An introduction to coding theory

Adrish Banerjee

Department of Electrical Engineering
Indian Institute of Technology Kanpur
Kanpur, Uttar Pradesh
India

Jan. 23, 2017

Lecture #3A: Syndrome, error correction and error detection

Linear Block Codes

Outline of the lecture

- Syndrome and error detection

Linear Block Codes

Outline of the lecture

- Syndrome and error detection
- Syndrome and error correction

Syndrome and error detection

- Let $\mathbf{v} = (v_0, v_1, \cdots, v_{n-1})$ be a codeword from a binary (n,k) linear block code with generator matrix $\mathbf{G}$ and parity check matrix $\mathbf{H}$.

Syndrome and error detection

- Let $\mathbf{v} = (v_0, v_1, \cdots, v_{n-1})$ be a codeword from a binary (n,k) linear block code with generator matrix $\mathbf{G}$ and parity check matrix $\mathbf{H}$.
- Assume $\mathbf{v}$ is transmitted over a BSC, then binary received sequence,

$$\begin{aligned}
\mathbf{r} = (r_0, r_1, \cdots, r_{n-1}) &= \mathbf{v} + \mathbf{e} \quad \text{(modulo-2)} \\
&= (v_0, v_1, \cdots, v_{n-1}) + (e_0, e_1, \cdots, e_{n-1}) \\
&= (v_0 + e_0, v_1 + e_1, \cdots, v_{n-1} + e_{n-1}),
\end{aligned}$$

where the binary vector $\mathbf{e} = (e_0, e_1, \cdots, e_{n-1})$ is the *error pattern*.

Adrish Banerjee Department of Electrical Engineering Indian Institute of Technology Kanpur Kanpur, Uttar Pradesh India
An introduction to coding theory

Syndrome and error detection

- Let $\mathbf{v} = (v_0, v_1, \cdots, v_{n-1})$ be a codeword from a binary (n,k) linear block code with generator matrix $\mathbf{G}$ and parity check matrix $\mathbf{H}$.
- Assume $\mathbf{v}$ is transmitted over a BSC, then binary received sequence,

$$\begin{aligned} \mathbf{r} = (r_0, r_1, \cdots, r_{n-1}) &= \mathbf{v} + \mathbf{e} \quad \text{(modulo-2)} \\ &= (v_0, v_1, \cdots, v_{n-1}) + (e_0, e_1, \cdots, e_{n-1}) \\ &= (v_0 + e_0, v_1 + e_1, \cdots, v_{n-1} + e_{n-1}), \end{aligned}$$

  where the binary vector $\mathbf{e} = (e_0, e_1, \cdots, e_{n-1})$ is the *error pattern*.
- The "1's" in $\mathbf{e}$ represent *transmission errors*, i.e.,

$$e_i = \begin{cases} 1 & \text{if } r_i \neq v_i \\ 0 & \text{if } r_i = v_i, \end{cases}$$

  and $e_i = 1$ indicates that the i^{th} position in $\mathbf{r}$ has an error.

Syndrome and error detection

- After receiving $\mathbf{r}$, the decoder must determine if $\mathbf{r}$ contains errors (error detection), and locate the errors in $\mathbf{r}$ (error correction).

Syndrome and error detection

- After receiving $\mathbf{r}$, the decoder must determine if $\mathbf{r}$ contains errors (error detection), and locate the errors in $\mathbf{r}$ (error correction).
- *Error detection* is achieved by computing the (n-k) tuple

$$\mathbf{s} = (s_0, s_1, \cdots, s_{n-k-1}) = \mathbf{r}\mathbf{H}^{\mathbf{T}} \quad (\textit{syndrome})$$

Syndrome and error detection

- After receiving $\mathbf{r}$, the decoder must determine if $\mathbf{r}$ contains errors (error detection), and locate the errors in $\mathbf{r}$ (error correction).
- *Error detection* is achieved by computing the (n-k) tuple

$$\mathbf{s} = (s_0, s_1, \cdots, s_{n-k-1}) = \mathbf{r}\mathbf{H}^{\mathbf{T}} \quad (\textit{syndrome})$$

- $\mathbf{r}$ is a codeword if and only if $\mathbf{s} = \mathbf{r}\mathbf{H}^{\mathbf{T}} = \mathbf{0}$.

