# CONVEX OPTIMIZATION: A SELECTIVE OVERVIEW

**Dimitri Bertsekas**
**M.I.T.**

**Taiwan**
**May 2010**

# OUTLINE

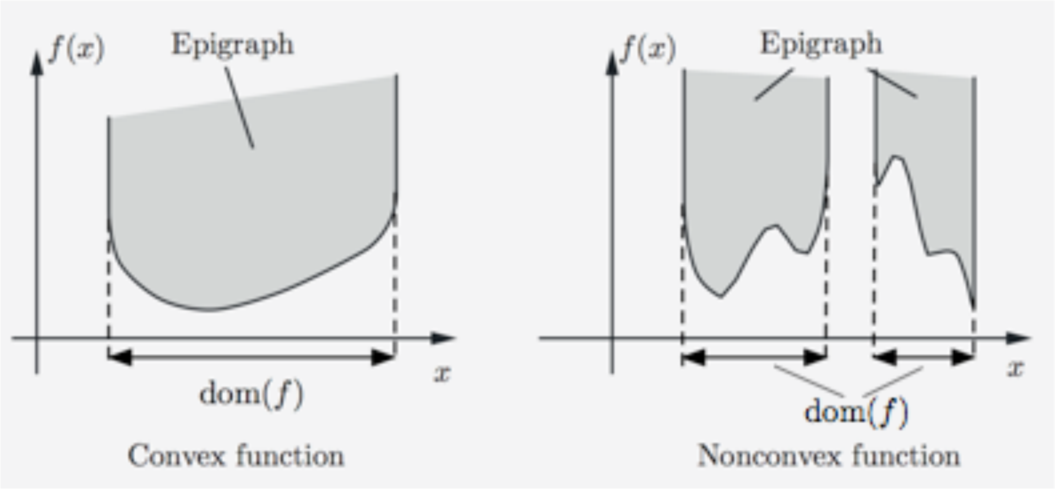
- Convexity issues in optimization
- Common geometrical framework for duality and minimax
- Unifying framework for existence of solutions and duality gap analysis
- Use of duality in algorithms

# SOME HISTORY

- Late 19th-Early 20th Century:
  - Caratheodory, Minkowski, Steinitz, Farkas
- 40s-50s: The big turning point
  - Game Theory: von Neumann
  - Optimization-related convexity: Fenchel
  - Duality: Fenchel, Princeton group (Nash, Gale, Kuhn, Tucker)
- 60s-70s: Consolidation
  - Rockafellar
- 80s-90s: Extensions to nonconvex optimization and nonsmooth analysis
  - Clarke, Mordukovich, Rockafellar-Wets
- 2000- … Rejuvenation: Many applications of large scale optimization using duality; resource allocation, combinatorial optimization, machine learning

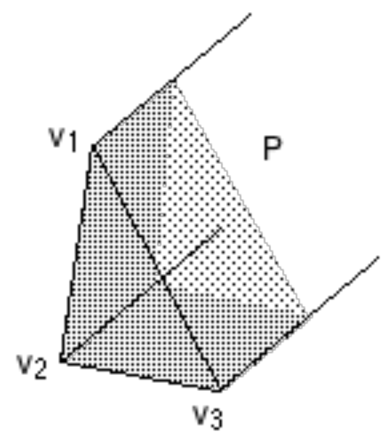
# WHY IS CONVEXITY IMPORTANT IN OPTIMIZATION I

- **A convex function has no local minima that are not global**
- **A nonconvex function can be "convexified" while maintaining the optimality of its minima**



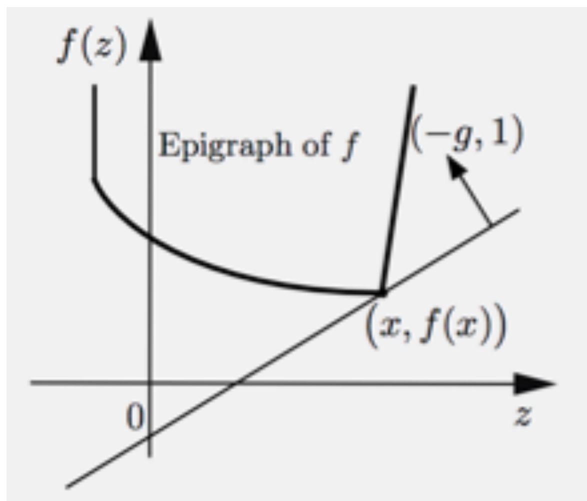
# WHY IS CONVEXITY IMPORTANT IN OPTIMIZATION II

