線型符号
Linear Code

「符号化技術」実験

October, 2018

1　概要

線型符号とは、誤り検出訂正に使われるブロック符号の種類を指す。線型符号は他の符号に比べて、符号化と復号が効率的であるという特徴を持つ。線型符号は、伝送路上を記号列を転送する方法に適用される。したがって通信中に誤りが発生しても、一部の誤りを受信側で検出することができる。線型符号の「符号」は記号のブロックであり、本来の送るべき記号列よりも多くの記号を使って符号化されている。長さ n の線型符号は、n 個の記号を含むブロックを転送する。

2　定義

q 個の元からなる有限体 $F = \mathbb{F}_q$ をとる。このとき n 次元線型空間 Fn の部分空間 C を線型符号という。また $k = dim_F\ C$ とするとき、線型符号 C のことを (n, k) 線型符号という。k 次元部分空間 C はその基底 $g_1, \cdots, g_k \in F^n$ を指定すれば定まる。これらを並べた $k \times n$ 行列 $G = (g_1^t, \cdots, g_k^t)^t$ を線型符号 C の生成行列という。定義から

$$C = \{\, xG \mid x \in F^k \,\}$$

が成り立つ。また k 次元部分空間 C は連立一次方程式で指定しても定まる。そこで

$$C = \{\, y \in F^n \mid yH^t = 0\}$$

となる $(n- k) \times n$ 行列 H を線型符号 C のパリティ検査行列という。定義から $GH^t = 0$ が成り立つ。これらの行列は適当に線型符号 C の基底を取りなおすことによって

$$G = (I|P), \quad H = (-P^t|I)$$

の形にできる。このような G, H を組織符号形式という。このとき符号化前の k 個の記号からなる情報系列がそのまま符号語に現れているので、容易に復号ができる。符号語の残り n − k 個の記号はパリティ検査記号と呼ばれる。

3 シンドローム復号

(n, k) 線型符号を C 、そのパリティ検査行列を H とする。受信語 $y \in F^n$ に対して yH^t をシンドロームという。剰余空間 F^n/C の完全代表系 $\{v1, \cdots, v_{q^{n-k}}\}$ は各 v_i が剰余類 $v_i + C$ のなかで最小の重みをもつとき、コセット・リーダという。このとき受信語 $y \in F^n$ は $yH^t = vH^t$ となるコセット・リーダ v をとると、符号語 $y- v \in C$ に復号される。これをシンドローム復号という。

4 例

次の行列 G を生成行列、あるいは H をパリティ検査行列とする 2 元体 $\mathbb{F}_2$ 上の $(7, 4)$ 線型符号を C とおく。この行列 G,H は組織符号形式である。またコセット・リーダ v とそのシンドローム vH^t は以下の表のようになる。

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

Table 1: $(7, 4)$ 線型符号 C のシンドローム表

コセット・リーダ v	シンドローム vH^t
(0, 0, 0, 0, 0, 0, 0)	(0, 0, 0)
(1, 0, 0, 0, 0, 0, 0)	(0, 1, 1)
(0, 1, 0, 0, 0, 0, 0)	(1, 0, 1)
(0, 0, 1, 0, 0, 0, 0)	(1, 1, 0)
(0, 0, 0, 1, 0, 0, 0)	(1, 1, 1)
(0, 0, 0, 0, 1, 0, 0)	(1, 0, 0)
(0, 0, 0, 0, 0, 1, 0)	(0, 1, 0)
(0, 0, 0, 0, 0, 0, 1)	(0, 0, 1)

たとえば送信したい情報系列を x = (0, 1, 0, 1) とすれば xG = (0, 1, 0, 1, 0, 1, 0) と符号化される。ここで符号語 xG のうち末尾の (0, 1, 0) がパリティ

検査記号である。通信中に先頭で誤りが起こり、受信語が $y = (1, 1, 0, 1, 0, 1, 0)$ になったとすると、そのシンドロームは $yH^t = (0, 1, 1)$ である。そこで上のシンドローム表から対応するコセット・リーダ $v = (1, 0, 0, 0, 0, 0, 0)$ を読み取ることで $xG = y - v$ と復号できる。生成行列 G は組織符号形式だったのでもとの情報系列はその先頭 (0, 1, 0, 1) を読み取ることでわかる。この符号 C は歴史的にはリチャード・ハミングによって 1947 年に初めて発見された誤り訂正符号のひとつである。

5 特性

- 線型符号の最小距離 $d = min_{x \neq y} d(x, y)$ と最小重み $w = min_{x \neq 0} w(x)$ は一致する
- (n, k) 線型符号の最小距離 d は不等式 $d \leq n - k + 1$ を満たす。これをシングルトンの限界式という。
- (n, k) 線型符号は $t < d/2$ 個の誤りを訂正できる。

6 利用

2 進線型符号は電子機器や記憶媒体などで広く使われている。例えば、コンパクトディスクにデジタルデータを格納する際には、リード・ソロモン符号が使われている。また 10 桁の国際標準図書番号（ISBN）$a_1 \cdots a_{10}$ は（$X = 10$ と見做して）11 元体上の一次方程式

$$1a_1 + 2a_2 + \cdots + 10a_{10} \equiv 0 \pmod{11}$$

で定まる線型符号である。