A Generalization of Komlós Theorem on Random Matrices

Arkadii Slinko

March 29, 2000

Abstract: In this paper we prove that, for any subset $\Delta \subset \mathbb{Z}$, the probability, that a random $\Delta_{n\times n}$ matrix is singular, is of order $O\left(1/\sqrt{n}\right)$.

1. Introduction

Let $M_n = (a_{ij})$ be a random $n \times n$ $(0,1)$-matrix, where $0, 1 \in \mathbb{R}$, and every entry a_{ij} is chosen independently with probabilities $\mathrm{Prob}\,(a_{ij}{=}0) = \mathrm{Prob}\,(a_{ij}{=}1) = 1/2$. P. Erdös conjectured and J. Komlós proved [?], that such a matrix is almost always regular, i.e., the probability of M_n to be regular, approaches 1 as n approaches infinity. Komlós [?] sharpened his result by proving that

$$\mathrm{Prob}\ (M_n \text{ is regular}) > 1 - O\left(1/\sqrt{n}\right). \tag{1}$$

The key tool was the so-called Littlewood-Offord lemma, which was proved in full generality by Erdös [?]. Recently, Komlós theorem was further improved by J. Kahn, J. Komlós and E. Szemerédi [?]: they showed that

$$\mathrm{Prob}\ (M_n \text{ is regular}) > 1 - O\left((1-\epsilon)^n\right). \tag{2}$$

for some $\epsilon > 0$. The latter result was obtained by using some rather involved techniques from combinatorial number theory and Fourier analysis and combinatorics.

Generalising his result, Komlós also proved [?] that a random $n \times n$ matrix $M_n = (a_{ij})$, where every entry a_{ij} is chosen independently from an arbitrary nondegenerate distribution, is almost always regular; but no estimates of the speed of the convergence of this probability to 1 were obtained. For, to obtain any estimates of this kind for more complicated alphabets, than $\{0, 1\}$, we need to find a substitute for the Littlewood-Offord lemma. Here we suggest such a substitute and prove that the inequality (1) also holds for a random matrix $M_n = (a_{ij})$ with $a_{ij} \in \Delta$, when Δ is arbitrary. The analogue of (2) is yet to be proved.

It should be noted that the problem of estimating the probability of regularity of the random $(0, 1)$ matrices turned out to be closely related to questions in a surprising variety of areas, e.g. geometry of the unit cube [?], threshold logic [?], and associative memories [?]. The result of this paper might have similar applications in relation to geometry of grids and q-valued logic.

1. The Number of Δ-Solutions to a Linear Diophantine Equation.

In the sequel by Δ^n we denote the set of all vectors $(x_1, \ldots, x_n)^t$, $x_i \in \Delta$, and by $\Delta_{m\times n}$ we denote the set of all $m \times n$ matrices with entries from Δ. In what follows, $q = |\Delta|$ is the cardinality of Δ.

Theorem 1 *Let $N_{q,n}$ be the maximal number of Δ-solutions to the equation*

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{3}$$

where $a_i, b \in \mathbb{Z}$, $a_i \neq 0$, $i = 1, \ldots, n$. Then there is a constant $C > 0$ which does not depend on q or n, such that

$$N_{q,n} \leq C\frac{q^{n-2}}{\sqrt{n}}. \tag{4}$$

Proof: Let $I = \{i \mid a_i > 0\}$ and $J = \{j \mid a_j < 0\}$. Let us define the following partial order on vectors from Δ^n. For any two vectors $\mathbf{x}, \mathbf{y} \in \Delta^n$, we write $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for all $i \in I$ and $x_j \geq y_j$ for all $j \in J$. We also write $\mathbf{x} < \mathbf{y}$ if $\mathbf{x} \leq \mathbf{y}$ and $x_k \neq y_k$ for at least one k. An important observation is: if $\mathbf{x}, \mathbf{y} \in \Delta^n$ and $\mathbf{x} < \mathbf{y}$, then

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n < a_1y_1 + a_2y_2 + \cdots + a_ny_n,$$

whence $\mathbf{x}$ and $\mathbf{y}$ cannot be both solutions. Therefore the number of solutions of (3) is not greater than the length of a maximal antichain in $(\Delta^n, \leq)$.

