A Unifying Framework for Duality and Minimax

Dimitri Bertsekas, Angelia Nedic, Asuman Ozdaglar

Electrical Engineering and Computer Science Dept.

Massachusetts Institute of Technology

November 19, 2002
Motivation

- **Minimax Theory:** Given $\phi : X \times Z \mapsto \mathbb{R}$, where $X \subset \mathbb{R}^n$, $Z \subset \mathbb{R}^m$, under what conditions do we have
  $$\sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z) ?$$

- **Optimization Duality:** Consider the problem

  minimize $f(x)$

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

  Define the Lagrangian function: $L(x, \mu) = f(x) + \sum_{j=1}^r \mu_j g_j(x)$.

  - Primal problem: $f^* = \inf_{x \in X} \sup_{\mu \geq 0} L(x, \mu)$
  - Dual problem: $q^* = \sup_{\mu \geq 0} q(\mu) = \sup_{\mu \geq 0} \inf_{x \in X} L(x, \mu)$

- All of (convex/concave) minimax theory and duality theory can be developed in terms of simple geometry of convex sets! (Need machinery from convex analysis)
Min Common/Max Crossing Problems

Let $M$ be a nonempty subset of $\mathbb{R}^{n+1}$

- **Min Common Point Problem**: Consider all vectors that are common to $M$ and the $(n+1)$st axis. Find one whose $(n+1)$st component is minimum.

- **Max Crossing Point Problem**: Consider nonvertical hyperplanes that contain $M$ in their “upper” closed halfspace. Find one whose crossing point of the $(n+1)$st axis is maximum.
Weak Duality

- Optimal value of the min common problem:

\[ w^* = \inf_{(0, w) \in M} w. \]

- Focus on hyperplanes with normals \((\mu, 1)\) whose crossing point \(\xi\) satisfies

\[ \xi \leq w + \mu' u, \quad \forall (u, w) \in M. \]

- Maximum crossing level over all hyperplanes with normal \((\mu, 1)\) is

\[ q(\mu) = \inf_{(u, w) \in M} \{w + \mu' u\} \]

Max crossing problem: maximize \(q(\mu)\)
subject to \(\mu \in \mathbb{R}^n\).

- Note that for all \((u, w) \in M\) and \(\mu \in \mathbb{R}^n\),

\[ q(\mu) = \inf_{(u, w) \in M} \{w + \mu' u\} \leq \inf_{(0, w) \in M} w = w^*, \]

\[ \sup_{\mu \geq 0} q(\mu) = q^* \leq w^*. \]
**Strong Duality**

**Question:** Under what conditions do we have $q^* = w^*$ and the supremum in the max crossing problem is attained?
Duality Theorems

- Assume that $w^*$ is finite and that the set

$$\tilde{M} = \{(u, w) \mid \exists \bar{w} \text{ with } \bar{w} \leq w \text{ and } (u, \bar{w}) \in M\}$$

is convex.

- **Min Common Max Crossing Theorem I:** We have $q^* = w^*$ if and only if for every sequence $\{(u_k, w_k)\} \subset M$ with $u_k \to 0$, we have $w^* \leq \lim\inf_{k \to \infty} w_k$. 

![Diagram showing the set $M$ and points $w^*$ and $q^*$]
Attainment of Optimum in Max Crossing Problem

- **Min Common Max Crossing Theorem II:** Assume that the set

  \[ D = \{ u \mid \text{there exists } w \in \mathbb{R} \text{ with } (u, w) \in \bar{M} \} \]

  contains the origin in its relative interior. Then \( q^* = w^* \) and there exists a vector \( \mu \in \mathbb{R}^n \) such that \( q(\mu) = q^* \).

- **Min Common Max Crossing Theorem III:** Involves polyhedral assumptions and guarantees \( q^* = w^* \) as well as attainment of the optimum of the max crossing problem.
Minimax Problems

- Given $\phi : X \times Z \mapsto \mathbb{R}$, where $X \subset \mathbb{R}^n$, $Z \subset \mathbb{R}^m$, under what conditions do we have
  \[ \sup_{z \in Z} \inf_{x \in X} \phi(x, z) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z) \]

- Introduce the perturbation function $p : \mathbb{R}^m \mapsto [-\infty, \infty]$
  \[ p(u) = \inf_{x \in X} \sup_{z \in Z} \{\phi(x, z) - u'z\}, \quad u \in \mathbb{R}^m. \]

- Apply min common/max crossing framework with $M = \text{epi}(p)$.

- Note that $w^* = \inf_{x \in X} \sup_{z \in Z} \phi(x, z)$.
  - Convexity in $x$ implies $M$ is a convex set.
  - Concavity/semicont. in $z$ implies $q^* = \sup_{z \in Z} \inf_{x \in X} \phi(x, z)$.
Minimax Theorems

- Assume that:
  - $X$ and $Z$ are convex and $p(0) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z)$ is finite.
  - For each $z \in Z$, the function $\phi(\cdot, z)$ is convex.
  - For each $x \in X$, the function $-\phi(x, \cdot)$ is closed and convex.

- **Minimax Thm. I**: The minimax equality holds iff the function $p$ is lower semicontinuous at $u = 0$.

- **Minimax Thm. II**: If 0 lies in the relative interior of $\text{dom}(p)$, then the minimax equality holds and the supremum in $\sup_{z \in Z} \inf_{x \in X} \phi(x, z)$ is attained by some $z \in Z$.

