LECTURE SLIDES ON

CONVEX OPTIMIZATION AND

DUALITY THEORY

TATA INSTITUTE FOR

FUNDAMENTAL RESEARCH

MUMBAI, INDIA

JANUARY 2009

PART I

BY DIMITRI P. BERTSEKAS

M.I.T.

http://web.mit.edu/dimitrib/www/home.html

LECTURE 1
INTRODUCTION/BASIC CONVEXITY CONCEPTS

LECTURE OUTLINE

- Convex Optimization Problems
- Why is Convexity Important in Optimization
- Multipliers and Lagrangian Duality
- Min Common/Max Crossing Duality
- Convex sets and functions
- Epigraphs
- Closed convex functions
- Recognizing convex functions

OPTIMIZATION PROBLEMS

- Generic form:

$$\begin{aligned} &\text{minimize} \;\; f(x) \\ &\text{subject to} \;\; x \in C \end{aligned}$$

Cost function $f : \Re^n \mapsto \Re$, constraint set C, e.g.,

$$\begin{aligned} C = X \cap \big\{x \mid h_1(x) = 0, \ldots, h_m(x) = 0\big\} \\ \cap \big\{x \mid g_1(x) \leq 0, \ldots, g_r(x) \leq 0\big\} \end{aligned}$$

- Examples of problem classifications:
  - Continuous vs discrete
  - Linear vs nonlinear
  - Deterministic vs stochastic
  - Static vs dynamic

- Convex programming problems are those for which f is convex and C is convex (they are continuous problems).

- However, convexity permeates all of optimization, including discrete problems.

WHY IS CONVEXITY SO SPECIAL?

- A convex function has no local minima that are not global
- A convex set has a nonempty relative interior
- A convex set is connected and has feasible directions at any point
- A nonconvex function can be "convexified" while maintaining the optimality of its global minima
- The existence of a global minimum of a convex function over a convex set is conveniently characterized in terms of directions of recession
- A polyhedral convex set is characterized in terms of a finite set of extreme points and extreme directions
- A real-valued convex function is continuous and has nice differentiability properties
- Closed convex cones are self-dual with respect to polarity
- Convex, lower semicontinuous functions are self-dual with respect to conjugacy

CONVEXITY AND DUALITY

- Consider the (primal) problem

$$\text{minimize} \quad f(x) \quad \text{s.t.} \qquad g_1(x) \le 0, \ldots, g_r(x) \le 0$$

- We introduce multiplier vectors $\mu = (\mu_1, \ldots, \mu_r) \ge 0$ and form the Lagrangian function

$$L(x, \mu) = f(x) + \sum_{j=1}^{r} \mu_j g_j(x), \qquad x \in \Re^n, \; \mu \in \Re^r.$$

- Dual function

$$q(\mu) = \inf_{x \in \Re^n} L(x, \mu)$$

- Dual problem: Maximize $q(\mu)$ over $\mu \ge 0$
- **Motivation**: Under favorable circumstances (strong duality) the optimal values of the primal and dual problems are equal, and their optimal solutions are related

KEY DUALITY RELATIONS

- Optimal primal value

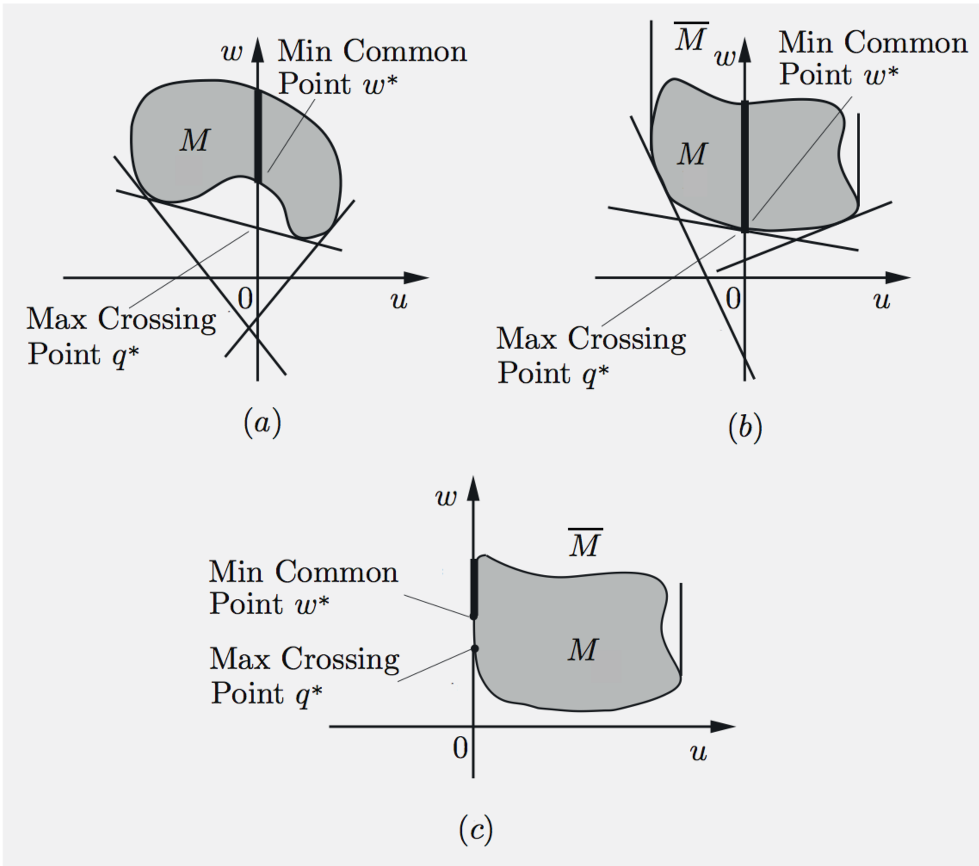
$$f^* = \inf_{g_j(x) \leq 0,\, j=1,\ldots,r} f(x) = \inf_{x \in \Re^n} \sup_{\mu \geq 0} L(x, \mu)$$

- Optimal dual value

$$q^* = \sup_{\mu \geq 0} q(\mu) = \sup_{\mu \geq 0} \inf_{x \in \Re^n} L(x, \mu)$$

- We always have $q^* \leq f^*$ (weak duality - important in discrete optimization problems).
- Under favorable circumstances (convexity in the primal problem, plus ...):
  - We have $q^* = f^*$ (strong duality)
  - If μ^* is optimal dual solution, all optimal primal solutions minimize $L(x, \mu^*)$
- This opens a wealth of analytical and computational possibilities, and insightful interpretations.
- Note that the equality of "sup inf" and "inf sup" is a key issue in minimax theory and game theory.

MIN COMMON/MAX CROSSING DUALITY



- All of duality theory and all of (convex/concave) minimax theory can be developed/explained in terms of this one figure.

- This is the novel aspect of the treatment (although the ideas are closely connected to conjugate convex function theory)

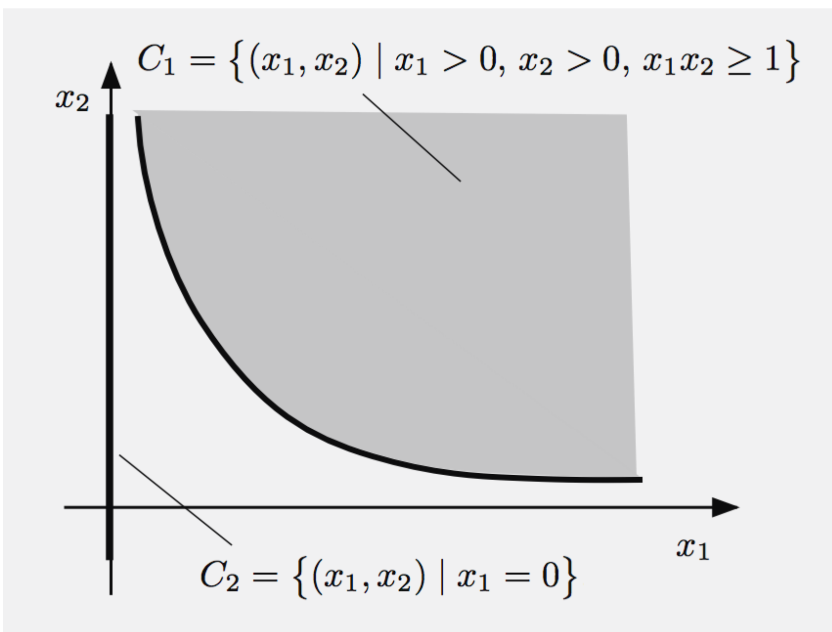
- The machinery of convex analysis is needed to flesh out this figure, and to rule out the exceptional/pathological behavior shown in (c).

EXCEPTIONAL BEHAVIOR

- If convex structure is so favorable, what is the source of exceptional/pathological behavior [like in (c) of the preceding slide]?

- **Answer:** Some common operations on convex sets do not preserve some basic properties.

- **Example:** A linearly transformed closed convex set need not be closed (contrary to compact and polyhedral sets).



- This is a major reason for the analytical difficulties in convex analysis and pathological behavior in convex optimization (and the favorable character of polyhedral sets).

COURSE OUTLINE

1) **Basic Convexity Concepts (2):** Convex sets and functions. Convex and affine hulls. Closure, relative interior, and continuity.

2) **More Convexity Concepts (2):** Directions of recession. Hyperplanes. Conjugate convex functions.

3) **Convex Optimization Concepts (1):** Existence of optimal solutions. Partial minimization. Saddle point and minimax theory.

4) **Min common/max crossing duality (1):** MC/MC duality. Special cases in constrained minimization and minimax. Strong duality theorem. Existence of dual optimal solutions.

5) **Duality applications (2):** Constrained optimization (Lagrangian, Fenchel, and conic duality). Subdifferential theory and optimality conditions. Minimax theorems. Nonconvex problems and estimates of the duality gap.

WHAT TO EXPECT FROM THIS COURSE

- We aim:
  - To develop insight and deep understanding of a fundamental optimization topic
  - To treat rigorously an important branch of applied math, and to provide some appreciation of the research in the field
- Mathematical level:
  - Prerequisites are linear algebra (preferably abstract) and real analysis (a course in each)
  - Proofs are important ... but the rich geometry helps guide the mathematics
- We will make maximum use of visualization and figures
- Applications: They are many and pervasive ... but don't expect much in this course. The book by Boyd and Vandenberghe describes a lot of practical convex optimization models (http://www.stanford.edu/ boyd/cvxbook.html)
- Handouts: Slides, 1st chapter, material in http://www.athenasc.com/convexity.html

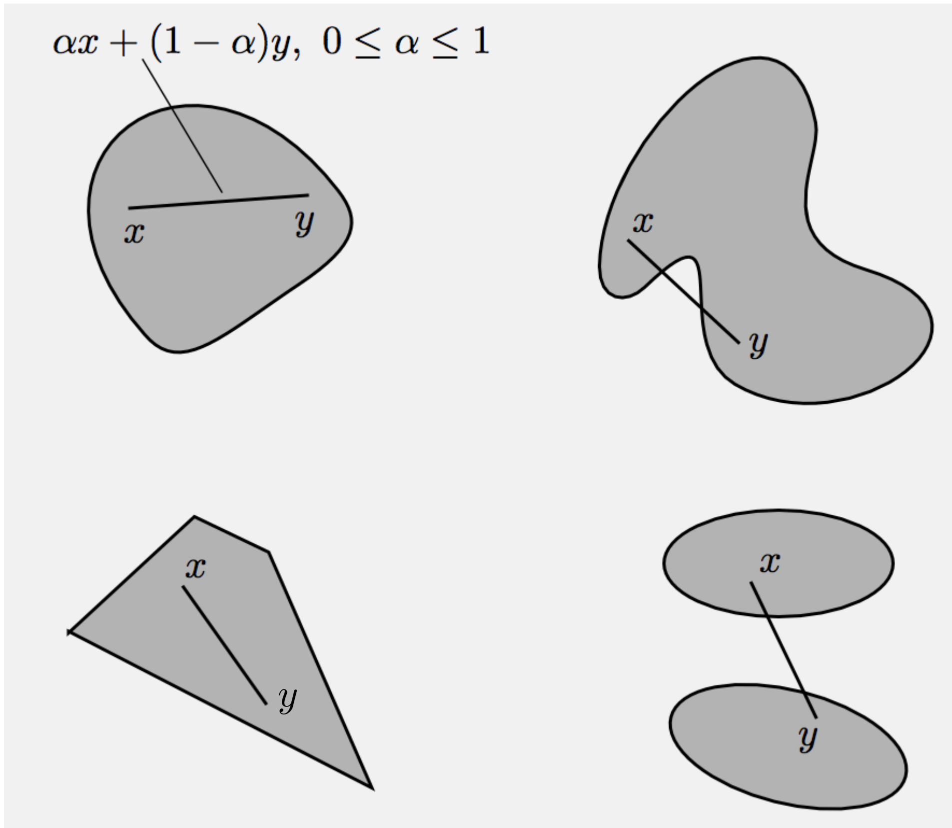
A NOTE ON THESE SLIDES

- These slides are a teaching aid, not a text
- Don't expect strict mathematical rigor
- The statements of theorems are fairly precise, but the proofs are not
- Many proofs have been omitted or greatly abbreviated
- Figures are meant to convey and enhance ideas, not to express them precisely
- The omitted proofs and a much fuller discussion can be found in the "Convex Optimization" textbook and handouts

SOME MATH CONVENTIONS

- All of our work is done in $\Re^n$: space of n-tuples $x = (x_1, \ldots, x_n)$
- All vectors are assumed column vectors
- "$'$" denotes transpose, so we use x' to denote a row vector
- $x'y$ is the inner product $\sum_{i=1}^n x_i y_i$ of vectors x and y
- $\|x\| = \sqrt{x'x}$ is the (Euclidean) norm of x. We use this norm almost exclusively
- See the appendix for an overview of the linear algebra and real analysis background that we will use

CONVEX SETS



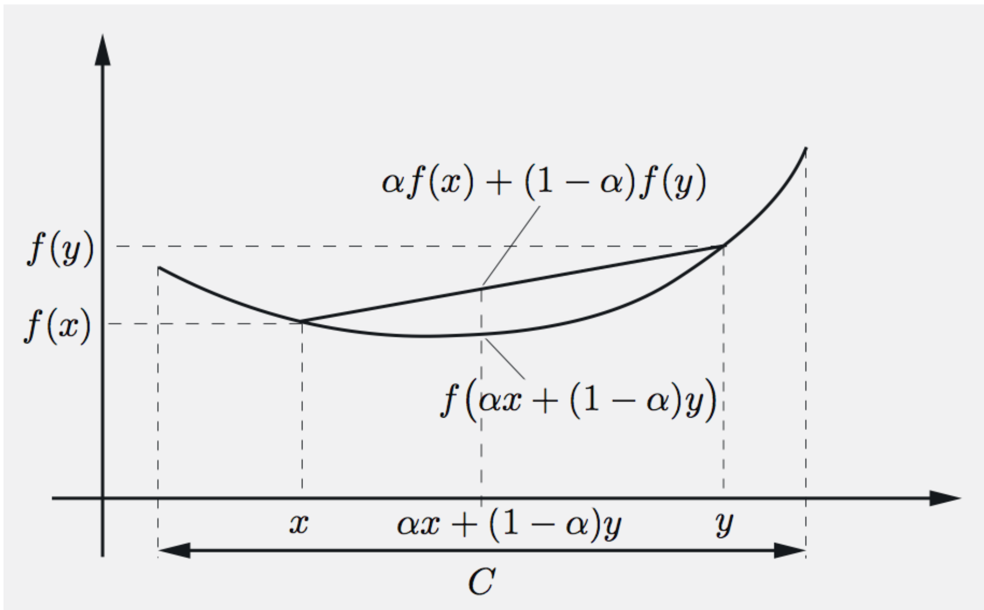
- A subset C of $\Re^n$ is called convex if

$$\alpha x + (1-\alpha)y \in C, \qquad \forall\ x, y \in C,\ \forall\ \alpha \in [0,1]$$

- Operations that preserve convexity
  - Intersection, scalar multiplication, vector sum, closure, interior, linear transformations

- Cones: Sets C such that $\lambda x \in C$ for all $\lambda > 0$ and $x \in C$ (not always convex or closed)

REAL-VALUED CONVEX FUNCTIONS



- Let C be a convex subset of $\Re^n$. A function $f : C \mapsto \Re$ is called *convex* if

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha) f(y)$$

for all $x, y \in C$, and $\alpha \in [0,1]$.

