RANDOMIZED SET SYSTEMS CONSTRAINED BY THE DISCRETE TOMOGRAPHY

Irina Arsenyan, Levon Aslanyan and Hasmik Sahakyan

***Abstract**: This article, in general, is devoted to a set of discrete optimization issues derived from the domain of pattern recognition, machine learning and data mining - specifically. The global objectives are the compactness hypotheses of pattern recognition, and the structural reconstruction of the discrete tomography.*

The driving force of the current research was the proof technic of the discrete isoperimetry problem. In proofs by induction the split technique was applied and then it is important to have some information about the sizes of the split compounds. Isoperimetry itself is a formalism of the compactness hypotheses. From one side knowledge on split sizes helps to find the compact structures and learning sets based on this, from the other side – split sizes help to prove the necessary relations. The pure combinatorial approaches [22-77] are not able at the moment to give an efficient description of the split sizes and – the weighted row-different matrices. The probabilistic method, as it is well-known, gives additional knowledge about the random subsets, and this may be useful as a complementary knowledge about a different objects or a situations concerned the properties of discrete structures – isoperimetry and tomography.

The discrete mathematical science deals with different types of discrete structures, studying their transformations and properties. In some problems we face the issues about the existence of structures under some special constraints, about the enumeration of structures under these constraints, and – on algorithmic optimization. Given a simple structure – in some cases, it can be even hard to compute some basic properties of it. Such are for example the graph chromatic number, the minimal set cover, the solution of the well-known SAT and plenty of other NP-complete problems. When structures are given, the mentioned parameters may be easily computable. To find a structure by the given parameters often becomes hard. We call such problems – inverse problems. Our special interest is in considering of simple (0,1) matrices and their row and column weights. Given a matrix we can compute the mentioned weights (direct problem). The inverse problem, -- when it is to find the construction with the given weights is not simple. At least there is not known polynomial algorithms for this problem. Moreover, the problem is known as the hypotheses posted by famous graph theorist C. Berge so that the problem is well known and unsolved.

Besides the logical and combinatorial analysis of the inverse type problems of discrete optimization, in several cases the probabilistic models were applied successfully. The idea of this paper is to use the probabilistic theory of combinatorial analysis to the discrete tomography problem given in terms of the (0,1) matrices. The paper tries to outline the models, relations and the methodology. Our research priority interest is to understand the opportunities, similarities and perspectives in this broad research area. The study is ongoing and the follow up publication will come soon.

***Keywords**: discrete tomography, discrete isoperimetry, probabilistic theory of combinatorial analysis, (0,1) matrices, discrete optimization.*

***ACM Classification Keywords**: G.2.1 Combinatorics*

Content of the paper

Introduction 205
A. The problem 208
B. The model 209
C. The technique 210
D. Example 1. Positive probability implies the existence. 211
E. Example 2. Probability distributions with the Chebyshev inequality give the asymptotics. 211
F. Example 3. Continuity theorem helps to describe the types of the Boolean functions. 212
Random Discrete Tomography 212
G. Series of compound trials. 212
H. The column weight probability and the average column weight. 213
I. The column weight variance. 214
J. Analysis of compound trials. 215
K. The row differences model. 218
L. Sperner families. 226
Conclusion 226
Bibliography 226
Authors' Information 231

Introduction

Our aim is to present the work done on composition and analysis of appropriate probabilistic models that support investigations of a group of combinatorial problems related to the basic one – the Discrete Tomography. The first articles in the domain of probabilistic theory of combinatorial analysis [1-4] by Erdős, Rényi and Gilbert, and [5-7] by Glagolev, Kospanov and Nechiporuk considered graph models with random vertices and edges, and random Boolean schemes and formulas -- evaluating the frequencies of appearance of different configurations given the specific properties. Since then, thousands of publications appeared in this domain. Today also the systematic presentation of the subject is accessible [8-10].

Our consideration is focused on study of different specifically constrained set systems, which are known alternatively as the problems of hypergraph theory, or the Boolean functions or the n-cube geometrical studies. As usual, we denote the set of all vertices of the n-dimensional unite cube by E^{n}. Subsets of E^{n} compose set systems with the base n-set, that represents the set of Boolean variables. Alternatively, we will also consider multi-sets, defined over the set of vertices of E^{n}.

Within the set theory, formally, a multi-set (or bag) is a 2-tuple (A, r) where A is some set of elements and $\mathrm{r}: \mathrm{A} \rightarrow \mathrm{Z}_{+}$ is a function from A to the set $\mathrm{Z}_{+} = \{1,2, \dots\}$ of positive natural numbers. For each $\mathrm{a} \in \mathrm{A}$ the multiplicity of a in (A, r) is the number $\mathrm{r}\ (\mathrm{a})$, that is the number of its occurrences. For a finite A multi-set (A, r) can be given by a list. If an underlying set U, wherefrom the elements of A are obtained is specified, then the definition can be simplified to just a multiplicity function $\mathrm{r}: \mathrm{U} \rightarrow \mathrm{Z}_{\geq 0}$ obtained by extending A to U with the use of the values 0 for all elements not in A.

The multi-sets $M = (A, r)$ with the universal set E^n and a multiplicity function $\mathrm{r}: E^n \rightarrow Z_{\geq 0}$ will be considered. Different interpretations concerning the structure (A, r) will be considered and used throughout this paper. In particular, A may denote the pairs of rows in a (0,1) matrix. The different rows compose a set of vertices of E^n. The number of repeated rows (or the number of pairs of equal rows) will denote multiplicity.

The probabilistic method used throughout this paper basically deals with the discrete probability spaces. We suppose that it is given a finite space A of elementary events (outcomes of a certain processes or experiments/trials) and that there are probabilities related to these events. If the total number of outcomes of experiment is n, then the probabilities $p_1, p_2, \dots, p_n$ are allocated to them so that $\sum p_i = 1$. The set theoretical relations such as union and intersection, and the inclusion exclusion principle regulate more complex events $X \subseteq A$ under this scheme. The probability space provides probability measures of these events, $p\ (X) = \sum_{a \epsilon X} p\ (a)$. Let us remind some of the basic probabilistic relations. Let A be the finite set of elementary events, X_1, and $X_2 \subseteq A$. Then $p\ (X_1 \cup X_2) = p\ (X_1) + p\ (X_2) - p\ (X_1 \cap X_2)$. Form the expression $p\ (X_1 \cup X_2) \leq p\ (X_1) + p\ (X_2)$ that is known as the Boole's inequality. If here we have the equality, $p\ (X_1 \cap X_2) = 0$, then we call X_1 and X_2 ***mutually exclusive***. X_1 and X_2 are ***independent***, iff $p\ (X_1 \cap X_2) = p\ (X_1) p\ (X_2)$. Further:

for arbitrary events X_i

$$p\ (X_1 \cup X_2 \cup \ldots \cup X_m) \leq p\ (X_1) + p\ (X_2) + \cdots + p\ (X_m), \qquad (2.1)$$

for independent events X_i

$$p\ (X_1 \cap X_2 \cap \ldots \cap X_m) = p\ (X_1) p\ (X_2) \ldots p\ (X_m) \qquad (2.2)$$

and

$$p\ (X_1 \cup X_2 \cup \ldots \cup X_m) = 1 - (1 - p\ (X_1))(1 - (1 - p\ (X_2)) \ldots (1 - (1 - p\ (X_m)). \qquad (2.3)$$

Repeated independent trials when there are only two outcomes in each trial and the probabilities of outcomes remind unchanged is known as the Bernoulli model.

In a succession of n Bernoulli trials number of successes can be $0, 1, \ldots, n$. If p and $q = 1 - p$ are the probabilities of success and failure correspondingly, then the probability that the n Bernoulli trials result in t success, exactly, equals $b(t, n, p) = C_n^t p^t q^{n-t}$. By this we obtained the probability of the random number t of successes in n Bernoulli trials that is known as the binomial probabilistic model. The maximum of $b(t, n, p)$ by t we call the central term. It is easy to check that when t goes from 0 to m the probability $b(t, n, p)$ first increase monotonically to the central term and then it decrease. The central term reaches at $t = (n + 1)p$ and $t - 1$ when these are integer, and at $[(n + 1)p]$ otherwise.

In numerical applications of Bernoulli trials, as a rule, n is large and p is small with the product $\lambda = np$ of a moderate magnitude. Consider the formula

$$C_n^t p^t q^{n-t} = \frac{n(n-1) \ldots (n-t+1)}{t!} p^t (1-p)^{n-t}.$$

Let $t^2 = o(n)$, then $n(n-1)\dots(n-t+1) \sim n^t$ and $(1-p)^{n-t} \sim (1-p)^n \sim e^{-np}$, and

$$b(t,n,p) = C_n^t p^t q^{n-t} \sim \frac{\lambda^t}{t!} e^{-\lambda}.$$

The right part of the last line determines the well-known Poisson probability distribution. The distribution, as demonstrated, is a convenient approximation to the binomial distribution.

