MINIMAX DESIGNS IN TWO DIMENSIONAL REGRESSION[1]

By Paul G. Hoel

University of California, Los Angeles

1. Summary. This paper studies the problem of how to space observations in regression so as to minimize the variance of an estimate of the regression function value at an arbitrary point in the domain of observations. Necessary and sufficient conditions are obtained for such a design, called a minimax design, in two dimensional polynomial regression of the type in which the regression function possesses a product structure. Such conditions are also obtained for minimax designs in one dimensional trigonometric and two dimensional spherical harmonics regression. Particular designs of the latter type are constructed.

2. Introduction. Let $f_1(x), \cdots, f_k(x)$ be a set of linearly independent continuous functions defined on a bounded compact domain X and let y_x denote a random variable associated with x whose mean is given by the regression value

$$E(y_x) = \beta_1 f_1(x) + \cdots + \beta_k f_k(x). \tag{1}$$

Let $x_1, \cdots, x_n$ denote a set of points in X at which observations are to be taken of the corresponding y's. It will be assumed that the y's are uncorrelated random variables with a common unknown variance σ^2. When X is one dimensional and the f's are the proper polynomials, this is the classical polynomial regression model.

A basic problem in regression theory is that of estimating in some optimum manner the value of (1) at an arbitrary point x in X. For a fixed set of x's this is usually accomplished by replacing the β's in (1) by their Markov estimates to yield an estimate, which will be denoted by $\hat{y}_x$, that possesses certain optimum properties. If this estimate is used, the design problem then becomes one of choosing the x's in some optimum manner. One approach to a solution, known as the minimax solution, consists in looking for a set of x's that will minimize the quantity $\max_{x \varepsilon X} V(\hat{y}_x)$, where V denotes the variance function. This problem of how to space observations in regression so as to yield a minimax solution has been studied in a number of papers [1], [2], [3], [4]. If the problem is restricted to polynomial regression, the optimum spacing is given by some surprisingly simple and elegant formulas. For more general regression functions, Kiefer and Wolfowitz [5] have obtained a general criterion for minimax optimality that seems very promising for studying other particular types of regression functions. This paper is concerned with applying this criterion to certain two dimensional regression problems and also to trigonometric regression in both one and two dimensions.

Received 27 July 1964; revised 1 February 1965.

[1] This research was supported by the Air Force Office of Scientific Research Grant No. AF-AFOSR-459-63.

The results obtained in this paper for two, and higher, dimensional polynomial regression generalize corresponding results for one dimensional polynomial regression. The two dimensional results should be of value, for example, in deciding where to take observations on a rectangular piece of land in order to estimate in an economical but efficient manner some physical property of it which can be expressed by means of a polynomial in two variables. The minimax criterion that is obtained for trigonometric regression is new and quite different. The two dimensional results should be of value, for example, to geophysicists in helping them to decide where to take observations on the surface of the earth so as to estimate such physical quantities as the steady state heat flow, or the elevation, at an arbitrary point on the surface of the earth. The classical regression functions for such problems are expansions in terms of surface spherical harmonics.

The design problems considered in this paper assume that the regression model (1) is correct with the value of k given in advance. No consideration is given to the problem of determining the errors that arise from choosing the wrong value of k or choosing the incorrect regression model. A striking feature of the trigonometric regression results, however, is that a design which is optimum for any particular order regression is also optimum for all lower order regressions. Thus, it is possible to take steps to guard against choosing too small a value of k for this type of regression.

