13 Linear Programming Complexity

13.1 Number of Simplex Iterations

Computational Efficiency of the Simplex Method

Observation

The computational efficiency of the simplex method is determined by

i the computational effort of each iteration;

ii the number of iterations.

Question: How many iterations are needed in the worst case?

Idea for negative answer (lower bound)

Describe

- a polyhedron with an exponential number of vertices;
- a path that visits all vertices and always moves from a vertex to an adjacent one that has lower costs.

Unit cube

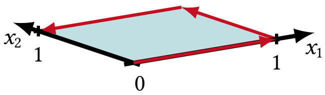
Consider the unit cube in $\mathbb{R}^n$, defined by the constraints

$$0 \le x_i \le 1, \quad i = 1, \ldots, n$$

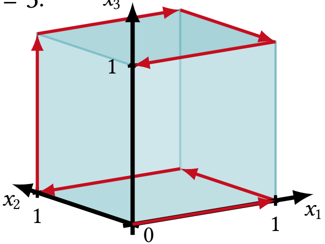
The unit cube has

- 2^n vertices;
- a *spanning path*, i.e., a path traveling along the edges of the cube visiting each vertex exactly once.

$n = 2$:



$n = 3$:



Efficiency of the Simplex Method (Cont.)

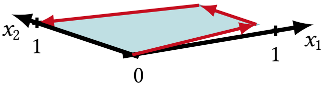
Klee-Minty cube

Consider a perturbation of the unit cube in $\mathbb{R}^n$, defined by the constraints

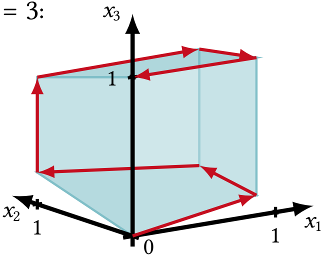
$$0 \leq x_1 \leq 1,$$
$$\varepsilon x_{i-1} \leq x_i \leq 1 - \varepsilon x_{i-1}, \quad i = 2, \dots, n$$

for some $\varepsilon \in (0, 1/2)$.

$n = 2$:



$n = 3$:



Efficiency of the Simplex Method (Cont.)

Klee-Minty cube

$$0 \le x_1 \le 1,$$

$$\varepsilon x_{i-1} \le x_i \le 1 - \varepsilon x_{i-1}, \quad i = 2, \ldots, n$$

Theorem 13.1 Consider the linear programming problem of minimizing $-x_n$ subject to the constraints above. Then,

a. the feasible set has 2^n vertices;
b. the vertices can be ordered so that each one is adjacent to and has lower cost than the previous one;
c. there exists a pivoting rule under which the simplex method requires $2^n - 1$ changes of basis before it terminates.

Remark

Such ‘bad’ instances exist for (almost) all popular pivoting rules.

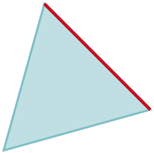
Diameter of Polyhedra and Polytopes

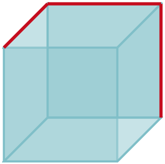
Definition 13.2 (Graph, combinatorial diameter) Let $P \subset \mathbb{R}^n$ be a polyhedron.

i. The graph (1-skeleton) $G(P)$ consists of the vertices and edges of P.
ii. For vertices v, w of P, the distance $\delta_P(v, w)$ is the minimum length of a path connecting v and w in $G(P)$.
iii. $\delta(P) := \max\{\delta_P(v, w) : v, w \text{ vertices of } P\}$ is called the (combinatorial) diameter of P.

Examples:

- $\delta(P) = 1$ for the n-dimensional simplex P
- $\delta(C_n) = n$ for the n-dimensional hypercube C_n





Diameter of Polyhedra and Polytopes (Cont.)

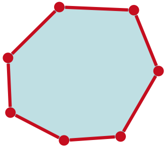
Observation: Diameter of the polyhedron of an LP is lower bound on # steps required by simplex method (no matter which pivoting rule).

Definition 13.3 For integers n and m let

$$\Delta(n, m) := \max\{\delta(P) \mid P \subseteq \mathbb{R}^n \text{ polytope given by } m \text{ inequalities}\}$$

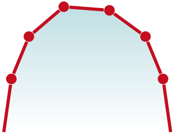
$$\Delta_u(n, m) := \max\{\delta(P) \mid P \subseteq \mathbb{R}^n \text{ polyhedron given by } m \text{ inequalities}\}$$

Examples.



$$\Delta(2, 7) = \left\lfloor \frac{7}{2} \right\rfloor = 3$$

$$\Delta(2, m) = \left\lfloor \frac{m}{2} \right\rfloor$$



$$\Delta_u(2, 7) = 7 - 2 = 5$$

$$\Delta_u(2, m) = m - 2$$

Remark. $\Delta(n, n + 1) = 1$ and $\Delta(n, m) \leq \Delta(n, m + 1)$.

Hirsch Conjecture (1957)

$$\Delta(n, m) \leq m - n$$

Remarks.

- Hirsch Conjecture is known to be true if $n \leq 3$ or $m \leq n + 5$.
- Hirsch Conjecture is false for unbounded polyhedra:

$$\Delta_u(n, m) \geq m - n + \left\lfloor \frac{n}{4} \right\rfloor \qquad \text{for } m \geq 2n.$$

- Known upper bounds:

$$\Delta(n, m) \leq m \cdot 2^{n-3} \qquad \text{(Barnette 1969; Larman 1970)}$$

$$\Delta(n, m) \leq m^{\log n + 2} \qquad \text{(Kalai 1992; Kalai \& Kleitman 1992)}$$

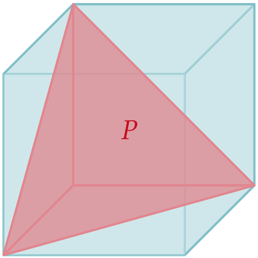
- Hirsch Conjecture disproven for $n = 43$, $m = 86$

  F. Santos, A counterexample to the Hirsch Conjecture, Ann. Math., 2012

Polynomial Hirsch Conjecture

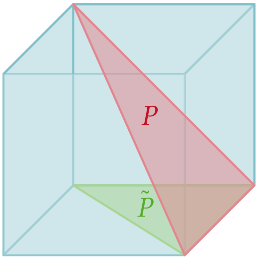
$$\Delta(n, m) \leq poly(m, n)$$

Definition 13.4 Let $[0,1]^n$ be the n-dimensional unit cube with vertices $\{0,1\}^n$. A polytope $P \subseteq \mathbb{R}^n$ is called 0/1-polytope if all vertices of P lie in $\{0,1\}^n$.



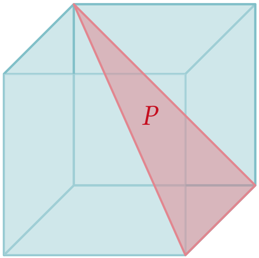
Lemma 13.5 Let $P \subseteq \mathbb{R}^n$ be 0/1-polytope with $\dim P \leq n-1$. There is a 0/1-polytope $\tilde{P}$ that is a projection of P such that $G(P)$ and $G(\tilde{P})$ are isomorphic.

