Telecom Paris ACCQ204, Coding Theory

ASSIGNMENT 2

Exercise 1. Determine the parameters (n, k, d) of the binary code

$$C = \{00001100, 00001111, 01010101, 11011101\}$$

Exercise 2 ($A(n, d)$, extending, puncturing, expurgating)**.** Define the intersection of length n binary vectors x and y to be the vector $x * y = (x_1y_1, x_2y_2, \ldots, x_ny_n)$.

1. Show that
$$wt(x + y) = wt(x) + wt(y) - 2wt(x * y)$$

2. Show that $A(n, d) \leq A(n - 1, d - 1)$ where $wt(x)$ denotes the Hamming weight of x. Hint: consider 'puncturing', that is removing a common coordinate from every codeword.

3. Show that $A(n, 2r - 1) = A(n + 1, 2r)$ where $A(n, d)$ denotes the largest number of length n codewords with minimum distance d. Hint: consider 'extending' codewords by adding a parity check bit, i.e., $x_1, x_2, \ldots, x_n$ becomes $x_1, x_2, \ldots, x_n, \sum x_i$.

4. Show that $A(n, d) \leq 2A(n - 1, d)$. Hint: consider dividing codewords into two classes, those beginning with a 0 and those beginning with a 1.

Exercise 3. For each of the following codes

$$C_1 = \{00000, 01010, 00001, 01011, 01001\}$$

$$C_2 = \{000000, 101000, 001110, 100111\}$$

$$C_3 = \{0000, 1100, 1010, 1001, 0110, 0101, 0011, 1111\}.$$

tell if it is linear and evaluate the parameters (n, k, d).

Exercise 4. The dual of an $[n, k]_q$ code $\mathcal{C}$ is the set

$$\mathcal{C}^{\perp} = \{c \in \mathbb{F}_q^n : \langle x, y\rangle = 0 \text{ for all } y \in \mathcal{C}\}$$

($\langle \cdot, \cdot \rangle$ denotes the standard "scalar" product).

Show that if G and H are the generator and parity matrices, respectively, of $\mathcal{C}$, then H and G are the generator and parity matrices, respectively, of $\mathcal{C}^{\perp}$.

Exercise 5. Let C_1 and C_2 be an $[n, k_1, d_1]$ and an $[n, k_2, d_2]$ code, respectively. Let $C_1|C_2$ be the code consisting of all codewords of the form

$$(u, u + v) = (u_1, u_2, \ldots, u_n, u_1 + v_1, u_2 + v_2, \ldots, u_n + v_n)$$

with $u = (u_1, u_2, \ldots, u_n) \in C_1$ and $v = (v_1, v_2, \ldots, v_n) \in C_2$. Show that $C_1|C_2$ is an $[2n, k_1 + k_2, \min\{2d_1, d_2\}]$ code. Hint. consider the cases $v = v'$ and $v \neq v'$. For the second case use the triangle inequality.

Exercise 6. In this exercise we show the existence of linear codes over $[q]$, $q \geq 2$, which achieve the Gilbert-Varshamov bound. To that aim we show the existence of a full rank generator matrix G of dimension $k \times n$ such that

$$k = (1 - H_q(\delta) - \varepsilon)n$$

and such that

$$wt(mG) \geq d$$

for any $m \in \mathbb{F}_q^k$.

1. Pick G randomly such that each of its elements is independently chosen with the uniform distribution over $[q]$. Fix $m \neq 0$. We first show that for such a random G, mG is a uniformly chosen vector over $[q]^n$.

   (a) Let X_i denote the i-th symbol of the n-vector mG. Show that X_i is independent of X_j for $i \neq j$.

   (b) Let $X_i = \sum_{j=1}^{k} m_j G_{ji}$. Since $m \neq 0$, at least one of its elements is non-zero. Say m_ℓ is the first non-zero element. Thus we can write $X_i = m_\ell G_{\ell i} + \sum_{j=\ell+1}^{k} m_j G_{ji}$. Using this, show that X_i is uniformly distributed over $[q]$ by conditioning over the possible realizations of $G_{\ell+1,i}, G_{\ell+2,i}, \ldots, G_{k,i}$.

2. Deduce that

$$Pr[wt(mG) < d] \leq \frac{q^{nH_q(\delta)}}{q^n}.$$

   Hint. $Vol_q(d-1, n) \leq q^{nH_q(\delta)}$.

3. Deduce that $Pr(\exists m : \ wt(mG) < d) \leq q^{-\varepsilon n}$ for some appropriate choice of k.

4. Conclude the proof.

Exercise 7 (Perfect codes)**.** A code is a perfect $t-$error correcting code if the set of $t-$spheres centered on the codewords fill the Hamming space $\{0,1\}^n$ without overlapping. Here we will show that such codes do not, in general, achieve the capacity of the BSC.

Consider a set of three codewords of length n. Let un denote the number of positions where the first codeword differs from both the second and the third codewords, let vn denote the number of positions where the second codeword differs from both the first and the third codewords, let wn denote the number of positions where the third codeword differs from both the first and the second codewords, and finally let zn denote the number of positions where the three codewords agree.

1. Argue that we can assume, without loss of generality, that one of them is the all-zero codeword.

2. Assuming that the code is $f \cdot n$-error correcting, give necessary conditions on u, v, w.

3. Show that for a certain range of f we must have $u + v + w > 1$ which is impossible.

4. Conclude that, for a certain range of f, perfect codes do not exist.

5. Reconcile this result with the Shannon's result which says that 'with high probability it is possibe to correct $f \cdot n$ errors with exponentially many codewords'.

Exercise 8. Is the code $C = \{000, 110, 011, 101\}$ MDS?

Exercise 9. Consider an $[n, k, d]$ MDS code over $\mathbb{F}_q$. Show that

1. the number of codewords of weight d is

$$N_d = \binom{n}{d}(q-1).$$

   Hint. Pick a subset of $k-1$ coordinates and fix the corresponding values to zero. Pick any other coordinate and let the symbol value in this coordinate run through all q symbols in $\mathbb{F}_q$.

2. Show that the number of codewords of weight $d+1$ is

$$N_{d+1} = \binom{n}{d+1}\left((q^2-1) - \binom{d+1}{d}(q-1)\right).$$