

14 集合概念の基礎

数学の歴史に集合が現れるのは19世紀の後半であり、それ以前の数学は集合を用いずに記述されていた。確かに、これまでに扱ってきた内容は連立1次方程式の解法や行列式の計算など式変形を主体とする数学であり、ことさらに集合概念を押し出す必要はなかった。もしかすると、これ以降に学ぶ内容についても集合を用いずに議論を展開することが、あるいは可能かもしれない。しかし、この立場に固執すれば、今後、より複雑な概念が縦横無尽に現れる中で、定義や命題をやや曖昧に述べざるをえなかったり、あるいは証明において数学的に重要ではない些細な部分には目をつぶるような判断力や数学的センスを読者に要求することになるだろう。しかし、それでは多くの読者を路頭に迷わせることになってしまう。

現代数学において集合を用いた表現が市民権を得たのは、その記法を用いると厳密に述べやすいことにつきる。数学は一部の選ばれた者のみに許された学問ではなく、万人に許される学であるとする立場において、数学的なセンスを問わずに誤解なく伝わる集合による表現はかかせない。本論もこの立場に身を置き、以降では集合と写像を用いた記述を採用する。そこで本節と19節では、先に1.4および1.5項で述べた集合と写像に関する概念の発展として、これらのより高度な使い方について解説する。ただし、線形代数学の文脈に現れる部分のみを取り上げるゆえ、それ以外の部分、例えば合併集合や共通部分、補集合といった集合演算などの扱いについては集合論の入門的な参考書を参照されたい。

14.1 集合の包含関係

包含関係は、二つの集合の一致を示す際に必須となる概念である。

定義 14.1.1. 二つの集合 A, B が与えられており、 A のいかなる元も B に属するとき、 A は B の部分集合 (subset) であるといい、 $A \subset B$ あるいは $B \supset A$ と表す。

とくに、 A 自身は A の部分集合である。 $A \subset B$ であるとき「 A は B に含まれる」と述べることもある。この表現は、 A が B の元であること (つまり $A \in B$) と誤解される恐れもあるゆえ注意したい。一方、 A が B の部分集合でないということは、 B に属さない A の元が存在することを意味する。

例 14.1.2. (1) $\{\text{りんご, スイカ}\} \subset \{\text{りんご, みかん, スイカ}\}$ である。

(2) $\mathbb{N} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$ である。これらの記号の定義は1.4を参照せよ。

(3) 集合 $A = \{2, 9, 11, 30\}$ は集合 $B = \{2, 9, 15, 26, 30, 37\}$ の部分集合ではない。何故なら $11 \in A$ および $11 \notin B$ であり、 A の全ての元が B に属するわけではないからである。

(4) $A = \{1\}$ (一点のみの集合) とし、 $X = \{1, A\}$ (自然数1と集合 A の二つの元からなる集合) とすれば、 $A \in X$ と $A \subset X$ が共に成り立つ。

次で定める特別な集合は全ての集合の部分集合となる:

定義 14.1.3. いかなる元も含まない集合を空集合 (empty set) とよび、これを \emptyset と書く。

命題 14.1.4. 集合 X に対して、空集合 \emptyset は X の部分集合である。

Proof. 背理法で示す。 \emptyset が X の部分集合でないと仮定しよう。このとき部分集合の定義により、 X に属さない \emptyset の元 a が存在する。とくに $a \in \emptyset$ であり、これは \emptyset が元を含まないことに矛盾する。ゆえに \emptyset は X の部分集合である。 \square

二つの集合 A, B が等しいとは、 A を構成する元と B を構成する元とが一致するということである。これは、 A の元は B の元でもあり、また B の元は A の元でもあることにほかならない。すなわち、次が成り立つ:

$$\text{集合 } A, B \text{ について, } \quad A = B \iff A \subset B \text{ かつ } B \subset A. \quad (14.1.1)$$

例 14.1.5. 集合 $A = \{\text{りんご}, \text{スイカ}\}$ と集合 $B = \{\text{りんご}, \text{スイカ}, \text{スイカ}\}$ は等しい. 実際, 包含関係の定義 14.1.1 によれば $A \subset B$ および $B \subset A$ が成り立つ. したがって式 14.1.1 より $A = B$ である. 集合 B にスイカが二つ入っているわけではないことに注意しよう. スイカを二つ含む集合を考えたいのであれば, 二つのスイカを区別する手段をあらかじめ与えておき (例えば, スイカ 1, スイカ 2 とラベルを貼る), $\{\text{りんご}, \text{スイカ 1}, \text{スイカ 2}\}$ と書けばよい.

例 14.1.6. 本節以前の議論においても, 式 14.1.1 の左向き「 $A \subset B$ かつ $B \subset A \implies A = B$ 」を暗黙裡のうちに何度か用いていた.

- (1) 例 1.6.3 およびその後の議論において, \mathbb{R}^2 から \mathbb{R}^2 への線形写像全体の集合 X と $f \begin{pmatrix} x \\ y \end{pmatrix} := \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$ なる形で表せる写像 $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ 全体の集合 Y が一致することを確認した. 例 1.6.3 において $Y \subset X$ を述べて, その後の議論において $X \subset Y$ を示している.
- (2) 命題 4.1.1 の説明では, 方程式 $A\mathbf{x} = \mathbf{b}$ の解全体の集合 X と一点からなる集合 $Y = \{B\mathbf{x}\}$ が一致することを述べている. まず, \mathbf{a} を方程式の解 (つまり $\mathbf{a} \in X$) とすれば, \mathbf{a} は $B\mathbf{x}$ に一致すること (つまり $\mathbf{a} \in Y$) を示した. これは $X \subset Y$ を示すことに相当している. 次に $B\mathbf{x}$ が方程式の解であることを確認した. これは $Y \subset X$ を示すことに他ならない.
- (3) 一般の連立 1 次方程式の解法 (4.4 項) における一般解の表示についても同様のことを行った. 方程式 $A\mathbf{x} = \mathbf{b}$ の解全体の集合 X と $\mathbf{a}_0 + c_1\mathbf{a}_1 + c_2\mathbf{a}_2 + c_3\mathbf{a}_3$ で表されるベクトル全体の集合 Y が一致することを, $X \subset Y$ および $Y \subset X$ の両方を確認することによって示している.

よりみち (集合が等しいとはどういうことか).

二つの集合が一致するとはどういうことか改めて考えてみると, 雲をつかむような, とりとめもない思索しかできないことに気づく. 先程, 式 14.1.1 が成立することをもっともらしく述べたが, 実は「 $A \subset B$ かつ $B \subset A \implies A = B$ 」の説明はしていない. これは本当に正しい事実だろうか. 例えば, 赤い袋 A の中に自然数 1 と 2 のみが入っていると, 青い袋 B にも 1 と 2 のみが入っているとしよう. この二つの袋は色が違っているにも関わらず一致していると言えるのだろうか. 「集合と考えている場合は一致する」と言いたいところではあるけれども, その根拠に式 14.1.1 を用いるわけにはいかない. 何故なら, いま式 14.1.1 を説明するための議論をしているからである. では, どうやって A と B が一致することを導けばよいのだろうか.

このように, 集合が一致することを説明するのは意外に難しいのである. そこで集合論では, 式 14.1.1 を公理として定め, 外延性公理と呼んでいる. より素朴な立場では, 式 14.1.1 が集合が等しいことの定義であると考えてもよいだろう.

14.2 集合の表し方

数学に限らず, 何かしらの概念を規定しようと思うと, 大きく分けて二通りの方法があることに気づく. 新たな概念 A を規定するにあたり, A であるものをすべて列挙する方法を外延的な定義といい, A が持っている性質によって規定する方法を内包的な定義という. 例えば, 正多面体の定義として, 次の二通りの述べ方がある:

- **外延的定義:** 正四面体および正六面体, 正八面体, 正十二面体, 正二十面体なる図形を総称して正多面体という.
- **内包的定義:** 各頂点が同じ数の面と接し, すべての面が合同な正多角形となる多面体¹⁵を正多面体という.

¹⁵より正確には, ここでは凸多面体のみを考えている.

集合の表し方についてもこのことは例外ではなく、外延的な記述と内包的な記述の両方が用いられる。例えば、次の定義は外延的な記法である：

$$X := \{ \text{りんご, みかん, スイカ} \}.$$

また、あらかじめ数列 a_n (ただし $n \in \mathbb{N}$) が与えられているとき、 a_1, a_2, a_3, \dots をすべて集めた集合のことを

$$\{ a_n \mid n \in \mathbb{N} \}$$

と表す。これも外延的な記法である。

集合を規定する場合、中括弧 “ $\{, \}$ ” を用いることが慣例となっている。また、上の式で用いた記号 “ \mid ” の代わりにコロン “ $:$ ” を用いる文献も多い。

例 14.2.1. 次の例はいずれも外延的な記法である。

(1) 正の偶数全体の集合は次のように表される：

$$\{ 2, 4, 6, 8, 10, 12, \dots \}, \quad \text{あるいは} \quad \{ 2n \mid n \in \mathbb{N} \}.$$

ただし、最初の記法は、12 以降の列に偶数がもれなく現れるという保証はどこにも書かれておらず、曖昧な表記と言える。出来る限り誤解を避けたいのであれば、後の記法のほうが望ましい。

(2) 写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) = x^2$ で定める。このとき、定義域 \mathbb{R} の元を f に代入した値 $f(x)$ の範囲を表す集合を次のように書く：

$$\{ x^2 \mid x \in \mathbb{R} \}, \quad \text{あるいは} \quad \{ f(x) \mid x \in \mathbb{R} \}.$$

この集合は 0 以上の実数全体の集合に一致する。

次に内包的な記法を説明しよう。あらかじめ集合 X が与えられており、更に、 X に属する元 x たちに関する条件 $P(x)$ が与えられているとする¹⁶。このとき、 X の元のうち $P(x)$ が成立する元のみを全て集めた集合、すなわち X において $P(x)$ が成立する範囲を

$$\{ x \in X \mid P(x) \}$$

と書く¹⁷。

例 14.2.2. 次の集合の表し方はいずれも内包的である。

- (1) 集合 $\{ x \in \mathbb{N} \mid x = 2m \text{ を満たす } m \in \mathbb{N} \text{ が存在する} \}$ は正の偶数全体の集合である。
- (2) (m, n) -行列 A および m 次列ベクトル \mathbf{b} に対して集合 $\{ \mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{b} \}$ は連立 1 次方程式 $A\mathbf{x} = \mathbf{b}$ の解全体からなる集合である。
- (3) \mathbb{R}^2 の部分集合 $\{ \mathbf{x} \in \mathbb{R}^2 \mid \mathbf{x} = (x, y) \text{ と置くと, } x > 0 \text{ かつ } y > 0 \}$ を第 1 象限という。
- (4) 集合 $\{ x \in \mathbb{R} \mid x \neq x \}$ は空集合 \emptyset に等しい。条件 $x \neq x$ を満たす $x \in \mathbb{R}$ が一つもないからである。

定義 14.2.3. a, b を実数とする。 a 以上かつ b 以下の実数をすべて集めた集合を $[a, b]$ と書き、これを閉区間と呼ぶ。 a より大きくかつ b 未満の実数をすべて集めた集合を (a, b) と書き、これを开区間と呼ぶ。また、 a 以上 b 未満の実数全体の集合を $[a, b)$ 、 a より大きく b 以下の実数全体の集合を $(a, b]$ と書き、これらを半开区間と呼ぶ。更に、 a 以上の実数全体を $[a, \infty)$ 、 a より大きい実数全体を (a, ∞) と書き、 a 以下の実数全体、および a 未満の実数全体をそれぞれ $(-\infty, a]$ 、 $(-\infty, a)$ と書く。

¹⁶ P の例は 3.2 項を参照せよ。

¹⁷ 数理論理学では、条件 “ $P(x)$ かつ $x \in X$ ” を改めて $Q(x)$ と置くことにより、 $\{ x \mid Q(x) \}$ と表す。

定義 14.2.3 に現れた集合に $\mathbb{R} = (-\infty, \infty)$ を加えたものを総称して区間と呼ぶ。一点集合 $\{a\} = [a, a]$ や空集合 $\emptyset = [2, 1]$ (2 以上かつ 1 以下の数は存在しない) も区間である。

練習 14.2.4. 定義 14.2.3 で与えたそれぞれの区間について、これを集合の内包的な表記で表せ。

解答例 (抜粋): $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$, $(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$, $[a, \infty) = \{a \in \mathbb{R} \mid a \geq 0\}$.

発展 (もう一つの区間の定義).

いま述べた区間の定義は、区間をすべて列挙した定め方であるから外延的と言える。一方で、条件「 $a, b \in I$ かつ $a < x < b \implies x \in I$ 」を満たす \mathbb{R} の部分集合として区間を定義する流儀もある (内包的な区間の定義)。これらの定義が一致することは明らかではなく、証明には実数の連続性の公理を要する。詳しくは解析の本を参照せよ。

14.3 外延的か内包的か

集合の表記が外延的なものか内包的なものかは文脈で判断すること。なかには外延的とも内包的ともとれる記法がある:

例 14.3.1. 次の記法はいずれも整数を成分とする 2 次正方行列全体のなす集合 $M_2(\mathbb{Z})$ を表す。(1) が外延的記法で (4) が内包的記法である。この中間の記法として (2) や (3) の記法もしばしば用いられる。これらを外延的とするか内包的とするかは記述者の立場に委ねられよう。

(1) 写像 $F: \mathbb{R}^4 \rightarrow M_2(\mathbb{R})$ を $F(a, b, c, d) := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ と定めておく。 $M_2(\mathbb{Z}) = \{F(\mathbf{x}) \mid \mathbf{x} \in \mathbb{Z}^4\}$.

(2) 上の \mathbf{x} や $F(\mathbf{x})$ を具体的に成分表示した表記: $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid (a, b, c, d) \in \mathbb{Z}^4 \right\}$.

(3) $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a, b, c, d \in \mathbb{Z} \right\}$, $M_2(\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in M_2(\mathbb{R}) \mid a, b, c, d \in \mathbb{Z} \right\}$.

(4) $M_2(\mathbb{Z}) = \left\{ A \in M_2(\mathbb{R}) \mid A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ とおくと, } a, b, c, d \in \mathbb{Z} \right\}$.

集合の表示が外延的か内包的かを厳密に分類したいのであれば、上の例における (2) や (3) のような表記を認めないことと約束し、記号 “|” の左側に現れる集合の元を表す記号が裸のまま用いられていれば内包的、文字が添え字づけられていれば外延的とすればよい。ここで、元を表す記号が裸であるとは、 $x \in X$ というように、元を表す記号が一つの文字 x のみからなる場合を指す。また、添え字づけられていたとは、 x_n あるいは x^2 , $f(x)$, x_f といったように、元そのものが複数の文字 (または記号) を用いて表されていることを指す。なお、添え字づけられた元は、ある写像によって代入された値であると見なすことができる¹⁸。

しかしながら、記述された集合が意味するところに誤解がないのであれば、例 14.3.1 における (2) あるいは (3) のような表記を認めても数学的議論に支障はなく、多くの文献でこのような記法が用いられている。本論もこれに準じる。

練習 14.3.2. 次の集合の記法は外延的か内包的か答えよ。

(1) $\{x \in \mathbb{Z} \mid x \text{ は奇数}\}$, (2) $\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$, (3) $\{(x, \sin x) \mid x \in \mathbb{R}\}$,

¹⁸例えば実数列 a_n の各項は、 $n \in \mathbb{N}$ に対して $a_n \in \mathbb{R}$ を対応させる写像 $f: \mathbb{N} \rightarrow \mathbb{R}$ を用いることで、 f に各自然数を代入した値と見なせる。

解答例: (1) と (2) が内包的表記であり, (3) が外延的表記である. ただし, (2) は例 14.3.1 でいうところの中間の記法ともみなせる. なお, (1) は奇数全体の集合, (2) は \mathbb{R}^2 上の原点を中心とする半径 1 の円周上の点全体の集合, (3) は $f(x) := \sin x$ で定められる関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ のグラフ上の点全体の集合を表す.

方程式を解くとは, 内包的記述を外延的記述に書き直す行為にほかならない:

練習 14.3.3. 次で与えられる集合 X の外延的表記を与えよ.

$$(1) X = \{x \in \mathbb{R} \mid x^2 - 1 = 0\}.$$

解答例: $x^2 - 1 = 0$ を満たす数は $x = 1$ および $x = -1$ である. ゆえに $X = \{1, -1\}$.

$$(2) X = \left\{ \mathbf{x} \in \mathbb{R}^5 \mid \begin{bmatrix} 0 & 1 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 4 \\ 5 \\ 0 \end{bmatrix} \right\}.$$

解答例: この連立 1 方程式の解法は 4.4 項で述べた通りであり, その解全体の集合は,

$$X = \left\{ \begin{bmatrix} 0 \\ 4 \\ 0 \\ 5 \\ 0 \end{bmatrix} + c_1 \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ -3 \\ 1 \\ 0 \\ 0 \end{bmatrix} + c_3 \begin{bmatrix} 0 \\ -2 \\ 0 \\ -1 \\ 1 \end{bmatrix} \mid c_1, c_2, c_3, c_4 \in \mathbb{R} \right\}.$$

14.4 集合論と逆理 (発展)

内包的な集合の記述を導入したことにより, 多様性に富んだ集合の表現が可能になった. しかし, そこには大きな落とし穴が潜んでいることが知られている. それは次の枠内における議論であり, ラッセルの逆理と呼ばれている.

ラッセルの逆理

考えられ得るすべての集合を集めた集合を U とする (つまり U は集合の集合である). 更に条件 $P(X)$ を “ $X \notin X$ ” と定め, 集合 Y を次で定義する:

$$Y = \{X \in U \mid X \notin X\}$$

このとき, 集合 Y 自身は Y の元となるであろうか. $Y \in Y$ および $Y \notin Y$ のいずれかが成立するはずである. どちらが正しいのか検討しよう.

(1) $Y \in Y$ と仮定すると, 定義によれば Y は条件 $P(X)$ を満たす集合の集まりであったから, その元である Y 自身も条件を満たす. すなわち $Y \notin Y$ が成立する. しかしこれは $Y \in Y$ と仮定したことに矛盾する.

(2) そこで, $Y \notin Y$ と仮定する. つまり Y は条件 $P(X)$ を満たしており, $P(X)$ を満たす集合全体の集まりが Y であったから, $Y \in Y$ である. しかしこれは $Y \notin Y$ という仮定に矛盾する.

こうして, いずれの場合も矛盾が生じてしまった.

数学における議論あるいは証明は, 一般の諸科学と比べても厳密性が非常に高いものであると多くの人が認識していることだろう. しかし, ラッセルの逆理によると, 数学の論理にも少々あやふやな部分があるということなのだろうか. また, そうではないとするのであれば, 上のような矛盾を排除するかたちで数学の理論 (特に集合論) を再構成することは可能なのだろうか. 現在では, 逆理を回避するための技

術が得られており、初学者がこの点について不安がる必要はないことになっている。本項では、この点について概略的な説明をしておこう。

逆理を回避するための技術論を検討するのであれば、そもそも「証明」とは何か再考する必要があるだろう。そのための模範となった理論はユークリッド幾何学である。ユークリッド幾何学では、いくつかの公理を前提として演繹的に数多くの定理を導いていた。これと同様にして集合論においても、集合に関するいくつかの公理を認め、それらの公理と論理的に正しい命題¹⁹および三段論法などの推論規則を有限回だけ用いて別の命題を導くことを「証明」と定めるのである。こうした立場で展開する集合論を公理的集合論と呼ぶ²⁰。

集合論の公理として何を採用すべきかという基準は、もちろん数学者各個人の価値観によって論点が変われることかもしれない。しかしながら、現在では ZFC²¹と呼ばれる公理系が多くの数学者の同意を得て、一般的に用いられている。ZFC の詳細を書く余裕はないが、この公理系においては集合全体の集合 U は構成できず、したがって枠内の議論における Y も定義できない(詳しい理由は本節末のコラムを見よ)。かくしてラッセルのパラドックスは避けられるのである。

ところで、良い公理系を導入した理論において、もはやラッセルの逆理は生じないにしても、それではラッセルの逆理とは別の矛盾も絶対に生じないという保証はあるのだろうか。もし、新たな矛盾論法が見つかってしまったならば、その矛盾を排除するようより頑強な公理系が作れるかどうかを検討せねばならない。こうした不安を解消するためにも、ZFC において矛盾が導かれることはない(これを無矛盾であるという)ことを証明しようという組織的な試みがなされた(この試みはヒルベルト・プログラムと呼ばれる)。ヒルベルト・プログラムにおける最終的な答えはゲーデルによって与えられており、彼によれば、自然数論を含む無矛盾ないかなる公理系においても、その公理系の内部で自身の無矛盾性を証明することは出来ない(ゲーデルの第 2 不完全性定理)というのである。無矛盾性が保証されることは決してないというゲーデルの回答は悲観すべきことだろうか。この議論は数学界の内部にとどまらず多くの人が興味を持ち、様々な論争を巻き起こすことになった。

ところで、自己言及によって矛盾を導くというラッセルの逆理と構造の似た逆理がいくつか知られている。例えば「この文は間違っている」という文は正しいか、それとも間違っているのか。正しいとすれば、その文面通り間違っているから矛盾であり、間違っているとすれば「間違っている」ことは間違いということで正しいことになり、やはり矛盾を得る。こうした日常言語の世界を我々はどう捉えるべきかという課題もある。論理学に加えて言語学や認識論といった様々な背景を抱えたこの難問も広く論じられ、その回答のうち代表的なものとして、例えばウィトゲンシュタインの『論理哲学論考』が挙げられる。

¹⁹ $A \Rightarrow A$ や $(A \Rightarrow B) \Rightarrow (\neg B \Rightarrow \neg A)$ といった命題は A, B にどんな論理式を代入しても真である。こういった論理式は恒真式と呼ばれ、数理論理学において厳密に定義される。

²⁰ これに対して、公理化せずに感覚的に集合を扱う立場を素朴集合論と言う。

²¹ ツェルメロ＝フレンケルの公理系 (ZF) に選択公理 (Axiom of Choice) を加えた公理系のこと。

発展 (集合とは何か).

公理的集合論における集合概念の厳密な定義とは何であろうか. これは数理論理学を専攻しなくても気になることであろう. 結論を先に述べてしまうと, 集合自体に確固たる定義などはない. 例えばユークリッド幾何学における「直線」については, 公理によっていくつかの性質が仮定されるのみであって, 直線そのものが定義されるわけではない. それと同様に, 集合論においても定義が厳密なのは公理であり集合自身ではないのである.

しかしそうすると, 今度は我々が数学で用いる集合らしきものが, 公理的集合論の文脈における集合であるのかどうかという不安にさいなまれるかもしれない. この点において, 大概是, 次の二点さえおさえておけば十分である. 一つは, 実数全体 \mathbb{R} や複素数全体 \mathbb{C} , ユークリッド空間 \mathbb{R}^n など日常的に現れる集合は集合論の公理を組み合わせることで構成できることが分かっており, これらが集合であること (集合として存在すること) を疑う余地はないということである. もう一つは, 我々が新たに構成する集合についての注意であり, これについては置換公理と分出公理をおさえておけばよい. 置換公理とは大まかに言えば, 外延的な記述によって集合を与えてよいという公理に相当し, 分出公理は, 内包的な記述によって集合 X の部分集合を定義してよいという公理に相当する. これらの公理のおかげで, 実質的な数学に現れる対象が集合であるかないかを我々が意識する必要はないのである.

分出公理についてもう少し詳しく言うと, それは次のように述べられる:

分出公理: X を集合, P を x に関する論理式とすれば, 集合 $A := \{x \in X \mid P(x)\}$ が存在する.

分出公理のかなめは文頭が「 X を集合とすれば～」となっている点である. 集合全体 U を集合であると仮定して上の X に適用するとラッセルの逆理によって矛盾が生じることから, したがって U は集合ではない, 集合全体をなすような集合は存在しないという結論に至る. これより詳しい事情については数理論理学の専門書に譲ろう.

15 線形空間

線形空間は線形代数学において主題となる代数構造である。公理化された代数構造を論じる理由は、2.4項で述べたように演算の定義にいちいち戻らなくても議論ができるという点、そして様々な空間を同時に論じることができる点にある。例えば、 \mathbb{R}^n において示された性質が写像のなす集合においても示され、それらの証明に使われた技法もほとんど同じというのであれば、それらを同時に証明できるような枠組みを与えておくと手間が省ける(命題 6.2.3 と 6.2.4 の類似性を思いだそう)。このように数学では、汎用性を重視して抽象的な代数構造を導入している。

15.1 ベクトル空間の公理

\mathbb{R}^n における和とスカラー倍の性質のなかで特に重要と思われる部分を抽出することで、我々は線形空間の定義を得る:

定義 15.1.1. 集合 V に対して和 $+$ とスカラー倍 \cdot の演算が定められており、さらに特別な元 $\mathbf{0} \in V$ が与えられているとする。これらの演算が次に述べるベクトル空間の公理を満たすとき、四つ組 $(V, \mathbf{0}, +, \cdot)$ を線形空間 (linear space) またはベクトル空間 (vector space) と呼ぶ。

ベクトル空間の公理

$\mathbf{a}, \mathbf{b}, \mathbf{c} \in V, r, s \in \mathbb{R}$ とする。

I. 各元 $\mathbf{a}, \mathbf{b} \in V$ に対して和 $\mathbf{a} + \mathbf{b} \in V$ が定まっており、次の性質を満たす:

$$(1) \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (2) (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (3) \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}.$$

II. 各元 $\mathbf{a} \in V$ および $r \in \mathbb{R}$ に対してスカラー倍 $r \cdot \mathbf{a}$ が定まっており、次の性質を満たす:

$$(4) r \cdot (s \cdot \mathbf{a}) = (rs) \cdot \mathbf{a}, \quad (5) (r + s) \cdot \mathbf{a} = (r \cdot \mathbf{a}) + (s \cdot \mathbf{a}) \quad (6) r \cdot (\mathbf{a} + \mathbf{b}) = (r \cdot \mathbf{a}) + (r \cdot \mathbf{b})$$

$$(7) 1 \cdot \mathbf{a} = \mathbf{a}, \quad (8) 0 \cdot \mathbf{a} = \mathbf{0}.$$

上の性質を満たす $\mathbf{0}$ のことを V の零元または零ベクトルと呼ぶ。また、線形空間の元を総称してベクトルと呼ぶ。

零ベクトルが V の元であることを強調し、これを $\mathbf{0}_V$ と書くこともある。慣例ではスカラー倍の演算記号 \cdot は省略して $r\mathbf{a}$ と書き、更に四つ組 $(V, \mathbf{0}, +, \cdot)$ を V と略記する。

ベクトル空間の公理 (1) から (8) を直ちに暗記しないと以後の線形代数学の理解に支障がでるかといえば、そのようなことはない。何故なら、線形空間における演算は \mathbb{R}^n における演算と同様に無意識のうちに処理されるからである。しかし、「 \mathbb{R}^n の演算と似たような演算をもつ集合」と曖昧に線形空間を定義するわけにもいかず、上のように形式的な定義を与えた。

例 15.1.2. ベクトル空間の公理における性質 (1) から (8) は、命題 2.4.1 における性質 (1), (3), (2), (11), (15), (14), (10), (9) に相当する²²。命題 2.4.1 から次が従う:

(1) \mathbb{R}^n における原点 $\mathbf{0} := (0, \dots, 0)$ を零元と見なすことにより、1.4 項で定めた \mathbb{R}^n における和とスカラー倍はベクトル空間の公理を満たす。したがって \mathbb{R}^n はベクトル空間である。

(2) (m, n) -行列全体のなす集合を $M_{m,n}(\mathbb{R})$ と書く。2.2 節で述べたように $M_{m,n}(\mathbb{R})$ には和とスカラー倍が定義されている。零行列 $O_{m,n}$ を零元と見なすことにより、 $M_{m,n}(\mathbb{R})$ における和とスカラー倍はベクトル空間の公理を満たし、したがって $M_{m,n}(\mathbb{R})$ は線形空間である。

我々は \mathbb{R}^n と $M_{n,1}(\mathbb{R})$ (あるいは $M_{1,n}(\mathbb{R})$) を同一視していた。また、行列式の項目において n 次正方行列全体の集合を $M_n(\mathbb{R})$ と書いていた。つまり、 $M_n(\mathbb{R}) = M_{n,n}(\mathbb{R})$ である。

²²したがって、多元環はベクトル空間である。多元環とは、ベクトル空間に分配法則と結合法則を満たす積を付加した代数構造のことを指す(詳しい定義は 2 節最後のコラムを見よ)。

15.2 線形空間の例

線形空間の例をいくつか挙げよう. はじめの例は \mathbb{R}^n の部分集合として具体的に表示できるものである. この例の一般化については次節で更に詳しく述べる.

例 15.2.1. $W_{[1,-1]} := \{(x, y) \in \mathbb{R}^2 \mid x - y = 0\}$ とすれば, W は直線 $y = x$ 上の点全体を表す. $W_{[1,-1]}$ における和とスカラー倍の演算を, \mathbb{R}^2 における和とスカラー倍の演算によって定めよう. すると, $\mathbf{0} \in \mathbb{R}^2$ が $W_{[1,-1]}$ に含まれること, および $W_{[1,-1]}$ の各元どうしの和や W の元のスカラー倍が再び $W_{[1,-1]}$ の元となることが次のように確認できる:

Proof. $0 - 0 = 0$ より $(x, y) = (0, 0)$ は条件「 $x - y = 0$ 」を満たす. ゆえに $\mathbf{0} = (0, 0) \in W$ である. 次に, $\mathbf{a} = (a_1, a_2) \in W, \mathbf{b} = (b_1, b_2) \in W, r \in \mathbb{R}$ と仮定し, $\mathbf{a} + \mathbf{b}, r\mathbf{a} \in W$ を示そう. $\mathbf{a} + \mathbf{b} = (a_1 + b_1, a_2 + b_2)$ が W に含まれることを示すには, $(x, y) = (a_1 + b_1, a_2 + b_2)$ が条件「 $x - y = 0$ 」を満たすことをいえばよい. $\mathbf{a}, \mathbf{b} \in W$ より, $(x, y) = (a_1, a_2)$ および $(x, y) = (b_1, b_2)$ について条件「 $x - y = 0$ 」が成立する. つまり $a_1 - a_2 = 0, b_1 - b_2 = 0$ である. したがって $(a_1 + b_1) - (a_2 + b_2) = (a_1 - a_2) + (b_1 - b_2) = 0 + 0 = 0$. したがって, $(x, y) = (a_1 + b_1, a_2 + b_2)$ は条件「 $x - y = 0$ 」を満たし, $\mathbf{a} + \mathbf{b} \in W$ である. また $r\mathbf{a} = (ra_1, ra_2)$ が W に含まれることを示すには, $(x, y) = (ra_1, ra_2)$ が条件「 $x - y = 0$ 」を満たすことをいえばよい. これも $(ra_1) - (ra_2) = r(a_1 - a_2) = r \cdot 0 = 0$ と直ちに分かる. \square

W における和とスカラー倍がベクトル空間の公理を満たすことは, \mathbb{R}^n における演算がそうであることから明らかであり, したがって W は線形空間である.

ユークリッド空間 \mathbb{R}^n とは n 本の座標軸を持つ空間のことであった. ここから類推される空間として, 無限個の座標軸を持つ空間を考えよう.

例 15.2.2. 無限個の実数の組 (x_1, x_2, x_3, \dots) 全体のなす集合を $\mathbb{R}^{\mathbb{N}}$ と書く. これは, 無限に続く実数列全体のなす集合とも考えられる. \mathbb{R}^n の場合と同様にして, $\mathbb{R}^{\mathbb{N}}$ にも次のように和とスカラー倍が定められる:

- $(x_1, x_2, x_3, \dots) + (y_1, y_2, y_3, \dots) := (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots),$
- $r(x_1, x_2, x_3, \dots) := (rx_1, rx_2, rx_3, \dots).$

すなわち, 各座標ごとに和とスカラー倍を取っている. これらの演算はベクトル空間の公理を満たし, $\mathbb{R}^{\mathbb{N}}$ はベクトル空間となる. なお, 組 (x_1, x_2, x_3, \dots) を $(x_n)_{n \in \mathbb{N}}$ と略記することもある. この記法を用いれば, 上の演算は次のように書ける: $(x_n)_{n \in \mathbb{N}} + (y_n)_{n \in \mathbb{N}} := (x_n + y_n)_{n \in \mathbb{N}}, r(x_n)_{n \in \mathbb{N}} := (rx_n)_{n \in \mathbb{N}}.$

次に多項式のなす空間を考えよう.

定義 15.2.3. 文字 x および非負整数 n , 実数 a_0, a_1, \dots, a_n を用いて

$$\sum_{i=0}^n a_i x^i = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

と表される式のことを実数を係数とする n 次以下の多項式 (**polynomial**) という. ここで, 形式的に $x^0 = 1$ と定める. 上の式において $a_n \neq 0$ となるとき, f の次数 (**degree**) を n と定め, これを記号 $\deg f$ で表す. また, このとき, f を n 次多項式という.

$m > n$ のとき, n 次以下の多項式は m 次以下の多項式でもある. 例えば $2x^2 + 3x - 1$ は 2 次多項式であり, これを $0x^3 + 2x^2 + 3x - 1$ と表すことで 3 次以下の多項式とも見なせる.

例 15.2.4. 実数を係数とする多項式全体の集合を $\mathbb{R}[x]$ と書き, これを実数係数多項式環という. また, その中で n 次以下の多項式全体のなす部分集合を $\mathbb{R}[x]_n$ と書く. 例えば $\mathbb{R}[x]_2 = \{ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ である. さて, $\mathbb{R}[x]$ において和とスカラー倍を次のように定めると, これらは再び多項式になる:

$$\left(\sum_{i=0}^n a_i x^i \right) + \left(\sum_{i=0}^n b_i x^i \right) := \sum_{i=0}^n (a_i + b_i) x^i, \quad r \left(\sum_{i=0}^n a_i x^i \right) := \sum_{i=0}^n (ra_i) x^i.$$

これらの演算がベクトル空間の公理を満たすことは容易に確かめられ、したがって $\mathbb{R}[x]$ は線形空間である。なお、 $\mathbb{R}[x]$ における零元とは、すべての係数 a_i がゼロなる多項式のことである。これを関数と見なせば、どんな数を代入しても 0 に値をとる定数関数を意味する。また、 n 次以下の多項式において和やスカラー倍を行うと、 n 次以下の多項式が得られることから、 $\mathbb{R}[x]_n$ も線形空間である。

多項式には、形式的な式と見なす立場と、 x を変数とする関数と見なす立場がある。前者は、二つの多項式 $p(x) = \sum_{i=0}^n a_i x^i$ と $q(x) = \sum_{i=0}^n b_i x^i$ が等しいことを、各 $i = 0, \dots, n$ について $a_i = b_i$ であると定める立場である。後者は、 $p(x) = q(x)$ がいかなる定義域の元 x に対しても成り立つことを $p = q$ の定義とする立場である。なお、後者の定義においてはあらかじめ定義域を宣言しておかねばならない。厳密には、これら二つの立場を区別すべきであるが、線形代数の初歩を学ぶにあたってはこだわる必要はないであろう。次の命題はその根拠となる。

命題 15.2.5. $n + 1$ 個以上の元を含む \mathbb{R} の部分集合 I を定義域とする二つの n 次以下の多項式関数 $p(x) = \sum_{i=0}^n a_i x^i$, $q(x) = \sum_{i=0}^n b_i x^i$ について、次は同値である。

$$(1) \text{ 各 } t \in I \text{ について } p(x) = q(x), \quad (2) \text{ 各 } i = 0, \dots, n \text{ について } a_i = b_i.$$

Proof. (2) \Rightarrow (1) は明らかゆえ (1) \Rightarrow (2) を示す。 $p(x) = q(x)$ より関数 $p(x) - q(x) = \sum_{i=0}^n (a_i - b_i)x^i$ は 0 に値を取る定数関数である。 I は $n + 1$ 個以上の元を含むことから、 $n + 1$ 個の異なる数 $t_1, \dots, t_{n+1} \in I$ を取ることができ、これらの数を $p(x) - q(x)$ に代入することで次の式を得る。

$$\begin{cases} t_1^n(a_n - b_n) + t_1^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_1(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \\ t_2^n(a_n - b_n) + t_2^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_2(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \\ & \vdots \\ t_n^n(a_n - b_n) + t_n^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_n(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \\ t_{n+1}^n(a_n - b_n) + t_{n+1}^{n-1}(a_{n-1} - b_{n-1}) + \dots + t_{n+1}(a_1 - b_1) + 1 \cdot (a_0 - b_0) & = 0 \end{cases} \quad (15.2.1)$$

上式に現れる各 $(a_i - b_i)$ にかかる係数を成分とする $n + 1$ 次正方行列 A を次で定める：

$$A := \begin{bmatrix} t_1^n & t_1^{n-1} & \dots & t_1 & 1 \\ t_2^n & t_2^{n-1} & \dots & t_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ t_n^n & t_n^{n-1} & \dots & t_n & 1 \\ t_{n+1}^n & t_{n+1}^{n-1} & \dots & t_{n+1} & 1 \end{bmatrix}.$$

すると式 15.2.1 は、 $\mathbf{x} = {}^t(a_n - b_n, \dots, a_0 - b_0)$ が $A\mathbf{x} = \mathbf{0}$ の解であることを意味している。ヴァンデルモンドの公式 (定理 11.2.2) より $|A| = \pm \prod_{1 \leq i < j \leq n+1} (t_j - t_i)$ であり、各 t_i は異なる数ゆえ $|A| \neq 0$ 。ゆえに A は可逆であり方程式 $A\mathbf{x} = \mathbf{0}$ の解は唯一解 $\mathbf{x} = \mathbf{0}$ のみである。したがって $a_i - b_i = 0$ ($i = 0, \dots, n$)。□

多項式を関数と見なす立場からは、次のような一般化が考えられる。

例 15.2.6. \mathbb{R} 上の区間 I を定義域とする実数値連続関数全体のなす集合を $C(I)$ と書く。 $C(I)$ において、和とスカラー倍を定義しよう。新たな関数 h を定義するという事は、定義域の各元を h に代入した値を定めることに他ならない。そこで、次のように演算を定義する。

- $f, g \in C(I)$ に対して関数 $(f + g) : I \rightarrow \mathbb{R}$ を次のように定める：
各 $x \in I$ について、 $(f + g)(x) := f(x) + g(x)$ 。
- $f \in C(I)$ および $r \in \mathbb{R}$ に対して関数 $(rf) : I \rightarrow \mathbb{R}$ を次のように定める：
各 $x \in I$ について、 $(rf)(x) := rf(x)$ 。

$f + g$ や rf が再び連続関数となること (すなわち $f + g, rf \in C(I)$) の証明は解析系の講義に譲ろう²³. $C(I)$ における零元は 0 に値をとる定数関数である. 上で定義された演算がベクトル空間の公理を満たすことは容易に確かめられ, したがって, $C(I)$ は線形空間である.

例 15.2.4 における多項式を区間 I を定義域とする関数とみなす立場においては, 多項式はいずれも連続関数であるから $\mathbb{R}[x] \subset C(I)$ となる. このとき, 例 15.2.4 における和とスカラー倍の定義と, 例 15.2.6 におけるそれは一致する.

例 15.2.7. (1) 関数の連続性は数列の収束概念を用いて定められていた. そこで, 収束概念が定まる空間を定義域とする関数についても連続性を定義することができる. 例えば \mathbb{R}^2 の点列 \mathbf{x}_n が点 $\mathbf{x} \in \mathbb{R}^2$ に収束することは, \mathbf{x}_n と \mathbf{x} の距離²⁴が 0 に収束することと定められる. 収束概念が定まる図形 X (例えば X を \mathbb{R}^2 の部分集合とすればよい) を定義域とする実数値連続関数全体のなす集合を $C(X)$ と書けば, $C(X)$ も例 15.2.6 と同様にして線形空間となる.

(2) 実は, 関数のなす空間を線形空間とみなすために, 関数を連続関数のみに制限する必要はない. より一般に, 集合 X を定義域とする実数値関数全体のなす集合を \mathbb{R}^X とすれば, 例 15.2.6 と同様にして和とスカラー倍を定めることで \mathbb{R}^X は線形空間となる.

なお, この例において $X = \mathbb{N}$ とする場合と例 15.2.2 における数列空間 $\mathbb{R}^{\mathbb{N}}$ の間には自然な 1 対 1 の対応があり, これらは同一の概念と見なすことができる. 実際, 数列空間の各元 (x_1, x_2, \dots) は $x(n) := x_n$ なる関数 $x : \mathbb{N} \rightarrow \mathbb{R}$ に対応し, 逆に関数 $y : \mathbb{N} \rightarrow \mathbb{R}$ は実数列 $y_n := y(n)$ (すなわち $(y(1), y(2), \dots)$) に対応する. この対応は, それぞれの和とスカラー倍の演算に関して整合的である. \mathbb{N} を定義域とする関数と数列の違いは, $x(1), x(2), \dots$ と書くか, あるいは x_1, x_2, \dots と書くかという, 僅かな記号上の違いしかないのである.

例 15.2.8 (発展). $C(X)$ には次のようにして積も定めることができる.

$f, g \in C(X)$ に対して関数 fg を次で定める: 各 $x \in X$ について, $(fg)(x) := f(x)f(g)$.

このとき, $C(X)$ における演算は, $1 \in \mathbb{R}$ に値を取る定数関数 $\mathbf{1}$ を単位元とすることで命題 2.4.1 における (1) から (18) すべての性質を満たす²⁵. すなわち, $C(X)$ は多元環となる. また, 多項式の積は再び多項式になることから, $\mathbb{R}[x]$ も多元環となる. $\mathbb{R}[x]_n$ は多元環にはならない. 何故なら, n 次多項式どうしの積は $2n$ 次の多項式となり, これは $\mathbb{R}[x]_n$ に含まれないからである.

実は, いかなる線形空間も, 無理矢理かけ算を導入することで多元環とみなすことができる. しかしながら, 大した意味もないのに人工的に積の演算を定めても実りある数学が得られることはないであろう. 線形代数学における行列のもつ性質, すなわち $M_n(\mathbb{R})$ の数学的な捉え方をいかに昇華するかという文脈において多元環の概念が意味を持つであり, この立場から遠く離れた枝葉末節的な対象については, 本論では言及しない.

15.3 体 K 上の線形空間 (発展)

これまでの議論において, 行列に現れる成分およびスカラー倍の係数, 連立 1 次方程式に現れる係数はいずれも実数であるとしていた. しかし, 扱う数を実数に限るべき確たる根拠はどこにもなかった. 仮にあるとすれば, 初学者にとって最もイメージがわかりやすいからということに尽きるだろう. そこで行列に現れる成分を有理数に限ったり, あるいは複素数も認めるといった状況を考えよう.

仮に成分を有理数に限定した場合, 行列演算を繰り返し展開しても成分に現れるのは有理数に限られ, 行列式の値も有理数である. また有理数係数の連立 1 次方程式の解も有理数である. 一方, 行列の成分に複素数を認める場合は, やはり行列演算後の成分や行列式の値は複素数となる. 複素数を係数とする連立

²³ 収束列による連続性の定義を採用すれば, 高校数学の範囲で示せることである.

²⁴ $\mathbf{x}_n = (a_n, b_n)$ と $\mathbf{x} = (a, b)$ の距離はピタゴラスの定理により $\sqrt{(a_n - a)^2 + (b_n - b)^2}$ と計算できる.

²⁵ 命題 2.4.1 では単位元の記号に $\mathbf{1}$ ではなく E を用いている.

1次方程式の解も複素数である。このように実数および有理数、複素数に共通した現象が生じる背景には、これらが四則演算で閉じているという共通項がある。

したがって、四則演算が定まる代数構造 K さえ与えられれば、行列の成分やスカラー倍の係数、連立1次方程式に現れる係数に K の元を取ることで、 K に関する線形代数の世界が考えられうる。四則演算が与えられる代数構造は体 (field) と呼ばれる。その形式的な定義は代数学の専門書を参照されたい²⁶。

例 15.3.1. (1) 実数全体 \mathbb{R} および有理数全体 \mathbb{Q} 、複素数全体 \mathbb{C} はそれぞれ体である。

(2) 整数全体 \mathbb{Z} は、整数どうしの割り算が整数になるとは限らないゆえ体ではない。整数を成分とする行列において、行列演算後の各成分や行列式は確かに整数となる。しかし、整数係数の連立1次方程式は、整数でない有理数を解に持つことがある。これは整数が割り算で閉じていないことに起因している。

(3) 無理数全体は体ではない。無理数どうしの和が無理数になるとは限らないからである。例えば $\sqrt{2}$ および $1 - \sqrt{2}$ は共に無理数であるが (練習 15.3.2), その和 $\sqrt{2} + (1 - \sqrt{2}) = 1$ は有理数である。

練習 15.3.2. $\sqrt{2}$ が無理数であることを認めたとえで $1 - \sqrt{2}$ が無理数であることを示せ。

解答例: 仮に $1 - \sqrt{2}$ が無理数でないとすればこれは有理数であり、 $\sqrt{2}$ は有理数どうしの引き算 $\sqrt{2} = 1 - (1 - \sqrt{2})$ で表せる。したがって $\sqrt{2}$ は有理数となり、これは $\sqrt{2}$ が無理数であることに反する。□

体 K に関する線形代数学では、次で定める線形空間を対象とする。

定義 15.3.3. K を体とする。集合 V に対して和 $+$ 、および K の元に関するスカラー倍 \cdot の演算が定められており、さらに特別な元 $\mathbf{0} \in V$ が与えられているとする。これらの演算が次の条件 (ベクトル空間の公理) をすべて満たすとき、四つ組 $(V, \mathbf{0}, +, \cdot)$ を体 K 上のベクトル空間または線形空間と呼ぶ。

I. 各元 $\mathbf{a}, \mathbf{b} \in V$ に対して和 $\mathbf{a} + \mathbf{b} \in V$ が定まっており、次の性質を満たす:

$$(1) \mathbf{a} + \mathbf{b} = \mathbf{b} + \mathbf{a}, \quad (2) (\mathbf{a} + \mathbf{b}) + \mathbf{c} = \mathbf{a} + (\mathbf{b} + \mathbf{c}), \quad (3) \mathbf{a} + \mathbf{0} = \mathbf{0} + \mathbf{a} = \mathbf{a}.$$

II. 各元 $\mathbf{a} \in V$ および $r \in K$ に対してスカラー倍 $r \cdot \mathbf{a}$ が定まっており、次の性質を満たす:

$$(4) r \cdot (s \cdot \mathbf{a}) = (rs) \cdot \mathbf{a}, \quad (5) (r + s) \cdot \mathbf{a} = (r \cdot \mathbf{a}) + (s \cdot \mathbf{a}) \quad (6) r \cdot (\mathbf{a} + \mathbf{b}) = (r \cdot \mathbf{a}) + (r \cdot \mathbf{b})$$

$$(7) 1 \cdot \mathbf{a} = \mathbf{a}, \quad (8) 0 \cdot \mathbf{a} = \mathbf{0}.$$

本論で主題とする線形空間は \mathbb{R} 上の線形空間である。しかし、体として \mathbb{R} 以外のものを採用しても、多くの場合に同等の議論が得られることを頭の片隅に留めておきたい。実は、体として複素数を採用したほうが実数の場合よりも理論が綺麗になる部分がある。この点についてより詳しいことは固有値の項目で述べよう。

²⁶可換な環 K から零元を除いた集合が積演算に関して群となるとき、 K を体という。

四則演算が成立する体 K 自身は、体の演算としての和と積を線形空間における和とスカラー倍とみなすことで線形空間になる。例えば実数直線 \mathbb{R} は座標軸が1つしかない線形空間である。次に、二つの体 K, L が与えられており、 $K \subset L$ が成り立つ場合を考えよう。このとき大きい体 L は、小さい体 K 上の線形空間とみなすこともできる。例えば $\mathbb{R} \subset \mathbb{C}$ であることから \mathbb{C} は \mathbb{R} 上の線形空間となる。このことは複素平面を通して、 \mathbb{C} は \mathbb{R}^2 と対応づけられることから分かる。一方で、 $\mathbb{Q} \subset \mathbb{R}$ について \mathbb{R} を \mathbb{Q} 上の線形空間とみたとき、この線形空間の性質を調べるのは容易でない (実際、無限次元となる)。一般に、大小関係のある二つの体の間にある対称性 (すなわち群) を調べる理論を体論 (ガロア理論) という。

複素数体 \mathbb{C} の定義を思いだそう。実数の世界に方程式 $x^2 = -1$ を満たす元を新たに加え、さらに四則演算が成り立つよう数空間を広げることで \mathbb{C} を得る。 $x^2 = -1$ を満たす数は二つ存在し (これを a, b としよう)、このうち一方を虚数単位として i と書き、このときもう一方は $-i$ と書かれる。ここで一つの疑問が現れる。ある人が虚数単位として数 a を選び参考書 A を書き、別の人が虚数単位に b を選んで参考書 B を書いたとすれば、参考書 A の i は参考書 B の $-i$ に相当する。したがって、二つの参考書で述べられている議論を比較しようと思えば面倒な翻訳作業が必要なはずである。しかし、現実にはこのような作業は必要なく、参考書 A, B を並行して読む際に翻訳を意識する必要はない。これは不思議なことではないか。

体および群の概念が生まれた背景には、四則演算と根号のみを用いた5次方程式の解の公式の非存在証明があった。方程式の解を付け加えた体を考えて、解の対称性と方程式の関係を見極めることで、解の公式の非存在性が理解されたのである。また、上で述べた翻訳作業が必要ないことは、 $x^2 = -1$ の解の対称性を通して理解されている。

16 いろいろな線形部分空間

例 15.2.1 における $W_{[1,-1]}$ と \mathbb{R}^2 の関係や、例 15.2.4 における $\mathbb{R}[x]_n$ と $\mathbb{R}[x]$ の関係のように、より大きな線形空間の部分集合として実現される線形空間の例がいくつも考えられる。これらを総称する概念として線形部分空間なる概念を得る。

本節にて部分空間の数多くの例を紹介する。このことから、線形空間の枠組みで論じることのできる対象がいかに豊富であるか理解されることと思う。

16.1 定義

線形空間 V の部分集合 W が線形空間となるための条件を考えよう。 W における和とスカラー倍の演算がベクトル空間の公理を満たすことは、既に V における演算がそうであることから直ちに得られる。ゆえに、 W が線形空間となるためには、 W 内での演算結果が再び W に含まれること（このことを W が演算で閉じているという）、および零元があればよいことが分かる。こうして我々は次の定義に至る：

定義 16.1.1. 線形空間 $V = (V, \mathbf{0}_V, +, \cdot)$ の部分集合 W が次の性質 (i) から (iii) をすべて満たすとき、 $W = (W, \mathbf{0}_V, +, \cdot)$ もまた線形空間となる。

$$(i) \mathbf{0}_V \in W, \quad (ii) \mathbf{a}, \mathbf{b} \in W \implies \mathbf{a} + \mathbf{b} \in W, \quad (iii) \mathbf{a} \in W, r \in \mathbb{R} \implies r\mathbf{a} \in W.$$

このとき W を V の線形部分空間 (linear subspace) または部分ベクトル空間という。本論では、これらを部分空間と略称で述べる²⁷。

例 16.1.2. 線形空間 V に対して、 V 自身は V の部分空間である。また V の零元のみからなる集合 $\{\mathbf{0}\}$ も V の部分空間である。これら二つの部分空間のことを、 V の自明な部分空間という。

部分空間の例は本節の後半で述べる。その前に、部分空間になるための条件 (i)~(iii) をよく理解するために、 \mathbb{R}^2 の部分集合のなかで部分空間にならない例を挙げよう。条件 (i)~(iii) のいずれか一つでも満たさなければ部分空間にはなり得ないことから、否定的例はいくらでも簡単に列挙できる。そこで、三つの性質のうち二つは満たすものの、残りの一つを満たさないような例、つまり、あと一歩で部分空間にならない例をここでは考える。

例 16.1.3. (1) 条件 (i) と (ii) を満たすが (iii) を満たさない例:

$$W := \{(x, y) \in \mathbb{R}^2 \mid x \geq 0 \text{ かつ } y \geq 0\} \text{ とすれば } W \text{ は (i) と (ii) を満たし, (iii) を満たさない.}$$

Proof. $(x, y) = (0, 0)$ が条件「 $x \geq 0$ かつ $y \geq 0$ 」を満たすことから $\mathbf{0} = (0, 0) \in W$ である。また、 $\mathbf{w} = (x_1, y_1)$, $\mathbf{v} = (x_2, y_2)$ とし、 $\mathbf{w}, \mathbf{v} \in W$ と仮定すれば、 $x_1, y_1, x_2, y_2 \geq 0$ であり、ゆえに $x_1 + x_2 \geq 0$, $y_1 + y_2 \geq 0$ である。したがって、 $(x, y) = (x_1 + x_2, y_1 + y_2)$ は条件「 $x \geq 0$ かつ $y \geq 0$ 」を満たす。よって $\mathbf{w} + \mathbf{v} = (x_1 + x_2, y_1 + y_2) \in W$ 。つまり W は (ii) を満たす。 W が (iii) を満たさないことを示すには、(iii) を満たさない反例の一つ挙げればよい。例えば $\mathbf{a} = (1, 0)$, $r = -1$ とすれば $\mathbf{a} \in W$, $r \in \mathbb{R}$ である。しかしながら $r\mathbf{a} = -1(1, 0) = (-1, 0)$ であり、 $(x, y) = (-1, 0)$ は条件「 $x \geq 0$ かつ $y \geq 0$ 」を満たさないゆえ $r\mathbf{a} \notin W$ 。□

(2) 条件 (i) と (iii) を満たすが (ii) を満たさない例:

$$W := \{(x, y) \in \mathbb{R}^2 \mid x = 0 \text{ または } y = 0\} \text{ とすれば } W \text{ は (i) と (iii) を満たし, (ii) を満たさない.}$$

Proof. $(x, y) = (0, 0)$ が条件「 $x = 0$ または $y = 0$ 」を満たすことから $\mathbf{0} = (0, 0) \in W$ である。また、 $\mathbf{w} = (a, b) \in W$, $r \in \mathbb{R}$ と仮定すれば、 a, b のうち少なくともいずれか一方は 0 である。ゆえに ra, rb のいずれか一方は 0 であり、 $(x, y) = (rx, ry)$ は条件「 $x = 0$ または $y = 0$ 」を満たす。し

²⁷数学では様々な空間概念が与えられており、ゆえに何を対象としているかによって部分空間の意味は異なる。

たがって $rw = (rx, ry) \in W$ である。 W が (ii) を満たさないことを示そう。例えば、 $\mathbf{a} = (1, 0)$, $\mathbf{b} = (0, 1)$ とすれば $\mathbf{a}, \mathbf{b} \in W$ である。 $\mathbf{a} + \mathbf{b} = (1, 0) + (0, 1) = (1, 1)$ であり、 $(x, y) = (1, 1)$ は条件「 $x = 0$ または $y = 0$ 」を満たさないゆえ $\mathbf{a} + \mathbf{b} \notin W$. \square

(3) 条件 (ii) と (iii) を満たすが (i) を満たさない例:

(iii) において $r = 0$ を適用することで (i) が得られるゆえ、このような例は存在しないと考えたいところである。しかし、実際には次の例が与えられる:

W を空集合とすれば、 W は条件 (ii) と (iii) を満たすが (i) を満たさない。

Proof. 空集合は元を含まない集合ゆえ、とくに零元も含まず、したがって W は (i) を満たさない。(ii) を満たすことは次のように背理法で示される。もし仮に (ii) を満たさないとすれば、それは $\mathbf{w}, \mathbf{v} \in W$ であるにも関わらず $\mathbf{w} + \mathbf{v} \notin W$ となる例があるということである。この例において、とくに $\mathbf{w} \in W$ であり、したがって W は元を含む。これは W が元を含まない集合であったことに反する。以上より W は (ii) を満たさねばならない。同様の論法を用いて、 W が (iii) を満たすことも示される。 \square

よりみち (前提が偽なる命題).

例 16.1.3 (3) において空集合が条件 (ii) を満たすことの説明として「(*) 前提が満たされない命題は、いかなる結論が述べられていても正しい」という論理の原則を持ち出すことが多い。いまの例では、前提となる $\mathbf{a}, \mathbf{b} \in W$ が成立しないゆえ (ii) は真であるという考え方である。しかしながら、この論理の原則を盲目的に認める立場に立って学ぶのであれば、それは迷信を信仰しているに等しい。原則 (*) が認められるゆえんは何か、自らの言葉で咀嚼することが学習者に望まれている。

ところで、前提が満たされない議論があるということは、何かの反例を挙げる際にも注意が必要ということである。何故なら、反例を挙げる手順を説明したつもりでいても、そのような手順を踏める対象が現実には存在しない可能性があるからである。空論にならないためにも、反例は具体的な例を挙げることによってなさねばならない。

さて、原則 (*) は次のようにして説明される。

命題 16.1.4. F を偽なる条件とする。任意の条件 P について「 F ならば P 」は成立する。

Proof. 「 F ならば P 」を示すために F を仮定しよう。すると「 F または P 」であることが認められる。すなわち、 F と P のうち少なくともいずれか一方が成立することになる。ところが F は偽なる条件であったゆえ成立せず、したがって P が成り立たねばならない。以上より P が導かれた。 \square

例 16.1.3(3) や、これと似た論法を用いた命題 14.1.4 では背理法による説明をしていた。他にも、対偶「 $(P$ でない) ならば $(F$ でない)」は結論が正しい命題ゆえ真であり、ゆえにもとの「 F ならば P 」も真、と説明することもできるだろう。論理の原則に更に踏み込んで、背理法による論法や対偶の同値性が認められるのは何故だろうか。そこには、条件「 A 」とその否定「 A でない」において、一方が成立しなければもう一方は成立すると考える立場 (これを排中律という) が背景にある。

部分空間となるための条件 (ii) と (iii) は次のようにまとめることができる。

命題 16.1.5. 定義 16.1.1 における条件 (ii) と (iii) が共に成立することと、次の条件が成立することは同値である:

$$(iv) \quad r, s \in \mathbb{R}, \mathbf{a}, \mathbf{b} \in W \implies r\mathbf{a} + s\mathbf{b} \in W.$$

Proof. (ii) と (iii) を仮定して (iv) を示そう. $r, s \in \mathbb{R}$, $\mathbf{a}, \mathbf{b} \in W$ とすれば, (iii) より $r\mathbf{a}, s\mathbf{b} \in W$ である. これに (ii) を適用し $r\mathbf{a} + s\mathbf{b} \in W$ を得る. すなわち (iv) が成り立つ. また, (iv) において $r = s = 1$ という特別な場合が (ii) に相当し, $s = 0$ なる場合が (iii) に相当する. すなわち (iv) ならば (ii) かつ (iii) である. \square

以降, 部分空間であることを確認する際は条件 (ii) と (iii) の代わりに条件 (iv) を用いよう. これにより証明が多少は短くなるであろう. また, 条件 (iv) からは更に次の性質が導かれる.

命題 16.1.6. W を線形空間 V の部分空間とすれば, 各 $\mathbf{v}_i \in W$ および $r_i \in \mathbb{R}$ について $\sum_{i=1}^{\ell} r_i \mathbf{v}_i \in W$.

Proof. 和の個数 ℓ に関する帰納法で示す. $\ell = 1$ の場合は部分空間の性質 (iii) に他ならない. 和の個数が ℓ のとき成立すると仮定し, 和の個数が $\ell + 1$ の場合を示そう. $\mathbf{v}_i \in W$, $r_i \in \mathbb{R}$ ($i = 1, \dots, \ell + 1$) とし, $\mathbf{u} = \sum_{i=1}^{\ell+1} r_i \mathbf{v}_i$ とすれば, $\mathbf{u} = (\sum_{i=1}^{\ell} r_i \mathbf{v}_i) + r_{\ell+1} \mathbf{v}_{\ell+1}$ と書ける, 帰納法の仮定より $\sum_{i=1}^{\ell} r_i \mathbf{v}_i \in W$ であり, 性質 (iii) より $r_{\ell+1} \mathbf{v}_{\ell+1} \in W$ である. よって, 性質 (ii) より $(\sum_{i=1}^{\ell} r_i \mathbf{v}_i) + r_{\ell+1} \mathbf{v}_{\ell+1} \in W$, すなわち $\mathbf{u} \in W$ である. \square

17 節で線形結合と呼ばれる概念を導入する. これを用いて, 上の命題で述べている W の性質は「 W は線形結合で閉じている」と呼ばれる.

16.2 \mathbb{R}^n の部分空間

\mathbb{R}^n の部分空間の外延的表示と内包的表示について論じる.

例 16.2.1. \mathbb{R}^3 において互いに平行²⁸でないベクトル $\mathbf{a}, \mathbf{b} \in \mathbb{R}^3$ を取ると (ただし $\mathbf{a}, \mathbf{b} \neq \mathbf{0}$), $W = \{r_1 \mathbf{a} + r_2 \mathbf{b} \mid r_1, r_2 \in \mathbb{R}\}$ によって \mathbb{R}^3 上の原点 $\mathbf{0}$ および \mathbf{a}, \mathbf{b} を含む平面が定まる. この W は \mathbb{R}^3 の部分空間である.

Proof. 部分空間となるための条件 (i) および (iv) が満たされることを確認すればよい.

(i): $r_1 = r_2 = 0$ とすることで $\mathbf{0} = r_1 \mathbf{a} + r_2 \mathbf{b}$ と表せる. ゆえに $\mathbf{0} \in W$ である.

(iv): $\mathbf{x}, \mathbf{y} \in W$ とすれば $x_1, x_2, y_1, y_2 \in \mathbb{R}$ を用いて $\mathbf{x} = x_1 \mathbf{a} + x_2 \mathbf{b}$, $\mathbf{y} = y_1 \mathbf{a} + y_2 \mathbf{b}$ と書ける. 各 $r, s \in \mathbb{R}$ について

$$r\mathbf{x} + s\mathbf{y} = r(x_1 \mathbf{a} + x_2 \mathbf{b}) + s(y_1 \mathbf{a} + y_2 \mathbf{b}) = (rx_1 + sy_1)\mathbf{a} + (rx_2 + sy_2)\mathbf{b}$$

である. すなわち, $r\mathbf{x} + s\mathbf{y}$ は $r_1 = rx_1 + sy_1$, $r_2 = rx_2 + sy_2$ とすることで $r_1 \mathbf{a} + r_2 \mathbf{b}$ と表せる. ゆえに $r\mathbf{x} + s\mathbf{y} \in W$ である. \square

例 16.2.1 にけるの証明において, \mathbf{a}, \mathbf{b} が平行でないことは用いられていない. \mathbf{a}, \mathbf{b} が平行である場合は, W は \mathbb{R}^3 の原点を通る直線になる. 一般に, \mathbb{R}^3 の部分空間とは, 原点を通る平面および原点を通る直線, 自明な部分空間 (\mathbb{R}^3 と $\{\mathbf{0}\}$) の四種に限られる. この分類は次元なる概念を通してなされる (例 22.4.2).

