Lecture 2: Convexity

Michael Patriksson

2010-10-27

Convexity of sets

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

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

A set S is convex if, from anywhere in S, all other points are "visible." (See below figure)

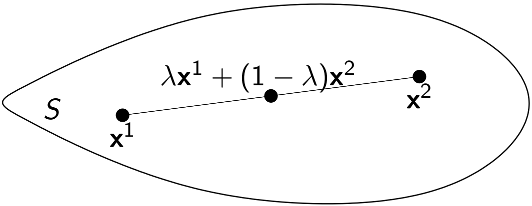


Figure: A convex set. (For the intermediate vector shown, the value of λ is $\approx 1/2$)

Examples

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

Two non-convex sets are shown in the below figure

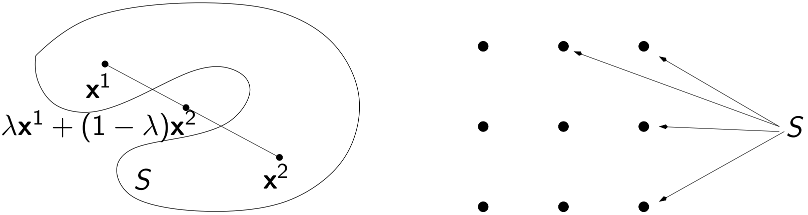


Figure: Two non-convex sets

Intersections of convex sets

Suppose that S_k, $k \in \mathcal{K}$, *is any collection of convex sets. Then, the intersection* $\cap_{k \in \mathcal{K}} S_k$ *is a convex set*

Proof.

Convex and affine hulls

The *affine hull* of a finite set $V = \{\mathbf{v}^1, \dots, \mathbf{v}^k\} \subset \mathbb{R}^n$ is the set

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

The *convex hull* of a finite set $V = \{\mathbf{v}^1, \dots, \mathbf{v}^k\} \subset \mathbb{R}^n$ is the set

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

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

Examples

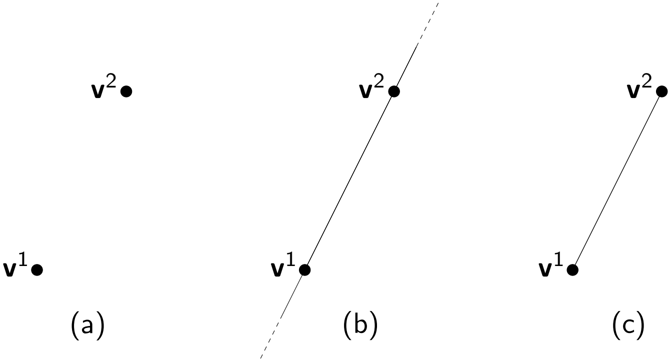


Figure: (a) The set V (b) The set $\mathrm{aff}\ V$ (c) The set $\mathrm{conv}\ V$

Carathéodory's Theorem

- The convex hull of $V \subset \mathbb{R}^n$ is the smallest convex set containing V
- *Let $V \subseteq \mathbb{R}^n$. Then,* $\operatorname{conv} V$ *is the set of all convex combinations of points of V*
- Every point of the convex hull of a set can be written as a convex combination of points from the set. How many do we need?
- [Carathéodory] *Let $\mathbf{x} \in \operatorname{conv} V$, where $V \subseteq \mathbb{R}^n$. Then $\mathbf{x}$ can be expressed as a convex combination of $n+1$ or fewer points of V*
- Proof by contradiction: if more than $n+1$ points are needed then these points must be affinely dependent $\Longrightarrow$ can remove at least one such point. Etcetera

Polytope

- A subset P of $\mathbb{R}^n$ is a *polytope* if it is the convex hull of finitely many points in $\mathbb{R}^n$
- The set shown in the below figure is a polytope

