CSE 421

Linear Programs

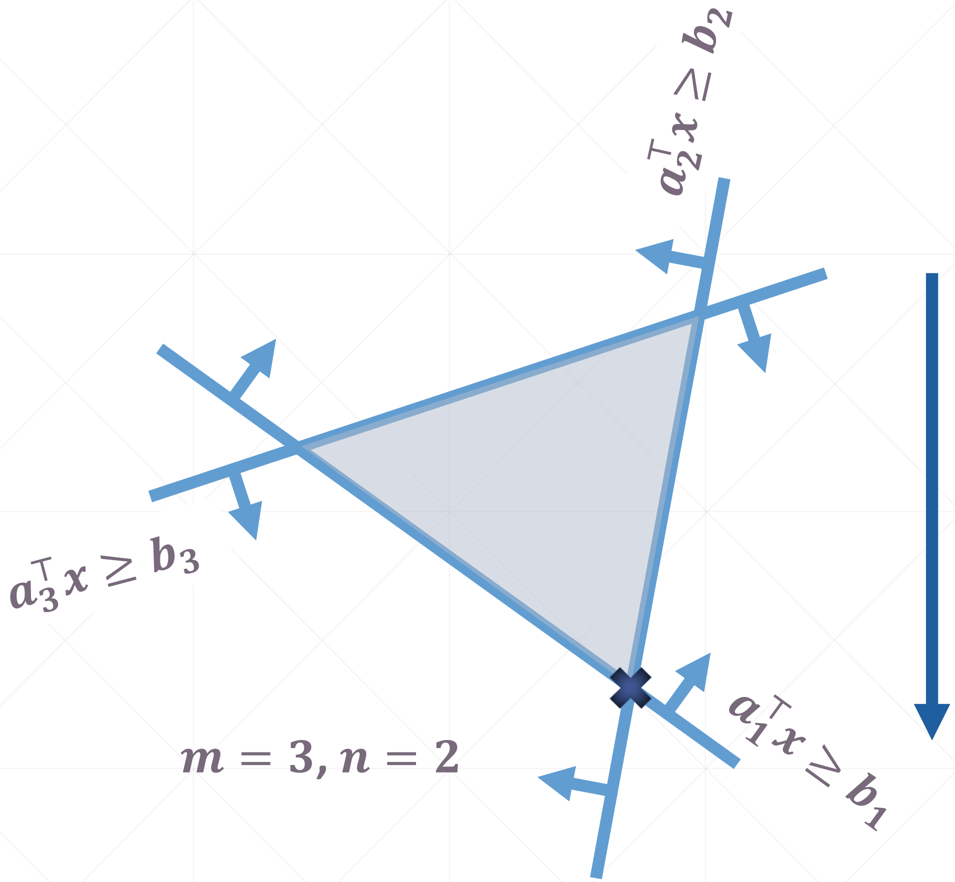
Yin Tat Lee

Linear Program

Consider the linear program (LP)

$$\min_{Ax \geq b} c^\top x$$

where A is a $m \times n$ matrix.



- m = the number of constraints
- n = the number of variables

- Example: Flight crew scheduling problem by American Airlines

$$m = 12{,}750{,}000, n = 837$$

Simplex Method [Dantzig 47]

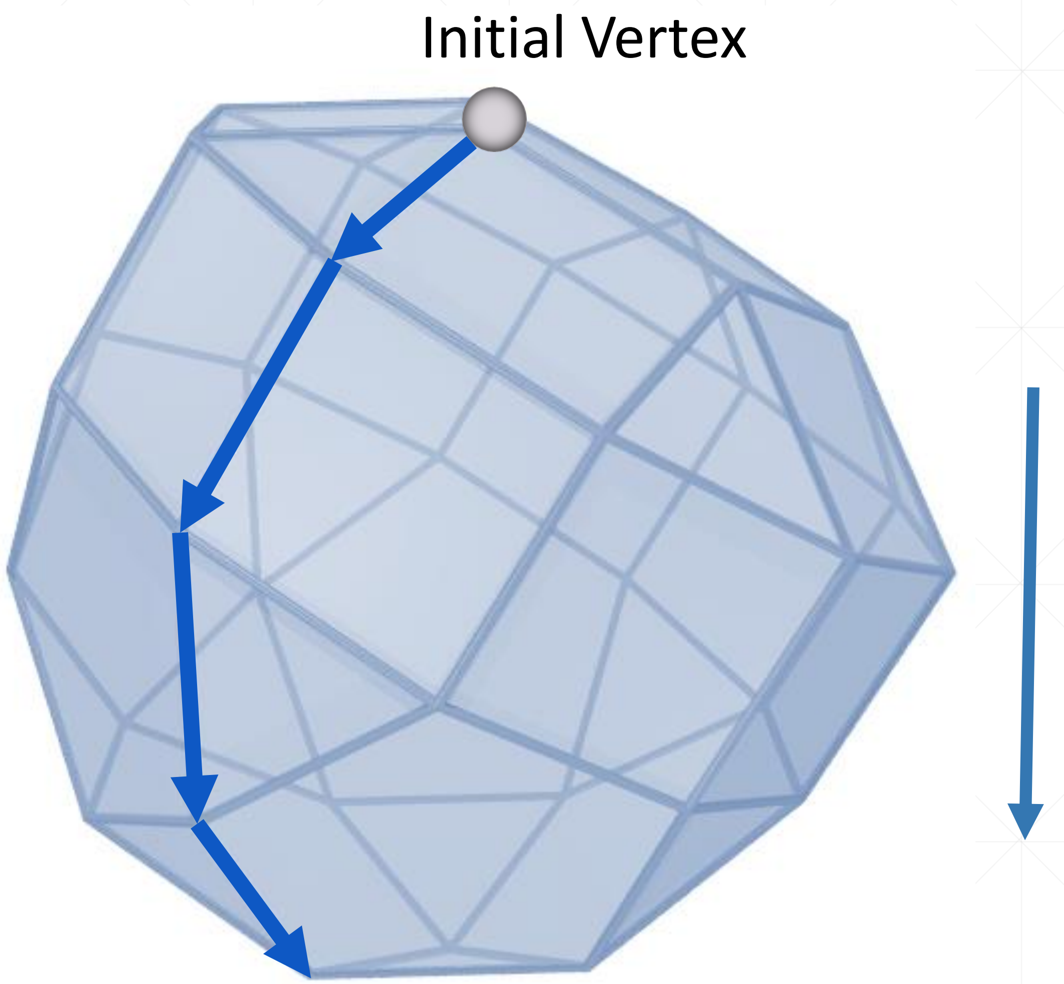
First generation of LP solver.

Efficient in practice

Exponential time in worst case

When applied to MaxFlow,

it is exactly augmenting path.



A Soviet Discovery Rocks World of Mathematics

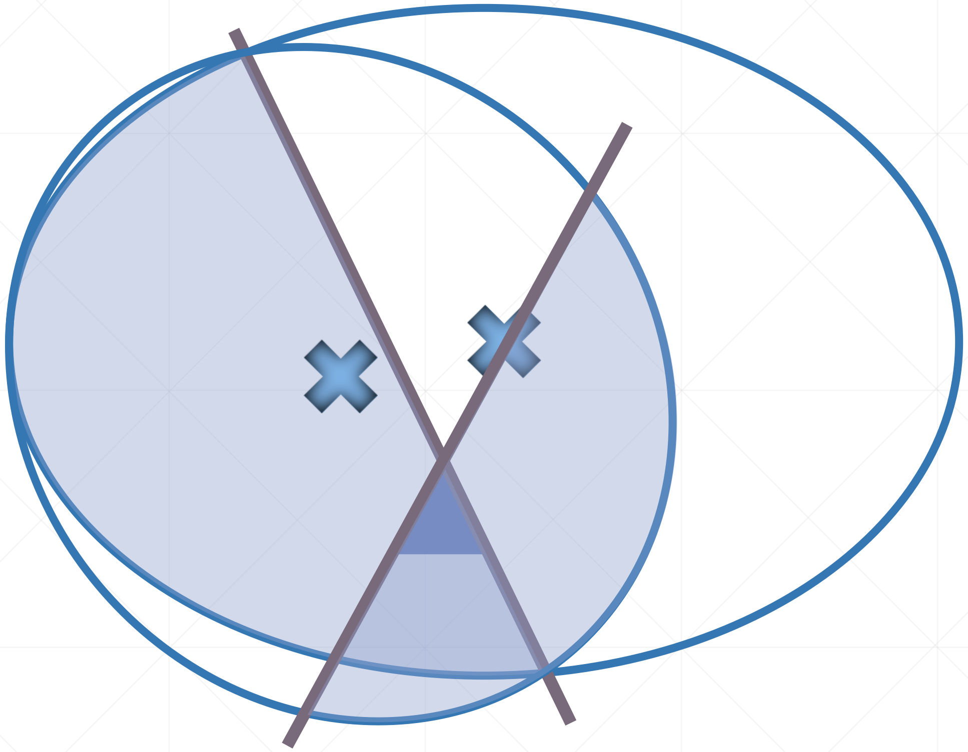
Source: New York Times

Ellipsoid Method [Shor; Nemirovski-Yudin; Khachiyan 70s]

First polynomial time algorithm for LP.

