CS 473: Algorithms

Ruta Mehta

University of Illinois, Urbana-Champaign

Spring 2018

Introduction to Linear Programming

Lecture 18

March 26, 2018

Some of the slides are courtesy Prof. Chekuri

Part I

Introduction to Linear Programming

Today...

Linear Programming and Standard Formulation

Geometry

Vertex Solution

Simplex Method

A Factory Example

Problem

Your factory can produce Laptop and iPhone using Copper.

1. One ton of Copper $\longrightarrow$ one Laptop
2. One ton of Copper $\longrightarrow$ one iPhone
3. We have 200 tons of Copper.
4. Laptop can be sold for **\$1** and iPhone for **\$6**.

A Factory Example

Problem

Your factory can produce Laptop and iPhone using Copper.

1. One ton of Copper $\longrightarrow$ one Laptop
2. One ton of Copper $\longrightarrow$ one iPhone
3. We have 200 tons of Copper.
4. Laptop can be sold for **$1** and iPhone for **$6**.

How many units of Laptop and iPhone should your factory manufacture to maximize profit?

A Factory Example

Problem

Your factory can produce Laptop and iPhone using Copper.

1. One ton of Copper $\longrightarrow$ one Laptop
2. One ton of Copper $\longrightarrow$ one iPhone
3. We have 200 tons of Copper.
4. Laptop can be sold for **\$1** and iPhone for **\$6**.

How many units of Laptop and iPhone should your factory manufacture to maximize profit?

Solution:

A Factory Example

Problem

Your factory can produce Laptop and iPhone using Copper.

1. One ton of Copper $\longrightarrow$ one Laptop
2. One ton of Copper $\longrightarrow$ one iPhone
3. We have 200 tons of Copper.
4. Laptop can be sold for **\$1** and iPhone for **\$6**.

How many units of Laptop and iPhone should your factory manufacture to maximize profit?

Solution: manufacture only iPhone

A Factory Example

Problem

Your factory can produce Laptop and iPhone using resources C, B, A.

1. One unit of A and C each $\longrightarrow$ One Laptop
2. One unit of B and C each $\longrightarrow$ One iPhone

A Factory Example

Problem

Your factory can produce Laptop and iPhone using resources C, B, A.

1. One unit of A and C each $\rightarrow$ One Laptop
2. One unit of B and C each $\rightarrow$ One iPhone
3. We have 200 units of A, 300 units of B, and 400 units of C.
4. Product Laptop can be sold for **\$1** and product iPhone for **\$6**.

A Factory Example

Problem

Your factory can produce Laptop and iPhone using resources C, B, A.

1. One unit of A and C each $\longrightarrow$ One Laptop
2. One unit of B and C each $\longrightarrow$ One iPhone
3. We have 200 units of A, 300 units of B, and 400 units of C.
4. Product Laptop can be sold for **\$1** and product iPhone for **\$6**.

How many units of Laptop and iPhone should your factory manufacture to maximize profit?

Solution:

A Factory Example

Problem

Your factory can produce Laptop and iPhone using resources C, B, A.

1. One unit of A and C each $\longrightarrow$ One Laptop
2. One unit of B and C each $\longrightarrow$ One iPhone
3. We have 200 units of A, 300 units of B, and 400 units of C.
4. Product Laptop can be sold for **\$1** and product iPhone for **\$6**.

How many units of Laptop and iPhone should your factory manufacture to maximize profit?

Solution: Formulate as a linear program.

A Factory Example

Problem

Can produce Laptop and iPhone, using resources A, B, C.

1. $A, C \rightarrow$ Laptop
2. $B, C \rightarrow$ iPhone
3. Have A: 200, B: 300 , and C: 400.
4. Price of L: **\$1**, and iP: **\$6**.

How many units to manufacture to max profit?

A Factory Example

Problem

Can produce Laptop and iPhone, using resources A, B, C.

1. $A, C \to$ Laptop
2. $B, C \to$ iPhone
3. Have A: 200, B: 300 , and C: 400.
4. Price of L: **\$1**, and iP: **\$6**.

How many units to manufacture to max profit?

Suppose x_1 units of Laptop and x_2 units of iPhone.

A Factory Example

Problem

Can produce Laptop and iPhone, using resources A, B, C.

1. $A, C \rightarrow$ Laptop
2. $B, C \rightarrow$ iPhone
3. Have A: 200, B: 300 , and C: 400.
4. Price of L: **\$1**, and iP: **\$6**.

How many units to manufacture to max profit?

Suppose x_1 units of Laptop and x_2 units of iPhone.

$$\begin{array}{lll} \textbf{max} & x_1 + 6x_2 & \\ \text{s.t.} & x_1 \leq 200 & (A) \\ & x_2 \leq 300 & (B) \\ & x_1 + x_2 \leq 400 & (C) \\ & x_1 \geq 0 & \\ & x_2 \geq 0 & \end{array}$$

Linear Programming Formulation

Let us produce x_1 units of Laptop and x_2 units of iPhone. Our profit can be computed by solving

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

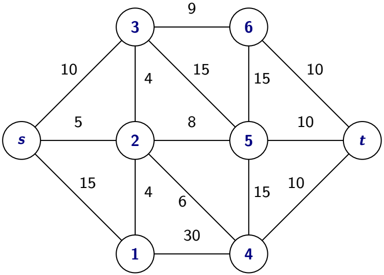
Linear Programming Formulation

Let us produce x_1 units of Laptop and x_2 units of iPhone. Our profit can be computed by solving

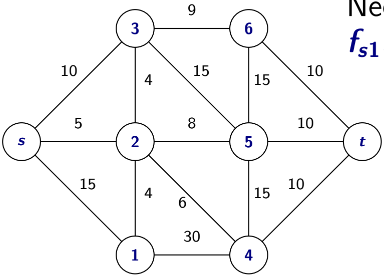
$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

What is the solution?

