Eigenvalue Bounds

A. J. Seary and W. D. Richards
School of Communication
Simon Fraser University
Burnaby BC Canada V5A1S6
email: seary@sfu.ca, richards@sfu.ca
Presented to INSNA Sunbelt XVIII
Draft - not for citation or distribution

Abstract: We view the (signed) eigenvalues produced by Correspondence Analysis as belonging to the the spectrum associated with the combinatorial Laplacian operator. We show how this spectrum can be used to provide bounds on distance and diameter of a graph.

For random and semi-random (Watts & Strogatz, 1998) graphs, the extreme eigenvalues provide bounds on diameter. When a graph is far from being random, the eigenvalues provide bounds on the distances between subsets of nodes, and thus provide bounds on the number of cohesive blocks (where the eigenvectors provide blocking information).

Introduction

In this paper we will assume that all graphs are undirected. Then the singular values and vectors produced by Correspondence Analysis (CA) are simply eigenvalues and eigenvectors. CA has been suggested as a method for finding block structure in graphs and for graph visualisation (Noma & Smith, 1986; Seary & Richards, 1995). The fact that the squared eigenvalues are a partition of χ^2 (Greenacre, 1984) may be used to argue that the corresponding eigenvectors measure "important" contributions to the structure of a graph (Richards and Seary, 1997) has led to criticisms of CA as a statistical method (Roberts, 1996). E.g., how do we deal with the "missing values" along the diagonal?

We avoid such problems completely by viewing CA as a technique of Spectral Graph Theory (SGT). In fact CA belongs to a closely-knit family as we describe below. SGT is the study of the relationship between spectra of graphs and graph invariants such as diameter, connectivity, expansion, cover time and many others. We will concentrate on diameter and distances between subsets, and show the relationship between these invariants and the eigenvalues.

Definitions

A good introduction to SGT is found in "Spectra of Graphs" by Cvetkovic, Doob and Sachs(CDS, 1980). Many of the results below come from the more recent "Spectral Graph Theory" by Fan Chung (Chung, 1995).

Assume G is an undirected connected loopless graph without multiple edges which is not complete. (The definitions below can be extended to weighted graphs, but for simplicity we will not consider multigraphs here. Assuming G is connected and not complete avoids certain trivial results). G has nodes V and edges E, with $|V| = n$.

The *Adjacency Matrix* **A**(G) of graph G is a binary matrix with

$\mathbf{A}(i,j) = 1$ if i is connected to j

$= 0$ otherwise

The eigenpairs of **A** are $(\alpha_i, \mathbf{a_i})$ such that $\mathbf{A}\mathbf{a_i} = \alpha_i\mathbf{a_i}$

If G is k-regular, then $\mathbf{a}_0 = \mathbf{1}/\sqrt{n}$, with $\alpha_0 = \max(\alpha_i) = k$.

The *Laplacian Matrix* **L'**(G) is a matrix with

$\mathbf{L'}(i,j) = -1$ if i is connected to j

$\mathbf{L'}(i,i) = \deg(i)$ where $\deg(i)$ is the degree of node i

$\mathbf{L'}(i,j) = 0$ otherwise

The eigenpairs of **L'** are $(\lambda'_i, \mathbf{l'_i})$ with $\lambda'_0 = \min(\lambda'_i) = 0$ and $\mathbf{l'_1} = \mathbf{1}/\sqrt{n}$

The $\mathbf{l'_i}$ are mutually orthogonal and $0=\lambda'_0 \le \lambda'_1 \le ,..., \le \lambda'_{n-1} \le n$

The multiplicity of 0 as an eigenvalue is equal to the number of components in G.

There are a number of other equvalent definitions of **L'** the simplest being:

$\mathbf{L'} = \mathbf{D} - \mathbf{A}$

where **D** is the diagonal matrix of node degrees.

The *CA matrix* **C**(G) is a matrix with

$\mathbf{C}(i,j) = 1/(\sqrt{(\deg(i)}\sqrt{\deg(j)})$ if i is connected to j

$\mathbf{C}(i,j) = 0$ otherwise

We may view **C** as $\mathbf{D^{-1/2}AD^{-1/2}}$. Let $(\gamma_i, \mathbf{c_i})$ be the eigenpairs of **C**.

