Convex Geometry

Convex Sets

A set $C \subset \mathbf{E}$ is said to be convex if

$$x, y \in C \text{ and } \lambda \in [0,1] \implies (1-\lambda)x + \lambda y \in C.$$

That is, C contains all line segments connecting points in C.

Convex Sets

A set $C \subset \mathbf{E}$ is said to be convex if

$$x, y \in C \text{ and } \lambda \in [0,1] \implies (1-\lambda)x + \lambda y \in C.$$

That is, C contains all line segments connecting points in C.

Examples:

– Subspaces and affine sets

– Half spaces $\{x \,|\, \langle a, x\rangle \leq \beta\}$ for all $a \in \mathbf{E} \setminus \{0\}$ and $\beta \in \mathbf{R}$.

– The unit ball $\mathbb{B} := \{x \,|\, \|x\| \leq 1\}$ and $\operatorname{intr}(\mathbb{B})$.

– The *unit simplex*

$$\Delta_n := \{\lambda \in \mathbf{R}^n : \textstyle\sum_{i=1}^n \lambda_i = 1, \lambda \geq 0\}.$$

Convexity Preserving Operations

Let $\mathcal{A} \in \mathbf{L}(\mathbf{E}, \mathbf{Y})$. If $C_1, C_2 \subset \mathbf{E}$ and $K \subset \mathbf{Y}$ are all convex, then so are the sets

– **Intersection:** $C_1 \cap C_2$

– **Scalar Multiplication:** $\mathbf{R}_+ K$ and $\lambda K \quad \forall\, \lambda \in \mathbf{R}$

– **Addition:** $C_1 + C_2$

– **Linear Image/Preimage:** $\mathcal{A}C_1$ and $\mathcal{A}^{-1}K$

– **Products:** $C_1 \times K$

– **Closure and Interior:** $\mathrm{cl}\, K$ and $\mathrm{intr}\, K$

– **Non-negative sums:** Let $Q \subset \mathbf{E}$ be convex and $\lambda_1, \lambda_2 \in \mathbf{R}_+$. Then

$$\lambda_1 Q + \lambda_2 Q = (\lambda_1 + \lambda_2) Q.$$

Polyhedra and Spectrahedra

A convex *polyhedron* is any set of the form

$$Q = \{x \in \mathbf{R}^n : Ax \geq c\},$$

for some $A \in \mathbf{R}^{m \times n}$ and $c \in \mathbf{R}^m$.
Equivalently, we may write Q as an intersection of finitely many half-spaces or as the preimage $A^{-1}(c + \mathbf{R}^n_+)$. Hence, a convex polyhedron is convex.

Polyhedra and Spectrahedra

A convex *polyhedron* is any set of the form

$$Q = \{x \in \mathbf{R}^n : Ax \geq c\},$$

for some $A \in \mathbf{R}^{m \times n}$ and $c \in \mathbf{R}^m$.
Equivalently, we may write Q as an intersection of finitely many half-spaces or as the preimage $A^{-1}(c + \mathbf{R}^n_+)$. Hence, a convex polyhedron is convex.

More generally, a *spectrahedron* is any set of the form

$$Q = \{x \in \mathbf{R}^n : x_1 A_1 + x_2 A_2 + \ldots + x_n A_n \succeq C\},$$

for some matrices $A_i \in \mathbf{S}^m$ and $C \in \mathbf{S}^n$. Equivalently, we may write Q as the preimage $\mathcal{A}^{-1}(C + \mathbf{S}^n_+)$ for the linear map $\mathcal{A}(x) = \sum_{i=1}^n x_i A_i$.

Spectrahedra

There are many more spectrahedra than polyhedra. For example, the elliptope is given by

$$\left\{ (x, y, z) \in \mathbf{R}^3 : \begin{pmatrix} 1 & x & y \\ x & 1 & z \\ y & z & 1 \end{pmatrix} \succeq 0 \right\}$$

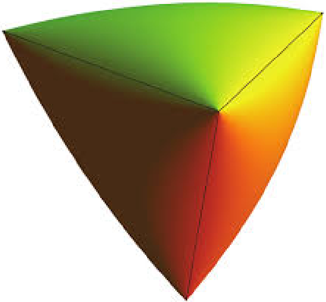


Figure: The elliptope

Convex Hulls and Convex Combinations

Convex Combinations: A point $x \in \mathbf{E}$ is a *convex combination* of points $x_1, \ldots, x_k \in \mathbf{E}$ if it can be written as $x = \sum_{i=1}^{k} \lambda_i x_i$ for some $\lambda \in \Delta_k$.

A convex combination $x = \sum_{i=1}^{k} \lambda_i x_i$ can be viewed as a weighted average of the points $x_1, \ldots, x_k$ with $\lambda_1, \ldots, \lambda_k$ as the corresponding weights.

Given a set $X \subset \mathbf{E}$, one can show that the set of all such convex combinations of points in X,

$$\left\{ \sum_{i=1}^{k} \lambda_i x_i \,|\, k \in \mathbb{N},\ \lambda \in \Delta_k, x_1, \ldots, x_k \in X \right\},$$

equals the convex hull of the set X, $\mathrm{conv}(X)$, i.e. the intersection of all convex sets containing X. Here $\mathbb{N} := \{1, 2, \ldots\}$ is the set of *natural numbers*.

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \text{conv}(Q)$ can be written as a convex combination of $n + 1$ or fewer points in Q.

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \text{conv}(Q)$ can be written as a convex combination of $n+1$ or fewer points in Q.

