Several questions about tensors

Fuzhen Zhang

Nova Southeastern University, Fort Lauderdale, Florida, USA

zhang@nova.edu

Presentation at AORC, Sungkyunkwan University, South Korea
May 2017

What is a tensor?

Tensor algebra

Tensor analysis

$\vdots$

What is a tensor?

A *Tensor* is an element of a *tensor space*
just like
a *vector* is an element of a *vector space*.

A vector in an n-dimensional space is represented by a one-dimensional array of length n with respect to a given basis:

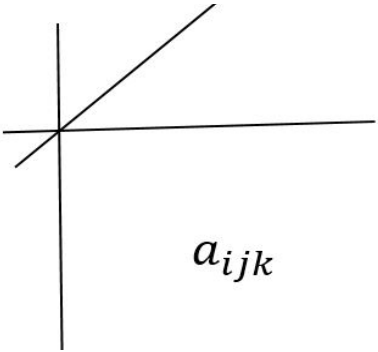
$$v = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n \longrightarrow (a_1, a_2, \ldots, a_n)$$

A tensor with respect to a basis is represented by a multi-dimensional array. For example, a linear transformation is represented in a basis as a two-dimensional square $n \times n$ array:

$$(a_{ij})$$

$n \times n \times n$ Tensors

$$A = (a_{ijk})$$



Like a *vector* in a vector space, a *tensor* is an element in a tensor space. In multilinear algebra, we begin with the **Cartesian space**

$$\begin{array}{lccccccccc} \text{Spaces}: & V_1 & \times & V_2 & \times & \cdots & \times & V_m & \mapsto & W \\ \text{Dim}: & n_1 & & n_2 & & \cdots & & n_m & & \\ \text{Bases}: & \{e_{1i_1}\} & & \{e_{2i_2}\} & & \cdots & & \{e_{mi_m}\} & & \end{array}$$

$$v = \left(\sum_{i_1=1}^{n_1} x_{1i_1} e_{1i_1}, \sum_{i_2=1}^{n_2} x_{2i_2} e_{2i_2}, \dots, \sum_{i_m=1}^{n_m} x_{mi_m} e_{mi_m} \right)$$

$$f(v) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_m=1}^{n_m} x_{1i_1} x_{2i_2} \cdots x_{mi_m} f(e_{1i_1}, e_{2i_2}, \dots, e_{mi_m})$$

$$f(v) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_m=1}^{n_m} x_{1i_1} x_{2i_2} \cdots x_{mi_m} f(e_{1i_1}, e_{2i_2}, \ldots, e_{mi_m})$$

$$x_{i_1 i_2 \ldots i_m} = x_{1i_1} x_{2i_2} \cdots x_{mi_m} \in \mathbb{F}$$

$$w_{i_1 i_2 \ldots i_m} = f(e_{1i_1}, e_{2i_2}, \ldots, e_{mi_m}) \in W$$

If f is a multilinear map s.t $\dim\langle \mathrm{Im}(f)\rangle = \prod_{t=1}^{m} n_t$, then f is said to be a *tensor map*, denoted by $\otimes$. The elements in $\langle \mathrm{Im}(\otimes)\rangle$ are called *tensors*. The elements in $\mathrm{Im}(\otimes)$ are *decomposable* tensors:

$$v_1 \otimes v_2 \otimes \cdots \otimes v_m \in W = V_1 \otimes V_2 \otimes \cdots \otimes V_m$$

$$(x_{i_1 i_2 \ldots i_m}), \quad 1 \le i_t \le n_t, \quad t = 1, 2, \ldots, m$$

$$f(v) = \sum x_{i_1 i_2 \ldots i_m} e_{1i_1} \otimes e_{2i_2} \otimes \cdots \otimes e_{mi_m}$$

Universal Factorization Property

$$\begin{tikzcd} V_1 \times \cdots \times V_m \arrow[r, "\otimes"] \arrow[dr, "f"'] & V_1 \otimes \cdots \otimes V_m \arrow[d, "T"] \\ & W \end{tikzcd}$$

$$f = T\otimes$$

In algebra, consider the **Cartesian product** $V_1 \times V_2 \times \cdots \times V_m$ as a **set**. Every set freely generates an R-module $\mathcal{F}$. Embed

$$V_1 \times V_2 \times \cdots \times V_m \hookrightarrow \mathcal{F}$$

Let

$$N = \langle\{(v_1, \ldots, \alpha v_k + \beta v_k', \ldots) - \alpha(v_1, \ldots, v_k, \ldots) - \beta(v_1, \ldots, v_k', \ldots)\}\rangle$$

