Bohemian Matrices: the Symbolic Computation Approach

L. Gonzalez-Vega,
J. Sendra & J. R. Sendra
CUNEF, UPM & UAH
Spain

With the help of
R. M. Corless, E. Y. S Chan & S. Thornton

OUTLINE

- Bohemian Matrices
- Bohemian Correlation Matrices:
  - **#($BCM_{n:\{0,1\}}$)**?
- Correlation Matrices: characterisation.
- **#($BCM_{n:\{-1,0,1\}}$)**?
- Final questions.

Bohemian matrices

http://www.bohemianmatrices.com

Bohemian matrices

A family of Bohemians [BOunded HEight Matrix of Integers] is a set of matrices where the free entries are from the finite population *P*.

The family of 6×6 matrices with population $\{-1, +1\}$. Here are twc instances out of the $2^{6^2} = 68.719.476.736$ possible:

$$\begin{bmatrix} 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 & 1 \\ -1 & 1 & -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 & 1 & 1 \end{bmatrix} \quad \begin{bmatrix} -1 & -1 & -1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & 1 & 1 \end{bmatrix}$$

Bohemian matrices

For a given dimension, population and characteristics, the set of Bohemian matrices is finite and these are examples of typical questions we want to answer:

- How many of them are singular?
- What is the maximum determinant?
- How many distinct characteristic polynomials does the family have?
- How many distinct eigenvalues does the family contain?
- What is the distribution of the number of different real eigenvalues? Patterns?