Maximum Flow in Network

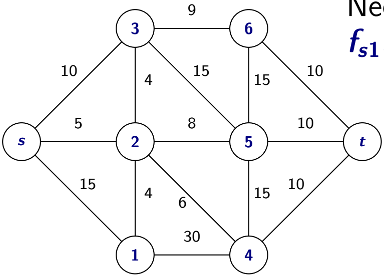


Maximum Flow in Network



Need to compute values $f_{s1}, f_{s2}, \ldots f_{25}, \ldots f_{5t}, f_{6t}$ such that

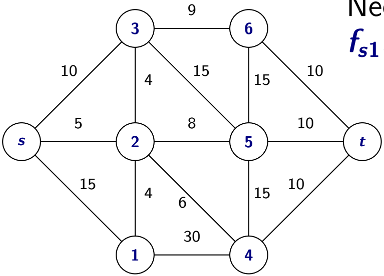
Maximum Flow in Network



Need to compute values $f_{s1}, f_{s2}, \ldots f_{25}, \ldots f_{5t}, f_{6t}$ such that

$$\begin{array}{lll} f_{s1} \leq 15 & f_{s2} \leq 5 & f_{s3} \leq 10 \\ f_{14} \leq 30 & f_{21} \leq 4 & f_{25} \leq 8 \\ f_{32} \leq 4 & f_{35} \leq 15 & f_{36} \leq 9 \\ f_{42} \leq 6 & f_{4t} \leq 10 & f_{54} \leq 15 \\ f_{5t} \leq 10 & f_{65} \leq 15 & f_{6t} \leq 10 \end{array}$$

Maximum Flow in Network



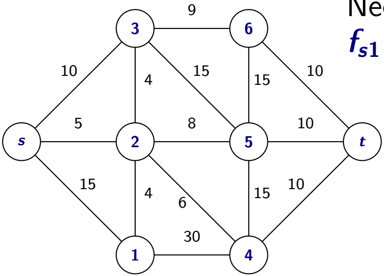
Need to compute values $f_{s1}, f_{s2}, \ldots f_{25}, \ldots f_{5t}, f_{6t}$ such that

$$
\begin{array}{lll}
f_{s1} \leq 15 & f_{s2} \leq 5 & f_{s3} \leq 10 \\
f_{14} \leq 30 & f_{21} \leq 4 & f_{25} \leq 8 \\
f_{32} \leq 4 & f_{35} \leq 15 & f_{36} \leq 9 \\
f_{42} \leq 6 & f_{4t} \leq 10 & f_{54} \leq 15 \\
f_{5t} \leq 10 & f_{65} \leq 15 & f_{6t} \leq 10
\end{array}
$$

and

$$
\begin{array}{lll}
f_{s1} + f_{21} = f_{14} & f_{s2} + f_{32} = f_{21} + f_{25} & f_{s3} = f_{32} + f_{35} + f_{36} \\
f_{14} + f_{54} = f_{42} + f_{4t} & f_{25} + f_{35} + f_{65} = f_{54} + f_{5t} & f_{36} = f_{65} + f_{6t}
\end{array}
$$

Maximum Flow in Network



Need to compute values $f_{s1}, f_{s2}, \ldots f_{25}, \ldots f_{5t}, f_{6t}$ such that

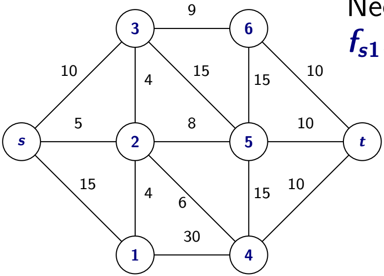
$$
\begin{array}{lll}
f_{s1} \le 15 & f_{s2} \le 5 & f_{s3} \le 10 \\
f_{14} \le 30 & f_{21} \le 4 & f_{25} \le 8 \\
f_{32} \le 4 & f_{35} \le 15 & f_{36} \le 9 \\
f_{42} \le 6 & f_{4t} \le 10 & f_{54} \le 15 \\
f_{5t} \le 10 & f_{65} \le 15 & f_{6t} \le 10
\end{array}
$$

and

$$
\begin{array}{lll}
f_{s1} + f_{21} = f_{14} & f_{s2} + f_{32} = f_{21} + f_{25} & f_{s3} = f_{32} + f_{35} + f_{36} \\
f_{14} + f_{54} = f_{42} + f_{4t} & f_{25} + f_{35} + f_{65} = f_{54} + f_{5t} & f_{36} = f_{65} + f_{6t}
\end{array}
$$

$$
f_{s1} \ge 0 \quad f_{s2} \ge 0 \quad f_{s3} \ge 0 \quad \cdots \quad f_{4t} \ge 0 \quad f_{5t} \ge 0 \quad f_{6t} \ge 0
$$

Maximum Flow in Network



Need to compute values $f_{s1}, f_{s2}, \ldots f_{25}, \ldots f_{5t}, f_{6t}$ such that

$$
\begin{array}{lll}
f_{s1} \le 15 & f_{s2} \le 5 & f_{s3} \le 10 \\
f_{14} \le 30 & f_{21} \le 4 & f_{25} \le 8 \\
f_{32} \le 4 & f_{35} \le 15 & f_{36} \le 9 \\
f_{42} \le 6 & f_{4t} \le 10 & f_{54} \le 15 \\
f_{5t} \le 10 & f_{65} \le 15 & f_{6t} \le 10
\end{array}
$$

and

$$
\begin{array}{lll}
f_{s1} + f_{21} = f_{14} & f_{s2} + f_{32} = f_{21} + f_{25} & f_{s3} = f_{32} + f_{35} + f_{36} \\
f_{14} + f_{54} = f_{42} + f_{4t} & f_{25} + f_{35} + f_{65} = f_{54} + f_{5t} & f_{36} = f_{65} + f_{6t}
\end{array}
$$

$$
f_{s1} \ge 0 \quad f_{s2} \ge 0 \quad f_{s3} \ge 0 \quad \cdots \quad f_{4t} \ge 0 \quad f_{5t} \ge 0 \quad f_{6t} \ge 0
$$

maximize: $f_{s1} + f_{s2} + f_{s3}$.

Maximum Flow as a Linear Program

For a general flow network $G = (V, E)$ with capacities c_e on edge $e \in E$, we have variables f_e indicating flow on edge e

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{e \text{ out of } s} f_e & \\
\text{subject to} \quad & f_e \leq c_e & \text{for each } e \in E \\
& \sum_{e \text{ out of } v} f_e - \sum_{e \text{ into } v} f_e = 0 & \forall v \in V \setminus \{s, t\} \\
& f_e \geq 0 & \text{for each } e \in E.
\end{aligned}
$$

Maximum Flow as a Linear Program

For a general flow network $G = (V, E)$ with capacities c_e on edge $e \in E$, we have variables f_e indicating flow on edge e

$$
\begin{aligned}
\text{Maximize} \quad & \sum_{e \text{ out of } s} f_e & \\
\text{subject to} \quad & f_e \leq c_e & \text{for each } e \in E \\
& \sum_{e \text{ out of } v} f_e - \sum_{e \text{ into } v} f_e = 0 & \forall v \in V \setminus \{s, t\} \\
& f_e \geq 0 & \text{for each } e \in E.
\end{aligned}
$$

Number of variables: m, one for each edge.
Number of constraints: $m + n - 2 + m$.

