Telecom Paris ACCQ204, Coding Theory

SOLUTIONS TO ASSIGNMENT 4

Exercise 1. Show that if $C_{out} = [N, K, D]_{q^k}$ and $C_{in} = [n, k, d]_q$ are linear block codes, then the concatenated code $C_{out} \circ C_{in}$ is a linear block code $[nN, kK, D']_q$ where $D' \geq dD$.

Solution. That $C_{out} \circ C_{in}$ is a linear code if C_{out} and C_{in} are linear follows from defining the generator matrix of $C_{out} \circ C_{in}$ in terms of the generator matrices of C_{out} and C_{in}.

It is easy to check that for the concatenated code, the codeword length is Nn and the message set is of size at least q^{kK}. Next we show that the minimum distance is at least dD. Consider messages $m_1 \neq m_2$. Let the set of positions in which $C_{out}(m_1)$ and $C_{out}(m_2)$ differ be denoted T. Then, by the property of the outer code, we have

$$|T| = \Delta(C_{out}(m_1), C_{out}(m_2)) \geq D.$$

For $i \in T$, we have

$$\Delta(C_{in}(C_{out}(m_1)_i), C_{in}(C_{out}(m_2)_i)) \geq d.$$

Summing over all i yields

$$\Delta(C_{in}(C_{out}(m_1)), C_{in}(C_{out}(m_2))) \geq dD.$$

□

Exercise 2 (Zyablov bound)**.** We will show a way to design an explicit code which achieves positive rate and relative minimum distance with "low complexity." By low complexity we mean subexponential in the block length.

From Exercise 6 Assignment 2 there exists linear codes over $[q]$ whose asymptotic rate $r = \lim_{n\to\infty} \frac{k(n)}{n}$ and relative minimum distance $\delta = \lim_{n\to\infty} \frac{d(n)}{n}$ satisfy

$$r \geq 1 - H_q(\delta).$$

1. Argue that to find a length n code whose rate and relative minimum distance satisfy

$$r \geq 1 - H_q(\delta) - \varepsilon$$

   it takes $q^{O(kn)}$ time, as opposed to $q^{O(q^k n)}$ time if the code has no structure.

2. Consider concatenating a linear code approaching the GV bound and a Reed Solomon code. Show that such a construction yields an asymptotic rate

$$\mathcal{R} \geq \sup_{r\geq 0} r \left(1 - \frac{\delta}{H_q^{-1}(1 - r - \varepsilon)}\right)$$

   for any $\varepsilon > 0$, where δ represents the relative minimum distance of the concatenated code and where r denotes the rate of the inner code. This bound is called the Zyablov bound.

3. Plot and compare the Zyablov bound and the Gilbert-Varshamov lower bounds (rate as a function of relative minimum distance).

4. Argue that it is possibe to construct an explicit code achieving the Zyablov bound with time complexity $\mathcal{N}^{\mathcal{O}(\log \mathcal{N})}$ where $\mathcal{N}$ denotes the length of the concatenated code.

   Hence, although the Zyablov bound is lower than the GV bound, it is easier to construct a code that achieves the Zyablov bound (by concatenation) than to construct a linear code achieving the GV bound (which takes $O(q^{\mathcal{N}})$ time).

Solution. 1. Given a $k \times n$ generator matrix of a linear code, it takes it takes $O(q^k kn)$ time to generate each codeword (there are q^k codewords and each of them takes $O(kn)$ to be written using the generator matrix). Therefore it takes $O(q^k kn)$ to evaluate the minimum distance of a linear code. Since there are $q^{O(kn)}$ possible matrices, it takes $q^{O(kn)}O(q^k kn) = q^{O(kn)}$ to find a code with the desired minimum distance

   Follows from the fact that a linear code is characterized by its generator $k \times n$ q-ary matrix.

2. Let C_{in} approach the GV bound, hence

$$\delta_{in} \geq H_q^{-1}(1 - r - \varepsilon).$$

   Let C_{out} be a RS code therefore satisfying

$$\delta_{out} = 1 - R.$$

   The concatenated code $(\mathcal{R}, \delta)$ thus satisfies

$$\mathcal{R} = rR$$

   and

$$\delta \geq (1 - R)H_q^{-1}(1 - r - \varepsilon).$$

   Expressing R as a function of δ and r we get

$$R \geq 1 - \frac{\delta}{H_q^{-1}(1 - r - \varepsilon)}.$$

   Therefore we can achieve

$$\mathcal{R} \geq r \left(1 - \frac{\delta}{H_q^{-1}(1 - r - \varepsilon)}\right)$$

   and maximizing over r yields the desired result.

3. The Zyablov bound (rate vs relative minimum distance) is lower than the GV bound for any relative minimum distance within $(0, 1/2)$.

