MATH11111: Fundamentals of Optimization **Fall 2022**

Lecture 25 Complementary Slackness Property

Lecturer: E. Alper Yıldırım *Week: 9*

25.1 Outline

- Complementary Slackness Property
- Review Problems

25.2 Quick Review

Consider the following pair of linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

- (P) is the primal problem and (D) is the dual problem.
- Denoting the optimal values of (P) and (D) by z^* and z_D^*, respectively, we have $z_D^* \leq z^*$.
- If both (P) and (D) have nonempty feasible regions, then $z_D^* = z^*$.
- In this lecture, we will establish the so-called complementary slackness property.

25.3 Complementary Slackness Property

Consider the following pair of primal-dual linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

- Recall that the matrix $A \in \mathbb{R}^{m \times n}$ can be expressed using its rows or columns:

$$A = \begin{bmatrix} (a^1)^T \\ \vdots \\ (a^m)^T \end{bmatrix}, \quad A = \begin{bmatrix} A^1 & \cdots & A^n \end{bmatrix},$$

where $a^i \in \mathbb{R}^n$ for each $i = 1, \ldots, m$, and $A^j \in \mathbb{R}^m$ for each $j = 1, \ldots, n$.

- A vector $x \in \mathbb{R}^n$ is feasible for (P) if and only if $b_i - (a^i)^T x = 0$ for each $i = 1, \dots, m$, and $x_j \geq 0$ for each $j = 1, \dots, n$.
- A vector $y \in \mathbb{R}^m$ is feasible for (D) if and only if $c_j - (A^j)^T y \geq 0$ for each $j = 1, \dots, n$.

Proposition 25.1 (Complementary Slackness)**.** *Let $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ be feasible solutions of (P) and (D), respectively. Then, $\bar{x}$ and $\bar{y}$ are optimal solutions of (P) and (D), respectively, if and only if*

$$\begin{aligned} \bar{x}_j \left(c_j - (A^j)^T \bar{y}\right) &= 0, \quad j = 1, \dots, n, \\ \bar{y}_i \left(b_i - (a^i)^T \bar{x}\right) &= 0, \quad i = 1, \dots, m. \end{aligned}$$

Proof. $\Rightarrow$: Let $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ be optimal solutions of (P) and (D), respectively. By the Strong Duality Theorem, we have $c^T \bar{x} = b^T \bar{y}$. Since $A\bar{x} = b$, we obtain

$$c^T \bar{x} = (A\bar{x})^T \bar{y} = \bar{x}^T A^T \bar{y},$$

which implies that $\bar{x}^T (c - A^T \bar{y}) = 0$, or equivalently,

$$\sum_{j=1}^{n} \bar{x}_j \left(c_j - (A^j)^T \bar{y}\right) = 0.$$

Since $\bar{x}_j \geq 0$ and $c_j - (A^j)^T \bar{y} \geq 0$ for each $j = 1, \dots, n$, it follows that

$$\bar{x}_j \left(c_j - (A^j)^T \bar{y}\right) = 0, \quad j = 1, \dots, n.$$

Since $A\bar{x} = b$, we obtain $(a^i)^T \bar{x} = b_i$, or equivalently, $b_i - (a^i)^T \bar{x} = 0$ for each $i = 1, \dots, m$. Therefore,

$$\bar{y}_i \left(b_i - (a^i)^T \bar{x}\right) = 0, \quad i = 1, \dots, m.$$

$\Leftarrow$: Let $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ be feasible solutions of (P) and (D), respectively. Suppose that

$$\bar{x}_j \left(c_j - (A^j)^T \bar{y}\right) = 0, \quad j = 1, \dots, n.$$

Then, adding up each of these expressions, we obtain

$$\sum_{j=1}^{n} \bar{x}_j \left(c_j - (A^j)^T \bar{y}\right) = 0.$$

Therefore, $c^T \bar{x} - \bar{x}^T A^T \bar{y} = 0$, i.e., $c^T \bar{x} = \bar{y}^T A \bar{x}$. By using $A\bar{x} = b$, we obtain $c^T \bar{x} = b^T \bar{y}$. By part (ii) of the Weak Duality Theorem, $\bar{x}$ and $\bar{y}$ are optimal solutions of (P) and (D), respectively. □

25.4 Discussion and Implications

Consider the following pair of primal-dual linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

Let $\bar{x} \in \mathbb{R}^n$ and $\bar{y} \in \mathbb{R}^m$ be feasible solutions of (P) and (D), respectively. Then, by the Complementary Slackness Property, $\bar{x}$ and $\bar{y}$ are optimal solutions of (P) and (D), respectively, if and only if

$$\begin{aligned} \bar{x}_j \left(c_j - (A^j)^T \bar{y}\right) &= 0, \quad j = 1, \ldots, n, \\ \bar{y}_i \left(b_i - (a^i)^T \bar{x}\right) &= 0, \quad i = 1, \ldots, m. \end{aligned}$$

