ARTICLE IN PRESS

Discrete Applied Mathematics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Discrete Applied Mathematics

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

Half-regular factorizations of the complete bipartite graph

Mark Aksen [a], Istvan Miklos [a,b,c,*], Kathleen Zhou [a]

[a] Budapest Semesters in Mathematics, Bethlen Gábor tér 2, Budapest, 1071, Hungary
[b] Alfréd Rényi Institute, Reáltanoda u 13-15, Budapest, 1053, Hungary
[c] Institute for Computer Science and Control, Lágymányosi út 11, Budapest, 1111, Hungary

ARTICLE INFO

Article history:
Received 22 February 2016
Received in revised form 21 April 2017
Accepted 5 June 2017
Available online xxxx

Keywords:
Degree sequences
Degree matrix
Graph factorization
Edge packing
Latin squares
Markov chain Monte Carlo

ABSTRACT

We consider a bipartite version of the color degree matrix problem. A bipartite graph $G(U, V, E)$ is *half-regular* if all vertices in U have the same degree. We give necessary and sufficient conditions for a bipartite degree matrix (also known as demand matrix) to be the color degree matrix of an edge-disjoint union of half-regular graphs. We also give necessary and sufficient perturbations to transform realizations of a half-regular degree matrix into each other. Based on these perturbations, a Markov chain Monte Carlo method is designed in which the inverse of the acceptance ratios is polynomial bounded.

Realizations of a half-regular degree matrix are generalizations of Latin squares, and they also appear in applied neuroscience.

1. Introduction

Consider an edge-weighted directed graph as a model of a neural network. Such a network can be built up using real life measurements (see, for example, [19]), and neuroscientists are interested in comparing this network with random networks. If the edges were not weighted, the typical approach would be to generate random graphs with prescribed in and out degrees. This topic has a tremendous literature, see, for example, [18,9,15,5,7].

If weights are introduced, one might want to generate a random graph that maintains not only the sum of the weights, but also weights of individual edges. The weights can be transformed into colors, and we would be looking for an edge-colored graph with prescribed in and out degrees for each color. This problem is known as finding edge packing [2], edge disjoint realizations [10], or degree constrained edge-partitioning [1]. The problem has also applications in discrete tomography [1,6]. In addition to finding one edge-colored graph with given constraints, it is important to know how to generate a "typical" solution, as we can see in the before mentioned neuroscientific problem.

Unfortunately, the general edge packing problem is NP-complete even for 2 colors [4,8,6]. Although the general problem is NP-complete, special cases are tractable. Such special cases include the 2-color case with the subgraph of one of the colors almost regular [17,16,3]. Another tractable case is when the graph is bipartite, the number of colors is 2, and there exist constants k_1 and k_2 such that for each vertex, the total number of edges in one of the vertex class is $k_1 - 1$, k_1 or $k_1 + 1$ and in the other vertex class is $k_2 - 1$, k_2 or $k_2 + 1$ [10]. When the number of colors is unlimited, tractable solutions exist if the graphs are forests for each color [1] or the graphs are forests for each color except one of the colors might contain one cycle [12].

A special case is when the bipartite graph is the union of n 1-factors. Such edge packings are simply the Latin squares. For any n, Latin squares exist, and even a Markov chain is known that explores the space of Latin squares for a fixed n [13]. This

* Corresponding author at: Alfréd Rényi Institute, Reáltanoda u 13-15, Budapest, 1053, Hungary.
E-mail addresses: maksen@princeton.edu (M. Aksen), miklos.istvan@renyi.mta.hu (I. Miklos), zhou.kat@gmail.com (K. Zhou).

http://dx.doi.org/10.1016/j.dam.2017.06.003
0166-218X/

Markov chain is conjectured to be rapidly mixing, that is, computationally efficient at sampling random Latin squares. The conjecture remains open, despite many efforts made over the past two decades.

In this paper, we consider another tractable case with unlimited colors. We require that the subgraphs for each color be half-regular, that is, the degrees are constant on one of the vertex classes, furthermore, this regular vertex class must be the same for all colors. We give sufficient and necessary conditions when this edge packing problem has a solution. We also give necessary and sufficient perturbations to transform solutions into each other. A Markov chain Monte Carlo method has been given using these perturbations, and we give a proof that the inverse of the acceptance ratio is polynomial bounded.

2. Preliminaries

In this paper, we will work with realizations of half-regular degree matrices, defined below.

Definition 2.1. A *bipartite degree sequence* $D = ((d_1, d_2, \ldots, d_n), (f_1, f_2, \ldots, f_m))$ is a pair of sequences of non-negative integers. A bipartite degree sequence is *graphical* if there exists a simple bipartite graph G whose degrees correspond exactly to D. We say that G is a *realization* of D.

