応用数学５ 2010 年 4 月 28 日

誤り訂正符号の導入，線形符号

講師: 安永憲司

1 誤り訂正問題

情報をメディアに記憶するときや通信するとき，雑音が発生すると，情報が変化してしまう．雑音が発生したとしても，元の情報を復元できる仕組みがあれば便利である．このような問題を解決する数学的道具として誤り訂正符号がある．

1.1 符号の例

情報 $x = (x_1, x_2, x_3, x_4) \in \{0,1\}^4$ をメディアに保存すること考える．このとき，1 ビットの誤りが発生したとしても元の情報が復元できる仕組みが欲しいとする．ただし，ここで誤りとは，0 が 1 に，1 が 0 に変化することを指す．

x をそのまま保存すると，1 ビットの誤りが発生すると，元の情報が復元できない．

3 回繰り返し符号

x の代わりに

$$(x_1, x_1, x_1, x_2, x_2, x_2, x_3, x_3, x_3, x_4, x_4, x_4)$$

を保存する．こうすれば，1 ビットの誤りがあっても必ずもとの情報を復元できる．

$x = (1,0,1,0)$ のとき，$(1,1,1,0,0,0,1,1,1,0,0,0)$ を保存する．1 ビット誤りが発生して $(1,0,1,0,0,0,1,1,1,0,0,0)$ となったとしても誤りは訂正できる．また，2 ビットまでの誤りが発生する場合を考えると，$(1,1,1,0,0,0,1,1,0,0,0,1)$ となった場合は誤り訂正できるが，$(1,1,1,0,0,0,1,1,1,1,0,1)$ となった場合は，$x = (1,0,1,1)$ なのか $(1,0,1,0)$ なのか区別できない．

Hamming の方法

1 ビット誤り訂正の方法として，Hamming は 1950 年に以下の方法を考案した．x の代わりに

$$(x_1, x_2, x_3, x_4, x_2 \oplus x_3 \oplus x_4, x_1 \oplus x_2 \oplus x_4, x_1 \oplus x_3 \oplus x_4)$$

を保存する．ただし，$\oplus$ は排他的論理和演算である．つまり，$0 \oplus 0 = 0, 0 \oplus 1 = 1, 1 \oplus 0 = 1, 1 \oplus 1 = 0$．排他的論理和は法を 2 とした足し算と考えればよい．

上記の変換は，x から xG への写像と等価である．

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

相異なる $x, y \in \{0,1\}^4$ に対して，xG と yG は 3 つ以上の位置で値が異なることが確認できる．このとき，1 ビット誤りが発生したとしても，それを必ず訂正できる．

2 基本用語の定義

有限アルファベット集合 Σ を考える.

定義 1 (Hamming 距離) $x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \Sigma^n$ に対し，x と y の Hamming 距離とは，$d(x, y) = |\{i : x_i \neq y_i\}|$.

命題 2 Hamming 距離は距離の公理を満たす．つまり，任意の $x, y, z \in \Sigma^n$ に対して以下が成り立つ.

1. 同一性: $d(x, y) = 0 \Leftrightarrow x = y$.
2. 対称性: $d(x, y) = d(y, x)$.
3. 三角不等式: $d(x, y) \leq d(x, z) + d(z, y)$.

定義 3 (符号) Σ 上の長さ n の符号とは，Σ^n の部分集合．符号 $C \subseteq \Sigma^n$ に対し，$c \in C$ を C の符号語，n を C の符号長と呼ぶ.

定義 4 (符号の最小距離) 符号 $C \subseteq \Sigma^n$ の最小距離 d とは,

$$d = \min_{\substack{c_1, c_2 \in C \\ c_1 \neq c_2}} d(c_1, c_2).$$

定義 5 (相対最小距離) 符号 $C \subseteq \Sigma^n$ の相対最小距離 δ とは，$\delta = \frac{d}{n}$．ただし，d は C の最小距離.

