Vietnam Journal of Mathematics **27**:4 (1999) 363–367

Vietnam Journal
of
MATHEMATICS

Short Communication

On Asymptotic Orders of n-Term Approximations and Non-Linear n-Widths

Dinh Dung

Institute of Information Technology, Nghia Do, Cau Giay, Hanoi, Vietnam

Received May 7, 1999

1. We investigate in the present paper the best n-term approximation by the family **V** formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel, and optimal continuous algorithms in non-linear approximation, especially n-term approximation in terms of non-linear n-widths, for functions from the Besov space of common mixed smoothness. A central question to be considered is: What, if any, are the advantages of non-linear approximation over approximation by linear manifolds, in the first place, for smoothness classes of functions? It was proved in previous papers that in contradistinction to such n-widths as Kolmogorov n-width and linear n-width, etc., which are connected with approximation by linear manifolds, all well-known, non-linear n-widths in the space $L_q(\mathbf{T}^d)$ of the unit ball of the classical smoothness Sobolev space W_p^α and Besov space $B_{p,\theta}^\alpha$ have, roughly speaking, the same (and better in certain cases) asymptotic order $n^{-\alpha/d}$ independently of the relations between p, q and θ, where $\mathbf{T}^d$ is the d-dimensional torus $[-\pi, \pi]^d$.

In non-linear approximation, the smoothness of functions to be approximated is more conveniently, and maybe more naturally, given by boundedness of Besov quasi-norm. In this paper we give the asymptotic orders of these non-linear n-widths and best n-term approximation by the family **V**, in the space $L_q(\mathbf{T}^d)$ of the unit ball $\mathbf{SB}_{p,\theta}^r$ of the Besov space of functions on $\mathbf{T}^d$ with common mixed smoothness r. Moreover, these asymptotic orders coincide and are achieved by a continuous algorithm of n-term approximation by **V**, which is explicitly constructed. The important point to note here is that if, for any $1 < p, q < \infty$, $0 < \theta, \tau \le \infty$ and $r > 1/p$, $\mathbf{SB}_{p,\theta}^r$ is a set of multivariate functions ($d > 1$) with given mixed Besov smoothness r, and the Besov scale $\theta < (r + (1/2))^{-1}$ (in particular, θ is small enough), then the asymptotic orders equal

$$n^{-r}(\log n)^{-\delta} = o(n^{-r}),$$

where $\delta = (d-1)(1/\theta - 1/2 - r) > 0$. This property is characteristic of multivariate non-linear approximation of functions with mixed Besov smoothness.

2. Let X be a quasi-normed linear space and $\Phi = \{\varphi_k\}_{k=1}^{\infty}$ a family of elements in X (a quasi-norm $\|\cdot\|$ is defined as a norm except that the triangle inequality is substituted by $\|f+g\| \leq C(\|f\|+\|g\|)$ with C an absolute constant). Denote by $\mathbf{M}_n(\Phi)$ the non-linear manifold of all linear combinations of the form

$$\varphi = \sum_{k\in Q} a_k\varphi_k,$$

where Q is a set of natural numbers with $|Q| = n$. Here and later, $|Q|$ denotes the cardinality of Q. We shall assume that some elements of Φ can coincide, in particular, Φ can be a finite set, i.e., the number of distinct elements of Φ is finite. Let W be a subset in X. The best n-term approximation $\sigma_n(W, \Phi, X)$ by the family Φ is given by

$$\sigma_n(W, \Phi, X) := \sup_{f\in W} \inf_{\varphi\in \mathbf{M}_n(\Phi)} \|f - \varphi\|.$$

3. A (continuous) algorithm in n-term approximation by Φ of the elements from W, is represented as a (continuous) mapping S from W into $\mathbf{M}_n(\Phi)$. Denote by $\mathcal{F}(X)$ the set of all bounded Φ whose intersection $\Phi \cap L$, with any finite-dimensional subspace L in X, is a finite set. The non-linear n-width $\tau_n(W, X)$ [3] is defined by

$$\tau_n(W, X) := \inf_{\Phi, S} \sup_{f\in W} \|f - S(f)\|,$$

where the infimum is taken over all continuous mappings S from W into $\mathbf{M}_n(\Phi)$ and all families $\Phi \in \mathcal{F}(X)$. Similar to $\tau_n(W, X)$ is the non-linear n-width $\tau_n'(W, X)$ which is defined in the same way as $\tau_n(W, X)$ but the infimum is taken over all continuous mappings S from W into a finite subset of $\mathbf{M}_n(\Phi)$ or, equivalently, over all continuous mappings S from W into $\mathbf{M}_n(\Phi)$ and all finite families Φ in X.