Proof: Let $x \in \text{conv}(Q)$.
1) WLOG

$$\bar{k} := \inf \left\{ k \in \mathbb{N} \,\middle|\, x = \sum_{i=1}^{k} \lambda_i x_i,\ x_1, \ldots, x_k \in Q,\ \lambda \in \Delta_k \right\} > n+1.$$

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \text{conv}(Q)$ can be written as a convex combination of $n+1$ or fewer points in Q.

Proof: Let $x \in \text{conv}(Q)$.
1) WLOG
$\overline{k} := \inf \left\{ k \in \mathbb{N} \,\middle|\, x = \sum_{i=1}^{k} \lambda_i x_i,\ x_1, \ldots, x_k \in Q,\ \lambda \in \Delta_k \right\} > n+1.$

2) $\exists\, x_1, \ldots, x_{\overline{k}} \in Q,\ \lambda \in \Delta_{\overline{k}}$ s.t. $x = \sum_{i=1}^{\overline{k}} \lambda_i x_i$ and $\lambda > 0$.

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \mathrm{conv}(Q)$ can be written as a convex combination of $n+1$ or fewer points in Q.

Proof: Let $x \in \mathrm{conv}(Q)$.
1) WLOG

$$\overline{k} := \inf\left\{k \in \mathbb{N} \,\middle|\, x = \sum_{i=1}^{k} \lambda_i x_i,\ x_1, \ldots, x_k \in Q,\ \lambda \in \Delta_k\right\} > n+1.$$

2) $\exists\, x_1, \ldots, x_{\overline{k}} \in Q,\ \lambda \in \Delta_{\overline{k}}$ s.t. $x = \sum_{i=1}^{\overline{k}} \lambda_i x_i$ and $\lambda > 0$.

3) Since $\overline{k} > n+1$, $\exists\, \mu_i,\ i = 2, \ldots, \overline{k}$ not all 0 s.t.

$$\{\mu_2, \ldots, \mu_{\overline{k}}\} \,:\, 0 = \sum_{i=2}^{\overline{k}} \mu_i (x_i - x_1) = \left(\sum_{i=2}^{\overline{k}} \mu_i x_i\right) - \left(\sum_{i=2}^{\overline{k}} \mu_i\right) x_1.$$

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \operatorname{conv}(Q)$ can be written as a convex combination of $n+1$ or fewer points in Q.

Proof: Let $x \in \operatorname{conv}(Q)$.
1) WLOG
$$\overline{k} := \inf\left\{k \in \mathbb{N} \,\middle|\, x = \sum_{i=1}^{k} \lambda_i x_i,\ x_1, \dots, x_k \in Q,\ \lambda \in \Delta_k \right\} > n+1.$$

2) $\exists\, x_1, \dots, x_{\overline{k}} \in Q,\ \lambda \in \Delta_{\overline{k}}$ s.t. $x = \sum_{i=1}^{\overline{k}} \lambda_i x_i$ and $\lambda > 0$.

3) Since $\overline{k} > n+1$, $\exists\, \mu_i,\ i = 2, \dots, \overline{k}$ not all 0 s.t.
$$\{\mu_2, \dots, \mu_{\overline{k}}\} \,:\, 0 = \sum_{i=2}^{\overline{k}} \mu_i (x_i - x_1) = \left(\sum_{i=2}^{\overline{k}} \mu_i x_i\right) - \left(\sum_{i=2}^{\overline{k}} \mu_i\right) x_1.$$

4) $\mu_1 := -\sum_{i=2}^{\overline{k}} \mu_i \implies \sum_{i=1}^{\overline{k}} \mu_i x_i = 0, \quad \sum_{i=1}^{\overline{k}} \mu_i = 0,\ \mu \notin \mathbf{R}_{-}^{\overline{k}}$.

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \text{conv}(Q)$ can be written as a convex combination of $n+1$ or fewer points in Q.

Proof: Let $x \in \text{conv}(Q)$.
1) WLOG
$$\overline{k} := \inf \left\{ k \in \mathbb{N} \,\middle|\, x = \sum_{i=1}^{k} \lambda_i x_i,\ x_1, \ldots, x_k \in Q,\ \lambda \in \Delta_k \right\} > n+1.$$

2) $\exists\, x_1, \ldots, x_{\overline{k}} \in Q,\ \lambda \in \Delta_{\overline{k}}$ s.t. $x = \sum_{i=1}^{\overline{k}} \lambda_i x_i$ and $\lambda > 0$.

3) Since $\overline{k} > n+1$, $\exists\, \mu_i,\ i = 2, \ldots, \overline{k}$ not all 0 s.t.
$$\{\mu_2, \ldots, \mu_{\overline{k}}\} \,:\, 0 = \sum_{i=2}^{\overline{k}} \mu_i (x_i - x_1) = \left(\sum_{i=2}^{\overline{k}} \mu_i x_i\right) - \left(\sum_{i=2}^{\overline{k}} \mu_i\right) x_1.$$

4) $\mu_1 := -\sum_{i=2}^{\overline{k}} \mu_i \implies \sum_{i=1}^{\overline{k}} \mu_i x_i = 0, \quad \sum_{i=1}^{\overline{k}} \mu_i = 0,\ \mu \notin \mathbf{R}_{-}^{\overline{k}}$.

5) $\forall\, \alpha \in \mathbf{R},\ x = \sum_{i=1}^{\overline{k}} (\lambda_i - \alpha \mu_i) x_i$ and $\sum_{i=1}^{\overline{k}} (\lambda_i - \alpha \mu_i) = 1$.

