CS23

CONVEX POLYNOMIAL APPROXIMATION

BY

BERNARD D. RUDIN

TECHNICAL REPORT CS23
JUNE 4, 1965

COMPUTER SCIENCE DEPARTMENT
School of Humanities and Sciences
STANFORD UNIVERSITY

PREPARED UNDER CONTRACT **Nonr-225(37)(NR-044-211)**
OFFICE OF NAVAL RESEARCH

PART 1: GENERAL CONSIDERATIONS

Introduction.

The problem to be considered is that of obtaining polynomial approximations to continuous functions or empirical data in such a way that the approximating polynomials are convex in some prescribed interval.

This problem arises naturally in connection with data smoothing and was in fact suggested to the author by a problem requiring the calculation of derivatives from data,

The difficulties arising from the use of interpolation and least squares methods for data smoothing by polynomial approximation are well known. There are excellent discussions in Lanczos [10]*, and Hamming [7]. There appears, however, to be very little literature which treats the problem of interest by methods of constrained polynomial approximation.

Such problems are usually posed in terms of minimizing functionals, which suggests treatment by variational methods. A similar problem has been so treated by Boltjanskii [3]. He examined the problem of approximating continuous functions with functions whose n-th derivatives satisfy a Lipschitz condition. By application of the Pontrjagin maximum principle [15], he obtained necessary conditions which solutions of the problem must satisfy. The problem of interest here can be formally stated in a manner similar to that of Boltjanskii, but such a representation does not appear to help in the study of means of computing best approximations.

Methods of polynomial approximation where the polynomial coefficient vectors are constrained to lie in a convex set are treated in the recent

---

*Numbers in brackets indicate references in the Bibliography.

paper of Rice [16]. He showed that the problem of obtaining best approximations to continuous functions with n-th degree polynomials whose k-th derivative is positive on [0,1] has solutions and gives conditions for uniqueness and location of these solutions. However, he does not find the problem of computing such approximations to be tractable.

The difficulties that Rice encountered are of two kinds. First, the problem is essentially nonlinear, having nonlinear constraints. Second, the geometry of the constraint set is difficult to deal with because it is not given explicitly.

In this work, both of the aforementioned difficulties will be dealt with. In this first part, after development of some necessary preliminaries, a theory for treating a class of nonlinear approximation problems is presented. This class is of interest not only because it includes the convex polynomial approximation problem, but because it provides a potentially useful generalization of the linear theory. A typical problem of this class is expressed as follows: Given an element f of a real normed linear space V, a set $\{x_i(z) : i = 1,\ldots,k\}$ of elements of V which are continuous functions of z in a subset S of E^n, and a set H in E^k, determine an element (y^*,z^*) in H x S so as to minimize

$$\|f - (y_1x_1(z) + \ldots + y_kx_k(z))\| .$$

To the writer's knowledge,this class of problems is being treated for the first time in this work. Questions of existence, uniqueness,

and location of best approximations of this type are discussed in Sections 5 and 6. Except as otherwise noted, the theorems given there appear to be new. One concludes from these results that under fairly general conditions, this class of problems exhibits most of the nice features of constrained linear approximation problems.

In the second part, the geometry of the set of polynomials convex on $[0,1]$ is developed. The theorems which represent the convex polynomials in such a manner that the results of Part 1 are applicable are given in Sections 10 and 11. These theorems appear to be new; in essence, they say that the problem of convex polynomial approximation on an interval can be reduced to a problem of minimizing a function subject to linear constraints. Further, the problem exhibits features which make it amenable to treatment by readily available computational procedures. In particular, the function to be minimized cannot have a relative minimum even though it need not be convex. It is also shown that under certain conditions solutions can only lie on the boundary of the set of constraints.

Computation of best least squares approximations by convex polynomials is illustrated in an appendix.

1. Definitions and Notation.

Throughout, V will denote a normed linear space over the real numbers with norm $\| \|$; elements of V will be denoted with letters f, g, x. E^n will denote n-dimensional Euclidean space. It will be convenient to have two means of referring to coordinate systems in Euclidean space: the element y of E^{n+1} will be written in either

of the forms

$$\cdot\ \cdot(y_1,\cdots,y_{n+1}) \text{ or } \cdot\ \cdot(y_o,\cdots,y_n) \ .$$

Sets or sequences of points are denoted by capital letters such as A, B. The elements of a set or sequence are indicated by enclosing them in brackets, and the customary procedure of writing "X is the set of all x which have property P" as

$$X = (\ x : x \text{ has property P})$$

is used.

The closed interval $0 \leq t \leq 1$ will be written as $[0,1]$ or I. The finite set of points on I given by $0 = t_o < t_1 < \cdots < t_N = 1$ will be denoted by T. The linear spaces of real-valued continuous functions f(t) on I or T will be written as C(1) and C(T) respectively. C(I) is a Banach space with norm

$$\|f\| = \max_{t \in I} |f(t)| \ ;$$

C(T) is a Banach space with norm

$$\|f\| = \max_{0 \leq i \leq N} |f(t_i)| \ .$$

Other norms can be put on these linear spaces to obtain Banach spaces. The spaces $C_p(I)$, $p \geq 1$, are obtained with definition

$$\|f\|_p = \left[\int_0^1 |f(t)|^p dt\right]^{1/p}$$

and the spaces $C_p(T)$ are obtained when

$$\|f\|_p = \left[\sum_{i=0}^{N} |f(t_i)|^p \right]^{1/p} .$$

The same notation is used for both norms, but context will always make the usage clear.

Inner product notation for sums of products will be used whenever convenient: $y \cdot x = y_o x_o + \cdots + y_n x_n$ for x, y in E^{n+1}. With the convention that $x''(t) = (1, t, \ldots, t^n)$, polynomials $p(t)$ of degree $< n$ can be written in the form

$$p(t) = y \cdot x^n(t) = y_o + y_1 t + \cdots + y_n t^n .$$

If $p(t) > 0$ on a set S, it will simply be called positive on S; if $p(t) > 0$, it will be called strictly positive on S.

It is now possible to state the convex polynomial approximation problem: Given an $f(t)$ in $C(1)$ or $C(T)$ normed in one of the ways given above, an integer n, and a set H in E^{n+1}, determine an element y^* in H such that

$$\|Y \cdot x^n - f\|$$

is minimized at $y = y^*$ subject to the condition that

$$\frac{d^2}{dt^2} Y \cdot x^n(t) = 2y_2 + 6y_3 t + \cdots + n(n-1) y_n t^{n-2} > 0$$

for all t in I.

2. The Problem is Reasonable.

It is worthwhile to inquire as to whether the problem posed at the end of the last section is reasonable in the following sense: Given a function convex on $[0,1]$, are there polynomials convex on $[0,1]$ which are arbitrarily close to the function in some norm? If the answer is yes, the problem will be considered reasonable.

The desired affirmative answer is contained in the following

Theorem 2.1: Let $f(t)$ be a function which has positive k-th derivative on $[0,1]$. Then given any $\epsilon > 0$, there is a polynomial $p(t)$ with positive k-th derivative on $[0,1]$ such that

$$\max_{0 \leq t \leq 1} |p(t) - f(t)| < \epsilon .$$

The desired result is the special case of this theorem with $k = 2$.

The theorem is proved as a consequence of two other theorems, the first of which is S. Bernstein's version of the Weierstrass approximation theorem.

For a function $f(t)$ defined on $[0,1]$, the expression

$$B_n[f(t)] = \sum_{m=1}^{n} f(\tfrac{m}{n})\binom{n}{m}t^m(1-t)^{n-m}$$

is called the Bernstein polynomial of order n of the function $f(t)$. With this definition, one can obtain

Theorem 2.2 (S. Bernstein): If $f(t)$ is continuous on $[0,1]$, then

$$\lim_{n \to \infty} B_n[f(t)] = f(t)$$

uniformly on $[0,1]$.

Proof can be found in Natanson [14].

