НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Модуль 3. Циклические коды

Коды МДР. Коды Рида-Соломона. Алгоритм декодирования РС кодов на базе расширенного алгоритма Евклида.

Иванов Ф. И.
к.ф.-м.н., доцент

Национальный исследовательский университет
«Высшая школа экономики»

12 июня 2020 г.

Коды МДР

Определение

Код МДР (код с максимально достижимым расстоянием) — это линейный код, для минимального расстояния которого выполняется соотношение $d = n - k + 1$, т. е. код лежит на границе Синглтона.

Информационные совокупности

Определение

Информационной совокупностью линейного (n, k)-кода над $GF(q)$ называется множество номеров компонент кодового вектора, в которых все q^k кодовых векторов различны.

Пример

Рассмотрим порождающую матрицу кода Хэмминга:

$$G = \left[\begin{array}{cccc:ccc} 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \end{array}\right].$$

Легко видеть, что номера 1, 2, 3, 4 составляют информационную совокупность, так как эти части строк матрицы линейно независимы, и все 16 векторов, порождаемых данной матрицей, в первых четырех компонентах различны.

Общие свойства кодов МДР

Теорема

Линейный код лежит на границе Синглтона тогда и только тогда, когда любая совокупность k номеров его компонент является информационной.

Доказательство

Необходимость. Пусть $d = n - k + 1$, но некоторая совокупность k номеров не является информационной. Тогда найдутся два кодовых вектора **u** *и* **v**, *которые в этих k компонентах совпадают. Значит, $d(\mathbf{u}, \mathbf{v}) \leq n - k$.*
Достаточность. Пусть любая совокупность k номеров является информационной, но $d < n - k + 1$. Тогда **u**,**v**, *что $d(\mathbf{u},\mathbf{v}) < n - k + 1$. В таком случае векторы* **u**,**v** *совпадают в $n - d > n - (n - k + 1) = k - 1$, т.е. по крайней мере, в k компонентах, совокупность номеров которых, таким образом, оказывается не информационной, вопреки условию.*

Двойственный код к МДР

Теорема

Код, двойственный коду МДР, есть также код МДР.

Доказательство

Любые k столбцов порождающей матрицы (n,k)-кода МДР линейно независимы, так как они образуют минор, строкам которого принадлежат компоненты информационной совокупности. Значит, расстояние двойственного кода есть $d' \geq k + 1 = n - k' + 1$. С другой стороны всегда $d' n - k' + 1$. Остается знак равенства: $d' = n - k' + 1$.

Свойства порождающей матрицы кода МДР

Теорема

Матрица $\mathbf{G} = \left(\mathbf{I}_{k\times k}, \mathbf{P}_{k\times(n-k)}\right)$ *порождает код МДР, если и только если любой минор матрицы* $\mathbf{P}_{k\times(n-k)}$ *отличен от нуля.*

Доказательство

Необходимость. Пусть минор L *порядка* $l(1 \leq l \leq k)$ *расположен на пересечении строк матрицы* $\mathbf{G}$ *с номерами* $r_1, r_2, ..., r_l$ *и столбцов матрицы* $\mathbf{P}_{k\times(n-k)}$ *с номерами* $c_1, c_2, ..., c_l$*. Если минор* $L = 0$*, то это значит, что вектор кода, равный некоторой линейной комбинации строк с номерами* $r_1, r_2, ..., r_l$*, имеет в точности* l *нулевых компонент с номерами* $c_1, c_2, ..., c_l$*. И этот же самый вектор имеет* $l' \leq l$ *отличных от нуля компонент на отрезке матрицы* $\mathbf{I}_{k\times k}$*. Таким образом, нашему коду МДР принадлежит вектор веса* $w = n - k - l + l' \leq n - k$*, что противоречит условию* $w \geq d = n - k + 1$*.*

Достаточность

На отрезке кодового вектора, отвечающем матрице $\mathbf{P}_{k\times(n-k)}$, любая совокупность k номеров компонент является информационной, так как произвольный минор матрицы $\mathbf{P}_{k\times(n-k)}$ отличен от нуля.
Совокупность k номеров компонент на отрезке, отвечающем матрице $\mathbf{I}_{k\times k}$, является информационной по определению. Если же совокупность k номеров компонент состоит из k_1 номеров на отрезке, отвечающем матрице $\mathbf{I}_{k\times k}$, и k_2, $(k_1 + k_2 = k)$ номеров на отрезке, отвечающем матрице $\mathbf{P}_{k\times(n-k)}$, то произведение отличного от нуля минора порядка k_2 на отличное от нуля его алгебраическое дополнение порядка k_1 само будет отлично от нуля, так как это алгебраическое дополнение в каждой строке и каждом столбце содержит в точности по одной единице. Таким образом, обсуждаемая совокупность также информационная.