(graphs $G=(V,E)$ and $\tilde{G}=(\tilde{V},\tilde{E})$ are isomorphic if there is a bijection $f : V \longrightarrow \tilde{V}$ such that $\{v,w\} \in E$ iff $\{f(v), f(w)\} \in \tilde{E}$)



Proof of Lemma 13.5

- let $P \subseteq \mathbb{R}^n$, $\dim(P) \leq n-1$ be a 0/1 polytope
- let $a \in \mathbb{R}^n \setminus \{0\}$, $\alpha \in \mathbb{R}$ be such that $a^\top x = \alpha$ for all $x \in P$, wlog $a_n \neq 0$
- let $\pi : \mathbb{R}^n \longrightarrow \mathbb{R}^{n-1}$ be defined as $(x_1, \ldots, x_{n-1}, x_n) \longmapsto (x_1, \ldots, x_{n-1})$
- let $\tilde{P} = \pi(P)$

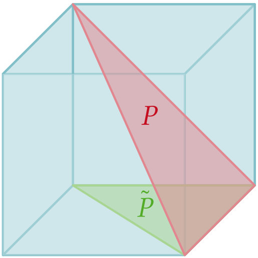


Claim: $\pi : P \longrightarrow \tilde{P}$ is a bijection

- for $y \in \tilde{P}$, the only $x \in P$ with $\pi(x) = y$ has $x_n = \frac{\alpha - \pi(a)^\top y}{a_n}$

Proof of Lemma 13.5 (Cont.)

Claim: x vertex of $P \iff \pi(x)$ vertex of $\tilde{P}$

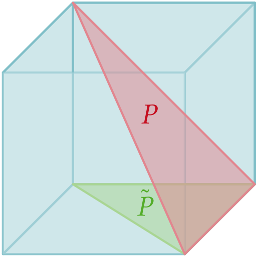


- we use the equivalence between vertices and extreme points of Theorem 3.21
- if there are $y, z \in P$, $\lambda \in (0,1)$ such that $x = \lambda y + (1-\lambda)z$, then $\pi(x) = \lambda\pi(y) + (1-y)\pi(z)$
- if there are $\tilde{y}, \tilde{z} \in \tilde{P}$ such that $\tilde{x} = \lambda\tilde{y} + (1-\lambda)\tilde{z}$, then

$$\begin{aligned}\pi^{-1}(\tilde{x}) &= \left(\tilde{x}_1, \ldots, \tilde{x}_{n-1}, \tfrac{\alpha - \pi(a)^\top \tilde{x}}{a_n}\right) \\ &= \lambda\left(\tilde{y}_1, \ldots, \tilde{y}_{n-1}, \tfrac{\alpha - \pi(a)^\top \tilde{y}}{a_n}\right) + (1-\lambda)\left(\tilde{z}_1, \ldots, \tilde{z}_{n-1}, \tfrac{\alpha - \pi(a)^\top \tilde{z}}{a_n}\right) \\ &= \lambda\pi^{-1}(\tilde{y}) + (1-\lambda)\pi^{-1}(\tilde{z})\end{aligned}$$

Proof of Lemma 13.5 (Cont.)

Claim: vertices x, y adjacent in P
$\iff$ vertices $\pi(x), \pi(y)$ adjacent in $\tilde{P}$



- let x, y be non-adjacent vertices in P
- $\frac{x+y}{2} = \sum_i \mu_{v_i} v_i$ for some vertices v_i of P and coefficients $\mu_i \geq 0$ with $\sum_i \mu_{v_i} = 1$ and $\mu_x + \mu_y < 1$
- then $\frac{\pi(x)+\pi(y)}{2} = \sum_i \mu_{v_i} \pi(v_i)$, i.e., $\pi(x)$ and $\pi(y)$ are non-adjacent in $\tilde{P}$
- similarly, one can show that if $\tilde{x}$ and $\tilde{y}$ are non-adjacent in $\tilde{P}$, then $\pi^{-1}(x)$ and $\pi^{-1}(y)$ are non-adjacent in P $\square$

Hirsch Conjecture for 0/1-Polytopes

D. Naddef, The Hirsch conjecture is true for (0, 1)-polytopes, Math. Program., 1989

Theorem 13.6

i Let $P \subseteq \mathbb{R}^n$ be a 0/1-polytope. Then $\delta(P) \leq \dim P$.

ii Let $P \subseteq \mathbb{R}^n$ be an n-dimensional 0/1-polytope with m facets. Then $\delta(P) \leq m - n$.

Proof: i by induction on n, $n = 1$ is trivial

- let v, w be two arbitrary vertices of P
- if there is a facet F of the hypercube $[0, 1]^n$ with $v, w \in F$, then by induction

$$\delta_P(v, w) \leq \delta(P \cap F) \leq \dim(P \cap F) \leq \dim(P)$$

- otherwise v and w differ in all coordinates
- v has a neighbor $\bar{v}$ contained in a common facet F of the hypercube $[0, 1]^n$ with w, then

$$\delta_P(v, w) \leq 1 + \delta_P(\bar{v}, w) \leq 1 + \delta(P \cap F) \leq 1 + \dim(P \cap F) \leq \dim(P)$$

Proof of Hirsch Conjecture for 0/1-Polytopes

ii $\delta(P) \leq m - n$

- by Lemma 13.5, it suffices to show this result for full-dimensional poytopes
- induction on n, $n = 1$ is trivial
- if $m \geq 2n$, then the result follows from i
- since P is full-dimensional, all facets have dimension $n - 1$ by Theorem 3.29
- every vertex of P is contained in at least n facets
- if $m < 2n$, all pairs of vertices u, v share a common facet F
- F has at most $m - 1$ facets
- by induction

$$\delta_P(v, w) \leq \delta(F) \leq (m-1) - (n-1) = m - n \qquad \square$$

Average Case Behavior of the Simplex Method

- Despite the exponential lower bounds on the worst case behavior of the simplex method (Klee-Minty cubes etc.), the simplex method usually behaves well in practice.
- The number of iterations is “typically” $O(m)$.
- There have been several attempts to explain this phenomenon from a more theoretical point of view. These results say that “on average” the number of iterations is $O(\cdot)$ (usually polynomial).
- One main difficulty is to come up with a meaningful and, at the same time, manageable definition of the term “on average”.
- Smoothed analysis: hybrid of worst-case and average-case analyses, measuring the expected performance of algorithms under slight random perturbations of worst-case inputs.

D. A. Spielman and S.-H. Teng, Smoothed analysis of algorithms: Why the simplex algorithm usually takes polynomial time, JACM, 2004

13 Linear Programming Complexity

13.2 Ellipsoid Method

Complexity of Linear Programming

- no variant of the simplex method has been shown to have a polynomial running time
- complexity of Linear Programming remained unresolved for a long time
- in 1979, the Soviet mathematician Leonid Khachiyan proved that the so-called ellipsoid method earlier developed for nonlinear optimization can be modified in order to solve LPs in polynomial time
- in November 1979, the New York Times featured Khachiyan and his algorithm in a front-page story

An Approach to Difficult Problems

Mathematicians disagree as to the ultimate practical value of Leonid Khachiyan's new technique, but concur that in any case it is an important theoretical accomplishment.

Mr. Khachiyan's method is believed to offer an approach for the linear programming of computers to solve so-called "traveling salesman" problems. Such problems are among the most intractable in mathematics. They involve, for instance, finding the shortest route by which a salesman could visit a number of cities without his path touching the same city twice.

Each time a new city is added to the route, the problem becomes very much more complex. Very large numbers of variables must be calculated from large numbers of equations using a system of linear programming. At a certain point, the compexity becomes so great that a computer would require billions of years to find a solution.

In the past, "traveling salesmen" problems, including the efficient scheduling of airline crews or hospital nursing staffs, have been solved on computers using the "simplex method" invented by George B. Dantzig of Stanford University.

