CentraleSupélec

ST7 – Optimization
Part V.2: Linear programming

jean-christophe@pesquet.eu

Standard form 标准型.

秩，变量个数.

Let $A \in \mathbb{R}^{K\times M}$ such that $\operatorname{rank} A = K < M$, let $b \in \mathbb{R}^K$, and let $d \in \mathbb{R}^M$. We want

$$\underset{z\in[0,+\infty[^M}{\text{minimize}} \quad \langle d \mid z \rangle \quad \text{s.t.} \quad Az = b.$$

$b \in \mathbb{R}^K$.

$\in \mathbb{R}^M$

$\operatorname{rank}(A) = K$,

Remark: Assuming that $\operatorname{rank} A = K$ is not restrictive since, if this condition is not met, some lines of A correspond to either redundant or incompatible equality constraints.

In addition, if $M = K$ then there exists a unique solution to the equation $Az = b$, which makes the problem trivial.

$K < M$. 不定解.

Extreme points 极点.

vertex, corner point.

Let C be a convex set.
$z \in C$ is an extreme point of C if

$$(\exists (u,v) \in C^2) \quad z = \frac{u+v}{2}$$

$$\Rightarrow \quad u = v.$$

Remark:

- if z is an extreme point of a convex C in a Hilbert space, then $z \in \mathrm{bd}(C)$ $(= \overline{C} \setminus \mathrm{int}\,(C))$
otherwise, $z \in C \setminus \mathrm{bd}(C) \Rightarrow z \in \mathrm{int}\,(C)$ and there exists an open ball centered at z included in C
- the extreme points of a polyhedron are called vertices.

多面体.

Vertices

Assume that the standard problem admits a solution.
Then one of the vertices of the feasible set is a solution.

Proof: There exists a solution $\overline{z}$ with a maximum of zero components. Assume that $\overline{z}$ is not a vertex of the feasible set $\widetilde{\mathcal{A}}$. Then

反证法

$$(\exists (u, v) \in \widetilde{\mathcal{A}}^2) \quad u \neq v \quad \text{and} \quad \overline{z} = \frac{u+v}{2}.$$

Since $\overline{z}$ is a solution to the standard problem,

$$\langle d \mid u \rangle \geq \langle d \mid \overline{z} \rangle \quad \text{and} \quad \langle d \mid v \rangle \geq \langle d \mid \overline{z} \rangle .$$

On the other hand, since

$$\langle d \mid \overline{z} \rangle = \frac{1}{2}(\langle d \mid u \rangle + \langle d \mid v \rangle),$$

we have $\langle d \mid u \rangle = \langle d \mid \overline{z} \rangle = \langle d \mid v \rangle$.

Vertices

Assume that the standard problem admits a solution.
Then one of the vertices of the feasible set is a solution.

Proof: For every $\lambda \in \mathbb{R}$, let

$$z_\lambda = \bar{z} + \lambda(u - v).$$

Then

$$\begin{aligned}
\langle d \mid z_\lambda \rangle &= \langle d \mid \bar{z} \rangle + \lambda(\langle d \mid u \rangle - \langle d \mid v \rangle) = \langle d \mid \bar{z} \rangle \\
Az_\lambda &= A\bar{z} + \lambda(Au - Av) = b.
\end{aligned}$$

Vertices

Assume that the standard problem admits a solution.
Then one of the vertices of the feasible set is a solution.

Proof:

- Let $\mathbb{K} = \{ i \in \{1, \ldots, M\} \mid \overline{z}^{(i)} = 0 \}$.
  $(\forall i \in \mathbb{K})\ u^{(i)} = v^{(i)} = 0 \Rightarrow z_\lambda^{(i)} = 0$.
- Let $\mathbb{J} = \{ j \in \{1, \ldots, M\} \mid u^{(j)} \neq v^{(j)} \}$.
  $(\forall j \notin \mathbb{J})$ such that $j \notin \mathbb{K}$, $z_\lambda^{(i)} = \overline{z}^{(i)} > 0$
- Let us now consider indices in $\mathbb{J}$.
  We know that $\mathbb{J} \neq \varnothing$ and $\mathbb{J} \cap \mathbb{K} = \varnothing$.
  Suppose, for example, that $(\exists j \in \mathbb{J})\ v^{(j)} > u^{(j)}$.
  Let $\lambda = \min_{\substack{j \in \mathbb{J} \\ v^{(j)} > u^{(j)}}} \frac{\overline{z}^{(j)}}{v^{(j)} - u^{(j)}} = \frac{\overline{z}^{(j_0)}}{v^{(j_0)} - u^{(j_0)}} > 0$.
  Then $(\forall j \in \mathbb{J} \setminus \{j_0\})\ z_\lambda^{(j)} \geq 0$ and $z_\lambda^{(j_0)} = 0$.

z_λ is thus a solution to the standard problem with at least one more zero component than $\overline{z}$, which is impossible.