Следствия

1. Все элементы матрицы $\mathbf{P}_{k\times(n-k)}$ отличны от нуля
2. Двоичных нетривиальных кодов МДР всего 2 и больше быть не может!
3. Это $(n{,}1{,}n)$ код с повторением и $(n{,}n-1{,}2)$ код с проверкой на четность
4. Все остальные коды МДР - недвоичные!

Определение

Кодом Рида-Соломона (РС) называется код БЧХ, если корни $\alpha^b, \alpha^{b+1}, ..., \alpha^{b+d-2}$ *его порождающего многочлена* $g(x) = g_0 + g_1 x + ... + g_r x^r$ *принадлежат тому же полю* $GF(q)$, *что и коэффициенты.*

С другой стороны $g(x)$ - это наименьшее общее кратное минимальных функций корней. Так как $m = 1$, то минимальная функция элемента α имеет вид $m_\alpha(x) = x - \alpha$, а потому $g(x)$ у РС-кода имеет вид:

$$g(x) = (x - \alpha^b)(x - \alpha^{b+1})...(x - \alpha^{b+d-2})$$

Основное свойство РС-кода

Теорема

Код Рида—Соломона есть код МДР.

Доказательство

$$g(x) = (x - \alpha^b)(x - \alpha^{b+1})...(x - \alpha^{b+d-2})$$

Степень $g(x)$ *равна* $d - 1$, *а по определению* $n - k$, *откуда:*

$$n - k = d - 1$$

Пример

Пусть $q = 5$, $n = q - 1 = 4$. Мультипликативная группа поля $GF(5)$ — это приведенная система вычетов по модулю 5. Положим $\alpha = 2, \alpha^2 = 4$ — корни порождающего многочлена. Тогда $g(x) = (x-2)(x-4) = 3 + 4x + x^2$. Порождающая матрица кода РС будет:

$$\mathbf{G} = \begin{pmatrix} 3 & 4 & 1 & 0 \\ 0 & 3 & 4 & 1 \end{pmatrix}$$

И ее каноническая форма:

$$\mathbf{G}_K = \begin{pmatrix} 1 & 0 & 3 & 4 \\ 0 & 1 & 3 & 2 \end{pmatrix}$$

Кодирование РС-кода

Теорема

Положим $a_i \in GF(q), (i = 0, 1..., k-1), \alpha \in GF(q)$, *и пусть* $\mathbf{a} = (a_0, a_1, ..., a_{k-1})$ *вектор информационных символов, а значит,* $a(x) = a_0 + a_1 x + ... + a_{k-1} x^{k-1}$ *— информационный многочлен. Тогда вектором кода РС будет*

$$\mathbf{u} = (a(1), a(\alpha), ..., a(\alpha^{q-2}))$$

Кодировать кодом РС можно без порождающего многочлена/ /порождающей матрицы! Все параметры и свойства кода РС определяются полностью двумя числами: $n = q - 1$ - длиной кода и $k < n$ - числом информационных символов!

Доказательство - начало

Рассмотрим многочлен

$$u(x) = a(1) + a(\alpha)x + ... + a(\alpha^{q-2})x^{q-2} =$$
$$= (a_0 + a_1 ... + a_{k-1}) +$$
$$+(a_0 + a_1\alpha ... + a_{k-1}\alpha^{k-1})x +$$
$$+(a_0 + a_1\alpha^2 ... + a_{k-1}\alpha^{2(k-1)})x^2 +$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$+(a_0 + a_1\alpha^{q-2} ... + a_{k-1}\alpha^{(q-2)(k-1)})x^{q-2} =$$

(изменив порядок суммирования)

$$= a_0(1 + x + x^2 + ... + x^{q-2}) +$$
$$+a_1(1 + +\alpha^2 x^2 + ... + (\alpha x)^{q-2}) +$$
$$+a_2(1 + \alpha^2 x + \alpha^4 x^2 + ... + (\alpha^2 x)^{q-2}) +$$
$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$
$$+a_{k-1}(1 + \alpha^{k-1}x + (\alpha^{k-1}x)^2 + ... + (\alpha^{k-1}x)^{q-2}).$$

