Linear Block Codes

Qi Zhang

Aarhus University School of Engineering

13/02/2014



Qi Zhang (ASE) Linear Block Codes 13/02/2014 1 / 29

1 Matrix Form

2 Linear Block Codes

3 Syndrome Error Detection



Matrix Form

- We know that linear combination of linearly independent vectors can generates space or subspace.
- If there are k linearly independent vectors of vector space V_n defined over GF(2), these k vectors can be written into a matrix form with size of $k \times n$ below.

$$G = \begin{bmatrix} g_{00} & g_{01} & \cdots & g_{0,n-1} \\ g_{10} & g_{11} & \cdots & g_{1,n-1} \\ \vdots & \vdots & & \vdots \\ g_{k-1,0} & g_{k-1,1} & \cdots & g_{k-1,n-1} \end{bmatrix}$$

- These k linearly independent vectors can generate 2^k possible linear combinations, i.e., becomes a k-dimension vector subspace.
- This subspace is also called the *row space* of **G**.

Example 2.6: In the following matrix G, the third row is replaced by addition of the second and third rows, and the first and second rows are permuted, generating matrix G':

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

verify that both matrices generate the same three-dimension subspace.



Qi Zhang (ASE) Linear Block Codes 13/02/2014 4 / 29

Solution:

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

The three-dimension subspace of vector space V_5 based on **G**:

$$0 \bullet (10110) \oplus 0 \bullet (01001) \oplus 0 \bullet (11011) = (00000)$$
 $0 \bullet (10110) \oplus 0 \bullet (01001) \oplus 1 \bullet (11011) = (11011)$
 $0 \bullet (10110) \oplus 1 \bullet (01001) \oplus 0 \bullet (11011) = (01001)$
 $0 \bullet (10110) \oplus 1 \bullet (01001) \oplus 1 \bullet (11011) = (10010)$
 $1 \bullet (10110) \oplus 0 \bullet (01001) \oplus 0 \bullet (11011) = (10110)$
 $1 \bullet (10110) \oplus 0 \bullet (01001) \oplus 1 \bullet (11011) = (01101)$
 $1 \bullet (10110) \oplus 1 \bullet (01001) \oplus 0 \bullet (11011) = (11111)$
 $1 \bullet (10110) \oplus 1 \bullet (01001) \oplus 1 \bullet (11011) = (00100)$

5 / 29

Dual subspace matrix

■ For the vector space S, which is the row space of $k \times n$ matrix **G**.

$$\mathbf{G} = egin{bmatrix} \mathbf{g}_0 \ \mathbf{g}_1 \ dots \ \mathbf{g}_{k-1} \end{bmatrix}$$

■ If S_d is the dual space of S, the dimension of S_d is n-k. S_d is the row space of matrix \mathbf{H} which is composed by n-k linearly independent vectors $\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_{n-k-1}$.

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_0 \\ \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_{n-k-1} \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & \dots & h_{0,n-1} \\ h_{10} & h_{11} & \dots & h_{1,n-1} \\ \vdots & \vdots & & \vdots \\ h_{n-k-1,0} & h_{n-k-1,1} & \dots & h_{n-k-1,n-1} \end{bmatrix}$$
AARHUS UNIVERSITET OF

■ Property: $\mathbf{g}_i \circ \mathbf{h}_i = 0$

Example 2.7: The vector subspace S is generated by matrix G. Verify that the generated vector subspace S_d by matrix H is the dual vector space of S.

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

Solution:

```
S = \{(00000), (11011), (10110), (01001), (10010), (11111), (01101), (00100)\}
S_d = \{(00000), (01001), (10010), (11011)\}
\mathbf{g}_i \circ \mathbf{h}_j = 0
...
(10110) \circ (01001) = 0 \quad (10110) \circ (10010) = 0
(01001) \circ (01001) = 0 \quad (01001) \circ (10010) = 0
...
(01101) \circ (01001) = 0 \quad (01101) \circ (10010) = 0
```



7 / 29

Introduction of linear block codes

- Message information is grouped into a *k*-bits block;
- There are 2^k possible messages;



- The generic denotation of a message: $\mathbf{m} = (m_0, m_1, \dots, m_{k-1});$
- The encoder encodes each k-bits source message into a n-bits codeword (or code vector): $\mathbf{c} = (c_0, c_1, \dots, c_{n-1})$;
- The encoding procedure is a bijective assignment between 2^k vectors of the message vector space and 2^k out of the 2^n possible vectors of the encoded vector space.
- The k bits are information bits, the n-k bits are redundancy. The coding rate R=k/n.
- **Definition 2.1**: A block code of length n and 2^k codewords are said to be a linear block code $C_b(n,k)$, if the 2^k codewords form a vector subspace, of dimension k, of the vector space V_n of all the vectors of length n with components in the field GF(2).
- **Property**: The sum of any two codewords is also a codeword.

