International Journal of Bifurcation and Chaos, Vol. 22, No. 4 (2012) 1250096 (6 pages)

DOI: 10.1142/S0218127412500964

A SPECTRAL RADIUS ESTIMATE AND ENTROPY OF HYPERCUBES

WILLIAM GELLER*, BRUCE KITCHENS† and MICHAŁ MISIUREWICZ‡
Department of Mathematical Sciences, Indiana University – Purdue University Indianapolis, 402 N. Blackford Street, Indianapolis, IN 46202, USA
**wgeller@math.iupui.edu*
†bkitchens@math.iupui.edu
‡mmisiure@math.iupui.edu

MICHAŁ RAMS
Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8, 00-956 Warszawa, Poland
rams@impan.pl

Received November 11, 2011

We consider a problem in Mathematical Biology that leads to a question in Graph Theory, which can be solved using an old but not widely known upper estimate of the spectral radius of a non-negative matrix. We provide a new proof of this estimate.

Keywords: Non-negative matrices; spectral radius; topological entropy; entropy of graphs; biological dynamics.

1. Introduction

One of the methods of studying systems arising in biology is to replace all values that can be attained by any given variable by just two values: below and above some threshold (see e.g. [Glass & Pasternack, 1978; Glass & Siegelmann, 2010; Perkins *et al.*, 2010; Wilds & Glass, 2009]). Then the evolution in time will be represented by a path on a corresponding graph: each variable can switch from one state to the other, depending on the values of all the remaining variables. If there are k variables, since generically only one variable can switch at any time, the graph that we should consider is the 1-skeleton of the k-dimensional hypercube. Since for each variable, the values of all other variables determine how this one will change, each edge of the graph has its direction. This way we get a digraph.

The complexity of our system (how many paths are possible?) can be measured by the entropy of the graph. Wilds *et al.* [2008] asked the question: what is the maximal entropy of the k-dimensional hypercube over all possible assignments of directions of edges? We learned about this question from Leon Glass. In [Wilds *et al.*, 2008] the authors conjecture, based on numerical experiments, that this maximal entropy is $\log \frac{k}{2}$ if k is even and $\log \frac{\sqrt{k^2-1}}{2}$ if k is odd. Here, we prove that indeed this is the case.

Note that if instead of a hypercube we consider a complete graph, a similar question has been studied extensively, see [Brualdi, 2010] (the digraphs one gets are called *tournaments*).

The entropy of a digraph is equal to the logarithm of the spectral radius of the transition (adjacency) matrix of this graph (or 0 if this spectral radius is 1 or less). Thus, we need to estimate the spectral radius of this matrix. This problem plays an important role in several branches of mathematics; apart from Linear Algebra and Graph Theory, we can list Dynamical Systems, Probability Theory and applications of mathematics. There are many different estimates, see, e.g. [Minc, 1988]. In particular, the following simple estimate is well known. Let M be an $n \times n$ non-negative matrix with row sums R_i, column sums C_i, and spectral radius λ. Then

$$\begin{aligned} &\min\{R_i : i = 1, \ldots, n\} \\ &\qquad \le \lambda \le \max\{R_i : i = 1, \ldots, n\}, \\ &\min\{C_i : i = 1, \ldots, n\} \\ &\qquad \le \lambda \le \max\{C_i : i = 1, \ldots, n\}. \end{aligned} \tag{1}$$

However, there are cases when we would like to get a better estimate, knowing the row and column sums of M. Such an estimate,

$$\lambda \le \max\{\sqrt{R_i C_i} : i = 1, \ldots, n\}, \tag{2}$$

has been obtained by Barankin [1945] and rediscovered by Kwapisz [1996]. It has been strengthened by Ostrowski [1951] to the following one:

$$\lambda \le \max\{R_i^p C_i^{1-p} : 1 \le i \le n\} \tag{3}$$

for every $p \in (0, 1)$. Note that we get only upper estimates; the lower ones similar to (1) do not hold.

