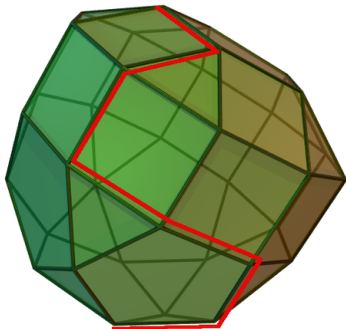
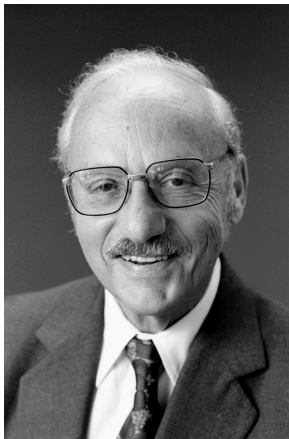


# Lecture 6 Simplex Algorithm

Can Li

ChE 597: Computational Optimization  
Purdue University

## History



Developed by George Dantzig in 1947.

Intuition: iterate over the extreme points of a polyhedron until an optimality condition is satisfied.

# Important results from Polyhedron theory

When  $P$  is nonempty and the optimal objective value is not  $-\infty$ , the optimal solution can always be obtained at an extreme point.  
extreme point  $\Leftrightarrow$  basic feasible solution

**Basic feasible solution** Consider a polyhedron  $P$  defined by linear equality and inequality constraints, and let  $x^*$  be an element of  $\mathbb{R}^n$ .

1. The vector  $x^*$  is a **basic solution** if:
  - All equality constraints are active;
  - Out of the constraints that are active at  $x^*$ , there are  $n$  of them that are linearly independent.
2. If  $x^*$  is a basic solution that satisfies all of the constraints, we say that it is a **basic feasible solution**.

## Basic solution for standard form polyhedron

**Theorem** Consider the constraints  $Ax = b$  and  $x \geq 0$  and assume that the  $m \times n$  matrix  $A$  has linearly independent rows. A vector  $x \in \mathbb{R}^n$  is a basic solution if and only if we have  $Ax = b$ , and there exist indices  $B(1), \dots, B(m)$  such that:

- (a) The columns  $A_{B(1)}, \dots, A_{B(m)}$  are linearly independent;
- (b) If  $i \neq B(1), \dots, B(m)$ , then  $x_i = 0$ .

Proof sketch: **Sufficiency** If (a) and (b) are satisfied, there are  $n$  linearly independent active constraints that uniquely defines a solution and thus is a basic solution.

**Necessity** For a basic solution, there must exist at least  $n - m$  active constraints  $x_i = 0$ , together with  $Ax = b$  forming  $n$  linearly independent constraints. Consider the active constraints forming the matrix (under permutation)  $\begin{bmatrix} B & N \\ 0 & I \end{bmatrix}$  being full rank, where  $A = [B \ N]$ ,  $I$  correspond to the indices of the  $x_i = 0$ .  $B$  must be full rank for the matrix to be full rank.

## Basic variables

**Definition** If  $x$  is a basic solution, the variables  $x_{B(1)}, \dots, x_{B(m)}$  are called **basic variables**; the remaining variables are called **nonbasic**. The columns  $A_{B(1)}, \dots, A_{B(m)}$  are called the **basic column** and, since they are linearly independent, they form a basis of  $\mathbb{R}^m$

## Basis matrix

**basis matrix**  $B \in \mathbb{R}^{m \times m}$  the matrix formed by arranging the  $m$  basic columns next to each other.

**basic variables** a vector  $x_B$  with the values of the basic variables

**nonbasic variables**  $x_N$

$$B = \left[ \begin{array}{c|c|c|c} & & & \\ A_{B(1)} & A_{B(2)} & \cdots & A_{B(m)} \\ & & & \end{array} \right], \quad x_B = \begin{bmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{bmatrix}.$$

$$x = [x_B; x_N], \quad A = [B \ N]$$

The basic variables are determined by solving the equation  $Bx_B = b$  whose unique solution is given by

$$x_B = B^{-1}b.$$

If  $x_B \geq 0$ , then it is a basic feasible solution

Recall that for LP with finite optimum, there exists an optimal basic feasible solution. Next, we will derive the conditions under which a basic feasible solution is optimal.

**Feasible direction** Let  $x$  be an element of a polyhedron  $P$ . A vector  $d \in \mathbb{R}^n$  is said to be a feasible direction at  $x$ , if there exists a positive scalar  $\theta$  for which  $x + \theta d \in P$ .

