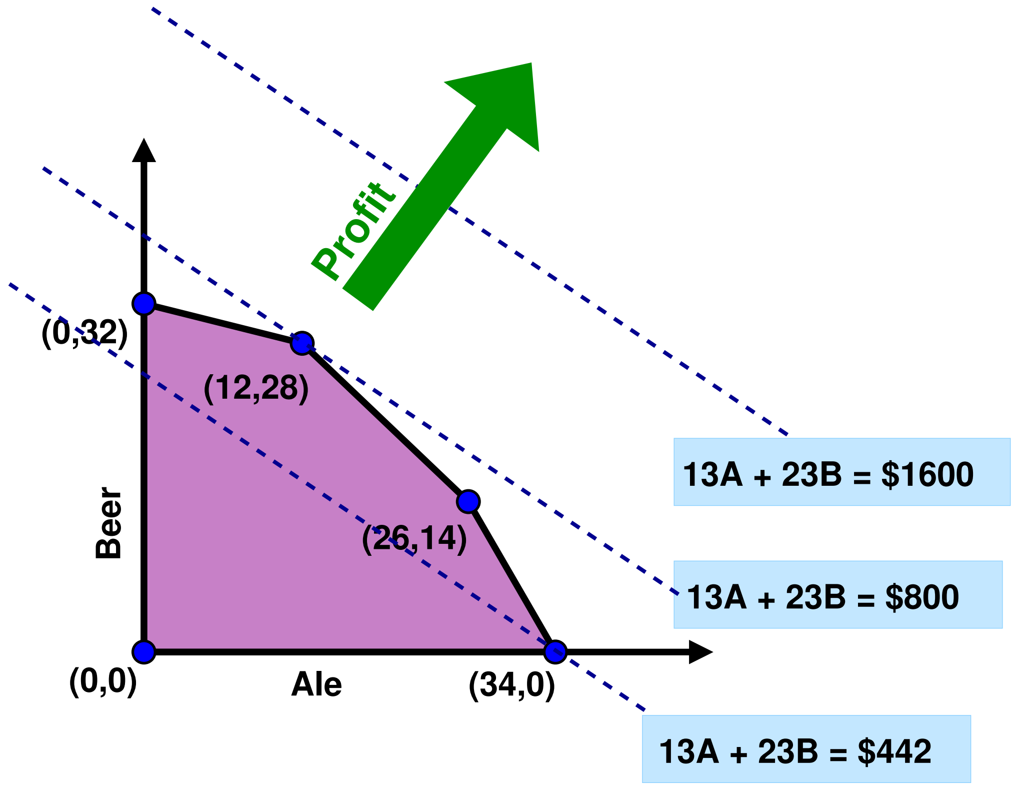
Chapter 3

Linear Programming



Linear Programming: What is it?

- Tool for optimal **allocation of scarce resources**, among a number of **competing activities**.
- Mathematical field of study concerned with such allocation questions, part of **operations research.**

Example: Small brewery produces ale and beer.

- Production limited by resources (raw materials) that are in short supply: corn, hops, barley malt.
- Recipes for *ale* and *beer* require **different proportions** of resources.

Beverage	Corn	Hops	Malt	**Profit ($)**
Ale	5	4	35	13
Beer	15	4	20	23
Quantity	480	160	1190	

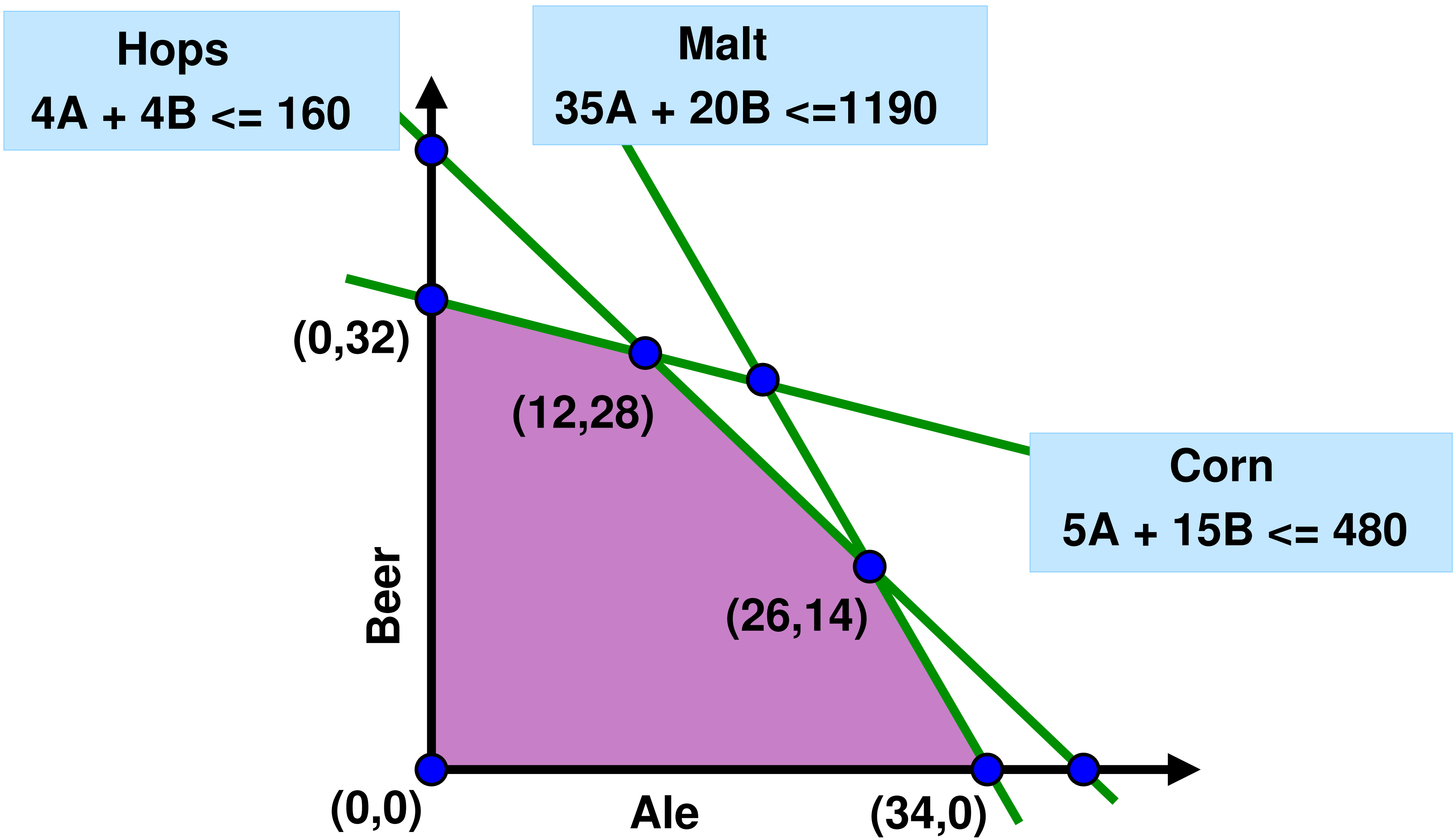
Robert Bland, *Allocation of Resources by Linear Programming*,Scientific American, Vol. 244, No. 6, June 1981

How can the brewer maximize profits?

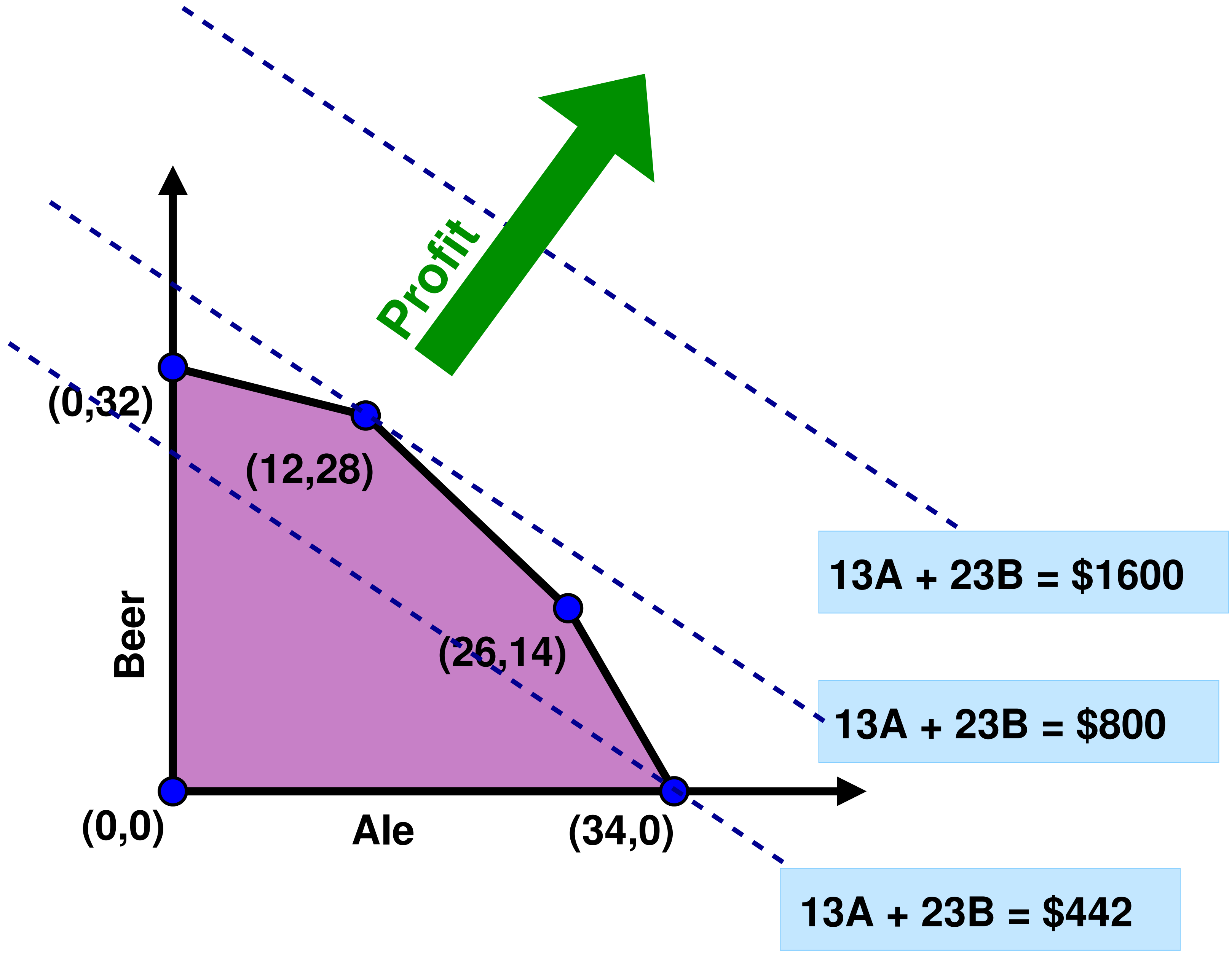
- Devote all resources to ale: 34 barrels of ale (all malt used up, long before supplies of hops and corn are exhausted): $A = 34 \Rightarrow \$442$.
- Devote all resources to beer: 32 barrels of beer (no more corn left): $B = 32 \Rightarrow \$736$.
- 7.5 barrels of ale, 29.5 barrels of beer $\Rightarrow \$776$.
- 12 barrels of ale, 28 barrels of beer (all corn and hops used) $\Rightarrow \$800$.

