Math 342. Notes for Lecture 10. February 6, 2018

**Back to the Plotkin bound**

Theorem: (a) If $d$ is even and $\frac{1}{2}n < d \leq n$, then

$$A_2(n, d) \leq 2\left\lfloor \frac{d}{2d - n} \right\rfloor.$$

(b) If $d$ is even, then $A_2(2d, d) \leq 4d$.

If $d$ is odd, we can use this bound in combination with $A_2(n, d) = A_2(n + 1, d + 1)$ (Corollary 2.8). Now $d + 1$ is even, and the Plotkin bound can be applied to $A_2(n + 1, d + 1)$, provided that

$$d + 1 \geq \frac{1}{2}(n + 1).$$

Recall the examples from Lecture 7.

Example 1. $A_2(7, 3) = A_2(8, 4) \leq 16$, which we know is optimal.

Example 2. $A_2(5, 3) = A_2(6, 4) \leq 2\left\lfloor \frac{4}{8 - 6} \right\rfloor = 4$. We have previously seen that $A_2(5, 3) = 4$. (Can you exhibit a binary $(5, 4, 3)$-code?)

Example 3. $A_2(11, 7) \leq 2\left\lfloor \frac{11}{14 - 11} \right\rfloor = 2\left\lfloor \frac{11}{3} \right\rfloor = 6$.

We have not yet proved the Plotkin bound. Here is a proof.

Proof of part (a): Let $C$ be an $(n, M, d)$-code. We want to show that $\frac{1}{2}n < d \leq n$, then

$$M \leq 2\left\lfloor \frac{d}{2d - n} \right\rfloor.$$
Denote the codewords in $C$ by $\bar{a}_1, \ldots, \bar{a}_M$, where

$$\bar{a}_i = (a_{i1}, \ldots, a_{in})$$

for each $i = 1, \ldots, M$. Here each digit $a_{ik}$ is 0 or 1. Consider the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{M1} & a_{M2} & \cdots & a_{Mn} \end{pmatrix},$$

whose rows are the words $\bar{a}_1, \ldots, \bar{a}_M$ of $C$.

The idea behind the proof is to estimate the sum $S$ of the distances $d(\bar{a}_i, \bar{a}_j)$ between all pairs of distinct codewords $\bar{a}_i$ and $\bar{a}_j$ from $C$ in two ways, one using the rows of the matrix $M$, and the other using the columns. Here $1 \leq i < j \leq M$.

On the one hand, there are $\binom{M}{2} = \frac{M(M - 1)}{2}$ such pairs, and each distance $d(\bar{a}_i, \bar{a}_j)$ is $\geq d$. Thus

$$S \geq \frac{M(M - 1)}{2} d. \quad (1)$$

On the other hand,

$$S = \sum_{1 \leq i < j \leq M} d(\bar{a}_i, \bar{a}_j) = \sum_{1 \leq i < j \leq M} \sum_{k=1}^{n} d(a_{ik}, a_{jk}).$$

Here, as usual,

$$d(a_{ik}, a_{jk}) = \begin{cases} 0, & \text{if } a_{ik} = a_{jk}, \\ 1, & \text{if } a_{ik} \neq a_{jk}. \end{cases}$$

Changing the order of summation, we obtain

$$S = \sum_{k=1}^{n} \sum_{1 \leq i < j \leq M} d(a_{ik}, a_{jk}).$$
Let us now fix $k$ and focus on the second sum
\[
\sum_{1 \leq i < j \leq M} d(a_{ik}, a_{jk}).
\]
The numbers $a_{1k}, \ldots, a_{Mk}$ come from the $k$th column of the matrix $A$. If there are $x$ zeros in this column and $M - x$ ones, then this sum is $x(M - x)$. (Why?) Writing
\[
x(M - x) = -(x^2 - Mx) = -(x - \frac{M}{2})^2 + \frac{M^2}{4},
\]
we see that the biggest value this quantity can attain is $\frac{M^2}{4}$ at $x = \frac{M}{2}$. Thus
\[
\sum_{1 \leq i < j \leq M} d(a_{ik}, a_{jk}) \geq \frac{M^2}{4} = \frac{M^2}{4}.
\]
Moreover, note that $x$ is an integer. If $M$ is odd, then $x - \frac{M}{2}$ is at least $\frac{1}{2}$ and thus
\[
x(M - x) = -(x - \frac{M}{2})^2 + \frac{M^2}{4} \leq \frac{M^2 - 1}{4}.
\]
Thus
\[
\sum_{1 \leq i < j \leq M} d(a_{ik}, a_{jk}) \leq \begin{cases} 
\frac{M^2}{4}, & \text{if } M \text{ is even}, \\
\frac{M^2 - 1}{4}, & \text{if } M \text{ is odd}.
\end{cases}
\]
and
\[
\sum_{k=1}^{n} \sum_{1 \leq i < j \leq M} d(a_{ik}, a_{jk}) \leq \begin{cases} 
\frac{nM^2}{4}, & \text{if } M \text{ is even}, \\
\frac{n(M^2 - 1)}{4}, & \text{if } M \text{ is odd}.
\end{cases}
\]
(2)
Let us now consider the cases, where $M$ is even and odd separately. Suppose $M$ is even. Combining (1) and (2), we obtain

