MATH 525
Section 3.1: Some Bounds on Codes

October 9, 2020

Objectives of Section 3.1

1. Be able to prove the existence or non-existence of a code, generally linear, with prescribed parameters: length, size, and minimum distance. This will be accomplished through the use of one of the following bounds:
   (i) Hamming bound;
   (ii) Singleton bound;
   (iii) Gilbert-Varshamov bound.
2. Study codes which meet an optimality criterion:
   (i) Maximum distance separable (MDS) codes.
   (ii) Perfect codes.

Hamming Bound

- Given $n =$ code length and $d =$ minimum distance, provide an *upper bound* for $|C|$, the code size. Suppose $d = 2t + 1$ or $d = 2t + 2$.
- For each codeword $c \in C$, let

$$S_t(c) = \{w \in K^n \mid d(w, c) \leq t\}.$$

- Claim: For every $c \in C$,

$$|S_t(c)| = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}.$$

- $S_t(c_1) \cap S_t(c_2) = \emptyset$ whenever $c_1, c_2 \in C$ and $c_1 \neq c_2$.
- Therefore,

$$|C| \cdot \left[\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}\right] \leq 2^n.$$

Hamming Bound

- The latter inequality shows that

$$|C| \leq \frac{2^n}{\sum_{i=0}^{t} \binom{n}{i}}.$$

- The above is known as the Hamming Bound or Sphere Packing Bound.

Remark: The Hamming bound shows that if a code of length n and distance d exists, then $|C|$ must satisfy the above upper bound. The bound can never be used to show the existence of a code. The Hamming bound is frequently used to show that codes with certain parameters do not exist.

Example

Let $n = 6$ and $d = 3$. Find an upper bound on $|C|$. We have

$$|C| \leq \frac{2^6}{\binom{6}{0} + \binom{6}{1}} = \frac{64}{7} = 9.14.$$

If C is linear, then $|C| \leq 8$ (why?).

Singleton Bound

- Let C be an (n, k, d) linear code with parity-check matrix H. The rank of H is at most $n - k$ (i.e., the number of columns of H). Therefore, any set of $n - k + 1$ rows of H is necessarily linearly dependent. Together with the theory from Section 2.9, this shows that C contains a nonzero codeword of weight at most $n - k + 1$.
- In conclusion, for an (n, k, d) linear code, $d \leq n - k + 1$.
- The above is known as the Singleton Bound.
- Like the Hamming bound, the Singleton bound is frequently used to show that codes with certain parameters do not exist. The bound can never be used to show the existence of a code.
- An (n, k, d) linear code is said to be maximum distance separable (MDS) if $d = n - k + 1$.

Example

Does a $(10, 6, 6)$ linear code exist?

Singleton Bound

Theorem (Theorem 3.1.8)

Let C be an (n, k, d)-linear code. Then:

(1) *C is MDS if and only if every $n - k$ rows of the parity-check matrix are linearly independent.*

(2) *C is MDS if and only if every k columns of the generator matrix are linearly independent.*

Proof.

(1) follows from Theorem 2.9.1 in the textbook and the Singleton bound. The proof of (2) is shown in the supplement (see Canvas). □

Corollary

The dual of an $(n, k, n - k + 1)$ MDS code is an $(n, n - k, k + 1)$ MDS code.

Gilbert-Varshamov Bound

- Contrary to the other two bounds (Hamming and Singleton), the GV bound is used to show that a code with certain given parameters does exist.
- **Question**: Is there a $(15, 6, 5)$-linear code? The parameters satisfy both the Hamming and the Singleton bounds, so the existence of such a code is not discarded.
- To construct a $(15, 6, 5)$-linear code, we will construct a 15×9 matrix H such that every set of four rows is linearly independent:

$$H = \left[\begin{array}{ccc} & I_9 & \\ \hline -- & \mathbf{r}_{10} & -- \\ -- & \mathbf{r}_{11} & -- \\ & \vdots & \\ -- & \mathbf{r}_{15} & -- \end{array}\right]$$

Gilbert-Varshamov Bound

$$H = \left[\begin{array}{ccc} & I_9 & \\ \hline -- & \mathbf{r}_{10} & -- \\ -- & \mathbf{r}_{11} & -- \\ & \vdots & \\ -- & \mathbf{r}_{15} & -- \end{array}\right]$$

- $\mathbf{r}_{10}$ is taken from K^9, but it cannot be a linear combination of three previous rows, including $\mathbf{0}$. There are

$$1 + \binom{9}{1} + \binom{9}{2} + \binom{9}{3} = 130$$

such linear combinations. We select any vector in K^9 that is not one of these linear combinations and let it be the 10th row of H.

