

Wolfe's Combinatorial Method is Exponential

Jamie Haddock

May 17, 2018

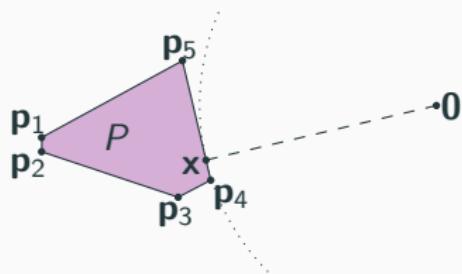
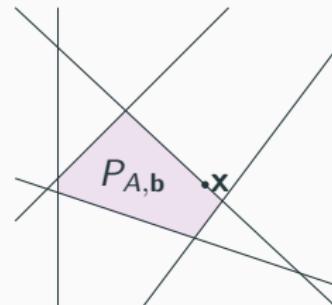
Graduate Group in Applied Mathematics
UC Davis

joint with Jesús De Loera and Luis Rademacher
<https://arxiv.org/abs/1710.02608>

Projection Algorithms for Convex and Combinatorial Optimization

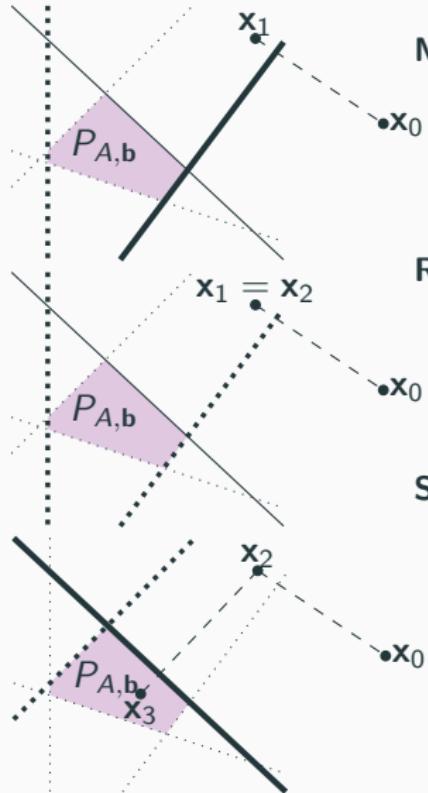
Two Problems

Linear Feasibility (LF): Given a rational matrix A and a rational vector \mathbf{b} , if $P_{A,\mathbf{b}} := \{\mathbf{x} : A\mathbf{x} \leq \mathbf{b}\}$ is nonempty, output a rational $\mathbf{x} \in P_{A,\mathbf{b}}$, otherwise output NO.



Minimum Norm Point (MNP): Given rational points $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m \in \mathbb{R}^n$ defining $P := \text{conv}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m)$, output rational $\mathbf{x} = \operatorname{argmin}_{\mathbf{q} \in P} \|\mathbf{q}\|^2$.

Iterative Projection Methods for LF



Motzkin's Method (MM)

- ▷ *On Motzkin's Method for Inconsistent Linear Systems* (joint with D. Needell)
<https://arxiv.org/abs/1802.03126>

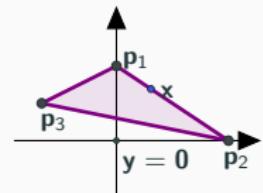
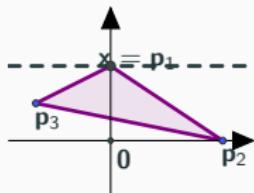
Randomized Kaczmarz (RK) Method

- ▷ *Randomized Projection Methods for Corrupted Linear Systems* (joint with D. Needell)
<https://arxiv.org/abs/1803.08114>

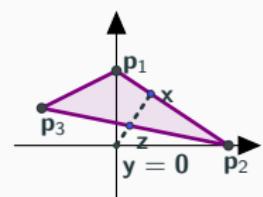
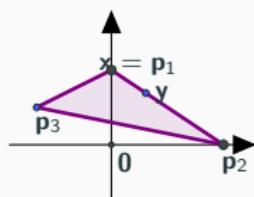
Sampling Kaczmarz-Motzkin (SKM) Methods

- ▷ *A Sampling Kaczmarz-Motzkin Algorithm for Linear Feasibility* (joint with J. A. De Loera and D. Needell)
SIAM Journal on Scientific Computing, 2017
<https://arxiv.org/abs/1605.01418>

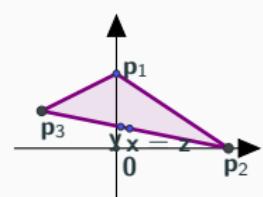
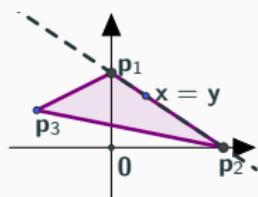
Wolfe's Combinatorial Methods for MNP



▷ *The Minimum Euclidean-Norm Point on
a Convex Polytope: Wolfe's
Combinatorial Algorithm is Exponential*
(joint J. A. De Loera and L. Rademacher)
STOC, 2018



<https://arxiv.org/abs/1710.02608>



Applications and Connections

LF:

- ▷ linear programming
- ▷ support vector machine
- ▷ linear equations

MNP:

- ▷ submodular function minimization
- ▷ colorful linear programming

Theorem (De Loera, H., Rademacher '17)

LF reduces to MNP on a simplex in strongly-polynomial time.

Minimum Norm Point ($\text{MNP}(P)$)

Minimum Norm Point in Polytope

We are interested in solving the problem ($\text{MNP}(P)$):

$$\min_{\mathbf{x} \in P} \|\mathbf{x}\|_2$$

where P is a polytope, and determining the minimum dimension face, F , which achieves distance $\|\mathbf{x}\|_2$.

Minimum Norm Point in Polytope

We are interested in solving the problem ($\text{MNP}(P)$):

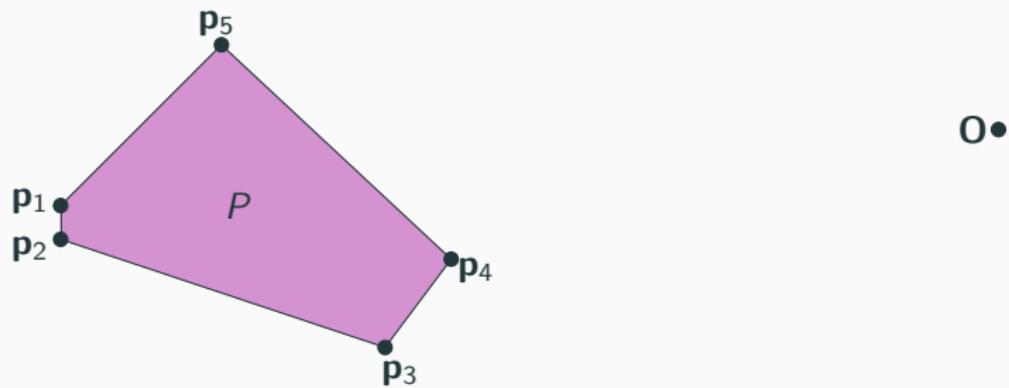
$$\min_{\mathbf{x} \in P} \|\mathbf{x}\|_2$$

where P is a polytope, and determining the minimum dimension face, F , which achieves distance $\|\mathbf{x}\|_2$.

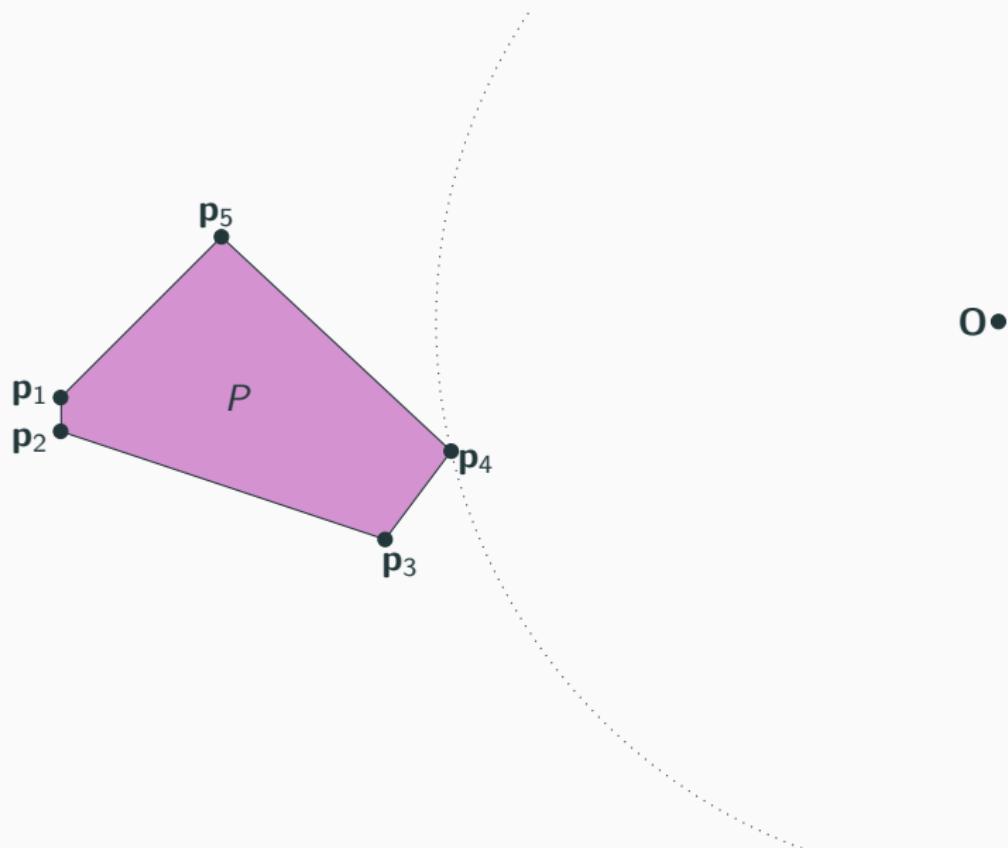
Reminder: A *polytope*, P , is the convex hull of points $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m$,

$$P = \left\{ \sum_{i=1}^m \lambda_i \mathbf{p}_i : \sum_{i=1}^m \lambda_i = 1, \lambda_i \geq 0 \text{ for all } i = 1, 2, \dots, m \right\}.$$

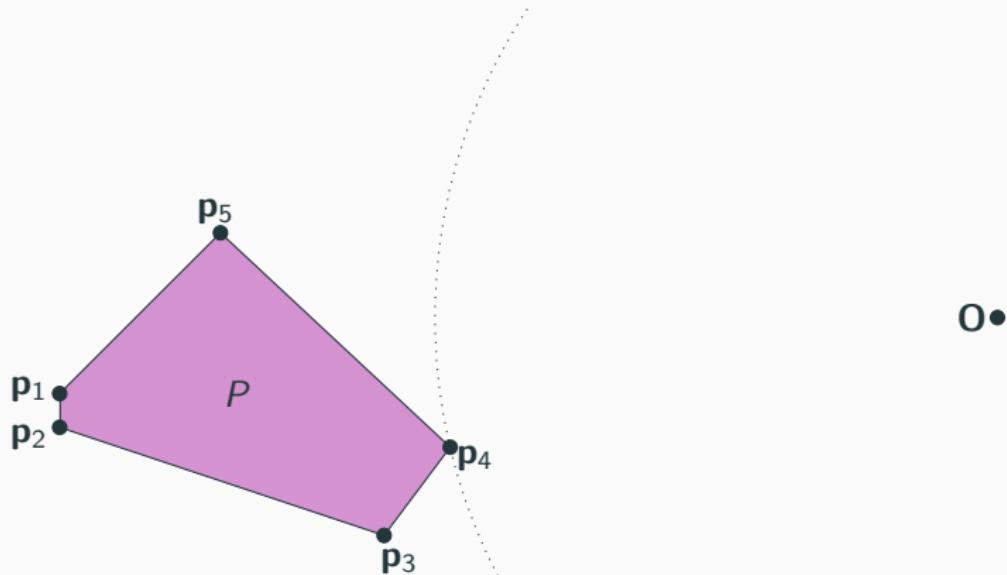
Minimum Norm Point in Polytope



Minimum Norm Point in Polytope



Minimum Norm Point in Polytope



▷ can be solved via interior-point methods

Applications

- arbitrary polytope projection

Applications

- arbitrary polytope projection
- nearest point problem for transportation polytopes

Applications

- arbitrary polytope projection
- nearest point problem for transportation polytopes
- subroutine in colorful linear programming

Applications

- arbitrary polytope projection
- nearest point problem for transportation polytopes
- subroutine in colorful linear programming
- subroutine in submodular function minimization

Applications

- arbitrary polytope projection
- nearest point problem for transportation polytopes
- subroutine in colorful linear programming
- subroutine in submodular function minimization
- machine learning - vision, large-scale learning

Applications

- arbitrary polytope projection
- nearest point problem for transportation polytopes
- subroutine in colorful linear programming
- subroutine in submodular function minimization
- machine learning - vision, large-scale learning
- compute distance to polytope

Applications

Theorem (De Loera, H., Rademacher '17)

Linear programming reduces to distance to a simplex in vertex-representation in strongly-polynomial time.