Basic solution 基础解. non degenerate. (非简并)

Let $(a_i)_{1\leq i\leq M}$ be the column vectors of A.
A solution $z = (z^{(i)})_{1\leq i\leq M}$ to the equation $Az = b$ is called a **basic solution** to this equation if $z = 0$ or if $\{a_i \mid i \in \mathbb{I}_0\}$ is a family of independent vectors where $\mathbb{I}_0 = \{i \in \{1, \dots, M\} \mid z^{(i)} \neq 0\}$.
If $\operatorname{card} \mathbb{I}_0 = K$ (resp. $\operatorname{card} \mathbb{I}_0 \neq K$), then z is said to be **non degenerate** (resp. **degenerate**).

Remark: If $z = (z^{(i)})_{1\leq i\leq M}$ is a basic solution, then there exists $\mathbb{I} \subset \{1, \dots, M\}$ such that

$$\begin{cases} (\forall i \in \{1, \dots, M\} \setminus \mathbb{I}) \quad z^{(i)} = 0 \\ A_{\mathbb{I}} = [a_i]_{i\in\mathbb{I}} \in \mathbb{R}^{K\times K} \text{ is invertible} \end{cases}$$

Basic solution

Let $(a_i)_{1\leq i\leq M}$ be the column vectors of A.
A solution $z=(z^{(i)})_{1\leq i\leq M}$ to the equation $Az=b$ is called a **basic solution** to this equation if $z=0$ or if $\left\{a_i \;\middle|\; i\in\mathbb{I}_0\right\}$ is a family of independent vectors where $\mathbb{I}_0=\left\{i\in\{1,\ldots,M\}\;\middle|\; z^{(i)}\neq 0\right\}$.
If $\operatorname{card}\mathbb{I}_0=K$ (resp. $\operatorname{card}\mathbb{I}_0\neq K$), then z is said to be **non degenerate** (resp. **degenerate**).

Remark: If $z=(z^{(i)})_{1\leq i\leq M}$ is a basic solution, then there exists $\mathbb{I}\subset\{1,\ldots,M\}$ such that

$$\begin{cases}(\forall i\in\{1,\ldots,M\}\setminus\mathbb{I}) & z^{(i)}=0\\ A_{\mathbb{I}}=[a_i]_{i\in\mathbb{I}}\in\mathbb{R}^{K\times K} & \text{is invertible}\end{cases}\Rightarrow \operatorname{card}\mathbb{I}=K.$$

可逆.

满秩.

Basic solution

Let $(a_i)_{1\le i\le M}$ be the column vectors of A.

A solution $z = (z^{(i)})_{1\le i\le M}$ to the equation $Az = b$ is called a basic solution to this equation if $z = 0$ or if $\{a_i \mid i \in \mathbb{I}_0\}$ is a family of independent vectors where $\mathbb{I}_0 = \{i \in \{1,\ldots,M\} \mid z^{(i)} \neq 0\}$.

If $\operatorname{card} \mathbb{I}_0 = K$ (resp. $\operatorname{card} \mathbb{I}_0 \neq K$), then z is said to be non degenerate (resp. degenerate).

Remark: If $z = (z^{(i)})_{1\le i\le M}$ is a basic solution, then there exists $\mathbb{I} \subset \{1,\ldots,M\}$ such that

$$\begin{cases} (\forall i \in \{1,\ldots,M\} \setminus \mathbb{I}) \quad z^{(i)} = 0 \\ A_{\mathbb{I}} = [a_i]_{i\in\mathbb{I}} \in \mathbb{R}^{K\times K} \text{ is invertible} \end{cases} \Rightarrow \quad \operatorname{card} \mathbb{I} = K.$$

Proof: Since $\operatorname{rank} A = K < M$, if z is a basic solution, then it suffices to complete $(a_i)_{i\in\mathbb{I}_0}$ with columns of A which correspond to zero components of z and are linearly independent.