	Ale	**Beer**		**(products)**
maximize	$13A$	$+23B$		**(profit)**
s.t.	$5A$	$+15B$	≤ 480	**(corn)**
	$4A$	$+4B$	≤ 160	**(hops)**
	$35A$	$+20B$	≤ 1190	**(malt)**
	$A,$	B	≥ 0	**(physical constraints)**

Brewery Problem: Feasible Region

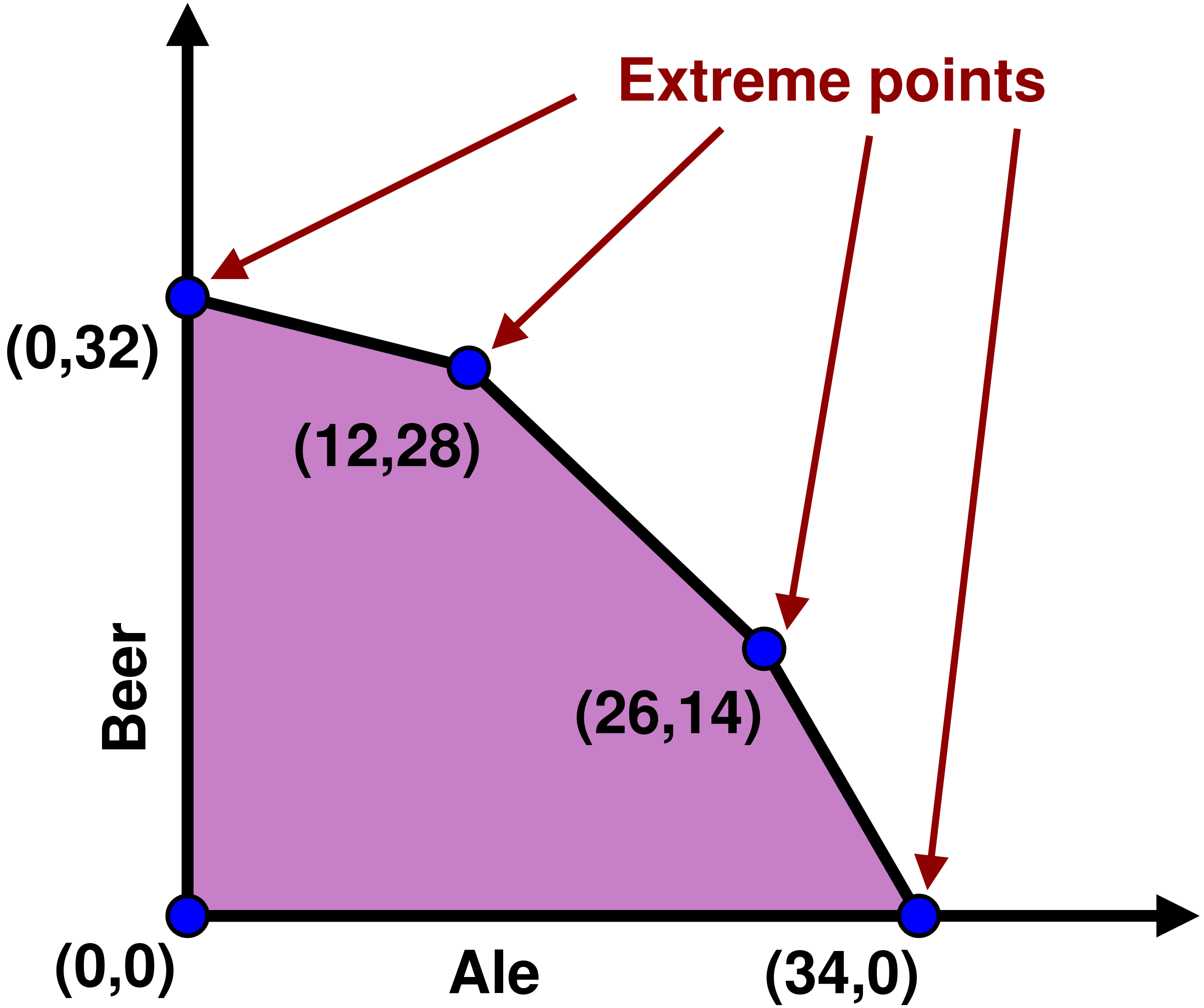


Brewery Problem: Objective Function



Brewery Problem: Geometry

Regardless of objective function coefficients, an optimal solution occurs at an **extreme point**.



Standard Form LP

- Input: real numbers c_j, b_i, a_{ij} .
- Output: real numbers x_j.
- $n = \#$ nonnegative variables, $m = \#$ constraints.
- Maximize linear objective function subject to linear equalities and physical constraints.

$$\begin{aligned} \max \quad & \sum_{j=1}^{n} c_j x_j \\ \text{s.t.} \quad & \sum_{j=1}^{n} a_{ij} x_j = b_i,\ 1 \le i \le m \\ & x_j \ge 0,\ 1 \le j \le n \end{aligned}$$

$$\begin{aligned} \max \quad & \boldsymbol{c}^t \boldsymbol{x} \\ \text{s.t.} \quad & A\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \ge \boldsymbol{0} \end{aligned}$$

Linear: Ressources needed and profit proportional to production.

Programming: Planning (not computer programming).

Brewery Problem: Converting to Standard Form

Original input:

$$\begin{aligned} \max \quad & 13A + 23B \\ \text{s.t.} \quad & 5A + 15B \leq 480 \\ & 4A + 4B \leq 160 \\ & 35A + 20B \leq 1190 \\ & A, B \geq 0 \end{aligned}$$

Standard form:

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

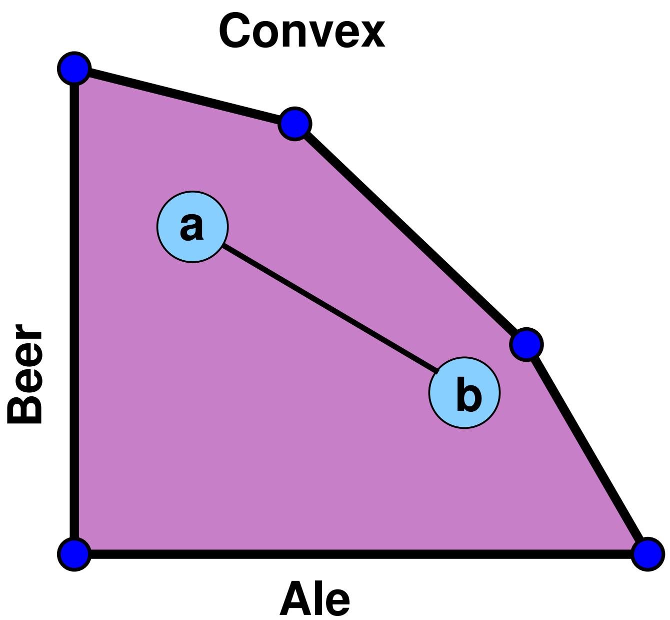
$$\begin{aligned} \max \quad & 13A + 23B \\ \text{s.t.} \quad & 5A + 15B + S_C = 480 \\ & 4A + 4B + S_H = 160 \\ & 35A + 20B + S_M = 1190 \\ & A, B, S_C, S_H, S_M \geq 0 \end{aligned}$$

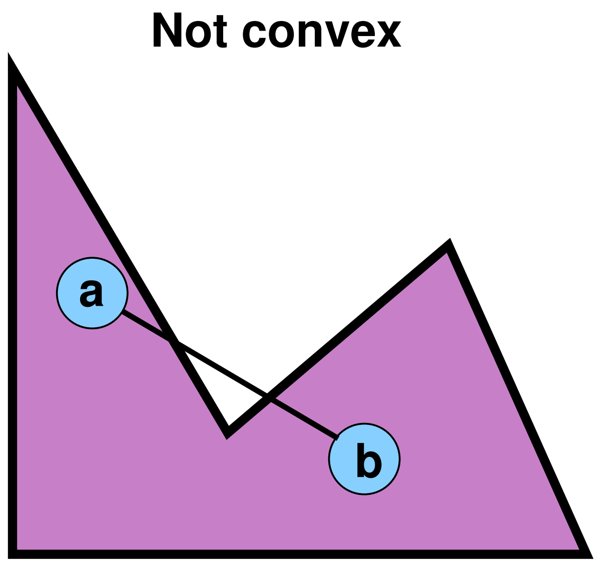
Geometry

- Inequalities: halfplanes (2D), hyperplanes.
- Bounded feasible region: convex polygon (2D), (convex) polytope.

Convex: if a and b are feasible solutions, then so is $(a+b)/2$.

Extreme point: feasible solution x that can't be written as $(a+b)/2$ for any two distinct feasible solutions a and b.



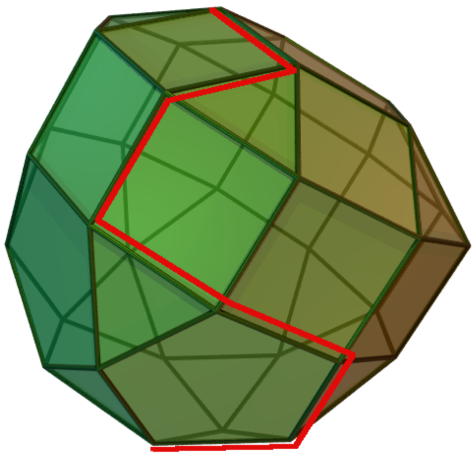


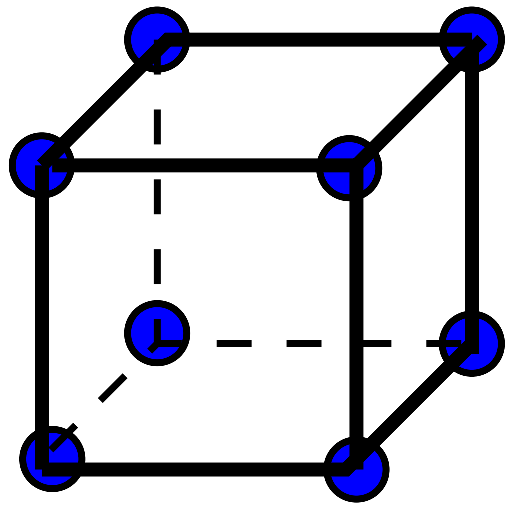
Geometry

Extreme point property. If there exists an optimal solution, then there exists one that is an extreme point. Only need to consider finitely many possible solutions.

Challenge. Number of extreme points can be exponential!
Consider n-dimensional hypercube: $2n$ equations, 2^n vertices.

Greed. Local optima are global optima. Extreme point is optimal if no neighboring extreme point is better.

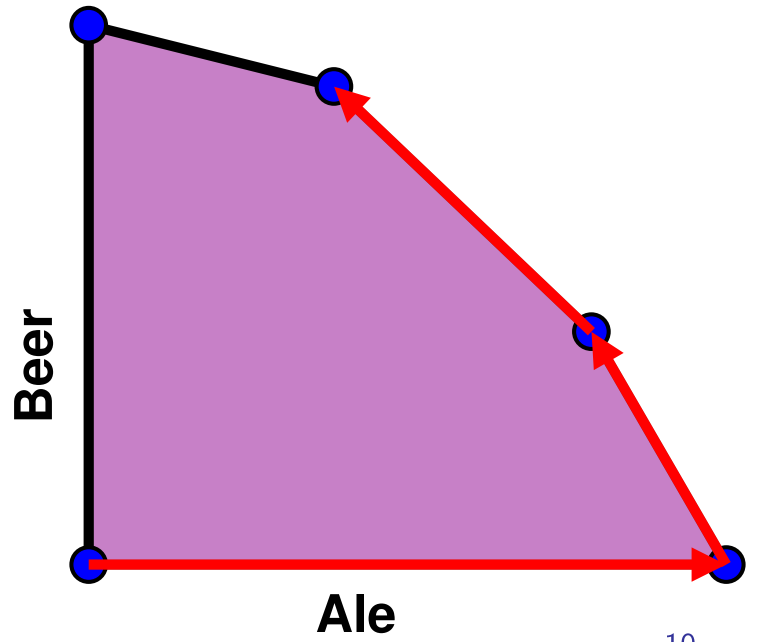




Simplex Algorithm (George Dantzig, 1947)

- Developed shortly after WWII in response to logistical problems.
- Generic algorithm, never decreases objective function.
- Start at some extreme point.
- Pivot from one extreme point to a neighboring one.
- Repeat until optimal.

