2012

Entropy Optimal Orthogonal Matrices

Akhilesh Pathak
Wright State University

Entropy Optimal Orthogonal Matrices

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

By

Akhilesh Pathak
B. Tech, Indian Institute of Technology Kanpur, 2009

2012
Wright State University

WRIGHT STATE UNIVERSITY
SCHOOL OF GRADUATE STUDIES

June 1, 2012

I HEREBY RECOMMEND THAT THE THESIS PREPARED UNDER MY SUPERVISION BY Akhilesh Pathak ENTITLED Entropy Optimal Orthogonal Matrices BE ACCEPTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF Master of Science.

K.T. Arasu, Ph.D.
Thesis Director

Weifu Fang, Ph.D.,
Chair
Department of
Mathematics and Statistics
College of Science and Mathematics

Committee on
Final Examination

K.T. Arasu, Ph.D.

David Miller, Ph.D.

Richard Mercer, Ph.D.

Andrew Hsu, Ph.D.
Dean, School of
Graduate Studies

ABSTRACT

Pathak, Akhilesh, M.S., Department of Mathematics and Statistics, Wright State University 2012, Entropy Optimal Orthogonal Matrices.

The entropy of an orthogonal matrix is defined by the Gadiyar, Maini, Padma and Sharatchandra who have re-defined Hadamard matrices as the orthogonal matrices that saturate the bound for entropy. They also presented numerical results for maximal entropy for dimension $n = 3, 5$. We prove the results analytically for $n \equiv 0 (\mathrm{mod}\ 4), n = 3$ and construct local extremums for $n = 5, 6, 10, 2p, 3p$, where p is prime. We also provide cojectures on necessary conditions for optimality and optimal matrices based on the prime factorization of the order.

CONTENTS

1. Introduction

The entropy of a random variable (or a probability distribution) is defined as the measure of uncertainty in Information Theory. As the uncertainty in the outcome(s) of an event decreases, we say that we have received some information. The idea of entropy was first introduced by Claude Shannon in his famous paper [1]. In signal processing and communication problems, specifying the probabilities in the cases where little information is available, is not solvable completely. This is a classical problem in probability theory. One can intuitively think that the optimal uncertainty could happen when all the outcomes are equiprobable, for which mathematicians Jacob Bernoulli and Laplace gave an intuitive principle of indifference [2]. George Boole, John Venn, and others gave it formally, named as the principle of insufficient reason [3]. During the 20th century, Information Theory provided advancement to classical notions by giving a unique and unambiguous criterion for the measure of uncertainty represented by a discrete random probability distribution, called entropy. The characterization of Shannon's measure of entropy was given as follows. For a discrete random variable taking finite number of possible values $x_1, x_2, \cdots, x_n$ with probabilities $p_1, p_2, \cdots, p_n$ such that $p_i > 0$ and $\sum_{i=1}^{n} p_i = 1$, the amount of uncertainty of the probability distribution is given by

$$H(X) = \sum_{i=1}^{n} p_i \log_2 \frac{1}{p_i}.$$

The minimum for the entropy defined above is zero for the case of certainty, i.e. for some j,

$$p_i = \begin{cases} 1 & for\ i = j \\ 0 & for\ i \neq j \end{cases}$$

The maximum for the entropy is $\log_2 n$ when all values are equiprobable, i.e.

$$p_i = \frac{1}{n}, \forall\ i = 1, 2, ..., n.$$

The entropy for an orthogonal matrix has been defined in [4]. We shall discuss bounds on the entropy for several values of n. The bound for an $n \times n$ orthogonal matrix has an interesting connection with Hadamard matrices. A detailed introduction to Hadamard matrices and its properties have been given in the next section 2.1. We also prove achievable sharper bounds with corresponding examples of an $n \times n$ orthogonal matrix for the case when Hadamard matrices do not exist. Certain symmetries have been found for various categories based on the prime factorization of n. The numerical computations done by [4] give entropy optimal matrices of order 3 and 5. We have analytically proved the results for orders 4k, 3, and have extended the results for order 6 and 10 using the Kronecker product for the orthogonal matrices. To prove the universal bound, we have used Jensen's inequality and for a sharper bound of order 3, we use the characterization given by Stewart [5].

1.1. **Entropy for an Orthogonal Matrix.** For an $n \times n$ orthogonal matrix M(n) with real entries a_{ij} at the $(i,j)^{th}$ position, the definition for entropy is given as below

$$E^R[M(n)] = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}^2 \ln \left(\frac{1}{a_{ij}^2} \right).$$

Another definition we give for a different class of matrices, namely complex inverse orthogonal matrices defined in [6].

1.2. **Complex Inverse Orthogonal Matrix.** An $n \times n$ complex inverse orthogonal matrix M is a matrix such that the product $(M)\left(\frac{1}{M}\right) = nI_n$, where M is a matrix with uni-modular columns and complex non-zero entries a_{ij} at the $(i,j)^{th}$ position and $\frac{1}{M}$ is the matrix with complex entries $\frac{1}{a_{ij}}$ at the $(i,j)^{th}$ position.

1.3. **Entropy for a Complex Inverse Orthogonal Matrix.** For an $n \times n$ inverse orthogonal matrix M with complex entries a_{ij} at the $(i,j)^{th}$ position, we define entropy as

$$E^C[M(n)] = \sum_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \ln \left(\frac{1}{|a_{ij}|^2} \right),$$

where $|a_{ij}|$ represents the absolute value of the complex number a_{ij}.

1.4. **Entropy Saturation.** We define the entropy saturation for an orthogonal matrix $M(n)$ of order n as the ratio of its entropy to the upper bound for n, denoted as

$$S[M(n)] = \frac{E[M(n)]}{nln(n)}$$

For a Hadamard matrix (if exists) of order n, $S[M(n)] = 1$.

2. Preliminaries

2.1. **Hadamard Matrices.** A Hadamard matrix H of order n is defined as an $n \times n$ square matrix with entries from 1, -1 such that

$$HH^T = nI_n,$$

where I_n is the identity matrix. In other words, Hadamard matrix is a matrix whose entries are either 1 or -1 and whose rows are mutually orthogonal. Clearly, the columns must also be mutually orthogonal, and the following also holds [7];

$$H^T H = nI_n.$$

From here onwards, we shall use the notation H_n for an $n \times n$ Hadamard matrix.

2.2. **Cayley's Parameterization of Orthogonal Matrices.** Any orthogonal matrix M which does not have -1 as an eigenvalue can be expressed as

$$M = (I + S)(I - S)^{-1},$$

for some suitable skew-symmetric matrix S. Conversely, any skew-symmetric matrix can be expressed in terms of a suitable orthogonal matrix M by

$$S = (M + I)^{-1}(M - I).$$

The above two forms set-up a one-to-one correspondence between orthogonal and skew-symmetric matrices. Proof is given in [11].