Let ξ be the random integer variable. The expected value $E\ (\xi)$ or mean and the variance $Var\ (\xi)$ can be computed as

$E\ [\xi] = \sum_t \mathrm{t} \cdot \Pr\ (\xi = t)$ and $Var\ [\xi] = \sum_t (\mathrm{t} - \mathrm{E}\ [\xi])^2 \cdot \Pr\ (\xi = t)$.

We differentiate several types of convergences of random variables. Let ξ_l be a sequence of random variables, and let their distribution functions be $F_l(x)$, respectively.

The first notion of convergence of a sequence of random variables is known as convergence in probability. The sequence ξ_l converges to a random variable ξ in probability, denoted $\xi_l \xrightarrow{p} \xi$ if for any $> 0 \lim_{l\to\infty} \Pr\ (|\xi_l - \xi| < \varepsilon) = 1$.

Note that this does not say that the difference between ξ_l and ξ becomes very small. What converges here is the probability that the difference between ξ_l and ξ becomes very small. It is, therefore, possible, although unlikely, for ξ_l and ξ to differ by a significant mount and for such differences to occur infinitely often.

A stronger kind of convergence, which does not allow such behavior, is called almost sure convergence or strong convergence. A sequence of random variables ξ_l converges to a random variable ξ almost surely, denoted $\xi_l \xrightarrow{s} \xi$ if for any $\varepsilon > 0 \lim_{L\to\infty} \Pr\ (\sup_{l\geq L} |\xi_l - \xi| < \varepsilon) = 1$.

Finally we like to remind the classical continuity theorem of distributions.

Suppose that for any fixed n the sequence

$$a_{0,n}, a_{1,n}, a_{2,n}, \dots$$

is a probability distribution of n-th trial, that is, $a_{k,n} \geq 0$ and $\sum_{k=0}^{\infty} a_{k,n} = 1$.

In order that a limit $a_k = \lim_{n\to\infty} a_{k,n}$ exists for every $k \geq 0$ it is necessary and sufficient that the limit

$$A(s) = \lim_{n\to\infty} \sum_{k=0}^{\infty} a_{k,n} \cdot s^k$$

exists for each s in the open interval $0 < s < 1$. In this case

$$A(s) = \sum_{k=0}^{\infty} a_k \cdot s^k.$$

Inequalities: First and Second Moment Methods.

We review some inequalities that play a considerable role in probabilistic analysis of algorithms. In particular, we discuss first and second moment methods that are 'bread-and-butter' of a typical probabilistic analysis.

Markov Inequality: For a nonnegative random variable ξ and $\varepsilon > 0$ the following holds: $\Pr(\xi \geq \varepsilon) \leq \frac{E(\xi)}{\varepsilon}$.

Indeed: let $I\ (A)$ be the indicator function of A (i.e., $I\ (A) = 1$ if A occurs, and zero otherwise). Then, $E\ [\xi] \geq E\ (\xi \cdot I\ (\xi \geq \varepsilon)) \geq \varepsilon\, \mathrm{E}(I\ (\xi \geq \varepsilon) = \varepsilon \Pr(\xi \geq \varepsilon)$.

Chebyshev Inequality: If one replaces in the Markov inequality ξ by $|\xi - \mathrm{E}\,[\xi]|$ then $\Pr(|\xi - \mathrm{E}\,[\xi]| > \varepsilon) \leq \frac{Var\,(\xi)}{\varepsilon^2}$.

A. The problem

Our investigation, in general, belongs to the combinatorial theory of finite and constraint set systems described above. Doing this r we will construct and investigate probabilistic models complementary to the combinatorial analysis, intending to achieve in an alternative way to gain the necessary knowledge on structures and properties of these set systems – of subsets of E^{n} representing the set systems of a universal n-element set. In this way we face 2 basic problems: first is how to obtain meaningful postulations over the probabilistic models, and second is how to transfer this knowledge to the combinatorial domain of the basic addressed problems. To give an expression about the structures and properties it is enough to note the following. Erdős and Rényi [1-2] initiated the domain of random graphs but they considered a larger domain of set systems with different properties: constrained intersection, special coverage and others.

Glagolev [5] initiated the probabilistic study of Boolean functions through the description of sub-cubes of the truth domain.

Aslanyan and Akopova (Arsenyan) [11-21] applied the probabilistic technique to the domain of discrete isoperimetry.

The Hungarian and Russian scientific schools were the well-recognized centers of intensive use of probabilistic models of combinatorics.

B. The model

Probabilistic model of a combinatorial problem consists the set of all structures of the problem and extends probabilities on this set. Allocate probabilities to the variables that are generating the n-dimensional unite cube. Boolean functions will appear with corresponding probabilities. If we allocate probabilities to the vertices of E^n to be the truth-value of random Boolean function f, then Boolean functions will appear with different probabilities. The probabilistic models that are used in combinatorics is very large. For the beginning, in our study of random discrete tomography we will use one of these models.

The probabilistic model $\mathfrak{M}_{\mathrm{pq}}$.

Consider random Boolean $m \times n$ matrix R_{mn} generated in the result of evaluation of column variables $x_j, j = 1, \cdots, n$, that independently and identically (i.i.) attain values 1 and 0 with respective none-zero probabilities p_j and $q_j = 1 - p_j$. The number of all matrices that will be generated in this scheme is 2^{mn} (this is the set of all possible $m \times n$ matrices) and the probability of a particular matrix being generated is tightly related to the number of "1" values in its columns. Consider a matrix, and let its column weights are as $s_1, s_2, \dots, s_n$, then, the probability of this matrix in $\mathfrak{M}_{\mathrm{pq}}$ is equal to $\prod_{j=1}^{n} p_j^{s_j} q_j^{m-s_j}$. There are $\prod_{j=1}^{n} C_m^{s_j}$ different matrices with column weights $s_1, s_2, \dots, s_n$ so that the probability that the random matrix obey column weights $s_1, s_2, \dots, s_n$ equals

$$\prod_{j=1}^{n} C_m^{s_j} p_j^{s_j} q_j^{m-s_j}. \tag{2.4}$$

It is convenient to consider this probabilistic model as a process, where m vertices i.i. distributed, as it is defined above, are dropped onto the E^n. Vertices can appear repeatedly and only in cases when there are no vertex repetitions we receive an m-subset of the E^n (vertices given by the rows of the matrix). In an approach, -- the indicator to the existence of such row different matrices can be the related nonzero probability - in the given model $\mathfrak{M}_{\mathrm{pq}}$.

In a special case we will suppose that $p_j = s_j/m$ and intend to prove asymptotically, when $n, m \to \infty$ the following:

$Z1$. Probabilities that column sums are equivalent/equal to $s_1, s_2, \dots, s_n$ are positive and/or tend to 1.

$Z2$. Probability that all rows are different tends to 1, or is straightly grader than 0.

Consider an arbitrary column j. Let M_j and D_j be the average value and the dispersion of the random weight of column j. We will treat $Z1$ and $Z2$ on this probabilistic basis, and will also combine the problems $Z1$ and $Z2$.

Additionally we will consider the issue:

$Z3$. Probability that a random set of $\mathfrak{M}_{\mathrm{pq}}$ is a Sperner Family (SF) is positive under the special constraints.

Alternatively, several more probabilistic models besides the $\mathfrak{M}_{\mathrm{pq}}$ can be involved into this study but we postpone this for the continuation.

C. The technique

The technique used in this domain is diverse. The first step in many studies is the calculation of averages for the target properties. The main value shows the existence of a value that is greater (smaller) than this value. Proving the existence of certain types of structures is achieved by showing that the averages/probabilities are positive. Next group of considerations are in computation and use of the second moments. By the Chebyshev inequality, when the main value and the variance are appropriate, conclusions are made in terms of probabilistic convergences, in particular in terms of “almost all” structures. Another scheme is in use of the means of convergences of probabilistic distributions. In particular, this is through the estimation of factorial moments of random numbers, and then the continuity of distributions in a couple of cases brings us to the resulting Poisson or some other distribution. For example, this is the case of the random number of isolated truth-values of the random Boolean functions.

D. **Example 1. Positive probability implies the existence.**

In a simplest application of the probabilistic method to prove that a structure with the desired properties exists we define an appropriate probability space and then show that the desired properties hold in this space with positive probability. Our first example belongs to this case. We also decided the first example to be graph theoretical.

- Let $G = (V, E)$ be a graph with n vertices and m edges. Then G contains a bipartite sub-graph with at least $m/2$ edges.

