Lecture 1

Convex Sets

(Basic definitions and properties; Separation theorems; Characterizations)

1.1 Definition, examples, inner description, algebraic properties

1.1.1 A convex set

In the school geometry a figure is called convex if it contains, along with any pair of its points x, y, also the entire segment $[x, y]$ linking the points. This is exactly the definition of a convex set in the multidimensional case; all we need is to say what does it mean "the segment $[x, y]$ linking the points $x, y \in \mathbf{R}^n$". This is said by the following

Definition 1.1.1 [Convex set]

1) *Let x, y be two points in $\mathbf{R}^n$. The set*

$$[x, y] = \{z = \lambda x + (1 - \lambda)y \mid 0 \le \lambda \le 1\}$$

is called a segment with the endpoints x, y.

2) *A subset M of $\mathbf{R}^n$ is called convex, if it contains, along with any pair of its points x, y, also the entire segment $[x, y]$:*

$$x, y \in M,\ 0 \le \lambda \le 1 \Rightarrow \lambda x + (1 - \lambda)y \in M.$$

Note that by this definition an empty set is convex (by convention, or better to say, by the exact sense of the definition: for the empty set, you cannot present a counterexample to show that it is not convex).

The simplest examples of nonempty convex sets are singletons – points – and the entire space $\mathbf{R}^n$. A much more interesting example is as follows:

Example 1.1.1 *The solution set of an arbitrary (possibly, infinite) system*

$$a_\alpha^T x \le b_\alpha,\ \alpha \in \mathcal{A}$$

of linear inequalities with n unknowns x – the set

$$M = \{x \in \mathbf{R}^n \mid a_\alpha^T x \le b_\alpha,\ \alpha \in \mathcal{A}\}$$

is convex.

In particular, the solution set of a finite system

$$Ax \le b$$

of m inequalities with n variables (A is $m \times n$ matrix) is convex; a set of this latter type is called polyhedral.

Indeed, let x, y be two solutions to the system; we should prove that any point $z = \lambda x + (1 - \lambda y)$ with $\lambda \in [0, 1]$ also is a solution to the system. This is evident, since for every $\alpha \in \mathcal{A}$ we have

$$\begin{array}{rcl} a_\alpha^T x & \le & b_\alpha \\ a_\alpha^T y & \le & b_\alpha, \end{array}$$

whence, multiplying the inequalities by nonnegative reals λ and $1 - \lambda$ and taking sum of the results,

$$\lambda a_\alpha^T x + (1 - \lambda) a_\alpha^T y \le \lambda b_\alpha + (1 - \lambda) b_\alpha = b_\alpha,$$

and what is in the left hand side is exactly $a_\alpha^T z$. ■

Remark 1.1.1 *Note that any set given by Example 1.1.1 is not only convex, but also closed (why?)*

Note that any plane in $\mathbf{R}^n$ (in particular, any linear subspace) is the set of all solutions to some system of linear *equations*. Now, a system of linear equations is equivalent to a system of linear inequalities (you can equivalently represent a linear equality by a pair of opposite linear inequalities). It follows that a plane is a particular case of a polyhedral set and is therefore convex. Of course, we could obtain this conclusion directly: convexity of a set means that it is closed with respect to taking certain *restricted* set of linear combinations of its members – namely, the *pair* combinations with *nonnegative* coefficients of unit sum. One can easily show that any plane or an *affine set* is closed with respect to taking linear combinations *not* obligatory positive of its elements with unit sum (please, try to do it!).

1.1.2 Inner description of convex sets: Convex combinations and convex hull

Convex combinations

Definition 1.1.2 *A convex combination of vectors $y_1, ..., y_m$ is their linear combination*

$$y = \sum_{i=1}^m \lambda_i y_i$$

with nonnegative coefficients with unit sum:

$$\lambda_i \geq 0, \ \sum_{i=1}^{m} \lambda_i = 1.$$

We have the following very useful though simple statement:

Proposition 1.1.1 *A set M in $\mathbf{R}^n$ is convex if and only if it is closed with respect to taking all convex combinations of its elements, i.e., if and only if any convex combination of vectors from M again is a vector from M.*

Proof.

"if" part: assume that M contains all convex combinations of the elements of M. Then, with any two points $x, y \in M$ and any $\lambda \in [0,1]$, M contains also the vector $\lambda x + (1-\lambda)y$, since it is a convex combination of x and y; thus, M is convex.

"only if" part: assume that M is convex; we should prove that then M contains any convex combination

$$(*) \quad y = \sum_{i=1}^{m} \lambda_i y_i$$

of vectors $y_i \in M$. The proof is given by induction in m. The case of $m = 1$ is evident (since the only 1-term convex combinations are of the form is $1 \cdot y_1 = y_1 \in M$). Assume that we already know that any convex combination of $m-1$ vectors, $m \geq 2$, from M is again a vector from M, and let us prove that this statement remains valid also for all convex combinations of m vectors from M. Let (*) be such a combination. We can assume that $1 > \lambda_m$, since otherwise there is nothing to prove (indeed, if $\lambda_m = 1$, then the remaining λ_i's should be zero, since all λ's are nonnegative with the unit sum, and we have $y = y_m \in M$). Assuming $\lambda_m < 1$, we can write

$$y = (1-\lambda_m)[\sum_{i=1}^{m-1} \frac{\lambda_i}{1-\lambda_m} y_i] + \lambda_m y_m.$$

What is in the brackets, clearly is a convex combination of $m-1$ points from M and therefore, by the inductive hypothesis, this is a point, let it be called z, from M; we have

$$y = (1-\lambda_m)z + \lambda_m y_m$$

with z and $y_m \in M$, and $y \in M$ by definition of a convex set M. ■

Convex hull

Proposition 1.1.2 [Convexity of intersections] *Let $\{M_\alpha\}_\alpha$ be an arbitrary family of convex subsets of $\mathbf{R}^n$. Then the intersection*

$$M = \cap_\alpha M_\alpha$$

is convex.

Indeed, if the endpoints of a segment $[x, y]$ belong to M, then they belong also to every M_α; due to the convexity of M_α, the segment $[x, y]$ itself belongs to every M_α, and, consequently, to their intersection, i.e., to M. ■

An immediate consequence of this Proposition is as follows:

Corollary 1.1.1 [Convex hull]
Let M be a nonempty subset in $\mathbf{R}^n$. Then among all convex sets containing M (these sets exist, e.g., $\mathbf{R}^n$ itself) there exists the smallest one, namely, the intersection of all convex sets containing M.

This set is called the <u>*convex hull*</u> *of M [notation:* $\mathrm{Conv}\,(M)$].

Proposition 1.1.3 [Convex hull via convex combinations] *For a nonempty $M \subset \mathbf{R}^n$:*

$$\mathrm{Conv}\,(M) = \{\textit{the set of all convex combinations of vectors from } M\}.$$

Proof. According to Proposition 1.1.1, any convex set containing M (in particular, $\mathrm{Conv}\,(M)$) contains all convex combinations of vectors from M. What remains to prove is that $\mathrm{Conv}\,(M)$ does not contain anything else. To this end it suffices to prove that the set of all convex combinations of vectors from M, let this set be called M^*, itself is convex (given this fact and taking into account that $\mathrm{Conv}\,(M)$ is the smallest convex set containing M, we achieve our goal – the inclusion $\mathrm{Conv}\,(M) \subset M^*$). To prove that M^* is convex is the same as to prove that any convex combination $\nu x + (1-\nu)y$ of any two points $x = \sum_i \lambda_i x_i$, $y = \sum_i \mu_i x_i$ of M^* – two convex combinations of vectors $x_i \in M$ – is again a convex combination of vectors from M. This is evident:

$$\nu x + (1-\nu)y = \nu \sum_i \lambda_i x_i + (1-\nu) \sum_i \mu_i x_i = \sum_i \xi_i x_i, \quad \xi_i = \nu\lambda_i + (1-\nu)\mu_i,$$

and the coefficients ξ_i clearly are nonnegative with unit sum. ■

Proposition 1.1.3 provides us with an inner ("worker's") description of a convex set. In the Convex Analysis they also show an extremely useful outer ("artist's") description of *closed* convex sets: we will prove that all these sets are given by Example 1.1.1 – they are exactly the sets of all solutions to systems (possibly, infinite) of nonstrict linear inequalities[1)].

1.1.3 More examples of convex sets: polytope and cone

"Worker's" approach to generating convex sets provides us with two seemingly new examples of them: – a *polytope* and a *cone*.