Now, define forward differences of $f(t)$ at $t = m/n$ by

$$\Delta f(\tfrac{m}{n}) = f(\tfrac{m+1}{n}) - f(\tfrac{m}{n}) \ ,$$

$$\Delta^k f(\tfrac{m}{n}) = \Delta(\Delta^{k-1} f(\tfrac{m}{n})) \ , \qquad k = 2,3,\ldots$$

Then, by direct differentiation and term rearrangement, the following expression for the k-th derivative of $B_n[f(t)]$ is obtained:

$$B_n^{(k)}[f(t)] = \frac{n!}{k!} \sum_{m=0}^{n-k} \Delta^k f(\tfrac{m}{n}) \binom{n-k}{m} t^m (1-t)^{n-m-k}$$

for $k = 1,2,\ldots,n$. If the k-th derivative of $f(t)$ is positive on $[0,1]$, then $\Delta^k f(\frac{m}{n})$ is positive for $0 < m \leq n - k$. This proves

Theorem 2.3: If $f(t)$ has positive k-th derivative on $[0,1]$, then the Bernstein polynomials of $f(t)$ have positive k-th derivative on $[0,1]$.

The proof of Theorem 2.1 follows directly from Theorems 2.2 and 2.3.

The above results are contained in Lorentz [11]. Convergence in the uniform norm implies convergence in any of the norms for $C[0,1]$ considered in Section 1.

Armed with the comforting knowledge that there are convex polynomials close to convex functions, it is now interesting to ask the following question: Among all n-th degree polynomials convex on $[0,1]$ is it possible to find best approximations to a given function (in particular, a convex function)? Here, "best" will mean the usual thing:

best in the sense that some norm is minimized.

To answer this question, the problem of Section 1 will be imbedded in a larger class of problems. To do this, and to facilitate discussion of the geometry of convex polynomials, the next two sections will be devoted to a development of results on convex sets and cones.

3. Convex Sets.

A set S in E^n is convex if for each pair of points y^1 and y^2 in S, the points $y = \theta y^1 + (1-\theta)y^2$ are in S, where $0 < \theta < 1$; that is, the line segment joining y^1 and y^2 lies in S. S is called strictly convex if $0 < \theta < 1$ causes the points y to lie in the interior of S. For a fixed vector x and a constant c, the plane E^n determined by

$$y \cdot x = y_1 x_1 + \cdots + y_n x_n = c$$

is called a supporting plane to S if the plane contains at least one point of S and S lies entirely in one of the half-spaces $\{y: y \cdot x \geq c\}, \{y: y \cdot x \leq c\}$. Any half-space containing S is called a supporting half-space to S. Theorems relating these concepts can be found in many places; for example, Karlin[8] proves:

Theorem 3.1: A closed convex set is the intersection of all of its supporting half-spaces, and every boundary point of the set lies on a supporting plane.

The dimension of a convex set S is defined as the dimension of the linear subspace of smallest dimension which contains S.

If the set S is convex, closed, bounded, and n-dimensional, it is called an n-dimensional convex body.

A supporting plane to a closed convex set S will be called proper if it contains exactly one point of S. In such a case, that point is called an extreme point of s. It is an immediate consequence of the definition that an extreme point cannot lie in the interior of a line segment joining two points of S.

The following theorems give relationships between convex sets, extreme points, and supporting planes. Their proofs can be found in Berge [2].

Theorem 3.2: If S is a compact non-empty convex set in E^n, it has an extreme point; further, every supporting plane of S contains an extreme point of S.

Theorem 3.3: A compact non-empty convex set S in E^n is the intersection of the closed convex sets containing the set of extreme points of S.

Theorem 3.4: If R and S are compact convex sets in E^n, each having at least one interior point, then R and S are homeomorphic.

4. Convex Cones.

A set K in E^n is a convex cone if for each pair of points y^1 and y^2 in K, the points $y = \alpha y^1 + \beta y^2$ are in K, where $\alpha, \beta > 0$. A convex cone is a convex set. The relationship of convex cones and supporting planes is shown by

Theorem 4.1: Let K be a closed convex cone in E^n. Then every

supporting plane to K contains the origin, and a supporting plane can be proper only at the origin.

Proof: Suppose $y \cdot x = c$ is a supporting plane to K at $y \neq 0$. Then $c = 0$, else there are points of K, namely αy for $\alpha < 1$ and $\alpha > 1$, on both sides of the plane. This shows that every supporting plane to K is of the form $y \cdot x = 0$ from which both parts of the theorem can be concluded.

For convenience, it will be supposed that any x which defines a supporting plane to a closed convex cone K is always taken so that K lies in the half-space given by $y \cdot x \geq 0$. The intersection of the translate $y \cdot x = 1$ of a supporting plane to K with K will be called a cross section of K. If K has a 'proper supporting plane, its corresponding cross section is called a proper cross section, and K is called pointed (the origin is an extreme point).

Theorem 4.2: Let K be a closed convex cone in E^n. A cross section of K is bounded if and only if it is proper.

Proof: For each fixed vector x, $y \cdot x$ is a continuous function on E^n. Define the set $S = \{y : y \in K, \|y\| = 1\}$ and let μ be the greatest lower bound of $y \cdot x$ on S. S is compact, so there exists a y in S for which $y \cdot x = \mu$. By convention, $y \cdot x > 0$ for each y in K, so $\mu \geq 0$. If $\mu > 0$, then for each y in S there is a number λ, $0 < \lambda \leq \frac{1}{\mu}$ such that $\lambda y \cdot x = 1$. This says that the cross section corresponding to x is bounded if and only if $\mu > 0$. If $\mu = 0$, there is a non-zero y in K such that $y \cdot x = 0$, which makes the supporting plane improper. The desired result follows immediately.

5. Existence.

Throughout this section and the next, $y = (y_1,\ldots,y_k)$ will denote a point of E^k, and $z = (z_1,\ldots,z_n)$ a point of E^n. The unit sphere

$$y_1^2 + \cdots + y_k^2 = 1$$

in E^K will be called U.

Achieser [1] gives the fundamental existence theorem for approximation in normed linear spaces as follows:

Theorem 5.1: Let $x_1,\ldots,x_k$ be k linearly independent elements of V. Then for any element f in V there exists a point y^* in E^k such that the function

$$\varphi(y) = \| y_1 x_1 + \cdots + y_k x_k - f \|$$

attains its greatest lower bound (and hence its minimum) at y^*.

Rice [16] shows that $\varphi(y)$ will also attain a minimum if y is constrained to lie in a closed set H in E^k.

The approximation problem under consideration involves the parameters nonlinearly. Thus, it would be useful to have an existence theorem which-covers the situationof interest and might also be applicable to other approximation problems. A rather general theorem is given by Young [22], and discussed by Rice [17], but appears difficult to apply. The theorem which will be given here is appropriate to the situation and is an extension of Theorem 5.1.

Definition 5.2: Let $x_1(z),\ldots,x_k(z)$ be k continuous functions on

E^n into V. Let S be a subset of E^n. If for each z in S, the set $B(z) = \{x_1(z),\ldots,x_k(z)\}$ is linearly independent in V, then the set $B(S) = \{B(z) : z \in S\}$ is called a basic set on S, or simply a basic set.

An example of a basic set is obtained by taking V = C(I), S the square in E^2 given by $0 < z_i \leq 1$, i = 1,2, and $x_1(z) = t^{z_1}$, $x_2(z) = t^{z_1+z_2}$.

Definition 5.3: Let B(S) be a basic set and define the function

$$\psi(y,z) = \|y \cdot x(z)\| = \|y_1 x_1(z) + \cdots + y_k x_k(z)\| .$$

Since $\psi(y,z)$ is positive, it has a greatest lower bound $\mu > 0$ on the set $U \times S$ in $E^k \times E^n$. If $\mu > 0$, B(S) is called an admissible basic set.

The example of a basic set given above is admissible. It would be tedious to show this by direct computation; however, the reason for the truth of the assertion is contained in the following

Lemma 5.4: Let B(S) be a basic set. If S is a compact set in E^n, then B(S) is admissible.

Proof: Since S is compact in E^n, $U \times S$ is compact in $E^k \times E^n$. The function $\psi(y,z)$ given in Definition 5.3 is continuous on $E^k \times E^n$ and hence attains its greatest lower bound μ on $U \times S$. Now, let (y^*,z^*) be a point in U x S such that

$$\psi(y^*,z^*) = \|y^* \cdot x(z^*)\| = \mu .$$

Since B(S) is a basic set, the $x_i(z^*)$ are linearly independent.

