# 数学 II 演習 ( 第 11 回 ) の略解

## 目 次


1 問 1 の解答 1

2 $n$ 行 $n$ 列のときにはどうなるのか 5

3 問 1 を見直すと 7

4 問 2 の解答 12

5 問 2 を見直すと 14

6 「行列の標準形の問題」との関係 20

7 最小多項式とは 24

8 Cayley-Hamilton の定理について 32

9 一変数多項式環のイデアルについて * 41

10 何故, イデアルの条件は自然なのか * 47


## 1 問 1 の解答

(1) 第 1 回の問 2 のときと同様に,

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

とすると, 行列 $A$ は,

$$A = \lambda I + N \tag{1}$$

というように表わせることが分かります. このとき, $N^2, N^3$ などを具体的に計算してみると,

$$N^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \ N^3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

となることが分かりますから,

$$N^k = 0, \quad (k \geq 3) \tag{2}$$

となることが分かります. そこで, $\lambda I$ と $N$ は可換であることに注意して, $(\lambda I + N)^n$ を二項展開してみると, (1) 式, (2) 式から,

$$\begin{aligned}
A^n &= (\lambda I + N)^n && \text{( (1) 式から )}\\
&= \lambda^n I + n\lambda^{n-1}N + \frac{n(n-1)}{2}\lambda^{n-2}N^2 + \cdots + N^n\\
&= \lambda^n I + n\lambda^{n-1}N + \frac{n(n-1)}{2}\lambda^{n-2}N^2 && \text{( (2) 式から )}\\
&= \lambda^n \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + n\lambda^{n-1} \cdot \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\\
&\qquad + \frac{n(n-1)}{2}\lambda^{n-2} \cdot \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\\
&= \begin{pmatrix} \lambda^n & n\lambda^{n-1} & \frac{n(n-1)}{2}\lambda^{n-2} \\ 0 & \lambda^n & n\lambda^{n-1} \\ 0 & 0 & \lambda^n \end{pmatrix}
\end{aligned}$$

となることが分かります.[1]

(2) (1) の結果から,

$$\begin{cases} A = \lambda I + N \\ A^2 = \lambda^2 I + 2\lambda N + N^2 \\ A^3 = \lambda^3 I + 3\lambda^2 N + 3\lambda N^2 \end{cases}$$

と表わせることが分かりますから, $f(A)$ は,

$$\begin{aligned}
f(A) &= aI + bA + cA^2 + dA^3\\
&= aI + b(\lambda I + N) + c(\lambda^2 I + 2\lambda N + N^2) + d(\lambda^3 I + 3\lambda^2 N + 3\lambda N^2)\\
&= (a + b\lambda + c\lambda^2 + d\lambda^3)I + (b + 2c\lambda + 3d\lambda^2)N + (c + 3d\lambda)N^2\\
&= \begin{pmatrix} a + b\lambda + c\lambda^2 + d\lambda^3 & b + 2c\lambda + 3d\lambda^2 & c + 3d\lambda \\ 0 & a + b\lambda + c\lambda^2 + d\lambda^3 & b + 2c\lambda + 3d\lambda^2 \\ 0 & 0 & a + b\lambda + c\lambda^2 + d\lambda^3 \end{pmatrix} \qquad (3)
\end{aligned}$$

となることが分かります. したがって,

$$f(A) = O \iff \begin{cases} a + b\lambda + c\lambda^2 + d\lambda^3 = 0 \\ b + 2c\lambda + 3d\lambda^2 = 0 \\ c + 3d\lambda = 0 \end{cases}$$

[1] もちろん, 上のように二項展開を用いて $A^n$ を計算しなければいけないということではありません. 例えば, $A^2, A^3, \cdots$ を具体的に計算してみることで $A^n$ の形を予想して, その予想が正しいことを数学的帰納法で証明するという方針でも構いません.

となることが分かります.

(3) 一般の多項式の場合を考えてみる前に, 様子を探ってみるために,

$$f(x) = a + bx + cx^2 + dx^3$$

として, (3) 式の右辺も $f(x)$ を用いて表わすことを考えてみます. すると,

$$\begin{cases} f(\lambda) = a + b\lambda + c\lambda^2 + d\lambda^3 \\ f'(\lambda) = b + 2c\lambda + 3d\lambda^2 \\ f''(\lambda) = 2c + 6d\lambda \end{cases}$$

となることが分かりますから, (3) 式は,

$$\begin{aligned} f(A) &= \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) \\ 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & f(\lambda) \end{pmatrix} \\ &= f(\lambda)I + f'(\lambda)N + \frac{1}{2}f''(\lambda)N^2 \end{aligned} \tag{4}$$

というように表わせることが分かります. すると, (4) 式から, 行列 $f(A)$ を多項式 $f(x)$ を用いて表わすには, $f(A)$ を $N$ のベキの形に整理して, それぞれの係数が $f(x)$ を用いてどのように表わせるのかということを考えてみると良さそうなことが分かります.

そこで, 一般の多項式

$$\begin{aligned} f(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ &= \sum_{k=0}^{n} a_k x^k \end{aligned}$$

の場合を考えてみます. いま (1) の結果から, 勝手な自然数 $k \in \mathbb{N}$ に対して,

$$A^k = \lambda^k I + k\lambda^{k-1}N + \frac{k(k-1)}{2}\lambda^{k-2}N^2 \tag{5}$$

というように表わせることが分かります. よって, (5) 式から,

$$\begin{aligned} f(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= \sum_{k=0}^{n} a_k A^k \\ &= \sum_{k=0}^{n} a_k \left\{ \lambda^k I + k\lambda^{k-1}N + \frac{k(k-1)}{2}\lambda^{k-2}N^2 \right\} \qquad (\text{ (5) 式から }) \\ &= \left\{ \sum_{k=0}^{n} a_k \lambda^k \right\} \cdot I + \left\{ \sum_{k=0}^{n} k a_k \lambda^{k-1} \right\} \cdot N \\ &\quad + \frac{1}{2}\left\{ \sum_{k=0}^{n} k(k-1) a_k \lambda^{k-2} \right\} \cdot N^2 \end{aligned} \tag{6}$$

というように表わせることが分かります. ここで,

$$\begin{cases} f(\lambda) = \sum_{k=0}^{n} a_k \lambda^k \\ f'(\lambda) = \sum_{k=0}^{n} k a_k \lambda^{k-1} \\ f''(\lambda) = \sum_{k=0}^{n} k(k-1) a_k \lambda^{k-2} \end{cases} \tag{7}$$

となることに注意すると, (6) 式, (7) 式から,

$$\begin{aligned} f(A) &= f(\lambda)I + f'(\lambda)N + \frac{1}{2}f''(\lambda)N^2 \\ &= \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) \\ 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & f(\lambda) \end{pmatrix} \end{aligned} \tag{8}$$

となることが分かります. したがって, (8) 式から,

$$f(A) = O \iff f(\lambda) = f'(\lambda) = f''(\lambda) = 0$$

となることが分かります.

(4) (1) と同様に,

$$N' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

とすると, 行列 $B$ は,

$$B = \lambda I + N' \tag{9}$$

というように表わせることが分かります. このとき, $(N')^2$ を具体的に計算してみると,

$$(N')^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{10}$$

となることが分かるので,

$$(N')^k = 0, \quad (k \geq 2)$$

となることが分かります. そこで, (1) と同様に, $\lambda I$ と $N'$ は可換であることに注意して, $(\lambda I + N')^n$ を二項展開してみると, (10) 式から,

$$\begin{aligned} B^n &= (\lambda I + N')^n && \text{( (9) 式から )} \\ &= \lambda^n I + n\lambda^{n-1}N' + \cdots + (N')^n \\ &= \lambda^n I + n\lambda^{n-1}N' && \text{( (2) 式から )} \\ &= \begin{pmatrix} \lambda^n & n\lambda^{n-1} & 0 \\ 0 & \lambda^n & 0 \\ 0 & 0 & \lambda^n \end{pmatrix} \end{aligned}$$

となることが分かります.[2]

(5) (3) と同様に考えると, 勝手な多項式 $f(x) \in \mathbb{C}[x]$ に対して,

$$\begin{aligned} f(B) &= f(\lambda)I + f'(\lambda)N' \\ &= \begin{pmatrix} f(\lambda) & f'(\lambda) & 0 \\ 0 & f(\lambda) & 0 \\ 0 & 0 & f(\lambda) \end{pmatrix} \end{aligned} \tag{11}$$

となることが分かります. したがって, (11) 式から,

$$f(B) = O \iff f(\lambda) = f'(\lambda) = 0$$

となることが分かります.

## 2 $n$ 行 $n$ 列のときにはどうなるのか

興味がある方がいるかもしれませんので, ここでは, 行列 $A$ が,

$$A = \underbrace{\begin{pmatrix} \lambda & 1 & 0 & \cdots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & \lambda & 1 \\ 0 & 0 & \cdots & 0 & \lambda \end{pmatrix}}_{n\text{ コ}}$$

という $n$ 行 $n$ 列の行列のときにはどうなるのかということを考えてみることにします. 問 1 と同様に,

$$N = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 0 \\ \vdots & \vdots & \ddots & 0 & 1 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n\text{ コ}}$$

とすると, 行列 $A$ は,

**行列 $A$ をスカラー行列とベキ零行列の和に分解する**

$$A = \lambda I + N$$

というように表わせることが分かります. このとき, $N^2, N^3, \cdots$ を具体的に計算してみ

[2] もちろん, 上のように二項展開を用いて $B^n$ を計算しなければいけないということではありません. 例えば, $B^2, B^3, \cdots$ を具体的に計算してみることで $B^n$ の形を予想して, その予想が正しいことを数学的帰納法で証明するという方針でも構いません.

ると,

$$N^2 = \underbrace{\begin{pmatrix} 0 & 0 & 1 & \cdots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & \ddots & \ddots & 1 \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n\ \text{コ}}, \cdots, \ N^{n-2} = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \ddots & 0 & 1 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n\ \text{コ}},$$

$$N^{n-1} = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n\ \text{コ}}, \ N^n = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{pmatrix}}_{n\ \text{コ}}$$

というように, **$N$ を掛ける毎に 1 が対角線の上の方へ一段づつ押し上げられて, 最後には,**

$$\boldsymbol{N^n = O}$$

**となる**ことが分かります.[3] よって, 問 1 の (1) と同様にして, 勝手な自然数 $k \in \mathbb{N}$ に対して,

$$\begin{aligned} A^k &= (\lambda I + N)^k \\ &= \lambda^k I + k\lambda^{k-1} N + \cdots + \frac{k(k-1)\cdots(k-n+2)}{(n-1)!}\,\lambda^{k-n+1} N^{n-1} \end{aligned} \tag{12}$$

というように表わせることが分かります. したがって,

$$\begin{aligned} f(x) &= a_0 + a_1 x + \cdots + a_n x^n \\ &= \sum_{k=0}^{n} a_k x^k \end{aligned}$$

とすると, 問 1 の (3) と同様に, (12) 式から,

$$\begin{aligned} f(A) &= a_0 I + a_1 A + \cdots + a_n A^n \\ &= \sum_{k=0}^{n} a_k A^k \\ &= \sum_{k=0}^{n} a_k \left\{ \lambda^k I + k\lambda^{k-1} N + \cdots + \frac{k(k-1)\cdots(k-n+2)}{(n-1)!}\,\lambda^{k-n+1} N^{n-1} \right\} \end{aligned}$$

[3] 例えば, 数学的帰納法を用いると, この事実をきちんと証明することができます.

$$
\begin{aligned}
&= \left\{ \sum_{k=0}^{n} a_k \lambda^k \right\} \cdot I + \left\{ \sum_{k=0}^{n} k a_k \lambda^{k-1} \right\} \cdot N \\
&\quad + \cdots + \frac{1}{(n-1)!} \left\{ \sum_{k=0}^{n} k(k-1)\cdots(k-n+2) a_k \lambda^{k-n+1} \right\} \cdot N^{n-1}
\end{aligned} \tag{13}
$$

というように表わせることが分かります．ここで，

$$
\begin{cases}
f(\lambda) = \sum_{k=0}^{n} a_k \lambda^k \\
f'(\lambda) = \sum_{k=0}^{n} k a_k \lambda^{k-1} \\
\qquad \vdots \\
f^{(n-1)}(\lambda) = \sum_{k=0}^{n} k(k-1)\cdots(k-n+2) a_k \lambda^{k-n+1}
\end{cases} \tag{14}
$$

となることに注意すると, (13) 式, (14) 式から,

$$
\begin{aligned}
f(A) &= f(\lambda) I + f'(\lambda) N + \cdots + \frac{f^{(n-1)}(\lambda)}{(n-1)!} N^{n-1} \\
&= \underbrace{\begin{pmatrix}
f(\lambda) & f'(\lambda) & \frac{f''(\lambda)}{2!} & \cdots & \frac{f^{(n-1)}(\lambda)}{(n-1)!} \\
0 & f(\lambda) & f'(\lambda) & \ddots & \vdots \\
0 & 0 & \ddots & \ddots & \frac{f''(\lambda)}{2!} \\
\vdots & \vdots & \ddots & f(\lambda) & f'(\lambda) \\
0 & 0 & \cdots & 0 & f(\lambda)
\end{pmatrix}}_{n\text{ コ}}
\end{aligned} \tag{15}
$$

となることが分かります．したがって, (15) 式から,

**多項式 $f(x)$ が行列 $A$ を「根」に持つための条件**

$$
f(A) = O \iff f(\lambda) = f'(\lambda) = \cdots = f^{(n-1)}(\lambda) = 0
$$

となることが分かります．

## 3　問 1 を見直すと

さて, $A$ を正方行列とすると, $A^2, A^3, \cdots$ などを考えることができますから, 勝手な多項式 $f(x)$ に対して, 変数 $x$ のところに $A$ という行列を代入して, $f(A)$ という行列を考えることができます．すなわち, 複素数を係数とする多項式全体の集合を,

**複素数を係数とする多項式全体の集合**

$$
\mathbb{C}[x] = \{ f(x) = a_0 + a_1 x + \cdots + a_n x^n \mid n \in \mathbb{N},\, a_0, a_1, \cdots, a_n \in \mathbb{C} \}
$$

として,

$$
f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{C}[x]
$$

に対して,

$$f(A) = a_0 I + a_1 A + \cdots + a_n A^n$$

という行列を考えることができます．このとき, 実は, **行列 $A$ の性質は, 行列 $A$ を「根」に持つような多項式に注目することで, より良く理解することができる**ということが分かっています．すなわち, いま, 行列 $A$ を「根」に持つような多項式全体の集合を,

**行列 $A$ を「根」に持つような多項式全体の集合**

$$I_A = \{ f(x) \in \mathbb{C}[x] \mid f(A) = O \}$$

という記号を用いて表わすことにすると, **行列 $A$ の性質は $I_A$ という集合に強く反映される**ということが分かっています．そこで, 行列 $A$ の (固有な) 性質が「見やすい形」で表現されるような「上手い「番地割り」を見つける」という**「行列の標準形の問題」**を考える上でも, **$I_A$ という集合の持つ性質をよく調べてみる**ということが大切になります．この演習でも, こうした事柄について順を追って説明していこうと思いますが, そのための準備として問 1 を出題してみました.

さて, 行列 $A$ が, 例えば,

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \tag{16}$$

というような特別な形をしているときには, $A^n$ などの計算を見通し良く行なうことができるということは, 第 1 回の問 2 のところでも見たことです．そこで, 行列 $A$ が, このような特別な形をしているときには, $I_A$ という集合も比較的容易に求めることができるのではないかと期待することは自然なことのように思われます．実際, 問 1 の (3) のところで見たように,

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

として, 勝手な多項式 $f(x) \in \mathbb{C}[x]$ に対して, (16) 式で与えられる行列 $A$ を多項式 $f(x)$ に代入することによって得られる行列 $f(A)$ が,

**(16) 式の行列 $A$ に対する行列 $f(A)$ の具体的な形**

$$\begin{aligned} f(A) &= f(\lambda)I + f'(\lambda)N + \frac{1}{2}f''(\lambda)N^2 \qquad (17) \\ &= \begin{pmatrix} f(\lambda) & f'(\lambda) & \frac{1}{2}f''(\lambda) \\ 0 & f(\lambda) & f'(\lambda) \\ 0 & 0 & f(\lambda) \end{pmatrix} \end{aligned}$$

というように具体的な式で表わせることが分かります．ただし, いきなり, (17) 式を持ち出すと, 皆さんに唐突な感じを与えてしまうかも知れないと考えて, 問 1 では, まず, $f(x)$ が三次式の場合に $f(A)$ という行列を具体的に求めてもらうことで, 少し感じをつかんでもらおうと思いました．そこで, ここでは, どうして (17) 式が成り立つのかということを見直してみることにします.

