

18.3 Separating Hyperplanes

Sections 01 and 02 looked at a convex set from the inside: which combinations of its points it contains, and how few points a combination needs. This section looks at it from the outside. A point that is not in a closed convex set can be cut off from it by a hyperplane, a set can be touched at its edge by a hyperplane that leaves the whole set on one side, and two disjoint convex sets can be pulled apart by one. All three facts come from a single construction, the point of the set nearest to a given point, and that construction is Chapter 11's orthogonal projection with the subspace replaced by a convex set.

Throughout, V is a finite-dimensional real inner product space, with inner product $\langle \cdot, \cdot \rangle$ (Definition 11.1.1), induced norm $\|\cdot\|$ (Definition 11.1.3), and $n = \dim V$. A sequence **converges**, $x_k \rightarrow x$, when $\|x_k - x\| \rightarrow 0$; **closed** means closed in the sense of Definition 18.1.6, so that a closed set contains the limit of every convergent sequence of its points. A set is **compact** if every sequence in it has a subsequence converging to a point of the set; this is the word of Corollary 16.2.5, which says that every closed bounded subset of V is compact. The field is $\mathbb{R}$ because the whole section is about inequalities.

18.3.1 The nearest point of a closed convex set

Chapter 11 answered a question about subspaces: given v and a finite-dimensional subspace U , which vector of U is closest to v ? Theorem 11.3.5 said there is exactly one, the orthogonal projection $P_U v$, and that $v - P_U v$ is perpendicular to U . Now replace U by a set with less structure, such as a disc, a box, or a half-plane, and ask the same question. Two small experiments in $\mathbb{R}^2$ show which hypotheses the answer needs.

- The **open** disc $\{y : \|y\| < 1\}$ and the point $x = (2, 0)$. The points $(1 - 1/k, 0)$ of the disc come closer and closer to x , but the distance 1 is never reached, since the only candidate $(1, 0)$ is missing. There is no nearest point. The set is convex but not closed.
- The **circle** $\{y : \|y\| = 1\}$ and the point $x = 0$. Every point of the circle is at distance 1, so every point is a nearest point. The set is closed but not convex.

So we ask for both properties, and then everything works.

Theorem 18.3.1. Nearest Point Theorem

Let C be a **nonempty, closed, convex** subset of V , and let $x \in V$.

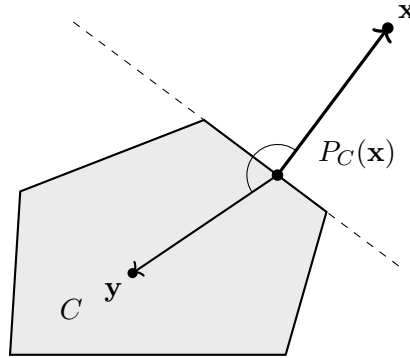
- There is **exactly one** point $p \in C$ with $\|x - p\| \leq \|x - y\|$ for every $y \in C$.
- A point $p \in C$ is this nearest point **if and only if**

$$\langle x - p, y - p \rangle \leq 0 \quad \text{for every } y \in C. \quad (18.3.1)$$

The nearest point is written $P_C(x)$, and the map $P_C: V \rightarrow C$ is the **projection onto C** . The number $d(x, C) = \|x - P_C(x)\|$ is the **distance** from x to C .

Idea. Three jobs, three tools. ① Existence: a nearest point is a minimum of the continuous function $y \mapsto \|x - y\|$, and a minimum exists on a compact set; C itself may be unbounded, so cut it down to the part of C within a fixed distance of x . ② The inequality Eq. 18.3.1 is the first-order condition: slide from p a little way towards another point y of C , which convexity allows, and ask that the distance not decrease. ③ Read backwards, the inequality

says that the angle at p between $x - p$ and $y - p$ is at least a right angle, and expanding $\|x - y\|^2$ then gives $\|x - y\|^2 \geq \|x - p\|^2 + \|y - p\|^2$. That inequality proves both that p is nearest and that nothing else is.



the angle at $P_C(\mathbf{x})$ between $\mathbf{x} - P_C(\mathbf{x})$ and $\mathbf{y} - P_C(\mathbf{x})$ is obtuse;
the dashed line through $P_C(\mathbf{x})$ is perpendicular to $\mathbf{x} - P_C(\mathbf{x})$

Proof. Existence. Since $C \neq \emptyset$, pick $y_0 \in C$ and put $R = \|x - y_0\|$. Let

$$K = \{y \in C : \|x - y\| \leq R\}.$$

Then $y_0 \in K$, and K is bounded, since $\|y\| \leq \|x\| + \|y - x\| \leq \|x\| + R$ for $y \in K$ by Corollary 11.1.6. It is closed: if $y_k \in K$ and $y_k \rightarrow y$, then $y \in C$ because C is closed, and

$$\|x - y\| \leq \|x - y_k\| + \|y_k - y\| \leq R + \|y_k - y\|$$

for every k ; the right-hand side tends to R , and a nonstrict inequality survives a limit, so $\|x - y\| \leq R$. By Corollary 16.2.5, the form in which Chapter 16 carried **the compactness of closed bounded sets, fact (A3) of Chapter 16's introduction**, over to every finite-dimensional space, K is compact. The function $f(y) = \|x - y\|$ is continuous, since $|f(y) - f(y')| \leq \|y - y'\|$ by Lemma 16.1.1. By **the extreme value theorem, fact (A4) of Chapter 16's introduction**, f attains a minimum on K at some $p \in K$. Every $y \in C \setminus K$ has $f(y) > R \geq f(p)$. So $\|x - p\| \leq \|x - y\|$ for every $y \in C$.

(b), ($\Rightarrow$). Let p be a nearest point and let $y \in C$. For $0 < t \leq 1$, the point $p + t(y - p) = (1 - t)p + ty$ lies in C by Definition 18.1.1. Hence, expanding with Theorem 11.1.2 (c) over $\mathbb{R}$,

$$\|x - p\|^2 \leq \|(x - p) - t(y - p)\|^2 = \|x - p\|^2 - 2t\langle x - p, y - p \rangle + t^2\|y - p\|^2.$$

Cancel $\|x - p\|^2$ and divide by $2t > 0$: $\langle x - p, y - p \rangle \leq \frac{t}{2}\|y - p\|^2$. This holds for every $t = 1/k$, and letting $k \rightarrow \infty$ gives [Eq. 18.3.1](#).

(b), ($\Leftarrow$), and uniqueness. Let $p \in C$ satisfy [Eq. 18.3.1](#), and let $y \in C$. The same expansion with $t = 1$ gives

$$\begin{aligned} \|x - y\|^2 &= \|x - p\|^2 - 2\langle x - p, y - p \rangle + \|y - p\|^2 \\ &\geq \|x - p\|^2 + \|y - p\|^2. \end{aligned} \tag{*}$$

So $\|x - y\| \geq \|x - p\|$, and p is a nearest point. Finally, the nearest point p found in the existence step satisfies [Eq. 18.3.1](#) by ($\Rightarrow$), so (*) applies to it; for $y \in C$ with $y \neq p$ it gives $\|x - y\|^2 \geq \|x - p\|^2 + \|y - p\|^2 > \|x - p\|^2$, and y is not a nearest point. This proves the theorem.



