Exact Augmented Lagrangian Duality in Mixed-Integer Nonlinear Optimization

Henri Lefebvre, Martin Schmidt
Trier University (Germany), Department of Mathematics

ALOP Research Seminar, 2024

Problem Setting

We consider general MI(N)LPs

$$
\begin{aligned}
z^* = \min_x \; & c^\top x \\
\text{s.t. } & Ax = b \\
& Bx \geq f \\
& x \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}
\end{aligned}
$$

Problem Setting

We consider general MI(N)LPs

$$
\begin{aligned}
z^* = \min_x \; & c^\top x \\
\text{s.t. } & Ax = b \\
& Bx \geq f \\
& x \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2}
\end{aligned}
$$

Our main interest is in “dualizing” constraints “$Ax = b$”

Problem Setting

We consider general MI(N)LPs

$$\begin{aligned} z^* = \min_x \ & c^\top x \\ \text{s.t. } & Ax = b \\ & Bx \geq f \\ & x \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2} \end{aligned}$$

Our main interest is in "dualizing" constraints "$Ax = b$"

Let $X = \{x \in \mathbb{R}^{n_1} \times \mathbb{Z}^{n_2} : Bx \geq f\}$

Lagrangian Duality

Primal Problem

$$z^* = \min_x \; c^\top x \quad \text{s.t. } Ax = b, \; x \in X$$

Lagrangian Dual Problem

$$z^{\text{LD}} = \sup_{\lambda \in \mathbb{R}^m} \; z^{\text{LR}}(\lambda)$$

$$z^{\text{LR}}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b)$$

Lagrangian Duality

Primal Problem

$$
\begin{aligned}
z^* = \min_x \ & c^\top x \\
\text{s.t. } & Ax = b \\
& x \in X
\end{aligned}
$$

Lagrangian Dual Problem

$$z^{\mathrm{LD}} = \sup_{\lambda \in \mathbb{R}^m} \ z^{\mathrm{LR}}(\lambda)$$

$$z^{\mathrm{LR}}(\lambda) = \min_{x \in X} \ c^\top x + \lambda^\top (Ax - b)$$

Strong duality does not hold in general, i.e., $z^* > z^{\mathrm{LD}}$

Augmented Lagrangian Duality

Primal Problem

$$z^* = \min_x \; c^\top x$$
$$\text{s.t. } Ax = b$$
$$x \in X$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\text{LD}+} = \sup_{\lambda \in \mathbb{R}^m} \; z_\rho^{\text{LR}+}(\lambda)$$

$$z_\rho^{\text{LR}+}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Augmented Lagrangian Duality

Primal Problem

$$z^* = \min_x \ c^\top x \quad \text{s.t. } Ax = b, \quad x \in X$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\mathrm{LD+}} = \sup_{\lambda \in \mathbb{R}^m} \ z_\rho^{\mathrm{LR+}}(\lambda)$$

$$z_\rho^{\mathrm{LR+}}(\lambda) = \min_{x \in X} \ c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Our main interest:

1. What can we say about the duality gap $\gamma = z^* - z_\rho^{\mathrm{LD+}}$?

Augmented Lagrangian Duality

Primal Problem

$$
\begin{aligned}
z^* = \min_x \ & c^\top x \\
\text{s.t. } & Ax = b \\
& x \in X
\end{aligned}
$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\text{LD}+} = \sup_{\lambda \in \mathbb{R}^m} \ z_\rho^{\text{LR}+}(\lambda)$$

$$z_\rho^{\text{LR}+}(\lambda) = \min_{x \in X} \ c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Our main interest:

1. What can we say about the duality gap $\gamma = z^* - z_\rho^{\text{LD}+}$?
2. When do we have $\gamma = 0$?

Augmented Lagrangian Duality

Primal Problem

$$z^* = \min_x \; c^\top x$$
$$\text{s.t. } Ax = b$$
$$x \in X$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\text{LD}+} = \sup_{\lambda \in \mathbb{R}^m} \; z_\rho^{\text{LR}+}(\lambda)$$

$$z_\rho^{\text{LR}+}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Our main interest:

1. What can we say about the duality gap $\gamma = z^* - z_\rho^{\text{LD}+}$?
2. When do we have $\gamma = 0$?
    2.1 For $\rho \to \infty$? (Asymptotic result)

Augmented Lagrangian Duality

Primal Problem

$$z^* = \min_x \; c^\top x$$
$$\text{s.t. } Ax = b$$
$$x \in X$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\text{LD}+} = \sup_{\lambda \in \mathbb{R}^m} \; z_\rho^{\text{LR}+}(\lambda)$$

$$z_\rho^{\text{LR}+}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho \psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Our main interest:

1. What can we say about the duality gap $\gamma = z^* - z_\rho^{\text{LD}+}$?
2. When do we have $\gamma = 0$?
   2.1 For $\rho \to \infty$? (Asymptotic result)
   2.2 For some $\rho < \infty$? (Exactness result)

Augmented Lagrangian Duality

Primal Problem

$$
\begin{aligned}
z^* = \min_x \ & c^\top x \\
\text{s.t. } & Ax = b \\
& x \in X
\end{aligned}
$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\text{LD+}} = \sup_{\lambda \in \mathbb{R}^m} z_\rho^{\text{LR+}}(\lambda)$$

$$z_\rho^{\text{LR+}}(\lambda) = \min_{x \in X} \ c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Our main interest:

1. What can we say about the duality gap $\gamma = z^* - z_\rho^{\text{LD+}}$?
2. When do we have $\gamma = 0$?
   - 2.1 For $\rho \to \infty$? (Asymptotic result)
   - 2.2 For some $\rho < \infty$? (Exactness result)
3. Is it hard to compute ρ?

