SOME ASPECTS OF ADAPTIVE WAVELET METHODS

D. Černá, V. Finěk

Department of Mathematics and Didactics of Mathematics, Technical University of Liberec

Abstract

Methods based on wavelets are an established tool in signal and image processing and a promising tool for the numerical solution of operator equations. In our contribution, we shortly describe all parts of adaptive wavelet methods and we mainly pay attention to approximate matrix-vector multiplication. Efficient approximate matrix-vector multiplication is enabled by a fast off-diagonal decay of entries of the stiffness matrix and a fast decay of the load vector in wavelet coordinates. We present numerical experiments to compare different approximate matrix-vector multiplication techniques.

1 Introduction

Wavelets have some interesting properties which may provide an advantage over classical methods such as finite element method. It is well-known fact that representations of smooth functions and also representations of a wide class of operators are sparse in wavelet coordinates. Further advantage of wavelet methods is an efficient diagonal preconditioning of stiffness matrices. Then the condition number of the preconditioned stiffness matrices does not depend on the size of matrices. And although wavelet stiffness matrices are only quasi sparse, an approximate multiplication of these matrices with given sparse vectors can be performed in linear complexity. In [3, 4], automatically adaptive and asymptotically optimal wavelet methods were proposed. They consists from the following four steps:

1. To transform a variational formulation into the well-conditioned infinite-dimensional l^2 problem.
2. To find a convergent iteration process for the l^2 problem which works with infinite vectors, the exact right hand side and exact matrix-vector multiplication.
3. To derive a finite dimensional version of above idealized iteration process with an inexact right hand side and approximate matrix-vector multiplication. The algorithm should provide an approximation of the unknown solution up to a given target accuracy ϵ.
4. To optimize a convergence rate which should match the rate of the best N-term approximation, and the associated computational work should be proportional to the number of unknowns.

Efficient approximate matrix-vector multiplication is enabled by a fast off-diagonal decay of entries of the stiffness matrix and a fast decay of the load vector in wavelet coordinates. In [3], a numerical routine **APPLY** was proposed which approximates the exact matrix-vector product with the desired tolerance ϵ and that has linear computational complexity, up to sorting operations. An optimized version of this approach was proposed in [6]. Authors optimize estimated number of matrix-vector multiplication subject to estimated multiplication error. To better utilize actual decay of matrix entries, a modified approach was proposed in [2]. Vector entries are not sorted with respect to their size but instead an actual decay of matrix entries is measured. Consequently in dependence on this decay, the multiplication is performed. Also this approach is asymptotically optimal. At the end, we present numerical experiments to compare different approximate matrix-vector multiplication techniques for wavelet bases proposed in [1]. Computations were carried out in MATLAB.

2 Wavelet Bases

First we introduce concepts from the wavelet theory and notations. We consider the domain $\Omega \subset \mathbb{R}^d$ and the space $L^2(\Omega)$ and we denote by $\langle \cdot, \cdot \rangle$ and $\|\cdot\|$ the L^2-inner product and the L^2-norm, respectively. Let $\mathcal{J}$ be at most countable index set and let each index $\lambda \in \mathcal{J}$ take the form $\lambda = (j,k)$, $j,k \in \mathbb{Z}$, where $|\lambda| := j$ is a *scale* or a *level*. Let

$$l^2(\mathcal{J}) := \left\{ v = \{v_\lambda\}_{\lambda \in \mathcal{J}} : v_\lambda \in \mathbb{R} \text{ and } \sum_{\lambda \in \mathcal{J}} |v_\lambda|^2 < \infty \right\}.$$

A family $\Psi := \{\psi_\lambda, \lambda \in \mathcal{J}\} \subset L^2(\Omega)$ is called a *wavelet basis* of $L^2(\Omega)$, if

i) Ψ is a Riesz basis for $L^2(\Omega)$, i.e. the closure of the linear span of Ψ is complete in $L^2(\Omega)$ and there exist constants $c, C \in (0,\infty)$ such that

$$c \|\mathrm{b}\|_{l^2(\mathcal{J})} \leq \left\| \sum_{\lambda \in \mathcal{J}} b_\lambda \psi_\lambda \right\| \leq C \|\mathrm{b}\|_{l^2(\mathcal{J})}, \quad \mathrm{b} := \{b_\lambda\}_{\lambda \in \mathcal{J}} \in l^2(\mathcal{J}). \tag{1}$$

Constants $c_\psi := \sup\{c : c \text{ satisfies } (1)\}$, $C_\psi := \inf\{C : C \text{ satisfies } (1)\}$ are called *Riesz bounds* and *cond* $\Psi = C_\psi / c_\psi$ is called the *condition number* of Ψ.

