

# 代数系と符号理論 第2回板書プリント

植松友彦 2007.4.13

## 参考書の追加

- 平澤、西島：符号理論入門、培風館、1999.
- 江藤、金子：誤り訂正符号とその応用、テレビジョン学会、1996.

## 1 生成行列と符号化 (補足)

例 1.1.2 の (7.4) 符号の生成行列  $G$  として、

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

を選んだとき、情報  $u = (1001)$  を符号化して得られる符号語  $c$  は

$$c = uG = (1001) \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} = (1001001)$$

となる。一方  $G$  の 1 行目を他の全ての行に加えて得られる

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

もまた、同一の符号の生成行列である (各自確かめよ)。このとき、情報  $u = (1001)$  を符号化すると

$$c = uG' = (1001) \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} = (0001111)$$

が得られ、生成行列  $G$  を用いたときと  $u$  に対応する符号語が異なることが分る。

## 2 パリティ検査行列

### 1) パリティ検査ベクトル

$(n, k)$  符号の生成行列  $G$  に対して、長さ  $n$  のベクトル  $h$  が

$$Gh^T = 0$$

を満足するとき、 $h$  をパリティ検査ベクトルという。このとき、

$$Gh^T = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_k \end{bmatrix} h^T = \begin{bmatrix} e_1 h^T \\ e_2 h^T \\ \vdots \\ e_k h^T \end{bmatrix} = 0$$

なので、パリティ検査ベクトル  $h$  は符号の全ての基底  $\{e_1, \dots, e_k\}$  に直交する。

パリティ検査ベクトルの全体は、 $V$  の部分空間（もっと正確には符号  $C$  の直交補空間）であり、線形代数の理論から、その次元は  $n - k$  である。

### 2) パリティ検査行列

$(n, k)$  符号  $C$  のパリティ検査行列  $H$  とは、その行が 1 次独立なパリティ検査ベクトルとなるような  $(n - k) \times n$  行列である。従って、

$$GH^T = 0$$

が成り立つ。

特に符号  $C$  の生成行列  $G$  が

$$G = [I_k, A] \quad (I_k \text{ は } k \times k \text{ 単位行列})$$

と書ける場合、

$$G[-A^T, I_{n-k}]^T = [I_k, A] \begin{bmatrix} -A \\ I_{n-k} \end{bmatrix} = [-I_k A + A I_{n-k}] = [-A + A] = 0$$

となるので、

$$H = [-A^T, I_{n-k}]$$

はパリティ検査行列である。

尚、(このようにして定めた) パリティ検査行列  $H$  は、同一符号  $C$  の任意の生成行列  $G'$  に対しても

$$G'H^T = 0$$

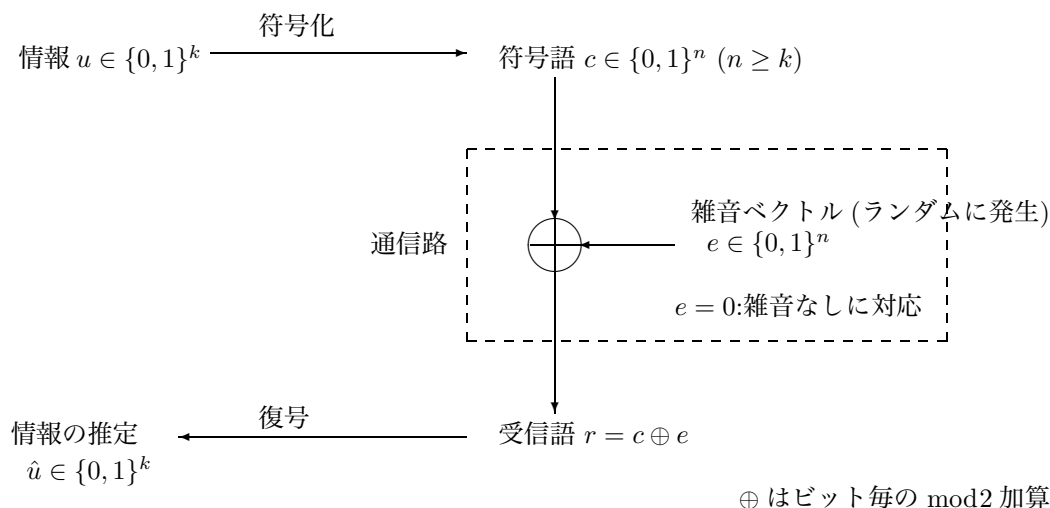
が成り立つことに注意しておく。従って、パリティ検査行列も生成行列同様に扱いやすいものを選ぶことが望ましい。更に、 $GH^T = 0$  は、 $0 = (GH^T)^T = HG^T$  も意味するので、任意の符号語  $c$  について、

$$Hc^T = 0$$

も成り立つ。

### 3) 情報伝送の数学的なモデルの例

情報に冗長性を付け加える操作 : 符号化  
誤りの検出や訂正を行なう操作 : 復号化 (復号)



