JOURNAL OF MAGNETIC RESONANCE 82, 150–155 (1989)

Rapid Singular Value Decomposition for Time-Domain Analysis of Magnetic Resonance Signals by Use of the Lanczos Algorithm

GLENN L. MILLHAUSER,* ALISON A. CARTER, DAVID J. SCHNEIDER,†‡
JACK H. FREED,† AND ROBERT E. OSWALD

Department of Pharmacology, NYS College of Veterinary Medicine, Cornell University, Ithaca, New York 14853

Received April 14, 1988; revised July 18, 1988

Time-series analysis has become an integral part of signal interpretation in magnetic resonance and other experiments. As an example, linear prediction with singular value decomposition (LPSVD) (*1, 2*) and the related method, Hankel SVD (HSVD) (*3*), have been recently applied to problems in NMR (*2*), 2D NMR (*4*), 2D electron spin-echo spectroscopy (2D ESE) (*5, 6*), Fourier transform ESR correlation spectroscopy (*7*), time-resolved femtosecond spectroscopy (*8*), and nonstationary neurological currents (*9*). The minimal goal of time-series analysis is to separate signal from noise. Spectral transform methods, such as FFT filtering (*10*), maximum entropy (MEM) (*11*), and LPZ (*12*), replace the original data with a new time series (or its spectrum) that has an enhanced signal-to-noise ratio. In contrast spectral decomposition methods, such as LPSVD or HSVD, which assume a certain functional form for the time series, provide both a recipe for separating signal from noise and a listing of the harmonic components contained in the original data. This harmonic list simplifies spectral analysis and can be used to (1) extend the noise-reduced time series, (2) selectively reconstruct certain regions of the harmonic spectrum, and (3) data-compress the signal. The substantial benefits of spectral decomposition are often offset by the requirements of considerable computational time and memory storage. For a time series of N data points spectral transform techniques usually require computational time of $O(N \log_2 N)$ to $O(N^2)$ and storage of $O(N)$ whereas spectral decomposition, such as HSVD, requires computational time of $O(N^3)$ and storage of $O(N^2)$.

We have investigated a new approach for solving the spectral decomposition problem which is based on the Lanczos algorithm (LA) (*13*). In the past the LA has provided an effective means for solving slow-motional magnetic resonance spectra (*14, 15*) as well as other complicated problems in chemical physics (*16, 17*). The LA is an extremely fast method for solving large eigenelement problems when either (1)

* Present address: Department of Chemistry, University of California, Santa Cruz, California 95064.
† Department of Chemistry, Cornell University, Ithaca, New York 14853.
‡ Present address: IBM Corporation, Data Systems Division, Dept. 48B/MS 428, Neighborhood Road, Kingston, New York 12401.

0022-2364/89 $3.00

all the eigenelements of a sparse matrix are required or (2) a small set of relevant eigenelements of a dense (or sparse) matrix of special structure are required. (A Lanczos–Prony method for time-series problems has been previously proposed (*18*), but it does not benefit from SVD as an intermediate step nor is it based on the LA that we employ.) The goal of this report is to discuss our recent findings on the applicability of the LA for spectral decomposition. We base our approach on the HSVD (*3*). The time-consuming singular value decomposition is performed with the LA and we find that the computational time for Lanczos-HSVD (LA-HSVD) is of $O(N^2)$ and the storage is of $O(N)$.

The HSVD, as developed by Barkhuijsen *et al.*, is based on the state space theory of Kung *et al.* (*19*). Given a time series of the form $\{x_0, x_1, x_2, \ldots, x_{N-1}\}$ the first step is to form the rectangular Hankel data matrix of order M which is given by

$$\mathbf{H} = \begin{bmatrix} x_0 & x_1 & x_2 & \cdots & x_{M-1} \\ x_1 & x_2 & x_3 & & x_M \\ \cdot & \cdot & \cdot & & \cdot \\ x_{N-M} & \cdot & \cdot & & x_{N-1} \end{bmatrix}. \qquad [1]$$

A singular value decomposition is then performed on $\mathbf{H}$,

$$\mathbf{H} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^{\mathrm{T}}, \qquad [2]$$

where $\boldsymbol{\Sigma}$ is the diagonal matrix of singular values, and $\mathbf{U}$ and $\mathbf{V}^{\mathrm{T}}$ are formed from the left and right singular vectors, respectively. The matrix of left singular vectors, $\mathbf{U}$, is used to find the frequencies and time constants of the time series. The details of this calculation are elaborated by Barkhuijsen *et al.* (*3*). By far the most time-consuming part of the total HSVD calculation is the decomposition in Eq. [2] and it is here where we apply the LA.