4. There are $q^{k^2/r}$ linear codes of rate $r = k/n$. Given such a code, it takes $O(q^k(k^2/r)k/r) = q^{O(k)}$ to generate all the codewords and compute their minimum weight. Therefore to find a linear code with desired rate and minimum distance it takes

$$q^{k^2/r}q^{O(k)} = q^{O(k^2)}$$

Since the linear code is used as an inner code we have $k = \log N$ where $N = q^t$ denotes the size of the RS code. Hence

$$q^{O(k^2)} = q^{O((\log N)^2)} = N^{O(\log N)}$$

which is upper bounded by $\mathcal{N}^{O(\log \mathcal{N})}$ where $\mathcal{N} = nN = N \log N$ denotes the length of the concatenated code.

□

Exercise 3 (Binary symmetric channel)**.** Let us examine the performance of linear codes against random errors. The binary symmetric channel with crossover probability $p < 1/2$ is defined by the following process: Given a codeword $\mathbf{c} \in \mathbb{F}_2^n$, we generate a random vector $\mathbf{y}$ where y_i is obtained by flipping c_i with probability p, independently of everything else. Equivalently,

$$\mathbf{y} = \mathbf{c} + \mathbf{z},$$

where $\mathbf{z}$ is a random vector whose components are independent and follow a Bernoulli(p) distribution. Here $\mathbf{y}$ is called the received vector, and $\mathbf{z}$ the noise vector.

We will measure the performance of a code $\mathcal{C} \subset \mathbb{F}_2^n$ of size 2^{nR} using the *average probability of error* under a minimum distance decoder $\mathrm{DEC}(\mathbf{y}) = \arg\min_{\mathbf{c}\in\mathcal{C}} d(\mathbf{y}, \mathbf{c})$:

$$\begin{aligned} P_e(\mathcal{C}) &= \frac{1}{2^{nR}} \sum_{\mathbf{c}\in\mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C}\backslash\{\mathbf{c}\} : \mathrm{DEC}(\mathbf{y}) = \mathbf{c}'] \\ &= \frac{1}{2^{nR}} \sum_{\mathbf{c}\in\mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C}\backslash\{\mathbf{c}\} : d(\mathbf{y}, \mathbf{c}') \le d(\mathbf{y}, \mathbf{c})], \end{aligned}$$

where $d(\cdot, \cdot)$ denotes Hamming distance. This is the average probability that there exists a codeword different from $\mathbf{c}$, that is closer to the received vector.

The goal of this and the next exercise is to show that for every $\epsilon > 0$ there exist linear codes of rate $R = 1 - H(p) - \epsilon$ whose probability of error is $2^{-\Omega(n)}$.

1. First, show that the Hamming distance between $\mathbf{y}$ and $\mathbf{c}$ is approximately np:

$$\Pr[d(\mathbf{c}, \mathbf{y}) > np(1 + \epsilon/2)] \le 2^{-\Omega(n)}$$

Hint: Find the probability that $\mathbf{z}$ has Hamming weight greater than $np(1 + \epsilon/2)$. You can use Chernoff bound, or directly compute the probability and then use Stirling's approximation.

2. Next, show that the probability of error can be bounded from above as $P_e(\mathcal{C}) \leq P_e^{(1)} + P_e^{(2)}$, where
$$P_e^{(1)} = \frac{1}{2^{nR}} \sum_{\mathbf{c} \in \mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C} \backslash \{\mathbf{c}\} : d(\mathbf{y}, \mathbf{c}') \leq np(1 + \epsilon/2)]$$
and
$$P_e^{(2)} = \Pr[d(\mathbf{c}, \mathbf{y}) > np(1 + \epsilon/2)] \leq 2^{-\Omega(n)}$$

3. Let us now find the probability of error for a random linear code obtained by choosing a generator matrix G uniformly. Show that for any two nonzero message vectors $\mathbf{u}_1 \neq \mathbf{u}_2$, the corresponding codeword $\mathbf{u}_1 G$ and $\mathbf{u}_2 G$ are statistically independent.

4. For fixed messages $\mathbf{u}_1 \neq \mathbf{u}_2$, show that
$$\Pr_{G,\mathbf{z}} \left[d(\mathbf{u}_1 G, \mathbf{u}_2 G + \mathbf{z}) < np(1 + \epsilon/2) \right] \leq 2^{-n(1 - H(p(1+\epsilon/2)) + o(1))}$$
Hint: First compute $\Pr_G \left[d(\mathbf{u}_1 G, \mathbf{x}) < np(1 + \epsilon/2) \right]$ for a fixed $\mathbf{x} \in \mathbb{F}_2^n$. Then average over $\mathbf{z}$.