定義 6 (レート) 符号 $C \subseteq \Sigma^n$ のレート R とは,

$$R = \frac{\log_{|\Sigma|} |C|}{n}.$$

例 7 前節の 3 回繰り返し符号は，符号長が 12 であり，符号語数は $2^4 = 16$ である．符号の最小距離は 3，相対最小距離は $\frac{3}{12} = \frac{1}{4}$，レートは $\frac{\log_2 16}{12} = \frac{1}{3}$ である．また，Hamming の方法は，符号長が 7 であり，符号語数と最小距離は 3 回繰り返し符号と同じである．したがって，相対最小距離は $\frac{3}{7}$ であり， レートは $\frac{4}{7}$ である.

命題 8 符号 $C \subseteq \Sigma^n$ について，以下はすべて等価である.

1. C の最小距離が d 以上である.
2. d が奇数のとき，C を使って任意の $(d-1)/2$ 個までの誤りを訂正できる.
3. C を使って任意の $d-1$ 個までの誤りを検出できる.
4. C を使って任意の $d-1$ 個までの消失を訂正できる.

証明: まず，C の最小距離が d 以上であると仮定して，その他が成り立つことを示す．d が奇数のとき，各符号語を中心とした半径 $(d-1)/2$ の球を考える．この球は Hamming 球と呼ばれ，それぞれは交わらないことがわかる．つまり，$(d-1)/2$ 個以下の誤りが発生したとしても，どの符号語を送信したかは一意に特定できる．検出について，$d-1$ 個以下の誤りが発生した場合，誤り数が 0 でない限りその系列は符号語ではないので，誤りの検出が可能である．消失について，$d-1$ 個以下のシンボルが消失した場合，残った系列が一致する符号語は一意に特定できる．もし二つの符号語と一致していたとすると，その二つの符号語の距離は $d-1$ 以下となり，最小距離の仮定に矛盾する.

次に，C の最小距離が $d-1$ 以下のとき，その他が成り立たないことを示す．ある二つの符号語 c_1 と c_2 の距離が $d-1$ であるとする．このとき，c_1 と c_2 どちらからの距離も $(d-1)/2$ 以下である系列 x が存在する．もし $(d-1)/2$ 個の誤りが発生して x となった場合，c_1 なのか c_2 なのか区別できず，誤り訂正できないことになる．検出について，$d-1$ 個の誤りによって c_1 を c_2 にすることができる．このとき，もともと c_2 だったのか，c_1 に誤りが発生して c_2 になったのか区別できず，誤り検出ができない．消失について，c_1 と c_2 の異なる記号の位置を消失した場合，c_1 と c_2 から同じ系列が作られる．つまり，c_1 なのか c_2 なのか区別できず，消失訂正はできない．■

補足 9 上記の命題において d が偶数の場合，$d/2-1$ 個までの誤りは訂正できるが，$d/2$ 個の誤りは訂正できない．したがって，符号が t 個までの誤りを訂正できるとき，その符号の最小距離は $2t+1$ 以上もしくは $2t+2$ 以上である．

3 線形符号

定義 10 (線形符号) Σ が体であり，$C \subseteq \Sigma^n$ がベクトル空間 Σ^n の線形部分空間であるとき，C は線形符号である．

3.1 有限体

定義 11 (体) 体とは，二つの演算，加法 $+$ と乗法 $\cdot$，が定義される要素の集合 $\mathbb{F}$ であり，以下の性質を満たすものである．特に，$\mathbb{F}$ が有限集合のとき，有限体という．任意の $a, b, c \in \mathbb{F}$ について，

1. $\mathbb{F}$ は $+$ と $\cdot$ に関して閉じている: $a+b \in \mathbb{F}, a \cdot b \in \mathbb{F}$.
2. 結合律: $a+(b+c) = (a+b)+c$, $a \cdot (b \cdot c) = (a \cdot b) \cdot c$.
3. 可換律: $a+b = b+a$, $a \cdot b = b \cdot a$.
4. 分配律: $a \cdot (b+c) = a \cdot b + a \cdot c$, $(a+b) \cdot c = a \cdot c + b \cdot c$.

