Telecom ParisTech — ACCQ204, Coding Theory
Teacher: Aslan Tchamkerten — December 2017

# Assignment 3 - Solutions

**Exercise 1.** Consider an $[n, k, d]$ MDS code over $\mathbb{F}_q$. Show that

1. the number of codewords of weight $d$ is

$$N_d = \binom{n}{d}(q-1).$$

   Hint. Pick a subset of $k-1$ coordinates and fix the corresponding values to zero. Pick any other coordinate and let the symbol value in this coordinate run through all $q$ symbols in $\mathbb{F}_q$.

2. Show that the number of codewords of weight $d+1$ is

$$N_{d+1} = \binom{n}{d+1}\left((q^2-1) - \binom{d+1}{d}(q-1)\right).$$

*Solution.*

1. Because the code is MDS, for any given $k$ coordinates, the components correspond to codewords in a one-to-one manner, that is they span every of the $q^k$ components. Now, pick arbitrary $k-1$ components and fix the corresponding values to zero. Because of the previous argument, this set of $k-1$ zero components is consistent with at least one other codeword. Now, pick another component. To any non-zero value of this component corresponds a unique codeword whose weight is at most $n-(k-1)$, but since the minimum weight is $d$, they all have weight $d$. Hence, for each subset of $k-1$ coordinates we get $q$ non-zero codewords of weight $d$. In total we thus have $(q-1)\binom{n}{k-1} = (q-1)\binom{n}{d}$.

2. Consider any subset of $d+1 = n-k+2$ coordinates. Take two of these coordinates and combine them with the remaining $k-2$ coordinates to form an information set. Fix the components in the $k-2$ coordinates to zero, and let the remaining two coordinates run freely through $\mathbb{F}_q$.These $q^2$ information set combinations must correspond to $q^2$ codewords. (In fact, we may view this subset of codewords as a shortened (d +1,2,d) MDS code.) One of these codewords must be the all-zero codeword, since the code is linear. The remaining $q^2-1$ codewords must have weight $d$ or $d+1$. Since there are $q-1$codewords of weight $d$ with support in any subset of d coordinate positions, the number of codewords of weight $d$ whose support is in any subset of $d+1$ coordinate positions is $\binom{d+1}{d}(q-1)$ (the number of codewords of weight $d$ in a $(d+1,2,d)$ MDS code). So the number of codewords of weight $d+1$ in any $d+1$ coordinate positions is

$$(q^2-1) - \binom{d+1}{d}(q-1).$$

   Since there are $n$ distinct subsets of $d+1$ coordinate positions, the given expression for $N_{d+1}$ follows.

□

**Exercise 2.** Construct an $RS(n = 4, k = 2)$ code. For the construction you may want to consider the irreducible polynomial $x^2 + x + 1$ over $\mathbb{F}_2$ and the evaluation points (to be justified) $\alpha_1 = 0$, $\alpha_2 = 1$, $\alpha_3 = x$, $\alpha_4 = x + 1 = x^2$.

*Solution.* Since $n = 4$ we need a base field with (at least) $4$ elements. So let's choose the base field $\mathbb{F}_4 = \mathbb{F}_2[X]/(x^2 + x + 1)$ whose elements are thus

$$\{0, 1, x, x + 1 = x^2\}.$$

Since $k = 2$, the message polynomials are of degree $k - 1 = 1$ and can be written as $f_0 + f_1 x$ with $f_0, f_1 \in \mathbb{F}_4$. Thus the mapping between information symbols and codewords is given by

$$(f_0, f_1) \to (f_0 + f_1\alpha_1, f_0 + f_1\alpha_2, f_0 + f_1\alpha_3, f_0 + f_1\alpha_4).$$