Because the parameter M is chosen to be much larger than the number of expected harmonics $\mathbf{H}$ must have a rank less than the value of M. The matrix, $\boldsymbol{\Sigma}$ (of dimension $\min(N - M + 1, M)$), which contains the singular values, $\{\sigma_i\}$, will have both nonzero and zero diagonal entries. The number of nonzero σ_i's is the rank of $\mathbf{H}$, $R(\mathbf{H})$, and the corresponding $R(\mathbf{H})$ vectors in $\mathbf{U}$ provide an orthonormal basis for the range of $\mathbf{H}$. If the time series is corrupted by noise then $\boldsymbol{\Sigma}$ will have $\min(N - M + 1, M) - R(\mathbf{H})$ small singular values, the magnitude of which will be related to the signal-to-noise ratio. Thus, the signal-associated σ_i's are used as a map to the relevant vectors contained in $\mathbf{U}$.

There are a number of methods for performing singular value decompositions and we briefly mention two of these. With the normal equations (NE) approach, the eigenelements of $\mathbf{H}\mathbf{H}^{\mathrm{T}}$ are used to determine the singular values and vectors (*20*). If M is chosen so that $\mathbf{H}$ is approximately square (i.e., $M \approx N/2$) the number of floating point operations (the FLOP count) will be $O(N^3)$ and the storage will be $O(N^2)$ (*20*). A slower, but typically more stable algorithm (*21*), is the Golub–Reinsch (GR) SVD which operates directly on the matrix $\mathbf{H}$ (and not $\mathbf{H}\mathbf{H}^{\mathrm{T}}$) and requires $O(N^3)$ FLOPs (with a multiplicative factor of about 3 greater than for the NE) and storage similar to that of the NE method.

The Lanczos algorithm is a three-term recursion for generating a select (or complete) set of eigenelements of a large matrix. The SVD problem can be reduced to an eigenelement problem by construction of the real symmetric matrix

$$\mathbf{B} = \begin{bmatrix} \mathbf{0} & \mathbf{H} \\ \mathbf{H}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}. \qquad [3]$$

The σ_i's of $\mathbf{H}$ will occur in $\pm$ pairs as the eigenvalues of $\mathbf{B}$. By applying the LA to $\mathbf{B}$ the recursion becomes

$$\beta_{2i}\mathbf{u}_i = \mathbf{H}^{\mathrm{T}}\mathbf{v}_i - \beta_{2i-1}\mathbf{u}_{i-1}$$

$$\beta_{2i+1}\mathbf{v}_{i+1} = \mathbf{H}\mathbf{u}_i - \beta_{2i}\mathbf{v}_i \qquad [4]$$

and the calculation is initiated with a random starting vector $\mathbf{v}_1$ and $\beta_1 = 0$ (*22*). Iteration of Eq. [4] produces two sets of orthogonal vectors, $\{\mathbf{u}_i\}$ and $\{\mathbf{v}_i\}$, and a set of coefficients $\{\beta_i\}$. After $j/2$ iterations the j-dimensional tridiagonal Lanczos matrix is defined as

$$\mathbf{T}_j = \begin{bmatrix} 0 & \beta_2 & 0 & \cdots & & 0 \\ \beta_2 & 0 & \beta_3 & \cdots & & 0 \\ 0 & \beta_3 & 0 & \cdots & & 0 \\ \cdot & \cdot & \cdot & & & \cdot \\ 0 & 0 & 0 & \cdots & & \beta_j \\ 0 & 0 & 0 & \cdots & \beta_j & 0 \end{bmatrix} \qquad [5]$$

and $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{j/2}\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{j/2}\}$ are essentially the Lanczos vectors. For j sufficiently large, diagonalization of $\mathbf{T}_j$ will yield eigenvalues (Ritz values) that are close approximations to those from $\mathbf{B}$. Likewise, the eigenvectors from $\mathbf{T}_j$ along with $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{j/2}\}$ and $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{j/2}\}$ can be used to construct a set of orthonormal vectors (Ritz vectors) which are close approximations to the eigenvectors of $\mathbf{B}$. These eigenvectors contain the left and right singular vectors of $\mathbf{H}$. Details of these calculations are reviewed elsewhere (*22*).

