Coding Theory 2019 Answers

Questions are SEEN or similar to seen unless otherwise stated.

NB Answers given here are sometimes expressed in brief form, but exam answers must form sentences.

1. (a) A set $V = \{v_1, v_2, ..., v_n\}$ of vectors is linearly independent if the only solution $(a_1, a_2, ..., a_n) \in F^n$ to the equation

$$\sum_i a_i v_i = 0$$

is $(a_1, a_2, ..., a_n) = (0, 0, ..., 0)$. (1 marks)

(b) $|\mathcal{M}_{n,m}(F)| = q^{nm}$ (1 marks)

(c) M generator if rows linearly independent (which implies $n \leq m$), and F finite.
Then M generates a $|F|$-ary [m,n]-code (dimension n, length m code over F). (2 marks)

(d) Many examples are possible. A simple one is

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

(1 marks)

(e)

$$C_1 = \{00000, 10100, 01100, 11000\}$$

(1 marks)

(f) S_1 not closed under +, so not a linear code.
S_2 is closed under linear combinations, so linear code.
S_3 is closed under linear combinations, so linear code.
S_4 is not closed under +. (4 marks)

(g) Let F denote the symbol set, so that each $x \in C$ is a string of 'entries' that are symbols from F. Then minimum weight

$$w(C) = min\{w(x)|x \in C \setminus \{\underline{0}\}\}$$

where $w(x)$ is weight of x (the number of entries in x different from the symbol $0 \in F$, the symbol set), and $\underline{0}$ denotes the zero vector.

Prove that, for a linear code, the minimum distance $d(C)$ is equal to $w(C)$.:
First recall that $d(x,y) = |\{i \ : \ x_i \neq y_i\}|$. Thus $d(x,y) = d(x-y,0)$; and by the above definition of $w(x)$ we have $d(x,y) = d(x-y,0) = w(x-y)$. Since C is linear we have $x - y \in C$; and since $d(C)$ is the minimum of $d(x,y)$ among all distinct codewords we have $d(C) = w(C)$ by the above identities. $\square$ (5 marks)

(h) Define $C^\perp$, the dual code to a linear code C.
$C \subset F_q^n$, $C^\perp = \{v \in F_q^n | v.x = 0 \forall x \in C\}$ where $v.x = \sum_i v_i x_i$ (over F).

Prove that $C^\perp$ is also a linear code:
Suppose $v, v' \in C^\perp$. Then $(\alpha v + \alpha' v').x = \alpha v.x + \alpha' v'.x = 0 + 0 = 0$ and hence the code is closed under linear combinations. (5 marks)

(i) Compute the dual of C_1 above, and hence or otherwise determine if it is self-dual.
Ignoring the last two digits (which are always zero in C_1) for now, we have
$(x,y,z).(1,0,1) = x + z = 0$
$(x,y,z).(0,1,1) = y + z = 0$
$(x,y,z).(1,1,0) = x + y = 0$
These imply $x = y = z$. The last two digits in the dual are not constrained, so $C^\perp = \{000, 111\} \times \mathbb{Z}_2^2$ (in the obvious notation)

(any equivalent, such as giving a PCM, is acceptable).
So $C_1^{\perp} \neq C_1$. So C_1 is not self-dual. (5 marks)

2. (a) The natural action of the symmetric group S_n on a space V^n permuting the standard ordered basis commutes with the Hamming distance function, and is of course invertible, so gives equivalent codes. The given matrices are related by, specifically $(45) \in S_5$. (Other answers acceptable.) (1 marks)

   (b) These matrices are related by row operations, so generate the same subspace. (1 marks)

   (c) Let $C \subset \mathbb{Z}_7^5$ be the linear code with generator matrix

$$G = \begin{pmatrix} 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 0 & 3 & 4 \\ 0 & 0 & 1 & 5 & 6 \end{pmatrix}$$

   i. Write down a parity check matrix H for C.

$$H = \begin{pmatrix} -1 & -3 & -5 & 1 & 0 \\ -2 & -4 & -6 & 0 & 1 \end{pmatrix}$$

   (1 marks)

   ii. Compute the matrix $G.H^t$ (where H^t is the transpose of H). Interpret your result.
   $GH^t = 0$ (show calculations)
   Columns of H^t are rows of H so GH^t assembles the various inner product calculations for C and $^\perp$, which must all be zero by definition. (1 marks)

   iii. Each vector has $w(v) = 3$. This tells us that $d(C)$ is at most 3. (1 marks)

   iv. Show that $d(C) = 3$.
   H has no zero or parallel columns, but $w(G_3) = 3$ so $d(C) \leq 3$.
   So $d(C) = 3$. (2 marks)

   v. How many of the coset leaders of C have weight 1?
   There are 7^2 coset leaders. There are $5 \times 6 = 30$ weight 1

vectors, none of which lie in C, and no distint pair of which have $x - y \in C$. So number = 30.
(full marks for any legitimate argument with right final answer) (3 marks)

vi. Codeword x is transmitted down a noisy channel, so that $y = 11254$ is received, with exactly one error having occured. What was the transmitted codeword x?

