OPTIMAL DESIGNS OF INDIVIDUAL REGRESSION COEFFICIENTS WITH A TCHEBYCHEFFIAN SPLINE REGRESSION FUNCTION

by

V. N. MURTY

Department of Statistics
Division of Mathematical Sciences

Mimeograph Series No. 191
June 1969

*This research was supported by the National Science Foundation Contract GP-8986.

CHAPTER I

INTRODUCTION

1. Formulation of the Optimal Design Problem

The study of optimal design of experiments involving statistical data reduces to the following setup. Let $f_o, f_1, \ldots, f_n$ denote continuous functions defined on a compact space X. The points of X are referred to as the possible levels of feasible experiments. For each level $x \in X$, some experiment may be performed, whose outcome is a random variable $y(x)$. It is assumed that the observation $y(x)$ is of the form

$$y(x) = \sum_{j=0}^{n} \theta_j f_j(x) + \eta(x) \tag{1.1}$$

where $\eta(x)$ is a random variable such that

$$E(\eta(x)) = 0$$

$$E(\eta(x) \cdot \eta(x')) = \begin{cases} 1 & x = x' \\ 0 & x \neq x' \end{cases}$$

and E denotes the expected value of the indicated random variable. The functions $f_o, f_1, \ldots, f_n$, called the regression functions, are known to the experimenter while the parameters, $\theta_o, \theta_1, \ldots, \theta_n$, called the regression coefficients are unknown. The experimenter is to estimate the parameters $\theta_o, \theta_1, \ldots, \theta_n$, or some function of these parameters, on the basis of N uncorrelated observations (1.1) allowing the possibility that different observations may correspond to different levels.

An <u>experimental design</u> specifies a probability measure ξ concentrating mass $p_1, p_2, \ldots, p_r$ at the points $x_1, x_2, \ldots, x_r$ where the values

$$p_i N = n_i \qquad i = 1,2,\ldots,r$$

are integers. The associated experiment involves taking n_i uncorrelated observations of the random variable $y(x_i)$; $i=1,2,\ldots,r$. Once a design is prescribed and the observations are made, a standard procedure is used for estimating the parameters $\theta_o, \theta_1, \ldots, \theta_n$. The problem confronting the experimenter is to choose his design so that it will possess certain optimal properties.

If the unknown parameter vector $\theta = (\theta_o, \theta_1, \ldots, \theta_n)$ is estimated by the method of least squares, obtaining a best linear unbiased estimate say $\hat{\theta}$, then the covariance matrix of $\hat{\theta}$ is given by

$$E(\hat{\theta} - \theta)(\hat{\theta} - \theta)' = \frac{1}{N} M^{-1}(\xi) \tag{1.2}$$

where

$$M(\xi) = ((m_{ij}(\xi)))_{i,j=0}^{n} \tag{1.3}$$

$$m_{ij}(\xi) = \int_X f_i f_j \, \xi(dx)$$

and ξ assigns mass $p_i = n_i/N$ to the points x_i; $i=1,2,\ldots,r$.

If the matrix $[M(\xi)]^{-1}$ is ''small'' in some sense, or $[M(\xi)]$ is ''large'' then roughly speaking $\hat{\theta}$ is close to θ. Most criteria for discerning optimality of an experimental design are based on maximizing some functional of the matrix $M(\xi)$, which is commonly called the <u>information matrix</u> of the design.

A linear form

$$(1.4) \qquad (c,\theta) = \sum_{i=0}^{n} c_i\theta_i; \ \sum_{i=0}^{n} c_i^2 > 0$$

is called <u>estimable</u> with respect to ξ if $c = (c_o, c_1, \ldots, c_n)$ is contained in the range of the matrix $M(\xi)$. The variance of the best linear unbiased estimate of (c,θ) is given by $N^{-1} V(c,\xi)$ where

$$V(c,\xi) = \sup \frac{(c,d)^2}{(d,M(\xi)d)}$$

and the sup is taken over the set of vectors d such that the denominator is nonzero. If c is not estimable with respect to ξ, we define $V(c,\xi) = \infty$. An arbitrary design ξ is called <u>c-optimal</u> if ξ minimizes $N^{-1}V(c,\xi)$ or $V(c,\xi)$. We are throughout concerned with the characterization of c_p-optimal designs where

$$c_p = (0,0,\ldots,0,1,0,\ldots,0)$$

with 1 in the $(p+1)^{st}$ co-ordinate position and zeros elsewhere i.e.

$$(c_p,\theta) = \theta_p; \ p = 0,1,\ldots,n.$$

2. Elfving's Theorem

The following theorem due to Elfving (1952) characterizes the c-optimal design for an arbitrary vector c.

Theorem 1.1

Let

$$\mathcal{R}_+ = \{f(x) = (f_0, f_1, \ldots, f_n) | x \in X\}$$

$$\mathcal{R}_- = \{-f(x) | x \in X\}$$

$$\mathcal{R} = \text{convex hull of } \mathcal{R}_+ \cup \mathcal{R}_-$$

A design ξ_0 is c-optimal if and only if there exists a measurable function $\phi(x)$ with

$$|\phi(x)| \equiv 1 \quad \forall\, x \in X$$

such that

(i) $$\int_X \phi(x)\, f(x)\, \xi(dx) = \beta c \quad \text{for some } \beta$$

and

(ii) βc is a boundary point of $\mathcal{R}$.

Moreover βc lies on the boundary of $\mathcal{R}$ if and only if

$$\beta^{-2} = v_0^{-1} \quad \text{where} \quad v_0 = \inf_{\xi} V(c,\xi).$$

It is thus seen from the Elfving theorem that one has to ascertain the boundary of $\mathfrak{R}$ in the search for c-optimal design and the following simple but useful lemma, due to Studden (1968), characterizes the boundary points of $\mathfrak{R}$.

3. Characterization of Boundary Points of $\mathfrak{R}$

Every vector $c \in \mathfrak{R}$ can be put in the form

$$c = \sum_{\nu=1}^{k} \epsilon_\nu \, p_\nu \, f(x_\nu) \tag{1.5}$$

where $\epsilon_\nu = \pm 1$, $p_\nu > 0$ and $\sum_1^k p_\nu = 1$. The integer k may always be taken to be at most $n+2$ and at most $n+1$ if c is a boundary point of $\mathfrak{R}$.

Lemma 1.1

A vector c of the form (1.5) lies on the boundary of $\mathfrak{R}$, if and only if there exists a nontrivial ''polynomial'' $u(x) = \sum_{\nu=0}^{n} a_\nu f_\nu(x)$ such that $|u(x)| \leq 1$ for $x \in X$; $\epsilon_\nu u(x_\nu) = 1$, $\nu = 1,2,\ldots,k$ and $\sum_{\nu=0}^{n} a_\nu c_\nu = u(c) = 1$.

4. Review of Earlier Related Work

Hoel and Levine (1964) showed that if $f_i(x) = x^i$, $i = 0,1,2,\ldots,n$ $X = [-1,1]$ and $c = f(x_0)$ with $|x_0| > 1$, then the c-optimal design is supported on the Tchebycheff points $s_\nu = \cos \frac{\nu\pi}{n}$, $\nu = 0,1,\ldots,n$. These are the points where $|T_n(x)| = 1$, $T_n(x)$ being the nth Tchebycheff polynomial of the first kind.

Kiefer and Wolfowitz (1965) considered more general systems of regression functions which form a Tchebycheff system and a related set of Tchebycheff points. Among other things they characterize certain sets of vectors c, for which the c-optimal design is supported on the entire set of Tchebycheff points.

Studden (1968) gives a slight generalization of the characterization of certain sets of vectors considered by Kiefer and Wolfowitz referred to above. Before stating the theorem of Studden we introduce the following notation, and an important property of Tchebycheff systems which is frequently referred to subsequently.

If the system of regression functions $\{f_i\}_o^n$ is a Tchebycheff system on $X \equiv [a,b]$, then there exists a unique polynomial (see Karlin and Studden (1966a), Theorem II .10.1) $W(x) = \sum_{i=0}^{n} a_i^* f_i(x)$ satisfying the properties

(i) $$|W(x)| \leq 1$$

(ii) There exist (n+1) points $a \leq s_o < s_1 \ldots < s_n \leqq b$ such that
$$W(s_i) = (-1)^{n-i}, \quad i = 0,1,\ldots,n.$$

Moreover when $U(x) \equiv 1$ is a polynomial, equality occurs in (i) only for $x = s_o,\ldots,s_n$ and $s_o = a$, and $s_n = b$. For any vector $c \neq 0$,

$$(1.6) \qquad D_\nu(c) = \begin{vmatrix} f_o(s_o) & \cdots & f_o(s_{\nu-1}) & f_o(s_{\nu+1}) & \cdots & f_o(s_n) & c_o \\ f_1(s_o) & \cdots & f_1(s_{\nu-1}) & f_1(s_{\nu+1}) & \cdots & f_1(s_n) & c_1 \\ \vdots \\ f_n(s_o) & \cdots & f_n(s_{\nu-1}) & f_n(s_{\nu+1}) & \cdots & f_n(s_n) & c_n \end{vmatrix}$$

$$(1.7) \quad F\binom{0\ ,1\ ,\ldots,n}{s_o,s_1,\ldots,s_n} = \begin{vmatrix} f_o(s_o) & f_o(s_1) & \cdots & f_o(s_n) \\ f_1(s_o) & f_1(s_1) & \cdots & f_1(s_n) \\ \vdots \\ f_n(s_o) & f_n(s_1) & \cdots & f_n(s_n) \end{vmatrix}$$

$$(1.8) \qquad L_\nu(x) = \frac{F\binom{0\ ,\ldots,\nu-1,\ \nu,\nu+1\ ,\ldots,n}{s_o,\ldots,s_{\nu-1},x,s_{\nu+1},\ldots,s_n}}{F\binom{0\ ,1\ ,\ldots,n}{s_o,s_1,\ldots,s_n}}$$

so that

$$(-1)^{n-\nu} D_\nu(c) = L_\nu(c)\, F\binom{0\ ,1\ ,\ldots,n}{s_o,s_1,\ldots,s_n}$$

For any vector c, and any polynomial $u(x) = \sum_i a_i f_i(x)$, $u(c)$ stands for $\sum_i a_i c_i$.

R denotes the class of vectors c, such that $\epsilon D_\nu(c) \geq 0$ for $\nu = 0,1,\ldots,n$ where ϵ is fixed to be +1 or -1 for a given c (i.e. the $D_\nu(c)$, $\nu = 0,1,\ldots,n$ all have the same sign in a weak sense). S denotes the class of vectors c, for which $\epsilon(-1)^\nu D_\nu(c) \leq 0$, $\nu = 0,1,\ldots,n$.

With this notation, the generalization of Kiefer and Wolfowitz (1965) theorem due to Studden (1968) is given below. See also Karlin and Studden (1966).

Theorem 1.2

Suppose that $\{f_i\}_0^n$ is a Tchebycheff system such that $U(x) \equiv 1$ is a polynomial.

