情報理論第二 (9)
線形符号

人間コミュニケーション学科
梶本　裕之
kajimoto@hc.uec.ac.jp

授業の流れ（予定）

第１週（10/ 8） 情報量と情報エントロピー
第２週（10/15） 相互情報量
第３週（10/22） 体育祭のため休講
第４週（10/29） 情報源符号化とデータ圧縮
第５週（11/ 5） ハフマン符号とデータ圧縮
第６週（11/12） 情報源符号化定理
第７週（11/19） 調布祭のため休講
第８週（11/26） 出張のため休講
第９週（12/ 3） マルコフ情報源モデル
第10週（12/10） 通信路のモデル化
第11週（12/17） 誤り検出・誤り訂正符号
第12週（12/24） 出張のため休講
第13週（ 1/ 7） 線形符号
第14週（ 1/14） ハミング符号
第15週（ 1/21） 秘密鍵暗号
第16週（ 1/28） 公開鍵暗号
第17週（ 2/ 4） 出張のため休講

(前回の復習)ハミング距離

- 同じ長さを持つ２つのビット列がどの程度はなれているかをあらわす量

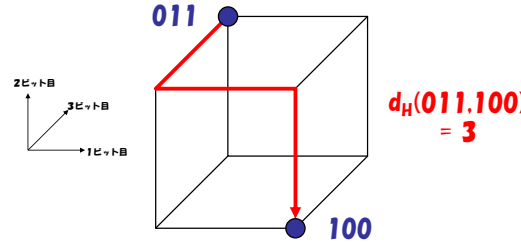
  $a = a_1a_2\ldots a_n$ ， $b = b_1b_2\ldots b_n$ について

$$d_H(a, b) = \Sigma_{i=1\sim n} | a_i - b_i |$$

  a と b の各桁を比較したときに異なっているビットの数

(前回の復習)ハミング距離のイメージ

- ３ビットの場合：空間で考えると

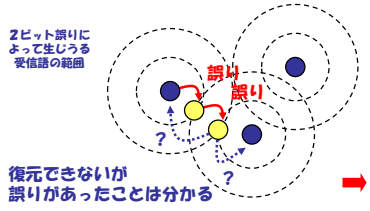


$d_H(011,100) = 3$

ハミング距離＝ｎ次元立方体上で何回で到達できるか

(前回の復習)誤り検出

- 誤りによって広がった各符号語の受信語の範囲が他の符号語を含まない限り，何らかの誤りが生じたことを検出できる

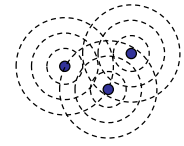


➡ 再送信を要求できる

ハミング距離と符号

- 一般に，各符号語間のハミング距離が最低でもｋと保証されているならば，その符号は最低でも
  - ｋより少ない誤りについて検出可能
  - ｋ/2 より少ない誤りについて訂正可能

となる



前回の小レポート(1)

- ハミング距離の間に三角不等式

$$f(a, b) + f(b, c) \geqq f(a, c)$$

が成立する理由を考えよ.

7

前回の小レポート(1)：回答例

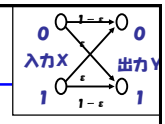
- $f(a, b)$: a, b 間で異なっているビットの個数.
- $f(b, c)$: b, c 間で異なっているビットの個数.
- $f(a, c)$: a, c 間で異なっているビットの個数.
- ある位 i で考えると，a_i, b_i, c_iの組み合わせは (0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1),(1,1,0),(1,1,1)の8通り.

	000	001	010	011	100	101	110	111
$f_i(a_i,b_i)$	0	0	1	1	1	1	0	0
$f_i(b_i,c_i)$	0	1	1	0	0	1	1	0
$f_i(a_i,c_i)$	0	1	0	1	1	0	1	0

- 全ての組み合わせで三角不等式が成立しているから，その合計である f で三角不等式が成り立つのは明らか

8

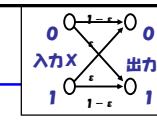
前回の小レポート(2)



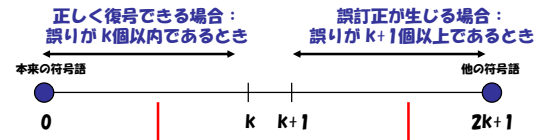
- 1ビットごとの誤り発生率がεであるような2元対称通信路でnビットの符号語を送る.
- この符号語間の最小距離が 2k+1 であった場合,
  - 実際に正しく復号できる確率,
  - 誤訂正が生じる確率
- を，それぞれ式で表せ.