The Lanczos calculation in exact arithmetic proceeds as follows. We assume that $\mathbf{v}_1$ has a nonzero projection on all of the singular vectors of $\mathbf{H}$. Successive applications of the Lanczos recursion generates new orthonormal vectors in this space. When a basis for the entire vector space generated by $\mathbf{H}$ is complete, no new orthonormal vectors can be created, and this is signaled by a value of zero for the most recent β_j. An N by j matrix, $\mathbf{Q}_j$, constructed from the Lanczos vectors along with Eq. [4] shows how the Lanczos procedure generates a similarity transform of $\mathbf{B}$,

$$\mathbf{B}\mathbf{Q}_j = \mathbf{Q}_j\mathbf{T}_j. \qquad [6]$$

(Note that if $\mathbf{H}$ is not square some of the zero eigenvalues of $\mathbf{B}$ will not correspond to σ_i's of $\mathbf{H}$.) For the purposes of the SVD problem it is a beneficial consequence of the LA that extremal well-separated eigenvalues of a matrix spectrum are the first projected for small j. The spectrum of $\mathbf{B}$ is bounded by $[-\sigma_1, \sigma_1]$ and most of the eigenvalues are clustered near zero (when $\mathbf{H}$ is overdetermined). Therefore, as j approaches $2R(\mathbf{H})$ only the signal-associated σ_i's will result from diagonalization of $\mathbf{T}_j$.

From the arguments above we can deduce the enormous computational savings

afforded by the LA. Examination of Eq. [4] shows that **H** is never modified (which is in contrast with other diagonalization schemes). Because there are only N independent elements in **H**, only these elements are required and the matrix vector products can be computed without actual generation of the matrix form of **H**. Furthermore, only the most recently calculated vectors, $\mathbf{v}_i$ and $\mathbf{u}_{i-1}$, are needed for each new iteration—all of the previous vectors can be stored on disk for later generation of the selected singular vectors. The storage requirements, therefore, are $O(N)$.

One matrix–vector multiplication in this case requires $O(M(N - M))$ FLOPs. As discussed above, iteration of Eq. [4] can be terminated when $j \geqslant 2R(\mathbf{H})$ so the computational time is of $O(R(\mathbf{H})M(N-M))$ or $O(R(\mathbf{H})N^2)$ when $M \approx N/2$. This computational savings over the standard NE or GR SVDs reflects the fact that only the signal-associated singular values and vectors are calculated (when the signal-to-noise ratio is sufficiently large).

Practical application of LA-SVD requires knowledge of when iteration of Eq. [4] should be terminated. When the signal-to-noise is reasonably high (>10) we find that this convergence is signaled by a loss of orthogonality of the most recent vector, $\mathbf{v}_j$, with respect to the starting vector. For low S/N signals one must monitor the large singular values calculated from $\mathbf{T}_j$ to determine when further iteration will not affect their values.

In Fig. 1 we show a plot of the singular values in descending order from a 99 point signal containing two harmonics and noise with a signal-to-noise ratio of 10. The first four σ_i's contain the signal information. We construct **H** as a square matrix because in past work we have found that this gives the most stable results. This choice also simplifies the arithmetic in Eq. [4] since $\mathbf{H}^{\mathrm{T}} = \mathbf{H}$. Application of the GR SVD generates 50 σ_i's and the first 20 are plotted. The LA SVD signaled convergence after seven iterations and the calculated σ_i's are plotted. It is clear that the signal-associated singular values are accurately calculated by the LA approach. In Table 1 we compare

