The Two-phase Simplex Method

Two-phase simplex method

1. Given an LP in standard from, first run phase I.
2. If phase I yields a basic feasible solution for the original LP, enter "phase II" (see above).

Possible outcomes of the two-phase simplex method

- i. Problem is infeasible (detected in phase I).
- ii. Problem is feasible but rows of A are linearly dependent (detected and corrected at the end of phase I by eliminating redundant constraints.)
- iii. Optimal cost is $-\infty$ (detected in phase II).
- iv. Problem has optimal basic feasible solution (found in phase II).

Remark: (ii) is not an outcome but only an intermediate result leading to outcome (iii) or (iv).

Big-M Method

Alternative idea: Combine the two phases into one by introducing sufficiently large penalty costs for artificial variables.

Big-M Method

Alternative idea: Combine the two phases into one by introducing sufficiently large penalty costs for artificial variables.

This way, the LP

$$\begin{array}{lrcl} \min & \sum_{i=1}^{n} c_i x_i & & \\ \text{s.t.} & A \cdot x & = & b \\ & x & \geq & 0 \end{array}$$

Big-M Method

Alternative idea: Combine the two phases into one by introducing sufficiently large penalty costs for artificial variables.

This way, the LP

$$\begin{array}{lrcl} \min & \sum_{i=1}^{n} c_i x_i & & \\ \text{s.t.} & A \cdot x & = & b \\ & x & \geq & 0 \end{array}$$

becomes:

$$\begin{array}{lrcrcl} \min & \sum_{i=1}^{n} c_i x_i & + & M \cdot \sum_{j=1}^{m} y_j & & \\ \text{s.t.} & A \cdot x & + & I_m \cdot y & = & b \\ & & & x, y & \geq & 0 \end{array}$$

Big-M Method

Alternative idea: Combine the two phases into one by introducing sufficiently large penalty costs for artificial variables.

This way, the LP

$$\begin{array}{lrcl} \min & \sum_{i=1}^{n} c_i\, x_i & & \\ \text{s.t.} & A \cdot x & = & b \\ & x & \geq & 0 \end{array}$$

becomes:

$$\begin{array}{lrcrcl} \min & \sum_{i=1}^{n} c_i\, x_i & + & M \cdot \sum_{j=1}^{m} y_j & & \\ \text{s.t.} & A \cdot x & + & I_m \cdot y & = & b \\ & & & x, y & \geq & 0 \end{array}$$

Remark: If M is sufficiently large and the original program has a feasible solution, all artificial variables will be driven to zero by the simplex method.

How to Choose M?

Observation

Initially, M only occurs in the zeroth row. As the zeroth row never becomes pivot row, this property is maintained while the simplex method is running.

How to Choose M?

Observation

Initially, M only occurs in the zeroth row. As the zeroth row never becomes pivot row, this property is maintained while the simplex method is running.

All we need to have is an order on all values that can appear as reduced cost coefficients.

How to Choose M?

Observation

Initially, M only occurs in the zeroth row. As the zeroth row never becomes pivot row, this property is maintained while the simplex method is running.

All we need to have is an order on all values that can appear as reduced cost coefficients.

Order on cost coefficients

$$a\,M + b < c\,M + d \quad :\Longleftrightarrow \quad (a < c) \;\vee\; (a = c \wedge b < d)$$

How to Choose M?

Observation

Initially, M only occurs in the zeroth row. As the zeroth row never becomes pivot row, this property is maintained while the simplex method is running.

All we need to have is an order on all values that can appear as reduced cost coefficients.

Order on cost coefficients

$$a\,M + b < c\,M + d \quad :\Longleftrightarrow \quad (a < c) \ \vee \ (a = c \wedge b < d)$$

In particular, $-a\,M + b < 0 < a\,M + b$ for any positive a and arbitrary b, and we can decide whether a cost coefficient is negative or not.

How to Choose M?

Observation

Initially, M only occurs in the zeroth row. As the zeroth row never becomes pivot row, this property is maintained while the simplex method is running.

All we need to have is an order on all values that can appear as reduced cost coefficients.

Order on cost coefficients

$$a\,M + b < c\,M + d \quad :\Longleftrightarrow \quad (a < c) \;\vee\; (a = c \wedge b < d)$$

In particular, $-a\,M + b < 0 < a\,M + b$ for any positive a and arbitrary b, and we can decide whether a cost coefficient is negative or not.

