Approximating ILP Core Points with Nonlinear Constraints

David Bremner[1] Naghmeh Shahverdi[2]

[1]University of New Brunswick, Canada

[2]Stanford University, USA

2023-07-07

But now I am six[ty], I'm as clever as clever, So I think I'll be six[ty] now for ever and ever – A.A. Milne

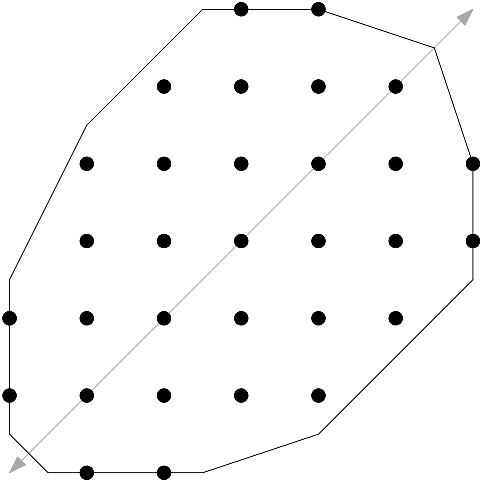
Symmetric Integer Linear Programs

Invertible matrix $g \in \mathrm{GL}_n(\mathbb{Z})$ is a (formulation) symmetry for ILP

$$\max\{ \langle \gamma, x \rangle \mid x \in P \cap \mathbb{Z}^n \}$$

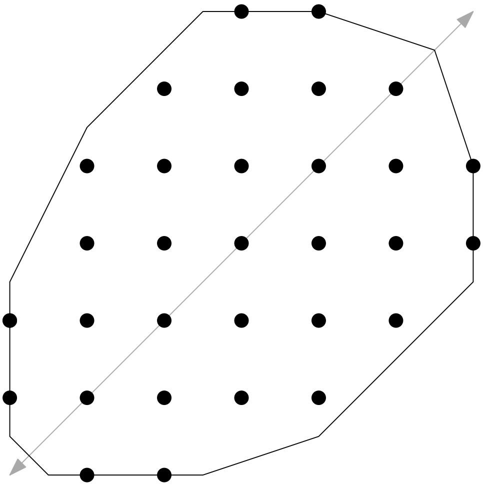
if $\forall x \in P$,

$$gx \in P \text{ and } \langle \gamma, gx \rangle = \langle \gamma, x \rangle$$



Symmetric Integer Linear Programs

208 of the 361 MIPLIB instances have a non-trivial symmetry group. (Rehn 2014)



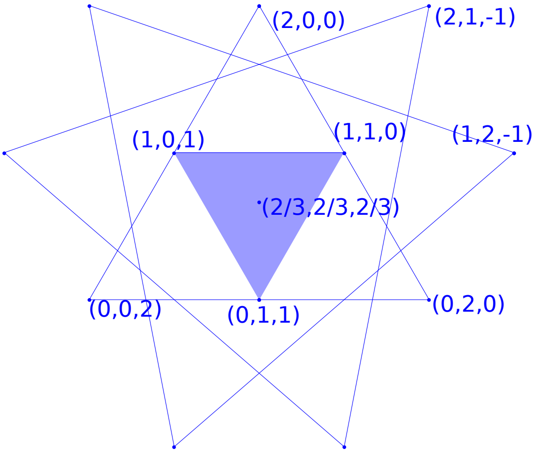
Core points

$$G \leq \mathrm{GL}_n(\mathbb{Z})$$
$$Gz = \{ gz \mid g \in G \}$$

Definition

$z \in \mathbb{Z}^n$ is called a core point for G if $(\operatorname{conv} Gz) \cap \mathbb{Z}^n = Gz$.

