What are high-dimensional permutation? How many are there?

Nati Linial and Zur Luria

IPAM Retreat June '11

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts.

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes.

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

- Random simplicial complexes.

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

- Random simplicial complexes.
- Extremal properties of simplicial complexes.

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

- Random simplicial complexes.
- Extremal properties of simplicial complexes.
- Random cover maps "random lifts of graphs".

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

- Random simplicial complexes.
- Extremal properties of simplicial complexes.
- Random cover maps "random lifts of graphs".
- Likewise for 3-manifolds,

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

- Random simplicial complexes.
- Extremal properties of simplicial complexes.
- Random cover maps "random lifts of graphs".
- Likewise for 3-manifolds, chord diagrams,

The general question

Many basic combinatorial structures have interesting high-dimensional counterparts. Graphs are one-dimensional simplicial complexes. Can we develop parallel theories of higher-dimensional simplicial complexes?

- Random simplicial complexes.
- Extremal properties of simplicial complexes.
- Random cover maps "random lifts of graphs".
- Likewise for 3-manifolds, chord diagrams, knots?

The basic setup

In particular, it is of interest to investigate analogs of

The basic setup

In particular, it is of interest to investigate analogs of

- Trees ("hypertrees")
- Cycles (nontrivial homology)

Today we want to seek high-dimensional counterparts of permutations.

The basic setup

In particular, it is of interest to investigate analogs of

- Trees ("hypertrees")
- Cycles (nontrivial homology)

Today we want to seek high-dimensional counterparts of permutations.
It is a recurring theme that in moving to higher dimensions many simple, even trivial facts that we are very used to, take on a new life and become richer and more geometric.

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:
For an n-vertex graph G with $n - 1$ edges TFAE.

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:
For an n-vertex graph G with $n-1$ edges TFAE. (i.e., G is a tree).

- G is connected.

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:
For an n-vertex graph G with $n - 1$ edges TFAE. (i.e., G is a tree).

- G is connected.
- G is acyclic.

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:
For an n-vertex graph G with $n - 1$ edges TFAE. (i.e., G is a tree).

- G is connected.
- G is acyclic.
- G is collapsible.

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:
For an n-vertex graph G with $n-1$ edges TFAE. (i.e., G is a tree).

- G is connected.
- G is acyclic.
- G is collapsible.

Recall: an elementary collapse is a step in which we remove the unique edge that touches a leaf vertex.

An illustration - the story of trees

We usually teach the following simple fact in undergraduate discrete math courses:
For an n-vertex graph G with $n - 1$ edges TFAE. (i.e., G is a tree).

- G is connected.
- G is acyclic.
- G is collapsible.

Recall: an elementary collapse is a step in which we remove the unique edge that touches a leaf vertex. We say that G is collapsible provided that all its edges can be eliminated through a series of elementary collapses.

A high-dimensional perspective of trees

Q: How should we capture the notion of graph connectivity

A high-dimensional perspective of trees

Q: How should we capture the notion of graph connectivity and acyclicity to facilitate the passage to high dimensions?
A: Using linear algebra. Let M be the $V \times E$ incidence matrix of G.

A high-dimensional perspective of trees

Q: How should we capture the notion of graph connectivity and acyclicity to facilitate the passage to high dimensions?
A: Using linear algebra. Let M be the $V \times E$ incidence matrix of G.

- G is connected iff the left kernel of M consists only of the all 1's vector.

A high-dimensional perspective of trees

Q: How should we capture the notion of graph connectivity and acyclicity to facilitate the passage to high dimensions?
A: Using linear algebra. Let M be the $V \times E$ incidence matrix of G.

- G is connected iff the left kernel of M consists only of the all 1's vector.
- G is acyclic iff M's right kernel is trivial.

A high-dimensional perspective of trees

Q: How should we capture the notion of graph connectivity and acyclicity to facilitate the passage to high dimensions?
A: Using linear algebra. Let M be the $V \times E$ incidence matrix of G.

- G is connected iff the left kernel of M consists only of the all 1's vector.
- G is acyclic iff M's right kernel is trivial.

By counting dimensions the two conditions are equivalent for an n-vertex graph with $n-1$ edges.

A high-dimensional perspective of trees (contd.)

In the d-dimensional case we consider the inclusion matrix M of the $(d-1)$ vs. d-dimensional faces of the simplicial complex X.

