Coding Theory: Exercises

Many of the following problems are taken from the following sources:

1. Introduction to Coding Theory, a book by *J. H. van Lint*
2. Essential Coding Theory, a book by *Venkatesan Guruswami, Atri Rudra, and Madhu Sudan*

---

1 Exercises on the basics of error-correcting codes

1. Let Σ be a finite set of alphabets. A function on $d : \Sigma^n \times \Sigma^n \to \mathbb{R}$ is called a *metric* (or, *distance function*) if the following conditions are satisfied: For all $\mathbf{u}, \mathbf{v} \in \Sigma^n$

   (a) $d(\mathbf{u}, \mathbf{v}) \geq 0$.

   (b) $d(\mathbf{u}, \mathbf{v}) = 0$ if and only if $\mathbf{u} = \mathbf{v}$.

   (c) $d(\mathbf{u}, \mathbf{v}) = d(\mathbf{v}, \mathbf{u})$.

   (d) For any $\mathbf{w} \in \Sigma$, $d(\mathbf{u}, \mathbf{w}) \leq d(\mathbf{u}, \mathbf{v}) + d(\mathbf{v} + \mathbf{w})$ (Triangular Inequality).

   For any $\mathbf{u}, \mathbf{v} \in \Sigma^n$, the *Hamming distance* between $\mathbf{u}$ and $\mathbf{v}$, denoted by $\Delta(\mathbf{u}, \mathbf{V})$, is the number of positions in which $\mathbf{u}$ and $\mathbf{v}$ differ. Show that Hamming distance is a metric.

2. Let C be a code with distance d for even d. Then argue that C can correct up to $d/2 - 1$ many errors but cannot correct $d/2$ errors. Using this, show that if a code C can correct at most t errors then it has a distance $2t + 1$ or $2t + 2$.

3. In the following, we will see that one can convert arbitrary codes into code with slightly different parameters:

   (a) Let C be an $(n, k, d)_2$ code with d odd. Then it can be converted into an $(n+1, k, d+1)_2$ code.

   (b) Let C be an $(n, k, d)_\Sigma$ code. Then it can be converted into an $(n-1, k, d-1)_\Sigma$ code.

4. In this problem we will consider a noise model that has both errors and erasures. In particular, let C be an $(n, k, d)_q$ code over an alphabet Σ. As usual a codeword $\mathbf{c} \in C$ is transmitted over a noisy channel, and the received word $\mathbf{y} \in (\Sigma \cup \{?\})^n$, where as before the special symbol ? denotes an erasure. We will use s to denote the number erasures in $\mathbf{y}$ and e to denote the number of (non-erasure) errors that occurred during transmission. To decode such a vector means, given $\mathbf{y}$ as input, to output a codeword $\mathbf{c} \in C$ such that the number positions where $\mathbf{c}$ disagree with $\mathbf{y}$ is the $n - s$ non-erased positions is at most e. For the rest of the problem assume that

$$2e + s < d \tag{1}$$

   (a) Argue that the output of the decoder for any C under Equation 1 is unique.

(b) For the code C, assume that there exists a decoder D that can correct from $< d/2$ many error in $T(n)$ time. Then under Equation 1, one can perform decoding in time $O(T(n))$.

5. (**Guessing hat problem**) 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 persons, but not their own. Each person is asked if they wish 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 at least one person does not abstain and all people who do not abstain guess their hat color correctly. They lose if all people abstain, or if some person guess their color incorrectly. Your goal below is to come up with a strategy that will allow the n people to win with pretty high probability. We begin with a simple warm-up.

   (a) Argue that n people can win with probability ate least $\frac{1}{2}$.
   Next we will see how one can really bump up the probability of success with some careful modeling, and some knowledge of Hamming codes.

   (b) Let us say that a directed graph G is a subgraph of the n-dimensional hypercube if its vertex set is $\{0,1\}^n$ and $u \to v$ is an edge in G, then u and v must 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 $K(G)/2^n$.

   (c) 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.

   (d) 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.

2 Exercises on the basics of probability

In the following probability problems, the *sample space* D of a random event is a finite set, the events are the subsets of the sample space D, a *probability distribution* $p : D \to [0,1]$ is a function such that

$$\sum_{e \in D} p(e) = 1.$$

A *random variable* X is a function $X : D \to \mathbb{R}$, and the *expectation* of X is defined as

$$\mathrm{E}[X] = \sum_{x \in D} p(x) \cdot X(x),$$

and the *variance* of X is defined as

$$\mathrm{Var}[X] = \mathrm{E}[(X - E[X])^2].$$

For an event $E \subseteq D$, $\Pr[E]$ denotes the probability of happening that event, that is,

$$\Pr[E] = \sum_{e \in E} p(e).$$

1. Let $E_1, E_2, \ldots, E_n$ be n events on a finite domain D with the probability distribution p. Then, show the following.

