Advanced Optimization

(10-801: CMU)

Lecture 6
Duality, Optimality

03 Feb, 2014

Suvrit Sra

Primal problem

Let $h_i : \mathbb{R}^n \to \mathbb{R}$ $(1 \le i \le m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t. } & h_i(x) \le 0, \quad 1 \le i \le m, \\ & x \in \{\operatorname{dom} f \cap \operatorname{dom} h_1 \cdots \cap \operatorname{dom} h_m\}. \end{aligned} \tag{P}$$

Primal problem

Let $h_i : \mathbb{R}^n \to \mathbb{R}$ $(1 \leq i \leq m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t. } & h_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & x \in \{\operatorname{dom} f \cap \operatorname{dom} h_1 \cdots \cap \operatorname{dom} h_m\}. \end{aligned} \tag{P}$$

Domain: The set $\mathcal{X} := \{\operatorname{dom} f \cap \operatorname{dom} h_1 \cdots \cap \operatorname{dom} h_m\}$

- We call (P) the **primal problem**
- The variable x is the **primal variable**

The reader will find no figures in this work. The methods which I set forth do not require either constructions or geometrical or mechanical reasonings: but only algebraic operations, subject to a regular and uniform rule of procedure.

—Joseph-Louis Lagrange
Preface to *Mécanique Analytique*

Lagrangian

To primal, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m_+ \to (-\infty, \infty]$,

$$\mathcal{L}(x, \lambda) := f(x) + \sum_{i=1}^{m} \lambda_i h_i(x).$$

Lagrangian

To primal, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m_+ \to (-\infty, \infty]$,

$$\mathcal{L}(x, \lambda) := f(x) + \sum_{i=1}^{m} \lambda_i h_i(x).$$

♠ Variables $\lambda \in \mathbb{R}^m_+$ called **Lagrange multipliers**

Lagrangian

To primal, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m_+ \to (-\infty, \infty]$,

$$\mathcal{L}(x, \lambda) := f(x) + \sum\nolimits_{i=1}^{m} \lambda_i h_i(x).$$

- Variables $\lambda \in \mathbb{R}^m_+$ called **Lagrange multipliers**
- Suppose x is feasible, and $\lambda \geq 0$. Then, we get the lower-bound:

$$f(x) \geq \mathcal{L}(x, \lambda) \qquad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+.$$

- In other words,

$$\sup_{\lambda \in \mathbb{R}^m_+} \mathcal{L}(x, \lambda) = \begin{cases} f(x), & \text{if } x \text{ feasible,} \\ +\infty & \text{otherwise.} \end{cases}$$

Lagrangian

Since, $f(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}_{+}^{m}$, **primal optimal**

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x, \lambda).$$

Lagrangian

Since, $f(x) \geq \mathcal{L}(x,\lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+$, **primal optimal**

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x,\lambda).$$

Proof:

- If x is not feasible, then some $h_i(x) > 0$

Lagrangian

Since, $f(x) \geq \mathcal{L}(x,\lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+$, **primal optimal**

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x,\lambda).$$

Proof:

- If x is not feasible, then some $h_i(x) > 0$
- In this case, inner $\sup$ is $+\infty$, so claim true by definition

Lagrangian

Since, $f(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+$, **primal optimal**

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x, \lambda).$$

Proof:

- If x is not feasible, then some $h_i(x) > 0$
- In this case, inner $\sup$ is $+\infty$, so claim true by definition
- If x is feasible, each $h_i(x) \leq 0$, so $\sup_\lambda \sum_i \lambda_i h_i(x) = 0$

Dual value

Primal value $\in [-\infty, +\infty]$

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x, \lambda).$$

Dual value

Primal value $\in [-\infty, +\infty]$

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x, \lambda).$$

Dual value $\in [-\infty, +\infty]$

$$d^* = \sup_{\lambda \geq 0} \inf_{x \in \mathcal{X}} \quad \mathcal{L}(x, \lambda).$$

Dual value

Primal value $\in [-\infty, +\infty]$

$$p^* = \inf_{x \in \mathcal{X}} \sup_{\lambda \geq 0} \quad \mathcal{L}(x, \lambda).$$

Dual value $\in [-\infty, +\infty]$

$$d^* = \sup_{\lambda \geq 0} \inf_{x \in \mathcal{X}} \quad \mathcal{L}(x, \lambda).$$

Dual function

$$g(\lambda) := \inf_{x \in \mathcal{X}} \mathcal{L}(x, \lambda).$$

Weak duality

- g is pointwise infimum of affine functions of λ
- Thus, g is concave (and may also take value $-\infty$)
- Maximizing concave $\implies$ minimizing convex $\implies$

Theorem (Weak duality.) $p^* \geq d^*$.