Syndrome and error detection

- After receiving $\mathbf{r}$, the decoder must determine if $\mathbf{r}$ contains errors (error detection), and locate the errors in $\mathbf{r}$ (error correction).
- *Error detection* is achieved by computing the (n-k) tuple

$$\mathbf{s} = (s_0, s_1, \cdots, s_{n-k-1}) = \mathbf{r}\mathbf{H}^\mathbf{T} \quad (\textit{syndrome})$$

- $\mathbf{r}$ is a codeword if and only if $\mathbf{s} = \mathbf{r}\mathbf{H}^\mathbf{T} = \mathbf{0}$.
- If $\mathbf{s} \neq \mathbf{0}$, $\mathbf{r}$ is not a codeword and transmission errors have been detected.

Syndrome and error detection

- After receiving $\mathbf{r}$, the decoder must determine if $\mathbf{r}$ contains errors (error detection), and locate the errors in $\mathbf{r}$ (error correction).
- *Error detection* is achieved by computing the (n-k) tuple

$$\mathbf{s} = (s_0, s_1, \cdots, s_{n-k-1}) = \mathbf{r}\mathbf{H}^\mathbf{T} \quad (\textit{syndrome})$$

- $\mathbf{r}$ is a codeword if and only if $\mathbf{s} = \mathbf{r}\mathbf{H}^\mathbf{T} = \mathbf{0}$.
- If $\mathbf{s} \neq \mathbf{0}$, $\mathbf{r}$ is not a codeword and transmission errors have been detected.
- If $\mathbf{s} = \mathbf{0}$, $\mathbf{r}$ is a codeword and no errors are detected. If $\mathbf{r}$ is a codeword other than the actual transmitted codeword, then an *undetected error* occurs. This happens whenever the error pattern $\mathbf{e}$ is a non-zero codeword.

Syndrome and error detection

- The syndrome $\mathbf{s}$ computed from the received vector $\mathbf{r}$ actually depends only on the error pattern $\mathbf{e}$, and not on the transmitted code word $\mathbf{v}$.

$$\mathbf{s} = \mathbf{r} \cdot \mathbf{H^T} = (\mathbf{v} + \mathbf{e})\mathbf{H^T} = \mathbf{v} \cdot \mathbf{H^T} + \mathbf{e} \cdot \mathbf{H^T}$$

Syndrome and error detection

- The syndrome $\mathbf{s}$ computed from the received vector $\mathbf{r}$ actually depends only on the error pattern $\mathbf{e}$, and not on the transmitted code word $\mathbf{v}$.

$$\mathbf{s} = \mathbf{r} \cdot \mathbf{H^T} = (\mathbf{v} + \mathbf{e})\mathbf{H^T} = \mathbf{v} \cdot \mathbf{H^T} + \mathbf{e} \cdot \mathbf{H^T}$$

- Since $\mathbf{v} \cdot \mathbf{H^T} = 0$,

$$\mathbf{s} = \mathbf{e} \cdot \mathbf{H^T}$$

Syndrome and error detection

Example 2.4: Consider a $(7, 4)$ linear code with parity-check matrix

$$\mathbf{H} = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right]$$

Let $\mathbf{r} = (0\ 1\ 0\ 0\ 0\ 0\ 1)$. The syndrome of $\mathbf{r}$ is

$$\begin{aligned} \mathbf{s} &= (s_0, s_1, s_2) = \mathbf{r} \cdot \mathbf{H^T} \\ &= (0\ 1\ 0\ 0\ 0\ 0\ 1) \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{bmatrix} = (1\ 1\ 1) \neq 0 \end{aligned}$$

Syndrome and error correction

- The syndrome $\mathbf{s}$ computed from the received vector $\mathbf{r}$ actually depends only on the error pattern $\mathbf{e}$, and not on the transmitted code word $\mathbf{v}$.

$$\mathbf{s} = \mathbf{r} \cdot \mathbf{H^T} = (\mathbf{v} + \mathbf{e})\mathbf{H^T} = \mathbf{e} \cdot \mathbf{H^T} \quad (\text{since } \mathbf{vH^T} = \mathbf{0})$$

