

1. TRUE/FALSE? Prove your answer!

- (a) There exists a binary one-error correcting code of length 9 with 52 codewords.
- (b) There exists a ternary one-error correcting code of length 6 with 10 codewords.

Solution : (a) **FALSE**. Let C be a binary one-error correcting code of length 9. Then it follows from the Sphere packing bound that

$$|C| \leq \frac{m^n}{\sum_{i=0}^t \binom{n}{i} (m-1)^i} = \frac{2^9}{\sum_{i=0}^1 \binom{9}{i} (2-1)^i} = \frac{2^9}{\binom{9}{0} + \binom{9}{1}} = \frac{512}{10} = 51.2$$

So $|C| \leq 51$.

Hence there does not exist a binary one-error correcting code of length 9 with 52 codewords.

(b) **TRUE**. By the Gilbert-Varshamov Bound, there exists a ternary one-error correcting code C of length 6 with

$$|C| \geq \frac{m^n}{\sum_{i=0}^{2t} \binom{n}{i} (m-1)^i} = \frac{3^6}{\sum_{i=0}^2 \binom{6}{i} (3-1)^i} = \frac{729}{\binom{6}{0} + 2\binom{6}{1} + 4\binom{6}{2}} = \frac{729}{73} > 9$$

So $|C| \geq 10$.

Hence there exists a ternary one-error correcting code of length 6 with 10 codewords. □

2. Let C be the binary code with generator matrix $\begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 1 \end{bmatrix}$.

- (a) Find a generator matrix in standard form for C .
- (b) Find a parity check matrix for C .

Solution : (a) We perform elementary row operations on the given generator matrix to find its row echelon form. This will be a generator matrix in standard form.

We begin by swapping R_2 and R_3 :

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

Next we replace R_1 by $R_1 - R_2$:

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

Finally, we replace R_1 by $R_1 - R_3$ and we replace R_2 by $R_2 - R_3$:

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

(b) Once we have a generator matrix in standard form for C , we can use Proposition 3.7 to find a parity check matrix for C . We get

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

3. Let C be the ternary code with parity check matrix $\begin{bmatrix} 2 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 2 & 1 \end{bmatrix}$.

- (a) Find the parameters for C (so n , k and d).
- (b) How many codewords does C have?

Solution : (a) Recall that a parity check matrix for an $[n, k, d]$ -code is an $(n - k) \times n$ -matrix. Hence $n = 5$ and $k = 3$. We use Proposition 3.8 to find d . Note that the given parity check matrix H has no zero-column. Hence every column of H is linearly independent. But

$$C_1 + C_4 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

So the first and fourth column of H are linearly dependent. Hence $d = 2$.

(b) In general, if C is a linear code over \mathbb{F} of dimension k then $|C| = |\mathbb{F}^k| = |\mathbb{F}|^k$. So here we get that

$$|C| = 3^3 = 27 \quad \square$$

4. Let C be the binary code word generator matrix $\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$.

- (a) Set up a standard array for C .
- (b) Write down the coset leaders of each coset.
- (c) Use your standard array to decode the following words :
 - 1111
 - 1110
 - 1010

Solution : (a) First, we find all the codewords. Recall that the codewords are all the linear combinations of the rows of the generator matrix. Hence

$$C = \{0000, 1001, 0110, 1111\}$$

This is the first row of the standard array.

Next, we fill the standard array : we find a word \mathbf{x} that is not in the array yet and add the coset $\mathbf{x} + C$ as a new row to the standard array. We make sure that a coset leader is in the first column. Let's say we start with $\mathbf{x} = 1000$. Then we get

$$\begin{array}{cccc} 0000 & 1001 & 0110 & 1111 \\ 1000 & 0001 & 1110 & 0111 \end{array}$$

We continue with $\mathbf{x} = 0100$ and find

$$\begin{array}{cccc} 0000 & 1001 & 0110 & 1111 \\ 1000 & 0001 & 1110 & 0111 \\ 0100 & 1101 & 0010 & 1011 \end{array}$$

Finally, we pick any remaining word, say $\mathbf{x} = 1100$ and get the following standard array :

0000 1001 0110 1111
 1000 0001 1110 0111
 0100 1101 0010 1011
 1100 0101 1010 0011

It's important to realize that this array is not unique. We could have picked 0001 as choice for a coset leader. Then the second row would have been

0001 1000 0111 1110

(b) Using our standard array, we easily get

Coset	Coset Leaders
{0000, 1001, 0110, 1111}	0000
{1000, 0001, 1110, 0111}	1000, 0001
{0100, 1101, 0010, 1011}	0100, 0010
{1100, 0101, 1010, 0011}	1100, 0101, 1010, 0011

Note that this is independent of the standard array we choose.