$\rightarrow$ There is no need to give M a fixed numerical value.

Example

Example:

$$
\begin{array}{lrcrcrcrcl}
\min & x_1 & + & x_2 & + & x_3 & & & & \\
\text{s.t.} & x_1 & + & 2\,x_2 & + & 3\,x_3 & & & = & 3 \\
 & -x_1 & + & 2\,x_2 & + & 6\,x_3 & & & = & 2 \\
 & & & 4\,x_2 & + & 9\,x_3 & & & = & 5 \\
 & & & & & 3\,x_3 & + & x_4 & = & 1 \\
 & & & & & & & x_1,\ldots,x_4 & \geq & 0
\end{array}
$$

Introducing Artificial Variables and M

make them zero

Auxiliary problem:

$$
\begin{array}{llllllllll}
\min & x_1 & +x_2 & +x_3 & & +M\,x_5 & +M\,x_6 & +M\,x_7 & & \\
\text{s.t.} & x_1 & +2\,x_2 & +3\,x_3 & & x_5 & & & = & 3 \\
& -x_1 & +2\,x_2 & +6\,x_3 & & & +x_6 & & = & 2 \\
& & 4\,x_2 & +9\,x_3 & & & & +x_7 & = & 5 \\
& & & 3\,x_3 & +x_4 & & & & = & 1 \\
& & & & & & & x_1,\ldots,x_4,x_5,x_6,x_7 & \geq & 0
\end{array}
$$

Note that this time the unnecessary artificial variable x_8 has been omitted.

We start off with $(x_5, x_6, x_7, x_4) = (3, 2, 5, 1)$.

Forming the Initial Tableau

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
0	1	1	1	0	M	M	M
3	1	2	3	0	1	0	0
2	−1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Forming the Initial Tableau

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
0	1	1	1	0	M	M	M
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Compute reduced costs by eliminating the nonzero entries for the basic variables.

Forming the Initial Tableau

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-3M$	$-M+1$	$-2M+1$	$-3M+1$	0	0	M	M
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Compute reduced costs by eliminating the nonzero entries for the basic variables.

Forming the Initial Tableau

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-5M$	1	$-4M+1$	$-9M+1$	0	0	0	M
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Compute reduced costs by eliminating the nonzero entries for the basic variables.

Forming the Initial Tableau

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-10M$	1	$-8M+1$	$-18M+1$	0	0	0	0
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Compute reduced costs by eliminating the nonzero entries for the basic variables.

Forming the Initial Tableau

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-10M$	1	$-8M+1$	$-18M+1$	0	0	0	0
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Compute reduced costs by eliminating the nonzero entries for the basic variables.

First Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-10M$	1	$-8M+1$	$-18M+1$	0	0	0	0
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

First Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-10M$	1	$-8M+1$	$-18M+1$	0	0	0	0
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

Reduced costs for x_2 and x_3 are negative.

First Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-10M$	1	$-8M+1$	$-18M+1$	0	0	0	0
3	1	2	3	0	1	0	0
2	-1	2	6	0	0	1	0
5	0	4	9	0	0	0	1
1	0	0	3	1	0	0	0

add $6M - \frac{1}{3}$ times

≤ 1, $\leq 1/3$, $\leq 5/9$, $x_3 \leq \frac{1}{3}$

Reduced costs for x_2 and x_3 are negative.

Basis change: x_3 enters the basis, x_4 leaves.

Second Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
$-4M-1/3$	1	$-8M+1$	0	$6M-1/3$	0	0	0	
2	1	2	0	-1	1	0	0	≤ 1
0	-1	2	0	-2	0	1	0	≤ 0
2	0	4	0	-3	0	0	1	$\leq \frac{1}{2}$
$1/3$	0	0	1	$1/3$	0	0	0	—

Second Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-4M - 1/3$	1	$-8M + 1$	0	$6M - 1/3$	0	0	0
2	1	2	0	-1	1	0	0
0	-1	2	0	-2	0	1	0
2	0	4	0	-3	0	0	1
$1/3$	0	0	1	$1/3$	0	0	0

Basis change: x_2 enters the basis, x_6 leaves.