- Complementary slackness property establishes useful relations between the value of the jth primal variable and the "slackness" of the jth dual constraint, as well as the value of the ith dual variable and the "slackness" of ith primal constraint.
- Note that the second set of conditions is satisfied by any feasible solution $\bar{x} \in \mathbb{R}^n$ of (P).
- Recall that we had motivated the vector $y \in \mathbb{R}^m$ as unit prices for the violation of primal equality constraints.
- By the complementary slackness property, if one uses the "correct" prices $\bar{y} \in \mathbb{R}^m$ for constraint violations, then the total violation cost is equal to zero at any primal optimal solution $\bar{x}$!
- Suppose that $\bar{x} \in \mathbb{R}^n$ is an optimal solution of (P).
- Let $\bar{B} = \{j \in \{1, \ldots, n\} : \bar{x}_j > 0\}$ and $\bar{N} = \{j \in \{1, \ldots, n\} : \bar{x}_j = 0\}$.
- For each $j \in \bar{B}$, any dual optimal solution $\bar{y} \in \mathbb{R}^m$ should satisfy $c_j - (A^j)^T \bar{y} = 0$.
- If $\bar{x}$ is a nondegenerate optimal vertex, then $|\bar{B}| = m$ and $A_{\bar{B}} \in \mathbb{R}^{m \times m}$ is invertible.
- Then, $c_j - (A^j)^T \bar{y} = 0$ for each $j \in \bar{B}$ if and only if $c_{\bar{B}} - (A_{\bar{B}})^T \bar{y} = \mathbf{0}$ if and only if $\bar{y} = \left((A_{\bar{B}})^T\right)^{-1} c_{\bar{B}} = \left((A_{\bar{B}})^{-1}\right)^T c_{\bar{B}}$ (see the proof of the strong duality theorem).
- If $\bar{y} \in \mathbb{R}^m$ is an optimal solution of (D), then, for each $j = 1, \ldots, n$ such that $c_j - (A^j)^T \bar{y} > 0$, any primal optimal solution should satisfy $\bar{x}_j = 0$.

25.5 Concluding Remarks

- Complementary slackness property allows us to construct a primal (dual) optimal solution by starting from a dual (primal) optimal solution.
- In the next lecture, we will discuss the economic interpretation of the dual variables.

Exercises

Question 25.1. *Consider the linear programming problem:*

$$\begin{array}{lllllllllll} (P) & \min & x_1 & + & x_2 & - & x_3 & & & & \\ & s.t. & & & & & & & & & \\ & & x_1 & + & 2x_2 & - & x_3 & + & x_4 & = & 2 \\ & & x_1 & - & x_2 & + & 2x_3 & + & x_4 & = & 1 \\ & & x_1 & , & x_2 & , & x_3 & , & x_4 & \geq & 0. \end{array}$$

Find an optimal solution of (P) without using the simplex method by relying on the information that $\bar{y} = [1/3, -1/3]^T$ *is an optimal solution of the dual problem (D).*

MATH11111: Fundamentals of Optimization	**Fall 2022**
Lecture 26	Economic Interpretation of the Dual Variables
Lecturer: E. Alper Yıldırım	*Week: 9*

26.1 Outline

- Optimal Dual Variables as Shadow Prices
- Review Problems

26.2 Quick Review and Setup

Consider the following pair of linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

- (P) is the primal problem and (D) is the dual problem.
- So far, we established the following properties between (P) and (D):
  - (i) Weak duality
  - (ii) Strong duality
  - (iii) Primal-dual symmetry
  - (iv) Complementary slackness property
- In this lecture, we will give an economic interpretation of the dual optimal solution.
- We assume that both (P) and (D) have nonempty feasible regions and that $A \in \mathbb{R}^{m \times n}$ has full row rank.
- We assume that $x^* \in \mathbb{R}^n$ is a *nondegenerate* optimal vertex of (P) with corresponding index sets B and N.

26.3 Dual Variables and Optimal Value

Recall the following primal-dual pair of linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

- **Question:** Suppose that we wish to replace the right-hand side parameter b_i by $b_i + \delta$ in (P), where $i = 1, \ldots, m$ and $\delta \in \mathbb{R}$ is sufficiently small. How does the optimal value of (P) change?

- Since $x^* \in \mathbb{R}^n$ is a *nondegenerate* optimal vertex of (P) with corresponding index sets B and N, then $x_B^* = (A_B)^{-1}b > \mathbf{0}$, $x_N^* = \mathbf{0}$, and $\bar{c}_j = c_j - c_B^T(A_B)^{-1}A^j \geq 0$, where $j \in N$ and $A^j \in \mathbb{R}^m$ denotes the jth column of A.

- Note that the reduced costs do not depend on the right-hand side vector b.

- The values of basic variables $x_B^* = (A_B)^{-1}b$ do depend on b.

- Since b is replaced by $b + \delta e^i$, the new values of basic variables are given by $x_B^*(\delta) = (A_B)^{-1}(b + \delta e^i) = (A_B)^{-1}b + \delta(A_B)^{-1}e^i \geq \mathbf{0}$ for sufficiently small $\delta \in \mathbb{R}$ since $x_B^* = x_B^*(0) = (A_B)^{-1}b > \mathbf{0}$. Note that $x_N^*(\delta) = x_N^* = \mathbf{0}$.

- Therefore, the new vertex corresponding to the index sets B and N remains feasible and optimal for the modified problem if $\delta \in \mathbb{R}$ is sufficiently small.

