Coding Theory and Cryptography: Session 3

Paridhi Latawa

March 16, 2021

Abstract

This material covers bounds on codes, perfect codes, and cyclic codes. Readings and practice problems covered are the following.

1. Module: Intro to Coding Theory via Hamming Codes. Problems: 22, 23, 25, 26.

2. Hankerson Section 3.1 - 3.7 and Chapter 4. Problems: 3.1.5ace, 3.3.10, 4.1.20, 21, 4.2.9, 4.3.4, 4.4.9, 4.5.5e.

3. Trappe and Washington Section 18.7. Problems: 12 and 13 on page 447.

4. Magma Software [separate HW doc]

1 Module: Intro to Coding Theory via Hamming Codes

22. Consider the code from Problem 7, that is $C = 1010, 0101, 1111$. Give an example of a vector that has the same distance to more than one codeword.

The vector 1111 has the same distance to both 1010 and 01010.
$d(1111, 1010) = d(1111, 0101) = 2$.

23. Let C be a code with an even minimum distance. Show that there exists a vector that has the same distance to more than one codeword.

Assume a code of length n, where $n > 2$. As the minimum distance is even, $d = 2m$ where m is any integer $m \geq 1$. For c to be linear, there must be a codeword that consists of all ones.

Assuming $n = 3$, the codewords with an even number of ones is $3c_2$, which comes out to be 3, or $2m + 1$ generalized.

These 3 codewords will be at equal distance from the codeword with all ones.

Assume $c = 000, 011, 110, 101, 111$, where there is an even minimum distance (meaning the weight is even).

$d(111, 011) = 1$
$d(111, 110) = 1$
$d(111, 1010) = 1$

$d(011, 110) = 2$
$d(011, 101) = 2$

For any codeword with even weight, $011, 110, or 101$, they will have the same distance to more than one codeword.

<u>**25.**</u> Show that the $[6, 3, 3]$ is not perfect and $[7, 4, 3]$ Hamming code is perfect using the sphere packing bound.

For a code C of length n and odd distance $d = 2t + 1$ to be perfect, C needs to attain the Hamming bound. This means that $|C| * \binom{n}{0} + \binom{n}{1} + ... + \binom{n}{t} \leq 2^n$

For $[6, 3, 3]$, we get

$2^3 * \binom{6}{0} + \binom{6}{1} \leq 2^6$

Simplifying, we get:

$2^3 * 1 + 6 \leq 2^6$

As the left side is less than and doesn't equal 2^6, $[6, 3, 3]$ is not perfect.

For $[7, 4, 3]$, we get

$2^4 * \binom{7}{0} + \binom{7}{1} \leq 2^7$

$2^4 * 1 + 7 \leq 2^7$

$2^4 * 2^3 \leq 2^7$

$2^7 = 2^7$

Thus, $[7, 4, 3]$ is perfect.

<u>**26.**</u> Show that the general Hamming codes are all perfect.

For a code C of length n and odd distance $d = 2t + 1$ to be perfect, C needs to attain the Hamming bound. This means that $|C| = \frac{2^n}{\binom{n}{0}+\binom{n}{1}+...+\binom{n}{t}}$

As per Hankerson section 3.3 page 73, we know that a Hamming code has a distance $d = 3$ and dimension $n = 2^r - 1$. So, $d = 3 = 2t + 1$ gives us $t = 1$.

Inputting this into the Hamming bound, we get: $\frac{2^n}{\binom{n}{0}+\ldots+\binom{n}{t}}$

As $t = 1$, we get $\frac{2^n}{\binom{n}{0}+\ldots+\binom{n}{1}}$

As $n = 2^r - 1$ and simplifying further, we get $\frac{2^{2^r-1}}{1+n}$

Inputting the value for n in the denominator, we get $\frac{2^{2^r-1}}{1+2^r-1}$

Simplifying further, we get 2^{2^r-1-r}

This confirms that any perfect code that has the length and distance mentioned above, which are characteristic of Hamming codes, has exactly 2^{2^r-1-r} codewords. This is also a power of 2, so the general Hamming codes are all perfect.

