ISSN 1990-4789, Journal of Applied and Industrial Mathematics, 2013, Vol. 7, No. 2, pp. 127–130. © Pleiades Publishing, Ltd., 2013.
Original Russian Text © S.V. Avgustinovich, E.V. Gorkunov, 2012, published in Diskretnyi Analiz i Issledovanie Operatsii, 2012, Vol. 19, No. 6, pp. 3-8.

On a Code Reconstruction by Correlation Coefficients of Its Subcodes

S. V. Avgustinovich[1,2*] and E. V. Gorkunov[1,2]**

[1] *Sobolev Institute of Mathematics, pr. Akad. Koptyuga 4, Novosibirsk, 630090 Russia*

[2] *Novosibirsk State University, ul. Pirogova 2, Novosibirsk, 630090 Russia*

Received March 27, 2012; in final form, October 17, 2012

Abstract—A generalization is obtained of the theorem on reconstruction of a binary code from dimensions of its subcodes. The notion is proposed of a correlation coefficient of the family of subcodes which, in the present consideration, is an analog for the dimension of a binary subcode.

DOI: 10.1134/S1990478913020014

Keywords: *code, strong isometry, reconstruction, equivalent codes*

INTRODUCTION

Let us consider a q-ary cube E_q^n, the set of words of length n on the alphabet $A = \{0, 1, 2, \dots, q-1\}$. A *Hamming distance* $d(x, y)$ between two words $x, y \in E_q^n$ is defined as the number of symbols in which x and y are distinct; i.e., $d(x, y) = |\{i \mid x_i \neq y_i\}|$. A *weight* $w(x)$ of the word $x \in E_q^n$ is defined as the number of its nonzero symbols, $w(x) = d(x, 0)$. The cube E_q^n together with the Hamming distance between its elements (vertices) forms a metric space. A *code* is defined as an arbitrary subset C of E_q^n. The elements of a code are called *codewords*. Two codes are called *equivalent* if there exists an isometry of E_q^n that maps one code to the other.

In the present article, we study some metric invariants of codes that generalize the notion of dimension of a binary code [1] and establish the sufficiency of the proposed invariants for reconstructing codes up to equivalence.

It is well known [1–5] that coincidence of some metric invariants of two codes does not mean that the codes are equivalent. It turns out that, for $q = 2$, if there is a bijection between two codes which preserves the dimension of each subcode then this bijection may be extended to an isometry of the whole space. In other words, the collection of dimensions of the subcodes of a binary code defines this code up to equivalence. Here, the *dimension of a code* means the dimension of a minimal face of E_2^n containing this code. Later, it has turned out [6] that the full collection of dimensions of the subcodes is superfluous, it suffices to consider only the subcodes of even powers; i.e., the bijection between codes preserving these dimensions may be extended to an isometry of E_2^n, and the codes turn out equivalent.

Generalization of this approach, applied to the binary codes, does not lead to success for an arbitrary $q > 2$. For example, there exist nonequivalent ternary codes that have a bijection preserving the dimensions of their subcodes. Thus, in the general case, some finer methods are necessary for q-ary codes. For clarity, we demonstrate the approach presented in this article on ternary codes. Generalization to an arbitrary integer q can be obtained naturally (see Section 3).

[*] E-mail: avgust@math.nsc.ru
[**] E-mail: evgumin@gmail.com

1. DEFINITIONS AND THE MAIN RESULT

Let C be a ternary code of length n. If all codewords of C have in some position the same symbol then we say that this position is *unessential* for C. If C_1 and C_2 are some disjoint subcodes of C then we denote by $K(C_1, C_2)$ the number of coordinate positions which are nonessential for both subcodes but in which the codewords of different codes have different values. Formally,

$$K(C_1, C_2) = |\{i \mid \exists\, a, b \in A, a \neq b : \forall x \in C_1\ x_i = a,\ \forall y \in C_2\ y_i = b\}|.$$

We call the value $K(C_1, C_2)$ the *correlation coefficient* of C_1 and C_2. Note some characteristic identities explaining this notion:

(i) $K(x, y) = d(x, y)$, where $x, y \in E_3^n$;

(ii) $K(\{x, y\}, \varnothing) = n - d(x, y)$;

(iii) $K(C, \varnothing) = n - \mathrm{Dim}(C)$, where $\mathrm{Dim}(C)$ is the dimension of C.

