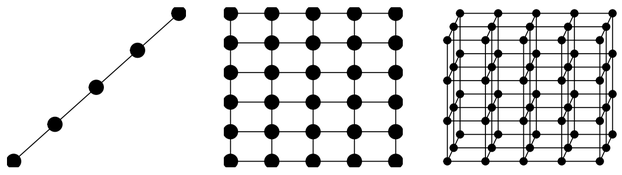

5th Polish Conference on Artificial Intelligence

Selected properties of Grid Graph Laplacians

Mieczysław Alojzy Kłopotek and Sławomir Tadeusz Wierzchoń

Institute of Computer Science, Polish Academy of Sciences, ul. Jana Kazimierza 5, 01-248 Warsaw, Poland, klopotek,stw@ipipan.waw.pl

19-21 April 2024

Abstract

Properties of the regular multidimensional grid graphs from the point of view of the analytical foundations for studying the clustering behavior of the GSA (Graph Spectral Analysis) clustering algorithms are investigated. From the theoretical point of view, the regular grid graph has no structure, and therefore the properties of spectra of the corresponding Laplacians may be indicative of a missing structure in the data. The range of eigenvalues is determined and it is shown that the eigenvalues are not distributed uniformly. This contradicts assumptions made in many GSA applications. The complexity of eigenvector elements distributions is shown.

Introduction

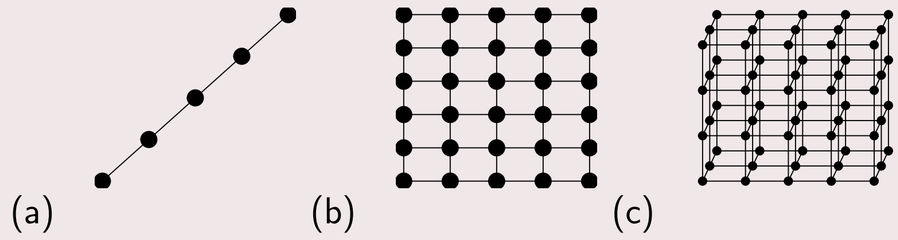

Figure: 1. Examples of grid graphs. (a) 1-dimensional grid graph, (b) 2-dimensional grid graph, (c) 3-dimensional grid graph.

Regular graph structures (Fig. 1) allow to study analytically Laplacian based graph algorithm (Notarstefano,2012). Applications include: (1) graph embedding without crossings on the spectrum of its Laplacian (Stankiewicz,2017), (2) gonality of curves via grid Laplacians (Cornelissen et al.,2015), (3) image processing (Cheung et al. ,2018), (4) chemistry studies (Merris, 1994). (5) mechanics (membrane vibration) (Cetkovic et al., 1980). (6) robotic swarms (Ramachandran and Berman, 2016).
We have investigated such graphs mainly from the point of view of properties of Graph Spectral Clustering algorithms. The study of regular grids is useful for the validation of some assumptions made for algorithms clustering via GSA large datasets, like Laplacian eigenvalue uniformity.

Method

Clustering via GSA: Be given a similarity matrix of objects $\boldsymbol{S}$. (1) Compute the combinatorial Laplacian as $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{S}$, where $\boldsymbol{D}$ is a diagonal matrix with components being row sums of $\boldsymbol{S}$. (2) Embed the objects in a low-dimensional Euclidean space spanned by lowest eigenvectors of $\boldsymbol{L}$. (3) Cluster in this embedding using e.g. $\boldsymbol{k}$-means algorithm. See e.g. (von Luxburg, 2007) for details.

Quest for Analytical Embedding: Define the path graph $\boldsymbol{P_n}$ as a tree with two nodes of vertex degree 1, and the other $\boldsymbol{n-2}$ nodes of vertex degree 2. A graph $\boldsymbol{G(n_1,\ldots,n_d)}$ is called a $\boldsymbol{d}$-dimensional grid if it is the Cartesian product of $\boldsymbol{d}$ paths $\boldsymbol{P_{n_1},\ldots,,,P_{n_d}}$, (Spie,2019). Such a graph can be embedded into an $\boldsymbol{d}$ dimensional Euclidean space with nodes having only consecutive integer coordinates (see Fig. 1). The nodes of a grid graph are linked with coordinates differing only by 1 at only one axis. Those graphs are of special interest because they can be used for practical implementation of parallel algorithms. In the similarity matrix, the entry is equal to 1, if there is a link between nodes.
We have shown that for such a regular grid structure, the combinatorial Laplacian eigenvalues $\boldsymbol{\lambda}$ and eigenvectors $\mathbf{v}$ are analytically given as:

$$\boldsymbol{\lambda}_{[z_1,\ldots,z_d]} = \sum_{j=1}^{d} \left(2\sin\left(\frac{\pi z_j}{2n_j}\right)\right)^2 \tag{1}$$

where, for each $\boldsymbol{j=1,\ldots,d}$, $\boldsymbol{z_j}$ is an integer such that $\boldsymbol{0 \le z_j \le n_j - 1}$.
Denoting $\boldsymbol{\lambda_{(j,z_j)}} = \left(2\sin\left(\frac{\pi z_j}{2n_j}\right)\right)^2$ the above formula can be rewritten as $\boldsymbol{\lambda_{[z_1,\ldots,z_d]}} = \sum_{j=1}^{d} \boldsymbol{\lambda_{(j,z_j)}}$. Define

$$\boldsymbol{\nu_{[z_1,\ldots,z_d],[x_1,\ldots,x_d]}} = \prod_{j=1}^{d} \cos\left(\frac{\pi z_j}{n_j}(x_j - 0.5)\right) \tag{2}$$

where, for each $\boldsymbol{j=1,\ldots,d}$, $\boldsymbol{x_j}$ is an integer such that $\boldsymbol{1 \le x_j \le n_j}$.
And finally define the $\boldsymbol{n}$ dimensional vector $\mathbf{v}_{[z_1,\ldots,z_d]}$ such that

$$\mathbf{v}_{[z_1,\ldots,z_d],i} = \boldsymbol{\nu}_{[z_1,\ldots,z_d],[x_1,\ldots,x_d]} \tag{3}$$

Theorem

If $\boldsymbol{z_j \in [-n_j+1,-1]}$, *then* $\mathbf{v}_{[z_1,\ldots,z_j,\ldots,z_d]} = \mathbf{v}_{[z_1,\ldots,z'_j,\ldots,z_d]}$ *where* $\boldsymbol{z'_j \in [0,n_j-1]}$, *and* $\boldsymbol{z'_j = -z_j}$. *If* $\boldsymbol{z_j = n_j}$, *then* $\mathbf{v}_{[z_1,\ldots,z_j,\ldots,z_d]} = \mathbf{0}$. *If* $\boldsymbol{z_j \in [n_j+1,2n_j-1]}$, *then* $\mathbf{v}_{[z_1,\ldots,z_j,\ldots,z_d]} = -\mathbf{v}_{[z_1,\ldots,z'_j,\ldots,z_d]}$ *where* $\boldsymbol{z'_j \in [0,n_j-1]}$, *and* $\boldsymbol{z'_j = 2n_j - z_j}$.

Results

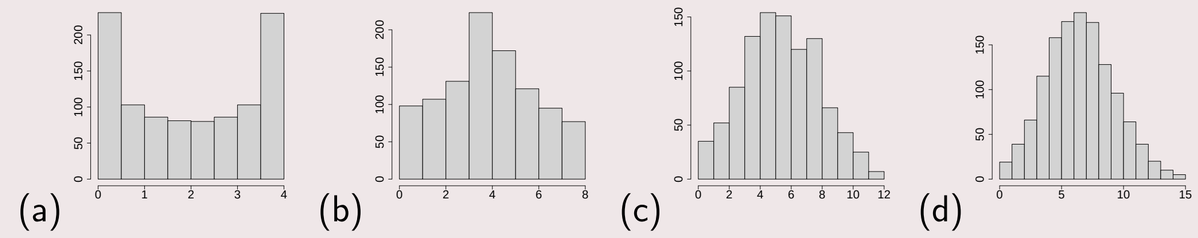

Figure: 2. The histograms of eigenvalues of combinatorial Laplacians of grid graphs of approximately 1,000 nodes. (a) 1-dimensional grid graph, (b) 2-dimensional grid graph, (c) 3-dimensional grid graph, (d) 4-dimensional grid graph.

Fig. 2 shows: (1) the distribution of eigenvalues differs significantly from uniform distribution, contrary to common assumptions, (2) one-dimensional grid distribution differs significantly in shape from higher dimensions. Same in Fig. 3.

