Lecture 2

Convexity

Emil Gustavsson
Department of Mathematical Sciences
Chalmers University of Technology and Göteborg University

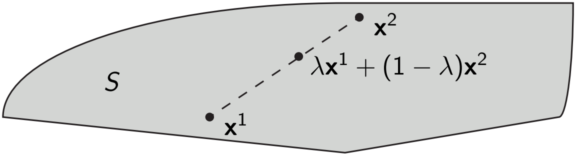
November 1, 2012

CHALMERS | GÖTEBORGS UNIVERSITET

Convex set

Let $S \subseteq \mathbb{R}^n$. The set S is **convex** if

$$\left.\begin{array}{c} \mathbf{x}^1, \mathbf{x}^2 \in S \\ \lambda \in (0,1) \end{array}\right\} \implies \lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2 \in S,$$



A set is thus convex if all convex combination of any two points in the set lies in the set. See the figure.

Examples

- The empty set is a convex set
- The set $\{\mathbf{x} \in \mathbb{R}^n \mid ||\mathbf{x}|| \leq a\}$ is convex for any $a \in \mathbb{R}$
- The set $\{\mathbf{x} \in \mathbb{R}^n \mid ||\mathbf{x}|| = a\}$ is non-convex for any $a > 0$
- The set $\{0, 1, 2, 3\}$ is non-convex

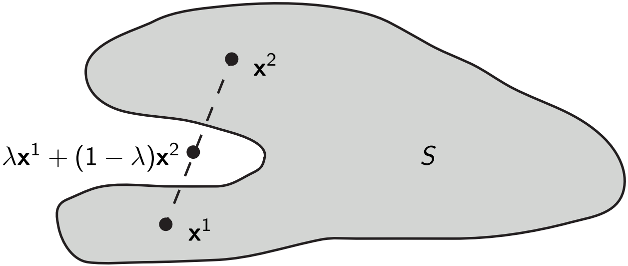


Figure: A non-convex set

Intersections of convex sets

Theorem

Let S_k, $k \in \mathcal{K}$ be a collection of convex sets. Then, the intersection $\cap_{k \in \mathcal{K}} S_k$ is a convex set.

Proof:

Affine and convex hull

We define the **affine hull** of a finite set $V = \{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^k\}$ as

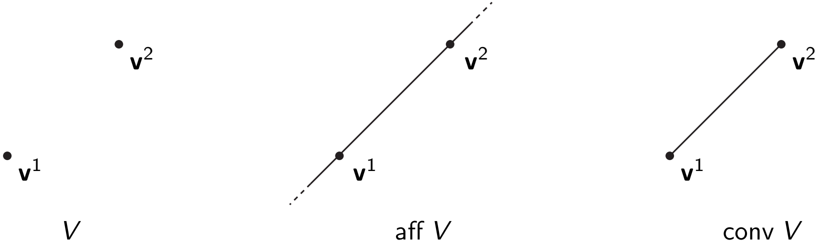
$$\text{aff } V := \left\{ \lambda_1 \mathbf{v}^1 + \cdots + \lambda_k \mathbf{v}^k \;\middle|\; \lambda_1, \ldots, \lambda_k \in \mathbb{R},\ \sum_{i=1}^{k} \lambda_i = 1 \right\}$$

We define the **convex hull** of a finite set $V = \{\mathbf{v}^1, \mathbf{v}^2, \ldots, \mathbf{v}^k\}$ as

$$\text{conv } V := \left\{ \lambda_1 \mathbf{v}^1 + \cdots + \lambda_k \mathbf{v}^k \;\middle|\; \lambda_1, \ldots, \lambda_k \geq 0,\ \sum_{i=1}^{k} \lambda_i = 1 \right\}$$

The sets are defined by all possible *affine (convex) combinations* of the k points.

Affine and convex hull



More on the convex hull

In general, we can define the **convex hull** of a set S as

- the unique minimal convex set containing S,
- the intersection of all convex sets containing S, or
- the set of all convex combinations of points in S.

