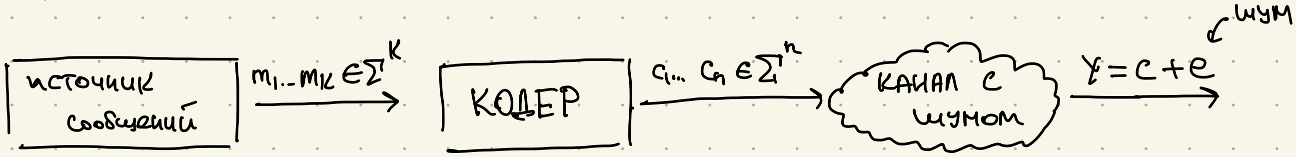
Лекция №1

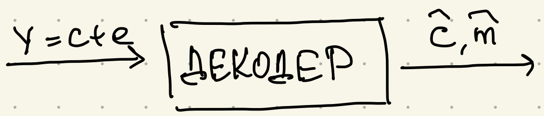
Линейные коды. Основные определения

① Модель канала с шумом

Σ – конечный алфавит (например, $\Sigma = \mathbb{F}_q$)

$m \in \Sigma^k$ – сообщение





Декодирование успешно, если $(\hat{c}, \hat{m}) = (c, m)$

$m \in \Sigma^k$ – сообщение

$c \in \Sigma^n$ – кодовое слово

② Основные определения

Опр. 1 ① Для строк $x, y \in \Sigma^n$, расстояние Хэмминга м/д x и y,

$\Delta(x, y)$ – кол-во позиций, в которых значения x и y различны)

$$\Delta(x, y) = |\{ i \mid x_i \neq y_i \}|$$

относительное расстояние Хэмминга: $\delta(x, y) = \dfrac{\Delta(x, y)}{n}$

Расстояние Хэмминга определяет метрику на Σ^n.

② Вес Хэмминга для $x \in \Sigma^n$ – кол-во ненулевых символов в x, т.е.

$$wt(x) = |\{ i \mid x_i \neq 0 \}|$$

Имеем,
$$wt(x - y) = \Delta(x, y)$$
$$wt(x) = \Delta(x, 0)$$

③ Блочный код C, исправляющий ошибки, длины n
(error correcting code)
над конечным алфавитом Σ — это подмножество Σ^n.

Элементы C – кодовые слова.

Если $\Sigma = \mathbb{F}_q$, то C – q-арный код

$\Sigma = \mathbb{F}_2$, то C – бинарный код

Кодирующее отображение $\mathcal{E}: M \to C$, $c \in \operatorname{Im}(\mathcal{E})$
$$m \mapsto c$$

④ Параметры кода $C \subseteq \Sigma^n$

- скорость кода (code rate) $R(C) = \dfrac{\lg |C|}{n \cdot \lg |\Sigma|}$

- размерность кода (code dimension) $\dfrac{\lg |C|}{\lg |\Sigma|}$

- минимальное расстояние C — мин. расстояние м/д двумя неодинаковыми кодовыми словами

$$d(C) = \min_{\substack{x, y \in C \\ x \neq y}} \Delta(x, y)$$

(т.е. $\forall c_1, c_2 \in C$ отличаются как мин. на $d(C)$ позиций)

относительное min. расстояние $\delta(C) = \dfrac{d(C)}{n}$

Примеры

① Код проверки на чётность на $\mathbb{F}_2$
(parity check code)

$$\{0,1\}^k \to \{0,1\}^{k+1}$$
$$m_1 \dots m_k \mapsto m_1 \dots m_k \, \| \sum_{i=1}^{k} m_i$$

$$R = \frac{k}{k+1}, \quad d = 2$$

② Код с повторением
(repetition code)

$$\{0,1\}^k \to \{0,1\}^n \quad (\exists\ \tfrac{n}{k} \in \mathbb{Z})$$

$$m_1 \dots m_k \mapsto \underbrace{m_1..m_k \| \dots \| m_1..m_k}_{\frac{n}{k} \text{ раз}}$$

$$R = \frac{k}{n},\ d = \frac{n}{k}$$

③ Код Хэмминга (1950)

$$\{0,1\}^4 \to \{0,1\}^7$$

$$m_1m_2m_3m_4 \mapsto m_1m_2m_3m_4 \| m_2 \oplus m_3 \oplus m_4 \| m_1 \oplus m_3 \oplus m_4 \| m_1 + m_2 + m_4$$

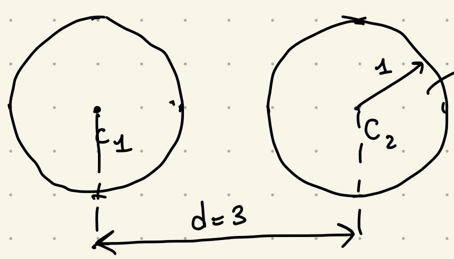
$R = \frac{4}{7}$, $d = 3$ (пока неочевидно)

исправление ошибки = нахождение вектора ошибки (e) $\Rightarrow$ нахождение кодового слова

Лемма 1 Код с минимальным расстоянием d исправляет $\lfloor \frac{d-1}{2} \rfloor$ ошибок.

