

# Numerical Optimization

## Linear Programming - Interior Point Methods

**Shirish Shevade**

Computer Science and Automation  
Indian Institute of Science  
Bangalore 560 012, India.

NPTEL Course on Numerical Optimization

## Computational Complexity of Simplex Algorithm

Example<sup>1</sup>

$$\min \quad -2^{n-1}x_1 - 2^{n-2}x_2 - \dots - 2x_{n-1} - x_n$$

$$\text{s.t.} \quad x_1 \leq 5$$

$$4x_1 + x_2 \leq 25$$

$$8x_1 + 4x_2 + x_3 \leq 125$$

$$\vdots$$

$$2^n x_1 + 2^{n-1} x_2 + \dots + 4x_{n-1} + x_n \leq 5^n$$

$$x_j \geq 0 \quad \forall j = 1, \dots, n$$

Simplex method, starting at  $\mathbf{x} = \mathbf{0}$ , would visit all  $2^n$  extreme points before reaching the optimal solution.

---

<sup>1</sup>V. Klee and G.J. Minty, *How good is the simplex algorithm?*. In O. Shisha, editor, *Inequalities*, II, pp. 159-175, Academic Press, 1971

$$\min \quad -4x_1 - 2x_2 - x_3$$

$$\text{s.t.} \quad x_1 \leq 5$$

$$4x_1 + x_2 \leq 25$$

$$8x_1 + 4x_2 + x_3 \leq 125$$

$$x_1, x_2, x_3 \geq 0$$

Iteration	Basic Vectors	Objective function
1	$x_4, x_5, x_6$	0
2	$x_1, x_5, x_6$	-20
3	$x_1, x_2, x_6$	-30
4	$x_4, x_2, x_6$	-50
5	$x_4, x_2, x_3$	-75
6	$x_1, x_2, x_3$	-95
7	$x_1, x_5, x_3$	-105
8	$x_4, x_5, x_3$	<b>-125</b>

**Simplex Algorithm is not a polynomial time algorithm** (number of computational steps grows as an exponential function of the number of variables, rather than as a polynomial function)

## Interior Point Methods for Linear Programming

- Points generated are in the “interior” of the feasible region
- Based on nonlinear programming techniques
- Some interior points methods:
  - Affine Scaling
  - Karmarkar’s Method

We consider the linear program,

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

where  $\mathbf{A} \in \mathbb{R}^{m \times n}$  and  $\text{rank}(\mathbf{A}) = m$ .

## Affine Scaling

### Idea:

(1) Use *projected steepest descent direction* at every iteration

Given a feasible interior point  $\mathbf{x}^k$  at current iteration  $k$ .

$$A\mathbf{x}^k = \mathbf{b}, \mathbf{x}^k \geq \mathbf{0}$$

Let  $\mathbf{d}$  denote a direction such that  $\mathbf{x}^{k+1} = \mathbf{x}^k + \alpha^k \mathbf{d}$ ,  $\alpha^k > 0$ .  
Therefore,

$$A\mathbf{x}^{k+1} = \mathbf{b} \Rightarrow A\mathbf{d} = \mathbf{0}$$

### Projected Steepest Descent Direction

Project the steepest descent direction,  $-\mathbf{c}$ , on the *null space of A*

Let  $\hat{\mathbf{c}} = -\mathbf{c} = \mathbf{p}_c + \mathbf{q}$  where

- $\mathbf{p}_c \in \text{Null Space}(\mathbf{A})$ .  $\therefore \mathbf{A}\mathbf{p}_c = \mathbf{0}$ .
- $\mathbf{q} \in \text{Row Space}(\mathbf{A})$ .  $\therefore \mathbf{q} = \mathbf{A}^T \boldsymbol{\lambda}$ .

$$\mathbf{A}\hat{\mathbf{c}} = \mathbf{A}\mathbf{A}^T \boldsymbol{\lambda} \Rightarrow \boldsymbol{\lambda} = (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A}\hat{\mathbf{c}}$$

$$\begin{aligned} \mathbf{p}_c &= \hat{\mathbf{c}} - \mathbf{q} \\ &= \hat{\mathbf{c}} - \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A}\hat{\mathbf{c}} \\ &= (\mathbf{I} - \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A}) \hat{\mathbf{c}} \\ &= -(\mathbf{I} - \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A}) \mathbf{c} \\ &= -\mathbf{P}\mathbf{c} \end{aligned}$$

where  $\mathbf{P}$  denotes the *projection matrix*.