Let us consider a multiset

$$A_{n,q} = \{\underbrace{1, \ldots, 1}_{q}, \underbrace{2, \ldots, 2}_{q}, \ldots, \underbrace{n, \ldots, n,}_{q}\} \tag{5}$$

with its subsets partially ordered by inclusion. Let $(\mathcal{P}(A_{n,q}), \subseteq)$ be the poset of all subsets of $A_{n,q}$. The posets $(\Delta^n, \leq)$ and $(\mathcal{P}(A_{n,q}), \subseteq)$ are isomorphic. Indeed, suppose $\Delta = \{\delta_1, \ldots, \delta_q\}$ and $\delta_1 < \delta_2 < \ldots < \delta_n$. Then

$$(\delta_{s_1}, \delta_{s_2}, \ldots, \delta_{s_n}) \to \{\underbrace{1, \ldots, 1}_{t_1}, \underbrace{2, \ldots, 2}_{t_2}, \ldots, \underbrace{n, \ldots, n,}_{t_n}\}, \tag{6}$$

where $t_i = s_i$ for $i \in I$ and $t_j = n - s_j$ for $j \in J$, is an isomorphism of these two posets. It is well-known that the collection of all subsets of $A_{n,q}$ of middle size, $\lfloor qn/2 \rfloor$, is a maximal antichain in $A_{n,q}$ [**?**, **?**]; if n and q are both odd, then the subsets of size $\lceil qn/2 \rceil$ also form a maximal antichain.

In [**?**] it was proved that the length $s(n)$ of a maximal antichain in $A_{n,q}$ satisfies the inequality

$$c\frac{q^{n-2}}{\sqrt{n}} \leq s(n) \leq C\frac{q^{n-2}}{\sqrt{n}},$$

for some constants c, C, not depending on q or n. This implies the result.

Comments: 1) It follows from [**?**] that

$$N_{q,n} \sim 3\sqrt{\frac{2}{\pi}}\frac{q^{n-2}}{\sqrt{n}};$$

2) For the alphabet $\Delta = \{0, 1, \ldots, q-1\}$, Illarionov [**?**] showed that

$$N_{q,n} < \frac{q^n}{\sqrt{n}}.$$

We will also need the following immediate corollary of this result.

Corollary 1 *Let $N_{q,n,s}$ be the maximal number of Δ-solutions to the equation*

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b, \tag{7}$$

where $a_i, b \in \mathbb{Z}$, $i = 1, \ldots, n$, and where exactly s of the a_i's are nonzero. Then

$$N_{q,n,s} \leq C \frac{q^{n-2}}{\sqrt{s}}. \tag{8}$$

Throughout this paper, C will denote the constant featuring in inequalities (4) and (10).

2. Determinants of Random Matrices

When speaking about random vectors from Δ^m, we assume that all vectors have the same probability q^{-m}. Similarly, when random matrices from $\Delta_{m\times n}$ are considered, it is assumed that all $m \times n$ matrices have equal probabilities q^{-mn}. A random $m \times n$ matrix can be viewed as consisting of m random rows or, equivalently, consisting of n random columns. The following result generalizes some results of Meshalkin [**?**] and Odlyzko [**?**].

