# Notes on Multi-criteria Optimization

D. Gueorguiev 10/7/23

Table of Contents

[Notes on Multi-criteria Optimization 1](#_Toc152306462)

[Orders and Cones (1.4) 1](#_Toc152306463)

[Classification of Multi-criteria Optimization Problems (1.5) 6](#_Toc152306464)

[Efficient Solutions and Non-Dominated Points (2.1) 7](#_Toc152306465)

[Weakly and Strictly Efficient Solutions (2.3) 7](#_Toc152306466)

[Appendix 11](#_Toc152306467)

[Affine Space 11](#_Toc152306468)

[Caratheodory’s Theorem 11](#_Toc152306469)

[Semi-continuity 12](#_Toc152306470)

[Radon’s Theorem (Johann Radon, 1921) 13](#_Toc152306471)

[Helly’s Theorem (Eduard Helly, 1922) 13](#_Toc152306472)

## Orders and Cones (1.4)

Definition: binary relation on a set is called

* *reflexive* if
* *irreflexive* if
* *symmetric* if
* *asymmetric* if
* *antisymmetric* if *and*
* *transitive* if *and*
* *negative transitive* if *and*
* *connected* if *or*
* *strongly connected (or total)* if *or*

**Definition**: *equivalence relation*

A binary relation on a set is an *equivalence relation* if it is reflexive, symmetric, and transitive.

**Definition**: *preorder*

A binary relation on a set is a *preorder* (*quasi-order*) if it is reflexive and transitive.

Given any preorder two other relations are closely associated with .

**Definition**: *strict preference* relation

(1.9)

**Definition**: *equivalence* (or *indifference*) relation

(1.10)

**Proposition 1.6**: Let be a preorder on . Then the relation defined in (1.9) is irreflexive and transitive and relation defined in (1.10) is an equivalence relation.

Proof: a) we consider first. This relation is reflexive because is. Furthermore, is symmetric by definition. Now let be such that and . Then using transitivity of

(1.11)

b) For , note that is irreflexive by definition. Suppose there are such that and . Then and from transitivity of , . To show that , assume . But since we get from transitivity of . This contradiction implies , i.e. .

**Proposition 1.7**: An asymmetric binary relation is irreflexive. A transitive, irreflexive binary relation is asymmetric.

Definition 1.8: a binary relation on is

* a *total preorder* if it is reflexive, transitive, and connected,
* a *total order* if it is an antisymmetric *total preorder*
* a strict weak order if it is asymmetric and negatively transitive

Notation:

is the *interior* of

is the *relative interior* of

is the *boundary* of

is the *closure* of

is the *convex hull* of

**Definition** *affine set*

The set is affine if

**Definition** *convex set*

The set is convex if

**Definition** *affine hull* or *affine span*

Affine hull of a set is the smallest affine set which contains . Equivalently, it is the intersection of all affine sets containing .

**Definition** *conical combination (weighted sum)*

Given the vectors in real vector space , a conical combination of those is the following element of :

where are non-negative numbers. The conical sum defines a cone in .

**Definition** *conical hull*

The set of all conical combinations for a given set is called the conical hull of and denoted . That is,

By taking , it follows the zero vector (origin) belongs to all conical hulls.

It can be easily shown that the conical hull of a set is a convex set. In fact, it is the intersection of all convex cones containing plus the origin.

Note: If is a compact set (in particular, when it is finite non-empty set of points), then the condition “plus the origin” is unnecessary.

If we discard the origin, we can divide all coefficients by their sum to see that a conical combination is a convex combination scaled by a positive factor.

A graph of a function

Description automatically generated

Figure: in the 2D plane, the conical hull of a circle passing through the origin is the open half-plane defined by the tangent line to the circle at the origin plus the origin.

**Definition** *convex hull*

*Informal definition*: The convex hull of a shape is the smallest convex set that contains the shape.

The convex hull can be defined as the intersection of all convex sets containing any given subset of Euclidean space, or equivalently as the set of all convex combinations of points in the subset. For a bounded subset of the plane, the convex hull may be visualized as the shape enclosed by a rubber band stretched around the subset.

1st definition of convex hull:

The convex hull of a given set is the (unique) minimal convex set containing .

2nd definition of convex hull:

The convex hull of a given set is the intersection of all convex sets containing .

