On a Problem of Chebyshev

by

W. J. Studden *

Department of Statistics
Purdue University
W. Lafayette, IN 47907

Department of Statistics
Division of Mathematical Sciences
Mimeograph Series #79-4

* Research supported by NSF Grant No. MCS76 - 08235 A01

On a Problem of Chebyshev

by

W. J. Studden *

Purdue University, W. Lafayette, IN 47907

1. Introduction.

The classical problem of Chebyshev mentioned in the title concerns the approximation of a power x^n by polynomials $Q_{n-1}(x)$ of degree n-1 using a sup norm over the interval $[-1,1]$. The polynomial $x^n-Q_{n-1}(x)$ in this case is the Chebyshev polynomial of the 1st kind $T_n(x) = \cos n\theta$ $(x = \cos\theta)$ with leading coefficient set equal to one. The problem considered here is to approximate the powers $x^{s+1}, s^{s+2}, \ldots, x^n$ simultaneously using the lower terms $1, x, \ldots, x^s$. Let $f'(x) = (1, x, \ldots, x^n)$ (primes will denote transposes), $f_1'(x) = (1, x, \ldots, x^s)$ and $f_2'(x) = (x^{s+1}, x^{s+2}, \ldots, x^n)$. Further let Q be an arbitrary $(n-s) \times (s+1)$ matrix and A be a positive definite $(n-s)\times(n-s)$ matrix with a fixed value, say one, for its determinant. It is required to find the value of both Q and A which will minimize the supremum over $[-1,1]$ of

$$d(x;Q,A) = (f_2(x) - Qf_1(x))' A(f_2(x) - Qf_1(x)) \tag{1.1}$$

Note that when $s = n-1$ we have the original problem of Chebyshev. The solution to the generalized problem arose from a problem in the optimal design of experiments. It is arrived at fairly simply using certain "canonical moments" of measures on $[-1,1]$. The simplicity of the solution seems to require minimizing over the matrix A as well as polynomial part Q.

A solution to the original problem using the canonical moments is described in the Section 2. Section 3 describes the general solution. Some

* Research supported by NSF Grant No. MCS76 - 08235 A01

examples are considered in the final section together with some properties of the general solution.

2. Solution of Original Problem.

In this section we give a solution to the original problem using canonical moments. Let q denote a vector of dimension n+1 with a one in the last component. The problem is then to minimize

$$\sup_{x\in[-1,1]} | q'f(x) |^2 \tag{2.1}$$

with respect to q. If ξ denotes an arbitrary probability measure on $[-1,1]$ then (2.1) may be replaced by

$$\sup_{\xi} \int (q'f(x))^2 d\xi(x)$$

$$= \sup_{\xi} q'M(\xi)q$$

where $M(\xi)$ is the $(n+1)x(n+1)$ matrix with elements

$$m_{ij} = \int x^{i+j} d\xi(x),\ i,\ j = 0,1,\ldots,n \quad .$$

Using game theoretic arguments it may be shown that

$$\rho = \inf_{q} \sup_{\xi} q'M(\xi)q = \sup_{\xi} \inf_{q} q'M(\xi)q \quad .$$

Letting $e' = (0,\ ,\ldots,0,1)$ it then follows that

$$\rho^{-1} = \inf_{\xi} \sup_{q} \frac{(e'q)^2}{q'M(\xi)q}$$

$$= \inf_{\xi} e'M^{-1}(\xi)e$$

The last equality uses Schwarz inequality. Note for later reference that equality is achieved for the supremum over q if and only if

$$q = M^{-1}(\xi)e/e'M^{-1}(\xi)e \tag{2.2}$$

The problem is now to minimize

$$e'M^{-1}(\xi)e = \frac{|M_{11}(\xi)|}{|M(\xi)|} \tag{2.3}$$

where $|M(\xi)|$ and $|M_{11}(\xi)|$ are the determinants of $M(\xi)$ and

$$M_{11}(\xi) = \int f_1(x)f_1'(x)d\xi(x).$$

The two determinants involved and their ratio have a simple expression in terms of the canonical moments of ξ. For any probability measure ξ on $[-1,1]$ let $c_i = \int x^i d\xi(x)$, $i = 0,1,\ldots$

Now let c_k^+ denote the maximize value of the k^{th} moment over measures μ having the same first k-1 moments as ξ. That is, consider those μ on $[-1,1]$ with $\int x^i d\mu(x) = c_i$ for $i = 0,1,\ldots,k-1$ then $c_k^+ = \sup_\mu \int x^k d\mu(x)$. Similarly let c_k^- denote the corresponding minimum. The canonical moments are defined by

$$p_k = \frac{c_k - c_k^-}{c_k^+ - c_k^-} \qquad k = 1,2,\ldots$$

Whenever $c_k^- = c_k^+$ we leave the p_k undefined.