   (a) (**Inclusion-Exclusion principle**)

   $$\Pr[\cup_{i=1}^{n} E_i] = \sum_{i=1}^{n} \Pr[E_i] - \sum_{1 \le i < j \le n} \Pr[E_i \cap E_j] + \cdots + (-1)^{n-1} \Pr[\cap_{i=1}^{n} E_i].$$

   (b) (**Union bound**)

   $$\Pr[\cup_{i=1}^{n} E_i] \le \sum_{i=1}^{n} \Pr[E_i].$$

2. Let E_1, E_2 be two events on a finite domain D with the probability distribution p. Then show that

   $$\Pr[E_1] = \Pr[E_1 \mid E_2] \cdot \Pr[E_2] + \Pr[E_1 \mid \overline{E_2}] \cdot \Pr[\overline{E_2}].$$

3. For a finite domain D, let u_D denotes the *uniform* distribution on D, i.e., for all $x \in D$, $u_D(x) = 1/|D|$.

   Let p_1 and p_2 be two probability distributions on the finite domains D_1 and D_2, respectively. Then, $p_1 \times p_2$ is a probability distribution on the domain $D_1 \times D_2$ defined as follows: For all $x \in D_1$ and $y \in D_2$, $p_1 \times p_2(x, y)$ is the probability of picking x from D_1 according to p_1 and picking y independently from D_2 according to p_2.

   Two distributions p_1 and p_2 over a finite domain D are called *identical* if for all $x \in D$, $p_1(x) = p_2(x)$.

   Then, show that for any positive integer m, the distribution $u_{D_1 \times D_2 \times \cdots \times D_m}$ is identical to the distribution $u_{D_1} \times u_{D_2} \times \cdots \times u_{D_m}$.

4. (**Linearity of Expectation**) Let $X_1, X_2, \ldots X_n$ be n random variables over a finite domain D with the probability distribution p. Then, show that

   $$\mathrm{E}[X_1 + \cdots + X_n] = \sum_{i=1}^{n} \mathrm{E}[X_i].$$

5. (**Indicator Random Variable**) Let D be a finite domain with the probability distribution p. A random variable $X : D \to \{0, 1\}$ is called *indicator random variable*. For any event E on D, let $\mathbf{1}_E$ denotes the following indicator random variable: For all $x \in D$,

   $$\mathbf{1}_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{otherwise .} \end{cases}$$

   Then for any event E, show that $\mathrm{E}[\mathbf{1}_E] = \Pr[E]$.

6. Let $x_1, x_2, \ldots x_k$ are k random numbers picked uniformly and independently from the set $[n] = \{1, 2, \ldots, n\}$. What is the expected number of *collisions*, i.e., unordered pairs $\{i, j\}$ such that $x_i = x_j$?

7. (**Markov Inequality**) Let D be a finite domain with the probability distribution p, and X is a *nonnegative* random variable on D. Then, show that for any $a > 0$,

   $$\Pr[X \ge a] \le \frac{\mathrm{E}[X]}{a}$$

8. (**Chebyshev Inequality**) Let D be a finite domain with the probability distribution p, and X is a random variable on D. Then,

$$\Pr[|X - \mathrm{E}[X]| \geq a] \leq \frac{\mathrm{Var}[X]}{a^2}.$$

In particular,

$$\Pr[|X - E[X]| \geq k \cdot \sqrt{\mathrm{Var}[X]}] \leq \frac{1}{k^2}.$$

9. (**Chernoff Bound**) Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed (i.i.d) binary random variables (i.e., the range of each of the random variable X_i is $\{0,1\}$) and $X = \sum_{i=1}^{n} X_i$. Then, for all $\epsilon \in (0,1]$, the *multiplicative* Chernoff bound states that,

$$\Pr[|X - \mathrm{E}[X]| > \epsilon \mathrm{E}[X]] < 2e^{-\frac{\epsilon^2 \mathrm{E}[X]}{3}},$$

and the *additive* Chernoff bound states that

$$\Pr[|X - \mathrm{E}[X]| > \epsilon n] < 2e^{-\frac{\epsilon^2 m}{2}}.$$

10. Let $G(V,E)$ be a random graph on n vertices constructed as follows: For all $\{u,v\}$, with probability $1/2$, $\{u,v\} \in E$. Let X be the random variable denoting the number of triangles in G. Compute the $\mathrm{E}[X]$ and $\mathrm{Var}[X]$. Calculate the best possible upper bound for $\Pr[X \geq (1+\epsilon)E[X]]$.

