НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Модуль 2. Линейные блоковые коды

Декодирование линейных кодов. Синдромное декодирование. Стандартное расположение. Граница Хэмминга. Операции над кодами.

Иванов Ф. И.
к.ф.-м.н., доцент

Национальный исследовательский университет
«Высшая школа экономики»

12 июня 2020 г.

Постановка задачи

Пусть по каналу связи отправлен кодовый вектор

$$\mathbf{u} = (u_1, u_2, ..., u_n), \mathbf{u} \in A$$

где A есть кодовое подпространство, и в канале произошла ошибка, изображаемая вектором

$$\mathbf{e} = (e_1, e_2, ..., e_n).$$

На приёмном конце принят вектор

$$\mathbf{v} = \mathbf{u} + \mathbf{e} = (u_1 + e_1, u_2 + e_2, ..., u_n + e_n), \mathbf{v} \in V.$$

и декодер вычисляет произведение $\mathbf{vH}^T$:

$$\mathbf{vH}^T = (\mathbf{u} + \mathbf{e})\mathbf{H}^T = \mathbf{uH}^T + \mathbf{eH}^T = \mathbf{eH}^T$$

Задача: по известным $\mathbf{S}$ и $\mathbf{H}$ найти такое $\mathbf{e}$ наименьшего веса, что

$$\mathbf{S} = \mathbf{eH}^T.$$

Поиск вектора ошибки

Пусть

$$V = A_0 \cup A_1 \cup A_2 \cup \ldots \cup A_{q^{n-k}-1}$$

есть разложение пространства V по подпространству $A = A_0$, и A_i — смежные классы. Каждый вектор $\mathbf{v}$ принадлежит одному и только одному смежному классу.

Теорема

Все векторы одного и того же смежного класса имеют одинаковые синдромы, и различным смежным классам отвечают различные синдромы.

Доказательство

Пусть $\mathbf{v}_1$ *и* $\mathbf{v}_2 \in A_i$, *и пусть* $\mathbf{v}_1\mathbf{H}^T = \mathbf{S}_1$, $\mathbf{v}_2\mathbf{H}^T = \mathbf{S}_2$;
$\mathbf{S}_1 - \mathbf{S}_2 = (\mathbf{v}_1 - \mathbf{v}_2)\mathbf{H}^T$. *Так как* $\mathbf{v}_1$ *и* $\mathbf{v}_2 \in A_i$, *то* $\mathbf{v}_1 - \mathbf{v}_2 \in A$ *откуда* $\mathbf{S}_1 = \mathbf{S}_2$. *Если же* $\mathbf{v}_1 \in A_{i_1}$, $\mathbf{v}_2 \in A_{i_2}$,*то* $\mathbf{v}_1 - \mathbf{v}_2 \notin A$, $\mathbf{v}\mathbf{H}^T \neq 0$,$\mathbf{v}_1\mathbf{H}^T \neq \mathbf{v}_2\mathbf{H}^T$, $\mathbf{S}_1 \neq \mathbf{S}_2$, *что и требовалось.*

Выбор вектора ошибки из смежного класса

Уравнение $\mathbf{S} = \mathbf{e}\mathbf{H}^T$ имеет в точности q^k решений.
Задача: какой вектор $\mathbf{e} \in A_i$ выбрать из q^k возможных векторов?
Соглашение, которое лежит в основе выбора вектора-ошибки, основано на принципе декодирования по максимальному правдоподобию. Оно состоит в том, что вектором-ошибкой считается вектор минимального веса в соответствующем смежном классе. Это означает, что либо в смежном классе содержится единственный вектор $\mathbf{e}$ минимального веса $w \leq t$, и тогда истинный вектор $\mathbf{u}$ получается в виде разности $\mathbf{u} = \mathbf{v} - \mathbf{e}$, либо в смежном классе имеется по крайней мере два вектора $\mathbf{e}_1$, $\mathbf{e}_2$ минимального веса w, и тогда нет оснований определить, какой из этих векторов является вектором-ошибкой.

