A simple view on convex analysis and its applications

Jan Brinkhuis
Erasmus University Rotterdam

Vladimir Tikhomirov
Moscow State University

3 October 2005

Econometric Institute Report EI-2005-37

Quand une situation, de la plus humble à la plus vaste, a été comprise dans les aspects essentiels, la démonstration de ce qui est compris (et du reste) tombe comme un fruit mûr à point. [...]

When a situation, from the most humble to the most immense, has been understood in the essential aspects, the proof of what is understood (and of the remainder) falls like a fruit that is just ripe. [...]

A. Grothendieck

Abstract.

Our aim is to give a simple view on the basics and applications of convex analysis. The essential feature of this account is the systematic use of the possibility to associate to each convex object—such as a convex set, a convex function or a convex extremal problem)—a *cone*, without loss of information. The core of convex analysis is the possibility of the dual description of convex objects, geometrical and algebraical, based on the duality of vector spaces; for each type of convex objects, this property is encoded in an *operator of duality*, and the name of the game is how to calculate these operators. The core of this paper is a unified presentation, for each type of convex objects, of the duality theorem and the complete list of calculus rules.

Now we enumerate the advantages of the 'cone'-approach. It gives a unified and transparent view on the subject. The intricate rules of the convex calculus all flow naturally from one common source. We have included for each rule a precise description of the weakest convenient assumption under which it is valid. This appears to be useful for applications; however, these assumptions are usually not given. We explain why certain convex objects have to be excluded in the definition of the operators of duality: the collections of associated cones of the target of an operator of duality need not be closed (here 'closed' is meant in an algebraic sense). This makes clear that the remedy is to take the closure of the target. As a byproduct of the cone approach, we have found the solution of the open problem of how to use the *polar* operator to give a dual description of arbitrary convex sets.

The approach given can be extended to the infinite-dimensional case.

1 Introduction

For many analytical problems, it is the *convexity* rather than the differentiability that is the vital property. Then it is natural to replace the methods of differential analysis by those of convex analysis.

In convex analysis, one studies convex sets, convex functions and convex extremal problems. The subject represents great geometrical beauty and analytical power; there are extremely many applications of the methods of convex analysis. The central property of convex objects is the possibility of their dual description, geometrical and algebraical, based on the duality of vector spaces. The official moment of birth of the subject is the publication of the monograph Convex Analysis by R.T. Rockafellar, but its origins can be traced back, for example to work of W. Fenchel on the dual description of convex functions, and even further back, to the work of H. Minkowski who started the systematic study of convexity. The earliest trace of it is the well-known observation that each statement in plane geometry about inclusion of points in lines has a dual statement, where the role of points and lines is interchanged. We recall the simplest example: the dual statement of *'through two different points runs precisely one line'* is *'two different, non-parallel lines have precisely one point in common'*. The need to add 'non-parallel' here, illustrates that the transition to the dual is not completely straightforward. The transition to the dual is the main subject of the present paper.

The operators of duality such as the *subdifferential*, the *polar operator*, the *conjugate function operator* (or *Young-Fenchel transformation*) and many others, allow a transition from one description to the other; the calculation of these operators can be carried out using certain rules of the convex calculus such as the theorems of *Moreau-Rockafellar*, *Fenchel-Moreau*, *Dubovitskii-Milyutin* and many others.

There is a parallel with differential analysis. There the calculation of the operator of differentiation, the derivative, can be carried out using the rules of the differential calculus such as the chain rule and the product rule. The explanation of this parallel is that the idea behind the operators of duality is the same as the one behind the operator of differentiation: in both cases one wants to profit from the available structure, convexity in one case and differentiability in the other one, in order to approximate a nonlinear object by a linear one.

This paper is organized as follows.

- Section 2: enumeration of all convex objects to be considered and the main binary operations on these convex objects.
- Section 3: presentation, for one type of convex objects, *cones*, of the theorem of duality and the four basic calculus rules for the operator of duality, *the conjugate cone operator*: for sum and intersection, and, more generally, for image and inverse image under a linear operator. Novel presentation of the proof of the main theorem of convex analysis in a geometrical spirit.
- Section 4: enumeration, for each type of convex object, of the duality theorem and the calculus rules for the operator of duality.
- Section 5: presentation of a systematic 'cone-approach', which gives automatically, for each type of convex objects, the duality theorem and the calculus rules for the operator of duality.

We conclude the paper by giving a selection of applications to various basic topics.

- Section 6: the calculus rules for the subdifferential of *convex functions at a point* are formulated; these rules are derived by *localization* from the calculus rules for sublinear functions.

- Section 7: a problem on the polar operator is settled: how to use this operator to give the dual description of a convex set that does *not* contain zero.
- Section 8: the results of separation of convex sets are given ((strict) separation, Hahn-Banach theorem, Farkas' lemma).
- Section 9: the multiplier theory for convex extremal problems is presented (convex Lagrange multiplier rule, theory of primal and dual problems, shadow price interpretation of multipliers and of solutions of dual problems).
- Section 10: for most types of convex objects, one application is given illustrating the duality theorem and the calculation of the operator of duality.

2 Basic concepts: convex objects and binary operations

We enumerate the main types of convex objects and the main binary operations on these types of convex objects. Let X be a vector space; **in this paper, this will mean that $X = \mathbb{R}^n$, the space of n-dimensional column vectors**. However, as we will explain, almost all results extend to the infinite dimensional case.

A **set** $A \subseteq X$ is called *convex*, if along with each two points $x_i,\ i = 1, 2$, it contains the whole segment

$$[x_1, x_2] = \{x \in X : x = \alpha_1 x_1 + \alpha_2 x_2,\ \alpha_i \geq 0,\ i = 1, 2,\ \alpha_1 + \alpha_2 = 1\}.$$

We denote the collection of all convex sets by $\mathrm{Co}(X)$. We single out the following subcollections of $\mathrm{Co}(X)$:

- the linear subspaces $\mathrm{Lin}(X)$, that is, the nonempty sets containing with each two points x_1 and x_2 all linear combinations $\alpha_1 x_1 + \alpha_2 x_2,\ \alpha_i \in \mathbb{R},\ i = 1, 2$;
- the affine subspaces $\mathrm{Aff}(X)$, that is, the translates of linear subspaces;
- the convex cones, or just *cones*, $\mathrm{Cone}(X)$, that is, the convex sets that are invariant under each homothety (multiplication by a positive scalar), and that contain zero ('the origin');
- the convex zero-sets $\mathrm{Co}_0(X)$, that is, the convex sets containing zero.

Some types of convex sets are special cases of other types of convex sets. Explicitly, one has the following inclusions:

$$\mathrm{Lin}(X) \subseteq \mathrm{Aff}(X) \subseteq \mathrm{Co}(X) \supseteq \mathrm{Co}_0(X) \supseteq \mathrm{Cone}(X) \supseteq \mathrm{Lin}(X).$$

One can make from two convex sets a new convex set by means of the following binary operations on $\mathrm{Co}(X)$:

- sum: $A_1 + A_2 = \{a_1 + a_2 : a_1 \in A_1,\ a_2 \in A_2\}$;
- intersection: $A_1 \cap A_2$;

- convex hull of the union: $A_1 \mathrm{co} \cup A_2 = \{\alpha a_1 + (1-\alpha)a_2,\ a_1 \in A_1,\ a_2 \in A_2,\ 0 \le \alpha \le 1\}$.

These three operations preserve the types listed above. Moreover, these operations—and other ones—can be viewed as special cases of the operations image and inverse image of convex sets under a linear operator. That is, the importance of these operations, to be defined now, is that these are the source of all operations.

Image and inverse image of convex sets. Let $Y = \mathbb{R}^m$ be another vector space and $\Lambda : X \to Y$ a linear operator. We recall that a linear operator is essentially a matrix: for each linear operator $\mathbb{R}^n \to \mathbb{R}^m$ there is a unique $m \times n$-matrix M such that $\Lambda(x) = Mx$ for all $x \in \mathbb{R}^n$. By $\Lambda A \in \mathrm{Co}(Y)$ we denote the *image of the set* $A \in \mathrm{Co}(X)$, that is,

$$\Lambda A = \{\Lambda x : x \in A\};$$

by $B\Lambda \in \mathrm{Co}(X)$ we denote the *inverse image of the set* $B \in \mathrm{Co}(Y)$, that is,

$$B\Lambda = \{x \in X : \Lambda x \in B\}.$$

For example, $A_1 + A_2$ is the image of the product $A_1 \times A_2$ under the addition mapping $+ : X \times X \to X$ and $A_1 \cap A_2$ is the inverse image of $A_1 \times A_2$ under the diagonal embedding $d : X \to X \times X$. The convex hull of the union requires an additional idea, the reduction of convex sets to convex cones, to be given later.

A **function** $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is called *convex* if its epigraph

$$\mathrm{epi} f = \{(x, \alpha) : x \in X,\ \alpha \in \mathbb{R},\ \alpha \ge f(x)\}$$

is a convex set in $X \times \mathbb{R}$. We denote the collection of all convex functions by $\mathrm{Co}^f(X)$. To each subcollection of $\mathrm{Co}(X)$ there corresponds a subcollection of $\mathrm{Co}^f(X)$. We single out the following subcollections of $\mathrm{Co}^f(X)$ (we make a slightly different choice here than for $\mathrm{Co}(X)$: we do not consider the type corresponding to $\mathrm{Co}_0(X)$, and we choose the type $\mathrm{Norm}(X)$, although we have not chosen the corresponding subcollection of $\mathrm{Co}(X)$):

- the linear functionals $\mathrm{Lin}^f(X)$;
- the affine functions $\mathrm{Aff}^f(X)$, that is, the functions $x \mapsto a(x) = \langle x', x\rangle + \alpha$ (where x' is a linear functional on X, $\langle x', x\rangle$ denotes its action on the element x, and $\alpha \in \mathbb{R}$);
- the *sublinear* functions $\mathrm{Cone}^f(X)$, that is, the functions $p : X \to \mathbb{R} \cup \{\pm\infty\}$ for which the epigraph is a cone, and $p(0) = 0$. To be more explicit, these are the convex functions $p(\cdot)$ on X with $p(\alpha x) = \alpha p(x),\ \alpha > 0,\ x \in X$, and $p(0) = 0$;
- the *norms* $\mathrm{Norm}(X)$, that is, the sublinear functions N that are positive real valued and symmetric, that is,

  $N(x) \in (0, +\infty)\ \forall x \in X \setminus \{0\}$;

  $N(-x) = N(x)\ \forall x \in X$.

Some types of convex functions are special cases of other types of convex functions. Explicitly, one has the following inclusions:

$$\mathrm{Lin}^f(X) \subseteq \mathrm{Aff}^f(X) \subseteq \mathrm{Co}^f(X) \supseteq \mathrm{Cone}^f(X) \supseteq \mathrm{Lin}^f(X),$$

and $\mathrm{Norm}(X) \subseteq \mathrm{Cone}^f(X)$.

One can make from two convex functions a new convex function by means of the following binary operations on $\mathrm{Co}^f(X)$:

- sum: $(f_1 + f_2)(x) = f_1(x) + f_2(x)$;
- maximum: $(f_1 \wedge f_2)(x) = \max(f_1(x), f_2(x))$;
- convex hull of minimum $(f_1 \mathrm{co} \vee f_2)(x) = \inf\{(1-\alpha)f_1(x_1) + \alpha f_2(x_2) : x = (1-\alpha)x_1 + \alpha x_2,\ 0 \leq \alpha \leq 1\}$;

These three operations preserve the types listed above. Moreover, these operations—and other ones—can be viewed as special cases of the operations image and inverse image of convex functions under a linear operator.

