□ **Karmarkar's Algorithm**

Motivation

Total Complexity of iterative algorithm =
(# of iterations) $\times$ (operations in each iteration)

Simplex Method

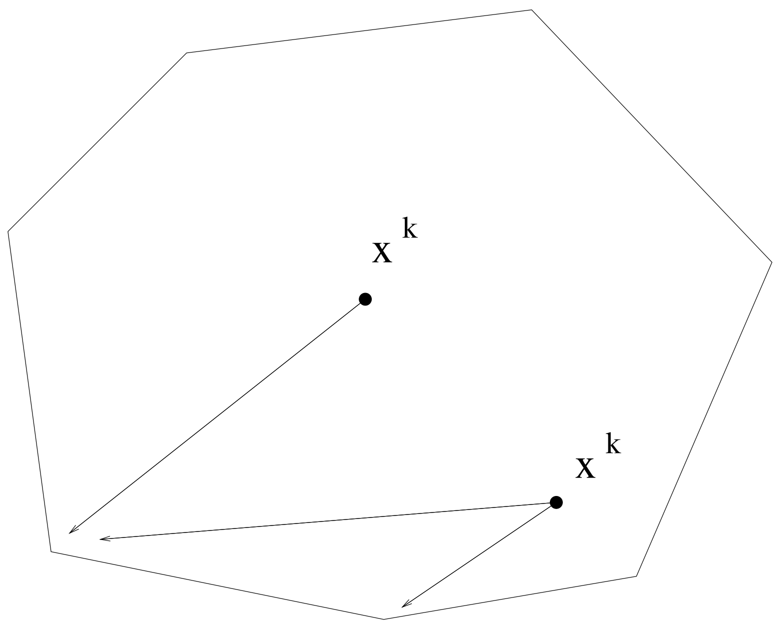
- Simple operations
  : Only check adjacent extreme points.
- May take many iterations
  : Klee-Minty example.

Karmarkar's Algorithm

- Complicated iterations
  : Check all directions for the best one.
- Take fewer iterations

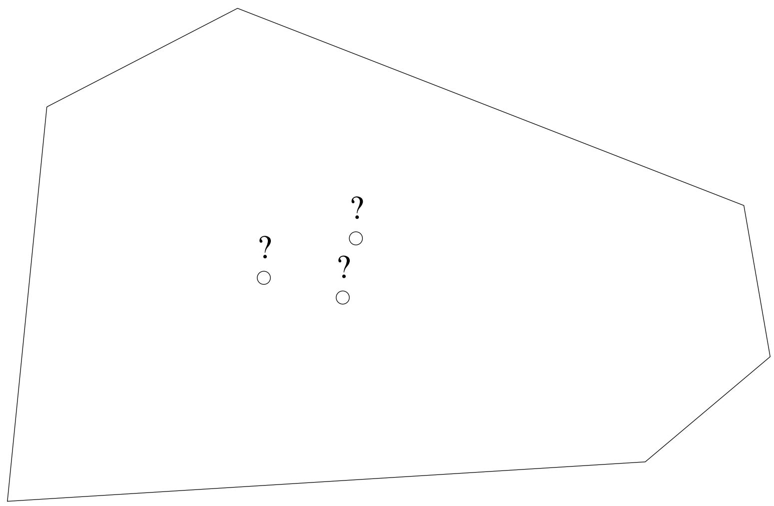
<u>Observations:</u>

(1) If a current solution is "near the center" of a polytope, it makes sense to move along the "steepest descent" direction.



(2) If $\mathbf{x}^k$ is not "near the center", we can "re-scale" the coordinate to transform it to be "near the center".

Questions: What's the center of a polytope?