Carathéodory's Theorem

Let $Q \subset \mathbf{E}$, where $\mathbf{E}$ is an n-dimensional Euclidean space. Then each point $x \in \text{conv}(Q)$ can be written as a convex combination of $n+1$ or fewer points in Q.

Proof: Let $x \in \text{conv}(Q)$.
1) WLOG
$\bar{k} := \inf \left\{ k \in \mathbb{N} \,\middle|\, x = \sum_{i=1}^{k} \lambda_i x_i,\ x_1, \ldots, x_k \in Q,\ \lambda \in \Delta_k \right\} > n+1.$

2) $\exists\, x_1, \ldots, x_{\bar{k}} \in Q,\ \lambda \in \Delta_{\bar{k}}$ s.t. $x = \sum_{i=1}^{\bar{k}} \lambda_i x_i$ and $\lambda > 0$.

3) Since $\bar{k} > n+1$, $\exists\, \mu_i,\ i = 2, \ldots, \bar{k}$ not all 0 s.t.
$\{\mu_2, \ldots, \mu_{\bar{k}}\} \,:\, 0 = \sum_{i=2}^{\bar{k}} \mu_i (x_i - x_1) = (\sum_{i=2}^{\bar{k}} \mu_i x_i) - (\sum_{i=2}^{\bar{k}} \mu_i) x_1.$

4) $\mu_1 := -\sum_{i=2}^{\bar{k}} \mu_i \implies \sum_{i=1}^{\bar{k}} \mu_i x_i = 0, \quad \sum_{i=1}^{\bar{k}} \mu_i = 0,\ \mu \notin \mathbf{R}_-^{\bar{k}}.$

5) $\forall\, \alpha \in \mathbf{R},\ x = \sum_{i=1}^{\bar{k}} (\lambda_i - \alpha\mu_i) x_i$ and $\sum_{i=1}^{\bar{k}} (\lambda_i - \alpha\mu_i) = 1$.

6) $\bar{\alpha} := \inf \left\{ \alpha \,\middle|\, \lambda_i - \alpha\mu_i \geq 0,\ i = 1, \ldots \bar{k} \right\} > 0$ and WLOG
$\lambda_{\bar{k}} - \bar{\alpha}\mu_{\bar{k}} = 0$ so $x = \sum_{i=1}^{\bar{k}-1} \bar{\lambda}_i x_i,\ \bar{\lambda} \in \Delta_{\bar{k}-1}$, where
$\bar{\lambda}_i := \lambda_i - \alpha\mu_i,\ i = 1, \ldots, \bar{k} - 1$. Contradiction.

Relative interior and Boundary

The *relative interior* of a set $Q \subset \mathbf{E}$, denoted $\operatorname{ri} Q$, is the interior of Q relative to $\operatorname{aff}(Q)$. That is,

$$\operatorname{ri} Q := \{x \in Q : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(x) \cap \operatorname{aff} Q \subseteq Q\}.$$

The *relative boundary* of Q is defined by $\operatorname{rb} Q := (\operatorname{cl} Q) \setminus (\operatorname{ri} Q)$.

Relative interior and Boundary

The *relative interior* of a set $Q \subset \mathbf{E}$, denoted $\operatorname{ri} Q$, is the interior of Q relative to $\operatorname{aff}(Q)$. That is,

$$\operatorname{ri} Q := \{x \in Q : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(x) \cap \operatorname{aff} Q \subseteq Q\}.$$

The *relative boundary* of Q is defined by $\operatorname{rb} Q := (\operatorname{cl} Q) \setminus (\operatorname{ri} Q)$.

The subspace parallel to $\operatorname{aff} Q$ is denoted $\operatorname{par} Q$. Observe that since $\operatorname{par} Q = \operatorname{aff} Q - x$ for all $x \in \operatorname{aff} Q$, it easily follows that

$$B_\epsilon(x) \cap \operatorname{aff} Q = x + B_\epsilon(0) \cap \operatorname{par} Q$$

for all $x \in \operatorname{aff} Q$.

Relative interior and Boundary

The *relative interior* of a set $Q \subset \mathbf{E}$, denoted $\operatorname{ri} Q$, is the interior of Q relative to $\operatorname{aff}(Q)$. That is,

$$\operatorname{ri} Q := \{x \in Q : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(x) \cap \operatorname{aff} Q \subseteq Q\}.$$

The *relative boundary* of Q is defined by $\operatorname{rb} Q := (\operatorname{cl} Q) \setminus (\operatorname{ri} Q)$.

The subspace parallel to $\operatorname{aff} Q$ is denoted $\operatorname{par} Q$. Observe that since $\operatorname{par} Q = \operatorname{aff} Q - x$ for all $x \in \operatorname{aff} Q$, it easily follows that

$$B_\epsilon(x) \cap \operatorname{aff} Q = x + B_\epsilon(0) \cap \operatorname{par} Q$$

for all $x \in \operatorname{aff} Q$.

$$[Q = \{x\}] \Longrightarrow [\operatorname{aff} Q = \{x\}, \quad \operatorname{par} Q = \{0\}, \quad Q = \operatorname{ri} Q \quad \text{and} \quad \operatorname{rb} Q = \emptyset]$$

Relative interior and Boundary

The *relative interior* of a set $Q \subset \mathbf{E}$, denoted $\operatorname{ri} Q$, is the interior of Q relative to $\operatorname{aff}(Q)$. That is,

$$\operatorname{ri} Q := \{x \in Q : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(x) \cap \operatorname{aff} Q \subseteq Q\}.$$

The *relative boundary* of Q is defined by $\operatorname{rb} Q := (\operatorname{cl} Q) \setminus (\operatorname{ri} Q)$.

