Convex Sets (I)

Lecture 1, Nonlinear Programming, (Part b)

National Taiwan University

September 13, 2016

Table of contents

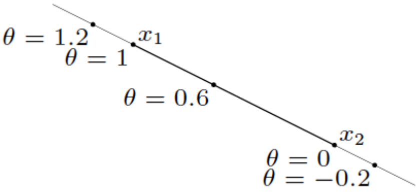
Line

Line

Let $x_1, x_2 \in \mathbf{R}^n$ and $x_1 \neq x_2$. The set of all points

$$\{\theta x_1 + (1-\theta)x_2 \mid \theta \in \mathbf{R}\}$$

is called a **line** passing through x_1 and x_2.



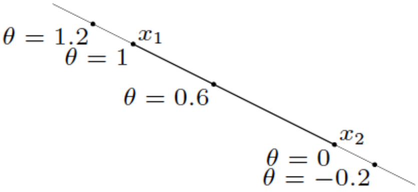
Line Segment

Line Segment

Let $x_1, x_2 \in \mathbf{R}^n$ and $x_1 \neq x_2$. The set of all points

$$\{\theta x_1 + (1-\theta)x_2 \mid \theta \in \mathbf{R}, 0 \leq \theta \leq 1\}$$

is called a **(closed) line segment** between x_1 and x_2.



Line and Line Segment

Line and Line Segment

Let $x_1, x_2 \in \mathbf{R}^n$ and $x_1 \neq x_2$. The set of all points

$$\{\theta x_1 + (1-\theta)x_2 \mid \theta \in \mathbf{R}\}$$

is called a **line** passing through x_1 and x_2. The set of all points

$$\{\theta x_1 + (1-\theta)x_2 \mid \theta \in \mathbf{R}, 0 \leq \theta \leq 1\}$$

is called a **(closed) line segment** between x_1 and x_2.

Another interpretation:

$$y = x_2 + \theta(x_1 - x_2)$$

is the sum of the **base point** x_2 and the **direction** $x_1 - x_2$ scaled by the parameter θ.

Affine Sets

Affine Sets

A set $C \subseteq \mathbf{R}^n$ is **affine** if the line through any two distinct points in C lies in C. That is,

$$x_1, x_2 \in C, \theta \in \mathbf{R} \quad \Rightarrow \quad \theta x_1 + (1-\theta)x_2 \in C.$$

Affine Combination

Let $x_1, x_2, \cdots, x_k \in \mathbf{R}^n$. Then, a point of the form

$$\theta_1 x_1 + \cdots + \theta_k x_k$$

with $\theta_1 + \cdots + \theta_k = 1$ is referred to as an **affine combination** of the points $x_1, x_2, \cdots, x_k$.

Affine Sets

Affine Sets and Subspaces

If C is an affine set and $x_0 \in C$, then the set

$$V = C - x_0 = \{x - x_0 \mid x \in C\}$$

is a subspace.

* Note that the subspace V associated with C does not depend on the choice of x_0.
Proof:

Dimension of Affine Set

Dimension of Affine Set

The **dimension** of an affine set C is defined as the dimension of the subspace $V = C - x_0$ where x_0 is any element of C.

Example: Solution set of linear equations (1/2)

Solution set of linear equations

The solution set of a system of linear equations

$$C = \{x \mid Ax = b\}$$

where $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$ is an affine set.

Proof:

Example: Solution set of linear equations (2/2)

Solution set of linear equations

The solution set of a system of linear equations

$$C = \{x \mid Ax = b\}$$

where $A \in \mathbf{R}^{m\times n}$ and $b \in \mathbf{R}^m$ is an affine set.

- The subspace associated with the affine set C is the nullspace of A.
- Converse: every affine set can be expressed as the solution set of system of linear equations.

Affine Hull

Affine Hull

The set of all affine combinations of points in some set $C \subseteq \mathbf{R}^n$ is called the **affine hull** of C, denoted **aff C**:

$$\textbf{aff } C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_1, \ldots, x_k \in C, \theta_1 + \cdots + \theta_k = 1\}.$$

The affine hull is the smallest affine set that contains C:

- If S is any affine set with $C \subseteq S$, then $\textbf{aff}\, C \subseteq S$.

Affine Dimension

Affine dimension

The affine dimension of C, a subset of $\mathbf{R}^n$, is defined by the dimension of its affine hull.

