On the spectral radius of a directed graph

Jaroslaw Kwapisz [1]
The Center for Dynamical Systems and Nonlinear Studies
Georgia Institute of Technology
Atlanta, GA 30332-0190
e-mail: jarek@math.gatech.edu

April 1994 [2]

Abstract. We provide upper estimates on the spectral radius of a directed graph. In particular we prove that the spectral radius is bounded by the maximum of the geometric mean of in-degree and out-degree taken over all vertices.

By a directed graph G we understand a quadruple (V, E, v_G^-, v_G^+) where V and E are finite sets and v_G^-, v_G^+ are functions from E to V. Elements of V are referred to as *vertices* and that of E as *edges*. For an edge $e \in E$ we denote by $v_G^-(e)$ *the tail* of v and by $v_G^+(e)$ *the head* of e. Notice that our definition allows for multiple edges between two vertices as well as for self loops (also possibly multiple). The number of edges that have a vertex v for the tail we call *the out-degree* of v and denote by $d_G^+(v)$. Analogously, by counting edges with heads at v we get $d_G^-(v)$, the *in-degree* of v. We shall drop subscript G in cases when it is clear what graph we are referring to.

Any sequence $e_1, ..., e_n$ of n edges with the property that $v^+(e_i) = v^-(e_{i+1})$, $i = 1, ..., n-1$, is a *n-walk*. The set of all n-walks in G is denoted by V^n. Given $\alpha \in V^n$ we write $v^+(\alpha)$ and $v^-(\alpha)$ for the head and the tail of the first and the last edge in α correspondingly.

An important number associated with G is the exponent characterizing the growth of $\#V^n$ as $n \to \infty$, namely

$$h(G) := \lim_{n\to\infty} \frac{1}{n} \ln \#V^n.$$

[1] Supported in part by Polish Academy of Sciences grant 210469101 “Iteracje i Fraktale”.

[2] Revised in June 1996.

In graph theory literature one usually encounters $\rho(G) := \exp(h(G))$, called *the index* or *spectral radius* of G. The later name is motivated by the fact that $\rho(G)$ is in fact the spectral radius of the adjacency matrix of G ([CDS80]).

For applications it is crucial to be able to compute or at least estimate $\rho(G)$ for a given graph G. This is a classical problem with numerous results pertaining to it (see [CDS80, CR90]); however, most of them are concerned with special classes of graphs or relate $\rho(G)$ to structural properties of G that are difficult to extract in practice. Therefore the classical estimates due to Frobenius, [3]

$$\inf\{d^+(v) : v \in V\} \le \rho(G) \le \max\{d^+(v) : v \in V\},$$

$$\inf\{d^-(v) : v \in V\} \le \rho(G) \le \max\{d^-(v) : v \in V\},$$

are often the easiest way to quickly estimate $\rho(G)$. The usefulness of these inequalities is greatly due to the fact that they require only the knowledge of degrees at the vertices, a kind of "local data" usually very easy to extract in applications. Nevertheless, for obvious reasons, these estimates work very poorly for graphs having big discrepancies between the out- and in-degrees. We provide "local" bounds that are typically much more efficient in such cases. Also, their simplicity may have some esthetical appeal.

Theorem 1 *If $G = (V, E, v^-, v^+)$ is a directed graph, then the following estimates hold*

(i) $\rho(G) \le d_G^{(-1)} := \max\{\sqrt{d^-(v^+(e)) \cdot d^+(v^-(e))} : e \in E\}$;

(ii) $\rho(G) \le d_G^{(0)} := \max\{\sqrt{d^-(v) \cdot d^+(v)} : v \in V\}$;

(iii) $\rho(G) \le d_G^{(k)} := \max\{\sqrt{d^-(v^-(\alpha)) \cdot d^+(v^+(\alpha))} : \alpha \in V^k\},\ k \in \mathbf{N}$.

Moreover,

(i bis) for any two vertices $u, v \in V$ and $n \in \mathbf{N}$, the number of n-walks starting at u and ending at v does not exceed $(d_G^{(-1)})^n$;

[3] Actually Frobenius is credited with proving the corresponding result in linear algebra (see Th. 1.1 Chpt. 3 in [Min74]).

