

A CONSTRUCTION OF CERTAIN NONLINEAR APPROXIMATING FAMILIES

HUGH E. WARREN

1. In discussing the nonlinear approximations of A. G. Vitushkin [2], G. G. Lorentz [1] points out a discrepancy between Vitushkin's definition of the degree of an algebraic polynomial in several variables and the usual definition of algebraic geometry. Lorentz observes that Vitushkin's reliance on a theorem of algebraic geometry makes this deviation unjustified. This paper is written to substantiate Lorentz and begin to fill the resulting gap in Vitushkin's theory.

Vitushkin considers, among others, approximating families F , in a real normed linear space L , of the form

$$(1) \quad F = \{p(t) : t \in R^n\},$$

where $p(t)$ is a polynomial in the co-ordinates t_i of t with coefficients in L . The ability of such an F to approximate other subsets of L is measured by the number of parameters n and the degree d of p as a polynomial. The problem lies in defining this d .

As defined in algebraic geometry and as used in this paper, the degree of a polynomial in several variables is the maximum sum of the exponents occurring in any one power product. Vitushkin uses, instead, the maximum exponent of any single variable. I will refer to this latter number as the "coarse degree" of a polynomial.

For a given normed linear space L and subsets K and F of L , let

$$D(K, F) = \sup_{x \in K} \inf_{y \in F} \|x - y\|.$$

Define $D_{nd}(K)$ to be the infimum of all $D(K, F)$ where F is of the form (1) and p has degree at most d . Define $V_{nd}(K)$ in the same way except let p have coarse degree at most d . Since this latter condition is less restrictive, we always have $V_{nd} \leq D_{nd}$ for any fixed K .

Vitushkin devotes the last three chapters of [2] to determining the asymptotic behavior of V_{nd} as n and d increase. In §2 we state his upper bound for V_{nd} as Theorem A. The lower bounds which Vitushkin gives ostensibly for V_{nd} are in fact lower bounds for D_{nd} , because they depend on results from algebraic geometry. As an example, the lower bound implicit in [2, Theorem 1, §36, p. 206] depends ultimately on [2, Lemma 3, §24, p. 154], which is a geometric result. For some sets K the upper bound of Theorem A is of the same

Received by the editors June 17, 1968.

order as the lower bound for $D_{nd}(K)$. If V_{nd} and D_{nd} were themselves of the same order, one would have a "best possible" result. §2, however, shows that V_{nd} can tend to zero much faster than D_{nd} .

§3 shows that in one classical case, Vitushkin's lower bounds on D_{nd} are nevertheless "best possible."

All logarithms in this paper are taken base two.

2. Let K be a compact subset of a real normed linear space L . Let $H_\epsilon(K)$ be the ϵ -entropy of K . The following result was proved by Vitushkin for L a Euclidean space [2, §29, p. 173, Theorem 1] and is easily extended.

THEOREM A. *If $n \log(d+1) \geq H_\epsilon(K)$, then $V_{nd}(K) < \epsilon$.*

PROOF. If a minimal ϵ -net for K contains h points, then by definition, $H_\epsilon(K) = \log h$. Exponentiating both sides of the inequality in the hypothesis gives $(d+1)^n \geq h$. Order the $(d+1)^n$ power products which occur in the general polynomial in n variables of coarse degree d as entries in a vector. Choose h points $\mathbf{t} = (t_1, \dots, t_n)$ such that evaluating this vector at the points \mathbf{t} gives h linearly independent vectors. This is possible since the power products are linearly independent functions.

Let $p(\mathbf{t})$ be as in (1) and have coarse degree d . Write down h equations $p(\mathbf{t}) = x$, where the x are the points of the minimal ϵ -net and the \mathbf{t} are some ordering of the points chosen above. These equations are linear in the coefficients of p , which are elements of the linear space L . Since by choice of the points \mathbf{t} the equations are independent, we can solve for the coefficients of p and obtain an approximating family of the form (1) which contains the given ϵ -net. This proves the theorem.

We now examine a particular K . Let L be l^2 . Denote by \mathbf{e}_n the sequence in l^2 with a one in the n th place and zeroes elsewhere. The set K will consist of the origin and all the points $\pm \mathbf{e}_n/n$.

THEOREM 1. *Let K be as defined above. If $n \geq 4$, then $V_{n1}(K) < 2^{2-n}$ and $D_{n1}(K) \geq n^{-3/2}$.*

PROOF. The $2m+1$ points $0, \pm \mathbf{e}_1/1, \dots, \pm \mathbf{e}_m/m$ are a $1/m$ -net for K . Put $m = 2^{n-1} - 1$, so that $2^n > 2m+1$. Theorem A, with $d=1$, shows that $V_{n1}(K) < 1/m$, and $n \geq 2$ implies $1/m \leq 2^{2-n}$.

The number $D_{n1}(K)$ is the infimum of all $D(K, F)$ where F is an n -dimensional linear manifold in l^2 . The set K is symmetric about the origin; thus for the infimum we need consider only subspaces of dimension n . Let \mathbf{R}^{n+1} be the subspace spanned by $\mathbf{e}_1, \dots, \mathbf{e}_{n+1}$. For any n -dimensional subspace F let F^* be an n -dimensional sub-

space of \mathbf{R}^{n+1} containing the orthogonal projection of F into \mathbf{R}^{n+1} . Let $\mathbf{v} = (v_i)$ be one of the two unit vectors in \mathbf{R}^{n+1} perpendicular to F^* . When $i \leq n+1$, the distance from \mathbf{e}_i/i to F^* is $|v_i|/i$. If s is the greatest of these $n+1$ distances, we have

$$s^2(1^2 + 2^2 + \cdots + (n+1)^2) \geq v_1^2 + \cdots + v_{n+1}^2 = 1,$$

because \mathbf{v} is a unit vector. Estimating the sum of the first $n+1$ squares, one sees that $s \geq n^{-3/2}$, $n \geq 4$.