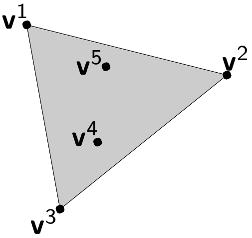


Figure: The convex hull of five points in $\mathbb{R}^2$

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

Extreme points

- A point $\mathbf{v}$ of a convex set P is called an *extreme point* if whenever $\mathbf{v} = \lambda \mathbf{x}^1 + (1-\lambda)\mathbf{x}^2$, where $\mathbf{x}^1, \mathbf{x}^2 \in P$ and $\lambda \in (0,1)$, then $\mathbf{v} = \mathbf{x}^1 = \mathbf{x}^2$

- Examples: The set shown in Figure 3(c) has the extreme points $\mathbf{v}^1$ and $\mathbf{v}^2$. The set shown in Figure 4 has the extreme points $\mathbf{v}^1$, $\mathbf{v}^2$, and $\mathbf{v}^3$. The set shown in Figure 3(b) does not have any extreme points

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

Polyhedra, I

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

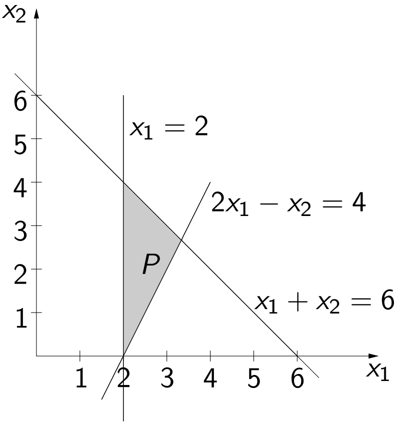
$$P = \{\, \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b} \,\}$$

- $\mathbf{A}\mathbf{x} \leq \mathbf{b} \iff \mathbf{a}_i\mathbf{x} \leq b_i$ for all i ($\mathbf{a}_i$ is row i of $\mathbf{A}$)

- Intersection of half-spaces. [*Hyperplane:* $\{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{a}_i\mathbf{x} = b_i \,\}$]

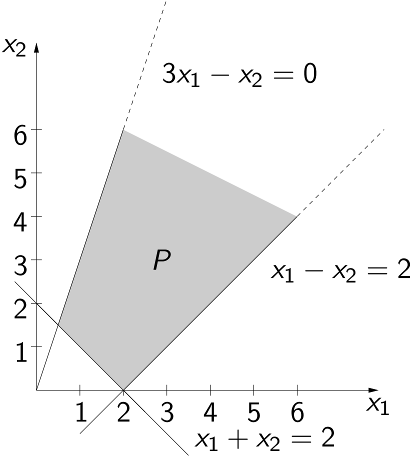
Polyhedra, II

- Examples: (a) The bounded polyhedron
  $P = \{ \mathbf{x} \in \mathbb{R}^2 \mid x_1 \geq 2;\ x_1 + x_2 \leq 6;\ 2x_1 - x_2 \leq 4 \}$



Polyhedra, III

- (b) The unbounded polyhedron
$P = \{ \mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 \geq 2;\; x_1 - x_2 \leq 2;\; 3x_1 - x_2 \geq 0 \}$



Algebraic characterizations of extreme points

- *Let* $\tilde{\mathbf{x}} \in P = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b} \}$*, where* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *with* $\operatorname{rank} \mathbf{A} = n$ *and* $\mathbf{b} \in \mathbb{R}^m$*. Further, let* $\tilde{\mathbf{A}}\tilde{\mathbf{x}} = \tilde{\mathbf{b}}$ *be the equality subsystem*[1] *of* $\mathbf{A}\tilde{\mathbf{x}} \leq \mathbf{b}$*. Then* $\tilde{\mathbf{x}}$ *is an extreme point of* P *if and only if* $\operatorname{rank} \tilde{\mathbf{A}} = n$
- Of great importance in Linear Programming (LP): for LP problems the matrix $\mathbf{A}$ always has full rank! Hence, can solve special subsystem of linear equalities to obtain an extreme point!
- Corollary: *The number of extreme points of* P *is finite*
- Since the number of extreme points is finite, the convex hull of the extreme points of a polyhedron is a polytope
- Consequence: Algorithm for linear programming!