**Basic direction** Let  $x_B = B^{-1}b$  be a BFS. We consider the possibility of moving away from  $x$ , to a new vector  $x + \theta d$ , by selecting a nonbasic variable  $x_j$ . Algebraically,  $d_j = 1$ , and  $d_i = 0$  for every nonbasic index  $i$  other than  $j$ . We require  $A(x + \theta d) = b$ , and since  $x$  is feasible, we also have  $Ax = b$ . Thus, we need  $Ad = 0$ . Recall now that  $d_j = 1$ , and that  $d_i = 0$  for all other nonbasic indices  $i$ . Then,

$$0 = Ad = \sum_{i=1}^n A_i d_i = \sum_{i=1}^m A_{B(i)} d_{B(i)} + A_j = B d_B + A_j.$$

$$d_B = -B^{-1}A_j, d_j = 1, d_i = 0 \quad \forall i \in N, i \neq j \text{ (jth basic direction)}$$

## When is a basic direction a feasible direction?

Since  $d_j = 1, d_i = 0 \forall i \in N$ , the nonbasic variables will remain feasible. We only need to make sure that the basic variables are still nonnegative.

(a)  $x$  is a nondegenerate basic feasible solution. Then,  $x_B > 0$ , from which it follows that  $x_B + \theta d_B \geq 0$ , and feasibility is maintained, when  $\theta$  is sufficiently small.  $d$  is a feasible direction.

(b)  $x$  is degenerate. Then,  $d$  is not always a feasible direction.

Indeed, it is possible that a basic variable  $x_{B(i)}$  is zero, while the corresponding component  $d_{B(i)}$  of  $d_B = -B^{-1}A_j$  is negative. In that case, if we follow the  $j$ th basic direction, the nonnegativity constraint for  $x_{B(i)}$  is immediately violated, and we are led to infeasible solutions;



## Reduced cost

Q: What will be the effect on the cost (objective) function if we move along the  $j$ th basic direction?

$$d_B = -B^{-1}A_j, d_j = 1, d_i = 0 \quad \forall i \in N, i \neq j \text{ (jth basic direction)}$$

$$\bar{c}_j = c^T d^j = c_j - c_B^T B^{-1} A_j$$

where  $d^j$  is the  $j$ th basic direction,  $c_B$  the vector of costs of the basic variables.  $\bar{c}_j$  is the **reduced cost** of the variable  $x_j$

## Optimality condition

**Theorem** Consider a basic feasible solution  $x$  associated with a basis matrix  $B$ , and let  $\bar{c}$  be the corresponding vector of reduced costs.

- (a) If  $\bar{c} \geq 0$ , then  $x$  is optimal.
- (b) If  $x$  is optimal and nondegenerate, then  $\bar{c} \geq 0$ .

## Optimality condition

Proof. (a) We assume that  $\bar{c} \geq 0$ , we let  $y$  be an arbitrary feasible solution, and we define  $d = y - x$ . Feasibility implies that  $Ax = Ay = b$  and, therefore,  $Ad = 0$ . The latter equality can be rewritten in the form

$$Bd_B + \sum_{i \in N} A_i d_i = 0$$

where  $N$  is the set of indices corresponding to the nonbasic variables under the given basis. Since  $B$  is invertible, we obtain

$$d_B = - \sum_{i \in N} B^{-1} A_i d_i$$

and

$$c^T d = c_B^T d_B + \sum_{i \in N} c_i d_i = \sum_{i \in N} (c_i - c_B^T B^{-1} A_i) d_i = \sum_{i \in N} \bar{c}_i d_i$$

For any nonbasic index  $i \in N$ , we must have  $x_i = 0$  and, since  $y$  is feasible,  $y_i \geq 0$ . Thus,  $d_i \geq 0$  and  $\bar{c}_i d_i \geq 0$ , for all  $i \in N$ . We conclude that  $c^T(y - x) = c^T d \geq 0$ , and since  $y$  was an arbitrary feasible solution,  $x$  is optimal.

## Optimality condition

(b) Suppose that  $x$  is a nondegenerate basic feasible solution and that  $\bar{c}_j < 0$  for some  $j$ . Since the reduced cost of a basic variable is always zero,  $x_j$  must be a nonbasic variable and  $\bar{c}_j$  is the rate of cost change along the  $j$ th basic direction. Since  $x$  is nondegenerate, the  $j$ th basic direction is a feasible direction of cost decrease, as discussed earlier. By moving in that direction, we obtain feasible solutions whose cost is less than that of  $x$ , and  $x$  is not optimal.