Third Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-4M-1/3$	$-4M+3/2$	0	0	$-2M+2/3$	0	$4M-1/2$	0
2	2	0	0	1	1	-1	0
0	$-1/2$	1	0	-1	0	$1/2$	0
2	2	0	0	1	0	-2	1
$1/3$	0	0	1	$1/3$	0	0	0

Current solution:

$x_5 = 2$

$x_2 = 0$

$x_4 = 2$

$x_3 = \frac{1}{3}$

$\text{Obj} = 4M + \frac{1}{3}$

Third Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-4M - 1/3$	$-4M + 3/2$	0	0	$-2M + 2/3$	0	$4M - 1/2$	0
2	2	0	0	1	1	-1	0
0	$-1/2$	1	0	-1	0	$1/2$	0
2	2	0	0	1	0	-2	1
$1/3$	0	0	1	$1/3$	0	0	0

Basis change: x_1 enters the basis, x_5 leaves.

Fourth Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-11/6$	0	0	0	$-1/12$	$2M-3/4$	$2M+1/4$	0
21	1	0	0	$1/2$	$1/2$	$-1/2$	0
$1/2$	0	1	0	$-3/4$	$1/4$	$1/4$	0
0	0	0	0	0	-1	-1	1
$1/3$	0	0	1	$1/3$	0	0	0

Fourth Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-11/6$	0	0	0	$-1/12$	$2M - 3/4$	$2M + 1/4$	0
21	1	0	0	$1/2$	$1/2$	$-1/2$	0
$1/2$	0	1	0	$-3/4$	$1/4$	$1/4$	0
0	0	0	0	0	-1	-1	1
$1/3$	0	0	1	$1/3$	0	0	0

Note that all artificial variables have already been driven to 0.

Fourth Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-11/6$	0	0	0	$-1/12$	$2M - 3/4$	$2M + 1/4$	0
21	1	0	0	$1/2$	$1/2$	$-1/2$	0
$1/2$	0	1	0	$-3/4$	$1/4$	$1/4$	0
0	0	0	0	0	-1	-1	1
$1/3$	0	0	1	$1/3$	0	0	0

≤ 42
—
—
≤ 1

Note that all artificial variables have already been driven to 0.

Basis change: x_4 enters the basis, x_3 leaves.

Fifth Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-7/4$	0	0	$1/4$	0	$2M - 3/4$	$2M + 1/4$	0
$1/2$	1	0	$-3/2$	0	$1/2$	$-1/2$	0
$5/4$	0	1	$9/4$	0	$1/4$	$1/4$	0
0	0	0	0	0	-1	-1	1
1	0	0	3	1	0	0	0

Fifth Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-7/4$	0	0	$1/4$	0	$2M - 3/4$	$2M + 1/4$	0
$1/2$	1	0	$-3/2$	0	$1/2$	$-1/2$	0
$5/4$	0	1	$9/4$	0	$1/4$	$1/4$	0
0	0	0	0	0	-1	-1	1
1	0	0	3	1	0	0	0

We now have an optimal solution of the auxiliary problem, as all costs are nonnegative (M presumed large enough).

Fifth Iteration

	x_1	x_2	x_3	x_4	x_5	x_6	x_7
$-7/4$	0	0	$1/4$	0	$2M-3/4$	$2M+1/4$	0
$1/2$	1	0	$-3/2$	0	$1/2$	$-1/2$	0
$5/4$	0	1	$9/4$	0	$1/4$	$1/4$	0
0	0	0	0	0	-1	-1	1
1	0	0	3	1	0	0	0

We now have an optimal solution of the auxiliary problem, as all costs are nonnegative (M presumed large enough).

By elimiating the third row as in the previous example, we get a basic feasible and also optimal solution to the original problem.

$x_1 = \frac{1}{2}$ $\quad$ $x_i = 0$ otherwise

$x_2 = \frac{5}{4}$

$x_7 = 0$ $\quad\quad$ opt $= \frac{4}{4}$

$x_4 = 1$

Computational Efficiency of the Simplex Method

Observation

The computational efficiency of the simplex method is determined by

i. the computational effort of each iteration;
ii. the number of iterations.

Computational Efficiency of the Simplex Method

Observation

The computational efficiency of the simplex method is determined by

1. the computational effort of each iteration;
2. the number of iterations.

Question: How many iterations are needed in the worst case?

Computational Efficiency of the Simplex Method

Observation

The computational efficiency of the simplex method is determined by

i. the computational effort of each iteration;
ii. the number of iterations.

Question: How many iterations are needed in the worst case?