Proof:

1. $f(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}$
2. $\forall x \in \mathcal{X}, \quad f(x) \geq \inf_{x'} \mathcal{L}(x', \lambda) = g(\lambda)$
3. Minimize over x on lhs to obtain

$$\forall\ \lambda \in \mathbb{R}^m_+ \qquad p^* \geq g(\lambda).$$

4. Thus, taking sup over $\lambda \in \mathbb{R}^m_+$ we obtain $p^* \geq d^*$.

Exercise

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t. } & h_i(x) \leq 0, \quad i = 1, \dots, m, \\ & k_i(x) = 0, \quad i = 1, \dots, p. \end{aligned}$$

Show that we get the Lagrangian dual

$$g : \mathbb{R}^m_+ \times \mathbb{R}^p : (\lambda, \nu) \mapsto \inf_x \quad \mathcal{L}(x, \lambda, \nu),$$

Lagrange variable ν corresponds to the equality constraints.
Prove that $p^* \geq \sup_{\lambda \geq 0, \nu \in \mathbb{R}^p} \; g(\lambda, \nu) = d^*$.

Strong duality

Duality gap

$$p^* - d^*$$

Duality gap

$$p^* - d^*$$

Strong duality if duality gap is zero: $p^* = d^*$

Duality gap

$$p^* - d^*$$

Strong duality if duality gap is zero: $p^* = d^*$

Several **sufficient** conditions known!

Duality gap

$$p^* - d^*$$

Strong duality if duality gap is zero: $p^* = d^*$

Several **sufficient** conditions known!

"Easy" necessary and sufficient conditions: unknown

General duality gap theorem

Theorem Let $v : \mathbb{R}^m \to \mathbb{R}$ be the *primal value function*

$$v(u) := \inf \{ f(x) \mid h_i(x) \leq u_i, \ 1 \leq i \leq m \} .$$

The following relations hold:

1. $p^* = v(0)$
2. $v^*(-\lambda) = \begin{cases} -g(\lambda) & \lambda \geq 0 \\ +\infty & \text{otherwise.} \end{cases}$
3. $d^* = v^{**}(0)$

So if $v(0) = v^{**}(0)$ we have strong duality

Conditions such as Slater's ensure $\partial v(0) \neq \emptyset$, which ensures v is finite and lsc at 0, whereby $v(0) = v^{**}(0)$ holds.

Slater's sufficient conditions

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

Slater's sufficient conditions

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t. } & h_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

Constraint qualification: There exists $x \in \operatorname{ri} \mathcal{X}$ s.t.

$$h_i(x) < 0, \qquad Ax = b.$$

That is, there is a **strictly feasible** point.

Slater's sufficient conditions

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.}\ & h_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & Ax = b. \end{aligned}$$

Constraint qualification: There exists $x \in \operatorname{ri} \mathcal{X}$ s.t.

$$h_i(x) < 0, \qquad Ax = b.$$

That is, there is a **strictly feasible** point.

Theorem Let the primal problem be convex. If there is a feasible point such that is strictly feasible for the non-affine constraints (and merely feasible for affine, linear ones), then strong duality holds. Moreover, in this case, the dual optimal is attained (i.e., $\partial v(0) \neq \emptyset$).

See BV §5.3.2 for a proof; (above, v is the primal value function)

Example with positive duality-gap

$$\min_{x,y} e^{-x} \quad x^2/y \leq 0,$$

over the domain $\mathcal{X} = \{(x, y) \mid y > 0\}$.

