JOURNAL OF APPROXIMATION THEORY 30, 197–202 (1980)

Non-Linear Families of Projections on $C[-1, 1]$

W. A. LIGHT

Department of Mathematics, University of Lancaster, Cartmel College, Bailrigg, Lancaster LA1 4YL, England

Communicated by E. W. Cheney

Received May 8, 1979

Many approximation processes can be regarded as defining linear projections on a suitable normed linear space, usually the space of continuous functions on some closed interval of the real line. In this case the norm of the projection gives an estimate for how well the process will perform in practice. Numerical evidence shows that amongst ultraspherical projections, the Chebyshev projection (arising from the truncated Chebyshev series) does not have minimal norm. In this paper we demonstrate this fact analytically by deriving first some general principles, and then applying these to the Chebyshev projection.

1. INTRODUCTION

A projection from a normed linear space X onto a subspace Y is a bounded linear operator $L: X \to Y$ having the property that $Ly = y$ for all $y \in Y$. Projections play an important role in approximation theory and numerical analysis, where their linearity is a powerful advantage. The use of projections in these areas is based on the acceptance of Lx as an approximation to $x \in X$ in the subspace Y. The error incurred in using this approximation can be estimated by means of the inequalities

$$\|x - Lx\| \leqslant \|I - L\| \operatorname{dist}(x, Y) \leqslant (1 + \|L\|) \operatorname{dist}(x, Y).$$

Here, $\operatorname{dist}(x, Y)$ signifies the infimum of $\|x - y\|$ as y ranges over the subspace Y. A projection with a small value for $\|L\|$ will provide a good approximation to x, which leads naturally to the question of finding a projection L^* such that $\|L^*\| = \min_{L: X \to Y} \|L\|$. Such a projection is called a minimal projection.

If $X = C[-1, 1]$ with the usual supremum norm and $Y = P_n[-1, 1]$ then the minimal projection of X onto Y is known to exist [1], although its form is not known for $n \geqslant 2$. In this paper we shall restrict ourselves to certain

0021-9045/80/110197-06$02.00/0

Copyright © 1980 by Academic Press, Inc.
All rights of reproduction in any form reserved.

special classes of projections from $C[-1, 1]$ onto $P_n[-1, 1]$ (which we shall denote henceforward by X and Y, repsectively). Let $w(x)$ be a non-negative Lebesgue integrable function for which $\int_{-1}^{1} w(t)\,dt > 0$. Then we may define a projection $L: X \to Y$ by

$$(Lx)(t) = \int_{-1}^{1} w(s)\,x(s)\,K(t, s)\,ds,$$

where $K(t, s) = \sum_{i=0}^{n} q_i(s)\,q_i(t)$, and the q_i are polynomials of degree i, orthonormal on $[-1, 1]$ with respect to w. Cheney, McCabe and Phillips [4] obtained results which can be applied to certain convex subclasses of projections of the form (1). Unfortunately, many of the interesting subclasses are not convex. For example, the ultraspherical projections where $w(s) = (1 - s^2)^{\alpha - 1/2}$ and $\alpha > -\frac{1}{2}$ do not form a convex class. Neither do the Jacobi projections, where $w(s) = (1 - s)^{\alpha}\,(1 + s)^{\beta}$ and $\alpha, \beta > -1$.

The convexity in [4] was critical since an appeal was made to the Kolmogorov criterion (or a modified form of this). In this paper we intend to relax the convexity assumption, and instead derive results which may be applied to the commonly occurring situations. It is possible to follow the arguments in [4], replacing all the linear approximation reasoning with results from non-linear approximation. In particular the critical-point theory of Braess [3] can be applied. However, it is less technical, and just as short, to prove the results directly.