Minimum Cost Flow with Lower Bounds

... as a Linear Program

For a general flow network $G = (V, E)$ with capacities c_e, lower bounds ℓ_e, and costs w_e, we have variables f_e indicating flow on edge e. Suppose we want a min-cost flow of value at least F.

$$\text{Minimize} \sum_{e \in E} w_e f_e$$

$$\text{subject to} \sum_{e \text{ out of } s} f_e \geq F$$

$$f_e \leq c_e \qquad f_e \geq \ell_e \qquad \text{for each } e \in E$$

$$\sum_{e \text{ out of } v} f_e - \sum_{e \text{ into } v} f_e = 0 \quad \text{for each } v \in V - \{s, t\}$$

$$f_e \geq 0 \qquad \text{for each } e \in E.$$

Minimum Cost Flow with Lower Bounds

... as a Linear Program

For a general flow network $G = (V, E)$ with capacities c_e, lower bounds ℓ_e, and costs w_e, we have variables f_e indicating flow on edge e. Suppose we want a min-cost flow of value at least F.

$$
\begin{aligned}
\text{Minimize} \quad & \sum_{e \in E} w_e f_e \\
\text{subject to} \quad & \sum_{e \text{ out of } s} f_e \geq F \\
& f_e \leq c_e \qquad f_e \geq \ell_e && \text{for each } e \in E \\
& \sum_{e \text{ out of } v} f_e - \sum_{e \text{ into } v} f_e = 0 && \text{for each } v \in V - \{s, t\} \\
& f_e \geq 0 && \text{for each } e \in E.
\end{aligned}
$$

Number of variables: m, one for each edge

Number of constraints: $1 + m + m + n - 2 + m = 3m + n - 1$.

Linear Programs

Problem

Find a vector $x \in \mathbb{R}^d$ that

$$
\begin{array}{lll}
\text{maximize/minimize} & \sum_{j=1}^{d} c_j x_j & \\
\text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i & \text{for } i = 1 \ldots p \\
 & \sum_{j=1}^{d} a_{ij} x_j = b_i & \text{for } i = p+1 \ldots q \\
 & \sum_{j=1}^{d} a_{ij} x_j \geq b_i & \text{for } i = q+1 \ldots n
\end{array}
$$

Linear Programs

Problem

Find a vector $x \in \mathbb{R}^d$ that

$$\begin{array}{lll} \text{maximize/minimize} & \sum_{j=1}^{d} c_j x_j & \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i & \text{for } i = 1 \ldots p \\ & \sum_{j=1}^{d} a_{ij} x_j = b_i & \text{for } i = p+1 \ldots q \\ & \sum_{j=1}^{d} a_{ij} x_j \geq b_i & \text{for } i = q+1 \ldots n \end{array}$$

Input is matrix $A = (a_{ij}) \in \mathbb{R}^{n \times d}$, column vector $b = (b_i) \in \mathbb{R}^n$, and row vector $c = (c_j) \in \mathbb{R}^d$

Canonical Form of Linear Programs

Canonical Form

A linear program is in canonical form if it has the following structure

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \quad \text{for } i = 1 \ldots n \end{array}$$

Canonical Form of Linear Programs

Canonical Form

A linear program is in canonical form if it has the following structure

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \quad \text{for } i = 1 \ldots n \end{array}$$

Conversion to Canonical Form

1. Replace $\sum_j a_{ij} x_j = b_i$ by

$$\sum_j a_{ij} x_j \leq b_i \quad \text{and} \quad -\sum_j a_{ij} x_j \leq -b_i$$

2. Replace $\sum_j a_{ij} x_j \geq b_i$ by $-\sum_j a_{ij} x_j \leq -b_i$

Matrix Representation of Linear Programs

A linear program in canonical form can be written as

$$\begin{aligned} \text{maximize} \quad & \mathbf{c} \cdot \mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned}$$

where $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{n \times d}$, column vector $\mathbf{b} = (b_i) \in \mathbb{R}^n$, row vector $\mathbf{c} = (c_j) \in \mathbb{R}^d$, and column vector $\mathbf{x} = (x_j) \in \mathbb{R}^d$

1. Number of variable is d
2. Number of constraints is n

Other Standard Forms for Linear Programs

$$\begin{aligned} \text{maximize} \quad & \boldsymbol{c} \cdot \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq \mathbf{0} \end{aligned}$$

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c} \cdot \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} \geq \boldsymbol{b} \\ & \boldsymbol{x} \geq \mathbf{0} \end{aligned}$$

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c} \cdot \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq \mathbf{0} \end{aligned}$$

Linear Programming: A History

1. First formal application to problems in economics by Leonid Kantorovich in the 1930s
    1. However, work was ignored behind the Iron Curtain and unknown in the West

Linear Programming: A History

1. First formal application to problems in economics by Leonid Kantorovich in the 1930s
   1. However, work was ignored behind the Iron Curtain and unknown in the West
2. Rediscovered by Tjalling Koopmans in the 1940s, along with applications to economics

Linear Programming: A History

1. First formal application to problems in economics by Leonid Kantorovich in the 1930s
    1. However, work was ignored behind the Iron Curtain and unknown in the West
2. Rediscovered by Tjalling Koopmans in the 1940s, along with applications to economics
3. First algorithm (Simplex) to solve linear programs by George Dantzig in 1947

Linear Programming: A History

1. First formal application to problems in economics by Leonid Kantorovich in the 1930s
   1. However, work was ignored behind the Iron Curtain and unknown in the West
2. Rediscovered by Tjalling Koopmans in the 1940s, along with applications to economics
3. First algorithm (Simplex) to solve linear programs by George Dantzig in 1947
4. Kantorovich and Koopmans receive Nobel Prize for economics in 1975 ; Dantzig, however, was ignored
   1. Koopmans contemplated refusing the Nobel Prize to protest Dantzig's exclusion, but Kantorovich saw it as a vindication for using mathematics in economics, which had been written off as "a means for apologists of capitalism"

Back to the Factory example

Produce x_1 units of product 1 and x_2 units of product 2. Our profit can be computed by solving

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

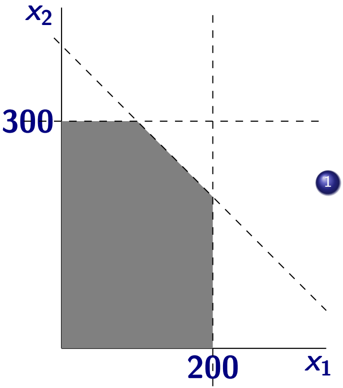
Back to the Factory example

Produce x_1 units of product 1 and x_2 units of product 2. Our profit can be computed by solving

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

What is the solution?