3rd definition of convex hull:

The set of all convex combinations (all coefficients sum up to one and are non-negative) of points in

4th definition of convex hull:

The union of all simplices with vertices in

**Note** on the practical visualization of convex hull

For bounded sets in the Euclidean plane, not all on one line, the boundary of the convex hull is the simple closed curve with minimum perimeter containing . Physical analogy which is often employed for visualizing the boundary is stretching rubber band so that it surrounds the entire set and then releasing it, allowing it to contract; when it becomes taut, itencloses the convex hull of .

However, this definition **does not necessarily generalize to higher dimensions**. For a finite set of points in 3D space, a neighborhood of a spanning tree of the points encloses them with arbitrarily small surface area, smaller than the surface of a convex hull. However, in higher dimensions, variants of [the obstacle problem](https://en.wikipedia.org/wiki/Obstacle_problem) of finding a minimum energy surface above a given shape can have a convex hull as their solution.

A blue triangle with red outline

Description automatically generated

Figure 1: the convex hull of the red set is the blue and red convex set.

A diagram of a hexagon with black dots and arrows

Description automatically generated

Figure 2: Convex hull of a bounded planar set: rubber band analogy

Equivalence of the definitions of convex hull

The first definition states that there should exist unique minimal convex set containing , for every . It is not obvious why this should be true.

Let us look into the second definition where the intersection of all convex sets containing is well defined – it is a subset of every other convex set that contains , because is included among the sets being intersected. Thus, it is the unique minimal convex set containing . Therefore, the two definitions are equivalent.

Each convex set containing must contain all convex combinations of points in , so the set of all convex combinations is contained in the intersection of all convex sets containing . Conversely, the set of all convex combinations is itself a convex set containing , so it also contains the intersection of all convex sets containing X, and therefore the second and third definition are equivalent.

According to the [Caratheodory’s Theorem](https://en.wikipedia.org/wiki/Carath%C3%A9odory%27s_theorem_(convex_hull)), if is a subset of a -dimensional Euclidean space, every convex combination of finitely many points from is also a convex combination of at most points in . The set of convex combinations of a -tuple of points is a simplex. Therefore, every convex combination of points of X belongs to a simplex whose vertices belong to X, and the third and fourth definitions are equivalent.

Note on a difference with respect to compactness between convex hull and conical hull

The convex hull of a compact set is also a compact set. This is not true for conical hull. For starters, the conical hull is unbounded. Moreover, it is not necessarily a closed set, here is a counterexample: sphere passing through the origin, with the conical hull being an open half-space plus the origin.

However, if is a non-empty convex compact set which does not contain the origin, then the convex conical hull of is closed set.

**Definition 1.11**: A subset is called a *cone*, if

**Definition 1.13**: A cone in is called

* *nontrivial* or *proper* if ,
* *convex* if for all
* *pointed* if for ,

**Theorem**: *closedness under addition is sufficient for convexity*

A cone is convex if for all we have . In such case the following would also be true: and since is a cone.

**Lemma**: the set is a convex cone if for all and

**Definition**: Set of non-negative elements on by

Given the order relation the set defined as

(1.15)

will be interpreted as “*the set of non-negative elements according to* ”

**Proposition 1.14**: Let be compatible with the scalar multiplication i.e., for all and all it holds that . Then defined in (1.15) is a cone.

*Proof*:

Let . Then for some with . Thus for all . Hence for all .

**Example 1.15**:

Let us consider a weak component-wise order on . Here *iff* for all or for all . Therefore, .

It is interesting to consider the definition (1.15) with fixed, i.e., . If is an order relation, is the set of elements of that is preferred to or that are dominated by .

A natural question to ask is: Under what conditions is the same for all ? In order to answer this question, we need another assumption on order relation .

**Definition**: *binary relation compatible with addition*

is said to be compatible with addition if for all and all .

**Lemma 1.16**: If is compatible with addition and then 0.

*Proof*:

Let . Then there are with such that . Using , compatibility with addition implies or equivalently .

**Lemma 1.16** means that if is compatible with addition then sets do not depend on .

**Theorem 1.17**: Let be a binary relation on which is compatible with scalar multiplication and addition. Then the following statements hold:

1. *iff* is reflexive
2. is pointed *iff* is antisymmetric
3. is convex *iff* is transitive

*Proof of 2*: Let be reflexive and let such that too. Then there are such that and as well as such that and . Thus, and there must be such that and . Therefore, compatibility with addition implies . Antisymmetry of now yields and therefore , i.e., is pointed.

