The Annals of Mathematical Statistics
1971, Vol. 42, No. 5, 1613–1621

ELFVING'S THEOREM AND OPTIMAL DESIGNS FOR QUADRATIC LOSS[1]

BY W. J. STUDDEN
Purdue University

1. Introduction. The purpose of this paper is to give a matrix analog of a geometric result of Elfving in the theory of optimal design of experiments. The connection with quadratic loss is indicated below.

Let $f = (f_1, \cdots, f_m)$ denote m linearly independent continuous functions on a compact set X and let $\theta = (\theta_1, \cdots, \theta_m)$ denote a vector of parameters. For each $x \in X$ an experiment can be performed. The outcome is a random variable $y(x)$ with mean value $\theta f'(x) = \sum_i \theta_i f_i(x)$ and a variance σ^2 independent of x. (Primes will denote transposes.) The functions $f_1, \cdots, f_m$, called the regression functions, are assumed known while $\theta = (\theta_1, \cdots, \theta_m)$ and σ^2 are unknown. An experimental design is a probability measure μ defined on a fixed σ-field of sets of X which include the one point sets. In practice, the experimenter is allowed N uncorrelated observations and the number of observations that he takes at each $x \in X$ is "proportional" to the measure μ. For a given μ let

$$m_{ij} = m_{ij}(\mu) = \int f_i f_j d\mu \text{ and } M(\mu) = \|m_{ij}\|_{i,j=1}^m.$$

The matrix $M(\mu)$ is called the information matrix of the design.

Suppose μ concentrates mass μ_i at the points x_i, $i = 1, \cdots, r$ and $N\mu_i = n_i$ are integers. If N uncorrelated observations are made, taking n_i observations at x_i, then the variance of the best linear unbiased estimate of $a\theta' = \sum_i a_i\theta_i$ is given by $\sigma^2 N^{-1} a M^{-1}(\mu) a'$. The inverse $M^{-1}(\mu)$ must be suitably defined if $M(\mu)$ is singular. A design μ is called a-optimal if μ minimizes $V(a, \mu) = aM^{-1}(\mu)a'$. The following geometric result was given by Elfving (1952); see also Karlin and Studden (1966).

THEOREM (Elfving). *Let R denote the smallest convex set in Euclidean m-space which is symmetric with respect to the origin and contains all of the vectors $f(x) = (f_1(x), \cdots, f_m(x))$, $x \in X$. A design μ_0 is a-optimal if and only if there exists a scalar valued function $\phi(x)$ satisfying $|\phi(x)| \equiv 1$ such that* (i) $\int \phi(x) f(x) d\mu_0(x) = \beta a$ *for some β and* (ii) *βa is a boundary point of R. Moreover βa lies on the boundary of R if and only if $\beta^2 = v^{-1}$ where $v = \min_\mu V(a, \mu)$.*

The quantity, analogous to $V(a, \mu)$, that we wish to consider is

$$V(A, \mu) = \operatorname{tr} A'M^{-1}(\mu)A = \operatorname{tr} M^{-1}(\mu)AA' \tag{1.1}$$

where A is an $m \times k$ matrix and tr denotes the trace. We thus wish to minimize the sum of quantities $V(a, \mu)$ where the a's are given by the columns of A.

Received July 13, 1970.

[1] This research was supported by the National Science Foundation Grants GP 8986 and GP 20306. Reproduction is permitted in whole or in part for any purposes of the United States Government.