

数学 II 演習 (第 10 回) の略解

目次

1 問 1 の解答	1
2 対称行列とは	4
3 直交行列とは	12
4 対称変換とは	20
5 直交補空間について	27
6 対称行列の固有ベクトル空間分解について	33
7 直交行列による対称行列の対角化について	37
8 問 1 を見直すと	40
9 多変数関数の合成関数に対する連鎖律について	47
10 問 2 の解答	51
11 線型空間上の内積とは	57
12 問 2 を見直すと *	66

1 問 1 の解答

(1) 関数 $f(x, y, z)$ のそれぞれの偏導関数を求めてみると,

$$\begin{cases} \frac{\partial f}{\partial x} = 3x - 2y + 2 \\ \frac{\partial f}{\partial y} = -2x + 2y - 2z + 2 \\ \frac{\partial f}{\partial z} = -2y + z \end{cases}$$

となることが分かります. したがって, 点 $p = (x, y, z) \in \mathbb{R}^3$ が関数 $f(x, y, z)$ の臨界点となるための条件は,

$$\begin{cases} 3x - 2y = -2 \\ x - y + z = 1 \\ 2y - z = 0 \end{cases} \quad (1)$$

という連立一次方程式で与えられることが分かります. そこで, (1) 式の連立一次方程式の解を求めてみると, 関数 $f(x, y, z)$ の臨界点 $p_0 \in \mathbb{R}^3$ は,

$$p_0 = (0, 1, 2) \in \mathbb{R}^3$$

となることが分かります.

- (2) 関数 $f(x, y, z)$ の二階偏導関数 $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial x \partial y}, \dots, \frac{\partial^2 f}{\partial z^2}$ を求めてみると, 関数 $f(x, y, z)$ の臨界点 p_0 におけるヘッシアンは,

$$A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix}$$

となることが分かります.¹

- (3) 行列 A の特性多項式 $\varphi_A(t) = \det(tI - A)$ を計算してみると,

$$\begin{aligned} \varphi_A(t) &= \begin{vmatrix} t-3 & 2 & 0 \\ 2 & t-2 & 2 \\ 0 & 2 & t-1 \end{vmatrix} \\ &= (t-3) \cdot \begin{vmatrix} t-2 & 2 \\ 2 & t-1 \end{vmatrix} - 2 \cdot \begin{vmatrix} 2 & 0 \\ 2 & t-1 \end{vmatrix} \quad (\text{1列目で展開}) \\ &= (t-3)\{(t-2)(t-1) - 4\} - 4(t-1) \\ &= (t-3)(t^2 - 3t - 2) - 4(t-1) \\ &= t^3 - 6t^2 + 3t + 10 \\ &= (t+1)(t-2)(t-5) \end{aligned} \quad (2)$$

となることが分かります. したがって, 行列 A の固有値 λ は,

$$\lambda = -1, 2, 5$$

となることが分かります. そこで, 第6回問1のときと同様にして, それぞれの固有値 λ に対する固有ベクトル空間 $V(\lambda)$ を求めてみると, 次のようになることが分かります.

¹今の場合, 関数 $f(x, y, z)$ は二次の多項式なので, $f(x, y, z)$ のすべての二階偏導関数は定数関数になり, ヘッシアンは, 点 $p = (x, y, z) \in \mathbb{R}^3$ によらず, 定数行列になることが分かります.

(i) $\lambda = -1$ のとき, 固有ベクトル空間 $V(-1)$ は,

$$\begin{aligned} V(-1) &= \left\{ t \cdot \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \in \mathbb{R}^3 \mid t \in \mathbb{R} \right\} \\ &= \mathbb{R} \cdot \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \end{aligned}$$

となることが分かります. 特に, 長さが 1 の固有ベクトルとして,

$$\mathbf{p}_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \in V(-1)$$

が取れることが分かります.²

(ii) $\lambda = 2$ のとき, 固有ベクトル空間 $V(2)$ は,

$$\begin{aligned} V(2) &= \left\{ t \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \in \mathbb{R}^3 \mid t \in \mathbb{R} \right\} \\ &= \mathbb{R} \cdot \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \end{aligned}$$

となることが分かります. 特に, 長さが 1 の固有ベクトルとして,

$$\mathbf{p}_2 = \frac{1}{3} \begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} \in V(2)$$

が取れることが分かります.³

(iii) $\lambda = 5$ のとき, 固有ベクトル空間 $V(5)$ は,

$$\begin{aligned} V(5) &= \left\{ t \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \in \mathbb{R}^3 \mid t \in \mathbb{R} \right\} \\ &= \mathbb{R} \cdot \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \end{aligned}$$

²もちろん, 長さが 1 の固有ベクトルとして, \mathbf{p}_1 の代わりに, $-\mathbf{p}_1$ というベクトルを考えても構いません.

³もちろん, 長さが 1 の固有ベクトルとして, \mathbf{p}_2 の代わりに, $-\mathbf{p}_2$ というベクトルを考えても構いません.

となることが分かります. 特に, 長さが 1 の固有ベクトルとして,

$$\mathbf{p}_3 = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} \in V(5)$$

が取れることが分かります.⁴

以上から, それぞれの固有値に対する長さが 1 の固有ベクトル $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}^3$ を並べてできる行列を,

$$P = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix} \quad (3)$$

とすると, P は直交行列で,

$$P^{-1}AP = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix}$$

となることが分かります. ここで, 固有値の順番や直交行列 P の取り方は一通りではありませんから, 皆さんの求めた行列 P が, 必ずしも (3) 式のような行列でなくとも, それぞれの列ベクトルがそれぞれの固有値に対応した長さが 1 の固有ベクトルになっていれば, もちろん構いません.

2 対称行列とは

第 9 回の問 1 のところでは, 「行列の対角化の問題」を取り上げて, 「対角化の問題」を解決するための一般的な戦略について考えてみました. また, 第 9 回の問 2 のところでは, 「行列の対角化の問題」を「固有ベクトル空間分解」という特定の座標軸の取り方によらない形で定式化できることを見ました. また, 零行列ではないベキ零行列の場合には「対角化の問題」を解決することができないことにも触れました. このように, 一般には, 「行列の対角化の問題」がいつでも解決するとは限らないのですが, 7 節で見ると, 行列 A が対称行列の場合には, いつでも「対角化の問題」が解決することが分かります.⁵ さらに, 対角化を実現する正則行列 P として直交行列が取れることも分かります. 皆さんに, こうした「直交行列による対称行列の対角化」の具体的な例に触れてもらおうと思って, 問 1 を出題してみました. ただし, 単純に対称行列 A をひとつ与えて, その行列を対角化するというような形より, こうした状況が自然に現われてくるような形で出題した方が, 皆さんにとって, 数学全体を有機的に理解する助けになるのではないかと考えて, 問 1 のように, 臨界点の近くでの多変数関数の様子を調べるという形で出題してみました.

そこで, 「直交行列による対称行列の対角化」という問題に入る前に, ここでは, 「対称行列とは何か」ということを, 「線型空間」という立場から見直してみることにします. 皆さ

⁴もちろん, 長さが 1 の固有ベクトルとして, \mathbf{p}_3 の代わりに, $-\mathbf{p}_3$ というベクトルを考えても構いません.

⁵対称行列とは, ${}^tA = A$ となる行列のことです.

ん良くご存じのように、 \mathbb{R}^n というユークリッド空間上では、ベクトル同士の内積を考えることができます。すなわち、 \mathbb{R}^n のベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、ベクトル \mathbf{u} の長さ $\|\mathbf{u}\|$ や、二つのベクトル \mathbf{u}, \mathbf{v} がなす角度を考えることができます。以下で見るように、「直交行列による対称行列の対角化」ということの意味をより良く理解するためには、 \mathbb{R}^n を、単に、「足し算」や「スカラー倍」のできる「線型空間」であると考えてではなく、さらに「内積」という構造も定まった「内積を持つ線型空間」であると考えて、「基底」や「線型写像」や「直和分解」などといった概念を「内積」との関係をもとに見直してみることが重要なポイントになります。

いま、 \mathbb{R}^n の二つのベクトル

$$\mathbf{u} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \mathbf{v} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

に対して、 \mathbf{u} と \mathbf{v} の間の内積を、

————— \mathbb{R}^n 上の標準的なユークリッド内積 —————

$$\langle \mathbf{u}, \mathbf{v} \rangle = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \quad (4)$$

という記号を用いて表わすことにします。⁶ このとき、内積 $\langle \mathbf{u}, \mathbf{v} \rangle$ は、行列の積を用いて、

$$\begin{aligned} \langle \mathbf{u}, \mathbf{v} \rangle &= \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \\ &= {}^t \mathbf{u} \cdot \mathbf{v} \end{aligned} \quad (5)$$

というように表わすことができることに注意します。

そこで、いま、対称行列とは限らない n 行 n 列の行列 A が、勝手にひとつ与えられているとします。このとき、 \mathbb{R}^n の二つのベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、 $A\mathbf{u}$ と \mathbf{v} の間の内積 $\langle A\mathbf{u}, \mathbf{v} \rangle$ を考えてみると、 $\langle A\mathbf{u}, \mathbf{v} \rangle$ は、

$$\begin{aligned} \langle A\mathbf{u}, \mathbf{v} \rangle &= {}^t(A\mathbf{u}) \cdot \mathbf{v} \\ &= {}^t \mathbf{u} {}^t A \cdot \mathbf{v} \\ &= {}^t \mathbf{u} \cdot {}^t A \mathbf{v} \\ &= \langle \mathbf{u}, {}^t A \mathbf{v} \rangle \end{aligned}$$

というように書き直せることが分かります。よって、勝手なベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、

————— 転置行列を特徴付ける式 —————

$$\langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, {}^t A \mathbf{v} \rangle \quad (6)$$

となることが分かります。

⁶ 「内積とは何か」ということや (4) 式の導出については、11 節を参照して下さい。

そこで、次に、与えられた行列 A に対して、

$$\langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, B\mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n \quad (7)$$

となるような行列 B が、どれだけ存在するのかということを考えてみることにします。すなわち、(6) 式より、

$$B = {}^tA$$

とすれば、(7) 式が成り立つことが分かりますが、(7) 式を満たすような行列 B は tA の他にも存在するのかどうかということを考えてみることにします。そこで、いま、(7) 式を満たすような行列 B が見つかったと仮定してみます。このとき、(7) 式の両辺から (6) 式の両辺を引き算してみると、

$$\begin{aligned} 0 &= \langle \mathbf{u}, B\mathbf{v} \rangle - \langle \mathbf{u}, {}^tA\mathbf{v} \rangle \\ &= \langle \mathbf{u}, B\mathbf{v} - {}^tA\mathbf{v} \rangle \\ &= \langle \mathbf{u}, (B - {}^tA)\mathbf{v} \rangle \end{aligned}$$

となることが分かりますから、

$$\langle \mathbf{u}, (B - {}^tA)\mathbf{v} \rangle = 0 \quad (8)$$

となることが分かります。すると、(8) 式が、勝手なベクトル $\mathbf{u} \in \mathbb{R}^n$ に対して成り立つのですから、

$$(B - {}^tA)\mathbf{v} = \mathbf{0} \quad (9)$$

でなければならぬことが分かります。⁷ さらに、(9) 式が、勝手なベクトル $\mathbf{v} \in \mathbb{R}^n$ に対して成り立つのですから、

$$B - {}^tA = \mathbf{0}$$

でなければならぬことが分かります。⁸ したがって、

$$B = {}^tA$$

⁷例えば、 $\mathbf{u} \in \mathbb{R}^n$ として、

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

というベクトルを取ってみると、 $(B - {}^tA)\mathbf{v} \in \mathbb{R}^n$ というベクトルの各成分が 0 でなければならぬことが分かります。あるいは、(8) 式において、 \mathbf{u} として、 $\mathbf{u} = (B - {}^tA)\mathbf{v}$ を取ることで、

$$0 = \|(B - {}^tA)\mathbf{v}\|^2$$

となることが分かりますが、このことから (9) 式を結論することもできます。

⁸例えば、 $\mathbf{v} \in \mathbb{R}^n$ として、

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

というベクトルを取ってみると、 $(B - {}^tA)$ という行列の各列ベクトルが零ベクトル $\mathbf{0} \in \mathbb{R}^n$ でなければならぬことが分かります。

でなければならないことが分かります.

以上から, 与えられた行列 A に対して, (7) 式を満たすような行列 B は,

$$B = {}^tA$$

しか存在しないということが分かりました. すなわち, 行列 A に対して, その転置行列 tA は, (6) 式を満たすような行列として一意的に特徴付けられるということが分かりました. いま, 行列 A が対称行列であるとは,

対称行列の定義式

$${}^tA = A$$

となるということですから, 上の転置行列の特徴付けと合わせて考えると,

対称行列の特徴付け

$$A \text{ が対称行列} \iff \langle Au, v \rangle = \langle u, Av \rangle, \quad \forall u, v \in \mathbb{R}^n \quad (10)$$

という形で対称行列を特徴付けることができることが分かります. この (10) 式が, 対称行列を考えるということの意味を, 「内積を持つ線型空間」という視点から見直したものであるということになります.

次に, 対称行列 A を特徴付ける式である (10) 式から, 対称行列 A の固有値や固有ベクトル空間に関して, どのようなことが分かるのかということを考えてみたいのですが, その前に, 問 1 の例について, 少し見返してみることにします. 問 1 では,

$$A = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix} \quad (11)$$

という対称行列を考えました. 問 1 の解答で見たように, このとき, 行列 A の特性多項式 $\varphi_A(t)$ は,

$$\varphi_A(t) = (t+1)(t-2)(t-5) \quad (12)$$

となることが分かりますから, 特に, 行列 A のすべての固有値は実数であるということが分かります. 問 1 の解答を見返してみると, このように対称行列 A の固有値がすべて実数であるお陰で, それぞれの固有値に対して, 複素数の世界にまで考察を広げなくとも, \mathbb{R}^3 という「 \mathbb{R} 上の線型空間」の中で固有ベクトルを見つけることができ, そうした固有ベクトルを用いて, 対称行列 A が (実数行列である) 直交行列 P により対角化されていることが分かります.

一般に, n 行 n 列の行列 A と n 行 n 列の正則行列 P に対して, $P^{-1}AP$ という行列の特性多項式 $\varphi_{P^{-1}AP}(t)$ は,

$$\begin{aligned} \varphi_{P^{-1}AP}(t) &= \det(tI - P^{-1}AP) \\ &= \det(tP^{-1}IP - P^{-1}AP) \\ &= \det(P^{-1}(tI - A)P) \end{aligned}$$

$$\begin{aligned}
&= \det(P^{-1}) \cdot \det(tI - A) \cdot \det P \\
&= (\det P)^{-1} \cdot \det(tI - A) \cdot \det P \\
&= \det(tI - A)
\end{aligned}$$

というように書き直せることが分かりますから、

————— 特性多項式の共役不変性 —————

$$\varphi_{P^{-1}AP}(t) = \varphi_A(t) \tag{13}$$

となることが分かります。したがって、例えば、3 行 3 列の行列 A が、3 行 3 列の正則行列 P を用いて、

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \tag{14}$$

というように対角化されると仮定すると、 A の特性多項式 $\varphi_A(t)$ は、

$$\begin{aligned}
\varphi_A(t) &= \varphi_{P^{-1}AP}(t) \\
&= \begin{vmatrix} t - \lambda_1 & 0 & 0 \\ 0 & t - \lambda_2 & 0 \\ 0 & 0 & t - \lambda_3 \end{vmatrix} \\
&= (t - \lambda_1)(t - \lambda_2)(t - \lambda_3)
\end{aligned} \tag{15}$$

となることが分かります。特に、 A も P も実数を成分にもつ行列であるとする、 $P^{-1}AP$ という行列も実数を成分にもつ行列になりますから、(14) 式より、 $\lambda_1, \lambda_2, \lambda_3$ は、すべて実数でなければならないことが分かります。よって、(15) 式と合わせて考えると、実数を成分にもつ正方行列 A が、実数を成分にもつ正則行列 P を用いて対角化されるとすると、行列 A のすべての固有値は実数でなければならないということが分かります。

我々は、「(実数を成分にもつ) 直交行列による (実数を成分にもつ) 対称行列の対角化」という問題を考えたいわけですが、上で見たことから、こうした問題が意味を持つためには、そもそも、対称行列のすべての固有値は実数であるということが保証されていなければ、お話にならないということが分かります。問 1 の例では、(12) 式のように、直接、行列 A のすべての固有値が実数になるということを確認したわけですが、実は、一般に、

————— 対称行列の持つ基本的な性質 (その 1) —————

$$A \text{ は対称行列} \implies A \text{ のすべての固有値は実数} \tag{16}$$

となることが分かります。この主張を確認するためには、実数を成分にもつ行列の場合でも、一般には、固有値は複素数になりうるということに注意して、すべての固有値に対して、自由に固有ベクトルを考えることができるように、一度、複素数の世界に対象を広げて考察を進める方がよいのですが、議論が混乱するといけないので、この確認作業は後にまわして、取り合えず、(16) 式の主張を認めることにします。すると、対称行列 A のすべての固有値は実数となりますから、それぞれの固有値 $\lambda \in \mathbb{R}$ に対して、

対称行列 A の (固有値 λ に対応した) 固有ベクトル空間

$$V(\lambda) = \{\mathbf{u} \in \mathbb{R}^n \mid A\mathbf{u} = \lambda\mathbf{u}\}$$

という固有ベクトル空間が \mathbb{R}^n という「 \mathbb{R} 上の線型空間」の中で意味を持つこととなります。

そこで、いま、対称行列 A の異なる固有値が二つ以上存在すると仮定して、二つの異なる固有値 $\lambda, \mu \in \mathbb{R}$, ($\lambda \neq \mu$) に対して、それぞれの固有値に対応した固有ベクトル

$$\mathbf{u} \in V(\lambda), \mathbf{v} \in V(\mu)$$

を、勝手にひとつずつ取ってきたとします。このとき、

$$\begin{aligned}\langle A\mathbf{u}, \mathbf{v} \rangle &= \langle \lambda\mathbf{u}, \mathbf{v} \rangle \\ &= \lambda \cdot \langle \mathbf{u}, \mathbf{v} \rangle\end{aligned}\tag{17}$$

$$\begin{aligned}\langle \mathbf{u}, A\mathbf{v} \rangle &= \langle \mathbf{u}, \mu\mathbf{v} \rangle \\ &= \mu \cdot \langle \mathbf{u}, \mathbf{v} \rangle\end{aligned}\tag{18}$$

となることが分かりますから、(10) 式, (17) 式, (18) 式から、

$$(\lambda - \mu) \cdot \langle \mathbf{u}, \mathbf{v} \rangle = 0\tag{19}$$

となることが分かります。いま、

$$\lambda \neq \mu$$

と仮定していましたが、(19) 式から、結局、

$$\langle \mathbf{u}, \mathbf{v} \rangle = 0$$

となることが分かりました。

以上から、対称行列の異なる固有値に対応した固有ベクトルは互いに直交するということが、すなわち、

$$\mathbf{u} \in V(\lambda), \mathbf{v} \in V(\mu) \implies \langle \mathbf{u}, \mathbf{v} \rangle = 0\tag{20}$$

となることが分かりました。ここで、(20) 式は、固有値 λ に対応した固有ベクトル空間 $V(\lambda)$ と固有値 μ に対応した固有ベクトル空間 $V(\mu)$ が、 \mathbb{R}^n の中で互いに直交する線型部分空間になっているということを意味していますが、このことを、「直交する」という記号「 \perp 」を用いて、

対称行列の持つ基本的な性質 (その 2)

$$\lambda \neq \mu \implies V(\lambda) \perp V(\mu)\tag{21}$$

このように表わしたします。

そこで、次に、(16) 式の主張を確かめるために、上で行なった考察を、複素数の世界に拡張して考えてみることにします。そのためには、 \mathbb{R}^n 上のユークリッド内積の代わりに、

$$\mathbf{u} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}, \mathbf{v} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$$

に対して、

————— \mathbb{C}^n 上の標準的なエルミート内積 —————

$$(\mathbf{u}, \mathbf{v}) = z_1 \bar{w}_1 + z_2 \bar{w}_2 + \cdots + z_n \bar{w}_n$$

という式によって定まる \mathbb{C}^n 上のエルミート内積を考えればよいということになります。⁹

そこで、 n 行 n 列の複素行列 A が勝手にひとつ与えられているとして、 \mathbb{R}^n 上のユークリッド内積 $\langle \cdot, \cdot \rangle$ を \mathbb{C}^n 上のエルミート内積 (\cdot, \cdot) に置き換えて、上の議論を繰り返してみると、勝手なベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ に対して、

————— 随伴行列を特徴付ける式 —————

$$(A\mathbf{u}, \mathbf{v}) = (\mathbf{u}, {}^t\bar{A}\mathbf{v}) \quad (22)$$

となることが分かります。一般に、行列 A に対して、 ${}^t\bar{A}$ という行列を A の随伴行列 (adjoint matrix) と呼びますが、前と同様にして、与えられた行列 A に対して、

$$(A\mathbf{u}, \mathbf{v}) = (\mathbf{u}, B\mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{C}^n \quad (23)$$

となるような行列 B は、

$$B = {}^t\bar{A}$$

しか存在しないということが分かります。すなわち、行列 A に対して、その随伴行列 ${}^t\bar{A}$ は、(22) 式を満たすような行列として一意に特徴付けられるということが分かります。一般に、

————— エルミート行列の定義式 —————

$$A = {}^t\bar{A}$$

となるような行列 A をエルミート行列と呼びますが、上の随伴行列の特徴付けと合わせて考えると、

————— エルミート行列の特徴付け —————

$$A \text{ がエルミート行列} \iff (A\mathbf{u}, \mathbf{v}) = (\mathbf{u}, A\mathbf{v}), \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{C}^n \quad (24)$$

⁹ここで、不幸なことに、 z と w のうちどちらに複素共役「 $\bar{\cdot}$ 」を付けてエルミート内積を定めるのかという慣習が、数学と物理とは逆になっています。この演習では、一応、数学の慣習に従うことにしましたが、物理の慣習の方に馴染みがあるという方は、そちらの慣習に読み替えて下さい。また、後の議論でいらない混乱をするといけないので、 \mathbb{R} 上のユークリッド内積に対しては $\langle \cdot, \cdot \rangle$ という記号を用い、 \mathbb{C} 上のエルミート内積に対しては (\cdot, \cdot) という記号を用いるというように、二種類の内積を区別して表わすことにしました。

という形でエルミート行列を特徴付けることができることが分かります。この (24) 式が、エルミート行列を考えるということの意味を、「(エルミート)内積を持つ(\mathbb{C} 上の)線型空間」という立場から見直したものであるということになります。

以上の準備のもとで、(24) 式という関係式から、エルミート行列 A の固有値や固有ベクトル空間に関して、どのようなことが分かるのかということを考えてみることにします。そこで、いま、エルミート行列 A の固有値 $\lambda \in \mathbb{C}$ に対して、 λ に対応するゼロでない固有ベクトル $0 \neq \mathbf{u} \in \mathbb{C}^n$ を、勝手にひとつ取ってくることにします。このとき、

$$\begin{aligned} (A\mathbf{u}, \mathbf{u}) &= (\lambda\mathbf{u}, \mathbf{u}) \\ &= \lambda \cdot (\mathbf{u}, \mathbf{u}) \end{aligned} \tag{25}$$

$$\begin{aligned} (\mathbf{u}, A\mathbf{u}) &= (\mathbf{u}, \lambda\mathbf{u}) \\ &= \bar{\lambda} \cdot (\mathbf{u}, \mathbf{u}) \end{aligned} \tag{26}$$