3. Optimality criteria. A general method for discovering a minimax solution to the design problem described in the preceding paragraph, as given in [5], can be expressed as follows. Let $\{x_i\}$, $\{w_i\}$ denote a set of points in X and a corresponding set of weights which satisfy $\sum_{i=1}^{p} w_i = 1$. A design then consists in agreeing to take the n available observations at the p points $x_1, x_2, \cdots, x_p$ in the proportions $w_1, w_2, \cdots, w_p$. Even though these proportions of n may not yield integer values for the corresponding frequencies, and hence not yield a realistic design, they will be treated as proportions of a valid design. As a result, it may happen that a design that is optimum can only be approximated in a real life situation. Let $g_1(x), \cdots, g_k(x)$ denote an orthonormal set of functions obtained from the f's by the standard orthogonalization procedure with respect to those points and weights. Thus, the g's satisfy

$$\sum_{i=1}^{p} g_\alpha(x_i)g_\beta(x_i)w_i = \delta_{\alpha\beta} . \tag{2}$$

Then it is shown in [5] that the sets $\{x_i\}$, $\{w_i\}$ will minimize the quantity $\max_{x \varepsilon X} V(\hat{y}_x)$ if, and only if, they make

$$\max_{x \varepsilon X} \sum_{\alpha=1}^{k} g_\alpha^2(x) = k. \tag{3}$$

Hereafter a minimax design of this type will be called an optimum design. It should be noted that X need not be one dimensional; in the following sections it will usually be two dimensional.

For certain problems it is possible to find designs that satisfy a more restrictive criterion than that given in (3). This criterion is formulated as a theorem.

THEOREM 1. *If the sets* $\{x_i\}$, $\{w_i\}$ *make* $\sum_{\alpha=1}^{k} g_\alpha^2(x)$ *independent of* x, *the design is optimum.*

PROOF. Assume that $\sum_{\alpha=1}^{k} g_\alpha^2(x) = c$. It then follows that

$$\sum_{\alpha=1}^{k} \sum_{i=1}^{p} g_\alpha^2(x_i)w_i = c\sum_{i=1}^{p} w_i .$$

But because of (2) this equality reduces to $k = c$, and therefore the design satisfies (3).

4. Regression in two dimensions. Criterion (3) will now be applied to certain two dimensional regression problems. Toward this end, let $f_{\alpha\beta}(x, y)$, $\alpha = 1, \cdots, r$, $\beta = 1, \cdots, s$, denote the functions to be used for two dimensional regression. If z denotes the basic regression random variable, the regression function value to be estimated may be expressed as

$$E[z(x, y)] = \sum_{\alpha=1}^{r} \sum_{\beta=1}^{s} c_{\alpha\beta} f_{\alpha\beta}(x, y).$$

Now suppose the basic regression functions $f_{\alpha\beta}(x, y)$ can be factored as follows:

$$f_{\alpha\beta}(x, y) = g_\alpha(x)h_\beta(y). \tag{4}$$

Let (x_i, y_j), $i = 1, \cdots, l$, $j = 1, \cdots, m$, denote a rectangular set of lm points to be selected in the x, y plane and let w_{ij} denote the weight assigned to the point (x_i, y_j). Form an orthonormal set of functions, denoted by $\tilde{g}_\alpha(x)$, $\alpha = 1, \cdots, r$, from the g's with respect to the $\{x_i\}$ and the weights $\{\xi_i\}$, where $\xi_i = \sum_j w_{ij}$. Similarly, let $\tilde{h}_\beta(y)$, $\beta = 1, \cdots, s$, denote an orthonormal set of h's with respect to the points $\{y_j\}$ and the weights $\{\eta_j\}$, where $\eta_j = \sum_i w_{ij}$. Thus, these functions satisfy

$$\sum_{i=1}^{l} \tilde{g}_\alpha(x_i)\tilde{g}_\beta(x_i)\xi_i = \delta_{\alpha\beta} \tag{5}$$
$$\sum_{j=1}^{m} \tilde{h}_\alpha(y_j)\tilde{h}_\beta(y_j)\eta_j = \delta_{\alpha\beta} .$$

Next, without using the factorability of $f_{\alpha\beta}(x, y)$, form an orthonormal set of functions with respect to the points $\{x_i, y_j\}$ and the weights $\{w_{ij}\}$. If these functions are denoted by $f_{\alpha\beta}^*(x, y)$, they must satisfy

$$\sum_{i=1}^{l} \sum_{j=1}^{m} f_{\alpha\beta}^*(x_i, y_j) f_{pq}^*(x_i, y_j) w_{ij} = \delta_{\alpha p}\delta_{\beta q} . \tag{6}$$

Because of assumption (4), it is easy to obtain the following optimum two dimensional design from one dimensional optimum designs.

