Convex Optimization

(EE227A: UC Berkeley)

Lecture 10

Duality, strong-duality

21 Feb, 2013

Suvrit Sra

Organizational

- ♠ Homework grading mechanism
- ♠ List of projects to be out soon
- ♠ Project timeline
  - ♡ Team lists due by end of Feb
  - ♡ Initial proposal by 14th March
  - ♡ Project midpoint review: 16th April
  - ♡ Project final paper, presentations: Finals week
- ♠ Midterm maybe around 21st March (in class, 3 hours, TBD)
- ♠ I hope to write lecture notes beginning March
- ♠ Email me any concerns, doubts, questions, feedback

Weak duality
Recap

Primal problem

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \leq i \leq m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \; & f_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & x \in \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\} \,. \end{aligned} \tag{P}$$

Primal problem

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \le i \le m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad 1 \le i \le m, \\ & x \in \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}. \end{aligned} \tag{P}$$

Def. **Domain:** The set $\mathcal{D} := \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}$

Primal problem

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \le i \le m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad 1 \le i \le m, \\ & x \in \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}. \end{aligned} \tag{P}$$

Def. Domain: The set $\mathcal{D} := \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}$

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

Primal problem

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \leq i \leq m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & x \in \{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f_1 \cdots \cap \mathrm{dom}\, f_m\}\,. \end{aligned} \tag{P}$$

Def. Domain: The set $\mathcal{D} := \{\mathrm{dom}\, f_0 \cap \mathrm{dom}\, f_1 \cdots \cap \mathrm{dom}\, f_m\}$

- We call (P) the **primal problem**
- The variable x is the **primal variable**
- We will attach to (P) a **dual problem**

Primal problem

Let $f_i : \mathbb{R}^n \to \mathbb{R}$ $(0 \le i \le m)$. Generic **nonlinear program**

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le 0, \quad 1 \le i \le m, \\ & x \in \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}. \end{aligned} \tag{P}$$

Def. Domain: The set $\mathcal{D} := \{\operatorname{dom} f_0 \cap \operatorname{dom} f_1 \cdots \cap \operatorname{dom} f_m\}$

- We call (P) the **primal problem**
- The variable x is the **primal variable**
- We will attach to (P) a **dual problem**
- In our initial derivation: no restriction to convexity.

Lagrangian

To the primal problem, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x).$$

Lagrangian

To the primal problem, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x).$$

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

Lagrangian

To the primal problem, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x).$$

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

♠ Suppose x is feasible, and $\lambda \geq 0$. Then, we get the lower-bound:

$$f_0(x) \geq \mathcal{L}(x, \lambda) \qquad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+.$$

Lagrangian

To the primal problem, associate **Lagrangian** $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$,

$$\mathcal{L}(x, \lambda) := f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x).$$

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

$$f_0(x) \geq \mathcal{L}(x, \lambda) \qquad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+.$$

- ♠ Lagrangian helps write problem in **unconstrained form**

Lagrangian

Claim: Since, $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+$, primal optimal

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

Lagrangian

Claim: Since, $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+$, 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 $f_i(x) > 0$

Lagrangian

Claim: Since, $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}^m_+$, 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 $f_i(x) > 0$
- In this case, inner $\sup$ is $+\infty$, so claim true by definition

Lagrangian

Claim: Since, $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}, \quad \lambda \in \mathbb{R}_+^m$, 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 $f_i(x) > 0$
- ♠ In this case, inner $\sup$ is $+\infty$, so claim true by definition
- ♠ If x is feasible, each $f_i(x) \leq 0$, so $\sup_\lambda \sum_i \lambda_i f_i(x) = 0$

Lagrange dual function

Def. We define the **Lagrangian dual** as

$$g(\lambda) := \inf_x \quad \mathcal{L}(x, \lambda).$$

Lagrange dual function

Def. We define the **Lagrangian dual** as

$$g(\lambda) := \inf_x \quad \mathcal{L}(x, \lambda).$$

Observations:

- g is pointwise inf of affine functions of λ
- Thus, g is concave; it may take value $-\infty$

Lagrange dual function

Def. We define the **Lagrangian dual** as

$$g(\lambda) := \inf_x \quad \mathcal{L}(x, \lambda).$$

Observations:

- g is pointwise inf of affine functions of λ
- Thus, g is concave; it may take value $-\infty$
- Recall: $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}$; thus

