Lecture 2: Convexity

Convexity of sets

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

$$\left. \begin{array}{l} \boldsymbol{x}^1, \boldsymbol{x}^2 \in S \\[2mm] \lambda \in (0, 1) \end{array} \right\} \implies \lambda \boldsymbol{x}^1 + (1 - \lambda) \boldsymbol{x}^2 \in S$$

A set S is convex if, from anywhere in S, all other points are "visible." (See Figure 1)
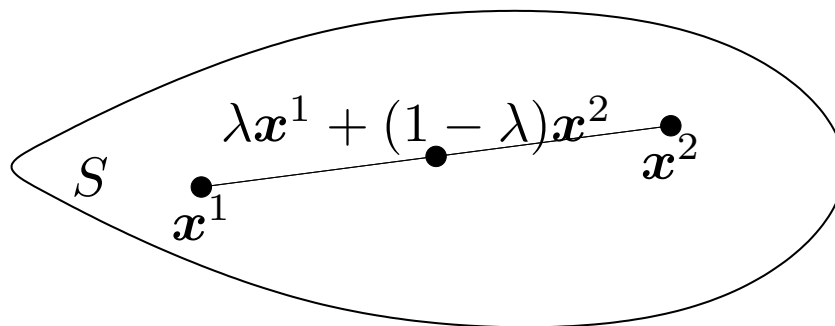
PSfrag replacements

Figure 1: 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 $\{\, \boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| \leq a \,\}$ is convex for every value of $a \in \mathbb{R}$

- The set $\{\, \boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x}\| = a \,\}$ is non-convex for every $a > 0$

- The set $\{0, 1, 2\}$ is non-convex

Two non-convex sets are shown in Figure 2

replacements

PSfrag replacements

$\boldsymbol{x}^1$

$\lambda\boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2$

S

$\boldsymbol{x}^2$

$\boldsymbol{x}^1$

$\boldsymbol{x}^2$

$\lambda\boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2$
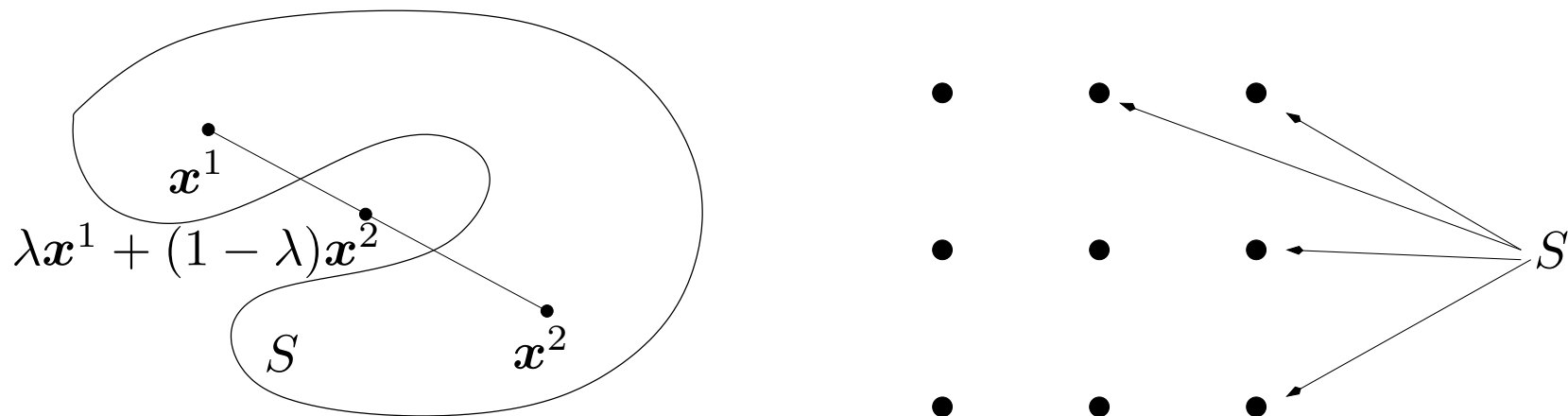
S

Figure 2: 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 $V = \{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k\} \subset \mathbb{R}^n$ is the set

