ISyE/Math/CS/Stat 525
Linear Optimization

3. The simplex method
part 2

Prof. Alberto Del Pia
University of Wisconsin-Madison

Based on the book *Introduction to Linear Optimization* by D. Bertsimas and J.N. Tsitsiklis

Outline

Sec. 3.4 We discuss how the simplex method avoids cycling in order to reach an optimal solution.

Sec. 3.5 We understand how it finds an initial basic feasible solution.

Sec. 3.6 The roots of the name “simplex method”.

Sec. 3.7 We discuss its running time.

3.4 Anticycling: lexicography and Bland's rule

Example 3.6

- We now see an example that shows that the simplex method can cycle.
- We consider a problem described in terms of the following initial tableau.

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-3/4$	20	$-1/2$	6	0	0	0
$x_5 =$	0	$1/4$	-8	-1	9	1	0	0
$x_6 =$	0	$1/2$	-12	$-1/2$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

We use the following pivoting rules:

(a) We select a nonbasic variable with the most negative reduced cost $\bar{c}_j$ to be the one that enters the basis.

(b) Out of all basic variables that are eligible to exit the basis, we select the one with the smallest subscript.

We then obtain the following sequence of tableaux:

Example 3.6

Initial tableau:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-3/4$	20	$-1/2$	6	0	0	0
$x_5 =$	0	$1/4$	-8	-1	9	1	0	0
$x_6 =$	0	$1/2$	-12	$-1/2$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Initial tableau:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-3/4$	20	$-1/2$	6	0	0	0
$x_5 =$	0	$1/4$	-8	-1	9	1	0	0
$x_6 =$	0	$1/2$	-12	$-1/2$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

First pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	-4	$-7/2$	33	3	0	0
$x_1 =$	0	1	-32	-4	36	4	0	0
$x_6 =$	0	0	4	$3/2$	-15	-2	1	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

First pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	-4	$-7/2$	33	3	0	0
$x_1 =$	0	1	-32	-4	36	4	0	0
$x_6 =$	0	0	4	$3/2$	-15	-2	1	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Second pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	0	-2	18	1	1	0
$x_1 =$	0	1	0	8	-84	-12	8	0
$x_2 =$	0	0	1	$3/8$	$-15/4$	$-1/2$	$1/4$	0
$x_7 =$	1	0	0	1	0	0	0	1

First pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	-4	$-7/2$	33	3	0	0
$x_1 =$	0	1	-32	-4	36	4	0	0
$x_6 =$	0	0	4	$3/2$	-15	-2	1	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Second pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	0	-2	18	1	1	0
$x_1 =$	0	1	0	8	-84	-12	8	0
$x_2 =$	0	0	1	$3/8$	$-15/4$	$-1/2$	$1/4$	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Second pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	0	0	-2	18	1	1	0
$x_1 =$	0	1	0	8	-84	-12	8	0
$x_2 =$	0	0	1	$3/8$	$-15/4$	$-1/2$	$1/4$	0
$x_7 =$	1	0	0	1	0	0	0	1

Third pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$1/4$	0	0	-3	-2	3	0
$x_3 =$	0	$1/8$	0	1	$-21/2$	$-3/2$	1	0
$x_2 =$	0	$-3/64$	1	0	$3/16$	$1/16$	$-1/8$	0
$x_7 =$	1	$-1/8$	0	0	$21/2$	$3/2$	-1	1

Example 3.6

Third pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$1/4$	0	0	-3	-2	3	0
$x_3 =$	0	$1/8$	0	1	$-21/2$	$-3/2$	1	0
$x_2 =$	0	$-3/64$	1	0	$3/16$	$1/16$	$-1/8$	0
$x_7 =$	1	$-1/8$	0	0	$21/2$	$3/2$	-1	1

Example 3.6

Fourth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-1/2$	16	0	0	-1	1	0
$x_3 =$	0	$-5/2$	56	1	0	2	-6	0
$x_4 =$	0	$-1/4$	$16/3$	0	1	$1/3$	$-2/3$	0
$x_7 =$	1	$5/2$	-56	0	0	-2	6	1

Third pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$1/4$	0	0	-3	-2	3	0
$x_3 =$	0	$1/8$	0	1	$-21/2$	$-3/2$	1	0
$x_2 =$	0	$-3/64$	1	0	$3/16$	$1/16$	$-1/8$	0
$x_7 =$	1	$-1/8$	0	0	$21/2$	$3/2$	-1	1

Example 3.6

Fourth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-1/2$	16	0	0	-1	1	0
$x_3 =$	0	$-5/2$	56	1	0	2	-6	0
$x_4 =$	0	$-1/4$	$16/3$	0	1	$1/3$	$-2/3$	0
$x_7 =$	1	$5/2$	-56	0	0	-2	6	1

Example 3.6

Fourth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-1/2$	16	0	0	-1	1	0
$x_3 =$	0	$-5/2$	56	1	0	2	-6	0
$x_4 =$	0	$-1/4$	$16/3$	0	1	$1/3$	$-2/3$	0
$x_7 =$	1	$5/2$	-56	0	0	-2	6	1

Fifth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-7/4$	44	$1/2$	0	0	-2	0
$x_5 =$	0	$-5/4$	28	$1/2$	0	1	-3	0
$x_4 =$	0	$1/6$	-4	$-1/6$	1	0	$1/3$	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Fifth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-7/4$	44	$1/2$	0	0	-2	0
$x_5 =$	0	$-5/4$	28	$1/2$	0	1	-3	0
$x_4 =$	0	$1/6$	-4	$-1/6$	1	0	$1/3$	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Sixth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-3/4$	20	$-1/2$	6	0	0	0
$x_5 =$	0	$1/4$	-8	-1	9	1	0	0
$x_6 =$	0	$1/2$	-12	$-1/2$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

Fifth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-7/4$	44	$1/2$	0	0	-2	0
$x_5 =$	0	$-5/4$	28	$1/2$	0	1	-3	0
$x_4 =$	0	$1/6$	-4	$-1/6$	1	0	$1/3$	0
$x_7 =$	1	0	0	1	0	0	0	1

Example 3.6

Sixth pivot:

$$\begin{array}{r|c|ccccccc|}
 & & x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & x_7 \\
\hline
 & 3 & -3/4 & 20 & -1/2 & 6 & 0 & 0 & 0 \\
\hline
x_5 = & 0 & 1/4 & -8 & -1 & 9 & 1 & 0 & 0 \\
x_6 = & 0 & 1/2 & -12 & -1/2 & 3 & 0 & 1 & 0 \\
x_7 = & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
\hline
\end{array}$$

Example 3.6

Sixth pivot:

		x_1	x_2	x_3	x_4	x_5	x_6	x_7
	3	$-3/4$	20	$-1/2$	6	0	0	0
$x_5 =$	0	$1/4$	-8	-1	9	1	0	0
$x_6 =$	0	$1/2$	-12	$-1/2$	3	0	1	0
$x_7 =$	1	0	0	1	0	0	0	1

- After six pivots, we have the same basis and the same tableau that we started with.
- At each basis change, we had $\theta^* = 0$.
- In particular, for each intermediate tableau, we had the same feasible solution and the same cost.
- The same sequence of pivots can be repeated over and over, and the simplex method never terminates.

Anticycling: lexicography and Bland's rule

- Next, we discuss anticycling rules under which the simplex method is guaranteed to terminate, thus extending Theorem 3.3 to degenerate problems.
- As an important corollary, we conclude that if the optimal cost is finite, then there exists an optimal basis, that is, a basis satisfying

$$B^{-1}b \geq 0$$

and

$$\bar{c}' = c' - c_B' B^{-1} A \geq 0'.$$

Lexicography

Lexicography

- We present here the lexicographic pivoting rule and see that it prevents the simplex method from cycling.
- We start with a definition.

Definition 3.5

A vector $u \in \mathbb{R}^n$ is said to be lexicographically larger (or smaller) than another vector $v \in \mathbb{R}^n$ if $u \neq v$ and the first component that is different in u and v is larger (or smaller, respectively) in u. Symbolically, we write

$$u >^L v \qquad \text{or} \qquad u <^L v.$$

Example:

$$(0, 2, 3, 0) >^L (0, 2, 1, 4),$$
$$(0, 4, 5, 0) <^L (1, 2, 1, 2).$$

Lexicography

- We present here the lexicographic pivoting rule and see that it prevents the simplex method from cycling.
- We start with a definition.

Definition 3.5
A vector $u \in \mathbb{R}^n$ is said to be lexicographically larger (or smaller) than another vector $v \in \mathbb{R}^n$ if $u \neq v$ and the first component that is different in u and v is larger (or smaller, respectively) in u. Symbolically, we write

$$u >^L v \qquad \text{or} \qquad u <^L v.$$

Remark: A vector $u \in \mathbb{R}^n$ is lexicographically positive if

$$u >^L 0,$$

i.e., if $u \neq 0$ and the first nonzero entry of u is positive.

Lexicography

Lexicographic pivoting rule

1. Choose an entering column A_j arbitrarily, as long as its reduced cost $\bar{c}_j$ is negative. Let $u = B^{-1}A_j$ be the jth column of the tableau.
2. For each i with $u_i > 0$, divide the ith row of the tableau (including the entry in the zeroth column) by u_i and choose the lexicographically smallest row. If row ℓ is lexicographically smallest, then the ℓth basic variable $x_{B(\ell)}$ exits the basis.

Example 3.7

- Consider the following tableau (the zeroth row is omitted), and suppose that the pivot column is the third one ($j = 3$).

$$
\begin{array}{r|c|cc c c|}
\hline
x_{B(1)} = & 1 & 0 & 5 & 3 & \cdots \\
x_{B(2)} = & 2 & 4 & 6 & -1 & \cdots \\
x_{B(3)} = & 3 & 0 & 7 & 9 & \cdots \\
\hline
\end{array}
$$

- There is a tie in trying to determine the exiting variable:
  - $x_{B(1)}/u_1 = 1/3$,
  - $x_{B(3)}/u_3 = 3/9 = 1/3$.

