Available online at www.sciencedirect.com

SciVerse ScienceDirect

Electronic Notes in Discrete Mathematics 38 (2011) 149–154

Electronic Notes in DISCRETE MATHEMATICS

www.elsevier.com/locate/endm

Spectra and structure of weighted graphs

Marianna Bolla [1,2]

Institute of Mathematics
Budapest University of Technology and Economics
Budapest, Hungary

Abstract

This article investigates relation between spectral and structural properties of large edge-weighted graphs. In social or biological networks we frequently look for partition of the vertices such that the induced subgraphs on them and the bipartite subgraphs between any pair of them exhibit regular behavior of information flow within or between the vertex subsets. We estimate the constants bounding the volume regularity of the cluster pairs by means of spectral gaps and classification properties of eigenvectors. We will focus on the more than two clusters case.

Keywords: Generalized random graphs, volume regularity, spectral clustering.

1 Preliminaries

Facing large networks, our purpose is to find some community structure in them, that is a partition of the vertices into clusters with homogeneous edge-densities within or between the clusters. For this purpose, the general framework of an edge-weighted graph will be used.

[1] Research supported by the Hungarian National Research Grants OTKA 76481 and OTKA-NKTH 77778.

[2] Email: marib@math.bme.hu

1571-0653/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.endm.2011.09.026

Let $G = (V, \mathbf{W})$ be a graph on n vertices, where the $n \times n$ symmetric *weight matrix* $\mathbf{W}$ has non-negative real entries and zero diagonal. The numbers $d_i = \sum_{j=1}^n w_{ij}$ $(i = 1, \dots, n)$ are the *generalized degrees* which constitute the main diagonal of the diagonal *degree matrix* $\mathbf{D}$. In [2], we investigated the spectral gap of the *normalized Laplacian* $\mathbf{L}_D = \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, where $\mathbf{I}$ denotes the identity matrix of appropriate size. As its spectrum is invariant under scaling the edge-weights, without loss of generality, $\sum_{i=1}^n \sum_{j=1}^n w_{ij} = 1$ will be supposed. The spectrum of $\mathbf{L}_D$ is in the [0,2] interval, and if G is connected ($\mathbf{W}$ is irreducible), 0 is a single eigenvalue with corresponding unit-norm eigenvector $\sqrt{\mathbf{d}} := (\sqrt{d_1}, \dots, \sqrt{d_n})^T$. For our convenience, we transform the normalized Laplacian into the so-called *normalized modularity matrix* defined by $\mathbf{B}_D = \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} - \sqrt{\mathbf{d}}\sqrt{\mathbf{d}}^T$. The spectrum of this matrix is in the [-1,1] interval (1 cannot be an eigenvalue if G is connected), and 0 is always an eigenvalue with eigenvector $\sqrt{\mathbf{d}}$. In fact, the so-called Expander Mixing Lemma of [6] can be generalized for edge-weighted graphs in terms of the spectral norm of $\mathbf{B}_D$. More generally, the eigenvalues of this matrix, separated from 0, together with the corresponding eigenvectors will play an important role in the community detection problem. To formulate this relation, weighted cuts, volumes, and volume regularity are defined.

The volume of $U \subset V$ is defined as $\mathtt{Vol}(U) = \sum_{i \in U} d_i$; further, for $X, Y \subset V$, $w(X, Y) = \sum_{i \in X} \sum_{j \in Y} w_{ij}$ is the weighted cut between X and Y. In this setup, the Expander Mixing Lemma is formulated in the following way: supposing $\mathtt{Vol}(V) = 1$, for all $X, Y \subset V$,

$$|w(X,Y) - \mathtt{Vol}(X)\mathtt{Vol}(Y)| \le \|\mathbf{B}_D\| \cdot \sqrt{\mathtt{Vol}(X)\mathtt{Vol}(Y)}.$$

As the spectral gap of G is $1 - \|\mathbf{B}_D\|$, a large spectral gap indicates a quasi-random property discussed in [5]. But what if there is a gap not at the ends of the spectrum? In this case we want to partition the vertices into clusters so that a relation similar to the above property for the edge-densities between the cluster pairs would hold. We will use a slightly modified version of the volume regularity's notion introduced by Alon, Coja-Oghlan, Han, Kang, Rödl, and Schacht [1].

Definition 1.1 Let $G = (V, \mathbf{W})$ be weighted graph with $\mathtt{Vol}(V) = 1$. The disjoint pair (A, B) is *α-volume regular* if for all $X \subset A$, $Y \subset B$ we have

$$|w(X,Y) - \rho(A,B)\mathtt{Vol}(X)\mathtt{Vol}(Y)| \le \alpha\sqrt{\mathtt{Vol}(A)\mathtt{Vol}(B)}, \tag{1}$$

where $\rho(A,B) = \frac{w(A,B)}{\mathtt{Vol}(A)\mathtt{Vol}(B)}$ is the relative inter-cluster density of (A, B).