If we then let

$$\eta_0 = q_0 = 1 \quad \eta_j = q_{j-1}p_j \quad j = 1,2,\ldots \quad (p_i+q_i = 1)$$

the determinant $|M(\xi)|$ is then (see Skibinsky(1969) or Studden (1979)) a multiple of two times the quantity

$$\prod_{i=1}^{n} (\eta_{2i-1}\eta_{2i})^{n+1-i}$$

The ratio in (2.3) then turns out to be a constant times

$$\frac{\prod_{i=1}^{s} (\eta_{2i-1}\eta_{2i})^{s+1-i}}{\prod_{i=1}^{n} (\eta_{2i-1}\eta_{2i})^{n+1-i}} \tag{2.4}$$

For $s = n-1$ this quantity is the inverse of

$$p_{2n} \prod_{i=1}^{2n-1} p_i q_i$$

which is maximized for

$$p_i = \frac{1}{2} \quad i = 1,2,\ldots,2n-1, p_{2n} = 1 \tag{2.5}$$

(The general solution is given in (3.4) below). Now the measure with density

$$\frac{1}{\pi\sqrt{1-x^2}} \tag{2.6}$$

has canonical moments $p_i \equiv 1/2$. See Skibinsky (1969) or Karlin and Studden (1966) p. 120 . Since the moments $c_0 = 1$, $c_1,\ldots,c_k$ and $p_1,p_2,\ldots p_k$ are in 1-1 correspondence the minimizing measure ξ_{n-1} corresponding to (2.5) has its $1^{\underline{st}}$ 2n-1 moments equal to those of the measure (2.6).

The solution to the original problem, namely that $T_n(x)$,with leading coefficient, minimizes (2.1) now follows. It can be readily seen that using the corresponding $q = q_{n-1}$ from (2.2) the polynomial $q'_{n-1}f(x)$ is orthogonal to $x^k, k = 0,1,\ldots,n-1$ with respect to the measure in (2.6).

The measure ξ_{n-1} corresponding to (2.5) is an "upper principal representation" for the measure (2.6). It concentrates mass proportional to 1:2:2:...:2:1 at the n+1 zeros of $(1-x^2)T_n'(x) = 0$. This can be verified by noting that ξ_{n-1} provides a quadrature formula corresponding to the measure (2.6) which is exact for polynomials of degree 2n-1. This quadrature formula is a classical

Bouzitat formula of the second kind. (See Ghizzetti and Ossicini (1970)). It may also be verified by noting that the support of ξ_{n-1} must be the points where $T_m^2(x)$ attains its supremum, i.e. the zeros of $(1-x^2)T_n'(x) = 0$. The corresponding weights at these points may be obtained by matching up the first n moments and requiring total mass equal to one.

3. The General Solution.

As indicated in the introduction the problem now is to find the Q and A which will minimize the supremum on $[-1,1]$ of the quantity $d(x;Q,A)$ defined in equation (1.1). A considerable simplification is obtained if we use some of the results from Karlin and Studden (1966) page 367. It is shown there that the minimizing Q and A are of a certain form. For any ξ we partition the matrix $M(\xi)$ according to f_1 and f_2 by defining

$$M_{11}(\xi) = M_{11} = \int f_1 f_1' d\xi,\ M_{22} = \int f_2 f_2' d\xi \text{ and } M_{12}' = M_{12} = \int f_1 f_2' d\xi$$

so that

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

The minimizing Q is shown to be of the form

$$Q = Q(\xi_s) = M_{21}(\xi_s) M_{11}^{-1}(\xi_s) \tag{3.1}$$

where ξ_s maximizes the determinant of the matrix

$$A^{-1}(\xi) = M_{22}(\xi) - M_{21}(\xi) M_{11}^{-1}(\xi) M_{12}(\xi) \tag{3.2}$$

The matrix A was normalized to have determinant equal to one. The minimizing A is the matrix $A(\xi_s)$ suitably normalized. Since the normalization does not change the problem we can restrict the matrix A to have determinant equal to that of $A(\xi_s)$.

