Week 4

Exercises (answers at end)

Version 2023-11-04. To accessible online version of these exercises

Exercise 4.1. Write down the weight enumerator of $Rep(n, \mathbb{F}_2)$, more generally of $Rep(n, \mathbb{F}_q)$.

Notation: below, $C \subseteq \mathbb{F}_q^n$ is a linear code, $d(C) = d$, and $t = \left[\frac{d-1}{2}\right]$.

Exercise 4.2. Prove that each vector $\underline{a}$ of weight $\leq t$ in the space $\mathbb{F}_q^n$ is a **unique coset leader** (that is, $w(\underline{a})$ is **strictly** less than weights of all other vectors in its coset $\underline{a} + C$).

Hint. If $\underline{a} \neq \underline{b}$ are in the same coset, show that $d \leq w(\underline{a}) + w(\underline{b})$. Then use $d - t > t$.

Exercise 4.3 (important fact about perfect linear codes — needed for exam)**.** Assume C is perfect. Use the Hamming bound to show that the number of cosets equals $\#S_t(\underline{0})$, i.e., there as many cosets as vectors of weight $\leq t$ in the space $\mathbb{F}_q^n$. Deduce that every coset has a unique coset leader, and that the coset leaders are exactly the vectors of weight $\leq t$.

Exercise 4.4 (not done in tutorial)**.** Find standard arrays for binary codes with each of the following generator matrices. For each code, determine whether every coset has a unique coset leader (i.e., if there is exactly one coset leader in each coset). Find the probability of an undetected / uncorrected error for $BSC(p)$ and argue whether the code is worth using for this channel, compared to transmitting unencoded information.

$$G_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad G_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad G_3 = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Exercise 4.5 (more weight enumerators — not done in tutorial)**.** (a) As usual, let $W_C(x, y)$ denote the weight enumerator of a q-ary linear code C. Show that $W_C(1, 0) = 1$ and that $W_C(1, 1) = q^k$ where $k = \dim C$.

(b) Show that the weight enumerator of the trivial binary code $\mathbb{F}_2^n$ is $W_{\mathbb{F}_2^n}(x, y) = (x + y)^n$. Can you write $W_{\mathbb{F}_q^n}(x, y)$ in a similar form?

(c) Write down $W_{E_3}(x, y)$. Can you suggest a compact way to write $W_{E_n}(x, y)$?

Week 4

Exercises — solutions

Version 2023-11-04. To accessible online version of these exercises

Exercise 4.1. Write down the weight enumerator of $Rep(n, \mathbb{F}_2)$, more generally of $Rep(n, \mathbb{F}_q)$.

Answer to E4.1. $Rep(n, \mathbb{F}_2)$ has one codevector of weight 0 and one codevector of weight n. Hence $W_{Rep(n,\mathbb{F}_2)}(x, y) = x^n + y^n$.

Exercise: show that $W_{\mathsf{Rep}(n,\mathbb{F}_q)}(x, y) = x^n + (q - 1)y^n$.

Notation: below, $C \subseteq \mathbb{F}_q^n$ is a linear code, $d(C) = d$, and $t = \left[\frac{d-1}{2}\right]$.

Exercise 4.2. Prove that each vector $\underline{a}$ of weight $\le t$ in the space $\mathbb{F}_q^n$ is a **unique coset leader** (that is, $w(\underline{a})$ is **strictly** less than weights of all other vectors in its coset $\underline{a} + C$).

Hint. If $\underline{a} \ne \underline{b}$ are in the same coset, show that $d \le w(\underline{a}) + w(\underline{b})$. Then use $d - t > t$.

Answer to E4.2. If $\underline{a}, \underline{b}$ are in the same coset, then by properties of cosets, $\underline{c} := \underline{a} - \underline{b}$ is a codevector. If $\underline{a} \ne \underline{b}$ then $\underline{c} \ne 0$ and so $d \le w(\underline{c}) = w(\underline{a} - \underline{b}) = d(\underline{a}, \underline{b})$. By the triangle inequality, $d(\underline{a}, \underline{b}) \le d(\underline{a}, \underline{0}) + d(\underline{0}, \underline{b}) = w(\underline{a}) + w(\underline{b})$. Thus, $d \le w(\underline{a}) + w(\underline{b})$ as claimed.