(a) For any design ξ

$$V(c,\xi) \geq \begin{cases} [W(c)]^2 & c \in R \\ [U(c)]^2 & c \in S \end{cases} \tag{1.9}$$

(b) Equality occurs in (1.9) for $\xi = \xi_o$ concentrating mass

$$p_\nu = \frac{|L_\nu(c)|}{\sum_{\nu=0}^{n} |L_\nu(c)|} = \frac{|D_\nu(c)|}{\sum_{\nu=0}^{n} |D_\nu(c)|}$$

at the points s_ν, $\nu = 0,1,\ldots,n$.

(c) The design ξ_o is the only design supported on $s_o < s_1 \ldots < s_n$ attaining equality in (1.9). If $c \in R$ then ξ_o is the only design attaining equality in (1.9).

For a general system of functions $\{f_i\}_o^n$, satisfying the following conditions

(i) $\{f_i\}_o^k$ for $k = n-2, n-1, n$ are T-systems on $X \equiv [-1,1]$

(ii) $f_o(x) \equiv 1$

(iii) $f_i(x) = (-1)^i f_i(-x) \quad i = 0,1,\ldots,n$

(iv) for every subset $i_1,i_2,\ldots,i_k$ of $0,1,\ldots,n$ the system $f_{i_1}(x), f_{i_2}(x),\ldots,f_{i_k}(x)$ is a T-system on the half open interval $(0,1]$.

(v) every polynomial $\sum_0^n a_i f_i$ either has fewer than n changes of direction on $(-1,1)$ or else is a constant on $(-1,1)$.

Studden (1968) showed that for $n \geq 1 \quad p \neq 0$

(a) if $n-p$ is even, $c_p \in R$ i.e. the unique c_p-optimal design is supported by the full set of Tchebycheff points $s_o,s_1,\ldots,s_n$ associated with the T-system $\{f_i\}_o^n$.

(b) for $n-p$ odd the unique c_p-optimal design is supported by the full set of Tchebycheff points $t_o,t_1,\ldots,t_{n-1}$ associated with the T-system $\{f_i\}_o^{n-1}$.

5. Problem Investigated in the Thesis

Spline functions have received considerable attention from mathematicians working in numerical analysis, interpolation and approximation theory (see Schoenberg, (1964) and Karlin, (1968) for further references). Studden and VanArman (1968) studied the problem of characterizing admissible designs, when the regression function is a polynomial spline with a finite number of fixed multiple knots. The problem investigated in this thesis is that of characterizing optimal designs of individual regression coefficients, when the regression function is a Tchebycheffian Spline Function (TSF), a general class that includes polynomial splines as a particular case.

Chapter II deals with the problem of explicit characterization of optimal designs of individual regression coefficients with a polynomial spline regression function defined on the interval $[-1,1]$ with a single multiple knot at the center of the interval. The results obtained in this case are similar to the results obtained by Studden (1968) and referred to earlier.

In Chapter III we consider the general class of <u>Tchebycheffian Spline Regression Functions</u> (TSF) defined on an interval $[a,b]$ and show that the optimal design for each individual regression coefficient is supported on the same set of points. This result led to the question of seeking necessary and sufficient conditions on the set of regression functions so as to ensure the optimal design of each individual regression coefficient to have its support on the same set of points. A sufficient condition has been obtained.

Classical Tchebycheff polynomials of first kind are taken as regression functions in Chapter IV and the support of the optimal design of each individual regression coefficient is explicitly given. When the experimenter is interested in more than one parameter in the regression model, and tries to obtain a design ξ that minimizes the maximum variance, he is looking for a minimax design with respect to single parameters, a concept introduced by Elfving (1959). In Chapter V we try to obtain this type of minimax design in the case of ordinary polynomial regression and present a partial solution.

CHAPTER II

POLYNOMIAL SPLINE REGRESSION WITH A SINGLE MULTIPLE KNOT AT THE CENTER

1. Oscillatory Polynomials W and W_1

As regression functions we consider the (2n-k+2) linearly independent and continuous functions $\{x^i\}_0^n \cup \{x_+^i\}_k^n$ defined on $[-1,1]$ where

$$x_+^i = \begin{cases} x^i & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \qquad i = k,k+1,\ldots,n$$

and k is an integer $\leq n$.

A ''polynomial'' is a linear combination of these (2n-k+2) functions. We show the existence and uniqueness of two polynomials denoted by $W(x)$, and $W_1(x)$ and study some of their properties.

Polynomial W(x):

Lemma 2.1

There exists a unique polynomial $W(x)$ satisfying

(i) $|W(x)| \leq 1$

(ii) The set $E = \{x : |W(x)| = 1\}$ contains precisely (2n-k+2) points.

(iii) $W(x)$ attains its supremum at each of the points of the set E with alternating signs and is of the form

$$\sum_{j=0}^{\frac{1}{2}(n-1)} a_{2j}\, x^{2j} + \sum_{j=0}^{\frac{1}{2}(n-k)} a_{k+2j}(x^{k+2j} - 2x_+^{k+2j}) \quad \text{when } n \text{ and } k \text{ are both odd}$$

$$\sum_{j=0}^{\frac{1}{2}(n-1)} a_{2j+1}\, x^{2j+1} + \sum_{j=0}^{\frac{1}{2}(n-k-1)} a_{k+2j}(x^{k+2j} - 2x_+^{k+2j}) \quad \text{when } n \text{ is odd and } k \text{ is even}$$

$$\sum_{j=0}^{\frac{1}{2} n} a_{2j}\, x^{2j} + \sum_{j=0}^{\frac{1}{2}(n-k-1)} a_{k+2j}(x^{k+2j} - 2x_+^{k+2j}) \quad \text{when } n \text{ is even and } k \text{ is odd}$$

$$\sum_{j=0}^{\frac{1}{2}(n-2)} a_{2j+1}\, x^{2j+1} + \sum_{j=0}^{\frac{1}{2}(n-k)} a_{k+2j}(x^{k+2j} - 2x_+^{k+2j}) \quad \text{when } n \text{ is even and } k \text{ is even}$$

and the coefficients a_j are $\neq 0$.

<u>Proof</u>:

We will prove the lemma for the case where n and k are both odd. The proof for the other cases is the same word for word.

Consider the function

$$f(x) = 2x_+^k - x^k$$

clearly $f(x)$ is an even function. Let V be the linear space spanned by

$$\{x^i\}_o^n \cup \{x_+^i\}_{k+1}^n$$

Then $g(x)\epsilon V$ implies $g(-x)\epsilon V$. Hence there exists a best approximation of $f(x)$, say $h(x)$ with respect to V which is also even (see Meinardus (1967) pp. 26-67). Hence $h(x)$ has the form

$$h(x) = \sum_{j=0}^{\frac{1}{2}(n-1)} \alpha_{2j}\, x^{2j} + \sum_{j=1}^{\frac{1}{2}(n-k)} \alpha_{k+2j}(2x_+^{k+2j} - x^{k+2j}).$$

We may thus consider only the space V_1 spanned by $\{x^{2j}\}_0^{\frac{1}{2}(n-1)} \cup \{2x_+^{k+2j} - x^{k+2j}\}_1^{\frac{1}{2}(n-k)}$. Each function in V_1 is clearly an even function, and f is even. Therefore a best approximation of f with respect to V_1 is also a best approximation of x^k with respect to the space V_2 spanned by $\{x^{2j}\}_0^{\frac{1}{2}(n-1)} \cup \{x^{k+2j}\}_1^{\frac{1}{2}(n-k)}$ on the interval $[0,1]$, and the dimension of V_2 is $n-\frac{1}{2}(k-1)$. But on the interval $[0,1]$ the spanning set of functions of the space V_2 is a Tchebycheff system with a unit element and hence best approximation of $f(x)$ with respect to V_2 is unique. i.e. $h(x)$ is unique and $f-h$ possesses precisely $n-\frac{1}{2}(k-3)$ extremal points including the end points 0 and 1 and $f-h$ attains its norm at these points with alternating signs. (See Meinardus (1967) pp. 29). Thus best approximation of f with respect to V_1 on the interval $[-1,1]$ is unique and has precisely $(2n-k+2)$ extremal points including -1, 0 and 1 at each of which $f(x)-h(x)$ attains its norm with alternating signs.

Let $W(x) = [f(x)-h(x)] \div ||f-h||$ where $||f-h|| = \sup_{-1\leq x\leq 1} |f(x)-h(x)|$.

It is now easily seen that $W(x)$ satisfies all the conditions of the lemma. Note that the $(2n-k+2)$ extreme points of $W(x)$ are symmetric about 0 and $n-\frac{1}{2}(k-1)$ are in $[-1,0)$, and $n-\frac{1}{2}(k-1)$ are in $(0,1]$.

Polynomial $W_1(x)$:

Lemma 2.2

There exists a unique polynomial $W_1(x)$ satisfying

(i) $|W_1(x)| \leq 1$

(ii) The set $E_1 = \{x: |W_1(x)| = 1\}$ consists of precisely $(2n-k+1)$ points.

(iii) $W_1(x)$ attains its supremum with alternating signs at each of the points of the set E_1 and is of the form

$$\sum_{j=0}^{\frac{1}{2}(n-1)} b_{2j+1} x^{2j+1} + \sum_{j=0}^{\frac{1}{2}(n-k-2)} b_{k+2j+1}(x^{k+2j+1} - 2x_+^{k+2j+1})$$ when n and k are both odd,

$$\sum_{j=0}^{\frac{1}{2}(n-1)} b_{2j}\, x^{2j} + \sum_{j=0}^{\frac{1}{2}(n-k-1)} b_{k+2j+1}(x^{k+2j+1} - 2x_+^{k+2j+1})$$ when n is odd and k is even,

$$\sum_{j=0}^{\frac{1}{2}(n-2)} b_{2j+1}\, x^{2j+1} + \sum_{j=0}^{\frac{1}{2}(n-k-1)} b_{k+2j+1}(x^{k+2j+1} - 2x_+^{k+2j+1})$$ when n is even and k is odd,

$$\sum_{j=0}^{\frac{1}{2}n} b_{2j}\, x^{2j} + \sum_{j=0}^{\frac{1}{2}(n-k-2)} b_{k+2j+1}(x^{k+2j+1} - 2x_+^{k+2j+1})$$ when n is even and k is even,

and the coefficients b_j are $\neq 0$. If k=n, the terms with $(x^{k+2j+1} - 2x_+^{k+2j+1})$ are omitted. The polynomial W_1 in this case is $T_n(x)$.

Proof:

If k=n, $W_1(x) = T_n(x)$ and properties (i) to (iii) are well known.

The construction for the case $k \leq n-1$, is similar to that of $W(x)$ except that we take

$$f_1(x) = 2x_+^{k+1} - x^{k+1}$$

and consider its best approximation with respect to V_1, spanned by $\{x^i\}_0^n \cup \{x_+^i\}_{k,i\neq k+1}^n$ and set $W_1(x) = [f_1(x)-h_1(x)] \div ||f_1-h_1||$ where h_1 is the unique best approximation of f_1. It is easy to verify that the W_1 so constructed satisfies the conditions stated in the lemma.