The inequality (*) is Pythagoras with the equality relaxed to an inequality. That is exactly the relation with Chapter 11, and it deserves to be said precisely.

This theorem generalizes Theorem 11.3.5. Let $C = U$ be a subspace of V . It is convex, and closed by Proposition 18.1.6 (c). For $p \in U$, the vector $y - p$ runs over **all** of U as y does, and with each $u \in U$ it also takes the value $-u$. So Eq. 18.3.1 says $\langle x - p, u \rangle \leq 0$ and $-\langle x - p, u \rangle \leq 0$, that is,

$$\langle x - p, u \rangle = 0 \quad \text{for every } u \in U.$$

For a subspace the variational inequality becomes an equality, the statement $x - p \in U^\perp$. Then $x = p + (x - p)$ is the splitting of Theorem 11.3.2 (a), so $P_C(x) = P_U x$ by Definition 11.3.2, and (*) becomes the equation of Theorem 11.1.3 used in the proof of Theorem 11.3.5. The projection onto a closed convex set is the orthogonal projection with “perpendicular” weakened to “at an obtuse angle”.

Example. Three projections

Find $P_C(x)$ in each case, and check Eq. 18.3.1.

- (a) C is the closed unit ball $\{y : \|y\| \leq 1\}$ of V , and $\|x\| > 1$.
- (b) C is the box $[0, 1]^3 = \{y \in \mathbb{R}^3 : 0 \leq y_i \leq 1 \text{ for each } i\}$ with the dot product, and $\mathbf{x} = (3, -1, 2)$.
- (c) C is the half-space $\{y : \langle a, y \rangle \leq c\}$, where $a \neq 0$, and $\langle a, x \rangle > c$.

Solution.

- (a) $P_C(x) = x/\|x\|$. With $p = x/\|x\|$ we have $\|p\| = 1$ and $x - p = (\|x\| - 1)p$, so for $y \in C$

$$\langle x - p, y - p \rangle = (\|x\| - 1)(\langle p, y \rangle - 1) \leq 0,$$

because $\|x\| - 1 > 0$ and $\langle p, y \rangle \leq \|p\|\|y\| \leq 1$ by Theorem 11.1.5. By Theorem 18.3.1 (b), p is the nearest point, and $d(x, C) = \|x\| - 1$.

- (b) Clamp each coordinate into $[0, 1]$: $\mathbf{p} = (1, 0, 1)$. Then $\mathbf{x} - \mathbf{p} = (2, -1, 1)$ and, for $\mathbf{y} \in C$,

$$\langle \mathbf{x} - \mathbf{p}, \mathbf{y} - \mathbf{p} \rangle = 2(y_1 - 1) - y_2 + (y_3 - 1) \leq 0,$$

since $y_1 \leq 1$, $y_2 \geq 0$, and $y_3 \leq 1$. So $P_C(\mathbf{x}) = (1, 0, 1)$ and $d(\mathbf{x}, C) = \sqrt{4 + 1 + 1} = \sqrt{6}$. The vector $\mathbf{a} = \mathbf{x} - \mathbf{p} = (2, -1, 1)$ also cuts $\mathbf{x}$ off: $\langle \mathbf{a}, \mathbf{x} \rangle = 9$ and $\|\mathbf{a}\|^2 = 6$, so $c = \langle \mathbf{a}, \mathbf{x} \rangle - \frac{1}{2}\|\mathbf{a}\|^2 = 6$, and the plane $2y_1 - y_2 + y_3 = 6$ strictly separates C from $\mathbf{x}$, since its left side is at most $2 + 0 + 1 = 3$ on C and equals 9 at $\mathbf{x}$. Theorem 18.3.2 below turns this recipe into a theorem.

- (c) Put $\lambda = (\langle a, x \rangle - c)/\|a\|^2$, which is positive, and $p = x - \lambda a$. Then $\langle a, p \rangle = \langle a, x \rangle - \lambda\|a\|^2 = c$, so $p \in C$, and for $y \in C$

$$\langle x - p, y - p \rangle = \lambda(\langle a, y \rangle - c) \leq 0.$$

So $P_C(x) = x - \lambda a$: drop the perpendicular to the boundary hyperplane, as Chapter 11 would have done. The distance is $\lambda\|a\| = (\langle a, x \rangle - c)/\|a\|$.

Two degenerate cases fix the scale. If $C = \{c\}$ is a single point, then $P_C(x) = c$ for every x , and Eq. 18.3.1 holds because $y - p = 0$. If $C = V$, then P_C is the identity. In every case $P_C(x) = x$ exactly when $x \in C$.

Quick check. Where did the proof of Theorem 18.3.1 use that C is closed, and where that it is convex? Match each use with one of the two failing experiments before the theorem.

Solution. Closedness was used in the existence step, to make K closed and hence compact; the open disc fails there, since the minimizing candidates converge to a point outside it. Convexity was used in $(\Rightarrow)$, to know that $(1 - t)p + ty \in C$, and uniqueness rests on that step. The circle fails there: with $\mathbf{x} = \mathbf{0}$, $\mathbf{p} = \mathbf{e}_1$, and $\mathbf{y} = -\mathbf{e}_1$, $\langle \mathbf{x} - \mathbf{p}, \mathbf{y} - \mathbf{p} \rangle = 2 > 0$.

Warning. **The proof needs an inner product, and uniqueness can fail without one.** In $\mathbb{R}^2$ with $\|\cdot\|_\infty$, let C be the x_1 -axis, which is closed and convex, and $\mathbf{x} = (0, 1)$. Then $\|\mathbf{x} - (t, 0)\|_\infty = \max(|t|, 1) = 1$ for every $t \in [-1, 1]$, so there is a whole segment of nearest points. Existence survives, since its proof used only (N3) and compactness. Uniqueness does not, because it came from the expansion of $\|\mathbf{x} - \mathbf{y}\|^2$, which is available only for induced norms.

18.3.2 Hyperplanes and half-spaces

A hyperplane in the sense of Chapter 5 passes through 0: it is a subspace of codimension 1 (Definition 5.5.2). To separate sets we need translates of hyperplanes, and Chapter 5 §06 called a coset $p + H$ of a hyperplane H (Definition 4.4.2) an **affine hyperplane**. In an inner product space these have a concrete form. By Theorem 5.5.3, $H = \ker \varphi$ for a nonzero $\varphi \in V^*$; by Theorem 11.5.1, $\varphi = \langle \cdot, a \rangle$ for a nonzero $a \in V$. Hence

$$p + H = \{y \in V : \langle a, y \rangle = c\}, \quad c = \langle a, p \rangle,$$

and conversely every such set with $a \neq 0$ is an affine hyperplane, with a as a **normal vector**. It cuts V into the two **closed half-spaces** $\{\langle a, \cdot \rangle \leq c\}$ and $\{\langle a, \cdot \rangle \geq c\}$. Both are convex, because $\langle a, \cdot \rangle$ is linear, and closed by Proposition 18.1.6 (c).