Inequalities (2) and (3) remain surprisingly unknown. Usually they do not appear even in the books with many different estimates of the spectral radii of non-negative matrices. The survey article of Brualdi [2010] mentions the result of Kwapisz, but does not of [Barankin, 1945] or [Ostrowski, 1951]. However, it cites a paper of Kolotilina [2002], which cites the paper of Ostrowski (and his result), which cites the paper of Barankin. This way, after we rediscovered (2) and (3), we learned that we were 66 and 60 years late.

Nevertheless, our proof of (3) ((2) is a special case of (3) for $p = 1/2$) is quite different, and simpler than the original proof of Ostrowski. We present it here, as well as the explanation of the ideas from Dynamical Systems that led to this proof. We should note that Kwapisz's proof also uses ideas from Dynamical Systems, but different to those used by us.

The paper is organized as follows. In Sec. 2, we give a simple algebraic proof of (3). In Sec. 3, we explain briefly how we got this proof using ideas from Dynamical Systems. This section is purely explanatory; it is not necessary to read it in order to understand the rest of the paper. In Sec. 4, we translate our results to the language of Graph Theory and derive some interesting consequences. In Sec. 5, we apply the results of the preceding section to the problem in Mathematical Biology that we mentioned. Finally, in Sec. 6, we give two examples. The first one shows that we cannot count on lower estimates of the spectral radius using estimates similar to (3). The second one shows that (3) is stronger than (2).

2. Spectral Radius Estimate

In this section we give a simple proof of the theorem of Ostrowski.

Theorem 1. *Let M be an $n \times n$ non-negative matrix with row sums R_i, column sums C_i, and spectral radius λ. Then for every $p \in (0, 1)$*

$$\lambda \le \max\{R_i^p C_i^{1-p} : 1 \le i \le n\}. \tag{4}$$

Proof. Let $M = (m_{ij})$ and assume first that M is positive. There are positive eigenvectors of M corresponding to the eigenvalue λ: left one (ℓ_i) and right one (r_i). That is,

$$\sum_{i=1}^{n} \ell_i m_{ij} = \lambda \ell_j, \quad \sum_{j=1}^{n} m_{ij} r_j = \lambda r_i. \tag{5}$$

We choose those eigenvectors in such a way that

$$\sum_{i=1}^{n} \ell_i r_i = 1. \tag{6}$$

By (5) and (6), we get

$$\begin{aligned} &\sum_{i,j=1}^{n} \frac{\ell_i m_{ij} r_j}{\lambda} \log \frac{\lambda r_i}{r_j} \\ &\qquad = \sum_{i,j=1}^{n} \frac{\ell_i m_{ij} r_j}{\lambda} \log \lambda + \sum_{i,j=1}^{n} \frac{\ell_i m_{ij} r_j}{\lambda} \log r_i \\ &\qquad\quad - \sum_{i,j=1}^{n} \frac{\ell_i m_{ij} r_j}{\lambda} \log r_j \end{aligned}$$

$$= \sum_{j=1}^{n} \ell_j r_j \log \lambda + \sum_{i=1}^{n} \ell_i r_i \log r_i - \sum_{j=1}^{n} \ell_j r_j \log r_j = \log \lambda. \tag{7}$$

By (5), for each i we have

$$\sum_{j=1}^{n} \frac{m_{ij} r_j}{\lambda r_i} = 1.$$

Thus, since the logarithmic function is concave, we get by Jensen's inequality

$$\sum_{j=1}^{n} \frac{m_{ij} r_j}{\lambda r_i} \log \frac{\lambda r_i}{r_j} \leq \log \sum_{j=1}^{n} \frac{m_{ij} r_j}{\lambda r_i} \cdot \frac{\lambda r_i}{r_j} = \log R_i. \tag{8}$$

Now, from (7) and (8), we get

$$\log \lambda = \sum_{i=1}^{n} \ell_i r_i \sum_{j=1}^{n} \frac{m_{ij} r_j}{\lambda r_i} \log \frac{\lambda r_i}{r_j} \leq \sum_{i=1}^{n} \ell_i r_i \log R_i. \tag{9}$$

