Solutions to problem set 4.

Mathematics 342, Term 2, 2018. Instructor: Reichstein.

(1) Suppose C is a binary linear code of length 6 with generator matrix

$$G = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

Find the minimal distance of C.

Solution: The minimal distance of a linear code is equal to the minimal weight of a non-zero word in the code. The elements of C are, by definition, the linear combinations

$$a_1\mathbf{R}_1 + a_2\mathbf{R}_2 + a_3\mathbf{R}_3,$$

where $\mathbf{R_1}$, $\mathbf{R}_2$ and $\mathbf{R}_3$ are the rows of G, and a_1, a_2, a_3 are 0 or 1. There are 8 words in C, corresponding to the 8 choices of (a_1, a_2, a_3); the seven non-zero words are

$$\begin{array}{cccccc} (1 & 1 & 1 & 1 & 1 & 1) \\ (0 & 1 & 1 & 1 & 1 & 0) \\ (1 & 0 & 1 & 0 & 1 & 1) \\ (1 & 0 & 0 & 0 & 0 & 1) \\ (0 & 1 & 0 & 1 & 0 & 0) \\ (1 & 1 & 0 & 1 & 0 & 1) \\ (0 & 0 & 1 & 0 & 1 & 0) \end{array}$$

The smallest weight of any of these words is 2; thus $d(C) = 2$.

(2) Give an example of a linear code C which does not have a generator matrix in standard form.

Solution: Let C be the code of length 3 with parity check matrix

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

That is, C consists of words of the form $(0, a, b)$, where a and b are arbitrary elements of F_q. Then C cannot have a generator matrix in standard form. Indeed, $\dim(C) = 2$, so a generator matrix in standard form would look like

$$G = \begin{pmatrix} 1 & 0 & x \\ 0 & 1 & y \end{pmatrix}.$$

But there is no word of the form $(1, 0, x)$ in C.

(3) Show that if a linear code C has a generator matrix G in standard form, such a generator matrix is unique. In other words, if G' is a generator matrix in standard form for C then $G' = G$.

Solution: Suppose C is a code, and G and G' are generator matrices in standard form. Let the rows of G be $g_1, \ldots, g_k$ and the rows of G' be $g'_1, \ldots, g'_k$. Here $k = \dim(C)$. We want to show that $g_i = g'_i$ for every $i = 1, \ldots, k$.

Since G is standard form, we have:

$$\begin{aligned} g_1 &= 100\cdots00**\cdots** \\ g_2 &= 010\cdots00**\cdots** \\ &\vdots \\ g_r &= 000\cdots01**\cdots** \end{aligned}$$

(the *** represent unknown values).

Since G is a generator matrix for C and g' is a codeword in C, we can write

$$g'_i = \lambda_1 g_1 + \cdots + \lambda_k g_k,$$

where each λ_i is an element of F_q. Equating the first k entries on both sides, we see that $\lambda_i = 1$ and $\lambda_j = 0$ for every $j \neq i$. In other words, $g'_i = g_i$. This completes the proof.

(4) Let $q = 5$ and consider the q-ary linear code C of length 4, consisting of all words (a_1, a_2, a_3, a_4) such that $a_1 + 2a_2 + 3a_3 + 4a_4 \equiv 0 \pmod 5$. Show that this is a linear code. Find a generator matrix in standard form for this code.

Solution: Let $\mathbf{a} = (a_1, a_2, a_3, a_4) \in C$, $\mathbf{b} = (b_1, b_2, b_3, b_4) \in C$, and let $\lambda \in F_q$. To show that C is linear we must show that $\mathbf{a} + \mathbf{b} \in C$ and $\lambda\mathbf{a} \in C$. Since $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2, a_3 + b_3, a_4 + b_4)$, and

$$\begin{aligned} (a_1 + b_1) + 2(a_2 + b_2) + 3(a_3 + b_3) + 4(a_4 + b_4) &= \\ (a_1 + 2a_2 + 3a_3 + 4a_4) + (b_1 + 2b_2 + 3b_3 + 4b_4) &= \\ 0 + 0 = 0. \end{aligned}$$

we have that $\mathbf{a} + \mathbf{b} \in C$. Secondly, $\lambda\mathbf{a} = (\lambda a_1, \lambda a_2, \lambda a_3, \lambda a_4)$, and

$$(\lambda a_1) + 2(\lambda a_2) + 3(\lambda a_3) + 4(\lambda a_4) = \lambda(a_1 + 2a_2 + 3a_3 + 4a_4) = \lambda(0) = 0$$

Thus $\lambda a \in C$ also, and thus C is a linear code.