2.3. **Exponential Parameterization over Special Orthogonal Group SO(N).** For every positive integer N, the orthogonal group O(N) is the group of $N \times N$ orthogonal matrices M satisfying

$$MM^T = M^T M = I_n, \ and \ M^* = M,$$

where M^* is the conjugate of M, because the determinant of an orthogonal matrix is either 1 or -1.

2.4. **The Optimization Problem.** The general optimization problem [14] over orthogonal matrices can be formulated as

$$max_{X \epsilon R^{(n \times n)}} F(X), \ s.t. \ X^T X = I_n,$$

Where I_n is the identity matrix and $F(X) : R^{n \times n} \rightarrow R$ is a differentiable function. The feasible set $M(n \times n) = \left\{ X : X \epsilon R^{n \times n} : X^T X = I_n \right\}$ is often referred to as the Stiefel Manifold.

This has wide applications in polynomial optimization, combinatorial optimization, eigenvalue problems, sparse PCA, p-harmonic flows, 1-bit compressive sensing, matrix rank minimization, etc. These problems are difficult because the constraints are not only non-convex, but numerically expensive to preserve during iteration. It is generally difficult to solve this problem since the orthogonality constraints can lead to many local maximizers and, in particular, several of these problems in special forms are NP-hard [12]. There are no known algorithms to obtain a global optimizer except for a few simple cases. Most existing constraint preserving algorithms use matrix reorthogonalization that requires matrix factorization such as SVDs. Other popular algorithms generate points along geodesics of $\mathrm{M}(n \times n)$ that compute matrix exponentials or solve PDEs. We will approach the problem using the algorithm given in [12] based on a simpler constraint preserving formula.

3. Known Results

As given in [4], we know the following results on the entropy function defined by [4], using numerical computation.

(1) For n = 2 and $n \equiv 0 (\text{mod } 4)$, the scaled Hadamard matrix saturates the entropy. E.g.

$$M = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -1/2 & 1/2 & -1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}$$

(2) For n = 3 the matrix below with rational entries gives maxima.

$$M = \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}$$

(3) For n = 5 the matrix below with rational entries gives maxima.

$$M = \begin{bmatrix} -3/5 & 2/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & -3/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & -3/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & -3/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & 2/5 & -3/5 \end{bmatrix}$$

4. Bounds On Entropy Of Real Orthogonal Matrix

4.1. Universal bound for all n:

Theorem 4.1. $E^R[M(n)] \leq nln(n)$.

Proof. Let $f : R \to R$ be a concave function, then the Jensen's inequality states that

$$\frac{\sum_{i=1}^{n} f(x_i)}{n} \leq f\left(\frac{\sum_{i=1}^{n} x_i}{n}\right)$$

And equality holds if and only if all the x_i's are equal.

For $f(x) = xln(1/x)$ defined for $0 < x \leq 1$, we can show that $f(x)$ is concave as $f''(x) = -1/x < 0$. For $x_i \in (0, 1]$ for all i such that $\sum_{i=1}^{n} x_i = 1$ using the Jensen's inequality we get,

$$\frac{\sum_{i=1}^{n} f(x_i)}{n} \leq f\left(\frac{1}{n}\right)$$
$$\frac{\sum_{i=1}^{n} x_i ln(1/x_i)}{n} \leq \frac{1}{n} ln(n)$$
$$\sum_{i=1}^{n} x_i ln(1/x_i) \leq ln(n)$$

The equality is achieved if and only if $x_i = \frac{1}{n}$ for all i. Now, consider an $n \times n$ orthogonal matrix M. For $j = 1, 2, ...n$ the j^{th} column has some of squares of the elements equal to 1, therefore, we have

$$\sum_{i=1}^{n} a_{ij}^2 \ln\left(\frac{1}{a_{ij}^2}\right) \leq ln(n)$$

for all j. Therefore,

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2 \ln\left(\frac{1}{a_{ij}^2}\right) \leq nln(n).$$

$$E^R[M(n)] \leq nln(n).$$

And the equality holds if and only if $a_{ij}^2 = \frac{1}{n}$, i.e. $a_{ij} = \pm\frac{1}{\sqrt{n}}$. Therefore the equality is achieved by $\frac{1}{\sqrt{n}}H_n$ where H_n is the Hadamard matrix of order n. □

Hence in the cases, where n is a positive integer such that the Hadmard matrix exists, the upper bound is achieved and by $\left(\frac{1}{\sqrt{n}}\right)H_n$ only. Therefore, we have an optimal matrix for n = 1, 2, 4, 8, 12, $\cdots$ i.e. in the cases when n = 1, 2 or 0 (mod 4). Hence the existence of optimal orthogonal matrix relies on the truth content of the Hadamard conjecture.

4.2. **Sharp bounds when n is not a multiple of 4:** In the cases when the Hadmard matrix does not exist, in the expression for the upper bound of the entropy, the inequality is strict; it cannot be achieved and therefore should be optimized to a sharper bound by several other techniques. We will describe the bounds for certain categories later in this paper. However, the bound $nln(n)$ is achievable for $E^C[M(n)]$ for all n. Before we proceed to other cases, we summerize the algorithms we use.

Given a feasible point X and the gradient $G := DF(X) = \left(\frac{\partial F(x)}{\partial X_{i,j}}\right)$, we define a skew-symmetric matrix A as either

$$A := GX^T - XG^T$$

or

$$A := (P_x G)X^T - X(P_x G)^T, \ where \ P_x := \left(I - \frac{1}{2}XX^T\right).$$

The new trial point is determined by the Crank-Nicolson-like scheme

$$Y(\tau) = X - \frac{\tau}{2}A(X + Y(\tau)),$$

where $Y(\tau)$ is given in the closed form: $Y(\tau) = QX$ and $Q := \left(I + \frac{\tau}{2}A\right)^{-1}\left(I - \frac{\tau}{2}A\right)$, as discussed in Cayley parameterization.

Optimality Conditions: We state the first-order and second order optimality conditions in the following two lemmas due to [12][14]. The Lagrangian function for the generalized optimization problem can be given as,

$$L(X, l) = F(X) - \frac{1}{2}tr(l(X^T X - 1)),$$

where l is the symmetric Lagrangian multiplier corresponding to $X^T X = I$.

Lemma 4.2. *If X is optimal, then X satisfies the first-order optimality conditions, $D_x L(X, l) = G - XG^T X = 0$ and $X^T X = 1$ and $X^T X = I$, with the associated Lagrangian multiplier $l = G^T X$. Define*

$$\triangledown F(X) := G - XG^T X, \ and \ A := GX^T - XG^T.$$

Then, $\triangledown F(X) = AX$. Moreover, $\triangledown F(X) = 0$, if and only if A=0

Lemma 4.3. *Second order necessary conditions: Suppose an $n \times n$ matrix X is optimal. Then there exists a Lagrange multiplier l such that the first order conditions are satisfied. Suppose also that*

$$tr(Z^T DD)(D(F(X))[z]) - tr(lZ^T Z) < 0,$$

for any matrix Z. Then X is a strict maxima.

