SOME ROBUST TYPE D-OPTIMAL DESIGNS IN POLYNOMIAL REGRESSION

by

W. J. Studden[1]
Purdue University, W. Lafayette, IN.

Mimeograph Series #81-26

Department of Statistics
Purdue University

July 1981
(Revised April 1982)

[1]This research was supported by NSF Grant No. 7901707.

SOME ROBUST TYPE D-OPTMAL DESIGNS IN POLYNOMIAL REGRESSION

by W. J. Studden[1]
Department of Statistics
Purdue University
W. Lafayette, IN. 47907

ABSTRACT

Consider a polynomial regression situation on an interval. A robustness type formulation of Stigler's is considered. A technique involving canonical moments is considered and some explicit solutions are given.

AMS 1970 Subject classification: 62K05, 62J05

Key Words: polynomial regression, optimal design, canonical moments

[1]This research was supported by NSF Grant No. 7901707.

1. Introduction. Consider a polynomial regression situation on [0,1]. For each x or "level" in [0,1] an experiment can be performed whose outcome is a random variable $y(x)$ with mean value $\sum_{i=0}^{m} \beta_i x^i$ and variance σ^2, independent of x. The parameters β_i, $i=0,1,\ldots,m$ and σ^2 are unknown. An experimental design is a probability measure ξ on [0,1]. If N observations are to be taken and ξ concentrates mass ξ_i at the points x_i, $i=1,2,\ldots,\ell$ and $N\xi_i = n_i$ are integers, the experimenter takes N uncorrelated observations, n_i at each x_i, $i=1,2,\ldots,\ell$. The covariance matrix of the least squares estimates of the parameters β_i is then given by $(\sigma^2/N)M^{-1}(\xi)$ where $M(\xi)$ is the information matrix of the design with elements $m_{ij} = \int_0^1 x^{i+j} d\xi(x)$. For an arbitrary probability measure or design some approximation would be needed in applications.

Various criteria have been used for determining a good design ξ. Typically one tries to minimize some functional $\psi(M(\xi))$ of the information matrix $M(\xi)$. Examples are $\psi_1(M(\xi)) = |M^{-1}(\xi)|$, $\psi_2(M(\xi)) = \sup_x d(x,\xi)$ where $d(x,\xi) = f'(x)M^{-1}(\xi)f(x)$, $f'(x) = (1,x,\ldots,x^m)$ or $\psi_3(M(\xi)) = c'M^{-1}(\xi)c$ for some c, etc. The solution to the first problem is called the D-optimal design and the second is called the G-optimal design. These are known to be equivalent, see Kiefer and Wolfowitz (1960).

In situations such as these, the model or the degree of the polynomial, is assumed known. Numerous papers have been devoted to this problem. Others have considered designing the experiment to include bias components. The object of the present paper is to describe a technique for solving a problem formulated in a paper by Stigler (1971). Some familiarity with Stigler's paper and the papers by Studden [1980][1981] is useful.

The situation is described generally as follows. Let $f'(x) = (1, x, \ldots, x^m) = (f_1'(x), f_2'(x))$ where $f_1'(x) = (1, x, \ldots, x^r)$, $f_2'(x) = (x^{r+1}, \ldots, x^m)$ and decompose $M(\xi)$ similarly as

$$M(\xi) = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

Here M_{11} has size r and M_{22} has size $s = m - r$. The model is thought to be of degree r but possibly the coefficients of the s higher powers $x^{r+1}, \ldots, x^m$ are not zero. We would then like to determine a design ξ which is D-optimal if the model is degree r, i.e. maximizes the determinant $|M_{11}(\xi)|$ subject however to the condition that there is some protection in being able to determine whether the coefficients $\beta_{r+1}, \beta_{r+2}, \ldots, \beta_m$ are zero. The covariance matrix for the LSE of $\beta_{r+1}, \ldots, \beta_m$ is proportional to the inverse of the matrix

$$\Sigma_s(\xi) = M_{22}(\xi) - M_{21}(\xi) M_{11}^{-1}(\xi) M_{12}(\xi). \tag{1.1}$$

We therefore consider the D_{rm}-problem; which is to maximize $|M_{11}(\xi)|$ subject to the condition $|\Sigma_s(\xi)| \geq C$. A solution to this problem will be called a D_{rm}-optimal design. This should provide a design which is good in the sense of being D-optimal if the model is of degree r and allows some protection in being able to test whether $\beta_{r+1}, \ldots, \beta_m$ are zero. The measure of protection is given somewhat indirectly by the constant C.

