NORTH-HOLLAND

An Efficient Lower Bound for the Generalized Spectral Radius of a Set of Matrices

Mohsen Maesumi*

Department of Mathematics
Lamar University
P.O. Box 10047
Beaumont, Texas 77710

Submitted by Richard A. Brualdi

ABSTRACT

The generalized spectral radius (GSR) is a fundamental concept in studying the regularity of compactly supported wavelets. Here we describe an efficient method for estimating a lower bound for the GSR. Let $\mathcal{M}_q$ be the set of all $q \times q$ matrices with complex entries. Suppose $\Sigma = \{A_0, \ldots, A_{m-1}\}$ is a collection of m matrices in $\mathcal{M}_q$. Let $\mathcal{L}_n$ be the set of all products of length n of the elements of Σ. Define $\rho_n(\Sigma) = \max_{A \in \mathcal{L}_n} [\rho(A)]^{1/n}$, where $\rho(A)$ is the spectral radius of A. The generalized spectral radius of Σ is then $\rho(\Sigma) = \limsup_{n\to\infty} \rho_n(\Sigma)$. The standard method for estimating $\rho(\Sigma)$, from below and at level n, is to calculate the spectral radii of all m^n products in $\mathcal{L}_n$ and select the largest. Here we use three elementary theorems from linear algebra, combinatorics, and number theory to show that the same result can be obtained with no more than m^n/n matrix calculations.

INTRODUCTION

The generalized spectral radius (GSR) is a fundamental concept in studying the regularity of compactly supported wavelets. Such wavelets can be constructed via infinite products of a special set of matrices. The GSR is the main measuring tool in determining the maximal growth rate of such products. In this respect, it generalizes the usual notion of spectral radius of a single matrix to a finite set of matrices. A similar notion, that of joint spectral radius (JSR), was first defined by Rota and Strang [1] and was shown to be equivalent to a third definition. Later, Daubechies and

*E-mail: maesumi@math.lamar.edu.

LINEAR ALGEBRA AND ITS APPLICATIONS 240:1–7 (1996)

0024-3795/96/$15.00
655 Avenue of the Americas, New York, NY 10010 SSDI 0024-3795(94)00171-8

Lagarias [2, 3] defined a generalized spectral radius (GSR) and conjectured that the two notions are equivalent. The conjecture was proved by Berger and Wang [4]. Hence for finite sets of matrices there is a well-defined notion of spectral radius. The finiteness conjecture, the statement that $\rho(\Sigma) = \rho_n(\Sigma)$ for some n, is investigated in [5]. Collela and Heil describe exact and approximate results for several special classes of matrices in [6]. Heil and Strang prove the continuity of the GSR as a function of matrix entries in [7].

The direct estimation of the GSR has an exponentially increasing cost. The branch-and-bound method [2, 3], originally defined for GSR, reduces the cost for estimating an upper bound. Gripenberg [8] has modified this method to include an estimate for a lower bound and improve efficiency. Here we describe an efficient method for estimating a lower bound for the GSR. This method can be combined with [8] to provide further speed up.

Let $\mathcal{M}_q^k$ be the set of k-tuples where all components are $q \times q$ matrices with complex entries. In this paper all matrices belong to an appropriate $\mathcal{M}_q$. Denote the spectral radius of a matrix B by $\rho(B)$, and define the average spectral radius of an element of $\mathcal{M}_q^k$ by $\overline{\rho}(B_1, B_2, \ldots, B_k) = [\rho(B_1 B_2 \cdots B_k)]^{1/k}$. Suppose $\Sigma = \{A_0, \ldots, A_{m-1}\}$ is a collection of m matrices in $\mathcal{M}_q$. Let $\mathcal{L}_n$ be the set of all products (or words) of length n of the elements of Σ. Define $\rho_n(\Sigma) = \max_{A \in \mathcal{L}_n} [\rho(A)]^{1/n} = \max_{A \in \mathcal{L}_n} \overline{\rho}(A)$. The generalized spectral radius of Σ is then $\rho(\Sigma) = \limsup_{n \to \infty} \rho_n(\Sigma)$. The standard method for estimating $\rho(\Sigma)$, from below and at level n, is to calculate the spectral radii of all m^n products in $\mathcal{L}_n$ and select the largest. Here we use three theorems from linear algebra, combinatorics, and number theory to show that the same result can be obtained with no more than m^n/n matrix calculations.

