Telecom ParisTech | ACCQ204, Coding Theory
Teacher: Aslan Tchamkerten | December 2017

ASSIGNMENT 2 - SOLUTIONS

Exercise 1 ($A(n,d)$, 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 $A(n,d) \leq A(n-1,d-1)$. 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 lenth 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.

Solution. 1. Immediate

2. If we delete a coordinate from an (n, M, d) code (n refers to the codeword length, M to the number of codewords, and $2r-1$ to the minimum distance), we get an $(n-1, M, \geq d-1)$ code, hence $A(n,d) \leq A(n-1,d-1)$.

3. Let $\mathcal{C}$ be an $(n, M, 2r-1)$ code. By adding an overall parity check bit we get an $(n+1, M, 2r)$ code since the minimum distance must be even by 1. and that adding a parity check cannot increase the minimum distance my more than 1. Therefore $A(n, 2r-1) \leq A(n+1, 2r)$. Conversely, deleting one coordinate gives an $(n, M, d \geq 2r-1)$ code (see 2.), hence $A(n, 2r-1) \geq A(n+1, 2r)$.

4. Consider and (n, M, d) code. Using the hint, consider removing the smallest of the two classes. The remaining class has at least $M/2$ codewords and its minimum distance is at least d. Therefore $A(n,d)/2 \leq A(n-1,d)$.

□

Exercise 2. Determine the parameters (n, k, d) of the binary code

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

Solution. $n = 8$, $k = 3$, $d = 2$ □

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 $\mathcal{C}^\perp$.

Solution. For any x, x' in the message spaces of $\mathcal{C}$ and $\mathcal{C}^\perp$, respectively, we have

$$\langle xG, x'H\rangle = xGH^T x' = 0$$

since $GH^T = 0$ (see Lemma in the course). Therefore H is the generator matrix of $\mathcal{C}^\perp$ and G its generator matrix (since $HG^T = (HG^T)^{TT} = (GH^T)^T = 0$ by the same lemma). □

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.

Solution. That $C_1|C_2$ has length $2n$ and dimension $k_1 + k_2$ is obvious. Let us consider the minimum distance. If $a = u, u + v$ and $b = u', u' + v'$ are different codewords then $d(a, b) = d(u, u') + d(u + v, u' + v')$. Using this we get

- If $v = v'$ then $d(a, b) \geq 2d_1$
- If $v \neq v'$ then

$$\begin{aligned} d(a, b) &= wt(u - u') + wt(u + v - u' - v') \\ &= wt(u' - u) + wt(u + v - u' - v') \\ &\geq wt(u - u' + u + v - u' - v') \\ &= wt(v - v') \\ &\geq d_2 \end{aligned}$$

This shows that the minimum distance of $C_1|C_2$ is $\geq \min\{2d_1, d_2\}$ and it is easy to check that this bound is indeed achievable. $\square$

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.

Solution. 1. (a) Holds since X_i and X_j involve different columns of G and that these columns are independent.
   (b) We have

$$\begin{aligned} P(X_\ell = x) &= \frac{1}{q^{k-\ell}} \sum_{(g_{\ell+1,i}, g_{\ell+2,i}, \cdots, g_{k,i})} P(X_\ell = x | (G_{\ell+1,i}, G_{\ell+2,i}, \ldots, G_{k,i}) = (g_{\ell+1,i}, g_{\ell+2,i}, \ldots, g_{k,i})) \\ &= \frac{1}{q^{k-\ell}} \sum_{(g_{\ell+1,i}, g_{\ell+2,i}, \cdots, g_{k,i})} \frac{1}{q} \\ &= \frac{1}{q}. \end{aligned}$$

2. Holds because of 1.

3. Holds by a union bound over m and by letting $k = (1 - H_q(\delta) - \varepsilon)n$.

4. By the previous step, and because the matrix G is uniformly distributed, as $n \to \infty$ the fraction of the matrices satisfying the desired property tends to one.

□

Exercise 7. Is the code $C = \{000, 110, 011, 101\}$ MDS?

Solution. $n = 3$, $k = 2$, $d = 2$, hence $d = n - k + 1$ and it is an MDS code. □

Exercise 8. Suppose we are in $\mathbb{F}_2$. Find

1. $\gcd(x^4 + x^2 + 1, x^2 + 1)$

2. $\gcd(x^6 + x^5 + x^3 + x + 1, x^4 + x^2 + 1)$

3. $\gcd(x^6 + x^5 + x^3 + x + 1, x^4 + x^3 + x + 1)$

Solution.

1. 1

2. $x^4 + x^2 + 1$

3. $x^2 + x + 1$

□

Exercise 9. Show that a Reed-Solomon code with 1 message symbol and n codeword symbols is an n times repetition code.

Solution. If we have a 1 message symbol, encoding polynomials are of degree zero (i.e., are constants) and evaluated n times. □