Now the identity

$$|M| = |M_{11}| \; |M_{22} - M_{21}M_{11}^{-1}M_{12}| \qquad (3.3)$$

shows that minimizing $|A(\xi)|$ is equivalent to minimizing (2.4) for general s. The minimizing measure ξ_s can readily be shown to have canonical moments

$$p_i = \frac{1}{2} \qquad i \text{ odd}$$

$$p_{2i} = \begin{cases} \frac{1}{2} & i = 1,2,\ldots,s \\ \frac{n-i+1}{2n-2i+1} & i = s+1,s+2,\ldots,n-1 \\ 1 & i = n \end{cases} \qquad (3.4)$$

One can now convert back to the measure ξ_s, then to the ordinary moments of ξ_s and then to the matrices Q and A. It is also possible to evaluate the ordinary moments of ξ_s used in Q and A directly from the canonical moments given in (3.4). These relationships, which originate with Stieltjes, are described more fully in Studden (1979) or Skibinsky (1969) and relate the power series generating the ordinary moments with its continued fraction expansion.

Let $\delta_0 = 1$ $\delta_i = q_{2i-2}p_{2i}$ $i = 1,2,\ldots$ and define U_{ij} recursively by $U_{0j} \equiv 1$, $j = 0,1,\ldots$ and for $i \leq j$

$$U_{ij} = \sum_{k=i}^{j} \delta_{k-i+1} U_{i-1,k} \quad i,j=1,2,\ldots \qquad (3.5)$$

Whenever $p_i = 1/2$ for i odd the ordinary moments $c_i = \int x^i d\xi$ are then given by

$$c_{2i-1} = 0, \quad c_{2i} = U_{ii} \quad i = 1,2,\ldots n \qquad (3.6)$$

The relationship between the p_i and the c_i is slightly more involved when the symmetry producing $p_{2i-1} = \frac{1}{2}$ is not present.

The minimizing measure ξ_s also has a simple description. See Studden (1979). The support of ξ_s consists of the points ± 1 and the n-1 zeros of

$$p'_{n-s}(x)t'_{s+1}(x) - \alpha_s p'_{n-s-1}(x)t'_s(x) = 0 \qquad (3.7)$$

where

$$\alpha_s = \frac{1}{2}\,\frac{(n-s+1)}{(2n-2s+3)} \qquad s = 0,1,\ldots,n-1$$

and p_i and i_i are the Legendre and Chebyshev polynomials $P_k(x)$ and $T_k(x)$ normalized so that their leading coefficients are one.

The prime here denotes differentiation. We note that $T'_k(x)$ is the Chebyshev polynomial of the second kind and $P'_k(x)$, $k = 1,2,\ldots$ are orthogonal to $(1-x^2)dx$, however we prefer to leave things in terms of P_k and T_k.

The weights that ξ_s assigns to each of the zeros x_i of (3.7) and ± 1 is given by

$$\xi_s(i) = \frac{2}{2n+1+U_{2s}(x_i)} \qquad i = 0,1,\ldots,n \qquad (3.8)$$

where $U_{2s}(x)$ is the Chebyshev polynomial of the $2^{\underline{nd}}$ kind,

$$U_k(x) = \frac{\sin(k+1)\theta}{\sin\theta} \quad , \quad x = \cos\theta \quad .$$

4. Examples and Further Properties.

The reduction of the minimizing A and Q to the form (3.1) and (3.2) is given in Karlin and Studden (1966). The same theorem also says that with the matrix A of the form (3.2) the quantity $d(x;Q,A)$, with the minimizing $Q_s = Q(\xi_s)$ and $A_s = A(\xi_s)$, satisfies the inequality

$$d(x;Q_s,A_s) \le n-s \tag{4.1}$$

For $s = n-1$ the expression $d(x;Q_s,A_s)$ reduces to $T_n^2(x)$. Equation (4.1) is then just the familiar fact the $T_n^2(x) \le 1$ for $x \in [-1,1]$.

