

LECTURE SLIDES ON
CONVEX ANALYSIS AND OPTIMIZATION
BASED ON 6.253 CLASS LECTURES AT THE
MASS. INSTITUTE OF TECHNOLOGY
CAMBRIDGE, MASS
SPRING 2010

BY DIMITRI P. BERTSEKAS

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

LECTURE 1

AN INTRODUCTION TO THE COURSE

LECTURE OUTLINE

- The Role of Convexity in Optimization
- Duality Theory
- Algorithms and Duality
- Course Organization

HISTORY AND PREHISTORY

- **Prehistory:** Early 1900s - 1949.
 - Caratheodory, Minkowski, Steinitz, Farkas.
 - Properties of convex sets and functions.
- **Fenchel - Rockafellar era:** 1949 - mid 1980s.
 - Duality theory.
 - Minimax/game theory (von Neumann).
 - (Sub)differentiability, optimality conditions, sensitivity.
- **Modern era - Paradigm shift:** Mid 1980s - present.
 - Nonsmooth analysis (a theoretical/esoteric direction).
 - Algorithms (a practical/high impact direction).
 - A change in the assumptions underlying the field.

OPTIMIZATION PROBLEMS

- Generic form:

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

Cost function $f : \mathbb{R}^n \mapsto \mathbb{R}$, constraint set C , e.g.,

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

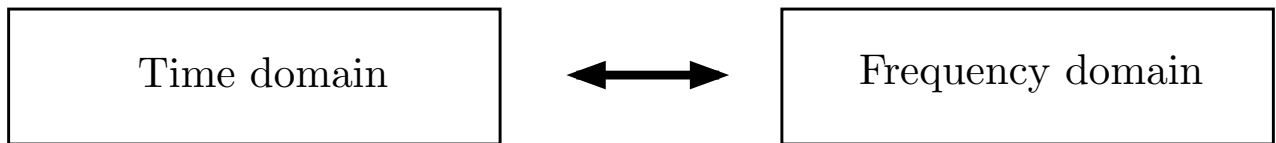
- **Continuous vs discrete problem distinction**
- Convex programming problems are those for which f and C are convex
 - They are continuous problems
 - They are nice, and have beautiful and intuitive structure
- However, **convexity permeates all of optimization**, including discrete problems
- Principal vehicle for continuous-discrete connection is **duality**:
 - The dual problem of a discrete problem is continuous/convex
 - The dual problem provides important information for the solution of the discrete primal (e.g., lower bounds, etc)

WHY IS CONVEXITY SO SPECIAL?

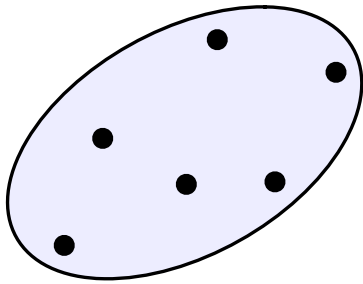
- A convex function has no local minima that are not global
- A nonconvex function can be “convexified” while maintaining the optimality of its global minima
- A convex set has a nonempty relative interior
- A convex set is connected and has feasible directions at any point
- The existence of a global 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

DUALITY

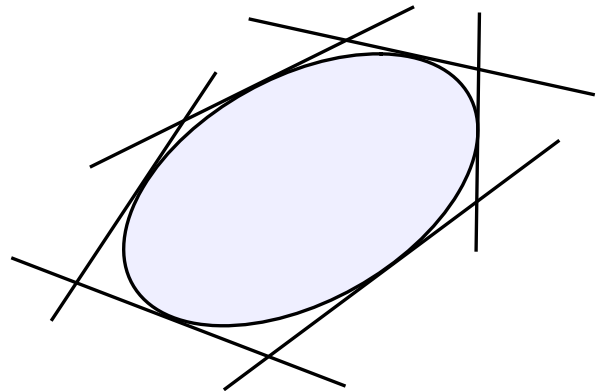
- Two different views of the same object.
- Example: Dual description of signals.