問 1 の (1) の解答で見たように, 行列 $A$ を,

**行列 $\boldsymbol{A}$ をスカラー行列 $\boldsymbol{\lambda I}$ とベキ零行列 $\boldsymbol{N}$ の和に分解する**

$$A = \lambda I + N$$

という形に表わしてみると, 勝手な自然数 $n \in \mathbb{N}$ に対して,

$$A^n = (\lambda I + N)^n \tag{18}$$

というように表わすことができますから, (18) 式の右辺を二項展開することで, $A^n$ を求めることができます. 第 1 回の問 2 のところで触れましたが, このような計算が許されるのは, 例えば,

$$\begin{aligned}(x+y)^2 &= (x+y)(x+y) \\ &= x^2 + xy + yx + y^2 \\ &= x^2 + 2xy + y^2\end{aligned}$$

という計算には, **$\boldsymbol{x, y}$ に対して,**

$$\boldsymbol{xy = yx}$$

**が成り立つという性質しか使っていない**ということに注意すると, **$\boldsymbol{X, Y}$ が互いに可換な正方行列の場合, すなわち, $\boldsymbol{X, Y}$ が,**

$$\boldsymbol{XY = YX}$$

**となるような正方行列である場合**でも, 例えば,

$$(X+Y)^2 = X^2 + 2XY + Y^2$$

という式のように, **二項展開の式がそのまま成り立つ**ことが分かるからでした. ここで, 例えば,

$$f(x) = x^n$$

とすると, $f(x+y)$ は,

$$\begin{aligned}f(x+y) &= (x+y)^n \\ &= x^n + nx^{n-1}y + \cdots + y^n \qquad (19)\end{aligned}$$

というように二項展開して表わすことができますが, こうして得られた (19) 式の右辺は, **$\boldsymbol{x}$ をひとつ固定したときに,**

$$\boldsymbol{f(x+y) = (x+y)^n}$$

**という「$\boldsymbol{y}$ の関数」を「$\boldsymbol{y}$ のべきの形」で表わした式である**とみなすこともできます. このように考えると, **二項展開とは,**

**多項式関数に対する Taylor 展開とは二項展開のことに他ならない**

$$\boldsymbol{f(x+y) = f(x) + f'(x)y + \frac{f''(x)}{2!}y^2 + \cdots + \frac{f^{(n)}(x)}{n!}y^n} \tag{20}$$

**という Taylor 展開に他ならない**ということが分かります.[4] 実際,

$$f(x) = x^n$$

とすると, 勝手な自然数 $k \in \mathbb{N}$ に対して,

$$f^{(k)}(x) = \begin{cases} n(n-1)\cdots(n-k+1)x^{n-k}, & (\ k \leq n \text{ のとき }) \\ 0, & (\ k > n \text{ のとき }) \end{cases}$$

となることが分かりますから, $k \leq n$ のとき

$$\begin{aligned} \frac{f^{(k)}(x)}{k!}y^k &= \frac{n(n-1)\cdots(n-k+1)}{k!}\, x^{n-k}y^k \\ &= \frac{n!}{k!(n-k)!}\, x^{n-k}y^k \end{aligned}$$

というように表わせることが分かります.

そこで, このことに注意して, $a_0, a_1, \cdots, a_n \in \mathbb{C}$ として, $n$ 次の多項式

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n$$

に,

$$A = \lambda I + N$$

という行列を代入するとどうなるかということを考えてみます. いま, $f(x)$ は $n$ 次の多項式ですから, $k \geq n+1$ のとき,

$$f^{(k)}(x) = 0$$

となることに注意して, $f(x)$ の Taylor 展開を考えてみると,

$$f(x+y) = f(x) + f'(x)\, y + \frac{f''(x)}{2!}\, y^2 + \cdots + \frac{f^{(n)}(x)}{n!}\, y^n \tag{21}$$

となることが分かります.[5] そこで, (21) 式を「多項式 $f(x+y)$ の各項を二項展開して, $y$ について整理して得られた式である」と考えてみることにします. すなわち, 例えば,

$$f(x) = 1 + x + 3x^2$$

[4] ここで, $f(x)$ が $n$ 次の多項式であるとすると, $k \geq n+1$ となる自然数 $k \in \mathbb{N}$ に対して, $f^{(k)}(x) = 0$ となりますから, (20) 式の右辺は有限和になることに注意します.

[5] ここで, $y \rightsquigarrow \Delta x$ と書き直して, (21) 式を,

$$f(x+\Delta x) = f(x) + f'(x)\, \Delta x + \frac{f''(x)}{2!}\, (\Delta x)^2 + \cdots + \frac{f^{(n)}(x)}{n!}\, (\Delta x)^n$$

と表わしてみると, より感じが出るかもしれません. あるいは, さらに, $x \rightsquigarrow a$, $\Delta x \rightsquigarrow (x-a)$ と書き直して,

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(x)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(x)}{n!}(x-a)^n$$

と表わしてみると, (21) 式が, 皆さんが良くご存じの Taylor 展開の式であることが納得できるかもしれません.

であるとすると,

$$\begin{cases} f'(x) = 1 + 6x, \\ f''(x) = 6 \end{cases}$$

となることが分かりますから,

$$f(x) + f'(x)\,y + \frac{f''(x)}{2!}\,y^2 = (1 + x + 3x^2) + (1 + 6x)y + 3y^2 \tag{22}$$

となることが分かりますが, (22) 式を,

$$\begin{aligned} f(x+y) &= 1 + (x+y) + 3(x+y)^2 \\ &= 1 + (x+y) + 3(x^2 + 2xy + y^2) \\ &= (1 + x + 3x^2) + (1 + 6x)y + 3y^2 \end{aligned}$$

というように「多項式 $f(x+y)$ の各項を二項展開して, $y$ について整理して得られた式である」と考えてみることにします. すると, 上で注意したように, このような計算には,

$$xy = yx$$

という性質しか用いていませんから, 前と同様に,

$$XY = YX$$

となるような正方行列 $X, Y$ に対しても, (21) 式と同じ式が成り立つはずです. したがって,

$$XY = YX$$

となる正方行列 $X, Y$ に対して,

$$f(X+Y) = f(X) + f'(X) \cdot Y + \frac{f''(X)}{2!} \cdot Y^2 + \cdots + \frac{f^{(n)}(X)}{n!} \cdot Y^n \tag{23}$$

という行列の間の等式が成り立つことが分かります. そこで,

$$N^3 = O$$

となることに注意して, (23) 式において,

$$X = \lambda I,\ Y = N$$

としてみると,

$$\begin{aligned} f(A) &= f(\lambda I + N) \\ &= f(\lambda I) + f'(\lambda I) \cdot N + \frac{f''(\lambda I)}{2!} \cdot N^2 \\ &= f(\lambda) I + f'(\lambda) I \cdot N + \frac{f''(\lambda)}{2!} I \cdot N^2 \end{aligned}$$

$$= f(\lambda)I + f'(\lambda)N + \frac{f''(\lambda)}{2!}N^2$$

となることが分かりますが, これは, (17) 式に他なりません. すなわち,

$$f(A) = f(\lambda I + N)$$

という行列を, 上のように Taylor 展開を用いて求めたのだと考えてみると, (17) 式は自然に理解できることが分かります. 興味のある方は, 問 1 の (5) で求めた

$$f(B) = f(\lambda I + N')$$

という行列に対する表示式や 2 節で求めた $f(A)$ という行列に対する表示式を, 同様の方法で求めてみて下さい.

このように考えてみると,「行列の世界」と「普通の数の世界」の間には,

**行列と普通の数の間の対応**

| 行列の世界 | | 普通の数の世界 |
|---|---|---|
| $A = \lambda I + N$ | $\longleftrightarrow$ | $a = \lambda + \delta a$ |
| スカラー行列 : $\lambda I$ | $\longleftrightarrow$ | 複素平面上の点 : $\lambda \in \mathbb{C}$ |
| ベキ零行列 : $N$ | $\longleftrightarrow$ | 無限小変位 : $\delta a$ |

というような対応があることが分かります.

## 4　問 2 の解答

(1) $f(x), g(x) \in I_A$ であるとすると, $I_A$ の定義から,

$$f(A) = g(A) = O$$

となることが分かりますから,

$$\begin{aligned} f(A) + g(A) &= O + O \\ &= O \end{aligned}$$

となることが分かります. よって,

$$f(x) + g(x) \in I_A$$

となることが分かります.

(2) $f(x) \in I_A$ であるとすると, $I_A$ の定義から,

$$f(A) = O$$

となることが分かりますから,

$$\begin{aligned} f(A)\,h(A) &= O \cdot h(A) \\ &= O \end{aligned}$$

となることが分かります. よって,

$$f(x)h(x) \in I_A$$

となることが分かります.

(3) $PP^{-1} = I$ となることに注意して,

$$A' = P^{-1}AP$$

に対して, $(A')^n$ を計算してみると,

$$\begin{aligned}
(A')^n &= \underbrace{(P^{-1}AP)(P^{-1}AP)\cdots(P^{-1}AP)}_{n\ \text{コ}} \\
&= P^{-1}A\,(PP^{-1})\,A\,(PP^{-1})\,A\,\cdots\,A\,(PP^{-1})\,AP \\
&= P^{-1}A^nP
\end{aligned}$$

となることが分かります.[6]

(4) (3) の結果より, 勝手な多項式

$$f(x) = a_0 + a_1x + \cdots + a_nx^n \in \mathbb{C}[x]$$

に対して,

$$\begin{aligned}
f(A') &= a_0I + a_1A' + \cdots + a_n(A')^n \\
&= a_0(P^{-1}IP) + a_1(P^{-1}AP) + \cdots + a_n(P^{-1}A^nP) \\
&= P^{-1}\,(\,a_0I + a_1A + \cdots + a_nA^n\,)\,P \\
&= P^{-1}f(A)\,P
\end{aligned}$$

となることが分かります. よって,

$$f(A') = P^{-1}f(A)\,P \tag{24}$$

となることが分かります.

(5) (24) 式から, $f(x) \in I_A$ に対して,

$$\begin{aligned}
f(A') &= P^{-1}f(A)\,P \\
&= P^{-1}\cdot O\cdot P \\
&= O
\end{aligned}$$

となることが分かります. よって,

$$f(x) \in I_A \implies f(x) \in I_{A'}$$

[6] もう少し厳密に証明したければ, $n$ に関する数学的帰納法を用いて証明することができます.

となることが分かりますから,

$$I_A \subset I_{A'} \tag{25}$$

となることが分かります.

また, $P$ は正則行列であることに注意して, (24) の両辺の左から $P$ を掛け算し, 右から $P^{-1}$ を掛け算すると,

$$f(A) = P\, f(A')\, P^{-1} \tag{26}$$

となることが分かります. すると, 前と同様にして, (26) 式から,

$$I_A \supset I_{A'} \tag{27}$$

となることが分かります. したがって, (25) 式と (27) 式から,

$$I_A = I_{A'}$$

となることが分かります.

## 5 問 2 を見直すと

3 節でも注意しましたが, **正方行列 $A$ の性質は, 行列 $A$ を「根」に持つような多項式全体の集合**

**行列 $A$ を「根」に持つような多項式全体の集合**

$$\boldsymbol{I_A = \{ f(x) \in \mathbb{C}[x] \mid f(A) = O \}}$$

**に注目することで, より良く理解することができる**ということが分かっています. そこで, こうした事柄を順番に考えてみるための第一歩として, 皆さんに, $I_A$ という集合の持つ基本的な性質を理解してもらおうと思って, 問 2 を出題してみました.

さて, 多項式同士を足したり掛けたりしても, やはり多項式が得られますから, 多項式全体の集合 $\mathbb{C}[x]$ は「足し算」や「掛け算」のできるような集合になっています.[7] 数学では, 多項式の集合のように,「足し算」や「掛け算」のできる集合のことを, 一般に, **環** ( ring ) と呼びます. 多項式全体の集合は, このような集合の代表例ですが,「足し算」や「掛け算」のできる集合であるということを強調して, **多項式環**と呼ばれています. ここで, 例えば, $x$ を $x^2 - 1$ で割り算してみると,

$$\frac{x}{x^2 - 1}$$

という有理関数になるというように, 多項式同士で「割り算」をしてみると, 一般には, 割り算の結果は多項式になるとは限りません.[8] したがって, 多項式全体の集合 $\mathbb{C}[x]$ の中で

[7] 一変数の多項式を複素数の世界で考えると, **代数学の基本定理**によって, すべての多項式は一次式の積の形に因数分解できて考察が簡単になるので, 以下でも, もっぱら, 複素数を係数に持つような多項式を考えることにしました.

[8] $f(x) \in \mathbb{C}[x]$ を多項式とすると, すべての複素数 $x \in \mathbb{C}$ に対して, $f(x) \in \mathbb{C}$ という値が定まりますが, 例えば,

$$g(x) = \frac{x}{x^2 - 1}$$

とすると, $g(x)$ は $x = \pm 1$ に対して値が定まらなくなってしまいます. したがって, $g(x)$ は多項式ではないことが分かります.

は「割り算」を自由に考えることができません．このように，**「環」と言うときには，必ずしも「割り算ができる」ということは要請していません**．一方，環であって，さらに，0 以外の元に対して，いつでも割り算ができるという場合には，これを**体**と呼んで区別したりします.[9] 例えば，実数全体の集合や複素数全体の集合，あるいは，有理数全体の集合は，こうした体の代表例で，それぞれ，**実数体，複素数体，有理数体**と呼ばれています．以下では，環の掛け算は可換であると仮定して話を進めることにします.[10] このような環を**可換環** ( commutative ring ) と呼んで，$n$ 行 $n$ 列の行列全体の集合のような**非可換環**と区別することがあります．

さて，「線型空間」に対して，「線型部分空間」という概念を考えることができるように，集合に付加的な構造が入っているときには,[11] このような構造を保つような部分集合を考えることは自然なことです．そこで，いま，環 $R$ が勝手にひとつ与えられているとして,[12] 環の場合に「構造を保つような部分集合」という概念を考えると，**部分環**を考えるということになります．すなわち，環 $R$ の部分環とは，それ自身が足し算や掛け算で閉じているような $R$ の部分集合 $R' \subset R$ のことです．これを，数式を使って表わせば，環 $R$ の部分集合 $R' \subset R$ が $R$ の部分環であるとは，

**部分環の条件**

$$x, y \in R' \implies x + y,\, xy \in R' \tag{28}$$

という条件が成り立つことであるということになります．

このように，環に対して部分環を考えるということは，線型空間に対して線型部分空間を考えるということのように自然なことなのですが，線型空間の場合と大きく異なる点は，**部分環の中には，特別な性質をもつ部分環，いわば，「偉い部分環」という概念が考えられる**ということです．こうした「偉い部分環」のことを，**イデアル** ( ideal ) と呼ぶのですが，これも数式を使って表わすと，(可換) 環 $R$ の部分集合 $I \subset R$ がイデアルであるとは，

**イデアルの条件**

$$x, y \in I,\ z \in R \implies x + y,\, xz \in I \tag{29}$$

という条件が成り立つことであるということになります．ここで，(28) 式の条件と (29) 式の条件を見比べてみると，**イデアルの条件では「掛け算 $xz$ において，少なくとも一方の元が $I$ の元であるならば，( 例えば，$x \in I$ であるならば，) もう一方の元が，必ずしも $I$ の元でなくとも，( 例えば，$z \notin I$ であったとしても，) 掛け算の結果 $xz$ は $I$ の元になる」という部分環の条件より強い条件を課している**ということが分かります．このように，一見，人工的に見えるイデアルという概念が，どうして部分環という概念より自然な概念であり，また，より重要な概念であると考えられているのかということについては，10 節で

[9]すなわち，「環」とは「加減乗」ができる集合のことであり，「体」とは「加減乗除」のできる集合のことです．

[10]すなわち，環の勝手な元 $x, y$ に対して，$xy = yx$ が成り立つということです．

[11]すなわち，いまの場合であれば，「足し算」や「掛け算」ができるとか，「スカラー倍」ができるとかいった代数的な構造のことです．

[12]「環」のことを英語で ring と言うので，環を表わすのに $R$ という記号が使われることが多いです．また，これでは抽象的で考えにくいと思われる方は，$R$ として一変数多項式環 $\mathbb{C}[x]$ だけを考えてもらっても構いません．

少し触れることにしようと思います．こうした言葉を用いると, 問 2 の (1), (2) で見たことは, **行列 $A$ を「根」に持つ多項式全体の集合 $I_A$ は, 一変数多項式環 $\mathbb{C}[x]$ のイデアルになる**ということになります.