[1] Strike out all rows i with $\mathbf{a}_i\tilde{\mathbf{x}} < b_i$; require equality for the rest

Cones

- A subset C of $\mathbb{R}^n$ is a *cone* if $\lambda \mathbf{x} \in C$ whenever $\mathbf{x} \in C$ and $\lambda > 0$
- Example: Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The set $\{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{0}^m \}$ is a cone
- The two figures below illustrate (a) a convex cone and (b) a non-convex cone in $\mathbb{R}^2$

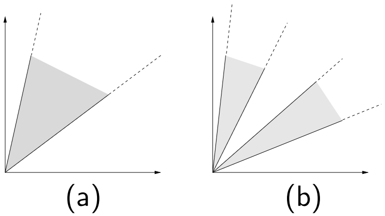


Figure: (a) A convex cone in $\mathbb{R}^2$ (b) A non-convex cone in $\mathbb{R}^2$

Representation Theorem, I

- *Let* $Q = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{b} \}$*,* P *be the convex hull of the extreme points of* Q*, and* $C := \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \leq \mathbf{0}^m \}$*.*
  If $\operatorname{rank} \mathbf{A} = n$ *then*
  $Q = P + C = \{ \mathbf{x} \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{u} + \mathbf{v}$ *for some* $\mathbf{u} \in P$ *and* $\mathbf{v} \in C \}$
  In other words, every polyhedron (that has at least one extreme point) is the sum of a polytope and a polyhedral cone

- Proof by induction on the rank of the subsystem matrix $\tilde{\mathbf{A}}$

Representation Theorem, II

- Central in Linear Programming. Can be used to establish: *Optimal solutions to LP problems are found at extreme points*!

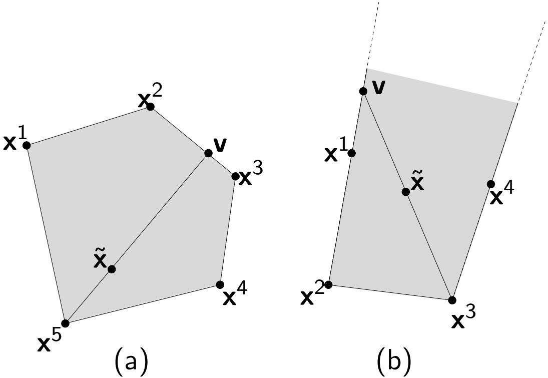


Figure: Illustration of the Representation Theorem (a) in the bounded case, and (b) in the unbounded case

Separation Theorem, I

- *"If a point* $\mathbf{y}$ *does not lie in a closed and convex set* C*, then there exists a hyperplane that separates* $\mathbf{y}$ *from* C*"*

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

- Proof later—requires existence and optimality conditions

- Consequence: *A set* P *is a polytope if and only if it is a bounded polyhedron.* [$\Longleftarrow$ trivial; $\Longrightarrow$ constructive]

Separation Theorem, II

- A *finitely generated* cone has the form

$$\operatorname{cone}\{\mathbf{v}^1,\dots,\mathbf{v}^m\} := \{\,\lambda_1\mathbf{v}^1+\cdots+\lambda_m\mathbf{v}^m \mid \lambda_1,\dots,\lambda_m \geq 0\,\}$$

- *A convex cone is finitely generated iff it is polyhedral*

Separation Theorem, III

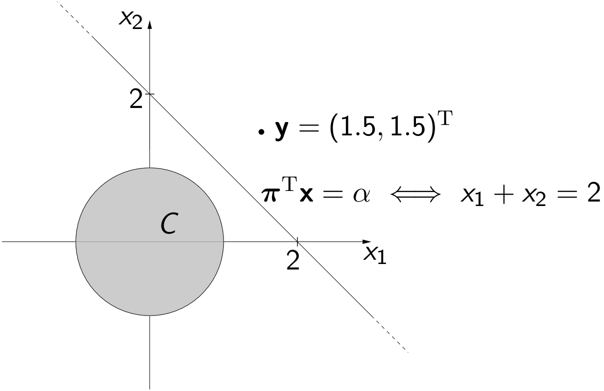


Figure: Illustration of the Separation Theorem: the unit disk is separated from $\mathbf{y}$ by the line $\{\, \mathbf{x} \in \mathbb{R}^2 \mid x_1 + x_2 = 2 \,\}$

