# Algorithmic and geometric aspects of combinatorial and continuous optimization





Antoine Deza, McMaster

## linear optimization

Given an *n*-dimensional vector *b* and an *n* x *d* (full row-rank) matrix *A* find, in any, a *d*-dimensional vector *x* such that :

 $Ax = b \qquad Ax = b \\ x \ge 0$ 

linear algebra

#### linear optimization

*"Can linear optimization be solved in strongly polynomial* time?" is listed by Smale (Fields Medal 1966) as one of the top problems for the XXI century

Polynomial : execution time bounded by a *polynomial* in *n*, *d*, and *input data length L* 

## linear optimization

Given an *n*-dimensional vector *b* and an *n* x *d* (full row-rank) matrix *A* find, in any, a *d*-dimensional vector *x* such that :

 $Ax = b \qquad Ax = b \\ x \ge 0$ 

linear algebra

#### linear optimization

*"Can linear optimization be solved in strongly polynomial* time?" is listed by Smale (Fields Medal 1966) as one of the top problems for the XXI century

**Strongly polynomial** : **polynomial** time; number of arithmetic operations bounded by a polynomial in the **dimension** of the problem (**independent** from the **input data length L**)

## linear optimization algorithms

Given an *n*-dimensional vector *b* and an *n* x *d* (full row-rank) matrix *A* and a *d*-dimensional cost vector *c*, solve : { max  $c^Tx : Ax = b, x \ge 0$  }

Simplex methods (Dantzig 1947) pivot-based, combinatorial, *not proven to be polynomial*, efficient in practice

Ellipsoid methods (Khachiyan 1979) polynomial ⇒ linear optimization is polynomial time solvable

Interior point methods (Karmarkar 1984) path-following, *polynomial*, efficient in practice

. . . . .

Primal-dual interior point (Kojima-Mizuno-Yoshise 1989)

Criss-cross (Terlaky 1983, Wang 1985, Chang 1979) Volumetric (Vaidya-Atkinson 1993, Anstreicher 1997) Monotonic build-up simplex (Anstreicher-Terlaky 1994)

## *linear optimization algorithms simplex methods*

Given an *n*-dimensional vector **b** and an *n* x **d** (full row-rank) matrix **A** and a **d**-dimensional cost vector **c**, solve : { max  $c^Tx : Ax = b, x \ge 0$  }

Simplex methods (Dantzig 1947): pivot-based, combinatorial, *not proven to be polynomial*, efficient in practice

- start from a *feasible basis*
- use a pivot rule
- ➢ find an optimal solution after a *finite number* of iterations
- most known pivot rules are known to be *exponential* (worst case); *efficient* implementations exist



## Given a a a

Simple

proven

> use

star

find

mos

(WOI

imp

 $\geq$ 

 $\geq$ 

 $\succ$ 

#### How Good Is the Simplex Algorithm?

VICTOR KLEE\*

Department of Mathematics, University of Washington, Seattle, Washington

AND

George J. Minty<sup>†</sup>

Department of Mathematics, Indiana University, Bloomington, Indiana

#### 1. INTRODUCTION

By constructing long "increasing" paths on appropriate convex polytopes, we show that the simplex algorithm for linear programs (at least with its most commonly used pivot rule, Dantzig [1]) is not a "good algorithm" in the sense of Jack Edmonds. That is, the number of pivots or iterations that may be required is not majorized by any polynomial function of the two parameters that specify the size of the program. In particular,  $2^d - 1$  iterations may be required in solving a linear program whose feasible region, defined by d linear inequality constraints in d nonnegative variables or by d linear equality constraints in 2d nonnegative variables, is projectively equivalent to a d-dimensional cube. Further, for each d there are positive constants  $\alpha_d$  and  $\beta_d$  such that

