

代数系と符号理論 (O) 第1回板書プリント

制作：植松友彦

2014.4.8

0 この講義の約束

0.1 禁止事項

授業中の食事は禁止する !!

0.2 成績の評価法

- 中間テスト (40 点) と期末テスト (40 点) を行う。
- 演習を提出する毎に、出席点を 2 点を与える。
- 3 回程度の宿題を出すので、期限までに必ず提出すること。A,B,C で評価し点数を与える。
 m 回分の宿題を期限までに提出しなかった場合、合計得点の 60 点を超える部分を 0.5^m 倍する。例えば、テストと出席と宿題で 100 点であったとしても、2 回提出物を出さなかった場合、成績は

$$(100 - 60) * (0.5)^2 + 60 = 70$$

となる。良い点が欲しければ、演習に出席し、宿題を期限までに提出すること。

0.3 教科書と参考文献

教科書: 植松 友彦著：代数系と符号理論、オーム社、2010.

- 参考書:
- 坂庭、渋谷: 代数系と符号理論入門、コロナ社、2010.
 - 今井: 符号理論、電子情報通信学会、1990.
 - 岩垂: 符号理論入門、昭晃堂、1992. (易しい)
 - 平澤、西島：符号理論入門、培風館、2000.
 - J. Justesen and T. Hoholdt：誤り訂正符号入門、森北出版、2005.
 - 和田山: 誤り訂正技術の基礎、森北出版、2010.
 - 情報理論とその応用学会編: 符号理論とその応用、培風館、2003.
 - W. W. Peterson and E. J. Weldon, Jr.: Error-Correcting Codes, MIT Press, 1971.
 - R. E. Blahut: Algebraic Codes for Data Transmission, Cambridge, 2003.
 - S. Lin and D. J. Costello, Jr. : Error Control Coding, Peason Prentice Hall, 2004.

1 線形代数と線形符号

1.1 符号理論とは

通信や記録を行う情報に冗長性を付け加えることで、情報通信システムや記録装置で生じた誤りを検出あるいは訂正し、信頼性を高める技術について研究する学問。

符号理論の広がり

- 情報理論 (通信の限界)
- 通信理論 (変復調方式、物理系)
- 数学 (代数学、確率、組み合わせ論)
- 計算機科学 (アルゴリズム、計算量)
- 計算機ハードウェア (RAID, メモリ構成)

符号理論の歴史

デジタル情報の様々な通信・記録システムにおいて誤り訂正符号が用いられている。

単一誤り訂正符号	1950	ハミング符号 (唯一個の誤りを訂正できる符号)
複数個の誤りを を訂正できる符号	1955	たたみ込み符号 (Elias)
	1959	BCH(Bose-Chaudhuri-Hocquenghem) 符号
	1960	RS(Reed-Solomon) 符号
	1963	Low Density Parity Check(LDPC) 符号 (Gallager)
衛星通信 への応用	1971	Marier 9 号 (火星探査衛星) 双直交符号
	1975	Viking(火星探査衛星) 双直交符号
	1977	Voyager(太陽系探査衛星)
		RS 符号+たたみ込み符号 (NASA の標準技術)
民生機 への応用	1980	音楽用 CD (RS 符号の実用化)
	1982	符号化変調方式 (Ungerboeck)
	1985	TV 文字放送 (差集合巡回符号)
	1988 頃	各種通信用モデム
	1992	MD (RS 符号)
	1993	デジタル携帯電話
	1996	CS デジタル放送 (RS 符号)
	1996	DVD (RS 符号)
	1999	1000Base-T (符号化変調方式)
	2004	10GBase-T (LDPC 符号)
新たな符号の発見	...	
	1993	Turbo 符号 (Berrou, Glavieux ら)
	1996	LDPC 符号の再発見 (MacKay と Neal)
	2009	Polar 符号 (Arikan)

日本における符号理論の研究の歴史については、下記の文献を参照されたい。
今井、萩原: “日本における符号理論の原点”, 電子情報通信学会 Fundamentals Review, vol.2, no.4, pp.9-15 (2009).