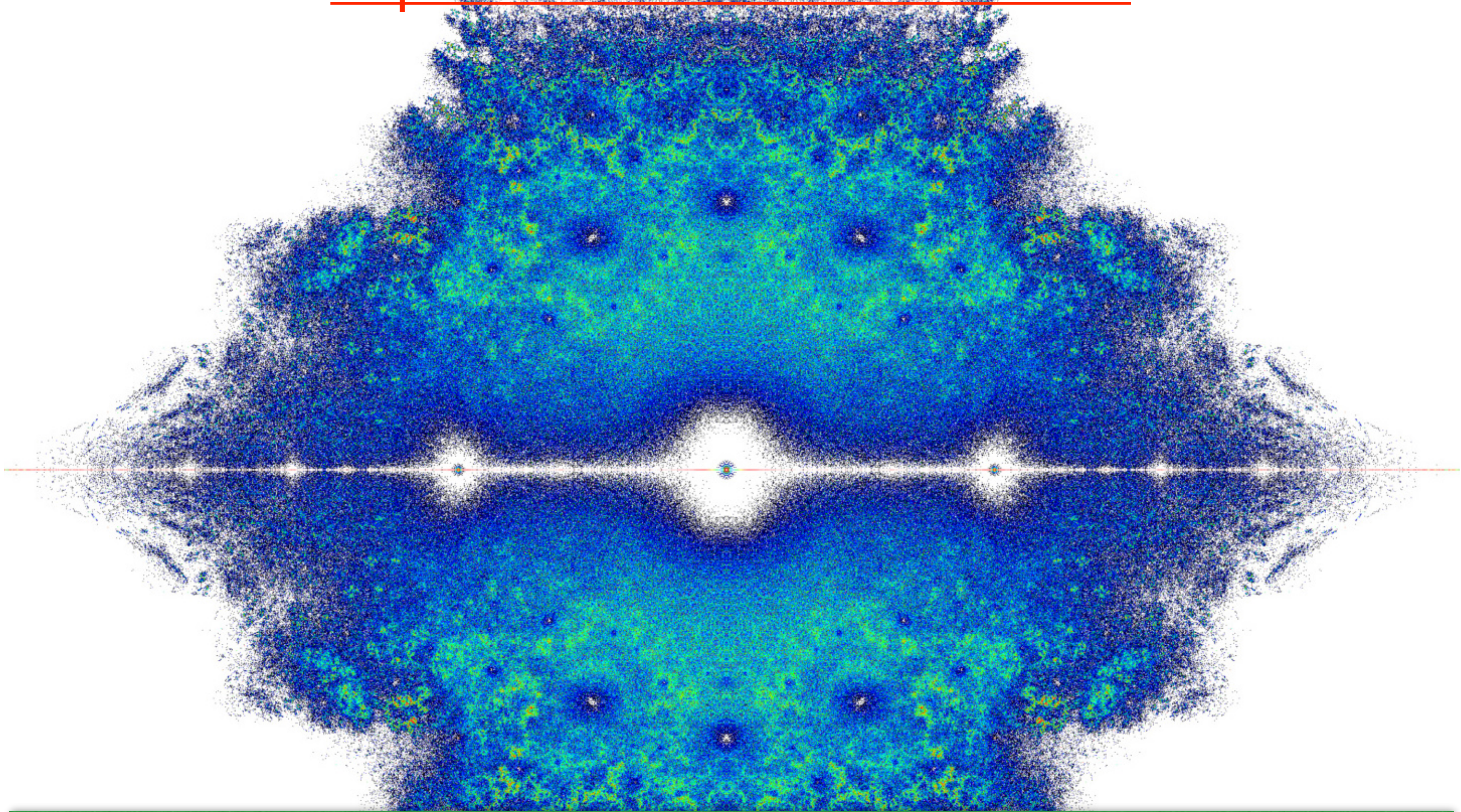
Bohemian matrices

By brute-force computation on all $2^{36} = 68.719.476.736$ matrices, there are **43.090.149.376** singular matrices.

Language	Time
Maple[5]	270 days
Matlab	10 days
Julia	31 hours
Python (sequential)	20 days
Python (batched)	17 hours
C++	4.75 hours
CPDB	**124ms**

CPDB: Characteristic Polynomial Database

http://www.bohemianmatrices.com



A density plot in the complex plane of the Bohemian eigenvalues of a sample of 100 million *15x15* tridiagonal matrices. The entries are sampled from *{-1, 0, 1}*. Color represents the eigenvalue density and the plot is viewed on *[-3-3i, 3+3i]*. Plot produced by Cara Adams.

Bohemian matrices

Motivations & Applications:

- Software testing. We have found bugs in major packages (Steven Thornton has computed many many many … eigenvalues).
- Understanding Random Matrices (Random Polynomials by A.T. Bharucha-Reid & M. Sambandham, Chapter 3, 1986).
- Our original Bohemian family, the Mandelbrot matrices invented by Piers Lawrence, has given rise to a genuinely new kind of companion matrix (Chan & Corless @ ELA 32 (2017)), and to what we now call *Algebraic Linearisations* (Chan et al @ LAA 563 (2019), 373–399) for *Non Linear Eigenvalue Problems* (solving *det(A(x))=0*).
- Many connections to combinatorics and graph theory.

On the number of Correlation Matrices in the set of NxN Bohemian matrices when N is fixed

$$\#(BCM_{n:\{0,1\}})$$

Problem at hand

Data:

- Population: *0* and *1* or *-1, 0* and *1 or -1* and *1*.
- Type of matrices: *n x n* Bohemian Matrices.
- $\mathbf{BM}_{n:\{0,1\}}$ or $\mathbf{BM}_{n:\{-1,0,1\}}$ or $\mathbf{BM}_{n:\{-1,1\}}$.

Output:

- For every *n*, compute the number of correlation matrices in the set $\mathbf{BM}_{n:\{0,1\}}$: $\mathbf{BCM}_{n:\{0,1\}}$

An $n \times n$ symmetric matrix A is a **correlation matrix** if

- It has ones on the diagonal.
- All its eigenvalues are nonnegative.

Applications:

- Population in the closed interval [-1,1].

Problem at hand

Looking for $\{0, 1\}$ correlation matrices.

n	total matrices	# correl	proportion correl
3	8	5	6.25e-1
4	64	15	2.34e-1
5	1024	52	5.08e-2
6	32768	203	6.20e-3
7	2097152	877	4.18e-4

Nick Higham's table from Manchester Bohemian Workshop, 2018

Out of the 2^binomial(**8**,2) = 268.435.456 possibilities, there are only **4140** correlation matrices giving a proportion of 1.54e-5. Computing time was 24 hours and a few minutes.

$\#(BCM_{8:\{0,1\}}) = 4140$

The solution

THE ON-LINE ENCYCLOPEDIA OF INTEGER SEQUENCES ®

founded in 1964 by N. J. A. Sloane

5, 15, 52, 203, 877, 4140 Search Hints

(Greetings from The On-Line Encyclopedia of Integer Sequences!)