Image and inverse image of convex functions. Let $X,\ Y$ be vector spaces and $\Lambda : X \to Y$ a linear operator. By $f\Lambda \in \mathrm{Co}^f(X)$ we denote the *inverse image of the function* $f \in \mathrm{Co}^f(Y)$, that is,

$$f\Lambda(x) = f(\Lambda x));$$

by $\Lambda f \in \mathrm{Co}^f(Y)$ we denote the *image of the function* $f \in \mathrm{Co}^f(X)$, that is,

$$\Lambda f(y) = \inf\{f(x) : \Lambda x = y\}.$$

An **extremal problem** (or **optimization problem**)

$$f(x) \to \min,\ x \in X,$$

is called *convex*, if $f : X \to \mathbb{R} \cup \{\pm\infty\}$ is a convex function. This includes the class of *convex programming problems*

$$f_0(x) \to \min,\ f_i(x) \leq 0,\ 1 \leq i \leq m,\ x \in A, \tag{Q}$$

where $f_i \in \mathrm{Co}^f(X),\ 0 \leq i \leq m$, and $A \subseteq \mathrm{Co}(X)$.

We single out the following subcollections of convex programming problems:

- If $X = \mathbb{R}^n,\ f_i \in \mathrm{Aff}^f(X),\ 0 \leq i \leq m$, and $A = \mathbb{R}^n_+$, the first orthant, then problem (Q) is called a *linear programming problem.*
- If X is the space of hermitian $n \times n$-matrices M ('hermitian' means complex matrix with $m_{lk} = \bar{m}_{kl}\,\forall k, l),\ f_i \in \mathrm{Aff}^f(X),\ 0 \leq i \leq m$, and A is the set of positive definite matrices in X ('positive definite' means $\bar{v}^T M v > 0$ for all nonzero vectors $v \in \mathbb{C}^n$), then problem (Q) is called a *semi-definite programming problem.*

3 Duality and calculus rules for convex cones

The aim of this section is to present the calculus for the operator of duality of one type of convex objects, *cones*. This paper is based on the idea that the easiest access to convex analysis is by means of this type of convex objects.

Let X be a vector space. By X' we denote its conjugate vector space, that is, the collection $\text{Lin}^f(X)$ of all linear functionals on X. By $\langle\cdot,\cdot\rangle : X' \times X \to \mathbb{R}$ we denote the bilinear form where $\langle x', x\rangle$ is the action of the linear functional $x' \in X'$ on the element $x \in X$. Explicitly, for $X = \mathbb{R}^n$ we have $X' = (\mathbb{R}^n)^T$ and $\langle x', x\rangle = \sum_{i=1}^n x'_i x_i$ is the standard inner product of the column-vectors $(x')^T$ and x. Note that $X'' = X$ if $X = \mathbb{R}^n$.

Operator of duality. We introduce the operator of duality for cones, associating to a cone $C \subseteq X$ a cone $C' \subseteq X'$, called *the conjugate cone* of C:

$$C' = \{x' \in X' : \langle x', x\rangle \geq 0 \ \forall x \in C\}. \qquad (*)$$

When it is not clear from the context that we view C as a cone in the vector space X, we will use the more precise notation $(C, X)'$ for the conjugate cone.

Geometrical interpretation. A *hyperplane through zero* in X is defined to be a linear subspace M in X of codimension one, that is, with $\dim X/M = 1$. Equivalently, a hyperplane through zero in X is a set of the form

$$M_{x'} = \{x \in X : \langle x', x\rangle = 0\}$$

for some nonzero $x' \in X'$. For each nonzero $x' \in C'$, the hyperplane $M_{x'}$ has the cone C completely on one of its two sides (where it is allowed that C has points in common with the hyperplane; C is even allowed to be contained entirely in the hyperplane). Conversely, a hyperplane through zero that has C completely on one of its sides—this side has to be *chosen* if C is contained in the hyperplane—determines a nonzero element of C' up to a nonzero positive scalar multiple.

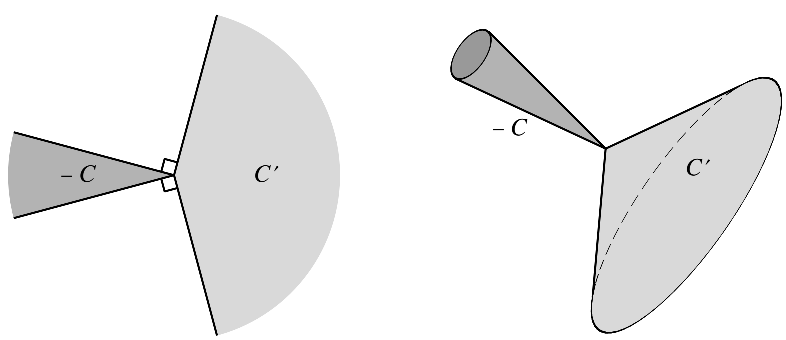

Figure 1: Conjugate cones.

The definition of conjugate cones is illustrated in Figure 1 for $X = \mathbb{R}^2$ and $X = \mathbb{R}^3$ (note that $(\mathbb{R}^n)'$ can be identified with $\mathbb{R}^n$: each $a \in \mathbb{R}^n$ can be viewed as the linear functional on $\mathbb{R}^n$ given by

$x \mapsto \sum_{i=1}^{n} a_i x_i$; therefore, a cone in $\mathbb{R}^n$ can be viewed as lying in the same space as its conjugate cone).

For later use we record the geometrical interpretation formally. For each set S in a vector space, the *linear span* $\mathrm{sp}S$ of S is the smallest linear subspace in X containing S. For each set S in a vector space X the *interior* $\mathrm{int}S$ is the largest open set contained in S.

Proposition 1 *For a cone $C \in \mathrm{Cone}(X)$, the following cases can be distinguished:*

1. ***Degenerate case:*** $\mathrm{sp}C \neq X$. *Then there exists a hyperplane through zero in X that contains C. Such a hyperplane, together with a choice of one of its two sides, corresponds to a nonzero element of C', up to a positive scalar multiple.*

2. ***Nondegenerate case:*** $\mathrm{sp}C = X$. *Then a hyperplane through zero that is disjoint from* $\mathrm{int}(C)$ *corresponds to a nonzero element of C', up to a positive scalar multiple. Moreover, the following criterion holds true for an element $\bar{c} \in \mathrm{int}C$:*

$$C \neq X \Leftrightarrow -\bar{c} \notin C.$$

Duality theory. We single out in $\mathrm{Cone}(X)$ the subcollection of the closed cones, denoted by $\mathrm{ClCone}(X)$.

The *conjugate* of a linear operator $\Lambda : X \to Y$ is the linear operator $\Lambda' : Y' \to X'$ defined by

$$\Lambda'(y')(x) = y'(\Lambda x) \ \forall x \in X \ \forall y' \in Y'.$$

We define for each set S in X its *relative interior* $\mathrm{ri}S$ to be the largest open subset of the affine hull of S that is contained in S. Note that for each $C \in \mathrm{Cone}(X)$, the cone C'' lies in the same space as C as $X'' = X$.

Theorem 1 ***Duality theorem and basic rules for cones.***

1. $C'' = C \Leftrightarrow C \in \mathrm{ClCone}(X)$,

2. $(C_1 + C_2)' = C_1' \cap C_2'$,

3. $(C_1 \cap C_2)' = C_1' + C_2'$ *if* $\mathrm{ri}C_1 \cap \mathrm{ri}C_2 \neq \emptyset$,

4. $(\Lambda C)' = C'\Lambda'$,

5. $(C\Lambda)' = \Lambda' C'$ *if* $\mathrm{ri}C \cap \Lambda X \neq \emptyset$.

Remark 1. The following weak looking property of cones follows readily from statement 1 of theorem 1:

$$C \in \mathrm{Cone}(X), \ C \neq X \ \Rightarrow C' \neq 0.$$

However, *this is in fact the key property of cones*; below it will be formally presented as theorem 2, it will be proved and then all statements of theorem 1 will be derived as consequences.

Remark 2. Statement 1 of theorem 1 gives the dual description of a closed cone C: the duality operator $'$ is an *involution* on the collection of closed cones, that is, if $C \in \text{ClCone}(X)$ and $D \in \text{ClCone}(X')$, then the following implication holds true:

$$D = C' \Rightarrow C = D'.$$

In geometrical terms, each closed cone C is the intersection of the following collection of closed half-spaces

$$\mathcal{H}_{x'} = \{x \in X : \langle x', x\rangle \geq 0\},\ x' \in C' \setminus \{0\}.$$

Statement 1 of theorem 1 can be formulated in the following alternative form:

$$C'' = \text{cl}C \text{ for all } C \in \text{Cone}(X).$$

Remark 3. Now we explain the relevance of theorem 1. A main task of the convex analysis is the explicit calculation of the operator of duality for convex objects. *Convex* analysis runs parallel to *differential* analysis. The operators of duality correspond to the operator of differentiation. The convex calculus rules for convex objects have a similar purpose as the well-known differential calculus rules for differentiable functions, such as the chain rule (resp. the product rule) for the derivative of a composition (resp. a product) of functions: in both cases the rules make it possible to calculate the operators. For *cones*, the convex objects under discussion in this section, the convex calculus rules consist of the statements 2.–5. of theorem 1. Note that the statements 2 and 3 are special cases of the statements 4 and 5 respectively. For other convex objects, all convex calculus rules will be derived from theorem 1.

Now we give a quick and novel presentation, in a geometrical spirit, of the proof of theorem 1. We will derive theorem 1 from the following result.

Theorem 2 ***Main theorem of the convex analysis.*** *For each cone $C \in \text{Cone}(X)$ for which $C \neq X$, the conjugate cone C' contains nonzero elements,*

$$C' \neq 0.$$

Proof. It suffices to give the proof in the nondegenerate case, clearly. The strategy of the proof is to show that the following set contains among its elements a hyperplane in X:

the set $\mathcal{S}$ of linear subspaces in X that are disjoint from int$C$.

Such a hyperplane must have the cone C on one of its sides, and so its existence proves the theorem.

We make two observations on the set $\mathcal{S}$.

1. The set $\mathcal{S}$ contains at least one nontrivial linear subspace, provided $\dim X > 1$.

   Indeed, if $\dim X > 1$, then we can choose a two dimensional linear subspace ('plane through zero') L in X that contains at least one point $\bar{c}$ from int$C$ and so one point that is not in C, for example $-\bar{c}$. Then, by applying the—easy—classification of cones in a plane to the intersection $C \cap L$, it follows that there exists a one dimensional linear subspace ('line through zero') in the plane L that is disjoint from int$C$.

2. The space X itself is not an element of $\mathcal{S}$.

   This follows from the assumption that int$C$ is nonempty.

Now we can finish the proof. Take an element M of $\mathcal{S}$ of maximal dimension. Let X/M be the quotient space and let $i : X \to X/M$ be the natural mapping $x \mapsto x + M$. Consider the cone $i(C) \subseteq X/M$. Use the following fact from linear algebra: the set of linear subspaces $L \subseteq X$ containing M corresponds bijectively to the set of linear subspaces $\tilde{L} \subseteq X/M$:

L corresponds to $\tilde{L}$ precisely if $i(L) = \tilde{L}$.

Therefore, it follows from the maximality property of M, that there are no nontrivial linear subspaces of the quotient space X/M that are disjoint from $\mathrm{int}(i(C)) = i(\mathrm{int}C)$. Therefore, by the first observation above, applied to the cone $i(C)$ in the vector space X/M, it follows that $\dim X/M \leq 1$. Moreover, by the second observation above, the case $\dim X/M = 0$ has to be excluded, and so $\dim X/M = 1$. That is, M is a hyperplane in X, as required. ■

Remark 4. It is remarkable that there is a completely different proof for this basic result. We give a sketch of this alternative.