より一般に, 線形空間 V および $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ について $W = \{\sum_{i=1}^n r_i \mathbf{u}_i \mid r_1, \dots, r_n \in \mathbb{R}\}$ は V の部分空間となる (命題 18.1.2). この例が部分空間の外延的表示であるのに対して, 内包的表示は次で与えられる. これは例 15.2.1 の一般化に相当している²⁹.

命題 16.2.2. (m, n) -行列 A に対して, 次で定められる W_A は \mathbb{R}^n の部分空間となる:

$$W_A := \{\mathbf{x} \in \mathbb{R}^n \mid A\mathbf{x} = \mathbf{0}\}.$$

²⁸線形空間における $\mathbf{0}$ でない二つのベクトル \mathbf{a}, \mathbf{b} が平行であるとは, $r\mathbf{a} = \mathbf{b}$ を満たす実数 r が存在することと定める.

²⁹例 15.2.1 は, $(1, 2)$ -行列 $A = [1, -1]$ の場合に相当する. なお, 例 15.2.1 では行ベクトルで表示していたが, ここでは列ベクトルによる表示を考えている.

Proof. 部分空間となるための条件 (i) および (iv) が満たされることを確認すればよい.

(i): $\mathbf{x} = \mathbf{0} \in \mathbb{R}^n$ とすれば, \mathbf{x} は条件 $A\mathbf{x} = \mathbf{0}$ を満たす. ゆえに $\mathbf{0} \in W_A$ である.

(iv): $\mathbf{a}, \mathbf{b} \in W_A, r, s \in \mathbb{R}$ とする. このとき, W_A の定義から $A\mathbf{a} = \mathbf{0}, A\mathbf{b} = \mathbf{0}$ が成り立つ. このとき $A(r\mathbf{a} + s\mathbf{b}) = r(A\mathbf{a}) + s(A\mathbf{b}) = r\mathbf{0} + s\mathbf{0} = \mathbf{0}$ ゆえ, $\mathbf{x} = r\mathbf{a} + s\mathbf{b}$ は条件 $A\mathbf{x} = \mathbf{0}$ を満たす. ゆえに $r\mathbf{a} + s\mathbf{b} \in W_A$ である. \square

上の W_A は斉次形連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ の解全体の空間に一致する. そこで, W_A は方程式 $A\mathbf{x} = \mathbf{0}$ の解空間と呼ばれる.

例 16.2.3. $A \neq B$ かつ $W_A = W_B$ なる例は山のようにある. 例えば, A, B が共に可逆正方行列ならば, それぞれの解空間 W_A, W_B は唯一解 $\mathbf{x} = \mathbf{0}$ のみからなる空間 $\{\mathbf{0}\}$ となる.

解空間の幾何的な意味を検討しよう. 列ベクトル $\mathbf{x} \in \mathbb{R}^n$ が $A\mathbf{x} = \mathbf{0}$ を満たすということは, A の各行と \mathbf{x} との行列としての積が 0 になるということである. これを \mathbb{R}^n 上のベクトルの内積と読み替えることにより, W_A の元であることは, A の各行と直交するベクトルであることと同値になることが分かる. この事実から, \mathbb{R}^n の任意の部分空間が W_A の形で表されることが示唆される. 例えば, 例 16.2.1 において, ベクトル \mathbf{a}, \mathbf{b} と共に直交するベクトル $\mathbf{c} \in \mathbb{R}^3$ を一つ取ると³⁰, \mathbf{c} の成分を横に並べた (1,3)-行列を A とすれば, $W = W_A$ となることが予想される. 詳しい説明は内積空間の節で述べよう.

練習 16.2.4. 次の \mathbb{R}^2 の部分空間 W について, $W = W_A$ なる行列 A を求めよ.

(1) $W = \{\mathbf{0}\}$. 解答例: 可逆行列ならば何でもよい. 例えば $A = E_2$ とせよ.

(2) $W = \mathbb{R}^2$. 解答例: $A = O_{2,2}, A = O_{1,2}$ などとすればよい.

(3) $W = \{r\mathbf{y} \mid r \in \mathbb{R}\}$, ただし $\mathbf{y} = \begin{bmatrix} a \\ b \end{bmatrix} \neq \mathbf{0}$ とする (W は原点 $\mathbf{0}$ と \mathbf{y} を通る直線上の点全体を表す).

解答例: $\mathbf{z} = \begin{bmatrix} -b \\ a \end{bmatrix}$ は \mathbf{y} に直交するベクトルである. \mathbf{z} の成分を横にならべた (1,2)-行列 $A = \begin{bmatrix} -b & a \end{bmatrix}$ について, $W = W_A$ となる. $W = W_A$ を示すために, $W \subset W_A$ および $W_A \subset W$ を示そう.

($W \subset W_A$): $\mathbf{x} \in W$ を勝手に取れば, 実数 r を用いて $\mathbf{x} = r\mathbf{y}$ と書ける. このとき,

$$A\mathbf{x} = A(r\mathbf{y}) = rA\mathbf{y} = r \begin{bmatrix} -b & a \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = r(-ba + ab) = r0 = 0.$$

ゆえに $\mathbf{x} \in W_A$ である.

($W_A \subset W$): $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in W$ を勝手に取れば, $A\mathbf{x} = \mathbf{0}$ が成り立っている. すなわち, $-bx_1 + ax_2 = 0$ である. $\mathbf{y} \neq \mathbf{0}$ ゆえ $a \neq 0$ または $b \neq 0$ である. $a \neq 0$ の場合は実数 $r = \frac{x_1}{a}$ について $x_1 = ra$ であり, これと先の式を合わせて $ax_2 = bx_1 = b(ra) = rba$. $a \neq 0$ ゆえ両辺を a で割り, $x_2 = rb$ を得る. 以上より, $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} ra \\ rb \end{bmatrix} = r\mathbf{y}$. ゆえに $\mathbf{x} \in W$ である. $b \neq 0$ の場合は, $r = \frac{x_2}{b}$ とおいて同様の計算をすれば $\mathbf{x} \in W$ を得る. \square

³⁰このような \mathbf{c} が取れるかどうかという問題も解決しなければならない. \mathbb{R}^3 に限定した話では, \mathbf{a} と \mathbf{b} の外積と呼ばれるベクトル $\mathbf{a} \times \mathbf{b}$ が, この例の一つに相当する.

16.3 部分空間の様々な例

より高度な部分空間の例を紹介しよう.

例 16.3.1. 例 15.2.2 で与えた数列空間 $\mathbb{R}^{\mathbb{N}}$ の部分集合で, 次の漸化式を満たす数列全体を F とする:

$$x_{n+2} = x_{n+1} + x_n. \quad (n \in \mathbb{N}) \quad (16.3.1)$$

このとき, F は $\mathbb{R}^{\mathbb{N}}$ の部分空間である.

Proof. (i): すべての項が 0 なる数列 $\mathbf{0} = (0, 0, 0, \dots)$ は漸化式 16.3.1 を満たす. ゆえに $\mathbf{0} \in F$.

(iv): $r, s \in \mathbb{R}$ とし, 数列 $\mathbf{x} = (x_n)_{n \in \mathbb{N}}$ および $\mathbf{y} = (y_n)_{n \in \mathbb{N}}$ が漸化式 16.3.1 を満たすと仮定する. すなわち, $x_{n+2} = x_{n+1} + x_n$ および $y_{n+2} = y_{n+1} + y_n$ が成立している. この二式をそれぞれ r 倍, s 倍して和をとることで

$$rx_{n+2} + sy_{n+2} = (rx_{n+1} + sy_{n+1}) + (rx_n + sy_n)$$

を得る. これは数列 $r\mathbf{x} + s\mathbf{y} = (rx_n + sy_n)_{n \in \mathbb{N}}$ が漸化式 16.3.1 を満たすことに他ならない. ゆえに $r\mathbf{x} + s\mathbf{y} \in F$. \square

漸化式 16.3.1 を満たす数列の中で, 初項が 0 かつ第 2 項が 1 なる数列

$$a_1 = 0, a_2 = 1, a_3 = 1, a_4 = 2, a_5 = 3, a_6 = 5, a_7 = 8, a_9 = 13, a_{10} = 21, \dots$$

をフィボナッチ数列 (Fibonacci sequence) という.

練習 16.3.2. 自然数 $k \geq 2$, および k 個の実数 a_0, a_2, \dots, a_{k-1} を固定する. 数列空間 $\mathbb{R}^{\mathbb{N}}$ の部分集合で, 次の漸化式を満たす数列全体を W とする:

$$x_{n+k} = a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \dots + a_1x_{n+1} + a_0x_n \quad (n \in \mathbb{N}) \quad (16.3.2)$$

このとき, W が $\mathbb{R}^{\mathbb{N}}$ の部分空間であることを示せ.

上の漸化式のことを線形漸化式と呼び, これを満たす数列のことを線形回帰数列という.

次に, 微分を用いた部分空間の例を挙げる.

例 16.3.3. (1) \mathbb{R} 上の开区間 I を定義域とする C^∞ -級関数³¹全体のなす集合を $C^\infty(I)$ と書く. 定数関数 $\mathbf{0}$ は C^∞ -関数であり, また C^∞ -級関数の和やスカラー倍は再び C^∞ -関数となる. もちろん C^∞ -級関数は連続関数である. 以上より $C^\infty(I)$ は例 15.2.6 で与えた $C(I)$ の部分空間である.

(2) $C^\infty(I)$ の元の中で, 次の微分方程式を満たす関数 $y = y(x)$ 全体を W とする:

$$y^{(2)}(x) = y^{(1)}(x) + y^{(0)}(x) \quad (16.3.3)$$

ここで, $y^{(n)}(x) = \frac{d^n}{dx^n}y(x)$ は y の n 階導関数のことであり, $y^{(0)}(x) = y(x)$ とする. このとき, W は $C^\infty(I)$ の部分空間である.

Proof. (i): 定数関数 $\mathbf{0}$ の微分は再び $\mathbf{0}$ となることから, $\mathbf{0}$ が式 16.3.3 を満たすことは明らかである. ゆえに $\mathbf{0} \in W$.

(iv) $f, g \in W$ とすれば関数 f, g は式 16.3.3 を満たす. すなわち, $f^{(2)}(x) = f^{(1)}(x) + f^{(0)}(x)$, $g^{(2)}(x) = g^{(1)}(x) + g^{(0)}(x)$ である. 各 $r, s \in \mathbb{R}$ に対して, この二式をそれぞれ r 倍, s 倍して和をとることで

$$\begin{aligned} rf^{(2)}(x) + sg^{(2)}(x) &= \left(rf^{(1)}(x) + sg^{(1)}(x) \right) + \left(rf^{(0)}(x) + sg^{(0)}(x) \right) \\ (rf(x) + sg(x))^{(2)} &= (rf(x) + sg(x))^{(1)} + (rf(x) + sg(x))^{(0)} \end{aligned}$$

を得る. これは関数 $rf(x) + sg(x)$ が微分方程式 16.3.3 を満たすことに他ならない. ゆえに $rf + sg \in W$. \square

³¹何回でも微分できる関数のことを C^∞ -関数という

練習 16.3.4. 自然数 $k \geq 2$, および k 個の実数 a_0, a_2, \dots, a_{k-1} を固定する. 次の微分方程式³²を満たす関数 $y = y(x)$ 全体を W とする. W が $C^\infty(I)$ の部分空間であることを示せ.

$$y^{(k)}(x) = a_{k-1}y^{(k-1)} + a_{k-2}y^{(k-2)} + \dots + a_1y^{(1)}(x) + a_0y^{(0)}(x). \quad (16.3.4)$$

式 16.3.3 の両辺を n 回微分することで $y^{(n+2)}(x) = y^{(n+1)}(x) + y^{(n)}(x)$ を得る. ここに漸化式 16.3.1 の類似性が伺えよう. 同様にして, 微分方程式 16.3.4 から

$$y^{(n+k)}(x) = a_{k-1}y^{(n+k-1)} + a_{k-2}y^{(n+k-2)} + \dots + a_1y^{(n+1)}(x) + a_0y^{(n)}(x)$$

が得られ, これは漸化式 16.3.2 と類似している. これらは式として単に似ているというだけではなく, 線形回帰数列の一般項の解法と線形常微分方程式の一般解の解法が線形代数学の枠組みにおいて並行して得られることが後に理解されるであろう.

例 16.3.5 (発展). (1) $\mathbb{R}[x]$ の中で, 多項式 $p(x) = x - 1$ で割り切れるもの全体を集めた集合を (p) と書く. すると (p) は $\mathbb{R}[x]$ の部分空間である. この議論において p が特別な多項式 $x - 1$ である必要はなく, 任意の多項式 $p(x)$ について同様の議論が成り立つ.

(2) \mathbb{R} 上の区間 I の元 t を一つ固定しておく, さらに, I を定義域とする連続関数 $f: I \rightarrow \mathbb{R}$ で, $f(t) = 0$ を満たすもの全体からなる集合を \mathcal{I}_t とおこう. このとき, \mathcal{I}_t は $C(I)$ の部分空間となる.

上の例において, $\mathbb{R}[x]$ の元と (p) の元の積を取ると再び (p) の元に含まれることがわかる. また, $C(I)$ の元と \mathcal{I}_t の元の積もまた \mathcal{I}_t に含まれる. このような特別な集合はイデアル (ideal) と呼ばれ, 環論と呼ばれる代数分野において広く調べられている.

³²方程式 16.3.4 は実数係数の線形常微分方程式と呼ばれる.

17 線形結合と線形独立性

線形空間の定義から直ちに導かれる基本的性質、すなわちベクトルの和とスカラー倍の性質について論じよう。本節で述べることは、形のうえではベクトルの式変形を繰り返すことに尽きる。しかし、これを単なる計算と見るのではなく、複数のベクトルの間の関係性と捉えることで、技法から理論へと考え方が昇華されるのである。これによって、より高い見地から線形空間を捉えられるようになる。なお、ベクトルたちの関係を調べるうえで技術的な部分のいくつかは連立1次方程式の掃き出し法による解法に帰着される。掃き出し法や行列の簡約化をしっかり復習したうえで本節に臨んでもらいたい。

17.1 線形結合

定義 17.1.1. 線形空間 V の元 $v \in V$ が³³, $u_1, u_2, \dots, u_n \in V$ および $a_1, a_2, \dots, a_n \in \mathbb{R}$ を用いて

$$v = a_1 u_1 + a_2 u_2 + \dots + a_n u_n = \sum_{i=1}^n a_i u_i$$

と書けるとき、 v は u_1, \dots, u_n の線形結合³³(linear combination) で書けるという。また、零ベクトルを線形結合で表す式 $\sum_{i=1}^n a_i u_i = \mathbf{0}$ が成立するとき、この等式を u_1, \dots, u_n による線形関係 (または1次関係) と呼ぶ。

線形結合の標語的な解釈

v が u_1, \dots, u_n の線形結合で書かれるとは、 v の情報が u_1, \dots, u_n の情報に分解されていると考えればよい。もう少し詳しく述べれば、 v の情報を得るには、 u_1, \dots, u_n の情報と線形結合に現れる係数 a_1, \dots, a_n の値さえ分かれば十分ということである。なお、ここでいう「情報」とは、分析すべき線形写像における値のことを意味する。

例 17.1.2. \mathbb{R}^n の任意のベクトルは、次の n 個のベクトルの線形結合で書くことができる:

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots, \quad e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

実際、各 $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} \in \mathbb{R}^n$ について $\begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ a_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \dots + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ a_n \end{bmatrix} = a_1 e_1 + \dots + a_n e_n = \sum_{i=1}^n a_i e_i$ で

ある (実は、同様の分解を式 12.4.1 において既に行っている)。上の e_1, \dots, e_n を \mathbb{R}^n の標準ベクトルまたは基本ベクトルという。

ベクトル v が u_1, \dots, u_n の線形結合で書けるとしよう。 v の情報を u_1, \dots, u_n たちの情報に還元する際に、 u_1, \dots, u_n の中でその情報が不要なものがあるかもしれない。例えば次のような状況が考えられる。

要不要論

³³線形結合は1次結合とも呼ばれる。

(1) $\mathbf{u}_1, \dots, \mathbf{u}_n$ のいずれかの情報が、ほかのベクトルたちの情報に分解できる場合.

例えば、 \mathbf{u}_n の情報が $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ に分解されるとき、 \mathbf{v} の情報は $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の情報だけから復元できることが示唆され、 \mathbf{u}_n は不要となる. 実際、 \mathbf{v} が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けており、更に \mathbf{u}_n が $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書けるならば、 \mathbf{v} は $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書ける:

$\mathbf{v} = \sum_{i=1}^n a_i \mathbf{u}_i$, $\mathbf{u}_n = \sum_{i=1}^{n-1} b_i \mathbf{u}_i$ と書けるならば、

$$\mathbf{v} = \sum_{i=1}^n a_i \mathbf{u}_i = \left(\sum_{i=1}^{n-1} a_i \mathbf{u}_i \right) + a_n \mathbf{u}_n = \left(\sum_{i=1}^{n-1} a_i \mathbf{u}_i \right) + a_n \left(\sum_{i=1}^{n-1} b_i \mathbf{u}_i \right) = \sum_{i=1}^{n-1} (a_i + a_n b_i) \mathbf{u}_i.$$

(2) 線形結合に現れる係数 a_i のうちのいくつかは 0 の場合.

このとき、 $a_i = 0$ に対応する \mathbf{u}_i は不要となる. 例えば、 $\mathbf{a} = (5, 3, 0) \in \mathbb{R}^3$ を標準ベクトルに分解すると、 $\mathbf{a} = 5\mathbf{e}_1 + 3\mathbf{e}_2 + 0\mathbf{e}_3 = 5\mathbf{e}_1 + 3\mathbf{e}_2$ ゆえ \mathbf{e}_3 は不要である. しかしこれは、特別なベクトルを考えたからたまたま \mathbf{e}_3 が不要になったのであり、 \mathbb{R}^3 の別のベクトルを分解しようと思えば、 \mathbf{e}_3 が必要になることもある. とくに、 \mathbb{R}^3 のすべてのベクトルを分解しようと思えば、 $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ のいずれも必須であることが分かる. この考察から、(2) の立場で \mathbf{u}_n が不要ということは、線形空間 V のいかなる元も \mathbf{u}_n を用いずに $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ のみによって分解できることと捉えるべきである. 仮にこの意味で \mathbf{u}_n が不要になる場合、 \mathbf{u}_n はとくに V の元であるから $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ たちの線形結合で書ける. すなわち、この議論の大筋については (1) に帰着される.

注意: 逆に、(1) は (2) に帰着するとも説明できる. 実際、(1) の最後の式は、 \mathbf{u}_n を用いない線形結合になっている.

ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に不要なものがあるかないかという状況を数学的な言葉で表すために、線形独立性なる概念を導入する.

17.2 線形独立性

以下、断りがなくとも V は線形空間であるとする.

定義 17.2.1. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ が線形独立³⁴(linearly independent) であるとは、

$$a_1, \dots, a_n \in \mathbb{R}, a_1 \mathbf{u}_1 + a_2 \mathbf{u}_2 + \dots + a_n \mathbf{u}_n = \mathbf{0} \implies a_1 = a_2 = \dots = a_n = 0$$

が成り立つことをいう. 線形独立でないベクトルの組を線形従属 (linearly dependent) であるという. すなわち、少なくともいずれか一つは 0 でないような実数の組 a_1, \dots, a_n を用いて、 $a_1 \mathbf{u}_1 + \dots + a_n \mathbf{u}_n = \mathbf{0}$ と表せることである.

線形従属の定義に現れた、少なくともいずれか一つは 0 でないような実数の組 a_1, \dots, a_n のことを自明でない組という. 実数の組 a_1, \dots, a_n が自明でないことは、 $(a_1, \dots, a_n) \neq \mathbf{0}$ であることに等しい. また、自明でない実数の組 a_1, \dots, a_n を係数とする線形関係を自明でないという. 更にこれらの否定概念として、自明な実数の組、および自明な線形関係を定める. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であるとは、それらによる線形関係が自明なものに限られることである. また、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形従属であるとは、自明でない線形関係が存在することである.

例 17.2.2. (1) \mathbb{R}^n の標準ベクトルの組 $\mathbf{e}_1, \dots, \mathbf{e}_n$ は線形独立である.

Proof. $a_1 \mathbf{e}_1 + \dots + a_n \mathbf{e}_n = \mathbf{0}$ とすれば $\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$ ゆえ $a_1 = \dots = a_n = 0$ である. □

³⁴線形独立 (線形従属) は 1 次独立 (1 次従属) とも呼ばれる.

- (2) m 次列ベクトル $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であることと, m 次行ベクトル ${}^t\mathbf{u}_1, \dots, {}^t\mathbf{u}_n$ が線形独立であることは同値である. 実際, 式 $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_{m,1}$ と式 $\sum_{i=1}^n a_i {}^t\mathbf{u}_i = \mathbf{0}_{1,m}$ は互いに両辺を転置した関係にある. ゆえに, これらのベクトルの組に関する線形関係は同等である.
- (3) $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に零ベクトルがあれば, これらは線形従属である. 例えば $\mathbf{u}_n = \mathbf{0}$ の場合, 自明でない実数の組 $0, \dots, 0, 1$ を係数とする線形関係 $0\mathbf{u}_1 + \dots + 0\mathbf{u}_{n-1} + 1\mathbf{u}_i = \mathbf{0}$ が成立する. また, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に重複がある場合も線形従属である. 例えば $\mathbf{u}_1 = \mathbf{u}_2$ の場合, 自明でない実数の組 $1, -1, 0, \dots, 0$ を係数とする線形関係 $1\mathbf{u}_1 + (-1)\mathbf{u}_2 + 0\mathbf{u}_3 + \dots + 0\mathbf{u}_i = \mathbf{0}$ が成立する.

例 17.2.3. 多項式環 $\mathbb{R}[x]$ における 3 つの元 $x^2, x, \mathbf{1}$ は線形独立である. ここで, $\mathbf{1}$ とはどんな数を代入しても $1 \in \mathbb{R}$ に値を取る定数関数とする.

Proof. $a_2, a_1, a_0 \in \mathbb{R}$, $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ と仮定し, $a_2 = a_1 = a_0 = 0$ を示そう. $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ の両辺に適当な数を三つほど代入して連立 1 次方程式を解くことで $a_2 = a_1 = a_0 = 0$ を示せるが (命題 15.2.5), ここでは微分による証明を紹介しよう. $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ の両辺を微分すると $2a_2x + a_1 = \mathbf{0}$ であり, これを更に微分することで $2a_2 = \mathbf{0}$ を得る. つまり, いかなる数を代入しても $2a_2$ に値を取る定数関数 (左辺) と 0 に値を取る定数関数 (右辺) は等しい. ゆえに $a_2 = 0$ である. これを $2a_2x + a_1 = \mathbf{0}$ に代入して $a_1 = 0$ を得る. これらをもとの式 $a_2x^2 + a_1x + a_0\mathbf{1} = \mathbf{0}$ に代入することで $a_0 = 0$ を得る. \square

練習 17.2.4. $n + 1$ 個の組 $x^n, x^{n-1}, \dots, x, \mathbf{1} \in \mathbb{R}[x]$ が線形独立であることを帰納法を用いて示せ.

次の命題は前項の最後で考察した要不要論と線形独立性 (従属性) の関係を述べている. すなわち, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に不要なものがあることと $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形従属性は同値である. また, それらの否定を取り, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中に不要なものがないことと $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性は同値である.

命題 17.2.5. n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ について次の (1) と (2) は同値である.

- (1) $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である,
- (2) $\mathbf{u}_1, \dots, \mathbf{u}_n$ のうち少なくとも一つのベクトルが他の $n - 1$ 個の線形結合で書ける.

Proof. (1) \Rightarrow (2): $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形従属であるとすれば, 自明でない実数の組 a_1, \dots, a_n を用いて $a_1\mathbf{u}_1 + \dots + a_n\mathbf{u}_n = \mathbf{0}$ とできる. このとき a_1, \dots, a_n のうちいずれか一つは 0 ではない. 例えば $a_n \neq 0$ として話を進めよう. このとき, 移項により $a_n\mathbf{u}_n = -a_1\mathbf{u}_1 - \dots - a_{n-1}\mathbf{u}_{n-1}$ を得る. $a_n \neq 0$ ゆえこの両辺を a_n で割れば $\mathbf{u}_n = \left(-\frac{a_1}{a_n}\right)\mathbf{u}_1 + \dots + \left(-\frac{a_{n-1}}{a_n}\right)\mathbf{u}_{n-1}$. ゆえに \mathbf{u}_n は, $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書ける. $a_i \neq 0$ ($i = 1, \dots, n - 1$) である場合も同様にして, \mathbf{u}_i が他の $n - 1$ 個の線形結合で書けることが示される.

(2) \Rightarrow (1): \mathbf{u}_n が $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ の線形結合で書ける場合を考えよう. このとき, ある実数の組 a_1, \dots, a_{n-1} を用いて $\mathbf{u}_n = a_1\mathbf{u}_1 + \dots + a_{n-1}\mathbf{u}_{n-1}$ と書ける. これを移項して $a_1\mathbf{u}_1 + \dots + a_{n-1}\mathbf{u}_{n-1} + (-1)\mathbf{u}_n = \mathbf{0}$ を得る. 自明でない線形関係が得られたゆえ, $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である. \mathbf{u}_n 以外のベクトルが他の $n - 1$ 個の線形結合で書ける場合についても同様の議論により $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形従属性を得る. \square

例 17.2.6. (1) \mathbb{R}^2 の $\mathbf{0}$ でない列ベクトル \mathbf{u}, \mathbf{v} が線形従属であるとは, 一方がもう一方の線形結合で書けるということであるから, これは実数 $r \neq 0$ を用いて $\mathbf{u} = r\mathbf{v}$ と書けること, すなわち \mathbf{u}, \mathbf{v} が原点を通る同一直線上にあることを意味する.

- (2) 一方, \mathbb{R}^2 の列ベクトル \mathbf{u}, \mathbf{v} が線形独立であるとは, \mathbf{u}, \mathbf{v} が原点を通る同一直線上にないこと, つまり \mathbf{u}, \mathbf{v} で張られる平行四辺形が面積を持つことを意味する. これは行列 $A = [\mathbf{u}, \mathbf{v}]$ が可逆であることにほかならない. この事実を n 次に一般化した場合の証明は命題 17.3.6 で与える.

- (3) \mathbb{R}^2 の 3 つのベクトル $\mathbf{u}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $\mathbf{u}_2 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, $\mathbf{u}_3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ は線形従属である. これは例えば, $\mathbf{u}_3 = \frac{2}{5}\mathbf{u}_1 + \frac{1}{5}\mathbf{u}_2$ と書けることから分かる. 実は, \mathbb{R}^n の $n + 1$ 個のベクトルの組は必ず線形従属になる (命題 17.3.5).

練習 17.2.7. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ について次を示せ.

- (1) n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立ならば, そこから一つ取り除いた $n-1$ 個の組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ も線形独立である.

解答例: $a_1\mathbf{u}_1 + \dots + a_{n-1}\mathbf{u}_{n-1} = \mathbf{0}$ とすれば, $a_n = 0$ とおくことで $a_1\mathbf{u}_1 + \dots + a_{n-1}\mathbf{u}_{n-1} + a_n\mathbf{u}_n = \mathbf{0}$ を得る. $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性より $a_1 = \dots = a_{n-1} = a_n = 0$ であり, 特に $a_1 = \dots = a_{n-1} = 0$. ゆえに $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ は線形独立である. \square

- (2) $n-1$ 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}$ が線形従属ならば, そこに新たなベクトル \mathbf{u}_n を加えた n 個のベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-1}, \mathbf{u}_n$ も線形従属である.

解答例: (2) は (1) の対偶にほかならない. \square

17.3 線形独立性の判定

\mathbb{R}^m のベクトルの組の線形独立性の判定法を与える.

命題 17.3.1. n 個の m 次列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ が線形独立であることと, (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ に関する斉次形連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ が唯一解を持つことは同値である.

Proof. $\mathbf{a}_1, \dots, \mathbf{a}_n$ が線形独立であるとする. n 次列ベクトル $\mathbf{x} = {}^t(r_1, \dots, r_n)$ が $A\mathbf{x} = \mathbf{0}$ を満たすならば,

$$r_1\mathbf{a}_1 + \dots + r_n\mathbf{a}_n = [\mathbf{a}_1, \dots, \mathbf{a}_n] \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} = A\mathbf{x} = \mathbf{0}.$$

$\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形独立性より $r_1 = \dots = r_n = 0$ である. つまり $\mathbf{x} = \mathbf{0}$ であり, 方程式 $A\mathbf{x} = \mathbf{0}$ の解は自明なものに限る.

次に, 方程式 $A\mathbf{x} = \mathbf{0}$ が自明な解しか持たないと仮定する. このとき $r_1\mathbf{a}_1 + \dots + r_n\mathbf{a}_n = \mathbf{0}$ ならば $\mathbf{x} = {}^t(r_1, \dots, r_n)$ について $A\mathbf{x} = r_1\mathbf{a}_1 + \dots + r_n\mathbf{a}_n = \mathbf{0}$ となる. つまり $A\mathbf{x} = \mathbf{0}$ であり, この方程式は自明な解しか持たないゆえ $\mathbf{x} = \mathbf{0}$ を得る. すなわち $r_1 = \dots = r_n = 0$ である. 以上より $\mathbf{a}_1, \dots, \mathbf{a}_n$ は線形独立である. \square

いまの議論を連立 1 次方程式の解法まで戻って詳しくみると, 線形独立性の判定だけではなく, 線形従属である場合にどのベクトルが他のベクトルの線形結合で書けるかも分かる. これを次の命題を通して見てみよう.

命題 17.3.2. (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ を行基本変形により $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$ に変形できるとする. $1 \leq n_1, n_2, \dots, n_\ell \leq n$ におよび $i = 1, \dots, n$ に対して次が成り立つ.

(1) $\sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} = \mathbf{0} \iff \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} = \mathbf{0}.$

(2) $\mathbf{a}_i = \sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k}$ と書ける $\iff \mathbf{b}_i = \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k}$ と書ける

Proof. (1): A を B に行基本変形できることから, $A' = [\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_\ell}]$ を $B' = [\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_\ell}]$ に行基本変形できる. また, $[A'|\mathbf{0}]$ を $[B'|\mathbf{0}]$ に行基本変形できる. ゆえに命題 4.2.1 より $A'\mathbf{x} = \mathbf{0}$ の解と $B'\mathbf{x} = \mathbf{0}$ の解は一致しており, $\mathbf{x} = {}^t(r_1, \dots, r_n)$ について

$$\sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} = \mathbf{0} \iff \mathbf{x} \text{ は } A'\mathbf{x} = \mathbf{0} \text{ の解} \iff \mathbf{x} \text{ は } B'\mathbf{x} = \mathbf{0} \text{ の解} \iff \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} = \mathbf{0}.$$

(2): $\mathbf{a}_i = \sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k}$ とすれば, $\sum_{k=1}^{\ell} r_k \mathbf{a}_{n_k} + (-1)\mathbf{a}_i = \mathbf{0}$ である. $\ell + 1$ 個のベクトルの組 $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_\ell}, \mathbf{a}_i$ について (1) を適用し, $\sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k} + (-1)\mathbf{b}_i = \mathbf{0}$ を得る. すなわち, $\mathbf{b}_i = \sum_{k=1}^{\ell} r_k \mathbf{b}_{n_k}$. 逆も同様に示される. \square

上の(1)は組 $\mathbf{a}_1, \dots, \mathbf{a}_k$ が線形独立(あるいは線形従属)であることと組 $\mathbf{b}_1, \dots, \mathbf{b}_k$ が線形独立(あるいは線形従属)であることの同値性を述べている. ベクトルの組が線形従属である場合, 命題 17.2.5 により, いずれかのベクトルが他のベクトルの線形結合で書ける. 上の命題を応用して, どのベクトルが他のベクトルの線形結合で書けるか調べてみよう:

例題 17.3.3. 次の列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_5$ が線形独立であるかどうか判定し, 線形従属の場合にはどのベクトルが他のベクトルの線形結合で書けるか答えよ.

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -1 \end{bmatrix}.$$

解答例: 与えられた列ベクトルを並べた行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_5]$ の簡約化を $B = [\mathbf{b}_1, \dots, \mathbf{b}_5]$ とすれば,

$$A = \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & -3 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix} \longrightarrow B = \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\mathbf{b}_1, \dots, \mathbf{b}_5$ が線形従属であることは成分を見れば明らかであり(線形従属性は $\text{rank } A = 3 \neq 5$ より方程式 $A\mathbf{x} = \mathbf{0}$ が自明でない解をもつことから分かる), ゆえに $\mathbf{a}_1, \dots, \mathbf{a}_5$ も線形従属である. 行列 B の成分を見れば $\mathbf{b}_3 = -\mathbf{b}_1 + \mathbf{b}_2$, $\mathbf{b}_5 = -2\mathbf{b}_1 + \mathbf{b}_2 - \mathbf{b}_4$ と書けることが分かる. したがって $\mathbf{a}_3 = -\mathbf{a}_1 + \mathbf{a}_2$, $\mathbf{a}_5 = -2\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_4$ である.

注意1: B の成分を見れば, $\mathbf{b}_2 = 2\mathbf{b}_1 + \mathbf{b}_4 + \mathbf{b}_5$ とも書けることが分かる. つまり, 他のベクトルで書けるものは $\mathbf{b}_3, \mathbf{b}_4$ に限るというわけではない. 上で $\mathbf{b}_3, \mathbf{b}_5$ を取り上げたのは, これ以外の主成分を含む列 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ が標準ベクトルであることから, $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ が線形独立であること, および $\mathbf{b}_3, \mathbf{b}_5$ が $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ の線形結合で書けることが直ちに分かるゆえである. 例えば, 組 $\mathbf{b}_1, \mathbf{b}_4, \mathbf{b}_5$ も線形独立であり, これ以外の $\mathbf{b}_2, \mathbf{b}_3$ を $\mathbf{b}_1, \mathbf{b}_4, \mathbf{b}_5$ の線形結合で表すこともできるが, それを示すのは標準ベクトルの組 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ に対して行うより骨が折れるであろう. より詳しい事情は次節の基底概念を通して説明される.

注意2: あらかじめ $\mathbf{a}_4, \mathbf{a}_5$ が線形独立であることが分かっており, これらの用いて他のベクトルを線形結合で表したい場合は, 列を並び替えて $\mathbf{a}_4, \mathbf{a}_5, \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ について同様の計算を行えばよい.

注意3: 行ベクトルについて同様の問題を考える場合は転置して列ベクトルの問題に変換し, 得られた答えを再び転置して行ベクトルに直せばよい.

(m, n) -行列 A による連立1次方程式 $A\mathbf{x} = \mathbf{0}$ の解が唯一解を持つかどうかは, 命題 6.2.1 により行列の階数を用いて判定できる. その条件は $\text{rank}[A|\mathbf{0}] = \text{rank } A = n$ である. 階数の定義から $\text{rank}[A|\mathbf{0}] = \text{rank } A$ であり, したがって条件 $\text{rank } A = n$ が $A\mathbf{x} = \mathbf{0}$ の解が唯一であるための同値条件である. 以上より次を得る:

系 17.3.4. \mathbb{R}^m の n 個の m 次列ベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ が線形独立であることと, $\text{rank}[\mathbf{a}_1, \dots, \mathbf{a}_n] = n$ であることは必要十分である.

命題 17.3.5. $m < n$ について, \mathbb{R}^m の n 個の列ベクトル $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である.

Proof. (m, n) -行列 $A = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ の階数は A の行の数 m 以下である(式 6.1.1). ゆえに $\text{rank } A \leq m < n$ であり, とくに $\text{rank } A \neq n$. 系 17.3.4 より $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である. \square

正方行列に現れるベクトルの組においては次が成り立ち, これらの条件を定理 6.2.2 につけ加えることができる.

定理 17.3.6. n 次正方行列 A について次は同値である.

(1) A は可逆である, (2) A の各列は線形独立である, (2) A の各行は線形独立である.

Proof. (1) \Leftrightarrow (2): 定理 6.2.2 および命題 17.3.1 から直ちに得られる:

$$A \text{ は可逆} \iff A\mathbf{x} = \mathbf{0} \text{ は唯一解をもつ} \iff A \text{ の各列は線形独立.}$$

(1) \Leftrightarrow (3): いま示した (1) と (2) の同値性および $|A| = |{}^tA|$ より得られる:

$$\begin{aligned} A \text{ は可逆} &\iff |A| \neq 0 \iff |{}^tA| \neq 0 \iff {}^tA \text{ は可逆} \\ &\iff {}^tA \text{ の各列は線形独立} \iff A \text{ の各行は線形独立.} \end{aligned}$$

□

18 基底

17.1 項で述べた要不要論を思い出そう. この議論の (1) に関係のある概念として線形独立性を前節で与えた. 本節では要不要論の (2) で論じたことと関係する, ベクトルの組による生成について述べる. これは, V の任意の元が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けるかどうかを定式化する概念である. これに線形独立性を合わせたものが基底であり, 一般の線形空間における基底は, ユークリッド空間における座標軸のような役割を果たすことになる.

18.1 ベクトルの組が生成する部分空間

V の任意の元が与えられた組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に分解できるかどうかはともかくとして, まず, $\mathbf{u}_1, \dots, \mathbf{u}_n$ に分解できるベクトルの範囲を表す記号を導入しよう

定義 18.1.1. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ において, $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの線形結合で書けるようなベクトルをすべて集めた V の部分集合を $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ と書く. すなわち, $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle := \{ \sum_{i=1}^n a_i \mathbf{u}_i \mid a_1, \dots, a_n \in \mathbb{R} \}$ である. 次の命題により, これは V の部分空間となる. $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ は, $\mathbf{u}_1, \dots, \mathbf{u}_n$ によって生成される部分空間と呼ばれる.

命題 18.1.2. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in V$ において, $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ は V の部分空間である.

Proof. 部分空間となるための条件 (i) および (iv) を確認すればよい. $\mathbf{0}$ は $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの線形結合で書けるゆえ $\mathbf{0} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ である. 次に, $\mathbf{x}, \mathbf{y} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ とすれば $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$, $\mathbf{y} = \sum_{i=1}^n b_i \mathbf{u}_i$ と書ける. このとき, 各 $r, s \in \mathbb{R}$ に対して $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^n (ra_i + sb_i) \mathbf{u}_i$ ゆえ $r\mathbf{x} + s\mathbf{y}$ も $\mathbf{u}_1, \dots, \mathbf{u}_n$ たちの線形結合で書ける. よって $r\mathbf{x} + s\mathbf{y} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ である. \square

次は, 要不要論の (1) で述べたことを, より一般的な状況に置き換えた主張である.

命題 18.1.3. V を線形空間とする. 各 $\mathbf{v}_1, \dots, \mathbf{v}_\ell \in V$ が組 $\mathbf{w}_1, \dots, \mathbf{w}_n \in V$ の線形結合で書けるならば, $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で書ける元は $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で書ける. すなわち,

$$\mathbf{v}_1, \dots, \mathbf{v}_\ell \in \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle \implies \langle \mathbf{v}_1, \dots, \mathbf{v}_\ell \rangle \subset \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle.$$

Proof. $W = \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle$ とおく. W は V の部分空間であり, 命題 16.1.6 より W は線形結合について閉じている. 仮定より $\mathbf{v}_1, \dots, \mathbf{v}_\ell \in W$ であるから, $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で書ける元は W に含まれる. ゆえに $\langle \mathbf{v}_1, \dots, \mathbf{v}_\ell \rangle \subset \langle \mathbf{w}_1, \dots, \mathbf{w}_n \rangle$. \square

上の証明では, $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で書いたときに現れる係数と $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で書いたときに現れる各係数の関係については論じなかった. これらの係数の関係は行列の積演算を通して得られる (25.1 項を見よ).

18.2 基底の定義と例

次で定める基底とは, 線形空間 V の各元を線形結合で表すときに過不足なく必要になるベクトルの組のことである.

定義 18.2.1. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle = V$ を満たすとき, $\mathbf{u}_1, \dots, \mathbf{u}_n$ は V を生成するという. また, V を生成するような線形独立な組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を V の基底 (basis) あるいは基という.

例 18.2.2. (1) \mathbb{R}^n の標準ベクトルの組 $\mathbf{e}_1, \dots, \mathbf{e}_n$ は \mathbb{R}^n を生成し, かつ線形独立である (例 17.1.2 および例 17.2.2). ゆえに \mathbb{R}^n の基底である.

(2) (i, j) -成分が 1 でそれ以外の成分がすべて 0 の (m, n) -行列を E_{ij} と書けば (ただし $1 \leq i \leq m$, $1 \leq j \leq n$), mn 個の行列の組 E_{ij} ($i = 1, \dots, m$, $j = 1, \dots, n$) は線形独立かつ $M_{m,n}(\mathbb{R})$ を生成する. ゆえにこれらは $M_{m,n}(\mathbb{R})$ の基底である.

(3) $\mathbb{R}[x]_n$ における $n+1$ 個の組 $x^n, x^{n-1}, \dots, x, 1$ は $\mathbb{R}[x]_n$ を生成し、かつ線形独立である (例 17.2.2(3)).
ゆえに $\mathbb{R}[x]_n$ の基底である.

(4) 約束として、自明な空間 $\{0\}$ は 0 個のベクトルの組からなる基底をもつとする

ユークリッド空間 \mathbb{R}^n には座標が定められていた. 基底とは、一般の線形空間において座標の代わりとなる概念である. V の各元 v について、基底による v の線形結合表示に現れる係数を v の位置データと見なすことにより、 v の座標表示を考えることができる. とくに \mathbb{R}^n の各元を標準基底 e_1, \dots, e_n の線形結合によって書いた際に現れる各係数は、その位置を示す座標の各成分に一致している. \mathbb{R}^n の通常の座標において、その表示が異なれば違う位置を示していたように、基底による線形結合の各係数に現れる実数の組が異なれば、線形結合が表す位置も当然異なっているべきであろう. このことは次の命題が保証している.

命題 18.2.3. $u_1, \dots, u_n \in V$ が線形独立であるとし、 $v \in V$ がこれらの線形結合で書けるとすれば、その表し方は一通りしかない. すなわち、 $v = \sum_{i=1}^n a_i u_i = \sum_{i=1}^n b_i u_i$ ならば $a_i = b_i$ ($i = 1, \dots, n$) である.

Proof. $\sum_{i=1}^n a_i u_i = \sum_{i=1}^n b_i u_i$ を移項すると線形関係 $\sum_{i=1}^n (a_i - b_i) u_i = 0$ を得る. u_1, \dots, u_n の線形独立性より $a_1 - b_1 = \dots = a_n - b_n = 0$. つまり $a_i = b_i$ ($i = 1, \dots, n$) である. \square

\mathbb{R}^n の座標と一般の基底の違いを述べておこう. \mathbb{R}^n の座標軸がそれぞれ直交するのに対して、基底を構成する各ベクトルは必ず直交するわけではない. そもそも、一般の線形空間においては直交なる概念が定まるとは限らないといえる. また、座標軸は \mathbb{R}^n の中で既に定まっているのに対して、次の例にあるように線形空間の基底の取り方は無数にある. 言い換えると、線形空間においては座標軸に似た概念 (すなわち基底) を多様に定めることができる. このため、基底による表示を用いた議論を行う場合、はじめにどんな基底を考えているか宣言する必要がある. 以後、本論における多くの命題もこのような形で述べることになる.

例 18.2.4. $A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ は可逆ゆえ命題 17.3.1 より列ベクトルの組 $u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ は線形独立である. \mathbb{R}^2 の任意のベクトル v は標準ベクトル e_1, e_2 の線形結合でかけ、また e_1, e_2 は u_1, u_2 の線形結合で書ける. 実際、 $e_1 = u_1 - u_2$, $e_2 = u_2$ である. このことは、 v が u_1, u_2 の線形結合で書けることを意味する (命題 18.1.3). すなわち u_1, u_2 は \mathbb{R}^2 を生成し、したがって \mathbb{R}^2 の基底である.

いまの例の一般化として定理 18.2.6 が得られる.

補題 18.2.5. ベクトルの組 $u_1, \dots, u_n \in V$ が線形独立であるとし、さらに $v \in V$ とすれば次が成り立つ.

(1) v が u_1, \dots, u_n の線形結合で表せないならば、ベクトルの組 u_1, \dots, u_n, v は線形独立である.

(2) ベクトルの組 u_1, \dots, u_n, v は線形従属ならば、 v は u_1, \dots, u_n の線形結合で表せる.

Proof. (2) は (1) の対偶ゆえ (1) のみ示せばよい. 線形独立性を示すために線形関係 $\sum_{i=1}^n a_i u_i + a_{n+1} v = 0$ を仮定しよう. このとき $a_{n+1} = 0$ でなければならない. 何故なら、もし $a_{n+1} \neq 0$ ならば移項により $v = \sum_{i=1}^n \frac{-a_i}{a_{n+1}} u_i$ となり、これは v が u_1, \dots, u_n の線形結合で表せないことに反する. ゆえに $a_{n+1} = 0$ であり、 $\sum_{i=1}^n a_i u_i = 0$ を得る. u_1, \dots, u_n の線形独立性より $a_1 = \dots = a_n = 0$ である. \square

定理 18.2.6. \mathbb{R}^n における n 個のベクトルの組 u_1, \dots, u_n において、これらが線形独立であることと \mathbb{R}^n の基底であることは同値である.

Proof. 列ベクトルに対して示そう. 行ベクトルの場合は転置をとって列ベクトルの場合に帰着させればよい. 列ベクトル u_1, \dots, u_n が線形独立であるとし、これが \mathbb{R}^n を生成することを背理法により示す. 仮に u_1, \dots, u_n の線形結合で書けないベクトル $v \in \mathbb{R}^n$ があるとすれば、 $n+1$ 個の組 $u_1, \dots, u_n, v \in \mathbb{R}^n$ は補題 18.2.5(1) より線形独立となる. しかしこのことは命題 17.3.5 に矛盾する. ゆえに、各 $v \in \mathbb{R}^n$ は u_1, \dots, u_n の線形結合で書けねばならない. \square

18.3 基底の探し方

有限個のベクトルの組で生成される線形空間における基底の探し方を検討しよう. 次の命題の証明では抽象的な基底の構成法が述べられている.

命題 18.3.1. V が零でないベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_m$ によって生成されているとすれば, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中からいくつかを取りだし V の基底とすることができる. とくに, 有限個のベクトルの組で生成される線形空間は基底を持つ.

Proof. V を生成するまで, 線形独立性が満たされるよう元を一つずつ加えていけばよい. これは次のような手続きによってなされる. まず \mathbf{u}_{n_1} として \mathbf{u}_1 を取る. この \mathbf{u}_{n_1} が V を生成するならば, \mathbf{u}_{n_1} は V の基底である. そうでない場合は $\mathbf{u}_2, \dots, \mathbf{u}_m$ のうち \mathbf{u}_{n_1} のスカラー倍で表せないものがある. 何故なら, もし $\mathbf{u}_1, \dots, \mathbf{u}_m \in \langle \mathbf{u}_{n_1} \rangle$ とすれば命題 18.1.3 より $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \subset \langle \mathbf{u}_{n_1} \rangle$ であり, V は \mathbf{u}_{n_1} によって生成されてしまう. そこで, \mathbf{u}_{n_1} のスカラー倍で表せないベクトルを仮に \mathbf{u}_{n_2} とすれば補題 18.2.5(1) より $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ は線形独立である. $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ が V を生成するならばこれは V の基底となる. そうでない場合は $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ を除いたベクトルのうちいずれかは $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ の線形結合で書けない. 何故なら, もし $\mathbf{u}_1, \dots, \mathbf{u}_m$ すべてが $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ の線形結合で書けるとすると, 命題 18.1.3 より $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \subset \langle \mathbf{u}_{n_1}, \mathbf{u}_{n_2} \rangle$ であり, V は $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ によって生成されてしまう. ゆえに $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}$ の線形結合で表せないベクトルがあり, これを \mathbf{u}_{n_3} とする. この作業を順次繰り返していくと, いずれ線形独立な組 $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}, \mathbf{u}_{n_3}, \dots, \mathbf{u}_{n_k}$ (ただし $k \leq m$) が V を生成することになる. 実際, 線形結合で書けない元を新たに付け加える操作は, 最大でも $\mathbf{u}_1, \dots, \mathbf{u}_m$ をすべて取りつくすことになる m 回までしか行えない. 以上の手続きにより, V の基底 $\mathbf{u}_{n_1}, \mathbf{u}_{n_2}, \mathbf{u}_{n_3}, \dots, \mathbf{u}_{n_k}$ が得られる. \square

命題 18.3.1 の証明における手順を改善すれば 次のような基底の構成もできる. これは, 基底の一部としたいベクトルがあらかじめ決まっているときに有効な手段となる. また, 例題 17.3.3 の注意 2 とも関連する話題である.

命題 18.3.2. 有限個のベクトルの組で生成される線形空間 V において, 線形独立な組 $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$ が与えられているとき, これらに新たなベクトルを付け加えて V の基底とすることができる.

Proof. V は有限個のベクトルで生成されていることから, $V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ と表せる. このとき,

$$V = \langle \mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m \rangle$$

でもあることに注意して, $n + m$ 個の組 $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m$ に対して命題 18.3.1 の証明を適用しよう. このとき先の証明において始めに \mathbf{u}_1 を選ぶところで, 代わりに $\mathbf{v}_1, \dots, \mathbf{v}_m$ を選び取ってしまうとよい. すなわち, $\mathbf{u}_{n_i} := \mathbf{v}_i$ ($i = 1, \dots, m$) とたうえて, 命題 18.3.1 の証明で述べた手続きを進めれば求める基底が得られる. \square

これまでに挙げてきた \mathbb{R}^n の部分空間については, 行列の簡約化の理論を通して基底を見つけることができる. これを次の例題を通して説明しよう.

例題 18.3.3. 次で定める $\mathbf{a}_1, \dots, \mathbf{a}_5$ について (例題 17.3.3 と同じもの), 次の問いに答えよ.

$$\mathbf{a}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 2 \\ 1 \end{bmatrix}, \mathbf{a}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \mathbf{a}_3 = \begin{bmatrix} -1 \\ 1 \\ 1 \\ -1 \\ 0 \end{bmatrix}, \mathbf{a}_4 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{a}_5 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ -3 \\ -1 \end{bmatrix}.$$

(1) $\langle \mathbf{a}_1, \dots, \mathbf{a}_5 \rangle$ の基底を求めよ.

解答例: $A = [\mathbf{a}_1, \dots, \mathbf{a}_5]$ の簡約化を $B = [\mathbf{b}_1, \dots, \mathbf{b}_5]$ とする (簡約化は例題 17.3.3 で行った). このとき $\langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4 \rangle = \langle \mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \mathbf{b}_4, \mathbf{b}_5 \rangle$ である. よって命題 17.3.2 より $\langle \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4 \rangle = \langle \mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{a}_4, \mathbf{a}_5 \rangle$ となる. また, 組 $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_4$ は \mathbb{R}^m の標準ベクトルゆえ線形独立である. ゆえに命題 17.3.2 より組 $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_4$ も線形独立であり, これらは $\langle \mathbf{a}_1, \dots, \mathbf{a}_5 \rangle$ の基底となる.

(2) 方程式 $A\mathbf{x} = \mathbf{0}$ の解空間 W_A の基底を求めよ.

解答例: $[A|\mathbf{0}]$ の簡約化は $[B|\mathbf{0}]$ であり, W_A の外延的表示を得るために次の方程式 $B\mathbf{x} = \mathbf{0}$ を解く:

$$\begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

掃き出し法により W_A は次のように表される:

$$W_A = \{c_1\mathbf{u}_1 + c_2\mathbf{u}_2 \mid c_1, c_2 \in \mathbb{R}\}. \quad (\text{ただし, } \mathbf{u}_1 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \mathbf{u}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ 1 \\ 1 \end{bmatrix})$$

ゆえに $W_A = \langle \mathbf{u}_1, \mathbf{u}_2 \rangle$ である. $\mathbf{u}_1, \mathbf{u}_2$ が線形独立性は, 主成分のある列に対応する行成分の情報を落とすことで理解できる. いまの例では主成分に対応する 1, 2, 4 行を目隠して

$$\mathbf{u}_1 = \begin{bmatrix} * \\ * \\ 1 \\ * \\ 0 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} * \\ * \\ 0 \\ * \\ 1 \end{bmatrix}$$

と見ると, これらが線形独立であることは標準ベクトルがそうであることと同程度に明らかであろう. 以上より $\mathbf{u}_1, \mathbf{u}_2$ 基底である.

上の例題で行った議論を一般的に述べると次の命題になる.

命題 18.3.4. (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ について次が成り立つ.

(1) \mathbb{R}^m の部分空間 $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ は $\text{rank } A$ 個のベクトルからなる基底を持つ.

(2) $A\mathbf{x} = \mathbf{0}$ の解空間 $W_A \subset \mathbb{R}^n$ は $n - \text{rank } A$ 個のベクトルからなる基底を持つ.

Proof. $k = \text{rank } A$ とおき, A の簡約化を $B = [\mathbf{b}_1, \dots, \mathbf{b}_n]$, B の各列のうち主成分を持つ列を $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ とし, 主成分を持たない列を $\mathbf{b}_{r_1}, \dots, \mathbf{b}_{r_{n-k}}$ とする.

(1): ベクトルの組 $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ は互いに異なる標準ベクトルからなるゆえ線形独立である. また, 簡約化の形から, B の各列は $\mathbf{b}_{n_1}, \dots, \mathbf{b}_{n_k}$ の線形結合で書ける. 命題 17.3.2 より組 $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ は線形独立であり, A の各列は $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ の線形結合で書ける. すなわち, $\mathbf{a}_1, \dots, \mathbf{a}_n \in \langle \mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k} \rangle$. これと命題 18.1.3 を合わせて $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle \subset \langle \mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k} \rangle$ を得る. 以上より $\mathbf{a}_{n_1}, \dots, \mathbf{a}_{n_k}$ は $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ の基底となる.

(2): 連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ を掃き出し法によって求めると, 任意定数の個数は B における主成分のない列の数 $n - k$ であるから, その一般解は $\mathbf{x} = \sum_{j=1}^{n-k} c_j \mathbf{u}_j$ と書ける. つまり $W_A = \langle \mathbf{u}_1, \dots, \mathbf{u}_{n-k} \rangle$ である. ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ が線形独立であることを示すために, 各 $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ の r_j 成分 ($j = 1, \dots, n - k$) に注目しよう. B の第 r_j 列 \mathbf{b}_{r_j} は主成分を含まない列であったことから, 掃き出し法で求めた一般解において r_j 成分は任意定数としていた. このことは, \mathbf{u}_j の r_j 成分は 1 であり, $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ のうち \mathbf{u}_j を除いた残りのベクトルの r_j 成分は 0 になっていることを意味する. (上の例題では $\mathbf{b}_{r_1} = \mathbf{b}_3$, $\mathbf{b}_{r_2} = \mathbf{b}_5$ となる. 確かに \mathbf{u}_1 の $r_1 = 3$ 成分は 1, $r_2 = 5$ 成分は 0 であり, \mathbf{u}_2 の $r_1 = 3$ 成分は 0, $r_2 = 5$ 成分は 1 となっている). ゆえに線形関係 $\sum_{j=1}^{n-k} c_j \mathbf{u}_j = \mathbf{0}$ を与えると, 各 $j = 1, \dots, n - k$ について左辺の第 r_j 成分は c_j となる. これが右辺の r_j 成分である 0 に等しいことから $c_j = 0$ ($j = 1, \dots, n - k$) を得る. すなわち, $\mathbf{u}_1, \dots, \mathbf{u}_{n-k}$ における線形関係は自明なものに限り, これらは線形独立である. \square

本項や 17.3.1 項では、ユークリッド空間 \mathbb{R}^n の部分空間における線形関係について例題を通して学んだ。一般の線形空間 V におけるベクトルの組の独立性の判定や部分空間 $W \subset V$ の基底選びは、 V に関する命題を \mathbb{R}^n に関する命題に翻訳したうえで行うことになる。実は、独立性の判定や基底選びに限らず V の分析は \mathbb{R}^n の分析を通してなされる。この翻訳の基本理念は、線形空間 V の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する条件を \mathbb{R}^n の標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ の条件で述べることにある。また、翻訳作業においては、 V のどの元が \mathbb{R}^n のどの元に対応しているのかを表示するために写像概念が用いられる。そこで次節からしばらくの間、写像に関する概念を整理することにしよう。

18.4 一般の基底

有限個のベクトルでは生成されない線形空間もある。そのような空間における基底概念について少しだけ補足しておこう。

例 18.4.1. 多項式環 $\mathbb{R}[x]$ において、有限個の多項式の組 f_1, \dots, f_n が $\mathbb{R}[x]$ を生成することはない。なぜなら、各多項式 f_1, \dots, f_n の中で最も高い次数を m とすれば、 $m+1$ 次多項式を f_1, \dots, f_n の線形結合で表すことはできないからである。

ベクトルの組による生成や基底概念は、次のようにして無限集合の場合にも拡張される。

定義 18.4.2. 線形空間 V の部分集合 $A \subset V$ において、 A の元による線形結合で書けるベクトルをすべて集めた集合を $\langle A \rangle$ と書く。 $\langle A \rangle = V$ となるとき、 A は V を生成するという。また、 A が線形独立であるとは、 A の中から有限個取りだしたベクトルの組が必ず線形独立になることをいう。さらに、 A が線形独立かつ V を生成するとき、これを V の基底あるいはハメル基底 (**Hamel basis**) という。

命題 18.4.3. 線形空間 V および空でない部分集合 $A \subset V$ に対して、 $\langle A \rangle$ は V の部分空間である。

Proof. A は空集合でないゆえ $\mathbf{a} \in A$ が取れる。このとき $\mathbf{0}_V = 0\mathbf{a}$ は A の元の線形結合で書けている。ゆえに $\mathbf{0}_V \in \langle A \rangle$ 。一方、 $\mathbf{x}, \mathbf{y} \in \langle A \rangle$ とすれば、 \mathbf{x}, \mathbf{y} は A の元の線形結合で書ける。ここで、 \mathbf{x}, \mathbf{y} それぞれの線形結合に現れる A の元の組は異なるかもしれないが、一部の係数を 0 とすることで共通の組 $\mathbf{a}_1, \dots, \mathbf{a}_n$ を用いて $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{a}_i$, $\mathbf{y} = \sum_{i=1}^n y_i \mathbf{a}_i$ と書くことができる。このとき、各 $r, s \in \mathbb{R}$ について $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^n (rx_i + sy_i) \mathbf{a}_i$ であるから $r\mathbf{x} + s\mathbf{y}$ は A の元の線形結合で書ける。ゆえに $r\mathbf{x} + s\mathbf{y} \in \langle A \rangle$ 。以上より $\langle A \rangle$ は部分空間である。□

とくに、ベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および集合 $A = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ について、

- $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle = \langle \{\mathbf{u}_1, \dots, \mathbf{u}_n\} \rangle$,
- $\mathbf{u}_1, \dots, \mathbf{u}_n$ が V を生成することと A が V を生成することは同値である,
- $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であることと A が線形独立であることは同値である,
- $\mathbf{u}_1, \dots, \mathbf{u}_n$ が V の基底であることと A が V の基底であることは同値である。

例 18.4.4. 多項式環 $\mathbb{R}[x]$ の部分集合 $A = \{x^n \mid n \in \mathbb{Z}_{\geq 0}\}$ は線形独立である。ここで、集合 $\mathbb{Z}_{\geq 0}$ は 0 以上の整数全体を表し、 $x^0 = \mathbf{1}$ (定数関数) とする。また、 $\mathbb{R}[x]$ の各元は A の元の線形結合で書ける。ゆえに A は $\mathbb{R}[x]$ の基底である。

いま、線形空間の任意の部分集合について、それが生成する部分空間を定めた。とくに部分集合として部分空間を取れば次を得る。

命題 18.4.5. (1) $W \subset V$ が V の部分空間であるとき、 $W = \langle W \rangle$ 。

(2) $A \subset V$ が V の部分集合であるとき、 $\langle \langle A \rangle \rangle = \langle A \rangle$ 。

Proof. (1): $W \subset \langle W \rangle$ は明らかゆえ $\langle W \rangle \subset W$ を示す. 各 $v \in \langle W \rangle$ は W の元による線形結合で書ける. また, W は線形結合で閉じている (命題 16.1.6) ゆえ $v \in W$.

(2): $W := \langle A \rangle$ について (1) を適用すればよい. □

$\mathbb{R}^{\mathbb{N}}$ のように基底を書き下すことが難しい空間もある.

例 18.4.6. 数列空間 $\mathbb{R}^{\mathbb{N}}$ において, 第 n 項が 1 でそれ以外の項がすべて 0 となる数列を e_n と書けば, $A = \{e_n \mid n \in \mathbb{N}\}$ は線形独立である. しかし, A は $\mathbb{R}^{\mathbb{N}}$ を生成しない. 実際, すべての項が 1 となる数列 $x = (1, 1, \dots)$ を A の元から有限個を取りだした組による線形結合で書くことはできない.

そもそも線形空間に必ず基底が存在するかどうかということ自体が明らかではない. ツォルンの補題³⁵と呼ばれる集合論における原理を適用することで, 次の定理が証明できることが知られている. 詳しい証明は集合論の入門的参考書を参照せよ.

定理 18.4.7. いかなるベクトル空間 V も基底 $A \subset V$ を持つ.

上の定理は基底の存在を超越的に示すものであり, 基底の形が明示的に書けることは意味しない. 基底の表示が与えられなければ, 本論全体を通しての主題でもある基底を用いた分析は行えない. このような線形空間を調べる際は, 線形空間に位相³⁶と呼ばれる構造を導入し, 極限操作を手掛かりに分析することになる.

³⁵多くの理工系学部の数学科では, 2 年次の集合論の講義で学ぶことになっている.

³⁶点列の収束発散を議論できるような枠組みを与える数学的構造を位相 (topology) という.

発展 (最小の部分空間).

部分集合 $A \subset V$ が生成する部分空間 $\langle A \rangle$ を次のように定義する流儀もある.

定義 18.4.8. 線形空間 V の部分集合 $A \subset V$ に対して, A を含む最小の V の部分空間を $\langle A \rangle$ とする.

ここでいう最小とは, 包含関係 \subset に関して最も小さいということである. 上の定義の利点は少ない言葉で済むこと, そして空集合 \emptyset は自明な部分空間 $\{0\}$ を生成することになり, \emptyset を $\{0\}$ の基底であると約束する手間が省けることにある. 一方, 欠点は, A を含む最小の部分空間はそもそも存在するかという疑問にあらかじめ答えておかねばならないことである. これは集合の共通部分をとる演算 \cap を用いて正当化される. この点について解説しよう.

