NONSTANDARD MATRIX MULTIPLICATION

Mladen Victor Wickerhauser

Numerical Algorithms Research Group
Department of Mathematics
Yale University
New Haven, Connecticut 06520

15 May 1990

Wave packets. Define wave packets over l^2 in the usual way. For a pair $P = \{p_i\}, Q = \{q_i\}$ of quadrature mirror filters (QMFs) satisfying the orthogonality and decay conditions stated in [CW], there is a unique solution to the functional equation

$$\phi(t) = \sqrt{2} \sum_{j \in \mathbf{Z}} p_j \phi(2t - j).$$

Put $w = w_{0,0,0} = \phi$, and define recursively

$$w_{2n,0,0}(t) = \sqrt{2} \sum_{j \in \mathbf{Z}} p_j w_{n,0,0}(2t - j),$$
$$w_{2n+1,0,0}(t) = \sqrt{2} \sum_{j \in \mathbf{Z}} q_j w_{n,0,0}(2t - j).$$

Then set $w_{nmk}(t) = 2^{m/2} w_{n00}(2^m t - k)$. Write $\mathcal{W}(\mathbf{R}) = \{w_{nmk} : n, m, k \in \mathbf{Z}\}$ for the collection of functions so defined, which we shall call *wave packets*.

The quadrature mirror filters P, Q may be chosen so that $\mathcal{W}(\mathbf{R})$ is dense in many common function spaces. With the minimal hypotheses of [CW], $\mathcal{W}(\mathbf{R})$ will be dense in $L^2(\mathbf{R})$. Using the Haar filters $P = \{1/\sqrt{2}, 1/\sqrt{2}\}, Q = \{1/\sqrt{2}, -1/\sqrt{2}\}$ produces $\mathcal{W}(\mathbf{R})$ which is dense in $L^p(\mathbf{R})$ for $1 < p < \infty$. Longer filters can generate smoother wave packets, so we can also produce dense subsets of Sobolev spaces, etc.

Research supported in part by ONR Grant N00014-88-K0020.

Basis subsets. Define a *basis subset* σ of the set of indices $\{(n,m,k) \in \mathbf{Z}^3\}$ to be any subcollection with the property that $\{w_{nmk} : (n,m,k) \in \sigma\}$ is a Hilbert basis for $L^2(\mathbf{R})$. We characterize basis subsets in [W1]. Abusing notation, we shall also refer to the collection of wave packets $\{w_{nmk} : (n,m,k) \in \sigma\}$ as a basis subset.

Since $L^2 \cap L^p$ is dense in L^p for $1 \leq p < \infty$, with certain QMFs a basis subset will also be a basis for L^p. Likewise, for nice enough QMFs, it will be a Hilbert basis for the various Sobolev spaces.

Since $L^2(\mathbf{R}) \otimes L^2(\mathbf{R})$ is dense in $L^2(\mathbf{R}^2)$, the collection $\{w_X \otimes w_Y : w_X \in \mathcal{W}(X), w_Y \in \mathcal{W}(Y)\}$ is dense in the space of Hilbert-Schmidt operators. Call $\sigma \subset \mathbf{Z}^6$ a basis subset if the collection $\{w_{n_X m_X k_X} \otimes w_{n_Y m_Y k_Y} : (n_X, m_X, k_X, n_Y, m_Y, k_Y) \in \sigma\}$ forms a Hilbert basis. Such two-dimensional basis subsets are characterized in [W2].

Ordering wave packets. Wave packets w_{nmk} can be totally ordered. We say that $w < w'$ if $(m,n,k) < (m',n',k')$. The triplets are compared lexicographically, counting the scale parameter m as most significant.

Tensor products of wave packets inherit this total order. Write $w_X = w_{n_X m_X k_X}$, etc. Then we will say that $w_X \otimes w_Y < w'_X \otimes w'_Y$ if $w_X < w'_X$ or else if $w_X = w'_X$ but $w_Y < w'_Y$. This is equivalent to $(m_X, n_X, k_X, m_Y, n_Y, k_Y) < (m'_X, n'_X, k'_X, m'_Y, n'_Y, k'_Y)$ comparing lexicographically from left to right.

