

§ 5. Линейные коды (продолжение)

Проверочная матрица кода.

Другой способ задания линейного подпространства $C \subset F^n$ размерности k состоит в указании $n - k$ линейных уравнений, которым удовлетворяют координаты векторов из C :

$$\sum_{j=1}^n h_{ij}x_j = 0 \quad (i = 1, \dots, n - k)$$

или $Hx^t = 0$ в матричной формулировке. Тогда $C = \{x \mid Hx^t = 0\}$. Матрица H называется *проверочной матрицей* кода C .

Если C — двоичный код, а x — кодовый вектор $C^\perp$, то $\langle c, x \rangle = 0 \quad \forall c \in C$; таким образом, x можно рассматривать как вектор, указывающий позиции для проверки на четность. Совершенно аналогично, в случае недвоичных кодов, любой вектор дуального кода можно рассматривать как вектор коэффициентов проверочного уравнения. Строки проверочной матрицы образуют базис проверочных уравнений кода.

Теорема. Порождающая матрица кода $C^\perp$ является проверочной матрицей кода C , и наоборот.

Доказательство. Пусть H — порождающая матрица кода $C^\perp$. Тогда $C^\perp = \{aH \mid a \in F^k\}$. В силу равенства предыдущей леммы $C^{\perp\perp} = C$ векторы кода C характеризуются равенством $\langle x, c \rangle = 0 \quad \forall c \in C^\perp$, равносильному равенству $\langle x, aH \rangle = 0 \quad \forall a \in F^k$. Так как $\langle x, c \rangle = xc^t$, последнее можно записать в виде $xH^t a^t = 0$. Будучи справедливым для всякого a , это равенство означает, что $xH^t = 0$, или $Hx^t = 0$.

Обратное утверждение вытекает из доказанного, если учесть лемму.

Любая матрица H , для которой $C = \{x \mid Hx^t = 0\}$, называется *контрольной* матрицей кода C . Проверочная матрица является частным случаем контрольной матрицы, если потребовать, чтобы ее строки были линейно независимыми.

Примеры. 1. В §1, п.5 был рассмотрен $[4, 2]_3$ -код Хэмминга, состоящий из 9 четверок $(a, b, c, d) \in \{0, 1, 2\}^4$, удовлетворяющих уравнениям $a+b = c$ и $b+c+d = 0$. Таким образом, данный код определяется посредством проверочной матрицы

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

Здесь первая строка матрицы дает уравнение

$$\langle (a, b, c, d), (1, 1, 2, 0) \rangle = a + b + 2c = 0,$$

т.е. $a + b = c$.

2. $[7, 4]_2$ -код Хэмминга из §1, п.3 можно определить посредством проверочной матрицы

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

Здесь каждый ненулевой двоичный 3-вектор встречается в точности один раз как столбец проверочной матрицы. Столбцы упорядочены таким образом, что i -й является двоичным представлением числа i .

Найдем минимальный вес $[7, 4]_2$ -кода Хэмминга. Если x — вектор веса 1, то произведение Hx^t представляет собой i -й столбец матрицы H , где i — позиция единственной единицы вектора x . Так как все столбцы матрицы H ненулевые, Hx^t не равно нулю, поэтому x не может быть кодовым словом. Если x — вектор веса 2, то Hx^t — сумма двух столбцов матрицы H , номера которых соответствуют позициям единиц вектора x . Так как в H никакие два столбца не равны, их сумма не может быть 0, поэтому и в этом случае x не может быть кодовым словом. С другой стороны, в H существуют три столбца, сумма которых равна 0 (например, первые три), поэтому код содержит слова веса 3 (например, $(1, 1, 1, 0, 0, 0, 0)$). Следовательно, минимальный вес рассматриваемого кода равен 3.

Лемма. Пусть C — линейный код с контрольной матрицей H . Множество столбцов матрицы H линейно независимо если, и только если в C существует ненулевое кодовое слово, все ненулевые символы которого стоят в позициях, соответствующих номерам данных столбцов. В частности, $d_{\min}(C) = d$ если, и только если в матрице H найдутся d линейно зависимых столбцов, но никакие $d - 1$ столбцов линейно не зависимы.

Доказательство. Заметим, что $Hx^t = \sum_{i=1}^n \mathbf{h}_i x_i$, где $\mathbf{h}_i$ — i -й столбец матрицы H . Поэтому символы x_i кодового слова x дают коэффициенты линейной зависимости столбцов матрицы H , причем только тех столбцов $\mathbf{h}_i$, для которых соответствующие коэффициенты x_i не равны 0.