 $\alpha_d n^{\lfloor d/2 \rfloor} < \Xi(d, n) < \beta_d n^{\lfloor d/2 \rfloor} \quad \text{for all} \quad n > d,$ 

where  $\Xi(d, n)$  is the maximum number of iterations required in solving nondegenerate linear programs whose feasible regions are d-dimensional

r ≥ 0 } not

(1)

htrix **A** 

## *linear optimization algorithms simplex methods*

Klee-Minty 1972: edge-path followed by the simplex method with Dantzig's rule visits the 2<sup>*d*</sup> vertices of a *combinatorial* cube (n = 2d)  $\Rightarrow 2^{d} - 1$  pivots required to reach the optimum

Zadeh 1973 : bad network problems

Zadeh 1980 : deformed products and least entered rule

Amenta-Ziegler 1999 : deformed products

Friedmann 2011 : least entered rule is superpolynomial

Surveys : Terlaky-Zhang 1993, Ziegler 2004, Meunier 2013

... Avis-Friedmann 2016...

Dear Victor,

Please post this offer of "1000 to the first person who can find a counterexample to the least ented rule or prove it to be polynomial. The least ented rule enter the improving variable which has been ented least often.

Sincerely,

Norman Zadeh

Zadeh's offer (Ziegler 2004) (Avis' postface to Zadeh 1980 report, 2009 reprint)



David Avis, Norman Zadeh, Oliver Friedmann, Russ Caflish (IPAM 2011)

## Linear optimization algorithms (central path following) interior point methods

Given an *n*-dimensional vector **b** and an *n* x **d** (full row-rank) matrix **A** and a **d**-dimensional cost vector **c**, solve : { max  $c^Tx : Ax = b, x \ge 0$  }

#### **Interior Point Methods** :

path-following, *polynomial*, efficient in practice

- start from the analytic center
- follow the central path
- > converge to an optimal solution in  $O(\sqrt{nL})$  iterations
  - (L: input data length)



$$\max \quad c^{\mathrm{T}}x - \mu \quad \sum_{i} \ln(b - Ax)_{i}$$

 $\mu$ : central path parameter  $x \in \mathbf{P}$ :  $Ax \le b$ 

## *linear optimization* (some) combinatorial and geometric parameters

Tardos 1985: algorithm polynomial in *n*, *d*, and  $L_A$  (size of *A*)  $\Rightarrow$  strongly polynomial for minimum cost flow, bipartite matching etc. ... Orlin 1986, Kitahara-Mizuno 2011, Mizuno 2014, Mizuno-Sukegawa-Deza 2015...

Ye 2011 : strongly polynomial simplex for Markov Decision Problem

Vavasis-Ye 1996 :  $O(d^{3.5} \log(d \chi_A))$  primal-dual interior point method ... Megiddo-Mizuno-Tsuchiya 1998, Monteiro-Tsuchiya 2003...

Bonifas-Summa-Eisenbrand-Hähnle-Niemeier 2014:  $O(\mathbf{d} \ ^{4}\Delta_{\mathbf{A}}^{2} \log(\mathbf{d} \ \Delta_{\mathbf{A}}))$ diameter ( $\Delta_{\mathbf{A}}$  largest sub-determinant norm; Dyer-Frieze 1994)

Dadush-Hähnle 2015:  $O(d^{3}/\delta_{A} \log(d/\delta_{A}))$  expected (shadow vertex) simplex pivots ( $\delta_{A}$  curvature ;  $1/\delta_{A} \leq d \Delta_{A}^{2}$ )

Diameter (of a polytope) :

lower bound for the number of iterations for *pivoting* **simplex methods** 

**Curvature** (of the central path associated to a polytope) :

large curvature indicates large number of iterations for *path following* **interior point methods** 







Polytope P defined by n inequalities in dimension d

polytope : *bounded* polyhedron



Polytope P defined by n inequalities in dimension d



Polytope P defined by n inequalities in dimension d



**Diameter**  $\delta(P)$ : smallest number such that **any two vertices**  $(v_1, v_2)$  can be connected by a **path with at most**  $\delta(P)$  edges