[1)] that a set of solutions to any system of nonstrict linear inequalities is closed and convex – this we already know from Example 1.1.1 and Remark 1.1.1. The point is, of course, to prove that any closed convex set is a solution to a system of nonstrict linear inequalities

A polytope is, by definition, the convex hull of a finite nonempty set in $\mathbf{R}^n$, i.e., the set of the form

$$\mathrm{Conv}\,(\{u_1, ..., u_N\}) = \{\sum_{i=1}^{N} \lambda_i u_i \mid \lambda_i \geq 0, \sum_i \lambda_i = 1\}.$$

An important case of a polytope is a *simplex* - the convex hull of $n+1$ affinely independent points $v_1, ..., v_{n+1}$ from $\mathbf{R}^n$:

$$M = \mathrm{Conv}\,(\{v_1, ..., v_{n+1}\}) = \{\sum_{i=1}^{n+1} \lambda_i v_i \mid \lambda_i \geq 0, \sum_{i=1}^{n+1} \lambda_i = 1\};$$

the points $v_1, ..., v_{n+1}$ are called *vertices* of the simplex. [2)]

A cone. A nonempty subset M of $\mathbf{R}^n$ is called *conic*, if it contains, along with every point $x \in M$, the entire ray $\mathbf{R}x = \{tx \mid t \geq 0\}$ spanned by the point:

$$x \in M \Rightarrow tx \in M \ \forall t \geq 0.$$

A convex conic set is called *a cone*[3)].

Proposition 1.1.4 $^+$ *A nonempty subset M of $\mathbf{R}^n$ is a cone if and only if it possesses the following pair of properties:*

- *is conic: $x \in M, t \geq 0 \Rightarrow tx \in M$;*
- *contains sums of its elements: $x, y \in M \Rightarrow x + y \in M$.*

As an immediate consequence, we get that a cone is closed with respect to taking linear combinations *with nonnegative coefficients* of the elements, and vice versa – a nonempty set closed with respect to taking these combinations is a cone.

Example 1.1.2 $^+$ *The solution set of an arbitrary (possibly, infinite) system*

$$a_\alpha^T x \leq 0, \ \alpha \in \mathcal{A}$$

of homogeneous linear inequalities with n unknowns x – the set

$$K = \{x \mid a_\alpha^T x \leq 0 \ \forall \alpha \in \mathcal{A}\}$$

– is a cone.

In particular, the solution set to a homogeneous finite system of m homogeneous linear inequalities

$$Ax \leq 0$$

(A is $m \times n$ matrix) is a cone; a cone of this latter type is called polyhedral.

[2)] In fact a polytope (which is defined via the “worker's” approach) is nothing but a *bounded polyhedral set*, i.e., is a bounded set given by finitely many linear inequalities (this is “artists's” description of a polytope). The equivalence of these two inner and outer definitions of a polytope is one of the deepest facts of Convex Analysis.

[3)] sometimes people call *cones* what we call conic sets and call *convex cones* what we call cones

Note that the cones given by systems of linear homogeneous nonstrict inequalities necessarily are closed. We will see in the mean time that, vice versa, every closed convex cone is the solution set to such a system, so that Example 1.1.2 is the generic example of a closed convex cone.

Cones form a very important family of convex sets, and one can develop theory of cones absolutely similar (and in a sense, equivalent) to that one of all convex sets. E.g., introducing the notion of *conic combination* of vectors $x_1, ..., x_k$ as a linear combination of the vectors with nonnegative coefficients, you can easily prove the following statements completely similar to those for general convex sets, with conic combination playing the role of convex one:

- A set is a cone if and only if it is nonempty and is closed with respect to taking all conic combinations of its elements;
- Intersection of any family of cones is again a cone; in particular, for any nonempty set $M \subset \mathbf{R}^n$ there exists the smallest cone containing M – its conic hull $\mathrm{Cone}\,(M)$, and this conic hull is comprised of all conic combinations of vectors from M.

In particular, the conic hull of a nonempty finite set $M = \{u_1, ..., u_N\}$ of vectors in $\mathbf{R}^n$ is the cone

$$\mathrm{Cone}\,(M) = \{\sum_{i=1}^{N} \lambda_i u_i \mid \lambda_i \geq 0,\ i = 1, ..., N\}.$$

A fundamental fact (cf. the above story about polytopes) is that this is the generic (inner) description of a polyhedral cone – of a set given by (outer description) finitely many homogeneous linear inequalities.

1.1.4 Algebraic properties of convex sets

The following statement is an immediate consequence of the definition of a convex set.

Proposition 1.1.5 $^+$ *The following operations preserve convexity of sets:*

- *Arithmetic summation and multiplication by reals: if $M_1, ..., M_k$ are convex sets in $\mathbf{R}^n$ and $\lambda_1, ..., \lambda_k$ are arbitrary reals, then the set*

$$\lambda_1 M_1 + ... + \lambda_k M_k = \{\sum_{i=1}^{k} \lambda_i x_i \mid x_i \in M_i,\ i = 1, ..., k\}$$

  is convex.

- *Taking the image under affine mapping: if $M \subset \mathbf{R}^n$ is convex and $x \mapsto \mathcal{A}(x) \equiv Ax + b$ is an affine mapping from $\mathbf{R}^n$ into $\mathbf{R}^m$ (A is $m \times n$ matrix, b is m-dimensional vector), then the set*

$$\mathcal{A}(M) = \{y = \mathcal{A}(x) \equiv Ax + b \mid x \in M\}$$

  is a convex set in $\mathbf{R}^m$;

- *Taking the inverse image under affine mapping: if $M \subset \mathbf{R}^n$ is convex and $y \mapsto Ay + b$ is an affine mapping from $\mathbf{R}^m$ to $\mathbf{R}^n$ (A is $n \times m$ matrix, b is n-dimensional vector), then the set*
$$\mathcal{A}^{-1}(M) = \{y \in \mathbf{R}^m \mid \mathcal{A}(y) \in M\}$$
is a convex set in $\mathbf{R}^m$.

1.1.5 Topological properties of convex sets

Convex sets and closely related objects - convex functions - play the central role in Optimization. To play this role properly, the convexity alone is insufficient; we need convexity plus closedness. From the analysis we already know about the most basic topology-related notions – convergence of sequences of vectors, closed and open sets in $\mathbf{R}^n$. Here are three more notions we will use:

The closure. It is clear from definition of a closed set that the intersection of any family of closed sets in $\mathbf{R}^n$ is also closed. From this fact it, as always, follows that for any subset M of $\mathbf{R}^n$ there exists the smallest closed set containing M; this set is called the *closure* of M and is denoted $\mathrm{cl}\, M$. In Analysis they prove the following inner description of the closure of a set in a metric space (and, in particular, in $\mathbf{R}^n$):

The closure of a set $M \subset \mathbf{R}^n$ is exactly the set comprised of the limits of all converging sequences of elements of M.

With this fact in mind, it is easy to prove that, e.g., the closure of the open Euclidean ball
$$\{x \mid |x - a| < r\} \quad [r > 0]$$
is the closed ball $\{x \mid |x - a| \le r\}$. Another useful application example is the closure of a set
$$M = \{x \mid a_\alpha^T x < b_\alpha,\ \alpha \in \mathcal{A}\}$$
given by strict linear inequalities: *if such a set is nonempty*, then its closure is given by the nonstrict versions of the same inequalities:
$$\mathrm{cl}\, M = \{x \mid a_\alpha^T x \le b_\alpha,\ \alpha \in \mathcal{A}\}.$$

Nonemptiness of M in the latter example is essential: the set M given by two strict inequalities
$$x < 0, \quad -x < 0$$
in $\mathbf{R}$ clearly is empty, so that its closure also is empty; in contrast to this, applying formally the above rule, we would get wrong answer
$$\mathrm{cl}\, M = \{x \mid x \le 0, x \ge 0\} = \{0\}.$$

The interior. Let $M \subset \mathbf{R}^n$. We say that a point $x \in M$ is an *interior* for M, if some neighborhood of the point is contained in M, i.e., there exists centered at x ball of positive radius which belongs to M:

$$\exists r > 0 \quad B_r(x) \equiv \{y \mid |y - x| \leq r\} \subset M.$$

The set of all interior points of M is called the *interior* of M [notation: $\mathrm{int}\, M$].

E.g.,

- The interior of an open set is the set itself;
- The interior of the closed ball $\{x \mid |x - a| \leq r\}$ is the open ball $\{x \mid |x - a| < r\}$ (why?)
- The interior of a polyhedral set $\{x \mid Ax \leq b\}$ with matrix A not containing zero rows is the set $\{x \mid Ax < b\}$ (why?)

The latter statement is not, generally speaking, valid for sets of solutions of infinite systems of linear inequalities. E.g., the system of inequalities

$$x \leq \frac{1}{n},\ n = 1, 2, ...$$

in $\mathbf{R}$ has, as a solution set, the nonpositive ray $\mathbf{R}_- = \{x \leq 0\}$; the interior of this ray is the negative ray $\{x < 0\}$. At the same time, strict versions of our inequalities

$$x < \frac{1}{n},\ n = 1, 2, ...$$

define the same nonpositive ray, not the negative one.