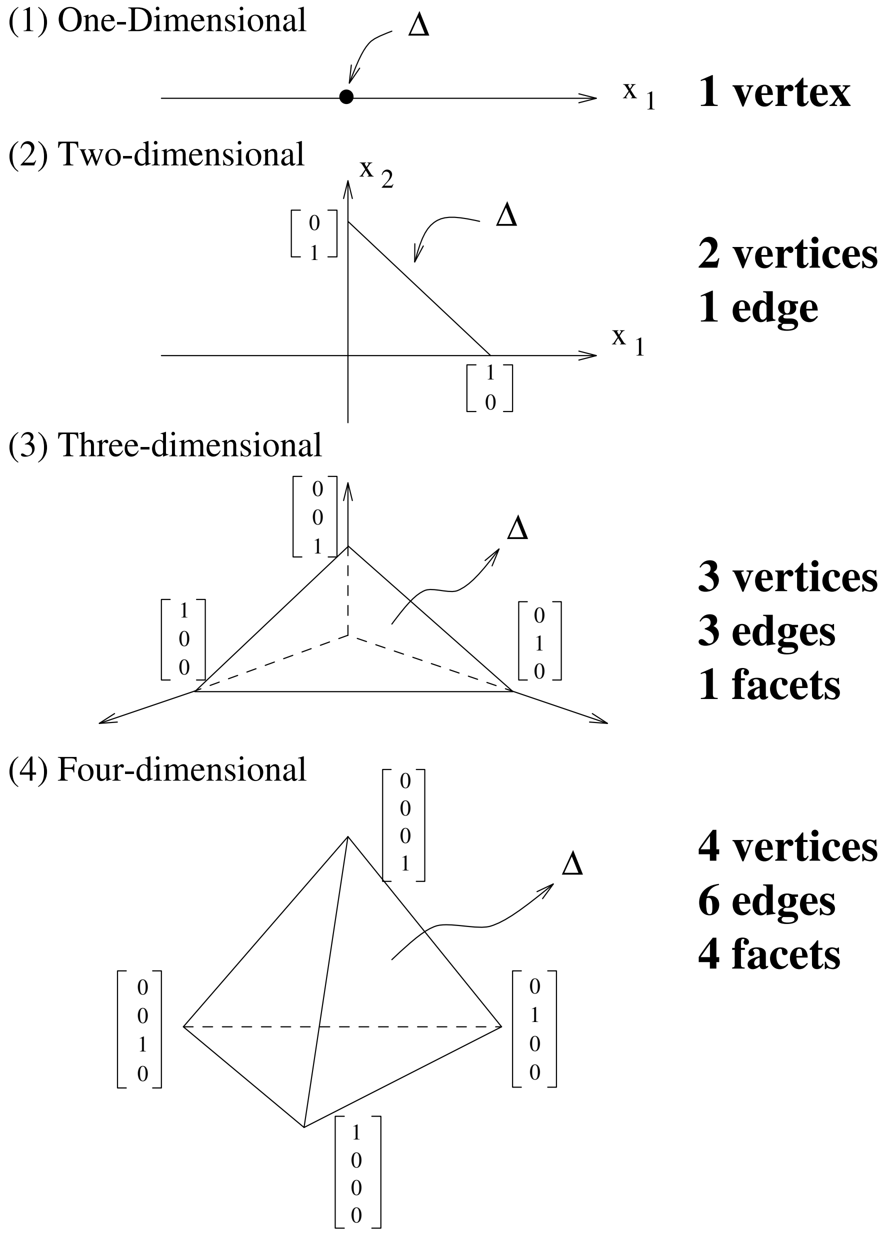


Simplex Structure

- Simplex regular polygon in R^n.

$$
\begin{aligned}
\Delta &= \{\mathbf{x} \in R^n \mid \textstyle\sum_{j=1}^{n} x_j = 1, x_j \geq 0\} \\
&= \{\mathbf{x} \in R^n \mid e^T \mathbf{x} = 1, \mathbf{x} \geq 0\}
\end{aligned}
$$

where $e = (1, \ldots, 1)^T \in R^n$.


(1) One-Dimensional
Δ
x 1
1 vertex
(2) Two-dimensional
x 2
Δ
2 vertices
1 edge
x 1
(3) Three-dimensional
Δ
3 vertices
3 edges
1 facets
(4) Four-dimensional
Δ
4 vertices
6 edges
4 facets

Questions:

For n-dimensional case,

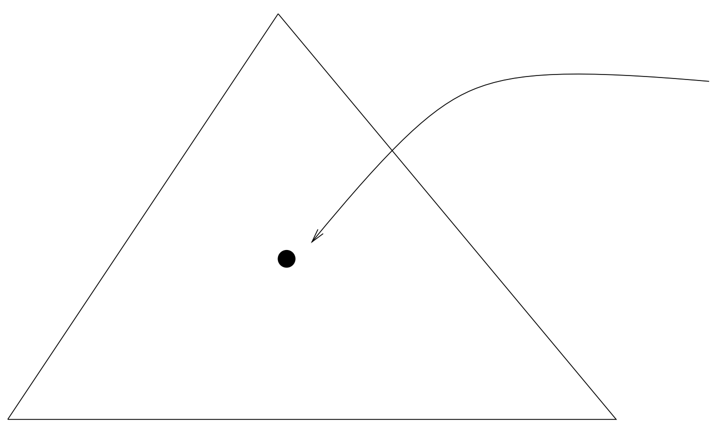
(1) How many vertices in Δ?

(2) How many edges in Δ?

(3) How many facets in Δ?

