Computational Method of Optimization

Convex Sets - Lecture II

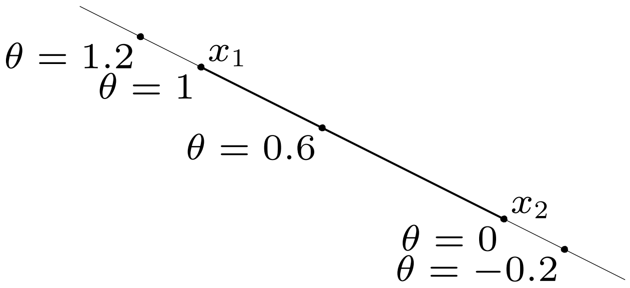
Convex Sets

- Topics covered
- Affine and convex sets
- Some important examples
- Operations that preserve convexity
- Generalized inequalities
- Separating and supporting hyperplanes
- Dual cones and generalized inequalities

Source: Convex Optimization — Boyd & Vandenberghe

Affine set

line through x_1, x_2: all points
$x = \theta x_1 + (1 - \theta)x_2 \quad (\theta \in R)$



affine set: contains the line through any two distinct points in the set
example: solution set of linear equations $\{x \mid Ax = b\}$

If we've got a point x_1 here and a point x_2 – when θ is zero, we're sitting right at x_2. When θ is one, we're sitting at x_1. It traces out the whole line like this.

An affine set can be described as a set where any two distinct points and the line through them are all inside the set.
So take any two points from an affine set, trace the line through them, and that line better be in the same set.

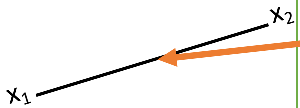
One example of an affine set is a solution set of linear equations, which you've all seen by now. There are lots of different ways of writing it. We can write it with a parameter.

We can write it using an equation on the left-hand side here, this is one of the ways you can represent any affine set

Convex set

line segment between x_1 and x_2: all points with $0 \leq \theta \leq 1$

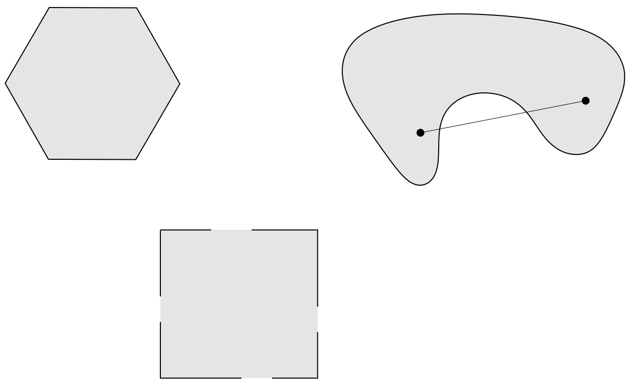
$x = \theta x_1 + (1 - \theta)x_2$



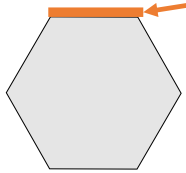
convex set: contains line segment between any two points in the set

$x_1, x_2 \in C, \; 0 \leq \theta \leq 1 \Rightarrow \theta x_1 + (1-\theta)x_2 \in C$

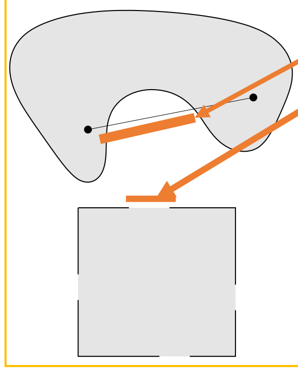
examples (one convex, two nonconvex sets)



Convex set is a set which takes any two points in the set – x_1 and x_2 and the line segment between those two points be in the set. That means that an affine set is a convex set, because an affine set contains the entire line through x_1 and x_2. It does contain the line segment between x_1 and x_2. first example of a convex set is the affine set from before. take x_1 and x_2 out of this convex set C. Vary between zero and one, and the convex combination of points here has to lie in the set C.



This one more technical example. This one looks like it should probably be convex but for some of the points where the boundary is missing. If I had picked a point here and a point here, this line actually is not inside the set, because it's only got a partial boundary around here.



Following on from convex combinations, before, we had just two points involved in our determination of whether a set was convex. We can actually extend this concept to a convex combination.

Convex combination and convex hull

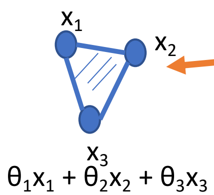
Convex combination of $x_1, \ldots, x_k$: any point x of the form

$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

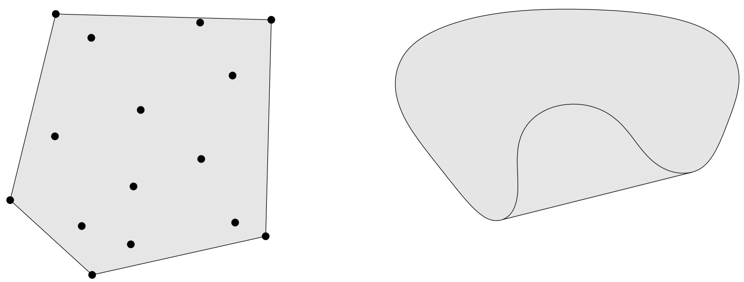
with $\theta_1 + \cdots + \theta_k = 1$, $\theta_i \geq 0$

Convex hull conv S: set of all convex combinations of points in S

If we picked points x_1 through x_k to start with, we can describe a convex combination of those points as being a set where we weight each of the points, x_1 through x_k, and the weights all have to be nonnegative and they have to add up to one.



If we had three points like this, we'll see that the set of all convex combinations actually looks something like this. x_1, x_2 and x_3, and you always have them nonnegative and you vary them so that they sum to one, the set of points that the sum traces out is called the convex hull of the set.



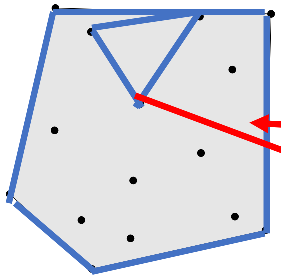
Convex combination and convex hull

Convex combination of $x_1, \ldots, x_k$: any point x of the form

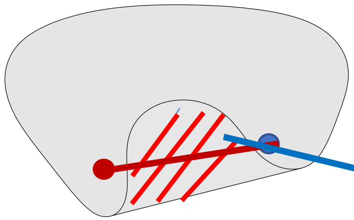
$$x = \theta_1 x_1 + \theta_2 x_2 + \cdots + \theta_k x_k$$

with $\theta_1 + \cdots + \theta_k = 1$, $\theta_i \geq 0$

Convex hull conv S: set of all convex combinations of points in S



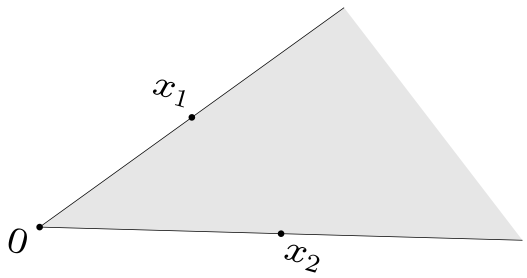
Supposed these black points and what is the convex hull of these points? It's formed by basically taking the points at the edges of the set and adding the line segments involved in those points. Then, you fill in the rest of the set. You could form it as a convex combination between these three points here. For any set, we can find the convex hull of that set.



Find the convex hull of this set, we need to just basically fill in this region here. That's the problematic region before. If we found two points like this, the line segment did not line entirely within the set.

