JOURNAL OF APPROXIMATION THEORY 16, 81–84 (1976)

n-Widths of Function Spaces

D. J. NEWMAN

Belfer Graduate School of Science, Yeshiva University, New York 10033

Communicated by Oved Shisha

A rather wide class of problems in approximation theory is concerned with determining just how well a given space of functions can be used to approximate "arbitrary" functions. We have learned that the appropriate measure of "how well" is given in terms of the associated modulus of continuity. Indeed, we may measure the "goodness" of an approximating space by the right ϵ that allows the conclusion

$$\text{distance (space}, f) \leqslant \omega_f(\epsilon),$$

ω_f being the modulus of continuity of f.

Furthermore, in very great generality, ϵ may be equivalently defined by restricting attention to differentiable functions. Hence, if we introduce T as the integral operator and φ as f', $(T\varphi = f)$, we see that $\omega_f(\epsilon)$ can be replaced by $\epsilon \cdot \| \varphi \|$ and the final definition we distill is

$$\epsilon = \epsilon(S) = \operatorname*{Sup}_{\varphi} \inf_{p \in S} \frac{\| T\varphi - p \|}{\| \varphi \|}. \tag{1}$$

However, this expression exists outside the confines of approximation theory. Very generally, we may have two normed spaces A and B and a (bounded) linear transformation $T: A \to B$. Then, if S is any subspace of B, we may form the usual quotient space B/S and reinterpret T as a mapping from $A \to B/S$. Our expression for $\epsilon(S)$ is then simply the norm of the operator $T: A \to B/S$, i.e.,

$$\epsilon(S) = \| T \|_{B/S}. \tag{2}$$

At times, we will speak of $\epsilon(S)$ as the S-width.

To get back to approximation theory, we now explore the "competition" between the subspaces S and call S_1 a *better* approximating space than S_2 if $\epsilon(S_1) < \epsilon(S_2)$. An important problem then becomes: What is the best n-dimensional approximating space?

Therefore, we are led to the following definition, which again may be given in the general framework of a $T: A \to B$.

$$n\text{-width} = \inf_S (S\text{-width}), \qquad \text{where} \quad \dim(S) = n. \tag{3}$$

To recover the approximation theory significance, we will use the terminology:

approximation–n-width for the n-width when $A = B$ is a space of functions on $[0, 1]$ and T is the integral operator $Tf(x) = \int_0^x f(t)\, dt$. (4)

(It is known, for example, that the approximation–n-width of $C[0, 1]$ is of the exact order of $1/n$.)

Our purpose is to provide a general lower bound for the n-width and thereby to show that $1/n$ is again the exact order for the approximation–n-width in $L^p[0, 1]$.

To achieve this lower bound, we introduce the notion of the n-breadth. Again, let A, B, T be given with $T: A \to B$. This time, choose R a subspace of A and set

$$n\text{-breadth} = \operatorname*{Sup}_{\dim R = n+1} \inf_{x \in R} \frac{\| Tx \|}{\| x \|}. \tag{5}$$

THEOREM. *We always have n-width $\geqslant$ n-breadth.*

Proof. Let S be any n-dimensional subspace of B and R any $n + 1$-dimensional subspace of A. We may assume, by a slight perturbation if necessary, that the unit spheres in these finite-dimensional spaces are strictly convex.

Now let x be given on the unit sphere in R and let $F(x)$ denote the nearest vector to Tx in the subspace S (measured, of course, in the norm of B). By our assumption of strict convexity, this $F(x)$ is uniquely defined and clearly depends continuously on x. Therefore, we have a continuous map of the n-sphere into n-space and so, by a celebrated theorem in topology [3], there must be two antipodal points that map into the same point!

This guarantees us an x_0 such that $F(-x_0) = F(x_0)$. However, since it is clear from the very definition of F that $F(-x) = -F(x)$, it follows that $F(x_0) = 0$ or, in other words, that $\inf_{p \in S} \| Tx_0 - p \| = \| Tx_0 \|$. Hence, we have

$$S\text{-width} = \sup_x \inf_{p \in S} \frac{\| Tx - p \|}{\| x \|} \geqslant \inf_{p \in S} \frac{\| Tx_0 - p \|}{\| x_0 \|} = \frac{\| Tx_0 \|}{\| x_0 \|} \geqslant \inf_{x \in R} \frac{\| Tx \|}{\| x \|},$$

and this holding for all S and R gives the desired inequality.

THEOREM 2. *For the space $L^p[0, 1]$ the approximation–n-width is of the exact order of magnitude of $1/n$.*

Proof. The upper bound, c/n, for the n-width is already provided by Jackson's theorem [2], which says that this order of approximation is achieved by the choice S = all polynomials of degree less than n. (An even simpler choice of S would be the step functions with steps at the points $1/n, 2/n,\ldots, (n-1)/n$.) To obtain the lower bound we apply Theorem 1, which reduces the problem to that of producing an $n+1$-dimensional subspace $L^p[0,1]$ throughout which $\| Tx \| \geqslant (c/n) \| x \|$. Here, Bernstein's inequality [1] leads to the construction. His inequality states, namely, that for the $L^p[0,1]$ norms, we have $\| P'(t) \| \leqslant 2\pi n \| P(t) \|$ for all nth degree trigonometric polynomials, i.e., sums $\sum_{-n}^{n} c_k e^{2\pi i k t}$. We may take R, e.g., as the span of $e^{-2\pi i t}, e^{2\pi i t}, e^{4\pi i t},\ldots, e^{2\pi i n t}$ and thereby obtain the desired bound with $C = 1/2\pi$. (Again, a more elementary choice for R is given by the step functions with steps at $1/(n+1), 2/(n+1),\ldots, n/(n+1)$.)