The proof of such proposition is very simple. It is to generate random subsets X of vertices with $pr(v \epsilon X) = 1/2$. Then the average of number of edges linked to X is equal to $m/2$, which implies the existence of the desired bipartite graph. Continue applying this construction recursively, we see that the graph edges will expire in $logm$ steps. We see an interesting extension of the “bisection” principle in a form - applicable to the arbitrary graphs.

E. **Example 2. Probability distributions with the Chebyshev inequality give the asymptotics.**

Next to the “positive probability” level model it comes the model based on second probabilistic moments. Let us bring an example from the field of Boolean functions. Consider ordinary Boolean functions. We intend to derive the complexity asymptotic formula $s(f)$ of reduced disjunctive normal form of random Boolean functions. Consider the appropriate model. Let the “thru” values are generated randomly and uniformly on Boolean vertices by probability p. For the average number of the k dimensional prime sub-cubes we obtain the formula

$$i_k(n,p) = m(i_k(f)) = C_n^k 2^{n-k} p^{2^k} (1 - p^{2^k})^{n-k}.$$

C_n^k denote the so called “directions” of the k-sub-cubes. On these directions the second probabilistic moments are calculated:

$$\begin{aligned} d(i_k(f)) &= 2^{n-k} p^{2^k} (1 - p^{2^k})^{n-k} + C_{n-k}^2 2^{n-k} p^{2^{k+1}} (1 - p^{2^k})^{2(n-k-1)} \\ &\quad - (1 + (n-k) + C_{n-k}^2) C_{n-k}^2 2^{n-k} p^{2^{k+1}} \left(1 - p^{2^k}\right)^{2(n-k)}. \end{aligned}$$

Then, continuing with the use of the Chebyshev inequality we obtain that simultaneously in all C_n^k directions with the probability tending to 1, and $n \to \infty$, $i_k(f) \sim m(i_k(f))$.

Let $p = 1/2$. There is a unique integer point k_0 of maximum of the function $i_k(n,p)$. The result achieved says that for a random Boolean function, with probability tending to 1 with $n \to \infty$

$$s(f) \sim m\big(i_{k_0}(f)\big) + m\big(i_{k_0+1}(f)\big).$$

F. **Example 3. Continuity theorem helps to describe the types of the Boolean functions.**
Consider the next model according to the continuity theorem of the generating functions of factorial moments. We construct the model around the well-known discrete isoperimetry problem.

As usual, we denote the set of points of the n-dimensional unite cube by E^n. For a subset $A \subseteq E^n$ we call a point α *interior* if $S_1^n(\alpha) \subseteq A$ where $S_1^n(\alpha)$ is the sphere of radius 1 with centre α in Hamming metrics ρ. Let $B\ (A)$ denote the set of all interior points of the subset A. $\Gamma(A) = A - B(A)$ is called the *boundary* of the subset A.

Discrete isoperimetry is the problem of finding the value

$$\Gamma(\mathrm{a}) = \min_{B \subseteq E^n, |B| = a} |\Gamma(\mathrm{B})|$$

for given a, $0 \leq a \leq 2^n$.

For the ratio τ_n of the subsets $A \subseteq E^n$ with $|A| = (1 - \alpha(n)n)2^{n-1}$ and k interior points the following holds:

$$\tau_n \to \frac{1}{k!}\left(\frac{e^\lambda}{2}\right)^k e^{-\frac{e^\lambda}{2}}$$

if $\lim_{n\to\infty} \alpha(n)n = \lambda$ $(k = 0,1,2, \dots)$.

Random Discrete Tomography

G. **Series of compound trials.**

$\mathfrak{M}_{\mathrm{pq}}$, is one of considered probabilistic distribution schemes over the set E^n. Let $x_1, x_2, \dots, x_n$ be the generating binary variables of E^n. Consider a compound trial, where the variables $x_1, x_2, \dots, x_n$ independently of each other accept values 1 by corresponding probabilities $p_1, p_2, \dots, p_n$. By the complementary probabilities $q_i = 1 - p_i, i = \overline{1,n}$ these variables accept the value 0. Series of m such compound trials will be considered. Here, in particular, the random x_i remain under the same distribution throughout the sequence of compound trials, and, because of in each trial there are only two possible outcomes and their probabilities are unchanged, we deal with the Bernoulli trials at $i = 1,2, \dots, n$. We use a compound trial consisting of n variables and get one vertex of E^n as the outcome of each compound trial.

Consider $x \in E^n$. Which is the probability of appearance of this vertex x in the model $\mathfrak{M}_{pq}$? Since the individual variable trials in a compound trial are independent, the probabilities of individual variables multiply. In order to calculate the probability of the vertex x it is to take the product obtained on replacing the symbols 1 and 0 by p and q (depending on i), respectively. Doing m sequential compound trials generates a distribution over the multi-sets on E^n.

H. **The column weight probability and the average column weight.**

Consider the random variable ξ_j representing the number of 1's in the column j (we call it also column sum, weight, projection) in the model $\mathfrak{M}_{\mathrm{pq}}$. The probability of the value ξ_j in the random matrix R_{mn} equals

$$C_m^{\xi_j} p_j^{\xi_j} q_j^{m-\xi_j} \tag{3.1}$$

On this base for the expectation of the weight of column j we obtain the formula

$$\begin{aligned} M_j = M(\xi_j) = \sum_{t=0}^{m} t C_m^t p_j^t q_j^{m-t} &= \sum_{t=1}^{m} t \frac{m!}{t!\,(m-t)!} p_j^t q_j^{m-t} \\ &= m \sum_{t=1}^{m} \frac{(m-1)!}{(t-1)!\,(m-1-(t-1))!} p_j^t q_j^{m-t}. \end{aligned}$$

Substituting $u = t - 1$ we obtain

$$M_j = m \sum_{u=0}^{m-1} C_{m-1}^u \, p_j p_j^u q_j^{m-1-u} = m p_j.$$

For example, when $p_j = q_j = 1/2$ we receive parameters of the usual homogeneous model of random Boolean functions on E^n that are easy interpretable. In the weighted model, the overall average sums by the set of coordinates/columns will be $mp_1, mp_2, \dots, mp_n$. And in utilization of $p_j = \frac{s_j}{m}, j = 1, \cdots, n$ we obtain that the average columns sums vector equals to $s_1, s_2, \dots, s_n$. In this concern the model $\mathfrak{M}_{\mathrm{pq}}$ is the most convenient to the problem $Z1$. But it does not fit well to the requirements of $Z2$.

At this point we have 2 known parameters – the probability of the given column weight and the average value of the weight. The column weight t probability $C_m^t p_j^t q_j^{m-t}$ for $p_j = \frac{s_j}{m}$ is straightly greater than 0, which indicates the existence of construction with this weight t. Although trivial, this same postulation does not follow from the notion of the main value. The main value instead shows that a greater and/or lesser values (an integrative event) exist. These notes are applicable also to compound values because of the independency of the coordinates.

I. **The column weight variance.**

The additional use of the variance in model $\mathfrak{M}_{\mathrm{pq}}$ brings more points. Combined with the Chebyshev inequality this gives intervals around the values s_j with a property that the random sums belong to these intervals (to polyhedrons) with a strongly positive probability.

Let us compute the variance of ξ_j. There are 2 ways. One is in direct use of formula $D_j = M(\xi_j^2) - M_j^2$. But, there is an easier way of computing $M(\xi_j(\xi_j - 1))$, having in mind that $M(\xi_j(\xi_j - 1)) = M(\xi_j^2) - M(\xi_j)$ and that we already computed the $M(\xi_j)$. The benefit of this choice is seen below.

$$M(\xi_j(\xi_j - 1)) = \sum_{t=0}^{m} t(t-1) C_m^t p_j^t\, q_j^{m-t} = \sum_{t=2}^{m} t(t-1) \frac{m!}{t!\,(m-t)!} p_j^t\, q_j^{m-t}$$

$$= m(m-1) \sum_{t=2}^{m} \frac{(m-2)!}{(t-2)!\,(m-2-(t-2))!} p_j^t\, q_j^{m-t}.$$

Performing substitution $u = t - 2$ we obtain

$$M(\xi_j(\xi_j - 1)) = m(m-1) \sum_{u=0}^{m-2} C_{m-2}^u\, {p_j}^2 p_j^u q_j^{m-2-u} = m(m-1) p_j^2.$$

Substituting the values finally we obtain that

$$D_j = m(m-1)p_j^2 + mp_j - \left(mp_j\right)^2 = mp_j\left(1-p_j\right) = mp_jq_j.$$

By the Chebyshev inequality we obtain the following probability estimate:

$$P\left(\left|\xi_j - M_j\right| \geq \varepsilon M_j\right) \leq \frac{D_j}{\varepsilon^2 M_j^2} = \frac{mp_jq_j}{\varepsilon^2 m^2 p_j^2} = \frac{q_j}{\varepsilon^2 mp_j}. \tag{3.1}$$

Let $\varepsilon \to 0$ with $m \to \infty$. q_j and p_j may depend on m, but let they have the same order. It is easy to choose ε in a way that $\frac{1}{\varepsilon^2 m} \to 0$. So the probability of the event $P\ \left(\left|\xi_j - M_j\right| < \varepsilon M_j\right)$ tends to 1 and the relation achieved says, "with probability tending to 1 the random weight ξ_j is equivalent to M_j". So, the probability, that the random variable ξ_j of the component j is equivalent to its average value mp_j when $m \to \infty$, is tending to 1.

