MAT4006: Introduction to Coding Theory

Lecture 02: Decoding Methods, Error Detection

Instructor: Zitan Chen
Scribe: Siqi Yao

Spring 2025

1 Maximum Likelihood Decoding (MLD)

Suppose a code C is used for communication over a channel. If a word $y \in B^n$ is received, the MLD rule will decode y to a codeword $\tilde{x} \in C$ such that:

$$\Pr(y \text{ received} \mid \tilde{x} \text{ sent}) = \max_{x \in C} \Pr(y \text{ received} \mid x \text{ sent}).$$

Ties are broken arbitrarily.

Question

What theorem is this method based on and what assumptions does it make about the prior distribution of x? (See the example last lecture)

Example (MLD on BSC)

Suppose a code C is used over a BSC with $p < \frac{1}{2}$. Then for $x, y \in \{0,1\}^n$, we have:

$$\Pr(y \text{ received} \mid x \text{ sent}) = \prod_{i=1}^{n} \Pr(y_i \text{ received} \mid x_i \text{ sent}) = (1-p)^{n-e} \cdot p^e,$$

where n is the block length, and $e := |\{i \mid x_i \neq y_i\}|$ is the number of bit errors.

Since $p < \frac{1}{2}$, we know that $p < 1-p$, and the quantity

$$(1-p)^{n-e} \cdot p^e$$

is **strictly decreasing** with respect to e.

Hence, the probability $\Pr(y \mid x)$ is maximized when $x \in C$ minimizes e, i.e., minimizes the Hamming distance to y.

Proof on Monotonicity

Let $f(e) = (1-p)^{n-e} \cdot p^e$. Then

$$\frac{df}{de} = (1-p)^{n-e} p^e \ln(p) - p^e (1-p)^{n-e} \ln(1-p) = f(e) \left(\ln p - \ln(1-p)\right).$$

Since $\ln p < \ln(1-p)$ when $p < \frac{1}{2}$, we have $\frac{df}{de} < 0$, so $f(e)$ decreases with e.

2 Hamming Distance

Definition 2.1 (Hamming Distance)

Let $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ be words of length n over an alphabet A. The *Hamming distance* from x to y, denoted by $d(x, y)$, is defined as:

$$d(x,y) = \sum_{i=1}^{n} d(x_i, y_i), \quad \text{where} \quad d(x_i, y_i) := \begin{cases} 1 & \text{if } x_i \neq y_i, \\ 0 & \text{if } x_i = y_i. \end{cases}$$

Examples

1. Let $A = \{0, 1\}$, and define

$$x = 01010, \quad y = 01101, \quad z = 11101.$$

   Then:

$$d(x,y) = 3, \quad d(y,z) = 1, \quad d(x,z) = 4.$$

2. Let $A = \{0, 1, 2, 3\}$, and define

$$x = 1234, \quad y = 1423.$$

   Then:

$$d(x,y) = 3.$$

Proposition 2.1 (The Hamming distance is a metric)

Let $x, y, z \in A^n$. Then:

1. $0 \leq d(x, y) \leq n$ (positivity)
2. $d(x, y) = 0 \iff x = y$
3. $d(x, y) = d(y, x)$ (symmetry)
4. $d(x, y) + d(y, z) \geq d(x, z)$ (triangle inequality)

In particular, $d(x, \mathbf{0})$ (i.e., the Hamming distance from x to the all-zero word) is called the **Hamming weight** of x.

3 Minimum Distance Decoding

Suppose a code C is used. If y is received, the minimum distance decoding rule decodes y to a codeword $\tilde{x} \in C$ such that:

$$\tilde{x} = \arg\min_{x \in C} d(x, y).$$

Theorem 2.2

Consider a q-ary symmetric channel where

$$p = \Pr(b \text{ received} \mid a \text{ sent}) \quad \text{for } a \neq b, \quad \text{with } p < \frac{1}{q}.$$

Then, the **maximum likelihood decoding (MLD)** rule is equivalent to the **minimum distance decoding** rule.

Proof. Let C be a code of length n. Suppose y is the received word.

For $e = 1, 2, \ldots, n$ and $x \in C$, the Hamming distance between x and y is e, i.e.,

$$d(x, y) = e,$$

if and only if the probability of receiving y given x was sent is:

$$\begin{aligned}\Pr(y \text{ received} \mid x \text{ sent}) &= \prod_{i=1}^{n} \Pr(y_i \text{ received} \mid x_i \text{ sent}) \\ &= p^e \left(1 - (q-1)p\right)^{n-e}.\end{aligned}$$

Since $p < \frac{1}{q}$, we have $p < 1 - (q-1)p$, and thus

$$p^e \left(1 - (q-1)p\right)^{n-e}$$

is a decreasing function in e.

Therefore, the MLD rule, which selects $\tilde{x} \in C$ maximizing $\Pr(y \mid x)$, is equivalent to selecting $\tilde{x} \in C$ minimizing $d(\tilde{x}, y)$. That is,

$$\tilde{x} = \arg\max_{x \in C} \Pr(y \text{ received} \mid x \text{ sent}) = \arg\min_{x \in C} d(x, y).$$

□

Definition 2.2 (Distance of Codes)

For a code C with $|C| \geq 2$, the (minimum) distance of C, denoted by $d(C)$, is

$$d(C) = \min\{d(x, y) \mid x, y \in C,\ x \neq y\}.$$

Such a code is often referred to as an (n, M, d)-code, where n is the block length, $M = |C|$ is the number of codewords, and $d = d(C)$ is the minimum distance.

Examples

Let $C = \{00000,\ 00111,\ 11111\}$ over $A = \{0, 1\}$. We compute the pairwise Hamming distances:

$$\begin{aligned}d(00000, 00111) &= 3, \\ d(00111, 11111) &= 2, \\ d(00000, 11111) &= 5.\end{aligned}$$

Therefore, the minimum distance is

$$d(C) = 2.$$

4 Error Detection

Consider C. Suppose y is received.

- If $y \notin C$, there are errors.
- If $y \in C$, we do not know if there are errors.

Definition 2.3 (Error Detection)

We say a code $C \subseteq A^n$ can detect u errors (where $u \geq 1$) if for any $x \in C$, $y \in A^n$, such that $0 < d(x, y) \leq u$, it holds that $y \notin C$.

Theorem 2.3

Let u be a positive integer. A code C can detect u errors if and only if

$$d(C) \geq u + 1.$$

Proof. Suppose $d(C) \geq u + 1$. Let $x \in C$ and y be the received word. If $d(x, y) = u \geq 1$, then $y \notin C$, so the error can be detected.

Conversely, suppose $d(C) \leq u$. Then there exist $x_1, x_2 \in C$ such that

$$d(x_1, x_2) = d(C) \leq u.$$

If x_1 is transmitted and x_2 is received, we cannot detect whether there has been an error. □

We will see error correction next class.