Augmented Lagrangian Duality

Primal Problem

$$z^* = \min_x \ c^\top x$$
$$\text{s.t. } Ax = b$$
$$x \in X$$

Augmented Lagrangian Dual Problem

$$z_\rho^{\text{LD}+} = \sup_{\lambda \in \mathbb{R}^m} \ z_\rho^{\text{LR}+}(\lambda)$$

$$z_\rho^{\text{LR}+}(\lambda) = \min_{x \in X} \ c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

with $\psi(u) > u$ if and only if $u \neq 0$, $\psi(0) = 0$

Our main interest:

1. What can we say about the duality gap $\gamma = z^* - z_\rho^{\text{LD}+}$?
2. When do we have $\gamma = 0$?
   2.1 For $\rho \to \infty$? (Asymptotic result)
   2.2 For some $\rho < \infty$? (Exactness result)
3. ~~Is it hard to compute ρ?~~

Outline

Literature Review

Asymptotic Result

Exact Penalty Parameters

Guarantees From a Given Penalty Parameter

Conclusion

Outline

Previous Works

Growing interest in the discrete community

	Asymptotic	Exactness	Poly. size	Poly. time	Opt. set
ILP (Boland and Eberhard 2014)	✓	✓	⇑		⇑
MILP (Feizollahi et al. 2016)	✓	✓	⇑		✓
MIQP (Gu et al. 2020)	✓	✓	✓		
MICP (Bhardwaj et al. 2024)	✓	✓			

Previous Works

Growing interest in the discrete community

	Asymptotic	Exactness	Poly. size	Poly. time	Opt. set
ILP (Boland and Eberhard 2014)	$\checkmark$	$\checkmark$	$\Uparrow$	$\Uparrow$	$\Uparrow$
MILP (Feizollahi et al. 2016)	$\checkmark$	$\checkmark$	$\Uparrow$	$\Uparrow$	$\checkmark$
MIQP (Gu et al. 2020)	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\Uparrow$
MICP (Bhardwaj et al. 2024)	$\checkmark$	$\checkmark$			$\Uparrow$
MINLP (our work)	$\checkmark$	$\checkmark$			$\checkmark$

Assumptions

Assumption (Compactness)

X is compact.

Assumptions

Assumption (Compactness)

X is compact.

Assumption (Penalty Function)

The penalty function $\psi : \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ is

1. *continuous, i.e., $\lim_{u \to u^*} \psi(u) = \psi(u^*)$;*
2. *positive definite, i.e., $\psi(u) > 0$ for all $u \neq 0$ and $\psi(0) = 0$.*

Outline

Bounding the Augmenting Term

Let $\lambda \in \mathbb{R}^m$. Let x_ρ denote any solution of

$$z_\rho^{\text{LR+}}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

Bounding the Augmenting Term

Let $\lambda \in \mathbb{R}^m$. Let x_ρ denote any solution of

$$z_\rho^{\mathrm{LR+}}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

$$\rho\psi(Ax_\rho - b) \leq \kappa$$

Bounding the Augmenting Term

Let $\lambda \in \mathbb{R}^m$. Let x_ρ denote any solution of

$$z_\rho^{\mathrm{LR+}}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

$$\rho\psi(Ax_\rho - b) \leq \kappa$$

Weak Duality

$$z_\rho^{\mathrm{LR+}}(\lambda) \leq z^*$$

Bounding the Augmenting Term

Let $\lambda \in \mathbb{R}^m$. Let x_ρ denote any solution of

$$z_\rho^{\mathrm{LR}+}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

$$\rho\psi(Ax_\rho - b) \leq \kappa$$

Weak Duality

$$z_\rho^{\mathrm{LR}+}(\lambda) \leq z^*$$

$$c^\top x_\rho + \lambda^\top (Ax_\rho - b) + \rho\psi(Ax_\rho - b) \leq c^\top x^*$$

Bounding the Augmenting Term

Let $\lambda \in \mathbb{R}^m$. Let x_ρ denote any solution of

$$z_\rho^{\text{LR+}}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

$$\rho\psi(Ax_\rho - b) \leq \kappa$$

Weak Duality

$$z_\rho^{\text{LR+}}(\lambda) \leq z^*$$

$$\begin{aligned} & c^\top x_\rho + \lambda^\top (Ax_\rho - b) + \rho\psi(Ax_\rho - b) \leq c^\top x^* \\ \iff \quad & \rho\psi(Ax_\rho - b) \leq c^\top x^* - c^\top x_\rho + \lambda^\top (b - Ax_\rho) \end{aligned}$$

Bounding the Augmenting Term

Let $\lambda \in \mathbb{R}^m$. Let x_ρ denote any solution of

$$z_\rho^{\text{LR+}}(\lambda) = \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\psi(Ax - b)$$

$$\rho\psi(Ax_\rho - b) \le \kappa$$

Weak Duality

$$z_\rho^{\text{LR+}}(\lambda) \le z^*$$

$$\begin{aligned}
& c^\top x_\rho + \lambda^\top (Ax_\rho - b) + \rho\psi(Ax_\rho - b) \le c^\top x^* \\
\iff \quad & \rho\psi(Ax_\rho - b) \le c^\top x^* - c^\top x_\rho + \lambda^\top (b - Ax_\rho) \\
\implies \quad & \rho\psi(Ax_\rho - b) \le \max_{y,z \in X} \left\{ c^\top y - c^\top z + \lambda^\top (b - Az) \right\}
\end{aligned}$$

