(Three) Cutting Planes for Signomial Programming

Liding Xu
Claudia D'Ambrosio, Sonia Haddad Vanier, Leo Liberti

OptimiX, LIX, École Polytechnique

April, 2022

Table of Contents

Problem definition: signomial terms

- Exponent vector: $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{R}^n$.
- Signomial term: $x^\alpha := \prod_{i=1}^n x_i^{\alpha_i}$, for $x \in \mathbb{R}^n_{++}$.
- Examples: $x_1^{-1} x_2^1, \quad x_1^{0.5} x_2^{-10} x_3^{1.2}$.
- When $\alpha \in \mathbb{Z}^n_+$, monomial term.
- Nonconvex.

Problem definition: Signomial Programming (SP)

$$
\begin{aligned}
\min \quad & c \cdot x & \text{(1a)} \\
\forall j \in [1:m] \quad & \sum_{k \in \mathcal{K}_j} a_{jk} x^{\alpha^k} \leq 0 & \text{(1b)} \\
\forall i \in [1:n] \quad & x_i \in [\underline{x}_i, \overline{x}_i] \subset \mathbb{R}_{++} & \text{(1c)}
\end{aligned}
$$

Solution methods of global solvers

An example.

- For example, $x_1^{1.9} x_2^{-0.1} x_3^4$.
- Introduce a lifting variable, $y = x_1^{1.9} x_2^{-0.1} x_3^4$;
- Introduce auxiliary variables $y_1 = x_1^{1.9}$, $y_2 = x_2^{-0.1}$, $y_3 = x_3^4$;

An example.

- For example, $x_1^{1.9}x_2^{-0.1}x_3^4$.
- Introduce a lifting variable, $y = x_1^{1.9}x_2^{-0.1}x_3^4$;
- Introduce auxiliary variables $y_1 = x_1^{1.9}$, $y_2 = x_2^{-0.1}$, $y_3 = x_3^4$;
- Construct convex relaxations for $y_1 = x_1^{1.9}$, $y_2 = x_2^{-0.1}$, $y_3 = x_3^4$;
- Relax multilinear constraint $y = y_1y_2y_3$.

Normalized formulation of signomial sets

W.l.o.g., we consider the signomial set as the hypograph set

$$\mathcal{S}_{\mathrm{s}} = \{(x, y) \in \mathbb{R}^n_+ \times \mathbb{R}_+ : y \leq x^{\alpha}\}.$$

α can have negative or positive entries.

Normalized formulation of signomial sets

W.l.o.g., we consider the signomial set as the hypograph set

$$\mathcal{S}_{\mathrm{s}} = \{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+ : y \leq x^\alpha\}.$$

α can have negative or positive entries. Decompose

$$\mathcal{S}_{\mathrm{s}} = \{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+ : y \leq x_-^{\alpha^-} x_+^{\alpha^+}\}.$$

Normalized formulation of signomial sets (continued)

Rearrange negative/positive power terms

$$\mathcal{S}_{\mathrm{s}} = \{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+ : y x_-^{-\alpha^-} \leq x_+^{\alpha^+}\}.$$

Normalized formulation of signomial sets (continued)

Rearrange negative/positive power terms

$$\mathcal{S}_{\mathrm{s}} = \{(x, y) \in \mathbb{R}_+^n \times \mathbb{R}_+ : yx_-^{-\alpha^-} \le x_+^{\alpha^+}\}.$$

Change of variables

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathbb{R}_+^h \times \mathbb{R}_+^l : u^\beta \le v^\gamma\},$$

$\beta > 0, \gamma > 0.$

Normalized formulation of signomial set

Scale β, γ such that $\max(\|\beta\|_1, \|\gamma\|_1) \leq 1$,

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathbb{R}_+^h \times \mathbb{R}_+^l : \; u^\beta \leq v^\gamma\}$$

Denote by $\psi^\alpha(x) = x^\alpha$,

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathbb{R}_+^h \times \mathbb{R}_+^l : \; \psi^\beta(u) - \psi^\gamma(v) \leq 0\}$$

A lift of nice properties

$$\mathcal{S}_{\rm s} = \{(u, v) \in \mathbb{R}^h_+ \times \mathbb{R}^l_+ : \psi^\beta(u) - \psi^\gamma(v) \le 0\}$$

- ψ^β, ψ^γ are concave, $\psi^\beta(u) - \psi^\gamma(v)$ is a difference-of-concave (DCC) function;
- If $\max(\|\beta\|_1, \|\gamma\|_1) = 1$, at least one of them is positively-homogeneous of degree-1;
- Assume both u, v are in box constraints, then ψ^β, ψ^γ are supermodular.