Any point $\mathbf{x} \in \operatorname{conv} S$, where $S \subseteq \mathbb{R}^n$ can thus be expressed as a convex combination of points in S.

How many do we need? 2 minutes

Caratheodory's Theorem
Let $\mathbf{x} \in \text{conv}\, S$, where $S \subseteq \mathbb{R}^n$. Then $\mathbf{x}$ can be expressed as a convex combination of $n + 1$ or fewer points of S.

Proof:

Polytope

A subset P of $\mathbb{R}^n$ is a **polytope** if it is the convex hull of finitely many points in $\mathbb{R}^n$.

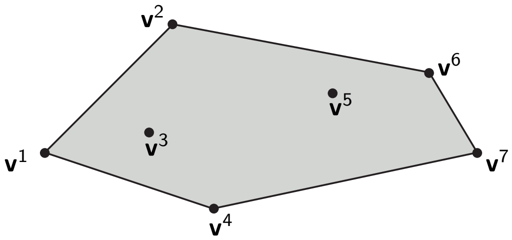


Figure: A polytope generated by seven points

A cube and a tetrahedron are examples of polytopes in $\mathbb{R}^3$.

Extreme points

A point $\mathbf{v}$ of a convex set P is a **extreme point** if whenever

$$\left.\begin{array}{r} \mathbf{v} = \lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2 \\ \mathbf{x}^1, \mathbf{x}^2 \in P \\ \lambda \in (0,1) \end{array}\right\} \implies \mathbf{v} = \mathbf{x}^1 = \mathbf{x}^2.$$

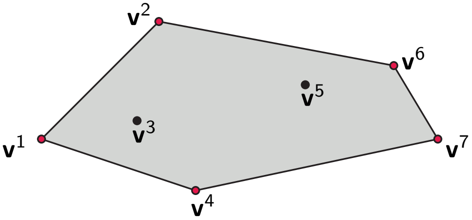


Figure: The red dots are the extreme points

An intuitive theorem then is

Theorem
Let P be the polytope conv V, where $V = \{\mathbf{v}^1, \ldots, \mathbf{v}^k\} \subset \mathbb{R}^n$. Then P is equal to the convex hull of its extreme points.

- We have up untill now described convex sets by a number of points – **interior representation**.
- We now look at convex sets described by linear constraints – **exterior representation**.

A subset P of $\mathbb{R}^n$ is a **polyhedron** if there exists a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ and a vector $\mathbf{b} \in \mathbb{R}^m$ such that

$$P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \le \mathbf{b}\}$$

- $\mathbf{A}\mathbf{x} \le \mathbf{b} \iff \mathbf{a}_i\mathbf{x} \le b_i,\ i = 1, \ldots, m.$ ($\mathbf{a}_i$ row i of $\mathbf{A}$)
- $\mathbf{a}_i\mathbf{x} \le b_i,\ i = 1, \ldots, m$ are half-spaces, so
- P is the intersection of m half-spaces.

Polyhedra, example I

Let $P = \{\mathbf{x} \in \mathbb{R}^2 \mid x_2 \geq 1,\ x_1 + 2x_2 \leq 6,\ 2x_1 - x_2 \geq 2\}$

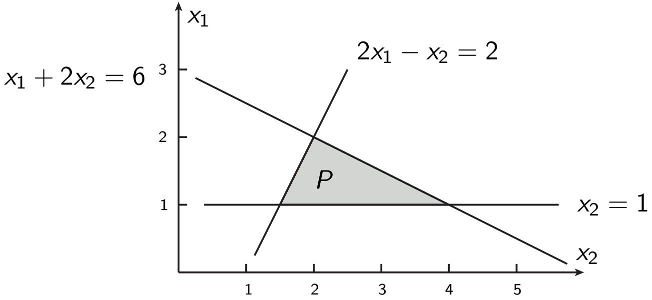


Figure: The polyhedron P, where $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -1 \\ 6 \\ -2 \end{pmatrix}$

Polyhedra, example II (unbounded)