Definition 2.2. A *bipartite degree matrix* $\mathcal{M} = (D, F)$ is a pair of $k \times n$ and $k \times m$ matrices of non-negative integers. A bipartite degree matrix is graphical if there exists an edge colored simple bipartite graph $G(U, V, E)$ such that for all color c_i and for all $u_j \in U$, the number of edges of u_j with color c_i is $d_{i,j}$ and for all $v_l \in V$, the number of edges of v_l with color c_i is $f_{i,l}$. Such graph is called a *realization* of $\mathcal{M}$. A bipartite degree matrix is *half-regular* if for all $i \leq k$ and $j, l \leq n$, $d_{i,j} = d_{i,l}$.

The rows of $\mathcal{M}$ are bipartite degree sequences that are also called *factor* s and the edge colored realization of $\mathcal{M}$ is also called an *$\mathcal{M}$-factorization*.

We consider two problems. One is the existence problem which asks if there is a realization of a given half-regular degree matrix. The other is the sampling problem, which considers the set of all realizations of a half-regular degree matrix and asks how to sample uniformly a realization from this set. Markov Chain Monte Carlo (MCMC) methods are generally applicable for such problems, and we are also going to introduce an MCMC for sampling realizations of half-regular degree sequences. Below we introduce the main definitions.

Definition 2.3. A *discrete time, finite Markov chain* is a random process that undergoes transitions from one state to another in a finite state space. The process has the Markov property which means that the distribution of the next state depends only on the current state. The transition probabilities can be described with a transition matrix $\mathbf{T} = \{t_{i,j}\}$, where $t_{i,j} := P(x_i|x_j)$, namely, the conditional probability that the next state is x_i given that the current state is x_j. When the state space is a large set of combinatorial objects, the transition probabilities are not given explicitly. Rather, a random algorithm is given that generates a random x_i by perturbing the current state x_j. Such algorithm is called *transition kernel*.

When the process starts in a state x_i, after t number of steps, it will be in a random state with distribution $\mathbf{T}^t\mathbf{1}_i$, where $\mathbf{1}_i$ is the column vector containing all 0's except for coordinate i, which is 1. The Markov chain Monte Carlo method is to tailor the transition probabilities so that the limit distribution $\lim_{t\to\infty}\mathbf{T}^t\mathbf{1}_i$ is a prescribed distribution. To fulfill this, it is necessary for the Markov chain to be irreducible, aperiodic and the transitions be reversible as defined below.

Definition 2.4. Given a discrete time, finite Markov chain on the state space $\mathcal{I}$, the *Markov graph* of the Markov chain is a directed graph $G(V, E)$, with vertex set $\mathcal{I}$ and an edge from v_i to v_j iff $P(v_j|v_i) \neq 0$. A Markov chain is *irreducible* iff its Markov graph is strongly connected. When the transition kernel of an irreducible Markov chain generates a class of perturbations, we also say that this class of perturbations is irreducible on the state space.

A Markov chain is *aperiodic* if the largest common divisor of cycle lengths of its Markov graph is 1. This automatically holds if there is a loop in the Markov chain, that is, a state x exists for which $P(x|x) \neq 0$.

Definition 2.5. The transition kernel of a Markov chain is reversible if for all x_i, x_j, $P(x_i|x_j) \neq 0 \Leftrightarrow P(x_j|x_i) \neq 0$.

The strength of the theory of Markov chain Monte Carlo is that any irreducible, aperiodic Markov chain with reversible transition kernel can be tailored into a Markov chain converging to a prescribed distribution as stated below.

Theorem 2.6 ([20,11])**.** *Let an irreducible, aperiodic Markov chain be given with reversible transition kernel T over the finite state space $\mathcal{I}$. Let π be a distribution over $\mathcal{I}$, for which $\forall x \in \mathcal{I}, \pi(x) \neq 0$. Then the following algorithm, called the Metropolis–Hastings algorithm, also defines a Markov chain that converges to π; namely, its transition kernel $\mathbf{T}'$ satisfies $\lim_{t\to\infty}\mathbf{T}'^t\mathbf{1}_i = \pi$ for all indices i.*

1. *Draw a random y following the distribution $T(\cdot|x_t)$ where x_t is the current state of the Markov chain after t steps.*
2. *Draw a random u following the uniform distribution over $[0, 1]$. The next state, x_{t+1}, will be y if*

$$u \leq \min\left\{1, \frac{\pi(y)T(x_t|y)}{\pi(x_t)T(y|x_t)}\right\}$$

and x_t otherwise.

Please cite this article in press as: M. Aksen, et al., Half-regular factorizations of the complete bipartite graph, Discrete Applied Mathematics (2017), http://dx.doi.org/10.1016/j.dam.2017.06.003.