We see, then, that in quite a number of instances, the lower bound of the breadth for the width is quite sharp, i.e., it gives the correct order of magnitude. There is some evidence that this holds in much more generality than just the integral operators in L^p. It can be shown, for example, if $A = B =$ Hilbert space and T is arbitrary, then the breadths are exactly equal to the widths. Another (trivial) observation is that the n-width $= 0$ whenever the n-breadth $= 0$. To dash any wild hopes, however, we now show that the breadth sometimes can be enormously smaller than the width.

LEMMA. *Let S be an n-dimensional subspace of E^N. We have:*

(I) *There exists a vector $v \in S$ with one of its components equal to* 1 *and with length* $\leqslant (N/n)^{1/2}$.

(II) *There exists a vector u, all of whose components are* ± 1, *such that* dist $(u, S) \geqslant (N-n)^{1/2}$.

Proof. From here on, we denote, for any vector $x \in E^N$, x_k to be its kth component and we write x^2 for $x_1^2 + x_2^2 + x_N^2$. Also, we let $v^1, v^2, v^3,\ldots, v^n$ denote an orthonormal basis for S.

(I) Form the N vectors $x^1, x^2,\ldots, x^N$ by setting $x^i = \sum_{j=1}^{n} v_i^j v^j$. If we then note that $(x^i)^2 = \sum_{j=1}^{n} (v_i^j)^2 = x_i^i$, we can conclude that one of the vectors x^i/x_i^i serves as our v. Each has some component equal to 1 (indeed, $(x^i/x_i^i)_i = 1$) and if we always had $x^{i^2}/x_i^{i^2} > N/n$, it would follow that $x_i^i/x_i^{i^2} > N/n$, or $x_i^i < n/N$. Summing over i would then give

$$n = N \cdot \frac{n}{N} > \sum_{i=1}^{N} x_i^i = \sum_{i=1}^{N} \sum_{j=1}^{n} (v_i^j)^2 = \sum_{j=1}^{n} \sum_{i=1}^{N} (v_i^j)^2 = \sum_{j=1}^{n} (v^j)^2 = n,$$

a contradiction.

(II) Let P be the operator of orthogonal projection onto S and let u range over the 2^N vectors that have every component ± 1. Since, identically, $\operatorname{dist}^2(x, S) + (Px)^2 = x^2$, it suffices for us to produce a u with $(Pu)^2 \leqslant n$. This will be accomplished by proving the interesting identity $\sum_u (Pu)^2 = n2^N$. We have, namely, $Pu = \sum_{k=1}^n (u, v^k)\, v^k$ so that $(Pu)^2 = \sum_{k=1}^n (u, v^k)^2$ and we need only to prove $\sum_u (u, v)^2 = 2^N$ for each of these v. But

$$\begin{aligned}\sum_u (u, v)^2 &= \sum_u \sum_{i,j} u_i v_i u_j v_j \\ &= \sum_u \sum_i 1 \cdot v_i^2 + \sum_u \sum_{i \neq j} u_i u_j v_i v_j \\ &= \sum_u 1 + \sum_{i \neq j} v_i v_j \sum_u u_i u_j = 2^N + 0\end{aligned}$$

and the proof is complete.

It is now trivial to give the promised example. Simply choose B to be E^N, A to be the same N-tuples normed by $\|(x_1, x_2, \dots, x_N)\| = \operatorname{Max} |x_i|$, and T to be the "identity" map. By part (II) of our lemma, for any n-dimensional S, the transformation $T\colon A \to B/S$ has norm $\geqslant (N - n)^{1/2}$ and so the n-width itself is $\geqslant (N - n)^{1/2}$. On the other hand, if R is $n + 1$-dimensional, then part (I) of the lemma, with $n + 1$ replacing n, provides a $v \in R$ such that $\| Tv \| / \| v \| \leqslant (N/(n + 1))^{1/2}$. Hence, the n-breadth is $\leqslant (N/(n + 1))^{1/2}$.

Of course, $(N/(n + 1))^{1/2} \leqslant (N - n)^{1/2}$, in line with Theorem 1, but the point is that it can be very much smaller, e.g., if $N = 2n$, n large, and our claim is verified.

References

1. S. N. Bernstein, Collected works, *Akad. Nauk SSSR*, (1952, 1954).
2. R. P. Feinerman and D. J. Newman, "Polynomial Approximation," pp. 139–141, Williams and Wilkins, Baltimore, 1974.
3. H. Seifert and W. Threlfall, "Lehrbuch der Topology," pp. 288–290, Chelsea, New York, 1947.