Example

Let $C = \left\{x \in \mathbf{R}^2 \mid x_1^2 + x_2^2 = 1\right\}$. What is the affine dimension of C?

Interior

Interior point

An element $x \in C \subseteq \mathbf{R}^n$ is called an **interior point** of C if there exists an $\epsilon > 0$ for which

$$\{y \mid ||y - x||_2 \leq \epsilon\}$$

is a subset of C.

Interior

The set of all points interior to C is called the **interior** of C, denoted **int C**:

$$\mathbf{int}\, C = \{y \mid y \in C \text{ and } y \text{ is an interior point of } C\}$$

Relative Interior

Consider a set $C \subseteq \mathbf{R}^n$ whose affine dimension is less than n. That is, $\mathbf{aff}\, C \neq \mathbf{R}^n$. What is the interior of C?

Relative Interior

The **relative interior** of the set C, denoted $\mathbf{relint}\, C$, is defined as its interior relative to **aff** C:

$$\mathbf{relint}\ C = \{x \in C \mid B(x,r) \cap \mathbf{aff}\ C \subseteq C \text{ for some } r > 0\}$$

where $B(x,r) = \{y \mid ||y - x||_2 \leq r\}$.

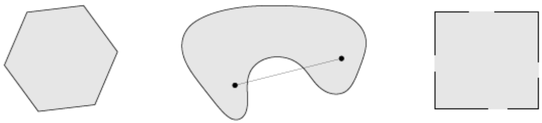
Convex Sets

Convex Set

A set C is **convex** if the line segment between any two points in C lies in C. That is, for any $x_1, x_2 \in C$ and any θ with $0 \leq \theta \leq 1$, we have

$$\theta x_1 + (1 - \theta) x_2 \in C.$$

Example: which of following is convex?



Example: Every affine set is also convex. Any line segment is also convex.

Convex Combination

Convex combination

A point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$, where $\theta_1 + \cdots + \theta_k = 1$ and $\theta_i \geq 0$, $i = 1, \ldots, k$, is called a convex combination of the points $x_1, \ldots, x_k$.

Property

A set is convex if and only if it contains every convex combination of its points.

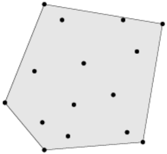
Convex Hull

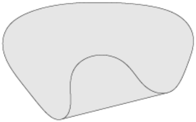
Convex Hull

The convex hull of a set C, denoted conv C, is the set of all convex combinations of points in C:

$$\mathbf{conv}\ C = \{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C,\ \theta_i \geq 0,\ i = 1, ..., k,\ \theta_1 + \cdots + \theta_k = 1\}.$$

Property: the convex hull **conv** C is always convex. It is the smallest convex set that contains C.





Generalized Definitions of Convex Combinations

- Infinite sum:
  - If C is convex and let $x_1, x_2, \ldots \in C$, then $\sum_{i=1}^{\infty} \theta_i x_i \in C$ where $\theta_i \geq 0, i = 1, 2, \ldots$ and $\sum_{i=1}^{\infty} \theta_i = 1$.
- Integral:
  - Let C be a convex set. Consider a function $p : \mathbf{R}^n \to \mathbf{R}$ that satisfies $p(x) \geq 0 \forall x \in C$ and $\int_C p(x) dx = 1$. Then $\int_C p(x) x \; dx \in C$.
- Probability distributions (most general form)
  - Suppose $C \subseteq \mathbf{R}^n$ is convex and x is a random vector with $x \in C$ with probability one. Then $\mathbf{E}[x] \in C$.

On Various Types of "Combinations"

Compare "linear combination," "affine combination," and "convex combination". All of these three types of combinations can be defined as the set $\{\theta_1 x_1 + \cdots + \theta_k x_k\}$ with certain constraints on the coefficients $\theta_1, \cdots, \theta_k$.

Type	Constraints on θ_i	Set of all combinations
linear combination	$\theta_1, ..., \theta_k \in \mathbf{R}$	span
affine combination	$\theta_1 + ... + \theta_k = 1$	affine hull
convex combination	$\theta_1 + ... + \theta_k = 1,\ \theta_i \geq 0$	convex hull

Cones

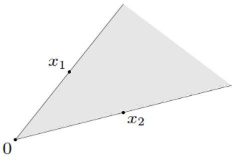
Cone

A set C is called a **cone** if for every $x \in C$ and $\theta \geq 0$ we have $\theta x \in C$. The set C is also said to be **nonnegative homogeneous**.