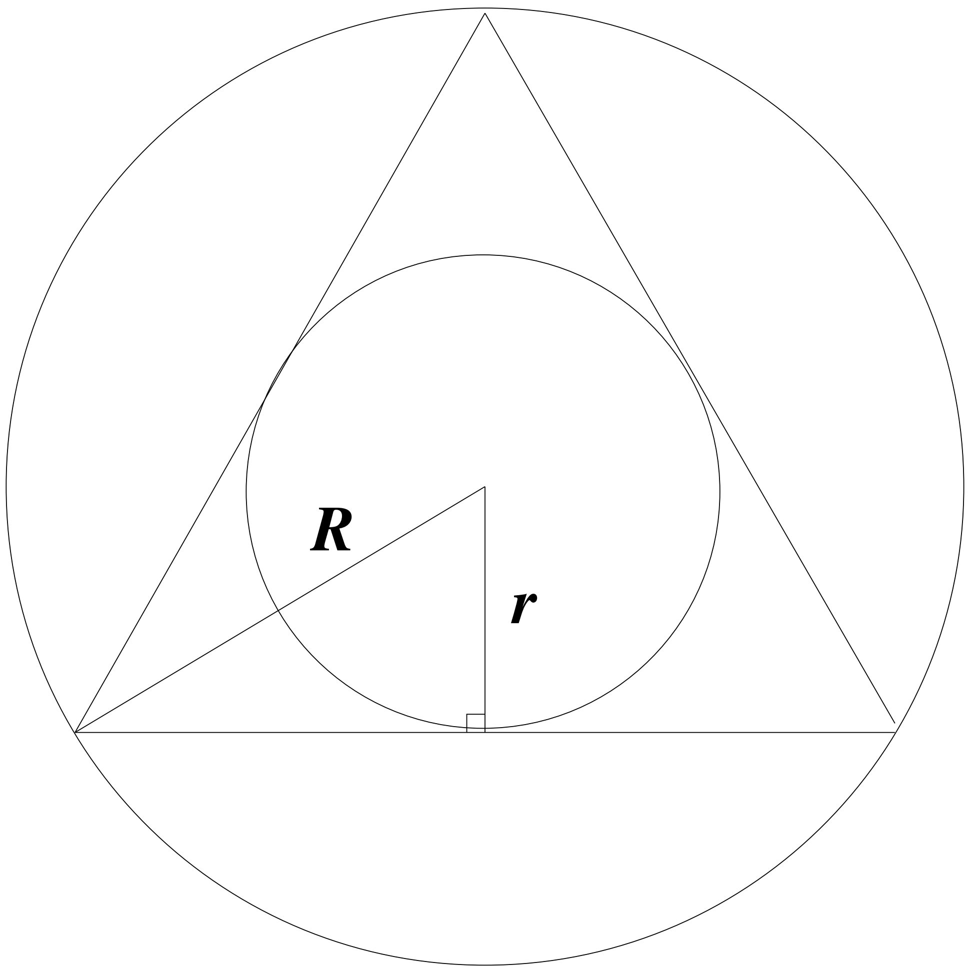
(4) Given $\mathbf{x} \in \Delta$, how can we identify it is a vertex? on an edge? or in the interior?

(5) What is the center of Δ?



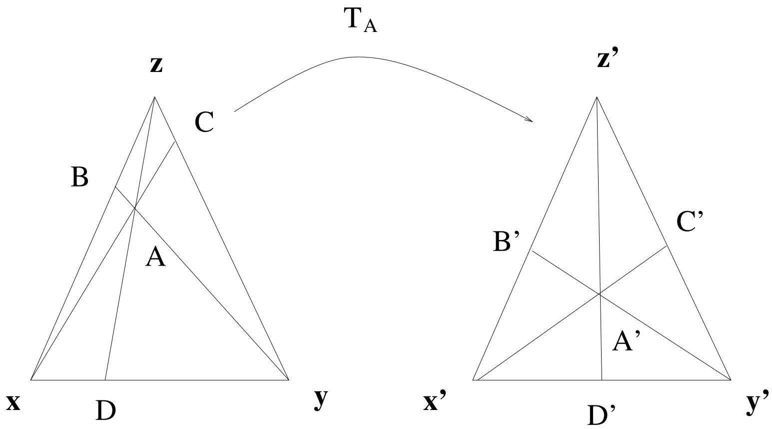
$$\begin{pmatrix} 1/n \\ 1/n \\ \vdots \\ 1/n \end{pmatrix} = \frac{e}{n}$$

(6) How far between the center to a vertex? to an facet?