J. **Analysis of compound trials.**

Consider series of m compound trials by the model $\mathfrak{M}_{\mathrm{pq}}$. Because of the component probabilities $p_1, p_2, \ldots, p_n$ in $\mathfrak{M}_{\mathrm{pq}}$ act separately/independently, we have that any set of events, each defined in terms of one individual coordinate probability are totally independent. Denote the event $\left|\xi_j - M_j\right| \geq \varepsilon M_j$ by $\mho_j$ and consider the compound event $\underline{\mho} = \mho_1 \cup \mho_2 \cup \ldots \cup \mho_n$. Then, by (2.1) $P\left(\underline{\mho}\right) \leq \sum_{j=1}^{n} P\left(\mho_j\right)$. Denote the event $\left|\xi_j - M_j\right| < \varepsilon M_j$, complementary to $\mho_j$ as $\neg\mho_j$. Then for $\underline{\neg\mho} = \neg\mho_1 \cup \neg\mho_2 \cup \ldots \cup \neg\mho_n$ it is true that $P\left(\underline{\neg\mho}\right) \leq \sum_{j=1}^{n} P\left(\neg\mho_j\right)$. These relations are true for arbitrary sets of events.

Events $\mho_j$ are independent by the notion above, and all they have the same structure of the probability. In this way, if $\overline{\mho} = \mho_1 \cap \mho_2 \cap \ldots \cap \mho_n$, then $P\left(\overline{\mho}\right) = \prod_{j=1}^{n} P\left(\mho_j\right)$. Also the events $\neg\mho_j$ are totally independent so that for $\overline{\neg\mho} = \neg\mho_1 \cap \neg\mho_2 \cap \ldots \cap \neg\mho_n$ it is valid the expression $P\left(\overline{\neg\mho}\right) = \prod_{j=1}^{n} P\left(\neg\mho_j\right)$.

By $\mathfrak{M}_{\mathrm{pq}}$ we have a very peculiar model that gives us more useful relations:

$P\left(\underline{\mho}\right) = 1 - \prod_{j=1}^{n}\left(1 - P\left(\mho_j\right)\right) = 1 - \prod_{j=1}^{n} P\left(\neg\mho_j\right)) = 1 - P\left(\overline{\neg\mho}\right)$

and

$P\left(\underline{\neg\mho}\right) = 1 - \prod_{j=1}^{n}(1 - P\left(\neg\mho_j\right)) = 1 - \prod_{j=1}^{n} P\left(\mho_j\right) = 1 - P\left(\overline{\mho}\right).$

We proceed to derive the “compound” generalization of point III.C. We use the set of Chebyshev inequalities for coordinates:

$$P\left(\mho_j\right) = P\left(\left|\xi_j - M_j\right| \geq \varepsilon M_j\right) \leq \frac{D_j}{\varepsilon^2 M_j^2} = \frac{q_j}{\varepsilon^2 m p_j}.$$

Let us mention also the equivalent forms

$$P\left(\neg\mho_j\right) = P\left(\left|\xi_j - M_j\right| < \varepsilon M_j\right) \geq 1 - \frac{D_j}{\varepsilon^2 M_j^2} = 1 - \frac{q_j}{\varepsilon^2 m p_j}.$$

Note that $\underline{\mho}$ means that there exists a more that ε divergence from the main value at least in one of the coordinates. And $\overline{\neg\mho}$ denotes the event wherein simultaneously in all coordinates deviations are less than ε. Concluding, -- our aim is to find the conditions when $P\left(\underline{\mho}\right)$ is small and/or when $P\left(\overline{\neg\mho}\right)$ is large (which is the same in our case of $\mathfrak{M}_{\mathrm{pq}}$).

Start with $P\left(\underline{\mho}\right) \leq \sum_{j=1}^{n} P\left(\mho_j\right)$. Substituting inequalities $P\left(\mho_j\right) \leq \frac{q_j}{\varepsilon^2 m p_j}$ into this formula we receive

$$P\left(\underline{\mho}\right) \leq \frac{1}{\varepsilon^2 m}\left(\frac{q_1}{p_1} + \frac{q_2}{p_2} + \cdots + \frac{q_n}{p_n}\right) \leq \frac{n}{\varepsilon^2 m} \frac{\max_{1\leq i\leq n} q_i}{1 - \max_{1\leq i\leq n} q_i}.$$

To get an applicable result let us suppose that $\frac{n}{\varepsilon^{2} \cdot m} \to 0$ asymptotically and that all factors $\frac{q_i}{p_i}$ are limited. In these conditions we receive that in the model $\mathfrak{M}_{\mathrm{pq}}$, having a series of m trials, the probability of a sensitive deviation from the mean value, -- at least in one of the coordinates tends to 0.

We may also use the formula $P\left(\underline{\mho}\right) = 1 - \prod_{j=1}^{n} P\left(\neg\mho_j\right))$ complemented with the $P\left(\neg\mho_j\right) \geq 1 - \frac{q_j}{\varepsilon^2 m p_j}$. To get an applicable postulation from this, it is to require that

$$\lambda = \prod_{j=1}^{n} (1 - \frac{q_j}{\varepsilon^2 m p_j}) \to 1.$$

$$\lambda \geq \prod_{j=1}^{n} \left(1 - \frac{\max_{1 \leq i \leq n} q_i}{\varepsilon^2 m \left(1 - \max_{1 \leq i \leq n} q_i\right)}\right) = \left(1 - \frac{\max_{1 \leq i \leq n} q_i}{\varepsilon^2 m \left(1 - \max_{1 \leq i \leq n} q_i\right)}\right)^n.$$

Apply the following formula: if $0 \leq x \leq 1/2$ and $0 \leq y$, then $\exp(-x(1-x)y) \leq (1-x)^y$. We get

$$\lambda \geq \exp\left(-\frac{n \cdot \max_{1 \leq i \leq n} q_i}{\varepsilon^2 m \left(1 - \max_{1 \leq i \leq n} q_i\right)}\left(1 - \frac{\max_{1 \leq i \leq n} q_i}{\varepsilon^2 m \left(1 - \max_{1 \leq i \leq n} q_i\right)}\right)\right).$$

We arrived to the same condition. To get an interpretable result it is to suppose that $\frac{n}{\varepsilon^{2} \cdot m} \to 0$ asymptotically, and that all factors $\frac{q_i}{p_i}$ are limited. In these conditions we receive that in the model $\mathfrak{M}_{\mathrm{pq}}$, having a series of m trials, the probability of a sensitive deviation from the mean value in at least in one of the coordinates tends to 0.

Recall out main target. Our interest is in a situation when column weights of an $\mathfrak{M}_{\mathrm{pq}}$ random matrix are close to the given $s_1, s_2, \dots, s_n$, and the rows of the matrix are all different. Ideally, $n \geq$logm is “satisfactory” for the row difference. And $\frac{q_i}{p_i}$ limited is also an acceptable condition (but not necessary), because of this is the case when each column participates in row differentiation.

This is our result for the point $Z1$. The domain described by the above intervals is a rectangular area in the space of all sum vectors space Ξ_m^n and the achieved property insists that there exist a proper random sum vector that belongs to the indicated rectangular area. Setting $s_1, s_2, \dots, s_n$ arbitrarily, we receive corresponding rectangular area of different size and probability (it can be also empty). Unless attractive, the property in this form is not yet useful, because of we do not know if the rows of random matrix that are different in this case.

The strategy at this point is:

α: n independent coordinates run in m-dimensional unit cubes each. If to consider the most transparent case $p = \frac{1}{2}$, then m-column-evaluations are equally probable with probability $\frac{1}{2^m}$. If $s = m/2$, then there are $\sim 2^m/\sqrt{m}$ evaluations with this s and the concluding probability is $1/\sqrt{m}$ in one direction, and, -- $m^{-n/2}$ in integration by the n coordinates. This is a small value, but still positive, that shows the existence of s-weighted vectors. All weights are possible but probabilities are different. And s is the central term of this distribution.