- Note that the feasible region of (D) does not depend on the right-hand side vector b.

- Recall that $y^* = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$ is an optimal solution of (D) and y^* does not depend on b.

- Since reduced costs do not depend on b and remain nonnegative for the modified primal problem, it follows that y^* remains an optimal solution of the dual of the modified problem.

- Denoting the optimal value of the modified problem by $z^*(\delta)$, we obtain

$$z^*(\delta) = c_B^T(A_B)^{-1}(b + \delta e^i) = (y^*)^T(b + \delta e^i) = z^* + \delta y_i^*,$$

  where $z^* = z^*(0)$ denotes the optimal value of the original problem (P) and (D).

- Therefore, if the right-hand side vector b is replaced by $b + \delta e^i$ in (P), where $i = 1, \ldots, m$ and $\delta \in \mathbb{R}$ is sufficiently small, then $z^*(\delta) = z^* + \delta y_i^*$.

- y_i^* gives the rate of change of the optimal value with respect to changes in the right-hand side parameter b_i.

- We therefore refer to y_i^* as the *shadow price* associated with the $i$the primal constraint (or *marginal cost per unit increase in the right-hand side parameter* b_i).

- The dual optimal solution $y^* \in \mathbb{R}^m$ provides information about how the optimal value will change as the right-hand side vector $b \in \mathbb{R}^m$ changes slightly.

- More generally, if b is replaced by $b + \delta b'$, where $\delta \in \mathbb{R}$ is sufficiently small and $b' \in \mathbb{R}^m$, then $z^*(\delta) = z^* + \delta(y^*)^T b'$.

26.4 Another Interpretation of Dual Variables

In this section, we give another useful interpretation of dual variables using a real-life example.

26.4.1 Manufacturer's Problem: Product Mix

- A manufacturing facility produces n different products using m different resources.
- Each unit of product j requires a_{ij} units of resource i, $i = 1, \ldots, m$; $j = 1, \ldots, n$.
- Each unit of product j produced yields c_j units of profit, $j = 1, \ldots, n$, and b_i units of resource i are available, $i = 1, \ldots, m$.
- We wish to determine the best mix of products so as to maximize the total profit.
- Assume that each product can be produced in any fractional amount and that demand for each product is unlimited.

Decision Variables:

x_j: number of units of product j to be produced, $j = 1, \ldots, n$.

Optimization Model:

$$\begin{array}{llrcll} \text{(MP)} & \max & \sum_{j=1}^{n} c_j x_j & & & \\ & \text{s.t.} & & & & \\ & & \sum_{j=1}^{n} a_{ij} x_j & \leq & b_i, & i = 1, \ldots, m, \\ & & x_j & \geq & 0, & j = 1, \ldots, n. \end{array}$$

26.4.2 Entrepreneur's Problem: Resource Pricing

- Suppose that an entrepreneur is interested in purchasing all available resources from the facility.
- **Question:** What are the "fair" unit prices for each of the m resources?
- By fair unit prices, we mean
  - (i) the entrepreneur should minimize her total purchasing cost;
  - (ii) the manufacturer should have an incentive to sell all of its resources at these prices.
- Let y_i denote the unit price of resource i, $i = 1, \ldots, m$.
- The total purchasing cost of the entrepreneur is given by $\sum_{i=1}^{m} b_i y_i$.
- By selling its resources, the manufacturer would give up c_j units of profit for each unit of product j not produced, $j = 1, \ldots, n$.
- Each unit of product j not produced would release a_{ij} units of resource i, $i = 1, \ldots, m$.
- Therefore, the unit prices y_i, $i = 1, \ldots, m$ should be such that

$$\sum_{i=1}^{m} a_{ij} y_i \geq c_j, \quad j = 1, \ldots, n.$$

Decision Variables:

y_i: unit price of resource $i,\ i = 1, \ldots, m$.

Optimization Model:

$$\begin{array}{llrcll} \text{(EP)} & \min & \sum_{i=1}^{m} b_i y_i & & & \\ & \text{s.t.} & & & & \\ & & \sum_{i=1}^{m} a_{ij} y_i & \geq & c_j, & j = 1, \ldots, n, \\ & & y_i & \geq & 0, & i = 1, \ldots, m. \end{array}$$

