A Geometric Framework for Nonconvex Optimization Duality and Penalty Methods

Asu Ozdaglar

Joint work with Angelia Nedić, UIUC

Conference on Complementarity, Duality, and Global Optimization

February, 2007

Department of Electrical Engineering & Computer Science

Massachusetts Institute of Technology

Introduction

$$\begin{array}{rl} \text{minimize} & f_0(x) \\ \text{subject to} & x \in X,\ f(x) = (f_1(x), \ldots, f_m(x)) \leq 0 \end{array}$$

Dual Problem: $\max_{\mu \in \mathbb{R}^m} q(\mu) = \inf_{x \in X} \{f_0(x) + \mu' f(x)\}$

Penalized problem: Solve a sequence of problems

$$q(c_k) = \inf_{x \in X} \{f_0(x) + c_k P(f(x))\}\ \text{, as } c_k \to \infty.$$

- Key Idea Common to Duality and Penalty
  - Relaxing the inequality constraints and augmenting the objective with some “constraint violation cost function”
- Key Issue Common to Duality and Penalty
  - Zero-gap between the optimal value of the original constrained problem and the dual/penalized problem.
- However, in the existing literature
  - They are treated separately
  - Zero-gap results under compactness and convexity assumptions.

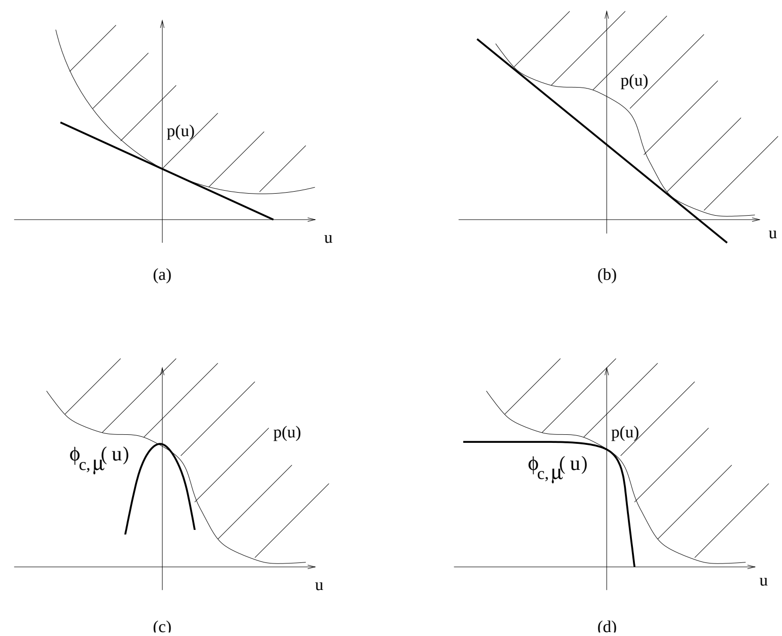
Our Work

- Our objective is to develop a unifying framework
  - For the analysis of both duality schemes and penalty methods
  - Applicable to a wide class of nonconvex problems
- Main Results:
  - A geometric framework defined in terms of augmenting functions
  - Separation results for nonconvex sets via general concave surfaces
  - Establishing necessary and sufficient conditions for zero duality gap in the geometric framework
  - Application of results to optimization duality and penalty methods using the primal function of the constrained optimization problem

Main Idea

Traditional duality relies on supporting the epigraph of the primal function using hyperplanes.

For nonconvex problems, support the epigraph using nonlinear surfaces, defined by augmenting functions



- **Related Literature:**
  - Rockafellar and Wets [98] use convex, nonnegative, and level-bounded augmenting functions and show zero-gap under coercivity assumptions.
  - Rubinov, Huang, Yang [02] study dual problems constructed by a family of augmenting functions satisfying peak at zero property.

