Parametric Inequalities in Mixed Integer Linear Optimization

Ted Ralphs[1]
Joint work with Suresh Bolusani[2]

[1]COR@L Lab, Department of Industrial and Systems Engineering, Lehigh University [2]Zuse Institute Berlin

INFORMS Annual Meeting, 17 October 2023

Attributions

Many current and former students contributed in various ways to the development of this long line of research.

Current/Former Students/Postdocs

- Scott DeNegre
- Menal Gúzelsoy
- Anahita Hassanzadeh
- Ashutosh Mahajan
- Sahar Tahernejad
- Yu Xie
- Federico Battista

Thanks!

Outline

Outline

Overview

- In many settings, we need either to
  - ⇒ solve a sequence of related mixed integer linear optimization problems (MILPs);
  - ⇒ analyze a parametric family of MILPs; or
  - ⇒ solve a problem with multiple stages in which the later-stage problems are parameterized on the solutions to earlier-stage problems.

Examples

- Warm starting or real-time optimization
- Decomposition-based algorithms (Lagrangean relaxation, Dantzig-Wolfe)
- Parametric/multiobjective optimization
- Multistage/multilevel optimization

- Branch-and-bound algorithms themselves consist of solving a sequence of related subproblems.
- Algorithms for solving MILPs depend heavily on the generation of valid inequalities, but such inequalities are typically only valid for a single instance.
- This talk presents some ideas on making the inequalities themselves parametric.

Setting: Mixed Integer Linear Optimization

- In this talk, we consider the generation of inequalities valid for an entire parametric family of MILPs.
- Throughout, we consider a base instance of the form

$$\min_{x \in \mathcal{P} \cap X} c^\top x, \qquad \text{(MILP)}$$

where, $c \in \mathbb{Q}^n$; $\mathcal{P} = \left\{x \in \mathbb{R}^n_+ \mid Ax = b\right\}$ with $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$, and $X = \mathbb{Z}^r \times \mathbb{R}^{n-r}$.

- Here, we study the parametric family obtained by letting the right-hand side (RHS) be a parameter $\beta \in \mathbb{R}^m$, as in

$$\min_{x \in \mathcal{P}(\beta) \cap X} c^\top x, \qquad \text{(MILP-}\beta\text{)}$$

where $\mathcal{P}(\beta) = \left\{x \in \mathbb{R}^n_+ \mid Ax = \beta\right\}$.

- Much of the theory can be extended to other parameterizations.

Parametric Valid Inequalities

Parametric Valid Inequality

A *parametric valid inequality* (PVI) is a pair (α, F), where $\alpha \in \mathbb{R}^n$ and $F : \mathbb{R}^m \to \mathbb{R}$ is a function such that

$$F(\beta) \leq \min_{x \in \mathcal{P}(\beta) \cap X} \alpha^\top x \;\; \forall \beta \in \mathbb{R}^m. \qquad \text{(PVI)}$$

- Then we have

$$\alpha^\top x \geq F(\beta) \; \forall x \in \mathcal{P}(\beta) \cap X.$$

- In this setting, the PVI corresponds to a parametric family of inequalities with the same left-hand side (LHS) α, but different RHSs.
- The RHS of a PVI is a *function* of the RHS of (MILP-β).
- These inequalities are related to subadditive inequalities, but here the function F does not need to be subadditive.

Simple Example: Parametric Chvátal Inequalities

Let

$$S^{\geq}(\beta) = \mathcal{P}^{\geq}(\beta) \cap \mathbb{Z}^n,$$

where

$$\mathcal{P}^{\geq}(\beta) = \left\{x \in \mathbb{R}^n_+ \mid Ax \geq \beta\right\}$$

with $A \in \mathbb{Q}^{m \times n}$ and $b \in \mathbb{Q}^m$. Then

For

- $u \in \mathbb{R}^m_+$ such that $\alpha := A^\top u \in \mathbb{Z}^n$, and
- $F : \mathbb{R}^m \to \mathbb{R}$ defined by $F(\beta) = \lceil u^\top \beta \rceil$,

we have that (α, F) is a PVI.