When we replace the matrix M by its transpose, ℓ_i's and r_i's will switch and R_i's will become C_i's. Thus,

$$\log \lambda \leq \sum_{i=1}^{n} \ell_i r_i \log C_i. \tag{10}$$

By (9) and (10), we get for every $p \in (0,1)$

$$\log \lambda \leq p \sum_{i=1}^{n} \ell_i r_i \log R_i + (1-p) \sum_{i=1}^{n} \ell_i r_i \log C_i = \sum_{i=1}^{n} \ell_i r_i (p \log R_i + (1-p) \log C_i). \tag{11}$$

In view of (6), this gives us

$$\log \lambda \leq \max\{p \log R_i + (1-p) \log C_i : 1 \leq i \leq n\}, \tag{12}$$

so (4) follows.

Now, if M is just non-negative instead of positive, for every $\varepsilon > 0$, we consider a positive matrix $M_\varepsilon = (m_{ij} + \varepsilon)$ with row sums $R_{\varepsilon,i}$, column sums $C_{\varepsilon,i}$ and spectral radius λ_ε. By what we proved, we have

$$\lambda_\varepsilon \leq \max\{R_{\varepsilon,i}^p C_{\varepsilon,i}^{1-p} : 1 \leq i \leq n\}. \tag{13}$$

Clearly, $\lambda \leq \lambda_\varepsilon$. Moreover, for every i

$$\lim_{\varepsilon \to 0} R_{\varepsilon,i}^p C_{\varepsilon,i}^{1-p} = R_i^p C_i^{1-p}.$$

Thus, we get (4) by taking the limit as $\varepsilon \to 0$ in (13). ■

3. Ideas From Dynamical Systems

Now we explain the source of the proof from the preceding section. This is not essential for the proof, so we will omit definitions. The reader who needs them can find them in many textbooks in Dynamical Systems (a standard one is [Walters, 1982]).

For a positive $n \times n$ non-negative matrix with row sums R_i, column sums C_i, and spectral radius λ, we consider the two-sided full shift on n symbols $\sigma : \Sigma \to \Sigma$. Denote the ith 1-cylinder by $[i]$ and the (i,j)th 2-cylinder by $[ij]$. Let $f : \Sigma \to \mathbb{R}$ be the function constant on 2-cylinders that takes value m_{ij} on $[ij]$. Then the topological pressure $P(\sigma, \log f)$ is equal to $\log \lambda$ and there is a probability measure μ on Σ, invariant for σ, which is the equilibrium state for $(\sigma, \log f)$. Moreover, the system (Σ, σ, μ) is a stationary Markov chain. If $p_i = \mu([i])$ and the transition probabilities are p_{ij}, then

$$P(\sigma, \log f) = -\sum_{i,j=1}^{n} p_i p_{ij} \log p_{ij} + \sum_{i,j=1}^{n} p_i p_{ij} \log m_{ij}.$$

Using an inequality that is standard in the proof of the Variational Principle (see, e.g. formula (6) of [Misiurewicz, 1976]), one gets

$$P(\sigma, \log f) \leq \sum_{i=1}^{n} p_i \log R_i.$$

Now instead of σ we consider σ^{-1}. This corresponds to replacing the matrix M by its transpose. They have the same spectral radii. Moreover, the pressure and metric entropy are preserved by taking the inverse of the map, and the system $(\Sigma, \sigma^{-1}, \mu)$ is also a stationary Markov chain with the same measures of cylinders $[i]$ and transition probabilities

$p_{ji}p_j/p_i$. However, the rows of M^T are the columns of M, so we get now

$$P(\sigma, \log f) \le \sum_{i=1}^{n} p_i \log C_i.$$

In this way we get (9) and (10), and the proof follows as in the preceding section.

4. Entropy of Graphs

A special case of non-negative matrices and their spectral radii, when the matrix has only entries 0 and 1, can be interpreted in terms of Dynamical Systems or Graph Theory.