Lemma 1 *Let F be a field and Δ be a finite subset of F. Let Δ^n be the set of all vectors of F^n with coordinates from Δ. Suppose that $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \Delta^n$. Then*

$$|\Delta^n \cap \mathit{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}| \leq q^k.$$

Proof: Since the result is trivial for $k = n$, we may assume that $k < n$. Without loss of generality we may assume that the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ form a linearly independent system of vectors, and that the first k rows of the matrix

$$X = (\mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_k),$$

composed of the vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$ as columns, are also linearly independent. Then $X = \begin{pmatrix} Y \\ Z \end{pmatrix}$, where Y is an invertible $k \times k$ matrix. Suppose now that $\mathbf{a} \in \Delta^n \cap \mathrm{Span}\,\{\mathbf{v}_1, \ldots, \mathbf{v}_k\}$. Let us write $\mathbf{a} = \begin{pmatrix} \mathbf{b} \\ \mathbf{c} \end{pmatrix}$, where $\mathbf{b} \in \Delta^k$. As

$$\mathbf{a} = u_1\mathbf{v}_1 + u_2\mathbf{v}_2 + \cdots + u_k\mathbf{v}_k = X\mathbf{u},$$

for some $\mathbf{u} = (u_1, \ldots, u_k)^t \in F^k$, we get $Y\mathbf{u} = \mathbf{b}$ and $Z\mathbf{u} = \mathbf{c}$. Since Y is invertible, we get $\mathbf{c} = ZY^{-1}\mathbf{b}$, and hence $\mathbf{c}$ is uniquely determined by $\mathbf{b}$. But there are only q^k possibilities to choose $\mathbf{b} \in \Delta^k$. □

Corollary 2 *Let $k \leq m$, and let $A \in \Delta_{m\times k}$ be a matrix of full rank, i.e., $r(A) = k$. If a vector $\mathbf{b} \in \Delta^m$ is taken at random, then for the augmented matrix $B = (A \mid \mathbf{b})$,*

$$\textit{Prob}\ (\, r\,(B){=}k\,) \leq q^{-m+k}.$$

Proof: Indeed, $\mathrm{r}\,(B) = k$ if and only if $\mathbf{b}$ is a linear combination of the columns of A. By Lemma 1 there are no more than q^k such linear combinations. Thus the required probability is not greater than $q^k/q^m = 1/q^{m-k}$. □

Lemma 2 *Let $k \leq m$, and let $A \in \Delta_{m\times k}$ be a matrix taken at random. Then*

$$\textit{Prob}\ (\, r\,(A){<}k\,) < 2q^{-m+k-1}.$$

Proof: We will form $A = (\mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_k)$ by taking subsequently k random columns $\mathbf{a}_1, \ldots, \mathbf{a}_k$. By Corollary 2, the probability that A has full rank, can be estimated as follows:

$$\mathrm{Prob}\ (\, \mathrm{r}\,(A){=}k\,) = \mathrm{Prob}\,(\mathbf{a}_1 \notin \{\mathbf{0}\}) \cdot \prod_{i=2}^{k} \mathrm{Prob}\,(\mathbf{a}_i \notin \mathrm{Span}\,\{\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}\}) \geq$$

$$\prod_{i=0}^{k-1}\left(1 - \frac{q^i}{q^m}\right) \geq 1 - \frac{1}{q^m}\sum_{i=0}^{k-1} q^i = 1 - \frac{q^k - 1}{q^m(q-1)} > 1 - 2q^{-m+k-1}. \square$$

Corollary 3 *Let $k, m \to \infty$ so that $k/m \to \gamma < 1$. Let A be a random $m \times k$ matrix. Then*

$$\textit{Prob}\,(\, r\,(A){<}k\,) = O\,(\delta^n)\,, \tag{9}$$

for some $0 < \delta < 1$. □

Let $S = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be a system of vectors. Let us define the *strong rank* of S, denoted $\mathrm{sr}\,(S)$, to be n if S is linearly independent; and k if any k of the $\mathbf{v}_i$'s are linearly independent but some $k{+}1$ of the vectors are linearly dependent. For a matrix A, by $\mathrm{sr}_c(A)$ and $\mathrm{sr}_r(A)$ we denote the strong rank of the system of columns and the strong rank of the system of rows of A, respectively.

