

Lecture slides by Kevin Wayne

## LINEAR PROGRAMMING I

- a refreshing example
- standard form
- geometry
- linear algebra
- ▶ simplex algorithm

## Linear programming

Linear programming. Optimize a linear function subject to linear inequalities.

(P) 
$$\max \sum_{j=1}^{n} c_j x_j$$
  
s.t.  $\sum_{j=1}^{n} a_{ij} x_j = b_i \quad 1 \le i \le m$   
 $x_j \ge 0 \quad 1 \le j \le n$ 

(P) 
$$\max c^T x$$
  
s.t.  $Ax = b$   
 $x \ge 0$ 

## Linear programming

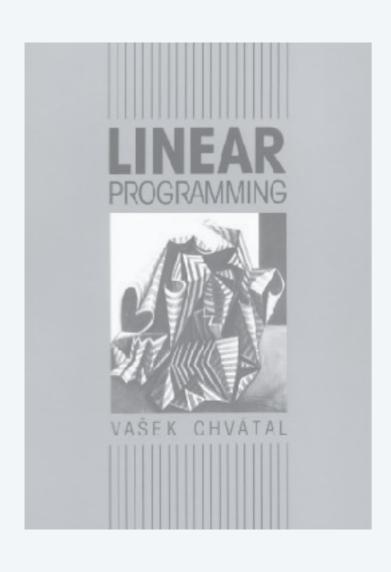
Linear programming. Optimize a linear function subject to linear inequalities.

Generalizes: Ax = b, 2-person zero-sum games, shortest path, max flow, assignment problem, matching, multicommodity flow, MST, min weighted arborescence, ...

#### Why significant?

- Design poly-time algorithms.
- Design approximation algorithms.
- Solve NP-hard problems using branch-and-cut.

Ranked among most important scientific advances of 20th century.



## LINEAR PROGRAMMING I

- a refreshing example
- > standard form
- ▶ geometry
- ► linear algebra
- ► simplex algorithm

Reference: The Allocation of Resources by Linear Programming, Scientific American, by Bob Bland

## Brewery problem

#### Small brewery produces ale and beer.

- Production limited by scarce resources: corn, hops, barley malt.
- Recipes for ale and beer require different proportions of resources.

Beverage	Corn (pounds)	Hops (ounces)	Malt (pounds)	Profit (\$)
Ale (barrel)	5	4	35	13
Beer (barrel)	15	4	20	23
constraint	480	160	1190	

#### How can brewer maximize profits?

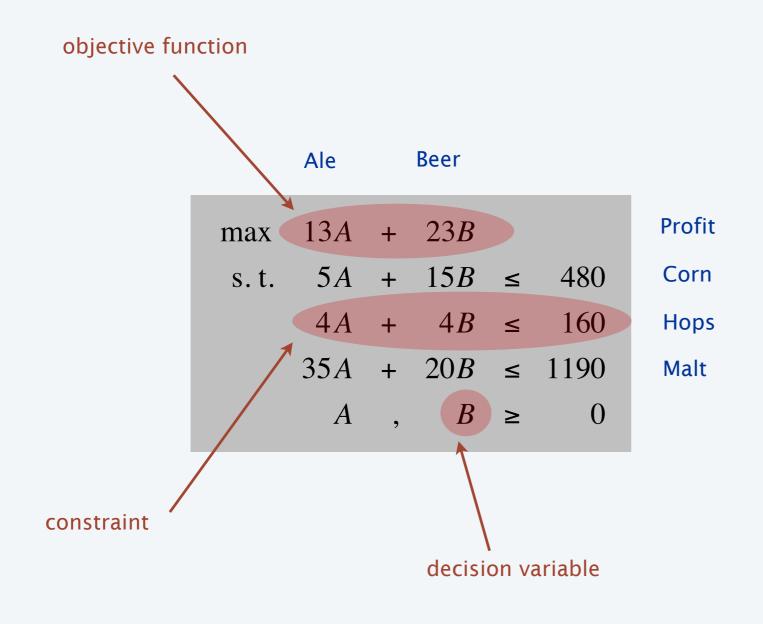
Devote all resources to ale: 34 barrels of ale ⇒ \$442

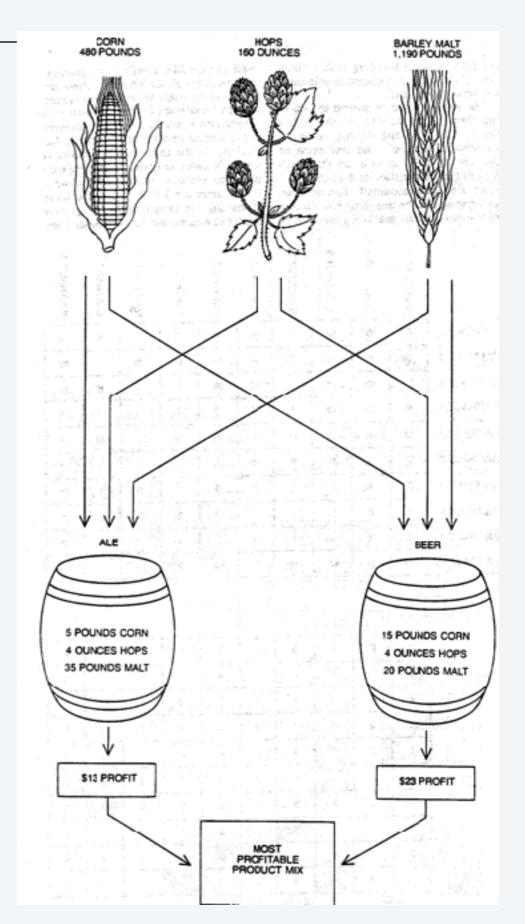
Devote all resources to beer: 32 barrels of beer ⇒ \$736