Hence $\mu = 0$ if and only if $y_i^* = 0$, all i. Since y^* is in U, $\mu \neq 0$ and B(S) is consequently admissible.

Theorem 5.5: Let f be an element of V, H a closed set in E^k, S a compact set in E^n. Let B(S) be a basic set in V with elements $x_1(z),\ldots,x_k(z)$. Then there exists an element (y^*,z^*) in $H \times S$ such that the function

$$\varphi(y,z) = \|y \cdot x(z) - f\|$$

attains its greatest lower bound on $H \times S$ at (y^*,z^*).

Proof: U is compact in E^k. Thus, for each fixed z in S, the continuous function

$$\psi(y,z) = \|y \cdot x(z)\|$$

attains its greatest lower bound $\mu(z)$ on U. By Lemma 5.4, $p(z) \geq \mu > 0$, where μ is the greatest lower bound of $\psi(y,z)$ on u x s. Also, observe that for any y in E^k and z in S,

$$\|y \cdot x(z)\| \geq \left(\sum_{i=1}^{k} y_i^2\right)^{1/2} \mu(z) \geq \left(\sum_{i=1}^{k} y_i^2\right)^{1/2} \mu \quad .$$

Now, let ρ be the greatest lower bound of $\varphi(y,z)$ on $H \times S$. By the inequality just derived,

$$\|y \cdot x(z) - f\| \geq \|y \cdot x(z)\| - \|f\| \geq \left(\sum_{i=1}^{k} y_i^2\right)^{1/2} \mu - \|f\| \quad .$$

Thus, if y is taken such that

$$\left(\sum_{i=1}^{k} y_i^2\right)^{1/2} > \frac{1}{\mu}(\rho + 1 + \|f\|) = r \, ,$$

then

$$\|y \cdot x(z) - f\| > \rho + 1 \quad .$$

This shows that only those y in the sphere

$$R = \left\{ y : \sum_{i=1}^{k} y_i^2 \leq r^2 \right\}$$

permit $\varphi(y,z)$ to approach ρ.

R is closed and bounded, so $R \cap H$ is closed and bounded and hence compact in E^k. S is compact in E^n, so $(R \cap H) \times S$ is compact in $E^k \times E^n$. Since $\varphi(y,z)$ is continuous on $E^k \times E^n$, it will attain its greatest lower bound on $(R \cap H) \times S$ which by the above argument is its greatest lower bound on $H \times S$, and this is what was to be proved.

Theorem 5.1 can be obtained from Theorem 5.5 as the special case when $x_1, \ldots, x_k$ are constant linearly independent elements of V.

Conditions under which solutions of approximation problems such as those under discussion are unique are discussed in the next section. Location of solutions is also discussed.

6. Uniqueness and Location of Solutions.

Achieser [1] proves a uniqueness theorem for linear approximation in a finite dimensional linear manifold in V under the condition that V is a strictly normalized space. This condition holds whenever equality in the expression

$$\|f + g\| \leq \|f\| + \|g\| \quad (f, g \neq 0)$$

holds only for $g = \alpha f (\alpha > 0)$.

Rice [16] gives more specific results. Let H be a closed set in E^k, and let $x_1, \ldots, x_k$ be fixed linearly independent elements of V. Let f be an element of V and suppose that

$$\min_{y \in E^k} \|y \cdot x - f\| < \min_{y \in H} \|y \cdot x - f\| .$$

Rice proves:

Theorem 6.1: (1) Every local minimum of $\|y \cdot x - f\|$ on H is a global minimum on H.

(2) If y^* minimizes $\|y \cdot x - f\|$ on H, then y^* is in the boundary of H.

(3) If H is strictly convex, then y^* is unique.

(4) If v is strictly normalized, then y^* is unique.

(Rice actually proves a slightly different statement than (4), but it is essentially the same in the present context.)

Some theorems similar to those of Rice can be proved under some assumptions on the nature of the mapping of H x S to the set of possible approximations in V.

Let Φ denote the mapping which associates the element $y \cdot x(z)$ in V with the element (y,z) in $E^k \times E^n$. Let $C = \Phi(H \times S)$. Henceforth, it will be assumed that C is a closed convex set in V. It will also be assumed that Φ sets up a 1 - 1 correspondence between H x S and C. If Φ is a homeomorphism between H x S and

C, then C will automatically be closed because H x S is closed.

Definition 6.2 (Riesz-Nagy [18]): A Banach space V is called uniformly convex if for each f, g in V such that $\|f\|, \|g\| \leq 1 + \epsilon$ and $\|f + g\| \geq 2$, then $\|f - g\| < \epsilon$.

It can be shown (Clarkson [5]) that of the spaces defined in Section 1, $C_2(I)$ is uniformly convex, but C(1) is not.

Theorem 6.3: If V is a uniformly convex space, then $\varphi(y,z) = \|y \cdot x(z) - f\|$ has a unique minimum in H x S.

Proof: Let $\{(y^n,z^n)\}$ be a minimizing sequence for φ. Let $g^n = y^n \cdot x(z^n)$ and ρ be the minimum of φ on H x S. Then given $\epsilon > 0$, there is an N sufficiently large so that for m, n > N,

$$\frac{\|g^n - f\|}{\rho}, \quad \frac{\|g^m - f\|}{\rho} < 1 + \epsilon .$$

Now, because C = Φ(H x S) is assumed convex, $\frac{1}{2}(g^n + g^m)$ is an element of C and

$$\left\| \frac{g^n + g^m}{2} - f \right\| \geq \rho ,$$

which implies

$$\left\| \frac{g^n - f}{\rho} + \frac{g^m - f}{\rho} \right\| \geq 2 .$$

By the assumed uniform convexity of V, it then follows that

$$\|g^n - g^m\| < \rho\epsilon ,$$

which shows that $\{g^n\}$, and hence any minimizing sequence in C, is a Cauchy sequence. By the completeness of V and the fact that C

is closed, this sequence converges to an element g of C. The element g is unique in C, for if $\{h^n\}$ is another minimizing sequence, then

$$g^1, h^1, g^2, h^2, \ldots, g^n, h^n, \ldots$$

is also a minimizing sequence which must converge to g. The assumption that Φ is a 1 - 1 correspondence implies the existence of a unique element (y,z) in $H \times S$ with $\varphi(y,z) = \rho$, which is what was to be proved.

Theorem 6.4: Let Φ be a homeomorphism. Then every local minimum of $\varphi(y,z)$ on $H \times S$ is a global minimum on $H \times S$.

Proof: Using the notation of the previous theorem, let g^1 and g^2 be elements of C such that $\|g^1-f\| < \|g^2-f\|$. The elements g of the line segment between g^1 and g^2 in C are given by the expression.

$$g = \theta g^1 + (1-\theta)g^2, \quad 0 \leq \theta \leq 1 \ .$$

By hypothesis, the points $\Phi^{-1}(g)$ lie on a continuous path from (y^1,z^1) to (y^2,z^2) in $H \times S$. Along this path, $\varphi(y,z)$ is monotone, since

$$\|\theta g^1 + (1-\theta)g^2 - f\| \leq \theta\|g^1-f\| + (1-\theta)\|g^2 - f\| \leq \|g^2 - f\| \ .$$

Now, let $\varphi(y,z)$ have a global minimum at (y^1,z^1) and a candidate for a local minimum at (y^2,z^2). Construct the path from (y^1,z^1) to (y^2,z^2) as indicated above. Because the path is continuous and φ is monotone along it, it is not possible for a relative minimum to

be at (y^2,z^2). This completes the proof.

There are several conditions which can cause Φ to be a homeomorphism. In particular, if H is compact, then Φ is a 1 - 1 continuous map from a compact space onto a Hausdorff space and hence a homeomorphism. Also, if C can be decomposed into a product A x B and Φ into a product $\Phi_1 \times \Phi_2$ such that Φ_1 is a homeomorphism of H onto A and Φ_2 is a homeomorphism of S onto B, then it can be shown that Φ is a homeomorphism.