**Diameter**  $\delta(\mathbf{P})$ : smallest number such that any two vertices can be connected by a path with at most  $\delta(\mathbf{P})$  edges

**Hirsch Conjecture** 1957 :  $\delta(\mathbf{P}) \leq \mathbf{n} - \mathbf{d}$ 



**Diameter**  $\delta(\mathbf{P})$ : smallest number such that any two vertices can be connected by a path with at most  $\delta(\mathbf{P})$  edges

Hirsch Conjecture 1957 :  $\delta(\mathbf{P}) \leq \mathbf{n} - \mathbf{d}$ 

disproved by Santos 2012 using construction with n = 2d



 $\lambda^{c}(\mathbf{P})$ : total curvature of the primal central path of { max  $\mathbf{c}^{\mathsf{T}}x : x \in \mathbf{P}$  }

 $\star \lambda^{c}(\mathbf{P})$ : redundant inequalities count



 $\lambda^{c}(\mathbf{P})$ : total curvature of the primal central path of { max  $\mathbf{c}^{\mathsf{T}}x : x \in \mathbf{P}$  }

 $\lambda(\mathbf{P})$ : largest total curvature  $\lambda^{\mathbf{c}}(\mathbf{P})$  over of all possible **c** 



 $\lambda^{c}(\mathbf{P})$ : total curvature of the primal central path of { max  $\mathbf{c}^{\mathsf{T}}x : x \in \mathbf{P}$  }

 $\lambda(\mathbf{P})$ : *largest* total curvature  $\lambda^{\mathbf{c}}(\mathbf{P})$  over of all possible **c** 

Continuous analogue of Hirsch Conjecture:  $\lambda(P) = O(n)$ (Deza-Terlaky-Zinchenko 2008)

✤ Dedieu-Shub 2005 hypothesis :  $\lambda(P) = O(d)$ 



 $\lambda^{c}(\mathbf{P})$ : total curvature of the primal central path of { max  $\mathbf{c}^{\mathsf{T}}x : x \in \mathbf{P}$  }

 $\lambda(\mathbf{P})$ : *largest* total curvature  $\lambda^{\mathbf{c}}(\mathbf{P})$  over of all possible **c** 

Continuous analogue of Hirsch Conjecture:  $\lambda(P) = O(n)$ (Deza-Terlaky-Zinchenko 2008)

✤ Dedieu-Shub 2005 hypothesis :  $\lambda(P) = O(d)$ 

✤ Deza-Terlaky-Zinchenko 2008 : polytope such that:  $\lambda(P) = \Omega(2^d)$ 



 $\lambda^{c}(\mathbf{P})$ : total curvature of the primal central path of { max  $\mathbf{c}^{\mathsf{T}}x : x \in \mathbf{P}$  }

 $\lambda(\mathbf{P})$ : largest total curvature  $\lambda^{\mathbf{c}}(\mathbf{P})$  over of all possible **c** 

Continuous analogue of Hirsch Conjecture:  $\lambda(P) = O(n)$ (Deza-Terlaky-Zinchenko 2008)

disproved by Allamigeon-Benchimol-Gaubert-Joswig 2014

Dedieu-Shub 2005 hypothesised  $\lambda(\mathbf{P}) = O(\mathbf{d})$ Dedieu-Malajovich-Shub 2005 proved it is true *on average* (de Loera-Sturmfels-Vinzant 2012)

Deza-Terlaky-Zinchenko 2008: **P** with exponential  $\lambda(\mathbf{P})$  and  $\mathbf{n} = \Omega(2^d)$ 

Continuous analogue of Hirsch Conjecture:  $\lambda(P) = O(poly(n,d))$ 

Allamigeon-Benchimol-Gaubert-Joswig 2014 : linear optimization instance  $(2n \approx 3d)$  for which central-path following methods require  $\Omega(2^{d/2})$  iterations