• 7.5 barrels of ale, 29.5 barrels of beer  $\Rightarrow$  \$776

• 12 barrels of ale, 28 barrels of beer ⇒ \$800

# Brewery problem



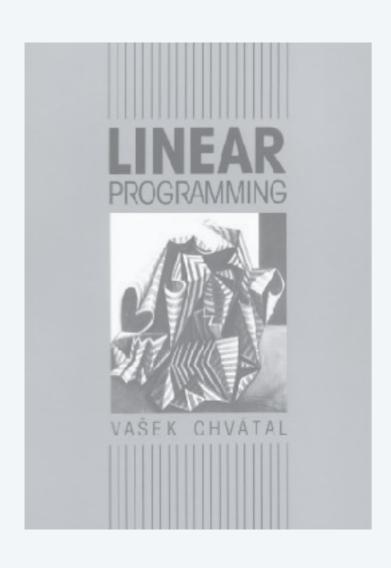


SCIENTIFIC AMERICAN JUNE 1981

# The Allocation of Resources by Linear Programming

Abstract, crystal-like structures in many geometrical dimensions can help to solve problems in planning and management. A new algorithm has set upper limits on the complexity of such problems

By Robert G. Bland



## LINEAR PROGRAMMING I

- ▶ a refreshing example
- > standard form
- ▶ geometry
- linear algebra
- ► simplex algorithm

## Standard form of a linear program

#### "Standard form" of an LP.

- Input: real numbers  $a_{ij}, c_j, b_i$ .
- Output: real numbers  $x_i$ .
- n = # decision variables, m = # constraints.
- Maximize linear objective function subject to linear inequalities.

(P) 
$$\max \sum_{j=1}^{n} c_j x_j$$
  
s.t.  $\sum_{j=1}^{n} a_{ij} x_j = b_i \quad 1 \le i \le m$   
 $x_j \ge 0 \quad 1 \le j \le n$ 

(P) 
$$\max c^T x$$
  
s. t.  $Ax = b$   
 $x \ge 0$ 

Linear. No  $x^2$ , xy, arccos(x), etc.

Programming. Planning (term predates computer programming).

## Brewery problem: converting to standard form

#### Original input.

max 
$$13A + 23B$$
  
s. t.  $5A + 15B \le 480$   
 $4A + 4B \le 160$   
 $35A + 20B \le 1190$   
 $A , B \ge 0$ 

#### Standard form.

- Add slack variable for each inequality.
- Now a 5-dimensional problem.

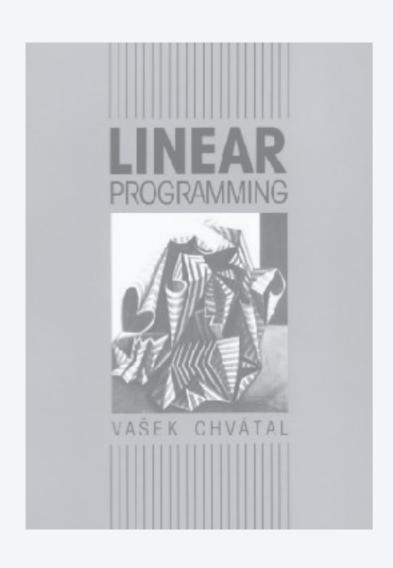
## Equivalent forms

Easy to convert variants to standard form.

(P) 
$$\max c^T x$$
  
s.t.  $Ax = b$   
 $x \ge 0$ 

Less than to equality.  $x + 2y - 3z \le 17 \Rightarrow x + 2y - 3z + s = 17, s \ge 0$ Greater than to equality.  $x + 2y - 3z \ge 17 \Rightarrow x + 2y - 3z - s = 17, s \ge 0$ Min to max. min  $x + 2y - 3z \Rightarrow \max -x - 2y + 3z$ Unrestricted to nonnegative. x unrestricted  $\Rightarrow x = x^+ - x^-, x^+ \ge 0, x^- \ge 0$ 

