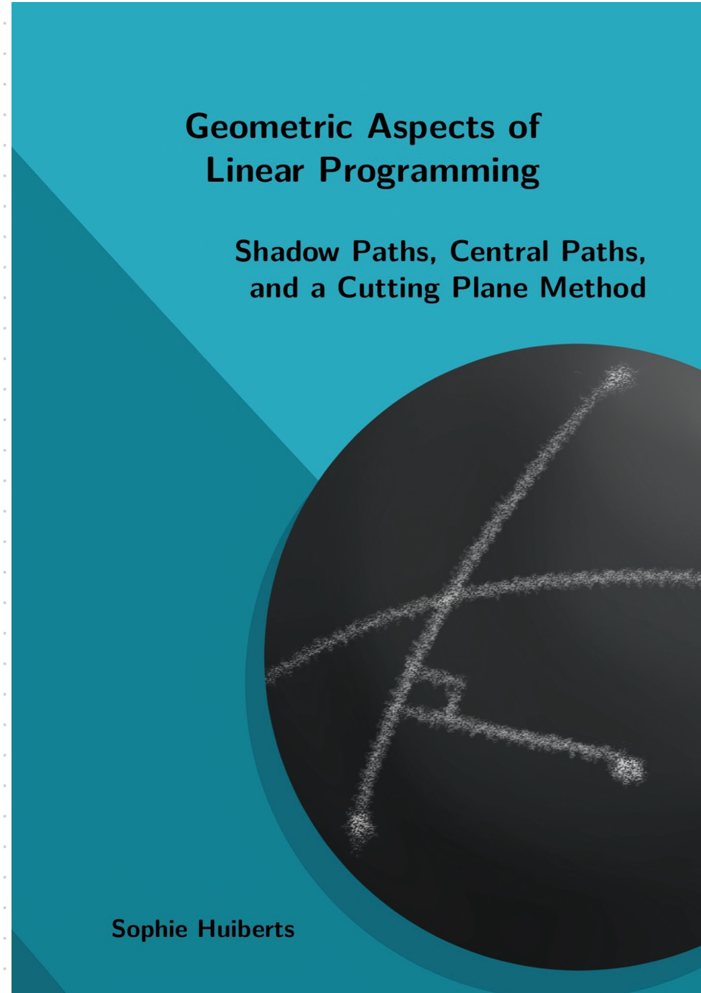
Gijs de Leve prize talk

Geometric Aspects of Linear Programming

Sophie Huiberts

CNRS & LIMOS



Linear Programming

maximize $c^T x$

subject to $Ax \le b$

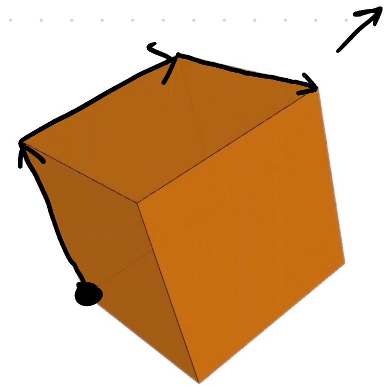
we get $A \in \mathbb{R}^{n \times d}$

$b \in \mathbb{R}^n$

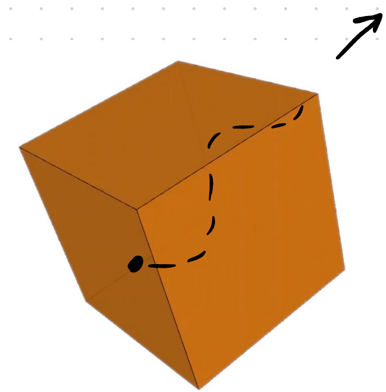
$c \in \mathbb{R}^d$

we compute $x \in \mathbb{R}^d$

simplex methods

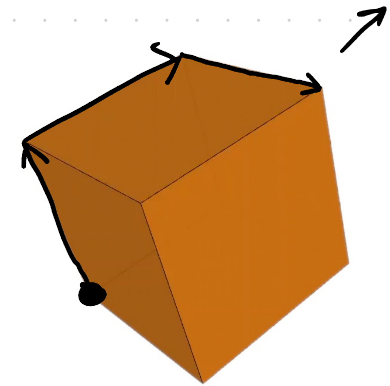


interior-point methods



Linear Programming

simplex methods



maximize $c^T x$
subject to $Ax \leq b$

we get $A \in \mathbb{R}^{n \times d}$
$b \in \mathbb{R}^n$
$c \in \mathbb{R}^d$

we compute $x \in \mathbb{R}^d$

In Practice

The simplex method visits $2(n+d)$ vertices before reaching an optimal one

Only a few documented cases where $> 10(n+d)$ iterations were performed

See, e.g., Shamir '87

Worst-case complexity

Theorem The simplex method has exponential worst-case complexity*

Klee Minty '72
Many, many others '72 - '23

*terms and conditions apply

Simplex method is

slow in theory

fast in practice

Average case analysis

Theorem There is a simplex method which,

Borgwardt '87

if the rows of A are iid uniform

from the sphere and $b=1$, visits

$d^2 n^{\frac{1}{d-1}}$ vertices in expectation

Average case analysis

Theorem There is a simplex method which,
Borgwardt '87
if the rows of A are iid uniform
from the sphere and $b=1$, visits
$d^2 n^{\frac{1}{d-1}}$ vertices in expectation

Theorem If $n \gg 2^{d^3}$ then this concentrates
Bonnet
Dadush
Grupel
Huiberts
Livshyts
'22
around the mean.