$$\frac{M(M - 1)}{2}d \leq S \leq n\frac{M^2}{4}.$$  

Multiplying both sides by 4 and dividing by $M$, we obtain

$$2(M - 1)d \leq nM,$$

or equivalently, $M(2d-n) \leq 2d$. Since we are assuming that $2d > n$, we see that $2d - n$ is a positive integer. Dividing both sides by $2(2d-n)$, we obtain $\frac{M}{2} \leq \frac{d}{2d-n}$. Since $M$ is even, $\frac{M}{2}$ is an integer; thus $\frac{M}{2} \leq \lfloor \frac{d}{2d-n} \rfloor$, or equivalently,

$$M \leq 2\lfloor \frac{d}{2d-n} \rfloor,$$

as desired.

Now suppose $M$ is odd. Then Combining (1) and (2), we obtain

$$\frac{M(M - 1)}{2}d \leq S \leq n\frac{M^2 - 1}{4}.$$  

Multiplying both sides by 4 and dividing by $M-1$, we obtain $2Md \leq n(M + 1)$ or equivalently, $M(2d-n) \leq n$ or $M \leq \frac{n}{2d-n}$.

Adding 1 from both sides and dividing by 2, we obtain

$$\frac{M + 1}{2} \leq \frac{1}{2} \left( \frac{n}{2d-n} + 1 \right) = \frac{1}{2} \cdot \frac{n + 2d-n}{2d-n} = \frac{d}{2d-n}.$$  

Since $M$ is odd, the left hand side is an integer. Thus

$$\frac{M + 1}{2} \leq \lfloor \frac{d}{2d-n} \rfloor.$$
and consequently,
\[ M \leq 2\left\lfloor \frac{d}{2d - n} \right\rfloor - 1, \]
as desired. This completes the proof of part (a).

For the proof of part (b), we need the following

Lemma: \[ A_q(n, d) \geq \frac{1}{q} A_q(n + 1, d). \]

Proof of the lemma: We need to show that if there exists a q-ary \((n + 1, M, d)\) code \(C\), then there exists a q-ary \((n, M', d')\)-code \(C'\), with \(M' \geq \frac{1}{q} M\) and \(d' \geq d\).

To construct \(C'\) from \(C\), consider the digits the last digit of the codewords in \(C\). Denote the symbol that appears in this position most frequently by \(\alpha\). Since there are \(q\) possible symbols, 0, 1, 2, \ldots, \(q - 1\), \(\alpha\) appears in the last position at least \(\frac{M}{q}\) times. Denote this number (i.e., the number of times \(\alpha\) appears in position \(n + 1\) in \(C\), by \(M'\).

We now define a q-ary code \(C'\) as follows. Start with the \(M'\) words from \(C\) with \(\alpha\) in the last position. Remove \(\alpha\) from each of these words. The resulting code \(C'\) is a q-ary code of length \(n\) with \(M'\) codewords. Denote \(d(C')\) by \(d'\). It remains to show that \(d' \geq d\). Indeed, if \(\bar{x}\) and \(\bar{y}\) are distinct words in \(C\) with \(\alpha\) in the last position, and \(\bar{x}'\) and \(\bar{y}'\) are words in \(C'\) obtained from \(\bar{x}\) and \(\bar{y}\) by dropping \(\alpha\), then
\[ d(\bar{x}', \bar{y}') = d(\bar{x}, \bar{y}) \geq d. \]
That is, the distance between any two words in \(C'\) is \(\geq d\), and the lemma follows.
Proof of part (b) of the Plotkin bound: By the lemma,

\[
\frac{1}{2} A_2(2d, d) \leq A_2(2d - 1, d),
\]

and by part (a),

\[
A_2(2d - 1, d) \leq 2\left\lfloor \frac{d}{2d - (2d - 1)} \right\rfloor 2\left\lfloor d \right\rfloor = 2d.
\]

Combining (3) and (4), we obtain

\[
A_2(2d, d) \leq 2A_2(2d - 1, d) \leq 4d,
\]
as desired. This completes the proof of the Plotkin bound.