It is also easily seen (this fact is valid for arbitrary metric spaces, not for $\mathbf{R}^n$ only), that

- the interior of an arbitrary set is open

The interior of a set is, of course, contained in the set, which, in turn, is contained in its closure:

$$\mathrm{int}\, M \subset M \subset \mathrm{cl}\, M. \tag{1.1.1}$$

The complement of the interior in the closure – the set

$$\partial M = \mathrm{cl}\, M \backslash \mathrm{int}\, M$$

– is called the *boundary* of M, and the points of the boundary are called *boundary points* of M (Warning: these points not necessarily belong to M, since M can be less than $\mathrm{cl}\, M$; in fact, all boundary points belong to M if and only if $M = \mathrm{cl}\, M$, i.e., if and only if M is closed).

The boundary of a set clearly is closed (as the intersection of two closed sets $\mathrm{cl}\, M$ and $\mathbf{R}^n \backslash \mathrm{int}\, M$; the latter set is closed as a complement to an open set). From the definition of the boundary,

$$M \subset \mathrm{int}\, M \cup \partial M \quad [= \mathrm{cl}\, M],$$

so that a point from M is either an interior, or a boundary point of M.

The relative interior. Many of the constructions to be considered possess nice properties in the interior of the set the construction is related to and may loose these nice properties at the boundary points of the set; this is why in many cases we are especially interested in interior points of sets and want the set of these points to be "enough massive". What to do if it is not the case – e.g., there are no interior points at all (look at a segment in the plane)? It turns out that in these cases we can use a good surrogate of the normal interior – the *relative interior*.

It is obvious (why?) that for every nonempty subset Y of $\mathbf{R}^n$ there exists the smallest affine set containing Y – the intersection of all affine sets containing Y. This smallest affine set containing Y is called the *affine hull* of Y (notation: $\mathrm{Aff}\,(Y)$). Intuitively, $\mathrm{Aff}\,(Y)$ is the "proper" space to look at Y: we simply cut from $\mathbf{R}^n$ "useless" dimensions (such that the projection of Y on these dimensions is a singleton).

Definition 1.1.3 [Relative interior] *Let $M \subset \mathbf{R}^n$. We say that a point $x \in M$ is relative interior for M, if M contains the intersection of a small enough ball centered at x with* $\mathrm{Aff}\,(M)$*:*

$$\exists r > 0 \quad B_r(x) \cap \mathrm{Aff}\,(M) \equiv \{y \mid y \in \mathrm{Aff}\,(M),\ |y-x| \le r\} \subset M.$$

The set of all relative interior points of M is called its relative interior [notation: $\mathrm{ri}\,M$*].*

E.g. the relative interior of a singleton is the singleton itself (since a point in the 0-dimensional space is the same as a ball of any positive radius); similarly, the relative interior of an affine set is the set itself. The interior of a segment $[x,y]$ ($x \neq y$) in $\mathbf{R}^n$ is empty whenever $n > 1$; in contrast to this, the relative interior is nonempty independently of n and is the interval (x,y) – the segment with deleted endpoints. Geometrically speaking, the relative interior is the interior we get when regard M as a subset of its affine hull (the latter, geometrically, is nothing but $\mathbf{R}^k$, k being the affine dimension of $\mathrm{Aff}\,(M)$).

We can play with the notion of the relative interior in basically the same way as with the one of interior, namely:

- since $\mathrm{Aff}\,(M)$, as any affine set, is closed and contains M, it contains also the smallest of closed sets containing M, i.e., $\mathrm{cl}\,M$. Therefore we have the following analogies of inclusions (1.1.1):
$$\mathrm{ri}\,M \subset M \subset \mathrm{cl}\,M \quad [\subset \mathrm{Aff}\,(M)]; \tag{1.1.2}$$

- we can define the *relative boundary* $\partial_{\mathrm{ri}}\,M = \mathrm{cl}\,M \backslash \mathrm{ri}\,M$ which is a closed set contained in $\mathrm{Aff}\,(M)$, and, as for the "actual" interior and boundary, we have
$$\mathrm{ri}\,M \subset M \subset \mathrm{cl}\,M = \mathrm{ri}\,M + \partial_{\mathrm{ri}}\,M.$$

Of course, if $\mathrm{Aff}\,(M) = \mathbf{R}^n$, then the relative interior becomes the usual interior, and similarly for boundary; this for sure is the case when $\mathrm{int}\,M \neq \emptyset$ (since then M contains a ball B, and therefore the affine hull of M is the entire $\mathbf{R}^n$, which is the affine hull of B).

Nice topological properties of a convex set

An arbitrary set M in $\mathbf{R}^n$ may possess very pathological topology: both inclusions in the chain

$$\mathrm{ri}\, M \subset M \subset \mathrm{cl}\, M$$

can be very "untight". E.g., let M be the set of rational numbers in the segment $[0,1] \subset \mathbf{R}$. Then $\mathrm{ri}\, M = \mathrm{int}\, M = \emptyset$ – since any neighbourhood of every rational real contains irrational reals – while $\mathrm{cl}\, M = [0,1]$. Thus, $\mathrm{ri}\, M$ is "incomparably smaller" than M, $\mathrm{cl}\, M$ is "incomparable larger", and M is contained in its relative boundary (by the way, what is this relative boundary?).

The following proposition demonstrates that the topology of a *convex* set M is much better than it might be for an arbitrary set.

Theorem 1.1.1 *Let M be a convex set in $\mathbf{R}^n$. Then*

(i) $^{+}$ *The interior* $\mathrm{int}\, M$*, the closure* $\mathrm{cl}\, M$ *and the relative interior* $\mathrm{ri}\, M$ *are convex;*

(ii) *If M is nonempty, then the relative interior* $\mathrm{ri}\, M$ *of M is nonempty*

(iii) *The closure of M is the same as the closure of its relative interior:*

$$\mathrm{cl}\, M = \mathrm{cl}\, \mathrm{ri}\, M$$

(in particular, every point of $\mathrm{cl}\, M$ *is the limit of a sequence of points from* $\mathrm{ri}\, M$*)*

(iv) *The relative interior remains unchanged when we replace M with its closure:*

$$\mathrm{ri}\, M = \mathrm{ri}\ \mathrm{cl}\, M.$$

Proof.

(ii): Let M be a nonempty convex set, and let us prove that $\mathrm{ri}\, M \neq \emptyset$. It suffices to consider the case when $\mathrm{Aff}\,(M)$ is the entire space $\mathbf{R}^n$. Indeed, by translation of M we always may assume that $\mathrm{Aff}\,(M)$ contains 0, i.e., is a linear subspace. Thus, in the rest of the proof of (ii) we assume that $\mathrm{Aff}\,(M) = \mathbf{R}^n$, and what we should prove is that the interior of M (which in the case in question is the same as relative interior) is nonempty.

Now, $\mathrm{Aff}\,(M) = \mathbf{R}^n$ possesses an affine basis $a_0, ..., a_n$ comprised of vectors from M. Since $a_0, ..., a_n$ belong to M *and M is convex*, the entire convex hull of the vectors – the simplex Δ with the vertices $a_0, ..., a_n$ – is contained in M. Consequently, an interior point of the simplex for sure is an interior point of M; thus, in order to prove that $\mathrm{int}\, M \neq \emptyset$, it suffices to prove that the interior of Δ is nonempty, as it should be according to geometric intuition.

The proof of the latter fact is as follows: since $a_0, ..., a_n$ is, by construction, an affine basis of $\mathbf{R}^n$, every point $x \in \mathbf{R}^n$ is affine combination of the points of the basis. The coefficients $\lambda_i = \lambda_i(x)$ of the combination – the barycentric coordinates of x with respect to the basis – are solutions to the following system of equations:

$$\sum_{i=0}^{n} \lambda_i a_i = x; \quad \sum_{i=0}^{n} \lambda_i = 1,$$

or, in the entrywise form,

$$\begin{array}{ccccccccc} a_{01}\lambda_0 & + & a_{11}\lambda_1 & + & ... & + & a_{n1}\lambda_n & = & x_1 \\ a_{02}\lambda_0 & + & a_{12}\lambda_1 & + & ... & + & a_{n2}\lambda_n & = & x_2 \\ ..... & & ..... & & ..... & & ..... & = & ... \\ a_{0n}\lambda_0 & + & a_{1n}\lambda_1 & + & ... & + & a_{nn}\lambda_n & = & x_n \\ \lambda_0 & + & \lambda_2 & + & ... & + & \lambda_n & = & 1 \end{array}; \tag{1.1.3}$$

(a_{pq} is q-th entry of vector a_p). This is a linear system of equations with $n+1$ equation and $n+1$ unknown. *The corresponding homogeneous system has only trivial solution* – indeed, a nontrivial solution to the homogeneous system would give us an equal to zero nontrivial linear combination of a_i with zero sum of coefficients, while from affine independence of $a_0, ..., a_n$ (they are affine independent since they form an affine basis) we know that no such a combination exists. It follows that the matrix of the system, let it be called A, is nonsingular, so that the solution $\lambda(x)$ linearly (consequently, continuously) depends on the right hand side data, i.e., on x.

