BALANCING GAUSSIAN VECTORS

KEVIN P. COSTELLO

Abstract. Let $x_1, \dots x_n$ be independent normally distributed vectors on R^d. We determine the distribution function of the minimum norm of the 2^n vectors $\pm x_1 \pm x_2 \cdots \pm x_n$.

1. Introduction

Let d be a fixed positive integer and $x_1, \dots, x_n$ be vectors in R^d of lengths at most 1. Bárány and Grinberg ([2], see also [1] for a generalized version), extending an earlier result of Spencer [5], proved that one can find a sign sequence $\epsilon_1, \dots, \epsilon_n$ such that

$$\|\epsilon_1 x_1 + \cdots + \epsilon_n x_n\| \leq 2d,$$

In this paper, we investigate the situation when $x_1, \dots, x_n$ are i.i.d Gaussian vectors (each coordinate having mean 0 and variance 1), motivated by a question of Vershynin. For a sign sequence $S = (\epsilon_1, \dots, \epsilon_n) \in \{-1, 1\}^n$, define

$$X_S := \sum_{i=1}^{n} \epsilon_i x_i.$$

We would like to study

$$Y := \min_{S \in \{-1,1\}^n} ||X_S||,$$

where $||.||$ is any norm on $\mathbf{R}^d$.

Using a standard large deviation argument, it is easy to see that most of the X_S have all coordinates of order $O(\sqrt{n})$. Thus, one expects that Y is typically of order $O(\frac{\sqrt{n}}{2^{n/d}})$, this being the lattice size in $\mathbf{R}^d$ so the ball of radius $\sqrt{n}$ contains 2^n lattice points (Here and later the asymptotic notation is used under the assumption that n tends to infinity and d remains fixed). Our first result confirms this intuition and provides a precise description of the distribution of Y.

Theorem 1.1. *For every fixed $\lambda > 0$,*

$$\mathbf{P}(Y > \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda) = (1+o(1))e^{-\lambda^d},$$

where V_d is the volume of the unit ball in $||.||$.

Remark 1.2. The constant implicit in the $o(1)$ notation here is dependent both on the specific norm used and on λ. In particular, we claim only weak convergence here.

This theorem is a special case of a more general theorem, whose presentation would be more natural with the following discussion. Let $X_S(i)$ denote the i^{th} coordinate of X_S. For any particular sequence S and any i, $X_S(i)$ has normal distribution with mean 0 and variance n. Therefore for each $a_i, b_i = o(n^{1/2})$ we have

$$\mathbf{P}(|X_S(i)| \in [a_i, b_i]) = (1+o(1))(b_i - a_i)\sqrt{\frac{1}{2\pi n}}. \qquad (1)$$

Taking products and integrating over the $||.||$ ball of radius ϵ, we obtain (so long as $\epsilon = o(n^{1/2})$):

$$\mathbf{P}(||X_S|| < \epsilon) = (1+o(1))V_d \epsilon^d (2\pi n)^{-d/2}.$$

Substituting $\epsilon = \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda$ yields,

$$\mathbf{P}(||X_S|| \leq \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda) = (1+o(1))\frac{\lambda^d}{2^{n-1}}.$$

Since $||X_S|| = ||X_{-S}||$ for any S, we will consider only those S whose first element is one. There are 2^{n-1} such sequences, so the expected number of such sequences with $||X_S|| \leq \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda$ is

$$(1+o(1))\frac{\lambda^d}{2^{n-1}}2^{n-1} = (1+o(1))\lambda^d.$$

Let $I_S(\lambda)$ be the indicator variable of the event that $||X_S|| \leq \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda$ and set $\Sigma_n(\lambda) := \sum_S I_S(\lambda)$. We have

$$\mathbf{P}(Y > \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda) = \mathbf{P}(\Sigma_n(\lambda) = 0).$$

The above argument shows that the expectation of $\Sigma_n(\lambda)$ is $(1+o(1)\lambda^d$, as n tends to infinity. Our main result shows that the distribution of $\Sigma_n(\lambda)$ tends to $Poi(\lambda^d)$, the Poisson distribution with mean λ^d.

