

Теория кодирования

МФТИ, осень 2013

Александр Дайняк

www.dainiak.com

Линейные коды: определение

Пусть q — степень простого числа и символы кодовых слов — элементы конечного поля \mathbb{F}_q .

$(n, M, d)_q$ -код C называется *линейным*, если он является линейным подпространством пространства \mathbb{F}_q^n , то есть линейная комбинация кодовых слов также является кодовым словом.

Если $\dim C = k$, то говорят, что задан

линейный $[n, k, d]_q$ -код

Линейные коды:

пример кода

Пример линейного двоичного кода — код с проверкой чётности:

$$C := \{(a_1, \dots, a_n) \in \mathbb{F}_2^n \mid \sum a_i = 0\}$$

Один из базисов этого кода:

$$(1, 0, 0, \dots, 0, 1)$$

$$(0, 1, 0, \dots, 0, 1)$$

$$(0, 0, 1, \dots, 0, 1)$$

$$\vdots$$

$$(0, 0, 0, \dots, 1, 1)$$

Линейные коды: порождающая матрица

Если выписать базис линейного $[n, k, d]_q$ -кода построчно в виде матрицы размера $k \times n$, получим *порождающую матрицу* кода.

Для задания $[n, k, d]_q$ -кода достаточно указать его порождающую матрицу $G \in \mathbb{F}_q^{k \times n}$.

Число линейных комбинаций k базисных векторов с коэффициентами из \mathbb{F}_q равно q^k , поэтому каждый $[n, k, d]_q$ -код является $(n, q^k, d)_q$ -кодом

Линейные коды: кодирование

Если линейный $[n, k, d]_q$ -кода задан порождающей матрицей G , а исходное сообщение представлено как вектор $\mathbf{x} \in \mathbb{F}_q^k$, то закодировать его можно быстро и просто:

$$\mathbf{x} \xrightarrow{\text{кодирование}} \mathbf{x}^T G \in \mathbb{F}_q^n$$

Обратно, если закодированное сообщение $\mathbf{a} \in \mathbb{F}_q^n$ было принято *без ошибок*, декодируем его, решая, например, методом Гаусса, систему:

$$\mathbf{x}^T G = \mathbf{a}$$

Правда, коды нам как раз были нужны, чтобы уметь декодировать сообщения *с ошибками*...

Линейные коды: канонический вид

Порождающую матрицу $G \in \mathbb{F}_q^{k \times n}$ линейными преобразованиями и перестановками строк и столбцов можно привести к *каноническому виду*:

$$\left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & \\ 0 & 1 & \cdots & 0 & \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \cdots & 1 & \end{array} \tilde{G} \right)$$

где $\tilde{G} \in \mathbb{F}_q^{k \times (n-k)}$.

Тогда кодирование будет *систематическим*:

$$(x_1, \dots, x_k) \xrightarrow{\text{кодирование}} (x_1, \dots, x_k \mid x\tilde{G})$$

Линейные коды: канонический вид

Если порождающая матрица кода задана в каноническом виде, то кодирование будет систематическим и слово (x_1, \dots, x_k) переходит в

$$(x_1, \dots, x_k \mid x\tilde{G}) = (x_1, \dots, x_k, \tilde{x}_1, \dots, \tilde{x}_{n-k})$$

- Разряды x_1, \dots, x_k — *информационные*
- Разряды $\tilde{x}_1, \dots, \tilde{x}_{n-k}$ — *проверочные*

Линейные коды: ЭКВИВАЛЕНТНЫЕ КОДЫ

Если матрица G исходно задана не в каноническом виде, а мы приводим её к каноническому виду, то получается в общем случае не тот же код, а *эквивалентный* ему.

Коды C_1 и C_2 *эквивалентны*, если существует перестановка π и константы $r_1, \dots, r_n \in \mathbb{F}_q \setminus \{0\}$, такие, что

$$(a_1, \dots, a_n) \in C_1 \Leftrightarrow (r_1 a_{\pi(1)}, \dots, r_n a_{\pi(n)}) \in C_2$$

Для эквивалентных кодов $d(C_1) = d(C_2)$.

Кодовое расстояние линейного кода

Для слова $\mathbf{a} \in \mathbb{F}_q^n$ будем через $\|\mathbf{a}\|$ обозначать *вес слова*, т.е. величину

$$\#\{i \mid a_i \neq 0\}$$

Утверждение.