Asymptotic Result (Part 1)

$$\rho \geq \frac{1}{\varepsilon}\kappa$$

Asymptotic Result (Part 1)

$$\begin{aligned} & \rho \geq \frac{1}{\varepsilon}\kappa \\ \Longrightarrow \quad & \rho\psi(Ax_\rho - b) \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \end{aligned}$$

Asymptotic Result (Part 1)

$$\begin{aligned} & \rho \geq \frac{1}{\varepsilon}\kappa \\ \implies \quad & \rho\psi(Ax_\rho - b) \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\ \implies \quad & \kappa \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \end{aligned}$$

Asymptotic Result (Part 1)

$$
\begin{aligned}
& \rho \geq \frac{1}{\varepsilon}\kappa \\
\implies \quad & \rho\psi(Ax_\rho - b) \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\
\implies \quad & \kappa \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\
\implies \quad & \varepsilon \geq \psi(Ax_\rho - b)
\end{aligned}
$$

Asymptotic Result (Part 1)

$$
\begin{aligned}
& \rho \geq \frac{1}{\varepsilon}\kappa \\
\implies \quad & \rho\psi(Ax_\rho - b) \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\
\implies \quad & \kappa \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\
\implies \quad & \varepsilon \geq \psi(Ax_\rho - b)
\end{aligned}
$$

Asymptotic Result (Part 1)

$$
\begin{aligned}
& \rho \geq \frac{1}{\varepsilon}\kappa \\
\implies \quad & \rho\psi(Ax_\rho - b) \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\
\implies \quad & \kappa \geq \frac{1}{\varepsilon}\kappa\psi(Ax_\rho - b) \\
\implies \quad & \varepsilon \geq \psi(Ax_\rho - b)
\end{aligned}
$$

Let $\varepsilon > 0$. Any $\rho \geq \frac{1}{\varepsilon}\kappa$ guarantees that $\psi(Ax_\rho - b) \leq \varepsilon$.

Asymptotic Result (Part 2)

Let $\varepsilon \to 0$ ($\rho \to \infty$)

There is a limit point of (a sub-sequence of) $(x_\rho)_{\rho>0}$, say $x_\infty \in X$

$$\varepsilon \geq \psi(Ax_\rho - b)$$

Asymptotic Result (Part 2)

Let $\varepsilon \to 0$ $(\rho \to \infty)$

There is a limit point of (a sub-sequence of) $(x_\rho)_{\rho>0}$, say $x_\infty \in X$

$$
\begin{aligned}
& \varepsilon \geq \psi(Ax_\rho - b) \\
\overset{\varepsilon \to 0}{\Longrightarrow} \quad & 0 \geq \liminf_{\rho \to \infty} \; \psi(Ax_\rho - b)
\end{aligned}
$$

Asymptotic Result (Part 2)

Let $\varepsilon \to 0$ ($\rho \to \infty$)

There is a limit point of (a sub-sequence of) $(x_\rho)_{\rho>0}$, say $x_\infty \in X$

$$
\begin{aligned}
& \varepsilon \geq \psi(Ax_\rho - b) \\
\overset{\varepsilon \to 0}{\Longrightarrow} \quad & 0 \geq \liminf_{\rho \to \infty} \ \psi(Ax_\rho - b) \\
\Longrightarrow \quad & 0 \geq \psi(Ax_\infty - b)
\end{aligned}
$$

Asymptotic Result (Part 2)

Let $\varepsilon \to 0$ ($\rho \to \infty$)

There is a limit point of (a sub-sequence of) $(x_\rho)_{\rho>0}$, say $x_\infty \in X$

$$
\begin{aligned}
& & \varepsilon &\geq \psi(Ax_\rho - b) \\
&\overset{\varepsilon\to 0}{\Longrightarrow} & 0 &\geq \liminf_{\rho\to\infty} \ \psi(Ax_\rho - b) \\
&\Longrightarrow & 0 &\geq \psi(Ax_\infty - b) \\
&\Longleftrightarrow & Ax_\infty &= b
\end{aligned}
$$

Asymptotic Result (Part 2)

Let $\varepsilon \to 0$ ($\rho \to \infty$)

There is a limit point of (a sub-sequence of) $(x_\rho)_{\rho>0}$, say $x_\infty \in X$

$$
\begin{aligned}
& & \varepsilon &\geq \psi(Ax_\rho - b) \\
&\overset{\varepsilon\to 0}{\Longrightarrow} & 0 &\geq \liminf_{\rho\to\infty} \; \psi(Ax_\rho - b) \\
&\Longrightarrow & 0 &\geq \psi(Ax_\infty - b) \\
&\Longleftrightarrow & Ax_\infty &= b
\end{aligned}
$$

Asymptotic Result (Part 2)

Let $\varepsilon \to 0$ ($\rho \to \infty$)

There is a limit point of (a sub-sequence of) $(x_\rho)_{\rho>0}$, say $x_\infty \in X$

$$\begin{aligned} & \varepsilon \geq \psi(Ax_\rho - b) \\ \overset{\varepsilon \to 0}{\Longrightarrow} \quad & 0 \geq \liminf_{\rho \to \infty} \; \psi(Ax_\rho - b) \\ \Longrightarrow \quad & 0 \geq \psi(Ax_\infty - b) \\ \Longleftrightarrow \quad & Ax_\infty = b \end{aligned}$$

This shows that x_∞ is feasible for the primal problem!

$$z^* = \lim_{\rho \to \infty} \; z_\rho^{\mathrm{LR}+}(\lambda)$$

Asymptotic Result (Part 3)