Convex Cone

A set C is called a **convex cone** if it is convex and is a cone. That is, for any $x_1, x_2 \in C$ and $\theta_1, \theta_2 \geq 0$ we have

$$\theta_1 x_1 + \theta_2 x_2 \in C.$$



Conic Combination

Conic combination

A point of the form $\theta_1 x_1 + \cdots + \theta_k x_k$ with $\theta_1, \ldots, \theta_k \geq 0$ is called a **conic combination** (or a **nonnegative linear combination**) of $x_1, x_2, \ldots, x_k$.

- Property: If x_i are in a convex cone C, then every conic combination of x_i is in C.
- Property: A set C is a convex cone if and only if it contains all conic combinations of its elements.
- Generalized definitions: the idea of conic combination can be generalized to infinite sums and integrals.

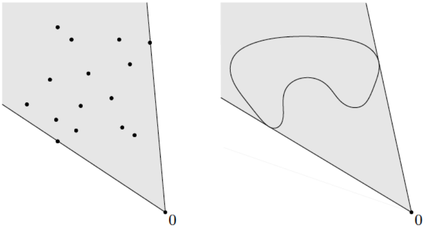
Conic Hull

Conic hull

The conic hull of a set C is the set of all conic combinations of points in C:

$$\{\theta_1 x_1 + \cdots + \theta_k x_k \mid x_i \in C,\ \theta_i \geq 0,\ i = 1, \ldots, k\}.$$

Property: The conic hull of a set C is the smallest convex cone that contains C.



Some Simple Examples of Affine / Convex Sets / Cones

- The empty set ϕ is affine (and hence convex).
- Any single point (i.e., **singleton**) $\{x_0\}$ is affine (and convex).
- The whole space $\mathbf{R}^n$ is affine (and convex).
- Any subspace is affine, and a convex cone.
- Any line is affine. If it passes through zero, it is a subspace, and also a convex cone.
- A line segment is convex, but is in general not affine.
- A **ray**, having the form $\{x_0 + \theta v \mid \theta \geq 0\}$, where $v \neq 0$, is convex but not affine. If $x_0 = 0$, then it is a convex cone.

Hyperplane

Hyperplane

A **hyperplane** is a set of the form

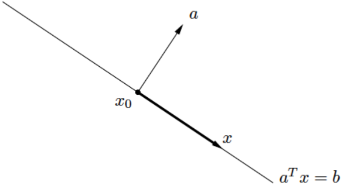
$$\left\{x \mid a^T x = b\right\}$$

where $a \in \mathbf{R}^n$, $a \neq 0$, and $b \in \mathbf{R}$.

- A hyperplane is the solution set of a nontrivial linear equation among components of x. So, a hyperplane is affine.
- The vector a is called the **normal vector** of the hyperplane. Every point in the hyperplane has a constant inner product with the normal vector a.
- The constant $b \in \mathbf{R}$ determines the offset of the hyperplane from 0.

Hyperplane

- The hyperplane $\{x \mid a^T x = b\}$ can be rewritten as $\{x \mid a^T(x - x_0) = 0\}$, where x_0 is any point in the hyperplane.



- Further, we can write

$$\left\{x \mid a^T(x - x_0) = 0\right\} = x_0 + a^{\perp}$$

where $a^{\perp}$ denotes the orthogonal complement of a: $a^{\perp} = \{v \mid a^T v = 0\}$.

Halfspaces

A hyperplane divides $\mathbf{R}^n$ into two **halfspaces**.

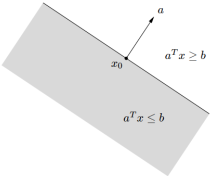
Halfspaces

A (closed) **halfspace** is a set of the form

$$\left\{x \mid a^T x \leq b\right\},$$

where $a \neq 0$.

- A halfspace is the solution set of one (nontrivial) linear inequality.
- Halfspaces are convex, but not affine.