How to implement?
Linear algebra.



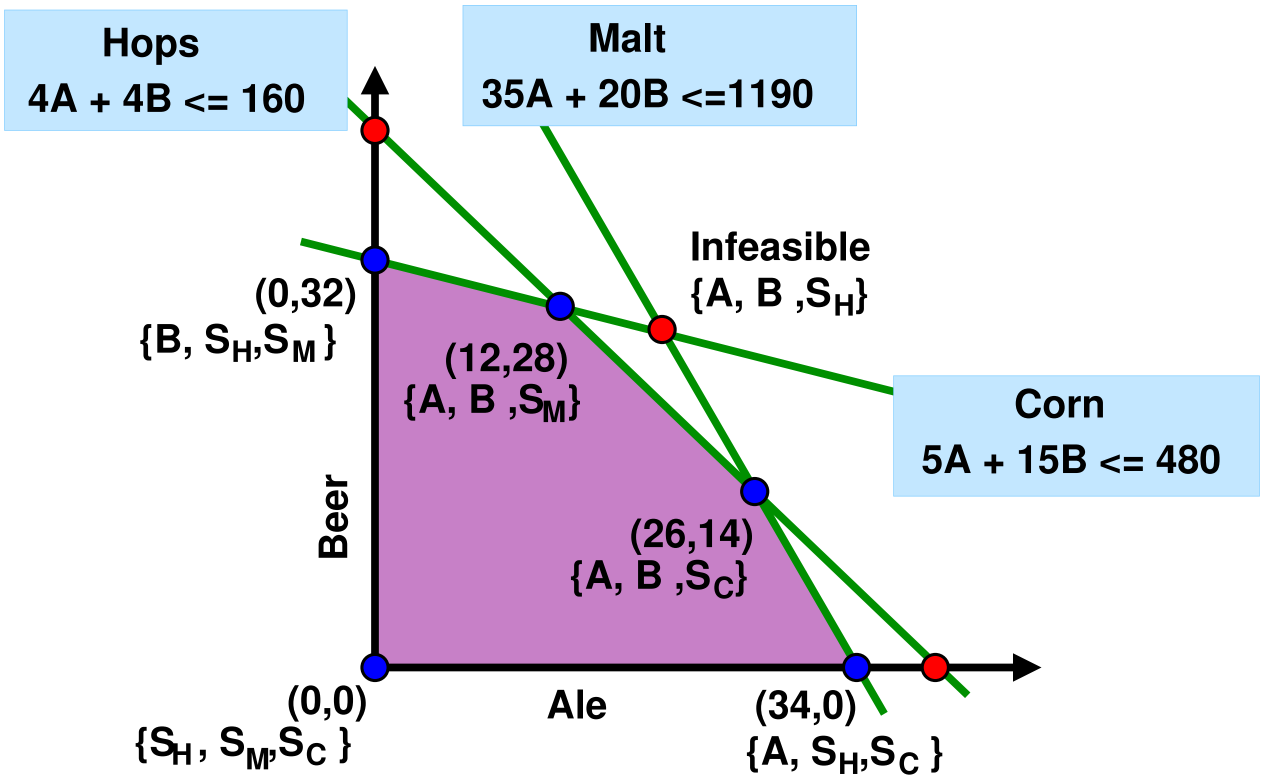
Simplex Algorithm: Basis

Basis: Subset of m of the $n' = n + m$ variables (n original $+$ m slack).

Basic feasible solution (BFS):

Set all $n' - m$ nonbasic variables to 0, solve for remaining m variables.

- Solve m equations in m unknowns.
- If **unique** and **feasible** solution $\Rightarrow$ BFS.
- BFS corresponds to **extreme point!** Simplex only considers BFS.



Simplex Algorithm: Pivot 1

$$\begin{array}{ll} \max & \text{obj} = 13A + 23B \\ \hline\hline \text{s.t.} & 5A + 15B + S_C = 480 \\ & 4A + 4B + S_H = 160 \\ & 35A + 20B + S_M = 1190 \\ & A, B, S_C, S_H, S_M \geq 0 \end{array}$$

Basis $= \{S_C, S_H, S_M\}$

$A = B = 0$

obj $= 0$

$S_C = 480$

$S_H = 160$

$S_M = 1190$

obj	$=$	0	$+$	13	A	$+$	23	B
S_C	$=$	480	$-$	5	A	$-$	15	B
S_H	$=$	160	$-$	4	A	$-$	4	B
S_M	$=$	1190	$-$	35	A	$-$	20	B

Which variable should enter next?

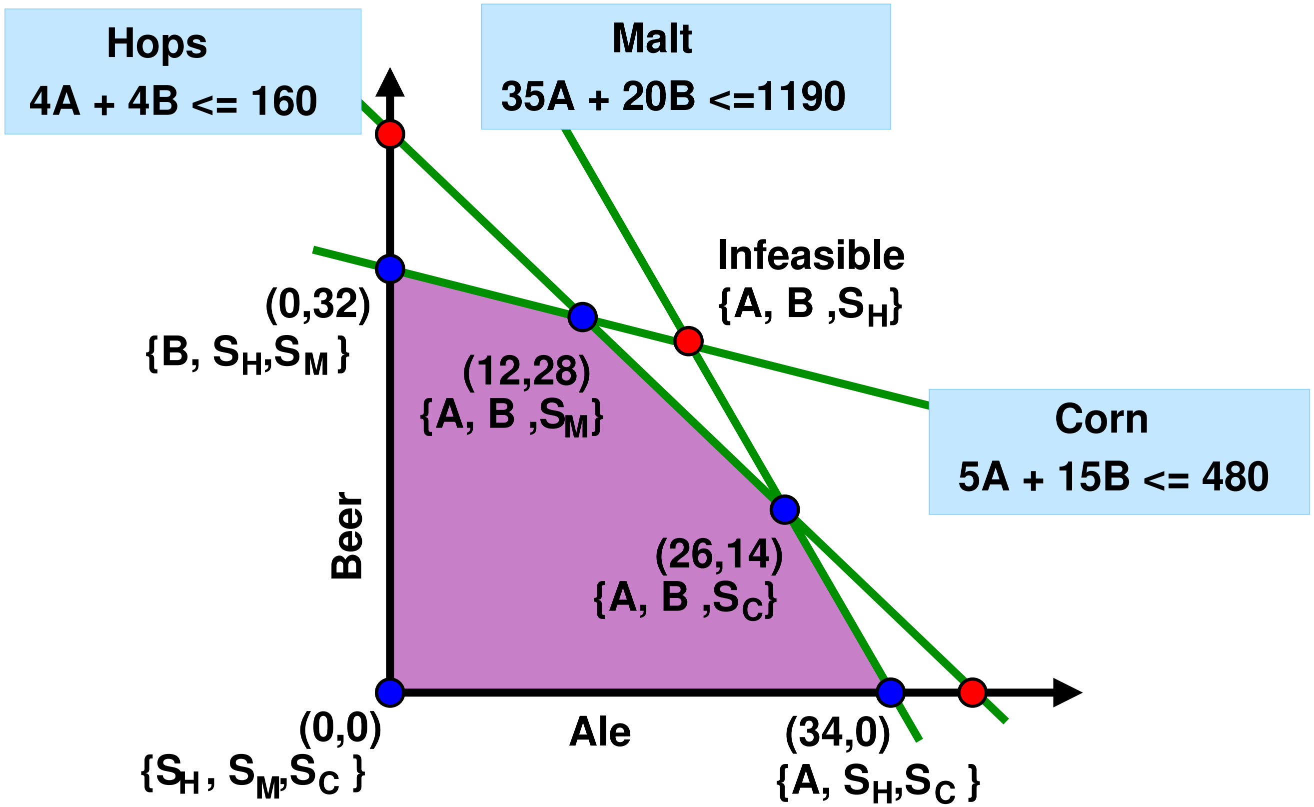
- Unit increase in $B \rightsquigarrow$ obj $+ \$23$.
- Letting A enter is also OK.

Simplex Algorithm: Selecting the Pivot Row

If B is increased, the first slack variable that becomes zero is S_C at $S_C = 480 - 15B = 0 \quad \Leftrightarrow \quad B = 480/15 = 32$ $\rightsquigarrow$ S_C **has to leave.**

What if S_H leaves (at $B = 160/4 = 40$)? Basis (B, S_C, S_M) outside the feasible region! Same problem if S_M leaves at $B = 1190/20 = 59.5$.

$\rightsquigarrow$ **Minimum ratio rule: min { 480/15, 160/4, 1190/20 }**



Simplex Algorithm: Pivot 1

obj	$=$	0	$+$	13	A	$+$	23	B
S_C	$=$	480	$-$	5	A	$-$	15	B
S_H	$=$	160	$-$	4	A	$-$	4	B
S_M	$=$	1190	$-$	35	A	$-$	20	B

B enters, S_C leaves $\rightsquigarrow$ solve pivot row $S_C = 480 - 5A - 15B$ for B:
Substitute $B = \frac{1}{15}(480 - 5A - S_C)$

obj	$=$	736	$+$	$16/3$	A	$+$	$-23/15$	S_C
B	$=$	32	$-$	$1/3$	A	$-$	$1/15$	S_C
S_H	$=$	32	$-$	$8/3$	A	$-$	$-4/15$	S_C
S_M	$=$	550	$-$	$85/3$	A	$-$	$-4/3$	S_C

Feasibility is preserved! (green highlights)

LP and Gauss-Jordan

$$\begin{bmatrix} 13 & 23 & 0 & 0 & 0 \\ 5 & 15 & 1 & 0 & 0 \\ 4 & 4 & 0 & 1 & 0 \\ 35 & 20 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ S_C \\ S_H \\ S_M \end{bmatrix} = \begin{bmatrix} \textbf{ojb} \\ 480 \\ 160 \\ 1190 \end{bmatrix} \underset{\text{augmented}}{\Rightarrow} \left[\begin{array}{cccccc} A & B & S_C & S_H & S_M & \\ \hline 13 & 23 & 0 & 0 & 0 & \textbf{obj} \\ 5 & 15 & 1 & 0 & 0 & 480 \\ 4 & 4 & 0 & 1 & 0 & 160 \\ 35 & 20 & 0 & 0 & 1 & 1190 \end{array}\right]$$

- Locate pivot element and save it: $piv = 15$
- Replace each row, except the pivot row, by that linear combination of itself and the pivot row which makes its pivot-column entry zero:

$$\left[\begin{array}{cccccc} A & B & S_C & S_H & S_M & \\ \hline 16/3 & \mathbf{0} & -23/15 & 0 & 0 & \textbf{obj} - 480 \cdot 23/15 \\ 5 & 15 & 1 & 0 & 0 & 480 \\ 8/3 & \mathbf{0} & 4/15 & 1 & 0 & 32 \\ 85/3 & \mathbf{0} & 4/3 & 0 & 1 & 550 \end{array}\right]$$

- Divide pivot row by piv: $1/3 \cdot A + B + 1/15 \cdot S_C + 0 + 0 = 32$

LP and Gauss-Jordan

- New basis (B, S_H, S_M):

$$\left[\begin{array}{cccccc} A & B & S_C & S_H & S_M & \\ \hline 16/3 & 0 & -23/15 & 0 & 0 & \textbf{obj} - 736 \\ 1/3 & 1 & 1/15 & 0 & 0 & 32 \\ 8/3 & 0 & 4/15 & 1 & 0 & 32 \\ 85/3 & 0 & 4/3 & 0 & 1 & 550 \end{array}\right]$$

- Corresponding tableau:

obj	=	736	+	16/3	A	+	−23/15	S_C
B	=	32	−	1/3	A	−	1/15	S_C
S_H	=	32	−	8/3	A	−	−4/15	S_C
S_M	=	550	−	85/3	A	−	−4/3	S_C

Simplex Algorithm: Pivot 2

obj	$=$	736	$+$	$16/3$	A	$+$	$-23/15$	S_C
B	$=$	32	$-$	$1/3$	A	$-$	$1/15$	S_C
S_H	$=$	32	$-$	$8/3$	A	$-$	$-4/15$	S_C
S_M	$=$	550	$-$	$85/3$	A	$-$	$-4/3$	S_C

Next pivot: A enters (only one magenta highlight left),
S_H leaves $\rightsquigarrow \min(32 \cdot 3, 32 \cdot 3/8, 330 \cdot 3/85)$:
Substitute $A = \frac{3}{8}(32 + \frac{4}{15}S_C - S_H)$

obj	$=$	800	$+$	-1	S_C	$+$	-2	S_H
B	$=$	28	$-$	$1/10$	S_C	$-$	$-1/8$	S_H
A	$=$	12	$-$	$-1/10$	S_C	$-$	$3/8$	S_H
S_M	$=$	210	$-$	$3/2$	S_C	$-$	$-85/8$	S_H

Feasibility is preserved!