Outline

Farkas Lemma

- A general theory of parametric inequalities can be derived based on duality, but here we provide a derivation beginning with the Farkas Lemma.
- Consider again the polyhedron

$$\mathcal{P} = \{x \in \mathbb{R}^n_+ \mid Ax = b\}$$

given in standard form with $A \in \mathbb{Q}^{m \times n}$, $b \in \mathbb{Q}^m$.

Farkas Lemma

$$\mathcal{P} = \emptyset \Leftrightarrow \exists u \in \mathbb{R}^m \text{ s.t. } u^\top A \leq 0, u^\top b > 0$$

- Equivalently, $\mathcal{P} = \emptyset$ if and only if we can separate b from the convex cone

$$\begin{aligned} C &= \{Ax \mid x \in \mathbb{R}^n_+\} \\ &= \{\beta \in \mathbb{R}^m \mid u^\top \beta \leq 0 \ \forall u \in C^*\}, \end{aligned}$$

where $C^* = \{u \in \mathbb{R}^m \mid u^\top A \leq 0\}$ (the *dual* of C).

Another Interpretation

- We lift the problem into a higher dimensional space by making b a vector of variables and homogenizing.

$$\mathcal{P}_\beta = \{x \in \mathbb{R}^n_+, \beta \in \mathbb{R}^m \mid Ax - I\beta = 0\}$$

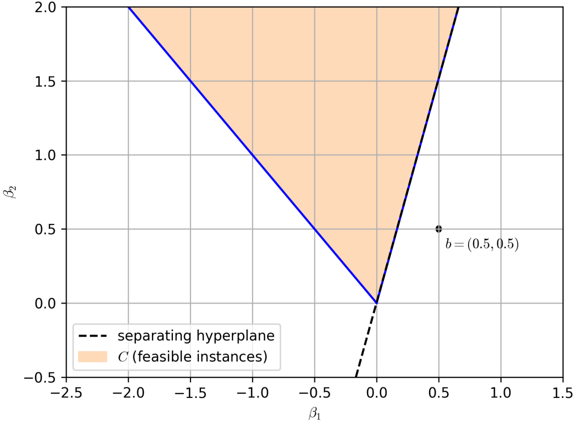
- Then project out the original variables to obtain C.

$$C = \text{Proj}_\beta(\mathcal{P}_\beta)$$

- In other words, C is just the set of values of β for which the linear system $Ax = \beta$ has a solution.
- Alternatively, C consists of the feasible members of a parametric family of linear optimization problems (LPs).
- Therefore, if we can separate b from C, we prove that $\mathcal{P} = \emptyset$ (corresponding instance is infeasible).
- The extreme rays of the dual cone correspond to the facets of C.
- These can thus be used to test feasibility of an entire parametric family of LPs.

Example 1

$$\mathcal{P}_\beta = \left\{ \begin{aligned} 2y_1 - 7y_2 + y_3 &= \beta_1 \\ 6y_1 + 7y_2 + 5y_3 &= \beta_2 \\ y_1, y_2, y_3 &\in \mathbb{R}_+ \end{aligned} \right\} \qquad C = \left\{ \begin{aligned} \beta_1 + \beta_2 &\geq 0 \\ -3\beta_1 + \beta_2 &\geq 0 \\ \beta &\in \mathbb{R}^2 \end{aligned} \right\}$$

