IE 521: Convex Optimization	**Spring 2017, UIUC**

Lecture 2: Geometry of Convex Sets – January 25

Instructor: Niao He *Scribe: Shuanglong Wang*

Courtesy warning: These notes do not necessarily cover everything discussed in the class. Please email TA (swang157@illinois.edu) if you find any typos or mistakes.

In this lecture, we cover the following topics

- Nice topological properties (cont'd)
- Representation Theorem (Caratheodory)
- Radon, Helley Theorems

2.1 Recall

- A set X is convex if $\forall x, y \in X, \lambda x + (1-\lambda)y \in X$ for any $\lambda \in [0,1]$.
- Convex Hull:
$$\text{Conv}(X) = \left\{ \sum_{i=1}^k \lambda_i x_i : k \in \mathbf{N}, \lambda_i \geq 0, \sum_{i=1}^k \lambda_i = 1, x_i \in X, \forall i = 1, \ldots, k \right\}.$$
Note that the convex hull of X is the smallest convex set that contains X.
- Affine Hull:
$$aff(X) = \left\{ \sum_{i=1}^k \lambda_i x_i : k \in \mathbf{N}, x_i \in X, \sum_{i=1}^k \lambda_i = 1, \forall i = 1, \ldots, k \right\}.$$
Notet that the affine hull of X is the smallest affine subspace that contains X. An affine space M is a shifted linear space, i.e. $M = \{a\} + L$, e.g. $\{x : Ax = b\} = x_0 + \{x : Ax = 0\}$ where x_0 is such that $Ax_0 = b$. The dimension of X: $\dim(X) = \dim($
$aff(X))$
- If $X \in \mathbf{R}^n$ is convex with non-empty interior, then
$$\forall x_0 \in \text{int}(X), x \in \text{cl}(X) \Rightarrow \lambda x_0 + (1-\lambda)x \in \text{int}(X).$$
Moreover, the interior of X if nonempty, is dense in $\text{cl}(X)$, i.e., $\text{cl}(\text{int}X) = \text{cl}(X)$

Questions: What if $\text{int}(X) = \emptyset$? For example, $X = Conv(e_1, e_2, e_3) = \{(x_1, x_2, x_3) : x_1, x_2, x_3 \geq 0, x_1 + x_2 + x_3 = 1\}$, $\text{int}(X) = \emptyset$,
$aff(X) = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 1\}$, $\dim(X) = 2$.

Definition 2.1 (Relative Inerior) $rint(X) = \{x : \exists r > 0, s.t. B(x,r) \bigcap Aff(X) \subseteq X\}$

Fact: If X is convex and nonempty, then rint(X) is always non-empty.
The proof is left as an exercise.

Proposition 2.2 *Let X be an nonempty convex set. Then*

a) $int(X)$, $cl(X)$, $rint(X)$ are non convex

b) If $x_0 \in rint(X), x \in cl(X)$, then $\lambda x_0 + (1-\lambda)x \in rint(X), \forall \lambda \in (0,1]$

c) $cl(rint(X)) = cl(X)$

d) $rint(cl(X)) = rint(X)$

Proof: a) is straightfoward, b), c) can be derived as an analogy to the case when int(X) is non-empty. d) is due to the following fact: $\forall x_0 \in \text{rint}(X)$, $x \in \text{rint}(\text{cl}(X))$, there exitts $y \in \text{cl}(X)$, s.t. $x \in (x_0, y)$. From *b*), this implies $x \in \text{rint}(X)$. ∎

Remark A convex set is perfectly well approximated by its relative interior or closure.

2.2 Representation Theorem

Theorem 2.3 *(Caratheodory) Let $X \subseteq \mathbf{R}^n$ be non empty and $\dim(X) = d \leq n$. Every point $x \in conv(X)$ is a convex combination of at most (d+1) points, i.e.*

$$Conv(X) = \left\{ \sum_{i=1}^{d+1} \lambda_i x_i : x_i \in X, \lambda_i \geq 0, \sum_{i=1}^{d+1} = 1 \right\}$$

Proof: Suppose the minimal representation of $x \in \text{Conv}(X)$ has $m \geq d+1$ terms

$$x = \sum_{i=1}^{m} \alpha_i x_i, \text{where } \alpha_i \geq 0, \sum_{i=1}^{m} \alpha_i = 1$$

The system of linear equations

$$\begin{cases} \sum_{i=1}^m \delta_i x_i = 0 \\ \sum_{i=1}^m \delta_i = 0 \end{cases}$$

has non trivial solution.
We can write $x = \sum_{i=1}^m (\alpha_i - t\delta_i)x_i$. Let $\lambda_i(t) = (\alpha_i - t\delta_i), i = 1, \ldots, m$, we have $\sum \lambda_i(t) = 1$. Let $t_* = \min \left\{ \frac{\alpha_i}{\delta_i}, \delta_i > 0 \right\} := \frac{\alpha_j}{\delta_j}$, then $\lambda_i(t_*) > 0, \forall i \neq j$ and $\lambda_j(t_*) = 0$. This leads to a smaller representation of x, contradiction! ∎

Example Suppose there are 100 different kinds of herbal tea, everyone of them is a blend of 25 herbs. Donald wants n particular mixture of all herbal teas with equal proportions. What's the least number of teas he should buy? **26**