V の部分集合たちを集めた集合 \mathcal{W} が与えられているとする (すなわち \mathcal{W} は集合の集合であり, このような集合は集合族と呼ばれる). このとき, 各 $W \in \mathcal{W}$ のいずれにも含まれている元をすべて集めた V の部分集合を $\bigcap \mathcal{W}$ と書く. すなわち,

$$\bigcup \mathcal{W} := \{v \in V \mid \text{各 } W \in \mathcal{W} \text{ について } v \in W\}$$

$$\left(\text{つまり, } x \in \bigcap \mathcal{W} \iff \text{各 } W \in \mathcal{W} \text{ について } x \in W\right).$$

さて, A を含む V の部分空間たち全体からなる集合族を \mathcal{W} としよう. \mathcal{W} は空集合ではない. 何故なら, V 自身は A を含む V の部分空間であるから $V \in \mathcal{W}$ である. このとき, $U := \bigcap \mathcal{W}$ と定めれば, U は A を含む最小の部分空間である.

Proof. 示すべきことは (1) $A \subset U$, および (2) U が部分空間であること, (3) U が A を含む部分空間の中で最小であることの三つである.

(1): $A \subset U$ を示すために任意に $a \in A$ を取る. $a \in U$ をいうには, 各 $W \in \mathcal{W}$ について $a \in W$ を示せばよい. 各 $W \in \mathcal{W}$ について W は A を含む部分空間 (つまり $A \subset W$) であった. $a \in A$ および $A \subset W$ ゆえ $a \in W$ である. 以上より「 $a \in A \implies a \in U$ 」が示された. つまり $A \subset U$.

(2): 各 $W \in \mathcal{W}$ は 0_V を含むゆえ $0_V \in U$ である. 次に $x, y \in U$ とすれば $rx + sy \in U$ となることを示そう. そのためには各 $W \in \mathcal{W}$ について $rx + sy \in W$ を示せばよい. $x, y \in U$ より各 $W \in \mathcal{W}$ において $x, y \in W$ であり, W が部分空間であることから $rx + sy \in W$ を得る. したがって $rx + sy \in U$. 部分空間となるための条件 (i) と (iv) が示されたゆえ, U は V の部分空間である.

(3): A を含む任意の V の部分空間 H について, $U \subset H$ となることを示せばよい. A を含む V の部分空間 H を勝手に取れば $H \in \mathcal{W}$ である. $U = \bigcap \mathcal{W}$ の定義により, 各 $W \in \mathcal{W}$ について $U \subset W$ であるから, とくに $U \subset H$. □

この U が定義 18.4.2 における $\langle A \rangle$ と一致することは次のように示される.

Proof. ($\langle A \rangle \subset U$): $A \subset U$ より $\langle A \rangle \subset \langle U \rangle = U$ (命題 18.4.5(1)).

($U \subset \langle A \rangle$): $\langle A \rangle$ は A を含む V の部分空間である (つまり $\langle A \rangle \in \mathcal{W}$). U の最小性より $U \subset \langle A \rangle$. □

19 写像概念の基礎

写像に関する概念のいくつかを述べる. 本節では概念をひたすら提示することに終始するゆえ, 読者はやや退屈に感じるかもしれない. そこで, 予告の意味を込めて, これらの概念が線形写像の性質とどう結びつくかを各項末で述べた.

19.1 像と逆像

集合 X から集合 Y への写像 $f: X \rightarrow Y$ における X のことを定義域 (domain) と呼ぶのであった. 定義域は始域 (source) とも呼ばれる. また, この f における Y のことを終域 (target) と呼ぶ. 定義域と始域の概念を更に細かく区別して用いる文献もある. 他方で, 定義域に対応する語句として値域 (range) を使う文献もあるが, これを終域の意味で使うのであれば高校数学における値域とは意味が異なる. こうした混乱を避けるため, 本論では始域および値域という呼称を控えよう. 高校数学において f の値域と呼んでいた概念を, 本論では像と呼ぶ:

定義 19.1.1. 写像 $f: X \rightarrow Y$ および定義域の部分集合 $A \subset X$ に対して, A の元を f に代入した値をすべて集めた Y の部分集合を A の f による像 (image) と呼び, これを $f(A)$ と書く. すなわち, 外延的記述をすれば

$$f(A) := \{f(x) \mid x \in A\}.$$

とくに $f(X)$ のことを f の像と呼ぶ.

練習 19.1.2. 次で与えられる関数 f の像を求めよ.

(1) $f: \mathbb{R} \rightarrow \mathbb{R}, f(x) = x^2$. 解答: $f(\mathbb{R}) = [0, \infty)$.

(2) $f: [0, 2] \rightarrow \mathbb{R}, f(x) = x^2$. 解答: $f([0, 2]) = [0, 4]$.

(3) $f: [-2, 2] \rightarrow [0, 4], f(x) = x^2$. 解答: $f([-2, 2]) = [0, 4]$.

(4) $f: [0, 2] \rightarrow [0, 4], f(x) = x^2$. 解答: $f([0, 2]) = [0, 4]$.

定義 19.1.3. 写像 $f: X \rightarrow Y$ および終域の部分集合 $B \subset Y$ に対して, f に代入すると B の元になるような元をすべて集めた X の部分集合を B の f による逆像 (inverse image) と呼び, これを $f^{-1}(B)$ と表す. すなわち, 内包的記述をすれば

$$f^{-1}(B) := \{x \in X \mid f(x) \in B\}.$$

また, 点 $b \in Y$ に対して, 一点集合 $\{b\}$ の逆像 $f^{-1}(\{b\})$ のことを中括弧を略して $f^{-1}(b)$ と書く.

一点の逆像の記号 $f^{-1}(b)$ は, f^{-1} なる写像に b を代入した値のことではない. $f^{-1}(b)$ は X の元ではなく, X の部分集合である. また, 一点集合になるとは限らず, 複数の点を含むこともあれば空集合になる場合もある.

例 19.1.4. 次で定められる関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ および $b \in \mathbb{R}$ について $f^{-1}(b)$ を求めよ.

(1) $f(x) = x^2, b = 2, -2$. 解答: $f^{-1}(2) = \{\sqrt{2}, -\sqrt{2}\}, f^{-1}(-2) = \emptyset$.

(2) $f(x) = \sin x, b = 0$. 解答: $f^{-1}(0) = \{n\pi \mid n \in \mathbb{Z}\}$.

(3) $f(x) = 2^x, b = 1$. 解答: $f^{-1}(1) = \{0\}$.

写像 $f, g: X \rightarrow Y$ が等しい ($f = g$) とは、定義域のいかなる元を代入しても一致すること、すなわち

$$\text{各 } x \in X \text{ について } f(x) = g(x) \quad (19.1.1)$$

が成り立つことに他ならない。上の条件 (19.1.1) が成立するとき、恒等的に f と g は等しいと言い、これを記号で $f(x) \equiv g(x)$ と書く。なお、誤解の恐れがない通常の多くの場合において、 $f(x) \equiv g(x)$ のことを $f(x) = g(x)$ と書く。また、 f が X の各元をある一点 $b \in Y$ に対応させる定置写像 (定数関数) であるとき (つまり $f(x) \equiv b$ であるとき)、これを $f(x) = b$ あるいは $f = b$ と書くことがある。後者の表記を認めれば、定数関数 $\mathbf{1}$ (0 次多項式) を $\mathbf{1}$ と書いてよいことになる。ただし、集合の元と写像を等号で結ぶこのような使い方には誤解が生じる恐れもあり、気をつける必要がある。線形代数の文脈においては、すべての元を零元に対応させる写像 $f: U \rightarrow V$ のことを $f = \mathbf{0}_V$ と書く。

定置写像 $f: X \rightarrow Y$ ($f = b$) において、 $f(X) = \{b\}$ および $f^{-1}(b) = X$ である。

以降で学ぶこと

写像の像と逆像は、線形代数の枠組みにおいても詳しく調べられる。その理由は線形写像による部分空間の像や逆像が再び部分空間となることにある。特に、線形写像 f の像には特別な記号が割り当てられ、 $\text{Im } f$ と書かれる。また、 f の原点による逆像 $f^{-1}(\mathbf{0})$ にも特別な記号 $\text{Ker } f$ が用いられる。行列 A において $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ を $T_A(\mathbf{x}) := A\mathbf{x}$ と定めれば A の階数が $\text{Im } T_A$ の次元に一致すること、および同次形連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ の解空間が $\text{Ker } T_A$ で表されること (例 22.1.3(2)) を通して次元公式が説明される (例 23.2.5)。

19.2 全射と単射

この項で述べる全単射性 (1 対 1 の対応) なる概念は、既に、行列式の性質の証明の際にも現れていた。そこで用いられていたように、証明の細部における技術面でこの概念が有用であることは理解されよう。しかし、数学のもっと根本的な部分において全単射性の概念は現れる。それは、異なる数学的対象³⁷ を対応付けて同等とみなす立場を記述する際に用いられる。例えば、2 節では \mathbb{R}^2 上の線形写像全体と 2 次正方行列全体 $M_2(\mathbb{R})$ が同一視できること、したがって線形写像の分析と行列計算の分析が同等であることを見た。何をもって同等とみなすべきか、それは考えている立場や価値観によって変わってくるだろう。しかしながらいずれにせよ、何かを同等とみなすとき、そこには全単射なる概念が自然に現れることになる。

定義 19.2.1. $f: X \rightarrow Y$ を写像とする。 x が重複なく X を動けば $f(x)$ も重複なく Y 上を動くとき、 f を単射 (injection) という。 x が X 全体を動けば $f(x)$ も Y 全体を動くとき、 f を全射 (surjection) あるいは上への写像 (onto map) という。単射性と全射性はそれぞれ次の条件に書き下すことができる:

- 単射性: 各 $x_1, x_2 \in X$ について、 $x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.
対偶をとれば次のようになる: 各 $x_1, x_2 \in X$ について、 $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$.
これは、各 $y \in f(X)$ について $f^{-1}(y)$ がちょうど 1 点からなる集合になることである。
つまり、各 $y \in Y$ について、 $f(x) = y$ をみたく $x \in X$ は高々一つしかない³⁸.
- 全射性: 各 $y \in Y$ に対して、 f に代入すると y になる元 $x \in X$ が存在する。
像を用いれば次のようになる: $f(X) = Y$ となること。

更に、全射かつ単射な写像を全単射 (あるいは 1 対 1 の対応, bijection) であるという。

³⁷ここでいう数学的対象とは、図形や空間であったり、あるいは何らかの代数構造を持つ数空間であったりと多岐にわたる。いずれにせよ、それらの多くは集合を用いて記述されるものである。

³⁸高々一つということは、一つもない可能性もある (つまり 0 個) ということである

射という字を用いることから単射性・全射性を弓矢に例えて説明すれば次のようになる. X を矢の集合, Y を的の位置 (標的) を表す集合とし, f を弓であると考え. いま, 弓 f を一つ固定し, 写像 $f: X \rightarrow Y$ とは, 弓 f を用いて矢 $x \in X$ を放つと $f(x) \in Y$ なる場所に矢が刺さると考える. このとき, 単射とは単発で当たるということである. 各々の的の位置に矢が刺さるとしても, 刺さる矢の数はせいぜい 1 本であり (矢が当たらないこともあり得る), 二本以上の矢が同じ場所に刺さることはない. 言い換えれば, もし矢 a, b がともに同じ位置に刺さった (つまり $f(a) = f(b)$) ならば, その位置に当たる矢の数は 1 本以下であるから, a と b は同一の矢ということになる. 全射とは的の全ての位置に矢が当たること, すなわち, どのような的の位置 $y \in Y$ においても, y に刺さる矢 $x \in X$ があること (つまり $f(x) = y$) を意味する. このとき, y に刺さる矢の数は 1 本以上であれば何本でも構わない.

例 19.2.2. 集合 X において X の元をまったく動かさない写像, すなわち $f(x) := x$ で定める写像 $f: X \rightarrow X$ を恒等写像 (identity map) といい, これを id_X と書く. id_X は全単射である.

写像が単射かどうか, あるいは全射かどうかは, f の定義式だけではなく, 定義域や終域に依存して決まるものである:

例 19.2.3. 練習 19.1.2 における写像の単射性および全射性は次のようになる.

(1) $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) = x^2$ と定めれば, これは単射でも全射でもない.

Proof. $x_1 = 1, x_2 = -1$ と置けば, これらは共に定義域の元であり, $x_1 \neq x_2$ である. ところが $f(x_1) = f(x_2) = 1$ ゆえ, 異なる元が f で同じ元に写されている. ゆえに f は単射ではない. また, f の像は $[0, \infty)$ であり, これは終域 \mathbb{R} に一致しない. ゆえに f は全射ではない. 実際, $y = -1$ は終域の元であるが, $f(x) = y$ を満たす定義域 \mathbb{R} の元 x は存在しない. \square

(2) $f: [0, 2] \rightarrow \mathbb{R}$ を $f(x) = x^2$ と定めれば, これは単射であり, かつ全射でない.

Proof. 単射性を示すために $a, b \in [0, 2]$ とし, $f(a) = f(b)$ と仮定しよう. このとき $a^2 = b^2$ である. これを移項して因数分解し $(a-b)(a+b) = 0$ を得る. ゆえに $a-b=0$ または $a+b=0$. $a-b=0$ ならば $a=b$ である. $a+b=0$ の場合は, $a=-b$ を得る. このとき, もし $a > 0$ とすれば $b < 0$ となり, これは $b \in [0, 2]$ (とくに $b \geq 0$) であることに矛盾する. ゆえに $a \leq 0$ であり, これと $a \geq 0$ を合わせて $a=0$ を得る. よって $b=-a=0$ であり, $a=b=0$. いずれの場合においても $a=b$ が示され, 以上より f は単射である. f が全射でないことは (1) と同様にして示される. \square

(3) $f: [-2, 2] \rightarrow [0, 4]$ を $f(x) = x^2$ と定めれば, これは単射はないが, 全射である.

Proof. f が単射でないことは (1) と同様にして示される. また f の像は $f([-2, 2]) = [0, 4]$ であり, これは終域に一致する. ゆえに f は全射である. \square

(4) $f: [0, 2] \rightarrow [0, 4]$ を $f(x) = x^2$ と定めれば, これは全単射である.

Proof. 単射性は (2) と同様にして示され, 全射性は (3) と同様にして示される. \square

定義 19.2.4. 写像 $f: X \rightarrow Y$ および部分集合 $A \subset X$ が与えられているとき, A の各元 $a \in A$ に対して Y の元 $f(a)$ を対応させる写像を f の A における制限 (restriction) と呼び, これを $f|_A: A \rightarrow Y$ と書く. また, 部分集合を定義域とする写像 $g: A \rightarrow Y$ に対して, 新たに定めた $\tilde{g}: X \rightarrow Y$ が $\tilde{g}|_A = g$ を満たすとき, \tilde{g} は g の拡張 (extension) であるという.

任意の写像 $f: X \rightarrow Y$ に対して、適当な部分集合 $A \subset X$ に f を制限すれば $f|_A$ は単射となる。例えば、 A を一点集合とすればよい。あるいはもっと極端に $A = \emptyset$ とすれば何も元を対応させない単射 $f|_{\emptyset}$ を得る。ほかにも、各 $y \in f(X)$ に対して $f^{-1}(y)$ の元を一つだけ取ってこれを x_y とし、 $A = \{x_y \mid y \in f(X)\}$ とすれば $f|_A$ は像が f の像に一致する単射となる³⁹。一方、 $f: X \rightarrow Y$ の終域を $f(X)$ に置き換えれば、写像 $f: X \rightarrow f(X)$ は全射となる。とくに単射 $f: X \rightarrow Y$ において終域を置き換えた写像 $f: X \rightarrow f(X)$ は全単射である。このことから、単射のことを1対1写像⁴⁰と呼ぶ文献もある。

練習 19.2.5. 写像 $f: X \rightarrow Y$ について次を示せ。

(1) f が単射ならば、部分集合 $A \subset X$ について $f^{-1}(f(A)) = A$.

解答例: 集合が一致することを示すには、両方の包含関係を確認すればよい。各 $a \in A$ に対して $f(a) \in f(A)$ であるから $a \in f^{-1}(f(A))$ 。つまり $A \subset f^{-1}(f(A))$ である。次に、任意に $\alpha \in f^{-1}(f(A))$ を取ろう。すると $f(\alpha) \in f(A)$ であるから、ある $a_0 \in A$ を用いて $f(a_0) = f(\alpha)$ と書ける。 f の単射性より $\alpha = a_0 \in A$ 。したがって $f^{-1}(f(A)) \subset A$ 。□

(2) f が全射ならば、部分集合 $B \subset Y$ について $f(f^{-1}(B)) = B$.

解答例: $\beta \in f(f^{-1}(B))$ を任意にとれば、ある $x \in f^{-1}(B)$ を用いて $\beta = f(x)$ と書ける。 $x \in f^{-1}(B)$ より $f(x) \in B$ 、つまり $\beta \in B$ である。ゆえに $f(f^{-1}(B)) \subset B$ 。次に $b \in B$ を任意に取れば、 f の全射性より、ある $x' \in X$ を用いて $f(x') = b$ と書ける。 $f(x') \in B$ より $x' \in f^{-1}(B)$ であり、ゆえに $b = f(x') \in f(f^{-1}(B))$ 。したがって $B \subset f(f^{-1}(B))$ 。□

一般の写像においては、上の (1) および (2) が成り立つとは限らない。例えば、 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) := x^2$ と定めれば、 $A = \{2\}$ および $B = [-5, 3]$ について

$$f^{-1}(f(A)) = f^{-1}(f(\{2\})) = f^{-1}(4) = \{2, -2\} \neq A,$$

$$f(f^{-1}(B)) = f(f^{-1}([-5, 3])) = f([0, \sqrt{3}]) = [0, 3] \neq B.$$

以降で学ぶこと

線形写像 T_A の単射性は連立1次方程式 $Ax = \mathbf{0}$ が唯一解を持つこと、つまり A の各列ベクトルの線形独立性で特徴づけられる (例 20.3.2 および命題 20.3.4)。全射性は A の列ベクトルで生成される部分空間が終域に一致すること、あるいは A の行の数と階数が一致することによって特徴づけられる。

19.3 逆写像

写像 $f: X \rightarrow Y$ が全射かつ単射であるとは、次を満たすことであった:

いかなる $y \in Y$ に対しても、 $f(x_y) = y$ を満たす $x_y \in X$ が唯一つ存在する。

定義 19.3.1. 写像 $f: X \rightarrow Y$ が全単射であるとき、上の性質を f は満たす。そこで、各 $y \in Y$ に対して上の x_y (つまり f に代入すると y になる定義域の元のこと) を対応させる写像 $g: Y \rightarrow X$ ($g(y) := x_y$) を f の逆写像 (inverse map) とよび f^{-1} と書く。

注意: f が全単射であり $f(x) = y$ ならば、上の x_y は x に相当し、 $f^{-1}(y) = x$ である。

f の逆関数が定義できるのは、 f が全単射のときに限る (命題 19.3.4)。以下、 f の逆関数について論じる際は、 f が全単射であることを暗黙のうちに前提として話を進めていると考えよ。

³⁹ $f|_A(A) = f(X)$ を満たす $A \subset X$ の存在を主張する命題は選択公理と呼ばれる (正確には選択公理と同値な主張となる)。

⁴⁰ 単射を1対1写像と呼び、全単射を1対1対応と呼んで区別している。

いま、全単射 $f: X \rightarrow Y$ の逆関数 f^{-1} が与えられているとし、 $x \in X, y = f(x)$ であるとしよう。このとき、記号 $f^{-1}(y)$ には二つの異なる意味が与えられている。一つは逆像のことであり、 $f^{-1}(\{y\})$ を略した表記のことであり、 f は全単射であるから、これは一点からなる X の部分集合 $f^{-1}(y) = \{x\}$ になる。もう一つの意味は、逆写像 $f^{-1}: Y \rightarrow X$ に y を代入した値のことであり、この立場では $f^{-1}(y)$ は X の元 x であり、部分集合ではない。記号 $f^{-1}(y)$ がどちらを意味しているかは文脈で判断しなければならないが、一点集合かあるいは一点集合の元かの違いは少なく、実質的な数学を理解するうえでは支障がないことが多い。これらの違いを厳密に区別する必要が生じるのは集合論においてのみである。一方、部分集合 $B \subset Y$ においても同様に記号 $f^{-1}(B)$ に二つの意味が与えられる。しかし、こちらは結果として同じ集合を表すことになり、どちらの意味で解釈しても構わない(練習 19.3.3)。

例 19.3.2. (1) 写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) := ax + b$ (ただし $a \neq 0$) と定めれば f は全単射であり、 $f^{-1}(y) = \frac{1}{a}y - \frac{b}{a}$ である。

(2) 写像 $f: [0, \infty) \rightarrow [0, \infty)$ を $f(x) = x^2$ と定めれば f は全単射であり、 $f^{-1}(y) = \sqrt{y}$ 。

(3) 写像 $f: (-\infty, 0] \rightarrow [0, \infty)$ を $f(x) = x^2$ と定めれば f は全単射であり、 $f^{-1}(y) = -\sqrt{y}$ 。

初等解析学では、関数に代入する変数に文字 x を用いることが多い。この慣習を踏襲すると逆関数に代入する変数も文字 x を用いることになる。例えば $f(x) = ax + b$ の逆関数は $f^{-1}(x) = \frac{1}{a}x - \frac{b}{a}$ と書かれる。この場合、 $f: X \rightarrow Y$ の逆関数 $f^{-1}: Y \rightarrow X$ に代入する x は Y の元である。つまり、集合 Y の元を表す文字に x を用いることになる。変数 x は必ず X の元であると勘違いしてはいけない。

練習 19.3.3. $f: X \rightarrow Y$ を全単射とし、 $B \subset Y$ とする。逆写像 f^{-1} の B による像 I と、 f の B による逆像 P が一致することを示せ(既に述べたように、 I, P はいずれも記号 $f^{-1}(B)$ で表される)。

解答例: ($I \subset P$): $i \in I$ とすれば、ある $b \in B$ を用いて $i = f^{-1}(b)$ と書ける。逆写像の定義により $f(i) = b$ である。 i を f に代入すると B の元になるゆえ、 B による f の逆像 P に i は含まれる。

($P \subset I$): $p \in P$ とすれば、 $f(p) \in B$ である。このとき $p = f^{-1}(f(p))$ ゆえ、 B による f^{-1} の像 I に p は含まれる。□

f^{-1} が逆関数と呼ばれるゆえんは、お互いに合成するとともに戻るという性質にある。次は定義 19.3.1 の下にある注意をより詳しく述べたものであり、この事実を前提として、下の命題文に現れる条件 (1) と (2) を満たす g のことを f の逆関数と定義する流儀もある。

命題 19.3.4. $f: X \rightarrow Y$ が全単射であることと、次の性質 (1), (2) を満たす $g: Y \rightarrow X$ が存在することは同値である。更に、この g は f の逆関数に一致する。

(1) いかなる $x \in X$ についても $g \circ f(x) = x$ が成り立つ(すなわち $g \circ f = \text{id}_X$)。

(2) いかなる $y \in f(X)$ についても $f \circ g(y) = y$ が成り立つ(すなわち $f \circ g = \text{id}_Y$)。

Proof. まず f が全単射であると仮定し、 $g = f^{-1}$ が性質 (1) と (2) を満たすことを示そう。(1) を示すために $x \in X$ を勝手に取る。 $y := f(x) \in Y$ とおこう。逆写像の定義よれば、 $g(y)$ とは f に代入すると y になる X の唯一の元、すなわち x のことであり、ゆえに $g(y) = x$ である。つまり、 $g \circ f(x) = g(f(x)) = g(y) = x$ 。次に (2) を示すために $y \in Y$ を勝手に取る。逆関数の定義によれば、 $g(y)$ は f に代入すると y になる元のことであり。つまり $f(g(y)) = y$ 、すなわち $f \circ g(y) = y$ 。

次に条件 (1) と (2) を満たす g が存在するとき、 f が全単射であることを示そう。ここでは対偶をとり、 f の単射性と全射性のいずれか一方でも欠けると (1) かつ (2) を満たす g が作れないことを示そう。 f が単射でないとする、ある $y \in f(X)$ について $f(x_1) = f(x_2) = y$ を満たす二つの異なる元 $x_1, x_2 \in X$ が存在する。条件 (1) を満たす g を定義しようとするとき、 $g(y)$ は x_1 と定めるべきだろうか、それとも x_2 と定めるべきだろうか。前者を採用すると $g \circ f(x_2) = g(f(x_2)) = g(y) = x_1$ となり、 g は (1) を満たさない。後者を採用しても $g \circ f(x_1) = x_2$ となり、やはり (1) が満たされない。もちろん、 x_1, x_2 以外の点 x_3 を採用しても $g \circ f(x_1) = x_3$ となり、やはり (1) は満たされない。すなわち、(1) を満たすように g

を定めることは出来ないことが分かる. f が全射でない場合は, f に代入した値にはなり得ない $y \in Y$ が存在しており, このとき $g: Y \rightarrow X$ をどう定義するにしても条件 (2) を満たすことはない. 何故なら, $f \circ g(y) = f(g(y))$, つまり $f \circ g(y)$ は f に $g(y)$ を代入した値である. ゆえに, f に代入した値にはなり得ない元 y と $f \circ g(y)$ は異なり, $f \circ g(y) \neq y$.

最後に, (1) と (2) を満たす関数 g が f^{-1} に一致することを示そう. f^{-1} も (1) と (2) を満たすことは既に示している. とくに (2) より各 $y \in Y$ に対して, $f \circ g(y) = y = f \circ f^{-1}(y)$ となる. すなわち $f(g(y)) = f(f^{-1}(y))$ であり, f の単射性より $g(y) = f^{-1}(y)$. つまり, いかなる Y の元についても, g および f^{-1} で写した値が等しいゆえ $g = f^{-1}$ である. \square

上の証明の第二段落では, 対偶をとらずに (1) と (2) から直接 f の全単射性を示すこともできる (練習 19.3.7). また, そのほうが証明はエレガントである. にも関わらず, ここで対偶による証明を採用したのは, 全単射でない写像が逆写像 (すなわち (1) と (2) を満たす写像) を持たない理由の具体的な説明になっているからである.

備考 19.3.5. $f: X \rightarrow X$ が全単射であることの必要十分条件は, $f \circ g = \text{id}_X = g \circ f$ を満たす $g: X \rightarrow X$ が存在することと同値である. これは, A が可逆であることの定義 (すなわち $AB = E = BA$ を満たす B が存在すること) の写像の言葉による言い換えに相当する.

練習 19.3.6. 全単射 h について次を示せ.

$$(i) \ h^{-1} \text{ も全単射である.} \quad (ii) \ (h^{-1})^{-1} = h.$$

解答例: $f = h^{-1}$, $g = h$ として命題 19.3.4 を適用すればよい. g は (1) および (2) を満たすゆえ $f = h^{-1}$ は全単射である. また, g は f の逆写像であるから $g = f^{-1} = (h^{-1})^{-1}$. すなわち $h = (h^{-1})^{-1}$. \square

練習 19.3.7 (発展). 命題 19.3.4 の (1) および (2) を満たす g が存在するならば f が全単射となることを直接証明せよ.

Proof. 単射性: $f(x_1) = f(x_2)$ とすれば, これらを g に代入し, $g(f(x_1)) = g(f(x_2))$ を得る. これと (1) を合わせれば $x_1 = g \circ f(x_1) = g(f(x_1)) = g(f(x_2)) = g \circ f(x_2) = x_2$. つまり $x_1 = x_2$.

全射性: 各 $y \in Y$ に対して, $x = g(y)$ とおくと, (2) より $y = f \circ g(y) = f(g(y)) = f(x)$. すなわち, f に代入すると y となる元 $x \in X$ が存在する. \square

グラフの対称性についても述べておこう. $A, B \subset \mathbb{R}$ および関数 $f: A \rightarrow B$ に対して, 次で与えられる \mathbb{R}^2 の部分集合を f のグラフという:

$$\Gamma_f := \{(x, f(x)) \mid x \in A\}$$

命題 19.3.8 (発展). X, Y を \mathbb{R} の部分集合, $f: X \rightarrow Y$ を全単射とし, 更に関数 $y = x$ が表す \mathbb{R}^2 の対角線を L とする. このとき, f のグラフと f^{-1} のグラフは, 直線 L を軸に線対称である.

Proof. 直線 L を軸に点 $(x, y) \in \mathbb{R}^2$ と対称な点は $(y, x) \in \mathbb{R}^2$ である. したがって, 次の同値性を示せばよい:

$$(x, y) \in \Gamma_f \iff (y, x) \in \Gamma_{f^{-1}}$$

(\Rightarrow) を示すために $(x, y) \in \Gamma_f$ を取れば, $y = f(x)$ である. ゆえに $f^{-1}(y) = x$ であり $(y, x) = (y, f^{-1}(y))$ は f^{-1} のグラフ上の点である. したがって $(y, x) \in \Gamma_{f^{-1}}$.

いま全単射 $h: Y \rightarrow X$ に関する命題 「 $(b, a) \in \Gamma_h \implies (a, b) \in \Gamma_{h^{-1}}$ 」 を示したと言ってもよい. この命題に $h = f^{-1}$, $(b, a) = (y, x)$ を適用することで, 「 $(y, x) \in \Gamma_{f^{-1}} \implies (x, y) \in \Gamma_f$ 」を得る. \square

関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ のグラフ上の点 $(\alpha, f(\alpha))$ における接線 l を表す関数が $g(x) = ax + b$ で与えられているとき、逆関数 f^{-1} のグラフ上の点 $(f(\alpha), \alpha)$ の接線 l' は、直線 L を軸に l と線対称である。つまり l' は $g^{-1}(y) = \frac{1}{a}y - \frac{b}{a}$ のグラフであり、その傾きは $1/a$ となる。このことから点 $\beta = f(\alpha)$ における逆関数の微分公式 $(f^{-1})'(\beta) = \frac{1}{f'(\alpha)}$ が示唆される。

以降で学ぶこと

T_A が全単射であることと A が可逆行列になることは同値になる。このとき、その逆写像は $T_{A^{-1}}$ で与えられる (命題 21.3.3)。とくに、 A が正方行列でなければ T_A は全単射でないことになる。この事実を道具立てを何もせずを示すことは意外に難しいのではないか。

19.4 写像の合成

全射性および単射性と写像の合成との関係について補足しておこう。

命題 19.4.1. 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ が与えられているとする。

- (1) f, g が共に単射ならば $g \circ f: X \rightarrow Z$ も単射である。
- (2) f, g が共に全射ならば $g \circ f: X \rightarrow Z$ も全射である。
- (3) f, g が共に全単射ならば $g \circ f: X \rightarrow Z$ も全単射である。

Proof. (1): $g \circ f(x_1) = g \circ f(x_2)$ とすれば $g(f(x_1)) = g(f(x_2))$ であり、 g の単射性より $f(x_1) = f(x_2)$ である。これに f の単射性を適用し $x_1 = x_2$ を得る。(2): 各 $z \in Z$ に対して、 g の全射性より $z = g(y)$ を満たす $y \in Y$ が存在する。また、この y に対して f の全射性を適用すると、 $y = f(x)$ を満たす $x \in X$ が存在する。このとき $z = g \circ f(x)$ 。(3): (1) および (2) より明らか。□

練習 19.3.7 で行った議論は次のように分解できる。

命題 19.4.2 (発展). 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ が与えられているとする。

- (1) $g \circ f: X \rightarrow Z$ が単射ならば f も単射である。
- (2) $g \circ f: X \rightarrow Z$ が全射ならば g も全射である。

Proof. (1): $f(x_1) = f(x_2)$ とすれば、これらを g に代入し $g \circ f(x_1) = g \circ f(x_2)$ を得る。これに $g \circ f$ の単射性を適用し $x_1 = x_2$ を得る。すなわち f は単射である。

(2): 各 $z \in Z$ に対して、 $g \circ f$ の全射性より $z = g \circ f(x)$ を満たす $x \in X$ が存在する。このとき、 $y := f(x)$ とおけば $y \in Y$ であり、 $z = g \circ f(x) = g(f(x)) = g(y)$ 。ゆえに g は全射である。□

練習 19.4.3. 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ を全単射とする。 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$ を示せ。

解答例: 写像の合成に関する結合律 (命題 3.2.5) から $(f^{-1} \circ g^{-1}) \circ (g \circ f) = \text{id}_X$ が容易に確かめられる。実際、

$$(f^{-1} \circ g^{-1}) \circ (g \circ f) = f^{-1} \circ (g^{-1} \circ (g \circ f)) = f^{-1} \circ ((g^{-1} \circ g) \circ f) = f^{-1} \circ (\text{id}_Y \circ f) = f^{-1} \circ f = \text{id}_X.$$

同様にして $(g \circ f) \circ (f^{-1} \circ g^{-1}) = \text{id}_Z$ も示され、ゆえに命題 19.3.4 より $f^{-1} \circ g^{-1}$ は $g \circ f$ の逆写像である。□

備考 19.4.4. 可逆行列 B, A に対して BA の逆行列は $A^{-1}B^{-1}$ であった。この事実を写像の言葉で言い換えたものが上の練習に他ならない。

逆像についても、練習 19.4.3 と類似の性質が成り立つ。

命題 19.4.5. 写像 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$, 部分集合 $C \subset Z$ に対して, $(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$.

Proof. $x \in (g \circ f)^{-1}(C) \iff g \circ f(x) \in C \iff g(f(x)) \in C \iff f(x) \in g^{-1}(C) \iff x \in f^{-1}(g^{-1}(C))$. □

上の性質を有限回適用することで, n 個の写像の合成についても同様の性質がなりたつ. 例えば三つの写像の合成については次のようになる:

$$(h \circ g \circ f)^{-1}(D) = (h \circ (g \circ f))^{-1}(D) = (g \circ f)^{-1}(h^{-1}(D)) = f^{-1}(g^{-1}(h^{-1}(D))).$$

以降で学ぶこと

19.2 で予告したことと命題 19.4.2 を合わせると直ちに次が導かれる:

命題 19.4.6. (m, n) -行列 A および (n, r) -行列 B について次が成り立つ.

- (1) 積 AB の各列が線形独立ならば B の各列も線形独立である.
- (2) 積 AB の各列が \mathbb{R}^m を生成するならば A の各列も \mathbb{R}^m を生成する.

練習 19.4.7. これまでに学習した知識から上の (1) を示せ.

解答例: 方程式 $Bx = \mathbf{0}$ が唯一解をもつことを示そう. ベクトル x を方程式 $Bx = \mathbf{0}$ の解とすれば, $ABx = A(Bx) = A\mathbf{0} = \mathbf{0}$ ゆえ x は $ABx = \mathbf{0}$ の解でもある. 仮定より AB の各列は線形独立であり, 命題 17.3.1 より $ABx = \mathbf{0}$ の解は唯一である. ゆえに $x = \mathbf{0}$. 以上より, $Bx = \mathbf{0}$ の解は自明なものに限る. 再び命題 17.3.1 より, B の各列は線形独立である. □

19.5 無限集合 (発展)

定義 8.1.1 をよく読むと, 置換を単射 $f: X_n \rightarrow X_n$ のことと定めている. しかし, 置換は逆写像 (逆置換) を持つゆえ全単射である. この点について補足しておこう.

命題 19.5.1. X, Y をともに n 点からなる集合とする. 写像 $f: X \rightarrow Y$ において次は同値である:

- (1) f は単射である,
- (2) f は全射である.

Proof. $X = \{x_1, \dots, x_n\}, Y = \{y_1, \dots, y_n\}$ において証明しよう.

(1) \Rightarrow (2): f は単射ゆえ $f(x_1), \dots, f(x_n)$ の中に重複はない. もし f が全射でないと仮定すれば, $f(x_1), \dots, f(x_n)$ のいずれでもない $y \in Y$ が存在する. このとき Y は $n+1$ 個の元 $f(x_1), \dots, f(x_n), y$ を含むことになり, これは Y が n 点集合であることに反する. ゆえに f は全射である.

(2) \Rightarrow (1): f を全射とする. もし f が単射でないと仮定すれば, $x_i \neq x_j$ かつ $f(x_i) = f(x_j)$ なる $i \neq j$ が取れる. このとき $f(x_1), \dots, f(x_n)$ のうち $f(x_i)$ と $f(x_j)$ は等しいゆえ, 集合 $f(X) = \{f(x_1), \dots, f(x_n)\}$ の点の数は $n-1$ 以下となる. f の全射性より $Y = f(X)$ であり, つまり Y の元の数は $n-1$ 以下である. これは Y が n 点集合であることに反する. ゆえに f は単射である. □

本論では「 n 点からなる集合」の正確な定義には踏み込まず, 素朴な立場で論じている⁴¹. 構成する元の個数がちょうど n 個の集合を「 n 点からなる集合」と定める立場において, 次の命題は, 二つの集合の元の個数が一致することが全単射の存在によって特徴づけられることを述べている:

命題 19.5.2. (1) 集合 X, Y が n 点からなるとすると, 全単射 $f: X \rightarrow Y$ が存在する.

(2) X が n 点からなるとし, 全単射 $f: X \rightarrow Y$ があれば Y も n 点集合になる.

⁴¹そもそも n 点集合なる概念を集合論的な道具立てのみでいかに定義するかを考えると, 「集合 $X_n := \{1, \dots, n\}$ との間」に全単射がある集合」と定めるしかない. その意味において命題 19.5.2 は明らかであり, これはナンセンスな主張である. ただし, この厳密な定義を採用する場合, n 点集合と $n-1$ 点集合の間に全単射が存在しないことは明らかではなく, 別途証明する必要がある.

(3) Y が n 点からなるとし、全単射 $g: X \rightarrow Y$ があれば X も n 点集合になる。

Proof. (1): X, Y がともに n 点からなるとすれば、重複のない表示 $x_1, \dots, x_n, y_1, \dots, y_n$ を用いて $X = \{x_1, \dots, x_n\}, Y = \{y_1, \dots, y_n\}$ と書ける。このとき、 $f: X \rightarrow Y$ を $f(x_i) := y_i$ ($i = 1, \dots, n$) と定めれば全単射である。

(2): X は n 点からなるゆえ重複のない表示 x_1, \dots, x_n を用いて $X = \{x_1, \dots, x_n\}$ と書ける。 $f: X \rightarrow Y$ を全単射とし、 $y_i := f(x_i)$ ($i = 1, \dots, n$) とおこう。 f の全射性より $Y = f(X) = \{f(x_1), \dots, f(x_n)\} = \{y_1, \dots, y_n\}$ である。また、 f の単射性は y_1, \dots, y_n の中に重複がないことを意味している。実際、もし仮に重複があり $y_i = y_j$ ($i \neq j$) となるならば、 $f(x_i) = f(x_j)$ と単射性より $x_i = x_j$ となり、これは x_1, \dots, x_n に重複がないことに反する。以上より、 $Y = \{y_1, \dots, y_n\}$ は n 点からなる。

(3): $g^{-1}: Y \rightarrow X$ は全単射であり、 $f = g^{-1}$ として (2) を適用すれば X は n 点集合である。□

そこで、集合の元の個数が有限でない場合についても、元の個数が一致するという概念を次で定めることができる:

定義 19.5.3. 集合 X, Y の間に全単射が存在するとき、 X と Y は対等である (あるいは濃度が等しい) という。

命題 19.5.4. X と Y が対等であり、 Y と Z が対等ならば X と Z も対等である。

Proof. 仮定より二つの全単射 $f: X \rightarrow Y$ および $g: Y \rightarrow Z$ が存在する。このとき $g \circ f: X \rightarrow Z$ は命題 19.4.1(3) より全単射であり、したがって X と Z は対等である。□

有限集合の場合と異なり、無限集合は自身の真部分集合⁴²と対等になり得る。このことは少なくともガリレオの時代には既に気づかれていた事実である。

例 19.5.5. 自然数全体 \mathbb{N} と正の偶数全体 $2\mathbb{N} = \{2n \mid n \in \mathbb{N}\}$ は対等である。実際、 $f: \mathbb{N} \rightarrow 2\mathbb{N}$ を $f(x) := 2x$ と定めればこれは全単射である。

一方で、すべての無限集合が対等というわけではない。

命題 19.5.6. \mathbb{N} と \mathbb{R} は対等ではない。

Proof. 全射 $f: \mathbb{N} \rightarrow \mathbb{R}$ が存在すると仮定すると矛盾が導けることを示そう。 $f: \mathbb{N} \rightarrow \mathbb{R}$ を全射とする。まず $f(1)$ を含まない閉区間 $[a_1, b_1]$ を取る (ただし $a_1 < b_1$)。次に、 $f(2)$ を含まない閉区間 $[a_2, b_2] \subset [a_1, b_1]$ を取る (ただし $a_2 < b_2$)。更に、 $f(3)$ を含まない区間 $[a_3, b_3] \subset [a_2, b_2]$ を取る (ただし $a_3 < b_3$)。この操作を順次繰り返していくと、単調増加数列 $a_1 \leq a_2 \leq a_3 \leq \dots$ および単調減少列 $b_1 \geq b_2 \geq b_3 \geq \dots$ を得る。このとき閉区間の取り方から、各 $n \in \mathbb{N}$ について $f(n) \notin [a_n, b_n]$ である。上に有界な単調増加数列は収束する (実数の連続性) ゆえ数列 a_n は収束し、この極限を $a = \lim_{n \rightarrow \infty} a_n$ とする。さて、 f の全射性より $f(n_0) = a$ を満たす $n_0 \in \mathbb{N}$ が存在する。このとき、各 $n > n_0$ について $a_{n_0} \leq a_n < b_n \leq b_{n_0}$ 、つまり $a_{n_0} \leq a_n \leq b_{n_0}$ である。したがって、その極限においても $a_{n_0} \leq a \leq b_{n_0}$ が成り立つ。すなわち $f(n_0) \in [a_{n_0}, b_{n_0}]$ である。ところが、閉区間 $[a_n, b_n]$ の取り方から $f(n_0) \notin [a_{n_0}, b_{n_0}]$ であり、矛盾を得た。□

例 19.5.7. \mathbb{Q} と \mathbb{N} は対等であることが知られている。また、 \mathbb{R} と対等な集合として無理数全体や \mathbb{C} 、あるいはユークリッド空間 \mathbb{R}^n などがある。

以上の事実から、無限集合の元の総数⁴³は一定ではないことが示唆される。とくに任意の無限集合 X が $X = \{x_1, x_2, \dots, \}$ と表せるわけではない。もしこのような表示ができるならば写像 $f: \mathbb{N} \rightarrow X$ を $f(n) := x_n$ と定めれば f は全射となる。更にあらかじめ部分列を取ることで x_1, x_2, \dots に重複がないようにしておけば f は全単射である。すなわち X と \mathbb{N} は対等である。いまの議論から $\mathbb{R} = \{x_1, x_2, \dots, \}$ と表せないことが分かる。

⁴²自分自身以外の部分集合のことを真部分集合 (proper subset) と呼ぶ。

⁴³無限集合も含めた文脈において、集合 X を構成する元の総数のことを X の濃度という。

よりみち (無限集合の不思議).

\mathbb{R} と \mathbb{R}^n は対等である. 次元の異なる空間の元の個数が等しいことを読者は不思議に感じるかもしれない. しかし, ここでいう全単射 $f: \mathbb{R} \rightarrow \mathbb{R}^n$ の存在性は, 代数構造や数列の極限など多くの数学的構造を無視したうえでの 1 対 1 対応があると述べているに過ぎないのである. 例えば線形空間としての演算を保つ写像に限れば \mathbb{R} と \mathbb{R}^n の間の 1 対 1 対応を作ることはできない (命題 21.2.7). また, 数列の発散・収束性を互いに保つ 1 対 1 対応 (これを同相写像という) も \mathbb{R} と \mathbb{R}^n の間には作れないことが知られている (22.1 項のコラムを参照). このように, 二つの対象を同じとみなす (つまり 1 対 1 の対応を与える) といっても, 様々な立場があり得る.

一方, 何の化粧もない単なる集合に限った場合, 無限集合の間の 1 対 1 対応はどこまで理解されているのだろうか. 実は, 実数の無限部分集合の大きさにどれくらいの種類があるかという基本的な問題ですら容易に理解されるものではなく, これは集合論の創始者であるカントールを生涯悩ませ続ける問題でもあった. すなわち, \mathbb{N} と \mathbb{R} と対等でない実数の部分集合 $X \subset \mathbb{R}$ は存在するかという問いである. このような X は存在しないという立場を連続体仮説といい, 連続体仮説 (あるいはその否定) の証明に彼は長い年月を費やしたが, いずれも証明することはできなかった. 現在では, 連続体仮説およびその否定のいずれも集合論の公理系からは導けないことが分かっており, これ以上この問題について論じるならば, 我々は連続体仮説とその否定のどちらか一方を公理として選択する必要に迫られることになる.

連続体仮説は微積分学とも無縁ではない. 逐次積分 (累次積分) の順序の入れ替え:

$$\int_a^b \int_c^d f(x, y) dx dy = \int_c^d \int_a^b f(x, y) dy dx.$$

について, 上の等式が成立するための f の条件を解析学では与えている. 一方, 上の等式が成立しない関数の例が存在するかという問題は授業ではあまり扱わないことが多い. 実は, この存在・非存在性も集合論の公理系からは導かれないこと, そして連続体仮説からは等号不成立の例が導けることが知られている.

集合論の公理系から導かれる不思議な事実についても述べておこう. それは, \mathbb{R}^3 の原点を中心とする半径 1 の球体を有限個の集合に分割し, これらを回転と並行移動によって上手く配置しなおすと半径 2 の球体になるという定理である (バナッハ・タルスキーの逆理). 一見するとこれは体積概念と矛盾するように思えるが, 体積が定義できないような複雑な集合に分割させることで, このような構成を実現させている.

以上, 無限集合の奥深さを示す例の一端をかけ足ながら取り上げた.

20 線形写像

線形代数学で扱う線形写像には二つの性格がある。一つは、分析すべき対象であり、そこには諸科学分野において現れる個々の具体的な線形写像をいかに理解するかが念頭にある。これこそが線形代数学の主題であるといってもよい。そしてもう一つは、一般のベクトル空間 V の言葉をユークリッド空間の言葉に翻訳するために与える対応 (線形同型) のことである。後者は前者を分析するための道具といえる。本節ではこれら二つの区別をせず、線形写像の定義から共通して得られる一般論を展開し、これらを区別した各論は次節以降に論じる。

ところで、線形写像の多くは行列によって表現され、以降では行列と標準ベクトル e_i の間の積に関する次の性質

$$Ae_i = A \text{ の第 } i \text{ 列目}$$

を断りなく用いる (本節では例 20.2.4 で用いた)。上式を頭の片隅に留めておいてもらいたい。

20.1 線形写像の基本的性質

1 節で述べた通り、線形写像とは比例関数の一般化に相当する概念である。

定義 20.1.1. 線形空間 V から線形空間 W への写像 $f: U \rightarrow V$ が次の性質 (i) および (ii) を満たすとき、 f を線形写像 (**linear map**) あるいは線形作用素 (**linear operator**) という:

- (i) すべての $x, y \in U$ に対して、 $f(x + y) = f(x) + f(y)$,
- (ii) すべての $x \in U$ および $r \in \mathbb{R}$ について、 $f(rx) = rf(x)$.

上の性質 (i), (ii) は線形性と呼ばれる。部分空間になるための条件がまとめられたように、線形性は次の性質 (iii) にまとめられる:

$$\text{線形性 (iii) すべての } x, y \in V \text{ および } a, b \in \mathbb{R} \text{ について、} f(ax + by) = af(x) + bf(y).$$

実際、(iii) において $a = b = 1$ とした場合が (i) であり、 $a = r, b = 0$ とした場合が (ii) である。また、(i) と (ii) を用いて (iii) は次のように導かれる: $f(ax + by) = f(ax) + f(by) = af(x) + bf(y)$ 。以上より、条件「(i) かつ (ii)」と条件 (iii) は同値である。以降、線形性の確認を (iii) によって判定することとしよう。

例 20.1.2. (m, n) -行列 A に対して写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ を $T_A(x) := Ax$ と定めれば、これは線形写像である。実際、線形性 (iii) は次のようにして確かめられる。

$$T_A(ax + by) = A(ax + by) = A(ax) + A(by) = a(Ax) + b(Ay) = aT_A(x) + bT_A(y).$$

上で定めた写像 T_A は今後頻繁に現れるゆえ忘れないこと。なお、ユークリッド空間の間の線形写像は必ず T_A の形で書ける (命題 21.3.4)。とくに写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ のうちで線形写像となるものは、 $(1, 1)$ -行列 $A = [a]$ を用いて表される比例関数 $T_A(x) = ax$ のみである。

練習 20.1.3. 次の写像は線形写像ではない。具体的な元を代入することにより、線形性 (i), (ii) のいずれも満たされないことを確認せよ。

- (1) 写像 $f: \mathbb{R} \rightarrow \mathbb{R}$ を $f(x) = x^2$ と定めれば、これは線形写像ではない。
- (2) 行列式を与える写像 $\det: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ は線形写像でない。

ユークリッド空間以外の線形空間における線形写像の例は本節の最後に述べるとして、しばらくは線形性から導かれる一般論を展開しよう。

命題 20.1.4. 線形写像 $f: U \rightarrow V$ において 0_U の行き先は 0_V である。すなわち $f(0_U) = 0_V$ 。

Proof. 各ベクトルを0倍すると零ベクトルになることを用いると, $f(\mathbf{0}_U) = f(0 \cdot \mathbf{0}_U) = 0 \cdot f(\mathbf{0}_U) = \mathbf{0}_V$. \square

線形性 (iii) は更に次のような有限和の形に一般化できる⁴⁴.

命題 20.1.5. $f : U \rightarrow V$ が線形性を満たすならば次も満たす:

$$\text{線形性 (iii)'} \quad \text{各 } \mathbf{u}_k \in U, r_k \in \mathbb{R} (k = 1, \dots, \ell) \text{ について, } f\left(\sum_{k=1}^{\ell} r_k \mathbf{u}_k\right) = \sum_{k=1}^{\ell} r_k f(\mathbf{u}_k).$$

Proof. 和の個数 ℓ に関する帰納法で示す. $\ell = 1$ の場合は線形性の性質 (ii) に他ならない. 和の個数が ℓ のときに等式が成立すると仮定し, 和の個数が $\ell + 1$ の場合について示そう.

$$\begin{aligned} f\left(\sum_{k=1}^{\ell+1} r_k \mathbf{u}_k\right) &= f\left(\left(\sum_{k=1}^{\ell} r_k \mathbf{u}_k\right) + r_{\ell+1} \mathbf{u}_{\ell+1}\right) \\ &= f\left(\sum_{k=1}^{\ell} r_k \mathbf{u}_k\right) + f(r_{\ell+1} \mathbf{u}_{\ell+1}) && \text{(線形性 (i))} \\ &= \sum_{k=1}^{\ell} r_k f(\mathbf{u}_k) + r_{\ell+1} f(\mathbf{u}_{\ell+1}) && \text{(帰納法の仮定と線形性 (ii))} \\ &= \sum_{k=1}^{\ell+1} r_k f(\mathbf{u}_k). \end{aligned}$$

\square

線形性 (iii)' は, 線形写像が線形関係を保存することを述べている:

命題 20.1.6. $f : U \rightarrow V$ を線形写像とし, $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ とする.

- (1) $\mathbf{x} \in U$ が $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合で書けるならば, $f(\mathbf{x})$ は $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形結合で書ける.
(すなわち, $\mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \implies f(\mathbf{x}) \in \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle$.)
- (2) $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_U \implies \sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}_V$.
- (3) $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ が線形独立ならば $\mathbf{u}_1, \dots, \mathbf{u}_n$ も線形独立である.
- (4) $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形従属ならば $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ も線形従属である.

Proof. (1): この主張は線形性 (iii)' の言い換えにすぎない. 実際, $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と書けるならば $f(\mathbf{x}) = f(\sum_{i=1}^n a_i \mathbf{u}_i) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$ であり, $f(\mathbf{x})$ は $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形結合で書ける.

(2): $\mathbf{x} = \mathbf{0}_U$ に対して (1) の証明と同等の計算をすればよい.

(3): 線形関係 $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_U$ を仮定すれば (2) より $\sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}_V$ である. 組 $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ の線形独立性より $a_1 = \dots = a_n = 0$. つまり $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である.

(4): これは (3) の対偶にほかならない. \square

次の例で見るように上の命題における (1) から (4) の逆はいずれも成り立たない. 逆が成り立つのは f が単射の場合に限る (命題 20.3.6).

例 20.1.7. $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ が定める線形写像 $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ について次が成り立つ:

- (1) $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ とする. このとき, $f(\mathbf{x}) = \mathbf{e}_1$, $f(\mathbf{e}_1) = \mathbf{e}_1$ である. よって, $f(\mathbf{x})$ は $f(\mathbf{e}_1)$ の線形結合で書ける. しかし \mathbf{x} を \mathbf{e}_1 の線形結合で書くことはできない.

⁴⁴ 命題 12.3.1 と同様の議論を行っている. 行列式を先に扱う都合上, 我々は線形性よりも複雑な多重線形性を先に論じていたのである.

(2) $0f(\mathbf{e}_1) + 1f(\mathbf{e}_2) = \mathbf{0} + \mathbf{0} = \mathbf{0}$ である. 一方で $0\mathbf{e}_1 + 1\mathbf{e}_2 = \mathbf{e}_2 \neq \mathbf{0}$ である.

(3) $\mathbf{e}_1, \mathbf{e}_2$ は線形独立であるが, $f(\mathbf{e}_1), f(\mathbf{e}_2)$ は線形独立でない.

次の命題は, 線形写像の値は基底の行き先によって決定されることを述べている.

命題 20.1.8. $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とし, $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ とする ($\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立である必要はなく, とくに重複があっても良い). このとき, $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす線形写像 $f: U \rightarrow V$ が唯一つ存在する.

Proof. U の各元 \mathbf{x} は実数の組 a_1, \dots, a_n を用いて $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と一意的に書ける (命題 18.2.3). f が線形性を満たすには $f(\mathbf{x}) := \sum_{i=1}^n a_i \mathbf{v}_i$ と定めるしかないが, 実際にこのように定めた f は線形性を満たす. 何故なら, 各 $\mathbf{x}, \mathbf{y} \in U$ を $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i, \mathbf{y} = \sum_{i=1}^n b_i \mathbf{u}_i$ と書けば, $r\mathbf{x} + s\mathbf{y}$ は $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形結合として $r\mathbf{x} + s\mathbf{y} = \sum_{i=1}^n (ra_i + sb_i) \mathbf{u}_i$ と一意的に書けており, f の定義から $f(\mathbf{x}) = \sum_{i=1}^n a_i \mathbf{v}_i, \mathbf{y} = \sum_{i=1}^n b_i \mathbf{v}_i, f(r\mathbf{x} + s\mathbf{y}) = \sum_{i=1}^n (ra_i + sb_i) \mathbf{v}_i$ である. したがって,

$$f(r\mathbf{x} + s\mathbf{y}) = \sum_{i=1}^n (ra_i + sb_i) \mathbf{v}_i = \sum_{i=1}^n ra_i \mathbf{v}_i + \sum_{i=1}^n sb_i \mathbf{v}_i = r \left(\sum_{i=1}^n a_i \mathbf{v}_i \right) + s \left(\sum_{i=1}^n b_i \mathbf{v}_i \right) = rf(\mathbf{x}) + sf(\mathbf{y}).$$

すなわち, 線形性 (iii) を満たす f は線形写像である.

次に f の一意性を示すために, 線形写像 $g: U \rightarrow V$ も $g(\mathbf{u}_i) = \mathbf{v}_i$ を満たすと仮定し $f = g$ を導こう. そのためには各 $\mathbf{x} \in U$ について $f(\mathbf{x}) = g(\mathbf{x})$ を示せばよい. $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と書けており, f の定義と g の線形性 (iii)' を用いて変形すると,

$$f(\mathbf{x}) = \sum_{i=1}^n a_i \mathbf{v}_i = \sum_{i=1}^n a_i g(\mathbf{u}_i) = g \left(\sum_{i=1}^n a_i \mathbf{u}_i \right) = g(\mathbf{x}).$$

ゆえに $f = g$ である. □

上の証明において, 各 $\mathbf{x} \in U$ が $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ と一意的に書けるとは限らないとすると f を簡単には定義できなくなる. 例えば, $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i = \sum_{i=1}^n b_i \mathbf{u}_i$ と二通りに書けるとすると $f(\mathbf{x})$ を $\sum_{i=1}^n a_i \mathbf{v}_i$ と $\sum_{i=1}^n b_i \mathbf{v}_i$ のいずれにすべきか, あらかじめ指定する必要が生じる. こうなると, 線形性 (iii) が満たされるように定めるのは絶望的である.

命題 20.1.9. 線形写像の合成はまた線形写像となる.

Proof. $f: U \rightarrow V$ および $g: V \rightarrow W$ を線形写像とし, これらの合成 $g \circ f: U \rightarrow W$ が線形性 (iii) を満たすことを示そう. $\mathbf{x}, \mathbf{y} \in U$ および $r, s \in \mathbb{R}$ に対して,

$$g \circ f(r\mathbf{x} + s\mathbf{y}) = g(f(r\mathbf{x} + s\mathbf{y})) = g(rf(\mathbf{x}) + sf(\mathbf{y})) = rg(f(\mathbf{x})) + sg(f(\mathbf{y})) = r(g \circ f)(\mathbf{x}) + s(g \circ f)(\mathbf{y}).$$

□

20.2 線形写像による像

写像 $f: X \rightarrow Y$ に対して, X の元を f に代入することで得られる Y の元をすべて集めた集合を f の像と呼び $f(X)$ と書くのであった. 線形写像における像には特別な記号 $\text{Im } f$ が用いられる.

定義 20.2.1. 線形空間 U から線形空間 V への線形写像 $f: U \rightarrow V$ において, f による U の像 $f(U)$ のことを $\text{Im } f$ と書く. すなわち, $\text{Im } f := \{f(\mathbf{u}) \mid \mathbf{u} \in U\}$.

命題 20.2.2. 線形写像 $f: U \rightarrow V$ の像 $\text{Im } f$ は V の部分空間である.

Proof. 部分空間になるための条件 (i) および (iv) を確認する.

(i): $f(\mathbf{0}_U) = \mathbf{0}_V$ ゆえ $\mathbf{0}_V \in \text{Im } f$ である.

(iv): $\mathbf{v}_1, \mathbf{v}_2 \in \text{Im } f$ とすれば, これらは f に代入して得られる元である. すなわち, $f(\mathbf{u}_1) = \mathbf{v}_1$, $f(\mathbf{u}_2) = \mathbf{v}_2$ を満たす $\mathbf{u}_1, \mathbf{u}_2 \in U$ が存在する. このとき, 各 $r, s \in \mathbb{R}$ について $r\mathbf{v}_1 + s\mathbf{v}_2 = rf(\mathbf{u}_1) + sf(\mathbf{u}_2) = f(r\mathbf{u}_1 + s\mathbf{u}_2)$ である. つまり, ベクトル $r\mathbf{v}_1 + s\mathbf{v}_2 \in V$ はベクトル $r\mathbf{u}_1 + s\mathbf{u}_2 \in U$ を f に代入することで得られる. ゆえに $r\mathbf{v}_1 + s\mathbf{v}_2 \in \text{Im } f$. \square

$f: U \rightarrow V$ を線形写像とし, $W \subset U$ を U の部分空間とする. f の W による像 $f(W)$ は, f の W への制限 $f|_W: W \rightarrow V$ の像とも見なせる. また, $f|_W$ も線形写像であることから, $f(W) = \text{Im } f|_W$ と書いてもよい. 上の命題から $f(W)$ も V の部分空間となることが分かる.

命題 20.2.3. 線形写像 $f: U \rightarrow V$ および組 $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ において,

$$f(\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle) = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle.$$

とくに $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U を生成するとき, $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は $\text{Im } f$ を生成する.

Proof.

$$\begin{aligned} f(\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle) &= \{ f(\mathbf{x}) \mid \mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \} \\ &= \left\{ f\left(\sum_{i=1}^n a_i \mathbf{u}_i\right) \mid a_1, \dots, a_n \in \mathbb{R} \right\} \quad (\text{この等号は後述する}) \\ &= \left\{ \sum_{i=1}^n a_i f(\mathbf{u}_i) \mid a_1, \dots, a_n \in \mathbb{R} \right\} \quad (f \text{ の線形性 (iii)' を用いた}) \\ &= \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle. \end{aligned}$$

上の変形で明らかでないのは $\{ f(\mathbf{x}) \mid \mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \} = \{ f(\sum_{i=1}^n a_i \mathbf{u}_i) \mid a_1, \dots, a_n \in \mathbb{R} \}$ のみゆえ, これを示そう. $A = \{ f(\mathbf{x}) \mid \mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle \}$, $B = \{ f(\sum_{i=1}^n a_i \mathbf{u}_i) \mid a_1, \dots, a_n \in \mathbb{R} \}$ とする. 集合の一致を示すには両方の包含関係を確認すればよい.

($A \subset B$): 各 $f(\mathbf{x}) \in A$ において, $\mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ より a_1, \dots, a_n を用いて $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ とかける. つまり $f(\mathbf{x}) = f(\sum_{i=1}^n a_i \mathbf{u}_i) \in B$.

($B \subset A$): 各 $f(\sum_{i=1}^n a_i \mathbf{u}_i) \in B$ において, $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ とおけば $\mathbf{x} \in \langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ である. つまり $f(\sum_{i=1}^n a_i \mathbf{u}_i)$ は, $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle$ の元 \mathbf{x} を用いて $f(\mathbf{x})$ と書ける. したがって $f(\sum_{i=1}^n a_i \mathbf{u}_i) \in A$. \square

上の命題より, 線形全射は「空間を生成する」という状況を保存する写像である.

例 20.2.4. (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ による線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($T_A(\mathbf{x}) := A\mathbf{x}$) において, 定義域 \mathbb{R}^n の標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ を f に代入した値 $A\mathbf{e}_1, \dots, A\mathbf{e}_n$ は A の各列ベクトル $\mathbf{a}_1, \dots, \mathbf{a}_n$ に一致する. したがって T_A の像は, A の各列ベクトルによって生成される \mathbb{R}^m の部分空間に等しい. すなわち, $\text{Im } T_A = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$.

20.3 線形写像による逆像

写像 $f: X \rightarrow Y$ および $B \subset Y$ に対して, f に代入すると B の元となるような X の元をすべて集めた集合を f の W による逆像と呼び, $f^{-1}(W)$ と書くのであった. 線形写像においては, 特別な逆像である核が調べられる.

定義 20.3.1. 線形空間 U から線形空間 V への線形写像 $f: U \rightarrow V$ において, f による $\mathbf{0}_V$ の逆像 $f^{-1}(\{\mathbf{0}_V\})$ のことを $\text{Ker } f$ と書き, これを f の核 (kernel) という. すなわち, $\text{Ker } f := \{ \mathbf{u} \in U \mid f(\mathbf{u}) = \mathbf{0}_V \}$.