$\rho \to z_\rho^{\mathsf{LR}+}(\lambda)$ is non-decreasing

Asymptotic Result (Part 3)

$\rho \to z_{\rho}^{\text{LR+}}(\lambda)$ is non-decreasing

$z_{\rho}^{\text{LR+}}(\lambda) \leq z^*$

Asymptotic Result (Part 3)

$\rho \to z_\rho^{\mathsf{LR+}}(\lambda)$ is non-decreasing

$z_\rho^{\mathsf{LR+}}(\lambda) \leq z^*$

$$z^* = \sup_{\rho > 0} \; z_\rho^{\mathsf{LR+}}(\lambda)$$

Outline

A Trivial Case

Assume that for all $x \in X$ such that $Ax \neq b$ it holds $\psi(Ax - b) \geq \delta > 0$.

A Trivial Case

Assume that for all $x \in X$ such that $Ax \neq b$ it holds $\psi(Ax - b) \geq \delta > 0$.

For any $\rho \geq 2\kappa/\delta$, $\qquad z^{\mathsf{LR+}}_{\rho}(\lambda) = z^*$

A Trivial Case

Assume that for all $x \in X$ such that $Ax \neq b$ it holds $\psi(Ax - b) \geq \delta > 0$.

For any $\rho \geq 2\kappa/\delta$, $\quad z_\rho^{\mathrm{LR+}}(\lambda) = z^*$

It holds

$$\psi(Ax_\rho - b) \leq \delta/2$$

A Trivial Case

Assume that for all $x \in X$ such that $Ax \neq b$ it holds $\psi(Ax - b) \geq \delta > 0$.

For any $\rho \geq 2\kappa/\delta$, $\qquad z_\rho^{\mathrm{LR+}}(\lambda) = z^*$

It holds

$$\psi(Ax_\rho - b) \leq \delta/2$$

So it must be that $\psi(Ax_\rho - b) \leq 0 \implies Ax_\rho = b$

A Trivial Case

Assume that for all $x \in X$ such that $Ax \neq b$ it holds $\psi(Ax - b) \geq \delta > 0$.

For any $\rho \geq 2\kappa/\delta$, $\quad z_\rho^{\text{LR+}}(\lambda) = z^*$

It holds

$$\psi(Ax_\rho - b) \leq \delta/2$$

So it must be that $\psi(Ax_\rho - b) \leq 0 \implies Ax_\rho = b$

Problem Assumption already fails for some MILPs (Feizollahi et al., 2016)

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

Lemma 1 The choice of the norm does not matter

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

Lemma 1 The choice of the norm does not matter

Say that $\rho < \infty$ is exact for $\|\cdot\|$

By equivalence of norms, there exists $\gamma > 0$ such that $\|\cdot\| \leq \gamma \|\cdot\|'$

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

Lemma 1 The choice of the norm does not matter

Say that $\rho < \infty$ is exact for $\|\cdot\|$

By equivalence of norms, there exists $\gamma > 0$ such that $\|\cdot\| \leq \gamma \|\cdot\|'$

$$
\begin{aligned}
z^* &= \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho \|Ax - b\| \\
&\leq \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \rho\gamma \|Ax - b\|' \\
&\leq z^*
\end{aligned}
$$

Thus $\rho\gamma$ is exact for $\|\cdot\|'$

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

Lemma 1 The choice of norm does not matter

Lemma 2 The choice of λ does not matter

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

Lemma 1 The choice of norm does not matter

Lemma 2 The choice of λ does not matter

Say $\rho < \infty$ is exact for $\|\cdot\|_2$ and $\lambda = 0$

Norm Penalty Functions

We now assume $\psi = \|\cdot\|$ for any norm $\|\cdot\|$.

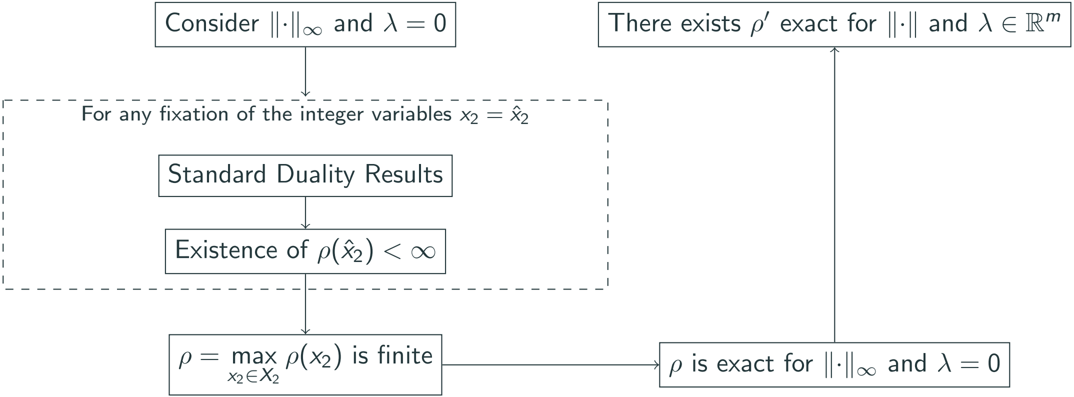
Lemma 1 The choice of norm does not matter

