# LINEAR PROGRAMMING

[V. CH4]: EFFICIENCY OF THE SIMPLEX METHOD

# Phillip Keldenich Ahmad Moradi

Department of Computer Science Algorithms Department TU Braunschweig

December 11, 2023

## PERFORMANCE MEASURES

WORST-CASE ANALYSIS OF THE SIMPLEX METHOD

EMPIRICAL AVERAGE PERFORMANCE

#### PERFORMANCE MEASURES

We saw that the simplex method (equipped with anti-cycling rules) will solve any linear programming problem for which an optimal solution exists. The question, now, is *how fast* it will solve such a problem.

We saw that the simplex method (equipped with anti-cycling rules) will solve any linear programming problem for which an optimal solution exists. The question, now, is *how fast* it will solve such a problem.

#### Performance measures are:

- worst case looks at all problems of a given "size" and asks how much effort is needed to solve the hardest.
- average case looks at the average amount of effort needed to solve all.

We saw that the simplex method (equipped with anti-cycling rules) will solve any linear programming problem for which an optimal solution exists. The question, now, is *how fast* it will solve such a problem.

### Performance measures are:

- worst case looks at all problems of a given "size" and asks how much effort is needed to solve the hardest.
- average case looks at the average amount of effort needed to solve all.
- → Worst-case analyses are generally easier. It needs to provide an upper bound on how much effort required + a specific example attaing this bound
- $\leadsto$  On an average case analysis, one needs a stochastic model of random problems and evaluate the amount of effort required to solve every problem in the sample space.

worst-case is more tractable, but less relevant while dealing with real problems

#### PERFORMANCE MEASURES

Before looking at the worst case, we must discuss:

**First:** how do we specify the size of a problem? recall that we have m (constraints), n (variables)

**First:** how do we specify the size of a problem? recall that we have m (constraints), n (variables)

 $\bullet$   $m \times n$ 

**First:** how do we specify the size of a problem? recall that we have m (constraints), n (variables)

- $\bullet$   $m \times n$
- number of non-zero elements

First: how do we specify the size of a problem? recall that we have m (constraints), n (variables)

- $\bullet$   $m \times n$
- number of non-zero elements
- the number of bits needed to store all data

First: how do we specify the size of a problem? recall that we have m (constraints), n (variables)

- $\bullet$   $m \times n$
- number of non-zero elements
- the number of bits needed to store all data

Even with the last one, "size" of data might still be ambiguous. (why?)

we shall simply focus on m and n to characterize the size of a problem.

#### PERFORMANCE MEASURES

**Second:** how one should measure the amount of work required to solve a problem? the number of seconds of computer time?

**Second:** how one should measure the amount of work required to solve a problem? the number of seconds of computer time?

Fortunately, a reasonable substitute:

Algorithms are generally iterative processes. So

time to solve a problem = number of iterations  $\times$  time required to do each iteration

**Second:** how one should measure the amount of work required to solve a problem? the number of seconds of computer time?

Fortunately, a reasonable substitute:

Algorithms are generally iterative processes. So  $time\ to\ solve\ a\ problem = number\ of\ iterations imes time\ required\ to\ do\ each\ iteration$ 

First factor is platform-independent and so a *resasonable suragate* for the acctual time (not when different algorithm are compared!)

We simply count the number of iterations (pivots).

WORST-CASE ANALYSIS OF THE SIMPLEX METHOD

## PERFORMANCE MEASURES

## WORST-CASE ANALYSIS OF THE SIMPLEX METHOD

EMPIRICAL AVERAGE PERFORMANCE

Since the simplex method operates by moving from one dictionary to another (without cycling), an upper bound on the number of iterations is simply the number of possible dictionaries:

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

For a fixed value of the sum n + m, this expression is maximized when m = n.

Since the simplex method operates by moving from one dictionary to another (without cycling), an upper bound on the number of iterations is simply the number of possible dictionaries:

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

For a fixed value of the sum n + m, this expression is maximized when m = n.