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

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

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

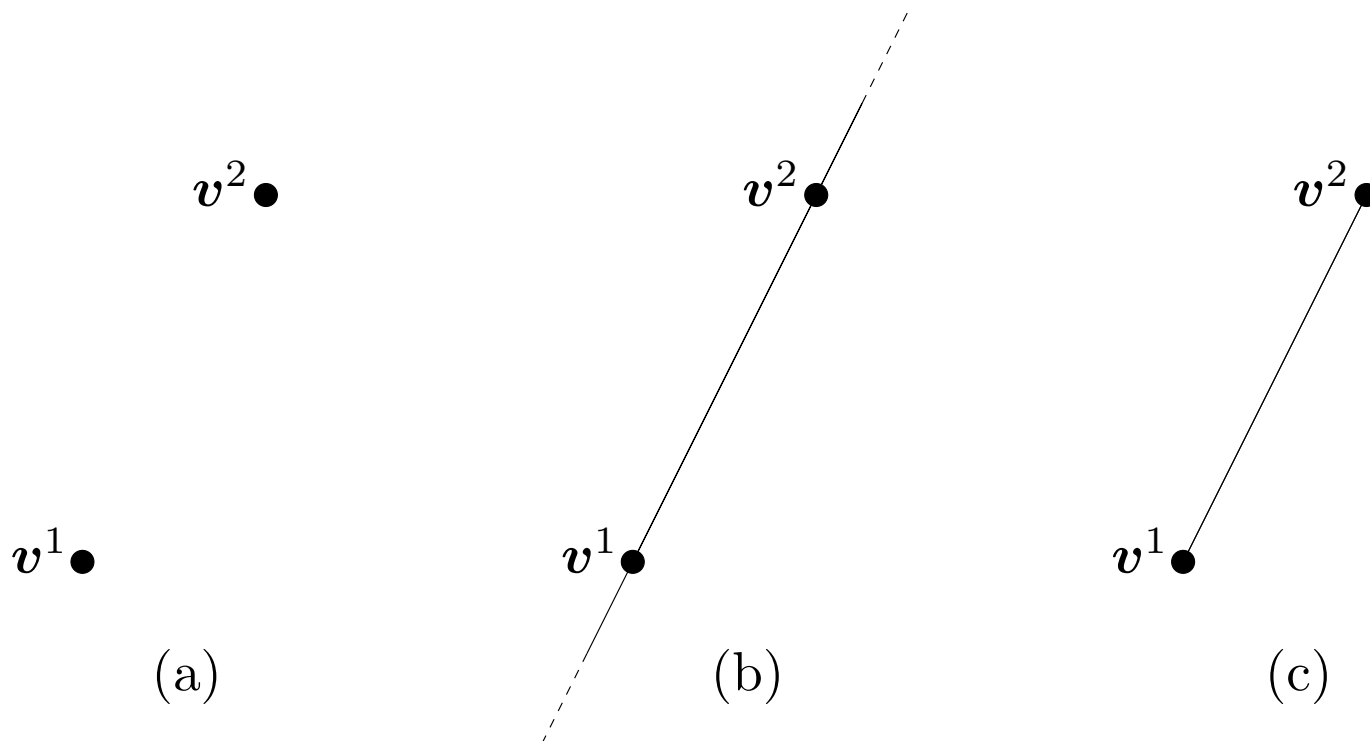The sets are defined by all possible *affine (convex) combinations* of the k points

Examples

Figure 3: (a) The set V (b) The set aff V (c) The set 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?

- [Car.:] *Let $\boldsymbol{x} \in \operatorname{conv} V$, where $V \subseteq \mathbb{R}^n$. Then $\boldsymbol{x}$ 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 $\implies$ 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 $\mathbb{R}^n$
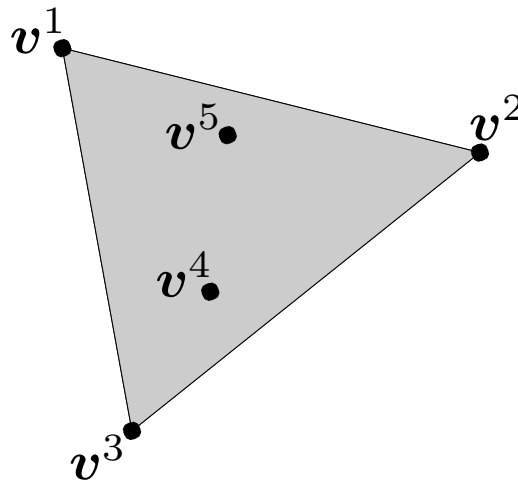
- The set shown in Figure 4 is a polytope

PSfrag replacements

Figure 4: 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 $\boldsymbol{v}$ of a convex set P is called an *extreme point* if whenever $\boldsymbol{v} = \lambda \boldsymbol{x}^1 + (1-\lambda)\boldsymbol{x}^2$, where $\boldsymbol{x}^1, \boldsymbol{x}^2 \in P$ and $\lambda \in (0,1)$, then $\boldsymbol{v} = \boldsymbol{x}^1 = \boldsymbol{x}^2$

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

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

Polyhedra

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

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

- $\boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \Longleftrightarrow \boldsymbol{a}_i\boldsymbol{x} \le b_i$ for all i ($\boldsymbol{a}_i$ is row i of $\boldsymbol{A}$)

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

- Examples: (a) Figure 5 shows the bounded polyhedron $P = \{\, \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 \ge 2;\ x_1 + x_2 \le 6;\ 2x_1 - x_2 \le 4 \,\}$

- (b) The unbounded polyhedron $P = \{\, \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 + x_2 \ge 2;\ x_1 - x_2 \le 2;\ 3x_1 - x_2 \ge 0 \,\}$ is shown in Figure 6
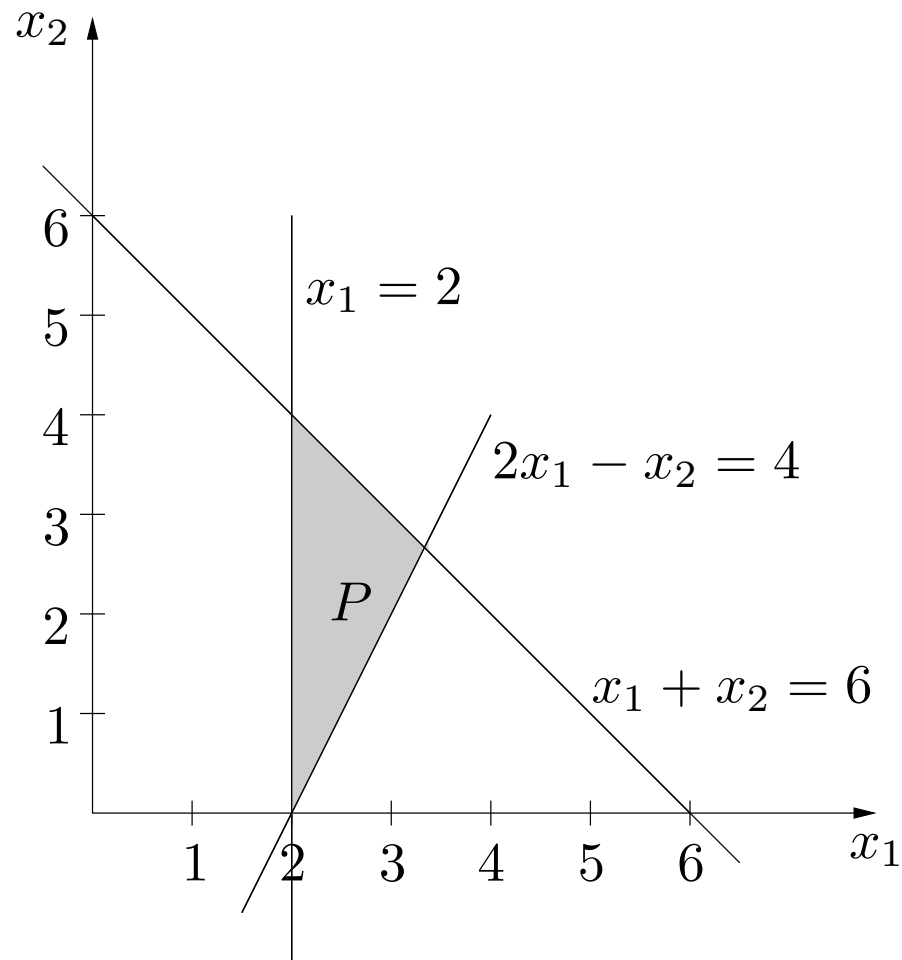
PSfrag replacements