- If f is a convex function, then all its level sets $\{x \in C \mid f(x) \le a\}$ and $\{x \in C \mid f(x) < a\}$, where a is a scalar, are convex.

EXTENDED REAL-VALUED FUNCTIONS

- The *epigraph* of a function $f : X \mapsto [-\infty, \infty]$ is the subset of $\Re^{n+1}$ given by

$$\text{epi}(f) = \big\{(x, w) \mid x \in X,\ w \in \Re,\ f(x) \le w\big\}$$

- The *effective domain* of f is the set

$$\text{dom}(f) = \big\{x \in X \mid f(x) < \infty\big\}$$

- We say that f is *proper* if $f(x) < \infty$ for at least one $x \in X$ and $f(x) > -\infty$ for all $x \in X$, and we will call f *improper* if it is not proper.

- Note that f is proper if and only if its epigraph is nonempty and does not contain a "vertical line."

- An extended real-valued function $f : X \mapsto [-\infty, \infty]$ is called *lower semicontinuous* at a vector $x \in X$ if $f(x) \le \liminf_{k\to\infty} f(x_k)$ for every sequence $\{x_k\} \subset X$ with $x_k \to x$.

- We say that f is *closed* if $\text{epi}(f)$ is a closed set.

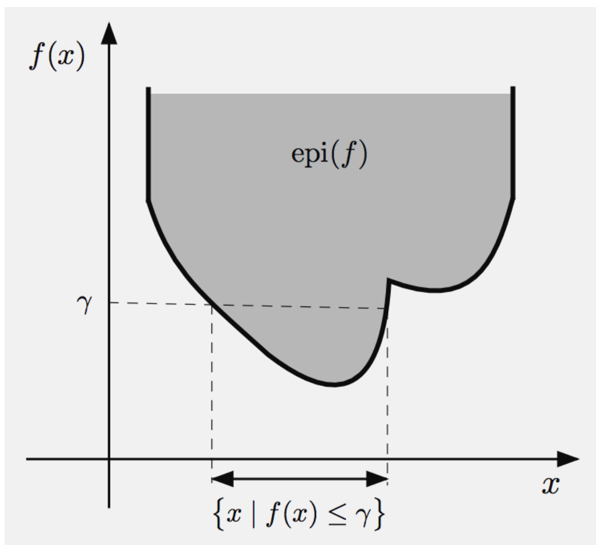
CLOSEDNESS AND SEMICONTINUITY

- *Proposition*: For a function $f : \Re^n \mapsto [-\infty, \infty]$, the following are equivalent:

   (i) $\{x \mid f(x) \leq a\}$ is closed for every scalar a.

   (ii) f is lower semicontinuous at all $x \in \Re^n$.

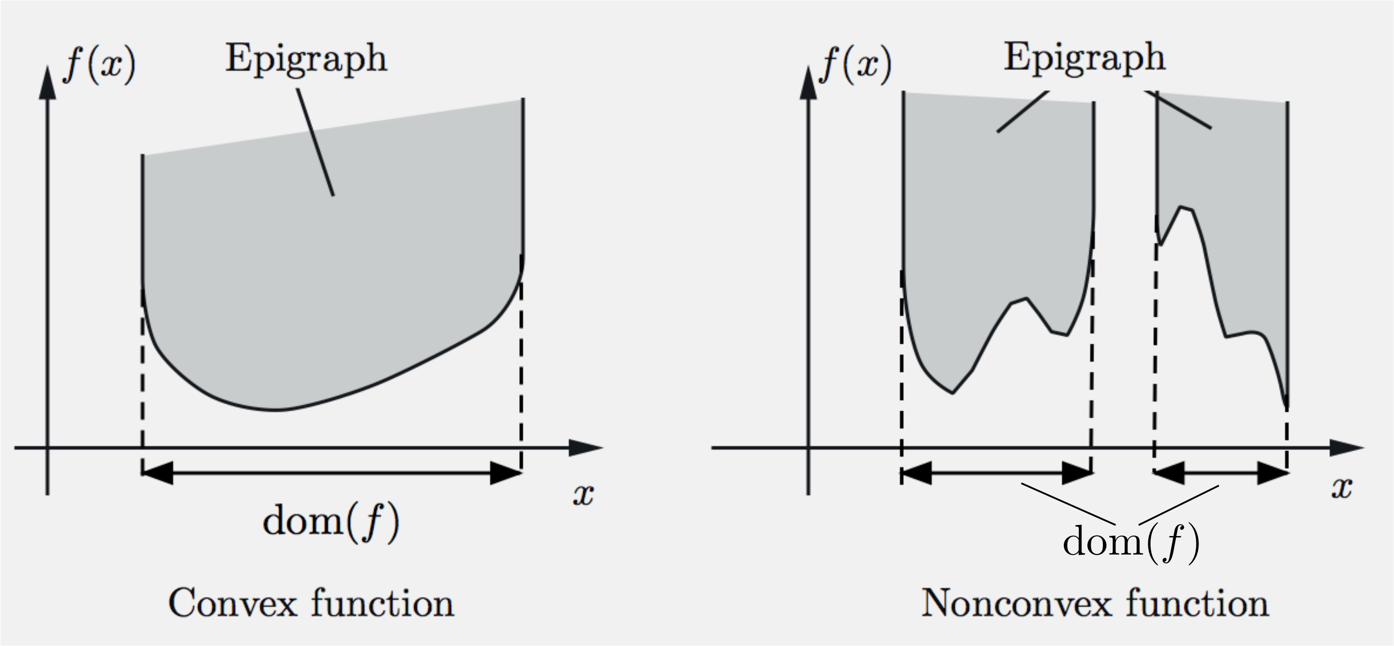
   (iii) f is closed.



- Note that:
   - If f is lower semicontinuous at all $x \in \text{dom}(f)$, it is not necessarily closed
   - If f is closed, $\text{dom}(f)$ is not necessarily closed

- *Proposition*: Let $f : X \mapsto [-\infty, \infty]$ be a function. If $\text{dom}(f)$ is closed and f is lower semicontinuous at all $x \in \text{dom}(f)$, then f is closed.

EXTENDED REAL-VALUED CONVEX FUNCTIONS



- Let C be a convex subset of $\Re^n$. An extended real-valued function $f : C \mapsto [-\infty, \infty]$ is called *convex* if $\text{epi}(f)$ is a convex subset of $\Re^{n+1}$.

- If f is proper, this definition is equivalent to

$$f\big(\alpha x + (1-\alpha)y\big) \le \alpha f(x) + (1-\alpha)f(y)$$

for all $x, y \in C$, and $\alpha \in [0,1]$.

- An improper *closed* convex function is very peculiar: it takes an infinite value (∞ or $-\infty$) at every point.

RECOGNIZING CONVEX FUNCTIONS

• Some important classes of elementary convex functions: Affine functions, positive semidefinite quadratic functions, norm functions, etc.

• *Proposition*: Let $f_i : \Re^n \mapsto (-\infty, \infty]$, $i \in I$, be given functions (I is an arbitrary index set).
(a) The function $g : \Re^n \mapsto (-\infty, \infty]$ given by

$$g(x) = \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x), \qquad \lambda_i > 0$$

is convex (or closed) if $f_1, \ldots, f_m$ are convex (respectively, closed).
(b) The function $g : \Re^n \mapsto (-\infty, \infty]$ given by

$$g(x) = f(Ax)$$

where A is an $m \times n$ matrix is convex (or closed) if f is convex (respectively, closed).
(c) The function $g : \Re^n \mapsto (-\infty, \infty]$ given by

$$g(x) = \sup_{i \in I} f_i(x)$$

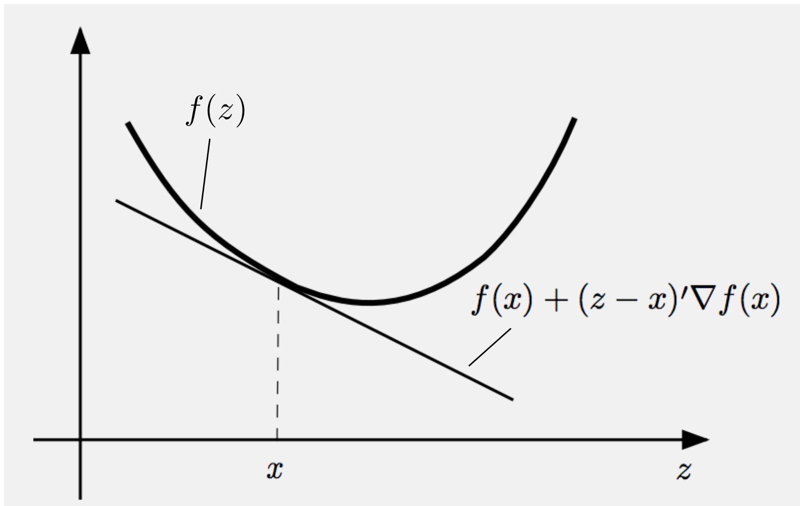
is convex (or closed) if the f_i are convex (respectively, closed).

LECTURE 2

LECTURE OUTLINE

- Differentiable Convex Functions
- Convex and Affine Hulls
- Caratheodory's Theorem
- Closure, Relative Interior, Continuity

DIFFERENTIABLE CONVEX FUNCTIONS



- Let $C \subset \Re^n$ be a convex set and let $f : \Re^n \mapsto \Re$ be differentiable over $\Re^n$.

  (a) The function f is convex over C iff

  $$f(z) \geq f(x) + (z - x)'\nabla f(x), \qquad \forall\, x, z \in C$$

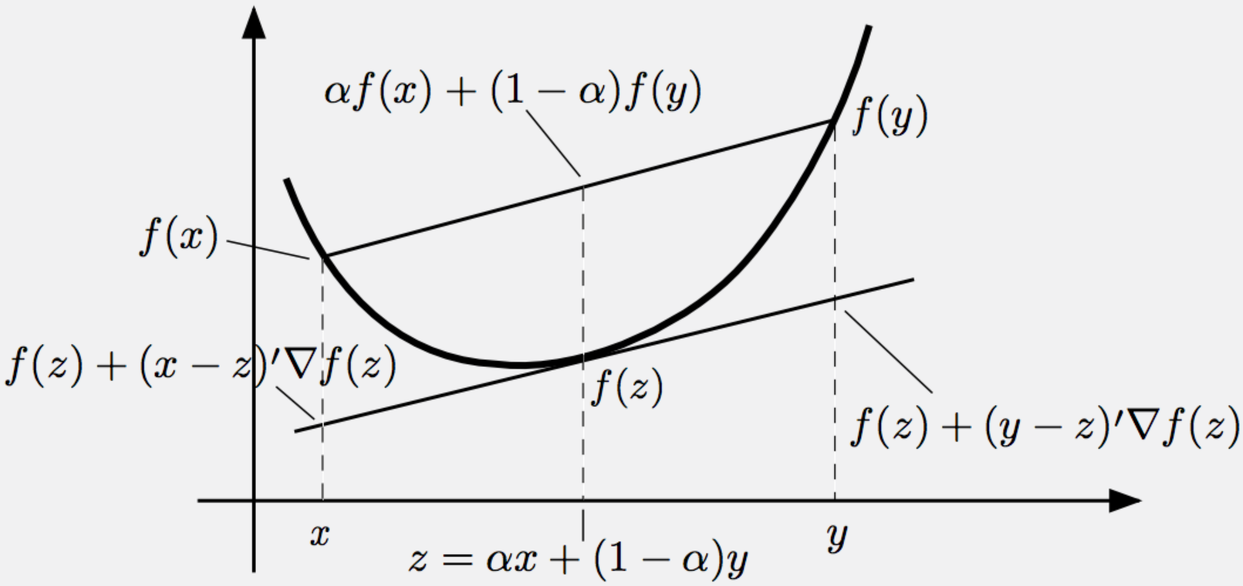
  Implies that x^* minimizes f over C iff

  $$\nabla f(x^*)'(x - x^*) \geq 0, \ \forall\, x \in C$$

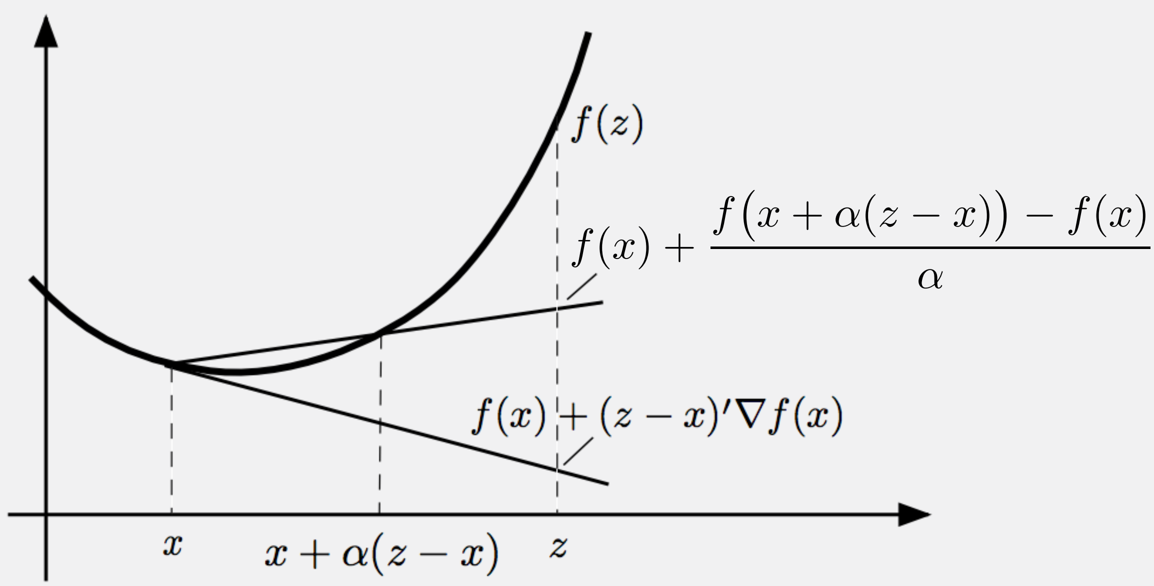
  (b) If the inequality is strict whenever $x \neq z$, then f is strictly convex over C, i.e., for all $\alpha \in (0, 1)$ and $x, y \in C$, with $x \neq y$

  $$f\big(\alpha x + (1 - \alpha)y\big) < \alpha f(x) + (1 - \alpha)f(y)$$

PROOF IDEAS



(a)



(b)

TWICE DIFFERENTIABLE CONVEX FUNCTIONS

- Let C be a convex subset of $\Re^n$ and let $f : \Re^n \mapsto \Re$ be twice continuously differentiable over $\Re^n$.

  (a) If $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$, then f is convex over C.

  (b) If $\nabla^2 f(x)$ is positive definite for all $x \in C$, then f is strictly convex over C.

  (c) If C is open and f is convex over C, then $\nabla^2 f(x)$ is positive semidefinite for all $x \in C$.