Farkas' Lemma

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

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \mathbf{b}, \\ \mathbf{x} &\geq \mathbf{0}^n, \end{aligned} \qquad \text{(I)}$$

  and

$$\begin{aligned} \mathbf{A}^{\mathrm{T}}\boldsymbol{\pi} &\leq \mathbf{0}^n, \\ \mathbf{b}^{\mathrm{T}}\boldsymbol{\pi} &> 0, \end{aligned} \qquad \text{(II)}$$

  has a feasible solution, and the other system is inconsistent
- Farkas' Lemma has many forms. "Theorems of the alternative"
- Crucial for LP theory and optimality conditions; used these days also to analyze and correct computer code!

*Relations between theorems

- Alg. Repr. of Extreme Pnts. (3.17) $\Longrightarrow$ Repr. Thm. (3.22)
- Repr. Thm. (3.22) + Sep. Thm. (3.24) $\Longrightarrow$ "P polytope $\Leftrightarrow$ P bounded polyhedron" (3.26)
- "Convex cone C finitely generated $\Leftrightarrow$ convex cone C is a polyhedron" (3.28): $\Longrightarrow$ from (3.26), $\Longleftarrow$ from (3.22)
- Sep. Thm. (3.24) $\Longrightarrow$ Farkas Lemma (3.30)
- Farkas Lemma (3.30) will later on be established much more simply by utilizing linear programming duality theory

Convexity of functions, I

- Suppose that $S \subseteq \mathbb{R}^n$ is convex. A function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is *convex at* $\bar{\mathbf{x}} \in S$ if

$$\left.\begin{array}{l} \mathbf{x} \in S \\ \lambda \in (0,1) \end{array}\right\} \Longrightarrow f(\lambda\bar{\mathbf{x}} + (1-\lambda)\mathbf{x}) \leq \lambda f(\bar{\mathbf{x}}) + (1-\lambda) f(\mathbf{x})$$

- The function f is *convex on* S if it is convex at every $\bar{\mathbf{x}} \in S$

- The function f is *strictly convex on* S if $<$ holds in place of $\leq$ above for every $\mathbf{x} \neq \bar{\mathbf{x}}$

Convexity of functions, II

- A convex function is such that a linear interpolation never is lower than the function itself. For a strictly convex function the linear interpolation lies above the function
- (Strict) concavity of $f \iff$ (strict) convexity of $-f$
- The below figure illustrates a convex function

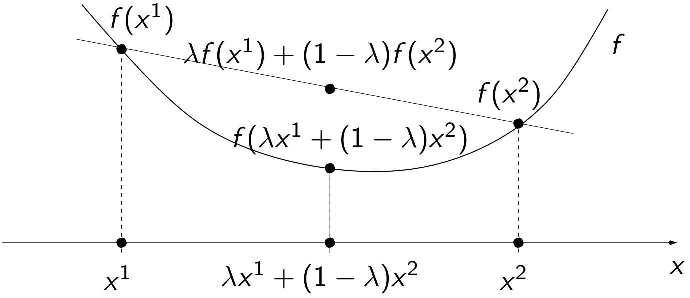


Figure: A convex function

Convexity of functions, III

- The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\mathbf{x}) := \|\mathbf{x}\|$ is convex on $\mathbb{R}^n$; $f(\mathbf{x}) := \|\mathbf{x}\|^2$ is strictly convex in $\mathbb{R}^n$
- Let $\mathbf{c} \in \mathbb{R}^n$. The linear function $\mathbf{x} \mapsto f(\mathbf{x}) := \mathbf{c}^{\mathrm{T}}\mathbf{x} = \sum_{j=1}^n c_j x_j$ is both convex and concave on $\mathbb{R}^n$
- The below figure illustrates a non-convex function