Example with positive duality-gap

$$\min_{x,y} e^{-x} \quad x^2/y \leq 0,$$

over the domain $\mathcal{X} = \{(x, y) \mid y > 0\}$.
Clearly, only feasible $x = 0$. So $p^* = 1$

Example with positive duality-gap

$$\min_{x,y} e^{-x} \quad x^2/y \leq 0,$$

over the domain $\mathcal{X} = \{(x, y) \mid y > 0\}$.
Clearly, only feasible $x = 0$. So $p^* = 1$

$$\mathcal{L}(x, y, \lambda) = e^{-x} + \lambda x^2/y,$$

so dual function is

$$g(\lambda) = \inf_{x,y>0} e^{-x} + \lambda x^2 y = \begin{cases} 0 & \lambda \geq 0 \\ -\infty & \lambda < 0. \end{cases}$$

Example with positive duality-gap

$$\min_{x,y} e^{-x} \quad x^2/y \leq 0,$$

over the domain $\mathcal{X} = \{(x, y) \mid y > 0\}$.
Clearly, only feasible $x = 0$. So $p^* = 1$

$$\mathcal{L}(x, y, \lambda) = e^{-x} + \lambda x^2/y,$$

so dual function is

$$g(\lambda) = \inf_{x,y>0} e^{-x} + \lambda x^2 y = \begin{cases} 0 & \lambda \geq 0 \\ -\infty & \lambda < 0. \end{cases}$$

Dual problem

$$d^* = \max_{\lambda} 0 \qquad \text{s.t. } \lambda \geq 0.$$

Example with positive duality-gap

$$\min_{x,y} e^{-x} \quad x^2/y \leq 0,$$

over the domain $\mathcal{X} = \{(x, y) \mid y > 0\}$.
Clearly, only feasible $x = 0$. So $p^* = 1$

$$\mathcal{L}(x, y, \lambda) = e^{-x} + \lambda x^2/y,$$

so dual function is

$$g(\lambda) = \inf_{x,y>0} e^{-x} + \lambda x^2 y = \begin{cases} 0 & \lambda \geq 0 \\ -\infty & \lambda < 0. \end{cases}$$

Dual problem

$$d^* = \max_{\lambda} 0 \qquad \text{s.t. } \lambda \geq 0.$$

Thus, $d^* = 0$, and gap is $p^* - d^* = 1$.

Example with positive duality-gap

$$\min_{x,y} e^{-x} \quad x^2/y \leq 0,$$

over the domain $\mathcal{X} = \{(x, y) \mid y > 0\}$.
Clearly, only feasible $x = 0$. So $p^* = 1$

$$\mathcal{L}(x, y, \lambda) = e^{-x} + \lambda x^2/y,$$

so dual function is

$$g(\lambda) = \inf_{x,y>0} e^{-x} + \lambda x^2 y = \begin{cases} 0 & \lambda \geq 0 \\ -\infty & \lambda < 0. \end{cases}$$

Dual problem

$$d^* = \max_{\lambda} 0 \qquad \text{s.t. } \lambda \geq 0.$$

Thus, $d^* = 0$, and gap is $p^* - d^* = 1$.
Here, we had no strictly feasible solution.

Fenchel and Lagrangian duality

$$\inf_{x \in \mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Fenchel and Lagrangian duality

$$\inf_{x \in \mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{u \in \mathcal{Y}} \quad f^*(-A^T u) + r^*(u).$$

Fenchel and Lagrangian duality

$$\inf_{x \in \mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{u \in \mathcal{Y}} \quad f^*(-A^T u) + r^*(u).$$

- Introduce new variable $z = Ax$

$$\inf_{x \in \mathcal{X}, z \in \mathcal{Y}} \quad f(x) + r(z), \qquad \text{s.t.} \quad z = Ax.$$

Fenchel and Lagrangian duality

$$\inf_{x \in \mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{u \in \mathcal{Y}} \quad f^*(-A^T u) + r^*(u).$$

- Introduce new variable $z = Ax$

$$\inf_{x \in \mathcal{X}, z \in \mathcal{Y}} \quad f(x) + r(z), \qquad \text{s.t.} \quad z = Ax.$$

- The (partial)-Lagrangian is

$$L(x, z; u) := f(x) + r(z) + u^T(Ax - z), \quad x \in \mathcal{X}, z \in \mathcal{Y};$$

Fenchel and Lagrangian duality

$$\inf_{x \in \mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{u \in \mathcal{Y}} \quad f^*(-A^T u) + r^*(u).$$

- Introduce new variable $z = Ax$

$$\inf_{x \in \mathcal{X}, z \in \mathcal{Y}} \quad f(x) + r(z), \qquad \text{s.t.} \quad z = Ax.$$