A high-dimensional perspective of trees (contd.)

In the d-dimensional case we consider the inclusion matrix M of the $(d-1)$ vs. d-dimensional faces of the simplicial complex X. (Viewed as a linear operator, this matrix is the boundary operator ∂_d.)

A high-dimensional perspective of trees (contd.)

In the d-dimensional case we consider the inclusion matrix M of the $(d-1)$ vs. d-dimensional faces of the simplicial complex X. (Viewed as a linear operator, this matrix is the boundary operator ∂_d.) Well, actually, we should be talking about the signed inclusion matrix to account for orientation, but let's ignore it. Alternatively we can work over $\mathbb{F}_2$ to do away with the signs.

Connectivity means that M's left kernel is trivial

Connectivity means that M's left kernel is trivial i.e., $H_{d-1}(X) = 0$ (the $(d-1)$-st homology of X vanishes).

Connectivity means that M's left kernel is trivial i.e., $H_{d-1}(X) = 0$ (the $(d-1)$-st homology of X vanishes).
Acyclicity means that M's right kernel vanishes,

Connectivity means that M's left kernel is trivial i.e., $H_{d-1}(X) = 0$ (the $(d-1)$-st homology of X vanishes).
Acyclicity means that M's right kernel vanishes, i.e., $H_d(X) = 0$ (the d-th homology of X vanishes).

Connectivity means that M's left kernel is trivial i.e., $H_{d-1}(X) = 0$ (the $(d-1)$-st homology of X vanishes).
Acyclicity means that M's right kernel vanishes, i.e., $H_d(X) = 0$ (the d-th homology of X vanishes).
If the simplicial complex X is d-dimensional, has n vertices, a full $(d-1)$-dimensional skeleton and $\binom{n-1}{d}$ d-faces, then again by couting dimensions, the two conditions are equivalent.

What about collapsibility?

The notion of an elementary collapse is easy to extend to general dimension: Find a $(d-1)$-dimensional face that is covered by

What about collapsibility?

The notion of an elementary collapse is easy to extend to general dimension: Find a $(d-1)$-dimensional face that is covered by exactly one d-dimensional face and remove that d-dimensional face.

There are various notions of collapsibility of the whole complex X. Here we consider a very simple one, namely that it is possible to eliminate all of X's d-dimensional faces through a series of elementary collapses.

What about collapsibility?

The notion of an elementary collapse is easy to extend to general dimension: Find a $(d-1)$-dimensional face that is covered by exactly one d-dimensional face and remove that d-dimensional face.

There are various notions of collapsibility of the whole complex X. Here we consider a very simple one, namely that it is possible to eliminate all of X's d-dimensional faces through a series of elementary collapses. (The order at which we collapse is immaterial).

But wait...

Don't these conditions (triviality of the left kernel, vanishing of the right kernel) depend on the underlying field?

But wait...

Don't these conditions (triviality of the left kernel, vanishing of the right kernel) depend on the underlying field?
In the one-dimensional case there is no such issue, since collapsibility is a purely combinatorial condition, and all three conditions for being a tree are equivalent.

But wait...

Don't these conditions (triviality of the left kernel, vanishing of the right kernel) depend on the underlying field?
In the one-dimensional case there is no such issue, since collapsibility is a purely combinatorial condition, and all three conditions for being a tree are equivalent. Consequently, the triviality of the kernels is independent of the underlying field.

But wait...

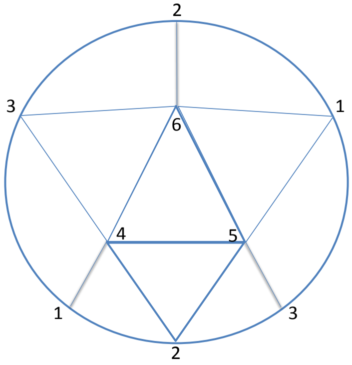
Don't these conditions (triviality of the left kernel, vanishing of the right kernel) depend on the underlying field?

In the one-dimensional case there is no such issue, since collapsibility is a purely combinatorial condition, and all three conditions for being a tree are equivalent. Consequently, the triviality of the kernels is independent of the underlying field.

The implication collapsibility $\to$ vanishing of the homology holds in all dimensions and over every base field.