How big is it? Not hard to show:

$$\frac{1}{2n}2^{2n} \le \binom{2n}{n} \le 2^{2n}$$

For n = 25:

$$2^{50} = 1.1259 \times 10^{15}$$

Our best chance for finding a bad example is to look at the case where m = n.

In 1972, V. Klee and G. J. Minty were the first to discover an example in which the simplex method using the largest coefficient rule requires  $2^n-1$  iterations to solve.

In 1972, V. Klee and G. J. Minty were the first to discover an example in which the simplex method using the largest coefficient rule requires  $2^n - 1$  iterations to solve.

lets look at the generic example:

Given constants 
$$1 = \beta_1 \ll \beta_2 \ll \cdots \ll \beta_n$$

$$\max_{x} \sum_{j=1}^{n} 2^{n-j} x_{j} - \frac{1}{2} \sum_{j=1}^{n} 2^{n-j} \beta_{j}$$
s.t. 
$$2 \sum_{j=1}^{i-1} 2^{i-j} x_{j} + x_{i} \le \sum_{j=1}^{i-1} 2^{i-j} \beta_{j} + \beta_{i} \quad i = 1, 2, \dots, n$$

$$x_{j} \ge 0 \qquad i = 1, 2, \dots, n$$

For n=3

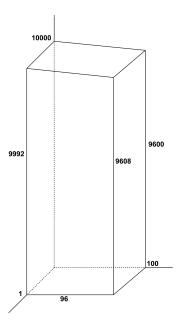
$$\begin{array}{lllll} \max_{x} & 4x_{1} + & 2x_{2} + & 1x_{3} - & 2\beta_{1} - \beta_{2} - \frac{1}{2}\beta_{3} \\ \text{s.t.} & 1x_{1} + & + & \leq \beta_{1} \\ & 4x_{1} + & 1x_{2} + & \leq 2\beta_{1} + \beta_{2} \\ & 8x_{1} + & 4x_{2} + & 1x_{3} \leq 4\beta_{1} + 2\beta_{2} + \beta_{3} \\ & x_{1}, & x_{2}, & x_{3} \geq 0 \end{array}$$

For n=3

$$\begin{array}{lllll} \max & 4x_1 + & 2x_2 + & 1x_3 - & 2\beta_1 - \beta_2 - \frac{1}{2}\beta_3 \\ \text{s.t.} & 1x_1 + & + & \leq \beta_1 \\ & 4x_1 + & 1x_2 + & \leq 2\beta_1 + \beta_2 \\ & 8x_1 + & 4x_2 + & 1x_3 \leq 4\beta_1 + 2\beta_2 + \beta_3 \\ & x_1, & x_2, & x_3 \geq 0 \end{array}$$

Assuming  $\beta_2 = 98, \beta_3 = 9800$ , the feasible region looks like :

$$\begin{array}{lll} 1x_1 & \leq 1 \\ 4x_1 + & 1x_2 & \leq 100 \\ 8x_1 + & 4x_2 + & 1x_3 \leq 10000 \\ x_1, & x_2, & x_3 \geq 0 \end{array}$$



Constraints represent a "minor" distortion to an 3-dimensional hypercube:

$$0 \le x1 \le 1$$

$$0 \leq x2 \leq 100$$

Constraints represent a "minor" distortion to an 3-dimensional hypercube:

$$\begin{aligned} 0 &\leq x1 \leq 1 \\ 0 &\leq x2 \leq 100 \\ 0 &\leq x3 \leq 10000 \end{aligned}$$

This region has  $2^{n-3}$  vertices and the idea is to trick the pivot rule so as to visit all of them.