Syndrome and error correction

- The syndrome $\mathbf{s}$ computed from the received vector $\mathbf{r}$ actually depends only on the error pattern $\mathbf{e}$, and not on the transmitted code word $\mathbf{v}$.

$$\mathbf{s} = \mathbf{r} \cdot \mathbf{H^T} = (\mathbf{v} + \mathbf{e})\mathbf{H^T} = \mathbf{e} \cdot \mathbf{H^T} \quad (\text{since } \mathbf{vH^T} = \mathbf{0})$$

- For error pattern $\mathbf{e} = \{e_0, e_1, \cdots, e_{n-1}\}$, and $\mathbf{H}$ given by

$$\begin{aligned}\mathbf{H} &= \left[\mathbf{I}_{n-k} : \mathbf{P^T}\right] \\ &= \left[\begin{array}{ccccc|cccc} 1 & 0 & 0 & \cdots & 0 & p_{0,0} & p_{1,0} & \cdots & p_{k-1,0} \\ 0 & 1 & 0 & \cdots & 0 & p_{0,1} & p_{1,1} & \cdots & p_{k-1,1} \\ 0 & 0 & 1 & \cdots & 0 & p_{0,2} & p_{1,2} & \cdots & p_{k-1,2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & p_{0,n-k-1} & p_{1,n-k-1} & \cdots & p_{k-1,n-k-1} \end{array}\right]\end{aligned}$$

the syndrome equations can be rewritten as

$$s_j = e_j + e_{n-k}p_{0j} + e_{n-k+1}p_{1j} + \cdots + e_{n-1}p_{k-1,j}, \qquad 0 \leq j \leq n-k$$

Syndrome and error correction

- This is a set of $n-k$ equations in n unknowns, $e_0, e_1, \cdots, e_{n-1}$.

Syndrome and error correction

- This is a set of $n-k$ equations in n unknowns, $e_0, e_1, \cdots, e_{n-1}$.
- The decoder must solve of these equations for the estimated error pattern, $\hat{\mathbf{e}}$. Estimated codeword is

$$\hat{\mathbf{v}} = \mathbf{r} + \hat{\mathbf{e}}$$

Syndrome and error correction

- This is a set of $n-k$ equations in n unknowns, $e_0, e_1, \cdots, e_{n-1}$.
- The decoder must solve of these equations for the estimated error pattern, $\hat{\mathbf{e}}$. Estimated codeword is

$$\hat{\mathbf{v}} = \mathbf{r} + \hat{\mathbf{e}}$$

- There are 2^k possible solutions to the syndrome equations and only one solution represents the true error pattern.

Syndrome and error correction

- This is a set of $n-k$ equations in n unknowns, $e_0, e_1, \cdots, e_{n-1}$.
- The decoder must solve of these equations for the estimated error pattern, $\hat{\mathbf{e}}$. Estimated codeword is

$$\hat{\mathbf{v}} = \mathbf{r} + \hat{\mathbf{e}}$$

- There are 2^k possible solutions to the syndrome equations and only one solution represents the true error pattern.
- To minimize the probability of a decoding error, the *the most probable* error pattern that satisfies the above equations is chosen as the true error vector.

Syndrome and error correction

- This is a set of $n-k$ equations in n unknowns, $e_0, e_1, \cdots, e_{n-1}$.
- The decoder must solve of these equations for the estimated error pattern, $\hat{\mathbf{e}}$. Estimated codeword is

$$\hat{\mathbf{v}} = \mathbf{r} + \hat{\mathbf{e}}$$

- There are 2^k possible solutions to the syndrome equations and only one solution represents the true error pattern.
- To minimize the probability of a decoding error, the *the most probable* error pattern that satisfies the above equations is chosen as the true error vector.
- Recall for BSC, the maximum likelihood decoder choose $\hat{\mathbf{v}}$ as the codeword $\hat{\mathbf{v}}$ that minimizes Hamming weight of the error pattern $\mathbf{e}$.