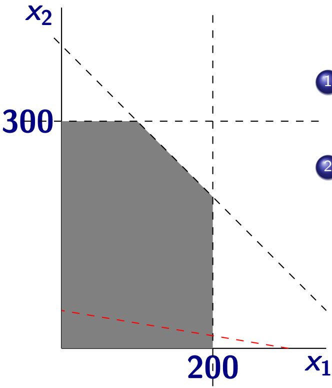
Solving the Factory Example



1. Feasible values of x_1 and x_2 are shaded region.

$$
\begin{array}{lccc}
\text{maximize} & & x_1 + 6x_2 & \\
\text{subject to} & x_1 \leq 200 & x_2 \leq 300 & x_1 + x_2 \leq 400 \\
& & x_1, x_2 \geq 0 &
\end{array}
$$

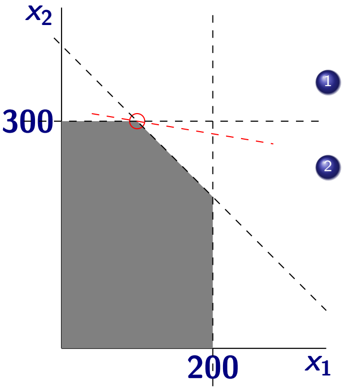
Solving the Factory Example



1. Feasible values of x_1 and x_2 are shaded region.
2. Objective (Cost) function is a direction — the line represents all points with same value of the function

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

Solving the Factory Example



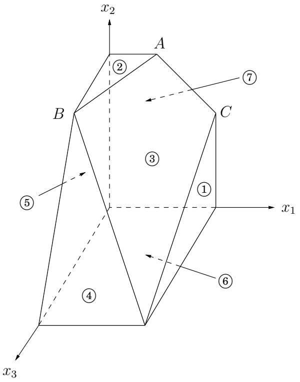
1. Feasible values of x_1 and x_2 are shaded region.
2. Objective (Cost) function is a direction — the line represents all points with same value of the function; moving the line until it just leaves the feasible region, gives optimal values.

$$\begin{array}{llll} \text{maximize} & & x_1 + 6x_2 & \\ \text{subject to} & x_1 \leq 200 & x_2 \leq 300 & x_1 + x_2 \leq 400 \\ & & x_1, x_2 \geq 0 & \end{array}$$

Linear Programming in 2-d

1. Each constraint a half plane
2. Feasible region is intersection of finitely many half planes — it forms a polygon.
3. For a fixed value of objective function, we get a line. Parallel lines correspond to different values for objective function.
4. Optimum achieved when objective function line just leaves the feasible region

An Example in 3-d



$$
\begin{aligned}
\max \quad x_1 + 6x_2 + 13x_3 & \\
x_1 \le 200 & \quad ① \\
x_2 \le 300 & \quad ② \\
x_1 + x_2 + x_3 \le 400 & \quad ③ \\
x_2 + 3x_3 \le 600 & \quad ④ \\
x_1 \ge 0 & \quad ⑤ \\
x_2 \ge 0 & \quad ⑥ \\
x_3 \ge 0 & \quad ⑦
\end{aligned}
$$

Polytope

Figure from Dasgupta etal book.

Part II

Simple Algorithm

Factory Example: Alternate View

Original Problem

Recall we have,

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

Factory Example: Alternate View

Original Problem

Recall we have,

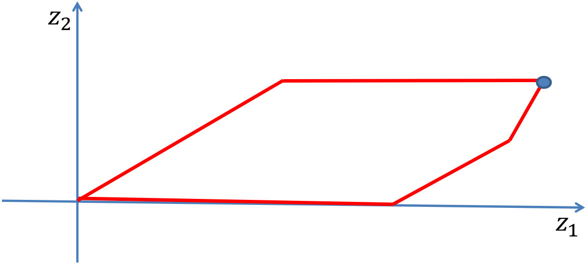
$$\begin{array}{lc} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 200 \quad x_2 \leq 300 \quad x_1 + x_2 \leq 400 \\ & x_1, x_2 \geq 0 \end{array}$$

Transformation

Consider new variable z_1 and z_2, such that $z_1 = x_1 + 6x_2$ and $z_2 = x_2$. Then $x_1 = z_1 - 6z_2$. In terms of the new variables we have

$$\begin{array}{lc} \text{maximize} & z_1 \\ \text{subject to} & z_1 - 6z_2 \leq 200 \quad z_2 \leq 300 \quad z_1 - 5z_2 \leq 400 \\ & z_1 - 6z_2 \geq 0 \qquad z_2 \geq 0 \end{array}$$

Transformed Picture



Feasible region rotated, and optimal value at the right-most point on polygon

Observations about the Transformation

Observations

1. Linear program can always be transformed to get a linear program where the optimal value is achieved at the point in the feasible region with highest x-coordinate
2. Optimum value attained at a vertex of the polygon
3. Since feasible region is convex, and objective function linear, every local optimum is a global optimum

A Simple Algorithm in 2-d

1. optimum solution is at a vertex of the feasible region
2. a vertex is defined by the intersection of two lines (constraints)

A Simple Algorithm in 2-d

1. optimum solution is at a vertex of the feasible region
2. a vertex is defined by the intersection of two lines (constraints)

Algorithm:

1. find all intersections between the n lines — at most n^2 points
2. for each intersection point $p = (p_1, p_2)$
    1. check if p is in feasible region (how?)
    2. if p is feasible evaluate objective function at p:
       $val(p) = c_1 p_1 + c_2 p_2$
3. Output the feasible point with the largest value

A Simple Algorithm in 2-d

1. optimum solution is at a vertex of the feasible region
2. a vertex is defined by the intersection of two lines (constraints)

Algorithm:

1. find all intersections between the n lines — at most n^2 points
2. for each intersection point $p = (p_1, p_2)$
   1. check if p is in feasible region (how?)
   2. if p is feasible evaluate objective function at p:
      $val(p) = c_1 p_1 + c_2 p_2$
3. Output the feasible point with the largest value

Running time: $O(n^3)$.