- **A polyhedral convex set is characterized by its extreme points and extreme directions**
- **Minima of linear functions over constraint sets can be found among extreme points**



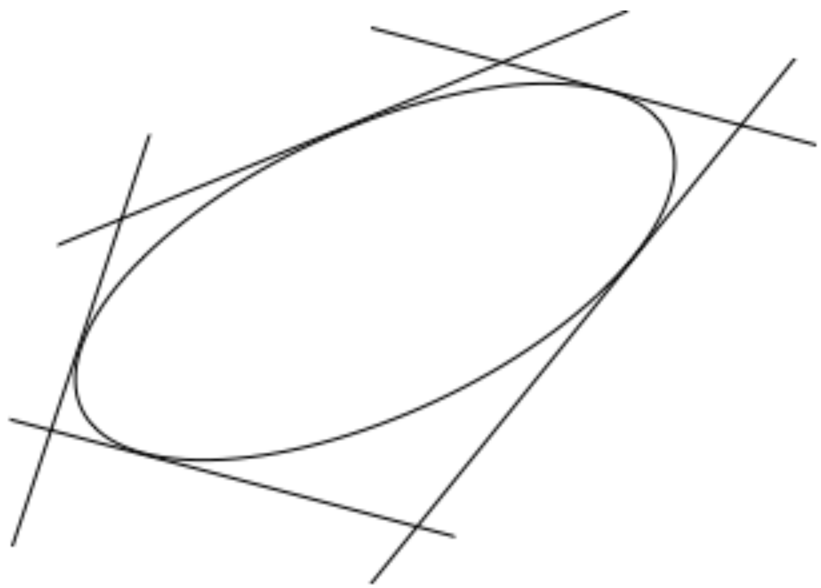
# WHY IS CONVEXITY IMPORTANT IN OPTIMIZATION III

- **A real-valued convex function is continuous and has nice differentiability properties**



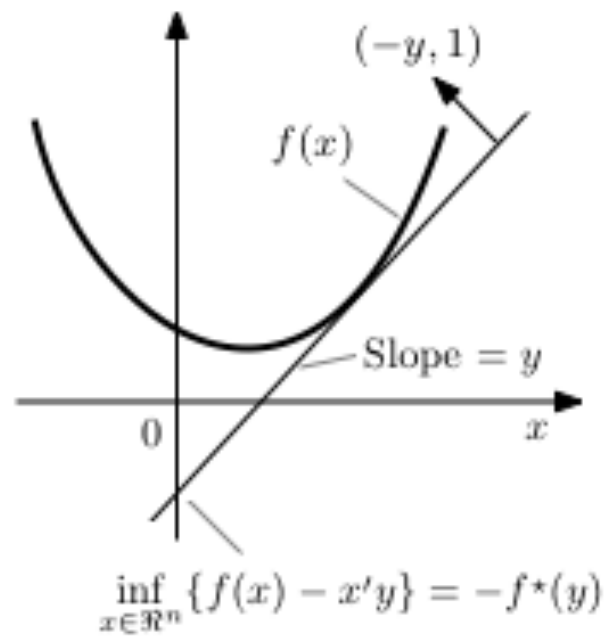

- **Convex functions arise prominently in duality (a different but equivalent view of the same object)**