However, in higher dimension these conditions are no longer equivalent

In particular, the underlying field cannot be ignored.



and now to the real subject - high dimensional permutations

A permutation can be encoded by means of a permutation matrix.

and now to the real subject - high dimensional permutations

A permutation can be encoded by means of a permutation matrix. As we all know, this is an $n \times n$ array of zeros and ones in which every line contains exactly one 1-entry.

and now to the real subject - high dimensional permutations

A permutation can be encoded by means of a permutation matrix. As we all know, this is an $n \times n$ array of zeros and ones in which every line contains exactly one 1-entry.
A line here means either a row or a column.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.
It is an array $[n] \times [n] \times \ldots \times [n] = [n]^{d+1}$ (with $d + 1$ factors) of zeros and ones in which every line contains exactly one 1-entry.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.
It is an array $[n] \times [n] \times \ldots \times [n] = [n]^{d+1}$ (with $d + 1$ factors) of zeros and ones in which every line contains exactly one 1-entry.
Whereas a matrix has two kinds of lines, namely rows and columns, now there are $d + 1$ kinds of lines.

A notion of high dimensional permutations

This suggests the following definition of a d-dimensional permutation on $[n]$.
It is an array $[n] \times [n] \times \ldots \times [n] = [n]^{d+1}$ (with $d + 1$ factors) of zeros and ones in which every line contains exactly one 1-entry.
Whereas a matrix has two kinds of lines, namely rows and columns, now there are $d + 1$ kinds of lines.
A line is a set of n entries in the array that are obtained by fixing d out of the $d + 1$ coordinates and the letting the remaining coordinate take all values from 1 to n.

The case $d = 2$.

The case $d = 2$.

The case $d = 2$. Don't I know you from somewhere?

According to our definition, a 2-dimensional permutation on $[n]$ is an $[n] \times [n] \times [n]$ array of zeros and ones in which every row every column and every shaft contains exactly one 1-entry.
An equivalent description can be achieved by using a topographical map of this terrain.

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows:

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows: The (i, j) entry in this array is k where k is the "height above the ground" of the unique 1-emtry in the shaft $(i, j, *)$.

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows: The (i, j) entry in this array is k where k is the "height above the ground" of the unique 1-emtry in the shaft $(i, j, *)$.
It is easily verified that the defining condition is that in this array every line (row or column) contains every entry $n \geq i \geq 1$ exactly once.

The two-dimensional case

Rather that an $[n] \times [n] \times [n]$ array of zeros and ones we can now consider an $[n] \times [n]$ array with entries from $[n]$, as follows: The (i, j) entry in this array is k where k is the "height above the ground" of the unique 1-emtry in the shaft $(i, j, *)$.
It is easily verified that the defining condition is that in this array every line (row or column) contains every entry $n \geq i \geq 1$ exactly once.
In other words: Two-dimensional permutations are synonymous with Latin Squares.

Where do we go from here

There are so many things we know about ("one-dimensional") permutations. Let's see if we can develop the analogous high-dimensional theory. We know

- How to count them.

Where do we go from here

There are so many things we know about ("one-dimensional") permutations. Let's see if we can develop the analogous high-dimensional theory. We know

- How to count them.
- How to generate them randomly and efficiently.

Where do we go from here

There are so many things we know about ("one-dimensional") permutations. Let's see if we can develop the analogous high-dimensional theory. We know

- How to count them.
- How to generate them randomly and efficiently.
- We know what they look like typically.

Where do we go from here

There are so many things we know about ("one-dimensional") permutations. Let's see if we can develop the analogous high-dimensional theory. We know

- How to count them.
- How to generate them randomly and efficiently.
- We know what they look like typically.
- Birkhoff von-Neumann Thm on doubly stochastic matrices.

Where do we go from here

There are so many things we know about ("one-dimensional") permutations. Let's see if we can develop the analogous high-dimensional theory. We know

- How to count them.
- How to generate them randomly and efficiently.
- We know what they look like typically.
- Birkhoff von-Neumann Thm on doubly stochastic matrices.
- Even trivial properties can turn into interesting questions in higher dimension.

