Vietnam Journal of Mathematics **26**:4 (1998) 367–371

Vietnam Journal of MATHEMATICS

Short Communication

On Best Continuous Methods in n-term Approximation

Dinh Dung

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

Received June 18, 1998

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 $\|f+g\| \le C(\|f\|+\|g\|)$ with C an absolute constant). Consider n-term approximation of elements $f \in X$ by the linear combinations of the form $\varphi = \sum_{k\in Q} a_k\varphi_k$, where Q is a set of natural numbers with cardinality n. It is convenient to assume that some elements of Φ can coincide, in particular, Φ can be a finite set, i.e., the number of the distinct elements of Φ is finite. Denote by $M_n(\Phi)$ the set of all these linear combinations. If $W \subset X$, we can put

$$\sigma_n(W, \Phi, X) := \sup_{f\in W} \sigma_n(f, \Phi, X) = \sup_{f\in W} \inf_{\varphi\in M_n(\Phi)} \|f-\varphi\|. \qquad (1)$$

There has recently been a great interest in both the theoretical and practical aspects of n-term approximation. However, the families Φ in definitions (1) are too general to make them useful. If X is separable and Φ is dense in the unit ball of X, then $\sigma_n(f, \Phi, X) = 0$ for any $f \in X$. One way to deal with this problem is to impose restrictions on Φ and/or on methods of approximation. A (continuous) method in n-term approximation by Φ is represented as a (continuous) mapping S from W into $M_n(\Phi)$. We can restrict the approximation by the elements of $M_n(\Phi)$ with only continuous algorithms and with only the families Φ from $\mathcal{F}(X)$ which consists of all such bounded Φ whose intersection with any finite-dimensional subspace in X is a finite set. The best n-term approximation with these restrictions leads to the non-linear n-width $\tau_n(W, X)$ which is given by

$$\tau_n(W, X) := \inf_{\Phi, S} \sup_{f\in W} \|f - S(f)\|, \qquad (2)$$

where the infimum is taken over all continuous mappings S from W into $M_n(\Phi)$ and all families $\Phi \in \mathcal{F}(X)$. Note that the restriction on the families Φ in the definition (2) is quite natural. All well-known approximation systems satisfy it. Similar to $\tau_n(W, X)$ is the non-linear n-width $\tau'_n(W, X)$ which is defined by formula (2) with the infimum taken over all continuous mappings S from W into $M_n(\Phi)$ and all finite families Φ in X.

A notion of non-linear n-width based on continuous methods of n-term approximation was introduced in [4]. 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 M_n the subset in l_∞ of all $x = \{x_k\}_{k=1}^\infty$ for which $x_k = 0,\ k \notin Q$ for some set of natural numbers Q with cardinality n. Consider the mapping R_Φ from the metric space 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 cardinality n. From the definitions, we can easily see that if the family Φ is bounded, then R_Φ is a continuous mapping from M_n into X and $M_n(\Phi) = R_\Phi(M_n)$. Thus, in this sense, $M_n(\Phi)$ is a non-linear manifold in X, parametrized continuously by M_n. On the other hand, any method 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 M_n. Therefore, if G is required to be continuous, then the method approximation S will also be continuous. The non-linear n-width $\alpha_n(W, X)$ [4] 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 M_n and all bounded families Φ in X.

There are other notions of non-linear n-width which are based on continuous methods of non-linear approximations different from n-term approximation. We would especially like to mention the well-known and very old Alexandroff non-linear n-width $a_n(W, X)$ (for the definition, see, e.g., [3, 8]), the non-linear manifold n-width $\delta_n(W, X)$ introduced in [2] (see also [6]), and the non-linear n-width $\beta_n(W, X)$ which is defined by

$$\beta_n(W, X) := \inf_{R, G} \sup_{f \in W} \|f - R(G(f))\|,$$

where the infimum is taken over all continuous mappings G from W into M_n and R from M_n into X. This non-linear n-width was introduced in [4].