- Dual description of **closed** convex sets



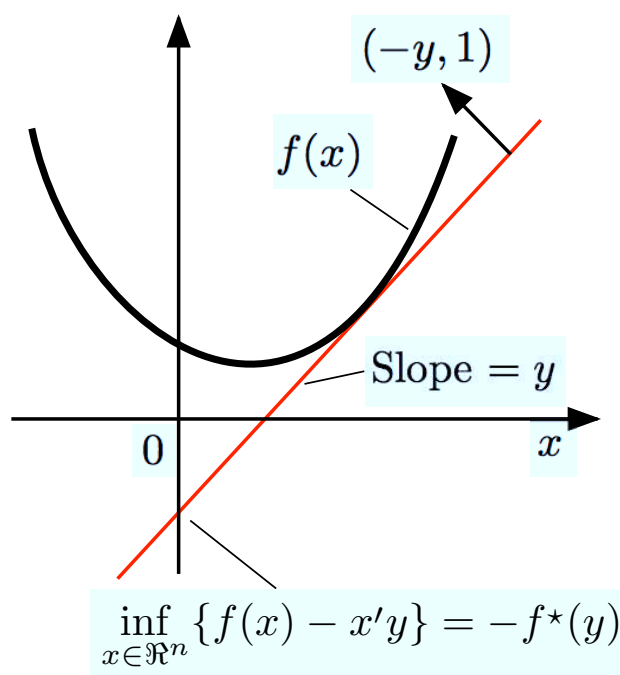
A union of points



An intersection of halfspaces

DUAL DESCRIPTION OF CONVEX FUNCTIONS

- Define a closed convex function by its epigraph.
- Describe the epigraph by hyperplanes.
- Associate hyperplanes with crossing points (the **conjugate function**).



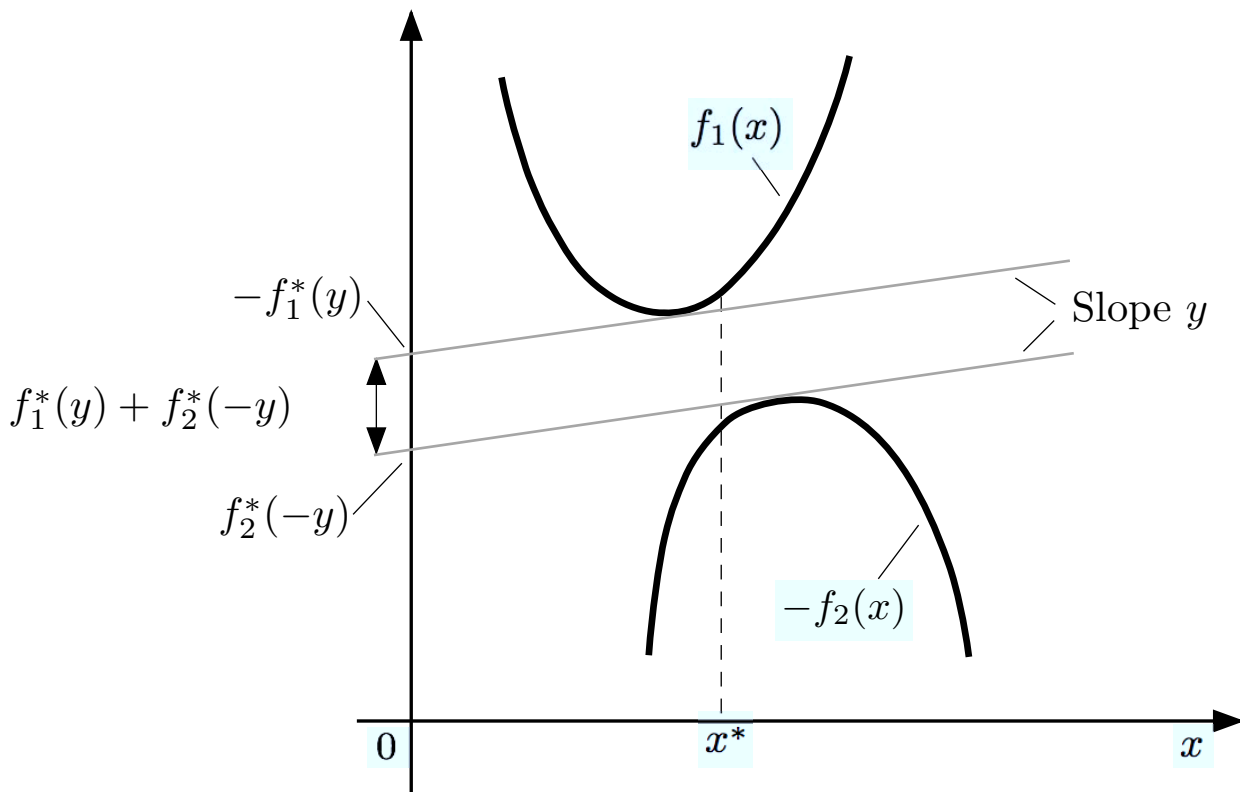
Primal Description

Values $f(x)$

Dual Description

Crossing points $f^*(y)$

FENCHEL PRIMAL AND DUAL PROBLEMS



Primal Problem Description
Vertical Distances

Dual Problem Description
Crossing Point Differentials

- Primal problem:

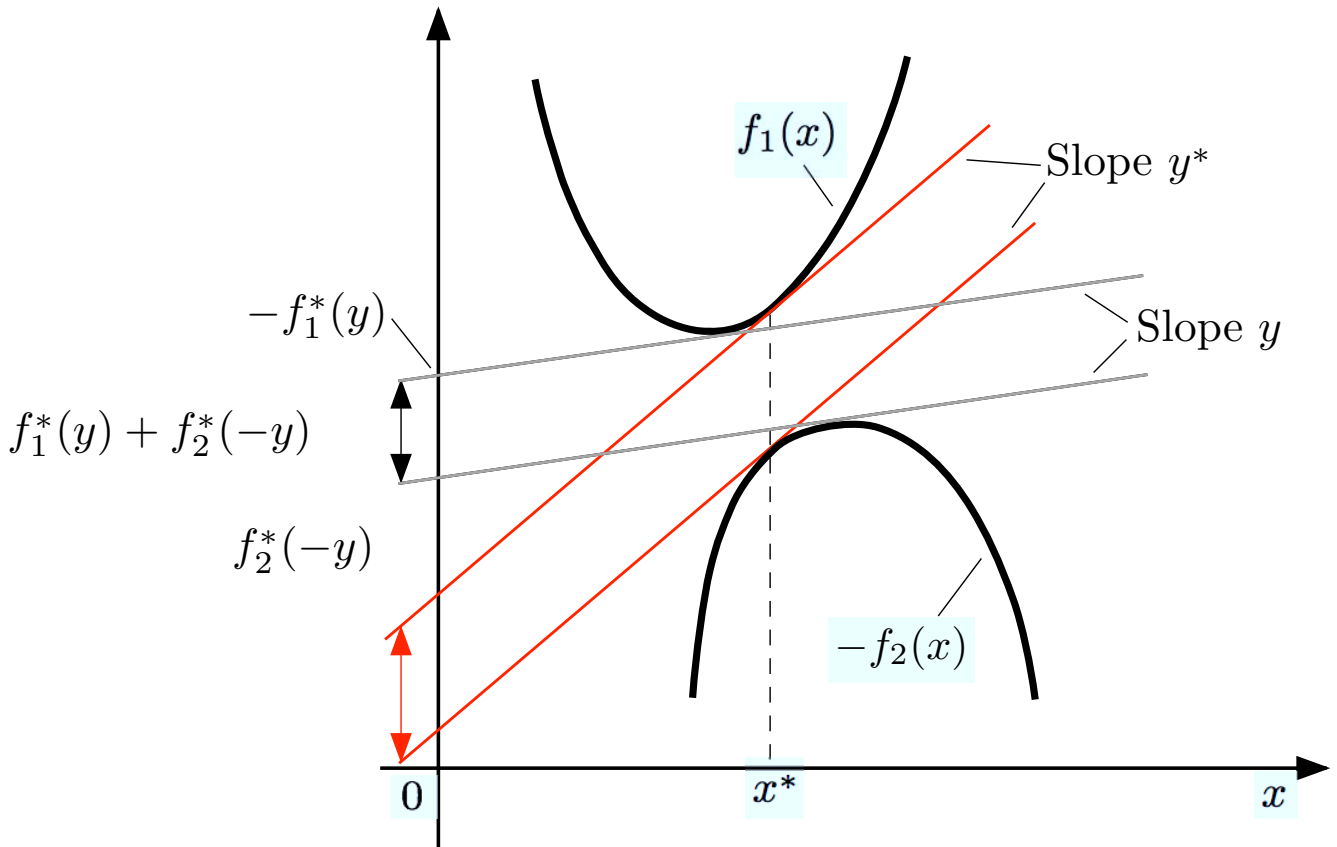
$$\min_x f_1(x) + f_2(x)$$

- Dual problem:

$$\max_y -f_1^*(y) - f_2^*(-y)$$

where f_1^* and f_2^* are the conjugates

FENCHEL DUALITY



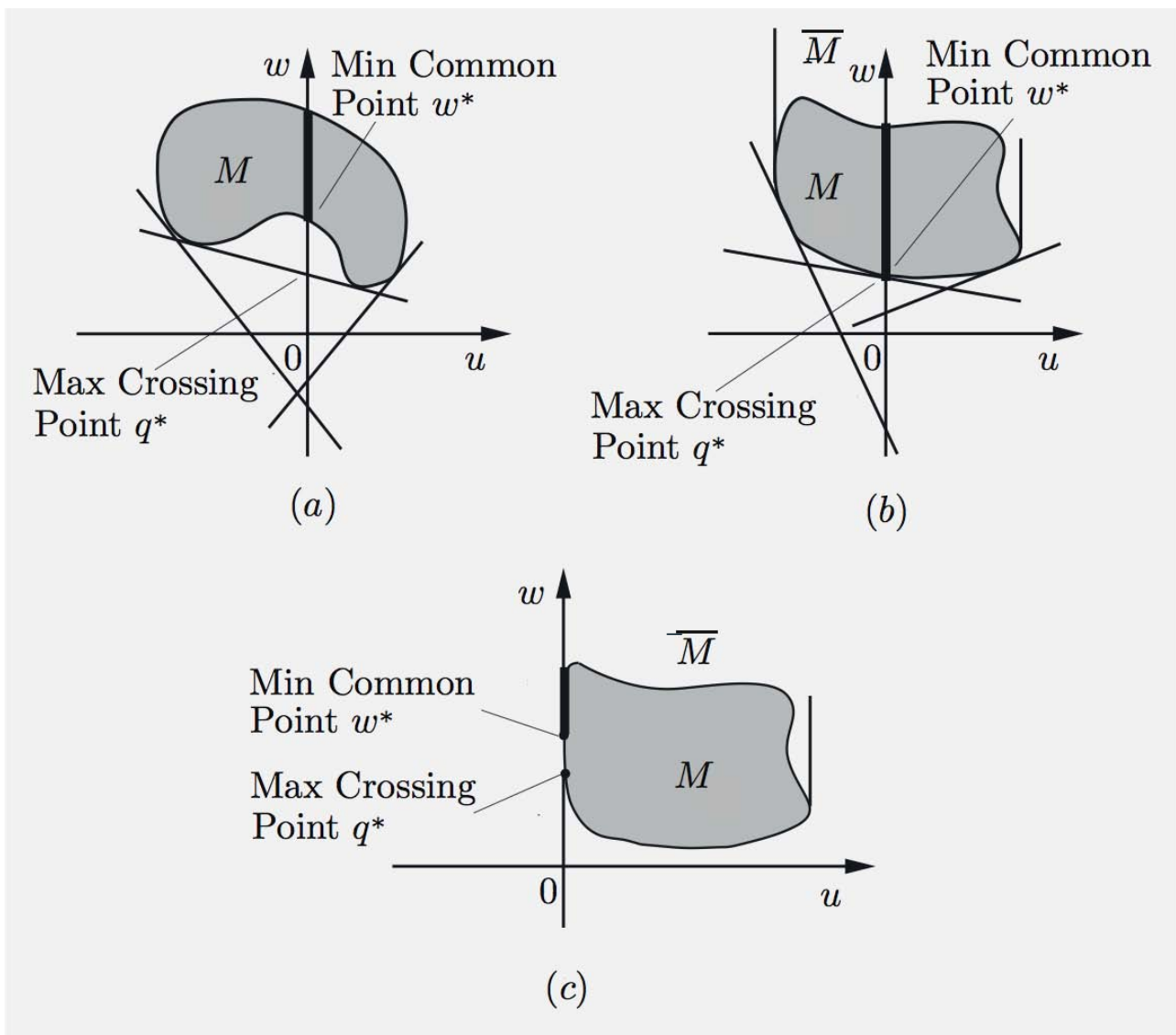
$$\min_x f_1(x) + f_2(x) = \max_y -f_1^*(y) - f_2^*(-y)$$