2. Zeros of a Polynomial

We need the following theorem concerning the zeros of a polynomial for subsequent use.

Theorem 2.1

Let $S(n,k;x) = \sum_{i=0}^{n} d_i x^i + \sum_{i=k}^{n} d_i' x_+^i$ with at least one of the d_i's=0 for some $i \geq k-1$; if $S(n,k;x)$ has $(2n-k+1)$ distinct zeros, and does not vanish identically in any interval containing two of these distinct zeros, then $S(n,k;x) \equiv 0$.

We first state and prove the following lemma.

Lemma 2.3

Theorem 2.1 is true for k=1.

Proof:

Since $d_i = 0$ for some $i \geq 0$ we consider two cases. (i) $d_0 = 0$. Then S can have at most (n-1) distinct zeros in $[-1,0)$, and at most (n-1) distinct zeros in (0,1]. Thus it can have at most (2n-1) distinct zeros, including 0. Hence if it has 2n distinct zeros it is clearly $\equiv 0$. (ii) If $d_o \neq 0$ and $d_i = 0$ for some $i \geq 1$ then S can have at most (n-1) distinct zeros in [-1,0), and at most n in (0,1], and thus can have at most (2n-1). Hence $S \equiv 0$, if it has 2n distinct zeros. This completes the proof of the lemma.

Proof of Theorem 2.1:

Let $k > 1$. Since S has (2n-k+1) distinct zeros and does not vanish identically in any interval containing two of these zeros, we claim that its derivative S', by Rolle's theorem has (2n-k) distinct zeros, and cannot vanish identically in between any two of these zeros. Suppose it vanishes between two zeros say z_1 and z_2. Then S is a constant on $[z_1,z_2]$ and has one of its distinct zeros in its interior and as such is $\equiv 0$ on $[z_1,z_2]$. If $z_2 < 0$, then $S = \sum_{i=k}^{n} d_i' x_+^i$ and can have at most only (n-k) distinct zeros, and does not vanish identically between them. As $2n-k+1 > n-k$, this implies $S \equiv 0$, a contradiction. Similar argument gives a contradiction when $z_1 > 0$ or $0 \in [z_1\ z_2]$. Differentiating S, (k-1) times we have,

$$S^{(k-1)} = d^*_{k-1} + d^*_k x + \ldots + d^*_n x^{n-k+1} + \sum_{i=1}^{n-k+1} d^{*\prime}_i x_+^i$$

and since $d_i = 0$ for some $i \geq k-1$ we also have $d^*_j = 0$ for some $j \geq k-1$ and $S^{k-1} = S(N,1;x)$ where $N = n-k+1$ and has by Rolle's theorem $(2n-k+1)-(k-1) = 2N$ distinct zeros and does not vanish identically in any interval containing two of these zeros. So $S^{k-1}(n,k;x) \equiv 0$. Hence $S(n,k;x) \equiv 0$.

3. Minimizing Property of the Polynomials W and W_1

Lemma 2.4

Among all polynomials u(x), with coefficient of x^j equal to unity, $W(x)/a_j$ minimizes

$$\sup_{-1 \leq x \leq 1} |u(x)|$$

for $j = 0,2,4,\ldots,k-1;\ k,k+1,\ldots,n$ (k odd)

$j = 1,3,5,\ldots,k-1;\ k,k+1,\ldots,n$ (k even).

$W_1(x)/b_j$ has the stated property for $j = 1,3,5,\ldots,k-2$ (k odd); $j = 0,2,4,\ldots,k-2$ (k even).

Proof:

Consider the case where n and k are both odd.

Let j be even and $0 \leq j \leq k-1$. Consider the space V, spanned by

$$\{x^i\}_{i=0,i\neq j}^{n} \cup \{x_+^i\}_k^n.$$

If $g(x)\epsilon V$, then so does $g(-x)$. Let $f(x) = x^j;\ x \epsilon [-1,1]$. f(x) is even and hence there exists a best approximation $P(x)$ of f with respect to V which is also even. If we consider the difference

$$W(x)/a_j - [x^j - P(x)]$$

either it vanishes at one of the extremal points of W(x) or it has (2n-k+1) distinct zeros in $[-1,1]$, and does not vanish identically in any interval containing two of these zeros. In the first case

$$||W(x)/a_j|| = |x^j - P(x)| = ||x^j - P(x)||$$

This implies that $W(x)/a_j$ minimizes

$$\sup_{-1 \leq x \leq 1} |u(x)|.$$

In the second case the difference is easily seen to be of the form

$$\beta_0 + \beta_2 x^2 + \ldots + \beta_{j-2} x^{j-2} + \beta_{j+2} x^{j+2} + \ldots + \beta_{n-1} x^{n-1} + \beta_k (x^k - 2x_+^k) + \ldots +$$

$$\beta_n (x^n - 2x_+^n),$$

and vanishes at $n - \frac{1}{2}(k-1)$ points in $[-1,0)$ and $n - \frac{1}{2}(k-1)$ points in $(0,1]$. But from Descartes' rule of signs, it can have at most $n - \frac{1}{2}(k+1)$ zeros in $(0, 1]$. Hence the difference vanishes identically and hence the result. The proof of the case k even and j odd integer $\leq k-1$ is similarly treated. Consider now the case $k \leq j \leq n$; In this case the difference $W(x)/a_j - [x^j - P(x)]$, is of the form

$$\sum_{\substack{0 \\ i \neq j}}^{n} d_i x^i + \sum_{i=k}^{n} d_i' x_+^i$$

and will either vanish at one of the extreme points of $W(x)$ or has $(2n-k+1)$ distinct zeros and does not vanish identically between any two of these. Hence from Theorem 2.1 it must be identically equal to zero and hence the result. The proofs for W_1 are similar.

<u>Lemma 2.5</u>

Among all polynomials with coefficient of x_+^{k+2j} equal to unity

$-W(x)/2a_{k+2j}$ minimizes

$$\sup_{-1 \leq x \leq 1} |u(x)|$$

for $j = 0,1,2,\ldots,\ell$ where $\ell = \frac{1}{2}(n-k)$ or $\frac{1}{2}(n-k-1)$ according as $(n-k)$ is even or odd.

Lemma 2.6

Among all polynomials with coefficient of x_+^{k+2j+1} equal to unity $-W_1(x)/2b_{k+2j+1}$ minimizes

$$\sup_{-1 \leq x \leq 1} |u(x)|$$

for $j = 0,1,2,\ldots,m$ where $m = \frac{1}{2}(n-k-2)$ or $\frac{1}{2}(n-k-1)$ according as $n-k$ is even or odd. We omit the proofs of these two lemmas as they are similar to that of lemma 2.4 which is treated in detail.

4. Optimal Designs of Individual Regression Coefficients

We now take as regression functions $\{f_i\}_o^n \cup \{g_j\}_{j=k}^n$ where

$$f_i = x^i; \quad i = 0,1,2,\ldots,n$$

$$g_j = x_+^j; \quad j = k, k+1,\ldots,n$$

defined on $[-1,1]$ and denote the corresponding regression coefficients by

$$\{\theta_i\}_o^n \cup \{\theta_j'\}_k^n$$

and state the following theorem concerning the optimal designs of the individual regression coefficients.

Theorem 2.2

(i) Optimal design for estimating θ_o is unique and is supported on $x = 0$.

(ii) For $k \geq 2$;

(a) If k is odd, the unique optimal designs for θ_j $(k \leq j \leq n)$; θ_j (j even and $\leq k-1$); and θ_j' (j odd) are supported on the full set E. The unique optimal designs for θ_j (j odd and $\leq k-2$) and θ_j' (j even) are supported on the full set E_1.

(b) If k is even, the unique optimal designs for θ_j $(k \leq j \leq n)$; θ_j (j odd and $\leq k-1$) and θ_j' (j even) are all supported on the full set E. The unique optimal designs for θ_j (j even and $\leq k-2$) and θ_j' (j odd) have for their support the full set E_1.

(iii) If $k = 1$, the unique optimal design for θ_j is supported on the full set of $(n+1)$ points of the set E in $[-1,0]$; the unique optimal design for θ_j' (j odd) is supported on the full set E and for θ_j' (j even) the support is the set E_1.

As for the proof of Theorem 2.2 we note that the minimizing properties of W and W_1 stated in section 3 and the lemma given below due to Kiefer and Wolfowitz (see Kiefer and Wolfowitz (1959)), immediately establish the supports stated in the theorem and hence the only part of the theorem to be proved relates to the statement that the respective supports are full which is dealt with in the next section.

Lemma 2.7

Let $\{f_i\}_o^n$ be continuous and linearly independent, defined on

$[-1,1]$ and we consider these as our regression functions, and $\{\theta_i\}_0^n$ are the respective regression coefficients.

$$\rho = \sup_{-1 \leq x \leq 1} \left|f_j - \sum_{\substack{i=0 \\ i \neq j}}^{n} \alpha_i f_i\right| = \min_{\beta} \sup_{-1 \leq x \leq 1} \left|f_j - \sum_{\substack{i=0 \\ i \neq j}}^{n} \beta_i f_i\right|$$

$$B = \{x : \left|f_j - \sum_{\substack{i=0 \\ i \neq n}}^{n} \alpha_i f_i\right| = \rho\} .$$

Optimal design for estimating θ_j satisfies $\xi^*(B) = 1$.

For a proof see Kiefer and Wolfowitz (1959).

5. Supports of the Optimal Designs

Let n and k be odd. Consider θ_j $j \leq k-1$; $j \neq 0$ and j even. From Theorem 2.2 we know that the optimal design for θ_j in this case has for its support the set E consisting of $(2n-k+2)$ points. Let $\{x_i\}_0^{2n-k+1}$ be these points with $x_0 = -1$, $x_{2n-k+1} = 1$, $x_{n-\frac{1}{2}(k-1)} = 0$, and the remaining are symmetric about zero. Moreover $x_0 < x_1 \ldots < x_{2n-k+1}$. Let $\{p_i\}_0^{2n-k+1}$ be the probabilities associated with these points by the optimal design. Then there exists a solution $\{\epsilon_\nu p_\nu\}$, by Elfving's Theorem, to the system of equations

$$\beta c_j = \sum_{\nu=0}^{2n-k+1} \epsilon_\nu \, p_\nu \, f(x_\nu); \quad \text{where}$$

$$\beta^{-1} = \text{coefficient of } x^j \text{ in } W(x).$$

Suppose $p_i = 0$ where $i > n - \frac{1}{2}(k-1)$. Then there exists a polynomial

$$P(x) = \sum_{i=0}^{n} d_i x^i + \sum_{i=k}^{n} d_i' \, x_+^i$$

such that (i) $\Sigma\, d_i^2 + \Sigma\, d_i'^2 > 0$, (ii) $d_j = 0$, and (iii) $P(x_\nu) = 0$ for $\nu \neq i$. Consider $Q(x) = P(x) + P(-x)$, and note that, $x_+^i - (-x)_+^i \equiv x^i$ if i is odd and $x_+^i + (-x)_+^i \equiv x^i$ if i is even. Then