Convex cone

Conic (nonnegative) combination of x_1 and x_2 : any point of the form $x=x_1\theta_1+ x_2\theta_2$; $\theta_1\geq 0$, $\theta_2\geq 0$



Convex cone: set that contains all conic combinations of points in the set

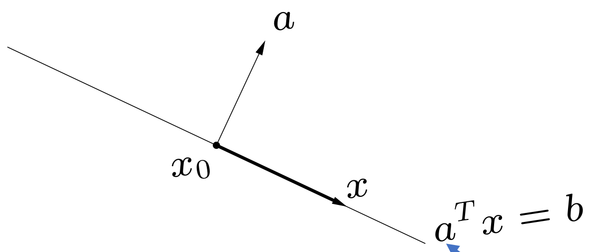
Let's have a look at convex cones now. Let's introduce a conic combination of points. This is again a specialized version of a linear combination of points. We take two points. We can have more than two. We have nonnegative weights on these points. That's described as a conic combination of these two points, x_1 and x_2.

Here, we have a figure where we have a point x_1 and a point x_2, and this is the set of points $\theta_1 x_1$ plus $\theta_2 x_2$ where θ_1 and θ_2 are nonnegative.

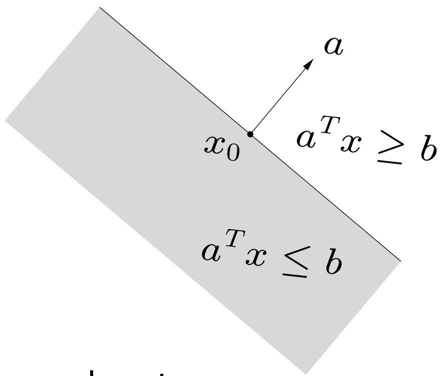
Obviously, it's a cone, and that's why we chose the name convex cone. It also has to be convex. The reason for that is if we look back to our definition of convex, there was a special case of these. We added another constraint on top of this. If we merely wanted to insure convexity, we'd only have to check that x was in this cone when θ_1 and θ_2 added up to one. Any convex cone has to be a convex set.

Hyperplanes and halfspaces

Hyperplane: set of the form {x| $a^T x = b$} ($a \neq 0$)



Halfspace: set of the form {x| $a^T x = b$} ($a \neq 0$)



- a is the normal vector
- hyperplanes are affine and convex;

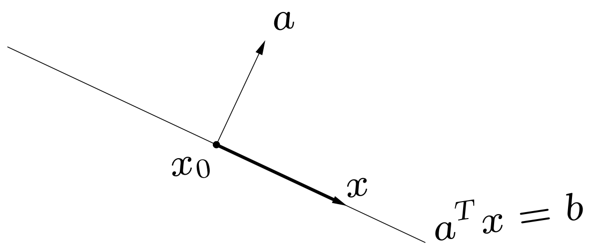
halfspaces are convex

Look at hyperplanes and halfspaces. A hyperplane is a set where a transpose x is equal to some constant b. a is nonzero, because otherwise, we could have any point in the whole series. This is an example hyperplane. We parameterize it with a point, say, x_0 and then we need to know that a transpose x_0 is equal to b, and then any other point that also has an inner product with a of b is also lying on this hyperplane

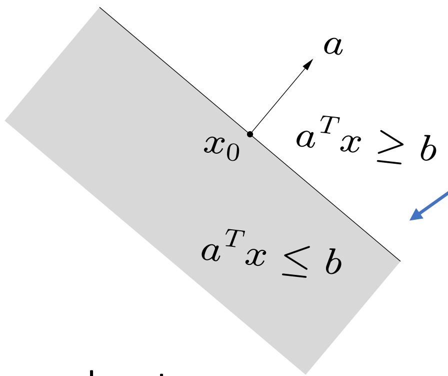
This is a hyperplane in two space, which is a line. In three space, you can think of it as looking something like this here. This would be a hyperplane in three space. It need not go through the origin, by the way. Following on from that, we have a halfspace, and this is similar to a hyperplane, but we just take the area on one side of that plane. We can have it either open or closed. This can be a strict inequality or a non-strict inequality.

Hyperplanes and halfspaces

Hyperplane: set of the form {x| $a^T x=b$} ($a \neq 0$)



Halfspace: set of the form {x| $a^T x=b$} ($a \neq 0$)



That's a hyperplane and a halfspace. They are both convex. A hyperplane is also affine, which ensures that it's convex, but a halfspace is merely convex and it's not affine. It's not affine because we would have to be able to contain the line through any two points. Let's pick these two points here and then draw a line between them.

- a is the normal vector
- hyperplanes are affine and convex;

halfspaces are convex

Euclidean balls and ellipsoids

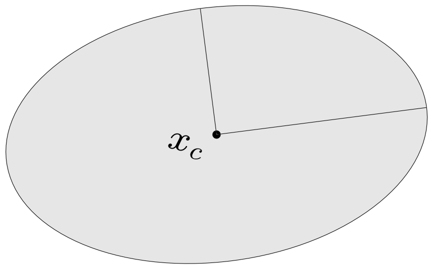
(Euclidean) ball with center x_c and radius r:

$$B(x_c, r)=\{x \mid \|x-x_c\|_2 \le r\}=\{x_c+ru \mid \|u\|_2 \le 1\}$$

Elliisod: set of the form

$$\{x \mid (x-x_c)^T P^{-1} (x-x_c) \le 1\}$$

with $P \in \boldsymbol{S}^n_{++}$ (i.e., P symmetric positive definite)



Other representation: $\{x_c + Au \mid \|u\|_2 \le 1\}$ with A square and non-singular

First of all, a Euclidean ball. Euclidean is going to mean that we use the two norm in the ball description. We parameterize a ball by picking a center point x_c and a radius for the ball of r. It's the set of any point that is in the Euclidean sense to x_c then r. Here's one way of expressing that. Here's another way of expressing that. This is a free parameter representation where you have a ball u and it's just a unit ball. You scale it by r, and then you center it around x_c

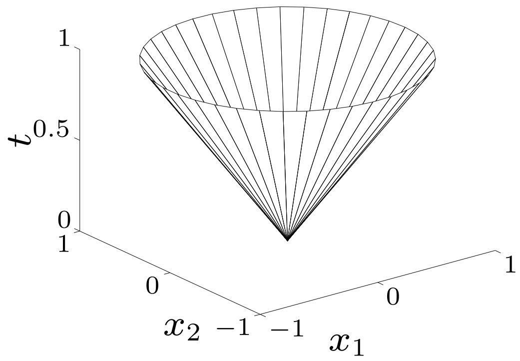
Take an ellipsoid, which is a generalization of the Euclidean ball, and that's points – we take a symmetric positive definite matrix P and we use this expression here. Ellipsoids and Euclidean balls are convex sets. One of the ways you could prove it is you could find two points and show that if those two points are in the set, then their convex combinations have to be in the set as well.

Norm balls and norm cones

Norm: a function $\| \cdot \|$ that satisfies
- $\|x\| \geq 0$; $\|x\| = 0$ if and only if $x=0$
- $\|tx\| = |t| \, \|x\|$ for $t \in R$
- $\|x+y\| \leq \|x\| + \|y\|$

notation: $\| \cdot \|$ is general (unspecified) norm;
$\| \cdot \|_{symb}$ is particular norm

Norm ball with center x_c and radius r: $\{x \mid \|x - x_c\| \leq r\}$
Norm cone: $\{(x, t) \mid \|x\| \leq t\}$
Euclidian norm cone is called second-order cone
norm balls and cones are convex

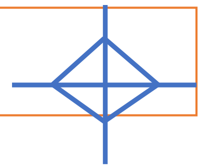


We simply have the norm symbol without any subscript, it does not necessarily mean the two norm, and in fact, it's usually used to represent a general norm.