ii) The functions are local in the sense that $\operatorname{diam}(\Omega_\lambda) \leq \mathrm{C} 2^{-|\lambda|}$ for all $\lambda \in \mathcal{J}$, where Ω_λ is the support of ψ_λ, and at a given level j the supports of only finitely many wavelets overlap in any point $x \in \Omega$.

iii) The primal scaling basis has polynomial exactness of order $N \in \mathbb{N}$ and the dual scaling basis has polynomial exactness of order $\tilde{N} \in \mathbb{N}$. It means that the space generated by an scaling basis contains the space of all algebraic polynomials on Ω of degree less than or equal to N ($\tilde{N}$).

By the Riesz representation theorem, there exists a unique family $\tilde{\Psi} = \left\{ \tilde{\psi}_\lambda, \lambda \in \mathcal{J} \right\}$ in $L^2(\Omega)$ biorthogonal to Ψ, i.e. $\left\langle \psi_{i,k}, \tilde{\psi}_{j,l} \right\rangle = \delta_{i,j}\delta_{k,l}$, for all $(i,k), (j,l) \in \mathcal{J}$. This family is also a Riesz basis for $L^2(\Omega)$. The basis Ψ is called a *primal* wavelet basis, $\tilde{\Psi}$ is called a *dual* wavelet basis.

3 Wavelet Discretization

Let H be a real Hilbert space with the inner product $(\cdot,\cdot)_H$ and the induced norm $\|\cdot\|_H$. Let $A : H \to H'$ be the selfadjoint and H- elliptic differential operator, i.e.

$$a(v,w) := (Av,w) \lesssim \|v\|_H \|w\|_H \qquad \text{and} \qquad a(v,v) \sim \|v\|_H^2.$$

Then, there exist positive constants c_A and C_A such that

$$c_A \|v\|_H \leq \|Av\|_{H'} \leq C_A \|v\|_H, \qquad v \in H$$

and the equation $Au = f$ has for any $f \in H'$ a unique solution. Further we assume that $\mathbf{D}^{-1}\Psi$, $\Psi = \{\psi_\lambda, \lambda \in I\}$, is a suitable wavelet (Riesz) basis in the energy space H and I an index set. Then, there exist positive constants c_ψ and C_ψ such that

$$c_\psi \|\mathbf{v}\|_2 \leq \left\| \mathbf{v}^T \mathbf{D}^{-1} \Psi \right\|_H \leq C_\psi \|\mathbf{v}\|_2, \qquad \mathbf{v} \in l^2(I) \tag{2}$$

and consequently

$$Au = f \qquad \Leftrightarrow \qquad \mathbf{A}\mathbf{u} = \mathbf{f},$$

where $\mathbf{D} := \operatorname{diag}(\omega_\lambda)_{\lambda\in I}$, $\omega_\lambda = \sqrt{(A\psi_\lambda, \psi_\lambda)}$, $\mathbf{A} = \mathbf{D}^{-1}(A\Psi, \Psi)\mathbf{D}^{-1}$ is a biinfinite diagonally preconditioned stiffness matrix, $u = \mathbf{u}^T\mathbf{D}^{-1}\Psi$ and $\mathbf{f} = \mathbf{D}^{-1}(f, \Psi)$. The condition number of matrix $\mathbf{A}$ satisfies

$$\kappa(\mathbf{A}) \leq \frac{C_\psi^2 C_A}{c_\psi^2 c_A} < +\infty \tag{3}$$

and the same holds (matrix $\mathbf{A}$ is positive definite) for all finite sections

$$\mathbf{A}_\Lambda := \mathbf{D}^{-1}(A\Psi_\Lambda, \Psi_\Lambda)\mathbf{D}^{-1}, \qquad \Psi_\Lambda := \{\psi_\lambda, \lambda \in \Lambda\}, \qquad \Lambda \subset I.$$