In case of several clusters we will assign vectors, so-called representatives to the vertices. The $(k-1)$ dimensional vertex representatives $\mathbf{r}_1, \dots, \mathbf{r}_n$ are row vectors of the $n \times (k-1)$ matrix $\mathbf{X}$ of column vectors $\mathbf{D}^{-1/2}\mathbf{u}_i$, where $\mathbf{u}_i$'s are unit-norm eigenvectors belonging to $k-1$ so-called *structural eigenvalues* of $\mathbf{B}_D$ well separated from 0. The weighted k-variance of the $(k-1)$-dimensional vertex representatives is defined by

$$(2) \qquad S_k^2(\mathbf{X}) = \min_{P_k \in \mathcal{P}_k} S_k^2(P_k, \mathbf{X}) = \min_{P_k=(V_1,\dots,V_k)} \sum_{a=1}^{k} \sum_{j \in V_a} d_j \|\mathbf{r}_j - \mathbf{c}_a\|^2,$$

where $\mathbf{c}_a = \frac{1}{\mathtt{Vol}(V_a)} \sum_{j \in V_a} d_j \mathbf{r}_j$ is the weighted center of cluster V_a $(a = 1, \dots, k)$ and $\mathcal{P}_k$ denotes the family of k-partitions of the vertices.

To investigate the ideal k-cluster case, let us consider the following generalized random simple graph model: given the partition $(V_1, \dots, V_k)$ of V, vertices $i \in V_a$ and $j \in V_b$ are connected with probability p_{ab}, independently of each other, $1 \le a, b \le k$. We can think of the probability p_{ab} as the inter-cluster density of the pair (V_a, V_b). Since generalized random graphs can be viewed as edge-weighted graphs with a special block-structure burdened with random noise, based on [3], we are able to give the following spectral characterization of them. Fixing k, and tending with n to infinity in such a way that the cluster sizes grow at the same rate, there exists a positive number $\theta \le 1$, independent of n, such that for every $0 < \tau < 1/2$ there are exactly $k-1$ eigenvalues of $\mathbf{B}_D$ greater than $\theta - n^{-\tau}$, while all the others are at most $n^{-\tau}$ in absolute value; further, the k-variance of the vertex representatives constructed by the $k-1$ transformed structural eigenvectors is $\mathcal{O}(n^{-2\tau})$, and the cluster pairs are α-volume regular with any small α, almost surely.

Generalized quasi-random graphs were introduced by Lovász and T. Sós [7] as deterministic counterparts of the generalized random graphs with the same spectral properties. In fact, the authors define so-called generalized quasi-random graph sequences by means of graph convergence that also implies the convergence of spectra. Though, the spectrum itself does not carry enough information for the cluster structure of the graph, together with some classification properties of the structural eigenvectors it does.

For general deterministic edge-weighted graphs, our result is that the existence of $k-1$ eigenvalues of $\mathbf{B}_D$ separated from 0 by ε, is indication of a k-cluster structure, while the eigenvalues accumulating around 0 are responsible for the pairwise regularities. The clusters themselves can be recovered by applying the k-means algorithm for the vertex representatives obtained by the eigenvectors corresponding to the structural eigenvalues. We will focus on the case $k > 2$: Theorem 2.1 bounds the volume regularity's constants of the different cluster pairs by means of ε and the k-variance of the vertex representatives (based on the structural eigenvectors). We are also able to give estimates for the intra-cluster densities. As for the case $k = 2$, due to [2], the k-variance itself can be estimated by the spectral gap, and hence, the estimation simplifies.

2 Results

Theorem 2.1 *Let $G = (V, \mathbf{W})$ be an edge-weighted graph on n vertices, with generalized degrees $d_1, \dots, d_n$ and degree matrix $\mathbf{D}$. Suppose that $\mathtt{Vol}(V) = 1$ and there are no dominant vertices: $d_i = \Theta(1/n)$, $i = 1, \dots, n$, as $n \to \infty$. Let the eigenvalues of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, enumerated in decreasing absolute values, be*

$$1 = \rho_1 > |\rho_2| \geq \ldots \geq |\rho_k| > \varepsilon \geq |\rho_i|, \quad i \geq k+1.$$

