Convex Optimization
A Journey of 60 Years

Dimitri P. Bertsekas

Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology

LIDS Paths Ahead Symposium, 2009

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.

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.

History and Prehistory

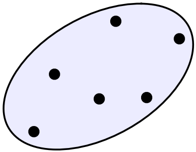
- Prehistory: Early 1900s - 1949.
  - Caratheodory, Minkowski, Steinitz, Farkas.
  - Properties of convex sets and functions.
- Fenchel - Rockafellar era: 1949 - mid 1980s.
  - Duality theory.
  - Minimax/game theory (von Neumann).
  - (Sub)differentiability, optimality conditions, sensitivity.
- Modern 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.

Duality

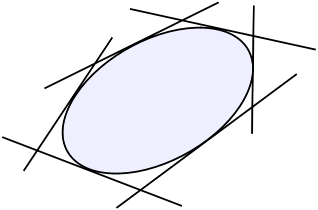
- Two different views of the same object.
- Example: Dual description of signals.

Time domain $\longleftrightarrow$ Frequency domain

- Dual description of closed convex sets



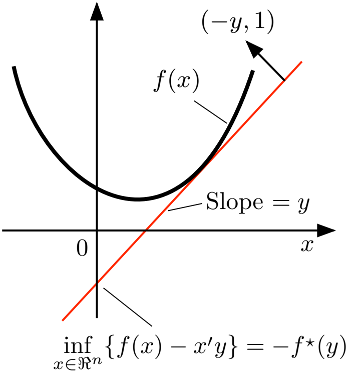
A union of points



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).

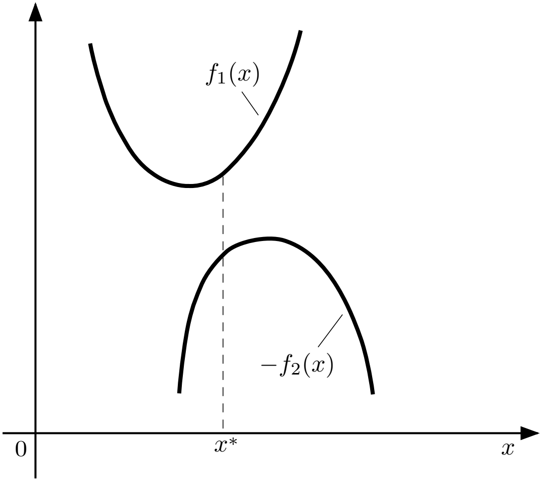


Primal Description

Values $f(x)$

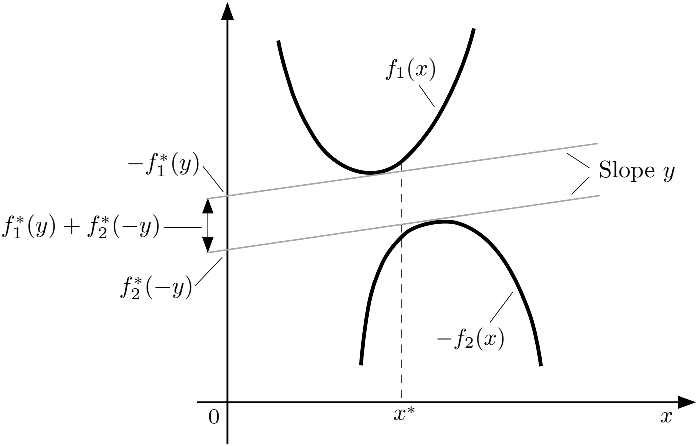
Dual Description

Crossing points $f^*(y)$



$$\min_x \left\{ f_1(x) + f_2(x) \right\}$$

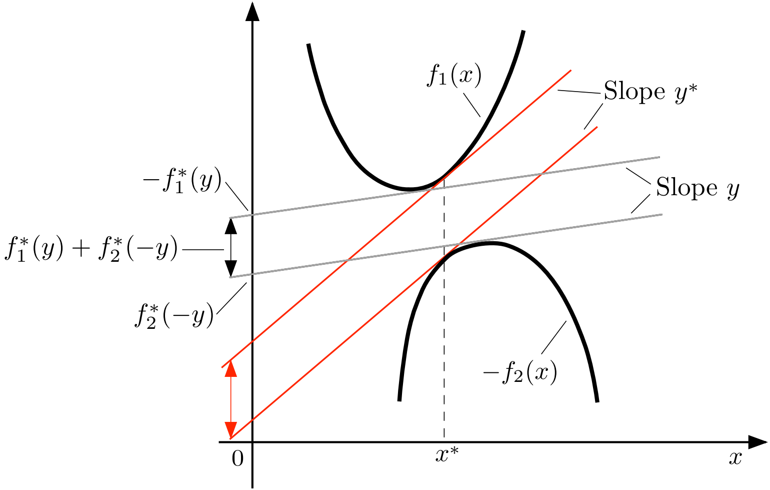
Fenchel Primal and Dual Problem Descriptions