THEOREM 2. *For two dimensional regression based on functions satisfying* (4), *an optimizing set of points and weights in the domain* $a \leqq x \leqq b$, $c \leqq y \leqq d$ *is given by the points* (x_i, y_j), $i = 1, \cdots, l$, $j = 1, \cdots, m$ *with corresponding weights* $w_{ij} = \xi_i\eta_j$ *that make the orthonormal functions* $\tilde{g}_\alpha(x)$ *and* $\tilde{h}_\beta(y)$ *satisfy, respectively,*

$$\max_{a \leqq x \leqq b} \sum_{\alpha=1}^{r} \tilde{g}_\alpha^2(x) = r, \qquad \max_{c \leqq y \leqq d} \sum_{\beta=1}^{s} \tilde{h}_\beta^2(y) = s.$$

PROOF. The functions $\tilde{f}_{\alpha\beta}(x_i, y_j) = \tilde{g}_\alpha(x_i)\tilde{h}_\beta(y_j)$ when substituted for $f_{\alpha\beta}^*(x_i, y_j)$ in the left side of (6) are readily seen to satisfy (6) because of (5).

Furthermore, because of the factorability of $\tilde{f}_{\alpha\beta}(x, y)$, it follows that

$$\max_{x,y} \sum_{\alpha=1}^{r} \sum_{\beta=1}^{s} \tilde{f}_{\alpha\beta}^{2}(x, y) = \max_x \sum_{\alpha=1}^{r} \tilde{g}_{\alpha}^{2}(x) \cdot \max_y \sum_{\beta=1}^{s} \tilde{h}_{\beta}^{2}(y).$$

But the points and weights were chosen to make the maxima on the right equal to r and s, respectively; consequently the maximum on the left must equal rs. Since there are rs functions in the regression formulation, criterion (3) is satisfied.

It is clear from this proof that Theorem 2 can be generalized to yield a corresponding theorem for higher dimensions.

As an illustration, consider polynomial regression in two dimensions of the type $E[z(x, y)] = c_1 + c_2x + c_3y + c_4x^2 + c_5xy + c_6y^2 + c_7x^2y + c_8xy^2 + c_9x^2y^2$. If the domain is $0 \leqq x \leqq 1$, $0 \leqq y \leqq 1$, an optimum solution is obtained by choosing equal weights at the nine points (x_i, y_j) given by $x_i = 0, \frac{1}{2}, 1$ and $y_j = 0, \frac{1}{2}, 1$. This follows from the fact that this weighting and spacing is known [4] to be optimum for each variable separately. More generally, an optimum design for two dimensional polynomials of this type is given by choosing equal weights and the Legendre points [2] for each variable. It should be noted that a polynomial of the type being considered here is not the traditional two dimensional polynomial of fixed degree.

5. Trigonometric regression in one dimension. Consider a regression function that is the partial sum of a Fourier series, namely,

$$E[y_\theta] = A_0 + \sum_{m=1}^{r} [A_m \cos m\theta + B_m \sin m\theta].$$

The domain here will be chosen to be $0 \leqq \theta \leqq 2\pi$. Now suppose that sets of points $\{\theta_i\}$ and weights $\{w_i\}$ have been chosen which make these trigonometric functions orthogonal and that corresponding normalizing coefficients a_m and b_m have been determined. Then in terms of earlier notation

$$(7) \qquad \sum_{\alpha=1}^{2r+1} g_\alpha^{2}(\theta) = a_0^2 + \sum_{m=1}^{r} [a_m^2 \cos^2 m\theta + b_m^2 \sin^2 m\theta].$$

For this type of regression Theorem 1 is readily applied to yield the following theorem.