It is now interesting to inquire about conditions which would force solutions to lie on the boundary of H x S. A set of conditions for this is given in

Theorem 6.5: Let Φ be a homeomorphism. Let H X S be closed, convex, and have interior points in E^k X E^n. Let (y^*,z^*) be a point of $E^k \times E^n$ such that (y^*,z^*) is not in H x S and $\varphi(y^*,z^*) < \varphi(y,z)$ for all (y,z) in H x S. Let V* be the smallest linear subspace of V which contains $g^* = y^*.x(z^*)$ and C. Then if C has an interior point in the relative topology in V*, the minimum points of φ on H x S must be on the boundary of H x S.

Proof: It is easily shown that Φ is a homeomorphism of H x S onto C considered as a subset of V*. Let $g^2 = y^2 \cdot x(z^2)$ be a candidate for a minimum in the relative interior of C corresponding to a point (y^2,z^2) in the interior of H X S (guaranteed by the homeomorphism). Construct the line segment from g^* to g^2. Because C is closed and convex with an interior, this line segment must meet the boundary of C in exactly one point which will be called g^1. By the same argument used in Theorem 6.4, $\|f-g\|$ is monotone along the

line segment g^* to g^2, and consequently is also monotone from g^1 to g^2. Under the homeomorphism, g^1 corresponds to a point (y^1,z^1) on the boundary of H x S and $\varphi(y^1,z^1) \leq \varphi(y^2,z^2)$. This completes the proof.

The remainder of this work is devoted to an example in which the foregoing theorems apply.

PART 2: CONVEX POLYNOMIALS

7. Methods of Representation.

Some kind of parametric representation of the set of polynomials of degree $< n$ which are convex on $[0,1]$ is needed before a computation of best convex polynomial approximation can be attempted. One such representation is suggested by Section 2: form Bernstein polynomials with coefficients whose second differences are positive. The second difference expressions will yield a finite set of linear inequalities which the coefficients must satisfy, which is desirable, but this method will be rejected since it can be shown that not all polynomials of degree $\leq n$ which are convex on $[0,1]$ can be represented exactly by Bernstein polynomials of degree $\leq n$ (see Section 12).

Another method would be the direct method of Section 1: make the polynomial $y \cdot x''(t)$ satisfy the infinite set of constraints

$$2y_2 + 6y_3 t + \cdots + n(n-1)y_n t^{n-2} \geq 0$$

for each t in $[0,1]$. This is the method found intractable by Rice [16].

The method which will be adopted here derives from the existence of a parametrization of the set of polynomials of degree $< n$ which are positive on $[0,1]$. It has the desirable property that the parameters must satisfy a finite set of linear constraints. This representation can be integrated twice to obtain a representation of the polynomials of degree $< n + 2$ which are convex on $[0,1]$.

8. The Cone of Positive Polynomials.

The results of this section and the next were obtained by Karlin and Shapley [9] by less direct means.

The point $y = (y_0, y_1, \ldots, y_n)$ in E^{n+1} representing the polynomial $y \cdot x^n(t) = y_0 + y_1 t + \cdots + y_n t^n$ corresponds to a polynomial positive on $[0,1]$ when $y \cdot x^n(t) \geq 0$ for each t in $[0,1]$. Let K^n denote the set of all y in E^{n+1} which have that property.

Theorem 8.1: K^n is a closed convex cone in E^{n+1} whose boundary consists of points representing polynomials of degree $\leq n$ which have roots in $[0,1]$ but are otherwise positive there.

Proof: If P_1 and P_2 are polynomials of degree $< n$ which are positive on $[0,1]$, then so also are the polynomials $\alpha p_1 + \beta p_2$ for all $\alpha, \beta \geq 0$; hence, K^n is a convex cone. Since a polynomial is a continuous function of its coefficients, a polynomial $p(t)$ which is strictly positive on $[0,1]$ will remain so in an open neighborhood about its coefficient point in E^{n+1}; hence, that point must lie in the interior of K^n. If $p(t)$ is positive but has a root at t_0 in $[0,1]$, then each open neighborhood of its coefficient point contains a point corresponding to a polynomial which is negative at t_0; hence, $p(t)$ corresponds to a boundary point of K^n. Since K^n contains its boundary, it is closed.

Corollary 8.2: The planes of the form

$$p(t_0) = y_0 + y_1 t_0 + \cdots + y_n t_0^n = 0 ,$$

where $p(t)$ is positive with a root at t_0 on $[0,1]$ are supporting planes to K^n.

Proof: If $q(t)$ is a positive polynomial of degree Cn on $[0,1]$, then $q(t) \geq p(t_o) = 0$, so K^n lies to one side of the plane $p(t_o) = 0$. By hypothesis, $p(t)$ corresponds to a point in the plane, so $p(t_o) = 0$ is a supporting plane.

If $p(t)$ has a root at t_o on $[0,1]$, then so does $\alpha p(t)$ for all $\alpha > 0$. Thus, the supporting planes of the form $p(t_o) = 0$ cannot be proper. K^n does have a proper supporting plane, however. This fact is used to prove

Theorem 8.3: K^n is pointed.

Proof: It will be shown that the plane

$$y_o + \frac{1}{2} y_1 + \cdots + \frac{1}{n+1} y_n = 0$$

is a proper supporting plane to K^n. First, the plane meets K^n at the origin. Second, if $y \neq 0$ is in K^n, then $p(t) = y \cdot x^n(t) > 0$ for t in $[0,1]$, but $p(t)$ is not identically zero, so

$$y_o + \frac{1}{2} y_1 + \cdots + \frac{1}{n+1} y_n = \int_0^1 p(t)dt > 0 .$$

The rest of the proof follows immediately from Theorem 4.1 and the definitions of Section 3 and 4.

9. The Cross Section P^n.

Theorem 8.3 implies that K^n has a proper cross section defined by the intersection of K^n with the plane

$$y_o + \frac{1}{2} y_1 + \cdots + \frac{1}{n+1} y_n = 1 .$$

This cross section will be called P^n and will be described in detail.

Theorem 9.1: P^n is an n-dimensional convex body.

Proof: It must be shown that P^n is convex, closed, bounded, and n-dimensional.

P^n is closed and convex because it is the intersection of two closed convex sets. By Theorem 4.2, P^n is bounded. To show that P^n is n-dimensional, observe that the points in P^n corresponding to the polynomials $1, 2t, 3t^2, \ldots, (n+1)t^n$ lie in the plane defining the cross section. Thus, the n vectors

$$\begin{aligned} &(-1,2,0,0,\ldots, 0), \\ &(-1,0,3,0,\ldots, 0), \\ &\qquad\cdots \\ &(-1,0,0,0,\ldots, 0, n+1), \end{aligned}$$

formed by subtracting the vector to the first point from those to the others, all lie in the plane of the cross section and are clearly linearly independent. The dimension of the plane must therefore be at least n. Since the dimension of the plane must also be $< n + 1$, the proof is completed.

Theorem 3.3 says that to describe P^n, it suffices to describe its set of extreme points. The nature of the extreme points of P^n is given by

Theorem 9.2: The extreme points of P^n correspond to polynomials which have n roots (counting multiplicities) on $[0,1]$.

Proof: Each polynomial corresponding to an extreme point of P^n must be of degree n exactly. To see this, suppose p(t) corresponds

to an extreme point but is of degree $< n$. Then the polynomials $tp(t)$ and $(1-t)p(t)$ are positive on $[0,1]$ and both are of degree $< n$. It is then clear that positive scaling factors a_1 and a_2 can be found so that $a_1 tp(t)$ and $a_2(1-t)p(t)$ correspond to points of P_n, and further, there will be a θ, $0 < \theta < 1$, so that

$$p(t) = \theta a_1 tp(t) + (1-\theta)\, a_2(1-t)p(t) ,$$

which contradicts the hypothesis that $p(t)$ corresponds to an extreme point.