*Proof of 3*: Let be transitive and let . Since is compatible with scalar multiplication, is a cone and we only need to show . By Lemma 1.16 we have 0 and 0. Compatibility with addition implies , transitivity yields , from which .

Let be convex and let be such that and . Then and . Because is convex, . By Lemma 1.16 we get 0 and by compatibility with addition .

**Example 1.18**:

1. The weak component wise order is compatible with addition and scalar multiplication. contains 0, is pointed and convex.
2. The max order is compatible with scalar multiplication, but not with addition e.g., , but this relation is reversed when adding . Furthermore, is reflexive, transitive, but not antisymmetric (e.g., and ).

We have defined a cone given relation . We can also use a cone to define an order relation. Let be acone. Define by:

(1.16)

**Proposition 1.19**: Let be a cone. Then defined in (1.16) is compatible with scalar multiplication and addition in .

*Proof:*

//TODO: furnish the proof

**Theorem 1.20**: Let be a cone and let be as defined in (1.16). Then the following statements hold:

1. is reflexive *iff* .
2. is antisymmetric *iff* is pointed.
3. is transitive *iff* is convex.

*Proof*:

1. Let and . Thus, and for all .

Let be reflexive. Then we have for all , i.e., .

1. Let and . Thus and . Adding d to the latter relation, compatibility with addition yields . Then antisymmetry implies .

Let be such that and . Therefore and . Since is pointed, , i.e., .

1. Let such that and . Therefore

//TODO: finish the notes on this section 1.4 Orders and Cones

## Classification of Multi-criteria Optimization Problems (1.5)

By the choice of an order on , we can finally define the meaning of “min” in the problem formulation:

(1.17)

The different interpretations of “min” pertaining to different orders are the foundation of the classification of multicriteria optimization problems.

With multiple objective functions we can evaluate objective value vectors . However, we have seen that these vectors are not always in objective space , directly.

Before, we have formulated the optimization problem

(1.18)

That is, we have used a mapping from objective space to , where the min in (1.18) is actually defined by the canonical order on .

In general, the objective function vectors are mapped from to an ordered space, e.g., where comparisons are made using the order relation . This mapping is called the *model map*.

The components of the Multicriteria Optimization Problem (MOP) are:

* the feasible set ,
* the objective function vector ,
* the objective space ,
* the ordered set ,
* the model map

The feasible set , the objective function vector , and objective space constitute *the data* of the MOP. The model map provides the link between objective space and ordered set, in which, finally, the meaning of the minimization is defined. Thus, with the three aspects – data, model map, and ordered set the classification

completely described a MOP.

**Definition 1.23**: A feasible solution is called an optimal solution of a MOP if there is no such such that

(1.22)

For an optimal solution is called an optimal value of the MOP.

The set of optimal solutions is denoted by .

The set of optimal values is denoted by .

Notes:

1. Since we are often dealing with orders which are not total, a positive definition of optimality, like

for all , is not possible in general.

1. For specific choices of and , specific names for optimal solutions and values are commonly used such as *efficient solutions* (see Definition 2.1) or *lexicographically optimal solutions*.

Definition 1.26: A multicriteria optimization class (MCO class) is *the set of all MOPs with the same model map and ordered set* and is denoted by:

For instance, will denote the class of all MOPs, where optimality is understood in the sense of efficiency.

## Efficient Solutions and Non-Dominated Points (2.1)

We consider multicriteria optimization problems of the class :

. (2.1)

**Definition 2.1**. A feasible solution is called *efficient* or *Pareto optimal* if there is no other such that . If is efficient, is called *nondominated point*. If and we say dominates and dominates . The set of all efficient solutions is denoted and called *the efficient set*. The set of all nondominated points , where , is denoted with and called *the nondominated set*.

Several other, equivalent, definitions of efficiency are frequently used, and we shall often refer to the one which is best suited in a given context. In particular, is efficient if

1. There is no such that for and for some ;
2. There is no such that ;
3. for all ;
4. ;
5. There is no with ;
6. for some implies

**Definition 2.1** and equivalent definitions , , and consider and check for images of feasible solutions to the left and below (in direction of ) of that point. In the equivalent definitions and , through , the set is translated so that the origin coincides with , and the intersection of the translated set with the negative orthant is checked. The intersection contains only if is efficient.