さて, $n$ 行 $n$ 列の行列 $A$ に対して, 行列 $A$ を掛け算することによって定まる線型空間 $\mathbb{C}^n$ 上の線型写像を,

$$L_A : \mathbb{C}^n \to \mathbb{C}^n$$

という記号を用いて表わすことにします.[13] このとき, 最初の行列 $A$ は,

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix},\ \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix},\ \cdots,\ \mathbf{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{C}^n$$

という線型空間 $\mathbb{C}^n$ の自然な基底に関する線型写像 $L_A$ の表現行列であると解釈することができます．そこで, いま, $\mathbb{C}^n$ の基底 $\{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n\}$ を, 勝手にひとつ取ってきて, $\{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n\}$ という基底を用いて, $\mathbb{C}^n$ の「番地割り」をやり直してみることにします. すると, 第 8 回の問 1 のところで見たように, このような「番地割り」の取り換えのもとで, 線型写像 $L_A$ の表現行列は, $\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n \in \mathbb{C}^n$ を列ベクトルとする正則行列を

$$P = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{pmatrix}$$

として,

$$A \rightsquigarrow P^{-1}AP$$

というように「姿」を変えることが分かります．このことは, **適当な正則行列 $P$ を用いて,**

$$\boldsymbol{A' = P^{-1}AP}$$

**という関係にあるような正方行列 $A$ と $A'$ は「本質的に同じ行列」である**ということを意味しています．すなわち, いま, 特定の座標軸に惑わされないように, $\mathbb{C}^n$ から最初の座標軸を消し去って, $\mathbb{C}^n$ を線型空間と思ったものを,

$$V = \mathbb{C}^n$$

と表わし, 行列 $A$ を掛け算することによって定まる線型写像の方も「$_A$」という添え字を省略して, 単に,

$$L : V \to V$$

と表わすことにすると, **これらの行列は, いずれも, $L : V \to V$ という「同じ線型写像」を表わしており, 行列としての姿が違って見えるのは, 単に,「異なる視点」から眺めているからに過ぎない**のだと考えることができます．このように考えてみると,**「線型空間 $V$ の「番地割り」の仕方に依存せずに理解できる線型写像 $L : V \to V$ の性質」**[14]**と**

[13]ここで, 多項式を表わす $f(x)$ と混同して余計な混乱を起こしてしまわないように, 行列 $A$ を掛け算することによって定まる線型空間 $\mathbb{C}^n$ 上の線型写像を, $f_A : \mathbb{C}^n \to \mathbb{C}^n$ ではなく, $L_A : \mathbb{C}^n \to \mathbb{C}^n$ という記号を用いて表わすことにしました. 英語で線型写像のことを linear map と言います.

[14]第 9 回の問 3 のところでは, こうした性質を「固有な性質」と呼びました.

**「$\boldsymbol{A \rightsquigarrow P^{-1}AP}$ という変換のもとで不変な行列 $\boldsymbol{A}$ の性質」とが対応している**ことが分かります.

第 9 回の問 3 のところで見たように, こうした「固有な性質」の代表例が, 固有値や固有ベクトル空間といった概念です. 例えば, 第 10 回の問 1 のところで見たように,

**特性多項式の共役不変性**

$$\varphi_A(x) = \varphi_{P^{-1}AP}(x)$$

というように, 行列 $A$ の特性多項式 $\varphi_A(x)$ と行列 $P^{-1}AP$ の特性多項式 $\varphi_{P^{-1}AP}(x)$ は等しくなることが分かりますから, これら二つの行列の固有値は等しくなることが分かります. このことは,

$$L(\mathbf{u}) = \lambda\mathbf{u}, \quad \mathbf{u} \in V$$

という方程式が, $\mathbf{u} \neq \mathbf{0}$ となるような解を持つような複素数 $\lambda \in \mathbb{C}$ として, すなわち,

$$\mathrm{Ker}(\,\lambda \cdot \mathrm{id}_V - L\,) \neq \{\mathbf{0}\}$$

となるような複素数 $\lambda \in \mathbb{C}$ として, **固有値という概念を $\mathbb{C}^n$ の特定の「番地割り」の仕方に依らずに理解できる**ということに対応しています.[15]

こうした観点から眺めると, 問 2 の (5) で見たことは, **$\boldsymbol{I_A}$ という集合も線型写像 $\boldsymbol{L}$ の「固有な性質」を表わす概念のひとつである**ということになります. 一般に, 二つの $\mathbb{C}$ 上の線型空間 $V$, $W$ の間の線型写像全体の集合を,

**線型空間 $\boldsymbol{V}$, $\boldsymbol{W}$ の間の線型写像全体の集合**

$$\mathrm{Hom}(V, W) = \{L : V \to W \mid L \text{ は線型写像 }\}$$

という記号を用いて表わすことにします.[16] すると, 第 7 回の問 1 のところで見たように, $L, M \in \mathrm{Hom}(V, W)$, $a \in \mathbb{C}$ として, $\mathbf{u} \in V$ に対して,

**線型写像同士の足し算やスカラー倍**

$$\begin{cases} (L + M)(\mathbf{u}) = L(\mathbf{u}) + M(\mathbf{u}) \\ (aL)(\mathbf{u}) = a \cdot L(\mathbf{u}) \end{cases}$$

という式によって, $\mathrm{Hom}(V, W)$ という集合上に「足し算」や「$\mathbb{C}$ 倍」の構造が入ることが分かります. すなわち, **$\mathbf{Hom}(\boldsymbol{V}, \boldsymbol{W})$ 自身がまた $\mathbb{C}$ 上の線型空間になる**ことが分かります.

ここで, さらに,

$$V = W$$

であるとすると, $L, M \in \mathrm{Hom}(V, V)$ に対して, $L$ と $M$ の合成写像

[15] ここで, $V$ 上の恒等写像を $\mathrm{id}_V : V \to V$ という記号を用いて表わしました.

[16] 線型写像とは, 線型空間の構造を保つような写像のことですが, 数学では, こうした数学的な構造を保つような写像のことを, 一般に, **準同型写像** ( <u>hom</u>omorphism ) と言います.

**$V = W$ のときには, さらに,「掛け算」の構造を考えることができる**

$$(L \circ M)(\mathbf{u}) = L(M(\mathbf{u})), \quad \mathbf{u} \in V$$

を考えることができますが, $L \circ M$ も線型写像となることが分かるので,

$$L \circ M \in \mathrm{Hom}(V, V)$$

となることが分かります. したがって, 合成写像 $L \circ M$ を $L$ と $M$ の「掛け算」であると解釈することにすれば, $\mathrm{Hom}(V, V)$ は, 線型空間としての構造の他に「掛け算」の構造も持つことが分かります. すると, こうした構造さえあれば, 勝手な多項式

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{C}[x]$$

に対して, $f(x)$ の変数 $x$ のところに, 勝手な線型写像 $L \in \mathrm{Hom}(V, V)$ を代入して, 線型写像

$$f(L) = a_0 \,\mathrm{id}_V + a_1 L + \cdots + a_n L^n : V \to V$$

考えてみることができることが分かります.[17] よって, 線型写像 $L \in \mathrm{Hom}(V, V)$ に対しても, $L$ を「根」に持つような多項式全体の集合

**線型写像 $L$ を「根」に持つような多項式全体の集合**

$$I_L = \{ f(x) \in \mathbb{C}[x] \mid f(L) = 0 \}$$

を考えることができることが分かります.

また, 第 7 回の問 1 のところで見たように, 線型空間 $V$ に基底を定めて「番地割り」して考えたときに, 線型写像 $L, M \in \mathrm{Hom}(V, V)$ の表現行列を, それぞれ, $A, B \in M_n(\mathbb{C})$ と表わすことにすると,[18] $a \in \mathbb{C}$ として,

**線型写像とその表現行列との間の対応 (その１)**

| 線型写像 | | 表現行列 |
|---|---|---|
| $\mathrm{Hom}(V, V)$ | $\cong$ | $M_n(\mathbb{C})$ |
| $L$ | $\longleftrightarrow$ | $A$ |
| $M$ | $\longleftrightarrow$ | $B$ |
| $L + M$ | $\longleftrightarrow$ | $A + B$ |
| $aL$ | $\longleftrightarrow$ | $aA$ |
| $L \circ M$ | $\longleftrightarrow$ | $AB$ |

というように, **線型写像における「足し算」や「スカラー倍」や「積」は, 行列における**

[17] 実際, 第 7 回の問 1 のところでは, 勝手な多項式 $f(x) \in \mathbb{R}[x]$ に「微分写像」$D = \frac{d}{dx}$ を代入して, 定数係数の線型常微分作用素

$$\begin{aligned} f(D) &= a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0 \,\mathrm{id} \\ &= a_n \frac{d^n}{dx^n} + a_{n-1} \frac{d^{n-1}}{dx^{n-1}} + \cdots + a_1 \frac{d}{dx} + a_0 \,\mathrm{id} \end{aligned}$$

を考えました.

[18] ここで, $n = \dim_{\mathbb{C}} V$ として, $n$ 行 $n$ 列の複素行列全体の集合を $M_n(\mathbb{C})$ という記号を用いて表わしました.

**「足し算」や「スカラー倍」や「積」に対応している**ことが分かります. よって,

$$\begin{aligned}\mathrm{Hom}(V,V) \ni f(L) = a_0\,\mathrm{id}_V + a_1 L + \cdots + a_n L^n \\ \longleftrightarrow \ f(A) = a_0 I + a_1 A + \cdots + a_n A^n \in M_n(\mathbb{C})\end{aligned}$$

というように対応することが分かりますから,

$$f(L) = 0 \in \mathrm{Hom}(V,V) \iff f(A) = O \in M_n(\mathbb{C}) \tag{30}$$

となることが分かります. したがって, (30) 式から,

$$I_L = I_A$$

というように, **行列 $A$ を「根」に持つような多項式全体の集合 $I_A$ も, 線型写像 $L$ を「根」に持つような多項式全体の集合 $I_L$ として, $\mathbb{C}^n$ の特定の「番地割り」の仕方に依らない形で理解できる**ことが分かります. このことが, 問 2 の (5) で確かめた

**イデアル $I_A$ の共役不変性**

$$I_A = I_{P^{-1}AP}$$

という事実と対応していることが分かります.

さて, いま, 線型空間 $V$ の基底を取り換えて, $V$ の「番地割り」を「旧番地割り」から「新番地割り」に取り換えたとします. このとき, 線型写像 $L : V \to V$ の「旧番地割り」に関する表現行列を $A$ として,「新番地割り」に関する表現行列を $A'$ とします. すると, 上で述べたことから, $n \in \mathbb{N}$, $f(x) \in \mathbb{C}[x]$ として, $L^n \in \mathrm{Hom}(V,V)$ という線型写像, あるいは, より一般に, $f(L) \in \mathrm{Hom}(V,V)$ という線型写像の「旧番地割り」に関する表現行列,「新番地割り」に関する表現行列は,

**線型写像とその表現行列との間の対応 (その２)**

| 表現行列 (「旧番地割り」) | | 線型写像 | | 表現行列 (「新番地割り」) |
|---|---|---|---|---|
| $M_n(\mathbb{C})$ | $\cong$ | $\mathrm{Hom}(V,V)$ | $\cong$ | $M_n(\mathbb{C})$ |
| $A$ | $\longleftrightarrow$ | $L$ | $\longleftrightarrow$ | $A'$ |
| $A^n$ | $\longleftrightarrow$ | $L^n$ | $\longleftrightarrow$ | $(A')^n$ |
| $f(A)$ | $\longleftrightarrow$ | $f(L)$ | $\longleftrightarrow$ | $f(A')$ |

というように, それぞれ, $A^n, (A')^n \in M_n(\mathbb{C})$, あるいは, $f(A), f(A') \in M_n(\mathbb{C})$ となることが分かります. ここで, 表現行列の変換公式から, $A$ と $A'$ の間の関係は, 適当な正則行列 $P$ を用いて,

$$A' = P^{-1}AP \tag{31}$$

という形で与えられることが分かりますが, **正則行列 $P$ は線型空間 $V$ の「番地割り」の取り換え方のみに依る行列であり, 線型写像 $L \in \mathrm{Hom}(V,V)$ には依らない行列である**ということに注意します. すなわち, **$L \in \mathrm{Hom}(V,V)$ がどのような線型写像であったとしても, $L$ の「旧番地割り」に関する表現行列と「新番地割り」に関する表現行列の間**

**には, (31) 式のような関係式が成り立つ**ことに注意します. そこで, $L \rightsquigarrow L^n$ と置き換えて, 線型写像 $L^n$ に対して, (31) 式の変換公式を用いると,

$$(A')^n = P^{-1}A^nP \tag{32}$$

となることが分かります. 全く同様に, $L \rightsquigarrow f(L)$ と置き換えて, 線型写像 $f(L)$ に対して, (31) 式の変換公式を用いると,

$$f(A') = P^{-1}f(A)P \tag{33}$$

となることが分かります. 問 2 の (3), (4) では, (32) 式, (33) 式が成り立つことを直接計算して確かめましたが, 上のように, これらの式は**「線型写像 $L^n$, $f(L)$ に対する表現行列の変換公式である」**と解釈してみると, 何も計算しなくとも, **最初から, (32) 式, (33) 式が成り立つことが分かる**というわけです.

## 6 「行列の標準形の問題」との関係

さて, 第 8 回の問 1 のところで見たように,「行列の標準形の問題」とは,

**行列の標準形の問題**

与えられた正方行列 $A$ に対して,

$$P^{-1}AP = \Lambda$$

となるような「見やすい形」の行列 $\Lambda$ と正則行列 $P$ を見つけよ.

という問題でした. また, この問題を「線型空間」という視点から考えると, 行列 $A$ を掛け算することによって定まる $\mathbb{C}^n$ 上の線型写像を,

$$L_A : \mathbb{C}^n \to \mathbb{C}^n$$

と表わすときに, **与えられた線型写像 $L_A$ の表現行列が ( 対角行列などの )「見やすい形」となるような $\mathbb{C}^n$ の「上手い基底」を見つける問題**として解釈することができるのでした. 実際, 第 10 回の問 1 のところでは,「内積を持つ $\mathbb{R}$ 上の線型空間」という観点から考察することで, **$A$ が対称行列の場合には「対称行列 $A$ に関して数ベクトル空間 $\mathbb{R}^n$ が固有ベクトル空間分解する」という形で「対角化の問題」が解決できる**ことを見ました. さらに, **対称行列 $A$ の異なる固有値に対応する固有ベクトル空間は互いに直交する**ということを確かめることで, **対角化を実現する正則行列 $P$ として直交行列が取れる**ということも見ました.