Proof: (a) By mean value theorem, for $x, y \in C$

$$f(y) = f(x)+(y-x)'\nabla f(x)+\tfrac{1}{2}(y-x)'\nabla^2 f\big(x+\alpha(y-x)\big)(y-x)$$

for some $\alpha \in [0, 1]$. Using the positive semidefiniteness of $\nabla^2 f$, we obtain

$$f(y) \geq f(x) + (y - x)'\nabla f(x), \qquad \forall\, x, y \in C$$

From the preceding result, f is convex.

(b) Similar to (a), we have $f(y) > f(x) + (y - x)'\nabla f(x)$ for all $x, y \in C$ with $x \neq y$, and we use the preceding result.

CONVEX AND AFFINE HULLS

- Given a set $X \subset \Re^n$:

- A *convex combination* of elements of X is a vector of the form $\sum_{i=1}^m \alpha_i x_i$, where $x_i \in X$, $\alpha_i \geq 0$, and $\sum_{i=1}^m \alpha_i = 1$.

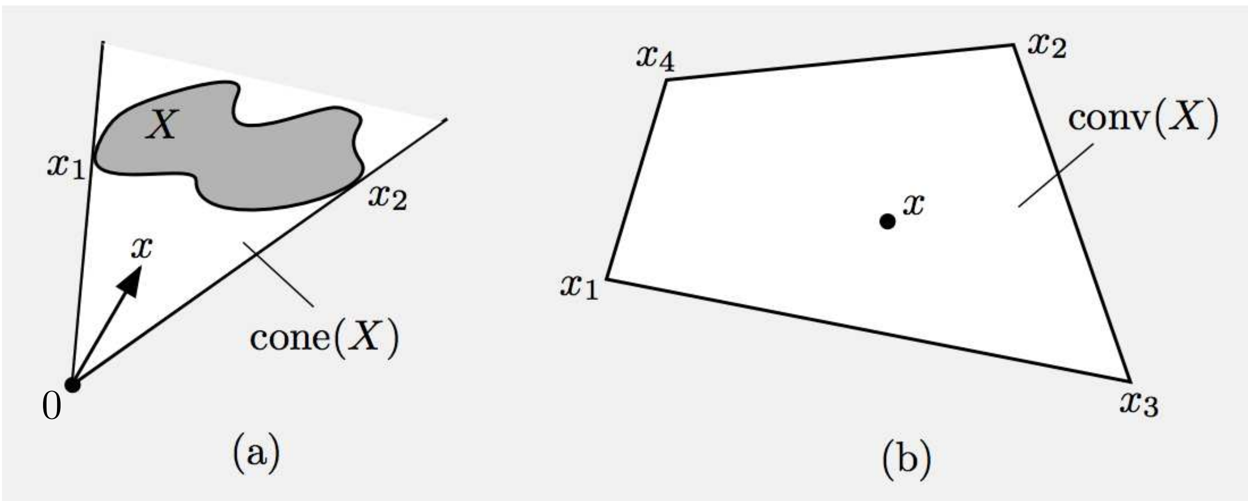
- The *convex hull* of X, denoted $\text{conv}(X)$, is the intersection of all convex sets containing X (also the set of all convex combinations from X).

- The *affine hull* of X, denoted $\text{aff}(X)$, is the intersection of all affine sets containing X (an affine set is a set of the form $\overline{x} + S$, where S is a subspace). Note that $\text{aff}(X)$ is itself an affine set.

- A *nonnegative combination* of elements of X is a vector of the form $\sum_{i=1}^m \alpha_i x_i$, where $x_i \in X$ and $\alpha_i \geq 0$ for all i.

- The *cone generated by* X, denoted $\text{cone}(X)$, is the set of all nonnegative combinations from X:
  - It is a convex cone containing the origin.
  - It need not be closed (even if X is compact).
  - If X is a finite set, $\text{cone}(X)$ is closed (nontrivial to show!)

CARATHEODORY'S THEOREM



- Let X be a nonempty subset of $\Re^n$.
  (a) Every $x \neq 0$ in cone(X) can be represented as a positive combination of vectors $x_1, \ldots, x_m$ from X that are linearly independent (so $m \leq n$).
  (b) Every $x \notin X$ that belongs to conv(X) can be represented as a convex combination of at most $n+1$ vectors.

PROOF OF CARATHEODORY'S THEOREM

(a) Let x be a nonzero vector in cone(X), and let m be the smallest integer such that x has the form $\sum_{i=1}^{m} \alpha_i x_i$, where $\alpha_i > 0$ and $x_i \in X$ for all $i = 1, \ldots, m$. If the vectors x_i were linearly dependent, there would exist $\lambda_1, \ldots, \lambda_m$, with

$$\sum_{i=1}^{m} \lambda_i x_i = 0$$

and at least one of the λ_i is positive. Consider

$$\sum_{i=1}^{m} (\alpha_i - \overline{\gamma}\lambda_i) x_i,$$

where $\overline{\gamma}$ is the largest γ such that $\alpha_i - \gamma\lambda_i \geq 0$ for all i. This combination provides a representation of x as a positive combination of fewer than m vectors of X – a contradiction. Therefore, $x_1, \ldots, x_m$, are linearly independent.

(b) Apply part (a) to the subset of $\Re^{n+1}$

$$Y = \{(z, 1) \mid z \in X\}$$

consider cone(Y), represent $(x, 1) \in$ cone(Y) in terms of at most $n + 1$ vectors, etc.

AN APPLICATION OF CARATHEODORY

- The convex hull of a compact set is compact.

Proof: Let X be compact. We take a sequence in $\mathrm{conv}(X)$ and show that it has a convergent subsequence whose limit is in $\mathrm{conv}(X)$.

By Caratheodory, a sequence in $\mathrm{conv}(X)$ can be expressed as $\left\{\sum_{i=1}^{n+1} \alpha_i^k x_i^k\right\}$, where for all k and i, $\alpha_i^k \geq 0$, $x_i^k \in X$, and $\sum_{i=1}^{n+1} \alpha_i^k = 1$. Since

$$\left\{(\alpha_1^k, \ldots, \alpha_{n+1}^k, x_1^k, \ldots, x_{n+1}^k)\right\}$$

is bounded, it has a limit point

$$\left\{(\alpha_1, \ldots, \alpha_{n+1}, x_1, \ldots, x_{n+1})\right\},$$

which must satisfy $\sum_{i=1}^{n+1} \alpha_i = 1$, and $\alpha_i \geq 0$, $x_i \in X$ for all i. Thus, the vector $\sum_{i=1}^{n+1} \alpha_i x_i$, which belongs to $\mathrm{conv}(X)$, is a limit point of the sequence $\left\{\sum_{i=1}^{n+1} \alpha_i^k x_i^k\right\}$, so $\mathrm{conv}(X)$ is compact. **Q.E.D.**

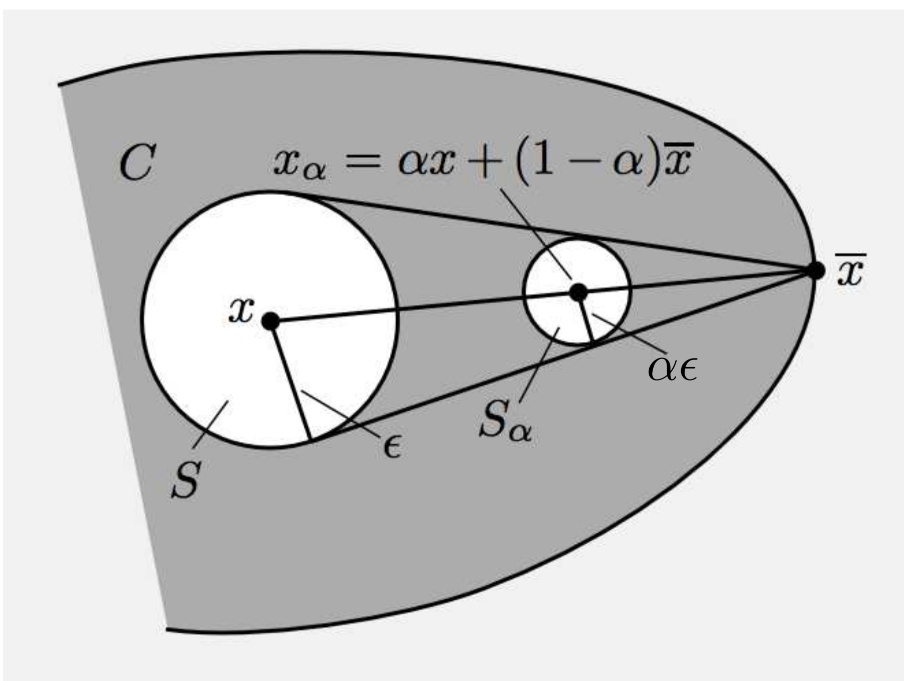
- Note the convex hull of a closed set need not be closed.

RELATIVE INTERIOR

- x is a *relative interior point* of C, if x is an interior point of C relative to $\text{aff}(C)$.

- $\text{ri}(C)$ denotes the *relative interior of* C, i.e., the set of all relative interior points of C.

- **Line Segment Principle**: If C is a convex set, $x \in \text{ri}(C)$ and $\overline{x} \in \text{cl}(C)$, then all points on the line segment connecting x and $\overline{x}$, except possibly $\overline{x}$, belong to $\text{ri}(C)$.

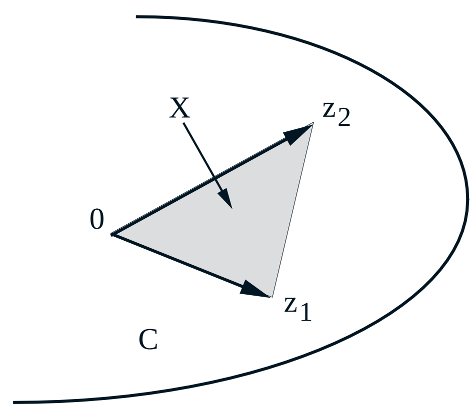


ADDITIONAL MAJOR RESULTS

- Let C be a nonempty convex set.

  (a) $\mathrm{ri}(C)$ is a nonempty convex set, and has the same affine hull as C.

  (b) $x \in \mathrm{ri}(C)$ if and only if every line segment in C having x as one endpoint can be prolonged beyond x without leaving C.



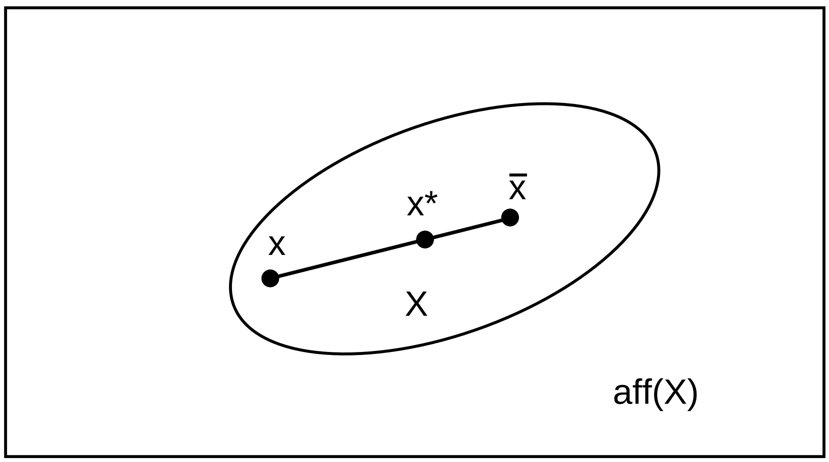
Proof: (a) Assume that $0 \in C$. We choose m linearly independent vectors $z_1, \ldots, z_m \in C$, where m is the dimension of $\mathrm{aff}(C)$, and we let

$$X = \left\{ \sum_{i=1}^{m} \alpha_i z_i \;\middle|\; \sum_{i=1}^{m} \alpha_i < 1,\ \alpha_i > 0,\ i = 1, \ldots, m \right\}$$

Then argue that $X \subset \mathrm{ri}(C)$.
(b) => is clear by the def. of rel. interior. Reverse: argue by contradiction; take any $\overline{x} \in \mathrm{ri}(C)$; use prolongation assumption and Line Segment Princ.

OPTIMIZATION APPLICATION

- A concave function $f : \Re^n \mapsto \Re$ that attains its minimum over a convex set X at an $x^* \in \mathrm{ri}(X)$ must be constant over X.



Proof: (By contradiction.) Let $x \in X$ be such that $f(x) > f(x^*)$. Prolong beyond x^* the line segment x-to-x^* to a point $\overline{x} \in X$. By concavity of f, we have for some $\alpha \in (0,1)$

$$f(x^*) \geq \alpha f(x) + (1-\alpha) f(\overline{x}),$$

and since $f(x) > f(x^*)$, we must have $f(x^*) > f(\overline{x})$ - a contradiction. **Q.E.D.**

- **Corollary:** A linear function can attain a mininum only at the boundary of a convex set.

CALCULUS OF RELATIVE INTERIORS: SUMMARY

- The relative interior of a convex set is equal to the relative interior of its closure.
- The closure of the relative interior of a convex set is equal to its closure.
- Relative interior and closure commute with Cartesian product and inverse image under a linear transformation.
- Relative interior commutes with image under a linear transformation and vector sum, but closure does not.
- Neither relative interior nor closure commute with set intersection.

CLOSURE VS RELATIVE INTERIOR

- Let C be a nonempty convex set. Then $\mathrm{ri}(C)$ and $\mathrm{cl}(C)$ are "not too different for each other."

- *Proposition*:

  (a) We have $\mathrm{cl}(C) = \mathrm{cl}\big(\mathrm{ri}(C)\big)$.

  (b) We have $\mathrm{ri}(C) = \mathrm{ri}\big(\mathrm{cl}(C)\big)$.

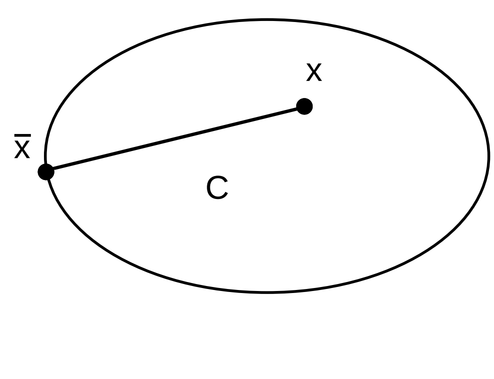
  (c) Let $\overline{C}$ be another nonempty convex set. Then the following three conditions are equivalent:

    (i) C and $\overline{C}$ have the same rel. interior.

    (ii) C and $\overline{C}$ have the same closure.

    (iii) $\mathrm{ri}(C) \subset \overline{C} \subset \mathrm{cl}(C)$.

Proof: (a) Since $\mathrm{ri}(C) \subset C$, we have $\mathrm{cl}\big(\mathrm{ri}(C)\big) \subset \mathrm{cl}(C)$. Conversely, let $\overline{x} \in \mathrm{cl}(C)$. Let $x \in \mathrm{ri}(C)$. By the Line Segment Principle, we have $\alpha x + (1 - \alpha)\overline{x} \in \mathrm{ri}(C)$ for all $\alpha \in (0, 1]$. Thus, $\overline{x}$ is the limit of a sequence that lies in $\mathrm{ri}(C)$, so $\overline{x} \in \mathrm{cl}\big(\mathrm{ri}(C)\big)$.



LINEAR TRANSFORMATIONS

- Let C be a nonempty convex subset of $\Re^n$ and let A be an $m \times n$ matrix.

  (a) We have $A \cdot \text{ri}(C) = \text{ri}(A \cdot C)$.

