MATH 525
Sections 2.10–2.12: Cosets and MLD for Linear Codes

October 5, 2020

Definition

Let C be a linear code of length n and let $u \in K^n$. The coset determined by u is the set

$$C + u = \{c + u \,|\, c \in C\}.$$

Example

Let C be the linear code

$$\{000000, 100101, 110011, 010110, 011001, 111100, 101010, 001111\}.$$

If $u = 100000$ then $C + u$ equals

$$\{100000, 000101, 010011, 110110, 111001, 011100, 001010, 101111\}.$$

If $u = 110010$ then $C + u$ equals

$$\{110010, 010111, 000001, 100100, 101011, 001110, 011000, 111101\}.$$

Remarks:

- C is always one of its cosets; in particular, C is the coset determined by $\mathbf{0}$ or by any of its codewords because $C = C + c$ for any $c \in C$.
- It is possible that $C + u = C + v$ even if $u \neq v$: Let C be the same code as in the previous example,

  $$\{000000, 100101, 110011, 010110, 011001, 111100, 101010, 001111\}.$$

  Then $C + 001110$ equals

  $$\{001110, 101011, 111101, 011000, 010111, 110010, 100100, 000001\}.$$

  Compare the last coset with the one determined by 110010 on the previous slide.

Properties of Cosets – Theorem 2.10.3:

Theorem

If C is a linear code of length n, then any two cosets of C have the same cardinality. Since C is one of its cosets, it follows that $|C + u| = |C|$ for any $u \in K^n$.

Theorem

If C is a linear code of length n, then any two cosets of C are either disjoint or coincide. More specifically, given $u, v \in K^n$, either $C + u = C + v$ or $(C + u) \cap (C + v) = \emptyset$.

Let C be a linear code of length n and dimension k. A few consequences of the above theorems are in order:

1. If two cosets share one word, then the two cosets coincide completely.
2. Every word in K^n belongs to exactly one coset of C.
3. $C + u = C + v$ if and only if $u + v \in C$.
4. The number of cosets of C equals 2^{n-k}.

Example

Write down all the cosets of the linear code $C = \{0000, 1011, 0101, 1110\}$.
To be worked out during the lecture.

Example (Cosets of a $(6,3)$ linear code)

$$
\begin{aligned}
C + 000000 &= \{000000, 100101, 110011, 010110, 011001, 111100, 101010, 001111\}\\
C + 100000 &= \{100000, 000101, 010011, 110110, 111001, 011100, 001010, 101111\}\\
C + 010000 &= \{010000, 110101, 100011, 000110, 001001, 101100, 111010, 011111\}\\
C + 110000 &= \{110000, 010101, 000011, 100110, 101001, 001100, 011010, 111111\}\\
C + 001000 &= \{001000, 101101, 111011, 011110, 010001, 110100, 100010, 000111\}\\
C + 000010 &= \{000010, 100111, 110001, 010100, 011011, 111110, 101000, 001101\}\\
C + 000001 &= \{000001, 100100, 110010, 010111, 011000, 111101, 101011, 001110\}\\
C + 000100 &= \{000100, 100001, 110111, 010010, 011101, 111000, 101110, 001011\}
\end{aligned}
$$

Maximum Likelihood Decoding for Linear Codes

- Let C be an (n, k) linear code with parity-check matrix H. The syndrome of the word $w \in K^n$ is defined as

$$s = w \cdot H.$$

  Note that the syndrome of w is a word in K^{n-k}.
- Important property:

$$wH = vH \Leftrightarrow wH + vH = 0 \Leftrightarrow (w + v)H = 0 \Leftrightarrow w + v \in C.$$

  In conclusion, $\mathrm{syn}(w) = \mathrm{syn}(v)$ if and only if w, v belong to the same coset of C. This is also saying that all the elements in a given coset share the same syndrome.
- There is a one-to-one correspondence between syndromes and cosets:

$$\text{syndromes} \longleftrightarrow \text{cosets} \quad \text{(bijection)}$$

- There exist 2^{n-k} different syndromes, one for each coset.

Maximum Likelihood Decoding for Linear Codes

- **Goal**: Given a received word, we want to decode it into the most likely sent codeword. Equivalently, we want to find the most likely error pattern. Note that if C is a code of length n, the error pattern associated with a received word r is a word in K^n.
- Suppose $v \in C$ is transmitted and the error pattern e occurs. Thus, the received word is $r = v + e$.
- $\mathrm{syn}(r) = s = rH = (v + e)H = eH$. This shows that the syndrome of r equals the syndrome of e.
- So, e belongs to the coset whose syndrome is s. Let that coset (whose syndrome equals s) be $C + u$.
- For MLD, e must be chosen as a word of minimum weight in $C + u$. For the smaller the weight of an error pattern, the most likely it is to have occurred.
- A word of minimum weight in a coset is called a coset leader.

Maximum Likelihood Decoding for Linear Codes

The above observations lead to a general decoding algorithm for linear codes. It is based on the one-to-one correspondence between coset leaders and syndromes. Usually, a table known as the standard decoding array (SDA) is formed: It consists of two columns, the first containing the 2^{n-k} coset leaders (error patterns) and the second containing their respective syndromes. We will illustrate it via a few examples.

Example

Let $C = \{0000, 1011, 0101, 1110\}$ be a linear code. Find a parity-check matrix for C, construct an SDA for C, and then decode the received word $r = 1010$.

Example

Let $C = \langle\{10101, 01110\}\rangle$. Find a parity-check matrix for C, construct an SDA for C, and then decode the received word $r = 11100$.

Reliability of IMLD for Linear Codes

Recall: Let C be a linear code and $v \in C$. Then $\theta_p(C, v) =$ probability that if v is sent over a BSC of reliability p, then IMLD correctly concludes that v was sent.

A coset leader is said to be *unique* if it is the only word of minimum weight in its coset. If $v \in C$ is transmitted, then IMLD will correctly conclude that v was sent if and only if an error pattern e equal to a unique coset leader occurs.

Thus,

$$\theta_p(C, v) = \theta_p(C, \mathbf{0}) = \sum_{w \in L} p^{n-\mathrm{wt}(w)}(1-p)^{\mathrm{wt}(w)}$$

where $L = L(\mathbf{0}) =$ set of coset leaders that are *unique*.

From now on for linear codes, we will just write $\theta_p(C)$ in place of $\theta_p(C, v)$.

Example

Find $\theta_p(C)$ where $C = \langle\{101010, 011011, 000111\}\rangle$.

- Let C be a linear code of length n and distance d. Given $u \in K^n$, is u a unique coset leader?
- We do not have an "easy" criterion to help us answer the above question, but in many situations the following result is useful:

  if $\mathrm{wt}(u) \leq \lfloor \frac{d-1}{2} \rfloor$*, then* u *is a unique coset leader.*
- u (of weight $\leq \lfloor \frac{d-1}{2} \rfloor$) must be a coset leader for otherwise u would be in the coset $C + v$ for some $v \neq u$ with $\mathrm{wt}(v) \leq \mathrm{wt}(u)$. Thus, $u + v \in C$. The weight of $u + v$ would be less than d, a contradiction. Finally, u is unique for otherwise there would exist $v \in C + u$ of weight equal to the weight of u. As a result, $u + v$ would again be a codeword of weight less than d, a contradiction.