THEOREM 3. *A necessary and sufficient condition that a design reduce* (7) *to a constant value is that the normalizing coefficients possess the values* $a_0 = 1$, $a_m = b_m = 2^{\frac{1}{2}}$, $m = 1, \cdots, r$. *A design having these coefficients is optimum not only for regression of order r but also for all lower order regressions.*

PROOF. Replace $\cos^2 m\theta$ by $1 - \sin^2 m\theta$ in (7). Then the reduction of (7) to a constant value requires that $\sum (-a_m^2 + b_m^2) \sin^2 m\theta$ reduce to a constant value. Since $\sin^2 m\theta$ can be expressed as a polynomial of degree m in $x = \sin^2 \theta$, the coefficients of the various powers of x must vanish in this sum. This requires that $a_r^2 = b_r^2$ because $\sin^2 r\theta$ is the only term that contains x^r. Next $a_{r-1}^2 = b_{r-1}^2$ because $\sin^2(r - 1)\theta$ and $\sin^2 r\theta$ are the only terms contributing to x^{r-1}, and since $a_r^2 = b_r^2$ the contribution from $\sin^2 r\theta$ cancels out. This argument can be continued to yield the result that

$$(8) \qquad a_m^2 = b_m^2, \qquad m = 1, \cdots, r.$$

These equalities are clearly both necessary and sufficient to reduce (7) to a constant value.

From the proof of Theorem 1 it follows that this constant value must equal the number of functions in the orthonormal set; hence it follows from (7) and (8) that

$$\sum_{\alpha=1}^{2r+1} g_\alpha^2(\theta) = \sum_{m=0}^{r} a_m^2 = 2r + 1. \tag{9}$$

If these same points and weights are used for a regression of order $r - 1$, the same orthonormal functions will be obtained except for the omission of the last two functions; hence it follows that the first equality in (9) will hold when r is replaced by $r - 1$. Because of Theorem 1, this proves that a design of this type which is optimum for regression of order r is also optimum for regression of order $r - 1$, and hence for all lower orders. Since the right side of (9) must equal $2r - 1$ for regression of order $r - 1$, it follows that $a_r^2 = 2$. The same type of argument requires that $a_m^2 = 2$, $m = 1, \cdots, r$. The value $a_0 = 1$ follows from the normalization of the constant function.

As an illustration, orthogonality considerations suggest that an optimum design for regression of order 3 will be obtained by choosing the points $\theta_i = i\pi/4$, $i = 0, 1, \cdots, 7$, and choosing equal weights. Calculations will show that this design yields the normalizing coefficients required by Theorem 3 and hence that the design is optimum.

6. Trigonometric regression in two dimensions. The generalization to two dimensions of the preceding results will be accomplished by means of surface spherical harmonics. These functions possess desirable orthogonality properties and form a natural basis for expanding a function on the surface of a sphere. The general expansion of a function $u(\theta, \varphi)$ to order r is given by

$$u(\theta, \varphi) = \sum_{n=0}^{r} \sum_{m=0}^{n} [A_{nm} \cos m\varphi + B_{nm} \sin m\varphi] P_{nm}(\cos \theta) \tag{10}$$

where $P_{nm}(\cos \theta)$ is the associated Legendre function defined by

$$P_{nm}(\cos \theta) = \sin^m \theta (d^m/d \cos \theta^m) P_n(\cos \theta)$$

and where $P_n(\cos \theta)$ is the Legendre polynomial of degree n in $\cos \theta$. The angles θ and φ are the co-latitude and azimuth, respectively, of a point on the surface of a unit sphere.