Back to the generic Klee-Minty example (n = 3), the first dictionary is

$\zeta = -$	$2\beta_1$ –	$1\beta_2$ –	$0.5\beta_3 +$	$4x_1 +$	$2x_2 +$	$1 x_3$
$w_1 = +$	$1\beta_1$		_	$1x_1$		
$w_2 = +$	$2\beta_1$ +	$1\beta_2$	_	$4x_1 -$	$1x_2$	
$w_3 = +$	$4\beta_1$ +	$2\beta_2$ +	$1\beta_3$ –	$8x_1$ -	$4x_2 -$	$1x_3$

which is feasible.

Back to the generic Klee-Minty example (n = 3), the first dictionary is

$\zeta = -$	$2\beta_1$ –	$1\beta_2$ –	$0.5\beta_3 +$	$4x_1 +$	$2x_2 +$	$1 x_3$
$w_1 = +$	$1\beta_1$		_	$1x_1$		
$w_2 = +$	$2\beta_1$ +	$1\beta_2$	_	$4x_1 -$	$1x_2$	
$w_3 = +$	$4\beta_1$ +	$2\beta_2$ +	$1\beta_3$ –	$8x_1$ -	$4x_2 -$	$1x_3$

which is feasible.

Using the largest-coefficient rule:  $(x_1, w_1)$  are the entering, leaving pair.

Back to the generic Klee-Minty example (n = 3), the first dictionary is

which is feasible.

Using the largest-coefficient rule:  $(x_1, w_1)$  are the entering, leaving pair.

Corresponding vertex on the Klee-Minty cube is: (0,0,0)

## After the first pivot

## After the first pivot

for the next pivot:  $(x_2, w_2)$ 

## After the first pivot

for the next pivot:  $(x_2, w_2)$ 

Corresponding vertex on the Klee-Minty cube is:  $(\beta_1, 0, 0)$ 

## After the second pivot

### After the second pivot

for the next pivot:  $(w_1, x_1)$ 

### After the second pivot

for the next pivot:  $(w_1, x_1)$ 

Corresponding vertex on the Klee-Minty cube is:  $(\beta_1, -2\beta_1 + 1\beta_2, 0)$ 

## After the third pivot

## After the third pivot

for the next pivot:  $(x_3, w_3)$ 

## After the third pivot

for the next pivot:  $(x_3, w_3)$ 

Corresponding vertex on the Klee-Minty cube is:  $(0, 2\beta_1 + 1\beta_2, 0)$ 

# After the fourth pivot

# After the fourth pivot

for the next pivot:  $(x_1, w_1)$ 

#### After the fourth pivot

for the next pivot:  $(x_1, w_1)$ 

Corresponding vertex on the Klee-Minty cube is:  $(0, 2\beta_1 + 1\beta_2, -4\beta_1 - 2\beta_2 + 1\beta_3)$ 

# After the fifth pivot

	$\zeta = +$	$2\beta_1$ –	$1\beta_2$ +	$0.5\beta_3$ $-$	$4w_1 + $	$2 w_2 -$	$1w_3$
$\boldsymbol{x}$	$r_1 = +$	$1\beta_1$		_	$1w_1$		
$\boldsymbol{x}$	$c_2 = -$	$2\beta_1$ +	$1\beta_2$	+	$4w_1$ -	$1w_2$	
a	$x_3 = +$	$4\beta_1$ –	$2\beta_2 +$	$1\beta_3$ –	$8w_1 + $	$4w_2 -$	$1w_3$

# After the fifth pivot

for the next pivot:  $(w_2, x_2)$ 

# After the fifth pivot

for the next pivot:  $(w_2, x_2)$ 

Corresponding vertex on the Klee-Minty cube is:  $(1\beta_1, -2\beta_1 + 1\beta_2, +4\beta_1 - 2\beta_2 + 1\beta_3)$ 

# After the sixth pivot

### After the sixth pivot

for the next pivot:  $(w_1, x_1)$ 

# After the sixth pivot

for the next pivot:  $(w_1, x_1)$ 

Corresponding vertex on the Klee-Minty cube is:  $(1\beta_1, 0, -4\beta_1 + 2\beta_2 + 1\beta_3)$ 

# And finally after the seventh pivot

# And finally after the seventh pivot

The optimal dictionary.