Lemma 2 The choice of λ does not matter

Say $\rho < \infty$ is exact for $\|\cdot\|_2$ and $\lambda = 0$

$$\begin{aligned} z^* &= \min_{x \in X} \; c^\top x + \rho \|Ax - b\|_2 \\ &= \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) - \lambda^\top (Ax - b) + \rho \|Ax - b\|_2 \\ &\leq \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + \|\lambda\|_2 \|Ax - b\|_2 + \rho \|Ax - b\|_2 \\ &= \min_{x \in X} \; c^\top x + \lambda^\top (Ax - b) + (\|\lambda\|_2 + \rho) \|Ax - b\|_2 \end{aligned}$$

Exact Penalty Parameter: How to Prove It



Exact Penalty Parameter: The Proof

Fix the integer part in the primal, say, $x_2 = \hat{x}_2$ with $x = (x_1, x_2)$

$$\begin{aligned} z^*(\hat{x}_2) = \min_{x_1} \; & c_1^\top x_1 + c_2^\top \hat{x}_2 \\ \text{s.t. } & Ax_1 = b - A\hat{x}_2 \\ & g(x_1, \hat{x}_2) \leq 0 \\ & x_1 \in \mathbb{R}^{n_1} \end{aligned}$$

We distinguish two cases

1. Feasible case: $z^*(x_2) < \infty$
2. Infeasible case: $z^*(x_2) = \infty$

Feasible Case

As per the previous lemma, let's use $\psi(\cdot) = \|\cdot\|_\infty$ and $\lambda = 0$

$$
\begin{aligned}
z^*(\hat{x}_2) - c_2^\top \hat{x}_2 = \quad \min_{x_1} \quad & c_1^\top x_1 \\
\text{s.t.} \quad & A_1 x_1 = b - A_2 \hat{x}_2 \\
& B_1 x_1 \geq f - B_2 \hat{x}_2
\end{aligned}
$$

Feasible Case

As per the previous lemma, let's use $\psi(\cdot) = \|\cdot\|_\infty$ and $\lambda = 0$

$$
\begin{aligned}
z^*(\hat{x}_2) - c_2^\top \hat{x}_2 = \min_{x_1} \quad & c_1^\top x_1 \\
\text{s.t.} \quad & A_1 x_1 = b - A_2 \hat{x}_2 \\
& B_1 x_1 \geq f - B_2 \hat{x}_2
\end{aligned}
\qquad
\begin{aligned}
= \max_{\mu,\lambda} \quad & (b - A_2 \hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \\
\text{s.t.} \quad & A_1^\top \mu + B_1^\top \lambda = c_1 \\
& \lambda \geq 0
\end{aligned}
$$

Feasible Case

As per the previous lemma, let's use $\psi(\cdot) = \|\cdot\|_\infty$ and $\lambda = 0$

$$z^*(\hat{x}_2) - c_2^\top \hat{x}_2 = \min_{x_1} \; c_1^\top x_1 \quad \text{s.t.} \quad A_1 x_1 = b - A_2 \hat{x}_2, \quad B_1 x_1 \geq f - B_2 \hat{x}_2$$

$$= \max_{\mu,\lambda} \; (b - A_2 \hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \quad \text{s.t.} \quad A_1^\top \mu + B_1^\top \lambda = c_1, \quad \lambda \geq 0$$

$$z_\rho^{\mathrm{LR+}}(\hat{x}_2) - c_2^\top \hat{x}_2 = \min_{x_1} \; c_1^\top x_1 + \rho \|A_1 x_1 + A_2 \hat{x}_2 - b\|_\infty \quad \text{s.t.} \quad B_1 x_1 \geq f - B_2 \hat{x}_2$$

Feasible Case

As per the previous lemma, let's use $\psi(\cdot) = \|\cdot\|_\infty$ and $\lambda = 0$

$$
z^*(\hat{x}_2) - c_2^\top \hat{x}_2 = \min_{x_1} \; c_1^\top x_1 \quad \text{s.t.} \quad A_1 x_1 = b - A_2\hat{x}_2, \; B_1 x_1 \geq f - B_2\hat{x}_2
$$

$$
= \max_{\mu,\lambda} \; (b - A_2\hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \quad \text{s.t.} \quad A_1^\top \mu + B_1^\top \lambda = c_1, \; \lambda \geq 0
$$

$$
z_\rho^{\text{LR}+}(\hat{x}_2) - c_2^\top \hat{x}_2 = \min_{x_1} \; c_1^\top x_1 + \rho\|A_1 x_1 + A_2\hat{x}_2 - b\|_\infty \quad \text{s.t.} \quad B_1 x_1 \geq f - B_2\hat{x}_2
$$

$$
= \max_{\mu,\lambda} \; (b - A_2\hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \quad \text{s.t.} \quad A_1^\top \mu + B_1^\top \lambda = c_1, \; \lambda \geq 0, \; \|\mu\|_1 \leq \rho
$$

Feasible Case

As per the previous lemma, let's use $\psi(\cdot) = \|\cdot\|_\infty$ and $\lambda = 0$

$$z^*(\hat{x}_2) - c_2^\top \hat{x}_2 = \begin{array}{rl} \min\limits_{x_1} & c_1^\top x_1 \\ \text{s.t.} & A_1 x_1 = b - A_2 \hat{x}_2 \\ & B_1 x_1 \geq f - B_2 \hat{x}_2 \end{array} = \begin{array}{rl} \max\limits_{\mu,\lambda} & (b - A_2 \hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \\ \text{s.t.} & A_1^\top \mu + B_1^\top \lambda = c_1 \\ & \lambda \geq 0 \end{array}$$

$$z_\rho^{\mathrm{LR}+}(\hat{x}_2) - c_2^\top \hat{x}_2 = \begin{array}{rl} \min\limits_{x_1} & c_1^\top x_1 + \rho \|A_1 x_1 + A_2 \hat{x}_2 - b\|_\infty \\ \text{s.t.} & B_1 x_1 \geq f - B_2 \hat{x}_2 \end{array} = \begin{array}{rl} \max\limits_{\mu,\lambda} & (b - A_2 \hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \\ \text{s.t.} & A_1^\top \mu + B_1^\top \lambda = c_1 \\ & \lambda \geq 0 \\ & \|\mu\|_1 \leq \rho \end{array}$$

Any $\rho > \|\mu^*\|_1$ is large enough!