Example 3.7

- Consider the following tableau (the zeroth row is omitted), and suppose that the pivot column is the third one ($j = 3$).

$$\begin{array}{r} x_{B(1)} = \\ x_{B(2)} = \\ x_{B(3)} = \end{array} \begin{array}{|c|ccc|c|} \hline 1 & 0 & 5 & 3 & \cdots \\ 2 & 4 & 6 & -1 & \cdots \\ 3 & 0 & 7 & 9 & \cdots \\ \hline \end{array}$$

- We divide the first and third rows of the tableau by $u_1 = 3$ and $u_3 = 9$, respectively, to obtain:

$$\begin{array}{r} x_{B(1)} = \\ x_{B(2)} = \\ x_{B(3)} = \end{array} \begin{array}{|c|ccc|c|} \hline 1/3 & 0 & 5/3 & 1 & \cdots \\ * & * & * & * & \cdots \\ 1/3 & 0 & 7/9 & 1 & \cdots \\ \hline \end{array}$$

Example 3.7

- Consider the following tableau (the zeroth row is omitted), and suppose that the pivot column is the third one ($j = 3$).

$$\begin{array}{r|c|cc c c|}
\cline{2-6}
x_{B(1)} = & 1 & 0 & 5 & 3 & \cdots \\
x_{B(2)} = & 2 & 4 & 6 & -1 & \cdots \\
x_{B(3)} = & 3 & 0 & 7 & 9 & \cdots \\
\cline{2-6}
\end{array}$$

- We divide the first and third rows of the tableau by $u_1 = 3$ and $u_3 = 9$, respectively, to obtain:

$$\begin{array}{r|c|cc c c|}
\cline{2-6}
x_{B(1)} = & 1/3 & 0 & 5/3 & 1 & \cdots \\
x_{B(2)} = & * & * & * & * & \cdots \\
x_{B(3)} = & 1/3 & 0 & 7/9 & 1 & \cdots \\
\cline{2-6}
\end{array}$$

- The tie between the first and third rows is resolved by performing a lexicographic comparison.

Example 3.7

- Consider the following tableau (the zeroth row is omitted), and suppose that the pivot column is the third one ($j = 3$).

$$\begin{array}{r|c|cc|c|c|}
\hline
x_{B(1)} = & 1 & 0 & 5 & 3 & \cdots \\
x_{B(2)} = & 2 & 4 & 6 & -1 & \cdots \\
x_{B(3)} = & 3 & 0 & 7 & 9 & \cdots \\
\hline
\end{array}$$

- We divide the first and third rows of the tableau by $u_1 = 3$ and $u_3 = 9$, respectively, to obtain:

$$\begin{array}{r|c|cc|c|c|}
\hline
x_{B(1)} = & 1/3 & 0 & 5/3 & 1 & \cdots \\
x_{B(2)} = & * & * & * & * & \cdots \\
x_{B(3)} = & 1/3 & 0 & 7/9 & 1 & \cdots \\
\hline
\end{array}$$

- Since $7/9 < 5/3$, the third row is chosen to be the pivot row, and the variable $x_{B(3)}$ exits the basis.

Lexicography

The lexicographic pivoting rule always leads to a unique choice for the exiting variable.

- Indeed, if this were not the case, two of the rows in the tableau would have to be proportional.
- Hence two rows of the matrix $B^{-1}A$ are proportional, and the matrix $B^{-1}A$ does not have linearly independent rows.
- Therefore, also A does not have linearly independent rows.
- This contradicts our standing assumption that A has linearly independent rows.

Lexicography

Theorem 3.4
Suppose that the simplex algorithm starts with all the rows in the simplex tableau, except the zeroth row, lexicographically positive. If the lexicographic pivoting rule is followed, then:

(a) Every row of the tableau, except the zeroth row, remains lexicographically positive throughout the algorithm.

(b) The zeroth row strictly increases lexicographically at each iteration.

(c) The simplex method terminates after a finite number of iterations.

Let's prove it!

Lexicography

Before we start, some simple properties of "$>^L$".

$$
\begin{aligned}
u >^L v &\quad\Leftrightarrow\quad u - v >^L 0 \\
\\
u >^L 0,\ \alpha > 0 &\quad\Rightarrow\quad \alpha u >^L 0 \\
u >^L 0,\ v >^L 0 &\quad\Rightarrow\quad u + v >^L 0 \\
u >^L 0 &\quad\Rightarrow\quad u + v >^L v.
\end{aligned}
$$

Lexicography

Theorem 3.4
Suppose that the simplex algorithm starts with all the rows in the simplex tableau, except the zeroth row, lexicographically positive. If the lexicographic pivoting rule is followed, then:

(a) Every row of the tableau, except the zeroth row, remains lexicographically positive throughout the algorithm.

Proof (a). Suppose that x_j enters the basis and that the pivot row is the ℓ-th. We have $u_\ell > 0$ and

$$\frac{(\text{old } \ell\text{-th row})}{u_\ell} <^L \frac{(\text{old } i\text{-th row})}{u_i}, \text{ if } i \neq \ell \text{ and } u_i > 0. \qquad (*)$$

- (new ℓ-th row) $= \frac{(\text{old } \ell\text{-th row})}{u_\ell}$, still lexicographically positive since $u_\ell > 0$.

Lexicography

Theorem 3.4
Suppose that the simplex algorithm starts with all the rows in the simplex tableau, except the zeroth row, lexicographically positive. If the lexicographic pivoting rule is followed, then:

(a) Every row of the tableau, except the zeroth row, remains lexicographically positive throughout the algorithm.

Proof (a). Suppose that x_j enters the basis and that the pivot row is the ℓ-th. We have $u_\ell > 0$ and

$$\frac{(\text{old } \ell\text{-th row})}{u_\ell} <^L \frac{(\text{old } i\text{-th row})}{u_i}, \text{ if } i \neq \ell \text{ and } u_i > 0. \qquad (*)$$

- (new i-th row) $=$ (old i-th row) $- \frac{u_i}{u_\ell}$(old ℓ-th row) for $i \neq \ell$.

Lexicography

Theorem 3.4
Suppose that the simplex algorithm starts with all the rows in the simplex tableau, except the zeroth row, lexicographically positive. If the lexicographic pivoting rule is followed, then:

(a) Every row of the tableau, except the zeroth row, remains lexicographically positive throughout the algorithm.

Proof (a). Suppose that x_j enters the basis and that the pivot row is the ℓ-th. We have $u_\ell > 0$ and

$$\frac{(\text{old } \ell\text{-th row})}{u_\ell} <^L \frac{(\text{old } i\text{-th row})}{u_i}, \text{ if } i \neq \ell \text{ and } u_i > 0. \qquad (*)$$

- (new i-th row) $=$ (old i-th row) $- \frac{u_i}{u_\ell}$(old ℓ-th row) for $i \neq \ell$.
  - If $u_i \leq 0$, then $-\frac{u_i}{u_\ell} \geq 0 \Rightarrow$ (new i-th row) is lexicographically positive.
  - If $u_i > 0$, then $(*) \Rightarrow$ (new i-th row) is lexicographically positive.

Lexicography

Theorem 3.4
Suppose that the simplex algorithm starts with all the rows in the simplex tableau, except the zeroth row, lexicographically positive. If the lexicographic pivoting rule is followed, then:

(b) The zeroth row strictly increases lexicographically at each iteration.

Proof (b).

- At the beginning of an iteration, the pivot element is positive and the reduced cost in the pivot column is negative.
- To make this reduced cost 0, we add a positive multiple of the pivot row, which is lexicographically positive.
- Thus, the zeroth row increases lexicographically.

Lexicography

Theorem 3.4

Suppose that the simplex algorithm starts with all the rows in the simplex tableau, except the zeroth row, lexicographically positive. If the lexicographic pivoting rule is followed, then:

(c) The simplex method terminates after a finite number of iterations.

Proof (c).

- The zeroth row is completely determined by the current basis.
- From (b), the zeroth row strictly increases lexicographically at each iteration, thus no basis can be repeated twice.
- Since there is a finite number of bases, the simplex method must terminate in a finite number of iterations. $\square$

Lexicography

- The lexicographic pivoting rule is straightforward to use if the simplex method is implemented in terms of the full tableau.
- It can also be used in conjunction with the revised simplex method, provided that the inverse basis matrix B^{-1} is formed explicitly.

Lexicography

- In order to apply the lexicographic pivoting rule, an initial tableau with lexicographically positive rows is required.
- Assume that an initial tableau is available (methods for obtaining an initial tableau are discussed in the next section).
- We can rename the variables so that the basic variables are the first m ones.
- This is equivalent to rearranging the tableau so that the first m columns of $B^{-1}A$ are the m unit vectors.
- The resulting tableau has lexicographically positive rows, as desired.

Bland's rule

Bland's rule

The smallest subscript pivoting rule, also known as Bland's rule, is as follows.

Smallest subscript pivoting rule

1. Find the smallest j for which the reduced cost $\bar{c}_j$ is negative and have the column A_j enter the basis.
2. Out of all variables x_i that are tied in the test for choosing an exiting variable, select the one with the smallest i.

Remark: Selecting the variable x_i with the smallest i is not the same as selecting the variable $x_{B(i)}$ with the smallest i.

Bland's rule

- This pivoting rule is compatible with an implementation of the revised simplex method in which the reduced costs of the nonbasic variables are computed one at a time, in the natural order, until a negative one is discovered.
- Under this pivoting rule, it is known that cycling never occurs and the simplex method is guaranteed to terminate after a finite number of iterations.

Theorem (Termination with Bland's rule)
If the simplex method uses Bland's rule, it terminates after a finite number of iterations.

Let's prove it!

3.5 Finding an initial basic feasible solution

Finding an initial basic feasible solution

- In order to start the simplex method, we need to find an initial basic feasible solution.
- Sometimes this is straightforward, like in Example 3.5.
- More generally, suppose that we are dealing with a problem involving constraints of the form

$$\begin{aligned} Ax &\leq b \\ x &\geq 0, \end{aligned}$$

  where $b \geq 0$.
- We can then introduce nonnegative slack variables s and rewrite the constraints in the form

$$\begin{aligned} Ax + s &= b \\ x, s &\geq 0. \end{aligned}$$

Finding an initial basic feasible solution

- In order to start the simplex method, we need to find an initial basic feasible solution.
- Sometimes this is straightforward, like in Example 3.5.
- More generally, suppose that we are dealing with a problem involving constraints of the form

$$\begin{aligned} Ax &\leq b \\ x &\geq 0, \end{aligned}$$

  where $b \geq 0$.
- We can then introduce nonnegative slack variables s and rewrite the constraints in the form

$$\begin{aligned} Ax + s &= b \\ x, s &\geq 0. \end{aligned}$$

- The vector (x, s) defined by $x = 0$ and $s = b$ is a basic feasible solution and the corresponding basis matrix is the identity.

