

Solutions to Exercise Sheet 6

Exercise 1. First note that the codewords of a linear code C form a closed linear subspace, hence the sum (modulo 2) and the difference (modulo 2) of two codewords are also codewords of the code:

$$\forall i, j, i \neq j, \exists l, m : \mathbf{w}_i + \mathbf{w}_j = \mathbf{w}_l, \mathbf{w}_i - \mathbf{w}_j = \mathbf{w}_m \in C. \quad (1)$$

The minimum Hamming distance is by definition

$$d = \min_{i,j} d(\mathbf{w}_i, \mathbf{w}_j), \quad (2)$$

where $d(\mathbf{a}, \mathbf{b})$ is the Hamming distance between \mathbf{a} and \mathbf{b} . By using the linearity (1) and the definition of the Hamming distance we can rewrite (2) as

$$d = \min_{i,j} d(\mathbf{w}_i, \mathbf{w}_j) = \min_{i,j} d(\mathbf{w}_i - \mathbf{w}_j, \mathbf{0}) = \min_k d(\mathbf{w}_k, \mathbf{0}). \quad (3)$$

Hence the minimum Hamming distance of the code is equal to the minimum Hamming distance between any codeword and the word $\mathbf{0}$ and the latter is the definition of the *minimum weight* of the code.

Then note that the conditions of the exercise correspond to the definition of the minimum weight. Hence the statement of the exercise is equivalent to : "the minimum Hamming distance of a code is equal to d if and only if its minimal weight is d ". The last statement is true because we have proven above the equality of the minimum distance and the minimum weight of the code. \square

Exercise 2. Suppose \mathbf{w} is a codeword, then from the definition of the parity matrix H , we have

$$H\mathbf{w} = \mathbf{0}. \quad (4)$$

Vector \mathbf{w} is a column vector with binary entries, we can rewrite the matrix product (4) as

$$\sum_i H_i \cdot w_i = \mathbf{0}, \quad (5)$$

where H_i denotes the i -th column of H and w_i is the i -th entry of \mathbf{w} .

Since the code corrects up to $e - 1$ errors and detects up to e errors, the minimum weight of the code is $d = 2e$, which means that there exists a codeword \mathbf{w}_d such that the number of nonzero entries in \mathbf{w}_d is $2e$. From (5), we can deduce that for \mathbf{w}_d , there are $2e$ columns in H (corresponding to the nonzero entries of \mathbf{w}_d) which are linearly dependent. As $2e$ is the minimum weight (therefore also the minimum number of linearly dependent columns), all sets of $2e - 1$ columns must be linearly independent. And conversely, if all sets of $2e - 1$ columns of H are linearly independent then the minimal weight of the code is at least $2e$. Then the code can detect up to e errors and correct $e - 1$ errors. \square

Exercise 3.

(a) The size of the matrix is given by $n = 6$ and $m = 4$. n corresponds to the size of the codewords. The rank of H is equal to $m = 4$ and corresponds to the number of parity bits. We define a codeword vector \mathbf{w} of components w_i , $i = 1, 2, \dots, n$. The condition $H\mathbf{w} = \mathbf{0}$ can be written in terms of the system of equations:

$$\begin{cases} w_1 + w_5 + w_6 = 0 \\ w_1 + w_2 + w_6 = 0 \\ w_2 + w_3 + w_6 = 0 \\ w_1 + w_4 + w_6 = 0. \end{cases}$$

By expressing w_1 from the first equation and inserting it into the second and the fourth ones we obtain

$$\begin{cases} w_1 = w_5 + w_6 \\ w_5 = w_2 \\ w_2 + w_3 + w_6 = 0 \\ w_5 = w_4. \end{cases}$$

Then we express w_2 from second and first equations and inserting it into the third one obtain

$$\begin{cases} w_1 = w_5 + w_6 \\ w_5 = w_2 \\ w_1 = w_3 \\ w_5 = w_4. \end{cases}$$

We have 4 equations and 6 unknown variables therefore, we can choose arbitrary two of them and obtain a solution. By choosing all four possible combinations for the pair of $\{w_1, w_2\}$ and inserting them into the last system of equations we find the following solutions:

$$\mathbf{w} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

It is easy to see that the minimum distance between the codewords is 3, therefore this code can correct one error.

- (b) To correct one error and detect two, the minimum Hamming distance has to be $d = 4$. This corresponds to having 3 linearly independent columns in H (see Exercise 2). In particular, the last column of H can neither be equal to another column of H nor to a linear combination of any two columns. There are 5 columns and the number of possible different linear combinations of two of them is $\frac{5 \cdot (5-1)}{2} = 10$. Remember that the column with all zeros is also forbidden, so this gives us $5 + 10 + 1 = 16$ different forbidden columns. Because the entries $h_{i,6}$ are bits, only $2^4 = 16$ different combinations are possible, and they are all excluded by the argument above. Therefore the Hamming distance d can not be 4.

Exercise 4.

- (a) The first 3 columns of G_1 are linearly independent and correspond to $k = 3$ bits of information, while the last two columns correspond to $m = 2$ parity bits. There are $2^k = 8$ codewords. The Hamming matrix has 2 rows (number of parity bits) and 5 columns (lengths of the codewords) and contains at least two linearly independent columns. H can be found by solving the equation $H\mathbf{w} = 0$. We can try a solution of the form:

$$\begin{pmatrix} h_{1,1} & h_{1,2} & h_{1,3} & 1 & 0 \\ h_{2,1} & h_{2,2} & h_{2,3} & 0 & 1 \end{pmatrix}$$

Multiplying each row in H by the three codewords from G_1 we get three equations for the three coefficients of each row of H which allow us to find:

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

By calculating all combinations of the three codes words from G_1 we get $\mathbf{w}_4 = \mathbf{w}_1 + \mathbf{w}_2$, $\mathbf{w}_5 = \mathbf{w}_1 + \mathbf{w}_3$, $\mathbf{w}_6 = \mathbf{w}_2 + \mathbf{w}_3$, $\mathbf{w}_7 = \mathbf{w}_1 + \mathbf{w}_2 + \mathbf{w}_3 = \mathbf{w}_4 + \mathbf{w}_3$ and by adding $\mathbf{w}_8 = (0, 0, 0, 0, 0)$ we obtain all the codewords (note that here indexes enumerate the code words and not the elements of a codeword). We can check that all codewords have at least two bits and therefore, the minimum distance of the code is $d = 2$. Hence, this code detects single errors without correcting them.

- (b) In this case, $n = 4$ and $k = 1$. The number of codewords is $2^k = 2$ and the number of parity bits is $m = n - k = 3$. Therefore we can write 3 (last) columns of H as the identity matrix while keeping the first column unknown. By inserting the codeword $\mathbf{w} = (1, 1, 1, 1)$ into equation $H\mathbf{w} = 0$ we find easily :

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

The code has only two codewords:

$$\mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

The Hamming distance is $d = 4$. This is a repetition code which corrects single errors and detects two errors.

Remark: The transmission rate is given by $R = k/n$. We observe that $R_{G_1} = 3/5$, but it does not permit to correct any error (can only detect single errors). In contrary $R_{G_2} = 1/4$ however it can correct single errors and detect double errors. There is a trade off between the transmission rate and the ability of the code to correct errors.