Докл. Акад. Наук СССР
Том 251 (1980), № 6

Soviet Math. Dokl.
Vol. 21 (1980), No. 2

COMBINATORIAL NUMBERS, GEOMETRIC CONSTANTS AND PROBABILISTIC INEQUALITIES

UDC 519

D. KATONA [GYULA KATONA] AND B. S. STEČKIN

In this note we summarize solutions of the problem of estimating variances of sums of random independent vectors of a normed linear space. The solution is attained by step-by-step increase in the complexity of the structures under consideration: systems of subsets–systems of vectors–systems of random vectors.

1. **Combinatorial numbers.** Suppose S_n is an unordered n-element ($|S| = n$) set. The Turán number $T(n, k, l)$ is the smallest m for which there exists a system of l-subsets $F = \{S_l^{(i)}\}_{1 \leq i \leq m}$, $S_l^{(i)} \subset S_n$, such that $\forall S_k \subset S_n\ \exists S_l^{(i)} \in F$: $S_l^{(i)} \subset S_k$. It is known (see [2] or [4]) that

$$\lim_{n\to\infty} T(n,k,2)n^{-2} = \frac{1}{(2(k-1))}, \qquad \lim_{n\to\infty} T(n,k,l)n^{-l} \geq \frac{(k-l)!}{k!}.$$

2. **Geometric constants.** Suppose X is a normed linear space and $\sigma_n = \{x_1, \ldots, x_n\}$ is a system of (i.e., a collection of not necessarily distinct) vectors $x_i \in X$, $\|x_i\| \geq 1$. The subsystem $\sigma_l \subset \sigma_n$ is an l-system $\sigma_l = \{x_{i_1}, \ldots, x_{i_l}\}$ such that $1 \leq i_1 < \cdots < i_l \leq n$. The geometric constant is defined as the number

$$\delta(l, k; X) = \inf_{\sigma_k \subset X} \max_{\sigma_l \subset \sigma_k} \Big\| \sum_{x \in \sigma_l} x \Big\|.$$

It is known (see [3] and [5]) that if H is at least a $(k-1)$-dimensional Hilbert space, then

$$(2) \qquad \delta(l, k; H) = \sqrt{\frac{l(k-l)}{k-1}}.$$

THEOREM 1. *Suppose $k > l$; then*

$$(3) \qquad \inf_X \delta(l, k; X) = \delta(l, k; l_\infty^k) = \frac{l}{2l-1},$$

where the infimum is taken over all normed linear spaces.

PROOF. Let $\sigma_{l+1} = \{x_1, \ldots, x_{l+1}\}$ and $y_j = x_1 + \cdots + x_{j-1} + x_{j+1} + \cdots + x_{l+1}$ for $j = 1, \ldots, l+1$. Then for every $x_i \in \sigma_{l+1}$ we have the identity

$$lx_i = \sum_{\substack{j=1\\ j\neq i}}^{l+1} y_j - (l-1)y_i,$$

from which, by the triangle inequality, we have

$$\max_j \|y_j\|(l + (l-1)) \geq l\|x_i\| \geq l,$$

1980 *Mathema··· Subject Classification.* Primary 05A05, 60B11, 60G50.

which also implies (3) as a lower bound due to the arbitrariness of σ_{l+1}. To prove the reverse inequality in the space l_∞^k (the space R^k with norm $\|x\| = \max_{1 \leq i \leq k} |x_i|$) we consider a system of k vectors of dimension k of the form

$$\left(-1, \frac{1}{2l-1}, \ldots, \frac{1}{2l-1}\right), \left(\frac{1}{2l-1}, -1, \ldots, \frac{1}{2l-1}\right), \ldots, \left(\frac{1}{2l-1}, \frac{1}{2l-1}, \ldots, -1\right).$$

Direct verification proves that the length of the sum of any l vectors of this system equals $l/(2l-1)$.

