Telecom Paris ACCQ204, Coding Theory
Teacher: Aslan Tchamkerten

Assignment 3 - Solutions

Exercise 1. 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 2. Show that a Reed-Solomon code with 2 message symbols and n codeword symbols is an n times repetition code.

Solution. If we have a 2 message symbols, encoding polynomials are of degree zero (i.e., are constants) and evaluated n times. □

Exercise 3. Construct an $RS(n = 4, k = 2)$ code. For the construction you may want to consider the irreducible polynomial $X^2 + X + 1$ over $\mathbb{F}_2$ and the evaluation points (to be justified) $\alpha_1 = 0$, $\alpha_2 = 1$, $\alpha_3 = x$, $\alpha_4 = x + 1 = x^2$.

Solution. Since $n = 4$ we need a base field with (at least) 4 elements. So let's choose the base field $\mathbb{F}_4 = \mathbb{F}_2[X]/(X^2 + X + 1)$ whose elements are thus

$$\{0, 1, x, x + 1 = x^2\}.$$

Since $k = 2$, the message polynomials are of degree $k - 1 = 1$ and can be written as $f_0 + f_1 x$ with $f_0, f_1 \in \mathbb{F}_4$. Thus the mapping between information symbols and codewords is given by

$$(f_0, f_1) \to (f_0 + f_1\alpha_1, f_0 + f_1\alpha_2, f_0 + f_1\alpha_3, f_0 + f_1\alpha_4).$$

The full mapping is thus

$$\begin{array}{llcllll@{\qquad}llcllll}
0 & 0 & \to & (0 & 0 & 0 & 0) & x & 0 & \to & (x & x & x & x) \\
0 & 1 & \to & (0 & 1 & x & x+1) & x & 1 & \to & (x & x+1 & 0 & 1) \\
0 & x & \to & (0 & x & x+1 & 1) & x & x & \to & (x & 0 & 1 & x+1) \\
0 & x+1 & \to & (0 & x+1 & 1 & x) & x & x+1 & \to & (x & 1 & x+1 & 0) \\
1 & 0 & \to & (1 & 1 & 1 & 1) & x+1 & 0 & \to & (x+1 & x+1 & x+1 & x+1) \\
1 & 1 & \to & (1 & 0 & x+1 & x) & x+1 & 1 & \to & (x+1 & x & 1 & 0) \\
1 & x & \to & (1 & x+1 & x & 0) & x+1 & x & \to & (x+1 & 1 & 0 & x) \\
1 & x+1 & \to & (1 & x & 0 & x+1) & x+1 & x+1 & \to & (x+1 & 0 & x & 1)
\end{array}$$

□

Exercise 4. Consider the following mapping from $(\mathbb{F}_q)^k$ to $(\mathbb{F}_q)^{k+1}$. Let $(f_0, f_1, \ldots, f_{k-1})$ be any k-tuple over $\mathbb{F}_q$, and define the polynomial $f(x) = f_0 + f_1 x + \ldots + f_{k-1}x^{k-1}$ of degree less than k. Map $(f_0, f_1, ..., f_{k-1})$ to the $(q+1)$-tuple $(\{f(\alpha_i), \alpha_i \in \mathbb{F}_q\}, f_{k-1})$—i.e., to the RS codeword corresponding to $f(x)$, plus an additional component equal to f_{k-1}.

Show that the $q^k (q+1)$-tuples generated by this mapping as the polynomial $f(z)$ ranges over all q^k polynomials over $\mathbb{F}_q$ of degree $< k$ form a linear $(n = q+1, k, d = n-k+1)$ MDS code over $\mathbb{F}_q$. [Hint:$f(x)$ has degree $< k-1$ if and only if $f_{k-1} = 0$.]

Solution. The code has length $n = q+1$. It is linear because the sum of codewords corresponding to$f(x)$ and $g(x)$ is the codeword corresponding to $f(x) + g(x)$, another polynomial of degree less than k. Its dimension is k because no polynomial other than the zero polynomial maps to the zero $(q+1)$-tuple.

To prove that the minimum weight of any nonzero codeword is $d = n-k+1$, use the hint and consider the two possible cases for f_{k-1}:

- If $f_{k-1} \neq 0$, then deg$f(x) = k-1$. By the fundamental theorem of algebra, the RS codeword corresponding to $f(x)$ has at most $k-1$ zeroes. Moreover, the f_{k-1} component is nonzero. Thus the number of nonzero components in the code $(q+1)$-tuple is at least $q-(k-1)+1 = n-k+1$.

- If $f_{k-1} = 0$ and $f(x) = 0$, then deg$f(x) \leq k-2$. By the fundamental theorem of algebra, the RS codeword corresponding to $f(x)$ has at most $k-2$ zeroes, so the number of nonzero components in the code $(q+1)$-tuple is at least $q-(k-2) = n-k+1$.

□

Exercise 5. Suppose we want to correct bursts of errors, that is error patterns that affect a certain number of consecutive bits. Suppose we are given an $[n, k]$ RS code over $\mathbb{F}_{2^t}$. Show that this code yields a binary code which can correct any burst of $(\lfloor (n-k) \rfloor /2 - 1)t$ bits.