Основные теоремы синдромного декодирования

Теорема

Линейный код A исправляет все независимые ошибки кратности t и менее тогда и только тогда, когда все векторы веса t и менее принадлежат различным смежным классам.

Теорема

Минимальное расстояние линейного кода равно $d = 2t + 1$ тогда и только тогда, когда все векторы веса $1, 2, \dots, t$ принадлежат различным смежным классам.

Алгоритм синдромного декодирования

- Вход: принятый вектор $\mathbf{v}$, проверочная матрица $\mathbf{H}$
- Выход: кодовое слово $\mathbf{u}$ или отказ от декодирования

Алгоритм:

1. Вычисляем $\mathbf{S} = \mathbf{v}\mathbf{H}^T = \mathbf{e}\mathbf{H}^T$
2. По вычисленному $\mathbf{S}$ находим смежный класс A_i кода A
3. Ищем вектор минимального веса в A_i: если их несколько, то возвращаем отказ от декодирования, если вектор $\mathbf{e}$ единственный, то переходим к следующему шагу
4. Возвращаем $\mathbf{u} = \mathbf{v} - \mathbf{e}$

Стандартное расположение

Алгоритм декодирования через стандартное расположение:

1. Выпишем все векторы кодового подпространства слева направо, начиная с нулевого вектора.
2. Расположим произвольный смежный класс под кодовым подпространством, начиная с вектора $\mathbf{e}$ минимального веса (лидер смежного класса), так чтобы под кодовым вектором $\mathbf{u}$ находился вектор $\mathbf{v} = \mathbf{u} + \mathbf{e}$ смежного класса.
3. Приняв вектор $\mathbf{v}$, и вчислив его синдром $\mathbf{S}$, определяем отвечающий ему смежный класс, а в нём находим вектор $\mathbf{v}$.
4. Непосредственно над ним в первой строке таблицы находится передававшийся вектор $\mathbf{u}$.

Пример

Пусть линейный код над $GF(2)$ имеет параметры $n = 5, k = 2, n - k = 3, d = 3$.

$$\mathbf{G} = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{pmatrix},$$

а проверочная

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

Тогда стандартное расположение имеет вид:

$$\mathbf{A_G^0} = (\mathbf{00000}), (\mathbf{10111}), (\mathbf{01011}), (\mathbf{11100})$$

$$\begin{array}{llll}
A_G^1 = A_G^0 + (10000) = \{(10000) & (00111) & (11011) & (01100)\}, \\
A_G^2 = A_G^0 + (01000) = \{(01000) & (11111) & (00011) & (10100)\}, \\
A_G^3 = A_G^0 + (00100) = \{(00100) & (10011) & (01111) & (11000)\}, \\
A_G^4 = A_G^0 + (00010) = \{(00010) & (10101) & (01001) & (11110)\}, \\
A_G^5 = A_G^0 + (00001) = \{(00001) & (10110) & (01010) & (11101)\}, \\
A_G^6 = A_G^0 + (10001) = \{10001) & (00110) & (11010) & (01101)\}, \\
A_G^7 = A_G^0 + (10010) = \{(10010) & (00101) & (11001) & (01110)\},
\end{array}$$

Пример - продолжение

A_G^i, $(i = 1, 2, \ldots, 7)$ — смежные классы. Отвечающие им синдромы

$$\mathbf{S}_1 = (111), \mathbf{S}_2 = (011), \mathbf{S}_3 = (100), \mathbf{S}_4 = (010),$$

$$\mathbf{S}_5 = (001), \mathbf{S}_6 = (110), \mathbf{S}_7 = (101).$$

5 первых смежных классов имеют своих лидеров - векторы веса 1, а значит код исправляет все одиночные ошибки. Смежные классы $6 - 7$ имеется по два вектора веса 2, остальные — большего веса. Лидеров в этих классах нет. Таким образом, целые два смежных класса не используются для исправления ошибок.

