Negative observations concerning approximations from spaces generated by scattered shifts of functions vanishing at ∞

Amos Ron

Abstract: Approximation by scattered shifts $\{\phi(\cdot - \alpha)\}_{\alpha \in A}$ of a basis function ϕ are considered, and different methods for localizing these translates are compared. It is argued in the note that the superior localization processes are those that employ the original translates only.

AMS (MOS) Subject Classifications: 41A15, 41A25, 41A63

Key Words and phrases: approximation order, Strang-Fix conditions, scattered shifts, radial basis functions, localization of splines.

Supported in part by the U.S. Army (Contract DAAL03-G-90-0090), by the National Science Foundation (grant DMS-9102857), and by the Israel-U.S. Binational Science Foundation (grant 9000220)

Negative observations concerning approximations from spaces generated by scattered shifts of functions vanishing at ∞

Amos Ron

1. Introduction

In recent years, approximation from spaces spanned by integer translates of one or several functions became a major theme in various areas of approximation theory (and in other fields, e.g., wavelets). In most of the examples one starts with a function $\phi : \mathbb{R}^d \to \mathbb{C}$ which grows (polynomially, exponentially) at infinity and **localizes** the function by an application of a difference operator, i.e., obtains a function ψ in the form

$$\psi = \sum_{\alpha\in\mathbb{Z}^d} a(\alpha)\phi(\cdot - \alpha), \tag{1.1}$$

with the infinite sum being convergent in some topology (e.g. the $C(\mathbb{R}^d)$-topology). The localized function ψ is either compactly supported (in the case of B-splines and box splines), or, alternatively, decays at some suitable rate at infinity (in the case of radial basis functions).

Approximation from spaces spanned by non-uniform (=scattered) translates of a single function is also a popular topic in multivariate approximation, with the major emphasis being put on radial basis functions (because of computational reasons). The most common setup starts with a set of **translation centers** $A \subset \mathbb{R}^d$, and attempts to write the approximant as a linear combination of

$$_A(\phi) := \{\phi(\cdot - \alpha)\}_{\alpha\in A}.$$

Localizations are very helpful in this more general setting, as well. Even when A is finite, a localization procedure can help improve the *stability* of the evaluation process of the approximant (cf. [DL] and [DLR]).

There are several different approaches for extending the idea of "localization then approximation" from the uniform case to the scattered case, and we describe below two essentially different approaches, whose comparison is important here. The first (referred to hereafter as *Method A*) starts with a function ϕ which grows at infinity, and adheres to using only $_A(\phi)$ (with A the set of translation centers as before). The idea here is to form, for each $\alpha \in A$, some linear combination

$$\psi_\alpha := \sum_{\beta\in A} c(\alpha,\beta)\phi(\cdot - \beta),$$

such that $c(\alpha, \beta)$ decays fast to zero as $\|\alpha - \beta\| \to \infty$ (to make the sum suitably convergent) and that ψ_α is a "bump function" with the bump centers around α. The basic strategy behind Method A is that, in essence, we want to use linear combinations of $_A(\phi)$, and use the localization process only to group these translates in a way that makes the computations more feasible. The drawback of this method is quite apparent: since A is non-uniform, we need, separately for each center α, to determine the "magic combination" that will generate the bump function ψ_α. Much of the possible computational simplicity of the original ϕ is lost in this way.

Because of that difficulty, one might be tempted to use a simpler alternative method named hereafter as *Method B*, that starts with a function ψ which already has distinguished decay properties (e.g., ψ is compactly supported, decay exponentially at ∞, etc.), and use the set of shifts ${}_A(\psi) = \{\psi(\cdot - \alpha)\}_{\alpha \in A}$ to construct a suitable approximation scheme. Candidates for ψ exist in abundance: for example, one can choose his favorite radial function ϕ, apply some *fixed* localization process (e.g., using integer centers) to obtain ψ, and then proceed to consider ${}_A(\psi)$. It is important to note that this variant of Method B incorporates, once more, suitable translates of ϕ into the approximation process, but in a different manner: since the set B of centers that are used in the localization process is determined independently of the set A of translation centers, the shifts of ϕ which are used in approximation schemes based on Method B are taken from the larger set $A + B$.

At present, a theory that unfolds the approximation properties of spaces generated by scattered shifts of a basis function is yet to be established; only partial preliminary results are known [P], [BP], [BDL], and [DR], all of them are based on Method A. The use of Method A is not accidental, since there is already strong enough evidence to support the claim that *Method B cannot, in general, yield good approximation schemes*.