このように, 対称行列など「内積と相性が良いような行列」の場合には,「見やすい形の行列」として「対角行列」を採用することで,「行列の標準形の問題」を解決できることが分かります. 一方, 第 9 回の問 3 のところで見たように, **$N \neq O$ となるようなベキ零行列 $N$ に対しては,「対角化の問題」はことごとく解決しないということ, すなわち, $P^{-1}NP$ が対角行列となるような正則行列 $P$ は存在しない**ことが分かります. したがって, **「行列**

**の標準形の問題」をいつでも解決できるようにするためには,「見やすい形」の行列を「対角行列」より少し一般化して考える必要がある**ことが分かります.

さて, 問 1 では,

**一般化された「見やすい形の行列」の例**

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \tag{34}$$

という形の行列を考えましたが, これらの行列も対角化できないということが, 例えば, 次のようにして分かります. いま, 問 1 の解答と同様に,

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

として,

$$A = \lambda I + N$$

というように表わしてみます. すると, ベキ零行列 $N$ は,

$$N = A - \lambda I \tag{35}$$

というように表わせることが分かります. そこで, いま,

$$P^{-1}AP = \Lambda \tag{36}$$

となる対角行列 $\Lambda$ と正則行列 $P$ が見つかったと仮定してみます. このとき, (35) 式の両辺で, 左から $P^{-1}$ を掛け算して, 右から $P$ を掛け算してみると,

$$\begin{aligned} P^{-1}NP &= P^{-1}(A - \lambda I)P \\ &= P^{-1}AP - \lambda I \\ &= \Lambda - \lambda I \end{aligned}$$

となることが分かりますから, ベキ零行列 $N$ が $(\Lambda - \lambda I)$ という対角行列に変形できることが分かります. ところが, 上で注意したように, ベキ零行列 $N$ は対角化できないことが分かりますから, これは矛盾です. よって, 行列 $A$ は対角化できないことが分かります. 全く同様にして, 行列 $B$ も対角化できないことが分かります.

このように, (34) 式で与えられる行列 $A, B$ は対角化できないのですが, 問 1 で見たように, これらの行列に対しては, **$\boldsymbol{n} \in \mathbb{N}$, $\boldsymbol{f(x)} \in \mathbb{C}[\boldsymbol{x}]$ として, 比較的簡単な計算で, 直接, $\boldsymbol{A^n}$ や $\boldsymbol{f(A)}$ などの行列を求めることができます.** その意味で, これらの行列も「十分見やすい形をしている」と考えることができます. そこで, **このように対角線の一段上だけに 1 がいくつか登場することを許すような「一般化された対角行列」も「見やすい形の行列」の仲間に入れる**ということが考えられました. このような「一般化された対角行列」のことを **Jordan 標準形**と呼ぶのですが,[19] **「見やすい形」の行列を「Jordan 標準形」と**

[19] Jordan 標準形の正確な定義については, 第 12 回のところで説明しようと思います.

**いう形で一般化してみると, 実は,「行列の標準形の問題」がいつでも解決する**ということが分かります. すなわち, **勝手にひとつ与えられた行列 $A$ に対して, $P^{-1}AP$ が Jordan 標準形となるような正則行列 $P$ が存在する**ということが分かります.

さて, 問 1 の結果から, (34) 式で与えられる Jordan 標準形 $A, B$ に対しては, $I_A, I_B$ という一変数多項式環 $\mathbb{C}[x]$ のイデアルが, それぞれ,

**(34) 式の行列に対するイデアル $I_A, I_B$ の具体的な記述**

$$I_A = \left\{ f(x) \in \mathbb{C}[x] \mid f(\lambda) = f'(\lambda) = f''(\lambda) = 0 \ \right\} \tag{37}$$

$$I_B = \left\{ f(x) \in \mathbb{C}[x] \mid f(\lambda) = f'(\lambda) = 0 \ \right\} \tag{38}$$

という形で与えられることが分かります. そこで, もうひとつ

**「見やすい形の行列」の例**

$$C = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \tag{39}$$

というスカラー行列を考えてみると, $f(x) \in \mathbb{C}[x]$ に対して,

$$f(C) = \begin{pmatrix} f(\lambda) & 0 & 0 \\ 0 & f(\lambda) & 0 \\ 0 & 0 & f(\lambda) \end{pmatrix}$$

となることが分かりますから, スカラー行列 $C$ に対応したイデアル $I_C$ は,

**(39) 式の行列に対するイデアル $I_C$ の具体的な記述**

$$I_C = \{ f(x) \in \mathbb{C}[x] \mid f(\lambda) = 0 \ \} \tag{40}$$

という形で与えられることが分かります. これらの Jordan 標準形に対する特性多項式は,

**行列 $A, B, C$ の特性多項式はすべて等しい**

$$\varphi_A(x) = (x - \lambda)^3$$
$$\varphi_B(x) = (x - \lambda)^3$$
$$\varphi_C(x) = (x - \lambda)^3$$

というように, いずれも $(x - \lambda)^3 \in \mathbb{C}[x]$ という**同じ多項式になる**ことが分かりますが,

$$\begin{cases} (x-\lambda)^2 \notin I_A \\ (x-\lambda)^2 \in I_B \end{cases}, \quad \begin{cases} (x-\lambda) \notin I_B \\ (x-\lambda) \in I_C \end{cases}$$

となることが分かりますから,

**行列 $A, B, C$ に対応するイデアル $I_A, I_B, I_C$ はすべて異なる**

$$I_A \subsetneq I_B \subsetneq I_C \tag{41}$$

というように, $I_A, I_B, I_C$ は**互いに異なるイデアルになる**ことが分かります. すると, 問 2 の (5) で見たように, これらのイデアルは, 例えば, $A \rightsquigarrow P^{-1}AP$ という変換のもとで不変に保たれることが分かりますから, (41) 式から, $A, B, C$ という Jordan 標準形は, $\mathbb{C}^3$ の「番地割り」の仕方を取り換えただけではお互いに写り合うことができない**本質的に異なる標準形**であることが分かります.

実は,「行列の標準形の問題」を一般的に解決してみると, **3 行 3 列の行列 $D$ で, その特性多項式 $\varphi_D(x)$ が,**

$$\boldsymbol{\varphi_D(x) = (x - \lambda)^3}$$

**となるようなものは, $\mathbb{C}^3$ の基底を取り替えることで, 上の $A, B, C$ のうちのいずれかの標準形に変換できる**ということが分かります. すなわち, **$P^{-1}DP$ という行列が, 上の $A, B, C$ のうちのいずれかの標準形と等しくなるような正則行列 $P$ が存在する**ということが分かります. このとき,

**イデアル $I_D$ をもとにして, 行列 $D$ の標準形に「当たり」をつける**

$$I_{P^{-1}DP} = I_D$$

という性質を用いると, $I_A, I_B, I_C$ というイデアルは, それぞれ, (37) 式, (38) 式, (40) 式によって具体的に記述されていますから, **$I_D$ を求めてから, これらの結果と比べることで, $P^{-1}DP$ という行列が, 上の $A, B, C$ のうちのいずれの標準形と等しくなるようにできるのかということが, すなわち, 行列 $D$ がどの形の標準形に変換できるのかということが判定できることになります**.

このように, Jordan 標準形 $J$ に対して, 行列 $J$ に対応した一変数多項式環 $\mathbb{C}[x]$ のイデアル

$$I_J = \{ f(x) \in \mathbb{C}[x] \mid f(J) = O \}$$

がすべてリストアップできているとすると, **与えられた正方行列 $A$ に対して, 行列 $A$ に対応したイデアル $I_A$ を計算して, $I_J$ たちと比べてみることで, 行列 $A$ がどのような形の標準形に変形できるのかという候補に, ある程度の見当を付けることができます**. 特に, **$A$ が 2 行 2 列の行列や 3 行 3 列の行列の場合には**, 行列のサイズが小さいために Jordan 標準形の種類が比較的少ないという特殊性から, **特性多項式 $\varphi_A(x)$ と $I_A$ というイデアルを調べるだけで, 行列 $A$ がどの形の標準形に変換できるのかということが完全に決定できる**ことが分かります. こうした事柄については, 次回以降の演習で順番に触れて行くことにしようと思います.

## 7　最小多項式とは

さて, 5 節, 6 節では, 正方行列 $A$ に対して, $A$ を「根」に持つような多項式全体からなる一変数多項式環 $\mathbb{C}[x]$ のイデアル

$$I_A = \{ f(x) \in \mathbb{C}[x] \mid f(A) = O \}$$

が, 行列 $A$ の「固有な性質」を反映していることを見ました. 問 1 のところで見たように, Jordan 標準形と呼ばれる「見やすい形」をした行列 $J$ の場合には, $f(x) \in \mathbb{C}[x]$ として, 比較的簡単な計算で, 直接, $f(J)$ という行列を求めることできます. また, それにより, 直接, $I_J$ というイデアルの具体的な記述を得ることができます. ところが, 一般には, 勝手にひとつ与えられた正方行列 $A$ に対して, $n \in \mathbb{N}$, $f(x) \in \mathbb{C}[x]$ として, $A^n$ や $f(A)$ という行列を直接計算して求めることは大変困難なことになります. そこで, 行列 $A$ に対応したイデアル $I_A$ を記述するためには「別な工夫」が必要になりますが, そのためのアイデアは**イデアル $I_A$ の抽象的な構造に注目する**ということです.

そこで, イデアル $I_A$ がどのような構造を持つのかということに「当たり」を付けてみるために, 手始めに, 行列 $A$ が 1 行 1 列の場合にはどうなるのかということを考えてみることにします. いま, $a \in \mathbb{C}$ として, 行列 $A$ が

$$A = \begin{pmatrix} a \end{pmatrix}$$

という 1 行 1 列の行列であるとします. このとき, 勝手な多項式 $f(x) \in \mathbb{C}[x]$ に対して,

$$f(A) = \begin{pmatrix} f(a) \end{pmatrix}$$

となることが分かりますから, イデアル $I_A$ は,

$$I_A = \{ f(x) \in \mathbb{C}[x] \mid f(a) = 0 \}$$

というように表わせることが分かります. すると,

$$f(a) = 0 \iff f(x) = (x - a)g(x), \ ( {}^{\exists} g(x) \in \mathbb{C}[x] ) \tag{42}$$

となることが分かりますから,

$$\begin{aligned} I_A &= \{ (x-a)g(x) \in \mathbb{C}[x] \mid g(x) \in \mathbb{C}[x] \} \\ &= (x - a) \cdot \mathbb{C}[x] \end{aligned}$$

というように表わせることが分かります. すなわち,

$$f(A) = O \iff f(x) \text{ は } (x - a) \text{ で割り切れる.} \tag{43}$$

となることが分かります. いま, 1 行 1 列の行列 $A$ とその行列成分 $a$ を同一視して考えることにすると, (43) 式は,

**多項式 $f(x)$ が $x = a$ を根に持つための条件**

$$f(a) = 0 \iff f(x) \text{ は } (x-a) \text{ で割り切れる.} \tag{44}$$

という皆さん良くご存じの「多項式 $f(x)$ が $x = a$ を根に持つための条件」に他なりません.

そこで, 次に, 問 1 の例について見返してみることにします. 問 1 では,

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \tag{45}$$

という「見やすい形」の行列を取り上げて, $f(x) \in \mathbb{C}[x]$ に対して, $f(A)$ という行列を具体的に求めることで, イデアル $I_A$ に対する

$$I_A = \left\{ f(x) \in \mathbb{C}[x] \mid f(\lambda) = f'(\lambda) = f''(\lambda) = 0 \right\}$$

という具体的な記述を得ました. すると,

$$f(\lambda) = f'(\lambda) = f''(\lambda) = 0 \iff f(x) = (x-\lambda)^3 g(x), \ ( {}^{\exists} g(x) \in \mathbb{C}[x] ) \tag{46}$$

となることが分かりますから,[20] イデアル $I_A$ は,

$$\begin{aligned} I_A &= \left\{ (x-\lambda)^3 g(x) \in \mathbb{C}[x] \mid g(x) \in \mathbb{C}[x] \right\} \\ &= (x-\lambda)^3 \cdot \mathbb{C}[x] \end{aligned} \tag{47}$$

というように表わせることが分かります. すなわち,

**多項式 $f(x)$ が (45) 式の行列 $A$ を「根」に持つための条件**

$$f(A) = O \iff f(x) \text{ は } (x-\lambda)^3 \text{ で割り切れる.} \tag{48}$$

となることが分かります.

全く同様に,

$$B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \tag{49}$$

という行列を考えると, 問 1 の結果から, イデアル $I_B$ は,

$$I_B = \left\{ f(x) \in \mathbb{C}[x] \mid f(\lambda) = f'(\lambda) = 0 \right\}$$

というように具体的に記述できることが分かります. すると, 前と同様に,

$$f(\lambda) = f'(\lambda) = 0 \iff f(x) = (x-\lambda)^2 g(x), \ ( {}^{\exists} g(x) \in \mathbb{C}[x] )$$

[20] 皆さん, 確かめてみて下さい. ここで, 多項式 $f(x)$ を, Taylor 展開を用いて,

$$f(x) = f(\lambda) + f'(\lambda)(x-\lambda) + \frac{f''(\lambda)}{2!}(x-\lambda)^2 + \cdots + \frac{f^{(n)}(\lambda)}{n!}(x-\lambda)^n$$

というように表わして考えると, (46) 式の「$\Longrightarrow$」が確かめやすくなるかもしれません.

となることが分かりますから, イデアル $I_B$ は,

$$\begin{aligned} I_B &= \{ (x-\lambda)^2 g(x) \in \mathbb{C}[x] \mid g(x) \in \mathbb{C}[x] \ \} \\ &= (x-\lambda)^2 \cdot \mathbb{C}[x] \end{aligned} \tag{50}$$

というように表わせることが分かります. すなわち,

**多項式 $f(x)$ が (49) 式の行列 $B$ を「根」に持つための条件**

$$f(B) = O \iff f(x) \text{ は } (x-\lambda)^2 \text{ で割り切れる.} \tag{51}$$

となることが分かります.

以上から, 問 1 で考えた (45) 式, (49) 式で与えられる行列 $A, B$ の場合でも,「多項式 $f(x)$ が, それぞれの行列 $A, B$ を「根」に持つための条件」が,

$$f(A) = O \iff f(x) \text{ は } (x-\lambda)^3 \text{ で割り切れる.}$$
$$f(B) = O \iff f(x) \text{ は } (x-\lambda)^2 \text{ で割り切れる.}$$

というように, (44) 式という皆さん良くご存知の事実を一般化した形で理解できることが分かります. 以下で見るように, 実は,「見やすい形」とは限らない一般の正方行列 $A$ に対しても, 行列 $A$ から定まる 0 と異なる適当な多項式 $\psi_A(x)$ が存在して,

**多項式 $f(x)$ が正方行列 $A$ を「根」に持つための条件**

$$f(A) = O \iff f(x) \text{ は } \psi_A(x) \text{ で割り切れる.} \tag{52}$$

というように, (44) 式を一般化した形で「多項式 $f(x)$ が正方行列 $A$ を「根」に持つための条件」を理解することができることが分かります.

そこで, (52) 式の事実を確かめる前に, (44) 式という皆さん良くご存じの事実を確かめてみることにします. まず,「$\Longrightarrow$」という主張について考えてみます. いま, $f(x) \in \mathbb{C}[x]$ として,

$$f(a) = 0 \tag{53}$$

となると仮定してみます. このとき, 確かめたいことは, $f(x)$ が $(x-a)$ で割り切れるということになります. そこで, 実際に割り切れるかどうかということは後で検討することにして, 取りあえず, $f(x)$ を $(x-a)$ で「割り算」してみると,「商」を $g(x) \in \mathbb{C}[x]$,「余り」を $r \in \mathbb{C}$ として, $f(x)$ は,

$$f(x) = (x-a)g(x) + r \tag{54}$$

という形に表わせることが分かります. ここで, (53) 式に注意して, (54) 式の両辺で, $x = a$ としてみると,

$$\begin{aligned} 0 &= f(a) && \text{( (53) 式から )} \\ &= (a-a)g(a) + r && \text{( (54) 式から )} \\ &= r \end{aligned}$$

となることが分かりますから,

$$r = 0$$

となることが分かります. よって, (54) 式と合わせて,

$$f(x) = (x-a)g(x)$$

となることが分かりますから, $f(x)$ は $(x-a)$ で割り切れることが分かります. したがって,「$\Longrightarrow$」という主張が成り立つことが分かります.

次に,「$\Longleftarrow$」という主張について考えてみます. いま, $f(x)$ が $(x-a)$ で割り切れると仮定してみます. このとき, $f(x)$ は, 適当な多項式 $g(x) \in \mathbb{C}[x]$ を用いて,

$$f(x) = (x-a)g(x) \tag{55}$$

というように表わされることになりますが, (55) 式の両辺で, $x = a$ としてみると,

$$\begin{aligned} f(a) &= (a-a)g(a) \\ &= 0 \end{aligned}$$

となることが分かります. よって,「$\Longleftarrow$」という主張が成り立つことが分かります. 以上から, (44) 式が成り立つことが分かりました.

