Parametric Valid Inequalities and the Solution of Multistage Optimization Problems

Ted Ralphs[1]

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

INFORMS Annual Meeting, Houston, TX, 25 October 2017

Attributions

Many students and collaborators contributed to development of this material.

Current and Former Ph.D Students

- Suresh Bolusani
- Scott DeNegre
- Menal Gúzelsoy
- Anahita Hassanzadeh
- Ashutosh Mahajan
- Sahar Tahernajad

Collaborators

- Stefano Coniglio
- Jacob Witzig
- ...

Thanks!

Outline

Outline

Overview

- In many settings, we need ether to
  - ⇒ solve a sequence of related MILPs;
  - ⇒ analyze a parametric family of MILPs; or
  - ⇒ solve a problem with multiple stages in which later-stage problems are parameterzed on the solutions to earlier stage problems.

Examples

- Decomposition-based algorithm (Lagrangean relaxation, Dantzig-Wolfe)
- Parametric optimization
- Multistage/multilevel Optimization

- Branch-and-bound algorithms themselves consist of solving a sequence of related subproblems!
- Algorithms for MILP depend heavily on the generation of valid inequalities, but such inequalities are typically only valid for a single instance.
- This talk is about some ideas about how to make the inequalities themselves parametric.

Setting: (Multi-Stage) Mixed Integer Linear Optimization

- In this talk, we'll initially consider (single-level) mixed integer linear optimization problems (MILPs).

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

where, $c \in \mathcal{R}^n$; $\mathcal{P}(\beta) = \{x \in \mathbb{R}^{n_1} \mid Ax \geq \beta\}$ with $A \in \mathbb{Q}^{m \times n}$, $b \in \mathcal{R}^m$; and $X = \mathbb{Z}^r_+ \times \mathbb{R}^{n-r}_+$.

- We'll later consider a multi-stage generalization.

MSMILP

$$\min_{x \in \mathcal{P}(b) \cap X} \left\{ c^\top x + \Xi(x) \right\}, \qquad \text{(MSMILP)}$$

where Ξ is the *risk function* that represents the impact of future uncertainty due to stochasticity or other unknowns.

Outline

Outline

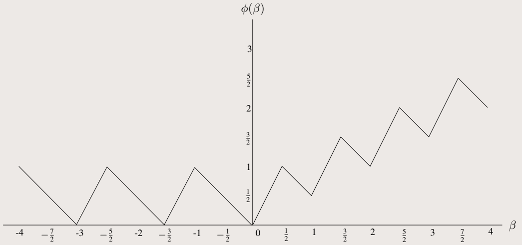
The Value Function

- The *value function* associated with (MILP) is

MILP Value Function

$$\phi(\beta) = \min_{x \in \mathcal{P}(\beta) \cap X} c^\top x \tag{VF}$$

for $\beta \in \mathcal{R}^m$. We let $\phi(\beta) = \infty$ if $\beta \in \Omega = \{\beta \in \mathbb{R}^m \mid \mathcal{S}(\beta) = \emptyset\}$.



The General Dual and Dual Functions

Dual Functions

A *dual function* $F : \mathbb{R}^m \rightarrow \mathbb{R}$ is one that satisfies $F(\beta) \leq \phi(\beta)$ for all $\beta \in \mathbb{R}^m$.

- The problem of finding a dual function for which $F(b) \approx \phi(b)$ is the *general dual problem* associated with (MILP).

$$\max \{F(b) : F(\beta) \leq \phi(\beta),\ \beta \in \mathcal{R}^m, F \in \Upsilon^m\} \qquad \text{(D)}$$

  where $\Upsilon^m \subseteq \{f \mid f : \mathcal{R}^m \rightarrow \mathcal{R}\}$

- 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 $\phi \in \Upsilon^m$

Outline

Relaxations

Relaxation

A relaxation of (MILP) is an optimization problem

$$\min_{x \in \mathcal{R}} f(x)$$

such that $\mathcal{R} \supseteq (\mathcal{P}(b) \cap X)$ and $f(x) \leq c^\top x \ \forall x \in \mathcal{P}(b) \cap X$.

- The linear optimization problem (LP) obtained by relaxing the requirement $x \in X$ is the *LP relaxation* of (MILP)

$$\min_{x \in \mathcal{P}(b)} \ c^\top x, \qquad \text{(LPR)}$$

- Solving the LP relaxation is the first step in many/most algorithms for solving (MILP).

