Convexity and Optimization

Given two points $u, v \in \mathbb{R}^n$, the interval spanned by u and v is defined as $[u, v] = \{\alpha u + (1-\alpha)v \mid 0 \leq \alpha \leq 1\} \subseteq \mathbb{R}^n$.

Definition 1 *A closed set $K \subseteq \mathbb{R}^n$ is called* convex *if for all $u, v \in K$ we have $[u, v] \subseteq K$.*

An easy to see equivalent definition is that a closed set $K \subseteq \mathbb{R}^n$ is convex iff for all $u, v \in K$ we have $\frac{1}{2}u + \frac{1}{2}v \in K$.

Lemma 1 *Let I be a finite set of indices. If $K_i \subseteq \mathbb{R}^n$, $i \in I$ are closed convex sets, then $\cap_{i \in I} K_i$ is also closed and convex.*

Definition 2 *If $x^i \in \mathbb{R}^n$ for $i = 1, ..., m$ and $\alpha_i \geq 0$, $i = 1, ..., m$ satisfying $\sum_{i=1}^m \alpha_i = 1$, then the vector $x = \sum_{i=1}^m \alpha_i x^i$ is called a* convex combination *of the vectors x^i, $i = 1, ..., m$.*

Definition 3 *Given a set $X \subseteq \mathbb{R}^n$, the* convex hull *of X (denoted by* $\text{conv}(X)$*) is the intersection of all convex sets $K \subseteq \mathbb{R}^n$ that contain X.*

Note that if X is finite then $Y = \text{conv}(X)$ is a closed convex set. We say in this case that Y is finitely generated. In the sequel we shall be concerned mostly with finitely generated closed convex sets.

Theorem 1 (Radon (1921)) *If $X \subseteq \mathbb{R}^n$ is of size $|X| \geq n+2$, then there exists a proper subset $\emptyset \neq S \subsetneq X$ such that* $\text{conv}(S) \cap \text{conv}(X \setminus S) \neq \emptyset$.

Proof. Assume $X = \{\mathbf{x}^i \mid i = 1, ...m\}$, $m \geq n+2$. Let us consider the set of equalities

$$\begin{aligned} \sum_{i=1}^m \tau_i &= 0 \\ \sum_{i=1}^m \tau_i \cdot \mathbf{x}^i &= 0 \end{aligned}$$

This homogeneous system contains $n+1$ equations in $m \geq n+2$ variables, and thus it has a nontrivial solution $0 \neq \tau^* \in \mathbb{R}^m$. Let us define $P = \{i \mid \tau_i^* > 0\}$ and $N = \{i \mid \tau_i^* \leq 0\}$. Note that we must have $P \neq \emptyset$ and $N \neq \emptyset$. Define further

$$\pi = \sum_{i \in P} \tau_i^*,$$

and $S = \{\mathbf{x}^i \mid i \in P\}$. Then we have

$$\mathbf{x}^* = \sum_{i \in P} \frac{\tau_i^*}{\pi} \mathbf{x}^i \; = \; \sum_{j \in N} -\frac{\tau_j^*}{\pi} \mathbf{x}^j$$

and thus $\mathbf{x}^* \in \operatorname{conv}(S) \cap \operatorname{conv}(X \setminus S)$, as claimed. $\square$

Theorem 2 (Helly (1913)) *If $K^i \subseteq \mathbb{R}^n$ are convex sets $i = 1, ...m$, $m \geq n+2$, such that for all $I \subseteq \{1, ...m\}$, $|I| \leq n+1$ we have $\cap_{i \in I} K^i \neq \emptyset$, then $\cap_{i=1}^m K^i \neq \emptyset$.*

Proof. Consider a smallest counter example. Then for all indices i we have a point $\mathbf{p}^i \in \cap_{\substack{j=1 \\ j \neq i}}^m K^j$. Since we must have $m \geq n+2$, by Radon's theorem there exists a proper subset $\emptyset \neq S \subsetneq \{1, ..., m\}$ such that $X = \operatorname{conv}(\{\mathbf{p}^i \mid i \in S\}) \cap \operatorname{conv}(\{\mathbf{p}^j \mid j \notin S\}) \neq \emptyset$. Since all $\mathbf{p}^i$ belongs to all sets K^j, $j \neq i$, we get that any point $\mathbf{x} \in X$ must belong to all sets K^i, $i = 1, ...m$. $\square$

Theorem 3 (Caratheodory (1907)) *If $X \subseteq \mathbb{R}^n$ is finite and $\mathbf{p} \in \operatorname{conv}(X)$, then there exists $S \subseteq X$, $|S| \leq n+1$ such that $\mathbf{p} \in \operatorname{conv}(S)$.*

