Optimum Unit Ball for Joint Spectral Radius: An Example from Four–Coefficient MRA

Mohsen Maesumi

Abstract. We give the exact value of joint spectral radius for certain matrices by explicitly constructing an optimum unit ball and an operator norm. Our example matrices are associated with 4–coefficient dilation equations which generate a multiresolution analysis. The method proposed here can be used in the search for orthogonal solution of the dilation equation with the highest Hölder exponent. It can also be used to decide if a given product of matrices is optimal, *i.e.*, it satisfies the Finiteness Conjecture. The optimal ball is generated as the convex hull of action of semigroup of matrices, normalized by their joint spectral radius, on extreme points of a set of unit diameter which is maximal, convex, symmetric, compact and invariant under the optimal product.

§1. Introduction

Joint spectral radius (JSR) of a family of matrices is the measure of maximal growth rate of long products of the family members. This concept has a natural application in regularity analysis of compactly supported wavelets, since the recursive algorithm also utilizes long products of a given set of matrices. One can obtain the Hölder exponent of such wavelets if JSR of an associated pair of matrices can be estimated. Calculation of JSR is a difficult task. Obtaining estimates is expensive and exact values, except for special cases, are rarely known. A recent paper [15] attempts to show that the general case is algorithmically unsolvable.

JSR is bounded from below by the normalized spectral radius of any product and from above by the maximum normalized consistent norm of products of length n for any n. In particular it is bounded from above by the maximum operator norm of all matrices in the family.

Approximation Theory VIII, Vol. 2: Wavelets and Multilevel Approximation
Charles K. Chui and Larry L. Schumaker (eds.), pp. 267–274.

ISBN 981-02-2972-0

The Finiteness Conjecture [6, 11] claims that there is a finite product whose normalized spectral radius is equal to JSR. If the attempt in [15] succeeds then this conjecture may not be true in all cases. However, in this report we demonstrate a simple method that can be used to verify if a given product indeed satisfies the conjecture.

Suppose a particular product is believed to satisfy the Finiteness Conjecture. Then we generate an operator norm in which the maximum norm of matrices equals the normalized spectral radius of the product (if the claim is true), thereby confirming the guess.

Our particular example comes from the study of four-coefficient dilation equations and the resulting multiresolution analyses (MRA). An issue of interest is calculation of the Hölder exponent of corresponding wavelets and identification of the smoothest one. Here we give the exact value of JSR for the matrices associated with a particular wavelet studied in [3] and confirm a conjecture by Colella and Heil [7] regarding its value. We also show that the corresponding MRA is not the smoothest one.

§2. Definitions and Preliminaries

Let $\mathcal{M}(q,\mathbb{C})$ be the set of all $q \times q$ matrices with complex entries. Denote operator, consistent and vector norms by $||\cdot||_o$, $||\cdot||_c$ and $||\cdot||_v$ respectively. Consistent (or matrix) norms are submultiplicative, $||AB||_c \leq ||A||_c||B||_c$. Operator (or induced or lub) norms are defined by $||A||_o = \max_{||x||_v=1} ||Ax||_v$, $x \in \mathbb{C}^q$. Operator norms are consistent but consistent norms are not necessarily induced. $||A||_v$ can be defined by viewing the matrix as a vector in $\mathbb{C}^{q^2}$. All norms on a finite dimensional space are equivalent, *i.e.*, given any two norms $||\cdot||$ and $||\cdot||'$ there are $b > a > 0$ such that $b||A||' \geq ||A|| \geq a||A||'$ for any A.