Geometry in d-dimentsion

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \quad \text{for } i = 1 \dots n \end{array}$$

Q: The set of points defined by a linear constraint

$$\{x \in \mathbb{R}^d \mid \sum_{j=1}^{d} a_{ij} x_j \leq b_i\} \text{ is,}$$

1. convex
2. non-convex

Geometry in d-dimentsion

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \quad \text{for } i = 1 \ldots n \end{array}$$

Q: The set of points defined by a linear constraint

$$\{x \in \mathbb{R}^d \mid \sum_{j=1}^{d} a_{ij} x_j \leq b_i\} \text{ is,}$$

1. convex
2. non-convex

This is also called a **halfspace**.

Geometry in d-dimentsion

$$
\begin{array}{ll}
\text{maximize} & \sum_{j=1}^{d} c_j x_j \\
\text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\
& \quad \text{for } i = 1 \ldots n
\end{array}
$$

Q: Intersection of a finitely many convex sets is,

1. convex
2. non-convex

Geometry in d-dimentsion

$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \quad \text{for } i = 1 \ldots n \end{array}$$

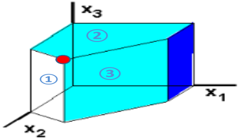
Q: Intersection of a finitely many convex sets is,

1. convex
2. non-convex

Thus feasible set, $\{x \mid \sum_{j=1}^{d} a_{ij} x_j \leq b_i \text{ for } i = 1 \ldots n\}$, is convex. Defines a polytope.

Geometry in d-dimentsion

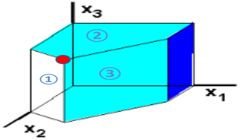
$$\begin{array}{ll} \text{maximize} & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \quad \text{for } i = 1 \ldots n \end{array}$$



Caratheodory Theorem. Every point x in a d-dimensional polytope can be written as a *convex combination* of $(d+1)$ vertices.

Geometry in d-dimentsion

maximize $\sum_{j=1}^{d} c_j x_j$
subject to $\sum_{j=1}^{d} a_{ij} x_j \leq b_i$
for $i = 1 \dots n$



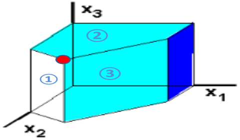
Caratheodory Theorem. Every point x in a d-dimensional polytope can be written as a *convex combination* of $(d+1)$ vertices.

Q: If x is a convex combination of vertices $v_1, \dots, v_k$, then for a constant vector c which of the following holds

1. $(c \cdot x) \geq \max_{i=1}^{k} (c \cdot v_i)$
2. $(c \cdot x) \leq \max_{i=1}^{k} (c \cdot v_i)$

Geometry in d-dimentsion

maximize $\sum_{j=1}^{d} c_j x_j$
subject to $\sum_{j=1}^{d} a_{ij} x_j \leq b_i$
for $i = 1 \ldots n$



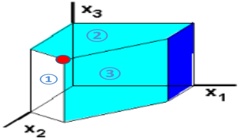
Caratheodory Theorem. Every point x in a d-dimensional polytope can be written as a *convex combination* of $(d+1)$ vertices.

Q: If x is a convex combination of vertices $v_1, \ldots, v_k$, then for a constant vector c which of the following holds

1. $(c \cdot x) \geq \max_{i=1}^{k} (c \cdot v_i)$
2. $(c \cdot x) \leq \max_{i=1}^{k} (c \cdot v_i)$

There exists a vertex solution.

maximize $\sum_{j=1}^{d} c_j x_j$
subject to $\sum_{j=1}^{d} a_{ij} x_j \leq b_i$
for $i = 1 \ldots n$



Caratheodory Theorem. Every point x in a d-dimensional polytope can be written as a *convex combination* of $(d+1)$ vertices.

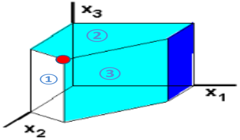
If x is a convex combination of vertices $v_1, \ldots, v_k$, then

1. $\min_{i=1}^{k}(c \cdot v_i) \leq (c \cdot x) \leq \max_{i=1}^{k}(c \cdot v_i)$

There exists a vertex solution.

Linear Programming in d-dimensions (Summary)

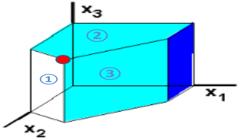
$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} \quad & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \text{for } i = 1 \dots n \end{aligned}$$



1. Each linear constraint defines a **halfspace** – Convex set.

Linear Programming in d-dimensions (Summary)

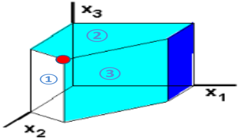
$$\begin{aligned} &\text{maximize} && \sum_{j=1}^{d} c_j x_j \\ &\text{subject to} && \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & && \text{for } i = 1 \dots n \end{aligned}$$



1. Each linear constraint defines a **halfspace** – Convex set.
2. Feasible region is an intersection of halfspaces – Convex **polytope**.

Linear Programming in d-dimensions (Summary)

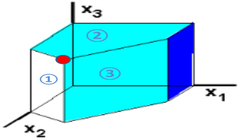
$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} \quad & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \text{for } i = 1 \ldots n \end{aligned}$$



1. Each linear constraint defines a **halfspace** – Convex set.
2. Feasible region is an intersection of halfspaces – Convex **polytope**.
3. Optimal value attained at a vertex of the polyhedron.
    - Using the Caratheodory Theorem. (Or the transformation)

Linear Programming in d-dimensions (Summary)

$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{d} c_j x_j \\ \text{subject to} \quad & \sum_{j=1}^{d} a_{ij} x_j \leq b_i \\ & \text{for } i = 1 \ldots n \end{aligned}$$



1. Each linear constraint defines a **halfspace** – Convex set.
2. Feasible region is an intersection of halfspaces – Convex **polytope**.
3. Optimal value attained at a vertex of the polyhedron.
    - Using the Caratheodory Theorem. (Or the transformation)
4. Tight inequality $\sum_{j=1}^{d} a_{ij} x_j = b_i$ defines hyperplane of $(d-1)$ dim.
5. A vertex is defined by intersection of d hyperplanes.
    - Solution of $\hat{A}x = \hat{b}$, where $\hat{A}$ is $d \times d$.
    - $\hat{A}$ has non-zero determinant – linear independence.