Geometric Framework

- **Geometric Primal Problem:** Given a nonempty (nonconvex) set $V \subset \mathbb{R}^m \times \mathbb{R}$ intersecting the w-axis, find the minimum value intercept of V and the w-axis, i.e.,

$$w^* = \inf_{(0,w)\in V} w.$$

- **Geometric Dual Problem:** Defined by an augmenting function

  Definition: A function $\sigma : \mathbb{R}^m \mapsto (-\infty, \infty]$ is called an **augmenting function** if it is convex, not identically equal to 0, and $\sigma(0) = 0$.

- Given an augmenting function σ, geometric dual problem considers concave surfaces $\{(u, \phi_{c,\mu}(u)) \mid u \in \mathbb{R}^m\}$ that lie below the set V, where

$$\phi_{c,\mu}(u) = -c\sigma(u) - \mu' u + \xi.$$

Geometric Dual Problem

- This surface is below V if and only if

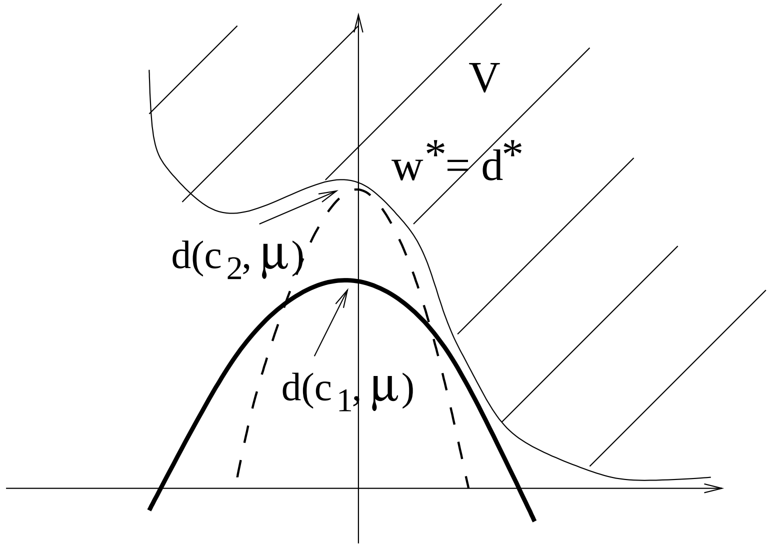
$$w + c\sigma(u) + \mu' u \geq \xi \quad \text{for all } (u, w) \in V.$$

The maximum intercept of such surface with the w-axis is given by

$$d(c, \mu) = \inf_{(u,w)\in V} \{w + c\sigma(u) + \mu' u\}.$$

The geometric dual problem consists of determining the maximum intercept of such surfaces over $c \geq 0$ and $\mu \in \mathbb{R}^m$, i.e.,

$$d^* = \sup_{c\geq 0,\ \mu\in\mathbb{R}^m} d(c, \mu).$$



Zero Duality Gap

- We are interested in conditions under which $d^* = w^*$, i.e., there is *zero duality gap*.
- Proposition (Weak Duality): The dual optimal value does not exceed the primal optimal value, $d^* \leq w^*$.
- To establish zero duality gap:
  - We study conditions on the set V and the augmenting function σ under which we can separate V from a vector $(0, w_0)$ that does not belong to $\mathrm{cl}(V)$.
- We say that the augmenting function σ strongly separates the set V and the vector $(0, w_0) \notin \mathrm{cl}(V)$ when for some $c \geq 0$ and $\xi \in \mathbb{R}$,

$$w + c\sigma(u) \geq \xi > w_0 \quad \text{for all } (u, w) \in V.$$

Notation and Terminology

- For a function $f : \mathbb{R}^n \mapsto [-\infty, \infty]$ and any scalar γ, we denote the (lower) γ-level set of f by $L_f(\gamma)$, i.e.,

$$L_f(\gamma) = \{x \in \mathbb{R}^n \mid f(x) \leq \gamma\}.$$

- We say that the function f is level-bounded when the set $L_f(\gamma)$ is bounded for every scalar γ.

