II. LINEAR CODES

ABSTRACT

Most of the important codes are special types of the so-called linear codes.

Linear codes have very concise description, nice properties, very easy encoding and quite easy decoding.

2. Linear codes

Very nice and useful properties have **linear codes**. They are special sets of words of length n over alphabet $\{0, \ldots, q-1\}$, where q is a power of prime.

Since now on sets of words F_q^n will be considered as the **vector space** $V(n, q)$ of vectors of length n with elements from the set $\{0, 1, \ldots, q-1\}$ and arithmetical operations will be taken modulo q.

The set $\{0, \ldots, q-1\}$ with operations $+$ and $\cdot$ modulo q is called also the Galois field $GF(q)$.

Definition A subset $C \subseteq V(n, q)$ is a **linear code** iff

(1) $u + v \in C$ for all $u, v \in C$

(2) $au \in C$ for all $u \in C, a \in GF(q)$

Example Codes C_1, C_2, C_3 are linear codes.

Lemma A subset $C \subseteq V(n, q)$ is a linear code iff one of the following conditions is satisfied

(1) C is a subspace of $V(n, q)$

(2) sum of any two codewords from C is in C (for the case $q = 2$)

If C is a k-dimensional subspace of $V(n, q)$, then C is called $[n, k]$-code. It has q^k codewords.

Linear codes are also called "**group codes**".

In the following we shall consider linear codes only for the case q is a power of prime.

Exercise Which of the following codes are linear?

$$C_1 = \{00, 01, 10, 11\}$$

$$C_2 = \{000, 011, 101, 110\}$$

$$C_3 = \{00000, 01101, 10110, 11011\}$$

$$C_5 = \{101, 111, 011\}$$

$$C_6 = \{000, 001, 010, 011\}$$

$$C_7 = \{0000, 1001, 0110, 1110\}$$

How to create a linear code

Notation If S is a set of vectors of a vector space, then let $\langle S\rangle$ be the set of all linear combinations of vectors from S.

Theorem For any subset S of a linear space, $\langle S\rangle$ is a linear space that consists of the following words:

- the zero word,
- all words in S,
- all sums of two or more words in S.

Example

$$S = \{0100, 0011, 1100\}$$

$$\langle S\rangle = \{0000, 0100, 0011, 1100, 0111, 1011, 1000, 1111\}.$$

Basic properties of linear codes

Notation $w(x)$ (weight of x) is the number of non-zero entries of x.

Lemma If $x, y \in V(n, q)$, then $h(x, y) = w(x - y)$.

Proof $x - y$ has non-zero entries in exactly those positions where x and y differ.

Theorem Let C be a linear code and let weight of C, notation $w(C)$ be the smallest of the weights of non-zero codewords of C. Then $h(C) = w(C)$.

Proof There are $x, y \in C$ such that $h(C) = h(x, y)$. Hence $h(C) = w(x - y) \geq w(C)$.

On the other hand for some $x \in C$

$$w(C) = w(x) = h(x, 0) \geq h(C).$$

Consequence

- If C is code with m codewords, then in order to determine $h(C)$ one has to make $\binom{m}{2} = \Theta(m^2)$ comparisons.
- If C is a linear code, then $m - 1$ comparisons is enough.

If C is a linear $[n, k]$-code, then it has a **basis** consisting of k codewords.

Example

Code

$$C_4 = \{0000000, 1111111, 1000101, 1100010, \\ 0110001, 1011000, 0101100, 0010110, \\ 0001011, 0111010, 0011101, 1001110, \\ 0100111, 1010011, 1101001, 1110100\}$$

has the basis

$$\{1111111, 1000101, 1100010, 0110001\}.$$

How many different bases has a linear code?

Theorem A binary linear code of dimension k has

$$\frac{1}{k!} \prod_{i=0}^{k-1} (2^k - 2^i)$$

bases.

Advantages and disadvantages of linear codes

Advantages- big.

