Journal of Computational Mathematics
Vol.41, No.4, 2023, 741–770.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2201-m2021-0130

GRAPH SPARSIFICATION BY UNIVERSAL GREEDY ALGORITHMS*

Ming-Jun Lai
Department of Mathematics, University of Georgia, Athens, GA 30602, USA
Email: mjlai@uga.edu
Jiaxin Xie[1)]
LMIB of the Ministry of Education, School of Mathematical Sciences, Beihang University, Beijing 100191, China
Email: xiejx@buaa.edu.cn
Zhiqiang Xu
LSEC, Inst. Comp. Math., Academy of Mathematics and System Science, Chinese Academy of Sciences, Beijing 100091, China;
School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
Email: xuzq@lsec.cc.ac.cn

Abstract

Graph sparsification is to approximate an arbitrary graph by a sparse graph and is useful in many applications, such as simplification of social networks, least squares problems, and numerical solution of symmetric positive definite linear systems. In this paper, inspired by the well-known sparse signal recovery algorithm called orthogonal matching pursuit (OMP), we introduce a deterministic, greedy edge selection algorithm, which is called the universal greedy approach (UGA) for the graph sparsification problem. For a general spectral sparsification problem, e.g., the positive subset selection problem from a set of m vectors in $\mathbb{R}^n$, we propose a nonnegative UGA algorithm which needs $O(mn^2 + n^3/\epsilon^2)$ time to find a $\frac{1+\epsilon/\beta}{1-\epsilon/\beta}$-spectral sparsifier with positive coefficients with sparsity at most $\lceil \frac{n}{\epsilon^2} \rceil$, where β is the ratio between the smallest length and largest length of the vectors. The convergence of the nonnegative UGA algorithm is established. For the graph sparsification problem, another UGA algorithm is proposed which can output a $\frac{1+O(\epsilon)}{1-O(\epsilon)}$-spectral sparsifier with $\lceil \frac{n}{\epsilon^2} \rceil$ edges in $O(m+n^2/\epsilon^2)$ time from a graph with m edges and n vertices under some mild assumptions. This is a linear time algorithm in terms of the number of edges that the community of graph sparsification is looking for. The best result in the literature to the knowledge of the authors is the existence of a deterministic algorithm which is almost linear, i.e. $O(m^{1+o(1)})$ for some $o(1) = O(\frac{(\log\log(m))^{2/3}}{\log^{1/3}(m)})$. Finally, extensive experimental results, including applications to graph clustering and least squares regression, show the effectiveness of proposed approaches.

Mathematics subject classification: 68W25, 05C50.
Key words: Spectral sparsification, Subset selection, Greedy algorithms, Graph clustering, Linear sketching.

1. Introduction

Graph sparsification aims to find a sparse subgraph from a dense graph G with n vertices and m edges (typically $m \gg n$) so that the sparsified subgraph can serve as a proxy for G in

* Received May 2, 2021 / Revised version received November 8, 2021 / Accepted January 27, 2022 / Published online March 29, 2023 /
1) Corresponding author

numerical computations for graph-based applications. In [4], Batson, Spielman, and Srivastava showed that for any undirected graph G one can find a sparse graph (sparsifier) whose graph Laplacian matrix can well preserve the spectrum of the original graph Laplacian matrix. Such a spectral graph sparsification plays increasingly important roles in many applications areas in mathematics and computer science [24, 30, 38]. A related research, known as *Laplacian Paradigm*, is illustrated as an emerging paradigm for the design of scalable algorithms in recent years. We refer the reader to [35,39,42] for excellent surveys on its background and applications.

Mathematically, we can state the graph sparsification problem as follows. Consider an undirected and weighted graph $G = (V, E, \mathbf{w})$, where V is a set of vertices, E is a set of edges, and $\mathbf{w}$ is a weight function that assigns a positive weight to each edge. The Laplacian matrix of the graph G is defined by

$$L_G = \sum_{(u,v)\in E} w_{(u,v)}(\mathbf{e}_u - \mathbf{e}_v)(\mathbf{e}_u - \mathbf{e}_v)^\top,$$

where $w_{(u,v)} \geq 0$ is the weight of edge (u, v) and $\mathbf{e}_u \in \mathbb{R}^{|V|}$ is the characteristic vector of vertex u (with a 1 on coordinated u and zeros elsewhere). In other words, for any $\mathbf{x} \in \mathbb{R}^n$,

$$\mathbf{x}^\top L_G \mathbf{x} = \sum_{(u,v)\in E} w_{(u,v)}\big(\mathbf{x}(u) - \mathbf{x}(v)\big)^2 \geq 0.$$

That is, L_G is positive semidefinite. *Spectral graph sparsification* is the process of approximating the graph G by a sparse (linear-sized) graph $H = (V, \tilde{E}, \tilde{\mathbf{w}})$ such that

$$a\mathbf{x}^\top L_G \mathbf{x} \leq \mathbf{x}^\top L_H \mathbf{x} \leq b\mathbf{x}^\top L_G \mathbf{x} \tag{1.1}$$

for all $\mathbf{x} \in \mathbb{R}^{|V|}$, where $b \geq a > 0$. Setting $\kappa := b/a$, H is called a κ-approximation of G or a κ-sparsifier of G. Actually, if we restrict the inequality in (1.1) only for all $\mathbf{x} \in \{0, 1\}^{|V|}$, one can obtain the cut sparsification [6]. Batson, Spielman and Srivastava [4, Theorem 1.1] proved that for every weighted graph G and every $\epsilon \in (0, 1)$ there exists a weighted graph H with at most $\lceil (n-1)/\epsilon^2 \rceil$ edges which is a $\frac{(1+\epsilon)^2}{(1-\epsilon)^2}$-approximation of L_G. More generally, let $V = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subset \mathbb{R}^n$ be a collection of vectors with $m \gg n$. We replace L_G by

$$B = \sum_{i=1}^m \mathbf{v}_i \mathbf{v}_i^\top, \tag{1.2}$$

which is clearly positive semidefinite. Let $\mathbb{R}^m_+$ denote the nonnegative orthant in $\mathbb{R}^m$ and let $\|\mathbf{s}\|_0$ denote the number of nonzero entries of the vector $\mathbf{s}$. In [4, Theorem 1.2], the authors proved that for any $\epsilon \in (0, 1)$, these exists an $\mathbf{s} = (s_1, \ldots, s_m) \in \mathbb{R}^m_+$ with $\|\mathbf{s}\|_0 \leq \lceil \mathrm{rank}(B)/\epsilon^2 \rceil$ such that

$$(1-\epsilon)^2 B \preceq \sum_{i=1}^m s_i \mathbf{v}_i \mathbf{v}_i^\top \preceq (1+\epsilon)^2 B. \tag{1.3}$$

Our aim is to use the ideas of the well-known orthogonal matching pursuit (OMP) ([33,40]) to study the spectral sparsification problem. We first focus on the following minimization problem:

$$\min_{\mathbf{s}\in\mathbb{R}^m, \mathbf{s}\geq 0} \|\mathbf{s}\|_0 \quad \text{s.t.} \quad (1-\epsilon)^2 B \preceq \sum_{i=1}^m s_i \mathbf{v}_i \mathbf{v}_i^\top \preceq (1+\epsilon)^2 B, \tag{1.4}$$

where $B = \sum_{i=1}^m \mathbf{v}_i \mathbf{v}_i^\top$ and $\{\mathbf{v}_1 \ldots, \mathbf{v}_m\} \subset \mathbb{R}^n$.