MATH11111: Fundamentals of Optimization **Fall 2022**

Lecture 15 Optimality Conditions

Lecturer: E. Alper Yıldırım *Week: 4*

15.1 Outline

- Feasible Directions
- Optimality Conditions
- Review Problems

15.2 Overview

In this lecture, we will start to discuss the building blocks of an algorithm, called the *simplex method*, for solving linear programming problems. Without loss of generality, we will assume that the linear programming problem is in standard form and that it satisfies the full row rank assumption.

15.2.1 Setup and Assumptions

Let (P) denote a linear programming problem in standard form.

$$\begin{array}{lll} (\mathrm{P}) & \min & c^T x \\ & \text{s.t.} & \\ & & \begin{array}{rcl} Ax & = & b, \\ x & \geq & \mathbf{0}, \end{array} \end{array}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$, and $x \in \mathbb{R}^n$.

Assumptions:

1. A has full row rank. (This assumption can always be ensured by preprocessing.)
2. We assume that the feasible region, denoted by $\mathcal{P}$, is nonempty. (We will revisit this assumption later on.)
3. We assume that we have computed a vertex $\hat{x}$ of $\mathcal{P}$. (We will revisit this assumption later on.)

15.3 Optimality Conditions

Remark 15.1. *By the Fundamental Theorem of Linear Programming, we know that (P) contains a vertex as an optimal solution if the optimal value is finite. Recall that $\hat{x}$ is a vertex by our assumption.*

Question 1. *Under what conditions would $\hat{x}$ be an optimal solution of (P)?*

- Note that $\hat{x}$ is an optimal solution of (P) if and only if $\mathcal{P}$ does not contain any other vector $\bar{x} \in \mathbb{R}^n$ such that $c^T\bar{x} < c^T\hat{x}$.
- Let $d = \bar{x} - \hat{x} \in \mathbb{R}^n$. Then, $c^Td = c^T(\bar{x} - \hat{x}) = c^T\bar{x} - c^T\hat{x}$.
- Therefore, $\hat{x}$ is an optimal solution of (P) if and only if $c^Td \geq 0$ for all $d \in \mathbb{R}^n$ such that $\hat{x} + d \in \mathcal{P}$.

Question 2. *How can we characterise such vectors $d \in \mathbb{R}^n$?*

15.3.1 Feasible Directions and Optimality Conditions

In this section, given an arbitrary $\bar{x} \in \mathcal{P}$ (not necessarily a vertex), we focus on the set of all directions $d \in \mathbb{R}^n$ such that $\bar{x} + d \in \mathcal{P}$.

If $\bar{x} + d \in \mathcal{P}$, then $\lambda(\bar{x} + d) + (1 - \lambda)\bar{x} = \bar{x} + \lambda d \in \mathcal{P}$ for each $\lambda \in [0, 1]$ since $\mathcal{P}$ is a convex set.

Definition 15.1. *Let $\bar{x} \in \mathcal{P}$. A vector $d \in \mathbb{R}^n$ is said to be a* feasible direction *at $\bar{x}$ if there exists $\lambda^* > 0$ such that $\bar{x} + \lambda^* d \in \mathcal{P}$.*

Remark 15.2.
1. *A vector $d \in \mathbb{R}^n$ is a feasible direction at $\bar{x}$ if, starting from $\bar{x}$, we can move in the direction of d at least for a while without leaving $\mathcal{P}$.*
2. *Note that $d = \mathbf{0} \in \mathbb{R}^n$ is a feasible direction at any feasible solution.*

The following proposition establishes the significance of feasible directions.