Applications

Theorem (De Loera, H., Rademacher '17)

Linear programming reduces to distance to a simplex in vertex-representation in strongly-polynomial time.

If a strongly-polynomial method for projection onto a polytope exists then this gives a **strongly-polynomial method for LP**.

Applications

Theorem (De Loera, H., Rademacher '17)

Linear programming reduces to distance to a simplex in vertex-representation in strongly-polynomial time.

If a strongly-polynomial method for projection onto a polytope exists then this gives a **strongly-polynomial method for LP**.

It was previously known that linear programming reduces to MNP on a polytope in weakly-polynomial time [Fujishige, Hayashi, Isotani '06].

Theorem (De Loera, H., Rademacher '17)

There exists a family of polytopes on which Wolfe's method requires exponential time to compute the MNP.

Wolfe's Optimality Condition

Theorem (Wolfe '74)

Let $P = \text{conv}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m)$. Then $\mathbf{x} \in P$ is MNP(P) if and only if

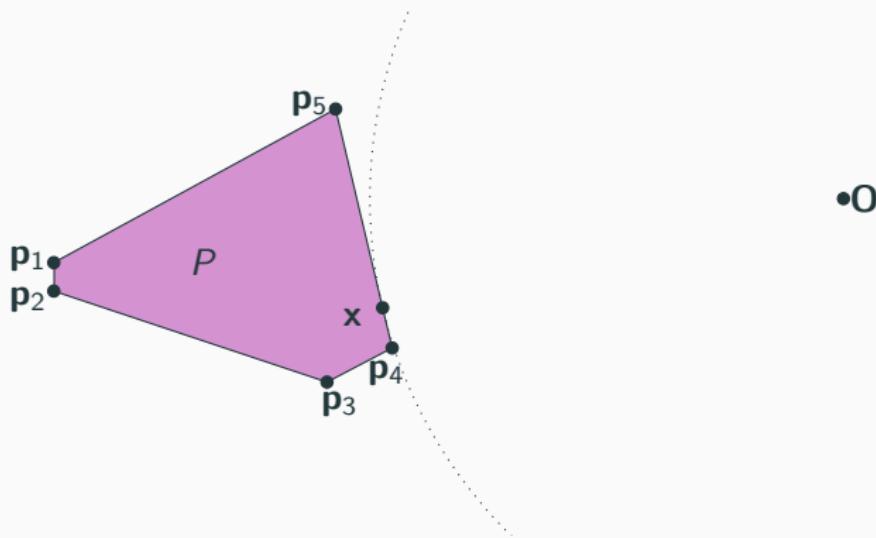
$$\mathbf{x}^T \mathbf{p}_j \geq \|\mathbf{x}\|_2^2 \text{ for all } j = 1, 2, \dots, m.$$

Wolfe's Optimality Condition

Theorem (Wolfe '74)

Let $P = \text{conv}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m)$. Then $\mathbf{x} \in P$ is MNP(P) if and only if

$$\mathbf{x}^T \mathbf{p}_j \geq \|\mathbf{x}\|_2^2 \text{ for all } j = 1, 2, \dots, m.$$

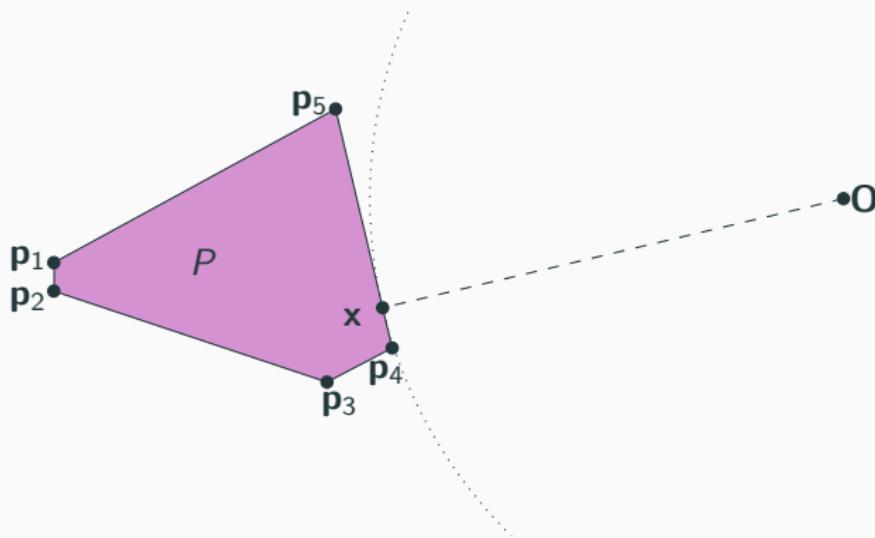


Wolfe's Optimality Condition

Theorem (Wolfe '74)

Let $P = \text{conv}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m)$. Then $\mathbf{x} \in P$ is MNP(P) if and only if

$$\mathbf{x}^T \mathbf{p}_j \geq \|\mathbf{x}\|_2^2 \text{ for all } j = 1, 2, \dots, m.$$

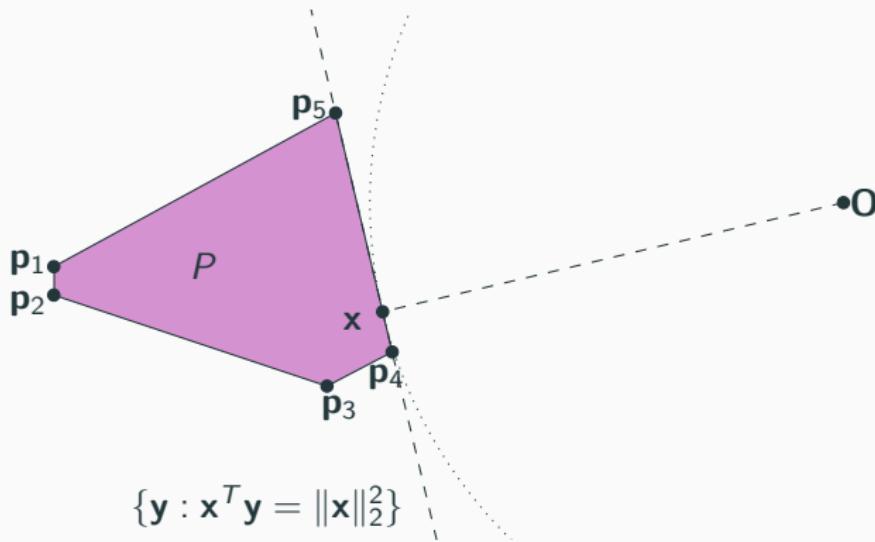


Wolfe's Optimality Condition

Theorem (Wolfe '74)

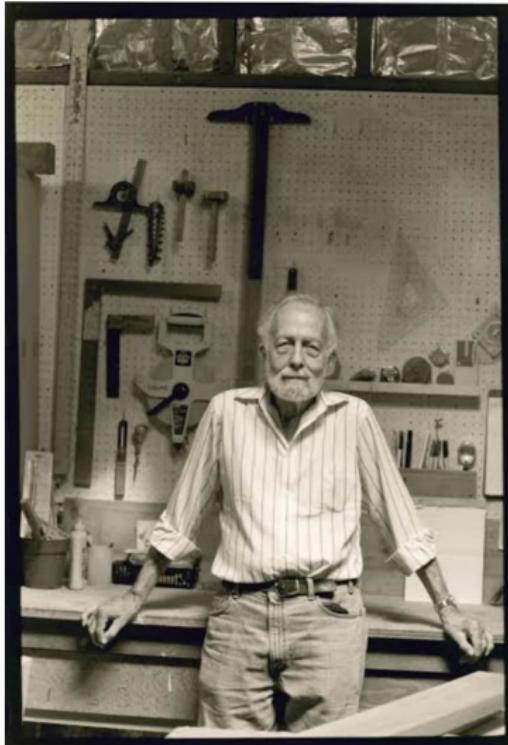
Let $P = \text{conv}(\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m)$. Then $\mathbf{x} \in P$ is MNP(P) if and only if

$$\mathbf{x}^T \mathbf{p}_j \geq \|\mathbf{x}\|_2^2 \text{ for all } j = 1, 2, \dots, m.$$



Wolfe's Method

Philip Wolfe



- Frank-Wolfe method
- Dantzig-Wolfe decomposition
- simplex method for quadratic programming

Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

Intuition and Definitions

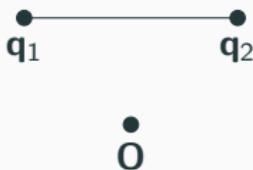
Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.

Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.



Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.



Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

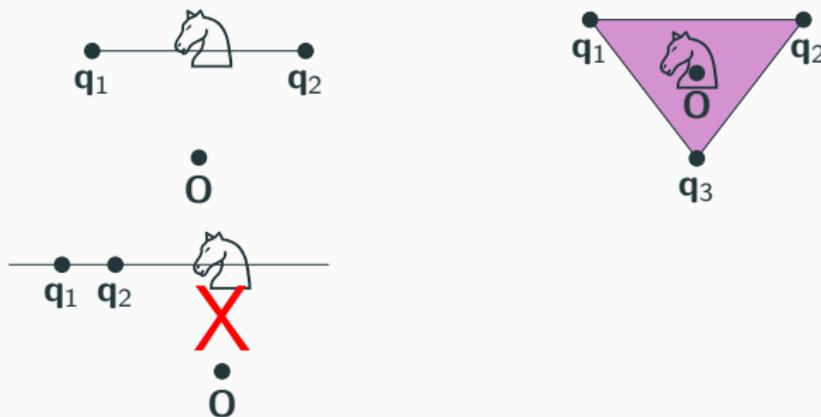
Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.



Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

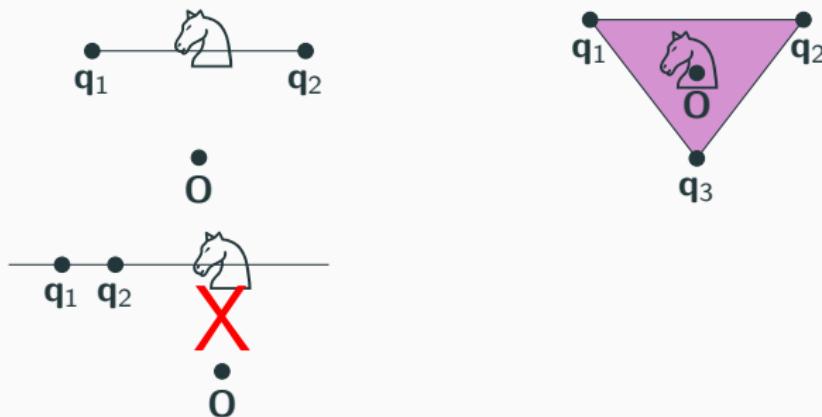
Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.



Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.

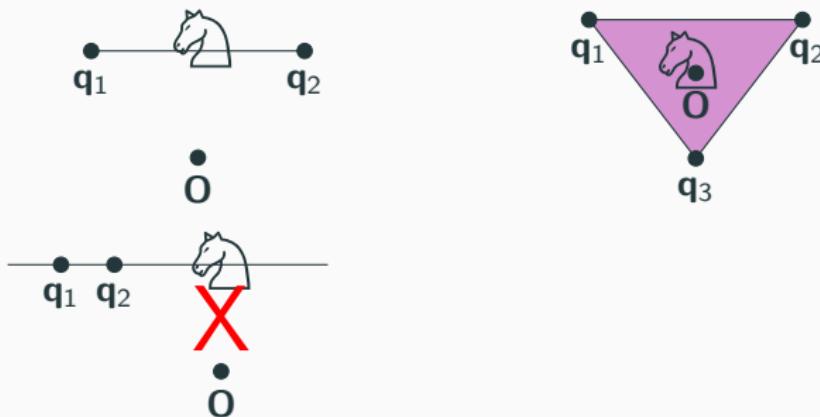


Note: Singletons are corrals.

Intuition and Definitions

Idea: Exploit linear information about the problem in order to progress towards the quadratic solution.

Def: An affinely independent set of points $Q = \{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}$ is a *corral* if $\text{MNP}(\text{aff}(Q)) \in \text{relint}(\text{conv}(Q))$.



Note: Singletons are corrals.

Note: There is a corral in P whose convex hull contains $\text{MNP}(P)$.