Suppose Σ is a collection of m matrices in $\mathcal{M}(q,\mathbb{C})$. Let $\mathcal{L}_n = \mathcal{L}_n(\Sigma)$ be the set of all m^n products of length n of the elements of Σ. Rota and Strang [14] defined *joint spectral radius* (JSR) as $\hat{\rho}(\Sigma) = \limsup_{n\to\infty} \hat{\rho}_n(\Sigma, ||\cdot||_v)$, where $\hat{\rho}_n(\Sigma, ||\cdot||_v)$, the maximum normalized norm of products of length n of Σ, is $\max_{A\in\mathcal{L}_n} ||A||_v^{1/n}$. Daubechies and Lagarias [6] defined *generalized spectral radius* (GSR) as $\check{\rho}(\Sigma) = \limsup_{n\to\infty} \check{\rho}_n(\Sigma)$, where $\check{\rho}_n(\Sigma) = \max_{A\in\mathcal{L}_n} \bar{\rho}(A)$, and $\bar{\rho}$, the normalized spectral radius for a product of length n, is $(\rho(A))^{1/n}$. Rota and Strang also gave another definition, which we refer to as *common spectral radius* (CSR), by $\tilde{\rho}(\Sigma) = \inf_{||\cdot||_c} \max_{A\in\Sigma} ||A||_c$, where the infimum is over all consistent norms. All of the various concepts of radius mentioned above, with the exception of $\hat{\rho}_n$, are invariant under similarity transformations.

In order to use the definition of CSR for computation it is advantageous to restrict the space of norms over which the optimization is performed.

The following lemmas show that we may take the infimum over the operator norms acting on a certain subspace of $\mathbb{C}^q$ instead of all consistent norms.

Lemma 1. *For every consistent norm $||\cdot||_c$ there is an operator norm $||\cdot||_o$ such that $||A||_o \le ||A||_c$ for all A. Therefore $\tilde{\rho}(\Sigma) = \inf_{||\cdot||_o} \max_{A\in\Sigma} ||A||_o$.*

Proof: Given a vector x define a matrix X whose columns are identical with x. Define $||x||_v = ||X||_c$ then the induced norm satisfies $||A||_o = \max_{x\neq 0} \frac{||Ax||_v}{||x||_v} = \max_{x\neq 0} \frac{||AX||_c}{||X||_c} \le ||A||_c$ as required. ■

Occasionally it happens that all elements of Σ have the same lower block triangular structure, perhaps after a similarity transformation. In that case the calculation of CSR may be segmented accordingly as explained below. Notice that if B is a $p \times p$ complex matrix then it can be mapped into a $2p \times 2p$ real matrix. In particular if B is real then the result of mapping can be permuted into a block diagonal matrix with two identical blocks. In any of above cases we apply the following lemma.

Lemma 2. *If each $A \in \Sigma$ has the same lower triangular block structure with diagonal blocks A_i, $i = 1\cdots k$, and $\Sigma_i = \{A_i, A \in \Sigma\}$ then $\tilde{\rho}(\Sigma) = \max_i \tilde{\rho}(\Sigma_i)$ where $\tilde{\rho}(\Sigma_i) = \inf_{||\cdot||_o} \max_{A_i\in\Sigma_i} ||A_i||_o$.*

Proof: This is essentially same as block triangularization lemmas of the same nature that are used for JSR or GSR [1]. The result follows from equivalence of CSR, JSR and GSR mentioned below. ■

Corollary 1. *If Σ is real then $\tilde{\rho}(\Sigma) = \inf_{||\cdot||_o} \max_{A\in\Sigma} ||A||_o$ where the operator norm is induced from a real vector norm $||A||_o = \max_{||x||=1, x\in\mathbb{R}^q} ||Ax||$.*

Proof: We map each $A \in \Sigma$ into a $2q \times 2q$ block diagonal matrix with identical blocks and apply Lemma 2. ■

Several papers have scrutinized the relationship between various approaches to spectral radius of a set of matrices. Here we describe some of the findings. The first theorem in this category was stated by Rota and Strang [14]. It asserted that JSR and CSR are equal, $\hat{\rho}(\Sigma) = \tilde{\rho}(\Sigma)$. The main step in the proof of the above theorem was the following. Suppose a set Σ and a consistent norm $||\cdot||$ are given. Then the necessary and sufficient condition that there exists a consistent norm $||\cdot||'$ such that $||A||' \le 1$ for all $A \in \Sigma$ is that Σ is product bounded, *i.e.*, there is a K such that $||A|| \le K$ for all n and $A \in \mathcal{L}_n(\Sigma)$.