Let $P = \{\mathbf{x} \in \mathbb{R}^2 \mid x_2 \geq 1,\ 2x_1 - x_2 \geq 2\}$

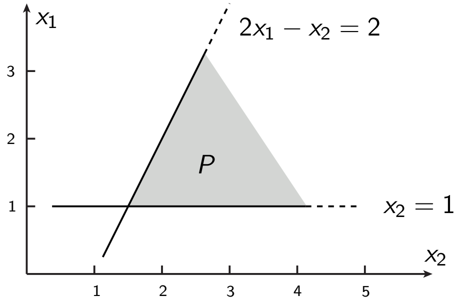


Figure: The polyhedron P, where $\mathbf{A} = \begin{pmatrix} 0 & -1 \\ -2 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$

Polytope vs. Polyhedron

To make a clear distinction between a polytope and a polyhedron.

- Polytope $=$ The convex hull of finitely many points.
- Polyhedron $=$ The intersection of finitely many half-spaces

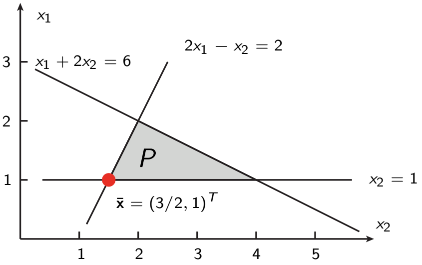
Theorem
A set P is a *polytope* if and only if it is a *bounded polyhedron*.

We can now define the extreme points of a polyhedron.

Let $\bar{\mathbf{x}} \in P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$, where $\mathbf{A} \in \mathbb{R}^{m \times n}$ with rank $\mathbf{A} = n$ and $\mathbf{b} \in \mathbb{R}^m$. Further, let $\bar{\mathbf{A}}\bar{\mathbf{x}} = \bar{\mathbf{b}}$ be the **equality subsystem** of $\mathbf{A}\bar{\mathbf{x}} \leq \mathbf{b}$. Then $\bar{\mathbf{x}}$ is an extreme point of P if and only if rank $\bar{\mathbf{A}} = n$.

To create the *equality subsystem*, strike out all rows i with $\mathbf{a}_i\bar{\mathbf{x}} < b_i$, require equality for the rest.

Algebraic representation, example I



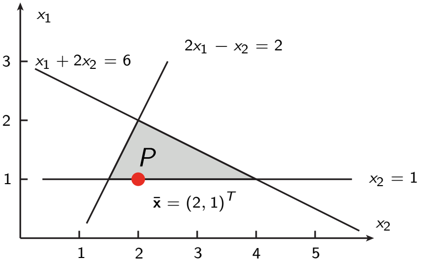
Let $P = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$, where

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & 1 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} -1 \\ 6 \\ -2 \end{pmatrix}$$

Look at $\bar{\mathbf{x}} = (3/2, 1)^T$.

- The second inequality is $1 \cdot (3/2) + 2 \cdot (1) < 6$, so we strike that row.
- The first and second constraints are fulfilled with equality, so
- $\bar{\mathbf{A}} = \begin{pmatrix} 0 & -1 \\ -2 & 1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -1 \\ -2 \end{pmatrix}$
- rank $\bar{\mathbf{A}} = 2$, so $\bar{\mathbf{x}}$ is an extreme point!

Algebraic representation, example II



Let $P = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$, where

$$\mathbf{A} = \begin{pmatrix} 0 & -1 \\ 1 & 2 \\ -2 & 1 \end{pmatrix} \text{ and } \mathbf{b} = \begin{pmatrix} -1 \\ 6 \\ -2 \end{pmatrix}$$

Look at $\bar{\mathbf{x}} = (2, 1)^T$.

- The second inequality is $1 \cdot (2) + 2 \cdot (1) < 6$, so we strike that row.
- The third inequality is $-2 \cdot (2) + 1 \cdot (1) < -2$, so we strike that row.
- $\bar{\mathbf{A}} = \begin{pmatrix} 0 & -1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} -1 \end{pmatrix}$
- rank $\bar{\mathbf{A}} = 1$, so $\bar{\mathbf{x}}$ is **not** an extreme point!