そこで, 行列 $A$ が一般の正方行列の場合にも (52) 式が成り立つことを確かめたいわけですが, そのために, まず, (52) 式が成り立つとすると, 多項式 $\psi_A(x)$ はどのような多項式でなければならないかということに「当たり」を付けてみることにします. いま, $f(x)$ が $\psi_A(x)$ で割り切れるということは, $f(x)$ が, 適当な多項式 $g(x) \in \mathbb{C}[x]$ を用いて,

$$f(x) = \psi_A(x)g(x) \tag{56}$$

という形に表わせるということですから, (52) 式が成り立つと仮定すると, イデアル $I_A$ は,

$$\begin{aligned} I_A &= \{\psi_A(x)g(x) \in \mathbb{C}[x] \mid g(x) \in \mathbb{C}[x]\ \} \\ &= \psi_A(x) \cdot \mathbb{C}[x] \end{aligned}$$

というように表わせることが分かります. また, (56) 式の両辺の多項式の次数を考えると, $g(x) \neq 0$ のとき,[21]

$$\begin{aligned} \deg f(x) &= \deg \psi_A(x) + \deg g(x) \\ &\geq \deg \psi_A(x) \qquad\qquad (\ \deg g(x) \geq 0 \text{ より }\ ) \end{aligned}$$

となることが分かりますから, **$\psi_A(x)$ は $I_A$ に属する 0 でない多項式のうち, 次数が最も低い多項式である**ということになります.

そこで, いま, $I_A$ に属する 0 でない多項式のうち, 次数が最も低い多項式を, 何でも良いから勝手にひとつ取ってきて,

$$\psi_A(x) \in I_A$$

[21] すなわち, $f(x) \neq 0$ のときということです.

と名前を付けることにします.[22] このとき, $A$ が 1 行 1 列の行列の場合と全く同様にして, こうして選んだ多項式 $\psi_A(x)$ に対して, (52) 式が成り立つことが, 次のように確かめられることが分かります.

まず,「$\Longrightarrow$」という主張について考えてみます. いま, $f(x) \in \mathbb{C}[x]$ として,

$$f(A) = O \tag{57}$$

となると仮定してみます. このとき, 確かめたいことは, $f(x)$ が $\psi_A(x)$ で割り切れるということになります. そこで, 実際に割り切れるかどうかということは後で検討することにして, 取りあえず, $f(x)$ を $\psi_A(x)$ で「割り算」してみると,「商」を $g(x) \in \mathbb{C}[x]$,「余り」を $r(x) \in \mathbb{C}[x]$ として, $f(x)$ は,

$$f(x) = \psi_A(x)g(x) + r(x), \quad (\deg r(x) < \deg \psi_A(x)) \tag{58}$$

という形に表わせることが分かります. ここで, (57) 式に注意して, (58) 式の両辺で, $x = A$ としてみると,

$$\begin{aligned} O &= f(A) && (\text{(57) 式から}) \\ &= \psi_A(A)g(A) + r(A) && (\text{(58) 式から}) \\ &= O \cdot g(A) + r(A) && (\psi_A(x) \in I_A \text{ から}) \\ &= r(A) \end{aligned}$$

となることが分かりますから,

$$r(x) \in I_A$$

となることが分かります. 一方, (58) 式から,

$$\deg r(x) < \deg \psi_A(x)$$

となることが分かりますが, $\psi_A(x)$ は $I_A$ に属する 0 でない多項式のうち, 次数が最も低い多項式として選んだのですから, このような多項式は 0 しかありえないことが分かります. よって,

$$r(x) = 0$$

となることが分かりますから, (58) 式と合わせて,

$$f(x) = \psi_A(x)g(x)$$

となることが分かります. すなわち, $f(x)$ は $\psi_A(x)$ で割り切れることが分かります. したがって,「$\Longrightarrow$」という主張が成り立つことが分かります.

次に,「$\Longleftarrow$」という主張について考えてみます. いま, $f(x)$ が $\psi_A(x)$ で割り切れると仮定してみます. このとき, $f(x)$ は, 適当な多項式 $g(x) \in \mathbb{C}[x]$ を用いて,

$$f(x) = \psi_A(x)g(x) \tag{59}$$

[22] 本当は, このような多項式 $\psi_A(x) \in I_A$ を取ることができるためには, $I_A \neq \{0\}$ となること, すなわち, $I_A$ に属する 0 以外の多項式が存在することを確かめる必要がありますが, 議論が混乱してしまわないように, この事実の確認作業は後回しにすることにします.

というように表わされることになりますが, (55) 式の両辺で, $x = A$ としてみると,

$$\begin{aligned} f(A) &= \psi_A(A) g(A) \\ &= O \cdot g(A) && (\ \psi_A(x) \in I_A \text{ から } ) \\ &= O \end{aligned}$$

となることが分かります. よって,「$\Longleftarrow$」という主張が成り立つことが分かります. 以上から, (52) 式が成り立つことが分かりました.

さて, 上で行なった議論では, $I_A$ に属する $0$ でない多項式のうち, 次数が最も低い多項式 (のうちのひとつ) を $\psi_A(x) \in I_A$ と定めましたが, 上でも注意したように, このような多項式 $\psi_A(x) \in I_A$ を取ることができるためには, $I_A \neq \{0\}$ となること, すなわち, $I_A$ に属する $0$ 以外の多項式が存在することを確かめる必要があります. そこで, この事実を確かめてみることにします. 後で見るように, **Cayley-Hamilton の定理**によれば, **このような多項式として, 行列 $A$ の特性多項式 $\varphi_A(x)$ を取ることができる**ということが分かるのですが, **このような多項式が存在するということだけ**であれば, 次のように議論することもできます.

いま, 勝手な自然数 $n \in \mathbb{N}$ に対して, 複素数を成分に持つ $n$ 行 $n$ 列の行列全体の集合を,

**$n$ 行 $n$ 列の複素行列全体の集合**

$$M_n(\mathbb{C}) = \{ A = (a_{ij})_{i,j=1,2,\cdots,n} \mid a_{ij} \in \mathbb{C} \ \}$$

という記号を用いて表わすことにします. このとき, $M_n(\mathbb{C})$ は $\mathbb{C}$ 上の線型空間となることが分かりますが, $n$ 行 $n$ 列の行列の成分は $n^2$ 個ありますから, その次元は,

$$\dim_{\mathbb{C}} M_n(\mathbb{C}) = n^2$$

となることが分かります. そこで, 与えられた行列 $A \in M_n(\mathbb{C})$ に対して, 例えば,

$$I = A^0,\ A^1,\ A^2,\ \cdots,\ A^{n^2}$$

という $(n^2+1)$ 個の行列を考えてみます. すると, $M_n(\mathbb{C})$ の次元は $n^2$ なのですから, これら $(n^2+1)$ 個の行列は線型独立ではあり得ないことが分かります. したがって,

$$\begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n^2} \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{60}$$

となる複素数 $a_0, a_1, a_2, \cdots, a_{n^2} \in \mathbb{C}$ が存在して,

$$a_0 I + a_1 A + \cdots + a_{n^2} A^{n^2} = O \tag{61}$$

となることが分かります. そこで,

$$f(x) = a_0 + a_1 x + \cdots + a_{n^2} x^{n^2} \in \mathbb{C}[x]$$

としてみると, (60) 式, (61) 式から,

$$f(x) \neq 0$$

であり,

$$\begin{aligned} f(A) &= a_0 I + a_1 A + \cdots + a_{n^2} A^{n^2} \\ &= O \end{aligned}$$

となることが分かります. よって,

$$0 \neq f(x) \in I_A$$

となることが分かりますから,

$$I_A \neq \{0\}$$

となることが分かります.

以上の議論から, 行列 $A$ を「根」に持つ多項式全体からなるイデアル $I_A$ は, $I_A$ に属する 0 でない多項式のうち, 次数が最も低い多項式 (のうちのひとつ) を $\psi_A(x)$ として,

**イデアル $I_A$ の記述**

$$\begin{aligned} I_A &= \{\psi_A(x) g(x) \in \mathbb{C}[x] \mid g(x) \in \mathbb{C}[x]\ \} \\ &= \psi_A(x) \cdot \mathbb{C}[x] \end{aligned} \tag{62}$$

というように記述できることが分かります. いま, $\hat{\psi}_A(x) \in I_A$ も, このような多項式であるとすると, (62) 式から, 適当な多項式 $g(x) \in \mathbb{C}[x]$ を用いて,

$$\hat{\psi}_A(x) = \psi_A(x) g(x)$$

と表わされることが分かりますが,

$$\deg \hat{\psi}_A(x) = \deg \psi_A(x)$$

であることに注意すると, $g(x)$ は定数でなければならないことが分かります. したがって, (62) 式を成り立たせるような多項式 $\psi_A(x)$ は 0 でない複素数を掛け算する不定性を除いて一意的に決まることが分かります. そこで, さらに, **多項式 $\psi_A(x)$ の $x$ に関する最高次の係数が 1 である**ということを要請すると,[23] このような多項式 $\psi_A(x)$ が唯一つ定まることが分かりますが, こうして定まる多項式 $\psi_A(x) \in \mathbb{C}[x]$ を行列 $A$ の**最小多項式**と呼びます.[24]

さて, 実際に, 行列 $A$ の最小多項式 $\psi_A(x)$ を求めるためには, $I_A$ の中で次数が最も低い多項式を探さないといけませんが, そのためには「別な工夫」が必要になります. その

[23] すなわち, $\deg \psi_A(x) = d$ として, 多項式 $\psi_A(x)$ が,

$$\psi_A(x) = x^d + a_{d-1} x^{d-1} + \cdots + a_0$$

という形をしているということです.

[24] 言葉の意味は, $f(A) = O$ となるような多項式の中で**次数が最小のもの**であるということです.

ための工夫のひとつが, 第 7 回の問 1 のところでも触れた **Cayley-Hamilton の定理**です. そのときにも述べましたが, Cayley-Hamilton の定理とは, **勝手にひとつ取ってきた正方行列 $A$ に対して, $A$ の特性多項式**

**行列 $A$ の特性多項式**

$$\boldsymbol{\varphi_A(x) = \det(xI - A)}$$

**に $A$ 自身を代入してみると, 常に,**

**Cayley-Hamilton の定理**

$$\boldsymbol{\varphi_A(A) = O}$$

**というように零行列になる**ということ, すなわち,

**Cayley-Hamilton の定理（言い換え）**

$$\boldsymbol{\varphi_A(x) \in I_A}$$

**となる**ということを主張する定理です. すると, (62) 式より, イデアル $I_A$ の元は, すべて, 最小多項式 $\psi_A(x)$ に適当な多項式を掛け算して得られることが分かりますから,

**特性多項式 $\varphi_A(x)$ は最小多項式 $\psi_A(x)$ で割り切れる**

$$\varphi_A(x) = g(x)\psi_A(x)$$

となるような多項式 $g(x) \in \mathbb{C}[x]$ が存在することが分かります. すなわち, **最小多項式 $\psi_A(x)$ は特性多項式 $\varphi_A(x)$ を割り切るような多項式である**ことが分かります.

この事実を用いると, 例えば, $A$ が 3 行 3 列の行列の場合というように, **行列 $A$ のサイズが余り大きくないときには, 特性多項式 $\varphi_A(x)$ を割り切るような多項式に行列 $A$ を代入して, その結果が零行列 $O$ になるかどうかを調べてみることで, 最小多項式を求めることができます**. 例えば, 問 1 の例では,

$$A = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad B = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix}$$

として,

$$\varphi_A(x) = (x - \lambda)^3$$
$$\varphi_B(x) = (x - \lambda)^3$$

となることが分かりますから, どちらの場合にも, 最小多項式は,

$$(x - \lambda),\ (x - \lambda)^2,\ (x - \lambda)^3 \in \mathbb{C}[x]$$

のうちのいずれかの多項式になることが分かります．ここで，

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

とすると，問 1 の (1) の解答で見たように，

$$\begin{aligned} (A - \lambda I) &= N \neq O, \\ (A - \lambda I)^2 &= N^2 \neq O, \\ (A - \lambda I)^3 &= N^3 = O \end{aligned}$$

となることが分かりますから，

$$\psi_A(x) = (x - \lambda)^3$$

となることが分かります．一方，

$$N' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

とすると，問 1 の (4) の解答で見たように，

$$\begin{aligned} (B - \lambda I) &= N' \neq O, \\ (B - \lambda I)^2 &= (N')^2 = O \end{aligned}$$

となることが分かりますから，

$$\psi_B(x) = (x - \lambda)^2$$

となることが分かります．したがって，

$$\begin{aligned} I_A &= (x - \lambda)^3 \cdot \mathbb{C}[x] \\ I_B &= (x - \lambda)^2 \cdot \mathbb{C}[x] \end{aligned}$$

となることが分かりますが，もちろん，これらの結果は，直接，$I_A, I_B$ を具体的に記述することにより求めた (47) 式，(50) 式と一致しています．

より一般には，**単因子論**というものを用いると，上のように特性多項式 $\varphi_A(x)$ を経由せずに，行列 $A$ から，直接，最小多項式 $\psi_A(x)$ を求めることもできるのですが，それを説明するためには少し準備が要りますから，ここでは省略することにします．

## 8 Cayley-Hamilton の定理について

そこで，次に，Cayley-Hamilton の定理について考えてみることにします．これまでにも何度か触れましたが，**Cayley-Hamilton の定理**とは，**勝手にひとつ取ってきた正方行列 $A$ に対して，$A$ の特性多項式**

**行列 $A$ の特性多項式**

$$\boldsymbol{\varphi_A(x) = \det(xI - A)} \tag{63}$$

**に行列 $A$ 自身を代入してみると,**

**Cayley-Hamilton の定理**

$$\boldsymbol{\varphi_A(A) = O} \tag{64}$$

**というように, いつでも零行列になる**ということを主張する定理です. 第 7 回の問 1 のところでも触れましたが, ここで, **(64) 式は, (63) 式の右辺に現われる $x$ のところに, 直接, 行列 $A$ を代入した**

**Cayley-Hamilton の定理に対するよくある誤解**

$$\boldsymbol{\det(A - A) = 0} \tag{65}$$

**という式とは意味が異なる**ということに注意して下さい. すなわち, **(65) 式は, 単に,**

$$\boldsymbol{(AI - A) = O}$$

**という零行列の行列式が 0 という「数」になるということを主張しているに過ぎない**のに対して, **(64) 式は, (63) 式の右辺の行列式を展開して, 特性多項式 $\varphi_A(x)$ を,**

$$\boldsymbol{\varphi_A(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad (\, a_0, a_1, \cdots, a_{n-1} \in \mathbb{C}\,)}$$

**というように多項式の姿に表わしてから, 行列 $A$ を代入することによって得られる**

$$\boldsymbol{\varphi_A(A) = A^n + a_{n-1}A^{n-1} + \cdots + a_1 A + a_0 I}$$

**という「行列」$\varphi_A(A)$ が $O$ という「零行列」になるということを主張している**わけです. このように, **(64) 式という Cayley-Hamilton の定理の主張は, 見かけほど当たり前の事実ではない**ことが分かります.

さて, 第 7 回の問 1 のところでは, $A$ が 2 行 2 列の行列である場合に, $\varphi_A(A)$ という行列を直接計算することにより, (64) 式の主張を確かめてみました. しかし, 一般のサイズの正方行列 $A$ に対して, $\varphi_A(A)$ という行列を直接計算することは大変困難ですから, (64) 式の主張を一般的に確かめるためには「別な工夫」が必要になります. そこで, ここでは, $P$ を正則行列として,

**特性多項式の共役不変性**

$$\varphi_{P^{-1}AP}(x) = \varphi_A(x) \tag{66}$$

となることに注目して,「行列の標準形の問題」の考え方を応用して, (64) 式を確かめてみることにします.