Basic solution

Let $(a_i)_{1\leq i\leq M}$ be the column vectors of A.
A solution $z=(z^{(i)})_{1\leq i\leq M}$ to the equation $Az=b$ is called a basic solution to this equation if $z=0$ or if $\{a_i \mid i\in\mathbb{I}_0\}$ is a family of independent vectors where $\mathbb{I}_0=\{i\in\{1,\ldots,M\}\mid z^{(i)}\neq 0\}$.
If $\operatorname{card}\mathbb{I}_0=K$ (resp. $\operatorname{card}\mathbb{I}_0\neq K$), then z is said to be non degenerate (resp. degenerate).

Remark: If $z=(z^{(i)})_{1\leq i\leq M}$ is a basic solution, then there exists $\mathbb{I}\subset\{1,\ldots,M\}$ such that

$$\begin{cases}(\forall i\in\{1,\ldots,M\}\setminus\mathbb{I}) \quad z^{(i)}=0\\ A_{\mathbb{I}}=[a_i]_{i\in\mathbb{I}}\in\mathbb{R}^{K\times K} \text{ is invertible}\end{cases} \Rightarrow \quad \operatorname{card}\mathbb{I}=K.$$

$\mathbb{I}$ is called a basic index set.

Basic solution 基础解.

z is a vertex of the feasible set of the standard problem if and only if $z \in [0, +\infty[^M$ and z is a basic solution.

Proof: Assume that $z \in [0, +\infty[^M$ is a basic solution.
Suppose that there exist $u = (u^{(i)})_{1 \le i \le M}$ and $v = (v^{(i)})_{1 \le i \le M}$ in the feasible set $\widetilde{\mathcal{A}}$ such that $z = (u + v)/2$. Let
$\mathbb{I}_0 = \{i \in \{1, \ldots, M\} \mid z^{(i)} > 0\}$.
If $i \in \{1, \ldots, M\} \setminus \mathbb{I}_0$, then $z^{(i)} = 0 \Rightarrow u^{(i)} = v^{(i)} = 0$.
In addition, when $\mathbb{I}_0 \neq \varnothing$,

$$\begin{aligned} Au = Av = b \quad &\Rightarrow \quad A(u - v) = 0 \\ &\Leftrightarrow \quad \sum_{i=1}^{M} (u^{(i)} - v^{(i)}) a_i = \sum_{i \in \mathbb{I}_0} (u^{(i)} - v^{(i)}) a_i = 0. \end{aligned}$$

Since $\{a_i \mid i \in \mathbb{I}_0\}$ is a family of independent vectors, for every $i \in \mathbb{I}_0$, $u^{(i)} = v^{(i)}$. Hence, z is a vertex of $\widetilde{\mathcal{A}}$.

Basic solution

z is a vertex of the feasible set of the standard problem if and only if $z \in [0, +\infty[^M$ and z is a basic solution.

Proof: Conversely, assume that z is a vertex of $\widetilde{\mathcal{A}}$ and it is not a basic solution. There thus exists $w = (w^{(i)})_{1 \leq i \leq M} \in \mathbb{R}^M \setminus \{0\}$ such that $(\forall i \in \{1, \ldots, M\} \setminus \mathbb{I}_0)\ w^{(i)} = 0$ and $Aw = \sum_{i \in \mathbb{I}_0} w^{(i)} a_i = 0$.

Let $\epsilon > 0$. First note that $A(z \pm \epsilon w) = Az = b$.

Furthermore,

$(\forall i \in \{1, \ldots, M\} \setminus \mathbb{I}_0)\ z^{(i)} \pm \epsilon w^{(i)} = 0.$

$(\forall i \in \mathbb{I}_0)\ z^{(i)} > 0$

If $w^{(i)} = 0$, then $z^{(i)} \pm \epsilon w^{(i)} > 0$.

If $w^{(i)} \neq 0$, then $z^{(i)} \pm \epsilon w^{(i)} \geq z^{(i)} - \epsilon |w^{(i)}|$ and $z^{(i)} - \epsilon |w^{(i)}| > 0 \Leftrightarrow \epsilon < z^{(i)}/|w^{(i)}|$.

In summary, provided that ϵ is small enough, $z \pm \epsilon w \in \widetilde{\mathcal{A}}$. In addition $z = \frac{1}{2}\big((z + \epsilon w) + (z - \epsilon w)\big)$, which contradicts the fact that z is a vertex of $\widetilde{\mathcal{A}}$.

Naive algorithm

We look for all the possible feasible basic solutions:

1. We extract all the possible invertible matrices $A_{\mathbb{I}}$ where $\mathbb{I} \subset \{1, \ldots, M\}$ and $\operatorname{card} \mathbb{I} = K$.
2. We check that the associated basic solution has nonnegative components.
3. We look for the vector with minimum cost among those.

