MATH 2210 Name: _____

Workshop 3: Hamming Codes ID: _____

Work in your group to complete the following exercises. You may print this handout, annotate the PDF or write your answer on paper. Make your grader's life easier by writing neatly and legibly!

Please include full explanations and write your answers using complete sentences (not just a bunch of mathematical symbols!). It is important to be able to explain your reasoning to someone else in writing.

Warmup

Question 1. Consider the following matrix:

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 2 & -5 & 0 \\ 1 & 5 & -1 \\ 0 & -4 & 3 \\ -2 & 1 & 1 \end{bmatrix}$$

(a) Show that A has rank 3.

(b) Determine whether each of the following vectors is in the column space of A.

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} -2 \\ -3 \\ 6 \\ -4 \\ -1 \end{bmatrix} \qquad \begin{bmatrix} 5 \\ 3 \\ -1 \\ 0 \\ 2 \end{bmatrix}$$

Linear Error-Correcting Codes

We rely on the internet for many important aspects of our lives. For example, some pacemakers send a signal to healthcare providers to allow for remote patient monitoring. In such settings, it is important to guarantee that the message (sequence of binary digits) received by the doctor is accurate. However, numerous external factors such normal background radiation can corrupt some bits of the message during transmission.

As we will explore, error-correcting codes provide a framework for us to deal with this issue: by adding a few extra bits to the message, we'll see that it will be possible to detect when a transmission error occurs. Moreover, by cleverly choosing the encoding, we'll see that we can locate and correct the error to recover the original message.

Question 2. In this worksheet, we'll consider binary messages. Therefore, we will want all of the entries in our matrices and vectors to be either 0 or 1, instead of allowing them to be all real numbers as we have done so far. All of the ideas (Gaussian elimination, linear independence, etc.) we have encountered in the course so far will also work in this scenario. This is because the digits 0 and 1 form a **finite field** (a set where the four basic arithmetical operations make sense), which we call $\mathbb{F}_2$. We can summarize the operations in $\mathbb{F}_2$ with the following tables.

$+$	0	1
0	0	1
1	1	0

$\times$	0	1
0	0	0
1	0	1

So, the operations are defined just like usual, except now $1 + 1 = 0$. We can think about having 0 represent even integers and 1 represent odd integers, since odd + odd = even.

(a) Perform Gaussian elimination to reduce the following matrix (over $\mathbb{F}_2$) into Echelon form.

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

(b) Determine the rank of B, and find bases for its column space and null-space.

A binary *error-correcting code* is a map $\mathbb{F}_2^k \mapsto \mathbb{F}_2^n$ from message words of length k (that is, column vectors with k entries) to code words of length n ($n > k$), similar to familiar real maps $\mathbb{R}^k \to R^n$. Such a code is *linear* if it is induced by multiplication by a $n \times k$ *generator matrix*, G.

Question 3. Assume that G has rank k. Argue that the set of code words, C, form a subspace of $\mathbb{F}_2^n$. What is its dimension?

Question 4. We'll consider the $(7, 4)$-Hamming code, so named because it maps 4-bit messages to 7-bit code words. Its generator matrix is,

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

(a) Determine the encoding of the following messages. (You can divide up the computations among your group members.)

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

(b) Examine your answer to part (a), or the structure of G. Given a valid code word (that is, a vector in the column space of G), what is an easy way to determine its corresponding message?

(c) For each of the following vectors, determine whether it is a valid code word. If it is, what message is it encoding?

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 \end{bmatrix}^\intercal \qquad \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}^\intercal$$

We would like a way to quickly determine whether a code word is valid, and to which message it corresponds. We can do this by using a second matrix, the *parity check matrix* H. This is an $n \times (n-k)$ matrix defined such that $H^\intercal c = 0$ for each code word c.

Question 5. Show that,

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

is a parity check matrix for the (7,4)-Hamming code.

Hint: By rank-nullity, it suffices to argue that $H^\intercal c = 0$ for four linearly-independent code words.

We can interpret the parity check matrix as enforcing three constraints on our code words:

1. There are an even number of 1s in positions $\{4, 5, 6, 7\}$ in the code word.
2. There are an even number of 1s in positions $\{2, 3, 6, 7\}$ in the code word.
3. There are an even number of 1s in positions $\{1, 3, 5, 7\}$ in the code word.

Hence, the name parity check matrix makes sense (we are checking the parity of the sum of these three sets).

Furthermore, if we look at the binary representations of 1 through 7, we see that the first set contains the numbers with a 1 in the third bit from the right (the 4s bit), the second set contains the numbers with a 1 in the second bit from the right (the 2s bit), and the third set contains the numbers with a 1 in the rightmost (1s) bit. Therefore, if our parity check vector $H^\intercal c \neq 0$, it tells us the (binary representation of the) index of the bit we need to flip to restore the proper parity.

Example:

The code word,

$$c = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}^\intercal$$

is not valid because it has parity check vector,

$$H^\intercal c = \begin{bmatrix} 1 & 1 & 0 \end{bmatrix}^\intercal.$$

110 is the binary representation of the number 6, and by flipping the 6th bit, we get the valid code word vector

$$\hat{c} = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}^\intercal.$$

The key property of the (7,4)-Hamming code is that it has *distance* 3, meaning that any two distinct code words differ in at least 3 bits [1]. Said another way, at least 3 bits in any valid code word need to be flipped in order to get to another valid code word.

Question 6.

(a) Explain how this allows us to detect if *at most 2* bits are flipped during transmission.

(b) Explain how this allows us to correct *at most 1* bit that was flipped during transmission.

Question 7. Let's put everything together. Each of the following vectors is a code word with 1 transmission error. Determine the location of the error, the corrected code word, and its corresponding message.

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

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

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

Question 8 (**Bonus**)**.** Try to find patterns in the generator and parity check matrices for the $(7, 4)$-Hamming code. Use these to write down the matrices for the $(15, 11)$-Hamming code.

[1] If you want, try to argue why this is the case. By linearity (why?), it suffices to argue that each non-zero code word contains at least 3 zeros. To do this, use Question 4b, and the discussion about the parity check matrix on the previous page.