いま, 与えられた正方行列 $A$ に対して,

$$P^{-1}AP = \Lambda \tag{67}$$

となる「見やすい形」の行列 $\Lambda$ と正則行列 $P$ が見つかったと仮定してみます. すると, 問 2 の (4) で見たように, 勝手な多項式 $f(x) \in \mathbb{C}[x]$ に対して,

$$P^{-1} \cdot f(A) \cdot P = f(\Lambda) \tag{68}$$

となることが分かりますから, (68) 式の両辺に左から $P$ を掛け算し, 右から $P^{-1}$ を掛け算することで,

$$f(A) = P \cdot f(\Lambda) \cdot P^{-1} \tag{69}$$

となることが分かります. そこで, (66) 式に注意して, $f(x)$ として,

$$\begin{aligned} f(x) &= \varphi_A(x) \\ &= \varphi_\Lambda(x) && \text{( (66) 式から )} \end{aligned}$$

と取ってみると, (66) 式から,

$$\begin{aligned} \varphi_A(A) &= P \cdot \varphi_A(\Lambda) \cdot P^{-1} \\ &= P \cdot \varphi_\Lambda(\Lambda) \cdot P^{-1} \end{aligned} \tag{70}$$

となることが分かります. よって,「見やすい形」の行列 $\Lambda$ に対して,

$$\varphi_\Lambda(\Lambda) = O \tag{71}$$

となることが分かれば, (70) 式, (71) 式から,

$$\begin{aligned} \varphi_A(A) &= P \cdot \varphi_\Lambda(\Lambda) \cdot P^{-1} && \text{( (70) 式から )} \\ &= P \cdot O \cdot P^{-1} && \text{( (71) 式から )} \\ &= O \end{aligned}$$

となることが分かりますから, (64) 式が成り立つことが分かります. したがって, **一般の正方行列 $A$ に対して, Cayley-Hamilton の定理が成り立つことを確かめるためには,「見やすい形」の行列 $\Lambda$ に対して, Cayley-Hamilton の定理が成り立つことを確かめればよい**ということが分かります.

このことは, 次のように解釈することができます. いま, 5 節と同様に, $\mathbb{C}^n$ から最初の座標軸を消し去って, $\mathbb{C}^n$ を線型空間と思ったものを,

$$V = \mathbb{C}^n$$

と表わし, $n$ 行 $n$ 列の行列 $A$ を掛け算することによって定まる線型写像を, 単に,

$$L : V \to V$$

と表わすことにします. また,

$$\begin{array}{ccc} (\mathbb{C}^n)_{\text{旧}} & \xrightarrow{A} & (\mathbb{C}^n)_{\text{旧}} \\ \wr\!\shortparallel & & \wr\!\shortparallel \\ V & \xrightarrow{L} & V \\ \wr\!\shortparallel & & \wr\!\shortparallel \\ (\mathbb{C}^n)_{\text{新}} & \xrightarrow{\Lambda} & (\mathbb{C}^n)_{\text{新}} \end{array}$$

というように, 線型空間 $V = \mathbb{C}^n$ の最初の「旧番地割り」に関する線型写像 $L$ の表現行列が $A$ で, 正則行列 $P$ の列ベクトル $\{\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_n\}$ を基底とする「新番地割り」に関する表現行列が $\Lambda$ であると考えてみることにします. ここで, $\varphi_A(x) = \varphi_\Lambda(x)$ を, 単に, $\varphi(x) \in \mathbb{C}[x]$ と表わすことにすると, 5 節で注意したように,

$$\begin{array}{ccc} (\mathbb{C}^n)_{\text{旧}} & \xrightarrow{\varphi(A)=\varphi_A(A)} & (\mathbb{C}^n)_{\text{旧}} \\ \wr\| & & \wr\| \\ V & \xrightarrow{\varphi(L)} & V \\ \wr\| & & \wr\| \\ (\mathbb{C}^n)_{\text{新}} & \xrightarrow{\varphi(\Lambda)=\varphi_\Lambda(\Lambda)} & (\mathbb{C}^n)_{\text{新}} \end{array}$$

というように, 線型写像 $\varphi(L)$ の「旧番地割り」に関する表現行列が $\varphi_A(A)$ となり,「新番地割り」に関する表現行列が $\varphi_\Lambda(\Lambda)$ となることが分かります. よって,

$$\varphi_A(A) = O \iff \varphi(L) = 0$$

となることが分かりますから,

**Cayley-Hamilton の定理（もともとの「番地割り」を用いた表示）**

$$\varphi_A(A) = O$$

という行列 $A$ に対する Cayley-Hamilton の定理が成り立つことを確かめるためには,

**Cayley-Hamilton の定理（特定の「番地割り」に依らない表示）**

$$\varphi(L) = 0 \tag{72}$$

という線型写像 $L : V \to V$ に対する Cayley-Hamilton の定理が成り立つことを確かめればよいということが分かります. さらに,

$$\varphi(L) = 0 \iff \varphi_\Lambda(\Lambda) = O$$

となることに注意すると, (72) 式が成り立つことを確かめるためには, 結局,

**Cayley-Hamilton の定理（「上手い番地割り」を用いた表示）**

$$\varphi_\Lambda(\Lambda) = O \tag{73}$$

というように,「見やすい形」の行列 $\Lambda$ に対して, Cayley-Hamilton の定理が成り立つことを確かめればよいということが分かります. このように, $\varphi(L) : V \to V$ という線型写像を仲立ちにして,

$$\begin{aligned} \varphi_A(A) = O &\iff \varphi(L) = 0 \\ &\iff \varphi_\Lambda(\Lambda) = O \end{aligned}$$

というように考えてみると, **一般の正方行列 $A$ に対して, Cayley-Hamilton の定理が成り立つことを確かめるためには,「見やすい形」の行列 $\Lambda$ に対して, Cayley-Hamilton**

**の定理が成り立つことを確かめればよい**ということがより納得しやすくなるかもしれません.

さて,「見やすい形」の行列 $\Lambda$ が対角行列のときには, (73) 式が成り立つことは簡単に確かめることができます. 例えば,

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

であるとすると, 行列 $\Lambda$ の特性多項式 $\varphi_\Lambda(x)$ は,

$$\varphi_\Lambda(x) = (x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$$

となることが分かります. よって,

$$\begin{aligned} \varphi_\Lambda(\Lambda) &= (\Lambda - \lambda_1 I)(\Lambda - \lambda_2 I)(\Lambda - \lambda_3 I) \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix} \begin{pmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & * \end{pmatrix} \begin{pmatrix} * & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & * & 0 \\ 0 & 0 & * \end{pmatrix} \begin{pmatrix} * & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

となることが分かりますから, (73) 式が成り立つことが分かります.[25]

このように, $\Lambda$ が対角行列のときには, (73) 式が成り立つことは簡単に確かめることができるわけですが, 第 9 回の問 3 のところや 6 節で見たように, 一般の正方行列 $A$ は対角化できるとは限りませんから, 勝手な正方行列 $A$ に対して, Cayley-Hamilton の定理が成り立つことを上のような戦略で確かめるためには,「見やすい形」をもう少し拡張して考える必要があります. 以下で見るように, そのためには,「見やすい形」の行列として,

**上三角行列**

$$\underbrace{\begin{pmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & * & * \\ 0 & 0 & \cdots & 0 & * \end{pmatrix}}_{n\ \text{コ}} \tag{74}$$

という**上三角行列**を採用すれば十分であることが分かります.

[25] ここで, 計算が見やすくなるように, 本質的でない行列成分は「$*$」という記号を用いて象徴的に表わしました.

そこで, **与えられた正方行列 $A$ に対して,**

$$\boldsymbol{P^{-1}AP = \Lambda}$$

**となる上三角行列 $\Lambda$ と正則行列 $P$ が存在する**ということの確認作業は後に回すことにして, まずは, 上三角行列 $\Lambda$ に対して, (73) 式が成り立つことを確かめてみることにします. 議論の本質は一般のサイズの上三角行列の場合でも同じですから, 話を具体的にするために, ここでは, $\Lambda$ が 3 行 3 列の上三角行列である場合に, (73) 式が成り立つことを確かめてみることにします.

いま,

$$\Lambda = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

であるとすると, 行列 $\Lambda$ の特性多項式 $\varphi_\Lambda(x)$ は,

$$\varphi_\Lambda(x) = (x-\lambda_1)(x-\lambda_2)(x-\lambda_3)$$

となることが分かります. よって,

$$\begin{aligned} \varphi_\Lambda(\Lambda) &= (\Lambda-\lambda_1 I)(\Lambda-\lambda_2 I)(\Lambda-\lambda_3 I) \\ &= \begin{pmatrix} 0 & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}\begin{pmatrix} * & * & * \\ 0 & 0 & * \\ 0 & 0 & * \end{pmatrix}\begin{pmatrix} * & * & * \\ 0 & * & * \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & * & * \\ 0 & * & * \\ 0 & 0 & * \end{pmatrix}\begin{pmatrix} * & * & * \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{aligned} \tag{75}$$

となることが分かりますから, (73) 式が成り立つことが分かります. 全く同様に, $\Lambda$ が一般のサイズの上三角行列の場合でも, (75) 式の右辺に現われる行列の積を後ろの方から計算してみると, 掛け算をするたびに,

$$\begin{pmatrix} 0 & 0 & \cdots & 0 \end{pmatrix}$$

という 0 だけが並んだ行が一段ずつ上の方に「侵攻」してきて, 最終的には,

$$\varphi_\Lambda(\Lambda) = O$$

となることが分かります.[26]

[26]皆さん, 確かめてみて下さい. 例えば, $\Lambda$ が $n$ 行 $n$ 列の上三角行列であるとして, $k = 0, 1, \cdots, (n-1)$ に対して,「掛け算を $k$ 回行なって得られる行列は, 下から $k+1$ 番目までの行の行列成分がすべて 0 になる」という主張を数学的帰納法を用いて証明するという方針を取るとスッキリと議論できるかもしれません.

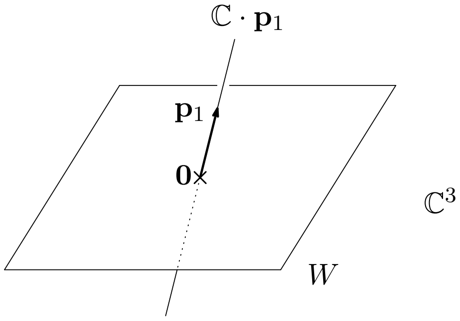



図 1: $\mathbb{C}^3$ を「$\mathbb{C}\cdot\mathbf{p}_1$ の方向」と「$W$ の方向」に直和分解してみる.

そこで, 次に, **与えられた正方行列 $A$ に対して,**

$$\boldsymbol{P^{-1}AP=\Lambda}$$

**となる上三角行列 $\boldsymbol{\Lambda}$ と正則行列 $\boldsymbol{P}$ が存在する**ということを確かめてみることにします. 上と同様に, 話を具体的にするために, ここでは, $A$ が 3 行 3 列の行列である場合に説明してみることにします.

そこで, いま, 3 行 3 列の行列 $A$ が, 勝手にひとつ与えられているとします. このとき, $A$ の固有値を, 勝手にひとつ選んできて,

$$\lambda_1\in\mathbb{C}$$

と名前を付けることにします. また, 固有値 $\lambda_1$ に対応した固有ベクトル

$$\mathbf{p}_1\in\mathbb{C}^3$$

を, 勝手にひとつ取ってきて, $\mathbf{p}_1$ の定める一次元の線型部分空間

$$\mathbb{C}\cdot\mathbf{p}_1=\{a\mathbf{p}_1\in\mathbb{C}^3\mid a\in\mathbb{C}\ \}\subset\mathbb{C}^3$$

と表わすことにします. さらに, 線型部分空間 $\mathbb{C}\cdot\mathbf{p}_1$ の「補空間」$W\subset\mathbb{C}^3$ を, 勝手にひとつ取ってきて, $\mathbb{C}^3$ を,

**$\mathbb{C}^3$ を「$\mathbb{C}\cdot\mathbf{p}_1$ の方向」と「それ以外の方向」に分解して考えてみる**

$$\mathbb{C}^3=\mathbb{C}\cdot\mathbf{p}_1\oplus W \tag{76}$$

というように「$\mathbb{C}\cdot\mathbf{p}_1$ の方向」と「$W$ の方向」に直和分解してみます ( 図 1 を参照 ). 例えば, 線型部分空間 $\mathbb{C}\cdot\mathbf{p}_1$ の「補空間」$W$ としては, 第 10 回の問 1 のところと同様にして, $\mathbb{C}^3$ 上の標準的なエルミート内積 ( , ) を用いて,

$$\begin{aligned}W&=(\mathbb{C}\cdot\mathbf{p}_1)^{\perp}\\&=\{\mathbf{u}\in\mathbb{C}^3\mid(\mathbf{u},\mathbf{p}_1)=0\ \}\end{aligned}$$

というように, $\mathbb{C}\cdot\mathbf{p}_1$ の直交補空間を取ることができます. あるいは, $\{\mathbf{p}_1,\mathbf{f}_2,\mathbf{f}_3\}$ が $\mathbb{C}^3$ の基底になるように, $\mathbf{p}_1$ に $\mathbb{C}^3$ のベクトル $\mathbf{f}_2,\mathbf{f}_3\in\mathbb{C}^3$ を, 勝手に二つ付け加えて,

$$W=\mathbb{C}\cdot\mathbf{f}_2\oplus\mathbb{C}\cdot\mathbf{f}_3$$

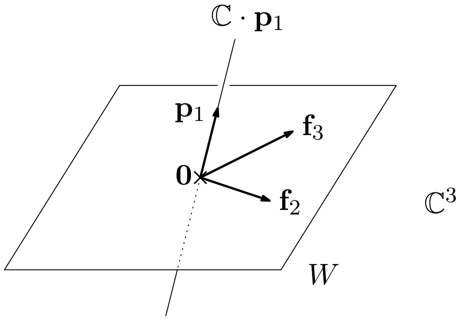



図 2: 「補空間」$W$ の基底 $\{\mathbf{f}_2, \mathbf{f}_3\}$ を, 勝手にひとつ選んできて, $\{\mathbf{p}_1, \mathbf{f}_2, \mathbf{f}_3\}$ という $\mathbb{C}^3$ の基底を考えてみる.

というように, $\mathbf{f}_2, \mathbf{f}_3$ が定める線型部分空間を取ることもできます.

そこで, (76) 式の直和分解のもとで,「補空間」$W$ の基底 $\{\mathbf{f}_2, \mathbf{f}_3\}$ を, 勝手にひとつ選んできて, $\mathbb{C}^3$ の基底として, $\{\mathbf{p}_1, \mathbf{f}_2, \mathbf{f}_3\}$ というベクトルたちを選んでみます ( 図 2 を参照 ). ここで, $\mathbf{p}_1$ は行列 $A$ の固有値 $\lambda_1$ に対応する固有ベクトルでしたから,

$$A\mathbf{p}_1 = \lambda_1\mathbf{p}_1 \tag{77}$$

となることが分かります. したがって,

**(76) 式の直和分解のもとで $L_A$ の表現行列の一列目が三角化される**

$$\begin{pmatrix} A\mathbf{p}_1 & A\mathbf{f}_2 & A\mathbf{f}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{f}_2 & \mathbf{f}_3 \end{pmatrix} \begin{pmatrix} \lambda_1 & * & * \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

となることが分かりますから, 行列 $A$ を掛け算することにより定まる線型写像を,

$$L_A : \mathbb{C}^3 \to \mathbb{C}^3$$

と表わすことにすると, この基底に関する線型写像 $L_A$ の表現行列は, 一列目の二行目以下がすべて 0 になるような行列となることが分かります.