Finding an initial basic feasible solution

- In general, finding an initial basic feasible solution is not easy.
- It can be done by solving an auxiliary LP problem.
- Let's see how!

Finding an initial basic feasible solution

- Consider the problem

$$\begin{aligned} \text{minimize} \quad & c'x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

- By possibly multiplying some of the equality constraints by -1, we can assume, without loss of generality, that $b \geq 0$.

Finding an initial basic feasible solution

- Consider the problem

$$\begin{aligned} \text{minimize} \quad & c'x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0. \end{aligned}$$

- By possibly multiplying some of the equality constraints by -1, we can assume, without loss of generality, that $b \geq 0$.
- We now introduce a vector $y \in \mathbb{R}^m$ of artificial variables and use the simplex method to solve the auxiliary problem

$$\begin{aligned} \text{minimize} \quad & y_1 + y_2 + \cdots + y_m \\ \text{subject to} \quad & Ax + y = b \\ & x \geq 0 \\ & y \geq 0. \end{aligned}$$

- Initialization is easy for the auxiliary problem: by letting $x = 0$ and $y = b$, we have a basic feasible solution and the corresponding basis matrix is the identity.

Finding an initial basic feasible solution

$$\begin{aligned} \text{minimize} \quad & c'x \\ \text{subject to} \quad & Ax = b \\ & x \geq 0 \end{aligned} \qquad \begin{aligned} \text{minimize} \quad & y_1 + y_2 + \cdots + y_m \\ \text{subject to} \quad & Ax + y = b \\ & x \geq 0 \\ & y \geq 0 \end{aligned}$$

- If x is a feasible solution to the original problem, this choice of x together with $y = 0$, yields a zero cost solution to the auxiliary problem.
- Therefore, if the optimal cost in the auxiliary problem is nonzero, we conclude that the original problem is infeasible.
- If we obtain a zero cost solution to the auxiliary problem, it must satisfy $y = 0$, and x is a feasible solution to the original problem.

Finding an initial basic feasible solution

- We have a method that either detects infeasibility or finds a feasible solution to the original problem.
- However, in order to initialize the simplex method for the original problem, we need
  - a basic feasible solution,
  - an associated basis matrix B, and
  - the corresponding tableau (depending on the implementation).

Finding an initial basic feasible solution

- All this is straightforward if the simplex method, applied to the auxiliary problem, terminates with a basis matrix B consisting exclusively of columns of A.

$\cancel{0}$	$\cancel{c' - c'_B B^{-1} A}$	$\cancel{c' - c'_B B^{-1}}$
$B^{-1} b$	$B^{-1} A$	$\cancel{B^{-1}}$

$$\downarrow$$

$-c'_B B^{-1} b$	$c' - c'_B B^{-1} A$
$B^{-1} b$	$B^{-1} A$

- We drop the columns that correspond to the artificial variables.
- We use B as the starting basis matrix.
- We recompute the zeroth row using the original cost vector c.
- We continue with the simplex method on the original problem.

Driving artificial variables out of the basis

Driving artificial variables out of the basis

The situation is more complex if:

- the original problem is feasible, and
- the simplex method applied to the auxiliary problem terminates with a feasible solution x^* to the original problem, where some of the artificial variables are in the final basis.

Since the final value of the artificial variables is zero, this implies that we have a degenerate basic feasible solution to the auxiliary problem.

- Our task is to obtain a different basis of x^* consisting only of columns of A.

Driving artificial variables out of the basis

- Let k be the number of columns of A that belong to the final basis ($k < m$) and, without loss of generality, assume that these are the columns $A_{B(1)}, \ldots, A_{B(k)}$.
- The columns $A_{B(1)}, \ldots, A_{B(k)}$ must be linearly independent since they are part of a basis.

Driving artificial variables out of the basis

- Let k be the number of columns of A that belong to the final basis ($k < m$) and, without loss of generality, assume that these are the columns $A_{B(1)}, \ldots, A_{B(k)}$.
- The columns $A_{B(1)}, \ldots, A_{B(k)}$ must be linearly independent since they are part of a basis.
- Under our standard assumption that the matrix A has full rank, the columns of A span $\mathbb{R}^m$, and we can choose $m-k$ additional columns $A_{B(k+1)}, \ldots, A_{B(m)}$ of A, to obtain a set of m linearly independent columns, that is, a basis consisting exclusively of columns of A.

Driving artificial variables out of the basis

- Let k be the number of columns of A that belong to the final basis ($k < m$) and, without loss of generality, assume that these are the columns $A_{B(1)}, \ldots, A_{B(k)}$.
- The columns $A_{B(1)}, \ldots, A_{B(k)}$ must be linearly independent since they are part of a basis.
- Under our standard assumption that the matrix A has full rank, the columns of A span $\mathbb{R}^m$, and we can choose $m - k$ additional columns $A_{B(k+1)}, \ldots, A_{B(m)}$ of A, to obtain a set of m linearly independent columns, that is, a basis consisting exclusively of columns of A.
- With this basis, all nonbasic components of x^* are at zero level, and it follows that x^* is the basic feasible solution associated with this new basis as well (cf. Theorem 2.4).
- At this point, the artificial variables and the corresponding columns of the tableau can be dropped.

Driving artificial variables out of the basis

- The procedure we have just described is called

  driving the artificial variables out of the basis,

  and depends crucially on the assumption that the matrix A has rank m.
- If A has rank less than m, constructing a basis for $\mathbb{R}^m$ using the columns of A is impossible and there exist redundant equality constraints that must be eliminated, as described by Theorem 2.5.
- All of the above can be carried out mechanically, in terms of the simplex tableau, in the following manner.

Driving artificial variables out of the basis

- Suppose that the ℓth basic variable is an artificial variable, which is in the basis at zero level.
- We examine the ℓth row of the tableau and search for some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\dots$	$\bar{c}_n$	$\dots$	0	$\dots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\dots$	$(B^{-1}A_n)_1$	$\dots$	0	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\dots$	$(B^{-1}A_n)_\ell$	$\dots$	1	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\dots$	$(B^{-1}A_n)_m$	$\dots$	0	$\dots$

- We either find this index j or not. We consider separately these two cases.

Driving artificial variables out of the basis

Case 1. We find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\ldots$	$\bar{c}_j$	$\ldots$	$\bar{c}_n$	$\ldots$	0	$\ldots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\ldots$	$(B^{-1}A_j)_1$	$\ldots$	$(B^{-1}A_n)_1$	$\ldots$	0	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\ldots$	$(B^{-1}A_j)_\ell$	$\ldots$	$(B^{-1}A_n)_\ell$	$\ldots$	1	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\ldots$	$(B^{-1}A_j)_m$	$\ldots$	$(B^{-1}A_n)_m$	$\ldots$	0	$\ldots$

Driving artificial variables out of the basis

Case 1. We find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\dots$	$\bar{c}_j$	$\dots$	$\bar{c}_n$	$\dots$	0	$\dots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\dots$	$(B^{-1}A_j)_1$	$\dots$	$(B^{-1}A_n)_1$	$\dots$	0	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\dots$	$(B^{-1}A_j)_\ell$	$\dots$	$(B^{-1}A_n)_\ell$	$\dots$	1	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\dots$	$(B^{-1}A_j)_m$	$\dots$	$(B^{-1}A_n)_m$	$\dots$	0	$\dots$

- We claim that A_j is linearly independent from the columns $A_{B(1)}, \dots, A_{B(k)}$.

Driving artificial variables out of the basis

Case 1. We find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\dots$	$\bar{c}_j$	$\dots$	$\bar{c}_n$	$\dots$	0	$\dots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\dots$	$(B^{-1}A_j)_1$	$\dots$	$(B^{-1}A_n)_1$	$\dots$	0	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\dots$	$(B^{-1}A_j)_\ell$	$\dots$	$(B^{-1}A_n)_\ell$	$\dots$	1	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\dots$	$(B^{-1}A_j)_m$	$\dots$	$(B^{-1}A_n)_m$	$\dots$	0	$\dots$

- To see this, note that $B^{-1}A_{B(i)} = e_i$, $i = 1, \dots, k$, and since $k < \ell$, the ℓth entry of these vectors is zero.
- Since the ℓth entry of $B^{-1}A_j$ is nonzero, this vector is not a linear combination of the vectors $B^{-1}A_{B(1)}, \dots, B^{-1}A_{B(k)}$.
- Equivalently, A_j is not a linear combination of the vectors $A_{B(1)}, \dots, A_{B(k)}$, which proves our claim.

Driving artificial variables out of the basis

Case 1. We find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\ldots$	$\bar{c}_j$	$\ldots$	$\bar{c}_n$	$\ldots$	0	$\ldots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\ldots$	$(B^{-1}A_j)_1$	$\ldots$	$(B^{-1}A_n)_1$	$\ldots$	0	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\ldots$	$(B^{-1}A_j)_\ell$	$\ldots$	$(B^{-1}A_n)_\ell$	$\ldots$	1	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\ldots$	$(B^{-1}A_j)_m$	$\ldots$	$(B^{-1}A_n)_m$	$\ldots$	0	$\ldots$

- We now bring A_j into the basis and have the ℓth basic variable exit the basis.

Driving artificial variables out of the basis

Case 1. We find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\ldots$	$\bar{c}_j$	$\ldots$	$\bar{c}_n$	$\ldots$	0	$\ldots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\ldots$	$(B^{-1}A_j)_1$	$\ldots$	$(B^{-1}A_n)_1$	$\ldots$	0	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\ldots$	$(B^{-1}A_j)_\ell$	$\ldots$	$(B^{-1}A_n)_\ell$	$\ldots$	1	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\ldots$	$(B^{-1}A_j)_m$	$\ldots$	$(B^{-1}A_n)_m$	$\ldots$	0	$\ldots$

- This is accomplished in the usual manner: perform those elementary row operations that replace $B^{-1}A_j$ by the ℓth unit vector.
- The only difference from the usual mechanics of the simplex method is that the pivot element could be negative.