Generator matrix **G**

- A linear block code $C_b(n, k)$ is a k-dimension vector subspace of the vector space V_n ;
- This k-dimension vector subspace is generated by k linearly independent *n*-components vectors, $\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_{k-1}$;
- **Each** possible codeword \mathbf{c} is a linear combination of the k vectors $\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_{k-1}$:

$$\mathbf{c} = m_0 \bullet \mathbf{g}_0 \oplus m_1 \bullet \mathbf{g}_1 \oplus \ldots \oplus m_{k-1} \bullet \mathbf{g}_{k-1}$$

■ Write the k linearly independent vectors $\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_{k-1}$ into matrix form:

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix} = \begin{bmatrix} g_{00} & g_{01} & \cdots & g_{0,n-1} \\ g_{10} & g_{11} & \cdots & g_{1,n-1} \\ \vdots & \vdots & & \vdots \\ g_{k-1,0} & g_{k-1,1} & \cdots & g_{k-1,n-1} \end{bmatrix}$$
AARHUS UNCHARRADIS INCOMPRISA



Generator matrix **G**

- The matrix G is called the generator matrix.
- The matrix mechanism for generating any code word of a message vector $\mathbf{m} = (m_0, m_1, \dots, m_{k-1})$:

$$\mathbf{c} = \mathbf{m} \circ \mathbf{G} = (m_0, m_1, \dots, m_{k-1}) \circ \begin{bmatrix} g_{00} & g_{01} & \dots & g_{0,n-1} \\ g_{10} & g_{11} & \dots & g_{1,n-1} \\ \vdots & \vdots & & \vdots \\ g_{k-1,0} & g_{k-1,1} & \dots & g_{k-1,n-1} \end{bmatrix}$$

$$= (m_0, m_1, \dots, m_{k-1}) \circ \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix}$$

$$= m_0 \bullet \mathbf{g}_0 \oplus m_1 \bullet \mathbf{g}_1 \oplus \dots \oplus m_{k-1} \bullet \mathbf{g}_{k-1}$$

■ The k linearly independent rows of the generator matrix **G** generate the linear block code $C_b(n, k)$.



Generator matrix **G**

■ Example 2.8: Consider the following generator matrix of size 4 × 7 and obtain the codeword corresponding the message vector $\mathbf{m} = (1001)$:

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \mathbf{g}_2 \\ \mathbf{g}_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution:

$$\mathbf{c} = \mathbf{m} \circ \mathbf{G} = 1 \bullet \mathbf{g}_0 \oplus 0 \bullet \mathbf{g}_1 \oplus 0 \bullet \mathbf{g}_2 \oplus 1 \bullet \mathbf{g}_3$$
$$= (1101000) \oplus (1010001) = (0111001)$$

Question: if the message vector $\mathbf{m} = (1110)$, what is the corresponding codeword?



11 / 29

Qi Zhang (ASE) Linear Block Codes 13/02/2014

Codewords of a linear block code $C_b(7,4)$

Messages	Messages Codewords					
0000	$0\ 0\ 0\ 0\ 0\ 0$					
0001	$1\ 0\ 1\ 0\ 0\ 0\ 1$					
0010	1110010					
0011	$0\ 1\ 0\ 0\ 0\ 1\ 1$					
0100	0110100					
0101	1100101					
0 1 1 0	1000110					
0 1 1 1	0010111					
1000	1101000					
1001	$0\ 1\ 1\ 1\ 0\ 0\ 1$					
1010	0011010					
1011	1001011					
1100	1011100					
1101	$0\ 0\ 0\ 1\ 1\ 0\ 1$					
1110	0101110					
$1\ 1\ 1\ 1$	1111111					



Codewords of a linear block code $C_b(7,4)$

```
Messages
          Codewords
0000
         000000
0001
         101 || 0001
0010
         111 || 0010
         010 || 0011
0011
0100
         011 || 0100
0101
         110 || 0101
0110
         100 \parallel 0110
0111
         001 || 0111
1000
         110 || 1000
1001
         011 || 1001
1010
         001 || 1010
1011
         100 || 1011
1 1 0 0
         101 || 1100
1 1 0 1
         000||1101
1 1 1 0
         0 1 0 || 1 1 1 0
1111
         111 \parallel 1111
```



- In the previous linear block code, the last four bits of each codeword are the same as the message bits;
- Namely, the message appears inside the codeword;
- The first three bits are the so-called parity check or redundancy bits.
- This particular form of the codeword is called **systematic form**.

$$n-k$$
 parity check bits k message bits

Note: the parity check bits can also be placed at the end of the codeword.



