代数系と符号理論　第3回板書プリント

植松友彦 2007.4.20

0.1　符号理論の研究テーマ

1. 最良符号はどのような性能を有するか (性能解析)
2. 良い符号はどのように設計法できるか (設計法)
3. そのような符号はどのように復号できるか (復号法)

1　VG 限界

[定理 **1.2.2**] **(Varshamov-Gilbert 限界)**
次の不等式が成り立つならば、符号長 n、パリティ検査記号数 $m(=n-k)$ 個以下で、最小距離 d 以上の 2 元符号が存在する。

$$1+\binom{n-1}{1}+\cdots+\binom{n-1}{d-2}<2^m$$

(証明) 任意の $d-1$ 個の列が一次独立となるような $m\times n$ 行列の作り方を考えよう。

1. 非ゼロの m 次元ベクトルを一つ選び、パリティ検査行列 H の第 1 列とし、$i=1$ とする。
2. 検査行列 H の第 1 列から第 i 列までのうち $d-2$ 個以下の列の 1 次結合を全て作る。
3. m 次元ベクトルで、2. で作った 1 次結合として現れてこないものが有れば、その 1 つを選んで、検査行列の $i+1$ 列として、$i\leftarrow i+1$ として 2. に戻る。そのような m 次元ベクトルがなければ終了する。

この構成法の 2. で作られる 1 次結合の数を L_i とすれば、

$$L_i=\binom{i}{0}+\binom{i}{1}+\binom{i}{2}+\cdots+\binom{i}{d-2}$$

となる。これらの 1 次結合として現れる異なる m 次元ベクトルの数は、L_i を越えないから、$L_i<2^m$ ならば、検査行列に第 $i+1$ 列を付け加えられることになる。このことと、

$$\binom{i}{0}+\binom{i}{1}+\binom{i}{2}+\cdots+\binom{i}{d-2}$$

が i について単調増加であることから、定理を得る。

Varshamov-Gilbert 限界は、大きな符号長 n に対して良い符号の存在を示すが、その具体的な構成法を示している訳ではない。

[定義 **1.2.5**] **(2** 元ハミング符号**)**
すべての非ゼロの 2 元 m 次元ベクトルを列とするパリティ検査行列によって定義される符号である。

[例] $m = 3$ の場合、

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

2 元ハミング符号のパラメータ

- 符号長：$n = 2^m - 1$
- 次元：$k = n - m = 2^m - m - 1$
- 最小距離：$d = 3$

[定義 **1.2.6**] **(2** 元拡大ハミング符号**)**
ハミング符号のパリティ検査行列に全零の列を追加した後、全 1 の行を追加したものをパリティ検査行列とする符号である。

[例] $m = 3$ の場合、

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

2 元拡大ハミング符号のパラメータ

- 符号長：$n = 2^m$
- 次元：$k = n - (m + 1) = 2^m - m - 1$
- 最小距離：$d = 4$

[定義 **1.2.7**] **(**陪直交符号**)**
2 元拡大ハミング符号の双対符号を陪直交符号という。

[例] $m = 3$ の場合、

$$\text{生成行列：}\ G = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

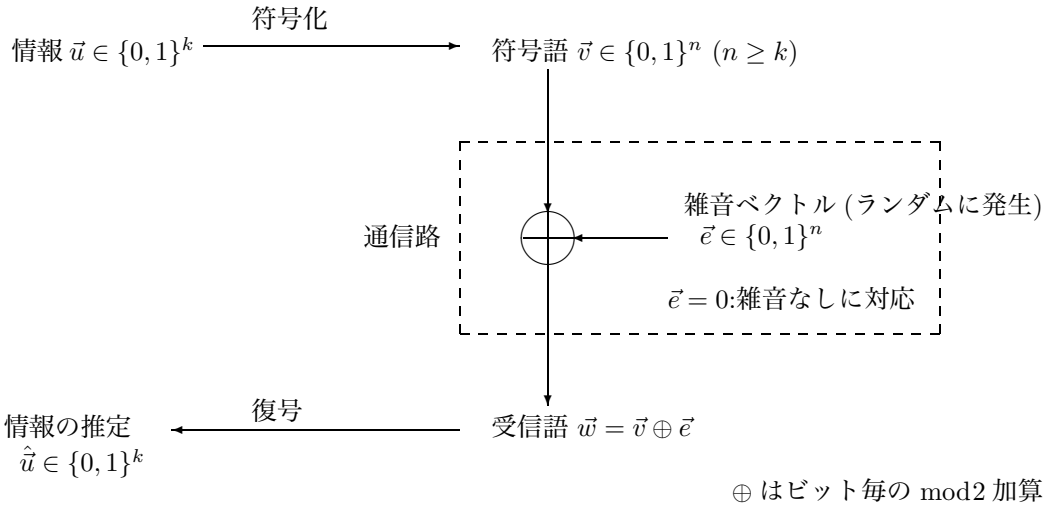
陪直交符号のパラメータ

- 符号長：$n = 2^m$
- 次元：$k = m + 1$
- 最小距離：$d = n/2$

2　限界距離復号法と最小距離復号法

2.1　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）



[定義 **1.3.1**] (限界距離復号器 (**bounded distance decoder**))
最小距離 d の符号 C と受信語 r が与えられたとき、$d(r,c) < d/2$ であるような符号語 $c \in C$ が存在すれば、符号語 c に復号し、もしそのような符号語がなければ復号不能と判定する復号器を限界距離復号器と呼ぶ。