Proposition 15.1. *Let $\bar{x} \in \mathcal{P}$. Then, $\bar{x}$ is an optimal solution of (P) if and only if $c^Td \geq 0$ for all feasible directions $d \in \mathbb{R}^n$ at $\bar{x}$.*

Proof. $\Rightarrow$: Suppose that $\bar{x}$ is an optimal solution of (P). Suppose, for a contradiction, that there exists a feasible direction $\bar{d} \in \mathbb{R}^n$ such that $c^T\bar{d} < 0$. Then, there exists a real number $\lambda^* > 0$ such that $\bar{x} + \lambda^* d \in \mathcal{P}$. We obtain $c^T(\bar{x} + \lambda^* d) = c^T\bar{x} + \lambda^* c^T\bar{d} < c^T\bar{x}$, contradicting the optimality of $\bar{x}$.

$\Leftarrow$: Suppose that $c^Td \geq 0$ for all feasible directions $d \in \mathbb{R}^n$ at $\bar{x}$. Suppose, for a contradiction, that $\bar{x}$ is not optimal. Then, there exists $x^* \in \mathcal{P}$ such that $c^Tx^* < c^T\bar{x}$. Then, let $d = x^* - \bar{x}$. Clearly, d is a feasible direction at $\bar{x}$ and $c^Td < 0$, contradicting our hypothesis. $\square$

15.3.2 Feasible Directions and Optimality Conditions at a Vertex

In this section, given a **vertex** $\hat{x} \in \mathcal{P}$, we aim to shed light on the set of feasible directions at $\hat{x}$.

Question 3. *Proposition 15.1 holds for any feasible solution $\bar{x} \in \mathcal{P}$. How can we specialise it to a vertex $\hat{x} \in \mathcal{P}$?*

Recall that $\hat{x}$ is a vertex if and only if there exist disjoint index sets $B \subseteq \{1, \ldots, n\}$ and $N \subseteq \{1, \ldots, n\}$ such that $|B| = m$, $|N| = n - m$, $A_B \in \mathbb{R}^{m \times m}$ is invertible, $\hat{x}_B = (A_B)^{-1}b \geq \mathbf{0}$, and $\hat{x}_N = \mathbf{0} \in \mathbb{R}^{|N|}$.

Let $d \in \mathbb{R}^n$ be a feasible direction at $\hat{x}$. Then, $\hat{x} + \lambda^* d \in \mathcal{P}$ for some $\lambda^* > 0$.

- Therefore, $A(\hat{x} + \lambda^* d) = A\hat{x} + \lambda^* Ad = b$, i.e., $Ad = \mathbf{0}$.
- $\hat{x} + \lambda^* d \geq \mathbf{0}$, i.e., $\hat{x}_B + \lambda^* d_B \geq \mathbf{0}$ and $\hat{x}_N + \lambda^* d_N = \lambda^* d_N \geq \mathbf{0}$.

We therefore obtain the following result: If $d \in \mathbb{R}^n$ is a feasible direction at $\hat{x}$, then $Ad = A_B d_B + A_N d_N = \mathbf{0}$, $\hat{x}_B + \lambda^* d_B \geq \mathbf{0}$, and $d_N \geq \mathbf{0}$.

Conversely, pick an arbitrary vector $d_N \in \mathbb{R}^{n-m}$ such that $d_N \geq \mathbf{0}$. In order to extend d_N to $d \in \mathbb{R}^n$ such that $Ad = A_B d_B + A_N d_N = \mathbf{0}$, set $d_B = -(A_B)^{-1} A_N d_N$.

- Then, for any $\lambda \in \mathbb{R}$, $A(\hat{x} + \lambda d) = A\hat{x} + \lambda Ad = b + \mathbf{0} = b$.
- Note that $\hat{x}_N + \lambda d_N = \lambda d_N \geq \mathbf{0}$ for all $\lambda \geq 0$.
- Consider $\hat{x}_B + \lambda d_B$. Note that d_B may not necessarily be a nonnegative vector!
  (i) **Case 1:** If there exists a real number $\lambda^* > 0$ such that $\hat{x} + \lambda^* d \in \mathcal{P}$ (i.e., $\hat{x}_B + \lambda^* d_B \geq \mathbf{0}$), then d is a feasible direction at $\hat{x}$.
  (ii) **Case 2:** Otherwise, d is not a feasible direction at $\hat{x}$.