- Under favorable conditions (convexity):
 - The optimal primal and dual values are equal
 - The optimal primal and dual solutions are related

A MORE ABSTRACT VIEW OF DUALITY

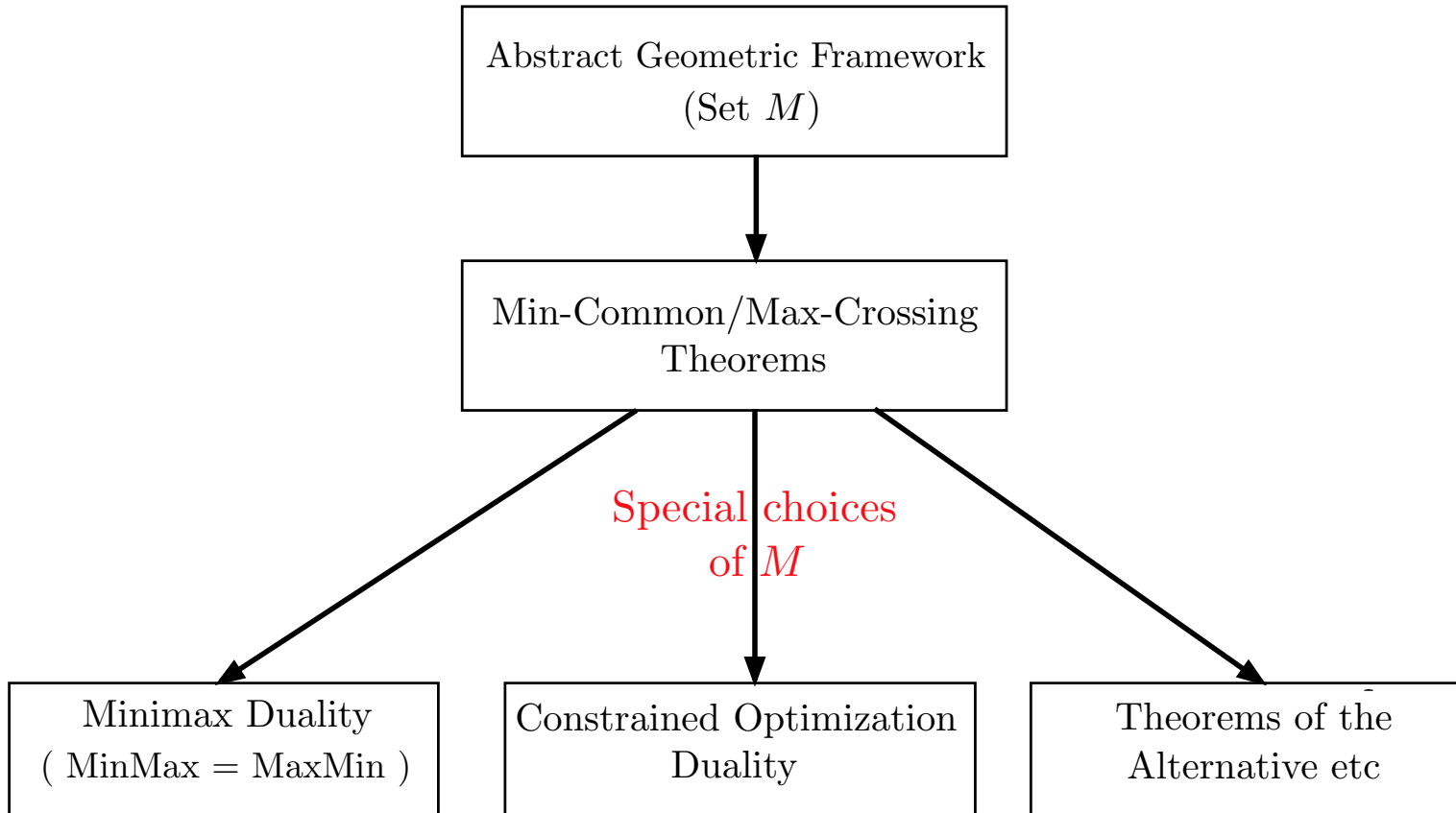
- Despite its elegance, the Fenchel framework is somewhat indirect.
- From duality of set descriptions, to
 - duality of functional descriptions, to
 - duality of problem descriptions.
- A more direct approach:
 - **Start with a set**, then
 - **Define two simple prototype problems** dual to each other.
- Avoid functional descriptions (a simpler, less constrained framework).