The count - An interesting numerology

As we all know (Stirling's formula)

$$n! = \left((1+o(1))\frac{n}{e}\right)^n$$

The count - An interesting numerology

As we all know (Stirling's formula)

$$n! = \left((1+o(1))\frac{n}{e}\right)^n$$

As we will discuss below the count of order-n Latin squares is

$$|\mathcal{L}_{\mathrm{n}}| = \left((1+\mathrm{o}(1))\frac{\mathrm{n}}{\mathrm{e}^2}\right)^{\mathrm{n}^2}$$

So, let us conecture

Conjecture

The number of d-dimensional permutations on [n] is

$$|S_n^d| = \left((1+o(1))\frac{n}{e^{d}}\right)^{n^{d}}$$

and what we actually know

At present we can only prove the upper bound

Theorem

The number of d-dimensional permutations on $[n]$ is

$$|S_n^d| \leq \left((1+o(1))\frac{n}{e^d}\right)^{n^d}$$

How do you prove the estimate for the number of Latin Squares?

Recall that the permanent of a square matrix is a "determinant without signs".

$$per(A) = \sum_{\sigma \in S_n} \prod a_{i,\sigma(i)}$$

This is a curious and fascinating mathematical object. E.g.

This is a curious and fascinating mathematical object. E.g.

- It counts perfect matchings in bipartite graphs.

This is a curious and fascinating mathematical object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.

This is a curious and fascinating mathematical object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.
- It is $\#$-P-hard to calculate the permanent exactly, even for a $0/1$ matrix.

This is a curious and fascinating mathematical object. E.g.

- It counts perfect matchings in bipartite graphs.
- In other words, it counts the generalized diagonals included in a $0/1$ matrix.
- It is $\#$-P-hard to calculate the permanent exactly, even for a $0/1$ matrix.
- On the other hand there is an efficient approximation scheme for permanents of nonnegative matrices.

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, it makes a lot of sense to seek bounds on it. We say that A is a doubly stochastic matrix provided that

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, it makes a lot of sense to seek bounds on it. We say that A is a doubly stochastic matrix provided that

- Its entries are nonnegative.

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, it makes a lot of sense to seek bounds on it. We say that A is a doubly stochastic matrix provided that

- Its entries are nonnegative.
- The sum of entries in every row is 1.

A lower bound on the permanent

Since the permanent is so mysterious and hard to compute, it makes a lot of sense to seek bounds on it. We say that A is a doubly stochastic matrix provided that

- Its entries are nonnegative.
- The sum of entries in every row is 1.
- The sum of entries in every column is 1.

By the marriage theorem, a doubly stochastic matrix has a positive permanent.

By the marriage theorem, a doubly stochastic matrix has a positive permanent. The set of doubly stochastic matrices is a convex polytope. The permanent is a continuous function, so: What is min $per\ A$ over $n \times n$ doubly-stochastic matrices?

By the marriage theorem, a doubly stochastic matrix has a positive permanent. The set of doubly stochastic matrices is a convex polytope. The permanent is a continuous function, so: What is min $per\ A$ over $n \times n$ doubly-stochastic matrices? As conjectured by van der Waerden in the 20's and proved over 50 years later by Falikman and by Egorichev, in the minimizing matrix all entries are $\frac{1}{n}$.

Theorem

The permanent of every $n \times n$ doubly stochastic matrix is $\geq \frac{n!}{n^n}$.

An upper bound on permanents

The following was conjectured by Minc and proved by Brégman

Theorem

Let A be an $n \times n$ $0/1$ matrix with r_i ones in the i-th row $i = 1, \ldots, n$. Then $perA \leq \prod_i (r_i!)^{1/r_i}$. The bound is tight.

Our work

The main part of our work is an extension of the Minc-Brégman theorem. In fact our work uses ideas from subsequent papers of Schrijver and Radhakrishnan.

Our work

The main part of our work is an extension of the Minc-Brégman theorem. In fact our work uses ideas from subsequent papers of Schrijver and Radhakrishnan.

This is why our result is an upper bound on the number of d-dimensional permutations.

Our work

The main part of our work is an extension of the Minc-Brégman theorem. In fact our work uses ideas from subsequent papers of Schrijver and Radhakrishnan.
This is why our result is an upper bound on the number of d-dimensional permutations.
What about a matching lower bound?

Our work

The main part of our work is an extension of the Minc-Brégman theorem. In fact our work uses ideas from subsequent papers of Schrijver and Radhakrishnan.
This is why our result is an upper bound on the number of d-dimensional permutations.
What about a matching lower bound?
We don't have it (yet....), but there is a reason.