We will discuss the existence and the properties of the efficient set and the nondominated set . It is convenient to consider and then use the properties of to derive results on . According to our definitions, is nondominated, if there is no such such that .

First, we show by means of an example that even for convex sets and the efficient set and the nondominated set might be empty or consists of isolated points.

Example 2.2 (*Goepfert* and *Nehse* (1990)).

//TODO: finish this example

**Proposition 2.3**.

*Proof*: The result is trivial if , because and the nondominated subsets of both are empty.

Let . First, assume that , but . In this case there are two possibilities.

If there is and such that .

Since we get , a contradiction. If there is such that . Let , which is in Therefore and again contradicting the assumption.

Hence, in either case .

Second, assume but . Then there is some with . Thus we can write with , and therefore with This implies that , contradicting the last assumption (Second). Hence, .

Figure: Non-dominated points of and are the same.

**Proposition 2.4.** . The set of non-dominated points of is a subset of the boundary of .

Proof:

Let and assume that . Therefore, and there exist an -neighborhood of (with ). Let , . Then we can choose some , such that . Now, with , i.e. which is a contradiction.

**Corollary 2.5** (follows from Propositions 2.3 and 2.4)

If is open or if is open .

**Proposition 2.6**.

Proof: Let . Then for some ,

Assuming it follows that there must be some and such that with from which we conclude that , contradicting the assumption.

Analogously, i.e.

**Proposition 2.62**. . The inclusion is not satisfied in general.

Proof: //TODO: find counterexample

**Proposition 2.7.** , for , .

Proof: //TODO: furnish a proof

**Definition 2.8.** Let be a preordered set, i.e. is reflexive and transitive. is inductively ordered, if every totally ordered subset of has lower bound. A totally ordered subset of is also called ***a chain***.

**Theorem 2.9 (*Zorn’s lemma*).** Let the preordered set be inductively ordered. Then contains a minimal element, i.e. there is such that implies .

**Theorem 2.10 (*Borwein*, 1983).** Let be a nonempty set and suppose there is some such that the section is compact; that is, contains a compact section. Then is nonempty.

Proof: We proceed like this: we use the compactness of to show that every chain in has lower bound. Thus is inductively ordered, and by Zorn’s Lemma contains a minimal element . Showing that is efficient in completes the proof.

Let be the compact section that exists by assumption and let , where is some index set, be a chain in . We prove that has a lower bound. To that end, let be the set of all finite subsets of index set . For all finiteness of and being a chain in imply that

exists and of course . Consider all sets , where . Obviously and is compact as a closed subset of the compact set . Furthermore, if , i.e. is finite, because it contains . Finally, by compactness of it follows that , which means there is some

(2.5)

In terms of the component-wise order this means for all , or, in other words, is a lower bound of , which is therefore inductively ordered.

Figure : is compact section of

Now we can apply Zorn’s Lemma (Theorem 2.9) to conclude that contains a minimal element . It remains to be shown that . Assume the contrary. Then there would be some with . For we have

(2.6)

The first inclusion holds because , the second is clear. Since this implies , so that contradicts minimality of in .

Note: in this Proof we have used the following fact about compact sets: if is compact and is a family of closed subsets of for some index set such that for all finite subsets of of then .

Now we will present another existence result which does not use a compact section but a condition on which is similar to the finite subcover property of compact sets: the *-semicompactness condition*, which considers open covers with special sets.

**Definition 2.11.** A set is called -*semicompact* if every open cover of of the form has a finite subcover. This means that whenever there exist and such that

(2.7)

Here denotes the complement of the set . Note that since is closed is open.

Based on Zorn’s Lemma again, we can prove that -semicompactness guarantees existence of efficient points – this fact has been established by Corley in 1980.

**Theorem 2.12.** (*Corley*, 1980) If is -semicompact then .

Proof: As with the proof of the Borwein’s Theorem we will show that is inductively ordered and consecutively apply Zorn’s Lemma. First, we will construct an open cover of as in Definition 2.11 and derive a contradiction when we assume that is not inductively ordered.

