Discrete Applied Mathematics 190–191 (2015) 56–74

Contents lists available at ScienceDirect

Discrete Applied Mathematics

journal homepage: www.elsevier.com/locate/dam

Spectral distances on graphs

Jiao Gu [a,b,c], Bobo Hua [c,d], Shiping Liu [c,e,*]

[a] School of Science, Jiangnan University, 214122, Wuxi, China
[b] Mathematics and Science college, Shanghai Normal University, 200234, Shanghai, China
[c] Max Planck Institute for Mathematics in the Sciences, 04103, Leipzig, Germany
[d] School of Mathematical Sciences, LMNS, Fudan University, Shanghai 200433, China
[e] Department of Mathematical Sciences, Durham University, DH1 3LE Durham, United Kingdom

ARTICLE INFO

Article history:
Received 28 October 2014
Received in revised form 20 March 2015
Accepted 16 April 2015
Available online 8 May 2015

Keywords:
Wasserstein distance
Spectral measure
Random rooted graph
Asymptotic behavior
Biological networks

ABSTRACT

By assigning a probability measure via the spectrum of the normalized Laplacian to each graph and using L^p Wasserstein distances between probability measures, we define the corresponding spectral distances d_p on the set of all graphs. This approach can even be extended to measuring the distances between infinite graphs. We prove that the diameter of the set of graphs, as a pseudo-metric space equipped with d_1, is one. We further study the behavior of d_1 when the size of graphs tends to infinity by interlacing inequalities aiming at exploring large real networks. A monotonic relation between d_1 and the evolutionary distance of biological networks is observed in simulations.

1. Introduction

One major interest in graph theory is to explore the differences of graphs in structure, that is, in the sense of graph isomorphism. In computational complexity theory, the subgraph isomorphism problem, like many combinational problems in graph theory, is NP hard. Therefore, a method that gives a quick and easy estimate of the difference between two graphs is desirable [34]. As we know, all the topological information of a graph can be found in its adjacency matrix. The spectral graph theory studies the relationship between the properties of graphs and the spectra of their representing matrices, such as adjacency matrices and Laplace matrices [14,18,17]. In particular, some important topological information of a graph can be extracted from its specific eigenvalue like the first or the largest one, see e.g. [18,17,39,11,25,12,10]. The approach of reading information from the entire spectrum of a graph was explored in [5–7,30,32] etc. In spite of the existence of co-spectral graphs (see [38, Chapter 3] for a general construction and the references therein), the spectra of graphs can support us one way on exploring problems that involve (sub-)graph isomorphism by the fast computation algorithms and the close relationship with the structure of graphs.

A spectral distance on the set of finite graphs of the same size, i.e. the same number of vertices, was suggested in a problem of Richard Brualdi in [37] to explore the so-called cospectrality of a graph. It was further studied in [26] using the spectra of adjacency matrices. Employing certain Gaussian measures associated to the spectra of normalized Laplacians and the corresponding L^1 distances, the first named author, Jost, the third named author and Stadler [21,20] explored a spectral distance well-defined on the set of all finite graphs without any constraint about sizes. In this paper, instead of

* Corresponding author at: Department of Mathematical Sciences, Durham University, DH1 3LE Durham, United Kingdom.
E-mail addresses: jiaogu@mis.mpg.de (J. Gu), bobohua@mis.mpg.de (B. Hua), shiping.liu@durham.ac.uk (S. Liu).

http://dx.doi.org/10.1016/j.dam.2015.04.011

the Gaussian measures, we assign Dirac measures to graphs through the spectra of normalized Laplacians and use the Wasserstein distances between probability measures to propose spectral distances between graphs. In fact, this notion of spectral distances provides a metrization of the notion of spectral classes of graphs introduced in [21] via the weak convergence of the corresponding Dirac measures. The spectral class can be considered as a weak notion of graph limits (see the concepts of graphon, graphing and related theories in the monograph of Lovász [33]). This notion of spectral distances is even adaptable for weighted infinite graphs. And we can prove diameter estimates with respect to these distances, which are sharp for certain cases.

A weighted graph G is a triple (V, E, θ) where V is the set of vertices, E is the set of edges and $\theta : E \to (0, \infty), (x, y) \mapsto \theta_{xy}$, is the (symmetric) edge weight function. We write $x \sim y$ if $(x, y) \in E$. We assume that for any vertex x, the weighted degree defined by $\theta_x := \sum_{y\sim x} \theta_{xy}$ is finite and $\theta_{xx} = 0$ (i.e. there is no self-loops).