Algebraic representation, example II

Why is this important?

- When considering linear programs (LPs), at least one of the extreme points are an optimal solution (if there exist one)
- How many extreme points can there exist?
  2 minutes
- To find an extreme point, just choose n linearly independent rows ($\Rightarrow$ rank $\bar{\mathbf{A}} = n$) and solve the equality subsystem
- i.e. there can at most be $\binom{m}{n}$
- This is the basic idea behind the **simplex algorithm**

Cones

A subset $C \subseteq \mathbb{R}^n$ is a **cone** if $\lambda \mathbf{x} \in C$ whenever $\mathbf{x} \in C$ and $\lambda > 0$.

Example: The set $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{0}\}$ is a polyhedral cone in $\mathbb{R}^n$.

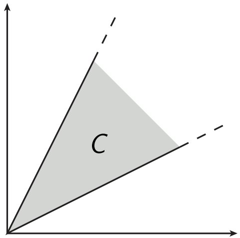


Figure: A convex cone

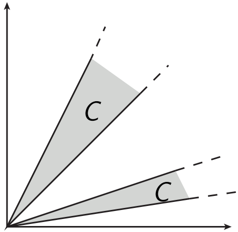


Figure: A non-convex cone

The representation theorem

Theorem

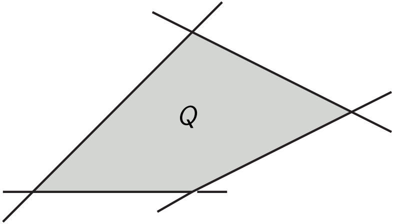
Let the polyhedron $Q = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \leq \mathbf{b}\}$ and let $\{\mathbf{v}^1, \ldots, \mathbf{v}^k\}$ be its extreme points.

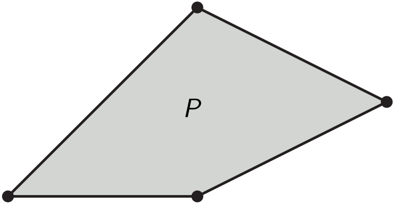
Define $P := \text{conv}\left(\{\mathbf{v}^1, \ldots, \mathbf{v}^k\}\right)$ and $C := \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \leq \mathbf{0}\}$.

Then $Q = P + C = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{u} + \mathbf{v} \text{ for some } \mathbf{u} \in P \text{ and } \mathbf{v} \in C\}$

Meaning that each polyhedron can be written as the sum of a polytope and a polyhedral cone.

The representation theorem, example I

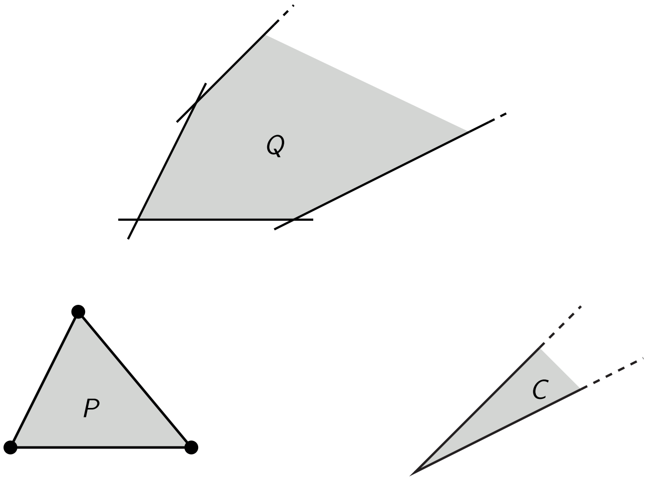




$C = \emptyset$

Bounded polyhedron = polytope

The representation theorem, example II



Unbounded polyhedron = polytope + polyhedral cone

Separation theorem

An intuitive result is:
"If a point $\mathbf{y}$ does not lie in a closed convex set S, then there exist a hyperplane separating the $\mathbf{y}$ from S".