Lemma 4.4.

(1) *Given any skew symmetric matrix W, the matrix $Q := (I+W)^{-1}(I-W)$ is well defined and orthogonal, i.e., $Q^T Q = I$.*

(2) *Given any skew-symmetric matrix W , the matrix $Y(\tau) = X - \frac{\tau}{2}A(X+Y(\tau))$ satisfies $Y(\tau)^T Y(\tau) = X^T X$, and $Y(\tau)$ is given by*

$$Y(\tau) = \left(I + \frac{\tau}{2}A\right)^{-1}\left(I - \frac{\tau}{2}A\right)X.$$

Also, its derivative with respect to τ is given by

$$Y'(\tau) = \left(I + \frac{\tau}{2}A\right)^{-1} W \left(\frac{X+Y(\tau)}{2}\right),$$

and in particular, $Y'(0) = -WX$.

(3) *Set $W = A = GX^T - XG^T$. Then $Y(\tau)$ is the descent curve at $\tau = 0$, i.e.*

$$F'_\tau(Y(0)) := \frac{\partial F(Y(\tau))}{\partial \tau}|_{\tau=0} = -\frac{1}{2}\left\|A\right\|_F^2,$$

where $F'_\tau(Y(\tau))$ denotes the derivative of $F_\tau(Y(\tau))$ with respect to τ.

Proof of (1) is well-known. And (2) and (3) are given in [12][14].

Monotone Curvilinear Search Algorithm

The curve $Y(\tau)$ defined above satisfies the following condition: The matrix W is continuous in X and satisfies

$$F'_\tau(Y(0)) = -\sigma \|A\|_F^2\,,\ \textit{where } \sigma > 0 \textit{ is constant.}$$

Algorithm Steps:

(1) Choose an initial point X_0, i.e. a matrix of order n.

(2) Initialize: set $k \leftarrow 0, \varepsilon \geq 0, and\ 0 < \rho_1 < \rho_2 < 1$.

(3) While true do:

  (i) Generate A

  (ii) Compute the step size τ_k

  (iii) Update $X_{k+1} \leftarrow Y(\tau_k)$

  (iv) Stop if $\|\bigtriangledown F_{k+1}\|$. Otherwise $k \leftarrow k+1$

Using the algorithms we cojecture the following observation. If M^* is the $n \times n$ orthogonal matrix such that $E^R[M(n)]$ is optimal, then the matrix M_2^* formed by squaring the elements of M^*, is a symmetric matrix up to elementary row transformation.

The results above will be confirmed in the following section.

5. Constructions and Proofs

5.1. **Case: n=3.** There are exactly three independent real variables in a 3×3 orthogonal matrix. Consider the following 3×3 orthogonal matrix

$$M = \begin{bmatrix} ? & x & y \\ ? & ? & z \\ ? & ? & ? \end{bmatrix}.$$

The element a_{11} can be determined by the fact that sum of squares of the elements in the first row is 1. Such a_{11} will have two values. Once a_{11} is determined the second row has two variables a_{21} *and* a_{22} and two corresponding equations for orthogonality i.e. the product of corresponding elements in the 1st and 2nd row is zero and the sum of squares of the elements is 1. This will determine at most four pairs of a_{21} *and* a_{22}, after that $a_{31}, a_{32},$ *and* a_{33} are determined. The third row having three unknowns and three equations for orthogonality, again, give at most eight solutions. Therefore, the triplet (x,y,z) determines at most 64 orthogonal matrices. The complexity of the total orthogonal matrices is the same as that of the suitable triplets, (x,y,z).

Another way to construct an orthogonal matrix with exactly three independent variables could be as below. Choosing three diagonal elements (having absolute value less than 1) independently, and determining the remaining six elements by the six orthogonality conditions.

$$M = \begin{bmatrix} x & ? & ? \\ ? & y & ? \\ ? & ? & z \end{bmatrix}$$

$$E^R(M) = [x^2 ln\left(\frac{1}{x^2}\right) + y^2 ln\left(\frac{1}{y^2}\right) + z^2 ln\left(\frac{1}{z^2}\right)] + \sum_{i \neq j} a_{ij}^2 ln\left(\frac{1}{a_{ij}^2}\right)$$

For any such fixed triplet, the quantity in the first bracket in the expression for E(M) is fixed. In order to optimize $E^R(M)$, we use Jensens Inequality, for the six elements a_{ij}^2 with $i \neq j$. The summation below is fixed.

$$\sum_{i \neq j} a_{ij}^2 = 3 - (x^2 + y^2 + z^2) = constant$$

Therefore, the optimal entropy occurs (for fixed diagonal), when a_{ij}^2, is a constant for all $i \neq j$ (If such an orthogonal matrix exists.) Hence, the optimal E(M) must occur when such a symmetric matrix exists. There is a characterization based on Householder reflections, given by Stewart (1980), where a symmetric orthogonal matrix can be represented by

$$A(u) = I - 2\frac{uu^T}{\|u\|^2},$$

where u is a non-zero vector and $\|.\|$ is the Euclidean norm. Any orthogonal matrix of size $n \times n$ can be constructed as a product of at most n such reflections. From the expression above for A, It can be shown that $A^T = A \ and \ AA^T = A^T A = I$. For the vector u=(p,q,r), with norm $\|u\|^2 = p^2 + q^2 + r^2 = d$, we have the expression for A:

$$A = \begin{bmatrix} 1 - 2p^2/d & -2pq/d & -2pr/d \\ -2pq/d & 1 - 2q^2/d & -2qr/d \\ -2pr/d & -2qr/d & 1 - 2r^2/d \end{bmatrix}.$$

Suppose the matrix M consists of 3 column vectors v_1, v_2, v_3 and define

$$f(v_1) = f(x_1, y_1, z_1) = x_1^2 ln\left(\frac{1}{x_1^2}\right) + y_1^2 ln\left(\frac{1}{y_1^2}\right) + z_1^2 ln\left(\frac{1}{z_1^2}\right).$$

Then, $E^R(M) = f(v_1) + f(v_2) + f(v_3)$. As f is a concave function, E(M) is also a concave function. Therefore, using Jensen's inequality for optimality, we need $f(v_1) = f(v_2) = f(v_3)$. Applying the condition above to the column vectors of the matrix A(p,q,r), we get p=q=r, for $u = (p, p, p)$, $\|u\|^2 = 3p^2 = d$, which gives

$$A(p,p,p) = \begin{bmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{bmatrix} \; or \; (-1) \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}.$$

This is the required 3×3 orthogonal matrix.