Now, if $p(t)$ is positive on $[0,1]$ but does not have all of its roots there, then its corresponding point in K^n cannot be an extreme point of P^n, for $p(t)$ must then have a root $a < 0$, a root $b > 1$, or a pair of complex roots $c \pm id$. This implies that $p(t)$ is expressible in one of the forms

$$p(t) = (t-a)u(t) = \frac{1}{2}(t-2a)u(t) + \frac{1}{2} tu(t) ,$$

$$p(t) = (b-t)v(t) = \frac{1}{2}(2b-1-t)v(t) + \frac{1}{2}(1-t)v(t) ,$$

$$p(t) = [(t-c)^2 + d^2]w(t) = (t-c)^2 w(t) + d^2 w(t) ,$$

where u, v, and w are polynomials positive on $[0,1]$. All three of the right-hand expressions can be scaled so that they are of the form $\theta p_1(t) + (1-\theta)p_2(t)$ with $0 \leq \theta \leq 1$ and p_1 and p_2 corresponding to points in P^n. This proves one half of the theorem.

Now, suppose $p(t)$ is a polynomial corresponding to an extreme point of P^n, and that there are polynomials $p_1(t)$, $p_2(t)$

corresponding to points of P^n and θ, $0 < \theta < 1$, so that $p(t) = \theta p_1(t) + (1-\theta)p_2(t)$. Because p_1 and p_2 must each have the same roots as p, they must be identical, for $p(t)$ already has the maximum possible number of roots. Thus, the supposed convex combination is impossible, and this completes the proof.

Knowing the permissible disposition of all of the roots makes it possible to write down polynomials proportional to those corresponding to extreme points of P^n. Any roots in the interior of $[0,1]$ must be of even order; Roots of odd order can occur only at 0 and 1. Hence for n even $(n=2m)$, the extreme polynomials are

$$\prod_{j=1}^{n} (t-z_{2j-1})^2 \quad \text{or} \quad t(1-t) \prod_{j=1}^{m-1} (t-z_{2j})^2 ,$$

and for n odd $(n=2m+1)$, they are

$$t \prod_{j=1}^{m} (t-z_{2j})^2 \quad \text{or} \quad (1-t) \prod_{j=1}^{m} (t-z_{2j-1})^2 ,$$

where the z_i are in $[0,1]$ and need not be distinct. The subscripts were taken as shown for later convenience.

One would expect that a convex linear combination of n + 1 extreme points would be required to represent an arbitrary point of P^n. However, it is a remarkable fact that every point in P^n, and hence any point of K^n, can be represented by a unique positive linear combination of at most two extreme points, and the extreme points can be chosen in a completely systematic manner. That this is so is stated in

Theorem 9.3 (Karlin-Shapley): Every polynomial corresponding to a point y of P^n has a unique representation by a pair of polynomials corresponding to extreme points of P^n as follows:

$$\sum_{i=0}^{n} y_i t^i = \alpha \prod_{j=1}^{m} (t-z_{2j-1})^2 + \beta t(1-t) \prod_{j=1}^{m-1} (t-z_{2j})^2$$

if $n = 2m$, and

$$\sum_{i=0}^{n} y_i t^i = \alpha t \prod_{j=1}^{m} (t-z_{2j})^2 + \beta(1-t) \prod_{j-1}^{m} (t-z_{2j-1})^2$$

if $n = 2m + 1$, with $\alpha > 0$, $\beta \geq 0$, $0 \leq z_1 \leq z_2 \cdots \leq z_{n-1} \leq 1$. Moreover, y is interior to P^n if and only if all of the inequalities are strict. Note that α and β are not independent. They are actually of the form $\alpha = \alpha' z_n$, $\beta = \beta'(1-z_n)$, $0 \leq z_n \leq 1$, where α' and β' are scaling factors which make the corresponding extreme points lie in P^n.

The proof of this theorem is too lengthy to repeat here. See Karlin and Shapley [9]. Note that each point in the simplex in E^{n-1} defined by $0 \leq z_1 \leq \cdots \leq z_{n-1} \leq 1$ generates two linearly independent polynomials proportional to polynomials corresponding to extreme points of P^n. A sketch of the cross section P^2 is shown in Figure 1.

Corollary 9.4: Every polynomial corresponding to a point y of K^n has a unique representation of the same type as that given in Theorem 9.3. Here, α and β may be regarded as independent.

Proof: Every element of K^n is a positive multiple of an element in P^n.

The plane of the page is

$$y_0 + \frac{1}{2} y_1 + \frac{1}{3} y_2 = 1.$$

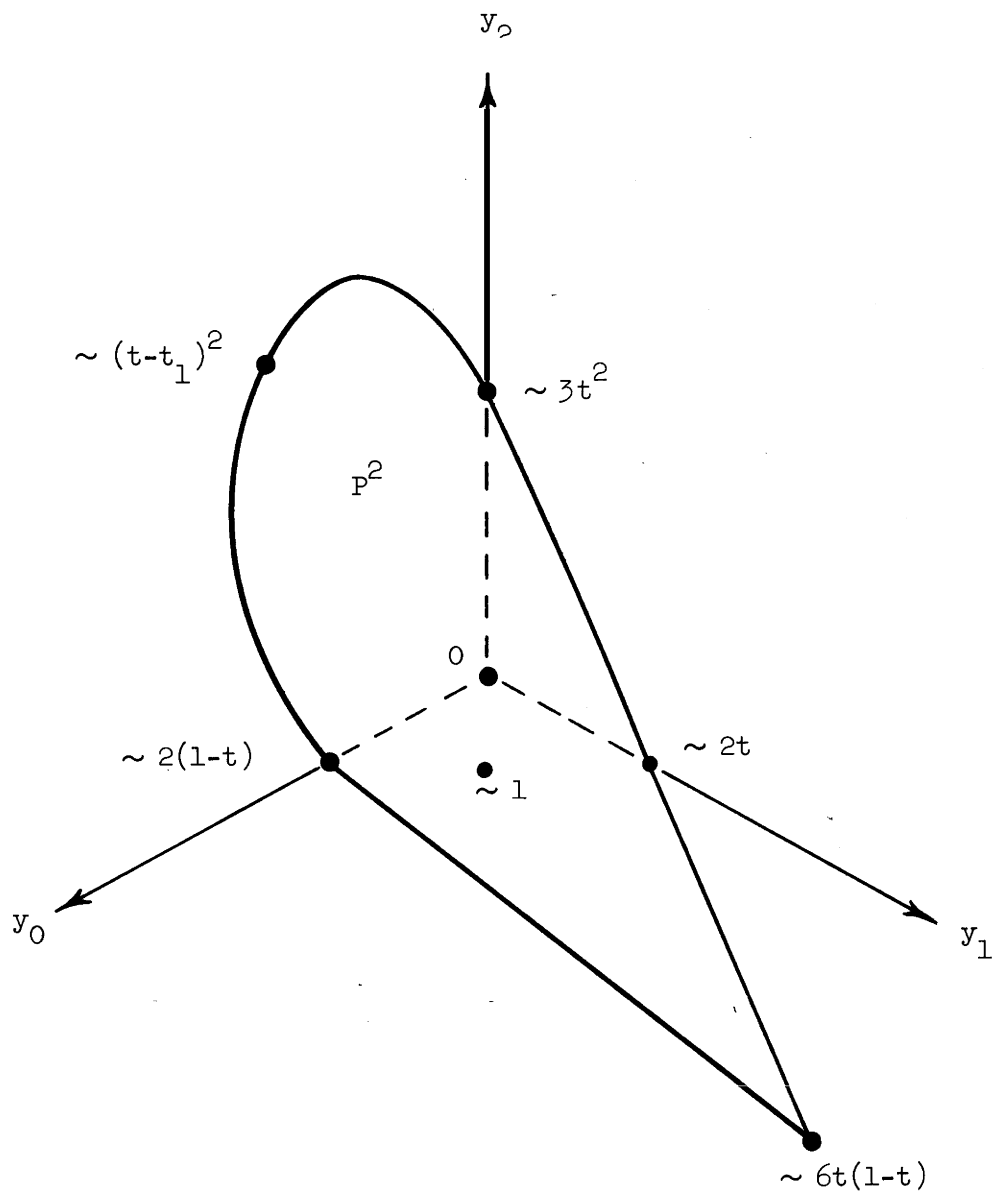

Figure 1

The cross section P^2

The representation of Theorem 9.3 will be used to generate the desired parametric representation of the convex polynomials.