## Idea:

- (2) Position the current point close to the centre of the feasible region

For example, one possible choice is the point:

$$\mathbf{1} = (1, 1, \dots, 1)^T$$

Given a point  $\mathbf{x}^k$  in the interior of the feasible region, define  $X^k = \text{diag}(\mathbf{x}^k)$ .

Define the transformation,  $\mathbf{y} = T(\mathbf{x}) = X^{k-1}\mathbf{x}$ .

$$\therefore \mathbf{y}^k = X^{k-1}\mathbf{x}^k = \mathbf{1} \text{ or } X^k\mathbf{y}^k = \mathbf{x}^k$$

## Affine Scaling Algorithm:

- Start with any interior point  $\mathbf{x}^0$
- while (stopping condition is not satisfied at the current point)
  - Transform the current problem into an equivalent problem in  $\mathbf{y}$ -space so that the current point is close to the centre of the feasible region
  - Use projected steepest descent direction to take a step in the  $\mathbf{y}$ -space without crossing the feasible set boundary
  - Map the new point back to the corresponding point in the  $\mathbf{x}$ -space

endwhile

## Stopping Condition for an Affine Scaling Algorithm

Consider the following primal and dual problems:

### Primal Problem (P)

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

### Dual Problem (D)

$$\begin{aligned} \max \quad & \mathbf{b}^T \boldsymbol{\mu} \\ \text{s.t.} \quad & \mathbf{A}^T \boldsymbol{\mu} \leq \mathbf{c} \end{aligned}$$

For any primal and dual feasible  $\mathbf{x}$  and  $\boldsymbol{\mu}$ ,

$$\mathbf{c}^T \mathbf{x} \geq \mathbf{b}^T \boldsymbol{\mu} \quad (\text{Weak Duality})$$

At optimality,

$$\mathbf{c}^T \mathbf{x} - \mathbf{b}^T \boldsymbol{\mu} = 0 \quad (\text{Strong Duality})$$

**Idea:** Use the duality gap,  $\mathbf{c}^T \mathbf{x} - \mathbf{b}^T \boldsymbol{\mu}$ , to check optimality

Q. How to get  $\boldsymbol{\mu}$ ?

Consider the primal problem:

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

Define the Lagrangian function,

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{c}^T \mathbf{x} + \boldsymbol{\mu}^T (\mathbf{b} - \mathbf{Ax}) - \boldsymbol{\lambda}^T \mathbf{x}$$

Assumption:  $\mathbf{x}$  is primal feasible and  $\boldsymbol{\lambda} \geq \mathbf{0}$

KKT conditions at optimality:

$$\begin{aligned} \nabla_{\mathbf{x}} \mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = \mathbf{0} & \Rightarrow \mathbf{A}^T \boldsymbol{\mu} + \boldsymbol{\lambda} = \mathbf{c} \\ \lambda_i x_i & = 0 \quad \forall i = 1, \dots, n \end{aligned}$$

Defining  $X = \text{diag}(x)$ , the KKT conditions are

$$X(\mathbf{c} - \mathbf{A}^T \boldsymbol{\mu}) = \mathbf{0}.$$

Solve the following problem to get  $\boldsymbol{\mu}$ :

$$\min_{\boldsymbol{\mu}} \|X\mathbf{c} - X\mathbf{A}^T\boldsymbol{\mu}\|^2$$

$$\therefore \boldsymbol{\mu} = (\mathbf{A}X^2\mathbf{A}^T)^{-1}\mathbf{A}X^2\mathbf{c}$$

Thus, at a given point  $\mathbf{x}^k$ ,

$$\text{Duality Gap} = \mathbf{c}^T\mathbf{x}^k - \mathbf{b}^T\boldsymbol{\mu}^k$$

where  $\boldsymbol{\mu}^k = (\mathbf{A}X^{k2}\mathbf{A}^T)^{-1}\mathbf{A}X^{k2}\mathbf{c}$  and  $X^k = \text{diag}(\mathbf{x}^k)$ .

Step 1: Equivalent problem formulation to get considerable improvement in the objective function

Given  $\mathbf{x}^k$ , define  $X^k = \text{diag}(\mathbf{x}^k)$ .

Define a transformation  $T$  as,

$$\mathbf{y} = T(\mathbf{x}) = X^{k-1} \mathbf{x}$$

Therefore,

$$X^k \mathbf{y}^k = \mathbf{x}^k \quad \text{and} \quad \mathbf{y}^k = \mathbf{1}.$$

Using this transformation,

$$\left. \begin{array}{l} \min \quad \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0} \end{array} \right\} \equiv \begin{array}{l} \min \quad \mathbf{c}^T X^k \mathbf{y} \\ \text{s.t.} \quad \mathbf{A}X^k \mathbf{y} = \mathbf{b}, \mathbf{y} \geq \mathbf{0} \end{array}$$

which can be written in standard form as

$$\begin{array}{l} \min \quad \bar{\mathbf{c}}^T \mathbf{y} \\ \text{s.t.} \quad \bar{\mathbf{A}} \mathbf{y} = \mathbf{b}, \mathbf{y} \geq \mathbf{0} \end{array}$$

where  $\bar{\mathbf{c}} = X^k \mathbf{c}$  and  $\bar{\mathbf{A}} = \mathbf{A}X^k$ .

