ARTICLE IN PRESS

Discrete Optimization ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Optimization

www.elsevier.com/locate/disopt

Projection results for the k-partition problem

Jamie Fairbrother[a], Adam N. Letchford[b,*]

[a] STOR-i Centre for Doctoral Training, Lancaster University, Lancaster LA1 4YF, United Kingdom
[b] Department of Management Science, Lancaster University Management School, Lancaster LA1 4YX, United Kingdom

ARTICLE INFO

Article history:
Received 22 February 2016
Received in revised form 9 August 2017
Accepted 11 August 2017
Available online xxxx

Keywords:
Graph partitioning
Polyhedral combinatorics
Branch-and-cut
Semidefinite programming

ABSTRACT

The *k-partition problem* is an $\mathcal{NP}$-hard combinatorial optimisation problem with many applications. Chopra and Rao introduced two integer programming formulations of this problem, one having both node and edge variables, and the other having only edge variables. We show that, if we take the polytopes associated with the 'edge-only' formulation, and project them into a suitable subspace, we obtain the polytopes associated with the 'node-and-edge' formulation. This result enables us to derive new valid inequalities and separation algorithms, and also to shed new light on certain SDP relaxations. Computational results are also presented.

1. Introduction

The *k-partition problem* (k-PP) is a strongly $\mathcal{NP}$-hard combinatorial optimisation problem, first defined in [1]. We are given an undirected graph G, with vertex set V and edge set E, a rational weight w_e for each edge $e \in E$, and an integer k with $2 \leq k \leq |V|$. The task is to partition V into k or fewer subsets (called "clusters"), such that the sum of the weights of the edges that have both end-vertices in the same cluster is minimised. The k-PP has applications in statistical clustering, numerical linear algebra, telecommunications, VLSI layout, sports team scheduling and statistical physics (see, e.g., [2–4]).

Note that the k-PP is equivalent to the *max-k-cut* problem, in which one wishes to maximise the sum of the weights of the edges that have *exactly one* end-vertex in the same cluster (see [5]). In particular, when $k = 2$, we have the well-known *max-cut* problem, which is known to be strongly $\mathcal{NP}$-hard (see [6]). Moreover, the problem of checking whether G is k-colourable can be reduced to the k-PP. Thus, the k-PP is strongly $\mathcal{NP}$-hard for all fixed k, and this is so even when $k = 3$ and G is planar (see again [6]). Not only that, but the special case of the k-PP in which G is a complete graph and $k = |V|$, called the *clique partitioning problem* (CPP), is strongly $\mathcal{NP}$-hard as well [7].

* Corresponding author.
E-mail addresses: j.fairbrother@lancaster.ac.uk (J. Fairbrother), A.N.Letchford@lancaster.ac.uk (A.N. Letchford).

http://dx.doi.org/10.1016/j.disopt.2017.08.001
1572-5286/

Please cite this article in press as: J. Fairbrother, A.N. Letchford, Projection results for the k-partition problem, Discrete Optimization (2017), http://dx.doi.org/10.1016/j.disopt.2017.08.001.

In their seminal paper [8], Chopra and Rao presented two different *0-1 linear programming* (0-1 LP) formulations of the k-PP. One of these formulations has both node and edge variables, whereas the other has only edge variables. For each formulation, several families of strong valid linear inequalities (a.k.a. cutting planes) have been discovered (e.g., [7–10]). For some of these families, we also have efficient separation algorithms (e.g., [11–14]). There is also a parallel literature concerned with *semidefinite programming* (SDP) relaxations of the k-PP (e.g., [2–5,15–17]).

The main result in this paper is the following. If we take the polytopes associated with the 'edge-only' formulations, and project them into a suitable subspace, we obtain the polytopes associated with the 'node-and-edge' formulations. Although this result is fairly easy to derive, it is very useful. Specifically, it leads to new valid inequalities and separation algorithms for the 'node-and-edge' formulations, and it sheds new light on the SDP relaxations given in [4,5,15].

The paper is structured as follows. A literature review is given in Section 2. The projection results are given in Section 3. The new inequalities and separation algorithms, along with our remarks on SDP relaxations, are presented in Section 4. Some computational results are given in Section 5. Finally, some concluding remarks are given in Section 6.

We use the following (standard) notation throughout the paper. The number of nodes and edges in G is denoted by n and m, respectively. For a given positive integer p, we let K_p denote the complete graph on p nodes. Its vertex set is $\{1, \dots, p\}$, which we denote by V_p. Its edge set is denoted by E_p. We also let I_p, e_p and J_p denote the identity matrix of order p, the all-ones vector with p components, and the (square) all-ones matrix of order p, respectively. Given a real symmetric matrix M, we write $M \succeq 0$ to indicate that M is positive semidefinite (psd).

We also use the following (again standard) terminology. A *clique* is a set of pairwise adjacent nodes. A set $C \subseteq E$ is a *cycle* if it induces a connected subgraph of G in which every node has degree 2. The nodes in the subgraph are denoted by $V(C)$. Two disjoint sets $R, S \subset E$ form a *wheel* if R is a cycle and there exists a node $h \in V \setminus V(R)$ such that $S = \big\{\{v,h\} : v \in V(R)\big\}$. (The set R is called the *rim*, the edges in S are called *spokes*, and h is called the *hub*.) Two disjoint sets $R, S \subset E$ and an edge $\{h, h'\} \in E \setminus R$ form a *bicycle wheel* if R is a cycle and $S = \big\{\{v,h\}, \{v,h'\} : v \in V(R)\big\}$.

2. Literature review

We now review the relevant literature. We cover the two 0-1 LP formulations in Sections 2.1 and 2.2, separation algorithms in Section 2.3, and SDP relaxations in Section 2.4.

2.1. The node-and-edge formulation

Chopra & Rao [8] present the following 0-1 LP formulation of the k-PP, which has $nk + m$ variables. For each $v \in V$ and for $c = 1, \dots, k$, let x_{vc} be a binary variable, indicating whether node v lies in the cth cluster. Also, for each $e \in E$, let y_e be a binary variable, indicating whether the end-nodes of e lie in the same cluster. Then:

$$\min \quad \sum_{e \in E} w_e y_e \tag{1}$$

$$\text{s.t.} \quad \sum_{c=1}^{k} x_{vc} = 1 \qquad (v \in V) \tag{2}$$

$$y_{uv} \ge x_{uc} + x_{vc} - 1 \;\; (\{u,v\} \in E,\, c = 1, \dots, k) \tag{3}$$

$$x_{uc} \ge x_{vc} + y_{uv} - 1 \;\; (\{u,v\} \in E,\, c = 1, \dots, k) \tag{4}$$

Please cite this article in press as: J. Fairbrother, A.N. Letchford, Projection results for the k-partition problem, Discrete Optimization (2017), http://dx.doi.org/10.1016/j.disopt.2017.08.001.