In this short note I will sketch two examples that illustrate the claim just made. The note was originated during discussions I had with Nira Dyn in which we found ourselves sharing the same opinion on Method B.

In preparing for this note, I tried to find some heuristic argument that explains the phenomenon revealed by the examples below. Here is the most "convincing" one I could get: In general, for a given center set A, only very few basis functions ϕ are "adequate" for A (in the sense of the existence of "good" approximation schemes based on ${}_A(\phi)$). A typical radial basis function ϕ, though, is *universally adequate*, i.e., is adequate for any center set A. Now, given a universally adequate basis function ϕ, a set of centers A, and a corresponding set of localized functions $\{\psi_\alpha\}_{\alpha \in A}$ (each of which is a linear combination of translates of ϕ), the test the functions $\{\psi_\alpha\}_{\alpha \in A}$ should stand is whether each (or at least most) of the original translates in ${}_A(\phi)$ can be reproduced by linear combinations of $\{\psi_\alpha\}_{\alpha \in A}$. In the case of Method A, since the centers in A are the only ones employed in the localization, the test amounts to the (linear) invertibility of some biinfinite matrix, a property that holds if the localization is done correctly. On the other hand, in the case of Method B, since many more translates of ϕ are used during the localization (viz., all centers in $A + B$), it is generally impossible to recover any translate of ϕ with the aid of ${}_A(\psi) = \{\psi(\cdot - \alpha)\}_{\alpha \in A}$. Therefore, the (possibly) good approximation qualities of $\mathrm{span}_A(\phi)$ are not necessarily inherited by $\mathrm{span}_A(\psi)$.

2. A discussion concerning uniform translation centers

Suppose that we start with some basis function ϕ and wish to obtain from ϕ a localized function ψ, in order to use

$$\psi(\cdot - \alpha), \quad \alpha \in \mathbb{Z}^d$$

for subsequent approximation schemes. The question at stake is what translates of ϕ to use in the localization process. Method A insists on using *integer* translates, while Method B, although might favor this choice over others, leaves room for other choices as well. Let's see the devastating effect of the choice of non-integer centers for localization.

We assume, for simplicity and without much loss of generality, that ψ decays at ∞ (at least) as fast as $\|\cdot\|_2^{-k}$ (with $\|\cdot\|_2$ being the Euclidean norm) for some $k > d$, and that ψ has a non-zero mean value (i.e., $\widehat{\psi}(0) \neq 0$). We further assume that the initial space of approximants we start with is refined by dilation (this is *stationary* refinement, cf. e.g., [R]). Under such (and even milder) assumptions, the approximation properties of our space (measured asymptotically by the criterion of *approximation orders*) are known to be governed by the so-called Strang-Fix conditions ([SF], [BJ], [BR], [BDR]). Roughly speaking, these conditions say that good approximation orders require $\widehat{\psi}$ to have a high order zero at each of the points of the set

$$2\pi \mathbb{Z}^d \backslash 0.$$

However, the Fourier transform of a typical basis function ϕ usually vanishes nowhere, and in particular does not vanish at any of the points in $2\pi\mathbb{Z}^d\backslash 0$. *The required zeros of $\widehat{\psi}$ are thus generated during the localization process.*

Since our original ϕ is (tacitly) assumed to grow at ∞, we expect its Fourier transform to be singular somewhere. Further, stationary refinements usually make sense only when the origin is the most singular point of $\widehat{\phi}$. In turn, if this is the case, then the difference scheme

$$T : f \mapsto \sum_{\beta \in B} c(\beta) f(\cdot - \beta),$$

which is used in the localization (i.e., the one that defines ψ as $T\phi$) should annihilate polynomials of sufficiently high degree, or equivalently, the (assumed to be) function

$$\widehat{T}(\omega) := \sum_{\beta \in B} c(\beta) e^{-i\beta\cdot\omega}$$

has a high order zero at the origin (more details can be found, e.g., in [DJLR] and [P]). Since