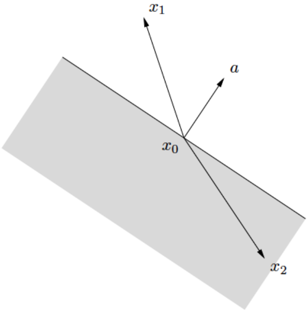


Halfspaces

- The halfspace $\{x \mid a^T x \leq b\}$ can also be rewritten as

$$\left\{x \mid a^T(x - x_0) \leq 0\right\},$$

where x_0 is any point on the associated hyperplane (i.e., $a^T x_0 = b$).



Halfspaces

- The boundary of the halfspace $\{x \mid a^T x \leq b\}$ is the hyperplane $\{x \mid a^T x = b\}$.
- The set
$$\left\{x \mid a^T x < b\right\}$$
is the interior of the halfspace $\{x \mid a^T x \leq b\}$. It is called an **open halfspace**.

Euclidean Balls

Euclidean ball

A **Euclidean ball** (or just **ball**) in $\mathbf{R}^n$ has the form

$$B(x_c, r) = \{x \mid ||x - x_c||_2 \leq r\} = \left\{x \mid (x - x_c)^T(x - x_c) \leq r^2\right\}$$

where $r > 0$ and $||\cdot||_2$ denotes the Euclidean norm.
The vector x_c is the **center** of the ball. The scalar r is its **radius**.

- $B(x_c, r)$ consists of all points within a distance r of the center x_c.
- The Euclidean ball can be rewritten as

$$B(x_c, r) = \{x_c + ru \mid ||u||_2 \leq 1\}.$$

Euclidean Balls

Property

A Euclidean ball is a convex set.

Proof:

Ellipsoid

Ellipsoid

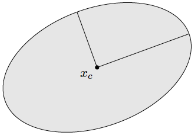
An ellipsoid has the form

$$\mathcal{E} = \left\{ x \mid (x - x_c)^T P^{-1} (x - x_c) \le 1 \right\},$$

where P is symmetric and positive definite: $P = P^T \succ 0$.
The vector $x_c \in \mathbf{R}^n$ is the center of the ellipsoid.

- The lengths of the semi-axes of $\mathcal{E}$ are given by $\sqrt{\lambda_i}$ where λ_i are the eigenvalues of P.
- A ball is an ellipsoid with $P = r^2 I$.
- An ellipsoid is convex.

Ellipsoid



- The ellipsoid $\mathcal{E} = \left\{x \mid (x - x_c)^T P^{-1}(x - x_c) \leq 1\right\}$ can be rewritten as

$$\mathcal{E} = \{x_c + Au \mid ||u||_2 \leq 1\}$$

  where A is square and nonsingular.

- W.l.o.g., we can assume A is symmetric and positive definite (by taking $A = P^{1/2}$).

Degenerate Ellipsoid

- If A is symmetric positive semidefinite but singular, then the set $\mathcal{E} = \{x_c + Au \mid ||u||_2 \leq 1\}$ is called a **degenerate ellipsoid**.
- Its affine dimension is rank A.
- Degenerate ellipsoids are also convex.

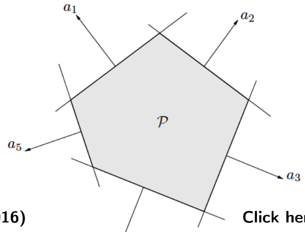
Polyhedra

Polyhedron

A polyhedron is defined as the solution set of a finite number of linear equations and linear inequalities:

$$\mathcal{P} = \left\{ x \mid a_j^T x \le b_j,\ j = 1, ..., m,\ c_j^T x = d_j, j = 1, ...p \right\}.$$

- A polyhedron is the intersection of a finite number of halfspaces and hyperplanes.



Polyhedra

- Polyhedra are convex sets.
- Affine sets (including subspaces, hyperplanes, and lines) are polyhedra.
- Rays, line segments, and hyperplanes are polyhedra.
- A bounded polyhedron is called a **polytope**.

Polyhedra

The polyhedron

$$\mathcal{P} = \left\{ x \mid a_j^T x \leq b_j,\ j = 1, \ldots, m,\ c_j^T x = d_j, j = 1, \ldots p \right\}$$

can be rewritten as

$$P = \{x \mid Ax \preceq b, Cx = d\}$$

where

$$A = \begin{bmatrix} a_1^T \\ \vdots \\ a_m^T \end{bmatrix} \text{ and } C = \begin{bmatrix} c_1^T \\ \vdots \\ c_p^T \end{bmatrix},$$

and the symbol $\preceq$ denotes **vector inequality** or **componentwise inequality** in $\mathbf{R}^m$: $u \preceq v$ means $u_i \leq v_i$ for $i = 1, \ldots, m$.