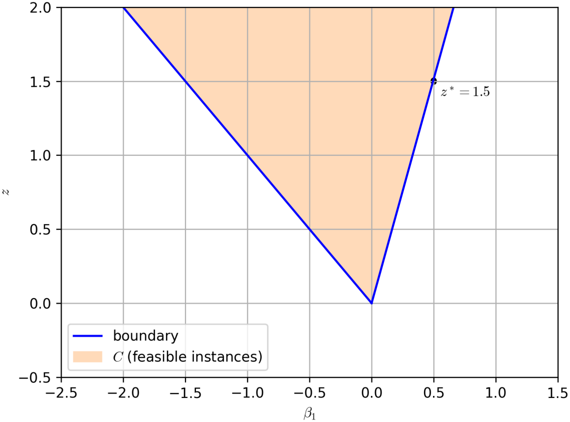


Farkas Proof of Optimality

- We now consider the case of an LP, constructed as follows.
  - Convert a^1 (coefficients of the first row of A) from a constraint to the objective.
  - Let $\bar{M} = \{2, \dots, m\}$ and $b_{\bar{M}} \in \mathbb{R}^{\bar{M}}$ be all but the first element of b.
  - The resulting LP is $\min_{x \in \mathbb{R}^n_+} \{a^1 x \mid A_{\bar{M}} x = b_{\bar{M}}\}$.
- The problem of finding the optimal value can then be recast as $b^* = \min\{b_1 \in \mathbb{R} \mid (b_1, b_{\bar{M}}) \in C\}$.
- To prove optimality, we need to show that $(b^*, b_{\bar{M}})$ is not only a member of C, but on its *boundary*.
- The LP optimality conditions imply $\exists u_{\bar{M}} \in \mathbb{R}^{\bar{M}}$ s.t. $u_{\bar{M}}^\top A_{\bar{M}} \leq a^1$, $u_{\bar{M}}^\top b_{\bar{M}} = b^*$.
- This is equivalent to $\exists u \in \mathbb{R}^m$ s.t. $u^\top A \leq 0$, $u^\top (b^*, b_{\bar{M}}) = 0$, $u_1 = -1$, implying
  - $(b^*, b_{\bar{M}})$ is on the boundary of C and
  - the boundary is one that is in the "right direction" ($u_1 < 0$).
- The vector u is a solution to the usual LP dual problem.

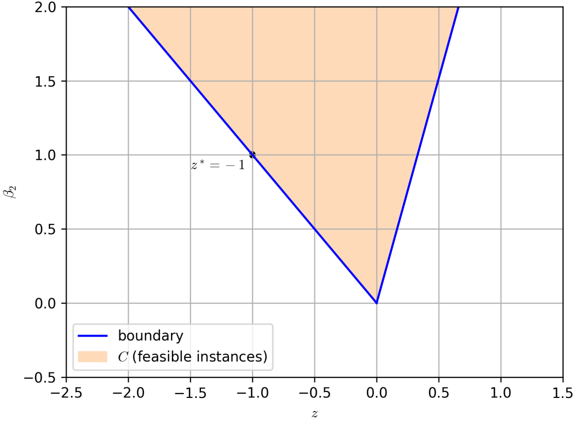
Example 2

$$z^* = \min 6y_1 + 7y_2 + 5y_3$$
$$\text{s.t. } 2y_1 - 7y_2 + y_3 = 1/2$$
$$y_1, y_2, y_3 \in \mathbb{R}_+$$



Example 3

- Note that our choice of objective was arbitrary and the same set C can yield proofs for other objectives.
- The figure shows that $\min_{y \in \mathbb{R}^3_+} \{2y_1 - 7y_2 + y_3 \mid 6y_1 + 7y_2 + 5y_3 = 1\} = -1$



Farkas Lemma for MILPs

- The very same logic extends easily to the MILP case.

$$
\begin{aligned}
\mathcal{S} &= \{x \in \mathbb{Z}^r_+ \times \mathbb{R}^{n-r}_+ \mid Ax = b\} \\
\mathcal{S}_\beta &= \{x \in \mathbb{Z}^r_+ \times \mathbb{R}^{n-r}_+, \beta \in \mathbb{R}^m \mid Ax - I\beta = 0\} \\
C &= \mathrm{Proj}_\beta(\mathcal{S}_\beta)
\end{aligned}
$$