## Optimal basis

A basis matrix  $B$  is said to be optimal if:

- (a)  $B^{-1}b \geq 0$ , and
- (b)  $\bar{c}^T = c^T - c_B^T B^{-1}A \geq 0$

# Simplex Algorithm

We assume all the basic feasible solutions are nondegenerate.

Under this assumption, the reduced costs being nonnegative are both necessary and sufficient for a BFS to be optimal.

If a reduced cost  $\bar{c}_j$  of a nonbasic variable  $x_j$  is negative, we can move toward the  $j$ th feasible direction  $d^j$  to further decrease the cost. It is desirable to move as far as possible:

$$\theta^* = \max\{\theta \geq 0 \mid x + \theta d \in P\}$$

The resulting cost change is  $\theta^* c^T d$ , which is the same as  $\theta^* \bar{c}_j$ .

Given that  $Ad = 0$ , we have  $A(x + \theta d) = Ax = b$  for all  $\theta$ , and the equality constraints will never be violated. Thus,  $x + \theta d$  can become infeasible only if one of its components becomes negative.

## Simplex Algorithm

Two cases can occur when we move along direction  $d$ :

(a) If  $d \geq 0$ , then  $x + \theta d \geq 0$  for all  $\theta \geq 0$ , the vector  $x + \theta d$  never becomes infeasible, and we let  $\theta^* = \infty$ .

(b) If  $d_i < 0$  for some  $i$ , the constraint  $x_i + \theta d_i \geq 0$  becomes  $\theta \leq -x_i/d_i$ . This constraint on  $\theta$  must be satisfied for every  $i$  with  $d_i < 0$ . Thus, the largest possible value of  $\theta$  is

$$\theta^* = \min_{\{i | d_i < 0\}} \left( -\frac{x_i}{d_i} \right).$$

Recall that if  $x_i$  is a nonbasic variable, then either  $x_i$  is the entering variable and  $d_i = 1$ , or else  $d_i = 0$ . In either case,  $d_i$  is nonnegative. Thus, we only need to consider the basic variables and we have the equivalent formula

$$\theta^* = \min_{\{i=1, \dots, m | d_{B(i)} < 0\}} \left( -\frac{x_{B(i)}}{d_{B(i)}} \right).$$

Note that  $\theta^* > 0$ , because  $x_{B(i)} > 0$  for all  $i$ , as a consequence of nondegeneracy.

## Simplex Algorithm

Once  $\theta^*$  is chosen, and assuming it is finite, we move to the new feasible solution  $y = x + \theta^* d$ . Since  $x_j = 0$  and  $d_j = 1$ , we have  $y_j = \theta^* > 0$ . Let  $\ell$  be a minimizing index of

$$-\frac{x_{B(\ell)}}{d_{B(\ell)}} = \min_{\{i=1,\dots,m \mid d_{B(i)} < 0\}} \left( -\frac{x_{B(i)}}{d_{B(i)}} \right) = \theta^*;$$

$$x_{B(\ell)} + \theta^* d_{B(\ell)} = 0.$$

We observe that the basic variable  $x_{B(\ell)}$  has become zero, whereas the nonbasic variable  $x_j$  has now become positive, which suggests that  $x_j$  should replace  $x_{B(\ell)}$  in the basis. The new basis matrix is

$$\bar{B}(i) = \begin{cases} B(i), & i \neq \ell \\ j, & i = \ell \end{cases}$$

In other words,  $\ell$  leaves the basis,  $j$  enters the basis.



## One simplex iteration (pivot) summary

1. In a typical iteration, we start with a basis consisting of the basic columns  $A_{B(1)}, \dots, A_{B(m)}$ , and an associated basic feasible solution  $x$ .
2. Compute the reduced costs  $\bar{c}_j = c_j - c_B^T B^{-1} A_j$  for all nonbasic indices  $j$ . If they are all nonnegative, the current basic feasible solution is optimal, and the algorithm terminates; else, choose some  $j$  for which  $\bar{c}_j < 0$ .
3. Compute  $u = -d_B = B^{-1} A_j$ . If no component of  $u$  is positive, we have  $\theta^* = \infty$ , the optimal cost is  $-\infty$ , and the algorithm terminates.
4. If some component of  $u$  is positive, let

$$\theta^* = \min_{\{i=1, \dots, m \mid u_i > 0\}} \frac{x_{B(i)}}{u_i}$$

5. Let  $\ell$  be such that  $\theta^* = x_{B(\ell)} / u_\ell$ . Form a new basis by replacing  $A_{B(\ell)}$  with  $A_j$ . If  $y$  is the new basic feasible solution, the values of the new basic variables are  $y_j = \theta^*$  and  $y_{B(i)} = x_{B(i)} - \theta^* u_i$ ,  $i \neq \ell$ .

