Statistics and Probability Letters 107 (2015) 356–361

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Asymptotic results for the number of Wagner's solutions to a generalised birthday problem

Alexey Lindo *, Serik Sagitov

Chalmers University of Technology and University of Gothenburg, SE-412 96 Gothenburg, Sweden

ARTICLE INFO

Article history:
Received 28 July 2015
Received in revised form 8 September 2015
Accepted 11 September 2015
Available online 25 September 2015

MSC:
60B20
60C05
60F05

Keywords:
Chen–Stein method
Functionals of random matrices

ABSTRACT

We study two functionals of a random matrix $\boldsymbol{A}$ with independent elements uniformly distributed over the cyclic group of integers $\{0, 1, \ldots, M-1\}$ modulo M. One of them, $V_0(\boldsymbol{A})$ with mean μ, gives the total number of solutions for a generalised birthday problem, and the other, $W(\boldsymbol{A})$ with mean λ, gives the number of solutions detected by Wagner's tree based algorithm.

We establish two limit theorems. Theorem 2.1 describes an asymptotical behaviour of the ratio λ/μ as $M \to \infty$. Theorem 2.2 gives bounds for the total variation distance between Poisson distribution and distributions of V_0 and W.

1. Introduction

Let (N, M, L) be three natural numbers larger than or equal to 2. Assume that we have a random matrix

$$\boldsymbol{A} = (a_{ij}), \quad 1 \le i \le L, \ 1 \le j \le N \tag{1}$$

with independent elements a_{ij} which are uniformly distributed on $\{0, 1, \ldots, M-1\}$. Let $\boldsymbol{J} = \{1, \ldots, L\}^N$ be the set of matrix positions, so that $|\boldsymbol{J}| = L^N$. For each $b \in \{0, 1, \ldots, M-1\}$, define $V_b \equiv V_b(\boldsymbol{A})$ as the number of vectors $\boldsymbol{i} = (i_1, \ldots, i_N) \in \boldsymbol{J}$ with

$$a_{i_1,1} + \cdots + a_{i_N,N} \stackrel{M}{=} b,$$

where the sign $\stackrel{M}{=}$ means equality modulo M. Clearly, $\sum_{b=0}^{M-1} V_b = L^N$, so that by the assumption of uniform distribution,

$$\mu := \mathrm{E}(V_0) = L^N M^{-1}. \tag{2}$$

The problem of finding all V_0 zero-sum vectors

$$\boldsymbol{a_i} = (a_{i_1,1}, \ldots, a_{i_N,N}), \quad \boldsymbol{i} = (i_1, \ldots, i_N) \in \boldsymbol{J} \tag{3}$$

for a given matrix $\boldsymbol{A}$, can be viewed as a generalised birthday problem. It arises naturally in a variety of situations including cryptography, see Wagner (2002) and references therein; ring linear codes (Greferath, 2009); abstract algebra, where in the theory of modules it is related to the notion of annihilators, see e.g. Lang (2002). This problem can be solved only by

* Corresponding author.
E-mail address: lindo@chalmers.se (A. Lindo).

http://dx.doi.org/10.1016/j.spl.2015.09.014
0167-7152/

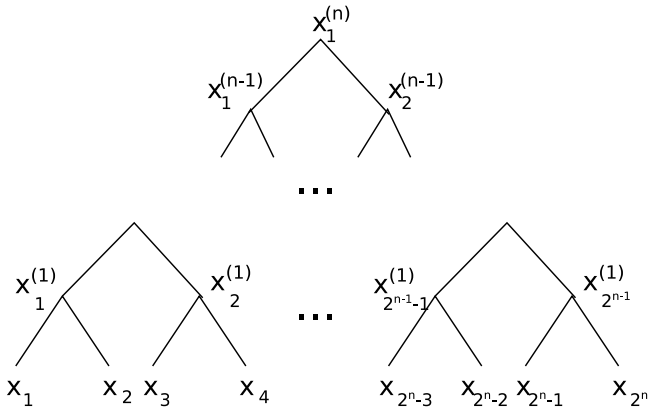

Fig. 1. Wagner's algorithm.

exhaustive search and is *NP*-hard (Schroeppel and Shamir, 1981). Wagner (2002) proposed a subexponential algorithm giving hope to quickly detect at least some of the solutions to these kinds of problems.