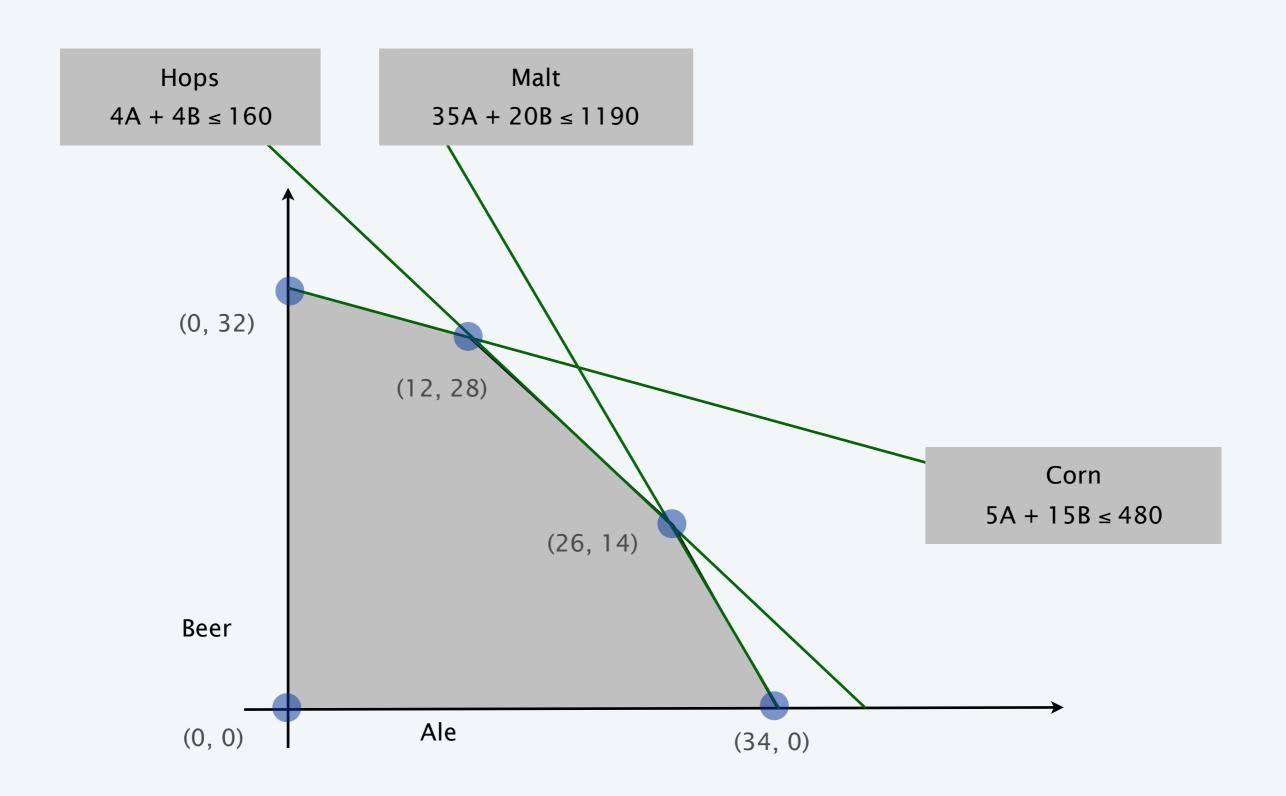
11



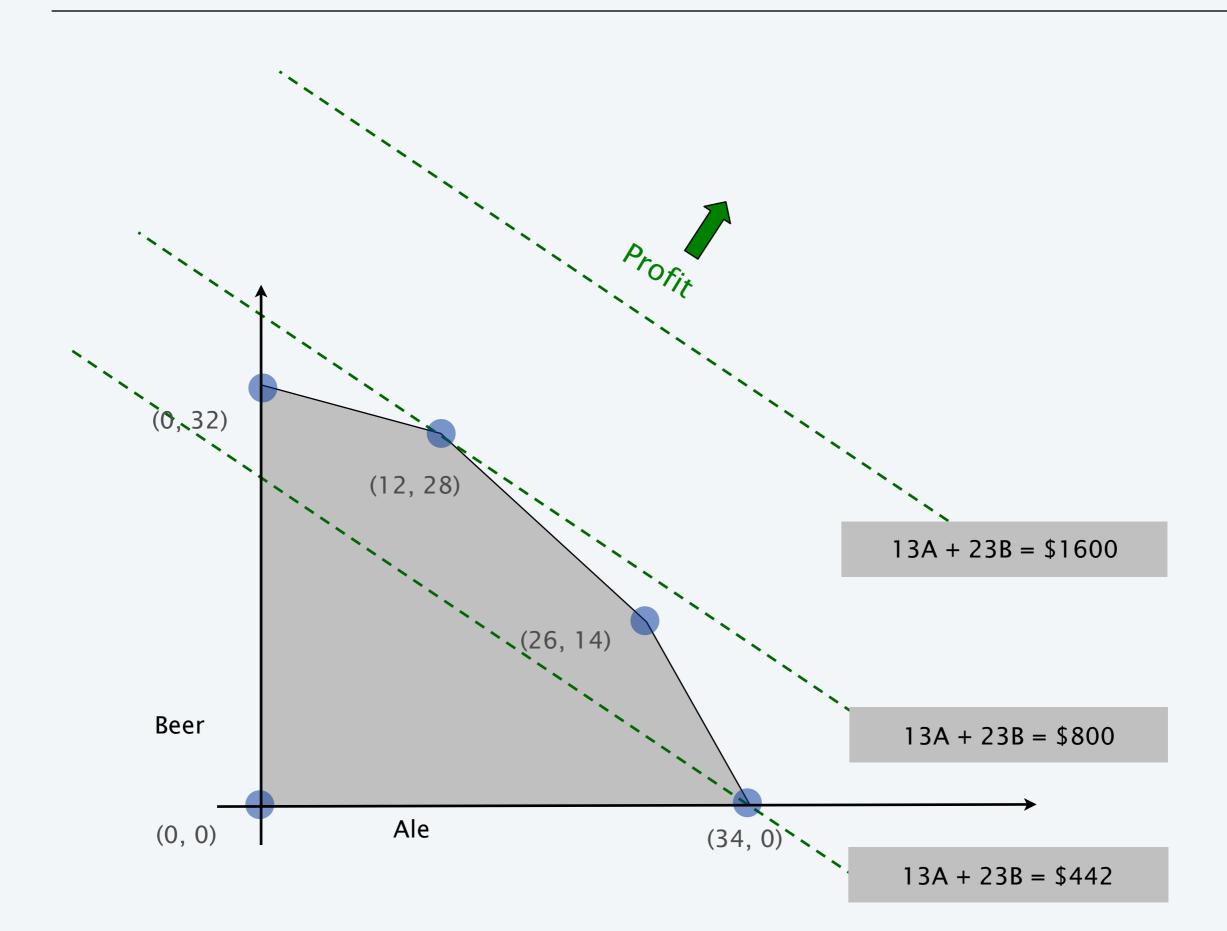
## LINEAR PROGRAMMING I

- ▶ a refreshing example
- > standard form
- geometry
- linear algebra
- ► simplex algorithm