Lemma 3 *(a) Let A be a random $m \times n$ matrix from $\Delta_{m\times n}$. Then*

$$Prob\,(\,sr_c(A)<k\,) \le \binom{n}{k} q^{-m+k}; \tag{10}$$

(b) Let α, β be any positive real numbers satisfying

$$\frac{h(\beta)}{\log_2 q} + \beta < \alpha < 1, \tag{11}$$

where $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$ is the entropy function. Let $k, m, n \to \infty$ with $m/n \to \alpha$, $k/n \to \beta$. Then for a random $m \times n$ matrix $A \in \Delta_{m\times n}$ there exists $0 < \delta < 1$ such that

$$Prob\,(\,sr_c(A)<k\,) = O\left(\delta^n\right). \tag{12}$$

(c) For every $q \ge 2$ and $0 < \alpha \le 1$ there exists $\beta > 0$ which satisfies (13).

Proof: (a) Follows immediately from Lemma 2. (b) follows from (a). Indeed, since $\binom{n}{\beta n} < 2^{nh(\beta)}$ (see, for example, [**?**]), then

$$\mathrm{Prob}\,(\,\mathrm{sr}_c(A)<\beta n\,) \lesssim \binom{n}{\beta n} q^{-(\alpha-\beta)n} < 2^{n[h(\beta)-(\alpha-\beta)\log_2 q]} = O\left(\gamma^n\right),$$

where $\gamma = 2^{h(\beta)-(\alpha-\beta)\log_2 q} < 1$ as a consequence of (13). (c) Let

$$g(x) = \frac{h(x)}{\log_2 q} + x.$$

Since this function is continuous and $g(0) = 0$, a positive number $\beta > 0$ exists which satisfies (13). $\Box$

Lemma 4 *Let $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \Delta^m$ be linearly independent vectors, and let $B = (\mathbf{v}_1 \mid \cdots \mid \mathbf{v}_k)$, and $sr_r(B) = s$. Then for a random vector $\mathbf{a} \in \Delta^m$*

$$Prob\,(\,r\,(\mathbf{v}_1, \ldots, \mathbf{v}_k, \mathbf{a}) = k\,) < C/q^{m-k-1}\sqrt{s}. \tag{13}$$

Proof: Let $B = (\mathbf{b}_1 \mid \cdots \mid \mathbf{b}_m)^t$, i.e., $\mathbf{b}_1^t, \ldots, \mathbf{b}_m^t$ are the rows of B. Without loss of generality we assume that $\mathbf{b}_1, \ldots, \mathbf{b}_k$ are linearly independent and that all other $\mathbf{b}_i$'s for $i > k$ are linearly dependent on these. In particular, we have

$$\sum_{i=1}^{k+1} \beta_i \mathbf{b}_i = 0 \tag{14}$$

for some $\beta_1, \ldots, \beta_k$ and $\beta_{k+1} = 1$. As $\mathrm{sr}_r(B) = s$, at least s of the coefficients $\beta_1, \ldots, \beta_k$ are nonzero.

Let $\mathbf{a} = (a_1, \ldots, a_m)^t$. In the event of $\mathbf{a}$ being dependent on $\mathbf{v}_1, \ldots, \mathbf{v}_k$, due to (15), we must have

$$\sum_{i=1}^{k+1} \beta_i a_i = 0, \tag{15}$$

hence by Corollary 1 $a_1, \ldots, a_{k+1}$ can be selected in at most $Cq^{k-1}/\sqrt{s}$ ways. Once $a_1, \ldots, a_k$ are determined, all coefficients $a_{k+2}, \ldots, a_m$ must be determined since the rows $\mathbf{b}_{k+2}, \ldots, \mathbf{b}_m$ are also dependent on the rows $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Thus the probability to choose $\mathbf{a}$ so that it linearly depends on $\mathbf{v}_1, \ldots, \mathbf{v}_k$ is at most $(Cq^{k+1}/\sqrt{s})/q^m = C/q^{m-k-1}\sqrt{s}$. $\Box$