Для любого линейного кода C имеем

$$d(C) = \min_{\substack{\mathbf{a} \in C \\ \mathbf{a} \neq \mathbf{0}}} \|\mathbf{a}\|$$

Кодовое расстояние линейного кода

Доказательство:

Поскольку $\mathbf{0} \in C$, то

$$d(C) \stackrel{\text{def}}{=} \min_{\substack{a, b \in C \\ a \neq b}} d(a, b) \leq \min_{\substack{a \in C \\ a \neq 0}} d(a, \mathbf{0}) = \min_{\substack{a \in C \\ a \neq 0}} \|a\|$$

В обратную сторону. Пусть кодовое расстояние достигается на паре слов a^*, b^* .

Тогда так как $(a^* - b^*) \in C$, получаем

$$d(C) = d(a^*, b^*) = \|a^* - b^*\| \geq \min_{\substack{a \in C \\ a \neq 0}} \|a\|$$

Проверочная матрица

Проверочная матрица H линейного кода — это матрица однородной системы линейных уравнений, которым удовлетворяет код.

Для любого кодового слова \mathbf{a} , по определению матрицы H , выполнено равенство $H\mathbf{a}^T = \mathbf{0}$.

Для $[n, k, d]_q$ -кода $H \in \mathbb{F}_q^{(n-k) \times n}$.

Пример: для кода с проверкой чётности
$$H = (11 \dots 1)$$

Проверочная матрица

Утверждение.

Если G и H — порождающая и проверочная матрицы линейного $[n, k, d]_q$ -кода, то

$$HG^T = \mathbf{0}^{(n-k) \times k}$$

Доказательство:

Достаточно заметить, что каждая строка матрицы G — это кодовое слово, а значит, она удовлетворяет системе, задаваемой H .

Двойственные коды

Если коды C_1 и C_2 таковы, что проверочная матрица C_1 является порождающей матрицей для C_2 , то эти коды называют *двойственными*.

Коды C_1 и C_2 двойственны т. и т.т., когда $\langle \mathbf{a}_1, \mathbf{a}_2 \rangle = 0$ для любых слов $\mathbf{a}_1 \in C_1, \mathbf{a}_2 \in C_2$.

Поэтому двойственные коды называются также *ортogonalными*.

Проверочная матрица и кодовое расстояние

Утверждение. (О связи кодового расстояния и проверочной матрицы)

Линейный код, определяемый проверочной матрицей H , имеет расстояние d т. и т.т., когда

- любые $(d - 1)$ столбцов H линейно независимы,
- найдутся d линейно зависимых столбцов в H .

Проверочная матрица и кодовое расстояние

Доказательство:

Пусть $\mathbf{a} = (a_1, \dots, a_n) \neq \mathbf{0}$ — произвольное кодовое слово кода с проверочной матрицей H .

Пусть H_1, \dots, H_n — столбцы H . Имеем

$$H\mathbf{a}^T = a_1 \cdot H_1 + \dots + a_n \cdot H_n$$

Если i_1, \dots, i_t — все ненулевые координаты \mathbf{a} , то

$$a_1 H_1 + \dots + a_n H_n = a_{i_1} H_{i_1} + \dots + a_{i_t} H_{i_t}$$

Так как \mathbf{a} — кодовое слово, то $H\mathbf{a}^T = \mathbf{0}$, то есть

$$a_{i_1} H_{i_1} + \dots + a_{i_t} H_{i_t} = \mathbf{0}$$

Проверочная матрица и кодовое расстояние

Продолжение доказательства:

Мы получили, что в коде есть слово веса не более t т. и т.т., когда некоторые t столбцов матрицы H линейно зависимы.

Осталось воспользоваться формулой, справедливой для любых линейных кодов:

$$d(C) = \min_{\substack{a \in C \\ a \neq 0}} \|a\|$$

Декодирование по синдрому

Пусть H — проверочная матрица кода C , и пусть $\mathbf{a} \in C$.

Если при передаче \mathbf{a} по каналу произошло t ошибок, на выходе из канала имеем слово \mathbf{b} .

Вектор $\mathbf{e} := \mathbf{b} - \mathbf{a}$ называется *вектором ошибок*.

Очевидно, $\|\mathbf{e}\| = t$.

Исправить ошибки в \mathbf{b} — то же самое, что найти вектор \mathbf{e} .

Используем равенство

$$H\mathbf{b}^T = H\mathbf{a}^T + H\mathbf{e}^T = H\mathbf{e}^T$$

Декодирование по синдрому

Получаем задачу: найти вектор \mathbf{e} , такой, что

$$\begin{cases} \|\mathbf{e}\| < \frac{d(C)}{2} \\ H\mathbf{e}^T = H\mathbf{b}^T \end{cases}$$

Утверждение.