2 Sections 3.1 - 3.7 and Chapter 4 of Hankerson

3.1.5: Find an upper bound for the size or dimension of a linear code with the given values of n and d.

a. $n = 8, d = 3$

From $d = 3 = 2t + 1$, we get $t = 1$.

The Hamming bound gives $|C| \leq \binom{n}{t} = \frac{2^8}{\binom{8}{0}+\binom{8}{1}} = \frac{256}{1+8} = \frac{256}{9}$.

But $|C|$ must be a power of 2, so $|C| \leq 16$, and thus $k \leq 4$.

So the upper bound for the size or dimension of a linear code with these values of n and d is 2^4.

b. $n = 7, d = 3$

From $d = 3 = 2t + 1$, we get $t = 1$.

The Hamming bound gives $|C| \leq \binom{n}{t} = \frac{2^7}{\binom{7}{0}+\binom{7}{1}} = \frac{128}{1+7} = \frac{128}{8}$.

This equals 16, which agrees with the condition that $|C|$ must be a power of 2, so $|C| \leq 16$, and thus $k \leq 4$.

So, the upper bound for the size or dimension of a linear code with these values of n and d is 2^4.

e. $n = 15, d = 5$

From $d = 5 = 2t + 1$, we get $t = 2$.

The Hamming bound gives $|C| \leq \binom{n}{t} = \frac{2^15}{\binom{15}{0}+\binom{15}{1}+\binom{15}{2}} = \frac{32768}{1+15+105} = \frac{32768}{121}$.

But $|C|$ must be a power of 2, so $|C| \leq 256$, and thus $k \leq 8$.

So the upper bound for the size or dimension of a linear code with these values of n and d is 2^8.

3.3.10. Use the Hamming code of length 7 in Example 3.3.1 and the message assignment in Exercise 2.6.12. Decode the message received: 1010111, 0110111, 1000010, 0010101, 1001011, 0010000, 1111100.

As per Example 3.3.1, we know a possibility for a parity check matrix for the Hamming code of length $7(r = 3)$ is H = $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

The syndrome of $wH = 1010111H = 101$, which is the third row of H. Thus the coset leader u is the third row of $I_7 : u = 0010000$. Thus w is decoded as $w + u = 1000111$.

The syndrome of $wH = 0110111H = 100$, which is the fifth row of H. Thus the coset leader u is the fifth row of $I_7 : u = 0000100$. Thus w is decoded as $w + u = 0110011$.

The syndrome of $wH = 1000010H = 101$, which is the third row of H. Thus the coset leader u is the third row of $I_7 : u = 0010000$. Thus w is decoded as $w + u = 1010010$.

The syndrome of $wH = 0010101H = 000$. This element is not any row of H. Thus, it is not decodable.

The syndrome of $wH = 1001011H = 111$, which is the first row of H. Thus the coset leader u is the first row of $I_7 : u = 1000000$. Thus w is decoded as $w + u = 0001011$.

The syndrome of $wH = 0010000H = 101$, which is the third row of H. Thus the coset leader u is the third row of $I_7 : u = 001000$. Thus w is decoded as $w + u = 0000000$. This is A.

The syndrome of $wH = 1111100H = 011$, which is the fourth row of H. Thus the coset leader u is the fourth row of $I_7 : u = 0001000$. Thus w is decoded as $w + u = 1110100$.

4.1.20: Let $h(x) = 1 + x^7$. Compute $f(X) mod h(x)$ and $p(x) mod h(x)$, and decide whether $f(X) \equiv p(x)(mod h(x))$.

a) $f(x) = 1 + x^3 + x^8, p(x) = x + x^3 + x^7$

$f(x) mod h(x) = (1 + x^3 + x^8) mod(1 + x^7).r(x) = x^3 - x + 1$

$p(x) mod h(x) = x + x^3 + x^7 mod(1 + x^7), r(x) = x^3 + x - 1$

$f(x) and \equiv p(x)(mod h(x))$ if and only if they have the same remainder when divided by $h(x)$