9

前回の小レポート(2) 回答例



- n個の符号でm個誤る確率は ${}_nC_m\,\varepsilon^m(1-\varepsilon)^{n-m}$
- 符号語間の最小距離が 2k+1 なので,



$$\Sigma_{i=0\sim k}\ {}_nC_i\,\varepsilon^i(1-\varepsilon)^{n-i}$$

$$\Sigma_{i=k+1\sim n}\ {}_nC_i\,\varepsilon^i(1-\varepsilon)^{n-i}$$

10

今日と次回のお題

- 符号語間の最小距離が常に3（＝2誤り検出・1誤り訂正可能）となるような符号を手続き的に生成する手法を紹介
  - 線形符号（背景的知識）
  - ハミング符号（線形符号の一種）

    情報源符号化におけるハフマン符号ほど本質的に重要なものではないが，符号設計の古典例として．．．

11

数学的準備：有限体

有限体

- 我々の知っている「整数」は無限小から無限大に続く無限集合である.

…,-5,-4,-3,-2,-1,0,1,2,3,4,5,…

- 「数」が有限個しかない世界を考えることも可能.
- (例) $F_3=\{0,1,2\}$
- 集合内で閉じた「 ＋ー×÷ 」を定義しよう

13

Mod(あまり) を使う

- ＋ー×÷ の結果も同じ集合の中に入っていなければならない→Mod（剰余）を使う
- (例) $F_3=\{0,1,2\}$: 0,1,2のみからなる世界
- 0+0=0, 0+1=1, 1x2=2,…
- 1+2=3=0 mod(3)
- 1−2 = (x + 2 = 1 (mod 3) となる x) = 2
- 2×2 = 4 = 1 (mod 3)
- 1÷2 = (x×2 = 1 (mod 3) となる x) = 2

14

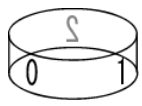
イメージ

- 無限の世界

-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5,

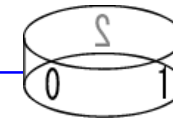
- Modを使った繰り返しの世界

0, 1, 2, 0, 1, 2, 0, 1, 2, 0, 1, 2, 0, 1, 2,



15

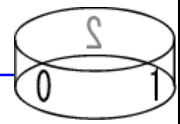
有限体



- 一般に q が素数の時,
- {0, 1, 2, … q−1} という集合は,
- mod q の条件下で
  要素数 q の有限体を形成
- (素数でなければダメ)

16

有限体：一般的な定義



- 「加減乗除」について閉じた
- 有限個の要素からなる集合

を有限体（ガロア体）と呼ぶ

- 「閉じた」とは，演算の結果が同じ集合の要素になること
- 集合の要素数が3の場合，この集合を F_3, GF(3) などと書く

17

参考

体：四則演算(＋－×÷)が定義できる集合.

整数は割り算で整数にならない場合がある
→Not 体.

1÷3=1/3→整数ではない

有理数は割り算をしても有理数→体.

18

例題

- {0, 1, 2} が mod 3 の条件下で四則演算について閉じていることを確認せよ.

（＋－×は明らか. ÷に関しては任意の要素 $x \neq 0$ について $1 \div x$ が定義できることを確認すれば十分. ）

- 上記は {0, 1, 2, 3} (mod 4) には当てはまらないことを確認せよ.

例題

- F_3 {0, 1, 2}
  - $1 \div 1 =$ [ ] $=$ [ ]
  - $1 \div 2 =$ [ ] $=$ [ ]
- F_4 {0, 1, 2, 3}
  - $1 \div 1 =$ [ ] $=$ [ ]
  - $1 \div 2 =$ [ ] $=$ [ ]
  - $1 \div 3 =$ [ ] $=$ [ ]

さらに詳しく知りたい人に

- 数学ガール
- 数学ガール2「フェルマーの最終定理」
- 暗号理論の基礎知識でもある。

記号 {0, 1} を有限体として考える

- これまで単なる記号だった {0, 1} という符号用の文字を整数値として考え直し，この中で閉じた四則演算体系を定義

$0+0=0$, $0-0=0$, $0\times0=0$
$0+1=1$, $0-1=$ [ ], $0\times1=0$, $0\div1=0$
$1+0=1$, $1-0=1$, $1\times0=0$
$1+1=$ [ ], $1-1=0$, $1\times1=1$, $1\div1=1$

