Hamming Codes

1 Linear Codes

1.1 Summary on linear codes

An (n, k, d_{min})-linear code C is a linear subspace of dimension k of F^n with minimum distance d_{min}

We can represent it by

- Generator matrix $\mathbf{G}(k \times n)$ used for encoding

  $$C = \left\{\mathbf{mG} : \mathbf{m} \in F^k\right\}$$

  mG means the message multiplied by the generator matrix

- Parity-check matrix $\mathbf{H}(n - k \times n)$ used for detecting errors

  $$C = \{\mathbf{c} \in F^n : \mathbf{cH}^\top = \mathbf{0}\}$$

$$cH^T \Leftrightarrow c_1 = c_2 = c_3 = c_4 = ...$$

The repetition code and Parity-Check codes are dual to each other, meaning that they have symmetry. The generator matrix of a repetition code is H of a parity-check code and vice versa.

1.2 Parameters of a linear code

A **linear code** is simply any subspace of F^n

Parameters:

- Length n
- Dimension k
- Redundancy $r = n - k$
- Rate $R = k/n$
- Minimum distance d_{min} - this shows how many errors can be corrected
- Error-correction capability $t = \lfloor (d_{min} - 1)/2 \rfloor$

We usually write (n, k, d_{min})-"name of code"

1.3 Parameters of parity-check and repetition codes

Parameter	Parity-check	Repetition
Length n	n	n
Dimension k	n-1	1
Rendundancy r	1	n-1
Rate R	1-1/n	1/n
Minimum distance d_{min}	2	n
Error-correction capability t	0	$\lfloor (n-1)/2 \rfloor$

1.4 Minimum distance of a linear code

Theorem: Viewing the columns of H as vectors in F^{n-k}, d_{min} is the minimum number of linearly dependent columns of H

E.g. for the $(5,1,5)$-repetition code

$$H_{\text{repetition}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

Any set of four columns is linearly independent, but the set of all five columns is linearly dependent, therefore $d_{min} = 5$

This is true for all repetition codes, you can see this adding all the columns together, remembering modulo 2, the linear combination is 0.

2 Hamming codes

2.1 The Hamming code of redundancy 3

Definition: This is the linear code with the parity check matrix

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

The columns are the integers 1 to 7 written in binary

Any two columns are linearly independent; the first 3 are linearly dependent.

Therefore $\mathbf{d_{min} = 3}$ this is the (7,4,3)-Hamming code

So for any matrix, the number of linearly dependent rows is equal to d_{min}

2.2 Generalisation

For any $r \geqslant 2$ we use the parity-check matrix whose columns are all the integers from 1 to $2^r - 1$ in binary

- For r=2: the (3,1,3)-repetition code
- For r=3: the (7,4,3)-Hamming code
- For r=4: the (15,11,4)-Hamming code with parity-check matrix

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

Note that $d_{min} = 3$ for all r

2.3 Parameters

In general a Hamming code has parameters

- Length $n = 2^r - 1$
- Dimension $k = 2^r - r - 1$
- Redundancy $r = n - k$
- Rate $R = 1 - r/(2^r - 1)$
- Minimum distance $d_{min} = 3$
- Error correction capability $t = 1$

2.4 Optimality of Hamming codes

The **sphere-packing bound**: If C is a code of length n which can correct at least one error (and hence $d_{min} = 3$), then

$$|C| \leq \frac{2^n}{n+1}$$

Consider again the spheres showing the errors that can be corrected.
The volume of a sphere is $\frac{4}{3}\pi r^3$
There are $|C|$ spheres which do not intersect.
There are $|C|(n+1)$ vectors in all the spheres
$|C|(n+1) \leqslant 2^n$ (can't be greater than the number of vectors)

3 CW

The hamming encoder is mG

Message decoder is just doing the opposite of creating the message in the first question.