Driving artificial variables out of the basis

Case 1. We find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\dots$	$\bar{c}_j$	$\dots$	$\bar{c}_n$	$\dots$	0	$\dots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\dots$	$(B^{-1}A_j)_1$	$\dots$	$(B^{-1}A_n)_1$	$\dots$	0	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	$(B^{-1}A_1)_\ell$	$\dots$	$(B^{-1}A_j)_\ell$	$\dots$	$(B^{-1}A_n)_\ell$	$\dots$	1	$\dots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\dots$	$(B^{-1}A_j)_m$	$\dots$	$(B^{-1}A_n)_m$	$\dots$	0	$\dots$

- Because $x_{B(\ell)} = 0$, adding a multiple of the ℓth row to the other rows does not change the values of the basic variables.
- This means that after the change of basis, we are still at the same basic feasible solution to the auxiliary problem, but we have reduced the number of basic artificial variables by one.

Driving artificial variables out of the basis

Case 2. We cannot find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\ldots$	$\bar{c}_n$	$\ldots$	0	$\ldots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\ldots$	$(B^{-1}A_n)_1$	$\ldots$	0	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	0	$\ldots$	0	$\ldots$	1	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\ldots$	$(B^{-1}A_n)_m$	$\ldots$	0	$\ldots$

- In this case, the ℓth row of $B^{-1}A$ is zero.
- Note that the ℓth row of $B^{-1}A$ is equal to $g'A$, where g' is the ℓth row of B^{-1}.
- Hence, $g'A = 0'$ for some nonzero vector g, and the matrix A has linearly dependent rows.

Driving artificial variables out of the basis

Case 2. We cannot find some j such that the ℓth entry of $B^{-1}A_j$ is nonzero.

0	$\bar{c}_1$	$\ldots$	$\bar{c}_n$	$\ldots$	0	$\ldots$
$x_{B(1)}$	$(B^{-1}A_1)_1$	$\ldots$	$(B^{-1}A_n)_1$	$\ldots$	0	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$x_{B(\ell)}$	0	$\ldots$	0	$\ldots$	1	$\ldots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$	
$x_{B(m)}$	$(B^{-1}A_1)_m$	$\ldots$	$(B^{-1}A_n)_m$	$\ldots$	0	$\ldots$

- Since we are dealing with a feasible problem (why?), we must also have $g'b = 0$.
- Thus, the constraint $g'Ax = g'b$ is redundant and can be eliminated (cf. Theorem 2.5 in Section 2.3).
- Since this constraint is the information provided by the ℓth row of the tableau, we can eliminate that row and continue from there.

Example 3.8

Consider the LP problem

$$\begin{aligned}
\text{minimize} \quad & x_1 + x_2 + x_3 \\
\text{subject to} \quad & x_1 + 2x_2 + 3x_3 = 3 \\
& x_1 - 2x_2 - 6x_3 = -2 \\
& 4x_2 + 9x_3 = 5 \\
& 3x_3 + x_4 = 1 \\
& x_1, \ldots, x_4 \geq 0.
\end{aligned}$$

Example 3.8

Consider the LP problem

$$\begin{aligned} \text{minimize} \quad & x_1 + x_2 + x_3 \\ \text{subject to} \quad & x_1 + 2x_2 + 3x_3 = 3 \\ & x_1 - 2x_2 - 6x_3 = -2 \\ & 4x_2 + 9x_3 = 5 \\ & 3x_3 + x_4 = 1 \\ & x_1, \ldots, x_4 \geq 0. \end{aligned}$$

In order to find a feasible solution, we form the auxiliary problem

$$\begin{aligned} \text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\ \text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\ & -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\ & 4x_2 + 9x_3 + x_7 = 5 \\ & 3x_3 + x_4 + x_8 = 1 \\ & x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0. \end{aligned}$$

Example 3.8

$$\begin{aligned} \text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\ \text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\ & -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\ & 4x_2 + 9x_3 + x_7 = 5 \\ & 3x_3 + x_4 + x_8 = 1 \\ & x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0. \end{aligned}$$

- A basic feasible solution to the auxiliary problem is obtained by letting

$$\begin{aligned} x_1, x_2, x_3, x_4 &= 0, \\ (x_5, x_6, x_7, x_8) &= b = (3, 2, 5, 1). \end{aligned}$$

- The corresponding basis matrix is the identity, and the cost of the solution is 11.

Example 3.8

$$
\begin{aligned}
\text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\
\text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& 4x_2 + 9x_3 + x_7 = 5 \\
& 3x_3 + x_4 + x_8 = 1 \\
& x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0.
\end{aligned}
$$

- Furthermore, we have $c' = [0' \mid 1']$, $c_B' = 1'$.
- The vector of reduced costs in the auxiliary problem is

$$c' - c_B' B^{-1} [A \mid I] = [0' \mid 1'] - 1' I [A \mid I]$$

Example 3.8

$$
\begin{aligned}
\text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\
\text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& 4x_2 + 9x_3 + x_7 = 5 \\
& 3x_3 + x_4 + x_8 = 1 \\
& x_1, \dots, x_4, x_5, \dots, x_8 \geq 0.
\end{aligned}
$$

- Furthermore, we have $c' = [0' \mid 1']$, $c'_B = 1'$.
- The vector of reduced costs in the auxiliary problem is

$$
\begin{aligned}
c' - c'_B B^{-1}[A \mid I] &= [0' \mid 1'] - 1' I[A \mid I] \\
&= [-1'A \mid 1' - 1']
\end{aligned}
$$

Example 3.8

$$\begin{aligned} \text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\ \text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\ & -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\ & 4x_2 + 9x_3 + x_7 = 5 \\ & 3x_3 + x_4 + x_8 = 1 \\ & x_1, \dots, x_4, x_5, \dots, x_8 \geq 0. \end{aligned}$$

- Furthermore, we have $c' = [0' \mid 1']$, $c'_B = 1'$.
- The vector of reduced costs in the auxiliary problem is

$$\begin{aligned} c' - c'_B B^{-1}[A \mid I] &= [0' \mid 1'] - 1' I [A \mid I] \\ &= [-1'A \mid 1' - 1'] = [-1'A \mid 0']. \end{aligned}$$

Example 3.8

$$
\begin{aligned}
\text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\
\text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& 4x_2 + 9x_3 + x_7 = 5 \\
& 3x_3 + x_4 + x_8 = 1 \\
& x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0.
\end{aligned}
$$

- Furthermore, we have $c' = [0' \mid 1']$, $c'_B = 1'$.
- The vector of reduced costs in the auxiliary problem is

$$
\begin{aligned}
c' - c'_B B^{-1}[A \mid I] &= [0' \mid 1'] - 1' I [A \mid I] \\
&= [-1'A \mid 1' - 1'] = [-1'A \mid 0'].
\end{aligned}
$$

- We form the initial tableau:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-11	0	-8	-21	-1	0	0	0	0
$x_5 =$	3	1	2	3	0	1	0	0	0
$x_6 =$	2	-1	2	6	0	0	1	0	0
$x_7 =$	5	0	4	9	0	0	0	1	0
$x_8 =$	1	0	0	3	1	0	0	0	1

- We bring x_4 into the basis and have x_8 exit the basis.
- The basis matrix B is still the identity and only the zeroth row of the tableau changes.
- We obtain:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-11	0	-8	-21	-1	0	0	0	0
$x_5 =$	3	1	2	3	0	1	0	0	0
$x_6 =$	2	-1	2	6	0	0	1	0	0
$x_7 =$	5	0	4	9	0	0	0	1	0
$x_8 =$	1	0	0	3	1	0	0	0	1

- We bring x_4 into the basis and have x_8 exit the basis.
- The basis matrix B is still the identity and only the zeroth row of the tableau changes.
- We obtain:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-10	0	-8	-18	0	0	0	0	1
$x_5 =$	3	1	2	3	0	1	0	0	0
$x_6 =$	2	-1	2	6	0	0	1	0	0
$x_7 =$	5	0	4	9	0	0	0	1	0
$x_4 =$	1	0	0	3	1	0	0	0	1

- We now bring x_3 into the basis and have x_4 exit the basis.
- The new tableau is:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-10	0	-8	-18	0	0	0	0	1
$x_5 =$	3	1	2	3	0	1	0	0	0
$x_6 =$	2	-1	2	6	0	0	1	0	0
$x_7 =$	5	0	4	9	0	0	0	1	0
$x_4 =$	1	0	0	3	1	0	0	0	1

- We now bring x_3 into the basis and have x_4 exit the basis.
- The new tableau is:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-4	0	-8	0	6	0	0	0	7
$x_5 =$	2	1	2	0	-1	1	0	0	-1
$x_6 =$	0	-1	2	0	-2	0	1	0	-2
$x_7 =$	2	0	4	0	-3	0	0	1	-3
$x_3 =$	$1/3$	0	0	1	$1/3$	0	0	0	$1/3$

- We now bring x_2 into the basis and x_6 exits.
- Note that this is a degenerate pivot with $\theta^* = 0$.
- The new tableau is:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-4	0	-8	0	6	0	0	0	7
$x_5 =$	2	1	2	0	-1	1	0	0	-1
$x_6 =$	0	-1	2	0	-2	0	1	0	-2
$x_7 =$	2	0	4	0	-3	0	0	1	-3
$x_3 =$	$1/3$	0	0	1	$1/3$	0	0	0	$1/3$

- We now bring x_2 into the basis and x_6 exits.
- Note that this is a degenerate pivot with $\theta^* = 0$.
- The new tableau is:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-4	-4	0	0	-2	0	4	0	-1
$x_5 =$	2	2	0	0	1	1	-1	0	1
$x_2 =$	0	$-1/2$	1	0	-1	0	$1/2$	0	-1
$x_7 =$	2	2	0	0	1	0	-2	1	1
$x_3 =$	$1/3$	0	0	1	$1/3$	0	0	0	$1/3$

- We now have x_1 enter the basis and x_5 exit the basis.
- We obtain the following tableau:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	-4	-4	0	0	-2	0	4	0	-1
$x_5 =$	2	2	0	0	1	1	-1	0	1
$x_2 =$	0	$-1/2$	1	0	-1	0	$1/2$	0	-1
$x_7 =$	2	2	0	0	1	0	-2	1	1
$x_3 =$	$1/3$	0	0	1	$1/3$	0	0	0	$1/3$

- We now have x_1 enter the basis and x_5 exit the basis.
- We obtain the following tableau:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	0	0	0	0	0	2	2	0	1
$x_1 =$	1	1	0	0	$1/2$	$1/2$	$-1/2$	0	$1/2$
$x_2 =$	$1/2$	0	1	0	$-3/4$	$1/4$	$1/4$	0	$-3/4$
$x_7 =$	0	0	0	0	0	-1	-1	1	0
$x_3 =$	$1/3$	0	0	1	$1/3$	0	0	0	$1/3$

- The cost in the auxiliary problem has dropped to zero: we have a feasible solution to the original problem.
- The artificial variable x_7 is still in the basis, at zero level.
- In order to obtain a basic feasible solution to the original problem and the corresponding basis, we need to drive x_7 out of the basis.