Very important in theory.

Slow in practice.



$$K = \{Ax \geq b, c^\top x \leq OPT\}$$

Interior Point Methods [Karmarkar 1984]

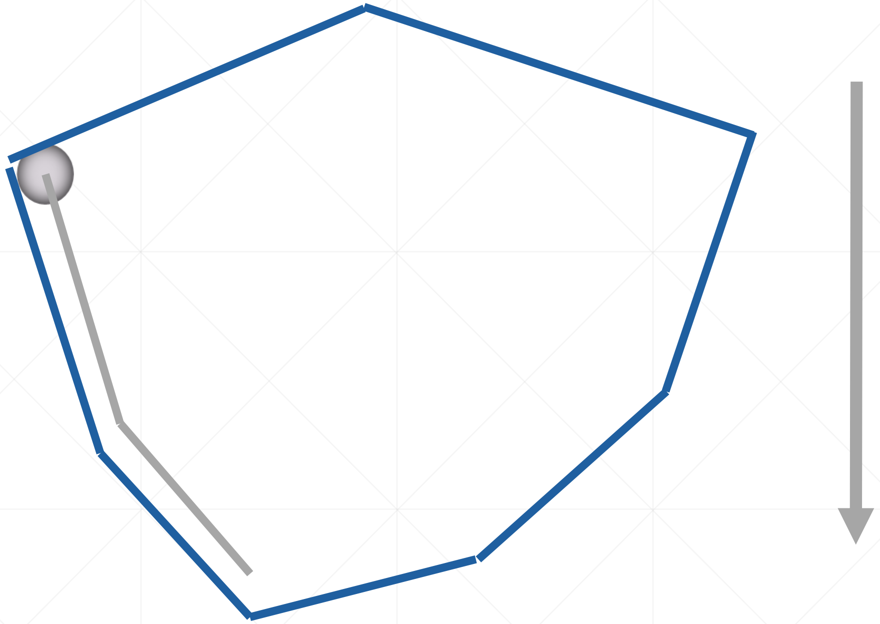
Karmarkar at Bell Labs: an equation to find a new way through the maze

Folding the Perfect Corner

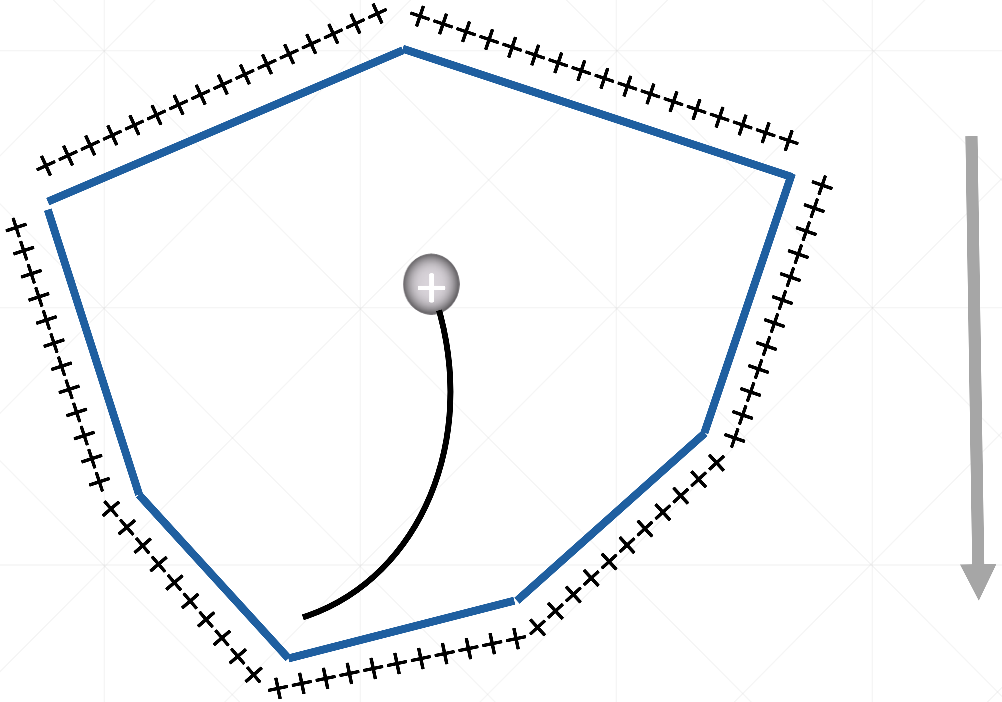
A young Bell scientist makes a major math breakthrough

Source: Time Magazine

Simplex Method

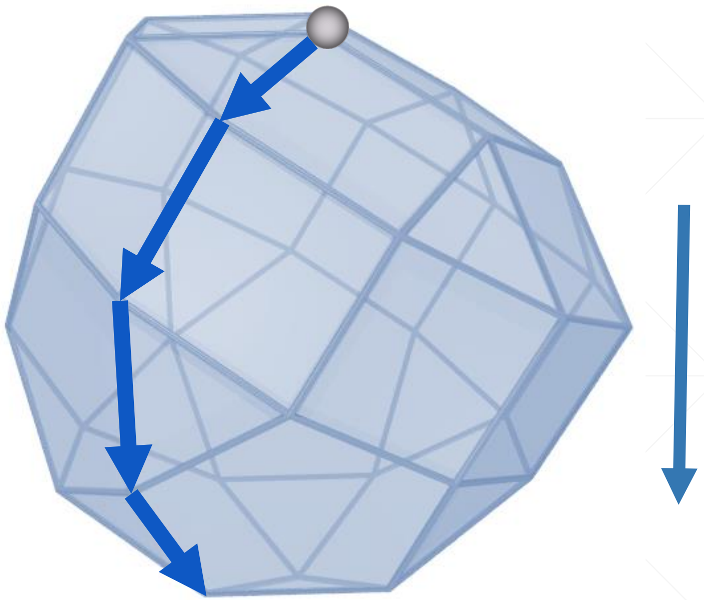


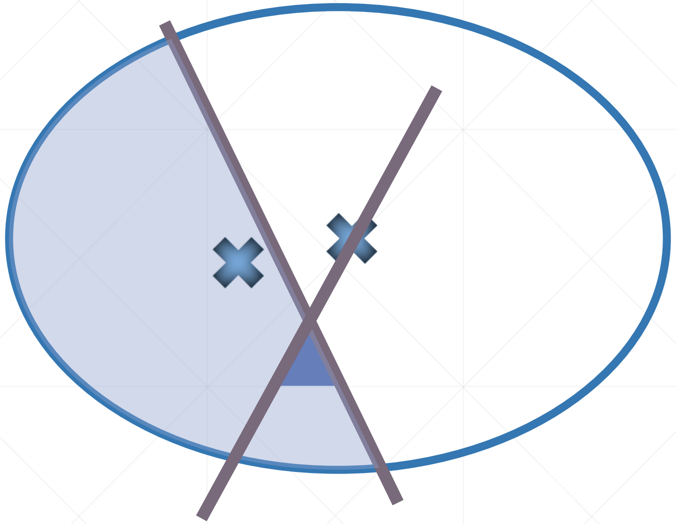
Interior Point Methods



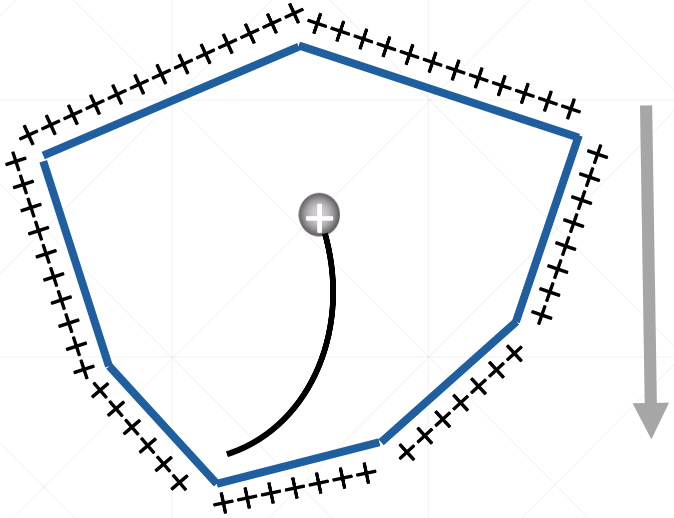
Best algorithms in theory and one of the best in practice