Figure 5: Illustration of the bounded polyhedron $P = \{\, \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 \geq 2;\ x_1 + x_2 \leq 6;\ 2x_1 - x_2 \leq 4\,\}$
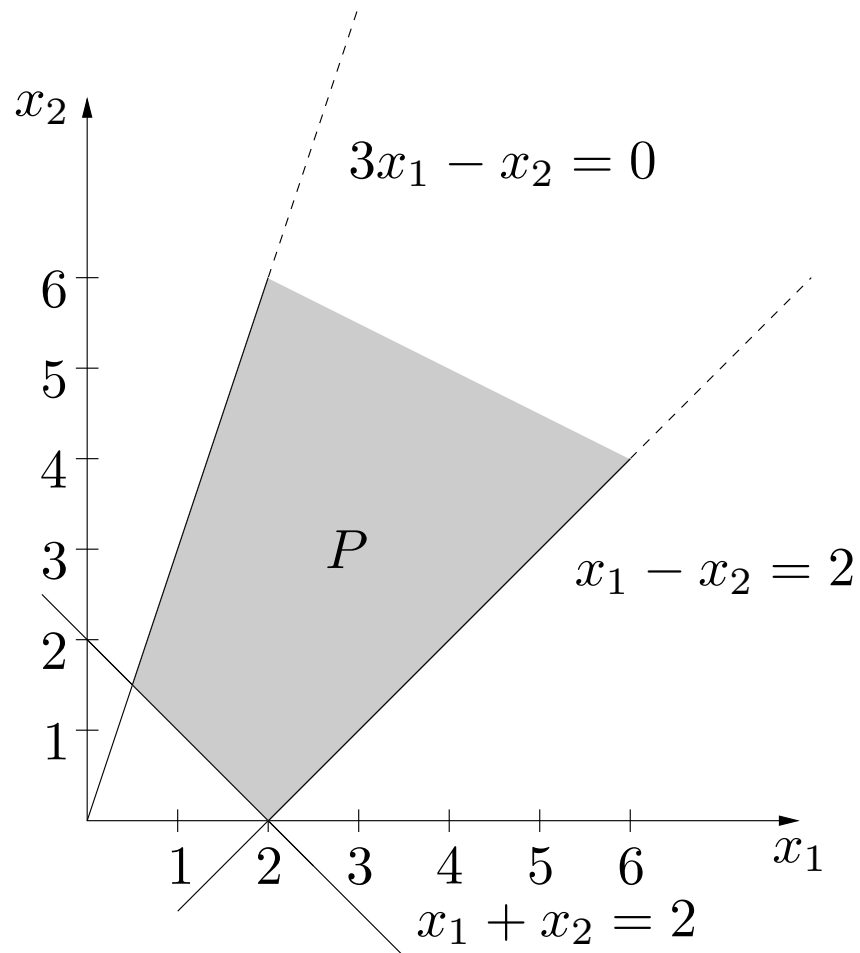
Figure 6: Illustration of the unbounded polyhedron $P = \{ \, \boldsymbol{x} \in \mathbb{R}^2 \mid x_1 + x_2 \geq 2; \; x_1 - x_2 \leq 2; \; 3x_1 - x_2 \geq 0 \, \}$

Algebraic characterizations of extreme points

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

- Of great importance in Linear Programming: A then 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!

Cones

- A subset C of $\mathbb{R}^n$ is a *cone* if $\lambda \boldsymbol{x} \in C$ whenever $\boldsymbol{x} \in C$ and $\lambda > 0$

- Example: Let $\boldsymbol{A} \in \mathbb{R}^{m \times n}$. The set $\{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{0}^m \,\}$ is a cone

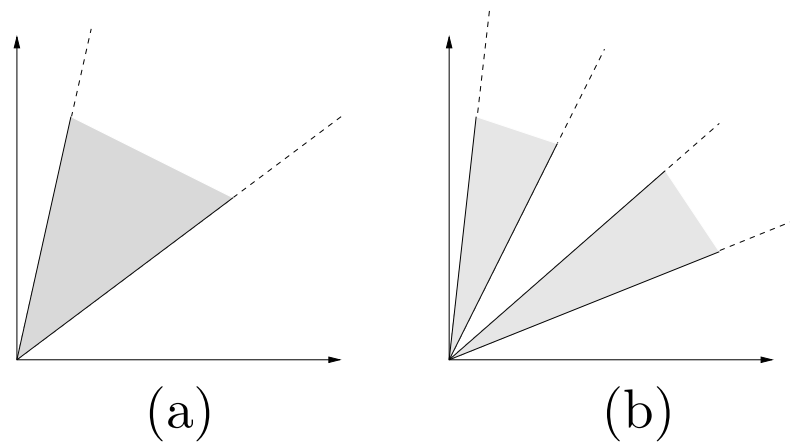- Figure 7(a) illustrates a convex cone and Figure 7(b) illustrates a non-convex cone in $\mathbb{R}^2$