Identify X with $\mathbb{R}^n$—and so also X' with $\mathbb{R}^n$, using the natural identification between $(\mathbb{R}^n)' = (\mathbb{R}^n)^T$ and $\mathbb{R}^n$ given by taking the transpose—and apply then the existence theorem of Weierstrass to show that the shortest distance problem for the set cl$C$, and a point $p \notin \mathrm{cl}C$, has a solution $\widehat{c} \in \mathrm{cl}C$. Writing out the optimality property of $\widehat{c}$ explicitly, gives that the nonzero vector $\widehat{c} - p$, viewed as an element of X', is a nonzero element of C'.

Now we give a relative version of theorem 2.

Corollary 1 ***Main theorem of the convex analysis (relative version).*** *For a vector space X, a convex cone C in X, and a linear subspace L in X with $\mathrm{ri}C \cap L \neq \emptyset$, the restriction mapping*

$$(C, X)' \to (L \cap C, L)',$$

defined by restricting linear functionals from X to L, is surjective.

Proof. To prove the corollary, it suffices to show that each nonzero $y' \in (L \cap C, L)'$ can be extended to an element of $(C, X)'$.

1. *Nondegenerate case*: $\mathrm{sp}\, C = X$. Then $\mathrm{int}\, C = \mathrm{ri}\, C$ and so it is nonempty. One can repeat the proof of theorem 2 in the nondegenerate case, with the set $\mathcal{S}$ replaced by the set $\mathcal{S}_{y'}$ of linear subspaces of X that are disjoint from int$C$ and contain $\ker y' = \{l \in L : \langle y', l \rangle = 0\}$.

2. *Degenerate case*: $\mathrm{sp}\, C \neq X$. Restrict y' to the linear subspace $L \cap \mathrm{sp}C$—this is an element of

   $$(L \cap C, L \cap \mathrm{sp}C)'$$

   —and extend the resulting functional to an element z' of $(C, \mathrm{sp}C)'$; this is possible as the natural mapping

   $$(C, \mathrm{sp}C)' \to (L \cap C, L \cap \mathrm{sp}C)'$$

is surjective by theorem 2 in the nondegenerate case, which has been established above. Let w' be the linear functional on $L + \mathrm{sp}C$ for which the restriction to $\mathrm{sp}C$ equals z' and for which the restriction to L equals y'. This is possible as the linear functionals $z' : \mathrm{sp}C \to \mathbb{R}$ and $y' : L \to \mathbb{R}$ agree on $L \cap \mathrm{sp}C$ by the choice of z'. Note that

$$w' \in (C, L \cap \mathrm{sp}C)'.$$

Choose an extension of w' to a linear functional on X; this is an extension of y' to an element of $(C, X)'$, as required.

■

Now we are ready to prove theorem 1.

Proof.

1. To prove this statement, it suffices to show that $C'' \subseteq C$ if $C \in \mathrm{ClCone}(X)$. To establish this inclusion, it suffices to show that for all $x \in X \setminus C$ there exists $x' \in C'$ with $\langle x', x \rangle < 0$. Choose $x \in X \setminus C$. We distinguish two cases.

   - $x \notin \mathrm{sp}C$.

     Take the linear functional on $\mathbb{R}x + \mathrm{sp}C$ that takes value zero on $\mathrm{sp}C$ and value -1 on x. Then extend this functional to a linear functional x' on X. This has the required properties.

   - $x \in \mathrm{sp}C$.

     Choose a point $v \in \mathrm{ri}C$ and a point y on the open interval (v, x) that does not belong to C. Then apply theorem 2 to the smallest cone containing C and $-y$ in the vector space $\mathrm{sp}C$: this is allowed as this cone cannot equal $\mathrm{sp}C$ by the choice of y: in particular, it does not contain the point y. Then extend the resulting nonzero element of

     $$(C + \mathbb{R}_+ \cdot y,\, \mathrm{sp}C)'$$

     to a linear functional x' on X. Note that

     $$x' \in C',\ \langle x', y \rangle \leq 0,\ \langle x', v \rangle > 0,$$

     (the last inequality holds as $v \in \mathrm{ri}C$ and $x' \neq 0$). It follows that $\langle x', x \rangle < 0$, taking into account that the point y lies on the open interval with endpoints v and x.

2. By the definitions.

3. Let Δ be the diagonal subspace $\{(x, x) : x \in X\}$ in $X \times X$. It suffices to prove that the restriction mapping

   $$(C_1 \times C_2, X \times X)' \to ((C_1 \times C_2) \cap \Delta, \Delta)',$$

   given by restriction of functionals from $X \times X$ to Δ, is surjective, as $((C_1 \times C_2) \cap \Delta, \Delta)'$ is isomorphic to $(C_1 \cap C_2, X)'$ by the definitions, and as this restriction mapping factorizes by way of the inclusion mapping from $(C_1,X)'+(C_2,X)'$ into $(C_1\cap C_2,X)'$. The desired surjectivity follows from corollary 1 with

$$X:=X\times X,\ C:=C_1\times C_2,\ L:=\Delta.$$

Indeed, choose $c\in \mathrm{ri}C_1\cap \mathrm{ri}C_2$, then

$$(c,c)\in \mathrm{ri}(C_1\times C_2)\cap\Delta,$$

so $\mathrm{ri}(C_1\times C_2)\cap\Delta\neq\emptyset$, and so corollary 1 can be applied.

4. By definition.

5. This is a formal consequence of what we have proved already. To be more precise, it follows from the following chain of equalities and natural isomorphisms:

$$(C\Lambda)'\overset{(i)}{=}(C\Lambda+\ker\Lambda)'\overset{(ii)}{=}(C\Lambda)'\cap(\ker\Lambda)'\overset{(iii)}{\simeq}(C\cap\Lambda X,\Lambda X)'\overset{(iv)}{\simeq}(C,Y)'/(\Lambda X,Y)'\overset{(v)}{\simeq}\Lambda'C',$$

((i): by the inclusion $\ker\Lambda\subseteq C\Lambda$; (ii): by statement 2 of theorem 1; (iii): apply the operator Λ and use the definitions; (iv): by corollary 1; (v): apply the operator Λ' and use the definitions.)

■

4 Duality theorem and calculus rules for convex objects

Plan. In this section we will do for all types of convex objects what we have done for *cones* in the previous section: we define an operator of duality, we present a duality theorem, and the calculus rules for the image and the inverse image and for all binary operations. This involves a lot of technical looking definitions. In the next section, we will see that *everything including these definitions flows naturally from the cone approach.*

We begin by enumerating the definitions of the set of regular objects $\mathcal{O}(X)^r$ for each type of convex objects $\mathcal{O}$ for which some elements have to be excluded in the definition of the operator of duality.

- $M\in \mathrm{Aff}(X)^r\Leftrightarrow 0\not\in M$,
- $A\in \mathrm{Co}(X)^r\Leftrightarrow A\neq\emptyset$,
- $f\in \mathrm{Co}^f(X)^r\Leftrightarrow f(x)>-\infty\ \forall x\in X$, and $\exists x\in X: f(x)<\infty$ (f is *proper*),
- $p\in \mathrm{Cone}^f(X)^r\Leftrightarrow p(x)>-\infty\ \forall x\in X$.

Then we enumerate the operators of duality for each type of convex objects.

- *annihilator* of $L\in \mathrm{Lin}(X)$:

$$L^\perp=\{x'\in X':\langle x',x\rangle=0\ \forall x\in L\}\in \mathrm{Lin}(X');$$

- *dual* of $M\in \mathrm{Aff}(X)^r$:

$$M^\bullet=\{x'\in X':\langle x',x\rangle=1\ \forall x\in M\}\in \mathrm{Aff}(X')^r;$$

- *support function* of $A \in \mathrm{Co}(X)^r$:

$$(x' \mapsto sA(x') = \sup\{\langle x', x\rangle : x \in A\}) \in \mathrm{Cone}^f(X');$$

- *polar* of $B \in \mathrm{Co}_0(X)$:

$$B^\circ = \{x' \in X' : \langle x', x\rangle \le 1 \ \forall x \in B\} \in \mathrm{Co}_0(X');$$

- *conjugate function* or *transformation of Young-Fenchel* of $f \in \mathrm{Co}^f(X)^r$:

$$(x' \mapsto f^*(x') = \sup_{x\in X}(\langle x', x\rangle - f(x))) \in \mathrm{Co}^f(X')^r;$$

- *subdifferential* of $p \in \mathrm{Cone}^f(X)^r$:

$$\partial p = \{x' : \langle x', x\rangle \le p(x) \ \forall x \in X\} \in \mathrm{Co}(X')^r;$$

- *dual norm* of $N \in \mathrm{Norm}(X)$:

$$(x' \mapsto N^*(x') = \sup_{N(x)=1} \langle x', x\rangle) \in \mathrm{Norm}(X').$$

Closed convex objects. The collection of *closed* convex objects of some type will be denoted by the prefix Cl: for example we will write $\mathrm{ClCo}(X)$ for the collection of closed convex sets in X.

Theorem 3 ***Duality theorem and calculus rules for convex objects.***

We enumerate the duality theorem and the calculus rules for all types of convex objects.

1. *Linear subspaces:*

   (a) $L^{\perp\perp} = L \Leftrightarrow L \in \mathrm{Lin}(X)$,
   (b) $(L_1 + L_2)^\perp = L_1^\perp \cap L_2^\perp$,
   (c) $(L_1 \cap L_2)^\perp = L_1^\perp + L_2^\perp$,
   (d) $(\Lambda L)^\perp = L^\perp \Lambda'$,
   (e) $(L\Lambda)^\perp = \Lambda' L^\perp$;

2. *Affine subspaces:*

   (a) $M^{\bullet\bullet} = M \Leftrightarrow M \in \mathrm{Aff}(X)^r$,
   (b) $(M_1 \mathrm{co}\bigcup M_2)^\bullet = M_1^\bullet \cap M_2^\bullet$,
   (c) $(M_1 \cap M_2)^\bullet = M_1^\bullet \mathrm{co}\bigcup M_2^\bullet$ *if* $M_1 \cap M_2 \neq \emptyset$,
   (d) $(\Lambda M)^\bullet = M^\bullet \Lambda'$,
   (e) $(M\Lambda)^\bullet = \Lambda' M^\bullet$ *if* $M \cap \Lambda X \neq \emptyset$;

3. *Convex sets:*

   (a) $\partial s A = A \Leftrightarrow A \in \mathrm{ClCo}(X)^r$,
   (b) $s(A_1 \mathrm{co} \bigcup A_2) = sA_1 \wedge sA_2$,
   (c) $s(A_1 \cap A_2) = sA_1 \mathrm{co} \vee sA_2$ *if* $\mathrm{ri} A_1 \cap \mathrm{ri} A_2 \neq \emptyset$,
   (d) $s(\Lambda A) = sA\Lambda'$,
   (e) $s(A\Lambda) = \Lambda' sA$ *if* $\mathrm{ri} A \cap \Lambda X \neq \emptyset$;

4. *Cones:*

   (a) $C'' = C \Leftrightarrow C \in \mathrm{ClCone}(X)^r$,
   (b) $(C_1 + C_2)' = C_1' \cap C_2'$,
   (c) $(C_1 \cap C_2)' = C_1' + C_2'$ *if* $\mathrm{ri} C_1 \cap \mathrm{ri} C_2 \neq \emptyset$,
   (d) $(\Lambda C)' = C'\Lambda'$,
   (e) $(C\Lambda)' = \Lambda' C'$ *if* $\mathrm{ri} C \cap \Lambda X \neq \emptyset$;