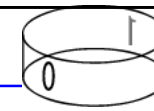
有限体を使うことのメリット

- $F_2 = \{0, 1\}$ を用いることにより
  符号化と復号化の過程を
  代数的な演算として簡潔に書ける

例：nビットの符号語の空間 F_2^n
ある符号語 $x \in F_2^n$
誤りベクトル $e \in F_2^n$
（誤りが起きた場所が1，他は0のベクトル）
実際の受信語 $y = x + e$

この授業ではこれ以降全て F_2 上の演算であると想定します

例：6ビット符号語と誤り



受信語 y　　元の符号語 x　　生じた誤り e

$$\begin{pmatrix}1\\0\\1\\0\\1\\1\end{pmatrix} = \begin{pmatrix}1\\0\\1\\1\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\0\\1\\0\\1\end{pmatrix}$$

注意

$$y = x + e$$

線形符号

線形符号

- 元の語 x ：kビット　　$x \in F_2^k$
- 符号語 y ：nビット　　$y \in F_2^n$
  $(n > k)$.

ある適当な生成行列G（n行k列）を用いて

$$y = Gx$$

と符号化する方法を線形符号という

26

線形符号の性質

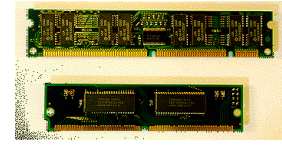
$$y = Gx$$

- 復号可能 $\Leftrightarrow$ xとyは1対1対応
- 0ベクトルは常に0ベクトルに符号化
- 元の語 x が空間F_2^kの全てにわたるとき，符号語 y の分布によって作られる空間はF_2^n内のk次元部分空間になる

27

例：（いわゆる）パリティビット方式

- 元の語3ビットに対して，
- そこに含まれる1が偶数個ならば0，
- 奇数個ならば1
- をパリティビットとして末尾に1桁付加するような符号の生成行列Gとは？



28

例：（いわゆる）パリティビット方式

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X \\ Y \\ Z \\ P \end{pmatrix}$$

生成行列G　元の語　符号語

X=X
Y=Y
Z=Z
P=X+Y+Z

P：XYZの「1の個数」が偶数なら0，奇数なら1

29

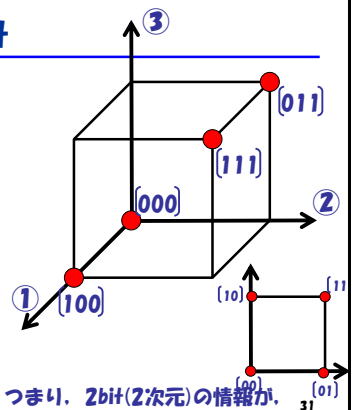
例

- 2ビットの語を3ビットの符号語に変換する線形符号を考える．
- Gの内容が以下のようであるとき，元の語と符号語との対応関係は？

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

30

例：2bit⇒3bit

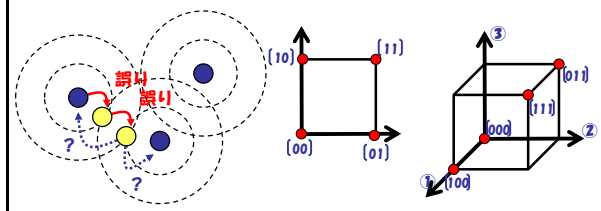


つまり，2bit(2次元)の情報が，3bit(3次元）空間に分散した．

31

例：2bit⇒3bit

元の語：2bit(2次元)の情報．４通り
符号語：3bit(3次元）．8つの格子点の中の４点
⇒できた「すき間」＝誤り検出，誤り訂正のための冗長性
つまり，この「すき間」の最小の間隔が重要！！



ただしこの例では，最小すき間が1しかないので誤り検出も不可能

32

生成行列の標準形

- 生成行列は，符号語空間の基底ベクトルの取り方を変えたり座標軸の並び順を換えたりすることによって，以下のような標準形に変形できることが知られている

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

k 行

n-k 行

33

標準形の生成行列による符号化

- 標準形をした生成行列:
- 元の語と付加されたパリティ部分とがきれいに分かれた符号語を生成できる！

符号語 $\begin{pmatrix} x \\ \hline Px \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & \cdots & 1 \\ \hline & P & \end{pmatrix}$ 元の語 $\begin{pmatrix} x \end{pmatrix}$