This observation motivates the following discussion. Given a vertex $\hat{x}$ of $\mathcal{P}$, let

$$\mathcal{D} = \{d \in \mathbb{R}^n : d_N \geq \mathbf{0}, \quad d_B = -(A_B)^{-1} A_N d_N\}.$$

Let $\hat{\mathcal{D}} \subseteq \mathbb{R}^n$ denote the set of all feasible directions at $\hat{x}$. We therefore obtain

$$\hat{\mathcal{D}} \subseteq \mathcal{D},$$

i.e., the set $\hat{\mathcal{D}}$ contains all feasible directions at a vertex $\hat{x}$ of $\mathcal{P}$. However, not every element of $\mathcal{D}$ is necessarily a feasible direction at $\hat{x}$ (i.e., $\mathcal{D}$ can be strictly larger than $\hat{\mathcal{D}}$). We will revisit this issue later on.

Corollary 15.2. *Let $\hat{x} \in \mathbb{R}^n$ be a vertex of $\mathcal{P}$. If $c^T d \geq 0$ for each $d \in \mathcal{D}$, then $\hat{x}$ is an optimal solution of (P).*

Proof. Since $\hat{\mathcal{D}} \subseteq \mathcal{D}$ and $c^T d \geq 0$ for each $d \in \mathcal{D}$, we obtain $c^T d \geq 0$ for all feasible directions $d \in \mathbb{R}^n$ at $\hat{x}$. Therefore, $\hat{x}$ is an optimal solution of (P) by Proposition 15.1. □

15.3.3 Optimality Conditions and Reduced Costs

Let $\hat{x} \in \mathbb{R}^n$ be a vertex of $\mathcal{P}$ and let $d \in \mathcal{D}$. Then, $d_N \geq \mathbf{0}$ and $d_B = -(A_B)^{-1} A_N d_N$. Therefore,

$$\begin{aligned} c^T d &= c_B^T d_B + c_N^T d_N \\ &= -c_B^T (A_B)^{-1} A_N d_N + c_N^T d_N \\ &= -\sum_{j \in N} c_B^T (A_B)^{-1} A^j d_j + \sum_{j \in N} c_j d_j \\ &= \sum_{j \in N} \underbrace{\left(c_j - c_B^T (A_B)^{-1} A^j\right)}_{\bar{c}_j} d_j. \end{aligned}$$

Definition 15.3. *For each index $j \in N$, the parameter $\bar{c}_j$ is called the* reduced cost *of the variable x_j.*

Using the reduced costs, we can arrive at a simplified set of optimality conditions at a vertex $\hat{x} \in \mathcal{P}$.

Corollary 15.4. *Let $\hat{x} \in \mathbb{R}^n$ be a vertex of $\mathcal{P}$. If $\bar{c}_j \geq 0$ for each $j \in N$, then $\hat{x}$ is an optimal solution of (P).*

Proof. Suppose that $\bar{c}_j \geq 0$ for each $j \in N$. Then, $c^T d = \sum_{j \in N} \bar{c}_j d_j \geq 0$ for each $d \in \mathcal{D}$ since $d_j \geq 0$ for each $j \in N$. The result follows from Corollary 15.2. □

Remark 15.3. *Therefore, by computing $n - m$ reduced costs and checking whether each one is nonnegative, we can verify the optimality of a vertex $\hat{x}$!*

Here are further questions that we will address in the following lectures:

- Corollary 15.4 states that the nonnegativity of all reduced costs at a vertex $\hat{x}$ is **sufficient** to ensure its optimality.
- Is this condition also **necessary** (i.e., is it true that all reduced costs should be nonnegative at an optimal vertex)?
- Related to the previous question, what if there is a negative reduced cost? What can we conclude?