2. The General Theorems

We begin the general setting of $C(T)$, where T is a compact Hausdorff space and (T, Σ, σ) is a measure space. Y is any n-dimensional subspace of $C(T)$, and L is a bounded linear projection from $C(T)$ to Y. Writing $\|L\| = \sup_{t \in T} \sup_{v \in M} (Lv)(t)$, where $M = \{v: v \in L_\infty(T) \text{ and } \|v\|_\infty \leqslant 1\}$ we claim that $\sup_{v \in M}(Lv)(t)$ is continuous on T, since L is bounded and Lv is continuous on T. Thus $\sup_{t \in T} \sup_{v \in M} (Lv)(t) = \sup_{v \in M} (Lv)(t_c)$ for some $t_c \in T$. Now $\hat{t} \circ L \in C(T)^*$, and is hence a measure on T. Consequently

$$\sup_{t \in T} \sup_{v \in M} (Lv)(t) = (Lv_c)(t_c)$$

for $t_c \in T$ and some $v_c \in L_\infty(T)$.

We shall now confine our attention to a 1-parameter family of projections from $C(T)$ to Y defined by L_λ, where $\lambda \in \mathbb{R}$. A pair (v, t) will be called critical for L_λ if

$$(L_\lambda v)(t) = \|L\|.$$

To indicate their dependence on λ we shall usually denote the critical pairs of L_λ by (v_λ, t_λ). We denote the set of all such critical pairs by C_λ. With this preamble, we have

THEOREM 2.2. *Let U be an open interval in $\mathbb{R}$, and let $L_\lambda, \lambda \in U$ be a 1-parameter family of projections from $C(T)$ onto Y. Then U contains a local maximum of the curve $\|L_\lambda\|$ at $\lambda = \lambda_0$ iplies $(L_\lambda v_{\lambda_0})(t_{\lambda_0})$ has a local maximum at $\lambda = \lambda_0$ for each critical pair $(v_{\lambda_0}, t_{\lambda_0}) \in C_{\lambda_0}$.*

Proof. Write $(L_\lambda v)(t) = L(\lambda, v, t)$. Then L can be regarded as a function defined: $U \times M \times T \to \mathbb{R}$. Then $\|L_\lambda\| = L(\lambda, v_\lambda, t_\lambda)$ has a local maximum at $\lambda = \lambda_0$ in U implies that L has a local maximum at $(\lambda_0, v_{\lambda_0}, t_{\lambda_0})$ in $U \times M \times T$. Hence the function $L(\lambda, v_{\lambda_0}, t_{\lambda_0})$ has a local maximum in U at $\lambda = \lambda_0$.

COROLLARY 2.3. *If $(dL/d\lambda)(\lambda, v, t)$ exists for all $(v, t) \in M \times I$, then the hypotheses of Theorem 2.2 imply that $(dL/d\lambda)(\lambda, v_{\lambda_0}, t_{\lambda_0})|_{\lambda=\lambda_0} = 0$.*

THEOREM 2.4. *Let U be an open interval of $\mathbb{R}$, in which $\|L_\lambda\|$ has a local minimum at $\lambda = \lambda_0$. Again L_λ is a 1-parameter family of projections from $C(T)$ to Y. Then $(L_\lambda v_{\lambda_0})(t_{\lambda_0})$ has a local minimum at $\lambda = \lambda_0$.*

Proof. This is a rewrite of the proof of Theorem 2.2.

COROLLARY 2.4. *Let U be an open interval of R in which $(d/d\lambda) L_\lambda v$ exists for all $v \in V$ and $\lambda \in U$, and $\|L_\lambda\|$ is differentiable on U. Suppose $(d/d\lambda)(L_\lambda \hat{v}_{\lambda_0})(\hat{t}_{\lambda_0}) \neq 0$ for some $\lambda_0 \in U$. Then there exists an open interval W such that $\lambda_0 \in W$ and $\|L_\lambda\|$ is either strictly increasing on W or strictly decreasing there.*

3. APPLICATION TO THE CHEBYSHEV PROJECTION

Numerical analysts have long favoured the truncated Chebyshev Series (see [5] for details) as a technique for approximating continuous functions in the maximum norm. However, numerical evidence in [7] shows that for the ultraspherical projection, where $w_\lambda(s) = (1 - s^2)^{\lambda - 1/2}$, $\lambda > -\frac{1}{2}$, the curve $\|L_\lambda\|$ is decreasing for $\lambda = 0$ (which corresponds to the Chebyshev projection). We shall now apply Theorem 2.4 to obtain a proof of this fact.