The quotient space is called the tensor product space of the V_i's:

$$\mathcal{F}/N = V_1 \otimes \cdots \otimes V_m$$

Quadratic form and tensor

Quadratic form

$$f(x) = x^t A x = \sum_{i,j=1}^{n} a_{ij} x_i x_j,$$

where $x = (x_1, \dots, x_n)$, $A = (a_{ij})$ is a symmetric matrix.
An n-dimensional homogeneous polynomial form of degree m, $f(x)$, is equivalent to the tensor product of a supersymmetric n-dimensional tensor A of order m, and the rank-one tensor x^m:

$$f(x) = Ax^m := \sum_{i_1,\dots,i_m=1}^{n} a_{i_1,\dots,i_m} x_{i_1} \cdots x_{i_m}$$

Tensors as mapping

$$\mathcal{A} : \langle n_1 \rangle \times \langle n_2 \rangle \times \cdots \times \langle n_m \rangle \mapsto \mathbb{F}$$

$$\mathcal{A}(i_1, i_2, \cdots i_m) = a_{i_1 i_2 \cdots i_m}$$

$$(a_0), \quad (a_i), \quad (a_{ij}), \quad (a_{ijk}), \quad \ldots$$

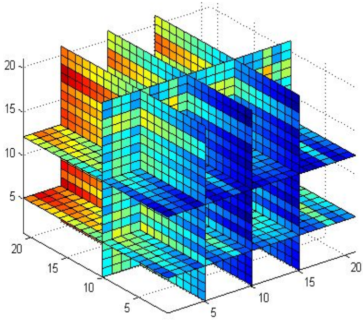
Tensors, hypermatrices, multidimensional array, cubes

$$\left(a_{i_1 i_2 \cdots i_m}\right)$$

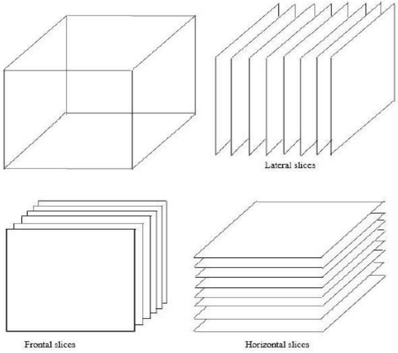
- **Hypermatrix**: Lim, Handbook of Linear Alg., 2nd ed., 2014
- **Tensor**: Cui, Li, Ng, SIAM. J. Matrix Anal. Appl., 2014
- K.C. Chang on **nonnegative tensors**, 2013
- L.Q. Qi research on tensors since 2000+
- Semi-magic **cube**: Ahmed et al, Discrete and Computational Geometry Algorithms and Combinatorics, 2003
- Stochastic **cubes**, Gupta and Nath, 1973
- **Multidimensional matrices**, Brualdi and Csima 1970s
- **Higher dimensional configurations,** Jurkat and Ryser 1968s

Applications of tensors

- Almost everywhere in Math and Physics
- Computer science
- Quantum computation and information
- Many more...



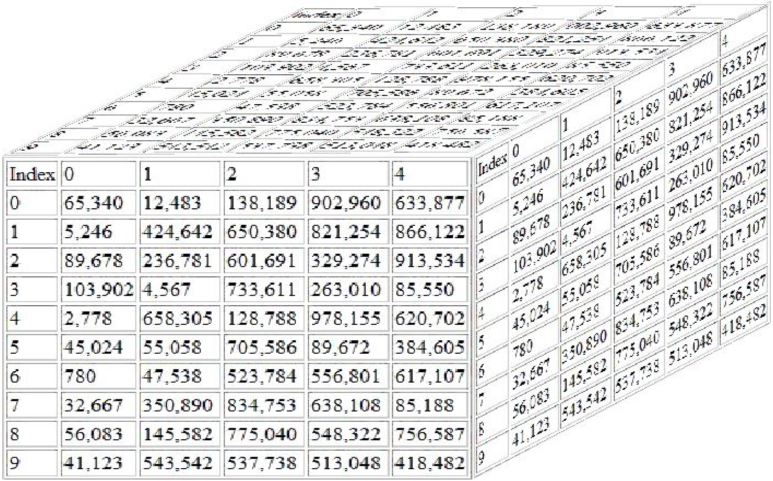
Ways to study: Divide stochastic cube into slices



Cube to slices. In a triply stochastic cube, every slice is doubly stochastic.