Доказательство - продолжение

Найдем отсюда $u(\alpha^{-i})$ для всех $i = 0, 1, ..., k-1$:

$$u(\alpha^{-i}) = a_0(1 + \alpha^{-i} + \alpha^{-2i} + ... + \alpha^{-i(q-2)}) +$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$+a_i(1 + \alpha^{-i}\alpha_i + \alpha^{-2i}\alpha^{2i} + ... + (\alpha^{-i}\alpha^{i}))^{q-2}) +$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$+a_{k-1}(1 + \alpha^{k-1}\alpha^{-i} + (\alpha^{k-1}\alpha^{-i})^2 + ... + (\alpha^{k-1}\alpha^{-i})^{q-2}) = -a_i$$

Действительно, только в одной из скобок все слагаемые обращаются в единицу; слагаемых в скобке ровно $q-1$ штук, величина q есть степень (быть может и первая) характеристики поля. Поэтому $q - 1 = -1$.

Доказательство - окончание

В каждой из остальных скобок содержится сумма членов геометрической прогрессии вида:

$$1 + \alpha^j + \alpha^{2j} + ... + \alpha^{(q-2)j} = \frac{\alpha^{j(q-1)} - 1}{\alpha^j - 1} = 0.$$

В самом деле, так как $\alpha \in GF(q)$, то α^j есть корень двучлена $x^{q-1} - 1$. Поэтому $\alpha^{j(q-1)} - 1 = 0$. Но $\alpha^j - 1 \neq 0$.
Как только $i \geq k$, больше не будет скобки с $q - 1$ единичными слагаемыми. Поэтому при $k \leq i \leq q - 2$ нулевой будет каждая скобка, и, значит, $u(\alpha^{-i}) = 0$.
Но $\alpha^{-i} = \alpha^{q-1-i}$. Следовательно, $u(\alpha^j) = 0$ при $j = 1, 2, ..., q - 1 - k$. Иначе говоря, вектор **u** есть вектор кода БЧХ с минимальным расстоянием $d = q - k = q - 1 - k + 1 = n - k + 1$, а значит, и кода РС.
Теорема доказана.

Поле $GF(3^2)$

Построим поле $GF(3^2)$ по модулю многочлена $x^2 + x + 2$. Если $\alpha^2 + \alpha + 2 = 0$, то таблица поля имеет вид:

$$
\begin{array}{rcrrl}
0 = 0 & & & & = (00), \\
\alpha^0 = & 1 & & & = (10), \\
\alpha^1 = & & & \alpha & = (01), \\
\alpha^2 = & 1 & & +2\alpha & = (12), \\
\alpha^3 = & 2 & & +2\alpha & = (22), \\
\alpha^4 = & 2 & & & = (20), \\
\alpha^5 = & & & 2\alpha & = (02), \\
\alpha^6 = & 2 & & +\alpha & = (21), \\
\alpha^7 = & 1 & & +\alpha & = (11), \\
\alpha^8 = & 1 & & & = (10).
\end{array}
$$

Пример кодирования

Пусть $a(x) = 1 + \alpha x + \alpha^2 x^2 + \alpha^3 x^3$, тогда подставляя ненулевые элементы поля, получим:

$$a(1) = \alpha^2, a(\alpha) = 0, a(\alpha^2) = \alpha^6, a(\alpha^3) = 0,$$

$$a(\alpha^4) = \alpha^5, a(\alpha^5) = 0, a(\alpha^6) = \alpha^7, a(\alpha^7) = 1.$$

Построен вектор кода РС с параметрами $n = 8, k = 4, d = 5$:

$$\mathbf{u} = (\alpha^2, 0, \alpha^6, 0, \alpha^5, 0, \alpha^7, 1).$$

Если же информационный многочлен есть $a(x) = 2 + \alpha x$, то $a(1) = 6, a(\alpha) = \alpha^5, a(\alpha^2) = \alpha^2, a(\alpha^3) = 1, a(\alpha^4) = \alpha^3, a(\alpha^5) = \alpha^7, a(\alpha^6) = \alpha, a(\alpha^7) = 0$, и вектор кода РС с параметрами $n = 8, k = 2, d = 7$ есть

$$\mathbf{u} = (\alpha^6, \alpha^5, \alpha^2, 1, \alpha^3, \alpha^7, \alpha, 0).$$