Qi Zhang (ASE) Linear Block Codes 13/02/2014 14 / 29

■ A systematic linear block code $C_b(n, k)$ is uniquely specified by a generator matrix of the form:

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_0 \\ \mathbf{g}_1 \\ \vdots \\ \mathbf{g}_{k-1} \end{bmatrix} = \begin{bmatrix} p_{00} & p_{01} & \dots & p_{0,n-k-1} & 1 & 0 & 0 & \dots & 0 \\ p_{10} & p_{11} & \dots & p_{1,n-k-1} & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{k-1,0} & p_{k-1,1} & \dots & p_{k-1,n-k-1} & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

$$= \begin{bmatrix} P & P & I_k \end{bmatrix}$$

- Submatrix P is of size $k \times (n-k)$;
- Submatrix I_k is of size $k \times k$.
- Generator matrix **G** is of size $k \times n$.



Qi Zhang (ASE) Linear Block Codes

Parity check equations:

$$c_j = m_0 \bullet p_{0,j} + m_1 \bullet p_{1,j} + \ldots + m_{k-1} \bullet p_{k-1,j} \quad 0 \le j < n-k$$

 $c_j = c_{n-k+j} = m_j \quad 0 \le i \le k-1, n-k \le j < n$



Qi Zhang (ASE) Linear Block Codes 13/02/2014 16 / 29

Example 2.9: List the parity check equations for the linear block code $C_b(7,4)$ blow:

$$\mathbf{c} = \mathbf{m} \circ \mathbf{G} = (m_0, m_1, m_2, m_3) \circ \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Solution:

$$c_0 = m_0 \oplus m_2 \oplus m_3$$
 $c_1 = m_0 \oplus m_1 \oplus m_2$
 $c_2 = m_1 \oplus m_2 \oplus m_3$
 $c_3 = m_0$
 $c_4 = m_1$
 $c_5 = m_2$
 $c_6 = m_3$



Parity check matrix **H**

■ We know $\mathbf{G} \Rightarrow \mathbf{S} \Rightarrow \mathbf{S}_d \Leftarrow \mathbf{H}$.

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & \dots & 0 & p_{00} & p_{1,0} & \dots & p_{k-1,0} \\ 0 & 1 & \dots & 0 & p_{01} & p_{1,1} & \dots & p_{k-1,1} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & p_{0,n-k-1} & p_{1,n-k-1} & \dots & p_{k-1,n-k-1} \end{bmatrix}$$

$$= \begin{bmatrix} I_{n-k} & P^T \end{bmatrix}$$

■ Property:

$$\mathbf{g}_{i} = (p_{i0}, \dots, p_{ij}, \dots, p_{i,n-k-1}, 0, \dots, \underbrace{1}_{i}, \dots, \underbrace{0}_{k-1})$$

$$\mathbf{h}_{j} = (0, \dots, \underbrace{1}_{j}, \dots, \underbrace{0}_{n-k-1}, p_{0j}, \dots, p_{ij}, \dots, p_{k-1,j})$$

$$\mathbf{g}_{i} \circ \mathbf{h}_{j} = p_{ij} \oplus p_{ij} = 0$$

$$\mathbf{G} \circ \mathbf{H}^{T} = \mathbf{0}$$



Parity check matrix **H**

As there is:

$$c = m \circ G$$

$$G \circ H^T = 0$$

hence

$$\mathbf{c} \circ \mathbf{H}^T = \mathbf{m} \circ \mathbf{G} \circ \mathbf{H}^T = \mathbf{0}$$

■ The codeword in systematic form is expressed as:

$$\mathbf{c} = (c_0, \ldots, c_i, \ldots, c_{n-k-1}, m_0, m_1, \ldots, m_{k-1})$$

$$\mathbf{h}_j = (0, \dots, \underbrace{1}_{j}, \dots, \underbrace{0}_{n-k-1}, p_{0j}, p_{1j}, \dots, p_{k-1,j})$$

Thus

$$\mathbf{c} \circ \mathbf{h}_{j} = c_{j} \oplus p_{0j} \bullet m_{0} \oplus p_{1j} \bullet m_{1}, \dots, p_{k-1,j} \bullet m_{k-1} = 0$$

$$c_{i} = p_{0i} \bullet m_{0} \oplus p_{1i} \bullet m_{1}, \dots, p_{k-1,i} \bullet m_{k-1}$$

■ It means parity check matrix **H** also specifies completely a given blocker code.