The polynomial T_n is orthogonal to $x^k, k = 0,1,\ldots,n-1$ with respect to (2.6). Since the minimizing measure ξ_{n-1} and (2.6) have the same moments $c_0,c_1,\ldots,c_{2n-1}$ it follows that t_n and x^k are orthogonal with respect to ξ_{n-1}. There seems to be no analog to (2.6) for the general measure ξ_s. However if we define

$$(g_{s+1},\ldots,g_n)' = A_s^{\frac{1}{2}} (f_2 - Q_s f_1)$$

then the polynomials g_i, $i = s+1,\ldots,n$ are orthonormal and orthogonal to $1,x,\ldots,x^s$ with respect to the measure ξ_s. See Kiefer (1962) or Karlin and Studden (1966).

As a specific example consider the case m = 3 . Using equation (3.4) we note that all the odd canonical moments are equal to $\frac{1}{2}$ while the even moments are

	p_2	p_4	p_6
s = 0	3/5	2/3	1
s = 1	1/2	2/3	1
s = 2	1/2	1/2	1

A few further calculations using (3.6) and (3.5) give the ordinary moments. The odd moments c_{2i-1} are zero while the even moments are given by the following:

	c_2	c_4	c_6
s = 0	3/5	13/25	63/125
s = 1	1/2	5/12	29/72
s = 2	1/2	3/8	11/32

Equation (4.1) for the three cases then gives

s = 2:

$$\left(x^3 - \frac{3}{4}x\right) \; 16 \; \left(x^3 - \frac{3}{4}x\right) \leq 1$$

s = 1:

$$\begin{pmatrix} x^2 - \frac{1}{2} \\ x^3 - \frac{5}{6}x \end{pmatrix}' \begin{pmatrix} 6 & 0 \\ 0 & 18 \end{pmatrix} \begin{pmatrix} x^2 - \frac{1}{2} \\ x^3 - \frac{5}{6}x \end{pmatrix} \leq 2$$

s = 0:

$$\begin{pmatrix} x \\ x^2 - \frac{3}{5} \\ x^3 \end{pmatrix}' \frac{5}{4} \begin{pmatrix} 63 & 0 & -13 \\ 0 & 5 & 0 \\ -13 & 0 & 15 \end{pmatrix} \begin{pmatrix} x \\ x^2 - \frac{3}{5} \\ x^3 \end{pmatrix} \leq 3$$

For $s = 0$ the measure ξ_0 has equal mass $1/(n+1)$ on the zeros of $(1 - x^2)P_n'(x) = 0$, which for $n = 3$ gives $x = \pm 1$ and $x = \pm 1\sqrt{5}$. For $s = n-1$ the measure ξ_{n-1} has mass on the zeros of $(1 - x^2)T_n'(x) = 0$ which are $x_\nu = \cos\frac{\nu\pi}{n}$ $s = 0,1,\ldots,n$. The interior zeros have weight $1/n$ while ± 1 have weight $1/2n$ each. For $s = 1$ and $n = 3$ there is weight 3/10 on $x = \pm 1/\sqrt{6}$ and 1/5 on $x = \pm 1$.

As a final remark observe that ξ_0 and ξ_{n-1} give essentially equal weight to the zeros of $(1-x^2)P_n'(x) = 0$ and $(1 - x^2)T_n'(x) = 0$ respectively. The zeros of all the classical polynomials distribute themselves according to the density given in (2.6). Therefore ξ_0 and ξ_{n-1} both converge (weakly) to the measure with density (2.6) as $n \to \infty$. The measures ξ_s, which also depend on n, can be shown to lie between ξ_0 and ξ_{n-1} so that they all converge to (2.6) uniformly in s.

[1] Ghizzetti, A. and Ossicini, A., "Quadrature Formulae," Academic Press, New York, 1970.

[2] Karlin, S. and Studden, W.J., "Tchebycheff Systems: with Applications in Analysis and Statistics," Interscience, New York, 1966.

[3] Kiefer, J., An extremum result, Can. J. Math 14 (1962), 597 - 601.

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

[5] Studden, W.J., D_s-optimal designs for polynomial regression using continued fractions, Dept. of Statistics Mimeo Series #79 - 2, Purdue University, Jan. 1979.