Совершеннный и квазисовершенный коды

Определение

Линейный код называется совершенным, если все q^{n-k} *смежных классов имеют своих лидеров, и ими являются все векторы веса* t *и менее.*

Определение

Линейный код называется квазисовершенным, если все q^{n-k} *смежных классов имеют своих лидеров, и ими являются все векторы веса* t *и менее, а также некоторые векторы веса* $t+1$.

Граница Хэмминга

Теорема

Параметры любого $(n,k,d = 2t+1)$ *кода удовлетворяют следующему неравенству:*

$$\sum_{i=0}^{t} \binom{n}{i} (q-1)^i \leq q^{n-k}$$

или

$$R = \frac{k}{n} \leq 1 - \frac{1}{n} \log_q \sum_{i=0}^{t} \binom{n}{i} (q-1)^i$$

Равенство достигается только для совершенных кодов

Операции над кодами

- **Расширение** двоичного (n, k, d)-кода — добавление проверки на четность
- **Выкалывание.** Эта операция — обратная расширению и состоит в удалении одного проверочного символа.
- **Выбрасывание.** Удаление из двоичного (n, k, d)-кода всех векторов нечетного веса.
- **Пополнение.** Добавление к двоичному (n, k, d)-коду сплошь единичного вектора, если он еще не принадлежит коду.
- **Удлинение.** Эта операция состоит в последовательном выполнении двух операций — пополнения и расширения.
- **Укорочение.** Фиксируем произвольный столбец (n, k, d)-кода и выбираем только те векторы, которые в данном столбце содержат 0.

Мажоритарное декодирование

Пусть проверочной матрицей двоичного линейного (6,3)-кода будет:

$$H = \left[\begin{array}{cccccc} 0 & 1 & 1 & \vdots & 1 & 0 & 0 \\ 1 & 1 & 0 & \vdots & 0 & 1 & 0 \\ 1 & 0 & 1 & \vdots & 0 & 0 & 1 \end{array}\right]$$

и пусть $\mathbf{u} = (a_1, a_2, a_3, a_4, a_5, a_6)$ — произвольный кодовый вектор, где первые три символа информационные. Система $\mathbf{uH}^T = \mathbf{0}$ равносильна:

$$a_2 + a_3 + a_4 = 0, a_1 + a_2 + a_5 = 0, a_1 + a_3 + a_6 = 0$$

Для a_1:

$$a_1 = a_2 + a_5,$$

$$a_1 = a_3 + a_6,$$

$$a_1 = a_1$$

Мажоритарное декодирование

В реальной передаче на приёмном конце имеют дело с вектором $\mathbf{v} = (a_1,a_2,a_3,a_4,a_5,a_6)$, где для любого $i = 1,2,...,6$. символ a_i может отличаться от a_i:

$$a_1' = a_2' + a_5',$$

$$a_1' = a_3' + a_6',$$

$$a_1' = a_1'$$

Если ошибок не было, то все 3 уравнения дадут одно и то же решение. Если ошибка была одна, то 2 из трех уравнений дадут одно значение, а третье - другое. Ошибка будет исправлена голосованием большинства.

Мажоритарное декодирование

Аналогичные системы для a_2 и a_3:

$$\begin{gathered} a_2 = a_3 + a_4, \\ a_2 = a_1 + a_5, \\ a_2 = a_2. \end{gathered}$$

$$\begin{gathered} a_3 = a_2 + a_4, \\ a_3 = a_1 + a_6, \\ a_3 = a_3. \end{gathered}$$

То есть код исправляет любую одну ошибку мажоритарным методом.

Пример $t = 3$

$$H = \left[\begin{array}{cccccccccccccccc}
0 & 0 & 1 & 1 & : & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & : & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & : & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 1 & : & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & : & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & : & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & : & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 0 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 1 & 1 & : & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{array}\right].$$

Рис.: (15,4) код, исправляющий 3 ошибки