Simplex Algorithm: Optimality

obj	=	800	+	-1	S_C	+	-2	S_H

obj	=	800	+	-1	S_C	+	-2	S_H
B	=	28	−	$1/10$	S_C	−	$-1/8$	S_H
A	=	12	−	$-1/10$	S_C	−	$3/8$	S_H
S_M	=	210	−	$3/2$	S_C	−	$-85/8$	S_H

When to stop pivoting? all coefficients in top row ≤ 0.

Why is resulting solution optimal?

- Any feasible solution satisfies system of equations in tableaux.
  in particular: obj $= 800 - S_C - 2S_H$
- Thus, optimal objective value obj ≤ 800 since $S_C, S_H \geq 0$.
- Current BFS has value 800 $\Rightarrow$ optimal (no further magenta highlights).
- At optimum: 28 (barrels of) Beer, 12 Ale, 210 units of Malt are left.

Simplex Algorithm: Problems and properties

Degeneracy. Pivot gives new basis, but same objective function value.

Cycling. A cycle is a sequence of degenerate pivots that returns to the first tableau in the sequence.

There exist pivoting rules for which **no cycling is possible**,
for instance **Bland's least index rule:**
"choose leftmost column with positive cost + min. ratio rule"

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

- Most pivot rules known to be **exponential in the worst-case.**
- No polynomial pivot rule known $\rightsquigarrow$ still an open question.

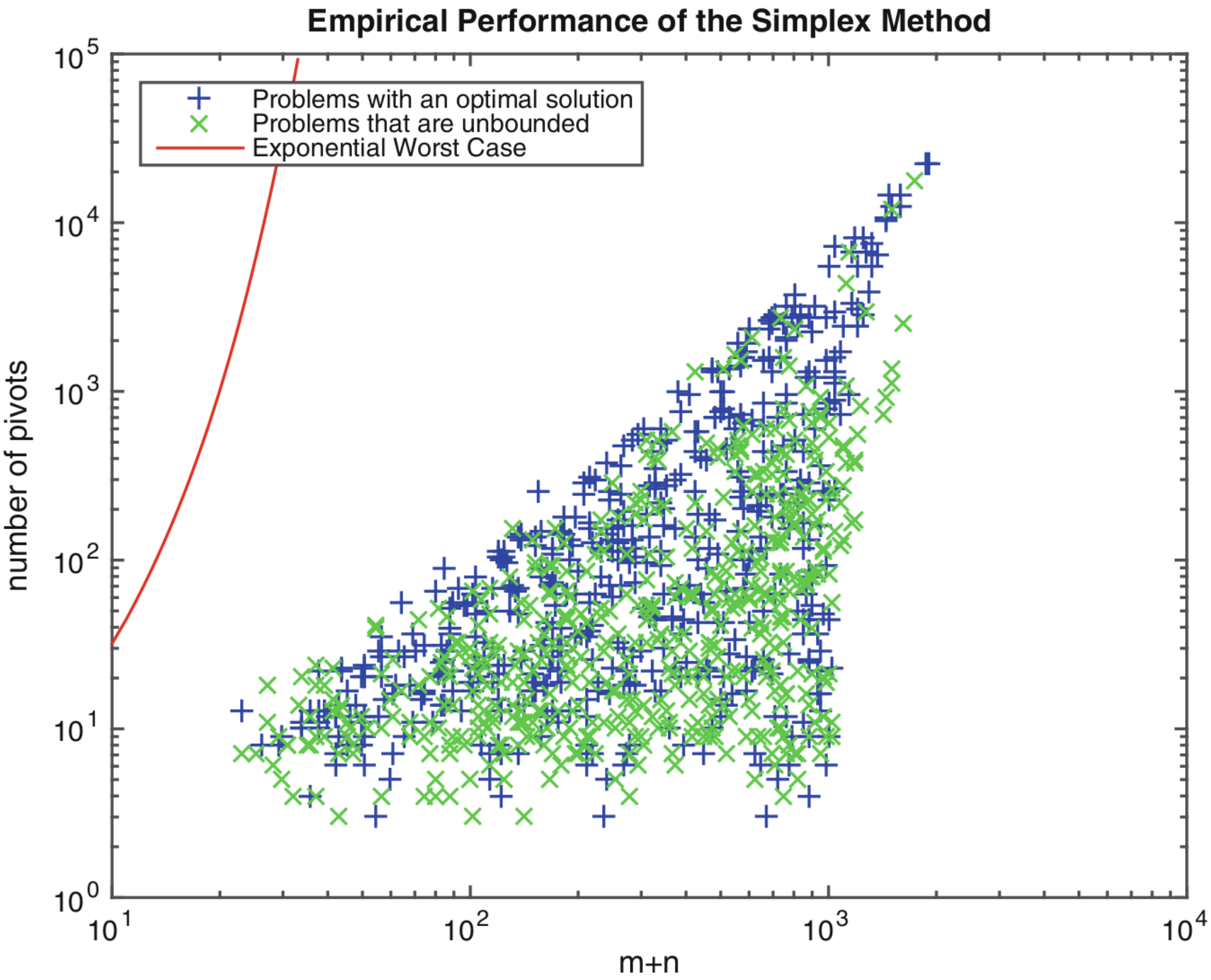


Figure 4.1 in Robert J. Vanderbei: Linear Programming, Springer. https://doi.org/10.1007/978-3-030-39415-8

Efficiency

Upper bound on the number of iterations is simply the number of basic feasible solutions, of which there can be at most

$$\binom{n+m}{m}$$

For fixed $n + m$, this expression is maximized when $m = n$.

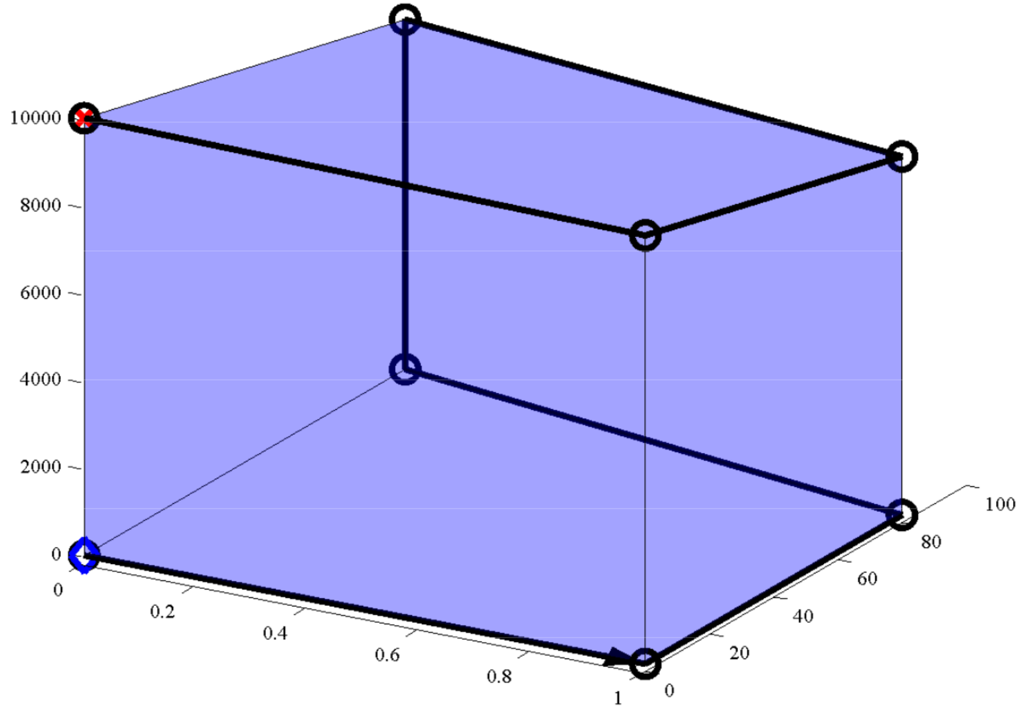
And how big is it? Exponentially big!
(simplified) Stirling's approximation: $\log n! \approx n \log n - n$

$$\log \binom{2n}{n} = \log \frac{(2n)!}{(n!)^2} = \log(2n)! - 2\log n! \approx 2n \log 2n - 2n \log n = 2n \log 2 = 2n$$

For LPs, there exist **Interior-Point algorithms** with guaranteed **polynomial runtime** (Karmarkar, '84). Researchers spent years trying to prove that the simplex worst-case complexity was polynomial...

Efficiency

...but the '72 **Klee-Minty counter-example** killed such hopes!
For most pivot rules there has been a KM-type counter-example.



No pivot rule guaranteed to yield worst-case polynomial time yet.
Yet practical performance definitely competitive (much better than most Interior Point methods!)

Efficiency: Different analysis concepts

- Let x be a problem instance, $T(x)$ the finishing time of Simplex alg. Think of "problem instance" as the matrix A in a LP problem. **Worst Case analysis:** $max_x T(x)$.
- Given random problems, what are the average finishing times? $\rightsquigarrow$ **Average Case analysis:** $E_{r \sim P(r)} T(r)$.

  Topic of intense study in 70' and 80's.
  Results: polynomal average case complexity.
- Given a problem that is randomly perturbed, what is the finishing time when averaged over all perturbations? $\rightsquigarrow$ **Smoothed analysis:** $max_x E_{r \sim P(r)} T(x + \epsilon r)$.
  - Interpolate between Worst Case and Average Case
  - Consider neighborhood of **every** input instance
  - If low, have to be unlucky to find bad input instance.

Efficiency: good news

Spielman-Teng '01: Coefficients of A perturbed by Gaussian noise with variance σ^2. Average complexity of solving such LP is at most a polynomial of n, m, σ^2 for every A.

You need to be very unlucky to find a bad LP input instance!