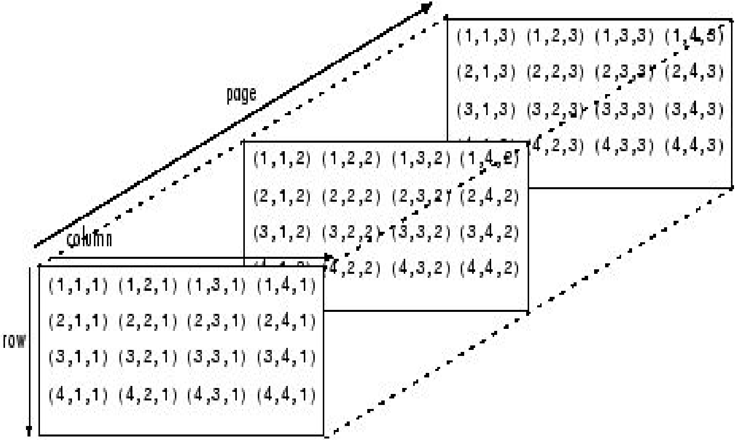
We will divide a (3D) stochastic cube into (2D) stochastic matrices - slices. Using the properties of stochastic matrices, we study the polytope of the stochastic cubes.

Data cube



Index	0	1	2	3	4
0	65,340	12,483	138,189	902,960	633,877
1	5,246	424,642	650,380	821,254	866,122
2	89,678	236,781	601,691	329,274	913,534
3	103,902	4,567	733,611	263,010	85,550
4	2,778	658,305	128,788	978,155	620,702
5	45,024	55,058	705,586	89,672	384,605
6	780	47,538	523,784	556,801	617,107
7	32,667	350,890	834,753	638,108	85,188
8	56,083	145,582	775,040	548,322	756,587
9	41,123	543,542	537,738	513,048	418,482

Tensor indices



Courtesy of mathworks

The $2 \times 2 \times 3$ and $3 \times 3 \times 3$ Tensors

$$A = (a_{ijk})$$

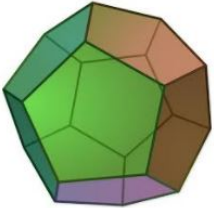
$i, j = 1, 2,\ k = 1, 2, 3$:

$$A = \left[\begin{array}{cc|cc|cc} a_{111} & a_{121} & a_{112} & a_{122} & a_{113} & a_{123} \\ a_{211} & a_{221} & a_{212} & a_{222} & a_{213} & a_{223} \end{array}\right]$$

$i, j, k = 1, 2, 3$:

$$A = \left[\begin{array}{c|c|c} 3 \times 3 & 3 \times 3 & 3 \times 3 \end{array}\right]$$

Combinatorial properties of tensors



Recall doubly stochastic matrices

Let $A = (a_{ij})$ be an $n \times n$ nonnegative matrix: $a_{ij} \geq 0$, $\forall i, j$.

If for every $i = 1, 2, \ldots, n$ (fix a row)

$$\sum_{j=1}^{n} a_{ij} = 1 \qquad \text{(row sum)}$$

and for every $j = 1, 2, \ldots, n$ (fix a column)

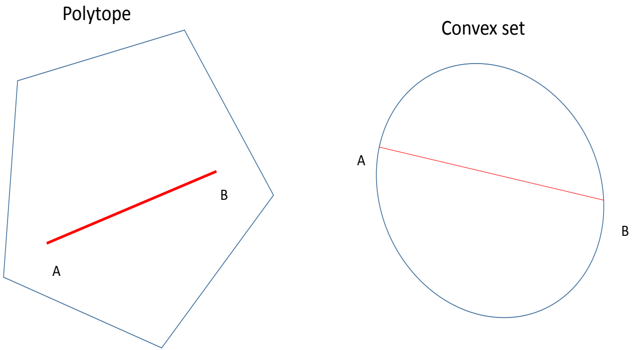
$$\sum_{i=1}^{n} a_{ij} = 1 \qquad \text{(column sum)}$$

then A is called a *doubly stochastic* matrix.