## Brewery problem: feasible region

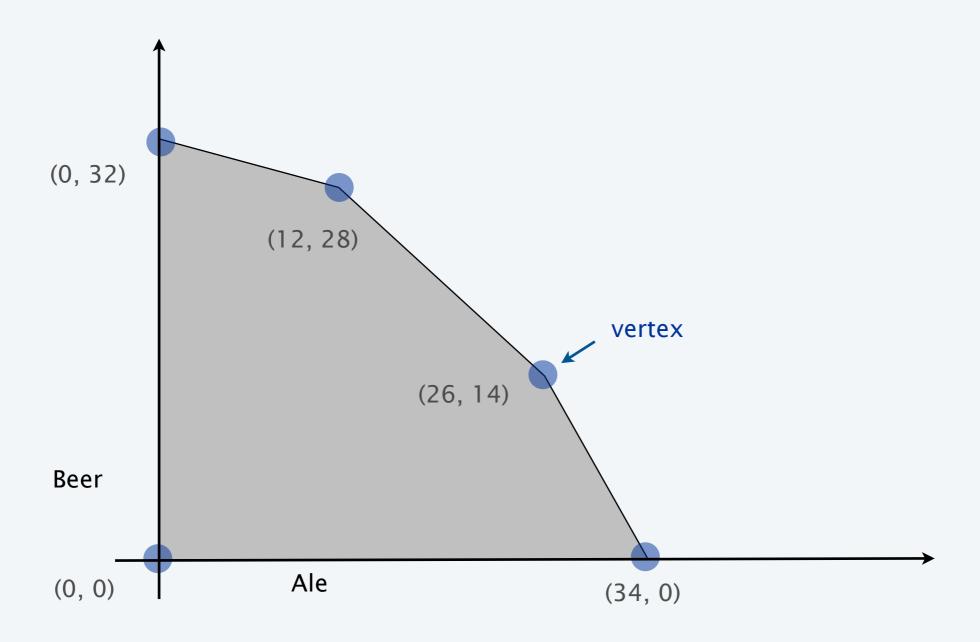


# Brewery problem: objective function



## Brewery problem: geometry

Brewery problem observation. Regardless of objective function coefficients, an optimal solution occurs at a vertex.

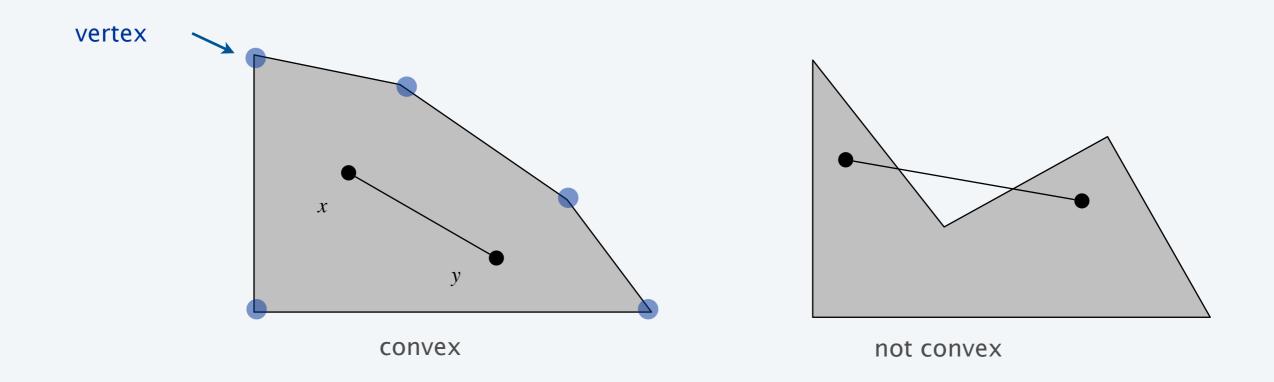


## Convexity

Convex set. If two points x and y are in the set, then so is  $\lambda x + (1-\lambda)y$  for  $0 \le \lambda \le 1$ .



Vertex. A point *x* in the set that can't be written as a strict convex combination of two distinct points in the set.



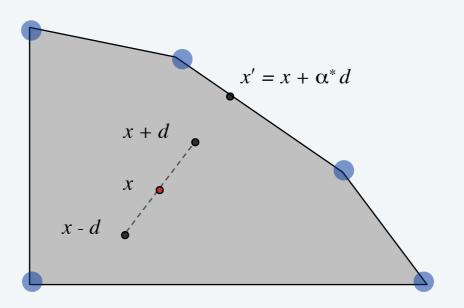
Observation. LP feasible region is a convex set.

## **Purificaiton**

Theorem. If there exists an optimal solution to (P), then there exists one that is a vertex.

(P) 
$$\max c^T x$$
  
s.t.  $Ax = b$   
 $x \ge 0$ 

Intuition. If x is not a vertex, move in a non-decreasing direction until you reach a boundary. Repeat.



#### **Purification**

Theorem. If there exists an optimal solution to (P), then there exists one that is a vertex.

#### Pf.

- Suppose x is an optimal solution that is not a vertex.
- There exist direction  $d \neq 0$  such that  $x \pm d \in P$ .
- Ad = 0 because  $A(x \pm d) = b$ .
- Assume  $c^{T}d \le 0$  (by taking either d or -d).
- Consider  $x + \lambda d$ ,  $\lambda > 0$ :

### Case 1. [ there exists j such that $d_j < 0$ ]

- Increase  $\lambda$  to  $\lambda^*$  until first new component of  $x + \lambda d$  hits 0.
- $x + \lambda^* d$  is feasible since  $A(x + \lambda^* d) = Ax = b$  and  $x + \lambda^* y \ge 0$ .
- $x + \lambda^* d$  has one more zero component than x.
- $c^{\mathrm{T}}x' = c^{\mathrm{T}}(x + \lambda^* d) = c^{\mathrm{T}}x + \lambda^* c^{\mathrm{T}}d \le c^{\mathrm{T}}x$ .