$$\widehat{\psi} = \widehat{T}\widehat{\phi},$$

high order zeros for $\widehat{\psi}$ on $2\pi\mathbb{Z}^d\backslash 0$ can be obtained if the following three conditions hold (a) $\widehat{\phi}$ has a high order singularity at 0 (forcing $\widehat{T}$ to have a high order zero there) (b) $\widehat{\phi}$ is not too singular on $2\pi\mathbb{Z}^d\backslash 0$ (for radial functions the origin is *assumed* to be the only point of singularity, cf. [R]) and (c) the high order zero $\widehat{T}$ has at the origin is projected on each of the points of $2\pi\mathbb{Z}^d\backslash 0$. While the first two conditions have to do with the choice of the original basis function ϕ, the last condition is directly connected to the localization process: *the zero $\widehat{T}$ has at the origin is reflected in the most effective way at each of the points in $2\pi\mathbb{Z}^d\backslash 0$ if $\widehat{T}$ is 2π-periodic, which is equivalent to having $B \subset \mathbb{Z}^d$.*

A beautiful concrete example for the above abstract discussion exists in box spline theory. We outline here the box spline situation and refer the reader to [RS] and the references therein for more details. We take the original "universally acceptable" function ϕ to be the truncated power function, [D], whose Fourier transform has the form

$$\widehat{\phi}(\omega) = \prod_{\xi \in \Xi} (-i\xi \cdot \omega)^{-1},$$

where Ξ is a finite collection of non-zero vectors in $\mathbb{R}^d$ which we assume here to be integers. Choosing a finite difference operator T by the rule

$$\widehat{T}(\omega) = \prod_{\xi \in \Xi} (e^{-ix_\xi \cdot \omega} - 1),$$

with x_ξ any non-zero constant multiple of ξ, $\xi \in \Xi$, results in a compactly supported ψ, known as *a box spline*. Choosing $\mathrm{span}_{\mathbb{Z}^d}(\psi)$ to be our space of approximants, it is known that the approximation orders that correspond to this space intimately depend not only on the original Ξ but also on the type of localization procedure used. In particular, for a fixed given Ξ, the best orders are obtained when each x_ξ is an integer vector (which implies that we use only integer translates for localizing), and the approximation order is zero (worst possible) if, for instance, each x_ξ is an irrational multiple of the integer ξ.

3. An example concerning scattered translation centers

While my personal understanding of the issue is primarily based on the observations detailed in the previous section, interested readers might argue that the discussion there considers only uniform translation centers, while our major objective here is the scattered case. In this section thus, we attempt to fill this gap. Because analogs of general principles such as the Strang-Fix conditions are yet to be discovered here, I am not able to maintain the level of generality of the discussion of §2, and concentrate on a modest, still quite illustrative, example.

Our function ψ is taken to be the univariate cardinal spline of order n. This function is supported on $[0, n]$ and is obtained as a localization of the truncated power x_+^{n-1}, using $\{0, 1, ..., n\}$ as the localization centers. We take A to be any set of translation centers with no accumulation points, and define S to be the space of all *infinite* linear combinations of ${}_A(\psi) = \{\psi(\cdot - \alpha)\}_{\alpha \in A}$. We refine this space by dilation, and approximate in the uniform norm. Under those conditions we have the following:

Proposition 3.1. *If S provides positive approximation orders to infinitely differentiable compactly supported functions, then for any bounded interval I there exist coefficients $\{c(\alpha, I)\}_{\alpha \in A}$ such that*

$$\sum_{\alpha \in A} c(\alpha, I)\psi(x - \alpha) = 1, \quad x \in I. \tag{3.2}$$

Proof: The statement "S provides positive approximation order to infinitely differentiable compactly supported functions" means, by definition, that for some $k > 0$ and for every smooth compactly supported f,

$$\mathrm{dist}(f(h\cdot), S) := \inf_{s \in S} \|f(h\cdot) - s\|_\infty = O(h^k). \tag{3.3}$$

We take f to be as above, and assume also that it is 1 on some neighborhood of the origin. For sufficiently small h, $f(h\cdot)$ is 1 on the interval I, and therefore (3.3) implies that

$$\inf_{s \in S} \|1 - s\|_{L_\infty(I)} = O(h^k).$$

Denoting by $S_{|_I}$ the restrictions of the functions in S to I, we conclude that 1 is in the closure of $S_{|_I}$, but since this latter space is finite dimensional, it is closed, and hence $1 \in S_{|_I}$, which is equivalent to the statement of the proposition. $\square$

The previous result shows that S cannot approximate well unless the constant function belongs, at least locally, to the space. At the same time, the next proposition shows that this latter property can hold only under very restrictive conditions on the distribution of the centers in A: either A contains a complete uniform set of centers, or that A contains at least two "almost complete" uniform sets of centers. In particular, the use of truly irregular sets A results in 0 approximation order, and thus using scattered translates of the B-spline ψ should be avoided. In contrast, we remind the reader of the classical fact (cf. e.g., [S]) that, for $\phi(x) := x_+^{n-1}$, the functions in ${}_A(\phi)$ can be easily localized by Method A to obtain the standard B-splines of order n that correspond to the knot sequence A, and these B-splines yield approximation order n.