Wolfe's method : combinatorial method for computing projection onto
a vertex-representation polytope (any dimension, any
number of points)

Wolfe's method : combinatorial method for computing projection onto
a vertex-representation polytope (any dimension, any
number of points)
- **pivots** between corrals which may contain $MNP(P)$

Wolfe's method : combinatorial method for computing projection onto
a vertex-representation polytope (any dimension, any
number of points)

- **pivots** between corrals which may contain $MNP(P)$
- **projects** onto affine hull of sets to check whether a corral

Wolfe's method : combinatorial method for computing projection onto
a vertex-representation polytope (any dimension, any
number of points)

- **pivots** between corrals which may contain $MNP(P)$
- **projects** onto affine hull of sets to check whether a corral
- optimality criterion **checks** if correct corral

Sketch of Method

$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$

$C = \{x\}$

while x is not MNP(P)

$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$

$C = C \cup \{\mathbf{p}_j\}$

$y = \text{MNP}(\text{aff}(C))$

while $y \notin \text{relint}(\text{conv}(C))$

$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\text{argmin}} \|z - y\|_2$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$x = z$

$y = \text{MNP}(\text{aff}(C))$

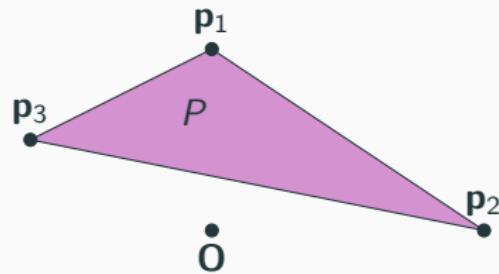
$x = y$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



Sketch of Method

$x \in P = \{p_1, p_2, \dots, p_m\}$

$C = \{x\}$

while x is not MNP(P)

$p_j \in \{p \in P : x^T p < \|x\|_2^2\}$

$C = C \cup \{p_j\}$

$y = \text{MNP}(\text{aff}(C))$

while $y \notin \text{relint}(\text{conv}(C))$

$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\text{argmin}} \|z - y\|_2$

$C = C - \{p_i\}$ where p_i, z
are on different faces of
 $\text{conv}(C)$

$x = z$

$y = \text{MNP}(\text{aff}(C))$

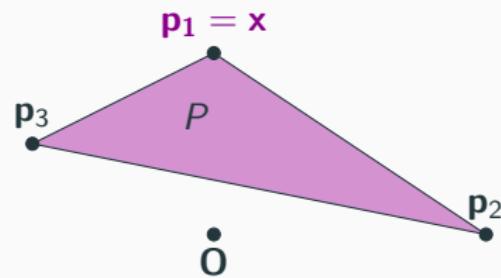
$x = y$

return x

$$p_1 = (0, 2)$$

$$p_2 = (3, 0)$$

$$p_3 = (-2, 1)$$



Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$\mathbf{C} = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

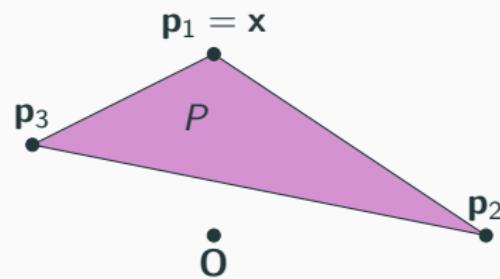
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$\mathbf{C} = \{\mathbf{p}_1\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

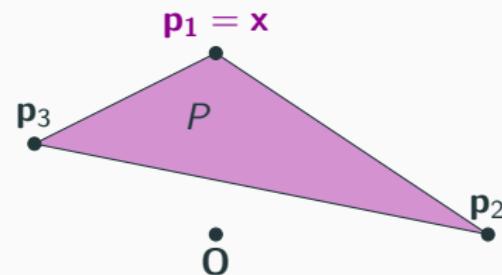
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

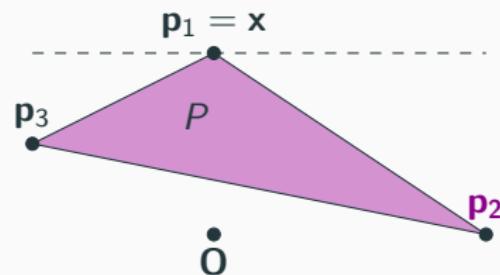
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$$

$$\mathbf{C} = \mathbf{C} \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$$C = C - \{\mathbf{p}_i\} \text{ where } \mathbf{p}_i, z$$

are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

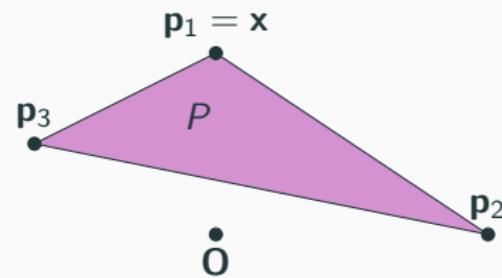
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$\mathbf{C} = \{\mathbf{p}_1, \mathbf{p}_2\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

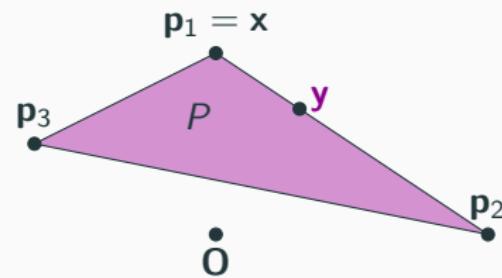
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint(conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\text{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

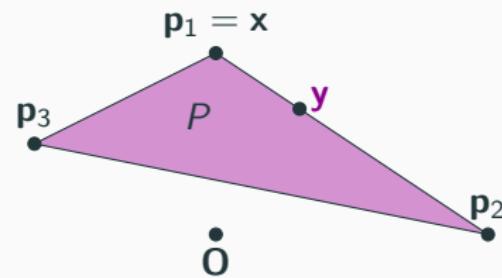
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

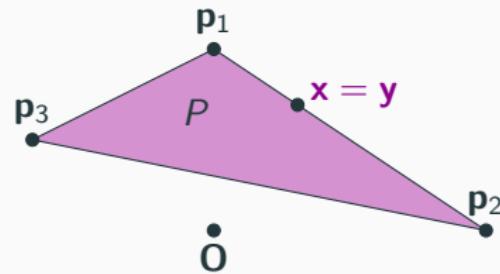
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

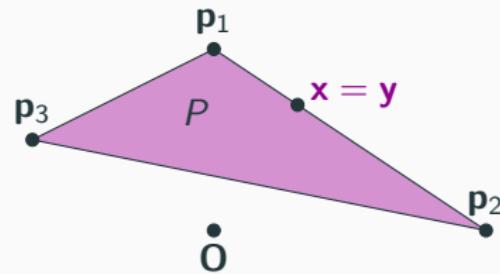
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

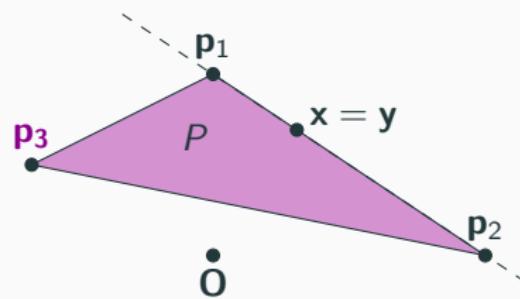
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

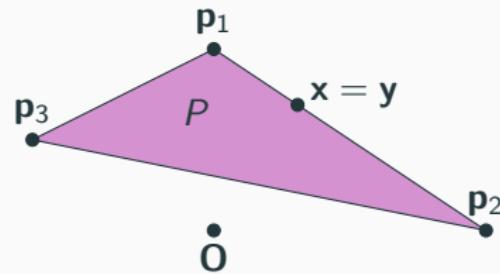
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$\mathbf{y} = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$\mathbf{z} = \underset{\mathbf{z} \in \text{conv}(C) \cap \overline{\mathbf{xy}}}{\operatorname{argmin}} \|\mathbf{z} - \mathbf{y}\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, \mathbf{z}
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

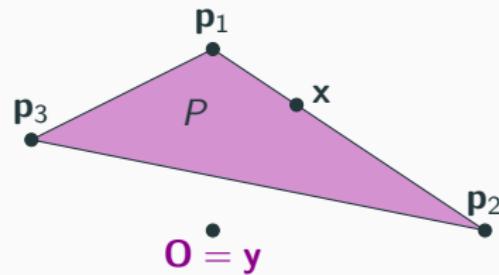
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint(conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\text{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

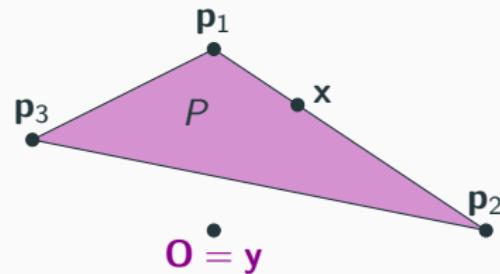
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\text{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

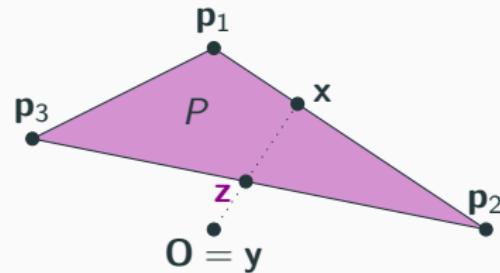
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

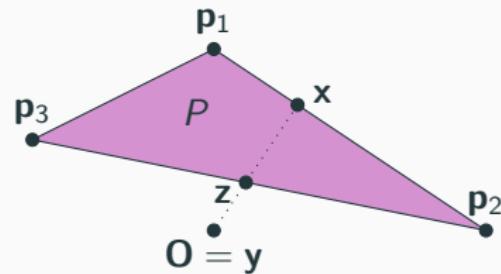
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

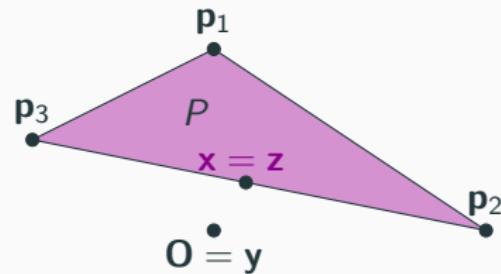
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

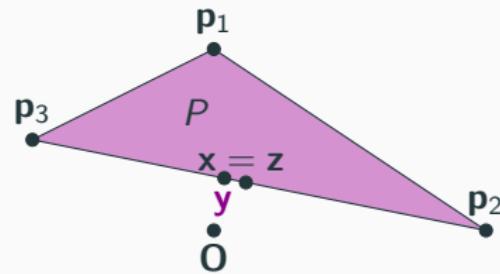
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint(conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

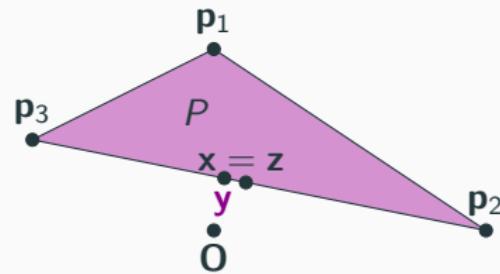
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

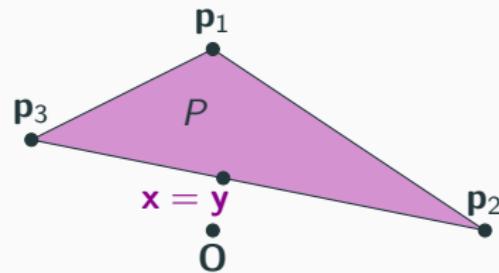
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : x^T \mathbf{p} < \|x\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$$C = C - \{\mathbf{p}_i\}$$
 where \mathbf{p}_i, z

are on different faces of

$$\text{conv}(C)$$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

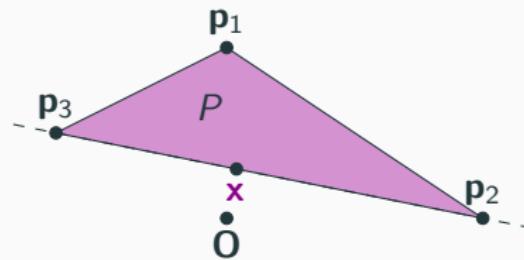
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Sketch of Method

$$x \in P = \{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_m\}$$

$$C = \{x\}$$

while x is not MNP(P)

$$\mathbf{p}_j \in \{\mathbf{p} \in P : \mathbf{x}^T \mathbf{p} < \|\mathbf{x}\|_2^2\}$$

$$C = C \cup \{\mathbf{p}_j\}$$

$$y = \text{MNP}(\text{aff}(C))$$

while $y \notin \text{relint}(\text{conv}(C))$

$$z = \underset{z \in \text{conv}(C) \cap \overline{xy}}{\operatorname{argmin}} \|z - y\|_2$$