Exercises

Question 15.1. *Let $\mathcal{P} = \{x \in \mathbb{R}^n : Ax = b, \quad x \geq \mathbf{0}\}$ be a nonempty polyhedron in standard form and let $\bar{x} \in \mathcal{P}$ be an arbitrary feasible solution. Show that the set of all feasible directions at $\bar{x}$ is a nonempty convex set.*

MATH11111: Fundamentals of Optimization	**Fall 2022**
Lecture 16	Degeneracy
Lecturer: E. Alper Yıldırım	*Week: 4*

16.1 Outline

- Necessity of Optimality Conditions
- Degeneracy in General Polyhedra
- Degeneracy in Standard Form Polyhedra
- Review Problems

16.2 Overview

In this lecture, we will discuss the concept of *degeneracy* and its effect on optimality conditions.

Let us recall our setup and assumptions. Let (P) denote a linear programming problem in standard form.

$$\begin{array}{llrcl} \text{(P)} & \min & c^T x & & \\ & \text{s.t.} & & & \\ & & Ax & = & b, \\ & & x & \geq & \mathbf{0}, \end{array}$$

where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$, and $x \in \mathbb{R}^n$.

Assumptions:

1. A has full row rank. (This assumption can always be ensured by preprocessing.)
2. We assume that the feasible region, denoted by $\mathcal{P}$, is nonempty. (We will revisit this assumption later on.)
3. We assume that we have computed a vertex $\hat{x}$ of $\mathcal{P}$. (We will revisit this assumption later on.)

16.2.1 Reduced Costs Revisited

Let $\mathcal{P}$ denote the feasible region of (P) and let $\hat{x}$ be a vertex of $\mathcal{P}$ with disjoint index sets B and N, where $B \subseteq \{1, \ldots, n\}$ and $N \subseteq \{1, \ldots, n\}$ such that $|B| = m$, $|N| = n - m$, $A_B \in \mathbb{R}^{m \times m}$ is invertible, $\hat{x}_B = (A_B)^{-1} b \geq \mathbf{0}$, and $\hat{x}_N = \mathbf{0} \in \mathbb{R}^{|N|}$.

- For each index $j \in N$, the reduced cost of the variable x_j, denoted by $\bar{c}_j$, is given by $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j$.

- If $\bar{c}_j \geq 0$ for each $j \in N$, then $\hat{x}$ is an optimal solution of (P).

Question 1. *If $\hat{x}$ is an optimal solution of (P), is it necessarily true that $\bar{c}_j \geq 0$ for each $j \in N$?*

In particular, we know from Lecture 15 that the nonnegativity of reduced costs at a vertex is sufficient to guarantee its optimality. We are interested in whether this condition is also necessary for optimality. If this is the case, then the nonnegativity of all reduced costs at a vertex would be equivalent to its optimality.

As illustrated by the following example, the answer turns out to be no. We will investigate the reasons. Later, we will see that the necessity of this condition holds under some additional assumptions on the vertex.

16.3 An Example

Consider the following linear programming problem with two variables:

$$\begin{array}{lrcrcl} \min & 3x_1 & - & 2x_2 & & \\ \text{s.t.} & & & & & \\ & -3x_1 & + & 3x_2 & \leq & 6 \\ & -x_1 & + & 2x_2 & \leq & 0 \\ & x_1 & , & x_2 & \geq & 0 \end{array}$$

Using the graphical method, you can verify that the unique optimal solution is given by $x^* = [0, 0]^T$ and the optimal value is $z^* = 0$.