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	0	0	0	0	0	2	2	0	1
$x_1 =$	1	1	0	0	1/2	1/2	−1/2	0	1/2
$x_2 =$	1/2	0	1	0	−3/4	1/4	1/4	0	−3/4
$x_7 =$	0	0	0	0	0	−1	−1	1	0
$x_3 =$	1/3	0	0	1	1/3	0	0	0	1/3

- x_7 is the third basic variable and the third row of $B^{-1}A$ is zero.
- This indicates that the matrix A has linearly dependent rows (Case 2).
- We remove the third row of the tableau, because it corresponds to a redundant constraint.
- The new tableau is:

Example 3.8

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
	0	0	0	0	0	2	2	0	1
$x_1 =$	1	1	0	0	$1/2$	$1/2$	$-1/2$	0	$1/2$
$x_2 =$	$1/2$	0	1	0	$-3/4$	$1/4$	$1/4$	0	$-3/4$
$x_3 =$	$1/3$	0	0	1	$1/3$	0	0	0	$1/3$

- There are no more artificial variables in the basis.
- Thus we can obtain an initial tableau for the original problem by removing all of the artificial variables.

Example 3.8

		x_1	x_2	x_3	x_4
	0	0	0	0	0
$x_1 =$	1	1	0	0	$1/2$
$x_2 =$	$1/2$	0	1	0	$-3/4$
$x_3 =$	$1/3$	0	0	1	$1/3$

- We compute the reduced costs of the original variables

$$\bar{c}' = c' - c_B' B^{-1} A.$$

- The original cost vector is $c = (1, 1, 1, 0)$, so $c_B = (1, 1, 1)$.
- The matrix $B^{-1}A$ is the tableau without the zeroth row and zeroth column.
- The vector of reduced costs is then $\bar{c} = (0, 0, 0, -1/12)$, and the cost of the solution $(1, 1/2, 1/3)$ is $11/6$.
- We fill in the zeroth row of the tableau and obtain:

Example 3.8

$$
\begin{array}{r|r|rrrr|}
 & & x_1 & x_2 & x_3 & x_4 \\
\cline{2-6}
 & -11/6 & 0 & 0 & 0 & -1/12 \\
\cline{2-6}
x_1 = & 1 & 1 & 0 & 0 & 1/2 \\
x_2 = & 1/2 & 0 & 1 & 0 & -3/4 \\
x_3 = & 1/3 & 0 & 0 & 1 & 1/3 \\
\cline{2-6}
\end{array}
$$

- We can now start executing the simplex method on the original problem.
- Exercise: Do it!

Example 3.8

$$\begin{aligned}
\text{minimize} \quad & x_5 + x_6 + x_7 + x_8 \\
\text{subject to} \quad & x_1 + 2x_2 + 3x_3 + x_5 = 3 \\
& -x_1 + 2x_2 + 6x_3 + x_6 = 2 \\
& 4x_2 + 9x_3 + x_7 = 5 \\
& 3x_3 + x_4 + x_8 = 1 \\
& x_1, \ldots, x_4, x_5, \ldots, x_8 \geq 0.
\end{aligned}$$

- We observe that in this example, the artificial variable x_8 was unnecessary.
- Instead of starting with $x_8 = 1$, we could have started with $x_4 = 1$ thus eliminating the need for the first pivot.
- More generally, whenever there is a variable that appears in a single constraint and with a positive coefficient, we can always let that variable be in the initial basis and we do not have to associate an artificial variable with that constraint.

The two-phase simplex method

The two-phase simplex method

We can now summarize a complete algorithm for LP problems in standard form.

Phase I:

1. By multiplying some of the constraints by -1, change the problem so that $b \geq 0$.
2. Introduce artificial variables $y_1, \ldots, y_m$, if necessary, and apply the simplex method to the auxiliary problem with cost $\sum_{i=1}^{m} y_i$.
3. If the optimal cost in the auxiliary problem is positive, the original problem is infeasible and the algorithm terminates.

The two-phase simplex method

Phase I:

4. If the optimal cost in the auxiliary problem is zero, a feasible solution to the original problem has been found. If no artificial variable is in the final basis, the artificial variables and the corresponding columns are eliminated, and a feasible basis for the original problem is available.
5. If the ℓth basic variable is an artificial one, examine the ℓth entry of the columns $B^{-1}A_j$, $j = 1, \ldots, n$.
   - If all of these entries are zero, the ℓth row represents a redundant constraint and is eliminated.
   - Otherwise, if the ℓth entry of the jth column is nonzero, apply a change of basis (with this entry serving as the pivot element): the ℓth basic variable exits and x_j enters the basis.

   Repeat this operation until all artificial variables are driven out of the basis.

The two-phase simplex method

Phase II:

1. Let the final basis obtained from **Phase I** be the initial basis for **Phase II**.
2. Let the initial tableau for **Phase II** be obtained from the final tableau of **Phase I** by discarding the columns corresponding to the artificial variables and the zeroth row.
3. Compute the cost of the feasible solution and the reduced costs of all variables for this initial basis, using the cost coefficients of the original problem.
4. Apply the simplex method to the original problem.

The two-phase simplex method

- The two-phase simplex algorithm is a complete method, in the sense that it can handle all possible outcomes.
- As long as cycling is avoided (due to either nondegeneracy, an anticycling rule, or luck), one of the following possibilities will materialize:
  (a) If the problem is infeasible, this is detected at the end of **Phase I**.
  (b) If the problem is feasible but the rows of A are linearly dependent, this is detected and corrected at the end of **Phase I**, by eliminating redundant equality constraints.
  (c) If the optimal cost is equal to $-\infty$, this is detected while running **Phase II**.
  (d) Else, **Phase II** terminates with an optimal solution.

3.6 Column geometry and the simplex method

Column geometry and the simplex method

- We introduce an alternative way of visualizing the workings of the simplex method.
- We consider the problem

$$\begin{aligned} \text{minimize} \quad & c'x \\ \text{subject to} \quad & Ax = b \\ & \sum_{i=1}^{n} x_i = 1 \qquad \text{convexity constraint} \\ & x \geq 0. \end{aligned}$$

  where A is an $m \times n$ matrix.
- This is a special type of a LP problem.
- However, every LP problem with a bounded feasible set can be brought into this form. (Exercise 3.28.)

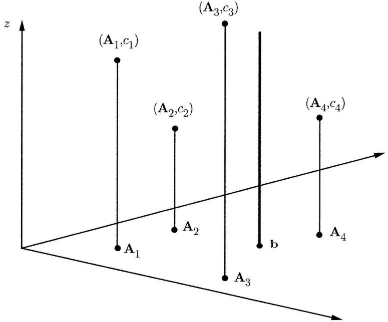
Column geometry and the simplex method

$$
\begin{aligned}
\text{minimize} \quad & c'x \\
\text{subject to} \quad & Ax = b \\
& \sum_{i=1}^{n} x_i = 1 \\
& x \geq 0.
\end{aligned}
$$

- We introduce an auxiliary variable z defined by $z = c'x$.
- Our problem can then be written in the form

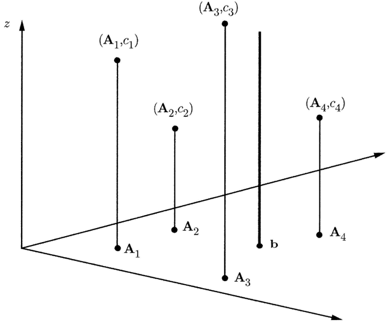
$$
\begin{aligned}
\text{minimize} \quad & z \\
\text{subject to} \quad & x_1 \begin{bmatrix} A_1 \\ c_1 \end{bmatrix} + x_2 \begin{bmatrix} A_2 \\ c_2 \end{bmatrix} + \cdots + x_n \begin{bmatrix} A_n \\ c_n \end{bmatrix} = \begin{bmatrix} b \\ z \end{bmatrix} \\
& \sum_{i=1}^{n} x_i = 1 \\
& x \geq 0.
\end{aligned}
$$

Column geometry and the simplex method



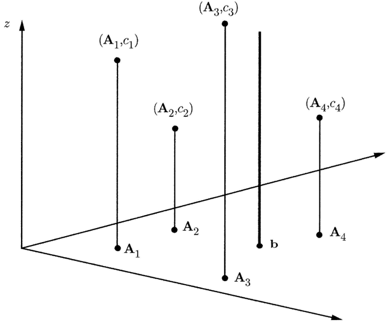
- We view the horizontal plane as an m-dimensional space containing the columns of A.
- We view the vertical axis as the one-dimensional space associated with the cost components c_i.
- Then, each point in the resulting three-dimensional space corresponds to a point (A_i, c_i).