Daubechies and Lagarias [6] showed $\check{\rho}_n(\Sigma) \le \check{\rho}(\Sigma) \le \hat{\rho}(\Sigma) \le \hat{\rho}_n(\Sigma, ||\cdot||_c)$. Therefore, if consistent norms are used in the definition of JSR then limsup may be replaced by lim or inf, as also stated in [14]. Similarly, in the definition of GSR limsup may be replaced by sup. They also conjectured that $\check{\rho}(\Sigma) = \hat{\rho}(\Sigma)$, and the conjecture was proved by Berger and Wang [1]. Hence the three definitions of spectral radius of a finite set of matrices agree. In fact the agreement exists for infinite but bounded sets as well. So we may talk of spectral radius of a set and denote it by $\rho(\Sigma) = \tilde{\rho}(\Sigma) = \hat{\rho}(\Sigma) = \check{\rho}(\Sigma)$. Heil and Strang [8] showed that $\rho(\Sigma)$ is a continuous function of Σ.

The Finiteness Conjecture [6, 11] states that for some finite n we have $\rho(\Sigma) = \check{\rho}_n(\Sigma)$. In the light of [15] the conjecture may not be true for all cases. However no counterexample has been given. Calculating spectral radius of a set of matrices by direct calculation of $\hat{\rho}_n(\Sigma)$ and $\check{\rho}_n(\Sigma)$ is extremely inefficient. The branch-and-bound method of Daubechies and Lagarias [5] significantly reduces the cost of upper estimates. A refinement in this method, including considerable savings for estimating lower bounds, has been proposed by Gripenberg [10]. Some savings can be realized by noticing the invariance of $\bar{\rho}(A)$, for $A \in \mathcal{L}_n(\Sigma)$, under exponentiation and cyclic permutation of elements of the product A [13].

§3. Generating the Optimum Unit Ball

In this section we show how to investigate the possibility that a product is optimal, *i.e.*, it satisfies the Finiteness Conjecture. First, we start with an example extensively studied by Colella and Heil [3]. Consider the two–scale real dilation equation in four coefficients

$$\phi(x) = c_0\phi(2x) + c_1\phi(2x-1) + c_2\phi(2x-2) + c_3\phi(2x-3). \qquad (1)$$

The pair of wavelet matrices whose infinite products produce $\phi(x)$ are

$$T_0 = \begin{pmatrix} c_0 & 0 & 0 \\ c_2 & c_1 & c_0 \\ 0 & c_3 & c_2 \end{pmatrix}, \qquad T_1 = \begin{pmatrix} c_1 & c_0 & 0 \\ c_3 & c_2 & c_1 \\ 0 & 0 & c_3 \end{pmatrix}. \qquad (2)$$

We assume $c_0 + c_2 = c_1 + c_3 = 1$, then the Hölder exponent of ϕ can be determined from two matrices obtained by restricting T_0 and T_1 to the space normal to the common left eigenvector $(1,1,1)$. These matrices are

$$S_0 = \begin{pmatrix} c_0 & 0 \\ -c_3 & 1-c_0-c_3 \end{pmatrix}, \qquad S_1 = \begin{pmatrix} 1-c_0-c_3 & -c_0 \\ 0 & c_3 \end{pmatrix}. \qquad (3)$$

Let $\Sigma = \{S_0, S_1\}$ and $\rho(\Sigma) < 1$ then ϕ is Hölder continuous with exponent $h \geq -\log_2 \rho(\Sigma) - \epsilon$ for any $\epsilon > 0$ [5]. To characterize the solutions that give a multiresolution [2, 3, 12] we restrict the coefficients by

$$(c_0, c_3) \neq (1,1), \qquad (4.a)$$

$$(c_0 - 1/2)^2 + (c_3 - 1/2)^2 = 1/2. \qquad (4.b)$$

The behavior of ϕ at the particular point $c^\star = (c_0^\star, c_3^\star) = (0.6, -0.2)$ has been scrutinized with the expectation it leads to the smoothest orthogonal scaling function [3] (note that $\check{\rho}_1(\Sigma)$ achieves its minimum at $c^\star$). Let $\Sigma^\star = \{S_0^\star, S_1^\star\}$ denote this particular value of Σ, where

$$S_0^\star = \begin{pmatrix} 0.6 & 0 \\ 0.2 & 0.6 \end{pmatrix}, \qquad S_1^\star = \begin{pmatrix} 0.6 & -0.6 \\ 0 & -0.2 \end{pmatrix}. \tag{5}$$

Colella and Heil carried extensive computations and conjectured [7]

$$\rho^\star = \rho(\Sigma^\star) = \max_{1 \le n \le 30} \check{\rho}_n(\Sigma^\star) = \rho(S_1^\star S_0^{\star 12})^{1/13} \approx 0.659679. \tag{6}$$

Gripenberg's method produced the same result for products of length up to 243. Here we confirm this conjecture and show that the corresponding scaling function or multiresolution analysis are not the smoothest ones.