Декодирование кодов РС - алгоритм Евклида для многочленов

Задача: для данных многочленов $a(z)$, $b(z)$ найти НОД$(a(z), b(z))$

Алгоритм:

$$
\begin{array}{ll}
 & \deg b(z) \le \deg a(z) \\
a(z) = b(z)q_0(z) + r_0(z), & \deg r_0(z) < \deg b(z) \\
b(z) = r_0(z)q_1(z) + r_1(z), & \deg r_1(z) < \deg r_0(z) \\
r_0(z) = r_1(z)q_2(z) + r_2(z), & \deg r_2(z) < \deg r_1(z) \\
\dots\dots\dots\dots & \dots\dots\dots\dots \\
r_k(z) = r_{k+1}(z)q_2(z) + r_{k+2}(z), & \deg r_{k+2}(z) < \deg r_{k+1}(z) \\
\dots\dots\dots\dots & \dots\dots\dots\dots \\
r_{n-2}(z) = r_{n-1}(z)q_n(z) & 0 = r_n(z)
\end{array}
$$

Откуда

$$(a(z), b(z)) = (b(z), r_0(z)) = (r_0(z), r_1(z)) = \dots = (r_{n-1}(z), 0) = r_{n-1}(z).$$

Расширенный алгоритм Евклида

Задача: для данных многочленов $a(z)$, $b(z)$ найти такую пару многочленов $s(z)$, $t(z)$, что

$$a(z)s(z) + b(z)t(z) = \text{НОД}(a(z), b(z))$$

Алгоритм:

Запустим итерационную процедуру:

$$\begin{aligned} u_{-2}(z) &= 0,\ u_{-1}(z) = 1. \\ v_{-2}(z) &= 1,\ v_{-1}(z) = 0, \end{aligned}$$

$$\begin{aligned} u_k(z) &= q_k(z)u_{k-1}(z) + u_{k-2}(z), \\ v_k(z) &= q_k(z)v_{k-1}(z) + v_{k-2}(z). \end{aligned}$$

и

$$\begin{aligned} v_{k-1}(z)r_k(z) + v_k(z)r_{k-1}(z) &= r_{-1}(z), \\ u_{k-1}(z)r_k(z) + u_k(z)r_{k-1}(z) &= r_{-2}(z), \\ v_k(z)u_{k-1}(z) - u_k(z)v_{k-1}(z) &= (-1)^k. \end{aligned}$$

Расширенный алгоритм Евклида

Тогда из последней системы для любого k справедливо:

$$r_k(z) = (-1)^k(-v_k(z)r_{-2}(z) + u_k(z)r_{-1}(z)).$$

Положив $r_{-2}(z) = a(z)$, $r_{-1}(z) = b(z)$ и $k = n - 1$ получим требуемое равенство.

Вывод ключевого уравнения

Пусть задан код РС над $GF(q)$ длины $n = q - 1$ и размерности $k < n$. Обозначения:

- $\mathbf{u} = (u_0, u_1, \ldots u_{n-1})$, $u_i \in GF(q)$ - кодовый вектор, переданный в канал связи,
- $\mathbf{v} = (v_0, v_1, \ldots, v_{n-1})$, $v_i \in GF(q)$ - полученный из канала вектор, в котором могут быть ошибки
- $\mathbf{e} = (e_0, e_1, \ldots, e_{n-1})$, $e_i \in GF(q)$ - вектор-ошибка, такой, что $\mathbf{v} = \mathbf{u} + \mathbf{e}$.

Синдром принятого вектора $\mathbf{S} = (S_1, S_2, \ldots S_{2t})$ имеет вид:

$$S_1 = Y_1 X_1 + Y_2 X_2 + \ldots + Y_t X_t,$$

$$S_2 = Y_1 X_1^2 + Y_2 X_2^2 + \ldots + Y_t X_t^2,$$

$$\ldots$$

$$S_{2t} = Y_1 X_1^{2t} + Y_2 X_2^{2t} + \ldots + Y_t X_t^{2t}.$$

Вывод ключевого уравнения

Сопоставим вектору $\mathbf{S} = (S_1, S_2, \ldots S_{2t})$ многочлен:

$$S(z) = \sum_{j=0}^{2t} S_j z^{j-1}$$

Можно показать, что:

$$S(z) = z^{2t} \sum_{i=1}^{t} Y_i X_i \frac{X_i^{2t}}{X_i z - 1} - \sum_{i=1}^{t} \frac{X_i Y_i}{X_i z - 1}$$

полагая

$$\sigma(z) = \prod_{i=1}^{t} (X_i z - 1),$$

$$\omega(z) = \sum_{i=1}^{t} X_i Y_i \prod_{l=1, l \neq i}^{t} (X_i z - 1), \quad \Phi(z) = \sum_{i=1}^{t} X_i^{2t+1} Y_i \prod_{l=1, l \neq i}^{t} (X_i z - 1)$$