Stigler has proposed considering a G-optimal version of the above problem and, at the end of his paper, of replacing the condition on $|\Sigma_s|$ by a sequence of constraints, one for each degree above r. The latter problem could be considered using our techniques. The G-optimal version seems to be somewhat more complicated than the D-optimal problem. Only the D_{rm}-problem is considered here.

One of the quantities used to measure the efficiency of a design is the D-efficiency given by

$$(1.2) \qquad e(\xi) = \left(\frac{|M(\xi)|}{\sup_{\eta}|M(\eta)|} \right)^{\frac{1}{m+1}}$$

Notice that the size of $M(\xi)$ is $m + 1$. Using (1.2) we rewrite the constant C as

$$(1.3) \qquad C = \rho^{s} \max_{\eta} |\Sigma_s(\eta)|$$

The D_{rm}-optimal design will have $|\Sigma_s(\xi)| = C$ so that

$$\rho = \left(\frac{\Sigma_s(\xi)}{\max_{\eta}|\Sigma_s(\eta)|} \right)^{1/s}$$

Thus ρ measures the D-efficiency in estimating the coefficients $\beta_{r+1},\ldots,\beta_m$. The case $\rho = 1$ corresponds to obtaining maximum information concerning $\beta_{r+1},\ldots,\beta_m$. The other extreme $\rho = 0$ gives rise to the D-optimal design for rth degree regression. The case $\rho = 1$ was studied in Studden (1980) with the aid of canonical moments. The purpose of the present paper is to illustrate the further use of canonical moments in handling the general D_{rm}-problem.

For a given design ξ the matrix $M(\xi)$ has entries $m_{ij} = \int x^{i+j} d\xi(x)$. For convenience let

$$(1.4) \qquad c_k = \int_0^1 x^k d\xi(x), \quad k=0,1,2,\ldots$$

For a given, fixed, set of moments $c_0,c_1,\ldots,c_{i-1}$ let c_i^+ denote the maximum of the ith moment $\int_0^1 x^i d\mu(x)$ over the set of measures μ having the given moments $c_0, c_1, \ldots, c_{i-1}$. Similarly let c_i^- denote the corresponding minimum. The canonical moments are defined by

(1.5)
$$p_i = \frac{c_i - c_i^-}{c_i^+ - c_i^-} \quad .$$

The p_i values are left undefined if $c_i^- = c_i^+$. As a simple example consider the first two canonical moments p_1, p_2 corresponding to c_1, c_2. The value of p_1 is simple c_1 since given $\int d\xi = 1$ the first moment can range over [0,1]. The set of all possible moments (c_1, c_2) is generated by taking the convex hull of the curve (x, x^2) for $0 \leq x \leq 1$. Thus for given c_1 the second moment c_2 is bounded between $c_2^- = c_1^2$ and $c_2^+ = c_1$. In this case

$$p_2 = \frac{c_2 - c_1^2}{c_1(1-c_1)} \quad .$$

It should be noted that measures ξ are on the unit interval [0,1]. The canonical moments for measures on an arbitrary interval are defined in precisely the same way. The canonical moments are invariant with respect to linear transformations and all the results and designs given here can be easily transformed to arbitrary intervals. The useful property of the canonical moments is that the p_i values range "independently" over the entire interval [0,1]. Problems defined in terms of the canonical moments have some chance of easy solution especially if the involved expressions are "relatively" simple. There are then known methods (which are described below) for converting back and forth between the ordinary moments c_i, the canonical moments p_i and the design ξ.

We give as a simple example the quadratic situation m = 2 which was used for illustration by Stigler (1971). More general cases are considered below. Suppose that m = 2 and r = 1. The model is thought to be linear but some protection against quadratic terms is required. Using the value of p_1 and p_2 just calculated from c_1 and c_2 we find that

$$(1.6)\qquad |M_{11}(\xi)| = c_2 - c_1^2 = p_1q_1p_2, \quad (q_i=1-p_i).$$

We show below that

$$(1.7)\qquad |M(\xi)| = (p_1q_1p_2)^2(q_2p_3q_3p_4).$$

Observe first that the D-optimal designs can be obtained very rapidly in terms of the p_i. For linear regression $|M_{11}(\xi)|$ is maximized for $p_1 = q_1 = 1/2$ and $p_2 = 1$ while the quadratic D-optimal design maximizing $|M(\xi)|$ in (1.7) has $p_1 = p_3 = 1/2$ $p_2 = 2/3$ and $p_4 = 1$. We mention in passing that symmetry of ξ about $x = 1/2$ is related to the odd canonical moments being 1/2.

