15-859V: Introduction to coding theory
Carnegie Mellon University

Spring 2010
Venkatesan Guruswami

PROBLEM SET 2
Due by Friday, March 19

INSTRUCTIONS

- This problem sets can be turned in groups of two people; i.e., a single write-up for each two person team suffices. If you prefer, you can also work alone (see the last bullet item for some "credit" for doing so). Solutions typeset in LaTeX are preferred.
- The generous time window of almost 4 weeks for the problem set should give you ample time to think carefully about the problems. Therefore, you are *strongly urged* to try and solve the problems without consulting *any* reference material other than the course notes and what we cover in class. If for some reason you feel the need to consult some source (such as a textbook or sources on the web), *please acknowledge the source* and try to articulate the difficulty you couldn't overcome before consulting the source and how it helped you overcome that difficulty. Alternatively, before consulting any such material, I encourage you to ask me for a hint, preferably by posting a comment on the blog post dedicated to this problem set, so all students can take advantage of any hints.
- Please use the comments section of the blog for any questions or clarifications about the problems.
- Please start work on the problem set early. The problem set has **six** problems and is worth a total of 100 points. As a rough estimate, scoring around $3/4$'th of the points, or $2/3$'rd of the points if you turn in solutions solo, suffices for an A on this problem set.

---

1. ($5+3+8=16$ points) For integers $1 \leqslant k \leqslant n$, call a subset $S \subseteq \{0,1\}^n$ to be k-wise independent if for every $1 \leqslant i_1 < i_2 < \cdots < i_k \leqslant n$ and $(a_1, a_2, \ldots, a_k) \in \{0,1\}^k$

$$\mathrm{Prob}_{x \in S}[x_{i_1} = a_1 \wedge x_{i_2} = a_2 \wedge \cdots \wedge x_{i_k} = a_k] = \frac{1}{2^k}$$

where the probability is over an element x chosen uniformly at random from S. In this problem, you will see how codes can be used to construct k-wise independent sets of small size.

   (a) Let $H \in \mathbb{F}_2^{t \times n}$ be the parity check matrix of an $[n, n-t, d]_2$ binary linear code of distance $d \geqslant k+1$. Define $S = \{x^T H \mid x \in \mathbb{F}_2^t\}$. Prove that S is a k-wise independent set of size 2^t.

   (b) Using the above, show how one can construct a k-wise independent subset of $\{0,1\}^n$ of size at most $(2n)^{\lceil \frac{k}{2} \rceil}$.

   (c) Prove an almost matching lower bound, namely any k-wise independent set $S \subseteq \{0,1\}^n$ satisfies

$$|S| \geqslant \sum_{i=0}^{\lfloor \frac{k}{2} \rfloor} \binom{n}{i}. \tag{1}$$

   Suggestion: Find an orthonormal set of vectors in $\mathbb{R}^{|S|}$ of cardinality at least the R.H.S of (1).

2. ($7+9=16$ points) In this exercise, you will prove that certain sparse and structured polynomials are irreducible.

(a) Let q be a prime. Prove that $f(X) = X^q - X - 1$ is an irreducible polynomial in $\mathbb{F}_q[X]$.

(b) Let $\mathbb{F}_q$ be any finite field and $\alpha \in \mathbb{F}_q$ be a non-zero element of order k (i.e., $\alpha^k = 1$ but $\alpha^i \neq 1$ for $1 \leqslant i < k$) which satisfies $\gcd(k, \frac{q-1}{k}) = 1$. Prove that $f(X) = X^k - \alpha$ is irreducible in $\mathbb{F}_q[X]$.

(Suggestion: In both cases, use the fact that if there is a factor of $f(X)$ of degree b, then there is a root ζ of the polynomial f satisfying $\zeta^{q^b} = \zeta$.)

3. (12 points) In this problem, you will prove that a certain "ultimate" form of Reed-Solomon decoding is NP-hard.

You may assume that the following problem is NP-hard.

Instance: A set $S = \{\alpha_1, \ldots, \alpha_n\} \subseteq \mathbb{F}_{2^m}$, an element $\beta \in \mathbb{F}_{2^m}$, and an integer $1 \leqslant k < n$.

Question: Is there a nonempty subset $T \subseteq \{1, 2, \ldots, n\}$ with $|T| = k + 1$ such that $\sum_{i \in T} \alpha_i = \beta$?

Consider the $[n, k]$ Reed-Solomon code C_{RS} over $\mathbb{F}_{2^m}$ obtained by evaluating polynomials of degree at most $k - 1$ at points in S. Define $y \in (\mathbb{F}_{2^m})^n$ as follows: $y_i = \alpha_i^{k+1} - \beta\alpha_i^k$ for $i = 1, 2, \ldots, n$.

Prove that there is a codeword of C_{RS} at Hamming distance at most $n - k - 1$ from y if and only if there is a set T as above of size $k + 1$ satisfying $\sum_{i \in T} \alpha_i = \beta$.

Conclude that finding the nearest codeword in a Reed-Solomon code over exponentially large fields is NP-hard.