26.4.3 Discussion

- You can easily verify that the entrepreneur's problem (EP) is the dual of the manufacturer's problem (MP).
- By weak duality, the total purchase cost should be at least as large as the maximum total profit.
- By strong duality, the minimum total purchase cost equals the maximum total profit.
- Therefore, for any optimal solution $y^* \in \mathbb{R}^m$ of (EP), y_i^* is the fair unit price of resource $i,\ i = 1, \ldots, m$.
- y_i^* is the minimum unit price of resource $i,\ i = 1, \ldots, m$ the manufacturer should be willing to sell.
- Conversely, the manufacturer should be willing to pay at most y_i^* per each additional unit of resource $i,\ i = 1, \ldots, m$.
- Let $x^* \in \mathbb{R}^n$ denote an optimal solution of (MP). Recall that x_j^* denotes the optimal production quantity of product $j,\ j = 1, \ldots, n$.
- The manufacturer is indifferent between producing the optimal quantities $x^* \in \mathbb{R}^n$ or selling all of its m resources at fair prices $y^* \in \mathbb{R}^m$.
- If $x_j^* > 0$, then $c_j = \sum_{i=1}^{m} a_{ij} y_i^*$, since if $c_j < \sum_{i=1}^{m} a_{ij} y_i^*$, the manufacturer would prefer selling the resources to producing product $j,\ j = 1, \ldots, n$.
- Similarly, if $c_j < \sum_{i=1}^{m} a_{ij} y_i^*$, then $x_j^* = 0,\ j = 1, \ldots, n$.
- Therefore, $x_j^* \left(c_j - \sum_{i=1}^{m} a_{ij} y_i^* \right) = 0,\ j = 1, \ldots, n$.
- If $b_i > \sum_{j=1}^{n} a_{ij} x_j^*$, then we have $y_i^* = 0$ (i.e., there is no incentive for the manufacturer to buy additional units of resource $i,\ i = 1, \ldots, m$).
- If $y_i^* > 0$, then we have $b_i = \sum_{j=1}^{n} a_{ij} x_j^*,\ i = 1, \ldots, m$.
- Therefore, $y_i^* \left(b_i - \sum_{j=1}^{n} a_{ij} x_j^* \right) = 0,\ i = 1, \ldots, m$.
- Observe that these are precisely the complementary slackness conditions.

Exercises

Question 26.1. *Consider the following pair of primal-dual linear programming problems:*

$$\begin{array}{ll} (P) & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ (D) & \max\{b^T y : A^T y \leq c\} \end{array}$$

Suppose that $x^* \in \mathbb{R}^n$ is a degenerate *optimal vertex of (P) with corresponding index sets B and N. Suppose that we wish to replace the right-hand side parameter b_i by $b_i + \delta$ in (P), where $i = 1, \ldots, m$ and $\delta \in \mathbb{R}$ is sufficiently small. Does the same analysis on the change of the optimal value as in the case of a nondegenerate optimal vertex remain valid? Why or why not?*

MATH11111: Fundamentals of Optimization **Fall 2022**

Lecture 27 A Dual Perspective on the Simplex Method

Lecturer: E. Alper Yıldırım *Week: 9*

27.1 Outline

- Simplex Method and Candidate Dual Solutions
- Review Problems

27.2 Quick Review and Setup

Consider the following pair of linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

- (P) is the primal problem and (D) is the dual problem.
- Recall that (P) can be solved using the simplex method.
- In this lecture, we will revisit the simplex method from the perspective of the dual problem (D).

27.2.1 Review of the Simplex Method

- Consider solving (P) using the simplex method in dictionary form. We assume that $A \in \mathbb{R}^{m \times n}$ has full row rank.
- Suppose that $\hat{x} \in \mathbb{R}^n$ is a vertex of the primal problem (P) with index sets B and N.
- We have $\hat{x}_N = \mathbf{0}$ and $\hat{x}_B = (A_B)^{-1}b \geq \mathbf{0}$. Therefore, $\hat{z} = c^T\hat{x} = c_B^T\hat{x}_B + c_N^T\hat{x}_N = c_B^T(A_B)^{-1}b$.
- The corresponding dictionary is given by

$$\begin{aligned} z &= c_B^T(A_B)^{-1}b + \sum_{j \in N} \underbrace{\left(c_j - c_B^T(A_B)^{-1}A^j\right)}_{\bar{c}_j} x_j \\ x_B &= (A_B)^{-1}b + \sum_{j \in N} \left(-(A_B)^{-1}A^j\right) x_j \end{aligned}$$

27.3 A Dual Perspective on the Simplex Method

In this section, we will develop a dual perspective on the simplex method. We will analyse an optimal vertex as well as an intermediate vertex encountered during the solution process of the primal problem (P).

27.3.1 Optimal Vertex

Recall the simplex method in dictionary form:

$$\begin{aligned} z &= c_B^T(A_B)^{-1}b + \sum_{j\in N} \underbrace{(c_j - c_B^T(A_B)^{-1}A^j)}_{\bar{c}_j} x_j \\ x_B &= (A_B)^{-1}b + \sum_{j\in N} \left(-(A_B)^{-1}A^j\right) x_j \end{aligned}$$