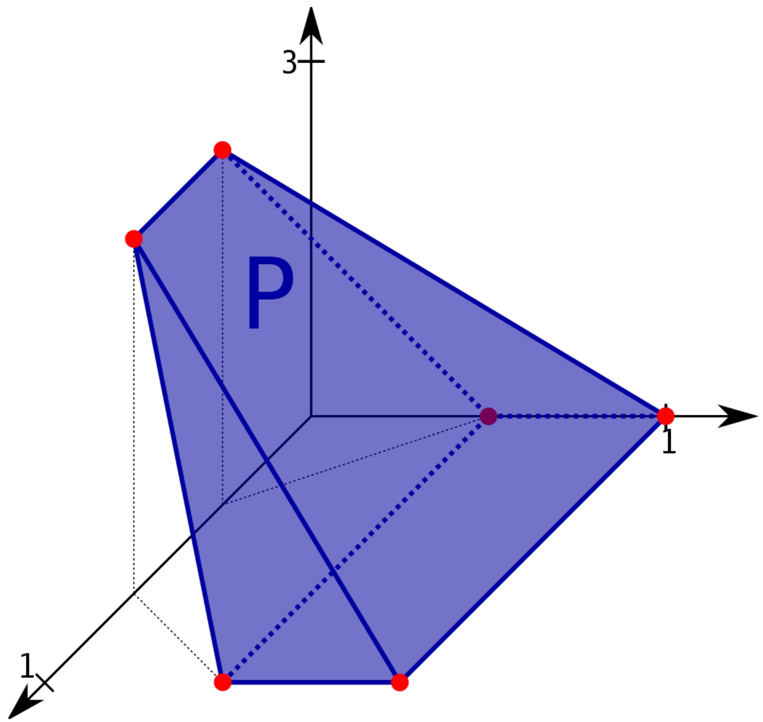
Smoothed Analysis of Algorithms: Why the Simplex Algorithm Usually Takes Polynomial Time

Daniel A. Spielman*
Department of Mathematics
M.I.T.
Cambridge, MA 02139
spielman@mit.edu

Shang-Hua Teng†
Akamai Technologies Inc. and
Department of Computer Science
University of Illinois at Urbana-Champaign
steng@cs.uiuc.edu

Further Questions

- **Unboundedness:** how can we check if optimal objective value is finite?
- **Initialization/infeasibility:** what to do if initial basis consisting of slack variables only is not feasible?
  $\rightsquigarrow$ **Phase-I / Phase-II Simplex Method**



- **LP Duality:** is there even more information in the final tableau?

Initialization

Recall our brewery problem: (Slack variables denoted by w_i):

$$\begin{array}{lrcrrcr} \text{maximize} & 13x_1 & + & 23x_2 & & & \\ \text{subject to} & 5x_1 & + & 15x_2 & +w_1 & = & 480 \\ & 4x_1 & + & 4x_2 & +w_2 & = & 160 \\ & 35x_1 & + & 20x_2 & +w_3 & = & 1190 \\ & & & x_1, x_2, & w_1, w_2, w_3 & \geq & 0 \end{array}$$

obj	$=$	0	$+$	13	x_1	$+$	23	x_2
w_1	$=$	480	$-$	5	x_1	$-$	15	x_2
w_2	$=$	160	$-$	4	x_1	$-$	4	x_2
w_3	$=$	1190	$-$	35	x_1	$-$	20	x_2

We were lucky...
...**positive values in constant column** show that the initial basis consisting of slack variables is **feasible** ($\Leftrightarrow$ for $x_1 = x_2 = 0$, all three slack variables are ≥ 0).

x_2 enters, w_1 leaves $\rightsquigarrow \min(480/15, 160/4, 1190/20)$.

The Brewery problem again

obj	$=$	736	$+$	$16/3$	x_1	$+$	$-23/15$	w_1
x_2	$=$	32	$-$	$1/3$	x_1	$-$	$1/15$	w_1
w_2	$=$	32	$-$	$8/3$	x_1	$-$	$-4/15$	w_1
w_3	$=$	550	$-$	$85/3$	x_1	$-$	$-4/3$	w_1

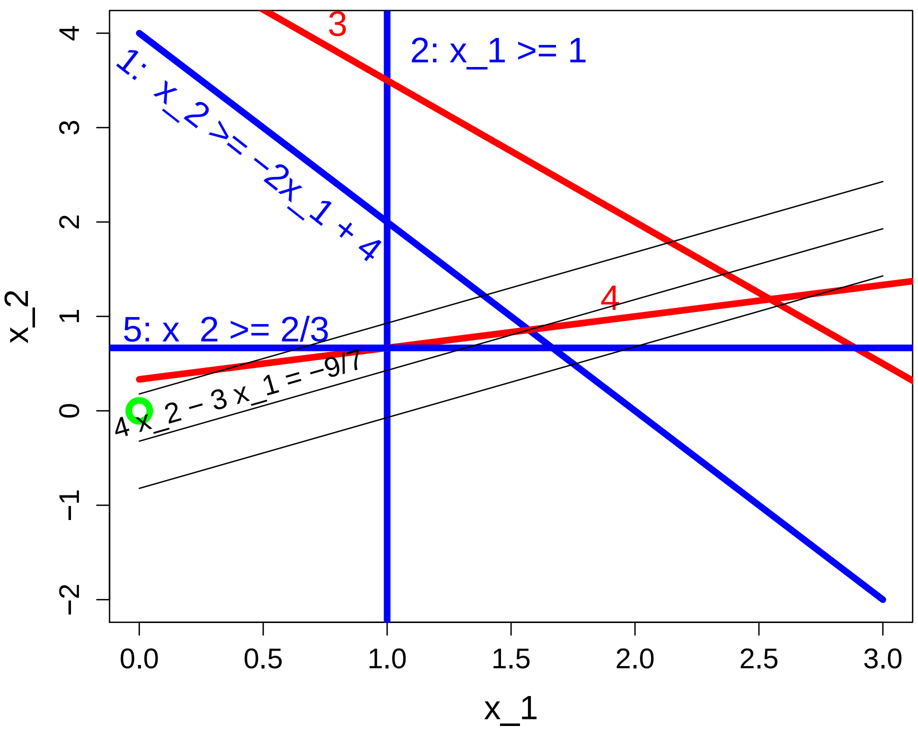
Feasibility is preserved!
x_1 enters, w_2 leaves $\rightsquigarrow \min(32 \cdot 3, 32 \cdot 3/8, 330 \cdot 3/85)$.

obj	$=$	800	$+$	-1	w_1	$+$	-2	w_2
x_2	$=$	28	$-$	$1/10$	w_1	$-$	$-1/8$	w_2
x_1	$=$	12	$-$	$-1/10$	w_1	$-$	$3/8$	w_2
w_3	$=$	210	$-$	$3/2$	w_1	$-$	$-85/8$	w_2

Feasibility is preserved!
Optimal! (no further magenta highlights in obj-row)

Initialization cont'd

$$
\begin{array}{lrcrcr}
\text{maximize} & -3x_1 & + & 4x_2 & & \\
\text{subject to} & -4x_1 & - & 2x_2 & \leq & -8 \\
 & -2x_1 & & & \leq & -2 \\
 & 3x_1 & + & 2x_2 & \leq & 10 \\
 & -x_1 & + & 3x_2 & \leq & 1 \\
 & & - & 3x_2 & \leq & -2 \\
 & & & x_1, x_2 & \geq & 0
\end{array}
$$



obj$_2$	$=$	0	$+$	-3	x_1	$+$	4	x_2
w_1	$=$	-8	$-$	-4	x_1	$-$	-2	x_2
w_2	$=$	-2	$-$	-2	x_1	$-$	0	x_2
w_3	$=$	10	$-$	3	x_1	$-$	2	x_2
w_4	$=$	1	$-$	-1	x_1	$-$	3	x_2
w_5	$=$	-2	$-$	0	x_1	$-$	-3	x_2

Initial basis is not feasible! $\rightsquigarrow$ Phase-I Problem

Phase-I Problem

Idea: Modify problem by subtracting a new variable, x_0, from each constraint and replace objective function with $-x_0$.

$$
\begin{array}{lrcrcrcr}
\text{maximize} & -x_0 & & & & & & \\
\text{subject to} & -x_0 & - & 4x_1 & - & 2x_2 & \leq & -8 \\
& -x_0 & - & 2x_1 & & & \leq & -2 \\
& -x_0 & + & 3x_1 & + & 2x_2 & \leq & 10 \\
& -x_0 & - & x_1 & + & 3x_2 & \leq & 1 \\
& -x_0 & & & - & 3x_2 & \leq & -2 \\
& & & & & x_0, x_1, x_2 & \geq & 0
\end{array}
$$

- Can always be made feasible: pick x_0 large, set $x_1 = 0$ and $x_2 = 0$.
- If optimal solution has **obj**$_1 = 0$, then the **original problem is feasible!** Note that obj$_1 = 0$ means that the "correction term" $x_0 = 0$, so the current point (x_1, x_2) must lie within the feasible region.
- Final phase-I basis can be used as initial phase-II basis (ignoring x_0 thereafter).
- If optimal solution has **obj**$_1 < 0$, then **original problem is infeasible!**

Initialization: First Pivot

obj$_2$	=	0	$+$	0	x_0	$+$	-3	x_1	$+$	4	x_2
obj$_1$	=	0	$+$	-1	x_0	$+$	0	x_1	$+$	0	x_2
w_1	=	-8	$-$	-1	x_0	$-$	-4	x_1	$-$	-2	x_2
w_2	=	-2	$-$	-1	x_0	$-$	-2	x_1	$-$	0	x_2
w_3	=	10	$-$	-1	x_0	$-$	3	x_1	$-$	2	x_2
w_4	=	1	$-$	-1	x_0	$-$	-1	x_1	$-$	3	x_2
w_5	=	-2	$-$	-1	x_0	$-$	0	x_1	$-$	-3	x_2

- Current basis is infeasible even for Phase-I.
- One pivot needed to get feasible.
- Entering variable is x_0 (there is no other choice, and we already know that the problem can be made feasible for large enough x_0...).
- Leaving variable is the one whose current value is most negative, i.e. the most violated constraint (here: w_1). This guarantees that after the first pivot all constraints are fulfilled.

Initialization: Second Pivot

obj_2	$=$	0	$+$	0	w_1	$+$	-3	x_1	$+$	4	x_2
obj_1	$=$	-8	$+$	-1	w_1	$+$	4	x_1	$+$	2	x_2
x_0	$=$	8	$-$	-1	w_1	$-$	4	x_1	$-$	2	x_2
w_2	$=$	6	$-$	-1	w_1	$-$	2	x_1	$-$	2	x_2
w_3	$=$	18	$-$	-1	w_1	$-$	7	x_1	$-$	4	x_2
w_4	$=$	9	$-$	-1	w_1	$-$	3	x_1	$-$	5	x_2
w_5	$=$	6	$-$	-1	w_1	$-$	4	x_1	$-$	-1	x_2

- Feasible!
- Focus on the yellow highlights.
- Let x_1 enter.
- Then w_5 must leave.
- After second pivot...

Initialization: Third Pivot

obj$_2$		$=$	-4.5	$+$	-0.75	w_1	$+$	0.75	w_5	$+$	3.25	x_2
obj$_1$		$=$	-2	$+$	0	w_1	$+$	-1	w_5	$+$	3	x_2
x_0		$=$	2	$-$	0	w_1	$-$	-1	w_5	$-$	3	x_2
w_2		$=$	3	$-$	-0.5	w_1	$-$	-0.5	w_5	$-$	2.5	x_2
w_3		$=$	7.5	$-$	0.75	w_1	$-$	-1.75	w_5	$-$	5.75	x_2
w_4		$=$	4.5	$-$	-0.25	w_1	$-$	-0.75	w_5	$-$	5.75	x_2
x_1		$=$	1.5	$-$	-0.25	w_1	$-$	0.25	w_5	$-$	-0.25	x_2

- x_2 must enter
- Then x_0 must leave.
- After third pivot...