We can define as we saw previously similar to the Euclidean ball here, for general norms, we can define a norm ball. This ball doesn't necessarily look like a ball. It's a ball in the general sense.

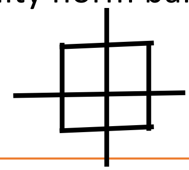
$\|x\|_1$ -->We have the one norm, which is the sum of absolute values.

$\|x\|_2 \rightarrow$ norm ball with the one norm looks like in R^2.
it's a diamond



We can also define a norm cone by adding another variable here and saying the norm cone is any point x, t where the norm of x is less than t. This is the Euclidean norm, and it's the norm contrast out. This is actually a solid figure. It might look like a wire frame, but it's actually filled in. Norm balls and norm cones is that they're convex.

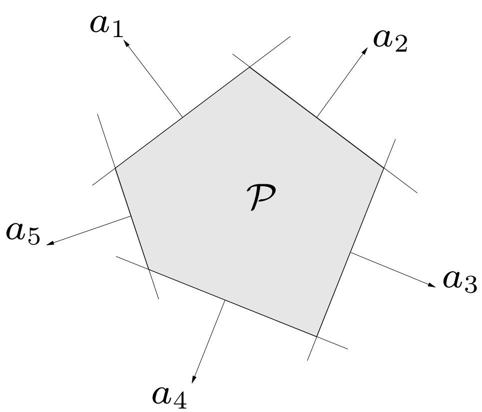
$\|x\|_\infty \rightarrow$ this the infinity norm ball.
Looks like a square



Polyhedra

Solution set of finitely many linear inequalities and equalities $Ax \preceq b$ $Cx = d$ $\{x \mid Ax \preceq b, \quad Cx = d\}$

($A \in R^{m \times n}$, $C \in R^{p \times n}$ $\preccurlyeq$ is component wise inequality)



Polyhedron is intersection of finite number of halfspaces and hyperplanes

A polyhedron is the solution set of finitely many linear inequalities and equalities. We can say a polyhedron is any x such that Ax is less than B and Cx equals d. That would be a polyhedron. We note here that there's a slightly different symbol for less than here, this curvy one.

The reason is because this here is what we got from R. If we say three is less than or equal to five, this is what we use. Here, AX, which is a vector, is component wise less than B, and to make this completely explicit, we're using this slightly different symbol here

Positive semidefinite cone

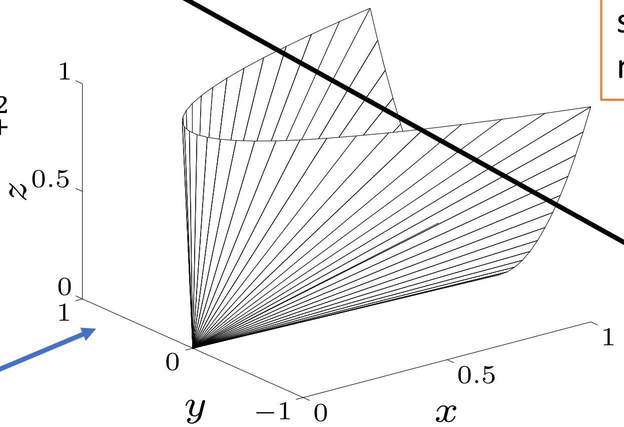
Notation:
- S^n is set of symmetric n × n matrices
- S_+^n ={X∈S^n|X≽0}: positive semidefinite n×n matrices

  X∈ $S_+^n \Longleftrightarrow$ $z^TXz \geq 0$ for all z

S_+^n is a convex cone

- S_{++}^n ={X∈S^n|X≽0}: positive semidefinite n×n matrices

Example: $\begin{pmatrix} x & y \\ y & z \end{pmatrix} \in S_+^2$



Look at the positive semidefinite cone. We say positive semidefinite matrices by now. S^n is the set of symmetric matrices. Suppose this is a convex set. How would we go about proving it? Here is convex if each of these things here is also in the set.
X∈ S^n $\boldsymbol{\theta}$≥0, Y∈ S^n $\boldsymbol{\theta}$≤1 and $\boldsymbol{\theta}$X+(1- $\boldsymbol{\theta}$)Y
Is that true for a symmetric matrix if you take a sum of symmetric matrices? Is it still symmetric? Yes, it is. That means that this is a convex set.

S_+^n set of positive semidefinite N by N matrices. We'll assume that they're symmetric. quadratic form $z^TXz \geq 0$ for all z, claim that S_+^n is a convex cone
X≥0, Y≥0 ; $\boldsymbol{\theta}_1$X+ $\boldsymbol{\theta}_2$Y; nonnegative combination of two of these items from the set, it also needs to be in the set. $z^TXz= z^T(\boldsymbol{\theta}_1X+ \boldsymbol{\theta}_2Y)z= \boldsymbol{\theta}_1z^TXz+ \boldsymbol{\theta}_2z^TYz$ - both are non-ve. The result is +ve semi definite

You could find out its x, y, and z coordinate plug it into the matrix and you see what the minimum value is. If it's ≥ zero, then it's in this cone. If it's < zero, it's outside the cone. If you did a grid search in here, you could build up a set of points which are inside the cone.

Operations that preserve convexity

Practical methods for establishing convexity of a set C

1. **Apply definition**

$$x_1,\ x_2 \in C,\ 0 \le \theta \le 1 \Longrightarrow \theta x_1 + (1-\theta) x_2 \in C$$

2. **Show that** C is obtained from simple convex sets (hyperplanes, halfspaces, norm balls, . . .) by operations that preserve convexity

- linear-fractional functions
- perspective function
- affine functions
- intersection

3. **Math Lab approach**

If a given set C is convex, we can just apply the definition. Sometimes, that'll be easy. Sometimes, it won't be easy. Another way we can do it is we can take operations that are known to preserve convexity. This is actually any number of intersections, and we look at that on the next slide. An affine function will preserve convexity. If we take a set C, put it through a function f which is affine, it's going to preserve convexity.

This is actually any number of intersections, and I think we look at that on the next slide. An affine function will preserve convexity. If we take a set C, put it through a function f which is affine, it's going to preserve convexity.

Say I want to know if a set is convex, and whether a given point is in the set or not in the set. Here's what I can do. I can pick a couple of points, say, x_1 and x_2 completely at random from the space. Then, I can see if x_1 and x_2 are in the convex set. Suppose they are. The requirement of convexity is that when I find this convex combination of these two points, that point itself has to lie in the convex set. If somebody gives you some random set. You make a tiny little window and run MatLab and you start generating random points from the set. You take a convex combination and test if it's still in the set. Now, as soon as you find some θ which is between zero and one where $\theta x_1+(1-\theta)x_2$ is not in the set and x_1 and x_2 are, it's not convex.