Let us first consider finite weighted graphs. The normalized Laplacian of $G = (V, E, \theta)$ is defined as, for any function $f : V \to \mathbb{R}$ and any $x \in V$,

$$\Delta_G f(x) = f(x) - \frac{1}{\theta_x}\sum_{y\sim x} f(y)\theta_{xy}. \tag{1}$$

This operator can be extended to an infinite weighted graph which has countable vertex set V but is not necessarily locally finite (see [27] or Section 2 below). As a matrix, Δ_G is unitarily equivalent to the Laplace matrix studied in [17].

If $x \in V$ is an isolated vertex, i.e. $\theta_x = 0$, (1) reads as $\Delta_G f(x) = f(x)$. This implies that an isolated vertex contributes an eigenvalue 1 to the spectrum of Δ_G, denoted by $\sigma(G)$. In this way, by the absence of the self-loops, the spectrum of any finite weighted graph $\sigma(G) = \{\lambda_i\}_{i=1}^N$, counting the multiplicity, satisfies the trace condition

$$\sum_{i=1}^{N} \lambda_i = N \tag{2}$$

where $N = |V|$. It is well-known that $\sigma(G)$ is contained in $[0, 2]$. We associate to $\sigma(G)$ a probability measure on $[0, 2]$ as follows:

$$\mu_{\sigma(G)} := \frac{1}{N}\sum_i \delta_{\lambda_i}, \tag{3}$$

where δ_{λ_i} is the Dirac measure concentrated on λ_i. We call $\mu_{\sigma(G)}$ the *spectral measure* for a finite weighted graph. (This is known as the empirical distribution of the eigenvalues in random matrix theory.) Denote by $P([0, 2])$ the set of probability measures on the interval $[0, 2]$. For any $\mu \in P([0, 2])$, the first moment of μ is defined as $m_1(\mu) := \int_{[0,2]} \lambda\, d\mu(\lambda)$. The trace condition (2) is then translated to

$$m_1(\mu_{\sigma(G)}) = 1. \tag{4}$$

This is a key property of the spectral measures for our further investigations.

Let d_p^W $(1 \le p < \infty)$ be the pth Wasserstein distance on $P([0, 2])$. That is, for any $\mu, \nu \in P([0, 2])$ (see e.g. [40]),

$$d_p^W(\mu, \nu) := \left(\inf_{\pi\in\Pi(\mu,\nu)} \int_{[0,2]\times[0,2]} d(x, y)^p d\pi(x, y)\right)^{1/p},$$

where $\Pi(\mu, \nu)$ denotes the collection of all measures on $[0, 2] \times [0, 2]$ with marginals μ and ν on the first and second factors respectively, i.e. $\pi \in \Pi(\mu, \nu)$ if and only if $\pi(A \times [0, 2]) = \mu(A)$ and $\pi([0, 2] \times B) = \nu(B)$ for all Borel subsets $A, B \subseteq [0, 2]$.

It is well-known that $(P([0, 2]), d_p^W)$ is a complete metric space for $p \in [1, \infty)$ which induces the weak topology of measures in $P([0, 2])$(see e.g. [40, Theorem 6.9]).

One can prove that $\mathrm{diam}(P([0, 2]), d_p^W) = 2$. Indeed, on one hand, for any $\mu, \nu \in P([0, 2])$ by the optimal transport interpretation of Wasserstein distance, $d_p^W(\mu, \nu) \le 2$. On the other hand, $d_p^W(\delta_0, \delta_2) = 2$. (Recall that δ_0, δ_2 are the Dirac measures concentrated on 0, 2, respectively.)

Definition 1.1. Given two finite weighted graphs $G = (V, E, \theta)$ and $G' = (V', E', \theta')$, the *spectral distance* between G and G' is defined as

$$d_p(G, G') := d_p^W(\mu_{\sigma(G)}, \mu_{\sigma(G')}). \tag{5}$$

We denote by $\mathcal{FG}$ the space of all finite weighted graphs. Then for any $1 \le p < \infty$, $(\mathcal{FG}, d_p)$ is a pseudo-metric space. This is not a metric space due to the existence of co-spectral graphs. However, in applications this spectral consideration leads to the simplification of measuring the discrepancy of graphs.