

# M3/4 P55 ALGEBRAIC COMBINATORICS

ABSTRACT.

## CONTENTS

1. What does the table of contents command do?	1
1.1. Check matrix	1
1.2. Correcting an error	2
1.3. Hamming Codes	2
1.4. Hamming bounds	4
1.5. Perfect Codes	5
1.6. The Golay Code	7

## 1. WHAT DOES THE TABLE OF CONTENTS COMMAND DO?

### 1.1. Check matrix.

**Definition 1.1.** Suppose  $A$  is a  $m \times n$  matrix over  $\mathbb{Z}_2$  and:

$$C = \{x \in \mathbb{Z}_2^n : Ax = 0\} \text{ We call } A \text{ a check matrix of the linear code } C$$

**Proposition 1.2.** Suppose the check matrix  $A$  of a linear code  $C$  satisfies

- (1)  $A$  has no zero column
- (2)  $A$  has no two equal columns

Then  $C$  corrects 1 error.

*Proof.* Suppose false. Then  $d(C) \leq 2$  by proposition 1.2. Hence by proposition 1.5  
 $\exists 0 \neq c \in C$  s.t  $wt(c) = 1$  or  $2$

Suppose  $wt(c) = 1$ . Then  $c = l_i (= 0 \dots 1 \dots)$  and

$$A_C = 0 \implies A l_i = 0 \text{ implies } i \text{th col of } A = 0 \text{ Contradiction}$$

Suppose  $wt(c) = 2$  then  $c = l_i + l_j$  so  $A c = 0 \implies A l_i + A l_j = 0 \implies$   
 $i \text{th col of } A = j \text{th col of } A \text{ contradiction} \quad \square$

Examples

(1)

$$C_3 = \{x \in \mathbb{Z}_2^6 : \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} x = 0\}$$

Corrects 1 error by 1.6

---

Date: DEADLINE AUGUST 26, 2011.

- (2) Suppose we want a code  $C$  which corrects 1 error and has  $3 \times n$  check matrix for some  $n$ . What is max dim of  $C$ ? Answer: By 1.6 need to find largest  $n$  s.t.  $\exists 3 \times n$  check matrix with distinct non zero cols (in  $\mathbb{Z}_2^3$ ). Such a matrix will have as cols all non zero vectors in  $\mathbb{Z}_2^3$  of which there are 7, eg:

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

this is a  $3 \times 7$  so in check matrix of code  $C$  of length 7 dim 4 (by rank nullity) correcting 1 error.

This sends 16 messages abcd using codewords abcdxyz where

$$x = a + b + c, y = a + b + d, z = a + c + d$$

This is called a Hamming code  $\text{Ham}(3)$

**1.2. Correcting an error.** Suppose a codeword  $c$  is sent and 1 error is made, so that received vector is  $c'$  which is not necessarily a code. How do we correct the error?

Well,  $c' = c + l_i$  for some  $i$  So

$$(1.3) \quad Ac' = A(c + l_i)$$

$$(1.4) \quad = Ac + Al_i$$

$$(1.5) \quad = Al_i$$

$$(1.6) \quad = i^{\text{th}} \text{ col of } A$$

E.g. Let  $C = \text{Ham}(3)$ . Suppose received vector is  $c = (1101000)^T$ .

Then

$$Ac' = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = 6^{\text{th}} \text{ column of } A$$

### 1.3. Hamming Codes.

**Definition 1.7.** Let  $k \geq 3$  A Hamming Code  $\text{Ham}(k)$  is a code for which the check matrix has as columns all the distinct non zero vectors in  $\mathbb{Z}_2^k$

**Proposition 1.8.** (1)  $\text{Ham}(k)$  has length  $2^k - 1$ , dim  $2^k - 1 - k$   
 (2)  $\text{Ham}(k)$  corrects 1 error

*Proof.* (1) Since there are  $2^k - 1$  non zero vectors in  $\mathbb{Z}_2^k$  check matrix of  $\text{Ham}(k)$  is  $k \times (2^k - 1)$  and rank  $k$

(2) Follows from 1.6

□

**Definition 1.9.** Let  $C, C' \subseteq \mathbb{Z}_2^n$ . Say  $C$  and  $C'$  are equivalent codes if there is a permutation of the coordinates sending codewords in  $C$  bijectively to codewords in  $C'$ . (This is equivalent to permuting the columns of the checkmatrices)

E.g all Hamming codes  $\text{ham}(k)$  are equivalent.