As a rule, the simplex method works well, but it offers no guarantee that after a certain number of computer steps it will always find an answer. Mr. Khachiyan's approach offers a way of telling right from the start whether or not a problem will be soluble in a given number of steps.

Two mathematicians conducting research at Stanford already have applied the Khachiyan method to develop a program for a pocket calculator, which has solved problems that would not have been possible with a pocket calculator using the simplex method.

Mathematically, the Khachiyan approach uses equations to create imaginary ellipsoids that encapsulate the answer, unlike the simplex method, in which the answer is represented by the intersections of the sides of polyhedrons. As the ellipsoids are made smaller and smaller, the answer is known with greater precision. MALCOLM W. BROWNE

New York Times, Nov. 27, 1979

Geometric Basics: Positive Definite Matrices

Definition 13.7 A symmetric matrix $D \in \mathbb{R}^{n \times n}$ is positive definite if

$$x^\top \cdot D \cdot x > 0 \qquad \text{for all } x \in \mathbb{R}^n \setminus \{0\}.$$

Lemma 13.8 For a symmetric matrix $D \in \mathbb{R}^{n \times n}$, the following statements are equivalent:

- i D is positive definite.
- ii D^{-1} exists and is positive definite.
- iii D has only real and positive eigenvalues.
- iv $D = B^\top \cdot B$ for a non-singular matrix $B \in \mathbb{R}^{n \times n}$.

Geometric Basics: Ellipsoids

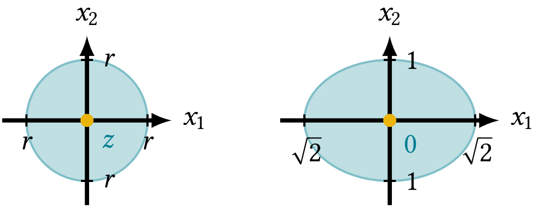
Definition 13.9 A set $E \subseteq \mathbb{R}^n$ of the form

$$E = E(z, D) := \{x \in \mathbb{R}^n \mid (x-z)^\top \cdot D^{-1} \cdot (x-z) \le 1\}$$

with $z \in \mathbb{R}^n$, $D \in \mathbb{R}^{n \times n}$ positive definite is called ellipsoid with center z.

Examples:

- $E(z, r^2 \cdot I) = \{x \in \mathbb{R}^n \mid (x-z)^\top \cdot (x-z) \le r^2\} = \{x \in \mathbb{R}^n \mid \|x-z\|_2 \le r\}$, is the ball of radius r at z, $\mathrm{Vol}(E(z, r^2 I)) = r^n \mathrm{Vol}(E(0, I)) = r^n \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}$
- $E\left(0, \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\right) = \left\{x \in \mathbb{R}^2 \mid \frac{1}{2}x_1^2 + x_2^2 \le 1\right\}$, $\mathrm{Vol}\left(E\left(0, \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\right)\right) = \sqrt{2}\pi$



- $\mathrm{Vol}(E(z, D)) = \sqrt{\det(D)}\mathrm{Vol}(E(0, I))$

Ellipsoid Method: Rough Idea

The ellipsoid method solves the following problem:

Given: $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$, polyhedron $P := \{x \in \mathbb{Q}^n \mid A \cdot x \geq b\}$.

Task: Find a point $x \in P$ or determine that P is empty.

Example:

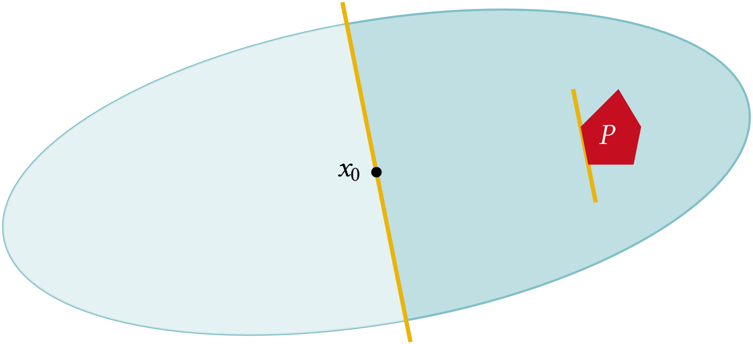
Ellipsoid Method: Rough Idea

The ellipsoid method solves the following problem:

Given: $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$, polyhedron $P := \{x \in \mathbb{Q}^n \mid A \cdot x \geq b\}$.

Task: Find a point $x \in P$ or determine that P is empty.

Example:



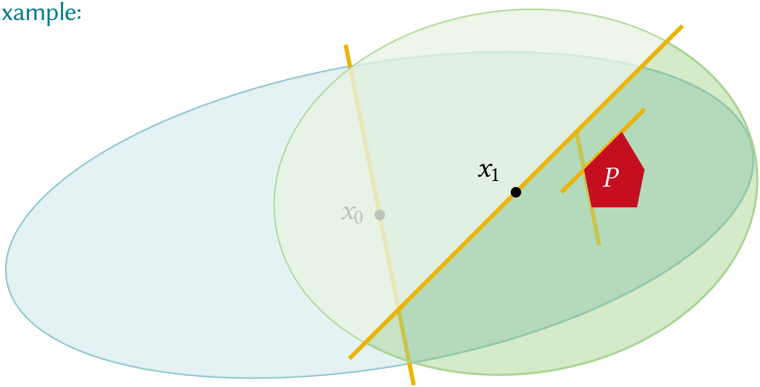
Ellipsoid Method: Rough Idea

The ellipsoid method solves the following problem:

Given: $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$, polyhedron $P := \{x \in \mathbb{Q}^n \mid A \cdot x \geq b\}$.

Task: Find a point $x \in P$ or determine that P is empty.

Example:



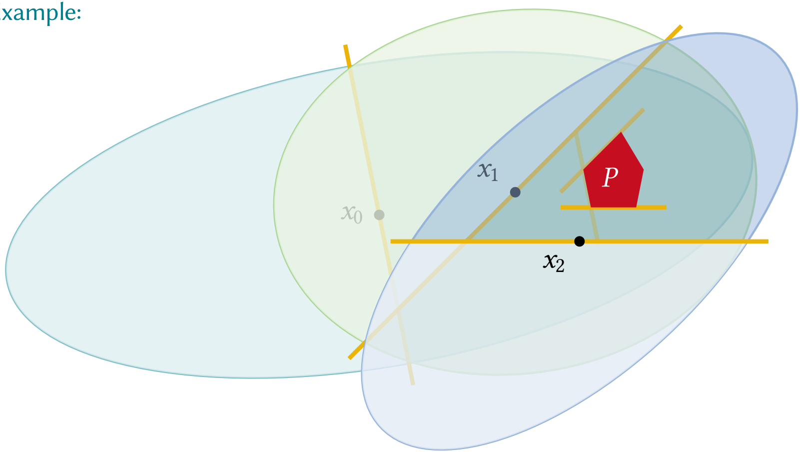
Ellipsoid Method: Rough Idea

The ellipsoid method solves the following problem:

Given: $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$, polyhedron $P := \{x \in \mathbb{Q}^n \mid A \cdot x \geq b\}$.

Task: Find a point $x \in P$ or determine that P is empty.