# DUAL DESCRIPTION OF CONVEX SETS



Primal description: Points    Dual description: Hyperplanes

# CONJUGACY: DUAL DESCRIPTION OF CONVEX FUNCTIONS




Primal description: Values $f(x)$ Dual description: Crossing points $f^\star(y)$

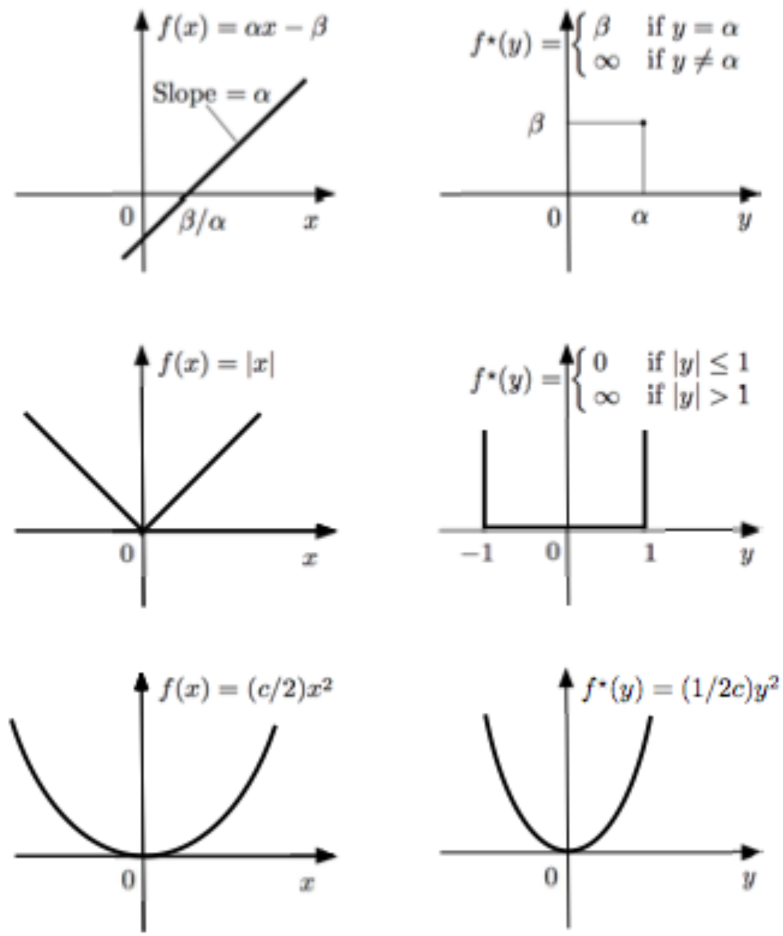
# CONJUGATE FUNCTION PAIRS

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

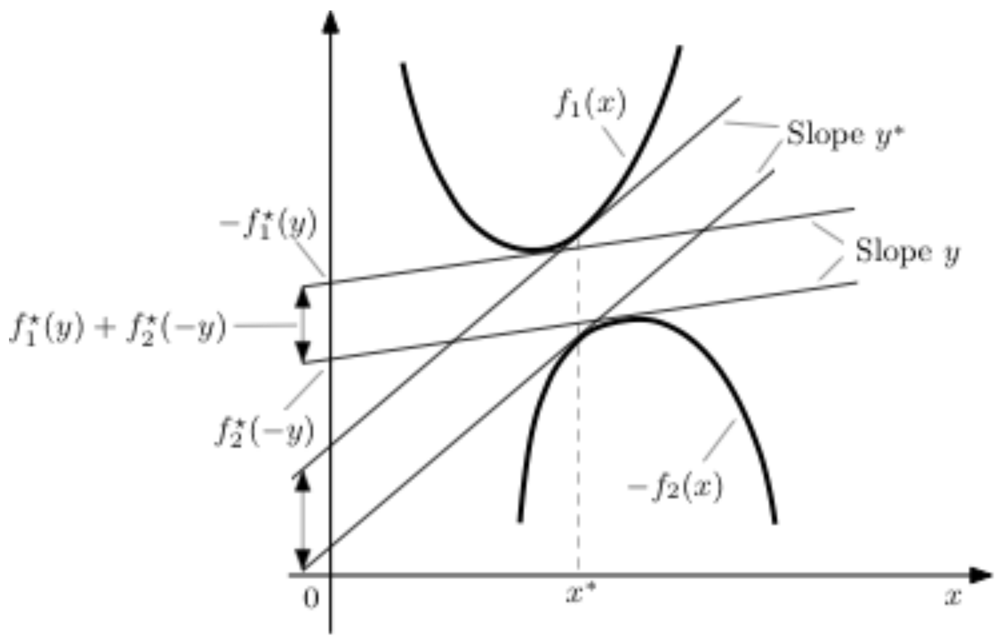
$$f^{\star\star}(x) = \sup_{y \in \Re^n} \{y'x - f^\star(y)\}$$