Idea for negative answer (lower bound)

Describe

- a polyhedron with an exponential number of vertices;
- a path that visits all vertices and always moves from a vertex to an adjacent one that has lower costs.

Computational Efficiency of the Simplex Method

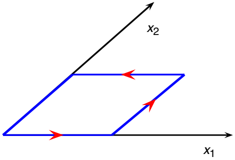
Unit cube

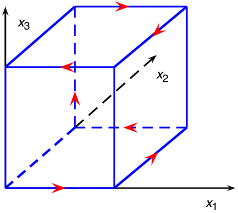
Consider the unit cube in $\mathbb{R}^n$, defined by the constraints

$$0 \le x_i \le 1, \quad i = 1, \ldots, n$$

The unit cube has

- 2^n vertices;
- a *spanning path*, i. e., a path traveling the edges of the cube visiting each vertex exactly once.





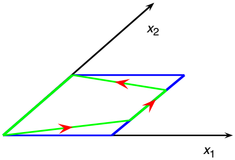
Computational Efficiency of the Simplex Method (cont.)

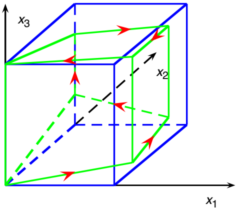
Klee-Minty cube

Consider a perturbation of the unit cube in $\mathbb{R}^n$, defined by the constraints

$$0 \leq x_1 \leq 1,$$
$$\epsilon x_{i-1} \leq x_i \leq 1 - \epsilon x_{i-1}, \quad i = 2, \ldots, n$$

for some $\epsilon \in (0, 1/2)$.





Computational Efficiency of the Simplex Method (cont.)

Klee-Minty cube

$$0 \le x_1 \le 1,$$
$$\epsilon x_{i-1} \le x_i \le 1 - \epsilon x_{i-1}, \quad i = 2, \dots, n, \epsilon \in (0, 1/2)$$

Computational Efficiency of the Simplex Method (cont.)

Klee-Minty cube

$$0 \le x_1 \le 1,$$
$$\epsilon x_{i-1} \le x_i \le 1 - \epsilon x_{i-1}, \quad i = 2, \ldots, n, \epsilon \in (0, 1/2)$$

Theorem 3.1.

Consider the linear programming problem of minimizing $-x_n$ subject to the constraints above. Then,

a. the feasible set has 2^n vertices;
b. the vertices can be ordered so that each one is adjacent to and has lower cost than the previous one;
c. there exists a pivoting rule under which the simplex method requires $2^n - 1$ changes of basis before it terminates.

Diameter of Polyhedra

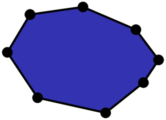
Definition 3.2.

- The distance $d(x, y)$ between two vertices x, y is the minimum number of edges required to reach y starting from x.
- The diameter $D(P)$ of polyhedron P is the maximum $d(x, y)$ over all pairs of vertices (x, y).
- $\Delta(n, m)$ is the maximum $D(P)$ over all polytopes in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.
- $\Delta_u(n, m)$ is the maximum $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.

Diameter of Polyhedra

Definition 3.2.

- The distance $d(x, y)$ between two vertices x, y is the minimum number of edges required to reach y starting from x.
- The diameter $D(P)$ of polyhedron P is the maximum $d(x, y)$ over all pairs of vertices (x, y).
- $\Delta(n, m)$ is the maximum $D(P)$ over all polytopes in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.
- $\Delta_u(n, m)$ is the maximum $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.

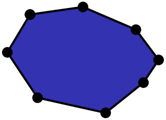


$$\Delta(2, 8) = \left\lfloor \frac{8}{2} \right\rfloor = 4$$

Diameter of Polyhedra

Definition 3.2.

- The distance $d(x, y)$ between two vertices x, y is the minimum number of edges required to reach y starting from x.
- The diameter $D(P)$ of polyhedron P is the maximum $d(x, y)$ over all pairs of vertices (x, y).
- $\Delta(n, m)$ is the maximum $D(P)$ over all polytopes in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.
- $\Delta_u(n, m)$ is the maximum $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.

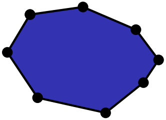


$$\Delta(2, 8) = \left\lfloor \frac{8}{2} \right\rfloor = 4$$
$$\Delta(2, m) = \left\lfloor \frac{m}{2} \right\rfloor$$

Diameter of Polyhedra

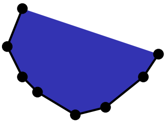
Definition 3.2.

