Теория кодирования.
Конспект практических занятий

Ерофей Башунов

3 сентября 2021

Линейные коды

Допустим, у нас есть некоторое векторное n-мерное пространство $\mathbb{F}_q^n$, где q — количество элементов. Выберем подпространство размерности k ($n > k$), это и будет являться линейным кодом размерности n и ранга k. Параметр q отвечает за арность кода (при $q = 2$ код является бинарным).

Отображение из k бит в n бит — добавление избыточной информации к векторам из подпространства. Это отображение необходимо для того, чтобы повысить устойчивость передачи информации к помехам путём добавления дополнительных $(n - k)$ «проверяющих» битов. Формально, оно записывается в виде $\mathbb{F}_q^k \to \mathbb{F}_q^n$.

Рассмотрим два вектора таких, что $a \in \mathbb{F}_q^k$ и $c \in \mathbb{F}_q^n$. Тогда преобразование можно рассматривать как умножение вектора на матрицу, а именно:

$$c = a \cdot \underset{k \times n}{G}$$

Матрица G называется порождающей матрицей.

Проверочная матрица

Проверочная матрица H^T — способ определить, является ли слово кодовым или нет. Необходимое и достаточное условие: $c \cdot H^T = 0$. Матрица находится с помощью уравнения:

$$\underset{k \times n}{G} \cdot \underset{n \times (n-k)}{H^T} = 0$$

Чтобы найти матрицу H^T, необходимо выполнить следующий алгоритм:

1. Путём линейных преобразований, приводим матрицу G к такому виду, чтобы в левой её части получилась единичная матрица. В процессе преобразований возможна перестановка столбцов (назовём матрицу перестановки P). Получится некоторая матрица $G' = G \cdot P$.

2. Найдём матрицу H' такую, что $G' \cdot H'^T = 0$. Пользуясь тем, что $G' = [I_k|S]$, получаем, что $H' = [-S^T|I_k]$

3. Так как $(G \cdot P) \cdot (P^{-1} \cdot H^T) = G \cdot H^T = 0$ и $(G \cdot P) \cdot H'^T = G'^T \cdot H'^T = 0$, то верно равенство $H'^T = P^{-1} \cdot H^T$, следовательно $H = H' \cdot P$.

Пример

Допустим, у нас есть матрица G вида

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

Путём линейных образований и перестановки 4 и 5 столбцов получаем матрицу G':

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

Отсюда находим матрицу H' как

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

А матрица H будет иметь вид:

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

Логарифмическое отношение правдоподобия

$$p(x) \cdot p(y|x) = p(x^y) = p(x|y) \cdot p(y)$$

$$p(x|y) = \frac{p(x) \cdot p(y|x)}{p(y)}$$

Выведем формулу логарифмического отношения правдоподобия. Помним, что $p(y|x) = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp -\frac{(y-x)^2}{2\sigma^2}$

$$L = \log \frac{P(c_i = 0|y_i)}{P(c_i = 1|y_i)} = \log \frac{\frac{1}{2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp -\frac{(y_i+1)^2}{2\sigma^2}}{p(y_i)} \cdot \frac{p(y_i)}{\frac{1}{2} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \exp -\frac{(y_i-1)^2}{2\sigma^2}} =$$

$$= \log \exp \frac{(y_i - 1)^2 - (y_i + 1)^2}{2\sigma^2} = \frac{(y_i - 1)^2 - (y_i + 1)^2}{2\sigma^2} = -\frac{4y_i}{2\sigma^2} = -\frac{2y_i}{\sigma^2}$$

10 сентября 2021

Двоичный симметричный код

У нас есть двоичный симметричный код длины n, исправляющий t ошибок. Вероятность ошибки составляет p.

Модель двоичного симметричного кода описывается схемой Бернулли, поэтому имеем формулу для вероятности наличия ошибок при декодировании n бит:

$$P(n) = \sum_{i=t+1}^{n} C_i^n p^i (1-p)^{n-i}$$

Граница Синглтона для линейных кодов

Теорема о границе Синглтона: Если код длины n над q-ичным полем, а минимальное расстояние между кодами равно d, то максимальная мощность такого кода составляет q^{n-d+1}.

Теорема о связи критериев экстремумов по отношению правдоподобия

Докажем, что максимизация по расстоянию Хэмминга эквивалентна минимизации в метрике Евклида:

$$\max_{c \in C} P(y|c) = \min_{c \in C} \sum (y_i - x_i)^2$$

Доказательство.

$$\max_{c \in C} P(y|c) = \max_{c \in C} \prod_{i=0}^{n-1} p(y_i|c_i) = \max_{c \in C} \prod_{i=0}^{n-1} \frac{\exp - \frac{(y_i - x_i)^2}{2\sigma^2}}{\sqrt{s\pi\sigma^2}} \sim$$

$$\sim - \max_{c \in C} \sum_{i=0}^{n-1} \frac{(y_i)^2}{2\sigma^2} \sim \min_{c \in C} \sum_{i=0}^{n-1} (y_i - x_i)^2$$

Ещё одно интересное утверждение о двоичных кодах

Если у нас есть двоичный код длины n размерности k, то количество кодовых слов из этого кода, где на $_i = 1$ либо 0, либо 2^{k-1}.

Рассмотрим ситуацию, где $c_i > 0$. Если слов, где $c_i = 0$ больше половины, то мы можем прибавить к этим кодам 1 и получить так же больше половины других кодов (что невозможно). Аналогично, если слов, где $c_i = 1$ так же не может быть больше половины. Следовательно, таких слов будет ровно половина, т.е. $\frac{2^k}{2} = 2^{k-1}$.