4. ($5 + 4 + 7 = 16$ points) Let C be an $[n, k, n - k + 1]_q$ linear code for some $1 \leqslant k < n$.

(a) Prove that $C^\perp$ is an $[n, n - k, k + 1]_q$ linear code.

(b) Prove that the number of codewords of C of weight $d = n - k + 1$ equals $(q - 1)\binom{n}{d}$.

(c) Let W_i be the number of codewords of C of weight i for $n - k + 1 \leqslant i \leqslant n$. Show that

$$\sum_{i=n-k+1}^{n-t} \binom{n-i}{t} W_i = \binom{n}{t}(q^{k-t} - 1), \quad 0 \leqslant t < k\,,$$

and that the solution to the above set of equations is unique.

(Suggestion: One approach is to use the q-ary MacWilliams identities (stated in Exercise 2 of Notes 5.1. You can assume these identities without proof.)

5. (20 points) In this problem, we will look at some binary "BCH-like" subfield subcodes of Reed-Solomon codes that meet the Gilbert-Varshamov bound.

Let $\mathbb{F} = \mathbb{F}_{2^m}$. Fix positive integers d, n with $(d-1)m < n < 2^m$, and a set $S = \{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ of n distinct nonzero elements of $\mathbb{F}$. For a vector $\mathbf{v} = (v_1, \ldots, v_n) \in (\mathbb{F}^*)^n$ of n not necessarily distinct nonzero elements from $\mathbb{F}$, define the *Generalized Reed-Solomon code* $\mathrm{GRS}_{S,\mathbf{v},d}$ as follows:

$$\mathrm{GRS}_{S,\mathbf{v},d} = \{(v_1 \cdot p(\alpha_1), v_2 \cdot p(\alpha_2), \ldots, v_n \cdot p(\alpha_n)) \mid p(X) \in \mathbb{F}[X] \text{ has degree } \leqslant n - d\}\,.$$

(a) Show that $\mathrm{GRS}_{S,\mathbf{v},d}$ is an $[n, n - d + 1, d]_{\mathbb{F}}$ linear code.

(b) Argue that $\mathrm{GRS}_{S,\mathbf{v},d} \cap \mathbb{F}_2^n$ is a binary linear code of rate at least $1 - \frac{(d-1)m}{n}$.

(c) Let $\mathbf{c} \in \mathbb{F}_2^n$ be a nonzero binary vector. Prove that (for every choice of d, S) there are at most $(2^m - 1)^{n-d+1}$ choices of the vector $\mathbf{v}$ for which $\mathbf{c} \in \mathrm{GRS}_{S,\mathbf{v},d}$.

(d) Prove that if the integer D satisfies

$$(2^m - 1)^{d-1} \geqslant \sum_{i=0}^{D-1} \binom{n}{i} ,$$

then there exists a vector $\mathbf{v} \in (\mathbb{F}^*)^n$ such that the minimum distance of the binary code $\mathrm{GRS}_{S,\mathbf{v},d} \cap \mathbb{F}_2^n$ is at least D.

(e) Using Parts (5b) and (5d), show how to conclude that the family of codes $\mathrm{GRS}_{S,\mathbf{v},d} \cap \mathbb{F}_2^n$ contains binary linear codes that meet the Gilbert-Varshamov bound.

6. (20 points) For a finite field $\mathbb{F}_{q^m}$, define the Trace map Tr as follows: for $x \in \mathbb{F}_{q^m}$

$$\mathrm{Tr}(x) = x + x^q + x^{q^2} + \cdots + x^{q^{m-1}} .$$

(a) Show that $\mathrm{Tr}(\alpha) \in \mathbb{F}_q$ for every $\alpha \in \mathbb{F}_{q^m}$.

(b) Prove that $\mathrm{Tr} : \mathbb{F}_{q^m} \to \mathbb{F}_q$ is a *surjective* $\mathbb{F}_q$-linear map.

(c) Let $C \subseteq \mathbb{F}_{q^m}^n$ be a linear code, and $C^\perp \subseteq \mathbb{F}_{q^m}^n$ its dual. Define $C_{|\mathbb{F}_q} = C \cap \mathbb{F}_q^n$ to be the subfield subcode of C.

Prove that

$$(C_{|\mathbb{F}_q})^\perp = \mathrm{Tr}(C^\perp)$$

where $\mathrm{Tr}(C^\perp) = \{\mathrm{Tr}(c) \mid c \in C^\perp\}$.

Suggestion: Prove both the set inclusions, starting with the easier inclusion $\mathrm{Tr}(C^\perp) \subseteq (C_{\mathbb{F}_q})^\perp$.

(d) Show that

$$\dim(C) \leqslant \dim(\mathrm{Tr}(C)) \leqslant m \cdot \dim(C) ,$$

and

$$\dim(C) - (m-1)(n - \dim(C)) \leqslant \dim(C_{|\mathbb{F}_q}) \leqslant \dim(C) ,$$

where for a linear space $X \subseteq \mathbb{F}^n$, $\dim(X)$ stands for its dimension as a $\mathbb{F}$-vector space.