PSfrag replacements

(a)   (b)

Figure 7: (a) A convex cone in $\mathbb{R}^2$ (b) A non-convex cone in $\mathbb{R}^2$

Representation Theorem

- Let $Q = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \,\}$, P be the convex hull of the extreme points of Q, and $C := \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{0}^m \,\}$. If rank $\boldsymbol{A} = n$ then
  $Q = P + C = \{\, \boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v} \text{ for some } \boldsymbol{u} \in P \text{ and } \boldsymbol{v} \in C \,\}$
  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 $\tilde{\boldsymbol{A}}$

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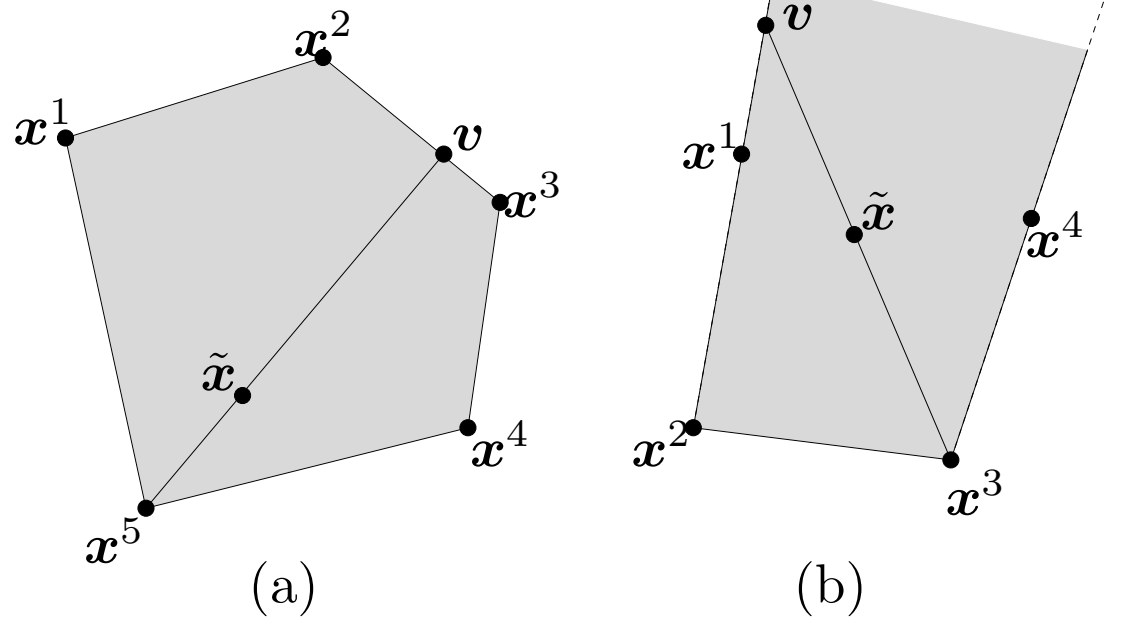 Optimal solutions to LP problems are found at extreme points!

Figure 8: Illustration of the Representation Theorem (a) in the bounded case, and (b) in the unbounded case

Separation Theorem

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

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

- 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$]$

- A *finitely generated* cone has the form

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

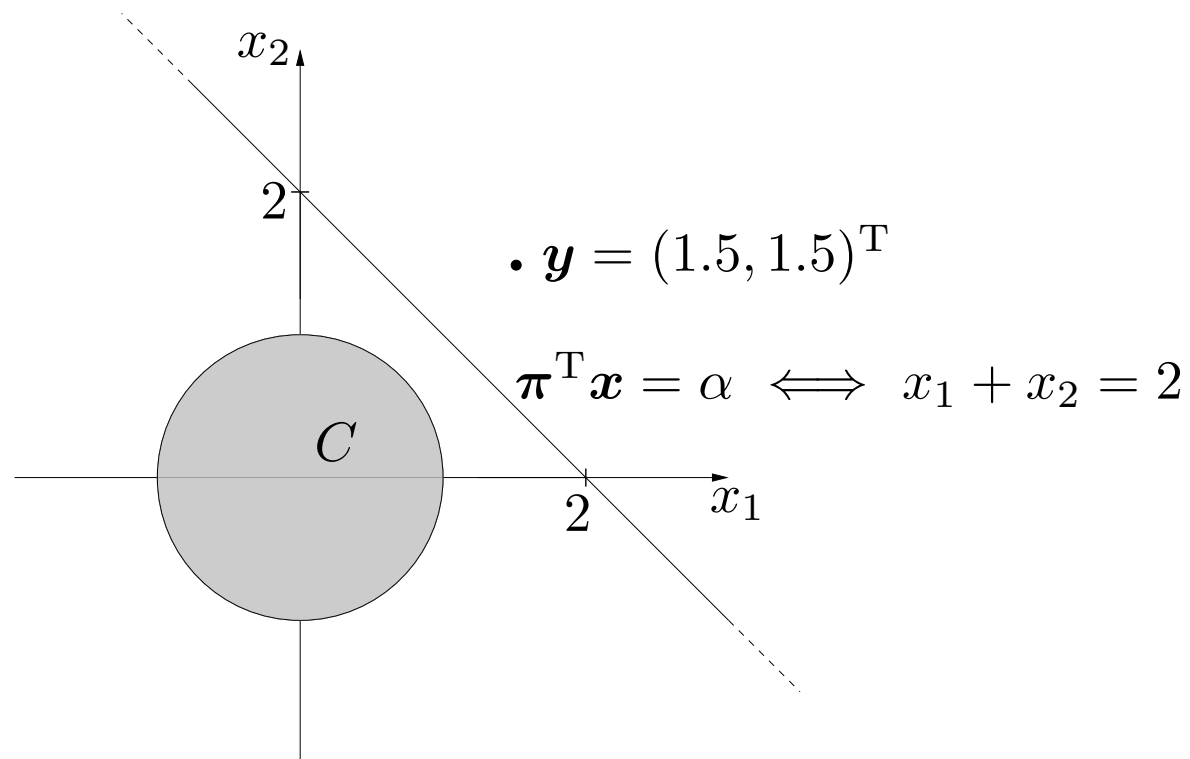- *A convex cone is finitely generated iff it is polyhedral*