5. *Convex zero-sets:*

   (a) $B^{\circ\circ} = B \Leftrightarrow B \in \mathrm{ClCo}_0(X)$,
   (b) $(B_1 \mathrm{co} \cup B_2)^\circ = B_1^\circ \cap B_2^\circ$,
   (c) $(B_1 \cap B_2)^\circ = B_1^\circ \mathrm{co} \cup B_2^\circ$ *if* $\mathrm{ri} B_1 \cap \mathrm{ri} B_2 \neq \emptyset$,
   (d) $(\Lambda B)^\circ = B^\circ \Lambda'$,
   (e) $(B\Lambda)^\circ = \Lambda' B^\circ$ *if* $\mathrm{ri} C \cap \Lambda X \neq \emptyset$;

6. *Convex functions:*

   (a) $f^{**} = f \Leftrightarrow f \in \mathrm{ClCo}^f(X)^r$,
   (b) $(f_1 \mathrm{co} \vee f_2)^* = f_1^* \wedge f_2^*$,
   (c) $(f_1 \wedge f_2)^* = f_1^* \mathrm{co} \vee f_2^*$ *if* $\mathrm{ridom} f_1 \cap \mathrm{ridom} f_1 \neq \emptyset$,
   (d) $(\Lambda f)^* = f^* \Lambda'$,
   (e) $(f\Lambda)^* = \Lambda' f^*$ *if* $\mathrm{ri}\,\mathrm{dom} f \cap \Lambda X \neq \emptyset$;

7. *Sublinear functions:*

   (a) $s\partial p = p \Leftrightarrow p \in \mathrm{ClCone}^f(X)^r$,
   (b) $\partial(p_1 \mathrm{co} \vee p_2) = \partial p_1 \cap \partial p_2$,
   (c) $\partial(p_1 \wedge p_2) = \partial p_1 \mathrm{co} \vee p_2$ *if* $\mathrm{ridom} p_1 \cap \mathrm{ridom} p_1 \neq \emptyset$,
   (d) $\partial(p\Lambda) = \Lambda' \partial p$,
   (e) $\partial(\Lambda p) = \partial p \Lambda'$, *if* $\mathrm{ri}\,\mathrm{dom} p \cap \Lambda X \neq \emptyset$;

8. *Norms:*

   (a) $N^{**} = N \Leftrightarrow N \in \mathrm{Norm}(X)$,
   (b) $(N_1 \mathrm{co} \vee N_2)^* = N_1^* \wedge N_2^*$,
   (c) $(N_1 \wedge N_2)^* = N_1^* \mathrm{co} \vee N_2^*$,
   (d) $(\Lambda N)^* = N^* \Lambda'$,
   (e) $(N\Lambda)^* = \Lambda' N^*$.

Three statements of this list have names: $5(a)$ is called *the theorem on the bipolar*, $6(c)$ is called *the theorem of Fenchel-Moreau* and $7(c)$ is called *the theorem of Dubovitskii-Milyutin*.

We emphasize again that the theorems of duality reflect the possibility of a dual (geometrical and algebraical) description. For example:

- a closed convex set A is the closed convex hull of a family of points, and, from the other side, the solution set of a family of nonhomogeneous linear inequalities in x of the form $\langle x', x\rangle \leq \alpha$.

- a convex function has a convex set as epigraph, and, from the other side, it is the upperbound of a family of affine functions.

Finally we note that the theorems of duality implies non-triviality of the image of convex objects under the operator of duality. For example, 'the polar of a convex zero-set $\neq X$ contains nonzero elements', and 'the subdifferential of a sublinear function is nonempty'.

5 Unified approach to the duality theorem and the calculus rules for convex objects

We will present a simple unified 'cone approach' to the relatively vast and complicated looking subject that has been presented in the last section. The main point is to derive all results of convex analysis from one result, theorem 2, which is possible by virtue of the following remarkable phenomenon:

Each result of the convex analysis can be reformulated as the statement that the conjugate of a suitable cone contains a nonzero element.

All results in this paper are examples of this phenomenon. For an explicit illustration see the proof of theorem 9 (Farkas' lemma).

To begin with, we will sketch this approach. For each type of convex objects $\mathcal{O}$ and each vector space X, we will proceed in the following systematic way.

1. We show how to turn an element $O \in \mathcal{O}(X)$ into a cone $K(O)$ in a suitable vector space, the *associated cone of* O, *without loss of information*.

2. We will calculate for all $O_1, O_2 \in \mathcal{O}$, the cones $K(O_1) + K(O_2)$ and $K(O_1) \cap K(O_2)$, and this will lead to the formulas
$$K(O_1) + K(O_2) = K(O_1 \boxplus O_2)$$

and

$$K(O_1) \cap K(O_2) = K(O_1 \Cap O_2)$$

for suitable binary operations $\boxplus$ ('addition-like') and $\Cap$ ('intersection-like'). The binary operations that arise in this way will in each case turn out to be well-known ones.

3. We will reformulate the condition $\mathrm{ri}K(O_1) \cap \mathrm{ri}K(O_2) \neq \emptyset$ in terms of O_1 and O_2.

4. We will calculate the conjugate cone $K(O)'$ and this will lead to the following result. For each type $\mathcal{O}$ there is a type $\tilde{\mathcal{O}}$, such that for all—or almost all– $O \in \mathcal{O}(X)$, there exists a—necessarily unique—element $D(O) \in \tilde{\mathcal{O}}(X')$ for which

$$K(O)' = K(D(O)). \qquad (*)$$

Let $\mathcal{O}(X)^r$ denote the set of $O \in \mathcal{O}(X)$ for which $D(O)$ is defined. It will be called the set of *regular elements*. It turns out that for each regular $O \in \mathcal{O}(X)$ its dual object $D(O) \in \tilde{\mathcal{O}}(X')$ is regular as well. That is, we get an operator of duality

$$D : \mathcal{O}(X)^r \to \tilde{\mathcal{O}}(X')^r.$$

This operator will in each case turn out to be a well-known one.

5. The statements of the duality theorem

$$D(D(O)) = O,$$

for closed objects O, and the rules

$$D(O_1 \boxplus O_2) = D(O_1) \Cap D(O_2),$$

and

$$D(O_1 \Cap O_2) = D(O_1) \boxplus D(O_2),$$

under suitable assumptions on O_1 and O_2, will then follow immediately from theorem 1, as well as rules for the dual of image and inverse image of a convex object under a linear operator. We emphasize that in this way the duality theorems and the basic rules for each type of convex objects are revealed to have one common source, theorem 2.

Associated cone. The main ideas of turning a convex object into a cone, called its *associated cone* are as follows (the details are given in theorem 4). It suffices to show how to turn a convex *set* into a cone: a convex *function* can be turned into a convex set without loss of information by virtue of the concept epigraph, and a convex *optimization problem* is given by a convex function. The geometrical idea of turning a convex set $A \in \mathrm{Co}(X)$ into a cone is illustrated in Figure 2: add a 'vertical' dimension to the space X, lift up $A_0 = A \times 0$ to $A_1 = A \times 1$, that is, to height one, and then take the cone $K(A) \in \mathrm{Cone}(X \times \mathbb{R})$ consisting of the rays from the origin through a point of A_1.

From an analytical point of view, this construction is also very natural. It is essentially *homogenization*: for example, homogenizing the polynomial $f(x_1, x_2) = x_2 - x_1^2$ gives

$$g(x_1, x_2, x_3) = x_3^2 f(\frac{x_1}{x_3}, \frac{x_2}{x_3}) = x_2 x_3 - x_1^2,$$

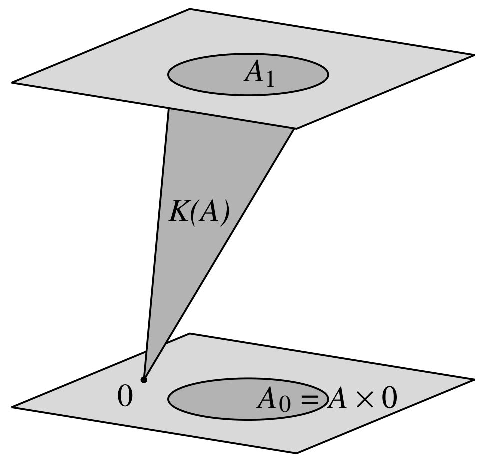

Figure 2: Associating a convex cone $K(A)$ to a convex set A.

and the associated cone of the convex set $A = \{x \in \mathbb{R}^2 : x_2 - x_1^2 \geq 0\}$ is the cone

$$K(A) = \{x \in \mathbb{R}^3 : x_2x_3 - x_1^2,\ x_3 \geq 0\}.$$

For some types of convex objects, one can replace this way of associating cones to convex sets and functions by a simpler one, and then we will do so: for example, the simplest way to turn a sublinear function into a cone is by taking its epigraph, which is a cone.

Need to enumerate. In order to display the results of the calculations, we have to enumerate, for all types of convex objects $\mathcal{O}$, the definitions of the following concepts: the set of regular elements $\mathcal{O}(X)^r$, the operator of duality $D : \mathcal{O}(X)^r \to \tilde{\mathcal{O}}(X')^r$ and the binary operations $\boxplus$ and $\Cap$. We will use the notation and terminology for the operators of duality and the binary operations which arise from the sum and intersection of associated cones that are well-established in the literature, instead of the notation D, $\boxplus$ and $\Cap$. This will lead to the appearance of minus-signs in some of the formulas in theorem 4. The main point of the following theorem is that all the complicated looking definitions of the dual operators and the main binary operations for convex objects arise naturally from the cone approach.

We let i denote the involution on X defined by $i(x) = -x\ \forall x \in X$, and j the involution on $X \times \mathbb{R}^2$ defined by $j(x, \alpha, \beta) = (x, \beta, \alpha)\ \forall x \in X, \forall \alpha, \beta \in \mathbb{R}$.

Theorem 4 ***The operator of duality and the two basic binary operations for associated cones of convex objects.*** *We enumerate, for each type of convex objects, the definition of the associated cone and the result of the calculation of the following cones:*

- *the conjugate of the associated cone,*
- *the sum of two associated cones,*
- *the intersection of two associated cones.*

1. *Linear subspaces:*

   (a) $K(L) = L$

   (b) $K(L)' = L^{\perp}$,

   (c) $K(L_1) + K(L_2) = K(L_1 + L_2)$,

   (d) $K(L_1) \cap K(L_2) = K(L_1 \cap L_2)$;

2. *Affine subspaces:*

   (a) $K(M) = \mathrm{sp}(M \times 1)$,

   (b) $K(M)' = K((-M)^{\bullet})$,

   (c) $K(M_1) + K(M_2) = K(M_1 \mathrm{co} \bigcup M_2)$,

   (d) $K(M_1) \cap K(M_2) = K(M_1 \cap M_2)$;

3. *Convex sets:*

   (a) $K(A) = \mathrm{cone}(A \times 1)$,

   (b) $K(A)' = K(s(-A))$,

   (c) $K(A_1) + K(A_2) = K(A_1 \mathrm{co} \bigcup A_2)$,

   (d) $K(A_1) \cap K(A_2) = K(A_1 \cap A_2)$;

4. *Convex zero-sets*

   (a) $K(B) = \mathrm{cone}(B \times 1)$,

   (b) $K(B)' = K((-B)^{\circ})$,

   (c) $K(B_1) + K(B_2) = K(B_1 \mathrm{co} \bigcup B_2)$,

   (d) $K(B_1) \cap K(B_2) = K(B_1 \cap B_2)$;

5. *Linear functions*

   (a) $K(l) = \mathrm{graph}\, l$,

   (b) $K(l)' = l \in X'$;

6. *Affine functions*

   (a) $K(a) = \mathrm{cone}((\mathrm{graph}\, a) \times 1)$,

   (b) $K(a)' = (a - a(0), a(0)) \in X' \times \mathbb{R}$;

7. *Convex functions*

   (a) $K(f) = \mathrm{cone}((\mathrm{epi} f) \times 1)$,

   (b) $K(f)' = j(K(f^* \circ i))$,

   (c) $K(f_1) + K(f_2) = K(f_1 \mathrm{co} \vee f_2)$,

   (d) $K(f_1) \cap K(f_2) = K(f_1 \wedge f_2)$;

8. *Sublinear functions*

   (a) $K(p) = \text{epi}\, p$,

   (b) $K(p)' = K(-\partial p)$,

   (c) $K(p_1) + K(p_2) = K(p_1 \text{co} \vee p_2)$,

   (d) $K(f_1) \cap K(f_2) = K(f_1 \wedge f_2)$;

9. *Norms*

   (a) $K(N) = \text{epi}\, N$,

   (b) $K(N)' = K(N^* \circ i)$,

   (c) $K(N_1) + K(N_2) = K(N_1 \text{co} \vee N_2)$,

   (d) $K(N_1) \cap K(N_2) = K(N_1 \wedge N_2)$.