### And finally after the seventh pivot

The optimal dictionary.

Corresponding vertex on the Klee-Minty cube is:  $(0,0,+4\beta_1+2\beta_2+1\beta_3)$ 

• It took  $7 = 2^3 - 1$  iterations.

Largest-coefficient rule can take  $2^n-1$  pivots to solve a problem in n variables and constraints.

• It took  $7 = 2^3 - 1$  iterations.

Largest-coefficient rule can take  $2^n-1$  pivots to solve a problem in n variables and constraints.

For 
$$n = 70, 2^n = 1.2 \times 10^{21}$$

On a machin with 1000 ips. It will take 40 billion years to solve! (The age of the universe is estimated to be 13.7 billion years)

• It took  $7 = 2^3 - 1$  iterations.

Largest-coefficient rule can take  $2^n-1$  pivots to solve a problem in n variables and constraints.

For 
$$n = 70, 2^n = 1.2 \times 10^{21}$$

On a machin with 1000 ips. It will take 40 billion years to solve! (The age of the universe is estimated to be 13.7 billion years)

Yet, problems with 10,000 to 100,000 variables/constraints are solved routinely every day. (Worst case analysis!)

• Note that the final dictionary could have been reached from the initial dictionary in just one pivot if we had selected  $x_3$  to be the entering variable. But the largest-coefficient rule dictated selecting  $x_1$  and the exponential worst case running time.

- Note that the final dictionary could have been reached from the initial dictionary in just one pivot if we had selected  $x_3$  to be the entering variable. But the largest-coefficient rule dictated selecting  $x_1$  and the exponential worst case running time.
- Indeed this is an open question:

Does there exist a varient of the Simplex method with polynomial worst case performance?

- Note that the final dictionary could have been reached from the initial dictionary in just one pivot if we had selected  $x_3$  to be the entering variable. But the largest-coefficient rule dictated selecting  $x_1$  and the exponential worst case running time.
- Indeed this is an open question:

Does there exist a varient of the Simplex method with polynomial worst case performance?

• However, For linear programs, we have other algorithm (called *interior point* methods) with *polynomial worstcase performance*. In contrast to the simplex algorithm, which finds an optimal solution by traversing the edges between vertices on a polyhedral set, interior-point methods move through the interior of the feasible region.

• Note that every pivot is the *swap* of an  $x_j$  with the corresponding  $w_j$ .

- ullet Note that every pivot is the *swap* of an  $x_j$  with the corresponding  $w_j$  .
- every dictionary looks just like the first one with the exception that the  $w_i$  's and the  $x_i$  's have become intertwined and *various signs have changed*.

- Note that every pivot is the *swap* of an  $x_j$  with the corresponding  $w_j$ .
- every dictionary looks just like the first one with the exception that the w<sub>i</sub> 's and the x<sub>i</sub> 's
  have become intertwined and various signs have changed.
- Objective function changes in a specific patern, can you guess how?

- ullet Note that every pivot is the *swap* of an  $x_j$  with the corresponding  $w_j$ .
- every dictionary looks just like the first one with the exception that the  $w_i$  's and the  $x_i$  's have become intertwined and *various signs have changed*.
- Objective function changes in a specific patern, can you guess how?

These are the necessary steps needed to take toward a proof of worst case exaponantial time of the generic Klee-Minty example. (Left as part of to the third homework)

EMPIRICAL AVERAGE PERFORMANCE

PERFORMANCE MEASURES

WORST-CASE ANALYSIS OF THE SIMPLEX METHOD

EMPIRICAL AVERAGE PERFORMANCE



Lets again consider the simple simplex implementation discussed previous week.