Note that by the definition of C,

$$H = \begin{pmatrix} 1 & 2 & 3 & 4 \end{pmatrix}$$

is a parity check matrix. Viewing H s a generator matrix for $C^{\perp}$ in standard form, we construct a parity check matrix for $C^{\perp}$ or equivalently, a generator matrix for C as follows:

$$G = \begin{pmatrix} -2 & 1 & 0 & 0 \\ -3 & 0 & 1 & 0 \\ -4 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}.$$

Interchanging the first and the third row and using row operations, we obtain a generator matrix in standard form

$$G_{stand} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}.$$

(5) Construct a standard array for the binary linear code with generator matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}.$$

Use this array to decode 01010 and 11101.

Solution: Once again, words in C are linear combinations

$$a_1\mathbf{R}_1 + a_2\mathbf{R}_2 + a_3\mathbf{R}_3,$$

where $\mathbf{R_1}$, $\mathbf{R}_2$ and $\mathbf{R}_3$ are the rows of G and a_1, a_2, a_3 are 0 or 1. Going through the 8 possible values of (a_1, a_2, a_3), we see that

$$C = \{00000, 10011, 01001, 00111, 11010, 10100, 01110, 11101\}.$$

C has $2^{5-3} = 4$ cosets. One possible choice of coset leaders leads to the following standard array:

00000	10011	01001	00111	11010	10100	01110	**11101**
10000	00011	11001	10111	**01010**	00100	11110	01101
01000	11011	00001	01111	10010	11100	00110	10101
00010	10001	01011	00101	11000	10110	01100	11111

Finding 01010 in the above table we decode it as 11010. (If the coset leader 00100 is chosen in the second coset instead of 10000 then the answer changes to 01110.)

Finding 11101 in the above table we decode it as 11101.

(6) Construct a standard array for the ternary linear code with generator matrix

$$G = \begin{pmatrix} 1 & 0 & 1 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

Using this array to decode 1211 and 2102.

Solution: Once again, words in C are linear combinations

$$a_1\mathbf{R}_1 + a_2\mathbf{R}_2,$$

where $\mathbf{R_1}$ and $\mathbf{R}_2$ are the rows of G and a_1, a_2 are 0, 1 or 2. Going through the 9 possible values of (a_1, a_2), we see that

$$C = \{0000, 1012, 0111, 2021, 0222, 1120, 1201, 2210, 2102\}.$$

The standard array is:

0000	1012	0111	2021	0222	1120	1201	2210	**2102**
1000	2012	1111	0021	1222	2120	2201	0210	0102
0100	1112	0211	2121	0022	1220	1001	2010	2202
0010	1022	0121	2001	0202	1100	**1211**	2220	2112
0001	1010	0112	2022	0220	1121	1202	2211	2100
2000	0012	2111	1021	2222	0120	0201	1210	1102
0200	1212	0011	2221	0122	1020	1101	2110	2002
0020	1002	0101	2011	0212	1110	1221	2200	2122
0002	1001	0100	2010	0211	1112	1220	2202	2121

Finding 1211 in the above table we decode it as 1201.
Finding 2102 in the above table we decode it as 2102.

(7) Let $q = 11$ and C be the q-ary linear code with parity check matrix

$$H = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{pmatrix}.$$

(a) What is the minimal distance of this code?
(Recall that Problem 5 on Midterm 1 asked you to show that $d(C) \leq 3$. This question is asking for the exact value of $d(C)$.)