Search: **seq:5,15,52,203,877,4140**

Displaying 1-10 of 40 results found. page 1 2 3 4

Sort: relevance | references | number | modified | created Format: long | short | data

A000110 Bell or exponential numbers: number of ways to partition a set of n labeled elements. (Formerly M1484 N0585) +30 954

1, 1, 2, **5, 15, 52, 203, 877, 4140**, 21147, 115975, 678570, 4213597, 27644437, 190899322, 1382958545, 10480142147, 82864869804, 682076806159, 5832742205057, 51724158235372, 474869816156751, 4506715738447323, 44152005855084346, 445958869294805289, 4638590332229999353, 49631246523618756274 (list; graph; refs; listen; history; text; internal format)

$$B_n = \frac{1}{e}\sum_{k=0}^{\infty}\frac{k^n}{k!} \qquad B_{n+1} = \sum_{k=0}^{n}\binom{n}{k}B_k \qquad \sum_{n=0}^{\infty}\frac{B_n}{n!}x^n = e^{e^x-1}$$

The solution

a(n+1) is the number of (symmetric) positive semidefinite n X n 0-1 matrices. These correspond to equivalence relations on {1,...,n+1}, where matrix element M[i,j] = 1 if and only if i and j are equivalent to each other but not to n+1. - Robert Israel, Mar 16 2011

How computations were performed?

- First approach with Matlab for computing eigenvalues giving wrong results for n=7 due to precision problems.

- Second approach: Maple but avoiding eigenvalue computations (see later).

Correlation Matrices: Characterisations

The characterisation

Theorem

Let A be a symmetric $n \times n$ matrix. Then we have:

- *A is positive definite $\Leftrightarrow D_k > 0$ for all leading principal minors*
- *A is negative definite $\Leftrightarrow (-1)^k D_k > 0$ for all leading principal minors*
- *A is positive semidefinite $\Leftrightarrow \Delta_k \geq 0$ for all principal minors*
- *A is negative semidefinite $\Leftrightarrow (-1)^k \Delta_k \geq 0$ for all principal minors*

Too many principal minors to check: *binomial(n,k)* for $1 \leq k \leq n$

The characterisation

All eigenvalues of A are nonnegative

if and only if

A is positive semidefinite

if and only if

Alternating signs for the coefficients of the characteristic polynomial of A

Proof: *A* is symmetric, real and Descartes Rule of Signs

The characterisation

For a sequence of real numbers $b_0, b_1, \ldots, b_n$, $\mathbf{Var}(b_0, b_1, \ldots, b_n)$ will denote the number of sign changes in $b_0, b_1, \ldots, b_n$ after dropping the zeros in the sequence.

Descartes' Rule of Signs:

Let P be the polynomial in $\mathbb{R}[x]$:

$$P(x) = \sum_{k=0}^{n} a_k x^k.$$

The number of positive real roots of $P(x) = 0$, counted with multiplicity, is equal to