さらに，二つの異なる要素 0（加法単位元）と 1（乗法単位元） が存在し，以下を満たす．

5. 加法単位元: 任意の $a \in \mathbb{F}$ に対して，$a+0 = a$.
6. 乗法単位元: 任意の $a \in \mathbb{F}$ に対して，$a \cdot 1 = a$.
7. 加法逆元の存在: 任意の $a \in \mathbb{F}$ に対して，$-a \in \mathbb{F}$ が存在して，$a+(-a) = 0$.
8. 乗法逆元の存在: 任意の $a \in \mathbb{F} \setminus \{0\}$ に対して，$a^{-1} \in \mathbb{F}$ が存在して，$a \cdot a^{-1} = 1$.

以降，体 $\mathbb{F}$ の要素集合は $\mathbb{F}$ で表し，要素数 q の有限体を $\mathbb{F}_q$ と表す．特に，$\mathbb{F}_2 = \{0, 1\}$ とし，加法は排他的論理和，乗法は $0 \cdot 1 = 0$, $1 \cdot 1 = 1$ とする．

例 12 有理数集合 $Q = \{\frac{a}{b} : a, b \text{ は整数}, b \neq 0\}$ や実数集合，複素数集合は体である．

定義 13 体 $\mathbb{F}$ に対して，1 を p 回足して 0 となるような正整数 p の最小値を $\mathbb{F}$ の標数という．そのような p が存在しないとき，標数は 0 とする．

命題 14 体の標数は 0 もしくは素数である．特に，有限体の標数は素数である．

命題 15 標数 p の有限体の要素数は，ある正整数 s に対して，p^s である．

例 16 正整数 m に対して，$\mathbb{Z}_m = \{0, 1, \ldots, m-1\}$ とする．$\mathbb{Z}_m$ 上での演算を，m を法とした加法・乗法と定義する．つまり，$a, b, c \in \mathbb{Z}_m$ に対し，$a + b = c \iff a + b \equiv c \pmod{m}$, $a \cdot b = c \iff a \cdot b \equiv c \pmod{m}$ と定義する．このとき，m が素数であることが，$\mathbb{Z}_m$ が体となるための必要十分条件である．また，素数 p に対して，$\mathbb{Z}_p$ の標数は p である．

3.2 ベクトル空間

体 $\mathbb{F}$ 上で定義される n 次元ベクトル空間 $\mathbb{F}^n$ を考える．ベクトル空間上での演算として，$x = (x_1, \ldots, x_n), y = (y_1, \ldots, y_n) \in \mathbb{F}^n, \alpha \in \mathbb{F}$ に対して，加法 $x + y = (x_1 + y_1, \ldots, x_n + y_n)$ とスカラー乗法 $\alpha x = (\alpha x_1, \ldots, \alpha x_n)$ を定義する．

定義 17 (ベクトル空間) 集合 V が体 $\mathbb{F}$ 上のベクトル空間（線形空間）であるとは，以下を満たすときである．任意の $x, y, z \in V$, $\alpha, \beta \in \mathbb{F}$ に対して，

1. V は加法についてアーベル群（可換群）である．つまり，
   (a) $x + y \in V$.
   (b) $(x + y) + z = x + (y + z)$.
   (c) $x + y = y + x$.
   (d) 単位元 $0 \in V$ が存在し， $0 + x = x$.
   (e) 逆元 $-x \in V$ が存在し，$x + (-x) = 0$.
2. $\alpha(x + y) = \alpha x + \alpha y$.
3. $(\alpha + \beta)x = \alpha x + \beta x$.
4. $(\alpha\beta)x = \alpha(\beta x)$.
5. $\mathbb{F}$ の乗法単位元 $1 \in \mathbb{F}$ に対して，$1x = x$.