1. $h(C)$ is easy to compute if C is a linear code.
2. Linear codes have simple specifications.

- To specify a non-linear code usually all codewords have to be listed.
- To specify a linear $[n, k]$-code it is enough to list k codewords.

Definition A $k \times n$ matrix whose rows form a basis of a linear $[n, k]$-code (subspace) C is said to be the **generator matrix** of C.

Example The generator matrix of the code

$$C_2 = \left\{ \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{matrix} \right\} \quad \text{is} \quad \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

and generator matrix of the code

$$C_4 \quad \text{is} \quad \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

3. There are simple encoding/decoding procedures for linear codes.

Disadvantages– small

1. Linear q-codes are not defined unless q is a prime power.
2. The restriction to linear codes might be a restriction to weaker codes than desired.

Equivalence of linear codes

Definition Two linear codes $GF(q)$ are called equivalent if one can be obtained from another by the following operations:
(a) permutation of the positions of the code;
(b) multiplication of symbols appearing in a fixed position by a non-zero scalar.

Theorem Two $k \times n$ matrices generate equivalent linear $[n, k]$-codes over $GF(q)$ if one matrix can be obtained from the other by a sequence of the following operations:
(a) permutation of the rows
(b) multiplication of a row by a non-zero scalar
(c) addition of one row to another
(d) permutation of columns
(e) multiplication of a column by a non-zero scalar

Proof Operations (a) – (c) just replace one basis by another. Last two operations convert a generator matrix to one of an equivalent code.

Theorem Let G be a generator matrix of an $[n, k]$-code. Then by operations (a) – (e) the matrix G can be transformed into the form: $[I_k|A]$ where I_k is the $k \times k$ identity matrix, and A is a $k \times (n-k)$ matrix.

Example

$$\begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 \end{matrix} \rightarrow \begin{matrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{matrix} \rightarrow$$

$$\begin{matrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{matrix} \rightarrow \begin{matrix} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \end{matrix} \rightarrow ?$$

Encoding with a linear code

is a vector $\times$ matrix multiplication

Let C be an $[n, k]$-code over $GF(q)$ with generator matrix G.

Theorem C has q^k codewords.

Proof Theorem follows from the fact that each codeword of C can be expressed uniquelly as a linear combination of the basis vectors.

Corollary The code C can be used to encode uniquely q^k messages. Let us identify messages with elements $V(k, q)$.

Encoding of a message $u = (u_1, \dots, u_k)$ with the code C:

$$u \to u \cdot G = \sum_{i=1}^{k} u_i r_i \quad \text{where } r_1, \dots, r_k \text{ are rows of } G.$$

Example Let C be the $[7, 4]$-code with the generator matrix

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

A message (u_1, u_2, u_3, u_4) is encoded as: ?
For example:

0 0 0 0 is encoded as ....?
1 0 0 0 is encoded as ....?
1 1 1 0 is encoded as ....?

Uniqueness of encodings

with linear codes

Theorem If G is a generator matrix for a binary linear code C of length n and dimension k, then

$$v = uG$$

ranges over all 2^k codewords of C as u ranges over all 2^k words of length k.

Therefore

$$C = \{uG \,|\, u \in \{0,1\}^k\}$$

Moreover

$$u_1 G = u_2 G$$

if and only if

$$u_1 = u_2.$$

Proof. If

$$0 = \sum_{i=1}^{n} u_{1,i} w_i - \sum_{i=1}^{n} u_{2,i} w_i = \sum_{i=1}^{n} (u_{1,i} - u_{2,i}) w_i$$

then, since w_i are linearly independent, $u_1 = u_2$.

Decoding of linear codes

Decoding problem If a codeword: $x_1 \dots x_n$ is sent and the word $y = y_1 \dots y_n$ is received, then $e = y - x = e_1 \dots e_n$ is said to be the **error vector.** The decoder must decide from y which x was sent, or, equivalently, which error e has occured.