In the proposition, Π_k stands for the space of univariate polynomials of degree $\le k$. Also, for any two sets C, D, the notation $C\backslash D$ should be understood as $C\backslash(C\cap D)$.

Proposition 3.4. *With ψ, A and S defined as in the second paragraph of this section, if, for every bounded I, there exist coefficients $\{c(\alpha, I)\}_{\alpha\in A}$ that satisfy (3.2), then one of the two following conditions holds:*

(a) There exists $\alpha \in \mathbb{R}$ such that $\alpha + \mathbb{Z} \subset A$.

(b) There exist $\alpha_k \in \mathbb{R}$, $k = 1, 2$, such that $\alpha_1 \ne \alpha_2 \pmod{\mathbb{Z}}$, and such that, for each k, $(\alpha_k+\mathbb{Z})\backslash A$ contains no more than $n-1$ points.

Proof: Let I be a bounded open interval. large enough to contain $\alpha \in A$ for which $c(\alpha) \ne 0$. By assumption, there exist I-dependent coefficients $(c(\alpha))_\alpha$ such that:

$$1 = \sum_{\alpha\in A} c(\alpha)\psi(x-\alpha), \quad x \in I. \tag{3.5}$$

In particular,

$$\sum_{\alpha\in A} c(\alpha)\psi(\cdot-\alpha) \text{ coincides on } I \text{ with a polynomial in } \Pi_{n-1}\backslash 0. \tag{3.6}$$

Picking any $\alpha_1 \in A\cap I$ for which $c(\alpha_1) \ne 0$, and denoting $J := \alpha_1 + \mathbb{Z}$, we define a sequence $d : J \to \mathbb{C}$ as follows:

$$d(j) := \begin{cases} c(j), & j \in A, \\ 0, & \text{otherwise.} \end{cases}$$

Since A has no accumulation points, we can find, for each $j \in J$, a neighborhood $V_j \subset (j-1/2, j+1/2)$ of j that satisfies

$$V_j \cap (\cup_{k=0}^n (A+k)) \subset \{j\}.$$

The choice of V_j is made to guarantee that, for $\beta \in A\backslash J$, none of the knots $\beta, \beta+1, ..., \beta+n$ of the spline $\psi(\cdot-\beta)$ lies in V_j, and hence, each $\psi(\cdot-\beta)$, $\beta \in A\backslash J$, coincides with some polynomial (necessarily of degree $< n$) on V_j. Also, for a purely technical reason, we choose each V_j small enough to ensure the relation

$$I \cap V_j \ne 0 \iff V_j \subset I. \tag{3.7}$$

From the above discussion, we conclude that the sum

$$\sum_{\beta \in A\setminus J} c(\beta)\psi(\cdot - \beta)$$

is a piecewise-polynomial with no breakpoints in any V_j. Invoking (3.6), we conclude that the same is true for

$$\sum_{\beta \in A\cap J} c(\beta)\psi(\cdot - \beta) = \sum_{j\in J} d(j)\psi(\cdot - j), \tag{3.8}$$

provided that $V_j \subset I$, or, what is equivalent, (3.8), that $V_j \cap I \neq \emptyset$. On the other hand, $J \subset \cup_{j\in J} V_j$ (since $j \in V_j$, $\forall j$), and the points in J are the only candidates for knots of the sum in (3.8), and, consequently, this sum has no breakpoints in I, i.e., it coincides on I with a polynomial p; also, necessarily, $p \in \Pi_{n-1}$. As is well known (see, e.g., the discussion about Marsden's identity in [B]), this implies that the sequence $d_{|_{I\cap J}}$ coincides with the restriction to $I \cap J$ of some $q \in \Pi_{n-1}$, with $\deg q = \deg p$. Also, $q \neq 0$, since $q(\alpha_1) = d(\alpha_1) = c(\alpha_1) \neq 0$, hence $p, q \in \Pi_{n-1}\setminus 0$. In particular, q can vanish at most $n-1$ times on J, and remembering that $j \in A$ whenever $d(j) = q(j) \neq 0$, we finally conclude that $\#((J\setminus A) \cap I) < n$. Since I can be chosen arbitrarily, it follows that $\#(J\setminus A) < n$.