Definition 18.3.1. Separation by a Hyperplane

Let $A, B \subseteq V$ be nonempty. A **nonzero** vector $a \in V$ and a number $c \in \mathbb{R}$ **separate** A and B if

$$\langle a, x \rangle \leq c \leq \langle a, y \rangle \quad \text{for all } x \in A \text{ and all } y \in B.$$

They **strictly separate** A and B if in addition there is a real $\varepsilon > 0$ with $\langle a, x \rangle \leq c - \varepsilon$ and $\langle a, y \rangle \geq c + \varepsilon$ for all $x \in A, y \in B$.

In words: A lies in one closed half-space of $\{\langle a, \cdot \rangle = c\}$ and B in the other, and strict separation puts a slab of width $2\varepsilon/\|a\|$ between them. Replacing (a, c) by $(-a, -c)$ swaps the roles of A and B . With $a = 0$ and $c = 0$ any two sets would be “separated”.

18.3.3 Cutting off a point

The first consequence of [Theorem 18.3.1](#) is the one the chapter uses most. If x is not in C , the vector from $P_C(x)$ to x is a normal vector for a hyperplane that cuts x off.

Theorem 18.3.2. Separating a Point from a Closed Convex Set

Let $C \subseteq V$ be nonempty, closed, and convex, and let $x \notin C$. Put $a = x - P_C(x)$. Then $a \neq 0$ and

$$\langle a, y \rangle \leq \langle a, x \rangle - \|a\|^2 \quad \text{for every } y \in C.$$

In particular a and $c = \langle a, x \rangle - \frac{1}{2}\|a\|^2$ strictly separate C and $\{x\}$.

Proof. Write $p = P_C(x)$. Since $p \in C$ and $x \notin C$, $a = x - p \neq 0$. For $y \in C$, [Eq. 18.3.1](#) gives $\langle a, y - p \rangle \leq 0$, hence

$$\langle a, y \rangle \leq \langle a, p \rangle = \langle a, x \rangle - \langle a, x - p \rangle = \langle a, x \rangle - \|a\|^2.$$

With $\varepsilon = \frac{1}{2}\|a\|^2 > 0$ this reads $\langle a, y \rangle \leq c - \varepsilon$, while $\langle a, x \rangle = c + \varepsilon$. This proves the theorem. ■

In the picture after [Theorem 18.3.1](#), the hyperplane of the theorem is the line parallel to the dashed one, halfway between $P_C(x)$ and x .

The theorem turns into a description of a closed convex set from the outside. Sections 01 and 02 built convex sets up from points; this builds them down from half-spaces.

Corollary 18.3.3. A Closed Convex Set Is an Intersection of Half-Spaces

Every nonempty closed convex subset C of V is the intersection of all the closed half-spaces $\{y : \langle a, y \rangle \leq c\}$, with $a \neq 0$, that contain C .

Proof. Each of these half-spaces contains C , so C lies in their intersection. Conversely, let $x \notin C$. By [Theorem 18.3.2](#), with $a = x - P_C(x) \neq 0$, the half-space $\{y : \langle a, y \rangle \leq \langle a, x \rangle - \|a\|^2\}$ contains C but not x . So x is not in the intersection. (If $C = V$, the family of half-spaces is empty and its intersection is V by convention.) ■

The corollary has a converse that is immediate: an intersection of closed half-spaces is closed and convex, because each half-space is and both properties pass to intersections. So the closed convex sets are **exactly** the intersections of closed half-spaces. §04 turns this into a formula, the support function, and §09 studies the case of finitely many half-spaces.

18.3.4 Supporting hyperplanes

A point on the edge of C cannot be cut off, because it belongs to C . The best one can hope for is a hyperplane through the point with all of C on one side, touching C without cutting it. To

say “on the edge” we need two more words.

Definition 18.3.2. Interior Point, Closure, Boundary Point

Let $S \subseteq V$. A point $x \in V$ is an **interior point** of S if there is a real $r > 0$ such that **every** y with $\|y - x\| < r$ lies in S ; the set of interior points is written $\text{int}(S)$. The **closure** $\bar{S}$ of S is the set of limits of all convergent sequences of points of S . A **boundary point** of S is a point of $\bar{S}$ that is not an interior point of S .

Constant sequences show $S \subseteq \bar{S}$, and by Definition 18.1.6, S is closed exactly when $\bar{S} = S$. By Theorem 16.2.2 neither notion depends on the norm used: a ball of radius r for one norm contains a ball of positive radius for any other. For the closed unit disc of $\mathbb{R}^2$, the interior points are those with $\|\mathbf{x}\| < 1$, and the boundary points form the unit circle. A segment in $\mathbb{R}^2$ has no interior points at all, so every one of its points is a boundary point; that is the case the next lemma isolates.

Lemma 18.3.4. A Convex Set Is Solid or Flat

Let $V \neq \{0\}$ and let $C \subseteq V$ be nonempty and convex. Then **either** C has an interior point, **or** C lies in an affine hyperplane: there are $a \neq 0$ and $c \in \mathbb{R}$ with $\langle a, y \rangle = c$ for every $y \in C$.

Idea. Fix $c_0 \in C$ and look at the directions $y - c_0$ available inside C . If they span a proper subspace, a normal vector to that subspace is the a we want. If they span V , then C contains $n + 1$ points in general position, hence the solid simplex they span, and the center of a solid simplex is an interior point: a point of the simplex is $\sum_i t_i v_i$ with $t_i \geq 0$ summing to 1, the t_i are determined by the point and depend continuously on it, and at the center they are all equal and positive, so every small enough move keeps them positive.

Proof. Fix $c_0 \in C$ and let $W = \text{span}\{y - c_0 : y \in C\}$.

Case 1: $W \neq V$. Then $\dim W < n$, so $\dim W^\perp = n - \dim W \geq 1$ by Theorem 11.3.2 (c), and there is $a \neq 0$ in $W^\perp$. For $y \in C$, $\langle a, y - c_0 \rangle = 0$, that is, $\langle a, y \rangle = \langle a, c_0 \rangle$. Take $c = \langle a, c_0 \rangle$.

Case 2: $W = V$. Choose a linearly independent list $(y_1 - c_0, \dots, y_m - c_0)$ with every $y_i \in C$ and m as large as possible; $m \leq n$ by Theorem 1.6.3. Every vector $y - c_0$ with $y \in C$ lies in the span of this list, since otherwise Lemma 1.5.3 would lengthen it. So the span contains every generator of W , hence equals V , and the list is a basis; thus $m = n$. Let $\varphi_i(v)$ be the i -th coordinate of v in this basis, a linear functional by Theorem 1.5.2, and write $\varphi_i = \langle \cdot, w_i \rangle$ by Theorem 11.5.1. Put $\mu_i(v) = \varphi_i(v - c_0)$. Then for every $v \in V$,

$$v = \left(1 - \sum_{i=1}^n \mu_i(v)\right) c_0 + \sum_{i=1}^n \mu_i(v) y_i,$$

and when $\mu_i(v) \geq 0$ for every i and $\sum_i \mu_i(v) \leq 1$, this is a convex combination of points of C , hence a point of C : by Theorem 18.1.3 it lies in $\text{conv } C$, and $\text{conv } C = C$ by Definition 18.1.4, since C is itself a convex set containing C .