10. Convex Polynomials.

Let Z^n denote the set of those z in E^n whose components satisfy the relations $0 \leq z_1 < \cdots \leq z_{n-1} \leq 1$, $0 \leq z_n \leq 1$. It is clear that Z^n is a compact convex set. Define the mapping ξ from Z^n to P^n by $\xi(z)$ = the element in P^n corresponding to the polynomial given by Theorem 9.3.

Now, define the mapping η from E^{n+1} to E^{n+3} as follows:

$$\eta(y_o, \blacksquare\blacksquare\blacksquare) = \left(0,0, \frac{1}{2\cdot 1} y_o, \frac{1}{3\cdot 2} y_1, \ldots, \frac{1}{(n+2)(n+1)} y_n\right) .$$

Under the corresponding relation between polynomials, the polynomials of degree < n are mapped into their indefinite double integrals. Let $Q^{n+2} = \eta(P^n)$.

Theorem 10.1: Q^{n+2} is an n-dimensional convex body homeomorphic to Z^n.

Proof: By Theorem 9.3, ξ is a 1 - 1 continuous map of Z^n onto P^n. Z^n is compact and P^n is Hausdorff, so ξ is a homeomorphism. Also, η is a linear 1 - 1 continuous map of P^n onto Q^{n+2} from which the rest of the proof follows.

The polynomials corresponding to points in Q^{n+2} can be realized as images of points in Z^n in the form $q(z,t) = \eta(\xi(z)).x^{n+2}(t)$, where $x^{n+2}(t) = (1,t,\ldots,t^{n+2})$. Let C^{n+2} denote the set of polynomials of degree $\leq n + 2$ which are convex on $[0,1]$.

Theorem 10.2: Each element of C^{n+2} has a unique representation of the

form

$$y_o + y_1 t + y_2 q(z,t)$$

with $q(z,t)$ defined as above and (y_o,y_1,y_2) a point in E^3 subject to the condition $y_2 \geq 0$.

Proof: Let $p(t)$ be convex on $[0,1]$. There is no loss of generality if it is supposed that the degree of $p(t)$ is exactly $n + 2$. Let $p'''(t)$ be the second derivative of $p(t)$. By Corollary 9.4, there is a scale factor $y_2 > 0$ so that $p''(t)/y_2$ corresponds to a point of P^n and a unique point z of Z^n which represents that element of P^n. By Theorem 10.1, the point z determines a unique element of Q^{n+2} and a corresponding polynomial $q(z,t)$. It follows that $y_2 q(z,t)$ agrees. with $p(t)$ except for the terms y_o and $y_1 t$ which are absent from $y_2 q(z,t)$. The rest of the proof follows easily.

In the proof of the last theorem, it is observed that the degree of any of the polynomials $q(z,t)$ is > 2. Thus, for each fixed z, the set $\{1,t,q(z,t)\}$ is linearly independent in the space of polynomials of degree $\leq n + 2$. Since Z^n is compact, this proves

Theorem 10.3: $\{1,t,q(z,t)\}$ is an admissible basic set.

Now define $H = \{y : y = (y_o,y_1,y_2)$ in E^3, $y_2 > 0\}$. H is closed. Define the mapping Φ from $H \times Z^n$ to $C(1)$ by the expression

$$\Phi(y,z) \equiv y_o \circ y_1 t \circ y_2 q(z,t) .$$

Theorem 10.4: The mapping Φ is a homeomorphism of $H \times Z^n$ onto C^{n+2}.

Proof: Let $E^+ = \{y_2 : y_2 \geq 0\}$, $E^1(t) = \{y_o + y_1 t : (y_o, y_1) \in E^2\}$, and C_o^{n+2} = the set of all polynomials of C^{n+2} with $y_o = y_1 = 0$. Then the mapping Φ can be considered as a mapping from $E^2 \times (E^+ \times Z^n)$ to $E^1(t) \times C_o^{n+2}$. Now, Φ can be decomposed into the form $\Phi_1 \times \Phi_2$ where Φ_1 maps E^2 to $E^1(t)$ and Φ_2 maps $E+ \times Z^n$ to C_o^{n+2}. By definition Φ_2 is 1 - 1, continuous, and onto C_o^{n+2}. A product of an open interval in E^+ and an open set in Z^n is mapped to an open set in C_o^{n+2} because Q^{n+2} is homeomorphic to Z^n. Thus, Φ_2 is an open mapping and consequently a homeomorphism. Φ_1 is a homeomorphism by definition From the remarks following Theorem 6.4, it follows that Φ is a homeomorphism.

11. Convex Polynomial Approximation.

Theorem 10.2 isolates the class of convex polynomials and Theorems 10.3 and 5.5 establish the fact that the best approximations exist within the class Furthermore, Theorems 10.4 and 6.4 give assurance that during computation of best convex approximations to f(t), if a local minimum of the function $\|y_o + y_1 t + y_2 q - f\|$ is found, then it is a solution to the problem.

Now, observe that with the definitions of H and Z^n given in Section 10, $H \times Z^n$ is a closed convex set with interior in E^{n+3}. Observe also that C^{n+2} is a convex set of dimension $n + 3$ in either C(1) or C(T). Thus, the linear subspace of either of these spaces generated by C^{n+2} is just the set of all polynomials of degree $\leq n + 2$, and in this subspace C^{n+2} has interior points (by an extension of Theorem 8.1). Thus, an immediate application of Theorem

6.5 yields

Theorem 11.1: Let f be an element of C(1) or C(T). Suppose the best approximation to f by polynomials of degree < n + 2 in one of the norms of Section 1 is not convex. Then the best convex polynomial approximation to f is obtained on the boundary of $H \times Z^n$.

In computational practice, one may as well allow H to be all of E^3, in which case either the best convex or the best concave polynomial approximation will be found. Since best approximations must occur in a compact part of E^3, application of Theorem 11.1 implies that all solutions are on the boundary of Z^n whenever the unconstrained best approximation is not already convex or concave.

Computational examples are described in the Appendix.

12. A Note on the Bernstein Polynomials; Some Unsolved Problems.

A look at Figure 1 shows that it is impossible to express the polynomial $(t-\frac{1}{2})^2$ as a positive linear combination of the polynomials t^2, $t(1-t)$, and $(1-t)^2$. Thus, it is not in general possible to obtain a best approximation by positive polynomials of degree $\leq n$ by taking positive linear combinations of the polynomials $t^k(1-t)^k$, $K = 0,1,\ldots,n$. The set of polynomials just referred to is linearly independent, so any polynomial of degree $\leq n$ can be represented as a linear combination of them. However, conditions on the coefficients making the polynomial positive are not known. This is an interesting problem which would bear investigation.

For reasons much the same as in the positive polynomial case, the attempt to represent all polynomials convex on [0,1] by linear

combinations of the same kind with a condition on the second differences of the coefficients will fail.

Another difficulty with the ordinary Bernstein polynomials is that no matter how many derivatives the parent function has, the order of convergence of $B_n(f)$ to f is $o(\frac{1}{n})$. See Voronowskaja [21] or Lorentz [11]. Butzer [4] has shown that certain linear combinations of the ordinary Bernstein polynomials converge to f like n^{-k} if f is bounded and has $2k$ derivatives on $[0,1]$. The question of whether Butzer's polynomials exhibit properties like that of the parent function is also open.

Now that best convex polynomial approximations can be computed, the problem of order of convergence estimation for these approximations becomes more interesting and should be investigated. However, no course of attack is immediately evident.

APPENDIX: COMPUTATIONAL EXAMPLES

The spaces $C_2(I)$ and $C_2(T)$ defined in Section 1 are uniformly convex, so best convex polynomial approximations in these spaces will be unique. Furthermore, the functions of the form $\|f - y \cdot x(z)\|_2^2$ which are to be minimized are differentiable functions of the parameters in the cases to be considered. One example will illustrate approximation in $C_2(I)$, the other in $C_2(T)$.