 $d_k = 0$  whenever  $x_k = 0$  because  $x \pm d \in P$ 

#### **Purification**

Theorem. If there exists an optimal solution to (P), then there exists one that is a vertex.

#### Pf.

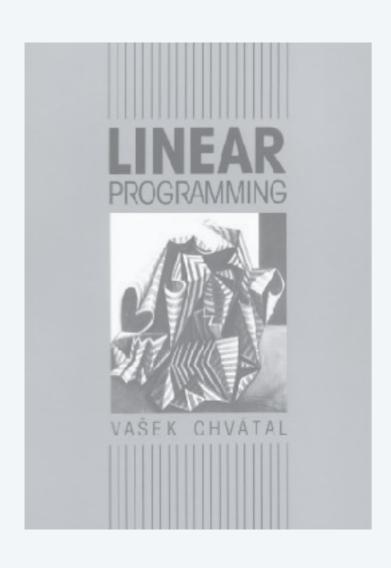
- Suppose x is an optimal solution that is not a vertex.
- There exist direction  $d \neq 0$  such that  $x \pm d \in P$ .
- Ad = 0 because  $A(x \pm d) = b$ .
- Assume  $c^{T}d \le 0$  (by taking either d or -d).
- Consider  $x + \lambda d$ ,  $\lambda > 0$ :

## Case 2. $[d_j \ge 0 \text{ for all } j]$

- $x + \lambda d$  is feasible for all  $\lambda \ge 0$  since  $A(x + \lambda d) = b$  and  $x + \lambda d \ge x \ge 0$ .
- As  $\lambda \to \infty$ ,  $c^{T}(x + \lambda d) \to \infty$  because  $c^{T}d < 0$ .



if  $c^{T}d = 0$ , choose d so that case 1 applies

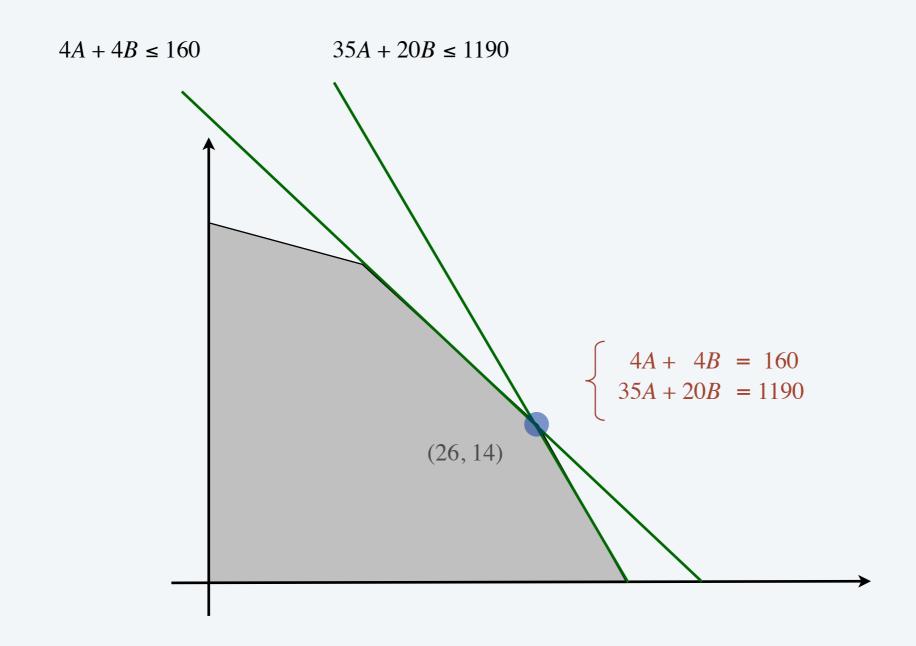


## LINEAR PROGRAMMING I

- ▶ a refreshing example
- > standard form
- ▶ geometry
- ▶ linear algebra
- ▶ simplex algorithm

## Intuition

Intuition. A vertex in  $\Re^m$  is uniquely specified by m linearly independent equations.



#### Basis

Basis: Subset of *m* of *n* variables

#### Basic feasible solution (BFS):

- Set n-m non-basic variables to zero, and solve for m basic variables
- Solve *m* equations for *m* unknowns
- If unique and feasible solution, it is BFS
- A BFS is a vertex

(P) 
$$\max \sum_{j=1}^{n} c_j x_j$$
  
s.t.  $\sum_{j=1}^{n} a_{ij} x_j = b_i \quad 1 \le i \le m$   
 $x_j \ge 0 \quad 1 \le j \le n$ 

Theorem. Let  $P = \{x : Ax = b, x \ge 0\}$ . For  $x \in P$ , define  $B = \{j : x_j > 0\}$ . Then, x is a vertex iff  $A_B$  has linearly independent columns.

Notation. Let B = set of column indices. Define  $A_B$  to be the subset of columns of A indexed by B.

Ex.

$$A = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 7 & 3 & 2 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}, b = \begin{bmatrix} 7 \\ 16 \\ 0 \end{bmatrix}$$

$$x = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad B = \{1, 3\}, \quad A_B = \begin{bmatrix} 2 & 3 \\ 7 & 2 \\ 0 & 0 \end{bmatrix}$$

Theorem. Let  $P = \{x : Ax = b, x \ge 0\}$ . For  $x \in P$ , define  $B = \{j : x_j > 0\}$ . Then, x is a vertex iff  $A_B$  has linearly independent columns.