Step 2: Find the projected steepest direction and step length at  $\mathbf{y}^k$  for the problem,

$$\begin{aligned} \min \quad & \bar{\mathbf{c}}^T \mathbf{y} \\ \text{s.t.} \quad & \bar{\mathbf{A}} \mathbf{y} = \mathbf{b}, \mathbf{y} \geq \mathbf{0} \end{aligned}$$

Given  $\mathbf{x}^k$ ,  $\mathbf{y}^k = \mathbf{X}^{k-1} \mathbf{x}^k = \mathbf{1}$ .

The projected direction of  $-\bar{\mathbf{c}}$  on the null space of  $\bar{\mathbf{A}}$  is,

$$\therefore \mathbf{d}^k = -(\mathbf{I} - \mathbf{X}^k \mathbf{A}^T (\mathbf{A} \mathbf{X}^{k2} \mathbf{A}^T)^{-1} \mathbf{A} \mathbf{X}^k) \mathbf{X}^k \bar{\mathbf{c}}.$$

Let  $\alpha^k (> 0)$  denotes the step length.

$$\begin{aligned} \mathbf{y}^{k+1} &= \mathbf{y}^k + \alpha^k \mathbf{d}^k \\ &= \mathbf{1} + \alpha^k \mathbf{d}^k \geq \mathbf{0} \end{aligned}$$

$$\text{Let } \alpha_{max} = \min_{j: \mathbf{d}_j^k < 0} -\frac{1}{\mathbf{d}_j^k} \text{ and set } \alpha^k = .9 * \alpha_{max}.$$

Step 3:  $\mathbf{x}^{k+1} = \mathbf{X}^k \mathbf{y}^{k+1}$

---

## Affine Scaling Algorithm (to solve an LP in Standard Form)

---

- (1) Input:  $\mathbf{A}, \mathbf{b}, \mathbf{c}, \mathbf{x}^0, \epsilon$
- (2) Set  $k := 0$ .
- (3)  $\mathbf{X}^k = \text{diag}(\mathbf{x}^k)$
- (4)  $\boldsymbol{\mu}^k = (\mathbf{A}\mathbf{X}^{k2}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{X}^{k2}\mathbf{c}$
- (5) **while**  $(\mathbf{c}^T\mathbf{x}^k - \mathbf{b}^T\boldsymbol{\mu}^k) > \epsilon$ 
  - (a)  $\mathbf{d}^k = -(\mathbf{I} - \mathbf{X}^k\mathbf{A}^T(\mathbf{A}\mathbf{X}^{k2}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{X}^k)\mathbf{X}^k\mathbf{c}$
  - (b)  $\alpha^k = .9 * \min_{j:\mathbf{d}_j^k < 0} -\frac{1}{\mathbf{d}_j^k}$
  - (c)  $\mathbf{x}^{k+1} = \mathbf{X}^k(1 + \alpha^k\mathbf{d}^k)$
  - (d)  $\mathbf{X}^{k+1} = \text{diag}(\mathbf{x}^{k+1})$
  - (e)  $\boldsymbol{\mu}^{k+1} = (\mathbf{A}\mathbf{X}^{k+12}\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{X}^{k+12}\mathbf{c}$
  - (f)  $k := k + 1$