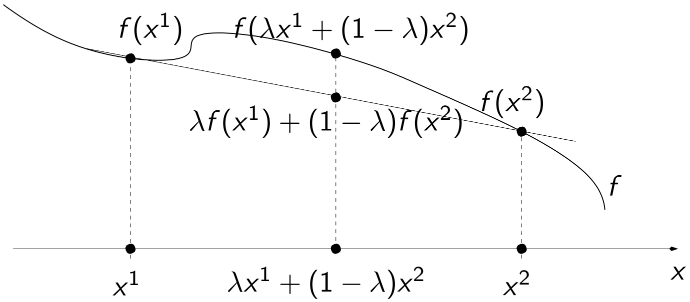


Figure: A non-convex function

Convexity of functions, IV

- Sums of convex functions are convex
- Composite function: $\mathbf{x} \mapsto f(g(\mathbf{x}))$
- *Suppose that $S \subseteq \mathbb{R}^n$ and $P \subseteq \mathbb{R}$. Let further $g : S \to \mathbb{R}$ be a function which is convex on S, and $f : P \to \mathbb{R}$ be convex and non-decreasing ($y \geq x \implies f(y) \geq f(x)$) on P. Then, the composite function $f(g)$ is convex on the set $\{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \in P \,\}$*
- The function $\mathbf{x} \mapsto -\log(-g(\mathbf{x}))$ is convex on the set $\{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) < 0 \,\}$

Epigraphs, I

- Characterize convexity of a function on $\mathbb{R}^n$ by the convexity of its *epigraph* in $\mathbb{R}^{n+1}$. [*Note:* the *graph* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the boundary of $\operatorname{epi} f$]

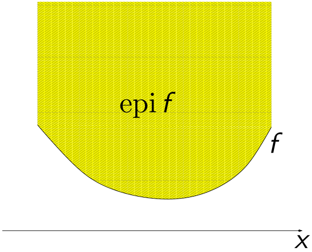


Figure: A convex function and its epigraph

- The epigraph of a function $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is the set

$$\operatorname{epi} f := \{ (\mathbf{x}, \alpha) \in \mathbb{R}^{n+1} \mid f(\mathbf{x}) \leq \alpha \}$$

Epigraphs, II

- Connection between convex sets and functions; in fact the definition of a convex function stems from that of a convex set:
- *The function* $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ *is convex on* $\mathbb{R}^n$ *if, and only if, its epigraph is a convex set in* $\mathbb{R}^{n+1}$

Convexity characterizations in C^1, I

- C^1: Differentiable once, gradient continuous
- *Let $f \in C^1$ on an open convex set S*
  (a) *f is convex on S* $\iff f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^{\mathrm{T}}(\mathbf{y}-\mathbf{x})$, $\mathbf{x}, \mathbf{y} \in S$
  (b) *f is convex on* $S \iff [\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})]^{\mathrm{T}}(\mathbf{x}-\mathbf{y}) \geq 0, \mathbf{x}, \mathbf{y} \in S$
- (a): "Every tangent plane to the function surface lies on, or below, the epigraph of f", or, that "a first-order approximation is below f"
- (b) ∇f is "monotone on S." [*Note:* when $n = 1$, the result states that f is convex if and only if its derivative f' is non-decreasing, that is, that it is monotonically increasing]
- Proofs use Taylor expansion, convexity and the Mean-value Theorem

Convexity characterizations in C^1, II

- The below figure illustrates part (a)