Intersection

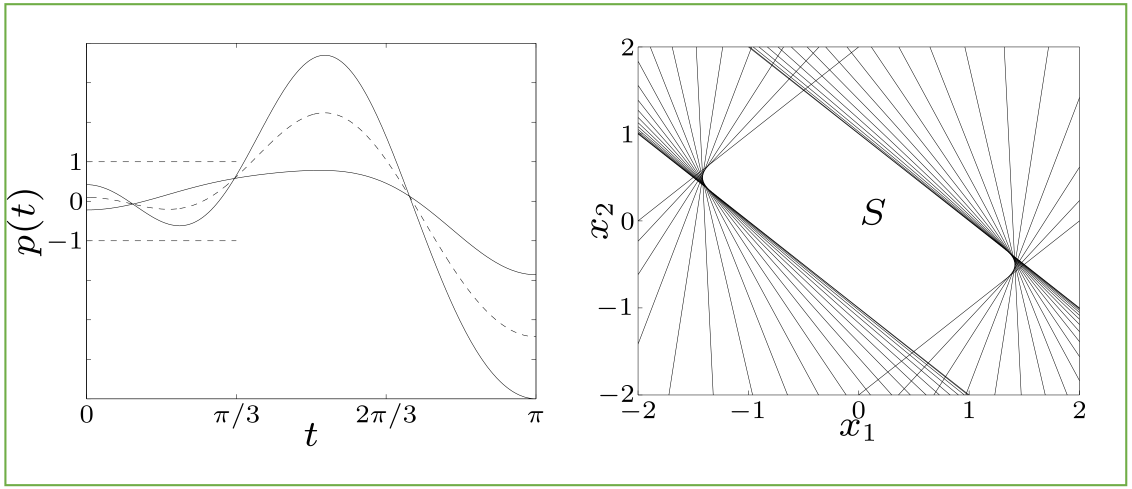
Intersection of (any number of) convex sets is convex

Example: $S=\{x \in R^m \mid |p(t)| \leq 1 \text{ for } |t| \leq \pi/3\}$

Where $p(t)=x_1 \cos t+x_2 \cos 2t+\cdots+x_m \cos mt$

For m=2

Let’s have a look in more detail at intersections. Here’s the claim – the intersection of any number of convex sets is convex. When I say any number, that means finite, infinite and unaccountably infinite. This is a very robust way of finding out whether something’s convex.



Here’s an example. Here, it’s an unaccountably infinite set. We just take this random thing here and we know that the individual elements are convex. We put them all together and we put up with this shape here.

Affine function

Suppose f: $R^n \rightarrow R^m$ is **affine** ($f(x)=Ax+b$ with $A \in R^{m \times n}$, $b \in R^m$)

- the image of a convex set under f is convex

  $S \subseteq R$ convex $\Longrightarrow f(S) = \{f(x) \mid x \in S\}$ convex

- the inverse image $f^{-1}(C)$ of a convex set under f is convex

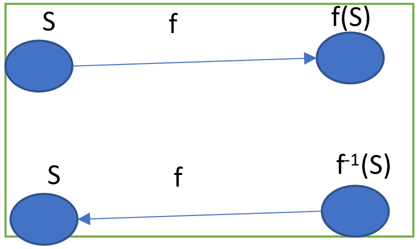
  $C \subseteq R$ convex $\Longrightarrow f^{-1}(C) = \{x \in R^n \mid f(x) \in C\}$ convex

Examples

- Scaling, translation, projection
- Solution of linear inequality $\{x \mid x_1A_1+x_2A_2+\ldots+x_mA_m \preceq B\}$ {with $A_i, B \in S^p$ }
- Hyperbolic cone $\{x \mid x^TPx \leq (c^Tx^2)\ c^Tx \geq 0\}$ with PS_+^n

Two simple examples are **scaling** and **translation**. If $S \subseteq R^n$ is convex, $\alpha \in R$, and $a \in R^n$, then the sets αS and $S + a$ are convex, where $\alpha S = \{\alpha x \mid x \in S\}$ (**scaling**), $S + a = \{x + a \mid x \in S\}$ **(translation)**

Suppose we have some function f that takes R^n into R^m, and it's affine. One way of looking at this is saying that if f(x) is Ax + b, A is a matrix and b is a vector.



Suppose set S is a convex set. Take this under a mapping f to some other set f of S. This is a convex set if f is an affine mapping.

Likewise, we can do the same for an inverse mapping. If we had another set but it could be any set. We can take this back via f and f is a convex set. After this mapping, it's also a convex set.

Projection could be really simple. We might have a set $(x_1, x_2) \in C$. This might be a point in a convex set C, and if this is convex, $\{x_1 \mid (x_1, x_2) \in C\}$ then the set of x_1 given that the sum of x_1 – for sum x_2, the set of just x_1's might be a convex set. That's a very simple prediction.

Affine function… contd.

$f(x) = B - A(x)$
$A(x) = A_1x_1 + \ldots + A_mx_m$
$f^{-1}(S_+^n) = \{x \in R^m \mid B - A(x) \in S_+^n\}$ equivalent to $B - A(x) \geq 0$
i.e., $B \leq A(x)$

A linear matrix inequality (LMI) is an affine mapping. If we have a set of matrices A(x) and they're all symmetric matrices, then the set x, where this weighted sum of symmetric matrices and the positive semidefinite sense is less than B, that is also a convex set.
S_+^n, semi positive definite matrix, is a convex set. If we take the cone – this function f of x – we'll make it equal to B minus A of x where A of x is $A_1x_1 + \ldots + A_mx_m$. This is an affine function, and it needs a little bit of hammering to get it in the same form as we saw above here. It is an affine function.

$\{(z,u) \mid \|z\| \leq u, u \geq 0\}$

$$f(x) = \begin{pmatrix} P^{1/2}x \\ C^T x \end{pmatrix}$$

Another one here is the hyperbolic cone.
We already saw that the set z, u, which is a cone when we restrict u. This is a norm cone, and we showed that this was a convex set. We know this is convex. You can see that if we apply the affine function f of x equal to P one half x and C transpose x, the inverse image of this mapping is also a convex set.

Perspective and linear-fractional function

Perspective function P : $R^{n+1} \rightarrow R^n$

$$P(x,t) = x/t, \text{ dom } P = \{(x,t) \mid t > 0\}$$

Images and inverse images of convex sets under perspective are convex

Linear-fractional function f : $R^n \rightarrow R^m$

$$f(x)=\frac{Ax+b}{c^Tx+d}, \text{ dom } f =\{x \mid c^Tx+d>0\}$$

Images and inverse images of convex sets under linear-fractional function are convex

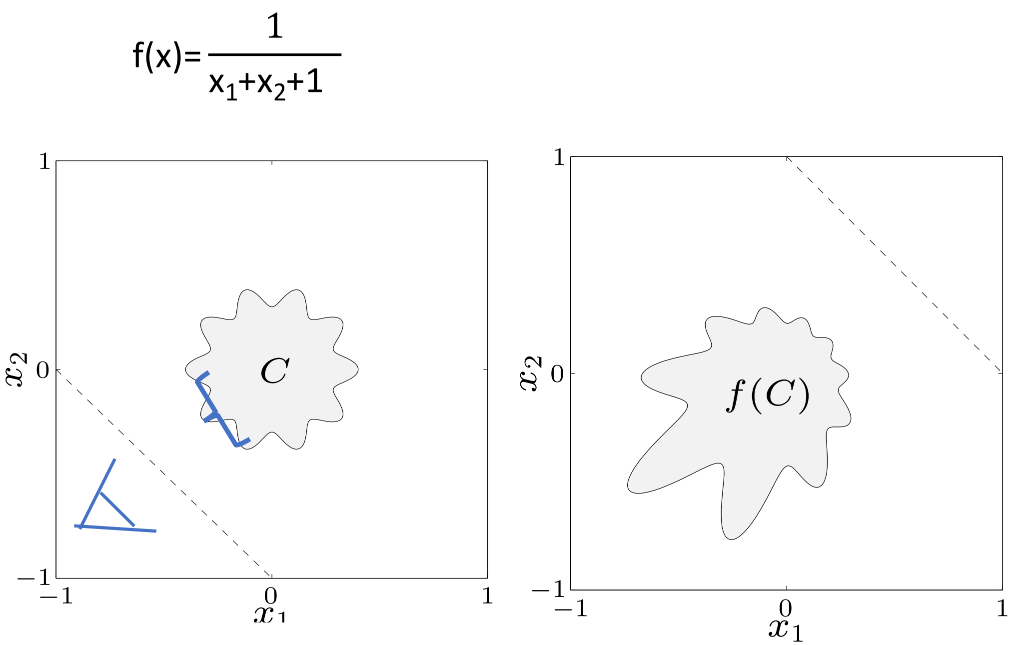
If we have some function P that reduces R^{N+1} to R^N by taking the last component and dividing – you can think about it as taking some vector like this and we pick off the last component. We divide each of these elements by that component and then we throw this away. It's a somewhat unusual operation, but one, two, three, I'd pick the three, and my new vector would be one third two thirds. That's a perspective function

$$\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \end{bmatrix}$$

A generalization of the perspective function is when we have a linear-fractional function. If we have some matrix A, a vector c and vectors b and d, then again, under both the image and the inverse images of a convex set under this transformation are themselves convex. There is a domain constraint on positivity of the denominator here as well

Perspective and linear-fractional function...contd.