$R = \frac{\sqrt{n-1}}{\sqrt{n}}$

$r = \frac{1}{\sqrt{n(n-1)}}$



$$X_A = \begin{pmatrix} 3/10 & 0 & 0 \\ 0 & 1/10 & 0 \\ 0 & 0 & 3/5 \end{pmatrix} \qquad X_A^{-1} = \begin{pmatrix} 10/3 & 0 & 0 \\ 0 & 10 & 0 \\ 0 & 0 & 5/3 \end{pmatrix}$$

$$\mathbf{x} = (1,0,0) \qquad \mathbf{x}' = (1,0,0)$$

$$\mathbf{y} = (0,1,0) \qquad \mathbf{y}' = (0,1,0)$$

$$\mathbf{z} = (0,0,1) \qquad \mathbf{z}' = (0,0,1)$$

$$A = (3/10, 1/10, 3/5) \qquad A' = (1/3, 1/3, 1/3)$$

$$B = (1/3, 0, 2/3) \qquad B' = (1/2, 0, 1/2)$$

$$C = (0, 1/7, 6/7) \qquad C' = (0, 1/2, 1/2)$$

$$D = (3/4, 1/4, 0) \qquad D' = (1/2, 1/2, 0)$$

Questions:

How to re-scale an interior point $\bar{\mathbf{x}} \in \Delta$ to the center?

Definitions:

(1) $\mathbf{x}^k \in \Delta$ is an *interior point* of Δ, if $\sum_{j=1}^n x_j^k = 1$ and $x_j^k > 0, \ \forall j$.

(2) matrix

$$X_k := \operatorname{diag}(\mathbf{x}^{\mathbf{k}}) = \begin{pmatrix} x_1^k & & \\ & \ddots & \\ & & x_n^k \end{pmatrix}$$

(3) projective transformation

$$\begin{aligned} T_{x^k} : \quad \Delta & \longrightarrow \quad \Delta \\ x^k & \longrightarrow \quad T_{x^k}(\mathbf{x}) = \frac{X_k^{-1}\mathbf{x}}{e^T X_k^{-1}\mathbf{x}} \end{aligned}$$

General Properties:

(1) T_{x^k} maps Δ to Δ.

(2) $\mathbf{x} \in \Delta$

$$\begin{array}{lcl} \text{a vertex} & \Longrightarrow & T_{x^k}(\mathbf{x}) \text{ a vertex} \\ \text{on an edge} & \Longrightarrow & T_{x^k}(\mathbf{x}) \text{ on an edge} \\ \text{on a facet} & \Longrightarrow & T_{x^k}(\mathbf{x}) \text{ on a facet} \\ \text{in interior} & \Longrightarrow & T_{x^k}(\mathbf{x}) \text{ in interior} \end{array}$$

(3) $T_{x^k}(x^k) = \frac{e}{n}$

(4) T_{x^k} is one-to-one and onto.

(5) $T_{x^k}^{-1}$ exists.

$$y = \frac{X_k^{-1} x}{e^T X_k{}^{-1} x}$$

$$\begin{tikzcd} \Delta \arrow[r, shift left, "T_{x^k}"] & \Delta \arrow[l, shift left, "T_{x_k}^{-1}"] \end{tikzcd}$$

$$T_{x^k}^{-1}(y) = \frac{X_k y}{e^T X_k y}$$

Karmarkar's Standard Form:

$$
(KLP) \quad \begin{array}{lrl}
\min & \mathbf{c}^T\mathbf{x} & \\
\text{s. t.} & \mathbf{A}\mathbf{x} = 0 & \rightarrow \text{Null space } \mathcal{N}(\mathbf{A}) \\
 & e^T\mathbf{x} = 1\} & \text{simplex} \\
 & \mathbf{x} \geq 0\} & \text{in } \Delta
\end{array}
$$

$$
\boxed{\min \mathbf{c}^T\mathbf{x} \text{ s. t. } \mathbf{x} \in \mathcal{N}(\mathbf{A}) \cap \Delta_n}
$$

$$
\begin{array}{ll}
\min & \mathbf{c}^T\mathbf{x} \\
\text{s. t.} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\
 & \vdots \\
 & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = 0 \\
 & x_1 + x_2 + \cdots + x_n = 1 \\
 & x_1, x_2, \ldots, x_n \geq 0
\end{array}
$$

Assumptions:

(A1) *Interior-Point Assumption*: $\mathbf{A}e = 0$, *i.e*
$\mathbf{x}^0 = (\frac{1}{n}, \dots, \frac{1}{n})$ is interior feasible.