Now we are done: let us take any $x = x^0$ with $\lambda_i(x^0) > 0$, e.g., $x^0 = (n+1)^{-1}\sum_{i=0}^{n} a_i$. Due to the continuity of $\lambda_i(\cdot)$'s, there is a neighborhood of x^0 – a centered at x^0 ball $B_r(x^0)$ of positive radius r - where the functions λ_i still are positive:

$$x \in B_r(x^0) \Rightarrow \lambda_i(x) \geq 0,\ i = 0, ..., n.$$

The latter relation means that every $x \in B_r(x^0)$ is an affine combination of a_i with positive coefficients, i.e., is a convex combination of the vectors, and therefore x belongs to Δ. Thus, Δ contains a neighborhood of x^0, so that x^0 is an interior point of Δ. ∎

(iii): We should prove that the closure of $\operatorname{ri} M$ is exactly the same that the closure of M. In fact we will prove even more:

Lemma 1.1.1 *Let $x \in \operatorname{ri} M$ and $y \in \operatorname{cl} M$. Then all points from the half-segment $[x, y)$,*

$$[x, y) = \{z = (1-\lambda)x + \lambda y \mid 0 \leq \lambda < 1\}$$

belong to the relative interior of M.

Proof of the Lemma. Let $\operatorname{Aff}(M) = a + L$, L being linear subspace; then

$$M \subset \operatorname{Aff}(M) = x + L.$$

Let B be the unit ball in L:

$$B = \{h \in L \mid |h| \leq 1\}.$$

Since $x \in \operatorname{ri} M$, there exists positive radius r such that

$$x + rB \subset M. \tag{1.1.4}$$

Since $y \in \operatorname{cl} M$, we have $y \in \operatorname{Aff}(M)$ (see (1.1.2)). Besides, for any $\epsilon > 0$ there exists $y' \in M$ such that $|y' - y| \leq \epsilon$; since both y' and y belong to $\operatorname{Aff}(M)$, the vector $y - y'$ belongs to L and consequently to ϵB. Thus,

$$(\forall \epsilon > 0): \ y \in M + \epsilon B. \tag{1.1.5}$$

Now let $z \in [x, y)$, so that

$$z = (1-\lambda)x + \lambda y$$

with some $\lambda \in (0,1)$; we should prove that z is relative interior for M, i.e., that there exists $r' > 0$ such that

$$z + r'B \subset M. \tag{1.1.6}$$

For any $\epsilon > 0$ we have, in view of (1.1.5),

$$z+\epsilon B \equiv (1-\lambda)x+\lambda y+\epsilon B \subset (1-\lambda)x+\lambda[M+\epsilon B]+\epsilon B = (1-\lambda)[x+\frac{\lambda\epsilon}{1-\lambda}B+\frac{\epsilon}{1-\lambda}B]+\lambda M \tag{1.1.7}$$

for all $\epsilon > 0$. Now, for the centered at zero Euclidean ball B and nonnegative t', t'' one has

$$t'B + t''B \subset (t'+t'')B$$

(in fact this is equality rather than inclusion, but it does not matter). Indeed, if $u \in t'B$, i.e., $|u| \leq t'$, and $v \in t''B$, i.e., $|v| \leq t''$, then, by the triangle inequality, $|u+v| \leq t'+t''$, i.e., $u+v \in (t'+t'')B$. Given this inclusion, we get from (1.1.7)

$$z+\epsilon B \subset (1-\lambda)\left[x + \frac{(1+\lambda)\epsilon}{1-\lambda}B\right] + \lambda M$$

for all $\epsilon > 0$. Setting ϵ small enough, we can make the coefficient at B in the right hand side less than r (see (1.1.4)); for this choice of ϵ, we, in view of (1.1.4), have

$$x + \frac{(1+\lambda)\epsilon}{1-\lambda}B \subset M,$$

and we come to

$$z + \epsilon B \subset (1-\lambda)M + \lambda M = M$$

(the concluding inequality holds true due to the convexity of M). Thus, $z \in \operatorname{ri} M$. ∎

Lemma immediately implies (iii). Indeed, $\operatorname{cl}\operatorname{ri} M$ clearly can be only smaller than $\operatorname{cl} M$: $\operatorname{cl}\operatorname{ri} M \subset \operatorname{cl} M$, so that all we need is to prove the inverse inclusion $\operatorname{cl} M \subset \operatorname{cl}\operatorname{ri} M$, i.e., to prove that every point $y \in \operatorname{cl} M$ is a limit of a sequence of points $\operatorname{ri} M$. This is immediate: of course, we can assume M nonempty (otherwise all sets in question are empty and therefore coincide with each other), so that by (ii) there exists a point $x \in \operatorname{ri} M$. According to Lemma, the half-segment $[x, y)$ belongs to $\operatorname{ri} M$, and y clearly is the limit of a sequence of points of this half-segment, e.g., the sequence $x_i = \frac{1}{n}x + (1-\frac{1}{n})y$. ∎

A useful byproduct of Lemma 1.1.1 is as follows:

Corollary 1.1.2 $^{+}$ *Let M be a convex set. Then any convex combination*

$$\sum_i \lambda_i x_i$$

of points $x_i \in \operatorname{cl} M$ where at least one term with positive coefficient corresponds to $x_i \in \operatorname{ri} M$ is in fact a point from $\operatorname{ri} M$.

(iv): The statement is evidently true when M is empty, so assume that M is nonempty. The inclusion $\mathrm{ri}\, M \subset \mathrm{ri}\,\mathrm{cl}\, M$ is evident, and all we need is to prove the inverse inclusion. Thus, let $z \in \mathrm{ri}\,\mathrm{cl}\, M$, and let us prove that $z \in \mathrm{ri}\, M$. Let $x \in \mathrm{ri}\, M$ (we already know that the latter set is nonempty). Consider the segment $[x, z]$; since z is in the relative interior of $\mathrm{cl}\, M$, we can extend a little bit this segment through the point z, not leaving $\mathrm{cl}\, M$, i.e., there exists $y \in \mathrm{cl}\, M$ such that $z \in [x, y)$. We are done, since by Lemma 1.1.1 from $z \in [x, y)$, with $x \in \mathrm{ri}\, M$, $y \in \mathrm{cl}\, M$, it follows that $z \in \mathrm{ri}\, M$. ■

> We see from the proof of Theorem 1.1.1 that to get a closure of a (nonempty) convex set, it suffices to subject it to the "radial" closure, i.e., to take a point $x \in \mathrm{ri}\, M$, take all rays in $\mathrm{Aff}\,(M)$ starting at x and look at the intersection of such a ray l with M; such an intersection will be a convex set on the line which contains a one-sided neighbourhood of x, i.e., is either a segment $[x, y_l]$, or the entire ray l, or a half-interval $[x, y_l)$. In the first two cases we should not do anything; in the third we should add y to M. After all rays are looked through and all "missed" endpoints y_l are added to M, we get the closure of M. To understand what is the role of convexity here, look at the *nonconvex* set of rational numbers from $[0, 1]$; the interior ($\equiv$ relative interior) of this "highly percolated" set is empty, the closure is $[0, 1]$, and there is no way to restore the closure in terms of the interior.

1.2 The Separation Theorem

In this section we answer the following question: assume we are given two convex sets S and T in $\mathbf{R}^n$. When can we separate them by a hyperplane, i.e., to find a nonzero linear form which at any point of S is less than or equal to its value at any point of T?

Let us start with definitions. A hyperplane M in $\mathbf{R}^n$ (an affine set of dimension $n-1$) is nothing but a level set of a nontrivial linear form:

$$\exists a \in \mathbf{R}^n, b \in \mathbf{R}, a \neq 0 : \quad M = \{x \in \mathbf{R}^n \mid a^T x = b\}.$$

Now we define the basic notion of *proper separation* of sets:

Definition 1.2.1 [Proper separation] *We say that a hyperplane*

$$M = \{x \in \mathbf{R}^n \mid a^T x = b\} \quad [a \neq 0]$$

properly separates (nonempty) convex sets S and T, if

(i) the sets belong to the opposite closed half-spaces into which M splits $\mathbf{R}^n$, and

(ii) at least one of the sets is not contained in M itself.