Polyhedra – An example

The set of nonnegative numbers

Let $\mathbf{R}_+$ denote the set of nonnegative numbers. Let $\mathbf{R}_{++}$ denote the set of positive numbers.

Nonnegative orthant

The **nonnegative orthant** in $\mathbf{R}^n$ is

$$\mathbf{R}^n_+ = \{x \in \mathbf{R}^n \mid x_i \geq 0,\ i = 1, ..., n\} = \{x \in \mathbf{R}^n \mid x \succeq 0\}.$$

- The nonnegative orthant is a polyhedron and a cone (called a **polyhedral cone**).

Simplexes – Another example of polyhedra

Affinely Independent

The $k+1$ points $v_0, v_1, \cdots, v_k \in \mathbf{R}^n$ are called **affinely independent** if $\{v_1 - v_0, \cdots, v_k - v_0\}$ is linearly independent.

Simplex

Suppose the $k+1$ points $v_0, v_1, \cdots, v_k \in \mathbf{R}^n$ are affinely independent. The **simplex** determined by these $k+1$ points is

$$C = \mathbf{conv}\ \{v_0, ..., v_k\} = \left\{\theta_0 v_0 + \cdots \theta_k v_k \mid \theta \succeq 0, \mathbf{1}^T\theta = 1\right\}$$

where $\mathbf{1}$ is the vector with all entries one.

- The above defined simplex is sometimes called a **k-dimensional simplex in $\mathbf{R}^n$**, since its affine dimension is k.

Examples of Simplexes

- A 1-dimensional simplex is a line segment.
- A 2-dimensional simplex is a triangle (including its interior).
- A 3-dimensional simplex is a tetrahedron.

Unit Simplex

The **unit simplex** in $\mathbf{R}^n$ is the n-dimensional simplex determined by the zero vector and the unit vectors: $\{0, e_1, \ldots, e_n\}$.

The unit simplex can be expressed as

$$\left\{x \mid x \succeq 0, \;\; \mathbf{1}^T x \leq 1\right\}.$$

Example of Simplexes – Probability Simplex

- The probability simplex in $\mathbf{R}^n$ is the $(n-1)$-**dimensional simplex** determined by the unit vectors $\{e_1, ..., e_n\}$.
- It can be expressed as

$$\left\{x \mid x \succeq 0, \;\; \mathbf{1}^T x = 1\right\}.$$

- Vectors in the probability simplex corresponds to probability distributions on a set with n elements.

Expressing A Simplex as A Polyhedron

- Consider the simplex

$$C = \mathbf{conv}\ \{v_0, \ldots, v_k\} = \left\{\theta_0 v_0 + \cdots \theta_k v_k \mid \theta \succeq 0, \mathbf{1}^T\theta = 1\right\}$$

- Let

$$B = \left[\begin{array}{ccc} v_1 - v_0 & \cdots & v_k - v_0 \end{array}\right] \in \mathbf{R}^{n\times k}$$

and $A = \left[\begin{array}{c} A_1 \\ A_2 \end{array}\right] \in \mathbf{R}^{n\times n}$ be a nonsingular matrix such that

$$AB = \left[\begin{array}{c} A_1 \\ A_2 \end{array}\right] B = \left[\begin{array}{c} I_k \\ 0_{(n-k)\times k} \end{array}\right].$$

- Then, we have $x \in C$ if and only if

$$A_2 x = A_2 v_0, \quad A_1 x \succeq A_1 v_0, \quad \mathbf{1}^T A_1 x \leq 1 + \mathbf{1}^T A_1 v_0.$$

(a form of a polyhedron)

Convex Hull Description of Polyhedra

- Consider the convex hull of the finite set $\{v_1, ..., v_k\}$,

$$
\begin{aligned}
& \mathbf{conv}\,\{v_1, ..., v_k\} \\
= \;& \{\theta_1 x_1 + \cdots + \theta_k x_k \mid \theta_i \geq 0,\ i = 1, ..., k,\ \theta_1 + \cdots + \theta_k = 1\} \\
= \;& \left\{\theta_1 x_1 + \cdots + \theta_k x_k \mid \theta \succeq 0,\ \mathbf{1}^T \theta = 1\right\}
\end{aligned}
$$