Qi Zhang (ASE) Linear Block Codes 13/02/2014 19 / 29

Parity check matrix **H**

Example 2.10: Determine the parity check matrix **H** for the linear block code $C_b(7,4)$ generated by the generator matrix:

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

Solution:

$$\mathbf{H} = \begin{bmatrix} I_{n-k} & P^T \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$



20 / 29

Qi Zhang (ASE) Linear Block Codes

- The components of the codeword $\mathbf{c} = (c_0, c_1, \dots, c_{n-1})$ are taken from GF(2), i.e., $c_i \in \text{GF}(2)$.
- The received vector is denoted by $\mathbf{r} = (r_0, r_1, \dots, r_{n-1})$, there is also $r_i \in \mathsf{GF}(2)$.
- Error pattern is modeled by $\mathbf{e} = (e_0, e_1, \dots, e_{n-1}), e_i \in \mathsf{GF}(2).$
- The error vector **e** has non-zero components in the positions when errors occur.
- What we are interested in is to detect error and correct the received vector.



Qi Zhang (ASE) Linear Block Codes 13/02/2014 21 / 29

There is

$$\mathbf{r} = \mathbf{c} \oplus \mathbf{e}$$

Hence

$$c = r \oplus e$$

Since any codeword fits the condition:

$$\mathbf{c} \circ \mathbf{H}^T = \mathbf{0}$$

an error-detection mechanism can be implemented based on the above expression:

$$s = r \circ H^{T}$$

$$= (c \oplus e) \circ H^{T}$$

$$= c \circ H^{T} \oplus e \circ H^{T} = e \circ H^{T}$$



22 / 29

- s is called syndrome vector;
- If syndrome vector is all-zero vector, then the received vector is a valid codeword;
- When the syndrome vector contains at least one non-zero component, an error is detected in the received vector.
- Note: it is possible that the syndrome vector can be the all-zero vector even though the errors occurs in the received vector.
- If error patterns are equal to one of the codewords, i.e., $\mathbf{e} = \mathbf{c}$, \mathbf{e} is not a all-zero component vector, there is

$$s = e \circ H^T = 0$$

- Q: How many undetectable non-zero error pattern exist?
- **A**: $2^k 1$



Example 2.11: For the linear block code $C_b(7,4)$, the parity check matrix is listed blow. Obtain the analytical expression of the syndrome vector bits.

$$\textbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$



Qi Zhang (ASE) Linear Block Codes 13/02/2014 24 / 29

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

- Solution:
- **Assuming** $\mathbf{r} = (r_0, r_1, r_2, r_3, r_4, r_5, r_6)$, then

$$\mathbf{s} = (s_0, s_1, s_2) = \mathbf{r} \circ \mathbf{H}^T = (r_0, r_1, r_2, r_3, r_4, r_5, r_6) \circ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix}$$

$$s_0 = r_0 \oplus r_3 \oplus r_5 \oplus r_6$$

$$s_1 = r_1 \oplus r_3 \oplus r_4 \oplus r_5$$

$$s_2 = r_2 \oplus r_4 \oplus r_5 \oplus r_6$$



Syndrome vector actually is dependent on the error vector. Thus the bits of the syndrome vector can be expressed as:



Qi Zhang (ASE) Linear Block Codes 13/02/2014 26 / 29

Example 2.12

Example 2.12: For the linear block code $C_b(7,4)$, the transmitted codeword \mathbf{c} is affected by channel noise and received as the vector $\mathbf{r} = (0001010)$. The syndrome vector is $\mathbf{s} = (001)$, so the syndrome bits can be expressed by components in the error vector as below. To decode the transmitted codeword \mathbf{c} .

$$\mathbf{H} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

$$\begin{array}{lll} 0 & = & e_0 \oplus e_3 \oplus e_5 \oplus e_6 \\ 0 & = & e_1 \oplus e_3 \oplus e_4 \oplus e_5 \\ 1 & = & e_2 \oplus e_4 \oplus e_5 \oplus e_6 \end{array}$$



Example 2.12

e_0	e_1	e_2	<i>e</i> ₃	e_4	e_5	e_6
0	0	1	0	0	0	0
1	1	1	1	0	0	0
1	0	0	0	0	0	1
0	1	0	1	0	0	1
0	0	0	1	0	1	0
1	1	0	0	0	1	0
0	1	1	0	0	1	1
1	0	1	1	0	1	1
0	1	0	0	1	0	0
1	0	0	1	1	0	0
1	1	1	0	1	0	1
0	0	1	1	1	0	1
1	0	1	0	1	1	0
0	1	1	1	1	1	1
1	1	0	1	1	1	1
0	0	0	0	1	1	1



Example 2.12

- There are $2^4 = 16$ different error patterns satisfy the equations;
- The probability of i errors occur is higher than that of i + 1 errors occur;
- In the channel like BSC, the error pattern with the smallest number of non-zero components is considered as the true error pattern.
- Therefore, for the previous case, $\mathbf{e} = (0010000)$ is considered as the true error pattern, so

$$\mathbf{c} = \mathbf{r} \oplus \mathbf{e} = (0001010) \oplus (0010000) = (0011010)$$



Qi Zhang (ASE) Linear Block Codes 13/02/2014 29 / 29