Discrete Mathematics **Lent 2009**

MA210

Solutions to Exercises 10

(1) Let C be the linear code of length n with check matrix

$$H = [\underbrace{1\ 1\ 1 \ldots\ 1}_{n}].$$

Show that C is the parity check code (defined in lectures).

Solution. To find all codewords $\bar{\boldsymbol{x}} = x_1 x_2 \ldots x_n$ in C, we must solve the equation

$$H\bar{\boldsymbol{x}} = \bar{\mathbf{0}},$$

where

$$\bar{\boldsymbol{x}}' = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

In our case, we obtain

$$x_1 + x_2 + \cdots + x_n = 0,$$

where the addition is modulo 2, that is, in $(\mathbb{Z}_2, +)$. Thus we obtain that

$$x_1 + x_2 + \cdots + x_{n-1} = x_n,$$

so x_n counts (modulo 2) the number of 1's among $x_1, x_2, \ldots, x_{n-1}$, where $x_1, x_2, \ldots, x_{n-1} \in \{0, 1\}$ are arbitrary. This is, however , the definition of the parity check code. □

(2) Let C be the d-repetition code of length n. Show that C is a linear code.

Solution. We have $n = \ell d$ for some natural number ℓ. Each word of C is formed by concatenation of ℓ copies of the same word of length d. Thus, if $\bar{\boldsymbol{x}}, \bar{\boldsymbol{y}} \in C$, then

$$\bar{\boldsymbol{x}} = \underbrace{\bar{\boldsymbol{u}}\bar{\boldsymbol{u}} \ldots \bar{\boldsymbol{u}}}_{\ell} \text{ and } \bar{\boldsymbol{y}} = \underbrace{\bar{\boldsymbol{v}}\bar{\boldsymbol{v}} \ldots \bar{\boldsymbol{v}}}_{\ell}.$$

However, then (we add two words by adding the corresponding bits of both words)

$$\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}} = \underbrace{\bar{\boldsymbol{w}}\bar{\boldsymbol{w}} \ldots \bar{\boldsymbol{w}}}_{\ell},$$

where $\bar{\boldsymbol{w}} = \bar{\boldsymbol{u}} + \bar{\boldsymbol{w}}$ is also a word of length d. So, $\bar{\boldsymbol{x}} + \bar{\boldsymbol{y}}$ is also a word from C and the code is linear. $\square$

(3) (a) Let C be the linear code with check matrix

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

Determine the length n of C, the dimension k of C, the minimum distance d of C. (We then refer to C as an $[n, k, d]$-code.)

(b) The following words were received:

$$11111, \quad 01101, \quad 01100.$$

Decide which of the above are codewords, and correct those which are not codewords, assuming that only one error has been made.

Solution. The length n of C is simply the number of columns of H, i.e., $n = 5$. This is because to find all codewords $\bar{\boldsymbol{x}} = x_1 x_2 \ldots x_n$ in C, we must solve the equation $H\bar{\boldsymbol{x}} = \bar{\mathbf{0}}$, so $\bar{\boldsymbol{x}}$ must have as many entries as H has columns.

From $H\bar{\boldsymbol{x}} = \bar{\mathbf{0}}$ we also obtain that

$$\begin{aligned} x_2 + x_3 &= 0, \\ x_1 + x_4 &= 0, \\ x_1 + x_2 + x_5 &= 0, \end{aligned}$$

that is,

$$\begin{aligned} x_2 &= x_3, \\ x_1 &= x_4, \\ x_1 + x_2 &= x_5. \end{aligned}$$

Hence, the choice of x_1 and x_2 uniquely determines x_3, x_4, x_5 and, consequently, the whole codeword $\bar{\boldsymbol{x}}$. There are 2^2 choices for x_1 and x_2, so the dimension of C is 2. (We know that every linear code must have 2^k codewords and we call k the dimension of C.)

By the theorem from the lecture, to determine the minimum distance in a linear code, we must find the smallest weight of a codeword not equal to $\bar{\mathbf{0}}$. In our case, there are three non-zero codewords: 10011, 01101, and 11110. So, the minimum distance $d = 3$.

