

Hyperplane separation theorem

In geometry, the **hyperplane separation theorem** is a theorem about disjoint convex sets in n -dimensional Euclidean space. There are several rather similar versions. In one version of the theorem, if both these sets are closed and at least one of them is compact, then there is a hyperplane in between them and even two parallel hyperplanes in between them separated by a gap. In another version, if both disjoint convex sets are open, then there is a hyperplane in between them, but not necessarily any gap. An axis which is orthogonal to a separating hyperplane is a **separating axis**, because the orthogonal projections of the convex bodies onto the axis are disjoint.

The hyperplane separation theorem is due to Hermann Minkowski. The Hahn–Banach separation theorem generalizes the result to topological vector spaces.

A related result is the supporting hyperplane theorem.

In the context of support-vector machines, the *optimally separating hyperplane* or *maximum-margin hyperplane* is a hyperplane which separates two convex hulls of points and is equidistant from the two.^{[1][2][3]}

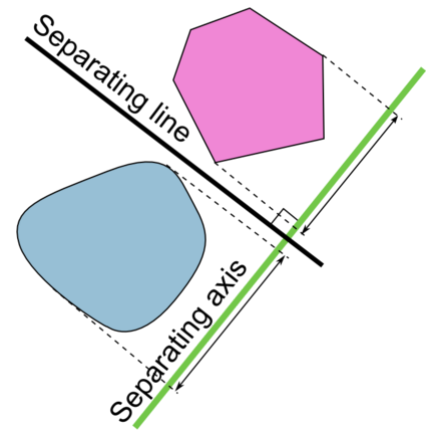


Illustration of the hyperplane separation theorem.

Contents

[Statements and proof](#)

[Converse of theorem](#)

[Counterexamples and uniqueness](#)

[Use in collision detection](#)

[See also](#)

[Notes](#)

[References](#)

[External links](#)

Statements and proof

Hyperplane separation theorem^[4] — Let A and B be two disjoint nonempty convex subsets of \mathbf{R}^n . Then there exist a nonzero vector v and a real number c such that

$$\langle x, v \rangle \geq c \text{ and } \langle y, v \rangle \leq c$$

for all x in A and y in B ; i.e., the hyperplane $\langle \cdot, v \rangle = c$, v the normal vector, separates A and B .

The proof is based on the following lemma:

Lemma — Let K be a nonempty closed convex subset of \mathbf{R}^n . Then there exists a unique vector in K of minimum norm (length).

Proof of lemma: Let $\delta = \inf\{|x| : x \in K\}$. Let x_j be a sequence in K such that $|x_j| \rightarrow \delta$. Note that $(x_i + x_j)/2$ is in K since K is convex and so $|x_i + x_j|^2 \geq 4\delta^2$. Since

$$|x_i - x_j|^2 = 2|x_i|^2 + 2|x_j|^2 - |x_i + x_j|^2 \leq 2|x_i|^2 + 2|x_j|^2 - 4\delta^2 \rightarrow 0$$

as $i, j \rightarrow \infty$, x_i is a Cauchy sequence and so has limit x in K . It is unique since if y is in K and has norm δ , then $|x - y|^2 \leq 2|x|^2 + 2|y|^2 - 4\delta^2 = 0$ and $x = y$. \square

Proof of theorem: Given disjoint nonempty convex sets A, B , let

$$K = A + (-B) = \{x - y \mid x \in A, y \in B\}.$$

Since $-B$ is convex and the sum of convex sets is convex, K is convex. By the lemma, the closure \overline{K} of K , which is convex, contains a vector v of minimum norm. Since \overline{K} is convex, for any n in K , the line segment

$$v + t(n - v), 0 \leq t \leq 1$$

lies in \overline{K} and so

$$|v|^2 \leq |v + t(n - v)|^2 = |v|^2 + 2t\langle v, n - v \rangle + t^2|n - v|^2.$$