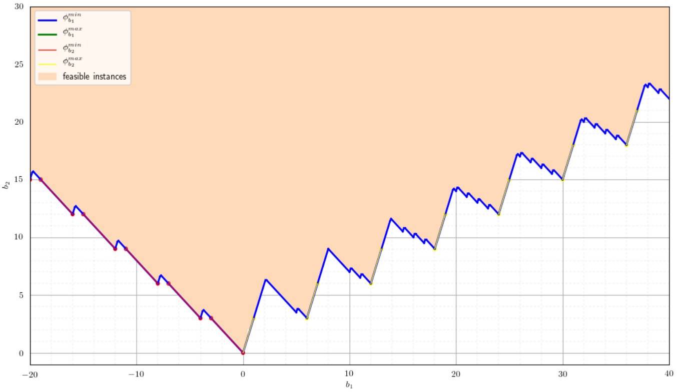
A Generalized Farkas Lemma

$$\mathcal{S} = \emptyset \Leftrightarrow b \notin C$$

- This is similar to other discrete Farkas lemmas [Bachem and Schrader, 1980, Blair and Jeroslow, 1982].
- It can be made more useful by replacing the condition $b \notin C$ with some relaxed conditions that can be verified in practice.

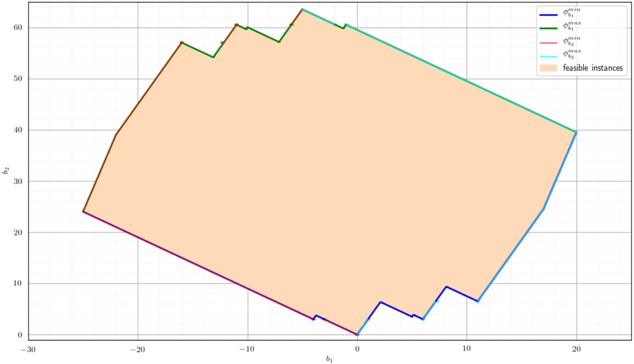
Example 4

$$\mathcal{S}_\beta = \left\{ \begin{array}{r} 6x_1 + 5x_2 - 4x_3 + 2x_4 - 7x_5 + x_6 = \beta_1 \\ 3x_1 + \frac{7}{2}x_2 + 3x_3 + 6x_4 + 7x_5 + 5x_6 = \beta_2 \\ x_1, x_2, x_3 \in \mathbb{Z}_+,\ x_4, x_5, x_6 \in \mathbb{R}_+ \end{array} \right\}$$



Example 5 (C Bounded)

$$\mathcal{S}_\beta = \left\{ \begin{aligned} 6x_1 + 5x_2 - 4x_3 + 2x_4 - 7x_5 + x_6 = \beta_1, \\ 3x_1 + \frac{7}{2}x_2 + 3x_3 + 6x_4 + 7x_5 + 5x_6 = \beta_2, \\ x_1, x_2, x_3 \in \{0,1\}, \\ 0 \le x_4, x_5, x_6 \le 3, \end{aligned} \right\}$$



Outline

Parametric Inequalities From Farkas

- Proofs of optimality are equivalent to proofs of validity for inequalities (but different parts of the boundary are relevant).
- The Farkas proof of optimality also shows that

$$a^1 x \geq b^* \ \ \forall \ \ x \in \mathcal{P}^{\bar{M}},$$

where

$$\mathcal{P}^{\bar{M}} = \{x \in \mathbb{R}^n_+ \mid A_{\bar{M}} x = b_{\bar{M}}\}.$$

- This proof of validity can be parameterized easily to obtain a PVI as follows.
- The largest valid RHSs for the first constraint as a function of the parametric RHS $\beta_{\bar{M}}$ for the remaining constraints is $\min\{\beta_1 \mid (\beta_1, \beta_{\bar{M}}) \in C\}$.
- Recalling the definition (PVI) with a^1 playing the role of α, it then follows that for $F : \mathbb{R}^{\bar{M}} \to \mathbb{R}$,

$$(a^1, F) \text{ is a PVI } \Leftrightarrow F(\beta_{\bar{M}}) \leq \min\{\beta_1 \mid (\beta_1, \beta_{\bar{M}}) \in C\}.$$

Example 4 (cont'd)

Let

$$a^1 = [6, 5, -4, 2, -7, 1]$$
$$a^2 = [3, \frac{7}{2}, 3, 6, 7, 5]$$

Then, from the figure in Example 4, we can read the following facts.