$\rightsquigarrow$ high computational cost of the order of C_M^K

Naïve.

Exercice 4

Let $M \in \mathbb{N} \setminus \{0, 1\}$. Solve the following problem:

$$\underset{(x^{(i)})_{1\leq i\leq M}\in[0,+\infty[^M}{\text{maximize}} \quad \sum_{i=1}^{M} i^2\, x^{(i)} \quad \boxed{\text{s.t.}} \quad \begin{cases} \displaystyle\sum_{i=1}^{M} x^{(i)} = 1 \\ \displaystyle\sum_{i=1}^{M} i\, x^{(i)} = 2. \end{cases}$$

最优化方法 - 凸集

Sep 17, 2015 in **Study** / Tagged in **Note, Optimization Methods**

最优化方法 - 凸集

凸集的定义、性质

设$S \subseteq E^n$，若对$\forall x^{(1)}, x^{(2)} \in S$及$\forall \lambda \in [0,1]$，都有$\lambda x^{(1)} + (1-\lambda)x^{(2)} \in S$，则称$S$为**凸集**。

设S_1和S_2是两个凸集，β实数，则

- $\beta S_1 = \{\beta x \mid x \in S_1\}$是凸集
- $S_1 + S_2 = \{x^{(1)} + x^{(2)} \mid x^{(1)} \in S_1, x^{(2)} \in S_2\}$是凸集
- $S_1 - S_2 = \{x^{(1)} - x^{(2)} \mid x^{(1)} \in S_1, x^{(2)} \in S_2\}$是凸集
- $S_1 \bigcap S_2$是凸集

极点和极方向的定义

- 极点

  设S是非空集合，$x \in S$，若x不能表示成S中两个不同点的凸组合，即若假设$x = \lambda x^{(1)} + (1-\lambda)x^{(2)}$，必推出$x = x^{(1)} = x^{(2)}$，则称$x$是凸集$S$的**极点**。

- 方向

  设S是闭凸集，d为非零向量，如果对S中的每一个x，有$\{x + \lambda d \mid \lambda \geq 0\} \subset S$，则称$d$是$S$的**方向**。

  > *设$d^{(1)}$和$d^{(2)}$是S的两个方向，若对任何正数λ，有$d^{(1)} \neq \lambda d^{(2)}$，则称$d^{(1)}$和$d^{(2)}$是两个不同的方向。*

> *设$S = \{x \mid Ax = b, x \geq 0\} \neq \emptyset$，$d$是非零向量，则$d$是$S$的方向 $\iff$ $d \geq 0$且$Ad = 0$。*

- 极方向

  若S的方向d不能表示成该集合的两个不同方向的正的线性组合，则称d为S的**极方向**。

例：设$S = \{(x_1, x_2)^T \mid x_2 \geq |x_1|\}, d^{(1)} = (1,1)^T, d^{(2)} = (-1,1)^T$，则$d^{(1)}, d^{(2)}$是$S$的极方向。

解：对$\forall x \in S, \forall \lambda \geq 0$，有

$x + \lambda d^{(1)} = (x_1, x_2)^T + \lambda(1,1)^T = (x_1 + \lambda, x_2 + \lambda)^T$

$x \in S \implies x_2 \geq |x_1|$

而$x_2 + \lambda \geq |x_1| + \lambda \geq |x_1 + \lambda|$,

$\implies \{x + \lambda d^{(1)} \mid \lambda \geq 0\} \subset S$

故$d^{(1)}$是S的方向。

设$d^{(1)} = \lambda_1(x_1, x_2)^T + \lambda_2(y_1, y_2)^T$，其中$\lambda_1, \lambda_2 > 0, (x_1, x_2)^T, (y_1, y_2)^T$是$S$的方向，则有

$\begin{cases} 1 = \lambda_1 x_1 + \lambda_2 y_1 \\ 1 = \lambda_1 x_2 + \lambda_2 y_2 \end{cases} \implies \lambda_1 x_1 + \lambda_2 y_1 = \lambda_1 x_2 + \lambda_2 y_2$

$\implies x_1 = \frac{\lambda_2}{\lambda_1}(y_2 - y_1) + x_2$

$(x_1, x_2)^T, (y_1, y_2)^T$是$S$的方向,

$\implies x_2 \geq |x_1|, y_2 \geq |y_1|, (x_1, x_2)^T \neq 0, (y_1, y_2)^T \neq 0$