- The (partial)-Lagrangian is

$$L(x, z; u) := f(x) + r(z) + u^T(Ax - z), \quad x \in \mathcal{X}, z \in \mathcal{Y};$$

- Associated dual function

$$g(u) := \inf_{x \in \mathcal{X}, z \in \mathcal{Y}} L(x, z; u).$$

Fenchel and Lagrangian duality

$$\inf_{x\in\mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{y\in\mathcal{Y}} \quad f^*(-A^T y) + r^*(y).$$

The infimum above can be rearranged as follows

$$g(y) \quad = \quad \inf_{x\in\mathcal{X}} \ f(x) + y^T Ax + \inf_{z\in\mathcal{Y}} \ r(z) - y^T z$$

Fenchel and Lagrangian duality

$$\inf_{x\in\mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{y\in\mathcal{Y}} \quad f^*(-A^T y) + r^*(y).$$

The infimum above can be rearranged as follows

$$\begin{aligned} g(y) &= \inf_{x\in\mathcal{X}} f(x) + y^T Ax + \inf_{z\in\mathcal{Y}} r(z) - y^T z \\ &= -\sup_{x\in\mathcal{X}} \left\{ -x^T A^T y - f(x) \right\} - \sup_{z\in\mathcal{Y}} \left\{ z^T y - r(z) \right\} \end{aligned}$$

Fenchel and Lagrangian duality

$$\inf_{x\in\mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{y\in\mathcal{Y}} \quad f^*(-A^T y) + r^*(y).$$

The infimum above can be rearranged as follows

$$\begin{aligned} g(y) &= \inf_{x\in\mathcal{X}} f(x) + y^T Ax + \inf_{z\in\mathcal{Y}} r(z) - y^T z \\ &= -\sup_{x\in\mathcal{X}} \left\{-x^T A^T y - f(x)\right\} - \sup_{z\in\mathcal{Y}} \left\{z^T y - r(z)\right\} \\ &= -f^*(-A^T y) - r^*(y) \quad \text{s.t. } y \in \mathcal{Y}. \end{aligned}$$

Fenchel and Lagrangian duality

$$\inf_{x\in\mathcal{X}} \quad f(x) + r(Ax) \quad \text{s.t.} \quad Ax \in \mathcal{Y}.$$

Dual problem

$$\inf_{y\in\mathcal{Y}} \quad f^*(-A^T y) + r^*(y).$$

The infimum above can be rearranged as follows

$$\begin{aligned} g(y) &= \inf_{x\in\mathcal{X}} f(x) + y^T Ax + \inf_{z\in\mathcal{Y}} r(z) - y^T z \\ &= -\sup_{x\in\mathcal{X}} \{-x^T A^T y - f(x)\} - \sup_{z\in\mathcal{Y}} \{z^T y - r(z)\} \\ &= -f^*(-A^T y) - r^*(y) \quad \text{s.t. } y \in \mathcal{Y}. \end{aligned}$$

Dual problem computes $\sup_{u\in\mathcal{Y}} g(u)$; so equivalently,

$$\inf_{y\in\mathcal{Y}} \quad f^*(-A^T y) + r^*(y).$$

Strong duality

$$\inf_x \{f(x) + r(Ax)\} = \sup_y \left\{-f^*(-A^T y) + r^*(y)\right\}$$

if either of the following conditions holds:

1. $\exists x \in \mathrm{ri}(\mathrm{dom}\, f)$ such that $Ax \in \mathrm{ri}(\mathrm{dom}\, r)$
2. $\exists y \in \mathrm{ri}(\mathrm{dom}\, r^*)$ such that $A^T y \in \mathrm{ri}(\mathrm{dom}\, f^*)$

Strong duality

$$\inf_x \{f(x) + r(Ax)\} = \sup_y \left\{-f^*(-A^T y) + r^*(y)\right\}$$

if either of the following conditions holds:

1. $\exists x \in \mathrm{ri}(\mathrm{dom}\, f)$ such that $Ax \in \mathrm{ri}(\mathrm{dom}\, r)$
2. $\exists y \in \mathrm{ri}(\mathrm{dom}\, r^*)$ such that $A^T y \in \mathrm{ri}(\mathrm{dom}\, f^*)$