The full mapping is thus

$$\begin{array}{cccccccccccccccc}
0 & 0 & \to & (0 & 0 & 0 & 0) & & x & 0 & \to & (x & x & x & x) \\
0 & 1 & \to & (0 & 1 & x & x+1) & & x & 1 & \to & (x & x+1 & 0 & 1) \\
0 & x & \to & (0 & x & x+1 & 1) & & x & x & \to & (x & 0 & 1 & x+1) \\
0 & x+1 & \to & (0 & x+1 & 1 & x) & & x & x+1 & \to & (x & 1 & x+1 & 0) \\
1 & 0 & \to & (1 & 1 & 1 & 1) & & x+1 & 0 & \to & (x+1 & x+1 & x+1 & x+1) \\
1 & 1 & \to & (1 & 0 & x+1 & x) & & x+1 & 1 & \to & (x+1 & x & 1 & 0) \\
1 & x & \to & (1 & x+1 & x & 0) & & x+1 & x & \to & (x+1 & 1 & 0 & x) \\
1 & x+1 & \to & (1 & x & 0 & x+1) & & x+1 & x+1 & \to & (x+1 & 0 & x & 1)
\end{array}$$

$\square$

**Exercise 3.** Consider the following mapping from $(\mathbb{F}_q)^k$ to $(\mathbb{F}_q)^{k+1}$. Let $(f_0, f_1, \ldots, f_{k-1})$ be any $k$-tuple over $\mathbb{F}_q$, and define the polynomial $f(x) = f_0 + f_1 x + \ldots + f_{k-1}x^{k-1}$ of degree less than $k$. Map $(f_0, f_1, ..., f_{k-1})$ to the $(q + 1)$-tuple $(\{f(\alpha_i), \alpha_i \in \mathbb{F}_q\}, f_{k-1})$—i.e., to the RS codeword corresponding to $f(x)$, plus an additional component equal to $f_{k-1}$.

Show that the $q^k (q+1)$-tuples generated by this mapping as the polynomial $f(z)$ ranges over all $q^k$ polynomials over $\mathbb{F}_q$ of degree $< k$ form a linear $(n = q + 1, k, d = n - k + 1)$ MDS code over $\mathbb{F}_q$. [Hint:$f(x)$ has degree $< k - 1$ if and only if $f_{k-1} = 0$.]

*Solution.* The code has length $n = q + 1$. It is linear because the sum of codewords corresponding to$f(x)$ and $g(x)$ is the codeword corresponding to $f(x) + g(x)$, another polynomial of degree less than $k$. Its dimension is $k$ because no polynomial other than the zero polynomial maps to the zero $(q + 1)$-tuple.

To prove that the minimum weight of any nonzero codeword is $d = n - k + 1$, use the hint and consider the two possible cases for $f_{k-1}$:

- If $f_{k-1} \neq 0$, then $\deg f(x) = k - 1$. By the fundamental theorem of algebra, the RS codeword corresponding to $f(x)$ has at most $k - 1$ zeroes. Moreover, the $f_{k-1}$ component is nonzero. Thus the number of nonzero components in the code $(q+1)$-tuple is at least $q - (k-1) + 1 = n - k + 1$.

- If $f_{k-1} = 0$ and $f(x) = 0$, then $\deg f(x) \leq k-2$. By the fundamental theorem of algebra, the RS codeword corresponding to $f(x)$ has at most $k-2$ zeroes, so the number of nonzero components in the code $(q+1)$-tuple is at least $q-(k-2) = n-k+1$.

$\square$

**Exercise 4.** Suppose we want to correct bursts of errors, that is error patterns that affect a certain number of consecutive bits. Suppose we are given an $[n,k]$ RS code over $\mathbb{F}_{2^t}$. Show that this code yields a binary code which can correct any burst of $(\lfloor(n-k)\rfloor/2-1)t$ bits.

*Solution.* Map each $2^t$ symbols of $\mathbb{F}_{2^t}$ into $t$ bits. The code can correct up to $(d-1)/2$ symbol errors which translates into an error correction capability of $(\lfloor(d-1)/2\rfloor - 1)t$ consecutive bits ($\lfloor(d-1)/2\rfloor t$ if the burst of errors starts at the beginning of a symbol). $\square$

**Exercise 5.** 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 subexponentially 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 $\delta$ 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 $\mathcal{N}^{O(\log \mathcal{N})}$ 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 $\delta$ 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.

$\square$