Decoding method– technicalities

Definition Suppose C is an $[n, k]$-code over $GF(q)$ and $a \in V(n, q)$. Then the set

$$a + C = \{a + x | x \in C\}$$

is called a **coset** of C.

Example Let $C = \{0000, 1011, 0101, 1110\}$ $\quad G = \begin{bmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}$

Cosets:

$$\begin{aligned} 0000 + C &= C, \\ 1000 + C &= \{1000, 0011, 1101, 0110\}, \\ 0100 + C &= \{0100, 1111, 0001, 1010\}, \\ 0010 + C &= \{0010, 1001, 0111, 1100\}. \end{aligned}$$

Are there some other cosets in this case?

Theorem Suppose C is a linear $[n, k]$-code over $GF(q)$. Then

(a) every vector of $V(n, k)$ is in some coset of C,

(b) every coset contains exactly q^k elements,

(c) two cosets are either disjoint or identical.

Each vector having minimum weight in a coset is called a **coset leader**.

Nearest neighbour decoding scheme:

1. Design a (Slepian) standard array for an $[n,k]$-code C –that is a $q^{n-k} \times q^k$ array of the form:

codewords 1	coset lead. 1	codeword 2	...	codeword 2^k
2	coset lead. 2	$+$	...	$+$
	$\vdots$	$+$	$+$	$+$
q^{n-k}	coset leader	$+$	...	$+$
	coset leaders			

Example

0000	1011	0101	1110
1000	0011	1101	0110
0100	1111	0001	1010
0010	1001	0111	1100

A word y is decoded as the codeword of the column in which y occurs.

The error vectors which will be corrected are precisely coset leaders!

In practice, this decoding method is too slow and requires too much memory.

Probability of error correction

What is the probability that a received vector will be decoded as the codeword sent (for binary linear codes)?

We know that the probability that the error vector is a given vector of weight i, in the case of binary symetric channel with probability p, is

$$p^i(1-p)^{n-i}.$$

Therefore, it holds.

Theorem Let C be a binary $[n, k]$-code, and for $i = 0, 1, \dots, n$ let α_i be the number of coset leaders of weight i. Then the probability $P_{corr}(C)$ that a received vector decoded by means of a standard array is the codeword which was sent is given by

$$P_{corr}(C) = \sum_{i=0}^{n} \alpha_i p^i (1-p)^{n-i}.$$

Example For the $[4, 2]$-code of the last example

$$\alpha_0 = 1, \alpha_1 = 3, \alpha_2 = \alpha_3 = \alpha_4 = 0.$$

Hence

$$P_{corr}(C) = (1-p)^4 + 3p(1-p)^3 = (1-p)^3(1+2p).$$

If $p = 0.01$, then $P_{corr} = 0.9897$

Probability of error detection

Suppose a binary linear code is used only for error detection.

The decoder will fail to detect errors which have occured iff the received word y is a codeword different from the codeword x which was sent, i. e. if the error vector $e = y - x$ is itself a non-zero codeword.

The probability $P_{undetect}(C)$ that an incorrect codeword is received is given by the following result.

Theorem Let C be a binary $[n, k]$-code and let A_i denote the number of codewords of C of weight i. Then, if C is used for error detection, the probability of an incorrect message being received is

$$P_{undetect}(C) = \sum_{i=1}^{n} A_i p^i (1-p)^{n-i}.$$

Example In the case of $[4, 2]$ code from the last example

$$A_2 = 1 \; A_2 = 2$$
$$P_{undetect}(C) = p^2(1-p)^2 + 2p^3(1-p) = p^2 - p^3.$$

For $p = 0.01$

$$P_{undetect}(C) = 0.000099.$$

Dual codes

Inner product of two vectors (words)

$$u = u_1 \dots u_n, \quad v = v_1 \dots v_n$$

in $V(n,q)$ is an element of $GF(q)$ defined by

$$u \cdot v = u_1 v_1 + \ldots + u_n v_n.$$

Example

In $V(4,2)$: $1001 \cdot 1001 = 0$
In $V(4,3)$: $2001 \cdot 1210 = 2$
$1212 \cdot 2121 = 2$