Both have the same remainder $r(X)$, so $f(x) and \equiv p(x)(mod h(x))$

b) $f(x) = x + x^5 + x^9, p(x) = x + x^5 + x^6 + x^13$

$f(x) mod h(x) = (x + x^5 + x^9) mod(1 + x^7).r(x) = x^5 - x^2 + x$

$p(x) mod h(x) = x + x^5 + x^6 + x^13 mod(1 + x^7), r(x) = x^5 + x$

Both do not have the same remainder so $f(x)$ is not equivalent to $p(x)$

c) $f(x) = 1 + x, p(x) = x + x^7$

$f(x) mod h(x) =$(1 + x)mod(1 + x^7). r(x) = 0

$p(x) mod h(x) = x + x^7 mod(1 + x^7), r(x) = x - 1$

Both do not have the same remainder so $f(x)$ is not equivalent to $p(x)$

4.1.21: Let $h(x) = 1 + x^7$. Compute $(f(x) + g(x))$ mod $h(x)$ and $(f(x)g(x))$ mod $h(x)$, where a) $f(x) = 1 + x^6 + x^8, g(x) = 1 + x$

$f(x) + g(x) = x^8 + x^6 + x + 2$ $(f(x) + g(x))$ mod $h(X) = x^6 + 2 = r(x)$

$f(X)g(x) = x^9 + x^8 + x^7 + x^6 + x + 1$ $(f(x)g(x))$ mod $h(X) = x^6 - x^2 = r(x)$

b) $f(x) = 1 + x^5 + x^9, g(x) = x + x^2 + x^7$

$f(x) + g(x) = 1 + x + x^2 + x^5 + x^7 + x^9$ $(f(x) + g(x))$ mod $h(X) = x^5 + x = r(x)$

$f(X)g(x) = x^{16} + x^{12} + x^{11} + x^{10} + 2x^7 + x^6 + x^2 + x$ $(f(x)g(x))$ mod $h(X) = x^6 - x^5 - x^4 - x^3 + 2x^2 + x - 2 = r(x)$

c) $f(x) = 1 + x^4 + x^5, g(x) = 1 + x + x^2$

$f(x)+g(x) = x^5+x^4+x^2+x+2$ $(f(x)+g(x)) \bmod h(X) = x^5+x^4+x^2+x+2 = r(x)$

$f(X)g(x) = x^7 + 2x^6 + 2x^5 + x^4 + x^2 + x + 1$ $(f(x)g(x)) \bmod h(X) = 2x^6 + 2x^5 + x^4 + x^2 + x = r(x)$

4.2.9: Find all words v of length 6 such that a) $pi^2(v) = v$ $v = 000000, 101010, 010101, 110011, 001100,$

b) $pi^3(v) = v$ $v = 000000, 0110110, 001001, 100100, 111111$

4.3.4) Let $g(x) = 1 + x^2 + x^3$ be the generator polynomial of a linear cyclic code of length 7.

a) Encode the following message polynomials: $1 + x^3, x, x + x^2 + x^3$ $k = 4$ $a(x) = 1 + x^3 . c(x) = a(x)g(x) c(x) = (1 + x^3)(1 + x^2 + x^3) = 1 + x^2 + 2x^3 + x^5 + x^6$ So, $c = 1010011$

$a(x) = x . c(x) = a(x)g(x) c(x) = (x)(1 + x^2 + x^3) = x + x^3 + x^4$ So, $c = 0101100$

$a(x) = x + x^2 + x^3 c(x) = a(x)g(x) c(x) = (x + x^2 + x^3)(1 + x^2 + x^3) = x + x^2 + 2x^3 + 2x^4 + 2x^5 + x^6$ So, $c = 1100010$

b) Find the message polynomial corresponding to the codewords $c(x) : x^2 + x^4 + x^5, 1 + x + x^2 + x^4, x^2 + x^3 + x^4 + x^6$

If $c(x) = x^2+x^4+x^5$, the corresponding message polynomial is $c(x)/g(x) = a(x) = x^2$. So, $c = 0010$