We say that S and T can be properly separated, if there exists a hyperplane which properly separates S and T, i.e., if there exists $a \in \mathbf{R}^n$ such that

$$\sup_{x \in S} a^T x \leq \inf_{y \in T} a^T y$$

and

$$\inf_{x\in S} a^T x < \sup_{y\in T} a^T y.$$

Sometimes we are interested also in a stronger notion of separation:

Definition 1.2.2 [Strong separation] *We say that nonempty sets S and T in $\mathbf{R}^n$ can be strongly separated, if there exist two distinct parallel hyperplanes which separate S and T, i.e., if there exists $a \in \mathbf{R}^n$ such that*

$$\sup_{x\in S} a^T x < \inf_{y\in T} a^T y.$$

It is clear that

Strong separation ⇒ Proper separation

We can immediately point out examples of sets which can be separated properly and cannot be separated strongly, e.g., the sets $S = \{x \in \mathbf{R}^2 \mid x_2 \geq x_1^2\}$ and $T = (x_1, 0)$.

Now we come to the important question:

when two given nonempty convex sets S and T in $\mathbf{R}^n$ can be separated [properly, strongly]?

The most important question is that on the possibility of proper separation. The answer is as follows:

Theorem 1.2.1 [Separation Theorem] *Two nonempty convex sets S and T in $\mathbf{R}^n$ can be properly separated if and only if their relative interiors do not intersect:*

$$\mathrm{ri}\, S \cap \mathrm{ri}\, T = \emptyset.$$

1.2.1 Necessity

The necessity of the indicated property (the "only if" part of the Separation Theorem) is more or less evident. Indeed, assume that the sets can be properly separated, so that for certain nonzero $a \in \mathbf{R}^n$ we have

$$\sup_{x\in S} a^T x \leq \inf_{y\in T} a^T y; \quad \inf_{x\in S} a^T x < \sup_{y\in T} a^T y. \tag{1.2.8}$$

We should lead to a contradiction the assumption that there exists a point $\bar{x} \in (\mathrm{ri}\, S \cap \mathrm{ri}\, T)$. Assume that it is the case; then from the first inequality in (1.2.8) it is clear that $\bar{x}$ maximizes the linear function $f(x) = a^T x$ on S and at the same time, it minimizes this form on T. Now, we have the following simple and important

> **Lemma 1.2.1** *A linear function $f(x) = a^T x$ can attain its maximum/minimum over a convex set M at a point $\bar{x} \in \mathrm{ri}\, M$ if and only if the function is constant on M.*

Proof. "if" part is evident. To prove the "only if" part, let $\bar{x} \in \operatorname{ri} M$ be, say, a minimizer of f over M and y be an arbitrary point of M; we should prove that $f(\bar{x}) = f(y)$. There is nothing to prove if $y = \bar{x}$, so let us assume that $y \neq \bar{x}$. Since $\bar{x} \in \operatorname{ri} M$, the segment $[y, \bar{x}]$, which is contained in M, can be extended a little bit through the point $\bar{x}$, not leaving M (since $\bar{x} \in \operatorname{ri} M$), so that there exists $z \in M$ such that $\bar{x} \in (y, z)$, i.e., $\bar{x} = (1-\lambda)y + \lambda z$ with certain $\lambda \in (0,1)$. Since f is linear, we have

$$f(\bar{x}) = (1-\lambda)f(y) + \lambda f(z);$$

since $f(\bar{x}) \leq \min\{f(y), f(z)\}$ and $0 < \lambda < 1$, this relation can be satisfied only when $f(\bar{x}) = f(y) = f(z)$. ∎

1.2.2 Sufficiency

Separation of a convex set and a point outside of the set. Consider the following

Proposition 1.2.1 $^{+}$ *Let M be a nonempty and closed convex set in $\mathbf{R}^n$, and let x be a point outside M ($x \notin M$). Consider the optimization program*

$$\min\{|x-y| \mid y \in M\}.$$

The program is solvable and has a unique solution y^*, and the linear form $a^T h$, $a = x - y^*$, strongly separates x and M:

$$\sup_{y \in M} a^T y = a^T y^* = a^T x - |a|^2.$$

The point y^* has a special name – it is called *projection of x on M*. Let us consider an example of application of the above proposition:

Corollary 1.2.1 *Let M be a convex set. Consider a function*

$$\psi_M(x) = \sup\{y^T x | \, y \in M\}.$$

Function $\psi_M(x)$ is called the support function of the set M.

Now, let M_1 and M_2 be two closed convex sets.

If for any $x \in \operatorname{Dom} \psi_{M_2}$ we have $\psi_{M_1}(x) \leq \psi_{M_2}(x)$ then $M_1 \subset M_2$.

Let $\operatorname{Dom} \psi_{M_1} = \operatorname{Dom} \psi_{M_2}$ and for any $x \in \operatorname{Dom} \psi_{M_1}$ we have $\psi_{M_1}(x) = \psi_{M_2}(x)$. Then $M_1 \equiv M_2$

Proof: assume, on the contrary, that there is $x_0 \in M_1$ and $x_0 \notin M_2$. Then in view of Proposition 1.2.1, there is $a \neq 0$ such that $a^T x_0 > a^T x$ for any $x \in M_2$. Hence, $a \in \operatorname{Dom} \psi_{M_2}$ and $\psi_{M_1}(a) > \psi_{M_2}(a)$, which is a contradiction.∎

Separating a convex set and a non-interior point Now we need to extend the above statement to the case when x^* is a point outside the relative interior of a convex (not necessarily closed) set. Here we, in general, loose strong separation, but still have the proper one:

Proposition 1.2.2 [Separation of a point and a set]
Let M be a convex set in $\mathbf{R}^n$, and let $x \notin \mathrm{ri}\, M$. Then x and M can be properly separated.

Proof. Let $\bar{M} = \mathrm{cl}\, M$; clearly, this is a closed convex set with the same relative interior as that of M. In particular, $x \notin \mathrm{ri}\, \bar{M}$. If also $x \notin \bar{M}$, then x and $\bar{M}$ (and also x and $M \subset \bar{M}$) can be strongly (and therefore properly) separated by the previous proposition. Thus, it remains to consider the case when x is a point of relative boundary of $\bar{M}$: $x \in \partial_{\mathrm{ri}}\, \bar{M} \equiv \bar{M} \backslash \mathrm{ri}\, \bar{M}$. Without loss of generality we may assume that $x = 0$ (simply by translation), so, let L be the linear span of M. Since $x = 0$ is not in the relative interior of $\bar{M}$, there is a sequence of points x_i not belonging to $\bar{M}$ and converging to 0. By the previous proposition, for each i, x_i can be strongly separated from $\bar{M}$: there exists a linear form $a_i^T x$ with the property

$$a_i^T x_i < \inf_{y \in \bar{M}} a_i^T y. \tag{1.2.9}$$

We can choose a_i to belong to L and normalize it - divide by its Euclidean norm; for the resulting unit vector $a_i \in L$ we still have (1.2.9). Since these vectors belong to the unit ball. i.e., to a compact set, we can extract from this sequence a converging subsequence; after renotation, we may assume that a_i themselves converge to certain vector $a \in L$. This vector is unit, since all a_i's are unit. Now, we have for every fixed $y \in \bar{M}$:

$$a_i^T x_i < a_i^T y.$$

As $i \to \infty$, we have $a_i \to a$, $x_i \to x = 0$, and passing to limit in our inequality, we get

$$a^T x = 0 \leq a^T y, \;\; \forall y \in \bar{M}.$$

Thus, we have established the main part of the required statement: the linear form $f(u) = a^T u$ separates $\{0\}$ and $\bar{M}$ (and thus $\{0\}$ and $M \subset \bar{M}$). It remains to verify that this separation is proper, which in our case simply means that M is not contained in the hyperplane $\{u \mid a^T u = 0\}$. But this is evident: if we assume that a is orthogonal to M, then due to the non-emptiness of $\mathrm{ri}\, M$, a is orthogonal to any vector from a small ball in L what implies that a is orthogonal to entire L, what contradicts our assumption that $a \in L$ and $|a| = 1$. ∎

The general case

Now we are ready to consider the general case. Thus, we are given two nonempty convex sets S and T with the non-intersecting relative interiors, and we should prove that the sets can be properly separated. Let $S' = \mathrm{ri}\, S$, $T' = \mathrm{ri}\, T$; these are two nonempty convex sets and they do not intersect. Let $M = T' - S'$ (here, by definition, $M = \{z = x - y \mid x \in T',\ y \in S'\}$); this again is a nonempty convex (as a sum of two nonempty convex sets) set which does not contain 0 (the latter - since S' and T' do not intersect). By the above proposition $\{0\}$ and M can be properly separated: there exists a such that

$$0 = a^T 0 \le \inf_{z\in M} a^T z \equiv \inf_{y\in T', x\in S'} a^T(y-x) = [\inf_{y\in T'} a^T y] - [\sup_{x\in S'} a^T x],$$

and

$$0 < \sup_{z\in M} a^T z \equiv \sup_{y\in T', x\in S'} a^T(y-x) = [\sup_{y\in T'} a^T y] - [\inf_{x\in S'} a^T x].$$