$$Q(x) = P(x)+P(-x) = \sum_{\substack{\nu=0 \\ \nu \neq j/2}}^{\frac{1}{2}(n-1)} d_{2\nu}\, x^{2\nu} + \sum_{\nu=1}^{\frac{1}{2}(n-k)} d'_{k+2\nu-1}\, x^{k+2\nu-1}$$

$$+ \sum_{\nu=0}^{\frac{1}{2}(n-k)} d'_{k+2\nu} \left[2x_+^{k+2\nu} - x^{k+2\nu} \right]$$

and for $x < 0$

$$Q(x) = \sum_{\substack{\nu=0 \\ \nu \neq j/2}}^{\frac{1}{2}(k-1)} d_{2\nu}\, x^{2\nu} + \sum_{\nu=1}^{\frac{1}{2}(n-k)} d_{k+2\nu-1}\, x^{2\nu+k-1} + \sum_{\nu=1}^{\frac{1}{2}(n-k)} d'_{k+2\nu-1}\, x^{k+2\nu-1}$$

$$- \sum_{\nu=0}^{\frac{1}{2}(n-k)} d'_{k+2\nu}\, x^{k+2\nu}$$

Therefore $Q(x)$ can at most have $n - \frac{1}{2}(k+3)$ zeros in $[-1,0)$. (Use Descartes' rule of signs, and note also that $d_o = 0$; as $P(0) = 0$). But actually it has $(n-1) - \frac{1}{2}(k-1)$ zeros in $[-1,0)$, and since $(n-1) - \frac{1}{2}(k-1) > n - \frac{1}{2}(k+3)$ we have

$$d_{2\nu} = 0; \quad \nu = 0,1,2,\ldots, \tfrac{1}{2}(k-1)$$

$$d_{k+2\nu-1} + d'_{k+2\nu-1} = 0; \quad \nu = 1,2,\ldots, \tfrac{1}{2}(n-k)$$

$$d'_{k+2\nu} = 0; \quad \nu = 0,1,\ldots, \tfrac{1}{2}(n-k).$$

But this implies that for $x < 0$

$$P(x) = \sum_{\nu=1}^{\frac{1}{2}(n-k)} d_{k+2\nu-1}\, x^{k+2\nu-1} + \sum_{\nu=0}^{\frac{1}{2}(n-k)} d_{k+2\nu}\, x^{k+2\nu}$$

and hence can have at most $n-k$ zeros in $[-1,0)$, but actually has $n-\frac{1}{2}(k-1)$, which implies that

$$d_{k+2\nu} = 0; \quad \nu = 0,1,\ldots, \tfrac{1}{2}(n-k)$$

$$d_{k+2\nu-1} = 0; \quad \nu = 1,2,\ldots, \tfrac{1}{2}(n-k)$$

i.e. $\qquad P(x) \equiv 0$, a contradiction.

The proof of the remaining cases is exactly similar to the one given above and when k is even we consider $P(x)-P(-x)$ and proceed as above.

It may be noted that the above method is exactly similar to the one used by Studden (1968). For those θ's whose support is on the set E_1, the proof of $p_i \neq 0$ is reduced to the earlier situation, by dropping the component corresponding to x_+^k, in the system of equations

$$\beta c_j = \sum_{\nu=0}^{2n-k} \epsilon_\nu \, p_\nu \, f(t_\nu) \; ;$$

β^{-1} is the coefficient of x^j in W_1, and $\{t_\nu\}_0^{2n-k}$ are the points of the set E_1.

6. Illustrations

Example 1

Let $n = 2$, $k = 1$ and the regression equation be denoted by

$$\theta_o + \theta_1 x + \theta_2 x^2 + \theta_1' x_+ + \theta_2' x_+^2 ; \quad x \in [-1,1]$$

Then the polynomials $W(x)$ and $W_1(x)$ are

$$W(x) = 1 + 8x + 8x^2 - 16x_+$$

$$W_1(x) = -\frac{2}{c}\, x - \frac{x^2}{c^2} + \frac{2}{c^2}\, x_+^2 \; ; \; c = \sqrt{2} - 1$$

The sets E and E_1 are

$$E = \{-1, -\tfrac{1}{2}, 0, \tfrac{1}{2}, 1\}$$

$$E_1 = \{-1, -c, \; c, 1\}$$

The optimal designs for θ_1, θ_2, and θ_1' are supported on the set E, with respective weights

$$(\,1/8\,,\ 4/8\,,\ 3/8\,,\ 0\,,\ 0\,);$$

$$(\,1/4\,,\ 2/4\,,\ 1/4\,,\ 0\,,\ 0\,);\ \text{and}$$

$$(\,1/16,\ 4/16\,,\ 6/16\,,\ 4/16\,,\ 1/16).$$

The optimal design for θ_2' is supported on the set E_1 with weights

$$\left\{\frac{c}{2(1+c)}\,,\ \frac{1}{2(1+c)}\,,\ \frac{1}{2(1+c)}\,,\ \frac{c}{2(1+c)}\right\}.$$

Example 2

Let $n = 2$, $k = 2$ and the regression equation be denoted by

$$\theta_o + \theta_1 x + \theta_2 x^2 + \theta_2' x_+^2\,; \quad x \in [-1,1]$$

The polynomial $W(x)$ is the same as $W_1(x)$ of the previous example, and the set E is the same as E_1 of the previous example. The optimal designs for θ_1, θ_2 and θ_2' are supported on the full set $E_1 = \{-1,-c,c,1\}$ with respective weights

$$\left\{\frac{c^2}{2(1+c^2)}\,,\ \frac{1}{2(1+c^2)}\,,\ \frac{1}{2(1+c^2)}\,,\ \frac{c^2}{2(1+c^2)}\right\},$$

$$\left\{\frac{c(2+c)}{2(1+c)^2}\,,\ \frac{1+2c}{2(1+c)^2}\,,\ \frac{1}{2(1+c)^2}\,,\ \frac{c^2}{2(1+c)^2}\right\},\ \text{and}$$

$$\left\{\frac{c}{2(1+c)}\,,\ \frac{1}{2(1+c)}\,,\ \frac{1}{2(1+c)}\,,\ \frac{c}{2(1+c)}\right\}.$$

Example 3

Let $n = 3$, $k = 3$ and the regression equation be denoted by

$$\theta_o + \theta_1 x + \theta_2 x^2 + \theta_3 x^3 + \theta_3' x_+^3 ; \quad x \in [-1,1].$$

The polynomials $W(x)$ and $W_1(x)$ are

$$W(x) = -1 + 27/2\, x^2 + 27/2\, x^3 - 27\, x_+^3$$

$$W_1(x) = -3x + 4\, x^3$$

The sets E and E_1 are

$$E = \{-1, \; -2/3, \; 0 \;, \; 2/3, \; 1\}$$

$$E_1 = \{-1, \; -1/2, \; 1/2 \;, \; 1\}$$

The optimal designs for θ_2, θ_3, and θ_3' are supported on the set E with respective weights

$$\{8/108, \; 27/108, \; 38/108, \; 27/108, \; 8/108\}$$

$$\{32/180, \; 63/180, \; 50/180, \; 27/180, \; 8/180\}$$ and

$$\{4/36 \;, \; 9/36 \;, \; 10/36 \;, \; 9/36 \;, \; 4/36\} .$$

The optimal design for θ_1 is supported on E_1 with weights $\{1/18, 8/18, 8/18, 1/18\}$.

CHAPTER III

OPTIMAL DESIGNS WITH A TCHEBYCHEFFIAN SPLINE REGRESSION FUNCTION (TSF)

1. Definition of a TSF

Starting with (n+1) functions $w_o, w_1, \ldots, w_n$ which are strictly positive on $[a,b]$ and such that w_k is of continuity class $C^{n-k}[a,b]$ we form the system

$$
(3.1)\qquad
\begin{aligned}
u_o(x) &= w_o(x) \\
u_1(x) &= w_o(x) \int_a^x w_1(\xi_1)\, d\xi_1 \\
&\vdots \\
u_n(x) &= w_o(x) \int_a^x w_1(\xi_1) \int_a^{\xi_1} w_2(\xi_2) \cdots \int_a^{\xi_{n-1}} w_n(\xi_n) d\xi_n \cdots d\xi_1
\end{aligned}
$$

It is shown (see Karlin and Studden (1966a) pp. 379 Theorem 1.2) that the functions $u_o, u_1, \ldots, u_n$ in (3.1) comprise an Extended Complete Tchebycheff (ECT) system on $[a,b]$ obeying the boundary conditions

$$
(3.2)\qquad u_k^{(p)}(a) = 0;\ p = 0,1,\ldots,k-1;\ k = 1,2,\ldots, n.
$$

A function s(x) is said to be a Tchebycheffian Spline Function (TSF) on $[a,b]$ of order (n+1) or degree n, with k knots $\{\eta_i\}_1^k$,

$$\eta_o = a < \eta_1 < \eta_2 < \ldots < \eta_k < b = \eta_{k+1}$$

provided (i) $s(x)$ reduces to a u-polynomial in the ECT system $\{u_i\}_o^n$ in each of the intervals (η_i, η_{i+1}); $i = 0,1,\ldots,k$. (ii) $s(x)$ has $n-1$ continuous derivatives.

The class of TSF's of degree n with k prescribed knots $\{\eta_i\}_1^k$ will be designated by $S_{n,k}(\eta_1,\eta_2,\ldots,\eta_k)$. Lemma 9.1 pp. 437 of Karlin and Studden (1966a) shows that $S_{n,k}(\eta_1,\eta_2,\ldots,\eta_k)$ is precisely the set of functions

$$(3.3) \qquad s(x) = \sum_{i=0}^{n} a_i u_i(x) + \sum_{j=1}^{k} a_{n+j}\, \phi_n\,(x;\eta_j)$$

where

$$(3.4) \quad \phi_n(x;\eta) = \begin{cases} w_o(x) \int_\eta^x w_1(\xi_1) \int_\eta^{\xi_1} w_2(\xi_2) \ldots \int_\eta^{\xi_{n-1}} w_n(\xi_n)\, d\xi_n \ldots d\xi_1 & \text{if } \eta \leq x \leq b \\ 0 & \text{if } a \leq x \leq \eta \end{cases}$$

Notice that $\phi_n(x;a) = u_n(x)$.

2. Preliminary Theorems and Lemmas on Best Approximation in the Uniform Norm by a TSF

In view of the representation (3.3) spline approximation problem with fixed knots $\{\eta_i\}_1^k$ reduces to the standard linear approximation problem of determining the best approximation of a given continuous function, in the uniform norm, by a linear combination of $(n+k+1)$

functions $\{u_i\}_o^n \cup \{\phi_n(x;\eta_j)\}_1^k$.

By the general linear theory (see Meinardus, G. (1967), pp. 1) we have the following theorem.