The non-linear n-widths τ_n, τ'_n, α_n, a_n, δ_n and β_n are different. However, they possess some common properties and are closely related. Let W be a compact subset in the quasi-normed linear space X. It was proved that

$$a_n(W, X) \leq \beta_n(W, X) \leq \alpha_n(W, X),$$

$$\delta_{2n+1}(W, X) \leq a_n(W, X) \leq \beta_n(W, X) \leq \delta_n(W, X),$$

(see [4]). We prove that

$$\tau_{n+1}(W,X) \le \tau'_{n+1}(W,X) \le a_n(W,X) \le \tau'_n(W,X),$$

and in addition

$$\alpha_n(W,X) = \tau_n(W,X) = \tau'_n(W,X)$$

for finite-dimensional X.

Our attention is focused on the quantities τ_n, τ'_n, α_n and σ_n which are directly based on n-term approximation. However, because of the close relationship between these non-linear n-widths and δ_n, β_n, a_n, and because in many cases they are asymptotically equivalent, it is quite useful to study them together. We restrict ourselves to consider continuous methods of n-term approximation and non-linear n-widths based on them for classes of functions with common mixed smoothness. The asymptotic degrees of α_n and β_n have been obtained in [4] for the well-known Sobolev and Besov classes of multivariate functions. Non-continuous methods of n-term approximation and the best n-term approximation for classes of functions with bounded mixed derivative or difference are considered in [5, 7]. To obtain lower bounds of $\sigma_n(W,\Phi,X)$ for well-known classes W of functions families Φ should be restricted by some "minimality properties". This approach was considered in [4, 5]. The reader can also consult [1] for a recent survey of various aspects of non-linear approximation and applications, especially n-term approximation.

A central problem in studying non-linear n-widths and the best n-term approximation $\sigma_n(W,\Phi,X)$ of classes of functions is to compute their asymptotic degree if these classes are defined by a common smoothness. In the present paper, we give the asymptotic degree of the above-mentioned non-linear n-widths and the best n-term approximation $\sigma_n(W,\mathcal{V}^d,X)$ by the dictionary $\mathcal{V}^d$ formed from the integer translates of the mixed dyadic scales of the tensor product multivariate de la Vallée Poussin kernel S_k, for Sobolev and Besov type classes of functions with a common mixed smoothness.

Let A be a finite subset in $\mathbf{R}^d$. For $0 < p, \theta \le \infty$, let $B^A_{p,\theta}$ denote the Besov type space of all such functions on the n-dimensional torus $\mathbf{T}^d := [0, 2\pi]^d$, for which

$$\|f\|_{B^A_{p,\theta}} := \|f\|_p + \sum_{\alpha \in A} |f|_{B^\alpha_{p,\theta}} \tag{3}$$

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

$$|f|_{B^\alpha_{p,\theta}} := \left(\int_{\mathbf{T}^d} \prod_{j=1}^d h_j^{-1-\theta\beta_j} \|\Delta^l_h f^{(s)}\|_p^\theta dh \right)^{1/\theta}, \ \theta < \infty,$$

(the integral changed to the supremum for $\theta = \infty$) for some triplet $l \in \mathbf{N}^d$ and $\beta, s \in \mathbf{R}^d$, satisfying the condition $\beta + s = \alpha$; $l_j > \beta_j > 0$, $j = 1, ..., d$, where Δ^l_h denotes the operator of the lth mixed difference with step h. We will use the abbreviation for the special case $A = \{0\} : B_{p,\theta} := B^{\{0\}}_{p,\theta}$. The Sobolev type space W^A_p is defined similarly by replacing $|f|_{B^\alpha_{p,\theta}}$ in (3) with $|f|_{W^\alpha_p} := \|f^{(\alpha)}\|_p$, where $f^{(\alpha)}$ is the mixed derivative in the sense of Weil of order α. Note that the Besov space and Sobolev space are special cases of $B^A_{p,\theta}$ and W^A_p. Denote by SX the unit ball in the space X. The main results of the present paper are the asymptotic degrees of non-linear n-widths and the best n-term approximation by the dictionary $\mathcal{V}^d$ of the Sobolev type class SW_p^A and the Besov type class $SB_{p,\theta}^A$ in the space L_q. It turns out that these asymptotic degrees are closely related to the linear problem