Column geometry and the simplex method



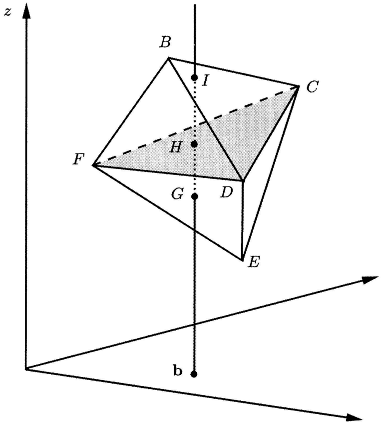
- Our objective is to construct a vector (b, z), which is a convex combination of the vectors (A_i, c_i), such that z is as small as possible.
- Note that the vectors of the form (b, z) lie on a vertical line, which we call the requirement line, and which intersects the horizontal plane at b.

Column geometry and the simplex method



- If the requirement line does not intersect the convex hull of the points (A_i, c_i), the problem is infeasible.
- If it does intersect it, the problem is feasible and an optimal solution corresponds to the lowest point in the intersection of the convex hull and the requirement line.

Column geometry and the simplex method



Example:

- The requirement line intersects the convex hull of the points (A_i, c_i).
- The point G corresponds to an optimal solution.
- The height of G is the optimal cost.

Column geometry and the simplex method

Definition 3.6

(a) A collection of vectors

$$y^1,\ y^2 \dots,\ y^{k+1} \in \mathbb{R}^n$$

are said to be affinely independent if the vectors

$$y^1 - y^{k+1},\ y^2 - y^{k+1}, \dots,\ y^k - y^{k+1}$$

are linearly independent. (Note that we must have $k \leq n$.)

(b) The convex hull of $k+1$ affinely independent vectors in $\mathbb{R}^n$ is called a k-dimensional simplex.

- Three points are either collinear or they are affinely independent and determine a two-dimensional simplex (a triangle).

Column geometry and the simplex method

Definition 3.6

(a) A collection of vectors

$$y^1,\ y^2 \ldots,\ y^{k+1} \in \mathbb{R}^n$$

are said to be affinely independent if the vectors

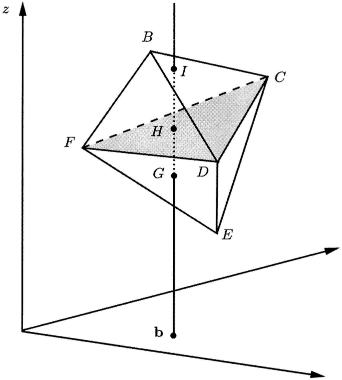
$$y^1 - y^{k+1},\ y^2 - y^{k+1}, \ldots,\ y^k - y^{k+1}$$

are linearly independent. (Note that we must have $k \leq n$.)

(b) The convex hull of $k+1$ affinely independent vectors in $\mathbb{R}^n$ is called a k-dimensional simplex.

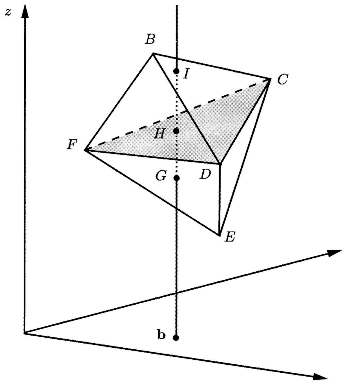
- Four points either lie on the same plane, or they are affinely independent and determine a three-dimensional simplex (a pyramid).

Column geometry and the simplex method



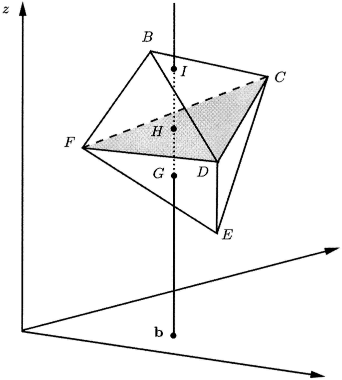
- We now give an interpretation of basic feasible solutions to our problem in this geometry.

Column geometry and the simplex method



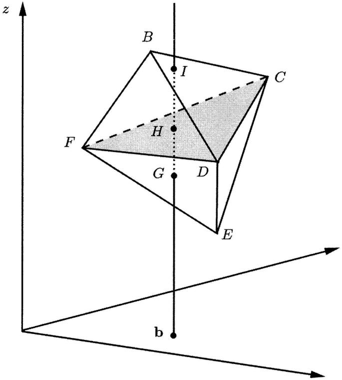
- In our original problem we have $m+1$ equality constraints.
- Thus, a basic feasible solution is associated with a collection of $m+1$ linearly independent columns $(A_i, 1)$ of our LP problem.

Column geometry and the simplex method



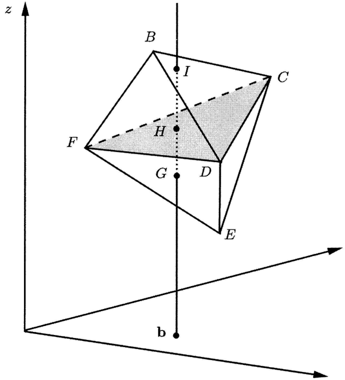
- These are in turn associated with $m + 1$ points (A_i, c_i), which we call basic points; the remaining points (A_i, c_i) are called the nonbasic points.
- Example: A possible choice of basic points is C, D, F.

Column geometry and the simplex method



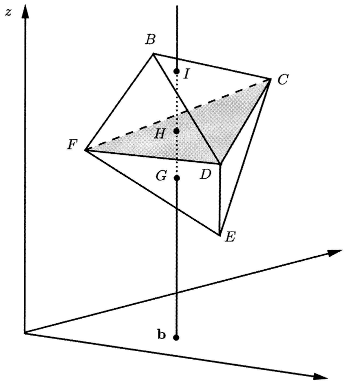
- The $m+1$ basic points are affinely independent (Exercise 3.29) and, therefore, their convex hull is an m-dimensional simplex, which we call the basic simplex.
- Example: The shaded triangle CDF is the basic simplex associated with the basic points C, D, F.

Column geometry and the simplex method



- Let the requirement line intersect the m-dimensional basic simplex at some point (b, z).
- The vector of weights x_i used in expressing (b, z) as a convex combination of the basic points, is the current basic feasible solution, and z represents its cost.

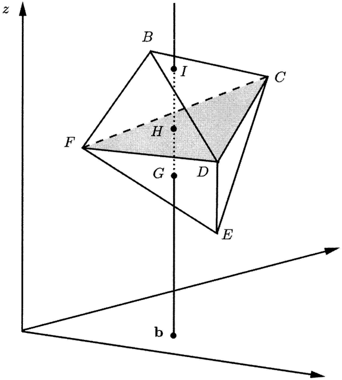
Column geometry and the simplex method



Example:

- The point H corresponds to the basic feasible solution associated with the basic points C, D, F.
- The height of H is the cost of this solution.

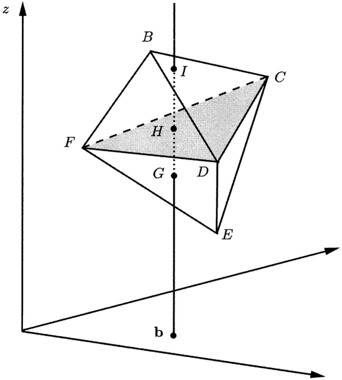
Column geometry and the simplex method



We now interpret a change of basis geometrically.
In a change of basis:

- A new point (A_j, c_j) becomes basic;
- One of the currently basic points is to become nonbasic.

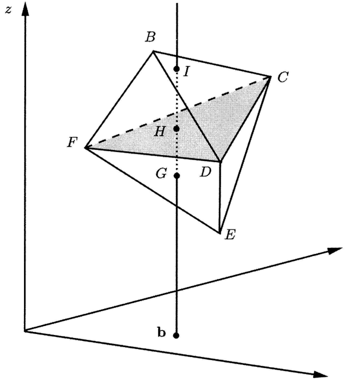
Column geometry and the simplex method



Example:

- Let C, D, F, be the current basic points,
- We could make point B basic, replacing F.
- The new basic simplex would be the convex hull of B, C, D.
- The new basic feasible solution would correspond to point I.

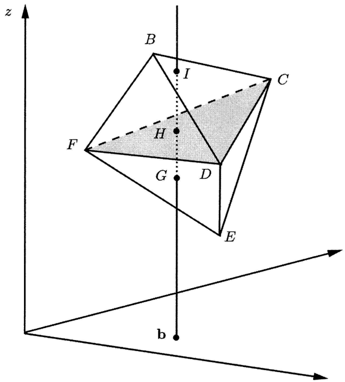
Column geometry and the simplex method



Example:

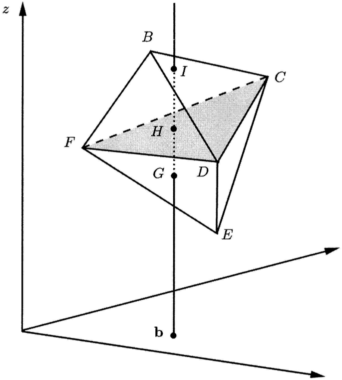
- Alternatively, we could make point E basic, replacing C.
- The new basic simplex would be the convex hull of D, E, F.
- The new basic feasible solution would correspond to G.

Column geometry and the simplex method



- The plane that passes through the basic points is called the dual plane.
- After a change of basis, the cost decreases, if and only if the new basic point is below the dual plane.

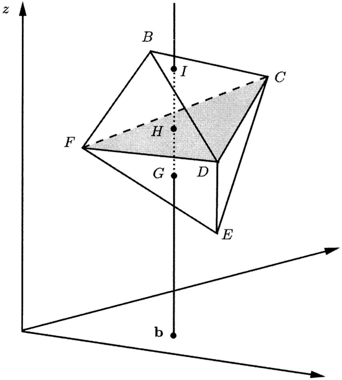
Column geometry and the simplex method



Example:

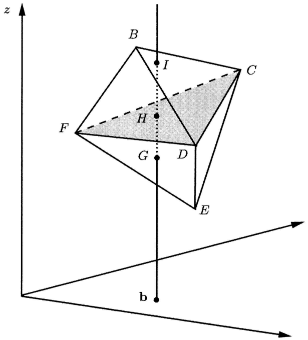
- Point E is below the dual plane and having it enter the basis is profitable.
- This is not the case for point B.