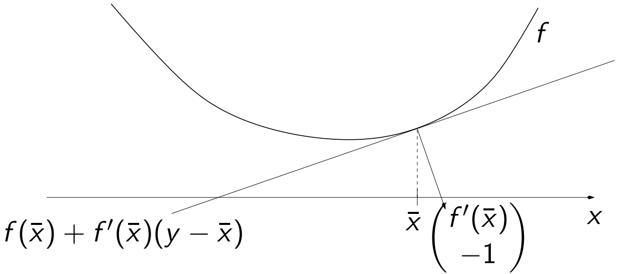


Figure: A tangent plane to the graph of a convex function

Convexity characterizations in C^2

- *Let f be in C^2 on an open, convex set $S \subseteq \mathbb{R}^n$*
  (a) *f is convex on $S \Longleftrightarrow \nabla^2 f(\mathbf{x})$ is positive semidefinite for all $\mathbf{x} \in S$*
  (b) *$\nabla^2 f(\mathbf{x})$ is positive definite for all $x \in S \Longrightarrow f$ is strictly convex on S*
- *Note:* $n = 1$, S is an open interval: (a) f is convex on S if and only if $f''(x) \geq 0$ for every $x \in S$; (b) f is strictly convex on S if $f''(x) > 0$ for every $x \in S$
- Proofs use Taylor expansion, convexity and the Mean-value Theorem
- Not the direction $\Longleftarrow$ in (b)! [$f(x) = x^4$ at $x = 0$]
- Difficult to check convexity; matrix condition for every $\mathbf{x}$
- Quadratic function: $f(\mathbf{x}) = (1/2)\mathbf{x}^{\mathrm{T}}\mathbf{Q}\mathbf{x} - \mathbf{q}^{\mathrm{T}}\mathbf{x}$ convex on $\mathbb{R}^n$ iff $\mathbf{Q}$ is psd ($\mathbf{Q}$ is the Hessian of f, and is independent of $\mathbf{x}$)

Convexity of feasible sets

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

$$\operatorname{lev}_g(b) := \{\, \mathbf{x} \in \mathbb{R}^n \mid g(\mathbf{x}) \leq b \,\}$$

- The below figure illustrates a level set of a convex function

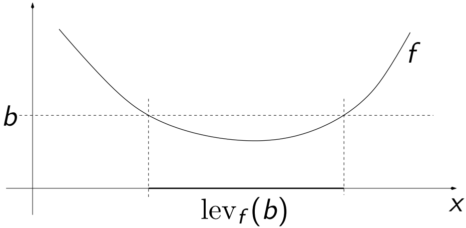


Figure: A level set of a convex function

- *Suppose that the function $g : \mathbb{R}^n \to \mathbb{R}$ is convex. Then, for every value of $b \in \mathbb{R}$, the level set $\operatorname{lev}_g(b)$ is a convex set. It is moreover closed*

*Euclidean projection, I

- The Euclidean projection of $\mathbf{w} \in \mathbb{R}^n$ is the nearest (in Euclidean norm) vector in S to $\mathbf{w}$. The vector $\mathbf{w} - \mathrm{Proj}_S(\mathbf{w})$ is *normal* to S

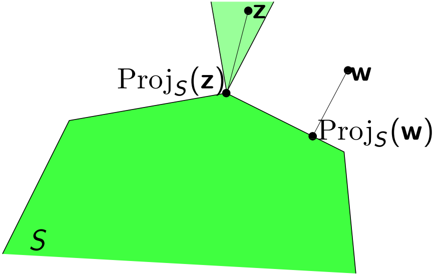


Figure: The projection of two vectors onto a convex set

*Euclidean projection, II

The distance function below is convex:

$$\mathrm{dist}_S(\mathbf{x}) := \|\mathbf{x} - \mathrm{Proj}_S(\mathbf{x})\|, \qquad \mathbf{x} \in \mathbb{R}^n$$

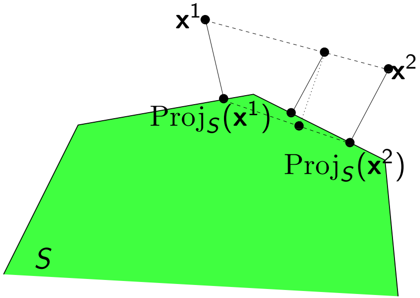


Figure: From the intermediate vector $\lambda\mathbf{x}^1 + (1-\lambda)\mathbf{x}^2$ shown the distance to the vector $\lambda\mathrm{Proj}_S(\mathbf{x}^1) + (1-\lambda)\mathrm{Proj}_S(\mathbf{x}^2))$ [dotted line segment] clearly is longer than to its projection on S [solid line]