A brief outline of this paper is as follows. An elementary theorem of linear algebra states that the characteristic polynomial of a product is invariant under the cyclic permutation of the elements of the product. The same applies to the spectral radius; e.g., $\rho(ABC) = \rho(BCA) = \rho(CAB)$. This gives the major savings in matrix calculations for the GSR. The number of cyclically different products (or words) in $\mathcal{L}_n$ is given by a classical combinatorical result due to MacMahon where the leading term is m^n/n. Further savings can be achieved by eliminating from calculations the products which are a "full power" (of a shorter product). For example, $ABAB = (AB)^2$; hence $\overline{\rho}(ABAB) = \overline{\rho}(AB)$, and we need not calculate the spectral radius for the product $ABAB$ if it has been calculated for AB. We show that the number of cyclically different products which are not a full power is given by the Dedekind-Liouville transform of MacMahon's formula. After this elimination, the number of matrix calculations at level n will be no more than m^n/n.

1. OPTIMIZING THE LOWER BOUND

The main saving in the calculation of the GSR comes from the following theorem [9]. (A completely algebraic proof of this theorem for the more general case of rectangular matrices is given in [9, Theorem 1.3.20].)

THEOREM 1. *Let A and B be square matrices of the same size. Then the characteristic polynomials of AB and BA are equal.*

Proof. Let $f(A, \lambda) = \det(A - \lambda)$ denote the characteristic polynomial of the matrix A. Assume B is an invertible matrix. Then we have

$$\begin{aligned} f(AB, \lambda) &= \det(AB - \lambda) = \det[(A - \lambda B^{-1})B] \\ &= \det[B(A - \lambda B^{-1})] = \det(BA - \lambda) = f(BA, \lambda), \end{aligned}$$

which proves the theorem. If B is not invertible, set $B_\epsilon = B - \epsilon$, where ϵ is a number different from any of the eigenvalues of B. Then B_ϵ is invertible; therefore $f(AB_\epsilon, \lambda) = f(B_\epsilon A, \lambda)$. In the limit as $\epsilon \to 0$ we obtain the statement of the theorem, since the determinant is a continuous function of the matrix entries. ■

COROLLARY 1. *The characteristic polynomial, and hence all the eigenvalues and the spectral radius of a product of matrices, are invariant under cyclic permutation of product entries. That is,*

$$f(C_1 C_2 \cdots C_i C_{i+1} \cdots C_n, \lambda) = f(C_{i+1} \cdots C_n C_1 C_2 \cdots C_i, \lambda)$$

and

$$\rho(C_1 C_2 \cdots C_i C_{i+1} \cdots C_n) = \rho(C_{i+1} \cdots C_n C_1 C_2 \cdots C_i).$$

Proof. This is immediate if we set $A = C_1 C_2 \cdots C_i$ and $B = C_{i+1} \cdots C_n$ in Theorem 1. ■

In order to count the matrix calculations we need the following theorem [10, 11].

THEOREM 2 (MacMahon). *The number of cyclically different words of length n from a set of m characters is*

$$M(n, m) = \frac{1}{n} \sum_{d \backslash n} m^d \phi\left(\frac{n}{d}\right),$$

where ϕ is the Euler totient function. ($\phi(l)$ is the number of integers in $\{0, \ldots, l-1\}$ which are relatively prime to l. $\phi(1) = 1$, and if $l = \Pi p_i^{\alpha_i}$

is a prime factorization of l, then $\phi(l) = \Pi p_i^{\alpha_i - 1}(p_i - 1)$. We use $d\backslash n$ to indicate (summation over) positive divisors, that is, $n = kd$ where k is an integer and d is a positive integer.)