The subspace parallel to $\operatorname{aff} Q$ is denoted $\operatorname{par} Q$. Observe that since $\operatorname{par} Q = \operatorname{aff} Q - x$ for all $x \in \operatorname{aff} Q$, it easily follows that

$$B_\epsilon(x) \cap \operatorname{aff} Q = x + B_\epsilon(0) \cap \operatorname{par} Q$$

for all $x \in \operatorname{aff} Q$.

$$[Q = \{x\}] \implies [\operatorname{aff} Q = \{x\}, \quad \operatorname{par} Q = \{0\}, \quad Q = \operatorname{ri} Q \quad \text{and} \quad \operatorname{rb} Q = \emptyset]$$

$$[Q = \mathbf{E}] \implies [\operatorname{aff} Q = \mathbf{E}, \quad \operatorname{par} Q = \mathbf{E}, \quad Q = \operatorname{ri} Q \quad \text{and} \quad \operatorname{rb} Q = \emptyset]$$

Relative interior and Boundary

The *relative interior* of a set $Q \subset \mathbf{E}$, denoted $\operatorname{ri} Q$, is the interior of Q relative to aff (Q). That is,

$$\operatorname{ri} Q := \{x \in Q : \exists \epsilon > 0 \text{ s.t. } B_\epsilon(x) \cap \operatorname{aff} Q \subseteq Q\}.$$

The *relative boundary* of Q is defined by $\operatorname{rb} Q := (\operatorname{cl} Q) \setminus (\operatorname{ri} Q)$.

The subspace parallel to aff Q is denoted $\operatorname{par} Q$. Observe that since $\operatorname{par} Q = \operatorname{aff} Q - x$ for all $x \in \operatorname{aff} Q$, it easily follows that

$$B_\epsilon(x) \cap \operatorname{aff} Q = x + B_\epsilon(0) \cap \operatorname{par} Q$$

for all $x \in \operatorname{aff} Q$.

$$[Q = \{x\}] \Longrightarrow [\operatorname{aff} Q = \{x\},\quad \operatorname{par} Q = \{0\},\quad Q = \operatorname{ri} Q \quad \text{and} \quad \operatorname{rb} Q = \emptyset]$$

$$[Q = \mathbf{E}] \Longrightarrow [\operatorname{aff} Q = \mathbf{E},\quad \operatorname{par} Q = \mathbf{E},\quad Q = \operatorname{ri} Q \quad \text{and} \quad \operatorname{rb} Q = \emptyset]$$

Theorem: For any nonempty convex set $Q \subset \mathbf{E}$, the relative interior $\operatorname{ri} Q$ is nonempty.

Relative interior and Boundary

Proof: WLOG $0 \in Q$ so $\operatorname{aff} Q = \operatorname{span}(Q) = \operatorname{par} Q$ is a subspace, set $k = \dim(\operatorname{aff} Q)$.

Let $d_1, \ldots, d_k$ be a basis for $\operatorname{aff} Q$ and define $\mathcal{A} \in \mathbf{L}(\mathbf{R}^k, \mathbf{E})$ by $\mathcal{A}\lambda := \sum_{i=1}^k \lambda_i d_i$ so that $\operatorname{aff} Q = \operatorname{ran} \mathcal{A}$.

Consequently, $\mathcal{A}$ maps the open set $\Omega := \left\{ \lambda \in \mathbf{R}^d_{++} \;\middle|\; \sum_{i=1}^k \lambda_i < 1 \right\}$ onto a subset of $\operatorname{aff} Q$ that is open relative to the subspace $\operatorname{aff} Q$ ($\mathcal{A}$ is a linear isomorphism between $\mathbf{R}^k$ and $\operatorname{aff} Q$). Consequently, $\mathcal{A}\Omega$ is open relative to $\operatorname{aff} Q$.

Observe that $\forall\, \lambda \in \Omega,\ \mathcal{A}\lambda = (\sum_{i=1}^k \lambda_i d_i) + (1 - \sum_{i=1}^k \lambda_i) \cdot 0 \in Q$ by convexity. Hence $\mathcal{A}\Omega \subset Q$ implying $\mathcal{A}\Omega \subset \operatorname{ri} Q$.

Access Theorem for Convex Sets

Theorem: Let $Q \subset \mathbf{E}$ be convex. Then $x \in \operatorname{ri} Q$ if and only if $\forall\, y \in \operatorname{cl} Q,\ [x, y) \subset \operatorname{ri} Q$.

Proof:($\Leftarrow$) Trivial. ($\Rightarrow$) Let $y \in \operatorname{cl} Q$, $x \in \operatorname{ri} Q$, and $\epsilon > 0$ be such that $B_\epsilon(x) \cap \operatorname{aff} Q \subset Q$. Then, for $x \in \operatorname{ri} Q$ and $\lambda \in (0, 1]$, convexity tells us that

$$\begin{aligned} Q &\supset \lambda(B_\epsilon(x) \cap \operatorname{aff} Q) + (1-\lambda)y \\ &= \lambda(x + B_\epsilon(0) \cap \operatorname{par} Q) + (1-\lambda)y \\ &= \lambda x + (1-\lambda)y + \lambda B_\epsilon(0) \cap \operatorname{par} Q \\ &= B_{(\lambda\epsilon)}((1-\lambda)x + \lambda y) \cap \operatorname{aff} Q. \end{aligned}$$

pic

Access Theorem for Convex Sets

Theorem: Let $Q \subset \mathbf{E}$ be convex. Then $x \in \operatorname{ri} Q$ if and only if $\forall\, y \in \operatorname{cl} Q,\ [x, y) \subset \operatorname{ri} Q$.