The partition $(V_1, \dots, V_k)$ of V is defined so that it minimizes the weighted k-variance $S_k^2(\mathbf{X})$ of the vertex representatives – defined in (2) – obtained as row vectors of the $n \times (k-1)$ matrix $\mathbf{X}$ of column vectors $\mathbf{D}^{-1/2}\mathbf{u}_i$, where $\mathbf{u}_i$ is the unit-norm eigenvector belonging to ρ_i $(i = 2, \dots, k)$. Suppose that there is a constant $0 < K \leq \frac{1}{k}$ such that $|V_i| \geq Kn$, $i = 1, \dots, k$. With the notation $s^2 = S_k^2(\mathbf{X})$, the (V_i, V_j) pairs are $\mathcal{O}(\sqrt{2k}s + \varepsilon)$-volume regular $(i \neq j)$ and for the clusters V_i $(i = 1, \dots, k)$ the following holds: for all $X, Y \subset V_i$,

$$|w(X,Y) - \rho(V_i)\mathtt{Vol}(X)\mathtt{Vol}(Y)| = \mathcal{O}(\sqrt{2k}s + \varepsilon)\mathtt{Vol}(V_i),$$

where $\rho(V_i) = \frac{w(V_i,V_i)}{\mathtt{Vol}^2(V_i)}$ is the relative intra-cluster density of V_i.

Proof. Recall that the spectrum of $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$ differs from that of $\mathbf{B}_D$ only in the following: it contains the eigenvalue $\rho_1 = 1$ with corresponding eigenvector $\mathbf{u}_1 = \sqrt{\mathbf{d}}$ instead of an eigenvalue 0 of $\mathbf{B}_D$ with the same eigenvector. If G is connected ($\mathbf{W}$ is irreducible), 1 is a single eigenvalue. The $(k-1)$-dimensional representatives of the vertices are row vectors of the matrix $\mathbf{X} = (\mathbf{x}_2, \dots, \mathbf{x}_k)$, where $\mathbf{x}_i = \mathbf{D}^{-1/2}\mathbf{u}_i$ $(i = 2, \dots, k)$. The representatives can as well be regarded as k-dimensional ones, as by inserting the vector $\mathbf{x}_1 = \mathbf{D}^{-1/2}\mathbf{u}_1 = \mathbf{1}$ will not change the k-variance $s^2 = S_k^2(\mathbf{X})$. Suppose that the minimum k-variance is attained on the k-partition $(V_1, \dots, V_k)$ of the vertices. By an easy analysis of variance argument it follows that $s^2 = \sum_{i=1}^k \mathtt{dist}^2(\mathbf{u}_i, F)$, where $F = \mathtt{Span}\,\{\mathbf{D}^{1/2}\mathbf{z}_1, \dots, \mathbf{D}^{1/2}\mathbf{z}_k\}$ with the so-called normalized partition vectors $\mathbf{z}_1, \dots, \mathbf{z}_k$ of coordinates $z_{ji} = \frac{1}{\sqrt{\mathtt{Vol}(V_i)}}$ if $j \in V_i$ and 0, otherwise $(i = 1, \dots, k)$. Note that the vectors $\mathbf{D}^{1/2}\mathbf{z}_1, \dots, \mathbf{D}^{1/2}\mathbf{z}_k$

form an orthonormal system. By [2], we can find another orthonormal system $\mathbf{v}_1, \dots, \mathbf{v}_k \in F$ such that $s^2 \le \sum_{i=1}^k \|\mathbf{u}_i - \mathbf{v}_i\|^2 \le 2s^2$. We approximate the matrix $\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} = \sum_{i=1}^n \rho_i \mathbf{u}_i \mathbf{u}_i^T$ by the rank k matrix $\sum_{i=1}^k \rho_i \mathbf{v}_i \mathbf{v}_i^T$ with the following accuracy (in spectral norm):

$$(3) \quad \left\| \sum_{i=1}^n \rho_i \mathbf{u}_i \mathbf{u}_i^T - \sum_{i=1}^k \rho_i \mathbf{v}_i \mathbf{v}_i^T \right\| \le \sum_{i=1}^k |\rho_i| \cdot \left\| \mathbf{u}_i \mathbf{u}_i^T - \mathbf{v}_i \mathbf{v}_i^T \right\| + \left\| \sum_{i=k+1}^n \rho_i \mathbf{u}_i \mathbf{u}_i^T \right\|,$$

which can be estimated from above with $\sum_{i=1}^k \sin \alpha_i + \varepsilon \le \sum_{i=1}^k \|\mathbf{u}_i - \mathbf{v}_i\| + \varepsilon \le \sqrt{2k}s + \varepsilon$, where α_i is the angle between $\mathbf{u}_i$ and $\mathbf{v}_i$, and for it, $\sin \frac{\alpha_i}{2} = \frac{1}{2}\|\mathbf{u}_i - \mathbf{v}_i\|$ holds, $i = 1, \dots, k$.