$Hy^t = \begin{pmatrix} -1 & -3 & -5 & 1 & 0 \\ -2 & -4 & -6 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 2 \\ 5 \\ 4 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$. Now find

coset leader: $H \begin{pmatrix} 0 \\ 0 \\ 0 \\ 5 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \end{pmatrix}$

(3 marks)

so $x = 11254 - 00050 = 11204$. (2 marks)

(d) Add/Mult. tables for $\mathbb{Z}_4$: BOOKWORK.
$\mathbb{Z}_4$ fails to form a field since there are not enough multiplicative inverses. (2 marks)

(e) Explain a way to construct a field of order 4.
ANSWER: Consider degree 2 polynomials over $\mathbb{Z}_2$. Quadratics are x^2+1, x^2+x+1, x^2, x^2+x. Only x^2+x+1 irreducible (check it is!), so extend $\mathbb{Z}_2$ by x obeying $x^2 + x + 1 = 0$. (4 marks)

Write down the addition and multiplication tables for this field.

$+$	0	1	x	$1+x$
0	0	1	x	$1+x$
1	1	0	$1+x$	x
x	x	$1+x$	0	1
$1+x$	$1+x$	x	1	0

$\times$	0	1	x	$1+x$
0	0	0	0	0
1	0	1	x	$1+x$
x	0	x	$1+x$	1
$1+x$	0	$1+x$	1	x

(4 marks)

3. (a) The Cartesian product of two sets, A, B, say, is the set of ordered pairs (a, b) with $a \in A$ and $b \in B$.
E.g. Σ_q^2: ordered 2-tuples from Σ_q. (1 marks)

(b) Associative: let us write $\mu(a, b) = ab$. Then $a(bc) = (ab)c$.
$(\mathbb{Z}, +)$ is associative.
$(\mathbb{Z}, -)$ is not. (E.g. $1 - (2 - 3) = 0$, $(1 - 2) - 3 = -4$.) (1 marks)

(c) Σ_q^n: ordered n-tuples from Σ_q. (1 marks)

$\{0, 1\}^3 = \{000, ..., 111\}$ (using $(i, j, k) \mapsto ijk$). (1 marks)

For the notation to be unambiguous, it is sufficient that each symbol is connected. For example sequences built from $\{0, 1\}$ are unambiguous but sequences built from $\{0, 10, 1\}$ are not. (Other answers acceptable.) (1 marks)

(d) $|P(\Sigma_q^n)| = 2^{(q^n)}$ (also accept count excluding empty set). (1 marks)

(e) i. Hamming distance $d(x, y) = \#\{i | x_i \neq y_i\}$ (1 marks)

ii. Weight 0/1: Vectors of form $(0, 0, ..., 0, X, 0, ..., 0)$. There are $n \times (q - 1) + 1$ of these.
Weight 2: Vectors of form $(0, 0, ..., 0, X, 0, .., Y, 0, ..., 0)$ with $X, Y \neq 0$. There are $\frac{n(n-1)}{2} \times (q - 1)^2$ of these. (1 marks)

iii. minimum distance $d(C) = min\{d(x, y) | x, y \in C,\ x \neq y\}$ (2 marks)

iv. $d(C) \geq 15$. (1 marks)

v. ball-packing bound on the size M of a q-ary (n, M, d)-code C:

$$M \sum_{r=0}^{t} \binom{n}{r} (q - 1)^r \leq q^n$$

where t such that $d \geq 2t + 1$. (2 marks)

(f) For each of the following triples (n, M, d) construct, if possible, a binary (n, M, d)-code:

$$(X, 2, X) \qquad (3, 8, 1) \qquad (4, 8, 2) \qquad (8, M, 3)$$

(for given values of X, M). If no such code exists, then prove it, stating any theorems used.
ANSWER: (X,2,X): $\{000000...0, 111111...1\}$.
(3,8,1): $\{000, 001, 010, 011, 100, 101, 110, 111\}$
(4,8,2): $\{0000, 0011, 0101, 0110, 1001, 1010, 1100, 1111\}$
(8,M,3): fails the BP bound if:

$$M(1 + 8) = 9 * M \not\leq 2^8 = 256$$

so fails for $M > 256/9$ (e.g. $M > 28$). (12 marks)

(/25 marks)

4. (a) $H(221000)^t = 0$ so $221000 \in C$. (1 marks)

(b) Confirm that G is a generator matrix for C:
1. rows linearly independent
2. $GH^t = ...calculation... = 0$
3. # rows = 6-3

All ok. (2 marks)

(c) Compute the encoded form of the letter U (or D):
(another example: E is 5th letter, so rep is 012 and encoding is 220112)
U is represented by 210 and its encoding is 202100
D is represented by 011 and its encoding is 020121 (3 marks)

(d) What is $d(C)$?
From words in G (say) we have $d(C) \leq 3$, but no column of H is zero or "parallel" to another, so $d(C) = 3$. (2 marks)