Если решение этой системы существует, то оно единственное.

Доказательство.

Пусть нашлись разные решения $\mathbf{e}_1 \neq \mathbf{e}_2$.

Тогда $\|\mathbf{e}_1 - \mathbf{e}_2\| \leq \|\mathbf{e}_1\| + \|\mathbf{e}_2\| < d(C)$

и $H(\mathbf{e}_1 - \mathbf{e}_2)^T = \mathbf{0}$ — противоречие.

Декодирование по синдрому

Можно составить таблицу решений системы

$$\begin{cases} \|\mathbf{e}\| < \frac{d(C)}{2} \\ H\mathbf{e}^T = \mathbf{s} \end{cases}$$

для всевозможных \mathbf{s} .

При получении из канала слова \mathbf{b} мы вычисляем $\mathbf{s} := H\mathbf{b}^T$ (синдром слова \mathbf{b}), и для этого \mathbf{s} смотрим в таблице соответствующий вектор \mathbf{e} .

Декодированное сообщение — решение системы

$$\mathbf{x}^T G = (\mathbf{b} - \mathbf{e})$$

Граница Варшамова—Гилберта

Теорема. (Р.Р. Варшамов, E.N. Gilbert)

Пусть числа $n, k, d' \in \mathbb{N}$ таковы, что

$$\sum_{j=0}^{d'-1} \binom{n-1}{j} < 2^{n-k}$$

Тогда существует $[n, k, d]$ -код, где $d > d'$.

Граница Варшамова—Гилберта

Доказательство:

Покажем, что в условиях теоремы можно построить матрицу $H \in \mathbb{F}_2^{(n-k) \times n}$, у которой любые d' столбцов линейно независимы.

Будем строить матрицу по столбцам.

Пусть уже выбраны t столбцов (где $t < n$) и требуется выбрать $(t + 1)$ -й столбец.

Этот новый столбец не должен образовывать нулевую линейную комбинацию с $(d' - 1)$ или менее из уже выбранных столбцов.

Граница Варшамова—Гилберта

Выбираемый столбец не должен образовывать нулевую линейную комбинацию с $(d' - 1)$ или менее из уже выбранных столбцов.

Так как у нас \mathbb{F}_2 , то это равносильно тому, что выбираемый столбец не равен сумме $(d' - 1)$ или менее уже выбранных столбцов.

Количество таких сумм равно

$$\sum_{j=0}^{d'-1} \binom{m}{j} \leq \sum_{j=0}^{d'-1} \binom{n-1}{j} < 2^{n-k}$$

Граница Варшамова—Гилберта

Итак, запрещённых для выбора столбцов у нас оказывается строго меньше 2^{n-k} , а всего векторов длины $(n - k)$ ровно 2^{n-k} .
Значит, найдётся вектор из \mathbb{F}_2^{n-k} , который можно добавить в качестве очередного столбца.

Двоичный код Хемминга (R.W. Hamming)

Двоичный код с проверкой чётности

$$\{(a_1, \dots, a_n) \in \mathbb{F}_2^n \mid \sum a_i = 0\}$$

может обнаруживать одну ошибку, т.к. если один разряд a_i заменить на противоположный, соотношение $\sum a_i = 0$ нарушится. Но исправить ошибку не удастся.

Хочется построить двоичный код, *исправляющий* хотя бы одну ошибку.

Для этого вместо «глобального» контроля чётности применим несколько «дихотомических» проверок на чётность...

Двоичный код Хемминга

Пример для $n = 7$. Рассмотрим код, удовлетворяющий соотношениям

$$a_4 + a_5 + a_6 + a_7 = 0$$

$$a_2 + a_3 + a_6 + a_7 = 0$$

$$a_1 + a_3 + a_5 + a_7 = 0$$

Проверочная матрица этого кода:

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

Двоичный код Хемминга

Проверочная матрица:

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

Столбцы матрицы — всевозможные ненулевые вектора высоты 3.
 j -й столбец суть двоичная запись числа j .

Любая пара столбцов л.н.з., значит $d(C) \geq 3$, значит, этот код исправляет одну ошибку и обнаруживает две.