Define the *adjoint order* $<^*$ by exchanging X and Y indices, namely $w_X \otimes w_Y <^* w'_X \otimes w'_Y$ if and only if $w_Y \otimes w_X <^* w'_Y \otimes w'_X$. This is also a total order.

Projections. Let $\mathcal{W}^1$ denote the space of bounded sequences indexed by the three wave packet indices n,m,k. With the ordering above, we obtain a natural isomorphism between l^∞ and $\mathcal{W}^1$. There is also a natural injection $J^1 : L^2(\mathbf{R}) \hookrightarrow \mathcal{W}^1$ given by $c_{nmk} = \langle v, w_{nmk} \rangle$, for $v \in L^2(\mathbf{R})$ and $w_{nmk} \in \mathcal{W}(\mathbf{R})$. If σ is a basis subset, then the composition J^1_σ of J^1 with projection onto the subsequences indexed by σ is also injective. J^1_σ is an isomorphism of $L^2(\mathbf{R})$ onto $l^2(\sigma)$, which is defined to be the square summable sequences of $\mathcal{W}^1$ whose indices belong to σ.

We also have a map $R^1 : \mathcal{W}^1 \to L^2(\mathbf{R})$ defined by $R^1 c(t) = \sum_{(n,m,k) \in \mathbf{Z}^3} c_{nmk} w_{nmk}(t)$.

This map is defined and bounded on the closed subspace of $\mathcal{W}^1$ isomorphic to l^2 under the natural isomorphism mentioned above. In particular, R^1 is defined and bounded on the range of J^1_σ for every basis subset σ. The related restriction $R^1_\sigma : \mathcal{W}^1 \to L^2(\mathbf{R})$ defined by $R^1_\sigma c(t) = \sum_{(n,m,k)\in\sigma} c_{nmk} w_{nmk}(t)$ is a left inverse for J^1 and J^1_σ. In addition, $J^1 R^1_\sigma$ is a projection of $\mathcal{W}^1$. Likewise, if $\sum_i \alpha_i = 1$ and $R^1_{\sigma_i}$ is one of the above maps for each i, then $J^1 \sum_i \alpha_i R^1_{\sigma_i}$ is also a projection of $\mathcal{W}^1$. It is an orthogonal projection on any finite subset of $\mathcal{W}^1$.

Similarly, writing $\mathcal{W}^2$ for $\mathcal{W}^1 \times \mathcal{W}^1$, the ordering of tensor products gives a natural isomorphism between l^∞ and $\mathcal{W}^2$. The space $L^2(\mathbf{R}^2)$, i.e., the Hilbert-Schmidt operators, inject into this sequence space $\mathcal{W}^2$ in the obvious way, namely $M \mapsto \langle M, w_{n_X m_X k_X} \otimes w_{n_Y m_Y k_Y}\rangle$. Call this injection J^2. If σ is a basis subset of $\mathcal{W}^2$, then the composition J^2_σ of J^2 with projection onto subsequences indexed by σ is also injective. J^2_σ is an isomorphism of $L^2(\mathbf{R}^2)$ onto $l^2(\sigma)$, the square summable sequences of $\mathcal{W}^2$ whose indices belong to σ.

The map $R^2 : \mathcal{W}^2 \to L^2(\mathbf{R}^2)$ given by $R^2 c(x,y) = \sum c_{XY} w_X(x) w_Y(y)$, is bounded on that subset of $\mathcal{W}^2$ naturally isomorphic to l^2. In particular, it is bounded on the range of J^2_σ for every basis subset σ.