$\langle(1,2,3)\rangle$ acting on $\{x \mid \mathbf{1}^T x = 2\}$



Core points

$$G \leq \mathrm{GL}_n(\mathbb{Z})$$
$$Gz = \{\, gz \mid g \in G \,\}$$

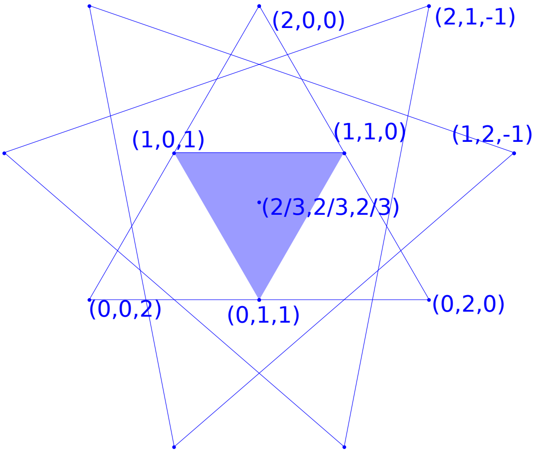
Definition

$z \in \mathbb{Z}^n$ is called a core point for G if $(\operatorname{conv} Gz) \cap \mathbb{Z}^n = Gz$.

Lemma

If G-symmetric convex set K contains an integer point, K contains a core point.

$\langle (1,2,3) \rangle$ acting on $\{x \mid \mathbf{1}^T x = 2\}$



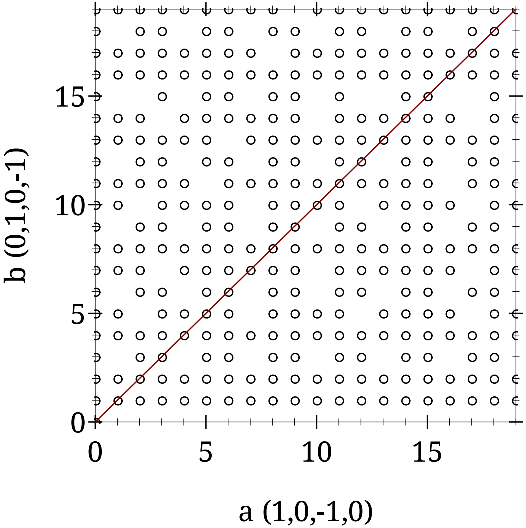
Core points for cyclic groups

$$\sigma \begin{bmatrix} c_0 \\ c_1 \\ \dots \\ c_{n-1} \end{bmatrix} = \begin{bmatrix} c_{n-1} \\ c_0 \\ \dots \\ c_{n-2} \end{bmatrix}$$

$$\mathcal{C}_n = \langle \sigma \rangle$$

Lemma (Rehn 2014)

For $a, b \in \mathbb{Z}$ *such that* $\gcd(2a+1, 2b) = 1$, $(1+a, b, -a, -b) \in \mathbb{Z}^4$ *is a core point for* $\mathcal{C}_4$.



Outer approximations of core points

Goal Find "useful" f s.t. $f(c) < 0$ for all core points c.

Example If G has exactly two invariant subspaces, core point $c \in \mathbb{R}^n$ with $\mathbf{1}^T c = k$

$$\operatorname{dist}(c, \operatorname{Fix}(G)) \leq (n-1)\sqrt{k(n-k)/n}\,. \qquad \text{(Rehn)}$$

where

$$\operatorname{Fix}_{\mathbb{R}}(G) := \{x \in \mathbb{R}^n \mid gx = x, \forall g \in G\}$$

"Approximately core" points

- For all core points c, $\forall z \in \mathbb{Z}^n \cap (\text{aff } Gc \setminus Gc)$, $z \notin \text{conv } Gc$

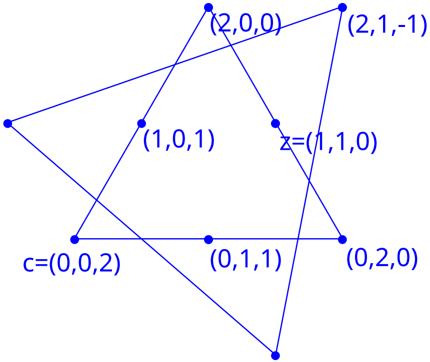
$$\lambda = (Gc)^{-1} z \not\geq 0$$

- For $E \subset \mathbb{Z}^n \setminus Gc$, we say that c is E-core if $E \cap \text{conv } Gc = \emptyset$.
- If $f(z) < 0$ for all E-core points z then $f(c) < 0$ for all core points c.