- For a given nonempty set X, the cone generated by the set X is denoted by $\text{cone}(X)$ and is given by

$$\text{cone}(X) = \{y \mid y = \lambda x \text{ for some } x \in X \text{ and } \lambda \geq 0\}.$$

- The asymptotic cone V^∞ of a nonempty set V is given by

$$V^\infty = \{d \mid \lambda_k x_k \to d \text{ for some } \{x_k\} \subset V \text{ and } \{\lambda_k\} \subset \mathbb{R} \text{ with } \lambda_k \downarrow 0\}.$$

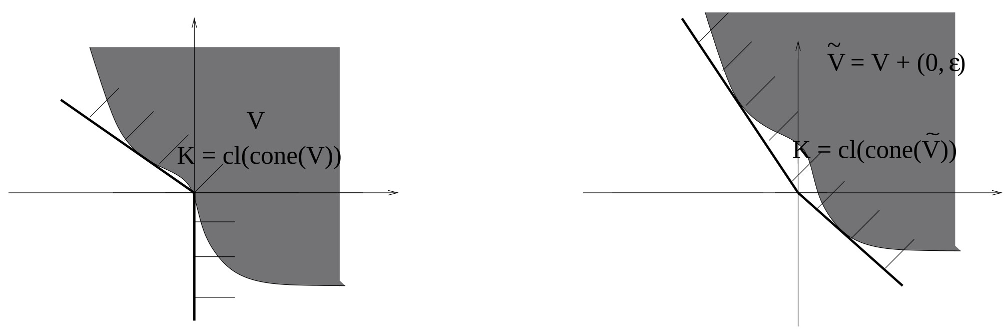
Properties of the set V

- Definition: We say that a set $V \subset \Re^m \times \Re$ is *extending upward in w-space (or u-space)* if for every vector $(\bar{u}, \bar{w}) \in V$, the half-line $\{(\bar{u}, w) \mid w \geq \bar{w}\}$ (or the cone $\{(u, \bar{w}) \mid u \geq \bar{u}\}$) is contained in V.
  - Both satisfied when V epigraph of a nonincreasing function.
- Lemma: Assume that $(0, -1)$ is not an asymptotic direc of V, i.e., $(0, -1) \notin V^\infty$. Let $(0, w_0) \notin \text{cl}(V)$. For a given $\epsilon > 0$, consider the set $\tilde{V}$ given by

$$\tilde{V} = \{(u, w) \mid (u, w - \epsilon) \in V\}, \tag{1}$$

and the cone generated by $\tilde{V}$, denoted by K. Then,

$$(0, w_0) \notin \text{cl}(K).$$



Separation Properties of Augmenting Function

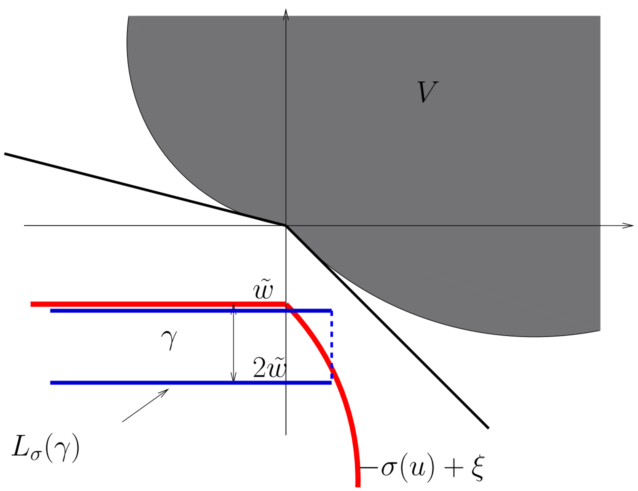
- Key Lemma: Let $\sigma : \mathbb{R}^m \mapsto (-\infty, \infty]$ be a nonnegative augmenting function. Let $C \subset \mathbb{R}^m \times \mathbb{R}$ be a nonempty cone, and let $\tilde{w}$ be a scalar with $\tilde{w} < 0$. Furthermore, let $\gamma > 0$ be a scalar such that

$$\{(u, \tilde{w}) \mid u \in L_\sigma(\gamma)\} \cap C = \emptyset.$$

Then, the set X defined by

$$X = \{(u, w) \in \mathbb{R}^m \times \mathbb{R} \mid w \leq -\frac{|\tilde{w}|}{\gamma}\sigma(u) + \tilde{w}\}$$

has no vector in common with the cone C.