We may also define the restrictions R^2_σ of R^2 to subsequences indexed by σ, defined by $R^2_\sigma c(x,y) = \sum_{(w_X,w_Y)\in\sigma} c_{XY} w_X(x) w_Y(y)$. There is one for each basis subset σ of $\mathcal{W}^2$. Then R^2_σ is a left inverse of J^2 and J^2_σ, and $J^2 R^2_\sigma$ is a projection of $\mathcal{W}^2$. As before, if $\sum_i \alpha_i = 1$ and σ_i is a basis subset for each i, then $J^2 \sum_i R^2_{\sigma_i}$ is also a projection of $\mathcal{W}^2$. It is an orthogonal projection on any finite subset of $\mathcal{W}^2$.

Applying operators to vectors. For definiteness, let X and Y be two named copies of $\mathbf{R}$. Let $v \in L^2(X)$ be a vector, whose coordinates with respect to wave packets form the sequence $J^1 v = \{\langle v, w_X\rangle : w_X \in \mathcal{W}(X)\}$.

Let $M : L^2(X) \to L^2(Y)$ be a Hilbert-Schmidt operator. Its matrix coefficients with respect to the complete set of tensor products of wave packets form the sequence $J^2 M = \{\langle M, w_X \otimes w_Y\rangle : w_X \in \mathcal{W}(X), w_Y \in \mathcal{W}(Y)\}$. We obtain the identity

$$\langle Mv, w_Y\rangle = \sum_{w_X \in \mathcal{W}(X)} \langle M, w_X \otimes w_Y\rangle \langle v, w_X\rangle$$

This identity generalizes to a linear action of $\mathcal{W}^2$ on $\mathcal{W}^1$ defined by

$$c(v)_{nmk} = \sum_{(n'm'k')} c_{nmkn'm'k'} v_{n'm'k'}.$$

Now, images of operators form a proper submanifold of $\mathcal{W}^2$. Likewise, images of vectors form a submanifold $\mathcal{W}^1$. We can lift the action of M on v to these larger spaces via the commutative diagram

$$\begin{array}{ccc} \mathcal{W}^1 & \xrightarrow{J^2_\sigma M} & \mathcal{W}^1 \\ {\scriptstyle J^1}\uparrow & & \downarrow{\scriptstyle R^1} \\ L^2(\mathbf{R}) & \xrightarrow[M]{} & L^2(\mathbf{R}) \end{array}$$

The significance of this lift is that by a suitable choice of σ we can reduce the complexity of the map $J^2_\sigma M$, and therefore the complexity of the operator application.

Composing operators. Let X, Y, Z be three named copies of $\mathbf{R}$. Suppose that $M : L^2(X) \to L^2(Y)$ and $N : L^2(Y) \to L^2(Z)$ are Hilbert-Schmidt operators. We have the identity

$$\langle NM, w_X \otimes w_Z \rangle = \sum_{w_Y \in \mathcal{W}(Y)} \langle N, w_Y \otimes w_Z \rangle \langle M, w_X \otimes w_Y \rangle.$$

This generalizes to an action of $\mathcal{W}^2$ on $\mathcal{W}^2$, which is defined by the formula

$$c(d)_{nmkn'm'k'} = \sum_{n''m''k''} d_{nmkn''m''k''} c_{n''m''k''n'm'k'},$$

where c and d are sequences in $\mathcal{W}^2$. Using J^2, we can lift multiplication by N to an action on these larger spaces via the commutative diagram

$$\begin{array}{ccc} \mathcal{W}^2 & \xrightarrow{J^2_\sigma N} & \mathcal{W}^2 \\ {\scriptstyle J^2}\uparrow & & \downarrow{\scriptstyle R^2} \\ L^2(\mathbf{R}^2) & \xrightarrow[N]{} & L^2(\mathbf{R}^2) \end{array}$$

Again, by a suitable choice of σ the complexity of the operation may be reduced to below that of ordinary operator composition.

Operation counts: transforming a vector. Suppose that M is a non-sparse operator of rank r. Ordinary multiplication of a vector by M takes at least $O(r^2)$ operations, with the minimum achievable only by representing M as a matrix with respect to the bases of its r-dimensional domain and range.