Proof. Assume $X = \{\mathbf{x}^i \mid i = 1, ...m\} \subseteq \mathbb{R}^n$ and that $\mathbf{p} = \sum_{i=1}^m \alpha_i \mathbf{x}^i$ such that $\alpha_i \geq 0$ for all $i = 1, ..., m$, and $\sum_{i=1}^m \alpha_i = 1$. Denote by $I = \{i \mid \alpha_i > 0\}$ and assume that $|I|$ is the smallest among all possible convex representations of $\mathbf{p}$. If $|I| \leq n+1$, then we are done with $S = \{\mathbf{x}^i \mid i \in I\}$. Otherwise by Radon's theorem we have $\beta_i \in \mathbb{R}$ such that $\sum_{i \in I} \beta_i = 0$, $\sum_{i \in I} \beta_i \mathbf{x}^i = 0$, and $\beta \neq 0$. Let us then choose an index $j \in I$ such that

$$0 > \frac{\alpha_j}{\beta_j} \geq \max_{k \in I : \beta_k < 0} \frac{\alpha_k}{\beta_k}.$$

Then we have

$$\mathbf{p} \; = \; \sum_{i \in I} \left[\alpha_i - \frac{\alpha_j}{\beta_j} \beta_i \right] \mathbf{x}^i$$

as another convex representation of $\mathbf{p}$, with fewer positive components than $|I|$, contradicting the choice of I. $\square$

Lemma 2 *Given $X \subseteq \mathbb{R}^n$, the set* $\mathrm{conv}\,(X)$ *is the set of convex combinations of finite subsets of X.*

Proof. Follows by Caratheodory's theorem. $\square$

Read more beautiful theorems about convexity in Danzer, Grünbaum and Klee: *Helly's theorem and its relatives*, 1921.

The idea of convex separation: for a vector $\mathbf{a} \in \mathbb{R}^n$ and real $b \in \mathbb{R}$ we denote by $H(\mathbf{a}, b) = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}^t\mathbf{x} \geq b\}$ the halfspace defined by $\mathbf{a}$ and b. For a convex set $K \subseteq \mathbb{R}^n$ and point $\mathbf{x}^0 \in \mathbb{R}^n$, $\mathbf{x}^0 \notin K$ we say that $H(\mathbf{a}, b)$ separates $\mathbf{x}^0$ from K if we have $K \subseteq H(\mathbf{a}, b)$ and $\mathbf{x}^0 \notin H(\mathbf{a}, b)$.

Lemma 3 *For all convex sets $K \subseteq \mathbb{R}^n$ and points $\mathbf{x}^0 \in \mathbb{R}^n \setminus K$ there exists a halfspace $H(\mathbf{a}, b)$ that separates $\mathbf{x}^0$ from K.*

$\square$

Note that a halfspace $H(\mathbf{a}, b)$ is a convex set itself. For a point $\mathbf{x}^0 \in \mathbb{R}^n$ and closed convex set $K \subseteq \mathbb{R}^n$ we define

$$d(\mathbf{x}^0, K) = \min_{\mathbf{x} \in K} d(\mathbf{x}^0, \mathbf{x}),$$

where $d(\mathbf{x}, \mathbf{y})$ is the Euclidean distance of points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$: $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(\mathbf{x} - \mathbf{y})^T(\mathbf{x} - \mathbf{y})}$. In particular, we have $d(\mathbf{x}^0, K) = 0$ iff $\mathbf{x}^0 \in K$.

The next theorem is a first example for a min-max theorem:

Theorem 4 *Given a closed convex body $K \subseteq \mathbb{R}^n$ and a point $\mathbf{x}^0 \in \mathbb{R}^n \setminus K$, we have*

$$\max_{\substack{H(\mathbf{a},b) \supseteq K \\ \mathbf{x}^0 \notin H(\mathbf{a},b)}} d(\mathbf{x}^0, H(\mathbf{a}, b)) \; = \; d(\mathbf{x}^0, K) \; = \; \min_{\mathbf{x} \in K} d(\mathbf{x}^0, \mathbf{x}).$$

$\square$

Polyhedra defined by systems of inequalities

Let us consider convex polyhedral regions defined by a system of inequalities:

$$P(A, \mathbf{b}) \;=\; \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} \leq \mathbf{b}\},$$

where $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. In the rest of this subsection we refer to $P(A, \mathbf{b})$ simply as P.

We usually assume that the nonnegativity of the variables is included among these inequalities, and thus in particular that $m \geq n$. As usual we use $[m] = \{1, ..., m\}$, and denote by $\mathbf{a}_i = (a_{i,1}, a_{i,2}, ..., a_{i,n}) \in \mathbb{R}^n$ the ith row of matrix A for $i = 1, ..., m$.