となることが分かりますが、(24) 式において、 $\mathbf{v} = \mathbf{u}$ としてみると、(25) 式、(26) 式から、

$$(\lambda - \bar{\lambda}) \cdot (\mathbf{u}, \mathbf{u}) = 0 \tag{27}$$

となることが分かります。いま、 $\mathbf{u} \neq 0$ と仮定していましたが、

$$(\mathbf{u}, \mathbf{u}) = \|\mathbf{u}\|^2 \neq 0$$

となることに注意すると、結局、(27) 式から、

$$\lambda = \bar{\lambda}$$

となることが分かります。¹⁰ すなわち、 $\lambda \in \mathbb{R}$ となることが分かりました。

以上より、

エルミート行列の持つ基本的な性質 (その 1)

$$A \text{ がエルミート行列} \implies A \text{ のすべての固有値は実数} \tag{28}$$

となることが分かりました。「実数を成分に持つ対称行列」は「実数を成分に持つエルミート行列」とみなすこともできますから、(28) 式の主張から、(16) 式の主張が従うことが分かります。

そこで、エルミート行列 A の固有値 $\lambda \in \mathbb{R}$ に対して、 λ に対応する固有ベクトル空間を、

エルミート行列 A の (固有値 λ に対応した) 固有ベクトル空間

$$V(\lambda)_{\mathbb{C}} = \{ \mathbf{u} \in \mathbb{C}^n \mid A\mathbf{u} = \lambda\mathbf{u} \}$$

という記号を用いて表わすことにします。¹¹ すると、固有値がすべて実数であるということに注意して、対称行列のときと同様に議論すると、(24) 式から、エルミート行列の異なる固有値に対応する固有ベクトル空間は互いに直交するということが、すなわち、

¹⁰ここで、ベクトル $\mathbf{u} \in \mathbb{C}^n$ の「長さ」を $\|\mathbf{u}\|$ という記号を用いて表わしました。

¹¹ここで、 \mathbb{C}^n の中で固有ベクトル空間を考えているということを強調するために、「 \mathbb{C} 」という添え字を付けて表わすことにしました。

$$\lambda \neq \mu \implies V(\lambda)_{\mathbb{C}} \perp V(\mu)_{\mathbb{C}} \quad (29)$$

となることが分かります。¹²

3 直交行列とは

さて、2節では、対称行列を考えるということの意味を「内積を持つ線型空間」という立場から見直してみることで、対称行列の持つ基本的な性質が理解できることを見ました。そこで、次に、直交行列とは何かということを考えてみることにします。

一般に、

直交行列の定義式

$${}^t P P = I \quad (30)$$

となるような n 行 n 列の正方行列 P を直交行列と呼びます。特に、(30) 式より、直交行列 P は正則行列であり、その逆行列は、

$$P^{-1} = {}^t P \quad (31)$$

という簡単な形で与えられることが分かります。このように逆行列を簡単に求めることができるということが、実際の計算にあたっては、直交行列を用いて対角化をするということの直接のご利益になります。ここでは、こうした直交行列を考えるということの意味を「内積を持つ線型空間」という立場から見直してみたいのですが、一般の場合でも話の本質は変わりませんから、話を具体的にするために、以下では、3行3列の行列の場合に説明することにします。

いま、3行3列の行列 P に対して、

行列 P の列ベクトルを $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}^3$ と表わす

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

というように、行列 P の列ベクトルを $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3 \in \mathbb{R}^3$ と表わすことにします。このとき、 ${}^t P P$ という行列を求めてみると、

$$\begin{aligned} {}^t P P &= \begin{pmatrix} {}^t \mathbf{p}_1 \\ {}^t \mathbf{p}_2 \\ {}^t \mathbf{p}_3 \end{pmatrix} \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \mathbf{p}_3 \end{pmatrix} \\ &= \begin{pmatrix} {}^t \mathbf{p}_1 \cdot \mathbf{p}_1 & {}^t \mathbf{p}_1 \cdot \mathbf{p}_2 & {}^t \mathbf{p}_1 \cdot \mathbf{p}_3 \\ {}^t \mathbf{p}_2 \cdot \mathbf{p}_1 & {}^t \mathbf{p}_2 \cdot \mathbf{p}_2 & {}^t \mathbf{p}_2 \cdot \mathbf{p}_3 \\ {}^t \mathbf{p}_3 \cdot \mathbf{p}_1 & {}^t \mathbf{p}_3 \cdot \mathbf{p}_2 & {}^t \mathbf{p}_3 \cdot \mathbf{p}_3 \end{pmatrix} \end{aligned}$$

¹²皆さん、確かめてみて下さい。

となることが分かります. ここで, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ として, \mathbb{R}^3 上の標準的なユークリッド内積は, 行列の掛け算を用いて,

$$\langle \mathbf{u}, \mathbf{v} \rangle = {}^t \mathbf{u} \cdot \mathbf{v}$$

というように表わされることに注意して, それぞれの行列成分を内積を用いて表わすことにすると, 結局, ${}^t P P$ という行列は,

$${}^t P P = \begin{pmatrix} \langle \mathbf{p}_1, \mathbf{p}_1 \rangle & \langle \mathbf{p}_1, \mathbf{p}_2 \rangle & \langle \mathbf{p}_1, \mathbf{p}_3 \rangle \\ \langle \mathbf{p}_2, \mathbf{p}_1 \rangle & \langle \mathbf{p}_2, \mathbf{p}_2 \rangle & \langle \mathbf{p}_2, \mathbf{p}_3 \rangle \\ \langle \mathbf{p}_3, \mathbf{p}_1 \rangle & \langle \mathbf{p}_3, \mathbf{p}_2 \rangle & \langle \mathbf{p}_3, \mathbf{p}_3 \rangle \end{pmatrix} \quad (32)$$

というように表わされること分かります. よって, (32) 式から, (30) 式という直交行列の条件は,

——— 行列 P が直交行列であるための条件 (言い換え) ———

$$\langle \mathbf{p}_i, \mathbf{p}_j \rangle = \begin{cases} 1, & i = j \text{ のとき} \\ 0, & i \neq j \text{ のとき} \end{cases}, \quad (i, j = 1, 2, 3) \quad (33)$$

という条件を「行列の積」という形でコンパクトに表現したものであることが分かります. 一般に, \mathbb{R}^n のいくつかのベクトル $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L \in \mathbb{R}^n$ が,

——— 正規直交系の定義式 ———

$$\langle \mathbf{p}_i, \mathbf{p}_j \rangle = \begin{cases} 1, & i = j \text{ のとき} \\ 0, & i \neq j \text{ のとき} \end{cases}, \quad (i, j = 1, 2, \dots, L) \quad (34)$$

という条件を満たすときに, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L\}$ を正規直交系と呼びます. そこで, いま, \mathbb{R}^n の正規直交系 $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L\}$ が, 勝手にひとつ与えられているとして, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L\}$ が生成する \mathbb{R}^n の線型部分空間を,

$$\begin{aligned} W &= \mathbb{R}\mathbf{p}_1 + \mathbb{R}\mathbf{p}_2 + \dots + \mathbb{R}\mathbf{p}_L \\ &= \{a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \dots + a_L\mathbf{p}_L \in \mathbb{R}^n \mid a_1, a_2, \dots, a_L \in \mathbb{R}\} \end{aligned}$$

と表わすことにします. すると, このとき, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L\}$ は W の基底になることが, 次のようにして分かります.

そのためには,

—— $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L\}$ が W の基底となるための条件 ——

(イ) 勝手なベクトル $\mathbf{u} \in W$ に対して,

$$\mathbf{u} = a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \dots + a_L\mathbf{p}_L$$

となるような実数 $a_1, a_2, \dots, a_L \in \mathbb{R}$ が存在する.

(ロ) $a_1, a_2, \dots, a_L \in \mathbb{R}$ として,

$$\mathbf{0} = a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \dots + a_L\mathbf{p}_L \implies a_1 = a_2 = \dots = a_L = 0$$

となる.

という二つの条件が満たされることが確かめられれば良いということになります. このうち, (イ) という条件については, W の定義から, 自動的に満たされることが分かりますから, 後は, (ロ) という条件が満たされることが確かめられればよいということになります.

そこで, いま, $a_1, a_2, \dots, a_L \in \mathbb{R}$ として,

$$\mathbf{0} = a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \dots + a_L\mathbf{p}_L \quad (35)$$

であると仮定してみます. このとき, (34) 式に注意して, $i = 1, 2, \dots, L$ に対して, \mathbf{p}_i と (35) 式の両辺との内積を考えると,

$$\begin{aligned} 0 &= \langle \mathbf{p}_i, a_1\mathbf{p}_1 + a_2\mathbf{p}_2 + \dots + a_L\mathbf{p}_L \rangle \\ &= \langle \mathbf{p}_i, a_1\mathbf{p}_1 \rangle + \langle \mathbf{p}_i, a_2\mathbf{p}_2 \rangle + \dots + \langle \mathbf{p}_i, a_L\mathbf{p}_L \rangle \\ &= a_1 \cdot \langle \mathbf{p}_i, \mathbf{p}_1 \rangle + a_2 \cdot \langle \mathbf{p}_i, \mathbf{p}_2 \rangle + \dots + a_L \cdot \langle \mathbf{p}_i, \mathbf{p}_L \rangle \\ &= a_i \end{aligned} \quad ((34) \text{ 式から})$$

となることが分かりますから,

$$a_1 = a_2 = \dots = a_L = 0$$

となることが分かります. よって, (ロ) という条件も満たされることが分かります.

以上の議論から, (イ), (ロ) という二つの条件が満たされることが分かりましたから, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_L\}$ は W の基底になることが分かりました. 特に, $L = n$ のときには,

$$W = \mathbb{R}^n$$

となることが分かりますから, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ は \mathbb{R}^n の基底になることが分かります. このとき, $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ を \mathbb{R}^n の正規直交基底と呼びます. これらの言葉を用いると, (32) 式より, P が直交行列となる条件を,

—— 直交行列を考えることの意味 (その1) ——

P が直交行列 $\iff P$ の列ベクトル $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ が \mathbb{R}^n の正規直交基底となる.

というように言い換えることができることが分かります.

こうして,

$${}^tPP = I$$

という行列 P が直交行列となる条件を「基底」という立場から見直せることが分かりましたが、次に、同じ条件を「線型写像」という立場から見直してみることになります。

いま、 n 行 n 列の行列 P が与えられているとして、 $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、 $P\mathbf{u} \in \mathbb{R}^n$ というベクトルと $P\mathbf{v} \in \mathbb{R}^n$ というベクトルの間の内積 $\langle P\mathbf{u}, P\mathbf{v} \rangle$ を考えてみます。そこで、前と同様に、 $\langle P\mathbf{u}, P\mathbf{v} \rangle$ という値を「行列の積」を用いて表わしてみると、

$$\begin{aligned} \langle P\mathbf{u}, P\mathbf{v} \rangle &= {}^t(P\mathbf{u}) \cdot P\mathbf{v} \\ &= {}^t\mathbf{u} {}^tP \cdot P\mathbf{v} \\ &= {}^t\mathbf{u} \cdot {}^tPP\mathbf{v} \\ &= \langle \mathbf{u}, {}^tPP\mathbf{v} \rangle \end{aligned} \tag{36}$$

となることが分かります。特に、

$${}^tPP = I$$

であるとすると、(36) 式から、

$$\begin{aligned} \langle P\mathbf{u}, P\mathbf{v} \rangle &= \langle \mathbf{u}, {}^tPP\mathbf{v} \rangle \\ &= \langle \mathbf{u}, I\mathbf{v} \rangle \\ &= \langle \mathbf{u}, \mathbf{v} \rangle \end{aligned}$$

となることが分かります。

逆に、勝手なベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、

$$\langle P\mathbf{u}, P\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle \tag{37}$$

という式が成り立っていると仮定してみます。このとき、(36) 式を用いて、

$$\begin{aligned} \langle P\mathbf{u}, P\mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle &= \langle \mathbf{u}, {}^tPP\mathbf{v} \rangle - \langle \mathbf{u}, \mathbf{v} \rangle && \text{((36) 式から)} \\ &= \langle \mathbf{u}, {}^tPP\mathbf{v} - \mathbf{v} \rangle \\ &= \langle \mathbf{u}, ({}^tPP - I)\mathbf{v} \rangle \end{aligned}$$

というように書き直してみると、(37) 式は、

$$\langle \mathbf{u}, ({}^tPP - I)\mathbf{v} \rangle = 0 \tag{38}$$

という形に書き直せることが分かります。すると、(38) 式が、勝手なベクトル $\mathbf{u} \in \mathbb{R}^n$ に対して成り立つのですから、

$$({}^tPP - I)\mathbf{v} = \mathbf{0} \tag{39}$$

でなければならないことが分かります。¹³ さらに、(39) 式が、勝手なベクトル $\mathbf{v} \in \mathbb{R}^n$ に対して成り立つのですから、

$${}^tPP - I = \mathbf{0}$$

¹³例えば、 $\mathbf{u} \in \mathbb{R}^n$ として、

$$\mathbf{u} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

でなければならないことが分かります.¹⁴ よって,

$${}^tPP = I$$

でなければならないことが分かりました.

以上から、「線型写像」という視点から見直すと、直交行列の条件は、

直交行列を考えることの意味 (その2)

$${}^tPP = I \iff \langle P\mathbf{u}, P\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n$$

というように言い換えられることが分かりました. すなわち、直交行列とは、内積を保つような線型写像を表わす行列であることが分かりました.

一般に、勝手な自然数 $n \in \mathbb{N}$ に対して、 \mathbb{R}^n の内積を保つような行列全体の集合を、

n 次の直交群

$$O(n) = \{P \in M_n(\mathbb{R}) \mid {}^tPP = I\}$$

という記号で表わして、 n 次の直交群 (orthogonal group) と呼んだりします.¹⁵ いま、 $P, Q \in O(n)$ であるとすると、

$$\begin{aligned} {}^t(PQ)(PQ) &= {}^tQ({}^tPP)Q \\ &= {}^tQQ && (P \in O(n) \text{ より}) \\ &= I && (Q \in O(n) \text{ より}) \end{aligned}$$

となることが分かりますから、

$$PQ \in O(n)$$

となることが分かります. また、この節の最初で注意したように、

$${}^tPP = I \tag{40}$$

となることから、 P は正則行列で、

$$P^{-1} = {}^tP$$

というベクトルを取ってみると、 $({}^tPP - I)\mathbf{v} \in \mathbb{R}^n$ というベクトルの各成分が 0 でなければならないことが分かります. あるいは、(38) 式において、 \mathbf{u} として、 $\mathbf{u} = ({}^tPP - I)\mathbf{v}$ を取ることで、

$$0 = \|({}^tPP - I)\mathbf{v}\|^2$$

となることが分かりますが、このことから (39) 式を結論することもできます.

¹⁴例えば、 $\mathbf{v} \in \mathbb{R}^n$ として、

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$$

というベクトルを取ってみると、 $({}^tPP - I)$ という行列の各列ベクトルが零ベクトル $\mathbf{0} \in \mathbb{R}^n$ でなければならないことが分かります.

¹⁵ここで、実数 を成分にもつ n 行 n 列の行列 (matrix) 全体の集合を $M_n(\mathbb{R})$ という記号を用いて表わしました.

となることが分かります。よって、

$$\begin{aligned} {}^t(P^{-1})P^{-1} &= {}^t({}^tP)P^{-1} \\ &= PP^{-1} \\ &= I \end{aligned}$$

となることが分かりますから、

$$P^{-1} \in O(n)$$

となることも分かります。第8回の問1のところでも触れたように、数学では、その集合の中で「掛け算」や「逆元」を考えることができるような集合のことを、一般に、群 (group) と呼びます。上で確かめたように、直交行列全体のなす集合 $O(n)$ はこのような集合になることが分かりますから、 $O(n)$ には直交群という名前が付いています。

また、一般に、 n 行 n 列の行列 P, Q に対して、

$$\begin{aligned} \det({}^tP) &= \det P \\ \det(PQ) &= \det P \cdot \det Q \end{aligned}$$

となることに注意して、(40) 式の両辺に現われる行列の行列式を考えてみると、

$$\begin{aligned} \det({}^tPP) &= \det({}^tP) \cdot \det P \\ &= \det P \cdot \det P \\ &= (\det P)^2 \\ \det I &= 1 \end{aligned}$$

となることが分かりますから、

$$(\det P)^2 = 1$$

となることが分かります。よって、

直交行列の持つ基本的な性質

$$P \text{ が直交行列} \implies \det P = \pm 1$$

となることが分かります。そこで、

$$\det P = 1$$

となるような直交行列全体の集合を、

n 次の回転群

$$SO(n) = \{P \in O(n) \mid \det P = 1\}$$

という記号で表わすことにすると、 $SO(n)$ も $O(n)$ の中だけで「掛け算」や「逆元」を考えることができるような集合になることが分かります。¹⁶ 特に、 $n = 3$ のときには、 $SO(3)$ のそれぞれの元は、 \mathbb{R}^3 の原点を通る直線を軸とする回転を表わすような行列であることを

¹⁶皆さん、確かめてみて下さい。

確かめることができるので、一般に、 $SO(n)$ のことを n 次の回転群とか、 n 次の特殊直交群 (special orthogonal group) とか呼んだりします。¹⁷

以上の議論をまとめると、「内積を持つ線型空間」という視点から見直すことで、直交行列の条件を、

直交行列の特徴付け

$$P \text{ が直交行列となる. } \iff P \text{ の列ベクトルが } \mathbb{R}^n \text{ の正規直交基底となる. } \quad (41)$$

$$\iff P \text{ は } \mathbb{R}^n \text{ の内積を保つ線型写像を定める. } \quad (42)$$

というように言い換えられることが分かりました。いま、

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

と表わすことにすると、 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ は \mathbb{R}^n の正規直交基底になることが分かります。このとき、 $i = 1, 2, \dots, n$ として、行列 P の第 i 列目の列ベクトル $\mathbf{p}_i \in \mathbb{R}^n$ は、

$$\mathbf{p}_i = P\mathbf{e}_i$$

というように表わせることに注意すると、 $i, j = 1, 2, \dots, n$ として、 P の列ベクトルたちの間の内積は、

$$\langle \mathbf{p}_i, \mathbf{p}_j \rangle = \langle P\mathbf{e}_i, P\mathbf{e}_j \rangle$$

というように表わせることが分かります。このことに注意すると、

$$\langle \mathbf{p}_i, \mathbf{p}_j \rangle = \begin{cases} 1, & i = j \text{ のとき} \\ 0, & i \neq j \text{ のとき} \end{cases}, \quad (i, j = 1, 2, \dots, n) \quad (43)$$

という直交行列に対する (41) 式の特徴付けと、

$$\langle P\mathbf{u}, P\mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^3 \quad (44)$$

という (42) 式の特徴付けとが、本質的に同じ内容を表わしているということが分かります。興味のある方は、実際に、二つの条件が同値であることをきちんと証明してみてください。

ここでは、 \mathbb{R}^n という「 \mathbb{R} 上の線型空間」をもとにして、直交行列について説明しましたが、 \mathbb{R}^n の代わりに \mathbb{C}^n という「 \mathbb{C} 上の線型空間」を考えて、 \mathbb{R}^n 上のユークリッド内積 $\langle \cdot, \cdot \rangle$ の代わりに \mathbb{C}^n 上のエルミート内積 (\cdot, \cdot) を考えることで、 \mathbb{C}^n に対して、全く同様の議論をすることもできます。

一般に、

¹⁷「特殊直交群」という言葉の意味は「 $\det P = 1$ 」という「特殊な条件」の付いた「直交行列」からなる「群」ということです。

ユニタリー行列の定義式

$${}^t\bar{U}U = I \quad (45)$$

となるような n 行 n 列の複素行列 U をユニタリー行列と呼びます. このとき, \mathbb{C}^n を「(エルミート)内積を持つ (\mathbb{C} 上の)線型空間」であると考えて, \mathbb{R}^n の場合と同様な議論をしてみると, U がユニタリー行列となる条件を,

ユニタリー行列の特徴付け

U がユニタリー行列となる. $\iff U$ の列ベクトルが \mathbb{C}^n の正規直交基底となる.
 $\iff U$ は \mathbb{C}^n の内積を保つ線型写像を定める.

というように言い換えることができることが分かります. また, 勝手な自然数 $n \in \mathbb{N}$ に対して, n 行 n 列のユニタリー行列全体の集合を,

n 次のユニタリー群

$$U(n) = \{U \in M_n(\mathbb{C}) \mid {}^t\bar{U}U = I\}$$

という記号を用いて表わすことにすると, $O(n)$ の場合と同様にして, $U(n)$ は群になることが分かるので, $U(n)$ のことを n 次のユニタリー群 (unitary group) と呼んだりします. さらに,

$$\begin{aligned} \det({}^t\bar{U}U) &= \det({}^t\bar{U}) \cdot \det U \\ &= \det(\bar{U}) \cdot \det U \\ &= \overline{\det U} \cdot \det U \\ &= |\det U|^2 \end{aligned}$$

となることに注意して, (45) 式の両辺に現われる行列の行列式を考えると,

$$|\det U|^2 = 1$$

となることが分かります. よって,

ユニタリー行列の持つ基本的な性質

U がユニタリー行列 $\implies \det U$ は絶対値 1 の複素数になる.

となることが分かります. そこで, $\det U = 1$ となるユニタリー行列全体の集合を,

n 次の特殊ユニタリー群

$$SU(n) = \{U \in U(n) \mid \det U = 1\}$$

という記号を用いて表わすことにすると, やはり, $SU(n)$ も群になることが分かるので, $SU(n)$ のことを n 次の特殊ユニタリー群 (special unitary group) と呼んだりします.¹⁸

¹⁸ 「特殊ユニタリー群」という言葉の意味は「 $\det U = 1$ 」という「特殊な条件」の付いた「ユニタリー行列」からなる「群」ということです.

興味のある方は、この節で \mathbb{R}^n に対して行なった議論を、 \mathbb{C}^n に対して繰り返してみることで、これらの事実を確かめてみて下さい。

4 対称変換とは

さて、第9回の問1のところでは、

行列の対角化の問題

与えられた正方行列 A に対して、

$$P^{-1}AP = \Lambda \quad (46)$$

となるような対角行列 Λ と正則行列 P を見つけよ。

という「行列の対角化の問題」について触れ、(46) 式を、

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

$$\iff A \begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{pmatrix} = \begin{pmatrix} \mathbf{p}_1 & \cdots & \mathbf{p}_n \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$$\iff \begin{cases} A\mathbf{p}_1 = \lambda_1\mathbf{p}_1 \\ A\mathbf{p}_2 = \lambda_2\mathbf{p}_2 \\ \vdots \\ A\mathbf{p}_n = \lambda_n\mathbf{p}_n \end{cases}$$

というように書き直して考えてみることで、

「行列の対角化の問題」を解決する戦略

(i) 特性多項式 $\varphi_A(t) = \det(tI - A)$ を計算して、 $\varphi_A(\lambda) = 0$ の解 $\lambda \in \mathbb{C}$ をすべて求める。(\implies 行列 A の固有値が求まる.)

(ii) それぞれの固有値 $\lambda \in \mathbb{C}$ に対して、

$$(A - \lambda I)\mathbf{u} = \mathbf{0}$$

という連立一次方程式を解いて、

$$V(\lambda) = \{ \mathbf{u} \in \mathbb{C}^n \mid A\mathbf{u} = \lambda\mathbf{u} \}$$

を求める。(\implies 行列 A の固有ベクトルが求まる.)

(iii) $V(\lambda)$ たちの中から、適当にベクトル \mathbf{p}_i , ($i = 1, 2, \dots, n$) を取り出して、正則行列 P を作る。

という戦略を立てることができることを見ました。ただし、一般には、この戦略がいつも上

手くいくとは限らないということ, 例えば, $N \neq 0$ となるベキ零行列 N はことごとく対角化可能ではないということも注意しました. このように, 一般には, 「対角化の問題」がいつでも解決するわけではないので,

行列の標準形の問題

与えられた行列 A に対して,

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

となるような「見やすい形」の行列 Λ と正則行列 P を見つけよ.

という「行列の標準形の問題」を解決するためには, 「見やすい形」の行列 Λ を対角行列より少し一般化して考える必要があります.