Let us convert this problem into standard form by defining x_3 and x_4 as the nonnegative slack variables for the first and the second constraints, respectively:

$$A = \begin{bmatrix} -3 & 3 & 1 & 0 \\ -1 & 2 & 0 & 1 \end{bmatrix}, \quad b = \begin{bmatrix} 6 \\ 0 \end{bmatrix}, \quad c = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 0 \end{bmatrix}$$

Since we obtain an equivalent linear programming problem in standard form, the unique optimal solution is $x^* = [0, 0, 6, 0]^T$ and $z^* = 0$. For this optimal solution, we obtain

$$\begin{aligned} \hat{B} &= \{j \in \{1, \dots, 4\} : x_j^* \neq 0\} = \{3\} \\ \hat{N} &= \{j \in \{1, \dots, 4\} : x_j^* = 0\} = \{1, 2, 4\}. \end{aligned}$$

Note that $m = 2$ and $n = 4$ in this example since there are two equality constraints and four decision variables in standard form. It is easy to verify that A has full row rank as its rows are linearly independent. Therefore, we need to find index sets $B \subseteq \{1, \dots, 4\}$ and $N \subseteq \{1, \dots, 4\}$ such that (i) $\hat{B} \subseteq B$, (ii) $N \subseteq \hat{N}$, (iii) $B \cup N = \{1, \dots, 4\}$, (iv) $B \cap N = \emptyset$, (v) $|B| = m = 2$ (therefore, $|N| = n - m = 4 - 2$), and (vi) $A_B \in \mathbb{R}^{2 \times 2}$ is invertible.

Let us pick $B = \{3, 4\}$. Note that $\hat{B} \subseteq B$, $|B| = 2$, and $A_B \in \mathbb{R}^{2 \times 2}$ is invertible since A_B consisting of the third and fourth columns of A is simply the 2×2 identity matrix. Then, $N = \{1, 2\}$.

However, we obtain

$$\bar{c}_2 = c_2 - c_B^T (A_B)^{-1} A^2 = -2 < 0$$

even though x^* is optimal and $2 \in N$.

This example shows that the nonnegativity of reduced costs is not necessary for the optimality of a vertex.

In the previous example, if we had used $B = \{2, 3\}$ instead, we would have still had $\hat{B} \subseteq B$, A_B is invertible, and we would have obtained $N = \{1, 4\}$. The reduced costs would then be given by

$$\begin{aligned} \bar{c}_1 &= c_1 - c_B^T (A_B)^{-1} A^1 = 2 \geq 0, \\ \bar{c}_4 &= c_4 - c_B^T (A_B)^{-1} A^4 = 0 \geq 0. \end{aligned}$$

With these choices of index sets B and N, we would be able to verify the optimality of $\hat{x}$!

In this example, note that $|\hat{B}| = 1$, whereas we need to have two indices in the set B since $m = 2$. Hence, we have some flexibility in choosing B. If we had a basic solution with $|\hat{B}| = 2$, then we would have only one choice given by $B = \hat{B}$.

16.4 Degeneracy

In this section, we introduce *degeneracy*, which is the underlying reason for the behaviour of reduced costs in the example above.

16.4.1 Degeneracy in General Polyhedra

Let $\mathcal{P} = \{x \in \mathbb{R}^n : (a^i)^T x \geq b_i,\ i \in M_1; (a^i)^T x \leq b_i,\ i \in M_2; (a^i)^T x = b_i,\ i \in M_3\}$ be a polyhedron and let $\hat{x} \in \mathbb{R}^n$. Recall that $I(\hat{x}) = \left\{i \in M_1 \cup M_2 \cup M_3 : (a^i)^T \hat{x} = b_i\right\}$. A point $\hat{x} \in \mathbb{R}^n$ is a basic solution of $\mathcal{P}$ if and only if $M_3 \subseteq I(\hat{x})$ and the set $\{a^i : i \in I(\hat{x})\}$ contains n linearly independent vectors.

Definition 16.1. *A basic solution $\hat{x} \in \mathbb{R}^n$ of $\mathcal{P}$ is said to be* degenerate *if $|I(\hat{x})| > n$ (i.e., if more than n constraints are active at $\hat{x}$). Otherwise (i.e., if $|I(\hat{x})| = n$), it is said to be* nondegenerate.