Birkhoff-von Neumann Polytope Theorem

- Birkhoff (1946) - von Neumann (1953): An $n \times n$ matrix is doubly stochastic if and only if it is a convex combination of some $n \times n$ permutation matrices.
- The van der Waerden conjecture (1926-1981): The permanent function defined on set of $n \times n$ doubly stochastic attains its minimum value $\frac{n!}{n^n}$ when all entries are equal to $\frac{1}{n}$.
- The Birkhoff polytope: Consider $n \times n$ matrices as elements in $\mathbb{R}^{n^2}$. The polytope of all $n \times n$ doubly stochastic matrices is generated by the permutation matrices. It has dimension $(n-1)^2$ with $n!$ vertices and n^2 facets.

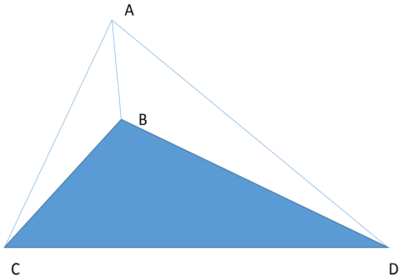
Polytope



$tA + (1-t)B, 0 \le t \le 1$

A polytope is a finitely generated convex set (hull)

Polytope



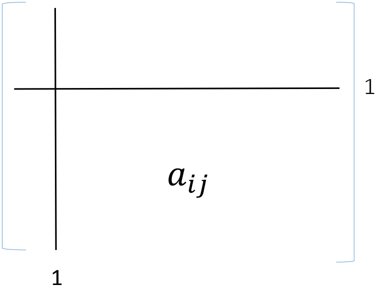
Vertices A, B, C, D, … (extreme points or 0-faces)
Edges AB, AC, AD, BC, BD, CD, … (1-faces)
Faces ABC, ABD, ACD, BCD, … (k-faces)

Polytope of tensors

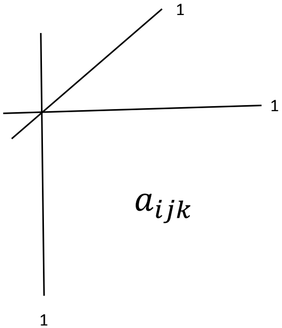
Study the polytopes of higher dimension (mainly $n \times n \times n$) tensors as subsets of $\mathbb{R}^m$ (resp. $m = n^3$)

- Shapes and relations of three polytopes
  1. 0-1 generated polytope Δ_n
  2. convex set of positive Per D_n
  3. and triply stochastic tensors Ω_n
- Number of vertices of triply stochastic tensors
- Line stochastic tensors vs plane stochastic tensors

From doubly stochastic to triply stochastic



Doubly stochastic



Triply stochastic

Higher dimensions

Consider a multidimensional array (hypermatrix, cube, tensor) of numerical values, $n \times n \times n$, say, satisfying:

$$A = (a_{ijk}), \quad a_{ijk} \geq 0$$

$$\sum_{i=1}^{n} a_{ijk} = 1, \quad \forall j, k$$

$$\sum_{j=1}^{n} a_{ijk} = 1, \quad \forall i, k$$

$$\sum_{k=1}^{n} a_{ijk} = 1, \quad \forall i, j$$

More generally, an $n_1 \times n_2 \times \cdots n_m$ tensor of order m

$$A = (a_{i_1 i_2 \cdots i_m}), \quad 1 \leq i_t \leq n_t, \quad t = 1, 2, \ldots, m$$

Warm-up question: Ranks of coefficient matrices?