Theorem 3.1

Let $a < \eta_1 < \eta_2 \ldots, \eta_k < b$ be fixed. Suppose $f(x) \in C[a,b]$. Then there exists a best approximation $s^*(x) \in S_{n,k}(\eta_1,\ldots,\eta_k)$. i.e. $s^*(x)$ satisfies

$$||s^*-f|| \leq ||s-f|| = \max_{a \leq x \leq b} |s(x)-f(x)|$$

for every $s(x) \in S_{n,k}(\eta_1,\ldots,\eta_k)$. The following two theorems are due to Schumaker (1967a and 1967b).

Theorem 3.2

Let $f \in C[a,b]$. Then there exists an $s(x) \in S_{n,k}(\eta_1,\ldots,\eta_k)$ such that $f-s$ alternates at least $(n+k+1)$ times on $[a,b]$. i.e. there exist $\{x_i\}_1^{n+k+2}$ points $a \leq x_1 < x_2 \ldots < x_{n+k+2} = b$, such that

$$f(x_i) - s(x_i) = \epsilon(-1)^i \max_{a \leq x \leq b} |f(x) - s(x)|$$

where $\epsilon = \pm 1$; for $i = 1,2,\ldots,n+k+2$.

Theorem 3.3

Suppose $f \in C[a,b]$. Then $s(x) \in S_{n,k}(\eta_1,\ldots,\eta_k)$ is the unique best approximation of f if there exist points $a \leq t_1 < t_2 \ldots < t_{n+k+2} \leq b$ with

$$t_{i+1} < \eta_i < t_{n+i+1}; \quad i = 1,2,\ldots,k$$

$$f(t_i)-s(t_i) = (-1)^i \zeta A_{n,k}; \quad i = 1,2,\ldots,(n+k+2)$$

where

$$A_{n,k} = \min_{s \in S_{n,k}(\eta_1,\ldots,\eta_k)} ||f-s||$$

$$\zeta = \pm 1.$$

Zero Structure of TSF's

The following lemma on the simple zeros of a TSF is due to Schumaker (1967b).

Lemma 3.1

Suppose $s(x) \in S_{n,k}(\eta_1,\ldots,\eta_k)$ possesses the zeros $x_1 < x_2 \ldots < x_{n+k}$ and $s(x)$ does not vanish identically on any interval containing two of these zeros. Then

$$x_i < \eta_i < x_{n+i}; \quad i = 1,2,\ldots,k. \tag{3.5}$$

Moreover $s \in S_{n,k}$ can have at most $(n+k)$ distinct zeros provided s does not vanish identically between any two of them.

One of the perversities of TSF's is that it is possible for a non-null TSF to vanish on an interval (simple examples already exist in the case of polynomial splines). We shall use the following conventions when counting zeros of a TSF. (See Karlin and Schumaker (1967) and Studden and VanArman (1968)).

(a) No zeros are counted on any open interval (η_i, η_j) if $s(x) \equiv 0$ there.

(b) The multiplicity of a zero $z \neq \eta_i$ $i = 1,2,\ldots,k$ is r if

$$s^{(j)}(z) = 0 \quad j = 0,1,\ldots,r-1; \quad s^{(r)}(z) \neq 0$$

(c) If $s(x) \equiv 0$ on (η_{i-1}, η_i) and $\not\equiv 0$ on (η_i, η_{i+1}) the zero at η_i is counted as in (b) using the right had derivatives. Similarly we use left had derivatives for $s(x) \not\equiv 0$ on (η_{i-1}, η_i) and $\equiv 0$ on (η_i, η_{i+1}).

(d) If $s(x) \not\equiv 0$ on (η_{i-1}, η_i) or (η_i, η_{i+1}) and

$$0 = s^{(j)}(\eta_i-) = s^{(j)}(\eta_i+) \quad j = 0,1,\ldots,r-1$$

$$A = s^{(r)}(\eta_i-),\ s^{(r)}(\eta_i+) = B \quad \text{and}$$

$$A \neq B,$$

then

$$\eta_i \text{ is a zero of order}$$

(i) r if $AB > 0$

(ii) $r+1$ if $AB < 0$

(iii) $r+1$ if $AB = 0$ and $B-A > 0$

(iv) $r+2$ if $AB = 0$ and $B-A < 0$

We let $Z(s)$ denote the number of zeros of s according to the above conventions.

<u>Lemma 3.2</u>

A non-trivial TSF $s(x) \in S_{n,k}(\eta_1,\ldots,\eta_k)$ has $Z(s) \leq n+k$.

For a proof see Studden and VanArman (1968) and Karlin and Schumaker (1967).

3. Uniqueness and Existence of the Oscillatory Polynomial $W(x)$

Utilizing the theorems and lemmas stated above we now state and prove

Theorem 3.4

Let $n \geq 2$ and $w_0(x)$ in the system (3.1) be $\equiv 1$. There exists a unique $W(x) \in S_{n,k}(\eta_1, \eta_2, \ldots, \eta_k)$ satisfying

(i) $|W(x)| \leq 1 \quad \forall\, x \in [a, b]$

(ii) The set of points $\{x : |W(x)| = 1\}$, consists of precisely $(n+k+1)$ points $\{x_i\}_1^{n+k+1}$

(iii) $W(x_i) = \epsilon(-1)^i$ where $\epsilon = \pm 1 \quad i = 1, 2, \ldots, n+k+1$

(iv) $x_{i+1} < \eta_i < x_{n+i}$; $i = 1, 2, \ldots, k$.

Proof:

Consider $f = \phi_n(x; \eta_k) \in C[a, b]$. Theorem 3.2 assures the existence of an $s^*(x) \in S_{n,k-1}(\eta_1, \eta_2, \ldots, \eta_{k-1})$ such that s^* is a best approximation of f with respect to the class $S_{n,k-1}(\eta_1, \eta_2, \ldots, \eta_{k-1})$, and $f - s^*$ alternates at least $(n+k)$ times. Hence there exist $(n+k+1)$ points $\{x_i\}_1^{n+k+1}$ where $a \leq x_1 < x_2 \ldots < x_{n+k+1} \leq b$ and

$$f(x_i) - s^*(x_i) = \epsilon(-1)^i \max_{a \leq x \leq b} |f(x) - s^*(x)|$$

$$i = 1, 2, \ldots, n+k+1$$

Set

$$W(x) = \frac{1}{||f - s^*||} [f(x) - s^*(x)]$$

$$= \sum_{i=0}^{n} a_i^* u_i(x) + \sum_{j=1}^{k} a_{n+j}^* \phi_n(x; \eta_j)$$

Clearly $$||W(x)|| = 1$$

Hence (i) of Theorem 3.4 is established. Also W(x) attains its norm with alternating signs at each of (n+k+1) points $\{x_i\}_1^{n+k+1}$. Hence (iii) of Theorem 3.4 is proved. $W'(x) \epsilon S_{n-1,k}(\eta_1, \eta_2, \ldots, \eta_k)$ and has at least (n+k-1) distinct zeros $\{x_i\}_2^{n+k}$ and does not vanish identically between any two of them. Hence from Lemma 3.1 we have

$$x_{i+1} < \eta_i < x_{n+i} \quad i = 1,2,\ldots,k. \tag{3.6}$$

Thus (iv) of Theorem 3.4 is established. If the set $\{x: |W(x)| = 1\}$ has at least (n+k+2) points, then $W'(x) \epsilon S_{n-1,k}$ must vanish at every interior such point so that W' will have at least (n+k) distinct zeros. If W' does not vanish identically between any two of these zeros, we have a contradiction, since W' cannot have more than (n+k-1) such zeros. Hence W' may vanish identically in (x_o, x_1) if the additional point x_o at which $|W(x_o)| = 1$ is in $[a, x_1)$ or in (x_{n+k+1}, x_o) if x_o is in $(x_{n+k+1}, b]$ or in (x_i, x_o) if x_o is in (x_i, x_{i+1}). In the first case we count x_1 as a zero of W' with multiplicity at least one according to our convention and together with the remaining (n+k-1) distinct zeros between no two of which it vanishes, $z(W') \geq n+k$ which is a contradiction. Similarly in the second case x_{n+k+1} will be a zero of W' with multiplicity at least one again leading to a contradiction. In the third case we have

$$W'(x) \equiv 0 \quad \text{in} \quad (x_i, x_o) \quad \text{and}$$

$$W'(x) \not\equiv 0 \quad \text{in} \quad (x_o, x_{i+1})$$

and $W'(x)$ being continuous $W'(x_o+) = 0$ and hence is a zero of multiplicity at least one, and hence we again have $Z(W') \geq n+k$ a contradiction. Hence the set $\{x: |W(x)| = 1\}$ consists of precisely $(n+k+1)$ points, $\{x_i\}_1^{n+k+1}$. Similar considerations establish

$$x_1 = a \; ; \quad \text{and} \quad x_{n+k+1} = b.$$

Finally (3.6) and Theorem 3.3 imply that $s^*(x)$ is the unique best approximation of f with respect to $S_{n,k-1}(\eta_1,\ldots,\eta_{k-1})$. This completes the proof of Theorem 3.4.

4. Optimal Designs of Individual Regression Coefficients with a TSF as Regression Function

Theorem 3.5

Let $n \geq 2$

$$E(y|x) = \sum_{i=0}^{n} \theta_i \, u_i(x) + \sum_{j=1}^{k} \theta_{n+j} \, \phi_n \, (x; \eta_j)$$

where $x \in [a, b]$ and $\{u_i\}_o^n$ is the ECT-system (3.1) with $w_o(x) \equiv 1$. Then the optimal design for estimating any $\theta_\ell (1 \leq \ell \leq n+k+1)$ is unique and is supported on the full set of extreme points of $W(x)$ obtained in Theorem 3.4, and the unique optimal design for estimating θ_o concentrates its entire mass at the point $x_1 = a$.

Proof:

Let

$$K(x, i) = \begin{cases} u_i(x) & i = 0,1,\ldots,n \\ \phi_n(x; s_i) & i = n+1,\ldots,n+k \end{cases}$$

where $s_{n+j} = \eta_j$; $j = 1,2,\ldots,k$ and let $f_i(x) = K(x; i)$; $i = 0,1,\ldots$ $n+k$. Then (3.6) and Theorem 2.2 pp. 514 of Karlin (1968) imply that

$$D_\nu(c_p) \quad \nu = 0, 1, 2, \ldots, n+k$$
$$p = 1, 2, \ldots, n+k$$

are all different from zero and have the same sign, where $D_\nu(c_p)$ has the same notation as in (1.6) with f_i's as defined and $s_i = x_{i+1}$ $i = 0,1,\ldots, n+k$ and $\{x_i\}_1^{n+k+1}$ is the set of points obtained in Theorem 3.4. Hence $c_p \in R$ for $p = 1,2,\ldots,n+k$. Hence the design $\xi = \xi_o$ concentrating mass

$$p_\nu = \frac{|D_\nu(c_p)|}{\sum_{\nu=0}^{n+k} |D_\nu(c_p)|}$$

at the points $x_{\nu+1}$; $\nu = 0,1,\ldots,n+k$ is the unique optimal design for estimating θ_j $(1 \leq j \leq n+k)$. Since $u_o(x_1) = 1$; $u_i(x_1) = 0$ $i = 1,2,$ $\ldots n$ and $\phi_n(x_1; \eta_j) = 0$; $j = 1,2,\ldots,k$; it is easily seen that the unique optimal design for estimating θ_o concentrates its entire mass at $x_1 = a$.