(A2) *Zero-Optimal Value Assumption:*
$z^* = \mathbf{c}^T \mathbf{x}^* = 0$

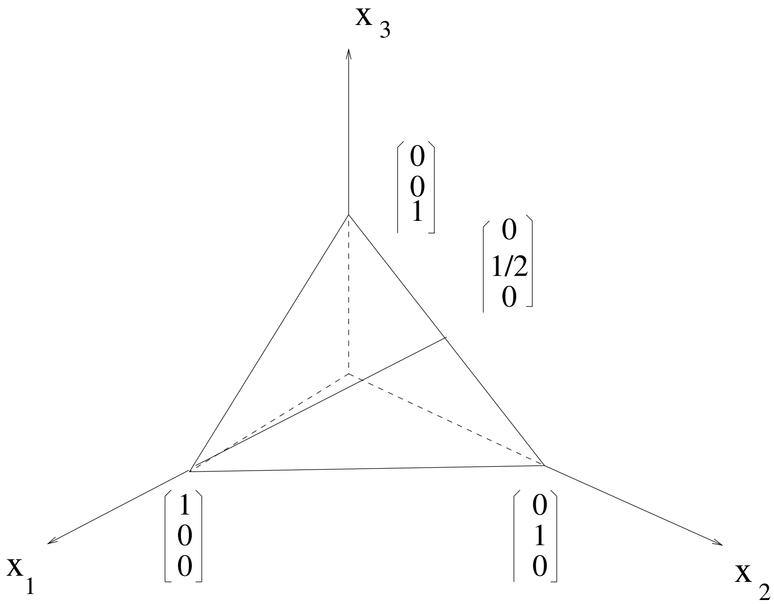
Definitions:

$\mathbf{x}$ is a feasible solution of (KLP) if
$\mathbf{A}\mathbf{x} = 0$ and $e^T \mathbf{x} = 1, \mathbf{x} \geq 0$.
It is an *interior feasible solution* of (KLP) if
$\mathbf{A}\mathbf{x} = 0,\ e^T \mathbf{x} = 1,\ \mathbf{x} > 0$.

Example:

$$
\begin{array}{llllll}
\min & -x_1 + & 1 & & & \\
\text{s. t.} & & x_2 - & x_3 & = 0 \\
& x_1 + & x_2 + & x_3 & = 1 \\
& x_1, & x_2, & x_3 & \geq 0
\end{array}
$$

$\mathbf{c}^T = (-1, 0, 0), \mathbf{A} = (0, 1, -1), \mathbf{x}^* = (1, 0, 0), z^* = 0$



<u>Observations:</u>

(1) (KLP) is always bounded. $P = \mathcal{N}(\mathbf{A}) \cap \boldsymbol{\Delta}_n$

(2) If $P \neq \emptyset$, then (KLP) has a finite optimum.

(3) Dual (KLP)

$$
\begin{array}{rl}
\max & w_{m+1} \\
\text{s. t.} & a_{11}w_1 + \cdots + a_{m1}w_m + w_{m+1} \leq c_1 \\
(DKLP) & a_{12}w_1 + \cdots + a_{m2}w_m + w_{m+1} \leq c_2 \\
& \vdots \\
& a_{1m}w_1 + \cdots + a_{mn}w_m + w_{m+1} \leq c_n
\end{array}
$$

(4) Dual (KLP) is always feasible. For example:

(a) $w_1 = w_2 = \cdots = w_m = 0$
$w_{m+1} = \min\{c_1, c_2, \ldots, c_n\}$ then
$\mathbf{w} = (w_1, \ldots, w_m, w_{m+1})^T$ is feasible to (DKLP).

(b) In fact for any $\mathbf{w} = (w_1, \ldots, w_m)^T \in R^m$ choose
$w_{m+1} = \min_{1 \leq j \leq n}\{c_j - \sum_{i=1}^m a_{ij} w_i\}$ then
$(w_1, \ldots, w_m, w_{m+1})^T$ is feasible to (DKLP).

(5) Without loss of generality, we may assume that $\mathbf{A}$ has full row rank. Then $(\mathbf{A}\mathbf{A}^T)$ is nonsingular.

Basic Ideas:

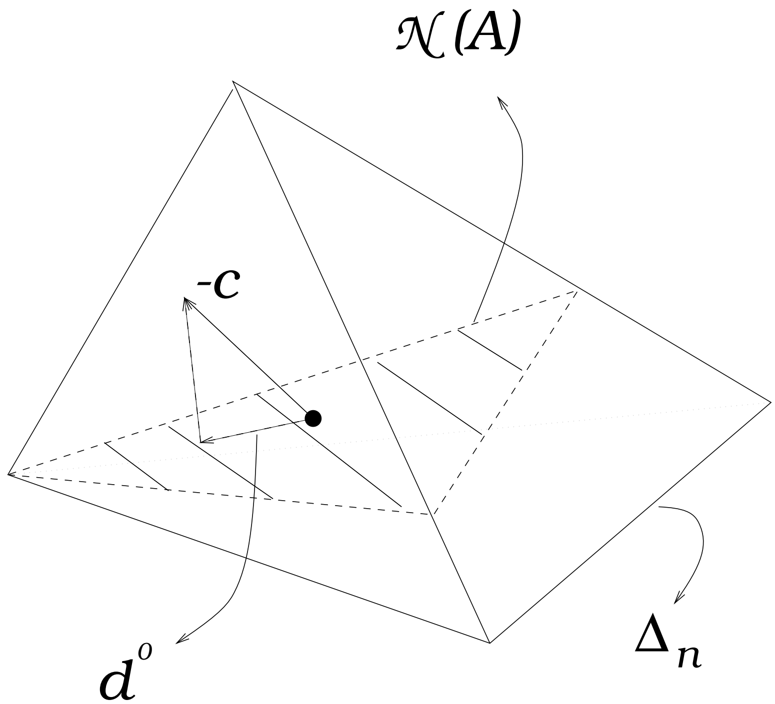
$$
\begin{array}{lrl}
\min & \mathbf{c}^T\mathbf{x} & \\
\text{s. t.} & \mathbf{A}\mathbf{x} = 0 & \rightarrow \mathcal{N}(\mathbf{A}) \\
& e^T\mathbf{x} = 1\} & \\
& \mathbf{x} \geq 0\} & \Delta_n
\end{array}
$$