5. Use part 4 to show that if $R < 1 - H(p) - \epsilon$, then $P_e^{(2)} = 2^{-\Omega(n)}$.

6. Combine everything to prove that there exists a linear code with rate $R \geq 1 - H(p) - \epsilon$ and $P_e = o(1)$.

Solution. 1. Easy application of Chernoff bound. The Hamming weight can be written as a sum of i.i.d. indicator random variables
$$\mathrm{wt}(\mathbf{z}) = \sum_{i=1}^{n} 1_{\{z_i = 1\}}.$$
The mean is equal to np. Using Chernoff bound,
$$\Pr[d(\mathbf{c}, \mathbf{y}) > np(1 + \epsilon/2)] \leq 2^{-\epsilon^2 n/3}.$$

2. The probability of error is
$$\begin{aligned}
&P_e(\mathcal{C}) \\
&= \frac{1}{2^{nR}} \sum_{\mathbf{c} \in \mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C} \backslash \{\mathbf{c}\} : \mathrm{DEC}(\mathbf{y}) = \mathbf{c}'] \\
&= \frac{1}{2^{nR}} \sum_{\mathbf{c} \in \mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C} \backslash \{\mathbf{c}\} : d(\mathbf{y}, \mathbf{c}') \leq d(\mathbf{y}, \mathbf{c}) | d(\mathbf{y}, \mathbf{c}) \leq np(1 + \epsilon/2)] \Pr[d(\mathbf{y}, \mathbf{c}) \leq np(1 + \epsilon/2)] \\
&\quad + \frac{1}{2^{nR}} \sum_{\mathbf{c} \in \mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C} \backslash \{\mathbf{c}\} : d(\mathbf{y}, \mathbf{c}') \leq d(\mathbf{y}, \mathbf{c}) | d(\mathbf{y}, \mathbf{c}) > np(1 + \epsilon/2)] \Pr[d(\mathbf{y}, \mathbf{c}) > np(1 + \epsilon/2)] \\
&\leq \frac{1}{2^{nR}} \sum_{\mathbf{c} \in \mathcal{C}} \Pr_{\mathbf{z}}[\exists \mathbf{c}' \in \mathcal{C} \backslash \{\mathbf{c}\} : d(\mathbf{y}, \mathbf{c}') \leq np(1 + \epsilon/2)] \\
&\quad + \Pr[d(\mathbf{y}, \mathbf{c}) > np(1 + \epsilon/2)]
\end{aligned}$$

3. If X and Y are independent random variables over $\mathbb{F}_2^n$ and X is uniformly distributed, then $X + Y$ is independent of Y and uniformly distributed. If $\mathbf{u}_1 \neq \mathbf{u}_2$, then there is at least one position where they differ. Therefore, $\mathbf{u}_1 G$ can be written as $\mathbf{u}_2 G + \mathbf{x}$, where $\mathbf{x}$ is a uniform random vector independent of $\mathbf{u}_2 G$.

   In a similar way, this can be extended to show that if $\mathbf{u}_1, \ldots, \mathbf{u}_k$ are linearly independent, then $\mathbf{u}_1 G, \ldots, \mathbf{u}_k G$ are all statistically independent and uniformly distributed.

4. If $\mathbf{u}_1 \neq \mathbf{u}_2$, we know from the previous part that $\mathbf{u}_1 G$ and $\mathbf{u}_2 G + \mathbf{z}$ are statistically independent. For any fixed $\mathbf{x}$,

$$\Pr_G \left[d(\mathbf{u}_1 G, \mathbf{x}) < np(1 + \epsilon/2) \right] = \binom{n}{np(1+\epsilon/2)} 2^{-n} \leq 2^{-n(1-H(p(1+\epsilon/2))+o(1))}$$

   Since this is true for every $\mathbf{x}$, we have

$$\begin{aligned}
&\Pr_{G,\mathbf{z}} \left[d(\mathbf{u}_1 G, \mathbf{u}_2 G + \mathbf{z}) < np(1 + \epsilon/2) \right] \\
&= \sum_{\mathbf{x}} \Pr_G \left[d(\mathbf{u}_1 G, \mathbf{x}) < np(1 + \epsilon/2) \middle| \mathbf{u}_2 G + \mathbf{z} = \mathbf{x} \right] \Pr[\mathbf{u}_2 G + \mathbf{z} = \mathbf{x}] \\
&\leq 2^{-n(1-H(p(1+\epsilon/2))+o(1))}
\end{aligned}$$

5. We have shown that for fixed $\mathbf{u}_1 \neq \mathbf{u}_2$,

$$\Pr_{G,\mathbf{z}} \left[d(\mathbf{u}_1 G, \mathbf{u}_2 G + \mathbf{z}) < np(1 + \epsilon/2) \right] \leq 2^{-n(1-H(p(1+\epsilon/2))+o(1))}$$

   But

$$P_e^{(2)} = \sum_{\mathbf{u} \in \mathbb{F}_2^{nR}} \frac{1}{2^{nR}} \Pr_{G,\mathbf{z}} \left[\exists \mathbf{u}_2 \neq \mathbf{u} : d(\mathbf{u}_1 G, \mathbf{u}_2 G + \mathbf{z}) < np(1 + \epsilon/2) \right]$$

   Taking union bound over all $\mathbf{u}_2 \in \mathbb{F}_2^{nR} \backslash \{\mathbf{u}\}$ gives

$$P_2^{(2)} \leq 2^{nR} 2^{-n(1-H(p(1+\epsilon/2))+o(1))} = 2^{-n(\epsilon - o(1))}$$

   if $R = 1 - H(p) - \epsilon$.

6. Follows from parts 1-5.