TUTORIAL X

1 Error-correcting VS error-detecting codes

Show that the following statements are equivalent for a code C:

1. C has minimum distance $d \geq 2$.
2. If d is odd, C can correct $(d-1)/2$ errors.
3. If d is even, C can correct $d/2 - 1$ errors.
4. C can detect $d-1$ errors.
5. C can correct $d-1$ erasures (in the erasure model, the receiver knows where the errors have occurred).

2 Generalized Hamming bound

Prove the following bound: for any $(n,k,d)_q$ code $C \subseteq (\Sigma)^n$ with $|\Sigma| = q$,

$$k \leq n - \log_q \left(\sum_{i=0}^{\lfloor \frac{(d-1)}{2} \rfloor} \binom{n}{i} (q-1)^i \right)$$

3 Parity check matrix

Let C be a $[n,k,d]_q$-linear code and $G \in \mathbb{F}_q^{k \times n}$ be a generator matrix. That is, $C = \{xG, x \in \mathbb{F}_q^k\}$. We call a parity check matrix of the code C a matrix $H \in \mathbb{F}_q^{(n-k)\times n}$ such that for all $c \in \mathbb{F}_q^n$ we have $cH^T = 0$ if and only if $c \in C$. The objective of this exercise is to show how to construct a parity check matrix from a generator matrix.

1. Show that H is a parity check matrix if and only if $GH^T = 0$ and $\text{rank}(H) = n-k$.
2. Show that, from G we can construct a generator matrix G' of the form $G' = [I_k|P]$ for some $P \in \mathbb{F}_q^{k\times(n-k)}$. (If n is not optimal, we may have to permute the coefficients of the vectors).
3. Construct a parity check matrix from G'.
4. Construct a parity check matrix of the code given by the generator matrix $G = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}$ in $\mathbb{F}_2$.

4 (Optional) Almost-universal hash-functions: link between almost-universal hash-functions and codes with a good distance

A hash function is generally a function from a large space to a small one. A desirable property for a hash function is that there are few collisions. A family of functions $\{f_y\}_{y\in\mathcal{Y}}$ from $f_y : \mathcal{X} \to \mathcal{Z}$ is called ϵ-almost universal if for any $x \neq x'$, we have $\underset{y}{\mathbf{P}}\{f_y(x) = f_y(x')\} \leq \epsilon$ for a uniformly chosen $y \in \mathcal{Y}$. In other words, for any $x \neq x'$,

$$|\{y \in \mathcal{Y} : f_y(x) = f_y(x')\}| \leq \epsilon|\mathcal{Y}| \, . \tag{1}$$

The objective of the exercise is to show that almost-universal hash-functions and codes with a good distance are equivalent: from one you can construct the other efficiently.

Definition 4.1. *Let $\mathcal{H} = \{f_1, \ldots, f_n\}$ be a family of hash-functions, where for each $1 \leq i \leq n$, $f_i : \mathcal{X} \to \mathcal{Z}$. We define the code $C_\mathcal{H} = \mathcal{X} \to \mathcal{Z}^n$ by*

$$C_\mathcal{H}(x) = (f_1(x), \ldots, f_n(x))$$

for all $x \in \mathcal{X}$.

On the contrary, given a code $C : \mathcal{X} \to \mathcal{Z}^n$, we define the family of hash-functions $\mathcal{H}_C = \{f_1, \ldots, f_n\}$, from $\mathcal{X}$ to $\mathcal{Z}$ by

$$f_i(x) = C(x)_i$$

where $x \in \mathcal{X}$ and $C(x)_i$ is the i-th letter of $C(x)$ in the alphabet $\mathcal{Z}$.

1. Let $\mathcal{H} = \{f_1, \ldots, f_n\}$ be a family of ϵ-almost universal hash-functions. Prove that $C_\mathcal{H}$ has minimum distance $(1-\epsilon)n$.

2. On the other way, let C be a code from $\mathcal{X}$ to $\mathcal{Z}^n$ with minimum distance δn, prove that $\mathcal{H}_C$ is a family of $(1-\delta)$-almost universal hash-functions.

5 Hamming riddle

There are n people in a room, each of whom is given a black/white hat chosen uniformly at random (and independent of the choices of all other people). Each person can see the hat color of all other people, but not their own. Each person is asked if (s)he wishes to guess their own hat color. They can either guess, or abstain. Each person makes their choice without knowledge of what the other people are doing. They either win collectively, or lose collectively. They win if all the people who don't abstain guess their hat color correctly and at least one person does not abstain. They lose if all people abstain, or if some person guesses their color incorrectly. The goal below is to come up with a strategy that will allow the n people to win with pretty high probability

1. Argue that the n people can win with probability at least $\frac{1}{2}$

2. Lets say that a directed graph G is a subgraph of the n-dimensional hypercube if its vertex set is $\{0,1\}^n$ and if $u \to v$ is an edge in G, then u and v differ in at most one coordinate. Let $K(G)$ be the number of vertices of G with in-degree at least one, and out-degree zero. Show that the probability of winning the hat problem equals the maximum, over directed subgraphs G of the n-dimensional hypercube, of $K(G)/2^n$

3. Using the fact that the out-degree of any vertex is at most n,show that $K(G)/2^n$ is at most $\frac{n}{n+1}$ for any directed subgraph G of the n-dimensional hypercube.

4. Show that if $n = 2^r - 1$, then there exists a directed subgraph G of the n-dimensional hypercube with $K(G)/2^n = \frac{n}{n+1}$.
Hint: This is where the Hamming code comes in.