We want codes that correct more than one error though. Ideally we would like to have a matrix condition that corrects lots of errors - we would like to generalize definition 1.6

**Proposition 1.10.** *Let  $d \geq 2$  and let  $C$  be a code with check matrix  $A$ .*

- (1) *Suppose every set of  $d - 1$  columns of  $A$  is linearly independent. If that is true, then the minimum distance  $d(C) \geq d$*
- (2) *Suppose in addition to (1) that  $\exists$  a set of  $d$  columns of  $A$  that are linearly dependent. then  $d(C) = d$*

*Proof.* (1) Suppose false, and  $d(C) \leq d - 1$ . Then  $\exists 0 \neq c \in C$  with  $wt(c) = r \leq d - 1$ . Write  $c$  as a sum of standard basis vectors:

$$c = e_{i_1} + \dots + e_{i_r}$$

So

$$\begin{aligned} 0 &= Ac = Ae_{i_1} + \dots + Ae_{i_r} \\ &= \text{col}_{i_1} + \dots + \text{col}_{i_r} \end{aligned}$$

This is a contradiction, since by the hypothesis of (1) any set of  $r \leq d - 1$  columns is linearly independent.

- (2) Suppose columns  $i_1 \dots i_d$  are linearly dependent, say

$$\lambda_1(\text{col})_{i_1} + \dots + \lambda_d(\text{col})_{i_d} = 0, \lambda_i \in \mathbb{Z}_2$$

As by (1) any  $d - 1$  columns are linearly independent, all of  $\lambda_i = 1 \forall i$ . Then

$$\begin{aligned} 0 &= \text{col}_{i_1} + \dots + \text{col}_{i_d} \\ &= A(e_{i_1} + \dots + e_{i_d}) \end{aligned}$$

Then  $c = e_{i_1} + \dots + e_{i_d} \in C$  and  $wt(c) = d$

□

E.g

Find a linear code of length 9 dimension 2 which corrects 2 errors. Answer: Check matrix  $A$  should be a  $7 \times 9$  matrix (of rank 7). Also need code  $C = \{x \in \mathbb{Z}_2^9 : Ax = 0\}$  to have  $d(C) \geq 5$  so by 1.8 want every set of 4 columns of  $A$  to be linearly independent.

Take

$$A = \begin{bmatrix} & & 1 & \dots & 0 \\ | & | & & \ddots & \\ & & 0 & \dots & 1 \end{bmatrix}$$

Consisting of an  $7 \times 7$  identity matrix and 2 columns  $c_1, c_2$

Need:

- (1)  $wt(c_1) \geq 4, wt(c_2) \geq 4$  (otherwise  $c_i$  and less than 3 columns of  $I_7$  would be linearly dependent)
- (2)  $wt(c_1 + c_2) \geq 3$  (otherwise  $c_1, c_2$  and  $\leq 2$  columns of  $I_7$  would be linearly dependent)

so take

$$A = \begin{bmatrix} 1 & 0 & \\ 1 & 0 & \\ 1 & 0 & \\ 1 & 1 & I_7 \\ 0 & 1 & \\ 0 & 1 & \\ 0 & 1 & \end{bmatrix}$$

This defines the code

$$\begin{aligned} C &= \{abaaa(a+b)bbb : a, b \in \mathbb{Z}_2\} \\ &= \{0^9, 101111000, 0100001111, 111110111\} \end{aligned}$$

**1.4. Hamming bounds.** Suppose a code  $C$  has length  $n$  and corrects  $e$  errors. How big can  $|C|$  be?

Recall:

$$\begin{aligned} &\text{for } v \in \mathbb{Z}_2^n \\ S_2(v) &= \{x \in \mathbb{Z}_2^n : d(x, v) \leq e\} \end{aligned}$$