ところが, 「対角化の問題」がいつでも解決することが理論的に保証される場合があります. その代表的な例が, 2 節で取り上げた「対称行列」や「エルミート行列」です. そこで見たように, これらの行列は,

対称行列の特徴付け

$$\langle Au, v \rangle = \langle u, Av \rangle, \quad \forall u, v \in \mathbb{R}^n, \quad (47)$$

あるいは,

エルミート行列の特徴付け

$$(Au, v) = (u, Av), \quad \forall u, v \in \mathbb{C}^n \quad (48)$$

という式によって特徴づけることができます. 2 節では, こうした「内積との相性の良さ」に注目することで,

対称行列やエルミート行列の基本的な性質

- (I) 固有値はすべて実数になる.
- (II) 相異なる固有値に対応する固有ベクトルは互いに直交する.

となることを見ました. 「対角化の問題」がいつでも解決できるということを理論的に保証する根拠も, やはり, (47) 式や (48) 式という「内積との相性の良さ」にあります. そこで, こうした「内積との相性の良さ」を用いて, 実際に「対角化の問題」が解決できることを確かめたいのですが, 議論がゴタゴタして何をやっているのが分からなくなってもいけませんから, 実際の確認作業は 6 節にまわすことにして, この節と, 5 節では, 6 節での確認作業において鍵となる事実について先に少し考えてみることにします. これまでの議論と同様に, \mathbb{R} 上の線型空間 \mathbb{R}^n を \mathbb{C} 上の線型空間 \mathbb{C}^n に置き換えることで, エルミート行列の場合でも全く同様の議論ができますから, 以下では, 対称行列の場合について説明することにします.

さて, 2 節で見たように, 対称行列の固有値や固有ベクトルに関する (I), (II) という性質

をより良く理解するためには,

$$A = {}^t A$$

という「数が並んだもの」としての対称行列の「姿」に注目するのではなく, (47) 式という「内積との相性の良さ」という性質に注目することが大切です. すなわち, 対称行列 A を「行列」と見るのではなく, 「線型写像」という視点から考察することが大切です. 6 節での確認作業においても事情は全く同様です. そこで, 「数が並んだもの」としての「行列」の「姿」に惑わされないように, 数ベクトル空間 \mathbb{R}^n から「座標軸」を消し去って, (\mathbb{R} 上の) 線型空間と思っただけの,

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

と表わすことにします. また, 対称行列 A を掛け算することにより定まる線型写像も, 単に,

$$f: V \rightarrow V$$

と表わすことにします. すると, (47) 式は,

$$\langle f(\mathbf{u}), \mathbf{v} \rangle = \langle \mathbf{u}, f(\mathbf{v}) \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in V$$

と表わすことができます. 一般に,

$$(V, \langle \cdot, \cdot \rangle_V)$$

を, \mathbb{R}^n のような「内積を持つ (\mathbb{R} 上の) 線型空間」として,¹⁹ V 上の線型写像 $f: V \rightarrow V$ が,

内積を持つ線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ 上の対称変換の定義式

$$\langle f(\mathbf{u}), \mathbf{v} \rangle_V = \langle \mathbf{u}, f(\mathbf{v}) \rangle_V, \quad \forall \mathbf{u}, \mathbf{v} \in V \quad (49)$$

という条件を満たすとき, 線型写像 f を V 上の対称変換と呼びます.

そこで, いま, 「内積を持つ線型空間」 $(V, \langle \cdot, \cdot \rangle_V)$ が, 勝手にひとつ与えられているとします.²⁰ このとき,

$$n = \dim_{\mathbb{R}} V$$

として, V の基底 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ を, 勝手にひとつ取ってきて,

$$V \ni \mathbf{u} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n \longleftrightarrow \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

¹⁹ 「内積を持つ線型空間」に関しては, 11 節を参照して下さい.

²⁰ これでは抽象的過ぎると思われる方は, $(V, \langle \cdot, \cdot \rangle_V)$ として, 数ベクトル空間 \mathbb{R}^n とその上の標準的なユークリッド内積 $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$, あるいは, \mathbb{R}^n の線型部分空間 $W \subset \mathbb{R}^n$ と, $\mathbf{u}, \mathbf{v} \in W$ に対して,

$$\langle \mathbf{u}, \mathbf{v} \rangle_W = \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n}$$

という式によって定まる W 上の内積 $\langle \cdot, \cdot \rangle_W$ だけを考えてもらっても構いません. また, $f: V \rightarrow V$ として, 対称行列 A を掛け算することによって定まる線型写像だけを考えてもらっても構いません. 実際, 6 節での確認作業で必要となるのは, このような場合だけです.

というように「番地割り」してみます. すると,

$$\begin{aligned}\mathbf{u} &= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n, \\ \mathbf{v} &= y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \cdots + y_n\mathbf{e}_n\end{aligned}$$

として, \mathbf{u} と \mathbf{v} の間の内積が,

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle_V &= \langle x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + \cdots + x_n\mathbf{e}_n, \mathbf{v} \rangle_V \\ &= \langle x_1\mathbf{e}_1, \mathbf{v} \rangle_V + \langle x_2\mathbf{e}_2, \mathbf{v} \rangle_V + \cdots + \langle x_n\mathbf{e}_n, \mathbf{v} \rangle_V \\ &= x_1 \cdot \langle \mathbf{e}_1, \mathbf{v} \rangle_V + x_2 \cdot \langle \mathbf{e}_2, \mathbf{v} \rangle_V + \cdots + x_n \cdot \langle \mathbf{e}_n, \mathbf{v} \rangle_V \\ &= \sum_{i=1}^n x_i \cdot \langle \mathbf{e}_i, \mathbf{v} \rangle_V \\ &= \sum_{i=1}^n x_i \cdot \langle \mathbf{e}_i, y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + \cdots + y_n\mathbf{e}_n \rangle_V \\ &= \sum_{i=1}^n x_i \cdot \{ \langle \mathbf{e}_i, y_1\mathbf{e}_1 \rangle_V + \langle \mathbf{e}_i, y_2\mathbf{e}_2 \rangle_V + \cdots + \langle \mathbf{e}_i, y_n\mathbf{e}_n \rangle_V \} \\ &= \sum_{i=1}^n x_i \cdot \{ y_1 \cdot \langle \mathbf{e}_i, \mathbf{e}_1 \rangle_V + y_2 \cdot \langle \mathbf{e}_i, \mathbf{e}_2 \rangle_V + \cdots + y_n \cdot \langle \mathbf{e}_i, \mathbf{e}_n \rangle_V \} \\ &= \sum_{i=1}^n x_i \cdot \left\{ \sum_{j=1}^n y_j \cdot \langle \mathbf{e}_i, \mathbf{e}_j \rangle_V \right\} \\ &= \sum_{i,j=1}^n x_i \cdot y_j \cdot \langle \mathbf{e}_i, \mathbf{e}_j \rangle_V \\ &= \sum_{i,j=1}^n x_i \cdot \langle \mathbf{e}_i, \mathbf{e}_j \rangle_V \cdot y_j\end{aligned}\tag{50}$$

というように表わせることが分かります. よって, (50) 式から,

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad G = \begin{pmatrix} \langle \mathbf{e}_1, \mathbf{e}_1 \rangle_V & \langle \mathbf{e}_1, \mathbf{e}_2 \rangle_V & \cdots & \langle \mathbf{e}_1, \mathbf{e}_n \rangle_V \\ \langle \mathbf{e}_2, \mathbf{e}_1 \rangle_V & \langle \mathbf{e}_2, \mathbf{e}_2 \rangle_V & \cdots & \langle \mathbf{e}_2, \mathbf{e}_n \rangle_V \\ \vdots & \vdots & \ddots & \vdots \\ \langle \mathbf{e}_n, \mathbf{e}_1 \rangle_V & \langle \mathbf{e}_n, \mathbf{e}_2 \rangle_V & \cdots & \langle \mathbf{e}_n, \mathbf{e}_n \rangle_V \end{pmatrix}$$

として,

————— V 上の内積の座標を用いた表示式 —————

$$\langle \mathbf{u}, \mathbf{v} \rangle_V = {}^t\mathbf{x} \cdot G \cdot \mathbf{y}\tag{51}$$

というように表わせることが分かります. 特に, $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ が V の正規直交基底のときには, すなわち,

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle_V = \begin{cases} 1, & i = j \text{ のとき} \\ 0, & i \neq j \text{ のとき} \end{cases}, \quad (i, j = 1, 2, \dots, n)$$

となるときには,

$$G = I$$

となることが分かりますから, (51) 式から,

$$\begin{aligned}\langle \mathbf{u}, \mathbf{v} \rangle_V &= {}^t \mathbf{x} \cdot \mathbf{y} \\ &= \langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}^n}\end{aligned}$$

となることが分かります. すなわち, この場合,

正規直交基底を用いると, 内積の構造まで込めて V と \mathbb{R}^n を同一視できる

$$(V, \langle \cdot, \cdot \rangle_V) \cong (\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathbb{R}^n}) \quad (52)$$

というように, 内積の構造まで込めて V と \mathbb{R}^n を同一視できることが分かります.

さらに, V 上の線型写像

$$f: V \rightarrow V$$

が, 勝手にひとつ与えられているとします. このとき, V の基底 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ に関する線型写像 f の表現行列 A を,

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

と表わすことにします. そこで, $i = 1, 2, \dots, n$ に対して,

$$\begin{aligned}f(\mathbf{e}_i) &= \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \cdots & \mathbf{e}_n \end{pmatrix} \begin{pmatrix} a_{1i} \\ a_{2i} \\ \vdots \\ a_{ni} \end{pmatrix} \\ &= a_{1i}\mathbf{e}_1 + a_{2i}\mathbf{e}_2 + \cdots + a_{ni}\mathbf{e}_n \\ &= \sum_{k=1}^n a_{ki}\mathbf{e}_k\end{aligned}$$

と表わせることに注意して,²¹

$$\mathbf{u} = \mathbf{e}_i, \quad \mathbf{v} = \mathbf{e}_j$$

として, (49) 式の両辺の値を考えてみると,

$$\begin{aligned}\langle f(\mathbf{e}_i), \mathbf{e}_j \rangle_V &= \left\langle \sum_{k=1}^n a_{ki}\mathbf{e}_k, \mathbf{e}_j \right\rangle_V \\ &= \sum_{k=1}^n a_{ki} \cdot \langle \mathbf{e}_k, \mathbf{e}_j \rangle_V\end{aligned} \quad (53)$$

²¹基底の元 \mathbf{e}_i の行き先 $f(\mathbf{e}_i)$ の「番地」が表現行列 A の i 列目の列ベクトルとなるのでした.

$$\begin{aligned}\langle \mathbf{e}_i, f(\mathbf{e}_j) \rangle_V &= \langle \mathbf{e}_i, \sum_{k=1}^n a_{kj} \mathbf{e}_k \rangle_V \\ &= \sum_{k=1}^n a_{kj} \cdot \langle \mathbf{e}_i, \mathbf{e}_k \rangle_V\end{aligned}\quad (54)$$

となることが分かります。したがって、 $f : V \rightarrow V$ が対称変換であるとする、(49) 式、(53) 式、(54) 式から、

$$\sum_{k=1}^n a_{ki} \cdot \langle \mathbf{e}_k, \mathbf{e}_j \rangle_V = \sum_{k=1}^n \langle \mathbf{e}_i, \mathbf{e}_k \rangle_V \cdot a_{kj}$$

となることが分かりますから、

$${}^tAG = GA \quad (55)$$

となることが分かります。特に、 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ が V の正規直交基底のときには、

$$G = I$$

となることが分かりますから、(55) 式から、

$${}^tA = A$$

となることが分かります。すなわち、正規直交基底に関する対称変換の表現行列は対称行列になることが分かります。

逆に、 V の正規直交基底 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ に関する線型写像 $f : V \rightarrow V$ の表現行列 A が対称行列であったと仮定してみます。このとき、(53) 式、(54) 式から、

$$\begin{aligned}\langle f(\mathbf{e}_i), \mathbf{e}_j \rangle_V &= a_{ji}, \\ \langle \mathbf{e}_i, f(\mathbf{e}_j) \rangle_V &= a_{ij}\end{aligned}$$

となることが分かりますから、

$$\langle f(\mathbf{e}_i), \mathbf{e}_j \rangle_V = \langle \mathbf{e}_i, f(\mathbf{e}_j) \rangle_V \quad (56)$$

となることが分かります。そこで、 $\mathbf{u}, \mathbf{v} \in V$ を、

$$\begin{aligned}\mathbf{u} &= x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n, \\ \mathbf{v} &= y_1 \mathbf{e}_1 + y_2 \mathbf{e}_2 + \dots + y_n \mathbf{e}_n\end{aligned}$$

と表わして、(49) 式の両辺の値を考えてみると、

$$\begin{aligned}\langle f(\mathbf{u}), \mathbf{v} \rangle_V &= \langle f\left(\sum_{i=1}^n x_i \mathbf{e}_i\right), \mathbf{v} \rangle_V \\ &= \left\langle \sum_{i=1}^n x_i \cdot f(\mathbf{e}_i), \mathbf{v} \right\rangle_V \\ &= \sum_{i=1}^n x_i \cdot \langle f(\mathbf{e}_i), \mathbf{v} \rangle_V\end{aligned}$$

$$\begin{aligned}
&= \sum_{i=1}^n x_i \cdot \langle f(\mathbf{e}_i), \sum_{j=1}^n y_j \mathbf{e}_j \rangle_V \\
&= \sum_{i,j=1}^n x_i \cdot y_j \cdot \langle f(\mathbf{e}_i), \mathbf{e}_j \rangle_V
\end{aligned} \tag{57}$$

$$\begin{aligned}
\langle \mathbf{u}, f(\mathbf{v}) \rangle_V &= \langle \mathbf{u}, f(\sum_{j=1}^n y_j \mathbf{e}_j) \rangle_V \\
&= \langle \mathbf{u}, \sum_{j=1}^n y_j \cdot f(\mathbf{e}_j) \rangle_V \\
&= \sum_{j=1}^n y_j \cdot \langle \mathbf{u}, f(\mathbf{e}_j) \rangle_V \\
&= \sum_{j=1}^n y_j \cdot \langle \sum_{i=1}^n x_i \mathbf{e}_i, f(\mathbf{e}_j) \rangle_V \\
&= \sum_{i,j=1}^n x_i \cdot y_j \cdot \langle \mathbf{e}_i, f(\mathbf{e}_j) \rangle_V
\end{aligned} \tag{58}$$

となることが分かります. よって, (56) 式, (57) 式, (58) 式から,

$$\langle f(\mathbf{u}), \mathbf{v} \rangle_V = \langle \mathbf{u}, f(\mathbf{v}) \rangle_V$$

となることが分かりますから, 線型写像 $f: V \rightarrow V$ は対称変換であることが分かります.

以上の議論から, 正規直交基底を用いて, 内積を持つ線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ に「番地割り」して考えると,

正規直交基底を用いた「番地割り」のもとでの対応

線型空間 V	\cong	数ベクトル空間 \mathbb{R}^n
V 上の内積: $\langle \cdot, \cdot \rangle_V$	\longleftrightarrow	\mathbb{R}^n 上の内積: $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$
対称変換: $f: V \rightarrow V$	\longleftrightarrow	対称行列: A

と対応していることが分かります. 第9回の問3のところで見たとおり, 一般に, 線型空間 V 上に線型写像 $f: V \rightarrow V$ が与えられているときに,

線型写像とその表現行列の固有値, 固有ベクトルの間の対応

線型写像		表現行列
線型写像 f の固有値	\longleftrightarrow	表現行列 A の固有値
線型写像 f の固有ベクトル	\longleftrightarrow	表現行列 A の固有ベクトル

というように, 線型写像 f の固有値, 固有ベクトルは, 表現行列 A の固有値, 固有ベクトルに対応することが分かります. したがって, 対称変換 f の固有値, 固有ベクトルは, その表現行列である対称行列 A の固有値, 固有ベクトルに対応することが分かります. すると, 2節で見たとおり, 対称行列のすべての固有値は実数になることが分かりますから, 対

称変換 f のすべての固有値も実数になることが分かります。また、対称行列の相異なる固有値に対応する固有ベクトルは互いに直交すること、上の同一視のもとで、 V 上の内積 $\langle \cdot, \cdot \rangle_V$ は \mathbb{R}^n 上の内積 $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ に対応していることが分かりますから、対称変換 f の相異なる固有値に対応する固有ベクトルも互いに直交することが分かります。²² 以上から、対称変換の場合にも、固有値、固有ベクトルに関して、

対称変換の基本的な性質

- (I) 対称変換の固有値はすべて実数になる。
- (II) 対称変換の相異なる固有値に対応する固有ベクトルは互いに直交する。

となることが分かりました。

5 直交補空間について

さて、第9回の問3のところでは、「直和」という概念について触れ、「行列の対角化の問題」を「固有ベクトル空間分解」という形で理解できることを見ました。そこで、ここでは、6節での確認作業の準備として、内積を持つ線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ に対して、 V の線型部分空間 $W \subset V$ が、勝手にひとつ与えられると、 V が「 W の方向」と「 W に直交する方向」に直分解するというものを見ておくことにします。

そこで、いま、内積を持つ線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ が、勝手にひとつ与えられているとします。²³ 一般に、 V の線型部分空間 $W \subset V$ に対して、 W のすべての元と直交するような V の元全体の集合

線型部分空間 W の直交補空間

$$W^\perp = \{ \mathbf{v} \in V \mid \langle \mathbf{w}, \mathbf{v} \rangle_V = 0, \forall \mathbf{w} \in W \}$$

を考えると、 $W^\perp \subset V$ も V の線型部分空間になることが分かります。²⁴ こうして定まる線型部分空間 W^\perp を (線型空間 V における) W の直交補空間と呼びます。

いま、

$$m = \dim_{\mathbb{R}} W$$

として、 W の基底 $\{ \mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m \}$ を、勝手にひとつ取ってきたとします。このとき、 $i = 1, 2, \dots, m$ に対して、

$$\mathbf{e}_i \in W$$

となりますから、

$$\mathbf{v} \in W^\perp \implies \langle \mathbf{e}_i, \mathbf{v} \rangle_V = 0 \tag{59}$$

²²あるいは、2節での議論と全く同様に、(49) を用いて、直接、対称変換 f の相異なる固有値に対応する固有ベクトルは互いに直交することを示すこともできます。

²³前と同様に、これでは抽象的過ぎると思われる方は、 V として、数ベクトル空間 \mathbb{R}^n や \mathbb{R}^n の線型部分空間 $V \subset \mathbb{R}^n$ だけを考慮してもらっても構いません。

²⁴皆さん、確かめてみて下さい。

となることが分かります. 逆に, $i = 1, 2, \dots, m$ に対して,

$$\langle \mathbf{e}_i, \mathbf{v} \rangle_V = 0 \quad (60)$$

となると仮定してみます. このとき, $\mathbf{w} \in W$ を,

$$\mathbf{w} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_m \mathbf{e}_m, \quad (a_1, a_2, \dots, a_m \in \mathbb{R})$$

というように表わしてみると, (60) 式から,

$$\begin{aligned} \langle \mathbf{w}, \mathbf{v} \rangle_V &= \langle a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_m \mathbf{e}_m, \mathbf{v} \rangle_V \\ &= a_1 \cdot \langle \mathbf{e}_1, \mathbf{v} \rangle_V + a_2 \cdot \langle \mathbf{e}_2, \mathbf{v} \rangle_V + \dots + a_m \cdot \langle \mathbf{e}_m, \mathbf{v} \rangle_V \\ &= 0 \end{aligned} \quad ((60) \text{ 式から})$$

となることが分かりますから,

$$\mathbf{v} \in W^\perp$$

となることが分かります. よって,

$$\langle \mathbf{e}_i, \mathbf{v} \rangle_V = 0 \implies \mathbf{v} \in W^\perp \quad (61)$$

となることが分かりました. したがって, (59) 式, (61) 式から,

$$\mathbf{v} \in W^\perp \iff \langle \mathbf{e}_i, \mathbf{v} \rangle_V = 0, \quad (i = 1, 2, \dots, m)$$

となることが分かりますから, W の直交補空間 W^\perp は,

W の基底を用いた直交補空間 W^\perp の表示

$$W^\perp = \{ \mathbf{v} \in V \mid \langle \mathbf{e}_i, \mathbf{v} \rangle_V = 0, \quad (i = 1, 2, \dots, m) \} \quad (62)$$

というようにも表わせることが分かります.

さて, V の線型部分空間 $W \subset V$ が, 勝手にひとつ与えられたとすると, V は,

$$V = W \oplus W^\perp$$

というように, 「 W の方向」と「 W に直交する方向」 W^\perp に直和分解することが, 次のようにして分かります. そのためには,

線型空間 V が W と W^\perp の直和となるための条件

(イ) 勝手な元 $\mathbf{u} \in V$ に対して,

$$\mathbf{u} = \mathbf{w} + \mathbf{v}$$

となるような元 $\mathbf{w} \in W, \mathbf{v} \in W^\perp$ が存在する.

(ロ) $\mathbf{w} \in W, \mathbf{v} \in W^\perp$ として,

$$\mathbf{0} = \mathbf{w} + \mathbf{v} \implies \mathbf{w} = \mathbf{v} = \mathbf{0}$$

となる.

という二つの条件が満たされることが確かめられればよいということになります.