The (obvious) analog of the van der Waerden conjecture fails in higher dimension

It is an easy consequence of the marriage theorem that if in a $0/1$ matrix A, all row sums and all column sums equal $k \geq 1$, then $perA > 0$.

The (obvious) analog of the van der Waerden conjecture fails in higher dimension

It is an easy consequence of the marriage theorem that if in a $0/1$ matrix A, all row sums and all column sums equal $k \geq 1$, then $perA > 0$.
The analogous statement is no longer true in higher dimensions.

The (obvious) analog of the van der Waerden conjecture fails in higher dimension

It is an easy consequence of the marriage theorem that if in a $0/1$ matrix A, all row sums and all column sums equal $k \geq 1$, then $perA > 0$.
The analogous statement is no longer true in higher dimensions.
Here is an example of a $4 \times 4 \times 4$ array with two zeros and two ones in every line which contains no 2-permutation.

An example

$$\begin{matrix} * & 0 & * & 0 \\ 0 & 1 & 0 & 1 \\ * & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{matrix} \qquad \begin{matrix} 0 & * & 0 & * \\ 0 & 0 & 1 & 1 \\ * & * & 0 & 0 \\ 1 & 0 & 1 & 0 \end{matrix}$$

$$\begin{matrix} 0 & 0 & * & * \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{matrix} \qquad \begin{matrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{matrix}$$

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dmensional permutation) A layer by layer.

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dmensional permutation) A layer by layer. Namely, A is an $n \times n \times n$ array of $0/1$ where every line has a single 1 entry.

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dmensional permutation) A layer by layer.
Namely, A is an $n \times n \times n$ array of $0/1$ where every line has a single 1 entry.
Note that every layer in A is a permutation matrix.

Approximately counting Latin squares

The general scheme: We consider a Latin square (= a 2-dmensional permutation) A layer by layer.
Namely, A is an $n \times n \times n$ array of $0/1$ where every line has a single 1 entry.
Note that every layer in A is a permutation matrix.
Given several layers in A, how many permutation matrices can play the role of the next layer?

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B.

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B. Therefore, there are exactly $perB$ possibilities for the next layer.

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B. Therefore, there are exactly $perB$ possibilities for the next layer.
The estimate for the number of Latin squares is attained by bounding at each step the number of possiblities for the next layer

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B. Therefore, there are exactly $perB$ possibilities for the next layer.
The estimate for the number of Latin squares is attained by bounding at each step the number of possiblities for the next layer
from below, using the van der Waerden's conjecture)

How many choices for the next layer?

Let B be a $0/1$ matrix where $b_{ij} = 1$ iff in all previous layers the ij entry is zero.
The set of all possible next layers coincides with the collection of generalized diagonals in B. Therefore, there are exactly *perB* possibilities for the next layer.
The estimate for the number of Latin squares is attained by bounding at each step the number of possiblities for the next layer
from below, using the van der Waerden's conjecture) and from above, using Minc-Brégman.

Back to basics - Reproving Brégman's theorem

One of the insights gained about Brégman's theorem is that it is useful to interpret it using the notion of entropy.

Back to basics - Reproving Brégman's theorem

One of the insights gained about Brégman's theorem is that it is useful to interpret it using the notion of entropy.
So let us review the basics of this method.

A quick recap of entropy

If X is a discrete random variable, taking the i-th value in its domain with probability p_i then its entropy is

$$H(X) = -\sum p_i \log p_i$$

A quick recap of entropy

If X is a discrete random variable, taking the i-th value in its domain with probability p_i then its entropy is

$$H(X) = -\sum p_i \log p_i$$

In particular if the range of X has cardinality N, then $H(X) \leq \log N$ with equality iff X is distributed uniformly.

A quick recap of entropy

If X is a discrete random variable, taking the i-th value in its domain with probability p_i then its entropy is

$$H(X) = -\sum p_i \log p_i$$

In particular if the range of X has cardinality N, then $H(X) \leq \log N$ with equality iff X is distributed uniformly.
All logarithms here are to base e. This is not the convention when it comes to entropy, but it will make things more convenient for us.

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y) H(X|Y = y).$$

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y) H(X|Y = y).$$

The chain rule is one of the fundamental properties of entropy:

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y)H(X|Y = y).$$

The chain rule is one of the fundamental properties of entropy: If $X_1, \ldots, X_n$ are discrete random variables defined on the same probability space, then

A quick recap of entropy (contd.)