Example:



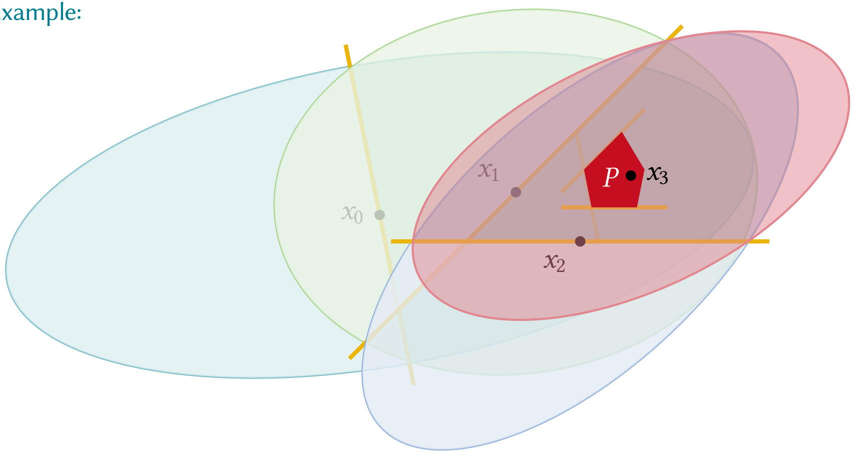
Ellipsoid Method: Rough Idea

The ellipsoid method solves the following problem:

Given: $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$, polyhedron $P := \{x \in \mathbb{Q}^n \mid A \cdot x \geq b\}$.

Task: Find a point $x \in P$ or determine that P is empty.

Example:



How to Find the Next Ellipsoid?

Theorem 13.10 Let $E = E(z, D)$ be an ellipsoid in $\mathbb{R}^n$ and $a \in \mathbb{R}^n \setminus \{0\}$. Consider the halfspace $H := \{x \in \mathbb{R}^n \mid a^\top \cdot x \geq a^\top \cdot z\}$ and set

$$\bar{z} := z + \frac{1}{n+1} \cdot \frac{D \cdot a}{\sqrt{a^\top \cdot D \cdot a}},$$

$$\bar{D} := \frac{n^2}{n^2 - 1} \cdot \left(D - \frac{2}{n+1} \cdot \frac{D \cdot a \cdot a^\top \cdot D}{a^\top \cdot D \cdot a} \right).$$

The matrix $\bar{D}$ is symmetric and positive definite. Thus, $\bar{E} := E(\bar{z}, \bar{D})$ is an ellipsoid. Moreover:

- i $E \cap H \subseteq \bar{E}$
- ii $\mathrm{Vol}(\bar{E}) < e^{-\frac{1}{2(n+1)}} \cdot \mathrm{Vol}(E)$

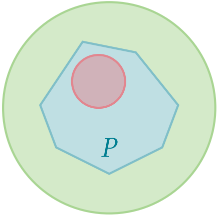
Proof: See Bertsimas & Tsitsiklis, Section 8.2. □

Simplifying Assumptions

Definition 13.11 A polyhedron $P \subseteq \mathbb{R}^n$ is full-dimensional if it has non-zero volume.

simplifying assumptions:

- polyhedron P is bounded (i.e., a polytope) and either empty or full-dimensional, i.e.,
  - $P \subseteq E(x_0, r^2 \cdot I) =: E_0$ with $r > 0$; $V := \mathrm{Vol}(E_0)$.
  - $\mathrm{Vol}(P) > v$ for some $v > 0$ (or P is empty).
  - assume that E_0, V, and v are known a priori.
- calculations (including square roots) can be made in infinite precision



We discuss these assumptions later in greater detail…

Ellipsoid Method

i set $t^* := \lceil 2(n+1)\log(V/v) \rceil$; $E_0 := E(x_0, r^2 \cdot I)$; $D_0 := r^2 \cdot I$; $t := 0$;

ii if $t = t^*$ then stop and output "P is empty";

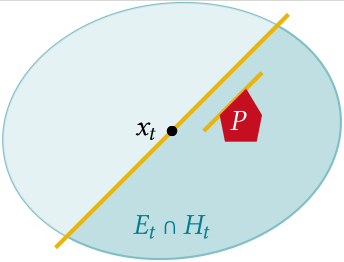
iii if $x_t \in P$ then stop and output x_t;

iv find violated constraint in $A \cdot x_t \geq b$, i.e., $a_i^\top \cdot x_t < b_i$ for some i;

v set $H_t := \{x \in \mathbb{R}^n \mid a_i^\top \cdot x \geq a_i^\top \cdot x_t\}$; (halfspace containing P)

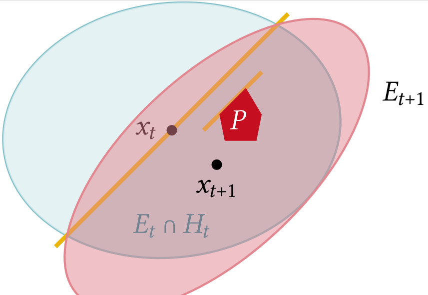
vi find ellipsoid $E_{t+1} \supseteq E_t \cap H_t$ by applying Theorem 13.10;

vii set $t := t + 1$ and go to step ii;



Ellipsoid Method

- i set $t^* := \lceil 2(n+1)\log(V/v) \rceil$; $E_0 := E(x_0, r^2 \cdot I)$; $D_0 := r^2 \cdot I$; $t := 0$;
- ii if $t = t^*$ then stop and output "P is empty";
- iii if $x_t \in P$ then stop and output x_t;
- iv find violated constraint in $A \cdot x_t \geq b$, i.e., $a_i^\top \cdot x_t < b_i$ for some i;
- v set $H_t := \{x \in \mathbb{R}^n \mid a_i^\top \cdot x \geq a_i^\top \cdot x_t\}$; (halfspace containing P)
- vi find ellipsoid $E_{t+1} \supseteq E_t \cap H_t$ by applying Theorem 13.10;
- vii set $t := t + 1$ and go to step ii;



Theorem 13.12 The ellipsoid method returns a point in P or decides correctly that $P = \emptyset$.

Proof: If $x_t \in P$ for some $t < t^*$, then the algorithm returns x_t.

Otherwise: By induction $P \subseteq E_k$ for $k = 0, 1, \dots, t^*$.

By Theorem 13.10 we get $\dfrac{\text{Vol}(E_{t+1})}{\text{Vol}(E_t)} < e^{-\frac{1}{2(n+1)}}$ for all t.

Thus, $\dfrac{\text{Vol}(E_{t^*})}{\text{Vol}(E_0)} < e^{-\frac{t^*}{2(n+1)}}$.

$$\Longrightarrow \quad \text{Vol}(E_{t^*}) < V \cdot e^{-\frac{\lceil 2(n+1)\log(V/\nu)\rceil}{2(n+1)}} \le V \cdot e^{-\log(V/\nu)} = \nu \,.$$

Since ν is a lower bound on the volume of non-empty P, the algorithm correctly decides that $P = \emptyset$. □

What if P is Unbounded?