**endwhile**

**Output :**  $\mathbf{x}^* = \mathbf{x}^k$

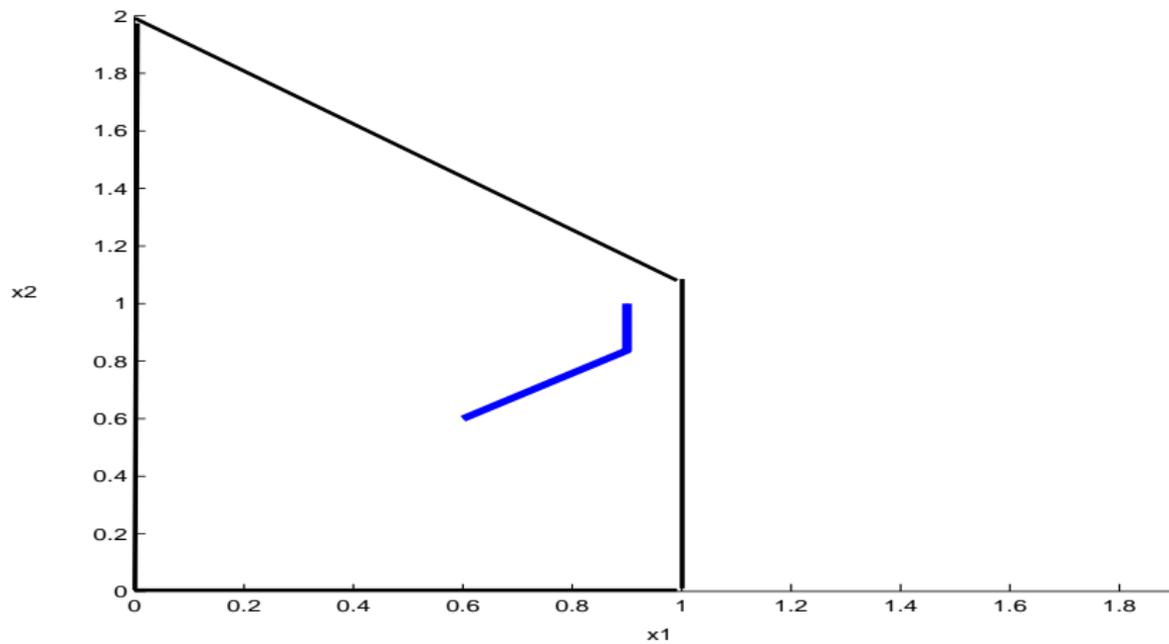
---

Application of Affine Scaling Algorithm to the problem,

$$\begin{aligned} \min \quad & -3x_1 - x_2 \\ \text{s.t.} \quad & x_1 + x_2 \leq 2 \\ & x_1 \leq 1 \\ & x_1, x_2 \geq 0 \end{aligned}$$

- $\mathbf{x}^* = (1, 1)^T$

Iteration	$\mathbf{x}^{kT}$	$\ \mathbf{x}^k - \mathbf{x}^*\ $
0	(.6, .6)	.57
1	(.87, .81)	.23
2	(.88, 1.00)	.12
3	(.90, .99)	.1026
4	(.90, 1.00)	.1001



## Karmarkar's Method

Assumptions:

- The problem is in *homogeneous* form:

$$\begin{aligned} \min \quad & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{0} \\ & \mathbf{1}^T \mathbf{x} = 1 \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

- Optimum objective function value is 0

Idea:

- (1) Use projective transformation to move an interior point to the centre of the feasible region
- (2) Move along projected steepest descent direction

## Projective Transformation:

Consider the transformation,  $T$ , defined by,

$$\mathbf{y} = T(\mathbf{x}) = \frac{X^{k-1}\mathbf{x}}{\mathbf{1}^T X^{k-1}\mathbf{x}}, \quad \mathbf{x} \neq \mathbf{0}$$

Remarks:

- $\mathbf{x} \mapsto \mathbf{y} \quad (\mathbf{y} = (\frac{1}{n}, \dots, \frac{1}{n})^T, \mathbf{1}^T \mathbf{y} = 1)$

- Inverse transformation:

$$\mathbf{x} = (\mathbf{1}^T X^{k-1}\mathbf{x})(X^k \mathbf{y}) \Rightarrow \mathbf{1}^T \mathbf{x} = (\mathbf{1}^T X^{k-1}\mathbf{x})\mathbf{1}^T (X^k \mathbf{y}).$$

$$\text{For every feasible } \mathbf{x}, \mathbf{1}^T \mathbf{x} = 1 \Rightarrow \mathbf{1}^T X^{k-1}\mathbf{x} = \frac{1}{\mathbf{1}^T X^k \mathbf{y}}$$

$$\therefore T^{-1}(\mathbf{y}) = \mathbf{x} = \frac{X^k \mathbf{y}}{\mathbf{1}^T X^k \mathbf{y}}$$

Using the transformation,

$$\mathbf{x} = \frac{X^k \mathbf{y}}{\mathbf{1}^T X^k \mathbf{y}}$$

$$\begin{array}{lll} \min \mathbf{c}^T \mathbf{x} & \min \frac{\mathbf{c}^T X^k \mathbf{y}}{\mathbf{1}^T X^k \mathbf{y}} & \min \mathbf{c}^T X^k \mathbf{y} \\ \text{s.t. } \mathbf{A} \mathbf{x} = \mathbf{0} & \text{s.t. } \mathbf{A} X^k \mathbf{y} = \mathbf{0} & \text{s.t. } \mathbf{A} X^k \mathbf{y} = \mathbf{0} \\ \mathbf{1}^T \mathbf{x} = 1 & \mathbf{1}^T \mathbf{y} = 1 & \mathbf{1}^T \mathbf{y} = 1 \\ \mathbf{x} \geq \mathbf{0} & \mathbf{y} \geq \mathbf{0} & \mathbf{y} \geq \mathbf{0} \end{array} \equiv$$