(c) Decoding using a standard array is quite simple : we look up the received word in the standard array and decode as the codeword (in the first row) that is in the same column as the received word. This clearly depends on the standard array we choose. We get

We decode 1111 as 1111.
 We decode 1110 as 0110.
 We decode 1010 as 0110.

5. Let C be the binary code with parity check matrix $\begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix}$.

- (a) Set up a syndrome table for C .
- (b) Use your syndrome table to decode the word 11010.

Solution : We pick words of small weight (first weight zero, then weight one, then weight two,...) and evaluate their syndrome until we get every single syndrome exactly once. We find

coset leader	syndrome
00000	000
10000	100
01000	110
00100	111
00010	001
00001	101
00101	010
10100	011

Note that the first six rows are unique since these cosets have exactly one coset leader. The last two rows are not unique since those cosets have more than one coset leader. We could have chosen 11000 instead of 00101 in the seventh row.

(b) First, we calculate

$$\text{syn}(11010) = \begin{bmatrix} 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = 011$$

Next, we look up this syndrome in the syndrome table. We decode by subtracting the coset leader from the received word. So we decode 11010 as

$$11010 - 10100 = 01110$$

Again, this depends on which coset leader we choose for our syndrome table. □

6. Decode the word

$$01000000000011$$

using the binary Hamming code $Ham(2, 4)$.

Solution : Recall that we have a simple decoding algorithm for $Ham(2, 4)$.

First, we calculate $\text{syn}(01000000000011)$. For $1 \leq i \leq 15$, let C_i be the i -th column of the special parity check matrix for $Ham(2, r)$ (so the i -th column is the binary expansion of the integer i). Then we get that (writing syndromes as a four-tuple)

$$\text{syn}(01000000000011) = C_2 + C_{14} + C_{15} = 0010 + 1110 + 1111 = 0011$$

Since this is the binary expansion of three, we decode by changing the third digit in the received word. So we decode 01000000000011 as

$$01100000000011$$

□

7. Prove there does not exist a binary two-error correcting code of length 8 with 5 codewords.

Proof : Suppose that there exists a binary two-error correcting code C of length 8 with 5 codewords. Then there are at least three codewords with the same digit in the first position, say \mathbf{x} , \mathbf{y} and \mathbf{z} . At least two of these five codewords have the same digit in the second position, say \mathbf{x} and \mathbf{y} .

Since C is two-error correcting, we know that $d(C) \geq 5$. So \mathbf{x} and \mathbf{y} are different in at least five positions. It turns out that \mathbf{z} will agree with either \mathbf{x} or \mathbf{y} in at least three of those five positions.

WLOG, we may assume that \mathbf{x} , \mathbf{y} and \mathbf{z} are of the following form ($z_2, \dots, z_8, x_8, y_8$ are unknown digits while \square and $*$ are specific digits) :

$$\begin{array}{rcccccccc} \mathbf{x} & = & \square & * & 0 & 0 & 0 & 0 & x_8 \\ \mathbf{y} & = & \square & * & 1 & 1 & 1 & 1 & y_8 \\ \mathbf{z} & = & \square & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 & z_8 \end{array}$$

So in the word $z_3z_4z_5z_6z_7$ the same symbol (say the symbol one) must show up at least three times. But then $d(\mathbf{y}, \mathbf{z}) \leq 4$, a contradiction.

Hence there does not exist a binary two-error correcting code of length 8 with 5 codewords. □

A more rigorous approach is as follows (instead of the part starting with WLOG).

Put $\Omega = \{3 \leq i \leq 8 : x_i \neq y_i\}$. Then $|\Omega| \geq 5$. Let $i \in \Omega$. Since $x_i, y_i, z_i \in \{0, 1\}$ and $x_i \neq y_i$, we conclude that either $z_i = x_i$ or $z_i = y_i$. Put $\Omega_{\mathbf{x}} = \{i \in \Omega : z_i = x_i\}$ and $\Omega_{\mathbf{y}} = \{i \in \Omega : z_i = y_i\}$. Then $\Omega = \Omega_{\mathbf{x}} \cup \Omega_{\mathbf{y}}$. Hence either $|\Omega_{\mathbf{x}}| \geq 3$ or $|\Omega_{\mathbf{y}}| \geq 3$. If $|\Omega_{\mathbf{x}}| \geq 3$ then $d_H(\mathbf{z}, \mathbf{x}) \leq 4$; if $|\Omega_{\mathbf{y}}| \geq 3$ then $d_H(\mathbf{z}, \mathbf{y}) \leq 4$. Either way we get a contradiction since $d(C) \geq 5$.