Lemma 13.13 Let $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{R}^m$, and let U be the largest absolute value of entries in A and b. Every extreme point of polyhedron $P = \{x \in \mathbb{R}^n \mid A \cdot x \geq b\}$ satisfies

$$-(nU)^n \leq x_j \leq (nU)^n \qquad \text{for } j = 1, \dots, n.$$

Proof: let x be an extreme point

- there are n linearly independent rows of $Ax \geq b$ that are active
- there is a submatrix $\bar{A} \in \mathbb{Z}^{n \times n}$, $\bar{b} \in \mathbb{R}^n$ such that $\bar{A}x = \bar{b}$
- $x = \bar{A}^{-1}\bar{b}$, by Cramer's rule $x_j = \frac{\det \bar{A}^j}{\det \bar{A}}$

$$|x_j| \leq \left|\det \bar{A}^j\right| = \left|\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} \bar{a}^j_{i,\sigma(s)}\right|$$

$$\leq \sum_{\sigma \in S_n} \prod_{i=1}^{n} \left|\bar{a}^j_{i,\sigma(i)}\right| \leq (n!)U^n \leq (nU)^n \qquad \square$$

Dealing with Unbounded P

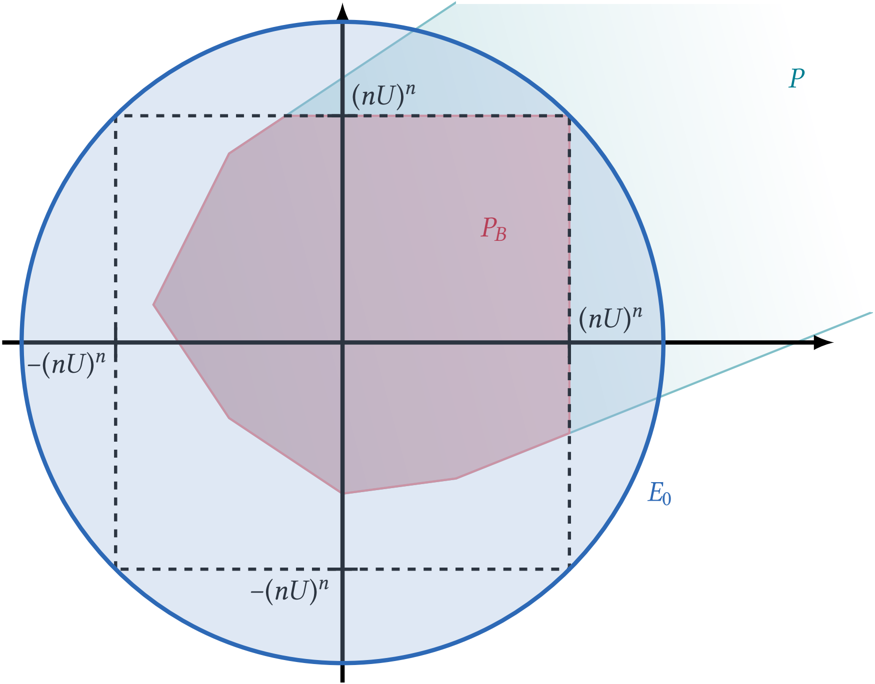
- let $P_B := \{x \in P \mid -(nU)^n \leq x_j \leq (nU)^n \text{ for all } j\}$
- under the assumption that $\text{rank}(A) = n$, we get

$$P \neq \emptyset \quad \Longleftrightarrow \quad P \text{ contains an extreme point} \quad \Longleftrightarrow \quad P_B \neq \emptyset$$

- thus, it suffices to look for a point in $P_B \subseteq P$ instead of P
- start with ellipsoid $E_0 := E(0, n(nU)^{2n} \cdot I)$, then

$$\begin{aligned} E(0, n(nU)^{2n} I) &= \left\{ x \in \mathbb{R}^n : \sum_{j=1}^{n} x_i^2 \leq n(nU)^{2n} \right\} \\ &\supseteq \left\{ x \in P : |x_j| \leq (nU)^n \text{ for all } j \right\} = P_B \end{aligned}$$

- $$\begin{aligned} \text{Vol}(E_0) &\leq (n(nU)^{2n})^{n/2} \text{Vol}(E(0, I)) \\ &= n^{n/2} (nU)^{n^2} \text{Vol}(E(0, I)) \\ &\leq n^{n/2} (nU)^{n^2} 2^n =: V \end{aligned}$$



P
$(nU)^n$
P_B
$(nU)^n$
$-(nU)^n$
E_0
$-(nU)^n$

What if P is not Full-Dimensional?

Lemma 13.14 Let $A \in \mathbb{Z}^{m \times n}$, $b \in \mathbb{Z}^m$, and let U be the largest absolute value of entries in A and b. Consider polyhedron $P = \{x \in \mathbb{R}^n \mid A \cdot x \geq b\}$, define

$$\varepsilon := \frac{\big((n+1)U\big)^{-(n+1)}}{2(n+1)}$$

and a new polyhedron

$$P_\varepsilon := \{x \in \mathbb{R}^n \mid A \cdot x \geq b - \varepsilon \cdot \mathbf{1}\}.$$

Then it holds that

- a $P = \emptyset \implies P_\varepsilon = \emptyset$
- b $P \neq \emptyset \implies P_\varepsilon$ is full-dimensional
- c Given a point in P_ε, a point in P can be obtained in polynomial time.

b follows since $x \in P$ implies $x + \delta e_i \in P_\varepsilon$ for all i, $\delta > 0$ small enough

c can be shown by rounding techniques (omitted here)

Proof of Lemma 13.14

a $P = \emptyset \implies P_\varepsilon = \emptyset$

$$\text{Primal: } \min \; 0^\top x \quad \text{s.t. } Ax \geq b \qquad \leftrightarrow \qquad \text{Dual: } \max \; p^\top b \quad \text{s.t. } p^\top A = 0^\top, \; p \geq 0$$

dual contains $p = 0$

infeasible — dual unbounded

- there is $p \geq 0$ with $p^\top A = 0^\top$ and $p^\top b = 1$
- by Lemma 13.13, there is a basic feasible solution $\hat{p}$ to $p^\top A = 0^\top$, $p^\top b = 1$, $p \geq 0$ such that $|\hat{p}_i| \leq ((n+1)U)^{n+1}$ for all i and at most $n+1$ components of $\hat{p}$ are non-zero
- $\hat{p}^\top(b - \varepsilon \mathbf{1}) = \hat{p}^\top b - \varepsilon \sum_{i=1}^m \hat{p}_i \geq 1 - \varepsilon(n+1)((n+1)U)^{n+1}$
  $= 1 - \frac{1}{2} = \frac{1}{2} > 0$
- $\max\{p^\top(b - \varepsilon\mathbf{1}) \text{ s.t. } p^\top A = 0^\top, p \geq 0\}$ unbounded
- $\min 0^\top x$ s.t. $Ax \geq b - \varepsilon\mathbf{1}$ infeasible □

Bounding the Minimum Volume

Lemma 13.15 If the polyhedron $P = \{x \in \mathbb{R}^n \mid A \cdot x \geq b\}$ is full-dimensional and bounded with U as above, then

$$\mathrm{Vol}(P) > n^{-n}(nU)^{-n(n+1)} .$$

Proof idea:

- P has $n+1$ extreme points $v^0, \ldots, v^n$
- so $\mathrm{Vol}(P) \geq \mathrm{Vol}\big(\mathrm{conv}(v^0, \ldots, v^n)\big) = \frac{1}{n!}\left|\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ v_0 & v_1 & \cdots & v_n \end{pmatrix}\right|$
- By Cramer's rule, $(v_i)_j = \frac{\det B_{ij}}{\det B_i}$, where B_i is a submatrix of A, and B_{ij} is obtained by replacing the jth column of B_i by elements of b.
- $\det\begin{pmatrix} 1 & 1 & \cdots & 1 \\ v_0 & v_1 & \cdots & v_n \end{pmatrix} = \frac{1}{\prod_{i=0}^{n} |\det B_i|}\left|\det\begin{pmatrix} \det B_0 & \det B_1 & \cdots & \det B_n \\ \det B_0 v_0 & \det B_1 v_1 & \cdots & \det B_n v_n \end{pmatrix}\right|$
- The big integer determinant above is ≥ 1, and $|\det B_i| \leq (nU)^n$. So we get

$$\mathrm{Vol}(P) \geq \frac{1}{n!}(nU)^{-n(n+1)} > n^{-n}(nU)^{-n(n+1)}$$

Bounding the number of iterations

Theorem 13.16

The number of iterations of the ellipsoid method can be bounded by $O(n^5 \log(nU))$. □