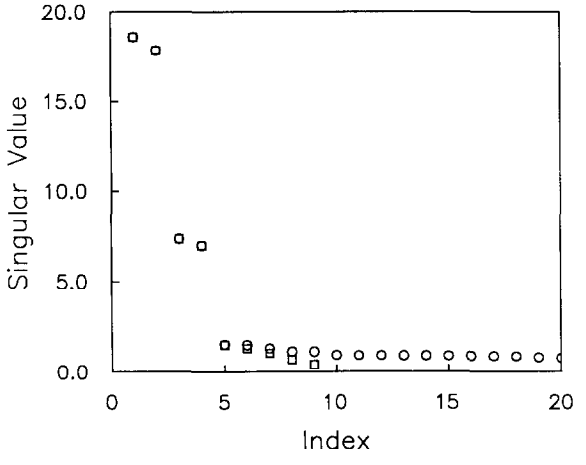

FIG. 1. Singular values plotted in decreasing order for a two-component signal. The GR SVD (circles) and Lanczos SVD (squares) are compared. Note that for the Lanczos approach only the signal-associated singular values need to be calculated.

TABLE 1

	Lanczos SVD		Gollub-Reinsch SVD	
Signal/noise	Component 1	Component 2	Component 1	Component 2
∞ (Amplitude)	1.0000	1.0000[a]	1.0000	1.0000
(Frequency)	2.5000	3.5000	2.5000	3.5000
(Time-constant)	2.0000	0.5000	2.0000	0.5000
(Phase in °)	0.0000	0.0000	0.0002	0.0006
1000	1.0000	1.0014[b]	1.0000	1.0014
	2.5006	3.4987	2.5006	3.4987
	2.0039	0.5001	2.0041	0.5001
	0.1217	0.1417	0.1220	0.1426
100	0.9955	1.0088[c]	0.9958	1.0089
	2.5019	3.4956	2.5018	3.4956
	2.0306	0.5009	2.0306	0.5009
	0.3911	0.4163	0.3922	0.4263
10	1.2092	0.8418[d]	1.2069	0.8404
	2.5096	3.5490	2.5096	3.5490
	1.1845	0.6431	1.1845	0.6430
	0.0918	3.6928	0.1107	3.6483
1	2.3336[e,f]		2.3342[f]	
	2.7878		2.7878	
	0.2310		0.2310	
	5.0464		5.0021	

Note. Number of required Lanczos iterations: [a] 5, [b] 5, [c] 6, [d] 7, and [e] 15.
[f] Only one component detected by SVD.

analysis of the signal, $\exp(-t/2)\cos(2\pi 2.5t) + \exp(-t/0.5)\cos(2\pi 3.5t)$, with Gaussian white noise added in the indicated proportions. For amplitudes, frequencies, and time constants (*2, 3*), the two methods give the same results to within better than 1%.

As a computational benchmark we applied both methods to a 250 point time series again containing two harmonic components. We perform our calculations on a Micro VAX II workstation and the arithmetic for the LA is double precision. The GR SVD requires approximately 54 s. Ten Lanczos iterations and diagonalization of the Lanczos matrix required 4 s. (In both cases the remainder of the calculation after the SVD required only 1 s.) In general, we find that we can now easily solve problems that are an order of magnitude larger than those solvable with the GR SVD. Furthermore, the structure of the LA allows for efficient vectorization and/or parallel processing so that large two-dimensional magnetic resonance problems may be rapidly solved with a large-scale computer or array processor. Even greater efficiency may be possible by taking advantage of the fast Hankel matrix–vector multiplication to reduce the complexity of the Lanczos matrix–vector multiplication from $O(N^2)$ to $O(N \log_2 N)$ (*23*), with hard-wiring of the calculation becoming feasible. The methods developed here can also be generalized for complex time series by substitution of

the Hermitian conjugate, $\mathbf{H}^H$, for $\mathbf{H}^T$ in Eqs. [3] and [4]. Furthermore, these methods are not restricted to square matrices.

A problem often encountered with the LA is a gradual loss of orthogonality of the Lanczos vectors due to round-off error. There are now several procedures for dealing with this problem including reorthogonalization (*21, 22*), and we plan to discuss them elsewhere. Based upon past experience with the LA (*14–17, 21, 22*) we anticipate that these methods will be applicable to cases of large M as well as large N. In this initial study we found that round-off error became a problem only after calculation of the important σ_i's in the examples considered here. We are greatly motivated by the enormous computer time and memory savings from the LA-HSVD approach, and more general Lanczos HSVD techniques will be the subject of future studies.