If $c(x) = 1 + x + x^2 + x^4$, the corresponding message polynomial is $c(x)/g(x) = a(x) = -1 + x$ with a remainder of $2x^2 + 2$. So, $c = 1100$

If $c(x) = x^2 + x^3 + x^4 + x^6$, the corresponding message polynomial is $c(x)/g(x) = a(x) = -2 + 2x - x^2 + x^3$ with a remainder of $4x^2 - 2x + 2$. So, $c = 0011$

4.4.9. Find a generator and a generating matrix for a linear code of length n and dimension k where

a) $n = 12, k = 5$ $1 + x^{12}$ factored is $(1 + x^6)^2 = (x^4 + 1)(x^8 - x^4 + 1)$. So, one generator for this linear code is simply $x^4 + 1$.

One generator matrix is the one in which the rows are the codewords that correspond to the generator polynomial and its first $k - 1$ cyclic shifts.

The generator matrix has size k x n

So if $g(x) = x^4 + 1$, this is also the same as 100010000000.

$xg(x) = 1x + x^5 = 010001000000$.

$x^2g(x) = 1x^2 + x^6 = 001000100000.$

$x^3g(x) = 1x^3 + x^7 = 000100010000.$

$x^4g(x) = 1x^4 + x^8 = 000010001000$

So a generating matrix for C is G $= \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \\ x^4g(x) \end{bmatrix} = \begin{bmatrix} 100010000000 \\ 010001000000 \\ 001000100000 \\ 000100010000 \\ 000010001000 \end{bmatrix}$

b) $n = 12, k = 7$ $1 + x^{12}$ factored is $(1 + x^6)^2 = (x^4 + 1)(x^8 - x^4 + 1)$. So, one generator for this linear code is simply $x^4 + 1$.

One generator matrix is the one in which the rows are the codewords that correspond to the generator polynomial and its first $k - 1$ cyclic shifts.

The generator matrix has size k x n

So if $g(x) = x^4 + 1$, this is also the same as 100010000000.

$xg(x) = 1x + x^5 = 010001000000.$

$x^2g(x) = 1x^2 + x^6 = 001000100000.$

$x^3g(x) = 1x^3 + x^7 = 000100010000.$

$x^4g(x) = 1x^4 + x^8 = 000010001000$

$x^5g(x) = 1x^5 + x^9 = 000001000100$

$x^6g(x) = 1x^6 + x^10 = 000000100010$

So a generating matrix for C is G $= \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \\ x^4g(x) \\ x^5g(x) \\ x^6g(x) \end{bmatrix} = \begin{bmatrix} 100010000000 \\ 010001000000 \\ 001000100000 \\ 000100010000 \\ 000010001000 \\ 000001000100 \\ 000000100010 \end{bmatrix}$

c) $n = 14, k = 5$ $1+x^{14}$ factored is $(1+x^7)^2 = (x^2+1)(x^12-x^10+x^8-x^6+x^4-x^2+1)$. So, one generator for this linear code is simply $x^2 + 1$.

One generator matrix is the one in which the rows are the codewords that correspond to the generator polynomial and its first $k - 1$ cyclic shifts.

The generator matrix has size k x n

So if $g(x) = x^2 + 1$, this is also the same as 10100000000000.

$xg(x) = x^3 + 1x = 01010000000000$

$x^2g(x) = x^4 + 1x^2 = 00101000000000$

$x^3g(x) = x^5 + 1x^3 = 00010100000000$

$x^4g(x) = x^6 + 1x^4 = 00001010000000$

So a generating matrix for C is G $= \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \\ x^4g(x) \end{bmatrix} = \begin{bmatrix} 10100000000000 \\ 01010000000000 \\ 00101000000000 \\ 00010100000000 \\ 00001010000000 \end{bmatrix}$

d) $n = 14, k = 6$

$1 + x^{14}$ factored is $(1 + x^7)^2 = (x^2 + 1)(x^12 - x^10 + x^8 - x^6 + x^4 - x^2 + 1)$. So, one generator for this linear code is simply $x^2 + 1$.