- The distance $d(x, y)$ between two vertices x, y is the minimum number of edges required to reach y starting from x.
- The diameter $D(P)$ of polyhedron P is the maximum $d(x, y)$ over all pairs of vertices (x, y).
- $\Delta(n, m)$ is the maximum $D(P)$ over all polytopes in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.
- $\Delta_u(n, m)$ is the maximum $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.



$$\Delta(2, 8) = \left\lfloor \frac{8}{2} \right\rfloor = 4$$

$$\Delta(2, m) = \left\lfloor \frac{m}{2} \right\rfloor$$

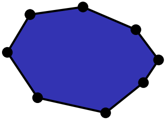


$$\Delta_u(2, 8) = 8 - 2 = 6$$

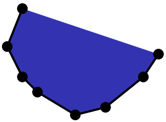
Diameter of Polyhedra

Definition 3.2.

- The distance $d(x, y)$ between two vertices x, y is the minimum number of edges required to reach y starting from x.
- The diameter $D(P)$ of polyhedron P is the maximum $d(x, y)$ over all pairs of vertices (x, y).
- $\Delta(n, m)$ is the maximum $D(P)$ over all polytopes in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.
- $\Delta_u(n, m)$ is the maximum $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of m inequality constraints.



$$\Delta(2, 8) = \left\lfloor \frac{8}{2} \right\rfloor = 4$$
$$\Delta(2, m) = \left\lfloor \frac{m}{2} \right\rfloor$$



$$\Delta_u(2, 8) = 8 - 2 = 6$$
$$\Delta_u(2, m) = m - 2$$

Hirsch Conjecture

Observation: The diameter of the feasible set in a linear programming problem is a lower bound on the number of steps required by the simplex method, no matter which pivoting rule is being used.

Hirsch Conjecture

Observation: The diameter of the feasible set in a linear programming problem is a lower bound on the number of steps required by the simplex method, no matter which pivoting rule is being used.

Polynomial Hirsch Conjecture

$$\Delta(n, m) \leq poly(m, n)$$

Hirsch Conjecture

Observation: The diameter of the feasible set in a linear programming problem is a lower bound on the number of steps required by the simplex method, no matter which pivoting rule is being used.

Polynomial Hirsch Conjecture

$$\Delta(n, m) \leq poly(m, n)$$

Remarks

- Known lower bounds: $\Delta_u(n, m) \geq m - n + \left\lfloor \frac{n}{5} \right\rfloor$

Hirsch Conjecture

Observation: The diameter of the feasible set in a linear programming problem is a lower bound on the number of steps required by the simplex method, no matter which pivoting rule is being used.

Polynomial Hirsch Conjecture

$$\Delta(n, m) \leq poly(m, n)$$

Remarks

- Known lower bounds: $\Delta_u(n, m) \geq m - n + \left\lfloor \frac{n}{5} \right\rfloor$
- Known upper bounds:

$$\Delta(n, m) \leq \Delta_u(n, m) < m^{1+\log_2 n} = (2n)^{\log_2 m}$$

Hirsch Conjecture

Observation: The diameter of the feasible set in a linear programming problem is a lower bound on the number of steps required by the simplex method, no matter which pivoting rule is being used.

Polynomial Hirsch Conjecture

$$\Delta(n, m) \leq poly(m, n)$$

Remarks

- Known lower bounds: $\Delta_u(n, m) \geq m - n + \left\lfloor \frac{n}{5} \right\rfloor$
- Known upper bounds:

$$\Delta(n, m) \leq \Delta_u(n, m) < m^{1+\log_2 n} = (2n)^{\log_2 m}$$

- The Strong Hirsch Conjecture

$$\Delta(n, m) \leq m - n$$

was disproven in 2010 by Paco Santos for $n = 43$, $m = 86$.

Average Case Behavior of the Simplex Method

- Despite the exponential lower bounds on the worst case behavior of the simplex method (Klee-Minty cubes etc.), the simplex method usually behaves well in practice.
- The number of iterations is “typically” $O(m)$.
- There have been several attempts to explain this phenomenon from a more theoretical point of view.
- These results say that “on average” the number of iterations is $O(\cdot)$ (usually polynomial).
- One main difficulty is to come up with a meaningful and, at the same time, manageable definition of the term “on average”.