If X and Y are two discrete random variables, then the conditional entropy

$$H(X|Y) := \sum_y Pr(Y = y) H(X|Y = y).$$

The chain rule is one of the fundamental properties of entropy: If $X_1, \ldots, X_n$ are discrete random variables defined on the same probability space, then

$$H(X_1, \ldots, X_n) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \ldots$$

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.
The number of generalized diagonals contained in A is exactly $perA$.

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i. The number of generalized diagonals contained in A is exactly $perA$. Let X be a random variable that takes values which are these generalized diagonals, with uniform distribution. Clearly,

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i. The number of generalized diagonals contained in A is exactly $perA$. Let X be a random variable that takes values which are these generalized diagonals, with uniform distribution. Clearly,

$$H(X) = \log(perA).$$

Proving Brégman's theorem using entropy

Let us fix an $n \times n$ $0/1$ matrix A in which there are exactly r_i 1-entries in the i-th row for every i.
The number of generalized diagonals contained in A is exactly $perA$. Let X be a random variable that takes values which are these generalized diagonals, with uniform distribution. Clearly,

$$H(X) = \log(perA).$$

Therefore, an upper bound on $H(X)$ yields an upper bound on $perA$, which is what we want.

We next express $X = (X_1, \ldots, X_n)$, where X_i is the index of the single 1-entry that is selected by the generalized diagonal X at the i-th row.

We next express $X = (X_1, \ldots, X_n)$, where X_i is the index of the single 1-entry that is selected by the generalized diagonal X at the i-th row. How should we interpret the relation

$$H(X) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \ldots$$

We next express $X = (X_1, \ldots, X_n)$, where X_i is the index of the single 1-entry that is selected by the generalized diagonal X at the i-th row. How should we interpret the relation

$$H(X) = H(X_1) + H(X_2|X_1) + H(X_3|X_1, X_2) + \ldots$$

In particular, what can we say about $H(X_i|X_1, \ldots, X_{i-1})$?

Let us fix a specific value for X, i.e., some generalized diagonal in A.

Let us fix a specific value for X, i.e., some generalized diagonal in A. There are r_i 1's in the i-th row.

Let us fix a specific value for X, i.e., some generalized diagonal in A. There are r_i 1's in the i-th row. Since X takes on a generalized diagonal, we know that there are r_i indices j for which the index X_j coincides with one of these r_i positions.

Let us fix a specific value for X, i.e., some generalized diagonal in A. There are r_i 1's in the i-th row. Since X takes on a generalized diagonal, we know that there are r_i indices j for which the index X_j coincides with one of these r_i positions. If such a j is smaller than i, we say that j is shading the i-th row.

Clearly X_i can take on only unshaded values.

Clearly X_i can take on only unshaded values.
If N_i is the (random) number of 1-entries in the i-th row that remain unshaded when the i-th row is reached, then clearly

$$H(X_i | X_1, \ldots, X_{i-1}) \leq \log N_i$$

Clearly X_i can take on only unshaded values.
If N_i is the (random) number of 1-entries in the i-th row that remain unshaded when the i-th row is reached, then clearly

$$H(X_i|X_1,\ldots,X_{i-1}) \leq \log N_i$$

since the conditioned random variable $(X_i|X_1,\ldots,X_{i-1})$ can take at most N_i values.

Clearly X_i can take on only unshaded values.
If N_i is the (random) number of 1-entries in the i-th row that remain unshaded when the i-th row is reached, then clearly

$$H(X_i|X_1, \ldots, X_{i-1}) \leq \log N_i$$

since the conditioned random variable $(X_i|X_1, \ldots, X_{i-1})$ can take at most N_i values.
Very nice. The troube is that we know very little about the random variable N_i.

A good trick

The way around this difficulty is not to sum the terms $H(X_i | X_1, \ldots, X_{i-1})$ in the normal order, but rather introduce σ, a random ordering of the rows.

A good trick

The way around this difficulty is not to sum the terms $H(X_i|X_1, \ldots, X_{i-1})$ in the normal order, but rather introduce σ, a random ordering of the rows. We add the corresponding terms in the order σ and finally we average over the random choice of σ.