The sup and inf here are taken over the relative interiors T', S' of the sets T and S; however, by (iii) of Theorem 1.1.1, $\mathrm{cl}\, T' = \mathrm{cl}\, T \supset T$, so that T' is dense in T, and similarly for S' and S. Therefore sup, inf over T' and S' of a linear form are the same as similar operations over S, T, and we come to

$$\sup_{x\in S} a^T x \le \sup_{y\in T} a^T y, \quad \inf_{x\in S} a^T x < \sup_{y\in T} a^T y,$$

which means proper separation of S and T ■

1.2.3 Strong separation

We know from the Separation Theorem what are simple necessary and sufficient conditions for proper separation of two convex sets - their relative interiors should be disjoint. There is also a simple necessary and sufficient condition for two sets to be strongly separated:

Proposition 1.2.3 *Two nonempty convex sets S and T in $\mathbf{R}^n$ can be strongly separated if and only if these sets are "at positive distance":*

$$\rho(S,T) = \inf_{x\in S, y\in T} |x-y| > 0.$$

This is, in particular, the case when one of the sets is compact, the other one is closed and the sets do not intersect.

Proof . The necessity - the "only if" part - is evident: if S and T can be strongly separated, i.e., for certain a one has

$$\alpha \equiv \sup_{x\in S} a^T x < \beta \equiv \inf_{y\in T} a^T y,$$

then for every pair (x,y) with $x\in S$ and $y\in T$ one has

$$|x-y| \ge \frac{\beta-\alpha}{|a|}$$

(since otherwise we would get from Cauchy's inequality

$$a^T y - a^T x = a^T(y-x) \le |a|\,|y-x| < \beta - \alpha,$$

which is impossible).

To prove the sufficiency - the "if" part - consider the set $\Delta = S - T$. This is a convex set (why?) which clearly does not contain vectors of the length less than $\rho(S,T) > 0$; consequently, it does not intersect the ball B of some positive radius r centered at the origin. Consequently, by the Separation Theorem M can be properly separated from B: there exists a such that

$$\inf_{z\in B} a^T z \geq \sup_{x\in S, y\in T} f^T(x-y) \ \& \ \sup_{z\in B} a^T z > \inf_{x\in S, y\in T} f^T(x-y). \tag{1.2.10}$$

From the second of these inequalities it follows that $a \neq 0$ (as it always is the case with proper separation); therefore $\inf_{z\in B} a^T z < 0$, so that the first inequality in (1.2.10) means that a strongly separates S and T.

The "in particular" part of the statement is a simple exercise from Analysis: two closed nonempty and non-intersecting subsets of $\mathbf{R}^n$ with one of them being compact are at positive distance from each other. ∎

With the Separation Theorem in our hands, we can get much more understanding of the geometry of convex sets.

1.2.4 Outer description of a closed convex set. Supporting planes

We can now prove the "outer" characterization of a closed convex set announced in the beginning of this section:

Theorem 1.2.2 *Any closed convex set M in $\mathbf{R}^n$ is the solution set of an (infinite) system of nonstrict linear inequalities.*

Geometrically: every closed convex set $M \subset \mathbf{R}^n$ which differs from the entire $\mathbf{R}^n$ is the intersection of closed half-spaces – namely, all closed half-spaces which contain M.

Proof is readily given by the Separation Theorem. Indeed, if M is empty, there is nothing to prove – an empty set is the intersection of two properly chosen closed half-spaces. If M is the entire space, there also is nothing to prove – according to our convention, this is the solution set to the empty system of linear inequalities. Now assume that M is convex, closed, nonempty and differs from the entire space. Let $x \notin M$; then x is at the positive distance from M since M is closed and therefore there exists a hyperplane strongly separating x and M (Proposition 1.2.3):

$$\forall x \notin M \ \exists a_x : a_x^T x > \alpha_x \equiv \sup_{y\in M} a_x^T y.$$

For every $x \notin M$ the closed half-space $H_x = \{y \mid a_x^T y \leq \alpha_x\}$ clearly contains M and does not contain x; consequently,

$$M = \cap_{x\notin M} H_x$$

and therefore M is not wider (and of course is not smaller) than the intersection of *all* closed half-spaces which contain M. ∎

Among the closed half-spaces which contain a closed convex and *proper* (i.e., nonempty and differing from the entire space) set M the most interesting are the "extreme" ones –

those with the boundary hyperplane touching M. The notion makes sense for an arbitrary (not necessary closed) convex set, but we will use it for closed sets only, and include the requirement of closedness in the definition:

Definition 1.2.3 [Supporting plane] *Let M be a convex closed set in $\mathbf{R}^n$, and let x be a point from the relative boundary of M. A hyperplane*

$$\Pi = \{y \mid a^T y = a^T x\} \quad [a \neq 0]$$

is called supporting to M at x, if it properly separates M and $\{x\}$, i.e., if

$$a^T x \geq \sup_{y \in M} a^T y \quad \& \quad a^T x > \inf_{y \in M} a^T y. \tag{1.2.11}$$

Note that since x is a point from the relative boundary of M and therefore belongs to $\mathrm{cl}\, M = M$, the first inequality in (1.2.11) in fact is equality. Thus, an equivalent definition of a supporting plane is as follows:

> *Let M be a closed convex set and x be a point of relative boundary of M. The hyperplane $\{y \mid a^T y = a^T x\}$ is called supporting to M at x, if the linear form $a(y) = a^T y$ attains its maximum on M at the point x and is nonconstant on M.*

E.g., the hyperplane $\{x_1 = 1\}$ in $\mathbf{R}^n$ clearly is supporting to the unit Euclidean ball $\{x \mid |x| \leq 1\}$ at the point $x = e_1 = (1, 0, ..., 0)$.

The most important property of a supporting plane is its existence:

Proposition 1.2.4 [Existence of supporting hyperplane] *Let M be a convex closed set in $\mathbf{R}^n$ and x be a point from the relative boundary of M. Then*

(i) *There exists at least one hyperplane which is supporting to M at x;*

(ii) *If Π is supporting to M at x, then the intersection $M \cap \Pi$ is of affine dimension less than the one of M (recall that the affine dimension of a set is, by definition, the affine dimension of the affine hull of the set, i.e., the linear dimension of the linear subspace L such that $\mathrm{Aff}\,(M) = x + L$).*

Proof. (i) is easy: if x is a point from the relative boundary of M, then it is outside the $\mathrm{ri}\, M$ and therefore $\{x\}$ and $\mathrm{ri}\, M$ can be properly separated by the Separation Theorem; the separating hyperplane is exactly the desired supporting to M at x hyperplane.

To prove (ii), note that if $\Pi = \{y \mid a^T y = a^T x\}$ is supporting to M at $x \in \partial_{\mathrm{ri}}\, M$, then the set $M' = M \cap \Pi$ is nonempty (it contains x) convex set, and the linear form $a^T y$ is constant on M' and therefore (why?) on the affine hull M'. At the same time, the form is nonconstant on M by definition of a supporting plane. Thus, $\mathrm{Aff}\,(M')$ is a proper (less than the entire $\mathrm{Aff}\,(M)$) subset of $\mathrm{Aff}\,(M)$, and therefore the affine dimension of M' is less than the affine dimension of M. [4]). ■

[4]) In the latter reasoning we used the following fact: if $P \subset M$ are two affine sets, then the affine dimension of P is $\leq$ the one of M, with $\leq$ being $=$ if and only if $P = M$. Please prove it.

1.3 Minimal representation of convex sets: extreme points

Supporting planes are useful tool to prove existence of *extreme points* of convex sets. Geometrically, an extreme point of a convex set M is a point in M which cannot be obtained as a convex combination of other points of the set; and the importance of the notion comes from the fact (which we will prove in the mean time) that the set of all extreme points of a "good enough" convex set M is the "shortest worker's instruction for building the set" – this is the smallest set of points for which M is the convex hull.

The exact definition of an extreme point is as follows:

Definition 1.3.1 [extreme points] *Let M be a nonempty convex set in $\mathbf{R}^n$. A point $x \in M$ is called an extreme point of M, if there is no nontrivial (of positive length) segment $[u, v] \in M$ for which x is an interior point, i.e., if the relation*

$$x = \lambda u + (1 - \lambda)v$$

with certain $\lambda \in (0, 1)$ and $u, v \in M$ is possible if and only if

$$u = v = x.$$

E.g., the extreme points of a segment are exactly its endpoints; the extreme points of a triangle are its vertices; the extreme points of a (closed) disk on the 2-dimensional plane are the points of the circumference.