Separation Theorem

- Assumption 1: (a) $w^* = \inf_{(0,w)\in V} w$ is finite.

  (b) The set V extends upward in u-space and w-space.

  (c) $(0,-1) \notin V^\infty$.

- Assumption 2: (a) The function σ is nonnegative, $\sigma(u) \geq 0$ for all u.

  (b) Given a sequence $\{u_k\} \subset \mathbb{R}^m$,

  $$\sigma(u_k) \to 0 \quad \Rightarrow \quad u_k^+ \to 0,$$

  where $u^+ = (\max\{0,u_1\},\ldots,\max\{0,u_m\})'$.

- Ex: $\sigma(u) = \max\{0, u_1, ..., u_m\}$, $\sigma(u) = \sum_{i=1}^m (\max\{0,u_i\})^\beta$, $\beta > 0$.

- Assumption 2(b) is equivalent to the following: for all $\delta > 0$, there holds

  $$\inf_{\{u \mid \operatorname{dist}(u,\mathbb{R}_-^m)\geq\delta\}} \sigma(u) > 0.$$

  – Related to the *peak at zero condition* studied by Rubinov et. al.

  – Satisfied by augmenting functions studied by Rockafellar-Wets and Huang-Yang.

Separation Theorem

- Theorem: Let Assumptions 1 and 2 hold. Then, the set V and a vector $(0, w_0) \notin \mathrm{cl}(V)$ can be strongly separated by the function σ, i.e., there exist scalars $c > 0$ and ξ such that

$$w + c\sigma(u) \geq \xi > w_0 \quad \text{for all } (u, w) \in V.$$

  Proof: Let K be the cone generated by upward translation of set V

  - Using Assumption 2(b) and the "northeast" extension property of set V, we show that there exists some γ s.t.

$$\{(u, w_0/2) \mid u \in L_\sigma(\gamma)\} \cap \mathrm{cl}(K) = \emptyset.$$

  - Using Key Lemma, we obtain $X \cap \mathrm{cl}(K) = \emptyset$, where

$$X = \{(u, w) \in \mathbb{R}^m \times \mathbb{R} \mid w \leq -\frac{|w_0|}{2\gamma}\sigma(u) + \frac{w_0}{2}\},$$

  - This implies that

$$w + \frac{|w_0|}{2\gamma}\sigma(u) > \frac{w_0}{2} \quad \text{for all } (u, w) \in \mathrm{cl}(K).$$

Necessary and Sufficient Conditions for Geometric Zero Duality Gap

- Proposition (Necessary Conditions): Let σ be an augmenting function that is continuous at the origin. Assume that there is zero duality gap, i.e., $d^* = w^*$. Then, for any sequence $\{(u_k, w_k)\} \subset V$ with $u_k \to 0$, we have

$$\liminf_{k \to \infty} w_k \geq w^*.$$

- Proposition (Sufficient Conditions): Let Assumptions 1 and 2 hold. Assume that for any sequence $\{(u_k, w_k)\} \subset V$ with $u_k \to 0$, we have

$$\liminf_{k \to \infty} w_k \geq w^*.$$

  Then, there is zero duality gap, i.e., $d^* = w^*$.

Constrained Optimization Duality

- We consider the following primal problem

$$\begin{aligned} f^* = \min_{x \in \mathbb{R}^n} \quad & f_0(x) \\ \text{s.t.} \quad & x \in X, \ f(x) = (f_1(x), \ldots, f_m(x)) \leq 0 \end{aligned}$$

- We define a dualizing parametrization function $\bar{f}$ as

$$\bar{f}(x, u) = \begin{cases} f_0(x) & \text{if } f(x) \leq u, \\ +\infty & \text{otherwise.} \end{cases}$$

- Given an augmenting function σ, we define the the augmented dual function as

$$q(c, \mu) = \inf_{x \in X} l(x, c, \mu) = \inf_{u \in \mathbb{R}^m} \{\bar{f}(x, u) + c\sigma(u) + \mu' u\}.$$