This implies no y with $w(y) = 1$ or 2 lies in C.
Now suppose x, y of wt 1 lie in $C + x$. Then $y - x$ lies in C. But $w(y - x) \leq 2$, so wt.1 vectors lie in distinct cosets.
There are $6 \times 2 = 12$ of them. Thus there are 12+1 vectors within 1 of 000000. Syndromes:
$S(000000) = 000$
$S(100000) = 100$
$S(200000) = 200$
$S(010000) = 010$
$S(001000) = 110$
$S(000100) = 210$
$S(000010) = 001$
etc (can use linearity)
$S(000001) = 101$
$S(000002) = 202$

(8 marks)

(e) Now:

200021.$H^t = 000$ so we have 001, which gives A;
112000.$H^t = 000$ so we have 200, which gives R;
112100.$H^t = 210$ so we have 112100-000100, which gives R;
220112.$H^t = 000$ so we have 012, which gives E;

...and so on (show all working; write in sentences), until

$$(022021)\begin{pmatrix} 100 \\ 010 \\ 110 \\ 200 \\ 001 \\ 101 \end{pmatrix} = (010) = S(010000)$$

giving 201, and hence S;

and so on, until:

$$(012022)\begin{pmatrix} 100 \\ 010 \\ 110 \\ 200 \\ 001 \\ 101 \end{pmatrix} = (101) = S(000001)$$

which thus corrects to 012022-000001=012021, giving 201, and hence S again.

Continuing until finally, 212012, again giving T.

Altogether we get:

ARREST THE STUDENT

(9 marks)

5. (a) Dim of ambient space is 4 so dim of self-dual code is 2 (since if dim is d then dim obeys $4-d=d$ by self-duality in Z_{11}^4, so $4-2=2$). Thus generator matrix takes form

$$\begin{pmatrix} 1 & 0 & a & b \\ 0 & 1 & c & d \end{pmatrix}$$

where for self-duality we have $1+a^2+b^2 \equiv 0$; $1+c^2+d^2 \equiv 0$ and $ac+bd \equiv 0$. Noting that $1+1+9=11$, try $a,b=1,3$ and $c,d=-3,1$. (Other correct answers exist.) (2 marks)

(b) $P(e=000...0)=(1-p)^n$ (1 marks)

If the actual error e associated to y is a coset leader then the check returns e and the decode method $y \mapsto y-e$ becomes $e+x \mapsto x$, which therefore works; and otherwise not. We therefore require to compute the probability that e is a coset leader. (1 marks)

(c) a standard array for $C=\{0000,1010,0101,1111\}$:
0000 1010 0101 1111
1000 0010 1101 0111
0100 1110 0001 1011
1100 0110 1001 0011
(any correctly formed standard array is acceptable)
(5 marks)

(d) Decode the received message 1101 using your array:
(IF the coset leaders are as above then)
the coset leader is 1000, so 1101-1000=0101 is the decoding.
(2 marks)

(e) Code C is transmitted down a binary symmetric channel with symbol error probability $p=0.01$, with the received vectors being decoded by the coset decoding method. ...Calculate $P_{err}(C)$, the word error probability of the code; and $P_{undetec}(C)$, the probability of there being an undetected error in a transmitted word.

ANSWER:
$P(e = 0000) = (1-p)^4 = .9606$ (4 sig figs)

(2 marks)

The next answer depends a little on the coset leaders in your standard array. It may simply be enough to compare with (a) and (b) above. Otherwise we may use the hint and then compare. That is, $P_{corr}(C)$ takes the given form since an error (including the null error) is corrected if it takes any of the forms appearing in the sum (or possibly takes any of another set of forms — generally, the coset leaders — but these have the same collection of probabilities as the given vectors); and is not corrected otherwise. Thus

$$P_{corr}(C) =$$

$$P(e = 0000) + P(e = 1000) + P(e = 0100) + P(e = 1100) =$$

$$(1-p)^4 + 2p(1-p)^3 + p^2(1-p)^2 = 0.9801$$

$$P_{err}(C) = 1 - P_{corr}(C) = 0.0199$$

(3 marks)

For there to be an undetected error in the transmitted word the received word would have to be in C, but in error. That means both transmitted word x and received word y are in C (and are different), so the error $e = x - y$ is also in C (and of course is not the zero word). Thus

$$P_{undetec}(C)|_{p=0.01} = P(e = 1010) + P(e = 0101) + P(e = 1111)$$

$$= 2 \times (0.01)^2(0.99)^2 + (0.01)^4 = 0.00019603$$

(3 marks)

(f) Code C is again transmitted down a binary symmetric channel with symbol error probability $p = 0.01$, but is now used only for error detection. If an error is detected in a received vector, the receiving device requests retransmission of the codeword. Calculate $P_{retrans}(C)$, the probability that a single codeword transmission will result in a request to retransmit.

$$P_{retrans} = 1 - P(no\ error\ detected)$$

$$= 1 - P(no\ error) - P(undetected\ error)$$

$$P_{undetec} = 2p^2(1-p)^2 + p^4$$

(3 marks)

so

$$P_{retrans}(C) = 1 - (1-p)^4 - 2p^2(1-p)^2 - p^4 \ = \ 4p + O(p^2)$$

so

$$P_{retrans}(C)|_{p=0.01} = etc. \sim 0.04$$

(3 marks)

(UNSEEN)