ACKNOWLEDGMENTS

We thank Dr. Jeff Gorcester for several helpful discussions. This work was supported by grants from the National Institutes of Health 1 F32 NS 07871-03 (G.L.M.), 1 R01 NS 18660-04 (R.E.O. and A.A.C.), and GM 25862 (J.H.F. and D.J.S.) and the National Science Foundation CHE 8703014 and DMR 86-04200 (J.H.F. and D.J.S.).

REFERENCES

1. R. KUMARESAN AND D. TUFTS, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-30,** 833 (1982).
2. H. BARKHUIJSEN, R. DE BEER, W. M. M. J. BOVEE, AND D. VAN ORMONDT, *J. Magn. Reson.* **61,** 465 (1985).
3. H. BARKHUIJSEN, R. DE BEER AND D. VAN ORMONDT, *J. Magn. Reson.* **73,** 553 (1987).
4. A. E. SCHUSSHEIM AND D. COWBURN, *J. Magn. Reson.* **71,** 371 (1987).
5. G. L. MILLHAUSER AND J. H. FREED, *J. Chem. Phys.* **85,** 63 (1986).
6. G. L. MILLHAUSER, J. GORCESTER, AND J. H. FREED, "Electronic Magnetic Resonance of the Solid State" (J. A. Weil, Ed.), Canadian Society for Chemistry, Ottawa, 1987.
7. J. GORCESTER AND J. H. FREED, *J. Magn. Reson.* **78,** 292 (1988).
8. F. W. WISE, M. J. ROSKER, G. L. MILLHAUSER, AND C. L. TANG, *IEEE J. Quantum Electron.* **QE-23,** 1116 (1987).
9. G. L. MILLHAUSER, S. M. SIMASKO, AND R. E. OSWALD, *Neurosci. Abstr.* **12,** 41 (1986).
10. T. INOUYE, T. HARPER, AND N. C. RASMUSSEN, *Nucl. Instrum. Methods* **67,** 125 (1969).
11. S. HAYKIN, "Adaptive Filter Theory," Prentice–Hall, Englewood Cliffs, New Jersey, 1986.
12. J. TANG AND J. R. NORRIS, *J. Chem. Phys.* **84,** 5210 (1986).
13. C. LANCZOS, *J. Res. Natl. Bur. Stand.* **45,** 255 (1950).
14. G. MORO AND J. H. FREED, *J. Chem. Phys.* **74,** 3757 (1981).
15. K. V. VASAVADA, D. J. SCHNEIDER, AND J. H. FREED, *J. Chem. Phys.* **86,** 647 (1987).
16. G. MORO AND J. H. FREED, "Large Scale Eigenvalue Problems" (J. Cullum and R. A. Willoughby, Eds.), Elsevier/North-Holland, New York, 1986.
17. D. J. SCHNEIDER AND J. H. FREED, *in* "Advances in Chemical Physics," Vol. 73, "Lasers, Molecules Methods" (J. O. Hirshfelder *et al.*, Eds.), p. 387, Wiley, New York, 1989.
18. D. TUFTS AND C. D. MELISSINOS, *IEEE Trans. Acoust. Speech Signal Process.* **ASSP-34,** 1046 (1986).
19. S. Y. KUNG, K. S. ARUN, AND D. V. BHASKAR RAO, *J. Opt. Soc. Am.* **73,** 1799 (1983).
20. H. BARKHUIJSEN, R. DE BEER, AND D. VAN ORMONDT, *J. Magn. Reson.* **64,** 343 (1985).
21. G. H. GOLUB AND C. F. VAN LOAN, "Matrix Computations," Johns Hopkins, Baltimore, 1983, and references therein.
22. J. K. CULLUM AND R. A. WILLOUGHBY, "Lanczos Algorithms for Large Symmetric Eigenvalue Computations," Vol. I, Birkhauser, Boston, 1985, and references therein.
23. W. B. GRAGG AND L. REICHEL, "IBM Bergen Scientific Center," Technical Report 87/17, November 1987.