Proof sketch:

- ellipsoid method requires $t^* = 2(n+1)\log(V/v)$ iterations
- first form bounded polyhedron P_B and then perturb it to $P_{B,\varepsilon}$
- for P_B largest entry is $U_B = (nU)^n$ and largest denominator is $1/\varepsilon_B = 2(n+1)((n+1)U_B)^{(n+1)}$
- re-normalizing to integers yields $\tilde{U} = 2(nU)^n(n+1)[(n+1)(nU)^n]^{n+1} \approx (nU)^{n^2}$
- $V = n^{n/2}2^n(nU)^{n^2} \approx (nU)^{n^2}$
- by Lemma 13.15, $v = n^{-n}(n\tilde{U})^{-n(n+1)} \approx (nU)^{-n^4}$
- $\log(\frac{V}{v}) \le \log\left(\frac{(2n)^n(nU)^{n^2}}{n^{-n}(n\tilde{U})^{-n(n+1)}}\right) = O(n^4\log(nU))$ □

Required Numeric Precision

Major problems:

- Bound number of arithmetic operations per iteration.
- How to take square roots?
- Only finite precision possible!

Theorem 13.17 Using only $O(n^3 \log U)$ binary digits of precision, the ellipsoid method still correctly decides whether P is empty in $O(n^6 \log(nU))$ iterations. Thus, the Linear Inequalities problem can be solved in polynomial time.

Solving LPs in Polynomial Time

Consider a pair of primal and dual LPs:

$$\begin{aligned} \min \quad & c^\top \cdot x \\ \text{s.t.} \quad & A \cdot x \geq b \\ & x \geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max \quad & p^\top \cdot b \\ \text{s.t.} \quad & p^\top \cdot A \leq c^\top \\ & p \geq 0 \end{aligned}$$

Solve the primal and dual LP by finding a point (x, p) in the polyhedron given by

$$\left\{ (x, p) \mid c^\top \cdot x = p^\top \cdot b,\ A \cdot x \geq b,\ p^\top \cdot A \leq c^\top,\ x, p \geq 0 \right\} .$$

Theorem 13.18 Linear programs can be solved in polynomial time. □

13 Linear Programming Complexity

13.3 Optimization vs. Separation

LPs with Exponentially Many Constraints

- The number of iterations of the ellipsoid method only depends on the dimension n and U, but not on the number of constraints m.
- Thus there is hope to solve LPs with exponentially many constraints (that are implicitly given) in polynomial time.

Example: Consider the following LP relaxation of the TSP (subtour LP):

$$\begin{aligned} \min \quad & \sum_{e \in E} c_e \cdot x_e \\ \text{s.t.} \quad & \sum_{e \in \delta(v)} x_e = 2 && \text{for all nodes } v \in V, \\ & \sum_{e \in \delta(X)} x_e \geq 2 && \text{for all subsets } \emptyset \neq X \subsetneq V, \qquad (13.1) \\ & 0 \leq x_e \leq 1 && \text{for all edges } e. \end{aligned}$$

Notice that there are $2^{n-1} - 1$ subtour elimination constraints (13.1).

LPs with Exponentially Many Constraints

- Describe polyhedron $P = \{x \in \mathbb{R}^n \mid A \cdot x \geq b\}$ by specifying n and an integer vector h of primary data of dimension $O(n^k)$ with k constant.
  Let $U_0 := \max_i |h_i|$.

- There is a mapping which, given n and h, defines $A \in \mathbb{Z}^{m \times n}$ and $b \in \mathbb{Z}^m$. Let

$$U := \max\{|a_{ij}|, |b_i| \mid i = 1, \dots, m,\ j = 1, \dots, n\} .$$

- We assume that there are constants C and ℓ such that

$$\log U \leq C \cdot n^\ell \cdot \log^\ell U_0 ,$$

  that is, U can be encoded polynomially in the input size.

- The number of iterations of the ellipsoid method is

$$O(n^6 \log(nU)) = O(n^6 \log n + n^{6+\ell} \log^\ell U_0)$$

  and thus polynomial in the input size of the primary problem data.

Separation Problem

In every iteration of the ellipsoid method, we have to solve the following problem:

Definition 13.19 Given a polyhedron $P \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$, the separation problem is to

- i either decide that $x \in P$, or
- ii find $d \in \mathbb{R}^n$ with $d^\top \cdot x < d^\top \cdot y$ for all $y \in P$.

Example: The subtour elimination constraints (13.1) can be separated in polynomial time by finding a minimum capacity cut.

Optimization is as Difficult as Separation

The following Theorem by Grötschel, Lovász, and Schrijver is a consequence of the ellipsoid method.

Theorem 13.20

i Given a family of polyhedra, if we can solve the separation problem in time polynomial in n and $\log U$, then we can also solve LPs over those polyhedra in time polynomial in n and $\log U$.

ii The converse is also true under some technical conditions. □

Example: The subtour LP for the TSP can be solved in polynomial time since the subtour elimination constraints (13.1) can be separated efficiently.

Path-Based LP Formulation for Flow Problems

Let $\mathcal{P}$ be the set of all s-t-dipaths in digraph D.

$$\begin{aligned} \max \quad & \sum_{P \in \mathcal{P}} y_P & \\ \text{s.t.} \quad & \sum_{P \in \mathcal{P} \,:\, a \in P} y_P \le u(a) & \text{for all } a \in A \\ & y_P \ge 0 & \text{for all } P \in \mathcal{P} \end{aligned}$$

Dual LP:

$$\begin{aligned} \min \quad & \sum_{a \in A} u(a) \cdot z_a & \\ \text{s.t.} \quad & \sum_{a \in P} z_a \ge 1 & \text{for all } P \in \mathcal{P} \quad (13.2) \\ & z_a \ge 0 & \text{for all } a \in A \end{aligned}$$

Since constraints (13.2) can be separated efficiently by a shortest path computation, the dual LP can be solved efficiently. Using complementary slackness conditions, also the primal LP can be solved in polynomial time.

13 Linear Programming Complexity

13.4 Interior Point Method

Primal-Dual Interior-Point Methods, Wright, Chapter 5

General idea: The central path

- Consider the LP

$$\begin{aligned} \min \quad & c^\top x \\ s.t. \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

- To get rid off the ≥ 0-constraints, we approximate the problem by using a log-barrier:

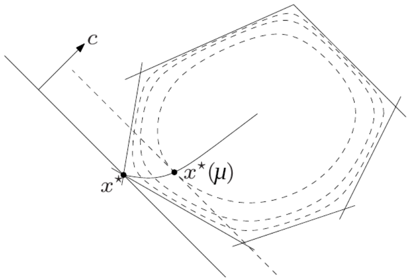
$$\begin{aligned} \min \quad & c^\top x - \mu \sum_i \log(x_i) \\ s.t. \quad & Ax = b \end{aligned} \qquad (P_\mu)$$

- We define the central path

$$\{x^*(\mu) \, : \, \mu > 0\},$$

  where x^* solves (P_μ).