One generator matrix is the one in which the rows are the codewords that correspond to the generator polynomial and its first $k - 1$ cyclic shifts.

The generator matrix has size k x n

So if $g(x) = x^2 + 1$, this is also the same as 10100000000000.

$xg(x) = x^3 + 1x = 01010000000000$

$x^2g(x) = x^4 + 1x^2 = 00101000000000$

$x^3g(x) = x^5 + 1x^3 = 00010100000000$

$x^4g(x) = x^6 + 1x^4 = 00001010000000$

$x^5g(x) = x^7 + 1x^5 = 00000101000000$

So a generating matrix for C is G $= \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \\ x^4g(x) \\ x^5g(x) \end{bmatrix} = \begin{bmatrix} 10100000000000 \\ 01010000000000 \\ 00101000000000 \\ 00010100000000 \\ 00001010000000 \\ 00000101000000 \end{bmatrix}$

e) $n = 14, k = 8$ $1+x^{14}$ factored is $(1+x^7)^2 = (x^2+1)(x^12-x^10+x^8-x^6+x^4-x^2+1)$. So, one generator for this linear code is simply $x^2 + 1$.

One generator matrix is the one in which the rows are the codewords that correspond to the generator polynomial and its first $k - 1$ cyclic shifts.

The generator matrix has size k x n

So if $g(x) = x^2 + 1$, this is also the same as 10100000000000.

$xg(x) = x^3 + 1x = 01010000000000$

$x^2g(x) = x^4 + 1x^2 = 00101000000000$

$x^3g(x) = x^5 + 1x^3 = 00010100000000$

$x^4g(x) = x^6 + 1x^4 = 00001010000000$

$x^5g(x) = x^7 + 1x^5 = 00000101000000$

$x^6g(x) = x^8 + 1x^6 = 00000010100000$

$x^7g(x) = x^9 + 1x^7 = 00000001010000$

So a generating matrix for C is G $= \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \\ x^4g(x) \\ x^5g(x) \\ x^6g(x) \\ x^7g(x) \end{bmatrix} \begin{bmatrix} 10100000000000 \\ 01010000000000 \\ 00101000000000 \\ 00010100000000 \\ 00001010000000 \\ 00000101000000 \\ 00000010100000 \\ 00000001010000 \end{bmatrix}$

4.5.5: Find the generator polynomial for the dual code of the cyclic code of length n having generator polynomial $g(x)$ where:

e) $n = 15, g(x) = 1 + x + x^4$

$\frac{1+x^{15}}{1+x+x^4} = x^{11} - x^8 - x^7 + x^5 + 2x^4 + x^3 - x^2 - 3x - 3 = h(x)$

The dual code generator polynomial would be $x^n h(x^{-1})$, when computed taking the values from above. This is called the reciprocal of H.

This could be inputted into Magma as a calculation.

So,

$h(X) = (x^15 - 1)/g(x)$. We can verify that G is a divisor of H by taking the modulus. When conducted, the result is 0, confirming that it is a divisor.

3 Trappe and Washington Section Page 447

12. Let $g(x) = 1 + x + x^3$ be a polynomial with coefficients in Z_2 a) Show that $g(x)$ is a factor of $x^7 - 1$ in $Z_2[x]$

$\frac{x^7-1}{1+x+x^3} = x^4 - x^2 - x + 1$, with a remainder

b) The polynomial $g(x)$ is the generating polynomial for a cyclic code $[7,4]$ code C. Find the generating matrix for C.

Knowing $n = 7$ and $k = 4$, we can conduct the following algorithm to get the generating matrix. $g(x) = 1 + x + x^3 = 1101000$

$xg(x) = 1x + x^2 + x^4 = 0110100$

$x^2g(x) = 1x^2 + x^3 + x^5 = 0011010$

$x^3g(x) = 1x^3 + x^4 + x^6 = 0001101$

So a generating matrix $G = \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \end{bmatrix} = \begin{bmatrix} 1101000 \\ 0110100 \\ 0011010 \\ 0001101 \end{bmatrix}$

c) Find a parity check matrix H for C.

