Journal of Computational Mathematics
Vol.33, No.5, 2015, 443–467.

http://www.global-sci.org/jcm
doi:10.4208/jcm.1506-m2014-0164

THE 1-LAPLACIAN CHEEGER CUT: THEORY AND ALGORITHMS*

K.C. Chang, Sihong Shao and Dong Zhang
LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China
Email: kcchang@math.pku.edu.cn, sihong@math.pku.edu.cn, dongzhang@pku.edu.cn

Abstract

This paper presents a detailed review of both theory and algorithms for the Cheeger cut based on the graph 1-Laplacian. In virtue of the cell structure of the feasible set, we propose a cell descend (CD) framework for achieving the Cheeger cut. While plugging the relaxation to guarantee the decrease of the objective value in the feasible set, from which both the inverse power (IP) method and the steepest descent (SD) method can also be recovered, we are able to get two specified CD methods. Comparisons of all these methods are conducted on several typical graphs.

Mathematics subject classification: 05C85, 05C50, 05C35, 58C40, 90C27, 35P30, 05C75, 90C26, 90C35
Key words: Spectral graph theory, Spectral clustering, 1-Laplace operator, Graph Laplacian, Eigenvalue problems, Cheeger constant, Graph cut, Optimization, Convergence

1. Introduction

Graph cut, partitioning the vertices of a graph into two or more disjoint subsets, is a fundamental problem in graph theory [1]. It is a very powerful tool in data clustering with wide applications ranging from statistics, computer learning, image processing, biology to social sciences [2]. There exist several kinds of balanced graph cut [3–5]. The Cheeger cut [6], which has recently been shown to provide excellent classification results [7–9], is one of them and its definition is as follows. Let $G = (V, E)$ denote a undirected and unweighted graph with vertex set $V = \{1, 2, \cdots, n\}$ and edge set E. Each edge $e \in E$ is a pair of vertices $\{i, j\}$. For any vertex i, the degree of i, denoted by d_i, is defined to be the number of edges passing through i. Let S and T be two nonempty subsets of V and use

$$E(S,T) = \Big\{\{i,j\} \in E : i \in S, j \in T\Big\}$$

to denote the set of edges between S and T. The edge boundary of S is $\partial S = E(S, S^c)$ (S^c is the complement of S in V) and the volume of S is defined to be $\mathrm{vol}(S) := \sum_{i\in S} d_i$. The number

$$h(G) = \min_{S\subset V, S\notin\{\emptyset, V\}} \frac{|\partial S|}{\min\{\mathrm{vol}(S), \mathrm{vol}(S^c)\}}$$

is called the *Cheeger constant*, and a partition (S, S^c) of V is called a *Cheeger cut* of G if

$$\frac{|\partial S|}{\min\{\mathrm{vol}(S), \mathrm{vol}(S^c)\}} = h(G),$$

* Received December 22, 2014 / Revised version received April 29, 2015 / Accepted June 29, 2015 / Published online September 18, 2015 /

where $|\partial S|$ is the cardinality of the set ∂S.

However, solving analytically the Cheeger cut problem is combinatorially NP-hard [7, 8]. Approximate solutions are required. The most well-known approach to approximate the Cheeger cut solutions is the spectral clustering method, which relaxes the original discrete combination optimization problem into a continuous function optimization problem through the graph Laplacian [10]. The (normalized) standard graph Laplacian (i.e. the 2-Laplacian) is defined as $L = I - D^{-1/2}AD^{-1/2}$, where I is the identity matrix, $D = \mathrm{diag}(d_1, \cdots, d_n)$ is a diagonal matrix and A the adjacency matrix of G. According to the linear spectral graph theory, the eigenvalues of L satisfy $0 = \lambda_1 \leq \cdots \leq \lambda_n \leq 2$ and the second eigenvalue λ_2 can be used to bound the Cheeger constant as follows

$$\frac{\lambda_2}{2} \leq h(G) \leq \sqrt{2\lambda_2}, \tag{1.1}$$

which is nothing but the Cheeger inequality [1]. Furthermore, the corresponding second eigenvector is also used to approximate the Cheeger cut, i.e. the 2-spectral clustering or the ℓ^2 relaxation. It should be noted that this second eigenvector is not the Cheeger cut, but only an approximation [10].

In order to achieve a better cut than the ℓ^2 relaxation, a spectral clustering based on the graph p-Laplacian defined by

$$(\Delta_p \boldsymbol{x})_i = \sum_{j \sim i} |x_i - x_j|^{p-1} \,\mathrm{sign}(x_i - x_j)$$

with small $p \in (1, 2)$ was proposed, in view of the fact that the cut by threshold the second eigenvector of the graph p-Laplacian tends to the Cheeger cut as $p \to 1^+$ [11]. Here $j \sim i$ denotes vertex j is adjacent to vertex i, $\sum_{j \sim i}$ means the summation is with respect to all vertices adjacent to vertex i, and $\mathrm{sign}(t)$ is the standard sign function which equals to 1 if $t > 0$, 0 if $t = 0$, and -1 if $t < 0$. The resulting ℓ^p relaxation

$$\frac{\sum\limits_{i \sim j} \big|x_i - x_j\big|^p}{\sum\limits_{i=1}^{n} d_i \big|x_i\big|^p} \tag{1.2}$$

is differentiable but nonconvex, so that standard Newton-like methods can be applied, but only local minimizers are obtained. In actual calculations, multiple runs with random initializations are taken to approximate the global minimizer.

All above mentioned p-spectral clustering for any $p \in (1, 2]$ are *indirect* methods.

Since the second (the first non-zero) eigenvalue of 1-Laplacian (see Definition 1.1) for connected graphs equals to the Cheeger constant, and the corresponding eigenvectors provide exact solutions of the Cheeger cut problem [7, 12]. We study the numerical solution of the second eigenvector of the graph 1-Laplacian.

However, the ℓ^1 nonlinear eigenvalue problem (the corresponding object function is obtained by setting $p = 1$ in Eq. (1.2)) is not only nonconvex but also nondifferentiable. Three types of algorithms have been proposed to minimize the 1-spectral clustering problem. They are: the Split-Bregman like ratio minimization algorithm [7], the inverse power (IP) method [8], and the steepest descent (SD) algorithm [13]. Unfortunately, all these methods fail to give global minimizers.