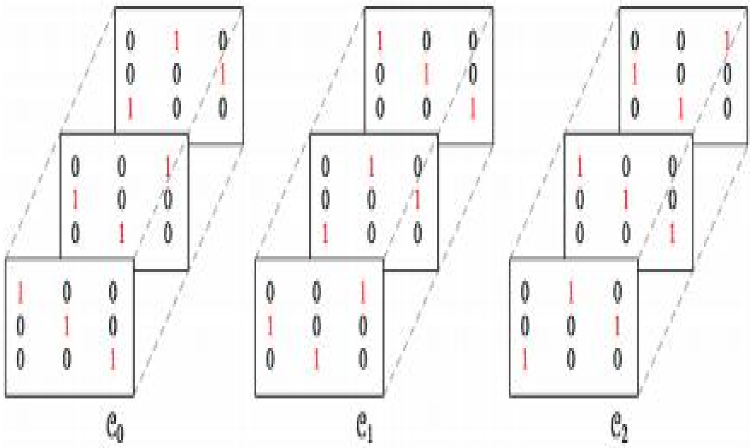
By a *matrix* approach, find the ranks of the coefficient matrices for

$$\sum_{i=1}^{n} x_{ij} = 1,\ j = 1, 2, \dots, n, \quad \sum_{j=1}^{n} x_{ij} = 1,\ i = 1, 2, \dots, n$$

and

$$\sum_{i=1}^{n} y_{ijk} = 1, \quad \sum_{j=1}^{n} y_{ijk} = 1, \quad \sum_{k=1}^{n} y_{ijk} = 1.$$

What is a permutation tensor?



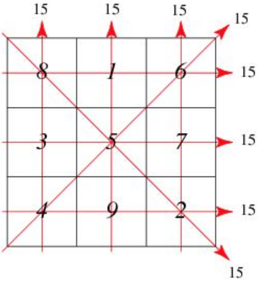
Courtesy of Xie, Jin, and Wei 2016 LAMA

Latin squares and permutation tensors

A	B	C
B	C	A
C	A	B

1	2	3
2	3	1
3	1	2

Magic square and Semi-magic Square



1	5	9
6	7	2
8	3	4

the 12 Latin squares of order three are given by

$$\begin{bmatrix}1&2&3\\2&3&1\\3&1&2\end{bmatrix},\begin{bmatrix}1&2&3\\3&1&2\\2&3&1\end{bmatrix},\begin{bmatrix}1&3&2\\2&1&3\\3&2&1\end{bmatrix},\begin{bmatrix}1&3&2\\3&2&1\\2&1&3\end{bmatrix},$$

$$\begin{bmatrix}2&1&3\\1&3&2\\3&2&1\end{bmatrix},\begin{bmatrix}2&1&3\\3&2&1\\1&3&2\end{bmatrix},\begin{bmatrix}2&3&1\\1&2&3\\3&1&2\end{bmatrix},\begin{bmatrix}2&3&1\\3&1&2\\1&2&3\end{bmatrix},$$

$$\begin{bmatrix}3&2&1\\1&3&2\\2&1&3\end{bmatrix},\begin{bmatrix}3&2&1\\2&1&3\\1&3&2\end{bmatrix},\begin{bmatrix}3&1&2\\1&2&3\\2&3&1\end{bmatrix},\begin{bmatrix}3&1&2\\2&3&1\\1&2&3\end{bmatrix},$$

Latin squares and permutation tensors

Fact:
L_n=# of $n \times n$ Latin square;
P_n= # of $n \times n \times n$ permutation tensors. Then

$$L_n = P_n$$

Proof. If (i, j)-entry of the Latin square is k, then let $p_{ijk} = 1$. $\square$

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix} \mapsto \left[\begin{array}{ccc|c|ccc|c|ccc} 1 & 0 & 0 & & 0 & 1 & 0 & & 0 & 0 & 1 \\ 0 & 0 & 1 & & 1 & 0 & 0 & & 0 & 1 & 0 \\ 0 & 1 & 0 & & 0 & 0 & 1 & & 1 & 0 & 0 \end{array}\right]$$

How many Latin squares?