Thus, by means of the correlation coefficients, we define, in particular, the distance between codewords.

In compliance with [6], a bijection $I\colon C_1 \to C_2$ between the codes $C_1, C_2 \subset E_3^n$ is called a *strong isometry* if it preserves the correlation coefficient of every pair of subcodes of C_1. Let M_1 be a code matrix of C_1, denote by $M_2 = I(M_1)$ the code matrix of C_2 obtained by applying I to each row of M_1. The main result of this article is the following

Theorem. *Each strong isometry of the ternary codes can be extended to an isometry of the ternary cube.*

Let us consider a code $C \subset E_3^n$ of power m, and let M be its code matrix. For an arbitrary column of M, the symbols of the alphabet A generates the *alphabet partition* of the set $\{1, \dots, m\}$ of numbers of the rows: each subset of the partition contains all numbers of the rows in which the elements of the column under consideration have the same value. For example, the partition $(\{1, 3\}, \{2, 4\}, \{5\})$ corresponds to the columns $(0, 1, 0, 1, 2)^\top$ and $(2, 1, 2, 1, 0)^\top$, and still more five ordered partitions differing from the given in order of writing the sets. For every 3-partition P, denote by $k(P)$ the number of columns of the matrix M for which P is one of the six alphabet partitions.

Further, we adhere to the notations:

$\mathcal{M} = \{1, \dots, m\}$ is the set of numbers of the rows of a matrix M;

$P = (P_0, P_1, P_2)$, $Q = (Q_0, Q_1, Q_2)$, and $R = (R_0, R_1, R_2)$ are the alphabet partitions of $\mathcal{M}$.

Given a subset $S \subseteq \mathcal{M}$, let $C(S)$ denote a subcode of C formed by the rows of M whose numbers are in S. If it is clear what code is under consideration then, for $P_1, P_2 \subseteq \mathcal{M}$, instead of $K(C(P_1), C(P_2))$ we write $K(P_1, P_2)$.

It is easy to show

Proposition 1. *Each bijection $I\colon C_1 \to C_2$ can be extended to an isometry of all cube if an only if, for every alphabet partition, the code matrices M_1 and $M_2 = I(M_1)$ contain the same number of columns with such a partition.*

On the set of alphabet partitions, we define the partial order $\preccurlyeq$ as follows:

$$(P_0, P_1, P_2) \preccurlyeq (Q_0, Q_1, Q_2) \iff P_0 \subseteq Q_0,\ P_1 \supseteq Q_1,\ P_2 \supseteq Q_2.$$

Note that, for a code $C \subset E_3^n$ of power m, its fixed code matrix M, and the alphabet partition $P = (P_0, P_1, P_2)$, we have

$$K(P_1, P_2) = \sum_{Q \preccurlyeq P} k(Q). \tag{1}$$

2. RECONSTRUCTION OF TERNARY CODES FROM THE CORRELATION COEFFICIENTS OF THEIR SUBCODES

The main step in the proof of the theorem is conversion of (1). In other words, our nearest aim is as follows: Having the correlation coefficients of subcodes, we try to calculate the number of columns of certain type in the code matrix.

Proposition 2. *For a ternary code with an alphabet partition P, the number of columns in the code matrix is equal to*

$$k(P) = \sum_{Q\preccurlyeq P}(-1)^{|P_0|-|Q_0|}K(Q_1, Q_2). \tag{2}$$

Proof. We proceed by induction on the power of P_0. By (1), if $P_0 = \varnothing$ then $K(P_1, P_2) = k(P)$. Thus, the base of induction is constructed. Assume that the assertion is true for $|P_0| < s$ and prove it for $|P_0| = s$.