そこで、まず、(イ) という条件について考えてみることにします。いま、 $\mathbf{u} \in V$ が、勝手にひとつ与えられているとします。このとき、

$$\mathbf{u} = \mathbf{w} + \mathbf{v} \quad (63)$$

となるような元 $\mathbf{w} \in W$, $\mathbf{v} \in W^\perp$ を見つけなければなりません。様子を探ってみるために、 $\mathbf{u} \in V$ が (63) 式のように表わされるとすると、 $\mathbf{w} \in W$ や $\mathbf{v} \in W^\perp$ はどのような元でなければならないのかということに「当たり」をつけてみることにします。そのために、上と同様に、

$$m = \dim_{\mathbb{R}} W$$

として、 W の正規直交基底 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_m\}$ を、勝手にひとつ取ってきて、 $a_1, a_2, \dots, a_m \in \mathbb{R}$ として、 $\mathbf{w} \in W$ を、

$$\mathbf{w} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_m \mathbf{e}_m \quad (64)$$

と表わして考えてみることにします。すると、(63) 式、(64) 式から、 $\mathbf{u} \in V$ は、

$$\mathbf{u} = a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_m \mathbf{e}_m + \mathbf{v} \quad (65)$$

というように表わせることになります。そこで、いま、 $i, j = 1, 2, \dots, m$ に対して、

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle_V = \begin{cases} 1, & i = j \text{ のとき} \\ 0, & i \neq j \text{ のとき} \end{cases} \quad (66)$$

となることと、 $\mathbf{v} \in W^\perp$ となることに注意して、 $\mathbf{e}_i \in W$ と (65) 式の両辺に現われる元の内積の値を考えてみます。すると、

$$\begin{aligned} \langle \mathbf{e}_i, \mathbf{u} \rangle_V &= \langle \mathbf{e}_i, a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \dots + a_m \mathbf{e}_m + \mathbf{v} \rangle_V \\ &= \langle \mathbf{e}_i, a_1 \mathbf{e}_1 \rangle_V + \langle \mathbf{e}_i, a_2 \mathbf{e}_2 \rangle_V + \dots + \langle \mathbf{e}_i, a_m \mathbf{e}_m \rangle_V + \langle \mathbf{e}_i, \mathbf{v} \rangle_V \\ &= a_1 \cdot \langle \mathbf{e}_i, \mathbf{e}_1 \rangle_V + a_2 \cdot \langle \mathbf{e}_i, \mathbf{e}_2 \rangle_V + \dots + a_m \cdot \langle \mathbf{e}_i, \mathbf{e}_m \rangle_V \quad (\mathbf{v} \in W^\perp \text{ より}) \\ &= a_i \quad ((66) \text{ 式より}) \end{aligned}$$

となることが分かりますから、 $i = 1, 2, \dots, m$ に対して、

$$a_i = \langle \mathbf{e}_i, \mathbf{u} \rangle_V \quad (67)$$

でなければならないことが分かります。こうして、 $\mathbf{u} \in V$ が (63) 式のように表わされるとすると、 $\mathbf{w} \in W$ がどのような元でなければならないのかということに「当たり」が付きしました。

そこで、いま、与えられた元 $\mathbf{u} \in V$ に対して、(67) 式によって、 $a_i \in \mathbb{R}$ の値を定めてみることにします。すると、(65) 式より、 \mathbf{v} は、

$$\mathbf{v} = \mathbf{u} - a_1 \mathbf{e}_1 - a_2 \mathbf{e}_2 - \dots - a_m \mathbf{e}_m \quad (68)$$

というように定めなければならないことが分かります。そこで、(68) 式によって、 $\mathbf{v} \in V$ を定めてみます。このとき、(66) 式、(67) 式、(68) 式から、 $i = 1, 2, \dots, m$ に対して、

$$\langle \mathbf{e}_i, \mathbf{v} \rangle_V = \langle \mathbf{e}_i, \mathbf{u} - a_1 \mathbf{e}_1 - a_2 \mathbf{e}_2 - \dots - a_m \mathbf{e}_m \rangle_V \quad ((68) \text{ 式より})$$

$$\begin{aligned}
&= \langle \mathbf{e}_i, \mathbf{u} \rangle_V - a_1 \cdot \langle \mathbf{e}_i, \mathbf{e}_1 \rangle_V - a_2 \cdot \langle \mathbf{e}_i, \mathbf{e}_2 \rangle_V \\
&\quad - \cdots - a_m \cdot \langle \mathbf{e}_i, \mathbf{e}_m \rangle_V \\
&= \langle \mathbf{e}_i, \mathbf{u} \rangle_V - a_i && \text{((66) 式より)} \\
&= 0 && \text{((67) 式より)} \tag{69}
\end{aligned}$$

となることが分かります. よって, (69) 式と W^\perp に対する (62) 式の表示から,

$$\mathbf{v} \in W^\perp$$

となることが分かります. 以上より, (67) 式によって定まる係数 $a_1, a_2, \dots, a_m \in \mathbb{R}$ を用いて,

$$\begin{aligned}
\mathbf{w} &= a_1 \mathbf{e}_1 + a_2 \mathbf{e}_2 + \cdots + a_m \mathbf{e}_m \in W \\
\mathbf{v} &= \mathbf{u} - a_1 \mathbf{e}_1 - a_2 \mathbf{e}_2 - \cdots - a_m \mathbf{e}_m \in W^\perp
\end{aligned}$$

と定めることで,

$$\mathbf{u} = \mathbf{w} + \mathbf{v}$$

と「成分分解」できることが分かりますから, (イ) という条件が満たされることが分かります.

次に, (ロ) という条件について考えてみます. そこで, いま, $\mathbf{w} \in W, \mathbf{v} \in W^\perp$ として,

$$\mathbf{0} = \mathbf{w} + \mathbf{v} \tag{70}$$

となると仮定してみます. このとき,

$$\langle \mathbf{w}, \mathbf{v} \rangle_V = 0 \tag{71}$$

となることに注意して, (70) 式の両辺と \mathbf{w} の間の内積を考えると,

$$\begin{aligned}
0 &= \langle \mathbf{w}, \mathbf{w} + \mathbf{v} \rangle_V \\
&= \langle \mathbf{w}, \mathbf{w} \rangle_V + \langle \mathbf{w}, \mathbf{v} \rangle_V \\
&= \|\mathbf{w}\|^2 && \text{((71) 式から)}
\end{aligned}$$

となることが分かりますから,

$$\mathbf{w} = \mathbf{0} \tag{72}$$

となることが分かります. 全く同様に, (70) 式の両辺と \mathbf{v} の間の内積を考えると,

$$\begin{aligned}
0 &= \langle \mathbf{v}, \mathbf{w} + \mathbf{v} \rangle_V \\
&= \langle \mathbf{v}, \mathbf{w} \rangle_V + \langle \mathbf{v}, \mathbf{v} \rangle_V \\
&= \|\mathbf{v}\|^2 && \text{((71) 式から)}
\end{aligned}$$

となることが分かりますから,

$$\mathbf{v} = \mathbf{0} \tag{73}$$

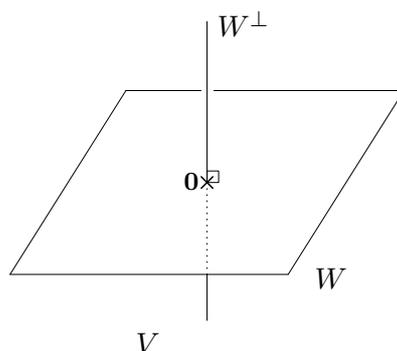


図 1: 線型空間 V の線型部分空間 $W \subset V$ が与えられると, V は「 W の方向」と「 W に直交する方向」 W^\perp に直和分解される.

となることが分かります. よって, (72) 式, (73) 式から,

$$\mathbf{w} = \mathbf{v} = \mathbf{0}$$

となることが分かりますから, (ロ) という条件も満たされることが分かります.

以上から, (イ), (ロ) という二つの条件が満たされることが分かりましたから, 線型空間 V は,

線型部分空間 W とその直交補空間 W^\perp による V の直和分解

$$V = W \oplus W^\perp \quad (74)$$

というように「 W の方向」と「 W に直交する方向」 W^\perp に直和分解されることが分かります (図 1 を参照). このように, W^\perp は, 線型空間 V の中で, 「 W の方向」だけではカバーできない方向を「 W に直交する方向」として「補っている」ので, W の直交補空間と呼ばれます.

そこで, さらに, V 上の対称変換

$$f: V \rightarrow V$$

が与えられているとして, 線型部分空間 W が,

線型写像 f の不変部分空間となるための条件

$$\mathbf{w} \in W \implies f(\mathbf{w}) \in W \quad (75)$$

という条件を満たしているとします. 一般に, 線型空間 V 上の線型写像 $f: V \rightarrow V$ が与えられているときに, (75) 式の条件を満たす V の線型部分空間 $W \subset V$ を線型写像 f の不変部分空間と呼びます.²⁵ 例えば, 線型写像 f の固有ベクトル空間

$$V(\lambda) = \{ \mathbf{w} \in V \mid f(\mathbf{w}) = \lambda \mathbf{w} \}$$

²⁵言葉の意味は「 W の勝手な元 $\mathbf{w} \in W$ に線型写像 f を施しても, その行き先 $f(\mathbf{w}) \in V$ は W の外に出て行かない」ということ, すなわち, 「線型写像 f を施すという操作で W は不変に保たれる」ということです.

は、こうした不変部分空間の代表例です。

そこで、 $W \subset V$ として、対称変換 f の不変部分空間を取ってくると、どのようなことになるのかということを考えてみます。すなわち、(74) 式から、線型空間 V が、

$$V = W \oplus W^\perp$$

というように直和分解することが分かりますが、 W が対称変換 f の不変部分空間であるときに、 W の直交補空間 W^\perp も対称変換 f の不変部分空間になるのかどうかということを考えてみます。

そこで、いま、 f が対称変換であるということに注目して、

$$\langle f(\mathbf{w}), \mathbf{v} \rangle_V = \langle \mathbf{w}, f(\mathbf{v}) \rangle_V, \quad \forall \mathbf{w}, \mathbf{v} \in V \quad (76)$$

という対称変換の定義式における $\mathbf{w}, \mathbf{v} \in V$ として、

$$\mathbf{w} \in W, \mathbf{v} \in W^\perp$$

となる元を取ってくると、どのようなことが分かるのかということを考えてみます。すると、 W は対称変換 f の不変部分空間であると仮定していましたが、 $f(\mathbf{w}) \in W$ となることに注意すると、(76) 式の左辺は、

$$\langle f(\mathbf{w}), \mathbf{v} \rangle_V = 0 \quad (f(\mathbf{w}) \in W, \mathbf{v} \in W^\perp \text{ より}) \quad (77)$$

となることが分かります。よって、(76) 式、(77) 式から、 $\mathbf{v} \in W^\perp$ であるとする、勝手な元 $\mathbf{w} \in W$ に対して、

$$\langle \mathbf{w}, f(\mathbf{v}) \rangle_V = 0$$

となることが分かりますから、

$$f(\mathbf{v}) \in W^\perp$$

となることが分かります。すなわち、

$$\mathbf{v} \in W^\perp \implies f(\mathbf{v}) \in W^\perp \quad (78)$$

となることが分かります。以上から、 W が対称変換 f の不変部分空間であるときに、 W の直交補空間 W^\perp も対称変換 f の不変部分空間になることが分かりました。

こうして、対称変換 f の不変部分空間 W を用いて、

$$V = W \oplus W^\perp$$

というように直和分解して考えることにより、「対称変換 $f : V \rightarrow V$ の様子を理解する問題」が「対称変換 f の定義域や値域を W や W^\perp に制限することによって得られる対称変換

$$\begin{aligned} f|_W &: W \rightarrow W \\ f|_{W^\perp} &: W^\perp \rightarrow W^\perp \end{aligned}$$

の様子を理解する問題」に帰着できることが分かります。このように、「サイズの大きな対称変換の様子を理解する問題」を「サイズがより小さな対称変換の様子を理解する問題」に帰着して考えることができるということが、6 節での確認作業における一番大きなポイントになります。

6 対称行列の固有ベクトル空間分解について

以上の準備のもとで、対称行列に対しては「対角化の問題」がいつでも解決できるということを理論的に保証することができるということを確認してみることになります。2節で見たように、対称行列 A の固有値や固有ベクトルの性質をより良く理解するためには、

$$A = {}^t A$$

という「数が並んだもの」としての対称行列の「姿」に注目するのではなく、対称行列 A を「線型写像」として眺めてみたときに、

対称行列の特徴付け

$$\langle A\mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{u}, A\mathbf{v} \rangle, \quad \forall \mathbf{u}, \mathbf{v} \in \mathbb{R}^n, \quad (79)$$

というように、「内積との相性が良い線型写像である」という性質に注目することが大切です。そこで、ここでも、「対称行列の対角化の問題」を座標軸の取り方によらない形で考察してみることにします。

第9回の問3のところでは、「行列の対角化の問題」をより良く理解するために、「対角化の問題」を座標軸の取り方によらない形で定式化することを考えました。その結果、「行列の対角化の問題」を「固有ベクトル空間分解」という形で定式化できることを見ました。そこで、ここでは、4節、5節の結果をもとにして、 n 行 n 列の対称行列 A が、勝手にひとつ与えられているときに、数ベクトル空間 \mathbb{R}^n が対称行列 A の固有ベクトル空間の直和に分解するということを確認してみることになります。

そこで、いま、対称行列 A の相異なる固有値を $\lambda_1, \lambda_2, \dots, \lambda_L \in \mathbb{R}$ として、 $i = 1, 2, \dots, L$ に対して、固有値 λ_i に対応する固有ベクトル空間を、

固有値 λ_i に対応する固有ベクトル空間

$$V(\lambda_i) = \{ \mathbf{u} \in \mathbb{R}^n \mid A\mathbf{u} = \lambda_i \mathbf{u} \}$$

と表わすことにします。このとき、確かめたいことは、

対称行列 A に関する \mathbb{R}^n の固有ベクトル空間分解

$$\mathbb{R}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \dots \oplus V(\lambda_L)$$

というように、数ベクトル空間 \mathbb{R}^n が対称行列 A の固有ベクトル空間の直和に分解するということになりますが、そのためには、

対称行列 A に関して, \mathbb{R}^n が固有ベクトル空間分解するための条件

(イ) 勝手なベクトル $\mathbf{u} \in \mathbb{R}^n$ に対して,

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L$$

となるようなベクトル $\mathbf{u}_1 \in V(\lambda_1), \mathbf{u}_2 \in V(\lambda_2), \dots, \mathbf{u}_L \in V(\lambda_L)$ が存在する.

(ロ) $\mathbf{u}_1 \in V(\lambda_1), \mathbf{u}_2 \in V(\lambda_2), \dots, \mathbf{u}_L \in V(\lambda_L)$ として,

$$\mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L \implies \mathbf{u}_1 = \mathbf{u}_2 = \cdots = \mathbf{u}_L = \mathbf{0}$$

となる.

という二つの条件が満たされることが確かめられれば良いということになります.

そこで, まず, (イ) という条件について考えてみることにします. いま, 固有ベクトル空間 $V(\lambda_1), V(\lambda_2), \dots, V(\lambda_L)$ の和を,

$$\begin{aligned} W &= V(\lambda_1) + V(\lambda_2) + \cdots + V(\lambda_L) \\ &= \{ \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L \in \mathbb{R}^n \mid \mathbf{u}_1 \in V(\lambda_1), \mathbf{u}_2 \in V(\lambda_2), \dots, \mathbf{u}_L \in V(\lambda_L) \} \end{aligned}$$

と表わすことにします.²⁶ このとき, 確かめたいことは,

$$\mathbb{R}^n = W$$

となるということになります.

そこで, いま,

$$W \subsetneq \mathbb{R}^n \tag{80}$$

であったと仮定してみます. このとき, 5 節で見たように,

$$\mathbb{R}^n = W \oplus W^\perp$$

というように, 数ベクトル空間 \mathbb{R}^n が「 W の方向」と「 W に直交する方向」 W^\perp に直和分解することが分かります. また, 対称行列 A を掛け算することによって定まる \mathbb{R}^n 上の対称変換を,

$$f_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$$

と表わすことにすると, $\mathbf{u}_1 \in V(\lambda_1), \mathbf{u}_2 \in V(\lambda_2), \dots, \mathbf{u}_L \in V(\lambda_L)$ として,

$$\mathbf{w} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L \in W$$

に対して,

$$\begin{aligned} f_A(\mathbf{w}) &= f_A(\mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L) \\ &= f_A(\mathbf{u}_1) + f_A(\mathbf{u}_2) + \cdots + f_A(\mathbf{u}_L) \end{aligned}$$

²⁶すなわち, W とは, $V(\lambda_1), V(\lambda_2), \dots, V(\lambda_L)$ の元を用いて「成分分解」ができる \mathbb{R}^n の元全体の集合のことです.

$$\begin{aligned}
&= A\mathbf{u}_1 + A\mathbf{u}_2 + \cdots + A\mathbf{u}_L \\
&= \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{u}_2 + \cdots + \lambda_L\mathbf{u}_L
\end{aligned}$$

となることが分かりますから,

$$f_A(\mathbf{w}) \in W$$

となることが分かります. よって,

$$\mathbf{w} \in W \implies f_A(\mathbf{w}) \in W$$

となることが分かりましたから, W は対称変換 f_A の不変部分空間であることが分かります. すると, 5節で見たことから, W の直交補空間 W^\perp も対称変換 f_A の不変部分空間となることが分かります.

そこで, 対称変換 f_A の定義域と値域を W^\perp に制限することによって得られる W^\perp 上の対称変換を,

$$g = f_A|_{W^\perp} : W^\perp \rightarrow W^\perp$$

と表わすことにします.²⁷ このとき, 対称変換 g の $\mathbf{0}$ と異なる固有ベクトルを, 何でもよいから勝手にひとつ取ってきて, $\mathbf{0} \neq \mathbf{v} \in W^\perp$ と表わすことにします. また, 固有ベクトル \mathbf{v} の固有値を $\lambda \in \mathbb{R}$ と表わすことにします. すると,

$$g(\mathbf{v}) = \lambda\mathbf{v} \tag{81}$$

となりますが, 対称変換 g の定義から,

$$g(\mathbf{v}) = A\mathbf{v} \tag{82}$$

と表わせることに注意すると, (81) 式, (82) 式から,

$$A\mathbf{v} = \lambda\mathbf{v}$$

となることが分かります. すなわち, \mathbf{v} は対称行列 A の固有ベクトルであることが分かります. いま, 対称行列 A の相異なる固有値を $\lambda_1, \lambda_2, \dots, \lambda_L \in \mathbb{R}$ としていましたから, λ は, $i_0 \in \{1, 2, \dots, L\}$ となる適当な番号 i_0 を用いて,

$$\lambda = \lambda_{i_0}$$

と表わせることが分かります. すると,

$$\mathbf{v} \in V(\lambda_{i_0}) \subset W$$

となることが分かりますから,²⁸

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

²⁷ 平たく言えば, 対称変換 g とは, 対称行列 A を掛け算する操作を W^\perp 上だけで考えたもののことです.

²⁸ $\mathbf{v} \in V(\lambda_{i_0})$ は, $i \neq i_0$ のとき, $\mathbf{u}_i = \mathbf{0} \in V(\lambda_i)$, $i = i_0$ のとき, $\mathbf{u}_{i_0} = \mathbf{v} \in V(\lambda_{i_0})$ として,

$$\mathbf{v} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L$$

と表わせることが分かります.

となることが分かります。よって、

$$\mathbf{0} \neq \mathbf{v} \in W \cap W^\perp$$

となることが分かりますから、

$$W \cap W^\perp \neq \{\mathbf{0}\}$$

となってしまう、 \mathbb{R}^n が、

$$\mathbb{R}^n = W \oplus W^\perp$$

というように直和分解されていたことに矛盾してしまいます。²⁹ そこで、何がいけなかったのかと反省してみると、

$$W^\perp \neq \{\mathbf{0}\}$$

と仮定していたこと、すなわち、(80) 式のように、

$$W \subsetneq \mathbb{R}^n$$

と仮定していたことが、上のような矛盾が導かれた原因であることが分かります。よって、

$$\begin{aligned} \mathbb{R}^n &= W \\ &= V(\lambda_1) + V(\lambda_2) + \cdots + V(\lambda_L) \end{aligned}$$

となることが分かりましたから、(イ) という条件が満たされることが分かります。

次に、(ロ) という条件について考えてみます。いま、 $\mathbf{u}_1 \in V(\lambda_1)$, $\mathbf{u}_2 \in V(\lambda_2)$, \dots , $\mathbf{u}_L \in V(\lambda_L)$ として、

$$\mathbf{0} = \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L \quad (83)$$

であると仮定してみます。このとき、2 節で見たように、

$$i \neq j \implies V(\lambda_i) \perp V(\lambda_j) \quad (84)$$

となることに注意して、 $i = 1, 2, \dots, L$ に対して、 \mathbf{u}_i と (84) 式の両辺との間の内積を考え、

$$\begin{aligned} 0 &= \langle \mathbf{u}_i, \mathbf{u}_1 + \mathbf{u}_2 + \cdots + \mathbf{u}_L \rangle \\ &= \langle \mathbf{u}_i, \mathbf{u}_1 \rangle + \langle \mathbf{u}_i, \mathbf{u}_2 \rangle + \cdots + \langle \mathbf{u}_i, \mathbf{u}_L \rangle \\ &= \langle \mathbf{u}_i, \mathbf{u}_i \rangle && ((84) \text{ 式から}) \\ &= \|\mathbf{u}_i\|^2 \end{aligned}$$

となることが分かりますから、

$$\mathbf{u}_i = \mathbf{0}$$

²⁹すなわち、 $W \cap W^\perp \neq \{\mathbf{0}\}$ であるとする、 $\mathbf{0} \neq \mathbf{v} \in W \cap W^\perp$ に対して、例えば、 $\mathbf{v} \in W$ と考えるか、 $\mathbf{v} \in W^\perp$ と考えるかに応じて、

$$\begin{aligned} \mathbf{v} &= \mathbf{v} + \mathbf{0} \in W \oplus W^\perp \\ &= \mathbf{0} + \mathbf{v} \in W \oplus W^\perp \end{aligned}$$

というように「成分分解」が何通りにもできてしまうことになるわけです。

となることが分かります. よって,

$$\mathbf{u}_1 = \mathbf{u}_2 = \cdots = \mathbf{u}_L = \mathbf{0}$$

となることが分かりますから, (ロ) という条件も満たされることが分かります.

以上から, (イ), (ロ) という二つの条件が満たされることが分かりましたから, \mathbb{R}^n は,

$$\mathbb{R}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_L)$$

というように, 対称行列 A の固有ベクトル空間の直和に分解することが分かりました.

7 直交行列による対称行列の対角化について

さて, 6 節では, n 行 n 列の対称行列 A が, 勝手にひとつ与えられたときに, 数ベクトル空間 \mathbb{R}^n が,

$$\mathbb{R}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_L) \quad (85)$$

というように, 対称行列 A の固有ベクトル空間の直和に分解することを見ました. そこで, ここでは, 6 節の結果を「対称行列の対角化の問題」に読み直してみることになります.

第 9 回の問 3 のところで見たとように, n 行 n 列の対称行列 A に対して,

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

となる正則行列 P と対角行列 Λ が存在するためには, (85) 式のように, \mathbb{R}^n が対称行列 A の固有ベクトル空間の直和に分解する必要がありますが, 逆に, \mathbb{R}^n が (85) 式のように固有ベクトル空間分解している場合には, 後は, それぞれの固有ベクトル空間 $V(\lambda_i)$, ($i = 1, 2, \dots, L$) の基底を勝手に一組ずつ取ってきて, それらの基底の元をすべて集めたものを列ベクトルとする行列

$$P = \begin{pmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{pmatrix} \quad (86)$$

を考えると, P は正則行列となり, $P^{-1}AP$ は対角行列となるのでした. こうして, 「対称行列に対する対角化の問題」はいつでも解決することができることが分かりました.

さて, 2 節で見たように, 行列 A が対称行列の場合には, $i, j = 1, 2, \dots, L$ に対して,

$$i \neq j \implies V(\lambda_i) \perp V(\lambda_j) \quad (87)$$

となることが分かりますから, (85) 式の右辺に現われる固有ベクトル空間は互いに直交していることが分かります. このことに注意して, それぞれの固有ベクトル空間 $V(\lambda_i)$ の基底として, 特に, 「正規直交基底」を取ってくることにすると, これらの基底の元をすべて集めてくることによって得られる \mathbb{R}^n の基底 $\{\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n\}$ も「正規直交基底」になることが分かります.³⁰ よって, 3 節で見たように, このとき, (86) 式で与えられる正則行列 P は直交行列になることが分かりますから, 行列 A が対称行列の場合には, 対角化の問題を解決する正則行列 P として「直交行列」が取れることが分かります.

³⁰ 皆さん, 確かめてみて下さい.

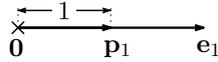


図 2: e_1 の長さを 1 にするように p_1 を定める.