Simple Algorithm in d Dimensions

Real problem: d-dimensions, n-constraints

Simple Algorithm in d Dimensions

Real problem: d-dimensions, n-constraints

1. optimum solution is at a vertex of the feasible region
2. a vertex is defined by the intersection of d hyperplanes
3. number of vertices can be

Simple Algorithm in d Dimensions

Real problem: d-dimensions, n-constraints

1. optimum solution is at a vertex of the feasible region
2. a vertex is defined by the intersection of d hyperplanes
3. number of vertices can be $\Omega(n^d)$

Running time: $O(dn^{d+1})$ which is not polynomial since problem size is at least nd. Also not practical.

How do we find the intersection point of d hyperplanes in $\mathbb{R}^d$?

Simple Algorithm in d Dimensions

Real problem: d-dimensions, n-constraints

1. optimum solution is at a vertex of the feasible region
2. a vertex is defined by the intersection of d hyperplanes
3. number of vertices can be $\Omega(n^d)$

Running time: $O(dn^{d+1})$ which is not polynomial since problem size is at least nd. Also not practical.

How do we find the intersection point of d hyperplanes in $\mathbb{R}^d$? Using Gaussian elimination to solve $\hat{A}x = \hat{b}$ where $\hat{A}$ is a $d \times d$ matrix and $\hat{b}$ is a $d \times 1$ matrix.

Simplex Algorithm

Simplex: Vertex hoping algorithm

Simplex Algorithm

Simplex: Vertex hoping algorithm

Moves from a vertex to its neighboring vertex

Simplex Algorithm

Simplex: Vertex hoping algorithm

Moves from a vertex to its neighboring vertex

Questions

- Which neighbor to move to?
- When to stop?
- How much time does it take?

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x_1^*, \ldots, x_d^*)$, then $c \cdot x^* > c \cdot \hat{x}$.

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x_1^*, \ldots, x_d^*)$, then $c \cdot x^* > c \cdot \hat{x}$.

How does $(c \cdot x)$ change as we move from $\hat{x}$ to x^* on the line joining the two?

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x_1^*, \ldots, x_d^*)$, then $c \cdot x^* > c \cdot \hat{x}$.

How does $(c \cdot x)$ change as we move from $\hat{x}$ to x^* on the line joining the two?

Strictly increases!

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x^*_1, \ldots, x^*_d)$, then $c \cdot x^* > c \cdot \hat{x}$.

How does $(c \cdot x)$ change as we move from $\hat{x}$ to x^* on the line joining the two?

Strictly increases!

- $d = x^* - \hat{x}$ is the direction from $\hat{x}$ to x^*.
- $x = \hat{x} + \delta d$. As δ goes from 0 to 1, x moves from $\hat{x}$ to x^*.

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x_1^*, \ldots, x_d^*)$, then $c \cdot x^* > c \cdot \hat{x}$.

How does $(c \cdot x)$ change as we move from $\hat{x}$ to x^* on the line joining the two?

Strictly increases!

- $d = x^* - \hat{x}$ is the direction from $\hat{x}$ to x^*.
- $x = \hat{x} + \delta d$. As δ goes from 0 to 1, x moves from $\hat{x}$ to x^*.
- $(c \cdot d) = (c \cdot x^*) - (c \cdot \hat{x}) > 0$.

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x^*_1, \ldots, x^*_d)$, then $c \cdot x^* > c \cdot \hat{x}$.

How does $(c \cdot x)$ change as we move from $\hat{x}$ to x^* on the line joining the two?

Strictly increases!

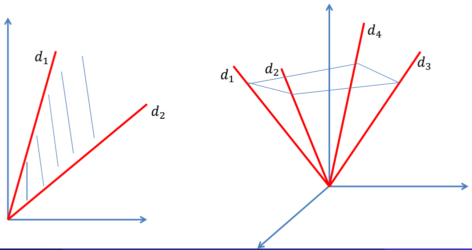
- $d = x^* - \hat{x}$ is the direction from $\hat{x}$ to x^*.
- $x = \hat{x} + \delta d$. As δ goes from 0 to 1, x moves from $\hat{x}$ to x^*.
- $(c \cdot d) = (c \cdot x^*) - (c \cdot \hat{x}) > 0$.
- $c \cdot x = c \cdot \hat{x} + \delta(c \cdot d)$. Strictly increasing with δ!
- Due to convexity, all of these are feasible points.

Cone

Definition

Given a set of vectors $D = \{d_1, \ldots, d_k\}$, the cone spanned by them is just their positive linear combinations, i.e.,

$$cone(D) = \{d \mid d = \sum_{i=1}^{k} \lambda_i d_i, \text{ where } \lambda_i \geq 0, \forall i\}$$



Cone (Contd.)

Lemma

If $\boldsymbol{d} \in \boldsymbol{cone(D)}$ *and* $(\boldsymbol{c} \cdot \boldsymbol{d}) > \boldsymbol{0}$, *then there exists* $\boldsymbol{d_i}$ *such that* $(\boldsymbol{c} \cdot \boldsymbol{d_i}) > \boldsymbol{0}$.

Proof.

To the contrary suppose $(c \cdot d_i) \leq 0, \ \forall i \leq k$.
Since d is a positive linear combination of d_i's,

$$\begin{aligned} (c \cdot d) &= (c \cdot \sum_{i=1}^{k} \lambda_i d_i) \\ &= \sum_{i=1}^{k} \lambda_i (c \cdot d_i) \\ &\leq 0 \end{aligned}$$

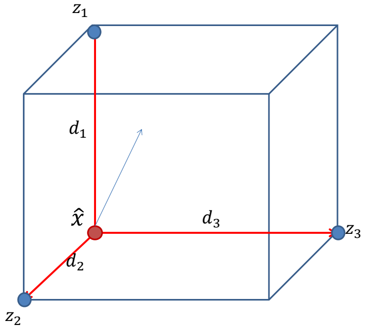
A contradiction! □

Improving Direction Implies Improving Neighbor

Let $z_1, \ldots, z_k$ be the neighboring vertices of $\hat{x}$. And let $d_i = z_i - \hat{x}$ be the direction from $\hat{x}$ to z_i.

Lemma

Any feasible direction of movement d from $\hat{x}$ is in the $cone(\{d_1, \ldots, d_k\})$.



Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x^*_1, \ldots, x^*_d)$, then $c \cdot x^* > c \cdot \hat{x}$.

- $d = x^* - \hat{x}$ is the direction from $\hat{x}$ to x^*.
- $(c \cdot d) = (c \cdot x^*) - (c \cdot \hat{x}) > 0$.