- If $\bar{c}_j \geq 0$ for each $j \in N$, then $\hat{x}$ is an optimal solution of (P) and we stop.
- Let $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$. The dual constraints are given by $(A^j)^T y \leq c_j,\ j = 1,\ldots,n$, or $c_j - (A^j)^T y \geq 0,\ j = 1,\ldots,n$.
- At $\hat{y}$, we have
  (i) $c_j - (A^j)^T\hat{y} = c_j - (A^j)^T\left((A_B)^{-1}\right)^T c_B = c_j - c_B^T(A_B)^{-1}A^j = \bar{c}_j \geq 0,\ j \in N$.
  (ii) $c_j - (A^j)^T\hat{y} = c_j - (A^j)^T\left((A_B)^{-1}\right)^T c_B = c_j - c_B^T(A_B)^{-1}A^j = \bar{c}_j = 0,\ j \in B$.
- Therefore, if $\hat{x}$ is an optimal solution of (P), then $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$ is a feasible solution of (D).
- Since $A\hat{x} = b$, we have $(a^i)^T\hat{x} = b_i$ or $b_i - (a^i)^T\hat{x} = 0$ for each $i = 1,\ldots,m$. Therefore, $\hat{y}_i\left(b_i - (a^i)^T\hat{x}\right) = 0$ for each $i = 1,\ldots,m$.
- Similarly, we have $\hat{x}_j\left(c_j - (A^j)^T\hat{y}\right) = 0$ for each $j = 1,\ldots,n$ since (i) $\hat{x}_j \geq 0$ and $\bar{c}_j = c_j - (A^j)^T\hat{y} = 0$ for each $j \in B$; and (ii) $\hat{x}_j = 0$ and $\bar{c}_j = c_j - (A^j)^T\hat{y} \geq 0$ for each $j \in N$.
- Therefore, $\hat{x}$ is feasible for (P), $\hat{y}$ is feasible for (D), and they satisfy the complementary slackness conditions.
- By the Complementary Slackness Conditions, we conclude that $\hat{x}$ is an optimal solution of (P), and $\hat{y}$ is an optimal solution of (D).

27.3.2 Intermediate Vertex

Again, recall the simplex method in dictionary form:

$$\begin{aligned} z &= c_B^T(A_B)^{-1}b + \sum_{j\in N} \underbrace{(c_j - c_B^T(A_B)^{-1}A^j)}_{\bar{c}_j} x_j \\ x_B &= (A_B)^{-1}b + \sum_{j\in N} \left(-(A_B)^{-1}A^j\right) x_j \end{aligned}$$

- Suppose now that $\hat{x}$ is an arbitrary intermediate vertex of (P) visited at one of the iterations of the simplex method, with index sets B and N.
- There is at least one index $j^* \in N$ such that $\bar{c}_{j^*} = c_{j^*} - c_B^T (A_B)^{-1} A^{j^*} < 0$.
- We can still define $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$.
- Note that $\hat{y}$ is not a feasible solution of (D) since $\bar{c}_{j^*} = c_{j^*} - c_B^T (A_B)^{-1} A^{j^*} = c_{j^*} - (A^{j^*})^T \hat{y} < 0$.
- We have $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$ and $\hat{y}$ violates at least one of the constraints of (D).
- For each $j \in B$, we have $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j = c_j - (A^j)^T \hat{y} = 0$.
- Therefore, for each $j \in B$, the dual constraint $c_j - (A^j)^T y \geq 0$ is active at $\hat{y}$.
- Since $A_B \in \mathbb{R}^{m \times m}$ is invertible, we obtain $\text{span}\{A^j : j \in B\} = \mathbb{R}^m$.
- Since (D) has no equality constraints, it follows that $\hat{y}$ is a basic but infeasible solution of (D).
- We have $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$ and $\hat{y}$ is a basic but infeasible solution of (D).
- For each $j \in B$, we have $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j = c_j - (A^j)^T \hat{y} = 0$. Therefore, we obtain $\hat{x}_j \left(c_j - (A^j)^T \hat{y}\right) = 0$ for each $j \in B$.
- Since $\hat{x}_j = 0$ for each $j \in N$, we obtain $\hat{x}_j \left(c_j - (A^j)^T \hat{y}\right) = 0$ for each $j \in N$.
- For each $i = 1, \ldots, m$, since $b_i - (a^i)^T \hat{x} = 0$, we obtain $\hat{y}_i \left(b_i - (a^i)^T \hat{x}\right) = 0$.
- Therefore, $\hat{x} \in \mathbb{R}^n$ and $\hat{y} \in \mathbb{R}^m$ satisfy the complementarity conditions, but not all complementary slackness conditions since $\hat{y}$ is not a feasible solution of (D).

27.4 Discussion and Concluding Remarks

- At each dictionary, the simplex method computes a basic feasible solution $\hat{x} \in \mathbb{R}^n$ of (P).
- A "candidate" dual solution is implicitly constructed by defining $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$.
- $\hat{y}$ is always a basic solution of (D). Furthermore, $\hat{x} \in \mathbb{R}^n$ and $\hat{y} \in \mathbb{R}^m$ satisfy the complementarity part of the complementary slackness conditions.
- The simplex method terminates at an optimal vertex $\hat{x}$ if and only if $\hat{y}$ is a basic feasible solution of (D).
- Otherwise, $\hat{y}$ is a basic but infeasible solution of (D) and the simplex method does not terminate at this dictionary.
- The simplex method maintains primal feasibility throughout each iteration, moving from one vertex of (P) to the next.
- At each dictionary, a candidate dual basic solution is constructed that satisfies the complementarity part of the complementary slackness conditions together with the current primal vertex.
- Therefore, the simplex method maintains primal feasibility and complementarity part of the complementary slackness conditions at each dictionary, and works towards dual feasibility.
- In the next lecture, we will study an alternative variant that maintains dual feasibility and complementarity part of the complementary slackness conditions at each dictionary, and works towards primal feasibility.