- Interior point methods "follow" the path and let $\mu \longrightarrow 0$.



Primal-dual central path

- We can form the log-penalized version of the dual problem, too

$$\begin{aligned} \min \quad & c^\top x \\ \text{s.t.} \quad & Ax = b \qquad (P) \\ & x \geq 0 \end{aligned} \qquad\qquad \begin{aligned} \max \quad & b^\top y \\ \text{s.t.} \quad & A^\top y + s = c \qquad (D) \\ & s \geq 0 \end{aligned}$$

$$\begin{aligned} \min \quad & c^\top x - \mu \sum_i \log(x_i) \qquad (P_\mu) \\ \text{s.t.} \quad & Ax = b \end{aligned} \qquad\qquad \begin{aligned} \max \quad & b^\top y + \mu \sum_i \log(s_i) \qquad (D_\mu) \\ \text{s.t.} \quad & A^\top y + s = c \end{aligned}$$

Theorem 13.21 Consider the pair of programs (P_μ) and (D_μ) for some $\mu > 0$.

- (P_μ) and (D_μ) are convex problems, and are dual from each other
- If (P) and (D) are *strictly feasible*, then the perturbed problems have a unique solution $x^*(\mu), y^*(\mu), s^*(\mu)$. These solutions solve the KKT system

$$\begin{cases} Ax = b\,; & x > 0, \;\; s > 0\,; \\ A^\top y + s = c\,; & x_i \cdot s_i = \mu, \; \forall i \in \{1, \dots, n\} \end{cases}$$

If (P) and (D) are *strictly feasible*, then the perturbed problems have a unique solution $x^*(\mu), y^*(\mu), s^*(\mu)$. These solutions solve the KKT system

$$\begin{cases} Ax = b\,; & x > 0, \quad s > 0\,; \\ A^\top y + s = c\,; & x_i \cdot s_i = \mu, \ \forall i \in \{1, \ldots, n\} \end{cases}$$

- primal simplex maintains primal feasibility & complementary slackness
  ⟶ progress to achieve dual feasibility
- dual simplex maintains dual feasibility & complementary slackness
  ⟶ progress to achieve primal feasibility
- In contrast, Interior Point methods maintain primal and dual feasibility
  ⟶ work towards achieving complementary slackness.

Theorem 13.22 Let $\mu = \frac{\varepsilon}{n}$ for some $\varepsilon > 0$. The optimal solutions $x^*(\mu), y^*(\mu), s^*(\mu)$ of (P_μ)-(D_μ) are ε-suboptimal for the original pair of problems (P)-(D).

Proof: Let x, y, s solve the KKT system.

- Primal and dual feasibility is clear
- The duality gap is $c^\top x - b^\top y = y^\top Ax + s^\top x - y^\top Ax = \sum_i s_i x_i = n\mu = \varepsilon$

□

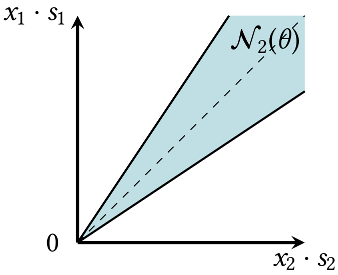
Path Neighborhood

- In practice, it is impossible to remain exactly on the path.
- Let $x \odot s := (x_1 \cdot s_1, \dots, x_n \cdot s_n)^\top$ and $e := (1, \dots, 1)^\top$.
- If (x, y, s) were on the central path, $x \odot s = \mu e \implies \mu = \frac{x^\top s}{n}$
- For $\theta \in [0, 1)$, we define the neighborhood

$$\mathcal{N}_2(\theta) = \left\{ (x, y, s) \,:\, Ax = b, A^\top y + s = c, x, s > 0, \left\| (x \odot s) - \left(\frac{x^\top s}{n} \right) e \right\| \le \theta \left(\frac{x^\top s}{n} \right) \right\}$$

A primal-dual feasible solution (x, y, s) lies

- on the central path if $x \odot s$ is on the dashed ray
- in $\mathcal{N}_2(\theta)$ if $x \odot s$ lies in some cone



- Simplifying assumption: We are given an initial solution $(x_0, y_0, s_0) \in \mathcal{N}_2(\theta)$.

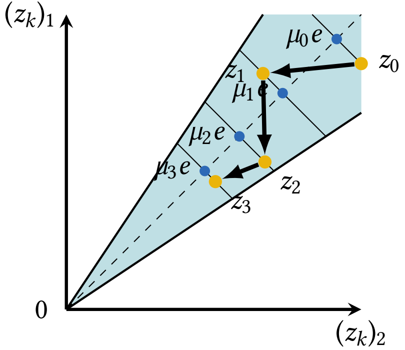
General Idea

For some fixed parameters $\theta, \sigma \in (0, 1)$:

- Given $(x_k, y_k, s_k) \in \mathcal{N}_2(\theta)$, with $\mu_k = \frac{x_k^\top s_k}{n}$.
- If $\mu_k \leq \varepsilon/n$: Stop
- Compute a direction $(\Delta x, \Delta y, \Delta s)$, and let

$$(x_{k+1}, y_{k+1}, s_{k+1}) := (x_k, y_k, s_k) + (\Delta x, \Delta y, \Delta s),$$

such that $(x_{k+1}, y_{k+1}, s_{k+1}) \in \mathcal{N}_2(\theta)$ and $\mu_{k+1} := \frac{\mu_{k+1}^\top x_{k+1}}{n} = \sigma \mu_k$



Path followed by
$z_k := (x_k \odot s_k)$

Computing the Newton direction

- Given $(x, y, s) \in \mathcal{N}_2(\theta)$ with $\frac{x^\top s}{n} = \mu$
- We want $(x + \Delta x, y + \Delta y, s + \Delta s)$ feasible and $(x + \Delta x) \odot (s + \Delta s) = \sigma\mu e$:

$$\left\{ \begin{aligned} A(x + \Delta x) &= b \\ A^\top(y + \Delta y) + s + \Delta s &= c \\ (x + \Delta x) \odot (s + \Delta s) &= \sigma\mu e \end{aligned} \right. \iff \left\{ \begin{aligned} A(\Delta x) &= 0 \\ A^\top(\Delta y) + (\Delta s) &= 0 \\ x \odot s + x \odot (\Delta s) + s \odot (\Delta x) & \\ +(\Delta x) \odot (\Delta s) &= \sigma\mu e \end{aligned} \right.$$

- Newton's method consists in linearizing the above equation by neglecting the quadratic term $(\Delta x) \odot (\Delta s)$.
- Setting $X := \mathrm{Diag}(x)$ and $S := \mathrm{Diag}(s)$, we obtain $(\Delta x, \Delta y, \Delta s)$ by solving:

$$\begin{pmatrix} A & 0 & 0 \\ 0 & A^\top & I \\ S & 0 & X \end{pmatrix} \cdot \begin{pmatrix} \Delta x \\ \Delta y \\ \Delta s \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \sigma\mu e - x \odot s \end{pmatrix}.$$

Correctness and Number of iterations

- Let $(x, y, s) \in \mathcal{N}_2(\theta)$ with $\mu = \frac{x^\top s}{n}$. We denote

$$(x^+, y^+, s^+) = (x + \Delta x, y + \Delta y, s + \Delta s), \quad \mu^+ := \frac{(x^+)^\top (s^+)}{n}.$$