Lagrange dual function

Def. We define the **Lagrangian dual** as

$$g(\lambda) := \inf_x \quad \mathcal{L}(x, \lambda).$$

Observations:

- g is pointwise inf of affine functions of λ
- Thus, g is concave; it may take value $-\infty$
- Recall: $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}$; thus
- $\forall x \in \mathcal{X}, \quad f_0(x) \geq \inf_{x'} \mathcal{L}(x', \lambda) = g(\lambda)$

Lagrange dual function

Def. We define the **Lagrangian dual** as

$$g(\lambda) := \inf_x \quad \mathcal{L}(x, \lambda).$$

Observations:

- g is pointwise inf of affine functions of λ
- Thus, g is concave; it may take value $-\infty$
- Recall: $f_0(x) \geq \mathcal{L}(x, \lambda) \quad \forall x \in \mathcal{X}$; thus
- $\forall x \in \mathcal{X}, \quad f_0(x) \geq \inf_{x'} \mathcal{L}(x', \lambda) = g(\lambda)$
- Now minimize over x on lhs, to obtain

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

Lagrange dual problem

$$\sup_{\lambda} g(\lambda) \qquad \text{s.t. } \lambda \geq 0.$$

Lagrange dual problem

$$\sup_{\lambda} g(\lambda) \qquad \text{s.t. } \lambda \geq 0.$$

- **dual feasible:** if $\lambda \geq 0$ and $g(\lambda) > -\infty$
- **dual optimal:** λ^* if $\sup$ is achieved

Lagrange dual problem

$$\sup_{\lambda} g(\lambda) \qquad \text{s.t. } \lambda \geq 0.$$

- **dual feasible:** if $\lambda \geq 0$ and $g(\lambda) > -\infty$
- **dual optimal:** λ^* if $\sup$ is achieved
- Lagrange dual is always concave, regardless of original

Weak duality

Def. Denote **dual optimal value** by d^*, i.e.,

$$d^* := \sup_{\lambda \geq 0} \quad g(\lambda).$$

Weak duality

Def. Denote **dual optimal value** by d^*, i.e.,

$$d^* := \sup_{\lambda \geq 0} \quad g(\lambda).$$

Theorem (Weak-duality): For problem (P), we have $p^* \geq d^*$.

Weak duality

Def. Denote **dual optimal value** by d^*, i.e.,

$$d^* := \sup_{\lambda \geq 0} \quad g(\lambda).$$

Theorem (Weak-duality): For problem (P), we have $p^* \geq d^*$.

Proof: We showed that for all $\lambda \in \mathbb{R}_+^m$, $p^* \geq g(\lambda)$.
Thus, it follows that $p^* \geq \sup g(\lambda) = d^*$.

Equality constraints

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t. } & f_i(x) \leq 0, \quad i = 1, \ldots, m, \\ & h_i(x) = 0, \quad i = 1, \ldots, p. \end{aligned}$$

Exercise: Show that we get the Lagrangian dual

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

where the Lagrange variable ν corresponding to the equality constraints is unconstrained.
Hint: Represent $h_i(x) = 0$ as $h_i(x) \leq 0$ and $-h_i(x) \leq 0$.

Equality constraints

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t. } & f_i(x) \le 0, \quad i = 1, \ldots, m, \\ & h_i(x) = 0, \quad i = 1, \ldots, p. \end{aligned}$$

Exercise: Show that we get the Lagrangian dual

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

where the Lagrange variable ν corresponding to the equality constraints is unconstrained.
Hint: Represent $h_i(x) = 0$ as $h_i(x) \le 0$ and $-h_i(x) \le 0$.

Again, we see that $p^* \ge \sup_{\lambda \ge 0, \nu} \; g(\lambda, \nu) = d^*$

Some duals

- Least-norm solution of linear equations: $\min x^T x$ s.t. $Ax = b$
- Linear programming standard form
- Study example (5.7) in BV (binary QP)

Strong duality

Duality gap

$$p^* - d^* \geq 0$$

Duality gap

$$p^* - d^* \geq 0$$

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

Notice: both p^* and d^* may be $+\infty$

Duality gap

$$p^* - d^* \geq 0$$

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

Notice: both p^* and d^* may be $+\infty$

Several **sufficient** conditions known!

Duality gap

$$p^* - d^* \geq 0$$

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

Notice: both p^* and d^* may be $+\infty$

Several **sufficient** conditions known!