Proposition 1. *Let $A = S_0^\star/\rho^\star$ and $B = S_1^\star/\rho^\star$. Then $\rho(\{A, B\}) = 1$. There is a neighborhood of $\Sigma^\star$ where $\rho(\Sigma) = \rho(S_1 S_0^{12})$, $\rho(\Sigma)$ is a strictly decreasing function of c_3 and the Hölder exponent of ϕ is a strictly increasing function of c_3.*

Proof: Let $P = BA^{12}$. Note P has an eigenvalue -1, denote the corresponding eigenvector by v. Define a polygonal unit ball $\mathcal{U}$ with 30 sides whose vertices, in counterclockwise direction, are labeled as $v_1, \cdots, v_{15}$ and $v_{-1}, \cdots, v_{-15}$, where $v_{-i} = -v_i$, $v_i = A^{i-1}v$ for $i = 1, \cdots, 14$, and $v_{15} = BA^{13}v$. One verifies that $\mathcal{U}$ is convex hence we can define a norm $||\cdot||_u$ based on it. Obviously, Av_i is a vertex of $\mathcal{U}$ for $i = 1, \cdots, 13$ and Bv_i is a vertex of $\mathcal{U}$ for $i = 13, 14$. One also verifies Av_i is in the interior of $\mathcal{U}$ for $i = 14, 15$ and Bv_i is in the interior of $\mathcal{U}$ for $i = 1, \cdots, 12, 15$. Therefore $\hat{\rho}_1(\{A, B\}, ||\cdot||_u) = 1$. On the other hand $\check{\rho}_{13}(\{A, B\}) \ge \rho(BA^{12})^{1/13} = 1$. We have $\check{\rho}_{13} \le \rho(\{A, B\}) \le \hat{\rho}_1$ therefore $\rho(\{A, B\}) = 1$. The convexity of the ball can be indicated by a system of inequalities of the form $F(V) \le 0$ where V is the vector of vertices of the ball, F is a vector of continuous functions and the inequality is component-wise. We say the ball has *slack* if the vertices satisfy $F(V) < 0$. If any component of $F(V)$ is zero then we say the ball is *critical.* A ball becomes critical for example if two adjacent sides are parallel or if two vertices coincide. Here $\mathcal{U}$ has slack and the entire construction of the ball remains stable under small changes in say c_3. In particular for $c_3 \in [-.2, -.2 + 10^{-5}]$ we may obtain the spectral radius of Σ through $\rho(S_1 S_0^{12})^{1/13}$. One verifies that $\rho(\Sigma)$ is a decreasing function of c_3 in the vicinity of $c^\star$ and hence Hölder exponent is an increasing function. ■

Remark. If we use any of the 13 cyclical permutations of BA^{12}, we will obtain the same unit ball (up to a scale). Given $0 \le n \le 12$, let

$P' = A^{12-n}BA^n$ and $P'v' = -v'$. Then the vertices are given by $v'_i = A^{i-1}v'$ for $i = 1, \cdots, n+2$, $v'_{n+3} = BA^{n+1}v'$, and $v'_i = A^{i-n-4}BA^n v'$ for $i = n+4, \cdots, 15$. Here v'_{n+4} corresponds to v_1 in Proposition 1.