さて, 第 10 回の問 1 のところで「対称行列の対角化の問題」を考えたときとは異なって, **一般には,「補空間」$W$ は, 行列 $A$ を掛け算する操作で不変であるとは限らない**ということに注意します. すなわち, 勝手なベクトル $\mathbf{v} \in W$ に対して,

$$A\mathbf{v} \in W$$

となるとは限らないということに注意します. そこで, $\mathbf{v} \in W$ に対して, $A\mathbf{v} \in \mathbb{C}^3$ というベクトルを「$\mathbb{C}\cdot\mathbf{p}_1$ の方向」と「$W$ の方向」に分解して,

**$A\mathbf{v} \in \mathbb{C}^3$ を「 $\mathbb{C}\cdot\mathbf{p}_1$ の方向」と「それ以外の方向」に分解して考えてみる**

$$A\mathbf{v} = \alpha(\mathbf{v})\cdot\mathbf{p}_1 + L'(\mathbf{v}) \ \in \mathbb{C}\cdot\mathbf{p}_1 \oplus W$$

というように表わすことにします ( 図 3 を参照 ). このとき, $\alpha$, $L'$ は, それぞれ,

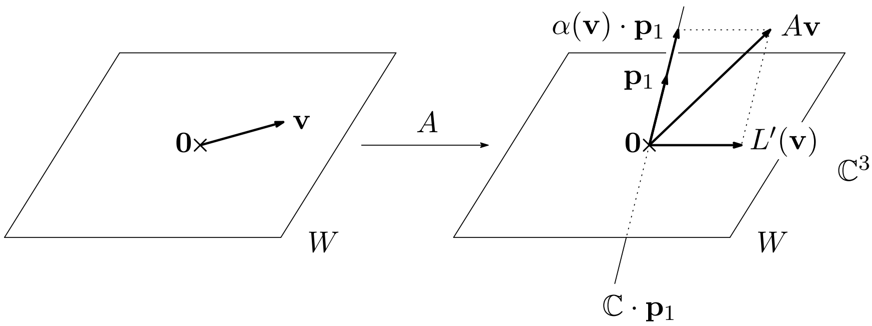



図 3: $\mathbf{v}\in W$ に対して, $A\mathbf{v}\in\mathbb{C}^3$ というベクトルを「$\mathbb{C}\cdot\mathbf{p}_1$ の方向」の成分 $\alpha(\mathbf{v})\cdot\mathbf{p}_1$ と「$W$ の方向」の成分 $L'(\mathbf{v})$ に分解してみる.

$$\begin{cases}\alpha : W \to \mathbb{C}\\ L' : W \to W\end{cases}$$

という線型写像となることが分かります.[27] そこで, いま, 線型写像

$$L' : W \to W$$

の固有値を, 勝手にひとつ取ってきて,

$$\lambda_2 \in \mathbb{C}$$

と名前を付けることにします. また, $L'$ の固有値 $\lambda_2$ に対応した固有ベクトル

$$\mathbf{p}_2 \in W$$

を, 勝手にひとつ取ってくることにします. さらに, $\{\mathbf{p}_2, \mathbf{p}_3\}$ が $W$ の基底になるように,

$$\mathbf{p}_3 \in W$$

を, 勝手にひとつ取ってくることにします. すると, このとき,

$$L'\mathbf{p}_2 = \lambda_2\mathbf{p}_2$$

となることが分かりますから,

$$\begin{aligned} A\mathbf{p}_2 &= \alpha(\mathbf{p}_2)\cdot\mathbf{p}_1 + L'(\mathbf{p}_2)\\ &= \alpha(\mathbf{p}_2)\cdot\mathbf{p}_1 + \lambda_2\mathbf{p}_2 \end{aligned} \tag{78}$$

となることが分かります. したがって, (77) 式, (78) 式から,

**基底 $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ に関して $L_A$ の表現行列は三角化される**

$$\begin{pmatrix} A\mathbf{p}_1 & A\mathbf{p}_2 & A\mathbf{p}_3 \end{pmatrix} = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{pmatrix}\begin{pmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & * \end{pmatrix} \tag{79}$$

となることが分かりますから, $\mathbb{C}^3$ の基底 $\{\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3\}$ に関する線型写像 $L_A$ の表現行列

[27] 皆さん, 確かめてみて下さい.

は, 上三角行列となることが分かります. よって, (79) 式から,

$$P = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{pmatrix}$$

として,

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

というように, 行列 $A$ は正則行列 $P$ を用いて,「上三角化」されることが分かります.

ここでは, $A$ が 3 行 3 列の行列の場合に説明しましたが, 例えば, **「$\mathbb{C}$ 上の線型空間 $\boldsymbol{V}$ と $\boldsymbol{V}$ 上の線型写像**

$$\boldsymbol{L : V \to V}$$

**に対して, 線型写像 $\boldsymbol{L}$ の表現行列が上三角行列となるような $\boldsymbol{V}$ の基底が存在する」**という形に問題を定式化してみると, $n = \dim_{\mathbb{C}} V$ に関する数学的帰納法を用いて, この主張を証明することができます.[28] したがって, $n$ 行 $n$ 列の行列 $A$ に対して,

$$\begin{aligned} V &= \mathbb{C}^n \\ L &= L_A : \mathbb{C}^n \to \mathbb{C}^n \end{aligned}$$

として, 上の主張を適用することで,

$$P^{-1}AP = \Lambda$$

となる上三角行列 $\Lambda$ と正則行列 $P$ が存在することが分かります. よって, 上で見たように,

$$\varphi_\Lambda(\Lambda) = O$$

となることから,

$$\varphi_A(A) = O$$

となることが分かりますから, 一般の正方行列 $A$ に対して, Cayley-Hamilton の定理が成り立つことが分かります.

## 9 一変数多項式環のイデアルについて *

さて, 7 節では, 正方行列 $A$ に対して, 行列 $A$ を「根」に持つような多項式全体からなるイデアル

$$I_A = \{ f(x) \in \mathbb{C}[x] \mid f(A) = O \}$$

が, 行列 $A$ の最小多項式 $\psi_A(x)$ を用いて,

$$\begin{aligned} I_A &= \{ \psi_A(x) g(x) \in \mathbb{C}[x] \mid g(x) \in \mathbb{C}[x] \} \\ &= \psi_A(x) \cdot \mathbb{C}[x] \end{aligned}$$

[28] 興味のある方は, 上で行なった議論を参考にして, 証明してみて下さい.

というように記述できることを見ました．そこで行なった議論を見返すと, 実は, 一変数多項式環 $\mathbb{C}[x]$ の勝手なイデアルに対して, 同様の議論を行なうことができることが分かります．そこで, この節では, **一変数多項式環 $\mathbb{C}[\boldsymbol{x}]$ のイデアルとはどのような部分集合であるのか**ということを考えてみることにします．そのために, まず, 一般に, 可換環 $R$ が, 勝手にひとつ与えられているとして, $R$ のイデアルの例を作る方法について考えてみることにします.[29]

いま, 可換環 $R$ のイデアル $I$ が, 勝手にひとつ与えられているとします．このとき,

$$y_0 \in I$$

という $I$ の元を, 勝手にひとつ取ってくると,

**イデアルの条件**

$$x, y \in I,\ z \in R \implies x + y,\, xz \in I \tag{80}$$

というイデアルの条件から, $R$ の勝手な元 $x \in R$ に対して,

$$xy_0 \in I$$

となることが分かります．すなわち, $y_0 \in R$ に対して,

**$\boldsymbol{y_0} \in \boldsymbol{R}$ が生成するイデアル**

$$\langle\, y_0 \,\rangle = \{xy_0 \in R \mid x \in R\,\}$$

という $R$ の部分集合を考えてみると,

**$\boldsymbol{y_0}$ が $\boldsymbol{I}$ の元なら, $\langle\, \boldsymbol{y_0} \,\rangle$ 全体がイデアル $\boldsymbol{I}$ に含まれる**

$$y_0 \in I \implies \langle\, y_0 \,\rangle \subset I \tag{81}$$

というように, $y_0$ が $I$ の元なら, $\langle\, y_0 \,\rangle$ 全体がすっぽりとイデアル $I$ に含まれてしまうことが分かります.

次に, より一般に, $y_0, y_1, \cdots, y_n \in R$ に対して,

**$\boldsymbol{y_0, y_1, \cdots, y_n} \in \boldsymbol{R}$ が生成するイデアル**

$$\langle\, y_0, y_1, \cdots, y_n \,\rangle = \{x_0y_0 + x_1y_1 + \cdots + x_ny_n \in R \mid x_0, x_1, \cdots, x_n \in R\,\} \tag{82}$$

という $R$ の部分集合を考えてみます．すなわち, $\langle\, y_0, y_1, \cdots, y_n \,\rangle$ は, $x_0, x_1, \cdots, x_n \in R$ を用いて,

$$x_0y_0 + x_1y_1 + \cdots + x_ny_n$$

という形に表わされるような $R$ の元全体からなる部分集合のことです．このとき, 前と同様に, $y_0, y_1, \cdots, y_n \in I$ であるとすると, (80) 式から, $R$ の勝手な元 $x_0, x_1, \cdots, x_n \in R$ に対して,

$$x_0y_0, x_1y_1, \cdots, x_ny_n \in I \tag{83}$$

[29] これでは抽象的過ぎると思われる方は, $R$ として, 一変数多項式環 $\mathbb{C}[x]$ だけを考えてもらっても構いません.

となることが分かります．すると，再び，(80) 式から，$I$ は足し算に関して閉じた集合でもありますから，(83) 式と合わせて，

$$x_0y_0 + x_1y_1 + \cdots + x_ny_n \in I$$

となることが分かります．したがって，

**$y_0, y_1, \cdots, y_n$ が $I$ の元なら，$\langle\, y_0, y_1, \cdots, y_n \,\rangle$ 全体がイデアル $I$ に含まれる**

$$y_0, y_1, \cdots, y_n \in I \implies \langle\, y_0, y_1, \cdots, y_n \,\rangle \subset I \tag{84}$$

というように，$y_0, y_1, \cdots, y_n$ が $I$ の元なら，$\langle\, y_0, y_1, \cdots, y_n \,\rangle$ 全体がすっぽりとイデアル $I$ に含まれてしまうことが分かります．一方，(82) 式という定義に戻って考えてみると，$\langle\, y_0, y_1, \cdots, y_n \,\rangle$ という部分集合自身がイデアルの条件を満たすことが分かりますから，[30] (84) 式と合わせて考えると，**$\langle\, y_0, y_1, \cdots, y_n \,\rangle$ とは，$y_0, y_1, \cdots, y_n \in R$ を含むようなイデアルの中で，包含関係 $\subset$ に関して最小のイデアルである**ことが分かります．この意味で，$\langle\, y_0, y_1, \cdots, y_n \,\rangle$ のことを，**$y_0, y_1, \cdots, y_n$ の生成するイデアル**と呼んだりします．このような**有限個の元で生成されるイデアル**が可換環 $R$ のイデアルの代表例であり，普通は，こうした**有限生成のイデアル**[31]しか扱いません．

以上の準備のもとで，**一変数多項式環 $\mathbb{C}[x]$ のイデアルとはどのような部分集合になるのか**ということを考えてみることにします．以下で見るように，この場合，状況はとても簡単になっていて，7 節で $I_A$ というイデアルを調べたときと同様に，実は，**一変数多項式環 $\mathbb{C}[x]$ のすべてのイデアル $I$ は，適当な多項式 $\psi(x) \in \mathbb{C}[x]$ を用いて，**

**$\mathbb{C}[x]$ のイデアル $I$ はすべて単項生成になる**

$$\begin{aligned} \boldsymbol{I} &= \langle\, \boldsymbol{\psi(x)} \,\rangle \\ &= \{\boldsymbol{\psi(x)g(x)} \in \mathbb{C}[\boldsymbol{x}] \mid \boldsymbol{g(x)} \in \mathbb{C}[\boldsymbol{x}]\ \} \\ &= \boldsymbol{\psi(x)} \cdot \mathbb{C}[\boldsymbol{x}] \end{aligned} \tag{85}$$

**という形に表わせる**ということが分かります．[32] このことは，例えば，多項式 $\varphi_0(x), \varphi_1(x) \in \mathbb{C}[x]$ を，勝手に二つ取ってきて，$\varphi_0(x), \varphi_1(x)$ の生成するイデアル $\langle\, \varphi_0(x), \varphi_1(x) \,\rangle$ を考えてみても，結局，$\langle\, \varphi_0(x), \varphi_1(x) \,\rangle$ は，適当な多項式 $\psi(x)$ を用いて，

$$\langle\, \varphi_0(x), \varphi_1(x) \,\rangle = \langle\, \psi(x) \,\rangle$$

という形に表わされてしまうということを意味しています．いま，

$$I = \{0\}$$

であるとすると，イデアル $I$ は，$0 \in \mathbb{C}[x]$ という定数関数を用いて，

$$I = \langle\, 0 \,\rangle$$

---

[30] 皆さん，確かめてみて下さい．

[31] すなわち，「有限個の元で生成されるイデアル」ということです．

[32] 記号が少し紛らわしいですが，多項式という感じを出すために，$y_0 \rightsquigarrow \psi(x)$ と書き換えて，$\langle\, y_0 \,\rangle$ を $\langle\, \psi(x) \,\rangle$ というように表わすことにしました．また，(85) 式の事実を，一変数多項式環 $\mathbb{C}[x]$ のイデアルはすべて**単項生成である**と言ったりします．単項生成とは「ひとつの元 $\psi(x) \in \mathbb{C}[x]$ で生成される」ということです．

という形に表わせることが分かります．したがって，この場合には，$\psi(x)=0$ として，(85) 式が成り立つことが分かりますから，以下では，もっぱら，**一変数多項式環 $\mathbb{C}[x]$ の零でないイデアル $I \neq \{0\}$ について考えてみる**ことにします．

そこで，いま，一変数多項式環 $\mathbb{C}[x]$ の零でないイデアル $I \neq \{0\}$ が，勝手にひとつ与えられているとします．このとき，7 節で $I_A$ というイデアルを調べたときと全く同様にして，イデアル $I$ に属する 0 でない多項式の中で，次数が最も低い多項式を，何でも良いから勝手にひとつ取ってきて，

**イデアル $I$ の 0 でない最も次数の低い元を $\psi(x)$ と名付ける**

$$\psi(x) \in I$$

とすると，イデアル $I$ は (85) 式のように記述できることが，次のようにして分かります．[33]

いま，イデアル $I$ に属する多項式 $f(x) \in I$ が，勝手にひとつ与えられているとします．このとき，実際に割り切れるかどうかということは後で検討することにして，取りあえず，$f(x)$ を $\psi(x)$ で「割り算」してみると，「商」を $g(x) \in \mathbb{C}[x]$，「余り」を $r(x) \in \mathbb{C}[x]$ として，$f(x)$ は，

$$f(x) = \psi(x)g(x) + r(x), \quad (\deg r(x) < \deg \psi(x)) \tag{86}$$

という形に表わせることが分かります．そこで，いま，(86) 式を，

$$r(x) = f(x) - \psi(x)g(x)$$

という形に書き直してみます．このとき，$\psi(x) \in I$ であるということと，$I$ はイデアルであるということから，

$$-\psi(x)g(x) = \psi(x) \cdot (-g(x)) \in I \tag{87}$$

となることが分かります．[34] よって，$f(x) \in I$ であることと合わせると，(87) 式から，

$$r(x) = f(x) + \{-\psi(x)g(x)\} \in I$$

となることが分かります．[35] すなわち，$r(x)$ もイデアル $I$ に属する多項式であることが分かります．一方，(86) 式から，

$$\deg r(x) < \deg \psi(x)$$

[33] 7 節で行なったイデアル $I_A$ に対する議論との唯一の違いは，7 節で「$f(A)=O$」という条件を用いて議論していた部分を，すべて「$f(x) \in I_A$」という条件で置き換えて議論するということです．

[34] $I = I_A$ の場合であれば，$\psi_A(A) = O$ となることから，

$$\begin{aligned} -\psi(A)g(A) &= O \cdot g(A) \\ &= O \end{aligned}$$

となることが分かるということです．

[35] これも，$I = I_A$ の場合であれば，$f(A) = O$，$-\psi(A)g(A) = O$ となることから，

$$\begin{aligned} r(A) &= f(A) + \{-\psi(A)g(A)\} \\ &= O + O \\ &= O \end{aligned}$$

となることが分かるということです．

となることが分かりますが, $\psi(x)$ はイデアル $I$ に属する 0 でない多項式の中で最も次数の低いものとして選んだのですから, このような多項式は 0 しかありえないことが分かります. よって,

$$r(x) = 0$$

となることが分かりますから, (86) 式と合わせて,

**イデアル $I$ に属する勝手な多項式 $f(x)$ は $\psi(x)$ で割り切れる**

$$f(x) = \psi(x)g(x) \tag{88}$$

となることが分かります. すなわち, **イデアル $I$ に属する勝手な多項式 $f(x)$ は $\psi(x)$ で割り切れる**ことが分かります. よって, (88) 式から,

$$f(x) \in I \implies f(x) \in \langle\, \psi(x) \,\rangle$$

となることが分かりますから,

$$I \subset \langle\, \psi(x) \,\rangle \tag{89}$$

となることが分かります. 一方, (81) 式より,

$$\langle\, \psi(x) \,\rangle \subset I$$

となることが分かりますから, (89) 式と合わせて,

$$I = \langle\, \psi(x) \,\rangle$$

となることが分かります. こうして, イデアル $I$ は (85) 式のように記述できることが分かりました.

