DCG, Wintersemester 2024/25 | Lecture 2 (5 pages)

Separation theorem and Caratheodory's theorem

Anne Driemel | updated: October 10, 2024

In the last lecture we defined the convex hull of a set $A \subseteq \mathbb{R}^d$ as the set of all convex combinations of points in A. The two theorems of this lecture lead to two different characterizations of the convex hull of a finite number of points. The separation theorem implies that the convex hull of a finite number of points $A \subseteq \mathbb{R}^d$ is the common intersection of all closed halfspaces of $\mathbb{R}^d$ that contain A. Caratheodory's theorem shows that we can write the convex hull of a finite number of points $A \subseteq \mathbb{R}^d$ as a union of simplices formed by all subsets of $d+1$ points of A.

Definition 2.1 (Simplex)**.** *A t-dimensional simplex in $\mathbb{R}^d$ is the convex hull of $t+1$ affinely independent points in $\mathbb{R}^d$.*

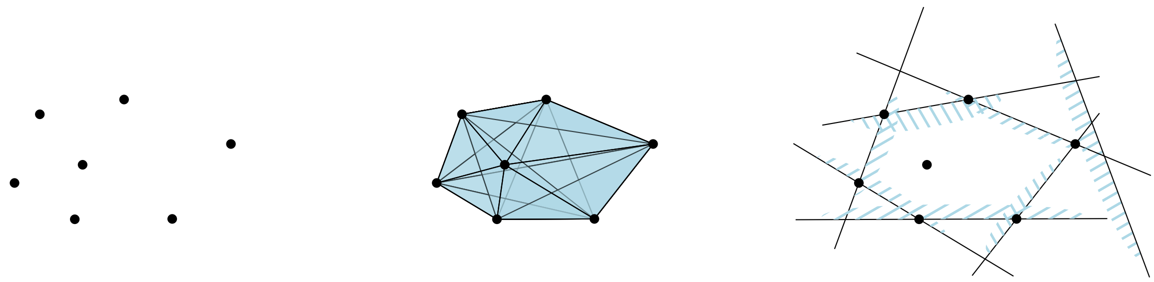


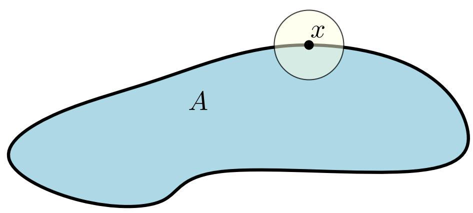
Figure 1: Two characterizations of the convex hull in the plane: (left) the point set; (middle) union of all triangles formed by three points from the set; (right) intersection of all halfplanes that contain the set.

1 Basic notions for sets

We denote with $\|\cdot\|$ the Euclidean norm.

Definition 2.2. *A set $A \subseteq \mathbb{R}^d$ is* bounded *if it is contained in a ball of finite radius. That is, there exist $p \in \mathbb{R}^d$ and $r \in \mathbb{R}$ such that $A \subseteq \{ \ x \in \mathbb{R}^d \ \mid \ \|x-p\| \leq r \ \}$.*

Definition 2.3. *A point $x \in \mathbb{R}^d$ is a* boundary point *of a set $A \subseteq \mathbb{R}^d$ if for any $\epsilon > 0$ there exist points $y \in A$ and $z \in \mathbb{R}^d \setminus A$ such that $\|x-y\| \leq \epsilon$ and $\|x-z\| \leq \epsilon$. We denote with $\partial(A)$ the set of all boundary points of A. We say A is* closed *if it contains its own boundary and A is* open*, if it contains no boundary points.*



Definition 2.4. *A set $A \subseteq \mathbb{R}^d$ is* convex *if for any two $x, y \in A$ it holds that the line segment $\{ \ tx + (1-t)y \ \mid \ t \in [0,1] \ \}$ is also contained in A.*

(a) not convex (b) convex

Observation 2.5. *The convex hull of any set is convex.*

This directly follows from the definitions. Indeed, let $x, y \in \text{conv}(A)$ for some set $A = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d$. Let $t \in [0, 1]$. By using convex combinations of A to rewrite x and y we have

$$tx + (1-t)y = t\left(\sum_{i=1}^{|A|} \alpha_i p_i\right) + (1-t)\left(\sum_{i=1}^{|A|} \beta_i p_i\right) = \sum_{i=1}^{|A|} (t\alpha_i + (1-t)\beta_i)p_i$$

Now we can choose values $\gamma_i = (t\alpha_i + (1-t)\beta_i)$. Clearly, $\gamma_1, \ldots, \gamma_{|A|}$ chosen this way satisfy the conditions of a convex combination of A.