#### **Pf.** ←

- Assume x is not a vertex.
- There exist direction  $d \neq 0$  such that  $x \pm d \in P$ .
- Ad = 0 because  $A(x \pm d) = b$ .
- Define  $B' = \{ j : d_j \neq 0 \}.$
- $A_{B'}$  has linearly dependent columns since  $d \neq 0$ .
- Moreover,  $d_i = 0$  whenever  $x_i = 0$  because  $x \pm d \ge 0$ .
- Thus  $B' \subseteq B$ , so  $A_{B'}$  is a submatrix of  $A_B$ .
- Therefore,  $A_B$  has linearly dependent columns.

Theorem. Let  $P = \{x : Ax = b, x \ge 0\}$ . For  $x \in P$ , define  $B = \{j : x_j > 0\}$ . Then, x is a vertex iff  $A_B$  has linearly independent columns.

#### Pf. $\Rightarrow$

- Assume  $A_B$  has linearly dependent columns.
- There exist  $d \neq 0$  such that  $A_B d = 0$ .
- Extend d to  $\Re^n$  by adding 0 components.
- Now, Ad = 0 and  $d_j = 0$  whenever  $x_j = 0$ .
- For sufficiently small  $\lambda$ ,  $x \pm \lambda d \in P \Rightarrow x$  is not a vertex. •

Theorem. Given  $P = \{x : Ax = b, x \ge 0\}$ , x is a vertex iff there exists  $B \subseteq \{1, ..., n\}$  such |B| = m and:

- $A_B$  is nonsingular.
- $x_B = A_{B^{-1}} b \ge 0.$
- $x_N = 0$ .

basic feasible solution

Pf. Augment  $A_B$  with linearly independent columns (if needed).

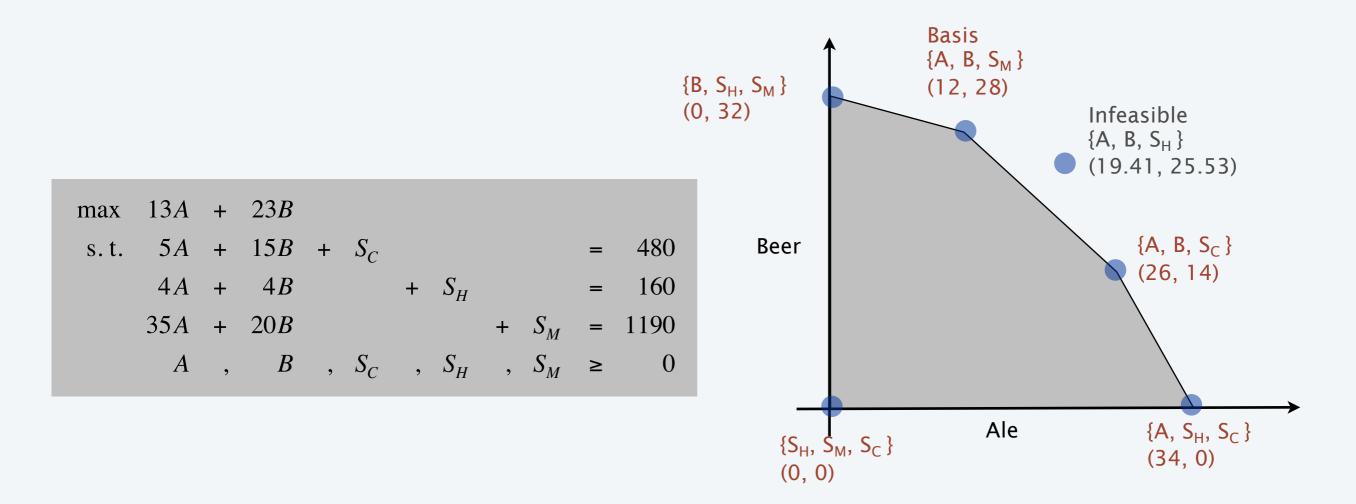
$$A = \begin{bmatrix} 2 & 1 & 3 & 0 \\ 7 & 3 & 2 & 1 \\ 0 & 0 & 0 & 5 \end{bmatrix}, b = \begin{bmatrix} 7 \\ 16 \\ 0 \end{bmatrix}$$

$$x = \begin{bmatrix} 2 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad B = \{1, 3, 4\}, \quad A_B = \begin{bmatrix} 2 & 3 & 0 \\ 7 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

Assumption.  $A \in \Re^{m \times n}$  has full row rank.

## Basic feasible solution: example

#### Basic feasible solutions.



## Fundamental questions

LP. For  $A \in \Re^{m \times n}$ ,  $b \in \Re^m$ ,  $c \in \Re^n$ ,  $\alpha \in \Re$ , does there exist  $x \in \Re^n$  such that: Ax = b,  $x \ge 0$ ,  $c^Tx \ge \alpha$ ?

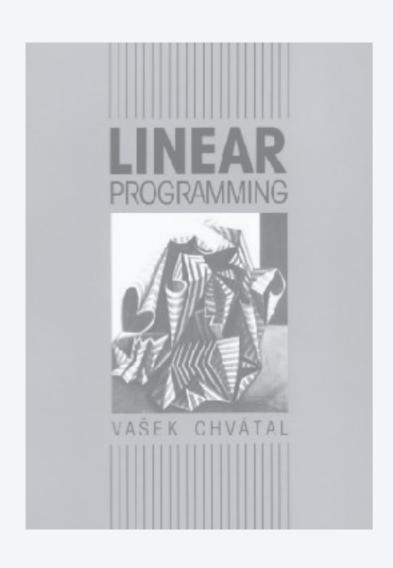
#### Q. Is LP in NP?