  (b) We have $A \cdot \text{cl}(C) \subset \text{cl}(A \cdot C)$. Furthermore, if C is bounded, then $A \cdot \text{cl}(C) = \text{cl}(A \cdot C)$.

Proof: (a) Intuition: Spheres within C are mapped onto spheres within $A \cdot C$ (relative to the affine hull).

(b) We have $A \cdot \text{cl}(C) \subset \text{cl}(A \cdot C)$, since if a sequence $\{x_k\} \subset C$ converges to some $x \in \text{cl}(C)$ then the sequence $\{Ax_k\}$, which belongs to $A \cdot C$, converges to Ax, implying that $Ax \in \text{cl}(A \cdot C)$.

To show the converse, assuming that C is bounded, choose any $z \in \text{cl}(A \cdot C)$. Then, there exists a sequence $\{x_k\} \subset C$ such that $Ax_k \to z$. Since C is bounded, $\{x_k\}$ has a subsequence that converges to some $x \in \text{cl}(C)$, and we must have $Ax = z$. It follows that $z \in A \cdot \text{cl}(C)$. **Q.E.D.**

Note that in general, we may have

$$A \cdot \text{int}(C) \neq \text{int}(A \cdot C), \qquad A \cdot \text{cl}(C) \neq \text{cl}(A \cdot C)$$

INTERSECTIONS AND VECTOR SUMS

- Let C_1 and C_2 be nonempty convex sets.

  (a) We have

$$\mathrm{ri}(C_1 + C_2) = \mathrm{ri}(C_1) + \mathrm{ri}(C_2),$$

$$\mathrm{cl}(C_1) + \mathrm{cl}(C_2) \subset \mathrm{cl}(C_1 + C_2)$$

  If one of C_1 and C_2 is bounded, then

$$\mathrm{cl}(C_1) + \mathrm{cl}(C_2) = \mathrm{cl}(C_1 + C_2)$$

  (b) If $\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) \neq \emptyset$, then

$$\mathrm{ri}(C_1 \cap C_2) = \mathrm{ri}(C_1) \cap \mathrm{ri}(C_2),$$

$$\mathrm{cl}(C_1 \cap C_2) = \mathrm{cl}(C_1) \cap \mathrm{cl}(C_2)$$

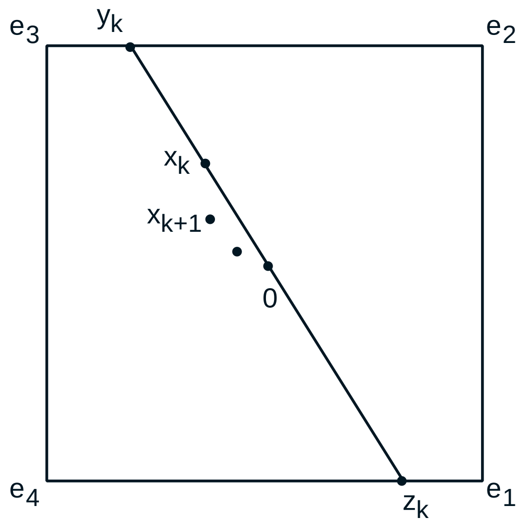
Proof of (a): $C_1 + C_2$ is the result of the linear transformation $(x_1, x_2) \mapsto x_1 + x_2$.

- Counterexample for (b):

$$C_1 = \{x \mid x \leq 0\}, \qquad C_2 = \{x \mid x \geq 0\}$$

CONTINUITY OF CONVEX FUNCTIONS

- If $f : \Re^n \mapsto \Re$ is convex, then it is continuous.



Proof: We will show that f is continuous at 0. By convexity, f is bounded within the unit cube by the maximum value of f over the corners of the cube.

Consider sequence $x_k \to 0$ and the sequences $y_k = x_k/\|x_k\|_\infty$, $z_k = -x_k/\|x_k\|_\infty$. Then

$$f(x_k) \le \big(1 - \|x_k\|_\infty\big) f(0) + \|x_k\|_\infty f(y_k)$$

$$f(0) \le \frac{\|x_k\|_\infty}{\|x_k\|_\infty + 1} f(z_k) + \frac{1}{\|x_k\|_\infty + 1} f(x_k)$$

Since $\|x_k\|_\infty \to 0$, $f(x_k) \to f(0)$. **Q.E.D.**

- Extension to continuity over ri(dom(f)).

CLOSURES OF FUNCTIONS

- The *closure* of a function $f : X \mapsto [-\infty, \infty]$ is the function $\mathrm{cl}\, f : \Re^n \mapsto [-\infty, \infty]$ with

$$\mathrm{epi}(\mathrm{cl}\, f) = \mathrm{cl}\big(\mathrm{epi}(f)\big)$$

- The *convex closure of* f is the function $\check{\mathrm{cl}}\, f$ with

$$\mathrm{epi}(\check{\mathrm{cl}}\, f) = \mathrm{cl}\Big(\mathrm{conv}\big(\mathrm{epi}(f)\big)\Big)$$

- *Proposition*: For any $f : X \mapsto [-\infty, \infty]$

$$\inf_{x \in X} f(x) = \inf_{x \in \Re^n} (\mathrm{cl}\, f)(x) = \inf_{x \in \Re^n} (\check{\mathrm{cl}}\, f)(x).$$

Also, any vector that attains the infimum of f over X also attains the infimum of $\mathrm{cl}\, f$ and $\check{\mathrm{cl}}\, f$.

- *Proposition*: For any $f : X \mapsto [-\infty, \infty]$:

  (a) $\mathrm{cl}\, f$ ($\check{\mathrm{cl}}\, f$) is the greatest closed (closed convex, resp.) function majorized by f.

  (b) If f is convex, then $\mathrm{cl}\, f$ is convex, and it is proper if and only if f is proper. Also,

$$(\mathrm{cl}\, f)(x) = f(x), \qquad \forall\, x \in \mathrm{ri}\big(\mathrm{dom}(f)\big),$$

  and if $x \in \mathrm{ri}\big(\mathrm{dom}(f)\big)$ and $y \in \mathrm{dom}(\mathrm{cl}\, f)$,

$$(\mathrm{cl}\, f)(y) = \lim_{\alpha \downarrow 0} f\big(y + \alpha(x - y)\big).$$

LECTURE 3

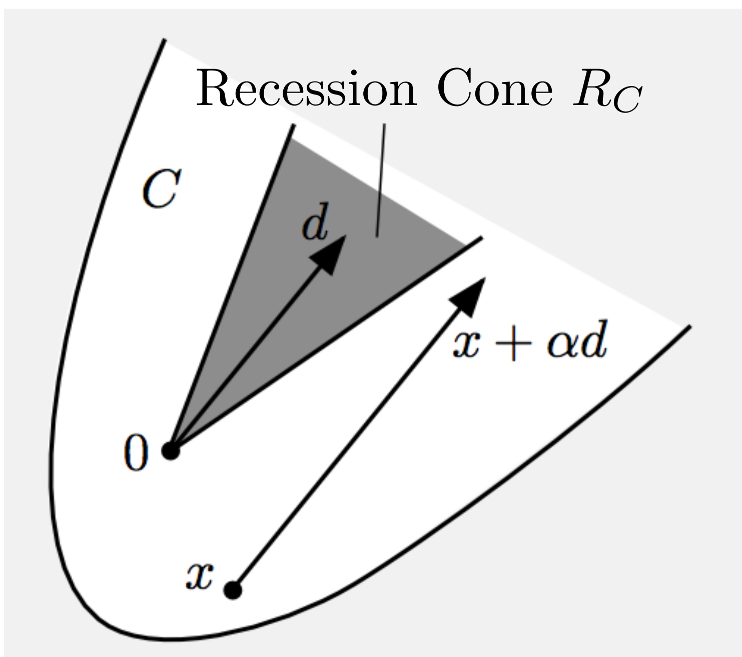
LECTURE OUTLINE

- Recession cones
- Directions of recession of convex functions
- Nonemptiness of closed set intersections
- Linear and Quadratic Programming
- Preservation of closure under linear transformation

RECESSION CONE OF A CONVEX SET

- Given a nonempty convex set C, a vector d is a *direction of recession* if starting at **any** x in C and going indefinitely along d, we never cross the relative boundary of C to points outside C:

$$x + \alpha d \in C, \qquad \forall\, x \in C, \;\; \forall\, \alpha \geq 0$$



- *Recession cone* of C (denoted by R_C): The set of all directions of recession.

- R_C is a cone containing the origin.

RECESSION CONE THEOREM

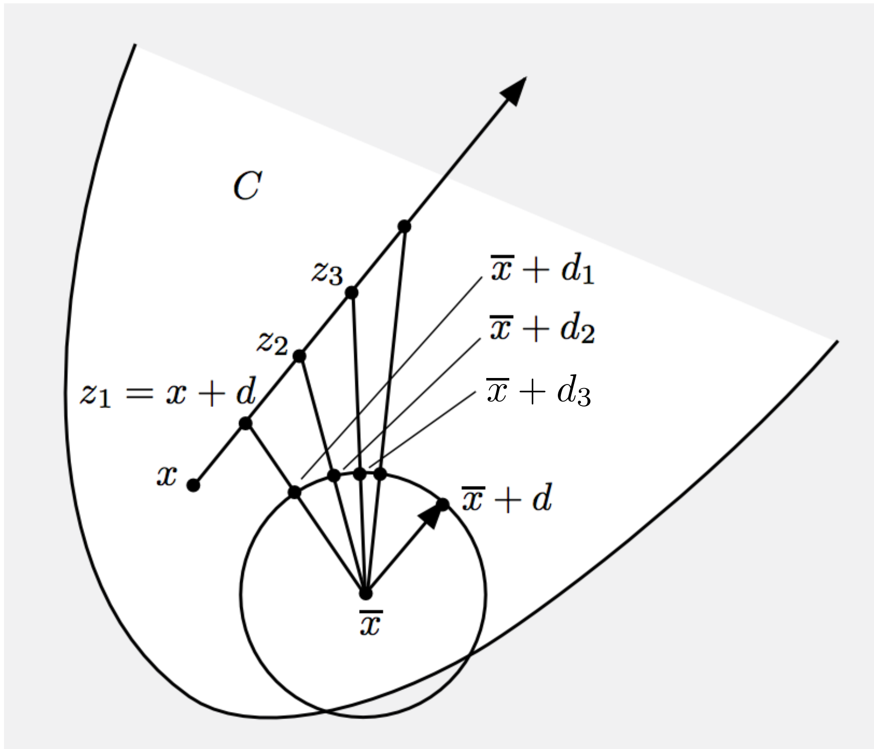
- Let C be a nonempty closed convex set.
  (a) The recession cone R_C is a closed convex cone.
  (b) A vector d belongs to R_C if and only if there exists a vector $x \in C$ such that $x + \alpha d \in C$ for all $\alpha \geq 0$.
  (c) R_C contains a nonzero direction if and only if C is unbounded.
  (d) The recession cones of C and $\mathrm{ri}(C)$ are equal.
  (e) If D is another closed convex set such that $C \cap D \neq \emptyset$, we have

$$R_{C \cap D} = R_C \cap R_D$$

More generally, for any collection of closed convex sets C_i, $i \in I$, where I is an arbitrary index set and $\cap_{i \in I} C_i$ is nonempty, we have

$$R_{\cap_{i \in I} C_i} = \cap_{i \in I} R_{C_i}$$

PROOF OF PART (B)



- Let $d \neq 0$ be such that there exists a vector $x \in C$ with $x + \alpha d \in C$ for all $\alpha \geq 0$. We fix $\overline{x} \in C$ and $\alpha > 0$, and we show that $\overline{x} + \alpha d \in C$. By scaling d, it is enough to show that $\overline{x} + d \in C$.

  Let $z_k = x + kd$ for $k = 1, 2, \ldots$, and $d_k = (z_k - \overline{x})\|d\|/\|z_k - \overline{x}\|$. We have

$$\frac{d_k}{\|d\|} = \frac{\|z_k - x\|}{\|z_k - \overline{x}\|} \frac{d}{\|d\|} + \frac{x - \overline{x}}{\|z_k - \overline{x}\|}, \quad \frac{\|z_k - x\|}{\|z_k - \overline{x}\|} \to 1, \quad \frac{x - \overline{x}}{\|z_k - \overline{x}\|} \to 0,$$

so $d_k \to d$ and $\overline{x} + d_k \to \overline{x} + d$. Use the convexity and closedness of C to conclude that $\overline{x} + d \in C$.

LINEALITY SPACE

- The *lineality space* of a convex set C, denoted by L_C, is the subspace of vectors d such that $d \in R_C$ and $-d \in R_C$:

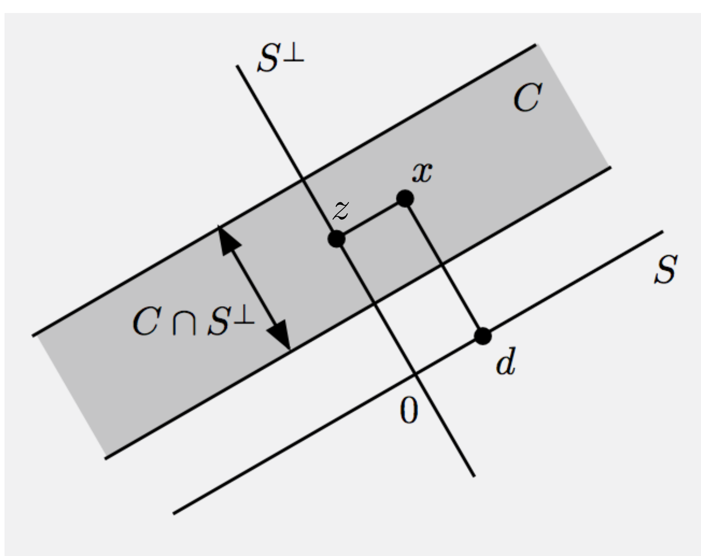
$$L_C = R_C \cap (-R_C)$$

- If $d \in L_C$, the entire line defined by d is contained in C, starting at any point of C.

- *Decomposition of a Convex Set*: Let C be a nonempty convex subset of $\Re^n$. Then,

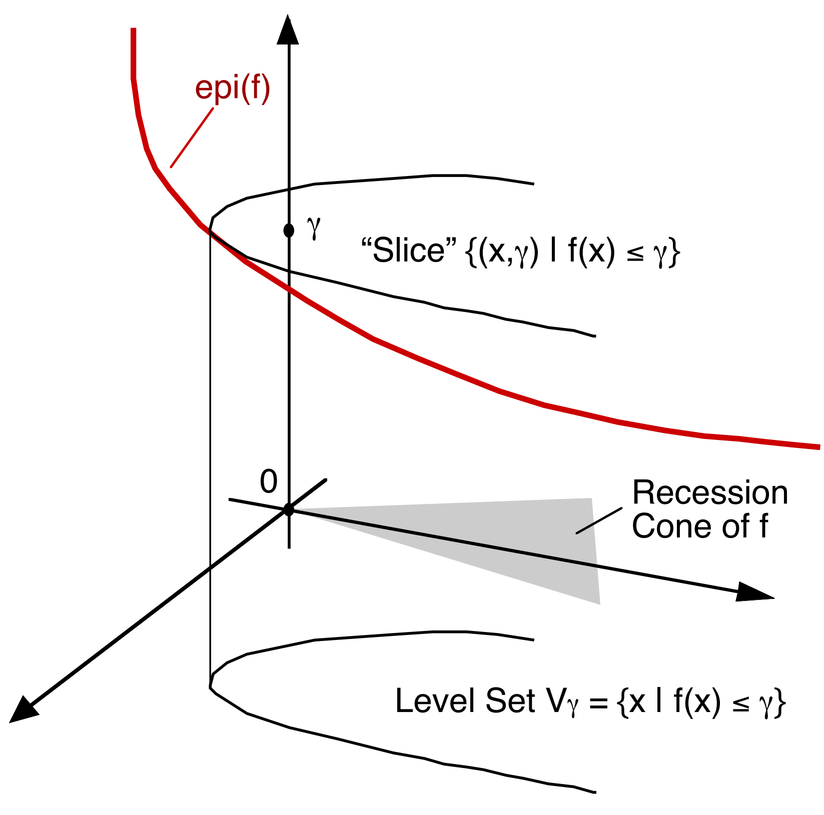
$$C = L_C + (C \cap L_C^\perp).$$

- True also if L_C is replaced by a subset $S \subset L_C$.