$$(\mathbf{1}, x) \rightarrow \sup, \; x \in A_+^o, \tag{4}$$

where $A_+^o := \{x \in \mathbf{R}^d : (\alpha, x) \le 1, \alpha \in A, \; x_j \ge 0, j = 1, \dots, d\}, \mathbf{1} := (1, 1, \dots, 1) \in \mathbf{R}^d$. Let $1/r$ be the optimal value of (4) and ν the linear dimensions of the set of solutions of (4), i.e.,

$$1/r := \sup\{(\mathbf{1}, x) : x \in A_+^o\}, \; \nu := \dim\{x \in A_+^o : (\mathbf{1}, x) = 1/r\}.$$

For $k \in \mathbf{Z}_+^d := \{k \in \mathbf{Z}^d : k_j \ge 0, \; j = 1, \dots, d\}$, we let the tensor product de la Vallée Poussin kernel S_k be defined by

$$S_k(x) := \prod_{j=1}^d (3k_j)^{-1} V_{k_j}(x_j),$$

where

$$V_\nu(t) := \frac{1}{2} + \sum_{k=1}^{\nu} \cos kt + \sum_{k=\nu+1}^{2\nu} \frac{2\nu - k}{\nu} \cos kt = \frac{\sin(\nu t)\sin(3\nu t/2)}{2\nu \sin^2(t/2)}$$

is the univariate de la Vallée Poussin kernel of order ν. Put $Q_k := \{s \in \mathbf{Z}_+^d : s_j < 3.2^{k_j+1}, j = 1, \dots, d\}$; $h^k := 3^{-1}\pi(2^{-k_1}, \dots, 2^{-k_d})$. We define the family $\mathcal{V}^d$ by

$$\mathcal{V}^d := \{\varphi_s^k\}_{s \in Q_k, k \in \mathbf{Z}_+^d}, \; \varphi_s^k := S_{2^{k+1}}(\cdot - sh^k).$$

Here, we use the notation: $2^x := (2^{x_1}, \dots, 2^{x_d})$ and $xy := (x_1 y_1, \dots, x_d y_d)$ for $x, y \in \mathbf{R}^d$. From well-known properties of de la Vallée Poussin kernels, it follows that $\mathcal{V}^d$ is bounded in L_q and $\mathcal{V}^d \in \mathcal{F}(L_q), 0 < q \le \infty$.

Put $a_+ := \max\{a, 0\}$ and $\mu(A) := \inf\{t > 0 : t\mathcal{E} \in \operatorname{conv}(A \cup \{0\})\}$, where conv$G$ denotes the covex hull of G and $\mathcal{E}$ the canonical basis in $\mathbf{R}^d$. Denote by γ_n any one of $\alpha_n, \tau_n, \tau_n', \beta_n, a_n$ and δ_n.

Theorem 1. *Let $1 < p, q < \infty$, $2 \le \theta \le \infty$ and let A be a finite subset of $\mathbf{R}^d$ with $\mu(A) > \max\{0, d/p - d/q, d/p - d/2\}$. Then we have*

$$\sigma_n(SB_{p,\theta}^A, \mathcal{V}^d, L_q) \asymp \gamma_n(SB_{p,\theta}^A, L_q) \asymp (n/\log^\nu n)^{-r} (\log^\nu n)^{1/2 - 1/\theta}, \tag{5}$$

$$\sigma_n(SW_p^A, \mathcal{V}^d, L_q) \asymp \gamma_n(SW_p^A, L_q) \asymp (n/\log^\nu n)^{-r}. \tag{6}$$

The asymptotic degrees (5) and (6) are achieved by an explicitly constructed positive homogeneous continuous mapping $G^* : Y \longrightarrow M_n$ such that

$$\sup_{f \in SY} \|f - R_{\mathcal{V}^d}(G^*(f))\|_q \asymp E(n),$$

where $E(n)$ is the right-hand side of either (5) or (6) and Y is either $B_{p,\theta}^A$ or W_p^A, respectively.