Now suppose that a set of points $\{\theta_i, \varphi_i\}$ and a corresponding set of weights $\{w_i\}$ have been found that make the functions of (10) orthogonal and that normalizing constants a_{nm} and b_{nm} have been determined with respect to those points and weights. If the resulting orthonormal functions are denoted by $g_\alpha(\theta, \varphi)$, then

$$\sum_\alpha g_\alpha^2(\theta, \varphi) = \sum_{n=0}^{r} \sum_{m=0}^{n} [a_{nm}^2 \cos^2 m\varphi + b_{nm}^2 \sin^2 m\varphi] P_{nm}^2(\cos \theta). \tag{11}$$

As in the one dimensional situation, consider the possibility of finding a design that reduces the right side of (11) to a constant value. By Theorem 1 such a design will be optimum. A first necessary condition for a design of this type is given by the following lemma.

LEMMA 1. *A design that reduces* (11) *to a constant value must yield normalizing constants satisfying* $a_{nm}^2 = b_{nm}^2$, $n = 1, \cdots, r$, $m = 1, \cdots, n$.

PROOF. Let $x = \sin^2 \theta$ and $y = \sin^2 \varphi$. It is well known that $P_{nm}^2(\cos \theta)$ is a polynomial of degree n in x and that $\cos^2 m\varphi$ and $\sin^2 m\varphi$ are polynomials of degree m in y. Thus, the right side of (11) is a polynomial in x and y; hence if (11) is to reduce to a constant, the coefficients of the nonconstant terms must vanish. Consider the terms in $x^r y^m$ and rewrite (11) in the form

$$(12)\qquad \sum_\alpha g_\alpha^2(\theta, \varphi) = G(\theta, \varphi) + \sum_{m=0}^r [a_{rm}^2 \cos^2 m\varphi + b_{rm}^2 \sin^2 m\varphi] P_{rm}^2(\cos \theta).$$

$G(\theta, \varphi)$ will contain no terms in $x^r y^m$. A term in $x^r y^r$ will be found only in the terms involving $\cos^2 r\varphi$, or $\sin^2 r\varphi$, and $P_{rr}^2(\cos \theta)$. If $\cos^2 r\varphi$ is replaced by $1 - \sin^2 r\varphi$, it follows that the coefficient of $x^r y^r$ will vanish if, and only if, $a_{rr}^2 = b_{rr}^2$. Next, the coefficient of $x^r y^{r-1}$ will involve terms in $P_{rr}^2(\cos \theta)$ and $P_{rr-1}^2(\cos \theta)$ only. Since $a_{rr}^2 = b_{rr}^2$ the contribution from terms involving $P_{rr}^2(\cos \theta)$ will drop out. Replacing $\cos^2(r - 1)\varphi$ by $1 - \sin^2(r - 1)\varphi$, it follows that the coefficient of $x^r y^{r-1}$ will vanish if, and only if, $a_{rr-1}^2 = b_{rr-1}^2$. This argument can be continued to yield the result

$$(13)\qquad a_{rm}^2 = b_{rm}^2, \qquad m = 1, \cdots, r.$$

The type of argument used to demonstrate (13) can be employed to demonstrate the more general result of the lemma.

For a design that reduces (11) to a constant, it follows from this lemma that (11) will reduce to

$$(14)\qquad \sum_\alpha g_\alpha^2(\theta, \varphi) = \sum_{n=0}^r \sum_{m=0}^n a_{nm}^2 P_{nm}^2(\cos \theta).$$

Additional information concerning the nature of the coefficients a_{nm}^2 can be obtained by studying the orthogonality properties of the functions $P_{n0}(\cos \theta)$, $n = 0, 1, \cdots, r$, which are the Legendre polynomials of degree n in $\cos \theta$. The following lemma is needed in this development.