Let $z = c_0 + \frac{1}{n+1} \sum_i (y_i - c_0)$, so that $\mu_i(z) = \frac{1}{n+1}$ for every i . Each $w_i \neq 0$, because $\varphi_i(y_i - c_0) = 1$; put $L = \max_i \|w_i\| > 0$ and $r = 1/(n(n+1)L)$. If $\|v - z\| < r$, then by

Theorem 11.1.5

$$|\mu_i(v) - \mu_i(z)| = |\langle v - z, w_i \rangle| \leq L\|v - z\| < \frac{1}{n(n+1)},$$

so $\mu_i(v) > \frac{1}{n+1} - \frac{1}{n(n+1)} \geq 0$ and $\sum_i \mu_i(v) < \frac{n}{n+1} + \frac{1}{n+1} = 1$. Hence $v \in C$, and z is an interior point of C . ■

The second ingredient says that from an interior point one can see the whole closure.

Lemma 18.3.5. Closure and Interior of a Convex Set

Let $C \subseteq V$ be convex.

- (a) $\overline{C}$ is closed and convex.
- (b) If $z \in \text{int}(C)$, $w \in \overline{C}$, and $0 < t \leq 1$, then $(1-t)w + tz \in \text{int}(C)$.

Idea. Part (a) pushes limits through: a limit of limits is a limit, and a convex combination of limits is the limit of convex combinations. For (b), picture a small ball around z inside C , and the cone from w over it. The cone contains a smaller ball around $(1-t)w + tz$; since w is only a limit, replace it by a nearby point w' of C , which shrinks the ball by half.

Proof.

- (a) Let $x_k \in \overline{C}$ with $x_k \rightarrow x$. Each x_k is a limit of points of C , so there is $c_k \in C$ with $\|c_k - x_k\| < 1/k$. Then $\|c_k - x\| \leq 1/k + \|x_k - x\| \rightarrow 0$, so $x \in \overline{C}$. For convexity, let $x = \lim c_k$ and $y = \lim d_k$ with $c_k, d_k \in C$, and $0 \leq s \leq 1$. The points $sc_k + (1-s)d_k$ lie in C and converge to $sx + (1-s)y$, by Corollary 11.1.6, so that point lies in $\overline{C}$.
- (b) Take $\rho > 0$ with every point within distance ρ of z in C , and put $p = (1-t)w + tz$. If $t = 1$ then $p = z$ and there is nothing to prove, so let $t < 1$. Choose $w' \in C$ with $(1-t)\|w - w'\| < t\rho/2$, possible because w is a limit of points of C . Let $\|q - p\| < t\rho/2$, and put $z' = (q - (1-t)w')/t$, so that $q = (1-t)w' + tz'$. Since $p = (1-t)w + tz$,

$$z' - z = \frac{(q - p) + (1-t)(w - w')}{t}, \quad \|z' - z\| < \frac{t\rho/2 + t\rho/2}{t} = \rho.$$

Hence $z' \in C$, and $q \in C$ by Definition 18.1.1. So every point within $t\rho/2$ of p lies in C , and $p \in \text{int}(C)$. ■

Now the theorem. It asks only that x_0 not be an interior point, which covers the boundary points of C and also the points outside $\overline{C}$.

Theorem 18.3.6. Supporting Hyperplane Theorem

Let $V \neq \{0\}$, let $C \subseteq V$ be convex, and let $x_0 \in V$ be a point that is **not** an interior point

of C . Then there is a **nonzero** $a \in V$ with

$$\langle a, y \rangle \leq \langle a, x_0 \rangle \quad \text{for every } y \in C.$$

If moreover $x_0 \in C$, the affine hyperplane $\{y : \langle a, y \rangle = \langle a, x_0 \rangle\}$ passes through x_0 and has C in one of its closed half-spaces; it is called a **supporting hyperplane** of C at x_0 .

Idea. Approach x_0 from outside $\overline{C}$ and cut off each approximating point with [Theorem 18.3.2](#). The normal vectors, scaled to length 1, live on the compact unit sphere, so a subsequence converges, and in the limit the strict inequalities become the nonstrict one we want. The only thing that could go wrong is that there are no points outside $\overline{C}$ near x_0 . [Lemma 18.3.4](#) and [Lemma 18.3.5](#) rule that out: if C is flat, its own hyperplane serves; if C is solid, pushing x_0 directly away from an interior point leaves $\overline{C}$ at once.

Proof. If $C = \emptyset$, any nonzero a works, and one exists because $V \neq \{0\}$. So let $C \neq \emptyset$.

Case 1: C has no interior point. By [Lemma 18.3.4](#) there are $a \neq 0$ and c with $\langle a, y \rangle = c$ for all $y \in C$. If $\langle a, x_0 \rangle \geq c$, then a works. Otherwise $-a$ works, since $\langle -a, y \rangle = -c < -\langle a, x_0 \rangle$.

Case 2: C has an interior point z . For $k \geq 1$ let $x_k = x_0 + \frac{1}{k}(x_0 - z)$. We claim $x_k \notin \overline{C}$. Indeed, with $t = \frac{1}{k+1}$,

$$(1-t)x_k + tz = \frac{k}{k+1}x_0 + \frac{1}{k+1}(x_0 - z) + \frac{1}{k+1}z = x_0,$$

so if x_k were in $\overline{C}$, then x_0 would be an interior point of C by [Lemma 18.3.5](#) (b), contrary to the hypothesis. By [Lemma 18.3.5](#) (a), $\overline{C}$ is nonempty, closed, and convex, so [Theorem 18.3.2](#), applied to $\overline{C}$ and x_k , gives $b_k \neq 0$ with $\langle b_k, y \rangle \leq \langle b_k, x_k \rangle$ for all $y \in \overline{C}$, in particular for all $y \in C$. Put $a_k = b_k / \|b_k\|$, a unit vector satisfying the same inequality.

The unit sphere of V is compact by Corollary 16.2.5, which rests on **fact (A3) of Chapter 16's introduction**. So a subsequence (a_{k_j}) converges to some a with $\|a\| = 1$; in particular $a \neq 0$. Fix $y \in C$. Then $\langle a_{k_j}, y - x_{k_j} \rangle \leq 0$ for every j , and $x_k \rightarrow x_0$ because $\|x_k - x_0\| = \|x_0 - z\|/k$. By Theorem 11.1.5,

$$\begin{aligned} |\langle a_{k_j}, y - x_{k_j} \rangle - \langle a, y - x_0 \rangle| &\leq |\langle a_{k_j} - a, y - x_{k_j} \rangle| + |\langle a, x_0 - x_{k_j} \rangle| \\ &\leq \|a_{k_j} - a\|(\|y - x_0\| + \|x_0 - z\|) + \|x_0 - x_{k_j}\|, \end{aligned}$$

which tends to 0. A nonstrict inequality survives the limit, so $\langle a, y - x_0 \rangle \leq 0$. This proves the theorem. ■

A supporting hyperplane need not be unique. At the corner $(1, 1)$ of the square $[-1, 1]^2$, every $\mathbf{a} = (a_1, a_2)$ with $a_1, a_2 \geq 0$, not both zero, works, since $a_1 y_1 + a_2 y_2 \leq a_1 + a_2$ on the square. At a point of the unit circle, the boundary of the disc, only the positive multiples of the point itself work. Corners admit many supporting hyperplanes and smooth points one; §08 builds its induction on dimension on these hyperplanes, and §04 uses them to prove that a norm is recovered from its dual.

