Abstract. A dual to a problem of optimally partitioning a given set of points in k-dimensional Euclidean space is presented. The dual possesses many properties similar to the dual in linear programming.

1. Introduction. A dual problem has been obtained for a functional version of the Neyman-Pearson problem by Francis and Wright [2], who also give an excellent review of the literature on the subject. The dual of a generalization of a version of the Neyman-Pearson problem, a problem of optimally partitioning a given set of points in k-dimensional Euclidean space, is presented in this paper. The results here are also related to work by McIlroy [4].

2. Formulation of the problem. Let $S$ be a bounded, Lebesgue-measurable set in $E_k$, k-dimensional Euclidean space, such that $a(S) = b$, and let $f_1, \ldots, f_n$ be real-valued functions integrable over $S$. Also, suppose that $b_1, \ldots, b_n$ are positive real numbers for which $\sum_{i=1}^{n} b_i = b$. For any Lebesgue-measurable subset $R$ of $E_k$, let $a(R) = \int_{R} 1$, where all integrals are Lebesgue integrals. As a further definition, if $R_1, \ldots, R_n$ are Lebesgue-measurable subsets of $S$ such that $a(R_i \cap R_j) = 0$ for $i \neq j$, then $R_1, \ldots, R_n$ will be said to be mutually nonoverlapping.

A partition of $S$ is now defined. If the sets $R_1, \ldots, R_n$ are mutually nonoverlapping, Lebesgue-measurable subsets of $S$ such that $\bigcup_{i=1}^{n} R_i = S$ and $a(R_i) = b_i$, $i = 1, \ldots, n$, then the n-tuple $(R_1, \ldots, R_n)$ will be called a feasible partition of $S$. Denote the class of all feasible partitions of $S$ by $P_n(S)$ and define

$$F_n(R_1, \ldots, R_n) = \sum_{i=1}^{n} \int_{R_i} f_i.$$  

With the above terminology, the following problem is called the primal problem:

(P1)  
$$\text{minimize } F_n(R_1, \ldots, R_n).$$

(R_1, \ldots, R_n) \in P_n(S)

Next let $g$ be an integrable, real-valued function defined on $S$, and denote $(u_1, \ldots, u_n) \in E_n$ by $U$. The notation $(U, g)$ will be used for such a point $U$ and function $g$. With $f_1, \ldots, f_n$ and $b_1, \ldots, b_n$ as in the primal, define

(1)  
$$G_n(U, g) = \sum_{i=1}^{n} u_i b_i + \int_{S} g$$

and

(2)  
$$W_n(S) = \{(U, g) | u_i + g(X) \leq f_i(X) \text{ for all } X \in S, i = 1, \ldots, n\}.$$
If \((U, g) \in W_n(S)\), then \((U, g)\) is said to be feasible to the dual. The dual problem may be stated as follows:

\[
(P2) \quad \maximize_{(U, g) \in W_n(S)} G_n(U, g).
\]

3. Remarks. The subsequent duality relationships between \((P1)\) and \((P2)\) are a consequence of the fact that \((P1)\) can be regarded as an “infinite transportation problem.” \((P2)\) is thus the dual of an “infinite transportation problem.” The function \(g\) in \((P2)\) corresponds to an infinite number of multipliers. To show how \((P2)\) relates to the dual of a transportation problem, let the integral \(\int_S g\) in (1) be replaced by the sum \(\sum_{j=1}^r v_j\). Also, in (2) replace the constraints \(u_i + g(X) \leq f_i(X)\) for all \(X \in S, i = 1, \cdots, n\), by \(u_i + v_j \leq c_{ij}, i = 1, \cdots, n, j = 1, \cdots, r\). With these changes, the dual problem \((P2)\) becomes

\[
\maximize \sum_{i=1}^n u_i b_i + \sum_{j=1}^r v_j 
\]

subject to

\[ u_i + v_j \leq c_{ij}, \quad i = 1, \cdots, n, \quad j = 1, \cdots, r, \]

which is the linear programming dual of a transportation problem [3].

The primal problem \((P1)\) occurs as a model in [1] for some districting, facility design, and warehouse layout problems. While efficient methods exist for solving \((P1)\), \((P2)\) might possibly lead to alternate solution procedures. Moreover, since \((P1)\) represents an extension of a version of the Neyman–Pearson lemma [5], \((P2)\) may have statistical implications.

4. Duality relationships. In developing the duality relationships between \((P1)\) and \((P2)\), two results proved in [1] are used. The two results are stated without proof below as Lemmas 1 and 2.