定義 18 (線形部分空間) ベクトル空間 V の部分集合 L が，V と同じ加法・スカラー乗法を用いてベクトル空間をなすとき，L を V の線形部分空間という．

命題 19 部分集合 $L \subseteq \mathbb{F}^n$ が $\mathbb{F}^n$ の線形部分空間であるための必要十分条件は，任意の $x, y \in L$ と $\alpha \in \mathbb{F}$ に対して，$x + y \in L, \alpha x \in L$ が成り立つことである．

上の命題より，$\mathbb{F}^n$ の線形部分空間は必ず，全零ベクトル $(0, \ldots, 0) \in \mathbb{F}^n$ を含むことがわかる．

定義 20 (線形結合) V を $\mathbb{F}$ 上のベクトル空間とする．ベクトル集合 $v_1, \ldots, v_r \in V$ とスカラー $\alpha_1, \ldots, \alpha_r \in \mathbb{F}$ に対し，$\alpha_1 v_1 + \cdots + \alpha_r v_r$ を $v_1, \ldots, v_r$ の線形結合という．

定義 21 (線形独立) V を $\mathbb{F}$ 上のベクトル空間とする．ベクトル集合 $v_1, \ldots, v_r \in V$ が線形独立であるとは，以下を満たすときである;

$$\alpha_1 v_1 + \cdots + \alpha_r v_r = 0 \iff \alpha_1 = \cdots = \alpha_r = 0.$$

ベクトル集合が線形独立でないとき，線形従属であるという．

定義 22 V を $\mathbb{F}$ 上のベクトル空間とする．ベクトル集合 $v_1, \ldots, v_r \in V$ によって張られる空間 $\mathrm{span}(V)$ とは，

$$\mathrm{span}(V) = \{\alpha_1 v_1 + \cdots + \alpha_r v_r : \alpha_1, \ldots, \alpha_r \in \mathbb{F}\}.$$

$\mathrm{span}(V)$ は V の線形部分空間である．

定義 23 (基底) V を $\mathbb{F}$ 上のベクトル空間とする．部分集合 $B = \{v_1, \ldots, v_k\} \subseteq V$ が V の基底であるとは，$\mathrm{span}(B) = V$ かつ B が線形独立のときである．

B がベクトル空間 V の基底であるとき，任意のベクトル $v \in V$ は B の線形結合として一意に表される．つまり，ある $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ が一意に存在し，

$$v = \alpha_1 v_1 + \cdots + \alpha_k v_k.$$

ベクトル空間 V には様々な基底が存在するが，基底に含まれるベクトルの数は同じである．この数のことを V の**次元**という．

命題 24 次元 k のベクトル空間 V に対し，線形独立である k 個のベクトル集合 $v_1, \ldots, v_k \in V$ は，V の基底である．

命題 25 $\mathbb{F}_q^n$ の線形部分空間 L の次元が k のとき，$|L| = q^k$.

定義 26 $u = (u_1, \ldots, u_n), v = (v_1, \ldots, v_n) \in \mathbb{F}^n$ とする．

1. u と v の**内積**とは，
$$\langle u, v \rangle = u_1 v_1 + \cdots + u_n v_n \in \mathbb{F}.$$
2. u と v が**直交する**とは，$\langle u, v \rangle = 0$ となるときである．
3. 線形部分空間 $S \subseteq \mathbb{F}^n$ の**直交補空間** $S^\perp$ とは，
$$S^\perp = \{v \in \mathbb{F}^n : \text{任意の } u \in S \text{ に対して } \langle u, v \rangle = 0\}.$$

上記で定義した内積は，スカラー積と呼ばれるものである．内積はスカラー積を一般化した概念であるが，ここではスカラー積を単に内積と呼ぶことにする．

線形符号 C として，$C = C^\perp$ を満たす符号がある．このような符号は自己直交符号と呼ばれる．

命題 27 線形部分空間 $S \subseteq \mathbb{F}^n$ の次元が k のとき，直交補空間 $S^\perp$ の次元は $n - k$ である．

例 28
1. $\{000, 111\}$ や $\{000, 011, 110, 101\}$ は $\mathbb{F}_2^3$ の線形部分空間である．
2. $\{011, 110\}$ や $\{001, 100, 010\}$ は線形独立であるが，$\{001, 010, 011\}$ は線形従属である．
3. $\mathbb{F}_2$ 上のベクトル集合 $\{1000, 1110\}$ によって張られる空間は $\{0000, 1000, 1110, 0110\}$ である．
4. $V = \mathbb{F}_q^4$ は，次元 4 のベクトル空間であり，$B = \{1000, 0100, 0010, 0001\}$ は V の基底である．
5. 線形符号 $C = \{0000, 1000, 1110, 0110\}$ に対し，$C^\perp = \{0000, 0110, 0001, 0111\}$ である．