$C = C - \{\mathbf{p}_i\}$ where \mathbf{p}_i, z
are on different faces of
 $\text{conv}(C)$

$$x = z$$

$$y = \text{MNP}(\text{aff}(C))$$

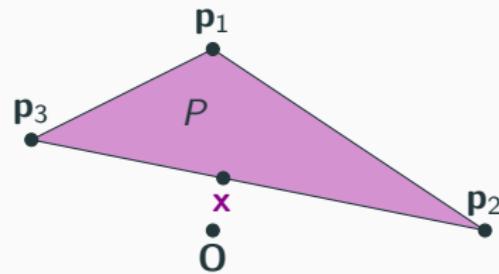
$$x = y$$

return x

$$\mathbf{p}_1 = (0, 2)$$

$$\mathbf{p}_2 = (3, 0)$$

$$\mathbf{p}_3 = (-2, 1)$$



$$C = \{\mathbf{p}_2, \mathbf{p}_3\}$$

Wolfe's Method

$\mathbf{x} = \mathbf{p}_i$ for some $i = 1, 2, \dots, m$, $\lambda = \mathbf{e}_i$

$C = \{i\}$

while $\mathbf{x} \neq \mathbf{0}$ and there exists \mathbf{p}_j with $\mathbf{x}^T \mathbf{p}_j < \|\mathbf{x}\|_2^2$

$C = C \cup \{j\}$

$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \|\sum_{i \in C} \alpha_i \mathbf{p}_i\|_2$, $\mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$

while $\alpha_i \leq 0$ for some $i = 1, 2, \dots, m$

$\theta = \min_{i: \alpha_i \leq 0} \frac{\lambda_i}{\lambda_i - \alpha_i}$

$\mathbf{z} = \theta \mathbf{y} + (1 - \theta) \mathbf{x}$

$i \in \{j : \theta \alpha_j + (1 - \theta) \lambda_j = 0\}$

$C = C - \{i\}$

$\mathbf{x} = \mathbf{z}$

solve $\mathbf{x} = P\lambda$ for λ

$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \|\sum_{i \in C} \alpha_i \mathbf{p}_i\|_2$, $\mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$

$\mathbf{x} = \mathbf{y}$

return \mathbf{x}

Wolfe's Method

$\mathbf{x} = \mathbf{p}_i$ for some $i = 1, 2, \dots, m$, $\lambda = \mathbf{e}_i$

$C = \{i\}$

while $\mathbf{x} \neq \mathbf{0}$ and there exists \mathbf{p}_j with $\mathbf{x}^T \mathbf{p}_j < \|\mathbf{x}\|_2^2$

$C = C \cup \{j\}$

$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \|\sum_{i \in C} \alpha_i \mathbf{p}_i\|_2$, $\mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$

while $\alpha_i \leq 0$ for some $i = 1, 2, \dots, m$

$\theta = \min_{i: \alpha_i \leq 0} \frac{\lambda_i}{\lambda_i - \alpha_i}$

$\mathbf{z} = \theta \mathbf{y} + (1 - \theta) \mathbf{x}$

$i \in \{j : \theta \alpha_j + (1 - \theta) \lambda_j = 0\}$

$C = C - \{i\}$

$\mathbf{x} = \mathbf{z}$

solve $\mathbf{x} = P\lambda$ for λ

$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \|\sum_{i \in C} \alpha_i \mathbf{p}_i\|_2$, $\mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$

$\mathbf{x} = \mathbf{y}$

return \mathbf{x}

Choice 1: Initial vertex.

Wolfe's Method

$\mathbf{x} = \mathbf{p}_i$ for some $i = 1, 2, \dots, m$, $\lambda = \mathbf{e}$;

$C = \{i\}$

while $\mathbf{x} \neq \mathbf{0}$ and there exists \mathbf{p}_j with $\mathbf{x}^T \mathbf{p}_j < \|\mathbf{x}\|_2^2$

$C = C \cup \{j\}$

$$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \|\sum_{i \in C} \alpha_i \mathbf{p}_i\|_2, \quad \mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$$

while $\alpha_i \leq 0$ for some $i = 1, 2, \dots, m$

$$\theta = \min_{i: \alpha_i \leq 0} \frac{\lambda_i}{\lambda_i - \alpha_i}$$

$$\mathbf{z} = \theta \mathbf{y} + (1 - \theta) \mathbf{x}$$

$$i \in \{j : \theta \alpha_j + (1 - \theta) \lambda_j = 0\}$$

$$C = C - \{i\}$$

$$\mathbf{x} = \mathbf{z}$$

solve $\mathbf{x} = P\lambda$ for λ

$$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \|\sum_{i \in C} \alpha_i \mathbf{p}_i\|_2, \quad \mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$$

$$\mathbf{x} = \mathbf{y}$$

return \mathbf{x}

Choice 1: Initial vertex.

Choice 2: Adding to corral.

Wolfe's Method

$\mathbf{x} = \mathbf{p}_i$ for some $i = 1, 2, \dots, m$, $\lambda = \mathbf{e}$;

$C = \{i\}$

while $\mathbf{x} \neq \mathbf{0}$ and there exists \mathbf{p}_j with $\mathbf{x}^T \mathbf{p}_j < \|\mathbf{x}\|_2^2$

$C = C \cup \{j\}$

$$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \left\| \sum_{i \in C} \alpha_i \mathbf{p}_i \right\|_2, \quad \mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$$

while $\alpha_i \leq 0$ for some $i = 1, 2, \dots, m$

$$\theta = \min_{i: \alpha_i \leq 0} \frac{\lambda_i}{\lambda_i - \alpha_i}$$

$$\mathbf{z} = \theta \mathbf{y} + (1 - \theta) \mathbf{x}$$

$$i \in \{j : \theta \alpha_j + (1 - \theta) \lambda_j = 0\}$$

$$C = C - \{i\}$$

$$\mathbf{x} = \mathbf{z}$$

solve $\mathbf{x} = P\lambda$ for λ

$$\alpha = \underset{\sum_{i \in C} \alpha_i = 1}{\operatorname{argmin}} \left\| \sum_{i \in C} \alpha_i \mathbf{p}_i \right\|_2, \quad \mathbf{y} = \sum_{i \in C} \alpha_i \mathbf{p}_i$$

$$\mathbf{x} = \mathbf{y}$$

return \mathbf{x}

Choice 1: Initial vertex.

Choice 2: Adding to corral.

Choice 3: Removing from corral.

Rules

Initial: `minnorm`

Insertion: `linopt` (select \mathbf{p}_j minimizing $\mathbf{x}^T \mathbf{p}_j$), `minnorm`

Rules

Initial: `minnorm`

Insertion: `linopt` (select \mathbf{p}_j minimizing $\mathbf{x}^T \mathbf{p}_j$), `minnorm`

- insertion rules have different benefits

Rules

Initial: `minnorm`

Insertion: `linopt` (select \mathbf{p}_j minimizing $\mathbf{x}^T \mathbf{p}_j$), `minnorm`

- insertion rules have different benefits
- behavior depends on choice of insertion rule

Rules

Initial: `minnorm`

Insertion: `linopt` (select \mathbf{p}_j minimizing $\mathbf{x}^T \mathbf{p}_j$), `minnorm`

- insertion rules have different benefits
- behavior depends on choice of insertion rule
- examples in which each insertion rule is better

Rules

Initial: `minnorm`

Insertion: `linopt` (select \mathbf{p}_j minimizing $\mathbf{x}^T \mathbf{p}_j$), `minnorm`

- insertion rules have different benefits
- behavior depends on choice of insertion rule
- examples in which each insertion rule is better
- a dropped vertex may be readded

Related Methods

- ▷ von Neumann's algorithm for linear programming

Related Methods

- ▷ von Neumann's algorithm for linear programming
- ▷ Frank-Wolfe method for convex programming (and variants)

Related Methods

- ▷ von Neumann's algorithm for linear programming
- ▷ Frank-Wolfe method for convex programming (and variants)
- ▷ Gilbert's procedure for quadratic programming

Related Methods

- ▷ von Neumann's algorithm for linear programming
 - ▷ Frank-Wolfe method for convex programming (and variants)
 - ▷ Gilbert's procedure for quadratic programming
-
- projection onto simple convex hull

Related Methods

- ▷ von Neumann's algorithm for linear programming
 - ▷ Frank-Wolfe method for convex programming (and variants)
 - ▷ Gilbert's procedure for quadratic programming
-
- projection onto simple convex hull
-
- ▷ Hanson-Lawson procedure for non-negative least-squares

Previous Results

- # iterations $\leq \sum_{i=1}^{n+1} i \binom{m}{i}$ with any rules (Wolfe '74)

Previous Results

- # iterations $\leq \sum_{i=1}^{n+1} i \binom{m}{i}$ with any rules (Wolfe '74)
- ϵ -approximate solution in $\mathcal{O}(nM^2/\epsilon)$ iterations with linopt insertion rule (Chakrabarty, Jain, Kothari '14)

Previous Results

- # iterations $\leq \sum_{i=1}^{n+1} i \binom{m}{i}$ with any rules (Wolfe '74)
- ϵ -approximate solution in $\mathcal{O}(nM^2/\epsilon)$ iterations with linopt insertion rule (Chakrabarty, Jain, Kothari '14)
- ϵ -approximate solution in $\mathcal{O}(\rho \log(1/\epsilon))$ iterations with linopt insertion rule (Lacoste-Julien, Jaggi '15)

Previous Results

- # iterations $\leq \sum_{i=1}^{n+1} i \binom{m}{i}$ with any rules (Wolfe '74)
- ϵ -approximate solution in $\mathcal{O}(nM^2/\epsilon)$ iterations with linopt insertion rule (Chakrabarty, Jain, Kothari '14)
- ϵ -approximate solution in $\mathcal{O}(\rho \log(1/\epsilon))$ iterations with linopt insertion rule (Lacoste-Julien, Jaggi '15)
 - ▷ pseudo-polynomial complexity

Exponential Behavior

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

- dimension and number of points grow linearly

Exponential Example

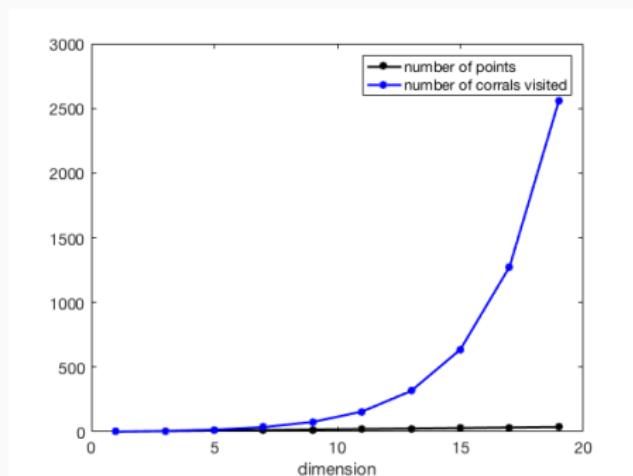
Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

- dimension and number of points grow linearly
- number of corrals visited grows exponentially

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

- dimension and number of points grow linearly
- number of corrals visited grows exponentially



Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Recursively Defined Instances

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Recursively Defined Instances

dim: $d - 2$

Instance: $P(d - 2)$

Points: $2d - 5$

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Recursively Defined Instances

dim: $d - 2$

Instance: $P(d - 2)$

Points: $2d - 5$

+2 dim
→

+4 points

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Recursively Defined Instances

dim: $d - 2$

Instance: $P(d - 2)$

Points: $2d - 5$

+2 dim
→

+4 points

dim: d

Instance: $P(d)$

Points: $2d - 1$

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Recursively Defined Instances

dim: $d - 2$

Instance: $P(d - 2)$

Points: $2d - 5$

$\xrightarrow{+2 \text{ dim}}$

+4 points

dim: d

Instance: $P(d)$

Points: $2d - 1$

$$P(1) := \{1\}$$

Exponential Example

Goal : build family of instances on which the number of iterations of Wolfe's method is at least exponential in the dimension and number of points

