Class 10 Exercises

CS250/EE387, Winter 2025

Warm-Up

1. What can you say (so far in this class) about the list-decodability of Reed-Solomon codes? That is, what is the best trade-off between R and ρ so that an RS code of rate R is (ρ, L)-list-decodable for, say, polynomial-sized L?

Solution

The best we can do (so far) is plug in the distance of RS codes into the Johnson bound. The distance of RS codes is $\delta = 1 - R$. The JB says that as long as

$$\rho < J_q(\delta) = (1 - 1/q)(1 - \sqrt{1 - q\delta/(q-1)}) \approx 1 - \sqrt{1-\delta} = 1 - \sqrt{R},$$

then an RS code of rate R is $(\rho, q\delta n^2)$-list-decodable.

One proof of the Johnson bound

Today we'll prove the (binary) Johnson bound. (You'll see a different proof on your homework). Recall from the lecture videos/notes that

$$J_2(\delta) = \frac{1}{2}(1 - \sqrt{1 - 2\delta}),$$

and that the Johnson bound says:

Theorem 1 (Johnson bound)**.** *Suppose that $\mathcal{C} \subseteq \{0,1\}^n$ is a code of relative distance at least δ. Suppose that $\rho \leq J_2(\delta)$. Then for any $z \in \{0,1\}^n$,*

$$|\mathcal{C} \cap B_2^n(z, \rho)| \leq \textit{something polynomial in } n$$

Towards proving Theorem 1, let $\mathcal{C} \subseteq \{0,1\}^n$ be a code of relative distance at least δ, and choose $\rho < J_2(\delta)$. let $z \in \{0,1\}^n$ be any vector. Suppose that $\mathcal{C} \cap B_2^n(z, \rho) = \{c_1, \ldots, c_M\}$. Our goal is to show that M is not too big.

2. Define a map $\phi : \{0,1\} \to \mathbb{R}^2$ by:

$$\phi(0) = (0,1) \qquad \phi(1) = (1,0).$$

Extend this to a map $\phi : \{0,1\}^n \to \mathbb{R}^{2n}$ in the natural way. That is,

$$\phi((x_1, \ldots, x_n)) = \phi(x_1) \circ \phi(x_2) \circ \cdots \circ \phi(x_n),$$

where $\circ$ denotes concatenation. Define

$$v := \alpha\phi(z) + \frac{1-\alpha}{2}\mathbf{1},$$

where $\alpha \in [0,1]$ is some parameter that we will define later, and where $\mathbf{1}$ is the all-ones vector of length $2n$.

What can you say about each of the following quantities? (That is, either simplify them or bound them). Your answers should be in terms of δ, ρ, α.

(a) $\langle \phi(c_i), \phi(c_j) \rangle$ for $i \neq j$. (Show that this is at most something, using the fact that the amount of agreement between two codewords is at most $(1-\delta)n$).

(b) $\langle v, \phi(c_i) \rangle$ for any $i = 1, \ldots, M$. (Show that this is at least something, using the fact that the agreement between any of the c_i and z is at least $(1-\rho)n$).

(c) $\langle v, v \rangle$. (Figure out exactly what this is equal to).

Solution

(a) For any $i \neq j$, by the distance of the code and the definition of ϕ, we have

$$\langle \phi(c_i), \phi(c_j) \rangle = \text{agreement}(c_i, c_j) \leq n(1-\delta).$$

(b) For any i, using the fact that $c_i \in B_2(z, \rho)$, we have

$$\begin{aligned}
\langle v, \phi(c_i) \rangle &= \alpha \langle \phi(z), \phi(c_i) \rangle + \frac{1-\alpha}{2} \langle \mathbf{1}, \phi(c_i) \rangle \\
&= \alpha \cdot \text{agreement}(z, c_i) + \frac{1-\alpha}{2} n \\
&\geq \alpha(1-\rho)n + \frac{1-\alpha}{2} n.
\end{aligned}$$

(c) From the definition of v,

$$\begin{aligned}
\langle v, v \rangle &= \alpha^2 \langle \phi(z), \phi(z) \rangle + \alpha(1-\alpha) \langle \phi(z), \mathbf{1} \rangle + \frac{(1-\alpha)^2}{4} \langle \mathbf{1}, \mathbf{1} \rangle \\
&= \left(\alpha^2 + \alpha(1-\alpha) + \frac{(1-\alpha)^2}{2} \right) n \\
&= \left(\alpha^2 + \alpha - \alpha^2 + 1/2 - \alpha + \alpha^2/2 \right) n \\
&= \left(\frac{1+\alpha^2}{2} \right) n
\end{aligned}$$

3. Choose $\alpha = \sqrt{1-2\delta}$. Show that

$$\langle v - \phi(c_i), v - \phi(c_j) \rangle \leq 0$$

for any $i \neq j$? (Hint, use (i) the previous part, (ii) the assumption that $\rho < J_2(\delta) = \frac{1-\alpha}{2}$ using our choice of α, and (iii) the fact that $(1-\alpha^2)/2 = \delta$ using again our choice of α).