Techniques





$$K = \{Ax \geq b, c^{\top}x \leq OPT\}$$

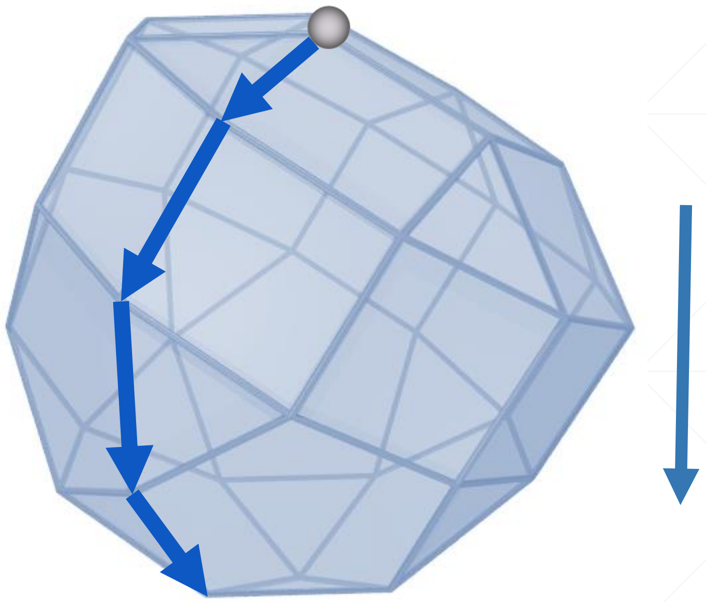


Simplex Method
is an example of
Iterative Method

Ellipsoid Method
is an example of
Divide and Conquer

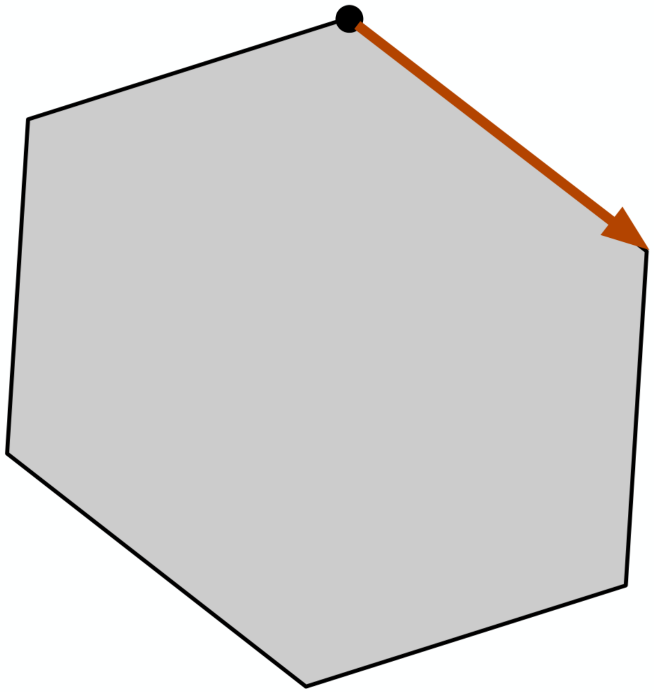
Interior Point Method
is an example of
Homotopy Method

Simplex Method



Simplex

If the polytope is perturbed,
# of steps $\leq$ roughly n^2.



Start with a vertex
In each step,
move to a lower vertex

Problem: Number of vertices
on this path can be
exponential!

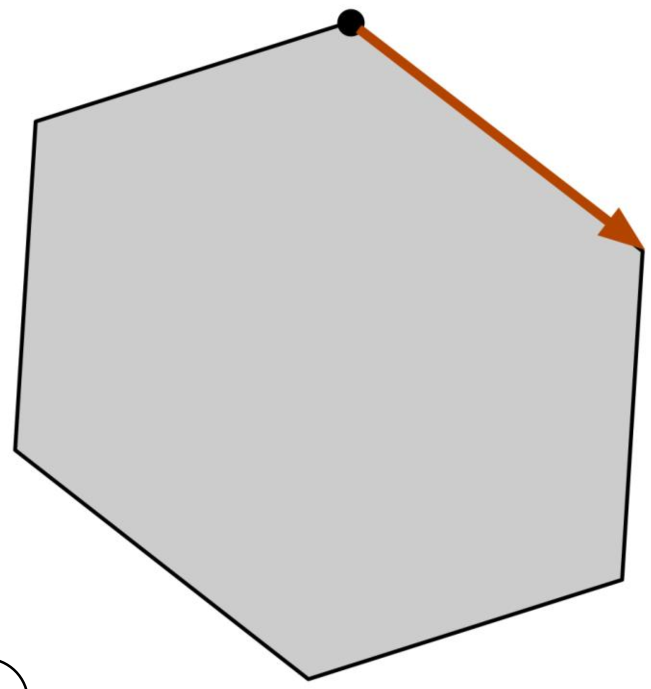
Simplex: how to find initial vertex?

maximize $c^\intercal x$
subject to
$Ax \leq b$
$x \geq 0$

Simplex: how to go to better vertex?

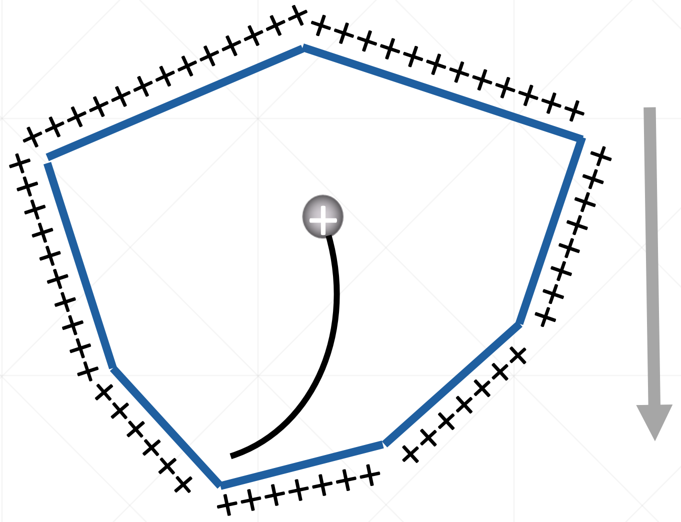
maximize $c^\intercal x$
subject to
$Ax \leq b$
$x \geq 0$

1. There must be $\hat{A}x = \hat{b}$.
2. Find y satisfying $n - 1$ of the equations, $c^\intercal y > 0$.
3. Change $x = x + \epsilon y$, until some new equation becomes tight.



Open Problem:
Can we have a rule to select new vertex such that # of steps are polynomially bounded?

Interior Point Method

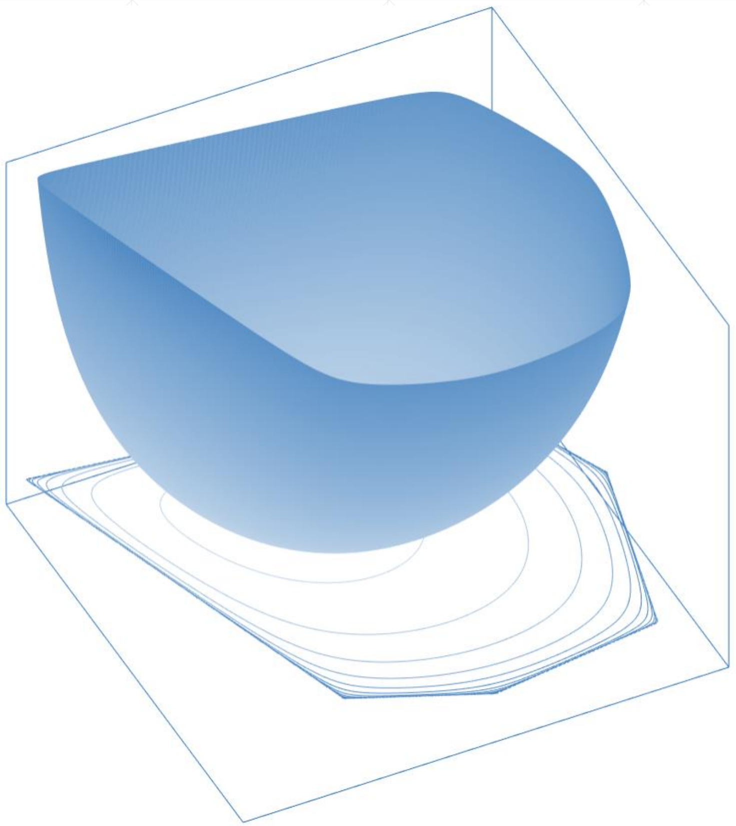


Constrained to Unconstrained

(Barrier Function)

$$\min_{Ax \geq b} c^\top x$$

- Difficulties lie in the polytope constraint $\{Ax \geq b\}$
- Smooth function is easier to minimize
- **Replace the constraint by a smooth function**