Extension to diameter

Theorem If $n \gg 2^{d^2}$ then the diameter is

Bonnet
Dadush
Grupel
Huiberts
Livshyts
'22

$$\geq d \, n^{\frac{1}{d-1}}$$

and

$$\leq d^2 \, n^{\frac{1}{d-1}}$$

with high probability

How Realistic is Average Case?

real world photo

random bitmap

Smooted analysis

Let $\bar{A} \in \mathbb{R}^{n \times d}$ have rows of norm ≤ 1.

$\bar{b} \in [-1, 1]^n$, $c \in \mathbb{R}^d$

Let $\hat{A}$, $\hat{b}$ have iid $N(0, \sigma^2)$ entries

Smooted analysis

Let $\bar{A} \in \mathbb{R}^{n \times d}$ have rows of norm ≤ 1.

$\bar{b} \in [-1, 1]^n$, $c \in \mathbb{R}^d$

Let $\hat{A}, \hat{b}$ have iid $N(0, \sigma^2)$ entries

Then

$$\max_{\bar{A}, \bar{b}, c} \mathbb{E}_{\hat{A}, \hat{b}} \left[\begin{array}{c} \text{time to solve} \\ \text{maximize } c^T x \\ \text{s.t } (\bar{A} + \hat{A}) x \leq \bar{b} + \hat{b} \end{array} \right] \leq \text{poly}(n, d, \sigma^{-1})$$

Why smoothed analysis?

independent measurement/numerical errors
do not conspire against your algorithm.

interpolate between worst case and
average case analysis.

shows algorithm is fast on average
in every large enough neighborhood

Smoothed analysis results

	Expected Number of Pivots
Spielman, Teng '01	$O(n^{86} d^{55} \sigma^{-30})$
Vershynin '09	$O(d^3 \log^3 n \sigma^{-4})$
Dadush, Huiberts '18	$O(d^2 \sqrt{\log n} \sigma^{-2})$
Huiberts, Lee, Zhang '23	$O(d^{13/4} \log^{7/4} n \sigma^{-3/2})$

Borgwardt '87	$\Omega(d^{3/2} \sqrt{\log n})$
Huiberts, Lee, Zhang '23	$\Omega(\min(2^d, \frac{1}{\sqrt{\sigma d \sqrt{\log n}}}))$

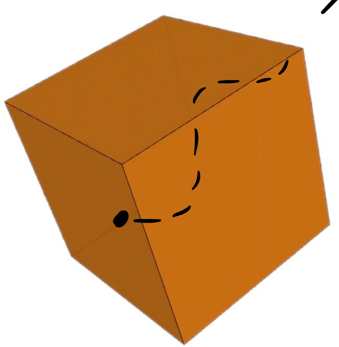
Linear Programming

maximize $c^T x$
subject to $Ax \leq b$

we get $A \in \mathbb{R}^{n \times d}$
$b \in \mathbb{R}^n$
$c \in \mathbb{R}^d$

we compute $x \in \mathbb{R}^d$

interior-point methods



Basic IPM analysis

Theorem An IPM can solve an LP using

$$\sqrt{n} \cdot L_{A,b,c}$$

linear system solves

$L_{A,b,c}$: # of bits to write down A, b, c

Basic IPMs are scale-invariant

For D positive diagonal, consider

$$\begin{aligned}&\text{maximize} && c^T x\\ &\text{subject to} && DAx \leq Db\end{aligned}$$

-> Same feasible region

-> Predictable change to central path

-> Same change to IPM path

Sophisticated IPM analysis

Theorem Specific IPMs can solve an LP using

$$\mathrm{poly}(n) \cdot L_A$$

linear system solves

L_A : # of bits to write down A

Question: Can we have both properties for a single IPM?

Monteiro, Tsuchiya 2003

Question: Can we have both properties for a single IPM?

Answer: Yes

Monteiro, Tsuchiya 2003 Dadush, Huiberts, Natura, Végh 2023

Separation oracle

$K \subseteq \mathbb{R}^d$ unknown convex.

$z + r\mathbb{B}_2^d \subseteq K \subseteq R\,\mathbb{B}_2^d$ with z, r unknown.

Can ask: is $x \in K$?