Now assume $w(\underline{a}) \le t$. Then $w(\underline{b}) \ge d - w(\underline{a}) \ge d - t$. But $t < \frac{d}{2}$ so $d - t > t$. We have $w(\underline{b}) \ge d - t > t \ge w(\underline{a})$. This shows that $\underline{a}$ has strictly minimal weight among the vectors in its coset, and so is the unique coset leader.

Exercise 4.3 (important fact about perfect linear codes — needed for exam)**.** Assume C is perfect. Use the Hamming bound to show that the number of cosets equals $\#S_t(\underline{0})$, i.e., there as many cosets as vectors of weight $\le t$ in the space $\mathbb{F}_q^n$. Deduce that every coset has a unique coset leader, and that the coset leaders are exactly the vectors of weight $\le t$.

Answer to E4.3. By the previous exercise, the vectors $\underline{a} \in S_t(\underline{0})$ are unique coset leaders of $\#S_t(\underline{0})$ distinct cosets. The total number of cosets is $\dfrac{q^n}{\#C}$.

Now if C is perfect, then $\#C = \dfrac{q^n}{\#S_t(\underline{0})}$ (the right-hand side is the Hamming bound), and so $\dfrac{q^n}{\#C} = \#S_t(\underline{0})$. Thus if C is perfect, cosets with a unique coset leader of weight $\leq t$ exhaust all cosets, as claimed.

Exercise 4.4 (not done in tutorial)**.** Find standard arrays for binary codes with each of the following generator matrices. For each code, determine whether every coset has a unique coset leader (i.e., if there is exactly one coset leader in each coset). Find the probability of an undetected / uncorrected error for $BSC(p)$ and argue whether the code is worth using for this channel, compared to transmitting unencoded information.

$$G_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad G_2 = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}, \qquad G_3 = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix}.$$

Answer to E4.4. G_1 generates the trivial binary code of length 2. Because the code is the whole space $\mathbb{F}_2^2$, its standard array consists of one row:

$$00 \quad 01 \quad 10 \quad 11$$

(the order of the codevectors after 00 is arbitrary). The only coset is the trivial coset which has only one coset leader, 00.

G_2 generates E_3, the even weight code of length 3. It has 4 codevectors and 2 cosets:

$$\begin{matrix} 000 & 101 & 011 & 110 \\ 001 & 100 & 010 & 111 \end{matrix}$$

Note that the non-trivial coset has three coset leaders; any of them could be put in column 1.

G_3: list all the 4 codevectors and then use the algorithm for constructing the standard array. One possible answer is given below:

$$\begin{matrix} 00000 & 10110 & 01011 & 11101 \\ 10000 & 00110 & 11011 & 01101 \\ 01000 & 11110 & 00011 & 10101 \\ 00100 & 10010 & 01111 & 11001 \\ 00010 & 10100 & 01001 & 11111 \\ 00001 & 10111 & 01010 & 11100 \\ 11000 & 01110 & 10011 & 00101 \\ 01100 & 11010 & 00111 & 10001 \end{matrix}$$

Coset leaders of weight 0 and 1 are the only coset leaders in their cosets. Coset leaders of weight 2 are not unique: e.g., 11000 and 00101 are coset leaders of the same coset.

Error probabilities. The code generated by G_1 is the trivial code, so using it is the same as sending unencoded information.

The code generated by G_2 has weight enumerator $W_{E_3}(x, y) = x^3 + 3xy^2$. Hence an undetected error occurs with probability

$$P_{\text{undetect}}(E_3) = W_{E_3}(1-p, p) - (1-p)^3 = 3(1-p)p^2 \sim 3p^2.$$

Note that this is of the same order as p^2 but at a rate of $2/3$ (recall the code considered in the chapter with worse rate $1/2$).