Theorem 1.3. *The distribution of $\Sigma_n(\lambda)$ tends to $Poi(\lambda^d)$ as n tends to infinity. In particular, for any fixed non-negative integer l and n tending to infinity*

$$\mathbf{P}(\Sigma_n(\lambda^d) = l) = (1+o(1))e^{-\lambda^d}\frac{\lambda^{dl}}{l!}.$$

Remark 1.4. Again, the $o(1)$ depends on λ, l, and the norm in question, though for each fixed λ and norm it tends to 0 for large n.

Theorem 1.1 is the case when $l = 0$.

2. Proof of Theorem 1.3

Our proof is combinatorial in nature, so it will be more convenient to work with sets than with sequences. Given a sequence $\epsilon_1, \dots, \epsilon_n$, the sum $\sum_{i=1}^n \epsilon_i x_i$ can be written as

$$\sum_{i\in S} x_i - \sum_{i\in \bar{S}} x_i,$$

where $S := \{i | 1 \le i \le n, \epsilon_i = 1\}$ and $\bar{S} := \{1, \dots, n\}\backslash S$. We abuse notation slightly and denote, for each subset S of $\{1, \dots, n\}$

$$X_S := \sum_{i\in S} x_i - \sum_{i\in \bar{S}} x_i.$$

We will be working only with those S containing 1. (This corresponds to the fact that we fix $\epsilon_1 = 1$). Let Ω be the collection of these sets. For an integer $1 \le k \le n$, we denote by Ω_k the collection of tuples of k distinct elements of Ω. (A common notation in combinatorics is $\Omega_k = \binom{\Omega}{k}$).

In order to prove Theorem 1.3, it suffices to show that for any fixed integer $k \ge 1$, the kth moment of $\Sigma_n(\lambda)$ converges to that of $Poi(\lambda^d)$, as n tends to infinity. This, via a routine calculation, is a direct consequence of the following

Proposition 2.1. *For any fixed integer $k \ge 1$,*

$$\mathbf{E}\Big(\sum_{(S_1,\dots,S_k)\in\Omega_k} I_{S_1}(\lambda)\dots I_{S_k}(\lambda)\Big) = (1+o(1))\frac{\lambda^{dk}}{k!}.$$

In the rest of this section, we are going to present four lemmas and then use these lemmas to conclude the proof of Proposition 2.1. The proofs of the lemmas follow in the later sections.

We are going to divide Ω_k into two parts based on the size of various intersections involving the sets in each tuple. The division will depend on a parameter $\delta > 0$, which one should think of as a very small positive constant. The following definition plays a critical role in the proof.

Definition 2.2. A $k-$tuple $(S_1, \dots S_k)$ is δ-**balanced** if every

$$(T_1, \dots T_k) \in \{S_1, \bar{S}_1\} \times \{S_2, \bar{S}_2\} \times \cdots \times \{S_k, \bar{S}_k\}$$

satisfies

$$(1-\delta)\frac{n}{2^k} \leq |T_1 \cap T_2 \cap \cdots \cap T_k| \leq (1+\delta)\frac{n}{2^k}.$$

Notation. Since λ is now fixed, we will write I_S instead of $I_S(\lambda)$ in the rest of the proof. Similarly, instead of saying a set is δ-balanced, we simply say that it is balanced. Set

$$\epsilon := \frac{\sqrt{2\pi n}}{(V_d 2^{n-1})^{1/d}}\lambda.$$

Our first lemma states that most k-tuples in Ω_k are $\delta-$ balanced.

Lemma 2.3. *The unbalanced k-tuples form an exponentially small proportion of Ω_k. To be more precise, the number of unbalanced k-tuples in Ω_k is at most $(2^k)(2e^{-\frac{2\delta^2 n}{2^k}})\binom{2^n}{k}$.*

The second lemma shows that if a tuple $(S_1, \dots, S_k)$ is $\delta-$balanced, then $\mathbf{E}(\prod_{i=1}^k I_{S_i})$ is close to $\prod_{i=1}^k \mathbf{E}(I_{S_i})$ and can be estimated almost precisely.

