MATROIDS AND THE CORRELATION CONSTANT

PEJMON SHARIATI

ABSTRACT. The correlation constant of a matrix, field, graph, or matroid explains how the edges in a graph or the column vectors in a vector space configuration are correlated. Our goal is to study different matrices and matroids and find correlation constants greater than 1 to improve the lower bound of the correlation constant for any field $\mathbb{F}$.

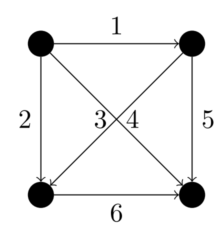
.

1. INTRODUCTION

. Let us start with M, a finite collection of vectors and b_i, b_{ij}, b stand for the number of bases for the column space of M that contain the distinct vector i, the distinct vectors i and j, and the total number of bases for the column space of M respectively. The formal definition of a correlation constant of a field $\mathbb{F}$ is the supremum of $\frac{bb_{ij}}{b_i b_j}$ over all pairs of distinct vectors i and j in finite vector configurations in vectors spaces over $\mathbb{F}$.

The Complete Graphs K_n. The first complete graph we will analyze is K_4. It can be described by the vertex-edge incidence matrix below.

$$K_4 = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & 0 & -1 & -1 \end{bmatrix}$$



Each row represents a vertex and each column represents an edge in the complete graph. The entries of this matrix indicated that this will also be a directed graph. If we look at all the entries of the second row we can analyze all the edges related to the second vertex. Edge 1 is directed towards the second vertex because the (2,1)-entry is -1. Edge's 4 and 5 are directed away from the second vertex because the (2,4) and (2,5)-entries are 1's. The (i,j)-entries that are 0 indicate that there are no edges incident on that vertex.

Key words and phrases. Matroids, Correlation Constant, Representable.

The Rank and The Bases. First we will determine the rank of the matrix in order to find the dimension of the column space. Let us choose to view each column as a column vector and using Gaussian elimination we can see that the rank of the matrix is 3, therefore we will find bases for the column space with 3 vectors.

This graph is represented over all the real numbers so after using Gaussian elimination to determine all the possible bases we can see that each vector is included in 8 bases. There is also a total of 16 bases.

Below, is the Hodge-Riemann form of the vector configuration of K_4 from Huh's paper [1].

$$HR(K_4) = \begin{bmatrix} 0 & 3 & 3 & 3 & 3 & 4 \\ 3 & 0 & 3 & 3 & 4 & 3 \\ 3 & 3 & 0 & 4 & 3 & 3 \\ 3 & 3 & 4 & 0 & 3 & 3 \\ 3 & 4 & 3 & 3 & 0 & 3 \\ 4 & 3 & 3 & 3 & 3 & 0 \end{bmatrix}$$

The entries of this matrix is described by

$$HR(M)_{ij} = \begin{cases} 0 & \text{if } i = j, \\ b_{ij} & \text{if } i \neq j. \end{cases}$$

Now that we have the various combinations of b_i, b_j and b_{ij} we can concoct a set of potential correlation constants and take the supremum to find the correlation constant of K_4

$$\sup \left\{ 0, \frac{3}{4}, 1 \right\} = 1$$

2. Patterns of the correlation constant

Now we would like to ask ourselves, what essentially affects the correlation constant? Does it depend on the characteristic of a field or is it an intrinsic property of a field? [1]

Modular Arithmetic. Let us see how the correlation constant of K_4 will be affected if we were to calculate it in modulo 2. We will call Z the matrix representing K_4 over $\mathbb{F}_2$.

$$Z = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Now that the (i,j)-entries of K_4 are in modulo 2 we can calculate the rank of Z and determine the bases for its column space. The rank of Z is 4 so we must use Gaussian elimination to determine 4 linearly independent vectors that span the column space. However, we must use Gaussian elimination via modulo 2 and from that we can determine that there are no possible bases for the column space. The correlation constant of K_4 via modulo 2 is 0. From these observations we can see that the correlation constant of K_4 decreased when represented over $\mathbb{F}_2$.

For the rest of this section we would like to answer the following questions:

(1) Is the correlation constant of every complete graph 1?
(2) Does the correlation constant always decrease with any complete graph when represented over $\mathbb{F}_2$?
(3) Is the correlation constant affected when a complete graph has orientation? Whether it is represented by an incidence matrix or an adjacency matrix?