16.4.2 Degeneracy in Standard Form Polyhedra

Let $\mathcal{P} = \{x \in \mathbb{R}^n : (a^i)^T x = b_i,\ i = 1, \ldots, m; x_j \geq 0,\ j = 1, \ldots, n\}$ be a polyhedron and let $\hat{x} \in \mathbb{R}^n$. Recall that $\hat{x}$ is a basic solution if and only if there exist disjoint index sets $B \subseteq \{1, \ldots, n\}$ and $N \subseteq \{1, \ldots, n\}$ such that $|B| = m$, $|N| = n - m$, $A_B \in \mathbb{R}^{m \times m}$ is invertible, $\hat{x}_B = (A_B)^{-1} b$ and $\hat{x}_N = \mathbf{0}$.

Recall that $\hat{B} = \{j \in \{1, \ldots, n\} : \hat{x}_j \neq 0\}$ and $\hat{N} = \{j \in \{1, \ldots, n\} : \hat{x}_j = 0\}$. Note that each of the m equality constraints is active at $\hat{x}$ and the inequality constraints $x_j \geq 0$ for each $j \in \hat{N}$ is active at $\hat{x}$.

Therefore,

$$|I(\hat{x})| = m + |\hat{N}| = m + n - |\hat{B}|.$$

Since $\hat{x}$ is a basic solution, we have $\hat{B} \subseteq B$, therefore, $|\hat{B}| \leq |B| = m$.

- **Case 1:** If $|\hat{B}| = |B| = m$ (i.e., $B = \hat{B}$), then $|I(\hat{x})| = m + n - |\hat{B}| = n$. Then, $\hat{x}$ is nondegenerate. Note that all basic variables are different from zero in this case.

- **Case 2:** If $|\hat{B}| < |B| = m$ (i.e., $\hat{B} \subset B$), then $|I(\hat{x})| = m + n - |\hat{B}| > n$. Then, $\hat{x}$ is degenerate. Note that at least one basic variable is equal to zero in this case.

Definition 16.2. *Let $\hat{x} \in \mathbb{R}^n$ be basic solution of $\mathcal{P}$ in standard form. If $|\hat{B}| < m$, then $\hat{x}$ is degenerate. Otherwise (i.e., if $|\hat{B}| = m$), then $\hat{x}$ is nondegenerate.*

Remarks

1. If $\hat{x} \in \mathbb{R}^n$ is a basic **feasible** solution of $\mathcal{P}$ in standard form with index sets B and N, then $\hat{x}_B \geq \mathbf{0}$ and $\hat{x}_N = \mathbf{0}$.
2. Therefore, $\hat{x}$ is a nondegenerate basic feasible solution if and only if $\hat{x}_B > \mathbf{0}$, i.e., each basic variable is strictly positive.
3. $\hat{x}$ is a degenerate basic feasible solution if and only if the value of at least one basic variable is equal to zero.

Exercises

Question 16.1. *Let $\mathcal{P} = \{x \in \mathbb{R}^n : Ax = b, \quad x \geq 0\}$ be polyhedron in standard form, where $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m, c \in \mathbb{R}^n$, and $x \in \mathbb{R}^n$, and let $\hat{x} \in \mathcal{P}$ be a degenerate vertex with k positive components, where $k < m$. How many different possible choices are there to define the index set B?*

MATH11111: Fundamentals of Optimization **Fall 2022**

Lecture 17 Optimality Conditions Under Nondegeneracy

Lecturer: E. Alper Yıldırım *Week: 5*

17.1 Outline

- Optimality Conditions Under Nondegeneracy
- Nondegeneracy vs Degeneracy
- Review Problems

17.2 Review and Setup

Let (P) denote a linear programming problem in standard form.

$$
\begin{array}{lll}
\text{(P)} & \min & c^T x \\
& \text{s.t.} & \\
& & \begin{array}{rcl} Ax & = & b, \\ x & \geq & \mathbf{0}, \end{array}
\end{array}
$$

where $A \in \mathbb{R}^{m \times n}$ has full row rank, $b \in \mathbb{R}^m, c \in \mathbb{R}^n$, and $x \in \mathbb{R}^n$.