3. The relationship between T and δ is essential below.

THEOREM 2 [1]. *In every system $\sigma_n \subset X$ there are at least $T(n, k, l)$ subsystems $\sigma_l \subset \sigma_n$ such that*

$$\Big\| \sum_{x \in \sigma_l} x \Big\| \geq \delta(l, k; X).$$

Let

$$\tau_X(n, \delta, l) = \inf_{\sigma_n \subset X} \left| \left\{ \sigma_l \subset \sigma_n : \Big\| \sum_{x \in \sigma_l} x \Big\| \geq \delta \right\} \right| .$$

Then Theorem 2 is equivalent to the inequality

$$\tau_X(n, \delta, l) \geq T(n, k, l). \tag{4}$$

4. Probabilistic inequalities. The basic result of this note is

THEOREM 3. *Suppose X is a separable normed linear space, and $\xi_1, \ldots, \xi_l$ are independent identically distributed random vectors in it. Then for every $\delta \in R^1$*

$$P\left\{ \Big\| \sum_{i=1}^{l} \xi_i \Big\| \geq x\delta \right\} \geq l! \lim_{n \to \infty} \tau_X(n, \delta, l) n^{-l} P\{\| \xi_i \| \geq x\}^l. \tag{5}$$

As a corollary to Theorems 2 and 3 we have

THEOREM 4. *Suppose X is separable, and $\xi_1, \ldots, \xi_l$ are independent identically distributed random vectors in X. Then for every natural number $k \geq l$*

$$P\left\{ \Big\| \sum_{i=1}^{l} \xi_i \Big\| \geq x\delta(l, k; X) \right\} \geq l! \lim_{n \to \infty} \frac{T(n, k, l)}{n^l} P\{\| \xi_i \| \geq x\}^l. \tag{6}$$

Indeed, it suffices to let $\delta = \delta(l, k; X)$ in (5) and use Theorem 2 in the form (4).

PROOF OF THEOREM 3. A *directed l-graph* on a set of vertices V is a pair $G = (V, E)$, where $E \subset V^l$; and an element $e \in E$, $e = (v_1, \ldots, v_l) \in V^l$, is called a *directed l-edge.* A *nondirected l-graph* is the result of factorization of all edges of a directed l-graph over all permutations of the elements of these edges, so that the Turán number for the directed l-graphs (in which all edges are present together with all their permutations) is not less than $l!T(n, k, l)$. If $G = (V, E)$ is a directed l-graph and $W \subset V$, then G_W denotes the proper subgraph of G induced by the vertices of W, i.e., $G_W = (W, E \cap W^l) = (W, E_W)$. The graph G^v is called a *duplication* of the graph G (on the vertex v) if on the vertices $(V - \{v\}) \cup \{v', v''\}$, G^v contains exactly those edges which can be obtained from the edges of G by replacing v by v' or v'' in them. The class of finite directed l-graphs $\mathcal{G}$ is said to be *duplicated* if every duplication of any graph from this class also belongs to this class. The class of graphs $\mathscr{G}$ is called *complete* if every proper subgraph of any graph of this class also belongs to this class.

Now suppose $H(n, \mathscr{G}) = \min_{G\in \mathscr{G}} |E|/n^l$, where $G = (V, E)$ is a directed l-graph on n vertices V, $|V| = n$.

Suppose $M = (X, \sigma, \mu)$ is a measure space. An infinite directed l-graph $G = (X, E)$ is said to be *measurable* if E is measurable in the product $M^l = (X^l, \sigma^l, \mu_l)$, and $\mu_l(E)$ is its measure. In [6] it was shown that if E is measurable, $\mathscr{G}$ is duplicated and complete, and all finite $E_W \in \mathscr{G}$, then the limit $\lim_{n\to\infty} H(n, \mathscr{G})$ exists and the following inequality is satisfied:

$$(7)\quad \frac{\mu_l(E)}{\mu^l(X)} \geqslant \lim_{n\to\infty} H(n, \mathscr{G}).$$

Now suppose $\mathscr{G}$ consists of all finite l-graphs of the form

$$\left(\{\{x_1, \ldots, x_n\}:\ x_i \in X,\ \|x_i\| \geqslant x\},\ \left\{(x_{i_1}, \ldots, x_{i_l}):\ 1 \leqslant i_1 \leqslant \ldots \leqslant i_l \leqslant n,\ \left\|\sum_{j=1}^{l} x_{i_j}\right\| \geqslant \delta x\right\}\right).$$