MIN COMMON/MAX CROSSING DUALITY



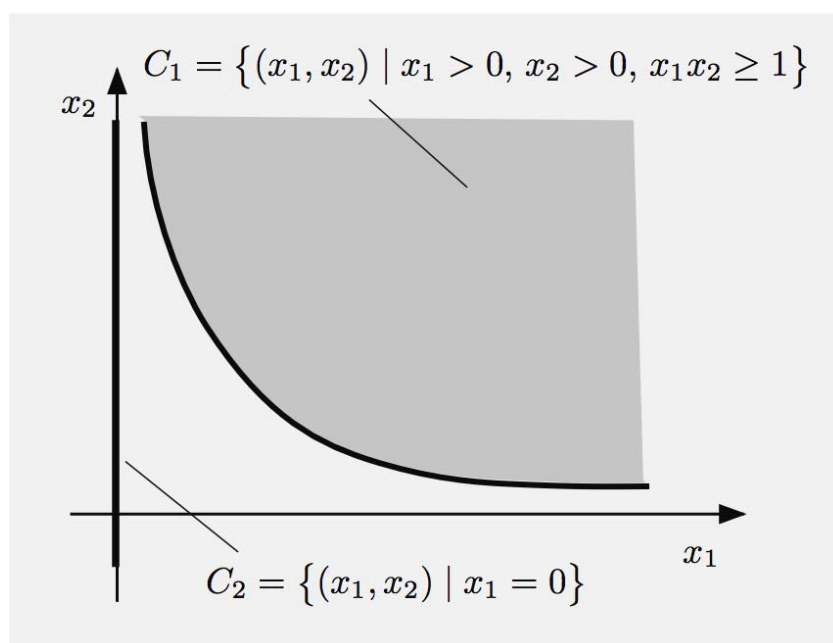
- All of duality theory and all of (convex/concave) minimax theory can be developed/explained in terms of this one figure.
- The machinery of convex analysis is needed to flesh out this figure, and to rule out the exceptional/pathological behavior shown in (c).

ABSTRACT/GENERAL DUALITY ANALYSIS



EXCEPTIONAL BEHAVIOR

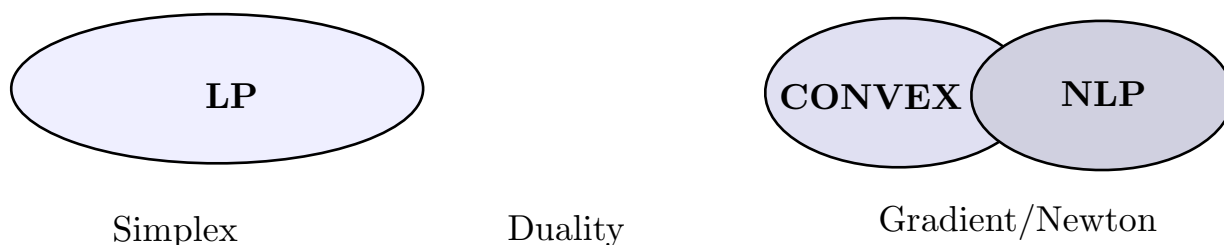
- If convex structure is so favorable, what is the source of exceptional/pathological behavior?
- **Answer:** Some common operations on convex sets do not preserve some basic properties.
- **Example:** A linearly transformed closed convex set need not be closed (contrary to compact and polyhedral sets).
 - Also the vector sum of two closed convex sets need not be closed.



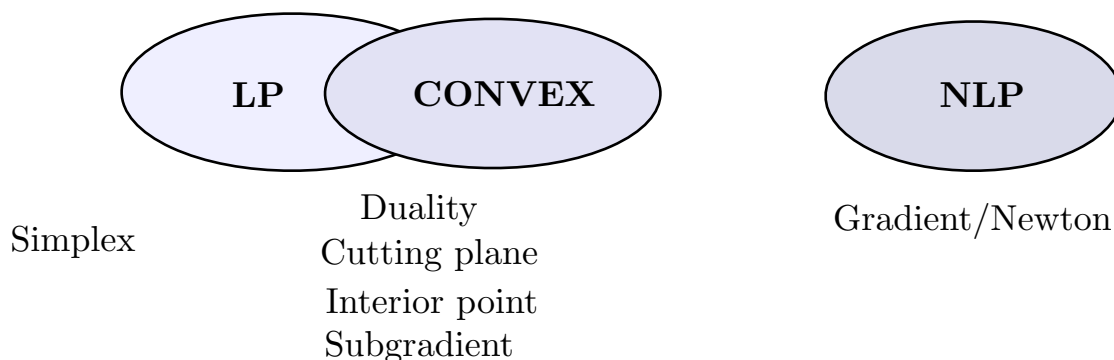
- This is a major reason for the analytical difficulties in convex analysis and pathological behavior in convex optimization (and the favorable character of polyhedral sets).