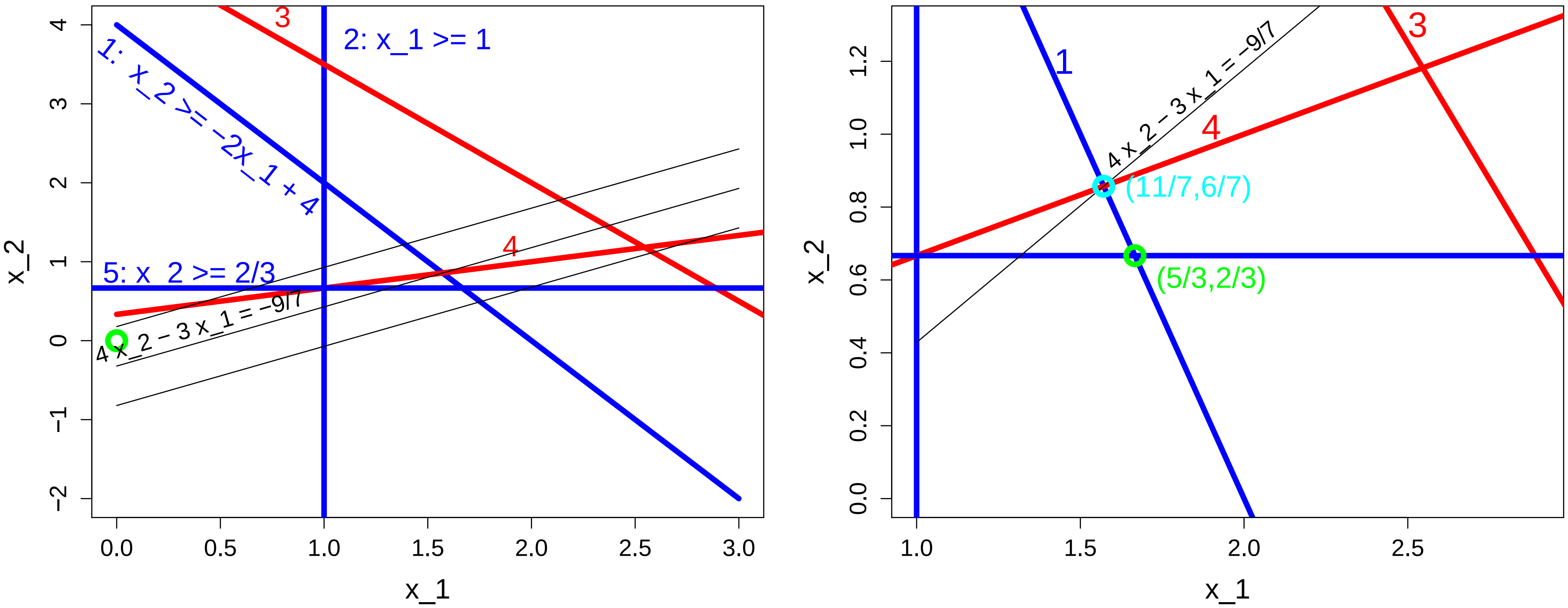
End of Phase-I, Begin of Phase-II

obj_2	$=$	$-\frac{7}{3}$	$+$	$-\frac{3}{4}$	w_1	$+$	$\frac{11}{6}$	w_5	$+$	0	x_0
obj_1	$=$	0	$+$	0	w_1	$+$	0	w_5	$+$	0	x_0
x_2	$=$	$\frac{2}{3}$	$-$	0	w_1	$-$	$-\frac{1}{3}$	w_5	$-$	0	x_0
w_2	$=$	$\frac{4}{3}$	$-$	$-\frac{1}{2}$	w_1	$-$	$\frac{1}{3}$	w_5	$-$	0	x_0
w_3	$=$	$\frac{11}{3}$	$-$	$\frac{3}{4}$	w_1	$-$	$\frac{1}{6}$	w_5	$-$	0	x_0
w_4	$=$	$\frac{2}{3}$	$-$	$-\frac{1}{4}$	w_1	$-$	$\frac{7}{6}$	w_5	$-$	0	x_0
x_1	$=$	$\frac{5}{3}$	$-$	$-\frac{1}{4}$	w_1	$-$	$\frac{1}{6}$	w_5	$-$	0	x_0

- Optimal for Phase-I (no yellow highlights).
- $\text{obj}_1 = 0$, therefore original problem is feasible.
- **For Phase-II:** Ignore column with x_0 and Phase-I objective row.
- w_5 must enter. w_4 must leave...

Phase-II: Optimal Solution

obj_2	$=$	$-\frac{9}{7}$	$+$	$-\frac{5}{14}$	w_1	$+$	$-\frac{11}{7}$	w_4
x_2	$=$	$\frac{6}{7}$	$-$	$-\frac{1}{14}$	w_1	$-$	$\frac{2}{7}$	w_4
w_2	$=$	$\frac{8}{7}$	$-$	$-\frac{3}{7}$	w_1	$-$	$-\frac{2}{7}$	w_4
w_3	$=$	$\frac{25}{7}$	$-$	$\frac{11}{14}$	w_1	$-$	$-\frac{1}{7}$	w_4
w_5	$=$	$\frac{4}{7}$	$-$	$-\frac{3}{14}$	w_1	$-$	$\frac{6}{7}$	w_4
x_1	$=$	$\frac{11}{7}$	$-$	$-\frac{3}{14}$	w_1	$-$	$-\frac{1}{7}$	w_4



Unboundedness

Consider the following tableau:

obj	$=$	0	$+$	2	x_1	$+$	-1	x_2	$+$	1	x_3
w_1	$=$	4	$-$	-5	x_1	$-$	3	x_2	$-$	-1	x_3
w_2	$=$	10	$-$	-1	x_1	$-$	-5	x_2	$-$	2	x_3
w_3	$=$	7	$-$	0	x_1	$-$	-4	x_2	$-$	3	x_3
w_4	$=$	6	$-$	-2	x_1	$-$	-2	x_2	$-$	4	x_3
w_5	$=$	6	$-$	-3	x_1	$-$	0	x_2	$-$	-3	x_3

- Could increase either x_1 or x_3 to increase obj.
- Consider increasing x_1.
- Which basic variable decreases to zero first?
- Answer: none of them, x_1 can grow without bound, and obj along with it.
- This is how we detect **unboundedness** with the simplex method.

The Two Phase Simplex Algorithm

Phase I: Formulate and solve the auxiliary problem.
Two outcomes are possible:

- The optimal value of x_0 in the auxiliary problem is positive.
  In this case the **original problem is infeasible.**

- The optimal value is zero and an **initial feasible tableau** for the original problem is obtained.

Phase II: If the original problem is feasible, apply the simplex algorithm to the initial feasible tableau obtained from Phase I above.
Again, two outcomes are possible:

- The LP is **unbounded**.

- An **optimal basic feasible solution** is obtained.

The Fundamental Theorem of linear Programming

Theorem: Every LP has the following three properties:

- If it has no optimal solution, then it is either infeasible or unbounded.
- If it has a feasible solution, then it has a basic feasible solution.
- If it is bounded, then it has an optimal basic feasible solution.

Proof: Phase I algorithm either proves that the problem is infeasible or produces a basic feasible solution. Phase II algorithm either discovers that the problem is unbounded or finds a basic optimal solution.
Assumption: no cycling occurs, guaranteed by several pivot rules.

Bland's rule:
Entering: choose the lowest-numbered nonbasic column with a positive coefficient.
Leaving: in case of ties in the ratio test, choose the leaving basic variable with the smallest index.

Primal problem: Ressource allocation

Brewer's problem: find optimal mix to maximize profits.

$$\begin{aligned} \max \quad & 13A + 23B \\ \text{s.t.} \quad & 5A + 15B \leq 480 \\ & 4A + \ \ 4B \leq 160 \\ & 35A + 20B \leq 1190, \quad A, B \geq 0 \end{aligned}$$

$$A^* = 12 \qquad B^* = 28 \qquad \text{OPT} = 800$$

General form: Find optimal allocation of m raw materials to n production processes. This is the primal $\mathcal{P}$: Given real numbers

- a_{ij} = units of raw material i needed to produce one unit of product j,
- b_i = ressource constraints for raw material i, $i = 1, \ldots, m$,
- c_j = profit per unit of product j, $j = 1, \ldots, n$,

$$\begin{array}{llll|lll} \text{maximize}_{\boldsymbol{x}} & \sum_{j=1}^{n} c_j x_j & & & \boldsymbol{c}^t \boldsymbol{x} & & \\ \text{subject to} & \sum_{j=1}^{n} a_{ij} x_j & \leq b_i, & i = 1, \ldots, m & A\boldsymbol{x} & \leq & \boldsymbol{b} \\ & x_j & \geq 0, & j = 1, \ldots, n & \boldsymbol{x} & \geq & \boldsymbol{0} \end{array}$$

The dual: Brewery example

- 5corn $+$ 4hops $+$ 35malt needed to brew one barrel of Ale (which would lead to profit of 13$\$$). If we produce one unit less of Ale, we free up $\{5/4/35\}$ units of {corn/hops/malt}.
- Selling for C, H, M dollars/unit yields $5C + 4H + 35M$ dollars.
- Only interested if this exceeds lost profit of 13$\$$: $5C+4H+35M \geq 13$. Similar for Beer: $15C + 4H + 20M \geq 23$.

Consider a buyer offering to purchase our entire inventory. Subject to above constraints, buyer wants to minimize cost. This is the dual $\mathcal{D}$:

Buyer's problem: Buy resources from brewer at minimum cost.

$$
\begin{aligned}
(\mathcal{D}) \quad \min \quad & 480C + 160H + 1190M \\
\text{s.t.} \quad & 5C + 4H + 35M \geq 13 \\
& 15C + 4H + 20M \geq 23 \\
& C, H, M \geq 0
\end{aligned}
$$

$$
\begin{aligned}
C^* &= 1 \\
H^* &= 2 \\
M^* &= 0 \\
\text{OPT} &= 800
\end{aligned}
$$

LP Duality

Every Problem $\mathcal{P}$: Given real numbers a_{ij}, b_i, c_j,

$$\begin{array}{llcll|lcl} \text{maximize}_{\boldsymbol{x}} & \sum_{j=1}^{n} c_j x_j & & & & \boldsymbol{c}^t\boldsymbol{x} & & \\ \text{subject to} & \sum_{j=1}^{n} a_{ij} x_j & \leq & b_i, & i=1,\ldots,m & A\boldsymbol{x} & \leq & \boldsymbol{b} \\ & x_j & \geq & 0, & j=1,\ldots,n & \boldsymbol{x} & \geq & \boldsymbol{0} \end{array}$$

Has a dual $\mathcal{D}$: Given real numbers a_{ij}, b_i, c_j,

$$\begin{array}{llcll|lcl} \text{minimize}_{\boldsymbol{y}} & \sum_{i=1}^{m} b_i y_i & & & & \boldsymbol{b}^t\boldsymbol{y} & & \\ \text{subject to} & \sum_{i=1}^{m} y_i a_{ij} & \geq & c_j, & j=1,\ldots,n & A^t\boldsymbol{y} & \geq & \boldsymbol{c} \\ & y_i & \geq & 0, & i=1,\ldots,m & \boldsymbol{y} & \geq & \boldsymbol{0} \end{array}$$

Duality Theorem (Dantzig-von Neumann 1947, Gale-Kuhn-Tucker 1951).
If ($\mathcal{P}$) and ($\mathcal{D}$) have feasible solutions, then max = min.

LP Duality: Economic Interpretation

Marginal (or Shadow-) prices:

Q. How much should brewer be willing to pay for additional supplies of scarce resources?

A.

obj	$=$	800	$+$	-1	S_C	$+$	-2	S_H

$\rightsquigarrow$ **Per unit changes in profit for changes in resources:**

$\rightsquigarrow$ corn $\$1$, hops $\$2$, malt $\$0$ (210 pounds of excess malt not utilized)

Q. New product "light beer" is proposed. It requires 2 corn, 5 hops, 24 malt. How much profit must be obtained from light beer to justify diverting resources from production of beer and ale?

A. Breakeven: 2 ($\$1$) $+$ 5 ($\$2$) $+$ 24 (0$\$$) $=$ $\$12$ / barrel.

How can we compute the shadow prices?

Simplex solves primal and dual simultaneously.

Top row of final simplex tableaux provides optimal dual solution!

Dual of Dual

Primal problem:

$$\begin{array}{lrcl} \text{maximize} & \boldsymbol{c}^t\boldsymbol{x} & & \\ \text{subject to} & A\boldsymbol{x} & \leq & \boldsymbol{b} \\ & x_j & \geq & 0, \quad j = 1, \ldots, n \end{array}$$

Note: A problem is defined by its data (notation used for the variables is arbitrary).