11. For a biased coin, let

$$|\Pr[\mathrm{HEAD}] - \Pr[\mathrm{TAIL}]| = \epsilon,$$

for some $\epsilon \in (0, 1/2)$. Using as minimum as possible coin tosses, design a random procedure such that it tells whether $\Pr[\mathrm{HEAD}] > \Pr[\mathrm{TAIL}]$ with probability $1/100$. Justify your answer.

12. Let $G(V,E)$ be an undirected graph $2n$ vertices and m edges. Then the vertex V can be partitioned into two sets A and B such that the number of edges across these two sets is at least $m/2$. Furthermore, show that V can even be partitioned into two sets A and B such that the number of edges across these two sets is at least

$$\frac{n}{2n-1}m.$$

13. A 3-CNF formula over variables $x_1, x_2, \ldots x_n$ is of the form

$$\Phi(x_1, x_2, \ldots, x_n) = \wedge_{i=1}^{m} \left(v_{i_1} \vee v_{i_1} \vee v_{i_3} \right),$$

where each v_{i_j} is either a variable x_i or its negation $\overline{x_i}$. The terms $(v_{i_1} \vee v_{i_1} \vee v_{i_3})$ in the formula Φ are called *clauses*. Show that given any such 3-CNF Φ over n-variables and m clauses, there exists an assignment $\mathbf{a} \in \{0,1\}^n$ on the variables such that it *satisfies* at least $7m/8$ clauses of Φ.

14. Let C be a coin such that the probability of showing head is p. Suppose that C is tossed m times, and Δ_p is the probability of obtaining an odd number of heads. Then, show the following:

    (a) $\Delta_p = \frac{1}{2} \cdot (1 - (1-2p)^m)$.

    (b) If m is odd, then Δ_p is a non-decreasing function of p.

    (c) Over $p \in [0, 1/2]$, Δ_p is a non-decreasing function of p.

    Note: Observe that $\Delta_p = \Pr[\mathrm{TAIL}] - \Pr[\mathrm{HEAD}] = 1 - 2p$, which can be thought as the bias of the coin, and the above problem describes a procedure to reduce that bias.

3 Exercises on the basics of finite fields and linear spaces

1. This exercise aims to prove that the multiplicative group of a finite field is *cyclic*. We will prove this via the following sequence of exercises.

   (a) For every element a in a finite group G, $a^{|G|} = 1$ [1].

   (b) Let G be a finite commutative abelian group. Let a and b be two elements in G such that the order[2] of a and b are m and n, respectively. Then, show that G has an element of order $\mathrm{lcm}(m, n)$.

   (c) Let G be a finite commutative abelian group such that for every positive integer n, the number of elements a with $a^n = 1$ is at most n. Then, G is cyclic.

   (d) Prove that the multiplicative group of any finite field is cyclic. Hence, for any element α in finite field $\mathbb{F}_q$, $\alpha^q = \alpha$.

2. Let $\mathbb{F}$ be a field and let $f(x)$ be an irreducible polynomial in $\mathbb{F}[x]$ of degree d. Then, show that for every polynomial $g(x)$ of degree less than d, there exists a polynomial $h(x)$ of degree less than d such that $g(x) \cdot h(x) = 1 \mod f(x)$. Using this, show that the set of all polynomials of degree less than d forms a field under polynomial addition and multiplication modulo $f(x)$.

3. For any prime q with $q \equiv 1 \mod 4$, show that $\mathbb{F}_q$ has an element $\alpha \in \mathbb{F}_q$ such that $\alpha^2 = -1$.

4. Over a finite field $\mathbb{F}_q$, an element $\alpha \in \mathbb{F}_q$ is called *quadratic residue* if $\alpha = \beta^2$ for some $\beta \in \mathbb{F}_q$. Otherwise, α is called *quadratic non-residue* in $\mathbb{F}_q$. Then, for any prime q with $q \equiv 3 \mod 4$, show that there exists two quadratic residues α and β in $\mathbb{F}_q$ such that $\alpha + \beta = -1$

4 Exercises on the basics of linear codes

1. Let G be a generator matrix of an $[n, k, d]_2$ code. Then, show that G has at least kd ones in it.

2. For any $[n, k, d]_2$ code, either all of the codewords begin with a zero or exactly half of the codewords begin with a zero.