**Lemma 1.** There exists a minimal partition to \((P1)\).

**Lemma 2.** Suppose that the functions \(f_1, \cdots, f_n\) are continuous almost everywhere on \(S\). If \((R_1, \cdots, R_n)\) is a minimal partition to \((P1)\) such that the set of boundary points of \(R_i, i = 1, \cdots, n\), has Lebesgue measure zero, then there exist real numbers \(u_1, \cdots, u_n\) for which

\[ f_j(X) - f_i(X) \leq u_i - u_j \quad \text{a.e. for } X \in R_i, \quad i, j = 1, \cdots, n. \]

Lemma 2 represents a necessary condition for a minimal partition \((R_1, \cdots, R_n)\) under the two assumptions that the functions \(f_1, \cdots, f_n\) are continuous almost everywhere on \(S\) and that the set of boundary points of each set \(R_i, i = 1, \cdots, n\), has Lebesgue measure zero. These assumptions are also required in Theorems 2 and 3 below, which depend upon Lemma 2. An alternative to the two assumptions would be the use of Jordon measure and the Riemann integral. The Jordan measure of the set of boundary points of a Jordan-measurable set is necessarily zero, and a function Riemann integrable over a Jordan-measurable set \(S\) is necessarily continuous almost everywhere on \(S\).

The duality relationships between \((P1)\) and \((P2)\) are now established.
THEOREM 1. If \((R_1, \cdots, R_n)\) is a feasible partition to the primal and \((U, g)\) is feasible to the dual, then \(G_n(U, g) \leq F_n(R_1, \cdots, R_n)\).

Proof. Since \((U, g)\) is feasible to the dual,

\[
u_i + g(X) \leq f_i(X) \quad \text{for all } X \in S, \quad i = 1, \cdots, n.
\]

Upon integrating (3) over \(R_i, i = 1, \cdots, n\), respectively, and adding,

\[
\sum_{i=1}^{n} u_i a(R_i) + \sum_{i=1}^{n} \int_{R_i} g \leq \sum_{i=1}^{n} \int_{R_i} f_i.
\]

But \((R_1, \cdots, R_n)\) is feasible to the primal, so

\[
a(R_i) = b_i, \quad i = 1, \cdots, n.
\]

For \(i = 1, \cdots, n\), multiply by \(u_i\) the corresponding equation in (5) and sum to obtain

\[
\sum_{i=1}^{n} u_i a(R_i) = \sum_{i=1}^{n} u_i b_i.
\]

Next using the fact that \(R_1, \cdots, R_n\) are mutually nonoverlapping and substituting (6) into (4) give

\[
\sum_{i=1}^{n} u_i b_i + \int_S g \leq \sum_{i=1}^{n} \int_{R_i} f_i.
\]

Thus, \(G_n(U, g) \leq F_n(R_1, \cdots, R_n)\), which proves the theorem.

From Theorem 1 it follows that if \((R_1, \cdots, R_n)\) is a feasible partition to the primal and \((U, g)\) is feasible to the dual, then the equality of \(F_n(R_1, \cdots, R_n)\) and \(G_n(U, g)\) establishes the optimality of both \((R_1, \cdots, R_n)\) and \((U, g)\). Under the assumptions of Lemma 2, Theorem 2 utilizes this fact to prove the existence of an optimal solution to the dual given a minimal partition to the primal. It also shows under these conditions that \(F_n(R_1, \cdots, R_n)\) and \(G_n(U, g)\) are always equal for optimal \((R_1, \cdots, R_n)\) and \((U, g)\).

THEOREM 2. Suppose that the functions \(f_1, \cdots, f_n\) are continuous almost everywhere on \(S\). If \((R_1, \cdots, R_n)\) is a minimal partition to the primal such that the set of boundary points of \(R_i, i = 1, \cdots, n\), has Lebesgue measure zero, then there exists an optimal solution \((U, g)\) to the dual. Furthermore, \(G_n(U, g) = F_n(R_1, \cdots, R_n)\).

Proof. From Lemma 2 there exist numbers \(u_1, \cdots, u_n\) for which

\[
f_i(X) - f_j(X) \leq u_i - u_j \quad \text{a.e. for } X \in R_i, \quad i, j = 1, \cdots, n.
\]

Let \(U = (u_1, \cdots, u_n)\). For \(i = 1, \cdots, n\) and \(X \in R_i\), let

\[
g(X) = f_i(X) - u_i \quad \text{if } f_i(X) - f_j(X) \leq u_i - u_j, \quad j = 1, \cdots, n,
\]

and

\[
g(X) = \min_{j=1,\cdots,n} \{f_j(X) - u_j\} \quad \text{otherwise}.
\]

It will be shown that \((U, g)\) so defined is feasible to the dual. Let \(X \in S\). Then there exists some \(k \in \{1, \cdots, n\}\) such that \(X \in R_k\). There are two cases to consider.
First suppose that  
\[ f_k(X) - f_i(X) \leq u_k - u_i, \quad i = 1, \ldots, n. \]  
From the definition of \( g \),  
\[ u_k + g(X) = f_k(X), \]  
so from (8) and (9),  
\[ u_i + g(X) \leq f_i(X), \quad i = 1, \ldots, n. \]  
Thus, in this case, \(( U, g)\) is feasible at the point \( X \). In the second case suppose that (8) does not hold. Again from the definition of \( g \), (10) holds. Hence \(( U, g)\) is feasible to the dual.