For

$$\bar{\boldsymbol{x}} = 11111, \quad \bar{\boldsymbol{y}} = 01101, \quad \bar{\boldsymbol{z}} = 01100,$$

we have

$$H\bar{\boldsymbol{x}}' = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad H\bar{\boldsymbol{y}}' = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad H\bar{\boldsymbol{z}}' = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Hence 01101 is a codeword and $11111, 01100$ are not. Since $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ corresponds to the last column of H and no two column of H are the same, we know that the error occurred at the last bit. (Again, we used a theorem from the lectures.) Thus, we decode 11111 as 11110 and 01100 as 01101. □

(4) Let C be the linear code with check matrix

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

If the word 110110 is received, and at most one error has been made, what was the intended codeword?

Solution. For $\bar{\boldsymbol{x}} = 110110$, we have

$$H\bar{\boldsymbol{x}}' = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix},$$

so, $\bar{\boldsymbol{x}}$ is not a codeword. By the theorem from the lectures, since no column of H is repeated and at most one error has been made, we need to change the bit corresponding to a column that equals to $\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$. This is the second column of H, so we must change the second bit of $\bar{\boldsymbol{x}}$. Thus, we decode 110110 as 100110. □

(5) (a) Let C be a code of length n. Suppose that C is 1-error-correcting. Prove that

$$|C| \leq \frac{2^n}{n+1}.$$

(b) Show there is no 1-error-correcting code of length 5 with $|C| = 6$.

Solution. For a word $\bar{\boldsymbol{x}} \in C$, let $N_1(\bar{\boldsymbol{x}})$ be the set of all possible words $\bar{\boldsymbol{y}}$ with (Hamming) distance at most 1 from $\bar{\boldsymbol{x}}$. For example, if $n = 4$, $\bar{\boldsymbol{x}} = 1101$, then $N_1(\bar{\boldsymbol{x}}) = \{1101, 0101, 1001, 1111, 1100\}$.

We make the following two observations:

(a) for every codeword $\bar{\boldsymbol{x}} \in C$, we have $|N_1(\bar{\boldsymbol{x}})| = n + 1$,

(b) for every two different codewords $\bar{\boldsymbol{x}}, \bar{\boldsymbol{z}} \in C$, we have $N_1(\bar{\boldsymbol{x}}) \cap N_1(\bar{\boldsymbol{z}}) = \emptyset$.

To see (a), we just realize that $\bar{\boldsymbol{x}}$ is at distance 0 from itself and that there are n words at distance 1 from $\bar{\boldsymbol{x}}$, each of them obtained by changing exactly one bit from n possible.

Part (b) follows from the fact that C is one-error-correcting, i.e., no possible word of length n can be within distance one from two codewords (otherwise, we couldn't decode such a word and the code would not be one-error-correcting).

So, what happens if we take the union of $N_1(\bar{\boldsymbol{x}})$ over all codewords $\bar{\boldsymbol{x}} \in C$?

Let $\bar{\boldsymbol{x}}_1, \bar{\boldsymbol{x}}_2, \ldots, \bar{\boldsymbol{x}}_{|C|}$ be all the codewords in C. It follows from (b) and (a) that the size of the union $N_1(\bar{\boldsymbol{x}}_1) \cup N_1(\bar{\boldsymbol{x}}_2) \cup \cdots \cup N_1(\bar{\boldsymbol{x}})$ is $|N_1(\bar{\boldsymbol{x}}_1)| + |N_1(\bar{\boldsymbol{x}}_2)| + \cdots + |N_1(\bar{\boldsymbol{x}})| = |C|(n+1)$ because these sets are pairwise disjoint. On the other hand, the number of words in the union cannot be bigger than the total number of words of length n which is 2^n. Hence, $|C|(n + 1) \leq 2^n$ must hold.

If there was an 1-error-correcting code of length 5 with $|C| = 6$, then $36 = 6(5 + 1) \leq 2^5 = 32$ would have to hold, but this is not possible. $\square$