"Easy" necessary and sufficient conditions: unknown

Slater's sufficient conditions

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & Ax = b. \end{aligned}$$

Slater's sufficient conditions

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t. } & f_i(x) \le 0, \quad 1 \le i \le m, \\ & Ax = b. \end{aligned}$$

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

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

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

Slater's sufficient conditions

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad 1 \leq i \leq m, \\ & Ax = b. \end{aligned}$$

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

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

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., $d^* > -\infty$).

Reading: Read BV §5.3.2 for a proof.

Counterexample

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

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

Counterexample

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

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

Counterexample

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

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

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

so dual function is

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

Counterexample

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

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

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

Counterexample

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

over the domain $\mathcal{D} = \{(x,y) \mid y > 0\}$.
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 $p^* - d^* = 1$.

Counterexample

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

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

Example: Maxent

$$\begin{aligned} \min \quad & \textstyle\sum_i x_i \log x_i \\ & Ax \leq b, \quad 1^T x = 1, \quad x > 0. \end{aligned}$$

Example: Maxent

$$\begin{array}{ll} \min & \sum_i x_i \log x_i \\ & Ax \leq b, \quad 1^T x = 1, \quad x > 0. \end{array}$$

Recall, convex conjugate of $f(x) = x \log x$ is $f^*(y) = e^{y-1}$.

Example: Maxent

$$\min \quad \sum_i x_i \log x_i$$
$$Ax \le b, \quad 1^T x = 1, \quad x > 0.$$

Recall, convex conjugate of $f(x) = x \log x$ is $f^*(y) = e^{y-1}$.

$$\max_{\lambda,\nu} \quad g(\lambda, \nu) = -b^T \lambda - v - \sum_{i=1}^{n} e^{-(A^T\lambda)_i - \nu - 1}$$
$$\text{s.t.} \quad \lambda \ge 0.$$

Example: Maxent

$$\min \quad \sum_i x_i \log x_i$$
$$Ax \le b, \quad 1^T x = 1, \quad x > 0.$$

Recall, convex conjugate of $f(x) = x \log x$ is $f^*(y) = e^{y-1}$.

$$\max_{\lambda,\nu} \quad g(\lambda,\nu) = -b^T\lambda - v - \sum_{i=1}^n e^{-(A^T\lambda)_i - \nu - 1}$$
$$\text{s.t.} \quad \lambda \ge 0.$$

If there is $x > 0$ with $Ax \le b$ and $1^T x = 1$, strong duality holds.

Example: Maxent

$$\min \quad \sum_i x_i \log x_i$$
$$Ax \leq b, \quad 1^T x = 1, \quad x > 0.$$

Recall, convex conjugate of $f(x) = x \log x$ is $f^*(y) = e^{y-1}$.

$$\max_{\lambda,\nu} \quad g(\lambda, \nu) = -b^T \lambda - v - \sum_{i=1}^n e^{-(A^T\lambda)_i - \nu - 1}$$
$$\text{s.t.} \quad \lambda \geq 0.$$

If there is $x > 0$ with $Ax \leq b$ and $1^T x = 1$, strong duality holds.
Exercise: Simplify above dual by optimizing out ν

Support vector machine

$$
\begin{aligned}
\min_{x,\xi} \quad & \tfrac{1}{2}\|x\|_2^2 + C\sum\nolimits_i \xi_i \\
\text{s.t.} \quad & Ax \ge 1 - \xi, \quad \xi \ge 0.
\end{aligned}
$$

Support vector machine

$$\begin{aligned} \min_{x,\xi} \quad & \tfrac{1}{2}\|x\|_2^2 + C\sum\nolimits_i \xi_i \\ \text{s.t.} \quad & Ax \geq 1 - \xi, \quad \xi \geq 0. \end{aligned}$$

$$L(x,\xi,\lambda,\nu) = \tfrac{1}{2}\|x\|_2^2 + C1^T\xi - \lambda^T(Ax - 1 + \xi) - \nu^T\xi$$