This result, in combination with theorem 1, will give the duality theorem and the basic rules for convex objects.

Completion of a type of convex objects $\mathcal{O}(X)$. Now we present a solution for the technical problem considered above, which is the reason that $D(O)$ can only be defined for *regular* $O \in \mathcal{O}(X)$. This solution is based on the following fact:

it turns out that for each $C \in \text{Cl}K(\mathcal{O}(X))$, *one has that* $C' \in \text{Cl}K(\tilde{\mathcal{O}}(X'))$.

This suggests to 'complete' each type of convex objects $\mathcal{O}(X)$ to $\mathcal{O}(X)^c$ by 'taking the closure': to be more precise, we go over from $K(\mathcal{O}(X))$ (resp. $K(\tilde{\mathcal{O}}(X')))$ to $\text{cl}K(\mathcal{O}(X))$ (resp. $\text{cl}K(\tilde{\mathcal{O}}(X')))$. Then the conjugate cone operator gives an operator of duality

$$D : \mathcal{O}(X)^c \to \tilde{\mathcal{O}}(X)^c,$$

which extends the operator $D : \mathcal{O}(X)^r \to \tilde{\mathcal{O}}(X)^r$, defined above. If we carry out this strategy for a concrete type $\mathcal{O}$, then calculations lead to explicit descriptions for the added elements and for their dual objects. We will restrict attention here to the following characteristic examples.

1. We enumerate the outcomes of calculating the conjugate of the cone associated to a nonregular convex object:

   - $K(M)' = M^\perp \times 0$ for all nonregular $M \in \text{Aff}(X)$;
   - $K(A)' = X' \times \mathbb{R}$ for the nonregular $A \in \text{Co}(X)$, that is, for $A = \emptyset$;
   - for all nonregular $f \in \text{Co}^f(X)$:

     $K(f)' = j(s(-\tilde{A}) \times 0)$ if $\tilde{A} = \{x : f(x) = -\infty\} \neq \emptyset$,

     $K(f)' = X' \times \mathbb{R}^2$ if $f \equiv \infty$;

- $K(p)' = \tilde{C}' \times 0$ for all nonregular $p \in \text{Cone}^f(X)$, where $\tilde{C} = \{x : p(x) = -\infty\}$.

2. We give, for the two most interesting cases, the outcome of calculating which cones are added to the set $K(\mathcal{O}(X))$ if we take the closure.

   - The cones $K(M)$ with $M \in \text{Aff}(X)$ are precisely the linear subspaces of $X \times \mathbb{R}$ that are 'not horizontal', that is, not contained in $X \times 0$. Therefore, we have to add the linear subspaces of $X \times 0$, the 'horizontal' linear subspaces in $X \times \mathbb{R}$. This is precisely the well-known construction of going over from the *affine space* to the *projective space*: the added objects are precisely the affine subspaces 'at infinity'.
   - The cones $K(A)$ with $A \in \text{Co}(X)$ are precisely the cones that are contained in the upper half space $X \times \mathbb{R}_+$ and that intersect the horizontal hyperplane $X \times 0$ only at zero. Therefore, we have to add the cones in $X \times \mathbb{R}$ that are contained in $X \times \mathbb{R}_+$ and that have at least one ray in common with the horizontal hyperplane $X \times 0$. These rays have a concrete interpretation: if $A \in \text{Co}(X)$ then the intersection of $\text{cl}K(A)$ and $X \times 0$ consists precisely of the *rays of recession* of the convex set A (half-lines that are contained in A).

Remark 6. We have turned all convex objects into cones. Here we give three other constructions for turning one type of convex objects into another:

- the epigraph operator $\text{epi} : \text{Co}^f(X) \to \text{Co}(X \times \mathbb{R})$ turns convex functions into convex sets.
- the indicator operator $\delta : \text{Co}(X) \to \text{Co}^f(X)$, defined by $\delta A(x) = 0$ if $x \in A$ and $= +\infty$ if $x \notin A$. The function $x \mapsto \delta A(x)$ is called the *indicator* of A, or the *indicator function* of A. The indicator operator turns convex sets into convex functions.
- the Minkowski operator $\mu : \text{Co}_0(X) \to \text{Cone}^f(X)$,

  $$\mu B(x) = \inf\{\alpha : \alpha^{-1}x \in B\}.$$

  The function $x \mapsto \mu B(x)$ is called the *function of Minkowski* of B. This operator turns zero-convex sets into sublinear functions. For example, each norm N on X is the image under the Minkowski operator of the open (or the closed) unit-ball in X with respect to the norm N.

Locally convex spaces. We have chosen to present the theorem of duality and the basic rules of convex objects in the context of vector spaces, without equipping these with a topology. Alternatively, one can give the theory for *locally convex topological spaces* X and their topological conjugate spaces X^* (that is, X^* is the collection of all *continuous* linear functionals on X). For the purposes of functional analysis it is required to equip vector spaces with a topology that makes it into a locally convex topological space. We make the point that this additional structure is itself a convex object. Indeed, one can prove that such a topology can be described as the weakest topology for which a certain family of *semi-norms* is continuous. We recall that a semi-norm is defined to be a sublinear function that takes nonnegative real values on the whole space X.

The infinite-dimensional case. All results given above extend to infinite dimensional vector spaces (without topology). Here we indicate how to do this. We have to restrict all results to convex sets that contain internal points and convex functions for which the epigraph contains internal points (in the finite dimensional case these asumptions can be omitted: they are automatically satisfied).

The concept 'internal point' is a substitute for the concept 'interior point' in the absence of a topology. Its definition is that a point s of a set S in a vector space X is an internal point if for each nonzero vector $v \in X$, the interval $[s, s + \frac{1}{n}v]$ is contained in S for a suitable $n \in \mathbb{N}$. In the proof of theorem 2, one has to use Zorn's lemma to show the existence of a linear subspace of X that is disjoint from the set of internal points of the cone C. Moreover, one has to use the concept 'algebraic closed set' as a substitute for the concept 'closed set'. Its definition is that a set S in X is algebraically closed if all points of the complement $X \setminus S$ are internal points of this complement.

6 Subdifferentials of convex functions

The aim of this section is to consider the precise convex analogue of the derivative of a differentiale function at a point: *the subdifferential* of a convex function at an internal point of its domain. The *domain* of a convex function $f : X \to \mathbb{R} \cup \{+\infty\}$ is defined to be

$$\mathrm{dom} f = \{x \in X : f(x) < +\infty\}.$$

The *subdifferential* of $f \in \mathrm{Co}^f(X)$ at an interior point $\widehat{x} \in X$ of the domain of f, is defined to be the following convex set

$$\partial f(\widehat{x}) = \{x' \in X' : f(\widehat{x}) + \langle x', h\rangle \leq f(\widehat{x} + h)\, \forall h \in X\},$$

in X'. The elements of this set are called the *subgradients* of f at $\widehat{x}$.

One can reduce subdifferentials of functions at points to subdifferentials of sublinear functions by means of *'localization of f at $\widehat{x}$'*. Indeed, for each pair $(f, \widehat{x})$, consisting of a convex function $f \in \mathrm{Co}^f(X)$ and a point $\widehat{x} \in \mathrm{int\,dom} f$, one can consider the sublinear function $p \in \mathrm{Cone}^f(X)$ for which the epigraph equals the closure of the conic hull of the set $\{v - (\widehat{x}, f(\widehat{x}) : v \in \mathrm{epi} f\}$. The following equality holds true:

$$\partial f(\widehat{x}) = \partial p.$$

Therefore, one can apply all the results on subdifferentials of sublinear functions to subdifferentials of convex functions at points. In particular, $\partial f(\widehat{x})$ is non-empty.

Theorem 5 ***Rules for the subdifferential.*** *Let $f_i : X \to \mathbb{R} \cup \{+\infty\},\ i = 1, 2$ be convex functions and let $\widehat{x} \in X$ be an interior point of the domain of f. We enumerate the calculus rules:*

1. $\partial(f_1 \mathrm{co} \vee f_2)(\widehat{x}) = \partial f_1(\widehat{x}) \cap \partial f_2(\widehat{x})$,
2. $\partial(f_1 \wedge f_2)(\widehat{x}) = \partial f_1(\widehat{x}) \mathrm{co} \vee f_2(\widehat{x})$,
3. $\partial(f\Lambda)(\widehat{x}) = \Lambda' \partial f(\widehat{x})$,
4. $\partial(\Lambda f)(\widehat{x}) = \partial f(\widehat{x})\Lambda'$, *if* $\mathrm{ri\,im} f \cap \Lambda X \neq \emptyset$.

7 Duality for polars of convex sets that do not contain the origin: looking at infinity and beyond.

In this section we show how to use the polar operator to describe the duality property of convex sets that do not necessarily contain zero.

The concept polar can be extended from convex zero-sets to arbitrary convex sets, by using the same definition: the polar set of $A \in \mathrm{Co}(X)$ is the following convex set $A^\circ \in \mathrm{Co}(X')$,

$$A^\circ = \{y \in X' : \langle x, y\rangle \leq 1 \ \forall x \in A\}.$$

The problem is that the formula $A^{\circ\circ} = \mathrm{cl}A$, which expresses the duality property if A contains zero, does not hold if A does not contain zero. The following example shows that the polar of a convex set that does not contain zero, need not even be a convex set.

Example 1 *Let A be the disk with center $(1,1)$ and radius 1. Then A° is the union of the part of the plane lying under the lower branch of the hyperbola $2(y_1+1)(y_2+1) = 1$ and the part of the plane consisting of the points between the upper branch of this hyperbola and the line $y_1 + y_2 = 1$.*

Solution problem (informal description). Again, the cone approach will clarify the situation completely. We begin with an informal discussion of the situation. To a set $A \in \mathrm{Co}(X)$, we associate the cone $K(A) = \mathrm{cone}(A \times 1)$ in $X \times \mathbb{R}$. The intersection of the conjugate cone $K(A)'$ with the horizontal hyperplane $X \times 1$ equals $(-A)^\circ \times 1$, by the definition of the polar operator. We observe that the conjugate cone $K(A)'$ is contained in the upper half space $X \times \mathbb{R}_+$, provided $0 \in A$. However, if $0 \notin A$, then this is not the case, and so we lose information if we intersect the conjugate cone $K(A)'$ with $X \times 1$. What we should do, in order to prevent loss of information, is to take the intersection with the horizontal hyperplane $X \times -1$ as well, and maybe also the intersection with $X \times 0$. Thus the *complete polar* of A consists of a collection of three convex sets in X, which are obtained by taking the intersection of $K(A)'$ with the three horizontal hyperplanes at levels 1, 0 and -1:

1. The first one is the usual polar.
2. The second one can be viewed as the part of the complete polar that lies *at infinity*.
3. The third one can be viewed as the part of the complete polar that lies *beyond infinity*.