Infeasible Case

It is sufficient to choose ρ such that

$$z_\rho^{\text{LR+}}(x_2) = \begin{cases} z^*(x_2), & \text{if } x_2 \text{ is feasible,} \\ \text{UB}, & \text{otherwise.} \end{cases}$$

Infeasible Case

It is sufficient to choose ρ such that

$$z_\rho^{\text{LR+}}(x_2) = \begin{cases} z^*(x_2), & \text{if } x_2 \text{ is feasible,} \\ \text{UB}, & \text{otherwise.} \end{cases}$$

If so,

$$z^* = z_\rho^{\text{LR+}} = \min_{x_2} z_\rho^{\text{LR+}}(\hat{x}_2)$$

Infeasible Case

It is sufficient to choose ρ such that

$$z_\rho^{\text{LR}+}(x_2) = \begin{cases} z^*(x_2), & \text{if } x_2 \text{ is feasible,} \\ \text{UB}, & \text{otherwise.} \end{cases}$$

If so,

$$z^* = z_\rho^{\text{LR}+} = \min_{x_2} z_\rho^{\text{LR}+}(\hat{x}_2)$$

Any ρ such that $z_\rho^{\text{LR}+}(\hat{x}_2) > \text{UB}$ is large enough!

Infeasible Case

We need to solve the "bilevel" problem

$$\begin{aligned} \min_{\rho} \quad & \rho \\ \text{s.t.} \quad & z_{\rho}^{\text{LR+}}(\hat{x}_2) \geq \text{UB} \end{aligned}$$

Infeasible Case

We need to solve the "bilevel" problem

$$\begin{aligned} \min_{\rho} \quad & \rho \\ \text{s.t.} \quad & z_{\rho}^{\mathrm{LR+}}(\hat{x}_2) \geq \mathrm{UB} \end{aligned} \qquad = \min_{\rho} \; \rho$$

$$\text{s.t.} \left\{ \begin{array}{ll} \min\limits_{x_1} & c_1^\top x_1 + \rho \|A_1 x_1 + A_2 \hat{x}_2 - b\|_\infty \\ \text{s.t.} & B_1 x_1 \geq f - B_2 \hat{x}_2 \end{array} \right\} \geq \mathrm{UB}$$

$$\rho \geq 0$$

Infeasible Case

We need to solve the "bilevel" problem

$$\begin{aligned} \min_{\rho} \quad & \rho \\ \text{s.t.} \quad & z_{\rho}^{\mathrm{LR}+}(\hat{x}_2) \geq \mathrm{UB} \end{aligned}$$

$$\begin{aligned} = \min_{\rho} \ & \rho \\ \text{s.t.} \ & \left\{ \begin{array}{ll} \min\limits_{x_1} & c_1^\top x_1 + \rho \| A_1 x_1 + A_2 \hat{x}_2 - b \|_\infty \\ \text{s.t.} & B_1 x_1 \geq f - B_2 \hat{x}_2 \end{array} \right\} \geq \mathrm{UB} \\ & \rho \geq 0 \end{aligned}$$

$$\begin{aligned} = \min_{\rho} \ & \rho \\ \text{s.t.} \ & \left\{ \begin{array}{ll} \max\limits_{\lambda,\mu} & (b - A_2 \hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \\ \text{s.t.} & A_1^\top \mu + B_1^\top \lambda = c_1 \\ & \lambda \geq 0 \\ & \|\mu\|_1 \leq \rho \end{array} \right\} \geq \mathrm{UB} \\ & \rho \geq 0 \end{aligned}$$

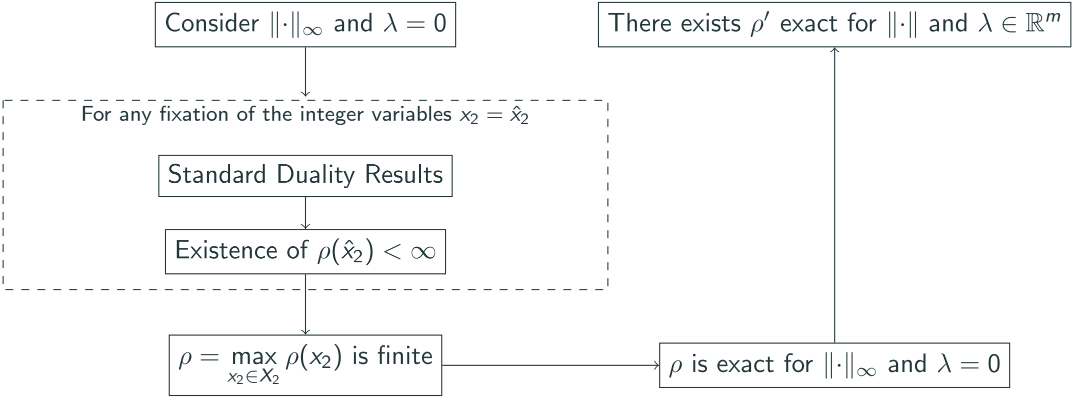
Infeasible Case

Single-level reformulation

$$
\begin{aligned}
\min_{\rho,\mu,\lambda} \quad & \rho \\
\text{s.t.} \quad & (b - A_2 \hat{x}_2)^\top \mu + (f - B_2 x_2)^\top \lambda \geq \text{UB} \\
& A_1^\top \mu + B_1^\top \lambda = c_1 \\
& \lambda \geq 0 \\
& \|\mu\|_1 \leq \rho
\end{aligned}
$$

This problem is feasible (Farkas Lemma) and lower bounded by zero.