β: The situation with low probabilities can be softened a bit. The way is in considering the equivalency classes to the s vectors [12-29]. For simplicity consider the n-cube when m is odd. This gives a 2^n multiplier (the size of the equivalency class) to the probability that now becomes $m^{1-n/2}$. Here we suppose (by bisection) that that the most acceptable value is close to the $n = \log_2 m$. This is a higher probability but the difference is not sensitive. The next step forward is:

γ: For each coordinate consider an interval of length $\sqrt{m}$. In the composed rectangular area there will be $m^{n/2}$ points and this gives a constant probability to the considered event. But this may only speak about the existence of a point nearby the vector s, in a rectangular area.

The further idea is to find a type of independency between the events related to the weights s, and the events of the row-difference. It is also to be able to apply the part of proof on row-difference to the parts of distribution by the weights, or to its central term, which we adopt to be the compound weight s.

K. The row differences model.

This point deals with the model $\mathfrak{M}_{\mathrm{pq}}$ considering random matrices R_{mn} obtained by $\mathfrak{M}_{\mathrm{pq}}$, evaluating probability of matrices under the constraint of having no repeated rows. For an arbitrary matrix R_{mn} we generate a correlated with it matrix $D_{C_m^2 n}$, that consists of all comparisons of pairs of rows of R_{mn} in the following way [6]. $D_{C_m^2 n}$ consists of $m-1$ separate parts, n-column sub-matrices, which are concatenated vertically. First sub-matrix has $m-1$ rows that represent coordinate wise $mod2$ summations of the first row of R_{mn} with the reminding rows $2,3,\dots,m$. Denote this sub-matrix by $D_{(m-1)\,n}$. The next sub-matrix $D_{(m-2)\,n}$ is composed by $m-2$ rows generated from $D_{(m-1)\,n}$ in the same way (first row with other rows). The last group $(m-1)$ will be a 1-row matrix, $D_{(m-(m-1))n} = D_{1n}$.

R_{mn} and $D_{C_m^2 n}$ are straightly related to each other by the following important properties. Rows $r \epsilon D_{C_m^2 n}$ correspond to pairs of rows of the matrix R_{mn} in a way that if s is the number of sub-matrix the row r belongs to, and t is the sequential number of this row in $D_{(m-s)\ n}$, then r have the property:

- The "1" value in a coordinate of r corresponds to the "difference" in the row pair (s, t) of R_{mn} in that coordinate, and
- The norm of r represents the Hamming distance of rows s and t.

These notes may have one more important interpretation:

- R_{mn} consists of all different pairs of rows if and only if $D_{C_m^2 n}$ does not include the row with all 0 coordinates.

The notes are reducing the problems with conditions of "difference" of all pairs of rows - to a specific set cover problems, with cover sub-sets composed from the rows of matrix $D_{C_m^2 n}$. In terms of R_{mn} "differences" must cover all pairs of rows. If to recall that the columns of R_{mn} are weighted, then the "differences" introduced by an individual column compose a bipartite graph, so that in fact the appearing set cover interpretation is very much specific and it works with a cover by a set of n bipartite graphs. Also it is to mention that when R_{mn} is a random matrix constructed directly, then its $D_{C_m^2 n}$ is a secondary construction, it is not given and not visible, so there is no direct way to check if it contains the all 0 row or not. Two frames are used to estimate the probabilities of matrices that have no repeated rows.

- ***Different Coordinates.*** Firstly, we prove that in the considered random generation columns of $D_{C_m^2 n}$ and its sub-matrices homogeneously appear with high weights (that represent row differences). Then we use the greedy estimation of the columns that are able to cover the rows of $D_{C_m^2 n}$ – in this way they cover all pairs of initial rows and the rows appear different.
- ***Different Vectors.*** In a second approach we study the probabilities of pairs of rows to be different. Extending this property to all pairs we get a lower estimate of the probability that m random rows are all different. This is valid/acceptable for some constraints over the m and n. And of course we follow with combining the postulations of this section with the ones about the column weights to get the proper estimations for the discrete tomography problem.

Different Coordinates

Next random variable of our consideration is the number of different coordinates in the pairs of rows. The expected number of different coordinates in the pairs of rows in the one-column model (having only one coordinate) can be calculated as:

$$\sum_{t=0}^{m} t(m-t)C_m^t p_j^t\, q_j^{m-t} = \sum_{t=1}^{m-1} t(m-t)C_m^t p_j^t\, q_j^{m-t} = m(m-1)p_j q_j \sum_{t=0}^{m-2} C_{m-2}^t p_j^{t-1}\, q_j^{m-2-t}$$
$$= m(m-1)p_j q_j.$$

Probabilities of random difference of coordinates are computed and estimated in a regular way, and their use is tightly correlated to the matrix $D_{C_m^2 n}$. The idea here is to follow the property of $D_{C_m^2 n}$ to have 1's in rows having large values of the coordinate differences. Details of this part partially repeat the above narration and preferred to be a subject of a separate publication.

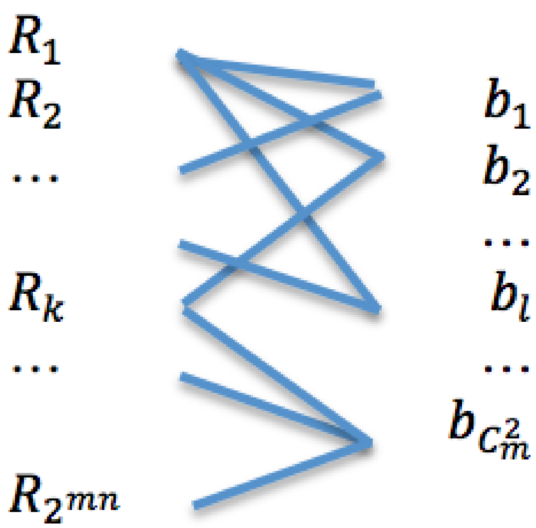
Different Vectors

Let us consider the random variable Δ_{mn}, -- the number of different pairs of rows in the matrices at the model $\mathfrak{M}_{\mathrm{pq}}$.

Let $R_1, \dots, R_k, \dots, R_{2^{mn}}$ are all $m \times n$ matrices that may appear randomly at the model $\mathfrak{M}_{\mathrm{pq}}$. Let $p\ (\mathrm{R}_k)$ is the probability of R_k, and let Δ_k is the number of "different" pairs of rows in R_k. Then the main value of the number of "different" pairs of rows at the random outcome of $\mathfrak{M}_{\mathrm{pq}}$ can be presented as:

$$M(\Delta_{mn}) = \sum_{\mathrm{R}_k \in \mathfrak{M}_{\mathrm{pq}}} p(\mathrm{R}_k) \cdot \Delta_k.$$

It is imperceptible how $p(\mathrm{R}_k)$ and Δ_k can be brought to a concise computable form. To simplify the formula we consider the following standard [5] "bipartite" scheme:



Right side vertices c_l represent all pairs of rows in the $m \times n$ matrices, listed in some fixed order. Edges, in this bipartite graph scheme connect R_k and b_l iff the pair b_l in R_k is “different” (consists of different rows). Vertex degrees at all b_l constantly are equal to $2^n(2^n-1)$. But edges are weighted, and the weight of link between R_k and b_l represents the probability that b_l as a “different” pair appears in R_k.

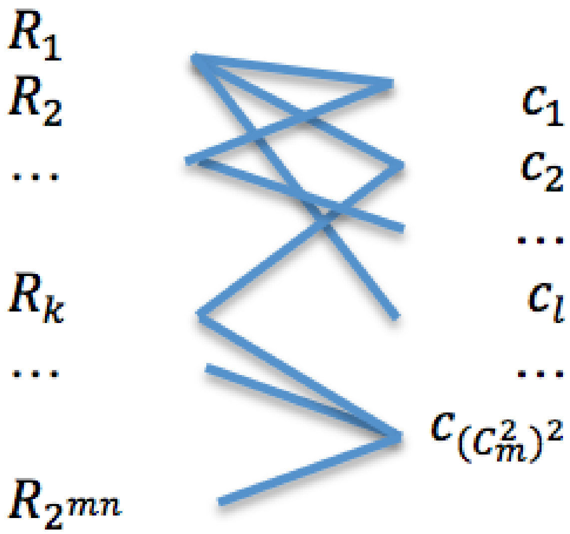
Consider an arbitrary pair of random rows. The probability that for a particular j-th coordinates on the considered pair of rows are identical is evidently $p_j^2+q_j^2$. Due to $1=p_j+q_j$ it is true that $p_j^2+q_j^2=1-2p_jq_j$ so that this forms will be equivalently exchanged in need in our narration. The probability that the entire rows are identical (all coordinates) equals to $\prod_{j=1}^{n}(1-2p_jq_j)$ and the complementary probability that this rows are different will be some $\alpha=1-\prod_{j=1}^{n}(1-2p_jq_j)$ (for example when $p_j=q_j=1/2$ then we receive the well known $\alpha=1-\frac{1}{2^n}$). We obtained that the probability that arbitrary b_l to be “different” is equal to $1-\prod_{j=1}^{n}(1-2p_jq_j)$ and finally, for the average number of “different” pairs of rows we obtain the following concise formula:

$$M(\Delta_{mn})=\sum_{\mathrm{R}_k\in\mathfrak{M}_{\mathrm{pq}}} p(\mathrm{R}_k)\cdot\Delta_k=C_m^2\left(1-\prod_{j=1}^{n}(1-2p_jq_j)\right).$$