**Conjugacy theorem:**

$$f^{\star\star} = f$$

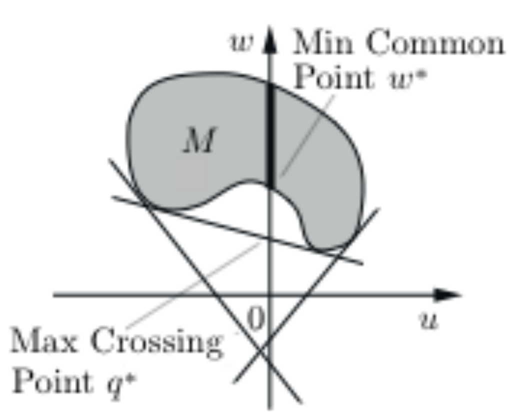


# FENCHEL DUALITY

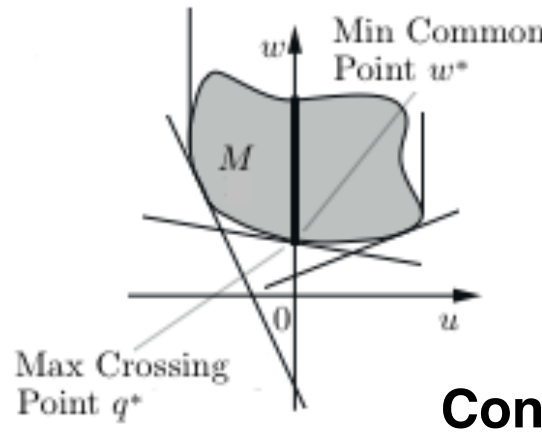



$$\min_x \{f_1(x) + f_2(x)\} = \max_y \{f_1^\star(y) + f_2^\star(-y)\}$$

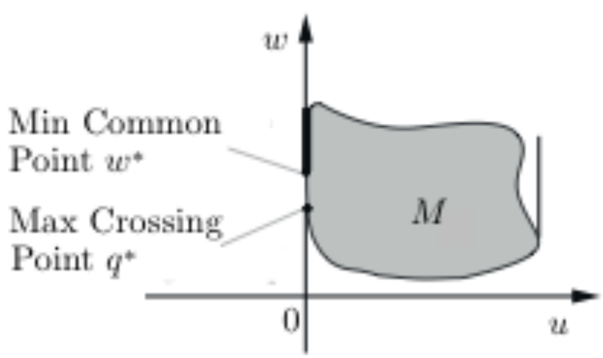
# MIN COMMON/MAX CROSSING DUALITY




Nonconvex



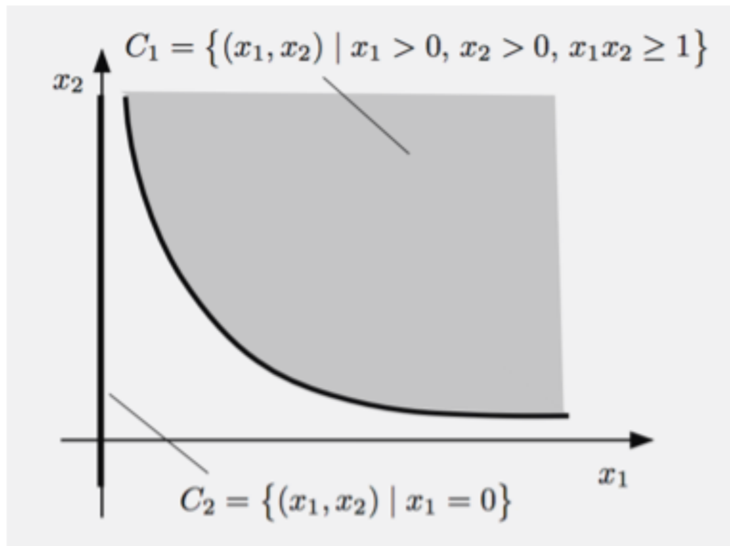

Convex
No duality gap




Convex
Duality gap

# A MAJOR ANOMALY OF CONVEX SETS

- **Linear transformations and vector sums need not preserve closure of convex sets**