Solution

With this choice of α, and the assumption that $\rho < J_2(\delta)$, all of these inner products are negative. To see this, we have

$$\begin{aligned}
\langle \phi(c_i) - v, \phi(c_j) - v \rangle &= \langle v, v \rangle - \langle v, \phi(c_i) \rangle - \langle v, \phi(c_j) \rangle + \langle \phi(c_i), \phi(c_j) \rangle \\
&\leq \left(\frac{1+\alpha^2}{2} \right) n - 2 \left(\alpha(1-\rho)n + \frac{1-\alpha}{2} n \right) + n(1-\delta) \\
&= n \left(1/2 + -\alpha + \alpha^2/2 + 2\alpha\rho - \delta \right).
\end{aligned}$$

Now we can plug in our assumption that $\rho < (1-\alpha)/2$, and get

$$\begin{aligned}\langle v-\phi(c_i), v-\phi(c_j)\rangle &< n\left(1/2+-\alpha+\alpha^2/2+\alpha(1-\alpha)-\delta\right)\\ &= n\left(\frac{1-\alpha^2}{2}-\delta\right)\\ &= n\left(\delta-\delta\right)=0,\end{aligned}$$

using the choice of α in the final line.

4. It turns out that you can't have too many vectors in $\mathbb{R}^D$ that are all at obtuse angles from each other. More precisely, we have the following fact:

   Fact 2. *Let $x_1, x_2, \ldots, x_M \in \mathbb{R}^D$ such that $\langle x_i, x_j\rangle \le 0$ for all $i \neq j$. Suppose further that there exists a non-zero vector $u \in \mathbb{R}^D$ so that $\langle u, v_i\rangle \ge 0$ for all $i = 1, \ldots, M$. Then $M \le 2D-1$.*

   Use the fact to prove Theorem 1.

   Solution

   We have

   $$\begin{aligned}\langle \phi(c_i)-v, v\rangle &= \langle v, \phi(c_i)\rangle - \langle v, v\rangle\\ &\ge \left(\frac{1+\alpha^2}{2}\right)n - \frac{1-\alpha}{2}n + \alpha(1-\rho)n\\ &\ge \left(\frac{1+\alpha^2}{2}\right)n - \frac{1-\alpha}{2}n + \alpha(1-(1-\alpha)/2)n\\ &= n\left(1/2-\alpha/2-1/2-\alpha^2/2+\alpha-\alpha/2+\alpha^2/2\right)\\ &= 0\end{aligned}$$

   where we have used part 1(b) and the fact that $\rho \le (1-\alpha)/2$.

   Thus, we can apply the fact with $x_i \leftarrow \phi(c_i) - v$ and $u \leftarrow v$, and conclude that

   $$M \le 4n-1.$$

5. **(Bonus).** Use this technique to prove the q-ary Johnson bound.

   Solution

   See "Extensions to the Johnson Bound", by Guruswami and Sudan. (Linked on website). The short version is that you take

   $$v \leftarrow \alpha\phi(z) + \frac{1-\alpha}{q}\mathbf{1}$$

   and choose

   $$\alpha \leftarrow \sqrt{1-\frac{q\delta}{q-1}}.$$

   Then do the same thing as above.

The Johnson bound is "tight"

We saw in the mini-lectures that there exist codes list-decodable well beyond the Johnson bound. But are there some codes for which the Johnson bound is essentially the best answer? Yes, there are, and now we'll prove it! (At least for $q = 2$; the proof for general q is similar).

6. Fix $\rho > 0$ and $\varepsilon > 0$.

   (a) Show that there exists a code $\mathcal{C} \subseteq \Sigma^n$ so that:

   - For all $c \in \mathcal{C}$, $\mathrm{wt}(c) \leq \rho n$
   - $\mathcal{C}$ has distance δ at least $\delta \geq 2\rho(1-\rho) - \varepsilon$.
   - $\mathcal{C}$ has rate at least $\frac{1}{100}\varepsilon^2$.