- The augmented dual problem is given by

$$\begin{aligned} q^* = \max \quad & q(c, \mu) \\ \text{s.t.} \quad & c \geq 0, \ \mu \in \mathbb{R}^m. \end{aligned}$$

Zero Duality Gap

- Consider the primal function $p : \mathbb{R}^m \mapsto [-\infty, \infty]$ of the optimization problem,

$$p(u) = \inf_{x \in X,\ f(x) \leq u} f_0(x).$$

- Let V be the epigraph of the primal function, $V = \text{epi}(p)$.

$$w^* = p(0) = f^*,$$

$$d(c, \mu) = \inf_{\{(u,w) \mid p(u) \leq w\}} \{w + c\sigma(u) + \mu' u\} = q(c, \mu).$$

- Proposition (Zero Duality Gap): Assume that f^* is finite and that $(0, -1) \notin (\text{epi}(p))^\infty$. Let σ satisfy Assumption 2. Assume further that $p(u)$ is lower semicontinuous at $u = 0$.

  Then, there is zero duality gap, i.e., $q^* = f^*$.

- The condition $(0, -1) \notin (\text{epi}(p))^\infty$ is satisfied, for example, when $\inf_{x \in X} f_0(x) > -\infty$.

Penalty Methods

- We consider the constrained optimization problem

$$f^* = \min_{x \in \mathbb{R}^n} \quad f_0(x)$$
$$\text{s.t.} \quad x \in X, \; f(x) = (f_1(x), \ldots, f_m(x)) \le 0$$

- We are interested in penalty methods of the form

$$\tilde{f}(c) = \min \quad \{f_0(x) + c\sigma(f(x))\}$$
$$\text{s.t.} \quad x \in X,$$

  where $c \ge 0$ is a penalty parameter that will ultimately increase to $+\infty$.

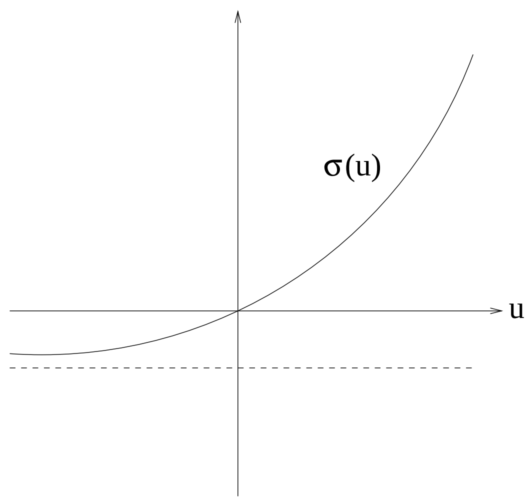
- Use the geometric framework with $V = \text{epi}(p)$ and $\mu = 0$.

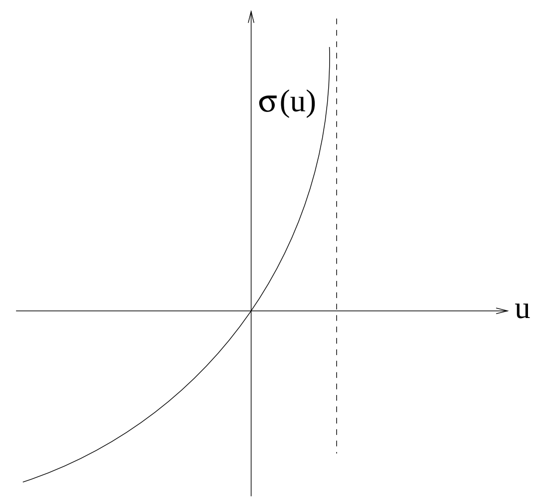
- Proposition (Sufficient Conditions for Penalty Convergence): Let the assumptions of zero duality gap proposition hold. Assume further that the augmenting function $\sigma(u)$ is nondecreasing in u. Then,

$$\lim_{c \to \infty} \tilde{f}(c) = f^*.$$

Extensions to Negative and Unbounded Augmenting Functions

- In recent work, we extended the geometric framework to:
  - Bounded-below augmenting functions; e.g. $\sigma(u) = a(e^u - 1)$
  - Unbounded augmenting functions; e.g. $\sigma(u) = -\log(1-u)$
  - Asymptotic augmenting functions; e.g. $\sigma(u) = -\log(1-u)$