Let l_∞ be the normed linear space of all bounded sequences of numbers $x = \{x_k\}_{k=1}^{\infty}$, equipped with the norm

$$\|x\|_\infty := \sup_{1\leq k<\infty} |x_k|,$$

and $\mathbf{M}_n$ the subset in l_∞ of all $x \in l_\infty$ for which $x_k = 0$, $k \notin Q$, for some set of natural numbers Q with $|Q| = n$. Consider the mapping R_Φ from the metric space $\mathbf{M}_n$ into X defined by

$$R_\Phi(x) := \sum_{k\in Q} x_k\varphi_k,$$

if $x = \{x_k\}_{k=1}^{\infty}$ and, $x_k = 0$, $k \notin Q$, for some Q with $|Q| = n$. From the definitions, we can easily see that if the family Φ is bounded, then R_Φ is continuous mapping from $\mathbf{M}_n$ into X and moreover, $\mathbf{M}_n(\Phi) = R_\Phi(\mathbf{M}_n)$. On the other hand, any algorithm S of n-term approximation of the elements in W by Φ can be treated as a composition

$$S = R_\Phi \circ G$$

for some mapping G from W into $\mathbf{M}_n$. Therefore, if G is required to be continuous, then the algorithm S will also be continuous. These preliminary remarks are a basis for the notion of the non-linear n-width $\alpha_n(W, X)$ which is given by

$$\alpha_n(W, X) := \inf_{\Phi, G} \sup_{f\in W} \|f - R_\Phi(G(f))\|,$$

where the infimum is taken over all continuous mappings G from W into $\mathbf{M}_n$ and all bounded families Φ in X (see [3]).

The continuity assumption in the definitions of the non-linear widths τ_n and α_n, which has the origin from the classical Alexandroff n-width, is quite natural: the closer objects are, the closer their reconstructions should be. On the one hand, the continuity assumption decreases the possibilities of approximation. On the other, it guarantees the lower bound of n-term approximations. Moreover, it does not weaken the rate of the corresponding n-term approximations for many well-known dictionaries and functions classes. Namely, it is known that the best n-term approximation and n-term approximation by continuous algorithm have the same asymptotic order. This is shown again in our paper for the family $\mathbf{V}$, and the unit ball of Sobolev and Besov spaces of functions with common mixed smoothness. As the continuity assumption on the algorithms of approximation by "complexes" leads to the Alexandroff n-width, the continuity assumption on the algorithms of n-approximation leads to various continuous non-linear n-widths.

There are other notions of non-linear n-widths which are based on continuous algorithms of non-linear approximations different from n-term approximation, and related to problems discussed in the present paper. They are the well-known and very old non-linear Alexandroff n-width $a_n(W, X)$ (see the definition, e.g., in [2–6]), the non-linear manifold n-width $\delta_n(W, X)$[1], and the non-linear n-width $\beta_n(W, X)$[2]. These non-linear n-widths are very close to $\tau_n(W, X)$ and $\alpha_n(W, X)$. From inequalities between them [2–5], it follows that they are asymptotically equivalent. More precisely, all these n-widths have the same asymptotic order for well-known smoothness classes of functions. The reader can also consult [1–3, 5] for various aspects of non-linear n-widths and n-term approximation for mixed smoothness classes of multivariate functions.

4. We now define Besov space of functions with mixed smoothness and the family $\mathbf{V}$ formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel.