We have $1=\gamma_0 = \max(\gamma_i) \ge \gamma_1 \ge ,..., \ge \gamma_{n-1} \ge -1$

Then CA calculates the pairs

$(\gamma_i, \mathbf{D}^{-1/2}\gamma_i\,\mathbf{c_i})$

Also, CA generally removes the (normalised) χ^2 expecteds, which removes the eigenpair belonging to eigenvalue 1 and produces a "trivial" vector of length 0.

Most versions of CA assume that G is not symmetric and ignore the signs of γ_i, though these are important.

The *Normal matrix* **N**(G) is **C**(G). The eigenpairs are

$$(\gamma_i , \mathbf{D}^{-1/2}\mathbf{c_i}) = (\nu_i , \mathbf{n_i})$$

This differs from CA only in the definition of the vectors, which remain normalized.

We have that

$$\mathbf{n_i}\mathbf{D}\mathbf{n_i} = \delta_{ij}$$

That is, the vectors are orthonormal in the **D** (or χ^2) metric. The Normal spectrum is referred to as the Q-spectrum in CDS.

An interesting property of **N**(G) is that we can add a constant value *c* along the diagonal without changing the eigenvectors. so that the eigenvalues become $c + \nu_i$ and the eigenvectors are unchanged (Seary & Richards, 1995). This corresponds to adding a constant *c*deg(i) to the original adjacency matrix **A**.

The cohesive (positive eigenvalue) on-diagonal blocks of G may be emphasized by shifting the spectrum as follows:

$$\nu' = (1+\nu) / 2$$

This makes all eigenvalues positive: positive ν_i become closer to 1, while negative ν_i become close to zero. This equation goes too far, since the most negative ν_{n-1} is only -1 for bipartite graphs. In general, we need only shift by $\nu_{n-1} \geq -1$ to ensure all eigenvalues are positive:

$$\nu' = (\nu_i - \nu_{n-1}) / 2$$

The shifting technique will be used extensively below.

The *combinatorial Laplacian* **L**(G) (Dodziuk & Kendall, 1986; Chung, Gregor'yan & Yau, 1996) is simply **I-N**(G). By the remark above, the eigenvalues are $\lambda_i = 1-\nu_i$. The eigenvectors are those of **L** so that

$$\mathbf{l_i}\mathbf{l_j} = \delta_{ij}$$

That is, the eigenvectors form an orthogonal system, which we will use below.

Note that while $\mathbf{n_0} = \mathbf{1}/\sqrt{n}$ is a constant vector, $\mathbf{l_0} = \mathbf{D^{1/2}1}$ is not. We prefer to use $\mathbf{n}_i$ since the components satisfy the χ^2 measure of distance: nodes with similar row-profiles have similar components.

We apologize for all these different matrices and spectra (though the last three are closely related). In fact, we can also define **L** in terms of **L'** as follows:

$$\mathbf{L} = \mathbf{D}^{-1/2}\mathbf{L'}\mathbf{D^{-1/2}}$$

We have

$$\{\lambda\} = \{1- \nu\} \text{ and } \{\nu\} = \{1 - \lambda\}$$

so that

$$0 = \lambda_0 = \min(\lambda_i) \le \lambda_1 \le ,\ldots, \le \lambda_{n-1} \le 2$$

Just as for **L'** the eigenvalues of **L** are all positive. Also:

$$\{\mathbf{l}\} = \{\mathbf{D}^{1/2}\,\mathbf{n}\} \text{ and } \{\mathbf{n}\} = \{\mathbf{D}^{-1/2}\,\mathbf{l}\}$$

One further definition (or redefinition):

The *volume* of a Graph G is

$$\text{vol } G = \Sigma_{\mathbf{i} \in \mathbf{V}} \deg(i) = 2|E|$$

The volume of any subset $S \subset V$ is just

$$\text{vol } S = \Sigma_{\mathbf{i} \in \mathbf{S}} \deg(i)$$

so that the volume of a node is just its node degree.