### 4) 双対符号

パリティ検査行列  $H$  の各行  $\{h_1, \dots, h_{n-k}\}$  は生成行列  $G$  の行に直交する。従って、生成行列の各行の線形結合である符号語にも直交する。そこで、 $H$  の行によって生成される符号を双対符号  $C^\perp$  と呼ぶ。明らかに、

$$\begin{aligned} C^\perp &\triangleq \{x \in F^n : x = \sum_{i=1}^{n-k} a_i h_i, \forall a_i \in F\} \\ &= \{x \in F^n : x \cdot c = 0, \forall c \in C\} \end{aligned}$$

であり、 $C^\perp$  の生成行列の一つが  $H$  である。

## 3 最小距離と最小重み

2つのベクトル  $x = (x_1, x_2, \dots, x_n)$  と  $y = (y_1, y_2, \dots, y_n)$  のハミング距離  $d_H(x, y)$  は

$$d_H(x, y) \triangleq \sum_{i=1}^n d(x_i, y_i)$$

によって定義される。但し、

$$d(x_i, y_i) \triangleq \begin{cases} 0 & x_i = y_i \\ 1 & x_i \neq y_i \end{cases}$$

ハミング距離：2つのベクトルの異なっている成分の数

(例 1)  $x = (1010111), y = (0110101)$  ならば  $d_H(x, y) = 3$ 。

(例 2) 2 元 (4, 2, 2) 符号

$$C = \{c_1 = (0000), c_2 = (1110), c_3 = (0011), c_4 = (1001)\}$$

$$d_H(c_1, c_2) = 3, d_H(c_1, c_3) = 2, d_H(c_1, c_4) = 2$$

$$d_H(c_2, c_3) = 3, d_H(c_2, c_4) = 3, d_H(c_3, c_4) = 2$$

最小距離  $d = 2$ 、符号化率  $r = 2/4 = 1/2$ 。

### 補題 1.2.1

$(n, k)$  線形符号の最小距離  $d_{min}$  は、非零の符号語の最小重み  $w_{min}$  に等しい。

(証明)

$$d_{min} = \min_{\substack{x, y \in C \\ x \neq y}} d_H(x, y) = \min_{\substack{x, y \in C \\ x \neq y}} w_H(x - y)$$

ここで、 $y$  を固定し  $x (\neq y)$  を変化したとき、 $z = x - y$  は零以外の全ての符号語を動けるから、

$$d_{min} = \min_{\substack{x, y \in C \\ x \neq y}} w_H(x - y) = \min_{\substack{z \in C \\ z \neq 0}} w_H(z) = w_{min}$$

が得られる。□

### 定理 1.2.1

$(n, k)$  符号が  $t$  個誤り訂正可能となるための必要十分条件は、 $t < d/2$  である。但し、 $d$  は符号の最小距離である。例えば  $d = 5 \mapsto t < 2.5$ 、 $d = 7 \mapsto t < 3.5$ 、 $d = 6 \mapsto t < 3$  である。