Proof:($\Leftarrow$) Trivial. ($\Rightarrow$) Let $y \in \operatorname{cl} Q$, $x \in \operatorname{ri} Q$, and $\epsilon > 0$ be such that $B_\epsilon(x) \cap \operatorname{aff} Q \subset Q$. Then, for $x \in \operatorname{ri} Q$ and $\lambda \in (0, 1]$, convexity tells us that

$$\begin{aligned} Q &\supset \lambda(B_\epsilon(x) \cap \operatorname{aff} Q) + (1-\lambda)y \\ &= \lambda(x + B_\epsilon(0) \cap \operatorname{par} Q) + (1-\lambda)y \\ &= \lambda x + (1-\lambda)y + \lambda B_\epsilon(0) \cap \operatorname{par} Q \\ &= B_{(\lambda\epsilon)}((1-\lambda)x + \lambda y) \cap \operatorname{aff} Q. \end{aligned}$$

pic

Corollaries:

1) For any nonempty convex set Q in $\mathbf{E}$,

$$\operatorname{cl}(\operatorname{ri} Q) = \operatorname{cl} Q \quad \text{and} \quad \operatorname{ri}(\operatorname{cl} Q)) = \operatorname{ri} Q.$$

2) $x \in \operatorname{ri} Q \iff \forall\, y \in Q\ \exists\, \overline{\lambda} > 1 \text{ s.t. } y + \lambda(x - y) \in Q\ \ \forall\, \lambda \in (0, \overline{\lambda}]$.

3) $\operatorname{intr}(Q + [\operatorname{par} Q]^\perp) = \operatorname{ri} Q + [\operatorname{par} Q]^\perp$.

Linear Images of the Relative Interior

Theorem: Let $Q \subset \mathbf{E}$ be convex and $A \in \mathbf{L}[\mathbf{E}, \mathbf{Y}]$. Then

$$\mathrm{ri}\,(AQ) = A(\mathrm{ri}\,Q) \text{ and } \mathrm{cl}\,(AQ) \supset A(\mathrm{cl}\,Q).$$

Proof: The closure inclusion follows from continuity.
Next observe that

$$\mathrm{cl}\,A(\mathrm{ri}\,Q) \supset A(\mathrm{cl}\,\mathrm{ri}\,Q) = A(\mathrm{cl}\,Q) \supset AQ \supset A(\mathrm{ri}\,Q) \qquad .$$

Hence, AQ and $A(\mathrm{ri}\,Q)$have the same closure and relative interior which tells us that $\mathrm{ri}\,(AQ) \subset A(\mathrm{ri}\,Q)$. For the reverse inclusion, let $z \in A(\mathrm{ri}\,Q)$ and $y \in \mathrm{ri}\,Q$ such that $z = Ay$. Then for all $w \in Q$, $[y, w) \subset Q$ which implies that for all $x \in AQ$, $[Ay, x) \subset AQ$. That is, $z = Ay \in \mathrm{ri}\,AQ$ which establishes the reverse inclusion.

The Relative Interior of the Sum

Let Q_1, $Q_2 \subset \mathbf{E}$ be convex and $\alpha, \beta \in \mathbf{R}$, then

$$\operatorname{ri}(\alpha Q_1 + \beta Q_2) = \alpha \operatorname{ri} Q_1 \ + \ \beta \operatorname{ri} Q_2 \ .$$

Proof: Let $A \in \mathbf{L}[\mathbf{E} \times \mathbf{E}, \mathbf{E}]$ be given by $A(x, y) := \alpha x + \beta y$. Then

$$\begin{aligned} \operatorname{ri}(\alpha Q_1 + \beta Q_2) &= \operatorname{ri} A(Q_1 \times Q_2) \\ &= A \operatorname{ri}(Q_1 \times Q_2) \\ &\overset{\text{why?}}{=} A(\operatorname{ri} Q_1 \times \operatorname{ri} Q_2) \\ &= \alpha \operatorname{ri} Q_1 \ + \ \beta \operatorname{ri} Q_2 \, . \end{aligned}$$

Separation Theorems

Separation theorems allow us to analyze the geometry of a convex set $Q \subset \mathbf{X}$ by studying how the elements of the dual space $\mathbf{X}^*$ act on Q. This is the essence of *duality theory* which provides the foundation of convex analysis.

In a Euclidean space, separation theorems can built on the notion of the distance to a set. Given a set $X \subset \mathbf{E}$, we define the distance to X by

$$\operatorname{dist}(z \,|X\,) := \inf_{x \in X} \|x - z\| \quad (= d_X(z)).$$

If X is closed and nonempty, then, for all $z \in \mathbf{E}$, there is a $x \in X$ such that $\|z - x\| = \operatorname{dist}(z \,|X\,)$. We call the set of such *closest points in* X *to* z the projection of z onto X and write

$$\operatorname{proj}_X(y) := \{x \in X : d_Q(y) = \|x - y\|\}.$$

The Projection Theorem for Convex Sets

For any nonempty, closed, convex set $Q \subset \mathbf{E}$, the set $\text{proj}_Q(y)$ is a singleton. Moreover, the closest point $z \in Q$ to y is characterized by the property:

$$\langle y - z, x - z \rangle \leq 0 \qquad \text{for all } x \in Q. \qquad (\Diamond)$$

Proof: If $z \in Q$ satisfies $(\Diamond)$, then, for all $x \in Q$,

$$\|y - x\|^2 = \|y - z\|^2 + 2\langle y - z,\, z - x \rangle + \|z - x\|^2 \geq \|y - z\|^2$$

with equality if and only if $z = x$. Hence, $(\Diamond)$ implies z is the unique element of $\text{proj}_Q(y)$.