18.3.5 Separating two convex sets

To separate two sets, separate one point from one set: the point 0 from the set of differences.

Theorem 18.3.7. Separating Hyperplane Theorem

Let $C_1, C_2 \subseteq V$ be **nonempty, disjoint**, and convex.

- (a) If C_1 is **compact** and C_2 is **closed**, then C_1 and C_2 are strictly separated: there are $a \neq 0$ and c with $\langle a, x \rangle < c < \langle a, y \rangle$ for all $x \in C_1, y \in C_2$, and indeed with a margin $\varepsilon > 0$ as in [Definition 18.3.1](#).
- (b) In general, C_1 and C_2 are separated: there are $a \neq 0$ and c with $\langle a, x \rangle \leq c \leq \langle a, y \rangle$ for all $x \in C_1, y \in C_2$.

Idea. The difference set $D = \{y - x : x \in C_1, y \in C_2\}$ is convex and misses 0, and a normal a with D on its positive side separates C_1 from C_2 . For (a), compactness of C_1 makes D closed, and [Theorem 18.3.2](#) cuts 0 off; for (b), 0 is not an interior point of D , and [Theorem 18.3.6](#) applies.

Proof. Let $D = \{y - x : x \in C_1, y \in C_2\}$. It is convex: for $0 \leq t \leq 1$,

$$t(y - x) + (1 - t)(y' - x') = (ty + (1 - t)y') - (tx + (1 - t)x'),$$

a difference of a point of C_2 and a point of C_1 . And $0 \notin D$, since $C_1 \cap C_2 = \emptyset$.

- (a) We show D is closed. Let $y_k - x_k \rightarrow d$ with $x_k \in C_1, y_k \in C_2$. Since C_1 is compact, a subsequence x_{k_j} converges to some $x \in C_1$. Then $y_{k_j} = (y_{k_j} - x_{k_j}) + x_{k_j} \rightarrow d + x$, since a subsequence of a convergent sequence has the same limit, part of the elementary algebra of limits that Chapter 16's introduction imports; so $d + x \in C_2$ because C_2 is closed, and $d = (d + x) - x \in D$. Now [Theorem 18.3.2](#), applied to D and the point $0 \notin D$, gives $b \neq 0$ with $\langle b, d \rangle \leq -\|b\|^2$ for all $d \in D$. With $a = -b$ this says

$$\langle a, x \rangle + \|a\|^2 \leq \langle a, y \rangle \quad \text{for all } x \in C_1, y \in C_2.$$

Fix $y \in C_2$. The numbers $\langle a, x \rangle, x \in C_1$, are bounded above by $\langle a, y \rangle - \|a\|^2$, so by **the completeness of $\mathbb{R}$, fact (A1) of Chapter 16's introduction**, they have a least upper bound s , and $s + \|a\|^2 \leq \langle a, y \rangle$ for every $y \in C_2$. Take $\varepsilon = \frac{1}{2}\|a\|^2$ and $c = s + \varepsilon$. Then $\langle a, x \rangle \leq s = c - \varepsilon$ and $\langle a, y \rangle \geq c + \varepsilon$.

- (b) If $V = \{0\}$, two nonempty subsets both contain 0 and cannot be disjoint, so $V \neq \{0\}$. The point 0 is not in D , so it is not an interior point of D . By [Theorem 18.3.6](#) there is $b \neq 0$ with $\langle b, d \rangle \leq \langle b, 0 \rangle = 0$ for all $d \in D$. With $a = -b$, this says $\langle a, x \rangle \leq \langle a, y \rangle$ for all $x \in C_1, y \in C_2$. As in (a), the least upper bound c of the numbers $\langle a, x \rangle, x \in C_1$, exists by (A1) and satisfies $\langle a, x \rangle \leq c \leq \langle a, y \rangle$. This proves the theorem. ■

18.3.6 Where strict separation fails

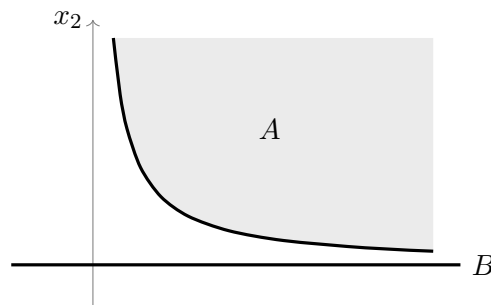
Part (a) asked one of the two sets to be compact. That hypothesis cannot be weakened to “closed”, and the standard witness is a hyperbola and its asymptote.

Example. A hyperbola and its asymptote

In $\mathbb{R}^2$ with the dot product, let

$$A = \{\mathbf{x} : x_1 > 0, x_2 > 0, x_1x_2 \geq 1\}, \quad B = \{\mathbf{x} : x_2 = 0\}.$$

Show that A and B are nonempty, disjoint, closed, and convex, that the x_1 -axis is the **only** hyperplane separating them, and that they are **not** strictly separated.



A and B are closed, convex and disjoint,
but the gap between them shrinks to zero

Solution. Nonempty and disjoint. $(1, 1) \in A$ and $\mathbf{0} \in B$; every point of A has $x_2 > 0$, and every point of B has $x_2 = 0$.

B is closed and convex. It is a subspace, so it is convex, and closed by Proposition 18.1.6 (c).

A is convex. Let $\mathbf{x}, \mathbf{z} \in A$ and $0 \leq t \leq 1$, and let $\mathbf{w} = t\mathbf{x} + (1-t)\mathbf{z}$. Its coordinates are positive. Expanding,

$$w_1w_2 = t^2x_1x_2 + (1-t)^2z_1z_2 + t(1-t)(x_1z_2 + x_2z_1).$$

By $(\sqrt{x_1z_2} - \sqrt{x_2z_1})^2 \geq 0$, we get $x_1z_2 + x_2z_1 \geq 2\sqrt{x_1x_2z_1z_2} \geq 2$, using $x_1x_2 \geq 1$ and $z_1z_2 \geq 1$. Hence $w_1w_2 \geq t^2 + (1-t)^2 + 2t(1-t) = 1$, and $\mathbf{w} \in A$.