1. Condition 1 ensures 'sup' attained at some y
2. Condition 2 ensures 'inf' attained at some x

Example: norm regularized problems

$$\min \quad f(x) + \|Ax\|$$

Example: norm regularized problems

$$\min \quad f(x) + \|Ax\|$$

Dual problem

$$\min_y \quad f^*(-A^T y) \quad \text{s.t. } \|y\|_* \le 1.$$

Example: norm regularized problems

$$\min \quad f(x) + \|Ax\|$$

Dual problem

$$\min_y \quad f^*(-A^T y) \quad \text{s.t. } \|y\|_* \leq 1.$$

Say $\|\bar{y}\|_* < 1$, such that $A^T\bar{y} \in \operatorname{ri}(\operatorname{dom} f^*)$, then we have strong duality—for instance if $0 \in \operatorname{ri}(\operatorname{dom} f^*)$

Example: variable splitting

$$\min \quad f(x) + h(x)$$

Exercise: Fill in the details for the following steps

$$\min_{x,z} \quad f(x) + h(z) \quad \text{s.t.} \quad x = z$$

Example: variable splitting

$$\min \quad f(x) + h(x)$$

Exercise: Fill in the details for the following steps

$$\min_{x,z} \quad f(x) + h(z) \quad \text{s.t.} \quad x = z$$

$$L(x, z, \nu) = f(x) + h(z) + \nu^T(x - z)$$

Example: variable splitting

$$\min \quad f(x) + h(x)$$

Exercise: Fill in the details for the following steps

$$\min_{x,z} \quad f(x) + h(z) \quad \text{s.t.} \quad x = z$$

$$L(x, z, \nu) = f(x) + h(z) + \nu^T(x - z)$$

$$g(\nu) = \inf_{x,z} L(x, z, \nu)$$

Minimax

Minimax problems

- Minimax theory treats problems involving a combination of minimization and maximization

Minimax problems

- Minimax theory treats problems involving a combination of minimization and maximization
- Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary nonempty sets

Minimax problems

- Minimax theory treats problems involving a combination of minimization and maximization
- Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary nonempty sets
- Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$

Minimax problems

- Minimax theory treats problems involving a combination of minimization and maximization
- Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary nonempty sets
- Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$
- **inf** over $x \in \mathcal{X}$, followed by **sup** over $y \in \mathcal{Y}$

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y)$$

Minimax problems

- Minimax theory treats problems involving a combination of minimization and maximization
- Let $\mathcal{X}$ and $\mathcal{Y}$ be arbitrary nonempty sets
- Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$
- **inf** over $x \in \mathcal{X}$, followed by **sup** over $y \in \mathcal{Y}$

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y)$$

- **sup** over $y \in \mathcal{Y}$, followed by **inf** over $x \in \mathcal{X}$

$$\inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

When are "inf sup" and "sup inf" equal?

Weak minimax

Theorem Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Then,

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y) \quad \leq \quad \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

Weak minimax

Theorem Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Then,

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y) \quad \leq \quad \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

Define:

$$f(x) := \sup_{y \in \mathcal{Y}} \phi(x, y) \qquad g(y) := \inf_{x \in \mathcal{X}} \phi(x, y).$$

Weak minimax

Theorem Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Then,

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y) \quad \leq \quad \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

Define:

$$f(x) := \sup_{y \in \mathcal{Y}} \phi(x, y) \qquad g(y) := \inf_{x \in \mathcal{X}} \phi(x, y).$$

$$f(x) \quad \geq \quad \phi(x, y) \quad \geq \quad g(y), \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}$$

Weak minimax

Theorem Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Then,

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x,y) \quad \leq \quad \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x,y)$$

Define:

$$f(x) := \sup_{y \in \mathcal{Y}} \phi(x,y) \qquad g(y) := \inf_{x \in \mathcal{X}} \phi(x,y).$$

$$\begin{aligned} f(x) &\geq \phi(x,y) \geq g(y), \quad \forall x \in \mathcal{X}, y \in \mathcal{Y} \\ \inf_x f(x) &= \inf_x \sup_y \phi(x,y) \geq \inf_x \phi(x,y) \geq g(y) \end{aligned}$$

Weak minimax

Theorem Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Then,

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y) \quad \leq \quad \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