**Proposition 1.11** (1.9).  $|S_e(v)| = \text{sum of binomial coefficients}$

*Proof.* Let:

$$\begin{aligned} d_i &= \text{no of: } x \in \mathbb{Z}_2^n \\ &\text{s.t } d(v, x) = i \end{aligned}$$

Then:

$$|S_e(v)| = d_i + d_1 + \dots + d + e$$

The vectors at distance  $i$  from  $v$  are those vector differing from  $v$  in  $i$  coordinates of which there are:  $\binom{n}{i}$  so  $d_i = \binom{n}{i}$   $\square$

**Theorem 1.12** (1.10, Hamming Bound). *Let  $C$  be a code of length  $n$ , correcting  $e$  errors.*

*Then*

$$|C| \leq \frac{2^n}{1 + n + \binom{n}{2} + \dots + \binom{n}{e}}$$

*Proof.* As  $C$  corrects  $e$  errors, the sphere  $S_e(c)$  for  $c \in C$  are all disjoint. Hence:

$$\begin{aligned} \left| \bigcup_{c \in C} S_e(c) \right| &= |C| |S_e(c)| \\ &= |C| (1 + n + \dots + \binom{n}{e}) \end{aligned}$$

Since  $\bigcup_{c \in C} S_e(c) \subseteq \mathbb{Z}_2^n$ , this gives  $|C|(1 + n + \dots + \binom{n}{e}) \leq 2^n$   $\square$

Eg. Let  $C$  be a linear code of length 9 correcting 2 errors. What is the maximum dimension of  $C$ ?

Ans. By hamming bound:

$$|C| \leq \frac{2^9}{1+9+\binom{9}{2}} = 2^9/46 < 2^4 \text{ Hence } \dim(C) \leq 3. \text{ We found such a } C \text{ of dim 2.}$$

is there one of dim 3?

To find one we need a  $6 \times 9$  check matrix with any 4 cols independent.

Taking

$$A = \begin{bmatrix} c_1 & c_2 & c_3 & \\ | & | & | & \\ & & & I_6 \end{bmatrix}$$

need  $c_1, c_2, c_3 \in \mathbb{Z}_2^6$  to satisfy:

- (1)  $wt(c_i) \geq 4 \quad \forall i$
- (2)  $wt(c_i + c_j) \geq 3 \quad \forall i \neq j$
- (3)  $wt(c_1 + c_2 + c_3) \geq 2$

Do  $\exists$  such  $c_1, c_2, c_3 \in \mathbb{Z}_2^6$ ?

Answer: No, see problem sheet 2

### 1.5. Perfect Codes.

**Definition 1.13.** A code  $C \subseteq \mathbb{Z}_2^n$  is *e-perfect* if  $C$  corrects  $e$  errors and

$$|C| = \frac{2^n}{1 + n + \dots + \binom{n}{e}}$$

Equivalently, the union of all the (disjoint) spheres  $S_e(c)$  ( $c \in C$ ) is the whole of  $\mathbb{Z}_2^n$ .

1-perfect codes

**Proposition 1.14** (1.11). Let  $C \subseteq \mathbb{Z}_2^n$ . Then

$$|C| = \frac{2^n}{1+n} \iff n = 2^k - 1, |C| = 2^{2^k - k}$$

for some  $k$

*Proof.*  $\Rightarrow$

If  $|C| = \frac{2^n}{1+n}$  then  $1+n = 2^k$  for some  $k$

$\Leftarrow$  Clear

□

Recall that Hamming code  $\text{Ham}(k)$  has length  $n = 2^k - 1$ , dimension  $n - k$  and corrects 1 error. Hence:

**Proposition 1.15** (1.12).  $\text{Ham}(k)$  is a 1-perfect code.

Are there any *e-perfect* codes for  $e \geq 2$

E.g.

For  $e = 2$ , we need  $1 + n + \binom{n}{2} = 2^k$  for some integer  $k$

This is quite rare, but does happen. (ask the number theory nerds)

Famous theorem (van-Lint, Tietravn, 1964)

**Theorem 1.16.** The only e-perfect codes are:

- (1)  $e = 1$ ,  $\text{Ham}(k)$
- (2)  $n = 2e + 1$   $C = \{0\dots 0, 1\dots 1\}$  of dim 1
- (3)  $e = 3, n = 23, \dim C = 12$ , the Golay code

Miraculous arithmetic:

$$1 + 23 + \binom{23}{2} + \binom{23}{3} = 2^{11}$$

Hamming bound is a result for non existence of codes  $C$  of length  $n$ , correcting  $e$  errors.