2 Caratheodory's theorem

Theorem 2.6. *Let $A \subseteq \mathbb{R}^d$, then any point in* $\text{conv}(A)$ *can be written as a convex combination of $d+1$ or fewer points of A.*

Proof. Let $q \in \text{conv}(A)$ and assume for the sake of contradiction that q can be written as a convex combination of $m \geq d+2$ points, but not fewer. Let $\alpha_1, \ldots, \alpha_m$ be the coefficients of this convex combination and let $p_1, \ldots, p_m$ be the corresponding points of A, such that

$$q = \sum_{i=1}^{m} \alpha_i p_i \tag{1}$$

The goal is to write q as a convex combination of $m-1$ points of A contradicting the initial statement.

Since $p_1, \ldots, p_m$ are at least $d+2$ points in $\mathbb{R}^d$ they are necessarily affinely dependent, so we can write one of them as an affine combination of the others. Let $i \in \{1, \ldots, m\}$ be the corresponding index of the point, such that

$$p_i = \sum_{j \neq i} \beta_j p_j$$

for some values β_j satisfying the conditions of an affine combination. We set $\beta_i = -1$ and obtain

$$0 = \sum_{j=1}^{m} \beta_j p_j \quad \text{with} \quad 0 = \sum_{j=1}^{m} \beta_j$$

Note that for any $\gamma \in \mathbb{R}$ we can now write

$$0 = \sum_{j=1}^{m} \gamma\beta_j p_j \quad \text{with} \quad 0 = \sum_{j=1}^{m} \gamma\beta_j \tag{2}$$

since the left side is 0 in both equalities. Now, we can subtract Equation (2) from Equation (1) and obtain

$$q - 0 = \sum_{i=1}^{m} \alpha_i p_i - \sum_{i=1}^{m} \gamma\beta_i p_i = \sum_{i=1}^{m} (\alpha_i - \gamma\beta_i) p_i$$

Define $\delta_i := \alpha_i - \gamma\beta_i$. Now we want to choose γ, such that one of the coefficients δ_i becomes 0 and the others satisfy the conditions of a convex combination of the remaining $m-1$ points of A. Note that the sum of coefficients will be 1, regardless of the choice of γ, since

$$\sum_{j=1}^{m} \alpha_i + \sum_{j=1}^{m} \gamma\beta_j = 1$$

Therefore, we only need to make sure that all coefficients are non-negative. This can be done by choosing $\gamma = \frac{\alpha_j}{\beta_j}$ for some $\beta_j > 0$ such that $\frac{\alpha_j}{\beta_j} \leq \frac{\alpha_i}{\beta_i}$ for all $\beta_i > 0$.

There are three cases for β_i:

(i) $\beta_i = 0$: $\delta_i = \alpha_i - \gamma\beta_i = \alpha_i \geq 0$

(ii) $\beta_i < 0$: $\delta_i = \alpha_i - \gamma\beta_i \geq 0$ for $\gamma \geq 0$

(iii) $\beta_i > 0$: $\delta_i = \alpha_i - \gamma\beta_i = \alpha_i - \frac{\alpha_j}{\beta_j}\beta_i \geq \alpha_i - \frac{\alpha_i}{\beta_i}\beta_i = 0$

In all cases, the coefficient δ_i is non-negative. Therefore, the coefficients chosen in this way satisfy the conditions of a convex combination. Moreover, at least one of the coefficients is zero, so the corresponding term can be dropped and we obtain a convex combination of $m-1$ points. A contradiction. $\square$

3 Separation theorem

Theorem 2.7. *Let $C, D \subseteq \mathbb{R}^d$ be sets that are each closed, bounded and convex and assume $C \cap D = \emptyset$. Then, there exists a hyperplane that strictly separates C and D. That is, there exist $a \in \mathbb{R}^d$ and $u \in \mathbb{R}$, such that*

(i) for any $x \in C$ it holds that $\langle a, x\rangle > u$, and

(ii) for any $x \in D$ it holds that $\langle a, x\rangle < u$.

Proof. (Sketch) Let $p \in C$ and $q \in D$ be two points minimizing $\|p - q\|$. That is,

$$\|p - q\| = \min_{x \in C} \min_{y \in D} \|x - y\| \tag{3}$$

Since C and D are both bounded and closed, this minimum exists and p and q are therefore well-defined. Now, consider the hyperplane H which passes through the midpoint $z = \frac{p+q}{2}$ of the line segment connecting p and q, and which has norm $p - q$.

We argue that H does not intersect C, using the same argument we can then show that it does not intersect D, by swapping the roles of C and D.