実際に、このような直交行列 P を求めるためには、それぞれの固有ベクトル空間 $V(\lambda_i)$ から正規直交基底を取ってくる必要がありますが、そのためには、最初に、固有ベクトル空間 $V(\lambda_i)$ の基底を勝手にひとつ取ってきて、それらの基底をもとにして、次のように $V(\lambda_i)$ の正規直交基底を構成すればよいことが分かります。一般の場合でも、考え方の本質は変わりませんから、以下では、

$$\dim_{\mathbb{R}} V(\lambda_i) = 3$$

であるとして、 $V(\lambda_i)$ の基底 $\{e_1, e_2, e_3\}$ を、勝手にひとつ取ってきたという場合に説明することにします。

そこで、問題は「基底 $\{e_1, e_2, e_3\}$ から $V(\lambda_i)$ の正規直交基底 $\{p_1, p_2, p_3\}$ を構成する」ということになりますが、構成のアイデアはとても単純で、「長さが 1 で、互いに直交する」という条件が満たされるように、 e_1 から始めて、順番に基底の元を取り換えてゆくということです。

そこで、まず、 e_1 の長さを 1 にすることを考えて、

$$p_1 = \frac{e_1}{\|e_1\|}$$

と定めることにします (図 2 を参照)。こうして、 $\{e_1, e_2, e_3\}$ の代わりに、 $\{p_1, e_2, e_3\}$ という基底を考えることにすれば、取りあえず、 p_1 の長さは 1 となることが分かります。

次に、 e_2 を p_1 と直交するように取り直すために、 $a_1 \in \mathbb{R}$ として、

$$e_2 = a_1 p_1 + f_2 \tag{88}$$

というように、ベクトル e_2 を「 p_1 方向の成分」 $a_1 p_1$ と「 p_1 に直交する方向の成分」 f_2 に分解することを考えてみます。そのために、5 節と同様に、

$$\langle p_1, f_2 \rangle = 0 \tag{89}$$

となることに注意して、 p_1 と (88) 式の両辺との間の内積を考えると、

$$\begin{aligned} \langle p_1, e_2 \rangle &= \langle p_1, a_1 p_1 + f_2 \rangle \\ &= a_1 \cdot \langle p_1, p_1 \rangle + \langle p_1, f_2 \rangle \\ &= a_1 \cdot \langle p_1, p_1 \rangle && \text{((89) 式から)} \\ &= a_1 && \text{(} \|p_1\| = 1 \text{ より)} \end{aligned}$$

となることが分かりますから、 e_1 は、

$$e_2 = \langle p_1, e_2 \rangle \cdot p_1 + f_2$$

というように分解することが分かります。そこで、

$$f_2 = e_2 - \langle p_1, e_2 \rangle \cdot p_1 \tag{90}$$

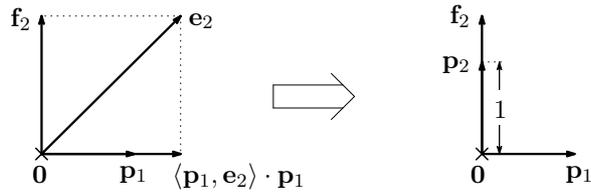


図 3: e_2 から p_1 に直交する成分 f_2 を取り出して, f_2 の長さを 1 にするように p_2 を定める.

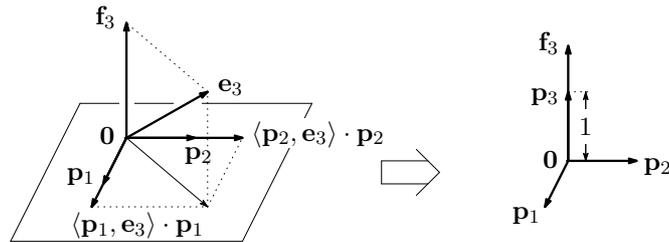


図 4: e_3 から p_1, p_2 に直交する成分 f_3 を取り出して, f_3 の長さを 1 にするように p_3 を定める.

と定めると, f_2 は p_1 と直交することが分かりますから, 後は, f_2 の長さを 1 にすることを考えて,

$$p_2 = \frac{f_2}{\|f_2\|}$$

と定めることにします (図 3 を参照). こうして, $\{p_1, e_2, e_3\}$ の代わりに, $\{p_1, p_2, e_3\}$ という基底を考えることにすれば, p_1, p_2 の長さは 1 で互いに直交することが分かります.

最後に, e_3 を p_1, p_2 と直交するように取り直すために, ベクトル e_3 を「 p_1 方向の成分」と「 p_2 方向の成分」と「 p_1, p_2 に直交する方向の成分」 f_3 に分解することを考えると, 前と同様にして,

$$e_3 = \langle p_1, e_3 \rangle \cdot p_1 + \langle p_2, e_3 \rangle \cdot p_2 + f_3$$

というように分解することが分かります. そこで,

$$f_3 = e_3 - \langle p_1, e_3 \rangle \cdot p_1 - \langle p_2, e_3 \rangle \cdot p_2 \quad (91)$$

と定めると, f_3 は p_1, p_2 と直交することが分かりますから, 後は, f_3 の長さを 1 にすることを考えて,

$$p_3 = \frac{f_3}{\|f_3\|}$$

と定めることにします (図 4 を参照). こうして, $\{p_1, p_2, e_3\}$ の代わりに, $\{p_1, p_2, p_3\}$ という基底を考えることにすれば, p_1, p_2, p_3 は長さが 1 で互いに直交することが分かりますから, めでたく, $\{p_1, p_2, p_3\}$ という正規直交基底が得られることが分かります. 一般に, 内積を持つ線型空間 V の勝手な基底 $\{e_1, e_2, \dots, e_n\}$ から始めて, 上のような形で正規直交基底 $\{p_1, p_2, \dots, p_n\}$ を構成する方法を Gram-Schmidt の直交化法と呼びます.

以上から, 「直交行列による対称行列の対角化の問題」を具体的に解決するためには,

「直交行列による対称行列の対角化の問題」を解決するための戦略

(i) 特性多項式 $\varphi_A(t) = \det(tI - A)$ を計算して, $\varphi_A(\lambda) = 0$ の解 $\lambda \in \mathbb{R}$ をすべて求める. (\implies 対称行列 A の固有値が求まる.)

(ii) それぞれの固有値 $\lambda \in \mathbb{R}$ に対して,

$$(A - \lambda I)\mathbf{u} = \mathbf{0}$$

という連立一次方程式を解いて,

$$V(\lambda) = \{ \mathbf{u} \in \mathbb{R}^n \mid A\mathbf{u} = \lambda\mathbf{u} \}$$

を求める. (\implies 対称行列 A の固有ベクトルが求まる.)

(iii) それぞれの固有ベクトル空間 $V(\lambda)$ の基底を勝手に一組ずつ取ってきて, 必要ならば, Gram-Schmidt の直交化法を用いて, それらの基底を正規直交基底に置き換える. さらに, それらの正規直交基底の元をすべて並べることで, 直交行列 P を作る.

という戦略にもとづいて計算を進めれば良いということになります.

ここでは「ユークリッド内積を持つ \mathbb{R} 上の線型空間」をもとにして, 「直交行列による対称行列の対角化の問題」について説明しましたが, 「エルミート内積を持つ \mathbb{C} 上の線型空間」をもとにして, 全く同様の考察を行なうことができます. すなわち, n 行 n 列のエルミート行列 A が, 勝手にひとつ与えられたときに, 上と同様の議論を行なうことで, 数ベクトル空間 \mathbb{C}^n が

エルミート行列 A に関する \mathbb{C}^n の固有ベクトル空間分解

$$\mathbb{C}^n = V(\lambda_1) \oplus V(\lambda_2) \oplus \cdots \oplus V(\lambda_L)$$

というように, エルミート行列 A の固有ベクトル空間の直和に分解することが分かります. また, それぞれの固有ベクトル空間 $V(\lambda_i)$ の正規直交基底を勝手に一組ずつ取ってきて, それらの基底の元をすべて集めたものを列ベクトルとする行列

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

を考えると, P はユニタリー行列となり, $P^{-1}AP$ は対角行列となることが分かります. すなわち, エルミート行列 A はユニタリー行列 P を用いて対角化できることが分かります. 興味のある方は, 対称行列に対する議論を参考にして, これらの事実を確かめてみて下さい.

8 問1を見直すと

さて, 問1では, 「直交行列による対称行列の対角化」という問題が自然に現われてくるような例として, 「臨界点の近くでの多変数関数の様子を調べる」という問題を取り上げました. そこで, ここでは, 臨界点の近くでの多変数関数の様子を調べるということについ

て少し考えてみようと思います。話を具体的にするために、以下では、問1で考えたような \mathbb{R}^3 上の関数の場合に説明してみることにします。

そこで、いま、 \mathbb{R}^3 上の関数 $f(x, y, z)$ が、勝手にひとつ与えられているとして、関数 f の様子を調べることを考えてみます。皆さんは、こうした事柄を微積分学の中で学んでいるわけですが、微積分学において、関数の様子を調べるときの基本的な考え方として、関数 f の様子を直接調べるのではなく、より簡単に様子が分かるような関数で近似して調べることがあります。例えば、 \mathbb{R}^3 上の関数の中で、多項式関数は比較的理解が容易な関数ですが、一般の滑らかな関数³¹を多項式関数で近似することを考えるということが Taylor 展開の基本的な考え方であるわけです。

Taylor 展開によれば、点 $p_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ を、勝手にひとつ取ってきたときに、点 $p = (x, y, z) \in \mathbb{R}^3$ が点 p_0 に十分近ければ、関数 f は、

関数 $f(x, y, z)$ の $p_0 \in \mathbb{R}^3$ のまわりでの Taylor 展開

$$f(x, y, z) \doteq f(p_0) + \frac{\partial f}{\partial x}(p_0)(x - x_0) + \frac{\partial f}{\partial y}(p_0)(y - y_0) + \frac{\partial f}{\partial z}(p_0)(z - z_0) + \dots$$

というように見えるという記述が得られるのでした。³² 特に、Taylor 展開の一次式の項までを考えると、

Taylor 展開を用いて、関数 f を一次式で近似してみる

$$f(x, y, z) - f(x_0, y_0, z_0) \doteq \frac{\partial f}{\partial x}(p_0)(x - x_0) + \frac{\partial f}{\partial y}(p_0)(y - y_0) + \frac{\partial f}{\partial z}(p_0)(z - z_0) \quad (92)$$

という近似が得られます。このように Taylor 展開の一次式の項までを考えた近似を用いて関数 f の様子を調べるということは、幾何学的には、関数 f のグラフに (三次元の) 接平面を描いて、その接平面の様子を調べることで関数 f の様子を調べるということを意味しています。皆さんは、高校以来、関数のグラフに接線を描き、その接線の様子を調べることで一変数関数の様子を調べることができるという考え方に慣れ親しんできたのではないかと思います。Taylor 展開の一次式の項までを考えた近似を取って調べるということは、こうしたアイデアの多変数関数の場合への自然な一般化になっています。

そこで、(92) 式の右辺を「線型写像」として解釈するために、点 $p_0 \in \mathbb{R}^3$ は、勝手にひとつ固定して考えることにして、 \mathbb{R}^3 の原点を点 p_0 までずらしたような線型空間 $T_{p_0}\mathbb{R}^3$ を考えてみます。すなわち、 $T_{p_0}\mathbb{R}^3$ とは、点 p_0 を始点に持つようなベクトル全体の集合です (図5を参照)。実は、 \mathbb{R}^3 のような「真っ直ぐな空間」上だけでなく、球面などの「曲がった空間」上でも微積分学を拡張して考えることができるのですが、こうした「曲がった空間上の微積分学」を扱う多様体論の言葉では、 $T_{p_0}\mathbb{R}^3$ のような線型空間を「点 p_0 における \mathbb{R}^3 の接空間 (tangent space)」と呼び、接空間 $T_{p_0}\mathbb{R}^3$ に属するベクトル $v \in T_{p_0}\mathbb{R}^3$ を点 p_0 における \mathbb{R}^3 の接ベクトル (tangent vector) と呼んだりします。³³ そこで、点 $p_0 \in \mathbb{R}^3$

³¹すなわち、何度でも微分できる関数ということです。

³²ここで、スペースを節約するため、及び、式を見やすくするために、 $f(x_0, y_0, z_0)$ などを $f(p_0)$ などと表わすことにしました。以下でも、同様の記法を用いることにします。

³³多様体とは「微積分学を展開することができるような曲がった空間」のことです。

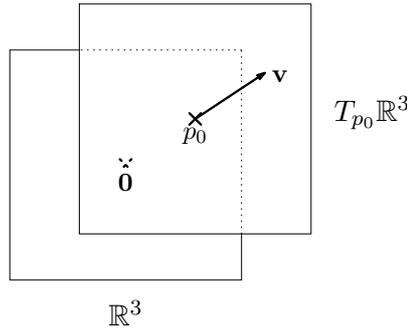


図 5: \mathbb{R}^3 の原点 0 を点 p_0 までずらしたような線型空間 \mathbb{R}^3 のコピー $T_{p_0}\mathbb{R}^3$ を考えてみる。ただし、イメージしやすいように、図では \mathbb{R}^3 などを二次元的に描いた。

における接ベクトル $\mathbf{v} \in T_{p_0}\mathbb{R}^3$ を,

$$\mathbf{v} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \in T_{p_0}\mathbb{R}^3 \cong \mathbb{R}^3$$

というように表わすことにして、 \mathbb{R}^3 の座標を (x, y, z) から (X, Y, Z) に取り替えて、(92) 式の右边を、

点 p_0 における関数 f の全微分

$$\begin{aligned} df_{p_0}(\mathbf{v}) &= \frac{\partial f}{\partial x}(p_0)X + \frac{\partial f}{\partial y}(p_0)Y + \frac{\partial f}{\partial z}(p_0)Z \\ &= \begin{pmatrix} \frac{\partial f}{\partial x}(p_0) & \frac{\partial f}{\partial y}(p_0) & \frac{\partial f}{\partial z}(p_0) \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \end{aligned}$$

というように表わすことにします。このとき、接空間 $T_{p_0}\mathbb{R}^3$ 上の線型関数

$$df_{p_0} : T_{p_0}\mathbb{R}^3 \rightarrow \mathbb{R}$$

を、点 p_0 における関数 f の全微分と呼びます。³⁴ いま、(92) 式の左辺は、

$$f(p) - f(p_0)$$

というように表わせることに注意して、関数 f の値域である \mathbb{R} の方も、 \mathbb{R} の原点を点 $f(p_0)$ までずらして、 \mathbb{R} の代わりに、点 $f(p_0)$ における \mathbb{R} の接空間 $T_{f(p_0)}\mathbb{R}$ を考えることにすると、点 p_0 における関数 f の全微分は、

関数 f の全微分は接空間の間の線型写像であると解釈できる

$$df_{p_0} : T_{p_0}\mathbb{R}^3 \rightarrow T_{f(p_0)}\mathbb{R} \cong \mathbb{R}$$

という線型写像として解釈できることとなります (図 6 を参照)。こうした記号を用いる

³⁴いま、点 $p_0 \in \mathbb{R}^3$ は勝手にひとつ固定して考えているので、 $\frac{\partial f}{\partial x}(p_0)$, $\frac{\partial f}{\partial y}(p_0)$, $\frac{\partial f}{\partial z}(p_0) \in \mathbb{R}$ は単なる数であることに注意して下さい。

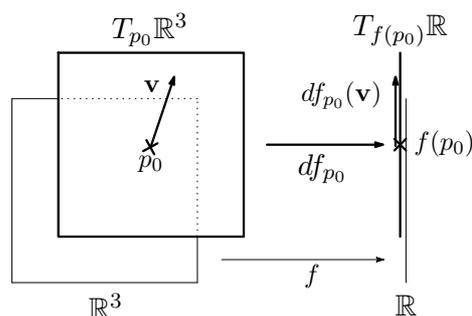


図 6: 点 p_0 での関数 f の全微分 df_{p_0} は, $T_{p_0}\mathbb{R}^3$ から $T_{f(p_0)}\mathbb{R}$ への線型写像と解釈できる.

と, (92) 式は,

Taylor 展開の一次式の項までを考えた近似式的全微分の記号を用いた表示

$$f(p_0 + \mathbf{v}) \doteq f(p_0) + df_{p_0}(\mathbf{v})$$

というように簡明な形で表わすことができます.

さて, 一変数関数に対して, 微分の値が消えるような点, すなわち, 接線の傾きが 0 になるような点に注目することで, 関数の大まかな様子を理解することができるということは, 皆さん良くご存じのことではないかと思います. このような点では, 例えば, 単調減少から単調増加に変わるというように, 関数の様子が大きく変わったりするので, 関数の大まかな様子を調べるときには, 特に重要な点であると考えられるわけです. こうした「微分が消える点」を, Taylor 展開の一次式の項である全微分が消えるという形で, 多変数関数の場合に拡張したものが臨界点という概念です. すなわち, ここで考えている \mathbb{R}^3 上の関数の場合であれば,

$$df_{p_0} = 0$$

となるような点 $p_0 \in \mathbb{R}^3$ を関数 f の臨界点と呼びます. 臨界点とは, Taylor 展開の一次式の項が消えるような点のことですから, 一次式の項までを考えた近似の範囲では, どの方向に進んでも関数の値が変わらないような点のことであると理解することができます. また, 幾何学的には, 「関数 f のグラフにおける (一般に高次元の) 接平面の傾きが零であるような点」であると理解することができます. 例えば, 問 1 の例では, (1) で見たように, 点 $p_0 = (0, 1, 2) \in \mathbb{R}^3$ が関数 f の臨界点になっていることが分かります. こうした臨界点は, 多変数関数の大まかな様子を調べるときにも, やはり重要な役を果たすことになります.

さて, 臨界点では, Taylor 展開の一次式の項は消えてしまいますから, 一次式の項までを考えた近似式からは, 関数 f の値は変わらないという情報しか得ることができません. そこで, もう少し詳しく関数 f の情報を得るためには, Taylor 展開の二次式の項までを考えた近似式を考察する必要があります.³⁵ この二次式の部分は, 一次式の部分と同様に, 行列の積の形に書き換えることができ, こうした書き換えをした形で表わすと, 関数 f は, 臨界点 p_0 の近くで,

³⁵ それでも駄目なら, さらに高次の項までを考えた近似式を考察する必要があります.

Taylor 展開を用いて、臨界点 p_0 のまわりで関数 f を二次式で近似してみる

$$f(p_0 + \mathbf{v}) \doteq f(p_0) + \frac{1}{2} \begin{pmatrix} X & Y & Z \end{pmatrix} \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p_0) & \frac{\partial^2 f}{\partial x \partial y}(p_0) & \frac{\partial^2 f}{\partial x \partial z}(p_0) \\ \frac{\partial^2 f}{\partial y \partial x}(p_0) & \frac{\partial^2 f}{\partial y^2}(p_0) & \frac{\partial^2 f}{\partial y \partial z}(p_0) \\ \frac{\partial^2 f}{\partial z \partial x}(p_0) & \frac{\partial^2 f}{\partial z \partial y}(p_0) & \frac{\partial^2 f}{\partial z^2}(p_0) \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (93)$$

というように見えるということになります。ここで、後の議論が見やすくなるように、前と同様に、 $(x, y, z) \in \mathbb{R}^3$ という座標の代わりに、

$$\mathbf{v} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \in T_{p_0} \mathbb{R}^3 \cong \mathbb{R}^3$$

という座標を用いて表わすことにしました。この (93) 式の右辺に現われた

関数 f の臨界点 p_0 におけるヘッセ行列

$$H_f(p_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(p_0) & \frac{\partial^2 f}{\partial x \partial y}(p_0) & \frac{\partial^2 f}{\partial x \partial z}(p_0) \\ \frac{\partial^2 f}{\partial y \partial x}(p_0) & \frac{\partial^2 f}{\partial y^2}(p_0) & \frac{\partial^2 f}{\partial y \partial z}(p_0) \\ \frac{\partial^2 f}{\partial z \partial x}(p_0) & \frac{\partial^2 f}{\partial z \partial y}(p_0) & \frac{\partial^2 f}{\partial z^2}(p_0) \end{pmatrix}$$

という行列を、関数 f の臨界点 p_0 におけるヘッセ行列と呼びます。何度でも微分できるような関数に対しては、偏導関数は微分する変数の順番には依りませんから、このような関数に対しては、ヘッセ行列は対称行列になるということに注意して下さい。ヘッセ行列を用いると、(93) 式は、

(93) 式のヘッセ行列を用いた表示

$$f(p_0 + \mathbf{v}) \doteq f(p_0) + \frac{1}{2} {}^t \mathbf{v} \cdot H_f(p_0) \cdot \mathbf{v} \quad (94)$$

というように簡明な形に表わすことができます。

例えば、問 1 の例では、問 1 の (2) の解答で見たように、臨界点 $p_0 = (0, 1, 2)$ における関数 f のヘッセ行列は、

$$H_f(p_0) = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix}$$

となることが分かりますから、

$$f(p_0) = f(0, 1, 2) = 1$$

であることに注意すると、関数 f は、臨界点 p_0 のまわりで、

$$f(p_0 + \mathbf{v}) \doteq 1 + \frac{1}{2} \begin{pmatrix} X & Y & Z \end{pmatrix} \begin{pmatrix} 3 & -2 & 0 \\ -2 & 2 & -2 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

$$= 1 + \frac{3}{2}X^2 - 2XY + Y^2 - 2YZ + \frac{1}{2}Z^2 \quad (95)$$

というような二次関数に見えるということになります。ところが、このままの形では、 $-2XY$ や $-2YZ$ などの項が登場しているために、(95) 式の右辺の二次関数がどのような関数であるのかということが余りハッキリとはしません。そこで、(95) の右辺に現われる二次関数を「より見やすい形」で表示することを念頭において、線型空間 $T_{p_0}\mathbb{R}^3 \cong \mathbb{R}^3$ 上の座標を取り替えることで、(95) の右辺の二次関数がどのように姿を変えるのかということを考えてみることにします。

そこで、いま、3 行 3 列の正則行列 P を、勝手にひとつ取ってきて、

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = P \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} \quad (96)$$

という式によって、 $T_{p_0}\mathbb{R}^3 \cong \mathbb{R}^3$ 上の新しい座標

$$\mathbf{v}' = \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix}$$

を定めてみます。すなわち、

$$\mathbf{v} = P\mathbf{v}' \quad (97)$$

という式により、線型空間 $T_{p_0}\mathbb{R}^3$ の「番地割り」を「旧番地」 $\mathbf{v} \in (\mathbb{R}^3)_{\text{旧}}$ から「新番地」 $\mathbf{v}' \in (\mathbb{R}^3)_{\text{新}}$ に取り替えることを考えてみます。³⁶ そこで、この新しい座標を用いて、(94) 式を書き直してみると、(97) 式を (94) 式に代入することで、

$$f(p_0 + \mathbf{v}') \doteq f(p_0) + \frac{1}{2} {}^t\mathbf{v}' \cdot {}^tPH_f(p_0)P \cdot \mathbf{v}' \quad (98)$$

となることが分かります。すなわち、(96) 式のような座標変換のもとで、ヘッセ行列 $H_f(p_0)$ は、

ヘッセ行列 $H_f(p_0)$ の変換公式

$$H_f(p_0) \rightsquigarrow {}^tPH_f(p_0)P$$

というように「姿」を変えることが分かります。

ここで、

$$H_f(p_0) \rightsquigarrow {}^tPH_f(p_0)P$$

という「ヘッセ行列の変換公式」は、「行列の標準形の問題」のときに考えた

$$H_f(p_0) \rightsquigarrow P^{-1}H_f(p_0)P$$

³⁶ここで、(94) 式の右辺に代入しやすくなるように、 $\mathbf{v} = P\mathbf{v}'$ という式によって、「新番地」 $\mathbf{v}' \in (\mathbb{R}^3)_{\text{新}}$ を定めるといふ言い方をしました。もちろん、これは、

$$\mathbf{v}' = P^{-1}\mathbf{v}$$

という式によって、「新番地」 $\mathbf{v}' \in (\mathbb{R}^3)_{\text{新}}$ を定めると言っても同じことです。

という「線型写像の表現行列の変換公式」と似ていますが、 P^{-1} の部分が tP に置き換わっているという違いがあります。ところが、上で注意したように、ヘッセ行列 $H_f(p_0)$ は対称行列でしたから、7節で見たように、対称行列 $H_f(p_0)$ を対角化するような正則行列 P として直交行列を取ることができることが分かります。すると、直交行列 P に対しては、

$${}^tP = P^{-1}$$

となることが分かりますから、大変都合の良いことに、

$${}^tPH_f(p_0)P = P^{-1}H_f(p_0)P$$

となることが分かります。したがって、ヘッセ行列 $H_f(p_0)$ を対角化するような直交行列 P を用いることで、 ${}^tPH_f(p_0)P$ という行列も対角行列に変形できることが分かります。

例えば、問1の例では、問1の(3)の解答で見たように、

$$P = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ 2 & -2 & 1 \end{pmatrix} \quad (99)$$

として、

$$\begin{aligned} {}^tPH_f(p_0)P &= P^{-1}H_f(p_0)P \\ &= \begin{pmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 5 \end{pmatrix} \end{aligned}$$

というように対角化できることが分かります。よって、(99)式で与えられる直交行列 P を用いて、(96)式のような座標変換をしてみると、関数 f は、臨界点 $p_0 \in \mathbb{R}^3$ の近くで、

$$f(p_0 + \mathbf{v}') \doteq 1 - \frac{1}{2}X'^2 + Y'^2 + \frac{5}{2}Z'^2 \quad (100)$$