A is closed. If $\mathbf{x}_k \rightarrow \mathbf{x}$ with $\mathbf{x}_k \in A$, then the coordinates converge, so $x_1, x_2 \geq 0$, and $x_{k,1}x_{k,2} \rightarrow x_1x_2$ by the algebra of limits, so $x_1x_2 \geq 1$. A product that is at least 1 has no zero factor, so $x_1, x_2 > 0$, and $\mathbf{x} \in A$.

The only separating hyperplane. Let $\mathbf{a} \neq \mathbf{0}$ and c separate A and B , in either order. On B , $\langle \mathbf{a}, (t, 0) \rangle = a_1t$ for all real t , and these values lie on one side of c ; that forces $a_1 = 0$. So $a_2 \neq 0$, and after replacing $(\mathbf{a}, c)$ by $(-\mathbf{a}, -c)$ we may take $a_2 > 0$. On B the value is 0. On A the values are a_2x_2 , and since $(1/s, s) \in A$ for every $s > 0$, they fill the interval $(0, \infty)$. These are unbounded above, so A is on the side $\geq c$ and B on the side $\leq c$: $0 \leq c \leq a_2s$ for every $s > 0$, which forces $c = 0$. The hyperplane $\{a_2x_2 = 0\}$ is the x_1 -axis, and it separates, since $a_2x_2 > 0$ on A and $= 0$ on B .

No strict separation. Strict separation would need the only separating hyperplane to miss

both sets, and B is that hyperplane.

The points $(t, 1/t) \in A$ and $(t, 0) \in B$ are $1/t$ apart, and [Exercise \(C2: Strict separation and distance\)](#) shows that this vanishing distance is exactly what rules out strict separation.

Warning. Disjoint closed convex sets need not be strictly separated. The hyperbola region and its asymptote above are the example to remember: both closed, both convex, disjoint, and every hyperplane that separates them contains one of them. “Disjoint” gives only the weak separation of [Theorem 18.3.7](#) (b). Strict separation needs more, such as compactness of one set in (a).

18.3.7 Extending linear functionals

Separation has an analytic twin. Instead of a set and a point, take a linear functional defined on a subspace and bounded there by a convex function; can it be extended to the whole space without breaking the bound? The bounds allowed are the following.

Definition 18.3.3. Sublinear Functional

Let V be a real vector space. A function $p: V \rightarrow \mathbb{R}$ is **sublinear** if, for **all** $x, y \in V$ and **all** real $t \geq 0$,

$$p(x + y) \leq p(x) + p(y) \quad \text{and} \quad p(tx) = tp(x).$$

Every norm is sublinear: (N3) is the first condition, and (N2) with $t \geq 0$ the second. Every linear functional is sublinear, with equality in the first condition. So is $p(\mathbf{x}) = \max(x_1, x_2)$ on $\mathbb{R}^2$, which is neither a norm nor linear: $p(-\mathbf{e}_1) = 0$. Taking $t = 0$ gives $p(0) = 0$. A nonexample by minimal change is $p(x) = \|x\|^2$: the first condition fails at $x = y \neq 0$ (it would give $4\|x\|^2 \leq 2\|x\|^2$), and so does the second, since $p(tx) = t^2p(x)$.

It turns out that sublinearity is exactly what makes one-step extensions possible.

Theorem 18.3.8. Hahn–Banach Theorem, Finite-Dimensional

Let V be a **finite-dimensional** real vector space, let p be a sublinear functional on V , let U be a subspace of V , and let $f \in U^*$ satisfy $f(u) \leq p(u)$ for every $u \in U$. Then there is $F \in V^*$ with

$$F(u) = f(u) \text{ for every } u \in U \quad \text{and} \quad F(v) \leq p(v) \text{ for every } v \in V.$$

Idea. Add one dimension at a time. On $U + \text{span}(v)$, a linear extension is fixed by one number $c = F(v)$. The requirement $F \leq p$ at the vectors $u + tv$ with $t > 0$ gives upper bounds on c , and at those with $t < 0$ it gives lower bounds. Sublinearity of p , together with linearity of f , is exactly what makes every lower bound at most every upper bound, so some c fits. Then induct on $\dim V - \dim U$. No inner product is needed.

Proof. We induct on $k = \dim V - \dim U$. If $k = 0$, then $U = V$ and $F = f$ works.

Let $k \geq 1$ and assume the theorem whenever the difference of dimensions is $k - 1$. Pick $v \in V \setminus U$ and let $W = U + \text{span}(v)$. Each $w \in W$ can be written as $u + tv$ with $u \in U$ and $t \in \mathbb{R}$, and only in one way: if $u + tv = u' + t'v$ with $t \neq t'$, then $v = (u' - u)/(t - t') \in U$, which is false; so $t = t'$ and $u = u'$. Appending v to a basis of U gives a basis of W by Lemma 1.5.3, so $\dim V - \dim W = k - 1$. For $c \in \mathbb{R}$, the formula $g(u + tv) = f(u) + tc$ therefore defines a linear functional on W that extends f .

For $u, u' \in U$, linearity of f , the hypothesis, and sublinearity of p give

$$f(u) + f(u') = f(u + u') \leq p(u + u') \leq p(u - v) + p(u' + v),$$

that is, $f(u) - p(u - v) \leq p(u' + v) - f(u')$. So every number in $L = \{f(u) - p(u - v) : u \in U\}$ is at most every number in $R = \{p(u' + v) - f(u') : u' \in U\}$. By **fact (A1) of Chapter 16's introduction**, the nonempty set L , bounded above by any element of R , has a least upper bound c , and $c \leq r$ for every $r \in R$. With this c , we check $g(u + tv) \leq p(u + tv)$ in three cases.

- $t = 0$: this is the hypothesis $f(u) \leq p(u)$.
- $t > 0$: $c \leq p(u/t + v) - f(u/t)$, the element of R at $u' = u/t$. Multiply by t and use $p(t \cdot) = t p(\cdot)$: $tc \leq p(u + tv) - f(u)$.
- $t = -s < 0$: $c \geq f(u/s) - p(u/s - v)$, since this is an element of L . Multiply by s : $sc \geq f(u) - p(u - sv)$, that is, $f(u) - sc \leq p(u + tv)$.

So $g \leq p$ on W . By the induction hypothesis applied to W and g , there is $F \in V^*$ extending g , hence f , with $F \leq p$ on V . This proves the theorem. ■

Beyond finite dimension. The induction above stops after $\dim V - \dim U$ steps. In an infinite-dimensional space there is no last step, and the one-step extension has to be combined with Zorn's lemma (Theorem 1.8.1), exactly as Chapter 1 §08 combined “append one vector” with Zorn's lemma to prove Theorem 1.8.3. The resulting theorem, the Hahn–Banach theorem of functional analysis, is outside the scope of this book, and nothing here depends on it.

The special case that matters most is a norm.

Corollary 18.3.9. Norm-Preserving Extension

Let V be a finite-dimensional real vector space with a norm $\|\cdot\|$, let U be a subspace, and let $f \in U^*$ satisfy $|f(u)| \leq \|u\|$ for every $u \in U$. Then there is $F \in V^*$ extending f with $|F(v)| \leq \|v\|$ for every $v \in V$.