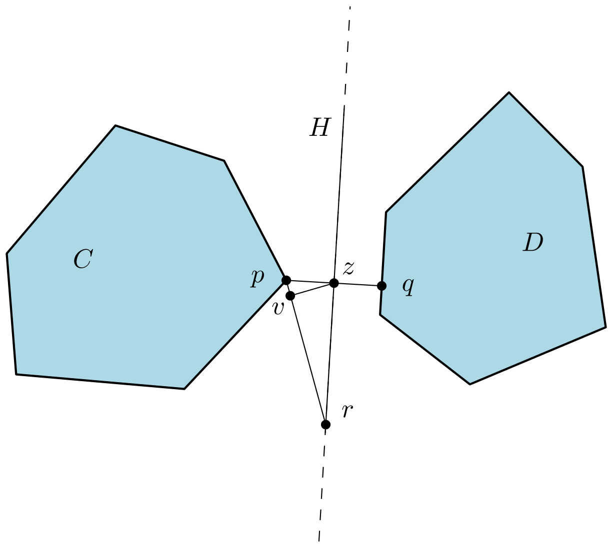
Assume for the sake of contradiction, that H would contain a point r of C, then we argue that we can find a point $v \in C$ such that

$$\|v - q\| < \|p - q\|,$$

contradicting the minimality of $\|p - q\|$ according to (3).

In particular, we choose the point on the line segment

$$\overline{pr} = \{\ (tp + (1-t)r \ \mid\ t \in [0,1]\ \}$$



that minimizes the distance to z. Since C is convex, $\overline{pr}$ must be contained in C, therefore, v is contained in C. Since the triangle (p, v, z) has a right angle at z, the point on the extended edge opposite z that minimizes the distance to z (the foot of the altitude) is contained in the edge. Note that this also holds in higher dimensions, since we only use the geometry of the plane that contains the triangle. It follows that

$$\|p - q\| = \|p - z\| + \|z - q\| < \|v - z\| + \|z - q\| \leq \|v - q\|$$

where the last inequality follows from the triangle inequality.

Therefore, we know that the hyperplane H does not intersect the sets C or D. At the same time, we chose two points $p \in C$ and $q \in D$ on either side of the hyperplane. If there would be a point p' in C on the other side of the hyperplane, then the line segment $\overline{p, p'}$ would intersect the hyperplane. By convexity the line segment is contained in C, so this contradicts what we just established. Likewise, we can argue for D that it lies entirely to one side of the hyperplane. $\square$

The following theorem gives a characterization of the convex hull using the separation theorem.

Theorem 2.8. *Let $A = \{p_1, \ldots, p_n\} \subseteq \mathbb{R}^d$ be a finite set, it holds that* $\operatorname{conv}(A)$ *is equal to the common intersection of all closed halfspaces that contain A.*

Proof. We show equality of the two sets by showing inclusion in both directions. We first show that any closed halfspace H that contains A also contains $\operatorname{conv}(A)$. Let $H(a, b)$ denote the halfspace $\{ x \in \mathbb{R}^d \mid \langle a, x \rangle \geq b \}$. Let $a \in \mathbb{R}^d$ and $b \in \mathbb{R}$ such that $A \subseteq H(a, b)$. Let x be any point in $\operatorname{conv}(A)$. By definition, x can be written as a convex combination defined by values $\alpha_1, \ldots, \alpha_{|A|}$. Now $A \subseteq H$ implies

$$\langle a, x \rangle = \left\langle a, \sum_{i=1}^{|A|} \alpha_i p_i \right\rangle = \sum_{i=1}^{|A|} \alpha_i \langle a, p_i \rangle \geq \sum_{i=1}^{|A|} \alpha_i b = b$$

which says that $x \in H(a, b)$. This implies that

$$\operatorname{conv}(A) \subseteq \bigcap_{\substack{a \in \mathbb{R}^d, b \in \mathbb{R} \\ A \subseteq H(a,b)}} H(a, b)$$

Next, we show the other direction. We claim that

$$\bigcap_{\substack{a\in\mathbb{R}^d, b\in\mathbb{R}\\ A\subseteq H(a,b)}} H(a,b) \subseteq \operatorname{conv}(A)$$

We show this claim by contradiction. Assume there exists an x that is contained in all closed halfspaces that contain A, but $x \notin \operatorname{conv}(A)$. We apply the separation theorem to the two sets $\{x\}$ and $\operatorname{conv}(A)$. In order to apply Theorem 2.7, we need to ensure that $\operatorname{conv}(A)$ is bounded, closed and convex. Convexity was observed above. The convex hull is certainly bounded, since it can be shown that

$$\operatorname{conv}(A) \subseteq \left\{ x \in \mathbb{R}^d \;\middle|\; \|x\| \leq \sqrt{d} \cdot C_{\max} \right\}$$

where $C_{\max}$ is an upper bound on the absolute value of any coordinate of any point in A.

The convex hull $\operatorname{conv}(A)$ is closed, since it can be described as a finite union of simplices. Any simplex is a closed set and the finite union of closed sets is also closed.

Now, the separation theorem implies that there exists a hyperplane that separates x from $\operatorname{conv}(A)$. This would imply that there exists a closed halfspace that contains A, but not x, which contradicts our initial definition of x being contained in all halfspaces that contain A. $\square$

References

- Jiří Matoušek, Chapter 1, Lectures on Discrete Geometry, Springer Graduate Texts in Mathematics.