On the other hand, the injection J^2 will require $O(r^2[\log r]^2)$ operations, and each of J^1 and R^1 require $O(r \log r)$ operations. For a fixed basis subset σ of $\mathcal{W}^2$, the application of $J^2_\sigma M$ to $J^1 v$ requires at most $\#|J^2_\sigma M|$ operations, where $\#|U|$ denotes the number of nonzero coefficients in U. We may choose our wavelet library so that $\#|J^2_\sigma M| = O(r^2)$. Thus the multiplication method described above costs an initial investment of $O(r^2[\log r]^2)$, plus at most an additional $O(r^2)$ per right-hand side. Thus the method has asymptotic complexity $O(r^2)$ per vector in its exact form, as expected for what is essentially multiplication by a conjugated matrix.

We can obtain lower complexity if we take into account the finite accuracy of our calculation. Given a fixed matrix of coefficients C, write C_δ for the same matrix with all coefficients set to 0 whose absolute values are less than δ. By the continuity of the Hilbert-Schmidt norm, for every $\epsilon > 0$ there is a $\delta > 0$ such that $\|C - C_\delta\|_{HS} < \epsilon$. Given M and ϵ as well as a library of wave packets, we can choose a basis subset $\sigma \subset \mathcal{W}^2$ so as to minimize $\#|(J^2_\sigma M)_\delta|$. The choice algorithm has complexity $O(r^2[\log r]^2)$, as shown in [W2]. For a certain class of operators, there is a library of wave packets such that for every fixed $\delta > 0$ we have

$$\text{(S)} \qquad \#|(J^2_\sigma M)_\delta| = O(r \log r),$$

with the constant depending, of course, on δ. We will characterize this class, give examples of members, and find useful sufficient conditions for membership in it. For the moment, call this class with property S the *sparsifiable* Hilbert-Schmidt operators $\mathcal{S}$. By the estimate above, finite-precision multiplication by sparsifiable rank-r operators has asymptotic complexity $O(r \log r)$.

Operation counts: composing two operators. Suppose that M and N are rank-r operators. Standard multiplication of N and M has complexity $O(r^3)$. The complexity of injecting N and M into $\mathcal{W}^2$ is $O(r^2[\log r]^2)$. The action of $J^2_\sigma N$ on $J^2 M$ has complexity $O(\sum_{nmk} \#|J^2_\sigma N_{YZ} : (n_Y, m_Y, k_Y) = (n,m,k)|\#|J^2 M_{XY} : (n_Y, m_Y, k_Y) = (n,m,k)|)$. The second factor is a constant $r \log r$, while the first when summed over all nmk is exactly $\#|J^2_\sigma N|$. Thus the complexity of the nonstandard multiplication algorithm, including the conjugation into the basis set σ, is $O(\#|J^2_\sigma N|\, r \log r)$. Since the first factor is r^2 in general, the complexity of the exact algorithm is $O(r^3 \log r)$ for generic matrices, reflecting the extra cost of conjugating into the basis set σ.

For the approximate algorithm, the complexity is $O(\#|(J^2_\sigma N)_\delta|\, r \log r)$. For the sparsifiable matrices, this can be reduced by a suitable choice of σ to a complexity of $O(r^2[\log r]^2)$ for the complete algorithm. Since choosing σ and evaluating J^2_σ each have this complexity, it is not possible to do any better by this method.

References

[CW] Ronald R. Coifman and M. Victor Wickerhauser, *Best-adapted wave packet bases*, preprint, Yale University (1990).

[CMQW] Ronald R. Coifman, Yves Meyer, Steven Quake and M. Victor Wickerhauser, *Signal processing and compression with wave packets*, preprint, Yale University (1990).

[W1] M. Victor Wickerhauser, *Acoustic signal compression with wave packets*, preprint, Yale University (1989).

[W2] M. Victor Wickerhauser, *Picture compression by best-basis sub-band coding*, preprint, Yale University (1990).

*Current address: Department of Mathematics, University of Georgia, Athens, Georgia 30602