MODERN VIEW OF CONVEX OPTIMIZATION

- **Traditional view:** Pre 1990s
 - LPs are solved by simplex method
 - NLPs are solved by gradient/Newton methods
 - Convex programs are special cases of NLPs



- **Modern view:** Post 1990s
 - LPs are often solved by nonsimplex/convex methods
 - Convex problems are often solved by the same methods as LPs
 - “Key distinction is not Linear-Nonlinear but Convex-Nonconvex” (Rockafellar)



THE RISE OF THE ALGORITHMIC ERA

- Convex programs and LPs connect around
 - Duality
 - Large-scale piecewise linear problems
- Synergy of:
 - Duality
 - Algorithms
 - Applications
- New problem paradigms with rich applications
- Duality-based decomposition
 - Large-scale resource allocation
 - Lagrangian relaxation, discrete optimization
 - Stochastic programming
- Conic programming
 - Robust optimization
 - Semidefinite programming
- Machine learning
 - Support vector machines
 - l_1 regularization/Robust regression/Compressed sensing

METHODOLOGICAL TRENDS

- New methods, renewed interest in old methods.
 - Interior point methods
 - Subgradient/incremental methods
 - Polyhedral approximation/cutting plane methods
 - Regularization/proximal methods
 - Incremental methods
- Renewed emphasis on complexity analysis
 - Nesterov, Nemirovski, and others ...
 - “Optimal algorithms” (e.g., extrapolated gradient methods)
- Emphasis on interesting (often duality-related) large-scale special structures

COURSE OUTLINE

- We will follow closely the textbook
 - Bertsekas, “Convex Optimization Theory,” Athena Scientific, 2009, including the on-line Chapter 6 and supplementary material at <http://www.athenasc.com/convexduality.html>
- Additional book references:
 - Rockafellar, “Convex Analysis,” 1970.
 - Boyd and Vanderbergue, “Convex Optimization,” Cambridge U. Press, 2004. (On-line at <http://www.stanford.edu/~boyd/cvxbook/>)
 - Bertsekas, Nedic, and Ozdaglar, “Convex Analysis and Optimization,” Ath. Scientific, 2003.
- Topics (the text’s design is modular, and the following sequence involves no loss of continuity):
 - **Basic Convexity Concepts:** Sect. 1.1-1.4.
 - **Convexity and Optimization:** Ch. 3.
 - **Hyperplanes & Conjugacy:** Sect. 1.5, 1.6.
 - **Polyhedral Convexity:** Ch. 2.
 - **Geometric Duality Framework:** Ch. 4.
 - **Duality Theory:** Sect. 5.1-5.3.
 - **Subgradients:** Sect. 5.4.
 - **Algorithms:** Ch. 6.

WHAT TO EXPECT FROM THIS COURSE

- Requirements: Homework (25%), midterm (25%), and a term paper (50%)
- We aim:
 - To develop insight and deep understanding of a fundamental optimization topic
 - To treat with mathematical rigor an important branch of methodological research, and to provide an account of the state of the art in the field
 - To get an understanding of the merits, limitations, and characteristics of the rich set of available algorithms
- Mathematical level:
 - Prerequisites are linear algebra (preferably abstract) and real analysis (a course in each)
 - Proofs will matter ... but the rich geometry of the subject helps guide the mathematics
- 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
 - You can do your term paper on an application area

A NOTE ON THESE SLIDES

- These slides are a teaching aid, not a text
- Don't expect a rigorous mathematical development
- 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 understanding of ideas, not to express them precisely
- The omitted proofs and a fuller discussion can be found in the “Convex Optimization Theory” textbook

MIT OpenCourseWare
<http://ocw.mit.edu>

6.253 Convex Analysis and Optimization
Spring 2010

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.