Remark: Note that the matrix above is optimal for every concave function applied element-wise and added over all the elements. It does not depend on the function as long as the function is concave.

Theorem 5.1. *The saturation value remains constant for the Kronecker products of matrices of order 3, i.e. the value* $S[M(3^n)] = 0.88$ *for all positive integers n, where* $M(3^n)$ *is given by the Kronecker product* $M(3^n) = M(3) \otimes M(3) \otimes ..... \otimes M(3)$

Proof. We use induction, since, $S[M(9)] = S[M(3)] = 0.88$ and

$$M(3^{n+1}) = M(3^n) \otimes M(3)$$

The entropy,

$$E^R[M(3^{n+1}] = 3\frac{n+1}{n}E^R[M(3^n)]$$

Therefore,

$$S[M(3^{n+1})] = 3\frac{n+1}{n}S[M(3^n)]\frac{3^n ln(3^n)}{3^{n+1}ln(3^{n+1})}$$

or

$$S[M(3^{n+1})] = S[M(3^n)]$$

Hence $S[M(3^n)] = 0.88$ is constant for all n. □

Verifying KKT point and Numerical results using the Algorithm: The optimal matrix of order 3 above satisfies the conditions in lemma 4.3 and 4.4 The numerical results using the algorithm described above gives the following results for step size 0.2:

Starting with $M_0 = \begin{bmatrix} 0 & -0.8 & -0.6 \\ -0.8 & -0.36 & 0.48 \\ 0.6 & 0.48 & -0.64 \end{bmatrix}$,

we get the 11^{th} iteration, $M_{11} = \begin{bmatrix} 0.33 & -0.66 & -0.66 \\ -0.66 & 0.33 & -0.66 \\ -0.66 & -0.66 & 0.33 \end{bmatrix}$,

which is the same as the result we have proved above. The algorithm works really fast for small values of n, as long as the saddle point (see below) is avoided. A saddle point is given by the following matrix:

$$M = \begin{bmatrix} 1/2 & -1/\sqrt{2} & -1/2 \\ 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 1/2 & -1/\sqrt{2} & 1/2 \end{bmatrix}.$$

The corresponding value of the entropy has a significant deviation from the bound $nln(n)$. For the matrix above we obtain,

$$E^R[M(3)] \approx 2.89 < 3ln(3) \approx 3.29.$$

Linear search

Given a feasible point X, we vary τ on real line, and evaluate $Y(\tau)$. It can be proved that $Y(\tau)$ is again an orthogonal matrix, in fact, more generally,

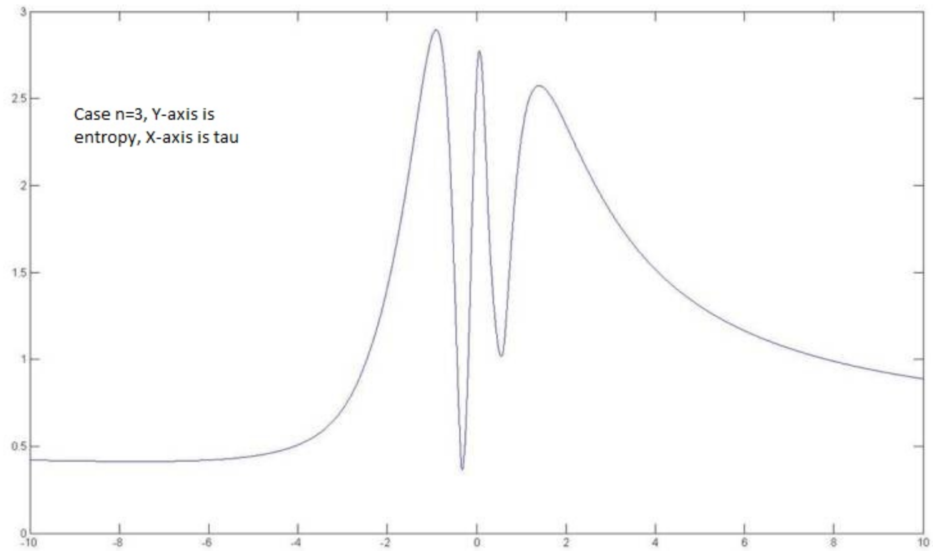
$$Y(\tau)Y(\tau)^T = XX^T$$

We propose two methods.

1. Fix X, vary τ over real line.

2. Recursive method: (i) Fix $\tau \rightarrow$ (ii) fix $X \rightarrow$ (iii) evaluate $Y(\tau) \rightarrow$ (iv) update $X = Y(\tau) \rightarrow$ go to step (ii).

In the picture above plot is the between $Y(\tau)$ on X-axis and Entropy on Y-axis.



The absolute maxima appears for $\frac{\tau}{2} = -0.905$, and the corresponding matrix is

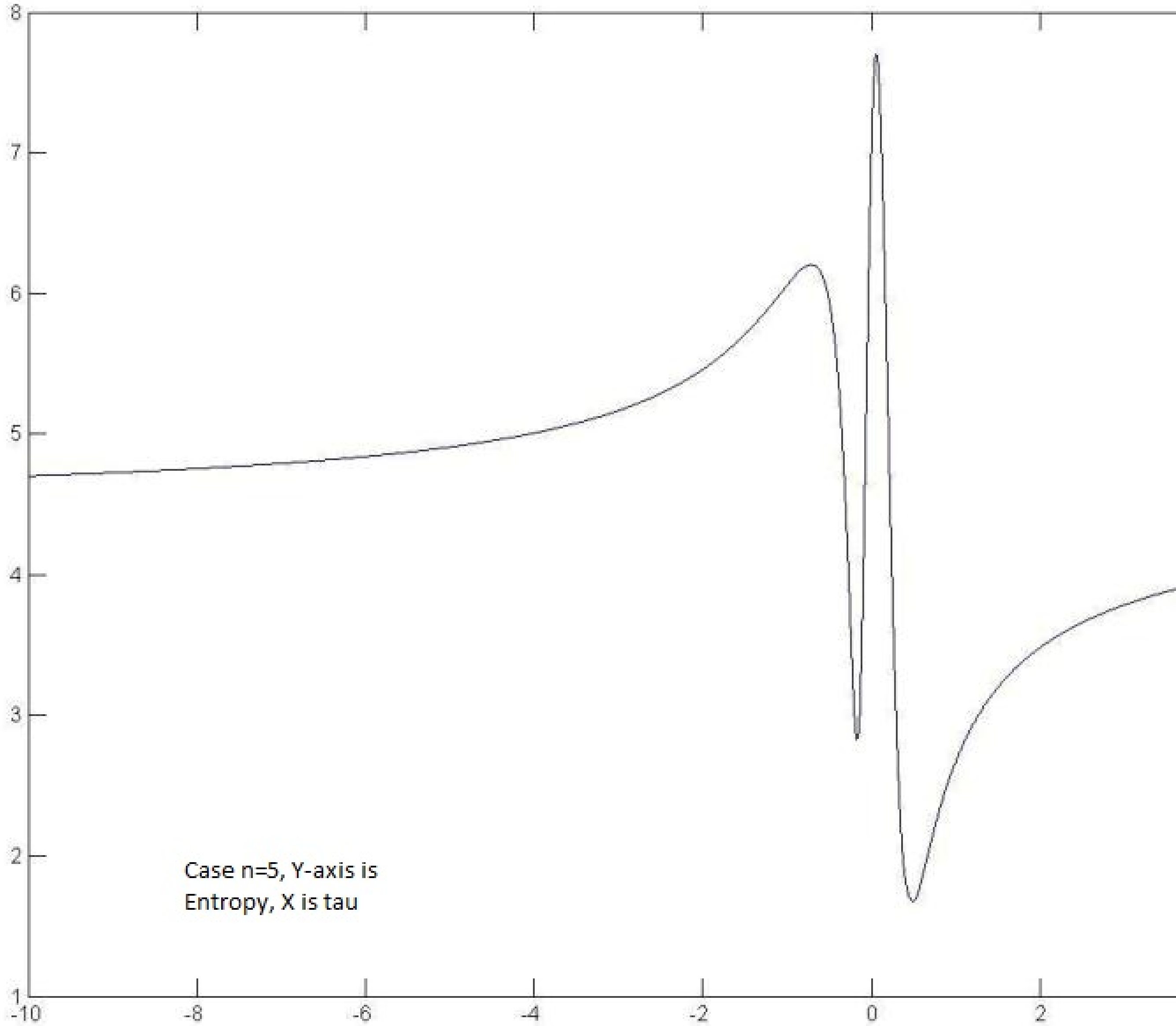
$$\begin{bmatrix} 0.668 & -0.666 & 0.333 \\ 0.668 & 0.334 & -0.666 \\ -0.333 & -0.666 & -0.668 \end{bmatrix}$$