It remains to show that \(( U, g)\) is optimal. From (7) and the definition of \( g \),  
\[ g(X) = f_i(X) - u_i \text{ a.e. for } X \in R_i, \quad i = 1, \ldots, n. \]  
Therefore, upon integrating,  
\[ \int_{R_i} g = \int_{R_i} f_i - u_i a(R_i), \quad i = 1, \ldots, n. \]  
Summing equation (11) over \( i = 1, \ldots, n \) and substituting \( a(R_i) = b_i \) give  
\[ \int_S g = \sum_{i=1}^n \int_{R_i} f_i - \sum_{i=1}^n u_i b_i, \]  
or  
\[ G_n(U, g) = F_n(R_1, \ldots, R_n). \]  
It then follows from (12) and Theorem 1 that \(( U, g)\) is an optimal solution to the dual, and the proof is complete.

The next property states a similar result when an optimal solution to the dual is given. The question of existence of an optimal solution to the primal is not considered in Theorem 3, however, since existence was established in Lemma 1.

**Theorem 3.** Suppose that \(( U, g)\) is an optimal solution to the dual and the functions \( f_1, \ldots, f_n \) are continuous almost everywhere on \( S \). If \(( R_1, \ldots, R_n)\) is a minimal partition to the primal such that the set of boundary points of \( R_i, i = 1, \ldots, n \), has Lebesgue measure zero, then  
\[ g(X) + u_i = f_i(X) \text{ a.e. for } X \in R_i, \quad i = 1, \ldots, n. \]  
Furthermore \( G_n(U, g) = F_n(R_1, \ldots, R_n) \).

**Proof.** It is first established that \( G_n(U, g) = F_n(R_1, \ldots, R_n) \). Since \(( R_1, \ldots, R_n)\) is a minimal partition to the primal and the hypotheses of Theorem 2 are satisfied, there exists an optimal solution \(( U', g')\) to the dual and \( F_n(R_1, \ldots, R_n) = G_n(U', g') \). But \(( U, g)\) is also optimal to the dual by hypothesis, so  
\[ G_n(U, g) = F_n(R_1, \ldots, R_n). \]  
To prove the first part of the theorem, assume (13) is false. Since \(( U, g)\) is feasible to the dual,  
\[ g(X) + u_i \leq f_i(X) \text{ for all } X \in R_i, \quad i = 1, \ldots, n. \]
Then from (15) and the assumption that (13) does not hold, there exists a nonempty index set \( I \subset \{1, \ldots, n\} \) defined as follows. For \( i = 1, \ldots, n \), let \( T_i \) be the set in \( R_i \) for which

\[
g(X) < f_i(X) - u_i.
\]

If \( a(T_i) > 0 \), let \( i \in I \). \( G_n(U, g) \) is next calculated under the assumption that (13) is false. Integrate \( g(X) + u_i \) over \( R_i \), \( i = 1, \ldots, n \), and sum to obtain

\[
G_n(U, g) = \sum_{i \in I} \left[ \int_{R_i} g + \int_{T_i} g \right] + \sum_{i \notin I} \int_{R_i} g + \sum_{i = 1}^{n} u_i b_i
\]

\[
= \sum_{i \in I} \left[ \int_{R_i - T_i} g + u_i \int_{R_i - T_i} 1 + \int_{T_i} g + u_i \int_{T_i} 1 \right] + \sum_{i \notin I} \left[ \int_{R_i} g + u_i \int_{R_i} 1 \right]
\]

\[
= \sum_{i \notin I} \int_{R_i} f_i + \sum_{i \in I} \int_{R_i - T_i} f_i + \sum_{i \in I} \left[ \int_{T_i} g + u_i \int_{T_i} 1 \right]
\]

\[
< \sum_{i = 1}^{n} \int_{R_i} f_i = F_n(R_1, \ldots, R_n).
\]

Hence, (14) is contradicted, and the theorem follows.

Theorem 3 is actually the statement of a complementary slackness condition associated with the primal and dual problems. Let \((U, g)\) be an optimal solution to the dual and \((R_1, \ldots, R_n)\) a minimal partition to the primal. If the hypotheses of Theorem 3 are satisfied, then \( g(X) = u_i = f_i(X) \) a.e. for \( X \in R_i \), \( i = 1, \ldots, n \). Thus if \( T_i = \{X | X \in R_i, g(X) + u_i < f_i(X)\} \) for \( i \in \{1, \ldots, n\} \), then \( a(T_i) = 0 \); that is, for an optimal dual solution \((U, g)\) and minimal partition \((R_1, \ldots, R_n)\), the set of points in \( R_i \) at which the \( i \)th dual constraint holds as a strict inequality has Lebesgue measure zero.

REFERENCES