The term volume is used because of an analogy to the continuous case (where the Laplacian is ∇^2). It turns out that many important results involving the continuous Laplacian on Riemannian manifolds can be translated to similar results on graphs (and vice-versa). A small cottage industry arose in the 80's doing just that, with many applications in network and algorithm design. The main goal was to explicitly construct graphs that had desirable properties of random graphs: small diameter, many disjoint paths, and sparsity. Because of this, the first attempts to bound diameter in terms of eigenvalues worked reasonably well for random-like graphs, but were very poor for graphs with block structure. Also many important results about k-regular graphs using the adjacency spectrum can be translated into results on general graphs using the Laplacian spectrum. For example:

For G k-regular, the *complexity* $\kappa(G) = 1/n\Pi_{i=1}(k - \alpha_i)$

For G in general, $\kappa(G) = 1/n\Pi_{i=1}\lambda'_i$

(This is a very old result: Kirchoff's Matrix-Tree Theorem).

In the 90's, another cottage industry arose with the realisation that Markov chains could be analysed by using the combinatorial Laplacian (Jerrum & Sinclair, 1989; Diaconis & Stroock, 1991): translating results from the Laplacian **L'** to combinatorial **L** as in

For G in general, $\kappa(G) = \Pi_{i=1}\lambda_i \Pi_i \text{ vol}(i)/\Sigma_{\mathbf{i}}\text{vol}(i)$

This result (Runge, 1976) illustrates that the translation requires careful bookkeeping about the degree of each vertex. The advantage is that we can make arguments in more detail than is possible for **L'** (and these same arguments apply almost trivially to the Normal spectrum).

Bounds on diameter

We will now present a series of eigenvalue bounds of diameter. In order to show the improvement in the bounds over the last decade, we will use the technique introduced by (Watts and Strogatz, 1998) to construct a series of graphs with constant edge density and variable randomness.

Their technique starts with a circulant graph of width k, and rewires edges at random with probability p. For details see (WS, 1998). WS used this process to show that clusterability (cohesion) decreases much more slowly than mean distance for increasing p, so that so-called "Small World" (semi-random) graphs with both clusters and short distances are easy to construct. We present results from their model in Figure 1, which shows cohesion (dotted) and diameter (solid) for a circulant graphs with 1000 nodes and width 10 over a range of

log(p) from -6 to 0. The results shown are means for 20 random graphs at each probability p.

We start with a simple example to illustrate the technique used to estimate bounds on diameter devised by (Chung, 1988).

We can naively calculate diameter by simply raising matrix $\mathbf{A}(G)$ to a power m for which all entries in $\mathbf{A}^m$ are >0. Then the diameter is m, since there is a path between every node in G.

However, for large G this is not very efficient. Instead, we express $\mathbf{A}^m$ in terms of the eigendecomposition. The following result applies only to regular graphs, but illustrates the methods we will use later. We know $\mathbf{a}_0 = \mathbf{1}/\sqrt{n}$ with $\alpha_0 = k = \max(\alpha_i)$ so that

$(\mathbf{A}^m)_{rs} = \Sigma_i(\alpha_i)^m(\mathbf{a_i a_i}^*)_{rs}$	eigendecomposition
$\geq k^m/n - \lvert\Sigma_{\mathbf{i>0}}(\alpha_i)^m\,(\mathbf{a_i})_\mathbf{r}(\mathbf{a_i}^*)_s\rvert$	force inequality
$\geq k^m/n - \lvert\alpha\rvert^m\{\Sigma_{\mathbf{i>0}}\lvert(\mathbf{a_i})_\mathbf{r}\rvert\,\lvert(\mathbf{a_i}^*)_s\rvert\}$	$\lvert\alpha\rvert=\lvert\alpha_1\rvert=\max_{i>0}\lvert\alpha_i\rvert$
$\geq k^m/n - \lvert\alpha\rvert^m\{\Sigma_{\mathbf{i>0}}\,(\mathbf{a_i})_\mathbf{r}^{\mathbf{2}}\}^{1/2}\{\Sigma_{\mathbf{i>0}}(\mathbf{a_i}^*)_s^{\,2}\}^{1/2}$	Cauchy-Schwartz
$= k^m/n - \lvert\alpha\rvert^m\{1\mathbf{-}(\mathbf{a_0})_\mathbf{r}^{\mathbf{2}}\}^{1/2}\{1\mathbf{-}(\mathbf{a_0}^*)_s^{\,2}\}^{1/2}$	constant eigenvector
$= k^m/n - \alpha^m(1\mathbf{-}\ 1/n)$	since $\mathbf{a}_0 = 1/\sqrt{n}$
> 0	if the two RHS terms cancel

Now, to make the two terms on the right cancel, choose

$$m = \lceil \ln(n-1)/\ln(k/\alpha)\rceil$$

so that

$$\mathrm{Diam}(G) \leq \lceil \ln(n-1)/\ln(k/\alpha)\rceil$$

The steps followed here are very similar for bounds developed below. The bound is very poor for the starting circulant (regular) graph used in the WS model.