which is same as $I - (2/3)J$ after interchanging the rows. Other critical points are as below,

For $\frac{\tau}{2} = -0.065$, it actually gives a saddle point, the corresponding matrix being,

$$\begin{bmatrix} 0.010 & -0.706 & -0.707 \\ 0.706 & -0.495 & 0.505 \\ 0.707 & 0.505 & -0.495 \end{bmatrix} \approx \begin{bmatrix} 0 & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{\sqrt{2}} & \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$$

Other critical points are as following,



Maxima for $\frac{\tau}{2} = 1.4$, given by,

$$\begin{bmatrix} 0.8271 & 0.1826 & -0.5316 \\ -0.1826 & -0.8072 & -0.5613 \\ 0.5316 & -0.5613 & 0.6342 \end{bmatrix}$$

Saddle for $\frac{\tau}{2} = 0.555$, given by,

$$\begin{bmatrix} 0.4292 & -0.0367 & -0.9025 \\ 0.0367 & -0.9976 & 0.0580 \\ 0.9025 & 0.0580 & 0.4268 \end{bmatrix}$$

Minima for $\frac{\tau}{2} = 0.32$, given by,

$$\begin{bmatrix} 0.0190 & -0.9993 & 0.0308 \\ 0.9993 & 0.0180 & -0.0314 \\ -0.0308 & -0.0314 & -0.9991 \end{bmatrix}$$

which is actually identity matrix after elementary row operations.

5.2. **Case: n=5.** We have the same theorem again as in case 3.

Theorem 5.2. *The saturation value remains constant for the Kronecker products of matrices of order 5, i.e. the value $S[M(5^n)]$ remains constant for all positive integers n, where $M(5^n)$ is given by the Kronecker product $M(5^n) = M(5) \otimes M(5) \otimes ..... \otimes M(5)$*

Proof. Same as for n=3. □

Using the algorithms we find the following matrices are optimal.

$$\begin{bmatrix} 0.7513 & -0.2263 & -0.3282 & -0.3602 & -0.3832 \\ -0.2263 & 0.7941 & -0.2986 & -0.3277 & -0.3487 \\ -0.3282 & -0.2986 & 0.5669 & -0.4754 & -0.5057 \\ -0.3602 & -0.3277 & -0.4754 & 0.4782 & -0.5551 \\ -0.3832 & -0.3487 & -0.5057 & -0.5551 & 0.4095 \end{bmatrix}$$

And

$$\begin{bmatrix} 0.5948 & -0.4009 & -0.4032 & -0.4012 & -0.4023 \\ -0.4009 & 0.6033 & -0.3990 & -0.3970 & -0.3981 \\ -0.4032 & -0.3990 & 0.5987 & -0.3993 & -0.4004 \\ -0.4012 & -0.3970 & -0.3993 & 0.6027 & -0.3984 \\ -0.4023 & -0.3981 & -0.4004 & -0.3984 & 0.6005 \end{bmatrix}$$

The matrix above is absolute maximum and indeed is $-[I - (2/5)J]$ i.e.

$$M = \begin{bmatrix} -3/5 & 2/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & -3/5 & 2/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & -3/5 & 2/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & -3/5 & 2/5 \\ 2/5 & 2/5 & 2/5 & 2/5 & -3/5 \end{bmatrix}$$

The matrices above for n=3, 5 give rise to the following conjecture,

Conjecture 1. *If M^* is the $n \times n$ orthogonal matrix such that $E^R[M(n)]$ is optimal, then the matrix M_2^* formed by squaring the elements of M^*, is a symmetric matrix up to row interchanges.*

5.3. **Case: n = prime** > 2. Again in this case, as a Hadamard matrix does not exist, we try for an orthogonal matrix with elements having at most two different absolute values. For a non-zero x, the value of $\binom{n}{x}$ and prime p are co-prime, therefore, the only way to construct such a matrix is when x=0. [Sylvester's conjecture]. Given that x=0, we assign the value of m as close as possible to $\frac{3p}{4}$. In this process, there will be almost $p - \frac{3p}{4} = \frac{p}{4}$ number of b's. With our assumption that x=0, more than one non-overlapping bs in any row need more than p rows. Therefore, the only possible value (p-m) can have is 1. With these conditions: x=0, p-m=1, we get $a = \pm\frac{2}{p},\ and\ b = \mp\frac{p-2}{p}$. The starting point for the iteration can be constructed in the following way,

$$M = \begin{bmatrix} -(p-2)/p & 2/p & ... & 2/p \\ 2/p & -(p-2)/p & ... & 2/p \\ ... & ... & ... & 2/p \\ 2/p & 2/p & ... & -(p-2)/p \end{bmatrix}.$$

The expression for entropy, in this case, can be given as below.

$$E^R[M(p)] = p\left[(p-1)\left(\frac{4}{p^2}\right) ln\left(\frac{p}{2}\right)^2 + \left(\frac{p-2}{p}\right)^2 ln\left(\frac{p}{p-2}\right)^2\right]$$

For large values of the prime p, the approximate value of entropy is given by

$$E^R[M(p)] \approx 8ln(\frac{p}{2}) + 4 \ll pln(p).$$

After iterations, the convergence and the 1st and second order conditions for local extrema guarantee that the value of the entropy is optimal. The matrix above does satisfy the and second order conditions given in lemma 4.3 and 4.4.