Disjunctive Relaxations

Valid Disjunction

A *valid disjunction* for (MILP) is a disjoint collection

$$X_1(b), X_2(b), \ldots, X_k(b) \tag{VD}$$

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

$$\mathcal{P}(b) \cap X \subseteq \bigcup_{1 \le i \le k} X_i(b)$$

- Any valid disjunction can be added to the LP relaxation to obtain a stronger relaxation of (MILP).

$$\min_{x \in \mathcal{P}(b) \cap (\bigcup_{1 \le i \le k} X_i(b))} c^\top x, \tag{DR}$$

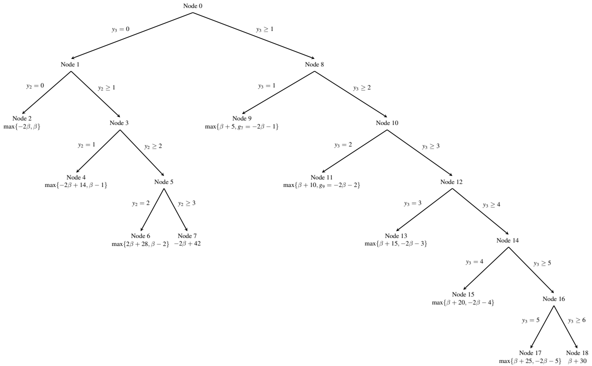
- Thus, the branch-and-bound tree encodes a relaxation of (MILP).
- We can also derive such disjunctions from problems structure (integrality).

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 sets $X_i(b)$ corresponds to a *subproblem*

$$\min_{x \in \mathcal{P}(b) \cap X \cap X_i(b)} c^\top x \tag{1}$$

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

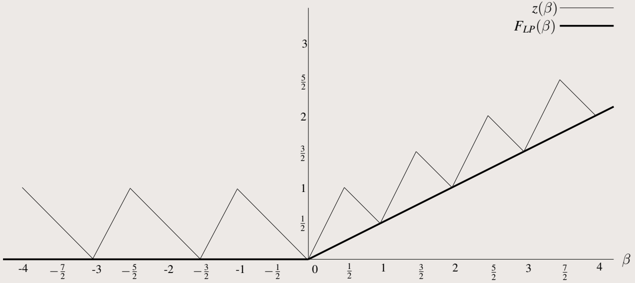


Dual Functions from Relaxations

- The value function of any relaxation is a valid dual function.
- The value function of the LP relaxation is

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

- To define the value function of (DR), we need the notion of a *parametric valid disjunction*, which we discuss next.



Dual Functions from Disjunctive Relaxation

Parametric Valid Disjunction

A *parametric valid disjunction* for (MILP) is a parametric family of disjoint collections

$$X_1(\beta), X_2(\beta), \ldots, X_k(\beta) \tag{PVD}$$

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

$$\mathcal{P}(\beta) \cap X \subseteq \bigcup_{1 \leq i \leq k} X_i(\beta)$$

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

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

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).

Outline

The Gauge Function

The *gauge function* associated with (MILP) is a function that returns the largest valid right-hand side for an inequality with left-hand side vector α.

MILP Gauge Function

$$\Gamma(\zeta) = \min_{x \in \mathcal{P}(b) \cap X} \zeta^\top x \; \forall \alpha \in \mathbb{R}^n \tag{GF}$$

Note that we have

$$\alpha x \geq \Gamma(\alpha) \;\; \forall x \in \mathcal{P}(b)$$

Valid Inequality

An *inequality* defined by $(\alpha, \eta) \in \mathbb{Q}^n \times \mathbb{Q}$ is *valid* for $\mathcal{P}(b)$ if $\eta \leq \Gamma(\alpha)$.

Inequalities from Relaxations

- Valid inequalities are often obtained more efficiently by considering a relaxation.
- Let $\mathcal{R} \supseteq \mathcal{P}(b)$ be the feasible set of a relaxation whose gauge function is

$$\Gamma_{\mathcal{R}}(\zeta) = \min_{x \in \mathcal{R}} \zeta^{\top} x \tag{RGF}$$

- Then $(\alpha, \Gamma_{\mathcal{R}}(\alpha))$ is valid for $\mathcal{P}(b)$.

Disjunctive Inequalities

- Let $\mathcal{R} = \mathcal{P}(\beta) \cap \left(\bigcup_{1 \leq i \leq k} X_i(\beta)\right)$ be the feasible set of a disjunctive relaxation.
- Then $(\alpha, \Gamma_{\mathcal{R}}(\alpha))$ is valid for $\mathcal{P}(b)$, where $\Gamma_{\mathcal{R}}$ is the gauge function (RGF).
- Inequalities derived in this way are known as *disjunctive inequalities*.
- Many known classes of valid inequalities can be derived by applying this framework to the relaxation (DR) for various classes of disjunction.