Fact (van Lint & Wilson, p.161):

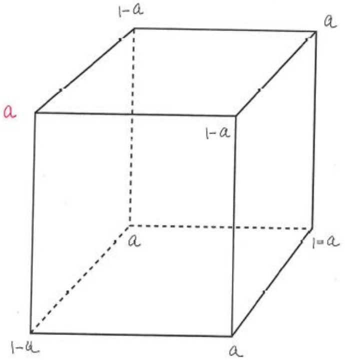
$$\prod_{k=1}^{n} (k!)^{n/k} \geq L_n \geq \frac{(n!)^{2n}}{n^{n^2}}.$$

Shao and Wei (1992):

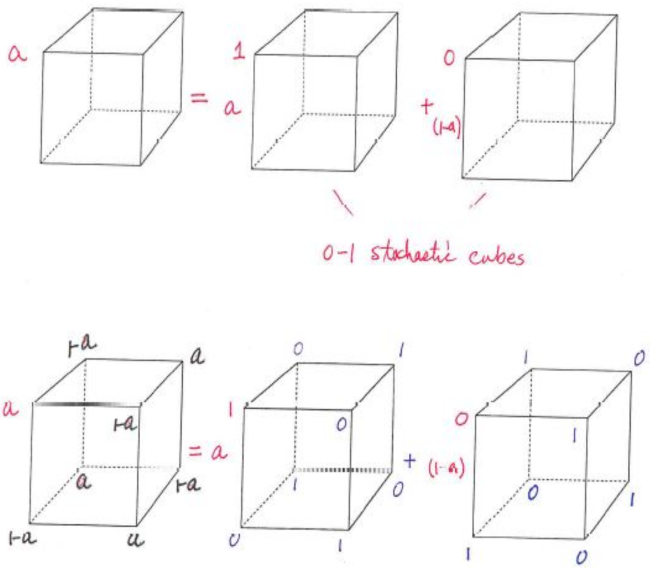
$$L_n = n! \sum_{A \in B_n} (-1)^{\sigma_0(A)} \binom{\operatorname{per} A}{n}$$

where B_n is the set of all 0-1 $n \times n$ matrices, $\sigma_0(A)$ is the number of zero entries in matrix A, and $\operatorname{per} A$ is the permanent of matrix A.

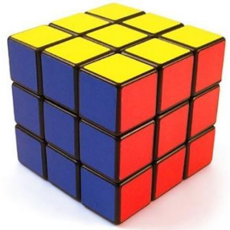
The $2 \times 2 \times 2$ stochastic tensors



The $2 \times 2 \times 2$ stochastic tensors



The $3 \times 3 \times 3$ stochastic tensors

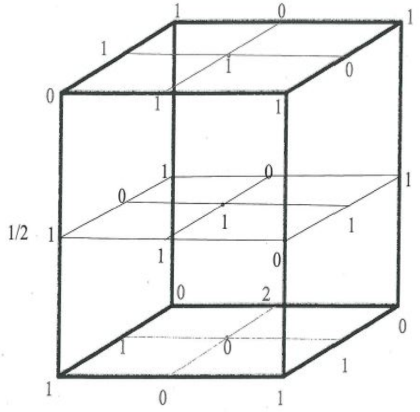


There are 12 0-1 $3 \times 3 \times 3$ permutation tensors.

Question: Can every $3 \times 3 \times 3$ stochastic cube be written as a convex combination of 0-1 $3 \times 3 \times 3$ stochastic cubes

12 0-1 permutation tensors as vertices + 54 non 0-1 vertices.

The $3 \times 3 \times 3$ case: An extreme pt with non 0-1 entries



Not a combination of 0-1 tensors; it's an extreme point

Let F be

$$\left(\begin{array}{ccc:ccc:ccc} 0 & \boxed{0.6} & 0.4 & \boxed{1} & 0 & 0 & 0 & 0.4 & \boxed{0.6} \\ \boxed{0.6} & 0 & 0.4 & 0 & 0.4 & \boxed{0.6} & 0.4 & \boxed{0.6} & 0 \\ 0.4 & 0.4 & \boxed{0.2} & 0 & \boxed{0.6} & 0.4 & \boxed{0.6} & 0 & 0.4 \end{array}\right)$$

If $F = x_1 P_1 + \cdots x_k P_k$, where each P_i is a permutation tensor, then every P_i takes the form below (same 0-1 pattern as F). There is only one such permutation cube. (Start with $\boxed{*}$).

