1 Recap

Recall from the last lecture that error-correcting codes are in fact injective maps from $k$ symbols to $n$ symbols in $\Sigma$, 

$$\text{Enc}: \Sigma^k \rightarrow \Sigma^n$$

where $k$ and $n$ are referred to as the message dimension and block length respectively. We also call the image of the encoding function code, which is usually denoted by $C$, i.e. $C = \text{Im}(\text{Enc})$; and an element $y \in C$ a codeword.

The minimum distance $d$ is defined as the smallest Hamming distance between two distinct codewords,

$$d = \min_{y_1 \neq y_2 \in C} \{\Delta(y_1, y_2)\} = \min_{y_1 \neq y_2 \in C} |\{i : y_{1i} \neq y_{2i}\}|$$

We want $d$ to be large so that more errors can be tolerated, but this makes the number of vertices we can put in $\Sigma^n$ smaller. Therefore we have to sacrifice the rate $\frac{k}{n}$ to generate the same number of codeword. In many ways, coding theory is about exploring a tradeoff.

2 Linear Codes

In coding theory, a linear code is an error-correcting code for which any linear combination of codewords is still a codeword. Linear codes have the following advantages: i. easy to figure out the minimum distance; and ii. simple encoding and decoding algorithms.

Definition 1. (Linear code) Let $\Sigma = \mathbb{F}_q$ be a finite field with $q$ elements, then $C$ is linear if $\forall y_1, y_2 \in C \subseteq \mathbb{F}_q^n$, $y_1 + y_2 \in C$. In other words, let $G \in \mathbb{F}_q^{n \times k}$ be a full rank $n \times k$ matrix (making the map injective), then $\text{Enc}: \mathbb{F}_q^k \rightarrow \mathbb{F}_q^n$ becomes $x \mapsto Gx$, which defines a linear code with its generator matrix $G$.

Example. Let $q = 2$, $n = 3$ and $k = 2$. Then the generator matrix $G = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}$, so that
\[ G \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_1 + x_2 \end{pmatrix} \]. Thus \( C = \text{Im}(G) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \right\} \).

Note that for linear codes, we introduce the following notation \([n, k, d]_q\) henceforth, where \(n\) is the block length, \(k\) is the message dimension, and \(d\) is the minimum distance if known.

**Definition 2.** (Hamming weight) The Hamming weight of \(x \in \mathbb{F}_q^n\) in a linear code is denoted by \(\text{wt}(x) = \Delta(x, 0)\).

**Fact 1.** In a linear code, the minimum distance \(d\) is equal to the minimum Hamming weight of a nonzero codeword.

**Proof.**
\[
d = \min_{y_1 \neq y_2 \in C} \{\Delta(y_1, y_2)\} = \min_{y_1 \neq y_2 \in C} \{\Delta(y_1 - y_2, 0)\} = \min_{y = y_1 - y_2 \neq 0 \in C} \{\text{wt}(y)\}
\]

**Definition 3.** (Dual code) Given \([n, k]_q\) code \(C\), denote the orthogonal space \(C^\perp \triangleq \{y \in \mathbb{F}_q^n : y^T x = 0, \forall x \in C\}\) as the dual code of \(C\). Note that \(C^\perp\) has parameters \([n, n-k]_q\).

**Definition 4.** (Parity check matrix) The parity check matrix \(H\) of \(C\) is defined as an \((n-k) \times n\) matrix such that \(C^\perp = \text{Im(Enc)}\), where \(\text{Enc} : \mathbb{F}_q^{n-k} \to \mathbb{F}_q^n\) maps \(x\) to \(H^T x\). In other words, \(H^T\) is the generator matrix of \(C^\perp\).

**Example.** Reconsider the previous example, in which
\[
C = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}
\]

Therefore \(C^\perp = \left\{ \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}\) and \(H = (1, 1, 1)\).

**Fact 2.** \(y \in C \iff H y = 0\). (re-express the code as null space of the parity check matrix)

**Proof.** Notice that \(H^T\) is the generator matirx of \(C^\perp\), i.e. \(C^\perp\) is the row span of \(H\). Let
\[
H = \begin{pmatrix} h^T_1 \\ h^T_2 \\ \vdots \\ h^T_{n-k} \end{pmatrix}, \text{ then } Hx = 0 \iff \begin{cases} h^T_1 x = 0 \\ h^T_2 x = 0 \\ \vdots \\ h^T_{n-k} x = 0 \end{cases} \iff \forall a_1, a_2, \cdots, a_{n-k} \in \mathbb{F}_q, \left( \sum_{i=1}^{n-k} a_i h^T_i \right) x = 0 \iff \forall y \in C^\perp, y^T x = 0 \iff x \in (C^\perp)^\perp = C\]
Corollary 3. The minimum distance $d$ is the minimum number of columns in $H$ that are linearly dependent.