というような二次関数に見えるということが分かります。ここで、(100)式では、 $X'Y'$ などの項が現われないために、(95)式と比較して、二次関数の様子がとても理解しやすい形になっていることに注意して下さい。

さて、一変数関数 $f(x)$ の場合には、臨界点 $x_0 \in \mathbb{R}$ における二階微分の値 $f''(x_0)$ の符号を見ることで、その臨界点での値が極大値であるか、極小値であるかということを判定できるということは、皆さん良くご存じのことだと思います。これを、Taylor 展開という立場から見直すと、Taylor 展開の二次式の項までを考えて、臨界点 x_0 の近くで、関数 f を、

Taylor 展開を用いて、臨界点のまわりで関数 f を二次式で近似してみる

$$f(x_0 + X) \doteq f(x_0) + \frac{f''(x_0)}{2}X^2 \quad (101)$$

というように近似して調べていると理解することができます。すなわち、 $f''(x_0) > 0$ のときには、(101)式の右辺に現われる二次関数のグラフは下に凸となるので、 $f(x_0)$ は極小値

になり、 $f''(x_0) < 0$ のときには、(101) 式の右辺に現われる二次関数のグラフは上に凸となるので、 $f(x_0)$ は極大値になると理解することができます。

多変数関数の場合にも、全く同様にして、極値判定の方法を考えることができます。ただし、上で見たように、多変数関数の場合には、必ずしも最初に与えられた座標系に対して、臨界点 p_0 におけるヘッセ行列 $H_f(p_0)$ が「見やすい形」になるとは限らないので、 ${}^tPH_f(p_0)P$ という行列が対角行列となるような正則行列 P を用いて、(96) 式のように座標変換して考える必要があります。例えば、ここで考えたように、関数 f が三変数関数であるとして、関数 f の臨界点 $p_0 \in \mathbb{R}^3$ におけるヘッセ行列が、3 行 3 列の正則行列 P を用いて、

$${}^tPH_f(p_0)P = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

というように対角化されると仮定すると、(96) 式のように座標変換することで、

関数 f の臨界点 p_0 の周りでの様子を「見やすい形」で記述する

$$f(p_0 + \mathbf{v}') \doteq f(p_0) + \frac{\lambda_1}{2}X'^2 + \frac{\lambda_2}{2}Y'^2 + \frac{\lambda_3}{2}Z'^2 \quad (102)$$

となることが分かります。すると、例えば、 $\lambda_1, \lambda_2, \lambda_3 > 0$ であるとする、(102) 式から、どの方向に進んでも関数 f の値が増えることが分かりますから、 $f(p_0)$ は極小値であることが分かります。また、 $\lambda_1, \lambda_2, \lambda_3 < 0$ であるとする、(102) 式から、どの方向に進んでも関数 f の値が減ることが分かりますから、 $f(p_0)$ は極大値になることが分かります。さらに、 $\lambda_1, \lambda_2 > 0, \lambda_3 < 0$ というように、符号の異なる固有値が登場する場合には、(102) 式から、関数 f の値が増える方向と値が減る方向が存在するので、 $f(p_0)$ は極値ではないことが分かります。

このように、多変数関数の場合には、臨界点におけるヘッセ行列 $H_f(p_0)$ の固有値の符号によって極値判定ができることとなります。ただし、いくつかの固有値が 0 となる場合、すなわち、 $\det H_f(p_0) = 0$ となる場合には、固有値 0 に対応した方向では「二次式の項までを考えた近似の範囲では関数 f の値は変わらない」ということしか分からないので、より詳しく関数の様子を調べるためには、さらに高い次数の項までを考えた近似式を用いて調べる必要があります。³⁷

9 多変数関数の合成関数に対する連鎖律について

さて、8 節では、 \mathbb{R}^3 上の関数 f に対して、点 $p_0 \in \mathbb{R}^3$ における関数 f の全微分 df_{p_0} を、

$$df_{p_0} : T_{p_0}\mathbb{R}^3 \rightarrow T_{f(p_0)}\mathbb{R}$$

というように接空間の間の線型写像として解釈しましたが、より一般に、ユークリッド空間の間の写像に対しても、同様の解釈を行なうことができます。話を具体的にするために、以下では、 \mathbb{R}^3 から \mathbb{R}^3 への写像

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

³⁷一変数関数の場合でも、 $f''(x_0) = 0$ となる場合には、このままでは極値判定はできないので、より詳しく関数の様子を調べるためには、さらに高い次数の項までを考えた近似式を用いて調べる必要があるわけです。

が、勝手にひとつ与えられているとして、点 $p_0 = (x_0, y_0, z_0) \in \mathbb{R}^3$ を、勝手にひとつ取ってきたときに、写像 f の点 p_0 の近くでの大まかな様子を調べることを考えてみることにします。

いま、 $(x, y, z) \in \mathbb{R}^3$ に対して、 $f(x, y, z) \in \mathbb{R}^3$ の成分を、

写像 f を成分を用いて記述する

$$f(x, y, z) = \begin{pmatrix} X(x, y, z) \\ Y(x, y, z) \\ Z(x, y, z) \end{pmatrix} \in \mathbb{R}^3$$

というように表わすことにします。すると、それぞれの成分 $X(x, y, z), Y(x, y, z), Z(x, y, z)$ は単なる関数ですから、それぞれの関数に対して、Taylor 展開の一次式の項までを考えた近似式を取ることで、写像 f のそれぞれの成分は、点 $p_0 \in \mathbb{R}^3$ のまわりで、

写像 f のそれぞれの成分を Taylor 展開を用いて記述してみる

$$\begin{cases} X(x, y, z) - X(p_0) \doteq \frac{\partial X}{\partial x}(p_0) \cdot (x - x_0) + \frac{\partial X}{\partial y}(p_0) \cdot (y - y_0) + \frac{\partial X}{\partial z}(p_0) \cdot (z - z_0) \\ Y(x, y, z) - Y(p_0) \doteq \frac{\partial Y}{\partial x}(p_0) \cdot (x - x_0) + \frac{\partial Y}{\partial y}(p_0) \cdot (y - y_0) + \frac{\partial Y}{\partial z}(p_0) \cdot (z - z_0) \\ Z(x, y, z) - Z(p_0) \doteq \frac{\partial Z}{\partial x}(p_0) \cdot (x - x_0) + \frac{\partial Z}{\partial y}(p_0) \cdot (y - y_0) + \frac{\partial Z}{\partial z}(p_0) \cdot (z - z_0) \end{cases} \quad (103)$$

というように記述できることが分かります。そこで、前と同様に、

$$\mathbf{v} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} \in T_{p_0}\mathbb{R}^3 \cong \mathbb{R}^3$$

として、(103) 式の右辺を、

$$\begin{aligned} df_{p_0}(\mathbf{v}) &= \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) \cdot X + \frac{\partial X}{\partial y}(p_0) \cdot Y + \frac{\partial X}{\partial z}(p_0) \cdot Z \\ \frac{\partial Y}{\partial x}(p_0) \cdot X + \frac{\partial Y}{\partial y}(p_0) \cdot Y + \frac{\partial Y}{\partial z}(p_0) \cdot Z \\ \frac{\partial Z}{\partial x}(p_0) \cdot X + \frac{\partial Z}{\partial y}(p_0) \cdot Y + \frac{\partial Z}{\partial z}(p_0) \cdot Z \end{pmatrix} \\ &= \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) & \frac{\partial X}{\partial z}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) & \frac{\partial Y}{\partial z}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) & \frac{\partial Z}{\partial z}(p_0) \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \end{aligned} \quad (104)$$

というように表わすことにすると、 df_{p_0} は、

点 $p_0 \in \mathbb{R}^3$ における写像 f の微分写像

$$df_{p_0} : T_{p_0}\mathbb{R}^3 \rightarrow T_{f(p_0)}\mathbb{R}^3$$

という線型写像として解釈することができます。この線型写像 df_{p_0} を点 p_0 における写像 f の微分写像と呼びます。また、(104) 式から、

$$T_{p_0}\mathbb{R}^3 \cong \mathbb{R}^3,$$

$$T_{f(p_0)}\mathbb{R}^3 \cong \mathbb{R}^3$$

という同一視のもとで,³⁸ 微分写像 df_{p_0} は,

$p_0 \in \mathbb{R}^3$ における写像 f の Jacobi 行列

$$J_f(p_0) = \begin{pmatrix} \frac{\partial X}{\partial x}(p_0) & \frac{\partial X}{\partial y}(p_0) & \frac{\partial X}{\partial z}(p_0) \\ \frac{\partial Y}{\partial x}(p_0) & \frac{\partial Y}{\partial y}(p_0) & \frac{\partial Y}{\partial z}(p_0) \\ \frac{\partial Z}{\partial x}(p_0) & \frac{\partial Z}{\partial y}(p_0) & \frac{\partial Z}{\partial z}(p_0) \end{pmatrix}$$

という行列で表現されるということになります. この行列 $J_f(p_0)$ を点 p_0 における写像 f の Jacobi 行列と呼んだりします. 微分写像を用いると, 関数のときと同様に, (103) 式は,

微分写像を用いた (103) 式の表示

$$f(p_0 + \mathbf{v}) \doteq f(p_0) + df_{p_0}(\mathbf{v}) \quad (105)$$

というように簡明な形で表わすことができます.

こうした微分写像の概念を用いると, 次のように, 多変数関数に対する合成関数の微分則を簡明な形で理解することができます. そこで, いま,

$$f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

という写像とは別に,

$$g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

という写像がもうひとつ与えられているとします. このとき,

$$f \circ g: \mathbb{R}^2 \rightarrow \mathbb{R}^3$$

という f と g の合成写像を考えることができますが, f の代わりに, 合成写像 $f \circ g$ に対して, (105) 式を適用してみると, $q_0 \in \mathbb{R}^2$, $\mathbf{u} \in T_{q_0}\mathbb{R}^2$ として,

$$(f \circ g)(q_0 + \mathbf{u}) \doteq (f \circ g)(q_0) + d(f \circ g)_{q_0}(\mathbf{u}) \quad (106)$$

となることが分かります. 一方, f の代わりに, 写像 g に対して (105) 式を適用してみると,

$$g(q_0 + \mathbf{u}) \doteq g(q_0) + dg_{q_0}(\mathbf{u}) \quad (107)$$

となることが分かりますが,

$$p_0 = g(q_0), \quad \mathbf{v} = dg_{q_0}(\mathbf{u}) \quad (108)$$

として, (107) 式を (105) 式に代入してみると,

$$\begin{aligned} (f \circ g)(q_0 + \mathbf{u}) &= f(g(q_0 + \mathbf{u})) \\ &\doteq f(g(q_0) + dg_{q_0}(\mathbf{u})) \end{aligned} \quad ((107) \text{ 式から})$$

³⁸すなわち, $T_{p_0}\mathbb{R}^3$ や $T_{f(p_0)}\mathbb{R}^3$ を, 線型空間 \mathbb{R}^3 を点 $p_0 \in \mathbb{R}^3$ まで平行移動したもの, あるいは, 線型空間 \mathbb{R}^3 を点 $f(p_0) \in \mathbb{R}^3$ まで平行移動したものであると考えて「番地割り」ということです.

$$\begin{aligned}
&= f(p_0 + \mathbf{v}) && \text{((108) 式から)} \\
&\doteq f(p_0) + df_{p_0}(\mathbf{v}) && \text{((105) 式から)} \\
&= f(g(q_0)) + df_{g(q_0)}(dg_{q_0}(\mathbf{u})) && \text{((108) 式から)} \\
&= (f \circ g)(q_0) + (df_{g(q_0)} \circ dg_{q_0})(\mathbf{u}) && \text{(109)}
\end{aligned}$$

となることが分かります。したがって、(106) 式と (109) 式を比べることで、

合成関数に対する微分則 (線型写像の立場から連鎖律を見直したもの)

$$d(f \circ g)_{q_0} = df_{g(q_0)} \circ dg_{q_0} \quad (110)$$

となることが分かります。³⁹ この (110) 式は、合成関数の微分写像は、それぞれの関数の微分写像の合成写像であるということを意味していますが、それぞれの微分写像を Jacobi 行列を用いて表わしてみると、

合成関数に対する連鎖律 ((110) 式を Jacobi 行列を用いて表わしたもの)

$$J_{f \circ g}(q_0) = J_f(g(q_0)) \cdot J_g(q_0) \quad (111)$$

ということになります。この (111) 式の等式を、それぞれの行列の行列成分を用いて表わしたものが、多変数関数の合成関数に対する微分則である連鎖律 (chain rule) に他なりません。興味のある方は、(111) 式の両辺に現われる行列の行列成分を求めてみて下さい。このように、微分写像という概念を用いることで、連鎖律を自然な形で理解できることが分かります。

さて、8 節の最初でも注意したように、Taylor 展開の一次式の項である微分写像を調べることで、関数や写像の大まかな様子を調べることができるのではないかとということが、微積分学における基本的な考え方なのですが、こうした期待が正しいことを裏付けるような定理がいくつか知られています。その代表的なものとして、陰関数定理や逆関数定理があります。例えば、逆関数定理とは、ここで考えたような写像

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$$

に対して、点 $p_0 \in \mathbb{R}^3$ における写像 f の微分写像

$$df_{p_0} : T_{p_0}\mathbb{R}^3 \rightarrow T_{f(p_0)}\mathbb{R}^3$$

が全単射であれば、もとの写像 f も点 p_0 の近くでは全単射になっているということを、したがって、点 $f(p_0) \in \mathbb{R}^3$ のまわりで局所的に写像 f の逆写像 f^{-1} を考えることができるということを主張する定理です。皆さんも、いずれこうした定理を学ぶことになると思いますが、こうした定理の存在は、関数や写像の大まかな様子を調べるときに、微分写像の様子を理解することがとても大切であるということを意味しています。皆さんが現在学ばれている線型代数学の知識は、8 節で述べたようなヘッセ行列の対角化に役立つだけでなく、こうした微分写像をより良く理解することを通して、微積分学全体の理解を助けてくれることとなります。

³⁹細かいことを言えば、この結論を得るために、Taylor 展開の一意性を用いています。

10 問2の解答

(1) いま, 2行2列の複素行列 X を, $a, b, c, d \in \mathbb{C}$ として,

$$X = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

というように表わすことにすると,

$${}^t\bar{X} = \begin{pmatrix} \bar{a} & \bar{c} \\ \bar{b} & \bar{d} \end{pmatrix}$$

となることが分かりますから,

$${}^t\bar{X} = X \iff \begin{cases} \bar{a} = a, \\ \bar{b} = c, \\ \bar{d} = d \end{cases} \quad (112)$$

となることが分かります. また,

$$\operatorname{tr} X = a + d$$

となることが分かりますから,

$$\operatorname{tr} X = 0 \iff a + d = 0 \quad (113)$$

となることが分かります. よって, (112) 式, (113) 式から,

$$X \in V \iff \begin{cases} a = -d \in \mathbb{R}, \\ b = \bar{c} \in \mathbb{C} \end{cases}$$

となることが分かります. そこで, $x, y, z \in \mathbb{R}$ として,

$$\begin{cases} b = \bar{c} = x + \sqrt{-1}y \\ a = -d = z \end{cases}$$

と表わすことにすれば, $X \in V$ は,

$$X = \begin{pmatrix} z & x + \sqrt{-1}y \\ x - \sqrt{-1}y & -z \end{pmatrix}$$

というように表わせることが分かります.

(2) $\{e_1, e_2, e_3\}$ が V の基底であることを示すためには,

(イ) 勝手な元 $X \in V$ に対して,

$$X = xe_1 + ye_2 + ze_3$$

となるような実数 $x, y, z \in \mathbb{R}$ が存在する.

(ロ) $x, y, z \in \mathbb{R}$ として,

$$\mathbf{0} = xe_1 + ye_2 + ze_3 \implies x = y = z = 0$$

となる.

という二つの条件が満たされることが確かめられればよいということになります.

そこで, まず, (イ) という条件について考えてみます. いま, (1) の結果から, 勝手な元 $X \in V$ は, $x, y, z \in \mathbb{R}$ として,

$$\begin{aligned} X &= \begin{pmatrix} z & x + \sqrt{-1}y \\ x - \sqrt{-1}y & -z \end{pmatrix} \\ &= x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & \sqrt{-1} \\ -\sqrt{-1} & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= xe_1 + ye_2 + ze_3 \end{aligned}$$

というように表わせることが分かりますから, (イ) という条件が満たされることが分かります. 次に, (ロ) という条件について考えてみます. いま, $x, y, z \in \mathbb{R}$ として,

$$\mathbf{0} = xe_1 + ye_2 + ze_3 \tag{114}$$

であると仮定してみます. すると, (114) 式から,

$$\begin{aligned} \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} &= x \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + y \begin{pmatrix} 0 & \sqrt{-1} \\ -\sqrt{-1} & 0 \end{pmatrix} + z \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} z & x + \sqrt{-1}y \\ x - \sqrt{-1}y & -z \end{pmatrix} \end{aligned}$$

となることが分かりますから,

$$x + \sqrt{-1}y = 0, \quad z = 0$$

となることが分かります. よって,

$$x = y = z = 0$$

となることが分かりますから, (ロ) という条件も満たされることが分かります. 以上から, (イ), (ロ) という二つの条件が満たされることが分かりましたから, $\{e_1, e_2, e_3\}$ が V の基底となることが分かります.

そこで, 内積の定義に戻って, 基底に属する元の間の内積の値を具体的に計算してみると, $i, j = 1, 2, 3$ に対して,

$$\langle e_i, e_j \rangle = \begin{cases} 1, & i = j \text{ のとき} \\ 0, & i \neq j \text{ のとき} \end{cases}$$

となることが分かります. したがって, $\{e_1, e_2, e_3\}$ は, V の正規直交基底になることが分かります. あるいは,

$$X = \begin{pmatrix} z & x + \sqrt{-1}y \\ x - \sqrt{-1}y & -z \end{pmatrix}, X' = \begin{pmatrix} z' & x' + \sqrt{-1}y' \\ x' - \sqrt{-1}y' & -z' \end{pmatrix} \in V$$

として, X と X' の間の内積 $\langle X, X' \rangle$ の値を具体的に計算してみると,

$$\begin{aligned} \langle X, X' \rangle &= \frac{1}{2} \operatorname{tr}(XX') \\ &= \frac{1}{2} \operatorname{tr} \left\{ \begin{pmatrix} z & x + \sqrt{-1}y \\ x - \sqrt{-1}y & -z \end{pmatrix} \begin{pmatrix} z' & x' + \sqrt{-1}y' \\ x' - \sqrt{-1}y' & -z' \end{pmatrix} \right\} \\ &= xx' + yy' + zz' \end{aligned}$$

となることが分かります. これより, $\{e_1, e_2, e_3\}$ という基底を用いて線型空間 V に「番地割り」をしてみると, 内積の構造まで込めて,

$$(V, \langle \cdot, \cdot \rangle) \cong (\mathbb{R}^3, \langle \cdot, \cdot \rangle_{\mathbb{R}^3})$$

というように同一視できることが分かりますが, このことから, $\{e_1, e_2, e_3\}$ が V の正規直交基底であることを結論しても構いません.⁴⁰

(3) 行列 A がユニタリー行列であること, すなわち,

$${}^t\bar{A}A = I \tag{115}$$

となることと, 一般に, 掛け算のできる二つの行列 B, C に対して,

$$\operatorname{tr}(BC) = \operatorname{tr}(CB) \tag{116}$$

となることに注意すると,

$$\begin{aligned} \langle T_A(X), T_A(Y) \rangle &= \frac{1}{2} \operatorname{tr}\{T_A(X)T_A(Y)\} \\ &= \frac{1}{2} \operatorname{tr}\{(AX{}^t\bar{A})(AY{}^t\bar{A})\} && (T_A \text{ の定義から}) \\ &= \frac{1}{2} \operatorname{tr}\{AX({}^t\bar{A}A)Y{}^t\bar{A}\} \\ &= \frac{1}{2} \operatorname{tr}(A(XY{}^t\bar{A})) && ((115) \text{ 式から}) \\ &= \frac{1}{2} \operatorname{tr}((XY{}^t\bar{A})A) && ((116) \text{ 式から}) \\ &= \frac{1}{2} \operatorname{tr}(XY) && ((115) \text{ 式から}) \\ &= \langle X, Y \rangle \end{aligned}$$

となることが分かります.

⁴⁰ここで, \mathbb{R}^3 上の標準的なユークリッド内積を $\langle \cdot, \cdot \rangle_{\mathbb{R}^3}$ という記号を用いて表わしました.

(4) 基底 $\{e_1, e_2, e_3\}$ に関する線型写像 $T_A: V \rightarrow V$ の表現行列 \hat{T}_A を求めるためには,

(イ) 基底の元の行き先 $T_A(e_1), T_A(e_2), T_A(e_3) \in V$ の「番地」を求める.

(\implies これらの「番地」を並べたものが表現行列 \hat{T}_A になる.)

(ロ) 基底 $\{e_1, e_2, e_3\}$ を用いた「番地割り」 $V \cong \mathbb{R}^3$ のもとで,

$$V \ni X = xe_1 + ye_2 + ze_3 \longleftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$$

と対応しているときに, $T_A(X) \in V$ の「番地」を,

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

を用いて表わす.

(\implies このとき,

$$V \ni T_A(X) \longleftrightarrow \hat{T}_A \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$$

と対応しているはず.)

という二つの方法を考えることができます. 表現行列 \hat{T}_A を求めるためには, どちらの方法を用いても構わないわけですが, 皆さんの参考のために, 以下では, それぞれの方法を用いて表現行列 \hat{T}_A を求めると, どのようなことになるのかということ順番に見てみることにします.

そこで, まず, (イ) という方法にもとづいて考えてみます. いま, 線型写像 T_A の定義にもとづいて, $T_A(e_1), T_A(e_2), T_A(e_3) \in V$ を求めてみると, それぞれ,

$$\begin{aligned} T_A(e_1) &= \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & e^{\sqrt{-1}\eta} \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\ &= \begin{pmatrix} 0 & e^{\sqrt{-1}\theta} \\ e^{\sqrt{-1}\eta} & 0 \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\ &= \begin{pmatrix} 0 & e^{\sqrt{-1}(\theta-\eta)} \\ e^{-\sqrt{-1}(\theta-\eta)} & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & \cos(\theta-\eta) + \sqrt{-1}\sin(\theta-\eta) \\ \cos(\theta-\eta) - \sqrt{-1}\sin(\theta-\eta) & 0 \end{pmatrix} \\ &= \cos(\theta-\eta)e_1 + \sin(\theta-\eta)e_2 + 0e_3 \tag{117} \\ T_A(e_2) &= \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & e^{\sqrt{-1}\eta} \end{pmatrix} \begin{pmatrix} 0 & \sqrt{-1} \\ -\sqrt{-1} & 0 \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} 0 & \sqrt{-1} e^{\sqrt{-1}\theta} \\ -\sqrt{-1} e^{\sqrt{-1}\eta} & 0 \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\
&= \begin{pmatrix} 0 & \sqrt{-1} e^{\sqrt{-1}(\theta-\eta)} \\ -\sqrt{-1} e^{-\sqrt{-1}(\theta-\eta)} & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & -\sin(\theta-\eta) + \sqrt{-1} \cos(\theta-\eta) \\ -\sin(\theta-\eta) - \sqrt{-1} \cos(\theta-\eta) & 0 \end{pmatrix} \\
&= -\sin(\theta-\eta)\mathbf{e}_1 + \cos(\theta-\eta)\mathbf{e}_2 + 0\mathbf{e}_3 \tag{118}
\end{aligned}$$

$$\begin{aligned}
T_A(\mathbf{e}_3) &= \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & e^{\sqrt{-1}\eta} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\
&= \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & -e^{\sqrt{-1}\eta} \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \\
&= 0\mathbf{e}_1 + 0\mathbf{e}_2 + 1\mathbf{e}_3 \tag{119}
\end{aligned}$$

となることが分かります. よって, (117) 式, (118) 式, (119) 式から, これらの元の「番地」は, それぞれ,

$$\begin{aligned}
V \ni T_A(\mathbf{e}_1) &\longleftrightarrow \begin{pmatrix} \cos(\theta-\eta) \\ \sin(\theta-\eta) \\ 0 \end{pmatrix} \in \mathbb{R}^3 \\
V \ni T_A(\mathbf{e}_2) &\longleftrightarrow \begin{pmatrix} -\sin(\theta-\eta) \\ \cos(\theta-\eta) \\ 0 \end{pmatrix} \in \mathbb{R}^3 \\
V \ni T_A(\mathbf{e}_3) &\longleftrightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \in \mathbb{R}^3
\end{aligned}$$

となることが分かりますから, 表現行列 \hat{T}_A は,

$$\hat{T}_A = \begin{pmatrix} \cos(\theta-\eta) & -\sin(\theta-\eta) & 0 \\ \sin(\theta-\eta) & \cos(\theta-\eta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

となることが分かります.