(http://w2.gakkai-web.net/gakkai/ieice/vol2no4pdf/vol2no4_09.pdf からダウンロード可能)

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い元の情報を復元する操作 : 復号化 (復号)

$$\begin{aligned} \text{情報 } \mathbf{u} \in \{0, 1\}^k &\longrightarrow \text{符号化} \longrightarrow \text{符号語 } \mathbf{v} \in \{0, 1\}^n \quad (n \geq k) \\ \text{符号語 } \mathbf{v} &\longrightarrow \text{通信路} \longrightarrow \text{受信語 } \mathbf{w} = \mathbf{v} \oplus \mathbf{e} \\ &\mathbf{e} \in \{0, 1\}^n : \text{誤りベクトル} \\ &\quad (\text{ランダムに発生, } 1 \text{ が誤りに対応}) \\ &\quad \mathbf{e} = 0 : \text{誤りなしに対応} \\ \text{受信語 } \mathbf{w} = \mathbf{v} \oplus \mathbf{e} &\longrightarrow \text{復号} \longrightarrow \mathbf{u} \in \{0, 1\}^k : \text{送信情報の推定値} \end{aligned}$$

但し、 $\oplus$ はビット毎の mod2 加算を表す。

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して設計できるか (設計法)
3. そのような符号はどのようにして効率的に復号できるか (復号法)

情報工学科では、これらの研究は、植松研、松本研、坂庭研、金子研、笠井研で行われている。

1.2 線形代数の復習

(a) ベクトル空間

集合 V が次の 2 条件を満足するときベクトル空間と呼ぶ。

(I) V の 2 元 $\mathbf{x}, \mathbf{y}$ に対して和と呼ばれる第三の元 $(\mathbf{x} + \mathbf{y})$ が定まり、次の法則が成り立つ。

- (1) $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$
- (2) $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$
- (3) ゼロベクトルと呼ばれる特別な元 0 がただ一つ存在し、 V の全ての元 $\mathbf{x}$ に対して、 $\mathbf{x} + 0 = \mathbf{x}$ が成り立つ。
- (4) V の全ての元 $\mathbf{x}$ に対して、 $\mathbf{x} + \mathbf{x}' = 0$ となる V の元 $\mathbf{x}'$ がただ一つ存在する。

(II) V の任意の元と任意のスカラー a に対し、 $\boldsymbol{x}$ の a 倍とよばれる V の元が定まり、次の法則が成り立つ。

$$(5) (a+b)\boldsymbol{x} = a\boldsymbol{x} + b\boldsymbol{x}$$

$$(6) a(\boldsymbol{x} + \boldsymbol{y}) = a\boldsymbol{x} + a\boldsymbol{y}$$

$$(7) (ab)\boldsymbol{x} = a(b\boldsymbol{x})$$

$$(8) 1\boldsymbol{x} = \boldsymbol{x} \quad \text{但し、} 1 \text{ はスカラーの乗法単位元である。}$$

線形代数の授業で習ったように、 n 次元の実ベクトルの集合は実数をスカラーとしてベクトル空間をなしている。

(b) ベクトル空間の基底

ベクトル空間 V の有限個のベクトル $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ が次の 2 条件を満足するとき、 $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ は V の基底であるという。また基底ベクトルの個数 k をベクトル空間 V の次元という。

(1) $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ は線形独立である。すなわち、

$$a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \dots + a_k\boldsymbol{e}_k = \mathbf{0}$$

となるようなスカラーの組 $a_1, a_2, \dots, a_k$ は $a_1 = a_2 = \dots = a_k = 0$ に限る。

(2) V の任意のベクトルは、 $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_k$ の線形結合として書ける。すなわち、任意の $\boldsymbol{x} \in V$ に対して、あるスカラーの組 $a_1, a_2, \dots, a_k$ が存在して、

$$\boldsymbol{x} = a_1\boldsymbol{e}_1 + a_2\boldsymbol{e}_2 + \dots + a_k\boldsymbol{e}_k$$

が成り立つ。

(c) 部分空間

ベクトル空間 V の部分集合 W が、 V と同じ演算に対して、ベクトル空間になるとき、 W を V の線形部分空間あるいは部分空間と呼ぶ。 W が V の部分空間であるための必要十分条件は、次の 2 条件である。

(1) 任意の $\boldsymbol{x}, \boldsymbol{y} \in W \implies \boldsymbol{x} + \boldsymbol{y} \in W$

(2) 任意のスカラー a と任意の $\boldsymbol{x} \in W \implies a\boldsymbol{x} \in W$

1.3 有限体上のベクトル空間

四則演算が定義できる数の集合を「体 (たい)」という。有理数の集合や実数の集合は体である。また、集合の要素数が有限個の体のことを「有限体」という。

2 元体 F_2

0 と 1 の元から成り、次のように加算と乗算を定めたものを 2 元体という。

$$\begin{aligned} \text{(加算)} \quad 0+0 &= 0 \\ 0+1 &= 1+0=1 \\ 1+1 &= 0 \\ \text{(乗算)} \quad 0\cdot 0 &= 0 \\ 0\cdot 1 &= 1\cdot 0=0 \\ 1\cdot 1 &= 1 \end{aligned}$$

