Convex Optimization for Machine Learning and Computer Vision

Lecture: Dr. Tao Wu
Exercises: Emanuel Laude, Zhenzhang Ye
Summer Semester 2018

Computer Vision Group
Institut für Informatik
Technische Universität München

Weekly Exercises 1

Room: 02.09.023
Wednesday, 25.04.2018, 12:15-14:00
Submission deadline: Monday, 23.04.2018, 16:15, Room 02.09.023

Theory: Convex Sets (12+8 Points)

Exercise 1 (4 Points)**.** Let $\mathcal{C}$ be a family of convex sets in $\mathbb{R}^n$, $C_1, C_2 \in \mathcal{C}$, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $\lambda \in \mathbb{R}$. Prove convexity of the following sets:

- $\bigcap_{C\in\mathcal{C}} C$
- $P := \{x \in \mathbb{R}^n : Ax \leq b\}$
- $C_1 + C_2 := \{x + y : x \in C_1, y \in C_2\}$ (the Minkowski sum of C_1 and C_2)
- $\lambda C_1 := \{\lambda x : x \in C_1\}$ (the λ-dilatation of C_1).

Solution.

- Let $x_1, x_2 \in \bigcap_{C\in\mathcal{C}} C$. Then $x_1, x_2 \in C$ for all $C \in \mathcal{C}$. Since any C is convex, $\mu x_1 + (1-\mu)x_2 \in C$ for all $\mu \in [0,1]$ and $C \in \mathcal{C}$ and therefore $\mu x_1 + (1-\mu)x_2 \in \bigcap_{C\in\mathcal{C}} C$.
- Let $x_1, x_2 \in P$, which means that $Ax_1 \leq b$ and $Ax_2 \leq b$. Let $\mu \in [0,1]$. Then, $A(\mu x_1 + (1-\mu)x_2) = \mu Ax_1 + (1-\mu)Ax_2 \leq \mu b + (1-\mu)b = b$. Therefore $\mu x_1 + (1-\mu)x_2 \in P$.
- Let $x, y \in C_1 + C_2$. Then there exist $x_1, y_1 \in C_1, x_2, y_2 \in C_2$ so that $x = x_1 + x_2$ and $y = y_1 + y_2$. Let $\mu \in [0,1]$. Then, since C_1, C_2 convex $\mu x + (1-\mu)y = \mu x_1 + \mu x_2 + (1-\mu)y_1 + (1-\mu)y_2 = \underbrace{\mu x_1 + (1-\mu)y_1}_{\in C_1} + \underbrace{\mu x_2 + (1-\mu)y_2}_{\in C_2} \in C_1 + C_2$.
- Let $x, y \in C_1$ and $\mu \in [0,1]$. Then, since C_1 convex, $\mu\lambda x + (1-\mu)\lambda y = \lambda \underbrace{(\mu x + (1-\mu)y)}_{\in C_1} \in \lambda C_1$.

Exercise 2 (4 Points)**.** Let $\emptyset \neq X \subset \mathbb{R}^n$. Prove the equivalence of the following statements:

- X is closed.

- Every convergent sequence $\{x_n\}_{n\in\mathbb{N}} \subset X$ attains its limit in X.

Solution. Let X be closed. By definition this means that the complement of X given as $X_C := \mathbb{R}^n \setminus X$ is open meaning that for all $x \in X_C$ there exists $\epsilon > 0$ s.t. the ball $B_\epsilon(x)$ is entirely contained in X_C:

$$B_\epsilon(x) \cap X = \emptyset.$$

Suppose that there exists a convergent sequence $X \supset \{x_n\}_{n\in\mathbb{N}} \to x$ with $x \notin X$. However, by definition of convergence for all $\epsilon > 0$ there exists $N \in \mathbb{N}$ s.t.

$$X \ni x_n \in B_\epsilon(x)$$

for all $n \geq N$, which contradicts the assumption. Let conversely X not be closed (not the same as open). That means there exists $x \notin X$ s.t. for all $\epsilon > 0$ it holds that $B_\epsilon(x) \cap X \neq \emptyset$. This means that for all $\epsilon_n := \frac{1}{n} > 0$ there exists $x_n \in B_\epsilon(x) \cap X$. By construction we have a sequence $\{x_n\}_{n\in\mathbb{N}}$ converging to $x \notin X$ but with elements in X.

Exercise 3 (4 Points)**.** Prove that if the set $C \subset \mathbb{R}^n$ is convex, then $\sum_{i=1}^N \lambda_i x_i \in C$ with $x_1, x_2, \ldots, x_N \in C$ and $0 \leq \lambda_1, \lambda_2, \ldots, \lambda_N \in \mathbb{R}$, $\sum_{i=1}^N \lambda_i = 1$.