The dashed curve shows an eigenvalue bound for distance based on combinatorial Laplacian eigenvalues, due to (Chung, Faber & Manteuffel, 1994). The bound is fair for p=1 but poor elsewhere. The method used for this bound is reminiscent of sparse matrix techniques for finding eigenpairs. Here we begin to use the detail available with $\mathbf{L}$ by calculating

$$\mathrm{Diam}(G) = \max(\mathrm{dist}(X,Y)) \quad X,Y \in V(G)$$

To bound the distances between all pairs of subsets, define vectors

$\psi_X(x) = 1$ if $x \in X$, 0 otherwise

$\psi_Y(y) = 1$ if $y \in Y$, 0 otherwise

Now express ψ_X and ψ_Y as Fourier series in $\mathbf{l}_i$

$$\mathbf{D}^{1/2}\psi_X = \Sigma\, a_i \mathbf{l}_i \quad , \quad \mathbf{D}^{1/2}\psi_Y = \Sigma\, b_i \mathbf{l}_i$$

Note that

$$a_0 = \mathrm{vol}X/\sqrt{\mathrm{vol}G}\ ,\ b_0 = \mathrm{vol}Y/\sqrt{\mathrm{vol}G}$$

and $\quad \Sigma_{i>0} a_i^2 = \|\mathbf{D}^{1/2}\psi_X\|^2 - a_0^2$

$= \mathrm{vol}X - (\mathrm{vol}X)^2/\mathrm{vol}G$

$= \mathrm{vol}X(\mathrm{vol}G\backslash X)/\mathrm{vol}G$

$=$ G\X (the *complement* of X) $= \dfrac{vol X\ vol \bar{X}}{vol G}\quad \bar{X}$

Similarly,

$$\sum_{i>0} b_i^2 = \frac{vol Y\ vol \bar{Y}}{vol G}$$

To estimate distance, form the inner product

$\mathrm{IP} = \langle \mathbf{D}^{1/2}\psi_Y\ ,\ p_t(\mathbf{L})\mathbf{D}^{1/2}\psi_X \rangle$

$= \Sigma_{\mathbf{i}}\, p_t(\lambda_i)\, a_i b_i$ inner product with $\mathbf{l_i l_j} = \delta_{ij}$

$= a_0 b_0 + \Sigma_{\mathbf{i>0}}\, p_t(\lambda_i)\, a_i b_i$ constant eigenvector

where $p_t(z)$ is some polynomial in z. We choose $p_t(z) = (1-z)^t$ so that, assuming

$1 - \lambda_1 \geq \lambda_{n-1} - 1$ (or $|\nu_1| \geq |\nu_{n-1}|$) (we will relax this restriction later),

then $\quad |p_t(\lambda_i)| \leq (1-\lambda)^t$ where $\lambda = \lambda_1$

So $\quad \mathrm{IP} \geq (\mathrm{vol}X\mathrm{vol}Y)/\mathrm{vol}G - (1-\lambda)^t \left(\Sigma_{\mathbf{i>0}} a_i^2 b_i^2\right)^{1/2}$ repeating the steps above

$$= \frac{vol X\ vol Y}{vol G} - (1-\lambda)^t \frac{\sqrt{vol X\ vol \bar{X}\ vol Y\ vol \bar{Y}}}{vol G}$$

Now to ensure that the two terms on the right cancel (so that IP>0), choose

$$t \geq \frac{\ln \sqrt{\dfrac{vol\,\bar{X}\; vol\,\bar{Y}}{vol\,X\; vol\,Y}}}{\ln \dfrac{1}{1-\lambda}}$$

so that

$$Diam(G) \leq \max \left\lceil \frac{\ln \sqrt{\dfrac{vol\,\bar{X}\;\; vol\,\bar{Y}}{vol\,X\;\; vol\,Y}}}{\ln \dfrac{1}{1-\lambda}} \right\rceil$$

To bound the diameter, take the maximum value by choosing X and Y as single nodes with the smallest degree.