Define:

$$f(x) := \sup_{y \in \mathcal{Y}} \phi(x, y) \qquad g(y) := \inf_{x \in \mathcal{X}} \phi(x, y).$$

$$\begin{aligned} f(x) &\geq \phi(x, y) \geq g(y), \quad \forall x \in \mathcal{X}, y \in \mathcal{Y} \\ \inf_x f(x) &= \inf_x \sup_y \phi(x, y) \geq \inf_x \phi(x, y) \geq g(y) \\ \inf_x \sup_y \phi(x, y) &\geq \sup_y g(y) = \sup_y \inf_x \phi(x, y). \end{aligned}$$

Weak minimax

Theorem Let $\phi : \mathcal{X} \times \mathcal{Y} \to \mathbb{R} \cup \{\pm\infty\}$ be any function. Then,

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y) \quad \leq \quad \inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

Define:

$$f(x) := \sup_{y \in \mathcal{Y}} \phi(x, y) \qquad g(y) := \inf_{x \in \mathcal{X}} \phi(x, y).$$

$$\begin{aligned} f(x) \quad &\geq \quad \phi(x, y) \quad \geq \quad g(y), \quad \forall x \in \mathcal{X}, y \in \mathcal{Y} \\ \inf_x f(x) \quad &= \quad \inf_x \sup_y \phi(x, y) \geq \inf_x \phi(x, y) \geq g(y) \\ \inf_x \sup_y \phi(x, y) \quad &\geq \quad \sup_y g(y) = \sup_y \inf_x \phi(x, y). \end{aligned}$$

Exercise: Derive weak duality from above minimax inequality.
Hint: Use $\phi = \mathcal{L}$ (Lagrangian) for suitably chosen y.

Strong minimax

- If "inf sup" equals "sup inf", common value called **saddle-value**
- Value exists if there is a **saddle-point**, i.e., pair (x^*, y^*)

$$\inf_{x \in \mathcal{X}} f(x) = \sup_{y \in \mathcal{Y}} g(y)$$

- That is

$$f(x^*) = \phi(x^*, y^*) = g(y^*)$$

Strong minimax

Def. Let ϕ be as before. A point (x^*, y^*) is a saddle-point of ϕ **if and only if** the infimum in the expression

$$\inf_{x \in \mathcal{X}} \sup_{y \in \mathcal{Y}} \phi(x, y)$$

is **attained** at x^*, and the supremum in the expression

$$\sup_{y \in \mathcal{Y}} \inf_{x \in \mathcal{X}} \phi(x, y)$$

is **attained** at y^*, and these two extrema are equal.

Strong minimax

Def. Let ϕ be as before. A point (x^*, y^*) is a saddle-point of ϕ **if and only if** the infimum in the expression

$$\inf_{x\in\mathcal{X}} \sup_{y\in\mathcal{Y}} \phi(x,y)$$

is **attained** at x^*, and the supremum in the expression

$$\sup_{y\in\mathcal{Y}} \inf_{x\in\mathcal{X}} \phi(x,y)$$

is **attained** at y^*, and these two extrema are equal.

$$x^* \in \operatorname*{argmin}_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \phi(x,y) \qquad y^* \in \operatorname{argmax}_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} \phi(x,y).$$

Strong minimax

- ♠ Classes of problems "dual" to each other can be generated by studying classes of functions ϕ,

Strong minimax

- Classes of problems "dual" to each other can be generated by studying classes of functions ϕ,
- More interesting question: Starting from the primal problem over $\mathcal{X}$, how to introduce a space $\mathcal{Y}$ and a "useful" function ϕ on $\mathcal{X} \times \mathcal{Y}$ so that we have a saddle-point?

Strong minimax

- ♠ Classes of problems "dual" to each other can be generated by studying classes of functions ϕ,
- ♠ More interesting question: Starting from the primal problem over $\mathcal{X}$, how to introduce a space $\mathcal{Y}$ and a "useful" function ϕ on $\mathcal{X} \times \mathcal{Y}$ so that we have a saddle-point?

Sufficient conditions for saddle-point

- ▶ Function ϕ is continuous, and
- ▶ It is convex-concave ($\phi(\cdot, y)$ convex for every $y \in \mathcal{Y}$, and $\phi(x, \cdot)$ concave for every $x \in \mathcal{X}$), and
- ▶ Both $\mathcal{X}$ and $\mathcal{Y}$ are convex; one of them is compact.