The probability of an uncorrected error here is $1 - P_{\text{corr}}(E_3) = 1 - (\alpha_0(1-p)^3 + \alpha_1 p(1-p)^2)$ where $\alpha_0 = 1$ (one coset leader of weight 0) and $\alpha_1 = 1$ (one coset leader of weight 1) . We have $1 - P_{\text{corr}}(E_3) = 1 - ((1-p)^3 + p(1-p)^2) = 1 - (1-p+p)(1-p)^2 = 1 - (1-p)^2 \sim 2p$.

The code E_3 does not improve the probability of incorrect decoding. Indeed, Hamming's theory says that E_3 has no error-correcting capability and can only be used for error detection.

The code generated by G_3 has weight enumerator $x^5 + 2x^2y^3 + xy^4$. Hence

$$P_{\text{undetect}} = 2(1-p)^2p^3 + (1-p)p^4 \sim 2p^3.$$

If $p = 0.01$, this is $\approx 2 \times 10^{-6}$, which is 5,000 times better than without encoding.

Furthermore, looking at the coset leaders, we find one coset leader of weight 0, $\alpha_0 = 1$; five coset leaders of weight 1, $\alpha_1 = 5$; two coset leaders of weight 2, $\alpha_2 = 2$. This gives

$$\begin{aligned} 1 - P_{\text{corr}} &= 1 - (\alpha_0(1-p)^5 + \alpha_1 p(1-p)^4 + \alpha_2 p^2(1-p)^3) \\ &= 1 - ((1-p)^2 + 5p(1-p) + 2p^2)(1-p)^3 \\ &= 8p^2 - 14p^3 + 9p^4 - 2p^5 \sim 8p^2. \end{aligned}$$

If $p = 0.01$, incorrect decoding occurs with probability $\approx 8 \times 10^{-4}$, which is 12.5 times better than without encoding.

Of course, this improvement in reliability comes at a price: the rate of the code is only 0.4, meaning that we have to transmit 2.5 times as much information.

Exercise 4.5 (more weight enumerators — not done in tutorial)**.** (a) As usual, let $W_C(x, y)$ denote the weight enumerator of a q-ary linear code C. Show that $W_C(1, 0) = 1$ and that $W_C(1, 1) = q^k$ where $k = \dim C$.

(b) Show that the weight enumerator of the trivial binary code $\mathbb{F}_2^n$ is $W_{\mathbb{F}_2^n}(x, y) = (x + y)^n$. Can you write $W_{\mathbb{F}_q^n}(x, y)$ in a similar form?

(c) Write down $W_{E_3}(x, y)$. Can you suggest a compact way to write $W_{E_n}(x, y)$?

Answer to E4.5. (a) Recall $W_C(x, y) = \sum_{\underline{c} \in C} x^{n-w(\underline{c})} y^{w(\underline{c})}$. If $y = 0$, the only non-zero term in this sum is the term without y which corresponds to the (unique) zero codevector of the linear code C; thus, $W_C(x, 0) = x^n$ and $W_C(1, 0) = 1$. Also, $W_C(1, 1) = \sum_{\underline{c} \in C} 1 = \#C = q^k$.

(b) To work out $W_{\mathbb{F}_q^n}(x,y)$, write it in the form $W_{\mathbb{F}_q^n}(x,y) = \sum_{i=0}^{n} A_i x^{n-i} y^i$ where $A_i = \#\{\underline{v} \in \mathbb{F}_q^n : w(\underline{v}) = i\}$. Note that $w(\underline{v}) = d(\underline{v}, \underline{0})$, and in the proof of the Hamming bound we calculated the number of words at distance i from $\underline{0}$ (or from any other fixed vector) to be $\binom{n}{i}(q-1)^i$. Hence

$$W_{\mathbb{F}_q^n}(x,y) = \sum_{i=0}^{n} \binom{n}{i} (q-1)^i x^{n-i} y^i = (x + (q-1)y)^n.$$

(c) The even weight code E_3 is $\{000, 011, 101, 110\}$, so that $W_{E_3}(x,y) = x^3 + 3xy^2$. The weight enumerator of E_n will be obtained in the lectures as an application of the MacWilliams identity.