Proof. $d = \min_{y \neq 0 \in C} \{ wt(y) \} = \min \{ wt(y) \mid y \neq 0, Hy = 0 \}$. □

3 Hamming Code

Hamming code [1] is defined by the case of linear code that $q = 2$, which has excellent rate $\frac{k}{n} \approx 1$ but lower distance as we will see later.

Definition 5. (Hamming code) Let $r \in \mathbb{N}^+$. Define the parity check matrix of a Hamming code as

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

i.e. $H \in \mathbb{F}_2^{r \times (2^r-1)}$, which is spanned by all distinct $2^r - 1$ nonzero column vectors.

Example. For $r = 2$, $H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$, and $C = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$.

Theorem 4. Hamming code is $[2^r - 1, 2^r - 1 - r, 3]_2$ code.

Proof. We only need to prove $d = 3$, which is equivalent to say the minimum number of linearly dependent column is 3. Since 0 is not a column of $H$, every 2 columns are linearly independent. But there exists obviously triple of linearly dependent columns, such as, $\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$. □

Remark. Let $n = 2^r - 1$, then Hamming code is $[n, n - \log_2(n + 1), 3]_2$ code.

Since the distance is 3, Hamming code is uniquely decodable for up to $\left\lfloor \frac{3}{2} \right\rfloor = 1$ error. In fact, we can correct one error easily. Let $y \in C$ be any codeword, and $z = y + e_i$ be the received message. Then

$$Hz = H(y + e_i) = He_i$$
which is just the $i$ the column of $H$. Otherwise $Hz = 0$ implies that $y$ is not modified. For example, with $y = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ and $z = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$, $Hz = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. This indicates that index 3 has changed.

**Definition 6.** (Perfect code) $C$ is a perfect code if Hamming balls centered at codewords of radius $t$ (i.e. max errors) can partition $\Sigma^n$ exactly.

**Theorem 5.** Hamming code is perfect.

*Proof.* $\forall x \in \mathbb{F}_2^n$, if $Hx = 0$, then $x \in C$. Otherwise $Hx = h_i$, where $h_i$ is the $i$-th column of $H$. Hence $H(x + e_i) = 0$ and therefore $x + e_i \in C$. \qed

## 4 Hadamard Code

The *Hadamard code* is a code with extremely low rate but high distance. It is always used for error detection and correction when transmitting messages over very noisy or unreliable channels.

**Definition 7.** (Hadamard Code) Let $r \in \mathbb{N}^+$. The generator matrix of Hadamard code is a $2^r \times r$ matrix where the rows are all possible binary strings in $\mathbb{F}_2^r$.

**Example.** For $r = 2$, we have $G = \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix}$, which maps the messages to $Gx = \{ \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$. \[\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \]

**Fact 6.** Hadamard code is a $[2^r, r, 2^r - 1]_2$ code.

*Proof.* It suffices to prove the minimum weight of a nonzero codeword is $2^r - 1$. Let $x \neq 0 \in \mathbb{F}_2^r$.
$\mathbb{F}_2^n$, i.e. $\exists k$ s.t. $x_k = 1$. Then

\[
\frac{wt(Gx)}{2^r} = \mathbb{P}_{i \in [2^r]}[g_i^T x = 1] = \mathbb{P}_{y \in \mathbb{F}_2^r}[y^T x = 1] = \mathbb{P}_{y^* \in \mathbb{F}_2^r \setminus \{k\}, y_k \in \mathbb{F}_2}[y_k x_k + \sum_{i \neq k} y_i^* x_i = 1] = \mathbb{E}_{y^* \in \mathbb{F}_2^r \setminus \{k\}}[\sum_{i : i \neq k} y_i^* x_i = 1 + y_k] = \frac{1}{2}
\]

where $g_i^T$ denote the $i$-th row of $G$. \hfill \Box

**Remark.** In other words, Hadamard code is $[n, \log_2 n, \frac{n}{2}]_2$ code with $n = 2^r$.

**Reference**