Here's a linear-fractional function, which you can think of as similar to a perspective function or a generalized perspective function. If we hold this figure on the side – if I hold it like this and I look at this figure from a different angle, I'll find that these parts here are closer to my eye. If I look at this, I'll find that these things loom closer in my vision than the other parts of the figure. This is quite a strange transformation, and C here is obviously not convex. In fact, why is C not convex?

The line segment is very much not in C. If this were a convex set, though, under this transformation, it would remain convex. The reason it's not is because you wouldn't be able to see it too well if it wasn't convex. what the dotted line is about. probably what it is – it's probably a line under the same transformation. That's what it is.

Generalized inequalities

A convex cone $K \subseteq R^n$ is a **proper cone** if

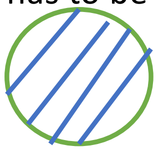
- K is closed (contains its boundary)
- K is solid (has nonempty interior)
- K is pointed (contains no line)

Examples

- nonnegative orthant $K = R_+^n = \{x \in R^n \mid x \geq 0, i=1,\dots,n\}$
- positive semidefinite cone $K = S_+^n$
- nonnegative polynomials on [0, 1]:

$K = \{x \in R^n \mid x_1 + x_2 t + x_3 t^2 + \cdots + x_n t^{n-1} \geq 0 \text{ for } t \in [0,1]\}$

Let's have a look at generalized inequalities. First of all, the cone has to be closed.



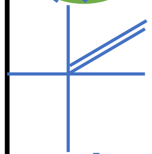
Closed means is if we have a shape and the set includes the border of that shape, then it's a closed set. If it doesn't include the boundary at every point, then it's not a closed set.

That's the first requirement to make a convex cone into a proper cone.

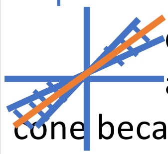
We also say that it's solid. It actually has a nonempty interior

Again, interior is not a concept. You can think of interior from an intuitive point of view as the inside of the set. The set not including its boundary.

If I just have this ray here, is that a proper cone? Why not?



This has an empty interior event though it's getting thicker by the second, but if I just have a ray out from the origin, it has no interior, so it's not a proper cone.



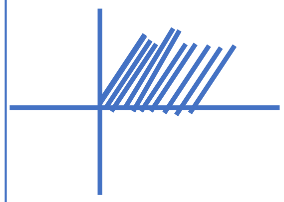
There's one more requirement, and that is that the cone cannot contain a line. That means that this set here – this is a cone, and it's also a convex cone, but it's not a proper cone because it contains this line here through the origin.

Generalized inequalities... contd.

Examples

- nonnegative orthant K= R_+^n ={x∈R^n |x ≥0, i=1,...,n}
- positive semidefinite cone K = S_+^n
- nonnegative polynomials on [0, 1]:

K={x∈R^n|x_1 +x_2t+x_3t^2 +⋯+x_nt^{n-1} ≥0 for t∈[0,1]}



Here are some examples of proper cones. First of all, the one you're most familiar with is the non negative orthant. If I have R^2, for example, the nonnegative orthant is this portion of R^2. That's a proper cone. It's convex. It's closed because it includes these boundaries. It's solid because there's certainly plenty of stuff in the interior. It's pointed because it contains no line, and so it's a proper cone.

Here's another example. The positive semidefinite cone is. That's probably slightly trickier to prove, but all of these concepts generalize into this cone here. They're not a problem. Here's another one, the set of nonnegative polynomials. We're going to find that this (second one) is extremely useful and used all the time. We're going to use this a lot. This one is a more interesting example, but we probably won't see it all that often. Still, it's a proper cone.

Generalized inequalities… contd.

Generalized inequality defined by a proper cone K

$x \preccurlyeq_K y \Leftrightarrow y - x \in K, \quad x \prec_K y \Leftrightarrow y - x \in \textbf{int}\, K$

Examples

Componentwise inequality ($K = \boldsymbol{R}_+^n$)

$x \preccurlyeq_{R_+^n} y \Leftrightarrow x_i \leq y_i$, i=1,…,n

matrix inequality ($K = S_+^n$)

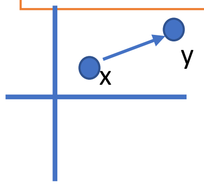
$X \preccurlyeq_{S_+^n} Y \Leftrightarrow Y - X$ positive semidefinite

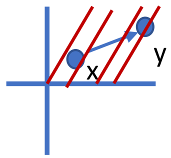
these two types are so common that we drop the subscript in $\preccurlyeq_K$

properties: many properties of $\preccurlyeq_K$ are similar to ≤ on R, e.g., $x \preccurlyeq_K y,\ u \preccurlyeq_K v \Longrightarrow x + u \preccurlyeq_K y + v$

These are so common that we drop the subscript. Not only that, but we often also drop the special symbol and just use an ordinary symbol like this ≤

Look now at generalized inequalities. Before, we had this symbol with the curvy lines. Now we add a K to it. If we take a proper cone K, we can parameterize a generalized inequality with this proper cone K. We describe it as being less than Y. If y-x is inside this cone K, we can strict inequalities in a similar way by saying that x is strictly this if y-x lies in the interior of the cone K.





Look at a quick example. The component wise inequality that we're familiar with – this says that x <y in the component wise sense if each of the components $x_1 \leq y_1$. Let's take a couple of points – this one here and this one here. Let's call this x and this y. x is described as being ≤ y in this proper cone K if y minus x lies in the cone.

If we had a vector 3,3 and a vector 1,1. Subtract them and we get 2,2. It points to y, and this is inside this proper cone here. That's a component wise inequality. We can say the same about a matrix inequality. We've seen these before. x less than y is this PSD cone here means that y minus x is positive semidefinite

Minimum and minimal elements

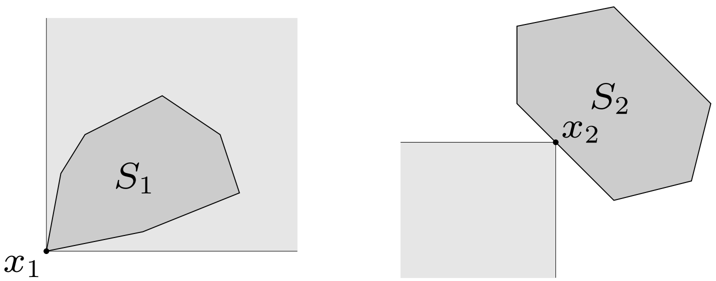
$\preccurlyeq_K$ is not in general a linear ordering: we can have $x \npreceq_K y$ and $y \npreceq_K x$

$x \in S$ is the **minimum element** of S with respect to $\preccurlyeq_K$ if $y \in S \Longrightarrow x \preccurlyeq_K y$