For $0 < t \leq 1$, we thus have:

$$0 \leq 2\langle v, n \rangle - 2|v|^2 + t|n - v|^2$$

and letting $t \rightarrow 0$ gives: $\langle n, v \rangle \geq |v|^2$. Hence, for any x in A and y in B , we have: $\langle x - y, v \rangle \geq |v|^2$. Thus, if v is nonzero, the proof is complete since

$$\inf_{x \in A} \langle x, v \rangle \geq |v|^2 + \sup_{y \in B} \langle y, v \rangle.$$

More generally (covering the case $v = 0$), let us first take the case when the interior of K is nonempty. The interior can be exhausted by a nested sequence of nonempty compact convex subsets $K_1 \subset K_2 \subset K_3 \subset \dots$ (namely, put $K_j \equiv [-j, j]^n \cap \{x \in \text{int}(K) : d(x, (\text{int}(K))^c) \geq \frac{1}{j}\}$). Since 0 is not in K , each K_n contains a nonzero vector v_n of minimum length and by the argument in the

early part, we have: $\langle x, v_n \rangle \geq 0$ for any $x \in K_n$. We can normalize the v_n 's to have length one. Then the sequence v_n contains a convergent subsequence (because the n -sphere is compact) with limit v , which is nonzero. We have $\langle x, v \rangle \geq 0$ for any x in the interior of K and by continuity the same holds for all x in K . We now finish the proof as before. Finally, if K has empty interior, the affine set that it spans has dimension less than that of the whole space. Consequently K is contained in some hyperplane $\langle \cdot, v \rangle = c$; thus, $\langle x, v \rangle \geq c$ for all x in K and we finish the proof as before. \square

The number of dimensions must be finite. In infinite-dimensional spaces there are examples of two closed, convex, disjoint sets which cannot be separated by a closed hyperplane (a hyperplane where a *continuous* linear functional equals some constant) even in the weak sense where the inequalities are not strict.^[5]

The above proof also proves the first version of the theorem mentioned in the lede (to see it, note that K in the proof is closed under the hypothesis of the theorem below.)

Separation theorem I — Let A and B be two disjoint nonempty closed convex sets, one of which is compact. Then there exist a nonzero vector v and real numbers $c_1 < c_2$ such that

$$\langle x, v \rangle > c_2 \text{ and } \langle y, v \rangle < c_1$$

for all x in A and y in B .

Here, the compactness in the hypothesis cannot be relaxed; see an example in the next section. This version of the separation theorem does generalize to infinite-dimension; the generalization is more commonly known as the Hahn–Banach separation theorem.

We also have:

Separation theorem II — Let A and B be two disjoint nonempty convex sets. If A is open, then there exist a nonzero vector v and real number c such that

$$\langle x, v \rangle > c \text{ and } \langle y, v \rangle \leq c$$

for all x in A and y in B . If both sets are open, then there exist a nonzero vector v and real number c such that

$$\langle x, v \rangle > c \text{ and } \langle y, v \rangle < c$$

for all x in A and y in B .

This follows from the standard version since the separating hyperplane cannot intersect the interiors of the convex sets.

Converse of theorem

Note that the existence of a hyperplane that only "separates" two convex sets in the weak sense of both inequalities being non-strict obviously does not imply that the two sets are disjoint. Both sets could have points located on the hyperplane.

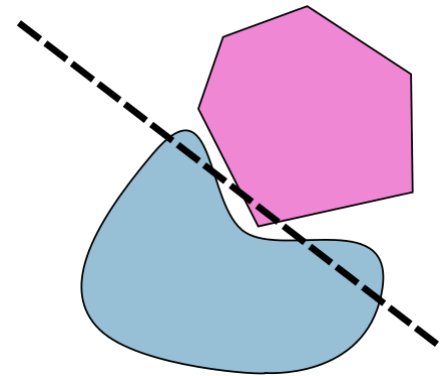
Counterexamples and uniqueness

If one of A or B is not convex, then there are many possible counterexamples. For example, A and B could be concentric circles. A more subtle counterexample is one in which A and B are both closed but neither one is compact. For example, if A is a closed half plane and B is bounded by one arm of a hyperbola, then there is no strictly separating hyperplane:

$$A = \{(x, y) : x \leq 0\}$$

$$B = \{(x, y) : x > 0, y \geq 1/x\}.$$

(Although, by an instance of the second theorem, there is a hyperplane that separates their interiors.) Another type of counterexample has A compact and B open. For example, A can be a closed square and B can be an open square that touches A .