получим ключевое уравнение:

$$S(z) = z^{2t} \frac{\Phi(z)}{\sigma(z)} - \frac{\omega(z)}{\sigma(z)}.$$

Ключевое уравнение

Выражение

$$S(z) = z^{2t}\frac{\Phi(z)}{\sigma(z)} - \frac{\omega(z)}{\sigma(z)}.$$

можно представить как:

$$-\omega(z) \equiv S(z)\sigma(z) \mod z^{2t}$$

- $\omega(z)$ — многочлен значений ошибок
- $\sigma(z)$ — многочлен локаторов ошибок

Решение ключевого уравнения

Из расширенного алгоритма Евклида:

$$r_k(z) = (-1)^k(-v_k(z)r_{-2}(z) + u_k(z)r_{-1}(z)).$$

или

$$r_k(z) \equiv (-1)^k u_k(z)r_{-1}(z) \mod r_{-2}(z)$$

Пусть $r_{-2}(z) = z^{2t}, r_{-1}(z) = S(z)$, тогда найдется такое k, что решение сравнения

$$r_k(z) \equiv (-1)^k u_k(z)S(z) \mod z^{2t}$$

даст $\xi u_k(z) = \sigma(z)$, $(-1)^k \xi r_k(z) = \omega(z)$. k выбирается так, чтобы $r_k(z)$ был многочленом, когда впервые выполнено $deg\ r_k(z) \leq t - 1$. Константа ξ выбирается так, чтобы $\sigma_0 = 1$

Основная теорема

Теорема

При $\sigma(0) = 1$, *выполнении условий* $\deg \sigma(z) \leq t$ *и* $\deg \omega(z) \leq t - 1$ *многочлены* $\omega(z)$, $\sigma(z)$, *получаемые в качестве решения ключевого уравнения, единственны, и многочлен* $\sigma(z)$ *имеет минимальную степень.*

Пример

Рассмотрим код РС над $GF(3^2)$ с корнями $\alpha, \alpha^2, \alpha^3, \alpha^4$ порождающего многочлена. Длина кода $n = 8$, и минимальное расстояние $d = 5$. Код исправляет любые ошибки кратности 1 и 2.
Пусть принят вектор $\mathbf{u} = (0{,}0{,}0{,}0{,}\alpha^5{,}0{,}\alpha^7{,}1)$
Компоненты синдрома: $S_1 = \alpha^7$, $S_2 = \alpha^2$, $S_3 = \alpha$, $S_4 = 0$.
Тогда

$$r_{-1}(z) = S(z) = \alpha z^2 + \alpha^2 z + \alpha^7,\ r_{-2}(z) = z^4.$$

Применяя алгоритм Евклида находим:

$$r_{-2}(z) = z^4,\ r_{-1}(z) = \alpha z^2 + \alpha^2 z + \alpha^7,\ q_0(z) = \alpha^7 z^2 + \alpha^4 z + \alpha^5,\ r_0(z) = \alpha^4$$

Это можно проверить:

$$z^4 = (\alpha z^2 + \alpha^2 z + \alpha^7)(\alpha^7 z^2 + \alpha^4 z + \alpha^5) + \alpha^4$$

При $k = 0$ тривиальным образом впервые выполняется неравенство $deg\ r_k(z) \leq t - 1$, где в нашем случае $t = 2$.
Имеем далее $u_{-2} = 0$; $u_{-1} = 1$, $u_0 = q_0 u_{-1} + u_{-2} = \alpha^7 z^2 + \alpha^4 z + \alpha^5$. Следует положить $\xi = \alpha^3$.
Окончательно получим многочлен локаторов ошибок $\sigma(z) = \alpha^2 z^2 + \alpha^7 z + 1$.

Нахождение значений ошибок. Формула Форни.

Если

$$\sigma(z) = \prod_{i=1}^{t}(X_i z - 1) = \sum_{i=0}^{t} \sigma_i z^i,$$

то

$$\sigma'(z) = \sum_{j=1}^{t} X_j \prod_{i=1, i\neq j} (X_j z - 1) = \sum_{i=1}^{t} i \sigma_i z^{i-1}$$

Подставим в

$$\omega(z) = \sum_{i=1}^{t} X_i Y_i \prod_{l=1, l\neq i}^{t} (X_i z - 1)$$

и $\sigma'(z)$ любой корень многочлена локаторов ошибок X_j^{-1}:

$$\omega(X_j^{-1}) = Y_j^{-1} \sigma'(X_j^{-1}),$$

откуда

$$Y_j = \frac{\omega(X_j^{-1})}{\sigma'(X_j^{-1})}.$$