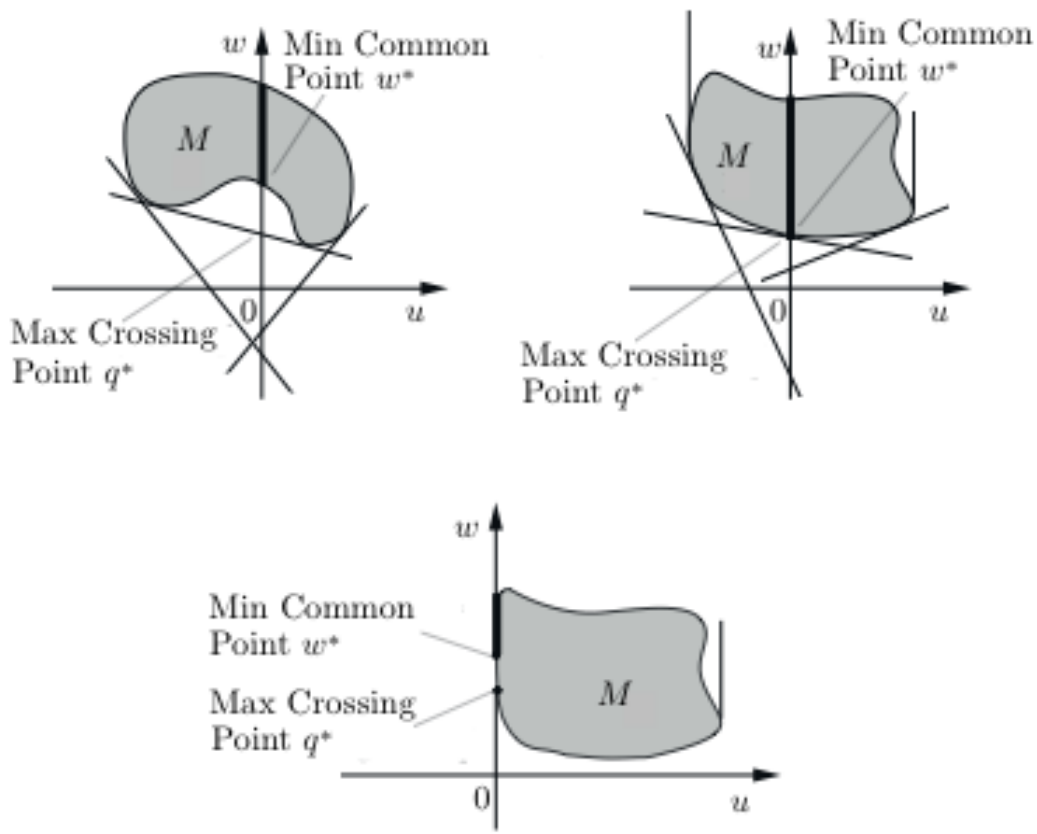

- **Source of duality gap**
- **Nice sets: Polyhedral, or convex and compact**

# BOOKS

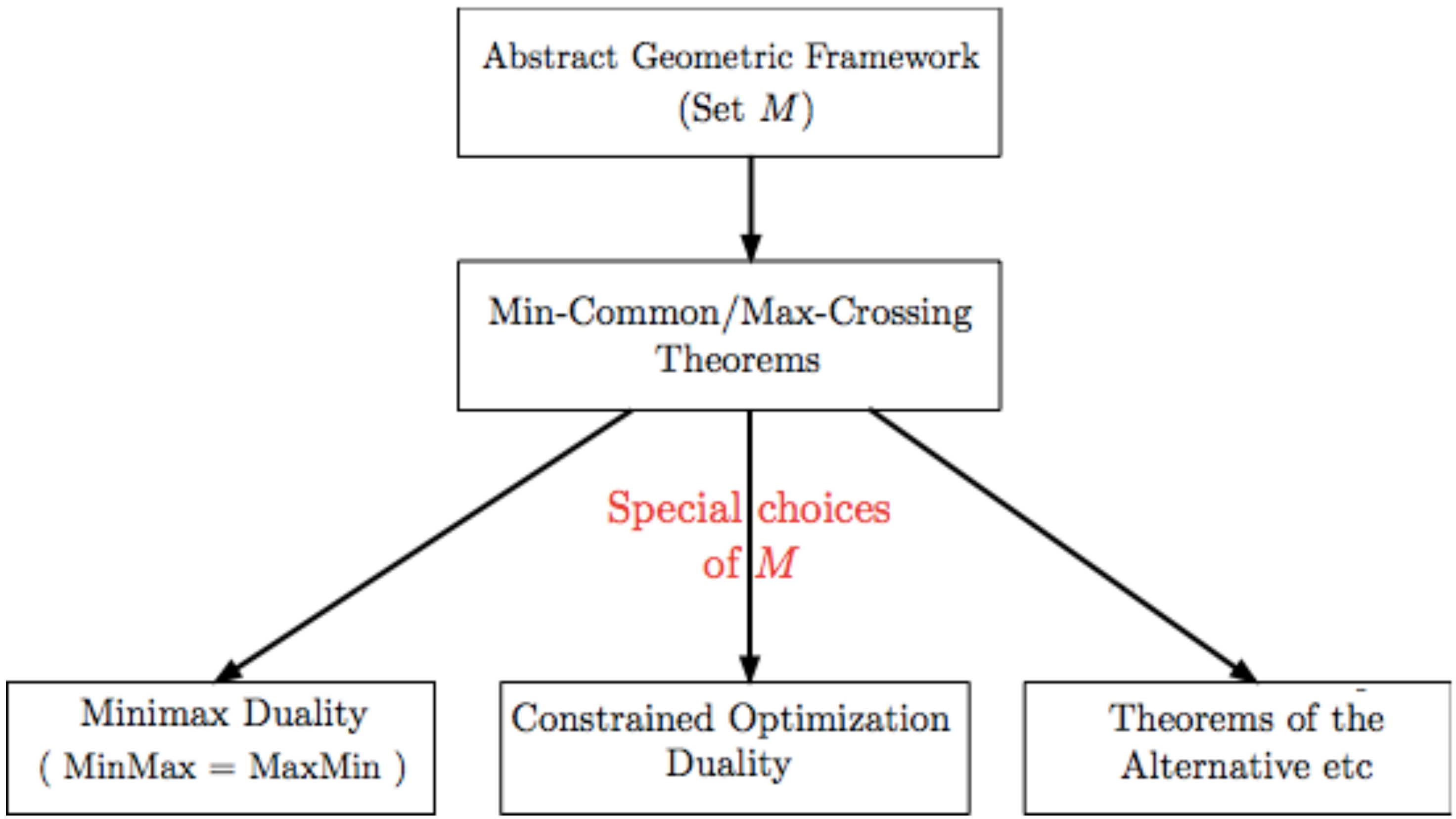
- Convex Analysis and Optimization, by D. P. Bertsekas, with A. Nedic and A. Ozdaglar (March 2003 - extends to nonconvex analysis - available in China)
- Convex Optimization Theory, by D. P. Bertsekas (short, more narrowly/deeply focused on convexity 2009 - with algorithms www supplement)
- Convex Optimization Algorithms (in preparation)

# I
# MIN COMMON/MAX CROSSING DUALITY

# GEOMETRICAL VIEW OF DUALITY



# ANALYTICAL APPROACH



# CONVEX PROGRAMMING DUALITY

- **Primal problem:**

  min $f(x)$ subject to $x \in X$ and $g_j(x) \leq 0$, $j=1,\ldots,r$

- **Dual problem:**

  max $q(\mu)$ subject to $\mu \geq 0$

  where the dual function is

$$q(\mu) = \inf_{x \in X} \{ f(x) + \mu' g(x) \}$$

- **Consider common/max crossing framework:**

$$M = \text{epi}(p), \qquad p(u) = \inf_{x \in X,\, g_j(x) \leq u_j} f(x)$$

# MINIMAX / ZERO SUM GAME THEORY ISSUES

- Given a function $\Phi(x,z)$, where $x \in X$ and $z \in Z$, under what conditions do we have

$$\inf_x \sup_z \Phi(x,z) = \sup_z \inf_x \Phi(x,z)$$

- Assume convexity/concavity, semicontinuity of $\Phi$
- Min common/max crossing framework:

M = epigraph of p

$$p(u) = \inf_x \sup_z \{\Phi(x,z) - u'z\}$$

$\inf_x \sup_z \Phi$ = Min common value

$\sup_z \inf_x \Phi$ = Max crossing value (can be shown)