5.4. **Case: n=6.** Again in this case, as a Hadamard matrix does not exist, we try for an orthogonal matrix with elements having at most two different absolute values. Similar to the case for n = 5, we get the following conditions on m, x, a, and b.

For n = 6, i.e. m = 3, 4, 5 we have

$$x + m \approx \frac{3n}{4} = 4.5.$$

The nearest integers to x + m are 4 and 5.

$$x \leq \frac{(6-m)^2}{6}$$

For m = 4 or m = 5, we have x = 0.

- Case: m = 4 and x = 0; we solve for a and b to get $a = \pm\frac{2}{3\sqrt{2}}$ and $b = \mp\frac{1}{3\sqrt{2}}$.

  The construction of such an orthogonal matrix is possible in the following way,

$$M = \begin{bmatrix} -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & 1/3\sqrt{2} & -2/3\sqrt{2} & -2/3\sqrt{2} \\ 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & -2/3\sqrt{2} & 1/3\sqrt{2} & -2/3\sqrt{2} \\ 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & -2/3\sqrt{2} & -2/3\sqrt{2} & 1/3\sqrt{2} \\ -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} \\ 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} \\ 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} & 2/3\sqrt{2} & 2/3\sqrt{2} & -1/3\sqrt{2} \end{bmatrix}.$$

  The value of the entropy for the matrix above is $E^R \approx 9.95 < 6ln(6) \approx 10.75$. The 1st and second order conditions for local extrema guarantee that the value of the entropy is optimal. The matrix above does satisfy the and second order conditions given in lemma 4.3 and 4.4.

Observation: The desirable matrix here in the 6×6 case is actually the Kronecker product of our results for 2×2 and 3×3, i.e. the Kronecker product of the two matrices below:

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \otimes \begin{bmatrix} -1/3 & 2/3 & 2/3 \\ 2/3 & -1/3 & 2/3 \\ 2/3 & 2/3 & -1/3 \end{bmatrix}.$$

We write M(6) = M(2) $\otimes$ M(3).

- Case: m = 5 and x = 0; we solve for a and b to get $a = \pm\frac{1}{3}$ and $b = \mp\frac{2}{3}$.

  The construction of such an orthogonal matrix is possible in the following way,

$$M = \begin{bmatrix} -2/3 & 1/3 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & -2/3 & 1/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & -2/3 & 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & -2/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 1/3 & -2/3 & 1/3 \\ 1/3 & 1/3 & 1/3 & 1/3 & 1/3 & -2/3 \end{bmatrix}$$

  The value of the entropy for the matrix above is $E^R \approx 9.48 < 9.95$.

- Case: m = 3 and x = 0; we solve for a and b to get $a = \pm\frac{1}{\sqrt{3}}$ and b = 0 (a and b must be positive).

- Case: m = 3 and x = 1; we solve for a and b to get $a^2 = \frac{1}{17-8\sqrt{3}}$ and $b^2 = \frac{7-4\sqrt{3}}{17-8\sqrt{3}}$.

  In this case (even if matrix exists), $E^R \approx 8.11 < 9.48 < 9.95$. Therefore, our conclusion in this case is that the matrix M(6) = M(2) ⊗ M(3) is optimal.

5.5. **Case: n = 2p, where p is prime > 2.** Again in this case, as a Hadamard matrix does not exist, we construct an orthogonal matrix with two distinct elements. For x = 0 and m = 2p - 1, we get $a = \pm\frac{1}{p}$ and $b = \mp\frac{p-1}{p}$.

The construction is as follows,

$$M = \begin{bmatrix} -(p-1)p & 1/p & ... & 1/p \\ 1/p & -(p-1)/p & ... & 1/p \\ ... & ... & ... & ... \\ 1/p & 1/p & ... & -(p-1)/p \end{bmatrix}.$$

Entropy for the matrix above is given by,

$$E^R = 2p\left[(2p-1)\left(\frac{1}{p^2}\right) ln\left(\frac{p}{1}\right)^2 + \left(\frac{p-1}{p}\right)^2 ln\left(\frac{p}{p-1}\right)^2\right].$$

For a large prime p, we have an approximate expression for entropy as

$$E^R \approx 8ln(p) + 8ln(2) + 4 \ll 2pln(2p).$$

In an alternate construction, we divide the $2p \times 2p$ matrix into four $p \times p$ sub-matrices. Optimizing each sub-matrix (of prime order) is equivalent to optimizing a $p \times p$ matrix as in the previous case.

For x = 0 and m = 2p - 2, we get $a = \pm\frac{2}{p\sqrt{2}}$ and $b = \mp\frac{p-2}{p\sqrt{2}}$.

The construction of such a matrix is possible in the following way,

$$M = \begin{bmatrix} -(p-2)/p\sqrt{2} & \cdots & 2/p\sqrt{2} & (p-2)/p\sqrt{2} & \cdots & -2/p\sqrt{2} \\ 2/p\sqrt{2} & \cdots & 2/p\sqrt{2} & -2/p\sqrt{2} & \cdots & -2/p\sqrt{2} \\ \cdots & \cdots & 2/p\sqrt{2} & \cdots & \cdots & -2/p \\ 2/p\sqrt{2} & \cdots & -(p-2)/p\sqrt{2} & -2/p\sqrt{2} & \cdots & (p-2)/p\sqrt{2} \\ -(p-2)/p\sqrt{2} & \cdots & 2/p\sqrt{2} & -(p-2)/p\sqrt{2} & \cdots & 2/p\sqrt{2} \\ 2/p\sqrt{2} & \cdots & 2/p\sqrt{2} & 2/p\sqrt{2} & \cdots & 2/p\sqrt{2} \\ \cdots & \cdots & 2/p\sqrt{2} & \cdots & \cdots & 2/p \\ 2/p\sqrt{2} & \cdots & -(p-2)/p\sqrt{2} & 2/p\sqrt{2} & \cdots & -(p-2)/p\sqrt{2} \end{bmatrix}$$

Observation: M(2p) = M(2) $\otimes$ M(p)

The equation for entropy in this case can be given as follows,

$$E^R[M(2p)] = 2p\left[(2p-2)\left(\frac{2}{p^2}\right)ln\left(\frac{p^2}{2}\right) + 2\left(\frac{(p-2)^2}{2p^2}\right)ln\left(\frac{2p^2}{(p-2)^2}\right)\right].$$

For large values of a prime p, the approximate value of entropy is given by

$$E^R[M(2p)] \approx 16ln(p) + (2p-8)ln(2) + 8 \ll 2pln(2p).$$

For large prime numbers, we compare the values in the two cases,

$$E^R[M(2p)] \approx 16ln(p) + (2p-8)ln(2) + 8 \gg E^R \approx 8ln(p) + 8ln(2) + 4 \ll 2pln(2p).$$

In this case, we conclude that the Kronecker product, M(2p) = M(2) $\otimes$ M(p) has entropy more than that of the matrix generated by m = 2p - 1 (because m = 2p - 2 is closer to 3(2p)/4 than that of 2p - 1).