If $u \cdot v = 0$ then words (vectors) u and v are **orthogonal**

Properties If $u, v, w \in V(n,q)$, $\lambda, \mu \in GF(q)$, then

$$u \cdot v = v \cdot u, \;\; (\lambda u + \mu v) \cdot w = \lambda(u \cdot w) + \mu(v \cdot w).$$

Given a linear $[n,k]$-code C, then **dual code** of C, denoted by $C^{\perp}$, is defined by

$$C^{\perp} = \{v \in V(n,q) \,|\, v \cdot u = 0 \text{ if } u \in C\}.$$

Lemma Suppose C is an $[n,k]$-code having a generator matrix G. Then for $v \in V(n,q)$

$$v \in C^{\perp} \Leftrightarrow vG^{\top} = 0,$$

where $G^{\top}$ denotes the transpose of G.

Proof Easy.

PARITE CHECKS versus ORTHOGONALITY

For understanding of the role parity checks play for linear codes it is important to understand relation between orthogonality and parity checks.

If words x and y are orthogonal, then the word y has even number of ones in positions determined by ones in x.

This implies that x is a parity check word for y and y is a parity check word for x.

Example Let the word

$$100001$$

is orthogonal to a set S of binary words of length 6. What can we say about words in S?

EXAMPLE

For the $[n, 1]$-repetition code C, with the generator matrix

$$G = (1, 1, \dots, 1)$$

the dual code

$$C^{\perp}$$

is $[n, n-1]$-code with the generator matrix $G^{\perp}$, descrribed by

$$G^{\perp} = \begin{pmatrix} 1 & 1 & 0 & 0 & \dots & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Parity check matrices

Example If

$$C_5 = \left\{ \begin{array}{cccc} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array} \right\}, \text{ then } C_5^{\perp} = C_5.$$

If

$$C_6 = \left\{ \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{array} \right\}, \text{ then } C_6^{\perp} = \left\{ \begin{array}{ccc} 0 & 0 & 0 \\ 1 & 1 & 1 \end{array} \right\}.$$

Theorem Suppose C is a linear $[n, k]$-code over $GF(q)$, then the dual code $C^{\perp}$ is a linear $[n, n-k]$-code.

Definition A **parity-check matrix** H for an $[n, k]$-code C is a generator matrix of $C^{\perp}$.

Definition A **parity-check matrix** H for an $[n, k]$-code C is a generator matrix of $C^{\perp}$.

Theorem If H is parity-check matrix of C, then

$$C = \{x \in V(n, q) \mid xH^{\top} = 0\},$$

and therefore any linear code is completely specified by a parity-check matrix.

Example Parity-check matrix for

$$C_5 \text{ is } \left\{ \begin{matrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{matrix} \right\}$$

and for

$$C_6 \text{ is } \{\, 1 \;\; 1 \;\; 1 \,\}.$$

The rows of a parity check matrix are **parity checks** on codewords. They say that certain linear combinations of the coordinates of every codeword are zero.

Syndrome decoding

Theorem If $G = [I_k|A]$ is the standard form generator matrix of an $[n,k]$-code C, then a parity check matrix for C is $H = [-A^\top|I_{n-k}]$.

Example Generator matrix $G = \left|\begin{array}{c|ccc} & 1 & 0 & 1 \\ & 1 & 1 & 1 \\ I_4 & 1 & 1 & 0 \\ & 0 & 1 & 1 \end{array}\right|$

$$\Rightarrow \text{parity check matrix } H = \left|\begin{array}{cccc|c} 1 & 1 & 1 & 0 & \\ 0 & 1 & 1 & 1 & I_3 \\ 1 & 1 & 0 & 1 & \end{array}\right|$$

Definition Suppose H is a parity-check matrix of an $[n,k]$-code C. Then for any $y \in V(n,q)$ the following word is called the **syndrome** of y:

$$S(y) = yH^\top.$$

Lemma Two words have the same syndrom iff they are in the same coset.