5. Basic Splines

Definition:

Let $a < \eta_1 < \eta_2 < \ldots < \eta_{n+k} < b$ be fixed. The functions $M_i(x;\ \eta_i,\eta_{i+1},\ldots,\eta_{i+n})$ $i = 1,2,\ldots,k$ (see Karlin (1968) pp. 522) are called the Basic Spline Functions and are given by

$$(3.7)\quad M_i(x;\eta_i,\eta_{i+1},\ldots,\eta_{i+n}) = \frac{\begin{vmatrix} u_o(\eta_i) & u_1(\eta_i) & \ldots, & u_{n-1}(\eta_i) & \phi_n(x;\eta_i) \\ u_o(\eta_{i+1}) & u_1(\eta_{i+1}) & \ldots, & u_{n-1}(\eta_{i+1}) & \phi_n(x;\eta_{i+1}) \\ \vdots & & & & \\ u_o(\eta_{i+n}) & u_1(\eta_{i+n}) & \ldots, & u_{n-1}(\eta_{i+n}) & \phi_n(x;\eta_{i+n}) \end{vmatrix}}{\begin{vmatrix} u_o(\eta_i) & u_1(\eta_i) & \ldots\ldots\ldots & u_n(\eta_i) \\ u_o(\eta_{i+1}) & u_1(\eta_{i+1}) & \ldots\ldots\ldots & u_n(\eta_{i+1}) \\ \vdots & & & \\ u_o(\eta_{i+n}) & u_1(\eta_{i+n}) & \ldots\ldots\ldots & u_n(\eta_{i+n}) \end{vmatrix}}$$

where $\{u_i\}_o^n$ and $\{\phi_n(x;\ \eta)\}$ are as given in (3.1) and (3.4).

If we take as the regression functions the basic spline functions M_i; $i = 1,2,\ldots,k$ and consider the regression model

$$E(y|x) = \sum_{i=1}^{k} \theta_i\ M_i(x;\eta_i,\eta_{i+1},\ldots,\eta_{i+n})$$

we can also characterize the optimal design for estimating θ_i $(i = 1,2,\ldots,k)$ as was done in the case of the Tchebycheffian spline regression function. Indeed one can obtain the unique oscillating polynomial W from Theorem 3.4, and using the strict total positivity result for the basic splines given in Lemma 4.2 pp. 524 of Karlin (1968), conclude that the optimal designs for estimating each θ_p, $p = 1,2,\ldots,k$ have the same support.

6. A Sufficient Condition for the Optimal Designs of Individual Regression Coefficients to have the Same Support

We have seen in the case of a Tchebycheffian spline regression function, that the optimal designs of each of the individual regression coefficients have the same support. The following theorem gives a sufficient condition for the optimal designs of individual regression coefficients to have the same support.

Theorem 3.6

Let the regression functions $\{f_i\}_o^n$ satisfy

(i) Continuous and linearly independent on $[a, b]$

(ii) $\{f_i\}_o^{n-1}$ is a Tchebycheff system on $[a, b]$

(iii) $\exists \{a_i\}_o^{n-1}$ with $\sum_o^{n-1} a_i^2 > 0 \ni \sum_o^{n-1} a_i f_i \equiv 1$

(iv) $a = s_o < s_1 < \ldots < s_n = b$ be the $(n+1)$ distinct points at which $f_n - h$ attains its norm with alternating signs where h is the best approximation of f_n with respect to the space spanned by $\{f_i\}_o^{n-1}$.

(v) $F\begin{pmatrix} i_1, i_2, \ldots, i_n \\ s_{j_1}, s_{j_2}, \ldots, s_{j_n} \end{pmatrix} > 0$

for every subset $(i_1, i_2, \ldots, i_n)$ and $(j_1, j_2, \ldots, j_n)$ of the set of integers $(0, 1, \ldots, n)$

(vi) $F\begin{pmatrix} 0, & 1, & \ldots, & n \\ s_0, & s_1, & \ldots, & s_n \end{pmatrix} \neq 0$

Then the c_p-optimal design for estimating θ_p in the regression model

$$E(y|x) = \sum_{j=0}^{n} \theta_j f_j$$

is unique and is supported on the full set of points $\{s_\nu\}_0^n$ for $p = 0,1,\ldots,n$.

Proof:

(ii) and (iii) guarantee the existence and uniqueness of an oscillatory polynomial

$$W(x) = \sum_{0}^{n} a_i^* f_i$$

with norm 1, attaining its norm at precisely (n+1) points $\{s_\nu\}_0^n$ with alternating signs with $s_0 = a$ and $s_n = b$

(vi) ensures that the polynomials

$$L_\nu(x); \quad \nu = 0,1,\ldots, n$$

associated with the set $\{s_\nu\}_0^n$ are well defined.

(v) ensures that $D_\nu(c_p) > 0$; $\nu = 0,1,\ldots,n$ for $p = 0,1,\ldots,n$. Hence each $c_p \in R$ and an appeal to Theorem 1.2 completes the proof.

Remark:

If $\{f_i\}_0^{n-1}$ is a Tchebycheff system on $(a,b]$ and $s_0 = a$, with $f(s_0) = c_0$, then optimal design for θ_0 is unique and concentrates its entire mass at s_0. Similarly if $\{f_i\}_0^{n-1}$ is a Tchebycheff system on (a,b) and $s_o = a$, and $s_n = b$ with $f(s_o) = c_o$ and $f(s_n) = c_n$, then optimal designs for estimating θ_o and θ_n are unique and are supported at the single points s_o and s_n respectively. As an example we have

$$f_i = \binom{n}{i} x^i (1-x)^{n-i}, \quad i = 0,1,\ldots,n$$

$$x \in [0, 1]$$

where $f(0) = c_o$, $f(1) = c_n$ and the kernel $K(x, i) = f_i(x)$ is Extended Totally Positive for $i = 0,1,2,\ldots,n$ and $0 < x < 1$. See Karlin (1968) pp. 287.

CHAPTER IV

OPTIMAL DESIGNS WITH TCHEBYCHEFF POLYNOMIALS OF THE FIRST KIND AS REGRESSION FUNCTIONS

1. Tchebycheff Polynomials of the First Kind

We consider the interval $[-1, 1]$. Tchebycheff Polynomials of the first kind denoted by $T_n(x)$ are defined as

$$T_n(x) = \cos (n \operatorname{arc} \cos x); \quad n = 0,1,\ldots$$

The polynomial $T_n(x)$ possesses $n+1$ extremal points in $[-1, 1]$ and the polynomial $T_n'(x)$ vanishes at each of the n extremal points in the interior of the interval.

The set of points at which $T_n(x)$ attains its norm with alternating signs, known as the Tchebycheff points are

$$s_\mu = -\cos \frac{\mu\pi}{n}, \quad \mu = 0,1,\ldots,n.$$

Explicitly

$$T_n(x) = \frac{1}{2} \sum_{\nu=0}^{[n/2]} (-1)^\nu \frac{n}{n-\nu} \binom{n-\nu}{\nu} (2x)^{n-2\nu} \quad n = 1,2,\ldots$$

$$T_o(x) \equiv 1$$

Moreover

$$T_n(-x) = (-1)^n T_n(x)$$

and finally we have the differential equation

$$(1-x^2) T_n''(x) - xT_n'(x) + n^2 T_n(x) = 0 \tag{4.1}$$

where $n \geq 2$.

The properties listed above are stated and proved in almost every textbook on linear approximation (see Meinardus (1967) pp. 31-33).

2. Some Results on Best Approximation with Tchebycheff Polynomials

Before proceeding to the problem of obtaining the optimal designs of individual regression coefficients, when Tchebycheff polynomials are taken as regression functions, we state and prove a few results on best approximation with these polynomials, which are needed.

Let θ_j be the regression coefficient associated with the regression function $f_j = T_j(x)$; $j = 0,1,\ldots,n$; $x \in [-1,1]$. The optimal design for estimating θ_j concentrates mass on the set of points $B = \{x: |T_j(x)-h(x)| = \sup_{-1 \leq x \leq 1} |T_j(x)-h(x)|\}$ where $h(x)$ is a best approximation of $T_j(x)$ with respect to the linear space spanned by $\{T_i(x)\}_{i=0, i \neq j}^{n}$. Thus the problem reduces to finding the polynomial $h(x)$, the set of points B, and the associated probabilities.

If $j = n$, then it is easily seen that $h(x)$ is $\equiv 0$ and the set B consists of the points

$$\{s_\mu = -\cos \mu\pi/n\}_0^n$$

and if p_μ denotes the probability associated with s_μ, then it is easily verified that $p_\mu = \frac{1}{2n}$ for $\mu = 0$ and n and $p_\mu = \frac{1}{n}$ for $\mu \neq 0$ and n. Thus $c_n \in R$ and the unique optimal design for estimating c_n concentrates mass on the full set of the Tchebycheff points. Similarly it is verified that $c_o \in S$ and has the same full support as c_n. As it belongs to the set S, from the Theorem of Studden (1968) (see Theorem 1.2, Chapter I) it is noted that the optimal design is not unique. It is interesting however to note that the optimal design for c_o having its support on $\{s_\mu\}$, assigns the same mass to s_μ as does the unique optimal design for c_n to s_μ. Before obtaining the optimal design for θ_j $(1 \leq j \leq n-1)$ we state and prove the following Theorems.

<u>Theorem 4.1</u>

Let V denote the linear space spanned by $\{T_{2r+1}(x)\}_1^{2k-1}$. Then the set

$$A = \{x: |T_1(x)-h(x)| = \sup_{-1 \leq x \leq 1} |T_1(x)-h(x)|\}$$

where $h(x)$ is a best approximation of $T_1(x)$ with respect to V, contains at least $2k$ points.

<u>Proof</u>:

If the set contains fewer than $2k$ points, then there exists an $h_1 \in V$ such that

$$h_1(x_i) = T_1(x_i) - h(x_i)$$

for all the extremal points (points of A) x_i. But then

$$8S \sin\theta \cdot \sin\phi = \cos(2k\theta-\phi)-\cos(2k\theta+\phi)+\cos(2k\phi-\theta)-\cos(2k\phi+\theta)$$

since

$$(2k\theta-\phi) - (2k\phi-\theta) = \pi$$

$$(2k\theta+\phi) + (2k\phi+\theta) = (2r+1)\pi$$

$$S = 0 \quad \text{for} \quad r = 1,2,\ldots,2k-1;\ k \geq 1.$$

Theorem 4.2

Let B denote the set of $2k$ points $\{x_\nu\}_1^k \cup \{-x_\nu\}_1^k$ where

$$x_\nu = \cos\frac{(2\nu-1)\pi}{4k+2}; \quad \nu = 1,2,\ldots,k.$$

There exists no function $h \in V$ such that either $h(x_\nu) > 0$ $\nu = 1,2,\ldots,k$ or $h(x_\nu) < 0$ $\nu = 1,2,\ldots,k$.