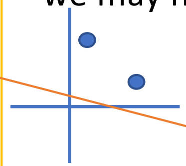
$x \in S$ is the **minimal element** of S with respect to $\preccurlyeq_K$ if $y \in S$, $y \preccurlyeq_K x \Longrightarrow y=x$

Example ($K = \boldsymbol{R}_+^2$)

x_1 is the minimum element of S_1
x_2 is the minimal element of S_2



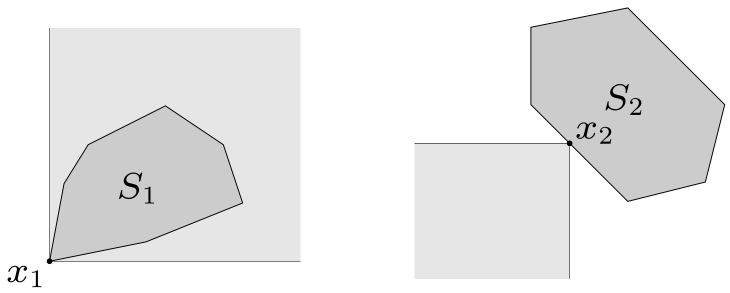
One of the things that doesn't carry over is we don't have a linear or a total ordering. That means that if we find x and y, x may be neither less than y or greater than y. Here's another way of expressing that. This generalized inequality here is not in general a linear ordering, we may have neither of these holding.

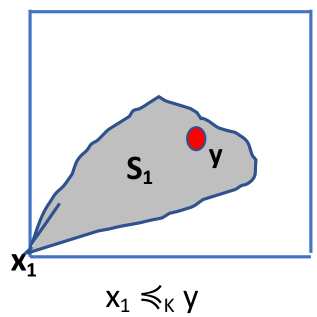


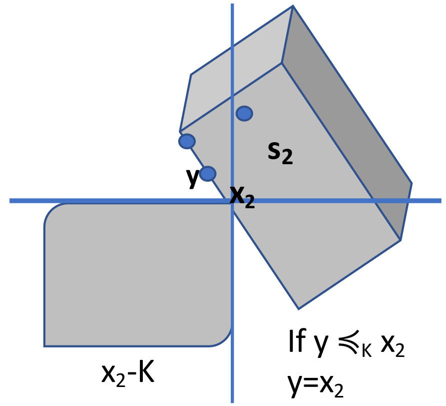
If I have a point here and a point here, then this point in the sense of component wise is neither less than this point nor is it more than this point. It's not a total ordering

X in a set is described as being the minimum element with respect to this generalized inequality if any other point inside the set is actually more than X in this case. That's one concept, and we'll have a look at a picture in a second. A second concept is a minimal element. We describe an element x as being minimal if any point that is less is **actually the same point**. Let's have a look at that.

Minimum and minimal elements







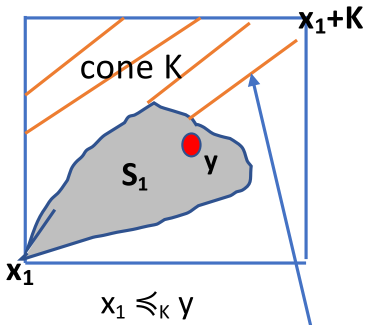
We're going to use the cone K is equal to $\boldsymbol{R}_+^2$, so component wise inequality in $\boldsymbol{R}^2$. This point x_1 here is the minimum element of S_1.

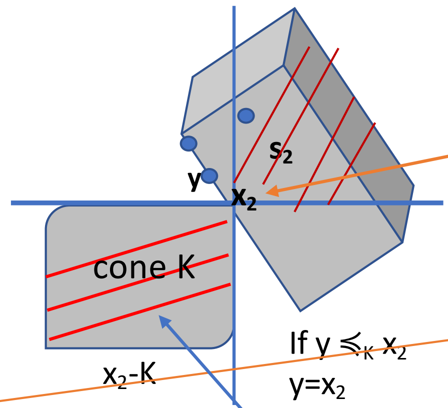
Let's have a look at the difference here in S_2. What we need here is that I picked x_2 and S_2. x_2 is a minimal element.

What that means is if I picked any other point y in S, then I can explicitly say that $x_1 \leq y$. That will hold for any point in here

if I pick some other point here, I know that if y is $\leq x_2$, then it means that y is equal to x_2.

Minimum and minimal elements





Here's a point, for example. Make that y. x_2 is not the minimum element, because for this way, I cannot write $x_2 \cancel{\preccurlyeq_K} y$ I can't actually do that because it's not a linear ordering. That's not allowed. But I can say that any point that is less is the same point. That's because -

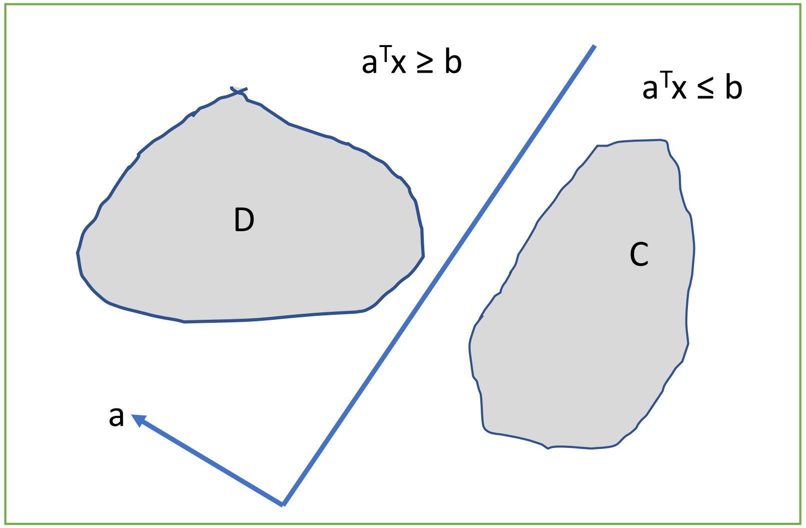
In this set, if I drew the same figure here, these are all of the points that are unambiguously more than x_2. This point here is not – this point y is not in the set, so I can't say that x_2 is unambiguously less than y.

-Here's another way of thinking about it. If I take this cone K, this set here is the set x_1+K, and that's the set of all points that are unambiguously more than x_1. They can always be described in reference to the cone K as being more than x_1 . x_1 is unambiguously least.

What I can say is I can form this other set x_2-K, and this is the set of points that are unambiguously less than x_2. Any point in here, I can say with full authority – this is a point that's less than x_2. That means because I find that there are no points inside the set S_2 that are unambiguously less, I can describe x_2 as being a minimal element.

Separating hyperplane theorem

If C and D are nonempty disjoint convex sets, there exist $a \neq 0$. s.t $a^T x \leq b$ for $x \in C$, $a^T x \geq b$ for $x \in D$

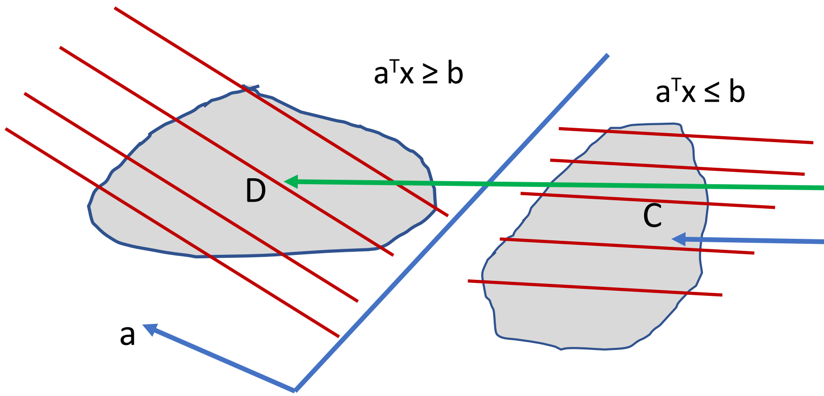


the hyperplane $\{x \mid a^T x = b\}$ separates C and D
Strict separation requires additional assumptions (e.g., C is closed, D is a singleton)

This one is the separating hyperplane theorem, and it says that if we take two general convex sets, C and D, and if they're disjoint – they have no intersection and C intersect D is the empty set. If C and D are disjoint, then there is some a, not zero, and there's b so that the halfspace $a^T x$ in C lies in a halfspace $a^T x \leq b$. The opposite holds for x and D

Strict separation requires more, so I won't go into the details here, but strict separation would be the hyperplane that actually passes through neither set. We can use this separating hyperplane theorem and turn it into another one called the supporting hyperplane theorem.