It remains to show that any $z \in \text{proj}_Q(y)$ must satisfy $(\Diamond)$. Define $\varphi(x) := \frac{1}{2}\|y - x\|^2$ so that $\nabla\varphi(x) = x - y$. If $z \in \text{proj}_Q(y)$, then, for all $x \in Q$,

$$\varphi'(z; x - z) = \lim_{t \downarrow 0} \frac{\varphi(z + t(x - z)) - \varphi(z)}{t} \geq 0 \quad \text{as } z + t(x - z) \in Q,\ t \in [0, 1].$$

So for all $x \in Q$, $0 \leq \varphi'(z; x - z) = \langle \nabla\varphi(z),\, x - z \rangle = \langle z - y,\, x - z \rangle$, which is $(\Diamond)$.

Strict Separation Theorem

Consider a nonempty, closed, convex set $Q \subset \mathbf{E}$ and a point $y \notin Q$. Then there exists a nonzero vector $z \in \mathbf{E}$ and a number $\beta \in \mathbf{R}$ satisfying

$$\langle z, x \rangle \leq \beta < \langle z, y \rangle \quad \text{for all } x \in Q.$$

Strict Separation Theorem

Consider a nonempty, closed, convex set $Q \subset \mathbf{E}$ and a point $y \notin Q$. Then there exists a nonzero vector $z \in \mathbf{E}$ and a number $\beta \in \mathbf{R}$ satisfying

$$\langle z, x \rangle \leq \beta < \langle z, y \rangle \qquad \text{for all } x \in Q.$$

Proof: Fix a point $y \notin Q$ and define the nonzero vector $z := y - \text{proj}_Q(y)$. Then for any $x \in Q$, the condition

$$\langle z, x - \text{proj}_Q(y) \rangle \leq 0 \qquad \text{for all } x \in Q$$

yields

$$\begin{aligned} \langle z, x \rangle \leq \langle z, \text{proj}_Q(y) \rangle &= \langle z, y \rangle + \left\langle z,\, \text{proj}_Q(y) - y \right\rangle \\ &= \langle z, y \rangle - \|z\|^2 < \langle z, y \rangle, \end{aligned}$$

as claimed, where $\beta := \langle z, \text{proj}_Q(y) \rangle$.

Supporting Hyperplanes to Points on the Relative Boundary

Theorem: Let $Q \subset \mathbf{E}$ be convex with $\bar{x} \in \operatorname{rb} Q$. Then there exists $\overline{z} \in \mathbf{E}$ such that

$$\langle \overline{z}, x \rangle \leq \langle \overline{z}, \bar{x} \rangle \ \forall\, x \in \operatorname{cl} Q \text{ and } \langle \overline{z}, x \rangle < \langle \overline{z}, \bar{x} \rangle \ \forall\, x \in \operatorname{ri} Q .$$

Proof: Set $\widehat{Q} := Q + [\operatorname{par} Q]^{\perp}$. Then $\operatorname{intr} \widehat{Q} = \operatorname{ri} Q + [\operatorname{par} Q]^{\perp} = \operatorname{ri} \widehat{Q}$. Since $\bar{x} \in \operatorname{rb} Q$, Q is not a single point and not all of $\mathbf{E}$ so $\widehat{Q} \neq \mathbf{E}$. Hence, there exists $\{x_k\} \subset \mathbf{E} \setminus \operatorname{cl} \widehat{Q}$ with $x_k \to x$. Let $\{z_k\} \subset \mathbf{E}$ be such that $\|z_k\| = 1$ and $\langle z_k, y \rangle \leq \langle z_k, x_k \rangle$ for all $y \in \widehat{Q}$, $k \in \mathbb{N}$. WLOG (why?) there is a $\overline{z} \in \mathbf{E}$ with $\|\overline{z}\| = 1$ such that $z_k \to \overline{z}$. Taking the limit, we have

$$\langle \overline{z}, y \rangle \leq \langle \overline{z}, x_k \rangle \ \forall\, y \in \operatorname{cl} \widehat{Q} \text{ and } \langle \overline{z}, y \rangle < \langle \overline{z}, x_k \rangle \ \forall\, y \in \operatorname{intr} \widehat{Q} .$$

Since $Q \subset \widehat{Q}$ and $\operatorname{ri} Q \subset \operatorname{intr} Q$, the result follows.

Dual Description of Convex Sets

Theorem: Given a nonempty set $Q \subset \mathbf{E}$, define the set of halfspaces

$$\mathcal{F}_Q := \{(a, b) \in \mathbf{E} \times \mathbf{R} : \langle a, x \rangle \leq b \quad \text{for all } x \in Q\}.$$

Then equality holds:

$$\operatorname{cl}\operatorname{conv}(Q) = \bigcap_{(a,b) \in \mathcal{F}_Q} \{x \in \mathbf{E} : \langle a, x \rangle \leq b\}. \tag{1}$$

Cones and Convex Cones

A set $K \subseteq \mathbf{E}$ is called a *cone* if the inclusion $\lambda K \subset K$ holds for any $\lambda \geq 0$.

In $\mathbf{R}^2$, the union of the x and y axes is a cone:

$$\{(x, 0) \,|\, x \in \mathbf{R}\} \cup \{(0, y) \,|\, y \in \mathbf{R}\}.$$

$\mathbf{R}^n_+$ and $\mathbf{S}^n_+$ are cones.