An equivalent definition of an extreme point is as follows:

Proposition 1.3.1 $^{+}$ *A point x in a convex set M is extreme if and only if the set $M \backslash \{x\}$ is convex.*

It is clear that a convex set M not necessarily possesses extreme points; as an example you may take the open unit ball in $\mathbf{R}^n$. This example is not interesting – the set in question is not closed; when replacing it with its closure, we get a set (the closed unit ball) with plenty of extreme points – these are all points of the boundary. There are, however, *closed* convex sets which do not possess extreme points – e.g., a line or an affine set of larger dimension. A nice fact is that the absence of extreme points in a closed convex set M always has the standard reason – the set contains a line. Thus, a closed and nonempty convex set M which does not contain lines for sure possesses extreme points. And if M is nonempty convex compact, it possesses a quite representative set of extreme points – their convex hull is the entire M. Namely, we have the following

Theorem 1.3.1 *Let M be a closed and nonempty convex set in $\mathbf{R}^n$. Then*

(i) *The set* $\mathrm{Ext}(M)$ *of extreme points of M is nonempty if and only if M does not contain lines;*

(ii) *If M is bounded, then M is the convex hull of its extreme points:*

$$M = \mathrm{Conv}\,(\mathrm{Ext}(M)),$$

so that every point of M is a convex combination of the points of $\mathrm{Ext}(M)$.

Note that part (ii) of this theorem is the finite-dimensional version of the celebrated *Krein-Milman Theorem*.

Proof. Let us start with (i). The "only if" part is easy, due to the following simple

Lemma 1.3.1 *Let M be a closed convex set in $\mathbf{R}^n$. Assume that for some $\bar{x} \in M$ and $h \in \mathbf{R}^n$ M contains the ray*

$$\{\bar{x} + th \mid t \geq 0\}$$

starting at $\bar{x}$ with the direction h. Then M contains also all parallel rays starting at the points of M:

$$(\forall x \in M) : \{x + th \mid t \geq 0\} \subset M.$$

In particular, if M contains certain line, then it contains also all parallel lines passing through the points of M.

Comment. For a convex set M, the directions h such that $x + th \in M$ for some (and thus for all) $x \in M$ and all $t \geq 0$ are called recessive for M.

Proof of the lemma is immediate: if $x \in M$ and $\bar{x} + th \in M$ for all $t \geq 0$, then, due to convexity, for any fixed $\tau \geq 0$ we have

$$\epsilon(\bar{x} + \frac{\tau}{\epsilon}h) + (1 - \epsilon)x \in M$$

for all $\epsilon \in (0, 1)$. As $\epsilon \to +0$, the left hand side tends to $x + \tau h$, and since M is closed, $x + \tau h \in M$ for every $\tau \geq 0$. ■

Lemma 1.3.1, of course, resolves all our problems with the "only if" part. Indeed, here we should prove that if M possesses extreme points, then M does not contain lines, or, which is the same, that if M contains lines, then it has no extreme points. But the latter statement is immediate: if M contains a line, then, by Lemma, there is a line in M passing through any given point of M, so that no point can be extreme. ■

Now let us prove the "if" part of (i). Thus, from now on we assume that M does not contain lines; our goal is to prove that then M possesses extreme points. Let us start with the following

Lemma 1.3.2 *Let M be a closed convex set, $\bar{x}$ be a relative boundary point of M and Π be a hyperplane supporting to M at $\bar{x}$. Then all extreme points of the nonempty closed convex set $\Pi \cap M$ are extreme points of M.*

Proof of the Lemma. First, the set $\Pi \cap M$ is closed and convex (as an intersection of two sets with these properties); it is nonempty, since it contains $\bar{x}$ (Π contains $\bar{x}$ due to the definition of a supporting plane, and M contains $\bar{x}$ due to the closedness of M). Second, let a be the linear form associated with Π:

$$\Pi = \{y \mid a^T y = a^T \bar{x}\},$$

so that

$$\inf_{x\in M} a^T x < \sup_{x\in M} a^T x = a^T \bar{x} \tag{1.3.12}$$

(see Proposition 1.2.4). Assume that y is an extreme point of $\Pi \cap M$; what we should do is to prove that y is an extreme point of M, or, which is the same, to prove that

$$y = \lambda u + (1-\lambda)v$$

for some $u, v \in M$ and $\lambda \in (0,1)$ is possible only if $y = u = v$. To this end it suffices to demonstrate that under the above assumptions $u, v \in \Pi \cap M$ (or, which is the same, to prove that $u, v \in \Pi$, since the points are known to belong to M). Indeed, we know that y is an extreme point of $\Pi \cap M$, so that the relation $y = \lambda u + (1-\lambda)v$ with $\lambda \in (0,1)$ *and* $u, v \in \Pi \cap M$ does imply $y = u = v$.

To prove that $u, v \in \Pi$, note that since $y \in \Pi$ we have

$$a^T y = a^T \bar{x} \geq \max\{a^T u, a^T v\}$$

(the concluding inequality follows from (1.3.12)). On the other hand,

$$a^T y = \lambda a^T u + (1-\lambda) a^T v;$$

combining these observations and taking into account that $\lambda \in (0,1)$, we conclude that

$$a^T y = a^T u = a^T v.$$

But these equalities imply that $u, v \in \Pi$. ∎

Equipped with the Lemma, we can easily prove (i) by induction on the dimension of the convex set M (recall that this is nothing but the affine dimension of $\mathrm{Aff}\,(M) = a + L$).

There is nothing to do if the dimension of M is zero, i.e., if M is a point - then, of course, $M = \mathrm{Ext}(M)$. Now assume that we already have proved the nonemptiness of $\mathrm{Ext}(T)$ for all nonempty closed and not containing lines convex sets T of certain dimension k, and let us prove that the same statement is valid for the sets of dimension $k+1$. Let M be a closed convex nonempty and not containing lines set of dimension $k+1$. Since the dimension of M is positive ($k+1$), its relative interior is not empty. Since M does not contain lines and is of positive dimension, it differs from entire $\mathrm{Aff}\,(M)$ and therefore it possesses a relative boundary point $\bar{x}$. Indeed, there exists $z \in \mathrm{Aff}\,(M)\backslash M$, so that the points

$$x_\lambda = x + \lambda(z - x)$$

(x is an arbitrary fixed point of M) do not belong to M for some $\lambda \geq 1$, while $x_0 = x$ belongs to M. The set of those $\lambda \geq 0$ for which $x_\lambda \in M$ is therefore nonempty and bounded from above; this set clearly is closed (since M is closed). Thus, there exists the largest $\lambda = \lambda^*$ for which $x_\lambda \in M$. We claim that x_{λ^*} is a relative boundary point of M. Indeed, by construction this is a point from M. If it were a point from the relative interior of M, all the points x_λ with close to λ^* and greater than λ^* values of λ would also belong to M, which contradicts the origin of λ^*.

According to Proposition 1.2.4, there exists a hyperplane $\Pi = \{x \mid a^T x = a^T \bar{x}\}$ which supports M at $\bar{x}$:

$$\inf_{x\in M} a^T x < \max_{x\in M} a^T x = a^T \bar{x}.$$

By the same proposition, the set $T = \Pi \cap M$ (which is closed, convex and nonempty) is of affine dimension less than that of M, i.e., of the dimension $\leq k$. T clearly does not contain lines (since even the larger set M does not contain lines). By Inductive Hypothesis, T possesses extreme points, and by Lemma 1.3.2 all these points are extreme also for M. The inductive step is complete, and (i) is proved. ∎

Now let us prove (ii). Thus, let M be nonempty, convex, closed and bounded; we should prove that

$$M = \text{Conv}\,(\text{Ext}(M)).$$

What is immediately seen is that the right hand side set is contained in the left hand side one. Thus, all we need is to prove that any $x \in M$ is a convex combination of points from $\text{Ext}(M)$. Here we again use induction on the dimension of M. The case of 0-dimensional set M (i.e., a point) is trivial. Assume that the statement in question is valid for all k-dimensional convex closed and bounded sets, and let M be a convex closed and bounded set of dimension $k+1$. Let $x \in M$; to represent x as a convex combination of points from $\text{Ext}(M)$, let us pass through x an arbitrary line $l = \{x + \lambda h \mid \lambda \in \mathbf{R}\}$ $(h \neq 0)$ in the affine hull M, which we can suppose to be $\mathbf{R}^{k+1}$. Moving along this line from x in each of the two possible directions, we eventually leave M (since M is bounded); as it was explained in the proof of (i), it means that there exist nonnegative λ_+ and λ_- such that the points

$$\bar{x}_\pm = x + \lambda_\pm h$$

both belong to the boundary of M. Let us verify that $\bar{x}_\pm$ are convex combinations of the extreme points of M (this will complete the proof, since x clearly is a convex combination of the two points $\bar{x}_\pm$). Indeed, M admits supporting at $\bar{x}_+$ hyperplane Π; as it was explained in the proof of (i), the set $\Pi \cap M$ (which clearly is convex, closed and bounded) is of less dimension than that one of M; by the inductive hypothesis, the point $\bar{x}_+$ of this set is a convex combination of extreme points of the set, and by Lemma 1.3.2 all these extreme points are extreme points of M as well. Thus, $\bar{x}_+$ is a convex combination of extreme points of M. Similar reasoning is valid for $\bar{x}_-$. ∎

1.3.1 Application: Extreme points of a polyhedral set

Consider a polyhedral set

$$K = \{x \in \mathbf{R}^n \mid Ax \leq b\},$$

A being a $m \times n$ matrix and b being a vector from $\mathbf{R}^m$. What are the extreme points of K? The answer is given by the following

Theorem 1.3.2 [Extreme points of polyhedral set]
Let $x \in K$. The vector x is an extreme point of K if and only if some n linearly independent (i.e., with linearly independent vectors of coefficients) inequalities of the system $Ax \leq b$ are equalities at x.