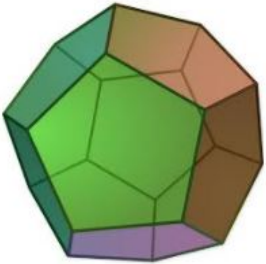
$$P_i = \left(\begin{array}{ccc:ccc:ccc} 0 & * & * & 1 & 0 & 0 & 0 & * & * \\ * & 0 & * & 0 & \boxed{*} & * & * & * & 0 \\ * & * & * & 0 & * & * & * & 0 & * \end{array}\right).$$

Upper bound for the number of vertices

Krein-Milman theorem: every compact convex polytope is the convex hull of its vertices.

The Birkhoff polytope (doubly stochastic matrices) is the convex hull of the $n!$ permutation matrices.

How many vertices (edges, i-faces, facets, etc) does Ω_n have?



Existing upper/lower bounds

Let $f_0(\Omega_n)$ be the number of vertices (0-face) of Ω_n.

Theorem (Ahmed 2003-Chang, Paksoy and Z. 2016, LZZ 2017)

$$\frac{(n!)^{2n}}{n^{n^2}} \leq f_0(\Omega_n) \leq \binom{n^3 - \lfloor \frac{(n-1)^3+1}{2} \rfloor}{3n^2 - 3n + 1} + \binom{n^3 - \lfloor \frac{(n-1)^3+2}{2} \rfloor}{3n^2 - 3n + 1}$$

The polytope Ω_n is an $(n-1)^3$-dimensional affine subspace of $\mathbb{R}^{n^3}$; it has exactly n^3 facets $F_{ijk} = \{x \in \Omega_n \mid x_{ijk} = 0\}, 1 \leq i, j, k \leq n$.

Question 0: Qs about the polytope Ω_n

- Over $\mathbb{R}$ - Convex Analysis, computational geometry
  1. What are *exactly* the vertices of Ω_n?
  2. Give *better* lower/upper bounds for # of vertices of Ω_n.
  3. What are *exactly* the vertices of Ω_n that are not 0-1 tensors?
  4. What are the k-faces (say, dim $=1$, edges) of Ω_n?
- Over $\mathbb{Q}$ - Algebraic Combinatorics
  1. Find the structures of the vertices of Ω_n.
  2. Find the number of vertices of Ω_n.
  3. Are there vertices of Ω_n that are not rational?

Questions 1: How many extreme points for $4 \times 4 \times 4$?

Case	lower	actual	upper
n=2	1	2	21318
n=3	2.37	66	$\frac{1}{27}\binom{65}{26}$
n=4	25.6	$f_0(\Omega_4)^*$	$\frac{1}{64}\binom{138}{63}$

Lower and upper bounds

* Ke, Li, and Xiao, 2016: $f_0(\Omega_4) = 225,216$
* R. Sze, email Dec. 30, 2016: $f_0(\Omega_4) = 37,081,728$

*

Question 2: Search for better bounds

Let L_n denote the number of $n \times n$ Latin squares.

Note that $L_n \geq \frac{(n!)^{2n}}{n^{n^2}}$ (see, e.g., van Lint&Wilson, p.162).

Every Latin square is interpreted as a 0-1 permutation tensor and every $n \times n \times n$ 0-1 permutation tensor is an extreme point of Ω_n:

$$\frac{(n!)^{2n}}{n^{n^2}} \leq L_n \leq f_0(\Omega_n)$$

A big gap between L_n and $f_0(\Omega_n)$! Need better bounds!!

Question 3: $K_n \leq f_0(\Omega_n)$?

Let L_n denote the number of $n \times n$ Latin squares.

It is known that (see also van Lint&Wilson's book).

$$\frac{(n!)^{2n}}{n^{n^2}} \leq L_n \leq \prod_{k=1}^{n} (k!)^{n/k} := K_n.$$

We would like to ask the question if $K_n \leq f_0(\Omega_n)$.

What is the permanent/determinant of a Latin square?

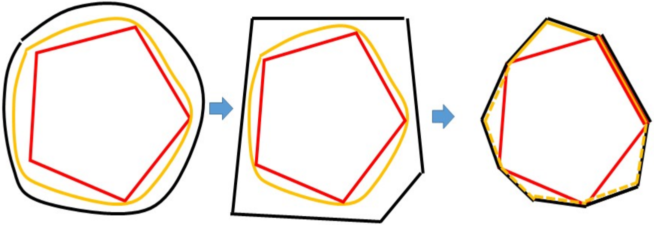
Question 4: What is the boundary of Ω_n?