Primal Description
Vertical Distances

Dual Description
Crossing Point Differentials

Fenchel Duality



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

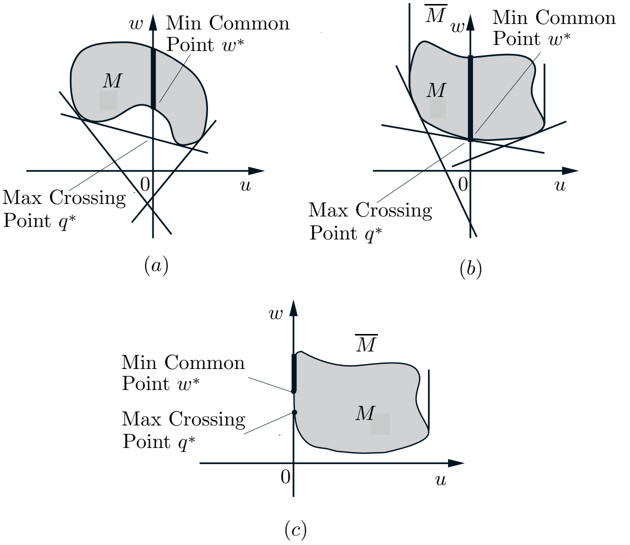
A More Abstract View of Duality

- Back to the primal and dual description of a set M.
- Two simple prototype problems dual to each other.

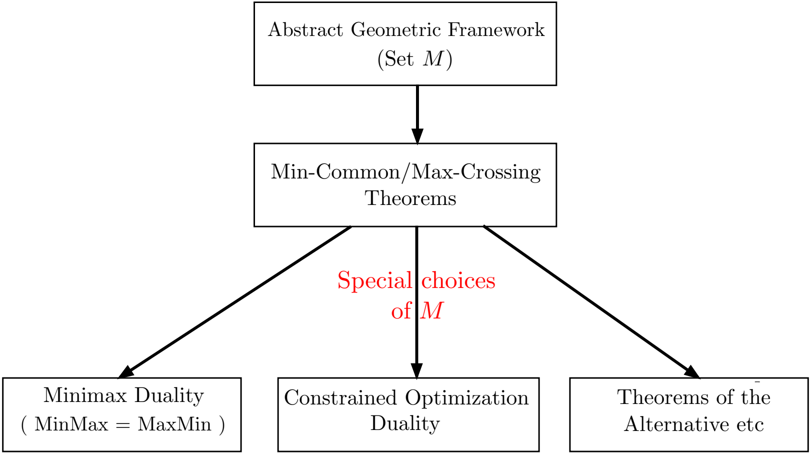
A More Abstract View of Duality

- Back to the primal and dual description of a set M.
- Two simple prototype problems dual to each other.

Min-Common/Max-Crossing Duality

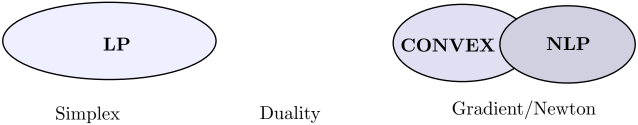


Abstract Framework for Duality Analysis

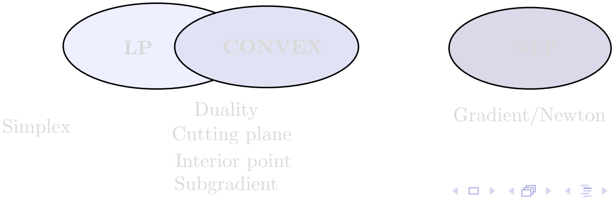


The Modern Era: Duality Coupled with Algorithms

- Traditional view: Pre 1990s
  - LPs are solved by simplex method (G. Dantzig view).
  - NLPs are solved by gradient/Newton methods (M. Powell view).
  - Convex programs are special cases of NLPs.