Theorem: A cone $K \subset \mathbf{E}$ is convex if and only if $K = K + K$.

Proposition: If $\subset \mathbf{E}$ is a convex cone, then aff $K = K - K$.

Cones and Polarity

The *polar cone* of a cone $K \subset \mathbf{E}$ is the set

$$K^\circ := \{v \in \mathbf{E} : \langle v, x\rangle \leq 0 \text{ for all } x \in K\}.$$

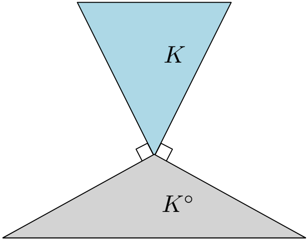


Figure: Polar cone

Cones and Polarity

The *polar cone* of a cone $K \subset \mathbf{E}$ is the set

$$K^\circ := \{v \in \mathbf{E} : \langle v, x \rangle \leq 0 \text{ for all } x \in K\}.$$

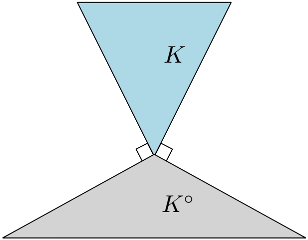


Figure: Polar cone

Polarity generalizes the notion of perpendicular subspaces: if S is a subspace, then $S^\circ = S^\perp$.

Cones and Polarity

The *polar cone* of a cone $K \subset \mathbf{E}$ is the set

$$K^\circ := \{v \in \mathbf{E} : \langle v, x \rangle \leq 0 \text{ for all } x \in K\}.$$

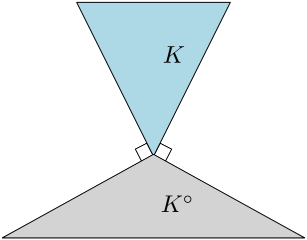


Figure: Polar cone

Polarity generalizes the notion of perpendicular subspaces: if S is a subspace, then $S^\circ = S^\perp$.

Theorem: [The Moreau Decomposition] Let $K \subset \mathbf{E}$ be a non-empty closed convex cone. Then for every $y \in \mathbf{E}$ there exists a unique pair $y_1 \in K$ and $y_2 \in K^\circ$ such that $y = y_1 + y_2$ with $\langle y_1, y_2 \rangle = 0$.

The Lineality of a Cone

Given a closed convex cone $K \subset \mathbf{E}$. The *lineality* of K, denoted $\operatorname{lin} L$, is the largest subspace contained in K.

The cone K is said to be *pointed* if $K \cap (-K) = \{0\}$, or equivalently, $\operatorname{lin} K = \{0\}$.

Show that $K^\circ \subset (\operatorname{lin} K)^\perp$.

Properties of the Polar

– For any nonempty cone $K \subset \mathbf{E}$, $(K^\circ)^\circ = \operatorname{cl}\operatorname{conv}(K)$.

– For any $\mathcal{A} \in \mathbf{L}[\mathbf{E}, \mathbf{Y}]$ and any nonempty cone $K \subset \mathbf{Y}$,
$$(\mathcal{A}K)^\circ = (\mathcal{A}^*)^{-1} K^\circ.$$

– For any two nonempty cones $K_1, K_2 \subset \mathbf{E}$, $(K_1 + K_2)^\circ = K_1^\circ \cap K_2^\circ$.

– Let $Q \subset \mathbf{E}$. We define the polar of Q to be the set
$$Q^\circ := \{z \,|\, \langle z, x\rangle \le 1 \ \forall\ x \in Q\}.$$
It is easy to see that if Q is a cone, this notion of polar coincides with cone polarity.

– For any nonempty $Q \subset \mathbf{E}$, $(Q^\circ)^\circ = \operatorname{cl}\operatorname{conv}(Q \cup \{0\})$.

– If $\mathbb{B}_\rho$ is the closed unit ball for some norm ρ, then $\mathbb{B}_\rho^\circ$ is the closed unit ball for its dual norm ρ^*, i.e. $\mathbb{B}_{\rho^*} = \mathbb{B}_\rho^\circ$.

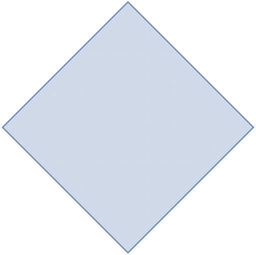
Visualizing the Polar of a Convex Set

Let $0 \in Q \subset \mathbf{E}$ and let K be the cone generated by $Q \times \{1\} \subset \mathbf{E} \times \mathbf{R}$, that is

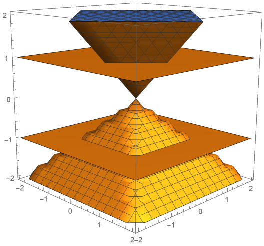
$$K = \{(\lambda x, \lambda) \in \mathbf{E} \times \mathbf{R} : x \in Q, \lambda \geq 0\}.$$

Since Q contains the origin, the polar cone K° is contained in $\mathbf{E} \times \mathbf{R}_-$. Then

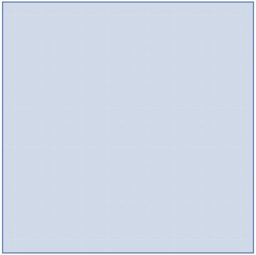
$$Q^\circ := \{x \in \mathbf{E} : (x, -1) \in K^\circ\}.$$