For example if there are two characters, A_0 and A_1, then the number of cyclically different words (or products) of length 6 is

$$M(6,2) = \frac{1}{6}\sum_{d\backslash 6} 2^d \phi\left(\frac{6}{d}\right) = \frac{1}{6}[2^1\phi(6) + 2^2\phi(3) + 2^3\phi(2) + 2^6\phi(1)] = 14.$$

These 14 products are

$$A_0^6, A_0^5A_1, A_0^4A_1^2, A_0^3A_1^3, A_0^3A_1A_0A_1, A_0^2A_1^2A_0A_1, \left(A_0^2A_1\right)^2, (A_0A_1)^3,$$

$$A_1^6, A_1^5A_0, A_1^4A_0^2, \underline{A_1^3A_0^3}, A_1^3A_0A_1A_0, A_1^2A_0^2A_1A_0, \left(A_1^2A_0\right)^2, \underline{(A_1A_0)^3}$$

where the underlined items are to be deleted, since they are cyclically equivalent to the products that are printed above them.

In practice one calculates the average spectral radius for all the products of length 1 through some n. In that case one can eliminate products which are a power of a smaller product from consideration. For example, from the above 14 products 5 are powers of products of shorter length. They are

$$A_0^6, \left(A_0^2A_1\right)^2, (A_0A_1)^3, A_1^6, \left(A_1^2A_0\right)^2.$$

Therefore, of all the $2^6 = 64$ products of length 6, only $14 - 5 = 9$ have average spectral radius distinct from each other and from the products of shorter length. Notice that the savings rate is somewhat better than the length of the product, e.g. $9 < 2^6/6$. In the rest of this paper we generalize the above example.

Assume $N(n,m)$ is the number of cyclically different products of length n which are not a power of a product of shorter length. Then we have

$$N(n,m) = M(n,m) - \sum_{\substack{d\backslash n \\ d<n}} N(d,m). \tag{1}$$

To see this let A be a product of length n. Suppose A can be written as a power of a product B of a shorter length d, and B itself is not a power of any product of a shorter length. Then necessarily $d\backslash n$, $d < n$. The number of matrices which share the same properties with B is $N(d,m)$, and each of these matrices can be used to eliminate exactly one entry from the list of products of length n.

We can write the above formula as

$$M(n,m)=\sum_{d\backslash n} N(d,m). \tag{2}$$

Now we can use the Dedekind-Liouville inversion principle [10, 12, 13] to find N in terms of M.

THEOREM 3 (Dedekind-Liouville). *We have*

$$g(n)=\sum_{d\backslash n} h(d) \iff h(n)=\sum_{d\backslash n}\mu(d)g\left(\frac{n}{d}\right),$$

where μ is the Möbius function. ($\mu(1)=1$, and if $l=\prod_{i=1}^{k}p_i^{\alpha_i}$ is prime factorization of l, then $\mu(l)=(-1)^k$ if all $\alpha_i=1$; otherwise $\mu(l)=0$.)

We can use the above theorem to deduce

$$N(n,m)=\sum_{d\backslash n}\mu(d)M\left(\frac{n}{d},m\right). \tag{3}$$

This formula gives the required number of calculations at level n. We can simplify this formula considerably if we use the following multi-index notation. Let $x=(x_1,x_2,\ldots,x_k)$, $\beta=(\beta_1,\beta_2,\ldots,\beta_k)$, and define $x^\beta=\prod_{i=1}^{k}x_i^{\beta_i}$. For a given polynomial $H(x)=\sum_\beta a_\beta x^\beta$ define "term by term exponentiation" in base m as

$$E(m,H)(x)=\sum_{\beta} a_\beta m^{x^\beta}.$$

Suppose $n>1$ is given. Let the prime factorization of n be written as $n=\prod_{i=1}^{k}p_i^{\alpha_i}$, and set $p=(p_1,p_2,\ldots,p_k)$. Furthermore, define a polynomial $Q(p)$ by

$$Q=Q(p)=\phi(n)=\prod_{i=1}^{k}p_i^{\alpha_i-1}(p_i-1).$$

Now we can give a brief formula for the Dedekind-Liouville transform of MacMahon's formula.