Для доказательства второго утверждения вспомним, что $d_{\min}(C) = w_{\min}(C)$. Но $w_{\min}(C) \leq d$ имеет место лишь при условии, что матрица H

содержит d линейно зависимых столбцов, а $w_{\min}(C) \geq d$ — при условии, что все наборы из $d - 1$ столбцов линейно независимы.

Граница Синглтона.

Теорема. (Граница Синглтона). Если C — линейный $[n, k]$ -код, то

$$d_{\min}(C) \leq n - k + 1.$$

Доказательство. Каждый минор размера $n - k \times n - k + 1$ проверочной матрицы имеет ранг не выше $n - k$, поэтому $n - k + 1$ столбцов проверочной матрицы линейно зависимы. Утверждение теоремы теперь следует из предыдущей леммы.

Код, для которого достигается граница Синглтона, т.е. $d = n - k + 1$, называется *кодом с максимальным расстоянием*. В таком коде всякое подмножество из $n - k + 1$ координатных позиций определяет кодовое слово.

Пример. $[4, 2]_3$ -код Хэмминга, как мы знаем, имеет минимальное расстояние 3 и, так как $3 = 4 - 2 + 1$, является кодом с максимальным расстоянием.

Систематические коды.

Теорема. Если порождающая матрица G кода C над полем F имеет вид

$$G = [I_{k \times k} | A_{k \times n-k}],$$

то проверочная матрица имеет вид:

$$H = [-A_{n-k \times k}^t | I_{n-k \times n-k}].$$

Доказательство.

Шаг 1. $xG = y \implies Hy^t = 0$.

Шаг 2. $Hy^t = 0 \implies \exists x \in F^n \ xG = y$.

Систематическим кодом называется код, порождающая матрица которого имеет вид, указанный в теореме, т.е. ее первые k столбцов образуют единичную $k \times k$ -матрицу. Иногда термин *систематическая порождающая матрица* используется для обозначения матрицы, среди столбцов которой найдутся образующие единичную $k \times k$ -матрицу, а матрицу вида $[I_{k \times k} | A_{k \times n-k}]$ называют *стандартной порождающей матрицей*. Подмножество координатных позиций линейного кода называется *информационным множеством*, если у этого кода

имеется порождающая матрица, систематическая по столбцам с номерами, соответствующими данным позициям. Позиции информационного множества содержат информационные символы, а остальные позиции занимают избыточные, или проверочные символы.

Примеры. Рассмотрим порождающие матрицы из примеров предыдущего параграфа, п. «Порождающая матрица кода».

1. Порождающая матрица кода повторения является тривиально стандартной, причем любая координатная позиция может служить информационным множеством.

2. Порождающая матрица кода проверки на четность также стандартная. Информационное множество образуют все координатные позиции, кроме последней.

3. Порождающие матрицы для $[7, 4]_2$ -кода Хэмминга и расширенного $[8, 4]_2$ -кода Хэмминга являются систематическими с информационным множеством $\{3, 5, 6, 7\}$. Хотя из определения это не очевидно, но множество позиций $\{1, 2, 3, 4\}$ также является информационным, что следует из возможности указать для данного кода стандартную порождающую матрицу

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

Не всякое множество из 4 позиций будет информационным. Проверочные уравнения гарантируют, что множество $\{4, 5, 6, 7\}$ — не информационное, так как каждое кодовое слово имеет четное число единиц в этих позициях.

4. Порождающая матрица троичного $[4, 2]_3$ -кода Хэмминга является стандартной. Любая пара позиций образует информационное множество.

Так как строки порождающей матрицы образуют базис линейного кода, очевидно, что можно выбрать базис таким образом, чтобы порождающая матрица оказалась систематической. Перестановка столбцов порождающей матрицы приводит к перестановке следования символов в кодовых словах. Два кода C и D называются *перестановочно эквивалентными*, если они совпадают с точности до перестановки их координатных позиций, т.е. если существует пере-

становка σ множества $\{1, \dots, n\}$, для которой

$$(x_1, x_2, \dots, x_n) \in C \iff (x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(n)}) \in D.$$

Задачи. 1. Доказать, что всякий линейный код перестановочно эквивалентен систематическому.

2. Доказать, что линейный $[n, k]$ -код C является кодом с максимальным расстоянием тогда, и только тогда, когда любое подмножество из k координатных позиций является информационным множеством.