Based on these considerations and relation between the cut norm and the spectral norm, the densities to be estimated in the defining formula (1) of volume regularity can be written in terms of stepwise constant vectors in the following way. The vectors $\mathbf{y}_i := \mathbf{D}^{-1/2}\mathbf{v}_i$ are stepwise constants on the partition $(V_1, \dots, V_k)$, $i = 1, \dots, k$. The matrix $\sum_{i=1}^k \rho_i \mathbf{y}_i \mathbf{y}_i^T$ is therefore a symmetric block-matrix on $k \times k$ blocks belonging to the above partition of the vertices. Let $\widetilde{w}_{ab}$ denote its entries in the (a, b) block $(a, b = 1, \dots, k)$. Using (3), the rank k approximation of the matrix $\mathbf{W}$ is performed with the following accuracy of the perturbation $\mathbf{E}$:

$$\|\mathbf{E}\| = \left\| \mathbf{W} - \mathbf{D}(\sum_{i=1}^k \rho_i \mathbf{y}_i \mathbf{y}_i^T)\mathbf{D} \right\| = \left\| \mathbf{D}^{1/2}(\mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2} - \sum_{i=1}^k \rho_i \mathbf{v}_i \mathbf{v}_i^T)\mathbf{D}^{1/2} \right\|.$$

Therefore, the entries of $\mathbf{W}$ – for $i \in V_a$, $j \in V_b$ – can be decomposed as $w_{ij} = d_i d_j \widetilde{w}_{ab} + \eta_{ij}$, where the cut norm of the $n \times n$ symmetric error matrix $\mathbf{E} = (\eta_{ij})$ restricted to $V_a \times V_b$ (otherwise it contains entries all zeroes) and denoted by $\mathbf{E}_{ab}$, is estimated as follows:

$$\|\mathbf{E}_{ab}\|_\Box \le c\sqrt{\mathtt{Vol}(V_a)}\sqrt{\mathtt{Vol}(V_b)}(\sqrt{2k}s + \varepsilon),$$

where the constant c does not depend on n. Consequently, for $a, b = 1, \dots, k$ and $X \subset V_a$, $Y \subset V_b$:

$$\begin{aligned} &|w(X,Y) - \rho(V_a, V_b)\mathtt{Vol}(X)\mathtt{Vol}(Y)| = \\ &\left| \sum_{i \in X} \sum_{j \in Y} (d_i d_j \widetilde{w}_{ab} + \eta_{ij}^{ab}) - \frac{\mathtt{Vol}(X)\mathtt{Vol}(Y)}{\mathtt{Vol}(V_a)\mathtt{Vol}(V_b)} \sum_{i \in V_a} \sum_{j \in V_b} (d_i d_j \widetilde{w}_{ab} + \eta_{ij}^{ab}) \right| = \\ &\left| \sum_{i \in X} \sum_{j \in Y} \eta_{ij}^{ab} - \frac{\mathtt{Vol}(X)\mathtt{Vol}(Y)}{\mathtt{Vol}(V_a)\mathtt{Vol}(V_b)} \sum_{i \in V_a} \sum_{j \in V_b} \eta_{ij}^{ab} \right| \le 2c(\sqrt{2k}s + \varepsilon)\sqrt{\mathtt{Vol}(V_a)\mathtt{Vol}(V_b)}, \end{aligned}$$

that gives the required statement both in the $a \neq b$ and $a = b$ case. □

Remark 2.2 The above theorem has only relevance if there is a remarkable spectral gap between $|\rho_k|$ and $|\rho_{k+1}|$. This is a necessary condition for s^2 to be "small". As it is not sufficient, the estimate is given in terms of s and ε, except the case $k = 2$, where s^2 can directly be estimated by the spectral gap. In this case we get that the pair (V_1, V_2) is $\mathcal{O}(\sqrt{\frac{1-\theta}{1-\varepsilon}})$-volume regular, where $\theta = |\rho_2|$, see [4].

Acknowledgement

The author wishes to thank Vera T. Sós, László Lovász, and Miklós Simonovits for their useful advices.

References

[1] Alon, N., Coja-Oghlan, A., Han, H., Kang, M., Rödl, V., and Schacht, M., *Quasi-randomness and algorithmic regularity for graphs with general degree distributions*, Siam J. Comput. **39** (6) (2010), 2336-2362.

[2] Bolla, M., and Tusnády, G., *Spectra and optimal partitions of weighted graphs*, Discrete Mathematics **128** (1994), 1-20.

[3] Bolla, M., *Noisy random graphs and their Laplacians*, Discrete Mathematics **308** (2008), 4221-4230.

[4] Bolla, M., *Beyond the expanders*, International Journal of Combinatorics, to appear.

[5] Chung, F., and Graham, R., *Quasi-random graphs with given degree sequences*, Random Structures and Algorithms **12** (2008), 1-19.

[6] Hoory, S., Linial, N., and Widgerson, A., *Expander graphs and their applications*, Bulletin (New series) of the American Mathematical Society **43** (4) (2006), 439-561.

[7] Lovász, L., and Sós, V. T., *Generalized quasirandom graphs*, J. Comb. Theory B **98** (2008), 146-163.