Strategy:

(a) Move along the "steepest descent" direction: $-\mathbf{c}$

(b) To keep feasibility: Project $-\mathbf{c}$ onto $\mathcal{N}(\tilde{A})$

$i.e\ d^0 = [I - \tilde{A}^T(\tilde{A}\tilde{A}^T)\tilde{A}](-\mathbf{c})$. Note $r > \frac{1}{n}$



$P = \Delta_n \cap \mathcal{N}(A)$.

Hence $\mathbf{x}^1 = \mathbf{x}^0 + \frac{\alpha}{n}\frac{d^0}{\|d^0\|} \in P$ for $\alpha \in (0, 1]$

$\mathbf{x}^0 = \frac{e}{n} \in P$ is the "center".

We want that $\mathbf{x}^1 = \mathbf{x}^0 + \alpha_0 d^0 \in P$ and

$\mathbf{c}^T\mathbf{x}^1 \leq \mathbf{c}^T\mathbf{x}^0$

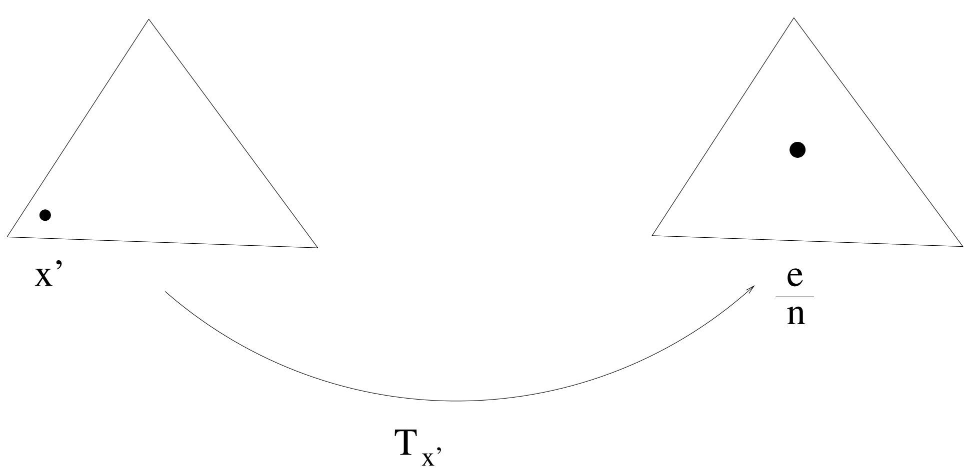
Question:

$\mathbf{x}'$ is no longer at the "center", so what's next?

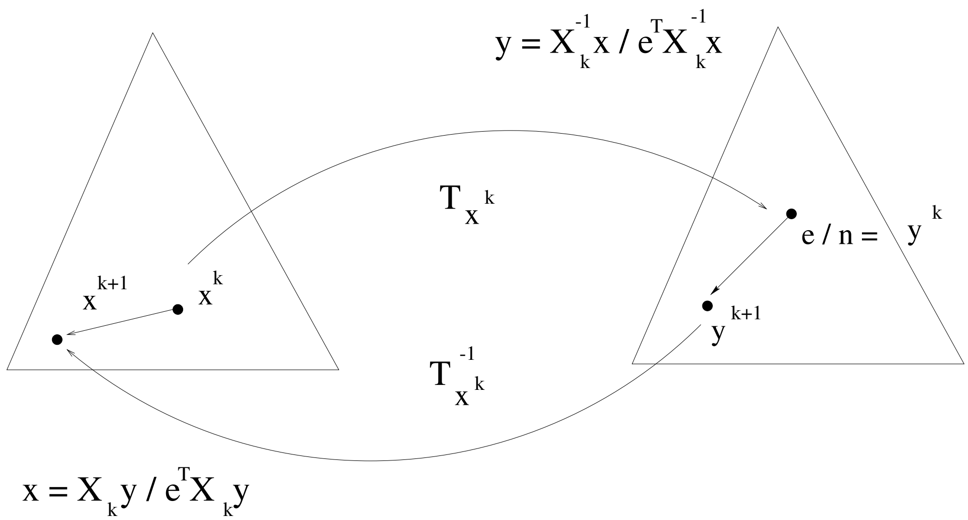
Answer:

For $0 \leq \alpha \leq 1$, $\mathbf{x}' \in P^i$ (the interior of P).