Thus, if $M = (m_{ij})$ is a 0–1 matrix, we can consider the space

$$\Sigma_M = \{(x_k)_{k=-\infty}^{\infty} : \forall_k\, x_k \in \{1, \dots, n\} \text{ and } m_{x_k x_{k+1}} = 1\}$$

with the topology of a subset of $\{1, \dots, n\}^{\mathbb{Z}}$ with the product topology (in each $\{1, \dots, n\}$, we take the discrete topology). Further, $\sigma_M : \Sigma_M \to \Sigma_M$ is the shift to the left, that is,

$$\sigma_M((x_k)) = (y_k), \quad \text{where } y_k = x_{k+1}.$$

The system (Σ_M, σ_M) is the *subshift of finite type* with *transition matrix* M. The logarithm of the spectral radius of M is equal to the exponential growth rate of the number of cylinders of length k (the sets of sequences with prescribed terms at coordinates $0, 1, \dots, k-1$), and is the *topological entropy* of σ_M.

Alternatively, we can think of a directed graph (*digraph*) G with n vertices, in which there is an arrow from the ith vertex to the jth one if and only if $m_{ij} = 1$. Then the logarithm of the spectral radius of M (we will call it the *entropy* of G and denote it $h(G)$) is the exponential growth rate of the number of paths of length k in G.

The number of arrows beginning in a given vertex v is the *outdegree* of v; the number of arrows ending at v is the *indegree* of v. Note that the outdegree of the ith vertex of G is the ith row sum of M, and the indegree of the ith vertex of G is the ith column sum of M. Thus, we can restate Theorem 1 as follows.

Theorem 2. *Let G be a digraph with n vertices. Let the outdegree of the ith vertex be d_i^+ and its indegree d_i^-. Then for every $p \in [0, 1]$*

$$h(G) \le \max\{p \log d_i^+ + (1-p) \log d_i^- : 1 \le i \le n\}, \tag{14}$$

(*unless the right-hand side is $-\infty$*).

Here we were able to add 0 and 1 as possible values of p to incorporate the trivial upper estimates of (1).

In fact, using (1), we get the following simple lemma.

Lemma 1. *If the outdegree of each vertex of G (or the indegree of each vertex of G) is k then $h(G) = \log k$.*

Let us define the *capacity* of a vertex of G as the geometric mean of its outdegree and indegree. Then from Theorem 2, we get immediately a version of the theorem of Barankin.

Corollary 1. *The entropy of a digraph is not larger than the maximum of the logarithms of the capacities of its vertices.*

Now we consider the following problem. For a digraph G let us denote by $|G|$ an undirected graph that is obtained from G by replacing each arrow by an edge. We assume that in G if there is an arrow from a vertex v to a vertex u then there is no arrow from u to v, so the graph $|G|$ is simple (i.e. has no multiple edges between a pair of vertices or self-loops at a vertex).

Problem. For a given simple undirected graph H, find

$$h_{\mathrm{dir}}(H) := \max\{h(G) : |G| = H\}.$$

Lemma 1 and Corollary 1 provide a quick solution in the case when H is a k-regular graph (the degree of each vertex of H is k) with k even.

Theorem 3. *If H is a k-regular simple undirected graph and k is even, then*

$$h_{\mathrm{dir}}(H) = \log \frac{k}{2}.$$

Proof. If G is any digraph with $|G| = H$, then the capacity of every vertex of G is the geometric mean of two non-negative integers whose sum is k. By the inequality between the geometric and arithmetic means, we see that this capacity is not larger than $k/2$, so by Corollary 1, $h(G) \le \log(k/2)$. Thus, $h_{\mathrm{dir}}(H) \le \log(k/2)$.

On the other hand, by the theorem stated by Euler in 1736 and proved by Hierholzer in 1873, since the degree of each vertex is even, there exists an Eulerian circuit (a closed path that follows each edge of H exactly once). When we follow it and orient edges accordingly, we get a digraph G with the indegree and outdegree of each vertex equal to $k/2$. Now, by Lemma 1, $h(G) = \log(k/2)$. Thus, $h_{\mathrm{dir}}(H) \geq \log(k/2)$. ■

In order to get a similar result for k odd, we have to add an assumption that H is bipartite. This means that the set of vertices of H can be divided into two subsets, such that there are no edges joining vertices from the same subset.