DIRECTIONS OF RECESSION OF A FUNCTION

- Some basic geometric observations:
  - The "horizontal directions" in the recession cone of the epigraph of a convex function f are directions along which the level sets are unbounded.
  - Along these directions the level sets $\{x \mid f(x) \leq \gamma\}$ are unbounded and f is monotonically nondecreasing.
- These are the *directions of recession* of f.



RECESSION CONE OF LEVEL SETS

- *Proposition*: Let $f : \Re^n \mapsto (-\infty, \infty]$ be a closed proper convex function and consider the level sets $V_\gamma = \{x \mid f(x) \leq \gamma\}$, where γ is a scalar. Then:

  (a) All the nonempty level sets V_γ have the same recession cone, given by

$$R_{V_\gamma} = \{d \mid (d, 0) \in R_{\text{epi}(f)}\}$$

  (b) If one nonempty level set V_γ is compact, then all nonempty level sets are compact.

Proof: For each fixed γ for which V_γ is nonempty,

$$\{(x, \gamma) \mid x \in V_\gamma\} = \text{epi}(f) \cap \{(x, \gamma) \mid x \in \Re^n\}$$

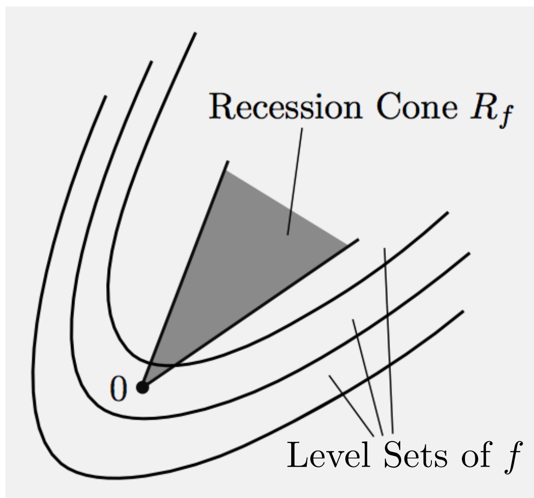
The recession cone of the set on the left is $\{(d, 0) \mid d \in R_{V_\gamma}\}$. The recession cone of the set on the right is the intersection of $R_{\text{epi}}(f)$ and the recession cone of $\{(x, \gamma) \mid x \in \Re^n\}$. Thus we have

$$\{(d, 0) \mid d \in R_{V_\gamma}\} = \{(d, 0) \mid (d, 0) \in R_{\text{epi}(f)}\},$$

from which the result follows.

RECESSION CONE OF A CONVEX FUNCTION

- For a closed proper convex function $f : \Re^n \mapsto (-\infty, \infty]$, the (common) recession cone of the nonempty level sets $V_\gamma = \{x \mid f(x) \leq \gamma\}$, $\gamma \in \Re$, is the *recession cone of* f, and is denoted by R_f.



- Terminology:
  - $d \in R_f$: a *direction of recession of* f.
  - $L_f = R_f \cap (-R_f)$: the *lineality space of* f.
  - $d \in L_f$: a *direction of constancy* of f.
- **Example:** For the pos. semidefinite quadratic

$$f(x) = x'Qx + a'x + b,$$

the recession cone and constancy space are

$$R_f = \{d \mid Qd = 0,\ a'd \leq 0\}, \quad L_f = \{d \mid Qd = 0,\ a'd = 0\}$$

RECESSION FUNCTION

- Function $r_f : \Re^n \mapsto (-\infty, \infty]$ whose epigraph is $R_{\mathrm{epi}(f)}$: the *recession function* of f.
- Characterizes the recession cone:

$$R_f = \big\{d \mid r_f(d) \le 0\big\}, \quad L_f = \big\{d \mid r_f(d) = r_f(-d) = 0\big\}$$

- Can be shown that

$$r_f(d) = \sup_{\alpha > 0} \frac{f(x + \alpha d) - f(x)}{\alpha} = \lim_{\alpha \to \infty} \frac{f(x + \alpha d) - f(x)}{\alpha}$$

- Thus $r_f(d)$ is the "asymptotic slope" of f in the direction d. In fact,

$$r_f(d) = \lim_{\alpha \to \infty} \nabla f(x + \alpha d)' d, \qquad \forall\, x, d \in \Re^n$$

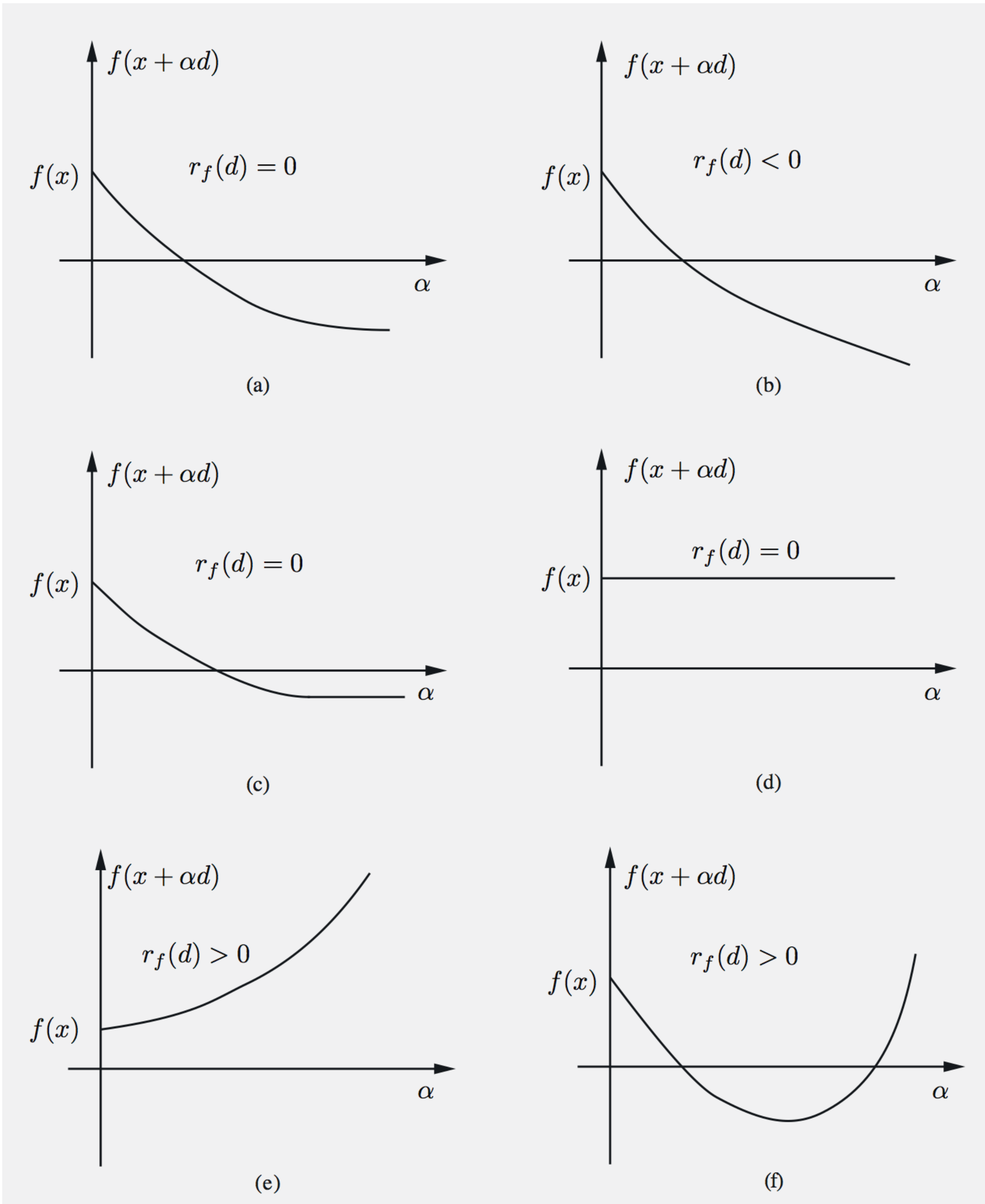
if f is differentiable.

- Calculus of recession functions:

$$r_{f_1 + \cdots + f_m}(d) = r_{f_1}(d) + \cdots + r_{f_m}(d)$$

$$r_{\sup_{i \in I} f_i}(d) = \sup_{i \in I} r_{f_i}(d)$$

DESCENT BEHAVIOR OF A CONVEX FUNCTION



- y is a direction of recession in (a)-(d).
- This behavior is *independent of the starting point* x, as long as $x \in \mathrm{dom}(f)$.

THE ROLE OF CLOSED SET INTERSECTIONS

- **A fundamental question:** Given a sequence of nonempty closed sets $\{C_k\}$ in $\Re^n$ with $C_{k+1} \subset C_k$ for all k, when is $\cap_{k=0}^{\infty} C_k$ nonempty?

- Set intersection theorems are significant in at least three major contexts, which we will discuss in what follows:

1. Does a function $f : \Re^n \mapsto (-\infty, \infty]$ attain a minimum over a set X? This is true iff the intersection of the nonempty level sets $\big\{x \in X \mid f(x) \leq \gamma_k\big\}$ is nonempty.

2. If C is closed and A is a matrix, is $A\,C$ closed? Special case:

– If C_1 and C_2 are closed, is $C_1 + C_2$ closed?

3. If $F(x, z)$ is closed, is $f(x) = \inf_z F(x, z)$ closed? (Critical question in duality theory.) Can be addressed by using the relation

$$P\big(\text{epi}(F)\big) \subset \text{epi}(f) \subset \text{cl}\Big(P\big(\text{epi}(F)\big)\Big)$$

where $P(\cdot)$ is projection on the space of (x, w).

ASYMPTOTIC DIRECTIONS

- Given nested sequence $\{C_k\}$ of closed convex sets, $\{x_k\}$ is an *asymptotic sequence* if

$$x_k \in C_k, \qquad x_k \neq 0, \qquad k = 0, 1, \ldots$$

$$\|x_k\| \to \infty, \qquad \frac{x_k}{\|x_k\|} \to \frac{d}{\|d\|}$$

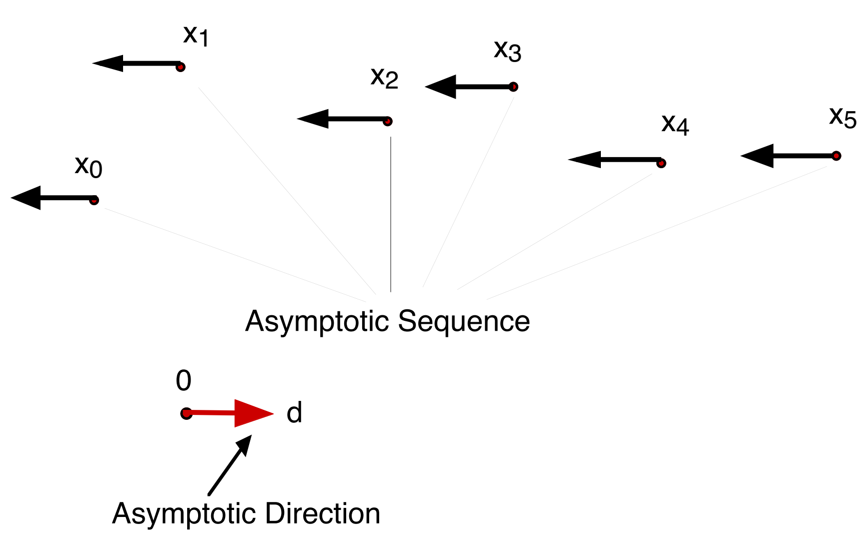
where d is a nonzero common direction of recession of the sets C_k.

- As a special case we define asymptotic sequence of a closed convex set C (use $C_k \equiv C$).

- Every unbounded $\{x_k\}$ with $x_k \in C_k$ has an asymptotic subsequence.

- $\{x_k\}$ is called *retractive* if for some $\overline{k}$, we have

$$x_k - d \in C_k, \quad \forall\, k \geq \overline{k}.$$



RETRACTIVE SEQUENCES

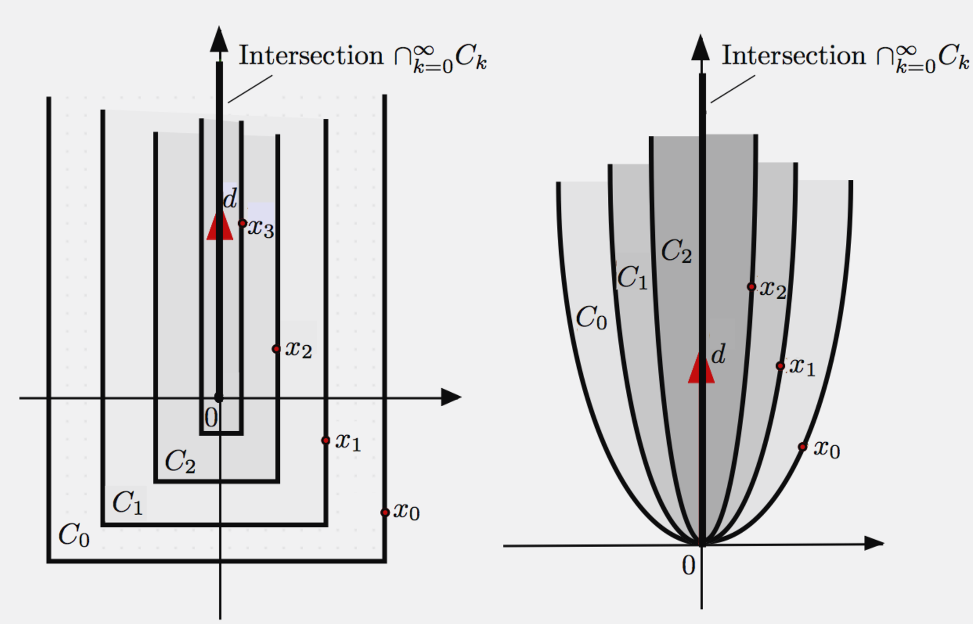
- A nested sequence $\{C_k\}$ of closed convex sets is *retractive* if all its asymptotic sequences are retractive.

- Intersections and Cartesian products of retractive set sequences are retractive.

- A closed halfspace (viewed as a sequence with identical components) is retractive.

- A polyhedral set is retractive. Also the vector sum of a convex compact set and a retractive convex set is retractive.

- Nonpolyhedral cones and level sets of quadratic functions need not be retractive.