In a similar way we may obtain the formula for the variance of Δ_{mn}. By the definition

$$D(\Delta_{mn}) = \sum_{\mathrm{R}_k \in \mathfrak{M}_{\mathrm{pq}}} p(\mathrm{R}_k) \cdot {\Delta_k}^2.$$

We apply again the formula $D(\Delta_{mn}) = M(\Delta_{mn}^2) - \left(M(\Delta_{mn})\right)^2$. And consider an analog of the bipartite scheme with random matrices. Matrices $R_1, \dots, R_k, \dots, R_{2^{mn}}$ appear by the scheme $\mathfrak{M}_{\mathrm{pq}}$ with the probabilities $p(\mathrm{R}_k)$.



Right part vertices c_l correspond to all pairs of pairs of vertices of matrices. First to note is the formula

$$M(\Delta_{mn}^2) = \sum_{c_{l'}, c_{l''}} p(c_{l'}, c_{l''}).$$

Split the set of all pairs $b_{l'}, b_{l''}$ into the classes:

- At first – consider the class of the pairs (pairs of pairs of rows) of the type b_l, b_l. Number of this pairs is C_m^2 and they all have the probability of “difference” 1-$\prod_{j=1}^{n}\left(1-2p_jq_j\right)$.
- The largest group consists of pairs $b_{l'}$ and $b_{l''}$ that have no common row. Number of this pairs is $C_m^2 C_{m-2}^2$ and they have probability of difference: $1 - \prod_{j=1}^{n}\left(1-2p_jq_j\right)^2$.
- The last group of pairs includes the 2 pairs of rows with one common row. Number of these fragments is equal to $C_m^2 2(m-2) = m\,(m-1)(m-2)$. And the probability of difference is: $1 - 2\prod_{j=1}^{n}\left(1-2p_jq_j\right) + \prod_{j=1}^{n}\left({p_j}^3 + {q_j}^3\right)$.

Combining into the general formula we obtain

$$M(\Delta_{mn}^2) = C_m^2\left(1-\prod_{j=1}^{n}(1-2p_jq_j)\right) + C_m^2C_{m-2}^2\left(1-\prod_{j=1}^{n}(1-2p_jq_j)^2\right) + C_m^2 2(m-2)\left(1-2\prod_{j=1}^{n}(1-2p_jq_j)+\prod_{j=1}^{n}({p_j}^3+{q_j}^3)\right) \tag{3.2}$$

We aim to apply the Chebyshev inequality. To obtain a "for almost all" type result it is to prove that with $m,n\to\infty$ it is true that $\frac{D}{\varepsilon^2M^2}\to 0$. In this case we check if $\frac{D}{M^2}\to 0$ and then it is simple to take a proper $\varepsilon,\varepsilon\to 0$ and this proves the required "almost all" type result. Consider the proposition $\frac{M(\Delta_{mn}^2)}{(M(\Delta_{mn}))^2}\to 1$ that is equivalent to $\frac{D}{M^2}\to 0$. For $p_j=q_j=1/2$ the sub-formula $\prod_{j=1}^{n}(1-2p_jq_j)=2^{-n}$ so that it tends to 0. In our case of arbitrary probabilities p_j, sub-formula $\prod_{j=1}^{n}(1-2p_jq_j)$ exceeds 2^{-n} but it is acceptable to seem that $\prod_{j=1}^{n}(1-2p_jq_j)\to 0$. In this case it is easy to be convinced that the first and the last summands of $M(\Delta_{mn}^2)$ are $o\ (M^2)$. For the analysis of the midterm, let us note that

$$\frac{1-\prod_{j=1}^{n}(1-2p_jq_j)^2}{\left(1-\prod_{j=1}^{n}(1-2p_jq_j)\right)^2}=\frac{1+\prod_{j=1}^{n}(1-2p_jq_j)}{1-\prod_{j=1}^{n}(1-2p_jq_j)}=1+\frac{2\prod_{j=1}^{n}(1-2p_jq_j)}{1-\prod_{j=1}^{n}(1-2p_jq_j)}.$$

Having this and $\prod_{j=1}^{n}(1-2p_jq_j)\to 0$ we conclude that the request $\frac{D}{M^2}\to 0$ is valid and that the Chebyshev inequality lets us to obtain the required "almost all" type postulation for the random number of the "different" rows at $\mathfrak{M}_{\mathrm{pq}}$. Denote

$$\chi=\prod_{j=1}^{n}(1-2p_jq_j).$$

Note 1.: $M(\Delta_{mn})$ have the form $C_m^2(1-\chi)$ with $\chi\to 0$. Consider the case when $C_m^2\chi\to 0$. As this is the average number, there must be an outcome of trial, -- the matrix R_k, so that the offset number of different rows is $\le C_m^2\chi$. As $C_m^2\chi$ becomes <1 this means a trivial thing – existence of an m-sub-set of the cube. Markov inequality does not help as well.

Claim 1.: Now we involve the variance into the game. By the Chebyshev inequality we have

$$P\left(|\Delta_{mn} - M(\Delta_{mn})| > \varepsilon M(\Delta_{mn})\right) \leq \frac{D(\Delta_{mn})}{\varepsilon^2 \left(M(\Delta_{mn})\right)^2}. \quad (3.3)$$

Consider the complementary to the $|\Delta_{mn} - M(\Delta_{mn})| \geq \varepsilon M(\Delta_{mn})$ event, and rewrite the inequality (3.3) in the form:

$$P\left((1-\varepsilon)M(\Delta_{mn}) < \Delta_{mn} < (1+\varepsilon)\, M(\Delta_{mn})\right) \geq 1 - D(\Delta_{mn})/\varepsilon^2\left(M(\Delta_{mn})\right)^2$$

and use its extension:

$$P\left((1-\varepsilon)M(\Delta_{mn}) < \Delta_{mn}\right) \geq 1 - \frac{D(\Delta_{mn})}{\varepsilon^2 \left(M(\Delta_{mn})\right)^2}. \quad (3.4)$$

In a way similar to the considerations of Note 1. we obtain that in proper selection of ε, and applying reasonable constraints on χ (this means – constraints on probabilities $p_1, p_2, \dots, p_n$), we may ensure that $(1-\varepsilon)M(\Delta_{mn}) > C_m^2 - 1$. This implies that (3.4) estimates the probability of $\Delta_{mn} = C_m^2$, i.e. the probability that the random R_{mn} will have all-different rows.

Note 2. We conclude in 2 steps. First – we further estimate the formula in (3.4). Then we consider the probability estimates of R_{mn} with weights $s_1, s_2, \dots, s_n$. We sum probabilities of these two events. When the sum becomes > 1 this implies that the two events are intersecting. Intersection means existence of a random outcome of trial $\mathfrak{M}_{\mathrm{pq}}$ with weights $s_1, s_2, \dots, s_n$ and with all-different rows. Ignoring the probability of this integrative event and just requiring that it be positive, we obtain the statement on existence of constructions with weights $s_1, s_2, \dots, s_n$ and with different rows.

Evaluate the right side formula of (3.4). We intend to add this formula to (2.4) finding out the conditions for this sum to be > 1. Then wee simply look for conditions of

$$\prod_{j=1}^{n} C_m^{s_j} {p_j}^{s_j} {q_j}^{m-s_j} > \frac{D(\Delta_{mn})}{\varepsilon^2 \big(M(\Delta_{mn})\big)^2}. \tag{3.5}$$

Our analysis is asymptotical, in $n, m \to \infty$. At this point we suppose that $\frac{D}{M^2} \to 0$ and in this condition we choose ε in a way that $\frac{D}{\varepsilon^2 M^2} \to 0$. We already checked these conditions for $p_j = 1/2$. Now let us consider the right side of (3.5).