A good trick

The way around this difficulty is not to sum the terms $H(X_i|X_1, \ldots, X_{i-1})$ in the normal order, but rather introduce σ, a random ordering of the rows. We add the corresponding terms in the order σ and finally we average over the random choice of σ.

What can we say about the expectation of $\log N_i^\sigma$?

This can be restated as follows: You are expecting r visitors who arrive at random, independently chosen times.

This can be restated as follows: You are expecting r visitors who arrive at random, independently chosen times.
One of the visitors is your guest of honor and you are interested in his (random) arrival rank among the r visitors.

Clearly his rank N is uniformly distributed over $1, \ldots, r$.

Clearly his rank N is uniformly distributed over $1, \dots, r$. In particular, the expectation

$$\mathbb{E}(\log N) = \frac{1}{r} \sum_{j=1,\dots,r} \log j = \frac{\log r!}{r} = \log(r!)^{1/r}.$$

Clearly his rank N is uniformly distributed over $1, \ldots, r$. In particular, the expectation

$$\mathbb{E}(\log N) = \frac{1}{r} \sum_{j=1,\ldots,r} \log j = \frac{\log r!}{r} = \log(r!)^{1/r}.$$

By summing over all rows, the Brégman bound is established

$$H(X) = \log(perA) \leq \sum_i \log(r_i!)^{1/r_i}.$$

Doing the d-dimensional case

If A is an $n^{d+1} = n \times n \times \ldots \times n$ array of $0/1$ we define $per_d(A)$ to be the number of d-permutations that are included in A. Let's consider all lines in A in the same direction, say lines of the form $l_{\mathbf{i}} = (i_1, \ldots, i_d, *)$. Let $r_{\mathbf{i}}$ be the number of 1's in the line $l_{\mathbf{i}}$.

Theorem

Let A be an $[n]^{d+1}$ *array of* $0/1$, *and let* $r_{\mathbf{i}}$ *be the number of* 1*'s in the line* $l_{\mathbf{i}}$, *as above. Then*

$$per_d(A) \leq \prod_{\mathbf{i}} \exp(f(d, r_{\mathbf{i}})).$$

The function $f(d, r)$ *is defined recursively, via* $f(0, r) = \log r$, *and*

$$f(d, r) = \frac{1}{r} \sum_{k=1,\ldots,r} f(d-1, k).$$

A few words about the proof

The general strategy remains the same, using entropy. The random variable X takes uniformly d-permutations that are contained in A.

A few words about the proof

The general strategy remains the same, using entropy. The random variable X takes uniformly d-permutations that are contained in A. The main new ingredient in the proof is the choice of the random ordering σ.

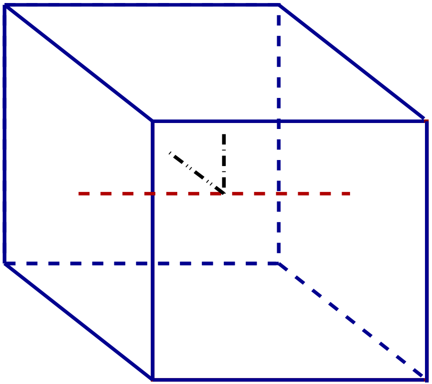
A few words about the proof

The general strategy remains the same, using entropy. The random variable X takes uniformly d-permutations that are contained in A. The main new ingredient in the proof is the choice of the random ordering σ. The expression

$$f(d,r) = \frac{1}{r} \sum_{k=1,\dots,r} f(d-1,k).$$

suggests that the story with the randomly-arriving guests is modified as follows:

Different shades, randomly-arriving visitors



Again we have r visitors one of whom is our guest of honor.

Again we have r visitors one of whom is our guest of honor.

The visitors are arriving at a random order. Only the guest of honor and everyone who came after him remain in the game and are invited again.

Again we have r visitors one of whom is our guest of honor.
The visitors are arriving at a random order. Only the guest of honor and everyone who came after him remain in the game and are invited again.
This procedure is repeated d times and we ask about N, the (random) number of guests who never showed up (in all d repeats) before the guest of honor.

What ordering σ?

This is accomplished by ordering the lines l_i as follows:

What ordering σ?

This is accomplished by ordering the lines l_i as follows: First choose a random ordering of the n layers,

What ordering σ?

This is accomplished by ordering the lines $l_{\mathbf{i}}$ as follows: First choose a random ordering of the n layers, then proceed recursively in each layer.