Figure 9: Illustration of the Separation Theorem: the unit disk is separated from $\boldsymbol{y}$ by the line $\{\,\boldsymbol{x} \in \mathbb{R}^2 \mid x_1 + x_2 = 2\,\}$

Farkas' Lemma

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

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \tag{I}$$

$$\boldsymbol{x} \geq \boldsymbol{0}^n,$$

and

$$\boldsymbol{A}^{\mathrm{T}} \boldsymbol{\pi} \leq \boldsymbol{0}^n, \tag{II}$$

$$\boldsymbol{b}^{\mathrm{T}} \boldsymbol{\pi} > 0,$$

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

- Simple proof later!

Convexity of functions

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

$$\left. \begin{array}{l} x \in S \\[2mm] \lambda \in (0,1) \end{array} \right\} \implies f(\lambda \bar{x} + (1-\lambda)x) \leq \lambda f(\bar{x}) + (1-\lambda)f(x)$$

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

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

- 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 $-f$
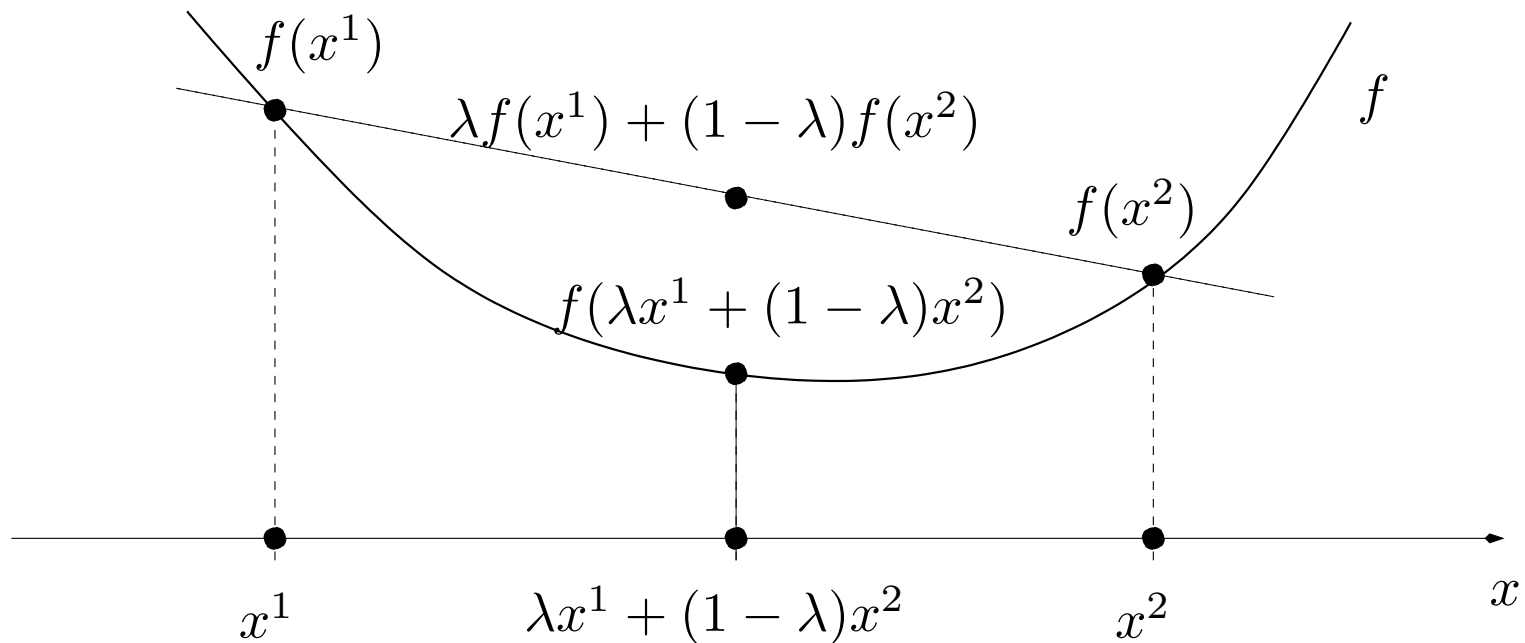
- Figure 10 illustrates a convex function

Figure 10: A convex function

- The function $f : \mathbb{R}^n \to \mathbb{R}$ defined by $f(\boldsymbol{x}) := \|\boldsymbol{x}\|$ is convex on $\mathbb{R}^n$; $f(\boldsymbol{x}) := \|\boldsymbol{x}\|^2$ is strictly convex in $\mathbb{R}^n$