THEOREM 4. *Suppose Q is defined as above then the number of matrix products of length n that have average spectral radius different from each other and from products of shorter length is given by*

$$N(n,m)=\frac{1}{n}E(m,Q)(p). \tag{4}$$

Proof. This can shown by applying MacMahon's formula and using induction on (1). ∎

For example if $n = p_1^2 p_2$ then $Q(p) = p_1(p_1-1)(p_2-1) = p_1^2p_2 - p_1^2 - p_1p_2 + p_1$ and

$$N(n,m) = \frac{1}{n}\left(m^{p_1^2p_2} - m^{p_1^2} - m^{p_1p_2} + m^{p_1}\right).$$

Similarly, if $n = p_1p_2$ then $Q(p) = (p_1-1)(p_2-1) = p_1p_2 - p_1 - p_2 + 1$ and

$$N(n,m) = \frac{1}{n}(m^{p_1p_2} - m^{p_1} - m^{p_2} + m).$$

The latter example is applicable to $n = 6$ and $m = 2$. We write $6 = 2 \times 3$; therefore $N(6,2) = \frac{1}{6}(2^6 - 2^3 - 2^2 + 2) = 9$, as was seen before.

In the next theorem we give an upper bound for N.

THEOREM 5. *We have* $N(1,m) = M(1,m) = m$ *and*

$$N(n,m) < \frac{m^n}{n} \qquad \text{for} \quad n > 1. \tag{5}$$

Proof. This is a simple consequence of the next, more general theorem. ∎

THEOREM 6. *Let* $H(x) = \sum_\beta a_\beta x^\beta$, *and assume* $H(x^k) \geq 0$ *for all* $k \geq 0$ *and all* x *in a certain set. Then for any* $m \geq 1$ *we have* $E(m,H)(x) \geq 0$ *for all* x *in the same set.*

Proof. We have

$$\begin{aligned} E(m,H)(x) = \sum_\beta a_\beta m^{x^\beta} &= \sum_{k=0}^{\infty} \frac{\ln^k(m)}{k!} \sum_\beta a_\beta x^{\beta k} \\ &= \sum_{k=0}^{\infty} \frac{\ln^k(m)}{k!} H(x^k) \geq 0. \end{aligned}$$

Which proves the theorem.

In order to prove Theorem 5 we only need to notice that for $n > 1$

$$Q = \prod_{i=1}^{k} p_i^{\alpha_i - 1}(p_i - 1) < \prod_{i=1}^{k} p_i^{\alpha_i} = n.$$

Moreover, this inequality remains valid when p_i is replaced by p_i^k. Now, upon term by term exponentiation of the two sides of the inequality, we obtain the result of Theorem 5. ∎

REFERENCES

1 G. C. Rota and G. Strang, A note on the joint spectral radius, *Kon. Nederl. Akad. Wet. Proc. A* 63:379–381 (1960).
2 I. Daubechies and J. Lagarias, Two-scale difference equations, II. Infinite matrix products, local regularity bounds and fractals, *SIAM J. Math. Anal.* 23:1031–1079 (1992).
3 I. Daubechies and J. Lagarias, Sets of matrices all infinite products of which converge, *Linear Algebra Appl.* 161:227–263 (1992).
4 M. A. Berger and Y. Wang, Bounded semi-groups of matrices, *Linear Algebra Appl.* 166:21–27 (1992).
5 J. C. Lagarias and Y. Wang, The finiteness conjecture for the generalized spectral radius of a set of matrices, *Linear Algebra Appl.*, to appear.
6 D. Collela and C. Heil, The characterization of continuous, four-coefficient scaling functions and wavelets, *IEEE Trans. Inform. Theory* 38:876–881 (1992).
7 C. Heil and G. Strang, Continuity of the joint spectral radius: Application to wavelets, in *Linear Algebra for Signal Processing* (A. Bojanczyk and G. Cybenko, Eds.), IMA Vol. Math. Appl., Springer-Verlag, New York.
8 G. Gripenberg, Computing the joint spectral radius, preprint.
9 R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge U.P., 1988.
10 R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics*, Addison-Wesley, 1994.
11 P. A. MacMahon, Application of a theory of permutations in circular procession to the theory of numbers, *Proc. London Math. Soc.* 23:305–313 (1892).
12 R. Dedekind, Abriß einer Theorie der höheren Congruenzen in Bezug auf einen reellen Primzahl-Modulus, *J. Reine Angew. Math.* 54:1–26 (1857). *Gesammelte Mathematische Werke*, Vol. 1, pp. 40–67.
13 J. Liouville, Sur l'expression $\phi(n)$, qui marque combien la suite $1, 2, 3, \ldots, n$ contient de nombres premiers à n, *J. Math. Pures Appl. Ser. 2* 2:110–112 (1857).

Received 17 June 1994; final manuscript accepted 18 July 1994