Assume is not inductively ordered. Then there is a totally ordered subset (a chain) of , say which has no lower bound. Therefore

(2.8)

Note: As seen in the proof of Borwein’s Theorem, any element in this intersection would be a lower bound of . (*it is easy to show that any element of the intersection is a lower bound of* ).

So assuming (2.8) then for each there exist some such that .

Since is closed, defines an open cover of . Moreover, if and only if and the sets of the cover are totally ordered by inclusion because is a chain. Also, the fact that is -semicompact implies there is a finite subcover of .

Combining the last two observations, it follows that there is a minimal set (with respect to inclusion) in the finite subcover and hence there exists a single such that . This implies for all and which is not possible. Therefore is inductively ordered. Further we proceed similarly to the proof of the Borwein’s Theorem to conclude that .

We apply Zorn’s Lemma to conclude that contains a minimal element . It remains to be shown that . Assume that . Then there would be some with . For we have

The first inclusion holds because . Since the statement that , so contradicts minimality of in . Thus, the last assumption cannot be true hence .

*for a chain*

Figure: constructing a cover from a chain

**Definition 2.13.** A set is called -compact, if for all the section is compact.

**Proposition 2.14.** If is -compact, then is -semicompact.

Proof: Let be an open cover of . For arbitrary take

(2.9)

(2.9) defines an open cover of , a compact set, since is -compact. But compactness implies that the cover in (2.9) contains a finite subcover of . This finite subcover together with yields a finite cover of , of the structure required for -semicompactness.

Note: it is always possible to cover the -compact set with open cover of the form - this is true because of the compactness property of .

**Corollary 2.15.** (*Hartley, 1978*) If is nonempty and -compact, then .

Proof: the result follows immediately from Theorem 2.12 and Proposition 2.14.

So far, we focused on existence of nondominated points. Let us now consider existence of efficient solutions, i.e. conditions that guarantee . We can use Theorem 2.12 and properties of to get an existence result for .

Theorem 2.19 below is a multicriteria analog to the well-known result that a lower semicontinuous function attains its minimum over a compact set.

**Definition 2.16.** A function is -semicontinuous if

(2.10)

Is closed for all , i.e. the preimage of the translated negative orthant is always closed.

//TODO: finish the section on Efficient Solutions and Nondominated points

## Weakly and Strictly Efficient Solutions (2.3)

Recall Table 1.2 defining the following orders:

Notation Definition Name

weak componentwise order

componentwise order

strict componentwise order

or lexicographic order

max-order

With the (weak, strict) componentwise orders, we defined the following subsets of as follows:

* , the nonnegative orthant of ;
* ;
* , the positive orthant of

Nondominated points are defined by the componentwise order on . When we use the weak and strict componentwise order instead, we obtain definitions of strictly and weakly nondominated points, respectively.

It will be proven an existence result for weakly nondominated points and weakly efficient solutions. We then give a geometric characterization of all three types of efficiency and some further results on the structure of weakly efficient solutions of convex multicriteria optimization problems.

**Definition 2.24**. A feasible solution is called *weakly efficient* (*weakly Pareto optimal*) if there is no such that , i.e., that for all . The point is then called *weakly nondominated*.

A feasible solution is called *strictly efficient* (*strictly Pareto optimal*) if there is no such that . The weakly (strictly) efficient and nondominated sets are denoted and , respectively.

**Notes**:

1. there is no such concept as *~~strict nondominance~~*  for sets .
2. By definition, strict efficiency prohibits solutions with i.e., strict efficiency is the multi-criteria analog on unique optimal solutions in scalar optimization. Thus,

and . (2.18)

1. All existence results for imply existence of as well.
2. can be nonempty even if is empty i.e. .

Obviously, a weakly nondominated point is a nondominated point with respect to , a notation that is quite convenient in the context of cone-efficiency and cone-nondominance.

From the definitions it follows:

(2.16)

and

(2.17)

As in the case of efficiency, weak efficiency has several equivalent definitions. A feasible solution is weakly efficient iff

1. there is no such that
2. .

**Theorem 2.25**. Let be nonempty and compact. Then .

*Proof*. Suppose . Then for all there is some such that . Taking the union over all we obtain

. (2.19)

Because

//TODO: finish the theorem proof

**Definition 2.28**. Let , , and .

(2.23)

is called the *level set* of at .

(2.24)

is called the *level curve* of at .

(2.25)

Is called the *strict level set* of at .