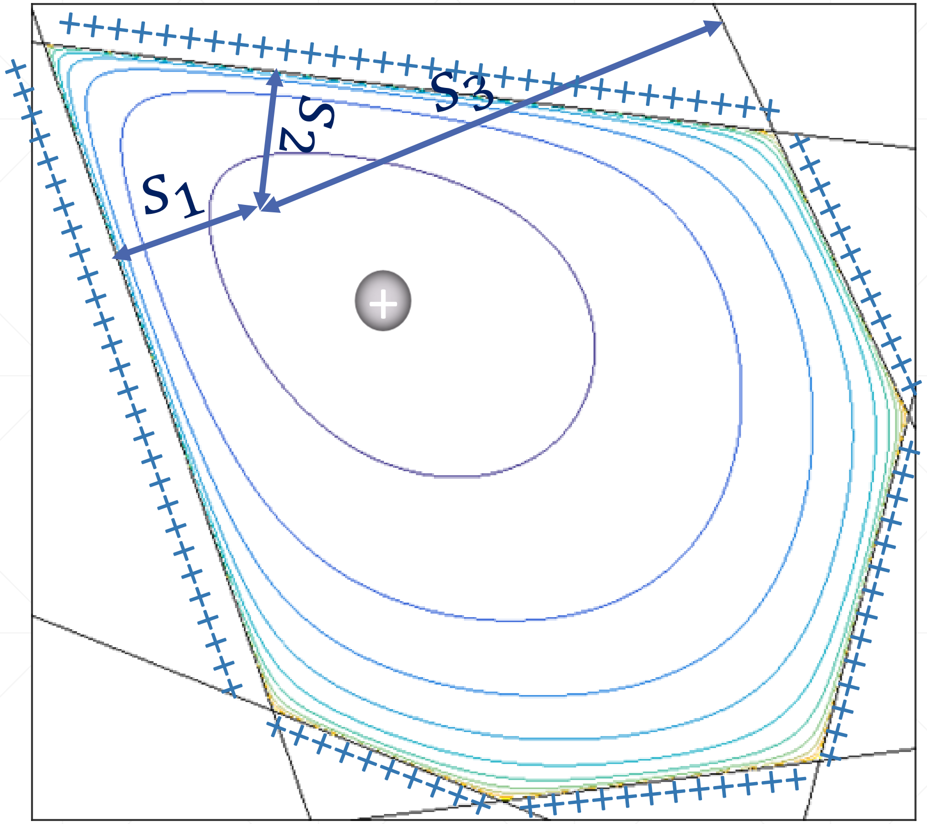
Requirements for barrier function:

- Smooth
- Blow up on the boundary

Example: Log Barrier Function

$$p(x) = \sum_{i=1}^{m} \log\left(\frac{1}{s_i(x)}\right)$$

- $s_i(x)$ is the distance from x to constraint i
- p blows up when x close to boundary



You can view this “physically”.

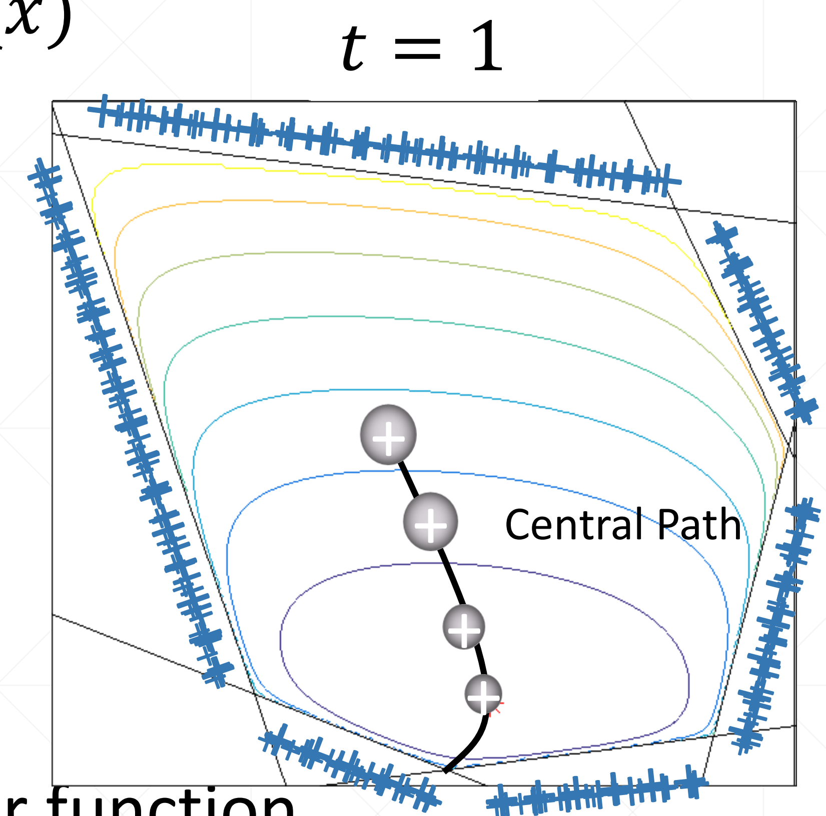
Central Path

(t = force)

$$\min_{Ax \geq b} c^\top x \sim \min_x c^\top x + \quad p(x)$$

Repeat

- Compute $\min_x c^\top x + t \cdot p(x)$
- Decrease t



What barrier function p should we pick?

- Karmarkar used log barrier function.
- Nesterov and Nemirovskii used universal barrier function.

Central Path

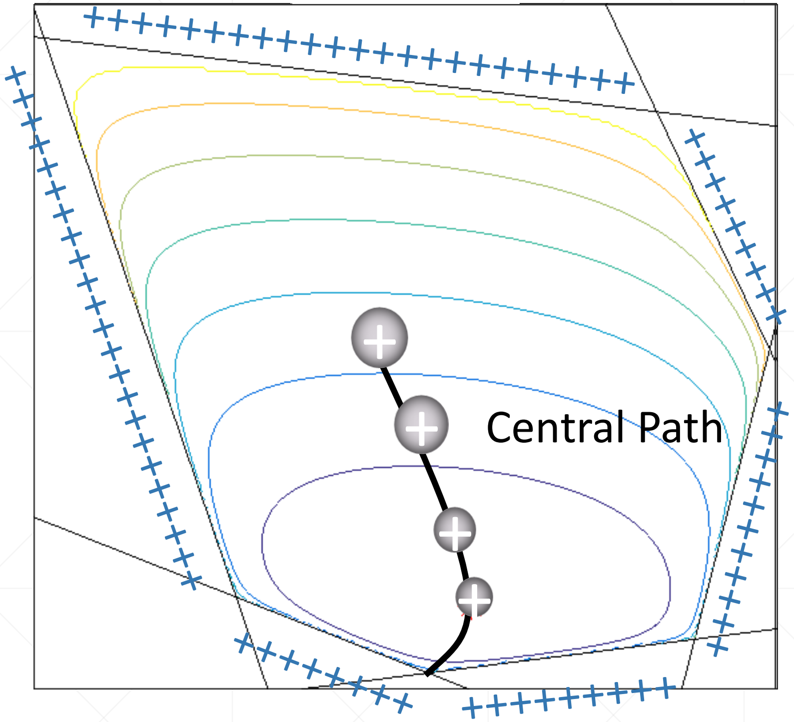
(t = force)

$$\min_{Ax \geq b} c^\top x \sim \min_x c^\top x + t \cdot p(x)$$

Repeat

- Compute $\min_x c^\top x + t \cdot p(x)$
- Decrease t

$t = 1$



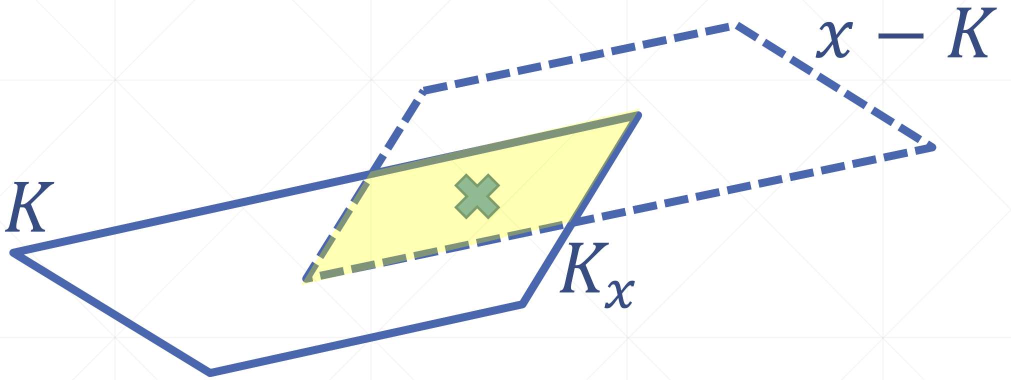
What barrier function p should we pick?

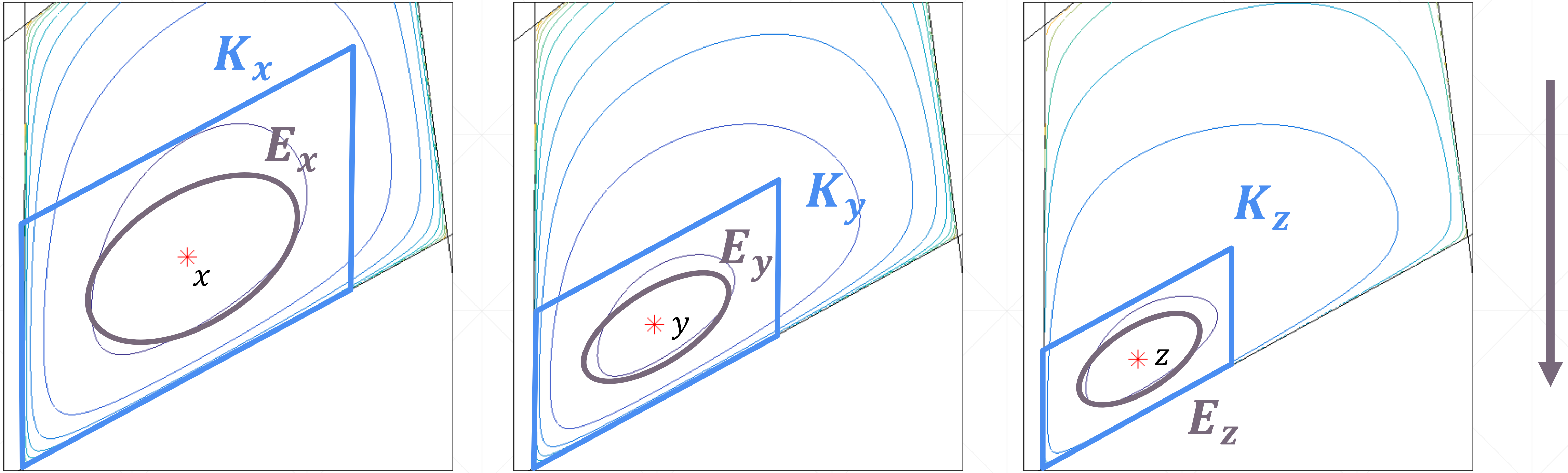
- p is easy to compute
- Central path is smooth.