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x_1^*, \ldots, x_d^*)$, then $c \cdot x^* > c \cdot \hat{x}$.

- $d = x^* - \hat{x}$ is the direction from $\hat{x}$ to x^*.
- $(c \cdot d) = (c \cdot x^*) - (c \cdot \hat{x}) > 0$.
- Let d_i be the direction towards neighbor z_i.
- $d \in Cone(\{d_1, \ldots, d_k\}) \Rightarrow \exists d_i,\ (c \cdot d_i) > 0$.

Observations

For Simplex

Suppose we are at a non-optimal vertex $\hat{x} = (\hat{x}_1, \ldots, \hat{x}_d)$ and optimal is $x^* = (x^*_1, \ldots, x^*_d)$, then $c \cdot x^* > c \cdot \hat{x}$.

- $d = x^* - \hat{x}$ is the direction from $\hat{x}$ to x^*.
- $(c \cdot d) = (c \cdot x^*) - (c \cdot \hat{x}) > 0$.
- Let d_i be the direction towards neighbor z_i.
- $d \in Cone(\{d_1, \ldots, d_k\}) \Rightarrow \exists d_i,\ (c \cdot d_i) > 0$.

Theorem

If vertex $\hat{x}$ is not optimal then it has a neighbor where the objective value $(c \cdot x)$ improves.

How Many Neighbors a Vertex Has?

Geometric view...

$A \in R^{n \times d}$ $(n > d)$, $b \in R^n$, the constraints are: $Ax \leq b$

Faces

- n constraints/inequalities. Each defines a hyperplane.
- Vertex: 0-dimensional face. Edge: 1D face. $\ldots$ Hyperplane: $(d-1)$D face.

How Many Neighbors a Vertex Has?

Geometric view...

$A \in R^{n \times d}$ $(n > d)$, $b \in R^n$, the constraints are: $Ax \leq b$

Faces

- n constraints/inequalities. Each defines a hyperplane.
- Vertex: 0-dimensional face. Edge: 1D face. $\ldots$ Hyperplane: $(d-1)$D face.
- r linearly independent hyperplanes forms $d-r$ dimensional face.

How Many Neighbors a Vertex Has?

Geometric view...

$A \in R^{n \times d}$ $(n > d)$, $b \in R^n$, the constraints are: $Ax \leq b$

Faces

- n constraints/inequalities. Each defines a hyperplane.
- Vertex: 0-dimensional face. Edge: 1D face. $\ldots$ Hyperplane: $(d-1)$D face.
- r linearly independent hyperplanes forms $d - r$ dimensional face.
- Vertices being of 0D, d L.I. hyperplanes form a vertex.

How Many Neighbors a Vertex Has?

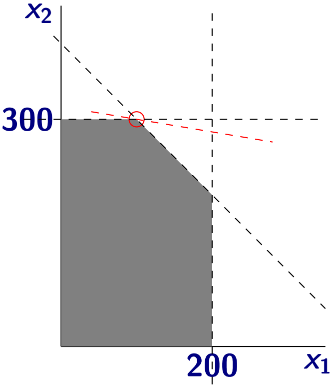
Geometric view...

$A \in R^{n \times d}$ $(n > d)$, $b \in R^n$, the constraints are: $Ax \leq b$

Faces

- n constraints/inequalities. Each defines a hyperplane.
- Vertex: 0-dimensional face. Edge: 1D face. $\ldots$ Hyperplane: $(d-1)$D face.
- r linearly independent hyperplanes forms $d - r$ dimensional face.
- Vertices being of 0D, d L.I. hyperplanes form a vertex.

In 2-dimension $(d = 2)$



How Many Neighbors a Vertex Has?

Geometric view...

$A \in R^{n \times d}$ $(n > d)$, $b \in R^n$, the constraints are: $Ax \leq b$

In 3-dimension $(d = 3)$

Faces

- n constraints/inequalities. Each defines a hyperplane.
- Vertex: 0-dimensional face. Edge: 1D face. ... Hyperplane: $(d - 1)$D face.
- r linearly independent hyperplanes forms $d - r$ dimensional face.
- Vertices being of 0D, d L.I. hyperplanes form a vertex.

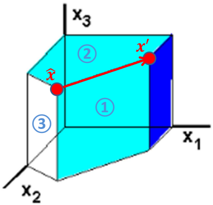


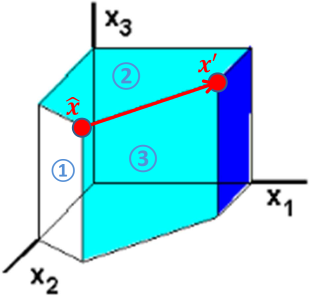
image source: webpage of Prof. Forbes W. Lewis

How Many Neighbors a Vertex Has?

Geometry view...

One neighbor per tight hyperplane. Therefore typically d.

- Suppose x' is a neighbor of $\hat{x}$, then on the edge joining is defined by $(d-1)$ hyperplanes.
- hx and x' also shares these $d-1$ hyperplanes
- In addition one more hyperplane, say $(Ax)_i = b_i$, is tight at $\hat{x}$. "Relaxing" this at $\hat{x}$ leads to x'.



Simplex Algorithm

Simplex: Vertex hoping algorithm

Moves from a vertex to its neighboring vertex

Questions + Answers

- Which neighbor to move to? One where objective value increases.

Simplex Algorithm

Simplex: Vertex hoping algorithm

Moves from a vertex to its neighboring vertex

Questions + Answers

- Which neighbor to move to? One where objective value increases.
- When to stop? When no neighbor with better objective value.

Simplex Algorithm

Simplex: Vertex hoping algorithm

Moves from a vertex to its neighboring vertex

Questions + Answers

- Which neighbor to move to? One where objective value increases.
- When to stop? When no neighbor with better objective value.
- How much time does it take? At most d neighbors to consider in each step.

Simplex in 2-d

Simplex Algorithm

1. Start from some vertex of the feasible polygon.
2. Compare value of objective function at current vertex with the value at **2** "neighboring" vertices of polygon.
3. If neighboring vertex improves objective function, move to this vertex, and repeat step 2.
4. If no improving neighbor (local optimum), then stop.

Simplex in Higher Dimensions

Simplex Algorithm

1. Start at a vertex of the polytope.
2. Compare value of objective function at each of the d "neighbors".
3. Move to neighbor that improves objective function, and repeat step 2.
4. If no improving neighbor, then stop.