3.3 生成行列

線形符号 C は，$\mathbb{F}_q^n$ の線形部分空間であるので，基底 $g_1, \ldots, g_k \in \mathbb{F}_q^n$ が存在する．ただし，k は C の次元である．任意の符号語 $c \in C$ は基底の線形結合で表現でき，逆に，基底によって張られる空間が符号 C となる．つまり，次元 k の線形符号 C は $k \times n$ 行列 $G = (g_1^T, \ldots, g_k^T)^T$ を用いて

$$C = \{xG : x \in \mathbb{F}_q^k\}$$

と表すことができる．ただし，行列 M に対して，M^T は M の転置行列を表す．このとき，G を符号 C の**生成行列**と呼ぶ．

定義 29 (生成行列と符号化) $C \subseteq \mathbb{F}_q^n$ を次元 k の線形符号とする．行列 $G \in \mathbb{F}_q^{k \times n}$ に対して，G の k 行が張る空間が C と一致するとき，G を C の生成行列と呼ぶ．生成行列は，メッセージ $x \in \mathbb{F}_q^k$ を $xG \in \mathbb{F}_q^n$ へ符号化する方法を与えている．つまり，線形符号とは，$E(x) = xG$ となるような線形写像 $E : \mathbb{F}_q^k \to \mathbb{F}_q^n$ を符号化関数としてもつ符号である．

要素数 q の有限体上の線形符号 C が次元 k であり最小距離が d のとき，C を $(n, k, d)_q$ 線形符号，もしくは q 元 (n, k, d) 線形符号という．有限体の要素数 q が明らかな場合や，最小距離 d が不明の場合には，q や d を省略して (n, k) 線形符号という場合もある．(n, k) 線形符号 C のレートは k/n である．また，一般に，符号長 n，レート R の符号を (n, Rn) 符号という．

定義 30 (Hamming 重み) $x \in \mathbb{F}_q^n$ の Hamming 重みとは $\mathrm{wt}(x) = |\{i : x_i \neq 0\}|$.

命題 31 $x, y \in \mathbb{F}_q^n$ に対して，$\mathrm{wt}(x - y) = d(x, y)$.

命題 32 線形符号の最小距離は，非零符号語の Hamming 重みの最小値である．

例 33 二元符号 $C = \{0000, 1110, 0111, 1001\}$ を考える．C は線形符号であるので，符号の最小距離は，非零符号語の最小 Hamming 重みに等しい．$\mathrm{wt}(0000) = 0, \mathrm{wt}(1110) = 3, \mathrm{wt}(0111) = 3, \mathrm{wt}(1001) = 2$ であるので，最小距離は 2 である．

例 34 行列 $G = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \end{pmatrix}$ は二元符号 $C = \{0000, 1110, 0111, 1001\}$ の生成行列である．一般に，G に対して行基本操作を施しても，符号は変化しない．つまり，$G' = \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$ も C の生成行列である．

生成行列 G に対し，行基本操作と列の置換を施すと，$G' = [I_k|M]$ という形にすることができる．ここで，I_k は $k \times k$ の単位行列であり，M は $k \times (n - k)$ のある行列である．このような形をした生成行列は**組織的**（または，**標準形**）であるという．組織的な生成行列をもつ符号では，$x \in \mathbb{F}_q^k$ を符号化すると，符号語の最初の k 個は，メッセージ x そのものになる．任意の線形符号は，列の置換を許せば，このような符号化をすることが可能である．