The approach used for the example problem mentioned above can be generalized. It can also be the basis of a method for estimating the spectral radius of a set of matrices. Here we list the notations that will be used. Let Σ^+ denote the set Σ augmented with identity, $\mathcal{L}_* = \mathcal{L}_*(\Sigma)$ the set of all products of Σ and $\mathcal{L}_*^+$ the generated semigroup. The action of a family of matrices $\mathcal{F}$ on a set of vectors W is the collection of vectors $\mathcal{F}W$. The Hermitian adjoint of P is shown by P^*. For a set of points D let $\mathcal{C}(D)$ indicate the convex hull of D. The closure of $\mathcal{C}(D)$ is shown by $\bar{\mathcal{C}}(D)$. Suppose x, y, z belong to a closed convex set and $0 < \alpha < 1$, if $z = \alpha x + (1-\alpha)y$ implies $x = y = z$ then z is called an *extreme* point of the set [9]. A compact set is the convex hull of its extreme points (Krein-Milman theorem). The vertices of a nondegenerate polyhedron are its extreme points. When a vertex becomes degenerate, *e.g.*, when all of its adjacent sides become parallel, then it is no longer an extreme point. But degenerate vertices are important for us since they indicate a ball has become critical.

Suppose $\rho(\Sigma) = 1$ then $P \in \mathcal{L}_*(\Sigma)$ is called an optimal product if $\rho(P) = 1$. We assume Σ is product bounded then $\{P^\nu, \nu = 1, 2, \cdots\}$ is bounded, each eigenvalue satisfies $|\lambda| \le 1$ and P has a full set of eigenvectors for eigenvalues with modulus one. We call $\mathcal{G}$ the *generator* for P if it is the maximal set with a positive diameter which is convex, symmetric, compact and invariant under P. In other words $\mathcal{G}$ is a ball (of a subspace with maximal dimension) for which P is an isometry. The generator of P may have polyhedral components $\{\pm P^m v, m = 1, \cdots, n\}$ corresponding to solutions of $P^n v = \pm v$ for some finite n, or ellipsoidal components $\{x, x^* S x = 1\}$ corresponding to positive definite matrices S as solutions to $P^* S P = S$. (The eigenvalues of P with modulus one determine the ergodicity of P and can be used to subdivide various cases. If Σ has block diagonal structure then one constructs a generator for each block.)

Suppose Σ is product bounded and $\mathcal{U}$ is a ball then $\mathcal{U}' = \bar{\mathcal{C}}(\mathcal{L}_*^+ \mathcal{U})$ is also a ball. Let $||\cdot||'$ be the norm in which $\mathcal{U}'$ is a unit ball, then $\max\{||A||', A \in \Sigma\} \le 1$. This is the *alternative construction for the norms* given in [14]. Here, however, we do not start with an arbitrary ball, instead we identify the generator set $\mathcal{G}$ of P and create the optimal ball $\mathcal{S} = \bar{\mathcal{C}}(\mathcal{L}_*^+ \mathcal{G})$ from it. The advantage of using $\mathcal{G}$ is that its extreme points generate the vertices (usually extreme points) of the optimum ball. In contrast, an arbitrary starting point will mark a set of nondescript points on the boundary of the optimal ball. (If $\mathcal{S}$ becomes a low dimensional ball of a subspace embedded in the current space then we restrict Σ to the complementary subspace and repeat the construction to get a subinvariant ball.)

Suppose $v \neq 0$, $\mathcal{F}_0 = \mathcal{C}(\{\pm v\})$ and $\mathcal{F}_n = \mathcal{C}(\Sigma^+ \mathcal{F}_{n-1})$ for $n \geq 1$. Let $V_0 = \{\pm v\}$ and for $n \geq 1$ define V_n, the *vertices* of $\mathcal{F}_n$, to be the points of $\Sigma^+ V_{n-1}$ which are on the boundary of $\mathcal{F}_n$.

Conjecture 1. *If there are $v' \in V_n$ and $m > n$ such that v' is in the interior of $\mathcal{F}_m$ then $\rho(\Sigma) > 1$.*

This conjecture implies that none of the vertices at a given stage of iteration will be overtaken by the next set of vertices. Consequently, if this occurs we do not have an optimal product. One case of this conjecture is simple, namely when v' is v itself. This is the subject of the next lemma.

Definition 1. Given a vector $v \neq 0$, a compact set of vectors V and a finite set of matrices Σ, we say v is *dominated by Σ acting on V* if there are $\alpha > 1$ and $w \in \mathcal{C}(\Sigma V)$ such that $v = w/\alpha$.