Lemma 2.4. *If $(S_1, S_2, \dots S_k)$ is a balanced $k-$tuple, then*

$$(1+o(1))(1+\delta)^{-dk}(\frac{\lambda^d}{2^{n-1}})^k \leq \mathbf{E}(\prod_{i=1}^k I_{S_i}) \leq (1+o(1))(1-\delta)^{-dk}(\frac{\lambda^d}{2^{n-1}})^k. \quad (2)$$

The first two lemmas together will imply that the contribution to the expectation from the balanced $k-$tuples is roughly the correct amount. Our next two lemmas together will imply that the contribution from the unbalanced $k-$tuples is negligible.

For each set $S \subset \{1, \dots, n\}$ assign the characteristic vector v_S as follows: v_S has length n and its ith coordinate is one if $i \in S$ and minus one otherwise.

Lemma 2.5. *Let r be a fixed positive integer and let $(S_1, S_2 \dots S_r)$ be a r-tuple in Ω_r with linearly independent characteristic vectors. Then*

$$\mathbf{E}(\prod_{i=1}^{r} I_{S_i}) = O(\epsilon^{dr}),$$

where the constant implicit in the O notation is dependent on r and the norm in question.

Lemma 2.6. *For any $j < k$, the number of $k-$tuples in Ω_k whose characteristic vectors span a j-dimensional space is $O(\left(\frac{3}{2^{k-j+2}}\right)^n \binom{2^{n-1}}{k})$.*

To conclude this section, we prove Proposition 2.1, assuming the above four lemmas.

We split the expectation in question into two parts: The contribution from balanced $k-$tuples and that from unbalanced $k-$tuples. By Lemma 2.3 the number of balanced $k-$tuples is

$$(1 + o(1))\binom{2^{n-1}}{k} = (1 + o(1))\frac{2^{(n-1)k}}{k!}.$$

Multiplying this bound by the contribution of each $k-$tuple (as bounded by Lemma 2.4), we conclude that the contribution from balanced tuples is between $(1 + o(1))(1+\delta)^{-dk}\frac{\lambda^{dk}}{k!}$ and $(1 + o(1))(1-\delta)^{-dk}\frac{\lambda^{dk}}{k!}$.

We next examine the contribution of those unbalanced $k-$tuples. We split this contribution further according to the dimension of the subspace spanned by the characteristic vectors of the sets in the tuple.

- Sets for which the subspace has dimension less than k. If the dimension is j, where $1 \le j < k$, then by Lemma 2.6 there are $O(\left(\frac{3}{2^{k-j+2}}\right)^n 2^{kn})$ such tuples. By Lemma 2.5, each such tuple contributes $O(\epsilon^{dj})$ towards the expectation. Therefore, the total contribution from this class is

$$\sum_{j=1}^{k-1} O(\left(\frac{3}{2^{k-j+2}}\right)^n 2^{kn}\epsilon^{dj}) = \sum_{j=1}^{k-1} O(\left(\frac{3}{4}\right)^n (2^n\epsilon^d)^j) = O((\frac{3}{4})^n n^{d(k-1)/2}) = o(1).$$

- Sets for which the subspace has dimension k. By Lemma 2.3, there is a $c_1 < 1$ such that the number of tuples in question is $O(c_1^n 2^{kn})$. By Lemma 2.5, the contribution of each such tuple is $O(\epsilon^{dk})$. Therefore the total contribution from this class is

$$O(c_1^n 2^{kn}\epsilon^{dk}) = O(c_1^n n^{k/2}) = o(1).$$

It follows that the contribution from unbalanced tuples is $o(1)$. Thus, the expectation in question is between $(1+o(1))(1+\delta)^{-dk}\frac{\lambda^{dk}}{k!}$ and $(1+o(1))(1-\delta)^{-dk}\frac{\lambda^{dk}}{k!}$, for any fixed δ and sufficiently large n. Letting δ tend to 0 concludes the proof.