Theorem 2 *Let A be a random $n \times n$ matrix from $\Delta_{n\times n}$. Then as $n \to \infty$*

$$\mathit{Prob}\,(\, r(A) < n\,) = O\left(1/\sqrt{n}\right).$$

Proof: Let α and β be any real numbers satisfying (13), - they exist due to Lemma 3(c). Let us view the first $n_0 = [\alpha n]$ columns of A as a separate random matrix B. Since the columns of A are statistically independent, we may assume that we fill B randomly first and then we fill the remaining $n - n_0$ columns of A. By Lemmata 2 and 3 we may assume that B has full rank and that $\mathrm{sr}_r(B) \geq \beta n$.

Let us denote by A_k, $k \geq n_0$, the random matrix obtained by filling the first k columns of A. Then by Lemma 4

$$\mathrm{Prob}\,(\,\mathrm{r}\,(A_k) = k) \geq \mathrm{Prob}\,(\,\mathrm{r}\,(A_{k-1}) = k-1) \times \left(1 - \frac{C}{q^{n-k}\sqrt{\beta n}}\right).$$

Since $A_{n_0} = B$ has full rank and $A = A_{n,q}$, we get that

$$\mathrm{Prob}\,(\,\mathrm{r}\,(A) = n) \geq \prod_{i=1}^{n-n_0} \left(1 - \frac{C}{q^i\sqrt{\beta n}}\right) \geq 1 - \frac{C}{\sqrt{\beta}(q-1)\sqrt{n}}.$$

The theorem is proved. □

3. Acknowledgements.

I am grateful to Marston Conder and Paul Hafner who kindly provided me with the references [1–3] when I was travelling overseas.

References

[1] I. Anderson, "Combinatorics of Finite Sets," Clarendon Press, Oxford, 1987.

[2] I. Anderson, A Variance Method in Combinatorial Number Theory, *Glasgow Math. J.* **10** (1969), 126–129.

[3] N.G. de Bruijn, C. van E. Tengbergen and D. Kruyswijk, On the set of divisors of a number, *Nieuw Arch. Wiskunde* (2) **23** (1951), 191–193.

[4] P. Erdös, On a lemma of Littlewood and Offord, *Bull. Amer. Math. Soc.* **51** (1945), 898–902.

[5] Z. Füredi, Random Polytopes in the d-Dimensional Cube, *Discrete Comput. Geom.* **1** (1986), 315–319.

[6] V.V. Illarionov, On estimates for the number of solutions of a certain equation in integers (Russian), *Mat. Zametki* **54** (1993), 57–64.

[7] J. Kahn, J. Komlós and E. Szemerédi, On the probability that a random ± 1 matrix is singular, *J. Amer. Math. Soc.* **8** (1995), 223–240.

[8] I. Kanter and H. Sompolinsky, Associative Recall of Memory Without Errors, *Phys. Rev.* (A) (3) **35** (1987), 380–392.

[9] J. Komlós, On the determinants of $(0,1)$ matrices, *Studia Sci. Math. Hungar.* **2** (1967), 7–21.

[10] J. Komlós, manuscript (1977). In: B. Bollobás (Ed.) "Random Graphs," Academic Press, New York/London, 1985, 347-350.

[11] J. Komlós, On the determinants of random matrices, *Studia Sci. Math. Hungar.* **3** (1968), 387–399.

[12] L.D. Meshalkin, To Substantiation of the Random Balance Method (Russian), *Zavodskaja Laboratorija* v. XXXVI, No 3 (1970), 316–318.

[13] W.W. Peterson and E.J. Weldon, “Error-Correcting Codes,” 2nd ed., Cambridge, MIT Press, 1972.

[14] A.M. Odlyzko, On Subspaces Spanned by Random Selections of ± 1 Vectors, *J. Combin. Theory Ser.* A **47** (1988), 124–133.

[15] Ju.A. Zuev, Methods of Geometry and Probabilistic Combinatorics in Threshold Logic, *Discrete Math. Appl.* **2** (1992), 427–438.