Recursively Defined Instances

dim: $d - 2$

Instance: $P(d - 2)$

Points: $2d - 5$

$\xrightarrow{+2 \text{ dim}}$

+4 points

dim: d

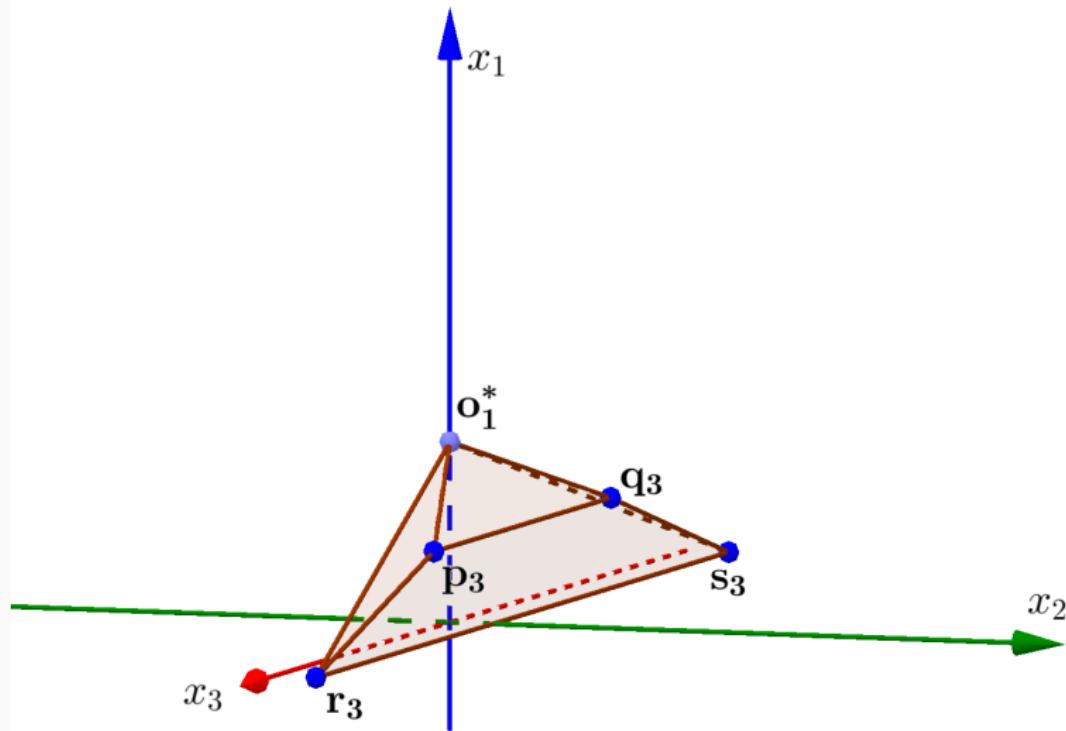
Instance: $P(d)$

Points: $2d - 1$

$$P(1) := \{1\}$$

$$P(3) := \{(1, 0, 0), \mathbf{p}_3, \mathbf{q}_3, \mathbf{r}_3, \mathbf{s}_3\}$$

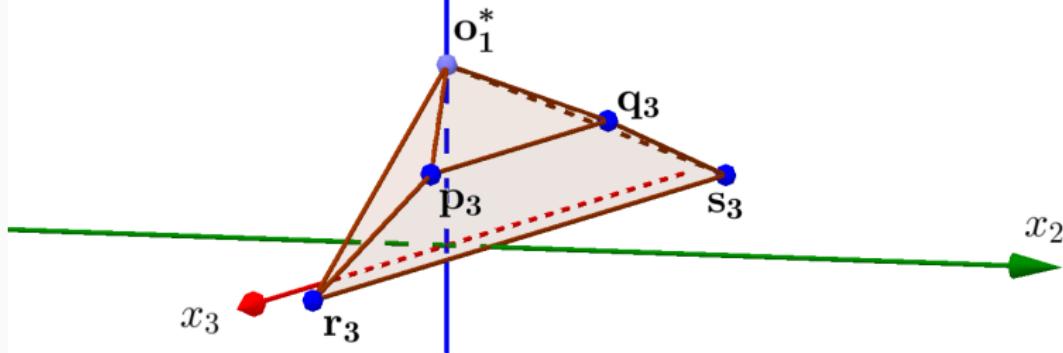
Exponential Example: dim 3



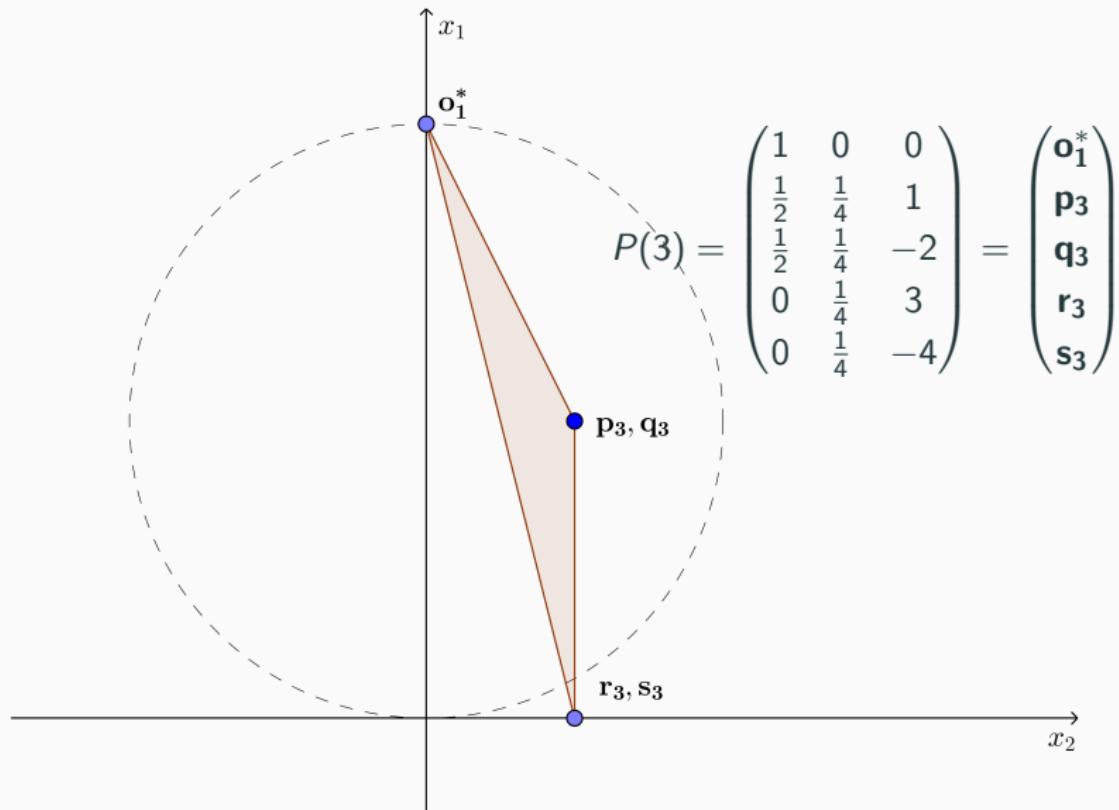
Exponential Example: dim 3

x_1

$$P(3) = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{4} & 1 \\ \frac{1}{2} & \frac{1}{4} & -2 \\ 0 & \frac{1}{4} & 3 \\ 0 & \frac{1}{4} & -4 \end{pmatrix} = \begin{pmatrix} \mathbf{o}_1^* \\ \mathbf{p}_3 \\ \mathbf{q}_3 \\ \mathbf{r}_3 \\ \mathbf{s}_3 \end{pmatrix}$$



Exponential Example: dim 3



Exponential Example

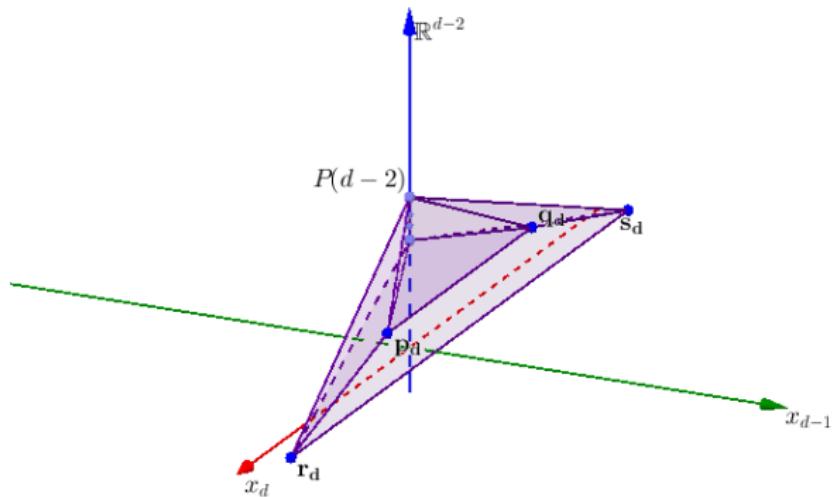
$$P(d) = \begin{pmatrix} P(d-2) & 0 & 0 \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & M_{d-2} \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & -(M_{d-2} + 1) \\ 0 & \frac{m_{d-2}}{4} & M_{d-2} + 2 \\ 0 & \frac{m_{d-2}}{4} & -(M_{d-2} + 3) \end{pmatrix}$$

\mathbf{o}_{d-2}^* : MNP($P(d-2)$)

$m_{d-2} = \|\mathbf{o}_{d-2}^*\|_\infty$

$M_{d-2} = \max_{\mathbf{p} \in P(d-2)} \|\mathbf{p}\|_1$

Exponential Example



$$P(d) = \begin{pmatrix} P(d-2) & 0 & 0 \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & M_{d-2} \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & -(M_{d-2} + 1) \\ 0 & \frac{m_{d-2}}{4} & M_{d-2} + 2 \\ 0 & \frac{m_{d-2}}{4} & -(M_{d-2} + 3) \end{pmatrix} \quad \downarrow \parallel \cdot \parallel$$

\mathbf{o}_{d-2}^* : MNP($P(d-2)$)
 $m_{d-2} = \|\mathbf{o}_{d-2}^*\|_\infty$
 $M_{d-2} = \max_{\mathbf{p} \in P(d-2)} \|\mathbf{p}\|_1$

Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the minnorm insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the minnorm insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Key Lemma: Sequence of Corrals

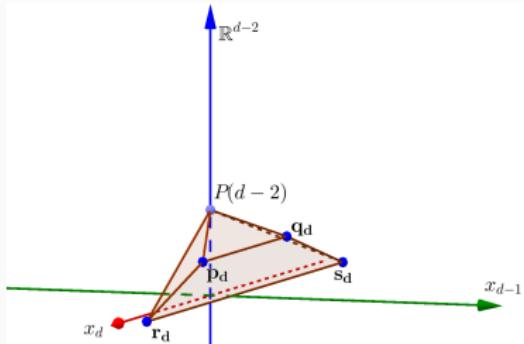
Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the *minnorm* insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Key Lemma: Sequence of Corrals

$$C(d - 2)$$



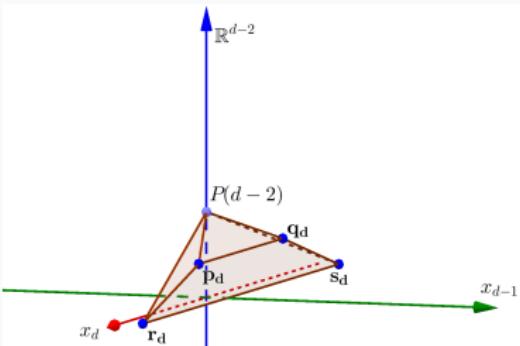
Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the *minnorm* insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Key Lemma: Sequence of Corrals

$$C(d-2) \longrightarrow$$



Exponential Example

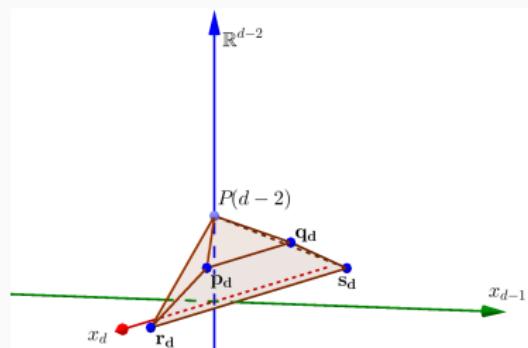
Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the *minnorm* insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Key Lemma: Sequence of Corrals

$$C(d-2) \longrightarrow$$

$$\begin{aligned} & C(d-2) \\ & O(d-2)\mathbf{p}_d \\ & \mathbf{p}_d\mathbf{q}_d \\ & \mathbf{q}_d\mathbf{r}_d \\ & \mathbf{r}_d\mathbf{s}_d \\ & C(d-2)\mathbf{r}_d\mathbf{s}_d \end{aligned}$$



Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the minnorm insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Sequence of Corrals: $\dim 1 \rightarrow \dim 3$

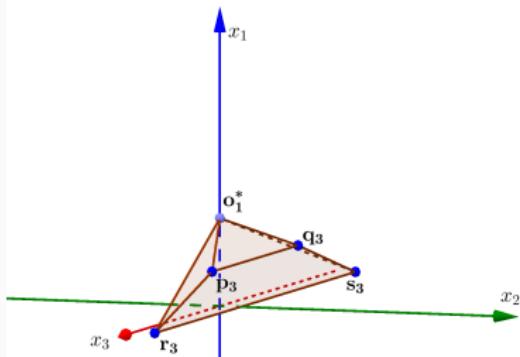
Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the *minnorm* insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Sequence of Corrals: $\dim 1 \rightarrow \dim 3$

1



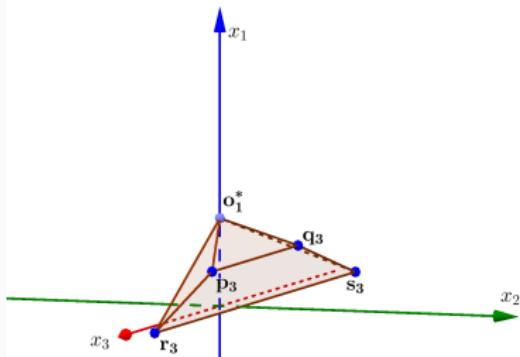
Exponential Example

Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the *minnorm* insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Sequence of Corrals: $\dim 1 \rightarrow \dim 3$

1 \longrightarrow



Exponential Example

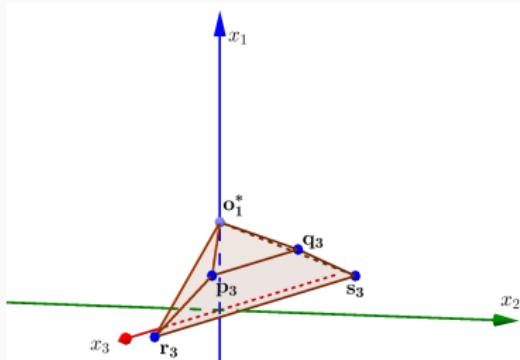
Theorem (De Loera, H., Rademacher '17)

Consider the execution of Wolfe's method with the *minnorm* insertion rule on input $P(d)$ where $d = 2k - 1$. Then the sequence of corrals, $C(d)$ has length $5 \cdot 2^{k-1} - 4$.