Intersection cuts: a standard procedure

Definition

Given $\mathcal{S} \in \mathbb{R}^p$, a closed set $\mathcal{C}$ is called $\mathcal{S}$-free, if the following conditions are satisfied:

1. $\mathcal{C}$ is convex;
2. $\text{inter}(\mathcal{C}) \cap \mathcal{S} = \emptyset$.

Intersection cut

Intersection cut is a framework.

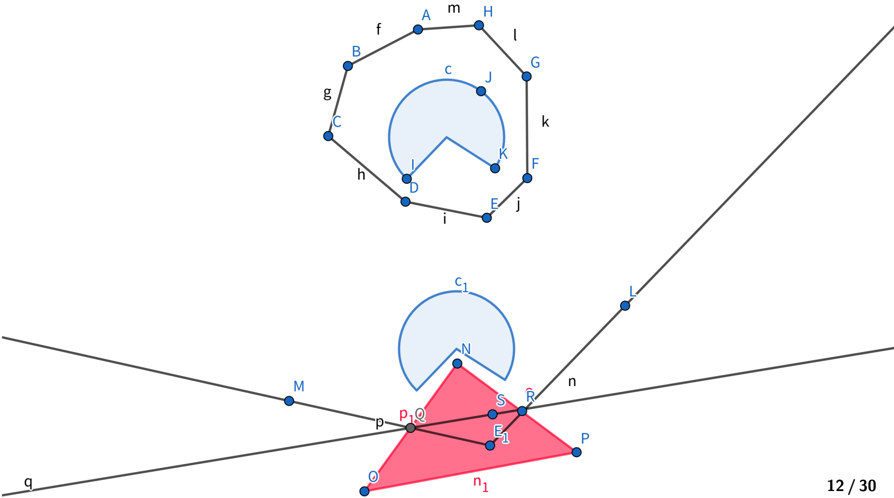
- Given a non-convex set $\mathcal{S}$, an $\mathcal{S}$-free set, and a corner polyhedron $\mathcal{P}$ containing $\mathcal{S}$.
- **Separation** (Zambelli et al., Integer Programming): intersect the corner polyhedron with the set $\mathcal{C}$.

Intersection cut

Intersection cut is a framework.

- Given a non-convex set $\mathcal{S}$, an $\mathcal{S}$-free set, and a corner polyhedron $\mathcal{P}$ containing $\mathcal{S}$.
- **Separation** (Zambelli et al., Integer Programming): intersect the corner polyhedron with the set $\mathcal{C}$.
- Nonlinear programming (Tuy 1964), $\mathcal{S}$ is the epigraph of a concave function;
- Recent development in MINLPs: outer product sets, bilinear sets, and quadratic sets.

The Geometry of intersection cuts



S-free sets of level sets in NLP

Theorem (Khamisov 1999,Serrano 2019)

Assume $\mathcal{S} := \{z \in \mathbb{R}^p : f_1(z) - f_2(z) \leq 0\}$, *where* f_1, f_2 *are concave functions. Then, for* $\check{z} \in \mathrm{dom}(f_2)$, $\mathcal{C}_{\hat{z}} := \{z \in \mathbb{R}^p : f_1(z) - f_2(\hat{z}) - \nabla f_2(\hat{z})^\top (z - \hat{z}) \geq 0\}$ *is a* $\mathcal{S}$*-free set.*

Theorem (Serrano 2021)

Assume $\mathcal{S} := \{z \in \mathbb{R}^p : f_1(z) - f_2(z) \leq 0\}$, *where* f_1, f_2 *are concave functions and positive-homogeneous of degree-1. Then, for* $\check{z} \in \mathrm{dom}(f_2)$, $\mathcal{C}_{\hat{z}} := \{z \in \mathbb{R}^p : f_1(z) - f_2(\hat{z}) - \nabla f_2(\hat{z})^\top (z - \hat{z}) \geq 0\}$ *is a maximal* $\mathcal{S}$*-free set.*

Signomial-free sets

Corollary

$\{(u, v) \in \mathbb{R}^h_+ \times \mathbb{R}^l_+ : \psi^\beta(u) - \psi^\gamma(\tilde{v}) - \nabla(\psi^\gamma(\tilde{v})) \cdot (v - \tilde{v}) \geq 0\}$ is a signomial-free set.

