§ 6. Декодирование линейных кодов

Словарное декодирование — тривиальная форма декодирования, при котором составляется список всех возможных принимаемых слов и каждому слову из списка сопоставляется кодовое слово (или несколько кодовых слов), к которому оно может быть декодировано. При декодировании каждое принятое слово ищется в составленном таким образом словаре и декодируется в одно из кодовых слов, перечисленных рядом.

Декодирование в ближайшее кодовое слово — обычно применяемое правило декодирования, согласно которому ищется первое кодовое слово c, ближайшее в смысле Хэмминга к принятому слову x, т.е. $c = \arg\min_y d_H(y, x)$. Данное правило базируется на предположении, что малое число ошибок встречается чаще, чем большее (§2, (D)).

В случае линейного кода C длины n над полем F принято считать, что канал передачи данных добавляет *вектор ошибки*. Пусть передается кодовое слово c и принято слово x; тогда шум в канале дает эффект добавления к c вектора ошибки $e = x - c \in F^n$, так что $x = c + e$. Задача декодирования заключается в вычислении по заданному x (одного из) векторов c и e. Вес вектора e равен числу позиций, в которых c и x различаются; поэтому в условиях канала mSC(p) (при $p < 1/m$), согласно принципу (D) мы должны искать вектор ошибки e минимального веса.

Введем отношение эквивалентности $\sim_C$ на F^n (относительно кода C). Будем говорить, что x и y эквивалентны относительно кода C ($x \sim_C y$), если $x - y \in C$. Классы эквивалентности (которые образуют разбиение пространства векторов), называются *смежными классами*. Вектор ошибки e для принятого вектора x должен находиться в том же самом смежном классе, поскольку $x \sim_C e$ (иначе говоря, $x + C = e + C$). Выделим в каждом смежном классе вектор наименьшего веса и назовем его *лидером* смежного класса. Если в смежном классе несколько векторов имеют наименьший вес, то выберем один из них. Вектор 0 является единственным лидером самого кода.

Декодирование по лидеру смежного класса представляет собой общую технику декодирования в случае линейных кодов. Когда принято слово x, нам неизвестна добавленная ошибка, но мы знаем, что она принадлежит смежному классу $x + C$. Таким образом, если $\hat{e}$ — лидер данного смежного класса, то $\hat{e}$ — один из наиболее вероятных векторов ошибки, и мы можем декодировать x в $\hat{c} = x - \hat{e}$. Рассмотрим два метода конкретной реализации декодирования по лидеру смежного класса.

Декодирование по стандартной таблице.

Пусть $k = |C|$ — размер C, $r = |F^n|/|C|$ — число различных смежных классов C. Перечислим все кодовые слова:

$$C = \{c_1 = 0, c_2, \ldots, c_k\}$$

и все лидеры смежных классов:

$$\{e_1 = 0, e_2, \ldots, e_r\},$$

(по одному для каждого смежного класса). Построим так называемую *стандартную таблицу*:

$$\begin{array}{cccccl}
c_1 & c_2 & \ldots & c_k & & \text{— строка кодовых слов} \\
e_2 + c_1 & e_2 + c_2 & \ldots & e_2 + c_k & & \text{— остальные смежные классы} \\
\ldots & \ldots & \ldots & \ldots & & \\
e_r + c_1 & e_r + c_2 & \ldots & e_r + c_k & &
\end{array}$$

в которой первая строка содержит кодовые слова, первый столбец содержит лидеры смежных классов, а на пересечении i-й строки и j-го столбца стоит элемент $c_j + e_i$. Каждый вектор из F^n содержится в стандартной таблице ровно один раз, причем ее i-я строка есть смежный класс $e_i + C$.

Процесс декодирования заключается в следующем. Если принято слово x, то найдем ее вхождение в таблицу: $x = e_i + c_j$, после чего будем считать, что вектор ошибки совпадает с e_i и декодируем x к $\hat{c} = x - e_i = c_j$.

Пример. Для двоичного $[5, 2]$-кода с матрицами

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

имеется 4 кодовых слова, которые разместим в первой строке стандартной таблицы:

$$\begin{array}{llll} 00000 & 10101 & 01110 & 11011 \\ 00001 & 10100 & 01111 & 11010 \\ 00010 & 10110 & & \\ 00100 & 10001 & & \\ 01000 & 11101 & & \\ 10000 & 00101 & & \\ 11000 & 01101 & & \\ 10010 & 00111 & & \end{array}$$

(заполните таблицу до конца!).

Метод декодирования по стандартной таблице имеет лишь теоретический и исторический интерес, но весьма небольшую практическую ценность, поскольку расточительно использует большой массив памяти.

Декодирование по синдрому.

Второй метод декодирования использует проверочную матрицу кода. Пусть C — линейный $[n,k]$-код, H — его проверочная матрица. Это значит, что $x \in C \iff Hx^t = 0$. Вектор-столбец Hx^t длины $r = n-k$ называется *синдромом* вектора x.

Поскольку синдром кодового слова есть 0, два вектора x и e, различающиеся на кодовое слово c, имеют один и тот же синдром:

$$Hx^t = H(c+e)^t = 0 + He^t = He^t$$

Иначе говоря, принятый вектор и соответствующий ему вектор ошибки имеют одинаковые синдромы.

Вместо всей стандартной таблицы нам достаточно составить таблицу (или, лучше сказать, словарь) синдромов, куда входят все возможные синдромы $\{s_1 = 0, \dots, s_r\}$ и соответствующие им лидеры $\{e_1 = 0, \dots, e_r\}$, такие, что $He_i^t = s_i$. При декодировании, приняв слово x, сначала вычисляем синдром $s = Hx^t$, затем ищем его в словаре синдромов: $s = s_i$, и, наконец, декодируем x к $\hat{c} = x - e_i$.

Пример. Рассмотрим $[4,2]_3$-код Хэмминга с проверочной матрицей

$$H = \begin{pmatrix} 1 & 1 & 2 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$$

Синдромами здесь являются вектор-столбцы длины 2. Например, вектор $x = (1, 2, 1, 1)$ имеет синдром

$$Hx^t = \begin{pmatrix} 1+2+2+0 \\ 0+2+1+1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Составим словарь синдромов (для удобства записывая синдромы в строку, а не в столбец):

синдром	лидер
00	0000
01	0001
02	0002
10	1000
11	0100
12	0002
20	2000
21	0010
22	0200

Для принятого вектора $x = (1, 2, 1, 1)$ ищем его синдром $(2, 1)$ в словаре и находим соответствующий ему лидер $\hat{e} = (0, 0, 1, 0)$. Предполагая, что это и есть ошибка, декодируем x в кодовое слово

$$\hat{c} = x - \hat{e} = (1, 2, 1, 1) - (0, 0, 1, 0) = (1, 2, 0, 1).$$

Задача. Составьте словарь синдромов для расширенного двоичного $[8, 4]$-кода Хэмминга с следующей проверочной матрицей (являющейся также порождающей матрицей, так как код самодуальный)

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

и декодируйте слова:

$$(0, 0, 1, 0, 0, 1, 1, 0) \quad \text{и} \quad (0, 1, 1, 1, 1, 1, 0, 1).$$