Theorem 4. *If H is a k-regular simple bipartite undirected graph and k is odd, then*

$$h_{\mathrm{dir}}(H) = \log \frac{\sqrt{k^2-1}}{2}.$$

Proof. If G is any digraph with $|G| = H$, then the capacity of every vertex of G is the geometric mean of two non-negative integers whose sum is k. Therefore, this capacity is not larger than $((k+1)/2) \cdot ((k-1)/2) = (k^2-1)/4$, so by Corollary 1, $h(G) \leq \log(\sqrt{k^2-1}/2)$. This proves that $h_{\mathrm{dir}}(H) \leq \log(\sqrt{k^2-1}/2)$.

Since the graph H is bipartite, there exists a partition of the set of vertices of H into two sets, X and Y, such that every edge of H joins a vertex from X with a vertex from Y. Since the degree of every vertex of H is k, by Theorem 2 of [Lenard, 2001] (a simple consequence of the Marriage Lemma), there is a one-to-one function $\varphi\colon X \to Y$, such that every $x \in X$ is joined by an edge with $\varphi(x)$. Similarly, there is a one-to-one function $\psi\colon Y \to X$, such that every $y \in Y$ is joined by an edge with $\psi(y)$. In particular, the cardinalities of X and Y are equal, so φ is also onto. When we erase all edges from x to $\varphi(x)$, we get a graph H' which is $(k-1)$-regular. Thus, as in the proof of Theorem 3, we can orient the edges of H' in such a way that each vertex has outdegree and indegree $(k-1)/2$. Now we orient the removed edges of H to go from X to Y, and in such a way we get a digraph G with $|G| = H$ such that all vertices from X have outdegree $(k+1)/2$ and indegree $(k-1)/2$, while all vertices from Y have outdegree $(k-1)/2$ and indegree $(k+1)/2$.

Let M be the transition matrix of G. Then M^2 has row sums

$$\frac{k+1}{2} \cdot \frac{k-1}{2} = \frac{k^2-1}{4},$$

so by (1), its spectral radius is $(k^2-1)/4$. Thus, the spectral radius of M is $\sqrt{k^2-1}/2$. This proves that $h_{\mathrm{dir}}(H) \geq \log(\sqrt{k^2-1}/2)$. ■

5. Application to Biological Systems

Now we consider the question mentioned in the Introduction: what is the maximal entropy of the k-dimensional hypercube over all possible assignments of directions of edges? This is our Problem in the case when H is the 1-skeleton of the k-dimensional hypercube. The results of the preceding section give immediately the answer.

Theorem 5. *If H is the 1-skeleton of the k-dimensional hypercube then*

(a) *if k is even, then $h_{\mathrm{dir}}(H) = \log \frac{k}{2}$;*
(b) *if k is odd, then $h_{\mathrm{dir}}(H) = \log \frac{\sqrt{k^2-1}}{2}$.*

Proof. The graph H is a k-regular simple undirected graph, so in view of Theorems 3 and 4, the only thing we have to prove is that if k is odd then H is bipartite. To show this, we embed the hypercube into $\mathbb{R}^k$ in a natural way, with vertices at $\{0,1\}^k$. Then we divide the set of vertices into two sets, according to whether the sum of coordinates of the vertex is even or odd. The endpoints of each edge differ exactly at one coordinate, so they belong to different sets of the partition. This completes the proof. ■

6. Examples

In this section we provide two examples. The first one shows that in Theorem 1 we do not get a lower estimate of the spectral radius by the minimum of $R_i^p C_i^{1-p}$, even for the matrices obtained from hypercubes, as in Sec. 5.

Consider the graph from Fig. 1. Four lower vertices have outdegree 2 and indegree 1, while four upper vertices have outdegree 1 and indegree 2. Thus, if $p \in (0,1)$, the values of $R_i^p C_i^{1-p}$ are 2^p and 2^{1-p}; all of them are positive. Nevertheless, all paths either stay in the bottom and follow a period 4 circuit, or go to the top and there follow the period 4 circuit. Therefore, their number grows linearly with the length of the path, so the entropy of the graph is 0.