Simplex in Higher Dimensions

Simplex Algorithm

1. Start at a vertex of the polytope.
2. Compare value of objective function at each of the d "neighbors".
3. Move to neighbor that improves objective function, and repeat step 2.
4. If no improving neighbor, then stop.

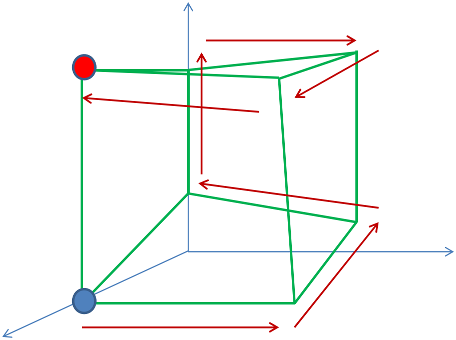
Simplex is a greedy local-improvement algorithm! Works because a local optimum is also a global optimum — convexity of polyhedra.

Solving Linear Programming in Practice

1. Naïve implementation of Simplex algorithm can be very inefficient

Solving Linear Programming in Practice

1. Naïve implementation of Simplex algorithm can be very inefficient – Exponential number of steps!



Solving Linear Programming in Practice

1. Naïve implementation of Simplex algorithm can be very inefficient
   1. Choosing which neighbor to move to can significantly affect running time
   2. Very efficient Simplex-based algorithms exist
   3. Simplex algorithm takes exponential time in the worst case but works extremely well in practice with many improvements over the years
2. Non Simplex based methods like interior point methods work well for large problems.

Polynomial time Algorithm for Linear Programming

Major open problem for many years: is there a polynomial time algorithm for linear programming?

Polynomial time Algorithm for Linear Programming

Major open problem for many years: is there a polynomial time algorithm for linear programming?
Leonid Khachiyan in 1979 gave the first polynomial time algorithm using the Ellipsoid method.

1. major theoretical advance
2. highly impractical algorithm, not used at all in practice
3. routinely used in theoretical proofs.

Polynomial time Algorithm for Linear Programming

Major open problem for many years: is there a polynomial time algorithm for linear programming?

Leonid Khachiyan in 1979 gave the first polynomial time algorithm using the Ellipsoid method.

1. major theoretical advance
2. highly impractical algorithm, not used at all in practice
3. routinely used in theoretical proofs.

Narendra Karmarkar in 1984 developed another polynomial time algorithm, the interior point method.

1. very practical for some large problems and beats simplex
2. also revolutionized theory of interior point methods

Polynomial time Algorithm for Linear Programming

Major open problem for many years: is there a polynomial time algorithm for linear programming?

Leonid Khachiyan in 1979 gave the first polynomial time algorithm using the Ellipsoid method.

1. major theoretical advance
2. highly impractical algorithm, not used at all in practice
3. routinely used in theoretical proofs.

Narendra Karmarkar in 1984 developed another polynomial time algorithm, the interior point method.

1. very practical for some large problems and beats simplex
2. also revolutionized theory of interior point methods

Following interior point method success, Simplex has been improved enormously and is the method of choice.

Degeneracy

1. The linear program could be infeasible: No points satisfy the constraints.
2. The linear program could be unbounded: Polygon unbounded in the direction of the objective function.
3. More than d hyperplanes could be tight at a vertex, forming more than d neighbors.

Infeasibility: Example

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 2 \quad x_2 \leq 1 \quad x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

Infeasibility has to do only with constraints.

Infeasibility: Example

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 2 \quad x_2 \leq 1 \quad x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

Infeasibility has to do only with constraints.

No starting vertex for Simplex.

Infeasibility: Example

$$\begin{array}{ll} \text{maximize} & x_1 + 6x_2 \\ \text{subject to} & x_1 \leq 2 \quad x_2 \leq 1 \quad x_1 + x_2 \geq 4 \\ & x_1, x_2 \geq 0 \end{array}$$

Infeasibility has to do only with constraints.

No starting vertex for Simplex. How to detect this?

Unboundedness: Example

$$\begin{aligned} \text{maximize} \quad & x_2 \\ x_1 + x_2 & \geq 2 \\ x_1, x_2 & \geq 0 \end{aligned}$$

Unboundedness depends on both constraints and the objective function.

Unboundedness: Example

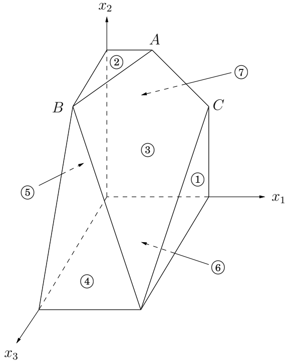
$$\begin{aligned} \text{maximize} \quad & x_2 \\ x_1 + x_2 &\geq 2 \\ x_1, x_2 &\geq 0 \end{aligned}$$

Unboundedness depends on both constraints and the objective function.

If unbounded in the direction of objective function, then Simplex detects it.

Degeneracy and Cycling

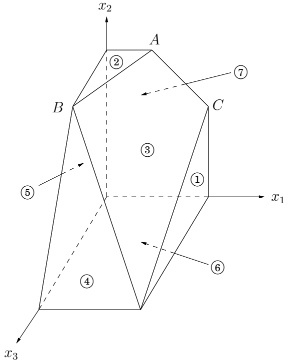
More than ***d*** inequalities tight at a vertex.



$$
\begin{aligned}
\max \quad & x_1 + 6x_2 + 13x_3 & \\
& x_1 \leq 200 & ① \\
& x_2 \leq 300 & ② \\
& x_1 + x_2 + x_3 \leq 400 & ③ \\
& x_2 + 3x_3 \leq 600 & ④ \\
& x_1 \geq 0 & ⑤ \\
& x_2 \geq 0 & ⑥ \\
& x_3 \geq 0 & ⑦
\end{aligned}
$$

Degeneracy and Cycling

More than ***d*** inequalities tight at a vertex.



$$
\begin{aligned}
\max \quad & x_1 + 6x_2 + 13x_3 & \\
& x_1 \le 200 & ① \\
& x_2 \le 300 & ② \\
& x_1 + x_2 + x_3 \le 400 & ③ \\
& x_2 + 3x_3 \le 600 & ④ \\
& x_1 \ge 0 & ⑤ \\
& x_2 \ge 0 & ⑥ \\
& x_3 \ge 0 & ⑦
\end{aligned}
$$

Depending on how Simplex is implemented, it may cycle at this vertex.

We will see how in the next lecture.