3. Proof of Lemma 2.3

We can restate the lemma probabilistically as follows:

Let $(S_1, S_2 \dots S_k)$ be chosen uniformly at random from Ω_k and let p_1 be the probability that $(S_1, \dots, S_k)$ is not balanced. Then $p_1 \leq (2^k)(2e^{-\frac{2\delta^2 n}{2^k}})$.

Consider another experiment where we choose k sets $T_1, \dots, T_k$ uniformly from Ω and let p_2 be the probability that $(T_1, \dots, T_k)$ is balanced. (The main difference here is that T_i and T_j maybe the same).

Since any $k-$tuple with two equal elements is unbalanced, $p_2 \leq p_1$, so it suffices to show that p_2 is large. To do so, we first note that in the second experiment, for each $(T_1, T_2, \dots T_k)$ and each coordinate except the first one, the probability that the coordinate has the desired sequence of signs is 2^{-k} (the first coordinate will always be equal to 1 by assumption). Thus we can express the number of coordinates having the desired sign sequence as

$$\chi(1 \in \bigcap_{i=1}^{n} T_i) + \sum_{i=2}^{n} y_i,$$

where the y_i are i.i.d random variables taking value one with probability 2^{-k} and zero with probability $1 - 2^{-k}$ and $\chi(1 \in \bigcap_{i=1}^{n} T_i)$ equal to one if 1 is contained in the intersection of the T_i and 0 otherwise.

The second term in this expression is a sum of independent binomial variables, and thus by Chernoff's bound ([3]) satisfies

$$\mathbf{P}(|\sum_{i=2}^{n} y_i - \mathbf{E}(\sum_{i=2}^{n} y_i)| > \beta) < 2e^{\frac{-2\beta^2}{n-1}}$$

In this case the expected sum of the y_i is either $2^{-k}n$ or $2^{-k}n+1$. Taking $\beta = \frac{\delta n}{2^k}$ in the Chernoff bound gives that the probability any particular sign sequence fails the criterion for balancedness is at most $2e^{-\frac{2\delta^2 n}{2^k}}$. By the union bound, the probability that at least one of the 2^k sign sequences fails the criterion for balancedness is at most $2^k(2e^{-\frac{2\delta^2 n}{2^k}})$.

4. Proof of Lemma 2.4

We first prove the following lemma, which is a slightly stronger version of Lemma 2.4 for the case $d = 1$.

Lemma 4.1. *Let $(S_1, \dots S_k)$ be a balanced $k-$tuple. Let $a_1 < b_1, a_2 < b_2, \dots a_k < b_k$ be any positive constants such that $|a_j| + |b_j| = o(n^{-1/2})$ for all j, and let A denote the event that X_{S_j} is contained in $[a_j, b_j]$ for all j. Then*

$$(1+o(1))\frac{\prod_{j=1}^{k}(b_j - a_j)}{(2\pi n)^{k/2}(1+\delta)^k} \le \mathbf{P}(A) \le (1+o(1))\frac{\prod_{j=1}^{k}(b_j - a_j)}{(2\pi n)^{k/2}(1-\delta)^k}, \tag{3}$$

where the constant implicit in the o notation depends only on n and not on the specific $k-$tuple, a_j, or b_j.

Proof

For each vector $v = (v_1, v_2, \dots v_k) \in \{-1, 1\}^k$, let $S_v = T_1 \cap T_2 \dots \cap T_n$, where T_i is S_i if $v_i = 1$, $\bar{S}_i$ otherwise. Let $y_v = \sum_{i \in S_v} x_i$. Each y_v is independently normally distributed with mean 0 and some variance σ_v^2. Let $c_v = 2^k \sigma_v^2 / n$. If all of the S_v were of the same size, all the c_v would be 1; what the balancedness of our $k-$tuple guarantees is that

$$(1-\delta) \le c_v \le (1+\delta).$$

We will estimate the probability that A occurs by first recasting A in terms of the S_v, then rescaling the S_v into new normal variables S'_v which have equal variance, making their joint distribution considerably easier to work with.