We also assume that A is of full column rank (that is the rank of A is n).

For a subset $I \subseteq [m]$ we define

$$P(I) \;=\; \left\{ \mathbf{x} \in \mathbb{R}^n \;\middle|\; \begin{array}{ll} \mathbf{a}_i^T\mathbf{x} & = b_i \quad \forall i \in I \\ \mathbf{a}_i^T\mathbf{x} & \leq b_i \quad \forall i \in [m] \setminus I \end{array} \right\}$$

Let us further denote by $\mathcal{T} = \mathcal{T}(A, \mathbf{b}) \subseteq 2^{[m]}$ the family of maximal subsets $I \subseteq [m]$ for which $P(I) \neq \emptyset$. In other words for all $I \in \mathcal{T}$ we have $P(I) \neq \emptyset$, but for all such subsets and elements $i \in [m] \setminus I$ we have $P(I \cup \{i\}) = \emptyset$. Finally, for a vector $\mathbf{x} \in P$ we define $T(\mathbf{x}) = \{i \in [m] \mid \mathbf{a}_i^T\mathbf{x} = b_i\}$.

Lemma 4 *We have $|P(I)| = 1$ for all $I \in \mathcal{T}(A, b)$.*

Proof. Assume indirectly that for some $I \in \mathcal{T}(A, \mathbf{b})$ we have $|P(I)| \geq 2$, and let $\mathbf{x}, \mathbf{y} \in P(I)$ be two distinct vectors. Note first that if we have $\mathbf{a}_i^T\mathbf{x} = 0$ for all $i \in [m] \setminus I$, the vector $\mathbf{x}$ is the solution of a system of equalities with A as coefficient matrix. Since A is of full column rank, $\mathbf{x}$ must be the unique solution, and thus we can assume w.l.o.g. that there exists an index $i \in [m] \setminus I$ such that $\mathbf{a}_i^T\mathbf{y} \neq 0$. Note also that by the maximality of the set I, we must have $\mathbf{a}_k^T\mathbf{x} > b_k$ and $\mathbf{a}_k^T\mathbf{y} > b_k$ for all indices $k \in [m] \setminus I$. Consequently, we have

$$\mathbf{z}(\lambda) = \mathbf{x} - \lambda\mathbf{y} \in P(I) \subseteq P$$

for all $-\delta \leq \lambda \leq \delta$ for some suitably small $\delta > 0$. Note finally that if $\epsilon = sign(\mathbf{a}_i^T\mathbf{y})$ then we have

$$\mathbf{a}_i^T\mathbf{z}(\lambda \cdot \epsilon) < b_i \quad \text{for all} \quad \lambda > \frac{\mathbf{a}_i^T\mathbf{x} - b_i}{\epsilon \cdot \mathbf{a}_i^T\mathbf{y}}.$$

Thus, there exists a largest $\lambda > 0$ such that $\mathbf{z}(\lambda \cdot \epsilon) \in P(I) \subseteq P$. For that value we must have

$$T(\mathbf{z}(\lambda \cdot \epsilon)) \supsetneq I$$

which would then contradict the maximality of I. This contradiction proves our claim. $\square$

Lemma 5 *If for a vector* $\mathbf{x} \in P$ *we have* $T(\mathbf{x}) \notin \mathcal{T}(A, \mathbf{b})$ *then* $\mathbf{x}$ *is not a vertex.*

Proof. By definition, $T(\mathbf{x}) \notin \mathcal{T}$ implies the existence of $I \in \mathcal{T}$ such that $I \supsetneq T(\mathbf{x})$, and thus by the above lemma we have $P(I) = \{\mathbf{y}\}$ for a vector $\mathbf{y} \neq \mathbf{x}$. Since $\mathbf{a}_i^T \mathbf{x} > b_i$ for all $i \in [m] \setminus T(\mathbf{x})$ which implies that $(1+\epsilon)\mathbf{x} - \epsilon\mathbf{y} \in P$ for all $-\delta \leq \epsilon \leq \delta$ for some suitably small $\delta > 0$. Thus in particular we have $\mathbf{u} = (1+\delta)\mathbf{x} - \delta\mathbf{y} \in P$ and $\mathbf{v} = (1-\delta)\mathbf{x} + \delta\mathbf{y} \in P$. Since these are distinct vectors (because we have $\mathbf{x} \neq \mathbf{y}$), and since $\mathbf{x} = \frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$ holds, $\mathbf{x}$ cannot be a vertex, as claimed. $\square$

Lemma 6 *If* $\mathbf{x} \in P$ *is not a vertex of* P*, then we have* $T(\mathbf{x}) \notin \mathcal{T}(A, \mathbf{b})$*.*