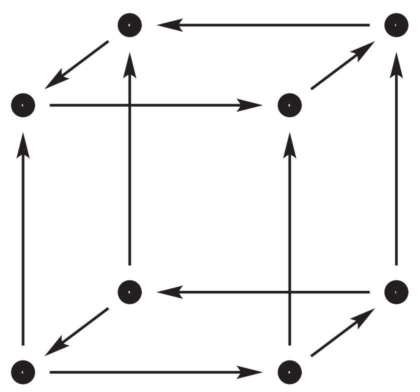

Fig. 1. Entropy zero digraph with positive capacities.

The second example shows that the estimate from Theorem 1 is stronger than its special case with $p = 1/2$. Let M be the following matrix:

$$\begin{pmatrix} 0 & 2 & 1 & 1 \\ 0 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The row sums are $4, 2, 2, 1$, and the column sums $1, 2, 4, 2$. The numbers $R_i^p C_i^{1-p}$ with $p = 1/2$ are 2, 2, $2\sqrt{2}$, $\sqrt{2}$, so their maximum is

$$2\sqrt{2} \approx 2.828427125.$$

On the other hand, the numbers $R_i^p C_i^{1-p}$ with $p = 2/3$ are $2\sqrt[3]{2}$, 2, $2\sqrt[3]{2}$, $\sqrt[3]{2}$, so their maximum is

$$2\sqrt[3]{2} \approx 2.519842100.$$

Of course, this is just an upper estimate; the spectral radius of M is equal to the largest root of the equation $x^3 - 2x - 5 = 0$, that is, approximately 2.094551482. A similar but more complicated example can be constructed on a hypercube.

Acknowledgment

M. Rams was partially supported by MNiSWN grant N201 607640 (Poland).

References

Barankin, E. W. [1945] "Bounds for the characteristic roots of a matrix," *Bull. Amer. Math. Soc.* **5**, 767–770.

Brualdi, R. A. [2010] "Spectra of digraphs," *Lin. Algeb. Appl.* **432**, 2181–2213.

Glass, L. & Pasternack, J. S. [1978] "Stable oscillations in mathematical models of biological control systems," *J. Math. Biol.* **6**, 207–223.

Glass, L. & Siegelmann, H. T. [2010] "Logical and symbolic analysis of robust biological dynamics," *Curr. Opin. Genet. Devel.* **20**, 644–649.

Kolotilina, L. Yu. [2002] "Bounds and inequalities for the Perron root of a nonnegative matrix," (in Russian), *Zap. Nauchn. Sem. POMI* **284**, 77–122 (English version: *J. Math. Sci.* **121** (2004), 2481–2507).

Kwapisz, J. [1996] "On the spectral radius of a directed graph," *J. Graph Th.* **23**, 405–411.

Lenard, A. [2001] "An application of the Marriage Lemma," *Math. Mag.* **74**, 234–238.

Minc, H. [1988] *Nonnegative Matrices* (John Wiley, NY).

Misiurewicz, M. [1976] "A short proof of the variational principle for a $\mathbb{Z}_+^N$ action on a compact space," *Astérisque* **40**, 147–157.

Ostrowski, A. [1951] "Ueber das Nichtverschwinden einer Klasse von Determinanten und die Lokalisierung der charakteristischen Wurzeln von Matrizen," *Compositio Math.* **9**, 209–226 (in German).

Perkins, T. J., Wilds, R. & Glass, L. [2010] "Robust dynamics in minimal hybrid models of genetic networks," *Phil. Trans. Roy. Soc. A* **368**, 4961–4975.

Walters, P. [1982] *An Introduction to Ergodic Theory* (Springer-Verlag, NY).

Wilds, R., Kauffman S. A. & Glass, L. [2008] "Evolution of complex dynamics," *Chaos* **18**, 033109.

Wilds, R. & Glass, L. [2009] "An atlas of robust, stable, high-dimensional limit cycles," *Int. J. Bifurcation and Chaos* **19**, 4055–4096.