(a) Retractive Set Sequence (b) Nonretractive Set Sequence

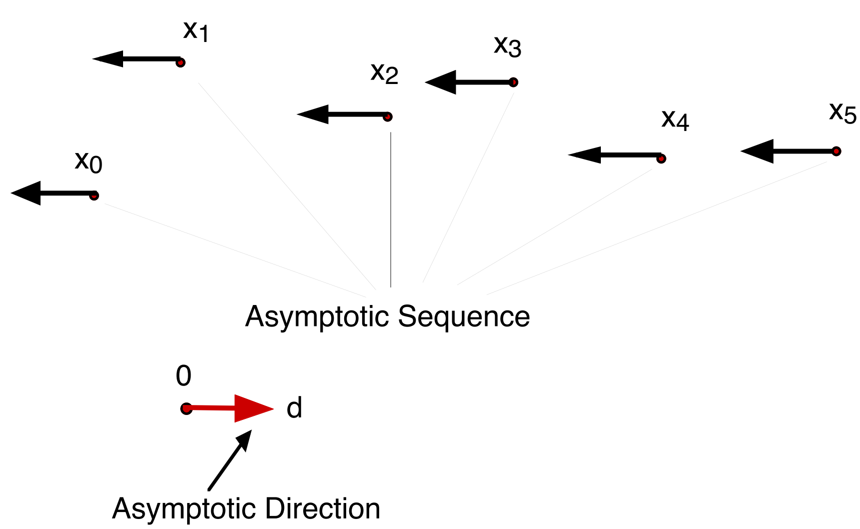
SET INTERSECTION THEOREM I

Proposition: If $\{C_k\}$ is retractive, then $\cap_{k=0}^{\infty} C_k$ is nonempty.

- Key proof ideas:

  (a) The intersection $\cap_{k=0}^{\infty} C_k$ is empty iff the sequence $\{x_k\}$ of minimum norm vectors of C_k is unbounded (so a subsequence is asymptotic).

  (b) An asymptotic sequence $\{x_k\}$ of minimum norm vectors cannot be retractive, because such a sequence eventually gets closer to 0 when shifted opposite to the asymptotic direction.



SET INTERSECTION THEOREM II

Proposition: Let $\{C_k\}$ be a nested sequence of nonempty closed convex sets, and X be a retractive set such that all the sets $\overline{C}_k = X \cap C_k$ are nonempty. Assume that

$$R_X \cap R \subset L,$$

where

$$R = \cap_{k=0}^{\infty} R_{C_k}, \qquad L = \cap_{k=0}^{\infty} L_{C_k}$$

Then $\{\overline{C}_k\}$ is retractive and $\cap_{k=0}^{\infty} \overline{C}_k$ is nonempty.

- Special case: $X = \Re^n$, $R = L$.

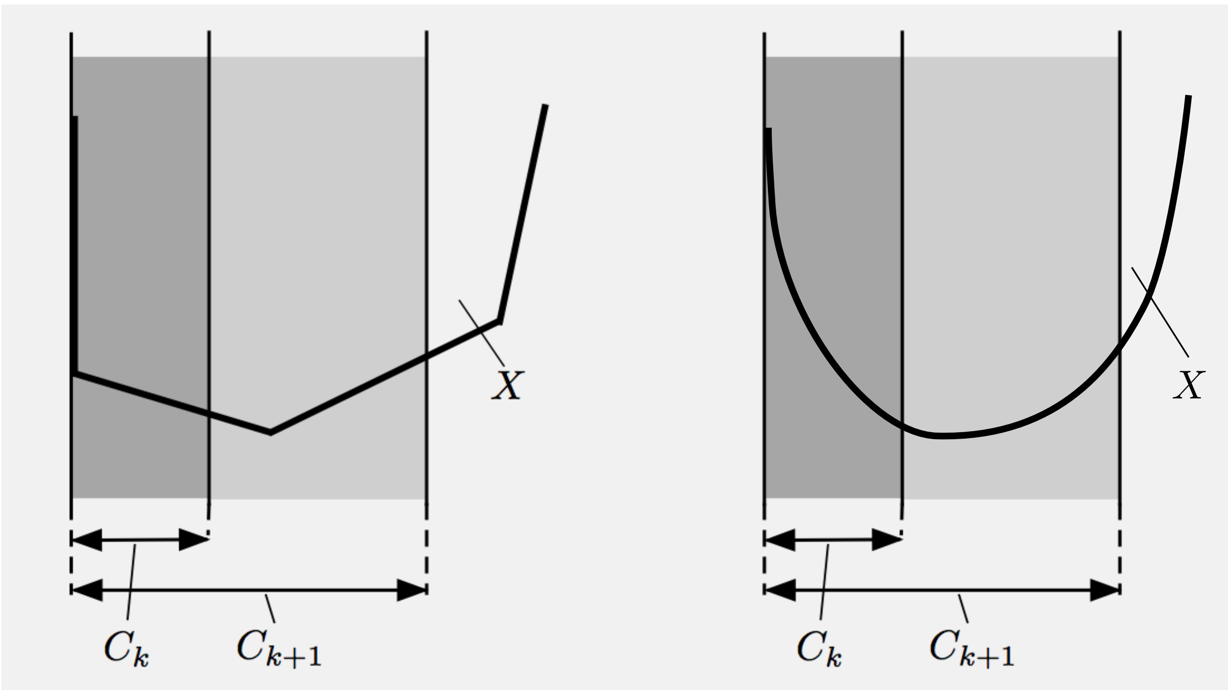
Proof: The set of common directions of recession of C_k is $R_X \cap R$. For any asymptotic sequence $\{x_k\}$ corresponding to $d \in R_X \cap R$:

(1) $x_k - d \in C_k$ (because $d \in L$)

(2) $x_k - d \in X$ (because X is retractive)

So $\{\overline{C}_k\}$ is retractive.

NEED TO ASSUME THAT X IS RETRACTIVE



Consider $\cap_{k=0}^{\infty} \overline{C}_k$, with $\overline{C}_k = X \cap C_k$

- The condition $R_X \cap R \subset L$ holds
- In the figure on the left, X is polyhedral.
- In the figure on the right, X is nonpolyhedral and nonretrative, and

$$\cap_{k=0}^{\infty} \overline{C}_k = \emptyset$$

LINEAR AND QUADRATIC PROGRAMMING

- **Theorem:** Let

$$f(x) = x'Qx + c'x, \quad X = \{x \mid a_j'x + b_j \leq 0,\ j = 1, \ldots, r\},$$

where Q is symmetric positive semidefinite. If the minimal value of f over X is finite, there exists a minimum of f over X.

Proof: (Outline) Write

$$\text{Set of Minima} = X \cap \{x \mid x'Qx + c'x \leq \gamma_k\}$$

with

$$\gamma_k \downarrow f^* = \inf_{x \in X} f(x).$$

Verify the condition $R_X \cap R \subset L$ of the preceding set intersection theorem, where R and L are the sets of common recession and lineality directions of the sets

$$\{x \mid x'Qx + c'x \leq \gamma_k\}$$

Q.E.D.

CLOSURE UNDER LINEAR TRANSFORMATIONS

- Let C be a nonempty closed convex, and let A be a matrix with nullspace $N(A)$.

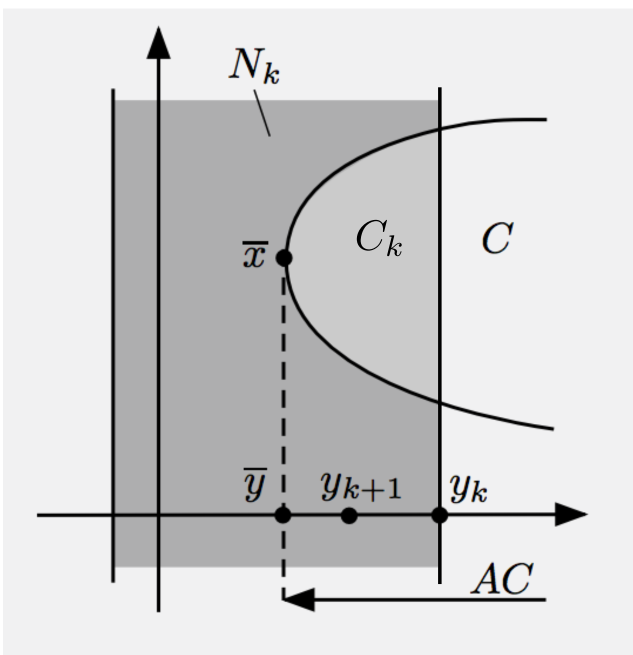
  (a) AC is closed if $R_C \cap N(A) \subset L_C$.

  (b) $A(X \cap C)$ is closed if X is a retractive set and

$$R_X \cap R_C \cap N(A) \subset L_C,$$

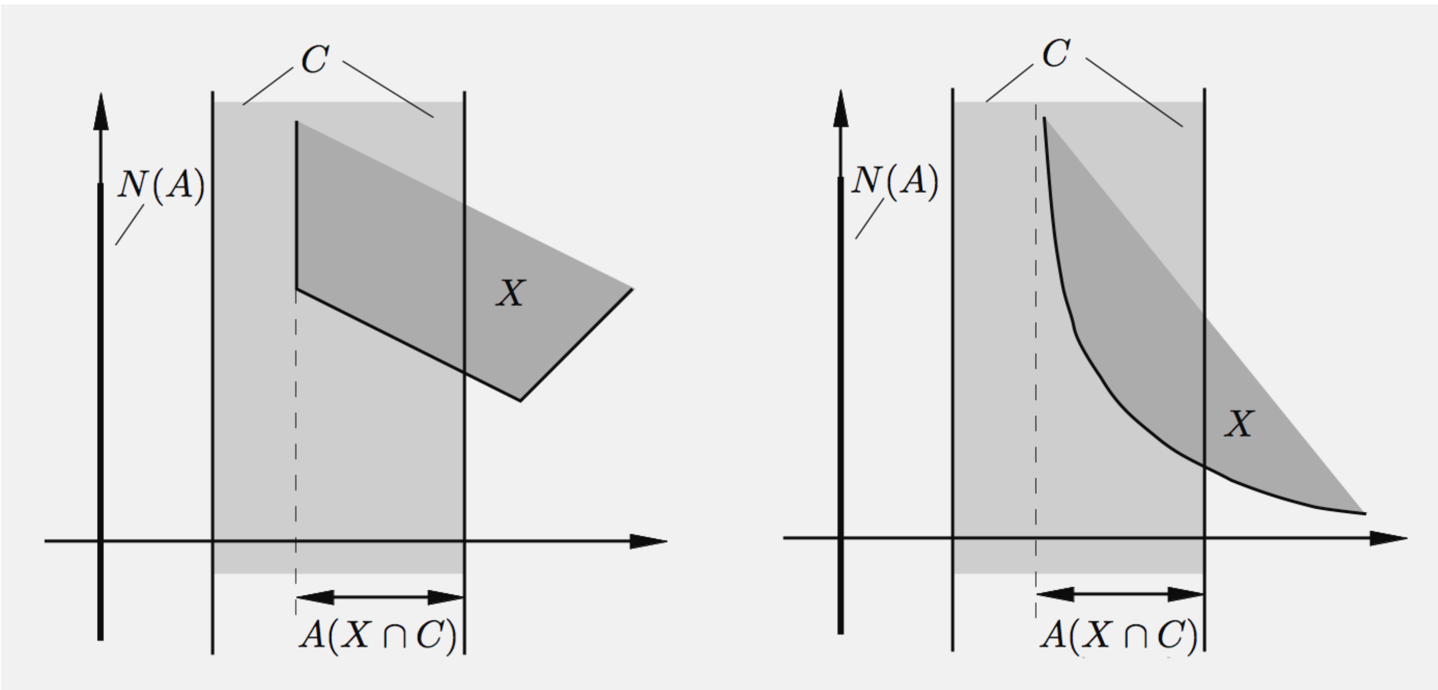
Proof: (Outline) Let $\{y_k\} \subset AC$ with $y_k \to \overline{y}$. We prove $\cap_{k=0}^{\infty} C_k \neq \emptyset$, where $C_k = C \cap N_k$, and

$$N_k = \{x \mid Ax \in W_k\}, \quad W_k = \{z \mid \|z - \overline{y}\| \leq \|y_k - \overline{y}\|\}$$



- **Special Case:** AX is closed if X is polyhedral.

NEED TO ASSUME THAT X IS RETRACTIVE



Consider closedness of $\quad A(X \cap C)$

- In both examples the condition

$$R_X \cap R_C \cap N(A) \subset L_C$$

is satisfied.

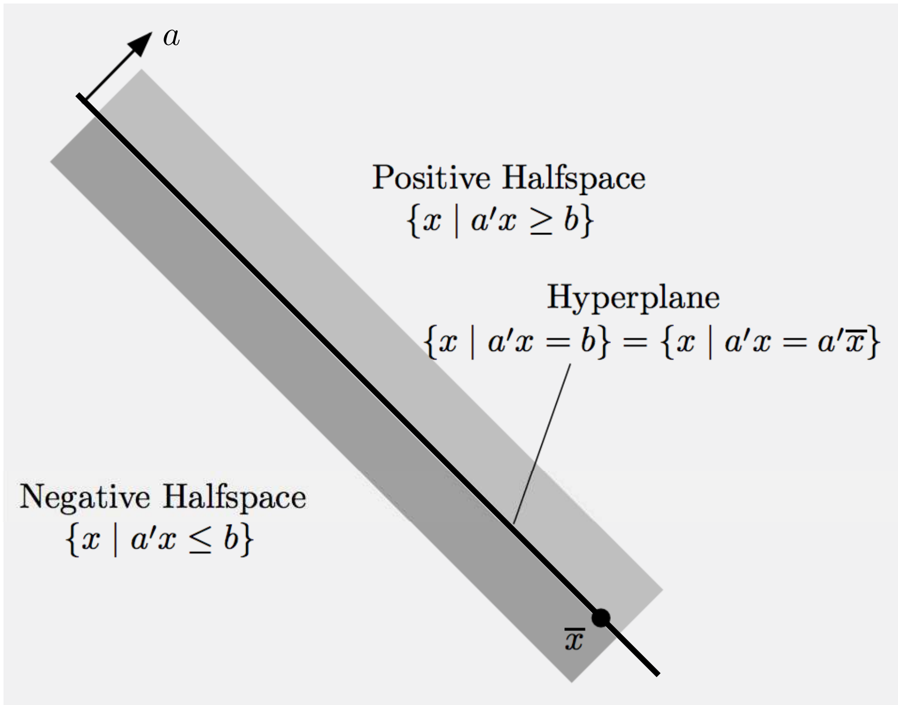
- However, in the example on the right, X is not retractive, and the set $A(X \cap C)$ is not closed.

LECTURE 4

LECTURE OUTLINE

- Hyperplane separation
- Proper separation
- Nonvertical hyperplanes
- Convex conjugate functions
- Conjugacy theorem
- Examples

HYPERPLANES



- A *hyperplane* is a set of the form $\{x \mid a'x = b\}$, where a is nonzero vector in $\Re^n$ and b is a scalar.

- We say that two sets C_1 and C_2 are *separated by a hyperplane* $H = \{x \mid a'x = b\}$ if each lies in a different closed halfspace associated with H, i.e.,

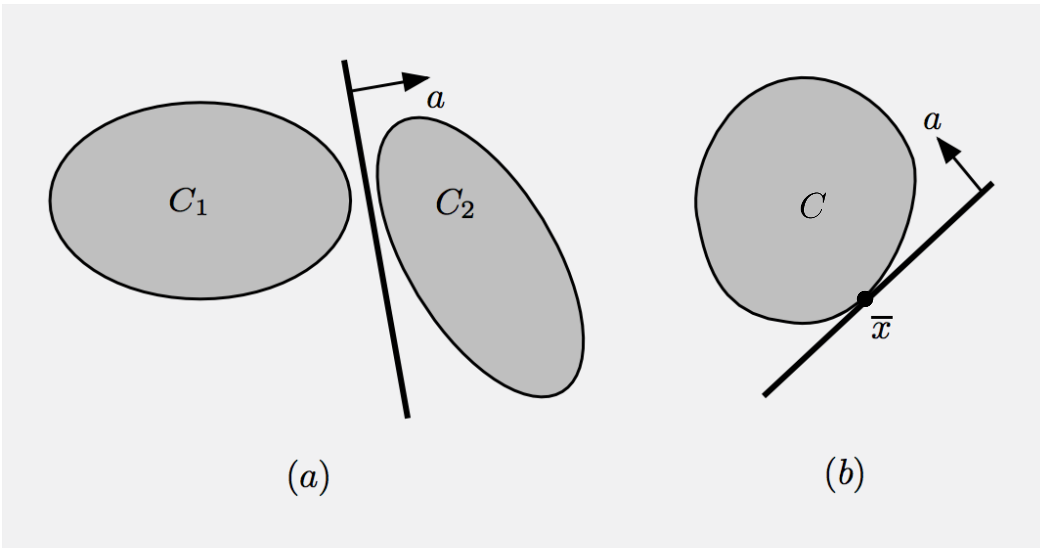
$$\text{either} \quad a'x_1 \leq b \leq a'x_2, \qquad \forall\, x_1 \in C_1,\ \forall\, x_2 \in C_2,$$

$$\text{or} \quad a'x_2 \leq b \leq a'x_1, \qquad \forall\, x_1 \in C_1,\ \forall\, x_2 \in C_2$$

- If $\overline{x}$ belongs to the closure of a set C, a hyperplane that separates C and the singleton set $\{\overline{x}\}$ is said be *supporting* C *at* $\overline{x}$.

