Линейные коды и их свойства

5 января 2023 г. 1:33

Линейным q-ичными кодом длины n с k информационных символов, или (n, k)-кодом над полем $GF(q)$, называется k-мерное подпространство линейного n-мерного пространства всех векторов над полем $GF(q)$.

Линейный (n, k)-код задаётся базисными векторами $\mathbf{g}_1 = (g_{11}, \ldots, g_{1n})$, $\mathbf{g}_2 = (g_{21}, \ldots, g_{2n})$, ..., $\mathbf{g}_k = (g_{k1}, \ldots, g_{kn})$, $g_{ij} \in GF(q)$ или порождающей матрицей

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \cdot & g_{1n} \\ g_{21} & g_{22} & \cdot & g_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ g_{k1} & g_{k2} & \cdot & g_{kn} \end{bmatrix}, \tag{13}$$

при этом кодовое слово $\mathbf{c} = (c_1, c_2, \ldots, c_n)$ является линейной комбинацией базисных векторов:

$$\mathbf{c} = m_1 \cdot \mathbf{g}_1 + \ldots + m_k \cdot \mathbf{g}_k = \mathbf{m} \cdot \mathbf{G}, \tag{14}$$

где $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ – информационная последовательность.

Если порождающая матрицы $\mathbf{G}$ может быть представлена как

$$\mathbf{G} = [\mathbf{I}_k | \mathbf{G}_2], \tag{15}$$

где $\mathbf{I}_k$ - единичная матрица, размером $k \times k$, то эта матрица имеет *левую каноническую форму*.
Если порождающая матрица имеет каноническую форму, то линейный код называется *систематическим*. Для систематического кода, кодовое слово представляется как

$$\mathbf{c} = \mathbf{m} \cdot \mathbf{G} = \mathbf{m} \cdot [\mathbf{I}_k | \mathbf{G}_2] = (\mathbf{m}, \mathbf{m} \cdot \mathbf{G}_2), \tag{16}$$

то есть, кодовое слово состоит из двух подслов: левое подслово это информационная последовательность $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ длины k, а правое подслово состоит из $r = n - k$ проверочных символов.

Пример. Пусть порождающая матрица с длиной кодового слова $n = 7$ и длиной информационной последовательности $k = 4$ имеет вид:

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Необходимо сформировать кодовое слово $\mathbf{c}$ для информационной последовательности $\mathbf{m} = (0101)$.

$$\mathbf{c} = \mathbf{m} \cdot \mathbf{G} = (0101) \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} = (0101010).$$

Свойства линейных кодов

Пусть $\mathbf{x}$ и $\mathbf{y}$ – два слова из X^n. *Расстоянием Хемминга* $d(\mathbf{x}, \mathbf{y})$ между $\mathbf{x}$ и $\mathbf{y}$ называется число позиций, в которых эти слова различаются.

Рассмотрим $\mathbf{c} = (\mathbf{c}_1, \ldots, \mathbf{c}_M)$ как кодовое слово длины n над алфавитом X. *Минимальное расстояние* d кода $\mathbf{C}$ это минимальное расстояние Хемминга между любыми парами кодовых слов из $\mathbf{C}$.

Теорема 1. *Минимальное расстояние линейного кода* $\mathbf{C}$ *равно минимуму из весов ненулевых кодовых слов.*
Theorem 2. Если любые $l \leq d - 1$ столбцов проверочной матрицы $\mathbf{H}$ линейного кода линейно независимы, то минимальное расстояние кода будет по меньшей мере d. Если при этом найдутся d линейно зависимых столбцов, то минимальное расстояние кода равно d.
Теорема 3. Код с минимальным расстоянием d исправляет любые ошибки крастности $t = (d - 1)/2$ и обнаруживает ошибки кратности $t \leq d - 1$.