(b) Suppose a word $\mathbf{x}$, transmitted using this code, is received as

$$\mathbf{y} = (1, 1, 0, 0, 0, 0, 0, 0, 0, 0).$$

Assuming at most one error could occur in transmission, find $\mathbf{x}$. Explain your answer.

Solution: (a) Answer: $d(C) = 3$. To prove this, we need to show:

(i) no two columns of H are linearly dependent.

(ii) H has three linearly dependent columns.

Denote the ith column of H by K_i. That is,

$$K_i = \begin{pmatrix} 1 \\ i \end{pmatrix}.$$

To prove (i), assume $aK_i + bK_j = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for some $a, b \in F_{11}$ and some $i \neq j$. Our goal is to show that $a = b = 0$. Indeed,

$$a \begin{pmatrix} 1 \\ i \end{pmatrix} + b \begin{pmatrix} 1 \\ j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

translates to $a + b = 0$ and $ia + jb = 0$. From the first equation, $b = -a$, and from the second $a(i - j) = 0$ in F_{11}. Since F_{11} is a field, and $i \neq j$, we conclude that $a = 0$ and thus $b = 0$ in F_{11}, as desired.

To prove (ii), note that $K_1 - 2K_2 + K_3 = 0$, so K_1, K_2 and K_3 are linearly dependent.

(b) Assume the error vector is of the form $\mathbf{e} = (0, \ldots, 0, m, 0, \ldots, 0)$, where m s in position p. That is a single error of magnitude m has occurred, in position p. Then $\mathbf{e}$ and $\mathbf{y}$ have the same syndrome.

$S(\mathbf{y}) = \mathbf{y} \cdot H^T = (2, 3)$, and $S(\mathbf{e}) = \mathbf{e} \cdot H^T = (m, mp)$.

Thus $m = 2$ and $mp = 3$ in F_{11}. Solving for p, we obtain, $2p = 3$ in F_{11} of $p = 3 \cdot 2^{-1} = 3 \cdot 6 = 7$ in F_{11}.

Thus $\mathbf{e} = (0, 0, 0, 0, 0, 0, 2, 0, 0, 0)$, and we decode $\mathbf{y}$ as

$\mathbf{x} = \mathbf{y} - \mathbf{e} = (1, 1, 0, 0, 0, 0, 9, 0, 0, 0)$.

(8) Let C be the binary linear code with parity check matrix

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

How many cosets does C have? Find the coset leaders by going through words of small weight and calculating their syndromes until you exhaust all possible syndromes. Decode received words (a) 1010101, (b) 1110001, (c) 0111111.

Solution: Since an $[n, k]$-code has an $(n - k) \times n$ parity check matrix, we know that $\dim(C) = k = 4$ and the length of C is $n = 7$. The standard array has $2^k = 16$ columns and $2^n = 128$ entries, thus it must have $128/16 = 8$ rows (i.e. C has 8 cosets).

Now let $\mathbf{e}_j$ be the error vector with 1 is position j and 0 in every other position. Then the syndrome $S(\mathbf{e}_j)$ is just the transpose of the jth column of H. We note that all the columns are distinct in H and that $\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$ is not among them. This means that $d(C) \geqslant 3$. Thus the 8 words of weight $\leqslant 1$, namely $\mathbf{0}$ and $\mathbf{e}_1, \ldots, \mathbf{e}_8$

are coset leaders, and each coset has at most one of them. Since there are 8 cosets, each coset has exactly one of these words.

(a) $S(1010101) = (100)$ is the transpose of the first column of H. That is, $S(\mathbf{e}_1) = (100)$, so assume that the error vector is $\mathbf{e}_1$ and decode (1010101) as $(1010101) - (1000000) = (0010101)$.

(b) Similarly, $S(1110001) = 000$ and $S(0000000) = 000$ so assume that no error occurred in tranmission and decode as (1110001) as $(1110001) - (0000000) = (1110001)$.

(c) $S(0111111) = (100)$ and $S(1000000) = (100)$ so we decode (0111111) as $(0111111) - (1000000) = (1111111)$.