Thus *the reason of the nonvalidity of the duality property for the polar operation on convex sets that do not contain zero is revealed. It is that the usual polar ignores the part of the complete polar lying at infinity and beyond*.

Solution problem (formal description). Formally, we define, for each convex set $A \in \mathrm{Co}(X)$, its *antipolar* ('polar beyond infinity'), by $A^\diamond = \{x' \in X' : \langle x', x\rangle \leq -1 \ \forall x \in A\}$ and its *polar at infinity*, by $A^\infty = \{x' \in X' : \langle x', x\rangle = 0\}$. In order to get a good duality theory, together with 'rules', one should start, instead of with convex sets $A \in \mathrm{Co}(X)$, with triples (A_+, A_0, A_-) of convex sets in X for which there exists a cone $C \in X \times \mathbb{R}$ for which

- $C \cap (X \times 1) = A_+$,
- $C \cap (X \times 0) = A_0$,
- $C \cap (X \times -1) = A_-$.

We will not display the details here.

8 Standard results on convex objects

In this section we illustrate how the standard results on convex objects can be derived from theorem 2.

The best-known and most-used results on convex sets are *the separation theorems*.

Two sets $A,\ B \subseteq X$ can be **separated** if there exists an affine hyperplane

$$M = \{x \in X : \langle x', x\rangle = \alpha\}$$

—with $x' \in X' \setminus \{0\}$ and $\alpha \in \mathbb{R}$—such that A and B lie on different sides of M.

Theorem 6 ***Separation of convex sets.*** *Two convex sets $A, B \in \mathrm{Co}(X)$ with $\mathrm{ri}A \cap \mathrm{ri}B = \emptyset$ can be separated.*

Proof. Choose $\bar{a} \in \mathrm{ri}A$ and $\bar{b} \in \mathrm{ri}B$. Consider the cone

$$C = \{\rho(a,1) - \sigma(b,1) : a \in A,\ b \in B,\ \rho, \sigma \in \mathbb{R}_+\}$$

in the space $X \times \mathbb{R}$. We want to apply theorem 2 to the cone C. To this end, we check that the assumptions of theorem 2 hold for C. Observe that $(\bar{a},1)-(\bar{b},1) \in \mathrm{ri}C \times \mathbb{R}$, and that $-((\bar{a},1)-(\bar{b},1)) \notin C$, as the equality

$$-((\bar{a},1) - (\bar{b},1)) = \rho(a,1) - \sigma(b,1)$$

with $a \in A, b \in B, \rho, \sigma \in \mathbb{R}_+$ would lead to the equality

$$(\bar{a} + \rho a, 1 + \rho) = (\bar{b} + \sigma b, 1 + \sigma)$$

and so to a contradiction with the assumption $\mathrm{ri}A \cap \mathrm{ri}B = \emptyset$. Therefore, theorem 2 can be applied to the cone C. It follows that $C' \neq 0$. Explicitly, there exists a nonzero $y' \in (X \times \mathbb{R})'$ for which

$$y'(\rho(a,1) - \sigma(b,1)) \geq 0$$

for all $a \in A, b \in B, \rho, \sigma \in \mathbb{R}_+$. This shows that A and B can be separated by the hyperplane $\{x \in X : \langle x', x\rangle = \gamma\}$, where $x' \in X'$ is defined by

$$\langle x', x\rangle = \langle y', (x, 0)\rangle\ \forall x \in X$$

and where γ is chosen to be a number from the interval

$$[\sup_{b\in B} x'(b), \inf_{a \in A} x'(a)].$$

■

Two sets $A, B \subseteq X$ can be **strictly separated** if there exist two parallel hyperplanes in X that separate A and B.

Theorem 7 ***Strict separation of convex sets.*** *Two convex sets $A, B \in \mathrm{Co}(X)$ can be strictly separated, if there exists $C \in \mathrm{Co}(X)$ with $0 \in \mathrm{int}C$, for which $C + \mathrm{ri}A$ is disjoint from $\mathrm{ri}B$.*

Proof. By the previous theorem, the sets $C + A$ and B can be separated. That is, there exists a nonzero $x' \in X'$ for which $\langle x', c+a\rangle \geq \langle x', b\rangle$ for all $a \in A,\ b \in B,\ c \in C$. As $x' \neq 0$ and as $0 \in \text{int}C$, it follows that there exists $\bar{c} \in C$ with $\langle x', \bar{c}\rangle < 0$. It follows that

$$\inf_{a\in A}\langle x', a\rangle > \sup_{b\in B}\langle x', b\rangle.$$

This shows that A and B can be strictly separated.

■

One of the central results of functional analysis is the following theorem (to be more precise, here we only consider the finite dimensional case).

Theorem 8 ***Hahn-Banach theorem.*** *Let $p : X \to \mathbb{R}$ be a sublinear function on a vector space X, and let $l : L \to \mathbb{R}$ be a linear functional on a linear subspace L in X such that*

$$\langle l, x\rangle \leq p(x)\ \forall x \in L.$$

Then there exists a linear functional $\Lambda : X \to \mathbb{R}$ on the whole space X that is an extension of l, that is,

$$\langle \Lambda, x\rangle = \langle l, x\rangle\ \forall x \in L,$$

and satisfies the inequality

$$\Lambda, x\rangle \leq p(x)\ \forall x \in X.$$

Proof. Apply corollary 1 (of theorem 2) with

$$X := X \times \mathbb{R},\ C := \text{epi}p,\ L := L \times \mathbb{R}.$$

Note that the assumption $\text{ri}C \cap L \neq \emptyset$ is satisfied, as $(0, 1) \in X \times \mathbb{R}$ is an interior point of epi$p$ as well as a point of L. Therefore, the conclusion of the corollary is seen to be precisely the conclusion of the Hahn-Banach theorem. ■

The following result is the central property of systems of linear inequalities. It shows that the *non-solvability* of such a system can in principle be demonstrated by producing a solution of a suitable other system. The proof below is an explicit illustration of the general principle that all results of convex analysis can be formulated as the statement that the conjugate of a suitable cone contains a nonzero element. We return to systems of linear inequalities in section 10: there we will prove the Farkas' lemma again, this time in order to illustrate the use of the rules of the convex calculus.

Theorem 9 ***Farkas' lemma.*** *The system of inequalities*

$$Ax \leq 0_m, \quad c^T \cdot x < 0$$

where A is an $m \times n$-matrix and $c \in \mathbb{R}^n$, and the system of inequalities

$$A^T y + c = 0_n, \quad y \geq 0_m,$$

are strong alternatives. That is, precisely one of the two is solvable.

Proof. We begin by rewriting the second alternative in the form

$$C' \neq 0$$

for a suitable cone C. We choose a sufficiently small $\varepsilon > 0$.

$$\begin{pmatrix} y \\ y_0 \end{pmatrix}^T \cdot \begin{pmatrix} Ax \\ c^T \cdot x \end{pmatrix} = 0, \ \forall x \in \mathbb{R}^n,$$

$$\begin{pmatrix} y \\ y_0 \end{pmatrix}^T \cdot z \geq 0, \ \forall z \in \mathbb{R}^{m+1}_+,$$

$$\begin{pmatrix} y \\ y_0 \end{pmatrix}^T \cdot u \geq 0,$$

for all $u \in \mathbb{R}^{m+1}$ that make an angle with

$$\begin{pmatrix} 0_m \\ 1 \end{pmatrix}$$

that is not larger than ε. To see that this is equivalent to the second alternative, note that the last condition is equivalent to the positivity of y_0. Then one can de-homogenize the system of linear inequalities, taking $y_0 = 1$. This gives precisely the system of the second alternative. Now we choose C to be the convex cone spanned by all vectors $\begin{pmatrix} Ax \\ c^T \cdot x \end{pmatrix}$ with $x \in \mathbb{R}^n$, all vectors from $\mathbb{R}^{m+1}_+$, and all vectors $u \in \mathbb{R}^{m+1}$ that make an angle with

$$\begin{pmatrix} 0_m \\ 1 \end{pmatrix}$$

that is not larger than ε. To prove Farkas' lemma, it remains to write out the condition $C \neq \mathbb{R}^{m+1}$, that is, as

$$\begin{pmatrix} 0_m \\ 1 \end{pmatrix}$$

is an interior point of C, the condition that minus this point does not belong to K. This gives the condition that the system

$$Ax \leq 0_m, \ c^T \cdot x \leq -1$$

has a solution. This is seen to be equivalent to the first alternative. ■

9 Duality and calculation of subdifferentials and multiplier methods

The basis of the use of convex analysis for the solution of unconstrained convex optimization problems is the following result.

Theorem 10 ***Convex Fermat theorem.*** *Let X be a vector space, $f : X \to \mathbb{R} \cup \{+\infty\}$ a convex function, and $\widehat{x}$ an internal point of the domain of f.*

Then each solution $\widehat{x}$ of the unconstrained convex optimization problem

$$f(x) \to \min, \ x \in X,$$

satisfies the following inclusion

$$0 \in \partial f(\widehat{x}).$$

Conversely, this condition is sufficient for global optimality of $\widehat{x}$*.*

Parallel with differentiable Fermat theorem. This result is a tautology. However, in combination with the calculus rules for subdifferentials of convex functions, it leads to an effective method for solving unconstrained convex optimization problems. This runs parallel to the method for solving unconstrained differentiable optimization problems by the differentiable Fermat theorem ('in optimum derivative is zero') in combination with the calculus rules for differentiation.

A convenient way to model a *constrained* convex optimization problem is as follows

$$f(x) = \langle x', x \rangle \to \min,\ x \in M,\ x \in C, \tag{P}$$

where $x' \in X',\ M \in \mathrm{Aff}(X)$, defined by

$$M = \{x \in X : \Lambda x = \bar{y}\}$$

—for a vector space Y, a linear operator $\Lambda : X \to Y$ and an element $\bar{y} \in Y$—and $C \in \mathrm{Cone}(X)$.

Example 2 *The linear programming problem in standard form is*

$$c^T x \to \min,\ Ax = b,\ x \geq 0,$$

where $c \in \mathbb{R}^n$*,* A *is an* $m \times n$*-matrix,* $b \in \mathbb{R}^m$ *and* x *is a variable vector in* $\mathbb{R}^n$*. This is the following special case of the problem* (P)*:*

$$X := \mathbb{R}^n,\ x' := x \mapsto c^T x,\ M := \{x \in \mathbb{R}^n : Ax = b\},\ C := \{x \in \mathbb{R}^n : x \geq 0\}.$$

The basis of the use of convex analysis for the solution of *constrained* optimization problems is the following result. We need some definitions. A *selection of Lagrange multipliers* is defined to be a nonzero vector $\lambda = (\mu_0, \mu)$ with $\mu_0 \in \mathbb{R}_+$ and $\mu \in Y'$. The *Lagrange function* is defined to be the function

$$\mathcal{L}(x, \lambda) = \mu_0 \langle x', x \rangle + \langle \mu, \Lambda x - \bar{y} \rangle.$$

Theorem 11 ***Convex Lagrange multiplier rule.*** *Each solution* $\widehat{x}$ *of the constrained convex optimization problem* (P) *satisfies the following inclusion*

$$0 \in \partial \mathcal{L}(\widehat{x}, \lambda)$$

for a suitable selection of Lagrange multipliers $\lambda = (\mu_0, \mu)$*. Moreover, this condition is sufficient for global optimality of* $\widehat{x}$ *provided* $\mu_0 \neq 0$*.*

Parallel with differentiable multiplier rule. In combination with the calculus for subdifferentials of convex functions, this result leads to an effective method for solving constrained convex problems. This runs parallel to the method for solving differentiable equality-constrained optimization problems by the differentiable Lagrange multiplier rule ('in optimum Lagrange function is stationary for a suitable selection of Lagrange multipliers'), in combination with the calculus rules for differentiation.