Corresponding to our definition in Section 2, we have

$$(Lx)(t) = \int_{-1}^{1} (1 - s^2)^{\lambda - 1/2} x(s) K_\lambda(t, s)\, ds, \qquad \lambda > -\tfrac{1}{2}$$

and

$$K_\lambda(t, s) = \sum_{i=1}^{n} \alpha_i C_i^{(\lambda)}(t)\, C_i^{(\lambda)}(s),$$

where the $C_i^{(\lambda)}$ are the ultraspherical polynomials normalized by $C_i^{(\lambda)}(1) = 1$ and α_i is the normalisation factor

$$\alpha_i^{-1} = \int_{-1}^{1} (1 - s^2)^{\lambda - 1/2}\, |C_i^{(\lambda)}(s)|^2\, ds.$$

It is easy to see that $(d/d\lambda) L_\lambda v$ exists for all $v \in V$, and that $\|L_\lambda\|$ is differentiable in an open interval contining zero.

LEMMA 3.1. (i) *A critical pair for* L_λ *is* $(\hat{v}_\lambda, 1)$, *where* $\hat{v}_\lambda(t) = \operatorname{sgn} K_\lambda(1, t)$

(ii) $\hat{v}_0(t) = \sum_{k=0}^{\infty}{}' d_k C_k^{(0)}(t)$, *where the prime denotes the first term being taken with weight* $\frac{1}{2}$, *and*

$$d_k = \frac{2}{\pi k} \tan\left(\frac{\pi k}{2n+1}\right).$$

(iii) $(L_\lambda \hat{v}_0)(1) = (L_0 \hat{v}_0)(1) + \sum_{r=1}^{n} b_r(\lambda)$, *where*

$$b_r(\lambda) = \sum_{j=1}^{\infty} a_r^{(n+2j)}(\lambda)\, d_{n+2j} \qquad n + r \text{ even}$$

$$= \sum_{j=1}^{\infty} a_r^{(n+2j-1)}(\lambda)\, d_{n+2j-1} \qquad n + r \text{ odd}$$

and $C_n^{(0)} = \sum_{r=1}^{n} a_r^{(n)}(\lambda)\, C_r^{(\lambda)}$.

Proof. Proofs or references to proofs of these facts have already been given in [6].

LEMMA 3.2. *If* $C_n^{(0)} = \sum_{r=1}^{n} a_r^{(n)}(\lambda)\, C_r^{(\lambda)}$, *then the following assertions are true*:

(i) $$\frac{d}{d\lambda} a_{n-2r}^{(n)}(\lambda) = -\frac{\sqrt{\pi}(n-2r)}{\Gamma(1/2)(n-r)\, r} \qquad \textit{for} \quad \lambda = 0$$

(ii) $$\left|\frac{d}{d\lambda} a_{n-2r}^{(n)}(\lambda)\right| > \left|\frac{d}{d\lambda} a_{n-2r}^{(n+2)}(\lambda)\right| \qquad \textit{for} \quad \lambda = 0.$$

Proof. The following formula (originally due to Gegenbauer) can be found in [1], although the normalisation here is $C_n^{(\lambda)}(1) = 1$;

$$a_{n-2r}^{(n)}(\lambda) = \frac{n(n-2r+\lambda)\,\Gamma(r-\lambda)\,\Gamma(n-r)}{2r\Gamma(-\lambda)\,\Gamma(n-r+\lambda+1)} \cdot \frac{\Gamma(n+2\lambda)}{\Gamma(n+1)\,\Gamma(2\lambda)}.$$

Taking the $\lim_{\lambda \to 0} (a_{n-2r}^{(n)}(\lambda)/\lambda)$ and using $\Gamma(2\lambda) = (2\pi)^{-1/2}\, 2^{2\lambda - 1/2}\Gamma(\lambda)$ $\Gamma(\lambda + \frac{1}{2})$ we obtain (i). Then (ii) follows trivially from (i).

THEOREM 3.1. *The curve* $\|L_\lambda\|$, $\lambda > -\frac{1}{2}$ *does not have a local minimum at* $\lambda = 0$.