$\implies x_2 \geq |x_1| = \left|\frac{\lambda_2}{\lambda_1}(y_2 - y_1) + x_2\right| \implies y_2 \leq y_1$

$y_2 \geq |y_1| \implies y_2 = y_1 \implies x_2 = x_1 \implies (x_1, x_2)^T = \frac{x_1}{y_1}(y_1, y_2)^T$

故$d^{(1)}$是S的极方向。

- 多面集的表示定理

  设$S = \{x \mid Ax = b, x \geq 0\}$为非空多面集，则有

  - 极点集非空，且存在有限个极点$x^{(1)}, \cdots, x^{(k)}$
  - 极方向集合为空集 $\iff$ S有界。若S无界，则存在有限个极方向$d^{(1)}, d^{(2)}, \cdots, d^{(l)}$
  - $x \in S \iff x = \sum_{j=1}^{k} \lambda_j x^{(j)} + \sum_{j=1}^{l} \mu_j d^{(j)}$

    其中$\lambda_j \geq 0, j = 1, 2, \cdots, k, \sum_{j=1}^{k} \lambda_j = 1$

    $\mu_j \geq 0, j = 1, 2, \cdots, l$

$\mu_j \geq 0, j = 1, 2, \cdots, l$

凸集分离定理

设S_1和S_2是E^n中两个非空集合，

$H = \{x \mid p^T x = \alpha\}$为超平面，

如果对$\forall x \in S_1$，都有$p^T x \geq \alpha$，

对$\forall x \in S_2$，都有$p^x \leq \alpha$，

则称超平面H分离集合S_1和S_2。

Farkas定理

设A为$m \times n$矩阵，c为n维列向量，

则$Ax \leq 0, c^T x > 0$有解，

$\iff A^T y = c, y \geq 0$无解。

证：$\implies$

设存在$y \geq 0$，使得$A^T y = c$

则$y^T A = c^T$

设$\overline{x}$为$Ax \leq 0, c^T x > 0$的一个解，

则有$A\overline{x} \leq 0, c^T \overline{x} > 0$

$\implies y^T A\overline{x} = c^T \overline{x} > 0 \quad (1)$

$y \geq 0, A\overline{x} \leq 0 \implies y^T A\overline{x} \leq 0$与$(1)$矛盾。

$\impliedby$

设$A^T y = c, y \geq 0$无解，令$S = \{z \mid z = A^T y, y \geq 0\}$，则$c \notin S$

可以证明S为闭凸集，由凸集分离定理知，

$\exists x \neq 0, \varepsilon > 0$，使得对

$\forall z \in S$，有$x^T c \geq \varepsilon + x^T z$

$\varepsilon > 0 \implies x^T c > x^T z$

$\implies c^T x > z^T x = y^T Ax$

即对任意的$y \geq 0$，有$c^T x > y^T Ax \quad (2)$

令$y = 0$，得$c^T x > 0$

$c^T x$为一定数，y的分量可取任意大

$\implies$ 由(2)，必有$Ax \leq 0$

故非零向量x是$Ax \leq 0, c^T x > 0$的解。

例题

例：设A是$m \times n$矩阵，B是$l \times n$矩阵，$c \in E^n$，证明下列两个系统恰有一个有解：

系1 $Ax \leq 0, Bx = 0, c^T x > 0$，对某些$x \in E^n$。

系2 $A^T y + B^T z = c, y \geq 0$，对某些$y \in E^n$和$z \in E^l$。

证：$Bx = 0$ 等价于 $\begin{cases} Bx \leq 0 \\ Bx \geq 0 \end{cases}$

故系统1有解，即

$\begin{bmatrix} A \\ B \\ -B \end{bmatrix} x \leq 0, c^T x > 0$有解。

由Farkas定理知，

$\begin{pmatrix} A^T & B^T & -B^T \end{pmatrix} \begin{bmatrix} y \\ u \\ v \end{bmatrix} = c, \begin{bmatrix} y \\ u \\ v \end{bmatrix} \geq 0$无解。

令$z = u - v$，则

$A^T y + B^T z = c, y \geq 0$无解。

即系统2无解。

反之，若系统2有解。即

$\begin{pmatrix} A^T & B^T & -B^T \end{pmatrix} \begin{bmatrix} y \\ u \\ v \end{bmatrix} = c, \begin{bmatrix} y \\ u \\ v \end{bmatrix} \geq 0$有解。