(ii bis) *for any two fixed edges* $e, f \in E$ *and* $n \in \mathbf{N}$, *the number of* n-*walks starting with* e *and ending with* f *does not exceed* $(d_G^{(0)})^{n-1}$;

(iii bis) *for* $k \in \mathbf{N}$, $n \geq k$, *and any two* k-*walks* α, β, *the number of* n-*walks starting with* α *and ending with* β *does not exceed* $(d_G^{(k)})^{n-k}$.

One may suspect that taking in the definitions of $d_G^{(k)}$-s min instead of max yields estimates of $\rho(G)$ from below, but this is not true even for graphs with irreducible adjacency matrices (see Example 1).

Frobenius's bounds have very simple intuitive proofs based on the fact that every $(n-1)$-walk in G can be extended to at least $\min d_G^+$ n-walks and no more than $\max d_G^+$ n-walks. However, we do not know an intuitive proof of Theorem 1 without using the methods of ergodic theory. An argument exploiting the variational principle for topological entropy and ideas from the theory of measure theoretic entropy is supplied in the end of this note. This approach does not seem to give (i) or the estimates on the number of n-walks stated in (i bis), (ii bis) and (iii bis). An elementary proof of Theorem 1 is based on the following result about matrices.

Lemma 1 *If* $A = (a_{i,j})_{i,j=1,\dots,n}$ *is an* $n \times n$ *matrix with complex entries, then for any* $x = (x_i)_{i=1,\dots,n} \in \mathbf{C}^n$ *we have*

$$\sum_i |\sum_j a_{i,j}x_j|^2 \leq \max\{\sum_p |a_{i,p}| \sum_q |a_{q,k}| : i,k \text{ such that } \ a_{i,k} \neq 0\} \cdot \sum_i |x_i|^2.$$

Proof of Lemma 1. We calculate:

$$\sum_i |\sum_j a_{i,j}x_j|^2 \leq \sum_i (\sum_j |a_{i,j}||x_j|)^2 = \sum_{i,j,k} |a_{i,j}||a_{i,k}||x_j||x_k| \leq$$

$$\sum_{i,j,k} |a_{i,j}||a_{i,k}| \frac{|x_j|^2 + |x_k|^2}{2} = \sum_{i,j,k} |a_{i,j}||a_{i,k}||x_k|^2 \leq \max_k\{\sum_{i,j} |a_{i,j}||a_{i,k}|\} \cdot \sum_k |x_k|^2.$$

Furthermore, we have

$$\max_k\{\sum_{i,j} |a_{i,j}||a_{i,k}|\} \leq$$

$$\max_k \{\sum_i |a_{i,k}| \cdot \max_{i:\ a_{i,k}\neq 0} \{\sum_j |a_{i,j}|\}\} = \max_{i,k:\ a_{i,k}\neq 0} \{\sum_q |a_{q,k}| \cdot \sum_p |a_{i,p}|\}.$$

The lemma follows. □

The maximum appearing in Lemma 1 deserves a compact notation. We set for any matrix $A = (a_{i,j})$

$$q(A) := \max\left\{ \sqrt{\sum_p |a_{i,p}| \sum_q |a_{q,k}|} : i,k \text{ such that } a_{i,k} \neq 0 \right\}.$$

By writing $\|\cdot\|$ for the Euclidean norm in $\mathbf{R}^n$ and by using the same notation for the corresponding operator norm, we may rephrase our lemma as follows

Corollary 1 *For $A = (a_{i,j})_{i,j=1,\ldots,n}$, $a_{i,j} \in \mathbf{C}$, we have*

$$\|A\| \leq q(A).$$

In particular, the spectral radius $\rho(A)$ of A satisfies

$$\rho(A) \leq q(A).$$

Proof of Theorem 1. The claims (i), (ii), (iii) follow from the definition of $\rho(G)$ and (i bis), (ii bis), (iii bis) respectively. We prove (i bis) first. Let $A = (a_{i,j})_{i,j\in V}$ be the adjacency matrix of G; by definition, $a_{i,j}$ is equal to the number of edges joining i to j. It is easy to see that $q(A) = d_G^{(-1)}$. Consider the n-th power $A^n =: (a_{i,j}^{(n)})_{i,j\in V}$ acting on $x = (0, \ldots, 1_{(j^{th}\text{ place})}, \ldots 0)$. Corollary 1 yields

$$a_{i,j}^{(n)} \leq \sqrt{\sum_k (a_{k,j}^{(n)})^2} = \|A^n \cdot x\| \leq \|A\|^n \leq q(A)^n = (d_G^{(-1)})^n.$$

The standard observation that $a_{i,j}^{(n)}$ is the number of n-walks from i to j ends the argument for (i bis).