The D_{12}-problem now reduces to maximizing $p_1q_1p_2$ subject to the condition that

$$(1.8)\qquad |\Sigma(\xi)| = \frac{|M|}{|M_{11}|} = p_1q_1p_2q_2p_3q_3p_4$$

$$\geq \rho \max_{\eta}|\Sigma(\eta)|$$

$$> \rho\, 2^{-6} .$$

It is evident that the solution involves $p_1 = p_3 = 1/2$, $p_4 = 1$ and we maximize p_2 subject to $p_2q_2 \geq \rho/4$ which gives

$$p_2 = \frac{1+\sqrt{1-\rho}}{2}$$

The canonical moment sequence $(1/2, p_2, 1/2, 1)$ is then converted back to the c_i, $i=0,1,2,3,4$ to get $M(\xi)$ and to the design ξ. It will be indicated shortly that the above sequence corresponds to a design on the points 0, 1/2, 1 with corresponding weights α, $1 - 2\alpha$, α where

$$\alpha = \frac{p_2}{2} = \frac{1+\sqrt{1-\rho}}{4} .$$

This is precisely the design encountered by Stigler (1971) on page 315. His value of C is related to our ρ by $4 = \rho C$.

The general formulation of the D_{rm}-design problem becomes quite simple once the value of $|M(\xi)|$ (and hence $|M_{11}(\xi)|$) is given. The expression below is taken from Skibinsky (1969). The value is also given in Wall (1948), Brezinski (1980) and is "well-known" in the theory of orthogonal polynomials and continued fractions. The results originate with Stieltjes.

If $\zeta_i = q_{i-1}p_i$, $i=1,2,\ldots$, where $q_0 = 1$ and $q_i = 1 - p_i$ then the value for $|M(\xi)|$ is given by

(1.9) $$|M(\xi)| = \prod_{i=1}^{m} (\zeta_{2i-1}\zeta_{2i})^{m+1-i} .$$

The value for $|M_{11}(\xi)|$ is, of course, just (1.9) with m replaced by r. To obtain $|\Sigma_s(\xi)|$, use is made of the fact that

$$|\Sigma_s(\xi)| = \frac{|M(\xi)|}{|M_{11}(\xi)|} .$$

The general D_{rm}-optimal design problem is thus relatively easy to state in terms of the canonical moments p_i.

The remaining sections are outlined briefly as follows. In section 2 we state a number of results allowing us to convert from the p_i to the c_i and the design ξ. These are taken from Skibinsky (1969) or the other sources listed above for (1.9). Proofs of most of these results will not be given here. The explicit solutions for the D_{1m}-problem and some calculations for the D_{2m}-problem are given in Section 3. In both cases some D-efficiencies as defined in (1.2) are calculated. It will be seen that the solution to the D_{12}-problem gives rise to an immediate solution to the

D_{1m}-problem. Similarly the D_{23}-problem gives the solution to the D_{2m}-problem, etc. Moreover the solution to D_{23} involves the solution to D_{12}, etc. so that the successive problems, as expected, are of more complexity.

2. Converting from p_i to c_i and ξ. There is a considerable amount of literature concerning the relationship between the sequences $\{p_i\}$ and $\{c_i\}$ and the design ξ. We will state here only those results which are pertinent to the D_{rm}-design problem.

In the D_{rm}-design problem the p_i values appear only through the determinants $|M|$ and $|M_{11}|$. These are given in (1.9) so there is no need to express the p_i values in terms of c_i or ξ. (The p_i values can be expressed as ratios of Hankel determinants involving the moments c_i. The $\zeta_i = q_{i-1}p_i$ values occur as the coefficients in the continued fraction expansion of the Stieltjes transform of the measure ξ and they occur in the three terms recursive relations for certain orthogonal polynomials related to ξ.)

The direction useful here is in going from the canonical moments p_i to the ordinary moments c_i and the design ξ. The c_i values are needed in calculating M or M^{-1} which leads to the covariance structure of the estimates of $\beta_1,\beta_2,\ldots,\beta_m$. The measure ξ is, of course, the design and is the principal object of study.

To go from the p_i to the c_i we have the following.

Lemma 2.1. If $S_{0j} = 1$, $j=0,1,\ldots$ and

$$S_{ij} = \sum_{k=i}^{j} \zeta_{k-i+1} S_{i-1k}, \quad i \geq j \tag{2.1}$$

then $c_m = S_{mm}$.