  - Proofs by applying min common max crossing theorems to the set
    $$M = \text{epi}(p).$$
Conditions for Attaining the Minimum

- Assume that:
  1. $X$ and $Z$ are convex and $p(0) = \inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty$.
  2. For each $x \in X$, the function $r_x(z) = -\phi(x, z)$ if $z \in Z$, and $\infty$ otherwise, is closed and convex.
  3. For each $z \in Z$, the function $t_z(x) = \phi(x, z)$ if $x \in X$, and $\infty$ otherwise, is closed and convex.
  4. The set of common directions of recession of all the functions $t_z$, $z \in Z$, consists of the zero vector only.

- Then, the minimax equality holds, and the infimum over $X$ is attained at a compact set of points.

- Special cases of Condition 4:
  - $X$ is compact.
  - $\exists$ a scalar $\gamma$ and $\bar{z} \in Z$ s.t. the level set $\{x \in X \mid \phi(x, \bar{z}) \leq \gamma\}$ is nonempty and compact.
Saddle Point Theorem

- Assume that:
  1. $X$ and $Z$ are convex and either $-\infty < \sup_{z \in Z} \inf_{x \in X} \phi(x, z)$, or $\inf_{x \in X} \sup_{z \in Z} \phi(x, z) < \infty$.
  2. For each $x \in X$, the function $r_x(z)$ is closed and convex, and the set of common directions of recession of all the functions $r_x$, $x \in X$, consists of the zero vector only.
  3. For each $z \in Z$, the function $t_z(x)$ is closed and convex, and the set of common directions of recession of all the functions $t_z$, $z \in Z$, consists of the zero vector only.

- Then, the minimax equality holds, and the set of saddle points of $\phi$ is nonempty and compact.

- Special cases:
  - $X$ and $Z$ are compact.
  - $Z$ is compact, and $\exists$ a scalar $\gamma$ and $\bar{z} \in Z$ such that the level set $\{x \in X \mid \phi(x, \bar{z}) \leq \gamma\}$ is nonempty and compact.
Optimization Duality

- Consider the **primal problem** (optimal value $f^*$)
  
  \[
  \begin{align*}
  &\text{minimize} & f(x) \\
  &\text{subject to} & x \in C, \ g_j(x) \leq 0, \ j = 1, \ldots, r,
  \end{align*}
  \]
  
  where $C$ is a convex set, $f$ and $g_j$ are convex over $C$.

- Consider also the **dual problem** (optimal value $q^*$)
  
  \[
  \begin{align*}
  &\text{maximize} & q(\mu) = \inf_{x \in C} \left\{ f(x) + \sum_{j=1}^{r} \mu_j g_j(x) \right\} \\
  &\text{subject to} & \mu \geq 0.
  \end{align*}
  \]

- **Main Question**: Under what conditions do we have
  
  \[ f^* = q^* = q(\mu^*)? \]

- Question can be addressed using min common max crossing framework.
Nonlinear Farkas’ Lemma

• Let \( C \) be convex, and \( f \) and the \( g_j \) be convex functions. Assume that

\[
f(x) \geq 0, \quad \forall \ x \in F = \{x \in C \mid g(x) \leq 0\},
\]

and one of the following conditions holds:

1. \( 0 \) is in the relative interior of the set

\[
D = \{u \mid g(x) \leq u \text{ for some } x \in C\}.
\]

2. The functions \( g_j, \ j = 1, \ldots, r \) are affine, and \( F \cap \text{ri}(C) \neq \emptyset \).

• Then, there exist scalars \( \mu_j^* \geq 0, \ j = 1, \ldots, r \), such that

\[
f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x^*) \geq 0, \quad \forall \ x \in C.
\]

– Reduces to Farkas’ Lemma if \( C = \mathbb{R}^n \), and \( f, g_j \) linear.
– Proofs by applying min common max crossing theorems to

\[
M = \{(u,w) \mid \text{there is } x \in C \text{ s.t. } g(x) \leq u, \ f(x) \leq w\}.
\]
Application to Convex Programming

- Consider the problem (optimal value $f^*$, assumed finite)

\[
\text{minimize} \quad f(x) \\
\text{subject to} \quad x \in C, \quad g_j(x) \leq 0, \quad j = 1, \ldots, r,
\]

where $C$ is convex, $f$ and $g_j$ are convex over $C$.

- Apply Farkas’ Lemma. There exist $\mu_j^* \geq 0$ such that

\[
f^* \leq f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x), \quad \forall \; x \in C.
\]

Since $F \subset C$ and $\mu_j^* g_j(x) \leq 0$ for all $x \in F$,

\[
f^* \leq \inf_{x \in F} \left\{ f(x) + \sum_{j=1}^{r} \mu_j^* g_j(x) \right\} \leq \inf_{x \in F} f(x) = f^*.
\]

Thus equality holds throughout above, and we have

\[
f^* = \inf_{x \in C} \{ f(x) + \mu^* g(x) \} = q(\mu^*).
\]
Reference

Convex Analysis and Optimization,
Dimitri Bertsekas, with Angelia Nedic and Asuman Ozdaglar.

- To be published January 2003.
  - Min Common/Max Crossing Duality
  - Existence of Solutions and Strong Duality
  - Pseudonormality and Lagrange Multipliers
  - Incremental Subgradient Methods