Support vector machine

$$\begin{aligned} \min_{x,\xi} \quad & \tfrac{1}{2}\|x\|_2^2 + C\sum\nolimits_i \xi_i \\ \text{s.t.} \quad & Ax \geq 1 - \xi, \quad \xi \geq 0. \end{aligned}$$

$$L(x, \xi, \lambda, \nu) = \tfrac{1}{2}\|x\|_2^2 + C1^T\xi - \lambda^T(Ax - 1 + \xi) - \nu^T\xi$$

$$\begin{aligned} g(\lambda, \nu) \quad & := \quad \inf L(x, \xi, \lambda, \nu) \\ & = \quad \begin{cases} \lambda^T 1 - \frac{1}{2}\|A^T\lambda\|_2^2 & \lambda + \nu = C\mathbf{1} \\ +\infty & \text{otherwise} \end{cases} \\ d^* \quad & = \quad \max_{\lambda \geq 0, \nu \geq 0} \quad g(\lambda, \nu) \end{aligned}$$

Exercise: Using $\nu \geq 0$, eliminate ν from above problem.

Example: regularized optimization

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

Example: regularized optimization

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

Dual problem

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

Example: regularized optimization

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

Dual problem

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

- Introduce new variable $z = Ax$

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

Example: regularized optimization

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

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

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

Example: regularized optimization

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

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

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

Regularized optimization

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

Dual problem

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

The infimum above can be rearranged as follows

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

Regularized optimization

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

Dual problem

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

The infimum above can be rearranged as follows

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

Regularized optimization

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

Dual problem

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

The infimum above can be rearranged as follows

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

Regularized optimization

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

Dual problem

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

The infimum above can be rearranged as follows

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

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

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

Regularized optimization

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:

Regularized optimization

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 $A^T y \in \mathrm{ri}(\mathrm{dom}\, f^*)$

Regularized optimization

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 $A^T y \in \mathrm{ri}(\mathrm{dom}\, f^*)$

- Condition 1 ensures 'sup' attained at some y
- 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 $A^T \bar{y} \in \mathrm{ri}(\mathrm{dom}\, f^*)$, then we have strong duality (e.g., for instance $0 \in \mathrm{ri}(\mathrm{dom}\, f^*)$)

Dual via Fenchel conjugates

$\min f(x)$ s.t. $f_i(x) \leq 0, Ax = b.$

$$\mathcal{L}(x, \lambda, \nu) \quad := \quad f_0(x) + \sum_i \lambda_i f_i(x) + \nu^T (Ax - b)$$

Dual via Fenchel conjugates

$\min f(x)$ s.t. $f_i(x) \leq 0, Ax = b.$

$$
\begin{aligned}
\mathcal{L}(x,\lambda,\nu) &:= f_0(x) + \sum_i \lambda_i f_i(x) + \nu^T(Ax - b) \\
g(\lambda,\nu) &= \inf_x \mathcal{L}(x,\lambda,\nu)
\end{aligned}
$$

Dual via Fenchel conjugates

$\min f(x)$ s.t. $f_i(x) \le 0, Ax = b$.

$$\begin{aligned}\mathcal{L}(x, \lambda, \nu) &:= f_0(x) + \sum_i \lambda_i f_i(x) + \nu^T(Ax - b) \\ g(\lambda, \nu) &= \inf_x \mathcal{L}(x, \lambda, \nu) \\ g(\lambda, \nu) &= -\nu^T b + \inf_x x^T A^T \nu + F(x)\end{aligned}$$

Dual via Fenchel conjugates

$\min f(x)$ s.t. $f_i(x) \leq 0, Ax = b$.

$$\begin{aligned}
\mathcal{L}(x, \lambda, \nu) &:= f_0(x) + \sum_i \lambda_i f_i(x) + \nu^T (Ax - b) \\
g(\lambda, \nu) &= \inf_x \mathcal{L}(x, \lambda, \nu) \\
g(\lambda, \nu) &= -\nu^T b + \inf_x x^T A^T \nu + F(x) \\
F(x) &:= f_0(x) + \sum_i \lambda_i f_i(x)
\end{aligned}$$

Dual via Fenchel conjugates

$\min f(x)$ s.t. $f_i(x) \leq 0, Ax = b.$

$$\begin{aligned}
\mathcal{L}(x, \lambda, \nu) &:= f_0(x) + \sum_i \lambda_i f_i(x) + \nu^T(Ax - b) \\
g(\lambda, \nu) &= \inf_x \mathcal{L}(x, \lambda, \nu) \\
g(\lambda, \nu) &= -\nu^T b + \inf_x x^T A^T \nu + F(x) \\
F(x) &:= f_0(x) + \sum_i \lambda_i f_i(x) \\
g(\lambda, \nu) &= -\nu^T b - \sup_x \langle x, -A^T \nu \rangle - F(x)
\end{aligned}$$