Condition For a Good Barrier Function

Observation: easy to optimize over ellipsoid.

Approximate the polytope by ellipsoids!

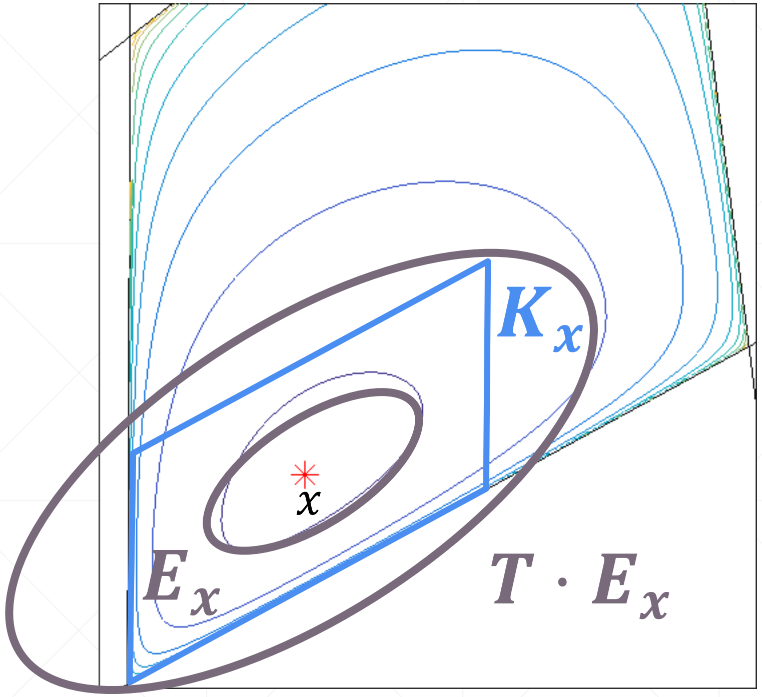




Condition For a Good Barrier Function

Observation: easy to optimize over ellipsoid.

Approximate the polytope by ellipsoids!



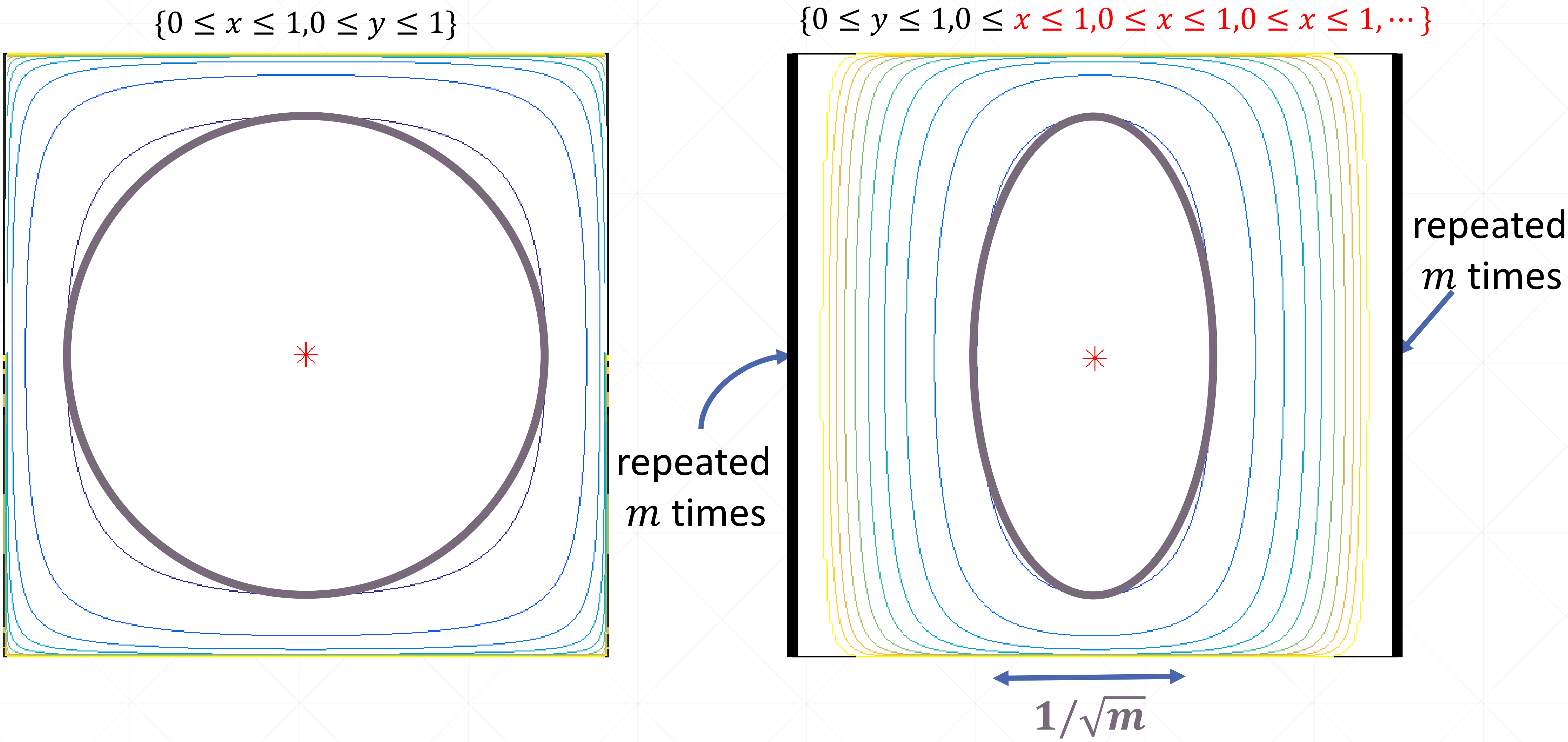
#iter depends on how well E approximates K.

Say E is a T-approximation of K if

$$E \subset K \subset T \cdot E$$

IPM takes ***T* iters** if E_x is a ***T*-approximation** of K_x for all x.

Log Barrier depends on representation



Condition For a Good Barrier Function

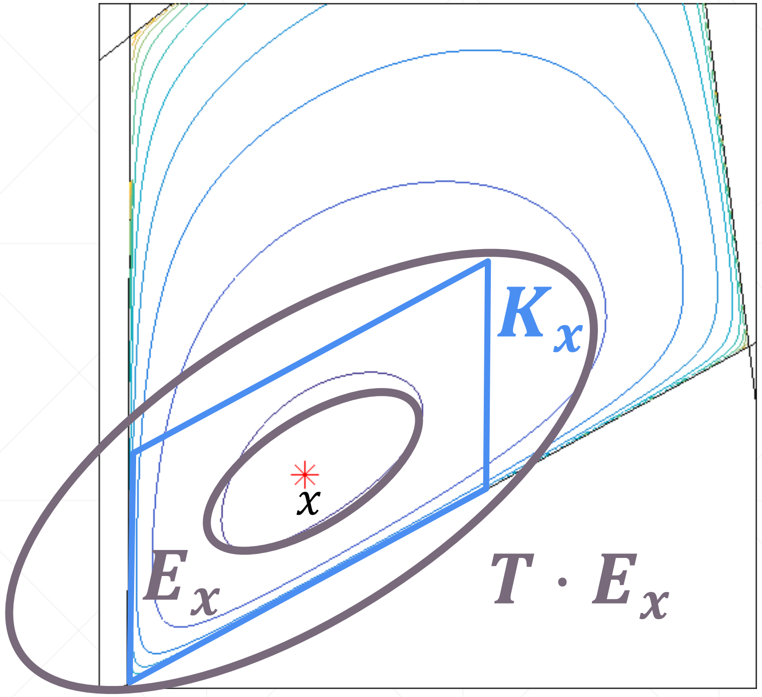
Observation: easy to optimize over ellipsoid.

Approximate the polytope by ellipsoids!

#iter depends on how well E approximates K.