Syndrom decoding Assume that a standard array is given and, in addition, in the last two columns the syndrom for each coset is given.

0 0 0 0	1 0 1 1	0 1 0 1	1 1 1 0	0	0
1 0 0 0	0 0 1 1	1 1 0 1	0 1 1 0	1	1
0 1 0 0	1 1 1 1	0 0 0 1	1 0 1 0	0	1
0 0 1 0	1 0 0 1	0 1 1 1	1 1 0 0	1	0

When a word y is received, compute $S(y) = yH^\top$, locate $S(y)$ in the "syndrom column", and then locate y in the same row and decode y as the codeword in the same column and in the first row.

KEY OBSERVATION for SYNDROM COMPUTATION

When designing a "syndrome decoding" we actually need to store only two columns: that for coset leaders and for syndromes.

Example

coset leaders $f(z)$	syndromes z
0000	00
1000	11
0100	01
0010	10

Decoding procedure

- **Step 1** Given y compute $S(y)$.
- **Step 2** Locate $z = S(y)$ in the syndrome column.
- **Step 3** Decode y as $y - f(z)$.

Example If $y = 1111$, then $S(y) = 01$ and the above decoding procedure produces

$$1111 - 0100 = 1011$$

.

Syndrom decoding is much fatser than searching for a nearest codeword to a received word. However, for large codes it is still too inefficient to be practical.

In general, the problem of finding the nearest neighbour in a linear code is $\mathbf{NP}$-complete. Fortunately, there are linear codes with really efficient decoding.

Hamming codes

An important family of simple error-correcting linear codes which are easy to encode and decode are the so-called Hamming codes.

Definition Let r be an integer and H be an $r \times (2^r - 1)$ matrix columns of which are non-zero distinct words from $V(r,2)$. The code having H as its parity-check matrix is called **binary Hamming code** and denoted by $Ham(r,2)$.

Example

$$\text{Ham}(2,2) = H = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \Rightarrow G = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$$

$$\text{Ham}(3,2) = H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \Rightarrow G = \begin{vmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{vmatrix}$$

Theorem Hamming code $Ham(r,2)$

- is $[2^r - 1, 2^r - 1 - r]$-code,
- has minimum distance 3,
- is a perfect code.

Properties Coset leaders are precisely words of weight ≤ 1. The syndrome of the word $0 \ldots 010 \ldots 0$ with 1 in j-th position and 0 otherwise is the transpose of the j-th column of H.

Decoding algorithm for the case the columns of H are arranged in the order of increasing binary numbers the columns represent.

- **Step 1** Given y compute syndrome $S(y) = yH^\top$.
- **Step 2** If $S(y) = 0$, then y is assumed to be the codeword sent.
- **Step 3** If $S(y) \neq 0$, then assuming a single error, $S(y)$ gives the binary position of the error.

Example

For the Hamming code given by the parity-check matrix

$$H = \begin{vmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{vmatrix}$$

and the received word

$$y = 110\ 1011,$$

we get syndrome

$$S(y) = 110$$

and therefore the error is in the sixth position.

Hamming code was discovered by Hamming (1950), Golay (1950).

It was conjectured for some time that Hamming codes and two so called Golay codes were the only non-trivial perfect codes.

Comment

Hamming codes were originally used to deal with errors in long-distance telephon calls.

ADVANTAGES of HAMMING CODE

Let a binary symmetric channel is used which with probability q correctly transfers a binary symbol.

If a 4-bit message is transmitted through such a channel, then correct transmission of the message occurs with probability q^4.

If Hamming $(7,4,3)$ code is used to transmitt a 4-bit message, then probbaility of correct decoding is

$$q^7 + 7(1-q)q^6.$$

In case $q = 0.9$ the probbaility of correct transmissio is 0.651 in the first case and 0.8503 in the case Hamming code is used – an essential improvement.