We claim that (ii bis) arises as a result of applying (i bis) to the line graph G^* which is obtained by "switching" the roles of edges and vertices in G (i.e. the edges of G are vertices of G^* and G^* has a directed edge between any two of them if and only if the head of the first coincides with the tail of the other). The statements (c) and (d) below assure that (i bis) for G^* indeed translates to (ii bis) for G.

Fact 1 *(a) Any edge e of G is simultaneously a vertex of G^* (also vice versa) and $d^+_{G^*}(e) = d^+_G(v^+_G(e))$ as well as $d^-_{G^*}(e) = d^-_G(v^-_G(e))$.*

(b) Any edge e^* in G^* determines a unique vertex $v = v^+_G(v^-_{G^*}(e^*))$ in G and $d^+_G(v) = d^+_{G^*}(v^-_{G^*}(e^*))$ as well as $d^-_G(v) = d^-_{G^*}(v^+_{G^*}(e^*))$.

(c) $d^{(0)}_G = d^{(-1)}_{G^*}$.

(d) There is a natural bijection between n-walks in G and $(n-1)$-walks in G^* (for $n > 1$).

(e) For any $k \in \mathbf{N}$, $d^{(k-1)}_G = d^{(k)}_{G^*}$.

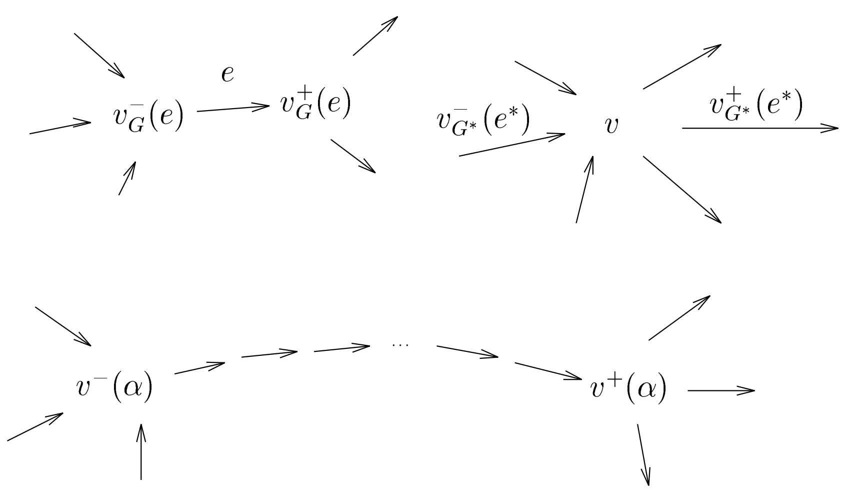

Figure 1.

Proof of Fact 1. The claims (a) and (b) can be easily verified by looking at Figure 1. Since all vertices v of G that do not correspond to an edge of G^* have $d^+_G(v) \cdot d^-_G(v) = 0$, the part (c) is a simple consequence of (b). To see the bijection in (d), observe that a n-walk in G, in terms of G^*, represents a sequence of n vertices each possessing an edge joining it to the next one. This uniquely determines a $(n-1)$-walk in G^* because G^* has at most one edge between any two vertices. Also vice versa, the n vertices visited by a $(n-1)$-walk in G^* form a sequence of n edges in G that constitutes a n-walk in G. Regarding (e), for $k = 1$, we get it directly from (a). For $k > 1$, we employ the bijection from (d) and then use (a). $\square$

To prove (iii bis) of Theorem 1 we proceed by induction. Having proven the claim for k its application to G^* yields the claim with $k + 1$ for G, as can be seen from (d), (e) in Fact 1. To start up ($k = 0$) we use the already established (ii bis). This ends the proof of Theorem 1. $\square$

Remark 1 Lemma 1 is valid even for infinite matrices as long as the absolute values of the entries in every column or row are summable. Such matrices correspond to (possibly infinite) directed graphs with finite in- and out-degrees at every vertex. Theorem 1 trivially extends to this class of graphs.