Функция Адамара

Есть функция $f : \mathbb{F}_2^n \to \mathbb{R}$. Есть функция Адамара высшего порядка $\hat{f}(u) = \sum_{v\in\mathbb{F}_2^n}(-1)^{uv} \cdot f(v)$. Тогда выполняется следующее утверждение:

$$\sum_{u\in C_\perp} f(u) = \frac{1}{|C|}\sum_{u\in C}\hat{f}(u)$$

.

Пояснение: $C_\perp$ — дуальный код для C.

Доказательство

$$\frac{1}{|C|}\sum_{u\in C}\hat{f}(u) = \frac{1}{|C|}\sum_{u\in C}\sum_{v\in\mathbb{F}_2^n}(-1)^{<u,v>}f(v) = \frac{1}{|C|}\sum_{v\in\mathbb{F}_2^n} f(v)\sum_{u\in C}(-1)^{<u,v>} =$$
$$= \frac{1}{|C|}\left(\sum_{v\in C_\perp} f(v)|C| + \sum_{v\in\mathbb{F}_2^n\setminus C_\perp} f(v)\sum_{u\in C}(-1)^{<u,v>}\right) =$$
$$= \frac{\sum_{v\in C_\perp} f(v)|C|}{|C|} = \sum_{v\in C_\perp} f(v)$$

Предпоследний переход верен, исходя из предыдущей теоремы.

1 октября 2021

Построение решётки по проверочной матрице

Задана проверочная матрица. Требуется:

1. Построить решётку для этой матрицы
2. Вычислить полином для решётки, задаваемый с помощью формулы:

$$A(k) = \sum_{p \prec k} A(p) \cdot [p \not\equiv k]$$

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

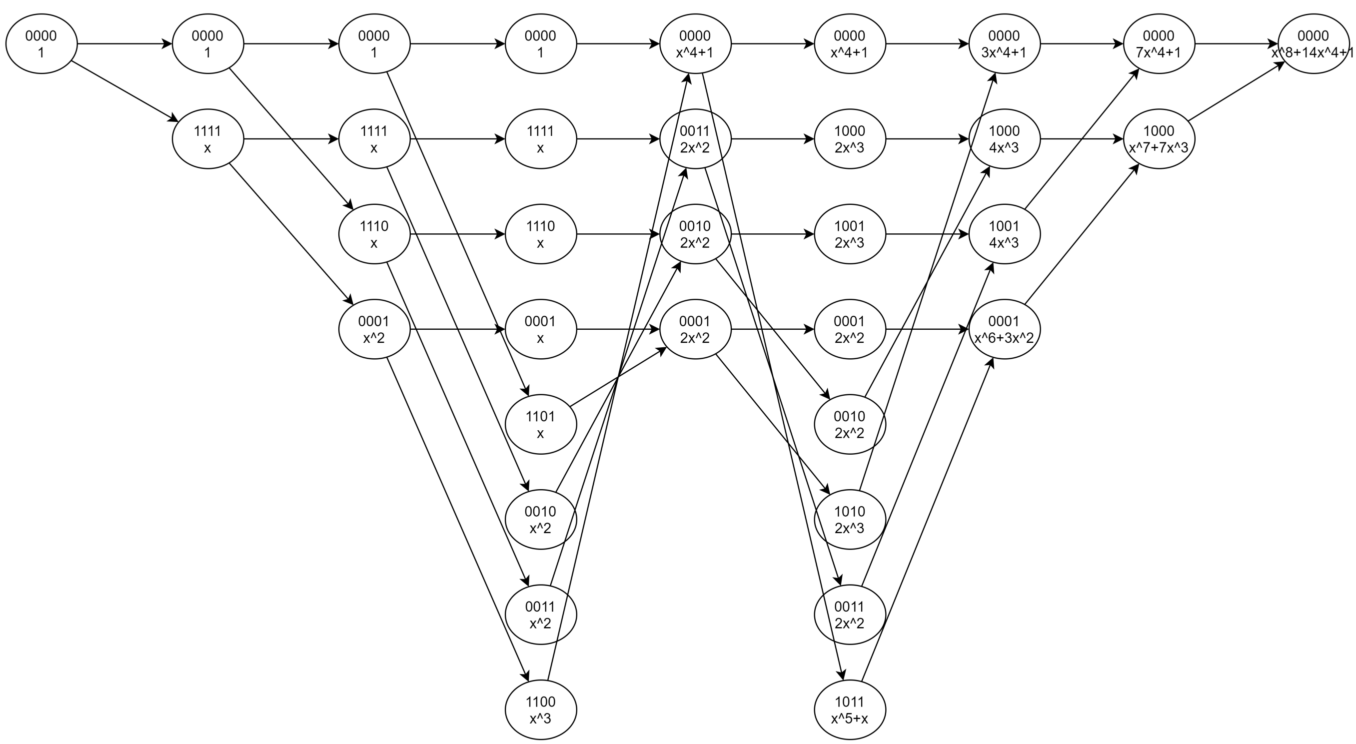

Рис. 1: Решётка для проверочной матрицы

Профиль решётки

Профиль решётки — это такая последовательность чисел $\{s_n\}$, что $s_i = \lfloor \log_2 C_i \rfloor$, где C_i — число вершин на i-м уровне этой решётки. Например, для рёшетки из предыдущего примера, профиль будет равен $\{0, 1, 2, 3, 2, 3, 2, 1, 0\}$.

Также можно вычислить, что количество операций в алгоритме Витерби в зависимости от профиля решётки составляет $O(\sum 2^{s_i})$.