補足 35 一般的な符号に比べ，線形符号は以下の利点をもつ．

1. 符号がベクトル空間であるので，基底を用いて符号を表現できる．
2. 符号の最小距離は，非零符号語の Hamming 重みの最小値に等しく，計算しやすい．

3.4 パリティ検査行列

$(n,k,d)_q$ 線形符号 C は，$\mathbb{F}_q^n$ の線形部分空間をなし，生成行列 G で定義されることを見てきた．ただし，G は $k \times n$ 行列であり，ランクは k である．線形符号を定義する方法として，直交補空間を用いることもできる．つまり，C に含まれるすべてのベクトルと直交するベクトル全体からなる部分空間 $C^\perp$ を用いて定義する．$(n-k) \times n$ 行列 H を $C^\perp$ の生成行列とする．次元 k の線形符号 C は H を用いて

$$C = \{c \in \mathbb{F}_q^n : cH^T = 0\}$$

と表すことができる．このとき，H を C の**パリティ検査行列**と呼ぶ．

定義 36 (パリティ検査行列) $C \subseteq \mathbb{F}_q^n$ を次元 k の線形符号とする．行についてフルランクである行列 $H \in \mathbb{F}_q^{(n-k)\times n}$ に対して，C に含まれるすべての符号語 c について $cH^T = 0$ であるとき，H を C パリティ検査行列と呼ぶ．パリティ検査行列は，系列 $x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$ が符号語であるための制約条件を与えている．線形符号とは，系列 x が符号語であるための制約条件が，$x_1, \ldots, x_n$ の線形方程式で与えられる符号である．

命題 37 パリティ検査行列 H で定義される線形符号 C の最小距離は，H の列のうちの d 個の列が線形従属であるような d の最小値と一致する．

3.5 Hamming 符号

$(7,4,3)_2$ Hamming 符号

生成行列

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

で定義される符号 $C_{\mathrm{Ham}} = \{xG : x \in \mathbb{F}_2^4\}$ は，$(7,4,3)_2$ Hamming 符号である．

また，以下の H を用いて $C_{\mathrm{Ham}} = \{x \in \mathbb{F}_2^7 : xH^T = 0\}$ と表すことができる．

$$H = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

Hamming 符号における 1 ビット訂正方法

符号語 $c \in C_{\mathrm{Ham}}$ を送信して，1 ビットの誤りが発生し，y になったとする．このとき，$y = c + e_i$ と表すことができる．ただし，e_i は i ビット目が 1 で，その他は 0 をとるベクトルである．

$(7,4,3)_2$ Hamming 符号は，1 ビット誤り訂正が可能なので，すべての符号語と y を比べて，距離が 1 の符号語を見つけることで，誤り訂正が可能である．しかし，より単純な方法で誤り訂正をする方法がある．

$$yH^T = (c + e_i)H^T = cH^T + e_iH^T = e_iH^T = (H \text{ の } i \text{ 番目の列})$$

となるので，誤りが発生した位置 i を計算することができる．

一般的な Hamming 符号

H_r を $r \times (2^r - 1)$ 行列で，各 i 番目の列が i の二進表現であるものとする．H_r はフルランクである．このとき，

$$C_{\mathrm{Ham}}^{(r)} = \{c \in \mathbb{F}_2^{2^r-1} : cH_r^T = 0\}$$

と定義した符号は二元 $(2^r - 1, 2^r - 1 - r)$ Hamming 符号である．

命題 38 Hamming 符号の最小距離は 3 である．

証明: 命題 37 により，H_r の列のうち，どの 2 つも線形従属ではなく，ある 3 つが線形従属であることを示せばよい．前者は，H_r の各列が互いに異なることからわかる．後者は，H_r の最初 3 つの列が，$1, 2, 3$ の二進表現であるので，足し合わせると 0 になることからわかる．■

Hamming 符号の最適性

命題 39 Hamming 符号は最小距離 3 の二元符号の中でレートに関して最適な符号である．

補題 40 符号長 n，最小距離 3 の二元符号 C に対して，

$$|C| \leq \frac{2^n}{n+1}.$$

証明: 符号語 $c \in C$ に対して，$N(c) = \{y \in \mathbb{F}_2^n : d(y, c) \leq 1\}$ と定義する．C は最小距離が 3 なので，任意の $c \neq c' \in C$ に対して，$N(c) \cap N(c') = \emptyset$ である．したがって，

$$2^n \geq \left| \bigcup_{c \in C} N(c) \right| = \sum_{c \in C} |N(c)| = |C| \cdot (n+1).$$