#### ⇒ path-following interior-point methods are not strongly polynomial

Result obtained using *tropical geometry*, which reduces the complexity analysis to a *combinatorial* problem



Arrangement A defined by *n* hyperplanes in dimension d



Simple arrangement:

*n* > *d* and any *d* hyperplanes **intersect** at a **unique distinct point** 



For a simple arrangement, the number of **bounded cells**  $I = \begin{pmatrix} n-1 \\ d \end{pmatrix}$ 



 $\lambda^{c}(\mathbf{A}) : \text{ average value of } \lambda^{c}(\mathbf{P}_{i}) \text{ over the bounded cells } \mathbf{P}_{i} \text{ of } \mathbf{A}:$  $\lambda^{c}(\mathbf{A}) = \underbrace{\sum_{i=1}^{i=I} \lambda^{c}(\mathbf{P}_{i})}_{I} \text{ with } I = \binom{n-1}{d}$ 

 $\star \lambda^{c}(P_{i})$ : redundant inequalities count



 $\lambda^{c}(A)$ : average value of  $\lambda^{c}(P_{i})$  over the bounded cells  $P_{i}$  of A:

 $\lambda(A)$  : largest value of  $\lambda^{c}(A)$  over all possible c



 $\lambda^{c}(A)$ : average value of  $\lambda^{c}(P_{i})$  over the bounded cells  $P_{i}$  of A:

 $\lambda(A)$  : largest value of  $\lambda^{c}(A)$  over all possible c

Dedieu-Malajovich-Shub 2005:  $\lambda(\mathbf{A}) \leq 2\pi \mathbf{d}$ 

(de Loera-Sturmfels-Vinzant 2012)

✤ A : simple arrangement



 $\delta(A)$ : average diameter of a bounded cell of A:

✤ A : simple arrangement



 $\delta(\mathbf{A}) : \text{ average diameter of a bounded cell of } \mathbf{A}:$   $\delta(\mathbf{A}) = \underbrace{\sum_{i=1}^{i=I} \delta(P_i)}_{I} \quad \text{with } I = \binom{n-1}{d}$ 

♦ δ(A): average diameter ≠ diameter of A
ex: δ(A)= 1.333...



 $\delta(\mathbf{A}) : \text{ average diameter of a bounded cell of } \mathbf{A}:$   $\delta(\mathbf{A}) = \underbrace{\sum_{i=1}^{i=I} \delta(P_i)}_{\mathbf{I}} \quad \text{with } \mathbf{I} = \binom{n-1}{d}$ 

\*  $\delta(\mathbf{P}_i)$ : only *active* inequalities count



 $\delta(A)$ : average diameter of a bounded cell of A:

**Conjecture** :  $\delta(A) \leq d$ (Deza-Terlaky-Zinchenko 2008)

(discrete analogue of Dedieu-Malajovich-Shub result)



Terlaky-Mut 2014 : Sonnevend curvature

Hirsch bound $\delta(P) \leq n - d \operatorname{imp}$	blies $\delta(A) \leq d \frac{n+1}{n-1}$	
Hirsch conjecture <i>holds</i> for $d = 2$ : $\delta(A) \le 2 \frac{n+1}{n-1}$		
Hirsch conjecture <i>holds</i> for $d = 3$ : $\delta(A) \leq 3 \frac{n+1}{n-1}$		
Larman 1970, Barnette 1974 δ <b>(Ρ) ≤ n</b> 2 <sup>d</sup> /12 (Labbé-Manneville-Santos 2015)		
Kalai-Kleitman 1992	$\delta(\boldsymbol{P}) \leq \boldsymbol{n}^{\log \boldsymbol{d}+2}$	
Todd 2014	$\delta(\boldsymbol{P}) \leq \left(\boldsymbol{n} - \boldsymbol{d}\right)^{\log \boldsymbol{d}}$	
Sukegawa-Kitahara 2015	$\delta(\boldsymbol{P}) \leq \left(\boldsymbol{n} - \boldsymbol{d}\right)^{\log(\boldsymbol{d} - 1)}$	