$$\mathbf{Var}(a_n, a_{n-1}, \ldots, a_0) - 2k$$

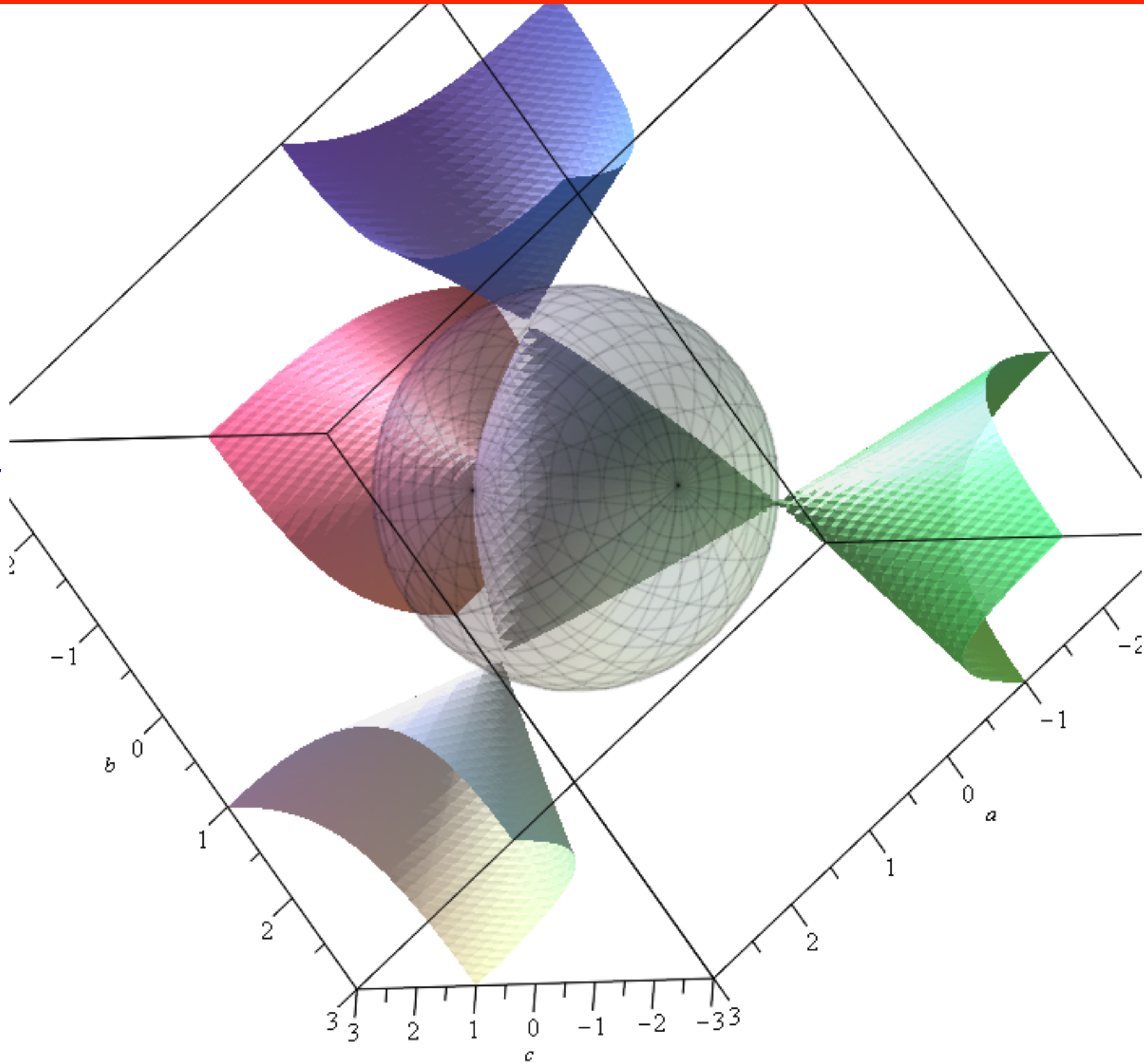
for some non–negative integer k.

When all roots are known to be real, Descartes' Rule of Signs is exact (taking into account the multiplicities). And this is the case !

Example: n=3

$$\begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}$$

$\lambda^3 - 3\lambda^2 -$



$c^2 - 1$

Example: n=3

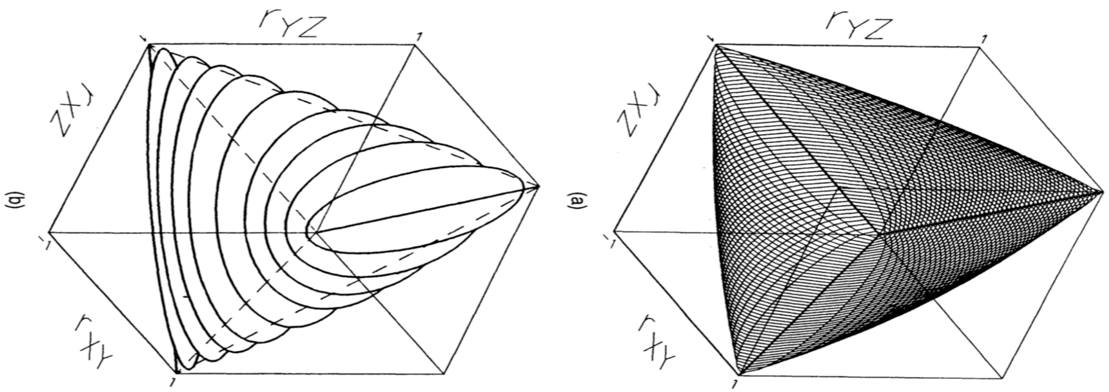
$$\begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}$$

$$\lambda^3 - 3\lambda^2 - (a^2 + b^2 + c^2 - 3)\lambda - 2abc + a^2 + b^2 + c^2 - 1$$

Correlation matrices. Characterisation when n=3:

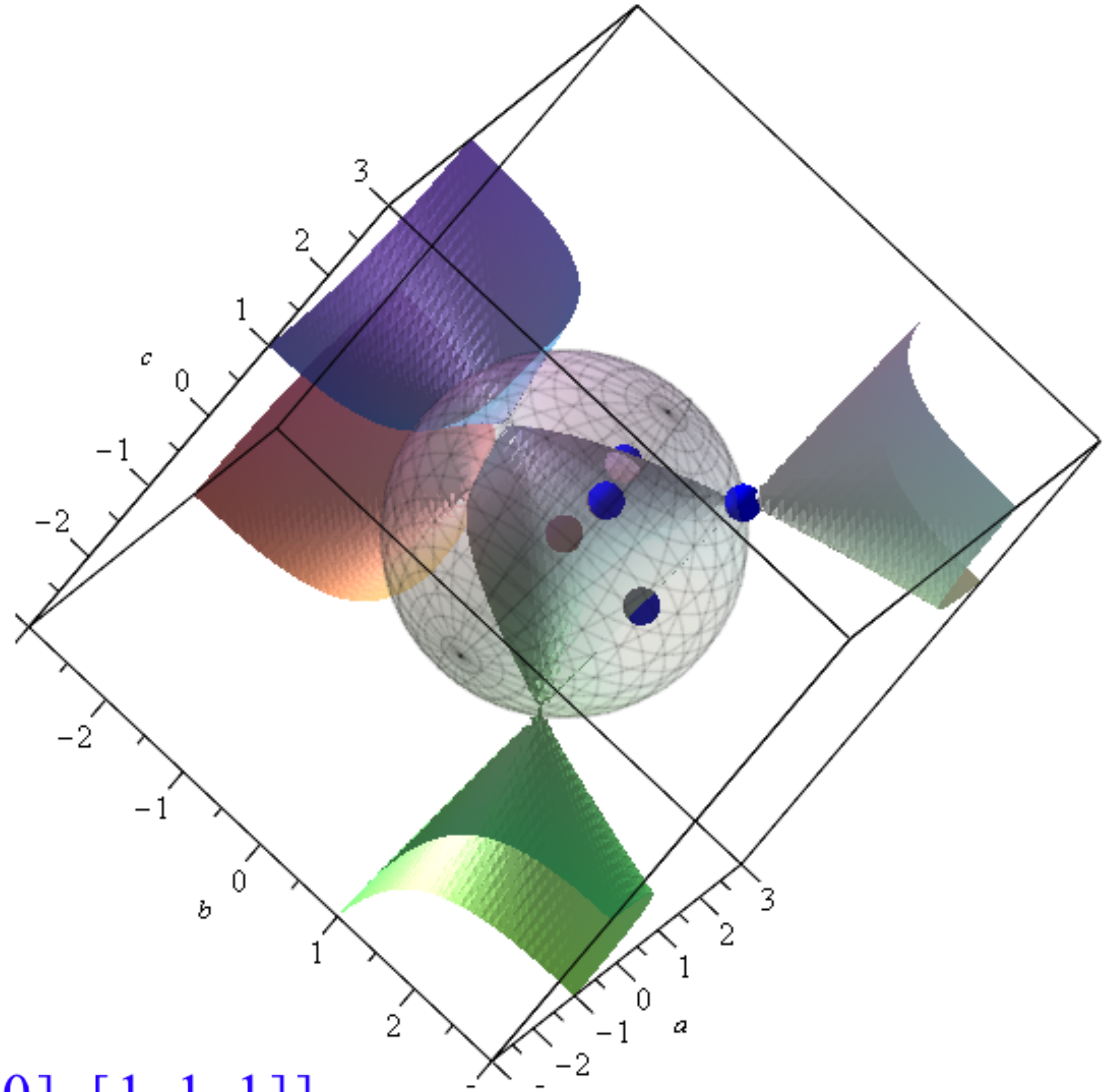
$$[a^2 + b^2 + c^2 - 3 \leq 0,\ -2abc + a^2 + b^2 + c^2 - 1 \leq 0]$$