Двоичный код Хемминга

Если ошибка случается в a_j , то можно вычислить левые части проверочных соотношений, и они дадут двоичную запись j , например:

$$\begin{array}{rcccccccl} & & & a_4 & + & \overline{a_5} & + & a_6 & + & a_7 & = & 1 \\ & & a_2 & + & a_3 & & & + & a_6 & + & a_7 & = & 0 \\ a_1 & & & + & a_3 & & + & \overline{a_5} & & + & a_7 & = & 1 \end{array}$$

Двоичный код Хемминга

Общий случай: $n := 2^m - 1$ для некоторого m .

Код Хемминга длины n определяется проверочной матрицей $H \in \mathbb{F}_2^{m \times n}$, столбцы которой — всевозможные двоичные вектора высоты m :

$$\begin{pmatrix} 1 & 0 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}$$

Код Хемминга имеет параметры

$$[2^m - 1, 2^m - m - 1, 3]$$

Двоичный код Хемминга

Граница Хемминга:

Для любого $(n, M, d)_q$ -кода имеем

$$M \leq \frac{q^n}{|S_{\lfloor (d-1)/2 \rfloor}(\mathbf{0})|}$$

В двоичном случае

$$M \leq \frac{2^n}{\sum_{k=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{k}}$$

Коды, достигающие эту границу, называются *совершенными* или *плотно упакованными*.

Двоичный код Хемминга

Граница Хемминга:

Для любого (n, M, d) -кода имеем

$$M \leq \frac{2^n}{\sum_{k=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{k}}$$

Утверждение.

Код Хемминга является совершенным.

Доказательство:

Для кода Хемминга имеем $n = 2^m - 1$, $M = 2^{2^m - m - 1}$ и $d = 3$.

Отсюда $\sum_{k=0}^{\lfloor (d-1)/2 \rfloor} \binom{n}{k} = n + 1 = 2^m = 2^n / M$.

Граница Синглтона для линейных кодов

Утверждение.

Для любого $[n, k, d]_q$ -кода имеем

$$k \leq n - d + 1$$

Доказательство:

По теореме Синглтона, для любого $(n, M, d)_q$ -кода выполнено $M \leq q^{n-d+1}$.

С другой стороны, для линейного кода $M = q^k$.

Остаточный код

Теорема. (G. Solomon, J.J. Stiffler)

Если существует $[n, k, d]_q$ -код, то существует и $[n - d, k - 1, d']_q$ -код, где $d' \geq d/q$.

Доказательство:

Пусть G — порождающая матрица некоторого $[n, k, d]_q$ -кода C .

Б.о.о. будем считать, что первая строка G содержит ровно d ненулевых элементов и имеет вид $(r_1 \dots r_d 0 \dots 0)$.

Остаточный код

Порождающая матрица кода C :

$$G = \begin{pmatrix} r_1 & \dots & r_d & 0 & \dots & 0 \\ & & \dots & G' & & \end{pmatrix}$$

Имеем $G' \in \mathbb{F}_q^{(k-1) \times (n-d)}$.

Покажем, что $\text{rk } G' = k - 1$. Допустим противное: некоторая нетривиальная линейная комбинация строк G' равняется $\mathbf{0}$.

Тогда линейная комбинация соответствующих строк G равна $(t_1 \dots t_d 0 \dots 0)$, где $\forall i (t_i \neq 0)$.

Остаточный код

$$G = \begin{pmatrix} r_1 \dots r_d & 0 \dots 0 \\ \dots & G' \end{pmatrix}$$

Линейная комбинация U некоторых строк G равна $(t_1 \dots t_d 0 \dots 0)$, где $\forall i (t_i \neq 0)$.

Т.к. \mathbb{F}_q поле, то $\exists s \in \mathbb{F}_q$, такой, что $st_d = -r_d$. Тогда

$$s \cdot U + (r_1 \dots r_d 0 \dots 0)$$

— линейная комбинация строк G , равная

$$((st_1 + r_1) \dots (st_{d-1} + r_{d-1}) 00 \dots 0)$$

Это противоречит условию $d(C) = d$.

Остаточный код

$$G = \begin{pmatrix} r_1 & \dots & r_d & 0 & \dots & 0 \\ & & & G' & & \end{pmatrix}$$

Итак, $G' \in \mathbb{F}_q^{(k-1) \times (n-d)}$ и $\text{rk } G' = k - 1$.

Значит G' является порождающей матрицей некоторого $[n - d, k - 1, d']_q$ -кода C' (этот код называется *остаточным* для исходного кода C).

Рассмотрим любой ненулевой вектор кода C'
 $\mathbf{a}' := (a'_1, \dots, a'_{n-d}) \neq \mathbf{0}$

такой, что $\|\mathbf{a}'\| = d'$.