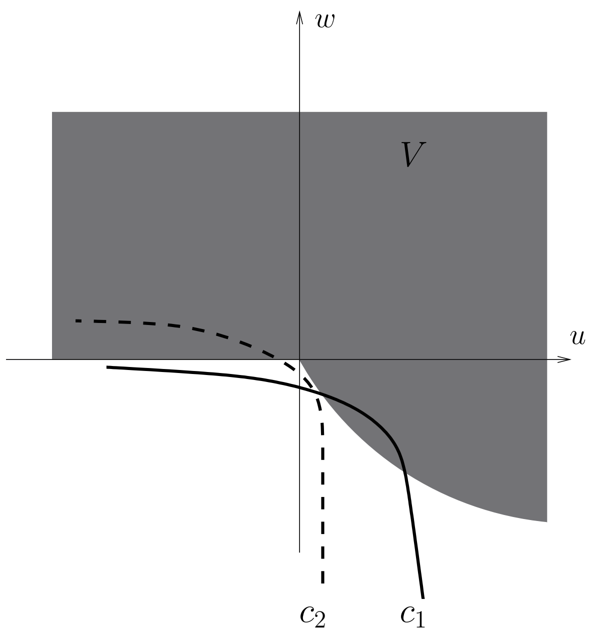
Geometric Framework for General Augmenting Functions

- It is immediate from the geometry that linear scaling by the penalty parameter c not sufficient for negative augmenting functions:

Separating the point $(0, w_0)$, with $w_0 < 0$, from the set V using concave surfaces of the form

$$-c_i(e^u - 1) + \xi,$$

with $c_2 > c_1$.



- We consider separation using concave surfaces defined by

$$\phi_{c,\mu}(u) = -\frac{1}{c}\,\sigma(cu) + \xi.$$

Separation Theorem for Bounded Augmenting Functions

- Assumption:
  - (a) The function σ is bounded-below, i.e., $\sigma(u) \geq \sigma_0$.
  - (b) For any sequence $\{u_k\} \subset \mathbb{R}^m$ and any positive scalar sequence $\{c_k\}$ with $c_k \to \infty$,

$$\limsup_{k \to \infty} \frac{\sigma(c_k u_k)}{c_k} < \infty \quad \Rightarrow \quad u_k^+ \to 0.$$

- Example: $\sigma(u) = \sum_{i=1}^m a_i(e^{u_i} - 1)$, $u \in \mathbb{R}^m$, $a_i > 0$. (Tseng and Bertsekas [93]).
- Theorem: Under the preceding assumptions, the set V and a vector $(0, w_0)$ that does not belong to the closure of V can be strongly separated by the function σ, i.e., there exist scalars $c_0 > 0$ and ξ_0 such that for all $c \geq c_0$,

$$w + \frac{1}{c}\sigma(cu) \geq \xi_0 > w_0 \quad \text{for all } (u, w) \in V.$$

Conclusions

- A unifying geometric framework for the analysis of general duality schemes and penalty methods.
- Separation results for general nonconvex sets using concave surfaces.
- Extensions to nonconvex augmenting functions.
- Conditions on the objective and constraint functions that guarantee the key assumption $(0, -1) \notin (\text{epi}(p))^{\infty}$.
- Zero duality gap results potentially useful for the development of dual algorithms for solving nonconvex constrained optimization problems.

References

- Nedić and Ozdaglar, "A Geometric Framework for Nonconvex Optimization Duality using Augmented Lagrangian Functions," to appear in *Journal of Global Optimization*, 2006.
- Nedić and Ozdaglar, "Separation of Nonconvex Sets with General Augmenting Functions," submitted for publication, 2006.
- Nedić, Ozdaglar, and Rubinov, "Abstract Convexity for Nonconvex Optimization Duality," submitted for publication, 2006.