Now s is $D(K \cap \mathbf{R}^{n+1}, F^*)$, which, since F^* contains the orthogonal projection of F into \mathbf{R}^{n+1} , is less than or equal to $D(K \cap \mathbf{R}^{n+1}, F)$. This number is in turn less than or equal to $D(K, F)$. Since F was an arbitrary subspace of dimension n , the theorem follows from $D(K, F) \geq n^{-3/2}$.

3. Let ω be a modulus of continuity function. Denote by Λ^ω those real valued functions on $[0, 1]$ which vanish at zero and have modulus of continuity less than or equal to ω . This section will establish

THEOREM 2. *Put*

$$\epsilon = 6\omega(1/n \log(d+1)).$$

In the space of real continuous functions on $[0, 1]$ with the uniform norm we have $D_{nd}(\Lambda^\omega) \leq \epsilon$.

PROOF. We will first construct a simple set E which closely approximates Λ^ω and then an approximating family F of the form (1) and degree d which contains E .

For $d \geq 2$, let k be the integer such that $2^k \leq d < 2^{k+1}$. If d is one, let k be one also. Take as the elements of E the 2^{nk} continuous functions on $[0, 1]$ which vanish at zero and have constant slope of $\pm nk\omega(1/nk)$ in each successive interval of length $1/nk$ beginning at zero. Given an f in Λ^ω we can determine a g in E a distance not more than $3\omega(1/nk)$ away. Indeed, if at the initial point of one of the successive intervals of length $1/nk$ we have g less than f , let g have positive slope in that interval, otherwise negative slope. We start this inductive construction at zero, where both f and g vanish.

To see that g never deviates from f by more than $3\omega(1/nk)$, let $[a, b]$ be one of the successive intervals of length $1/nk$ on which g is linear. Fix a point x in $[a, b]$. The modulus of continuity function ω dominates the modulus of continuity of f , so

$$f(a) - \omega(x-a) \leq f(x) \leq f(a) + \omega(x-a).$$

Since ω is monotone, we have $\omega(x-a) \leq \omega(1/nk)$. Supposing that $g(a)$ is less than $f(a)$, we have

$$f(x) - g(x) = f(x) - g(a) - nk\omega(1/nk)(x - a).$$

From this expression and the two preceding inequalities, it follows that $f(a) - g(a) - 2\omega(1/nk) \leq f(x) - g(x) \leq f(a) - g(a) + \omega(1/nk)$ and more specifically that

$$f(a) - g(a) - 2\omega(1/nk) \leq f(b) - g(b) \leq f(a) - g(a).$$

We see that if f and g differ by at most $2\omega(1/nk)$ at the initial point of the interval, then they differ by no more than that at the final point and by no more than $3\omega(1/nk)$ in between. A similar argument works when $g(a) \geq f(a)$. Now f and g are both zero at zero, and we can use the above reasoning, interval by interval, to show that g is no more than $3\omega(1/nk)$ from f .

At this point, note that $3\omega(1/nk) \leq \epsilon$. For by the subadditivity of ω , $3\omega(1/nk) \leq 6\omega(1/2nk)$. The desired inequality now follows from the monotonicity of ω and the relation $n \log(d+1) \leq 2nk$, which comes from the definition of k .

It remains only to construct the family F containing E . For each i from 1 to n apply the technique of Theorem A to get a family $F_i(t)$ of the form (1) with $t \in \mathbf{R}^1$ and degree of the polynomial p equal to d . This F_i is to contain the $2^k \leq d+1$ continuous functions on $[0, 1]$ which

- (1) vanish at and to the left of $(i-1)/n$;
- (2) have constant slope $\pm nk\omega(1/nk)$ in each of the k succeeding intervals of length $1/nk$;
- (3) are constant to the right of i/n .

For $t = (t_1, \dots, t_n)$, put

$$F(t) = F_1(t_1) + \dots + F_n(t_n).$$

To obtain any function g in the set E as a member of F , choose t_1 to get the right sequence of slopes in $[0, 1/n]$, t_2 for the slopes in $[1/n, 2/n]$, and so on.

This theorem amounts to one half of determining the asymptotic behavior of $D_{nd}(\Lambda^\omega)$. The other half, after one clarifies the matter of "degree", can be read from [2, Theorem 1, §38, p. 219]. Namely, there is a constant B independent of n , d and ω such that $D_{nd}(\Lambda^\omega) \geq B\epsilon$, where ϵ is the quantity in the statement of Theorem 2.

REFERENCES

1. G. G. Lorentz, *Metric entropy and approximation*, Bull. Amer. Math. Soc. **72** (1966), 926.
2. A. G. Vitushkin, *Estimation of the complexity of the tabulation problem*, Fizmatgiz, Moscow, 1959; English transl., *Theory of transmission and processing of information*, Pergamon Press, New York, 1961.

UNIVERSITY OF OREGON