Proof. A norm is sublinear, so [Theorem 18.3.8](#) with $p = \|\cdot\|$ gives F extending f with $F(v) \leq \|v\|$ for all v . Applying this to $-v$ gives $-F(v) = F(-v) \leq \|-v\| = \|v\|$ by (N2). Hence $|F(v)| \leq \|v\|$. ■

In particular, for any $x \neq 0$, taking $U = \text{span}(x)$ and $f(tx) = t\|x\|$ produces a functional F

with $F(x) = \|x\|$ and $|F| \leq \|\cdot\|$ everywhere: a linear functional that “sees” the full length of x without ever exceeding the norm. §04 meets this functional again, as the vector at which a dual norm is attained.

Example. Many extensions

On $\mathbb{R}^2$ with $\|\cdot\|_\infty$, let $U = \text{span}((1, 1))$ and $f(s(1, 1)) = s$. Find every F as in [Corollary 18.3.9](#), and compare with the interval produced by the proof of [Theorem 18.3.8](#) for $\mathbf{v} = \mathbf{e}_1$.

Solution. First, $|f(s(1, 1))| = |s| = \|s(1, 1)\|_\infty$, so the hypothesis holds. An extension has $F(\mathbf{e}_1) = c$ for some c , and then $F(\mathbf{e}_2) = F((1, 1)) - c = 1 - c$, so $F(\mathbf{x}) = cx_1 + (1 - c)x_2$.

If $0 \leq c \leq 1$, then $|F(\mathbf{x})| \leq c|x_1| + (1 - c)|x_2| \leq \|\mathbf{x}\|_\infty$. If $c > 1$, then $\mathbf{x} = (1, -1)$ gives $F(\mathbf{x}) = 2c - 1 > 1 = \|\mathbf{x}\|_\infty$. If $c < 0$, then $\mathbf{x} = (-1, 1)$ gives $F(\mathbf{x}) = 1 - 2c > 1$. So the extensions are exactly $F(\mathbf{x}) = cx_1 + (1 - c)x_2$ with $c \in [0, 1]$: infinitely many.

In the proof, with $p = \|\cdot\|_\infty$, $\mathbf{u} = s(1, 1)$ and $\mathbf{v} = \mathbf{e}_1$, the elements of L are $s - \max(|s - 1|, |s|)$, which is 0 for $s \geq \frac{1}{2}$ and $2s - 1 < 0$ for $s < \frac{1}{2}$; so $\sup L = 0$. The elements of R are $\max(|s + 1|, |s|) - s$, which is 1 for $s \geq -\frac{1}{2}$ and $-2s > 1$ for $s < -\frac{1}{2}$; so $\inf R = 1$. Every c between $\sup L$ and $\inf R$ works, and that is the interval $[0, 1]$ found directly. The proof picked its left end.

The corners of the square B_∞ are why the extension is not unique: U meets the unit sphere at the corner $(1, 1)$, and every supporting line of the square at that corner gives an extension.

18.3.8 Exercises

18.3.8.1 A. Check your understanding

Exercise. A1

- State the characterization [Eq. 18.3.1](#) of the nearest point, and say what it becomes when C is a subspace.
- Which two facts of Chapter 16’s introduction does the existence of $P_C(x)$ use, and to which set are they applied?
- Determine whether the following statement is correct, and justify your answer: every nonempty closed subset of $\mathbb{R}^2$ has a unique point nearest to $\mathbf{0}$.
- Determine whether the following statement is correct, and justify your answer: two disjoint nonempty closed convex subsets of $\mathbb{R}^2$ can always be strictly separated.

Solution.

- A point $p \in C$ is $P_C(x)$ exactly when $\langle x - p, y - p \rangle \leq 0$ for every $y \in C$ ([Theorem 18.3.1](#) (b)). For a subspace U it becomes the equality $\langle x - p, u \rangle = 0$ for every $u \in U$, so $P_C(x) = P_U x$ and [Theorem 11.3.5](#) is recovered.
- The compactness of closed bounded sets (A3), through [Corollary 16.2.5](#), and the extreme value theorem (A4). They are applied to $K = \{y \in C : \|x - y\| \leq R\}$, the part of C within the distance R of x attained at one chosen point of C , and to the function

$$y \mapsto \|x - y\|.$$

- (c) Incorrect. The unit circle is closed and nonempty, and every one of its points is at distance 1 from $\mathbf{0}$. Uniqueness needs convexity.
- (d) Incorrect. The region A above the hyperbola $x_1x_2 = 1$ in the positive quadrant and its asymptote B , the x_1 -axis, are disjoint, closed, and convex, but the only hyperplane separating them is B itself (Example (A hyperbola and its asymptote)).

18.3.8.2 B. Practice

Exercise. B1: Projection onto a half-plane

In $\mathbb{R}^2$ with the dot product, let $C = \{\mathbf{y} : y_1 + y_2 \leq 1\}$ and $\mathbf{x} = (2, 3)$. Find $P_C(\mathbf{x})$ and $d(\mathbf{x}, C)$, and verify Eq. 18.3.1.

Solution. Here $\mathbf{a} = (1, 1)$, $c = 1$, and $\langle \mathbf{a}, \mathbf{x} \rangle = 5 > 1$. By Example (Three projections) (c), $\lambda = (5 - 1)/2 = 2$ and $P_C(\mathbf{x}) = (2, 3) - 2(1, 1) = (0, 1)$. Check: $0 + 1 = 1$, so $\mathbf{p} = (0, 1) \in C$, and $\mathbf{x} - \mathbf{p} = (2, 2)$, so $\langle \mathbf{x} - \mathbf{p}, \mathbf{y} - \mathbf{p} \rangle = 2y_1 + 2(y_2 - 1) = 2(y_1 + y_2 - 1) \leq 0$ for $\mathbf{y} \in C$. By Theorem 18.3.1 (b) this is the nearest point, and $d(\mathbf{x}, C) = \|(2, 2)\| = 2\sqrt{2}$.

Exercise. B2: Cutting off a point from a diamond

Let $C = B_1 = \{\mathbf{y} \in \mathbb{R}^2 : |y_1| + |y_2| \leq 1\}$, with the dot product, and $\mathbf{x} = (2, 2)$.

- (a) Show that $P_C(\mathbf{x}) = (\frac{1}{2}, \frac{1}{2})$.
- (b) Hence find the strictly separating hyperplane of Theorem 18.3.2, and check it directly.
- (c) Find every supporting hyperplane of C at $(\frac{1}{2}, \frac{1}{2})$.

Solution.

- (a) Let $\mathbf{p} = (\frac{1}{2}, \frac{1}{2})$, which lies in C . Then $\mathbf{x} - \mathbf{p} = (\frac{3}{2}, \frac{3}{2})$ and, for $\mathbf{y} \in C$,

$$\langle \mathbf{x} - \mathbf{p}, \mathbf{y} - \mathbf{p} \rangle = \frac{3}{2}(y_1 + y_2 - 1) \leq 0,$$

because $y_1 + y_2 \leq |y_1| + |y_2| \leq 1$. By Theorem 18.3.1 (b), $P_C(\mathbf{x}) = \mathbf{p}$.