For a non-negative integer r, the univariate symmetric difference operator Δ_h^s, $h \in \mathbf{T}$, is defined inductively by $\Delta_h^s := \Delta_h^1 \Delta_h^{s-1}$, starting from the operator $\Delta_h^1 f := f(\cdot + (h/2)) - f(\cdot - (h/2))$. For s a natural number and $e \subset E := \{1, 2, ...d\}$, we let the multivariate mixed sth difference operator $\Delta_h^s(e)$, $h \in \mathbf{T}^d$, be defined by

$$\Delta_h^s(e) f := \prod_{j \in e}^{d} \Delta_{h_j}^s f,$$

where the univariate operator $\Delta_{h_j}^s$, is applied to the variable x_j (in particular, $\Delta_h^s(\emptyset) f \equiv f$). For $r > 0$ and $0 < p, \theta \leq \infty$, let $\mathbf{B}_{p,\theta}^r$ denote the Besov space of all functions on $\mathbf{T}^d$, for which the quasi-norm

$$\|f\|_{\mathbf{B}_{p,\theta}^r} := \sum_{e \subset E} |f|_{B_{p,\theta}^{r,e}}$$

is finite, where $\|\cdot\|_p$ is the usual p-integral norm in $L_p := L_p(\mathbf{T}^d)$ and

$$|f|_{B_{p,\theta}^{r,e}} := \left(\int_{\mathbf{T}^d} \prod_{j \in e} |h_j|^{-1-\theta r} \|\Delta_h^s(e) f\|_p^\theta dh \right)^{1/\theta}, \theta < \infty,$$

(the integral is changed to the supremum for $\theta = \infty$) for some $s > r$.

For $\nu \in \mathbf{N}^d$, we let the multivariate tensor product de la Vallée Poussin kernel S_ν of order ν be defined by

$$S_\nu(x) := \prod_{j=1}^{d} \frac{V_{\nu_j}(x_j)}{3\nu_j},$$

where

$$V_{\nu_j}(x_j) := 1 + 2\sum_{k=1}^{\nu_j} \cos kx_j + 2\sum_{k=\nu_j+1}^{2\nu_j} \frac{2\nu_j - k}{\nu_j} \cos kx_j.$$

We define the family $\mathbf{V}$ formed from the integer translates of the mixed dyadic scales of S_ν, by

$$\mathbf{V} := \{\varphi_s^k\}_{s\in Q_k, k\in \mathbf{N}^d}, \quad \varphi_s^k := S_{2^k}(\cdot - sh^k),$$

where $2^x = (2^{x_1}, ..., 2^{x_d})$ and $xy = (x_1y_1, ..., x_dy_d)$ for $x, y \in \mathbf{R}^d$;

$$Q_k := \{s \in \mathbf{Z}_+^d : s_j < 3 \times 2^{k_j+1} \; j = 1, ..., d\}; \quad h^k := 3^{-1}\pi(2^{-k_1}, .., 2^{-k_d}).$$

5. We use the notation $F \asymp F'$ if $F \ll F'$ and $F' \ll F$, and $F \ll F'$ if $F \le CF'$ with C an absolute constant. Let

$$\mathbf{SB}^r_{p,\theta} := \{f \in \mathbf{B}^r_{p,\theta} : \|f\|_{\mathbf{B}^r_{p,\theta}} \le 1\}$$

be the unit ball in $\mathbf{B}^r_{p,\theta}$. Denote by γ_n any one of α_n, τ_n, τ'_n, β_n, a_n, and δ_n. The main results of the present paper read as follows:

Theorem. *Let $1 < p, q < \infty$, $0 < \theta \le \infty$ and $r > 1/p$. Then we have*

$$\gamma_n(\mathbf{SB}^r_{p,\theta}, L_q) \asymp \sigma_n(\mathbf{SB}^r_{p,\theta}, \mathbf{V}, L_q) \asymp n^{-r}(\log n)^{(d-1)(r+1/2-1/\theta)}.$$

In addition, we can explicitly construct a finite subset $\mathbf{V}^*$ of V and a positive homogeneous continuous mapping $G^* : \mathbf{B}^r_{p,\theta} \to \mathbf{M}_n$ such that the asymptotic order of $\gamma_n(\mathbf{SB}^r_{p,\theta}, L_q)$ is achieved by the continuous algorithm $S^* = R_{\mathbf{V}^*} \circ G^*$ of n-term approximation by $\mathbf{V}^*$, i.e.,

$$\sup_{f\in \mathbf{SB}^r_{p,\theta}} \|f - S^*(f)\|_q \ll n^{-r}(\log n)^{(d-1)(r+1/2-1/\theta)}.$$