Separating hyperplane theorem



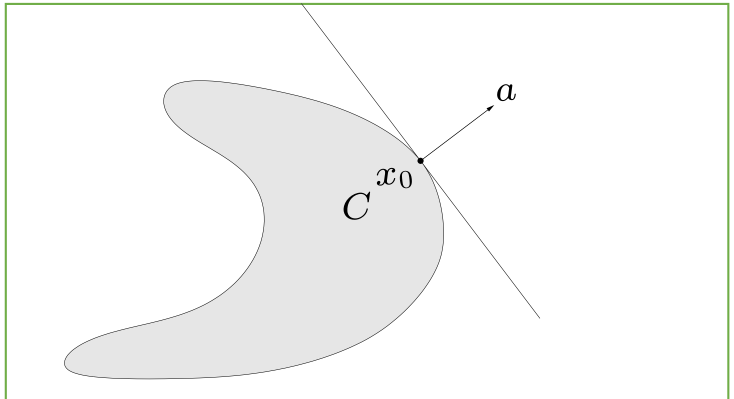
Here's a couple of sets. Here is a hyperplane between the sets, and if F is in R^2, then it's a line. You can also think a prediction of R^2. This says that we can find some hyperplane so that D lies entirely in this halfspace here, and C, if you like, lies entirely in this halfspace here. It's called a separating hyperplane for quite obvious reasons. This theorem says that any two sets like this that are disjoint – there has to be a separating hyperplane.

Supporting hyperplane theorem

Supporting hyperplane to set C at boundary point x_0 :

$$\{x \mid a^T x = a^T x_0\}$$

Where $a \neq 0$ and $a^T x \leq a^T x_0$ for all $x \in C$



Supporting hyperplane theorem: if C is convex, then there exists a supporting hyperplane at every boundary point of C

A supporting hyperplane we define as being one where – this is described as being a supporting hyperplane if the only point that it touches is – if it goes through x_0 and all of the set is on one side. You can say that if C is a convex set, then at every boundary point of C, there will be a supporting hyperplane.

Connect these two new ideas together. We have a separating hyperplane theorem which says if we have two sets and they're disjoint, we can find the hyperplane that separates them, and one which says that if we have a convex set and a boundary point, we can find a supporting hyperplane. We can actually use the separating hyperplane theorem to prove this one. We can actually say – we'll form a set x_0, and we'll also form a set interior of C

We take a convex set C. We'll assume that it's convex by looking at this part of it. We'll say that x_0 is one set and it's convex because it's just a single point. The interior of C is a convex set if C is convex. The interior of C and x_0 do not intersect if x_0 is on the boundary, and then we can find a separating hyperplane between x_0 and the interior of the set. That will be a supporting hyperplane.

Dual cones and generalized inequalities

Dual cone of a cone K:

$K^* = \{y \mid y^T x \geq 0 \text{ for all } x \in K\}$

Examples:

K= $\boldsymbol{R}_+^n$: $K^* = \boldsymbol{R}_+^n$

K= $\boldsymbol{S}_+^n$: $K^* = \boldsymbol{S}_+^n$

K= $\{(x,t) \mid \|x\|_2 \leq t\}$: $K^* = \{(x,t) \mid \|x\|_2 \leq t\}$

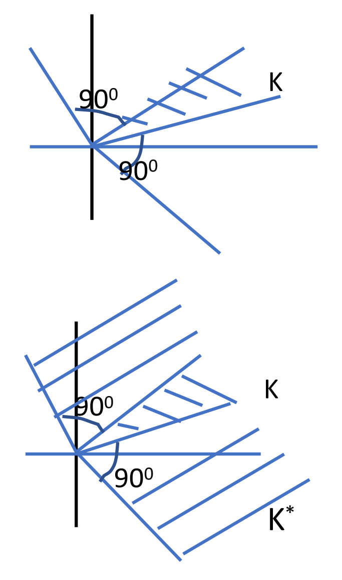
K= $\{(x,t) \mid \|x\|_1 \leq t\}$: $K^* = \{(x,t) \mid \|x\|_\infty \leq t\}$

first three examples are self-dual cones

dual cones of proper cones are proper, hence define generalized inequalities:

$y \succcurlyeq_{K*} 0 \iff y^T x \geq 0 \text{ for all } x \succcurlyeq_K 0$

Dual cone is defined for any cone K – it doesn't have to be convex. It surely doesn't have to be proper. We define the dual cone, K*, as equal to any point y with inner product ≥ 0 with a point x in the cone K.



Let's have a look and see what that means. If I have a cone that looks something like this – I'll call this K. It happens to be proper, but that's fine. K* in this case is any vector y where $y^T x$ for all x and K is nonnegative. Let's take this line here. Draw out a right angle here. That's a right angle right there.

I take this one here and I do a similar thing. This set here is the dual cone of K. This one here is K*. That's the set of – if I take any vector in here, its angle to all elements in the set here is less than or equal to 90 degrees.

If I look at this vector here or any vector in this dual cone K*, its angle to all elements of K is less than or equal to 90 degrees. This is the dual cone.

Dual cones and generalized inequalities

Dual cone of a cone K:

$K^* = \{y \mid y^T x \geq 0 \text{ for all } x \in K\}$

Examples:

K= $\boldsymbol{R}_+^n$: K^*= $\boldsymbol{R}_+^n$

K= $\boldsymbol{S}_+^n$: K^*= $\boldsymbol{S}_+^n$

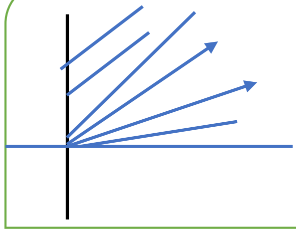
K= $\{(x,t) \mid \|x\|_2 \leq t\}$: K^*= $\{(x,t) \mid \|x\|_2 \leq t\}$

K= $\{(x,t) \mid \|x\|_1 \leq t\}$: K^*= $\{(x,t) \mid \|x\|_\infty \leq t\}$

first three examples are self-dual cones

dual cones of proper cones are proper, hence define generalized inequalities:

$y \succcurlyeq_{K*} 0 \iff y^T x \geq 0 \text{ for all } x \succcurlyeq_K 0$



A few examples – if I look at the nonnegative orthant from before – that's the set here – the dual cone of the nonnegative orthant is the nonnegative orthant itself. It's described as being self-dual.

It's described as being self-dual. You can see that quite easily if I have two nonnegative vectors y and x, y transpose x has to be greater than or equal to zero, and you should be able to go through a proof of that quite easily.