Remark 2 In many applications one encounters directed graphs with real positive weights assigned to every edge. With every n-walk we associate then a weight that is the product of weights of its edges. By counting all edges and n-walks with these weights one gets a natural generalization of Theorem 1. We supply here a rough sketch of the proof. In the case when the weights are natural numbers, we can replace every edge by as many edges joining the same vertices as its weight is. Application of Theorem 1 as stated to thus obtained graph gives us the desired conclusions about the original weighted graph. Now, notice that under the scaling of all weights by the same multiplicative factor $N > 0$ the quantities $d^+, d^-, d^{(k)}, \rho(G)$ transform also to their multiples by N. This solves the problem for rational weights since we can take for N the common denominator of all weights. Finally, the general case of real positive weights is taken care of by an approximation argument based on continuity of $d^{(k)}$ and $\rho(G)$ as functions of weights.

Example 1.

Consider the graph G defined by Figure 2. We can see that, even though both out-degree d^+ and in-degree d^- assume all the values in the set $\{1, 2, 4\}$, the geometric mean $\sqrt{d^+ \cdot d^-}$ is equal to $\sqrt{4} = 2$ for every vertex. By Theorem 1 the spectral radius $\rho(G)$ is not greater than 2. With the aid of Mathematica we checked that the leading eigenvalue of the adjacency matrix

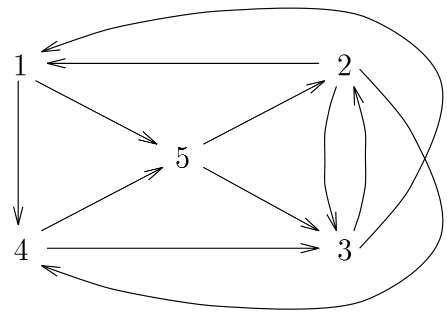

Figure 2.

of G,

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 2 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \end{pmatrix},$$

is equal to $1.90... < 2$. Hence, our remark after the statement of Theorem 1 is motivated at least in the context of (ii). The graph G^* (not pictured) provides an appropriate example for (i). We did not seek for examples related to (iii). [4]

Let us now sketch an argument that proves (ii) part of Theorem 1 on the grounds of ergodic theory. All facts and definitions used without comment are standard and can be found either in [Wal82] or in the first three chapters of Parry's book ([Par69]). For simplicity we will assume that there are no multiple edges in G so that any walk is uniquely determined by the sequence of vertices it passes through. The graph G defines a *subshift of finite type* (Λ, σ) where $\Lambda \subset V^{\mathbf{Z}}$ consists of all bi-infinite sequences of vertices traced by bi-infinite walks in G, and σ is a shift transformation $(\sigma((v_i)) := (v_{i+1})$, see [Wal82]. In this context, $\rho(G) = \exp(h(\sigma))$, with $h(\sigma)$ standing for the topological entropy of σ. Let $M(\Lambda, \sigma)$ be the set of all probabilistic measures

[4] A simpler example is provided by the graph with two vertices and three edges each joining a different ordered pair of vertices. The corresponding adjacency matrix is *reducible* ([Min74]), which is not the case for A.

on Λ invariant under σ. The variational principle ([Wal82]) says that

$$h(\sigma)=\sup\{h_\mu(\sigma):\mu\in M(\Lambda,\sigma)\}.$$

Thus we will be done after proving that for an arbitrary $\mu\in M(\Lambda,\sigma)$ we have

$$h_\mu(\sigma)\le \ln d_G^{(0)}.$$

Denote by $\mathcal{V}$ the partition of Λ into sets $V_v=\{(x_i)_{i\in\mathbf{Z}}\in\Lambda : x_0=v\}$, where v runs over all vertices of G. The partition $\mathcal{V}$ is a generator for σ (i.e. $\bigvee_{i=-\infty}^{\infty}\sigma^i\mathcal{V}$ coincides with $\sigma-$algebra of all Borel subsets of Λ) so we can compute $h_\mu(\sigma)$ as the conditional entropy ([Par69]),

$$h_\mu(\sigma)=H(\sigma\mathcal{V}/\bigvee_{i=0}^{\infty}\sigma^{-i}\mathcal{V}):=\int I(\mathcal{V}/\bigvee_{i=0}^{\infty}\sigma^{-i}\mathcal{V})d\mu.$$