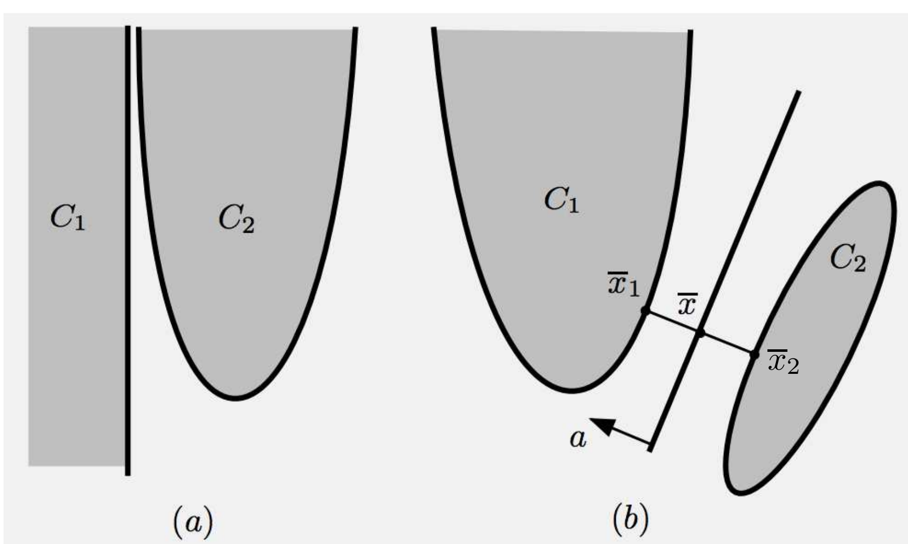
VISUALIZATION

- Separating and supporting hyperplanes:



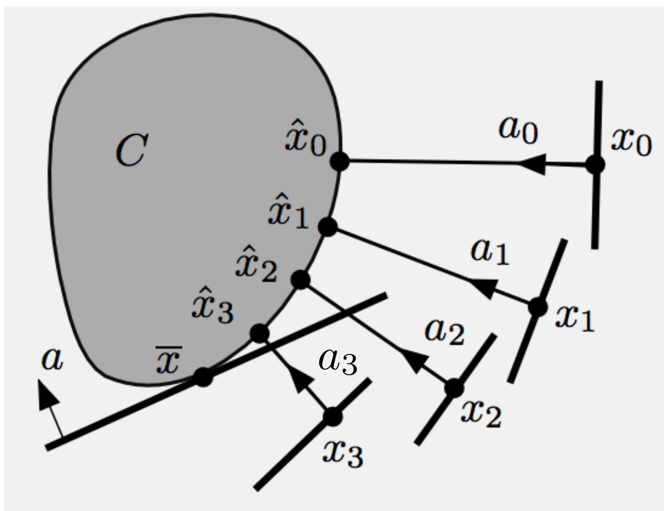
- A separating $\{x \mid a'x = b\}$ that is disjoint from C_1 and C_2 is called *strictly* separating:

$$a'x_1 < b < a'x_2, \qquad \forall\ x_1 \in C_1,\ \forall\ x_2 \in C_2$$



SUPPORTING HYPERPLANE THEOREM

- Let C be convex and let $\overline{x}$ be a vector that is not an interior point of C. Then, there exists a hyperplane that passes through $\overline{x}$ and contains C in one of its closed halfspaces.



Proof: Take a sequence $\{x_k\}$ that does not belong to $\mathrm{cl}(C)$ and converges to $\overline{x}$. Let $\hat{x}_k$ be the projection of x_k on $\mathrm{cl}(C)$. We have for all $x \in \mathrm{cl}(C)$

$$a_k' x \geq a_k' x_k, \qquad \forall\, x \in \mathrm{cl}(C),\ \forall\, k = 0, 1, \ldots,$$

where $a_k = (\hat{x}_k - x_k)/\|\hat{x}_k - x_k\|$. Let a be a limit point of $\{a_k\}$, and take limit as $k \to \infty$. **Q.E.D.**

SEPARATING HYPERPLANE THEOREM

- Let C_1 and C_2 be two nonempty convex subsets of $\Re^n$. If C_1 and C_2 are disjoint, there exists a hyperplane that separates them, i.e., there exists a vector $a \neq 0$ such that

$$a'x_1 \leq a'x_2, \qquad \forall\, x_1 \in C_1,\ \forall\, x_2 \in C_2.$$

Proof: Consider the convex set

$$C_1 - C_2 = \{x_2 - x_1 \mid x_1 \in C_1,\, x_2 \in C_2\}$$

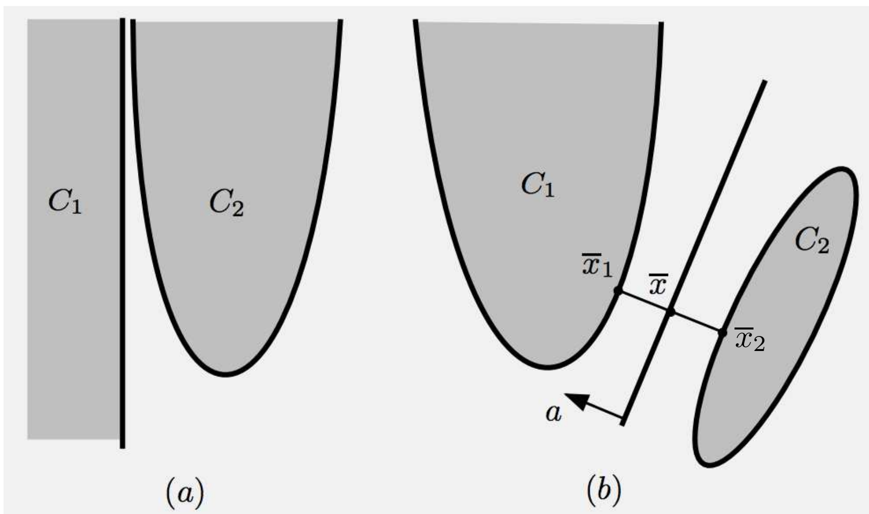
Since C_1 and C_2 are disjoint, the origin does not belong to $C_1 - C_2$, so by the Supporting Hyperplane Theorem, there exists a vector $a \neq 0$ such that

$$0 \leq a'x, \qquad \forall\, x \in C_1 - C_2,$$

which is equivalent to the desired relation. **Q.E.D.**

STRICT SEPARATION THEOREM

- **Strict Separation Theorem**: Let C_1 and C_2 be two disjoint nonempty convex sets. If C_1 is closed, and C_2 is compact, there exists a hyperplane that strictly separates them.



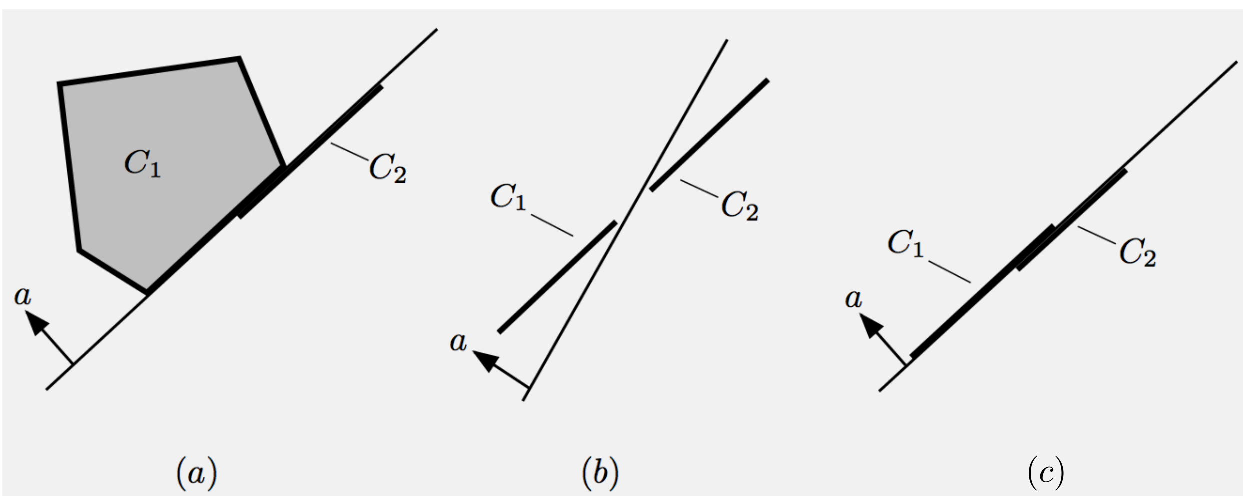
Proof: (Outline) Consider the set $C_1 - C_2$. Since C_1 is closed and C_2 is compact, $C_1 - C_2$ is closed. Since $C_1 \cap C_2 = \emptyset$, $0 \notin C_1 - C_2$. Let $\overline{x}_1 - \overline{x}_2$ be the projection of 0 onto $C_1 - C_2$. The strictly separating hyperplane is constructed as in (b).

- **Note**: Any conditions that guarantee closedness of $C_1 - C_2$ guarantee existence of a strictly separating hyperplane. However, there may exist a strictly separating hyperplane without $C_1 - C_2$ being closed.

ADDITIONAL THEOREMS

- **Fundamental Characterization**: The closure of the convex hull of a set $C \subset \Re^n$ is the intersection of the closed halfspaces that contain C. (Proof uses the strict separation theorem.)

- We say that a hyperplane *properly separates C_1 and C_2* if it separates C_1 and C_2 and does not fully contain both C_1 and C_2.



- **Proper Separation Theorem**: Let C_1 and C_2 be two nonempty convex subsets of $\Re^n$. There exists a hyperplane that properly separates C_1 and C_2 if and only if

$$\mathrm{ri}(C_1) \cap \mathrm{ri}(C_2) = \emptyset$$

PROPER POLYHEDRAL SEPARATION

- Recall that two convex sets C and P such that

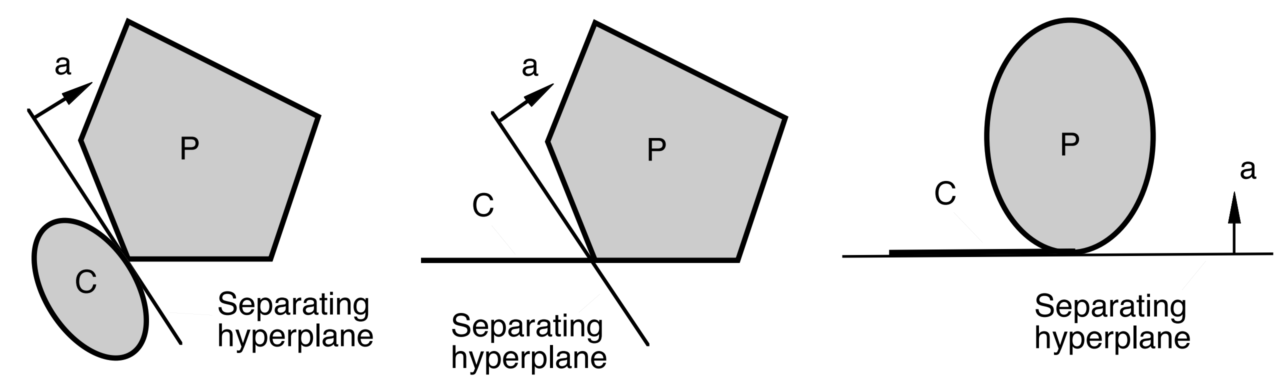
$$\mathrm{ri}(C) \cap \mathrm{ri}(P) = \emptyset$$

can be properly separated, i.e., by a hyperplane that does not contain both C and P.

- If P is polyhedral and the slightly stronger condition

$$\mathrm{ri}(C) \cap P = \emptyset$$

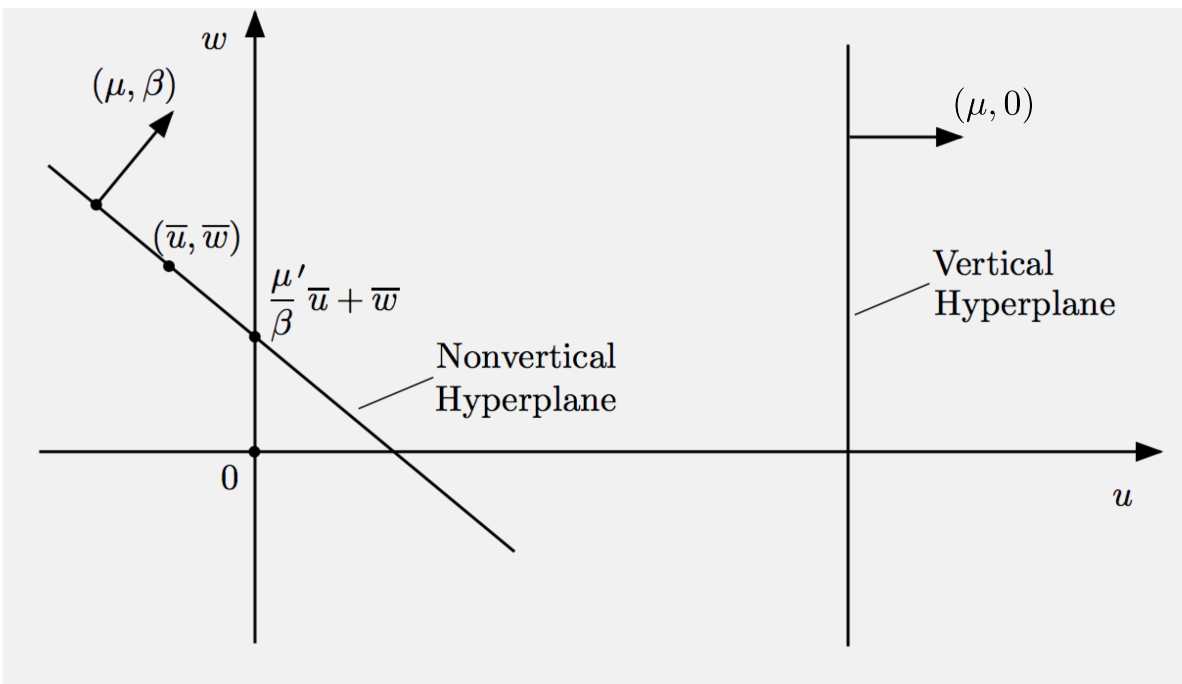
holds, then the properly separating hyperplane can be chosen so that it does not contain the non-polyhedral set C while it may contain P.