P. J. Rousseeuw and G. Molenberghs: *The Shape of Correlation Matrices*. The American Statistician 48, 276-279, 1994.



Example: n=3

$$\begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}$$

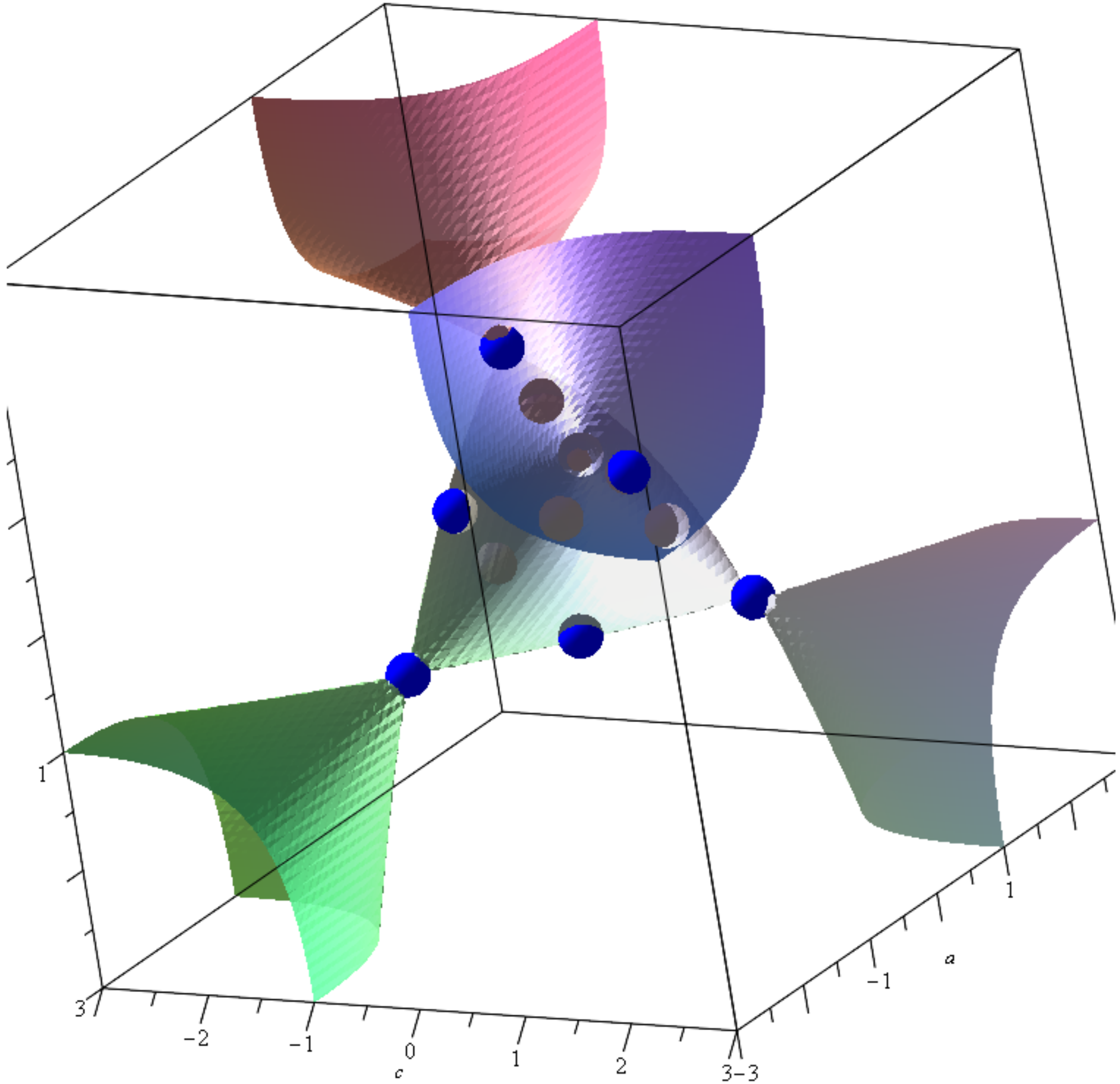


BCM:{0,1}

[[0, 0, 0], [0, 0, 1], [0, 1, 0], [1, 0, 0], [1, 1, 1]]

Example: n=3

$$\begin{bmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{bmatrix}$$



BCM$_{3:\{-1,0,1\}}$

$$[[-1, -1, 1], [-1, 0, 0], [-1, 1, -1], [0, -1, 0], [0, 0, -1], [0, 0, 0], [0, 0, 1], [0, 1, 0], [1, -1, -1], [1, 0, 0], [1, 1, 1]]$$