Lemma 3. *If v is dominated by Π, a finite subset of $\mathcal{L}_*(\Sigma)$, acting on $\{v\}$ then $\rho(\Sigma) > 1$.*

Proof: First assume $\Pi = \Sigma$ and consider any norm in which $||v|| = 1$. Then for $w' \in \mathcal{C}(\Sigma v)$ we have $\max ||w'|| \geq \alpha$. On a compact convex set the maximum of a convex function is attained at an extremum point. Norm function is convex hence $\max ||w'||$ occurs at $||Av||$ for an $A \in \Sigma$. Hence $\max ||A|| \geq \alpha > 1$. Since this occurs for any norm, by the definition of CSR, we have $\rho(\Sigma) > 1$. For general Π we have $\mathcal{L}_*(\Pi) \subset \mathcal{L}_*(\Sigma)$ hence $\rho(\Sigma) \geq \rho(\Pi) > 1$. ■

Suppose numerical evidence suggests that Σ is product bounded, $\rho(\Sigma) = 1$ and P is the optimal product. To verify the guess, one starts with $\mathcal{G}_0$, the generator of P, and constructs $\mathcal{G}_n = \mathcal{C}(\Sigma^+ \mathcal{G}_{n-1})$ for $n \geq 1$. If this process terminates, *i.e.*, $\mathcal{G}_n = \mathcal{G}_{n-1}$ for some $n \geq 1$, then P is indeed an optimal product. When $\mathcal{G}_0$ is ellipsoidal (with no polyhedral subset, *i.e.*, P is ergodic on the ellipsoid) we need only to check $\mathcal{G}_1 = \mathcal{G}_0$. For the general case we have the following conjecture.

Conjecture 2. *If P is optimal then $\mathcal{G}_n = \mathcal{G}_{n-1}$ for some finite n.*

In this preliminary report we have emphasized the importance of analysis of spectral radius from a geometrical point of view. In a future article we will report on several related issues, *e.g.*, properties of the optimal product, an adaptive branch-and-bound method where the norm changes with each iteration of the ball and a larger table of values of Hölder exponent for the example problem.

Acknowledgments. The author thanks Christopher Heil for pointing out recent references.

References

1. Berger, M. A. and Y. Wang, Bounded semi-groups of matrices, Linear Algebra Appl. **166** (1992), 21–27.
2. Cohen, A., Ondelettes, analyses multirésolutions et filtres miroirs en quadrature, Ann. Inst. H. Poincare **7** (1990), 439–459.
3. Colella, D. and C. Heil, The characterization of continuous, 4-coefficient scaling functions and wavelets, IEEE Trans. Inf. Th. **38** (1992), 876-881.
4. Daubechies I. and J. Lagarias, Two-scale difference equations, I. existence and global regularity of solutions, SIAM J. Math. Anal. **22** (1991), 1388–1410.
5. Daubechies, I. and J. Lagarias, Two-scale difference equations, II. Infinite matrix products, local regularity bounds and fractals, SIAM J. Math. Anal. **23** (1992), 1031–1079.
6. Daubechies, I. and J. Lagarias, Sets of matrices all infinite products of which converge, Linear Algebra Appl. **161** (1992), 227–263.
7. Heil C. and D. Colella, Dilation equations and the smoothness of compactly supported wavelets, in *Wavelets: mathematics and applications*, J. J. Benedetto and M. W. Frazier (eds.), CRC Press, Boca Raton, FL, 1993, 161–200.
8. Heil, C. 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, New York.
9. Horn, R. and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1988.
10. Gripenberg, G., Computing the joint spectral radius, Linear Algebra Appl., to appear.
11. Lagarias J. C. and Y. Wang, The finiteness conjecture for the generalized spectral radius of a set of matrices, Linear Algebra Appl., to appear.
12. Lawton, W., Necessary and sufficient conditions for constructing orthonormal wavelet bases, J. Math. Phys. **32** (1991), 57–61.
13. Maesumi, M., An optimization for calculation of generalized spectral radius, Linear Algebra Appl., to appear.
14. Rota, G. C. and G. Strang, A note on the joint spectral radius, Kon. Nederl. Akad. Wet. Proc. A **63** (1960), 379–381.
15. Toker, O., On the algorithmic unsolvability of asynchronous iterative processes, American Control Conference, 1995, preprint.

Mohsen Maesumi
Lamar University, P. O. Box 10047, Beaumont, TX 77710
maesumi@math.lamar.edu