From part a, we get the generator matrix $G = \begin{bmatrix} 1101000 \\ 0110100 \\ 0011010 \\ 0001101 \end{bmatrix}$

Row reducing the generator matrix, we get $G = \begin{bmatrix} I_k \mid P \end{bmatrix} = \left[\begin{array}{c|c} 1000 & 110 \\ 0100 & 011 \\ 0010 & 111 \\ 0001 & 101 \end{array}\right]$

The standard definition of a parity check matrix is $H = [-P^T|I_{n-k}]$

Applying these transformations to the generator matrix, we get $H = \left[\begin{array}{c|c} 1011 & 100 \\ 1110 & 010 \\ 0111 & 001 \end{array}\right]$

d) Show that $G'H^T = 0$, where $G' = = \begin{bmatrix} 1101000 \\ 0110100 \\ 0011010 \\ 0111001 \end{bmatrix}$

$$H = \left[\begin{array}{c|c} 1011 & 100 \\ 1110 & 010 \\ 0111 & 001 \end{array}\right]$$ The transpose of H is $H^T = \begin{bmatrix} 110 \\ 011 \\ 111 \\ 101 \\ 100 \\ 010 \\ 001 \end{bmatrix}$

$$G'H^T = \begin{bmatrix} 1101000 \\ 0110100 \\ 0011010 \\ 0111001 \end{bmatrix} \begin{bmatrix} 110 \\ 011 \\ 111 \\ 101 \\ 100 \\ 010 \\ 001 \end{bmatrix} = 0$$

e) Show that the rows of G' generate C

G' has the same first three rows as G, and the last row is generated by the linear combination of row two and four from G. The generator matrix is the basis of the linear code, so G' is a generator matrix for C.

f) Show that a permutation of the columns of G' gives the generating matrix for the Hamming $[7, 4]$ code, and therefore these two codes are equivalent.

We know that $n = 7$ and $k = 4$ So, we can find irreducible factors of $1 + x^7$, which are $(x + 1)(x^6 - x^5 + x^4 - x^3 + x^2 - x + 1)$

So, one generator for this linear code is $x + 1$

A generator matrix is a matrix in which the rows are the codewords that correspond to the generator polynomial and its first $k - 1$ cyclic shifts.

The generator matrix has size k x n

So if $g(x) = x + 1$, this is also the same as 1100000.

$xg(x) = x^3 + 1x = 0110000$

$x^2g(x) = x^4 + 1x^2 = 0011000$

$x^3g(x) = x^5 + 1x^3 = 0001100$

So a generating matrix for C is G $= \begin{bmatrix} g(x) \\ xg(x) \\ x^2g(x) \\ x^3g(x) \end{bmatrix} \begin{bmatrix} 1100000 \\ 0110000 \\ 0011000 \\ 0001100 \end{bmatrix}$ This row reduces to the matrix $\begin{bmatrix} 1000100 \\ 0100100 \\ 0010100 \\ 0001100 \end{bmatrix}$

The G', when arranged such that I_4 is the beginning four columns, will form a generating matrix. While this will not match the above generating matrix, a Hamming code can have multiple generating matrices.

13. Let C be the cyclic binary code of length 4 with generating polynomial $g(x) = x^2 + 1$. Which of the following polynomials correspond to elements of C?

$f_1(x) = 1 + x + x^3 f_2(x) = 1 + x + x^2 + x^3 f_3(x) = x^2 + x^3$

As written on page 430 of Trappe and Washington, if $m(x)$ corresponds to an element of C, then $m(x) = g(x)f(x)$ or $h(x)m(x) = 0 mod(x^n - 1)$

$\frac{1+x+x^3}{x^2+1} = x + \frac{1}{x^2+1} \frac{1+x+x^2+x^3}{x^2+1} = x + 1 \frac{x^2+x^3}{x^2+1} = x + 1 + \frac{-x-1}{x^2+1}$

$f_2(x) = 1 + x + x^2 + x^3$ is the only polynomial that does not have a remainder so it is the only polynomial that corresponds to an element of C.