- Equivalence based on the assumption: Optimal objective function value is 0 and  $\mathbf{1}^T X^k \mathbf{y} > 0$

Consider the problem,

$$\begin{aligned} \min \quad & \mathbf{c}^T X^k \mathbf{y} \\ \text{s.t.} \quad & \mathbf{A} X^k \mathbf{y} = \mathbf{0} \\ & \mathbf{1}^T \mathbf{y} = 1 \\ & \mathbf{y} \geq \mathbf{0} \end{aligned}$$

- Step 1: Find a projected steepest descent direction in the  $\mathbf{y}$ -space ( Projection of  $-X^k \mathbf{c}$  onto the subspace of  $\{\mathbf{d} : \mathbf{A} X^k \mathbf{d} = \mathbf{0}, \mathbf{1}^T \mathbf{d} = 0, \mathbf{d} \geq \mathbf{0}\}$ )

Find  $\mathbf{d}$  by solving

$$\begin{aligned} \min \quad & \frac{1}{2} \|X^k \mathbf{c} - \mathbf{d}\|^2 \\ \text{s.t.} \quad & \mathbf{A} X^k \mathbf{d} = \mathbf{0}, \end{aligned}$$

projecting it onto the null space of  $\mathbf{1}^T$  and ensure  $\mathbf{d} \geq \mathbf{0}$ .

Consider the problem,

$$\begin{aligned} \min \quad & \frac{1}{2} \|X^k \mathbf{c} - \mathbf{d}\|^2 \\ \text{s.t.} \quad & \mathbf{A}X^k \mathbf{d} = \mathbf{0} \end{aligned}$$

$$\mathcal{L}(\mathbf{d}, \boldsymbol{\mu}) = \frac{1}{2} \|X^k \mathbf{c} - \mathbf{d}\|^2 + \boldsymbol{\mu}^T \mathbf{A}X^k \mathbf{d}$$

$$\begin{aligned} \nabla_{\mathbf{d}} \mathcal{L}(\mathbf{d}, \boldsymbol{\mu}) = \mathbf{0} \Rightarrow & -(X^k \mathbf{c} - \mathbf{d}) + X^k A^T \boldsymbol{\mu} = \mathbf{0} \\ \Rightarrow & \mathbf{d} = X^k \mathbf{c} - X^k A^T \boldsymbol{\mu} \end{aligned}$$

$$\mathbf{0} = \mathbf{A}X^k \mathbf{d} = \mathbf{A}X^{k^2} \mathbf{c} - \mathbf{A}X^{k^2} A^T \boldsymbol{\mu} \Rightarrow \mathbf{A}X^{k^2} \mathbf{c} = \mathbf{A}X^{k^2} A^T \boldsymbol{\mu}$$

Projection of  $-\mathbf{d}$  on the null space of  $\mathbf{1}^T$ :

$$\mathbf{d}^k = -\left(\mathbf{I} - \frac{1}{n} \mathbf{1}\mathbf{1}^T\right)(X^k \mathbf{c} - X^k A^T \boldsymbol{\mu})$$

---

## Karmarkar's Projective Scaling Algorithm (for homogeneous Linear Program)

---

- (1) Input: Homogeneous LP,  $A, c, \epsilon$
  - (2) Set  $k := 0, \mathbf{x}^k = \frac{1}{n}\mathbf{1}$
  - (3)  $X^k = \text{diag}(\mathbf{x}^k)$
  - (4) **while**  $c^T \mathbf{x}^k > \epsilon$ 
    - (a) Find the projected steepest descent direction  $\mathbf{d}^k$
    - (b)  $\mathbf{y}^{k+1} = \frac{1}{n}\mathbf{1} + \frac{\delta}{\sqrt{n(n-1)} \|\mathbf{d}^k\|} \mathbf{d}^k \quad (\delta = \frac{1}{3})$
    - (c)  $\mathbf{x}^{k+1} = T^{-1}(\mathbf{y}^{k+1})$
    - (d)  $X^{k+1} = \text{diag}(\mathbf{x}^{k+1})$
    - (e)  $k := k + 1$
- endwhile**

**Output :**  $\mathbf{x}^* = \mathbf{x}^k$

---