Obviously, and .

**Theorem 2.30** (*Ehrgott et al. (1997)*). Let be a feasible solution and define . Then

1. is *strictly efficient* if and only if

. (2.28)

1. is *efficient* if and only if

. (2.29)

1. is *weakly efficient* if and only if

. (2.30)

*Proof*.

1. is *strictly efficient*

there is no , such that

there is no , such that for all

there is no , such that

1. is *efficient*

there is no such that both for all

and for some

there is no such that both and for some

1. is *weakly efficient*

there is no such that for all

there is no such that

**Theorem 2.30** shows that not all the criteria are always needed to see if a feasible solution is weakly or strictly efficient. Once the intersection of some level sets contains only , or the intersection of some strict level sets is empty, it will remain so when intersected with more (strict) level sets. Question: how many objectives are actually needed to determine if a feasible solution is (strictly, weakly) efficient or not?

Let and denote by the objective function vector that only contains .

**Corollary 2.32**. Let be nonempty and let . Then the following statements hold:

1. If is a weakly efficient solution of it is also a weakly efficient solution of
2. If is a strictly efficient solution of it is also a strictly efficient solution of

Corollary 2.32 says that weak or strict efficiency of some solution for a problem with a subset of the objectives implies weak (strict) efficiency for a problem with all objectives.

**Question**: Is it possible to find all weakly (strictly) efficient solutions by solving only problems with less than objectives?

**Answer**: For weakly efficient solutions this is possible for convex functions.

**Assumption for this section**: is a convex set and that are convex functions.

This implies that all level sets are convex.

**Theorem 2.33** (*Helly, 1923*). Let and let be convex sets. Then

iff for all collections of sets

Equivalently stated, we can say that

Iff there is a subset of sets , such that

.

**Proposition 2.34**. Consider the multicriteria optimization problem , where is convex, are convex and . Then is weakly efficient iff there is a subset , such that is weakly efficient solution of

//TODO: finish the section on weakly and strictly efficient solutions

## The Weighted Sum Method and Related Topics (3)

We will investigate to what extent an MOP

(3.1)

of the Pareto class

can be solved (i.e. its efficient solutions be found) by solving single objective problems of the type:

(3.2)

which in terms of the classification of MOPs presented in Section 1.5is written as

(3.3)

where denotes the scalar product in . We call the single objective (or scalar) optimization problem (3.2) a *weighted sum scalarization* of the MOP (3.1).

## Appendix

### Affine Space

A diagram of a triangle with arrows

Description automatically generated

Figure: the origins from Alice’s and Bob’s perspectives. Vector computation from Alice’s perspective is in red, whereas that from Bob’s is in blue.

Informal Definition of Affine Space:

Affine space is what is left from a vector space after one has forgotten which point is the origin. Imagine that Alice knows that certain point is the actual origin, but Bob believes that another point – call it – is the origin. Two vectors, and , are to be added. Bob draws an arrow from point to point and another arrow from point to point , and completes the parallelogram to find what Bob thinks is , but Alice knows that he actually has computed . Similarly, Alice and Bob may evaluate any linear combination of a and b or of any finite set of vectors and will, generally, get different answers. However, if the sum of the coefficients in a linear combination is 1, then Alice and Bob will arrive at the same answer.

If Alice travels to then Bob can similarly travel to

Under this condition, for all coefficients , Alice and Bob describe the same point with the same linear combination, despite using different origins. While only Alice knows the “linear structure”, both Alice and Bob know the “affine structure” – i.e., the values of affine combinations, defined as linear combinations in which the sum of the coefficients is 1. A set with an affine structure is an affine space.

Definition affine space

//TODO: finish Affine space discussion

### Compact Space

//TODO: finish Compact space discussion

### Caratheodory’s Theorem

If a point lies in the convex hull of a set , then x can be written as the convex combination of at most *extremal* points in , as non-extremal points can be removed from without changing the membership of in the convex hull.

**Conical Combination Theorem** (*equivalent to Caratheodory’s Theorem which is for convex hulls*)

If a point lies in the conical hull of a set , then x can be written as a conical combination of at most points.

**Example**:

Caratheodory’s theorem in 2D

We can construct a triangle consisting of points from that encloses any point in the convex hull of .

Let . The convex set of this set is a square. Let in the convex hull of . We can then construct a set , the convex hull of which is a triangle and encloses .