The **theorem does not apply** if one of the bodies is not convex.

In the first version of the theorem, evidently the separating hyperplane is never unique. In the second version, it may or may not be unique. Technically a separating axis is never unique because it can be translated; in the second version of the theorem, a separating axis can be unique up to translation.

Use in collision detection

The **separating axis theorem** (SAT) says that:

Two convex objects do not overlap if there exists a line (called axis) onto which the two objects' projections do not overlap.

SAT suggests an algorithm for testing whether two convex solids intersect or not.

Regardless of dimensionality, the separating axis is always a line. For example, in 3D, the space is separated by planes, but the separating axis is perpendicular to the separating plane.

The separating axis theorem can be applied for fast collision detection between polygon meshes. Each face's normal or other feature direction is used as a separating axis. Note that this yields possible separating axes, not separating lines/planes.

In 3D, using face normals alone will fail to separate some edge-on-edge non-colliding cases. Additional axes, consisting of the cross-products of pairs of edges, one taken from each object, are required.^[6]

For increased efficiency, parallel axes may be calculated as a single axis.

See also

- Dual cone

- [Farkas's lemma](#)
- [Kirchberger's theorem](#)
- [Optimal control](#)

Notes

1. [Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome \(2008\). *The Elements of Statistical Learning : Data Mining, Inference, and Prediction* \(<https://web.stanford.edu/~hastie/Papers/ESLII.pdf#page=148>\) \(PDF\) \(Second ed.\). New York: Springer. pp. 129–135.](#)
2. [Witten, Ian H.; Frank, Eibe; Hall, Mark A.; Pal, Christopher J. \(2016\). *Data Mining: Practical Machine Learning Tools and Techniques* \(<https://books.google.com/books?id=1SylCgAAQBAJ&pg=PA253>\) \(Fourth ed.\). Morgan Kaufmann. pp. 253–254. ISBN 9780128043578.](#)
3. [Deisenroth, Marc Peter; Faisal, A. Aldo; Ong, Cheng Soon \(2020\). *Mathematics for Machine Learning*. Cambridge University Press. pp. 337–338. ISBN 978-1-108-45514-5.](#)
4. [Boyd & Vandenberghe 2004](#), Exercise 2.22.
5. [Haïm Brezis, *Analyse fonctionnelle : théorie et applications*, 1983, remarque 4, p. 7.](#)
6. https://docs.godotengine.org/en/stable/tutorials/math/vectors_advanced.html#collision-detection-in-3d

References

- [Boyd, Stephen P.; Vandenberghe, Lieven \(2004\). *Convex Optimization* \(\[https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf\]\(https://web.stanford.edu/~boyd/cvxbook/bv_cvxbook.pdf\)\) \(PDF\). Cambridge University Press. ISBN 978-0-521-83378-3.](#)
- [Golshtein, E. G.; Tretyakov, N.V. \(1996\). *Modified Lagrangians and monotone maps in optimization*. New York: Wiley. p. 6. ISBN 0-471-54821-9.](#)
- [Shimizu, Kiyotaka; Ishizuka, Yo; Bard, Jonathan F. \(1997\). *Nondifferentiable and two-level mathematical programming*. Boston: Kluwer Academic Publishers. p. 19. ISBN 0-7923-9821-1.](#)

External links

- [Collision detection and response \(http://www.metanetsoftware.com/technique/tutorialA.html\)](http://www.metanetsoftware.com/technique/tutorialA.html)

Retrieved from "https://en.wikipedia.org/w/index.php?title=Hyperplane_separation_theorem&oldid=1050344372"

This page was last edited on 17 October 2021, at 08:10 (UTC).

Text is available under the Creative Commons Attribution-ShareAlike License 3.0; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.