Assume that $N = 2^n$, $n \geq 1$ and $M = 2^m + 1$, $m \geq n$. It will be convenient to use the symmetric form

$$D_m := \{-2^{m-1}, \ldots, -1, 0, 1, \ldots, 2^{m-1}\}$$

of $\{0, 1, \ldots, M-1\}$ as the set of possible values for a_{ij}. Wagner's algorithm has a binary tree structure, see Fig. 1, starting from N leaves at level n and moving towards the top of the tree at level 0. For a given a vector $\boldsymbol{x} = (x_1, \ldots, x_{2^n})$ with $x_j \in D_m$ the algorithm searches for the value

$$H_n(\boldsymbol{x}) := x_1^{(n)} \in D_{m-n} \cup \{\Delta\}, \tag{4}$$

obtained recursively in a way explained next (the special state Δ indicates that the algorithm is terminated and a solution is not found). Put $x_j^{(0)} \equiv x_j$. For $h = 1, \ldots, n$ and $j = 1, \ldots, 2^{n-h}$, let $x_j^{(h)} = b$ if there exists such a $b \in D_{m-h}$ that

$$x_{2j-1}^{(h-1)} + x_{2j}^{(h-1)} \stackrel{M}{=} b,$$

and put $x_j^{(h)} = \Delta$ otherwise. In particular, if $x_k^{(h-1)} = \Delta$ for at least one of the two indices $k \in \{2j-1, 2j\}$, then $x_j^{(h)} = \Delta$.

A vector $\boldsymbol{x}$ will be called a Wagner's solution to the generalised birthday problem, if $H_n(\boldsymbol{x}) = 0$. The total number $W \equiv W(\boldsymbol{A})$ of Wagner's solutions among the vectors (3) has mean

$$\lambda := \mathrm{E}(W) = L^N p_{n,m},$$

where

$$p_{n,m} := \mathrm{P}(H_n(\boldsymbol{a_i}) = 0), \quad \boldsymbol{i} \in \boldsymbol{J}.$$

The proportion of Wagner's solutions can be characterised by the ratio of the means

$$R_{n,m} := \lambda/\mu = (2^m + 1)p_{n,m}, \tag{5}$$

where μ, defined by (2), is the mean total number of solutions. Clearly, $R_{n,m}$ is the conditional probability of a given zero-sum random vector to be Wagner's solution.

There is a growing number of papers studying the properties of various tree based algorithms with some of them, in particular (Minder and Sinclair, 2012), suggesting further developments of Wagner's approach. The main results of this paper are stated in the next section. Theorem 2.1 gives an integral recursion for calculating the limit for the key ratio (5). Theorem 2.2 gives an upper bound for the total variation distance between Poisson distribution $Po(\mu)$ and $\mathcal{L}(V_0)$, distribution of V_0, as well as a bound for the total variation distance between $Po(\lambda)$ and $\mathcal{L}(W)$. Recall that the total variation distance between the distributions of $\mathbb{Z}_+$-valued random variables X and Y, where $\mathbb{Z}_+ = \{0, 1, 2, \ldots\}$, is given by

$$d_{\mathrm{TV}}(\mathcal{L}(X), \mathcal{L}(Y)) = \sup_{A \subseteq \mathbb{Z}_+} |\mathrm{P}(X \in A) - \mathrm{P}(Y \in A)|.$$

Among related results concerning speed of convergence for functional of random matrices over finite algebraical structures we can only name a recent paper (Fulman and Goldstein, 2015).