From (1), transferring the terms, we obtain

$$k(P) = K(P_1, P_2) - \sum_{Q\prec P} k(Q).$$

Further, using the inductive assumption, insert (2) in the last equation and change the order of summation:

$$\begin{aligned} k(P) &= K(P_1, P_2) - \sum_{Q\prec P}\sum_{R\preccurlyeq Q}(-1)^{|Q_0|-|R_0|}K(R_1, R_2) \\ &= K(P_1, P_2) - \sum_{R\prec P}(-1)^{-|R_0|}K(R_1, R_2)\sum_{R\preccurlyeq Q\prec P}(-1)^{|Q_0|}. \end{aligned} \tag{3}$$

Fixing the partition $R \prec P$, denote $|R_0| = t$ where $0 \leqslant t < s$. In order to obtain the partition Q from the interval between R and P such that $|Q_0| = t + i$, it is necessary and sufficient that i elements of the set $P_0 \setminus R_0$ be joined to R_0. Here, i must assume any integer value from 0 to $s - t - 1$. Thus,

$$\sum_{R\preccurlyeq Q\prec P}(-1)^{|Q_0|} = \sum_{i=0}^{s-t-1} C^i_{s-t}(-1)^{t+i} = (-1)^t(1-1)^{s-t} - (-1)^t(-1)^{s-t} = -(-1)^s.$$

Taking into account (3), we continue and obtain

$$k(P) = K(P_1, P_2) + \sum_{R\prec P}(-1)^{s-|R_0|}K(R_1, R_2) = \sum_{R\preccurlyeq P}(-1)^{|P_0|-|R_0|}K(R_1, R_2).$$

The proof of Proposition 2 is complete. □

Proof of the theorem. Proposition 2 implies that the matrices M_1 and M_2 have the same, up to permutation, families of the alphabet partitions. It remains to apply Proposition 1 to complete the proof. □

Corollary. *Strong isometric codes are equivalent.*

3. REMARKS

In conclusion, we present some judgments concerning generalization of the results of our article.

1. The theorem can be generalized to the codes over an alphabet of a power q as follows: For such a code, it is necessary to take into account its correlation coefficient for an arbitrary $q - 1$ pairwise nonintersecting subcodes. Such passage to the general case does not change the statement of the theorem essentially.

2. The main goal in the proof of extendability of strong isometry, as it stated in the present article, is to determine the column structure of the code matrices of the codes under consideration. Given the code matrix of a ternary code of weight m we can find 3^m different columns (for the codes of sufficiently large length). Therefore, generally speaking, to determine the column structure of a code matrix, it is necessary to calculate the 3^m values which characterize the number of entrances of each m-column in the given code matrix. Obviously, for this goal, the knowledge is superfluous of the correlation coefficients of each pair of subcodes. Thus, the problem arises of finding a minimal family of the correlation coefficients which is sufficient for reconstructing the code matrix and establishing the equivalence of codes.

ACKNOWLEDGMENTS

The authors were supported by the Russian Foundation for Basic Research (project no. 10–01–00424) and the Target Program of the Siberian Division of the Russian Academy of Sciences for 2012–2014 (Integration project no. 14).

REFERENCES

1. S. V. Avgustinovich, “On the Strong Isometry of Binary Codes,” Diskret. Anal. Issled. Oper. Ser.1. **7** (3), 3–5 (2000).
2. Zh. K. Abdurakhmanov “On the Geometric Structure of Error-Correcting Codes,” Candidate’s Dissertation in Mathematics and Physics (Tashkent, 1991).
3. S. V. Avgustinovich and F. I. Solov’eva, “To the Metric Rigidity of Binary Codes,” Problemy Peredachi Informatsii **39** (2), 23–28 (2003) [Problems Inform. Transmission **39** (2), 178–183 (2003)].
4. V. Yu. Krasin, ”On the Weak Isometries of the Boolean Cube,” Diskret. Anal. Issled. Oper. Ser.1. **13** (4), 26–32 (2006) [J. Appl. Indust. Math. **1** (4), 463–467 (2007)].
5. F. I. Solov’eva, S. V. Avgustinovich, T. Honold, and W. Heise, “On the Extendability of Code Isometries,” J. Geom. **61**, 3–16 (1998).
6. E. V. Gorkunov and S. V. Avgustinovich, “On the Reconstruction of Binary Codes from the Dimensions of Their Subcodes,” Diskret. Anal. Issled. Oper. **17** (5), 15–21 (2010) [J. Appl. Indust. Math. **5** (3), 348–351 (2011)].