例 20.3.2. (m, n) -行列 A による線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($T_A(\mathbf{x}) := A\mathbf{x}$) の核とは, $A\mathbf{x} = \mathbf{0}$ を満たす点 $\mathbf{x} \in \mathbb{R}^n$ 全体からなる集合である. ゆえに $\text{Ker } T_A$ は連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ の解空間 W_A に等しい.

命題 16.2.2 において斉次形連立 1 次方程式の解空間が部分空間になることを見た。一般の線形写像においても、その核は定義域の部分空間になる。より一般に、次が成り立つ。

命題 20.3.3. 線形写像 $f: U \rightarrow V$ および終域の部分空間 $W \subset V$ に対して、 f の W による逆像 $f^{-1}(W)$ は U の部分空間である。

Proof. 部分空間になるための条件 (i) および (iv) を確認する。

(i): $f(\mathbf{0}_U) = \mathbf{0}_V \in W$ ゆえ $\mathbf{0}_U$ を f に代入すると W の元となる。すなわち、 $\mathbf{0}_U \in f^{-1}(W)$ である。

(iv): $\mathbf{u}_1, \mathbf{u}_2 \in f^{-1}(W)$ とすれば、 $f(\mathbf{u}_1), f(\mathbf{u}_2) \in W$ である。このとき W が部分空間であることから、各 $r, s \in \mathbb{R}$ に対して、 $f(r\mathbf{u}_1 + s\mathbf{u}_2) = rf(\mathbf{u}_1) + sf(\mathbf{u}_2) \in W$ である。すなわち、 $r\mathbf{u}_1 + s\mathbf{u}_2 \in f^{-1}(W)$ 。□

命題 20.3.4. 線形写像 $f: U \rightarrow V$ において次は同値である：

- (1) f は単射である。すなわち、いかなる $\mathbf{v} \in \text{Im } f$ においても、その逆像 $f^{-1}(\mathbf{v})$ が 1 点からなる、
- (2) $\text{Ker } f = \{\mathbf{0}_U\}$ 。

Proof. (1) \Rightarrow (2): (2) は、(1) における \mathbf{v} として $\mathbf{0}_V = f(\mathbf{0}_U) \in \text{Im } f$ を考えた特別な場合である。

(2) \Rightarrow (1): $f(\mathbf{x}) = f(\mathbf{y})$ を仮定し、 $\mathbf{x} = \mathbf{y}$ を示そう。ベクトル $\mathbf{x} - \mathbf{y}$ を f に代入すると $f(\mathbf{x} - \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) = \mathbf{0}_V$ 。つまり $\mathbf{x} - \mathbf{y} \in \text{Ker } f$ であり、仮定より $\text{Ker } f$ は零ベクトル $\mathbf{0}_U$ のみからなるとしていたから $\mathbf{x} - \mathbf{y} = \mathbf{0}_U$ である。これを移項して $\mathbf{x} = \mathbf{y}$ を得る。□

例 20.3.5. (m, n) -行列 A において、 T_A が単射であることと $A\mathbf{x} = \mathbf{0}$ が唯一解を持つことは同値である。

単射線形写像においては命題 20.1.6 の逆も成り立つ。すなわち、単射線形写像は線形独立性を保存する写像である：

命題 20.3.6. $f: U \rightarrow V$ を線形写像とし、 $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ とする。 f が単射であるとき次が成り立つ：

- (1) $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i \iff f(\mathbf{x}) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$
- (2) $\sum_{i=1}^n a_i \mathbf{u}_i = \mathbf{0}_U \iff \sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}_V$ 。
- (3) $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である $\iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形独立である。
- (4) $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形従属である $\iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形従属である。
- (5) $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V の基底である $\implies \mathbf{u}_1, \dots, \mathbf{u}_n$ は U の基底である。

Proof. まず(1)を示そう。(1) \Rightarrow は命題20.1.6で得られているゆえ(\Leftarrow)のみ証明する。 $f(\mathbf{x}) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$ とする。このとき、 $\mathbf{y} = \sum_{i=1}^n a_i \mathbf{u}_i$ とおき、 $\mathbf{x} - \mathbf{y}$ を f に代入すると、

$$f(\mathbf{x} - \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) = f(\mathbf{x}) - f\left(\sum_{i=1}^n a_i \mathbf{u}_i\right) = \sum_{i=1}^n a_i f(\mathbf{u}_i) - \sum_{i=1}^n a_i f(\mathbf{u}_i) = \mathbf{0}.$$

ゆえに $\mathbf{x} - \mathbf{y} \in \text{Ker } f$ である。 f は単射ゆえ $\text{Ker } f$ は $\mathbf{0}_U$ のみからなる。ゆえに $\mathbf{x} - \mathbf{y} = \mathbf{0}_U$ 。すなわち $\mathbf{x} = \mathbf{y} = \sum_{i=1}^n a_i \mathbf{u}_i$ 。

(2) は、(1) において $\mathbf{x} = \mathbf{0}_U$ とした特別な場合である。(3) は (2) より直ちに得られる。(4) は、(3) における両条件の否定をとった条件ゆえ、これらも同値である。

最後に (5) を示す。 $\mathbf{u}_1, \dots, \mathbf{u}_n$ の線形独立性は (3) より分かっている。各 $\mathbf{x} \in U$ に対して $f(\mathbf{x}) \in V$ であり、 $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ が V を生成することから $f(\mathbf{x}) = \sum_{i=1}^n a_i f(\mathbf{u}_i)$ と書ける。(1) より $\mathbf{x} = \sum_{i=1}^n a_i \mathbf{u}_i$ 。すなわち $\mathbf{u}_1, \dots, \mathbf{u}_n$ は U を生成する。□

命題 20.3.6(5) の逆が成り立つためには、 $\text{Im } f = V$ となる必要がある。つまり、 f が全単射でなければならない(命題 21.1.5)。

練習 20.3.7. 命題 20.1.8 の設定において v_1, \dots, v_n が線形独立であるとき, $f(u_i) = v_i$ を満たす線形写像 $f: U \rightarrow V$ は単射である. これを示せ.

解答例: $\text{Ker } f = \{0_U\}$ を示せばよい. そこで $x \in \text{Ker } f$ とする. u_1, \dots, u_n が U を生成することから $x = \sum_{i=1}^n a_i u_i$ と書けば,

$$0 = f(x) = f\left(\sum_{i=1}^n a_i u_i\right) = \sum_{i=1}^n a_i f(u_i) = \sum_{i=1}^n a_i v_i.$$

v_1, \dots, v_n は線形独立ゆえ $a_1 = \dots = a_n = 0$ であり, $x = \sum_{i=1}^n 0 u_i = 0_U$ を得る. すなわち $\text{Ker } f$ の元は 0_U に限る. \square

20.4 様々な線形写像の例

例 20.4.1. \mathbb{R}^n のベクトルに対して第 i 座標を対応させる写像 $p_n: \mathbb{R}^n \rightarrow \mathbb{R}$ ($f(x_1, \dots, x_n) := x_i$) は第 i 座標への射影 (**projection**) と呼ばれる. p_i は線形写像である.

例 20.4.2. n 次正方行列 $A = [a_{ij}]$ に対して A の対角成分の和 $\text{tr } A := \sum_{i=1}^n a_{ii}$ を A の跡 (トレース) という. 写像 $\text{tr}: M_n(\mathbb{R}) \rightarrow \mathbb{R}$ は線形写像である.

例 20.4.3. (1) 数列空間 $\mathbb{R}^{\mathbb{N}}$ において列を左にずらす写像, すなわち $S(x_1, x_2, x_3, \dots) := (x_2, x_3, \dots)$ で定められる写像 $S: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ はシフト作用素と呼ばれる. S は線形写像である.

(2) S は全射である. 何故なら, 各数列 $y = (y_1, y_2, y_3, \dots)$ に対して, $x = (0, y_1, y_2, \dots)$ と定めれば $S(x) = y$ となる. また, $S(x) = 0$ なる数列 x は, 初項を除いてすべて 0 なる数列であり, $\text{Ker } S = \{(x, 0, 0, \dots) \mid x \in \mathbb{R}\}$.

(3) 数列 $x = (x_n)_{n \in \mathbb{N}}$ が線形漸化式 16.3.2:

$$x_{n+k} = a_{k-1}x_{n+k-1} + a_{k-2}x_{n+k-2} + \dots + a_1x_{n+1} + a_0x_n$$

を満たすならば, 数列 $S(x) = (x_{n+1})_{n \in \mathbb{N}}$ もまた上の漸化式を満たす.

(4) 漸化式 $x_{n+2} = x_{n+1} + x_n$ を満たす数列全体 F は $\mathbb{R}^{\mathbb{N}}$ の部分空間である (例 16.3.1). S を F に制限した $S: F \rightarrow F$ は全単射である.

Proof. 全射性: 各 $y = (y_1, y_2, y_3, \dots)$ に対して, $x = (y_2 - y_1, y_1, y_2, y_3, \dots)$ と定めれば $x \in F$ であり $S(x) = y$ となる. なお, $x = S(y) - y$ である. 実際, これらの列の初項は等しい. また, 第 2 項以降について, 左辺の第 n 項は y_{n-1} , 右辺は $y_{n+1} - y_n$ であり, 数列 y が漸化式 $y_{n+1} = y_n + y_{n-1}$ ($n \geq 2$) を満たすことからこれらも等しい⁴⁵.

単射性: $\text{Ker } S = \{0\}$ を示せばよい. $x = (x_1, x_2, x_3, \dots) \in F$, $S(x) = 0$ とすれば (2) より x の第 2 項以降はすべて 0 であり, これに漸化式の条件 $x_3 = x_2 + x_1$ を合わせて $x_1 = 0$ を得る. つまり $x = 0$ である. \square

例 20.4.4. (1) I を开区間とし, I 上の実数値 C^∞ -級関数全体を $C^\infty(I)$ とする. $C^\infty(I)$ の元 f は何回でも微分できるゆえ, f の微分 f' もまた何回も微分できる, すなわち $f' \in C^\infty(I)$. そこで, $D: C^\infty(I) \rightarrow C^\infty(I)$ を $D(f) := f'$ と定めれば D は線形写像である.

⁴⁵実は, 初項と第 2 項が等しいことさえ分かれば, F の元は漸化式 $x_{n+2} = x_{n+1} + x_n$ を満たすゆえ第 3 項以降もすべて等しいことが分かる.

(2) D は全射である. 実際, 各 $f \in C^\infty(I)$ に対して f の原始関数を F とすれば, F を 1 回微分すると f になり, また f は何回でも微分できる. すなわち F も何回でも微分可能であり $F \in C^\infty(I)$. $D(F) = F' = f$ より D の全射性を得る. 一方, 微分すると $\mathbf{0} \in C^\infty(I)$ になる関数は定数関数しかないことから⁴⁶, $\text{Ker } D$ は定数関数全体のなす集合になる.

(3) 関数 $y \in C^\infty(I)$ が線形微分方程式 16.3.4:

$$y^{(k)}(x) = a_{k-1}y^{(k-1)} + a_{k-2}y^{(k-2)} + \cdots + a_1y^{(1)}(x) + a_0y^{(0)}(x)$$

を満たすならば, 上の両辺を微分することで関数 $D(y)$ もまた上の微分方程式を満たすことが分かる.

(4) 微分方程式 $y^{(2)} = y^{(1)} + y^{(0)}$ を満たす C^∞ -級関数全体 W は $C^\infty(I)$ の部分空間である (例 16.3.3). D を W に制限した $D : W \rightarrow W$ は全単射である.

Proof. 全射性: 各 $y \in W$ に対して $F := D(y) - y$ (つまり $F(x) = y'(x) - y(x)$) と定めれば W は部分空間であったから $F \in W$ である. このとき $D(F) = y$, すなわち $F'(x) = y(x)$ が成り立つ. 実際, y が微分方程式 $y^{(2)}(x) = y^{(1)}(x) + y(x)$ を満たすことから, $F'(x) = (y'(x) - y(x))' = y^{(2)}(x) - y^{(1)}(x) = y(x)$.

単射性: $\text{Ker } D = \{\mathbf{0}\}$ を示せばよい. $y \in W$, $D(y) = \mathbf{0}$ とすれば (2) より y は定数関数である. つまり $y' = \mathbf{0}$, $y'' = \mathbf{0}$ であり, これに微分方程式の条件 $y''(x) = y'(x) + y(x)$, すなわち $\mathbf{0} = \mathbf{0} + y$ を合わせて $y = \mathbf{0}$ を得る. \square

微分を用いた線形写像は微分作用素と呼ばれ, とくに上の写像 D は通常 $\frac{d}{dx}$ で表す.

例 20.4.5 (発展). $a \in \mathbb{R}$ を固定し, 写像 $T : C^\infty(\mathbb{R}) \rightarrow \mathbb{R}^\mathbb{N}$ を

$$T(f) := (f(a), f^{(1)}(a), f^{(2)}(a), \dots)$$

と定めれば, これは線形写像である. また, 例 20.4.3 におけるシフト作用素 S および例 20.4.4 における微分作用素 D について $S \circ T = T \circ D$ が成り立つ. これを図式で表すと次のようになる.

$$\begin{array}{ccccc} C^\infty(\mathbb{R}) & \ni & f & \xrightarrow{T} & (f(a), f^{(1)}(a), f^{(2)}(a), f^{(3)}(a), \dots) & \in & \mathbb{R}^\mathbb{N} \\ D \downarrow & & \downarrow & & \downarrow & & \downarrow S \\ C^\infty(\mathbb{R}) & \ni & f^{(1)} & \xrightarrow{T} & (f^{(1)}(a), f^{(2)}(a), f^{(3)}(a), f^{(4)}(a), \dots) & \in & \mathbb{R}^\mathbb{N} \end{array}$$

補足: 点 a の周りで冪級数展開 (テーラー展開) 可能であり, かつ収束半径が無大となる関数に限れば T は単射になり, 上の図式は微分作用素とシフト作用素がほぼ同等であることを示唆している.

例 20.4.6 (発展). I を \mathbb{R} 上の開区間とし, I 上の実数値連続関数全体のなす線形空間を $C(I)$ とする. また $a, b \in I$ を固定しておく.

(1) $S_{a,b} : C(I) \rightarrow \mathbb{R}$ を $S_{a,b}(f) := \int_a^b f(x)dx$ と定めれば $S_{a,b}$ は線形写像である. $\text{Ker } S_{a,b}$ の元は区間 $[a, b]$ における定積分が 0 となる関数である. $\text{Im } S_{a,b}$ は区間 $[a, b]$ が幅を持つかどうかで異なる. $a = b$ の場合, 区間 $[a, b]$ における積分はゼロゆえ, $\text{Im } S_{a,b} = \{0\}$. 言い換えると $\text{Ker } S_{a,b} = C(I)$ である. 一方 $a \neq b$ ならば積分は様々な値を取り得る. 実際, 各 $r \in \mathbb{R}$ に対して, 定数関数 $f(x) = \frac{r}{b-a}$ を取れば $S_{a,b}(f) = r$ である. すなわち, $S_{a,b}$ は全射であり, $\text{Im } S_{a,b} = \mathbb{R}$.

⁴⁶実際, 対偶「 f が定数関数でないならば $f' \neq \mathbf{0}$ 」が次のように示される: f が定数関数でないならば, $f(x) \neq f(y)$ をみたす異なる二点 $x, y \in I$ がある. $a = \frac{f(x)-f(y)}{x-y}$ とおけば ($x \neq y$ より分母は 0 でない), $f(x) \neq f(y)$ より $a \neq 0$ である. 平均値の定理より $f'(t) = a$ を満たす $t \in I$ が存在し, とくに f' はゼロ値定数関数 $\mathbf{0}$ ではない.

- (2) $C(I)$ の元 f に対して次で定義される $C(I)$ の元 $I_a(f) : I \rightarrow \mathbb{R}$ を対応させる写像 $I_a : C(I) \rightarrow C(I)$ を考える.

$$I_a(f)(x) := \int_a^x f(t) dt.$$

I_a は線形写像である. 微分積分学の基本定理より, $I_a(f)$ は f の原始関数のうちの一つである. $\text{Im } I_a$ は $F(a) = 0$ を満たす C^1 -級関数⁴⁷ F たち全体に一致する.

Proof. $V = \{F \in C(I) \mid F(a) = 0 \text{ かつ } F \text{ は } C^1\text{-級}\}$ とし, $\text{Im } I_a = V$ を示そう.

($\text{Im } I_a \subset V$): $\text{Im } I_a$ の各元 $F = I_a(f)$ が $F(a) = 0$ を満たすことは定義から直ちにわかる (区間 $[a, a]$ 上の定積分は 0). また, F の微分が連続関数 f になることから, F は C^1 -級である. ゆえに $F \in V$.

($V \subset \text{Im } I_a$): $F \in V$ とする. $f := F'$ とすれば, F は C^1 -級ゆえ $f \in C(I)$ である. $G := I_a(f)$ とおけば $G' = f$ である. ゆえに $(F - G)' = F' - G' = f - f = \mathbf{0}$. 微分が $\mathbf{0}$ になる関数は定数関数のみであるから⁴⁸, $F - G$ は定数関数であり, これに a を代入すると $(F - G)(a) = F(a) - G(a) = 0 - 0 = 0$. ゆえに関数 $F - G$ は 0 値定数関数であり, $F - G = \mathbf{0}$. 以上より $F = G = I_a(f) \in \text{Im } I_a$. \square

$I_a(f) = I_a(g)$ とすれば両辺を微分して $f = g$ となるゆえ I_a は単射であり $\text{Ker } I_a = \{\mathbf{0}\}$.

- (3) $\text{Im } I_a$ の各元は微分可能な関数であり, ゆえにそれらは微分作用素 $D = \frac{d}{dx}$ に代入することができる. 微分積分学の基本定理より $D \circ I_a = \text{id}_{C(I)}$ である.

例 20.4.7 (発展). X, Y を \mathbb{R} の部分集合とする (あるいはより一般に $X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m$ としてもよい). 連続写像 $F : X \rightarrow Y$ を一つ与えると, 次のような写像 $T_F : C(Y) \rightarrow C(X)$ が定義できる (写像の向きが F と逆になっていることに注意せよ).

$$T_F(f) := f \circ F, \quad (\text{ここで } f \in C(Y), \text{ つまり } f : Y \rightarrow \mathbb{R} \text{ は連続関数}).$$

- (1) T_F は線形写像である.

Proof. まず $T_F(f) \in C(X)$ であること, すなわち $f \circ F : X \rightarrow \mathbb{R}$ が連続であることを示さねばならないが, これは解析学の講義に譲り (一般に, 連続写像による合成は連続である), 線形性 (iii) のみ確認しよう. 各 $f, g \in C(Y)$ および $r, s \in \mathbb{R}$ について

$$T_F(rf + sg) = rT_F(f) + sT_F(g)$$

を示したい. これらは共に X を定義域とする写像であるから, 左右の両写像に各 $x \in X$ を代入した値が等しいことが分かれば上の等式を得る. $C(X)$ における和とスカラー倍の定義を思いだしながら計算すると,

$$\begin{aligned} T_F(rf + sg)(x) &= (rf + sg) \circ F(x) = (rf + sg)(F(x)) = (rf)(F(x)) + (sg)(F(x)) \\ &= r \cdot f(F(x)) + s \cdot g(F(x)). \end{aligned}$$

$$\begin{aligned} (rT_F(f) + sT_F(g))(x) &= (rT_F(f))(x) + (sT_F(g))(x) = (r(f \circ F))(x) + (s(g \circ F))(x) \\ &= r \cdot (f \circ F)(x) + s \cdot (g \circ F)(x) = r \cdot f(F(x)) + s \cdot g(F(x)). \end{aligned}$$

ゆえに $T_F(rf + sg) = rT_F(f) + sT_F(g)$ が成り立つ. \square

- (2) T_F は $C(X)$ の積演算とも整合的な写像である. すなわち, $T_F(fg) = T_F(f)T_F(g)$.

⁴⁷ 導関数が連続となる関数を C^1 -級関数という.

⁴⁸ 「 $g' = \mathbf{0}$ ならば g は定数関数」は先程示した「 g は定数関数でないならば $g' \neq \mathbf{0}$ 」の対偶にあたる.

Proof. 各 $x \in X$ を代入した値が一致することを示せばよい.

$$\begin{aligned} T_F(fg)(x) &= (fg) \circ F(x) = (fg)(F(x)) = f(F(x)) \cdot g(F(x)). \\ (T_F(f)T_F(g))(x) &= (T_F(f))(x) \cdot (T_F(g))(x) = (f \circ F)(x) \cdot (g \circ F)(x) = f(F(x)) \cdot g(F(x)). \end{aligned}$$

□

21 線形空間の同一視

認識した対象を区別する、あるいは分類するという行為は、その対象を理解するための基本的な手段である。分類(類別)とは、似たものをどうしを集め、また著しく異なっているものを分けることをいう。線形空間の分類の場合、何を基準とするのが妥当であろうか。本節では、まずはじめに、線形代数的な性質の相互翻訳が可能で 1 対 1 の対応(線形同型写像)について論じる。そして、二つの線形空間の間にこの対応があるかどうかを同一視の基準と定め、その定義の妥当性について考察する。

21.1 線形同型写像

次で与える対応によって、線形代数的な性質が相互に翻訳されることを見よう。

定義 21.1.1. 線形写像 $f: U \rightarrow V$ が全単射であるとき、これを線形同型 (linear isomorphism) あるいは単に同型という。

例 21.1.2. 単射線形写像 $f: U \rightarrow V$ において、終域を置き換えた写像 $f: U \rightarrow \text{Im } f$ は線形同型である。

例 21.1.3. U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$ とする。このとき、命題 20.1.8 より $f(\mathbf{u}_i) = \mathbf{e}_i$ を満たす線形写像 $f: U \rightarrow \mathbb{R}^n$ が存在する。この f は線形同型である。

Proof. 命題 20.2.3 より $\text{Im } f = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle = \langle \mathbf{e}_1, \dots, \mathbf{e}_n \rangle = \mathbb{R}^n$ 。すなわち f は全射である。また、単射性は練習 20.3.7 による。□

線形写像 f が全単射ならば次の命題により f^{-1} も自動的に線形写像となる。例 19.3.6(i) より f^{-1} は全単射であり、したがって f^{-1} も線形同型である。

命題 21.1.4. 線形同型 $f: U \rightarrow V$ の逆写像 $f^{-1}: V \rightarrow U$ は線形写像である。

Proof. f^{-1} が線形性 (iii) を満たすことを示そう。そこで $\mathbf{x}, \mathbf{y} \in V$ および $r, s \in \mathbb{R}$ を任意に取る。 $\mathbf{a} = f^{-1}(\mathbf{x})$, $\mathbf{b} = f^{-1}(\mathbf{y})$ とおく。逆写像の定義から $f(\mathbf{a}) = \mathbf{x}$, $f(\mathbf{b}) = \mathbf{y}$ である。また f の線形性より $f(r\mathbf{a} + s\mathbf{b}) = rf(\mathbf{a}) + sf(\mathbf{b}) = r\mathbf{x} + s\mathbf{y}$ である。この両辺をそれぞれ f^{-1} に代入することで

$$\begin{aligned} f^{-1}(f(r\mathbf{a} + s\mathbf{b})) &= f^{-1}(r\mathbf{x} + s\mathbf{y}) \\ r\mathbf{a} + s\mathbf{b} &= f^{-1}(r\mathbf{x} + s\mathbf{y}) \\ rf^{-1}(\mathbf{x}) + sf^{-1}(\mathbf{y}) &= f^{-1}(r\mathbf{x} + s\mathbf{y}). \end{aligned}$$

上の最後の式は f^{-1} が線形性 (iii) を満たすことを意味している。□

命題 20.3.6 から更に踏み込んで、線形同型では次が成立する。

命題 21.1.5. $f: U \rightarrow V$ が線形同型であるとき、 $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ について次が成り立つ。

$$\mathbf{u}_1, \dots, \mathbf{u}_n \text{ は } U \text{ の基底である} \iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \text{ は } V \text{ の基底である.}$$

Proof. (\Leftarrow) は命題 20.3.6(5) より得られているゆえ (\Rightarrow) のみを示せばよい。 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする。命題 20.3.6(3) より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形独立である。また、 f は全射ゆえ $\text{Im } f = V$ であり、命題 20.2.3 より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V を生成する。□

逆に、基底が対応し合う線形写像は同型である:

命題 21.1.6. U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$, $f: U \rightarrow V$ を線形写像とすると、次が成り立つ。

$$f \text{ は線形同型である} \iff f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \text{ は } V \text{ の基底である.}$$

Proof. (\Rightarrow) は命題 21.1.5 より明らか。(\Leftarrow) は例 21.1.3 における証明とほとんど同じ論法で得られる。□

命題 21.1.7. 線形同型の合成は線形同型である。

Proof. $f: U \rightarrow V$ および $g: V \rightarrow W$ を共に線形同型とすれば命題 20.1.9 より $g \circ f$ は線形写像であり、また命題 19.4.1(3) より全単射である。□

21.2 同型な線形空間

二つの線形空間 U, V が本質的に同じであるとはどういうことか考えよう。それは、線形空間の枠組みにおいてそれらを区別できないことと定めるのが妥当である。すなわち、線形空間上のあらゆる命題において U と V における真偽が一致することに他ならない。言い換えれば、 U で成り立つことと同等の現象が必ず V においても成り立ち、またその逆も言えるということである。線形空間の定義に立ち戻れば、それは和とスカラー倍の演算の性質のみを規定した対象であったから、線形空間の現象として述べるのできるのは、それらの演算に関する言明のみである。したがって、 U と V の間に和とスカラー倍の演算が保存されるような (つまり演算について整合的な) 1 対 1 対応があるとき、線形空間の枠組みで語れる現象のみを用いて U と V を区別することはできないことになる。以上の考察から、次の線形同型なる概念 (本質的に同じものとみなすこと) を得る。

定義 21.2.1. 線形空間 U, V の間に線形同型写像 $f: U \rightarrow V$ が存在するとき、 U と V は線形同型である (**linearly isomorphic**) あるいは単に同型であるという。このとき、 $U \simeq V$ と書く。

例 21.2.2. 2 次行ベクトル全体 $U = M_{1,2}(\mathbb{R})$ と 2 次列ベクトル全体 $V = M_{2,1}(\mathbb{R})$ について考えよう。前者は横に成分を並べたベクトルの集合であり、後者は縦に成分を並べたそれである。ゆえにこれらは見た目の上では異なっているとも考えられる。しかし、線形空間的な性質 (つまり演算に関する性質) における違いのみを用いてこれらを区別することはできない。実際、写像 $f: U \rightarrow V$ を $f(\mathbf{x}) := {}^t \mathbf{x}$ と定めればこれは線形同型であり、 U において $\mathbf{u}_1, \dots, \mathbf{u}_n \in U$ に関する線形空間的な性質 $P(\mathbf{u}_1, \dots, \mathbf{u}_n)$ が成立するとき、 V において $P(f(\mathbf{x}_1), \dots, f(\mathbf{x}_n))$ が成立することが示唆される⁴⁹。すなわち、 U で成り立ち V で成り立たないような線形空間の枠組みにおける現象を挙げることはできない。

いま、線形空間を同一視する基準を与えた。ここで細かい理屈をこねれば、 U 自身は U と同一視できるか、あるいは U と同型な線形空間と同型な線形空間は U と同型か、といった素朴な疑問が生じよう。学問上のあらゆる理論は、こうした些細な疑問にも答えられるよう構築されねばならない。

命題 21.2.3. 線形空間 U, V, W について次が成り立つ。

(反射律) $U \simeq U$, (対称律) $U \simeq V \implies V \simeq U$, (推移律) $U \simeq V, V \simeq W \implies U \simeq W$.

Proof. (反射律): 恒等写像 $\text{id}_U: U \rightarrow U$ は線形同型ゆえ $U \simeq U$ である。(対称律): $U \simeq V$ とすれば線形同型 $f: U \rightarrow V$ が存在し、その逆写像 $f^{-1}: V \rightarrow U$ も線形同型であることから $V \simeq U$ を得る。(推移律): $U \simeq V, V \simeq W$ とすれば線形同型 $f: U \rightarrow V, g: V \rightarrow W$ が存在し、これらの合成 $g \circ f: U \rightarrow W$ が線形同型であることから $U \simeq W$ を得る。□

よりみち (同一視すること)。

一般に、二つの対象を結ぶ記号 (集合論では、これを関係と呼ばれる概念を用いて定義する) が反射律および対称律、推移律を満たすとき、これを同値関係 (**equivalence relation**) と言う。線形代数学に限らず何らかの立場で二つの対象を同一のものとみなすとき、その同一性は同値関係になることが望ましい。これを認めるならば、考えるべき対象に同値関係を与えるということは、同一視する基準を与えることと言い換えてもよい。

同値関係となるような概念はこれまでもいくつか学んでいる。

例 21.2.4. (1) 平面または空間上の図形 A, B が合同であるとき、 $A \equiv B$ と書く。また、図形 A, B が相似であるとき $A \sim B$ と書く (記号 \sim は国際標準ではない)。合同 \equiv および相似 \sim はそれぞれ同値関係である。

(2) 二つの集合 A, B の間に全単射が存在するとき A と B は対等であるというのであった。対等は同値関係である。反射律は例 19.2.2 により、対称律は例 19.3.6(i) に相当する。また、推移律は命題 19.5.4 で述べた。

⁴⁹ここでは「示唆される」と述べるに留め、証明は行わない。任意の性質 P についてこの事実を示すには、数理論理的な枠組みにおいて「線形空間的な性質 (論理式)」を再帰的に定義し、論理式の長さに関する帰納法を用いればよい。

同型な線形空間の例を挙げよう.

例 21.2.5. (1) 線形空間 U および V において同じ個数からなる基底が取れるとき, U と V は線形同型である. 実際, U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$, V の基底を $\mathbf{v}_1, \dots, \mathbf{v}_n$ とすれば, 命題 20.1.8 により $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす線形写像 $f: U \rightarrow V$ が取れる. この写像は命題 21.1.6 より線形同型である.

(2) $M_{m,n}(\mathbb{R})$ と \mathbb{R}^{mn} は共に mn 個のベクトルからなる基底を持つ. ゆえにこれらは線形同型である.

例 21.2.6. 例 15.2.2 および例 15.2.7(2) において, 異なる集合に同一の記号 $\mathbb{R}^{\mathbb{N}}$ を与えていた. 前者の数列空間を U とし, 後者の写像空間を V としよう. このとき, 例 15.2.7(2) で与えた次の対応:

$$f: U \rightarrow V, \quad f(x_1, x_2, \dots) := "x(n) := x_n \text{ で定義される写像 } x: \mathbb{N} \rightarrow \mathbb{R}"$$

は線形同型となる. したがって, 線形空間の枠組みにおいてこれらは同一のものと見なすことができ, それゆえ, これら二つの空間に同じ記号 $\mathbb{R}^{\mathbb{N}}$ を与えている.

$M_{m,n}(\mathbb{R})$ と \mathbb{R}^{mn} は線形同型であるにもかかわらず同じ記号を用いることはない. これは, 線形空間としての構造だけではなく, 積の構造も加味した代数構造として $M_{m,n}(\mathbb{R})$ を扱うためである. $M_{m,n}(\mathbb{R})$ の元と $M_{n,r}(\mathbb{R})$ の元の間に行列としての積が定義されている一方で, \mathbb{R}^{mn} の元と \mathbb{R}^{nr} の元の間には自然な積が定まらない. ゆえに, これらを同一の記号で表すのは好ましくない.

補足(発展): ちなみに \mathbb{R}^m の元どうしの積を座標ごとの掛け算によって定めることができる. しかしながら $m = nn$ におけるこの積演算と $M_n(\mathbb{R})$ の積演算(行列としての積)の間に自然な対応は与えられない. 前者は可換である, すなわち常に $\mathbf{xy} = \mathbf{yx}$ が成り立つのに対し, 後者はそうではないからである. なお, \mathbb{R}^m と, m 点集合 $X_m = \{1, \dots, m\}$ 上の関数(連続関数)全体 $C(X_m) = \mathbb{R}^{X_m}$ の間には積演算を込めた意味での(すなわち多元環としての)自然な 1 対 1 対応がつく. 実際, 有限数列(第 m 項までの数列)について例 21.2.6 と同様の対応を考えれば, $C(X_m)$ との間には積演算についても整合的な 1 対 1 対応が得られる.

次は線形同型でない例である.

命題 21.2.7. 自然数 $m < n$ について \mathbb{R}^m と \mathbb{R}^n は線形同型でない.

Proof. もし線形同型 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ が存在するとすれば, 定義域 \mathbb{R}^n の標準ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_n$ について命題 21.1.5 (より根本的には命題 20.3.6(3)) より $f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)$ は \mathbb{R}^m における線形独立な組となり, これは命題 17.3.5 に矛盾する. ゆえに \mathbb{R}^m と \mathbb{R}^n の間に線形同型写像は存在せず, \mathbb{R}^m と \mathbb{R}^n は線形同型でない. \square

命題 21.2.7 の証明で用いた f の性質は単射性のみである. つまり, 自然数 $m < n$ について線形単射 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ は存在しない. 一方, 逆向きの線形全射 $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ の非存在性も示すことができる(系 22.4.6).

21.3 線形写像のなす空間

1 節で予告したように, 行列表とは線形写像の数値化にほかならない. すなわち, 線形写像のなす空間と行列のなす空間は同一視できる. 1.6 項において \mathbb{R}^2 の間で定まる線形写像について紹介した事実の一般化を本項で述べよう. そのためには, (m, n) -行列表全体 $M_{m,n}(\mathbb{R})$ に対応する線形写像のなす集合に記号を与えておかねばならない. $M_{m,n}(\mathbb{R})$ には線形空間の構造が入っていたゆえ, 対応する線形写像の集合も線形空間になることが示唆される.

定義 21.3.1. 線形空間 U から線形空間 V への線形写像全体のなす集合を $\text{Hom}(U, V)$ と書く. とくに $U = V$ のとき, これを $\text{End}(U)$ と書く. $\text{Hom}(U, V)$ には次のように和とスカラー倍が定まり, 線形空間となる:

- $f, g \in \text{Hom}(U, V)$ に対して $(f + g) : U \rightarrow V$ を次の写像として定める:
各 $\mathbf{x} \in U$ について, $(f + g)(\mathbf{x}) := f(\mathbf{x}) + g(\mathbf{x})$.
- $f \in \text{Hom}(U, V)$ および $r \in \mathbb{R}$ に対して $(rf) : U \rightarrow V$ を次の写像として定める:
各 $\mathbf{x} \in U$ について, $(rf)(\mathbf{x}) := rf(\mathbf{x})$.

$\text{Hom}(U, V)$ における零ベクトルは, U の各元に対して $\mathbf{0}_V$ を対応させる定値写像 $f = \mathbf{0}_V$ である.

記号 Hom は準同型写像 (homomorphism) に由来する. 線形空間に限らず, 与えられた代数構造に関する演算と相性のよい写像, すなわち演算との合成が交換可能 (可換) になる写像のことを準同型と言う. また, ある代数構造からそれ自身への準同型のことを自己準同型 (endomorphism) と言う. 線形写像は線形空間に関する準同型であり, とくに線形空間上の自己準同型は線形変換と呼ばれる.

練習 21.3.2. (1) 上で定めた和とスカラー倍において, 線形写像の和が線形写像になること, および線形写像のスカラー倍が線形写像になることを示せ. すなわち, $f, g \in \text{Hom}(U, V)$, $r \in \mathbb{R}$ ならば $f + g, rf \in \text{Hom}(U, V)$ ということである.

解答例: $f + g$ が線形写像となることを示そう. 線形性 (iii) は次のように確認できる:

$$\begin{aligned} (f + g)(s\mathbf{x} + t\mathbf{y}) &= f(s\mathbf{x} + t\mathbf{y}) + g(s\mathbf{x} + t\mathbf{y}) = (sf(\mathbf{x}) + tf(\mathbf{y})) + (sg(\mathbf{x}) + tg(\mathbf{y})) \\ &= s(f(\mathbf{x}) + g(\mathbf{x})) + t(f(\mathbf{y}) + g(\mathbf{y})) = s(f + g)(\mathbf{x}) + t(f + g)(\mathbf{y}). \end{aligned}$$

□

(2) $\text{Hom}(U, V)$ における和とスカラー倍の演算がベクトル空間の公理を満たすことを確かめよ.

(m, n) -行列 A において, 線形写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ を $T_A(\mathbf{x}) := A\mathbf{x}$ と定めるのであった. 行列における演算と T_A との関係についてまとめておこう.

命題 21.3.3. (1) サイズの等しい行列 A, B について $T_A = T_B \iff A = B$.

(2) サイズの等しい行列 A, B および $r, s \in \mathbb{R}$ について $T_{rA+sB} = rT_A + sT_B$.

(3) 行列 B と A の間に積が定まるとき, $T_{BA} = T_B \circ T_A$.

(4) A が可逆 $\iff T_A$ は全単射. また, このとき, $(T_A)^{-1} = T_{A^{-1}}$.

Proof. (1): (\Leftarrow) は明らかゆえ (\Rightarrow) のみ示す. A, B の各列が等しいことが示されれば $A = B$ である. $T_A(\mathbf{e}_i)$ および $T_B(\mathbf{e}_i)$ はそれぞれ A, B の i 列目であり, 仮定より $T_A(\mathbf{e}_i) = T_B(\mathbf{e}_i)$ である. ゆえに A, B の各列は等しい.

(2): $T_{rA+sB}(\mathbf{x}) = (rA + sB)(\mathbf{x}) = r(A\mathbf{x}) + s(B\mathbf{x}) = rT_A(\mathbf{x}) + sT_B(\mathbf{x})$.

(3): $T_{BA}(\mathbf{x}) = (BA)\mathbf{x} = B(A\mathbf{x}) = BT_A(\mathbf{x}) = T_B(T_A(\mathbf{x})) = T_B \circ T_A(\mathbf{x})$.

(4): A を可逆とすれば, (3) より $T_A \circ T_{A^{-1}} = T_E = \text{id}_{\mathbb{R}^n}$, $T_{A^{-1}} \circ T_A = \text{id}_{\mathbb{R}^m}$ である. したがって命題 19.3.4 より T_A は全単射であり, その逆写像は $T_{A^{-1}}$ である. 逆に T_A が全単射であると仮定し, A の可逆性を示そう. 線形写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ が同型であることから命題 21.2.7 より $m = n$ ゆえ A は正方行列である. また, 標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ について命題 20.3.6(3) より $T_A(\mathbf{e}_1), \dots, T_A(\mathbf{e}_n)$ は線形独立である. すなわち, 正方行列 A の各列ベクトルによる組は線形独立であり, 定理 17.3.6 より A は可逆である. □

定理 21.3.4. 任意の線形写像 $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ は (m, n) -行列 A_f を用いて $f = T_{A_f}$ と一意的に表せる. 更に, $\mathcal{T} : M_{m,n}(\mathbb{R}) \rightarrow \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ を $(A) := T_A$ と定めれば, これは線形同型写像である.

Proof. $A_f := [f(\mathbf{e}_1), \dots, f(\mathbf{e}_n)]$ と定めれば $f = T_{A_f}$ である. 実際, $T_{A_f}(\mathbf{e}_i) = \text{“}A_f \text{の} i \text{列目”} = f(\mathbf{e}_i)$ ゆえ T_{A_f} と f における標準基底の行き先は等しい. したがって命題 20.1.8 より $T_{A_f} = f$ である.

いま示したことは, \mathcal{T} の全射性にほかならない. また, \mathcal{T} の単射性は命題 21.3.3(1) による. \mathcal{T} の線形性 (iii) は命題 21.3.3(2) による. □

上の定理における \mathcal{T} の逆写像 $\mathcal{S} : \text{Hom}(\mathbb{R}^n, \mathbb{R}^m) \rightarrow M_{m,n}(\mathbb{R})$ は, $f \in \text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ に対して $A_f = [f(e_1), \dots, f(e_n)]$ を対応させる写像である.

備考 21.3.5 (発展). $\text{End}(U)$ の各元 f, g に対して, これらの積 fg を合成 $f \circ g$ によって定める. すると, 既に定めていた和とスカラー倍および今の積演算は, $I = \text{id}_U$ を単位元として命題 2.4.1 における性質をすべて満たす (練習 21.3.6). すなわち, $\text{End}(U)$ は多元環となる. また, 前定理において $m = n$ とするとき, 命題 21.3.3(3) により, $T : M_m(\mathbb{R}) \rightarrow \text{End}(\mathbb{R}^m)$ は積の演算についても整合的な 1 対 1 対応になる. すなわち, $M_m(\mathbb{R})$ と $\text{End}(\mathbb{R}^m)$ は多元環として同一視できる.

練習 21.3.6. $F, G, H \in \text{End}(U)$ および $r, s \in \mathbb{R}$ について次を確認せよ:

$$(13) (rF)(sG) = (rs)(FG), \quad (16) F(G + H) = FG + FH, \quad (17) (F + G)H = FH + GH.$$

注意: 積に関する結合律: (18) $(FG)H = F(GH)$ は命題 3.2.5 において, より一般的な立場から既に示している.

解答例: 積演算が写像の合成を意味していることに注意し, 各 $\mathbf{u} \in U$ を代入した値が一致することを確認すればよい. (16) のみ示そう.

$$\begin{aligned} F(G + H)(\mathbf{u}) &= F \circ (G + H)(\mathbf{u}) = F((G + H)(\mathbf{u})) = F(G(\mathbf{u}) + H(\mathbf{u})) \\ &= F(G(\mathbf{u})) + F(H(\mathbf{u})) = F \circ G(\mathbf{u}) + F \circ H(\mathbf{u}) = (F \circ G + F \circ H)(\mathbf{u}) = (FG + FH)(\mathbf{u}) \\ & \quad (\uparrow F \text{ の線形性を用いた}) \end{aligned}$$

以下, id_U を $\text{End}(U)$ における単位元とみなすとき, これを I と書く. また, $F \in \text{End}(U)$ の n 回の合成を F^n と書く:

$$F^n := \underbrace{F \circ F \cdots \circ F}_{n \text{ 個の積}}.$$

計算に現れる式を簡略にするため便宜上 $F^0 = I$ とすることもある.

21.4 線形変換と多項式 (発展)

前項において $\text{End}(U)$ が多元環の構造を持つこと, すなわち, 線形変換たちの間で正方行列と類似の演算ができることを見た. ここで, 写像あるいは行列による演算が複雑化したときに有効な記法を導入しよう.

定義 21.4.1. 実数係数の n 次多項式 $\Phi(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$ および線形変換 $F : U \rightarrow U$ に対して, $\Phi(t)$ における変数 t を F に置き換え, また定数項 a_0 を $a_0 I$ に置き換えると次のような線形変換が得られる:

$$a_n F^n + a_{n-1} F^{n-1} + \cdots + a_1 F + a_0 I.$$

これを略して $\Phi(F)$ と書く. また正方行列 A に対しても同様に, $\Phi(A)$ を次の正方行列として定める:

$$a_n A^n + a_{n-1} A^{n-1} + \cdots + a_1 A + a_0 E.$$

注意: この定義はより一般に, 多元環の元に対して定義される概念である.

上の定義はあくまで形式的なものであって, $\Phi(t)$ に F や A を代入しているわけではないことに注意せよ.

例 21.4.2. $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ を $D = \frac{d}{dx}$ と定めれば, $y \in C^\infty(\mathbb{R})$ に関する条件としての次の三つの式はすべて同値である.

$$\begin{aligned} y^{(k)}(x) + a_{k-1} y^{(k-1)}(x) + a_{k-2} y^{(k-2)}(x) + \cdots + a_1 y^{(1)}(x) + a_0 y^{(0)}(x) &= 0, \\ D^k(y) + a_{k-1} D^{k-1}(y) + a_{k-2} D^{k-2}(y) + \cdots + a_1 D(y) + a_0 I(y) &= \mathbf{0}, \\ (D^k + a_{k-1} D^{k-1} + a_{k-2} D^{k-2} + \cdots + a_1 + a_0 I)(y) &= \mathbf{0}. \end{aligned}$$

多項式 $\Phi(t)$ を $\Phi(t) := t^n + a_{n-1}t^{n-1} + \cdots + a_1t + a_0$ と定めれば, 上の式は $\Phi(D)(y) = 0$ と略記できる. この $\Phi(t)$ を上の常微分方程式の特性多項式と呼ぶ.

練習 21.4.3 (ケーリー・ハミルトンの定理). $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ について, $\Phi_A(t) = t^2 - (a+b)t + (ad-bc)$ とおくと $\Phi(A) = O$ となることを確かめよ.

多項式は, $\Phi(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$ のように展開した形以外にも, 一部の項を括弧でくくったり, 因数分解したりと無数の表示を持つ. すると, 多項式を変形してから変数 t を線形変換 F に置き換えても問題がないかという疑念が浮かぶかもしれない. 例えば, $\Phi(t) = (t - \lambda)^n$ と因数分解される場合, 定義 21.4.1 による線形変換 $\Phi(F)$ (すなわち展開した式について t を F に置き換えたもの) と $(F - \lambda I)^n$ は等しいだろうか. これは $(t - \lambda)^n$ と $(F - \lambda I)^n$ をそれぞれ二項定理を用いて実際に展開することで確かめられるだろう. より一般に, 多項式がどのような形をしていても, t を F に置き換えた写像は定義 21.4.1 のそれと一致する. この事実を次の練習において確認しよう. いまから示すことは, 一般の多元環について成り立つ事実であり, とくに正方行列 A についても成り立つ.

練習 21.4.4. 線形変換 $F : U \rightarrow U$ および多項式 $\Phi, \Phi_1, \Phi_2, \Psi, \Psi_k$ について次を示せ.

(1) 実数 λ について, $\Phi(t) = \lambda t^k \Psi(t) \implies \Phi(F) = \lambda F^k \Psi(F)$.

解答例: $\Psi(t) = \sum_{i=0}^n a_i t^i$ とおけば, $\Phi(t) = \sum_{i=0}^n a_i \lambda t^{i+k}$ であるから,

$$\begin{aligned} \Phi(F) &= \text{“多項式 } a_n \lambda t^{i+n} + \cdots + a_0 \lambda t^k \text{ の } t \text{ を } F \text{ に置き換えた写像”} \\ &= a_n \lambda F^{i+n} + \cdots + a_0 \lambda F^k = \sum_{i=0}^n \lambda F^k a_n F^i = \lambda F^k \sum_{i=0}^n a_n F^i = \lambda F^k \Psi(F). \end{aligned}$$

(2) $\Phi(t) = \sum_{k=1}^r \Psi_k(t) \implies \Phi(F) = \sum_{k=0}^r \Psi_k(F)$.

解答例: $\Psi_k(t) = \sum_{i=0}^n a_{k,i} t^i$ ($k = 1, \dots, r$) とおけば, 例 3.2.4 により

$$\Phi(t) = \sum_{k=1}^r \left(\sum_{i=0}^n a_{k,i} t^i \right) = \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} t^i \right) = \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} \right) t^i$$

となるから,

$$\begin{aligned} \Phi(F) &= \text{“多項式 } \left(\sum_{k=1}^r a_{k,n} \right) t^n + \cdots + \left(\sum_{k=1}^r a_{k,1} \right) t + \left(\sum_{k=1}^r a_{k,0} \right) \text{ の } t \text{ を } F \text{ に置き換えた写像”} \\ &= \left(\sum_{k=1}^r a_{k,n} \right) F^n + \cdots + \left(\sum_{k=1}^r a_{k,1} \right) F + \left(\sum_{k=1}^r a_{k,0} \right) I \\ &= \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} \right) F^i = \sum_{i=0}^n \left(\sum_{k=1}^r a_{k,i} F^i \right) = \sum_{k=1}^r \left(\sum_{i=0}^n a_{k,i} F^i \right) = \sum_{k=1}^r \Psi_k(F). \end{aligned}$$

(3) $\Phi(t) = \Phi_1(t)\Phi_2(t) \implies \Phi(F) = \Phi_1(F)\Phi_2(F)$

解答例: $\Phi_1(t) = \sum_{i=0}^n a_i t^i$ とおけば, $\Phi(t) = \sum_{i=0}^n a_i t^i \Phi_2(t)$ である. ここで $\Psi_i(t) = a_i t^i \Phi_2(t)$ ($i = 0, \dots, n$) とおけば, (1) より $\Psi_i(F) = a_i F^i \Phi_2(F)$ が成り立つ. また, $\Phi(t) = \sum_{i=0}^n \Psi_i(t)$ であるから (2) より $\Phi(F) = \sum_{i=0}^n \Psi_i(F)$ である. 以上より,

$$\Phi(F) = \sum_{i=0}^n \Psi_i(F) = \sum_{i=0}^n a_i F^i \Phi_2(F) = \left(\sum_{i=0}^n a_i F^i \right) \Phi_2(F) = \Phi_1(F) \Phi_2(F).$$

上の性質を有限回適用することで、多項式 $\Phi(t)$ の表し方によらず、 $\Phi(t)$ における t を F に置き換えた写像は一致することが分かる。とくに $\Phi(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_k)^{n_k}$ と因数分解されるとき、

$$\Phi(F) = (F - \lambda_1 I)^{n_1}(F - \lambda_2 I)^{n_2} \cdots (F - \lambda_k I)^{n_k}$$

である。

命題 21.4.5. $F : U \rightarrow U$ を線形変換とし、 $\Phi(t)$ と $\Psi(t)$ を多項式とする。このとき、二つの線形変換 $\Phi(F)$ および $\Psi(F)$ は可換である。すなわち

$$\text{各 } \mathbf{u} \in U \text{ について, } \quad \Phi(F) \circ \Psi(F)(\mathbf{u}) = \Psi(F) \circ \Phi(F)(\mathbf{u}).$$

Proof. $\Theta(t) = \Phi(t)\Psi(t)$ とおく。すると $\Theta(F) = \Phi(F)\Psi(F)$ である。一方、 $\Theta(t) = \Psi(t)\Phi(t)$ であるから $\Theta(F) = \Psi(F)\Phi(F)$ でもある。以上より、 $\Phi(F)\Psi(F) = \Theta(F) = \Psi(F)\Phi(F)$ 。□

命題 21.4.6. $\mathcal{S} : \text{End}(U) \rightarrow M_n(\mathbb{R})$ が線形写像であり、さらに $\mathcal{S}(I) = E$ および $\mathcal{S}(FG) = \mathcal{S}(F)\mathcal{S}(G)$ をみたすならば⁵⁰、各多項式 $\Psi(t)$ について $\mathcal{S}(\Psi(F)) = \Psi(\mathcal{S}(F))$ が成り立つ。

Proof. $\Psi(t) = a_n t^n + a_{n-1} t^{n-1} + \cdots + a_1 t + a_0$ とおく。仮定より $\mathcal{S}(F^n) = \mathcal{S}(F)^n$ が成り立つことに注意すれば、

$$\begin{aligned} \mathcal{S}(\Psi(A)) &= \mathcal{S}(a_n F^n + a_{n-1} F^{n-1} + \cdots + a_1 F + a_0 I) \\ &= a_n \mathcal{S}(F^n) + a_{n-1} \mathcal{S}(F^{n-1}) + \cdots + a_1 \mathcal{S}(F) + a_0 \mathcal{S}(I) \\ &= a_n \mathcal{S}(F)^n + a_{n-1} \mathcal{S}(F)^{n-1} + \cdots + a_1 \mathcal{S}(F) + a_0 E = \Psi(\mathcal{S}(F)). \end{aligned}$$

□

上の命題において定義域や終域を $\text{End}(U)$ および $M_n(\mathbb{R})$ に限定する必要はなく、より一般に、多元環の間の写像に関する主張として上の命題は成立する。

例 21.4.7. 前項で与えた $\mathcal{T} : M_n(\mathbb{R}) \rightarrow \text{End}(U)$ についても前命題と同様の主張が成り立つ。したがって、多項式 $\Psi(t)$ および n 次正方行列 A について、 $T_{\Psi(A)} = \Psi(T_A)$ 。

⁵⁰このような写像を多元環に関する準同型という

22 線形空間の次元

前節において、線形空間を類別する手段、すなわち線形同型なる概念を与えた。本節では線形空間が同型であるかどうかをはかる指標として次元とよばれる量を導入する。これにより、線形空間の完全な分類を得る。

22.1 次元の定義

二つの線形空間 U と V が同型でないとは、線形同型写像 $f: U \rightarrow V$ が存在しないということである。つまり、 U と V が同型でないことを示すには非存在証明を行わねばならない。一般的な経験則として、存在証明に比べて非存在証明は難しい。非存在証明が難しいのは、長時間探し続けたが見つからなかったというだけでは証明ならず、どんなに才能あふれる者が探しても見つからないことを示さねばならない点にある。この事情と関連して、多くの非存在証明は背理法によってなされている。命題 21.2.7 においてもそうであった。

さて、命題 21.2.7 における非存在証明の鍵となる概念が何であったか振り返ろう。それは \mathbb{R}^m と \mathbb{R}^n では座標軸の数、すなわち基底におけるベクトルの個数が異なることが本質をついていた。基底におけるベクトルの個数は線形空間たちが同型であるか否かを決定する重要な量であり、この量を次元と呼ぶ。次の命題を前提として次元の定義を得る：

命題 22.1.1. 線形空間 V の基底におけるベクトルの個数は、基底の取り方によらずに一定である。

Proof. 二つのベクトルの組 $\mathbf{u}_1, \dots, \mathbf{u}_m$ および $\mathbf{v}_1, \dots, \mathbf{v}_n$ がそれぞれ V の基底であることを仮定し、 $m = n$ を示せばよい。基底 $\mathbf{u}_1, \dots, \mathbf{u}_m$ に関して例 21.2.5(1) を適用すれば V と \mathbb{R}^m は同型となる。また、基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に対しても同様の考察を行い V と \mathbb{R}^n は同型となる。すなわち、 $V \simeq \mathbb{R}^m$ 、 $V \simeq \mathbb{R}^n$ であり、これに命題 21.2.3(推移律) を合わせて $\mathbb{R}^m \simeq \mathbb{R}^n$ を得る。命題 21.2.7 より $\mathbb{R}^m \simeq \mathbb{R}^n$ となるのは $m = n$ のときに限る。□

定義 22.1.2. 線形空間 V の有限個のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$ が V の基底となるとき、この基底の個数 n を V の次元 (**dimension**) と呼び $\dim V$ で表す。また、基底となる有限個のベクトルの組を持つ線形空間のことを有限次元であるといい $\dim V < \infty$ と書く。そのような基底を持たない線形空間を無限次元であるといい $\dim V = \infty$ と書く。

仮に命題 22.1.1 が認められなければ、一つの線形空間 V が m 個の元からなる基底と n 個の元からなる基底を持つ可能性があり ($m \neq n$)、これでは V の次元を定められない。次元の定義に命題 22.1.1 が必要なのはこのためである。行列の階数や置換の符号の定義の際にも類似の議論を行ったのを覚えているだろうか。

命題 18.3.1 により、線形空間が有限個のベクトルで生成されることと有限次元であることは同値である。

例 22.1.3. (1) $\dim \mathbb{R}^n = n$, $\dim M_{m,n}(\mathbb{R}) = mn$, $\dim \mathbb{R}[x]_n = n + 1$.

(2) (m, n) -行列 $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ について、例 20.2.4 および 20.3.2 において $\text{Im } T_A = \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$, $\text{Ker } T_A = W_A$ となることを見た。また、これらの基底におけるベクトルの個数は命題 18.3.4 によって得られている。すなわち、 $\dim \text{Im } T_A = \dim \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle = \text{rank } A$, $\dim \text{Ker } T_A = \dim W_A = n - \text{rank } A$ である。この事実から一般の線形写像 f においても $\dim \text{Im } f$ のことを線形写像 f の階数と呼び $\text{rank } f$ と書く。また、 $\dim \text{Ker } f$ にも特別な記号が与えられており、これを $\text{null } f$ と書き、 f の退化次数と呼ぶ。あまり概念を増やして情報が錯綜しても困るゆえ、本論ではこれらの語句の使用は控える。

(3) 例 18.4.1 より $\mathbb{R}[x]$ は有限個のベクトルで生成されることはない。したがって無限次元である。

(4) $\mathbb{R}^{\mathbb{N}}$ は無限次元である。この事実は次項で示す (例 22.3.5)。

次元を比べるだけで線形空間どうしが同型かどうかを判定できる. とくに次の定理の (3) は, 線形同型写像 $f: U \rightarrow V$ の非存在性を示すための手段となる.

定理 22.1.4. 線形空間 U, V において次が成り立つ.

(1) U が有限次元であるとき, $U \simeq V \iff \dim U = \dim V$.

(2) $U \simeq V \implies \dim U = \dim V$.

(3) $\dim U \neq \dim V \implies U \not\simeq V$.

Proof. (1): (\Leftarrow) は例 21.2.5(1) より得られているゆえ (\Rightarrow) を示せばよい. $f: U \rightarrow V$ を同型とする. いま U は有限次元であると仮定していたから, 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が取れる (つまり $\dim U = n$). 命題 21.1.5 より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は V の基底であり, $\dim V = n$. 以上より $\dim U = \dim V$.

(2): U が有限次元の場合は (1) により示されている. また V が有限次元の場合は U と V の立場を入れ替えることで (1) に帰着できる. U, V いずれも無限次元の場合は $\dim U = \dim V = \infty$ である.

(3): (2) の対偶である. □

発展 (不変量).

二つの対象 X, Y が本質的に異なることを理解するには, それらに関する特別な量 $i(X), i(Y)$ が異なることを見ればよいのではないか. このような視点で与えられる量 i のことを一般に**不変量 (invariant)** という. 次元は線形空間に関する不変量である (定理 22.1.4(3)). 数学の多くの場面において, 対象としている数学的構造を保つ 1 対 1 対応の非存在性を示す際に不変量は鍵となる役割を担う. 例えば位相幾何学 (トポロジー) では, 点列の収束・発散を保つ 1 対 1 対応 (同相写像) があることを二つの空間が同一視できることの基準としている. 線形空間における同一視を同型と呼んだのに対して, 位相幾何における空間の同一視を同相という. 同相は同値関係である. 位相幾何的な立場でも次元という量が定義され (ルベークの被覆次元), 被覆次元の異なる空間は同相ではない. すなわち, 被覆次元は位相幾何における不変量である. とくに \mathbb{R}^n の被覆次元が n であることが分かっており, このことから $m \neq n$ ならば \mathbb{R}^m と \mathbb{R}^n は同相ではないことが示される.

ところで, 通常の不変量は $i(X) = i(Y)$ だからといって, X と Y が同一視できるかどうかは分からない. 例えば曲面の被覆次元は当然すべて 2 であり, 被覆次元は曲面の分類には適さない. このことは, 状況に応じて臨機応変に不変量を使い分ける必要があることを示唆している. 一方で, $i(X) = i(Y)$ から X, Y の同一視が導けるような強力な不変量のことを**完全不変量**という. 例えば, 有限次元の線形空間において次元は完全不変量である (定理 22.1.4(1)). 一般に, 完全不変量の一致は同一視の仕方についての定義の変形に過ぎないことが多く, 完全不変量から何かが分かるという可能性は低い. 言い換えると, いま考えている立場において不要な情報を捨てて有益な情報のみを残す不変量が望まれるのであり, このような不変量は情報を捨てたことから完全不変量にはなりえない. 結局, どんな量が必要とされるかを見極める力が求められている. こうした考え方は数学に限らず諸科学の分析において共通に認められるものである.

ちなみに, 曲面を分類する際の有効な不変量として, オイラー標数やベッチ数, ホモロジー群などが知られている. 最後に群が登場したように不変量を数に限る必要はなく, 様々な数学的対象をその候補に挙げることができる.

22.2 連立 1 次方程式の任意定数の個数

4.5 項にて予告していた連立 1 次方程式の一般解の表示に現れる任意定数の個数について再考する. 方程式 $A\mathbf{x} = \mathbf{b}$ の解全体の集合 W は, 解が存在するとすれば斉次方程式 $A\mathbf{x} = \mathbf{0}$ の解空間 W_A を平行移動した集合であった. したがって, 連立 1 次方程式の任意定数の個数について論ずるのであれば斉次方程式

についてのみ考察すれば十分である. いま, 方程式 $Ax = \mathbf{0}$ の解空間 W_A が次のように表されているとする.

$$W_A = \{c_1 \mathbf{a}_1 + c_2 \mathbf{a}_2 + \cdots + c_k \mathbf{a}_k \mid c_1, \dots, c_k \in \mathbb{R}\} = \langle \mathbf{a}_1, \dots, \mathbf{a}_k \rangle.$$

4.5 項で問題提起した次の 3 点について答えよう.

- (i) 任意定数の個数に水増しがないかどうかをどうやって判定すればよいのか. 特に, 掃き出し法による一般解の表示において任意定数の水増しはないか.

回答: 水増しがないとは $\mathbf{a}_1, \dots, \mathbf{a}_k$ の中に不要なベクトルがないということであり, これは $\mathbf{a}_1, \dots, \mathbf{a}_k$ が線形独立であることに他ならない (命題 17.2.5). つまり, 水増しがないことを示すには $\mathbf{a}_1, \dots, \mathbf{a}_k$ の線形独立性を示せばよい. また, 掃き出し法による一般解の表示において $\mathbf{a}_1, \dots, \mathbf{a}_k$ が線形独立になることは, 命題 18.3.4(2) およびその証明において得られている. すなわち掃き出し法による一般解の表示において任意定数の水増しはない.

- (ii) 任意定数の水増しのない異なる二つの一般解を与えたときに, 二つ解の任意定数の個数は必ず一致するか.

回答: 命題 22.1.1 より W_A の基底におけるベクトルの個数は一定である. $\mathbf{a}_1, \dots, \mathbf{a}_k$ に水増しがないければこれは W_A の基底であり, その数は $\dim W_A$ に一致する. すなわち, 任意定数に水増しのない一般解の表示において, 任意定数の個数は必ず $\dim W_A$ 個になる.

- (iii) 任意定数の定め方には, どれくらいの種類が考えられるのか.

回答: この質問は $\mathbf{a}_1, \dots, \mathbf{a}_k$ の選び方がどの程度あるかを問うているのであり, それは W_A の基底の選び方のぶんだけ任意性がある.

連立 1 次方程式の解法として掃き出し法を特別視しない立場において, 一般解の表示における任意定数の個数を定義しようと思えば次のようになる.

定義 22.2.1. (m, n) -行列 A による連立 1 次方程式 $Ax = \mathbf{b}$ の解が存在するとき, その一般解の表示における任意定数の個数とは, \mathbb{R}^n の部分空間 W_A (斉次方程式の解空間) の次元のことであり, その数は $n - \text{rank } A$ に等しい.

22.3 線形独立な最大個数

次元は次の数で表すこともできる (命題 22.3.3(2) を参照せよ). この概念を通して, 部分空間の次元や生成系との関係について考察しよう.