$y^T x \geq 0$

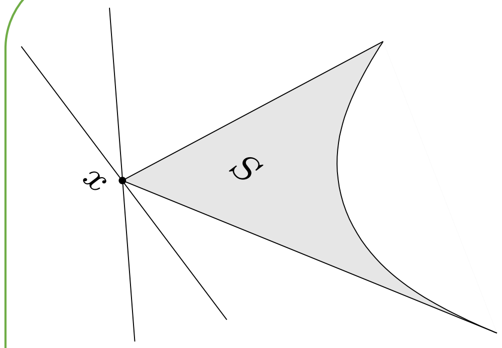
$\sum y_i x_i \geq 0$

We could just do y_i, x_i and so if the components are nonnegative, so is the product and so is the sum.

The positive semidefinite cone – that's also self-dual, so in this particular case, we actually use a slightly different definition of inner product. If you find the set here, it's going to be the set itself. Euclidean norm cone – it's self dual, but finally, we get to one that's not self dual, and in fact, most of them aren't. A lot of the useful ones are, but here's one that isn't. It's the one norm cone. The dual cone of a one norm cone is the infinity norm cone.

Minimum and minimal elements via dual inequalities

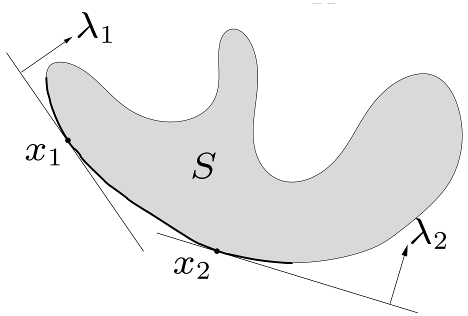
Minimum element w.r.t $\preccurlyeq_K$
x is minimum element of S iff for all $\boldsymbol{\lambda} \succ_{K^*} 0$, x is the unique minimizer of $\boldsymbol{\lambda}^T z$ over S



We talked about minimum and minimal elements before, and here is one of the ways that those come into play. We have some set S. We know nothing about the set. Let's say that x is the minimum element of S. That is true wrt the generalized inequality. x is the minimum iff for all $\boldsymbol{\lambda}$ that are strictly in the dual cone if x is the unique minimizer of $\boldsymbol{\lambda}$ transpose z over S, then x is the minimum element.

Minimal element w.r.t $\preccurlyeq_K$

- If x minimizes $\boldsymbol{\lambda}^T z$ over S for some $\boldsymbol{\lambda} \succ_{K^*} 0$, x is the minimal
- if x is a minimal element of a convex set S, then there exists a nonzero $\boldsymbol{\lambda} \succ_{K^*} 0$, such that x minimizes $\boldsymbol{\lambda}^T z$ over S



If x minimizes $\boldsymbol{\lambda}$ transpose z over S and there is some $\boldsymbol{\lambda}$ – if we can find any $\boldsymbol{\lambda}$ that's in the dual cone, then x is minimal. Finally, we can say that if x is minimal, there is some $\boldsymbol{\lambda}$ such that $\boldsymbol{\lambda}$ transpose z has a minimizer x.

Minimum and minimal elements via dual inequalities

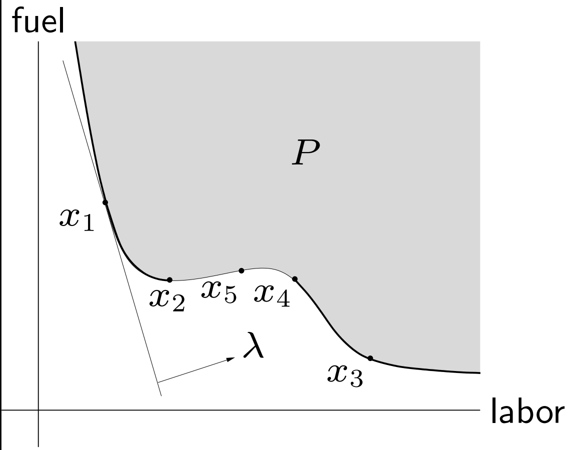
Optimal production frontier

- different production methods use different amounts of resources $x \in R^n$
- production set P: resource vectors x for

all possible production methods

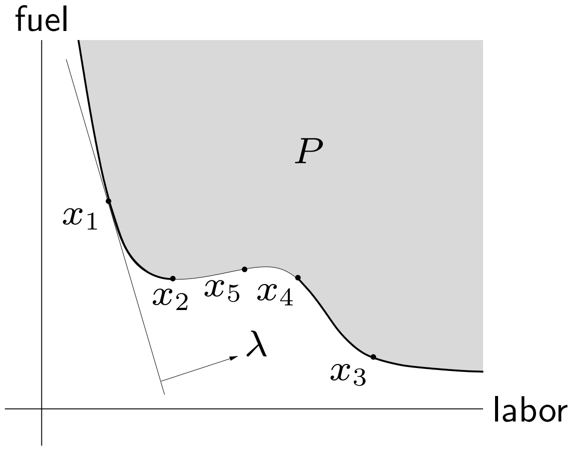
- efficient (Pareto optimal) methods correspond to resource vectors x that are minimal wrt $\boldsymbol{R}_+^n$

Example (n = 2)
x_1, x_2 ,x_3 are efficient; x_4 ,x_5 are not



In optimal production frontier, P is a set of possible production methods. On this axis, we have fuel, and on this axis, we have labor. Say we're testing out different production methods in our factor. If a point is in the set P, it says that there is some production method, which, say, produces 100 milk bottles and takes this amount of labor and this amount of fuel.

Minimum and minimal elements via dual inequalities

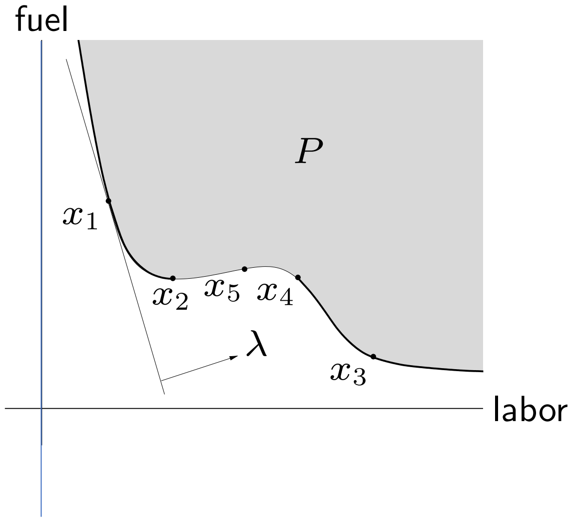


If x is in P, there is some possible production method that uses that amount of labor and this amount of fuel. This is a very simple example in R^2. If x is in P, there is some possible production method that uses that amount of labor and this amount of fuel. This is a very simple example in R^2. This set we could find.
First of all, what would it mean if we had a minimum element? Let's say we had x and P, and we're using this generalized inequality with R^+ to the N. If x was a minimum in the set, it would mean that if I had x here, say this was in the set.

The whole of P would lie within this. It would say that any production method requires more labor, more fuel or both than x. That would be x being minimal. Minimum is a slightly different quality

If I had x_1 here, this is a minimum element. What it means is that there is no method that requires simultaneously less fuel and less labor than x_1. It's efficient. Any vector along here – there is no method that is unambiguously better. There's none that simultaneous require less of both. Labor might be more expensive than fuel, so we might want to move along the curve and trade off between them, but we don't want to move off this curve. Likewise, this part of the curve is also efficient methods.

Minimum and minimal elements via dual inequalities



The problem with x_5 is we can produce at x_5, but we could produce at x_2 and we would use less labor and simultaneously less fuel. There'd be no point in picking the method x_5 if these were all the parameters involved. x_1, x_2, and x_3 are efficient, and x_4 and x_5 are not. You can think about these as being minimal with respect to R plus to the N. There are other ways. We might be interested in a different cone, and there might, in fact, be a minimum for a different set, but this is just tying at l east one of these things into a practical application.