Dual via Fenchel conjugates

$\min f(x)$ s.t. $f_i(x) \leq 0, Ax = b.$

$$\begin{aligned}
\mathcal{L}(x,\lambda,\nu) &:= f_0(x) + \sum_i \lambda_i f_i(x) + \nu^T(Ax - b) \\
g(\lambda,\nu) &= \inf_x \mathcal{L}(x,\lambda,\nu) \\
g(\lambda,\nu) &= -\nu^T b + \inf_x x^T A^T \nu + F(x) \\
F(x) &:= f_0(x) + \sum_i \lambda_i f_i(x) \\
g(\lambda,\nu) &= -\nu^T b - \sup_x \langle x, -A^T\nu\rangle - F(x) \\
g(\lambda,\nu) &= -\nu^T b - F^*(-A^T\nu).
\end{aligned}$$

Not so useful! F^* hard to compute.

Dual via Fenchel conjugates

- Introduce new variables!

Dual via Fenchel conjugates

Introduce new variables!

$$\min f(x) \quad \text{s.t.} \quad f_i(x_i) \le 0, Ax = b$$
$$x_i = z, i = 1, \ldots, m.$$

Dual via Fenchel conjugates

Introduce new variables!

$$\min f(x) \quad \text{s.t.} \quad f_i(x_i) \leq 0, Ax = b$$
$$x_i = z, i = 1, \ldots, m.$$

$$\mathcal{L}(x, x_i z, \lambda, \nu, \pi_i)$$
$$:= f(x) + \sum\nolimits_i \lambda_i f_i(x_i) + \nu^T(Ax - b) + \sum\nolimits_i \pi_i^T(x_i - z)$$

Dual via Fenchel conjugates

Introduce new variables!

$$\min f(x) \quad \text{s.t.} \qquad f_i(x_i) \leq 0, Ax = b$$
$$x_i = z, i = 1, \ldots, m.$$

$$\mathcal{L}(x, x_i z, \lambda, \nu, \pi_i)$$
$$:= f(x) + \sum\nolimits_i \lambda_i f_i(x_i) + \nu^T(Ax - b) + \sum\nolimits_i \pi_i^T(x_i - z)$$
$$g(\lambda, \nu, \pi_i) = \inf_{x, x_i, z} \mathcal{L}(x, x_i, z, \lambda, \nu, \pi_i)$$

Dual via Fenchel conjugates

Introduce new variables!

$$\min f(x) \quad \text{s.t.} \quad f_i(x_i) \leq 0, Ax = b$$
$$x_i = z, i = 1, \ldots, m.$$

$$\mathcal{L}(x, x_i z, \lambda, \nu, \pi_i)$$
$$:= f(x) + \sum\nolimits_i \lambda_i f_i(x_i) + \nu^T(Ax - b) + \sum\nolimits_i \pi_i^T(x_i - z)$$
$$g(\lambda, \nu, \pi_i) = \inf_{x, x_i, z} \mathcal{L}(x, x_i, z, \lambda, \nu, \pi_i)$$
$$= -\nu^T b + \inf_x f(x) + \nu^T Ax + \inf_z \sum\nolimits_i -\pi_i^T z$$
$$+ \sum\nolimits_i \inf_{x_i} \pi_i^T x_i + \lambda_i f_i(x_i)$$

Dual via Fenchel conjugates

Introduce new variables!

$$\min f(x) \quad \text{s.t.} \quad f_i(x_i) \le 0, Ax = b$$
$$x_i = z, i = 1, \ldots, m.$$

$$\mathcal{L}(x, x_i z, \lambda, \nu, \pi_i)$$
$$:= f(x) + \sum\nolimits_i \lambda_i f_i(x_i) + \nu^T(Ax - b) + \sum\nolimits_i \pi_i^T(x_i - z)$$
$$g(\lambda, \nu, \pi_i) = \inf_{x, x_i, z} \mathcal{L}(x, x_i, z, \lambda, \nu, \pi_i)$$
$$= -\nu^T b + \inf_x f(x) + \nu^T Ax + \inf_z \sum\nolimits_i -\pi_i^T z$$
$$+ \sum\nolimits_i \inf_{x_i} \pi_i^T x_i + \lambda_i f_i(x_i)$$
$$= \begin{cases} -\nu^T b - f^*(-A^T\nu) - \sum_i (\lambda_i f_i)^*(-\pi_i) & \text{if } \sum_i \pi_i = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

Example

Exercise: Derive the Lagrangian dual in terms of Fenchel conjugates for the following linearly constrained problem:

$$\min \quad f(x) \quad \text{s.t. } Ax \leq b, \quad Cx = d.$$

Hint: No need to introduce extra variables.