定義 22.3.1. 線形空間 V の部分集合 $A \subset V$ において, A の中に n 個のベクトルからなる線形独立な組があり, さらに A の中のいかなる $n+1$ 個のベクトルからなる組も線形従属になるとする. このとき, A における線形独立な最大個数を n と定める.

「 A における線形独立な最大個数が n であること」と「 A の元による線形独立な組におけるベクトルの個数は必ず n 以下であり, かつ A の中に n 個のベクトルの組で線形独立なものがあること」は同値である.

次の命題は, 線形独立な最大個数をとる範囲を生成元だけに制限しても, あるいは全体まで広げてもよいことを述べている.

命題 22.3.2. 線形空間 V および $A \subset V$ について次が成り立つ.

- (1) A における線形独立な最大個数が n であるとき, 線形独立な組 $\mathbf{a}_1, \dots, \mathbf{a}_n \in A$ は $\langle A \rangle$ の基底である.
- (2) $\langle A \rangle$ における線形独立な最大個数が n であるとき, 線形独立な組 $\mathbf{w}_1, \dots, \mathbf{w}_n \in \langle A \rangle$ は $\langle A \rangle$ の基底である.

(3) A における線形独立な最大個数が n であるとき, $\dim\langle A \rangle = n$.

(4) $\langle A \rangle$ における線形独立な最大個数が n であるとき, $\dim\langle A \rangle = n$.

Proof. (1): まず A の各元がいずれも $\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形結合で書けることを示そう. 仮に $\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形結合で書けないベクトル $\mathbf{a} \in A$ があるとすれば $n+1$ 個のベクトルの組 $\mathbf{a}_1, \dots, \mathbf{a}_n, \mathbf{a}$ は補題 18.2.5(1) により線形独立であり, これは線形独立な最大個数が n であることに反する. ゆえに $A \subset \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ である. V の部分空間 $\langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ は線形結合で閉じていることから $\langle A \rangle \subset \langle \mathbf{a}_1, \dots, \mathbf{a}_n \rangle$ であり, 以上より $\mathbf{a}_1, \dots, \mathbf{a}_n$ は $\langle A \rangle$ を生成する.

(2): いま示した 「 B における線形独立な最大個数が n であるとき, 線形独立な組 $\mathbf{a}_1, \dots, \mathbf{a}_n \in B$ は $\langle B \rangle$ の基底である」 における B として $\langle A \rangle$ を取れば, 命題 18.4.5 より $\langle B \rangle = B$ であり, (2) を得る.

(3) と (4) は, それぞれ (1) と (2) および次元の定義から明らか. \square

命題 22.3.3. 有限次元線形空間 V において次が成り立つ.

(1) V において線形独立な最大個数が定まり, その数は $\dim V$ に等しい.

(2) V における $\dim V$ 個の線形独立な組は V の基底である.

(3) $\dim V$ 個のベクトルの組 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{\dim V}$ が V を生成するならば, これは V の基底である.

Proof. $\dim V = n$ とし, $\mathbf{v}_1, \dots, \mathbf{v}_n$ を V の基底とする.

(1): V における $n+1$ 個のベクトルからなる組が線形従属になることを背理法によって示そう. 仮に $\mathbf{u}_1, \dots, \mathbf{u}_{n+1} \in V$ が線形独立であるとすれば, 命題 18.3.2 により, これらに更にいくつかのベクトルを加えることで V の基底とすることができる. すなわち, $n+1$ 個以上の元からなる V の基底が存在することになり, これは $\dim V = n$ に反する. したがって, $\mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ は線形従属である. また, n 個の組 $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立であり, V における線形独立な最大個数は n である.

(2): $A := \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ について, いま示した (1) と命題 22.3.2(2) より得る.

(3): $\mathbf{u}_1, \dots, \mathbf{u}_n$ が線形独立であることを示そう. そのためには, $\mathbf{u}_1, \dots, \mathbf{u}_n$ における線形独立な最大個数を k とするとき, k がちょうど n であることを示せばよい. そこで, $\mathbf{u}_1, \dots, \mathbf{u}_n$ の中から線形独立な組 $\mathbf{u}_{n_1}, \dots, \mathbf{u}_{n_k}$ を取る. 命題 22.3.2(1) より $\mathbf{u}_{n_1}, \dots, \mathbf{u}_{n_k}$ は $\langle \mathbf{u}_1, \dots, \mathbf{u}_n \rangle = V$ の基底となる. つまり $\dim V = k$ であり, ゆえに $n = k$. \square

補足: 上の (1) の証明において命題 18.3.2 を用いたが, これが証明に本質的に必要なわけではない. 命題 18.3.2 を用いずとも $\mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ の線形従属性は次のようにして示すことができる.

仮にこれらが線形独立であるとしよう. 命題 20.1.8 により, 各 \mathbf{u}_i を \mathbf{u}_i 自身に写す線形写像 $f: \langle \mathbf{u}_1, \dots, \mathbf{u}_{n+1} \rangle \rightarrow V$ が存在し, これは練習 20.3.7 より単射である. $\dim V = n$ より \mathbb{R}^n と V は同型であり, また $\langle \mathbf{u}_1, \dots, \mathbf{u}_{n+1} \rangle$ と \mathbb{R}^{n+1} は同型である. これらを結ぶ同型写像を $g: V \rightarrow \mathbb{R}^n$, $h: \mathbb{R}^{n+1} \rightarrow \langle \mathbf{u}_1, \dots, \mathbf{u}_{n+1} \rangle$ とすれば, 単射線形写像 $g \circ f \circ h: \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ を得る. しかし, \mathbb{R}^{n+1} から \mathbb{R}^n への線形単射が存在しないことは命題 21.2.7 の下に述べた通りである. ここに矛盾を得ることから, $\mathbf{u}_1, \dots, \mathbf{u}_{n+1}$ は線形従属でなければならない. \square

一方で, 線形空間が無限次元であることは次のように言い換えることができる.

命題 22.3.4 (発展). 線形空間 V において次は同値である.

(1) V は無限次元である.

(2) V は線形独立な無限部分集合を含む.

(3) V において, いくらでも多くのベクトルからなる線形独立な組を取ることができる.

(すなわち, 任意の自然数 n について, 線形独立な組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ が存在する.)

Proof. (1) \Rightarrow (2): 各 $k \in \mathbb{N}$ において組 v_1, \dots, v_k が線形独立となるようにベクトルの無限列 v_n を見つけたい. このような v_n が帰納的に取れることを示そう. いま, v_k までが得られているとし, v_1, \dots, v_k は線形独立であるとする. V は有限次元でないゆえ, これらは V の基底ではない. したがって, v_1, \dots, v_k の線形結合で表せない元 $v_{k+1} \in V$ が存在する. このとき, $k+1$ 個の組 v_1, \dots, v_{k+1} は補題 18.2.5(1) より線形独立である. こうして帰納的に選んだ v_n を集めた無限集合 $\{v_n \mid n \in \mathbb{N}\}$ は線形独立である.

(2) \Rightarrow (3): 明らか.

(3) \Rightarrow (1): 対偶を示そう. V が有限次元であるならば, 命題 22.3.3(1) より線形独立な最大個数は $\dim V$ であり, V の中から $\dim V + 1$ 個のベクトルからなる線形独立な組を取ることはできない. とくに, V の中からいくらかでも多くのベクトルからなる線形独立な組を取ることはできない. \square

例 22.3.5. 例 18.4.6 より数列空間 $\mathbb{R}^{\mathbb{N}}$ は線形独立な無限集合 $\{e_n \mid n \in \mathbb{N}\}$ を持つゆえ無限次元である.

22.4 次元から分かること

次元を用いて部分空間の分類や, 線形単射や線形全射の存在性について考察しよう.

命題 22.4.1. 有限次元線形空間 V の部分空間 W は有限次元であり, $\dim W \leq \dim V$. また, 等号成立は $W = V$ のときに限る.

Proof. $\dim V = n$ とおく. V における線形独立な最大個数は n であったから, W における線形独立な最大個数は n を越えることはない. そこでこの数を m とし, 線形独立な組 $w_1, \dots, w_m \in W$ を取ろう. $A = W$ について命題 22.3.2(2) を適用すれば w_1, \dots, w_m は W の基底であり, $\dim W = m \leq n = \dim V$. また, 等号 $n = m$ が成立する場合, 命題 22.3.3(2) より w_1, \dots, w_m は V の基底となり, $W = V$ を得る. \square

例 22.4.2. \mathbb{R}^3 の部分空間 W の次元は前命題により 3 以下であり, 次のように分類される.

- $\dim W = 0$ なる部分空間は $\{0\}$ のみである.
- $\dim W = 1$ なる部分空間の基底は一つのベクトル $u \neq 0$ からなり, $W = \{ru \mid r \in \mathbb{R}\}$ は u と 0 を通る直線に一致する.
- $\dim W = 2$ なる部分空間は, 基底 $u_1, u_2 \in W$ を持つ. $W = \{ru_1 + su_2 \mid r, s \in \mathbb{R}\}$ であり, これは $u_1, u_2, 0$ を通る平面に一致する.
- $\dim W = 3$ なる部分空間は前命題により $W = \mathbb{R}^3$ のみに限る.

練習 22.4.3 (発展). (1) t_1, \dots, t_{n+1} を相異なる実数とする. 任意の実数 r_1, \dots, r_{n+1} に対して, $p(t_i) = r_i$ ($i = 1, \dots, n+1$) を満たす n 次多項式 $p \in \mathbb{R}[x]_n$ が存在することを示せ.

解答: $I := \{t_1, \dots, t_{n+1}\} \subset \mathbb{R}$ とおけば, 命題 15.2.5 より $\mathbb{R}[x]_n \subset C(I)$ とみなせる. 写像 $p: I \rightarrow \mathbb{R}$ を $p(t_i) := r_i$ ($i = 1, \dots, n+1$) と定めれば $p \in C(I)$ である. p が $\mathbb{R}[x]_n$ の元でもあることを示せば主張を得る. $C(I) = \mathbb{R}^I \simeq \mathbb{R}^{n+1}$ ゆえ $\dim C(I) = n+1$ である⁵¹. また, $\dim \mathbb{R}[x]_n = n+1$ であるから命題 22.4.1 より $C(I) = \mathbb{R}[x]_n$ であり, $p \in \mathbb{R}[x]_n$. \square

(2) 上の証明の脚注で触れた f_i を多項式で表せ. すなわち, $f_i(t_i) = 1, f_i(t_j) = 0$ ($j \neq i$) を満たす n 次多項式を求めよ (この答えを通して, (1) で存在を示した多項式 p の具体的な式を書き下せることが分かる).

解答: $i = 1$ のみ記す.

$$f_1(x) := \frac{1}{(t_1 - t_2)(t_1 - t_3) \dots (t_1 - t_{n+1})} (x - t_2)(x - t_3) \dots (x - t_{n+1}).$$

⁵¹具体的に次のような基底があることから分かる: $f_i: I \rightarrow \mathbb{R}$ を $f_i(t_i) := 1, f_i(t_j) := 0$ (ただし $j \neq i$) と定めれば, f_1, \dots, f_{n+1} は $C(I)$ の基底である. 実際, $\sum_{i=1}^{n+1} a_i f_i(x) = 0$ とすれば両辺に t_i を代入することで $a_i = 0$ を得るゆえ線形独立である. また, 各 $g \in C(I)$ は $g(x) = \sum_{i=1}^{n+1} g(t_i) f_i(x)$ と書けるゆえ $C(I)$ を生成している.

線形単射や線形全射の存在は、次元の大小によって特徴づけられる。

命題 22.4.4. 有限次元の線形空間 U, V について次が成り立つ。

- (1) $\dim U \leq \dim V$ ならば線形単射 $f: U \rightarrow V$ が存在する。
- (2) $\dim U \geq \dim V$ ならば線形全射 $f: U \rightarrow V$ が存在する。

Proof. (1) $\dim U = n$ とおき、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする。 $\dim V \geq n$ より、 n 個のベクトルからなる線形独立な組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ が取れる。このとき、例 21.2.5(1) より $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす線形同型 $f: U \rightarrow \langle \mathbf{v}_1, \dots, \mathbf{v}_n \rangle$ があり、とくに f は単射である。 f の終域を V と解釈し、線形単射 $f: U \rightarrow V$ を得る。

(2) $\dim V = m$ および $\dim U = m + k$ (ただし $k \geq 0$) とおく。また、 $\mathbf{v}_1, \dots, \mathbf{v}_m$ を V の基底とし、 $\mathbf{u}_1, \dots, \mathbf{u}_m, \mathbf{u}_{m+1}, \dots, \mathbf{u}_{m+k}$ を U の基底とする。命題 20.1.8 により

$$f(\mathbf{u}_i) = \mathbf{v}_i \quad (i = 1, \dots, m), \quad f(\mathbf{u}_{m+j}) = \mathbf{0}_V \quad (j = 1, \dots, k)$$

を満たす線形写像 $f: U \rightarrow V$ を取れば、 $\text{Im } f = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_{m+k}) \rangle = \langle \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{0}_V, \dots, \mathbf{0}_V \rangle = V$ ゆえ f は全射である。 \square

命題 22.4.5. 有限次元線形空間の間の線形写像 $f: U \rightarrow V$ について次が成り立つ。

- (1) f が単射ならば $\dim U \leq \dim V$ 。
- (2) f が全射ならば $\dim U \geq \dim V$ 。

Proof. (1): U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$ とすれば命題 20.3.6(3) より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \in V$ は線形独立である。 $\dim V$ は、 V における線形独立な最大個数に等しいゆえ $\dim V \geq n = \dim U$ 。

(2): V の基底を $\mathbf{v}_1, \dots, \mathbf{v}_m$ とすれば、 f の全射性より $f(\mathbf{u}_i) = \mathbf{v}_i$ を満たす $\mathbf{u}_i \in U$ ($i = 1, \dots, m$) が存在する。命題 20.1.6(3) より $\mathbf{u}_1, \dots, \mathbf{u}_m$ は線形独立である。 $\dim U$ は、 U における線形独立な最大個数に等しいゆえ $\dim U \geq m = \dim V$ 。 \square

上の命題の対偶をとると、命題 21.2.7 の後に述べたことが得られる。

系 22.4.6. 有限次元線形空間 U および V について次が成り立つ。

- (1) $\dim U > \dim V$ ならば線形単射 $f: U \rightarrow V$ は存在しない。
- (2) $\dim U < \dim V$ ならば線形全射 $f: U \rightarrow V$ は存在しない。

単射性や全射性は次元からも判断できる。

練習 22.4.7. 有限次元線形空間 U, V の間の線形写像 $f: U \rightarrow V$ について、次を示せ。

- (1) f は単射である $\iff \dim U = \dim \text{Im } f$.

解答例. (\implies): $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とすれば、 $\text{Im } f = \langle f(\mathbf{u}_1), \dots, f(\mathbf{u}_n) \rangle$ である。 f の単射性より $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ は線形独立であり、したがってこれは $\text{Im } f$ の基底である。ゆえに $\dim \text{Im } f = n = \dim U$ 。

(\impliedby): $f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)$ を $\text{Im } f$ の基底とすれば、 f の線形性より $\mathbf{u}_1, \dots, \mathbf{u}_n$ は線形独立である。仮定より $\dim U = \dim \text{Im } f = n$ である。これと命題 22.3.3(2) を合わせれば、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U の基底であることが分かる。したがって、命題 21.1.6 より $f: U \rightarrow \text{Im } f$ は全単射である。とくに f は単射である。 \square

- (2) f は全射である $\iff \dim V = \dim \text{Im } f$.

解答例. (\implies): f が全射ならば $\text{Im } f = V$ より $\dim \text{Im } f = \dim V$ である。

(\impliedby): $\text{Im } f \subset V$ ゆえ $W = \text{Im } f$ について命題 22.4.1 を適用すれば $\text{Im } f = V$ 。すなわち f は全射である。 \square

22.5 無限次元の空間も含めた一般論 (発展)

これまで無限集合の線形独立性と線形写像との関係については控えていたが、ここでまとめておこう。有限個のベクトルからなる組の場合における証明を参考にすることで、次の命題群を証明することができる。これらは余力のある読者への演習問題として残すこととし、一般の読者には軽く読み流して次節に進むことを勧める。

命題 22.5.1 (命題 20.1.6 参照). $f: U \rightarrow V$ を線形写像とし, $A \subset U$ とする.

- (1) $x \in \langle A \rangle \implies f(x) \in \langle f(A) \rangle$.
- (2) $f(A)$ が線形独立ならば A も線形独立である.
- (3) A が線形従属ならば $f(A)$ も線形従属である.

命題 22.5.2 (命題 20.1.8 参照). $A \subset U$ を線形空間 U の基底とし, あらかじめ写像 $f: A \rightarrow V$ を与えておく. このとき, f の拡張である線形写像 $\tilde{f}: U \rightarrow V$ が唯一つ存在する.

命題 22.5.3 (命題 20.2.3 参照). 線形写像 $f: U \rightarrow V$ および $A \subset U$ において, $f(\langle A \rangle) = \langle f(A) \rangle$. とくに A が U を生成するとき, $f(A)$ は $\text{Im } f$ を生成する.

命題 22.5.4 (命題 20.3.6 参照). $f: U \rightarrow V$ を線形写像とし, $A \subset U$ とする. f が単射であるとき次が成り立つ:

- (1) A は線形独立である $\iff f(A)$ は線形独立である.
- (2) A は線形従属である $\iff f(A)$ は線形従属である.
- (3) $f(A)$ は V の基底である $\implies A$ は U の基底である.

命題 22.5.5 (命題 21.1.5 参照). $f: U \rightarrow V$ が線形同型であるとき, $A \subset U$ について次が成り立つ.

$$A \text{ は } U \text{ の基底である} \iff f(A) \text{ は } V \text{ の基底である.}$$

命題 22.5.6 (命題 21.1.6 参照). U の基底を $A \subset U$, $f: U \rightarrow V$ を線形写像とすると, 次が成り立つ.

$$f \text{ は線形同型である} \iff f(A) \text{ は } V \text{ の基底である.}$$

次は命題 22.1.1 に対応する主張である。この定理の証明は本論の枠を越えており、証明には無限集合の濃度演算について学ぶ必要がある。

定理 22.5.7. 線形空間 V の基底におけるベクトルの個数 (濃度) は, 基底の取り方によらずに一定である。

上の事実から、有限次元でない線形空間の次元を一括して ∞ と書くのではなく、基底の元の個数 (濃度) によって細かく分類する案が考えられる。この案による次元は、無限次元線形空間を含めた枠組みにおける完全不変量となる:

定理 22.5.8 (定理 22.1.4 参照). 線形空間 U, V において, $U \simeq V$ であることと U の基底と V の基底が対等であることは同値である。

Proof. $U \simeq V$ であると仮定し, $f: U \rightarrow V$ を線形同型とせよ. A を U の基底とすれば命題 22.5.6 により $f(A)$ は V の基底であり, また, f の単射性より A と $f(A)$ は対等である.

逆に, U の基底 A と, V の基底 B が対等であるとすれば, 全単射 $f: A \rightarrow B$ が存在する. このとき命題 22.5.2 より, f の拡張である線形写像 $\tilde{f}: U \rightarrow V$ が唯一つ存在する. \tilde{f} が線形同型であることは命題 22.5.6 より得る. \square

次は 22.4 項で述べた命題に対応する主張である。これらの証明も余力ある読者への演習として残そう。

命題 22.5.9 (命題 22.4.4 参照). 線形空間 U, V の基底をそれぞれ A, B とすれば次が成り立つ.

- (1) $f: A \rightarrow B$ が単射ならば, その拡張である線形単射 $\tilde{f}: U \rightarrow V$ が存在する.
- (2) $f: A \rightarrow B$ が全射ならば, その拡張である線形全射 $\tilde{f}: U \rightarrow V$ が存在する.

命題 22.5.10 (命題 22.4.5 参照). 線形空間 U, V の基底をそれぞれ A, B とする.

- (1) 線形単射 $f: U \rightarrow V$ が存在するならば, 単射 $p: A \rightarrow B$ が存在する.
- (2) 線形全射 $f: U \rightarrow V$ が存在するならば, 全射 $p: A \rightarrow B$ が存在する.

発展 ($\mathbb{R}[x]$ と $\mathbb{R}^{\mathbb{N}}$ は線形同型か).

本論では, 最も単純な無限次元空間の例として $\mathbb{R}[x]$ と $\mathbb{R}^{\mathbb{N}}$ を挙げているが, これらが線形同型かどうかについては言及しなかった. 結論を先に述べればこれらは同型ではない. この事実を示すには, 定理 22.5.8 を念頭に, $\mathbb{R}[x]$ の基底と $\mathbb{R}^{\mathbb{N}}$ の基底が対等でないことを示す戦略が考えられよう. 例 18.4.4 で述べたように $\mathbb{R}[x]$ の基底 $\{x^n \mid n \in \mathbb{Z}_{\geq 0}\}$ は非負整数と 1 対 1 の対応がつき, とくに \mathbb{N} と対等である. したがって, $\mathbb{R}^{\mathbb{N}}$ の基底が \mathbb{N} と対等でないことをいかに導くかが鍵となる. $\mathbb{R}^{\mathbb{N}}$ に基底が存在することは定理 18.4.7 によって保証されているものの, その具体像が不明であることが問題を難しくしており, 解決の糸口は極限操作を導入することにある.

$\mathbb{R}^{\mathbb{N}}$ のベクトルの間に距離を定めれば, ベクトルの列の収束発散が定まり, これを用いて $\mathbb{R}^{\mathbb{N}}$ 上の写像の連続性を定義できる. 距離には様々な定め方があるものの, 線形空間の構造と相性のよい距離ということであれば和やスカラー倍の演算を写像とみなしたときにこれらが連続写像となるような距離を導入することが望ましい. このような距離が入った線形空間を線形距離空間という. 線形距離空間の理論を進めると, \mathbb{N} と対等な基底をもつ線形距離空間は完備ではないことが示される. ここで, コーシー点列が必ず収束するような空間を完備であるという (微積分学でも \mathbb{R}^n の完備性を学ぼう). 一方, $\mathbb{R}^{\mathbb{N}}$ を完備な線形距離空間にできることが知られており, 以上の事実を総合すると, $\mathbb{R}^{\mathbb{N}}$ の基底は \mathbb{N} と対等でないことが分かる.

このように, 無限次元の線形空間の理論においては, 極限概念を導入することで対象を解析する手段を広げている. 線形空間への極限概念の導入の仕方も一通りではなく, その一端として大学初年級の線形代数学においても内積空間 (計量ベクトル空間) やノルム空間を学ぶことになる.

23 次元公式と商空間

連立1次方程式と斉次形方程式の解の関係、および線形常微分方程式とその斉次方程式の解の関係には類似性が認められていた(命題6.2.3および6.2.4)。これら類似する性質が線形写像に関する命題として一般的な立場から証明できることを本節で述べる。また、この性質を通して線形写像の次元公式を示す。次元公式とは、線形写像による空間の分解を次元の視点から述べた式のことである。

一方、線形空間の分解それ自体を記述するための概念として、商空間を定める。これは、次のような要請に応じて導入されるものである: ある線形空間において、いくつかのベクトル方向があまり重要ではないと判断されたとしよう。このとき、これらの方向を捨象した空間概念をいかに与えればよいか。この要請に満足する空間が、不要と思われる方向と並行な直線を集めた集合(つまり直線の集合)、あるいは、与えられた平面と並行な平面を集めた集合(つまり平面の集合)のような形で実現されることを本節の後半で見る。

ところで、本節および次節において、 $\mathbf{u}-\mathbf{v}$ と表記すればよいところをわざわざ読みにくい表記で $-\mathbf{v}+\mathbf{u}$ と書いた箇所がいくつかある。これは、可換性を満たさない代数構造においても同様の主張が成り立ち、その際の証明を見越したことによるものである。後に群論を学ぶ際の助けになるだろう。

23.1 空間の並行移動

連立1次方程式の解の集合の表示において集合の平行移動について言及していた。ここで改めて正確な定義を述べておこう。

定義 23.1.1. 線形空間 U の部分集合 $A \subset U$ およびベクトル $\mathbf{u} \in U$ に対して、 A を \mathbf{u} 方向に平行移動した集合を $A + \mathbf{u}$ あるいは $\mathbf{u} + A$ とかく(図4)。すなわち、

$$A + \mathbf{u} := \{\mathbf{a} + \mathbf{u} \mid \mathbf{a} \in A\}, \quad \mathbf{u} + A := \{\mathbf{u} + \mathbf{a} \mid \mathbf{a} \in A\}.$$

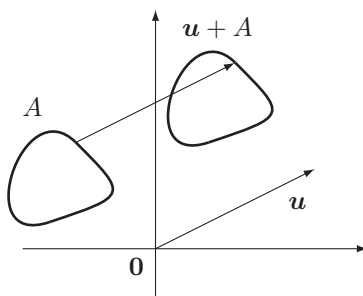


図4: 図形 A の \mathbf{u} 方向への平行移動

明らかに $A + \mathbf{u} = \mathbf{u} + A$ であり、この事実は和の演算の可換性 ($\mathbf{a} + \mathbf{u} = \mathbf{u} + \mathbf{a}$) に由来するものである。以降では $\mathbf{u} + A$ という表記を主に用いることにする⁵²。

命題 23.1.2. A と $\mathbf{u} + A$ は対等である。すなわち、これらの元の個数は等しい。

Proof. 写像 $f: A \rightarrow U$ を $f(\mathbf{a}) := \mathbf{u} + \mathbf{a}$ で定めれば、 $f(A) = \mathbf{u} + A$ が成り立つ。 f の単射性は明らかであり、したがって A と $f(A)$ は対等である。□

本論では、上の集合 A として U の部分空間を主に考える。

練習 23.1.3. 部分空間 $W \subset U$ および $\mathbf{u} \in W$ において、 $\mathbf{u} + W = W$ を示せ。

解答例: W が和の演算について閉じていることから、 $\mathbf{u} + W \subset W$ を得る。また、各 $\mathbf{v} \in W$ に対して $\mathbf{w} := -\mathbf{u} + \mathbf{v}$ とおけば $\mathbf{w} \in W$ であり、 $\mathbf{v} = \mathbf{u} + \mathbf{w} \in \mathbf{u} + W$ 。つまり $W \subset \mathbf{u} + W$ である。□

⁵²実は、 $A + \mathbf{u}$ のほうを採用していれば、冒頭で述べた $-\mathbf{v} + \mathbf{u}$ という表記は必要でなくなる。しかしながら、ここでは慣例に従い $\mathbf{u} + A$ と書くことにしたい。

次に述べる命題は、命題 6.2.3 および 6.2.4 で論じたことを線形写像の言葉で統一的に述べなおした主張に相当している⁵³。したがって、読者も直ちに証明方針を予想できることと思う。

命題 23.1.4. $f : U \rightarrow V$ を線形写像とし、 $\mathbf{b} \in V$ とする。更に $f(\mathbf{a}) = \mathbf{b}$ を満たす $\mathbf{a} \in U$ を一つ取って固定しよう (すなわち $\mathbf{a} \in f^{-1}(\mathbf{b})$)。このとき次が成り立つ:

- (1) 任意の $\mathbf{z} \in \text{Ker } f$ に対し、 $\mathbf{a} + \mathbf{z} \in f^{-1}(\mathbf{b})$ である。
- (2) 任意の $\mathbf{y} \in f^{-1}(\mathbf{b})$ は、ある $\mathbf{z} \in \text{Ker } f$ を用いて $\mathbf{y} = \mathbf{a} + \mathbf{z}$ と表せる。
- (3) $f^{-1}(\mathbf{b}) = \mathbf{a} + \text{Ker } f$

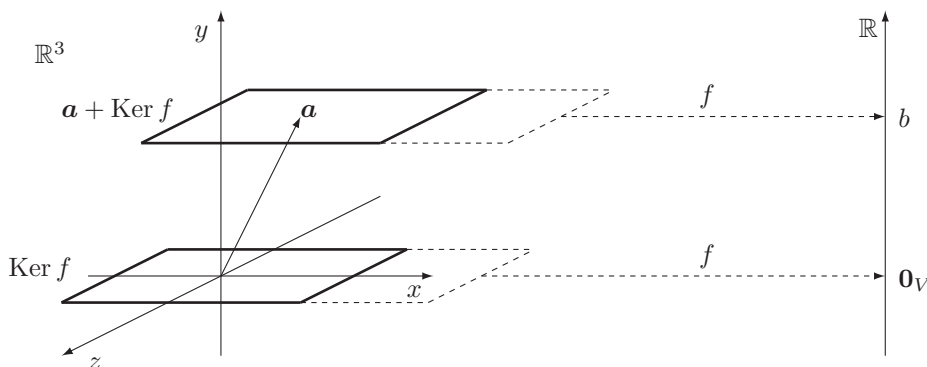


図 5: 線形写像の核とその平行移動

図 5 は線形写像 $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ ($f(x, y, z) := y$) について命題 23.1.4(3) を模式的に説明したものである。 $\text{Ker } f$ は x - z 平面に一致する。また、 $\mathbf{a} := (x_1, b, z_1)$ とすれば $f(\mathbf{a}) = b$ である (つまり $\mathbf{a} \in f^{-1}(b)$)。 $\text{Ker } f$ を \mathbf{a} 方向にずらしたものが $\mathbf{a} + \text{Ker } f$ であり、図では平行四辺形が右斜め上にずれたように見えているが、実際には x - z 平面と並行な方向に無限に広がる平面である。つまり、 $\text{Ker } f$ を真上に持ち上げたもの ($(0, b, 0)$ 方向に平行移動させた平面) とも一致する。命題 23.1.4(3) によれば $f^{-1}(b) = \mathbf{a} + \text{Ker } f$ であり、 $\mathbf{a} + \text{Ker } f$ に属する各元を f に代入すれば、その値はすべて b となる。

命題 23.1.4 の証明. (1) : $\mathbf{z} \in \text{Ker } f$ とすると、 $f(\mathbf{a} + \mathbf{z}) = f(\mathbf{a}) + f(\mathbf{z}) = b + \mathbf{0} = b$ 。ゆえに $\mathbf{a} + \mathbf{z} \in f^{-1}(b)$ である。

(2) : $\mathbf{y} \in f^{-1}(b)$ とする。 $\mathbf{z} := -\mathbf{a} + \mathbf{y}$ とおこう。このとき $f(\mathbf{z}) = f(-\mathbf{a} + \mathbf{y}) = -f(\mathbf{a}) + f(\mathbf{y}) = -b + b = \mathbf{0}$ より $\mathbf{z} \in \text{Ker } f$ である。また、 \mathbf{z} の定め方から $\mathbf{y} = \mathbf{a} + \mathbf{z}$ であり、我々は主張を得た。

(3) : 両方の包含関係 $f^{-1}(b) \subset \mathbf{a} + \text{Ker } f$ および $f^{-1}(b) \supset \mathbf{a} + \text{Ker } f$ を示せばよい。しかし、これらは (2) および (1) の主張をそれぞれ言い換えたものに過ぎない。 □

命題 23.1.2 および 23.1.4(3) より、 $f^{-1}(b) \neq \emptyset$ であるとき $f^{-1}(b)$ と $\text{Ker } f$ は対等である。とくに、線形写像の空でない逆像は互いに元の個数が等しい⁵⁴。

系 23.1.5. 線形写像 $f : U \rightarrow V$ および各 $\mathbf{b}_1, \mathbf{b}_2 \in \text{Im } f$ について、 $f^{-1}(\mathbf{b}_1)$ と $f^{-1}(\mathbf{b}_2)$ は対等である。

Proof. $f^{-1}(\mathbf{b}_1)$ と $f^{-1}(\mathbf{b}_2)$ はそれぞれ $\text{Ker } f$ と対等であり、したがって $f^{-1}(\mathbf{b}_1)$ と $f^{-1}(\mathbf{b}_2)$ も対等である。 □

⁵³ 実際、 (m, n) -行列 A による線形写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ に対して命題 23.1.4 を適用したものが命題 6.2.3 である。また、 $T(f) := \sum_{k=0}^n \alpha_k \frac{d^k}{dx^k} f$ (ただし各 $\alpha_k \in C^\infty(\mathbb{R})$ はあらかじめ決めておいた関数) で定められる線形写像 $T : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ に対して命題 23.1.4 を適用したものが命題 6.2.4 である。

⁵⁴ 「元の個数」といっても $\text{Ker } f$ は線形空間ゆえ、これらは一点集合でなければ無限集合である。しかし、より一般の代数構造 (群の準同型) においても同様の主張が成り立ち、その場合は n 点集合として元の個数が等しくなることがある。

23.2 線形写像の次元公式

有限次元の線形空間を定義域とする線形写像の次元公式を述べる。像と核がともに有限次元であることを一応ながら確認しておく。

命題 23.2.1. 有限次元線形空間を定義域とする線形写像の像と核は有限次元である。

Proof. この写像の像は命題 20.2.3 より有限個のベクトルによって生成される。ゆえに有限次元である。また、この写像の核は有限次元線形空間の部分空間であるから、命題 22.4.1 より有限次元である。□

定理 23.2.2. 有限次元線形空間 U を定義域とする線形写像 $f : U \rightarrow V$ において、 z_1, \dots, z_n が $\text{Ker } f$ の基底であり、 $u_1, \dots, u_m \in U$ について $f(u_1), \dots, f(u_m)$ が $\text{Im } f$ の基底であるならば、 $m + n$ 個の組 $u_1, \dots, u_m, z_1, \dots, z_n$ は U の基底である。

Proof. まず線形独立性を示そう。そこで線形関係 $\sum_{i=1}^m r_i u_i + \sum_{j=1}^n s_j z_j = \mathbf{0}_U$ を仮定する。これらに f をほどこすと

$$\begin{aligned} \sum_{i=1}^m r_i f(u_i) + \sum_{j=1}^n s_j f(z_j) &= f(\mathbf{0}_U) \\ \sum_{i=1}^m r_i f(u_i) + \sum_{j=1}^n s_j \mathbf{0}_V &= \mathbf{0}_V \\ \sum_{i=1}^m r_i f(u_i) &= \mathbf{0}_V. \end{aligned}$$

$f(u_1), \dots, f(u_m)$ の線形独立性より $r_1 = \dots = r_m = 0$ を得る。つまり $\sum_{j=1}^n s_j z_j = \mathbf{0}_U$ であり、 z_1, \dots, z_n の線形独立性より $s_1 = \dots = s_n = 0$ を得る。以上より $u_1, \dots, u_m, z_1, \dots, z_n$ は線形独立である。

次に各 $y \in U$ が $u_1, \dots, u_m, z_1, \dots, z_n$ の線形結合で書けることを示そう。 $\mathbf{b} := f(y)$ とおくと $\mathbf{b} \in \text{Im } f$ ゆえ $\mathbf{b} = \sum_{i=1}^m r_i f(u_i)$ と書ける。このとき $\mathbf{a} = \sum_{i=1}^m r_i u_i$ とおくと、 $f(\mathbf{a}) = f(\sum_{i=1}^m r_i u_i) = \sum_{i=1}^m r_i f(u_i) = \mathbf{b}$ より $\mathbf{a} \in f^{-1}(\mathbf{b})$ である。この \mathbf{a}, \mathbf{b} に対して命題 23.1.4(2) を適用すれば、 $y \in f^{-1}(\mathbf{b})$ は $z \in \text{Ker } f$ をもちいて $y = \mathbf{a} + z$ と書ける。 z_1, \dots, z_n は $\text{Ker } f$ の基底であったから $z = \sum_{j=1}^n s_j z_j$ と書け、 $y = \mathbf{a} + z = \sum_{i=1}^m r_i u_i + \sum_{j=1}^n s_j z_j$ を得る。□

いまの定理より直ちに次の公式を得る。

定理 23.2.3 (線形写像の次元公式). 有限次元線形空間 U を定義域とする線形写像 $f : U \rightarrow V$ において、

$$\dim U = \dim \text{Ker } f + \dim \text{Im } f.$$

上の公式は、 U および $\text{Im } f, \text{Ker } f$ の三つの空間のうち二つの次元が分かっていたら、残りの一つの次元も分かることを述べている。つまり、三つの空間のうち二つの空間がよく分かっていたら、残りの空間も分かるということである。

備考 23.2.4. 定理 23.2.2 の状況のもとで、 $H := \langle u_1, \dots, u_m \rangle$ とおけば、 $f|_H : H \rightarrow \text{Im } f$ は基底を基底に写す写像ゆえ線形同型写像である (命題 21.1.6)。

例 23.2.5. (m, n) -行列 A による線形写像 $T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ において、例 22.1.3(2) より $\dim \text{Im } T_A = \text{rank } A$ 、 $\dim \text{Ker } f = n - \text{rank } A$ である。ゆえに

$$\dim \text{Ker } f + \dim \text{Im } f = (n - \text{rank } A) + \text{rank } A = n = \dim \mathbb{R}^n$$

であり、確かに T_A において次元公式は成立している。

定理 23.2.2 の設定が成り立つ状況について補足しておこう。例えば、 $\text{Ker } f$ の基底を拡張して U の基底を得たならば、これらは定理 23.2.2 の設定を満たす：

練習 23.2.6. U を有限次元とし、 $f : U \rightarrow V$ を線形写像とする。いま、 z_1, \dots, z_n が $\text{Ker } f$ の基底であるとすれば、命題 18.3.2 より、これにいくつかのベクトルを付け加えて $z_1, \dots, z_n, u_1, \dots, u_m$ を U の基底とすることができる。このとき、 $f(u_1), \dots, f(u_m)$ が $\text{Im } f$ の基底となることを示せ。

解答例：次元公式により $\dim \text{Im } f = \dim U - \dim \text{Ker } f = (n + m) - n = m$ である。命題 20.2.3 より

$$\text{Im } f = \langle f(z_1), \dots, f(z_n), f(u_1), \dots, f(u_m) \rangle = \langle \mathbf{0}_V, \dots, \mathbf{0}_V, f(u_1), \dots, f(u_m) \rangle = \langle f(u_1), \dots, f(u_m) \rangle.$$

ゆえに $f(u_1), \dots, f(u_m)$ は $\text{Im } f$ を生成する。更に、命題 22.3.3(3) より、 $\dim \text{Im } f$ 個の組 $f(u_1), \dots, f(u_m)$ は $\text{Im } f$ の基底である。□

一方、 $f(u_1), \dots, f(u_m)$ が $\text{Im } f$ の基底であるとき、線形独立な組 u_1, \dots, u_m を拡張して U の基底を得たとしても、その際に加えたベクトルの組が $\text{Ker } f$ の基底になるとは限らない。

例 23.2.7. 行列 $A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ による線形写像 $T_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ において、 $\text{Im } T_A = \langle e_1 \rangle = \langle T_A(e_1) \rangle$ および $\text{Ker } T_A = \langle e_2 \rangle$ である。よって、 $T_A(e_1)$ は $\text{Im } T_A$ の基底であり、 e_1 に $u := e_1 + e_2$ を付け加えた e_1, u は \mathbb{R}^2 の基底となる。このとき、 $u \notin \text{Ker } T_A$ ゆえ u は $\text{Ker } T_A$ の基底ではない。

23.3 商空間

代入元 u を動かすことにより写像 $f : U \rightarrow V$ の値 $f(u)$ を変化させるにはどうすればよいか、あるいは逆に $f(u)$ を変化させないような代入元 u の動かし方を考えよう。こうした問いは数学の応用とも無縁ではない。例えば写像 f が利益やリスクと関係のある量ならば、これをできる限り都合のよい量に変化させたいと思うことは当然であろう。また、ある都合のよい性質を f の値が特徴づけるのであれば、これらの性質を保ちながら u を変化させることにどの程度の自由度があるのか知っておくに越したことはない。さて、 f が線形写像である場合、この問題の答えは次の通りである。

命題 23.3.1. 線形写像 $f : U \rightarrow V$ および $u, z \in U$ において、 $f(u) = f(u + z)$ となるための必要十分条件は $z \in \text{Ker } f$ となることである。

Proof. $b := f(u)$ とおこう。 $z \in \text{Ker } f$ を仮定すれば、命題 23.1.4(1) より $f(u + z) = b$ (すなわち $f(u) = f(u + z)$) である。逆に、 $f(u) = f(u + z) = b$ ならば $f(z) = f(-u + u + z) = -f(u) + f(u + z) = -b + b = \mathbf{0}$ より $z \in \text{Ker } f$ を得る。□

つまり、 $\text{Ker } f$ と並行な方向にベクトル u を移動させるか否かで $f(u)$ が変化するかどうかが決まるのである。とくに f の値の変化を望む立場からは $\text{Ker } f$ 方向への移動は無意味である。そこで、 U の中で $\text{Ker } f$ 方向の情報を捨象する概念として商空間を与えよう。ここでは $\text{Ker } f$ に限らず、一般の部分空間 $W \subset U$ に対して商空間 U/W を定める。

定義 23.3.2. U を線形空間とし、 W をその部分空間とする。 W を並行移動させた集合 (これは U の部分集合である) をすべて集めた集合 (つまり集合族) を U の商空間 (quotient space) と呼び、 U/W と書く。すなわち、

$$U/W := \{ u + W \mid u \in U \}.$$

集合 $u + W$ を商空間 U/W の元とみなすとき、これを略して $[u]$ と書く。

$u = u + \mathbf{0} \in u + W = [u]$ より $u \in [u]$ であることに注意せよ。また、 $[0] = \mathbf{0} + W = W$ である。

例 23.3.3. \mathbb{R}^3 における x - z 平面を $W = \{ (x, y, z) \in \mathbb{R}^3 \mid y = 0 \}$ とすれば、 \mathbb{R}^3/W は W と並行な平面によって構成される集合である。図 6 では有限個の平面しか描かれていないが、実際は無数の平面が連続的に並んでいる。

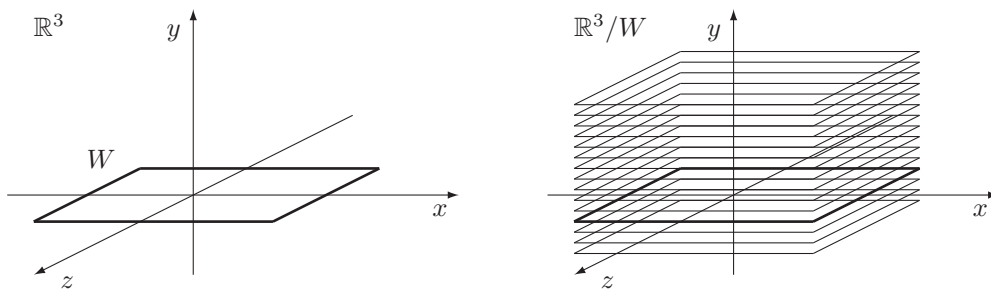


図 6: \mathbb{R}^3 上の平面 W による商空間 \mathbb{R}^3/W

U の元 u の代わりに U/W の元 $[u]$ を考えるにあたって, この対応に関する写像は商写像と呼ばれる:

定義 23.3.4. U の各元 $u \in U$ に対して, $[u] \in U/W$ を対応させる写像 $q: U \rightarrow U/W$ を商写像 (quotient map) と言う.

商写像 $q: U \rightarrow U/W$ は全射である. 実際, U/W の各元 $u + W$ に対して, $q(u) = u + W$ である.

次の命題は, u と $u + w$ が商空間 U/W において同じ元を表すこと (すなわち $[u] = [u + w]$) と $w \in W$ であることの同値性を述べている (系 23.3.6). つまり, U の元 u の代わりに $[u]$ を考えるということは, W と並行な方向を捨象することを意味している.

命題 23.3.5. 各 $u, v \in U$ について, $[u] = [v] \iff -v + u \in W \iff u - v \in W$.

Proof. (\implies) を示すために $[u] = [v]$ (すなわち $u + W = v + W$) を仮定しよう. すると $u \in [u] = v + W$ ゆえ, ある $w \in W$ を用いて $u = v + w$ と書くことができる. この両辺に $-v$ を加えて $-v + u = w \in W$ を得る.

次に (\impliedby) を示すために $-v + u \in W$ を仮定しよう. $[u] = [v]$ を示すには $[u] \subset [v]$ および $[v] \subset [u]$ を示せばよい. まず $[u] \subset [v]$ を示すために $a \in [u]$ を任意に取る. すると, ある $w \in W$ を用いて $a = u + w$ と書くことができる. このとき, 別の $w' \in W$ を用いて $a = v + w'$ となることを示せばよいのであるが, これは次の計算により得られる.

$$a = u + w = (v - v) + u + w = v + (-v + u) + w.$$

つまり, $w' := (-v + u) + w$ とすれば, $(-v + u), w \in W$ および W が和の演算について閉じていることから $w' \in W$ であり, $a = v + w' \in v + W$ を得る. いま $[u] \subset [v]$ であることが分かった. u と v の立場を入れ替えることで $[v] \subset [u]$ も同様にして示すことができる. この証明において, W が部分空間であることから $-u + v = -(-v + u) \in W$ であることに注意せよ. \square

系 23.3.6. 各 $w \in U$ について, $[u + w] = [u] \iff w \in W$.

Proof. 前命題の u および v として, $u + w$ および u を取れば主張を得る. \square

23.4 商空間の例

例 23.4.1. (1) \mathbb{R}^2 における x 軸上の点のなす直線 $X = \{(x, y) \in \mathbb{R}^2 \mid y = 0\}$ に対して, \mathbb{R}^2/X は X と並行な直線からなる集合 (直線を元とする集合) である. \mathbb{R}^2/X と y 軸上の点のなす集合 Y には自然な 1 対 1 対応がある. 実際, \mathbb{R}^2/X の元である直線 $\ell = \{(x, y) \in \mathbb{R}^2 \mid y = a\}$ に対して, ℓ と Y との交点 $(0, a)$ を対応させる写像 $\mathbb{R}^2/X \rightarrow Y$ は全単射である.

(2) 上の例において \mathbb{R}^2/X との間には 1 対 1 対応を与える直線を Y に限る必要はなく, 直線 X と並行でない任意の直線 L について, \mathbb{R}^2/X と L の間に 1 対 1 対応がつく. 実際, 各直線 $\ell \in \mathbb{R}^2/X$ に対して, ℓ と L の交点を対応させれば (ℓ と L は平行でないゆえ一点で交わる), この対応 $\mathbb{R}^2/X \rightarrow L$ は全単射である.

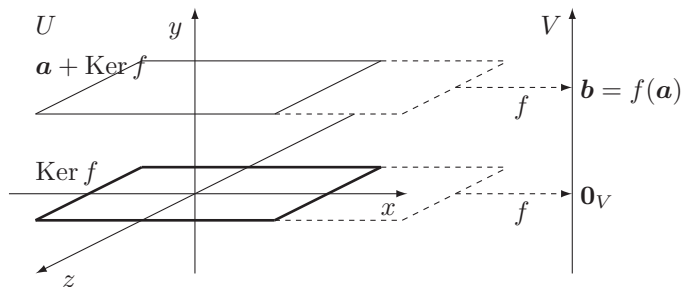
(3) \mathbb{R}^2 における直線 $y = \frac{1}{2}x$ 上の点全体を L とすれば, \mathbb{R}^2/L と y 軸との間でも 1 対 1 対応がつく.

(4) \mathbb{R}^3 における x 軸上の点のなす直線 X に対して, \mathbb{R}^3/X と y - z 平面の間に 1 対 1 対応がつく.

上の例から, 商空間も線形空間となることが示唆される. 商空間にどのように和とスカラー倍を定めるかは, 次節で詳しく述べよう.

線形写像 $f: U \rightarrow V$ に対して, f の値が動かないような代入元 \mathbf{u} の移動とは, $\text{Ker } f$ と並行な方向に移動させることに他ならなかった. そこで, $\text{Ker } f$ 方向を捨象した $U/\text{Ker } f$ を定義域とする f の代わりとなる写像 $F: U/\text{Ker } f \rightarrow V$ を考えることができる.

例 23.4.2. $f: U \rightarrow V$ を線形写像とする. 各 $\mathbf{b} \in \text{Im } f$ に対して $f(\mathbf{a}) = \mathbf{b}$ なる $\mathbf{a} \in U$ を取れば, 命題 23.1.4(3) より $f^{-1}(\mathbf{b}) = \mathbf{a} + \text{Ker } f = [\mathbf{a}] \in U/\text{Ker } f$ である.



そこで, 各 $f^{-1}(\mathbf{b}) = [\mathbf{a}] \in U/\text{Ker } f$ に対して $\mathbf{b} \in V$ を対応させる写像を $F: U/\text{Ker } f \rightarrow V$ と定めよう. 写像の値が変化しない方向を潰したことから, L が $U/\text{Ker } f$ 上を動けば $F(L)$ の値も必ず変化し, したがって F は単射となる.

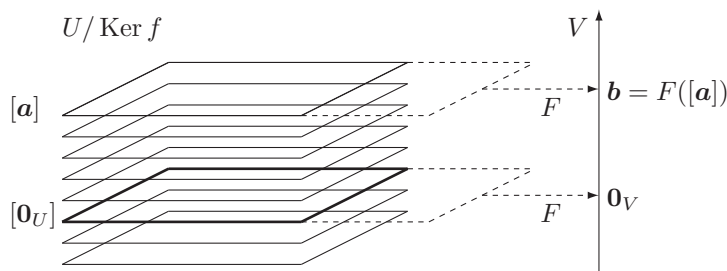


図 7: 線形写像 f が誘導する写像 $F: U/\text{Ker } f \rightarrow V$

実は, 上の F は線形写像となる. F の線形性および単射性の証明は定理 24.2.1(準同型定理) をみよ.

23.5 同値関係と商集合 (発展)

商空間 U/W の各元は W と平行な集合であるゆえ, 互いに交わらないはずである. これを証明によって確認しておこう.

命題 23.5.1. 商空間 U/W において次が成り立つ.

- (1) 各 $\mathbf{u} \in U$ に対して, 商空間 U/W の要素で \mathbf{u} を含むものは $[\mathbf{u}]$ 唯一のみである.
- (2) 各 $L \in U/W$ および $\mathbf{u} \in L$ について $[\mathbf{u}] = L$. とくに $\mathbf{w} \in W$ について $[\mathbf{w}] = W = [\mathbf{0}]$.
- (3) 各 $L \in U/W$ および $\mathbf{u} \in U$ について, $[\mathbf{u}] = L \iff \mathbf{u} \in L$.
- (4) U/W の任意の二つの元は, 完全に一致するか交わらないかのいずれかである.

(すなわち, $L, L' \in U/W$ とすれば $L \cap L' = \emptyset$ または $L = L'$ のいずれかが成り立つ⁵⁵.)

⁵⁵ 集合演算記号 \cap は共通部分を表す. すなわち, 集合 A, B のいずれにも含まれる元をすべて集めた集合を A と B の共通部分と呼び, これを $A \cap B$ で表す.

Proof. (1): U/W の二つの元 $[a] = a + W$ と $[b] = b + W$ がともに u を含むと仮定する. すると, ある $w, w' \in W$ を用いて $u = a + w = b + w'$ と書くことができる. このとき, $-b + a = w' - w \in W$ であり, 命題 23.3.5 より $[a] = [b]$. すなわち, u を含む U/W の要素は一つしかない (注意: $[u]$ は u を含む U/W の要素ゆえ $[u] = [a] = [b]$ が成り立っている).

(2): $L \in U/W$ および $u \in L$ について, $[u]$ と L はともに u を元として含む U/W の要素である. このような U/W の要素は (1) より唯一つしかないことから, $[u] = L$ である.

(3): (\Leftarrow) は (2) で示している. また, $[u] = L$ とすれば $u \in [u] = L$ ゆえ $u \in L$ である.

(4): $L, L' \in U/W$ および $L \cap L' \neq \emptyset$ とすれば $u \in L \cap L'$ が取れる. このとき, (1) より $L = L' = [u]$ である. \square

上の命題から, 商空間 U/W は U の各元を互いに交わらないグループに分類していることが分かる. $u \in U$ の属するグループが $[u]$ という具合である. このような交わりのない分類 (グループ分け) は, 線形代数の枠組みに縛られない一般論として次のように定められる:

定義 23.5.2. 集合 X の元に対して同値関係⁵⁶ \simeq が与えられているとき, 各 $x \in X$ の同値類 $[x] \subset X$ を次のように定める.

$$[x] := \{y \in X \mid y \simeq x\}.$$

また, X の部分集合からなる次の集合族 X/\simeq を同値関係 \simeq に関する商集合という:

$$X/\simeq := \{[x] \mid x \in X\}.$$

商集合は X の各元を互いに交わらないグループに分類しており, 命題 23.3.5 および 23.5.1 と同様の主張が一般の商集合においても成り立つ:

命題 23.5.3. X 上の同値関係 \simeq による商集合 X/\simeq において次が成り立つ.

(1) 各 $x, y \in X$ について, $[x] = [y] \iff x \simeq y$.

(2) 各 $x \in X$ に対して, 商集合 X/\simeq の要素で x を含むものは $[x]$ 唯一つのみである.

(3) 各 $L \in X/\simeq$ および $x \in X$ について, $[x] = L \iff x \in L$.

(4) X/\simeq の任意の二つの元は, 完全に一致するか交わらないかのいずれかである.

Proof. (1): (\Rightarrow) を示すために $[x] = [y]$ を仮定しよう. $x \simeq x$ (反射律) より $x \in [x]$ であり, $[x] = [y]$ ゆえ $x \in [y]$. つまり, $x \simeq y$ である. 次に (\Leftarrow) を示すために $x \simeq y$ を仮定しよう. $[x] = [y]$ を示すには $[x] \subset [y]$ および $[y] \subset [x]$ を示せばよい. まず $[x] \subset [y]$ を示すために $a \in [x]$ を任意に取る. このとき $a \simeq x$ であり, これと $x \simeq y$ から $a \simeq y$ を得る (推移律). つまり $a \in [y]$ である. いま $[x] \subset [y]$ であることが分かった. x と y の立場を入れ替えることで $[y] \subset [x]$ も同様に示すことができる. この証明において, 対称律から $y \simeq x$ であることに注意せよ.

(2): X/\simeq の二つの元 $[a]$ と $[b]$ がともに x を含むと仮定する. すると, $x \simeq a$ および $x \simeq b$ が成り立っている. よって対称律と推移律より $a \simeq b$ であり, (1) より $[a] = [b]$. すなわち, x を含む X/\simeq の要素は一つしかない (注意: $[x]$ は x を含む X/\simeq の要素ゆえ $[x] = [a] = [b]$ が成り立っている).

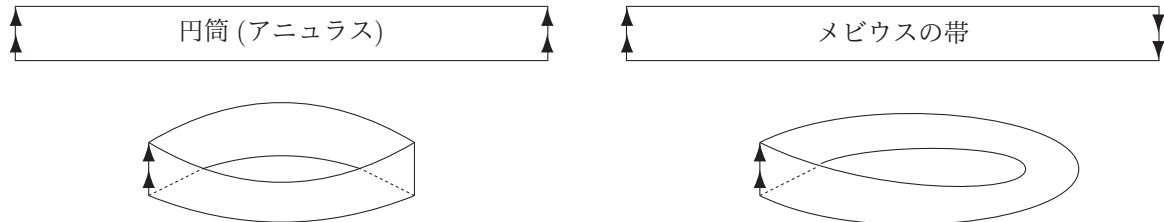
(3) および (4) は, 命題 23.5.1(2) および (3), (4) と同様に示される. \square

例 23.5.4. 線形空間 U およびその部分空間 W に対して, $u \simeq v$ を「 $-v + u \in W$ を満たすこと」と定めれば, これは同値関係である. また, この同値関係による商集合 U/\simeq と本節で与えた商空間 U/W は一致する. すなわち, 各 $u \in U$ について u の同値類と集合 $u + W$ は同じ集合である.

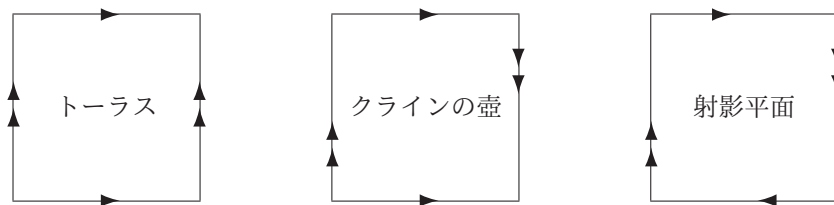
⁵⁶同値関係の定義は 21.2 項のコラムを見よ.

よりみち (図形の貼り合わせ).

商集合は、対象の同一視を記述するための言葉である。代数学以外の文脈でも用いられる概念であり、例えば幾何学における商集合に図形の貼り合わせがある。最も単純な貼り合わせの例は、線分の両端点を同一視することで得る円周であろう。ここから一つ次元を挙げて、長方形 (内側も含む) の左右の両辺を貼り合わせる (すなわち両辺を同一視する) 操作を考えよう。両辺の貼り合わせ方には、向きをそろえる場合と逆にする場合の二通りが考えられ、前者による貼り合わせからは円筒 (円柱の側面) が得られ、後者からはメビウスの帯が得られる。



さらに、長方形の上下の辺どうしも貼り合わせるとすれば、次の3通りが考えられる:



左のトーラスとは、いわゆるドーナツの表面のような図形のことであり、細長い円筒の両端を自然に貼り合わせることでトーラスが得られる。この両端の貼り合わせを逆向きに行えばクラインの壺になるのであるが、折り紙などを用いて実際に貼り合わせようと思うと、どうしても面が重なりあって上手くいかないはずである。射影平面はさらに複雑な図形のように思えるだろう。実は、右の二つの図形は \mathbb{R}^3 上の図形として実現できないことが知られている。また、これらはメビウスの帯を含むことから、表裏が定まらない曲面でもある。このように、貼り合わせによって構成した新しい図形の性質を調べる際に、証明が求められる数学では貼り合わせの厳密な定義が必要となる。この要請に応える概念が同値関係および商集合なのである。

上の構成では射影平面のイメージがつかみにくいだらうから、別の方法による構成を紹介しよう。 \mathbb{R}^3 の単位球面

$$S^2 = \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1 \}$$

にある点に対して、 $x \simeq y$ を「 $x = y$ または $x = -y$ であること」と定めると \simeq は同値関係となる。これは、球の中心について点対称な位置にある互いの点を同一視することに相当する。先程貼り合わせによって作った射影平面と商集合 S^2/\simeq との間には自然な 1 対 1 対応がある。このことは、北半球面 $\{ (x, y, z) \in S^2 \mid z \geq 0 \}$ において今の同値関係による同一視を行うと理解し易い。北半球面に限ると、同一視すべき異なる二点は赤道にしか現れない。そして、赤道上的点における同一視は、ちょうど上で定めた射影平面における長方形の各辺の同一視と対応付けられる。なお、球面の赤道付近の帯

$$A = \left\{ (x, y, z) \in S^2 \mid |z| \leq \frac{1}{10} \right\}$$

において、 A/\simeq はメビウスの帯となる。いまの例とは別に、射影平面には次のような構成もある。

練習 23.5.5. \mathbb{R}^3 上の原点を通る直線全体の集合を P とおくと、 P と S^2/\simeq の間に自然な 1 対 1 対応があることを確認せよ。

解答例: 各 $s \in S^2/\simeq$ に対して、同値類 s は S^2 の二点からなる部分集合であり、これらの点は互いに原点对称な位置にある。そこで、 s 上の 2 点を結ぶ直線を $F(s)$ とすれば、これは原点を通る直線である。対応 $F : S^2/\simeq \rightarrow P$ は全単射である。□

24 準同型定理と短完全列 (発展)

前節において、線形空間を分解する概念として次元公式と商空間を与えた。本節では、これらの関係について論じ、次元公式を商空間の立場から述べた主張である準同型定理を導く。実は、線形空間に限らず多くの代数構造 (群や環, 加群, 多元環など) においても準同型定理は認められる。つまり、この定理は代数構造を分解して理解するうえで基本となる考え方であり、代数学では至る所で用いられるものである。これと関連して、準同型定理による代数構造の分解を図式で表す短完全列についても少しだけ触れる。

24.1 商空間の線形構造

商空間における和とスカラー倍の定義を与えよう。

補題 24.1.1. 商空間 U/W において和とスカラー倍の演算を次のように定めることができる。

各 $[u], [v] \in U/W$ および $r \in \mathbb{R}$ に対して, $[u] + [v] := [u + v]$, $r[u] := [ru]$.

ここで、「定めることができる」と書いたのは次の点に考慮せねばならないからである: U/W の元 L を U の元を用いて表す方法は命題 23.5.1 に見るように何通りもあり (実際 L の元の個数ぶんだけある), 例えば $L = [u] = [u']$ としよう。このとき L のスカラー倍 rL は, $[ru]$ と $[ru']$ のいずれであると定めるべきだろうか。もし $[ru] \neq [ru']$ ならば L の代表として u と u' のいずれを選ぶのか (あるいは u, u' のいずれとも異なる L の元を代表に選ぶのか) あらかじめ決めておかなければ, 上は定義として認められない (読み手に定義が伝わらない) ことになる。しかしながら, 実際には $[ru] = [ru']$ が成り立ち, L の中のど

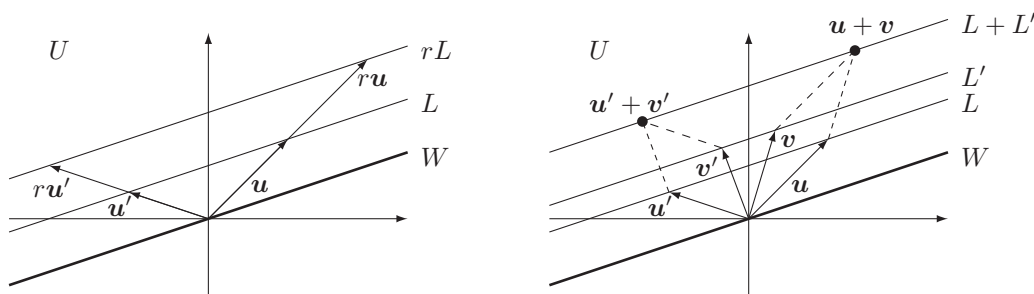


図 8: $L \in U/W$ のスカラー倍および $L' \in U/W$ との和

のベクトルを代表に選んでも定義が変わることはない。つまり, L の代表元に何をを選ぶか指定する必要はないのである。和の定義についても同様のことが言える (図 8)。

補題 24.1.1 の証明. 上で説明したように, 示すべき事は和とスカラー倍の定義が代表元の取り方によらないこと, すなわち次の二つである:

$$(1) [u] = [u'] \implies [ru] = [rv]. \quad (2) [u] = [u'] \text{ かつ } [v] = [v'] \implies [u + v] = [u' + v'].$$

(1): $[u] = [u']$ とすれば命題 23.3.5 より $-u' + u \in W$ であり, ゆえに $-ru' + ru = r(-u' + u) \in W$. 再び命題 23.3.5 より $[ru] = [ru']$ を得る。

(2): 仮定より $-u' + u, -v' + v \in W$ であり, ゆえに

$$-(u' + v') + (u + v) = (-u' + u) + (-v' + v) \in W$$

である⁵⁷. したがって $[u + v] = [u' + v']$. □

⁵⁷いまの式変形を非可換な場合も見据えた形で行うと次のようになる:

$$\begin{aligned} -(u' + v') + (u + v) &= (-v' - u') + (u + v) = -v' + (-u' + u) + v \\ &\in -v' + (W + v) = -v' + (v + W) = (-v' + v) + W = [-v' + v] = W. \end{aligned}$$

非可換の世界では一般に $(W + v) = (v + W)$ は成立せず, 上の式変形を行うには条件 $(W + v) = (v + W)$ を仮定する必要がある。この条件と関連して, 正規部分群なる概念を群論で学ぶことになる。

上の補題により, 任意の $\mathbf{u}, \mathbf{v} \in U$ および $r \in \mathbb{R}$ について次の式変形が認められることになる.

$$[\mathbf{u} + \mathbf{v}] = [\mathbf{u}] + [\mathbf{v}], \quad [r\mathbf{u}] = r[\mathbf{u}].$$

$[0] = W \in U/W$ を零元とみなせば, 補題 24.1.1 で定めた U/W 上の演算がベクトル空間の公理を満たすことは容易に確かめられる. したがって U/W は線形空間となる.