The first few moments are

$$c_1 = p_1 = \zeta_1$$

$$c_2 = p_1(p_1+q_1p_2) = \zeta_1(\zeta_1+\zeta_2)$$

$$c_3 = \zeta_1[\zeta_1(\zeta_1+\zeta_2)+\zeta_2(\zeta_1+\zeta_2+\zeta_3)] .$$

To describe the design ξ we need the support and the corresponding weights.

Lemma 2.2. If $\zeta_1,\zeta_2,\ldots,\zeta_k$ are not zero and $\zeta_{k+1} = 0$ then the corresponding design ξ concentrates its mass on the zeros of the polynomial

$$(2.2) \qquad D_k(x) = \begin{vmatrix} x & -1 & 0 & & & 0 \\ -\zeta_1 & 1 & -1 & & & \\ 0 & -\zeta_2 & x & & & \\ & & & \cdot & & \\ & & & & \cdot & \\ & & & & \cdot & -1 \\ 0 & & & & -\zeta_k & \tau(x) \end{vmatrix}$$

where $\tau(x) = x$ or 1 according as k is even or odd.

Note that $D_k(x)$ is a tri-diagonal matrix and is roughly of degree $k/2$. The expression for $D_k(x)$ can be expanded by the last row to show that

$$(2.3) \qquad D_k(x) = \tau(x)D_{k-1}(x) - \zeta_k D_{k-2}(x).$$

The case $k = 2m$, where $p_{2m} = 1$, $\zeta_{2m} = q_{2m-1}$ and $\zeta_{2m+1} = 0$ will be of particular interest to us. In this case

$$(2.4) \qquad D_{2m}(x) = x^{m+1} + \sum_{i=1}^{m-1} (-1)^i a_i x^{m-i}$$

where $a_1 = \zeta_1 + \zeta_2 + \ldots + \zeta_m$, $a_2 = \sum_{i_1<i_2} \zeta_{i_1}\zeta_{i_2+1}$ and in general a_j is a sum of products of j terms; the sum being over the terms with all subscripts at least two units apart. For example

$$D_4(x) = x^3 - (\zeta_1+\zeta_2+\zeta_3+\zeta_4)x^2 + (\zeta_1\zeta_3+\zeta_1\zeta_4+\zeta_2\zeta_4)x \, .$$

Further simplification occurs if $p_{2i+1} = 1/2$ and also from noting that $x = 1$, as well as $x = 0$ is a root of $D_{2m}(x) = 0$.

The weights on the various points can be calculated in a number of ways. For our purposes we shall resort to simply calculating the c_i and setting up the linear equations involving the weights and the ordinary moments. That is, if $x_0, x_1, \ldots, x_k$ are the required support points then

(2.5) $$c_j = \sum_{i=0}^{k} w_i x_i^j, \quad j=0,1,\ldots,k.$$

These equations are solved for $w_0, \ldots, w_k$. In all our cases $p_{2i+1} = 1/2$ and the solution is symmetric.

<u>Lemma 2.3.</u> (a) The design corresponding to $(1/2, p_2, 1/2, 1)$ concentrates mass α, $1 - 2\alpha$, α on the points 0, 1/2 1 respectively where $\alpha = p_2/2$.

(b) The design corresponding to $(1/2, p_2, 1/2, p_4, 1/2, 1)$ concentrates mass α, $1/2 - \alpha$, $1/2 - \alpha$, α on the points 0, $1 - t$, t, 1 respectively where $\alpha = p_2p_4/2(q_2+p_2p_4)$, $t = (1+\sqrt{p_2q_4})/2$.

<u>Proof.</u> The proof in each case follows from Lemma 2.2 and (2.5). In each case the expression (2.4) can be used to show that the points of support are correct. To get the weights we use (2.5) and the symmetry. For example in the first case the symmetric weights will give $c_1 = p_1 = 1/2$. The second moment is $c_2 = p_1(p_1+q_1p_2) = (1+p_2)/4$ so that using (2.5) we require

$$\frac{(1+p_2)}{4} = (1/2)^2(1-2\alpha) + \alpha$$

which gives $\alpha = p_2/2$. The situation in part (b) is similar.

3. The cases r=1 and r=2 and general m. We first discuss $r = 1$, where a simple linear regression model, $\beta_0 + \beta_1 x$, is being considered and some protection is desired for the terms $\beta_2, \beta_3, \ldots, \beta_m$.