Example: n=4

$$\begin{bmatrix} 1 & a & b & d \\ a & 1 & c & e \\ b & c & 1 & f \\ d & e & f & 1 \end{bmatrix}$$

Correlation matrices. Characterisation when *n* = *4*:

$$-a^2 - b^2 - c^2 - d^2 - e^2 - f^2 + 6 \geq 0$$

$$-2\,a\,b\,c - 2\,a\,d\,e - 2\,b\,d\,f - 2\,c\,e\,f + 2\,a^2 + 2\,b^2 + 2\,c^2 + 2\,d^2 + 2\,e^2 + 2\,f^2 - 4 \leq 0$$

$$a^2 f^2 - 2\,a\,b\,e\,f - 2\,a\,c\,d\,f + b^2 e^2 - 2\,b\,c\,d\,e + c^2 d^2 + 2\,a\,b\,c + 2\,a\,d\,e + 2\,b\,d\,f + 2\,c\,e\,f - a^2 - b^2 - c^2 - d^2 - e^2 - f^2 + 1 \geq 0$$

Example: n=4

Correlation matrices. Characterisation when *n*=*4*:

$$-a^2 - b^2 - c^2 - d^2 - e^2 - f^2 + 6 \geq 0$$

$$-2\,a\,b\,c - 2\,a\,d\,e - 2\,b\,d\,f - 2\,c\,e\,f + 2\,a^2 + 2\,b^2 + 2\,c^2 + 2\,d^2 + 2\,e^2 + 2\,f^2 - 4 \leq 0$$

$$a^2 f^2 - 2\,a\,b\,e\,f - 2\,a\,c\,d\,f + b^2 e^2 - 2\,b\,c\,d\,e + c^2 d^2 + 2\,a\,b\,c + 2\,a\,d\,e + 2\,b\,d\,f + 2\,c\,e\,f - a^2 - b^2 - c^2 - d^2 - e^2 - f^2 + 1 \geq 0$$

What about the geometry of this set ?

Experiments (continued)

Maple based:

- First case not in Nick's table: *n=8*.
- The coefficients of the characteristic polynomial are easy to compute and **easier to evaluate** when population are integer numbers, rational numbers, ….. .

On the number of Correlation Matrices in the set of NxN Bohemian matrices when N is fixed

$\#(\mathbf{BCM}_{n:\{-1,0,1\}})$, $\#(\mathbf{BCM}_{n:\{-1,1\}})$, ...

More experiments

n	$BM_{n:\{-1,0,1\}}$	$BCM_{n:\{-1,0,1\}}$	%
3	27	11	40.74%
4	729	49	6.72%
5	59049	257	0.43%
6	14348907	1282	0.0089 %

11,49,257,1282 Search Hints

(Greetings from The On-Line Encyclopedia of Integer Sequences!)

Search: **seq:11,49,257,1282**

Sorry, but the terms do not match anything in the table.

More properties

- All matrices in $\mathbf{BCM}_{n:\{-1,0,1\}}$ except the identity are singular.
- The eigenvalues of the matrices in $\mathbf{BCM}_{n:\{-1,0,1\}}$ belong to the set

$$\{0, 1, 2, \ldots, n\}$$

  and there is always, at least, one multiple eigenvalue.
- All matrices in $\mathbf{BCM}_{n:\{-1,1\}}$ have the same characteristic polynomial:

$$\lambda^n - n\,\lambda^{n-1} = \lambda^{n-1}(\lambda - n)$$

- $\#(\mathbf{BCM}_{n:\{-1,1\}}) = 2^{n-1}$.
- $\mathbf{BCM}_{n:\{1,0\}}$ encodes Bell numbers (partitions of a set) and their characteristic polynomials encode the partitions of n.

Final Questions

Ongoing work

Many questions to answer yet !

- $\#(\mathbf{BCM}_{n:\{-1,0,1\}})$?
- To understand the "inequalities"?
- How to use these inequalities to deal with the Correlation Matrix Completion Problem?

- The Bohemian Matrices and Applications Workshop (Manchester 2018 organised by Rob Corless and Nick Higham):
  - https://www.maths.manchester.ac.uk/~higham/conferences/bohemian.php
- The Bohemian Matrix Minisymposium at ICIAM 2019 (Valencia) organised by Rob Corless and Nick Higham.

http://www.bohemianmatrices.com

Thanks !