Orbits and circulant matrices

For $c \in \mathbb{R}^n$, $\mathrm{Cir}(c)$ is the $n \times n$ circulant matrix with column k equal $\sigma^k(c)$

$$\mathrm{Cir}(c) = \begin{bmatrix} c_0 & c_{n-1} & \cdots & c_2 & c_1 \\ c_1 & c_0 & c_{n-1} & \cdots & c_2 \\ \vdots & c_1 & c_0 & \ddots & \vdots \\ c_{n-2} & & \ddots & \ddots & c_{n-1} \\ c_{n-1} & c_{n-2} & \cdots & c_1 & c_0 \end{bmatrix}$$



$$z = \mathrm{Cir}(c)\lambda$$

$$\lambda = \mathrm{Cir}(c)^{-1} z$$

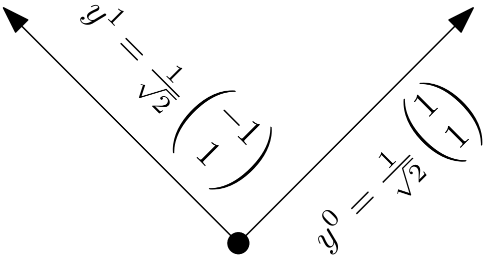
Eigenvectors of circulant matrices

$$\mathrm{Cir}(c)y^m = \psi_m y^m \qquad \text{(eigen)}$$
$$w_n^n = 1 \qquad \text{(unity)}$$

$$w_n = e^{\frac{2\pi i}{n}}$$

$$y_k^m = n^{-1/2} w_n^{-mk}$$
$$\psi_m = \langle y^m, c \rangle$$

i.e. For any c, same y^m.



Determinants of circulant matrices

$$\det \operatorname{Cir}(c) = \prod_{0 \le m \le \lceil (n-1)/2 \rceil} \ell_m(c)$$

$$\ell_m(c) = \begin{cases} \langle V_m, c \rangle & m \in \{0, n/2\} \\ \operatorname{len}_m(c) & \text{otherwise} \end{cases}$$

$$\operatorname{len}_m(c) = \langle V_m, c \rangle^2 + \langle U_m, c \rangle^2$$

where

$$V_m + iU_m = y^m$$

Inverses of circulant matrices

Theorem

Let $c \in \mathbb{Z}^n$, $\operatorname{rank} \operatorname{Cir}(c) = n$; *then* $\operatorname{Cir}(c)^{-1} = \operatorname{Cir}(\hat{T}(c))$ *where*

$$\hat{T}(c) = \frac{1}{n} \begin{bmatrix} \langle c, \mathbf{1} \rangle^{-1} & + T_0(c) \\ \vdots & \\ \langle c, \mathbf{1} \rangle^{-1} & + T_{n-1}(c) \end{bmatrix}$$

where

$$T_k(c) = 2 \sum_{m=1}^{\lfloor (n-1)/2 \rfloor} \frac{\langle \sigma^k(V_m), c \rangle}{\langle V_m, c \rangle^2 + \langle U_m, c \rangle^2} + ((n+1) \bmod 2) \frac{\cos(k\pi)}{\langle V_{\frac{n}{2}}, c \rangle}$$