Proof:

Let $h(x) = \sum_{r=1}^{2k-1} \alpha_r T_{2r+1}(x) \in V$ Then

$$h(x_i) = \sum_{r=1}^{2k-1} \alpha_r \cos\frac{(2r+1)(2i-1)\pi}{4k+2} \qquad i = 1,2,\ldots,k$$

consider

$$A = \sum_{j=1}^{k} h(x_j) \cos\frac{(2j-1)\pi}{4k+2}$$

$$= \sum_{r=1}^{2k-1} \alpha_r \sum_{j=1}^{k} \cos\frac{(2r+1)(2j-1)\pi}{4k+2} \cdot \cos\frac{(2j-1)\pi}{4k+2}$$

$$= 0 \quad \text{(From Lemma 4.1)}$$

if $h(x_i) > 0$ for $i = 1,2,\ldots,k$ or < 0 clearly $A \neq 0$. This contradiction proves the theorem.

Theorem 4.3

There exists a unique $h(x) \in V$ such that

(i) $T_1(x_i) - h(x_i) = T_1(x_1) - h(x_1)$; $i = 2,3,\ldots,k$

(ii) $T_1'(x_i) - h'(x_i) = 0$; $i = 1,2,\ldots,k$

where $x_\nu = \cos \frac{(2\nu-1)\pi}{4k+2}$; $\nu = 1,2,\ldots,k$.

Proof:

Any $h(x) \in V$ is of the form

$$\sum_{r=1}^{2k-1} \alpha_r T_{2r+1}(x)$$

and hence is determined by (2k-1) parameters $\{\alpha_r\}_1^{2k-1}$. Conditions (i) and (ii) give (2k-1) linear equations in the (2k-1) unknown parameters $\{\alpha_r\}_1^{2k-1}$ and it is easily checked that the matrix of this system of linear equations is non-singular and as such has a unique solution .

Theorem 4.4

The unique $h(x)$ of Theorem 4.3 is the best approximation of $T_1(x)$ with respect to V.

Proof:

Let

$$\phi(x) = T_1(x) - h(x)$$

$$= x - \sum_{r=1}^{2k-1} \alpha_r T_{2r+1}(x)$$

$\phi(x)$ is an odd function and has the same value at each of the points

$$x_\nu = \cos \frac{(2\nu-1)\pi}{4k+2} ; \quad \nu = 1,2,\ldots,k.$$

Also $\phi'(x_\nu) = 0$; $\nu = 1,2,\ldots,k$. But

$$\phi'(x) = 1 - \sum_{r=1}^{2k-1} \alpha_r T'_{2r+1}(x)$$

$$\phi''(x) = - \sum_{r=1}^{2k-1} \alpha_r T''_{2r+1}(x)$$

Using the differential equation (4.1) we get

$$\phi''(x) = \frac{x}{1-x^2}[\phi'(x)-1] + \frac{1}{1-x^2} h_o(x)$$

where

$$h_o(x) = \sum_{r=1}^{2k-1} \beta_r T_{2r+1}(x) \in V;$$

$$\beta_r = (2r+1)^2 \alpha_r; \quad r = 1,2,\ldots,2k-1.$$

Hence

$$\phi''(x_\nu) = -\frac{x_\nu}{1-x_\nu^2} + \frac{1}{1-x_\nu^2} h_o(x_\nu).$$

As $h_o \in V$, from Theorem 4.2 we see that $h_o(x_\nu) \leq 0$ for some ν i.e. $\phi''(x_\nu) < 0$ for some ν; $1 \leq \nu \leq k$. Also $\phi(x_\nu) > 0$ $\nu = 1,2,\ldots,k$. Thus at one of the points x_ν, $\phi(x)$ has a local maximum $\phi(x)$ being a polynomial of degree 4k-1, and an odd function $\phi'(x)$ is an even polynomial of degree 4k-2 and we already note its (2k-1) zeros in $(0,1]$ which are precisely the points $\{x_\nu\}_1^k$ and $\{\xi_\nu\}_1^{k-1}$ where $\xi_\nu \in (x_\nu, x_{\nu+1})$. If we consider the closed interval $[x_\nu, x_{\nu+1}]$ where x_ν is the point at which $\phi(x)$ has a local maximum and as $\phi'(x)$ vanishes at only one point ξ_ν in the interior of $[x_\nu, x_{\nu+1}]$ it is clear that $\phi(x)$ attains its supremum on the interval $[x_\nu, x_{\nu+1}]$ at the point x_ν and hence also at $x_{\nu+1}$, and its minimum at ξ_ν. From this we conclude

$$\sup_{-1\leq x\leq 1} |\phi(x)| = |\phi(x_\nu)| = |\phi(-x_\nu)| \qquad \nu = 1,2,\ldots,k.$$

Thus the set of 2k points $\{x_\nu\}_1^k \cup \{-x_\nu\}_1^k$ are the extreme points of $\phi(x)$. As there exists no function in V which is positive at all the points x_ν; $\nu = 1,2,\ldots,k$ it follows that $h(x)$ is the best approximation of $T_1(x)$ with respect to V.

Remark:

$h(x)$ is also a best approximation of $T_1(x)$ with respect to the space spanned by $\{T_j(x)\}_{\substack{j=0 \\ j\neq 1}}^{n}$ where $n = 4k$ or $4k-1$.

Theorem 4.5

Let $C = \{x_\nu\}_0^k \cup \{-x_\nu\}_0^k$ where $x_\nu = \cos \nu\pi/2k+2$; $\nu = 0,1,2,\ldots,k$ The set C consisting of (2k+2) points is precisely the set of extremal points of $T_1(x)-h(x)$ where $h(x)$ is a best approximation of $T_1(x)$

with respect to the space spanned by $\{T_j(x)\}_{\substack{j=0\\ j\neq 1}}^{n}$, where $n = 4k+1$ or $4k+2$. Proof of Theorem 4.5 is analogous to the proofs of Theorems 4.2 to 4.4 and hence is omitted.

3. Optimal Design for θ_1

With the establishment of the best approximation of $T_1(x)$ and its extreme points, we are now in a position to state formally the theorem concerning the optimal design for estimating θ_1.

Theorem 4.6

Optimal design for estimating θ_1 concentrates mass on the set of 2k points $\{x_\nu\}_1^k \cup \{-x_\nu\}_1^k$, where

$$x_\nu = \cos \frac{(2\nu-1)\pi}{4k+2}$$

when $n = 4k-1$ or $4k$ and concentrates mass on the set of $(2k+2)$ points $\{x_\nu\}_0^k \cup \{-x_\nu\}_0^k$ where

$$x_\nu = \cos \frac{\nu\pi}{2k+2}$$

if $n = 4k+1$ or $4k+2$.

4. Optimal Design for θ_j $(1 < j < n)$

Given a j such that $1 < j < n$ there exists a unique integer $p \geq 0$ such that either

(i) $(4p-1)\, j \leq n < (4p+1)j$ or

(ii) $(4p+1)j \leq n < (4p+3)j$.

Let the integer j, corresponding to θ_j be such that (i) holds. Clearly $p > 0$. Let $h(x) = \sum_{r=1}^{2p-1} \alpha_r T_{2r+1}$ be the best approximation of $T_1(x)$ with respect to the space spanned by $\{T_i(x)\}_{i=0;\ i\neq 1}^{4p-1}$. Then from Theorem 4.4 the extreme points of $T_1(x)-h(x)$ consists of 2p points $\{x_\nu\}_1^p \cup \{-x_\nu\}_1^p$ where $x_\nu = \cos \frac{(2\nu-1)\pi}{4p+2}$. Since

$$\sup_{-1\leq x\leq 1} |T_j(x) - \sum_{r=1}^{2p-1} \alpha_r T_{(2r+1)j}(x)| = \sup_{-1\leq x\leq 1} |T_1(x) - \sum_{r=1}^{2p-1} \alpha_r T_{2r+1}(x)|$$

the extreme points of

$$\phi(x) = T_j(x) - \sum_{r=1}^{2p-1} \alpha_r T_{(2r+1)j}(x)$$

are given by the set of 2jp points

$$D = \begin{cases} \cos \frac{1}{j}[r\pi - \frac{\nu\pi}{4p+2}]; & \nu = 1,3,\ldots,(2p-1) \quad r = 0,1,\ldots,j \\ \\ \cos \frac{1}{j}[r\pi + \frac{\nu\pi}{4p+2}]; & \nu = 1,3,\ldots,(2p-1) \quad r = 1,2,\ldots,(j-1) \end{cases}$$

As we proved in the case of $T_1(x)$, it can be shown that there exists no function belonging to the space spanned by $\{T_i(x)\}_{i=0,\ i\neq j}^{n}$, which has the same signs as $\phi(x)$ on the set D. Hence by Kolmogoroff's criterion $\sum_{r=1}^{2p-1} \alpha_r T_{(2r+1)j}(x)$ is a best approximation of $T_j(x)$ and thus the optimal design for θ_j has for its support the set D.

If the integer j is such that $(4p+1)j \leq n < (4p+3)j$ then either $p = 0$ or $p > 0$ if $p = 0$, i.e. $j < n < 3j$ then the best approximation of $T_j(x)$ is the function which is identically zero and hence the support of the optimal design for θ_j consists of the (j+1) Tchebycheff points $\{-\cos \frac{\nu\pi}{j}\}_0^j$.

If $p > 0$, it can be shown as in the case (i) that the support of the optimal design for θ_j is on the set of $(2p+1)\ j+1$ points

$$E = \begin{cases} \cos \frac{1}{j}[r\pi - \frac{\nu\pi}{2p+2}]; & \nu = 0,1,\ldots,p \quad r = 0,1,\ldots,j \\ \\ \cos \frac{1}{j}[r\pi + \frac{\nu\pi}{2p+2}]; & \nu = 0,1,\ldots,p \quad r = 1,2,\ldots,(j-1) \end{cases}$$

Finally we prove that the optimal design for θ_j is supported on the full set of points. Details of proof are given for the case $j=1$ and $n = 4k-1$; the proof for the general case is exactly the same.

Theorem 4.6 and Elfving's Theorem guarantee the existence of a solution $\{\epsilon_\nu\ p_\nu\}$ for the system of equations

$$\sum_{\nu=1}^{2k} \epsilon_\nu\ p_\nu\ f(x_\nu) = \beta c_1$$

where $\epsilon_\nu = \pm 1$; $p_\nu \geq 0$; $\sum_{\nu=1}^{2k} p_\nu = 1$ and $f(x) = (T_o, T_1, \ldots, T_{4k-1})$. Let p_ν be the weight associated with

$$x_\nu = \cos \frac{(2\nu-1)\pi}{4k+2}\ ; \quad \nu = 1,2,\ldots,k.$$

From symmetry, the system of equations reduces to

$$
\begin{aligned}
p_1 \cos \theta_1 + p_2 \cos \theta_2 &+ \cdots + p_k \cos \theta_k &= \beta/2 \\
p_1 \cos 3\theta_1 + p_2 \cos 3\theta_2 &+ \cdots + p_k \cos 3\theta_k &= 0 \\
&\cdots\cdots\cdots\cdots\cdots\cdots & \\
p_1 \cos(2k-1)\theta_1 + p_2 \cos(2k-1)\theta_2 &+ \cdots + p_k \cos(2k-1)\theta_k &= 0
\end{aligned}
$$

where $\theta_\nu = \frac{(2\nu-1)\pi}{4k+2}$; $\nu = 1,2,\ldots,k$. The coefficient matrix of the above system being a Vandermonde type, it is easily seen that $p_\nu \neq 0$; $\nu = 1,2,\ldots,k$.