5.6. **Case: n = 3p, where p is prime > 3.** Again in this case, as a Hadamard matrix does not exist. We construct an orthogonal matrix with two distinct elements. For x = 0 and m = 3p - 1, we get $a = \pm\frac{2}{3p}$ and $b = \mp\frac{3p-2}{3p}$.

The construction is as follows,

$$M = \begin{bmatrix} -(3p-2)/3p & 2/3p & ... & 2/3p \\ 2/3p & -(3p-2)/3p & ... & 2/3p \\ ... & ... & ... & ... \\ 2/3p & 2/3p & ... & -(3p-2)/3p \end{bmatrix}.$$

Entropy for the matrix above is given by,

$$E^R = 3p\left[(3p-1)\left(\frac{4}{9p^2}\right) ln\left(\frac{3p}{2}\right)^2 + \left(\frac{3p-2}{3p}\right)^2 ln\left(\frac{3p}{3p-2}\right)^2\right].$$

For a large prime p, we have an approximate expression for entropy as

$$E^R \approx 8ln(3p) - 8ln(2) + 4 \ll 3pln(3p).$$

In an alternate construction, we divide the 3×3 matrix into nine $p \times p$ sub-matrices. Optimizing each sub-matrix (of prime order) is equivalent to optimizing a $p \times p$ matrix as in the previous case.

For x=0, we use four different entries as

$$a = \pm\frac{4}{3p},\ b = \mp\frac{2}{3p}, c = \mp\frac{2(p-2)}{3p}, d = \pm\frac{p-2}{3p}.$$

Where (a,b,c,d) occur ($2p-2$, $p-1$, 2, 1) times respectively. In fact, we obtain M(3p) = M(3) $\otimes$ M(p). The expression for entropy, in this case, is given below.

$$\begin{aligned} E^R[M(3p)] =& 3p\left[(2p-2)\left(\frac{16}{9p^2}\right)ln\left(\frac{9p^2}{16}\right) + (p-1)\left(\frac{4}{9p^2}\right)ln\left(\frac{9p^2}{4}\right)\right. \\ &\left. +2\left(\frac{2(p-2)}{3p}\right)^2 ln\left(\frac{3p}{2(p-2)}\right)^2 + \left(\frac{p-2}{3p}\right)^2 ln\left(\frac{3p}{p-2}\right)^2\right] \end{aligned}$$

For large values of a prime p, the approximate value of entropy is given by

$$E^R[M(3p)] \approx 24ln(p) + 6pln\left(\frac{3}{2}\right) + 12.$$

For large prime numbers, we compare the values in the two cases,

$$E^R[M(3p)] \approx 24ln(p) + 6pln\left(\frac{3}{2}\right) + 12 \gg E^R \approx 8ln(3p) - 8ln(2) + 4.$$

In this case also, we conclude that the Kronecker product M(3p) = M(3) $\otimes$ M(p) has more entropy than that of the matrix generated by m = 3p - 1.

5.7. **General Case: $\mathbf{n = p_1p_2}$, where $\mathbf{p_1}$ and $\mathbf{p_2}$ are primes with $\mathbf{p_1 < p_2}$.** A construction for x = 0 and $m = p_1p_2 - 1$, can be done using $a = \pm\frac{2}{p_1p_2}$ and $b = \mp\frac{p_1p_2-2}{p_1p_2}$ as follows,

$$M = \begin{bmatrix} -(p_1p_2-2)/p_1p_2 & 2/p_1p_2 & \cdots & 2/p_1p_2 \\ 2/p_1p_2 & -(p_1p_2-2)/p_1p_2 & \cdots & 2/p_1p_2 \\ \cdots & \cdots & \cdots & \cdots \\ 2/p_1p_2 & 2/p_1p_2 & \cdots & -(p_1p_2-2)/p_1p_2 \end{bmatrix}.$$

Entropy for the matrix above is given by,

$$E^R = p_1p_2\left[(p_1p_2-1)(\frac{2}{p_1p_2})^2 ln\left(\frac{p_1p_2}{2}\right)^2 + \left(\frac{p_1p_2-2}{p_1p_2}\right)^2 ln\left(\frac{p_1p_2}{p_1p_2-2}\right)^2\right].$$

For large primes, we have an approximate expression for entropy as

$$E^R \approx 8ln(p_1p_2/2) + 4 \ll p_1p_2 ln(p_1p_2).$$

In our alternate construction, we divide the $p_1p_2 \times p_1p_2$ matrix into p_1^2 $p_2 \times p_2$ sub-matrices. Optimizing each sub-matrix (of prime order) is equivalent to optimizing a $p_2 \times p_2$ matrix as in the previous cases. We use four different entries as

$$a = \pm\frac{4}{p_1p_2}, b = \mp\frac{2(p_1-1)}{p_1p_2}, c = \mp\frac{2(p_2-2)}{p_1p_2}, d = \pm\frac{(p_1-2)(p_2-2)}{p_1p_2},$$

where (a, b, c, d) occur $\{(p_1-2)(p_2-2), p_2-1, p_1-1, 1\}$ times respectively. We obtain $M(p_1p_2) = M(p_1) \otimes M(p_2)$. The expression for entropy, in this case, can be given as

$$\begin{aligned} E^R[M(3p)] =& p_1p_2 \Bigg[(p_1-2)(p_2-2)(\frac{4}{p_1p_2})^2 ln\left(\frac{p_1p_2}{4}\right)^2 + \\ & (p_2-1)\left(\frac{2(p_1-1)}{p_1p_2}\right)^2 ln\left(\frac{p_1p_2}{2(p_1-1)}\right)^2 + \\ & (p_1-1)\left(\frac{2(p_2-1)}{p_1p_2}\right)^2 ln\left(\frac{p_1p_2}{2(p_2-1)}\right)^2 + \\ & \left(\frac{(p_1-2)(p_2-2)}{p_1p_2}\right)^2 ln\left(\frac{p_1p_2}{(p_1-2)(p_2-2)}\right)^2 \Bigg]. \end{aligned}$$

For large values of the primes, the approximate value of entropy is given by

$$E^R[M(p_1p_2)] \approx 32ln(p_1p_2) + 8p_1 ln\left(\frac{p_2}{2}\right) + 8p_2 ln\left(\frac{p_1}{2}\right) + 4(p_1+p_2)$$

This value of entropy is significantly larger than $8ln(p_1p_2/2) + 4$.