- $a^1 x \geq -5 \;\; \forall x \in \{x \in \mathbb{Z}^3_+ \times \mathbb{R}^3_+ \mid a^2 x = 4\}$.
- $a^2 x \geq 7 \;\; \forall x \in \{x \in \mathbb{Z}^3_+ \times \mathbb{R}^3_+ \mid a^1 x = 10\}$.
- $\min\{a^2 x \mid x \in \mathbb{Z}^3_+ \times \mathbb{R}^3_+, a^1 x = 3\} = 3$

All of these involve locating points on the boundary of C.

Deriving Parametric Inequalities

- The following alternative conditions for (a^1, F) to be a PVI for $F : \mathbb{R}^{\bar{M}} \to \mathbb{R}$ follow directly.

$$(a^1, F) \text{ is a PVI } \Leftrightarrow C \subseteq \text{epi}(F).$$

- Note the similarity to conditions for validity of a standard (non-parametric) valid inequality.
- Informally, this result states that a PVI is an *outer approximation* of C.
- To efficiently derive such outer approximating functions, we must understand the structure of the boundary of C.

Characterizing the Boundary of C

- Let us consider the function describing the boundary of C in the direction of minimization of β_1.
- The following function describes the boundary.

$$\phi^{\bar{M}}(\beta_{\bar{M}}) = \min\{\beta_1 \mid (\beta_1, \beta_{\bar{M}}) \in C\}$$

- On the other hand, it is straightforward to see that $\phi^{\bar{M}}$ is the *value function* of the MILP

$$\min_{x \in \mathcal{P}^{\bar{M}}} a^1 x. \qquad \text{(MILP-}\mathcal{P}^{\bar{M}}\text{)}$$

- We thus get yet another condition for (a^1, F) to be a PVI.

$$\begin{aligned} (a^1, F) \text{ is a PVI} \quad &\Leftrightarrow \quad F \text{ bounds } \phi^{\bar{M}} \text{ from below} \\ &\Leftrightarrow \quad \text{is a } \textit{dual function} \text{ for (MILP-}\mathcal{P}^{\bar{M}}\text{)} \end{aligned}$$

The General Dual and Dual Functions

- For the remainder of the talk, we focus on (MILP-β) rather than the equivalent (MILP-$\mathcal{P}^{\bar{M}}$) for simplicity.

Dual Functions

A *dual function* $F : \mathbb{R}^m \to \mathbb{R}$ for (MILP-β) is one that satisfies

$$F(\beta) \leq \phi(\beta) := \min_{x \in \mathcal{P}(\beta) \cap X} c^\top x \;\; \forall \beta \in \mathbb{R}^m.$$

- The problem of finding a dual function for which $F(b) \approx \phi(b)$ for given $b \in \mathbb{Q}^m$ is the *general dual problem* associated with (MILP-β) [Tind and Wolsey, 1981].

$$\max \{F(b) \mid F(\beta) \leq \phi(\beta),\; \beta \in \mathbb{R}^m, F \in \Upsilon^m\},$$

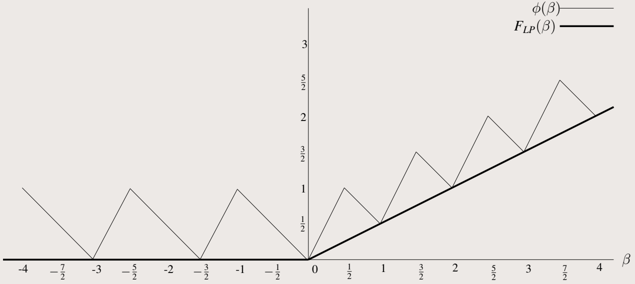
where $\Upsilon^m \subseteq \{f \mid f : \mathbb{R}^m \to \mathbb{R}\}$.