Assume $\max(\|\beta\|_1, \|\gamma\|_1) = 1$ (one function is positive-homogeneous of degree-1), we prove a stronger result

Theorem

$\{(u, v) \in \mathbb{R}^h_+ \times \mathbb{R}^l_+ : \psi^\beta(u) - \psi^\gamma(\tilde{v}) - \nabla\psi^\gamma(\tilde{v}) \cdot (v - \tilde{v}) \geq 0\}$ is a maximal signomial-free set.

Signomial set on box constraint

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathbb{R}^h_+ \times \mathbb{R}^l_+ : \psi^\beta(u) \le \psi^\gamma(v)\},$$

Recall that $\psi^\beta(u), \psi^\gamma(v)$ are concave.

Signomial set on box constraint

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathbb{R}_+^h \times \mathbb{R}_+^l : \psi^\beta(u) \le \psi^\gamma(v)\},$$

Recall that $\psi^\beta(u), \psi^\gamma(v)$ are concave.
We assume $u \in \mathcal{U} := [\underline{u}, \overline{u}]$ is a box constraint. Consider now

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathcal{U} \times \mathbb{R}_+^l : \psi^\beta(u) - \psi^\gamma(v) \le 0\}.$$

Convex under-estimators

Constructing convex under-estimators of $\psi^{\beta}(u)$?

- Using supermodularity: supermodular relaxation.
- Factorization and relaxation: factorization relaxation.

Supermodular relaxation: supermodular functions

Definition

Given $D = \prod_{1 \leq i \leq n} D_i$ ($D_i \subset \mathbb{R}$), a function $f : D \to \mathbb{R}$ is supermodular on D, if for every $x, y \in D$,
$f(x) + f(y) \leq f(\max\{x, y\}) + f(\min\{x, y\})$.

Lemma

ψ^β is supermodular on $\mathcal{U}$.

Supermodular relaxation: transformation

Definition

Define

$$g : [0,1]^h \to \mathbb{R} : w \to g(w) := \prod_{1 \leq j \leq h} ((\overline{u_j} - \underline{u_j})w_j + \underline{u_j})^{\beta_j} - \underline{u}^{\beta},$$

$$\pi : [0,1]^h \to \mathcal{U} : w \to \pi(w) := ((\overline{u_1} - \underline{u_1})w_1 + \underline{u_1}, \cdots, (\overline{u_h} - \underline{u_h})w_h + \underline{u_h}),$$

and

$$\pi^{-1} : \mathcal{U} \to [0,1]^h : u \to \pi^{-1}(u) := (\frac{u_1 - \underline{u_1}}{\overline{u_1} - \underline{u_1}}, \cdots, \frac{u_h - \underline{u_h}}{\overline{u_h} - \underline{u_h}}).$$

Supermodular relaxation: after transformation

Theorem

The transformed function g is concave and supermodular on $[0, 1]^h$, and convex-extendable from vertices.

Supermodular relaxation: affine underestimating functions for g

Theorem

Let $H := \{1, \cdots, h\}$, let $\chi : 2^H \to \{0,1\}^H$ be the indicator function over subsets of H, and define $\rho(j,S) := g(\chi_{S\cup\{j\}}) - g(\chi_S)$ ($j \in H, S \subseteq H$) the increment function of g. Then,

$$g(\chi_S) + \sum_{j\in H\setminus S} \rho(j,S)w_j - \sum_{j\in S} \rho(j, N \setminus \{j\})(1-w_j) \leq g(w),$$
$$g(\chi_S) + \sum_{j\in H\setminus S} \rho(j,\varnothing)w_j - \sum_{j\in S} \rho(j, S \setminus \{j\})(1-w_j) \leq g(w), \quad S \subseteq H, \tag{2}$$

Not an envelope. Separation can be done by a heuristic (Nemhauser 79).

Supermodular relaxation: the formulation

The proposition leads to the `Supermodular Relaxation`:

$$\mathcal{S}_{\text{sup}} = \left\{ (u, v) \in \mathcal{U} \times \mathbb{R}^{I} : \left(\pi^{-1}(u), \psi^{\gamma}(v) - \psi^{\beta}(\underline{u}) \right) \text{ satisfies (2)} \right\}.$$

Factorization relaxation

$$\mathcal{S}_{\mathrm{s}} = \{(u, v) \in \mathcal{U} \times \mathbb{R}^{l}_{+} : \psi^{\beta}(u) - \psi^{\gamma}(v) \leq 0\}.$$