4 Approximate Matrix-Vector Multiplications

In [3], authors exploited an off-diagonal decay of entries of the wavelet stiffness matrices and a decay of entries of the load vector in wavelet coordinates to design a numerical routine **APPLY** which approximates the exact matrix-vector product with the desired tolerance ϵ and that has linear computational complexity, up to sorting operations. An example of the decay of matrix entries, we can observe in the Table 1. The idea of **APPLY** for one dimensional problems is the following: To truncate $\mathbf{A}$ in scale by zeroing $a_{i,j}$ whenever $\delta(i,j) > k$ (δ represents the level difference of two functions in the wavelet expansion) and denote resulting matrix by $\mathbf{A_k}$. At the same time, vector entries $\mathbf{v}$ are sorted with respect to the size of their absolute values. One obtains $\mathbf{v_k}$ by retaining 2^k biggest coefficients in absolute values of $\mathbf{v}$ and setting all other equal to zero. The maximum value of k is determined to reach a desired accuracy of approximation. Then one computes an approximation of $\mathbf{Av}$ by

$$\mathbf{w} := \mathbf{A_k v_0} + \mathbf{A_{k-1}}(\mathbf{v_1} - \mathbf{v_0}) + \ldots + \mathbf{A_0}(\mathbf{v_k} - \mathbf{v_{k-1}}) \tag{4}$$

with the aim to balance both accuracy and computational complexity.

k	(3,3)	(3,5)	(4,4)	(4,6)
1	0.19035062	0.39681820	0.24313996	0.20842228
2	0.03148893	0.07742380	0.05924896	0.03254208
3	0.00941647	0.00425257	0.00752139	0.00519580
4	0.00332923	0.00145692	0.00135003	0.00081173
5	0.00117706	0.00051510	0.00023865	0.00014350
6	0.00041615	0.00018211	0.00004219	0.00002537
7	0.00014713	0.00006439	0.00000746	0.00000448
8	0.00005202	0.00002276	0.00000132	0.00000062

Table 1: Computed $\max_{\delta(i,j)=k} a_{i,j}$ for several wavelet basis from [1].

In [5], binning and approximate sorting was used to eliminate sorting costs and then an algorithm with linear complexity was obtained. The idea is following: Reorder the elements of $\mathbf{v}$ into the sets $V_0, \ldots, V_q$, where $v_\lambda \in V_i$ if and only if

$$2^{-i-1}\|\mathbf{v}\|_2 < v_\lambda < 2^{-i}\|\mathbf{v}\|_2, \quad 0 \leq i < q.$$

And then to generate vectors $\mathbf{v_k}$ by successively extracting 2^k elements from $\bigcup_i V_i$, starting from V_0 and when it is empty continuing with V_1 and so forth. Finally the scheme (4) is applied.

An optimized version of the above approach was proposed in [6]. The indices of $\mathbf{v}$ are stored in buckets, depending on the modulus of the corresponding wavelet coefficients in this way:

$$v_\lambda \in \mathbf{v_k} \qquad \Longleftrightarrow \qquad 2^{-(k+1)/2}\|\mathbf{v}\|_\infty < |v_\lambda| \leq 2^{-k/2}\|\mathbf{v}\|_\infty.$$

Then $\forall\, \mathbf{v} \in l_2(\Lambda)$, we compute the approximate matrix vector product by $\sum_{k=0} \mathbf{A}_{j_k}\mathbf{v_k}$, where $j_k \in \mathbf{N}_0$ are solutions of

$$\sum_{k=0} c_{j_k} \# \mathbf{v_k} \longrightarrow \min!, \qquad \sum_{k=0} e_{j_k} ||\mathbf{v_k}|| \leq \epsilon - \delta, \tag{5}$$

$$\text{and} \quad \delta = ||\mathbf{A}|| \left|\left| v - \sum_{k=0} \mathbf{v_k} \right|\right| \leq \epsilon/2, \tag{6}$$

where $\mathbf{A_j}$ and e_j are matrices and constants such that $||\mathbf{A} - \mathbf{A_j}|| \leq e_j$ and c_j are upper bounds for the number of non-zero entries in each column of $\mathbf{A_j}$. So, they try to optimize number of arithmetic operations.

To better utilize the actual decay of matrix and vector entries, in [2] a different approach was designed. We are not searching for 2^k biggest vector entries in absolute value but instead we trace actual decay of matrix and vector entries and then the actual number of entries in $\mathbf{v_k}$ depends on these decays. Let us denote $S_{Ak} := \max\{|a_{i,j}|,\, \delta(i,j) = k\}$. Then, we multiply matrix $\mathbf{A_0}$ with vector entries which are greater than given tolerance ϵ, matrix $\mathbf{A_1} - \mathbf{A_0}$ with vector entries which are greater than ϵ/S_{A1}, ..., and matrix $\mathbf{A_k} - \mathbf{A_{k-1}}$ with vector entries which are greater than ϵ/S_{Ak}. In [2], an asymptotic optimality of this multiplication algorithm was proved.