- We call F^* *strong* for this instance if F^* is a *feasible* dual function and $F^*(b) = \phi(b)$.
- This dual instance always has a solution F^* that is strong if $\phi \in \Upsilon^m$.

Dual Functions from Relaxations

- A straightforward way to derive dual functions is to take the value function of a relaxation.
- For example, the value function of the LP relaxation is

$$\phi_{LP}(\beta) = \min_{x \in \mathcal{P}(\beta)} c^\top x.$$



Dual Functions from Disjunctive Relaxation

Parametric Valid Disjunction

A *parametric valid disjunction* for (MILP-β) is a parametric family

$$X^1(\beta), X^2(\beta), \ldots, X^k(\beta) \tag{PVD}$$

of disjoint collections of subsets of $\mathbb{R}^n$ such that

$$\mathcal{P}(\beta) \cap X \subseteq \bigcup_{1 \leq i \leq k} X^i(\beta) \ \forall \beta \in \mathbb{R}^m.$$

With any parametric valid disjunction of the form (PVD), we can associate the following value function.

$$\begin{aligned} \phi_D(\beta) &= \min_{x \in \mathcal{P}(\beta) \cap (\bigcup_{1 \leq i \leq k} X^i(\beta))} c^\top x, \\ &= \min_{1 \leq i \leq k} \phi_D^i(\beta), \end{aligned}$$

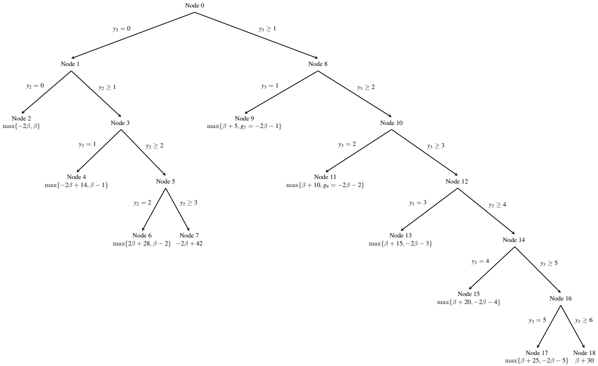
where $\phi_D^i(\beta) = \min_{x \in \mathcal{P}(\beta) \cap X^i(\beta)} c^\top x$ (value function of relaxation of subproblem i).

Disjunctions via Branch and Bound

- Branch and bound can be viewed as an algorithm for iteratively constructing and solving disjunctive relaxations.
- In the context of branch and bound, each set $X^i(b)$ corresponds to a *subproblem*

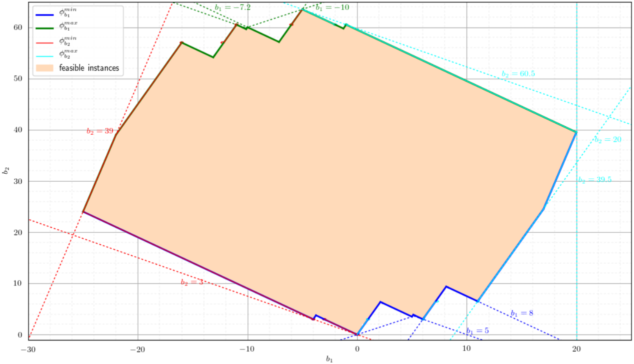
$$\min_{x \in \mathcal{P}(b) \cap X \cap X^i(b)} c^\top x$$

associated with a leaf node of the branch-and-bound tree.



Outer Approximating C

Using the machinery described so far, we can outer approximate C with dual functions obtained from, e.g., (partial) branch-and-bound trees.



More Examples

Disjunctive Inequalities

- Let $\{X^i\}_{i=1}^k$ be a disjunction valid for all instances in (MILP-β) family, where $X^i = \{x \in \mathbb{R}^n \mid D^i x \geq d^i\}$ for $D^i \in \mathbb{R}^{l^i \times n}, d^i \in \mathbb{R}^{l^i}$.
- Further, let $(u^i, v^i) \in \mathbb{R}^m \times \mathbb{R}_+^{l^i}$ be such that ${u^i}^\top A + {v^i}^\top D^i \leq \alpha^\top$.
- Then with $F(\beta) = {u^i}^\top \beta + {v^i}^\top D^i$, (α, F) is a PVI.