(a) $Q = \{x : \|x\|_1 \leq 1\}$



(b) Homogenization



(c) $Q^\circ = \{x : \|x\|_\infty \leq 1\}$

The Tangent Cone

The *tangent cone* to a set $Q \subset \mathbf{E}$ at a point $\bar{x} \in Q$ is the set

$$T_Q(\bar{x}) := \left\{ \lim_{i \to \infty} \tau_i^{-1}(x_i - \bar{x}) : x_i \to \bar{x} \text{ in } Q,\ \tau_i \searrow 0 \right\}.$$

The Tangent Cone

The *tangent cone* to a set $Q \subset \mathbf{E}$ at a point $\bar{x} \in Q$ is the set

$$T_Q(\bar{x}) := \left\{ \lim_{i\to\infty} \tau_i^{-1}(x_i - \bar{x}) : x_i \to \bar{x} \text{ in } Q,\ \tau_i \searrow 0 \right\}.$$

Proposition: If $Q \subset \mathbf{E}$ is convex, then

$$T_Q(\bar{x}) = \text{cl } \mathbf{R}_+(Q - \bar{x}) \qquad \forall\, x \in Q.$$

The Tangent Cone

The *tangent cone* to a set $Q \subset \mathbf{E}$ at a point $\bar{x} \in Q$ is the set

$$T_Q(\bar{x}) := \left\{ \lim_{i \to \infty} \tau_i^{-1}(x_i - \bar{x}) : x_i \to \bar{x} \text{ in } Q,\ \tau_i \searrow 0 \right\}.$$

Proposition: If $Q \subset \mathbf{E}$ is convex, then

$$T_Q(\bar{x}) = \mathrm{cl}\ \mathbf{R}_+(Q - \bar{x}) \qquad \forall\, x \in Q.$$

Corollary: If $Q \subset \mathbf{E}$ is polyhedral convex, then

$$T_Q(\bar{x}) = \mathbf{R}_+(Q - \bar{x}) \qquad \forall\, \bar{x} \in Q.$$

The Normal Cone

The *normal cone* to a set $Q \subset \mathbf{E}$ at a point $\bar{x} \in Q$ is the set

$$N_Q(\bar{x}) := \{v \in \mathbf{E} : \langle v, x - \bar{x} \rangle \leq o(\|x - \bar{x}\|) \quad \text{as } x \to \bar{x} \text{ in } Q\},$$

i.e., $v \in N_Q(\bar{x})$ if and only if

$$\limsup_{x \xrightarrow{Q} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0,$$

where the notation $x \xrightarrow{Q} \bar{x}$ means that x tends to $\bar{x}$ in Q.

The Normal Cone

The *normal cone* to a set $Q \subset \mathbf{E}$ at a point $\bar{x} \in Q$ is the set

$$N_Q(\bar{x}) := \{v \in \mathbf{E} : \langle v, x - \bar{x} \rangle \leq o(\|x - \bar{x}\|) \quad \text{as } x \to \bar{x} \text{ in } Q\},$$

i.e., $v \in N_Q(\bar{x})$ if and only if

$$\limsup_{x \xrightarrow{Q} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0,$$

where the notation $x \xrightarrow{Q} \bar{x}$ means that x tends to $\bar{x}$ in Q.

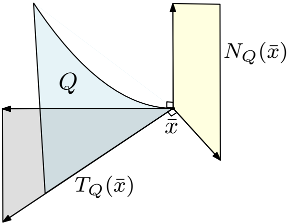


Figure: Illustration of the tangent and normal cones for nonconvex sets.

$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ$

Lemma: For any set $Q \subset \mathbf{E}$ and a point $\bar{x} \in Q$, the polarity relationship holds:

$$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ.$$

$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ$

Lemma: For any set $Q \subset \mathbf{E}$ and a point $\bar{x} \in Q$, the polarity relationship holds:

$$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ.$$

Corrolary: If $Q \subset \mathbf{E}$ is convex, then
$$N_Q(\bar{x}) = \{z \,|\, \langle z,\, x - bx \rangle \leq 0 \;\forall\, x \in Q\}.$$

$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ$

Lemma: For any set $Q \subset \mathbf{E}$ and a point $\bar{x} \in Q$, the polarity relationship holds:

$$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ.$$

Corrolary: If $Q \subset \mathbf{E}$ is convex, then
$$N_Q(\bar{x}) = \{z \,|\, \langle z,\, x - bx \rangle \leq 0 \;\forall\, x \in Q\,\}.$$

Corrolary: Let $Q \subset \mathbf{E}$ be convex with $\bar{x} \in Q$. Then $v \in N_Q(\bar{x})$ if and only if $\bar{x} = \mathrm{proj}_Q(\bar{x} + \lambda v)$ for some (or all) $\lambda > 0$.

$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ$

Lemma: For any set $Q \subset \mathbf{E}$ and a point $\bar{x} \in Q$, the polarity relationship holds:

$$N_Q(\bar{x}) = (T_Q(\bar{x}))^\circ.$$

Corrolary: If $Q \subset \mathbf{E}$ is convex, then
$$N_Q(\bar{x}) = \{z \,|\, \langle z,\, x - bx \rangle \leq 0 \;\forall\, x \in Q\}.$$

Corrolary: Let $Q \subset \mathbf{E}$ be convex with $\bar{x} \in Q$. Then $v \in N_Q(\bar{x})$ if and only if $\bar{x} = \mathrm{proj}_Q(\bar{x} + \lambda v)$ for some (or all) $\lambda > 0$.

Corrolary: Let $Q \subset \mathbf{E}$ be convex with $\bar{x} \in Q$. Then $v \in N_Q(\bar{x})$ if and only if $\bar{x} \in \mathrm{argmax}_{x \in Q} \langle v,\, x \rangle$.