- Let $\boldsymbol{c} \in \mathbb{R}^n$. The linear function $\boldsymbol{x} \mapsto f(\boldsymbol{x}) := \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} = \sum_{j=1}^{n} c_j x_j$ is both convex and concave on $\mathbb{R}^n$

- Figure 11 illustrates a non-convex function
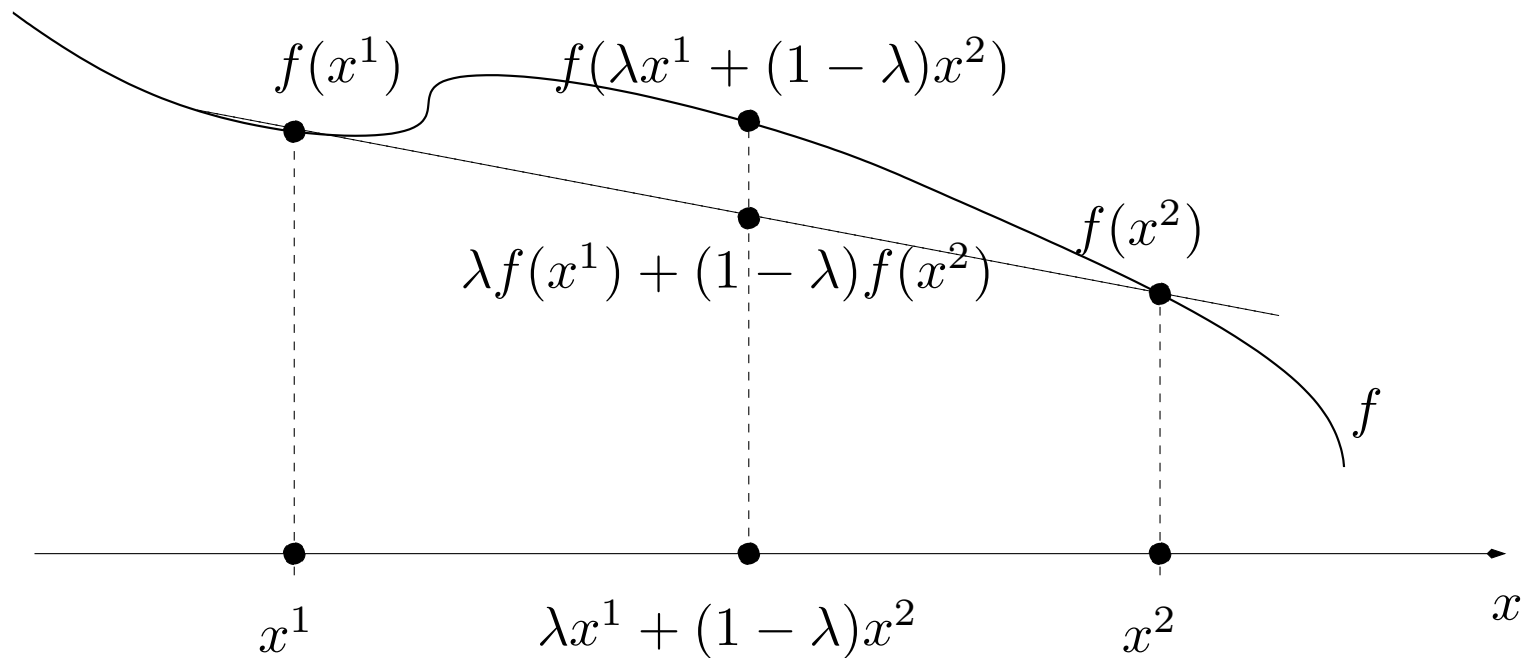
ag replacements

Figure 11: A non-convex function

- Sums of convex functions are convex

- Composite function: $\boldsymbol{x} \mapsto f(g(\boldsymbol{x}))$

- *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 $(y \geq x \Longrightarrow f(y) \geq f(x))$ on P. Then, the composite function $f(g)$ is convex on the set*
  $$\{\, \boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) \in P \,\}$$

- The function $\boldsymbol{x} \mapsto -\log(-g(\boldsymbol{x}))$ is convex on the set
  $$\{\, \boldsymbol{x} \in \mathbb{R}^n \mid g(\boldsymbol{x}) < 0 \,\}$$

Epigraphs

- 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 epi f]
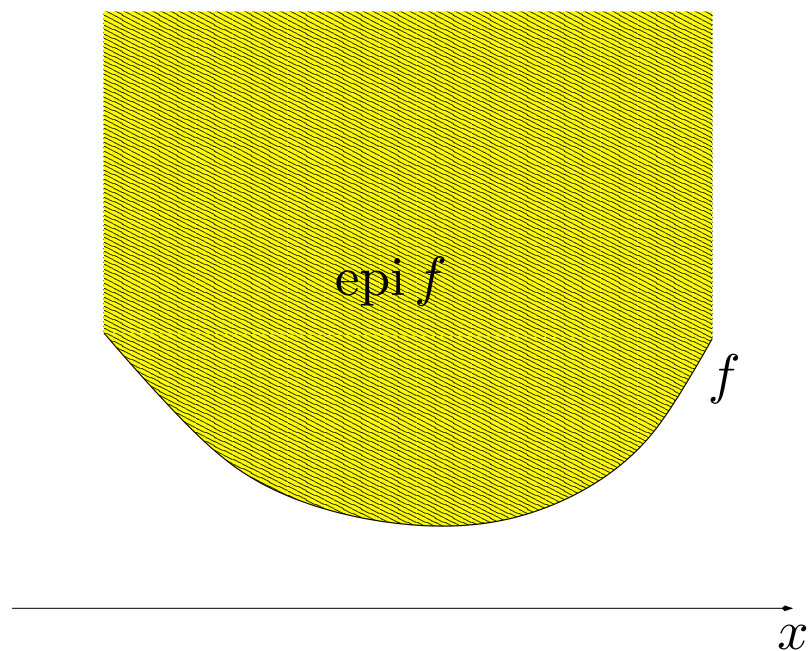
PSfrag replacements

epi f

f

x

Figure 12: A convex function and its epigraph

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

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

  The epigraph of the function f restricted to the set $S \subseteq \mathbb{R}^n$ is

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

- Connection between convex sets and functions; in fact the definition of a convex function stems from that of a convex set!