- It is a polyhedra and is bounded. (why?)
- How can we express $\mathbf{conv}\,\{v_1, ..., v_k\}$ in the form

$$\mathcal{P} = \left\{x \mid a_j^T x \leq b_j,\ j = 1, ..., m,\ c_j^T x = d_j, j = 1, ...p\right\}?$$

Convex Hull Description of Polyhedra

- Conversely, how do we express a polyhedron

$$\mathcal{P} = \left\{ x \;\middle|\; a_j^T x \leq b_j,\ j = 1, ..., m,\ c_j^T x = d_j,\ j = 1, ...p \right\}$$

  in the form of convex hull description $\mathbf{conv}\{v_1, ..., v_k\}$?
- Example: consider

$$C = \{x \mid |x_i| \leq 1,\ i = 1, ..., n\}$$

  (with $2n$ linear inequalities). Then we have

$$C = \mathbf{conv}\{v_1, ..., v_{2^n}\},$$

  where $v_1, ..., v_{2^n}$ are the 2^n vectors whose components are all 1 or -1.

Notations for Sets of Symmetric Matrices

- The notation $\mathbf{S}^n$ denotes the set of symmetric $n \times n$ matrices:

$$\mathbf{S}^n = \left\{ X \in \mathbf{R}^{n \times n} \mid X = X^T \right\}.$$

- $\mathbf{S}^n$ is a vector space with dimension $n(n+1)/2$.
- The notation $\mathbf{S}_+^n$ denotes the set of symmetric positive semidefinite matrices:

$$\mathbf{S}_+^n = \{ X \in \mathbf{S}^n \mid X \succeq 0 \}.$$

- The notation $\mathbf{S}_{++}^n$ denotes the set of symmetric positive definite matrices:

$$\mathbf{S}_{++}^n = \{ X \in \mathbf{S}^n \mid X \succ 0 \}.$$

Positive Semidefinite Cone

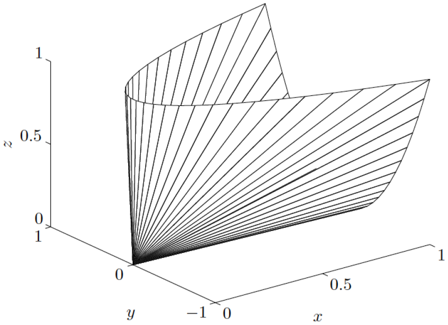
Convexity of Positive Semidefinite Cones

The set $\mathbf{S}_+^n$ is a convex cone:
if $\theta_1, \theta_2 \geq 0$ and $A, B \in \mathbf{S}_+^n$, then $\theta_1 A + \theta_2 B \in \mathbf{S}_+^n$.

Proof:

Positive Semidefinite Cone in $\mathbf{S}^2$

$$X = \begin{bmatrix} x & y \\ y & z \end{bmatrix} \in \mathbf{S}_+^2 \iff x \geq 0, \quad z \geq 0, \quad xz \geq y^2.$$



Norm balls and Norm Cones

Norm balls and Norm Cones

- Suppose $||\cdot||$ is a norm on $\mathbf{R}^n$.
- It can be shown that a **norm ball** of radius r and center x_c, given by $\{x \mid ||x - x_c|| \leq r\}$, is convex.
- The **norm cone** associated with the norm $||\cdot||$ is the set

$$C = \{(x,t) \mid ||x|| \leq t\} \subseteq \mathbf{R}^{n+1}.$$

Second-Order Cone

The **second-order cone** is the norm cone for the Euclidean norm, i.e.,

$$\begin{aligned} C &= \left\{ (x,t) \in \mathbf{R}^{n+1} \,\middle|\, \|x\|_2 \leq t \right\} \\ &= \left\{ \begin{bmatrix} x \\ t \end{bmatrix} \,\middle|\, \begin{bmatrix} x \\ t \end{bmatrix}^T \begin{bmatrix} I & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ t \end{bmatrix} \leq 0, \quad t \geq 0 \right\}. \end{aligned}$$

It is also known as the quadratic cone, the Lorentz cone, or ice-cream cone.