#### A. Yes.

- Number of vertices  $\leq C(n, m) = \binom{n}{m} \leq n^m$ .
- Cramer's rule ⇒ can check a vertex in poly-time.

Cramer's rule. For  $B \in \Re^{n \times n}$  invertible,  $b \in \Re^n$ , the solution to Bx = b is given by:

$$x_i = \frac{\det(B_i)}{\det(B)}$$
 replace *i*th column of *B* with *b*



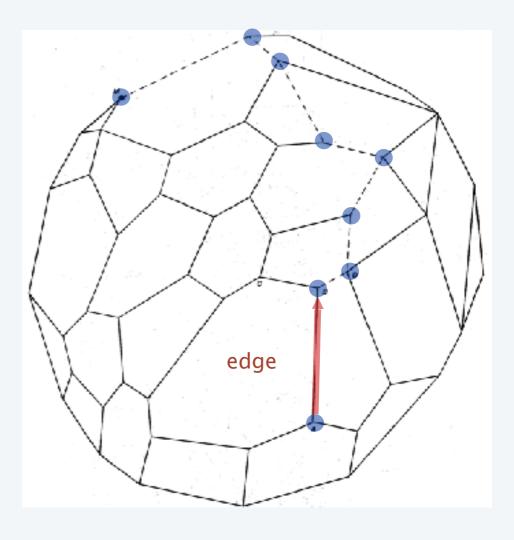
## LINEAR PROGRAMMING I

- a refreshing example
- > standard form
- ▶ geometry
- Inear algebra
- ▶ simplex algorithm

## Simplex algorithm: intuition

Simplex algorithm. [George Dantzig 1947] Move from BFS to adjacent BFS, without decreasing objective function.

replace one basic variable with another



Greedy property. BFS optimal iff no adjacent BFS is better. Challenge. Number of BFS can be exponential!

# Simplex algorithm: initialization

max 2	Z su	bject t	O								
13 <i>A</i>	+	23 <i>B</i>						_	Z	=	0
5 <i>A</i>	+	15 <i>B</i>	+	$S_C$						=	480
4 <i>A</i>	+	4 <i>B</i>			+	$S_H$				=	160
35 <i>A</i>	+	20 <i>B</i>					+	$S_{M}$		=	1190
A	,	В	,	$S_C$	,	$S_H$	,	$S_{M}$		≥	0

Basis = 
$$\{S_C, S_H, S_M\}$$
  
 $A = B = 0$   
 $Z = 0$   
 $S_C = 480$   
 $S_H = 160$   
 $S_M = 1190$ 

## Simplex algorithm: pivot 1

Basis = 
$$\{S_C, S_H, S_M\}$$
  
 $A = B = 0$   
 $Z = 0$   
 $S_C = 480$   
 $S_H = 160$   
 $S_M = 1190$ 

Substitute:  $B = 1/15 (480 - 5A - S_C)$ 

Basis = 
$$\{B, S_H, S_M\}$$
  
 $A = S_C = 0$   
 $Z = 736$   
 $B = 32$   
 $S_H = 32$   
 $S_M = 550$ 

## Simplex algorithm: pivot 1

Basis = 
$$\{S_C, S_H, S_M\}$$
  
 $A = B = 0$   
 $Z = 0$   
 $S_C = 480$   
 $S_H = 160$   
 $S_M = 1190$ 

- Q. Why pivot on column 2 (or 1)?
- A. Each unit increase in B increases objective value by \$23.
- Q. Why pivot on row 2?
- A. Preserves feasibility by ensuring RHS  $\geq 0$ .

min ratio rule: min { 480/15, 160/4, 1190/20 }

## Simplex algorithm: pivot 2

Basis = 
$$\{B, S_H, S_M\}$$
  
 $A = S_C = 0$   
 $Z = 736$   
 $B = 32$   
 $S_H = 32$   
 $S_M = 550$ 

Substitute:  $A = 3/8 (32 + 4/15 S_C - S_H)$ 

Basis = 
$$\{A, B, S_M\}$$
  
 $S_C = S_H = 0$   
 $Z = 800$   
 $B = 28$   
 $A = 12$   
 $S_M = 110$ 

## Simplex algorithm: optimality

- Q. When to stop pivoting?
- A. When all coefficients in top row are nonpositive.
- Q. Why is resulting solution optimal?
- A. Any feasible solution satisfies system of equations in tableaux.
  - In particular:  $Z = 800 S_C 2 S_H$ ,  $S_C \ge 0$ ,  $S_H \ge 0$ .
  - Thus, optimal objective value  $Z^* \le 800$ .
  - Current BFS has value 800 ⇒ optimal.

max Z subjec	t to								
	-	$S_C$	_	$2 S_H$		_	Z	=	-800
В	+	$\frac{1}{10} S_C$	+	$\frac{1}{8}$ $S_H$				=	28
A	-	$\frac{1}{10} S_C$	+	$\frac{3}{8}$ $S_H$				=	12
	-	$\frac{25}{6}$ $S_C$	_	$\frac{85}{8} S_H$	+	$S_{M}$		=	110
A , $B$	,	$S_C$	,	$S_H$	,	$S_{M}$		≥	0

Basis = 
$$\{A, B, S_M\}$$
  
 $S_C = S_H = 0$   
 $Z = 800$   
 $B = 28$   
 $A = 12$   
 $S_M = 110$ 

## Simplex tableaux: matrix form

#### Initial simplex tableaux.

$$c_B^T x_B + c_N^T x_N = Z$$

$$A_B x_B + A_N x_N = b$$

$$x_B , x_N \ge 0$$

#### Simplex tableaux corresponding to basis *B*.

$$(c_N^T - c_B^T A_B^{-1} A_N) x_N = Z - c_B^T A_B^{-1} b \qquad \text{subtract } c_B^T A_B^{-1} \text{ times constraints}$$
 
$$I x_B + A_B^{-1} A_N x_N = A_B^{-1} b \qquad \text{multiply by } A_B^{-1}$$
 
$$x_B \quad , \qquad x_N \geq 0$$

$$x_B = A_{B^{-1}}b \ge 0$$
  
 $x_N = 0$   $c_N^T - c_B^T A_{B^{-1}} A_N \le 0$ 

basic feasible solution

optimal basis

## Simplex algorithm: corner cases

Simplex algorithm. Missing details for corner cases.