**A blue and pink square with black dots

Description automatically generated**

Figure: An illustration of Caratheodory’s Theorem for a square in

Proof of Caratheodory’s Theorem:

Note: We will use the fact that is an ordered field, that is a field on which it can be imposed total order. Thus, the theorem can be applied to any field , together with total order.

**Theorem** (*Caratheodory’s Theorem*)

If then

1. is the non-negative sum of at most points of .
2. is the convex sum of at most points of .

We will prove the Caratheodory’s Theorem in the finite case. This reduction to the finite case is possible because is the set of *finite* convex combination of elements of .

**Lemma**:

If then and at most of them are nonzero.

Proof of the Lemma:

When the proof is trivial. If we can prove it for all then by induction, we have proved it for all . Thus,

//Finish Caratheodory theorem overview

### Semi-continuity

Property of extended real-valued functions that is weaker than continuity. An extended real-valued function f is *upper* (resp. *lower*) *semicontinuous* at a point if the function values for arguments near are not much higher (resp. lower) than .

A function is continuous *iff* it is both upper and lower semicontinuous.

**Definition**: Assume that is a topological space and is a function with values in the extended real numbers . Then is called *upper semicontinuous at a point* if for all . Equivalently, is *upper semicontinuous* at *iff*

.

**Lemma**: A function is called *upper semicontinuous* if it satisfies the following equivalent conditions:

1. The function is upper semicontinuous at every point of its domain.
2. All sets with are open in , where
3. All superlevel sets with are closed in
4. The hypograph is closed in .

//TODO: Finish semi-continuity overview

### Radon’s Theorem (Johann Radon, 1921)

Any set of points in can be partitioned into two sets whose convex hulls intersect. A point in the intersection of these convex hulls is called a *Radon point* of the set.

For example, in the case of , any set of four points in the Euclidean plane can be partitioned in one of two ways. It may form a triple and a singleton, where the convex hull of the triple (a triangle) contains the singleton; alternatively, it may form two pairs of points that form the endpoints of two intersecting line segments.

A couple of triangles with red and green lines

Description automatically generated

Figure: Two sets of four points in the plane (the vertices of a square and an equilateral triangle with its centroid), the multipliers solving the system of three linear equations for these points, and the Radon partitions formed by separating the points with positive multipliers from the points with negative multipliers.

Proof and construction:

Consider any set . Then there exists a set of multipliers , not all of which are zero, solving the system of linear equations

, ,

because there are unknowns (the multipliers) but only equations that they must satisfy (one for each coordinate of the points, together with the final equation requiring the sum of the multipliers to be zero). Fix some particular nonzero solution . Let be the set of points with positive multipliers, and let be the set of points with multipliers that are negative or zero. Then and form the required partition of the points into two subsets with intersecting convex hulls.

The convex hulls of and must intersect, because they both contain the point

,

where

The left-hand side of the formula for expresses this point as a convex combination of the points in , and the right-hand side expresses it as a convex combination of the points in . Therefore, p belongs to both convex hulls.

QED

Note: This proof method allows for the efficient construction of a Radon point, in an amount of time that is polynomial in the dimension by using Gaussian elimination or other efficient algorithms to solve the system of equations for the multipliers.

### Helly’s Theorem (Eduard Helly, 1922)

Let be a finite collection of convex subsets of , with . If the intersection of every of these sets is nonempty, then the whole collection has nonempty intersection; that is,

.

For infinite collections one has to assume compactness:

Let be a collection of compact convex subsets of , such that every subcollection of cardinality at most

has nonempty intersection. Then the whole collection has nonempty intersection.

A diagram of different shapes

Description automatically generatedFigure: Helly’s theorem for the Euclidean plane: if a family of convex sets has a nonempty intersection for every triple of sets , then the whole family has a nonempty intersection.

Proof:

We prove the finite version using Radon’s theorem. The infinite version then follows by the finite intersection property characterization of compactness: a collection of closed subsets of a compact space has a non-empty intersection iff every finite subcollection has a non-empty intersection (once you fix a single set, the intersection of all others with it are closed subsets of a fixed compact space).

The proof is by induction:

*Base case*: Let . By our assumptions, for every there is a point that is in the common intersection of all with the possible exception of . Now we apply Radon’s theorem to the set