Dual in usual LP form:

$$\begin{array}{lrcl} \text{maximize} & (-\boldsymbol{b}^t)\boldsymbol{y} & & \\ \text{subject to} & (-A^t)\boldsymbol{y} & \leq & (-\boldsymbol{c}) \\ & y_i & \geq & 0, \quad i = 1, \ldots, m \end{array}$$

Dual is **negative transpose** of primal.

Theorem: Dual of dual is primal. Proof:

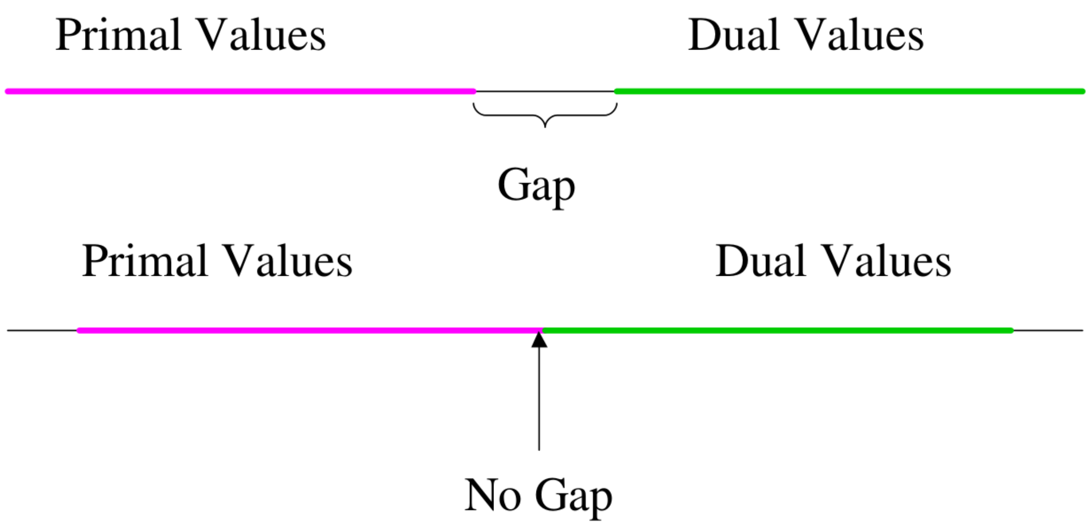
$$\boxed{\begin{array}{lrcl} \text{minimize} & (-\boldsymbol{c})^t\boldsymbol{x} & & \\ \text{subject to} & (-A^t)^t\boldsymbol{x} & \geq & (-\boldsymbol{b}) \end{array}} = \boxed{\begin{array}{lrcl} \text{maximize} & \boldsymbol{c}^t\boldsymbol{x} & & \\ \text{subject to} & A\boldsymbol{x} & \leq & \boldsymbol{b} \end{array}}$$

Weak Duality Theorem

If $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^t$ is feasible for the primal and $\boldsymbol{y} = (y_1, y_2, \ldots, y_m)^t$ is feasible for the dual, then $\boldsymbol{c}^t\boldsymbol{x} \le \boldsymbol{b}^t\boldsymbol{y}$

Proof: $\boldsymbol{c}^t\boldsymbol{x} \le \boldsymbol{y}^t A\boldsymbol{x} \le \boldsymbol{y}^t\boldsymbol{b}$.

An important question: Is there a gap between the largest primal value and the smallest dual value?



Answer is provided by the **Strong Duality Theorem:**

If ($\mathcal{P}$) and ($\mathcal{D}$) have feasible solutions, then $\max_{\mathcal{P}} = \min_{\mathcal{D}}$**.**

Simplex Method and Duality

A primal problem:

obj	$=$	0	$+$	-3	x_1	$+$	2	x_2	$+$	1	x_3
w_1	$=$	0	$-$	0	x_1	$-$	-1	x_2	$-$	2	x_3
w_2	$=$	3	$-$	3	x_1	$-$	4	x_2	$-$	1	x_3

Its dual:

obj	$=$	0	$+$	0	y_1	$+$	-3	y_2
z_1	$=$	3	$-$	0	y_1	$-$	-3	y_2
z_2	$=$	-2	$-$	1	y_1	$-$	-4	y_2
z_3	$=$	-1	$-$	-2	y_1	$-$	-1	y_2

Notes:

- Dual is negative transpose of primal.
- Primal is feasible, dual is not.

Use primal to choose pivot: x_2 enters, w_2 leaves.

Make analogous pivot in dual: z_2 leaves, y_2 enters.

Second Iteration

After First Pivot:

obj	$=$	$3/2$	$+$	$-3/2$	x_1	$+$	$-1/2$	w_2	$+$	$1/2$	x_3
w_1	$=$	$3/4$	$-$	$-3/4$	x_1	$-$	$1/4$	w_2	$-$	$9/4$	x_3
x_2	$=$	$3/4$	$-$	$-3/4$	x_1	$-$	$1/4$	w_2	$-$	$1/4$	x_3

Primal (feasible)

obj	$=$	$-3/2$	$+$	$-3/4$	y_1	$+$	$-3/4$	z_2
z_1	$=$	$3/2$	$-$	$3/4$	y_1	$-$	$3/4$	z_2
y_2	$=$	$1/2$	$-$	$-1/4$	y_1	$-$	$-1/4$	z_2
z_3	$=$	$-1/2$	$-$	$-9/4$	y_1	$-$	$-1/4$	y_2

Dual (still not feasible)

Note: negative transpose property intact.
Again, use primal to pick pivot: x_3 enters, w_1 leaves.
Make analogous pivot in dual: z_3 leaves, y_1 enters.

After Second Iteration

obj	=	$5/3$	$+$	$-4/3$	x_1	$+$	$-5/9$	w_2	$+$	$-2/9$	w_1
x_3	=	$1/3$	$-$	$-1/3$	x_1	$-$	$1/9$	w_2	$-$	$4/9$	w_1
x_2	=	$2/3$	$-$	$-2/3$	x_1	$-$	$2/9$	w_2	$-$	$-1/9$	w_1

Primal is optimal

obj	=	$-5/3$	$+$	$-1/3$	z_3	$+$	$-2/3$	z_2
z_1	=	$4/3$	$-$	$1/3$	z_3	$-$	$2/3$	z_2
y_2	=	$5/9$	$-$	$-1/9$	z_3	$-$	$-2/9$	z_2
y_1	=	$2/9$	$-$	$-4/9$	z_3	$-$	$1/9$	y_2

Dual: negative transpose property remains intact, is optimal.

Conclusion: Simplex method applied to primal problem (two phases, if necessary), solves both the primal and the dual.

This is the essence of the **strong duality theorem**:

If the primal problem has an optimal solution, $\boldsymbol{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)^t$, then the dual also has an optimal solution, $\boldsymbol{y}^* = (y_1^*, y_2^*, \ldots, y_m^*)^t$, and $\boldsymbol{c}^t\boldsymbol{x}^* = \boldsymbol{b}^t\boldsymbol{y}^*$.

Strong Duality Theorem

Theorem. If the primal problem has an optimal solution, $\boldsymbol{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)^t$, then the dual also has an optimal solution, $\boldsymbol{y}^* = (y_1^*, y_2^*, \ldots, y_m^*)^t$, and $\boldsymbol{c}^t\boldsymbol{x}^* = \boldsymbol{b}^t\boldsymbol{y}^*$

Proof idea: Combine the Fundamental Theorem of LP with the Weak Duality Theorem:

The **Fundamental LP Theorem** says that an optimal basic feasible solution exists.

Negative transpose property: primal optimality implies dual feasibility. The optimal tableau gives $\boldsymbol{x}^*$ and $\boldsymbol{y}^*$ with $\boldsymbol{c}^t\boldsymbol{x}^* = \boldsymbol{b}^t\boldsymbol{y}^*$ and $\boldsymbol{x}^*$ feasible for $\mathcal{P}$, $\boldsymbol{y}^*$ feasible for $\mathcal{D}$

Weak Duality theorem: Feasibility implies $\boldsymbol{b}^t\boldsymbol{y}^* \geq \boldsymbol{c}^t\boldsymbol{x}^*$, so equality implies that $\boldsymbol{y}^*$ solves $\mathcal{D}$.

Duality Gap

Direct consequence:

If primal has an optimal solution, there is no duality gap.

In general, there are four possibilities:

- Primal optimal, dual optimal (no gap).
- Primal unbounded, dual infeasible (no gap).
- Primal infeasible, dual unbounded (no gap).
- Primal infeasible, dual infeasible (infinite gap).

Example of infinite gap:

$\mathcal{P}$	$\mathcal{D}$
$\begin{array}{lrcrcr} \text{max} & 2x_1 & - & x_2 & & \\ \text{s.t} & x_1 & - & x_2 & \leq & 1 \\ & -x_1 & + & x_2 & \leq & -2 \\ & & & x_1, x_2 & \geq & 0. \end{array}$	$\begin{array}{lrcrcr} \text{max} & -y_1 & + & 2y_2 & & \\ \text{s.t.} & -y_1 & + & y_2 & \leq & -2 \\ & y_1 & - & y_2 & \leq & 1 \\ & & & y_1, y_2 & \geq & 0. \end{array}$
$x_1 \leq 1 + x_2$ $x_1 \geq 2 + x_2$	$y_1 \geq 2 + y_2$ $y_1 \leq 1 + y_2$

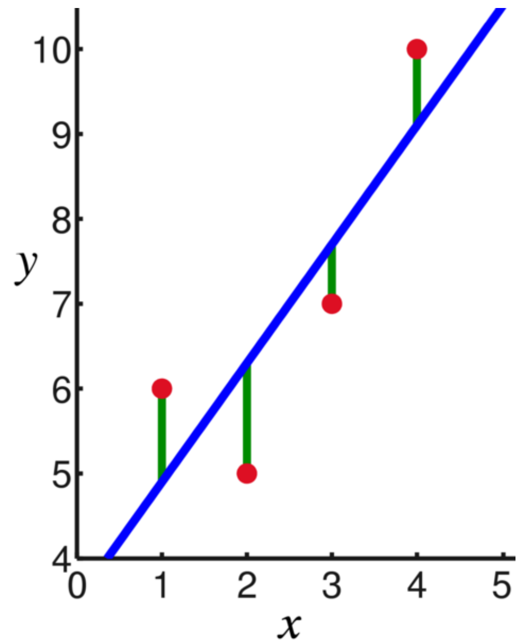
Complementary slackness

- Primal slack: $\boldsymbol{w} = \boldsymbol{b} - A\boldsymbol{x}$
- Weak Duality: $\boldsymbol{x}$, $\boldsymbol{y}$ feasible $\Rightarrow \boldsymbol{c}^t\boldsymbol{x} \leq \boldsymbol{y}^t A\boldsymbol{x} \leq \boldsymbol{y}^t\boldsymbol{b}$
- Strong Duality: $\boldsymbol{x}^*$, $\boldsymbol{y}^*$ optimal $\Rightarrow \boldsymbol{c}^t\boldsymbol{x}^* = (\boldsymbol{y}^*)^t A\boldsymbol{x} = (\boldsymbol{y}^*)^t\boldsymbol{b}$

$$
\begin{aligned}
(\boldsymbol{y}^*)^t\boldsymbol{w} &= (\boldsymbol{y}^*)^t\boldsymbol{b} - (\boldsymbol{y}^*)^t A\boldsymbol{x} = 0 \\
&= \sum_{j=1}^{m} y_j^* w_j, \quad \text{with } y_j^* \geq 0, w_j \geq 0,\ \forall j
\end{aligned}
$$

If there are "leftovers" (positive slack $w_k > 0$), additional quantities of this raw material have no value (shadow price $y_k^* = 0$).