Theorem 1 is obtained from the following:

Theorem 2. *Let $1 \le p, q \le \infty$, $0 < \tau \le \theta \le \infty$ and let A be a finite subset of $\mathbf{R}^d$ with $\mu(A) > (d/p - d/q)_+$. Then we have*

$$\sigma_n(SB^A_{p,\theta}, \mathcal{V}^d, B_{q,\tau}) \asymp \gamma_n(SB^A_{p,\theta}, B_{q,\tau}) \asymp E_{\theta,\tau}(n), \tag{7}$$

where

$$E_{\theta,\tau}(n) := (n/\log^\nu n)^{-r}(\log^\nu n)^{1/\tau - 1/\theta}.$$

The asymptotic degrees (7) are achieved by an explicitly constructed positive homogeneous continuous mapping $G^* : B^A_{p,\theta} \longrightarrow M_n$ such that

$$\sup_{f \in SB^A_{p,\theta}} \|f - R_{\mathcal{V}^d}(G^*(f))\|_q \asymp E_{\theta,\tau}(n).$$

The asymptotic degree $\gamma_n(SW^A_p, L_q) \asymp (n/\log^\nu n)^{-r}$ has been proved in [3] for a_n and δ_n. The well-known Littlewood–Paley theorem plays a central role in the proofs of Theorems 1 and 2. In these proofs, we employed some methods of discretization based on quasi-norm equivalences for the spaces $B^A_{p,\theta}$ and W^A_p, estimates for non-linear n-widths and a continuous method of n-term approximation by the canonical basis, in spaces of sequences with mixed norm. Establishing the lower bound in Theorem 2 also involves the following lemma.

Lemma 1. *Let the linear space L be quasi-normed by two quasi-norms $\|\cdot\|_X$ and $\|\cdot\|_Y$, and W a subset of L. If Φ is a family of elements in X such that $\sigma_m(W, \Phi, X) > 0$, we have*

$$\sigma_{n+m}(W, \Phi, Y) \le \sigma_n(SX, \Phi, Y)\sigma_m(W, \Phi, X).$$

If $\|\cdot\|_X$ and $\|\cdot\|_Y$ are equivalent, W is compact in these quasi-norms and $\gamma_m(W, X) > 0$, we have

$$\gamma_{n+m}(W, Y) \le \gamma_n(SX, Y)\gamma_m(W, X).$$

References

1. R. DeVore, Nonlinear approximation, *Acta Numerica* **7** (1998) 51–150.
2. R. DeVore, R. Howard, and C. Micchelli, Optimal non-linear approximation, *Manuscripta Mathematica* **63** (1989) 469–478.
3. Dinh Dung, Optimal non-linear approximation of functions with a mixed smoothness, *East J. on Approximations* **4** (1998) 101–112.
4. Dinh Dung, On nonlinear n-widths and n-term approximation, *Vietnam J. Math.* **26** (1998) 165–176.
5. B. Kashin and V. Temlyakov, On best m-term approximation and the entropy of sets in the space L^1, *Mat. Zametki* **56** (1994) 57–86; *Math. Notes* **56** (1994) 1137–1157 (English).
6. P. Mathé, s-Number in information-based complexity, *J. Complexity* **6** (1990) 41–66.
7. V. Temlyakov, Greedy algorithms with regard to the multivariate systems with a special structure (manuscript).
8. V. Tikhomirov, *Some Topics in Approximation Theory*, Moscow State University, Moscow, 1976.