Exercises

Question 27.1. *Consider the following pair of primal-dual linear programming problems:*

$$\begin{array}{ll} (P) & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ (D) & \max\{b^T y : A^T y \leq c\} \end{array}$$

Assume that $A \in \mathbb{R}^{m\times n}$ has full row rank. Let $\hat{x} \in \mathbb{R}^n$ be a vertex of (P) with corresponding index sets B and N. Let $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$ denote the corresponding candidate dual basic solution. Prove the following statement:

> *$\hat{y}$ is a degenerate basic solution of (D) if and only if there exists at least one index $j^* \in N$ such that $\bar{c}_{j^*} = c_{j^*} - c_B^T (A_B)^{-1} A^{j^*} = 0$.*

MATH11111: Fundamentals of Optimization	**Fall 2022**
Lecture 28	The Dual Simplex Method
Lecturer: E. Alper Yıldırım	*Week: 10*

28.1 Outline

- The Dual Simplex Method
- Review Problems

28.2 Quick Review and Setup

Consider the following pair of linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

- (P) is the primal problem and (D) is the dual problem.
- Recall that (P) can be solved using the simplex method.
- The simplex method for solving (P) maintains primal feasibility and the complementarity part of complementary slackness at each dictionary, and works towards dual feasibility.
- In this lecture, we will study an alternative variant, referred to as *the dual simplex method*, that maintains dual feasibility and the complementarity part of complementary slackness at each dictionary, and works towards primal feasibility.
- We assume that $A \in \mathbb{R}^{m \times n}$ has full row rank.
- Let $B \subseteq \{1, \ldots, n\}$ and $N \subseteq \{1, \ldots, n\}$ be two disjoint index sets such that $|B| = m$, $|N| = n - m$, and $A_B \in \mathbb{R}^{m \times m}$ is invertible.
- Let $\hat{x} \in \mathbb{R}^n$ be such that $\hat{x}_N = \mathbf{0}$ and $\hat{x}_B = (A_B)^{-1} b$. Note that $\hat{x}$ is a basic solution of (P).
- Suppose that $\hat{x}$ is infeasible for (P), i.e., there exists at least one $j \in B$ such that $\hat{x}_j < 0$.
- Suppose also that $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j \geq 0$ for each $j \in N$.

28.2.1 Corresponding Dictionary

Under the aforementioned assumptions, consider the corresponding dictionary given by

$$\begin{aligned} z &= c_B^T (A_B)^{-1} b + \sum_{j \in N} \underbrace{\left(c_j - c_B^T (A_B)^{-1} A^j\right)}_{\bar{c}_j} x_j \\ x_B &= (A_B)^{-1} b + \sum_{j \in N} \left(-(A_B)^{-1} A^j\right) x_j \end{aligned}$$

- Note that the current dictionary does not correspond to a primal vertex since $\hat{x}_B = (A_B)^{-1} b \not\geq \mathbf{0}$.
- However, all reduced costs are nonnegative since $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j \geq 0$ for each $j \in N$.
- Let $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$. Using the dual perspective, this dictionary is primal infeasible (i.e., $\hat{x}$ is a basic but infeasible solution of (P)), dual feasible (i.e., $\hat{y}$ is a basic feasible solution of (D)), and the complementarity part of complementary slackness conditions are satisfied by $\hat{x}$ and $\hat{y}$.
- **Question:** Is there a variant of the simplex method that maintains dual feasibility, the complementarity part of complementary slackness, and works towards primal feasibility?

28.3 Development of the Dual Simplex Method

Recall the the corresponding dictionary given by

$$\begin{aligned} z &= c_B^T (A_B)^{-1} b + \sum_{j \in N} \underbrace{\left(c_j - c_B^T (A_B)^{-1} A^j\right)}_{\bar{c}_j} x_j \\ x_B &= (A_B)^{-1} b + \sum_{j \in N} \left(-(A_B)^{-1} A^j\right) x_j \end{aligned}$$

- To achieve primal feasibility, we need to make sure that the values of all basic variables are nonnegative.
- To maintain dual feasibility, we need to make sure that all Row 0 coefficients remain nonnegative at each dictionary.
- Recall that the complementarity part of the complementary slackness comes for free at each dictionary by the definition $\hat{y} = \left((A_B)^{-1}\right)^T c_B \in \mathbb{R}^m$.

28.3.1 A Closer Look

Let us analyse the current dictionary in more detail:

$$\begin{aligned} z &= \hat{z} + \sum_{j \in N} \bar{c}_j x_j \\ x_i &= \hat{x}_i + \sum_{j \in N} \bar{a}_{ij} x_j, \quad i \in B. \end{aligned}$$

- Note that $\hat{z}$ denotes the objective function value and $\hat{x}_i,\ i \in B$ denotes the values of the basic variables at the current dictionary.
- We denote by $\bar{a}_{ij}$ the right-hand side coefficient of the nonbasic variable x_j corresponding to the row in which x_i is the basic variable.
- Suppose that $\hat{x}_p < 0$ for some $p \in B$.
- This suggests an incorrect choice of $p \in B$. We therefore would like to replace $p \in B$ with one of the indices $q \in N$.
- We have $x_p = \hat{x}_p + \sum_{j \in N} \bar{a}_{pj} x_j$ and $\hat{x}_p < 0$ for some $p \in B$. We wish to replace $p \in B$ with one of the indices $q \in N$:

$$x_q = -\frac{\hat{x}_p}{\bar{a}_{pq}} + \frac{1}{\bar{a}_{pq}} x_p - \sum_{j \in N \setminus \{q\}} \frac{\bar{a}_{pj}}{\bar{a}_{pq}} x_j$$