Say E is a T-approximation of K if

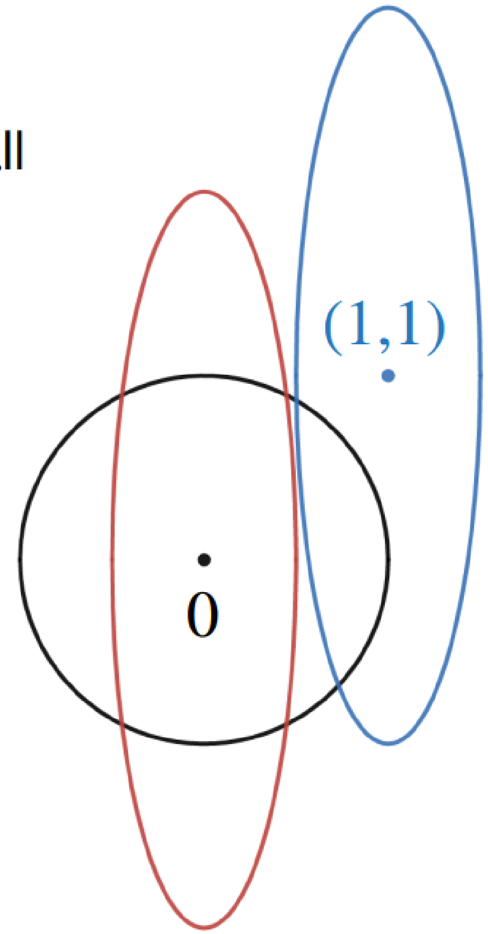
$$E \subset K \subset T \cdot E$$



IPM takes ***T* iters** if E_x is a ***T*-approximation** of K_x for all x.

Ellipsoid

Ellipsoid: a squished ball



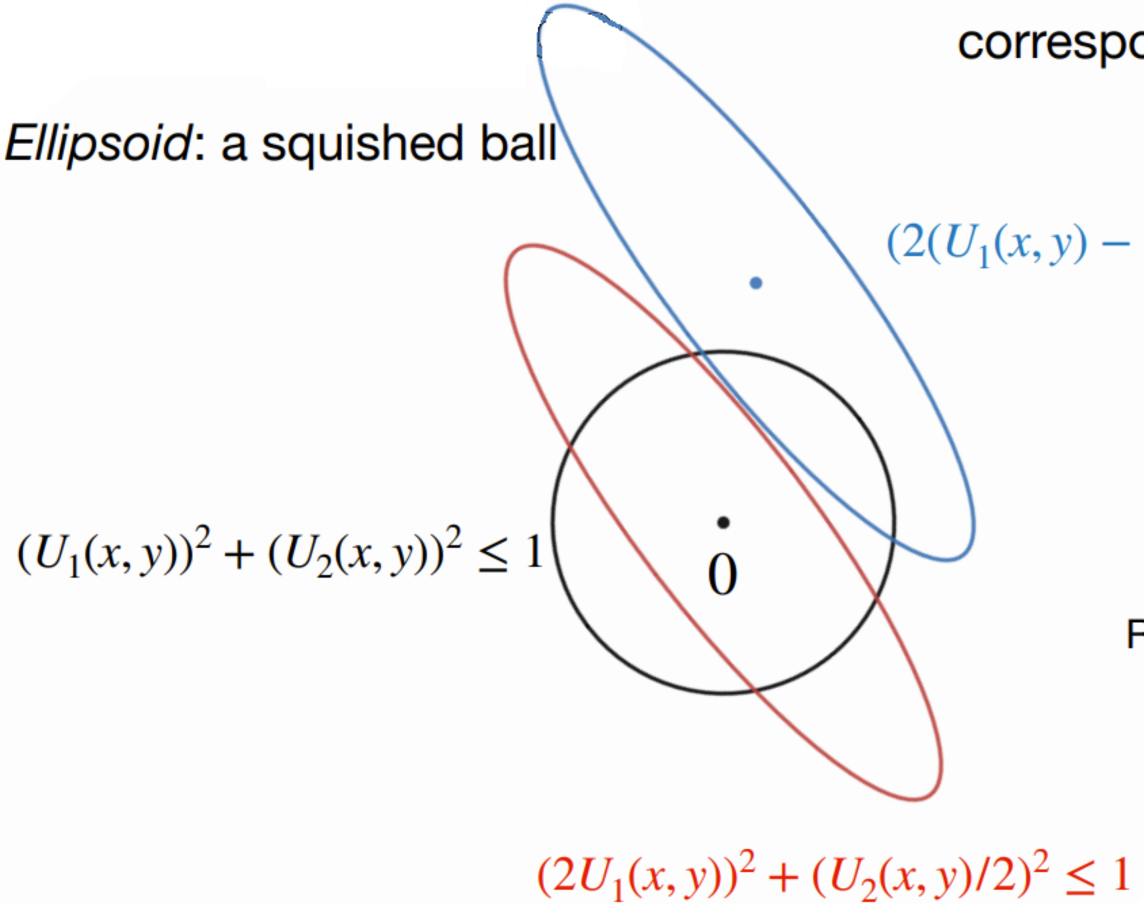
Ratio of area of ellipsoid to sphere:

$$\frac{1}{2} \cdot \frac{2}{1} = 1$$

Ellipsoid

Let U^{-1} be the linear transformation corresponding to a rotation.

Ellipsoid: a squished ball

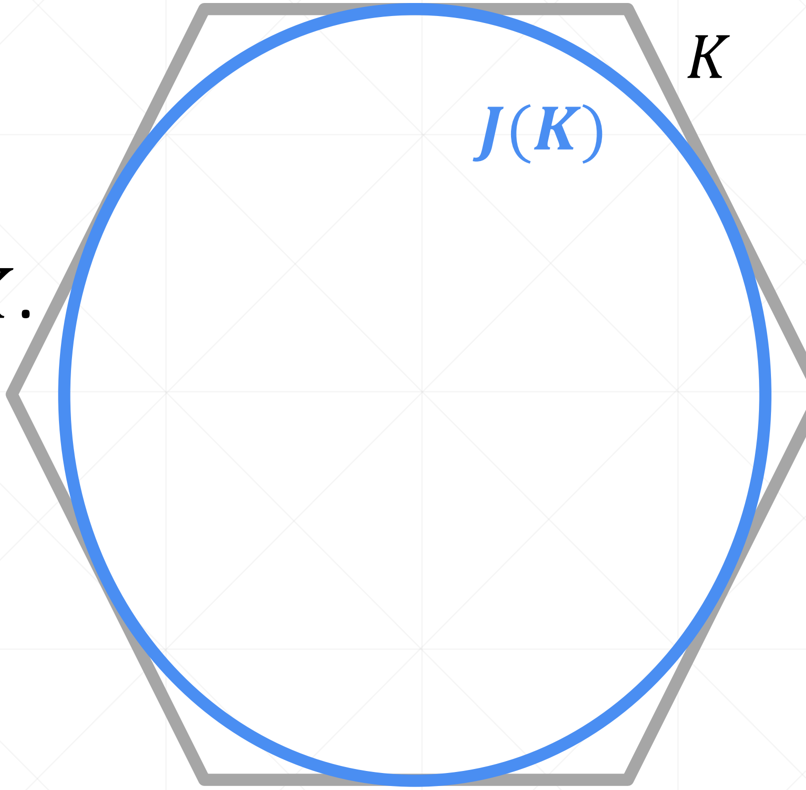


Ratio of area of ellipsoid to sphere:

$$\frac{1}{2} \cdot \frac{2}{1} = 1$$

John Ellipsoid

For any polytope K, let $J(K)$ be the maximum ellipsoid inside K.



<u>Theorem</u>

If K is symmetric, $J(K)$ is a $\sqrt{n}$ **approximation** of K.

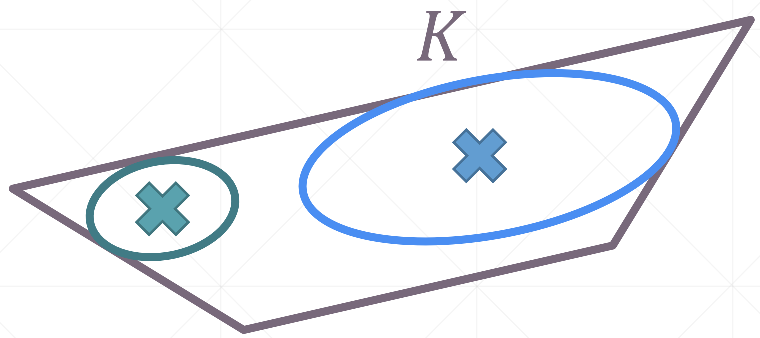
Furthermore,

$J(K)$ can be approximated by
solving $\tilde{O}(1)$ linear systems.