次に, **(85) 式を成り立たせるような $I$ の生成元 $\psi(x)$ の取り方にどれだけの自由度があるのか**ということを考えてみることにします. そこで, いま, $\psi_1(x), \psi_2(x) \in I$ として, イデアル $I$ が,

$$\begin{aligned} I &= \langle\, \psi_1(x) \,\rangle \\ &= \langle\, \psi_2(x) \,\rangle \end{aligned}$$

というように二通りに記述できたと仮定してみます. このとき,

$$\psi_2(x) \in I = \langle\, \psi_1(x) \,\rangle$$

であると考えてみると, $\langle\, \psi_1(x) \,\rangle$ の定義から, 適当な多項式 $g_1(x) \in \mathbb{C}[x]$ が存在して,

$$\psi_2(x) = \psi_1(x)g_1(x) \tag{90}$$

というように表わせることが分かります. 一方,

$$\psi_1(x) \in I = \langle\, \psi_2(x) \,\rangle$$

であると考えてみると, $\langle\, \psi_2(x) \,\rangle$ の定義から, 適当な多項式 $g_2(x)$ が存在して,

$$\psi_1(x) = \psi_2(x) g_2(x) \tag{91}$$

というように表わせることが分かります. そこで, (90) 式を (91) 式に代入してみると,

$$\psi_1(x) = \psi_1(x) g_1(x) g_2(x)$$

となることが分かりますから,

$$\psi_1(x)(g_1(x)g_2(x) - 1) = 0 \tag{92}$$

となることが分かります. いま,

$$I \neq \{0\}$$

と仮定していたので,

$$\psi_1(x) \neq 0$$

となることに注意すると, (92) 式より,

$$g_1(x) g_2(x) = 1 \tag{93}$$

となることが分かります. ここで, (93) 式の両辺に現われる多項式の次数を比べてみると,

$$\deg g_1(x) + \deg g_2(x) = 0 \tag{94}$$

となることが分かりますが,

$$\deg g_1(x),\ \deg g_2(x) \geq 0$$

となることに注意すると, (94) 式から,

$$\deg g_1(x) = \deg g_2(x) = 0$$

となることが分かります. したがって, $g_1(x), g_2(x)$ はどちらも定数であることが分かりますが, (93) 式に注意すると, これらの定数は 0 でない定数であることが分かります.

逆に, $0 \neq c \in \mathbb{C}$ とすると, 複素数全体の集合 $\mathbb{C}$ は定数関数として多項式全体の集合 $\mathbb{C}[x]$ に含まれていると見なせますから,

$$c\psi(x) \in \langle\, \psi(x) \,\rangle$$

となることが分かります. したがって, (81) 式から,

$$\langle\, c\psi(x) \,\rangle \subset \langle\, \psi(x) \,\rangle \tag{95}$$

となることが分かります. 一方,

$$\psi(x) = \frac{1}{c} \cdot c\psi(x) \in \langle\, c\psi(x) \,\rangle$$

と考えてみると, やはり, (81) 式から,

$$\langle\, \psi(x) \,\rangle \subset \langle\, c\psi(x) \,\rangle \tag{96}$$

となることが分かります．よって, (95) 式と (96) 式から,

$$\langle\, \psi(x)\,\rangle = \langle\, c\psi(x)\,\rangle$$

となることが分かります．以上より, $I$ の生成元 $\psi(x)$ に 0 でない複素数 $0 \neq c \in \mathbb{C}$ を掛け算しても, $\psi(x)$ と $c\psi(x)$ は同じイデアルを定めるということが分かります.

以上の議論を纏めると, 一変数多項式環 $\mathbb{C}[x]$ の零でないイデアル $I \neq \{0\}$ が, 勝手にひとつ与えられたときに, イデアル $I$ は,

**一変数多項式環 $\mathbb{C}[x]$ の零でないイデアル $I$ の記述**

(I) $I$ に属する 0 でない多項式の中で, 次数が最も低い多項式を $\psi(x) \in I$ とすると,

$$I = \langle\, \psi(x)\,\rangle \tag{97}$$

というように表わせる.

(II) (97) 式の右辺に現われるイデアル $I$ の生成元 $\psi(x)$ は, 0 でない複素数を掛け算する不定性を除いて一意的に定まる.

というように記述できることが分かりました.

## 10 何故, イデアルの条件は自然なのか *

さて, 可換環 $R$ に対して,

**部分環の条件**

$$x, y \in R' \implies x + y,\, xy \in R'$$

という部分環 $R' \subset R$ の定義は, 皆さんにとっても違和感はないのではないかと思います. 一方,

**イデアルの条件**

$$x, y \in I,\ z \in R \implies x + y,\, xz \in I \tag{98}$$

というイデアル $I \subset R$ の定義に対しては, 少なからず違和感を感じた方も多いのではないかと思います．そこで, 数学者が, イデアルという概念を用いて, どのようなことをイメージしているのかということを少しだけ説明してみることにします.

第 7 回の問 1 のところで注意したように, 特定の座標軸を持たないような線型空間の代表例が集合上の関数全体のなす集合です．そこで, いま, 勝手にひとつ取ってきた集合 $M$ に対して, $M$ 上の複素数値関数全体の集合を,

**$M$ 上の複素数値関数全体の集合**

$$\mathscr{F}_M = \{ f : M \to \mathbb{C} \}$$

という記号を用いて表わすことにします.[36] すると, 関数は「足し算」や「スカラー倍」ができるだけでなく,「掛け算」もできますから, $\mathscr{F}_M$ という集合を, 単に, 線型空間であると考えるよりは, 可換環であると考える方がより自然なことのように思われます.

以下では, $M$ として, 単なる集合ではなく,

$$M = \mathbb{R}^3, \mathbb{C}$$

というように,「空間」としてイメージできるようなものを考えて, これを幾何学的な対象であると考えることにします. また, 集合 $M$ のことを空間 $M$ と呼んだりすることにします. すると, 関数の集合を考えることによって, **幾何学的な対象である空間 $M$ に対して, 代数的な対象である可換環 $\mathscr{F}_M$ を対応させることができる**ということになります. このようにして, 空間 $M$ を色々と取り換えてみることで, 可換環の例が色々とできるわけですが, 現代数学では, この見方を逆転させて, **「可換環 $R$ があれば, それは, いつでも何らかの「空間」$M$ の上の関数全体のなす環 $\mathscr{F}_M$ となっている」**と考えてみるということが, すなわち, **「すべての可換環に対して, 対応する「空間」が存在する」**と考えてみるということが, 極めて有効な物の見方であることが分かっています. そして, そうした認識のもとで, **代数的な対象である「可換環」と幾何学的な対象である「空間」とは, 実は同じ物である**とみなすことが常識的な考え方になっています.[37]

そこで, 皆さんの参考のために, 幾何学的な描像と代数的な描像の対応について, 少しだけ説明してみることにします. こうした見方は, 慣れてしまえば, とても便利な見方なのですが, 頭に馴染むまで少し時間がかかると思いますので, すぐに納得できないような点があっても余り気にせずに, 軽い気持ちで眺めてみて下さい.

さて, 幾何学的な描像では, 空間 $M$ に対して,**「部分空間」**$N \subset M$ という概念を考えることができます.[38] そこで, **代数的な描像において「部分空間」に対応する概念はどのようなものであるのか**ということを考えてみます. このとき, 安直に考えると, それは「部分環」だろうという気がするわけですが, それで良いのかということを考えてみるために, **部分空間 $N \subset M$ が, 勝手にひとつ与えられたときに, $N$ に対して, 自然に定めることのできる $\mathscr{F}_M$ の部分集合が存在するかどうか**ということを考えてみます. そこで, いろいろと考えてみるのですが, 部分空間 $N$ に制限すると 0 になってしまうような $M$ 上の関数全体の集合

**部分空間 $N$ に対応したイデアル**

$$I_N = \{ f \in \mathscr{F}_M \mid x \in N \text{ のとき}, f(x) = 0 \text{ となる.} \}$$

という部分集合以外に, 部分空間 $N$ に自然に対応させることができるような集合は, どうも見当たらないことが分かります.[39]

そこで, 部分空間 $N \subset M$ に対して, $I_N \subset \mathscr{F}_M$ という部分集合を対応させて考えるということが自然なことであるということは, 取りあえず認めることにして, $I_N$ という部分

[36]英語で, 関数のことを function と言います.

[37]ただし, 上では, $\mathscr{F}_M$ として, 複素数 $\mathbb{C}$ に値を取る関数を考えましたが, 一般には,「関数の値が属すべき集合が空間の各点によって異なり得る」という形で「関数」の概念を少し拡張して考える必要があります.

[38]一般には, $M$ や $N$ として, 球面のような「曲がった空間」をイメージしています.

[39]これでは納得できないと思われる方は, もっと良い対応が考えられるかどうか, じっくりと考えてみて下さい.

集合がどのような代数的な性質を持っているのかということを考えてみます．すると, $I_N$ は, 単に「部分環」の条件を満たしているだけではなく, それより強い「イデアル」の条件を満たしていることが分かります．すなわち,

**イデアルに特徴的な条件**

$$x \in I, y \in R \implies xy \in I$$

という, 一見, 不自然に見えるイデアルの条件とは,**「部分空間 $N$ 上で 0 になるような関数 $f(x)$ に勝手な関数 $g(x)$ を掛け算してできる $f(x)g(x)$ という関数は, やはり, 部分空間 $N$ 上で 0 になるような関数である」**ということをイメージした条件であることが分かります.

このように, 部分空間 $N$ がどんなものであったとしても, **$N$ に対応して現われる $I_N$ という $\mathscr{F}_M$ の部分集合は常にイデアルになる**のですから, もし,

**幾何学的対象と代数的対象の間の対応**

$$\begin{array}{ccc} \text{幾何学} & & \text{代数学} \\ \text{空間：} M & \longleftrightarrow & \text{可換環：} R = \mathscr{F}_M \\ \text{部分空間：} N & \longleftrightarrow & \text{イデアル：} I = I_N \end{array} \tag{99}$$

という対応関係を信じるとすれば,**「部分環」より「イデアル」の方がより自然な概念である**ということになります．このとき, 上の対応は, 二つの部分空間の間に $N' \subset N$ という包含関係があると,

**(99) の対応は包含関係を逆転させる**

$$N' \subset N \iff I_{N'} \supset I_N$$

というように, 代数的な描像では包含関係が逆転するような対応になっていることに注意して下さい．特に,

$$N = M$$

とすると, $M$ 全体で 0 になるような関数とは 0 しかありませんから,

$$M \longleftrightarrow I_M = \{0\}$$

というように, 可換環 $\mathscr{F}_M$ の中の最小のイデアル $\{0\}$ が対応することが分かります．すると, 一見, 上の対応は, $M$ に対して, $\mathscr{F}_M$ という可換環を対応させる対応と, $I_M = \{0\}$ という可換環を対応させる対応と, 二重の対応を考えているように見えますが, 次のように考えると, これら二つの対応は本質的に同じ物であることが分かります.

いま, $M$ 上の関数 $f \in \mathscr{F}_M$ に対して, $f$ の定義域を部分集合 $N$ 上に制限したものを,

**$M$ 上の関数 $f \in \mathscr{F}_M$ の部分集合 $N$ への制限**

$$f|_N : N \to \mathbb{C}$$

という記号を用いて表わすことにします．すると, $I_N$ の定義によって, イデアル $I_N$ に属

するような関数 $f \in I_N$ は, 部分空間 $N$ 上に制限すると,

$$f|_N = 0$$

というように 0 という定数関数になってしまうのですから, $M$ 上の関数から $N$ 上の関数を作るという観点から考えると, $I_N$ の分だけの「無駄」があることが分かります. そこで, この「無駄」を省いて, $\mathscr{F}_M$ の元のうちイデアル $I_N$ に属するような元を 0 と思ったような集合を,

**可換環 $\mathscr{F}_M$ のイデアル $I_N$ による商集合 ( 可換環になる )**

$$\mathscr{F}_M / I_N$$

という記号で表わすことにします.[40] さらに, **部分集合 $N$ 上の関数は, すべて $\mathscr{F}_M$ の元を部分集合 $N$ 上に制限することによって得られる**ということを仮定すると, すなわち, **部分集合 $N$ 上の関数は, すべて $M$ 上の関数に拡張できる**ということを仮定すると,

$$\mathscr{F}_N \ni f|_N \ \longleftrightarrow\ [f] \in \mathscr{F}_M / I_N$$

という対応によって,

**部分空間 $N$ 上の関数環 $\mathscr{F}_N$ のイデアル $I_N$ を用いた記述**

$$\mathscr{F}_N \ \cong\ \mathscr{F}_M / I_N \tag{100}$$

というように同一視できることが分かります.[41] そこで, (100) 式によって,

$$\mathscr{F}_N \ \longleftrightarrow\ I_N$$

というように対応させて考えることにすると, いま, $\mathscr{F}_M$ はひとつ固定して考えていますから, 可換環 $\mathscr{F}_N$ を考えることと, イデアル $I_N$ を考えることとは, 本質的に同じことであると解釈できることが分かります.

実際には, (99) **式のような対応を完全なものにするためには, 可換環 $R = \mathscr{F}_M$ に応じて, どういうものを部分空間と認めるのかということをきちんと考えないといけません**. 例えば, 前節までの主役である一変数多項式環 $\mathbb{C}[x]$ の場合には, 対応する空間 $M$ として, 複素平面 $\mathbb{C}$ を取ることは自然なことですが, 9 節で述べた一変数多項式環 $\mathbb{C}[x]$ のイデアルの分類を考えると, 部分空間としては, 多項式の零点集合である有限個の点からなる部分集合のみを考えて,[42]

**一変数多項式環 $\mathbb{C}[x]$ の場合の対応**

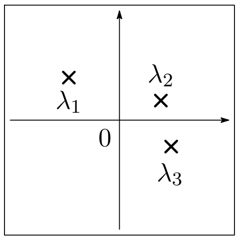
| | | |
|---|---|---|
| 空間 : $M = \mathbb{C}$ | $\longleftrightarrow$ | 可換環 : $R = \mathscr{F}_M = \mathbb{C}[x]$ |
| 部分空間 : $N = \{\lambda_1, \cdots, \lambda_m\}$ | $\longleftrightarrow$ | イデアル : $I_N = \langle \prod_{k=1}^{m} (x - \lambda_k) \rangle$ |

というように対応させてみると, ピッタリと対応することが分かります ( 図 4 を参照 ).

[40] 本当は, 可換環 $R$ とその中のイデアル $I \subset R$ に対して, $R/I$ という集合がどのような集合なのかということを, きちんと定義する必要がありますが, 少し説明を要しますので, ここでは省略することにします.

[41] ここで, $f \in \mathscr{F}_M$ を $\mathscr{F}_M / I_N$ の元と見なしているということを表わすために, $f$ に括弧「 $[\cdot]$ 」を付けて, $[f]$ という記号を用いて表わしました.

[42] ただし, 多項式が重根を持つ場合に対応して, 一般には, 点の重複も許す必要があります.




一対一対応 $\longleftrightarrow$

$I_N = \langle\, \psi(x)\, \rangle$
$\psi(x) = (x - \lambda_1)(x - \lambda_2)(x - \lambda_3)$

空間 : $\mathbb{C}$　　　　可換環 : $\mathbb{C}[x]$

図 4: 複素平面 $\mathbb{C}$ 内の有限個の点の集合 $\{\lambda_1, \cdots, \lambda_m\}$ に対して, それらを根全体の集合に持つような多項式 $\psi(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_m)$ を対応させることができる.

さて, 整数全体の集合 $\mathbb{Z}$ も可換環の代表的な例ですが,「多項式の次数」の代わりに「整数の絶対値」で置き換えて, 9 節の議論をそっくりそのまま繰り返してみると,「整数環 $\mathbb{Z}$ のイデアルとはどのような部分集合であるのか」ということが分かります. 興味のある方は, この作業を行なってみて, 上のような対応表を得るためには, $\mathbb{Z}$ に対応する「空間」として, 何を考えるのが良さそうかということを考えてみて下さい. また, 多項式の世界で Taylor 展開にあたることが, 整数の世界では何にあたっているのかということも考えてみて下さい.