- Clearly, we need $\bar{a}_{pq} \neq 0$. Furthermore, if $\bar{a}_{pq} < 0$, then the value of x_q would be negative.
- Therefore, we need to pick $q \in N$ such that $\bar{a}_{pq} > 0$ since we would like to work towards primal feasibility.
- We have $\hat{x}_p < 0$ and need to pick $q \in N$ such that $\bar{a}_{pq} > 0$:

$$x_q = -\frac{\hat{x}_p}{\bar{a}_{pq}} + \frac{1}{\bar{a}_{pq}} x_p - \sum_{j \in N \setminus \{q\}} \frac{\bar{a}_{pj}}{\bar{a}_{pq}} x_j$$

- Substitute this expression for x_q in the right-hand side of the rows corresponding to the other basic variables $i \in B \setminus \{p\}$:

$$\begin{aligned} x_i &= \hat{x}_i + \bar{a}_{iq} x_q + \sum_{j \in N \setminus \{q\}} \bar{a}_{ij} x_j \\ &= \left(\hat{x}_i - \frac{\bar{a}_{iq} \hat{x}_p}{\bar{a}_{pq}}\right) + \frac{\bar{a}_{iq}}{\bar{a}_{pq}} x_p + \sum_{j \in N \setminus \{q\}} \left(\bar{a}_{ij} - \frac{\bar{a}_{iq} \bar{a}_{pj}}{\bar{a}_{pq}}\right) x_j. \end{aligned}$$

- We have $\hat{x}_p < 0$ and need to pick $q \in N$ such that $\bar{a}_{pq} > 0$.
- For each $i \in B \setminus \{p\}$, we obtain:

$$x_i = \left(\hat{x}_i - \frac{\bar{a}_{iq} \hat{x}_p}{\bar{a}_{pq}}\right) + \frac{\bar{a}_{iq}}{\bar{a}_{pq}} x_p + \sum_{j \in N \setminus \{q\}} \left(\bar{a}_{ij} - \frac{\bar{a}_{iq} \bar{a}_{pj}}{\bar{a}_{pq}}\right) x_j.$$

- Note that the value of x_i in the next dictionary may be larger, smaller, or remain the same depending on the sign of $\bar{a}_{iq}$.
- We have $\hat{x}_p < 0$ and need to pick $q \in N$ such that $\bar{a}_{pq} > 0$:

$$x_q = -\frac{\hat{x}_p}{\bar{a}_{pq}} + \frac{1}{\bar{a}_{pq}} x_p - \sum_{j \in N \setminus \{q\}} \frac{\bar{a}_{pj}}{\bar{a}_{pq}} x_j$$

- Now substitute this expression for x_q in the right-hand side of Row 0:

$$\begin{aligned} z &= \hat{z} + \bar{c}_q x_q + \sum_{j \in N \setminus \{q\}} \bar{c}_j x_j \\ &= \left(\hat{z} - \frac{\bar{c}_q \hat{x}_p}{\bar{a}_{pq}}\right) + \frac{\bar{c}_q}{\bar{a}_{pq}} x_p + \sum_{j \in N \setminus \{q\}} \left(\bar{c}_j - \frac{\bar{c}_q \bar{a}_{pj}}{\bar{a}_{pq}}\right) x_j. \end{aligned}$$

- We have $\hat{x}_p < 0$ and need to pick $q \in N$ such that $\bar{a}_{pq} > 0$. In Row 0, we obtain

$$z = \left(\hat{z} - \frac{\bar{c}_q \hat{x}_p}{\bar{a}_{pq}}\right) + \frac{\bar{c}_q}{\bar{a}_{pq}} x_p + \sum_{j \in N \setminus \{q\}} \left(\bar{c}_j - \frac{\bar{c}_q \bar{a}_{pj}}{\bar{a}_{pq}}\right) x_j$$

- In order to maintain dual feasibility, we need to ensure that the new Row 0 coefficients are all nonnegative.
- Since $\bar{c}_q \geq 0$ and $\bar{a}_{pq} > 0$, the Row 0 coefficient of the new nonbasic variable x_p is nonnegative.
- For each $j \in N \setminus \{q\}$, we need to ensure that

$$\bar{c}_j - \frac{\bar{c}_q \bar{a}_{pj}}{\bar{a}_{pq}} \geq 0 \Longleftrightarrow \bar{c}_j \bar{a}_{pq} \geq \bar{c}_q \bar{a}_{pj}, \quad j \in N \setminus \{q\}.$$