Sequence of Corrals: $\dim 1 \rightarrow \dim 3$

1 \longrightarrow

$$\begin{aligned} & (1, 0, 0) \\ & (1, 0, 0)\mathbf{p}_3 \\ & \mathbf{p}_3\mathbf{q}_3 \\ & \mathbf{q}_3\mathbf{r}_3 \\ & \mathbf{r}_3\mathbf{s}_3 \\ & (1, 0, 0)\mathbf{r}_3\mathbf{s}_3 \end{aligned}$$



Adding Point to Corral

Lemma

Let $P \subseteq \mathbb{R}^d$ be a finite set of points that is a corral. Let \mathbf{x} be the minimum norm point in $\text{aff } P$. Let $\mathbf{q} \in \text{span}(\mathbf{x}, \text{span}(P)^\perp)$, and assume $\mathbf{q}^T \mathbf{x} < \min\{\|\mathbf{q}\|_2^2, \|\mathbf{x}\|_2^2\}$. Then $P \cup \{\mathbf{q}\}$ is a corral. Moreover, the minimum norm point \mathbf{y} in $\text{conv}(P \cup \{\mathbf{q}\})$ is a (strict) convex combination of \mathbf{q} and the minimum norm point of P : $\mathbf{y} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{q}$ with $\lambda = \mathbf{q}^T (\mathbf{q} - \mathbf{x}) / \|\mathbf{q} - \mathbf{x}\|_2^2$.

Adding Point to Corral

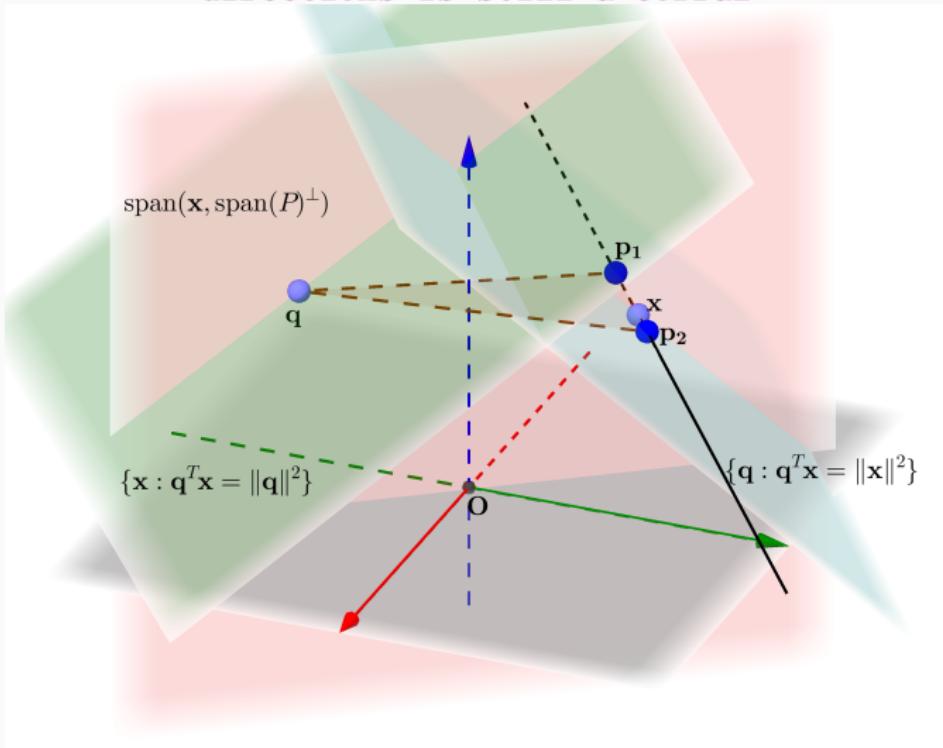
Lemma

Let $P \subseteq \mathbb{R}^d$ be a finite set of points that is a corral. Let \mathbf{x} be the minimum norm point in $\text{aff } P$. Let $\mathbf{q} \in \text{span}(\mathbf{x}, \text{span}(P)^\perp)$, and assume $\mathbf{q}^T \mathbf{x} < \min\{\|\mathbf{q}\|_2^2, \|\mathbf{x}\|_2^2\}$. Then $P \cup \{\mathbf{q}\}$ is a corral. Moreover, the minimum norm point \mathbf{y} in $\text{conv}(P \cup \{\mathbf{q}\})$ is a (strict) convex combination of \mathbf{q} and the minimum norm point of P : $\mathbf{y} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{q}$ with $\lambda = \mathbf{q}^T (\mathbf{q} - \mathbf{x}) / \|\mathbf{q} - \mathbf{x}\|_2^2$.

a corral with a point made from MNP and orthogonal directions is still a corral

Adding Point to Corral

a corral with a point made from MNP and orthogonal directions is still a corral



Orthogonal Corrals

Lemma

Let $A \subseteq \mathbb{R}^d$ be a proper linear subspace. Let $P \subseteq A$ be a non-empty finite set. Let $Q \subseteq A^\perp$ be another non-empty finite set. Let \mathbf{x} be the minimum norm point in $\text{aff } P$. Let \mathbf{y} be the minimum norm point in $\text{aff } Q$. Let \mathbf{z} be the minimum norm point in $\text{aff}(P \cup Q)$. We have:

1. \mathbf{z} is the minimum norm point in $[\mathbf{x}, \mathbf{y}]$ and therefore

$$\mathbf{z} = \lambda\mathbf{x} + (1 - \lambda)\mathbf{y} \text{ with } \lambda = \frac{\|\mathbf{y}\|_2^2}{\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2}.$$

2. If $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$, then \mathbf{z} is a strict convex combination of \mathbf{x} and \mathbf{y} .

3. If $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$ and P and Q are corrals, then $P \cup Q$ is also a corral.

Orthogonal Corrals

Lemma

Let $A \subseteq \mathbb{R}^d$ be a proper linear subspace. Let $P \subseteq A$ be a non-empty finite set. Let $Q \subseteq A^\perp$ be another non-empty finite set. Let \mathbf{x} be the minimum norm point in $\text{aff } P$. Let \mathbf{y} be the minimum norm point in $\text{aff } Q$. Let \mathbf{z} be the minimum norm point in $\text{aff}(P \cup Q)$. We have:

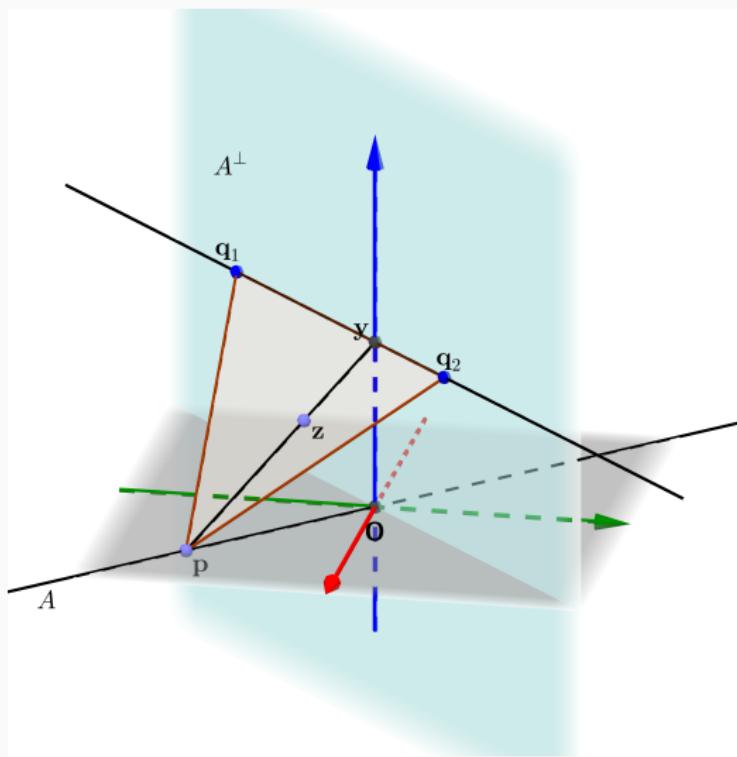
1. \mathbf{z} is the minimum norm point in $[\mathbf{x}, \mathbf{y}]$ and therefore

$$\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \text{ with } \lambda = \frac{\|\mathbf{y}\|_2^2}{\|\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2}.$$

2. If $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{y} \neq \mathbf{0}$, then \mathbf{z} is a strict convex combination of \mathbf{x} and \mathbf{y} .
3. If $\mathbf{x} \neq \mathbf{0}$, $\mathbf{y} \neq \mathbf{0}$ and P and Q are corrals, then $P \cup Q$ is also a corral.

the union of orthogonal corrals is still a corral

Orthogonal Corrals



the union of orthogonal corrals is still a corral

Wolfe's Criterion under Addition of Orthogonal Point

Lemma

For a point \mathbf{z} define $H_{\mathbf{z}} = \{\mathbf{w} \in \mathbb{R}^n : \mathbf{w} \cdot \mathbf{z} < \|\mathbf{z}\|_2^2\}$. Suppose that we have an instance of the minimum norm point problem in \mathbb{R}^d as follows: Some points, P , live in a proper linear subspace A and some, Q , in A^\perp . Let \mathbf{x} be the minimum norm point in $\text{aff } P$ and \mathbf{y} be the minimum norm point in $\text{aff}(P \cup Q)$. Then $H_{\mathbf{y}} \cap A = H_{\mathbf{x}} \cap A$.

Wolfe's Criterion under Addition of Orthogonal Point

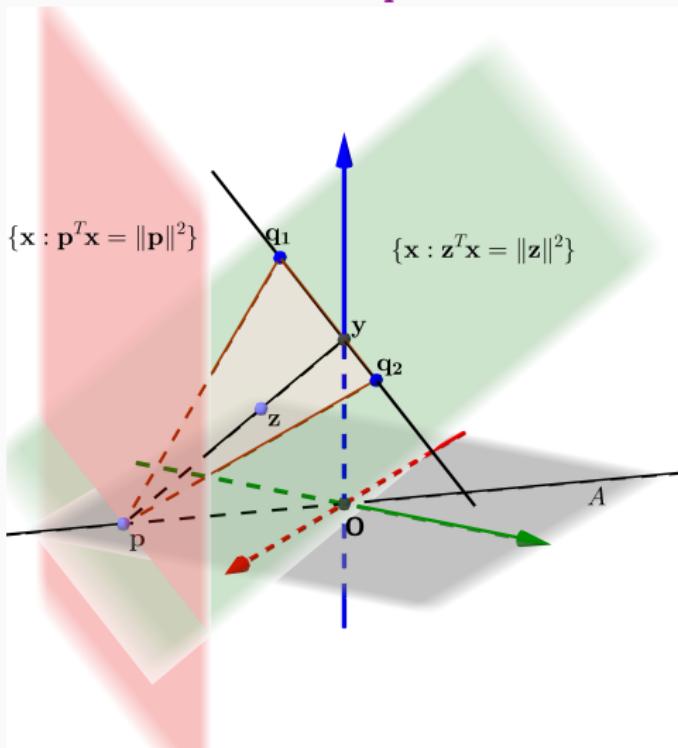
Lemma

For a point \mathbf{z} define $H_{\mathbf{z}} = \{\mathbf{w} \in \mathbb{R}^n : \mathbf{w} \cdot \mathbf{z} < \|\mathbf{z}\|_2^2\}$. Suppose that we have an instance of the minimum norm point problem in \mathbb{R}^d as follows: Some points, P , live in a proper linear subspace A and some, Q , in A^\perp . Let \mathbf{x} be the minimum norm point in $\text{aff } P$ and \mathbf{y} be the minimum norm point in $\text{aff}(P \cup Q)$. Then $H_{\mathbf{y}} \cap A = H_{\mathbf{x}} \cap A$.