Example: variable splitting

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

Example: variable splitting

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

Exercise: Fill in the details for the following steps

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

Example: variable splitting

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

Exercise: Fill in the details for the following steps

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

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

Example: variable splitting

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

Exercise: Fill in the details for the following steps

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

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

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

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

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

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

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

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

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

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,

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

Proof:

$$\forall x, y, \quad \inf_{x' \in \mathcal{X}} \phi(x', y) \;\; \leq \;\; \phi(x, y)$$

Weak minimax

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

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

Proof:

$$\forall x, y, \quad \inf_{x' \in \mathcal{X}} \phi(x', y) \;\leq\; \phi(x, y)$$

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

Weak minimax

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

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

Proof:

$$\forall x, y, \quad \inf_{x' \in \mathcal{X}} \phi(x', y) \le \phi(x, y)$$

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

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

Weak minimax

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

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

Proof:

$$\begin{aligned}
\forall x, y, \quad & \inf_{x' \in \mathcal{X}} \phi(x', y) && \leq && \phi(x, y) \\
\forall x, y, \quad & \inf_{x' \in \mathcal{X}} \phi(x', y) && \leq && \sup_{y' \in \mathcal{Y}} \phi(x, y') \\
\forall x, \quad & \sup_{y \in \mathcal{Y}} \inf_{x' \in \mathcal{X}} \phi(x', y) && \leq && \sup_{y' \in \mathcal{Y}} \phi(x, y') \\
\Longrightarrow \quad & \sup_{y \in \mathcal{Y}} \inf_{x' \in \mathcal{X}} \phi(x', y) && \leq && \inf_{x \in \mathcal{X}} \sup_{y' \in \mathcal{Y}} \phi(x, y').
\end{aligned}$$

Weak minimax

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

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

Proof:

$$\begin{aligned}
\forall x, y, \quad & \inf_{x' \in \mathcal{X}} \phi(x', y) &\leq& \quad \phi(x, y) \\
\forall x, y, \quad & \inf_{x' \in \mathcal{X}} \phi(x', y) &\leq& \quad \sup_{y' \in \mathcal{Y}} \phi(x, y') \\
\forall x, \quad & \sup_{y \in \mathcal{Y}} \inf_{x' \in \mathcal{X}} \phi(x', y) &\leq& \quad \sup_{y' \in \mathcal{Y}} \phi(x, y') \\
\implies \quad & \sup_{y \in \mathcal{Y}} \inf_{x' \in \mathcal{X}} \phi(x', y) &\leq& \quad \inf_{x \in \mathcal{X}} \sup_{y' \in \mathcal{Y}} \phi(x, y').
\end{aligned}$$

Exercise: Show that weak duality is follows from above minimax inequality. *Hint:* Use $\phi = \mathcal{L}$ (Lagrangian), and suitably choose 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^*)

$$\phi(x, y^*) \geq \phi(x^*, y^*) \geq \phi(x^*, y) \quad \text{for all } x \in \mathcal{X}, y \in \mathcal{Y}.$$

Exercise: Verify above inequality!

Strong minimax

Def. Let ϕ be as before. A point (x^*, y^*) is a saddle-point of ϕ (min over $\mathcal{X}$ and max over $\mathcal{Y}$) **iff** 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 (min over $\mathcal{X}$ and max over $\mathcal{Y}$) **iff** 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

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

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

Example: Lasso-like problem

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

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

Saddle-point formulation

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

Example: Lasso-like problem

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

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

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

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

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

Saddle-point formulation

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

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

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

Nonconvex QP – I (TRS)

Trust region subproblem (TRS)

$$\min \quad x^T A x + 2b^T x \qquad x^T x \le 1.$$

A is symmetric but not necessarily semidefinite!

Nonconvex QP – I (TRS)

Trust region subproblem (TRS)

$$\min \quad x^T A x + 2b^T x \qquad x^T x \le 1.$$

A is symmetric but not necessarily semidefinite!

Theorem TRS always has zero duality gap.