2.3 Radon and Helley Theorem

Theorem 2.4 *(Radon) Let S be a collection of N points in $\mathbf{R}^n$ with $N \geq n+2$. Then we can write $S = S_1 \cup S_2$ s.t. $S_1 \cap S_2 = \emptyset$, and $Conv(S_1) \cap Conv(S_2) \neq \emptyset$*

Proof: Let $S = \{x_1, ..., x_N\}$ Consider the linear system

$$\begin{cases} \sum_{i=1}^N \gamma_i x_i = 0 \\ \sum_{i=1}^N \gamma_i = 0 \end{cases} \implies \text{(n+1)equations, but } N \geq (n+2) \text{ unknowns}$$

There exists a non-zero solution $\gamma_1, ..., \gamma_N$.
Let $I = \{i : \gamma_i \geq 0\}$, $J = \{j : \gamma_j < 0\}$ and $a = \sum_{i \in I} \gamma_i = -\sum_{j \in J} \gamma_j$, then

$$\sum_{i \in I} \gamma_i x_i = \sum_{j \in J} (-\gamma_j) x_j \implies \sum_{i \in I} \frac{\gamma_i}{a} x_i = \sum_{j \in J} \frac{-\gamma_j}{a} x_j$$

The partition $S_1 = \{x_i, i \in I\}$ and $S_2 = \{x_j : j \in J\}$ gives the desired result. ■

Theorem 2.5 *(Helley) Let $S_1, ..., S_N$ be a collection of convex sets in $\mathbf{R}^n$. Assume every (n+1) sets of them have a point in common, then all the sets have a point in common.*

Proof: Let's prove this by induction on N
Base case: $N = n+1$, obviously true.
Induction step: Assume that the collection of $N(\geq n+1)$ sets have common point if every $(n+1)$ of them have common point. We want to show that this holds true for a collection of $N+1$ sets.

From the assumption, there exits $x_i \in S_1 \cap ... \cap S_{i-1} \cap S_{i+1} \cap ... \cap S_{N+1} \neq \emptyset$. Hence, we obtain $(N+1) \geq (n+2)$ points $\{x_1, ..., x_{N+1}\}$.

By Radom's theorem, we can split $\{x_1, x_2, ...x_{N+1}\}$ into disjoint sets, whose convex hulls have nonempty intersection. Without loss of generality, let's assume the two disjoint sets are $\{x_1, \dots, x_k\}$ and $\{x_{k+1}, \dots, x_N\}$, and

$$\text{Conv}(\{x_1, ..., x_k\}) \cap \text{Conv}(\{x_{k+1}, ..., x_{N+1}\}) \neq \emptyset.$$

Let $z \in \text{Conv}(\{x_1, ..., x_k\}) \cap \text{Conv}(\{x_{k+1}, ..., x_{N+1}\})$. Since $\{x_1, ..., x_k\} \subseteq S_{k+1} \cap ... \cap S_{N+1} \implies z \in \text{Conv}(\{x_1, ..., x_k\}) \subseteq S_{k+1} \cap ... \cap S_{N+1}$.Since $\{x_{k+1}, ..., x_{N+1}\} \subseteq S_1 \cap ... \cap S_k \implies z \in \text{Conv}(\{x_{k+1}, ..., x_{N+1}\}) \subseteq S_1 \cap ... \cap S_k$. Therefore, $z \in S_1 \cap ... \cap S_{N+1}$. ■

Remark

- The theorem is not true for infinite collection: e.g. $S_i = [i, \infty), \cap_{i=1}^{+\infty} S_i = \emptyset$
- The theorem is not true if reduce to $(n+1)$ sets to n sets.

Corollary 2.6 *(Helley) Let $\mathscr{F}$ be any collection of compact convex sets in $\mathbf{R}^n$. If every $(n+1)$ sets have common point, then all sets have n points in common*

Remark Helley's theorem have many applications, especially for uniform approximation.

Example Consider the optimization problem

$$p_* = \min_{x \in \mathbf{R}^{10}} g_0(x), \quad \text{s.t. } g_i(x) \le 0, i = 1, ..., 521$$

Suppose $\forall t \in \mathbf{R}$, $X_0 = \left\{x \in \mathbf{R}^{10} : g_0(x) \le t\right\}$ is convex, $X_i = \left\{x \in \mathbf{R}^{10} : g_i(x) \le 0\right\}$ is convex. How many constraints can you drop without affecting the optimal value?

Answer: You can drop as many as 521 - 11 = 510 constraints. i.e. you just need to keep 11 constraints.

Proof: Suppose every 11 constraint relaxation will change the optimal value. $\forall \{i_1, i_2, ...i_n\} \subseteq 1, ...521$,

$$\min_{x \in \mathbf{R}^{10}} \{g_o(x) : g_{i_1}(x) \le 0, ..., g_{i_n}(x) \le 0\} = p(i_1, ..., i_N) < p_*$$

Since there are only finite combinations, let p_{max} be the largest among $p(i_1, ..., i_N)$, $p_{max} < p_*$. Consider the collection of sets, $i = 1, ..., 521$

$$S_i = \left\{x \in \mathbf{R}^{10} : g_0(x) \le p_{max}, g_i(x) \le 0\right\}$$

Hence, we have

(i) S_i is nonempty and convex, and

(ii) every 11 sets of them have non empty intersection

By Helley's theorem, $S_1 \cap ... \cap S_{521} \neq \emptyset$ i.e. $\exists x \in \mathbf{R}^{10}$ s.t. $g_0(x) < p_*$ and $g_i(x) \le 0, \forall i$ Contradiction!
■