If, for all sufficiently large I, the I-dependent polynomial q is of degree 0, then $J \cap I \subset A$, all I, and hence $J \subset A$. Thus, to complete the proof, we assume that $\deg q > 0$ for some large I, and prove that in such a case we can find $\alpha_2 \in A$, such that $\alpha_2 \notin J = \alpha_1 + \mathbb{Z}$, and such that A contains $\alpha_2 + \mathbb{Z}$ with the possible exception of $n-1$ centers. To do this, we consider on I the sum

$$\sum_{\beta \in (A\setminus J)} c(\beta)\psi(\cdot - \beta) = 1 - p. \tag{3.9}$$

Since $\deg p = \deg q > 0$, $1 - p \neq 0$, and consequently, we have verified that (3.6) holds even when A is replaced there by $A\setminus J$. Thus, with α_2 any number in $A\setminus J$, the first part of the proof can be invoked to conclude that $\#((\alpha_2 + \mathbb{Z})\setminus A) < n$. $\square$

References

[B] C. de Boor, Splines as linear combinations of B–splines, a survey, in *Approximation Theory II*, G. G. Lorentz, C. K. Chui, and L. L. Schumaker eds., Academic Press (New York) 1976, 1–47.

[BP] R.K. Beatson and M.J.D. Powell, Univariate multiquadric approximation: quasi-interpolation to scattered data, Constructive Approx. **8** (1992), 275–288.

[BDL] M.D. Buhmann, N. Dyn, and D. Levin, On quasi-interpolation by radial basis functions with scattered centers, DAMTP 1992/NA6, Silver St., Cambridge, England CB3 9EW, 1992.

[BDR] C. de Boor, R.A. DeVore and A. Ron, Approximation from shift-invariant subspaces of $L_2(\mathbb{R}^d)$, Trans. Amer. Math. Soc., to appear.

[BJ] C. de Boor and R.Q. Jia, Controlled approximation and a characterization of the local approximation order. Proc. Amer. Math. Soc., **95** (1985) 547-553.

[BR] C. de Boor and A. Ron, Fourier analysis of approximation power of principal shift-invariant spaces, Constructive Approximation, **8** (1992), 427–462.

[D] W. Dahmen, On multivariate B-splines, SIAM J. Numer. Anal. **17** (1980), 179-191.

[DJLR] N. Dyn, I.R.H. Jackson, D. Levin, A. Ron, On multivariate approximation by the integer translates of a basis function, Israel Journal of Mathematics **78** (1992), 95–130.

[DL] N. Dyn and D. Levin, Bell shaped basis functions for surface fitting, in Approximation Theory and Applications (Z. Ziegler Ed.), Academic Press, New York, (1981), 113–129.

[DLR] N. Dyn, D. Levin and S. Rippa, Numerical procedures for global surface fitting of scattered data by radial functions, SIAM J. Sci. and Stat. Computing **7** (1986), 639–659.

[DR] N. Dyn and A. Ron, Approximation by scattered shifts of a radial basis function, ms.

[P] M. J. D. Powell, The theory of radial basis function approximation in 1990, in: *Advances in Numerical Analysis Vol. II: Wavelets, Subdivision Algorithms and Radial Basis Functions*, W.A. Light, ed., Oxford University Press, (1992), 105–210.

[R] A. Ron, Approximation orders from principal shift-invariant spaces generated by a radial basis function, in *Numerical Methods of Approximation Theory* Vol. 9, D. Braess & L.L. Schumaker eds., International Series of Numerical Mathematics Vol. 105, Birkhäuser Verlag, Basel, 1992, 245–268.

[RS] A. Ron and N. Sivakumar, The approximation order of box spline spaces, Proc. Amer. Math. Soc., to appear.

[S] L.L. Schumaker, *Spline Functions:* Basic Theory, New York, John Wiley & Sons, 1981.

[SF] G. Strang and G. Fix, A Fourier analysis of the finite element variational method, C.I.M.E. II Ciclo 1971, in *Constructive Aspects of Functional Analysis* G. Geymonat, ed., 1973, 793-840.