Exact Penalty Parameter: How to Prove It



Outline

The Perturbed Problem

Primal Problem

$$
\begin{aligned}
z^* = \min_x \; & c^\top x \\
\text{s.t. } & Ax = b \\
& x \in X
\end{aligned}
$$

Perturbed Primal Problem

$$
\begin{aligned}
\tilde{z}^*_\varepsilon = \min_x \; & c^\top x \\
\text{s.t. } & \|Ax - b\|_\varphi \leq \varepsilon \\
& x \in X
\end{aligned}
$$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_\varphi \leq \varphi^{-1}(\varepsilon)$$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_{\varphi} \leq \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\qquad \psi(u) \leq \varepsilon \implies \|u\|_2 \leq \sqrt{2\varepsilon}$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_\varphi \leq \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\quad \psi(u) \leq \varepsilon \implies \|u\|_2 \leq \sqrt{2\varepsilon}$

Let $\rho \geq \kappa/\varphi(\varepsilon)$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_\varphi \leq \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\qquad \psi(u) \leq \varepsilon \implies \|u\|_2 \leq \sqrt{2\varepsilon}$

Let $\rho \geq \kappa/\varphi(\varepsilon)$

$$\psi(Ax_\rho - b) \leq \varphi(\varepsilon)$$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_{\varphi} \leq \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\qquad \psi(u) \leq \varepsilon \implies \|u\|_2 \leq \sqrt{2\varepsilon}$

Let $\rho \geq \kappa/\varphi(\varepsilon)$

$$\psi(Ax_\rho - b) \leq \varphi(\varepsilon) \qquad \implies \qquad \|Ax_\rho - b\|_{\varphi} \leq \varphi^{-1}(\varphi(\varepsilon)) = \varepsilon$$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \le \varepsilon \implies \|u\|_\varphi \le \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\qquad \psi(u) \le \varepsilon \implies \|u\|_2 \le \sqrt{2\varepsilon}$

Let $\rho \ge \kappa/\varphi(\varepsilon)$

$$\psi(Ax_\rho - b) \le \varphi(\varepsilon) \qquad \implies \qquad \|Ax_\rho - b\|_\varphi \le \varphi^{-1}(\varphi(\varepsilon)) = \varepsilon$$

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_\varphi \leq \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\quad \psi(u) \leq \varepsilon \implies \|u\|_2 \leq \sqrt{2\varepsilon}$

Let $\rho \geq \kappa/\varphi(\varepsilon)$

$$\psi(Ax_\rho - b) \leq \varphi(\varepsilon) \qquad \implies \qquad \|Ax_\rho - b\|_\varphi \leq \varphi^{-1}(\varphi(\varepsilon)) = \varepsilon$$

x_ρ is feasible for the perturbed primal problem!

Relation to a Perturbed Problem

Assume that ψ has "some relation" to a norm:

$$\psi(u) \leq \varepsilon \implies \|u\|_\varphi \leq \varphi^{-1}(\varepsilon)$$

Example $\psi = \frac{1}{2}\|\cdot\|_2^2$, $\quad \psi(u) \leq \varepsilon \implies \|u\|_2 \leq \sqrt{2\varepsilon}$

Let $\rho \geq \kappa/\varphi(\varepsilon)$

$$\psi(Ax_\rho - b) \leq \varphi(\varepsilon) \qquad \implies \qquad \|Ax_\rho - b\|_\varphi \leq \varphi^{-1}(\varphi(\varepsilon)) = \varepsilon$$

x_ρ is feasible for the perturbed primal problem!

For any $\rho \geq \kappa/\varphi(\varepsilon)$,

$$\tilde{z}_\varepsilon^* \leq z_\rho^{\mathrm{LR+}}(\lambda) \leq z^*$$

From Exact Penalty Parameter to Sensitivity

Let x^ε denote an optimal point of the perturbed problem

Let $\psi = \|\cdot\|_2$ and let $\rho^* < \infty$ be such that

$$z^* = z_{\rho^*}^{\mathsf{LR}+}(\lambda)$$

From Exact Penalty Parameter to Sensitivity

Let x^ε denote an optimal point of the perturbed problem

Let $\psi = \|\cdot\|_2$ and let $\rho^* < \infty$ be such that

$$\begin{aligned} & z^* = z_{\rho^*}^{\mathrm{LR+}}(\lambda) \\ \iff \quad & z^* = \min_{x \in X} \left(c^\top x + \lambda^\top (Ax - b) + \rho^* \psi(Ax - b) \right) \end{aligned}$$

From Exact Penalty Parameter to Sensitivity

Let x^{ε} denote an optimal point of the perturbed problem

Let $\psi = \|\cdot\|_2$ and let $\rho^* < \infty$ be such that

$$
\begin{aligned}
& z^* = z_{\rho^*}^{\mathrm{LR+}}(\lambda) \\
\iff \quad & z^* = \min_{x \in X} \left(c^\top x + \lambda^\top (Ax - b) + \rho^* \psi(Ax - b) \right) \\
\implies \quad & z^* \le c^\top x^{\varepsilon} + \lambda^\top (Ax^{\varepsilon} - b) + \rho^* \psi(Ax^{\varepsilon} - b)
\end{aligned}
$$