34

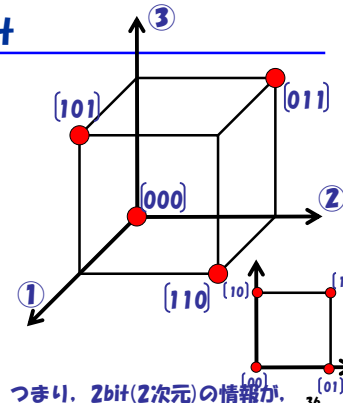
例：2bit->3bitの（いわゆる）パリティビット方式

$$\underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}}_{\text{生成行列}G} \underbrace{\begin{pmatrix} X \\ Y \end{pmatrix}}_{\text{元の語}} = \underbrace{\begin{pmatrix} X \\ Y \\ P \end{pmatrix}}_{\text{符号語}}$$

X=X
Y=Y
P=X+Y
P：XYZの「1の個数」が偶数なら0，奇数なら1

35

例：2bit⇒3bit

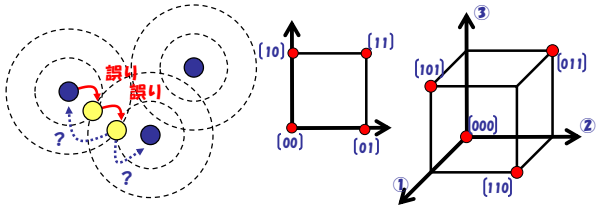


つまり，2bit(2次元)の情報が，3bit(3次元）空間に分散した．

36

2bit⇒3bitのいわゆるパリティ方式

元の語：2bit(2次元)の情報．4通り
符号語：3bit(3次元)．8つの格子点の中の4点
⇒できた「すき間」＝誤り検出，誤り訂正のための冗長性
つまり，この「すき間」の最小の間隔が重要！！



この例では，最小すき間が2あるので誤り検出可能

例題

・生成行列Gの内容が以下のような標準形であるとき，元の語と符号語との対応関係を求めよ．

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

例題

$$G = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$ ：単位行列

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} =$$

例題

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \quad (0\,0\,0) \Rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \quad (0\,0\,1) \Rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \quad (0\,1\,0) \Rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \quad (0\,1\,1) \Rightarrow$$

例題

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \quad (1\,0\,0) \Rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \quad (1\,0\,1) \Rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = \quad (1\,1\,0) \Rightarrow$$

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \quad (1\,1\,1) \Rightarrow$$

検査行列

・下記の生成行列Gに対して，
右のような行列Hをつくることができる

$$G = \begin{pmatrix} 1 & & 0 \\ & 1 & \\ 0 & \cdots & 1 \\ \hline & P & \end{pmatrix}$$ （k列，k行，n-k行）

$$H = \begin{pmatrix} -P & \begin{matrix} 1 & 0 \\ & \ddots \\ 0 & 1 \end{matrix} \end{pmatrix}$$ （k列，n-k列，n-k行）

検査行列という

検査行列の性質

- 生成行列Gによって作られた符号語をHに適用すると，結果が0になる！！

$$H(Gx) = (HG)x = 0x = 0$$

$$\begin{pmatrix} -P & \begin{matrix} 1 & 0 \\ & \ddots \\ 0 & 1 \end{matrix} \end{pmatrix} \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \\ \hline & P & \end{pmatrix} \begin{pmatrix} \\ x \\ \\ \end{pmatrix} = (-P+P)x$$

受信語にHをかけてみて0になるかどうかで
パリティ部分に誤りが生じたかどうかを検査できる

例題

- 以下の生成行列Gに対応する検査行列Hをつくり，HGが確かに0となることを確認せよ．（{0, 1}の有限体では－P＝Pとなることに注意）

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

44

例題

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

$$\begin{pmatrix} & \\ & \end{pmatrix} \begin{pmatrix} & \\ & \end{pmatrix} = \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix}$$

45

例題

- 元の語が[100]の場合，
- （1）生成行列Gによって生じる符号語は？
- （2）検査行列Hで0になるかどうか確認せよ
- （3）1bitのエラーが生じたとき，検査行列でエラーを検出できるかどうか．

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

46

例題

$$Gx = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} \\ \\ \\ \\ \\ \\ \end{pmatrix} \quad \cdots \text{符号語}$$

$$H(Gx) = \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \\ \\ \\ \\ \\ \\ \end{pmatrix} = \begin{pmatrix} \\ \\ \end{pmatrix}$$

エラーなし

例題

1bitエラーが生じたとする

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$ 検査の結果0にならない

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$ 検査の結果0にならない

48

レポート課題

- 元の語が[010]の場合，
- （1）生成行列Gによって生じる符号語は？
- （2）検査行列Hで0になるかどうか確認せよ
- （3）1bitのエラーが生じたとき，検査行列でエラーを検出できるかどうか，すべて（6とおり）調べよ．

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

49