Lemma 13.23 $\mu^+ = \sigma\mu$

Proof:

- From $A^\top(\Delta y) + (\Delta s) = 0$ and $A(\Delta x) = 0$, we have $\Delta s^\top \Delta x = -(\Delta y)^\top A(\Delta x) = 0$.
- $x \odot (\Delta s) + s \odot (\Delta x) = \sigma\mu e - x \odot s \implies x^\top(\Delta s) + s^\top(\Delta x) = \sigma\mu n - x^\top s = n\mu(\sigma - 1)$
- $\mu^+ = \frac{(x+\Delta x)^\top(s+\Delta s)}{n} = \frac{1}{n}(x^\top s + n\mu(\sigma - 1)) = \sigma\mu.$

Theorem 13.24 Let $\theta = 0.4$. If $(x, y, s) \in \mathcal{N}_2(\theta)$ and $\sigma = 1 - \frac{\theta}{\sqrt{n}}$, then $(x^+, y^+, s^+) \in \mathcal{N}_2(\theta)$. As a result, the algorithm finds an ε-suboptimal solution after $O(\sqrt{n}\log(\frac{n\mu_0}{\varepsilon}))$ iterations.

Proof: If the iterates remain in $\mathcal{N}_2(\theta)$, $\mu_k = (1 - \frac{\theta}{\sqrt{n}})^k \mu_0 \implies \mu_K \leq \frac{\varepsilon}{n}$ for

$$K = \left\lceil \frac{\log(\epsilon/(n\mu_0))}{\log(1 - \theta/\sqrt{n})} \right\rceil = O\left(\sqrt{n}\log\left(\frac{n\mu_0}{\varepsilon}\right)\right).$$

Iterates stay in the neighborhood (1/3)

Lemma 13.25 Let $u, v \in \mathbb{R}^n$ with $u \geq 0, v \geq 0$. Then

$$\|u \odot v\| \leq 2^{-\frac{3}{2}} \|u + v\|^2$$

Proof: See Wright, Chapter 5.

Lemma 13.26 If $\|z - \mu e\| \leq \theta\mu$, then $\min_i z_i \geq (1 - \theta)\mu$.

Proof: Let $i \in \{1, \dots, n\}$.

- $|z_i - \mu| \leq \|z - \mu e\|_\infty \leq \|z - \mu e\|_2 \leq \theta\mu$.
- $z_i - \mu \geq -\theta\mu \implies z_i \geq \mu(1 - \theta)$. □

Lemma 13.27 $\|x^+ \odot s^+ - \mu^+ e\| \leq \frac{\theta^2 + n(1-\sigma)^2}{2^{\frac{3}{2}}(1-\theta)}\mu$

Proof:

- Let $u, v \in \mathbb{R}^n$ with $u_i = \sqrt{\frac{x_i}{s_i}}(\Delta x)_i$ and $v_i = \sqrt{\frac{s_i}{x_i}}(\Delta s)_i$.
- From $x_i(\Delta s)_i + s_i(\Delta x)_i = \sigma\mu - x_i s_i$, we have $(u + v)_i = \frac{1}{\sqrt{x_i s_i}}(\sigma\mu - x_i s_i)$
- $\|\Delta x \odot \Delta s\| = \|u \odot v\| \leq 2^{-\frac{3}{2}} \|u + v\|^2 \leq 2^{-\frac{3}{2}} \frac{\|\sigma\mu e - x \odot s\|^2}{\min_i x_i s_i}$

Iterates stay in the neighborhood (2/3)

- We have shown $\|\Delta x \odot \Delta s\| \leq 2^{-\frac{3}{2}} \frac{\|\sigma\mu e - x \odot s\|^2}{\min_i x_i s_i}$
- By Lemma 13.26 applied to $z = x \odot s$, the denominator is

$$\min_i z_i \geq \mu(1-\theta)$$

- To bound the numerator, we write

$$\begin{aligned} \|\sigma\mu e - x \odot s\|^2 &= \|(\mu e - x \odot s) - (1-\sigma)\mu e\|^2 \\ &= \underbrace{\|\mu e - x \odot s\|^2}_{\leq \theta^2\mu^2} + (1-\sigma)^2\mu^2 n - 2(1-\sigma)\mu \underbrace{e^\top(\mu e - x \odot s)}_{=\mu n - x^\top s = 0} \\ &\leq \mu^2(\theta^2 + (1-\sigma)^2 n) \end{aligned}$$

- Putting all together,

$$\|\Delta x \odot \Delta s\| \leq \frac{\mu(\theta^2 + (1-\sigma)^2 n)}{2^{\frac{3}{2}}(1-\theta)}$$

- Finally, we have $x \odot s + x \odot \Delta s + s \odot \Delta x - \sigma\mu e = 0$, which implies

$$x^+ \odot s^+ - \mu^+ e = (\Delta x) \odot (\Delta s).$$

Iterates stay in the neighborhood (3/3)

Theorem 13.28 Let the parameters $\theta \in (0,1)$ and $\sigma \in (0,1)$ be such that

$$\frac{\theta^2 + n(1-\sigma)^2}{2^{3/2}(1-\theta)} \leq \sigma\theta. \qquad (\star)$$

(in particular, this works for all $n \geq 1$ with $\theta = 0.4$ and $\sigma = 1 - \frac{\theta}{\sqrt{n}}$). Then

$$(x, y, s) \in \mathcal{N}_2(\theta) \implies (x^+, y^+, s^+) \in \mathcal{N}_2(\theta).$$

Proof: (Sketch)

- $Ax^+ = b$ and $A^\top y^+ + s^+ = c$ are clear from the definition of $(\Delta x, \Delta y, \Delta s)$.
- $\|x^+ \odot s^+ - \mu^+ e\| \leq \theta\mu^+$ follows from Lemma (13.27) and $(\star)$.
- So we only have to show that $x^+ > 0$ and $s^+ > 0$.
- Let $i \in \{1, \dots, n\}$. By Lemma (13.26),

$$x_i^+ \cdot s_i^+ > (1-\theta)\mu^+ = (1-\theta)\sigma\mu > 0$$

- More generally, one can show that $(x + \alpha\Delta x)_i(s_i + \alpha\Delta s)_i > 0$ holds for all $\alpha \in [0,1]$. So there is no $\alpha \in [0,1]$ such that $(x + \alpha\Delta x)_i = 0$ or $(s + \alpha\Delta s)_i = 0$. Thus, $x_i^+ > 0$ and $s_i^+ > 0$. □

Final Remarks

- The first polynomial-time interior point method for linear programming was discovered by Karmarkar in 1984.
- The main work in each iteration is to solve a linear system of size $O(n+m)$.
- To obtain a polynomial time algorithm, we would need to show that
  - a an ϵ−subobtimal solution can be rounded to an exact solution in $O(n^3)$ if $\epsilon \leq 2^{-2L}$, for an instance with integer data of bitlength L;
  - b we can obtain in polynomial time an initial point with $\mu_0 \leq 2^{\beta L}$ for some constant β.
- a can be shown, similar as for the ellipsoid method; By using (variants) of the studied path-following method, we can achieve b. This yields a total iteration complexity of $O(\sqrt{n}L)$.
- The parameters θ and σ used in this presentation are selected to make the method run in polynomial time; In practice, one can use other parameters to achieve faster convergence.
- State-of-the art (academic and commercial) solvers use a mix of the simplex algorithm and some interior point method.