3 convex sets
2 polytopes

Δ: 0-1

L: per>0

Ω: each line sum = 1



Question 5: When is a tensor stochastic?

$$A = (a_{ijk})_{n\times n\times n}$$

$$A = (a_{i_1 i_2 \cdots i_d})_{n_1\times n_2\times\cdots\times n_d}$$

Line-stochastic: each sum w.r.t. one index $=1$

Plane-stochastic: each sum w.r.t. two indices $=1$

k-hyperplane stochastic: each sum w.r.t. k indices $=1$

Does there exist some sort of stochastic tensor of size $2\times 2\times 2\times 2$? 0-1 tensor of size $2\times 2\times 2\times 2$?

From det, per, GMF of matrices to tensors

$$A = (a_{ij})_{n\times n}$$

$$\det A = \frac{1}{2}\sum_{\alpha,\beta\in S_n} (-1)^{sgn(\alpha)sgn(\beta)} \prod_i a_{\alpha(i)\beta(i)}$$

$$\operatorname{per} A = \frac{1}{2}\sum_{\alpha,\beta\in S_n} \prod_i a_{\alpha(i)\beta(i)}$$

$$d_G^{\chi}A = \frac{1}{2}\sum_{\alpha,\beta\in G} \chi(\alpha)\chi(\beta) \prod_i a_{\alpha(i)\beta(i)}$$

$$\downarrow$$

$$A = (a_{ijk})_{n\times n\times n}$$

Hyperdeterminant

Cayley (1849-):

$$A = \left(a_{i_1 i_2 \cdots i_d}\right)_{n\times n\times\cdots\times n}$$

$$\det A = \frac{1}{n!}\sum_{\pi_1,\ldots,\pi_d\in S_n} sgn(\pi_1)\ldots sgn(\pi_d)\prod_{i=1}^{n} a_{\pi_1(i)\cdots\pi_d(i)}$$

$\det(A) = 0$ if d is odd ......

Gelfand et al (1992) ......

L.-H. Lim (Chapter 15 in Handbook of Lin. Alg., CRC, 2013)

Hyperpermanent or permanent for d-dim arrarys

Cayley (1849-):

$$A = \left(a_{i_1 i_2 \cdots i_d}\right)_{n \times n \times \cdots \times n}$$

$$\operatorname{per} A = \frac{1}{n!} \sum_{\pi_1, \ldots, \pi_d \in S_n} \prod_{i=1}^{n} a_{\pi_1(i) \cdots \pi_d(i)}$$

$$\operatorname{per} A = \sum_{\pi_2, \ldots, \pi_d \in S_n} \prod_{i=1}^{n} a_{i \pi_2(i) \cdots \pi_d(i)}$$

Tensors

$$A = (a_{ijk})_{n \times n \times n}$$

$$A = (a_{ij\cdots k})_{n_1 \times n_2 \times \cdots \times n_d}$$

Question 6: Find bounds of the permanent of a 0-1 tensor

Let A be a 0-1 tensor. Then

$$? \leq \operatorname{per} A \leq ?$$

References

1. M. Ahmed, J. De Loera, and R. Hemmecke, *Polyhedral Cones of Magic Cubes and Squares,* in Disc. Comput. Geo. Algo. Comb., Vol. **25**, pp. 25–41 (eds B. Aronov et al), 2003, Springer.
2. H. Chang, V.E. Paksoy, and F. Zhang, *Polytopes of Stochastic Tensors*, Ann. Funct. Analysis, Vol. 7, No. 3 (2016), 386–393.
3. L.-B. Cui, W. Li, and M.K. Ng, *Birkhoff–von Neumann Theorem for Multistochastic Tensors,* SIMAX. **35** (2014) 956–973.
4. Z. Li, F. Zhang and X.-D. Zhang, On the number of vertices of the stochastic tensor polytope, LAMA online, 2017.
5. R. Ke, W. Li and M. Xiao, Characterization of Extreme Points of Multi-Stochastic Tensors, Comput. Methods Appl. Math. 2016.
6. G.M. Ziegler, *Lectures on Polytopes,* Springer, 1995.