由Farkas定理，知

$\begin{bmatrix} A \\ B \end{bmatrix} x \leq 0, c^T x > 0$无解。

$\begin{bmatrix} -B \end{bmatrix}$

即$Ax \leq 0, Bx = 0, c^T x > 0$无解，亦即系统1无解。

综上可得，两个系统恰有一个有解。

Gordan定理

设A为$m \times n$矩阵，

则$Ax < 0$有解，

$\iff A^T y = 0, y \geq 0 (y \neq 0)$无解。

证： $\implies$

设存在$\bar{x}$，使得$A\bar{x} < 0$

若存在非零向量$y \geq 0$，使得$A^T y = 0$

则有$y^T A = 0$， $\implies y^T A\bar{x} = 0$

$A\bar{x} < 0 \implies y$的各分量不可能为非负数，与$y \geq 0$矛盾。

$\impliedby$

（证等价命题）即若$Ax < 0$无解，则存在非零向量$y \geq 0$，使得$A^T y = 0$

设$Ax < 0$无解，令$S_1 = \{z \mid z = Ax, x \in E^n\}, S_2 = \{z \mid z < 0\}$

$Ax < 0$无解 $\implies S_1 \bigcap S_2 = \emptyset$

由分离定理知，存在非零向量y，使得对$\forall x \in E^n, \forall z \in S_2$，有$y^T Ax \geq y^T z \quad (1)$

特别地，当$x = 0$时，有$y^T z \leq 0$。

$z < 0$，它的分量可取任意负数 $\implies y \geq 0$

在(1)中令$z \to 0$，则对$\forall x \in E^n$，有

$y^T Ax \geq 0 \quad (2)$

令$x = -A^T y$，代入(2)，得$-y^T AA^T y \geq 0$

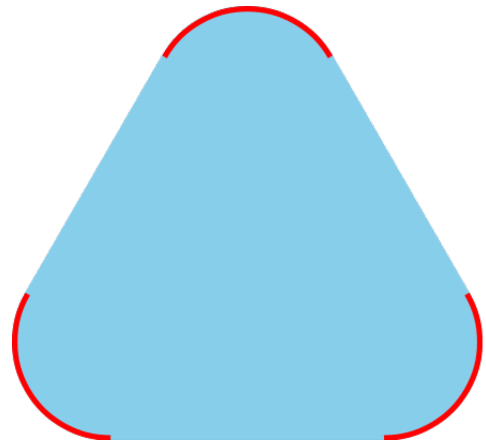
即$-\|A^T y\| \geq 0 \implies A^T y = 0$

故存在非零向量$y \geq 0$，使得$A^T y = 0$

WIKIPEDIA

Extreme point

In mathematics, an **extreme point** of a convex set S in a real vector space is a point in S which does not lie in any open line segment joining two points of S. In linear programming problems, an extreme point is also called vertex or corner point of S.[1]



A convex set in light blue, and its extreme points in red.

Contents

Definition

Throughout, it is assumed that S is a real or complex vector space.

For any $x, x_1, x_2 \in S$, say that x **lies between**[2] x_1 and x_2 if $x_1 \neq x_2$ and there exists a $0 < t < 1$ such that $x = tx_1 + (1 - t)x_2$.

If K is a subset of S and $x \in K$, then x is called an **extreme point**[2] of K if it does not lie between any two *distinct* points of K. That is, if there does *not* exist $x_1, x_2 \in K$ and $0 < t < 1$ such that $x_1 \neq x_2$ and $x = tx_1 + (1 - t)\, x_2$. The set of all extreme points of K is denoted by $\operatorname{extreme}(K)$.

Characterizations

The **midpoint**[2] of two elements x and y in a vector space is the vector $\frac{1}{2}(x + y)$.

For any elements x and y in a vector space, the set $[x, y] := \{tx + (1 - t)y : 0 \leq t \leq 1\}$ is called the **closed line segment** or **closed interval** between x and y. The **open line segment** or **open interval** between x and y is $(x, x) := \varnothing$ when $x = y$ while it is $(x, y) := \{tx + (1 - t)y : 0 < t < 1\}$ when $x \neq y$.[2] The points x and y are called the **endpoints** of these interval. An interval is said to be **non-degenerate** or **proper** if its endpoints are distinct. The **midpoint** of an interval is the midpoint of its endpoints.

Note that $[x, y]$ is equal to the convex hull of $\{x, y\}$ so if K is convex and $x, y \in K$, then $[x, y] \subseteq K$.

If K is a nonempty subset of X and F is a nonempty subset of K, then F is called a **face**[2] of K if whenever a point $p \in F$ lies between two points of K, then those two points necessarily belong to F.