次に, (ロ) という方法にもとづいて考えてみます. いま,

$$V \ni X = x\mathbf{e}_1 + y\mathbf{e}_2 + z\mathbf{e}_3 \longleftrightarrow \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3$$

と対応しているとして、線型写像 T_A の定義にもとづいて、 $T_A(X) \in V$ を求めてみます。すると、

$$w = x + \sqrt{-1}y$$

として、

$$\begin{aligned} T_A(X) &= AX {}^t\bar{A} \\ &= \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & e^{\sqrt{-1}\eta} \end{pmatrix} \begin{pmatrix} z & w \\ \bar{w} & -z \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\ &= \begin{pmatrix} e^{\sqrt{-1}\theta}z & e^{\sqrt{-1}\theta}w \\ e^{\sqrt{-1}\eta}\bar{w} & -e^{\sqrt{-1}\eta}z \end{pmatrix} \begin{pmatrix} e^{-\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\eta} \end{pmatrix} \\ &= \begin{pmatrix} z & e^{\sqrt{-1}(\theta-\eta)}w \\ e^{-\sqrt{-1}(\theta-\eta)}\bar{w} & -z \end{pmatrix} \end{aligned}$$

となることが分かります。ここで、

$$\begin{aligned} e^{\sqrt{-1}(\theta-\eta)}w &= \{\cos(\theta-\eta) + \sqrt{-1}\sin(\theta-\eta)\} \cdot (x + \sqrt{-1}y) \\ &= \cos(\theta-\eta) \cdot x - \sin(\theta-\eta) \cdot y \\ &\quad + \sqrt{-1} \{\sin(\theta-\eta) \cdot x + \cos(\theta-\eta) \cdot y\} \end{aligned}$$

となることに注意すると、

$$\begin{aligned} (T_A)(X) &= \{\cos(\theta-\eta) \cdot x - \sin(\theta-\eta) \cdot y\} \mathbf{e}_1 \\ &\quad + \{\sin(\theta-\eta) \cdot x + \cos(\theta-\eta) \cdot y\} \mathbf{e}_2 + z \mathbf{e}_3 \end{aligned} \quad (120)$$

となることが分かります。よって、(120) 式から、

$$\begin{aligned} V \ni T_A(X) &\longleftrightarrow \begin{pmatrix} \cos(\theta-\eta) \cdot x - \sin(\theta-\eta) \cdot y \\ \sin(\theta-\eta) \cdot x + \cos(\theta-\eta) \cdot y \\ z \end{pmatrix} \\ &= \begin{pmatrix} \cos(\theta-\eta) & -\sin(\theta-\eta) & 0 \\ \sin(\theta-\eta) & \cos(\theta-\eta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \end{aligned}$$

と対応することが分かりますから、表現行列 \hat{T}_A は、

$$\hat{T}_A = \begin{pmatrix} \cos(\theta-\eta) & -\sin(\theta-\eta) & 0 \\ \sin(\theta-\eta) & \cos(\theta-\eta) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

となることが分かります。⁴¹ 第1回の問3のところで見たとように、これは、 z 軸まわりの $(\theta-\eta)$ 回転を表わす行列に他なりません。

⁴¹もちろん、これは、(イ) という方法にもとづいて求めた表現行列 \hat{T}_A と同じものです。

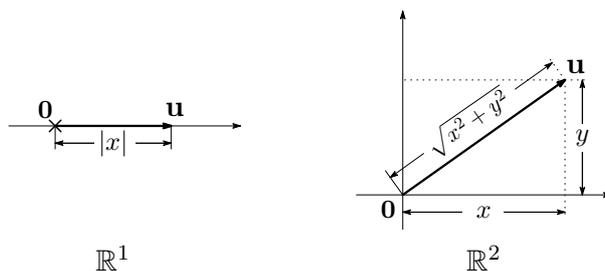


図 7: $n = 1, 2$ のとき, ベクトル $\mathbf{u} \in \mathbb{R}^n$ の長さ $\|\mathbf{u}\|$ は, それぞれ, $|x|$, $\sqrt{x^2 + y^2}$ で与えられる.

11 線型空間上の内積とは

皆さん良くご存じのように, \mathbb{R}^3 のようなユークリッド空間上では, ユークリッド内積を考えることにより, 「ベクトルの長さ」や「二つのベクトルの間の角度」といった概念を意味付けることができます. こうした内積の概念は, 一般の線型空間にも拡張して考えることができ, 内積をひとつ定めることによって, ユークリッド空間の場合と同様に, 「線型空間の元の長さ」や「二つの元の間の角度」といった概念を意味付けることができます. 皆さんに, こうした「内積を持つ線型空間」の例や「内積を保つ線型写像」の例などに触れてもらおうと思って, 問 2 を出題してみました.

そこで, 一般の線型空間上の内積について考えてみる前に, \mathbb{R}^n 上のユークリッド内積の持つ性質について少し反省してみることにします. そのために, $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ として, ベクトル \mathbf{u} の長さ $\|\mathbf{u}\|$ やベクトル \mathbf{u} とベクトル \mathbf{v} の間の角度が, \mathbb{R}^n の座標を用いてどのように表わすことができるのかということを考えてみることにします.

そこで, まず, ベクトル $\mathbf{u} \in \mathbb{R}^n$ の長さ $\|\mathbf{u}\|$ について考えてみます. すると, $n = 1$ のときには,

$$\mathbf{u} = \begin{pmatrix} x \end{pmatrix} \in \mathbb{R}^1$$

として, ベクトル \mathbf{u} の長さは,

$$\begin{aligned} \|\mathbf{u}\| &= |x| \\ &= \sqrt{x^2} \end{aligned}$$

という式で与えられることが分かります (図 7 を参照). また, $n = 2$ のときには, 「三平方の定理」から,

$$\mathbf{u} = \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2$$

として, ベクトル \mathbf{u} の長さは,

$$\|\mathbf{u}\| = \sqrt{x^2 + y^2}$$

という式で与えられることが分かります (図 7 を参照). さらに, $n = 3$ のときには,

$$\mathbf{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \mathbf{u}_1 = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \in \mathbb{R}^3$$

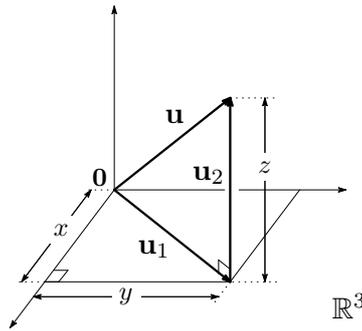


図 8: $n = 3$ のとき, ベクトル $\mathbf{u} \in \mathbb{R}^3$ の長さ $\|\mathbf{u}\|$ は $\sqrt{x^2 + y^2 + z^2}$ で与えられる.

として, \mathbb{R}^3 のベクトル \mathbf{u} を,

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$$

と分解してから, 「三平方の定理」を用いると, ベクトル \mathbf{u} の長さは,

$$\begin{aligned} \|\mathbf{u}\| &= \sqrt{\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2} \\ &= \sqrt{(x^2 + y^2) + z^2} \\ &= \sqrt{x^2 + y^2 + z^2} \end{aligned}$$

という式で与えられることが分かります (図 8 を参照). 全く同様にして, 勝手な自然数 $n \in \mathbb{N}$ に対して,

$$\mathbf{u} = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}, \mathbf{u}_1 = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ 0 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

として, \mathbb{R}^n のベクトル \mathbf{u} を,

$$\mathbf{u} = \mathbf{u}_1 + \mathbf{u}_2$$

と分解してから, 「三平方の定理」を用いると, ベクトル \mathbf{u} の長さは,

$$\begin{aligned} \|\mathbf{u}\| &= \sqrt{\|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2} \\ &= \sqrt{\|\mathbf{u}_1\|^2 + (x_n)^2} \end{aligned} \tag{121}$$

と表わせることが分かりますから, 後は, (121) 式と n に関する数学的帰納法を用いることで, 一般に, \mathbb{R}^n のベクトル \mathbf{u} の長さは,

$$\|\mathbf{u}\| = \sqrt{(x_1)^2 + (x_2)^2 + \cdots + (x_n)^2} \tag{122}$$

という式で与えられることが分かります.⁴²

⁴² 皆さん, 確かめてみて下さい.

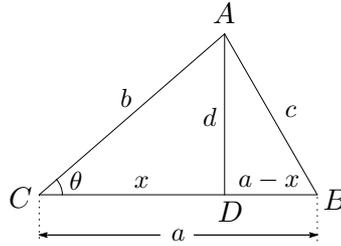


図 9: 三角形 $\triangle ABC$ の様子.

次に, \mathbb{R}^n の二つのベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して, ベクトル \mathbf{u} とベクトル \mathbf{v} の間の角度について考えてみます. いま, 平面上の三角形 $\triangle ABC$ に対して, 点 A から辺 BC に下ろした垂線の足を点 D とします. このとき, それぞれの線分の長さを,

$$\begin{aligned} |\overline{BC}| &= a, \quad |\overline{AC}| = b, \quad |\overline{AB}| = c, \\ |\overline{AD}| &= d, \quad |\overline{CD}| = x, \quad |\overline{BD}| = a - x \end{aligned}$$

と表わし, 点 C における三角形の内角を θ と表わすことにします (図 9 を参照). このとき, 直角三角形 $\triangle ACD, \triangle ABD$ に対して, それぞれ, 「三平方の定理」を用いると,

$$\begin{cases} d^2 = b^2 - x^2 \\ d^2 = c^2 - (a - x)^2 \end{cases}$$

となることが分かりますから,

$$b^2 - x^2 = c^2 - (a - x)^2 \quad (123)$$

となることが分かります. よって, (123) 式から,

$$2ax = a^2 + b^2 - c^2 \quad (124)$$

となることが分かります. 一方, x は,

$$x = b \cos \theta \quad (125)$$

と表わせることが分かりますから, (124) 式, (125) 式から,

$$2ab \cos \theta = a^2 + b^2 - c^2 \quad (126)$$

となることが分かります. もちろん, (126) 式は, 皆さん良くご存じの「三角形の余弦定理」に他なりません.

そこで, いま, \mathbb{R}^n の二つのベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して, ベクトル \mathbf{u} とベクトル \mathbf{v} の間の角度を θ として, \mathbf{u} と \mathbf{v} を二辺とする三角形に対して, (126) 式を適用してみます (図 10 を参照). すると, この場合,

$$a = \|\mathbf{u}\|, \quad b = \|\mathbf{v}\|, \quad c = \|\mathbf{u} - \mathbf{v}\| \quad (127)$$

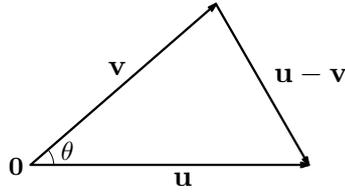


図 10: 二つのベクトル \mathbf{u} と \mathbf{v} を二辺とする三角形を考えてみる.

となることが分かりますから, (126) 式, (127) 式から,

$$\|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos \theta = \frac{1}{2} \cdot \{\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2\} \quad (128)$$

となることが分かります. 皆さん, 良くご存じのように, (128) 式の左辺に現われた量を, \mathbb{R}^n の二つのベクトル \mathbf{u}, \mathbf{v} の間のユークリッド内積と呼びます. ここでは, \mathbb{R}^n 上の標準的なユークリッド内積を,

\mathbb{R}^n 上の標準的なユークリッド内積の幾何学的な定義式

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos \theta \quad (129)$$

という記号を用いて表わすことにします.

いま, (122) 式を用いると,

$$\mathbf{u} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \mathbf{v} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$$

として,

$$\begin{aligned} \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 - \|\mathbf{u} - \mathbf{v}\|^2 &= \{(x_1)^2 + (x_2)^2 + \cdots + (x_n)^2\} \\ &\quad + \{(y_1)^2 + (y_2)^2 + \cdots + (y_n)^2\} \\ &\quad - \{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2\} \\ &= 2 \cdot (x_1 y_1 + x_2 y_2 + \cdots + x_n y_n) \end{aligned} \quad (130)$$

と表わせることが分かりますから, (128) 式, (129) 式, (130) 式から, \mathbb{R}^n 上の標準的なユークリッド内積は, 座標を用いて,

\mathbb{R}^n 上の標準的なユークリッド内積の座標を用いた定義式

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \quad (131)$$

というように表わせることが分かります.

さて, (131) 式の表示を用いると, ユークリッド内積 $\langle \cdot, \cdot \rangle_{\mathbb{R}^n}$ は,

— \mathbb{R}^n 上の標準的なユークリッド内積の持つ基本的な性質 —

(イ) 双線型性：勝手なベクトル $\mathbf{u}, \mathbf{u}', \mathbf{v}, \mathbf{v}' \in \mathbb{R}^n$ と勝手な実数 $c \in \mathbb{R}$ に対して、次が成り立つ.

$$\begin{cases} \langle \mathbf{u} + \mathbf{u}', \mathbf{v} \rangle_{\mathbb{R}^n} = \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} + \langle \mathbf{u}', \mathbf{v} \rangle_{\mathbb{R}^n} \\ \langle c\mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} = c \cdot \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} \\ \langle \mathbf{u}, \mathbf{v} + \mathbf{v}' \rangle_{\mathbb{R}^n} = \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} + \langle \mathbf{u}, \mathbf{v}' \rangle_{\mathbb{R}^n} \\ \langle \mathbf{u}, c\mathbf{v} \rangle_{\mathbb{R}^n} = c \cdot \langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} \end{cases}$$

(ロ) 対称性：勝手なベクトル $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ に対して、次が成り立つ.

$$\langle \mathbf{u}, \mathbf{v} \rangle_{\mathbb{R}^n} = \langle \mathbf{v}, \mathbf{u} \rangle_{\mathbb{R}^n}$$

(ハ) 正值性：勝手なベクトル $\mathbf{u} \in \mathbb{R}^n$ に対して、

$$\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbb{R}^n} \geq 0$$

となり、かつ、次が成り立つ.

$$\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbb{R}^n} = 0 \iff \mathbf{u} = \mathbf{0}$$

という三つの性質を持つことが分かります.⁴³ また、ベクトル $\mathbf{u} \in \mathbb{R}^n$ の長さは、ユークリッド内積を用いて、

— \mathbb{R}^n 上の標準的なユークリッド内積に関するベクトル $\mathbf{u} \in \mathbb{R}^n$ の長さ —

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_{\mathbb{R}^n}}$$

というように表わせることも分かります.

さて、2 節、4 節、6 節、7 節などでも見たように、内積という概念は、色々な数学的な現象を理解しようと思ったときに、とても便利な概念です。そこで、一般の線型空間 V に対しても、上の三つの条件を満たすような写像

$$\langle \cdot, \cdot \rangle_V : V \times V \rightarrow \mathbb{R} \tag{132}$$

として、内積を定義することが考えられました。⁴⁴ すなわち、(132) 式のような写像が、

⁴³もちろん、これら三つの性質は、(129) 式という幾何学的な定義式にもとづいても導くことができます。興味のある方は考えてみて下さい。

⁴⁴後の議論で余計な混乱を生じないように、線型空間 V 上の内積を $\langle \cdot, \cdot \rangle_V$ というように添え字を付けて表わすことにしました。

— \mathbb{R} 上の線型空間 V 上の内積 $\langle \cdot, \cdot \rangle_V$ の条件 —

(イ) 双線型性：勝手な元 $\mathbf{u}, \mathbf{u}', \mathbf{v}, \mathbf{v}' \in V$ と勝手な実数 $c \in \mathbb{R}$ に対して、次が成り立つ。

$$\begin{cases} \langle \mathbf{u} + \mathbf{u}', \mathbf{v} \rangle_V = \langle \mathbf{u}, \mathbf{v} \rangle_V + \langle \mathbf{u}', \mathbf{v} \rangle_V \\ \langle c\mathbf{u}, \mathbf{v} \rangle_V = c \cdot \langle \mathbf{u}, \mathbf{v} \rangle_V \end{cases}$$

$$\begin{cases} \langle \mathbf{u}, \mathbf{v} + \mathbf{v}' \rangle_V = \langle \mathbf{u}, \mathbf{v} \rangle_V + \langle \mathbf{u}, \mathbf{v}' \rangle_V \\ \langle \mathbf{u}, c\mathbf{v} \rangle_V = c \cdot \langle \mathbf{u}, \mathbf{v} \rangle_V \end{cases}$$

(ロ) 対称性：勝手な元 $\mathbf{u}, \mathbf{v} \in V$ に対して、次が成り立つ。

$$\langle \mathbf{u}, \mathbf{v} \rangle_V = \langle \mathbf{v}, \mathbf{u} \rangle_V$$

(ハ) 正値性：勝手な元 $\mathbf{u} \in V$ に対して、

$$\langle \mathbf{u}, \mathbf{u} \rangle_V \geq 0$$

となり、かつ、次が成り立つ。

$$\langle \mathbf{u}, \mathbf{u} \rangle_V = 0 \iff \mathbf{u} = \mathbf{0}$$

という三つの条件を満たすとき、 $\langle \cdot, \cdot \rangle_V$ を線型空間 V 上の内積と呼びます。特に、(ハ) という条件から、勝手な元 $\mathbf{u} \in V$ に対して、 \mathbf{u} の長さを、

— 内積を持つ線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ の元 $\mathbf{u} \in V$ の長さ —

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle_V}$$

という式によって定めることができることが分かります。また、 $\mathbf{u}, \mathbf{v} \in V, t \in \mathbb{R}$ として、

$$t\mathbf{u} + \mathbf{v} \in V$$

という元に対して、(ハ) という条件を当てはめると、

$$\begin{aligned} 0 &\leq \langle t\mathbf{u} + \mathbf{v}, t\mathbf{u} + \mathbf{v} \rangle_V && ((\text{ハ}) \text{より}) \\ &= t^2 \cdot \langle \mathbf{u}, \mathbf{u} \rangle_V + t \cdot \langle \mathbf{u}, \mathbf{v} \rangle_V + t \cdot \langle \mathbf{v}, \mathbf{u} \rangle_V + \langle \mathbf{v}, \mathbf{v} \rangle_V && ((\text{イ}) \text{より}) \\ &= \|\mathbf{u}\|^2 \cdot t^2 + 2\langle \mathbf{u}, \mathbf{v} \rangle_V \cdot t + \|\mathbf{v}\|^2 && ((\text{ロ}) \text{より}) \end{aligned} \quad (133)$$

となることが分かります。これが、勝手な実数 $t \in \mathbb{R}$ に対して成り立つためには、(133) 式の右辺に現われる二次式の判別式 D は、

$$D \leq 0$$

でなければならないので、

$$\{\langle \mathbf{u}, \mathbf{v} \rangle_V\}^2 \leq \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2 \quad (134)$$

となることが分かります。よって、(134) 式から、

Schwarz の不等式

$$|\langle \mathbf{u}, \mathbf{v} \rangle_V| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\| \quad (135)$$

となることが分かります. この (135) 式を Schwarz の不等式と呼びます. すると, (135) 式より,

内積と二つの元 \mathbf{u}, \mathbf{v} のなす角度の関係

$$\langle \mathbf{u}, \mathbf{v} \rangle_V = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \cdot \cos \theta \quad (136)$$

となるような実数 $\theta \in \mathbb{R}$ が存在することが分かりますから, このような実数 θ として, 二つの元 \mathbf{u} と \mathbf{v} のなす角度が定義できることが分かります.

さて, このような内積を持つ線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ が, 勝手にひとつ与えられているときに, $\dim_{\mathbb{R}} V = n$ として, V 上の基底 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ を, 勝手にひとつ取ってきたとします. このとき, 7 節で見たように, 必要なら, 基底 $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ に対して, Gram-Schmidt の直交化法を施すことによって, V 上の正規直交基底 $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ を構成することができます. また, 4 節で見たように, 正規直交基底 $\{\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n\}$ を用いて線型空間 V に「番地割り」して考えることにすると, 内積の構造まで込めて,

正規直交基底を用いると, 内積の構造まで込めて V と \mathbb{R}^n を同一視できる

$$(V, \langle \cdot, \cdot \rangle_V) \cong (\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathbb{R}^n})$$

というように同一視ができることが分かります. このことを逆に考えると, 一見, とても抽象的に見える線型空間 $(V, \langle \cdot, \cdot \rangle_V)$ も, ユークリッド空間 $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_{\mathbb{R}^n})$ から座標軸を消し去った「真っ直ぐな空間」としてイメージすることができるということが分かります. また, 線型空間 V の元の「長さ」や二つの元のなす「角度」も, ユークリッド空間上のベクトルの「長さ」や二つのベクトルのなす「角度」としてイメージすることができるということが分かります.

さて, 第 7 回の問 1 のところでは, 集合上の「関数の空間」が「特定の座標軸を持たない線型空間」の代表的な例であることを述べました. また, 第 9 回の問 1 のところでは, 特に, 三点からなる集合 $S = \{0, 1, 2\}$ を考えて, S 上の関数の空間

$$V_S = \{f : S \rightarrow \mathbb{R}\}$$

について考えてみました.⁴⁵ そこで, いま, $f, g \in V_S$ に対して,

集合 S 上の実数値関数全体のなす線型空間 V_S 上の自然な内積

$$\langle f, g \rangle_{V_S} = \sum_{x \in S} f(x)g(x) \quad (137)$$

という式によって定まる写像

$$\langle \cdot, \cdot \rangle_{V_S} : V_S \times V_S \rightarrow \mathbb{R}$$

⁴⁵ 第 9 回の問 1 では, V_S として, S 上の複素数値関数全体の集合を考えましたが, ここでは, S 上の実数値関数全体の集合を考えることにします.

を考えてみます. このとき, $\langle \cdot, \cdot \rangle_{V_S}$ は, V_S 上の内積を定めることが分かります.⁴⁶ 実際, 第9回の問1のときと同様に,

$$V_S \ni f \longleftrightarrow \begin{pmatrix} f(0) \\ f(1) \\ f(2) \end{pmatrix} \in \mathbb{R}^3$$

というように「番地割り」して考えてみると, 内積の構造まで込めて,

$$(V_S, \langle \cdot, \cdot \rangle_{V_S}) \cong (\mathbb{R}^3, \langle \cdot, \cdot \rangle_{\mathbb{R}^3})$$

というように同一視されることが分かります. その意味で, $\langle \cdot, \cdot \rangle_{V_S}$ のような内積は関数の空間 V_S 上の自然な内積であると考えられます.