Если d-чётное, C исправляет $\frac{d-2}{2}$ ошибок

т.е. если $d = 3$, то код исправляет 1 ошибку.



$B(c_2, r) = \{x : \Delta(x, c_2) \le r\}$

$\forall y \in B(c, 1)$ – c - ближайшее кодовое слово

③ Линейный код (q-простое, $\Sigma = \mathbb{F}_q$)

Определения:

1. Для $n \geq k > 1$, $[n,k]_q$ -линейный код C – подпространство в $\mathbb{F}_q^n$ размерности k.

$\Downarrow$

$\exists\ c_1 \dots c_k$ – лин. незав. над $\mathbb{F}_q^n$ вектора, образующие базис C

2. $C \subseteq \mathbb{F}_q^n$ – лин. код размерности k.

Матрица $G \in \mathbb{F}_q^{k\times n}$ называется порождающей / образующей (generator matrix) кода C, если строки (!) G образуют базис C $\Rightarrow$

экв. опр-ие лин. кода: $C = \{c \in \mathbb{F}_q^n : c = u \cdot G,\ u \in \mathbb{F}_q^k\}$

$$C = \; \vdash u \dashv \; \boxed{k \quad \overset{n}{G}}$$

Обозначение: лин. $[n,k]_q$-код с min. расстоянием d обозначается $[n,k,d]$-кодом.

Пример 3 Код Хэмминга – $[7,4,3]_2$-код с порождающей матр.

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

Так как $\mathrm{rank}(G) = k$, то $\exists$ k столбцов $G_{i_1} \dots G_{i_k}$, т.ч. $[G_{i_1} | \dots | G_{i_k}] \in \mathbb{F}_q^{k\times k}$ – обратимая $\Rightarrow$ $\forall$ G имеет т.н. систематическую форму

$$G = [\, I_k \mid \underset{\in \mathbb{F}_q^{k\times(n-k)}}{A} \,]$$

③ Проверочная матрица (parity-check) $[n,k,d]_q$-кода C – это матрица $H \in \mathbb{F}_q^{(n-k)\times k}$ т.ч. $H \cdot c = 0$ $\forall c \in C$

экв. опр-ие кода C: $C = \ker(H)$

$$\mathrm{rank}(H) = n - \dim \ker H = n - k$$

Имеем равенства $HG^T = 0$, $GH^T = 0$

Пример 3. Проверочная матрица кода Хэмминга $H = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$

Лемма 2 (min. расстояние через проверочную матрицу)

$\exists H$ — проверочная матрица лин. кода $C \neq \{0\}$. Тогда $d(C)$ — наибольшее целое d, т.ч. любые $(d-1)$ столбцы H лин. независимы.

$\triangleleft$ $\exists$ $H = \begin{bmatrix} | & & | \\ h_1 & \dots & h_n \\ | & & | \end{bmatrix}$, $\quad c = (c_1 \dots c_n)$, $wt(c) > 0$

$Hc = 0$

$\forall$ J - мн-во индексов т.ч. $c_j \neq 0$, $j \in J$, $|J| = wt(c)$

$\sum c_j \vec{h}_j = 0 \Rightarrow$ мы нашли $wt(c)$ лин. завис. столбцов в H.

Обратно, $\exists$ t лин. завис. столбцов в H $\Rightarrow$ коэфф-ты лин. комбинации этих столбцов обр. ненулевое кодовое слово веса t.

Т.к. d — min. возможное значение для t $\Rightarrow \nexists$ кодового слова веса $d-1$ (хотя как min. одно кодовое слово веса d существует) $\blacktriangleright$.

Вывод: $d(\text{Хэмминга}) = 3$

Теорема 3 (Граница Хэмминга) C — бинарный код длины n р-ти k.

Тогда

$$|C| \leq \frac{2^n}{\sum\limits_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i}} \qquad (1)$$

$\triangleleft$ $\exists$ $d = 3$.

Для $c \in C$, определим шар $B(c, \lfloor \frac{d-1}{2} \rfloor = 1) = \{ y \in \{0,1\}^n \mid \Delta(c, y) \leq 1 \}$

$B(c) \cap B(c') = \emptyset$, т.к. $d = 3$ $\forall c \neq c'$ $\Rightarrow$

$\Rightarrow 2^n \geq \left| \bigcup\limits_{c \in C} B(c) \right| = \sum\limits_{c \in C} |B(c)| = |C| \cdot (n+1)$

В общем случае, $\left| B(c, \lfloor \frac{d-1}{2} \rfloor) \right| = \sum\limits_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i}$ $\blacktriangleright$

Коды, для которых граница (1) выполняется с равенством, называются **совершенными**.