Example: Lasso-like problem

$$p^* := \min_x \quad \|Ax - b\|_2 + \lambda \|x\|_1.$$

Example: Lasso-like problem

$$p^* := \min_x \quad \|Ax - b\|_2 + \lambda \|x\|_1.$$

$$\|x\|_1 = \max \left\{ x^T v \mid \|v\|_\infty \leq 1 \right\}$$
$$\|x\|_2 = \max \left\{ x^T u \mid \|u\|_2 \leq 1 \right\}.$$

Example: Lasso-like problem

$$p^* := \min_x \quad \|Ax - b\|_2 + \lambda \|x\|_1.$$

$$\|x\|_1 = \max \left\{ x^T v \mid \|v\|_\infty \leq 1 \right\}$$
$$\|x\|_2 = \max \left\{ x^T u \mid \|u\|_2 \leq 1 \right\}.$$

Saddle-point formulation

$$p^* \quad = \quad \min_x \max_{u,v} \left\{ u^T (b - Ax) + v^T x \mid \|u\|_2 \leq 1, \quad \|v\|_\infty \leq \lambda \right\}$$

Example: Lasso-like problem

$$p^* := \min_x \quad \|Ax - b\|_2 + \lambda\|x\|_1.$$

$$\|x\|_1 = \max\left\{x^T v \mid \|v\|_\infty \le 1\right\}$$
$$\|x\|_2 = \max\left\{x^T u \mid \|u\|_2 \le 1\right\}.$$

Saddle-point formulation

$$\begin{aligned} p^* &= \min_x \max_{u,v} \left\{u^T(b - Ax) + v^T x \mid \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda\right\} \\ &= \max_{u,v} \min_x \left\{u^T(b - Ax) + x^T v \mid \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda\right\} \end{aligned}$$

Example: Lasso-like problem

$$p^* := \min_x \quad \|Ax - b\|_2 + \lambda \|x\|_1.$$

$$\|x\|_1 = \max \left\{ x^T v \mid \|v\|_\infty \le 1 \right\}$$
$$\|x\|_2 = \max \left\{ x^T u \mid \|u\|_2 \le 1 \right\}.$$

Saddle-point formulation

$$\begin{aligned} p^* &= \min_x \max_{u,v} \left\{ u^T(b - Ax) + v^T x \mid \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda \right\} \\ &= \max_{u,v} \min_x \left\{ u^T(b - Ax) + x^T v \mid \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda \right\} \\ &= \max_{u,v} u^T b \qquad A^T u = v, \ \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda \end{aligned}$$

Example: Lasso-like problem

$$p^* := \min_x \quad \|Ax - b\|_2 + \lambda \|x\|_1.$$

$$\|x\|_1 = \max \left\{ x^T v \mid \|v\|_\infty \le 1 \right\}$$
$$\|x\|_2 = \max \left\{ x^T u \mid \|u\|_2 \le 1 \right\}.$$

Saddle-point formulation

$$\begin{aligned}
p^* &= \min_x \max_{u,v} \left\{ u^T(b - Ax) + v^T x \mid \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda \right\} \\
&= \max_{u,v} \min_x \left\{ u^T(b - Ax) + x^T v \mid \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda \right\} \\
&= \max_{u,v} u^T b \qquad A^T u = v, \ \|u\|_2 \le 1, \quad \|v\|_\infty \le \lambda \\
&= \max_u u^T b \qquad \|u\|_2 \le 1, \quad \|A^T v\|_\infty \le \lambda.
\end{aligned}$$

Optimality via minimax

$$x^* \in \operatorname*{argmin}_{x\in\mathcal{X}} \max_{y\in\mathcal{Y}} \phi(x,y) \qquad y^* \in \operatorname{argmax}_{y\in\mathcal{Y}} \min_{x\in\mathcal{X}} \phi(x,y).$$

Point (x^*, y^*) is a saddle-point if and only if

$$0 \in \partial\phi(x^*, y^*) = \partial_x\phi(x^*, y^*) \times \partial_y\phi(x^*, y^*)$$

Optimality via minimax

$$x^* \in \operatorname*{argmin}_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \phi(x, y) \qquad y^* \in \operatorname{argmax}_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \phi(x, y).$$

Point (x^*, y^*) is a saddle-point if and only if

$$0 \in \partial\phi(x^*, y^*) = \partial_x \phi(x^*, y^*) \times \partial_y \phi(x^*, y^*)$$

When ϕ is of "convex-concave" form, yields KKT conditions.