8. Let \mathbb{F} be a field and $n \in \mathbb{N}$. Use the Triangle Inequality (regarding the Hamming distance) to prove the following :

(a) $w(\mathbf{a} - \mathbf{b}) \leq w(\mathbf{a}) + w(\mathbf{b})$ for all $\mathbf{a}, \mathbf{b} \in \mathbb{F}^n$.

(b) $w(\mathbf{x} - \mathbf{y}) \geq |w(\mathbf{x}) - w(\mathbf{y})|$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$.

Proof : Recall that

$$w(\mathbf{x} - \mathbf{y}) = d_H(\mathbf{x}, \mathbf{y}) \quad \text{for all } \mathbf{x}, \mathbf{y} \in \mathbb{F}^n$$

In particular, we get that

$$w(\mathbf{x}) = w(\mathbf{x} - \mathbf{0}) = d_H(\mathbf{x}, \mathbf{0}) \quad \text{for all } \mathbf{x} \in \mathbb{F}^n$$

(a) Let $\mathbf{a}, \mathbf{b} \in \mathbb{F}^n$. By the Triangle Inequality, we get that

$$d_H(\mathbf{a}, \mathbf{b}) \leq d_H(\mathbf{a}, \mathbf{0}) + d_H(\mathbf{0}, \mathbf{b})$$

Hence

$$w(\mathbf{a} - \mathbf{b}) \leq w(\mathbf{a}) + w(\mathbf{b})$$

(b) Let $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$. By (a), we get that

$$w(\mathbf{y}) = w(\mathbf{x} - (\mathbf{x} - \mathbf{y})) \leq w(\mathbf{x}) + w(\mathbf{x} - \mathbf{y})$$

Hence

$$w(\mathbf{x} - \mathbf{y}) \geq w(\mathbf{y}) - w(\mathbf{x})$$

Similarly, we find that

$$w(\mathbf{x}) = w(\mathbf{y} - (\mathbf{y} - \mathbf{x})) \leq w(\mathbf{y}) + w(\mathbf{y} - \mathbf{x}) = w(\mathbf{y}) + w(\mathbf{x} - \mathbf{y})$$

and so

$$w(\mathbf{x} - \mathbf{y}) \geq w(\mathbf{x}) - w(\mathbf{y})$$

Since $|w(\mathbf{x}) - w(\mathbf{y})| \in \{w(\mathbf{x}) - w(\mathbf{y}), w(\mathbf{y}) - w(\mathbf{x})\}$, we get that $w(\mathbf{x} - \mathbf{y}) \geq |w(\mathbf{x}) - w(\mathbf{y})|$. □

9. Let C be a linear t -error correcting code of length n over the field \mathbb{F} and $\mathbf{x} \in \mathbb{F}^n$ with $w(\mathbf{x}) \leq t$. Prove that \mathbf{x} is the unique coset leader of its coset.

Proof : First, we prove the following claim :

If \mathbf{y} is a coset leader of the coset containing \mathbf{x} then $\mathbf{y} = \mathbf{x}$.

Indeed, let \mathbf{y} be a coset leader of the coset containing \mathbf{x} . Then $w(\mathbf{y}) \leq t$ and $\mathbf{y} \in \mathbf{x} + C$. Hence $\mathbf{y} = \mathbf{x} + \mathbf{c}$ for some $\mathbf{c} \in C$. Using Ex#8a, we get that

$$w(\mathbf{c}) = w(\mathbf{y} - \mathbf{x}) \leq w(\mathbf{y}) + w(\mathbf{x}) \leq t + t = 2t < 2t + 1 \leq d(C) = w(C)$$

So $\mathbf{c} \in C$ and $w(\mathbf{c}) < w(C)$. This is only possible if $\mathbf{c} = \mathbf{0}$. Thus $\mathbf{y} = \mathbf{x} + \mathbf{c} = \mathbf{x} + \mathbf{0} = \mathbf{x}$, which proves the claim.

It follows immediately from this claim that the coset of x has at most one coset leader (namely x). Since every coset has a coset leader, it follows from the claim that \mathbf{x} is indeed a coset leader of the coset containing \mathbf{x} . So \mathbf{x} is the unique coset leader of the coset containing \mathbf{x} . □