- Yes
- No because:

$a^T x > b$

$a^T y \leq b$ for all $y \in K$

Gradient oracle

$f: \mathbb{R}^d \to \mathbb{R}$ unknown convex

$\|\nabla f(y)\| \leq L$ for all $y \in K$

Can query $x \in \mathbb{R}^d$:

- value $f(x)$
- gradient $\nabla f(x)$

We want to minimize $f(x)$

over $x \in K$

with a small # of queries.

Ellipsoid method

1. Have an ellipsoid containing all optimal points
2. Query center point
3. if not in K:
   find smaller ellipsoid using cut
4. if in K:
   find smaller ellipsoid using gradient
5. go to 1.

Ellipsoid method

Pros

- Convergence guarantee

Cons

- Slow in practice

LP based loop

1. have a set S of valid constraints.
2. solve minimize $c^T x$
   subject to $a^T x \le b$
   for all $(a,b) \in S$
3. query LP optimal solution
4. if not in K
   add (a,b) to S

LP based loop

Pros

- mostly fast in practice
- LP solves are cheap

Cons

- no convergence guarantee
- only lower bounds no feasible points

Our new algorithm

1. have a set S of valid constraints.
2. solve a quadratic program
3. query its solution x^t
4. if not in K
   add (a,b) to S
5. if in K
   add $(\nabla f(x_t), \langle \nabla f(x_t), x_t \rangle)$ to S

Our new algorithm

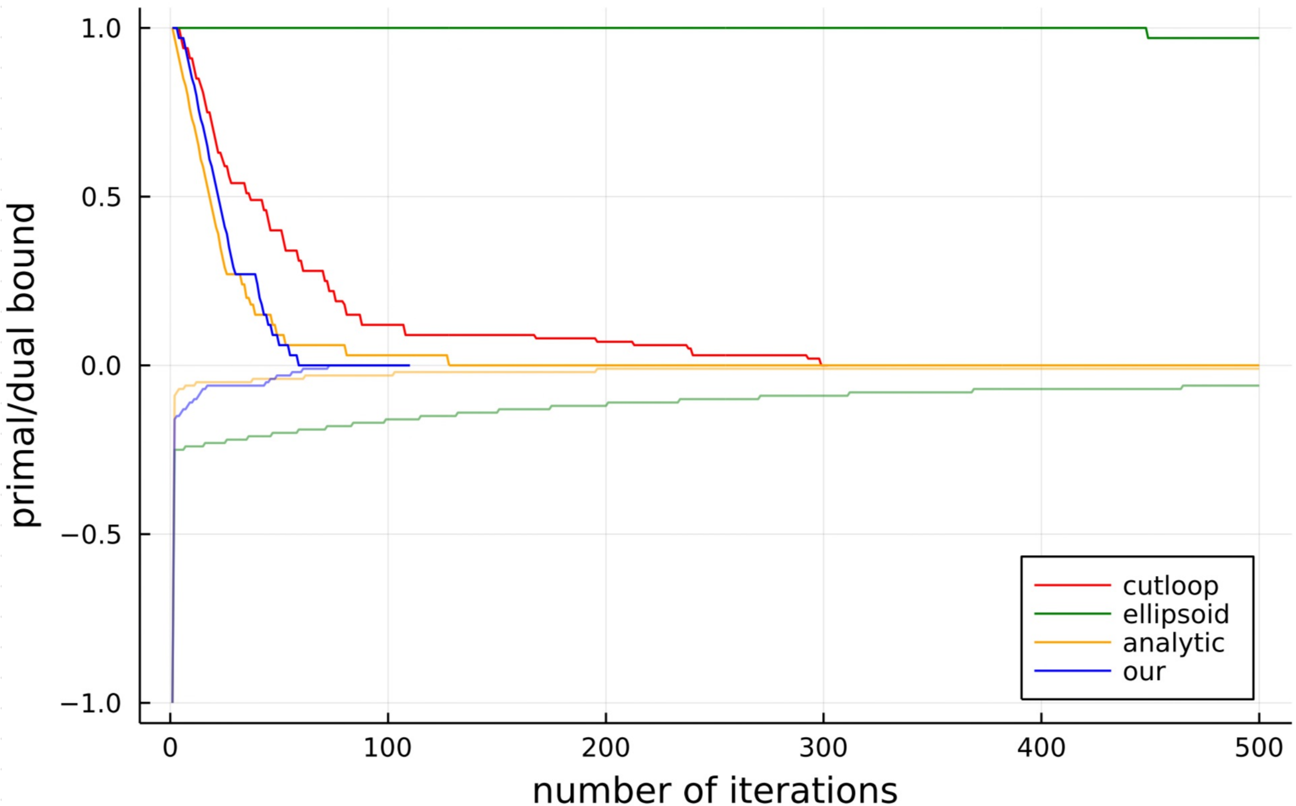
Pros

- Convergence guarantee
- good in experiments

Cons

- Need to solve a quadratic program every round

Experimental results



Linear programming

- diameter of random polyhedra
- smoothed analysis of simplex method
- scale-invariant IPM

Oracle model

- new cutting plane method