LEMMA 2. *If a set of points* $\{\theta_i, \varphi_i\}$ *and weights* $\{w_i\}$ *orthogonalize the functions in* (10), *then*

$$(15)\qquad \sum_i w_i(\cos \theta_i)^{2m} = 1/(2m + 1), \qquad m = 1, \cdots, r - 1.$$

PROOF. The proof will be by induction. The orthogonality of $P_{00}(\cos \theta) = 1$ and $P_{20}(\cos \theta) = (3 \cos^2 \theta - 1)/2$ verifies the formula for $m = 1$. Now assuming that the formula holds for $m = 1, \cdots, s$, it will be shown that it holds for $m = s + 1$.

Write $P_{n0}(\cos \theta) = P_n(x)$ for the Legendre polynomial of degree n. The orthogonality of these polynomials is given by

$$(16)\qquad \int_{-1}^1 P_\alpha(x) P_\beta(x)\, dx = 0, \qquad \alpha \neq \beta.$$

Let α and β be any two unequal integers satisfying $\alpha + \beta = 2(s + 1)$ and let

$P_\alpha(x)$ be written in the form

$$P_\alpha(x) = \sum_{j=0}^{[\alpha/2]} c_j x^{\alpha-2j}$$

where $[\alpha/2]$ denotes the largest integer in $\alpha/2$. Then (16) yields

$$\tfrac{1}{2}\int_{-1}^{1} \sum_j c_j x^{\alpha-2j} \sum_l d_l x^{\beta-2l}\, dx = 0. \tag{17}$$

The corresponding equation for the orthogonality of $P_{\alpha_0}(\cos\theta) = P_\alpha(x)$ and $P_{\beta_0}(\cos\theta) = P_\beta(x)$ with respect to the points $\{\theta_i\}$ and weights $\{w_i\}$ is

$$\sum_i w_i \sum_j c_j x_i^{\alpha-2j} \sum_l d_l x_i^{\beta-2l} = 0. \tag{18}$$

But since (15) is assumed to hold for $m = 1, \cdots, s$,

$$\sum_i w_i x_i^{2m} = 1/(2m+1) = \tfrac{1}{2}\int_{-1}^{1} x^{2m}\, dx, \qquad m = 1, \cdots, s. \tag{19}$$

If the sums in (17) and the corresponding sums in (18) are multiplied out and the individual terms integrated and summed, respectively, the corresponding individual terms will be identical because of (19), except for the highest power term in each, namely, the one with exponent $\alpha + \beta$. Equating (17) and (18) then shows that these terms must also be equal; hence since $\alpha + \beta = 2(s+1)$,

$$\sum w_i x_i^{2(s+1)} = 1/[2(s+1)+1].$$

This completes the induction step. The upper limit of $m = r - 1$ in (15) arises from the fact that $r - 1$ is the highest power of x that can be obtained from orthogonality considerations.

These lemmas will now be used to prove the main theorem.

THEOREM 4. *A necessary and sufficient condition that a design reduce* (11) *to a constant value is that the normalizing coefficients possess the values*

$$a_{n0}^2 = 2n+1, \quad a_{nm}^2 = b_{nm}^2 = 2(2n+1)(n-m)!/(n+m)!, \quad m \neq 0. \tag{20}$$

A design having these coefficients is optimum not only for regression of order r but also for all lower order regressions.

PROOF. Consider the necessity; the sufficiency will be obvious in the course of the necessity proof. For $m \neq 0$, it follows from Lemma 1 that

$$\begin{aligned} 1/a_{nm}^2 &= \tfrac{1}{2}(1/a_{nm}^2 + 1/b_{nm}^2) = \tfrac{1}{2}\sum_i w_i[\cos^2 m\varphi_i + \sin^2 m\varphi_i]P_{nm}^2(\cos\theta_i) \\ &= \tfrac{1}{2}\sum_i w_i P_{nm}^2(\cos\theta_i). \end{aligned} \tag{21}$$