■

証明 (命題 39): $C_{\mathrm{Ham}}^{(r)}$ は，補題 40 の不等式の等号を，符号長 $2^r - 1$ の場合に対して満たしている．つまり，符号長 $2^r - 1$，最小距離 3 の符号の中で，最大の符号語数をもつ符号である．■

定理 41 (Hamming 限界，球充填限界 (shpere-packing bound)) 符号長 n，最小距離 d の二元符号 C に対して，

$$|C| \leq \frac{2^n}{\sum_{i=0}^{\lfloor \frac{d-1}{2} \rfloor} \binom{n}{i}}.$$

Hamming 限界を達成する符号は**完全符号**と呼ばれる．二元線形符号で完全符号なのは，Hamming 符号，$(n, 1, n)_2$ 符号である $C = \{0 \ldots 0, 1 \ldots 1\}$ （ただし n は奇数）と，$(23, 12, 7)_2$ Golay 符号だけであることが示されている．

3.6 双対符号

定義 42 (双対符号) 線形符号 $C \subseteq \mathbb{F}_q^n$ の双対符号 $C^\perp$ とは,

$$C^\perp = \{c \in \mathbb{F}_q^n : \text{すべての } c' \in C \text{ に対して } \langle c, c' \rangle = 0\}.$$

つまり，C の直交補空間が C の双対符号である．

命題 43 C が $(n,k)_q$ 線形符号のとき，以下が成り立つ．

1. $C^\perp$ は $(n, n-k)_q$ 線形符号である．
2. $(C^\perp)^\perp = C$.
3. H が C のパリティ検査行列のとき，H は $C^\perp$ の生成行列である．また，G が C の生成行列のとき，G は $C^\perp$ のパリティ検査行列である．

Hamming 符号の双対符号は，$(2^r-1, r)_2$ 符号であり，シンプレックス符号と呼ばれる．シンプレックス符号の生成行列の列には，長さ r の非零ベクトルがすべて現れる．生成行列に零ベクトルを含めたとき，$(2^n, r)_2$ 符号となり，Hadamard 符号と呼ばれる．

Hamming 符号は最適なレートをもつことを示したが，相対最小距離は $3/2^r$ であり，符号長を大きくすると 0 に近づく．一方，Hadamard 符号やシンプレックス符号はレートが $r/2^r$ であり，符号長を大きくすると 0 に近づく．しかし，大きな最小距離をもつことがわかる．

命題 44 Hadamard 符号とシンプレックス符号の相対最小距離は $1/2$ である．

3.7 符号族，漸近的によい符号

定義 45 (符号族) $C = \{C_i : i \in \mathbb{N}\}$ に対して，各 C_i が $(n_i, k_i, d_i)_{q_i}$ 符号であり，$n_i > n_{i-1}$, $q_i \geq q_{i-1}$ であるとき，C を符号族という．C のレートは,

$$R(C) = \liminf_{i\to\infty} \left\{ \frac{k_i}{n_i} \right\}.$$

C の相対最小距離は,

$$\delta(C) = \liminf_{i\to\infty} \left\{ \frac{d_i}{n_i} \right\}.$$

Hamming 符号は最適なレートをもつが相対最小距離が小さい．Hadamard 符号やシンプレックス符号はレートは小さいが相対最小距離は大きい．では，レートと相対最小距離がどちらもよい符号は存在するだろうか．つまり，レートがある正の実数定数 $R > 0$ 以上，相対最小距離がある正の実数定数 $\delta > 0$ 以上となる符号族は存在するだろうか．レートと相対最小距離をともに大きくすることは矛盾する要求であり，これらの間にはトレードオフが存在する．どのような (R, δ) の組合せならば存在し，また存在しないのか．さらに，現実に応用することを考えた場合，その符号を使った誤り訂正は効率的にできるのか．以上の点について，今後の授業で議論する．