3. For some fixed $c \geq \mathbb{Z}_{>2}$, let $\{G_n\}_{n \in \mathbb{Z}_{>c}}$ be a family of c-regular connected graphs with the number of vertices of G_n is n and the girth [3] is ℓ_n. Then, using the graph family, construct a code family $\mathcal{C} = \{C_i\}_{i \in \mathbb{Z}_{>c}}$ where each C_i is an $[n_i, k_i, d_i]_2$ code with

$$n_i = \frac{ci}{2}, \quad k_i = \left(\frac{c}{2} - 1\right) i + 1, \quad \text{and } d_i = \ell_i.$$

   Now we show that such a code family $\mathcal{C}_{i \in \mathbb{Z}_{>c}}$ can *not* be asymptotically good. In particular, we will show that $\ell_n = O(\log n)$ [4], which will imply that the relative distance of $\mathcal{C}$ is zero.

   When $\ell_n = 2t + 1$: Show the following.

   (a) For any vertex u of G_n, the induced subgraph by all the vertices within distance at most $t - 1$ from u is a tree.

[1] Here, 1 denotes the identity element in G.

[2] The order of an element in G is the smallest positive integer i such that a^i is the identity element in G

[3] The length of the shortest cycle in a graph is called its *girth*.

[4] This is known as *Moore bound*.

(b) $n \geq 1 + c\Sigma_{k=1}^{t-1}(c-1)^k$.

(c) $\ell_n = O(\log n)$.

When $\ell_n = 2t$: Show the following.

(a) For any edge $\{u, v\}$ of G, let T_u be the induced subgraph in the graph in the $G - v$ [5] formed by all the vertices within distance at most $t-1$. Similarly, define T_v. Then, both T_u and T_v are tree. Furthermore, the vertex sets of T_u and T_v are disjoint.

(b) $n \geq 2\sum_{k=0}^{t-1}(c-1)^k$.

(c) $\ell_n = O(\log n)$.

4.1 Exercises on derived codes

4. The set of all $n_2 \times n_1$ matrices over $\mathbb{F}_2$ forms a vector space V of dimension $n_1 n_2$. For $i = 1, 2$, let C_i be an $[n_i, k_i, d_i]_2$ linear code over $\mathbb{F}_2$. Let C be the subsets of V consisting of those matrices for which every column, respectively every row, is a codeword in C_1, respectively C_2. Show that C is an $[n_1 n_2, k_1 k_2, d_1 d_2]_2$ code. The code C is called *direct product* of C_1 and C_2.

5. Let C be an $[n, k]_q$ code. Define a function $f : C \to \mathbb{F}_q^{n^m}$ as follows: For $\mathbf{c} = (c_1, c_2, \ldots, c_n)$,

$$f(\mathbf{c}) = \left(c_{i_1} + c_{i_2} + \cdots + c_{i_m}\right)_{i_1, i_2, \ldots, i_m \in [n]}.$$

Then, show that

$$f(C) = \{f(\mathbf{c}) \mid \mathbf{c} \in C\}$$

is an $[n^m, k]_q$ code. Furthermore, given a generator matrix G, describe a generator matrix for $f(C)$.

4.2 Exercises on dual codes

6. For any $[n, k, n-k+1]_q$ code, show that its dual is an $[n, n-k, k+1]_q$ code.

7. Let $S \subseteq \mathbb{F}_q^n$. For any $I \subseteq [n]$ with $|I| = t$, $\phi_I : S \to \mathbb{F}_q^t$ is defined as follows: for all $(c_1, c_2, \ldots, c_n) \in S$,

$$\phi_I : (c_i)_{i \in [n]} \mapsto (c_i)_{i \in I}.$$

The set S is called *t-wise independent* if for all $I \subseteq [n]$ with $|I| = t$ and for all $\mathbf{a}, \mathbf{b} \in \mathbb{F}_q^t$,

$$|\phi_I^{-1}(\mathbf{a})| = |\phi_I^{-1}(\mathbf{b})|.$$

In other words, if one picks a vector $(s_1, s_2, \ldots, s_n)$ from S uniformly at random, then for any $I = \{i_1, i_2, \ldots, i_t\} \subseteq [n]$, the $s_{i_1}, s_{i_2}, \ldots, s_{i_t}$ are uniformly and independently random over $\mathbb{F}_q$.

For that for any C whose dual $C^\perp$ has distance $d^\perp$ is $(d^\perp - 1)$-wise independent.

[5] It denotes the graph we obtain after deleting the vertex v from the graph G.