命題 24.1.2. 商写像 $q : U \rightarrow U/W$ は線形全射であり, $\text{Ker } q = W$.

Proof. 線形性 (iii) は直ちに確認できる:

$$q(r\mathbf{u} + s\mathbf{v}) = [r\mathbf{u} + s\mathbf{v}] = [r\mathbf{u}] + [s\mathbf{v}] = r[\mathbf{u}] + s[\mathbf{v}] = rq(\mathbf{u}) + sq(\mathbf{v}).$$

また, $[0] = W$ に注意すると,

$$\mathbf{u} \in \text{Ker } q \iff q(\mathbf{u}) = [0] \iff [\mathbf{u}] = W \iff \mathbf{u} \in W \quad (\text{命題 23.5.3(3)}).$$

すなわち, $\text{Ker } q = W$. □

命題 24.1.3. 有限次元線形空間 U および部分空間 $W \subset U$ について, $\dim U/W = \dim U - \dim W$.

Proof. 商写像 $q : U \rightarrow U/W$ に対して次元公式を適用すればよい. q の全射性より $U/W = \text{Im } q$ であること, また $\text{Ker } q = W$ に注意すれば,

$$\begin{aligned} \dim \text{Im } q + \dim \text{Ker } q &= \dim U \\ \dim U/W + \dim W &= \dim U \\ \dim U/W &= \dim U - \dim W. \end{aligned}$$

□

練習 24.1.4. U を有限次元線形空間とし, W をその部分空間とする. W の基底 $\mathbf{w}_1, \dots, \mathbf{w}_n$ にいくつかのベクトルを付け加えて $\mathbf{w}_1, \dots, \mathbf{w}_n, \mathbf{u}_1, \dots, \mathbf{u}_m$ を U の基底とすれば $[\mathbf{u}_1], \dots, [\mathbf{u}_m]$ は U/W の基底となる. これを示せ.

解答例: 練習 23.2.6 と同様の論法をたどればよい. 前命題より $\dim U/W = \dim U - \dim W = (n + m) - n = m$ である. 命題 20.2.3 より

$$\text{Im } q = \langle q(\mathbf{w}_1), \dots, q(\mathbf{w}_n), q(\mathbf{u}_1), \dots, q(\mathbf{u}_m) \rangle = \langle [0], \dots, [0], [\mathbf{u}_1], \dots, [\mathbf{u}_m] \rangle = \langle [\mathbf{u}_1], \dots, [\mathbf{u}_m] \rangle.$$

ゆえに $[\mathbf{u}_1], \dots, [\mathbf{u}_m]$ は $\text{Im } q = U/W$ を生成する. 更に, 命題 22.3.3(3) より, $\dim U/W$ 個の組 $[\mathbf{u}_1], \dots, [\mathbf{u}_m]$ は U/W の基底である. □

24.2 準同型定理

例 23.4.2 で述べた事実の詳細は次の通りである.

定理 24.2.1 (準同型定理). 線形写像 $f : U \rightarrow V$ について, $U/\text{Ker } f$ と $\text{Im } f$ は線形同型である.

Proof. 写像 $F : U/\text{Ker } f \rightarrow \text{Im } f$ を $F([\mathbf{u}]) := f(\mathbf{u})$ と定めることができる. この定義が代表元の取り方によらないこと, すなわち「 $[\mathbf{u}] = [\mathbf{v}]$ ならば $f(\mathbf{u}) = f(\mathbf{v})$ 」であることを確認しよう. 実際, $[\mathbf{u}] = [\mathbf{v}]$ ならば $-\mathbf{v} + \mathbf{u} \in \text{Ker } f$ であり, これは $f(-\mathbf{v} + \mathbf{u}) = \mathbf{0}_V$ を意味する. つまり $-f(\mathbf{v}) + f(\mathbf{u}) = f(-\mathbf{v} + \mathbf{u}) = \mathbf{0}_V$ であり, $f(\mathbf{u}) = f(\mathbf{v})$ を得る.

いま定めた F が線形同型であることを示そう. まず線形性 (iii) は次のように確認できる:

$$F(r[\mathbf{u}] + s[\mathbf{v}]) = F([r\mathbf{u} + s\mathbf{v}]) = f(r\mathbf{u} + s\mathbf{v}) = rf(\mathbf{u}) + sf(\mathbf{v}) = rF([\mathbf{u}]) + sF([\mathbf{v}]).$$

また, 各 $f(\mathbf{u}) \in \text{Im } f$ に対して, $F([\mathbf{u}]) = f(\mathbf{u})$ であるから F は全射である. 最後に単射性を示そう. $[\mathbf{u}] \in \text{Ker } F$ とすれば $F([\mathbf{u}]) = \mathbf{0}_V$, すなわち $f(\mathbf{u}) = \mathbf{0}_V$ であり, したがって $\mathbf{u} \in \text{Ker } f$ である. ゆえに命題 23.5.3(2) より $[\mathbf{u}] = \text{Ker } f = [0]$ であり, $\text{Ker } F$ の元は零元のみであることが分かった. すなわち F は単射である. □

練習 24.2.2. 次元公式を用いて, U が有限次元の場合における線形写像 $f: U \rightarrow V$ の準同型定理の別証明を与えよ.

解答例: 命題 24.1.3 および次元公式より $\dim U / \text{Ker } f = \dim U - \dim \text{Ker } f = \dim \text{Im } f$. したがって $U / \text{Ker } f$ と $\text{Im } f$ の次元は等しく, ゆえにこれらは同型である. \square

有限次元の空間に限れば, 上のようにして準同型定理の主張する線形同型性は次元定理から直ちに分かってしまう. したがって, 準同型定理の真の御利益は何かと問われれば, それは無限次元空間を対象とするか, あるいは定理の証明に現れた同型写像 F を本質的に必要とするような例を挙げねばならない. しかし, 何の数学理論の予備知識も仮定せずに, そのような応用例を説明するのは残念ながら困難なことである. そこで, ここでは次の例を述べるに留めておく.

例 24.2.3 (不定積分の線形性). 例 20.4.4(2) において, 微分作用素 $D: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ ($D(f) := f'$) の核 $\text{Ker } D$ が定数関数全体に一致することを見た. つまり $C^\infty(I) / \text{Ker } D$ は, 定数関数だけの差を無視した関数の空間である. すなわち, 関数 $f \in C^\infty(\mathbb{R})$ および定数 $C \in \mathbb{R}$ について,

$$[f(x)] = [f(x) + C] \in C^\infty(\mathbb{R}) / \text{Ker } D.$$

D に準同型定理を適用すれば, 線形同型 $\mathcal{D}: C^\infty(I) / \text{Ker } D \rightarrow C^\infty(I)$ を得る. \mathcal{D} の逆写像 $\mathcal{I}: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R}) / \text{Ker } D$ は, $f \in C^\infty(\mathbb{R})$ に対してその原始関数 F の同値類

$$\mathcal{I}(f) = F + \text{Ker } D \in C^\infty(\mathbb{R}) / \text{Ker } D$$

を対応させる写像である. 原始関数には定数関数の個数だけ任意性があるものの, $\mathcal{I}(f)$ は原始関数の選び方に依らずに決まる $C^\infty(\mathbb{R}) / \text{Ker } D$ の元である. $\mathcal{I}(f)$ は f の不定積分とよばれ, 微積分学では次の式で表される:

$$\int f(x) dx.$$

こうして我々は, 不定積分が $C^\infty(\mathbb{R}) / \text{Ker } D$ への線形写像として実現されること, とくに不定積分の線形性を得る. なお, 例 20.4.6(2) で与えた単射 $I_a: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ は原始関数の一つを与える線形写像であったゆえ, $q: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R}) / \text{Ker } D$ を商写像とすれば $\mathcal{I} = q \circ I_a$ が成り立つ.

24.3 完全系列と短完全列

定義 24.3.1. 線形写像の列

$$\dots \xrightarrow{f_{n-2}} U_{n-1} \xrightarrow{f_{n-1}} U_n \xrightarrow{f_n} U_{n+1} \xrightarrow{f_{n+1}} U_{n+2} \xrightarrow{f_{n+2}} \dots \quad (24.3.1)$$

が各 $n \in \mathbb{Z}$ について $\text{Im } f_n = \text{Ker } f_{n+1}$ を満たすとき, これを完全系列 (exact sequence) という. また, 次のような特別な線形写像の列

$$\dots \xrightarrow{\text{id}} \{\mathbf{0}\} \xrightarrow{\text{id}} \{\mathbf{0}\} \xrightarrow{f_0} W \xrightarrow{f_1} U \xrightarrow{f_2} V \xrightarrow{f_3} \{\mathbf{0}\} \xrightarrow{\text{id}} \{\mathbf{0}\} \xrightarrow{\text{id}} \dots$$

を略して次のように書く⁵⁸:

$$0 \xrightarrow{f_0} W \xrightarrow{f_1} U \xrightarrow{f_2} V \xrightarrow{f_3} 0 \quad (24.3.2)$$

この列が完全系列であるとき, これを短完全列 (short exact sequence) という.

列 24.3.1 が完全系列であるとき, 各 $\mathbf{u} \in U_n$ について $f_n(\mathbf{u}) \in \text{Im } f_n = \text{Ker } f_{n+1}$ ゆえ $f_{n+1}(f_n(\mathbf{u})) = \mathbf{0}_{U_{n+2}}$ である. したがって合成写像 $f_{n+1} \circ f_n: U_n \rightarrow U_{n+2}$ は, 定義域のすべての元を零元に写す写像である (つまり $f_{n+1} \circ f_n = \mathbf{0}_{U_{n+2}}$). また, 線形写像の列 24.3.2 において, 両端の写像 f_0 および f_3 はともに自明な写像ゆえ記号を割りふらないことが多い. f_0 は零元を零元に写す線形写像であり (つまり $f_0 = \mathbf{0}_W$), f_3 はすべての元を零元に写す線形写像である (つまり $f_3 = \mathbf{0}$).

短完全列の定義から次の性質が直ちに従う.

⁵⁸代数学では, 自明な空間 $\{\mathbf{0}\}$ を 0 と略するのが慣例となっている.

補題 24.3.2. 線形写像の列 24.3.2 において次が成り立つ. とくに列 24.3.2 が短完全列であるとき, これらの性質が成り立っている.

(1) $\text{Im } f_0 = \text{Ker } f_1 \iff f_1 : W \rightarrow U$ は単射である (つまり W と $\text{Im } f_1$ は同型).

(2) $\text{Im } f_2 = \text{Ker } f_3 \iff f_2 : U \rightarrow V$ は全射である.

Proof. (1): $\text{Im } f_0 = \{\mathbf{0}_W\}$ ゆえ, $\text{Im } f_0 = \text{Ker } f_1 \iff \text{Ker } f_1 = \{\mathbf{0}_W\} \iff f_1$ は単射.

(2): $\text{Ker } f_3 = V$ ゆえ, $\text{Im } f_2 = \text{Ker } f_3 \iff \text{Im } f_2 = V \iff f_2$ は全射. □

線形写像があると, 対応する短完全列を与えることができる:

例 24.3.3. (1) 任意の線形写像 $f : U \rightarrow V$ について, 次は短完全列である.

$$0 \longrightarrow \text{Ker } f \xrightarrow{\text{id}} U \xrightarrow{f} \text{Im } f \longrightarrow 0$$

(2) U の部分空間 W および商写像 $q : U \rightarrow U/W$ について, 次は短完全列である.

$$0 \longrightarrow W \xrightarrow{\text{id}} U \xrightarrow{q} U/W \longrightarrow 0$$

Proof. (1) で与えられた写像列が短完全列であることを示すには, 前補題より次の三つの性質: $\text{id} : \text{Ker } f \rightarrow U$ が単射であること, および $\text{id}(\text{Ker } f) = \text{Ker } f$, $f : U \rightarrow \text{Im } f$ が全射であることを示せばよい. しかし, これらはすべて明らかである. また, 商写像 q について (1) を適用すると, 命題 24.1.2 より $\text{Ker } q = W$ ゆえ (2) を得る. □

短完全列において準同型定理を適用してみよう.

命題 24.3.4. 短完全列

$$0 \longrightarrow W \xrightarrow{f_1} U \xrightarrow{f_2} V \longrightarrow 0$$

において, f_1 の単射性より W は U の部分空間とみなすことができる ($W \simeq \text{Im } f_1 \subset U$). そこで W と $\text{Im } f_1$ を同一視すれば, U/W と V は同型である.

Proof. $\text{Im } f_1 = \text{Ker } f_2$ および f_2 の全射性に注意すると, f_2 に関する準同型定理により次を得る:

$$V = \text{Im } f_2 \simeq U / \text{Ker } f_2 = U / \text{Im } f_1 = U/W$$

. □

完全系列の標語的な解釈

完全系列の特別な場合である短完全列

$$0 \longrightarrow W \xrightarrow{f_1} U \xrightarrow{f_2} V \longrightarrow 0$$

を次元公式の文脈で読んでみよう. 全射線形写像 $f_2 : U \rightarrow V$ に関する次元公式により,

$$\begin{aligned} \dim U &= \dim \text{Ker } f_2 + \dim \text{Im } f_2 \\ &= \dim \text{Im } f_1 + \dim \text{Im } f_2 = \dim W + \dim V. \end{aligned}$$

この式から, 短完全列の中央に位置する空間 U が左右の空間 W および V によって分解されていると見れる. 実際, w_1, \dots, w_n を W の基底, v_1, \dots, v_m を V の基底とすれば, f_2 の全射性より $f(u_i) = v_i$ ($i = 1, \dots, m$) を満たす $u_i \in U$ が取れる. このとき, 定理 23.2.2 により $u_1, \dots, u_n, f_1(w_1), \dots, f_1(w_n)$ は U の基底である. すなわち, W の基底と V の基底を通して U の基底を与えることができる.

また, 例 24.3.3(2) および命題 24.3.4 によれば, 短完全列を与えることは商空間を与えることの言い換えに他ならない. 商空間を与えることは, 線形空間の各元を互いに交わらないグループに分類することを

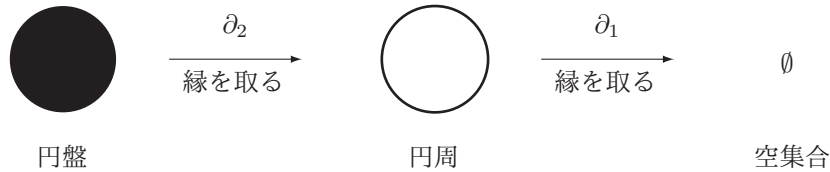
意味し、これは文字通りの空間の分解である。短完全列から導かれる分解を荒っぽく述べれば、 V の元の個数ぶんの W と合同な図形によって U は分解されている。

以上のことから、短完全列は代数構造の分解を表す図式であることが分かる。一般の完全系列についても、短完全列ほどの単純明快さはないものの、ある種の分解を与えていると考えられよう。実際、複雑な現象をより単純な対象に分解して理解するという人間の基本的な考え方のもとで、代数学を援用する多くの数学分野において完全系列が扱われている。

よりみち (合成すると消える写像).

完全系列が満たすべき性質の一つに $f_{n+1} \circ f_n = \mathbf{0}$ がある. このように合成すると消えてしまう写像列の例について, ここで二つほど紹介しておこう.

ある空間上の基本的な図形に対して, その縁 (境界) を対応させる操作を考える. 一般に n 次元の図形の縁は $n-1$ 次元になる. つまり, n 次元の各図形 Δ に対して, その縁を $\partial_n(\Delta)$ と書けば, $\partial_n(\Delta)$ は $n-1$ 次元の図形である. 例えば円盤に対してこの操作を二回ほどこすと次のようになる:



他の多くの図形に対しても縁を対応させる操作を 2 回繰り返すと消えてしまうことが分かり, 空集合に対応する図形を $\mathbf{0}$ と書くとすれば, $\partial_{n-1} \circ \partial_n = \mathbf{0}$ を得る.

上の対応 ∂_n を線形代数の文脈に無理矢理持ち込むこともできる. 空間 X に配置された n 次元の各図形を基底とする線形空間を $C_n(X)$ とすると, 基底の間の写像を線形写像として拡張することにより (命題 20.1.8 あるいは 22.5.2), 線形写像 $\partial_n : C_n(X) \rightarrow C_{n-1}(X)$ を得る (これを境界作用素と呼ぶ). このとき線形写像の列

$$\cdots \xrightarrow{\partial_{n+2}} C_{n+1}(X) \xrightarrow{\partial_{n+1}} C_n(X) \xrightarrow{\partial_n} C_{n-1}(X) \xrightarrow{\partial_{n-1}} \cdots$$

は $\partial_n \circ \partial_{n+1} = \mathbf{0}$ を満たす. しかしながら, 一般には上の列は完全系列にはならず (つまり $\text{Im } \partial_{n+1} \neq \text{Ker } \partial_n$), $\text{Im } \partial_{n+1} \subset \text{Ker } \partial_n$ であるに過ぎない. そこで, 上の写像の列がどれだけ完全系列から離れているかを調べる指標として商空間 $H_n(X) = \text{Ker } \partial_n / \text{Im } \partial_{n+1}$ が与えられる. $H_n(X)$ はホモロジー群と呼ばれ, 空間 X にどのくらい穴があいているかを計る量であることが知られている.

ところで, 合成すると消える写像のうち, 理工系学科の教育で最初に学ぶものといえば何であろうか. おそらくそれは電磁気学で学ぶ次の式である.

$$\text{rot} \circ \text{grad } f = \mathbf{0}, \quad \text{div} \circ \text{rot } F = \mathbf{0}. \quad (24.3.3)$$

ここで, 勾配ベクトル場 $\text{grad } f$ およびベクトル場の回転 $\text{rot } F$, ベクトル場の発散 $\text{div } g$ は次で定められるのであった:

$$\begin{aligned} \text{grad } f &:= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right), \\ \text{rot } F &:= \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right), \\ \text{div } g &:= \frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z}, \end{aligned}$$

ただし, 上に現れる関数はすべて変数 $(x, y, z) \in \mathbb{R}^3$ をパラメータとし, f は C^∞ -級関数, $F = (f_1, f_2, f_3)$ および $g = (g_1, g_2, g_3)$ は C^∞ -級ベクトル場である. 実は, これらの写像は図形の縁を対応させる写像とも無縁ではない. 勾配および回転, 発散は外微分と呼ばれる線形写像によって統一的に記述されることを後にベクトル解析を通して学ぶことになる. そして, 境界作用素と外微分の関係を行列の転置操作 (双対性) と関連づけて理解することになる (ド・ラームの定理). これ以上詳しいことは, 多様体論および微分形式の専門書を参照されたい.

練習 24.3.5. 上の定義をもとに式 24.3.3 を確認せよ.

ヒント: C^2 -級関数ゆえ偏微分の順序交換ができること (例えば $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$) を用いよ.

25 線形結合の行列表示

n 次元線形空間 U はユークリッド空間 \mathbb{R}^n と同型であったゆえ、 U において述べられる線形空間に関する現象はすべて線形同型 $T: U \rightarrow \mathbb{R}^n$ を用いて \mathbb{R}^n の現象に書き換えられるはずである。本節では、 U の上での線形結合に関する情報を \mathbb{R}^n の情報に読みかえる技術について解説する。

25.1 線形結合の組と行列

ベクトルの間の線形結合を繰り返し計算し続けると \sum 記号が二重三重に現れ、読む側には非常に複雑に見えてしまう。これを回避する手段として、ベクトルの線形結合を行列を用いて表す記法を導入しよう。

\mathbb{R}^m の n 個のベクトルの組 $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ について $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ は (m, n) -行列ゆえ (n, ℓ) -行列 $A = [a_{ij}]$ との積を取ることができる:

$$[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\ell} \\ a_{21} & a_{22} & \dots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\ell} \end{bmatrix} = \left[\sum_{i=1}^n a_{i1} \mathbf{u}_i \quad \sum_{i=1}^n a_{i2} \mathbf{u}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell} \mathbf{u}_i \right].$$

上の右辺には A の第 j 列成分を係数とする $\mathbf{u}_1, \dots, \mathbf{u}_n$ による線形結合の ℓ 個の組が現れている。そこで、 \mathbb{R}^m の元とは限らない一般の線形空間 U におけるベクトルを並べた列⁵⁹ $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ と (n, ℓ) -行列 $A = [a_{ij}]$ に対して、これらの積を上と同様に定める。

$$[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1\ell} \\ a_{21} & a_{22} & \dots & a_{2\ell} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{n\ell} \end{bmatrix} := \left[\sum_{i=1}^n a_{i1} \mathbf{v}_i \quad \sum_{i=1}^n a_{i2} \mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell} \mathbf{v}_i \right]. \quad (25.1.1)$$

より抽象的に書けば、

$$[\mathbf{v}_h]_{h=1, \dots, n} [a_{ij}]_{\substack{i=1, \dots, n, \\ j=1, \dots, \ell}} = \left[\sum_{i=1}^n a_{ij} \mathbf{v}_i \right]_{j=1, \dots, \ell}.$$

式 25.1.1 における $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ および右辺は U 上のベクトルを並べた組であって、 $U = \mathbb{R}^m$ でない限りこれらは行列ではないことに注意せよ。また、 A が列ベクトルである場合は、ベクトルの組と A の積は線形結合を意味する⁶⁰:

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^n x_i \mathbf{u}_i.$$

例 25.1.1. (1) $[\mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} 3 & 2 & 1 \\ 1 & -1 & 4 \end{bmatrix} = [3\mathbf{v}_1 + \mathbf{v}_2 \quad 2\mathbf{v}_1 - \mathbf{v}_2 \quad \mathbf{v}_1 + 4\mathbf{v}_2].$

(2) $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] E = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n].$

ベクトルの組における和とスカラー倍も定めておこう:

- $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] + [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n] := [\mathbf{v}_1 + \mathbf{w}_1, \mathbf{v}_2 + \mathbf{w}_2, \dots, \mathbf{v}_n + \mathbf{w}_n].$
- $r[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n] := [r\mathbf{v}_1, r\mathbf{v}_2, \dots, r\mathbf{v}_n].$

⁵⁹ ベクトルを並べる際にカンマで区切るかどうかはこだわらないことにする。

⁶⁰ $\sum_{i=1}^n \mathbf{u}_i x_i$ と書いてもよいが、慣例ではスカラー係数を左側に書く。このような事情から、ベクトルのスカラー係数を右側に書く流儀もある。

25.2 ベクトルの組と行列の演算の基本性質 (付録)

前項で定めたベクトルの組と行列の積が分配法則や結合法則, スカラー律を満たすことを退屈ではあるが確認しておこう. 実際の計算では, これらの性質は意識することなく用いられるであろう.

命題 25.2.1. 線形空間 V のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$, および (n, ℓ) -行列 A , $r \in \mathbb{R}$ に対して

$$[\mathbf{v}_1, \dots, \mathbf{v}_n](rA) = (r[\mathbf{v}_1, \dots, \mathbf{v}_n])A = r([\mathbf{v}_1, \dots, \mathbf{v}_n]A).$$

Proof. $A = [a_{ij}]$ とおくと $rA = [ra_{ij}]$ であり,

$$\begin{aligned} (\text{左辺}) &= \left[\sum_{i=1}^n ra_{i1}\mathbf{v}_i \quad \sum_{i=1}^n ra_{i2}\mathbf{v}_i \quad \dots, \quad \sum_{i=1}^n ra_{i\ell}\mathbf{v}_i \right] \\ &= \left[\sum_{i=1}^n a_{i1}(r\mathbf{v}_i) \quad \sum_{i=1}^n a_{i2}(r\mathbf{v}_i) \quad \dots, \quad \sum_{i=1}^n a_{i\ell}(r\mathbf{v}_i) \right] \quad (\text{この式は上の中央の式に相当する}) \\ &= \left[r \sum_{i=1}^n a_{i1}\mathbf{v}_i \quad r \sum_{i=1}^n a_{i2}\mathbf{v}_i \quad \dots, \quad r \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] \\ &= r \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i \quad \sum_{i=1}^n a_{i2}\mathbf{v}_i \quad \dots, \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] = (\text{右辺}). \end{aligned}$$

□

命題 25.2.2. 線形空間 V のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$, および (n, ℓ) -行列 A, B に対して

$$[\mathbf{v}_1, \dots, \mathbf{v}_n]A + [\mathbf{v}_1, \dots, \mathbf{v}_n]B = [\mathbf{v}_1, \dots, \mathbf{v}_n](A + B).$$

Proof. $A = [a_{ij}]$, $B = [b_{ij}]$ とおくと,

$$\begin{aligned} (\text{左辺}) &= \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i \quad \sum_{i=1}^n a_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] + \left[\sum_{i=1}^n b_{i1}\mathbf{v}_i \quad \sum_{i=1}^n b_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n b_{i\ell}\mathbf{v}_i \right] \\ &= \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i + \sum_{i=1}^n b_{i1}\mathbf{v}_i \quad \sum_{i=1}^n a_{i2}\mathbf{v}_i + \sum_{i=1}^n b_{i2}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i + \sum_{i=1}^n b_{i\ell}\mathbf{v}_i \right] \\ &= \left[\sum_{i=1}^n (a_{i1} + b_{i1})\mathbf{v}_i \quad \sum_{i=1}^n (a_{i2} + b_{i2})\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n (a_{i\ell} + b_{i\ell})\mathbf{v}_i \right] = (\text{右辺}) \end{aligned}$$

□

命題 25.2.3. 線形空間 V のベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n$, および (n, ℓ) -行列 A , (ℓ, r) -行列 B に対して

$$([\mathbf{v}_1, \dots, \mathbf{v}_n]A)B = [\mathbf{v}_1, \dots, \mathbf{v}_n](AB).$$

Proof. $A = [a_{ij}]$, $B = [b_{jk}]$ とおけば, (n, r) -行列 $AB = [z_{ik}]$ の各成分は行列の積の定義より $z_{ik} = \sum_{j=1}^{\ell} a_{ij}b_{jk}$ である.

$$\begin{aligned} (\text{左辺}) &= \left[\sum_{i=1}^n a_{i1}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^n a_{i\ell}\mathbf{v}_i \right] B \\ &= \left[\mathbf{u}_1 \quad \dots \quad \mathbf{u}_\ell \right] B \quad (\mathbf{u}_j := \sum_{i=1}^n a_{ij}\mathbf{v}_i \quad (j = 1, \dots, \ell) \text{ とおいた}) \\ &= \left[\sum_{j=1}^{\ell} b_{j1}\mathbf{u}_j \quad \dots \quad \sum_{j=1}^{\ell} b_{jr}\mathbf{u}_j \right] = \left[\sum_{j=1}^{\ell} b_{jk}\mathbf{u}_j \right]_{k=1, \dots, r} = \left[\sum_{j=1}^{\ell} b_{jk} \left(\sum_{i=1}^n a_{ij}\mathbf{v}_i \right) \right]_{k=1, \dots, r} \\ &= \left[\sum_{j=1}^{\ell} \left(\sum_{i=1}^n a_{ij}b_{jk}\mathbf{v}_i \right) \right]_{k=1, \dots, r} = \left[\sum_{i=1}^n \left(\sum_{j=1}^{\ell} a_{ij}b_{jk}\mathbf{v}_i \right) \right]_{k=1, \dots, r} \quad (\text{例 3.2.4 を用いた}) \\ &= \left[\sum_{i=1}^n \left(\sum_{j=1}^{\ell} a_{ij}b_{jk} \right) \mathbf{v}_i \right]_{k=1, \dots, r} = \left[\sum_{i=1}^n z_{ik}\mathbf{v}_i \right]_{k=1, \dots, r} = [\mathbf{v}_1, \dots, \mathbf{v}_n][z_{ik}] = (\text{右辺}). \end{aligned}$$

□

25.3 線形結合再考

線形結合に関するいくつかの基本的な性質について、25.1 項で定めた行列による記法を用いる立場から再考する。例えば、命題 18.1.3 は次のように示される。

命題 18.1.3 (再掲). 各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ が組 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で表され、また各 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ が組 $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で表されるならば、各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ は $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で表される。

Proof. 各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ が $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ の線形結合で表されることから、 (ℓ, r) -行列 B を用いて $[\mathbf{u}_1, \dots, \mathbf{u}_r] = [\mathbf{v}_1, \dots, \mathbf{v}_\ell]B$ と書ける。また、各 $\mathbf{v}_1, \dots, \mathbf{v}_\ell$ が $\mathbf{w}_1, \dots, \mathbf{w}_n$ 個の線形結合で表されることから、 (n, ℓ) -行列 A を用いて $[\mathbf{v}_1, \dots, \mathbf{v}_\ell] = [\mathbf{w}_1, \dots, \mathbf{w}_n]A$ と書ける。ゆえに、

$$[\mathbf{u}_1, \dots, \mathbf{u}_r] = [\mathbf{v}_1, \dots, \mathbf{v}_\ell]B = ([\mathbf{w}_1, \dots, \mathbf{w}_n]A)B = [\mathbf{w}_1, \dots, \mathbf{w}_n](AB).$$

この式は各 $\mathbf{u}_1, \dots, \mathbf{u}_r$ が $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で表されることを意味している。□

いまの証明から、 \mathbf{u}_k ($k = 1, \dots, r$) を $\mathbf{w}_1, \dots, \mathbf{w}_n$ の線形結合で書くときに現れる係数は、行列 AB の第 k 列成分に等しいことが分かる。

次は、各成分ごとに命題 18.2.3 を主張するものに他ならない。

練習 25.3.1. $\mathbf{v}_1, \dots, \mathbf{v}_m \in U$ が線形独立であるとし、 A, B を (m, n) -行列とする。このとき、 $[\mathbf{v}_1, \dots, \mathbf{v}_m]A = [\mathbf{v}_1, \dots, \mathbf{v}_m]B$ ならば $A = B$ となることを示せ。

解答例: 移項して $[\mathbf{v}_1, \dots, \mathbf{v}_m](A - B) = [\mathbf{0}_V, \dots, \mathbf{0}_V]$ 。 $A - B = [x_{ij}]$ とおき、各成分を比較すれば $j = 1, \dots, n$ について $\sum_{i=1}^m x_{ij}\mathbf{v}_i = \mathbf{0}_V$ であり、 $\mathbf{v}_1, \dots, \mathbf{v}_m$ の線形独立性より $x_{ij} = 0$ 、つまり $A = B = O$ を得る。□

次は、各成分ごとに線形性 (iii)' を並べた式である。

練習 25.3.2. A を (m, n) -行列、 $f: U \rightarrow W$ を線形写像とする。 $\mathbf{u}_j, \mathbf{v}_i \in U$ について次を示せ。

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A \implies [f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [f(\mathbf{v}_1), \dots, f(\mathbf{v}_m)]A$$

解答例: $A = [a_{ij}]$ とすれば、仮定より、

$$[\mathbf{u}_1, \dots, \mathbf{u}_n] = \left[\sum_{i=1}^m a_{i1}\mathbf{v}_i \quad \dots \quad \sum_{i=1}^m a_{in}\mathbf{v}_i \right] \quad \text{であるから、}$$

$$\begin{aligned} [f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] &= \left[f\left(\sum_{i=1}^m a_{i1}\mathbf{v}_i\right) \quad \dots \quad f\left(\sum_{i=1}^m a_{in}\mathbf{v}_i\right) \right] \\ &= \left[\sum_{i=1}^m a_{i1}f(\mathbf{v}_i) \quad \dots \quad \sum_{i=1}^m a_{in}f(\mathbf{v}_i) \right] = [f(\mathbf{v}_1), \dots, f(\mathbf{v}_m)]A. \end{aligned}$$

□

次の命題は、命題 17.3.5 を一般の線形空間に拡張したものである。

命題 25.3.3. $\mathbf{v}_1, \dots, \mathbf{v}_n \in \langle \mathbf{u}_1, \dots, \mathbf{u}_m \rangle$ とする。このとき、 $m < n$ ならば $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形従属である。

Proof. 仮定より (m, n) -行列 A を用いて $[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_m]A$ と書ける。 $n > m$ より連立 1 次方程式 $A\mathbf{x} = \mathbf{0}$ は自明でない解 ${}^t[c_1, \dots, c_n] \neq \mathbf{0}$ を持つ。このとき、

$$\begin{aligned} c_1\mathbf{v}_1 + \dots + c_n\mathbf{v}_n &= [\mathbf{v}_1, \dots, \mathbf{v}_n] \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = [\mathbf{u}_1, \dots, \mathbf{u}_m]A \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \\ &= [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} = 0\mathbf{u}_1 + \dots + 0\mathbf{u}_m = \mathbf{0}. \end{aligned}$$

自明でない線形関係 $\sum_{j=1}^n c_j\mathbf{v}_j = \mathbf{0}$ があるゆえ $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形従属である。□

練習 25.3.4. 線形空間 U の基底におけるベクトルの個数が基底の取り方によらないこと (命題 22.1.1) を前命題を用いて示せ.

解答例: $\mathbf{u}_1, \dots, \mathbf{u}_m$ および $\mathbf{v}_1, \dots, \mathbf{v}_n$ をそれぞれ U の基底とする. 仮に $m < n$ とすれば前命題より $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形従属になってしまう. ゆえに $m \geq n$ でなければならない. $\mathbf{u}_1, \dots, \mathbf{u}_m$ と $\mathbf{v}_1, \dots, \mathbf{v}_n$ の立場を入れ替えることで $n \geq m$ も示され, したがって $m = n$. \square

25.4 線形独立性の判定 (2)

一般のベクトル空間 U における線形結合を行列表示することで, U における線形関係をユークリッド空間上の線形関係に読み替えられることを見よう. $\mathbf{u}_1, \dots, \mathbf{u}_m$ を U の基底とし, $F(\mathbf{u}_j) = \mathbf{e}_j$ なる線形同型 $F: U \rightarrow \mathbb{R}^m$ を取る.

$$\begin{array}{ccc} U & : & \mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m & & \mathbf{v} = \sum_{i=1}^m r_i \mathbf{u}_i \\ F \downarrow & & \downarrow \downarrow \downarrow & & \downarrow \\ \mathbb{R}^m & : & \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m & & \sum_{i=1}^m r_i \mathbf{e}_i = {}^t[r_1, \dots, r_m] \end{array}$$

U において調べたいベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in U$ が与えられたとき, これに対応する $T(\mathbf{v}_1), \dots, T(\mathbf{v}_n) \in \mathbb{R}^m$ を調べるというのが基本方針である. いま $\mathbf{u}_1, \dots, \mathbf{u}_m$ は基底であるから各 $\mathbf{v}_1, \dots, \mathbf{v}_n$ は $\mathbf{u}_1, \dots, \mathbf{u}_m$ の線形結合で書ける:

$$\mathbf{v}_1 = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad \mathbf{v}_2 = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad \mathbf{v}_n = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

これをまとめて書くと, (m, n) -行列 $A = [a_{ij}] = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ を用いて

$$[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_m] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

このとき, $[F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)] = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ である. 実際, 上式に F をほどこせば, 練習 25.3.2 により

$$[F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)] = [F(\mathbf{u}_1), \dots, F(\mathbf{u}_m)]A = EA = A = [\mathbf{a}_1, \dots, \mathbf{a}_n].$$

命題 20.3.6 より $\mathbf{v}_1, \dots, \mathbf{v}_n$ の線形関係と $F(\mathbf{v}_1), \dots, F(\mathbf{v}_n)$ (すなわち $\mathbf{a}_1, \dots, \mathbf{a}_n$) の線形関係は同等である. つまり, 各 $\mathbf{w} \in U$ および $s_1, \dots, s_n \in \mathbb{R}$ について

$$\mathbf{w} = \sum_{j=1}^n s_j \mathbf{v}_j \iff F(\mathbf{w}) = \sum_{j=1}^n s_j \mathbf{a}_j. \quad (25.4.1)$$

が成り立つ. とくに, $n = m$ の場合は A は正方行列であり, 定理 17.3.6 により $\mathbf{a}_1, \dots, \mathbf{a}_n$ の線形独立性 (したがって $\mathbf{v}_1, \dots, \mathbf{v}_n$ の線形独立性) は A の可逆性に帰着される. 以上の事実をまとめると:

命題 25.4.1. $\mathbf{u}_1, \dots, \mathbf{u}_m$ を U の基底とする. ベクトルの組 $\mathbf{v}_1, \dots, \mathbf{v}_n \in U$ について $[\mathbf{v}_1, \dots, \mathbf{v}_n] = [\mathbf{u}_1, \dots, \mathbf{u}_m]A$ とすれば次が成り立つ.

(1) $A = [\mathbf{a}_1, \dots, \mathbf{a}_n]$ とおけば, $\sum_{i=1}^n c_i \mathbf{v}_i = \mathbf{0}_U \iff \sum_{i=1}^n c_i \mathbf{a}_i = \mathbf{0}$.

(2) $n = m$ のとき, $\mathbf{v}_1, \dots, \mathbf{v}_n$ は線形独立 (したがって U の基底である) $\iff A$ は可逆.

注意: (2) において v_1, \dots, v_n が U を生成することは命題 22.3.3 から直ちに分かることであるが, 次のように直接示すこともできる:

$[v_1, \dots, v_n] = [u_1, \dots, u_n]A$ の両辺に右から A^{-1} を掛ければ $[u_1, \dots, u_n] = [v_1, \dots, v_n]A^{-1}$. したがって, 基底 u_1, \dots, u_n の各ベクトルは v_1, \dots, v_n の線形結合で書けるゆえ, 命題 18.1.3 より v_1, \dots, v_n は U を生成する.

本項で述べた事実を用いて, 一般の線形空間におけるベクトルの組の線形独立性を判定できる:

例題 25.4.2. u_1, \dots, u_5 を U の基底とする. 次で与えるベクトルの組 v_1, \dots, v_5 が線形独立かどうか答えよ. また, これらのベクトルで生成される部分空間 $\langle v_1, \dots, v_5 \rangle$ の基底を求めよ.

$$\begin{aligned} v_1 &= u_1 - u_3 + 2u_4 + u_5, & v_2 &= u_2 + u_4 + u_5, & v_3 &= -u_1 + u_2 + u_3 - u_4, \\ v_4 &= u_3, & v_5 &= -2u_1 + u_2 + u_3 - 3u_4 - u_5. \end{aligned}$$

解答例: 上の線形結合を行列を用いた表示に書き直せば,

$$[v_1, \dots, v_5] = [u_1, \dots, u_5] \begin{bmatrix} 1 & 0 & -1 & 0 & -2 \\ 0 & 1 & 1 & 0 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ 2 & 1 & -1 & 0 & -3 \\ 1 & 1 & 0 & 0 & -1 \end{bmatrix}.$$

上式に現れる行列は例題 17.3.3 における $A = [a_1, \dots, a_5]$ に一致している. a_1, \dots, a_5 が線形独立でないことは例題 17.3.3 の通りであり, したがって, 命題 25.4.1(1) より v_1, \dots, v_5 は線形独立でない.

また, 例題 17.3.3 によれば a_1, a_2, a_4 は線形独立であり, これ以外のベクトルは $a_3 = -a_1 + a_2$, $a_5 = -2a_1 + a_2 - a_4$ と書ける. したがって, 再び命題 25.4.1(1) より v_1, v_2, v_4 は線形独立であり, 更に式 (25.4.1) から $v_3 = -v_1 + v_2$, $v_5 = -2v_1 + v_2 - v_4$ と書けることが分かる. 以上により, v_1, v_2, v_4 は $\langle v_1, \dots, v_5 \rangle$ の基底である.

25.5 基底の変換行列

前項にて, U 上の各ベクトル v を数値化する方法を述べた. ここで注意すべきは, 数値化された情報は線形同型 F の取り方に依存すること, 言い換えれば U の基底の取り方に依存することである. 基底を取りかえるたびに以前の基底による情報が役に立たなくなるようで甚だ不便であるから, 基底の取り換えにより v の数値化がどうに変化するかを調べておこう.

いま, U は n 次元であるとし, U において二組の基底 u_1, \dots, u_n および u'_1, \dots, u'_n が与えられているとする. このとき, 各 $u'_1, \dots, u'_n \in U$ は基底 u_1, \dots, u_n の線形結合で書ける. これを式で書けば:

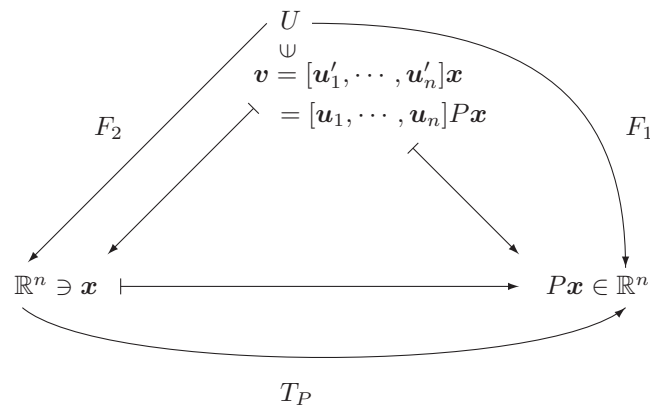
$$\begin{aligned} [u'_1, u'_2, \dots, u'_n] &= \left[\sum_{i=1}^n a_{i1} u_i \quad \sum_{i=1}^n a_{i2} u_i \quad \dots \quad \sum_{i=1}^n a_{in} u_i \right] \\ &= [u_1, u_2, \dots, u_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}. \end{aligned}$$

定義 25.5.1. 上式に現れる行列 $P = [a_{ij}]$ を基底 u_1, \dots, u_n による基底 u'_1, \dots, u'_n の変換行列という.

備考 25.5.2. 命題 25.4.1(2) により変換行列 P は可逆である. そこで $[u'_1, u'_2, \dots, u'_n] = [u_1, u_2, \dots, u_n]P$ の両辺に右から P^{-1} をかけることで $[u_1, u_2, \dots, u_n] = [u'_1, u'_2, \dots, u'_n]P^{-1}$ 得る. すなわち, 基底 u'_1, \dots, u'_n による基底 u_1, \dots, u_n の変換行列は P^{-1} である.

例 25.5.3. $w_1, \dots, w_n \in \mathbb{R}^n$ を基底とする. 標準基底 e_1, \dots, e_n による基底 w_1, \dots, w_n の変換行列は, $[w_1, \dots, w_n]$ である.

備考 25.5.4. 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とする. $F_1: U \rightarrow \mathbb{R}^n$ を \mathbf{u}_i を \mathbf{e}_i に写す線形同型とし, $F_2: U \rightarrow \mathbb{R}^n$ を \mathbf{u}'_i を \mathbf{e}_i に写す線形同型とする.



各ベクトル $\mathbf{v} = [\mathbf{u}'_1, \dots, \mathbf{u}'_n]\mathbf{x} \in U$ (ただし $\mathbf{x} \in \mathbb{R}^n$) に対応する \mathbb{R}^n の元は, 同型 F_2 を用いれば $F_2(\mathbf{v}) = \mathbf{x}$ であり, 同型 F_1 を用いれば, $\mathbf{v} = [\mathbf{u}'_1, \dots, \mathbf{u}'_n]\mathbf{x} = [\mathbf{u}_1, \dots, \mathbf{u}_n]P\mathbf{x}$ より $F_1(\mathbf{v}) = P\mathbf{x}$ である. このとき,

$$F_1 = T_P \circ F_2, \quad T_P = F_1 \circ F_2^{-1}, \quad F_2 = T_{P^{-1}} \circ F_1$$

が成り立つことは上の図式から直ちに分かる.

練習 25.5.5. 上の備考において $T_P = F_1 \circ F_2^{-1}$ であることを確認せよ.

解答例: 各標準基底 $\mathbf{e}_j \in \mathbb{R}^n$ について $F_1 \circ F_2^{-1}(\mathbf{e}_j) = T_P(\mathbf{e}_j)$ であることを確かめればよい. $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ と置けば, 変換行列の定義から $\mathbf{u}'_j = [\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{p}_j$ である. したがって,

$$\begin{aligned} F_1 \circ F_2^{-1}(\mathbf{e}_j) &= F_1(\mathbf{u}'_j) = F_1([\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{p}_j) = [F(\mathbf{u}_1), \dots, F(\mathbf{u}_n)]\mathbf{p}_j \\ &= [\mathbf{e}_1, \dots, \mathbf{e}_n]\mathbf{p}_j = E\mathbf{p}_j = \mathbf{p}_j = P\mathbf{e}_j = T_P(\mathbf{e}_j). \end{aligned}$$

□

26 線形写像の表現行列

線形写像 $f: U \rightarrow V$ をユークリッド空間の間の線形写像 (すなわち行列) に対応させる方法について述べる. 前節の対応と同様に, ここで与える対応も基底の取り方に依存することに注意しなければならない.

26.1 表現行列

U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$, V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ をあらかじめ与えておく. 線形写像 $T: U \rightarrow V$ を (m, n) -行列 A による線形写像 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ に翻訳する方法を考えよう. $\mathbf{v}_1, \dots, \mathbf{v}_m$ が基底であることから, 各 $T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)$ は $\mathbf{v}_1, \dots, \mathbf{v}_m$ の線形結合で書ける:

$$T(\mathbf{u}_1) = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad T(\mathbf{u}_2) = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad T(\mathbf{u}_n) = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix}.$$

つまり, (m, n) -行列 $A = [a_{ij}]$ を用いてまとめて書けば,

$$[T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}. \quad (26.1.1)$$

定義 26.1.1. 上の設定における式 (26.1.1) を満たす行列 $A = [a_{ij}]$ を, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ に関する $T: U \rightarrow V$ の表現行列または行列表示という. A の定義を次のように言い換えても良い:

$$A \text{ の第 } j \text{ 列} = \text{線形結合 } T(\mathbf{u}_j) = \sum_{i=1}^m a_{ij} \mathbf{v}_i \text{ に現れる係数を並べた列ベクトル.}$$

注意: 練習 25.3.1 により, 式 (26.1.1) を満たす行列 A は唯一つ存在する.

例 26.1.2. (m, n) -行列 B について, \mathbb{R}^n の標準基底および \mathbb{R}^m の標準基底に関する $T_B: \mathbb{R}^n \rightarrow \mathbb{R}^m$ ($T_B(\mathbf{x}) := B\mathbf{x}$) の表現行列は B に一致する.

例 26.1.3. (1) $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とする. 定義域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および終域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する恒等写像 $\text{id}_U: U \rightarrow U$ の表現行列は単位行列 E である.

(2) $\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とし, $\mathbf{v}_1, \dots, \mathbf{v}_n$ を V の基底とする. 各 \mathbf{u}_i を \mathbf{v}_i に対応させる線形同型 $f: U \rightarrow V$ について, U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に関する f の表現行列は単位行列 E である.

例題 26.1.4. 次の線形写像および基底について, 表現行列を求めよ.

$$(1) T_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \quad (T_A(\mathbf{x}) := A\mathbf{x}), \quad \text{ただし } A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}.$$

基底: 定義域と終域ともに $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ とせよ.

解答例:

$$\begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix} = -2\mathbf{v}_1 + 0\mathbf{v}_2, \quad \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 3 \end{bmatrix} = 0\mathbf{v}_1 + 3\mathbf{v}_2.$$

ゆえに $[T_A(\mathbf{v}_1), T_A(\mathbf{v}_2)] = [\mathbf{v}_1, \mathbf{v}_2] \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ であり, T_A の表現行列は $\begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$.

(2) $D: \mathbb{R}[x]_3 \rightarrow \mathbb{R}[x]_3$ ($D(p(x)) := \frac{d}{dx}p(x)$). 基底: 定義域と終域ともに $\mathbf{1}, x, x^2, x^3$ とせよ.

解答例:

$$\begin{aligned} D(\mathbf{1}) &= 0\mathbf{1} + 0x + 0x^2 + 0x^3, & D(x) &= 1 \cdot \mathbf{1} + 0x + 0x^2 + 0x^3, \\ D(x^2) &= 0\mathbf{1} + 2x + 0x^2 + 0x^3, & D(x^3) &= 0\mathbf{1} + 0x + 3x^2 + 0x^3. \end{aligned}$$

ゆえに

$$[D(\mathbf{1}), D(x), D(x^2), D(x^3)] = [\mathbf{1}, x, x^2, x^3] \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

であり, 求める表現行列は $\begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

(3) $I_{a,b}: \mathbb{R}[x]_3 \rightarrow \mathbb{R}$ ($I_{a,b}(p(x)) := \int_a^b p(x) dx$). ただし $a, b \in \mathbb{R}$ は定数とする.

$\mathbb{R}[x]_3$ の基底: $x^3, x^2, x, \mathbf{1}$. \mathbb{R} の基底: 1 .

解答例:

$$\begin{aligned} I_{a,b}(x^3) &= \left[\frac{1}{4}x^4 \right]_a^b = \frac{1}{4}(b-a)^4, & I_{a,b}(x^2) &= \left[\frac{1}{3}x^3 \right]_a^b = \frac{1}{3}(b-a)^2, \\ I_{a,b}(x) &= \left[\frac{1}{2}x^2 \right]_a^b = \frac{1}{2}(b-a)^2, & I_{a,b}(\mathbf{1}) &= \left[x \right]_a^b = (b-a) \end{aligned}$$

したがって

$$[I_{a,b}(x^3), I_{a,b}(x^2), I_{a,b}(x), I_{a,b}(\mathbf{1})] = [1] \left[\frac{1}{4}(b-a)^4, \frac{1}{3}(b-a)^2, \frac{1}{2}(b-a)^2, b-a \right]$$

であり, 求める表現行列は $(1, 4)$ -行列 $\left[\frac{1}{4}(b-a)^4, \frac{1}{3}(b-a)^2, \frac{1}{2}(b-a)^2, b-a \right]$.

備考: とくに $I_{a,b}(c_1x^3 + c_2x^2 + c_3x + c_4\mathbf{1}) = c_1I_{a,b}(x^3) + c_2I_{a,b}(x^2) + c_3I_{a,b}(x) + c_4I_{a,b}(\mathbf{1}) = \frac{c_1}{4}(b-a)^4 + \frac{c_2}{3}(b-a)^2 + \frac{c_3}{2}(b-a)^2 + c_4(b-a)$. 実際に積分計算を行う際も, このように各項ごとに積分するのであった.

例 26.1.5 (発展). 例 18.4.6 にて与えた, 数列空間 $\mathbb{R}^{\mathbb{N}}$ における線形独立な無限部分集合 $B = \{e_n | n \in \mathbb{N}\}$ について, $V = \langle B \rangle$ と定める. 次の二つのシフト作用素

$$\begin{aligned} S_+ : V &\rightarrow V & S_+(x_1, x_2, x_3, \dots) &:= (0, x_1, x_2, x_3, x_4, \dots) \\ S_- : V &\rightarrow V & S_-(x_1, x_2, x_3, \dots) &:= (x_2, x_3, x_4, x_5, x_6, \dots) \end{aligned}$$

の基底 B に関する表現行列は, 13.2 項のコラムで与えた A と B にそれぞれ等しい:

$$A = \begin{bmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{bmatrix}.$$

$S_- \circ S_+ = \text{id}_V$ ゆえ BA は単位行列となる. 一方, $S_+ \circ S_-(x_1, x_2, x_3, \dots) = (0, x_2, x_3, \dots)$ ゆえ $S_+ \circ S_- \neq \text{id}_V$. つまり AB は単位行列ではない. このようなことが起きる背景には, 基底が無限集合であること (つまり無限次元であること), そして有限集合の場合 (命題 19.5.1) とは異なり, 無限集合 B から B 自身への全射 (あるいは単射) が全単射になるとは限らないという事情がある. 実際, $S_-|_B : B \rightarrow B$ は全射であり $S_+|_B : B \rightarrow B$ は単射であるが, これらは全単射ではない.

表現行列の捉え方

$\mathbf{u}_1, \dots, \mathbf{u}_n$ を U の基底とし, $\mathbf{v}_1, \dots, \mathbf{v}_m$ を V の基底とする. また, これらの基底に関する線形写像 $T: U \rightarrow V$ の表現行列を A とする (つまり $[T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A$). $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$ とすれば, T はベクトル $[\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{x} \in U$ をベクトル $[\mathbf{v}_1, \dots, \mathbf{v}_m]A\mathbf{x} \in V$ に写す写像である. 実際, 線形性 (iii)' より

$$T([\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{x}) = [T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)]\mathbf{x} = [\mathbf{v}_1, \dots, \mathbf{v}_m]A\mathbf{x}.$$

つまり, 次のような図式を得る:

$$\begin{array}{ccccccc} U & \ni & [\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{x} & \xrightarrow{T} & [\mathbf{v}_1, \dots, \mathbf{v}_m]A\mathbf{x} & \in & V \\ F \downarrow & & \downarrow & & \downarrow & & \downarrow G \\ \mathbb{R}^n & \ni & \mathbf{x} & \xrightarrow{T_A} & A\mathbf{x} & \in & \mathbb{R}^m \end{array}$$

ここで, $F: U \rightarrow \mathbb{R}^n$ は $F(\mathbf{u}_j) = \mathbf{e}_j$ を満たす線形同型写像, $G: V \rightarrow \mathbb{R}^m$ は $G(\mathbf{v}_i) = \mathbf{e}_i$ を満たす線形同型写像である. $\mathbf{u}_1, \dots, \mathbf{u}_n$ が U の基底であることから, U の任意の元は $[\mathbf{u}_1, \dots, \mathbf{u}_n]\mathbf{x}$ の形で表せることに注意すれば, 上の図式から直ちに次を得る:

$$\text{各 } \mathbf{u} \in U \text{ について, } T_A \circ F(\mathbf{u}) = G \circ T(\mathbf{u}), \quad \text{すなわち, } T_A \circ F = G \circ T.$$

$T_A \circ F = G \circ T$ が成立するとき, これを次のような図式で表す.

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ F \downarrow & & \downarrow G \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^m \end{array} \quad (26.1.2)$$

このような図式は可換図式と呼ばれる.

26.2 Hom(U, V) と $M_{m,n}(\mathbb{R})$

$\text{Hom}(\mathbb{R}^n, \mathbb{R}^m)$ と $M_{m,n}(\mathbb{R})$ が同型であることを 21.3 項にて既に見たが, これらが $\text{Hom}(U, V)$ (ただし $\dim U = n, \dim V = m$) と同型になることも容易に示唆される. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ を与えておき, これらの基底による表現行列を対応させる写像 $\mathcal{T}: \text{Hom}(U, V) \rightarrow M_{m,n}(\mathbb{R})$ が線形同型を与えることを見よう.

各線形写像 $f: U \rightarrow V$ の表現行列を $\mathcal{T}(f)$ とする. つまり, $\mathcal{T}(f)$ は

$$[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(f)$$

を満たす (m, n) -行列である. また, これとは逆に, 各 (m, n) -行列 A に対して, 写像 $\mathcal{S}(A): U \rightarrow V$ を,

$$[\mathcal{S}(A)(\mathbf{u}_1), \dots, \mathcal{S}(A)(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A$$

を満たす線形写像と定める (このような線形写像は命題 20.1.8 により存在する). 次は定義より明らかである:

- 各 $A \in M_{m,n}(\mathbb{R})$ について $\mathcal{T}(\mathcal{S}(A)) = A$ (つまり $\mathcal{T} \circ \mathcal{S} = \text{id}_{M_{m,n}(\mathbb{R})}$).

Proof. $\mathcal{S}(A): U \rightarrow V$ は, $[\mathcal{S}(A)(\mathbf{u}_1), \dots, \mathcal{S}(A)(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A$ を満たす線形写像であり, この写像 $\mathcal{S}(A)$ の表現行列は A である. すなわち, $\mathcal{T}(\mathcal{S}(A)) = A$. □

- 各 $f \in \text{Hom}(U, V)$ について $\mathcal{S}(\mathcal{T}(f)) = f$ (つまり $\mathcal{S} \circ \mathcal{T} = \text{id}_{\text{Hom}(U, V)}$).

Proof. 行列 $\mathcal{T}(f)$ とは, $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(f)$ を満たす行列である. 一方, $\mathcal{S}(\mathcal{T}(f)) : U \rightarrow V$ は $[\mathcal{S}(\mathcal{T}(f))(\mathbf{u}_1), \dots, \mathcal{S}(\mathcal{T}(f))(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(f)$ を満たす線形写像であり,

$$[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathcal{S}(\mathcal{T}(f))(\mathbf{u}_1), \dots, \mathcal{S}(\mathcal{T}(f))(\mathbf{u}_n)]$$

となる. $\mathcal{S}(\mathcal{T}(f))$ と f に U の各基底 \mathbf{u}_j を代入した値はそれぞれ等しいゆえ, $\mathcal{S}(\mathcal{T}(f)) = f$. \square

したがって, 命題 19.3.4 より \mathcal{T}, \mathcal{S} は全単射かつ $\mathcal{T}^{-1} = \mathcal{S}$ である. これらが線形写像であることは容易に確認できる:

練習 26.2.1. 上で定めた $\mathcal{T} : \text{Hom}(U, V) \rightarrow M_{m,n}(\mathbb{R})$ および $f, g \in \text{Hom}(U, V)$, $r, s \in \mathbb{R}$ について次を示せ.

$$\text{線形性 (iii): } \mathcal{T}(rf + sg) = r\mathcal{T}(f) + s\mathcal{T}(g).$$

解答例: $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(f)$ および $[g(\mathbf{u}_1), \dots, g(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(g)$ であり, このとき

$$\begin{aligned} [(rf + sg)(\mathbf{u}_1), \dots, (rf + sg)(\mathbf{u}_n)] &= r[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] + s[g(\mathbf{u}_1), \dots, g(\mathbf{u}_n)] \\ &= r[\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(f) + s[\mathbf{v}_1, \dots, \mathbf{v}_m]\mathcal{T}(g) \\ &= [\mathbf{v}_1, \dots, \mathbf{v}_m](r\mathcal{T}(f) + s\mathcal{T}(g)). \end{aligned}$$

以上より, 線形写像 $rf + sg$ の表現行列は $r\mathcal{T}(f) + s\mathcal{T}(g)$ である. すなわち, $\mathcal{T}(rf + sg) = r\mathcal{T}(f) + s\mathcal{T}(g)$. \square

線形写像の合成と表現行列の間には次の関係が成立する.

命題 26.2.2. $f : U \rightarrow V, g : V \rightarrow W$ を線形写像とし, U, V, W の基底をそれぞれ $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m, \mathbf{w}_1, \dots, \mathbf{w}_\ell$ とし, これらの基底に関する f, g の表現行列をそれぞれ A, B とする. このとき, これらの基底に関する $g \circ f$ の表現行列は BA である.

Proof. 仮定より,

$$[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_m]A, \quad [g(\mathbf{v}_1), \dots, g(\mathbf{v}_m)] = [\mathbf{w}_1, \dots, \mathbf{w}_\ell]B$$

である. 練習 25.3.2 に注意すれば,

$$[g(f(\mathbf{u}_1)), \dots, g(f(\mathbf{u}_n))] = [g(\mathbf{v}_1), \dots, g(\mathbf{v}_m)]A = [\mathbf{w}_1, \dots, \mathbf{w}_\ell]BA.$$

ゆえに $g \circ f$ の表現行列は BA である. \square

系 26.2.3. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および V の基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に関する線形同型 $f : U \rightarrow V$ の表現行列を A とすれば, これらの基底に関する $f^{-1} : V \rightarrow U$ の表現行列は A^{-1} である.

Proof. $W = U$ として前命題を適用する. B を $f^{-1} : V \rightarrow U$ の表現行列とすれば,

$$E = \text{“id}_U : U \rightarrow U \text{ の表現行列”} = \text{“}f^{-1} \circ f : U \rightarrow U \text{ の表現行列”} = BA.$$

同様にして $AB = E$ も示される. つまり, B は A の逆行列である. \square

系 26.2.3 は、表現行列の定義から直接に示すこともできる:

系 26.2.3 の別証明. $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_n]A$ に逆写像 f^{-1} をほどこせば、練習 25.3.2 より $[\mathbf{u}_1, \dots, \mathbf{u}_n] = [f^{-1}(\mathbf{v}_1), \dots, f^{-1}(\mathbf{v}_n)]A$ である. この両辺に右から A^{-1} をかけて $[f^{-1}(\mathbf{v}_1), \dots, f^{-1}(\mathbf{v}_n)] = [\mathbf{u}_1, \dots, \mathbf{u}_n]A^{-1}$ を得る. すなわち、 A^{-1} は f^{-1} の表現行列である. \square

$U = V$ とし、 U, V の基底として共に $\mathbf{u}_1, \dots, \mathbf{u}_n$ を取るとき、命題 26.2.2 は同型 $\mathcal{T} : \text{End}(U) \rightarrow M_n(\mathbb{R})$ が積演算とも整合的であること、すなわち $\mathcal{T}(gf) = \mathcal{T}(g)\mathcal{T}(f)$ を意味している. 命題 21.4.6 より直ちに次が従う.

系 26.2.4. $\mathcal{T} : \text{End}(U) \rightarrow M_n(\mathbb{R})$ を上で与えた同型とすれば、任意の多項式 $\Psi(t)$ および $f \in \text{End}(U)$ について $\mathcal{T}(\Psi(f)) = \Psi(\mathcal{T}(f))$. すなわち、 U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば、 $\Psi(f) \in \text{End}(U)$ の表現行列は $\Psi(A)$ である.

26.3 基底の取りかたによる表現行列の違い

線形写像 $T : U \rightarrow V$ の表現行列が基底の取り換えによってどう変化するか考察しよう.

$$\begin{aligned} U \text{ の基底: } & \mathbf{u}_1, \dots, \mathbf{u}_n \quad \text{および} \quad \mathbf{u}'_1, \dots, \mathbf{u}'_n, \\ V \text{ の基底: } & \mathbf{v}_1, \dots, \mathbf{v}_m \quad \text{および} \quad \mathbf{v}'_1, \dots, \mathbf{v}'_m, \end{aligned}$$

とし、 $\mathbf{u}_1, \dots, \mathbf{u}_n$ と $\mathbf{v}_1, \dots, \mathbf{v}_m$ に関する T の表現行列を A 、 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ と $\mathbf{v}'_1, \dots, \mathbf{v}'_m$ に関する T の表現行列を B とする. また、基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P 、基底 $\mathbf{v}_1, \dots, \mathbf{v}_m$ による基底 $\mathbf{v}'_1, \dots, \mathbf{v}'_m$ の変換行列を Q とする. このとき、

命題 26.3.1. 上の設定のもとで $B = Q^{-1}AP$.