Theorem 3.1. For $r = 1$ and general m the D_{1m}-optimal design has canonical moments

$$p_{2i-1} = 1/2, \quad i=1,2,\ldots,m$$

$$p_2 = \frac{1+\sqrt{1-\rho}}{2} \tag{3.1}$$

$$p_{2i} = \frac{m-i+1}{2m-2i+1}, \quad i=2,3,\ldots,m-1$$

$$p_{2m} = 1.$$

Corollary 3.1. The D_{12}-optimal design ξ_{12} has mass α_1, $1 - 2\alpha_1$, α_1 on the points 0, 1/2, 1 respectively where

$$\alpha_1 = p_2/2 = (1+\sqrt{1-\rho})/4 .$$

Corollary 3.2. The D_{13}-optimal design ξ_{13} has mass α_2, $1/2 - \alpha_2$, $1/2 - \alpha_2$, α_2 on the four points 0, $1 - t$, t, 1 where

$$t = (1+\sqrt{p_2 q_4})/2, \quad \alpha_2 = 2p_2/(2p_2+3q_3),$$

$$p_2 = (1+\sqrt{1-\rho})/2, \quad q_4 = 1/3.$$

We should reemphasize that the definition of ρ in equation (1.3) makes use of the number s of extra parameters to be guarded against. In the present case $r = 1$ and $s = m - 1$. With this definition some of the lower order canonical moments and efficiences considered below are independent of m. Thus, ρ (or $1 - \rho$) measures how much is taken away from the D-optimality of the basic model with r parameters.

The proof of Theorem 3.1 makes use of (1.9) and the problem easily reduces to maximizing p_2 subject to the condition that $p_2q_2 \geq \rho/4$. In the process we use that fact that p^iq^j is maximized for $p = i/(i+j)$. The resulting value of p_2 is then given in equation (3.1). The corollaries follow immediately from Lemma 2.3.

Some of the pertinent quantities of ξ_{12} and ξ_{13} from Corollaries 3.1 and 3.2 are given in the following table.

Table 3.1

ρ	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
p_2	1	.974	.947	.918	.887	.854	.816	.774	.724	.658	.500
$1-2\alpha_1$	0	.026	.053	.082	.113	.146	.184	.226	.276	.342	.500
t	.789	.785	.781	.777	.772	.767	.761	.754	.746	.734	.704
$1-2\alpha_2$	0	.038	.078	.118	.160	.204	.252	.304	.364	.438	.600

The values of $1 - 2\alpha_1$ and $1 - 2\alpha_2$ are listed, as measures of the proportion of the observations which are taken on the interior of the interval and not at the endpoints 0 and 1. Thus for $\rho = .5$ about 15% are to be taken at the midpoint 1/2 to guard against β_2. Note that the two interior points for ξ_{13} stay close to approximately $t = 3/4$ and the weight changes for varying ρ. The D-optimal design for quadratic regression has $p_2 = 2/3$ and $p_4 = 1$ corresponding to $\rho = 8/9$ while the D-optimal design for cubic regression has $p_2 = 3/5$, $p_4 = 2/3$ and $p_6 = 1$ corresponding to $\rho = .96$. (These designs are symmetric and $p_{2i+1} = 1/2$.)

We might mention here in passing that if $m \to \infty$ the resulting limiting design has canonical moments $p_i = 1/2$, $i \neq 2$ and $p_2 = (1+\sqrt{1-\rho})/2$. It can be shown, somewhat difficult arguments involving Stieltjes transforms, that the limiting design, denoted by $\xi_{1\infty}$, has density on $[0,1]$ given by

$$\frac{\rho}{\pi(x(1-x))^{1/2}[2-\rho-2\sqrt{1-\rho}+16\sqrt{1-\rho}\ x(1-x)]} \tag{3.2}$$

For $\rho = 1$ this is the arc-sin law while for $\rho \to 0$ the density converges to the D-optimal design for linear regression with masses 1/2 and 1/2 at the endpoints 0 and 1.

The case $r = 2$ and general m can be formulated in terms of canonical moments very easily. For $r = 2$ and $m = 3$ the problem reduces to

$$\begin{cases} \text{maximize } p_2^2 q_2 p_4 \\ \text{subject to } p_2 q_2 p_4 q_4 \geq \rho/16 \ . \end{cases} \tag{3.3}$$

For given p_2 the solution for p_4 is

$$p_4 = 1/2[1+(1-\frac{\rho}{4p_2q_2})^{1/2}] \ . \tag{3.4}$$

If we substitute this value back into $p_2^2 q_2 p_4$ and maximize with respect to p_2 we find that p_2 is the root of

$$\rho(1-2p)^2 + 16(2-3p)(pq-\frac{\rho}{4})(p-1) = 0 \tag{3.5}$$

that lies between 1/2 and $\max\left\{2/3, \frac{1+\sqrt{1-\rho}}{2}\right\}$. Equation (3.5) is expressed in the form given since an analysis of (3.3) leads to a consideration of the position of the solution p_2 which is related to (3.5). The expression in (3.5) reduces to

$$48p^4 - 128p^3 + 16(\rho+7)p^2 - 8(4+3\rho)p + q\rho = 0. \tag{3.6}$$

The solution for $r = 2$ and general m follows the same pattern as the case $r = 1$ and general m in that p_2 and p_4 remain the same and the higher moments p_{2i} are given by (3.1) for $i \geq 3$.