尚、2 元体では 1 の加法逆元 (-1 によって表す) は $1+1=0$ から 1 であるので、 $-1=1$ であることに注意しておく。

n 次元ベクトル空間 F^n

体 F に対して、集合

$$F^n = \{(x_0, x_1, \dots, x_{n-1}) : x_i \in F, i = 0, 1, \dots, n-1\}$$

によって定める。このとき、 F^n の元 $\mathbf{x} = (x_0, x_1, \dots, x_{n-1})$ と $\mathbf{y} = (y_0, y_1, \dots, y_{n-1})$ に対して、和を

$$\mathbf{x} + \mathbf{y} = (x_0 + y_0, x_1 + y_1, \dots, x_{n-1} + y_{n-1})$$

によって、また $a \in F$ に対して $\mathbf{x}$ の a 倍を

$$a\mathbf{x} = (ax_0, ax_1, \dots, ax_{n-1})$$

によって定める。このとき、 F^n は n 次元ベクトル空間となることは容易に分る (各自やってみる)。以下、この授業ではベクトル空間として、 F^n のみを取り扱う。

1.4 ブロック符号

[定義 1.1](ブロック符号) 体 F 上のブロック符号 C とは、 n 次元ベクトル空間 F^n の部分集合

$$C = \{\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_M\} \subset F^n$$

であり、 $\mathbf{c}_i$ ($i = 1, 2, \dots, M$) を「符号語 (codeword)」といい、 n を符号長 (code length) という。特に 2 元体上のブロック符号のことを「2 元符号 (binary code)」という。また、

$$R = \frac{1}{n} \log_q M$$

を符号の「伝送速度 (rate)」という。但し、 q は有限体 F の元の個数である。

[定義 1.2](線形符号) 有限体 F 上の (n, k) 線形符号 (linear code) C とは、ベクトル空間 F^n の k 次元部分空間である。また、線形符号の伝送速度 R は、 $R = k/n$ である。

線形符号の性質

- (1) 任意の 2 つの符号語 c_i, c_j について、 $c_i + c_j$ は符号語である。
- (2) 任意の符号語 c_i とスカラー $a \in F$ について、 ac_i は符号語である。
- (3) 零ベクトルは常に符号語である。
- (4) 符号語の総数は $M = q^k$ に等しい。但し、 q は体 F の元の個数である。

[例 1.3] 符号長が 4 であり、 F_2^4 の 4 個のベクトルによって定まる次のブロック符号 C_1 を考えよう。

$$C_1 = \{(1000), (0100), (0010), (0001)\}$$

これは線形符号ではない。なぜなら、 $(1000) + (0100) = (1100)$ は符号語ではないからである。

[例 1.4] 要素として F_2 の元をもつ 4 個のベクトルによって定まる符号 C を考えよう。

$$C_2 = \{(0000), (1100), (0011), (1111)\}$$

どの 2 つの符号語の和も符号語になっている。スカラーは 0 と 1 なので、明らかに任意の符号語のスカラー倍は零ベクトルかその符号語自身である。従って、この符号は線形符号である。

線形符号を特定するには、すべての符号語のリストを与える代わりに、 k 個の 1 次独立な符号語からなる基底を与えることで定めることができる。このとき、任意の符号語はこれら k 個の 1 次独立な符号語の線形結合として表すことができる。

[例 1.5] 例 1.4 の線形符号の基底は、 $\{(1100), (0011)\}$ である。任意の符号語はこの基底の線形結合で表すことができる。事実、残りの 2 つの符号語は

$$\begin{aligned}(0000) &= 0(1100) + 0(0011) \\ (1111) &= 1(1100) + 1(0011)\end{aligned}$$

によって表される。尚、基底の選び方は 1 通りには限らず、4 つのベクトル $\{(0000), (1100), (0011), (1111)\}$ の中から 1 次独立な 2 つのベクトルを選べば基底になり、上記以外に $\{(0011), (1111)\}$ ならびに $\{(1100), (1111)\}$ も基底である。

今後の日程

- 4 月 9 日 (水) 講義 2
- 4 月 16 日 (水) 演習 1
- 4 月 22 日 (火) 講義 3
- 4 月 23 日 (水) 講義 4
- 4 月 30 日 (水) 演習 2