Recall: Linear curve fitting



- **Notation:** n objects at locations $\boldsymbol{x}_i \in \mathbb{R}^p$.
  Every object has measurement $y_i \in \mathbb{R}$.
- **Approximate** "regression targets" y as a **parametrized function** of $\boldsymbol{x}$.
- Consider a 1-dim problem initially.
- Start with n data points (x_i, y_i), $i = 1, \ldots, n$.
- Choose d **basis functions** $g_0(x), g_1(x), \ldots$.
- Fitting to a **line** uses **two** basis functions $g_0(x) = 1$ and $g_1(x) = x$. In most cases $n \gg d$.
- **Fit function = linear combination of basis functions:**
  $f(x; \boldsymbol{w}) = \sum_j w_j g_j(x) = w_0 + w_1 x$.
- $f(x_i) = y_i$ exactly is (usually) **not possible**, so approximate $f(x_i) \approx y_i$
- n **residuals** are defined by $r_i = y_i - f(x_i) = y_i - (w_0 + w_1 x_i)$.

Recall: Basis functions

X has as many columns as there are basis functions. Examples:

- **High-dimensional linear functions**
  $\boldsymbol{x} \in \mathbb{R}^p$, $g_0(\boldsymbol{x}) = 1$ and $g_1(\boldsymbol{x}) = x_1, g_2(\boldsymbol{x}) = x_2, \ldots, g_p(\boldsymbol{x}) = x_p$.

$$X_{i\bullet} = \boldsymbol{g}^t(\boldsymbol{x}_i) = (1, \text{---}\ \boldsymbol{x}_i^t\ \text{---}), \quad (i\text{-th row of } X)$$

$$f(\boldsymbol{x}; \boldsymbol{w}) = \boldsymbol{w}^t \boldsymbol{g} = w_0 + w_1 x_1 + \cdots + w_p x_p.$$

- **Document analysis:** Assume a fixed collection of words:

$$\begin{aligned}
\boldsymbol{x} &= \text{text document} \\
g_0(\boldsymbol{x}) &= 1 \\
g_i(\boldsymbol{x}) &= \#(\text{occurences of } i\text{-th word in document}) \\
f(\boldsymbol{x}; \boldsymbol{w}) &= \boldsymbol{w}^t \boldsymbol{g} = w_0 + \sum_{i \in \text{words}} w_i g_i(\boldsymbol{x}).
\end{aligned}$$

Least absolute deviations regression

Least squares regression: $\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \|\boldsymbol{r}\|_2^2$

LAD-regression is less sensitive to outliers than least squares regression is. It is defined by minimizing the ℓ_1-norm of the residual vector.

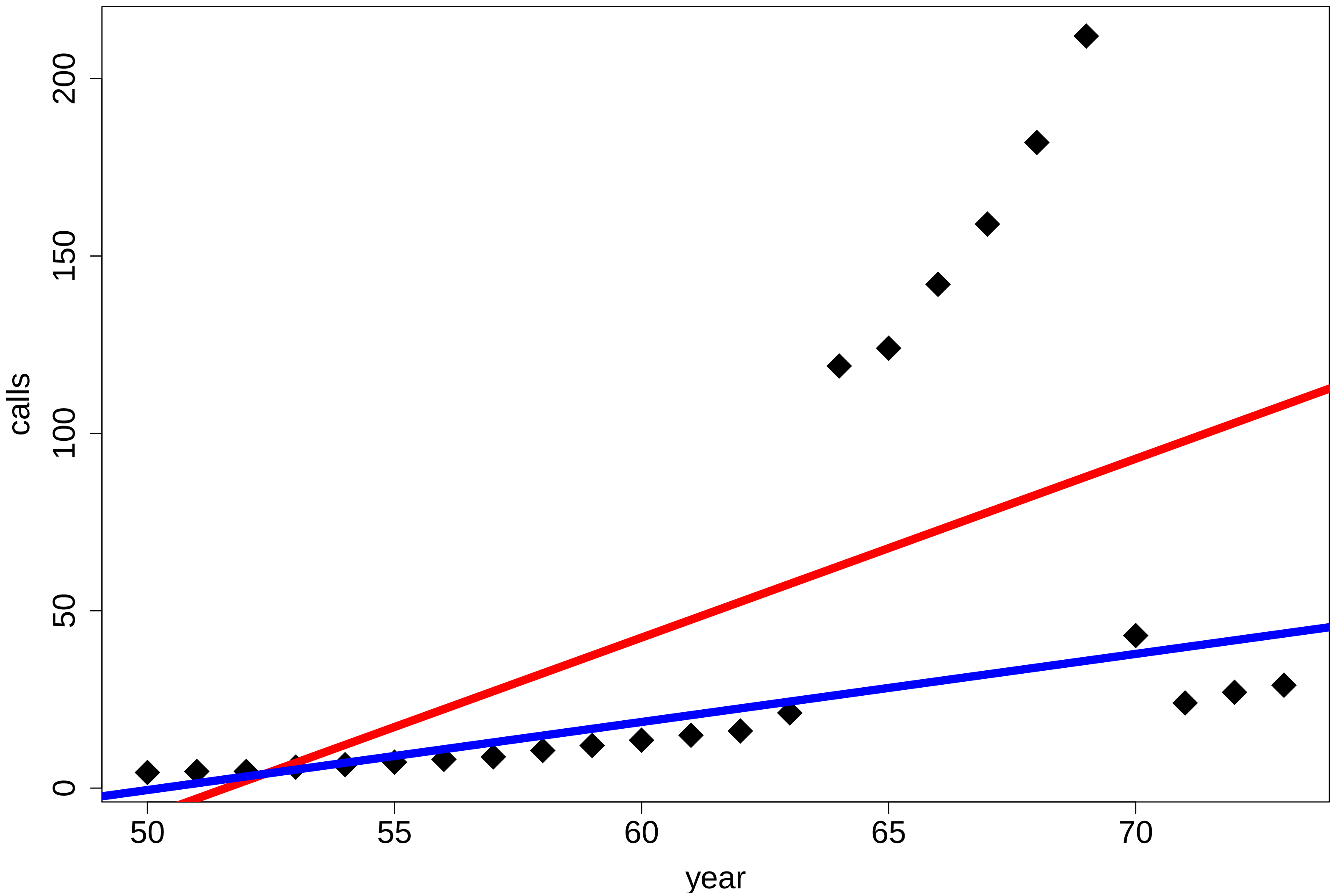
$$\hat{\boldsymbol{w}} = \arg\min_{\boldsymbol{w}} \|\boldsymbol{r}\|_1 = \arg\min_{\boldsymbol{w}} \|\boldsymbol{y} - X\boldsymbol{w}\|_1 = \arg\min_{\boldsymbol{w}} \sum_{i=1}^{n} \Big| y_i - \sum_{j=1}^{d} x_{ij} w_j \Big|.$$

Unlike for least squares regression, there is no explicit formula for the solution. However, the problem can be reformulated as:

$$\begin{aligned} \text{minimize} \quad & \sum_i t_i \\ \text{s.t.} \quad & t_i - \Big| y_i - \sum_j x_{ij} w_j \Big| = 0, \ i = 1, \ldots, n, \end{aligned}$$

which already looks similar to a LP...

Least absolute deviations regression



Rousseeuw & Leroy (1987) give data on annual numbers of Belgian telephone calls. Their investigations showed that for 1964-69 the total length of calls (in minutes) had been recorded rather than the number. Red: least-squares, blue: least absolute deviations.

Least absolute deviations regression

This is equivalent to the following problem:

$$
\begin{aligned}
\text{minimize} \quad & \sum_i t_i \\
\text{s.t.} \quad & t_i = y_i - \sum_j x_{ij} w_j, \quad \text{if} \quad y_i - \sum_j x_{ij} w_j \geq 0. \\
& -t_i = y_i - \sum_j x_{ij} w_j, \quad \text{else.}
\end{aligned}
$$

Note that in the first case, we can relax the constraint to $t_i \geq y_i - \sum_j x_{ij} w_j$, since we minimize over the t_i anyway.
Similarly, in the second case: $-t_i \leq y_i - \sum_j x_{ij} w_j$.
Both cases can be combined into range constraints:

$$
-t_i \leq y_i - \sum_j x_{ij} w_j \leq t_i, \; i = 1, \dots, n.
$$

Least absolute deviations regression

Finally, LAD-regression is equivalent to the following LP problem:

$$\begin{array}{ll} \text{minimize} & \sum_i t_i \\ \text{s.t.} & -t_i \leq y_i - \sum_j x_{ij} w_j \leq t_i,\ i = 1, \ldots, n. \end{array}$$

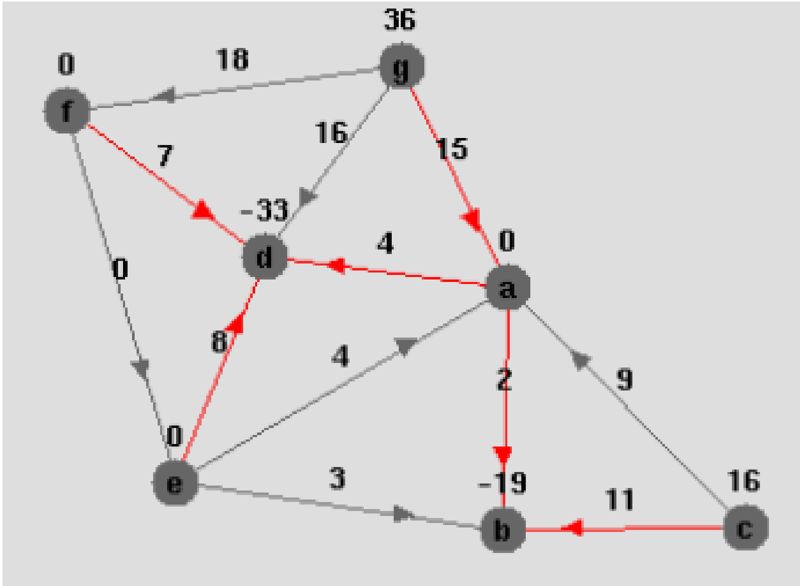
Range constraints can be transformed to standard notation by replication:

$$\begin{aligned} -\sum_j x_{ij} w_j - t_i &\leq -y_i,\ i = 1, \ldots, n. \\ \sum_j x_{ij} w_j - t_i &\leq y_i,\ i = 1, \ldots, n. \end{aligned}$$

$\Rightarrow$ Use Simplex algorithm on joint variable set $\{\boldsymbol{w}, \boldsymbol{t}\}$ to find optimal $\hat{\boldsymbol{w}}$:

$$\max \left(\ \mathbf{0}^t, -\mathbf{1}^t \ \right) \begin{pmatrix} \boldsymbol{w} \\ \boldsymbol{t} \end{pmatrix} \quad \text{s.t.} \quad -X\boldsymbol{w} - \boldsymbol{t} \leq -\boldsymbol{y}, X\boldsymbol{w} - \boldsymbol{t} \leq \boldsymbol{y}.$$

Network Flow Problems



- Network consists of set of nodes $\mathcal{N}$ and set of directed arcs $\mathcal{A}$.
- $b_i,\ i \in \mathcal{N}$: supply at node i.
- Demands are recorded as negative supplies.
- $c_{ij},\ (i,j) \in \mathcal{A}$: cost of shipping 1 unit along arc (i,j).
- Variables $x_{ij},\ (i,j) \in \mathcal{A}$: quantity to ship along arc (i,j).
- Objective: minimize $\sum_{(i,j)\in\mathcal{A}} c_{ij}x_{ij}$.

Constraints

- Mass conservation (flow balance):
  inflow(k) - outflow(k) $=$ demand(k) $=$ -supply(k), $k \in \mathcal{N}$

$$\sum_{i:(i,k)\in\mathcal{A}} x_{ik} - \sum_{j:(k,j)\in\mathcal{A}} x_{kj} = -b_k$$

- Nonnegativity: $x_{ij} \geq 0,\ (i,j) \in \mathcal{A}$

- Final LP: min $\boldsymbol{c}^t\boldsymbol{x}$ s.t. $A\boldsymbol{x} = -\boldsymbol{b},\ \boldsymbol{x} \geq \boldsymbol{0}$.