Column geometry and the simplex method



- In fact, the vertical distance from the dual plane to a point (A_j, c_j) is equal to the reduced cost of the associated variable x_j. (Exercise 3.30.)
- Requiring the new basic point to be below the dual plane is therefore equivalent to requiring the entering column to have negative reduced cost.

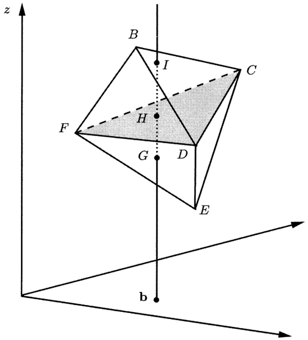
Column geometry and the simplex method



We discuss next the selection of the basic point that will exit the basis.

- Each possible choice of the exiting point leads to a different basic simplex.
- These m basic simplices, together with the original basic simplex (before the change of basis) form the boundary (the faces) of an $(m+1)$-dimensional simplex.

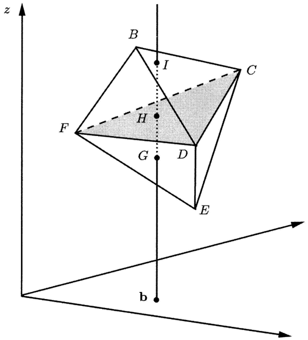
Column geometry and the simplex method



Example:

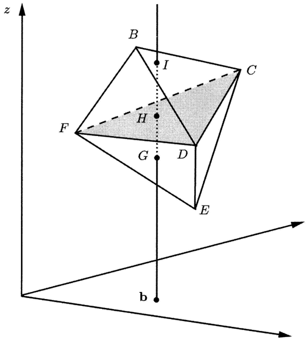
- The basic points C, D, F, determine a two-dimensional basic simplex.
- If point E is to become basic, we obtain a three-dimensional simplex (pyramid) with vertices C, D, E, F.

Column geometry and the simplex method



- The requirement line exits this $(m+1)$-dimensional simplex through its top face and must therefore enter it by crossing some other face.
- This determines which one of the potential basic simplices will be obtained after the change of basis.

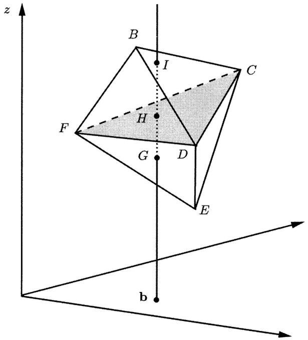
Column geometry and the simplex method



Example:

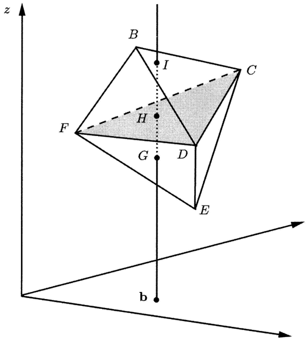
- The requirement line exits the pyramid through its top face with vertices C, D, F.
- It enters the pyramid through the face with vertices D, E, F.
- D, E, F is the new basic simplex and C exits the basis.

Column geometry and the simplex method



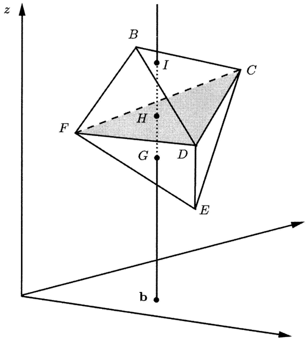
- We can now visualize pivoting through the following physical analogy.
- Think of the original basic simplex with vertices C, D, F, as a solid object anchored at its vertices.

Column geometry and the simplex method



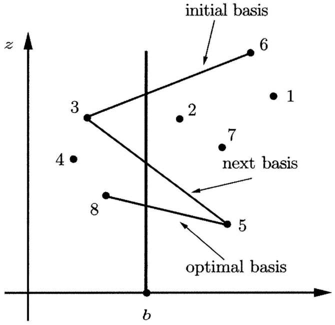
- Grasp the corner of the basic simplex at the vertex C leaving the basis, and pull the corner down to the new basic point E.
- While so moving, the simplex will hinge, or pivot, on its anchor and stretch down to the lower position.

Column geometry and the simplex method



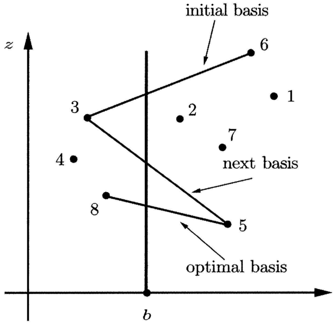
- The terms "simplex" and "pivot" associated with the simplex method have their roots in this column geometry.

Example 3.10



- In this problem we have $m = 1$.
- We use the following pivoting rule: choose a point (A_i, c_i) below the dual plane to become basic, whose vertical distance from the dual plane is largest.
- Exercise 3.30: this is identical to the pivoting rule that selects an entering variable with the most negative reduced cost.

Example 3.10



- Initial basic simplex: $3, 6$.
- Next basic simplex: $3, 5$.
- Next basic simplex: $5, 8$.

3.7 Computational efficiency of the simplex method

Computational efficiency of the simplex method

The computational efficiency of the simplex method is determined by two factors:

(a) The computational effort at each iteration.

(b) The number of iterations.

Computational efficiency of the simplex method

- The computational requirements of each iteration have already been discussed in Section 3.3.
- For example, the full tableau implementation needs $O(mn)$ arithmetic operations per iteration.
- The same is true for the revised simplex method in the worst case.
- We now turn to a discussion of the number of iterations.

The number of iterations in the worst case

The number of iterations in the worst case

- The number of extreme points of the feasible set can increase exponentially with the number of variables and constraints.
- However, it has been observed in practice that the simplex method typically takes only O(m) pivots to find an optimal solution.
- Unfortunately, however, this practical observation is not true for every LP problem.
- We will describe shortly a family of problems for which an exponential number of pivots may be required.

The number of iterations in the worst case

- Recall that for nondegenerate problems, the simplex method always moves from one vertex to an adjacent one, each time improving the value of the cost function.
- We will now describe a polyhedron that has an exponential number of vertices, along with a path that visits all vertices, by taking steps from one vertex to an adjacent one that has lower cost.
- Once such a polyhedron is available, then the simplex method – under a pivoting rule that traces this path – needs an exponential number of pivots.

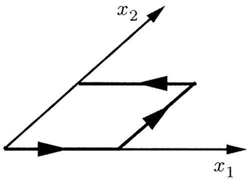
The number of iterations in the worst case

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

$$0 \leq x_i \leq 1, \qquad i = 1, \ldots, n.$$

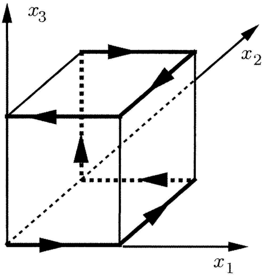
- The unit cube has 2^n vertices: all binary vectors.
- Furthermore, there exists a path that travels along the edges of the cube and which visits each vertex exactly once; we call such a path a spanning path.
- Let's see how a spanning path can be constructed.

The number of iterations in the worst case



This is a spanning path p_2 in the two-dimensional cube.

The number of iterations in the worst case



A spanning path p_3 in the three-dimensional cube can be obtained as follows:

- Split the three-dimensional cube into two two-dimensional cubes (one in $x_3 = 0$ and one in $x_3 = 1$).
- Follow path p_2 in one of them.
- Move to the other cube and follow p_2 in the reverse order.

This construction generalizes and provides a recursive definition of a spanning path for the general n-dimensional cube.

The number of iterations in the worst case

Let us now introduce the cost function $-x_n$.

- Half of the vertices of the cube have zero cost and the other half have a cost of -1.
- Thus, the cost does not decrease strictly with each move along the spanning path.

We do not yet have the desired example!

The number of iterations in the worst case

- However, we choose some $\epsilon \in (0, 1/2)$ and consider the perturbation of the unit cube defined by the constraints

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

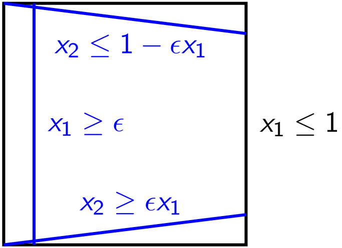
- Then it can be verified that the cost function decreases strictly with each move along a suitably chosen spanning path. Which one?

The number of iterations in the worst case

- However, we choose some $\epsilon \in (0, 1/2)$ and consider the perturbation of the unit cube defined by the constraints

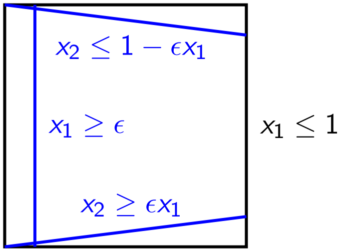
$$\epsilon \le x_1 \le 1,$$
$$\epsilon x_{i-1} \le x_i \le 1 - \epsilon x_{i-1}, \qquad i = 2, \ldots, n.$$

- Then it can be verified that the cost function decreases strictly with each move along a suitably chosen spanning path. Which one?



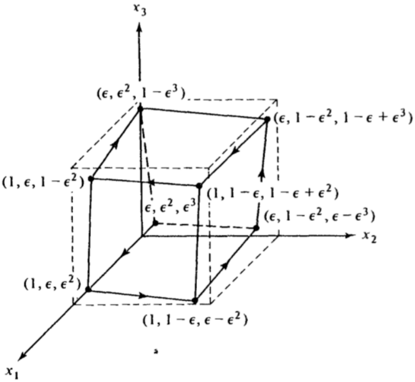
The number of iterations in the worst case

- If we start the simplex method at the first vertex on that spanning path and if our pivoting rule is to always move to the next vertex on that path, then the simplex method will require $2^n - 1$ pivots.
- We summarize this discussion in the following theorem.



The number of iterations in the worst case

- If we start the simplex method at the first vertex on that spanning path and if our pivoting rule is to always move to the next vertex on that path, then the simplex method will require $2^n - 1$ pivots.
- We summarize this discussion in the following theorem.