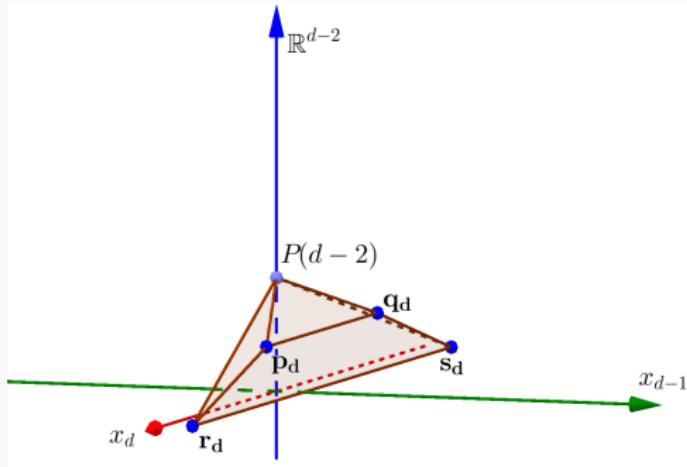
adding orthogonal points to the corral doesn't create any available points

Wolfe's Criterion under Addition of Orthogonal Point

adding orthogonal points to the corral doesn't create any available points



Sketch of Proof of Sequence $C(d)$: $C(d - 2)$



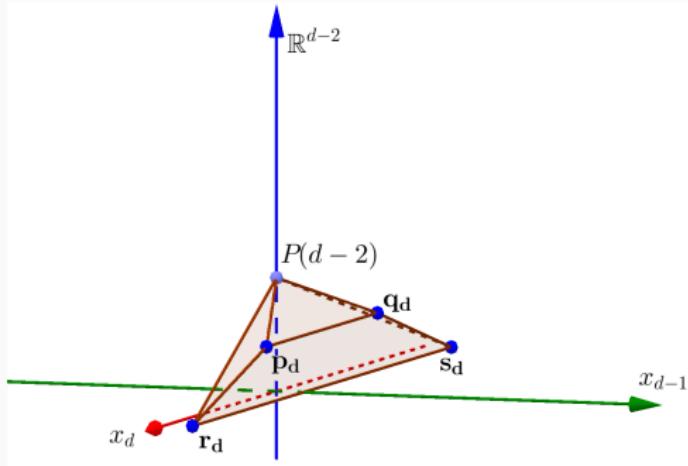
$$P(d) = \begin{pmatrix} P(d-2) & 0 & 0 \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & M_{d-2} \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & -(M_{d-2} + 1) \\ 0 & \frac{m_{d-2}}{4} & M_{d-2} + 2 \\ 0 & \frac{m_{d-2}}{4} & -(M_{d-2} + 3) \end{pmatrix}$$

\mathbf{o}_{d-2}^* : MNP($P(d-2)$)

$m_{d-2} = \|\mathbf{o}_{d-2}^*\|_\infty$

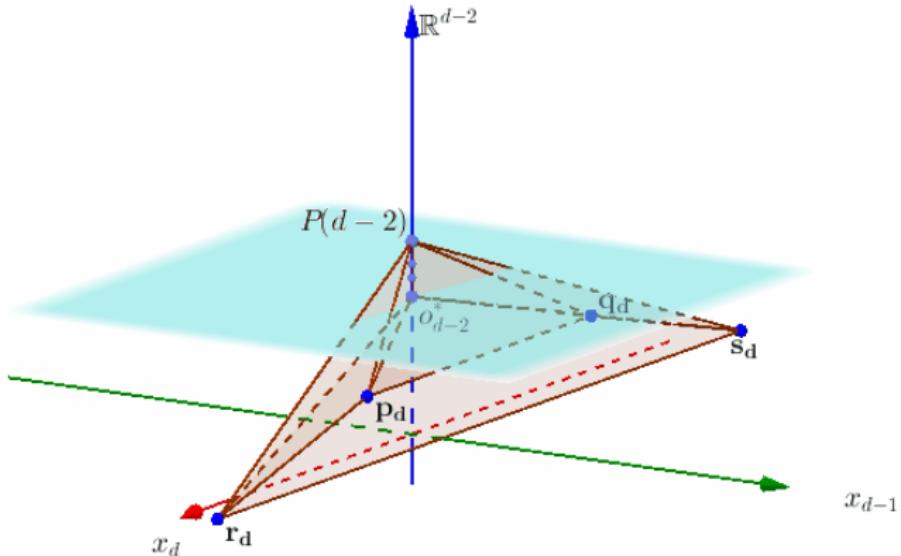
$M_{d-2} = \max_{\mathbf{p} \in P(d-2)} \|\mathbf{p}\|_1$

Sketch of Proof of Sequence $C(d)$: $C(d - 2)$

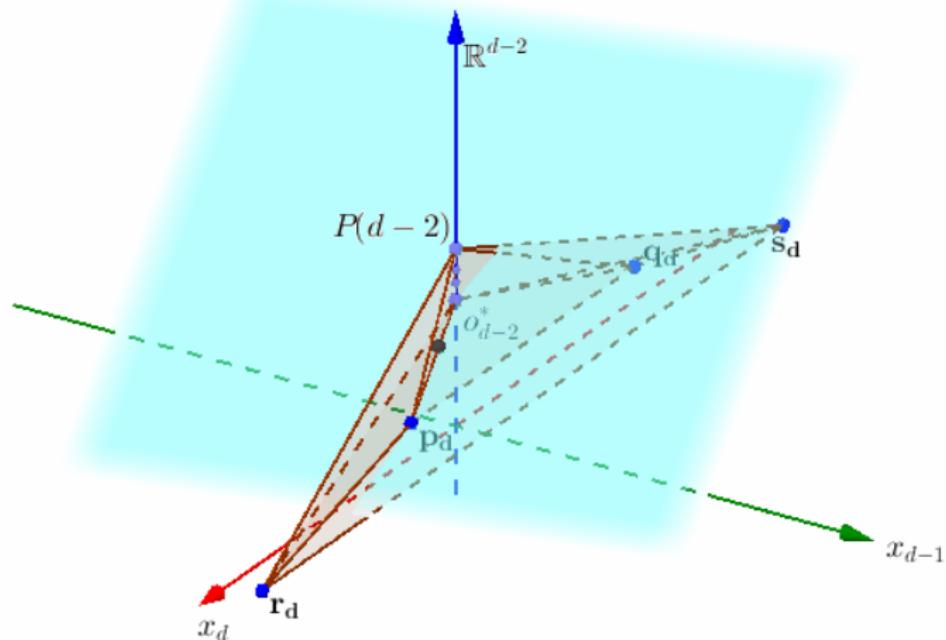


$$P(d) = \begin{pmatrix} P(d-2) & 0 & 0 \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & M_{d-2} \\ \frac{1}{2}\mathbf{o}_{d-2}^* & \frac{m_{d-2}}{4} & -(M_{d-2} + 1) \\ 0 & \frac{m_{d-2}}{4} & M_{d-2} + 2 \\ 0 & \frac{m_{d-2}}{4} & -(M_{d-2} + 3) \end{pmatrix} \quad \downarrow \parallel \parallel \quad \begin{aligned} \mathbf{o}_{d-2}^* &\text{: MNP}(P(d-2)) \\ m_{d-2} &= \|\mathbf{o}_{d-2}^*\|_\infty \\ M_{d-2} &= \max_{\mathbf{p} \in P(d-2)} \|\mathbf{p}\|_1 \end{aligned}$$

Sketch of Proof of Sequence $C(d)$: $O(d - 2)\mathbf{p}_d$

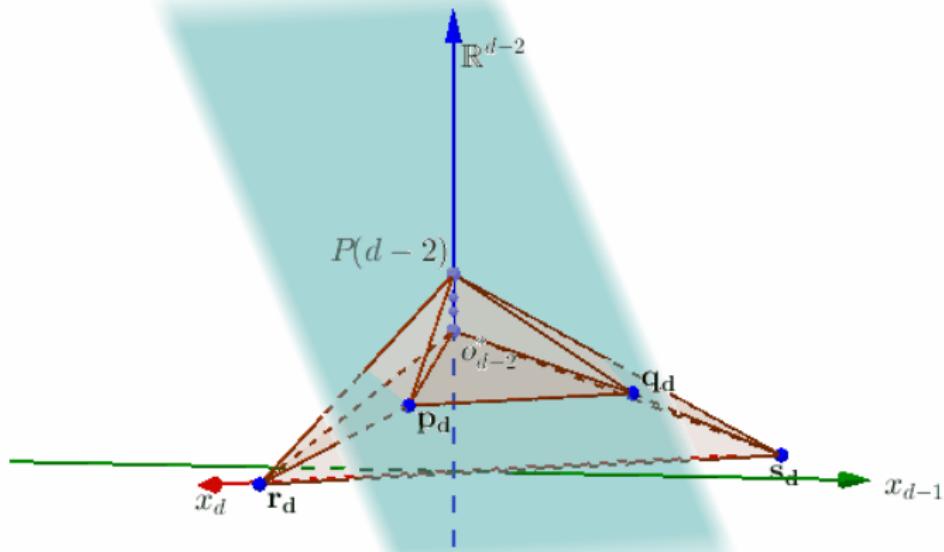


Sketch of Proof of Sequence $C(d)$: $O(d - 2)\mathbf{p}_d$

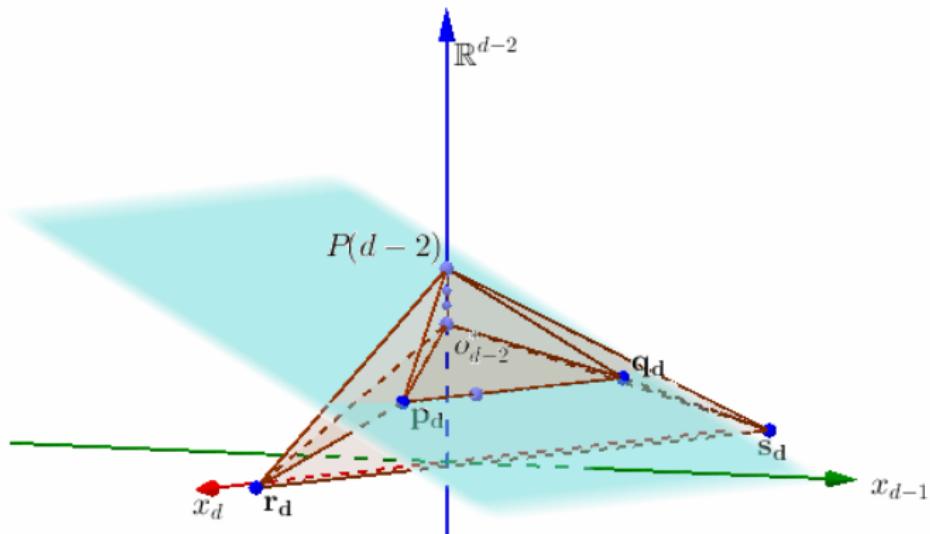


a corral with a point made from MNP and orthogonal directions is still a corral

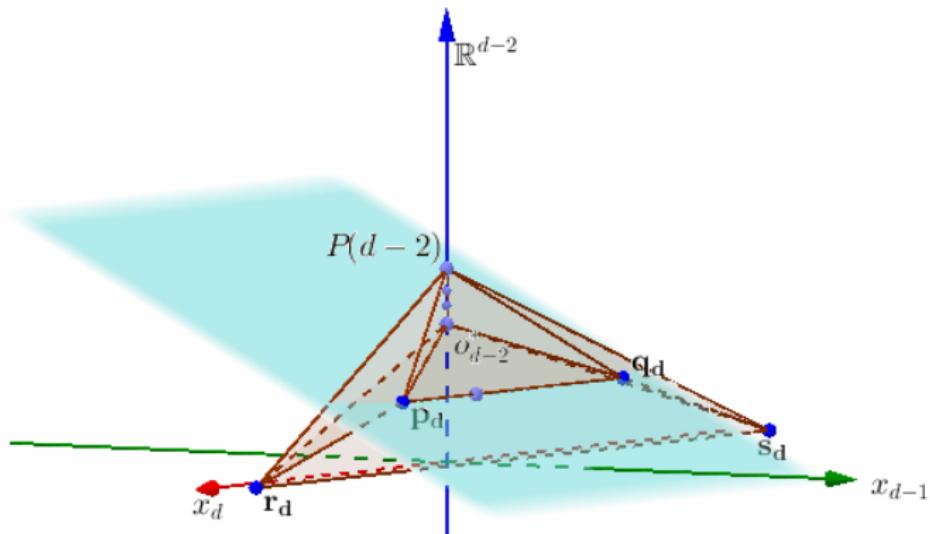
Sketch of Proof of Sequence $C(d)$: $\mathbf{p}_d \mathbf{q}_d$