> **Theorem**[2] — Let K be a non-empty convex subset of a vector space X and let $p \in K$. Then the following are equivalent:
>
> 1. p is an extreme point of K;
> 2. $K \setminus \{ p \}$ is convex;
> 3. p is not the midpoint of a non-degenerate line segment contained in K;
> 4. for any $x, y \in K$, if $p \in [x, y]$ then $x = p$ or $y = p$;
> 5. if $x \in X$ is such that both $p + x$ and $p - x$ belong to K, then $x = 0$;
> 6. $\{ p \}$ is a face of K.

Examples

- If $a < b$ are two real numbers then a and b are extreme points of the interval $[a, b]$. However, the open interval (a, b) has no extreme points.[2]
- An injective linear map $F : X \to Y$ sends the extreme points of a convex set $C \subseteq X$ to the extreme points of the convex set $F(C)$.[2] This is also true for injective affine maps.
- The perimeter of any convex polygon in the plane is a face of that polygon.[2]
- The vertices of any convex polygon in the plane $\mathbb{R}^2$ are the extreme points of that polygon.
- The extreme points of the closed unit disk in $\mathbb{R}^2$ is the unit circle.
- Any open interval in $\mathbb{R}$ has no extreme points while any non-degenerate closed interval not equal to $\mathbb{R}$ does have extreme points (i.e. the closed interval's endpoint(s)). More generally, any open subset of finite-dimensional Euclidean space $\mathbb{R}^n$ has no extreme points.

Properties

The extreme points of a compact convex form a Baire space (with the subspace topology) but this set may *fail* to be closed in X.[2]

Theorems

Krein–Milman theorem

The Krein–Milman theorem is arguably one of the most well-known theorems about extreme points.

> **Krein–Milman theorem** — If S is convex and compact in a locally convex space, then S is the closed convex hull of its extreme points: In particular, such a set has extreme points.

For Banach spaces

These theorems are for Banach spaces with the Radon–Nikodym property.

A theorem of Joram Lindenstrauss states that, in a Banach space with the Radon–Nikodym property, a nonempty closed and bounded set has an extreme point. (In infinite-dimensional spaces, the property of compactness is stronger than the joint properties of being closed and being bounded).[3]

> **Theorem** (Gerald Edgar) — Let E be a Banach space with the Radon-Nikodym property, let C be a separable, closed, bounded, convex subset of E, and let a be a point in C. Then there is a probability measure p on the universally measurable sets in C such that a is the barycenter of p, and the set of extreme points of C has p-measure 1.[4]

Edgar's theorem implies Lindenstrauss's theorem.

k-extreme points

More generally, a point in a convex set S is ***k*-extreme** if it lies in the interior of a k-dimensional convex set within S, but not a $k+1$-dimensional convex set within S. Thus, an extreme point is also a 0-extreme point. If S is a polytope, then the k-extreme points are exactly the interior points of the k-dimensional faces of S. More generally, for any convex set S, the k-extreme points are partitioned into k-dimensional open faces.

The finite-dimensional Krein-Milman theorem, which is due to Minkowski, can be quickly proved using the concept of k-extreme points. If S is closed, bounded, and n-dimensional, and if p is a point in S, then p is k-extreme for some $k < n$. The theorem asserts that p is a convex combination of extreme points. If $k = 0$, then it's trivially true. Otherwise p lies on a line segment in S which can be maximally extended (because S is closed and bounded). If the endpoints of the segment are q and r, then their extreme rank must be less than that of p, and the theorem follows by induction.

See also

- Choquet theory

Citations

1. Saltzman, Matthew. "What is the difference between corner points and extreme points in linear programming problems?" (https://www.quora.com/What-is-the-difference-between-corner-points-and-extreme-points-in-linear-programming-problems).
2. Narici & Beckenstein 2011, pp. 275-339.
3. Artstein, Zvi (1980). "Discrete and continuous bang-bang and facial spaces, or: Look for the extreme points". *SIAM Review*. **22** (2): 172–185. doi:10.1137/1022026 (https://doi.org/10.1137%2F1022026). JSTOR 2029960 (https://www.jstor.org/stable/2029960). MR 0564562 (https://www.ams.org/mathscinet-getitem?mr=0564562).
4. Edgar GA. A noncompact Choquet theorem. (http://www.ams.org/journals/proc/1975-049-02/S0002-9939-1975-0372586-2/S0002-9939-1975-0372586-2.pdf) Proceedings of the American Mathematical Society. 1975;49(2):354-8.