- *Suppose that $S \subseteq \mathbb{R}^n$ 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 $\mathbb{R}^{n+1}$*

Convexity characterizations in C^1

- C^1: Differentiable once, gradient continuous

- *Let $f \in C^1$ on an open convex set S*
  (a) *f is convex on $S \Longleftrightarrow f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^{\mathrm{T}}(\boldsymbol{y} - \boldsymbol{x}), \boldsymbol{x}, \boldsymbol{y} \in S$*
  (b) *f is convex on $S \Longleftrightarrow [\nabla f(\boldsymbol{x}) - \nabla f(\boldsymbol{y})]^{\mathrm{T}}(\boldsymbol{x} - \boldsymbol{y}) \geq 0, \boldsymbol{x}, \boldsymbol{y} \in S$*

- (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 Mean-value Theorem
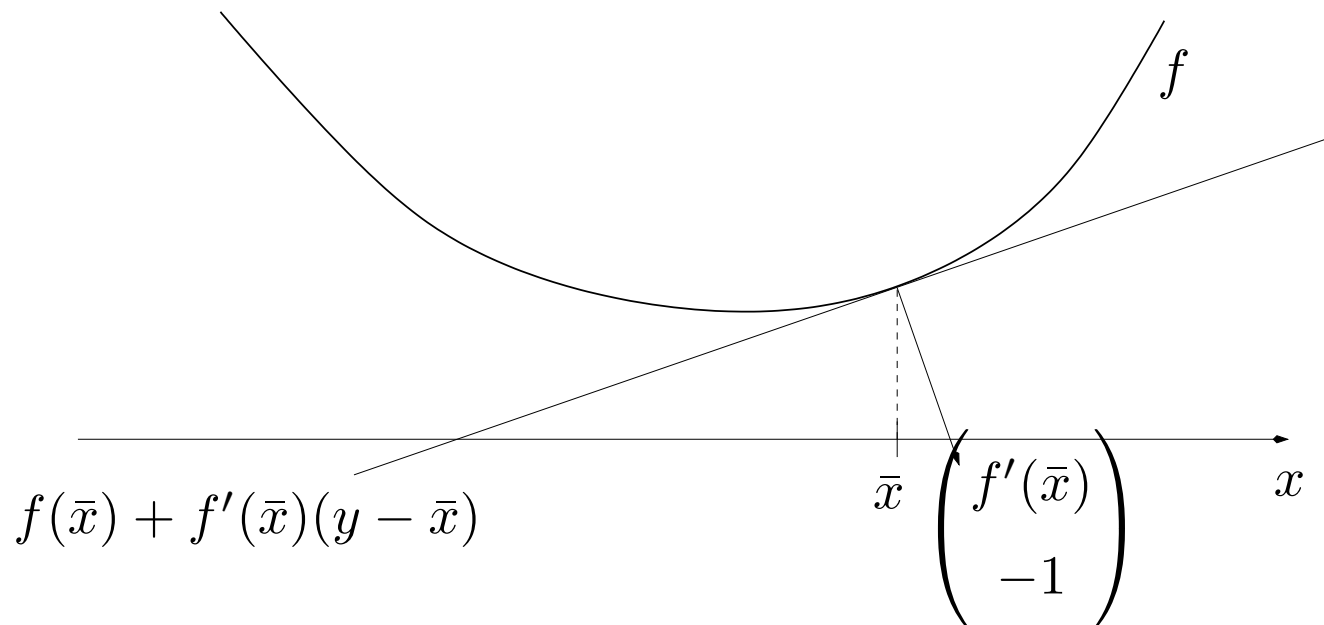
- Figure 13 illustrates part (a)

f

$f(\bar{x}) + f'(\bar{x})(y - \bar{x})$

$\bar{x}$

$\begin{pmatrix} f'(\bar{x}) \\ -1 \end{pmatrix}$

x

Figure 13: 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 $S \iff \nabla^2 f(\boldsymbol{x})$ is positive semidefinite for all $\boldsymbol{x} \in S$
  (b) $\nabla^2 f(\boldsymbol{x})$ is positive definite for all $x \in S \implies f$ 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 Mean-value Theorem

- Not the direction $\impliedby$ in (b)! $[f(x) = x^4$ at $x = 0]$

- Difficult to check convexity; matrix condition for every $\boldsymbol{x}$

- Quadratic function: $f(\boldsymbol{x}) = (1/2)\boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x} - \boldsymbol{q}^{\mathrm{T}}\boldsymbol{x}$ convex on $\mathbb{R}^n$ iff $\boldsymbol{Q}$ is psd ($\boldsymbol{Q}$ is the Hessian of f, and is independent of $\boldsymbol{x}$)

