Lecture 14: Linear Codes: Examples and Properties
Given a field \((F, +, \cdot)\)

- We consider the set of all \(n\)-tuples with entries in \(F\)
- That is, we will consider the set \(F^n\)
- The total number of elements in the set if \(|F|^n\)
Basic Terminology

- A code $C \subseteq \mathbb{F}^n$
- A codeword $c = (c_1, \ldots, c_n) \in C$ such that all $c_1, \ldots, c_n \in \mathbb{F}$
- Block length is $n$
- Weight of a codeword: $\text{wt}(c) = |\{i : c_i \neq 0\}|$
- (Hamming) Distance $d_H(c, c') = \text{wt}(c - c')$
- Distance of a code: $d(C) = \min_{c, c' \in C} \min_{c \neq c'} \text{wt}(c - c')$
- $(N, K, d)$-code: $|C| = K$, $|\mathbb{F}|^n = N$ and $d(C) = d$
Given fixed $\mathbb{F}$ and $n$

We want to maximize $|C|$ and $d(C)$

- $|C|$ determines how much information can be transmitted over the channel, and
- $d(C)$ determines the robustness of the encoding (because, to force the maximum likelihood decoding algorithm to output an incorrect codeword, the channel needs to introduce at least $\lceil d(C)/2 \rceil$ errors)

We will see later that these two parameters are conflicting and there is a trade-off of these two parameters
Linear Codes

- Linear Code: If $C$ is a vector subspace of $\mathbb{F}^n$
- Suppose $(c_1, \ldots, c_k)$ is a basis of the vector subspace $C$
- $[n, k, d]_\mathbb{F}$ code: A code $C$ that is a vector subspace of $\mathbb{F}^n$, of dimension $k$, and $d(C) = d$
- The generator matrix $G$ of a code $C$ is defined to be a matrix in $\mathbb{F}^{k \times n}$ as defined below.

$$
G = \begin{pmatrix}
    c_1 \\
    c_2 \\
    \vdots \\
    c_k 
\end{pmatrix}
$$

- Note that $\begin{pmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{pmatrix} \cdot G$ generates all codewords in $C$, where $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$
Distance of a Linear Code

**Claim**

\[ d(C) = \min_{c \in C} \text{wt}(c) \]

**Proof.**

- Let \( d(C) \) is realized by the distance between the codewords \( c \) and \( c' \)
- Note that \( c - c' \) is also a codeword (because \( C \) is a vector space)
- Note that \( \text{wt}(c - c') = d(C) \)
- If there exists \( \tilde{c} \) such that \( \text{wt}(\tilde{c}) < d(C) \) then \( d_H(0, \tilde{c}) < d(C) \) (which is a contradiction)
- Therefore, we have the claim
The repetition code \( \{0^n, 1^n\} \) has generator matrix

\[
\begin{pmatrix}
1 & 1 & \cdots & 1 \\
n \text{-times}
\end{pmatrix}
\]

It is an \([n, 1, n]_2\) code
Let $G$ be the generator matrix of a code $C$

Let $G'$ be the generator matrix obtained by replacing the row $G_{i,*}$ in the matrix $G$ by the row $\alpha G_{i,*}$, for $\alpha \in \mathbb{F}^*$, then $G'$ generates the same code as $G$

Let $G'$ be the generator matrix obtained by replacing the row $G_{i,*}$ in the matrix $G$ by $G_{i,*} + \alpha G_{j,*}$, for $i \neq j$ and $\alpha \in \mathbb{F}^*$, then $G'$ generates the same code as $G$

We write $G \equiv G'$
Suppose $G'$ is a generator matrix obtained by swapping two columns of the generator matrix $G$

Then the code generated by $G'$ is similar to the code generated by $G$

We write $G \sim G'$

If $G$ is the generator matrix of an $[n, k, d]_F$ code, then $G'$ also generates an $[n, k, d]_F$ code (the codewords can be bijectively mapped where the mapping swaps the $i$-th and the $j$-th coordinate of the codeword)
Let $G$ be a generator matrix of an $[n, k, d]_\mathbb{F}$ code.