A1. Convex Cubic Approximation in $C_2(I)$.

This case can be solved exactly. This is facilitated by the use of the Legendre polynomials on the interval $[0,1]$, the first four of which are (see Milne [13])

$$P_0(t) = 1 ,$$

$$P_1(t) = 1 - 2t ,$$

$$P_2(t) = 1 - 6t + 6t^2 ,$$

$$P_3(t) = 1 - 12t + 30t^2 - 20t^3 .$$

These polynomials are orthogonal on $[0,1]$; in fact, they satisfy the relationship

$$\int_0^1 P_i(t)P_j(t)dt = \begin{cases} 0 , & i \neq j \\ (2j+1)^{-1}, & i = j . \end{cases}$$

They are linearly independent, forming a complete orthogonal set; hence any polynomial of degree n can be written as a unique linear combination of the first $n + 1$ of them.

Polynomial approximations of the third degree to $f(t)$ on $[0,1]$ are obtained by minimizing

$$\left\| f - \sum_{i=0}^{3} y_i P_i \right\|_2^2 = \int_0^1 [f(t)]^2 dt - 2 \sum_{i=0}^{3} y_i \int_0^1 f(t) P_i(t) dt + \sum_{i=0}^{3} \frac{y_i^2}{2i+1} ,$$

where the right hand side has been obtained by using the orthogonality relations. This expression is quadratic in the y_i, and by completing squares it is easily shown that its minimum value is

$$\int_0^1 [f(t)]^2 dt - \sum_{i=0}^{3} \frac{y_i^2}{2i+1} , \qquad (A1)$$

which is obtained for

$$y_i = \frac{\int_0^1 f(t) P_i(t) dt}{1/2i+1} ; \quad i = 0,1,2,3 . \qquad (A2)$$

Now, let it be required that the approximation be convex on $[0,1]$. This condition is expressed as

$$\frac{d^2}{dt^2} \sum_{i=0}^{3} y_i P_i(t) = 12y_2 + (60-120t)y_3 > 0 ,$$

or,

$$y_2 + 5(1-2t)y_3 \geq 0 ; \quad 0 \leq t \leq 1 .$$

What this means geometrically is shown in Figure 2, where the shaded region is the intersection of all of the half-spaces given by the constraints. The boundary lines of the cone of possible solutions are given by $y_2 \pm 5y_3 = 0$.

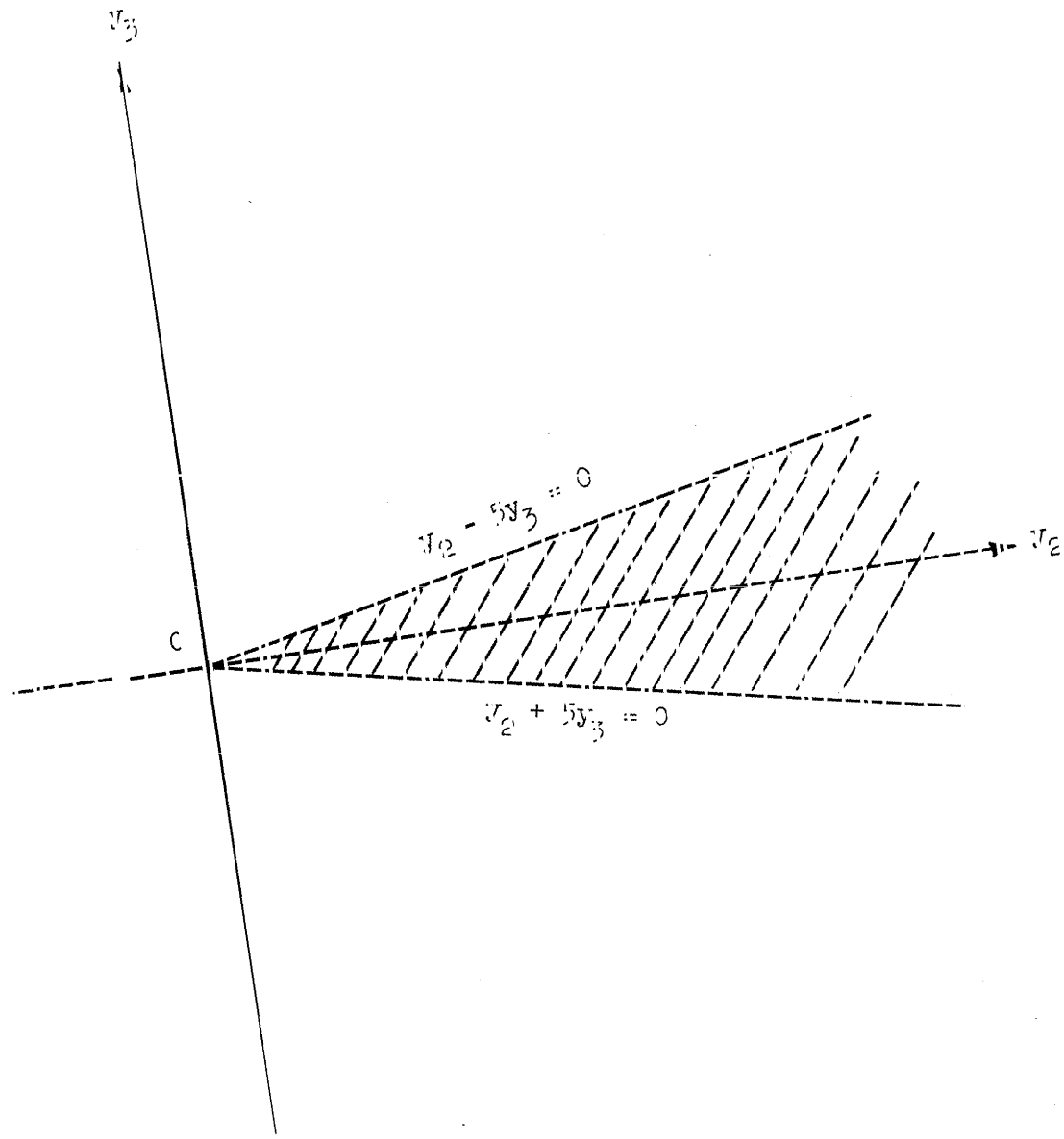

Figure 2

The shaded area represents convex polynomials

Now, supposing it is known that the best unconstrained least squares cubic approximation to f(t) is not convex. Then one can conclude by applying Theorem 11.1 or Theorem 6.1(2) that the solution must lie on the boundary lines. The expression to minimize then becomes

$$\int_0^1 \left\{f(t) - [a_o P_o(t) + a_1 P_1(t) + y_2 P_2(t) \pm \tfrac{1}{5}\, y_2 P_2(t)]\right\}^2 dt \ ,$$

and again it is easy to show by completing squares that the minimum is

$$\int_0^1 [f(t)]^2 dt - \left\{a_o^2 + \tfrac{1}{3} a_1^2 + \tfrac{1}{5} a_2^2 + \tfrac{1}{7} (\tfrac{1}{5} \pm a_2)^2\right\} \ , \qquad \text{(A3)}$$

and that the minimum is obtained for

$$y_i = (2i+1) \int_0^1 f(t) P_i(t) dt \ , \qquad i = 0,1 \ ;$$

$$Y2 = \frac{\int_0^1 f(t)\, [P_2(t) \pm \frac{1}{5} P_3(t)]dt}{\frac{1}{5} + \frac{1}{7}(\pm)5} \ . \qquad \text{(A4)}$$

Two solutions are possible from equations (A3) and (A4); the correct one is that which gives the smallest value in (A3).

To illustrate, consider the problem of obtaining the best convex cubic approximation on [0,1] to $f(t) = t^4$. Using equations (A2) it is found that the best approximation is

$$\tfrac{1}{5} P_o - \tfrac{2}{5} P_1 + \tfrac{2}{7} P_2 - \tfrac{1}{10} P_3 \ ,$$

or,

$$-\frac{1}{70} + \frac{2}{7}t - \frac{9}{7}t^2 + 2t^3 \ .$$

The mean square error from equation (A2) is found to be approximately 2.3×10^{-5}. It is easily shown that this approximation is not convex.