IMPORTANT CODES

- Hamming $(7, 4, 3)$-code. It has 16 codewords of length 7, it can be used to send $2^7 = 128$ messages and can be used to correct 1 error.
- Golay $(23, 12, 7)$ code. It has 4 096 codewords, it can be used to transmitt 8 3888 608 messages and can correct 3 errors.
- Quadratic residue $(47, 24, 11)$-code. It has

  $$16777216 \text{ codewords}$$

  and can be used to transmit

  $$140737488355238 \text{ messages.}$$

  and correct 5 errors.

GOLAY CODES – DESCRIPTION

Golay codes $\mathcal{G}_{24}$ and $\mathcal{G}_{23}$ were used by *Voyager I* and *Voyager II* to transmit color pictures of Jupiter and Saturn. Generating matrix for $\mathcal{G}_{24}$ has the form

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

$\mathcal{G}_{24}$ is $(24, 12, 8)$-code and the weights of all codewords is multiple of 4. $\mathcal{G}_{23}$ is obtained from $\mathcal{G}_{24}$ by deleting last last symbols of each codeword of $\mathcal{G}_{24}$. $\mathcal{G}_{23}$ is $(23, 12, 7)$-code.

GOLAY CODE – CONSTRUCTION

Matrix G for Golay code $\mathcal{G}_{24}$ has actually a simple and regular construction.

The first 12 columns are formed by a unitary matrix I_{12}, next column has all 1's.

Rows of the last 11 columns are cyclic permutations of the first row which has 1 at those positions that are squares modulo 11, that is

$$0, 1, 3, 4, 5, 9.$$

SINGLETON BOUND

If C is a linear $(n.k,d)$-code, then $n-k \geq d-1$ (Singleton bound).

To show the above bound we can use the following lemma.

Lemma If u is a codeword of a linear code C of weight s, then there is a dependence relation among s columns of any parity check matrix of C, and conversely, any dependence relation among s columns of a parity check matrix of C yields a codeword of weight s in C.

Proof Let H be a parity check matrix of C. Since u is orthogonal to each row of H, the s components in u that are nonzero are the coefficients of the dependence relation of the s columns of H corresponding to the s nonzero components. The converse holds by the same reasoning.

Corolllary If C is a linear code, then C has minimum weight d iff d is the largest number so that every $d-1$ columns of any parity check matrix of C are independent.

Corollary For a linear (n,k,d) it holds $n-k \geq d-1$.

A linear (n,k,d)-code is called **maximum distance separable** (MDS code) if $d = n-k+1$.

MDS codes are codes with maximal possible minimum weight.

REED-MULLER CODES

Reed-Muller codes form a family of codes defined recursively with interesting properties and easy decoding.

If D_1 is a binary $[n, k_1, d_1]$-code and D_2 is a binary $[n, k_2, d_2]$-code, a binary code C of length $2n$ is defined as follows $C = \{|u|u+v|, \text{where } u \in D_1, v \in D_2\}$.

Lemma C is $[2n, k_1 + k_2, \min\{2d_1, d_2\}]$-code and if G_i is a generator matrix for D_i, $i = 1, 2$, then $\begin{pmatrix} G_1 & G_2 \\ 0 & G_2 \end{pmatrix}$ is a generator matrix for C.

Reed-Muller codes $R(r, m)$, with $0 \leq r \leq m$ are binary codes of length $n = 2^m$. $R(m, m)$ is the whole set of words of length n, $R(0, m)$ is the repetition code.

If $0 < r < m$, then $R(r+1, m+1)$ is obtained from codes $R(r+1, m)$ and $R(r, m)$ by the above construction.

Theorem The dimension of $R(r, m)$ equals $1 + \binom{m}{1} + \ldots + \binom{m}{r}$. The minimum weight of $R(r, m)$ equals 2^{m-r}. Codes $R(m-r-1, m)$ and $R(r, m)$ are dual codes.