Trick #1: Set $\lambda = 2\lambda/(\lambda_{n-1} + \lambda_1)$ i.e, shift the spectrum

Then the bound for diameter becomes

$$Diam(G) \leq \max \left\lceil \frac{\ln \sqrt{\dfrac{vol\,\bar{X}\;\; vol\,\bar{Y}}{vol\,X\;\; vol\,Y}}}{\ln \dfrac{\lambda_{n-1} + \lambda_1}{\lambda_{n-1} - \lambda_1}} \right\rceil$$

This is the lighter dashed curve shown in Figure 1.

Trick #2: $p_t(z) = (1-z)^t$ is not the best polynomial. A better choice is the Chebyshev polynomial $T_t(z)$ usually defined as

$$T_t(z) = \cos(t \cos^{-1}(z))$$
$$= \cosh(t \cosh^{-1}(z))$$

The last expression is not as familiar, but note that $\cosh^{-1}(z)$ is very close to $\ln(z)$ for large z. We will use the polynomial

$$S_t(z) = \frac{T_t\left(\dfrac{\lambda_{n-1} + \lambda_1 - 2z}{\lambda_{n-1} - \lambda_1}\right)}{T_t\left(\dfrac{\lambda_{n-1} + \lambda_1}{\lambda_{n-1} - \lambda_1}\right)}$$

By the minimax property of Chebyshev polynomials,

$$\max_{z\in[\lambda_1,\lambda_{n-1}]} S_t(\lambda_1) \geq 1/T_t((\lambda_{n-1} + \lambda_1)/(\lambda_{n-1} - \lambda_1))$$

Now the inequality becomes

$$IP \geq (volX volY)/volG - S_t(\lambda)\,(volX volY(volG\backslash X)(volG\backslash Y))^{1/2}/volG$$

To make the RHS equal to 0 choose

$$t \geq \frac{\cosh^{-1}\sqrt{\dfrac{vol\,\bar{X}\ vol\,\bar{Y}}{vol\,X\ vol\,Y}}}{\cosh^{-1}\dfrac{\lambda_{n-1}+\lambda_1}{\lambda_{n-1}-\lambda_1}}$$

Then the bound on diameter becomes

$$Diam(G) \leq \max\left[\frac{\cosh^{-1}\sqrt{\dfrac{vol\,\bar{X}\ vol\,\bar{Y}}{vol\,X\ vol\,Y}}}{\cosh^{-1}\dfrac{\lambda_{n-1}+\lambda_1}{\lambda_{n-1}-\lambda_1}}\right]$$

where the maximum is ensured by choosing X and Y to be nodes with the smallest degrees. This gives the dash-dot curve in figure 1. The bound is much better at low p, because

$\cosh^{-1}(z)$ behaves quite differently from ln(z) around z=1 (or λ_1 near 0).

Can the bound be improved at low p? In (Chung, 1996) the numerator is replaced by

$\cosh^{-1}(volG/volXvolY)$ which is much better at p=0 and quite tight at p=1. This is shown by the lighter dash-dot curve in Figure 1.

Bounds on distance between subsets

We can use very similar methods to find bounds on distances among many subsets (Chung,1995). Let $X_i \subset V$, i = 1 ,..., k+1 be disjoint subsets of G.

Then

$$\min_{i\neq j}\ dist\,(X_i, X_j) \leq \max_{i\neq j}\left[\frac{\ln\sqrt{\dfrac{vol\,\bar{X}_i\ vol\,\bar{X}_j}{vol\,X_j\ vol\,X_j}}}{\ln\dfrac{\lambda_{n-1}+\lambda_k}{\lambda_{n-1}-\lambda_k}}\right]$$

For example, choose 2 nodes: the distance between them is bound by the diameter (using λ_1), as we have seen.