Lagrange Cuts

- Let $R \subseteq M := \{1, \ldots, m\}$ define a partition of the constraints into two subsets and $u \in \mathbb{R}^{|R|}$ be given.
- Then with $F(\beta) = \min_{x \in X}\{(c^\top - u^\top A_R)x \mid A_{M \setminus R} x = \beta_{M \setminus R}\}$, $(c^\top - u^\top A_R, F)$ is a PVI.

Benders' Cuts

- Benders' cuts are a particularly important special case that arises frequently in practice.
- The setup is slightly different, but the underlying procedure is the same.
- Let $T \subseteq N := \{1, \ldots, n\}$ define a partition of the variables into two subsets.
- Let $F : \mathbb{R}^m \to \mathbb{R}$ be such that

$$F(\beta) \leq \min_{x_T \in \mathcal{P}_T(\beta) \cap X^T} \alpha_T^\top x_T, \qquad \text{(MILP-Benders-}\alpha\text{)}$$

  where $\mathcal{P}_T(\beta) = \{x \in X^T \mid A_T x_T = \beta\}$ and $X^T = \{x \in \mathbb{R}^T \mid x_i \in \mathbb{Z} \text{ for } i \leq r\}$.
- Then $\alpha_T^\top x_T \geq F(b - A_{N \setminus T} x_{N \setminus T}) \ \ \forall x \in \mathcal{P}$.
- With $\alpha = c$, (α, F) is (a generalized version of) the standard Benders' cut.
- Note, however, that this cut can be applied without projecting out the variables indexed by T.

Applications

Bound inequalites in bilevel optimization [Tahernejad, 2019]

- The second-level optimality conditions are of the form

$$d^2 y \leq \phi(b^2 - A^2 x),$$

where x, y are the first- and second-level variables, d^2 is the second-level objective ector, and ϕ is the value function of the second-level problem.
- Approximating ϕ from above yields a relaxed version that is a PVI.

Parametric disjunctive inequalities [Kelly et al., 2023]

- Such inequalities are used to warm start solution of sequences of MILPs

Multiobjective optimization [Fallah et al., 2023]

- The boundary of C is very closely related to the efficient frontier of an associated *multiobjective MILP*.
- Methods for approximating C can also be used to construct the efficient frontier.

Conclusions and Future Work

- We have only scratched the surface of the theory of PVIs in this talk.
- The concept has a wide range of applications, though practical implementation is a challenge.
- They have already been applied successfully in some limited contexts.
- There are many interesting lines of research to be pursued.

Generalizations

Many generalizations are possible.

- Parametric inequalities can also be defined using similarly defined *primal functions*, which arise from *restrictions*.
- Different parameterizations can be considered.
- The theory can be "easily" extended to settings beyond MILP.

References I

A. Bachem and R. Schrader. Minimal equalities and subadditive duality. *Siam J. on Control and Optimization*, 18(4):437–443, 1980.

C.E. Blair and R.G. Jeroslow. The value function of an integer program. *Mathematical Programming*, 23:237–273, 1982.

Samira Fallah, Ted K. Ralphs, and Natashia L Boland. On the Relationship Between the Value Function and the Efficient Frontier of a Mixed Integer Linear Optimization Problem. Technical report, COR@L Laboratory, Lehigh University, 2023. URL `https://arxiv.org/abs/2303.00785`.

S. Kelly, Ralphs. T.K., and A.M. Kazachkov. Warm starting of mixed integer linear optimization problems via parametric disjunctive cuts. 2023. MIP Workshop 2023 poster competition abstract.

S. Tahernejad. *Two-stage Mixed Integer Stochastic Bilevel Linear Optimization*. PhD, Lehigh University, 2019.

J. Tind and L. Wolsey. An elementary survey of general duality theory in mathematical programming. *Mathematical Programming*, 21:241–261, 1981.