On the capacity for almost disjunctive codes.

D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.

Vorobyev I.V.

Lomonosov Moscow State University, Faculty of Mechanics and Mathematics

ACCT 2014

Basic Notations

Let $\triangleq$ denote equality by definition, $|A|$ be the cardinality of a set A, $[n] \triangleq 1, 2, \ldots, n$ be the set of positive integers from 1 to n.

A binary matrix $X \triangleq ||x_i(j)||, \quad i \in [N], \quad j \in [t]$, is called a *code* with t codewords(columns) $\boldsymbol{x}(1)$, $\boldsymbol{x}(2)$, . . ., $\boldsymbol{x}(t)$ of length N.

We say that a binary vector $\boldsymbol{x}$ *covers* binary vector $\boldsymbol{y}$ $(\boldsymbol{x} \succeq \boldsymbol{y})$ if $x_i \geq y_i$ for any $i \in [N]$.

Definition 1.

An s-subset of columns $\boldsymbol{x}(\mathcal{S})$, $|\mathcal{S}| = s$, of a code X is said to be an *s-bad* subset of columns in the code X if there exists an another column $\boldsymbol{y}$, such that the disjunctive sum

$$\bigvee_{i \in \mathcal{S}} \boldsymbol{x}(i) \succeq \boldsymbol{y}.$$

Otherwise, the s-subset $\boldsymbol{x}(\mathcal{S})$ is called *s-good* subset of columns in the code X.

Example: $s = 2$.

$\boldsymbol{x}(1)$	$\boldsymbol{x}(2)$	$\boldsymbol{x}(3)$	$\boldsymbol{x}(4)$	$\boldsymbol{x}(5)$
1	1	0	0	1
0	0	1	0	1
1	0	1	1	0
0	1	0	1	0

Set $\{\boldsymbol{x}(3), \boldsymbol{x}(4)\}$ is good.
Set $\{\boldsymbol{x}(2), \boldsymbol{x}(3)\}$ is bad.

Definition 2.

A code X is said to be a *disjunctive (s, ε)-code*, if the number of all s-good subsets of columns of the code X is at least $(1 - \varepsilon) \cdot \binom{t}{s}$.

Example: In case $\varepsilon = 0$, we obtain s-disjunctive code.

Definition 3.

Let $t_\varepsilon(N, s)$ be the maximal size of $(s,\ \varepsilon)$-codes of length N and let $N_\varepsilon(t, s)$ be the minimal length of $(s,\ \varepsilon)$-codes of size t. If $\varepsilon = 0$, then the number

$$R(s) \triangleq \overline{\lim_{N\to\infty}} \frac{\log_2 t_0(N, s)}{N} = \overline{\lim_{t\to\infty}} \frac{\log_2 t}{N_0(t, s)}$$

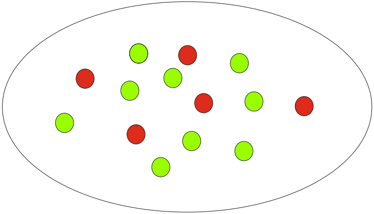
is called the *rate* of s-codes.

Define the number

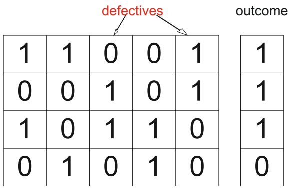
$$C(s) \triangleq \overline{\lim_{\varepsilon\to 0}}\ \overline{\lim_{N\to\infty}} \frac{\log_2 t_\varepsilon(N, s)}{N} = \overline{\lim_{\varepsilon\to 0}}\ \overline{\lim_{t\to\infty}} \frac{\log_2 t}{N_\varepsilon(t, s)}$$

called the *capacity* of almost disjunctive s-codes.

Combinatorial Group Testing



Problem: identify all defective elements among t items.
Assumption: at most s defectives.
Solution: use disjunctive codes.



$$\begin{array}{|c|c|c|c|c|}\hline 1 & 1 & 0 & 0 & 1 \\ \hline 0 & 0 & 1 & 0 & 1 \\ \hline 1 & 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 1 & 0 \\ \hline \end{array} \qquad \begin{array}{|c|}\hline 1 \\ \hline 1 \\ \hline 1 \\ \hline 0 \\ \hline \end{array}$$

Outcome covers only s defective columns.

In case of almost disjunctive codes parameter ε can be treated as a probability of failure.