## Compute matrix inverse faster

**Motivation:** Taking matrix inverse ( $B^{-1}$ ) at each iteration is expensive ( $O(m^3)$ ). We would like a more efficient implementation.

**Observation:** The new basis  $\bar{B}$  and the old basis  $B$  only differ at the  $\ell$ th column

$$B = [A_{B(1)} \cdots A_{B(m)}]$$

$$\bar{B} = [A_{B(1)} \cdots A_{B(\ell-1)} \quad A_j \quad A_{B(\ell+1)} \cdots A_{B(m)}]$$

**Idea:** Exploit the information of  $B^{-1}$  to compute  $\bar{B}^{-1}$

## Row operations

$$B^{-1}\overline{B} = \begin{bmatrix} \begin{array}{c|c|c|c|c|c|c} & & & & & & \\ e_1 & \cdots & e_{\ell-1} & u & e_{\ell+1} & \cdots & e_m \end{array} \\ \hline 1 & & u_1 & & & & \\ & \ddots & \vdots & & & & \\ & & u_\ell & & & & \\ & & \vdots & \ddots & & & \\ & & u_m & & 1 & & \end{array} \right]$$

where  $u = B^{-1}A_j$ .

Let us apply a sequence of elementary row operations that will change the above matrix to the identity matrix.

(a) For each  $i \neq \ell$ , we add the  $\ell$  th row times  $-u_i/u_\ell$  to the  $i$  th row. (Recall that  $u_\ell > 0$ .) This replaces  $u_i$  by zero.

(b) We divide the  $\ell$  th row by  $u_\ell$ . This replaces  $u_\ell$  by one.

Let these row operations be matrix  $Q$ .  $QB^{-1}\overline{B} = I$ , which yields

$QB^{-1} = \overline{B}^{-1}$ . This shows that if we apply the same sequence of row operations to the matrix  $B^{-1}$ , we obtain  $\overline{B}^{-1}$ .

# Full tableau implementation

$$B^{-1}[b \mid A]$$

with columns  $B^{-1}b$  and  $B^{-1}A_1, \dots, B^{-1}A_n$ .

- This matrix is called the **simplex tableau**.
- $B^{-1}b$  is called the zeroth column
- The column  $u = B^{-1}A_j$  corresponding to the variable that enters the basis is called the **pivot column**.
- If the  $\ell$  th basic variable exits the basis, the  $\ell$  th row of the tableau is called the **pivot row**.

At the end of each iteration, we need to update the tableau  $B^{-1}[b \mid A]$  and compute  $\bar{B}^{-1}[b \mid A]$ . This can be accomplished by left-multiplying the simplex tableau with a matrix  $Q$  satisfying  $QB^{-1} = \bar{B}^{-1}$

## Full tableau implementation

It is customary to add the objective and the reduced cost as a zeroth row.

$-c_B^T B^{-1} b$	$c^T - c_B^T B^{-1} A$
$B^{-1} b$	$B^{-1} A$

We've already shown how to update the rows of  $B^{-1}$  to pivot to  $\bar{B}^{-1}$ .

It can be shown that the rule for updating the zeroth row turns out to be identical to the rule used for the other rows of the tableau: add a multiple of the pivot row to the zeroth row to set the reduced cost of the entering variable to zero.

## An iteration of the full tableau implementation

1. A typical iteration starts with the tableau associated with a basis matrix  $B$  and the corresponding basic feasible solution  $x$ .
2. Examine the reduced costs in the zeroth row of the tableau. If they are all nonnegative, the current basic feasible solution is optimal, and the algorithm terminates; else, choose some  $j$  for which  $\bar{c}_j < 0$ .
3. Consider the vector  $u = B^{-1}A_j$ , which is the  $j$  th column (the pivot column) of the tableau. If no component of  $u$  is positive, the optimal cost is  $-\infty$ , and the algorithm terminates.
4. For each  $i$  for which  $u_i$  is positive, compute the ratio  $x_{B(i)}/u_i$ . Let  $\ell$  be the index of a row that corresponds to the smallest ratio. The column  $A_{B(\ell)}$  exits the basis and the column  $A_j$  enters the basis.
5. Add to each row of the tableau a constant multiple of the  $\ell$  th row (the pivot row) so that  $u_\ell$  (the pivot element) becomes one and all other entries of the pivot column become zero.