В коде C есть вектор вида

$$(a_1, \dots, a_d, a'_1, \dots, a'_{n-d})$$

Остаточный код

$$G = \begin{pmatrix} r_1 & \dots & r_d & 0 & \dots & 0 \\ & & \dots & G' & & \end{pmatrix}$$

В коде C есть вектор вида

$$(a_1, \dots, a_d, a'_1, \dots, a'_{n-d})$$

Пусть f_1, \dots, f_q — все элементы поля \mathbb{F}_q .

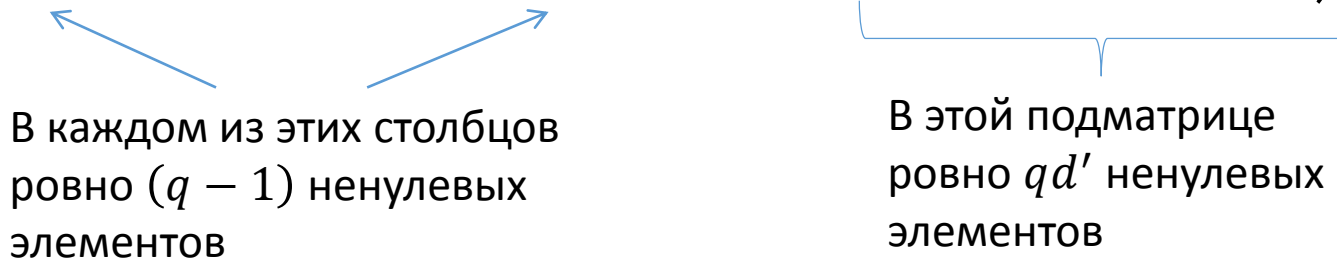
Коду C принадлежат все вектора вида

$$((a_1 - f_i r_1), \dots, (a_d - f_i r_d), a'_1, \dots, a'_{n-d})$$

Запишем эти q векторов построчно в виде матрицы и оценим количество ненулевых элементов в ней.

Остаточный код

Рассмотрим матрицу:

$$\begin{pmatrix} (a_1 - f_1 r_1) & \dots & (a_d - f_1 r_d) & a'_1 & \dots & a'_{n-d} \\ (a_1 - f_2 r_1) & \dots & (a_d - f_2 r_d) & a'_1 & \dots & a'_{n-d} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (a_1 - f_q r_1) & \dots & (a_d - f_q r_d) & a'_1 & \dots & a'_{n-d} \end{pmatrix}$$


В каждом из этих столбцов
ровно $(q - 1)$ ненулевых
элементов

В этой подматрице
ровно qd' ненулевых
элементов

Итого в матрице $d(q - 1) + qd'$ ненулевых элементов.

Остаточный код

В рассмотренной матрице $d(q - 1) + qd'$ ненулевых элементов.

Каждая строка матрицы — ненулевой вектор кода C , значит в матрице не менее чем qd ненулевых элементов.

Отсюда

$$d(q - 1) + qd' \geq qd$$

и следовательно

$$d' \geq d/q$$

Граница Грайсмера—Соломона—Штиффлера

Теорема. (J.H. Griesmer, G. Solomon, J.J. Stiffler)

Для любого $[n, k, d]_q$ -кода имеем

$$n \geq \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil$$

Доказательство: индукция по k .

Утверждение очевидно при $k = 1$. Предположим, что оно выполнено для кодов с размерностью $\leq k - 1$ и докажем его для $[n, k, d]_q$ -кодов.

Граница Грайсмера—Соломона—Штиффлера

Обозначим через $N(k, d)$ минимальную длину слов u кода с размерностью k и расстоянием d .

Пусть C — какой-нибудь $[N(k, d), k, d]_q$ -код.

Остаточный для C код имеет параметры

$$[N(k, d) - d, k - 1, d']_q$$

и по предположению индукции, для него справедливо неравенство

$$N(k, d) - d \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^i} \right\rceil$$

Граница Грайсмера—Соломона—Штиффлера

Из неравенства

$$N(k, d) - d \geq \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^i} \right\rceil$$

с учётом того, что $d' \geq d/q$, получаем

$$N(k, d) \geq d + \sum_{i=0}^{k-2} \left\lceil \frac{d'}{q^i} \right\rceil \geq d + \sum_{i=0}^{k-2} \left\lceil \frac{d}{q^{i+1}} \right\rceil = \sum_{i=0}^{k-1} \left\lceil \frac{d}{q^i} \right\rceil$$