この命題は、それぞれの表現行列に関する可換図式をまとめた次の可換図式からほとんど明らかとも言える.

$$\begin{array}{ccccc} \mathbb{R}^n & \xleftarrow{F_2} & U & \xrightarrow{F_1} & \mathbb{R}^n \\ T_B \downarrow & & T \downarrow & & \downarrow T_A \\ \mathbb{R}^m & \xleftarrow{G_2} & V & \xrightarrow{G_1} & \mathbb{R}^m \end{array} \quad (26.3.1)$$

ここで、 F_1, F_2, G_1, G_2 は次の対応を意味する線形同型写像である:

$$\begin{aligned} F_1: F_1(\mathbf{u}_j) &= \mathbf{e}_j \in \mathbb{R}^n \text{ を満たす写像,} & G_1: G_1(\mathbf{v}_i) &= \mathbf{e}_i \in \mathbb{R}^m \text{ を満たす写像,} \\ F_2: F_2(\mathbf{u}'_j) &= \mathbf{e}_j \in \mathbb{R}^n \text{ を満たす写像,} & G_2: G_2(\mathbf{v}'_i) &= \mathbf{e}_i \in \mathbb{R}^m \text{ を満たす写像.} \end{aligned}$$

上の図式において $F_1 \circ F_2^{-1} = T_P$ および $G_1 \circ G_2^{-1} = T_Q$ である. これは変換行列の定義から直ちに得られる (詳しくは備考 25.5.4 を見よ). 可換図式 (26.3.1) における左上の \mathbb{R}^n から左下の \mathbb{R}^m への写像が近道と遠回りでも同等なことから $T_B = T_{Q^{-1}} \circ T_A \circ T_P$ であり、これは $B = Q^{-1}AP$ であることに他ならない.

命題 26.3.1 の証明. $F_1 \circ F_2^{-1} = T_P$ および $G_1 \circ G_2^{-1} = T_Q$ であり、次の可換図式が成り立つ:

$$\begin{array}{ccc} \mathbb{R}^n & \xrightarrow{T_P} & \mathbb{R}^n \\ T_B \downarrow & & \downarrow T_A \\ \mathbb{R}^m & \xrightarrow{T_Q} & \mathbb{R}^m \end{array}$$

可換図式 (26.3.1) から上の可換図式が導かれること (つまり $T_A \circ T_P = T_Q \circ T_B$) はほとんど明らかではあるが、一応確認しておこう. $G_1 \circ T = T_A \circ F_1$ の両辺に左から G^{-1} を合成することで $T = G_1^{-1} \circ T_A \circ F_1$

を得る. また, 同様に $T = G_2^{-1} \circ T_B \circ F_2$ でもある. $G_1^{-1} \circ T_A \circ F_1 = G_2^{-1} \circ T_B \circ F_2$ の両辺に左から G_1 を右から F_2^{-1} それぞれ合成すると

$$\begin{aligned} G_1 \circ G_1^{-1} \circ T_A \circ F_1 \circ F_2^{-1} &= G_1 \circ G_2^{-1} \circ T_B \circ F_2 \circ F_2^{-1} \\ \text{id}_{\mathbb{R}^m} \circ T_A \circ T_P &= T_Q \circ T_B \circ \text{id}_{\mathbb{R}^n} \\ T_A \circ T_P &= T_Q \circ T_B. \end{aligned}$$

ゆえに命題 21.3.3 より $T_{AP} = T_{QB}$ および $AP = BQ$ を得る. この両辺に左から Q^{-1} をかけて $B = Q^{-1}AP$ である. \square

定理の主張への理解を促すため, 上では図式を用いて説明した. 計算結果が正しければそれでよいというのであれば, 次のような証明もある.

命題 26.3.1 の別証明. $[T(\mathbf{u}'_1), \dots, T(\mathbf{u}'_n)]$ を二通りの方法で計算する. $[\mathbf{u}'_1, \dots, \mathbf{u}'_n] = [\mathbf{u}_1, \dots, \mathbf{u}_n]P$ に T をほどこすと練習 25.3.2 より

$$[T(\mathbf{u}'_1), \dots, T(\mathbf{u}'_n)] = [T(\mathbf{u}_1), \dots, T(\mathbf{u}_n)]P = [\mathbf{v}_1, \dots, \mathbf{v}_m]AP.$$

一方で,

$$[T(\mathbf{u}'_1), \dots, T(\mathbf{u}'_n)] = [\mathbf{v}'_1, \dots, \mathbf{v}'_m]B = [\mathbf{v}_1, \dots, \mathbf{v}_m]QB.$$

以上より $[\mathbf{v}_1, \dots, \mathbf{v}_m]AP = [\mathbf{v}_1, \dots, \mathbf{v}_m]QB$ である. 練習 25.3.1 より $AP = QB$. したがって $Q^{-1}AP = B$. \square

線形変換 $T: U \rightarrow U$ において, T で写すことによりベクトルが元の位置からどう変化するかを見るのであれば, 定義域と終域において同一の基底を取っておくことが望ましい. そこで, 定義域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および終域の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する T の表現行列のことを, 以下では単に基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する T の表現行列と呼ぶことにしよう. 前命題の特別な場合として次を得る.

系 26.3.2. $T: U \rightarrow U$ を線形変換とし, U の二組の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ および $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ が与えられているとする. このとき, $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する T の表現行列を A , $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ に関する T の表現行列を B , $\mathbf{u}_1, \dots, \mathbf{u}_n$ による $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とすれば, $B = P^{-1}AP$ である.

定義 26.3.3. 同じサイズの正方行列 A, B において, $B = P^{-1}AP$ を満たす可逆行列 P が存在するとき, A と B は相似 (similar) であるという.

線形写像を分析する立場からは, 表現行列に現れる成分があまり複雑でないことが望ましい. そのためには基底を上手く選ぶ必要がある. 次節以降では, 与えられた線形変換と相性のよい基底, すなわち表現行列が複雑にならないような基底の探し方を考察する.

練習 26.3.4. 二つの行列が相似であるという関係は同値関係である. これを示せ.

26.4 1対1の対応と可換図式

一般の線形写像 T を表現行列 A を用いて調べるにあたって, T が持つ性質と T_A が持つ性質が同等であることを理解しておく必要がある. 例えば, 次のような性質を T が持つことと T_A が持つことは同値である:

単射性, 全射性, 像の次元が l , 核の次元が r .

これらの同値性が可換図式 (26.1.2) における全単射 F, G を通して導かれることを確認しておこう. 上の四つの性質のほかにも, $\lambda \in \mathbb{R}$ が固有値となること, および固有値 λ に関する固有空間の次元, あるいは広義固有空間の次元について同様の対応を次節以降で論ずることになる.

線形空間の枠組みの外でも通用する基本的な事実として, 次の命題が成り立つ.

命題 26.4.1. 次の可換図式が成り立っているとする (すなわち $G \circ T = S \circ F$ が成り立つ). ここで, U, V, R_n, R_m は集合であり (線形空間でなくてもよい), T, S, F, G は写像である.

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ F \downarrow & & \downarrow G \\ R_n & \xrightarrow{S} & R_m \end{array}$$

F および G が全単射であるとき次が成り立つ.

(1) $T = G^{-1} \circ S \circ F, \quad S = G \circ T \circ F^{-1}.$

(2) T は単射 $\iff S$ は単射, T は全射 $\iff S$ は全射, T は全単射 $\iff S$ は全単射.

(3) $X \subset U$ について, $T(X) = G^{-1}(S(F(X)))$.

(4) $Y \subset V$ について, $T^{-1}(Y) = F^{-1}(S^{-1}(G(Y)))$.

Proof. (1): $G \circ T = S \circ F$ の両辺に左から G^{-1} を合成すれば $T = G^{-1} \circ S \circ F$ を得る. また, 両辺に右から F^{-1} を合成すれば $S = G \circ T \circ F^{-1}$.

(2): 単射の合成は単射であること, および全射の合成が全射となることから, (1) より直ちに従う.

(3): 写像 $T = G^{-1} \circ S \circ F$ における X による像として, $T(X) = G^{-1} \circ S \circ F(X) = G^{-1}(S(F(X)))$ である.

(4): 逆写像 G^{-1} の Y による逆像について $(G^{-1})^{-1}(Y) = G(Y)$ であることに注意して, $T = G^{-1} \circ S \circ F$ における Y の逆像を取ると, 命題 19.4.5 より

$$T^{-1}(Y) = (G^{-1} \circ S \circ F)^{-1}(Y) = F^{-1}(S^{-1}((G^{-1})^{-1}(Y))) = F^{-1}(S^{-1}(G(Y))).$$

□

備考: (3) および (4) の証明に F の全単射性は実は不要であり, これらは G の単射性のみから導かれる:

G の単射性のみを用いた (3) と (4) の別証明.

(3): G の単射性より各 $B \subset V$ について $G^{-1}(G(B)) = B$ である (練習 19.2.5(1)). $S \circ F = G \circ T$ より $S \circ F(X) = G \circ T(X) = G(T(X))$. この両辺による G の逆像をとれば, $G^{-1}(S \circ F(X)) = G^{-1}(G(T(X))) = T(X)$.

(4): 両方の包含関係を示す. 各 $x \in T^{-1}(Y)$ に対して, $T(x) \in Y$ より $(S \circ F)(x) = G \circ T(x) = G(T(x)) \in G(Y)$. ゆえに $x \in (S \circ F)^{-1}(G(Y)) = F^{-1}(S^{-1}(G(Y)))$. 一方, 各 $w \in F^{-1}(S^{-1}(G(Y))) = (S \circ F)^{-1}(G(Y))$ に対して, $S \circ F(w) \in G(Y)$ より, ある $y \in Y$ を用いて $S \circ F(w) = G(y)$ と書ける. $G(T(w)) = G \circ T(w) = S \circ F(w) = G(y)$ より $G(T(w)) = G(y)$ であり, G の単射性から $T(w) = y \in Y$. したがって $w \in T^{-1}(Y)$ である. □

上の命題の応用として, 一般の線形写像 $T : U \rightarrow V$ の像と核が表現行列についての線形写像の像と核にそれぞれ対応することを見よう.

命題 26.4.2. U の基底 u_1, \dots, u_n および V の基底 v_1, \dots, v_m に関する $T : U \rightarrow V$ の表現行列を A とすれば, 次が成り立つ.

(1) $\text{Im } T \simeq \text{Im } T_A, \quad (2) \text{ Ker } T \simeq \text{Ker } T_A$

Proof. 可換図式 (26.1.2):

$$\begin{array}{ccc} U & \xrightarrow{T} & V \\ F \downarrow & & \downarrow G \\ \mathbb{R}^n & \xrightarrow{T_A} & \mathbb{R}^m \end{array}$$

について前命題を適用する.

(1): $X = U$ について前命題 (3) を適用すると

$$\operatorname{Im} T = T(U) = G^{-1}\left(T_A(F(U))\right) = G^{-1}\left(T_A(\mathbb{R}^n)\right) = G^{-1}(\operatorname{Im} T_A).$$

(↑ここで前命題 (3) を用いた)

すなわち、線形同型写像 $G^{-1} : \operatorname{Im} T_A \rightarrow \operatorname{Im} T$ が存在するゆえ、これらは同型である。

(2): $Y = \{\mathbf{0}_V\}$ について前命題 (4) を適用すれば、

$$\operatorname{Ker} T = T^{-1}(\{\mathbf{0}_V\}) = F^{-1}\left(T_A^{-1}(G(\{\mathbf{0}_V\}))\right) = F^{-1}\left(T_A^{-1}(\{\mathbf{0}\})\right) = F^{-1}(\operatorname{Ker} T_A).$$

(↑ここで前命題 (4) を用いた)

線形同型写像 $F^{-1} : \operatorname{Ker} T_A \rightarrow \operatorname{Ker} T$ により、これらは同型である。 □

$T_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ における次元定理は、例 23.2.5 でみたように連立 1 次方程式の解法との関係からすぐに分かる。これと上の主張を合わせることで、一般の写像 $f : U \rightarrow V$ における次元公式の別証明を得る。

27 固有値と固有ベクトル

特別な基底によるベクトルの分解を考えよう. 分解を行う一義的な理由は, 複雑なベクトルをより単純なベクトルに分解し理解を容易にすることにある. しかし, 単純なベクトルと一言でいっても, 何を基準に単純なのかという疑問もあろう. そこで, ここでは次の二通りの方法を提案しよう. 一つは, 標準的と考えられるベクトルへの分解であり, もう一つは与えられた線形写像と相性のよいベクトルへの分解である.

前者の最も典型的な例は \mathbb{R}^n のベクトルを標準基底に分解するというものである. また, 解析学で学ぶ関数の冪級数展開もそのような分解の一つと言えるだろう. 例えば, テーラーの定理の剰余項が 0 に収束することを確認することで, 次のような冪級数展開 (テイラー展開) を得る:

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}, \quad \cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n}, \quad e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

上の三つの冪級数は \mathbb{C} 上でも収束することが確認され, $x = i\theta$ に対してこれらを適用することで次の等式が導かれる. ここで, i は虚数単位 (すなわち $i^2 = -1$ を満たす数) とする.

定理 27.0.3 (オイラーの公式). $e^{i\theta} = \cos \theta + i \sin \theta$.

このように, 関数の多項式による分解によって三角関数と指数関数の相互理解が深められる. 一方で, 冪級数展開 $f(x) = \sum_{n=0}^{\infty} a_n x^n$ について次のような計算

$$\frac{d}{dx} f(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}, \quad \int f(x) dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} + C \quad (C \text{ は積分定数})$$

も期待できるのであるが, この計算は x の冪がずれるゆえ少々計算が複雑になる. n 階微分になれば複雑さは増す一方である.

そこで, 微分作用素とより相性の良い関数による分解が考えられる. 例えば, 周期 2π の関数⁶¹の多くは次のように分解できることが知られている (フーリエ展開)⁶²:

$$f(x) = \sum_{n=-\infty}^{\infty} z_n e^{inx}, \quad \text{ここで, 各 } z_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} dt \text{ は複素数.}$$

このとき, $\frac{d}{dx} e^{ax} = a e^{ax}$ が複素数値関数としても成り立つことをあらかじめ確認しておけば,

$$\frac{d}{dx} f(x) = \sum_{n=-\infty}^{\infty} i n z_n e^{inx}, \quad \frac{d^n}{dx^n} f(x) = \sum_{n=-\infty}^{\infty} (in)^n z_n e^{inx}$$

となることが期待され, 指数関数による分解は多項式による分解よりも微分作用素と相性がよいことが示唆される.

このように, 与えられた線形写像と相性のよいベクトルによる分解をいかに見つけるかが本節の主題となる. また, この問題は線形変換の表現行列をいかに簡単に表すか, というもう一つの基本的な問題と

⁶¹ $f(x+2\pi) = f(x)$ を満たす関数のこと

⁶² f が実数値関数の場合は, オイラーの公式により e^{inx} を三角関数に分解して表示するのが一般的である. 実際, 定義から z_{-n} は z_n の共役な複素数であり, $a_n = z_n + z_{-n}$ および $b_n = (z_n - z_{-n})i$ は共に実数になる. このとき

$$z_n e^{inx} + z_{-n} e^{-inx} = a_n \cos(nx) + b_n \sin(nx)$$

であることは容易に確かめられ, 三角級数展開

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

を得る.

も密接に関連している。まずはこの二つの問題の関係から論じよう。以下では、線形空間 U から U 自身への線形変換、すなわち $\text{End}(U)$ の元のみを考えることにする⁶³。

27.1 固有ベクトル

上で述べた微分作用素に対する指数関数のように $f(\mathbf{u}) = a\mathbf{u}$ を満たすベクトル \mathbf{u} についての議論は、そうでないベクトルについての議論よりも単純であろうことは明白である。このようなベクトルを固有ベクトルと言う。

定義 27.1.1. 線形変換 $f: U \rightarrow U$ において $f(\mathbf{v}) = \lambda\mathbf{v}$ を満たす $\lambda \in \mathbb{R}$ および零ベクトルでない $\mathbf{v} \in U$ が存在するとき、 λ を f における固有値 (eigenvalue) といい、 \mathbf{v} を固有値 λ に関する f の固有ベクトル (eigenvector) という。また、 n 次正方行列 A による線形変換 $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ の固有値および固有ベクトルをそれぞれ A における固有値および固有ベクトルと呼ぶ。

$\mathbf{u} \in U$ が f の固有値 λ_i に関する固有ベクトル \mathbf{v}_i に分解されるとき、すなわち $\mathbf{u} = \sum_{i=1}^n a_i \mathbf{v}_i$ とかけるとき $f(\mathbf{u})$ の値は次のように簡単に計算できる:

$$\begin{aligned} f(\mathbf{u}) &= f\left(\sum_{i=1}^n a_i \mathbf{v}_i\right) = \sum_{i=1}^n a_i f(\mathbf{v}_i) = \sum_{i=1}^n a_i \lambda_i \mathbf{v}_i, \\ f^2(\mathbf{u}) &= f\left(\sum_{i=1}^n a_i \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^n a_i \lambda_i f(\mathbf{v}_i) = \sum_{i=1}^n a_i \lambda_i^2 \mathbf{v}_i, \\ f^3(\mathbf{u}) &= \sum_{i=1}^n a_i \lambda_i^3 \mathbf{v}_i, \quad \dots, \quad f^n(\mathbf{u}) = \sum_{i=1}^n a_i \lambda_i^n \mathbf{v}_i. \end{aligned}$$

例 27.1.2. (1) n 次対角行列 $B = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ および \mathbb{R}^n の標準ベクトル $\mathbf{e}_1, \dots, \mathbf{e}_n$ について、 $B\mathbf{e}_i = \lambda_i \mathbf{e}_i$ である。ゆえに λ_i は B の固有値であり、 \mathbf{e}_i は B の固有値 λ_i に関する固有ベクトルである。

(2) U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ について、各 \mathbf{u}_i が線形変換 $f \in \text{End}(U)$ の固有値 λ_i に関する固有ベクトルであるとすると、このとき、 $f(\mathbf{u}_i) = \lambda_i \mathbf{u}_i$ であるから、基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列は対角行列

$$\text{列} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \text{である.}$$

例 27.1.3. ある自然数 n について $f^n = \mathbf{0}$ を満たす線形変換を冪零 (nilpotent) であるという。また、 $A^n = O$ を満たす行列を冪零であるという。冪零変換の固有値は 0 のみに限る。

Proof. まず 0 が f の固有値となることを示そう。 $\mathbf{x} \neq \mathbf{0}$ なるベクトルを一つとれば、 $f^n(\mathbf{x}) = \mathbf{0}$ ゆえ、ある非負整数 m について $f^m(\mathbf{x}) \neq \mathbf{0}$ かつ $f^{m+1}(\mathbf{x}) = \mathbf{0}$ となる。このとき、 $\mathbf{y} := f^m(\mathbf{x}) \neq \mathbf{0}$ とすれば、 $f(\mathbf{y}) = \mathbf{0} = 0 \cdot \mathbf{y}$ ゆえ \mathbf{y} は固有値 0 に関する固有ベクトルである。

一方、 $\mathbf{x} \neq \mathbf{0}$ を固有値 λ に関する f の固有ベクトルとすれば、

$$\mathbf{0} = f^n(\mathbf{x}) = f^{n-1} \circ f(\mathbf{x}) = f^{n-1}(\lambda\mathbf{x}) = \lambda f^{n-1}(\mathbf{x}) = \dots = \lambda^n \mathbf{x}.$$

いま $\mathbf{x} \neq \mathbf{0}$ としていたから $\lambda^n = 0$ である⁶⁴。ゆえに $\lambda = 0$ 。□

⁶³定義域の基底と終域の基底を別々に取ることが許されるならば、表現行列は簡約行列に取ることができ、したがって多くを論じる必要はなくなる。例 26.1.3(2) がその典型的な場合である。

⁶⁴仮に $\lambda^n \neq 0$ とすると、その逆数倍を $\lambda^n \mathbf{x} = \mathbf{0}$ の両辺にほどこせば $\mathbf{x} = \mathbf{0}$ となり矛盾してしまう。

例 27.1.4. (1) シフト作用素 $S: \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^{\mathbb{N}}$ ($S((x_n)_{n \in \mathbb{N}}) := (x_{n+1})_{n \in \mathbb{N}}$) において, 初項 a 等比 λ なる等比数列 $\mathbf{x} = (a\lambda^{n-1})_{n \in \mathbb{N}}$ は S の固有値 λ に関する固有ベクトルである.

(2) $y = e^{\lambda x}$ とおくと $y' = \lambda e^{\lambda x}$ である. したがって, 微分作用素 $D: C^\infty(I) \rightarrow C^\infty(I)$ ($D(y) := y'$) において関数 $e^{\lambda x}$ は D の固有値 λ に関する固有ベクトルである.

固有値が存在しない線形変換もある.

例 27.1.5. (1) 平面 \mathbb{R}^2 における 90° 回転を表す行列 $A = \begin{bmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ において, 各 $\mathbf{x} \in \mathbb{R}^2$ (ただし $\mathbf{x} \neq \mathbf{0}$) は $A\mathbf{x}$ と並行でないゆえ $A\mathbf{x} \neq \lambda\mathbf{x}$ である. ゆえに \mathbb{R} 上の線形空間としての線形写像 $T_A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ に固有ベクトルは存在しない.

(2) 列ベクトル成分やスカラー倍に複素数を許すとすれば, \mathbb{C} 上の線形空間としての線形写像 $T_A: \mathbb{C}^2 \rightarrow \mathbb{C}^2$ は固有ベクトルを持つ. 実際, $\mathbf{x} = \begin{bmatrix} i \\ 1 \end{bmatrix}$ は A の固有値 i に関する固有ベクトルである:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} i \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ i \end{bmatrix} = i \begin{bmatrix} i \\ 1 \end{bmatrix}.$$

27.2 固有ベクトルからなる基底と行列の対角化

各ベクトルを固有ベクトルの線形結合で表すと f^n が容易に計算できることを前節で見た. これと同様に, 対角行列 B において B^n を求めることは容易である:

$$\begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} = \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix}, \quad \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}^n = \begin{bmatrix} \lambda_1^n & & \\ & \ddots & \\ & & \lambda_n^n \end{bmatrix}.$$

実は, 線形変換の固有ベクトルからなる基底を求めることと, 与えられた正方行列 A に対して $P^{-1}AP$ が対角行列となるような可逆行列 P を探すことが, 次の議論によって同等であることが分かる.

- 正方行列 A に対して, 標準基底 $\mathbf{e}_1, \dots, \mathbf{e}_n$ に関する $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ の表現行列は A である. ここで, T_A の固有ベクトルからなる基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ があるとすれば, 基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ に関する T_A の表現行列 B は例 27.1.2(2) により対角行列である. また, $\mathbf{e}_1, \dots, \mathbf{e}_n$ による $\mathbf{v}_1, \dots, \mathbf{v}_n$ の変換行列は $P = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ であり, このとき系 26.3.2 より $B = P^{-1}AP$ である.

- U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列 A に対して, $P^{-1}AP = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$ となるならば, $[\mathbf{u}'_1, \dots, \mathbf{u}'_n] := [\mathbf{u}_1, \dots, \mathbf{u}_n]P$ で定められる新たな基底⁶⁵ $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ に関する f の表現行列は, 系 26.3.2 により対角行列 $P^{-1}AP$ に等しい. 表現行列の定義から $f(\mathbf{u}'_i) = \lambda_i \mathbf{u}'_i$ ($i = 1, \dots, n$) であり, ゆえに各 \mathbf{u}'_i は固有値 λ_i に関する固有ベクトルである.

$B = P^{-1}AP$ が対角行列であるとき, B^n が容易に計算できることから, A^n も次のように計算できる:

$$A^n = (PBP^{-1})^n = \underbrace{PBP^{-1} \cdot PBP^{-1} \cdot PBP^{-1} \cdots PBP^{-1}}_n = PB^nP^{-1}. \quad (27.2.1)$$

定義 27.2.1. 正方行列 A に対して, $P^{-1}AP$ が対角行列となるような可逆行列 P が存在するとき, A は対角化可能であるといい, $P^{-1}AP$ を A の対角化 (diagonalization) という.

行列の対角化可能性については次節で詳しく解説する. まずは線形変換の固有ベクトルをいかに求めるかを論じよう.

⁶⁵命題 25.4.1(2) により $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ は U の基底である.

27.3 特性多項式

線形空間 U における恒等写像 $\text{id}_U \in \text{End}(U)$ を I と書くのであった。線形変換 $f : U \rightarrow U$ および $\lambda \in \mathbb{R}$ について、 $\mathbf{x} \in U$ (ただし $\mathbf{x} \neq \mathbf{0}$) が λ に関する固有ベクトルとなる条件は次のように言い換えることができる:

$$\begin{aligned} f(\mathbf{x}) = \lambda \mathbf{x} &\iff \lambda \mathbf{x} - f(\mathbf{x}) = \mathbf{0} \iff \lambda I(\mathbf{x}) - f(\mathbf{x}) = \mathbf{0} \\ &\iff (\lambda I - f)(\mathbf{x}) = \mathbf{0} \iff \mathbf{x} \in \text{Ker}(\lambda I - f). \end{aligned}$$

したがって、固有値 λ に関する固有ベクトルが存在することは、 $\text{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$ であることと同値である:

命題 27.3.1. 線形空間 U および $f \in \text{End}(U)$, $\lambda \in \mathbb{R}$ について次は同値である。

- (1) λ は f の固有値である。
- (2) $\text{Ker}(\lambda I - f)$ は自明でない, すなわち $\text{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$.
- (3) $\lambda I - f$ は単射でない。

とくに $U = \mathbb{R}^n$ および $f = T_A$ について、 $(\lambda I - T_A)(\mathbf{x}) = (\lambda E - A)\mathbf{x}$ であり、 $\text{Ker}(\lambda I - T_A)$ が自明でないことは、方程式 $(\lambda E - A)\mathbf{x}$ が自明でない解をもつことに他ならない。これは、定理 6.2.2 により A が可逆でないこと、すなわち $|\lambda E - A| = 0$ と同値である。

また、一般の n 次元線形空間 U においても、 U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば、 $\lambda I - f$ の表現行列は $\lambda E - A$ である (1 次多項式 $\Psi(t) = \lambda - t$ について系 26.2.4 を適用せよ)。 $\lambda I - f$ が単射でないことは $T_{\lambda E - A}$ が単射でないことに同値であり (命題 26.4.1), これは $|\lambda E - A| = 0$ と同値である。

以上の考察から次が成り立つ:

系 27.3.2. $n \in \mathbb{N}$ および $\lambda \in \mathbb{R}$ とする。

- (1) n 次正方行列 A について、次が成り立つ: λ は A の固有値である $\iff |\lambda E - A| = 0$.
- (2) n 次元線形空間 U のある基底に関する $f \in \text{End}(U)$ の表現行列を A とすれば,

$$\lambda \text{ は } f \text{ の固有値である } \iff |\lambda E - A| = 0.$$

以上により、固有値を求めるには t に関する方程式 $|tE - A| = 0$ を解けばよいことが分かった。

定義 27.3.3. 正方行列 A に対して多項式 $\Phi_A(t) := |tE - A|$ を A の特性多項式あるいは固有多項式と呼ぶ。また、線形変換 $f \in \text{End}(U)$ の固有多項式 $\Phi_f(t)$ をある基底に関する f の表現行列の固有多項式と定める。次の命題により、 $\Phi_f(t)$ は f の表現行列の選び方に依らずに決まる多項式である。特性多項式の零点を表す方程式 $\Phi_A(t) = 0$ および $\Phi_f(t) = 0$ を特性方程式と呼ぶ。

命題 27.3.4. $f : U \rightarrow U$ を線形変換とし、基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による f の表現行列を A , 基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ による f の表現行列を B とすれば $\Phi_A = \Phi_B$.

Proof. 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とすれば系 26.3.2 より $B = P^{-1}AP$ が成り立つ。このとき

$$\begin{aligned} \Phi_B(t) &= |tE - B| = |tP^{-1}EP - P^{-1}AP| = |P^{-1}(tE)P - P^{-1}AP| = |P^{-1}((tE)P - AP)| \\ &= |P^{-1}(tE - A)P| = |P^{-1}| \cdot |tE - A| \cdot |P| = |tE - A| = \Phi_A(t). \end{aligned}$$

□

例 27.3.5. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ について

$$\begin{aligned} \Phi_A(t) &= \begin{vmatrix} t-a & -b \\ -c & t-d \end{vmatrix} = (t-a)(t-d) - (-b)(-c) \\ &= t^2 - (a+b)t + (ad-bc) = t^2 - (\operatorname{tr} A)t + \det A. \end{aligned}$$

例 27.3.6. 上三角行列 $A = \begin{bmatrix} a_1 & & * \\ & a_2 & * \\ O & \ddots & \\ & & a_n \end{bmatrix}$ について, 例 11.1.2 より

$$\Phi_A(t) = \begin{vmatrix} t-a_1 & & * \\ & t-a_2 & * \\ O & \ddots & \\ & & t-a_n \end{vmatrix} = (t-a_1)(t-a_2)\cdots(t-a_n).$$

すなわち, A の対角成分は A の固有値である.

練習 27.3.7. n 次正方形行列 A について, $\Phi_A(t)$ は n 次多項式であることを確認せよ.

解答例: $tE - A = [b_{ij}]$ の行列式 $|B| = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) b_{1\sigma(1)} b_{2\sigma(2)} \cdots b_{n\sigma(n)}$ の各項のうち, t の n 次多項式が現れるのは $\sigma = \operatorname{id}$ に関する項 $\operatorname{sgn}(\operatorname{id}) b_{11} b_{22} \cdots b_{nn} = (t-a_{11})(t-a_{22}) \cdots (t-a_{nn})$ のみであり, 他の置換に関する項はすべて $n-1$ 次以下の多項式となる. ゆえにこれらの和は n 次多項式である. \square

練習 27.3.8. A を m 次正方形行列, B を n 次正方形行列とする. $m+n$ 次正方形行列 $X = \begin{bmatrix} A & * \\ O & B \end{bmatrix}$ について $\Phi_X(t) = \Phi_A(t)\Phi_B(t)$ を示せ.

練習 27.3.9. 正方形行列 A について $\Phi_{tA} = \Phi_A$ を示せ.

27.4 固有空間

線形変換の固有ベクトルをすべて求める方法を考えよう. 前項で論じたように, 線形変換 $f: U \rightarrow U$ および $\lambda \in \mathbb{R}$ について, $\operatorname{Ker}(\lambda I - f) \neq \{\mathbf{0}\}$ であるとき, $\operatorname{Ker}(\lambda I - f)$ から $\mathbf{0}$ ベクトルを除いた集合は f の固有値 λ に関する固有ベクトル全体に一致する. そこで,

定義 27.4.1. $\lambda \in \mathbb{R}$ を $f \in \operatorname{End}(U)$ の固有値とする. このとき, U の部分空間 $\operatorname{Ker}(\lambda I - f)$ を f の固有値 λ における固有空間と呼び, 本論ではこれを $W(\lambda, f)$ で表す. また, n 次正方形行列 A について, $T_A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ の固有値 λ における固有空間を $W(\lambda, A)$ と書く. すなわち, $W(\lambda, A)$ は連立 1 次方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ の解空間 $W_{\lambda E - A}$ である.

例題 27.4.2. 行列 $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$ について, 次の問いに答えよ.

(1) A の特性多項式 $\Phi_A(t)$ を求めよ.

解答例:

$$\begin{aligned} \Phi_A(t) &= |tE - A| = \left| \begin{bmatrix} t & 0 \\ 0 & t \end{bmatrix} - \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix} \right| = \begin{vmatrix} t-8 & 10 \\ -5 & t+7 \end{vmatrix} \\ &= (t-8)(t+7) - (-5) \cdot 10 = t^2 - t - 56 + 50 = t^2 - t - 6 = (t-3)(t+2). \end{aligned}$$

(2) A の固有値をすべて求めよ.

解答例: 方程式 $\Phi_A(t) = 0$ を解いて $t = 3, -2$. 系 27.3.2 により, 3 と -2 が A の固有値である.

(3) A の各固有値に関する固有空間を求めよ (固有空間をその基底による外延的表記で表せ).

解答例: $\lambda = -2, 3$ の場合について連立 1 次方程式 $(\lambda E - A)\mathbf{x} = \mathbf{0}$ をそれぞれ解けばよい. いま $\Phi_A(\lambda) = 0$ に注意して簡約化を行うと

$$\lambda E - A = \begin{bmatrix} t-8 & 10 \\ -5 & t+7 \end{bmatrix} \rightarrow \begin{bmatrix} -5 & t+7 \\ t-8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{t+7}{5} \\ t-8 & 10 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -\frac{t+7}{5} \\ 0 & \frac{(t+7)(t-8)}{5} + 10 \end{bmatrix}$$

ここで, 最後の (2, 2)-成分は

$$\frac{(t+7)(t-8)}{5} + 10 = \frac{1}{5}((t+7)(t-8) + 50) = \frac{1}{5}\Phi_A(\lambda) = 0.$$

したがって $\lambda = -2, 3$ における $(\lambda E - A)\mathbf{x} = \mathbf{0}$ の解はそれぞれ次のようになる:

$$W(-2, A) = \left\{ r \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mid r \in \mathbb{R} \right\}, \quad W(3, A) = \left\{ r \begin{bmatrix} 2 \\ 1 \end{bmatrix} \mid r \in \mathbb{R} \right\}.$$

上の例題における各固有空間の基底による組 $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ は線形独立である. これは偶然ではなく, 一般論として命題 28.1.1 にて示す. $\mathbf{v}_1, \mathbf{v}_2$ は \mathbb{R}^2 の基底であり, この基底に関する T_A の表現行列を B とすれば例題 26.1.4(1) より $B = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ となる. また, 標準基底 $\mathbf{e}_1, \mathbf{e}_2$ による基底 $\mathbf{v}_1, \mathbf{v}_2$ に関する

変換行列は $P = [\mathbf{v}_1, \mathbf{v}_2] = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ である. ゆえに系 26.3.2 より $B = P^{-1}AP$ を得る.

例題 27.4.3. $A = \begin{bmatrix} 8 & -10 \\ 5 & -7 \end{bmatrix}$ について, A^n を求めよ.

解答例: $B = P^{-1}AP$ とすれば, 式 (27.2.1) より $A^n = PB^nP^{-1}$ である. $P = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$ とすれば上で論じたことから $B = \begin{bmatrix} -2 & 0 \\ 0 & 3 \end{bmatrix}$ である. また, P の逆行列を計算すれば $P^{-1} = \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix}$ である. よって,

$$A^n = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} (-2)^n & 0 \\ 0 & 3^n \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} (-2)^n & 2 \cdot 3^n \\ (-2)^n & 3^n \end{bmatrix} \begin{bmatrix} -1 & 2 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} -(-2)^n + 2 \cdot 3^n & 2(-2)^n - 2 \cdot 3^n \\ -(-2)^n + 3^n & 2(-2)^n - 3^n \end{bmatrix}.$$

練習 27.4.4. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ において $\Phi_A(t) = 0$ の解が λ_1, λ_2 であるとき, A の固有空間を求めよ.

解答例: $c \neq 0$ の場合と $c = 0$ の場合に分けて考える.

(1) $c \neq 0$ の場合:

$$\begin{aligned} \lambda_i E - A &= \begin{bmatrix} \lambda_i - a & -b \\ -c & \lambda_i - d \end{bmatrix} \rightarrow \begin{bmatrix} -c & \lambda_i - d \\ \lambda_i - a & -b \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 1 & -(\lambda_i - d)/c \\ \lambda_i - a & -b \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -(\lambda_i - d)/c \\ 1 & -b + (\lambda_i - a)(\lambda_i - b)/c \end{bmatrix} \end{aligned}$$

$\Phi_A(t) = (t-a)(t-b) - bc$ より, 上の行基本変形における最後の (2, 2)-成分は

$$-b + \frac{(\lambda_i - a)(\lambda_i - b)}{c} = \frac{-bc + (\lambda_i - a)(\lambda_i - b)}{c} = \frac{\Phi_A(\lambda_i)}{c} = \frac{0}{c} = 0.$$

ゆえに固有値 λ_i ($i = 1, 2$) に関する固有空間は, $W(\lambda_i, A) = \left\langle \begin{bmatrix} (\lambda_i - d)/c \\ 1 \end{bmatrix} \right\rangle$.

(2) $c = 0$ の場合: $\Phi(t) = (t - a)(t - d)$ より $\lambda_1 = a, \lambda_2 = d$ としてよい. b が 0 かどうかで更に場合分けをする.

- $b = 0$ ならば A は対角行列であり, a, d が A の固有値となる. $a = d$ の場合は T_A は各ベクトルを a 倍する写像に他ならず, 固有値 a に関する固有空間は全空間 $W(a, A) = \mathbb{R}^2$ である. $a \neq d$ の場合は, $W(\lambda_i, A) = \langle \mathbf{e}_i \rangle$ ($i = 1, 2$) となる.
- $b \neq 0$ のとき, 固有値 a に関する固有ベクトルは

$$aE - A = \begin{bmatrix} 0 & -b \\ 0 & a - d \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 \\ 0 & a - d \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

より $W(a, A) = \langle \mathbf{e}_1 \rangle$. 更に $a \neq d$ の場合は

$$dE - A = \begin{bmatrix} d - a & -b \\ 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -b/(d - a) \\ 0 & 0 \end{bmatrix}$$

よって, $W(d, A) = \left\langle \begin{bmatrix} b/(d - a) \\ 1 \end{bmatrix} \right\rangle$.

27.5 一般の線形写像の固有空間

一般の線形写像の固有値も, 系 27.3.2 により特性方程式を解くことで求められる. f の固有ベクトルと, その表現行列の固有ベクトルがいかに対応づけられるかを確認しておこう.

命題 27.5.1. U を n 次元線形空間とし $f: U \rightarrow U$ を線形変換とする. 更に, U の基底 $\mathbf{v}_1, \dots, \mathbf{v}_n$ による f 表現行列を A とすれば次が成り立つ.

- (1) $\mathbf{x} \in \mathbb{R}^n$ が T_A の固有値 λ に関する T_A の固有ベクトルであることと, $\mathbf{y} = [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} \in U$ が固有値 λ に関する f の固有ベクトルであることは同値である.
- (2) λ を f あるいは A の固有値とする. $\mathbf{y}_i = [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x}_i$ ($i = 1, \dots, k$) とすれば, $\mathbf{y}_1, \dots, \mathbf{y}_k$ が $W(\lambda, f)$ の基底であることと $\mathbf{x}_1, \dots, \mathbf{x}_k$ が $W(\lambda, A)$ の基底であることは同値である.
- (3) $\dim W(\lambda, f) = \dim W(\lambda, A)$.

Proof. (1): 各 $\mathbf{v}_i \in U$ を $\mathbf{e}_i \in \mathbb{R}^n$ に対応させる線形同型を $F: U \rightarrow \mathbb{R}^n$ とする. このとき $g = \lambda I - f$ の表現行列は $\lambda E - A$ であり (系 26.2.4), 練習 26.4.2(2) の証明によれば $\text{Ker } g = W(\lambda, f)$ と $\text{Ker}(\lambda E - A) = W(\lambda, A)$ は, F によって 1 対 1 に対応する. 25.4 項で述べたことによれば, F は $[\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x}$ を \mathbf{x} に対応させる同型であり, すなわち, $\mathbf{x} \in W(\lambda, A)$ と $[\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} \in W(\lambda, f)$ は同値である.

(2) および (3): 対応 $F: U \rightarrow \mathbb{R}^n$ ($[\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} \mapsto \mathbf{x}$) が線形同型であることから明らか. \square

補足: 上の (1) の証明が抽象的過ぎると感じる場合は, 次のような同値変形を行って確認してもよい:

$\mathbf{x} = {}^t[x_1, \dots, x_n]$ とおく. $\mathbf{y} = [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} = \sum_{i=1}^n x_i \mathbf{v}_i$ であること, および表現行列の定義より $[f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)] = [\mathbf{v}_1, \dots, \mathbf{v}_n]A$ に注意すると,

$$\mathbf{y} \text{ は } \lambda \text{ に関する } f \text{ の固有ベクトルである} \iff f\left(\sum_{i=1}^n x_i \mathbf{v}_i\right) = \lambda \mathbf{y}$$

$$\iff \sum_{i=1}^n x_i f(\mathbf{v}_i) = \lambda [\mathbf{v}_1, \dots, \mathbf{v}_n]\mathbf{x} \iff [f(\mathbf{v}_1), \dots, f(\mathbf{v}_n)]\mathbf{x} = [\mathbf{v}_1, \dots, \mathbf{v}_n](\lambda \mathbf{x})$$

$$\iff [\mathbf{v}_1, \dots, \mathbf{v}_n]A\mathbf{x} = [\mathbf{v}_1, \dots, \mathbf{v}_n](\lambda \mathbf{x})$$

$$\iff A\mathbf{x} = \lambda \mathbf{x} \iff \mathbf{x} \text{ は } \lambda \text{ に関する } A \text{ の固有ベクトルである.}$$

(↑ここで $\mathbf{v}_1, \dots, \mathbf{v}_n$ の線形独立性および練習 25.3.1 を用いた.)

例 27.5.2. 線形写像 $D: \mathbb{R}[x]_3 \rightarrow \mathbb{R}[x]_3$ を $D(p(x)) := \frac{d}{dx}p(x)$ と定める. 基底 $\mathbf{1}, x, x^2, x^3$ に関する D の

表現行列は, 例題 26.1.4(2) より $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ である. したがって $\Phi_D(t) = |tE - A| = t^4$ であり,

D の固有値は 0 のみである. また, A の固有空間は, 連立 1 次方程式 $(0E - A)\mathbf{x} = \mathbf{0}$ を解いて $W(0, A) =$

$\left\{ r \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \mid r \in \mathbb{R} \right\}$ である. これに対応する $\mathbb{R}[x]_3$ の部分空間が D の固有空間 $W(0, D)$ となる. $W(0, A)$

の各元 $\begin{bmatrix} r \\ 0 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^4$ に対応する $\mathbb{R}[x]_3$ の元は命題 27.5.1 より $[1, x, x^2, x^3] \begin{bmatrix} r \\ 0 \\ 0 \\ 0 \end{bmatrix} = r\mathbf{1} + 0x + 0x^2 + 0x^3 = r\mathbf{1}$

(定数関数) である. つまり, $W(0, D) = \langle \mathbf{1} \rangle$.

28 固有空間分解と行列の対角化

前節において、線形変換 $f: U \rightarrow U$ の固有値および固有ベクトルを求める方法を与えた。また、 U の各ベクトル \mathbf{u} を固有ベクトルに分解すると $f(\mathbf{u})$ の計算が容易になることを見た。しかしながら、そのためには任意のベクトルが固有ベクトルの線形結合で書けるかどうかが問題になる。残念ながら一般の線形変換についてこのような分解ができるわけではなく、本節では、この分解が可能であるための必要十分条件を提示する。

ところで、固有値を求めるには特性方程式 $\Phi_f(t) = 0$ を解くことが前提となっており、方程式が実数解を持たない場合にスカラー係数の範囲を複素数に広げたほうが都合がよい場合がある。しかしながら、スカラー係数を実数に限った話なのか複素数まで広げているのか議論が曖昧になっては困るので、本節においては体 \mathbb{K} を \mathbb{R} または \mathbb{C} のいずれかとし、線形空間とは体 \mathbb{K} 上の線形空間のこと、すなわちスカラー係数として \mathbb{K} 上の元を取るとする。また、行列とは \mathbb{K} の元を各成分に持つ行列のこととする。もちろん、線形写像あるいは正方行列の固有値は \mathbb{K} の元であり、多項式とは \mathbb{K} の元を係数に持つ多項式のことである。また、多項式 $\Phi(x)$ が因数分解されるとは、 $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ を用いて $\Phi(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ と書けることと定める。以上の設定のもとで、 $\mathbb{K} = \mathbb{R}$ の場合として前節までに述べてきたことは $\mathbb{K} = \mathbb{C}$ の場合においても成り立つ。

以降の内容において、実数または複素数のいずれかに限定するものについては、その都度、 \mathbb{K} がどちらであるか明確にした上で説明する。

28.1 固有ベクトルの線形独立性

次の命題は、異なる固有値に関する固有ベクトルがどのように配置されているかを理解するうえで基本となる事実である。そこで、本節では3回にわたってこれを証明しよう。まずは固有ベクトルの定義さえ知っていれば済む証明を与える。

命題 28.1.1. $f: U \rightarrow U$ を線形変換とし、その相異なる固有値を $\lambda_1, \dots, \lambda_r$ とする。また $\dim W(\lambda, f) = n_\lambda$ ($\lambda = \lambda_1, \dots, \lambda_r$) とする。このとき、各固有値 λ について固有空間 $W(\lambda, f)$ の基底 $\mathbf{u}_{\lambda,1}, \dots, \mathbf{u}_{\lambda,n_\lambda}$ を一つ選べば、これらを集めたベクトルの組 $\mathcal{B} = \{\mathbf{u}_{\lambda,i} \mid \lambda = \lambda_1, \dots, \lambda_r, i = 1, \dots, n_\lambda\}$ は線形独立である。

Proof. \mathcal{B} が線形従属であると仮定し、矛盾を示そう。 $\mathbf{b}_1, \dots, \mathbf{b}_\ell \in \mathcal{B}$ を $\langle \mathcal{B} \rangle$ の基底とする。 \mathcal{B} 自身は線形従属ゆえ、 $\mathbf{b}_1, \dots, \mathbf{b}_\ell$ のいずれでもない $\mathbf{b}_{\ell+1} \in \mathcal{B}$ が取れる。ここで、各 \mathbf{b}_i ($i = 1, \dots, \ell$) は固有値 γ_i に関する固有ベクトルであるとし (γ_i には重複がある可能性もある)、また \mathbf{b} を固有値 γ に関する固有ベクトルとする。 $\mathbf{b} = \sum_{i=1}^{\ell} c_i \mathbf{b}_i$ とし、この両辺に f ほどこすと、

$$\begin{aligned} f(\mathbf{b}) &= c_1 f(\mathbf{b}_1) + \cdots + c_\ell f(\mathbf{b}_\ell) \\ \gamma \mathbf{b} &= c_1 \gamma_1 \mathbf{b}_1 + \cdots + c_\ell \gamma_\ell \mathbf{b}_\ell \\ \gamma(c_1 \mathbf{b}_1 + \cdots + c_\ell \mathbf{b}_\ell) &= c_1 \gamma_1 \mathbf{b}_1 + \cdots + c_\ell \gamma_\ell \mathbf{b}_\ell \\ \sum_{i=1}^{\ell} c_i(\gamma - \gamma_i) \mathbf{b}_i &= \mathbf{0}. \end{aligned}$$

$\mathbf{b}_1, \dots, \mathbf{b}_\ell$ の線形独立性より、 $c_i(\gamma - \gamma_i) = 0$ ($i = 1, \dots, \ell$) であり、とくに γ_i と γ が異なる固有値であれば $\gamma - \gamma_i \neq 0$ ゆえ $c_i = 0$ を得る。したがって、線形結合 $\mathbf{b} = \sum_{i=1}^{\ell} c_i \mathbf{b}_i$ に現れる係数 c_i のうち、0でないものは $\gamma_i = \gamma$ の場合、すなわち $\mathbf{b}_i \in W(\gamma, f)$ の場合に限る。このような \mathbf{b}_i および \mathbf{b} は、 $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ のいずれかのベクトルである。命題 17.2.5 より $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ は線形従属となり、これは $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ を $W(\gamma, f)$ の基底として取ったことに矛盾する。□

系 28.1.2. A を n 次正方行列とし、その相異なる固有値を $\lambda_1, \dots, \lambda_r$ とすれば、 $\sum_{k=1}^r \dim W(\lambda_k, T_A) \leq n$.

備考 28.1.3. 実は、 f の特性多項式が $\Phi_f(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$ と因数分解される場合、 $\dim W(\lambda_k, f) \leq n_k$ が成立する。この事実の詳細は次節以降で述べる (定理 29.3.1)。

背理法による証明には、命題が成立する様子がいま一つ掴めないところがある。そこで、命題 28.1.1 を背理法を用いずに示してみよう。

命題 28.1.1 の別証明

$$\sum_{\lambda=\lambda_1, \dots, \lambda_r} \left(\sum_{i=1, \dots, n_\lambda} c_{\lambda, i} \mathbf{u}_{\lambda, i} \right) = \mathbf{0}$$

と仮定し、 $c_{\lambda, i} = 0$ を示す。各 $\lambda = \lambda_1, \dots, \lambda_r$ について $\mathbf{v}_\lambda = \sum_{i=1, \dots, n_\lambda} c_{\lambda, i} \mathbf{u}_{\lambda, i}$ とおけば、 $\sum_{\lambda=\lambda_1, \dots, \lambda_r} \mathbf{v}_\lambda = \mathbf{0}$ である。このとき $\mathbf{v}_{\lambda_k} \in W(\lambda_k, f)$ ゆえ $f(\mathbf{v}_{\lambda_k}) = \lambda_k \mathbf{v}_{\lambda_k}$ であることに注意すると、 $\sum_{k=1}^r \mathbf{v}_{\lambda_k} = \mathbf{0}$ の両辺に f をほどこすことで $\sum_{k=1}^r \lambda_k \mathbf{v}_{\lambda_k} = \mathbf{0}$ を得る。この式に更に f をほどこせば、 $\sum_{k=1}^r \lambda_k^2 \mathbf{v}_{\lambda_k} = \mathbf{0}$ であり、この操作を順次繰り返すことで次の等式たちを得る：

$$\begin{aligned} \mathbf{v}_{\lambda_1} + \mathbf{v}_{\lambda_2} + \dots + \mathbf{v}_{\lambda_r} &= \mathbf{0}, \\ \lambda_1 \mathbf{v}_{\lambda_1} + \lambda_2 \mathbf{v}_{\lambda_2} + \dots + \lambda_r \mathbf{v}_{\lambda_r} &= \mathbf{0}, \\ \lambda_1^2 \mathbf{v}_{\lambda_1} + \lambda_2^2 \mathbf{v}_{\lambda_2} + \dots + \lambda_r^2 \mathbf{v}_{\lambda_r} &= \mathbf{0}, \\ &\vdots \\ \lambda_1^{r-1} \mathbf{v}_{\lambda_1} + \lambda_2^{r-1} \mathbf{v}_{\lambda_2} + \dots + \lambda_r^{r-1} \mathbf{v}_{\lambda_r} &= \mathbf{0}. \end{aligned}$$

$$\text{つまり、} P = \begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{r-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{r-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \lambda_r & \lambda_r^2 & \dots & \lambda_r^{r-1} \end{bmatrix} \quad \text{とおけば} \quad [\mathbf{v}_{\lambda_1}, \dots, \mathbf{v}_{\lambda_r}] P = [\mathbf{0}, \dots, \mathbf{0}].$$

各 λ_k は相異なる固有値としていたゆえ、 $1 \leq p < q \leq r$ ならば $\lambda_q - \lambda_p \neq 0$ である。したがってヴァンデルモンドの行列式 (定理 11.2.2) より $|P| \neq 0$ 、つまり P は可逆であり、 $[\mathbf{v}_{\lambda_1}, \dots, \mathbf{v}_{\lambda_r}] P = [\mathbf{0}, \dots, \mathbf{0}]$ の両辺に右から P^{-1} をかけることで $[\mathbf{v}_{\lambda_1}, \dots, \mathbf{v}_{\lambda_r}] = [\mathbf{0}, \dots, \mathbf{0}]$ を得る。つまり各 $\lambda = \lambda_1, \dots, \lambda_r$ について $\sum_{i=1, \dots, n_\lambda} c_{\lambda, i} \mathbf{u}_{\lambda, i} = \mathbf{0}$ である。 $\mathbf{u}_{\lambda, 1}, \dots, \mathbf{u}_{\lambda, n_\lambda}$ の線形独立性から $c_{\lambda, i} = 0$ を得る。□

上の証明の前半は強力な公式 (ヴァンデルモンドの公式) に帰着させている。この公式を用いない方法は本節の後半で与える (補題 28.4.1)。

28.2 対角化可能条件

次の定理は正方行列の対角化可能性の必要十分条件を与えるとともに、その証明において対角化を求める方法も提示している。

定理 28.2.1. A を n 次正方行列とすれば次は同値。

- (1) A の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とすれば、 $\sum_{k=1}^r \dim W(\lambda_k, A) = n$.
(注意: この条件は系 28.1.2 より $\sum_{k=1}^r \dim W(\lambda_k, A) \geq n$ と同値である.)
- (2) A の固有ベクトルからなる \mathbb{R}^n の基底が存在する.
- (3) A は対角化できる. すなわち $P^{-1}AP$ が対角行列となるような可逆行列 P が存在する.

Proof. (1)⇒(2): 各 $\lambda = \lambda_1, \dots, \lambda_r$ について $n_\lambda = \dim W(\lambda, A)$ とおく。 $W(\lambda, A)$ の基底を一つ取り、これらを集めて $B = \{ \mathbf{u}_{\lambda, i} \mid \lambda = \lambda_1, \dots, \lambda_r, i = 1, \dots, n_\lambda \}$ とすれば、命題 28.1.1 より B は線形独立である。また、仮定より B の元の個数は n に等しいゆえ B は \mathbb{R}^n の基底となる。

(2)⇒(3): $\mathbf{u}_1, \dots, \mathbf{u}_n$ を A の固有ベクトルからなる \mathbb{R}^n の基底とし、 \mathbf{u}_i は固有値 γ_i に関する固有ベクトルとする (つまり $A\mathbf{u}_i = \gamma_i \mathbf{u}_i$)。 $P := [\mathbf{u}_1, \dots, \mathbf{u}_n]$ と定めれば、定理 17.3.6 より P は可逆である。こ

のとき $P^{-1}AP$ が対角行列となることは次のように確かめられる: $Pe_i = \mathbf{u}_i$ および, この両辺に左から P^{-1} をかけて得られる $e_i = P^{-1}\mathbf{u}_i$ に注意すれば

$$“P^{-1}AP \text{ の } i \text{ 列目}” = P^{-1}APe_i = P^{-1}A\mathbf{u}_i = P^{-1}\gamma_i\mathbf{u}_i = \gamma_iP^{-1}\mathbf{u}_i = \gamma_ie_i.$$

ゆえに

$$P^{-1}AP = [\gamma_1\mathbf{e}_1, \dots, \gamma_n\mathbf{e}_n] = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix}.$$

(3) \Rightarrow (1):

$$P^{-1}AP = \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix} = B \quad (\text{したがって } A = PBP^{-1})$$

であると仮定し, $P = [\mathbf{u}_1, \dots, \mathbf{u}_n]$ とおけば, P は可逆ゆえ $\mathbf{u}_1, \dots, \mathbf{u}_n$ は \mathbb{R}^n の基底である. \mathbf{u}_i が固有値 γ_i の固有ベクトルであることは, 次のように確認できる ($Pe_i = \mathbf{u}_i$ に注意する).

$$\begin{aligned} A\mathbf{u}_i &= PBP^{-1}\mathbf{u}_i = PBP^{-1}(Pe_i) = PBe_i \\ &= P \begin{bmatrix} \gamma_1 & & \\ & \ddots & \\ & & \gamma_n \end{bmatrix} e_i = P(\gamma_i e_i) = \gamma_i P e_i = \gamma_i \mathbf{u}_i. \end{aligned}$$

したがって, A の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とすれば

$$\sum_{k=1}^r \dim W(\lambda_k, A) \geq \sum_{k=1}^r \left(“\mathbf{u}_i \in W(\lambda_k, A) \text{ を満たす } i \text{ の個数}” \right) = n.$$

上の式と系 28.1.2 を合わせて $\sum_{k=1}^r \dim W(\lambda_k, A) = n$ を得る. □

上の定理は, \mathbb{R}^n の各ベクトルが A の各固有空間の元によって分解できることと A が対角化可能であることが必要十分であることを述べている. 一般の行列については, 任意のベクトルが固有ベクトルに分解できるわけではない. しかしながら, 固有空間を少し広げた広義固有空間の元によって \mathbb{R}^n の任意のベクトルが分解できる場合がある. これについては次節以降に論じる.

系 28.2.2. n 次正方行列 A が n 個の相異なる固有値 $\lambda_1, \dots, \lambda_n$ を持つならば, A は対角化可能である.

Proof. 各固有値 λ_k の固有ベクトルは少なくとも一つは存在する. したがって $\dim W(\lambda, T_A) \geq 1$ であり, $\sum_{k=1}^n \dim W(\lambda_k, T_A) \geq n$. ゆえに前定理より A は対角化できる. □

備考 28.2.3. $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ において $\Phi_A(t) = 0$ の解が λ_1, λ_2 であるとしよう. $\lambda_1 \neq \lambda_2$ の場合: 系 28.2.2 より A は対角化可能である. 一方, $\lambda_1 = \lambda_2$ の場合は, A が対角化可能であるためには $\dim W(\lambda_1, A) = 2$, つまり $W(\lambda_1, A) = \mathbb{R}^2$ でなければならない. そのためには, A が元々対角行列でなければならないことが練習 27.4.4 から分かる.

ここで, 行列の対角化の方法についてまとめておこう.

まとめ (対角化の仕方).

n 次正方行列 A の対角化可能性の判定, および A の対角化は次の手順によって得られる:

- (1) 特性多項式 $\Phi_A(t) = |tE - A|$ を求め, 特性方程式 $\Phi_A(t) = 0$ を解く.
- (2) $\Phi_A(t) = 0$ の相異なる解 $\lambda_1, \dots, \lambda_k$ について, 各固有空間 $W(\lambda_k, A) = \text{Ker}(\lambda_k I - T_A)$ の次元を求める. これは連立 1 次方程式 $(\lambda_k E - A)\mathbf{x} = \mathbf{0}$ の解空間のことであり, この基底は命題 18.3.4 の証明と同様にして求められる.
- (3) 定理 28.2.1 により, $\sum_{k=1}^r \dim W(\lambda_k, A) = n$ ならば A は対角化可能であり, そうでなければ A は対角化可能でない.
- (4) A が対角化可能であるとき, (2) で求めた各固有空間 $W(\lambda_k, A)$ の基底を並べれば, これは n 個のベクトルの組 $\mathbf{p}_1, \dots, \mathbf{p}_n$ であり, 定理 28.2.1 における (2) \Rightarrow (3) の証明によれば $P = [\mathbf{p}_1, \dots, \mathbf{p}_n]$ について $B = P^{-1}AP$ は対角行列となる.
- (5) このとき対角行列 B の (i, i) -成分は, \mathbf{p}_i に対応する固有値に等しい.

以上の手順から, A を対角化するための可逆行列 P の取り方には, 各固有ベクトルの基底の並べ方ぶんの自由度があることが分かる.

ところで, 方程式 $\Phi_A(t) = 0$ を解くには, n 次方程式の解の公式を用いるか, $\Phi_A(t)$ を上手く因数分解しなければならない. 一般に, 4 次以下の n 次方程式については根号と四則演算のみによる解の公式があり, 5 次以上についてはそのような公式は存在しないことが知られている. 3 次および 4 次方程式の解の公式の導出や, 5 次以上の方程式の解をいかに求めるかという問題は線形代数学の枠を超えるため, 本論では上の事実に言及するに留め, これ以上は論じないことにする.

例題 28.2.4. 次の行列が対角化可能かどうか答え, 対角化可能ならば対角化せよ.

$$(1) A = \begin{bmatrix} 5 & 6 & 0 \\ -1 & 0 & 0 \\ 1 & 2 & 2 \end{bmatrix}.$$

まず固有値を求め, 各固有値に関する固有空間の次元を求める.

$$\begin{aligned} \Phi_A(t) = |tE - A| &= \begin{vmatrix} t-5 & -6 & 0 \\ 1 & t & 0 \\ -1 & -2 & t-2 \end{vmatrix} = \begin{vmatrix} t-5 & -6 \\ 1 & t \end{vmatrix} (t-2) \quad (\text{定理 12.5.1 を用いた}) \\ &= ((t-5)t - (-6) \cdot 1)(t-2) = (t^2 - 5t + 6)(t-2) = (t-2)^2(t-3). \end{aligned}$$

特性方程式 $\Phi_A(\lambda) = 0$ を解いて, $\lambda = 3, 2$ が A の固有値である.

- $\lambda = 3$ の固有空間 $W(3, A)$: 方程式 $(2E - A)\mathbf{x} = \mathbf{0}$ を解くために $3E - A$ を簡約化すると

$$3E - A = \begin{bmatrix} -2 & -6 & 0 \\ 1 & 3 & 0 \\ -1 & -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 0 \\ -2 & -6 & 0 \\ -1 & -2 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 3 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -3 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $(3E - A)\mathbf{x} = \mathbf{0}$ の解空間は, $W(3, A) = \left\langle \begin{bmatrix} 3 \\ -1 \\ 1 \end{bmatrix} \right\rangle$ であり, $\dim W(3, A) = 1$.

- $\lambda = 2$ の固有空間 $W(2, A)$:

$$2E - A = \begin{bmatrix} -3 & -6 & 0 \\ 1 & 2 & 0 \\ -1 & -2 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $(2E - A)\mathbf{x} = \mathbf{0}$ の解空間は、 $W(2, A) = \left\langle \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle$ であり、 $\dim W(2, A) = 2$.

以上より、 $\dim W(3, A) + \dim W(2, A) = 3$ である。したがって A は対角化可能であり、各固有空間の基底を並べた正方行列を

$$P \begin{bmatrix} 3 & -2 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \quad \text{とすれば} \quad P^{-1}AP = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

注意: これまでの議論により、 P^{-1} を求めなくても A の対角化が分かる。途中の計算にミスがないか心配な場合は、対角化 $B = P^{-1}AP$ の両辺に左から P をかけた等式 $PB = AP$ が成り立つかどうかを確認するとよい。

$$(2) \quad A = \begin{bmatrix} 1 & 3 & 2 \\ 0 & -1 & 0 \\ 1 & 2 & 0 \end{bmatrix}.$$

$$\begin{aligned} \Phi_A(t) = |tE - A| &= \begin{vmatrix} t-1 & -3 & -2 \\ 0 & t+1 & 0 \\ -1 & -2 & t \end{vmatrix} = (-1) \begin{vmatrix} -1 & t-1 & -2 \\ t+1 & 0 & 0 \\ -2 & -1 & t \end{vmatrix} \quad (1 \text{ 列と } 2 \text{ 列の入れ替え}) \\ &= (-1)^2 \begin{vmatrix} t+1 & 0 & 0 \\ -1 & t-1 & -2 \\ -2 & -1 & t \end{vmatrix} = (t+1) \begin{vmatrix} t-1 & -2 \\ -1 & t \end{vmatrix} \\ &= (t+1) \left((t-1)t - (-2)(-1) \right) = (t+1)(t^2 - t - 2) = (t+1)^2(t-2). \end{aligned}$$

● $\lambda = 2$ の固有空間 $W(2, A)$:

$$2E - A = \begin{bmatrix} 1 & -3 & -2 \\ 0 & 3 & 0 \\ -1 & -2 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & -2 \\ 0 & 3 & 0 \\ 0 & -5 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & -3 & -2 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに $(2E - A)\mathbf{x} = \mathbf{0}$ の解空間は、 $W(2, A) = \left\langle \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} \right\rangle$ であり、 $\dim W(2, A) = 1$.