Outline

Parametric Valid Inequalities

Parametric Valid Inequality

A *parametric valid inequality* is a pair (α, F), where $\alpha \in \mathbb{R}^n$ and F is a dual function for the MILP

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

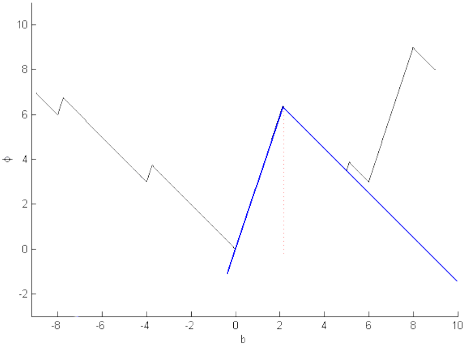
- The right-hand side of a parametric valid inequality is a *function*.
- The parametric inequality corresponds to a parametric family of inequalities with the same left-hand side.
- We have

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

- Note that (α, ϕ_α) is a valid parametric inequality, where ϕ_α is the value function of (MILP-α).
- These inequalities are related to subadditive inequalities.

Dual Functions from Branch and Bound

- Recall that a *dual function* $F : \mathbb{R}^m \to \mathbb{R}$ is one that satisfies $F(\beta) \leq \phi(\beta)$ for all $\beta \in \mathbb{R}^m$.
- Observe that any branch-and-bound tree yields a lower approximation of the value function.



Dual Functions from Branch-and-Bound

Let T be set of the terminating nodes of the tree. Then in a terminating node $t \in T$ we solve:

$$\begin{aligned} \phi^t(\beta) = \min\ & c^\top x \\ \text{s.t. } & Ax = \beta, \\ & l^t \leq x \leq u^t, x \geq 0 \end{aligned} \tag{2}$$

The dual at node t:

$$\begin{aligned} \phi^t(\beta) = \max\ & \{\pi^t \beta + \underline{\pi}^t l^t + \bar{\pi}^t u^t\} \\ \text{s.t. } & \pi^t A + \underline{\pi}^t + \bar{\pi}^t \leq c^\top \\ & \underline{\pi} \geq 0, \bar{\pi} \leq 0 \end{aligned} \tag{3}$$

We obtain the following strong dual function:

$$\min_{t \in T} \{\hat{\pi}^t \beta + \hat{\underline{\pi}}^t l^t + \hat{\bar{\pi}}^t u^t\}, \tag{4}$$

where $(\hat{\pi}^t, \hat{\underline{\pi}}^t, \hat{\bar{\pi}}^t)$ is an optimal solution to the dual (3).

Outline

The Risk Function

Recall problem (MSMILP) from earlier. The general class of functions we consider are of the following form.

Canonical Risk Function

$$\Xi(x) = \mathbb{E}_{\omega \in \Omega}\left[\Xi_\omega(x)\right], \tag{RF}$$

Scenario Risk Function

$$\Xi_\omega(x) = \min\left\{d^1 y \mid y \in \operatorname{argmin}\{d^2 y \mid y \in \mathcal{P}_2(A^2_\omega x) \cap Y\}\right\} \tag{2LRF}$$

- ω is a random variable over a finite probability space $(\Omega, \mathscr{F}, \mathcal{P})$;
- Realizations of ω are called *scenarios*;
- $\mathcal{P}_2(\beta) = \{y \in \mathbb{R}^{n_2}_+ \mid G^2 y \geq A^2_\omega x\}$;
- $Y = \mathbb{Z}^{r_2}_+ \times \mathbb{R}^{n_2 - r_2}_+$; and
- $G^2 \in \mathbb{Q}^{m_2 \times n_2}$, $A^2_\omega \in \mathbb{Q}^{m_2 \times n_1}$ for $\omega \in \Omega$.

Value Function Reformulation

One way to reformulate (MSMILP) as a single-level problem is as follows.

$$
\begin{aligned}
\min \quad & c^1 x + \sum_{\omega \in \Omega} p_\omega d^1 y^\omega && (5)\\
\text{subject to} \quad & A^1 x \geq b^1 && (6)\\
& G^2 y^\omega \geq A^2_\omega x \;\forall \omega \in \Omega && (7)\\
& d^2 y^\omega \leq \phi(A^2_\omega x) \;\forall \omega \in \Omega && (8)\\
& x \in X && (9)\\
& y^\omega \in Y \;\forall \omega \in \Omega && (10)
\end{aligned}
$$

where ϕ is the value function of the second-stage problem. This is, in principle, a standard mathematical optimization problem.

Bound Inequalities

- One relaxation is to drop the value function constraint (8) and replace it with the linear inequality

$$d^2 y \leq \eta,$$

where η is such that

$$\eta \geq \max_{(x,y)\in\mathcal{F}} \quad d^2 y, \tag{11}$$

where F is the feasible region of the bilevel problem.

- One way to derive η is to solve the problem on the left-hand side of (11) by branch-and-bound.
- We can stop the branch and bound after a given time limit and extract a dual function from the tree.
- With the dual function, we may even be able to tighten the inequality after branching.

Conclusions

- The concept of a parametric inequality has a wide range of applications, of which we only scraped the surface in this talk.
- Parametric inequalities can also be defined using similarly defined *primal functions*, which arise from *restrictions*.
- They have already been applied successfully in some limited contexts.
- If we can generate them effectively, this will open up many interesting lines of research.