John Ellipsoid Barrier

$$p(x) = -\log(\mathrm{vol}(J(K_x))$$

(illustration only)

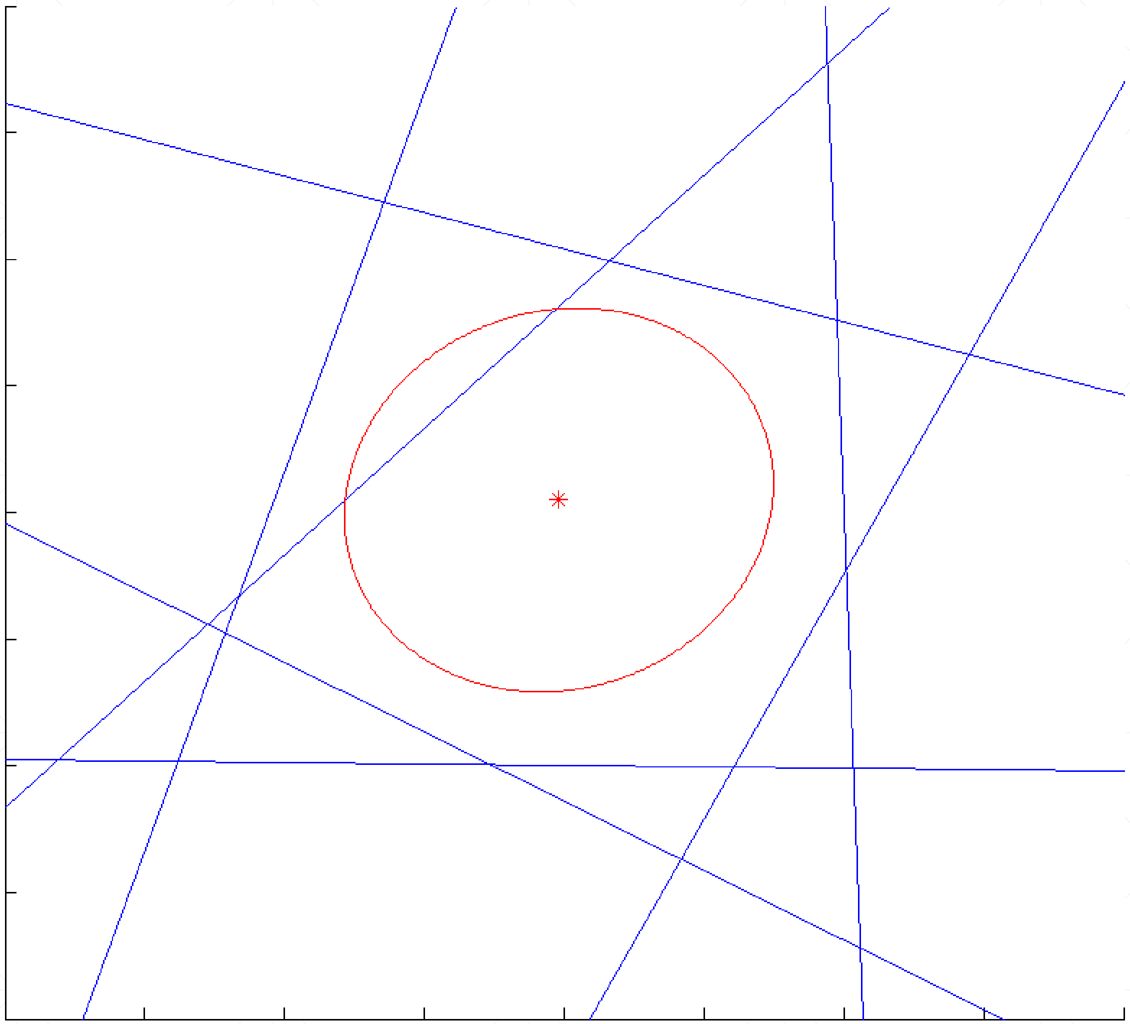


- $J(K_x)$ is the largest ellipsoid centered at x.
- It does not depend on representation.
- It involves volume computation but is volume of ellipsoid!
  - Universal barrier involves volume of polytope which is much **harder**.

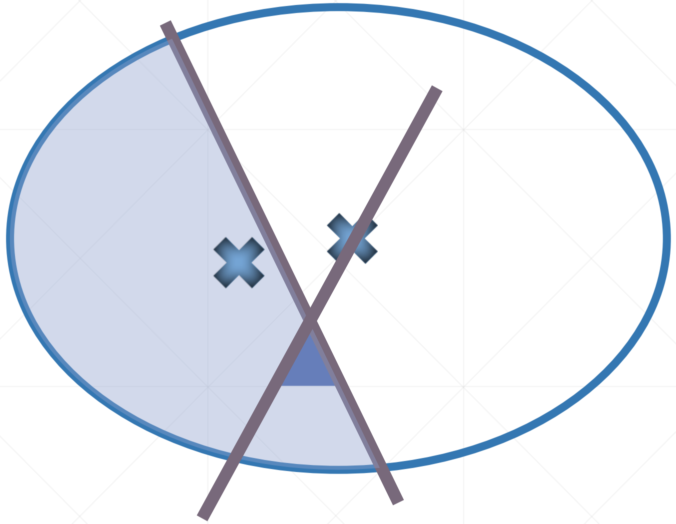
Alternative View

Algorithm:

- Update John ellipsoid
- Push the cost constraint
- Repeat



Ellipsoid Method

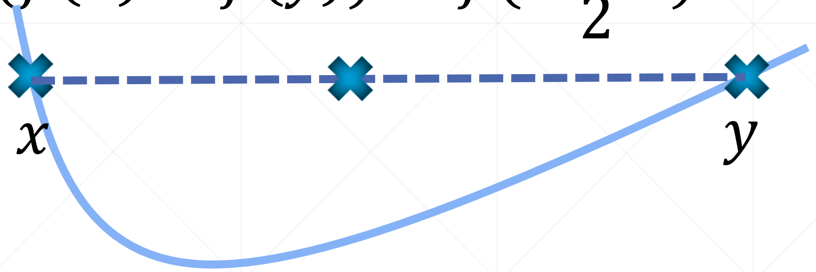


$K = \{Ax \geq b, c^\top x \leq OPT\}$

Convex Minimization

f is convex if

$$\frac{1}{2}(f(x) + f(y)) \geq f(\frac{x+y}{2})$$



Minimize a general convex function f

Assume:

- Access the function by computing gradient
- f is not exponentially large or the solution is not exponentially far away.

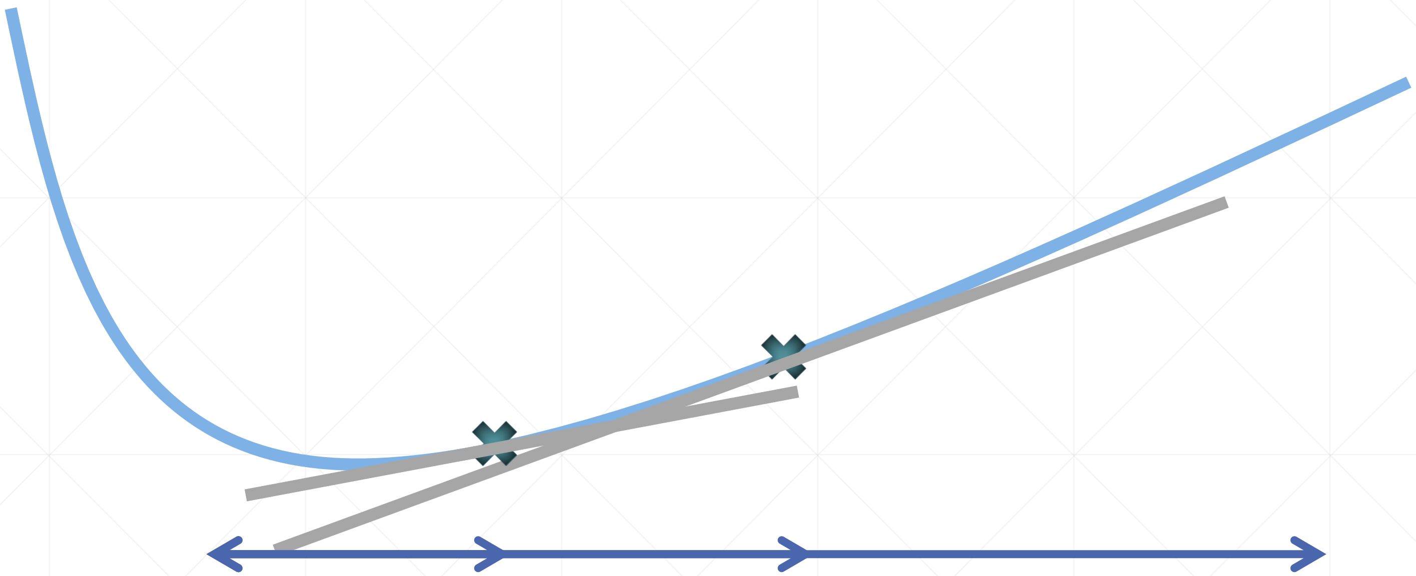
Goal: decrease the error $f(x_k) - \mathrm{OPT}$ exponentially.