What ordering σ?

This is accomplished by ordering the lines $l_{\mathbf{i}}$ as follows: First choose a random ordering of the n layers, then proceed recursively in each layer.
In the shade terminology, we first eliminate the ones in row $\mathbf{i}$ that are shaded in direction $d+1$.

What ordering σ?

This is accomplished by ordering the lines $l_{\mathbf{i}}$ as follows: First choose a random ordering of the n layers, then proceed recursively in each layer.

In the shade terminology, we first eliminate the ones in row $\mathbf{i}$ that are shaded in direction $d+1$.

(This is the first round of invitation).

What ordering σ?

This is accomplished by ordering the lines $l_{\mathbf{i}}$ as follows: First choose a random ordering of the n layers, then proceed recursively in each layer.
In the shade terminology, we first eliminate the ones in row $\mathbf{i}$ that are shaded in direction $d + 1$.
(This is the first round of invitation). Shades that come from other directions correspond to our subsequent rounds of inviting the guests.

Let's try to explain the expression $\left(\frac{n}{e^d}\right)^{n^d}$.

Let's try to explain the expression $(\frac{n}{e^d})^{n^d}$.
Of course the arithmetic mean of the numbers from 1 to r is $(1 + o(1))\frac{r}{2}$.

Let's try to explain the expression $(\frac{n}{e^d})^{n^d}$.
Of course the arithmetic mean of the numbers from 1 to r is $(1+o(1))\frac{r}{2}$.
Stirling's formula says that the geometric mean of these numbers is $(1+o(1))\frac{r}{e}$.

Let's try to explain the expression $(\frac{n}{e^d})^{n^d}$.
Of course the arithmetic mean of the numbers from 1 to r is $(1 + o(1))\frac{r}{2}$.
Stirling's formula says that the geometric mean of these numbers is $(1 + o(1))\frac{r}{e}$.
In other words, the arithmetic mean of the numbers $\log j$ over $j = 1, \ldots, r$ is $\log r - 1$.
When we repeat the above process of inviting guests and re-inviting only the late arrivals (what a strange idea???) we lose this 1 term d times.

Is this the ultimate notion of a d-dimensional permutation?

While we find this definition very appealing and the open questions are fascinating, this is by no means the final word on the subject.

Is this the ultimate notion of a d-dimensional permutation?

While we find this definition very appealing and the open questions are fascinating, this is by no means the final word on the subject.
In ongoing work with A. Morgenstern we consider high-dimensional analogs of tournaments (a notion first considered by I. Leader and his students).

Is this the ultimate notion of a d-dimensional permutation?

While we find this definition very appealing and the open questions are fascinating, this is by no means the final word on the subject.
In ongoing work with A. Morgenstern we consider high-dimensional analogs of tournaments (a notion first considered by I. Leader and his students).
An acyclic high-dimensional tournament seems like another good analog of a permutation.

Is this the ultimate notion of a d-dimensional permutation?

While we find this definition very appealing and the open questions are fascinating, this is by no means the final word on the subject.
In ongoing work with A. Morgenstern we consider high-dimensional analogs of tournaments (a notion first considered by I. Leader and his students).
An acyclic high-dimensional tournament seems like another good analog of a permutation.
There is still too little to report from that front.

and there are many more open questions than answers...

- Can we prove a matching lower bound?

and there are many more open questions than answers...

- Can we prove a matching lower bound?
- We know a few things about the polytope of multi-stochastic arrays, but we have only started scratching the ground there.

and there are many more open questions than answers...

- Can we prove a matching lower bound?
- We know a few things about the polytope of multi-stochastic arrays, but we have only started scratching the ground there.
- Similar questions for related concepts: tournaments, STS, 1-factorizations ...

and there are many more open questions than answers...

- Can we prove a matching lower bound?
- We know a few things about the polytope of multi-stochastic arrays, but we have only started scratching the ground there.
- Similar questions for related concepts: tournaments, STS, 1-factorizations ...
- Efficient random generation of such objects.

and there are many more open questions than answers...

- Can we prove a matching lower bound?
- We know a few things about the polytope of multi-stochastic arrays, but we have only started scratching the ground there.
- Similar questions for related concepts: tournaments, STS, 1-factorizations ...
- Efficient random generation of such objects.
- Investigating the typical and extremal properties of d-dimensional permutations.

That's all folks