Results

Lower bound

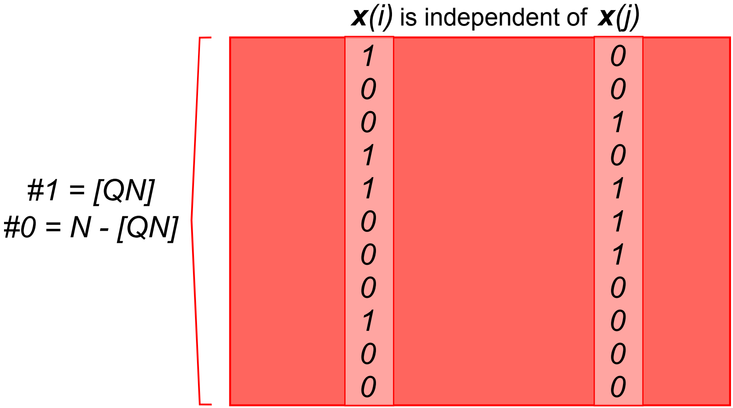
$$C(s) \geq \underline{C}(s) \triangleq \max_{0<Q<1} C(s,Q),$$

$$C(s,Q) \triangleq h(Q) - [1-(1-Q)^s]\, h\left(\frac{Q}{1-(1-Q)^s}\right),$$

$$h(a) \triangleq -a\log_2 a - (1-a)\log_2(1-a).$$

Sketch of the Proof.

The proof is based on the random coding method over the ensemble of constant-weight binary codes.



$\boldsymbol{x}(i)$ is independent of $\boldsymbol{x}(j)$

#1 = [QN]
#0 = N - [QN]

1	0
0	0
0	1
1	0
1	1
0	1
0	1
0	0
1	0
0	0
0	0

We calculate the probability of bad subset as

$$\sum_{k} \Pr\left\{ \boldsymbol{x}(S) \text{ is } s\text{-bad in } X \;\Big/\; \left| \bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i) \right| = k \right\} \mathcal{P}^{(N)}(s, Q, k),$$

where we applied the total probability formula and introduced the notation

$$\mathcal{P}^{(N)}(s, Q, k) \triangleq \Pr\left\{ \left| \bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i) \right| = k \right\}.$$

Further we bound $\Pr\left\{ \boldsymbol{x}(S) \text{ is } s\text{-bad in } X \;/\; \left| \bigvee_{i\in\mathcal{S}} \boldsymbol{x}(i) \right| = k \right\}$ by $\min\left\{ 1;\ (t-s)\frac{\binom{k}{\lfloor QN \rfloor}}{\binom{N}{\lfloor QN \rfloor}} \right\}$ and proceed to the optimization problem.

Asymptotics

$$\underline{C}(s) = \frac{\ln 2}{s}(1+o(1)), \text{ at } Q(s) = \frac{\ln 2}{s}(1+o(1)).$$

Comparison with Disjunctive Codes

The rate of disjunctive codes

$$\frac{4\log_2 s}{e^2 s^2}(1+o(1)) \leq R_0(s) \leq \frac{2\log_2 s}{s^2}(1+o(1)).$$

The capacity of almost disjunctive codes

$$\frac{\ln 2}{s}(1+o(1)) \leq C(s) \leq \frac{1}{s}.$$

s	2	3	4	5	6
$\underline{C}(s)$	0.3832	0.2455	0.1810	0.1434	0.1188
$\underline{R}_0(s)$	0.1825	0.0787	0.0439	0.0279	0.0194
$\underline{R'}_0(s)$	0.1281	0.0821	0.0566	0.0420	0.0325

Bound $\underline{R}_0(s)$ is taken from paper [1989, D'yachkov A.G., Rykov V.V., Rashad A.M.].

Bound $\underline{R'}_0(s)$ is taken from paper [2013, D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.].

Constructions

1. Construction based on shortened Reed-Solomon codes for classical disjunctive code.[2000, D'yachkov A.G., Macula A.J., Rykov V.V.]
2. MDS codes [2013, Bassalygo, Rykov].
   Let $t = 2^{\frac{q}{\log_2 q}}, \quad N = q(q+1), \quad s = \sigma q$, where q – prime power, σ – some positive constant. If $\sigma < \ln 2$ then $\varepsilon(q) \to 0$ exponentially and rate $R = \frac{\log_2 t}{N} \to \frac{\sigma}{s}$.

Bibliography

[1] *D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.*, "Bounds on the Rate of Disjunctive Codes"// 2014 IEEE International Symposium on Information Theory, Honolulu, USA, July 2014.

[2] *Dyachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.*, "Almost Disjunctive List-Decoding Codes"// arXiv:1407.2482 [cs.IT], 2014.

[3] *D'yachkov A.G., Vorobyev I.V., Polyanskii N.A., Shchukin V.Yu.*, "Bounds on the Rate of Disjunctive Codes"// Problems of Information Transmission, vol. 50, no. 1, pp. 27-56, 2014.

[4] *D'yachkov A.G., Rykov V.V., Rashad A.M.*, "Superimposed Distance Codes"// Probl. Control Inform. Theory. 1989. V. 18. no 4. P 237-250.

[5] *D'yachkov A.G., Macula A.J., Rykov V.V.* New Constructions of Superimposed Codes. // IEEE Trans. Inform. Theory. 2000. V.46. n 1. P. 284-290.

[6] *Bassalygo L. A.*, *Rykov V. V.*, "Multiple-access hyperchannel"// Problems of Information Transmission, vol. 49, no. 4, pp. 299-307, 2013.

Thank you for your attention!