Other types of convex problems. Convex programming problems can be modelled as problems of type (P). In order to do this, one has to make use of the reduction of convex objects to cones, which plays a central role in this paper. Applying then the convex multiplier rule given above leads to the well-known conditions of Karush-Kuhn-Tucker, as we will see in section 10.

The convex Lagrange multiplier rule can be presented in an alternative way: in terms of *primal and dual problems*. The problem (P) above is then called the primal problem. The dual problem is defined to be the problem

$$\varphi(y') = \langle y', \bar{y} \rangle \to \max,\ x' - \Lambda' y' \in C' \qquad (D)$$

Example 3 *For the special case of the LP-problem in standard form, the dual problem according to the definition above gives*

$$b^T \cdot y \to \max,\ c \geq A^T y,$$

keeping in mind that $(\mathbb{R}^n_+)' = (\mathbb{R}^n_+)^T$. *This is the usual dual LP-problem.*

Theorem 12 ***Primal-dual convex problems.*** *Consider the pair of primal-dual problems* (P) *and* (D). *Assume that* (P) *has a Slater point* $\tilde{x}$ $(\Lambda\tilde{x} = \bar{y},\ \tilde{x} \in \mathrm{int}C)$.

1. *Then the minimal value* $v(P)$ *of* (P) *is not smaller than the maximal value* $v(D)$ *of* (D),

$$v(D) \geq v(P).$$

2. *In particular, for a pair of primal-dual admissible elements* $(\widehat{x}, \widehat{y}')$ *the following criterion holds true*

$$f(\widehat{x}) = \varphi(\widehat{y}') \Leftrightarrow \widehat{x} \in \mathrm{absmin}(P),\ \widehat{y}' \in \mathrm{absmax}(D),\ v(P) = v(D) \in \mathbb{R}.$$

Relation convex Lagrange multiplier rule and primal-dual problems. Under assumption of the existence of a Slater point, primal-dual admissible pairs $(\widehat{x}, \widehat{y}')$ with $f(\widehat{x}) = \varphi(\widehat{y}')$ correspond precisely to solutions $(\widehat{x}, \lambda)$ of the multiplier rule with $\mu_0 = 1$: here $\widehat{y}'$ corresponds to μ. That is, the two theorems above represent the same result.

The Lagrange multipliers have the following concrete *shadow price* interpretation. Introduce parameters in the problem (P) by considering the family of problems $(P_y)_{y \in Y}$, defined by

$$f(x) = \langle x', x \rangle \to \min,\ \Lambda x = y,\ x \in C,$$

for all $y \in \mathbb{R}^m$. Let $S(y)$ be the minimal value of (P_y) for all $y \in Y$.

Theorem 13 ***Shadow price interpretation multipliers convex problems/solutions dual problems.*** *Assume that there exists a Slater point for the problem* (P). *Then the function* $S : Y \to \mathbb{R} \cup \{+\infty\}$ *is convex and* $\bar{y}$ *is an interior point of the domain of* S. *Moreover, if* $(\widehat{x}, \widehat{y'})$ *is a primal-dual admissible pair satisfying the equivalent conditions of the previous theorem, then*

$$S(y) \geq S(\bar{y}) + \langle y', y - \bar{y} \rangle$$

for all $y \in Y$.

This theorem is the motivation for calling y' a shadow price. Indeed, if we view (P_y) as a problem of minimizing costs, then if we put a price of y' for preventing a change from the constraint $\Lambda x = \bar{y}$ into $\Lambda x = y$, then this price will be attractive for all y as $S(y) \geq S(\bar{y}) + \langle y', y - \bar{y} \rangle$. Moreover, usually we will have

$$S(y) \approx S(\bar{y}) + \langle y', y - \bar{y} \rangle$$

for all $y \approx \bar{y}$, which provides additional motivation for the terminology shadow price.

On proofs. The three theorems in this section are immediate consequences of the separation of convex sets.

10 Applications

In this section we illustrate, for most types of convex objects, the duality theorem and the calculus rules for its operator of duality.

10.1 Duality of linear subspaces, calculation of annihilators, and the theory of linear equations

The duality relation for annihilators is as follows

$$L^{\perp\perp} = L \Leftrightarrow L \in \mathrm{ClLin}(X).$$

The calculation of annihilators consists of the following formulas:

$(L_1 + L_2)^\perp = L_1^\perp \cap L_2^\perp \; (i); (L_1 \cap L_2)^\perp = L_1^\perp + L_2^\perp \; (ii); (\Lambda L)^\perp = L^\perp \Lambda^* \; (iii); (L\Lambda)^\perp = \Lambda^* L^\perp \; (iv).$

Theorem 14 *(on solvability of linear equations) Let* $X = \mathbb{R}^n, \; Y = \mathbb{R}^m, \; \Lambda \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

1. *For the solvability of the system of linear equations* $\Lambda x = y$ *it is necessary and sufficient that* $y \in \ker \Lambda^T$.
2. *If* $X = Y$, *then the following alternative holds true: or the equation* $\Lambda x = y$ *is solvable for an arbitrary righthand side, or* $\ker \Lambda \neq \{0\}$.

Proof. We have

$$(\Lambda X)^\perp \overset{(iii)}{=} L^\perp \Lambda^T \; (v) \Leftrightarrow \Lambda X \overset{(0)}{=} (\Lambda X)^{\perp\perp} \overset{Id}{(} \; (\Lambda X)^\perp)^\perp \overset{(v)}{=} (X^\perp \Lambda^T)^\perp = (\{0\}\Lambda^T)^\perp + (\ker \Lambda^T)^\perp.$$

As a consequence $\Lambda X = Y \Leftrightarrow \ker \Lambda^T = \{0\}$. Moreover $\det\Lambda^T \neq 0 \Leftrightarrow \det\Lambda \neq 0 \Leftrightarrow \ker \Lambda = \{0\}$.

■

10.2 Duality of cones, calculation of conjugate cones, and the theory of linear inequalities.

The duality relation for the conjugacy operator is as follows:

$C'' = C \Leftrightarrow C \in \mathrm{ClCone}(X)$ if $\mathrm{ri}C \neq \emptyset$,

The calculation of the conjugacy operator consists of the following formulas:

$(C_1 + C_2)' = C_1' \cap C_2'\,(i)$; $(C_1 \cap C_2)' = C_1' + C_2'$ if $\mathrm{ri}C_1 \cap \mathrm{ri}C_2 \neq \emptyset\,(ii)$; $(\Lambda C)' = C'\Lambda^*\,(iii)$; $(C\Lambda)^* = \Lambda^* C'$ if $\mathrm{ri}C \cap \Lambda X \neq \emptyset\,(iv)$.

If $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, C_i are polyhedral cones, then in (ii) and (iv) one gets equalities.

Theorem 15 *(on the solvability of linear inequalities in the finite dimensional case).*

1. *Let A be a matrix of size $m \times n$ (m rows and n columns) and $b \in \mathbb{R}^m$. In order for the system $Ax = b$, $x \geq 0$ to have a solution, it is necessary and sufficient that for an arbitrary $y \in \mathbb{R}^{m'}$ for which $yA \geq 0$, the inequality $y \cdot b \geq 0$.*

2. *The inequality $Ax \leq b$ is solvable if and only if $y \cdot b \geq 0$ for an arbitrary $y \geq 0$ for which $yA = 0$.*

3. *The inequality $Ax \leq b$, $x \geq 0$ is solvable if and only if $y \cdot b \geq 0$ for all y for which $y \geq 0$ and $yA \geq 0$.*

Proof.

1. The cone $A\mathbb{R}^n_+$ is clearly polyhedral, and as a consequence

$$A\mathbb{R}^n_+ \overset{(0)}{=} ((A\mathbb{R}^n_+)')' \overset{(iii)}{=} (\mathbb{R}^n_+ A)' \Rightarrow 1).$$

2. The solvability of the inequality $Ax \leq b$ is equivalent(?) with $b \in A\mathbb{R}^n + \mathbb{R}^n_+$. The sum of two polyhedral cones in a finite-dimensional space is a polyhedral cone, as can be seen easily. In this way, $A\mathbb{R}^n + \mathbb{R}^m_+$ is a polyhedral cone, that is,

$$A\mathbb{R}^n + \mathbb{R}^m_+ \overset{(0)}{=} ((A\mathbb{R}^n + \mathbb{R}^m_+)')' \overset{(i)}{=} ((\mathbb{R}^n)'A' \cap (\mathbb{R}^m_+)')' = (0A' \cap \mathbb{R}^m_+)'.$$

   But $0A' = \ker A'$ and so the proved equality is equivalent to 2).

3. $A\mathbb{R}^n_+ + \mathbb{R}^m_+$ is a polyhedral cone in $\mathbb{R}^m$. As a consequence,

$$A\mathbb{R}^n_+ + \mathbb{R}^m_+ \overset{(0)}{=} ((A\mathbb{R}^n_+ + \mathbb{R}^m_+)')' \overset{(i),(iii)}{=} (\mathbb{R}_+ A' \cap \mathbb{R}^m_+)',$$

   which is equivalent to 3).

■

Comments. Statement 1) was proved by Minkowski (1896) and Farkas (1901), statement 2) by Ky Fan (1956) and statement 3) by Fenchel (1960).

10.3 Duality of convex zero-sets, calculation of polars, and finite dimensional convex geometry.

The duality relation for the polar operator is as follows:

$$B^{\circ\circ} = B \Leftrightarrow B \in \mathrm{ClCo}_0(X),$$

The calculation of polar operators consists of the following formulas:

$(B_1 \mathrm{co} \cup B_2)^\circ = B_1^\circ \cap B_2^\circ \,(i)$; $(B_1 \cap B_2)^\circ = B_1^\circ \mathrm{co} \cup B_2^\circ$ if $\mathrm{ri}B_1 \cap \mathrm{ri}B_2 \neq \emptyset \,(ii)$; $(\Lambda B)^\circ = B^\circ \Lambda^* \,(iii)$; $(B\Lambda)^\circ = \Lambda^* B^\circ$ if $\mathrm{ri}C \cap \Lambda X \neq \emptyset \,(iv)$.

Here we will consider all in the euclidian space $\mathbb{E}^n$. To begin with, we recall some examples of convex sets containing zero and their polars: for $n = 3$ the regular polyhedra (tetraeder, cube, octaeder, dodecaeder and icosaeder), for general n the polyhedra

$$B_\infty^n = \{x \in \mathbb{E}^n : |x_i| \leq 1 \ \forall i\},$$

and

$$B_1^n = \{x \in \mathbb{E}^n : \sum_{i=1}^n |x_i| \leq 1\}.$$

More generally, one can take the unit ball of any norm. For $n = 3$ the polar of a regular polyheder is again a regular polyheder. For example the polar of a cube is an octaeder. For general n, the polar of the polyheder B_∞^n is the polyheder B_1^n. More generally, the polar of the unit ball with respect to some norm is the unit ball with respect to the dual norm.