Then there exists $P \in \mathbb{F}^{k \times (n-k)}$ such that

$$G \sim \left[ I_{k \times k} \mid P \right],$$

where $I_{k \times k}$ is the identity matrix of dimension $k \times k$.

Proof Outline: Row rank = Column rank, swap $k$ independent columns of $G$ into its first $k$ columns, and perform Gaussian Elimination.
Define the matrix $H$ in $\mathbb{F}^{(n-k) \times n}$ as follows

$$H = \begin{bmatrix} -P^\top & I_{(n-k) \times (n-k)} \end{bmatrix}$$

**Claim**

The inner product of any row $G_{i,*}$ and any row $H_{j,*}$ is always 0.

**Proof.**

- Note that $G_{i,*} = (\delta_i, P_{i,*})$
- Note that $H_{i,*} = (-P^\top_{*,j}, \delta_j)$
- Their inner product is $-P_{i,j} + P_{i,j} = 0$
Claim

*Every codeword in the code generated by $G$ is orthogonal to every codeword in the code generated by $H$*
Proof.

- A codeword in the code generated by $G$ looks like
  \[
  \sum_{i=1}^{k} \alpha_i \cdot G_{i,*}, \text{ for } \alpha_1, \ldots, \alpha_k \in \mathbb{F}
  \]
- A codeword in the code generated by $H$ looks like
  \[
  \sum_{j=1}^{n-k} \beta_j \cdot H_{j,*}, \text{ for } \beta_1, \ldots, \beta_{n-k} \in \mathbb{F}
  \]
- Now, we have
  \[
  \left\langle \sum_{i=1}^{k} \alpha_i G_{i,*}, \sum_{j=1}^{n-k} \beta_j H_{j,*} \right\rangle = \sum_{i=1}^{k} \sum_{j=1}^{n-k} \alpha_i \beta_j \langle G_{i,*}, H_{j,*} \rangle
  \]
  \[
  = \sum_{i=1}^{k} \sum_{j=1}^{n-k} \alpha_i \beta_j \cdot 0 = 0
  \]
Let $C$ be the code generated by $G$

We denote the code generated by $H$ as $C^\perp$ (dual of $C$)

Show that $(C^\perp)^\perp = C$

We represent the $d(C^\perp)$ by $d^\perp$

Let $t(H)$ represent the minimum number of columns of $H$ that can be (non-trivially) linearly combined to yield the 0 column

Claim

\[
d(C) = t(H)
\]

Proof is left as an exercise
The columns of the generator matrix $G$ has all binary strings of length $r$.

For example, for $r = 3$, we have

\[
G = \begin{pmatrix}
0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 1
\end{pmatrix}
\]

This generates a $[2^r, r, 2^{r-1}]_2$ code (Prove this)
From the generator matrix of the Hadamard Code, we remove all those columns that have a 0 as their top-most entry.

For example, for \( r = 3 \), we have:

\[
G = \begin{pmatrix}
1 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
\end{pmatrix}
\]

This generates a \([2^{r-1}, r, 2^{r-2}]_2\) code (Prove this).
Simplex Code

- From the generator matrix of the Hadamard Code, we remove the all-0 column
- For example, for $r = 3$, we have
  \[
  G = \begin{pmatrix}
  0 & 0 & 0 & 1 & 1 & 1 & 1 \\
  0 & 1 & 1 & 0 & 0 & 1 & 1 \\
  1 & 0 & 1 & 0 & 1 & 0 & 1
  \end{pmatrix}
  \]
- This generates a $[2^r - 1, r, 2^{r-1}]_2$ code (Prove this)

Claim

Let $G$ be the generator matrix of the Simplex Code. We have $t(G) = 3$.

Prove this.
Hamming Code is the dual of the Simplex code
Therefore, it is a $[2^r - 1, 2^r - r - 1, 3]_2$ code (Prove this)
Write down the generator matrix for $r = 3$ case