Solution. Map each 2^t symbols of $\mathbb{F}_{2^t}$ into t bits. The code can correct up to $(d-1)/2$ symbol errors which translates into an error correction capability of $(\lfloor (d-1)/2 \rfloor - 1)t$ consecutive bits $(\lfloor (d-1)/2 \rfloor t$ if the burst of errors starts at the beginning of a symbol). □

Exercise 6 (Secret sharing)**.** Throughout, we let $\mathcal{C}$ be a binary linear code of length n. We say that a codeword v' *covers* a codeword v if the non-zero components of v are a subset of the non-zero components of v'. A non-zero codeword v is said to be minimal if it covers no other codeword.

1. Let v' be a non-zero non-minimal codeword of $\mathcal{C}$. Argue that v' covers some minimal codeword which we denote as $v(1)$.

   Solution. Since v' is non-minimal it covers at least some other codeword. □

2. Argue that $v' - v(1)$ is another codeword with weight strictly less than v'.

   Solution. By linearity $v' - v(1)$ is another codeword. It has weight strictly less than v' since a set of non-zero coordinates of v' are flipped. □

3. Deduce that $v' - v(1) - v(2) - \ldots - v(s) = 0$ for some minimal codewords $v(1), \ldots, v(s)$.

   Solution. By recursion. □

4. Secret sharing: Let $\mathcal{C}$ be an $[n, k]$ binary linear code. An information set $\mathcal{I}$ is a set of k components whose values entirely specify any codeword (for instance, for an MDS code, any k components is an information set). Show that there always exists an information set that contains the first component, unless all codewords have their first component equal to zero.

   Solution. If the first component of every codeword is always 0, the first component provides no information about the codeword. Thus, the first component cannot be included in any information set, because otherwise the information set would be of size $k - 1$. Alternatively, observe that in this case the first column of the generator matrix is zero.

   If there exists at least one codeword with $c_1 = 1$, the first component contributes information and may be part of an information set as we describe next. To find an information set that contains the first component it suffices to select k indices corresponding to k columns of the generator that are linearly independent and that contain the first column—the resulting $k \times k$ submatrix of the generator matrix is nonsingular.

   □

5. Pick $v_1 \in \{0, 1\}$ uniformly at random, this will be our "secret". Assign uniformly random values from $\{0, 1\}$ to all $k-1$ components $v_j, j \in \mathcal{I} \backslash \{1\}$, independently of v_1. From $\{v_j, j \in \mathcal{I}\}$ compute the full codeword $v = v_1, v_2, \ldots, v_n$. Distribute digits $v_2, v_3, \ldots, v_n$ to $n - 1$ distinct persons.

   We now provide secrecy analysis for this scheme and analyze the sets of persons that are able to recover the secret v_1.

   (a) A set of t persons, with combined knowledge of $v_{j_1}, v_{j_2}, \ldots, v_{j_t}$, represents a *critical set* if they can recover the secret v_1 without error, but any proper subset of these persons recovers the value of v_1 only with probability $1/2$. Show that if a set of t persons, with combined knowledge of $v_{j_1}, v_{j_2}, \ldots, v_{j_t}$, represents a *critical set*, then

   $$v_1 = v_{j_1} + \ldots + v_{j_t} \mod 2.$$

   Hint: consider the parity check matrix representation of $\mathcal{C}$

   Solution. Because the code is linear, there can't be any nonlinear dependency between codeword components. To see this, pick two codewods $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n)$, let $w = u + v$, and suppose $u_1 = f(u_2, \ldots, u_n)$ and $v_1 = f(v_2, \ldots, v_n)$. By linearity of the code we have that

   $$w_1 = f(w_2, \ldots, w_n) = f(u_2 + v_2, \ldots, u_n + v_n)$$

   must be equal to

   $$u_1 + u_2 = f(u_2, \ldots, u_n) + f(v_2, \ldots, v_n),$$

and similarly for scalar multiplication. Therefore the function f must be linear.

Hence, if $(v_{j_1}, \ldots, v_{j_t})$ is a critical set, then v_1 is a linear function of $v_{j_1}, \ldots, v_{j_t}$. Therefore,

$$v_1 = \sum_{i=1}^{t} a_i v_{j_i}$$

with the a_i's in $\{0, 1\}$. Finally, these a_i's must be equal to one because otherwise the set of t variables would not be critical. □

(b) Deduce that the codeword with zeros everywhere except at positions 1 and $\{j_i, i = 1, \ldots, t\}$ belongs to the dual code $\mathcal{C}^\perp$ of $\mathcal{C}$.

Solution. For any $c \in \mathcal{C}$ we have $\langle c, v' \rangle = v_1 + v_{j_1} + \ldots + v_{j_t} \mod 2 = 0$. So $v' \in \mathcal{C}^\perp$. □

(c) Deduce that any critical set of persons corresponds to a minimal codeword in $\mathcal{C}^\perp$ whose first component is a 1, and such that the persons indices correspond to the components of the non-zero entries of the codeword, after the first component.

Solution. Follows from a. and b. □

(d) We now illustrate the secret sharing scheme through an example. Consider the code whose parity-check matrix is

$$H = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix}$$

It can be checked that positions 1,2,4 form an information set. Fix the first digit of a codeword, v_1, our secret, then choose the second and fourth positions uniformly at random, and compute the full codeword v. Give the digits in positions 2,3,4, and 5 to Alice, Bob, Carol, and David, respectively. What are the critical sets that can recover the secret v_1?

Solution. $\mathcal{C}^\perp = \{00000, 11100, 01011, 10111\}$. The codewords starting with a 1 are 11100 and 10111, and they are minimal. 11100 corresponds to Alice and Bob, and 10111 corresponds to Bob, Carol, and David. □