Sukegawa 2016, 2018

Borgwardt-de Loera-Finhold 2016 (Hirsch holds for transportation polytopes)



Haimovich's probabilistic analysis of shadow-vertex simplex method, Borgwardt 1987
Forge-Ramírez Alfonsín 2001: counting *k*-face cells of *A*\*

**Diameter** (of a polytope) :

lower bound for the number of iterations for the **simplex method** (*pivoting methods*)

lower bound :  $(1 + \varepsilon) (n - d)$  upper bound:  $(n - d)^{\log d}$ 

**Curvature** (of the central path associated to a polytope) :

large curvature indicates large number of iteration for *central path following* **interior point methods** 

**lower bound** :  $\Omega(2^{d/2})$  with  $2^n \approx 3^d$  upper bound:  $2\pi d \binom{n-1}{d}$ 

Allamigeon-Benchimol-Gaubert-Joswig 2018 exponential lower bound for  $\lambda(\mathbf{P})$  contrasts with the belief that a polynomial upper bound for  $\delta(\mathbf{P})$  might exist, e.g.  $\delta(\mathbf{P}) \leq d(n - d)/2$ 

 $\Delta(d, n)$ : largest diameter over all *d*-dimensional polytopes with *n* facets



 $\Delta(4,10) = 5, \Delta(5,11) = 6$  Goodey 1972

 $\Delta(d, n)$ : largest diameter over all *d*-dimensional polytopes with *n* facets



 $\Delta(4,11) = \Delta(6,12) = 6$  Bremner-Schewe 2011

 $\Delta(d, n)$ : largest diameter over all *d*-dimensional polytopes with *n* facets



 $\Delta(4,12) = \Delta(5,12) = 7$  Bremner-Deza-Hua-Schewe 2013

 $\Delta(d, n)$ : largest diameter over all *d*-dimensional polytopes with *n* facets

Characterize all combinatorial types of paths of length *k* 

Find necessary conditions for a (chirotope of a) polytope to admit an embedding of a *k*-path on its boundary (without shortcuts)

If *no* such (chirotope of a) polytope exists:  $\Delta(d, n) \neq k$ 



in the dual setting



in the dual setting



#### in the dual setting



#### in the dual setting



#### in the dual setting

a vertex path of a simple polytope becomes a simplicial facet path



(1, 4) (2, 5) (3, 6) ... (10, 12) (9, 13)

#### in the dual setting



#### in the dual setting

a vertex path of a simple polytope becomes a simplicial facet path



(1, 4) (2, 5) (3, 6) ... (10, 12) (9, 13)

generating non-revisiting path from restricted growth strings

[1, 2, 3, 1, 4, 3, 5, 4, 6] length *k d* symbols

generating non-revisiting path from restricted growth strings

[1, 2, 3, 1, 4, 3, 5, 4, 6] length **k d** symbols

obtainable from set partitionings

{ { 1, 4 }, { 2 }, { 3, 6 }, { 5, 8 }, { 7 }, { 9 } } *k* elements *d* subsets

generating non-revisiting path from restricted growth strings

[1, 2, 3, 1, 4, 3, 5, 4, 6] length **k d** symbols

obtainable from set partitionings

{ { 1, 4 }, { 2 }, { 3, 6 }, { 5, 8 }, { 7 }, { 9 } } *k* elements *d* subsets



generating revisiting path from non-revisiting path (by identifying all possible revisits in a non-revisiting path, and avoiding introducing an extra edge)



upper bounds on the number of revisits and drops to consider

for paths length *k* involving *i* revisits and *j* drops:

i - j = k + d - n $0 \le i \le k - d$  $0 \le j \le n - 2d$ 

#### chirotopes

chirotope  $\chi$  is a function from  $E^r$  to {-, 0, +} ( $E = \{ 1, ..., n \}, r = d + 1$ )