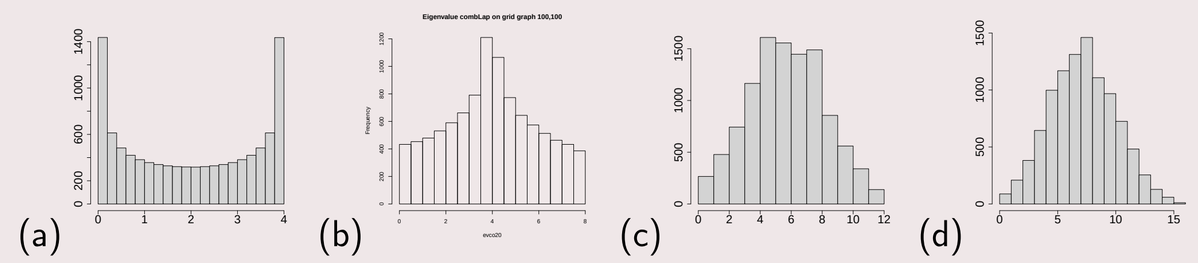

Figure: 3. The histograms of eigenvalues of combinatorial Laplacians of grid graphs of approximately 10,000 nodes. (a) 1-dimensional grid graph, (b) 2-dimensional grid graph, (c) 3-dimensional grid graph, (d) 4-dimensional grid graph.

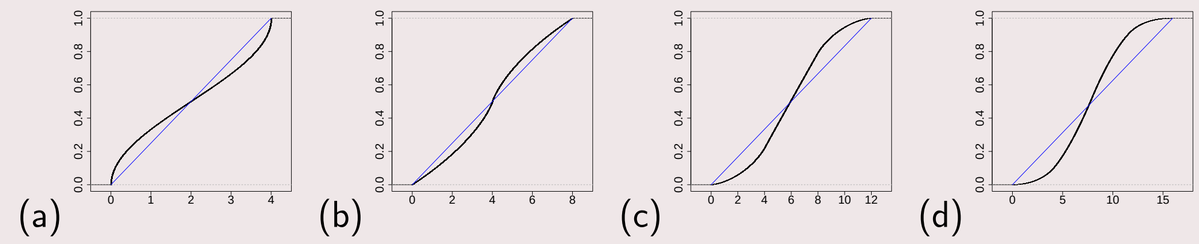

Figure: 4. The in the limit cumulative distribution function of eigenvalues of combinatorial Laplacians of grid graphs. (a) 1-dimensional grid, (b) 2-dimensional grid, (c) 3-dimensional grid, (d) 4-dimensional grid graph. Blue line – the uniform distribution function

Cumulative distribution functions of eigenvalues of combinatorial Laplacians of 1, 2, 3, and 4-dimensional grid graphs with the number of nodes "in the limit" are depicted in Figure 4. The multidimensional grid graph exhibits no similarity to *uniform eigenvalue distribution* though it is structureless.

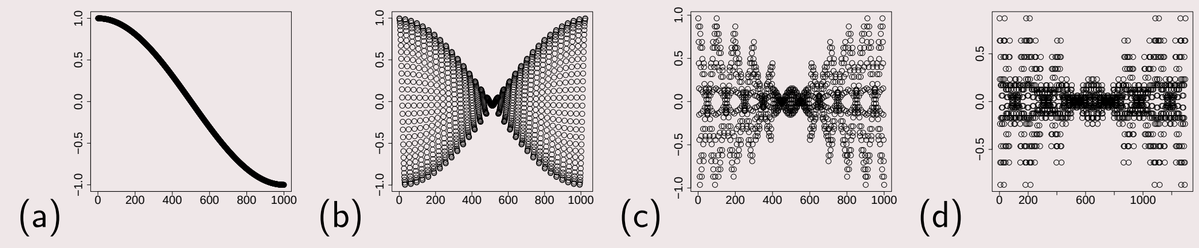

Figure: 5. The plots of sample eigenvectors of combinatorial Laplacians of grid graphs of approximately 1,000 nodes. (a) 1-dimensional grid graph, $\mathbf{z}=[1]$, (b) 2-dimensional grid graph, $\mathbf{z}=[1,1]$, (c) 3-dimensional grid graph, $\mathbf{z}=[1,1,1]$, (d) 4-dimensional grid graph, $\mathbf{z}=[1,1,1,1]$. X coordinates: identifiers of nodes.

Figure 5 visualizes the shapes of sample eigenvectors of the grids elaborated on in Fig.2. A clear cosine shape can be recognized for a 1-dimensional graph. One recognizes the cosine product for two-dimensional grids. The patterns are not so easily classified by eye inspection in higher dimensions. This behavior *questions assumptions* of some GSA on *piece-wise linearity of eigenvector entries*.

Conclusions

Shapes of eigenvalue spectra seem to be kept with growing number of nodes and they are in no way uniform. Such a study should be a clear warning sign for those who develop approximate procedures for clustering large scale graphs, like CSC (Tremblay,2016), as the common assumption of uniformity is not valid.
Entries in eigenvectors form a clear cosine line for 1-dimensional graph, but shapes in higher dimensions do not seem to be obvious. This should be a concern when utilizing GSA methods for large graphs like Nyström like approaches, (Li,2015).