Gilbert-Varshamov Bound

- $\mathbf{r}_{11}$ is taken from K^9; it cannot be a linear combination of three previous rows; also, it cannot be the zero row. There are

$$1 + \binom{10}{1} + \binom{10}{2} + \binom{10}{3} = 176$$

such linear combinations. We select any vector in K^9 that is not one of these linear combinations and let it be the 11th row of H.

- $\cdots$
- Finally, $\mathbf{r}_{15}$ is taken from K^9; it cannot be a linear combination of three previous rows; also, it cannot be the zero row. There are

$$1 + \binom{14}{1} + \binom{14}{2} + \binom{14}{3} = 470$$

such linear combinations. We select any vector in K^9 that is not one of these linear combinations and let it be the 15th row of H.

Gilbert-Varshamov Bound

In conclusion, it is possible to select 15 rows in K^9 such that every set containing four of them is linearly independent. Thus, we have just produced a parity-check matrix for a $(15, 6, d)$-linear code with $d \geq 5$. In other words, a $(15, 6, 5)$-linear code exists. In general, we have

Theorem (Gilbert-Varshamov Bound)

If $2^{n-k} > \binom{n-1}{0} + \binom{n-1}{1} + \cdots + \binom{n-1}{d-3} + \binom{n-1}{d-2}$, *then a binary, linear* (n, k, d)*-code exists.*

Example

Is there a $(9, 2, 5)$ linear code?

Gilbert-Varshamov Bound

Corollary

If $n \neq 1, d \neq 1$, then there exists a linear code C of length n and distance at least d with

$$|C| \geq \frac{2^{n-1}}{\binom{n-1}{0} + \binom{n-1}{1} + \cdots + \binom{n-1}{d-3} + \binom{n-1}{d-2}}.$$

Remarks:

1. The corollary indicates that "good" linear codes exist.
2. The converse of the theorem on the previous slide is false: Geometric Goppa codes exceed the GV-bound. Although

$$1 + \binom{14}{1} + \binom{14}{2} + \binom{14}{3} = 470 > 2^{15-7} = 256,$$

a $(15, 7, 5)$-linear code does exist.

Example

Determine a lower and an upper bound on the size of a linear code with $n = 9$ and $d = 5$. The bounds must be "efficient" in the sense that the lower bound is as large as possible and the upper bound is as low as possible.

* * *

Perfect Codes

Definition

A code C of length n and distance $d = 2t + 1$ is said to be a perfect code if

$$|C| = \frac{2^n}{\binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t}}.$$

Remarks:

1. C is a perfect code if and only if C attains the Hamming bound.
2. Suppose C is a perfect code and $w \in K^n$. Then $w \in S_t(c)$ for some $c \in C$. Recall: $S_t(c) = \{w \in K^n \mid d(w, c) \leq t\}$.

Example

The universe code K^n is a perfect code.

Example

Let n be odd. The repetition code $\{00\ldots0, 11\ldots1\}$ of length n is a perfect code.

Remarks:

1. The above two codes are examples of trivial perfect codes.
2. The Hamming and Golay codes (to be studied in this chapter) are both perfect codes. If a code is perfect, then it must have the same parameters as either the Hamming, Golay, or the trivial codes. This result was a major breakthrough in coding theory proved by van Lint and Tietäväinen in 1971 and 1973, respectively.

Theorem

If C is a perfect code of length n and distance $d = 2t + 1$, then C corrects all error patterns of weight $\leq t$, and no other error patterns.

Proof.

By Theorem 1.12.9, C corrects all error patterns of weight $\leq t$. Now suppose $u \in C$ is transmitted and an error pattern e of weight $> t$ occurs. Then $d(u, u + e) = \mathrm{wt}(e) > t$. However, since C is perfect, $u + e$ belongs to $S_t(v)$ for some $v \in C$, that is, $d(u + e, v) \leq t$. In other words,

$$d(u, u + e) > d(u + e, v),$$

so C does not correct e. □

Example

The linear code with generating matrix

$$G = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

has distance equal to 3. Therefore, it is a 1-error-correcting code. Nevertheless, it corrects the error pattern $e = (1, 0, 0, 0, 0, 0, 1)$ of weight equal to 2.