Syndrome and error correction

Example 3.1: Let

$$\mathbf{H} = \left[\begin{array}{ccc|cccc} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \end{array}\right]$$

Suppose $\mathbf{v} = (1\ 0\ 0\ 1\ 0\ 1\ 1)$ is transmitted and $\mathbf{r} = (1\ 0\ 0\ 1\ 0\ 0\ 1)$ is received. Then the syndrome of $\mathbf{r}$ is

$$\mathbf{s} = (s_0, s_1, s_2) = \mathbf{r}.\mathbf{H}^\mathbf{T} = (1\ 1\ 1)$$

Let $\mathbf{e} = (e_0, e_1, e_2, e_3, e_4, e_5, e_6)$ be the error pattern.

Syndrome and error correction

Since

$$\mathbf{s} = \mathbf{e}.\mathbf{H}^\mathbf{T}$$

we have the following 3 equations:

$$\begin{aligned} 1 &= e_0 + e_3 + e_5 + e_6 \\ 1 &= e_1 + e_3 + e_4 + e_5 \\ 1 &= e_2 + e_4 + e_5 + e_6 \end{aligned}$$

Adrish Banerjee Department of Electrical Engineering Indian Institute of Technology Kanpur Kanpur, Uttar Pradesh India
An introduction to coding theory

Syndrome and error correction

- The solutions are:

$$
\begin{array}{cc}
(0\ 0\ 0\ 0\ 0\ 1\ 0) & (1\ 0\ 1\ 0\ 0\ 1\ 1) \\
(1\ 1\ 0\ 1\ 0\ 1\ 0) & (0\ 1\ 1\ 1\ 0\ 1\ 1) \\
(0\ 1\ 1\ 0\ 1\ 1\ 0) & (1\ 1\ 0\ 0\ 1\ 1\ 1) \\
(1\ 0\ 1\ 1\ 1\ 1\ 0) & (0\ 0\ 0\ 1\ 1\ 1\ 1) \\
(1\ 1\ 1\ 0\ 0\ 0\ 0) & (0\ 1\ 0\ 0\ 0\ 0\ 1) \\
(0\ 0\ 1\ 1\ 0\ 0\ 0) & (1\ 0\ 0\ 1\ 0\ 0\ 1) \\
(1\ 0\ 0\ 0\ 1\ 0\ 0) & (0\ 0\ 1\ 0\ 1\ 0\ 1) \\
(0\ 1\ 0\ 1\ 1\ 0\ 0) & (1\ 1\ 1\ 1\ 1\ 0\ 1)
\end{array}
$$

Syndrome and error correction

- The solutions are:

$$
\begin{array}{cc}
(0\ 0\ 0\ 0\ 0\ 1\ 0) & (1\ 0\ 1\ 0\ 0\ 1\ 1) \\
(1\ 1\ 0\ 1\ 0\ 1\ 0) & (0\ 1\ 1\ 1\ 0\ 1\ 1) \\
(0\ 1\ 1\ 0\ 1\ 1\ 0) & (1\ 1\ 0\ 0\ 1\ 1\ 1) \\
(1\ 0\ 1\ 1\ 1\ 1\ 0) & (0\ 0\ 0\ 1\ 1\ 1\ 1) \\
(1\ 1\ 1\ 0\ 0\ 0\ 0) & (0\ 1\ 0\ 0\ 0\ 0\ 1) \\
(0\ 0\ 1\ 1\ 0\ 0\ 0) & (1\ 0\ 0\ 1\ 0\ 0\ 1) \\
(1\ 0\ 0\ 0\ 1\ 0\ 0) & (0\ 0\ 1\ 0\ 1\ 0\ 1) \\
(0\ 1\ 0\ 1\ 1\ 0\ 0) & (1\ 1\ 1\ 1\ 1\ 0\ 1)
\end{array}
$$

- Note that the true error pattern,

$$
\begin{aligned}
\mathbf{e} &= \mathbf{r} + \mathbf{v} \\
&= (1\ 0\ 0\ 1\ 0\ 0\ 1) + (1\ 0\ 0\ 1\ 0\ 1\ 1) \\
&= (0\ 0\ 0\ 0\ 0\ 1\ 0)
\end{aligned}
$$

is one of the 16 possible solutions. It is also the most probable solution.