For $m = 0$ the normalizing coefficient is obtained directly from definition. Let $x = \cos\theta$ and let $P_{nm}^2(x)$, which is a polynomial of degree n in x^2, be expressed as

$$P_{nm}^2(x) = \sum_{j=0}^{n} h_j x^{2j}. \tag{22}$$

Then (21), (22) and (15) will yield

$$1/a_{nm}^2 = \tfrac{1}{2}\sum_{j=0}^{n} h_j \sum_i w_i x_i^{2j} = \tfrac{1}{2}\sum_{j=0}^{n} [h_j/(2j+1)], \; n = 1, \cdots, r-1. \tag{23}$$

A standard formula for associated Legendre functions is

$$\tfrac{1}{2}\int_{-1}^{1} P_{nm}^2(x)\,dx = (2n+1)^{-1}(n+m)!/(n-m)!.$$

If both sides of (22) are multiplied by $\frac{1}{2}$ and integrated from -1 to 1, this formula will yield the result

$$(2n+1)^{-1}(n+m)!/(n-m)! = \sum_{j=0}^{n} [h_j/(2j+1)].$$

Applying this result to (23) gives (20) for all values of n except $n = r$. Now consider its verification for $n = r$.

There is an addition formula for associated Legendre functions [6] which when properly applied yields the identity

$$P_{n0}^2(x) + 2\sum_{m=1}^{n} [(n-m)!/(n+m)!]P_{nm}^2(x) = 1. \tag{24}$$

If formula (20) is applied to (14) for each value of $n < r$ and then formula (24) introduced, it will be observed that

$$\begin{aligned}\sum_{m=0}^{n} a_{nm}^2 P_{nm}^2(x) &= 2(2n+1)\sum_{m=1}^{n} [(n-m)!/(n+m)!]P_{nm}^2(x) \\ &\quad + (2n+1)P_{n0}^2(x) \\ &= 2n+1, \qquad n < r.\end{aligned} \tag{25}$$

Since this holds for all $n < r$, (14) will reduce to

$$\sum_{\alpha} g_{\alpha}^2(\theta, \varphi) = \sum_{n=0}^{r-1} (2n+1) + \sum_{m=0}^{r} a_{rm}^2 P_{rm}^2(\cos\theta). \tag{26}$$

But if the right side is to reduce to a constant it must reduce to the value $(r+1)^2$ because that is the number of functions, $g_{\alpha}(\theta, \varphi)$, in (11). Since the first sum on the right of (26) is r^2, this requires that

$$\sum_{m=0}^{r} a_{rm}^2 P_{rm}^2(\cos\theta) = 2r + 1. \tag{27}$$

From (25) this equality will be satisfied if a_{rm}^2 is given by formula (20). However, because of the nature of the polynomials $P_{rm}^2(\cos\theta)$, it is easily shown that there must be a unique set of constants a_{rm}^2 satisfying (27) and hence that they must be given by (20). This completes the proof.