The Kronecker product $M(p_1p_2) = M(p_1) \otimes M(p_2)$ has entropy more than the matrix generated by $m = p_1p_2 - 1$. It is easy to show that the Kronecker product obtained in this case is an orthogonal matrix.

For two orthogonal matrices $M_1 = M(p_1)$, and $M_2 = M(p_2)$ we have that, $M_1M_1^T = I_{p_1}$ and $M_2M_2^T = I_{p_2}$. The Kronecker product $M = M(p_1p_2) = M(p_1) \otimes M(p_2) = M_1 \otimes M_2$.

$$MM^T = (M_1 \otimes M_2)(M_1 \otimes M_2)^T = (M_1 \otimes M_2)(M_1^T \otimes M_2^T) = (M_1M_1^T) \otimes (M_1M_1^T) = I_{p_1} \otimes I_{p_1} = I_{p_1p_2}$$

Similarly, $M^TM = I_{p_1p_2}$, therefore the matrix $M = p_1p_2$ is an orthogonal matrix. In general the Kronecker product of two matrices is not commutative; however the value of entropy of a Kronecker product is independent of the order, i.e.

$$E^R[M(p_1p_2)] = E^R[M(p_1) \otimes M(p_2)] = E^R[M(p_2) \otimes M(p_1)]$$

5.8. **A Generalization for a Square Free Number:**

Conjecture 2. *Let n be a square free number. Let $n = (p_1)(p_2)\cdots(p_k)$, where $p_1, p_2, \cdots, p_k$ are distinct prime numbers. The optimal matrix of order n is given by*

$$E^R[M(n)] = E^R[M(p_1) \otimes M(p_2) \otimes ... \otimes M(p_k)].$$

5.9. **Construction of Entropy Optimal Complex Inverse Orthogonal Matrix:** Applying Jensen's inequality, as in the case of real orthogonal matrices, we get similar bounds on the entropy of a complex inverse orthogonal matrix,

$$0 \leq E^c[M(n)] \leq nlogn.$$

This is a general bound that works for all n, and equality is achieved if and only if $|a_{ij}|^2 = \frac{1}{n}$ Or in polar form, we write $a_{ij} = \frac{1}{\sqrt{n}} e^{i\theta}$, where $e^{i\theta}$ is a root of unity.

For n = 2, we have the following optimal matrix which achieves the bound $2ln(2)$.

$$M(2) = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

For n = 3, we construct the optimal matrix as follows,

$$M(3) = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & \omega/\sqrt{3} & \omega^2/\sqrt{3} \\ 1/\sqrt{3} & \omega^2/\sqrt{3} & \omega/\sqrt{3} \end{bmatrix}.$$

For the matrix above, $E^c[M(3)] = 3ln(3)$.

For general n: The optimal value $nln(n)$ is achieved by the $n \times n$ scaled Vandermonde matrix (V_n).

$$M(n) = \left(\frac{1}{\sqrt{n}}\right) V_n$$

Where V_n is given as the matrix below,

$$M = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \alpha & \alpha^2 & \cdots & \alpha^{n-1} \\ 1 & \alpha^2 & \alpha^4 & \cdots & \alpha^{2n-2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & \alpha^{n-1} & \alpha^{2n-2} & \cdots & \alpha^{(n-1)^2} \end{bmatrix},$$

where α is the n^{th} root of unity ($\alpha \neq 1$). It is easy to check that the optimal entropy is achievable for all n, for the matrix above. $E^c[M(n)] = nln(n)$.

6. Acknowledgements

I thank Dr. Miller and Dr. Mercer for participating in this committee, and for the support they have given. I especially thank Dr. Arasu for his advice, his help in these 2 years, and the passion with which he teaches. I am honored to have had him as an advisor, and feel incredibly lucky to have had the opportunity to work with such an amazing mentor. I would like to thank the entire WSU Department of Mathematics and Statistics for their support in the last two years. I also thank NSF for the support in my research.

References

[1] Shannon, Claude E. (July/October 1948). *A Mathematical Theory of Communication.* Bell System Technical Journal 27 (3): 379423.

[2] Keynes, John Maynard (1921), *Chapter IV. The Principle of Indifference.* A Treatise on Probability, 4, Macmillan and Co., pp. 4164.

[3] *The Principle of Sufficient Reason: A Reassessment by Alexander R. Pruss*

[4] H Gopalkrishna Gadiyar1, K M Sangeeta Maini1, R Padma1, and H S Sharatchandra2. *Entropy and Hadamard matrices*, Journal of Physics A: Mathematical and General, 2003.

[5] TW Anderson, I Olkin, LG Underhill. *Generation of Random Orthogonal Matrices.* Technical Report No. 6, August 1985.

[6] *Equivalence Classes of Inverse Orthogonal and Unit Hadamard Matrices*, R. Craigen, Bull. Austral. Math. Soc. Vol. 44 (1991) [109-115]

[7] Beth, Thomas; Jungnickel, Dieter; Lenz, *Hanfried Design Theory. Vol. I, Second edition.* Encyclopedia of Mathematics and its Applications, 69. Combridge University Press, Cambridge, 1999.

[8] Maz'ya, Vladimir; Shaposhnikova, T. O. (1999). *Jacques Hadamard: A Universal Mathematician.* AMS. pp. 383ff.. ISBN 0-8218-1923-2.

[9] okovi, Dragomir (2008). *Hadamard matrices of order 764 exist.* Combinatorica 28 (4): 487489.

[10] Trinh, Q. K.; Fan, P.; Gabidulin, E. M., *Multilevel Hadamard matrices and zero correlation zone sequences.* Electron. Lett. 42 (2006), no. 13, 748-750.

[11] Parametrization of an orthogonal matrix in terms of generalized eulerian angles. Richard C. Raffenetti, Klaus Ruedenberg, Article first published online: 18 JUN 2009.

[12] Raymond Puzio, yark, Chi Woo. *Cayley's parameterization of orthogonal matrices.* (version 17). PlanetMath.org. Freely available at http://planetmath.org/CayleysParameterizationOfOrthogonalMatrices.html.

[13] *Special Orthogonal Group SO(N), Wave Equations in Higher Dimensions* Springer, Dong, S.-H.

[14] *A Feasible Method For Optimization With Orthogonality Constraints.* Zaiwen Wen and Wotao Yin, 2010. In Rice Univ. Technical Report10.

[15] Crank, J.; Nicolson, P. (1947). *A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type.* Proc. Camb. Phil. Soc. 43 (1): 5067.