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

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

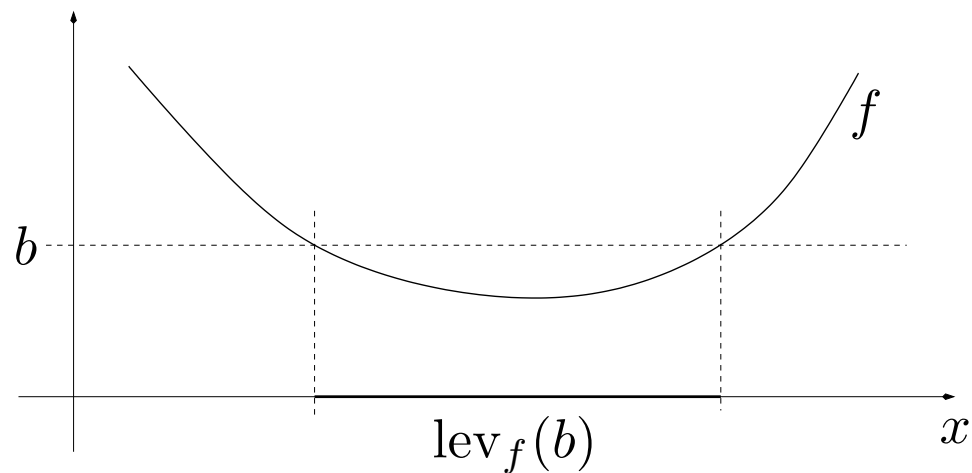- Figure 14 illustrates a level set of a convex function

PSfrag replacements

Figure 14: 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 $\text{lev}_g(b)$ is a convex set. It is moreover closed*

  Proof.

- We speak of a *convex problem* when f is convex (minimization) and for constraints $g_i(\boldsymbol{x}) \le 0$, the functions g_i are convex; and for constraints $h_j(\boldsymbol{x}) = 0$, the functions h_j are affine

*Euclidean projection

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

PSfrag replacements

$\boldsymbol{z}$

$\mathrm{Proj}_S(\boldsymbol{z})$
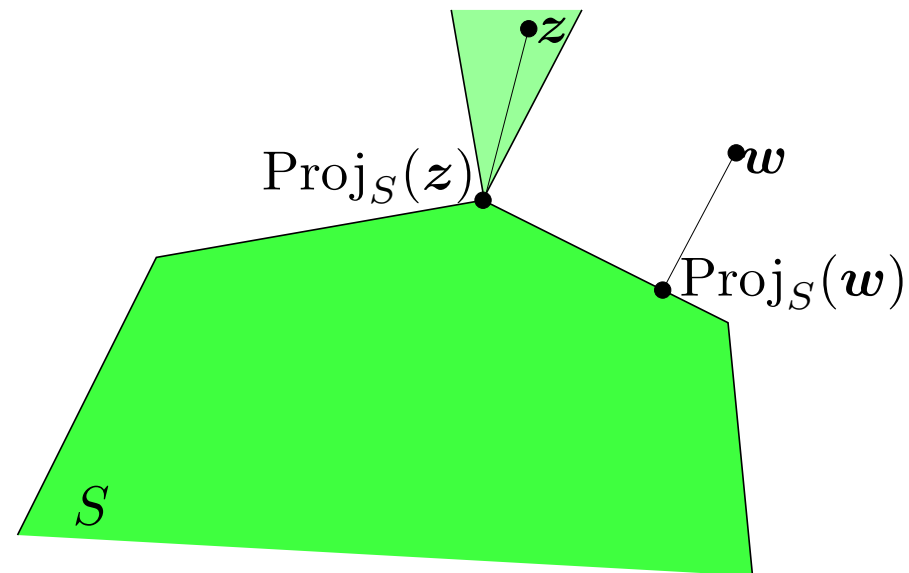
$\boldsymbol{w}$

$\mathrm{Proj}_S(\boldsymbol{w})$

S

Figure 15: The projection of two vectors onto a convex set

The distance function below is convex:

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

PSfrag replacements
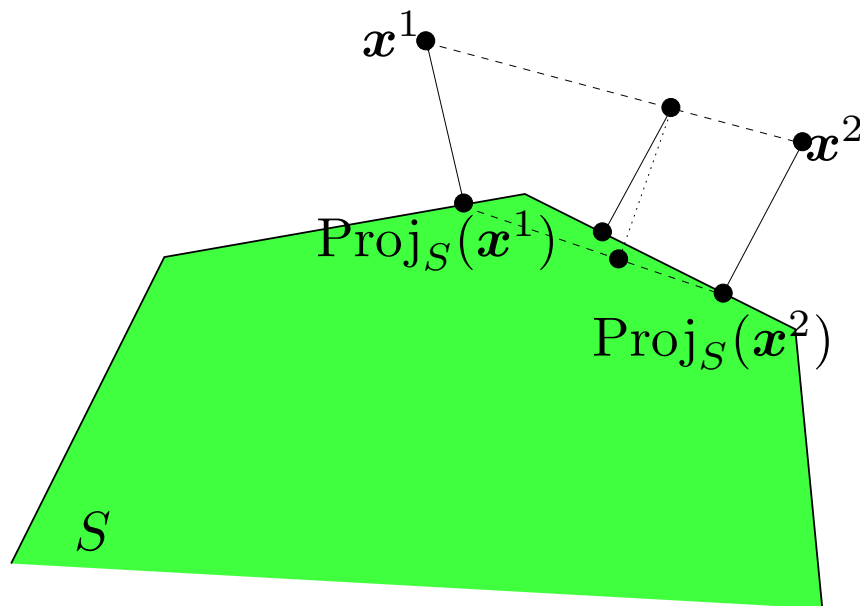
Figure 16: From the intermediate vector $\lambda\boldsymbol{x}^1 + (1 - \lambda)\boldsymbol{x}^2$ shown the distance to the vector $\lambda\mathrm{Proj}_S(\boldsymbol{x}^1) + (1 - \lambda)\mathrm{Proj}_S(\boldsymbol{x}^2))$ [dotted line segment] clearly is longer than to its projection on S [solid line]