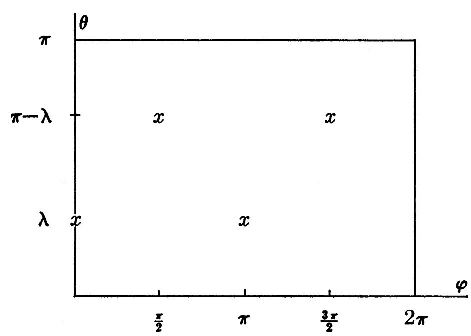

FIG. 1

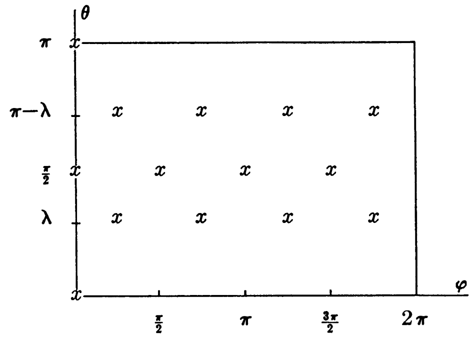

FIG. 2

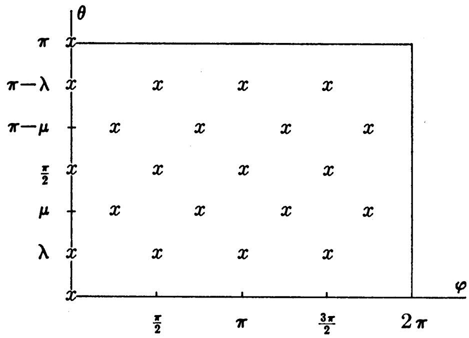

FIG. 3

From (27) it follows that when r is replaced by $r - 1$, (11) will still reduce to a constant because its value will be the original value, namely $(r + 1)^2$, minus the value given by (27), namely $2r + 1$. Thus, a design of this type which is optimum for regression of order r is also optimum for all lower order regressions.

7. Particular optimum designs. The results of the preceding section will be applied to construct optimum designs for regression of orders 1, 2, and 3. These regressions, given by (10), will contain 4, 9, and 16 terms, respectively.

First, consider regression of order 1, namely

$$E[z(\theta, \varphi)] = c_1 + c_2 \cos \theta + c_3 \sin \theta \cos \varphi + c_4 \sin \theta \sin \varphi.$$

Considerations of symmetry and orthogonality suggest trying the design as shown in Figure 1. Calculations will show that the four functions here are orthogonal on this set of points if equal weights are used, regardless of the value of λ chosen. If λ is chosen to satisfy $\cos^2 \lambda = \frac{1}{3}$, the normalizing constants will satisfy (20) and hence the design will be optimum.

Next, consider regression of order 2. Symmetry and orthogonality suggest the design as shown in Figure 2. Orthogonality requirements, together with formulas (20), required that λ satisfy $\cos^2 \lambda = \frac{1}{3}$ and that the weights be chosen as $\frac{1}{15}$ for all points for which $\theta = 0, \pi/2, \pi$ and $\frac{3}{40}$ for all other points. It is relatively easy to construct other optimum designs using more points. For example, if two angles, λ and μ, are introduced, a design based on 22 points with equal weights can be constructed.

Finally, consider regression of order 3. The usual symmetry and orthogonality considerations led to the design based on 22 points as shown in Figure 3. Here it will be found that it is necessary to choose λ and μ to satisfy $\sin^2 \lambda = \frac{3}{7}$ and $\sin^2 \mu = \frac{6}{7}$ and to choose weights that are proportional to 440 for $\theta = 0$ and π, 594 for $\theta = \pi/2$, and 539 for $\theta = \lambda, \mu, \pi - \lambda$, and $\pi - \mu$.

Although Theorem 4 does not give a constructive method of finding optimum solutions, it would appear that symmetry and orthogonality considerations, together with Theorem 4, should enable one with the help of modern computing techniques to find optimum designs for higher order regressions.

I wish to express my appreciation to my former colleague Professor Henrici for suggesting where I might find a formula that would yield formula (24).

REFERENCES

[1] GUEST, P. G. (1958). The spacing of observations in polynomial regression. *Ann. Math. Statist.* **29** 294–299.

[2] HOEL, P. G. (1958). Efficiency problems in polynomial estimation. *Ann. Math. Statist.* **29** 1134–1145.

[3] HOEL, P. G. and LEVINE, A. Optimal spacing and weighting in polynomial prediction. *Ann. Math. Statist.* **35** 1553–1560.

[4] KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. *Ann. Math. Statist.* **30** 271–294.

[5] KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. *Canad. J. Math.* **12** 363–366.

[6] ERDÉLYI, A. editor. (1953). *Higher Transcendental Functions*, **1** McGraw-Hill, New York, 168.