Obviously $\mathscr{G}$ is complete. Clearly $\mathscr{G}$ is also duplicated: if x_n is duplicated, then the resulting graph coincides with the graph

$$\left(\{x_1, \ldots, x_n, x_n\},\ \left\{(x_{i_1}, \ldots, x_{i_l}),\ 1 \leqslant i_1 \leqslant \ldots \leqslant i_l \leqslant n:\ \left\|\sum_{j=1}^{l} x_{i_j}\right\| \geqslant \delta x\right\}\right),$$

which by definition lies in $\mathscr{G}$. It is also clear that

$$H(n, \mathscr{G}) \geqslant \frac{l!\,\tau_X(n, \delta, l)}{n^l}.$$

The measure space has the form

$$(\{x_i \in X:\ \|x_i\| \geqslant x\},\ \sigma,\ P),$$

$$E = \left\{(x_{i_1}, \ldots, x_{i_l}).\ 1 \leqslant i_1 \leqslant \ldots \leqslant i_l \leqslant n,\ \left\|\sum_{j=1}^{l} x_{i_j}\right\| \geqslant \delta x,\ x_{i_j} \in X,\ \|x_{i_j}\| \geqslant x\right\}.$$

Thus direct use of (7) leads to the inequality

$$\frac{P\left\{\left\|\sum_{i=1}^{l} \xi_i\right\| \geqslant \delta x\right\}}{P\{\|\xi_i\| \geqslant x\}^l} \geqslant l!\lim_{n\to\infty} \frac{\tau_X(n, \delta, l)}{n^l}.$$

Now use of Theorems 1, 4 and (1), (2) yields

COROLLARY. *In every separable X*

$$(8)\quad P\left\{\left\|\sum_{i=1}^{l} \xi_i\right\| \geqslant \frac{xl}{2l-1}\right\} \geqslant \frac{1}{l+1} P\{\|\xi_i\| \geqslant x\}^l;$$

$$(9)\quad P\{\|\xi + \eta\| \geqslant \tfrac{2}{3}x\} \geqslant \tfrac{1}{2} P\{\|\xi\| \geqslant x\}^2.$$

In Hilbert space

$$(10)\quad P\left\{\left|\sum_{i=1}^{l} \xi_i\right| \geqslant x\sqrt{\frac{l(k-l)}{k-1}}\right\} \geqslant \frac{1}{\binom{k}{l}} P\{|\xi_i| \geqslant x\}^l;$$

$$(11)\quad P\left\{|\xi+\eta| \geqslant x\sqrt{\frac{2(k-2)}{k-1}}\right\} \geqslant \frac{1}{k-1} P\{|\xi| \geqslant x\}^2.$$

5. REMARK. The natural question of how sharp our inequalities are is related to the same question for (4): in a number of cases (4) is not best possible. Inequality (11) cannot be improved with respect to the coefficient on the right side (or even with respect to the exponent, if $k = 3$ and the dimension $\geqslant 2$). See [3] for small-dimensional spaces. It is possible to extend these inequalities to algebraic structures, for example, commutative semigroups.

The authors express their gratitude to all those who furthered the gradual progress in solving this problem over an extended time period, namely: V. V. Arestov, V. I. Berdyšev, Gábor Fejes Tóth, Marek Kanter, A. Sidorenko, Gábor Szász and Gábor Tusnády.

Steklov Institute of Mathematics
Academy of Sciences of the USSR

Received 24/NOV/79

BIBLIOGRAPHY

1. B. S. Stečkin, Zb. Rad. Mat. Inst. Beograd (N. S.) **2 (10)** (1977), 129. (Russian)

2. Paul Erdös and Joel Spencer, *Probabilistic methods in combinatorics,* Academic Press, 1974.

3. G. O. H. Katona, Teor. Verojatnost. i Primenen. **22** (1977), 466; English transl. in Theory Probability Appl. **22** (1977).

4. Gyula Katona, Tibor Nemetz and Miklós Simonovits, Mat. Lapok **15** (1964), 228. (Hungarian)

5. Gyula Katona, Mat. Lapok **20** (1969), 123. (Hungarian)

6. ———, Combinatorics (Proc. Fifth Hungarian Colloq., Keszthely, 1976), Vol. II (Colloq. Math. Soc. János Bolyai, vol. 18), North-Holland, 1978, p. 679.

Translated by L. ROSENBLATT