From Exact Penalty Parameter to Sensitivity

Let x^ε denote an optimal point of the perturbed problem

Let $\psi = \|\cdot\|_2$ and let $\rho^* < \infty$ be such that

$$
\begin{aligned}
& z^* = z^{\mathrm{LR}+}_{\rho^*}(\lambda) \\
\iff \quad & z^* = \min_{x \in X} \left(c^\top x + \lambda^\top (Ax - b) + \rho^* \psi(Ax - b) \right) \\
\implies \quad & z^* \leq c^\top x^\varepsilon + \lambda^\top (Ax^\varepsilon - b) + \rho^* \psi(Ax^\varepsilon - b) \\
\implies \quad & z^* \leq \tilde{z}^*_\varepsilon + \|\lambda\|_2 \|Ax^\varepsilon - b\|_2 + \rho^* \|Ax^\varepsilon - b\|_2
\end{aligned}
$$

From Exact Penalty Parameter to Sensitivity

Let x^ε denote an optimal point of the perturbed problem

Let $\psi = \|\cdot\|_2$ and let $\rho^* < \infty$ be such that

$$
\begin{aligned}
& z^* = z_{\rho^*}^{\mathrm{LR+}}(\lambda) \\
\iff \quad & z^* = \min_{x \in X} \left(c^\top x + \lambda^\top (Ax - b) + \rho^* \psi(Ax - b) \right) \\
\implies \quad & z^* \le c^\top x^\varepsilon + \lambda^\top (Ax^\varepsilon - b) + \rho^* \psi(Ax^\varepsilon - b) \\
\implies \quad & z^* \le \tilde{z}_\varepsilon^* + \|\lambda\|_2 \|Ax^\varepsilon - b\|_2 + \rho^* \|Ax^\varepsilon - b\|_2 \\
\implies \quad & z^* \le \tilde{z}_\varepsilon^* + (\|\lambda\|_2 + \rho^*)\varepsilon
\end{aligned}
$$

From Exact Penalty Parameter to Sensitivity

Let x^ε denote an optimal point of the perturbed problem

Let $\psi = \|\cdot\|_2$ and let $\rho^* < \infty$ be such that

$$
\begin{aligned}
& z^* = z_{\rho^*}^{\mathrm{LR}+}(\lambda) \\
\iff \quad & z^* = \min_{x \in X} \left(c^\top x + \lambda^\top (Ax - b) + \rho^* \psi(Ax - b) \right) \\
\implies \quad & z^* \leq c^\top x^\varepsilon + \lambda^\top (Ax^\varepsilon - b) + \rho^* \psi(Ax^\varepsilon - b) \\
\implies \quad & z^* \leq \tilde{z}_\varepsilon^* + \|\lambda\|_2 \|Ax^\varepsilon - b\|_2 + \rho^* \|Ax^\varepsilon - b\|_2 \\
\implies \quad & z^* \leq \tilde{z}_\varepsilon^* + (\|\lambda\|_2 + \rho^*)\varepsilon
\end{aligned}
$$

$$z^* - (\|\lambda\|_2 + \rho^*)\varepsilon \leq \tilde{z}_\varepsilon^*$$

Approximation Guarantees

Combining

- $\tilde{z}^*_\varepsilon \leq z^{\mathsf{LR+}}_\rho(\lambda) \leq z^*$, for any $\rho \geq \kappa/\varphi(\varepsilon)$
- $z^* - (\|\lambda\|_2 + \rho^*)\varepsilon \leq \tilde{z}^*_\varepsilon$

Approximation Guarantees

Combining

- $\tilde{z}_\varepsilon^* \leq z_\rho^{\mathrm{LR+}}(\lambda) \leq z^*$, for any $\rho \geq \kappa/\varphi(\varepsilon)$
- $z^* - (\|\lambda\|_2 + \rho^*)\varepsilon \leq \tilde{z}_\varepsilon^*$

- If $\psi = \|\cdot\|$, then
$$z^* - o\left(\frac{1}{\rho}\right) \leq z_\rho^{\mathrm{LR+}}(\lambda) \leq z^*$$
- If $\psi = \frac{1}{2}\|\cdot\|_2^2$, then
$$z^* - o\left(\frac{1}{\sqrt{\rho}}\right) \leq z_\rho^{\mathrm{LR+}}(\lambda) \leq z^*$$

This generalizes Theorem 2 of Feizollahi et al. (2016)

Outline

Conclusion

In this Talk

- Asymptotic zero-duality gap for MILPs
- Exact penalty parameters when using norms
- Approximation guarantees

Conclusion

In this Talk

- Asymptotic zero-duality gap for MILPs
- Exact penalty parameters when using norms
- Approximation guarantees

In our Preprint

- Extend these results to
  - Convex problems under $\sim$ Slater Constraint Qualification
  - Nonconvex problems under $\sim$ Mangasarian–Fromovitz Constraint Qualification
- Show that computing ρ for MILPs can by done in polynomial time

Conclusion

In this Talk

- Asymptotic zero-duality gap for MILPs
- Exact penalty parameters when using norms
- Approximation guarantees

In our Preprint

- Extend these results to
    - Convex problems under $\sim$ Slater Constraint Qualification
    - Nonconvex problems under $\sim$ Mangasarian–Fromovitz Constraint Qualification
- Show that computing ρ for MILPs can by done in polynomial time

In the Future

- Use these results to derive algorithms for multi-level optimization problems