## Example

$$\begin{array}{ll}\text{minimize} & -10x_1 - 12x_2 - 12x_3 \\ \text{subject to} & x_1 + 2x_2 + 2x_3 \leq 20 \\ & 2x_1 + x_2 + 2x_3 \leq 20 \\ & 2x_1 + 2x_2 + x_3 \leq 20 \\ & x_1, x_2, x_3 \geq 0.\end{array}$$

Add slack variables to transform to the standard form.

$$\begin{array}{ll}\text{minimize} & -10x_1 - 12x_2 - 12x_3 \\ \text{subject to} & x_1 + 2x_2 + 2x_3 + x_4 = 20 \\ & 2x_1 + x_2 + 2x_3 + x_5 = 20 \\ & 2x_1 + 2x_2 + x_3 + x_6 = 20 \\ & x_1, \dots, x_6 \geq 0.\end{array}$$

Note that  $x = (0, 0, 0, 20, 20, 20)$  is a basic feasible solution and can be used to start the algorithm. Let accordingly,  $B(1) = 4$ ,  $B(2) = 5$ , and  $B(3) = 6$ .

## Example

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	0	-10	-12	-12	0	0	0
$x_4 =$	20	1	2	2	1	0	0
$x_5 =$	20	2*	1	2	0	1	0
$x_6 =$	20	2	2	1	0	0	1

$\bar{c}_1 < 0$ . choose  $x_1$  to enter the basis

$$\frac{x_{B(1)}}{u_1} = \frac{x_4}{u_1} = \frac{20}{1} = 20$$

$$\frac{x_{B(2)}}{u_2} = \frac{x_5}{u_2} = \frac{20}{2} = 10$$

$$\frac{x_{B(3)}}{u_3} = \frac{x_6}{u_3} = \frac{20}{2} = 10$$

The second row can be selected as the pivot row.

Row operations:

Add 5 times the second row to the zeroth row.

Add  $-\frac{1}{2}$  times the second row to the first row.

Add  $-1$  times the second row to the third row.

Divide the second row by 2.



## Example

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	100	0	-7	-2	0	5	0
$x_4 =$	10	0	1.5	1*	1	-0.5	0
$x_1 =$	10	1	0.5	1	0	0.5	0
$x_6 =$	0	0	1	-1	0	-1	1

choose  $x_3$  to enter the basis.

$$\frac{x_{B(1)}}{u_1} = \frac{x_4}{u_1} = \frac{10}{1} = 10$$

$$\frac{x_{B(2)}}{u_2} = \frac{x_1}{u_2} = \frac{10}{1} = 10$$

$u_3$  is negative.

Choose the first row to be the pivot row.

Add 2 times the first row to the zeroth row.

Add -1 times the first row to the second row.

Add 1 times the first row to the third row.

Divide the first row by 1.

## Example

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	120	0	-4	0	2	4	0
$x_3 =$	10	0	1.5	1	1	-0.5	0
$x_1 =$	0	1	-1	0	-1	1	0
$x_6 =$	10	0	2.5*	0	1	-1.5	1

$x_2$  has negative reduced cost and enters the basis.

$$\frac{x_{B(1)}}{u_1} = \frac{x_3}{u_1} = \frac{10}{1.5}$$

$u_2$  is negative.

$$\frac{x_{B(3)}}{u_3} = \frac{x_6}{u_3} = \frac{10}{2.5}$$

Choose the third row to be the pivot row.

Add  $\frac{4}{2.5}$  times the third row to the zeroth row.

Add  $-\frac{1.5}{2.5}$  times the third row to the first row.

Add  $\frac{1}{2.5}$  times the third row to the second row.

Divide the third row by 2.5.

## Example

		$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
	136	0	0	0	3.6	1.6	1.6
$x_3 =$	4	0	0	1	0.4	0.4	-0.6
$x_1 =$	4	1	0	0	-0.6	0.4	0.4
$x_2 =$	4	0	1	0	0.4	-0.6	0.4

Optimal solution found. Optimal objective is -136.

$$x^* = (4, 4, 4, 0, 0, 0)$$

## Reference

1. Chapter 3. Bertsimas, D., & Tsitsiklis, J. N. (1997).  
Introduction to linear optimization (Vol. 6, pp. 479-530).  
Belmont, MA: Athena scientific.