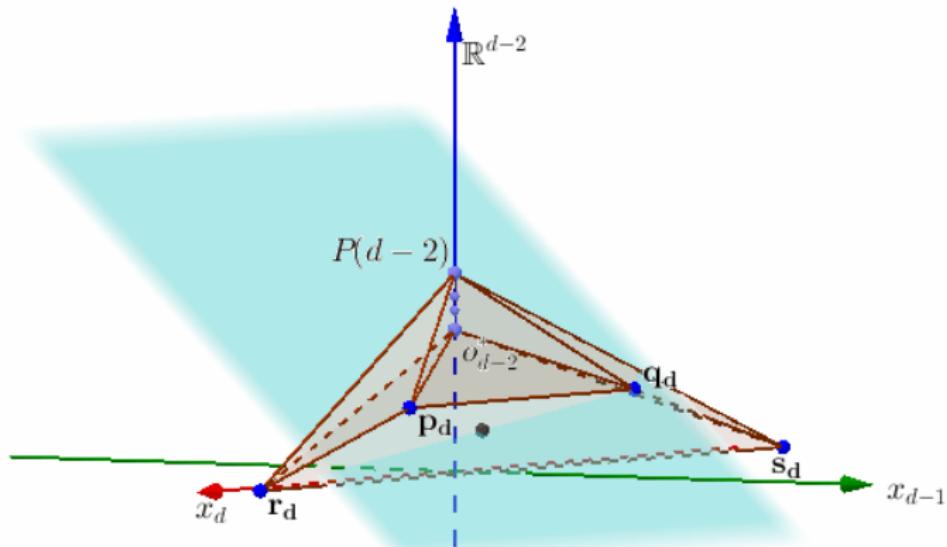
Sketch of Proof of Sequence $C(d)$: $\mathbf{p}_d \mathbf{q}_d$



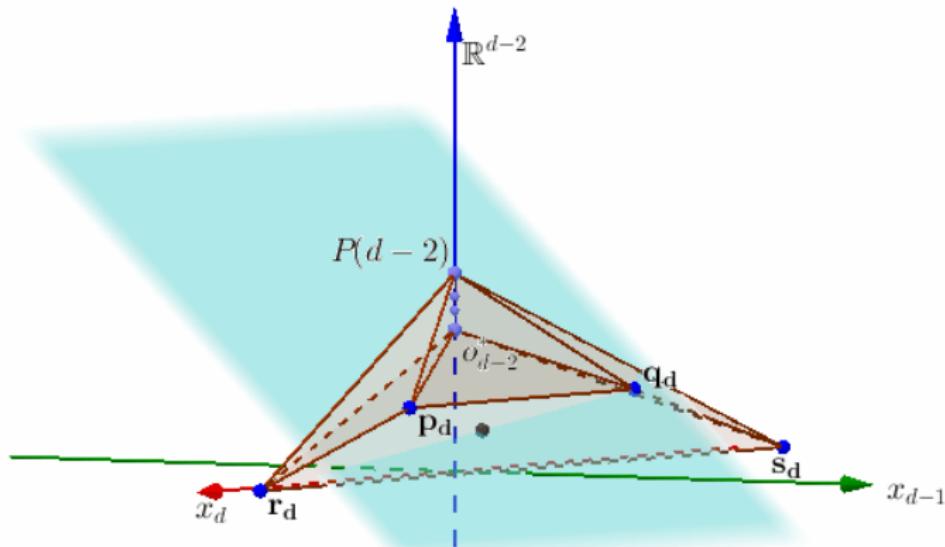
Sketch of Proof of Sequence $C(d)$: $\mathbf{q}_d \mathbf{r}_d$



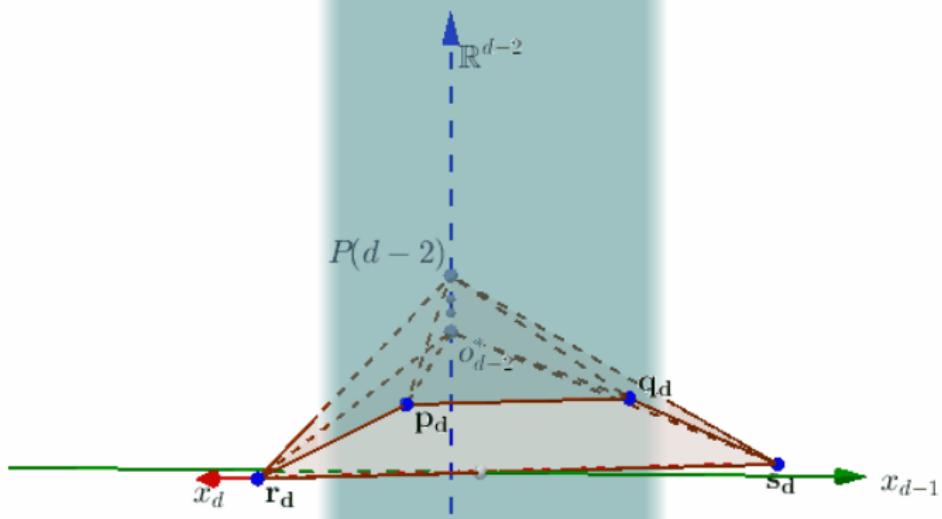
Sketch of Proof of Sequence $C(d)$: $\mathbf{q}_d \mathbf{r}_d$



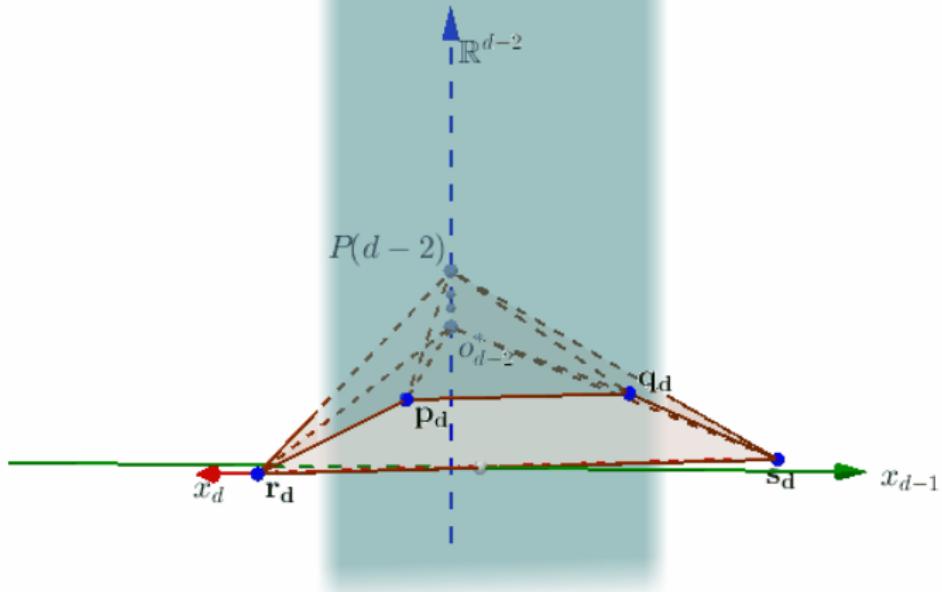
Sketch of Proof of Sequence $C(d)$: $r_d s_d$



Sketch of Proof of Sequence $C(d)$: $r_d s_d$



Sketch of Proof of Sequence $C(d)$: $C(d-2)\mathbf{r}_d\mathbf{s}_d$



- the union of orthogonal corrals is still a corral
- adding orthogonal points to the corral doesn't create any available points

Conclusions

Future Directions

Future Directions

1. Find an exponential example for Wolfe's method with linopt insertion rule.

Future Directions

1. Find an exponential example for Wolfe's method with linopt insertion rule.
2. Search for types of polytopes where Wolfe's method is polynomial (e.g. base polytopes).

Future Directions

1. Find an exponential example for Wolfe's method with linopt insertion rule.
2. Search for types of polytopes where Wolfe's method is polynomial (e.g. base polytopes).
3. Give an average (or smoothed) analysis of Wolfe's method.

Thanks...

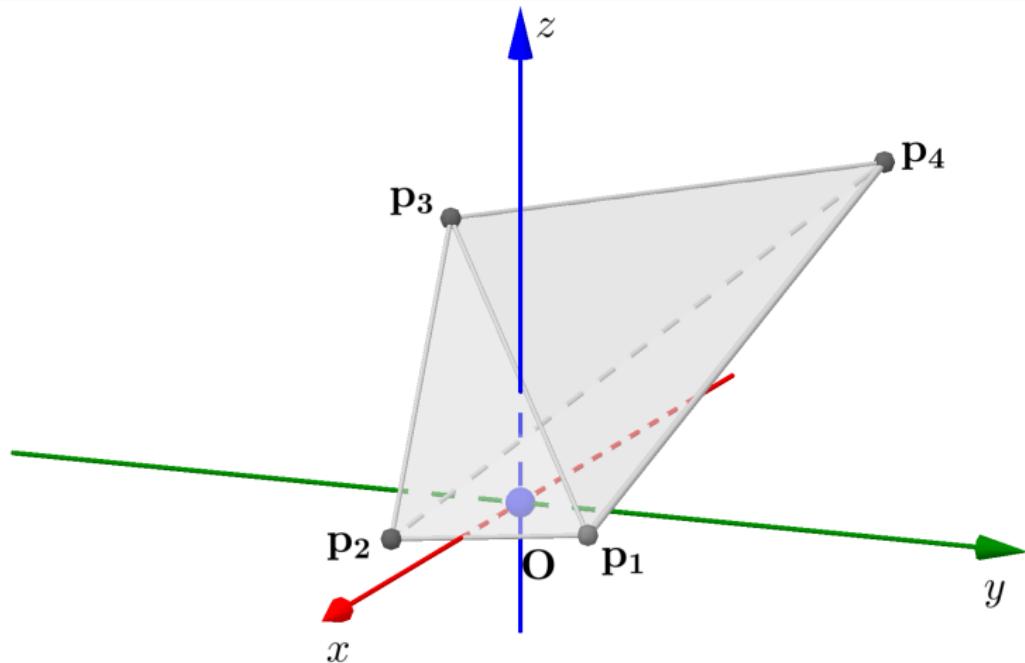


Questions?

- [1] I. Bárány and S. Onn.
Colourful linear programming and its relatives.
Mathematics of Operations Research, 22(3):550–567, 1997.
- [2] D. Chakrabarty, P. Jain, and P. Kothari.
Provable submodular minimization using wolfe's algorithm.
CoRR, abs/1411.0095, 2014.
- [3] J. A. De Loera, J. Haddock, and L. Rademacher.
**The minimum Euclidean-norm point on a convex polytope:
Wolfe's combinatorial algorithm is exponential.**
2017.
- [4] S. Fujishige, T. Hayashi, and S. Isotani.
**The minimum-norm-point algorithm applied to submodular
function minimization and linear programming.**
Citeseer, 2006.

Example: minnorm < linopt

$$P = \text{conv}\{(0.8, 0.9, 0), (1.5, -0.5, 0), (-1, -1, 2), (-4, 1.5, 2)\} \subset \mathbb{R}^3$$



Example: minnorm < linopt

Major Cycle	Minor Cycle	C
0	0	{ p_1 }
1	0	{ p_1, p_2 }
2	0	{ p_1, p_2, p_3 }
3	0	{ p_1, p_2, p_3, p_4 }
3	1	{ p_1, p_2, p_4 }

Major Cycle	Minor Cycle	C
0	0	{ p_1 }
1	0	{ p_1, p_4 }
2	0	{ p_1, p_4, p_3 }
2	1	{ p_1, p_3 }
3	0	{ p_1, p_3, p_2 }
4	0	{ p_1, p_2, p_3, p_4 }
4	1	{ p_1, p_2, p_4 }

Example: minnorm < linopt

Major Cycle	Minor Cycle	C
0	0	$\{p_1\}$
1	0	$\{p_1, p_2\}$
2	0	$\{p_1, p_2, p_3\}$
3	0	$\{p_1, p_2, p_3, p_4\}$
3	1	$\{p_1, p_2, p_4\}$

minnorm < linopt {

Major Cycle	Minor Cycle	C
0	0	$\{p_1\}$
1	0	$\{p_1, p_4\}$
2	0	$\{p_1, p_4, p_3\}$
2	1	$\{p_1, p_3\}$
3	0	$\{p_1, p_3, p_2\}$
4	0	$\{p_1, p_2, p_3, p_4\}$
4	1	$\{p_1, p_2, p_4\}$