Proof. We have already remarked that the conditions of Theorem 2.4 are satisfied and we need only show that $(d/d\lambda)(L_\lambda \hat{v}_0)(t_0) = 0$. From Lemma 3.1(iii) it will be sufficient to show that $(d/d\lambda)\, b_r(\lambda) > 0$ and so it will suffice to prove that the sums

$$\sum_{j=1}^{\infty} \frac{d}{d\lambda} a_r^{(n+2j)}(\lambda)\, d_{n+2j} \qquad \text{and} \qquad \sum_{j=1}^{\infty} \frac{d}{d\lambda} a_r^{(n+2j-1)}(\lambda)\, d_{n+2j-1}$$

(where $n + r$ is respectively even and odd) exist and are positive for $\lambda = 0$ and $0 \leqslant r \leqslant n$. The argument which establishes this was given in [6], but we repeat it here in the case $n + r$ even, for the sake of completeness. Firstly, since the Chebyshev series converges for $t = 1$, the sum $\sum_{j=1}^{\infty} d_{n+2j}$ converges. Also it can be written as

$$\sum_{j=1}^{\infty} d_{n+2j} = \sum_{p=2}^{\infty} A_p, \qquad \text{where} \quad A_p = \sum_{1-n \leqslant j \leqslant n-1} d_{p(n+1/2)+j}.$$

Now by inspection of the formula for d_k, we can deduce that each A_p is negative, since each positive $d_{p(n+1/2)+j}$ has a corresponding $d_{p(n+1/2)-j}$ of greater modulus. Finally multiplying each term in the series by $(d/d\lambda)$ $a_r^{(n+2j)}(\lambda)$, which are all negative and decrease in modulus from Lemma 3.2, will cause the sums

$$B_p = \sum_{\substack{1-n \leqslant j \leqslant n-1 \\ j \text{ even}}} \frac{d}{d\lambda} a_r^{(p(n+1/2)+j)}(\lambda)\, d_{p(n+1/2)+j}$$

to be positive while $\sum B_p$ is still convergent. Thus we may conclude that $(d/d\lambda)\, b_t(\lambda) > 0$ for $\lambda > 0$ and so the proof is complete.

4. Comments

Computational experience from [6] suggests that the curve $\|L_\lambda\|$ as defined in section three increases monotonically for $\lambda > 0$. However, it is not possible to apply the results of Section 2 for any value of λ other than zero, since, as can be seen in the application in section three, considerable information is needed about the expansion of critical pairs $(\hat{t}_\lambda, \hat{v}_\lambda)$ in terms of the corresponding ultraspherical polynomials. Such information is not at present available, nor will it be easy to obtain.

From the analysis presented in [6], one can deduce that the curve $\|L_\lambda\|$ is *increasing* in the region of $\lambda = 0$ so that for sufficiently small negative λ we have $\|L_\lambda\| < \|L_0\|$. It is worth noting here that the performance of the Chebyshev expansion as an approximation process is not its only attractive feature. One of its great stengths is its computational simplicity, which ultraspherical expansions for $\lambda < 0$ do not possess.

Acknowledgments

The authors would like to thank Professor C. W. Clenshaw for many helpful discussions on the above material, and the referee for his helpful comments on Section 2.

References

1. R. Askey, "Orthogonal Polynomials and Special Functions," SIAM, Philadelphia, 1975.
2. J. Blatter and E. W. Cheney, On the existence of extremal projections, *J. Approx. Theory* **6** (1972), 72–79.
3. D. Braess, Non-linear approximation, *in* "Approximation Theory II" (C. G. Lorentz, C. K. Chui, and L. L. Schumaker, Eds.), Academic Press, New York, 1976.
4. E. W. Cheney, J. H. McCabe, and G. M. Phillips, A mixed norm bivariate approximation problem, preprint, Mathematical Institute of the University of St. Andrews, Scotland, 1977.
5. L. Fox and I. B. Parker, "Chebyshev Polynomials in Numerical Analysis," Oxford Univ. Press, London/New York, 1968.
6. W. A. Light, Norms of some projections on C a, b, *J. Approx. Theory* **20** (1977), 51–56.
7. W. A. Light, A comparison between Chebyshev and ultraspherical expansions, *J. Inst. Math. Appl.* **21** (1978), 000–000.