Therefore, we can "re-center" it at $\frac{e}{n}$ by the projective transformation $T_{x'}$



Karmarkar's Algorithm



$$
(P)\quad \begin{array}{ll} \min & \mathbf{c}^T\mathbf{x} \\ \text{s. t.} & \mathbf{A}\mathbf{x} = 0 \\ & e^T\mathbf{x} = 1 \\ & \mathbf{x} \geq 0 \end{array}
$$

$$
\begin{array}{ll} \min & \frac{\mathbf{c}^T\mathbf{X}_k\mathbf{y}}{e^T\mathbf{X}_k\mathbf{y}} \\ \text{s. t.} & \mathbf{A}\mathbf{X}_k\mathbf{y} = 0 \\ & e^T\mathbf{y} = 1 \\ & \mathbf{y} \geq 0 \end{array}
$$

$$
(P')\quad \begin{array}{ll} \min & (\mathbf{c}^T\mathbf{X}_k)\mathbf{y} \\ \text{s. t.} & \left[\begin{smallmatrix} \mathbf{A}\mathbf{X}_k \\ e^T \end{smallmatrix}\right]\mathbf{y} = \left[\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right] \\ & \mathbf{y} \geq 0 \end{array}
$$

$$
\begin{aligned} \mathbf{c}^k &= \mathbf{X}_k\mathbf{c} \\ \mathbf{B}_k &= \left[\begin{smallmatrix} \mathbf{A}\mathbf{X}_k \\ e^T \end{smallmatrix}\right] \end{aligned}
$$

$$
\begin{array}{ccl}
 & T_{x^k} & \\
\boxed{\mathbf{x}^k} & \longrightarrow & \boxed{\mathbf{y}^k} = \frac{e}{n} \\
 & & \mathbf{d}^k = -[I - \mathbf{B}_k^T(\mathbf{B}_k\mathbf{B}_k^T)^{-1}\mathbf{B}_k]\mathbf{c}^k \\
 & & = -[I - \mathbf{B}_k^T(\mathbf{B}_k\mathbf{B}_k^T)^{-1}\mathbf{B}_k]\mathbf{X}_k\mathbf{c} \\
 & & \mathbf{y}^{k+1} = \mathbf{y}^k + \frac{\alpha}{n}(\frac{d^k}{\|d^k\|}) \\
 & & = \frac{e}{n} + \frac{\alpha}{n}(\frac{d^k}{\|d^k\|}) \\
 & & 0 < \alpha \leq 1 \\
\boxed{\mathbf{x}^{k+1}} & \longleftarrow & \boxed{\mathbf{y}^{k+1}} \\
 & T_{x^k}^{-1} &
\end{array}
$$

$$
\begin{aligned}
\mathbf{x}^{k+1} &= \frac{\mathbf{X}_k\mathbf{y}^{k+1}}{e^T\mathbf{X}_k\mathbf{y}^{k+1}} \\
k &\longleftarrow k+1
\end{aligned}
$$

Karmarkar's Algorithm

Step 1: (Initialization)

$k \longleftarrow 0,\ \epsilon > 0$ be sufficiently small.

$\mathbf{x}^0 = \frac{e}{n},\ 0 < \alpha \leq 1.$

Step 2: (Optimality Check)

If $\mathbf{c}^T\mathbf{x}^k \leq \epsilon$, then STOP!

$\mathbf{x}^k$ is optimal.

Step 3: (Improvement)

Let

$$\mathbf{X}_k = \begin{pmatrix} x_1^k & & \\ & \ddots & \\ & & x_n^k \end{pmatrix},\ \mathbf{B}_k = \begin{pmatrix} \mathbf{AX}_k \\ e^T \end{pmatrix}$$

Calculate

$$\begin{aligned} \mathbf{d}^k &= -[I - \mathbf{B}_k^T(\mathbf{B}_k\mathbf{B}_k^T)^{-1}\mathbf{B}_k]\mathbf{X}_k\mathbf{c} \\ \mathbf{y}^{k+1} &= \tfrac{e}{n} + \tfrac{\alpha}{n}\left(\tfrac{d^k}{\|d^k\|}\right) \\ \mathbf{x}^{k+1} &= \tfrac{X_k y^{k+1}}{e^T X_k y^{k+1}} \end{aligned}$$