尚、最小距離が d なので、受信語からの距離が $d/2$ より小さい符号語は高々1 個しか存在しないことに注意！

[定義] (最小距離復号器 (**minimum distance decoder**))
受信語 r が与えられたとき、$d(r,c)$ を最小にするような符号語 c がただ一つ存在すれば、符号語 c に復号し、もしそのような符号語が 2 個以上存在すれば復号不能と判定する復号器を最小距離復号器と呼ぶ。

3　シンドローム復号法

[定義 **1.3.2**] (コセット)
C が (n,k) 線形符号で、$a \in F^n$ とする。a を含むコセットとは、全ての符号語にベクトル a を加えてできる F^n の部分集合

$$a + C = \{a + c | c \in C\}$$

のことである。

[補題 **1.3.1**] 2 つのベクトルが同一のコセットに含まれる為の必要十分条件は、それらが同一のシンドロームを有することである。

(証明) H を符号のパリティ検査行列とする。$x, y \in F^n$ が同一のコセットに含まれているとする。すなわち、ある符号語 c_1, c_2 に対して、$x = a + c_1, y = a + c_2$ とする。このとき、

$$Hx^T = H(a + c_1)^T = Ha^T = H(a + c_2)^T = Hy^T$$

が成り立つので、これらの語は同一のシンドロームを持つ。

また逆に、x, y が同一のシンドロームを持てば、$Hx^T = Hy^T$ から $H(x - y)^T = 0$ が得られ、$x - y$ が符号語である。従って、$x = y + c_1$ となる符号語が存在するので、x, y は同一のコセットに含まれる。

[定義] (シンドローム復号法)
受信語 r に対して、シンドローム $s = Hr^T$ を求め、シンドローム s を有する最小個数の誤りを持つ誤りパターン f を選び出し、受信語を $r - f$ に復号する復号法をシンドローム復号法と呼ぶ。すなわち、シンドローム復号法では、r の属するコセットの中で、最小重みのベクトル f(コセット代表元という) を求め、$r - f$ に復号する。

[例] パリティ検査行列

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

を持つ (6,3) ハミング符号 C について、シンドローム復号法を行おう。まず、準備として、シンドロームとコセット代表元の対応表を作る
(1) まず、左上を零ベクトルとして、全ての符号語を一行に書く。この場合、

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

なので、

シンドローム								
000	000000	100110	010101	001011	110011	101101	011110	111000

(2) 次に 1 行目に出現しなかった 2 元ベクトルで、重みが最小のものを一つ選び、一行目のベクトルに加えたものを二行目に書く。選んだベクトル (コセット代表元であり、これが誤り訂正可能な誤りパターンとなる) を 100000 とすると、

シンドローム								
$(000)^T$	000000	100110	010101	001011	110011	101101	011110	111000
$(110)^T$	100000	000110	110101	101011	010011	001101	111110	011000

(3) これまでに出現しなかった 2 元ベクトルで、重みが最小のものを一つ選び、一行目のベクトルに加えたものを三行目に書く。選んだベクトルを 010000 とすると、

シンドローム								
$(000)^T$	000000	100110	010101	001011	110011	101101	011110	111000
$(110)^T$	100000	000110	110101	101011	010011	001101	111110	011000
$(101)^T$	010000	110110	000101	011011	100011	111101	001110	101000

(4) 上記の操作を繰り返し、全ての 2 元ベクトルが表に書かれるまで続けると、例えば次の表が得られる。

シンドローム								
$(000)^T$	000000	100110	010101	001011	110011	101101	011110	111000
$(110)^T$	100000	000110	110101	101011	010011	001101	111110	011000
$(101)^T$	010000	110110	000101	011011	100011	111101	001110	101000
$(011)^T$	001000	101110	011101	000011	111011	100101	010110	110000
$(100)^T$	000100	100010	010001	001111	110111	101001	011010	111100
$(010)^T$	000010	100100	010111	001001	110001	101111	011100	111010
$(001)^T$	000001	100111	010100	001010	110010	101100	011111	111001
$(111)^T$	100001	000111	110100	101010	010010	001100	111111	011001

シンドローム	コセット代表元 (訂正可能な誤りパターン)
$(000)^T$	(000000)
$(110)^T$	(100000)
$(101)^T$	(010000)
$(011)^T$	(001000)
$(100)^T$	(000100)
$(010)^T$	(000010)
$(001)^T$	(000001)
$(111)^T$	(100001)

従って、$r = (100010)$ を受信した場合、シンドローム復号法では、シンドローム

$$s = Hr^T = (100)^T$$

に対応するコセット代表元は、上の表から $f = (000100)$ となり $c = r - f = (100110)$ に復号される。

[補題 1.3.2] 線形符号 C が t 重誤り訂正可能であるための必要十分条件は、重みが t 以下の全ての (誤り) ベクトルがコセット代表元となることである。

(証明は教科書参照)

[定理 1.3.1](ハミング限界) C が要素数 q の有限体上の (n, k) 符号であり、t 重誤り訂正可能であれば、

$$\sum_{j=0}^{t} (q-1)^j \binom{n}{j} \leq q^{n-k}$$

でなければならない。

(証明)　重み t 以下の語の総数が、

$$1 + (q-1)\binom{n}{1} + (q-1)^2\binom{n}{2} + \cdots + (q-1)^t\binom{n}{t}$$

であり、これらに対応するシンドロームは全て異なっていなければならないことと、シンドロームの総数が q^{n-k} であることから導かれる。□

特に、2元ハミング符号については、$n = 2^m - 1, k = n - m, t = 1$ から

$$\binom{n}{0} + \binom{n}{1} = 1 + n = 2^m = 2^{n-k}$$

となり、ハミング限界の等号が成り立つ。等号を満たす符号のことを完全符号という。