Hint: Use induction to prove.

Solution. When N=2, it directly follows the definition of convex set.
Assume it holds for N. Now consider N+1 case:

$$\sum_{i=1}^{N+1} \lambda_i x_i = \sum_{i=1}^{N} \lambda_i x_i + \lambda_{N+1} x_{N+1}$$

If there exists a certain i such that $\lambda_i = 0$, it will be N case which is assumed to hold. Therefore, all $\lambda_i > 0$ and above equation turns into:

$$(1-\lambda_{N+1}) \sum_{i=1}^{N} \frac{\lambda_i}{1-\lambda_{N+1}} x_i + \lambda_{N+1} x_{N+1}$$

Using our assumption, $\sum_{i=1}^N \frac{\lambda_i}{1-\lambda_{N+1}} x_i$ is an element in C. Therefore, the convexity is proved.

Definition (Convex Hull)**.** The convex hull $\mathrm{conv}(S)$ of a finite set of points $S \subset \mathbb{R}^n$ is defined as

$$\mathrm{conv}(S) := \left\{ \sum_{i=1}^{|S|} a_i x_i : x_i \in S, \sum_{i=1}^{|S|} a_i = 1, a_i \geq 0 \right\}$$

Exercise 4 (8 Points)**.** Prove the following statement: Let $n \in \mathbb{N}$ and let $A \subset \mathbb{R}^n$ contain $n+2$ elements: $|A| = n+2$. Then there exists a partition of A into two disjoint sets A_1, A_2

$$A = A_1 \dot{\cup} A_2,$$

(meaning that $A_1 \cap A_2 = \emptyset$) so that the convex hulls of A_1 and A_2 intersect:

$$\operatorname{conv}(A_1) \cap \operatorname{conv}(A_2) \neq \emptyset.$$

You may use the following hint. Don't forget to prove the hint!

Hint: Let $x_1, \ldots, x_{n+2} \in \mathbb{R}^n$. Then the set $\{x_1 - x_{n+2}, \ldots, x_{n+1} - x_{n+2}\}$ is linearly dependent and there exist multipliers $a_1, \ldots, a_{n+2}$, not all of which are zero, so that

$$\sum_{i=1}^{n+2} a_i x_i = 0, \quad \sum_{i=1}^{n+2} a_i = 0.$$

The desired partition is formed via all points corresponding with $a_i \geq 0$ and all points with $a_i < 0$.

Solution. Let $A := \{x_1, x_2, \ldots, x_{n+2}\} \subset \mathbb{R}^n$. Since $n+1$ vectors in $\mathbb{R}^n$ are always linearly dependent there exist scalars $a_1, \ldots, a_{n+1}$, not all of which are zero so that

$$\sum_{i=1}^{n+1} a_i (x_i - x_{n+2}) = \sum_{i=1}^{n+1} a_i x_i + \underbrace{\left(- \sum_{i=1}^{n+1} a_i \right)}_{=: a_{n+2}} x_{n+2} = 0.$$

Then, by construction $\sum_{i=1}^{n+2} a_i = 0$. Define $A_1 := \{x_i : a_i > 0\}$ and $A_2 := \{x_j : a_j \leq 0\}$. Clearly, $A = A_1 \dot{\cup} A_2$ forms a partition and A_1, A_2 are both nonempty. Suppose A_2 was empty. Then $a_i > 0$ for all $1 \leq i \leq n+2$. But $a_{n+2} := -\sum_{i=1}^{n+1} a_i < 0$ contradicts this assumption (The same holds for A_1). We have that

$$0 = \sum_{\{i: a_i < 0\}} a_i x_i + \sum_{\{j: a_j \geq 0\}} a_j x_j \iff \sum_{\{i: a_i < 0\}} \underbrace{-a_i}_{\geq 0} x_i = \sum_{\{j: a_j \geq 0\}} a_j x_j,$$

and on the other hand

$$0 = \sum_{\{i: a_i < 0\}} a_i + \sum_{\{j: a_j \geq 0\}} a_j \iff \sum_{\{i: a_i < 0\}} -a_i = \sum_{\{j: a_j \geq 0\}} a_j =: w > 0.$$

Altogether this yields

$$\underbrace{\sum_{\{i: a_i < 0\}} \frac{-a_i}{w} x_i}_{\in \operatorname{conv}(A_1)} = \underbrace{\sum_{\{j: a_j \geq 0\}} \frac{a_j}{w} x_j}_{\in \operatorname{conv}(A_2)},$$

which completes the proof. The theorem is called Radon's Theorem.