Apply $\prod_{j=1}^{n}\left({p_j}^2 + {q_j}^2\right) \geq \prod_{j=1}^{n}\left({p_j}^3 + {q_j}^3\right)$ on the last term of (3.2), and combining the first ad last terms of this formula we obtain that this sum is:

$$\leq C_m^2(2m-3)\left(1 - \prod_{j=1}^{n}\left(1 - 2p_j q_j\right)\right) = a.$$

Compose the following difference with the midterm of (3.2)

$$(C_m^2)^2\left(1 - \prod_{j=1}^{n}\left(1 - 2p_j q_j\right)^2\right) - C_m^2 C_{m-2}^2\left(1 - \prod_{j=1}^{n}\left(1 - 2p_j q_j\right)^2\right)$$
$$= C_m^2(2m-3)\left(1 - \prod_{j=1}^{n}\left(1 - 2p_j q_j\right)^2\right) = b.$$

Additionally, denote

$$C_m^2 C_{m-2}^2\left(1 - \prod_{j=1}^{n}\left(1 - 2p_j q_j\right)^2\right) = c.$$

In (3.2) we want to delete the minor term a keeping only the major term c. The objective is in keeping (3.2) increasing or, -- in the same order. The mentioned change is possible due to $a \ll c$. It is also challenging replacing c by the term $(C_m^2)^2\left(1 - \prod_{j=1}^{n}\left(1 - 2p_j q_j\right)^2\right)$ but this is not acceptable due to $a < b$.

This analysis helps to correctly use the (3.5) for particular $p_1, p_2, \dots, p_n$. The concise estimation of (3.5) can be done for some regular examples of $p_1, p_2, \dots, p_n$. Two things are to be elaborated: does the probabilistic method bring some knowledge on tomography, and how accurate is this result. To get an idea on the last point let us consider the example of $p_j = 1/2$ treating the case $s_j = m/2$. Above, for this case we accepted the constraint $m^2 \ll 2^n$. That is, $2\log m \ll n$. It is evident that for $s_j = m/2$ and $\log m \le n$ the required tomography matrix exist. So if successful the probabilistic method requires al least 2 times more columns to differentiate the rows. The hope is that for nontrivial $p_1, p_2, \dots, p_n$ this, even in approximation, may bring an additional knowledge about the tomographic property.

L. **Sperner families.**

In a short note consider case of Sperner families. Many of the existence issues about the Sperner families are already resolved – the maximal Sperner family, almost all Sperner families, weighted Sperner families (by recursive apply of Kruskal-Katona theorem [60]). To understand the relation between the random sets and the Sperner property consider the bipartite graph with left side, including all m-subsets, and with right side, that consists of all pairs of comparable vertices of E^n. Compute the average number of comparable pairs of vertices in all Sperner families:

$$\varepsilon = \frac{\sum_{i=0}^{n} C_n^i \left(2^i - 1\right) C_{2^n-2}^{m-2}}{C_{2^n}^m} = \frac{m(m-1)}{2^n(2^n-1)}(3^n - 2^n)$$

Let $m^2 = o(1.333\dots^n), \mathrm{n} \longrightarrow \infty$ then ε is nearly zero value that indicates that there exist a Sperner family of size m, or in more precise, that the random subset of this size is a Sperner family.

Conclusion

Considerations above intend to get an additional knowledge about the row-different matrices of the discrete tomography problem, using the probabilistic theory of combinatorics [1-21]. The objective is reasonable because the pure combinatorial approaches [22-77] are not able at the moment to give an efficient description of the column weighted row-different matrices. The probabilistic method gives knowledge on random subsets, which might be useful as a complementary knowledge about a different object or a situation concerned to discrete tomography.

Bibliography

1. Erdős P., Graph theory and probability, *Canad. J. Math.,* volume 11, pp. 34-38, 1959, doi:10.4153/CJM-1959-003-9.MR 0102081.
2. Erdős P., Rényi A., On Random Graphs, I, Publicationes Mathematicae, 6, pp. 290–297, 1959.
3. Gilbert E. N., Random Graphs, Annals of Mathematical Statistics 30 (4), pp. 1141–1144, 1959.
4. Erdős P., Rényi A., On the evaluation of random graphs, Publications of the Mathematical Institute of the Hungarian Academy of Sciences, V, pp. 17-61, 1960.

5. Glagolev V., Some estimates of disjunctive normal forms in the algebra of logic, *Problemy Kybernetiki* (in Russian), vol. 19, Nauka, Moscow, pp. 75-94, 1967.
6. Kospanov E., On algorithm of synthesis of rather simple tests, Discrete Analysis (in Russian), vol. 8, pp. 43-47, 1966.
7. Нечипорук Э. И., О топологических принципах самокорректирования, Проблемы кибернетики, 21, стр. 5-102, Наука, Москва, 1969.
8. Дискретная математика и математические вопросы кибернетики. Т.1., М., Наука, 1974.
9. Erdős P., Spencer J., Probabilistic Methods in Combinatorics, Akademia Kiado – Budapest, 1974.
10. Alon N., Spencer J., The Probabilistic Method (3rd edition), Wiley-Interscience, 352p., 2008.
11. Асланян Л.А., О сложности сокращенной дизъюнктивной нормальной формы частичных булевых функций. I., Ученые Записки, Естественные науки, ЕрГУ, 1, 1974, pp. 11-18.
12. Асланян Л.А., О сложности сокращенной дизъюнктивной нормальной формы частичных булевых функций. II., Ученые Записки, Естественные науки, ЕрГУ, 3, 1974, pp.16-23.
13. Асланян Л.А., О применениях сокращенной дизъюнктивной нормальной формы в задачах доопределения частичных булевых функций, Молодой научный работник, Естественные науки, ЕрГУ, 20(2), 1974, pp. 65-75.
14. Асланян Л.А., Об одном методе распознавания, основанном на разделении классов дизъюнктивными нормальными формами, Кибернетика, 5, Киев, 1975, pp. 103-110.
15. Асланян Л.А., Алгоритмы распознавания с логическими отделителями, Сборник работ по Математической кибернетике, вып. 1, ВЦ АН СССР, Москва, 1976, pp. 116-131.
16. Асланян Л.А., И.А.Акопова, Докозательства некоторых оценок сокращенных дизъюнктивных нормальных форм булевых функций, Ученые Записки, Естественные науки, ЕрГУ, 1,1980, 14-23.
17. Aslanyan L.A., I.A.Akopova, On the distribution of the number of interior points in subsets of the n-dimensional unit cube, Colloquia Mathematika Societatis Yanos Bolyai, 37, Finite and Infinite Sets, (Eger) Hungary, 1981, pp. 47-58.
18. Асланян Л.А., О длине кратчайшей дизъюнктивной нормальной формы слабо определенных булевых функций, Прикладная математика, ЕрГУ, 2, 1982, pp. 32-42.
19. Асланян Л.А., Дискретная изопериметрическая проблема - асимптотический случай, Доклады Академии наук, Арм.ССР, 74 (3), Ереван, 1982, pp. 99-103.
20. Асланян Л.А., К вопросу минимизации систем слабо определенных булевых функций. Некоторые задачи автоматизации проектирования, SzTAKI, MTA, Tanulmanyok, 135, Budapest, 1982, pp. 51-86.

21. Асланян Л.А., О единственности компактных булевых функций с точностью до изоморфизма, Working Paper, MTA SzTAKI,E/17 Budapest, 1983, pp. 1-11.

22. Aslanyan L., Sahakyan H., Numerical characterization of n-cube subset partitioning, Electronic Notes in Discrete Mathematics, Volume 27, Pages 3-4/110, October 2006, Elsevier B.V.,ODSA 2006 - Conference on Optimal Discrete Structures and algorithms.

23. Aslanyan, H. Sahakyan, A. Hovsepyan On constrained convexity tomography and Lagrangean approximations, Mathematical Problems of Computer Science, XXX, 2008, ISSN 0131-4645, pp.111-122.

24. H. Sahakyan, Numerical characterization of n-cube subset partitioning, Discrete Applied Mathematics, 157 (2009), pp. 2191-2197.

25. K. Ivanova, I. Mitov, K. Markov, P. Stanchev, K. Vanhoof, L. Aslanyan, H. Sahakyan, Metric Categorization Relations Based on Support System Analysis, Computer Science and Information Technologies Conference, Yerevan, Sept. 28 – Oct. 2, pp. 85-88, 2009.

26. Sahakyan Hasmik, “(0,1)-matrices with different rows”, Ninth International Conference on Computer Science and Information Technologies, Revised Selected Papers, IEEE conference proceedings (November, 2013).

27. Sahakyan Hasmik, “On Hypergraph Degree Sequence Characterization”, Forty-Fourth Southeastern International Conference on Combinatorics, Graph Theory, and Computing, March 4-8, 2013, Florida Atlantic University.