Bibliography

- Adasch, Norbert; Ernst, Bruno; Keim, Dieter (1978). *Topological Vector Spaces: The Theory Without Convexity Conditions*. Lecture Notes in Mathematics. **639**. Berlin New York: Springer-Verlag. ISBN 978-3-540-08662-8. OCLC 297140003 (https://www.worldcat.org/oclc/297140003).
- Bourbaki, Nicolas (1987) [1981]. *Sur certains espaces vectoriels topologiques* (http://www.numdam.org/item?id=AIF_1950__2__5_0) [*Topological Vector Spaces: Chapters 1–5*]. *Annales de l'Institut Fourier*. Éléments de mathématique. **2**. Translated by Eggleston, H.G.; Madan, S. Berlin New York: Springer-Verlag. ISBN 978-3-540-42338-6. OCLC 17499190 (https://www.worldcat.org/oclc/17499190).
- Paul E. Black, ed. (2004-12-17). "extreme point" (https://xlinux.nist.gov/dads/HTML/extremepoint.html). *Dictionary of algorithms and data structures*. US National institute of standards and technology. Retrieved 2011-03-24.
- Borowski, Ephraim J.; Borwein, Jonathan M. (1989). "extreme point". *Dictionary of mathematics*. Collins dictionary. Harper Collins. ISBN 0-00-434347-6.
- Grothendieck, Alexander (1973). *Topological Vector Spaces* (https://archive.org/details/topologicalvecto0000grot). Translated by Chaljub, Orlando. New York: Gordon and Breach Science Publishers. ISBN 978-0-677-30020-7. OCLC 886098 (https://www.worldcat.org/oclc/886098).
- Jarchow, Hans (1981). *Locally convex spaces*. Stuttgart: B.G. Teubner. ISBN 978-3-519-02224-4. OCLC 8210342 (https://www.worldcat.org/oclc/8210342).
- Köthe, Gottfried (1969). *Topological Vector Spaces I*. Grundlehren der mathematischen Wissenschaften. **159**. Translated by Garling, D.J.H. New York: Springer Science & Business Media. ISBN 978-3-642-64988-2. MR 0248498 (https://www.ams.org/mathscinet-getitem?mr=0248498). OCLC 840293704 (https://www.worldcat.org/oclc/840293704).
- Köthe, Gottfried (1979). *Topological Vector Spaces II*. Grundlehren der mathematischen Wissenschaften. **237**. New York: Springer Science & Business Media. ISBN 978-0-387-90400-9. OCLC 180577972 (https://www.worldcat.org/oclc/180577972).
- Narici, Lawrence; Beckenstein, Edward (2011). *Topological Vector Spaces*. Pure and applied mathematics (Second ed.). Boca Raton, FL: CRC Press. ISBN 978-1584888666. OCLC 144216834 (https://www.worldcat.org/oclc/144216834).

- Robertson, Alex P.; Robertson, Wendy J. (1980). *Topological Vector Spaces*. Cambridge Tracts in Mathematics. **53**. Cambridge England: Cambridge University Press. ISBN 978-0-521-29882-7. OCLC 589250 (https://www.worldcat.org/oclc/589250).
- Rudin, Walter (1991). *Functional Analysis* (https://archive.org/details/functionalanalys00rudi). International Series in Pure and Applied Mathematics. **8** (Second ed.). New York, NY: McGraw-Hill Science/Engineering/Math. ISBN 978-0-07-054236-5. OCLC 21163277 (https://www.worldcat.org/oclc/21163277).
- Schaefer, Helmut H.; Wolff, Manfred P. (1999). *Topological Vector Spaces*. GTM. **8** (Second ed.). New York, NY: Springer New York Imprint Springer. ISBN 978-1-4612-7155-0. OCLC 840278135 (https://www.worldcat.org/oclc/840278135).
- Schechter, Eric (1996). *Handbook of Analysis and Its Foundations*. San Diego, CA: Academic Press. ISBN 978-0-12-622760-4. OCLC 175294365 (https://www.worldcat.org/oclc/175294365).
- Trèves, François (2006) [1967]. *Topological Vector Spaces, Distributions and Kernels*. Mineola, N.Y.: Dover Publications. ISBN 978-0-486-45352-1. OCLC 853623322 (https://www.worldcat.org/oclc/853623322).
- Wilansky, Albert (2013). *Modern Methods in Topological Vector Spaces*. Mineola, New York: Dover Publications, Inc. ISBN 978-0-486-49353-4. OCLC 849801114 (https://www.worldcat.org/oclc/849801114).