The number of iterations in the worst case

Theorem 3.5
Consider the LP problem of minimizing $-x_n$ subject to the constraints

$$\begin{aligned} \epsilon \le x_1 &\le 1, \\ \epsilon x_{i-1} \le x_i &\le 1 - \epsilon x_{i-1}, \qquad i = 2, \ldots, n. \end{aligned}$$

Then:

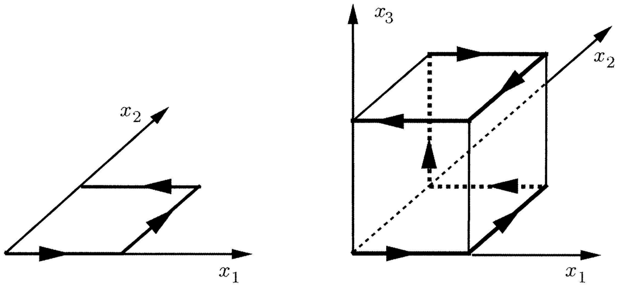
(a) The feasible set has 2^n vertices.

(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.

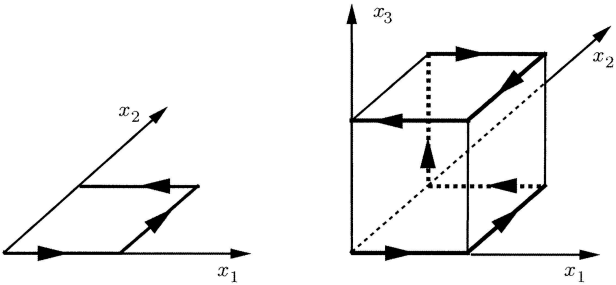
Proof: Exercise 3.32.

The number of iterations in the worst case



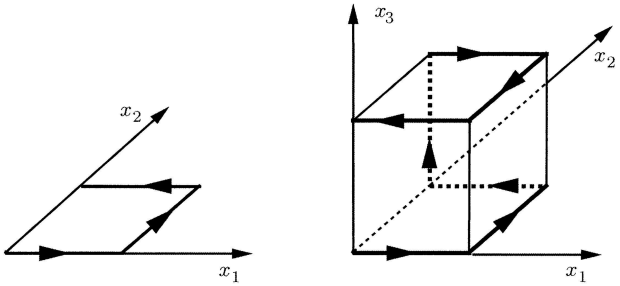
- We observe in the figure the first and the last vertex in the spanning path are adjacent.
- This property persists in the perturbed polyhedron as well.
- Thus, with a different pivoting rule, the simplex method could terminate with a single pivot.

The number of iterations in the worst case



- We are thus led to the following question: is it true that for every pivoting rule there are examples where the simplex method takes an exponential number of iterations?
- For several popular pivoting rules, such examples have been constructed.

The number of iterations in the worst case

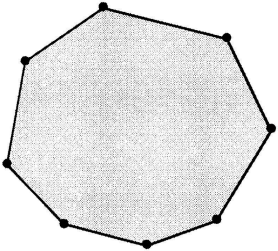


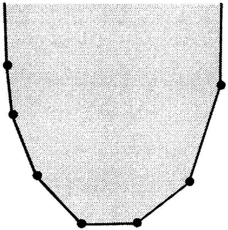
- However, these examples cannot exclude the possibility that some other pivoting rule might fare better.
- This is one of the most important open problems in the theory of LP.
- In the next subsection, we address a closely related issue.

The diameter of polyhedra and the Hirsch conjecture

The diameter of polyhedra and the Hirsch conjecture

- The preceding discussion leads us to the notion of the diameter of a polyhedron P, which is defined as follows.
- Suppose that from any vertex of the polyhedron, we are only allowed to jump to an adjacent vertex.
- We define the distance $d(x, y)$ between two vertices x and y as the minimum number of such jumps required to reach y starting from x.



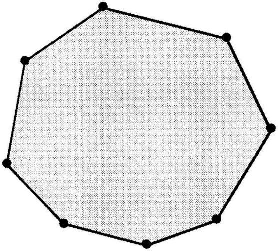


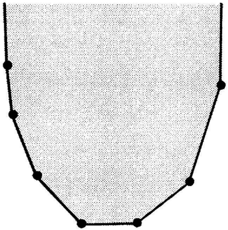
The diameter of polyhedra and the Hirsch conjecture

- The diameter $D(P)$ of the polyhedron P is then defined as the maximum of $d(x, y)$ over all pairs (x, y) of vertices.
- For example, any polyhedron P in $\mathbb{R}^2$ that is represented in terms of m linear inequality constraints is such that:

$D(P) \leq \left\lfloor \frac{m}{2} \right\rfloor$
if P is bounded

$D(P) \leq m - 2$
if P is unbounded





The diameter of polyhedra and the Hirsch conjecture

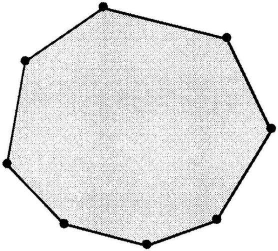
- Suppose that the feasible set P in a LP problem has diameter d.
- Let x and y be two vertices of P such that the distance between x and y is equal to d.
- If the simplex method is initialized at x, and if y happens to be the unique optimal solution, then at least d steps will be required.

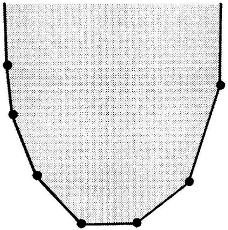
The diameter of polyhedra and the Hirsch conjecture

- We define $\underline{\Delta(n, m)}$ as the maximum of $D(P)$ over all polyhedra in $\mathbb{R}^n$ that are represented in terms of m linear inequality constraints.

$D(P) \leq \left\lfloor \frac{m}{2} \right\rfloor$
if P is bounded

$D(P) \leq m - 2$
if P is unbounded





- Therefore, $\Delta(2, m) = m - 2$.

The diameter of polyhedra and the Hirsch conjecture

- Now, if $\Delta(n, m)$ increases exponentially with n and m, this implies that there exist examples for which the simplex method takes an exponentially increasing number of steps, no matter which pivoting rule is used.
- Thus, in order to have any hope of developing pivoting rules under which the simplex method requires a polynomial number of iterations, we must first establish that $\Delta(n, m)$ grows with n and m at the rate of some polynomial.

The diameter of polyhedra and the Hirsch conjecture

- The practical success of the simplex method has led to the conjecture that $\Delta(n, m)$ does not grow exponentially fast.
- In fact, the following, much stronger, conjecture has been advanced:

Hirsch Conjecture (1957)
$\Delta(n, m) \leq m - n$.

The diameter of polyhedra and the Hirsch conjecture

- The practical success of the simplex method has led to the conjecture that $\Delta(n, m)$ does not grow exponentially fast.
- In fact, the following, much stronger, conjecture has been advanced:

Hirsch Conjecture (1957)
$\Delta(n, m) \leq m - n$.

Bad news: The Hirsch conjecture is false.

- Klee and Walkup, 1967. (Due to unbounded polyhedra.)
- Santos, 2010. (Only considering polytopes.)

The diameter of polyhedra and the Hirsch conjecture

- Even though the Hirsch conjecture is false, we do not know whether the growth of $\Delta(n, m)$ is polynomial or exponential.
- The following (weaker) conjecture has been advanced:

Polynomial Hirsch Conjecture
$\Delta(n, m)$ is bounded above by a polynomial of n and m.

- Despite the significance of $\Delta(n, m)$, we are far from establishing the polynomial Hirsch conjecture.

The diameter of polyhedra and the Hirsch conjecture

- Even though the Hirsch conjecture is false, we do not know whether the growth of $\Delta(n, m)$ is polynomial or exponential.
- The following (weaker) conjecture has been advanced:

Polynomial Hirsch Conjecture
$\Delta(n, m)$ is bounded above by a polynomial of n and m.

- Despite the significance of $\Delta(n, m)$, we are far from establishing the polynomial Hirsch conjecture.
- Question: If the polynomial Hirsch conjecture is true, then is the simplex method a polynomial-time algorithm?

The diameter of polyhedra and the Hirsch conjecture

- Regarding upper bounds, it has been established that the worst-case diameter grows slower than exponentially.
- But the available upper bound grows faster than any polynomial.
- In particular, the following bound is known:

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

The average case behavior of the simplex method

The average case behavior of the simplex method

- Our discussion has been focused on the worst-case behavior of the simplex method, but this is only part of the story.
- Even if every pivoting rule requires an exponential number of iterations in the worst case, this is not necessarily relevant to the typical behavior of the simplex method.
- For this reason, there has been a fair amount of research aiming at an understanding of the average behavior of the simplex method.

The average case behavior of the simplex method

- On average $O(n)$ iterations seem to suffice.
- This has been supported by so-called probabilistic analysis, though it is very hard to come up with a representative probabilistic model for random feasible LP-instances.
- The smoothed analysis of the simplex method shows that bad instances are very non-dense in the set of all possible instances, since tiny random perturbations of the coefficients gives a polynomial number of iterations in expectation.

The average case behavior of the simplex method

- The main difficulty in studying the average behavior of any algorithm lies in defining the meaning of the term "average."
- Basically, one needs to:
    1. Define a probability distribution over the set of all problems of a given size.
    2. Take the mathematical expectation of the number of iterations required by the algorithm, when applied to a random problem drawn according to the postulated probability distribution.
- Unfortunately, there is no natural probability distribution over the set of LP problems.
- Nevertheless, a fair number of positive results have been obtained for a few different types of probability distributions.

The average case behavior of the simplex method

- In one such result, a set of vectors $c, a_1, \ldots, a_m \in \mathbb{R}^n$ and scalars $b_1, \ldots, b_m$ is given.
- For $i = 1, \ldots, m$, we introduce either constraint

$$a_i' x \leq b_i \qquad \text{or} \qquad a_i' x \geq b_i,$$

  with equal probability.
- We then have 2^m possible LP problems, and suppose that L of them are feasible.
- Haimovich (1983) has established that under a rather special pivoting rule, the simplex method requires no more than $n/2$ iterations, on the average over those L feasible problems.
- This linear dependence on the size of the problem agrees with observed behavior.