Theorem
Suppose that the set $S \subseteq \mathbb{R}^n$ is closed and convex, and that the point $\mathbf{y}$ does not lie in S. Then there exist a $\alpha \in \mathbb{R}$ and $\boldsymbol{\pi} \neq \mathbf{0}^n$ such that $\boldsymbol{\pi}^T \mathbf{y} > \alpha$ and $\boldsymbol{\pi}^T \mathbf{x} \leq \alpha$ for all $\mathbf{x} \in S$.

Proof: Later in the course.

Separation theorem, example

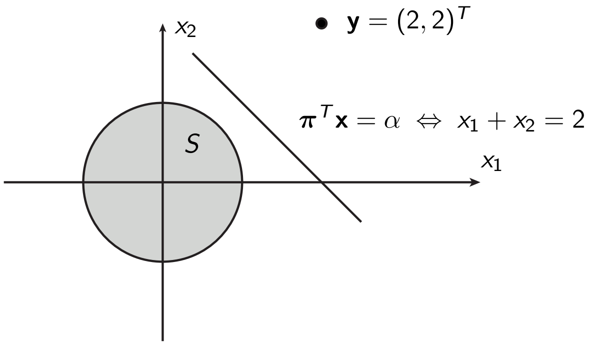


Figure: The hyperplane $\boldsymbol{\pi}^T\mathbf{x} = \alpha$ separates the point $\mathbf{y} = (2,2)^T$ from the unit disc. $\boldsymbol{\pi} = (1,1)^T$ and $\alpha = 2$.

Farkas' Lemma

A very important result is the following

Theorem

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. Then exactly one of the systems

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b}, \\ \mathbf{x} &\geq \mathbf{0}, \end{aligned} \tag{1}$$

and

$$\begin{aligned} \mathbf{A}^T \boldsymbol{\pi} &\leq \mathbf{0}, \\ \mathbf{b}^T \boldsymbol{\pi} &> 0, \end{aligned} \tag{2}$$

has a feasible solution, and the other is inconsistent.

Farkas' Lemma

What Farkas' Lemma says is the following:

- Either the vector **b** lies in the cone spanned by the columns of **A**, i.e.

$$\mathbf{b} = \mathbf{A}\mathbf{x}, \quad \mathbf{x} \geq \mathbf{0},$$

- or the vector **b** does not lie in the cone, meaning that there exist a hyperplane $\boldsymbol{\pi}$ separating the vector from the cone, i.e.

$$\mathbf{A}^T\boldsymbol{\pi} \leq \mathbf{0}, \quad \mathbf{b}^T\boldsymbol{\pi} > 0.$$

Proof of Farkas' Lemma:

Definition

Suppose $S \subseteq \mathbb{R}^n$ is **convex**. A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex on S if

$$\left.\begin{array}{c} \mathbf{x}^1, \mathbf{x}^2 \in S \\ \lambda \in (0,1) \end{array}\right\} \implies f\left(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2\right) \leq \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2)$$

- A function is **strictly convex** on S if $<$ holds in place of $\leq$ for all $\mathbf{x}^1 \neq \mathbf{x}^2$.
- A function f is **concave** if $-f$ is convex.

$$f\left(\lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2\right) \leq \lambda f(\mathbf{x}^1) + (1-\lambda) f(\mathbf{x}^2)$$

The linear interpolation between two points on the function never is lower than the function.