Theorem 3.2. The D_{2m}-optimal design has canonical moments p_2 and p_4 which are the solution of (3.3) given by (3.5) and (3.4), $p_{2i-1} = 1/2$, $p_{2m} = 1$ and

$$p_{2i} = \frac{m-i}{2m-2i+1}, \quad i=3,\ldots,m-1.$$

The actual design for m = 3 is given by Lemma 2.3. For higher values of m Lemma 2.2 may be used. Some of the pertinant quantities from Lemma 2.3(b) are given in Table 3.2 for m = 3.

Table 3.2.

ρ	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
p_2	.667	.663	.660	.655	.650	.644	.636	.625	.611	.587	.500
p_4	1.0	.971	.941	.909	.874	.837	.797	.752	.699	.634	.500
α	.333	.328	.323	.317	.309	.301	.291	.278	.262	.237	.167
t	.500	.569	.599	.622	.643	.662	.680	.697	.714	.732	.750

Note that for $\rho = 0$ the design begins at the D-optimal design for quadratic regression and ends at $\rho = 1$ giving maximal information for β_3 where it has mass proportional to 1:2:2:1 on the points 0, 1/4, 3/4, 1.

4. Efficiencies.

In this section some D-efficiencies and some G-efficiences are given for the designs described in Section 3. Recall that in equation (1.2) the D-efficiency of ξ for regression with m + 1 parameters was defined by

$$e_m(\xi) = e_m^D(\xi) = \left(\frac{|M_m(\xi)|}{\sup_\eta |M_m(\eta)|} \right)^{1/(m+1)} .$$

The quantity $e_m(\xi)$ measures how well ξ performs in a D-optimal sense for polynomial regression of degree m. The values for the supremum in the denominator can easily be calculated using the expression for $|M_m(\xi)|$ from equation (1.9). The maximum occurs for $p_{2i-1} = 1/2, p_{2i} = (m-i)/(2m-2i+1)$, $i=1,\ldots,m-1$ and $p_{2m} = 1$. Thus the values of $e_m(\xi)$ are relatively simple expressions in terms of the canonical moments of ξ.

The first three efficiencies are given by

$$e_1(\xi) = (p_2)^{1/2}$$

(4.1)
$$e_2(\xi) = 3\left(\frac{p_2^2 q_2 p_4}{4}\right)^{1/3}$$

$$e_3(\xi) = 1/2(5^5 p_2^3 q_2^2 p_4^2 q_4 p_6)^{1/4} .$$

Using the appropriate p_i values for the various designs we find that

$$e_1(\xi_{1m}) = \left(\frac{1+\sqrt{1-\rho}}{2}\right)^{1/2}$$

$$e_2(\xi_{12}) = 3/2\left(\frac{(1+\sqrt{1-\rho})\rho}{4}\right)^{1/3}$$

(4.2)

$$e_2(\xi_{1m}) = \left(\frac{m-1}{2m-3}\right)^{1/3} e_2(\xi_{12})$$

$$e_3(\xi_{13}) = \left(\frac{5^5}{2^7 3^3}(1+\sqrt{1-\rho})\rho^2\right)^{1/4}$$

In Table 4.1 some of the efficiencies are given for various ρ. A plot of some of the efficiencies for ξ_{1m} is drawn in Figure 4.1. The table can be used to plot the efficiencies for ξ_{23} if desired.

Note that the efficiencies that are zero for $\rho = 0$ rise fairly rapidly for increasing ρ and those that are 1 for $\rho = 0$ decrease rather slowly. Thus by increasing ρ one appears to lose a small amount of D-efficiency for the original model for a significant gain in over-all D-efficiency for the higher order models. Recall that ρ itself measures the D-efficiency for

