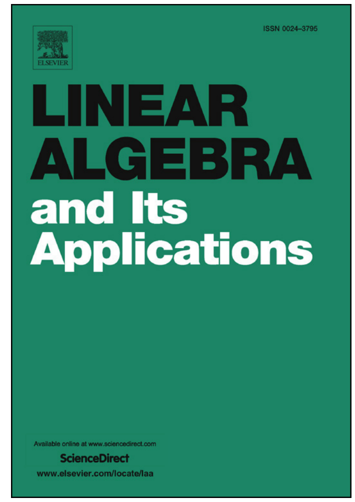
Accepted Manuscript

Completion of tree metrics and rank 2 matrices

Daniel Irving Bernstein



PII: S0024-3795(17)30436-6
DOI: http://dx.doi.org/10.1016/j.laa.2017.07.016
Reference: LAA 14262

To appear in: *Linear Algebra and its Applications*

Received date: 23 March 2017
Accepted date: 14 July 2017

Please cite this article in press as: D.I. Bernstein, Completion of tree metrics and rank 2 matrices, *Linear Algebra Appl.* (2017), http://dx.doi.org/10.1016/j.laa.2017.07.016

COMPLETION OF TREE METRICS AND RANK 2 MATRICES

DANIEL IRVING BERNSTEIN

Abstract. Motivated by applications to low-rank matrix completion, we give a combinatorial characterization of the independent sets in the algebraic matroid associated to the collection of $m \times n$ rank-2 matrices and $n \times n$ skew-symmetric rank-2 matrices. Our approach is to use tropical geometry to reduce this to a problem about phylogenetic trees which we then solve. In particular, we give a combinatorial description of the collections of pairwise distances between several taxa that may be arbitrarily prescribed while still allowing the resulting dissimilarity map to be completed to a tree metric.

Keywords: low-rank matrix completion, algebraic matroids, tropical geometry, tree-metric completion

MSC Classes: 14T05, 52B40, 52C25

1. introduction

Given a matrix where only some of the entries are known, the low-rank matrix completion problem is to determine the missing entries under the assumption that the matrix has some low rank r. One can also assume additional structure such as (skew) symmetry or positive definiteness. Practical applications of the low-rank matrix completion problem abound. A well-known example is the so-called "Netflix Problem" of predicting an individual's movie preferences from ratings given by several other users. A brief survey of other applications appears in [2].

Singer and Cucuringu show how ideas from rigidity theory can be applied to this problem in [16]. Jackson, Jordán, and Tibor further develop these ideas in [8, 9]. Király, Theran, and Tomioka incorporate ideas from algebraic geometry into this rigidity-theoretic framework in [11] and Király, Theran and Rosen further develop these ideas in [12]. We add tools from tropical geometry to this picture.

Let V be a determinantal variety over some algebraically closed field $\mathbb{K}$. The results in this paper concern the cases were $V = \mathcal{S}_r^n(\mathbb{K})$, the collection of $n \times n$ skew-symmetric $\mathbb{K}$-matrices of rank at most r, or $V = \mathcal{M}_r^{m\times n}(\mathbb{K})$, the collection of $m \times n$ $\mathbb{K}$-matrices of rank at most r. A *masking operator* corresponding to some $S \subseteq \binom{[n]}{2}$ in the skew symmetric case, or $S \subseteq [m] \times [n]$ in the rectangular case, is a map $\Omega_S : V \to \mathbb{K}^S$ that projects a matrix M onto the entries specified by S. In the case of skew-symmetric $n \times n$ matrices, we view S as the edge set of a graph on vertex set $[n]$, which we denote $G(S)$. In the case of rectangular matrices, we view S as the edge set of a bipartite graph on partite sets of size m and n which we also denote $G(S)$. Context will make the proper interpretation of $G(S)$ clear.

Low-rank matrix completion problems can now be phrased as: given $\Omega_S(M)$ can we recover M if we know $M \in V$? For generic M the answer to this question only depends on the observed entries S and not the particular values observed. Namely, given $\Omega(M)$ for generic $M \in V$, M may be recovered up to finitely many choices if and only if S is a spanning set of the algebraic matroid associated to V. Hence it is useful to find