- Since $\bar{c}_j \geq 0$ and $\bar{a}_{pq} > 0$, we obtain $\bar{c}_j \bar{a}_{pq} \geq 0$.
- Since $\bar{c}_q \geq 0$, we only need to worry about $j \in N \setminus \{q\}$ such that $\bar{a}_{pj} > 0$.
- We have $\hat{x}_p < 0$ and need to pick $q \in N$ such that $\bar{a}_{pq} > 0$.
- We therefore need $\frac{\bar{c}_j}{\bar{a}_{pj}} \geq \frac{\bar{c}_q}{\bar{a}_{pq}}$ for each $j \in N \setminus \{q\}$ such that $\bar{a}_{pj} > 0$. Since $\bar{a}_{pq} > 0$, we obtain

$$\frac{\bar{c}_q}{\bar{a}_{pq}} = \min_{j \in N : \bar{a}_{pj} > 0} \frac{\bar{c}_j}{\bar{a}_{pj}}.$$

- Therefore, if we pick $q \in N$ accordingly, then we ensure that the next dictionary remains dual feasible.
- We have $\hat{x}_p < 0$ and need to pick $q \in N$ such that $\bar{a}_{pq} > 0$.
- In Row 0, we obtain

$$z = \left(\hat{z} - \frac{\bar{c}_q \hat{x}_p}{\bar{a}_{pq}}\right) + \frac{\bar{c}_q}{\bar{a}_{pq}} x_p + \sum_{j \in N \setminus \{q\}} \left(\bar{c}_j - \frac{\bar{c}_q \bar{a}_{pj}}{\bar{a}_{pq}}\right) x_j$$

- Since $\bar{c}_q \geq 0, \hat{x}_p < 0$ and $\bar{a}_{pq} > 0$, the new objective function value either remains the same or increases!
- This is expected since we would like to move to a better dual basic feasible solution and (D) is a **maximization** problem.

28.4 The Dual Simplex Method

Consider the following pair of primal-dual linear programming problems:

$$\begin{array}{ll} \text{(P)} & \min\{c^T x : Ax = b, \quad x \geq \mathbf{0}\} \\ \text{(D)} & \max\{b^T y : A^T y \leq c\} \end{array}$$

Initialisation

- We assume that $A \in \mathbb{R}^{m \times n}$ has full row rank.
- Let $B \subseteq \{1, \ldots, n\}$ and $N \subseteq \{1, \ldots, n\}$ be two disjoint index sets such that $|B| = m$, $|N| = n - m$, and $A_B \in \mathbb{R}^{m \times m}$ is invertible.
- Let $\hat{x} \in \mathbb{R}^n$ be such that $\hat{x}_N = \mathbf{0}$ and $\hat{x}_B = (A_B)^{-1} b$. Note that $\hat{x}$ is a basic solution of (P).
- Suppose that $\hat{x}$ is infeasible for (P), i.e., there exists at least one $j \in B$ such that $\hat{x}_j < 0$.
- Suppose also that $\bar{c}_j = c_j - c_B^T (A_B)^{-1} A^j \geq 0$ for each $j \in N$.

Algorithm

1. **Leaving Variable:** Choose $p \in B$ such that $\hat{x}_p < 0$.
2. **Entering Variable:** Choose $q \in N$ such that

$$\frac{\bar{c}_q}{\bar{a}_{pq}} = \min_{j \in N : \bar{a}_{pj} > 0} \frac{\bar{c}_j}{\bar{a}_{pj}}.$$

3. Move x_q to the left-hand side and x_p to the right-hand side in the row in which x_p is the basic variable. Substitute this expression for x_q in the other rows including Row 0.
   (a) If the new values of basic variables are all nonnegative, then stop. We have an optimal dictionary.
   (b) Otherwise, $B \leftarrow (B \backslash \{p\}) \cup \{q\}$ and $N \leftarrow (N \backslash \{q\}) \cup \{p\}$. Go to Step 1.

28.5 Discussion and Concluding Remarks

- The dual simplex method can be used to solve (P) and (D) starting from a basic but infeasible solution of (P) with a corresponding feasible candidate dual solution.
- Dual feasibility and the complementarity part of the complementary slackness are maintained at each dictionary, and progress is made towards primal feasibility.
- The objective function value remains the same or increases since we are making progress in the dual problem (D).
- We first determine the leaving variable by looking at the values of basic variables in the current dictionary.

- We then determine the entering variable by performing a minimum ratio test using the ratios of reduced costs and coefficients of nonbasic variables in the right-hand side of the row corresponding to the leaving basic variable.
- The dual simplex method can be very useful for reoptimization of the primal problem (P) after replacing $b \in \mathbb{R}^m$ by $b' \in \mathbb{R}^m$ (see our discussion on sensitivity analysis and reoptimization in the subsequent lectures).

Exercises

Question 28.1. *Solve the following linear programming problem using the dual simplex method:*

$$\begin{array}{lrcrcrcrcr}
\min & x_1 & + & 2x_2 & + & x_3 & & & & \\
s.t. & & & & & & & & & \\
 & 3x_1 & - & x_2 & - & x_3 & & & \leq & -3 \\
 & x_1 & & & & & - & 4x_4 & \leq & -2 \\
 & -3x_1 & + & 2x_2 & + & x_3 & + & 2x_4 & \leq & 6 \\
 & x_1 & , & x_2 & , & x_3 & , & x_4 & \geq & 0
\end{array}$$