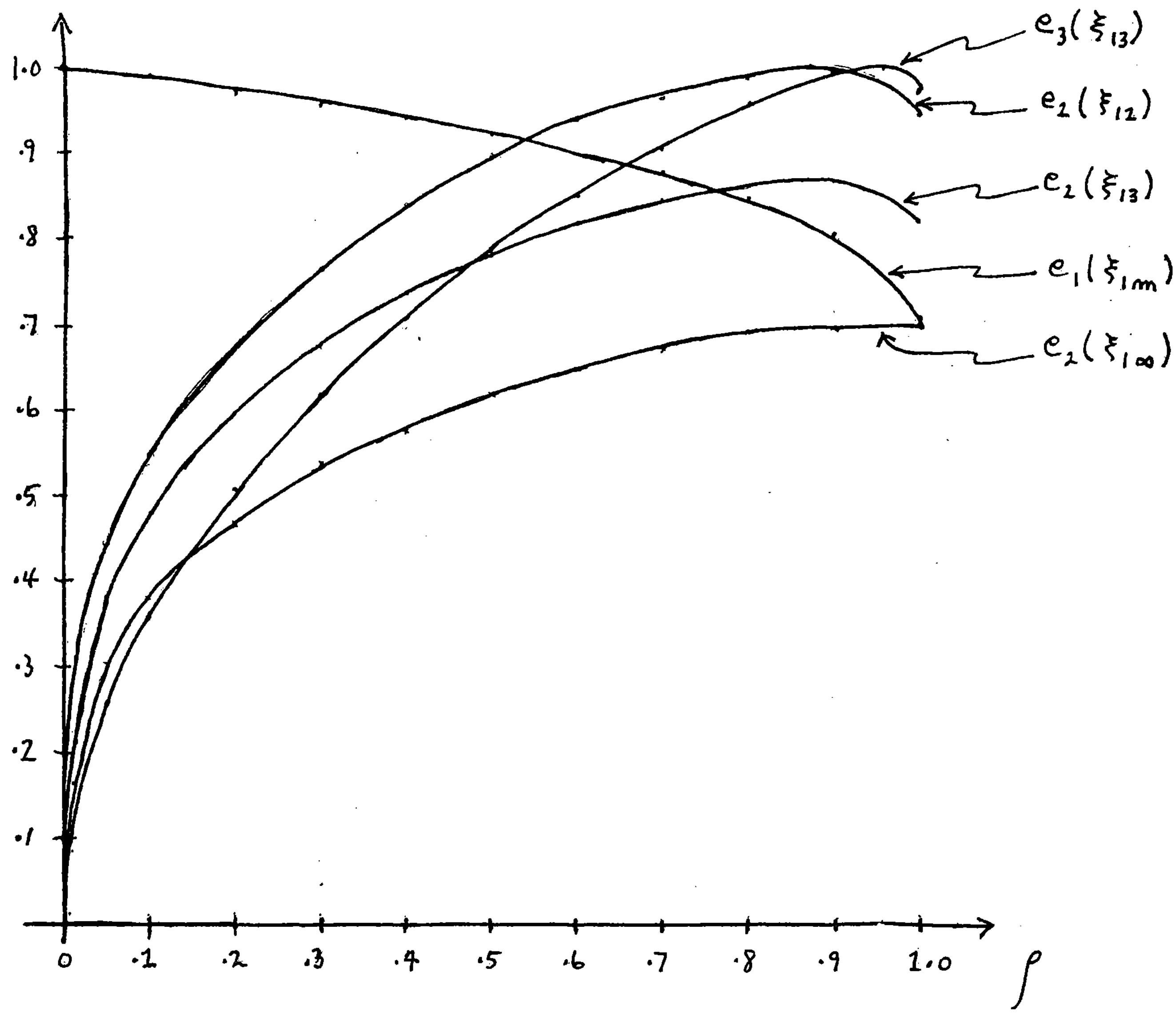
Table 4.1

ρ	0	.05	.1	.2	.3	.4	.5	.6	.7	.8	.9	1.0
$e_1(\xi_{1m})$	1	.994	.987	.973	.958	.942	.924	.904	.880	.851	.811	.707
$e_2(\xi_{12})$	0	.440	.548	.684	.775	.843	.896	.938	.971	.992	1.000	.945
$e_2(\xi_{13})$	0	.385	.479	.597	.677	.736	.783	.820	.848	.867	.873	.825
$e_2(\xi_{1\infty})$	0	.305	.380	.474	.537	.584	.622	.651	.673	.688	.693	.655
$e_3(\xi_{13})$	0	.257	.364	.512	.622	.712	.788	.854	.910	.957	.991	.975
$e_1(\xi_{23})$	.817	.816	.814	.812	.809	.806	.802	.797	.791	.782	.766	.707
$e_2(\xi_{23})$	1.0	1.0	1.0	1.0	1.0	.999	.999	.998	.996	.993	.987	.945
$e_3(\xi_{23})$	0	.546	.647	.764	.838	.891	.932	.963	.985	.998	.998	.935