In terms of the y_v, the event $\{S_j \in [a_j, b_j] \forall j\}$ transforms to the k simultaneous inequalities

$$a_j \le \sum_{v_j = 1} y_v - \sum_{v_j = -1} y_v \le b_j, \tag{4}$$

where j runs from 1 to k and each summation is taken over all v having the required value of v_j.

The system (4) defines a $(2^k - k)$-dimensional slab in the 2^k-dimensional coordinate system whose axes are the y_v. Orthogonal cross-sections to this slab are k-dimensional boxes whose edges are parallel to vectors whose y_v^{th} coordinate is v_j for some j. Each such side has length $\frac{b_j - a_j}{\sqrt{2^k}}$, so each orthogonal cross section has area $\frac{\prod_{j=1}^{k}(b_j - a_j)}{2^{k^2/2}}$ in this coordinate system.

We now apply the transformation $y'_v = y_v / c_v$ to each coordinate. Since each coordinate is divided by a factor between $(1-\delta)$ and $(1+\delta)$, the volume of any $k-$dimensional object in the y'_v coordinate system is between $(1+\delta)^{-k}$ and $(1-\delta)^{-k}$ times what its was in the y_v system. In particular, this applies to the cross sectional area of the slab of solutions to (4) (as this area is the minimum over all $k-$dimensional subspaces W of the volume of the intersection of W with the transformed slab). Thus the area of an orthogonal cross section to the solution (4) lies between

$$(1+\delta)^{-k}\frac{\prod_{j=1}^k (b_j - a_j)}{2^{k^2/2}}$$

and

$$(1-\delta)^{-k}\frac{\prod_{j=1}^k (b_j - a_j)}{2^{k^2/2}}.$$

Since the y'_v are all normal with equal standard deviation, their joint distribution is rotationally invariant. This implies that the probability the y'_v satisfy (4) is the same as it would be if the slab of solutions were rotated so its orthogonal cross sections lay parallel to $\{y'_{v_1}, \dots y'_{v_k}\}$ for some $v_1, \dots v_k$.

Since the a_j and b_j are sufficiently close to zero, the joint density function of $\{y'_1, \dots y'_k\}$ inside the rotated slab is $(1+o(1))(\sqrt{2^k/2n\pi})^k$. Combining this with the our bounds on the cross sectional area of the solution set, we obtain our desired bound. ■

Application of this lemma to each coordinate in turn along with the independence of the coordinates yields that for any a_{ij} and b_{ij} such that $|a_{ij}| + |b_{ij}| = o(n^{-1/2})$ we have:

$$\begin{aligned}(1+o(1))\frac{\prod_{i=1}^d \prod_{j=1}^k (b_{ij} - a_{ij})}{(2\pi n)^{kd/2}(1+\delta)^{kd}} &\leq \mathbf{P}(\bigwedge_{i=1}^d \bigwedge_{j=1}^k \{S_j(i) \in [a_{ij}, b_{ij}]\}) \\ &\leq \frac{\prod_{i=1}^d \prod_{j=1}^k (b_{ij} - a_{ij})}{(2\pi n)^{kd/2}(1+\delta)^{kd}}.\end{aligned}$$

Integration over the product of balls of $||.||$ radius ϵ in $\mathbf{R}^d \times \mathbf{R}^d \times \dots \times \mathbf{R}^d$ now yields

$$\begin{aligned}(1+o(1))\frac{\epsilon^{dk} V_d^k}{(2\pi n)^{kd/2}(1+\delta)^{kd}} &\leq \mathbf{P}(\wedge_{j=1}^k \{||S_j|| < \epsilon\}) \\ &\leq \frac{\epsilon^{dk} V_d^k}{(2\pi n)^{kd/2}(1-\delta)^{kd}}\end{aligned}$$

which simplifies to the desired bounds.