# II
# UNIFICATION OF EXISTENCE AND NO DUALITY GAP ISSUES

# INTERSECTIONS OF NESTED FAMILIES OF CLOSED SETS

- **We will connect two basic problems in optimization**
  - Attainment of a minimum of a function f over a set X
  - Existence of a duality gap
- **The 1st question is a set intersection issue:**

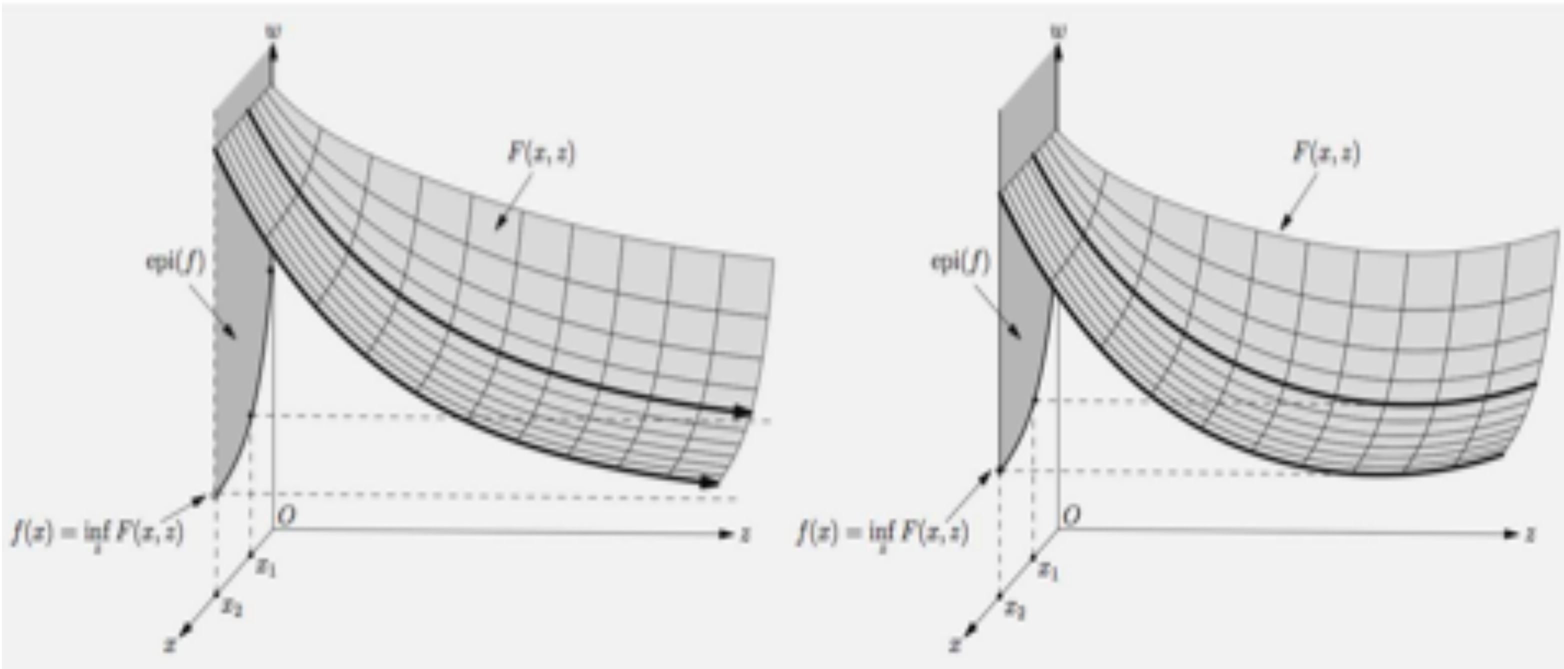
  **The set of minima is the intersection of the nonempty level sets $\{x \in X \mid f(x) \leq \gamma\}$**
- **The 2nd question is also a set intersection issue (but not obvious). It is related to another fundamental question:**

