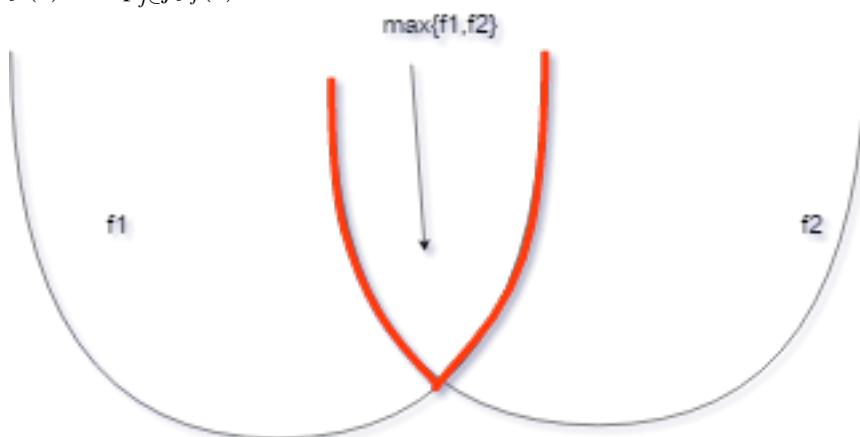
Take any $\alpha > 0$. $\alpha(x, t) = (\alpha x, \alpha t)$.

$$\alpha \sigma_C(x) = \alpha \sigma_C(x) \leq \alpha t. \alpha(x, c) \in epi\sigma_C$$



1.4 Operations Preserve Convexity of Functions

- Positive affine transformation
 $f_1, f_2, \dots, f_k \in \text{cvx} \mathbb{R}^n$
 $f = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_k f_k$
- Supremum of functions. Let $\{f_i\}_{i \in I}$ be arbitrary family of functions. If $\exists x \sup_{j \in J} f_j(x) < \infty \Leftrightarrow f(x) = \sup_{j \in J} f_j(x)$



- Composition with linear map.
 $f \in \text{cvx} \mathbb{R}^n, A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a linear map. $f \circ A(x) = f(Ax) \in \text{cvx} \mathbb{R}^n$

$$\begin{aligned}f \circ A(x) &= f(A(\alpha x + (1 - \alpha)y)) \\&= f(A\alpha x + (1 - \alpha)Ay) \\&\leq \alpha f(Ax) + (1 - \alpha)f(Ay)\end{aligned}$$