28. Sahakyan Hasmik, “On the set of simple hypergraph degree sequences”, Seventh Czech-Slovak International Symposium on Graph Theory, Combinatorics, Algorithms and Applications, 7-13 July, 2013, Kosice, Slovakia.

29. Levon Aslanyan, Hans-Dietrich Gronau, Hasmik Sahakyan, Peter Wagner, Constraint Satisfaction Problems on Specific Subsets of the n-Dimensional Unit Cube, CSIT 2015, Revised Selected Papers, IEEE conference proceedings, p.47-52, DOI: 10.1109/CSITechnol.2015.7358249

30. Aslanyan L., The discrete isoperimetry problem and related extremal problems of discrete spaces, Problemy Kybernetiki, 36, pp. 85-127, 1979.

31. Aslanyan L., Recognition method based on classification by disjunctive normal forms, *Cybernetics*, no. 5, pp. 103-110, 1975.

32. Aslanyan L., Danoyan H., Complexity of Hash-Coding type search algorithms with perfect codes, Journal of Next Generation Information Technology, volume 5, number 4, November 2014, pp. 26-35.

33. Aslanyan L., Danoyan H., On the optimality of the Hash-Coding type nearest neighbour search algorithm, Selected Revised Papers of 9th CSIT conference, IEEE Xplore, 2013.

34. Aslanyan L., Castellanos J., Logic based pattern recognition - ontology content (1), *International Journal Information Technologies and Applications*, volume 1, pp. 206-210, 2007.

35. Aslanyan L. and Ryazanov V., Logic based pattern recognition - ontology content (2), *Information Theories and Applications*, volume 15, number 4, pp. 314-318, 2008.

36. Barcucci E., Del Lungo A., Nivat M., Pinzani R., Reconstructing convex polyominoes from horizontal and vertical projections, Theor. Comput. Sci., 155, pp. 321–347, 1996.

37. Behrens Sarah, Erbes Cathy, Ferrara Mike, Hartke Stephen, Reinger Ben, Spinoza Hannah, Tomilson Charles, New results on degree sequences of uniform hypergraphs, Electronic Journal of Combinatorics, 20, number 4, 2013.

38. Berge C., Graphs and Hypergraphs, North Holland Publishing Company, Amsterdam, 1973.

39. Berge C., Hypergraphs: Combinatorics of Finite Sets, North-Holland, 1989.

40. Billington D., Conditions for degree sequences to be realisable by 3-uniform hypergraphs, The Journal of Combinatorial Mathematics and Combinatorial Computing, 3, pp. 71-91, 1988.

41. Billington D., Lattices and Degree Sequences of Uniform Hypergraphs, Ars Combinatoria, 21A, pp. 9-19, 1986.

42. Billington D., The (s,p)-exchange property, Ars Combinatoria, 23, pp. 185-200, 1987.

43. Bhanu Murthy N. L., Murali K. Srinivasan, The polytope of degree sequences of hypergraphs, Linear Algebra Appl. 350 pp. 147–170, 2002.

44. Bollobás B., Random Graphs (2nd ed.), Cambridge University Press, 2001.

45. Boonyasombat V., Degree sequences of connected hypergraphs and hypertrees, Lecture Notes in Math., 1073, pp. 236-247, 1984.

46. Brualdi R. A., Matrices of zeros and ones with fixed row and column sum vectors, Linear Algebra Appl. 33 pp. 159-231, 1980.

47. Chrobak, M., Durr, C., Reconstructing hv-convex polyominoes from orthogonal projections, Inform. Process. Lett. 69(6), pp. 283–289, 1999.

48. Colbourn Charles J., Kocay W.L. and Stinson D.R., Some NP-complete problems for hypergraph degree sequences, Discrete Applied Mathematics 14, pp. 239-254, 1986.

49. Demange M., Monnot J., An introduction to inverse combinatorial problems, *Paradigms of Combinatorial Optimization: Problems and New Approaches*, John Wiley & Sons Inc., Hoboken, NJ, USA, volume 2, 2013.

50. Engel K.. Sperner Theory, Encyclopaedia of Mathematics and its Applications, volume 65, Cambridge University Press, New York, 1997, 417p.

51. Feller W., An Introduction to Probability Theory and its Applications, volume I, 3rd (revised) edition, ISBN: 9780471257080.

52. Garey M. R., Johnson D. S., Computers and Intractability: A Guide to the Theory of NP Completeness, Freeman and Company, 1979.

53. Glagolev V., Some estimates of disjunctive normal forms in the algebra of logic, *Problemy Kybernetiki* (in Russian), vol. 19, Nauka, Moscow, pp. 75-94, 1967.

54. Gnedenko B., On a local limit theorem of probability theory, Uspekhi Math. Nauk, 1948, v. 3, n. 3(25), pp. 187-194.

55. Gronau H.-D. O. F. and Proske Chr., Maximal families of restricted subsets of a finite set, Acta Cybernetica, Tom.5, Fasc.4, 1982.

56. Gua Xiaofeng, Lai Hong-Jian, Realizing degree sequences with k-edge-connected uniform hypergraphs, Discrete Mathematics 313 pp. 1394–1400, 2013.

57. Hansel G., Sur le nombre des fonctiones booleennes monotones de n variables, C. R. Acad. Sci., Paris 262, pp. 1080-1090, 1966.

58. Harper L. H., Global Methods for Combinatorial Isoperimetric Problems, Cambridge University Press, Cambridge, 2004.

59. Heuberger C., Inverse combinatorial optimization: a survey on problems, methods and results, *Journal of Combinatorial Optimization*, vol. 8, pp. 329–361, 2004.

60. Katona G., A theorem of finite sets, *Theory of Graphs*, Academy of Sciences, Budapest, pp. 187-207, 1968.

61. Katerinochkina N., On sets, containing maximal number of pair wise incomparable n-dimensional k-valued vectors, Mathematical notes, v. 24, no, 3, Sept. 1978.

62. Kocay William and Li Pak Ching, On 3-hypergraphs with equal degree sequences, Ars Combin. 82 pp. 145–157, 2007, MR 2292314 (2007k:05049).

63. Koren Michael, Pairs of sequences with a unique realization by bipartite graphs, J. Combin. Theory, B21, 3, pp. 224-234, 1976.

64. Koren Michael, Extreme degree sequences of simple graphs, J. Combinatorial Theory, Ser. B, 15:213–224, 1973.

65. Korn G., Korn T., Mathematical Handbook for Scientists and Engineers, McGraw-Hill company, 1961.

66. Korobkov V., On monotone functions of algebra of logic, Problemy Kybernetiki, 13, 1965.

67. Nigmatullin R. G., Variational principle in algebra of logic, Discrete Analysis, Novosibirsk 10 (1967) 69-89.

68. Paschos Vangelis Th. (Editor), Paradigms of Combinatorial Optimization: Problems and New Approaches, 2nd Edition, ISBN: 978-1-119-01519-2, August 2014, Wiley-ISTE.

69. Ryser H. J., Combinatorial Mathematics, 1963.

70. Ryser H. J., Combinatorial properties of matrices of zeros and ones, Canad. J. Math. 9 (1957) 371–377.

71. Sally C., Chris N. Potts, Barbara M. Smith, Constraint satisfaction problems: Algorithms and applications, Invited Review, European Journal of Operational Research 119 (1999) 557-581.

72. Tsang Ed, *Foundations of Constraint Satisfaction*, Academic Press, 1993.

73. Woeginger G. J., The reconstruction of polyominoes from their orthogonal projections, Inform. Process. Lett., 77:225-229, 2001.

74. Zhuravlev Yu., *Elected Research Works* (in Russian), Magister, Moscow, 1998.

75. Zhuravlev Yu., L. Aslanyan, and V. Ryazanov, Analysis of a training sample and classification in one recognition model, *Pattern Recognition and Image Analysis*, vol. 24, no. 3, pp. 347-352, 2014.

76. Тараканов В. Е., Комбинаторные задачи и (0,1)-матрицы, М., Наука, 1985.

77. Тышкевич Р. И., Характеризация (0,1)-матриц, определяемых числом единиц в строках и столбцах, и униграфических последовательностей, Доклады АН БССР, 1978, 22, 7, 592-595.

Authors' Information

Irina Arsenyan – Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia

Major Fields of Scientific Research: Boolean functions, Probabilistic Combinatorial Analysis.

Levon Aslanyan – Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia

Major Fields of Scientific Research: Discrete Mathematics, Pattern Recognition.

Hasmik Sahakyan – Institute for Informatics and Automation Problems of the National Academy of Sciences of Armenia

Major Fields of Scientific Research: Discrete Tomography, Combinatorial Analysis.