Branching on singularity

Reformulate singular case using binary r_m

$$\det(\mathrm{Cir}(x)) = \prod_{0 \le m \le \lceil (n-1)/2 \rceil} \ell_m(x) = 0$$

$$-Mr_m \le \ell_m(x) \le Mr_m$$

$$1 \le \sum_{n=0}^{\lceil \frac{n-1}{2} \rceil} (1 - r_m)$$

To enforce nonsingular case, we need constraints

$$|\ell_m(c)| > 0 \quad 1 \le m \le \left\lceil \frac{n-1}{2} \right\rceil$$

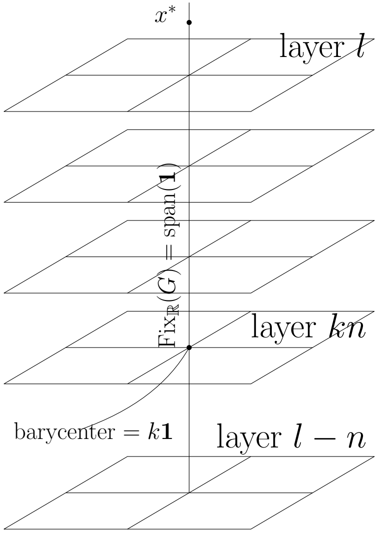
Transitive case: searching layer by layer

The *kth-layer* is

$$\text{layer}(n,k) := \{z \in \mathbb{Z}^n \mid \langle z, \mathbf{1} \rangle = k\}$$

If G acts by permuting coordinates:

$$\frac{j\mathbf{1}}{n} \in \text{Fix}_{\mathbb{R}}(G) \cap \text{layer}(n,j)$$



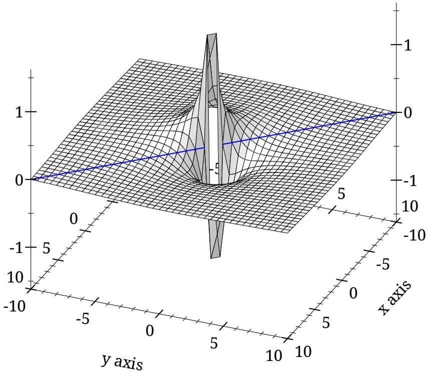
Constraints for the non-singular case

Lemma

Given $c, z \in \text{layer}(n, k)$ *such that* $\det \text{Cir}(c) \neq 0$.

$$z \notin \text{conv}\, Gc \Leftrightarrow \langle T(c), z + \rho\mathbf{1} \rangle < 0$$

(Hints: $z \in \text{aff}\, Gc$, relabelling, $\langle T(c), \mathbf{1} \rangle = 0$)



Algorithm for first k-variables active

$$\mathrm{OPT}_\infty(\ldots) := \text{opt. solution for IP and ..., or } -\infty$$

Set f^* to $\mathrm{OPT}_\infty(\det \mathrm{Cir}(x) = 0)$.

for $i = 1 \ldots k$ **do**

  $\Gamma := \sum_{j=0}^{k-1} x_j = i \pmod{k}$

  Choose $\widehat{E}^i = \{z^1, \ldots, z^{m_i}\} \in \{0, \pm 1, \pm 2\}^{m_i \times n}$.

  for $j = 1, \ldots, m_i$ **do**

    $f^* \leftarrow \max\{f^*, \mathrm{OPT}_\infty(\Gamma, x = z^j + \rho \mathbf{1})\}$.

  $f^* \leftarrow \max\{f^*, \mathrm{OPT}_\infty(\Gamma, \mathrm{conv}\, Gx \cap \widehat{E}^i = \emptyset)\}$.

return f^*

Hard examples for branch and bound

- Symmetric lattice free polytopes (infeasible IPs) from
  - compute simplex Gc for core point c
  - cut off vertices
- all take more than 1h to be solved in Gurobi 8.1, CPLEX 12.10 and GLPK 4.6 on an Intel Core-i5, 1.4 GHz and 8 GB RAM
- same machine, Knitro

Name	GAP Id	Longest cycle	Dimension	Time (s)
P1	(5,1)	5	5	3.1
P2	(15,2)	5	15	17
P3	(21,2)	7	21	15
P4	(45,1)	8	45	335

Conclusions and future work

Conclusions

- Initial results on small synthetic problems look promising
- Same techniques can be extended for products of cyclic groups

Future Work

- Rewrite using "less black box" solvers for better reproducibility, extensibility
- More practical problems (hopefully) in progress
- Generalization to other groups needs "nice" inverse of orbit matrix.
- Informed convexification/linearization