The conditional information function $I(\sigma\mathcal{V}/\bigvee_{i=0}^{\infty}\sigma^{-i}\mathcal{V})$ above is defined as $\sum_{U\in\sigma\mathcal{V}}-\ln E(\chi_U/\bigvee_{i=0}^{\infty}\sigma^{-i}\mathcal{V})$, where $E(.../...)$ denotes the conditional expectation ([Par69]). Now, since $\bigvee_{i=0}^{\infty}\sigma^{-i}\mathcal{V}$ is finer than $\mathcal{V}$, we have [5]

$$E\left(I(\sigma\mathcal{V}/\bigvee_{i=0}^{\infty}\ \mathcal{V})/\mathcal{V}\right)\le I(\sigma\mathcal{V}/\mathcal{V}).$$

The function $I(\sigma\mathcal{V}/\mathcal{V})$ is $\mathcal{V}$-measurable and, by using the standard estimate of entropy by the logarithm of cardinality of the partition, [6] we see that for every $v\in V$

$$I(\sigma\mathcal{V}/\mathcal{V})(x)\le \ln d^-(v),\quad x\in V_v.$$

The last three inequalities assemble to the following

$$h_\mu(\sigma)\le\int I(\sigma\mathcal{V}/\mathcal{V})d\mu\le\sum_{v\in V}\mu(V_v)\cdot\ln d^-(v).$$

---

[5] It is generally true that if $\mathcal{A}$ and $\mathcal{C}$ are finite partitions and $\mathcal{D}$ is a σ-algebra finer than $\mathcal{C}$, then $E(I(\mathcal{A}/\mathcal{D})/\mathcal{C})\le I(\mathcal{A}/\mathcal{C})$. This follows from the standard fact that $H(\mathcal{A}_1/\mathcal{D}_1)\le H(\mathcal{A}_1)$ for any partition $\mathcal{A}_1$ and σ-algebra $\mathcal{D}_1$ (see 1.23 in [Par69]). To see that, fix $C\in\mathcal{C}$ and consider it as a probability space with respect to the normalized restriction of μ. Let $\mathcal{A}_1$ and $\mathcal{D}_1$ to be the restrictions to C of $\mathcal{A}$ and $\mathcal{D}$ respectively. The later inequality coincides with the first restricted to the atom C.

[6] As before, we really apply this fact to the restriction of our probability space and the partition $\sigma\mathcal{V}$ to every atom of $\mathcal{V}$.

Analogous considerations for the inverse of σ lead to

$$h_\mu(\sigma^{-1}) \le \int I(\sigma^{-1}\mathcal{V}/\mathcal{V})d\mu \le \sum_{v\in V} \mu(V_v)\cdot \ln d^+(v).$$

Since $h_\mu(\sigma) = h_\mu(\sigma^{-1})$ ([Wal82]), the two inequalities above yield

$$h_\mu(\sigma) \le \sum_{v\in V} \mu(V_v)\cdot \frac{\ln d^-(v) + \ln d^+(v)}{2} \le \ln d_G^{(0)}.$$

This ends the proof of part (ii) of Theorem 1. Part (iii), recall, follows from (ii). As far as (i) is concerned, we do not know how to treat it using similar argument to the one presented above.

References

[CDS80] D. Cvetkovic, M. Doob, and H. Sachs, *Spectra of graphs*, Academic Press, New York, 1980.

[CR90] D. Cvetkovic and P. Rowlinson, *The largest eigenvalue of a graph*, Linear and Multilinear Algebra **28** (1990), 3–33.

[Min74] H. Minc, *Nonnegative matrices*, Technion, Israel Institute of Technology, Haifa, 1974.

[Par69] W. Parry, *Entropy and generators in ergodic theory*, W. A. Benjamin, New York, 1969.

[Wal82] P. Walters, *An introduction to ergodic theory*, Graduate texts in mathematics, vol. 79, Springer-Verlag, New York, 1982.