We will start by answering the first question, and we will look at the incidence matrix of K_3 which is essentially a triangle.

$$K_3 = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$$

After repeating the same process as described earlier in this paper we can see that the correlation constant is $\frac{3}{4}$. Now let us see how the correlation constant is affected when this matrix is represented over $\mathbb{F}_2$. Let V denote the matrix representing K_3 over $\mathbb{F}_2$

$$V = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

We can see that the correlation constant of V is the same as K_3, therefore we can answer the second question. The correlation constant does not always decrease when the matrix is represented over $\mathbb{F}_2$. Although this appeared to be a promising pattern it is a lot more difficult to find one than it seems. Now let's see what happens when K_4 and K_3 are represented by an adjacency matrix.

Adjacency matrices are square matrices that represent undirected graphs with 0's and 1's as its entries. The dimensions of the adjacency matrix representing a complete graph K_n is $n \times n$. The following are the adjacency matrices for K_4 and K_3.

$$K_4 = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}, K_3 = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

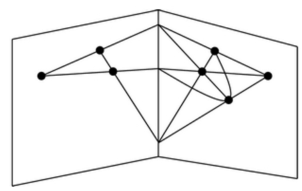
Using the same process as before to calculate the bases we can see that the correlation constant of these two matrices are 1. As you calculate the correlation constant of more and more adjacency matrices representing complete graphs we can see that it is always 1. Now that we have seen the incidence and adjacency matrices for K_4 and K_3 we can answer questions 1 and 3. The correlation constant of every complete graph K_n is always 1 when represented by an adjacency matrix (undirected), but only sometimes when represented by an incidence matrix (directed). This leads us to postulate that orientation and direction does affect the correlation constant of a graph.

3. Interesting Matroids

Now that we have gained an understanding of what a correlation constant is, how to calculate it, and its properties we can start to look at more complex matroids.

Let us look at the following matroid S_8 from Oxley's book [2].

$$S_8 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$



This matroid is only representable if and only if that field has characteristic two, therefore it is represented over $GF(2)$. We must now determine the correlation constant of this matroid by using the previous processes, but with Gaussian elimination via modulo 2. We will use Sage to easily calculate the bases of this matroid [4].From this we can determine a set of potential correlation constants and take the supremum of that set to get the correlation constant which is $\frac{(48)(12)}{(20)(28)} = \frac{36}{35}$. We obtained these numbers when $i = 4$ and $j = 8$. This is the first correlation constant we have seen greater than 1 so we have accomplished what we wanted to in the abstract.

Our next matroid is called $M_{4,2}$ which can be constructed from the conditions set forth in Schröter's paper [3]. This was the first matroid discovered by Seymour and Welsh to have a correlation constant greater than 1.

$$M_{4,2} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \end{bmatrix}$$

Using Gaussian elimination via modulo 2 [4] we can see that the correlation constant of this matroid is also $\frac{36}{35}$. We obtained these numbers when $i = 1$ and $j = 5$. $M_{4,2}$ and S_8 are represented by two different matrices, so it is curious to see that their correlation constants are the same. However, the first four columns of each matrix were part of the I_4 identity matrix which means the other 4 column vectors are linear combinations of the first four.

4. Hypotheses and Conclusions

From what we have seen so far we can see that to find correlation constants greater than 1 we do not have to construct large matrices, nor make the entries of these matrices large. We have seen two examples of matrices completely comprised of only 1's and 0's that involve the identity matrix and their correlation constants were greater than 1. We now hope to find more matroids with correlation constants greater than 1 and $\frac{36}{35}$ to improve its lower bound. We have established a solid background, found few patterns, and seen real examples of these special matroids. We hope to find more and have the tools to do so.

Special thanks to Professor Tarasca for his mentorship and introducing me to the topic of matroids and the correlation constant. Special thanks to Lazaros Gallos for involving me in the REU DIMACS program.

References

[1] J. Huh. Combinatorial applications of the Hodge-Riemann relations. *ArXiv e-prints*, November 2017.

[2] J. G. Oxley. *Matroid theory*. Oxford University Press, 2011.

[3] B. Schröter. The correlation constant of a field. *ArXiv e-prints*, April 2018.

[4] The Sage Developers. *SageMath, the Sage Mathematics Software System (Version 8.2.1)*, 2018. `http://www.sagemath.org`.

Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08901

Current address: Department of Mathematics, Rutgers University, New Brunswick, New Jersey 08901

E-mail address: `Pejmon.Shariati@tufts.edu`