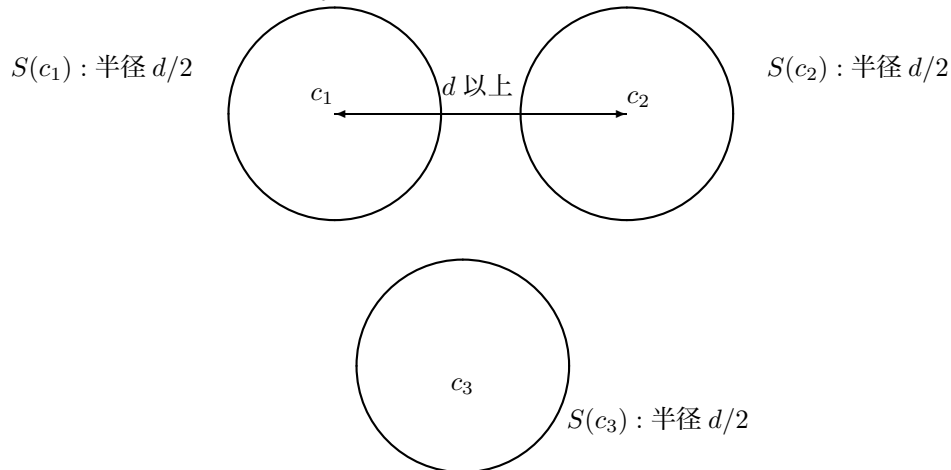
(証明) 符号  $C$  の各符号語  $c_i \in C$  について、復号領域  $S(c_i)$  を

$$S(c_i) \triangleq \{y \in F^n : d_H(c_i, y) < d/2\}$$

によって定める。このとき、最小距離が  $d$  なので、 $S(c_i)$  は共通部分を持たない。すなわち

$$S(c_i) \cap S(c_j) = \emptyset \quad i \neq j$$

符号語  $c_m \in C$  を送信し、 $r$  を受信したとする。このとき、 $r \in S(c_j)$  となる復号領域  $S(c_j)$  があれば、符号語  $c_j$  が送信されたと判定し、そのような復号領域がなければ誤りを検出したと判定する (送信符号語を推定しない)。このような復号法 (限界距離復号法と呼ぶ) によって、 $d/2$  個未満の誤りが直せるのは明らかである。



一方、 $d/2$  個の誤りを (全て) 訂正できないことは次のようにして分かる。一般性を失うことなく  $d$  が偶数であるとする。このとき、 $d_H(x, y) = d$  であるような符号語  $x, y$  については、 $w_H(x - y) = d$  であるので、

$$e_1 + e_2 = x - y, w_H(e_1) = w_H(e_2) = d/2$$

となるような誤りベクトル  $e_1, e_2$  が存在する。このとき、

$$x - e_1 = y + e_2$$

であり、符号語  $x$  に  $d/2$  個の誤りが生じた場合と、符号語  $y$  に  $d/2$  個の誤りが生じた場合を区別できないことになる。従って、 $d/2$  個の誤りを全て訂正できない。□

(例) 繰り返し符号  $C = \{(00000), (11111)\}$

この符号の最小距離は 5 である。従って、2 個までの誤りを訂正できる。事実 (00000) を送信して (01001) を受信したとすれば、最小距離復号法によって (00000) に復号される。

## 補題 1.2.2

$C$  を  $(n, k)$  符号であるとし、 $H$  を  $C$  のパリティ検査行列であるとする。このとき、 $j$  個の列が一次従属であるならば、 $C$  はこれらの列の幾つかに対応する成分が非零である符号語を含む。また、 $C$  が重み  $j$  の符号語を含むならば、 $H$  の  $j$  個の一次従属な列が存在する。

(証明) 前半:  $H = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_n]$  として、 $\mathbf{h}_1, \dots, \mathbf{h}_j$  が一次従属だったとする。すなわち、

$$a_1 \mathbf{h}_1 + a_2 \mathbf{h}_2 + \dots + a_j \mathbf{h}_j = 0$$

が成り立つ全ては零ではない係数  $a_i \in F$  が存在するとする。このとき、

$$H(a_1, a_2, \dots, a_j, 0, \dots, 0) = a_1 \mathbf{h}_1 + a_2 \mathbf{h}_2 + \dots + a_j \mathbf{h}_j = 0$$

から、 $(a_1, a_2, \dots, a_j, 0, \dots, 0)$  は符号語である。

後半: 他方、 $(a_1, a_2, \dots, a_j, 0, \dots, 0)$  が重み  $j$  の符号語だとすれば、すぐ上の式から直ちに、 $\{\mathbf{h}_1, \dots, \mathbf{h}_j\}$  が一次従属であることが分かる。□

## $H$ の列ベクトルと最小距離の関係

2 元符号の場合、次のことを知っているると便利である。

- $H$  の列ベクトルに零ベクトルがなければ、最小距離は 2 以上である。
- $H$  の列ベクトルに零ベクトルがなく、全て異なっていれば、最小距離は 3 以上である。