5 Numerical Experiments

At the end, we present numerical comparison of different approximate matrix-vector multiplication techniques proposed in [3], [6], and in [2]. In numerical experiments, we employ the stiffness matrix corresponding to the discretization of the one dimensional Poisson equation in wavelet coordinates using the wavelet basis (4,4) proposed in [1]. As a testing vector, we used the expansion of function $e^x(x^2 + 4x + 1)$ in a dual basis on interval $[-1, 1]$.

	CF		DSS		CDV	
ϵ	#	$\|\mathbf{Au} - \mathbf{w}\|_{l_2}$	#	$\|\mathbf{Au} - \mathbf{w}\|_{l_2}$	#	$\|\mathbf{Au} - \mathbf{w}\|_{l_2}$
4^{-1}	283	0.01125725	416	0.02180305	555	0.00440601
4^{-2}	466	0.00362875	640	0.00426107	1004	0.00072592
4^{-3}	717	0.00076026	1070	0.00008879	1334	0.00072560
4^{-4}	1100	0.00028663	1755	0.00004870	2748	0.00012386
4^{-5}	2097	0.00004649	3057	0.00002449	5839	0.00002194
4^{-6}	3402	0.00001238	6374	0.00001290	13314	0.00000388
4^{-7}	6684	0.00000296	9135	0.00000625	14103	0.00000388
4^{-8}	13704	0.00000045	17360	0.00000046	31013	0.00000068

Table 2: Errors of matrix-vector multiplications $\mathbf{Au} - \mathbf{w}$ for (4,4) wavelet.

In the Table 2, the first column represents the required precision ϵ. Columns denoted by CF contain results obtained by the method proposed in [2], columns denoted by DSS contain results obtained by approach (5) proposed in [6], and finally columns denoted by CDV contain results obtained by approach proposed in [3]. In case of method [2], we gradually updated bins until the desired estimated precision has been reached. In case of method [3], we gradually increased the controlling parameter k in (4) until the desired estimated precision has been reached. The error estimates were computed in the similar way as in [6]. Only difference consists in the absence of restriction (6). Presented results affirm that the approximate matrix-vector multiplication technique proposed in [2] is more efficient than methods proposed in [3, 6]. In Table 3, we present results obtained by the method proposed in [2] where the errors were approximated and not estimated as before. One can observe that the prescribed errors were achieved with substantially smaller numbers of operations.

	CF_{aprox}	
ϵ	#	$\|\mathbf{Au}-\mathbf{w}\|_{l_2}$
4^{-1}	58	0.18904028
4^{-2}	209	0.05715800
4^{-3}	418	0.00871160
4^{-4}	652	0.00210840
4^{-5}	861	0.00088876
4^{-6}	1366	0.00024345
4^{-7}	2644	0.00005249
4^{-8}	4513	0.00001456

Table 3: Errors of matrix-vector multiplications $\mathbf{Au}-\mathbf{w}$ for (4,4) wavelet.

Acknowledgements: This research has been supported by the project ESF "Constitution and improvement of a team for demanding technical computations on parallel computers at TU Liberec" No. CZ.1.07/2.3.00/09.0155.

References

[1] ČERNÁ, D.; FINĚK, V. *Construction of Optimally Conditioned Cubic Spline Wavelets on the Interval*. Adv. Comput. Math., vol. 34, 2011, pp. 519-552.

[2] ČERNÁ, D.; FINĚK, V. *Approximate Multiplication in Adaptive Wavelet Methods*. Submitted, 2010.

[3] COHEN, A.; DAHMEN, W.; DEVORE, R. *Adaptive Wavelet Schemes for Elliptic Operator Equations - Convergence Rates*. Math. Comput., vol. 70, no. 233, 2001, pp. 27-75.

[4] COHEN, A.; DAHMEN, W.; DEVORE, R. *Adaptive Wavelet Methods II - Beyond the elliptic case*. Found. Math., vol. 2, no. 3, 2002, pp. 203-245.

[5] STEVENSON, R. *Adaptive solution of operator equations using wavelet frames*. SIAM J. Numer. Anal., vol. 41, no. 3, 2003, pp. 1074 - 1100.

[6] DIJKEMA, T. J., SCHWAB Ch.; STEVENSON R., *An adaptive wavelet method for solving high-dimensional elliptic PDEs*, Constructive approximation, **30** (2009), no. 3, 423-455.

Dana Černá
Department of Mathematics and Didactics of Mathematics
Technical University of Liberec
Studentská 2
Liberec 461 17, Czech Republic
dana.cerna@tul.cz

Václav Finěk
Department of Mathematics and Didactics of Mathematics
Technical University of Liberec
Studentská 2
Liberec 461 17, Czech Republic
vaclav.finek@tul.cz