- (b) With $\mathbf{a} = (\frac{3}{2}, \frac{3}{2})$, we have $\|\mathbf{a}\|^2 = \frac{9}{2}$ and $\langle \mathbf{a}, \mathbf{x} \rangle = 6$, so $c = 6 - \frac{9}{4} = \frac{15}{4}$. The hyperplane $\frac{3}{2}(y_1 + y_2) = \frac{15}{4}$ is the line $y_1 + y_2 = \frac{5}{2}$. Directly: on C , $y_1 + y_2 \leq 1 < \frac{5}{2}$, while $x_1 + x_2 = 4 > \frac{5}{2}$.
- (c) Let $\mathbf{a} \neq \mathbf{0}$ satisfy $\langle \mathbf{a}, \mathbf{y} \rangle \leq \langle \mathbf{a}, \mathbf{p} \rangle = \frac{1}{2}(a_1 + a_2)$ for all $\mathbf{y} \in C$. Taking $\mathbf{y} = \mathbf{e}_1$ and $\mathbf{y} = \mathbf{e}_2$ gives $a_1 \leq \frac{1}{2}(a_1 + a_2)$ and $a_2 \leq \frac{1}{2}(a_1 + a_2)$, that is, $a_1 \leq a_2$ and $a_2 \leq a_1$; so $a_1 = a_2$. Taking $\mathbf{y} = -\mathbf{e}_1$ gives $-a_1 \leq a_1$, so $a_1 \geq 0$, and $a_1 \neq 0$ since $\mathbf{a} \neq \mathbf{0}$. Conversely $\mathbf{a} = (a, a)$ with $a > 0$ works, by the inequality in (a). So the only supporting hyperplane at $\mathbf{p}$ is the line $y_1 + y_2 = 1$ through the edge containing $\mathbf{p}$.

Exercise. B3: A unique extension

On $\mathbb{R}^2$ with $\|\cdot\|_1$, let $U = \text{span}((1, 2))$ and $f(t(1, 2)) = 3t$. Show that f satisfies the hypothesis of [Corollary 18.3.9](#), and find every extension F that the corollary allows.

Solution. $|f(t(1, 2))| = 3|t| = |t| + 2|t| = \|t(1, 2)\|_1$, so the hypothesis holds. Every $F \in (\mathbb{R}^2)^*$ is $F(\mathbf{x}) = c_1x_1 + c_2x_2$ (Theorem 5.1.1), and it extends f exactly when $F((1, 2)) = 3$, that is, $c_1 + 2c_2 = 3$.

Next, $|F(\mathbf{x})| \leq \|\mathbf{x}\|_1$ for all $\mathbf{x}$ holds exactly when $|c_1| \leq 1$ and $|c_2| \leq 1$. Indeed, if both hold then $|c_1x_1 + c_2x_2| \leq |x_1| + |x_2|$; and if, say, $|c_1| > 1$, then $\mathbf{x} = \mathbf{e}_1$ gives $|F(\mathbf{e}_1)| > 1 = \|\mathbf{e}_1\|_1$.

So we need $c_1 = 3 - 2c_2$ with $|3 - 2c_2| \leq 1$ and $|c_2| \leq 1$. The first gives $1 \leq c_2 \leq 2$, the second $c_2 \leq 1$; hence $c_2 = 1$ and $c_1 = 1$. The only extension is $F(\mathbf{x}) = x_1 + x_2$.

18.3.8.3 C. Going deeper

Exercise. C1: The projection does not stretch

Let $C \subseteq V$ be nonempty, closed, and convex. Prove that $\|P_C(x) - P_C(y)\| \leq \|x - y\|$ for all $x, y \in V$.

Hint: write [Eq. 18.3.1](#) once at x , tested with the point $P_C(y)$, and once at y , tested with $P_C(x)$, and add.

Solution. Put $p = P_C(x)$ and $q = P_C(y)$. By [Theorem 18.3.1](#) (b), applied at x with the point $q \in C$ and at y with the point $p \in C$,

$$\langle x - p, q - p \rangle \leq 0, \quad \langle y - q, p - q \rangle \leq 0.$$

The second is $\langle q - y, q - p \rangle \leq 0$. Adding, $\langle (x - y) + (q - p), q - p \rangle \leq 0$, that is,

$$\|q - p\|^2 \leq \langle y - x, q - p \rangle \leq \|y - x\| \|q - p\|$$

by Theorem 11.1.5. If $q = p$ there is nothing to prove; otherwise divide by $\|q - p\| > 0$. This proves $\|p - q\| \leq \|x - y\|$.

Exercise. C2: Strict separation and distance

Let $C_1, C_2 \subseteq V$ be nonempty, closed, and convex, and put $\delta = \inf\{\|y - x\| : x \in C_1, y \in C_2\}$.

- Prove that if C_1 and C_2 are strictly separated, then $\delta > 0$.
- Prove that if $\delta > 0$, then C_1 and C_2 are strictly separated.
- Hence deduce, without the computation of [Example \(A hyperbola and its asymptote\)](#), that the sets A and B there are not strictly separated.

Hint for (b): replace the difference set D of [Theorem 18.3.7](#) by its closure.

Solution.

- (a) Let $a \neq 0$, c , and $\varepsilon > 0$ strictly separate the sets, with C_1 on the side $\leq c - \varepsilon$. For $x \in C_1$ and $y \in C_2$, Theorem 11.1.5 gives

$$2\varepsilon \leq \langle a, y \rangle - \langle a, x \rangle = \langle a, y - x \rangle \leq \|a\| \|y - x\|,$$

so $\|y - x\| \geq 2\varepsilon/\|a\|$. Hence $\delta \geq 2\varepsilon/\|a\| > 0$.

- (b) Let $D = \{y - x : x \in C_1, y \in C_2\}$, which is convex by the computation in the proof of [Theorem 18.3.7](#). Every $d \in D$ has $\|d\| \geq \delta$, and so does every point of $\overline{D}$: if $d_k \rightarrow d$ with $d_k \in D$, then $\|d\| \geq \|d_k\| - \|d_k - d\| \geq \delta - \|d_k - d\|$ for every k , and the right side tends to δ . So $0 \notin \overline{D}$, and $\overline{D}$ is nonempty, closed, and convex by [Lemma 18.3.5 \(a\)](#). By [Theorem 18.3.2](#) there is $b \neq 0$ with $\langle b, d \rangle \leq -\|b\|^2$ on $\overline{D} \supseteq D$. With $a = -b$, this is $\langle a, x \rangle + \|a\|^2 \leq \langle a, y \rangle$ for all $x \in C_1, y \in C_2$, and the last paragraph of the proof of [Theorem 18.3.7 \(a\)](#) turns it into strict separation with $\varepsilon = \frac{1}{2}\|a\|^2$.
- (c) The points $(t, 1/t) \in A$ and $(t, 0) \in B$ are at distance $1/t$, so $\delta \leq 1/t$ for every $t > 0$, and $\delta = 0$. By (a), A and B are not strictly separated.