Proof. By definition, we must have $\mathbf{u}, \mathbf{v} \in P$, $\mathbf{u} \neq \mathbf{v}$ such that $\mathbf{x} = \frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$. Note that this implies $T(\mathbf{x}) = T(\mathbf{u}) \cap T(\mathbf{v})$, and thus by Lemma 4 we must have $T(\mathbf{u}) \supsetneq T(\mathbf{x})$ proving our claim. $\square$

Corollary 1 *The set*

$$V(P) \;=\; \{\mathbf{x} \in P \mid T(\mathbf{x}) \in \mathcal{T}(A, \mathbf{b})\}$$

is the set of vertices of P*.*

Proof. Immediate by lemmas 5 and 6. $\square$

Polyhedra defined by a system of equalities in the positive orthant

Let us consider convex polyhedral regions defined by a system of equalities and nonnegativity:

$$Q(A, \mathbf{b}) \;=\; \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{b}, \;\; \mathbf{x} \geq 0\},$$

where $A \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. In this case we assume that $n \geq m$ and that A is of full row rank, that is the rank of A is m. In the rest of this subsection we refer to $Q(A, \mathbf{b})$ simply as Q. As usual, we denote by $\mathbf{a}^j = (a_{1,j}, a_{2,j}, ..., a_{m,j}) \in \mathbb{R}^m$ the jth column of A, for $j \in [n]$.

For a vector $\mathbf{x} \in Q$ we define $S(\mathbf{x}) = \{j \in [n] \mid \mathbf{x}_j > 0\}$, sometimes called the *support* of $\mathbf{x}$ (since we have $x_j = 0$ for all $j \in [n] \setminus S(\mathbf{x})$.) For a subset $J \subseteq [n]$ we simply refer to the set of column vectors $\{\mathbf{a}^j \mid j \in J\}$ as $A(J)$.

Lemma 7 *If a vector* $\mathbf{x} \in Q$ *is not a vertex of* Q*, then* $A(S(\mathbf{x}))$ *is a set of linearly dependent vectors.*

Proof. If $\mathbf{x}$ is not a vertex, then by definition we have $\mathbf{u}, \mathbf{v} \in Q$, $\mathbf{u} \neq \mathbf{v}$ such that $\mathbf{x} = \frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$. This implies that $S(\mathbf{x}) = S(\mathbf{u}) \cup S(\mathbf{v})$. Furthermore we have $\mathbf{z} = \mathbf{u} - \mathbf{v} \neq 0$ and $A\mathbf{z} = 0$ implying our claim, since $A\mathbf{z}$ is a nontrivial linear combination of the vectors $A(S(\mathbf{x}))$. $\square$

Lemma 8 *If for a vector* $\mathbf{x} \in Q$ *the set* $A(S(\mathbf{x}))$ *is linearly dependent, then* $\mathbf{x}$ *is not a vertex.*

Proof. The linear dependence of $A(S(\mathbf{x}))$ implies the existence of a vector $\mathbf{z} \in \mathbb{R}^n$, $\mathbf{z} \neq 0$ such that $A\mathbf{z} = 0$ and $z_j = 0$ for all $j \in [n] \setminus S(\mathbf{x})$. Thus we have $x_j > 0$ for all indices for which $z_j \neq 0$. This implies that $\mathbf{x} + \lambda\mathbf{z} \in Q$ for all $-\delta \leq \lambda \leq \delta$ for a suitably small $\delta > 0$. Thus in particular we have that $\mathbf{u} = \mathbf{x} + \delta\mathbf{z}$ and $\mathbf{v} = \mathbf{x} - \delta\mathbf{z}$ are vectors from Q. Since $\mathbf{x} = \frac{1}{2}\mathbf{u} + \frac{1}{2}\mathbf{v}$, the vector $\mathbf{x}$ cannot be a vertex, proving our claim. $\square$

Corollary 2 *The set* $V(Q) = \{\mathbf{x} \in Q \mid A(S(\mathbf{x}))$ *is linearly independent* $\}$ *is the set of vertices of* Q*.*

Proof. Immediate by Lemmas 7 and 8. $\square$

Let us add finally, that since any set of linearly independent columns of $A(S(\mathbf{x}))$ for some $\mathbf{x} \in V(Q)$ can be extended to a basic subset B of columns (of cardinality m, since we assumed that $rank(A) = m$), and since for a basis the equations $B\mathbf{y} = b$ have a unique solution, we must have $\mathbf{x}_B = \mathbf{y}$, and thus we can view $\mathbf{x}$ as a basic feasible solution to the system $A\mathbf{x} = b$, $\mathbf{x} \geq 0$, corresponding to the basis B of A. Thus, all vertices of Q can be viewed as basic feasible solutions. Conversely, any basic feasible solution is a vertex by the above Lemmas.