We can thus summarize the result in the following theorem.

Theorem 4.7

If $j < n < 3j$, then the optimal design for estimating θ_j has for its support the full set $\{-\cos \frac{\nu\pi}{j}\}_0^j$.

If $(4p-1)\, j \leq n < (4p+1)j$, optimal design for θ_j has for its support the full set D.

If $(4p+1)j \leq n < (4p+3)j$, optimal design for θ_j has for its support the full set E.

~~examples~~

CHAPTER V

MINIMAX DESIGNS

1. Introduction

In this chapter we will confine to polynomial regression on the interval $[-1, 1]$, so that our regression functions are $f_i = x^i$; $i = 0,1,\ldots,n$. Elfving (1959) defines a design ξ^* to be minimax s.p. (with respect to the single parameters) when it minimizes $\max_{0\leq i\leq n} V(c_i,\xi)$. We obtain explicitly the design ξ^* when $n \leq 12$ $(n \neq 11)$. We are not able to obtain a general solution of this problem. But the results obtained indicate the direction in which one could look for a possible general solution.

From the results obtained by Studden (1968) we see that the optimal design for estimating θ_p $(p \neq 0$; $n-p$ even), has for its support the full set of points $\{-\cos \frac{\nu\pi}{n}\}_0^n$ with respective masses $\{|D_\nu(c_p)|/\sum_{\nu=0}^{n}|D_\nu(c_p)|\}_0^n$ and when $n-p$ is odd, the support is on the full set of points $\{-\cos \frac{\nu\pi}{n-1}\}_0^{n-1}$. Thus we know explicitly the optimal design for estimating any of the individual regression coefficients.

2. Characterization of the Minimax Design

The following simple theorem characterizes the minimax design.

Theorem 5.1

The unique optimal design ξ_k^* for estimating θ_k is minimax s.p. if

$$\max_{0 \leq i \leq n} V(c_i, \xi_k^*) = \inf_{\xi} V(c_k, \xi).$$

Proof:

$$\max_i V(c_i, \xi_k^*) \geq \inf_{\xi} \max_i V(c_i, \xi) \geq \max_i \inf_{\xi} V(c_i, \xi) \geq \inf_{\xi} V(c_k, \xi)$$

Hence if the two end terms are equal, i.e. if

$$\max_i V(c_i, \xi_k^*) = \inf_{\xi} V(c_k, \xi)$$

then we have

$$\inf_{\xi} \max_i V(c_i, \xi) = V(c_k, \xi_k^*)$$

i.e. ξ_k^* is minimax s.p.

3. Minimax Designs for $n \leq 12$

When $n=1$, it is easily seen that

$$M^{-1}(\xi_1^*) = M(\xi_1^*) = I_2$$

so that $\max_i V(c_i, \xi_1^*) = 1 = \inf_{\xi} V(c_1, \xi)$. Hence ξ_1^* is minimax s.p. $n=2$;
A direct computation yields

$$M^{-1}(\xi_2^*) = \begin{pmatrix} 2 & 0 & -2 \\ 0 & 2 & 0 \\ -2 & 0 & 4 \end{pmatrix}$$

$$[T_1(x_i)-h(x_i)]\ h_1(x_i) = |T_1(x_i)-h(x_i)|^2 > 0$$

and by Kolmogoroff's criterion (see Meinardus (1967) pp. 15) $h(x)$ cannot be a best approximation. This contradiction proves the theorem.

Lemma 4.1

$$S = \sum_{j=1}^{k} \cos \frac{(2r+1)(2j-1)\pi}{4k+2} \cdot \cos \frac{(2j-1)\pi}{4k+2} = 0$$

for $r = 1,2,\ldots,2k-1;\ k \geq 1.$

Proof:

$$2S = \sum_{j=1}^{k} \cos\ (2j-1)\ \theta + \sum_{j=1}^{k} \cos\ (2j-1)\ \phi$$

where

$$\theta = \frac{(2r+2)}{4k+2} \cdot \pi; \quad \phi = \frac{2r\pi}{4k+2}$$

so that

$$\theta = \phi + \frac{2\pi}{4k+2} .$$

But

$$\sum_{j=1}^{k} \cos\ (2j-1)\ \theta = \sin 2k\theta/2 \sin \theta$$

$$\sum_{j=1}^{k} \cos\ (2j-1)\ \phi = \sin 2k\phi/2 \sin \phi$$

Hence

Hence

$$\max_i V(c_i, \xi_2^*) = 4 = \inf_\xi V(c_2, \xi)$$

ξ_2^* is minimax s.p. n=3; In this case also one can obtain easily

$$M^{-1}(\xi_3^*) = \begin{pmatrix} 3 & 0 & -4 & 0 \\ 0 & 11 & 0 & -12 \\ -4 & 0 & 8 & 0 \\ 0 & -12 & 0 & 16 \end{pmatrix}$$

OK

$$M = \begin{matrix} 1 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 9/24 \\ 1/2 & 0 & 9/24 & 0 \\ 0 & 9/24 & 0 & 11/32 \end{matrix}$$

Hence ξ_3^* is minimax s.p.

The above three cases indicate that in the search for minimax design one should look into the $\max_i \inf_\xi V(c_i, \xi)$ and consider the design ξ_k^* if this max is attained for $i = k$. However if we consider the case n=4 we readily see that $\max_i \inf_\xi V(c_i, \xi)$ is attained for i=2 and 4. Thus neither ξ_2^* nor ξ_4^* will be minimax s.p. From what has been stated earlier we see that ξ_2^* concentrates mass on $\{-\cos \nu\pi/4\}_0^4$ with respective probabilities $\frac{1}{16}, \frac{4}{16}, \frac{6}{16}, \frac{4}{16}$ and $\frac{1}{16}$, and ξ_4^* concentrates mass on the same set of points with respective probabilities $\frac{1}{8}, \frac{2}{8}, \frac{2}{8}, \frac{2}{8}, \frac{1}{8}$. The actual minimax s.p design in this case is $\frac{1}{2}(\xi_2^* + \xi_4^*)$ which concentrates mass on the set $\{-\cos \frac{\nu\pi}{4}\}_0^4$ with probabilities $\frac{3}{32}, \frac{8}{32}, \frac{10}{32}, \frac{8}{32}, \frac{3}{32}$.

The case n=11 is similar to the case n=4. $\max_i \inf_\xi V(c_i, \xi)$ is attained for i=7 and 9. But it has not been possible to work out the minimax s.p. design in this case, as the computations involved became too tedious.

For values of n between 5 and 12, excluding n=11, the minimax s.p. design was obtained by inverting the matrix $M(\xi)$ using a computer and the solution obtained is presented below.

$n = 5; \quad k = 3$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos \nu\pi/5\}_0^5$ with respective probabilities 0.060, 0.176, 0.264, 0.264, 0.176, and 0.060.

$n = 6; \quad k = 4$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos \nu\pi/6\}_0^6$ with respective probabilities 0.056, 0.139, 0.194, 0.222, 0.194, 0.139, and 0.056.

$n = 7; \quad k = 5$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos \nu\pi/7\}_0^7$ with respective probabilities, 0.051, 0.117, 0.152, 0.180, 0.180, 0.152, 0.117, and 0.051.

$n = 8; \quad k = 6$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos \nu\pi/8\}_0^8$ with respective probabilities, 0.047, 0.103, 0.125, 0.147, 0.156, 0.147, 0.125, 0.103, 0.047.

$n = 9; \quad k = 7$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos \nu\pi/9\}_0^9$ with respective probabilities, 0.044, 0.092, 0.106, 0.124, 0.134, 0.134, 0.124, 0.106, 0.092, 0.044.

$n = 10; \quad k = 8$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos\ \nu\pi/10\}_o^{10}$ with respective probabilities, 0.040, 0,084, 0.094, 0.106, 0.116, 0.120, 0.116, 0.106, 0.094, 0.084, 0.040.

$n = 12; \quad k = 8$

ξ_k^* is minimax s.p. concentrating mass on $\{-\cos\ \nu\pi/12\}_o^{12}$ with respective probabilities, 0.028, 0.058, 0.066, 0.080, 0.096, 0.111, 0.117, 0.111, 0.096, 0.080, 0.066, 0.058 and 0.028.

LIST OF REFERENCES

1. Elfving, G. (1952). Optimum allocation in linear regression theory. Ann. Math. Stat. 23, 255-262.

2. Elfving, G. (1959). Design of linear experiments. Probability and Statistics, Harold Cramer Volume, edited by Ulf Grenander, John Wiley and Sons, New York. 58-74.

3. Hoel, P.G. and Levine, A. (1964). Optimal spacing and weighing in polynomial prediction. Ann. Math. Stat. 35, 1553-1560.

4. Karlin, S. and Studden, W.J. (1966a). Tchebycheff Systems: with Applications in Analysis and Statistics. Interscience, New York.

5. Karlin, S. and Studden, W.J. (1966). Optimal experimental designs. Ann. Math. Stat. 37, 783-815.

6. Karlin, S. and Schumaker, L.L. (1967). The fundamental theorem of algebra for Tchebycheffian monosplines. Jour. d'Analyse Math. 20, 233-270.

7. Karlin, S. (1968). Total Positivity, Vol. I. Stanford University Press, California.

8. Kiefer, J. and Wolfowitz, J. (1959). Optimum designs in regression problems. Ann. Math. Stat. 30, 271-294.

9. Kiefer, J. and Wolfowitz, J. (1965). On a theorem of Hoel and Levine on extrapolation. Ann. Math. Stat. 36, 1627-1655.

10. Meinardus, G. (1967). Approximation of functions: Theory and Numerical Methods. Translated by Larry L. Schumaker. Springer-Verlag, New York.

11. Schoenberg, I.J. (1964). On interpolation by spline functions and its minimal properties. International Series Numerical Math. 5, 109-129.

12. Schumaker, L.L. (1967a). Uniform approximation by Tchebycheffian spline functions-I, Fixed knots. Mathematics Research Centre Summary Report No. 768, Madison, Wisconsin.

13. Schumaker, L.L. (1967b). Uniform approximation by Tchebycheffian spline functions-II, Free knots. Mathematics Research Centre Summary Report No. 810. Madison, Wisconsin.

14. Studden, W.J. (1968). Optimal designs on Tchebycheff points. <u>Ann. Math. Stat.</u> 39, 1435-1447.

15. Studden, W.J. and VanArman, D.J. (1968). Admissible designs for polynomial spline regression. <u>Ann. Math. Stat.</u> 40, (to appear in Oct. 1969).