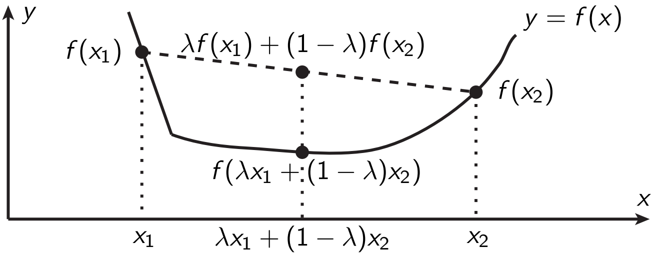


Figure: A convex function

Example, II

Two important examples:

- $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + d$, where $\mathbf{c} \in \mathbb{R}^n$, $d \in \mathbb{R}$ is both convex and concave.
- $f(\mathbf{x}) = ||\mathbf{x}||$ is convex, $f(\mathbf{x}) = ||\mathbf{x}||^2$ is strictly convex.

Proposition (sum of convex functions)
The non-negative linear combination of convex functions is convex.

Proposition (composite functions)
Suppose $S \subseteq \mathbb{R}^n$ and $P \subseteq \mathbb{R}$. Let $g : S \to \mathbb{R}$ be convex on S and $f : P \to \mathbb{R}$ be convex and non-decreasing on P.
Then the composite function $f(g)$ is convex on $\{\mathbf{x} \in S \mid g(\mathbf{x}) \in P\}$.

The **epigraph** of a function $f : \mathbb{R}^n \to \mathbb{R}^n \cup \{+\infty\}$ is defined as

$$\operatorname{epi} f := \{(\mathbf{x}, \alpha) \in \mathbb{R}^n \times \mathbb{R} \mid f(\mathbf{x}) \leq \alpha\}.$$

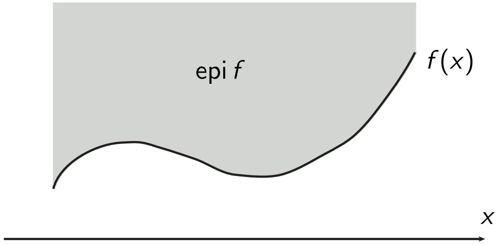


Figure: The epigraph of a non-convex function. Note that the boundary of epi f is the graph of f.

Epigraph, II

The **epigraph** of a function $f : \mathbb{R}^n \to \mathbb{R}^n \cup \{+\infty\}$ restricted to the set $S \subseteq \mathbb{R}^n$ is defined as

$$\operatorname{epi}_S f := \{(\mathbf{x}, \alpha) \in S \times \mathbb{R} \mid f(\mathbf{x}) \leq \alpha\}.$$

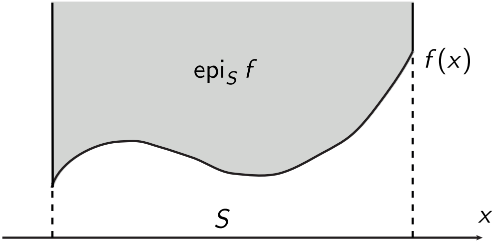


Figure: The epigraph of a non-convex function restricted to S

Epigraph, III

We now make a connection between convex functions and convex sets.

> Suppose $S \subseteq \mathbb{R}^n$ is a convex set. Then, the function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ is convex on S if and only if its epigraph restricted to S is a convex set in $\mathbb{R}^{n+1}$.

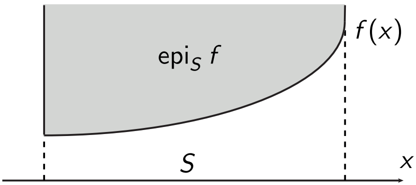
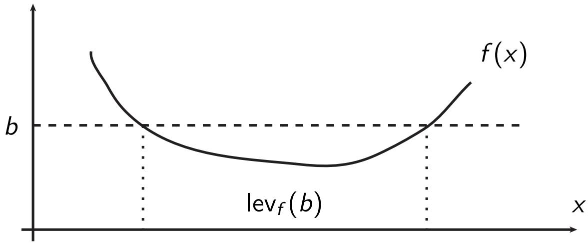


Figure: The epigraph of a convex function restricted to S

Level set

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. The **level set** of f with respect to the b is the set

$$\text{lev}_f(b) : \{\mathbf{x} \in \mathbb{R}^n \mid f(\mathbf{x}) \leq b\}$$



If f is convex, then for every value of $b \in \mathbb{R}$, $\text{lev}_f(b)$ is *closed* and *convex*.