- Let $\hat{x}$ be a vertex of $\mathcal{P}$ with corresponding index sets B and N.
- For each index $j \in N$, the reduced cost of the nonbasic variable x_j, denoted by $\bar{c}_j$, is given by $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j$.
- If $\bar{c}_j \geq 0$ for each $j \in N$, then $\hat{x}$ is an optimal solution of (P).
- However, we may have $\bar{c}_j < 0$ for some $j \in N$ if $\hat{x}$ is a degenerate optimal solution of (P).

In this lecture, we will consider the effect of nondegeneracy on the optimality conditions in linear programming.

17.3 Optimality Conditions Under Nondegeneracy

The next proposition illustrates that we can establish stronger optimality conditions under the nondegeneracy assumption.

Proposition 17.1. *Let (P) be a linear programming problem in standard form and let $\hat{x}$ be a nondegenerate vertex of $\mathcal{P}$ with corresponding index sets B and N. Then, $\hat{x}$ is an optimal solution of (P) if and only if $\bar{c}_j \geq 0$ for each $j \in N$.*

Proof. $\Leftarrow$: Follows from Corollary 15.4.

$\Rightarrow$: Let $\hat{x}$ be a nondegenerate vertex of $\mathcal{P}$ with corresponding index sets B and N such that $\hat{x}$ is optimal. Suppose, for a contradiction, that there exists $j^* \in N$ such that $\bar{c}_{j^*} < 0$. Let us construct a feasible direction $d \in \mathbb{R}^n$ at $\hat{x}$ as follows. We set $d_{j^*} = 1$ and $d_j = 0$ for each $j \in N \backslash \{j^*\}$. Note that $d_N \geq \mathbf{0}$. Let $d_B = -(A_B)^{-1} A_N d_N \in \mathbb{R}^m$ so that $Ad = A_B d_B + A_N d_N = \mathbf{0}$. We claim that d is a feasible direction at $\hat{x}$. We need to show that there exists some real number $\lambda^* > 0$ such that $\hat{x} + \lambda^* d \in \mathcal{P}$. Clearly, $A(\hat{x} + \lambda d) = A\hat{x} + \lambda A d = b + \mathbf{0} = b$ for any $\lambda \in \mathbb{R}$, and $\hat{x}_N + \lambda d_N = \lambda d_N \geq \mathbf{0}$ for any $\lambda \geq 0$. Consider $\hat{x}_B + \lambda d_B$.

Case 1: If d_B has at least one negative component, then define $\lambda^* = \min\limits_{j \in B : d_j < 0} \frac{-\hat{x}_j}{d_j}$. Since $\hat{x}$ is nondegenerate, we obtain $\hat{x}_j > 0$ for each $j \in B$. Therefore, $\lambda^* > 0$ and it is finite. Note that $\hat{x}_B + \lambda d_B \geq \mathbf{0}$ for each $\lambda \in [0, \lambda^*]$. Therefore, $\hat{x} + \lambda d \in \mathcal{P}$ for each $\lambda \in [0, \lambda^*]$. In addition, by setting $\lambda = \lambda^*$, we obtain $c^T(\hat{x} + \lambda^* d) = c^T \hat{x} + \lambda^* c^T d = c^T \hat{x} + \lambda^* (c_B^T d_B + c_N^T d_N) = c^T \hat{x} + \lambda^* (-c_B^T (A_B)^{-1} A_N d_N + c_N^T d_N) = c^T \hat{x} + \lambda^* \sum\limits_{j \in N} \bar{c}_j d_j = c^T \hat{x} + \lambda^* \bar{c}_{j^*} d_{j^*} = c^T \hat{x} + \lambda^* \bar{c}_{j^*} < c^T \hat{x}$, which contradicts the optimality of $\hat{x}$.