To show this, let X and Y be two distinct subsets among the X_i, and consider

$$\langle \mathbf{D}^{1/2}\psi_Y , (\mathbf{I}-\mathbf{L})^t\mathbf{D}^{1/2}\psi_X\rangle \geq a_0 b_0 + \Sigma_{\mathbf{i<k}}(1 - \lambda_i)^t a_i b_i - \Sigma_{\mathbf{i>k}}(1 - \lambda_i)^t a_i b_i$$

In (Chung, Grigor'yan and Yau, 1996) it is shown that we can always choose X, Y such that

$$\Sigma_{\mathbf{i<k}}(1 - \lambda_i)^t a_i b_i \geq 0$$

so that

$$\langle \boldsymbol{D}^{½}\psi_Y , (\boldsymbol{I}-\boldsymbol{L})^t \boldsymbol{D}^{½}\psi_X \rangle > \frac{vol X\, vol Y}{vol G} - (1-\lambda)^t \frac{\sqrt{vol X\ vol Y\ vol \bar{X}\ vol \bar{Y}}}{vol G}$$

Now choose

$$t \geq \frac{\ln \sqrt{\dfrac{vol\ \bar{X}\ vol\ \bar{Y}}{vol X\ vol Y}}}{\ln \dfrac{\lambda_{n-1} + \lambda_k}{\lambda_{n-1} - \lambda_k}}$$

to ensure that the terms on the RHS cancel.

Can these bounds be improved? In (Chung, 1996) the numerator is replaced by

ln(volG/volXvolY)

This is some improvement, but better bounds would result by replacing the denominator. As before, we can find the maximum for *all* k by fixing the subsets in the numerator to be the nodes with smallest degree. The result implies that for large λ_k the distances are small. Another interpretation is that if the upper bound on distance is large for a given k, then *some* pair may be distant, but if *all* distance bounds are small for a given λ_k, there are too many subsets to allow this. Therefore, we can use these distance bounds to get an upper bound in the number of (cohesive, on-diagonal) blocks to look for.

References

(Chung, 1988) F.R.K. Chung, "Diameters and Eigenvalues", *J. Am. Math. Soc.* 1988 **2** (2) ,187-196.

(Chung, Faber & Manteuffel, 1994) F.R.K. Chung, V. Faber, & T.A. Manteuffel, "An upper bound in the diameter of a graph from eigenvalues associated with its Laplacian", *SIAM J. Discrete Math.* 1994 **7** (3) 443-457.

(Chung, 1995) F.K.R. Chung, *Spectral Graph Theory,* CBMS Lecture Notes, AMS Publication, 1995.

(Chung, 1996) F.K.R. Chung, "Eigenvalues of graphs", *Proceedings of ICM,* Zürich, 1994, Birkhäuser Verlag, Berlin, 1333-1342.

(Cvetkovic, Doob & Sachs, 1980) D.M. Cvetkovic, M. Doob, & H. Sachs, *Spectra of Graphs, Theory and Application,* Academic Press, 1980.

(Diaconis & Stroock, 1991) P. Diaconis, D.W. Stroock, "Geometric bounds for eigenvalues of Markov chains", *Ann. Appl. Prob.* 1991 **1** (1) 36-61.

(Dodziuk & Kendall, 1986) J. Dodziuk, W. S. Kendall, "Combinatorial Laplacians and isoperimetric inequality", *From local times to global geometry",* (Ed: K.D. Ellworthy). Pittman Research Notes in Mathematics **150** 68-74.

(Greenacre, 1984) M. Greenacre, *Theory and Application of Correspondence Analysis*, Academic Press, 1984.

(Jerrum & Sinclair, 1989) M. Jerrum, A.J. Sinclair, "Approximating the permanent", *SIAM J. Computing* **18** (1989) 1149-1178.

(Richards & Seary, 1997) W.D. Richards, A.J. Seary, "Convergence analysis of communication networks", *Organizational Communication: Emerging Perspectives* **5.**

(Ed: G. Barnett & L. Thayer), Ablex, Norwood NJ, 1997, 141-189.

(Runge, 1976) see (Cvetkovic, Doob & Sachs, 1980) p. 49.

(Seary & Richards, 1995) A.J. Seary, W.D. Richards, "Partitioning networks by eigenvectors", *Proc. International conference on Social Networks,* Vol. 1, London, 1995, 47-58.

(Watts & Strogatz, 1998) D.J. Watts, S.H. Strogatz, "Collective dynamics of small-world networks", *Nature*, **393**, 440-442 (1998).