Algorithmic Coding Theory: Midsem

Instructions

You may assume basic facts of linear algebra, combinatorics, and finite fields.

Each question carries 20 marks.

Questions

1. (Erasure) Let C be a $(n, k, d)_\Sigma$ code. A codeword $\mathbf{c} \in C$ is transmitted over a channel and the received word $\mathbf{y} \in (\Sigma \cup \{?\})^n$, where the symbol "?" denotes an erasure. Let s be the number of erasures in $\mathbf{y}$ and let e be the number of non-erasure errors that occurred during the transmission. To decode $\mathbf{y}$ means to output a codeword $\mathbf{c} \in C$ such that the number of positions where $\mathbf{c}$ disagree with $\mathbf{y}$ in the non-erased positions is at most e.

   (a) (5 points) Assume that
   $$2e + s < d. \tag{1}$$
   Argue that the output of the decoder for any code C under (1) is unique.

   (b) (15 points) Let $\mathcal{C}$ be an $[n, k, d = n - k + 1]_q$ Reed-Solomon code. Use the Berlekamp-Welch algorithm to give a polynomial time algorithm that decodes $\mathcal{C}$ from e errors and s erasures as long as $2e + s < d$.

2. (20 points) Consider the bivariate version of the Reed-Solomon code, which encodes a polynomial $f \in \mathbb{F}_q[x, y]$ with degree less than k in both x and y by its evaluations at all q^2 points $(\alpha, \beta) \in \mathbb{F}_q \times \mathbb{F}_q$

   (a) (6 points) What are the block length, dimension, and minimum distance of this code?

   (b) (14 points) Describe how one can efficiently decode this code up to (almost) half its minimum distance.

3. (Graph-theoretic Proof for the GV Bound): Let $q \geq 2$ and $1 \leq d \leq n(1 - 1/q)$. Recall that the GV bound shows that for any $0 < \epsilon \leq 1 - H_q(d/n)$, there exists a code with rate $R \geq 1 - H_q(d/n) - \epsilon$ and distance d. The proof taught in the class was based on greedy construction. In this problem, we are going to prove the GV bound using graph-theoretic approach.

   Let us start with some definitions. Consider the graph $G_{n,d,q} = (V, E)$ where the vertex set $V = [q]^n$, *i.e.*, the set of all vectors. For any two distinct vector $\mathbf{u} \neq \mathbf{v} \in V$, there is an edge $(\mathbf{u}, \mathbf{v}) \in E$ if and only if $\Delta(\mathbf{u}, \mathbf{v}) < d$.

   (a) (5 points) Recall that an *independent set* of a graph G is a subset of vertices such that there is no edge between any pair of them. Show that any independent set C of $G_{n,d,q}$ is a q-ary code of distance d.

(b) (8 points) Recall that the *degree* of a vertex in a graph G is the number of edges incident to that vertex. Let Δ be the largest degree in a graph $G = (V, E)$, show that G has an independent set of size at least $\frac{|V|}{\Delta+1}$.

(c) (7 points) Prove the GV bound using the above graph-theoretic view of a code.

4. (Chinese remainder code): In this problem, we will consider the number-theoretic counterpart of Reed-Solomon codes. Let $1 \leq k < n$ be integers and let $p_1 < p_2 < \cdots < p_n$ be n distinct primes. Denote $K = \prod_{i=1}^{k} p_i$ and $N = \prod_{i=1}^{n} p_i$. The notation $\mathbb{Z}_M$ stands for integers modulo M, i.e., the set $\{0, 1, \ldots, M-1\}$. Consider the *Chinese remainder code* defined by the encoding map $E : \mathbb{Z}_K \to \mathbb{Z}_{p_1} \times \mathbb{Z}_{p_2} \times \cdots \times \mathbb{Z}_{p_n}$ defined by:

$$E(m) = (m \mod p_1,\ m \mod p_2,\ \ldots,\ m \mod p_n).$$

(Note that this is not a code in the usual sense we have been studying since the symbols at different positions belong to different alphabets. Still, notions such as distance of this code make sense and are studied in the question below.)

(a) (10 points) Suppose that $m_1 \neq m_2$. For $1 \leq i \leq n$, define the indicator variable $b_i = 1$ if $E(m_1)_i \neq E(m_2)_i$ and $b_i = 0$ otherwise. Prove that $\prod_{i=1}^{n} p_i^{b_i} > N/K$.

(b) (10 points) Use the above to deduce that when $m_1 \neq m_2$, the encodings $E(m_1)$ and $E(m_2)$ differ in at least $n - k + 1$ locations.

5. (20 points) Different proofs of singleton bound:

(a) (10 points) If $\mathcal{C}$ is an $[n, k]$ linear code over the field $\mathbb{F}$, then distance of the code is $\leq n - k + 1$. Give two different proofs of this fact using parity-check and generator matrices respectively. You may assume basic linear algebraic facts.

(b) (6 points) Can you prove the same bound if $\mathcal{C}$ is not a linear code?

(c) (4 points) You have known the following fact since kindergarten. Any $n + 1$ vectors in $\mathbb{F}^n$ are always linearly dependent over field $\mathbb{F}$. Can you prove this over $\mathbb{Q}$ using the pigeon-hole principle? (Hint: First, deal with the finite field case.)