Consequence: can handle constraints

Example: linear program, logistic regression

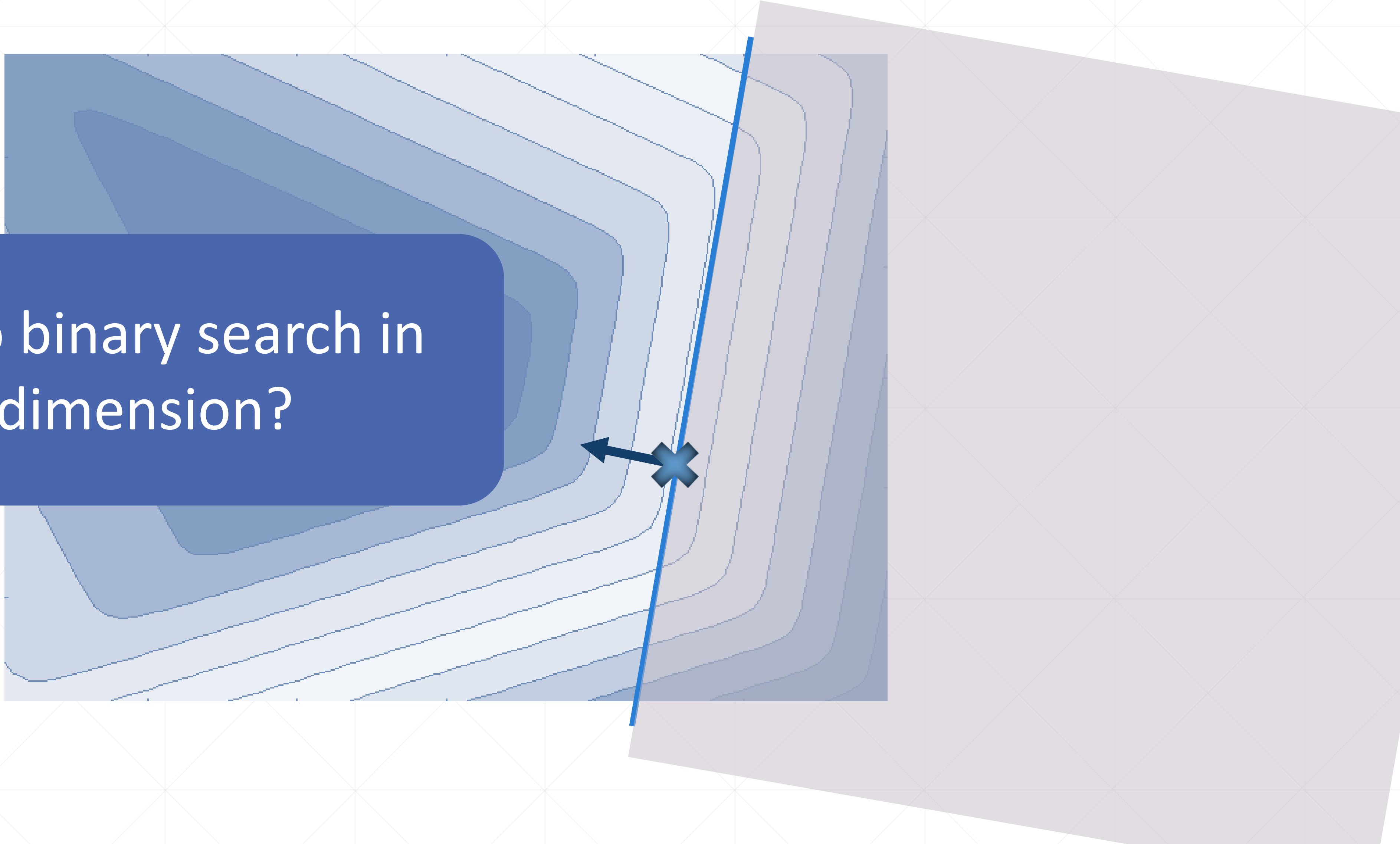
Why linear program is a convex function?

Why can we solve in 1 dimension?



Binary search!

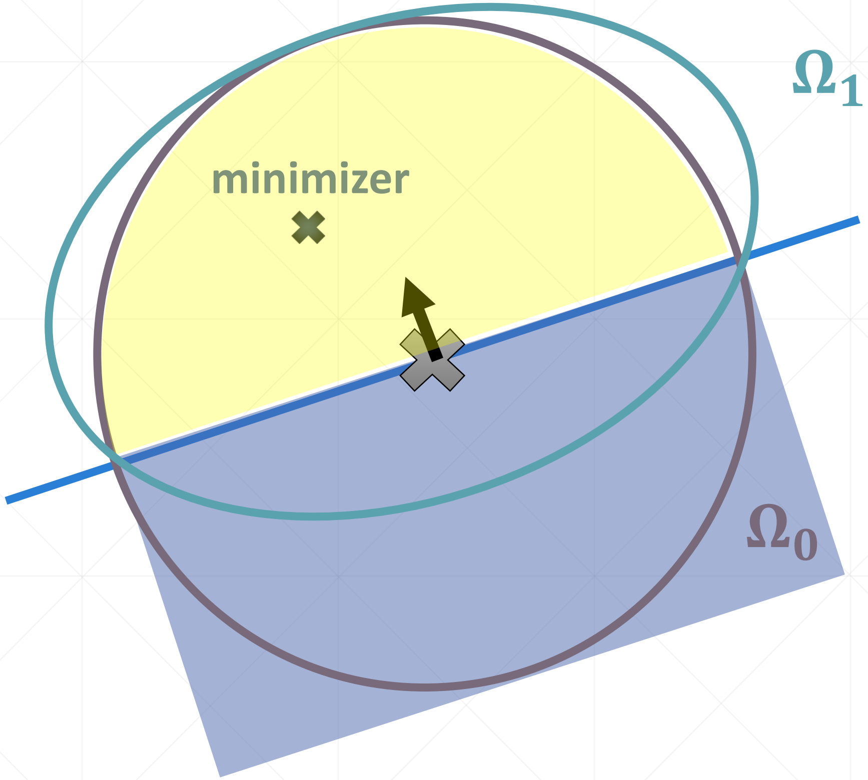
Convexity Allows Us to Cut



How to do binary search in high dimension?

Ellipsoid Method [Shor; Nemirovski-Yudin; Khachiyan 70s]

- For $k = 0,1,\cdots$
  - Compute gradient at center of Ω_k
  - Let Ω_{k+1} be the smallest ellipsoid containing
    $\Omega_k \cap \{\text{Half Space}\}$.



New volume is at most $1 - \frac{1}{n}$ of old volume.

$O(n^2 \log(1/\varepsilon))$ iterations.

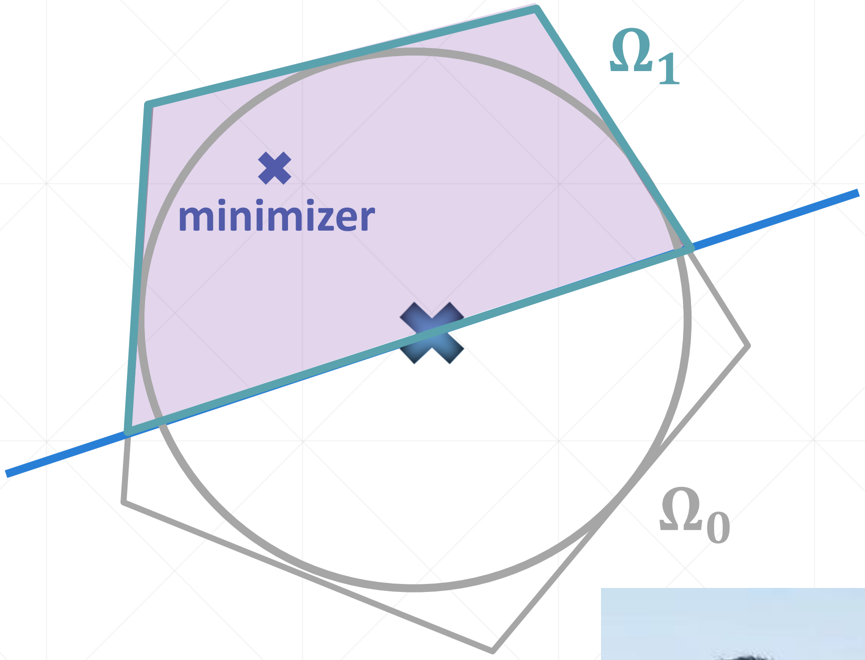
$O(n^2)$ time per iteration.

John Ellipsoid Method

John Ellipsoid Method
never throws away information.

- For $k = 0,1,\cdots$
  - Compute largest ellipsoid inside Ω_k (John Ellipsoid)
  - Compute gradient at the center.
  - Let $\Omega_{k+1} = \Omega_k \cap \{\text{Half Space}\}$.

New volume is at most 0.9 of old volume.

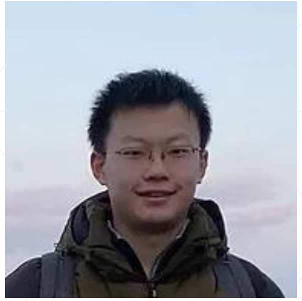


$O(n \log(1/\varepsilon))$ iterations. [Khachiyan 88]

$O(n^{2.88})$ time per iteration. [Nesterov-Nemirovskii 89]

Improved to $O(n^{2.38})$ using slightly different ellipsoid. [Vaidya 89]

Further improved to $O(n^2)$. [2020]



Haotian
UW PhD