Applying equations (A4) with minus sign (which is seen to be correct by plotting the point of best unconstrained approximation in Figure 2), one obtains

$$\frac{1}{5}P_0 - \frac{2}{5}P_1 + \frac{7}{24}P_2 - \frac{7}{120}P_3 \ ,$$

or,

$$\frac{1}{30} - \frac{7}{20}t + \frac{7}{6}t^3 \ .$$

From equation (A3), the mean square error obtained is approximately $2.8 \cdot 10^{-4}$.

Approximation in $C_2(T)$ can be handled in essentially the same manner using the orthogonal polynomials described by Forsythe [6].

A2. Convex Quartic Approximation in $C_2(T)$.

By application of Theorems 9.3 and 10.2 for the case n = 2, every polynomial of degree < 4 which is convex on $[0,1]$ can be represented in the form

$$p(y,z,t) = y_0 + y_1 t + Y2 \int_0^{\tau} \int_0^{t} [z_2(t-z_1)^2 + (1-z_2)t(1-t)]dt^2 \ , \quad (A5)$$

with $y_2 \geq 0, \quad 0 \leq z_1 < 1, \quad 0 \leq z_2 < 1.$

Thus, best least squares convex approximations to functions f(t) in $C_2(T)$ are obtained by minimizing

$$\|f - p\|_2^2 = \sum_{i=1}^{N} [f(t_i) - p(y,z,t_i)]^2 \qquad (A6)$$

subject to the constraints. In Section 11 it was pointed out that the constraint $y_2 \geq 0$ need not be applied in practice, so only the bounds on , z_1 and z_2 will be used.

One might now proceed by trying to solve the problem using the method of Lagrange multipliers.

However, the equations so obtained will be non-linear and difficult to solve, thus it seems worthwhile to use a numerical procedure from the start. Fortunately, such procedures are available, and many are programmed for digital computers The method to be employed here is the gradient projection method of Rosen [19]. It has been programmed for use on the IBM 7090 computer by Merrill [12]. For use on the problem at hand, a subprogram for evaluating expression (A6) and its gradient on the parameter space must be supplied. The program is already ahle to handle the constraints. A subprogram has been written for the following test problem:

$$T = \{t_i : t_i = 0.1i \ ; \ i = 0,1,2,\ldots,10\} \ ,$$

$$f(t_i) = e^{-7t_i} \ .$$

For purposes of comparison, and to obtain starting approximations for the gradient projection code, best unconstrained quartic approximations for this test case were computed. This was done using the

method described in Forsythe [6] and an IBM 7094 computer code based on the program described in Rudin [20]. The second, third, and fourth degree approximations and the corresponding sums of the squared errors were computed as follows:

$$\text{Second degree,} \qquad \Sigma\, \epsilon_i^2 = 0.092146842 \ ,$$

$$0.82273361 - 2.5890284t + 1.8636677t^2 \ .$$

$$\text{Third degree,} \qquad \Sigma\, \epsilon_i^2 = 0.013453531 \ ,$$

$$0.95122132 - 4.6305554t + 7.2173224t^2 - 3.5691031t^3 \ .$$

$$\text{Fourth degree,} \qquad \Sigma\, \epsilon_i^2 = 0.0012569747 \ ,$$

$$0.99040337 - 5.9910430t + 14.019760t^2 - 14.453004t^3 + 5.4419509t^4 \ .$$

The third and fourth degree approximations are not convex. Thus, the best convex approximations in these cases must lie on the boundary of the constraint set.

However, in the first application of the gradient projection method, the solutions were not constrained to lie on the boundary of Z^2 (see Section 10), but allowed to range over all of Z^2. No other constraints were applied. As a starting guess, the above second degree approximation was used, for it is convex.

Convergence towards a minimum was very slow, despite various accelerating options in the program that were applied. After some 3500 iterations, the following result was obtained:

$$y_o = 0.92700600 ,$$

$$y_1 = -4.7174867 ,$$

$$y_2 = 29.390549 ,$$

$$z_1 = 0.78380506 ,$$

$$z_2 = 1 .$$

, The corresponding polynomial and sum of squared errors are:

$$p(y,z,t) = 0.92700600 - 4.7174867t + 9.0280473t^2 - 7.6788203t^3 + 2.4492124t^4 ,$$

$$\Sigma \epsilon_i^2 = 0.0092285508 .$$

This result is better than the best third degree (unconstrained) polynomial approximation, but not as good as the best fourth degree approximation, which gives a lower bound for the error. Note that this result is on the boundary of Z^2.

Some subsequent computations were made forcing the solutions to lie on the boundary of Z^2, but unless the starting approximation was close to the one found above, convergence was also quite slow. It appears that slow convergence is the price that one must pay for the lack of convexity of the expression (A6).

It should be mentioned, however, that the long computation referred to above took 12 minutes on the IBM 7090 computer.

Perhaps better (e.g., faster) computational procedures can be found; however, the principal aim here has been to demonstrate the possibility of solving such problems by practical means. This has been accomplished.

BIBLIOGRAPHY

1. Achieser, N. I., Theory of Approximation, Frederick Ungar, 1956, Chapter 1.

2. Berge, C., Topological Spaces, Macmillan, 1963, Chapter 8.

3. Boltjanskii, V. G., "Application of the theory of optimal processes to problems of approximation of functions," (Russian) Trudy Mat. Inst. Steklov. 60 (1961), 82-95.

4. Butzer, P. L., "Linear combinations-of Bernstein polynomials," Canadian J. Math., 5 (1963) pp. 559-567.

5. Clarkson, J. A., "Uniformly convex spaces,," Trans. Amer. Math. Soc., 40 (1936), pp. 396-414.

6. Forsythe, G. E., "Generation and use of orthogonal polynomials for data-fitting with a digital computer," J. Soc. Indust. Appl. Math., 5 (1957), pp. 74-88.

7. Hamming, R. W., Numerical Methods for Scientists and Engineers, McGraw-Hill, 1962, Chapters 17, 21.

8. Karlin, S., Mathematical Methods and Theory in Games, Programming, and Economics, Vol. I, Addison-Wesley, 1959, Appendix B.

9. Karlin, S. and L. S. Shapley, Geometry of Moment Spaces, Memoirs Amer. Math. Soc., No. 12, 1953, Secs. 9, 10.

10. Lanczos, C., Applied Analysis, Prentice-Hall, 1956, Chapter 5.

11. Lorentz, G. G., Bernstein Polynomials, Univ. Toronto Press, 1953, Chapter 1.

12. Merrill, R. P., "Gradient Projection (GP90)," SHARE Code GP90 SHARE Secretary's Distribution No. 1399, 4 January 1963.

13. Milne, W. E., Numerical Calculus, Princeton Univ. Press, 1949, Chapter 9.

14. Natanson, I. P., Konstruktivnaya Teoriya Funktsii (Russian), Gostekhizdat, 1949, Chapter 1.

15. Pontrjagin, L. S., V. G. Boltjanskii, R. V. Gamkrelidze, and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, Interscience, 1962, pp. 197-213.

16. Rice, J. R., "Approximation with convex constraints," J. Soc. Indust. Appl. Math., 11 (1963), pp. 15-32.

17. Rice, J. R., Approximation of Functions, Vol. I, Addison-Wesley, 1964, Chapter 1.

18. Riesz, F. and B. Sz.-Nagy, Functional Analysis, Frederick Ungar, 1955, Sec. 145.

19. Rosen, J. B., "The gradient projection method for nonlinear programming, Part 1, linear constraints," J. Soc. Indust. Appl. Math., 8 (1960), pp. 181-217.

20. Rudin, B. D., 'On routines for least squares curve fitting with orthogonal polynomials," Lockheed Missile and Space Co. Technical Report, LMSD 310956, 1957.

21. Voronowskaja, E., "De/termination de la forme asymptotique d'approximation des fonctions par les polynomes de M. Bernstein," C. R. Acad. Sci. U.R.S.S. (1932) pp. 79-85.

22. Young, J. W., "General theory of approximation by functions involving a given number of arbitrary parameters," Trans. Amer. Math. Soc., 8 (1907), pp. 331-344.