5. Proof of Lemma 2.5

Since all norms on $\mathbf{R}^d$ are equivalent, there is a fixed C (dependent on our choice of norm) such that for any epsilon the ball of radius ϵ in our norm is contained in the ball of radius $C\epsilon$ in the l^∞ norm. Therefore it suffices to prove that

$$\mathbf{P}(\wedge_{j=1}^{r}\{||S_j||_\infty < \epsilon\}) = O(\epsilon^{dr}),$$

and by symmetry and the independence of coordinates it is thus sufficient to prove

$$\mathbf{P}(\wedge_{j=1}^{r}\{|S_j(1)| < \epsilon\}) = O(\epsilon^{r}).$$

Let A be the matrix formed by the ± 1 characteristic vectors of the S_i. Without loss of generality we can assume that A has the form

$$A = \left[\begin{array}{c|c} A_1 & A_2 \end{array}\right],$$

where A_1 is a nonsingular $r \times r$ matrix. Since

$$\mathbf{P}(\wedge_{j=1}^{r}\{|S_j(1)| < \epsilon\}) \leq \sup_{x_{r+1}\dots x_n} \mathbf{P}(\wedge_{j=1}^{r}\{|S_j(1)| < \epsilon\}|x_{r+1},\dots,x_n),$$

it suffices to prove the bound conditioned on $x_{r+1},\dots x_n$, treating the remaining r variables as random. For any fixed values of the last $n-r$ variables, the matrix equation becomes

$$A_1 \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_r \end{pmatrix} \in H \tag{5}$$

where H is some r dimensional cube of side length 2ϵ in $\mathbf{R}^r$. Since A_1 has determinant at least 1 (it is a nonsingular matrix with integer entries), it follows that the $r-$dimensional volume of the set of $x_1,\dots x_r$ for which (5) holds is at most $(2\epsilon)^r$. Since the density function of the joint distribution function of $(x_1,\dots x_r)$ is bounded from above by 1, the claim follows.

6. Proof of Lemma 2.6

As in the proof of Lemma 2.3, we begin by wording the lemma probabilistically as follows:

Let $(S_1, S_2, \dots S_k)$ be chosen uniformly from Ω_k, and let w_i be the characteristic vector of S_i. Then the probability the w_i span a $k-j$ dimensional space is $O(\left(\frac{3}{2^{j+2}}\right)^n)$, for any $j < k$.

Lemma 6.1. *The probability that w_{k-j+1} lies in the span of $w_1, w_2, \dots w_{k-j}$ is $O(\left(\frac{3}{8}\right)^n)$, where the constant in the O depends only on $k-j$.*

With Lemma 6.1, we can finish the proof as follows. With the cost of a factor at most $\binom{k}{j} = O(1)$ we can assume that $w_1, \dots, w_{k-j}$ are independent. The probability that $w_1, \dots, w_k$ span a space of dimension $k-j$ is then at most

$$O(1)\mathbf{P}(w_{k-j+1} \in \text{ Span } \{w_1, \dots, w_{k-j})$$
$$\times \prod_{l=k-j+1}^{k} \mathbf{P}\Big(w_{l+1} \in \text{ Span } \{w_1, \dots, w_{k-j}\} | w_{k-j+1}, \dots, w_l \in \text{ Span } \{w_1, \dots, w_{k-j}\}\Big).$$

By Lemma 6.1, $\mathbf{P}(w_{k-j+1} \in \text{ Span } \{w_1, \dots, w_{k-j}\}) = O((3/8)^n)$. Furthermore, any subspace of dimension d contains at most 2^d vectors in $\{-1,1\}^n$ (as some subset of d coordinates uniquely determines each vector in the subspace), so

$$\mathbf{P}\Big(w_{l+1} \in \text{ Span } \{w_1, \dots, w_{k-j}\} | w_{k-j+1}, \dots, w_l \in \text{ Span } \{w_1, \dots, w_{k-j}\}\Big) \le 2^{k-j}/2^{n-1} = O(1/2^n).$$

The bound follows immediately by multiplying the bound for each l.

To conclude, we present the proof of Lemma 6.1.