Theorem 16 *The following two definitions of compact polyhedra in finite-dimensional space are equivalent:*

1. *the convex hull of a finite collection of points,*
2. *a compact that is the intersection of a finite collection of half-spaces.*

Proof. Let M_1 be a compact in $\mathbb{R}^n$ that is the intersection of a finite collection of half-spaces. By the theorem of Krein-Milman it is the convex hull of its boundary points. Each boundary point is the intersection of n hyperplanes bounding one of the finite collection of halfspaces. Therefore the set of boundary points is finite.

Conversely, let $M_2 \subset \mathbb{E}^n$ be the convex hull of a finite collection of points $\{x_i\}_{i=1}^N$. Then, without restricting the generality of the argument, one can assume that 0 is an interior point of $M_2 = \mathrm{co}\{x_i\}_{i=1}^N$ (otherwise we consider M_2 in the affine hull $\{x_i\}_{i=1}^N$).

Then the polyheder M_2^0 is a convex compact (as a closed set, contained in a certain ball: the polar of a ball with center 0 contained in M_2. Then M_2^0 is the intersection of a finite collection of halfspaces $\langle x_i, y_i \rangle \leq 1$. By the theorem on the bipolar $M_2 = M_2^{\circ\circ}$ and so, M_2 is the intersection of a finite number of halfspaces. ■

10.4 Duality of convex functions, calculation of the Young-Fenchel transformation, and duality of convex problems

The duality relation for the Young-Fenchel transformation is as follows

$$f^{**} = f \Leftrightarrow f \in \mathrm{ClCo}^f(X)^r,$$

The calculation of Young-Fenchel transformation consists of the following formulas:

$(f_1 \mathrm{co} \vee f_2)^* = f_1^* \wedge f_2^*\,(i)$; $(f_1 \wedge f_2)^* = f_1^* \mathrm{co} \vee f_2^*$ if $\mathrm{ridom} f_1 \cap \mathrm{ridom} f_1 \neq \emptyset\,(ii)$; $(\Lambda f)^* = f^* \Lambda^*\,(iii)$; $(f\Lambda)^* = \Lambda^* f^*$ if $\mathrm{ri}\,\mathrm{dom} f \cap \Lambda X \neq \emptyset\,(iv)$.

We begin the subject duality of convex problems with the general duality scheme. Let X be a locally convex space and X^* its conjugate space. We consider the problem

$$f(x) \to \min,\ x \in X, \tag{$\bar{P}$}$$

where $f : X \to \mathbb{R} \cup \{\pm\infty\}$.

We embed $(\bar{P})$ in a family of problems, depending on a parameter $y \in Y$, where Y is also a locally convex topological space, considering a function $F : X \times Y \to \mathbb{R} \cup \{\pm\infty\}$ for which $F(x, 0) = f(x)$. The family

$$F(x, y) \to \min,\ x \in X \tag{$\bar{P}$}$$

is called a *perturbation of the problem* $(\bar{P})$, and the function $S : Y \to \mathbb{R} \cup \{\pm\infty\}$, associating to each $y \in Y$ the value $S(y)$ of the problem $(\bar{P}_y)$ is called the S-function of the family $(\bar{P}_y)$.

We find the conjugate of the function S:

$$S^*(y^*) = \sup_{y \in Y}(\langle y^*, y\rangle_{(Y,Y^*)} - \inf F(x, y)) = \sup_{x \in X, y \in Y}(\langle x^*, 0\rangle_{(X,X^*)} + \langle y^*, y\rangle_{(Y,Y^*)} - F(x, y)) = F^*(0, y^*).$$

If $S \in \mathrm{ClCo}^f(Y)$, then by virtue of (0) we come to the problem dual to $(\bar{P})$:

$$-F^*(0, y^*) \to \max,\ y \in Y^*.$$

Applying this construction to the problem of linear programming in symmetric form:

$$c \cdot x \to \min,\ Ax \geq b,\ x \geq 0,\ x \in \mathbb{R}^n,\ c \in \mathbb{R}^{n'},\ A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \tag{P_1}$$

the inequalities between the vectors are meant coordinate-wise. We consider the following perturbation of problem (P_1):

$$c \cdot x \to \min,\ Ax \geq y,\ x \geq 0. \tag{P_{1y}}$$

The epigraph of the S-function of the problems (P_{1y}) is a polyhedral cone, and therefore, its graph is closed (that is, the S-function is closed) and by the theorem of Fenchel-Moreau we obtain that $S = S^{**}$, in particular $S(b) = S^{**}(b)$. It is easy to see that

$$S^*(z) = \max\{z \cdot \lambda : A^T\lambda \leq c,\ \lambda \geq 0\},$$

that is, if $|S(y)| < \infty$, then $S(b)$ is the value of the problem

$$b \cdot \lambda \to \max,\ A^T\lambda \leq c,\ \lambda \geq 0\} \tag{${P_1}^*$}$$

This leads to the following result:

Theorem 17 *For the problems* (P_1) *and* (P_1^*) *the following alternative holds true:* $|S(b)| = \infty$ *or* $|S(b)| < \infty$*. In the first case either the value of the problem is* $-\infty$ *or the set of admissible elements is empty. In the second case both problems* (P_1) *and* (P_1^*) *have the same value and are both solvable. Therefore, for admissible elements* $\widehat{x}$ *and* $\widehat{\lambda}$ *the following criterion for both to be optimal holds true:*

$$\langle \widehat{x}, c \rangle = \langle \widehat{\lambda}, b \rangle.$$

10.5 Duality of norms, calculation of dual norms, and subdifferentials

The duality relation for norms is as follows

$$N^{**} = N \Leftrightarrow N \in \mathrm{Norm}(X).$$

The calculation of dual norms consists of the following formulas:
$(N_1 \mathrm{co} \vee N_2)^* = N_1^* \wedge N_2^*$ (i); $(N_1 \wedge N_2)^* = N_1^* \mathrm{co} \vee N_2^*$ (ii);
$(\Lambda N)^* = N^* \Lambda^*$ (iii); $(N\Lambda)^* = \Lambda^* N^*$ (iv).

The best-known example is the l_p-norm on $\mathbb{R}^n$, $1 \leq p <\leq \infty$ given by

$$\|x\|_{l_p^n} = (\sum_{i=1}^{n} |x_i|^p)^{1/p}$$

for $p \neq \infty$ and by

$$\|x\|_{l_p^\infty} = \max_{1 \leq i \leq n} |x_i|.$$

The dual of the l_p-norm turns out to be the $l_{p'}$-norm, where $1/p + 1/p' = 1$ (with the convention that $1/\infty = 0$).

The main application is to the calculation of the subdifferential of convex functions. We give a numerical example.

Example 4 *Solve* $f(x_1, x_2) = (x_1^2 + x_2^2)^{\frac{1}{2}} - \frac{1}{2}x_1 + \frac{1}{3}x_2 \rightarrow \min.$

Solution. This is a convex minimization problem; writing down the stationarity conditions does not lead to the solution. Let us instead compute the subdifferential at the point of non-differentiability $(0, 0)$: this is seen to be the disk with center $(-\frac{1}{2}, \frac{1}{3})$ and radius one. This disk contains the origin $(0, 0)$. It follows that the point of non-differentiability $(0, 0)$ is the unique solution of the problem.

10.6 Duality of convex functions at a point, calculation of subdifferentials and the Karush-Kuhn-Tucker conditions

The calculation of subdifferentials consists of the following formulas:
$\partial(f_1 \mathrm{co} \vee f_2)(\widehat{x}) = \partial f_1(\widehat{x}) \cap \partial f_2(\widehat{x})$ (i);
$\partial(f_1 \wedge f_2)(\widehat{x}) = \partial f_1(\widehat{x}) \mathrm{co} \vee f_2(\widehat{x})$ (ii);
$\partial(f\Lambda)(\widehat{x}) = \Lambda^* \partial f(\widehat{x})$ (iii);
$\partial(\Lambda f)(\widehat{x}) = \partial f(\widehat{x})\Lambda'$, if $\mathrm{ri}\,\mathrm{im} f \cap \Lambda X \neq \emptyset$ (iv);
$\partial(f_1 + f_2)(\widehat{x}) = \partial f_1(\widehat{x}) + \partial f_2(\widehat{x})$ if ... (v);
$\partial(f_1 \oplus f_2)(\widehat{x}) = \partial f_1(\widehat{x}) \cap \partial f_2(\widehat{x})$ (vi).

Theorem 18 ***Karush-Kuhn-Tucker theorem.*** *Let in the problem of convex programming*

$$f_0(x) \to \min,\ f_i(x) \le 0,\ 1 \le i \le m,\ x \in A \tag{P}$$

where X is a locally convex space, $f_i : X \to \mathbb{R}$ convex and continuous at points $x \in A$, where A is a convex subset of X. Then, if $\widehat{x}$ is a solution of (P), then there exists a nonzero selection of Lagrange multipliers $\lambda = (\lambda_0, \dots, \lambda_m)$ such that (a) $\lambda \ge 0, (b)$ $\lambda_i f_i(\widehat{x}) = 0,\ 1 \le i \le m$ and the minimum principle holds (for the Lagrange function $\mathcal{L}(x, \lambda) = \sum_{i=1}^{m} \lambda_i f_i(x)$):

$$(c)\ \min \mathcal{L}(x, \lambda) = \mathcal{L}(\widehat{x}, \lambda).$$

and if for the problem (P) the conditions $(a) - (c)$ are satisfied with $\lambda_0 \ne 0$, then $\widehat{x}$ is a solution of (P).

Proof. Let $\widehat{x}$ be a solution of (P). Then this same element is also a solution of the following problem (without constraints):

$$f(x) \to \min, x \in X,$$

where $f(x) = \max\{f_0(x) - f_0(\widehat{x}), f_1(x), \dots, f_m(x)\} + \delta A(x) = g(x) + \delta A(x)$. By the Fermat theorem $0 \in \partial f(\widehat{x})$, by the theorem of Moreau-Rockafellar there exists an element $x^* \in X^*$ such that

$$x^* \in \partial g(\widehat{x}),\ -x^* \in \partial \delta A(\widehat{x}). \tag{i}$$

Let us assume that $f_k(\widehat{x}) = 0,\ 1 \le k \le m',\ f_i(\widehat{x}) < 0,\ k \ge m'+1$. Then, according to the theorem of Dubovitsky-Milyutin there exist elements $x_i^* \in \partial f_i(\widehat{x})$ and numbers $\lambda_i \ge 0,\ 0 \le i \le m',\ \sum_{i=1}^{m'} \lambda_i = 1$ such that

$$x^* = \sum_{i=1}^{m'} \lambda_i x_i^* \tag{ii}$$

From the second inclusion in (i) it follows that

$$\langle x^*, \widehat{x} \rangle = sA(-x^*) \qquad (iii).$$

We put $\lambda_i = 0,\ i \ge m'+1$ and from $(ii), (iii)$ and the definition of the subdifferential we obtain that

$$0 = \inf_{x \in A} \sum_{i=0}^{m} \lambda_i \langle x_i^*, x - \widehat{x} \rangle \le \inf_{x \in A} (\lambda_0 (f_0(x) - f_0(\widehat{x})) + \sum_{i=1}^{m} \lambda_i f_i(x)).$$

The sufficiency is obvious. ■

Bibliography

G.G. Magaril-Il'yaev, V.M. Tikhomirov, *Convex Analysis: Theory and Applications*, Translations of Mathematical Monographs, Volume 222, American Mathematical Society, (2000).

R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NY, (1970)

W. Fenchel, *On conjugate convex functions*, Can. J. Math., 1, 73-77 (1949).

Minkowski, *Theorie der konvexe Körper*, in Gesammelte Abhandlungen, Vol. II, Teubner, Leipzig (1910).