   Hint: You may use the following statement: Suppose that $S \subseteq [n]$ is a random subset of size ρn. Let S' be an independent copy of S. Let $S \triangle S'$ denote the *symmetric difference* between S and S' (that is, $S \triangle S' = (S \cup S') \setminus (S \cap S')$). Then

$$\Pr[|S \triangle S'| \geq \mathbb{E}|S \triangle S'| + \varepsilon n] \leq 2^{-\frac{1}{25}\varepsilon^2 n}.$$

   Another hint: Choose a bunch of random codewords of weight ρn. How many can you choose so that it's still very likely that they have distance at least $\delta \geq 2\rho(1-\rho) - \varepsilon$?

Solution

Following the hint, choose $c^{(1)}, \ldots, c^{(M)} \in \{0,1\}^n$ independently at random so that they have weight exactly ρn. Let $S^{(j)}$ denote the support of $c^{(j)}$, so $|S^{(j)}| = \rho n$.
For any $i \neq j$, we have

$$\mathbb{E}[\Delta(c^{(i)}, c^{(j)})] = \sum_{\ell=1}^{n} \mathbb{E}\mathbf{1}[c_\ell^{(i)} \neq c_\ell^{(j)}] = n \cdot 2\rho(1-\rho),$$

since for each $\ell \in [n]$, the probability that $c_\ell^{(i)} \neq c_\ell^{(j)}$ is $2\rho(1-\rho)$. Note that $\Delta(c^{(i)}, c^{(j)}) = |S^{(i)} \triangle S^{(j)}|$, so by the concentration bound from before,

$$\Pr[\Delta(c^{(i)}, c^{(j)}) \geq 2n\rho(2-\rho) + n\varepsilon] \leq 2^{-\varepsilon^2 n/25}.$$

Thus, as long as $M \leq 2^{\varepsilon^2 n/100}$, by a union bound the probability that there exists any pair $i \neq j$ so that $\Delta(c^{(i)}, c^{(j)}) \geq 2n\rho(1-\rho) + n\varepsilon$ is at most

$$\binom{M}{2} 2^{-\varepsilon^2 n/25} \leq 2^{2\varepsilon^2 n/100} 2^{-\varepsilon^2 n/25} = 2^{-\varepsilon^2 n/25}.$$

In particular, there *exists* a choice of $2^{\varepsilon^2 n/100}$ codewords $c^{(j)}$ so that all of them have distance at least δ from each other. By definition, such a code has rate at least $\varepsilon^2/100$.

   (b) Recall that $J_2(\delta) = \frac{1}{2}(1 - \sqrt{1-2\delta})$. Show that, for any $\delta \in (0, 1/2)$ and any $\varepsilon > 0$, there is a binary code $\mathcal{C} \in \{0,1\}^n$ with distance at least δ so that there is some $z \in \{0,1\}^n$ with

$$|\{c \in \mathcal{C} : \Delta(c, z) \leq J_2(\delta) + C\varepsilon\}| \geq 2^{\varepsilon^2 n/100},$$

   where C is some constant that doesn't depend on n, δ, ε.
   In particular, this shows that the Johnson bound can be tight in some cases.
   Hints: Use the previous part. What happens if you take $\delta = 2\rho(1-\rho) + \varepsilon$ and plug it into $J_2(\delta)$? It might be relevant that $1 - 4\rho(1-\rho) = (1-2\rho)^2$. It might also be relevant that $\sqrt{a - \varepsilon} = \sqrt{a} - O(\varepsilon)$ for small ε.

Solution

Fix $\delta \in (0, 1/2)$. Choose $\rho = J_2(\delta) + C\varepsilon$ for some C that we will choose in a moment. We claim that we can take $\delta = 2\rho(1-\rho) + \varepsilon$ and find some C so that this is consistent. Indeed, with that choice of δ,

$$J_2(\delta) = \frac{1}{2}(1 - \sqrt{1 - 4\rho(1-\rho) - 2\varepsilon}) = \frac{1}{2}(1 - \sqrt{(1-2\rho)^2 - 2\varepsilon}) = \frac{1}{2}(2\rho + O(\varepsilon)) = \rho + O(\varepsilon).$$

So choose C so that this works out. But we just saw in the previous part that we can find a code $\mathcal{C}$ of size at least $2^{\varepsilon^2 n/100}$ that actually *all* lies inside a ball of radius ρ about 0. So we take $z = 0$, and $\mathcal{C}$ to be that code, and we're done.