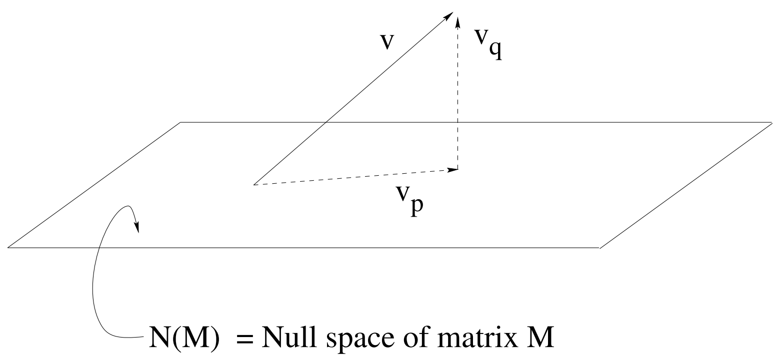
$k \longleftarrow k+1$ GO TO Step 2.

Note:

(1) To make sure the algorithm stops in $5nL$ steps. Karmarkar chose $\alpha = 1/3$ for theoretical proof.

(2) In practice, α can be chosen as large as $0.9, 0.99, 0.999, \ldots$

(3) Formula for projection: $v = v_p + v_q$



$$
\begin{aligned}
\mathcal{N}(M) &= \{ \ \mathbf{x} \mid M\mathbf{x} = 0 \ \} \\
v_p &= [\ I - M^T(MM^T)^{-1}M\ ]v \\
v_q &= M^T(MM^T)^{-1}Mv
\end{aligned}
$$

$$\underline{\text{Example}} \quad \begin{array}{lrl} \min & -x_1 + 1 & \\ \text{s.t.} & x_2 - x_3 & = 0 \\ & x_1 + x_2 + x_3 & = 1 \\ & x_1, x_2, x_3 & \geq 0 \end{array}$$

$$\mathbf{x}^0 = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} \quad X_0 = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & 1/3 \end{pmatrix}$$

$$\mathbf{A} = (0,\ 1,\ -1),\ \mathbf{AX}_0 = (0,\ 1/3,\ -1/3)$$

$$\mathbf{B}_0 = \begin{pmatrix} \mathbf{AX}_0 \\ e^T \end{pmatrix} = \begin{pmatrix} 0 & 1/3 & -1/3 \\ 1 & 1 & 1 \end{pmatrix},$$

$$\mathbf{c} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix}$$

$$\mathbf{d}^0 = -[I - \mathbf{B}_0^T(\mathbf{B}_0\mathbf{B}_0^T)^{-1}\mathbf{B}_0]\mathbf{X}_0\mathbf{c} = \begin{pmatrix} 2/9 \\ -1/9 \\ -1/9 \end{pmatrix}$$

$$\|\mathbf{d}^0\| = \sqrt{(2/9)^2 + (1/9)^2 + (1/9)^2} = \sqrt{6}/9$$

Note that

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix} = \begin{pmatrix} 2/3 \\ -1/3 \\ -1/3 \end{pmatrix} = 3\mathbf{d}^0$$

i.e. $\mathbf{d}^0$ points to the optimal solution!

Convergence in Polynomial-Time

Basic concept:

(1) To converge, we want

$$\mathbf{c}^T\mathbf{x}^k \searrow 0, \qquad k = 1, 2, \cdots$$

(2) To converge in polynomial-time,

$$\mathbf{c}^T\mathbf{x}^k \searrow 0 \text{ very fast!}$$

(3) For example, if

$$\mathbf{c}^T\mathbf{x}^k \leq e^{-k/5n}(\mathbf{c}^T\mathbf{x}^0), \quad k = 1, 2, \cdots$$

For $L > 0$ large enough s.t.

$$2^{-L}(\mathbf{c}^T\mathbf{x}^0) \approx 0$$

we want

$$e^{-k/5n}(\mathbf{c}^T\mathbf{x}^0) \leq 2^{-L}(\mathbf{c}^T\mathbf{x}^0) < \epsilon$$

or $e^{-k/5n} \leq 2^{-L}$

Then

$$-k/5n \leq log_e 2^{-L} < log_2 2^{-L} = -L$$

i.e., If $k > 5nL$, then

$$\mathbf{c}^T\mathbf{x}^k \leq e^{-k/5n}(\mathbf{c}^T\mathbf{x}^0) < \epsilon$$

Hence $\{\mathbf{c}^T\mathbf{x}^k\} \searrow 0$ in $O(nL)$ iterations.

(4) In Karmarkar's algorithm, the most time-consuming job is to find

$$(B_k B_k^T)^{-1}$$

A direct implementation takes $O(n^3)$ elementary operations. Therefore the total complexity is

$$O(n^4 L)$$

By rank-one updating, can be reduced to

$$O(n^{3.5} L)$$