Convex sets, I

Another important connection between convex functions and convex sets.

Proposition

Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is a convex function. Then the set $\{\mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \leq 0\}$ is a convex set.

Proof

Convex sets, II

Another important connection between convex functions and convex sets.

Proposition
Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is an affine function. Then the set $\{\mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) = 0\}$ is a convex set.

Proof

Case: $f \in C^1$

We now consider the special case when $f \in C^1$, meaning that f is differentiable and has a continuous gradient.

Theorem

$f \in C^1$ is convex on the open, convex set S if and only if

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^T (\mathbf{x} - \bar{\mathbf{x}}), \text{ for all } \mathbf{x}, \bar{\mathbf{x}} \in S.$$

- Meaning that the every tangent plane to the graph of f lies on, or below, the epigraph of f, or
- that each first-order approximation of f lies below f.

Case: $f \in C^1$, Example

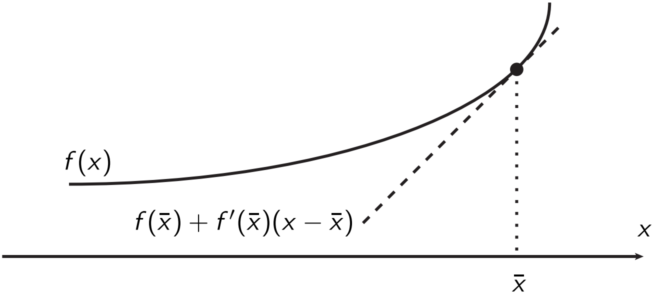


Figure: A convex function. The first-order approximation lies below the function.

Theorem
$f \in C^1$ is convex on the open, convex set S if and only if

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^T(\mathbf{x} - \bar{\mathbf{x}}), \text{ for all } \mathbf{x}, \bar{\mathbf{x}} \in S.$$

Proof:

Case: $f \in C^1$

Another equivalent way of writing this is the following.

Theorem

$f \in C^1$ is convex on the open, convex set S if and only if

$$[\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})]^T (\mathbf{x} - \mathbf{y}) \geq 0$$

- Meaning that the gradient of f is monotone on S, or
- that the angle between $\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})$ and $\mathbf{x} - \mathbf{y}$ should be between $-\pi/2$ and $\pi/2$.

Case: $f \in C^2$

Now the case when $f \in C^2$, meaning that it is twice differentiable with continuous Hessian. Then the following hold

Theorem

Let $f \in C^2$ on an open, convex set $S \subseteq \mathbb{R}^n$. Then

a) f is convex on $S \iff \nabla^2 f(\mathbf{x}) \succeq 0$ for all $\mathbf{x} \in S$

b) $\nabla^2 f(\mathbf{x}) \succ 0$ for all $\mathbf{x} \in S \implies f$ is strictly convex on S.

Proof sketch: Use Taylor expansion and mean-value theorem.

Note that in b), "$\Longleftarrow$" does not hold. Take for example $f(x) = x^4$.

Case: $f \in C^2$

An important example of a function in C^2 is the quadratic function

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} - \mathbf{q}^T\mathbf{x},$$

which is convex on $\mathbb{R}^n$ if and only if $\mathbf{Q} \succeq 0$. This because $\nabla^2 f(\mathbf{x}) = \mathbf{Q}$ independent of $\mathbf{x}$.

Convex optimization problem

We can now say that the optimization problem

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}), && (3a) \\ \text{subject to} \quad & g_i(\mathbf{x}) \leq 0, \quad i \in \mathcal{I}, && (3b) \\ & g_i(\mathbf{x}) = 0, \quad i \in \mathcal{E}, && (3c) \\ & x \in X, && (3d) \end{aligned}$$

is convex if

- f is a convex function,
- g_i, $i \in \mathcal{I}$ are convex functions,
- g_i, $i \in \mathcal{E}$ are affine functions, and
- X is a convex set.