- Q. What if min ratio test fails?
- Q. How to find initial basis?
- Q. How to guarantee termination?

## Unboundedness

#### Q. What happens if min ratio test fails?



max Z subject to					
	+	$2x_4$	+	$20x_{5}$	-Z = 2
$x_1$	_	$4x_4$	_	$8x_5$	= 3
$x_2$	+	$5x_4$	_	$12x_{5}$	= 4
$x_3$					= 5
$x_1$ , $x_2$ , $x_3$	,	$x_4$	,	$x_5$	≥ 0

A. Unbounded objective function.

$$Z = 2 + 20x_5 \qquad \begin{vmatrix} x_2 \\ x_3 \\ x_4 \end{vmatrix} = \begin{vmatrix} 4 + 12x_5 \\ 5 \\ 0 \end{vmatrix}$$

 $3 + 8x_5$ 

## Phase I simplex method

Q. How to find initial basis?

(P) 
$$\max c^T x$$
  
s.t.  $Ax = b$   
 $x \ge 0$ 

A. Solve (P'), starting from basis consisting of all the  $z_i$  variables.

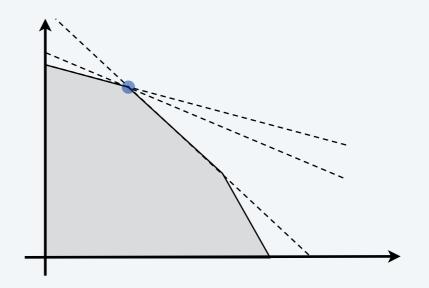
(P') max 
$$\sum_{i=1}^{m} z_i$$
s. t.  $Ax + Iz = b$ 

$$x, z \ge 0$$

- Case 1:  $min > 0 \Rightarrow (P)$  is infeasible.
- Case 2: min = 0, basis has no  $z_i$  variables  $\Rightarrow$  okay to start Phase II.
- Case 3a: min = 0, basis has  $z_i$  variables. Pivot  $z_i$  variables out of basis. If successful, start Phase II; else remove linear dependent rows.

## Simplex algorithm: degeneracy

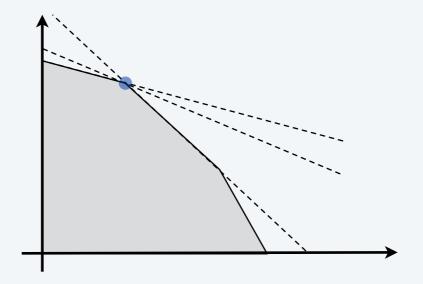
Degeneracy. New basis, same vertex.



Degenerate pivot. Min ratio = 0.

## Simplex algorithm: degeneracy

Degeneracy. New basis, same vertex.



Cycling. Infinite loop by cycling through different bases that all correspond to same vertex.

#### Anti-cycling rules.

- Bland's rule: choose eligible variable with smallest index.
- Random rule: choose eligible variable uniformly at random.
- Lexicographic rule: perturb constraints so nondegenerate.

## Lexicographic rule

Intuition. No degeneracy  $\Rightarrow$  no cycling.

Perturbed problem.

(P') max 
$$c^T x$$
  
s.t.  $Ax = b + \varepsilon$  where  $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ , such that  $\varepsilon_1 \succ \varepsilon_2 \succ \cdots \succ \varepsilon_n$ 

**Lexicographic rule.** Apply perturbation virtually by manipulating ε symbolically:

$$17 + 5\varepsilon_1 + 11\varepsilon_2 + 8\varepsilon_3 \le 17 + 5\varepsilon_1 + 14\varepsilon_2 + 3\varepsilon_3$$

much much greater, say  $\varepsilon_i = \delta^i$  for small  $\delta$ 

## Lexicographic rule

Intuition. No degeneracy  $\Rightarrow$  no cycling.

Perturbed problem.

(P') max 
$$c^T x$$
  
s. t.  $Ax = b + \varepsilon$  where  $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$ , such that  $\varepsilon_1 \succ \varepsilon_2 \succ \cdots \succ \varepsilon_n$ 

Claim. In perturbed problem,  $x_B = A_{B^{-1}}(b + \varepsilon)$  is always nonzero. Pf. The  $j^{th}$  component of  $x_B$  is a (nonzero) linear combination of the components of  $b + \varepsilon \Rightarrow$  contains at least one of the  $\varepsilon_i$  terms.

Corollary. No cycling.

43

much much greater, say  $\varepsilon_i = \delta^i$  for small  $\delta$ 

which can't cancel

## Simplex algorithm: practice

Remarkable property. In practice, simplex algorithm typically terminates after at most 2(m + n) pivots.



but no polynomial pivot rule known

#### Issues.

- Avoid stalling.
- Choose the pivot.
- Maintain sparsity.
- Ensure numerical stability.
- Preprocess to eliminate variables and constraints.

Commercial solvers can solve LPs with millions of variables and tens of thousands of constraints.