Theorem

Given the power function $\psi^\beta(u)$, let $\bar\beta \in \mathbb{R}_+^{h+1}$ satisfy that $\bar\beta_{[h]} = \beta_{[h]}$ and $\bar\beta_0 = 1 - \sum_{j\in[h]} \beta_j$, let $\zeta \in \mathbb{R}_+^h$ satisfy that $\zeta_j = \bar\beta_j / (\sum_{i\in[0:j]} \bar\beta_i)$. Denote

$$\mathcal{E}_\beta := \{(u,t) \in \mathbb{R}_+^h \times \mathbb{R} : \exists s \in \mathbb{R}_+^{h+1}\ s_{h+1} = t\ s_1 = 1 \quad \forall j \in [0:h]\ u_j^{\zeta_j} s_j^{1-\zeta_j} \leq s_{j+1}\}.$$

Then, $\operatorname{epi}_{\mathbb{R}_+^h}(\psi^\beta) = \mathcal{E}_\beta$.

Factorization relaxation: convexification

$$\mathcal{E}_\beta := \{(u,t) \in \mathbb{R}^h_+ \times \mathbb{R} : \exists s \in \mathbb{R}^{h+1}_+ \; s_{h+1} = t \; s_1 = 1 \\ \forall j \in [0:h] \; u_j^{\zeta_j} s_j^{1-\zeta_j} \leq s_{j+1}\}.$$

Compute bounds on s, replace the concave term $u_j^{\zeta_j} s_j^{1-\zeta_j}$ by its convex envelope f^j (a piece-wise function) over the box constraint.

$$\mathcal{E}_\beta := \{(u,t) \in \mathbb{R}^h_+ \times \mathbb{R} : \exists s \in \mathbb{R}^{h+1}_+ \; s_{h+1} = t \; s_1 = 1 \\ \forall j \in [0:h] \; f^j(u_j, s_j) \leq s_{j+1}\}.$$

Factorization relaxation: the extended formulation

Get rid of multilinear terms.
A convex relaxation in an extended formulation

$$\mathcal{S}_{\mathrm{lc}} := \{(u, v) \in \mathcal{U} \times \mathbb{R}^l : \exists s \in S \; \psi^\gamma(v) \geq s_{h+1} \; s_1 = 1 \\ \forall j \in [0 : h] \; f^j(u_j, s_j) \leq s_{j+1}\}.$$

Factorization relaxation: projection

Define recursively $F^j(u) := f^j(u_j, f^{j-1}(u_{j-1}, \cdots))$ ($j \in [0:h]$). F^h is a polyhedral convex function.

$$\mathcal{S}_{\mathrm{lc}} = \{(u, v) \in \mathcal{U} \times \mathbb{R}^l : \psi^\gamma(v) \geq F^h(u)\}.$$

The gradient of F^h can be computed in a linear time.

Development environment

- Software: **SCIP** 8.0.0, CPLEX 22.1, and Ipopt 3.14.7.
- Hardware: Intel Xeon W-2245 CPU @ 3.90GHz, 126GB main memory.
- Data: From MINLPLib, **C** (68 continuous instances), **MI** (189 mixed-integer instances), and **All**.

Numerical results

- Default: SCIP's default;
- ICUT: only intersection cuts;
- SOCUT: only outer approximation cuts from supermodular relaxation;
- POCUT: only outer approximation cuts from facotrization relaxations;

Closed root gap function: 0-100%, the larger, the better.

Numerical results

Benchmark	Default		ICUT			SOCUT			POCUT		
	solved	closed	solved	closed	relative	solved	closed	relative	solved	closed	relative
C-clean	11/68	0.38	13/68	0.47	1.24	11/68	0.4	1.06	13/67	0.41	1.09
C-affected	4/47	0.4	6/47	0.53	1.34	4/47	0.43	1.08	6/46	0.45	1.13
MI-clean	12/189	0.32	12/189	0.35	1.1	16/187	0.33	1.01	11/183	0.33	1.04
MI-affected	7/120	0.44	7/120	0.49	1.12	11/118	0.45	1.02	6/114	0.46	1.06
All-clean	23/257	0.34	25/257	0.38	1.14	27/255	0.35	1.03	24/250	0.35	1.05
All-affected	11/167	0.43	13/167	0.5	1.18	15/165	0.44	1.04	12/160	0.46	1.08

Table: Summary of the closed root gaps and relative improvement to the Default setting

- Intersection cuts are strongest;
- Supermodular outer approximation is weak;
- Facotrization outer approximation is a good alternative to the conventional factorization.

Future development

- Implementing a signomial term handler in a solver may be useful, to detect signomials, reformulate (eliminate intermediate variables), cut (initial LP estimation + separation);
- (Unexploited): both sides of the DCC formulation are monotone, useful for propagation and presolving?