This theorem was announced in [3] and proved in [4] for the case $\theta \ge 2$, replacing the condition $r > (1/p)$ by the weaker condition $r > \max\{0, 1/p - 1/q, 1/p - 1/2\}$. However, the most interesting case is when $d > 1$ and $\theta < (r + (1/2))^{-1}$. Particularly, if $\mathbf{SB}^r_{p,\theta}$ is a set of multivariate functions ($d > 1$) with given mixed smoothness r, this takes place for a θ small enough. As mentioned above, the asymptotic order of γ_n and σ_n in this case is

$$\gamma_n(\mathbf{SB}^r_{p,\theta}, L_q) \asymp \sigma_n(\mathbf{SB}^r_{p,\theta}, \mathbf{V}, L_q) \asymp n^{-r}(\log n)^{-\delta} = o(n^{-r}),$$

with $\delta = (d-1)(1/\theta - 1/2 - r) > 0$.

6. For $0 < p \le \infty$, denote by l_p^m the space of all sequences $x = \{x_k\}_{k=1}^m$ of (complex) numbers, equipped with the quasi-norm $\|x\|_{l_p^m} := \left(\Sigma_{k=1}^m |x_k|^p\right)^{1/p}$ with the change to the max norm when $p = \infty$.

Let $0 < p, \theta \le \infty$, $\mathbf{N} = \{N_k\}_{k \in Q}$ be a sequence of natural numbers with Q a finite set of indices. Denote by $\mathbf{b}^{\mathbf{N}}_{p,\theta}$ the space of all such sequences $x = \{x^k\}_{k \in Q} = \{\{x^k_j\}_{j=1}^{N_k}\}_{k \in Q}$, for which the mixed quasi-norm

$$\|x\|_{\mathbf{b}^{\mathbf{N}}_{p,\theta}} := \Big(\sum_{k \in Q} \|x^k\|^\theta_{X^k}\Big)^{1/\theta}, \ \theta < \infty,$$

is finite (the sum is changed to supremum for $\theta = \infty$), where $X^k := l_p^{N_k}$. Let $\mathbf{S}^{\mathbf{N}}_{p,\theta}$ be the unit ball in $\mathbf{b}^{\mathbf{N}}_{p,\theta}$.

We employed in the proof of our results, in particular, the following:

Lemma. *Let* $0 < p, \theta, \tau \le \infty$, *and* $p \le \theta$. *Then, for any* $n < m$, *we can explicitly construct a positive homogeneous continuous mapping* $G : \mathbf{b}^{\mathbf{N}}_{p,\theta} \to \mathbf{M}_n$ *such that*

$$\sigma_n(\mathbf{S}^{\mathbf{N}}_{p,\theta}, \mathcal{E}, \mathbf{b}^{\mathbf{N}}_{\infty,\tau}) \le \sup_{x \in \mathbf{S}^{\mathbf{N}}_{p,\theta}} \|x - S(x)\|_{\mathbf{b}^{\mathbf{N}}_{\infty,\tau}} \le n^{-1/p} |Q|^{1/p + 1/\tau - 1/\theta},$$

where $m = \sum_{k \in Q} N_k$ *and* $S := R_{\mathcal{E}} \circ G$ *and* $\mathcal{E}$ *is the canonical basis in* $\mathbf{b}^{\mathbf{N}}_{\infty,\tau}$.

References

1. R. DeVore, R. Howard, and C. Micchelli, Optimal non-linear approximation, *Manuscripta Math.* **63** (1989) 469–478.
2. R. DeVore, G. Kyriazis, D. Leviatan, and V. Tikhomirov, Wavelet compression and non-linear n-widths, *Adv. Comp. Math.* **1** (1993) 194–214.
3. Dinh Dung, On best continuous methods in n-term approximation, *Vietnam J. Math.* **26** (1998) 367–371.
4. Dinh Dung, Continuous algorithms in n-term approximation and non-linear widths, *J. of Approx. Theory* **102** (2000) 217–242.
5. Dinh Dung and Vu Quoc Thanh, On non-linear n-widths, *Proc. Amer. Math. Soc.* **124** (1996) 2757–2765.
6. V. Tikhomirov, *Some Topics in Approximation Theory*, Moscow State University, Moscow, 1976 (Russian).