Case 2: If $d_B \geq \mathbf{0}$, then $\hat{x}_B + \lambda d_B \geq \mathbf{0}$ for any $\lambda \geq 0$. Therefore, $\hat{x} + \lambda d \in \mathcal{P}$ for each $\lambda \geq 0$. Furthermore, $c^T(\hat{x} + \lambda d) = c^T \hat{x} + \lambda \bar{c}_{j^*} \to -\infty$ as $\lambda \to \infty$, which implies that (P) is unbounded, which contradicts the optimality of $\hat{x}$. Therefore, such an index $j^* \in N$ cannot exist, i.e., we should have $\bar{c}_j \geq 0$ for each $j \in N$. $\square$

Remark 17.1. *This result implies that the nonnegativity of all reduced costs is not only sufficient but also necessary for optimality under the nondegeneracy assumption.*

17.4 Nondegeneracy vs Degeneracy

17.4.1 Nondegenerate Case

Let $\hat{x}$ be a nondegenerate vertex of $\mathcal{P}$ (i.e., $\hat{x}_j > 0$ for each $j \in B$). Let $\hat{\mathcal{D}}$ denote the set of feasible directions at $\hat{x}$ and let $\mathcal{D} = \{d \in \mathbb{R}^n : d_N \geq \mathbf{0}, \quad d_B = -(A_B)^{-1} A_N d_N\}$. We always have

$$\hat{\mathcal{D}} \subseteq \mathcal{D}.$$

.

Conversely, for any $d \in \mathcal{D}$, we have $d_N \geq \mathbf{0}$.

1. **Case 1:** If $d_B = -(A_B)^{-1} A_N d_N \geq \mathbf{0}$, then $\hat{x} + \lambda d \in \mathcal{P}$ for each $\lambda \geq 0$. Therefore, $d \in \hat{\mathcal{D}}$.
2. **Case 2:** If $d_B = -(A_B)^{-1} A_N d_N$ has at least one negative component, then recall that $\lambda^* = \min\limits_{j \in B : d_j < 0} \frac{-\hat{x}_j}{d_j} > 0$. We obtain that $\hat{x} + \lambda d \in \mathcal{P}$ for each $\lambda \in [0, \lambda^*]$. Therefore, $d \in \hat{\mathcal{D}}$.

Therefore, we arrive at the following important observation:

> If $\hat{x}$ is nondegenerate, then $\hat{\mathcal{D}} = \mathcal{D}$.

In particular, this observation is the underlying reason for the result stated in Proposition 17.1.

17.4.2 Degenerate Case

On the other hand, let $\hat{x}$ be a degenerate vertex of $\mathcal{P}$ (i.e., $\hat{x}_j = 0$ for some $j \in B$) and let $d \in \mathcal{D}$. Consider the case in which $d_B = -(A_B)^{-1} A_N d_N$ has at least one negative component and recall that $\lambda^* = \min_{j \in B: d_j < 0} \frac{-\hat{x}_j}{d_j}$.

Therefore, if $d_j < 0$ for an index $j \in B$ such that $\hat{x}_j = 0$, we may have $\lambda^* = 0$. We can no longer guarantee that d is a feasible direction at $\hat{x}$. Therefore, we may have $\hat{\mathcal{D}} \subset \mathcal{D}$, which implies that we may have have $c^T d \geq 0$ for all $d \in \hat{\mathcal{D}}$ (i.e., $\hat{x}$ is optimal), but $c^T \bar{d} < 0$ for a vector $\bar{d} \in \mathcal{D} \backslash \hat{\mathcal{D}}$ (in particular, see the example in Section 16.3 of Lecture 16). This is why the nonnegativity of reduced costs of nonbasic variables is not necessary for the optimality of a degenerate vertex.

We will revisit the issue of degeneracy later.

Exercises

Question 17.1. *Can you construct a polyhedron $\mathcal{P} \subset \mathbb{R}^2$ that has at least one nondegenerate vertex and at least one degenerate vertex?*