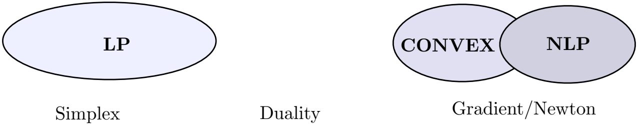


- Modern view: Post 1990s
  - LPs are often solved by nonsimplex/convex methods.
  - Convex problems are often solved by the same methods as LPs.
  - "Key distinction is not Linear-Nonlinear but Convex-Nonconvex" (Rockafellar)

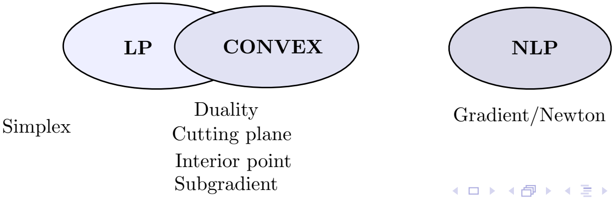


The Modern Era: Duality Coupled with Algorithms

- Traditional view: Pre 1990s
  - LPs are solved by simplex method (G. Dantzig view).
  - NLPs are solved by gradient/Newton methods (M. Powell view).
  - 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)



Methodological Trends

- Convex programs and LPs connect around duality and large-scale piecewise linear problems.
- New methods, renewed interest in old methods

  Interior point methods
  Subgradient methods
  Polyhedral approximation/cutting plane methods
  Regularization/proximal methods
- Renewed emphasis on complexity analysis

  Nesterov, Nemirovski, and others ...
  Extrapolated gradient methods

Methodological Trends

- Convex programs and LPs connect around duality and large-scale piecewise linear problems.
- New methods, renewed interest in old methods

  Interior point methods
  Subgradient methods
  Polyhedral approximation/cutting plane methods
  Regularization/proximal methods

- Renewed emphasis on complexity analysis

  Nesterov, Nemirovski, and others ...
  Extrapolated gradient methods

Methodological Trends

- Convex programs and LPs connect around duality and large-scale piecewise linear problems.
- New methods, renewed interest in old methods

  Interior point methods
  Subgradient methods
  Polyhedral approximation/cutting plane methods
  Regularization/proximal methods

- Renewed emphasis on complexity analysis

  Nesterov, Nemirovski, and others ...
  Extrapolated gradient methods

Synergy Between Duality, Algorithms, and 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
  Incremental methods

Synergy Between Duality, Algorithms, and 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
  Incremental methods

Synergy Between Duality, Algorithms, and 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
  Incremental methods

Speculation - What's Next?

- Very large problems/new applications.

  Problems with network overlays (e.g., smart grids).
  Huge data sets in machine learning.

- New approaches to large size and complexity.

  Approximate dynamic programming paradigm (e.g., LP-based dynamic programming).
  Reduced space approximations.
  Sampling mechanisms.

- Better hardware/better algorithms multiplier effect?

- A new paradigm?

Speculation - What's Next?

- Very large problems/new applications.

  Problems with network overlays (e.g., smart grids).
  Huge data sets in machine learning.

- New approaches to large size and complexity.

  Approximate dynamic programming paradigm (e.g., LP-based dynamic programming).
  Reduced space approximations.
  Sampling mechanisms.

- Better hardware/better algorithms multiplier effect?

- A new paradigm?

Speculation - What's Next?

- Very large problems/new applications.

  Problems with network overlays (e.g., smart grids).
  Huge data sets in machine learning.

- New approaches to large size and complexity.

  Approximate dynamic programming paradigm (e.g., LP-based dynamic programming).
  Reduced space approximations.
  Sampling mechanisms.

- Better hardware/better algorithms multiplier effect?

- A new paradigm?

Speculation - What's Next?

- Very large problems/new applications.

  Problems with network overlays (e.g., smart grids).
  Huge data sets in machine learning.

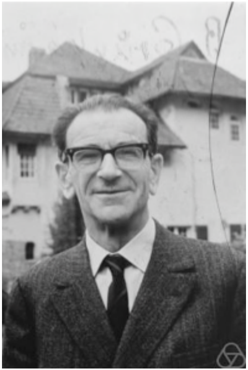
- New approaches to large size and complexity.

  Approximate dynamic programming paradigm (e.g., LP-based dynamic programming).
  Reduced space approximations.
  Sampling mechanisms.

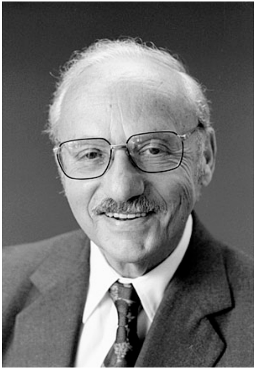
- Better hardware/better algorithms multiplier effect?

- A new paradigm?

Fenchel, Dantzig, Rockafellar



Werner Fenchel



George Dantzig

Terry Rockafellar

Paul Tseng, 1959-2009