Proof. Let a_i, $i = 1, ..., m$, be the the rows of A.

The "only if" part: let x be an extreme point of K, and let I be the set of those indices i for which $a_i^T x = b_i$; we should prove that the set F of vectors $\{a_i \mid i \in I\}$ contains n linearly independent vectors, or, which is the same, that linear span of F is $\mathbf{R}^n$. Assume that it is not the case; then the orthogonal complement to F contains a nonzero vector h (since the dimension of $F^\perp$ is equal to $n-$ the dimension of the linear span of F, and is therefore positive). Consider the segment $\Delta_\epsilon = [x - \epsilon h, x + \epsilon h]$, $\epsilon > 0$ being the parameter of our construction. Since h is orthogonal to the "active" vectors a_i – those with $i \in I$, all points y of this segment satisfy the relations $a_i^T y = a_i^T x = b_i$. Now, if i is a "nonactive" index – one with $a_i^T x < b_i$ – then $a_i^T y \leq b_i$ for all $y \in \Delta_\epsilon$, provided that ϵ is small enough. Since there are finitely many nonactive indices, we can choose $\epsilon > 0$ in such a way that all $y \in \Delta_\epsilon$ will satisfy all "nonactive" inequalities $a_i^T x \leq b_i$, $i \notin I$. Since $y \in \Delta_\epsilon$ satisfies, as we have seen, also all "active" inequalities, we conclude that with the above choice of ϵ we get $\Delta_\epsilon \subset K$, which is a contradiction: $\epsilon > 0$ and $h \neq 0$, so that Δ_ϵ is a nontrivial segment with the midpoint x, and no such segment can be contained in K, since x is an extreme point of K. ∎

To prove the "if" part, assume that $x \in K$ is such that among the inequalities $a_i^T x \leq b_i$ which are equalities at x there are n linearly independent, say, those with indices $1, ..., n$, and let us prove that x is an extreme point of K. This is immediate: assuming that x is not an extreme point, we would get the existence of a nonzero vector h such that $x \pm h \in K$. In other words, for $i = 1, ..., n$ we would have $b_i \pm a_i^T h \equiv a_i^T (x \pm h) \leq b_i$, which is possible only if $a_i^T h = 0$, $i = 1, ..., n$. But the only vector which is orthogonal to n linearly independent vectors in $\mathbf{R}^n$ is the zero vector (why?), and we get $h = 0$, which was assumed not to be the case. ∎.

Corollary 1.3.1 *The set of extreme points of a polyhedral set is finite.*

Indeed, according the above theorem, every extreme point of a polyhedral set $K = \{x \in \mathbf{R}^n \mid Ax \leq b\}$ satisfies the equality version of certain n-inequality subsystem of the original system, the matrix of the subsystem being nonsingular. Due to the latter fact, an extreme point is uniquely defined by the corresponding subsystem, so that the number of extreme points does not exceed the number C_m^n of $n \times n$ submatrices of the matrix A and is therefore finite. ∎

Note that C_m^n is nothing but an upper (ant typically very conservative) bound on the number of extreme points of a polyhedral set given by m inequalities in $\mathbf{R}^n$: some $n \times n$ submatrices of A can be singular and, what is more important, the majority of the nonsingular ones normally produce "candidates" which do not satisfy some of the remaining inequalities.

1.4 Exercises: convex sets

Exercise 1.4.1 *Which of the following sets are convex:*

- $\{x \in \mathbf{R}^n \mid \sum_{i=1}^n x_i^2 = 1\}$
- $\{x \in \mathbf{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1\}$
- $\{x \in \mathbf{R}^n \mid \sum_{i=1}^n x_i^2 \geq 1\}$
- $\{x \in \mathbf{R}^n \mid \max_{i=1,\dots,n} x_i \leq 1\}$
- $\{x \in \mathbf{R}^n \mid \max_{i=1,\dots,n} x_i \geq 1\}$
- $\{x \in \mathbf{R}^n \mid \max_{i=1,\dots,n} x_i = 1\}$
- $\{x \in \mathbf{R}^n \mid \min_{i=1,\dots,n} x_i \leq 1\}$
- $\{x \in \mathbf{R}^n \mid \min_{i=1,\dots,n} x_i \geq 1\}$
- $\{x \in \mathbf{R}^n \mid \min_{i=1,\dots,n} x_i = 1\}$

Exercise 1.4.2 *Which of the following pairs (S, T) of sets are (a) properly separated and (b) strongly separated by the linear form $f(x) = x_1$:*

- $S = \{x \in \mathbf{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1\}$, $T = \{x \in \mathbf{R}^n \mid x_1 + x_2 \geq 2, x_1 - x_2 \geq 0\}$;
- $S = \{x \in \mathbf{R}^n \mid \sum_{i=1}^n x_i^2 \leq 1\}$, $T = \{x \in \mathbf{R}^n \mid x_1 + x_2 \geq 2, x_1 - x_2 \geq -1\}$;
- $S = \{x \in \mathbf{R}^n \mid \sum_{i=1}^n |x_i| \leq 1\}$, $T = \{x \in \mathbf{R}^n \mid x_1 + x_2 \geq 2, x_1 - x_2 \geq 0\}$;
- $S = \{x \in \mathbf{R}^n \mid \max_{i=1,\dots,n} x_i \leq 1\}$, $T = \{x \in \mathbf{R}^n \mid x_1 + x_2 \geq 2, x_1 - x_2 \geq -1\}$;
- $S = \{x \in \mathbf{R}^n \mid x_1 = 0\}$, $T = \{x \in \mathbf{R}^n \mid x_1 \geq \sqrt{x_2^2 + \dots + x_n^2}\}$;
- $S = \{x \in \mathbf{R}^n \mid x_1 = 0\}$, $T = \{x \in \mathbf{R}^n \mid x_1 = 1\}$;
- $S = \{x \in \mathbf{R}^n \mid x_1 = 0, x_2^2 + \dots + x_n^2 \leq 1\}$, $T = \{x \in \mathbf{R}^n \mid x_1 = 0, x_2 \geq 100\}$;
- $S = \{x \in \mathbf{R}^2 \mid x_1 > 0, x_2 \geq 1/x_1\}$, $T = \{x \in \mathbf{R}^2 \mid x_1 < 0, x_2 \geq -1/x_1\}$.

Exercise 1.4.3 *Prove the following important result*

Homogeneous Farkas Lemma *Let $a_1, ..., a_N$ be vectors of $\mathbf{R}^n$.*

Vector a is a conic combination of $a_1, ..., a_N$ (in other words, there are $\lambda_1 \geq 0, ..., \lambda_N \geq 0$ such that $a = \sum_{i=1}^N \lambda_i a_i$) if and only if *for all x such that $x^T a_i \geq 0$, $i = 1, ..., N$, we also have $x^T a \geq 0$.*

Otherwise, linear inequality $a^T x \geq 0$ is a consequence of the system of linear inequalities $Ax \geq 0$ (i.e. this inequality holds for any solution x of the system) if and only if *there is a nonnegative $\lambda \in \mathbf{R}^N$ such that $a = A^T \lambda$.*

Hint: of course, the "only if" statement is harder (why? what happens if we sum inequalities with nonnegative coefficients?). Let $K = \{\sum \lambda_i a_i | \lambda_i \geq 0\}$. Use the Separation Theorem to show that if $a \in \mathbf{R}^n$ does not belong to K then there exists x such that $a^T x < 0$ and $a_i^T x \geq 0$, and derive Homogeneous Farcas Lemma from this fact.