On the left, the separating hyperplane can be chosen so that it does not contain C. On the right where P is not polyhedral, this is not possible.

NONVERTICAL HYPERPLANES

- A hyperplane in $\Re^{n+1}$ with normal (μ, β) is nonvertical if $\beta \neq 0$.
- It intersects the $(n+1)$st axis at $\xi = (\mu/\beta)'\overline{u} + \overline{w}$, where $(\overline{u}, \overline{w})$ is any vector on the hyperplane.



- A nonvertical hyperplane that contains the epigraph of a function in its "upper" halfspace, provides lower bounds to the function values.
- The epigraph of a proper convex function does not contain a vertical line, so it appears plausible that it is contained in the "upper" halfspace of some nonvertical hyperplane.

NONVERTICAL HYPERPLANE THEOREM

- Let C be a nonempty convex subset of $\Re^{n+1}$ that contains no vertical lines. Then:

  (a) C is contained in a closed halfspace of a nonvertical hyperplane, i.e., there exist $\mu \in \Re^n$, $\beta \in \Re$ with $\beta \neq 0$, and $\gamma \in \Re$ such that $\mu' u + \beta w \geq \gamma$ for all $(u, w) \in C$.

  (b) If $(\overline{u}, \overline{w}) \notin \text{cl}(C)$, there exists a nonvertical hyperplane strictly separating $(\overline{u}, \overline{w})$ and C.

Proof: Note that cl(C) contains no vert. line [since C contains no vert. line, ri(C) contains no vert. line, and ri(C) and cl(C) have the same recession cone]. So we just consider the case: C closed.

(a) C is the intersection of the closed halfspaces containing C. If all these corresponded to vertical hyperplanes, C would contain a vertical line.

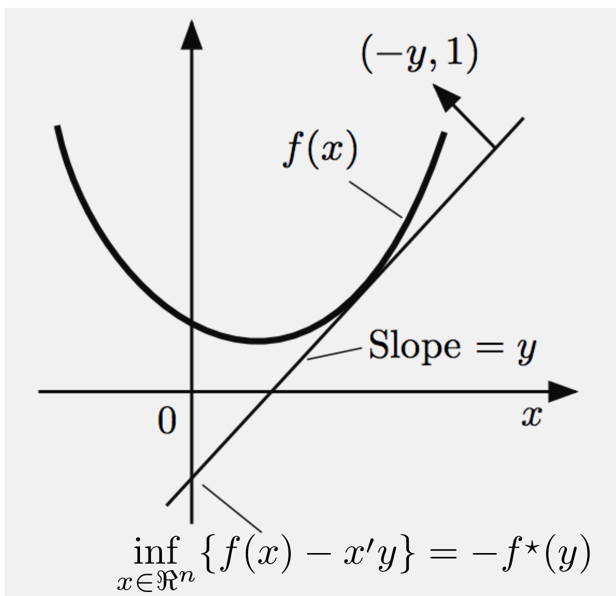
(b) There is a hyperplane strictly separating $(\overline{u}, \overline{w})$ and C. If it is nonvertical, we are done, so assume it is vertical. "Add" to this vertical hyperplane a small ϵ-multiple of a nonvertical hyperplane containing C in one of its halfspaces as per (a).

CONJUGATE CONVEX FUNCTIONS

- Consider a function f and its epigraph

Nonvertical hyperplanes supporting epi(f)
$\mapsto$ Crossing points of vertical axis

$$f^\star(y) = \sup_{x\in\Re^n}\big\{x'y - f(x)\big\}, \qquad y \in \Re^n.$$

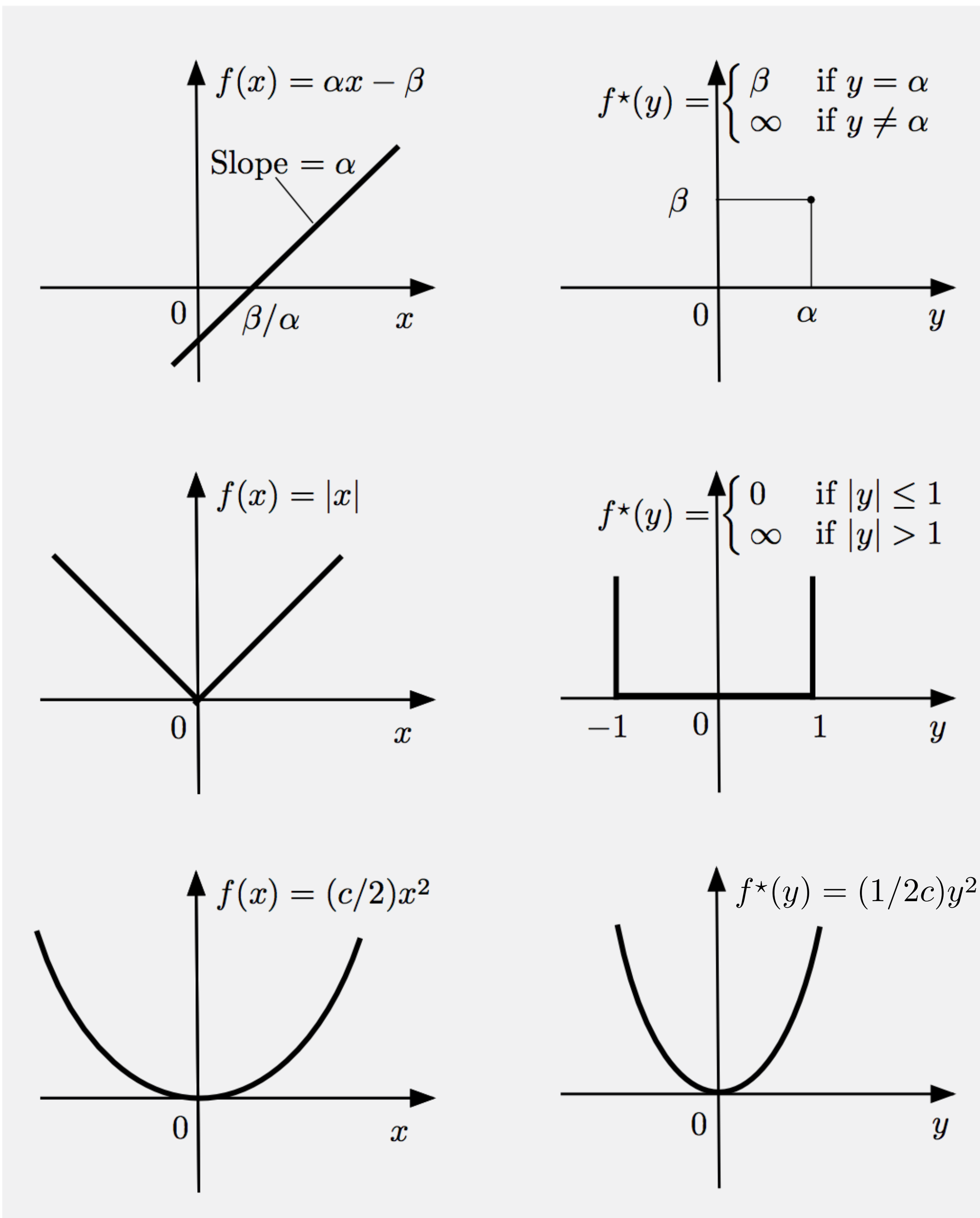


- For any $f : \Re^n \mapsto [-\infty,\infty]$, its *conjugate convex function* is defined by

$$f^\star(y) = \sup_{x\in\Re^n}\big\{x'y - f(x)\big\}, \qquad y \in \Re^n$$

EXAMPLES

$$f^\star(y) = \sup_{x\in\Re^n} \big\{x'y - f(x)\big\}, \qquad y \in \Re^n$$



CONJUGATE OF CONJUGATE

- From the definition

$$f^\star(y) = \sup_{x \in \Re^n} \big\{x'y - f(x)\big\}, \qquad y \in \Re^n,$$

note that h *is convex and closed*.

- Reason: $\text{epi}(f^\star)$ is the intersection of the epigraphs of the linear functions of y

$$x'y - f(x)$$

as x ranges over $\Re^n$.

- Consider the conjugate of the conjugate:

$$f^{\star\star}(x) = \sup_{y \in \Re^n} \big\{y'x - f^\star(y)\big\}, \qquad x \in \Re^n.$$

- $f^{\star\star}$ is convex and closed.

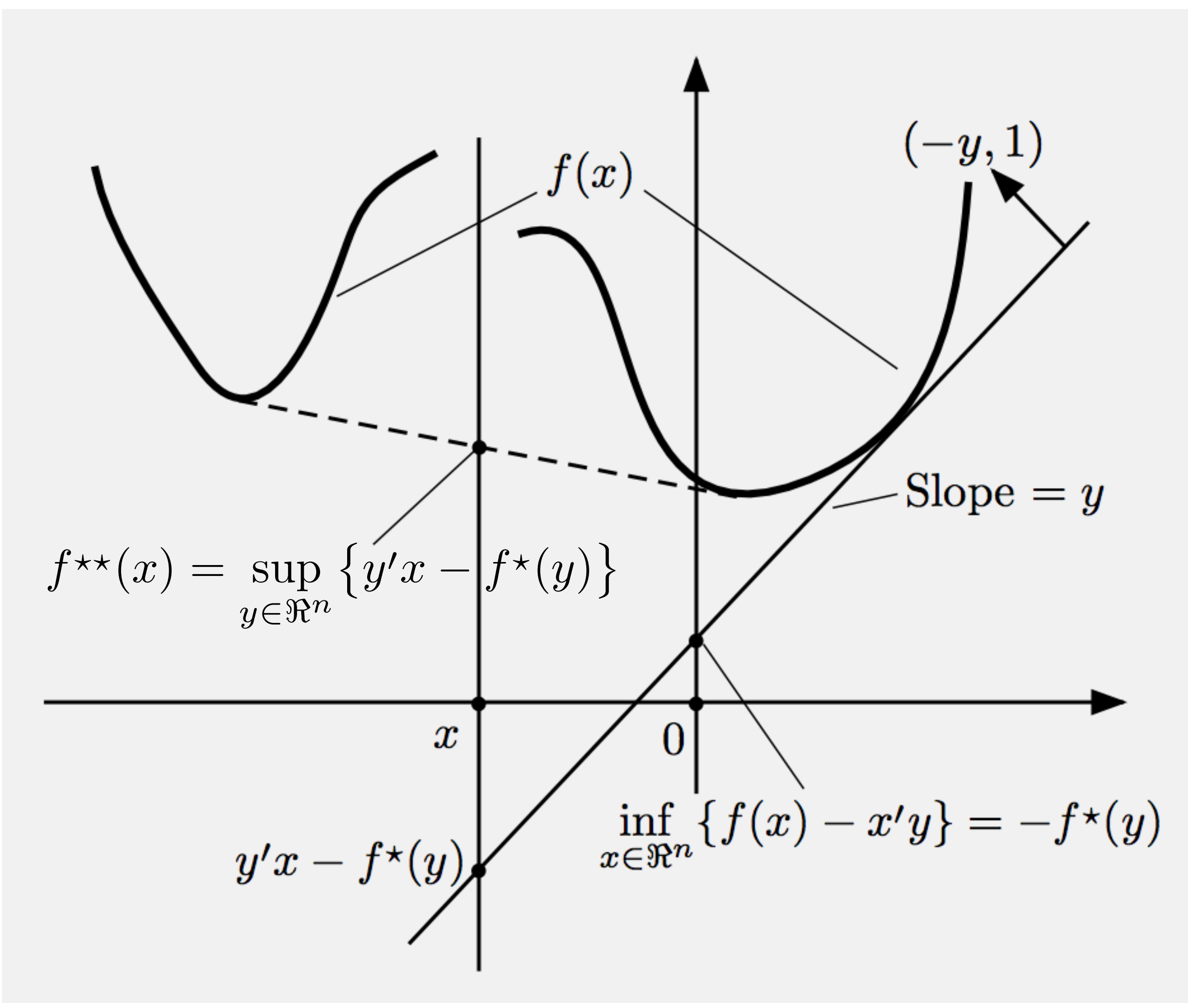
- **Important fact/Conjugacy theorem:** If f is closed proper convex, then $f^{\star\star} = f$.

CONJUGACY THEOREM - VISUALIZATION

$$f^\star(y) = \sup_{x \in \Re^n} \big\{ x'y - f(x) \big\}, \qquad y \in \Re^n$$

$$f^{\star\star}(x) = \sup_{y \in \Re^n} \big\{ y'x - f^\star(y) \big\}, \qquad x \in \Re^n$$

- If f is closed convex proper, then $f^{\star\star} = f$.



CONJUGACY THEOREM

- Let $f : \Re^n \mapsto (-\infty, \infty]$ be a function, let $\check{\mathrm{cl}}\, f$ be its convex closure, let $f^\star$ be its convex conjugate, and consider the conjugate of $f^\star$,

$$f^{\star\star}(x) = \sup_{y \in \Re^n} \big\{ y'x - f^\star(y) \big\}, \qquad x \in \Re^n$$

(a) We have

$$f(x) \geq f^{\star\star}(x), \qquad \forall\, x \in \Re^n$$

(b) If f is convex, then properness of any one of f, $f^\star$, and $f^{\star\star}$ implies properness of the other two.

(c) If f is closed proper and convex, then

$$f(x) = f^{\star\star}(x), \qquad \forall\, x \in \Re^n$$

(d) If $\check{\mathrm{cl}}\, f(x) > -\infty$ for all $x \in \Re^n$, then

$$\check{\mathrm{cl}}\, f(x) = f^{\star\star}(x), \qquad \forall\, x \in \Re^n$$

A COUNTEREXAMPLE

- A counterexample (with closed convex but improper f) showing the need to assume properness in order for $f = f^{\star\star}$:

$$f(x) = \begin{cases} \infty & \text{if } x > 0, \\ -\infty & \text{if } x \leq 0. \end{cases}$$

We have

$$f^{\star}(y) = \infty, \qquad \forall\, y \in \Re^n,$$

$$f^{\star\star}(x) = -\infty, \qquad \forall\, x \in \Re^n.$$

But

$$\check{\mathrm{cl}}\, f = f,$$

so $\check{\mathrm{cl}}\, f \neq f^{\star\star}$.

A FEW EXAMPLES

- l_p and l_q norm conjugacy, where $\frac{1}{p} + \frac{1}{q} = 1$

$$f(x) = \frac{1}{p}\sum_{i=1}^{n} |x_i|^p, \qquad f^\star(y) = \frac{1}{q}\sum_{i=1}^{n} |y_i|^q$$

- Conjugate of a strictly convex quadratic

$$f(x) = \frac{1}{2}x'Qx + a'x + b,$$

$$f^\star(y) = \frac{1}{2}(y-a)'Q^{-1}(y-a) - b.$$

- Conjugate of a function obtained by invertible linear transformation/translation of a function p

$$f(x) = p\big(A(x-c)\big) + a'x + b,$$

$$f^\star(y) = q\big((A')^{-1}(y-a)\big) + c'y + d,$$

where q is the conjugate of p and $d = -(c'a + b)$.

SUPPORT FUNCTIONS

- Conjugate of indicator function δ_X of set X

$$\sigma_X(y) = \sup_{x \in X} y'x$$

is called the *support function of* X.

- $\text{epi}(\sigma_X)$ is a closed convex cone.

- The sets X, $\text{cl}(X)$, $\text{conv}(X)$, and $\text{cl}\big(\text{conv}(X)\big)$ all have the same support function (by the conjugacy theorem).

- To determine $\sigma_X(y)$ for a given vector y, we project the set X on the line determined by y, we find $\hat{x}$, the extreme point of projection in the direction y, and we scale by setting

$$\sigma_X(y) = \|\hat{x}\| \cdot \|y\|$$