Set $q = k-j+1$ and consider $w_1, \dots, w_q$ as the row vectors of a q by n matrix, whose entries are $A_{i,m}$, $1 \le i \le q, 1 \le m \le n$. Notice that $A_{i,m}$ are either one or minus one. For the rows to be dependent, there must be numbers $a_1, \dots, a_q$ (not all zero) such that for every column $1 \le m \le n$ the entries $A_{i,m}$ in that column satisfy

$$\sum_{i=1}^{q} a_i A_{i,m} = 0 \tag{6}$$

Since no row is a zero vector, at least two of the coefficients a_i must be non-zero. Furthermore, as the sets S_i are different elements of Ω, no two rows are equal or opposite of each other, so one cannot have exactly 2 non-zero coefficients.

Now we show that one cannot have exactly three non-zero coefficients, either. Assume, for a contradiction, that a_1, a_2, a_3 are the only non-zero coefficients. Since the first and second rows are not equal and not opposite of each other, we can assume, without loss of generality, that the first two elements in the first column are $1, 1$ and the first two elements in the second column are $1, -1$. Now look at the first two elements of the third row. If they are $1, 1$, we end up with two equalities

$$a_1 + a_2 + a_3 = 0 \text{ and } a_1 - a_2 + a_3 = 0,$$

from which we can deduce $a_2 = 0$, a contradiction. The remaining three cases lead to similar contradictions.

Thus, we conclude that there are at least 4 non-zero coefficients. We are going to need the following solution of Erdös[4], to the so-called Littlewood-Offord problem.

Lemma 6.2. *Let $a_1, \dots, a_q$ be real numbers with at least k non-zero. Then the number of vectors $v = (v_1, \dots, v_q) \in \{-1, 1\}^q$ satisfying $\sum_{i=1}^n a_i v_i = 0$ is at most $\frac{\binom{k}{\lfloor k/2 \rfloor}}{2^k} 2^q$.*

Applying Lemma 6.2 with $k = 4$, we conclude that there are at most $\left(\frac{3}{8}\right) 2^q$ distinct columns satisfying (6), given any set $a_1, \dots, a_q$ with at least 4 non-zero elements. It follows that the probability that the rows are dependent is bounded from above by the probability of the event B_q that there are at most $\left(\frac{3}{8}\right) 2^q$ distinct columns in the matrix.

Now let $w_1', w_2', \dots w_q'$ be vectors chosen independently with respect to the uniform distribution over the set of all $(-1, 1)$ vectors of length n with first coordinate 1. Let B_q' be the event that the q by n matrix formed by the w_i' has at most $\frac{3}{8} 2^q$ distinct columns. Clearly, $\mathbf{P}(B_q') \geq \mathbf{P}(B_q)$, as some of the w_i' can be the same. Therefore, it suffices to show $\mathbf{P}(B_q') = O(\left(\frac{3}{8}\right)^n)$. The matrix formed by the w_i' has the first column equal the all one vector. The remaining $n - 1$ columns are i.i.d. random vectors from $\{-1, 1\}^q$.

Let $\mathbf{F}_q$ be the collection of all sets of $\left(\frac{3}{8}\right) 2^q$ distinct $\{-1, 1\}$ vectors of length q. By the union bound,

$$\mathbf{P}(B_q') \leq |\mathbf{F}_q| (3/8)^{n-1} = O((3/8)^n),$$

since $|\mathbf{F}_q| = O(1)$ as it depends only on $q = O(1)$.

Acknowledgement. We would like to thank V. Vu, T. Tao, R. Vershynin, A. Naor and J. H. Kim for enlightening conversations.

References

[1] I. Bárány and B. Doerr, Balanced Partitions of Vector Sequences, *Linear Algebra Appl.* **414** (2006) 464-469

[2] I. Bárány and V.S. Grinberg, On some Combinatorial Questions in Finite Dimensional Spaces, *Linear Algebra Appl.* **41** (1981) pp. 1-9

[3] H. Chernoff, A Measure of Asymptotic Efficiency for Tests of a Hypothesis Based on the Sum of Observations, *Annals of Mathematical Statistics* **23** (1952) pp. 493-507

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

[5] J. Spencer, Balancing Games, *J. Combin. Theory B.* **23** (1977) 68-74

Department of Mathematics, Rutgers, Piscataway, NJ 08854

E-mail address: kcostell@math.rutgers.edu