  **When is the function**

  $$f(x) = \inf_z F(x,z)$$

  **lower semicontinuous, assuming F(x,z) is convex and lower semicontinuous?**

# VISUALIZATION OF PARTIAL MINIMIZATION $f(x)=\inf_z F(x,z)$



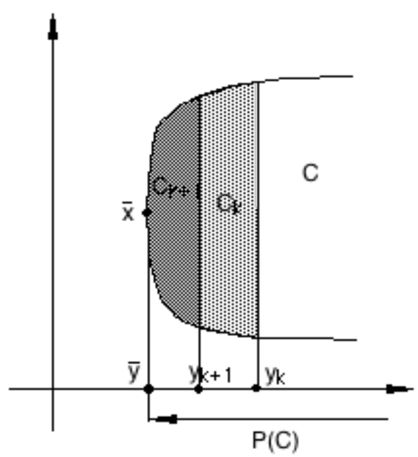

$$P\big(\text{epi}(F)\big) \subset \text{epi}(f) \subset \text{cl}\Big(P\big(\text{epi}(F)\big)\Big)$$

# PRESERVATION OF SEMICONTINUITY UNDER PARTIAL MINIMIZATION

- **Key observation: For $f(x) = \inf_z F(x,z)$, we have**

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

**where P(·) is projection on the space of z. So if projection preserves closedness, f is l.s.c.**




**Given C, when is P(C) closed?**

**If $y_k$ is a sequence in P(C) that converges to $\bar{y}$, we must show that the intersection of the $C_k$ is nonempty**

# UNIFIED TREATMENT OF EXISTENCE OF SOLUTIONS AND DUALITY GAP ISSUES

Results on nonemptiness of intersection of a nested family of closed sets (use of directions of recession)

No duality gap results In convex programming

inf sup $\Phi$ = sup inf $\Phi$

Existence of minima of f over X

# III
# PROBLEM STRUCTURES AND ALGORITHMS

# SEPARABLE PROBLEMS & DECOMPOSITION

$$\text{minimize} \quad \sum_{i=1}^{n} f_i(x_i)$$
$$\text{subject to} \quad a_j' x \le b_j, \quad j = 1, \ldots, r,$$

$$\text{maximize} \quad \sum_{i=1}^{n} q_i(\mu) - \sum_{j=1}^{r} \mu_j b_j \qquad q_i(\mu) = \inf_{x_i \in \Re} \left\{ f_i(x_i) + x_i \sum_{j=1}^{r} \mu_j a_{ji} \right\}$$
$$\text{subject to} \quad \mu \in \Re^r,$$

**Dual function calculated by decomposition**

# DISCRETE/INTEGER LP

$$\begin{aligned} &\text{minimize} \quad c'x \\ &\text{subject to} \ \ Ax \le b, \quad x_i = 0 \text{ or } 1, \quad i = 1, \ldots, n \end{aligned}$$

- Typically there is duality gap
- Solution of dual problem provides a lower bound
- Use in Lagrangian relaxation and branch-and-bound
- Dual problem is nondifferentiable/polyhedral

# ADDITIVE COST PROBLEMS

$$f(x) = \sum_{i=1}^{m} f_i(x)$$

- **Huge number m of terms**
- **This is a common structure:**
  - **Dual problems of separable problems**
  - **Expected values (e.g. in stochastic programming) have this structure**
- **Calculation of gradient or subgradient of the sum is very time consuming**
- **Need for an incremental algorithmic approach**
  - **Move x along the gradient/subgradient of a single component $f_i$**

# LARGE NUMBER OF CONSTRAINTS

$$\begin{aligned} &\text{minimize} && f(x) \\ &\text{subject to} && a_j'x \le b_j, \quad j = 1, \ldots, r, \end{aligned}$$

- Calls for approximation of the constraint set
- Outer approximation
- Inner approximation
- Inner and outer approx are dual to each other
- An alternative: Penalty approach converts to a minimization of a large sum

$$\text{minimize} \quad f(x) + c\sum_{j=1}^{r} P(a_j'x - b_j)$$

# CLASSES OF ALGORITHMS

- **Descent methods**
  - Subgradient methods and incremental versions
  - Optimal algorithms (gradient methods with extrapolation) - Nesterov's methods
- **Approximation methods**
  - Cutting plane
  - Simplicial decomposition
  - Proximal and bundle method
  - Interior point methods
- **All these methods rely on convexity concepts**

# CONCLUDING REMARKS

- Optimization has become a universal tool in applications
- Convexity is the "soul" of optimization
- Geometry is the "soul" of convexity
- Very few simple geometric ideas are sufficient to unify/clarify most of convex optimization
- Theoretical/algorithmic research on convex optimization is still very active

**Thank you!**