This time we will concern ourselves with an existence result

Gilbert-Varshamov bound

**Example 1.17.** Let  $C$  be a linear code of length 15, correcting 2 errors. What is the maximum dimension of  $C$ ?

Ans:

Hamming bound gives

$$|C| \leq \frac{2^{15}}{1 + 15 + \binom{15}{2}} = \frac{2^{15}}{|2|} < 2^9$$

Hence  $\dim C \leq 8$

More on this later.

**Theorem 1.18** (G-V bound). [1.12] Let  $n, k, d$  be positive integers such that

$$1 + n - 1 + \binom{n-1}{2} \dots + \binom{n-1}{d-2} < 2^{n-k}$$

Then there exists a linear code of length  $n$ , dimension  $k$  with  $d(C) \geq d$

Eg. take  $n = 15, d = 5$

$$1 + 14 + \binom{14}{2} + \binom{14}{3} = 1 + 14 + 91 + 364 < 512 = 2^9 = 2^{15-6}$$

So G-V bound tells us that such code  $C$  of dim 6 exists.

There may or may not exist such codes of dim 7 or 8. Sadly neither Hamming bound or G-V bound give us anything about the answer to this.

*Proof.* Assume the G-V bound equation. We want to construct a check matrix  $A$  such that:

- (1)  $A$  is  $(n-k) \times n$  (of rank  $n-k$ )
- (2) any  $d-1$  columns of  $A$  are linearly independent

We construct such a matrix inductively, column by column.

Start by choosing the first  $n-k$  columns:

$$[e_1 \dots e_{n-k}]$$

(inductive step) Suppose we've chosen  $i$  columns  $c_1, \dots, c_i \in \mathbb{Z}_2^{n-k}$  Where  $n-k \leq i \leq n-1$  s.t any  $d-1$  columns from  $c_1 \dots c_i$  are linearly independent. Then:

$$A_i = (c_1, \dots, c_i)$$

is  $(n-k) \times i$  and satisfies (2)

For the inductive step we need to choose a further column  $c_{i+1}$  so that  $A_{i+1} = (c_1, \dots, c_i, c_{i+1})$  still satisfies 2

How many "bad" vectors are there - vectors in  $\mathbb{Z}_2^{n-k}$  which are the sum of  $\leq d-2$  of the vectors from  $c_1, \dots, c_i$

There are at most  $1 + i + \binom{i}{2} + \binom{i}{3} \dots + \binom{i}{d-2}$  such vectors.

But since  $i$  is at most  $n - 1$ , this is less than  $2^n - k$  by the G-V bound. So therefore there is a vector in  $\mathbb{Z}_2^{n-k}$  that is not a sum of  $\leq d - 2$  of the vectors  $c_1, \dots, c_i$ . Hence the matrix

$$A_{i+1} = (c_1, \dots, c_i, c_{i+1})$$

satisfies property (2)

By this inductive step we construct  $A_i$  for  $i = n - k, \dots, n$ . The matrix  $A = A_n$  is the required check matrix.  $\square$

**1.6. The Golay Code.** This is a 3-perfect code of length 23, dimension 12

To construct it we first construct the *extended* Golay code  $G_{24}$  Start with  $H = \text{Ham}(3)$ , check matrix:

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

And its reverse  $K$ , with check matrix

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

Add a parity check bit (= sum of bits) to  $H, K$  to get length 8 codes  $H', K'$

Note.1  $H', K'$  are linear codes of length 8 dim 4. Note.2 All codewords are have weight 0, 8 or 4.

Taking the 14 codewords of weight 4 in  $H'$  you'll see that you can define a collection of blocks, forming a 3-design. ( $v = 8$  points,  $k = 4$  (size of block))

**Proposition 1.19** (1.13).  $H \cap K = \{0^7, 1^7\}$  &  $H' \cap K' = \{0^8, 1^8\}$