● $\lambda = -1$ の固有空間 $W(-1, A)$:

$$-E - A = \begin{bmatrix} -2 & -3 & -2 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 2 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 1 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

ゆえに方程式 $(-E - A)\mathbf{x} = \mathbf{0}$ の解空間は、 $W(-1, A) = \left\langle \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\rangle$ であり、 $\dim W(-1, A) = 1$.

以上より、 $\dim W(2, A) + \dim W(-1, A) \neq 3$ である。したがって A は対角化可能でない。

補足: 3次正方行列 A の特性多項式が $\Phi_A(t) = (t - \lambda)(t - \gamma)^2$ (ただし $\lambda \neq \gamma$) である場合、備考 28.1.3において予告したことを認めると、 $1 \leq \dim W(\lambda, A) \leq 1$ および $1 \leq \dim W(\gamma, A) \leq 2$ であることが分かる。つまり、 $\dim W(\lambda, A) = 1$ となることは計算せずとも分かり、 A が対角化可能かどうかは、 $\dim W(\gamma, A)$ のみによって決まる。したがって、先に $\dim W(\gamma, A)$ の次元から確認するほうが筋がよい。

例 27.1.5 に現れた行列について改めて考えてみよう.

練習 28.2.5. $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ が対角化可能かどうか答えよ.

解答例: $\Phi_A(t) = t^2 + 1$ より方程式 $\Phi_A(t) = 0$ は実数解を持たない. したがって, A の固有ベクトルは \mathbb{R}^2 上には存在せず, 実数を成分とする行列 P を用いて $P^{-1}AP$ を対角行列にすることはできない. 一方, 成分に複素数を認める場合は $\Phi_A(t) = 0$ の解 $t = \pm i$ が A の固有値であり, その固有空間はそれぞれ $W(i, A) = \left\langle \begin{bmatrix} i \\ 1 \end{bmatrix} \right\rangle$, $W(-i, A) = \left\langle \begin{bmatrix} -i \\ 1 \end{bmatrix} \right\rangle$ である⁶⁶. ゆえに $P = \begin{bmatrix} i & -i \\ 1 & 1 \end{bmatrix}$ とすれば $P^{-1}AP = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$.

28.3 一般の線形変換の場合

これまでの内容を一般の線形空間の表現行列に関する主張に書き換えると次のようになる. この定理の証明の一部は 27.2 項で論じたことと重複するが, ここで改めて述べておく.

定理 28.3.1. n 次元線形空間 U 上の線形変換 $f: U \rightarrow U$ について次は同値である.

- (i) f の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とすれば $\sum_{k=1}^r \dim W(\lambda_k, f) = n$.
- (ii) f の固有ベクトルからなる U の基底が存在する.
- (iii) f の任意の表現行列は対角化可能である.
- (iv) f は対角行列となるような表現行列を持つ.

Proof. (i) \Leftrightarrow (ii): U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列を A とすれば, 上の条件 (i), (ii) はそれぞれ定理 28.2.1 における条件 (1), (2) と同値になる. 実際, (i) \Leftrightarrow (1) および (ii) \Leftrightarrow (2) は命題 27.5.1 による. 条件 (1), (2) の同値性より, 条件 (i), (ii) も同値である.

(i) \Rightarrow (iii): U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を一つ取り, この基底に関する f の表現行列を A とする. いま条件 (i) を仮定しているゆえ, 命題 27.5.1(3) より A は定理 28.2.1 における条件 (1) を満たし, したがって対角化できる.

(iii) \Rightarrow (i): U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を一つ取り, この基底に関する f の表現行列を A とする. 仮定より (iii) は対角化可能であり, したがって A は定理 28.2.1 における条件 (1) を満たす. 命題 27.5.1(3) より, これは f が (i) を満たすことに他ならない.

(iii) \Rightarrow (iv): 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列 A が対角化できるとする. つまり, $B = P^{-1}AP$ が対角行列となるような可逆行列 P が存在する. このとき, n 個のベクトル $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ を $[\mathbf{u}'_1, \dots, \mathbf{u}'_n] := [\mathbf{u}_1, \dots, \mathbf{u}_n]P$ と定めれば, これらは U の基底であり (命題 25.4.1(2)), 基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ による f の表現行列は, 系 26.3.2 より $B = P^{-1}AP$ である. ゆえに f は対角行列になる表現行列を持つ.

(iv) \Rightarrow (iii): U の基底 $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ に関する f の表現行列が対角行列 B であるとする. U の任意の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ に関する f の表現行列 A が対角化できることを示そう. 基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による $\mathbf{u}'_1, \dots, \mathbf{u}'_n$ の変換行列を P とすれば, 系 26.3.2 より $B = P^{-1}AP$ であり, A は対角化可能である. \square

練習 28.3.2. 例 27.5.2 で定めた線形変換 $D: \mathbb{R}[x]_3 \rightarrow \mathbb{R}[x]_3$ ($D(p(x)) := \frac{d}{dx}p(x)$) は, 対角行列となるような表現行列を持つかどうか答えよ.

解答例: D の固有値は 0 のみであり, その固有空間は 1 次元である. $\dim W(0, D) \neq \dim \mathbb{R}[x]_3 = 4$ より, 対角行列となるような D の表現行列は存在しない. \square

⁶⁶これらの計算は練習 27.4.4(1) から直ちに分かる.

基底 $\mathbf{1}, x, x^2, x^3$ に関する上の D の表現行列は、例題 26.1.4(2) より $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ であり、 $A^4 = O$

である (これは 3 次多項式を 4 回微分すると必ず $\mathbf{0}$ になることに相当している). このような行列は、次のような一般論によって対角化できないことが分かる.

練習 28.3.3 (発展). 対角化可能な冪零行列は O に限ることを示せ.

解答例: n 次冪零行列 A が対角化できるならば、 A の固有ベクトル $\mathbf{p}_1, \dots, \mathbf{p}_n$ からなる \mathbb{R}^n の基底がある. 例 27.1.3 より A の固有値は 0 のみであるゆえ、 A の対角化は $P^{-1}AP = O$ である. この両辺に左から P を、右から X^{-1} をかけて $A = O$. \square

28.4 対角化可能な行列に関するケーリー・ハミルトンの定理

命題 28.1.1 の別証明の前半において、次の補題と同等の主張を証明していた. これを別の視点からもう一度証明してみよう. ここでの議論は、後に学ぶ広義固有空間分解およびケーリー・ハミルトンの定理への理解の助けとなるだろう.

補題 28.4.1. f の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とする. $\mathbf{u}_k \in W(\lambda_k, f)$ ($k = 1, \dots, r$) かつ $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ ならば、 $\mathbf{u}_1 = \dots = \mathbf{u}_r = \mathbf{0}$.

Proof. $(f - \lambda_k I)(\mathbf{u}_k) = f(\mathbf{u}_k) - \lambda_k \mathbf{u}_k = \lambda_k \mathbf{u}_k - \lambda_k \mathbf{u}_k = \mathbf{0}$ に注意しよう.. いまから、 $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に線形変換 $(f - \lambda_k I)$ ($k = 2, \dots, r$) を順次ほどこすことで $\mathbf{u}_2, \dots, \mathbf{u}_r$ の項が消えて \mathbf{u}_1 の項のみが残ることを見よう. $r - 1$ 次多項式 $\Psi_1(t) = (t - \lambda_2) \cdots (t - \lambda_r)$ は、各 $k = 2, \dots, r$ について別の多項式 $\Theta_k(t)$ を用いて $\Psi_1(t) = \Theta_k(t)(t - \lambda_k)$ と分解できる. このとき、

$$\begin{aligned} \Psi_1(f)(\mathbf{u}_k) &= \left(\Theta_k(f)(f - \lambda_k I) \right) (\mathbf{u}_k) = \Theta_k(f) \circ (f - \lambda_k I)(\mathbf{u}_k) \\ &= \Theta_k(t) \left((f - \lambda_k I)(\mathbf{u}_k) \right) = \Theta_k(t)(\mathbf{0}) = \mathbf{0} \end{aligned}$$

である. ゆえに $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に線形変換 $\Psi_1(f)$ をほどこせば

$$\begin{aligned} \Psi_1(f)(\mathbf{u}_1 + \dots + \mathbf{u}_r) &= \mathbf{0} \\ \sum_{k=1}^r \Psi_1(f)(\mathbf{u}_k) &= \mathbf{0} \\ \Psi_1(f)(\mathbf{u}_1) + \sum_{k=2}^r \Psi_1(f)(\mathbf{u}_k) &= \mathbf{0} \\ \Psi_1(f)(\mathbf{u}_1) &= \mathbf{0}. \end{aligned}$$

また、各 $k = 2, \dots, r$ について $(f - \lambda_k I)(\mathbf{u}_1) = \lambda_1 \mathbf{u}_1 - \lambda_k \mathbf{u}_1 = (\lambda_1 - \lambda_k) \mathbf{u}_1$ である. これを順次繰り返すことで

$$\begin{aligned} \mathbf{0} &= \Psi_1(f)(\mathbf{u}_1) = \left((f - \lambda_2 I) \cdots (f - \lambda_r I) \right) (\mathbf{u}_1) \\ &= (f - \lambda_2 I) \circ \cdots \circ (f - \lambda_{r-1} I) \circ (f - \lambda_r I) (\mathbf{u}_1) \\ &= (f - \lambda_2 I) \circ \cdots \circ (f - \lambda_{r-1} I) \left((\lambda_1 - \lambda_r) \mathbf{u}_1 \right) \\ &= (f - \lambda_2 I) \circ \cdots \circ (f - \lambda_{r-2} I) \left((\lambda_1 - \lambda_{r-1})(\lambda_1 - \lambda_r) \mathbf{u}_1 \right) \\ &= \cdots = (\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_r) \mathbf{u}_1. \end{aligned}$$

各 λ_k は相異なる数であったゆえ $(\lambda_1 - \lambda_2) \cdots (\lambda_1 - \lambda_r) \neq 0$ であり、したがって $\mathbf{u}_1 = \mathbf{0}$ である.

次に $\mathbf{u}_1 + \dots + \mathbf{u}_r = \mathbf{0}$ の両辺に $(f - \lambda_k I)$ ($k = 1, 3, \dots, r$) を順次ほどこせば、同様の理由により $\mathbf{u}_2 = \mathbf{0}$ を得る. これらと類似の議論を $\mathbf{u}_3, \dots, \mathbf{u}_r$ についても行い、 $\mathbf{u}_1 = \dots = \mathbf{u}_r = \mathbf{0}$ を得る. \square

上の議論から次の命題の証明が容易に推察できる。証明の鍵は、各線形変換 $(f - \lambda_k I)$ が互いに可換になることにある。

命題 28.4.2. n 次正方行列 A が対角化可能であるとき、 A の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とすれば、多項式 $\phi(t) = (t - \lambda_1) \cdots (t - \lambda_r)$ について $\phi(A) = O$ 。

Proof. A の対角化可能性から、 \mathbb{K}^n の各元は固有ベクトルの線形結合で表せる。したがって、各固有ベクトル $\mathbf{v}_k \in W(\lambda_k, A)$ について $\phi(A)\mathbf{v}_k = \mathbf{0}$ を示せば十分である。 $\phi(t)$ は、 $r - 1$ 次多項式 $\Theta(t)$ を用いて $\phi(t) = \Theta(t)(t - \lambda_k)$ と書けること、および $(A - \lambda_k E)\mathbf{v}_k = A\mathbf{v}_k - \lambda_k \mathbf{v}_k = \mathbf{0}$ から

$$\phi(A)\mathbf{v}_k = \Theta(A)(A - \lambda_k E)\mathbf{v}_k = \Theta(A)\mathbf{0} = \mathbf{0}.$$

以上より $\phi(A) = O$. □

次項で証明する因数定理を用いれば、次の主張を得る。

命題 28.4.3 (ケーリー・ハミルトンの定理). n 次正方行列 A が対角化可能であるとき、 $\Phi_A(A) = O$ 。

Proof. $\phi(t) = (t - \lambda_1) \cdots (t - \lambda_r)$ とすれば、系 28.5.2 より $\Phi_A(t) = \Psi(t)\phi(t)$ と書ける。ゆえに $\Phi_A(A) = \Psi(A)\phi(A) = \Psi(A)O = O$. □

一般の正方行列におけるケーリー・ハミルトンの定理については 29.4 項をみよ。

28.5 多項式と方程式の解に関する基本的性質 (付録)

x を変数とする \mathbb{K} 係数の n 次方程式

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0 \quad \text{ただし } a_0, a_1, \dots, a_{n-1} \in \mathbb{K}$$

の解に関するいくつかの性質について簡単に解説しておく。

定理 28.5.1 (因数定理). \mathbb{K} 係数 n 次多項式 $\Phi(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ について、 $\lambda \in \mathbb{K}$ を $\Phi(x) = 0$ の解とすれば、 $\Phi(x)$ は $x - \lambda$ で割り切れる。すなわち、ある \mathbb{K} 係数 $n - 1$ 次多項式 $\Psi(x)$ を用いて、 $\Phi(x) = (x - \lambda)\Psi(x)$ と書ける。

Proof. $\Phi(x)$ を $x - \lambda$ で割った商を $\Psi(x)$ 、余りを $R(x)$ とする。いま $x - \lambda$ が 1 次式であることから、商は $n - 1$ 次式、余りは 0 次式、すなわち定数 $r \in \mathbb{K}$ である。つまり、 $\Phi(x) = (x - \lambda)\Psi(x) + r$ と書ける。この両辺に $x = \lambda$ を代入し $r = 0$ を得る。以上より $\Phi(x) = (x - \lambda)\Psi(x)$. □

系 28.5.2. $\Phi(x)$ を n 次多項式とする。相異なる数 $\lambda_1, \dots, \lambda_r$ がそれぞれ方程式 $\Phi(x) = 0$ の解ならば、ある $n - r$ 次多項式 $\Psi(x)$ を用いて $\Phi(x) = (x - \lambda_1) \cdots (x - \lambda_r)\Psi(x)$ と書ける。

Proof. $\Phi(\lambda_1) = 0$ に因数定理を適用すれば、 $n - 1$ 次多項式 $\Psi_1(x)$ を用いて、 $\Phi(x) = (x - \lambda_1)\Psi_1(x)$ と書ける。この両辺に λ_2 を代入すると $(\lambda_2 - \lambda_1)\Psi_1(\lambda_2) = 0$ であり、仮定より $\lambda_2 - \lambda_1 \neq 0$ であるゆえ $\Psi_1(\lambda_2) = 0$ 。これに因数定理を適用すれば、 $n - 2$ 次多項式 $\Psi_2(x)$ を用いて $\Psi_1(x) = (x - \lambda_2)\Psi_2(x)$ と書ける。したがって、 $\Phi(x) = (x - \lambda_1)(x - \lambda_2)\Psi_2(x)$ である。これを順次繰り返し、 $\Phi(x) = (x - \lambda_1) \cdots (x - \lambda_r)\Psi_r(x)$ を得る。 □

次の定理は方程式論において基本となる定理である。この定理を示すには \mathbb{R}^2 に関する若干の位相的な知識、あるいは微分可能な複素数値関数 (これを正則関数という) に関する知識が必要となるため本論では証明を割愛する。

定理 28.5.3 (代数学の基本定理). 任意の \mathbb{C} 係数多項式 $\Phi(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ について、方程式 $\Phi(x) = 0$ は複素数の解を必ず持つ。

代数学の基本定理を認めると、任意の多項式が因数分解できることが分かる:

系 28.5.4. 任意の \mathbb{C} 係数多項式 $\Phi(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ は $\Phi(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ と因数分解できる.

Proof. n 次多項式 $\Phi(x)$ は代数学の基本定理により解 $\lambda_1 \in \mathbb{C}$ を持つ. ゆえに因数定理から $n-1$ 次多項式 $\Psi_1(x)$ を用いて $\Phi(x) = (x - \lambda_1)\Psi_1(x)$ と書ける. また, $n-1$ 次方程式 $\Psi_1(x) = 0$ は解 $\lambda_2 \in \mathbb{C}$ を持つゆえ $n-2$ 次多項式 $\Psi_2(x)$ を用いて $\Psi_1(x) = (x - \lambda_2)\Psi_2(x)$ と書けるゆえ, $\Phi(x) = (x - \lambda_1)(x - \lambda_2)\Psi_2(x)$ である. これを順次繰り返す, 因数分解 $\Phi(x) = (x - \lambda_1)(x - \lambda_2) \cdots (x - \lambda_n)$ を得る. \square

複素数を成分とする行列について, 系 28.2.2 を次のように言い換えてもよい.

系 28.5.5. n 次正方行列 A における特性方程式 $\Phi_A(t) = 0$ が \mathbb{C} において重解を持たないならば, A は複素数の範囲で対角化可能である.

Proof. 系 28.5.4 より $\Phi_A(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ と因数分解できる. 仮定より, $\lambda_1, \dots, \lambda_n$ は相異なる複素数ゆえ系 28.2.2 より A は対角化できる. \square

実数係数の n 次多項式における複素解とその共役についても述べておこう. 複素数 $z = a + bi$ (ただし $a, b \in \mathbb{R}$) に対して, $\bar{z} := a - bi$ を z の共役な複素数という. $z + \bar{z}$ および $z\bar{z}$ は共に実数である. また, $\overline{z + w} = \bar{z} + \bar{w}$ および $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$, $\overline{z^n} = \bar{z}^n$ が成り立つ. また, 実数 r について $\bar{r} = r$ である.

命題 28.5.6. \mathbb{R} 係数の n 次方程式 $\Phi(x) = 0$ が複素解 $z \in \mathbb{C}$ を持つならば, その共役な複素数 \bar{z} も $\Phi(x) = 0$ の解である.

Proof. $\Phi(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0$ とおく. $\Phi(z) = 0$ と仮定し, $\Phi(\bar{z}) = 0$ を導こう.

$$\begin{aligned} \Phi(\bar{z}) &= \bar{z}^n + a_{n-1}\bar{z}^{n-1} + \cdots + a_1\bar{z} + a_0 = \overline{z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0} \\ &= \overline{z^n + a_{n-1}z^{n-1} + \cdots + a_1z + a_0} = \overline{\Phi(z)} = \overline{0} = 0. \end{aligned}$$

\square

29 広義固有空間分解 (発展)

これまで、線形変換 $f: U \rightarrow U$ の表現行列を与える際に、上手く基底を選んでより複雑でない表現行列を得る方法について論じてきた。定理 28.2.1(1) の条件のもとでは、表現行列は対角行列に取れる。一方で、定理 28.2.1(1) の条件を満たさない線形写像に対して、どこまで表現行列を簡単にできるのだろうか。 f の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による表現行列を B としよう。 A は次で定義される行列であった。

$$A \text{ の第 } j \text{ 列} = \text{線形結合 } T(\mathbf{u}_j) = \sum_{i=1}^n a_{ij} \mathbf{u}_i \text{ に現れる係数を並べた列ベクトル.}$$

したがって、 A をより簡単な行列にせよという課題は、 $T(\mathbf{u}_j) = \sum_{i=1}^n a_{ij} \mathbf{u}_i$ に現れる係数の多くをいかに 0 にできるかという問題に帰着される。本節では、この問題への自然なアプローチとして不変部分空間および冪零部分空間の概念が導かれることを見る。

29.1 不変部分空間

U の基底を $\mathbf{u}_1, \dots, \mathbf{u}_n$ とする。このとき、 j 番目のベクトル \mathbf{u}_j として $f \in \text{End}(U)$ の固有ベクトルを選ぶと、 $f(\mathbf{u}_j) = \lambda \mathbf{u}_j$ ゆえ $f(\mathbf{u}_j)$ は \mathbf{u}_j 自身のみによる線形結合で書け、したがって表現行列の第 j 列は簡単な形 λe_j になる。ここで、 \mathbf{u}_j が固有ベクトルであるための条件が次のように書き換えられることに注意しよう。

$$f(\langle \mathbf{u}_j \rangle) \subset \langle \mathbf{u}_j \rangle.$$

そこで上の条件を一般化し、 U の基底の一部として次の条件を満たす組 $\mathbf{u}_l, \mathbf{u}_{l+1}, \dots, \mathbf{u}_{l+k}$ を考える：

$$f(\langle \mathbf{u}_l, \mathbf{u}_{l+1}, \dots, \mathbf{u}_{l+k} \rangle) \subset \langle \mathbf{u}_l, \mathbf{u}_{l+1}, \dots, \mathbf{u}_{l+k} \rangle$$

すると、各 \mathbf{u}_{l+j} ($j = 0, \dots, k$) は $\mathbf{u}_l, \mathbf{u}_{l+1}, \dots, \mathbf{u}_{l+k}$ のみによる線形結合で書けるゆえ、表現行列の第 l 列から第 $l+k$ 列は比較的簡単な成分になる (実際、 $l, l+1, \dots, l+k$ 成分以外は 0 となる)。この着想を一般的な立場から述べようとすれば次の定義に至る：

定義 29.1.1. $f: U \rightarrow U$ を線形変換とし、 W を U の部分空間とする。 W が $f(W) \subset W$ を満たすとき、 W は f -不変であると言う。

U がいくつかの f -不変部分空間に分解できるならば、次の命題に述べるような表現行列が得られる。これまでの考察からこの命題の主張は明らかであるが、一応証明を述べておこう。

命題 29.1.2. U を有限次元線形空間とし、 $f \in \text{End}(U)$ とする。各 W_1, \dots, W_r が U の f -不変部分空間であり、 W_γ ($\gamma = 1, \dots, r$) の基底 $\mathbf{u}_{\gamma,1}, \dots, \mathbf{u}_{\gamma,n_\gamma}$ をそれぞれ一つ選び、これらをすべて集めたベクトルの組 $\mathcal{B} = \{ \mathbf{u}_{\gamma,i} \mid \gamma = 1, \dots, r, i = 1, \dots, n_\gamma \}$ が U の基底になるとする。このとき、基底 \mathcal{B} に関する f の表現行列 A は次の形になる：

$$A = \begin{bmatrix} A_1 & & & \\ & A_2 & & \\ & & \ddots & \\ & & & A_r \end{bmatrix}$$

ここで、各 A_γ ($\gamma = 1, \dots, r$) はサイズ $\dim W_\gamma = n_\gamma$ の正方行列である。

Proof. $\dim U = n$ とする。 \mathcal{B} が U の基底となることから $n = \sum_{\gamma=1}^r n_\gamma$ である。そこで $n_0 = 0$ と置けば、各 $j = 1, \dots, n$ は次のように書ける：

$$j = n_0 + n_1 + n_2 + \dots + n_{\gamma-1} + k \quad (\text{ただし, } 1 \leq \gamma \leq r \text{ かつ } 1 \leq k \leq n_\gamma)$$

我々が示すべき事は A の第 j 列目の成分のうち第 $n_0 + \dots + n_{\gamma-1} + 1$ 成分から第 $n_0 + \dots + n_{\gamma-1} + n_\gamma$ 成分のほかがすべて 0 になることである。基底 \mathcal{B} における $j = n_0 + \dots + n_{\gamma-1} + k$ 番目のベクトル

は $\mathbf{u}_{n_\gamma, k} \in W_\gamma$ であり, W_γ が f -不変なことから $f(\mathbf{u}_{n_\gamma, k}) \in W_\gamma$ である. つまり, 基底 \mathcal{B} の線形結合で $f(\mathbf{u}_{n_\gamma, k})$ を表示した際に現れる係数のうち 0 でないものは $\mathbf{u}_{\gamma, 1}, \dots, \mathbf{u}_{\gamma, n_\gamma}$ の係数に限られる. したがって, A の第 j 列目に現れる成分のうち 0 でないものは $\mathbf{u}_{\gamma, 1}, \dots, \mathbf{u}_{\gamma, n_\gamma}$ に対応する成分, すなわち第 $n_0 + \dots + n_{\gamma-1} + 1$ から第 $n_0 + \dots + n_{\gamma-1} + n_\gamma$ 成分に限られる. \square

備考 29.1.3. 上の命題における各 A_γ は, 基底 $\mathbf{u}_{\gamma, 1}, \dots, \mathbf{u}_{\gamma, n_\gamma}$ に関する $f|_{W_\gamma} : W_\gamma \rightarrow W_\gamma$ の表現行列に等しい.

例 29.1.4. U が $f \in \text{End}(U)$ の固有ベクトルからなる基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を持つとき, 各 $W_j = \langle \mathbf{u}_j \rangle$ は U の f -不変部分空間である. これらに対して前命題を適用すれば, 各 A_j は $(1, 1)$ -行列であり, f の表現行列は対角行列となる.

29.2 冪零部分空間

本節の始めに提示した問題を前項とは別の視点から論じよう. 前項では, 望ましい基底の性質を導きだし, その性質のもとで表現行列が比較的簡単になることを見た. 本項ではこれとは逆の立場を取り, 表現行列が簡単な形をしていると仮定し, そのときに基底が満たすべき性質は何かを調べよう.

さて, 論ずべきことは U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ を上手く取ることで $T(\mathbf{u}_j) = \sum_{i=1}^n a_{ij} \mathbf{u}_i$ に現れる係数の多くをいかに 0 にできるかであった. そして, 最も都合が良い場合とは, \mathbf{u}_j が固有ベクトルであるとき, すなわち $f(\mathbf{u}_j)$ が \mathbf{u}_j 自身の線形結合で書ける場合であった. これが望めないとするならば, 次に考える最も単純な形は, \mathbf{u}_j 自身と, 基底をなす別のもう一つのベクトル \mathbf{u}_i との線形結合によって $f(\mathbf{u}_j)$ が表せる場合であろう. ここで更に踏み込んで, 各 $f(\mathbf{u}_j)$ が, \mathbf{u}_j と一つ隣のベクトル \mathbf{u}_{j-1} の線形結合で書ける場合, すなわち

$$f(\mathbf{u}_j) = s_{j-1} \mathbf{u}_{j-1} + r_j \mathbf{u}_j \quad (j = 2, \dots, n)$$

となる場合を考えよう. また, 簡単のために $f = T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ の場合を考えるとすれば, 代数学の基本定理により f の固有ベクトルは必ず存在する. ゆえに基底における最初の \mathbf{u}_1 を固有ベクトルである (つまり $f(\mathbf{u}_1) = r_1 \mathbf{u}_1$) としてよい. このとき f の表現行列は次のようになる.

$$A = \begin{bmatrix} r_1 & s_1 & & & & \\ & r_2 & s_2 & & & \\ & & & \ddots & & \\ & & & & r_{n-1} & s_{n-1} \\ & & & & & r_n \end{bmatrix}$$

ここで, $s_{j-1} = 0$ ならば \mathbf{u}_j は固有ベクトルである. 一方 $s_{j-1} \neq 0$ の場合は \mathbf{u}_j の代わりに $\frac{1}{s_{j-1}} \mathbf{u}_j$ を基底として取れば,

$$f\left(\frac{1}{s_{j-1}} \mathbf{u}_j\right) = \frac{1}{s_{j-1}} f(\mathbf{u}_j) = \frac{1}{s_{j-1}} (s_{j-1} \mathbf{u}_{j-1} + r_j \mathbf{u}_j) = \mathbf{u}_{j-1} + r_j \left(\frac{1}{s_{j-1}} \mathbf{u}_j\right)$$

であるから, $s_{j-1} = 0$ または $s_{j-1} = 1$ としてよい. また, A は上三角行列であるから, 各 r_j は A の固有値になっている (例 27.3.6). そこで, 列 r_1, \dots, r_n が各固有値ごとに行儀よく並んでいるような更に特別な場合を検討しよう. すなわち, f の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とし, U の基底 $\{\mathbf{u}_{\lambda_k, j} \mid k = 1, \dots, r, j = 1, \dots, n_k\}$ に関する表現行列 A が次のようになる場合である:

$$A = \begin{bmatrix} A_{\lambda_1} & & & & \\ & A_{\lambda_2} & & & \\ & & \ddots & & \\ & & & & A_{\lambda_r} \end{bmatrix}, \quad A_{\lambda_k} = \begin{bmatrix} \lambda_k & s_{\lambda_k, 1} & & & \\ & \lambda_k & s_{\lambda_k, 2} & & \\ & & & \ddots & \\ & & & & \lambda_k & s_{\lambda_k, n_k-1} \\ & & & & & \lambda_k \end{bmatrix} \quad (29.2.1)$$

ここで各 A_{λ_k} は n_k 次正方形行列であり, $s_{\lambda_k, j} = 0$ または $s_{\lambda_k, j} = 1$ である.

注意: 各固有値 λ_k の固有ベクトルは少なくとも一つは存在する. そこで A_{λ_k} の第 1 列目は固有ベクトルに対応する列であるとしてよく, ゆえに A_{λ_k} における $(1, 1)$ -成分の一つ上の成分は 0 であるとしてよい.

このとき, $\mathbf{u}_{\lambda_k, j}$ ($j = 2, \dots, n_k$) が固有ベクトルでなければ $s_{\lambda_k, j-1} = 1$ ゆえ $f(\mathbf{u}_{\lambda_k, j}) = \mathbf{u}_{\lambda_k, j-1} + \lambda_k \mathbf{u}_{\lambda_k, j}$ である. ゆえに,

$$\mathbf{u}_{\lambda_k, j-1} = f(\mathbf{u}_{\lambda_k, j}) - \lambda_k \mathbf{u}_{\lambda_k, j} = (f - \lambda_k I)(\mathbf{u}_{\lambda_k, j}).$$

つまり, $\mathbf{u}_{\lambda_k, j}$ に線形写像 $f - \lambda_k I$ を繰り返すことで $\mathbf{u}_{\lambda_k, j-1}, \mathbf{u}_{\lambda_k, j-2}, \dots$ が次々と得られ, これを続けると最後には λ_k に関する固有ベクトル $\mathbf{u}_{\lambda_k, j-\ell}$ を得る⁶⁷. また,

$$(f - \lambda_k I)(\mathbf{u}_{\lambda_k, j-\ell}) = f(\mathbf{u}_{\lambda_k, j-\ell}) - \lambda_k \mathbf{u}_{\lambda_k, j-\ell} = \lambda_k \mathbf{u}_{\lambda_k, j-\ell} - \lambda_k \mathbf{u}_{\lambda_k, j-\ell} = \mathbf{0}$$

であるから

$$\mathbf{u}_{\lambda_k, j} \xrightarrow{f - \lambda_k I} \mathbf{u}_{\lambda_k, j-1} \xrightarrow{f - \lambda_k I} \mathbf{u}_{\lambda_k, j-2} \xrightarrow{f - \lambda_k I} \cdots \xrightarrow{f - \lambda_k I} \mathbf{u}_{\lambda_k, j-\ell} \xrightarrow{f - \lambda_k I} \mathbf{0}.$$

すなわち,

$$\mathbf{u}_{\lambda_k, j} \in \text{Ker}(f - \lambda_k I)^{\ell+1}. \quad (29.2.2)$$

こうして我々は冪零部分空間の概念に至る:

定義 29.2.1. 有限次元線形空間 U 上の線形変換 $g : U \rightarrow U$ に対して, 部分空間の増大列 $\text{Ker } g \subset \text{Ker } g^2 \subset \text{Ker } g^3 \subset \dots$ は下の補題により十分大きい N について $\text{Ker } g^N = \text{Ker } g^{N+1} = \text{Ker } g^{N+2} = \dots$ を満たす. このとき, $\text{Ker } g^N$ を g の冪零部分空間という. また, 線形変換 $f : U \rightarrow U$ の固有値 λ に対して, 線形写像 $(f - \lambda I) : U \rightarrow U$ の冪零部分空間を, f の固有値 λ に関する広義固有空間あるいは一般固有空間 (**generalized eigenspace**) という. 本論では, これを記号 $\widetilde{W}(\lambda, f)$ で表す. 更に, 正方形行列 A について, $\widetilde{W}(\lambda, T_A)$ を $\widetilde{W}(\lambda, A)$ と表す.

備考 29.2.2. $W(\lambda, f) = \text{Ker}(\lambda I - f) = \text{Ker}(f - \lambda I)$ より, f の固有値 λ に関する固有空間は, 固有値 λ に関する広義固有空間の部分空間である.

以下では $g \in \text{End}(U)$ を一般の線形変換として扱うものの, $g = f - \lambda I$ のことと考えて読むと何を論じているのかイメージが湧くことと思う.

補題 29.2.3. 有限次元線形空間 U 上の線形変換 $g : U \rightarrow U$ において次が成り立つ.

- (1) $\text{Ker } g \subset \text{Ker } g^2 \subset \text{Ker } g^3 \subset \dots$.
- (2) $\text{Ker } g^n = \text{Ker } g^{n+1}$ ならば, $\text{Ker } g^n = \text{Ker } g^{n+1} = \text{Ker } g^{n+2} = \text{Ker } g^{n+3} = \dots$.
- (3) $\dim U = N$ ならば, $\text{Ker } g^N = \text{Ker } g^{N+1} = \text{Ker } g^{N+2} = \dots$.

Proof. (1): $\mathbf{u} \in \text{Ker } g^n$ とすれば, $g^n(\mathbf{u}) = \mathbf{0}$. よって, $g^{n+1}(\mathbf{u}) = g(g^n(\mathbf{u})) = g(\mathbf{0}) = \mathbf{0}$ ゆえ $\mathbf{u} \in \text{Ker } g^{n+1}$.

(2): $\text{Ker } g^{n+1} = \text{Ker } g^{n+2}$ を示せば, あとは帰納的に各 $\text{Ker } g^{n+m}$ がすべて一致することが分かる. $\text{Ker } g^{n+1} \subset \text{Ker } g^{n+2}$ は (1) で示したゆえ $\text{Ker } g^{n+2} \subset \text{Ker } g^{n+1}$ を示そう. 各 $\mathbf{u} \in \text{Ker } g^{n+2}$ について, $\mathbf{0} = g^{n+2}(\mathbf{u}) = g^{n+1}(g(\mathbf{u}))$ より $g(\mathbf{u}) \in \text{Ker } g^{n+1} = \text{Ker } g^n$. よって, $g(\mathbf{u}) \in \text{Ker } g^n$ ゆえ $g^n(g(\mathbf{u})) = \mathbf{0}$. つまり, $\mathbf{u} \in \text{Ker } g^{n+1}$ である.

(3): g が単射ならば $\{\mathbf{0}\} = \text{Ker } g = \text{Ker } g^2 = \text{Ker } g^3 = \dots$ である. そこで, 単射でない場合を考えよう. このとき $\dim \text{Ker } g \geq 1$ である. $\text{Ker } g^n$ が $\text{Ker } g^{n+1}$ よりも真に大きくなるとき, それらの次元も真に大きくなる. 仮に, 各 $n = 1, \dots, N$ について $\text{Ker } g^{n+1}$ が $\text{Ker } g^n$ よりも真に大きくなるとすれば,

$$\dim \text{Ker } g^{N+1} \geq \dim \text{Ker } g^N + 1 \geq (\dim \text{Ker } g^{N-1} + 1) + 1 \geq \dots \geq \dim \text{Ker } g + N \geq 1 + N$$

となり $\dim U = N$ に矛盾してしまう. したがって, $n = 1, \dots, N$ のいずれかにおいて $\text{Ker } g^n = \text{Ker } g^{n+1}$ とならなければならない. 以降は (2) よりすべて一致する. \square

⁶⁷少なくとも $\mathbf{u}_{\lambda_k, 1}$ は λ_k に関する固有ベクトルであるから, $f - \lambda_k I$ を j 回ほどこす間に λ_k に関する固有ベクトルが必ず得られる.

これまでの議論をまとめると、式 (29.2.1) のような形の表現行列を得るためには、式 (29.2.2) により U の基底として各固有値に関する広義固有空間の元を取ることが少なくとも必須であることが分かった。実は、任意の $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ について、これが可能であることが知られている (定理 29.3.1)。

f の各固有値に関する広義固有空間が f -不変であることを確認しておこう。

補題 29.2.4. $f, g : U \rightarrow U$ を可換な線形変換とする (すなわち $g \circ f = f \circ g$)。このとき、各 $n \in \mathbb{N}$ について $W = \text{Ker } g^n$ は f -不変部分空間である。

Proof. 各 $\mathbf{u} \in W$ について、 $f(\mathbf{u}) \in W$ を示したい。そのためには $g^n(f(\mathbf{u})) = \mathbf{0}$ を言えばよい。 g^n と f が可換になること、および $g^n(\mathbf{u}) = \mathbf{0}$ から、

$$g^n(f(\mathbf{u})) = g^n \circ f(\mathbf{u}) = f \circ g^n(\mathbf{u}) = f(g^n(\mathbf{u})) = f(\mathbf{0}) = \mathbf{0}.$$

□

線形変換 g 自身は g と可換である。また、線形変換 f および $f - \lambda I$, $f - \gamma I$ はそれぞれ可換である (命題 21.4.5)。これらの事実と先の補題から次を得る。

系 29.2.5. $f, g : U \rightarrow U$ を線形変換とする。各 $\lambda, \gamma \in \mathbb{K}$ について次が成り立つ。

- (1) $g : U \rightarrow U$ の冪零部分空間 W は g -不変部分空間である。
- (2) $f - \lambda I$ の冪零部分空間 \widetilde{W} は f -不変部分空間である。とくに、 f の固有値 λ に関する広義固有空間 $\widetilde{W}(\lambda, f)$ は f -不変である。
- (3) $f - \lambda I$ の冪零部分空間 \widetilde{W} は $(f - \gamma I)$ -不変部分空間である。

29.3 広義固有空間分解

表現行列として対角行列を取れる線形変換においては、各ベクトルが固有ベクトルの線形結合で表されるのであった。これに対して一般の線形変換については、その特性多項式が因数分解できるならば各ベクトルを広義固有空間の元の線形結合で書くことができる。この事実を主張する次の定理の証明は次節で与えるとして、本節の後半ではこの定理から導かれる基本的事実について解説しよう。

定理 29.3.1 (広義固有空間分解). 有限次元線形空間上の線形変換 $f : U \rightarrow U$ の相異なる固有値を $\lambda_1, \dots, \lambda_r$ とし、 f の特性多項式が

$$\Phi_f(t) = (t - \lambda_1)^{n_1} (t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$$

であるとする。このとき、 $\dim \widetilde{W}(\lambda_k, f) = n_k$ ($k = 1, \dots, r$) が成り立つ。また、各 $\widetilde{W}(\lambda_k, f)$ の基底 $\mathbf{u}_{\lambda_k, 1}, \dots, \mathbf{u}_{\lambda_k, n_k}$ をそれぞれ一つ選び、これらをすべて集めてベクトルの組

$$\mathcal{B} = \{ \mathbf{u}_{\lambda_k, j} \mid k = 1, \dots, r, j = 1, \dots, n_k \}$$

を作れば、 \mathcal{B} は U の基底になる。

例 29.3.2. (1) 定理 29.3.1 の設定のもとで、次のような $\widetilde{W}(\lambda_k, f)$ の基底の選び方を考えよう：

$\mathbf{u}_{\lambda_k, 1}, \dots, \mathbf{u}_{\lambda_k, n_k}$ の選び方として、まず $\text{Ker}(f - \lambda_k I)$ の基底を選び、それに $\text{Ker}(f - \lambda_k I)^2$ の元を付け加えて $\text{Ker}(f - \lambda_k I)^2$ の基底とし、更に $\text{Ker}(f - \lambda_k I)^3$ の元を付け加えて $\text{Ker}(f - \lambda_k I)^3$ の基底とし… という操作を繰り返して $\widetilde{W}(\lambda_k, f)$ の基底 $\mathbf{u}_{\lambda_k, 1}, \dots, \mathbf{u}_{\lambda_k, n_k}$ を得る。ここで $\mathbf{u}_{\lambda_k, j} \in \text{Ker}(f - \lambda_k I)^\ell$ とすれば、

$$\begin{aligned} (f - \lambda_k I)(\mathbf{u}_{\lambda_k, j}) &\in \text{Ker}(f - \lambda_k I)^{\ell-1} \\ f(\mathbf{u}_{\lambda_k, j}) - \lambda_k \mathbf{u}_{\lambda_k, j} &\in \text{Ker}(f - \lambda_k I)^{\ell-1} \\ f(\mathbf{u}_{\lambda_k, j}) &\in \text{Ker}(f - \lambda_k I)^{\ell-1} + \lambda_k \mathbf{u}_{\lambda_k, j}. \end{aligned}$$

上式は、各 $f(\mathbf{u}_{\lambda_k, j})$ が $\mathbf{u}_{\lambda_k, 1}, \mathbf{u}_{\lambda_k, 2}, \dots, \mathbf{u}_{\lambda_k, j}$ の線形結合で書けることを述べている。したがって、上で与えた基底に対して命題 29.1.2 を適用すれば各 A_k は上三角行列となり、したがって基底 B による f の表現行列は上三角行列となる。

- (2) X を n 次正方行列とする。 $T_X : \mathbb{C}^n \rightarrow \mathbb{C}^n$ に対して (1) を適用し、(1) で与えた基底を各固有値ごとに並べた行列を P とすれば $A = P^{-1}XP$ である。つまり、任意の正方行列は複素数の範囲において、ある上三角行列と相似になることが分かる。なお、内積空間の単元においては、別の文脈から直交行列を用いた行列の上三角化が論じられる。

備考 29.3.3. 実は、巧妙に $W(\lambda_k, f)$ の基底を選ぶことにより、表現行列を式 29.2.1 のような形できることが知られており、これをジョルダン標準形と言う。この詳しい構成については次節で述べる。

具体的なくつかの線形写像について、どのようなベクトルが広義固有空間の元となるか見ておこう。ここで述べる例は、線形漸化式の一般項や線形常微分方程式の一般解の構造の理解への助けとなるものである。

まず、一般論として次の補題を用意する。これは、例 17.2.3 で述べたことの一般化に他ならない。

補題 29.3.4. 線形変換 $g : U \rightarrow U$ および $\mathbf{u} \in U$ について、 $g^n(\mathbf{u}) \neq \mathbf{0}$ かつ $g^{n+1}(\mathbf{u}) = \mathbf{0}$ ならば、 $n+1$ 個のベクトルからなる組 $\mathbf{u}, g(\mathbf{u}), g^2(\mathbf{u}), \dots, g^n(\mathbf{u})$ は線形独立である。

Proof. 各 $k = 0, \dots, n$ について、 $k+1$ 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-k}(\mathbf{u})$ が線形独立であることを k に関する帰納法で示そう。 $k \leq n$ を満たす自然数 k について、 k 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-(k-1)}(\mathbf{u})$ が線形独立であると仮定し、 $k+1$ 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-k}(\mathbf{u})$ の線形独立性を示す。線形関係 $\sum_{i=0}^k r_i g^{n-i}(\mathbf{u}) = \mathbf{0}$ を仮定し、この両辺に g をほどこせば、

$$\begin{aligned} r_0 g^{n-0+1}(\mathbf{u}) + r_1 g^{n-1+1}(\mathbf{u}) + r_2 g^{n-2+1}(\mathbf{u}) + \dots + r_{k-1} g^{n-k+1}(\mathbf{u}) &= \mathbf{0} \\ r_1 g^n(\mathbf{u}) + r_2 g^{n-1}(\mathbf{u}) + \dots + r_{k-1} g^{n-(k-1)}(\mathbf{u}) &= \mathbf{0} \end{aligned}$$

$g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-(k-1)}(\mathbf{u})$ は線形独立であったゆえ $r_1 = r_2 = \dots = r_{k-1} = 0$ 。ゆえに $r_0 g^n(\mathbf{u}) = \mathbf{0}$ であり、 $g^n(\mathbf{u}) \neq \mathbf{0}$ より $r_0 = 0$ 。以上より $k+1$ 個の組 $g^n(\mathbf{u}), g^{n-1}(\mathbf{u}), \dots, g^{n-k}(\mathbf{u})$ は線形独立である。□

次の例において $\mathbb{R} = \mathbb{C}$ の場合は複素数値関数 (正則関数) についての解析の知識が必要となる。ここでは、複素数値関数においても実数値関数の場合と同様の微分公式が満たされることを既知の事実であるとして話を進めよう⁶⁸。

例 29.3.5. $D : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ を微分作用素 $D(y) = y'$ とし、また $\lambda \in \mathbb{R}$ とする。

- (1) 非負整数 n について $x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^{n+1}$ かつ $x^n e^{\lambda x} \notin \text{Ker}(D - \lambda I)^n$ 。
(2) $e^{\lambda x}, x e^{\lambda x}, x^2 e^{\lambda x}, \dots, x^n e^{\lambda x}$ は線形独立である。

Proof. (1): n に関する帰納法で示す。 $n = 0$ について $e^{\lambda x}$ は固有値 λ の固有ベクトルゆえ $e^{\lambda x} \in \text{Ker}(D - \lambda I)$ 。次に、 $x^{n-1} e^{\lambda x} \in \text{Ker}(D - \lambda I)^n$ かつ $x^{n-1} e^{\lambda x} \notin \text{Ker}(D - \lambda I)^{n-1}$ を仮定して $x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^{n+1}$ および $x^n e^{\lambda x} \notin \text{Ker}(D - \lambda I)^n$ を示そう。

$$\begin{aligned} (D - \lambda I)(x^n e^{\lambda x}) &= (x^n e^{\lambda x})' - \lambda x^n e^{\lambda x} = (n x^{n-1} e^{\lambda x} + x^n \lambda e^{\lambda x}) - \lambda x^n e^{\lambda x} \\ &= n x^{n-1} e^{\lambda x} \in \text{Ker}(D - \lambda I)^n. \end{aligned}$$

ゆえに $x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^{n+1}$ である。また、 $n x^{n-1} e^{\lambda x} \notin \text{Ker}(D - \lambda I)^{n-1}$ ゆえ $x^n e^{\lambda x} \notin \text{Ker}(D - \lambda I)^n$ である⁶⁹。

⁶⁸ 正則関数に関する解析学のことを関数論という。正則関数における微分公式の証明については、関数論の入門書を参照されたい。

⁶⁹ 何故なら、仮に $\mathbf{u} = x^n e^{\lambda x} \in \text{Ker}(D - \lambda I)^n$ と仮定すれば $(D - \lambda I)(\mathbf{u}) \in \text{Ker}(D - \lambda I)^{n-1}$ となり、これは $(D - \lambda I)(\mathbf{u}) = n x^{n-1} e^{\lambda x} \notin \text{Ker}(D - \lambda I)^{n-1}$ に矛盾してしまう。

(2): $g = D - \lambda I$ および $\mathbf{u} = x^n e^{\lambda x}$ について前補題を適用すると, $N + 1$ 個の組

$$x^n e^{\lambda x}, nx^{n-1} e^{\lambda x}, n(n-1)x^{n-2} e^{\lambda x}, \dots, n! e^{\lambda x}$$

は線形独立である. ゆえに, これらにスカラー倍をほどこした組 $x^n e^{\lambda x}, x^{n-1} e^{\lambda x}, x^{n-2} e^{\lambda x}, \dots, e^{\lambda x}$ も線形独立である. \square

例 29.3.6. $S : \mathbb{K}^{\mathbb{N}} \rightarrow \mathbb{K}^{\mathbb{N}}$ をシフト作用素 $S((x_n)_{n \in \mathbb{N}}) := (x_{n+1})_{n \in \mathbb{N}}$ とする. $\lambda \neq 0$ および非負整数 N について, 数列 $\mathbf{x}_N^\lambda \in \mathbb{R}^{\mathbb{N}}$ を次で定める.

$$\mathbf{x}_N^\lambda := (1, 2^N \lambda, 3^N \lambda^2, 4^N \lambda^3, \dots) \quad (\text{つまり } \mathbf{x}_N^\lambda = (n^N \lambda^{n-1})_{n \in \mathbb{N}}).$$

(1) $\mathbf{x}_N^\lambda \in \text{Ker}(S - \lambda I)^{N+1}$ かつ $\mathbf{x}_N^\lambda \in \text{Ker}(S - \lambda I)^N$.

(2) $\mathbf{x}_0^\lambda, \mathbf{x}_1^\lambda, \dots, \mathbf{x}_N^\lambda$ は線形独立である.

Proof. ここでは \mathbf{x}_N^λ を \mathbf{x}_N と略すことにする.

(1): N に関する帰納法で示す. $N = 0$ の場合, \mathbf{x}_0 は初項 1 公比 λ の等比数列であり, これは S の λ に関する固有ベクトルであるから $\mathbf{x}_0 \in \text{Ker}(S - \lambda I)$ である. 次に各 $\mathbf{x}_0, \dots, \mathbf{x}_{N-1}$ について $\mathbf{x}_k \in \text{Ker}(S - \lambda I)^{k+1}$ ($k = 0, \dots, N-1$), および $\mathbf{x}_{N-1} \notin \text{Ker}(S - \lambda I)^{N-1}$ (つまり $(S - \lambda I)^{N-1}(\mathbf{x}_{N-1}) \neq \mathbf{0}$) を仮定して, $\mathbf{x}_N \in \text{Ker}(S - \lambda I)^{N+1}$ および $\mathbf{x}_N \notin \text{Ker}(S - \lambda I)^N$ を示そう.

$$\begin{aligned} (S - \lambda I)(\mathbf{x}_N) &= (S - \lambda I)\left((n^N \lambda^{n-1})_{n \in \mathbb{N}}\right) = \left((n+1)^N \lambda^n\right)_{n \in \mathbb{N}} - \lambda \left(n^N \lambda^{n-1}\right)_{n \in \mathbb{N}} \\ &= \left(\left((n+1)^N - n^N\right) \lambda^n\right)_{n \in \mathbb{N}} = \lambda \left(\left((n+1)^N - n^N\right) \lambda^{n-1}\right)_{n \in \mathbb{N}}. \end{aligned}$$

ここで, $(n+1)^N - n^N$ は n に関する $N-1$ 次の多項式であるから $(n+1)^N - n^N = \sum_{k=0}^{N-1} a_k n^k$ と書ける (ここで $a_{N-1} \neq 0$). ゆえに

$$\begin{aligned} \lambda \left(\left((n+1)^N - n^N\right) \lambda^{n-1}\right)_{n \in \mathbb{N}} &= \lambda \left(\left(\sum_{k=0}^{N-1} a_k n^k\right) \lambda^{n-1}\right)_{n \in \mathbb{N}} = \lambda \left(\sum_{k=0}^{N-1} a_k n^k \lambda^{n-1}\right)_{n \in \mathbb{N}} \\ &= \lambda \sum_{k=0}^{N-1} (a_k n^k \lambda^{n-1})_{n \in \mathbb{N}} = \lambda \sum_{k=0}^{N-1} a_k \mathbf{x}_k \in \text{Ker}(S - \lambda I)^N. \end{aligned}$$

以上より, $(S - \lambda I)(\mathbf{x}_N) \in \text{Ker}(S - \lambda I)^N$ ゆえ $\mathbf{x}_N \in \text{Ker}(S - \lambda I)^{N+1}$. また,

$$\begin{aligned} (S - \lambda I)^N(\mathbf{x}_N) &= (S - \lambda I)^{N-1}\left((S - \lambda I)(\mathbf{x}_N)\right) = (S - \lambda I)^{N-1}\left(\lambda \sum_{k=0}^{N-1} a_k \mathbf{x}_k\right) \\ &= \lambda \sum_{k=0}^{N-1} a_k (S - \lambda I)^{N-1}(\mathbf{x}_k) = \lambda a_{N-1} (S - \lambda I)^{N-1}(\mathbf{x}_{N-1}) \neq \mathbf{0}. \end{aligned}$$

ゆえに $\mathbf{x}_N \notin \text{Ker}(S - \lambda I)^N$.

(2): $g = S - \lambda I$ について前補題を適用すれば, $\mathbf{x}_N, g(\mathbf{x}_N), \dots, g^N(\mathbf{x}_N)$ は線形独立である. また, (1) の証明で行った計算によれば

$$\langle \mathbf{x}_N, g(\mathbf{x}_N), \dots, g^N(\mathbf{x}_N) \rangle \subset \langle \mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0 \rangle$$

である. したがって

$$N + 1 = \dim \langle \mathbf{x}_N, g(\mathbf{x}_N), \dots, g^N(\mathbf{x}_N) \rangle \leq \dim \langle \mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0 \rangle \leq N + 1$$

より $\dim \langle \mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0 \rangle = N + 1$. これと命題 22.3.3(3) を合わせて, $\mathbf{x}_N, \mathbf{x}_{N-1}, \dots, \mathbf{x}_0$ の線形独立性を得る. \square

29.4 ケーリー・ハミルトンの定理

定理 29.3.1 を認めると、直ちにケーリー・ハミルトンの定理が導かれる。証明の筋書きが命題 28.4.2 と類似していることにお気づきだろうか。

定理 29.4.1 (ケーリー・ハミルトンの定理). n 次正方行列 A について $\Phi_A(A) = O$.

Proof. 複素数の範囲において $\Phi_A(t) = (t - \lambda_1)^{n_1}(t - \lambda_2)^{n_2} \cdots (t - \lambda_r)^{n_r}$ と因数分解しよう。ここで、 $\lambda_1, \dots, \lambda_r$ は相異なる A の固有値である。 $T_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ に対して定理 29.3.1 を適用すれば、 \mathbb{C}^n の各元は広義固有空間の元の線形結合で表せる。したがって、各 $\mathbf{v}_k \in \widetilde{W}(\lambda_k, A)$ について $\Phi_A(A)\mathbf{v}_k = \mathbf{0}$ を示せば十分である。定理 29.3.1 より $\dim \widetilde{W}(\lambda_k, A) = n_k$ であり、線形写像

$$g = (T_A - \lambda_k I)|_{\widetilde{W}(\lambda_k, A)} : \widetilde{W}(\lambda_k, A) \rightarrow \widetilde{W}(\lambda_k, A)$$

に対して補題 29.2.3(3) を適用すれば $\widetilde{W}(\lambda_k, A) = \text{Ker}(T_A - \lambda_k I)^{n_k}$ 。つまり

$$(A - \lambda_k E)^{n_k} \mathbf{v}_k = (T_A - \lambda_k I)^{n_k}(\mathbf{v}_k) = \mathbf{0}$$

である。 $\Phi(t)$ は、 $n - n_k$ 次多項式 $\Theta(t)$ を用いて $\Phi_A(t) = \Theta(t)(t - \lambda_k)^{n_k}$ と書けることから、

$$\Phi_A(A)\mathbf{v}_i = \Theta(A)(A - \lambda_k E)^{n_k} \mathbf{v}_k = \Theta(A)\mathbf{0} = \mathbf{0}.$$

以上より $\Phi_A(A) = O$. □

系 29.4.2. 有限次元線形空間上の線形変換 $f : U \rightarrow U$ について、 $\Phi_f(f) = \mathbf{0}_U$.

Proof. U のある基底に関する表現行列を A とすれば、 $\Phi_f(A) = \Phi_A(A) = O$ 。系 26.2.4 により線形変換 $\Phi_f(f)$ の表現行列は $\Phi_f(A) = O$ であり、したがって $\Phi_f(f)$ は任意のベクトルを $\mathbf{0}_U$ に写す定値写像である。 □

f が実数体 \mathbb{R} 上の線形空間 U における線形変換であり、実数の範囲において特性多項式が因数分解されない場合や、特に特性方程式が実数解を持たない場合 (つまり固有ベクトルが存在しない場合) においても系 29.4.2 は成り立っている。上の議論は、実数に限った話題であっても複素数に範囲を広げておくことで理解が容易になる典型例と言えるだろう。もちろん、ケーリー・ハミルトンの定理や系 29.4.2 の証明において、複素数における議論が必須というわけではない。本節の最後に、実数の範囲に限って論じた系 29.4.2 の別証明を紹介しておこう。

よりみち (ケーリー・ハミルトンの定理の間違った理解).

一見すると次の式変形によってケーリー・ハミルトンの定理が示されたかに思えるが、これは正しくない:

$$\Phi_A(t) = |tE - A| \text{ に } t = A \text{ を代入して、 } \Phi_A(A) = |AE - A| = |O| = 0.$$

そもそも上の式変形における左辺 $\Phi_A(A)$ は n 次正方行列であり、右辺の 0 は数であるから、これでは行列と数という異なる概念を等式で結んでしまったことになる。勘違いの原因は、 \mathbb{K} の元と $M_n(\mathbb{K})$ の元を混同したことにある。あるいは、多項式 $\Phi_A(t)$ に行列 X を代入したものが $\Phi_A(X)$ であると誤認してしまったと言ってもよい。 $\Phi_A(X)$ の正確な定義は、定義 21.4.1 で述べたように、 $\Phi_A(t)$ における t^n を X^n に置き換え、定数項 a を aE に置き換えた式で表される行列のことであった。

上のコラムにおける勘違いを反省し、また、ケーリー・ハミルトンの定理の別証明のための準備として、 $\text{End}(U)$ の元を成分とする行列を考えよう。 $\text{End}(U)$ の間には和と積が定義されていたゆえ、 \mathbb{K} を成分とする行列の場合と同様にして $\text{End}(U)$ 成分の行列の和、積、行列式などが定義できる。例えば、 $X = [g_{ij}]$ を $\text{End}(U)$ 成分の n 次正方行列とすれば、その行列式

$$\det X := \sum_{\sigma \in \mathfrak{S}_n} \text{sgn}(\sigma)(g_{1\sigma(1)} \circ g_{2\sigma(2)} \circ \cdots \circ g_{n\sigma(n)})$$

は $\text{End}(U)$ の元, すなわち何らかの線形変換である.

さて, \mathbb{R} 成分の行列について証明したことを並行して論じることで, $\text{End}(U)$ 成分の行列に関する多くの命題が得られる. ここで注意すべきことは, 行列式の性質などの証明において各成分どうしの積の可換性 ($ab = ba$) を用いた部分がいくつかあり, 可換性が成り立つとは限らない $\text{End}(U)$ 成分の行列においては, これらを適用できない点である. しかし, 互いに可換な元のみを成分とする特別な行列のみを考えるのであれば, その限りではない. 例えば, 各 $g_{ij} \in \text{End}(U)$ が互いに可換であるとき, 正方行列 $X = [g_{ij}]$ について, その余因子行列を \tilde{X} とすれば, 13.2 項で論じたことと同様にして

$$\tilde{X}X = \begin{bmatrix} \det X & & \\ & \ddots & \\ & & \det X \end{bmatrix}$$

が成り立つ.

系 29.4.2 の別証明. U の基底 $\mathbf{u}_1, \dots, \mathbf{u}_n$ による f の表現行列を $A = [a_{ij}]$ とすれば, $[f(\mathbf{u}_1), \dots, f(\mathbf{u}_n)] = [\mathbf{u}_1, \dots, \mathbf{u}_n]A$ である. すなわち,

$$\begin{aligned} f(\mathbf{u}_1) &= a_{11}\mathbf{u}_1 + a_{21}\mathbf{u}_2 + \dots + a_{n1}\mathbf{u}_n \\ f(\mathbf{u}_2) &= a_{12}\mathbf{u}_1 + a_{22}\mathbf{u}_2 + \dots + a_{n2}\mathbf{u}_n \\ &\dots \\ f(\mathbf{u}_n) &= a_{1n}\mathbf{u}_1 + a_{2n}\mathbf{u}_2 + \dots + a_{nn}\mathbf{u}_n \end{aligned}$$

である. 上の連立式を移項し, また恒等写像 I をあえて用いれば次のように書ける:

$$\begin{aligned} (f - a_{11}I)(\mathbf{u}_1) + (-a_{21}I)(\mathbf{u}_2) + \dots + (-a_{n1}I)(\mathbf{u}_n) &= \mathbf{0}_U \\ (-a_{12}I)(\mathbf{u}_1) + (f - a_{22}I)(\mathbf{u}_2) + \dots + (-a_{n2}I)(\mathbf{u}_n) &= \mathbf{0}_U \\ &\dots \\ (-a_{1n}I)(\mathbf{u}_1) + (-a_{2n}I)(\mathbf{u}_2) + \dots + (f - a_{nn}I)(\mathbf{u}_n) &= \mathbf{0}_U \end{aligned}$$

ここで, $\text{End}(U)$ の元 g と, U の元 \mathbf{u} の積を $g\mathbf{u} := g(\mathbf{u})$ と定め, $\text{End}(U)$ の元を成分とする行列と, ベクトルを成分とする列の間の積を導入すれば, 上式は次のように表せる:

$$\begin{bmatrix} f - a_{11}I & -a_{21}I & \dots & -a_{n1}I \\ -a_{12}I & f - a_{22}I & \dots & -a_{n2}I \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n}I & -a_{2n}I & \dots & f - a_{nn}I \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \mathbf{u}_2 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0}_U \\ \mathbf{0}_U \\ \vdots \\ \mathbf{0}_U \end{bmatrix}.$$

上の左辺に現れる行列を X とすれば, X の各成分は互いに可換である (命題 21.4.5). また, $\text{End}(U)$ 成分の行列 Y, Z , および U 成分の列ベクトル \mathbf{x} の間の積について結合律 $Y(Z\mathbf{x}) = (YZ)\mathbf{x}$ が成り立つことが \mathbb{R} 成分行列の場合と同様にして示され, したがって上の両辺に左から X の余因子行列 \tilde{X} をかけることで

$$\begin{bmatrix} \det X & & \\ & \ddots & \\ & & \det X \end{bmatrix} \begin{bmatrix} \mathbf{u}_1 \\ \vdots \\ \mathbf{u}_n \end{bmatrix} = \begin{bmatrix} \mathbf{0}_U \\ \vdots \\ \mathbf{0}_U \end{bmatrix} \quad (29.4.1)$$

を得る. 一方, 次の計算により $\det X = \Phi_f(f)$ である:

$$\begin{aligned} \det X &= \text{“多項式 } \Phi_{tA}(t) = |tE - {}^tA| \text{ における変数 } t \text{ を } f \text{ に置き換えた写像”} \\ &= \Phi_{tA}(f) = \Phi_A(f) = \Phi_f(f). \end{aligned}$$

したがって式 (29.4.1) は, 各 $j = 1, \dots, n$ について $\Phi_f(f)(\mathbf{u}_j) = \mathbf{0}_U$ であることを意味する. すなわち, 線形変換 $\Phi_f(f)$ はすべての元を $\mathbf{0}_U$ に写す写像である. \square