the neglected part or the higher order terms. The designs themselves for increasing ρ move rather slowly away from the D-optimal design for the lower order model. For example, consider the case $r = 1$, $m = 3$ where our model is roughly linear and protection is desired for β_2, β_3. For $\rho = .5$ the D_{13}-optimal design takes about $1 - 2\alpha_2 = .20$ of the observations (see Table 3.2) away from the endpoints and puts it at approximately 1/4 and 3/4. The resulting design has 50% efficiency for $\beta_2\beta_3$, 92.4% D-efficiency for the linear model and 78.3% and 78.8% overall D-efficiency for the quadratic and cubic models. For $\rho = .8$ the corresponding values are $e_1(\xi_{13}) = .811$ $e_2(\xi_{13}) = .873$, $e_3(\xi_{13}) = .991$. The design in this case moves 36.4% of the observations away from 0 and 1.

Some comparisons with other designs can be made by calculating their canonical moments. For example it can be shown that the design ξ_n which

Figure 4.1



Plot of ρ = D-efficiency for guarding against higher coefficients vs. $e_k(\xi)$ = D-efficiency for degree k.

puts equal weight on n equally spaced points on [0,1] has

(4.3) $$p_2 = 1/3\left(\frac{n+1}{n-1} \right) \text{ and } p_4 = 2/5\left(\frac{n+2}{n-1} \right).$$

For example, $e_1(\xi_{10}) = .638$ and $e_2(\xi_{10}) = .707$. For $n \to \infty$ the corresponding values are $e_1 = .577$ and $e_2 = .585$.

The G or A efficiencies of the designs considered above can also be calculated using canonical moments, the expressions being relatively simple for the linear and quadratic cases. For example let $d_r(x,\xi) = f_1'(x)M_{11}^{-1}(\xi)f_1(x)$ denote the normalized variance for estimating the response function at the point x. It is known that the D-optimal design ξ_r for degree r minimizes $\sup_x d_r(x,\xi)$ and $\sup_x d_r(x,\xi_r) = r + 1$. The G-efficiency $e_r^G(\xi)$, for degree r, of the design ξ is given by

$$e_r^G(\xi) = \frac{r+1}{\sup_x d_r(x,\xi)} .$$

The design ξ_{12} has $p_4 = 1$ so that from (4.3) we find that

$$e_2^G(\xi_{12}) = \begin{cases} 3(1-p_2) & p_2 \geq \frac{2}{3} \quad \text{or} \quad \rho \leq \frac{8}{9} \\ \frac{3}{2}\, p_2 & p_2 \leq \frac{2}{3} \quad \text{or} \quad \rho \geq \frac{8}{9} \end{cases}$$

The design ξ_{13} has $p_4 = \frac{2}{3}$ (from Corollary 3.2) so that

$$e_2^G(\xi_{13}) = \begin{cases} \dfrac{6(1-p_2)}{2+p_2} & p_2 \geq \frac{5}{8} \\ \dfrac{6p_2}{5-p_2} & p_2 \leq \frac{5}{8} \end{cases}$$

A short table of some G-efficiencies is given below.

ρ	0	.1	.2	.3	.4	.5	.6	.7	.8	.9	1
p_2	1	.974	.947	.918	.887	.854	.816	.774	.724	.658	.500
$e_1^G(\xi_{1m})$	1	.987	.973	.957	.940	.921	.899	.873	.840	.794	.667
$e_2^G(\xi_{12})$	0	.078	.159	.246	.339	.438	.552	.678	.828	.987	.750
$e_2^G(\xi_{13})$	0	.052	.108	.169	.235	.307	.392	.489	.608	.772	.667

References

Brezinski, Claude (1980). Padé-Type Approximation and General Orthogonal Polynomials, ISNM 50, Birkhäuser, Boston.

Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. Canad. J. Math. 12, 363-366.

Skibinsky, Morris (1969). Some striking properties of binomial and beta moments. Ann. Math. Statist. 40, 1753-1764.

Stigler, Stephen M. (1971). Optimal experimental design for polynomial regression. JASA, 66, 311-318.

Studden, W. J. (1980). D_s-optimal designs for polynomial regression using continued fractions. Ann. Statist. 8, 1132-1141.

Studden, W. J. (1982). Optimal designs for weighted polynomial regression using canonical moments. To appear in Statistical Decision Theory and Related Topics III.

Wall, H. S. (1948). Analytic Theory of Continued Fractions. Van Nostrand, New York.