(simplicial polytope) chirotope  $\Rightarrow \chi : E^r \rightarrow \{-, +\}$  (for rank *r* tuples)

geometric interpretation:

 $\chi$  determines the orientation of a set of vertices



#### chirotopes

chirotope  $\chi$  is a function from  $E^r$  to { -, 0, + } ( $E = \{ 1, ..., n \}, r = d + 1$ )

(simplicial polytope) chirotope  $\Rightarrow \chi : E^r \rightarrow \{-, +\}$  (for rank *r* tuples)

geometric interpretation:

 $\chi$  determines the orientation of a set of vertices



necessary condition for a (chirotope of a) polytope

 $\chi$  alternates (swap switches the sign)  $\chi$  satisfies the 3-term Grassmann-Plücker relations

necessary condition for a (chirotope of a) polytope

 $\chi$  alternates (swap switches the sign)  $\chi$  satisfies the 3-term Grassmann-Plücker relations



#### necessary condition for a (chirotope of a) polytope

 $\chi$  alternates (swap switches the sign)  $\chi$  satisfies the 3-term Grassmann-Plücker relations



 $\chi(\sigma, x_1, x_2)\chi(\sigma, x_3, x_4)$  $\chi(\sigma, x_1, x_3)\chi(\sigma, x_4, x_2)$  $\chi(\sigma, x_1, x_4)\chi(\sigma, x_2, x_3)$ 

one positive, one negative

#### necessary condition for a (chirotope of a) polytope

for any given **k**-path

satisfies the Grassmann-Plücker constraints embeds the *k*-path on the boundary of a polytope (without shortcut )

#### necessary condition for a (chirotope of a) polytope

for any given **k**-path

satisfies the Grassmann-Plücker constraints embeds the *k*-path on the boundary of a polytope (without shortcut )

use a satisfiability solver

SAT variable true / false

chirotope sign plus / minus

#### necessary condition for a (chirotope of a) polytope

$$16\binom{n}{d-1}\binom{n-d+1}{4}$$

for any given *k*-path

constraints

satisfies the Grassmann-Plücker constraints embeds the *k*-path on the boundary of a polytope (without shortcut )

use a *satisfiability solver* 

SAT variable true / false

chirotope sign plus / minus

#### embeds the **k**-path on the boundary of a polytope

every element of path complex lies on the boundary (i.e. no 'facet' cuts through the interior of the polytope)



(1, 2, 4) forms a facet on the boundary

as do (2, 3, 4) and (2, 3, 5)

but (1, 2, 3) is *not* a facet

#### embeds the **k**-path on the boundary of a polytope

every element of path complex lies on the boundary (i.e. no 'facet' cuts through the interior of the polytope)



embed the *k*-path on the boundary of a polytope *without shortcut* 



ensure that the *k*-path is a shortest path (i.e. no other shorter paths)

embed the *k*-path on the boundary of a polytope *without shortcut* 



ensure that the *k*-path is a shortest path (i.e. no other shorter paths)

embed the *k*-path on the boundary of a polytope *without shortcut* 



ensure that the *k*-path is a shortest path (i.e. no other shorter paths)

## $\delta(4,12) \neq 8$

# revisits / drops	# completed
0	160
1	1258
2	5168
3	7398
4	1512

#### recent progresses



 $\delta(4,12) = \delta(5,12) = 7$  Bremner-D.-Hua-Schewe (2013)

Pivot path: 123 - 234 - 345 - 456 - 467

Column presentation: 123 423 453 456 476

1<sup>st</sup> column changes (replaced by next available number) 2<sup>nd</sup> column changes (replaced by next available number)...

{1,2,3,2,...}

1 in position 1 2 in positions 2 and 4 3 in position 3

 $[\{1\},\!\{2,\!4\},\!\{3\},\ldots]$