また, 「 $\sum_{x \in S}$ 」という和を適当な意味に解釈することで, (137) 式は, いろいろな集合 S 上の関数の空間の場合に拡張して考えることができます. 例えば, $S = \mathbb{R}$ の場合であれば, 「 $\sum_{x \in \mathbb{R}}$ 」とは積分のことであると解釈するのが自然ですから,

集合 \mathbb{R} 上の実数値関数全体のなす線型空間 $V_{\mathbb{R}}$ 上の自然な内積

$$\langle f, g \rangle_{V_{\mathbb{R}}} = \int_{\mathbb{R}} f(x)g(x)dx$$

という内積を考えることになります.⁴⁷ このような内積が関数空間 V_S 上の内積の代表的な例になります. 興味のある方は, $S = [0, 2\pi]$ として,

$$\langle f, g \rangle_{V_{[0, 2\pi]}} = \int_0^{2\pi} f(x)g(x)dx$$

という内積を考えると,

$$\cos mx, \sin nx, \quad (m, n \in \mathbb{N})$$

という三角関数たちの間の内積の値がどうなるのかということを考えてみて下さい.

ここでは, \mathbb{R} 上の線型空間に対する内積を考えましたが, \mathbb{C}^n 上のエルミート内積を一般化して, \mathbb{C} 上の線型空間に対する内積を考えることもできます. すなわち, \mathbb{C} 上の線型空間 $V_{\mathbb{C}}$ に対して,

$$(\cdot, \cdot)_{V_{\mathbb{C}}} : V_{\mathbb{C}} \times V_{\mathbb{C}} \rightarrow \mathbb{C}$$

という写像が,

⁴⁶皆さん, $\langle \cdot, \cdot \rangle_{V_S}$ が (イ), (ロ), (ハ) という三つの性質を持つことを確かめてみて下さい.

⁴⁷実際には, $\int_{\mathbb{R}} f(x)g(x)dx$ という積分の値がきちんと定まる必要がありますから, 上のような内積を考えるためには, $V_{\mathbb{R}}$ の適当な線型部分空間に制限して考える必要があります.

— \mathbb{C} 上の線型空間 $V_{\mathbb{C}}$ 上の内積 $(\cdot, \cdot)_{V_{\mathbb{C}}}$ の条件 —

(イ) “双線型性”：勝手な元 $\mathbf{u}, \mathbf{u}', \mathbf{v}, \mathbf{v}' \in V_{\mathbb{C}}$ と勝手な複素数 $\zeta \in \mathbb{C}$ に対して、次が成り立つ。

$$\begin{cases} (\mathbf{u} + \mathbf{u}', \mathbf{v})_{V_{\mathbb{C}}} = (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} + (\mathbf{u}', \mathbf{v})_{V_{\mathbb{C}}} \\ (\zeta \mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} = \zeta \cdot (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} \\ (\mathbf{u}, \mathbf{v} + \mathbf{v}')_{V_{\mathbb{C}}} = (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} + (\mathbf{u}, \mathbf{v}')_{V_{\mathbb{C}}} \\ (\mathbf{u}, \zeta \mathbf{v})_{V_{\mathbb{C}}} = \bar{\zeta} \cdot (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} \end{cases}$$

(ロ) 対称性：勝手な元 $\mathbf{u}, \mathbf{v} \in V_{\mathbb{C}}$ に対して、次が成り立つ。

$$(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} = \overline{(\mathbf{v}, \mathbf{u})_{V_{\mathbb{C}}}}$$

(ハ) 正値性：勝手な元 $\mathbf{u} \in V_{\mathbb{C}}$ に対して、

$$(\mathbf{u}, \mathbf{u})_{V_{\mathbb{C}}} \geq 0$$

となり、かつ、次が成り立つ。

$$(\mathbf{u}, \mathbf{u})_{V_{\mathbb{C}}} = 0 \iff \mathbf{u} = \mathbf{0}$$

という三つの条件を満たすときに、 $(\cdot, \cdot)_{V_{\mathbb{C}}}$ を線型空間 $V_{\mathbb{C}}$ 上の内積と呼びます。ここで、(イ) という条件では、 $\mathbf{u} \in V_{\mathbb{C}}$ を勝手にひとつ固定したときに、

$$(\mathbf{u}, \cdot)_{V_{\mathbb{C}}} : V_{\mathbb{C}} \rightarrow \mathbb{C}$$

という写像は線型写像ではなく、勝手な複素数 $\zeta \in \mathbb{C}$ に対して、 ζ 倍を外に出すときに、 $\bar{\zeta}$ 倍という複素共役に化けるので、(イ) という条件を「 $\bar{\cdot}$ 」を付けて、「双線型性」というように表現しました。⁴⁸ 2 節でも注意しましたが、不幸なことに、(イ) という条件において、第一成分に対して反線型性の条件を課すのか、第二成分に対して反線型性の条件を課すのかという慣習が、数学と物理では逆になっています。ここでは、数学の慣習に従うことにしましたが、物理の慣習に慣れている方は、そちらの慣習に読み替えて理解して下さい。

また、(ロ) という条件から、勝手な元 $\mathbf{u} \in V_{\mathbb{C}}$ に対して、

$$(\mathbf{u}, \mathbf{u})_{V_{\mathbb{C}}} \in \mathbb{R}$$

となることが分かるので、(ハ) という条件が意味を持つことに注意して下さい。これより、勝手な元 $\mathbf{u} \in V_{\mathbb{C}}$ に対して、 \mathbf{u} の長さを、

— 内積を持つ線型空間 $(V_{\mathbb{C}}, (\cdot, \cdot)_{V_{\mathbb{C}}})$ の元 $\mathbf{u} \in V_{\mathbb{C}}$ の長さ —

$$\|\mathbf{u}\| = \sqrt{(\mathbf{u}, \mathbf{u})_{V_{\mathbb{C}}}}$$

という式によって定めることができます。

そこで、前と同様に、 $\mathbf{u}, \mathbf{v} \in V_{\mathbb{C}}$ 、 $\zeta \in \mathbb{C}$ として、

$$\zeta \mathbf{u} + \mathbf{v} \in V$$

⁴⁸ こうした「線型写像もどき」の写像のことを反線型写像 (anti-linear map) と呼んだりします。

という元に対して, (八) という条件を当てはめると,

$$\begin{aligned}
 0 &\leq (\zeta \mathbf{u} + \mathbf{v}, \zeta \mathbf{u} + \mathbf{v})_{V_{\mathbb{C}}} && ((八) \text{より}) \\
 &= |\zeta|^2 \cdot (\mathbf{u}, \mathbf{u})_{V_{\mathbb{C}}} + \zeta \cdot (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} + \bar{\zeta} \cdot (\mathbf{v}, \mathbf{u})_{V_{\mathbb{C}}} + (\mathbf{v}, \mathbf{v})_{V_{\mathbb{C}}} && ((イ) \text{より}) \\
 &= \|\mathbf{u}\|^2 \cdot |\zeta|^2 + (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} \cdot \zeta + \overline{(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}} \cdot \bar{\zeta} + \|\mathbf{v}\|^2 && ((ロ) \text{より}) \quad (138)
 \end{aligned}$$

となることが分かります. ここで, (138) 式の右辺に現われる $(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}$ という値を, その絶対値 $|(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}|$ に置き換えるために, $(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} \in \mathbb{C}$ という値の複素数としての極表示を,

$$(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} = |(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}| e^{\sqrt{-1}\theta_0}$$

というように表わすことにして, 勝手な実数 $t \in \mathbb{R}$ に対して,

$$\zeta = t e^{-\sqrt{-1}\theta_0}$$

としてみます. すると,

$$\begin{aligned}
 (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} \cdot \zeta &= (\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}} \cdot e^{-\sqrt{-1}\theta_0} \cdot t \\
 &= |(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}| \cdot t
 \end{aligned}$$

となることが分かりますから, (138) 式から,

$$0 \leq \|\mathbf{u}\|^2 \cdot t^2 + 2|(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}| \cdot t + \|\mathbf{v}\|^2$$

となることが分かります. これが, 勝手な実数 $t \in \mathbb{R}$ に対して成り立つのですから, 前と同様に,

Schwarz の不等式

$$|(\mathbf{u}, \mathbf{v})_{V_{\mathbb{C}}}| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\| \quad (139)$$

となることが分かります. この (139) 式も Schwarz の不等式と呼ばれています.

12 問2を見直すと*

さて, 問2では, トレースが0となる2行2列のエルミート行列全体からなる線型空間 V を取り上げて, V 上の内積やその内積を保つような線型写像について考えてみました. そこで, 問2の数学的な内容を見返してみる前に, そもそも, どうして問2の問題文中で与えられているような内積が V 上の内積として自然なものなのかということを考えてみることにします.

いま, 2行2列の複素行列全体の集合を,

2行2列の複素行列全体の集合

$$M_2(\mathbb{C}) = \left\{ X = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \mid z_{11}, z_{12}, z_{21}, z_{22} \in \mathbb{C} \right\}$$

という記号を用いて表わすことにします.⁴⁹ このとき,

$$M_2(\mathbb{C}) \ni X = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \longleftrightarrow \begin{pmatrix} z_{11} \\ z_{12} \\ z_{21} \\ z_{22} \end{pmatrix} \in \mathbb{C}^4$$

というように対応させることで, 線型空間として,

線型空間として $M_2(\mathbb{C})$ は \mathbb{C}^4 と同一視できる

$$M_2(\mathbb{C}) \cong \mathbb{C}^4 \quad (140)$$

というように同一視できることが分かります.⁵⁰ そこで, (140) 式という同一視のもとで, \mathbb{C}^4 上の標準的なエルミート内積 $(\cdot, \cdot)_{\mathbb{C}^4}$ を $M_2(\mathbb{C})$ 上の内積 (\cdot, \cdot) として読み替えてみます. すると, この内積は, $X, Y \in M_2(\mathbb{C})$ に対して,

$M_2(\mathbb{C})$ 上の標準的なエルミート内積

$$(X, Y) = \text{tr}(X {}^t\bar{Y}) \quad (141)$$

という式によって与えられることが分かります.⁵¹ したがって, $M_2(\mathbb{C})$ に対して, (141) 式で与えられるような内積を考えることは, とても自然なことであることが分かります.

さて, 問2では,

$$V = \{X \in M_2(\mathbb{C}) \mid {}^t\bar{X} = X, \text{tr } X = 0\}$$

という $M_2(\mathbb{C})$ の部分集合を考えました. ここで,

$${}^t\bar{X} = X$$

という条件は複素数を掛け算する操作で不変ではありませんから, V は複素数を掛け算する操作で閉じた部分集合にはならないことに注意します. したがって, $M_2(\mathbb{C})$ を \mathbb{C} 上の線型空間と見なしたときには, V は $M_2(\mathbb{C})$ の線型部分空間にはならないことが分かります.⁵² 一方, V は実数を掛け算する操作で閉じた部分集合にはなりますから, $M_2(\mathbb{C})$ を \mathbb{R} 上の線型空間と見なしたときには, V は $M_2(\mathbb{C})$ の線型部分空間になることが分かります. また, $X, Y \in V$ とすると,

$${}^t\bar{Y} = Y$$

⁴⁹英語で行列のことを「matrix」と言います.

⁵⁰ここで, 数ベクトル空間 \mathbb{C}^4 には自然な積構造は存在しないことに注意して下さい. その意味で, $M_2(\mathbb{C})$ という記号で表わすときには, $M_2(\mathbb{C})$ を, 単なる線型空間としてではなく, 線型空間としての構造の他に, 行列の積により定まる積構造まで入った集合として考えていることが多いです.

⁵¹皆さん, (141) 式の右辺を具体的に計算してみることで, このことを確かめてみて下さい.

⁵²例えば, $X \in V$ として, $Y = \sqrt{-1}X$ という行列を考えると,

$$\begin{aligned} {}^t\bar{Y} &= -\sqrt{-1} \cdot {}^t\bar{X} \\ &= -\sqrt{-1} \cdot X && (X \in V \text{ より}) \\ &= -Y \end{aligned}$$

となることが分かりますから, $X = O$ でない限り, $\sqrt{-1}X \in V$ とはならないことが分かります.

となりますから,

$$(X, Y) = \text{tr}(XY) \quad (142)$$

というように表わせることが分かります. そこで, (142) 式と 11 節で見たエルミート内積に対する (□) という条件を用いると,

$$\begin{aligned} (X, Y) &= \text{tr}(XY) && ((142) \text{ 式より}) \\ &= \text{tr}(YX) \\ &= (Y, X) && ((142) \text{ 式より}) \\ &= \overline{(X, Y)} && ((\square) \text{ より}) \end{aligned}$$

となることが分かりますから, $(\langle \cdot, \cdot \rangle)$ というエルミート内積を, V 上に制限したものを,

$$\langle \cdot, \cdot \rangle = (\cdot, \cdot)|_{V \times V}$$

という記号を用いて表わすことにすると, $\langle \cdot, \cdot \rangle$ は,

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$$

という写像を定めることが分かります. さらに, (\cdot, \cdot) が満たすべき \mathbb{C} 上の線型空間に対するエルミート内積の条件から, $\langle \cdot, \cdot \rangle$ は \mathbb{R} 上の線型空間に対するユークリッド内積の条件を満たすことが分かります.⁵³ こうして, $M_2(\mathbb{C})$ 上の自然なエルミート内積を V 上に制限することによって, 問 2 の問題文にあるようなユークリッド内積が得られることが分かりました. ただし, 問 2 の (1) で考えたような同一視

$$V \cong \mathbb{R}^3$$

のもとで, $\langle \cdot, \cdot \rangle$ が \mathbb{R}^3 上の標準的なユークリッド内積 $\langle \cdot, \cdot \rangle_{\mathbb{R}^3}$ に対応するように, あらかじめ $\frac{1}{2}$ 倍するという「細工」をしておきました.

次に, 問 2 の数学的な内容について見直してみることにします. 問 2 の (2) で見たように, $\{e_1, e_2, e_3\}$ という正規直交基底を用いて線型空間 V に「番地割り」してみると, 内積の構造まで込めて,

$$(V, \langle \cdot, \cdot \rangle) \cong (\mathbb{R}^3, \langle \cdot, \cdot \rangle_{\mathbb{R}^3}) \quad (143)$$

というように同一視できることが分かります. また, 問 2 の (3) では, 2 行 2 列のユニタリ行列 $A \in U(2)$ を用いて, $X \in V$ に対して,

$$\begin{aligned} T_A(X) &= AXA^{-1} \\ &= AX {}^t \bar{A} \end{aligned}$$

という式によって定まる V 上の線型写像

$$T_A : V \rightarrow V$$

⁵³ 皆さん, 確かめてみて下さい.

について考えてみました。⁵⁴ 第8回の問1や第9回の問1のところで見たとように、 T_A という線型写像は「表現行列の変換公式」として、数学の色々な場所で登場してきます。⁵⁵

さて、3節で見たように、2行2列のユニタリー行列とは、 \mathbb{C}^2 上の標準的なエルミート内積 $(\cdot, \cdot)_{\mathbb{C}^2}$ を保つような線型写像として特徴付けることができますが、問2の(3)で見たことから、 A がこのようなユニタリー行列である場合には、 T_A という V 上の線型写像も V 上の内積 $\langle \cdot, \cdot \rangle$ を保つような線型写像になるということが分かります。そこで、いま、(143)式のように、内積の構造まで込めて V と \mathbb{R}^3 を同一視して考えることにします。すると、3節で見たように、 \mathbb{R}^3 上の標準的なユークリッド内積 $\langle \cdot, \cdot \rangle_{\mathbb{R}^3}$ を保つような線型写像として、3行3列の直交行列を特徴付けることができますから、 T_A は直交行列であると見なせることが分かります。⁵⁶ よって、2行2列のユニタリー行列 $A \in U(2)$ に対して、3行3列の直交行列 $T_A \in O(3)$ を対応させることで、

————— T は $U(2)$ の元を $O(3)$ の元に写すような写像と解釈できる —————

$$T : U(2) \rightarrow O(3)$$

という写像が得られることが分かります。⁵⁷ 例えば、問2の(4)で見たように、勝手な実数 $\theta, \eta \in \mathbb{R}$ に対して、

$$A = \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & e^{\sqrt{-1}\eta} \end{pmatrix} \in U(2) \quad (144)$$

という行列を考えると、 T_A は、

$$T_A = \begin{pmatrix} \cos(\theta - \eta) & -\sin(\theta - \eta) & 0 \\ \sin(\theta - \eta) & \cos(\theta - \eta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SO(3) \quad (145)$$

という z 軸まわりの $(\theta - \eta)$ 回転を表わす行列になることが分かります。

さて、3節で注意したように、 $U(2)$ や $O(3)$ は、その集合の中で掛け算や逆元を考えることができるような集合になります。⁵⁸ いま、 $A, B \in U(2)$ 、 $X \in V$ に対して、

$$\begin{aligned} T_{AB}(X) &= (AB)X(AB)^{-1} \\ &= ABXB^{-1}A^{-1} \\ &= A(BXB^{-1})A^{-1} \\ &= AT_B(X)A^{-1} \\ &= T_A(T_B(X)) \end{aligned}$$

⁵⁴ここで、3節で説明した n 行 n 列のユニタリー行列全体の集合を表わす $U(n)$ という記号を用いました。

⁵⁵少し記号が紛らわしいですが、 $X \rightsquigarrow A$ 、 $A \rightsquigarrow P^{-1}$ と書き直してみると、

$$T_{P^{-1}}(A) = P^{-1}AP$$

と表わすことができます。

⁵⁶正確には、(143)式のような「番地割り」のもとで、線型写像 T_A の表現行列 \hat{T}_A を考えると、表現行列 \hat{T}_A が直交行列になるということです。以下の議論では、線型写像 T_A とその表現行列 \hat{T}_A を同一視して考えることにします。

⁵⁷ここで、3節で説明した n 行 n 列の直交行列全体の集合を表わす $O(n)$ という記号を用いました。

⁵⁸数学では、このような集合を群と呼ぶということも、3節で注意しました。

となることが分かりますから,

T は $U(2)$ の掛け算を $O(3)$ の掛け算に写す

$$T_{AB} = T_A \circ T_B \quad (146)$$

となることが分かります. すなわち, T は $U(2)$ の掛け算を $O(3)$ の掛け算に写すような写像であることが分かります. また, 勝手な自然数 $n \in \mathbb{N}$ に対して, n 行 n 列の単位行列を I_n という記号を用いて表わすことにすると, $X \in V$ に対して,

$$\begin{aligned} T_{I_2}(X) &= I_2 X I_2^{-1} \\ &= I_2 X I_2 \\ &= X \end{aligned}$$

となることが分かりますから,

T は $U(2)$ の単位元を $O(3)$ の単位元に写す

$$T_{I_2} = I_3 \quad (147)$$

となることが分かります. すなわち, T は $U(2)$ の単位元を $O(3)$ の単位元に写すような写像であることが分かります. さらに, (147) 式に注意して, (146) 式において,

$$B = A^{-1}$$

としてみると,

$$T_A \circ T_{A^{-1}} = I_3$$

となることが分かりますから,

T は $U(2)$ における逆元を $O(3)$ における逆元に写す

$$T_{A^{-1}} = (T_A)^{-1}$$

となることが分かります. すなわち, T は $U(2)$ の逆元を $O(3)$ の逆元に写すような写像であることが分かります. 以上から, T は 掛け算を掛け算に写し, 単位元を単位元に写し, 逆元を逆元に写すような写像であることが分かります.⁵⁹

そこで, 以上の準備のもとで, (144) 式, (145) 式を見返してみることにします. いま, $\theta = \eta \in \mathbb{R}$ である場合には,

$$c = e^{\sqrt{-1}\theta} \in \mathbb{C}$$

として,

$$A = cI_2 \quad (148)$$

というように表わせませんが, この場合には, (145) 式から,

$$T_{cI_2} = I_3 \quad (149)$$

⁵⁹数学では, T のように, 群の構造を保つような写像のことを群の準同型写像と呼びます.

となることが分かります. すなわち, $U(2)$ の元のうち, スカラー行列は, 写像 T によって, すべて単位元に写されることが分かります. このことは, 写像 T の定義に戻って考えると, A がスカラー行列のときには, 勝手な元 $X \in V$ に対して,

$$AX = XA$$

となること, すなわち, スカラー行列はすべての行列と可換であるという事実の反映であることが分かります.

さらに, (146) 式と (149) 式を合わせて考えると, $A \in U(2)$, $c = e^{\sqrt{-1}\theta}$, ($\theta \in \mathbb{R}$) として,

$$\begin{aligned} T_{cA} &= T_{cI_2A} \\ &= T_{cI_2} \circ T_A \\ &= I_3 \circ T_A \\ &= T_A \end{aligned} \tag{150}$$

となることが分かります. したがって, 直交行列を構成するという観点から考えると, T という写像には「 $A \in U(2)$ に絶対値が 1 の複素数を掛け算しても直交行列 T_A 自体は変わらない」という「無駄」があることが分かります. 3 節で見たように, 一般に, ユニタリー行列の行列式は, 絶対値が 1 の複素数になりますから, ユニタリー行列 $A \in U(2)$ が, 勝手にひとつ与えられたときに, その行列式 $\det A$ は, 適当な実数 $\theta_0 \in \mathbb{R}$ を用いて,

$$\det A = e^{\sqrt{-1}\theta_0}$$

というように表わせることが分かります. そこで,

$$c = e^{-\frac{\sqrt{-1}\theta_0}{2}}$$

として, $cA \in U(2)$ という元を考えてみると,

$$\begin{aligned} \det(cA) &= c^2 \cdot \det A \\ &= e^{-\sqrt{-1}\theta_0} \cdot e^{\sqrt{-1}\theta_0} \\ &= 1 \end{aligned}$$

となることが分かります. このとき, (150) 式で見たように,

$$T_A = T_{cA}$$

となるのですから, 直交行列を構成するという観点から考えると, 写像 T の定義域を, 行列式の値が 1 となるようなユニタリー行列全体の集合 $SU(2)$ に制限して,

写像 T の定義域を $SU(2)$ に制限して考える

$$T : SU(2) \rightarrow O(3) \tag{151}$$

という写像を考えれば十分であることが分かります.

そこで、以下では、(151) 式の写像について少し調べてみることにします。そのために、まず、 $SU(2)$ という集合がどのような集合なのかということを考えてみることにします。いま、 $SU(2)$ の元 $A \in SU(2)$ を、勝手にひとつ取ってきて、

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{C}$$

と表わしてみます。また、行列 A の列ベクトルを、

$$\mathbf{p}_1 = \begin{pmatrix} a \\ b \end{pmatrix}, \quad \mathbf{p}_2 = \begin{pmatrix} c \\ d \end{pmatrix} \in \mathbb{C}^2$$

と表わすことにします。このとき、3 節で見たように、

$$A \in U(2) \iff \{\mathbf{p}_1, \mathbf{p}_2\} \text{ が } \mathbb{C}^2 \text{ の正規直交基底になる}$$

となることに注意します。すると、 $\|\mathbf{p}_1\| = 1$ となることから、

$$|a|^2 + |b|^2 = 1 \tag{152}$$

となることが分かります。次に、

$$(\mathbf{p}_1, \mathbf{p}_2)_{\mathbb{C}^2} = 0$$

となる条件を考えてみると、

$$\bar{a}c + \bar{b}d = 0 \tag{153}$$

となることが分かります。⁶⁰ いま、(152) 式から、 $a \neq 0$ 、あるいは、 $b \neq 0$ のうち、少なくとも一方が成り立つことが分かりますが、 $a \neq 0$ であるとする、(153) 式から、

$$c = -\frac{\bar{b}d}{\bar{a}}$$

となることが分かりますから、

$$\begin{aligned} \begin{pmatrix} c \\ d \end{pmatrix} &= \begin{pmatrix} -\frac{\bar{b}d}{\bar{a}} \\ d \end{pmatrix} \\ &= \frac{d}{\bar{a}} \cdot \begin{pmatrix} -\bar{b} \\ \bar{a} \end{pmatrix} \end{aligned} \tag{154}$$

と表わせることが分かります。また、 $b \neq 0$ であるとする、(153) 式から、

$$d = -\frac{\bar{a}c}{\bar{b}}$$

となることが分かりますから、

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} c \\ -\frac{\bar{a}c}{\bar{b}} \end{pmatrix}$$

⁶⁰ここで、以下の議論が見やすくなるように、 $(\mathbf{p}_1, \mathbf{p}_2)_{\mathbb{C}^2} = 0$ ではなく、 $(\mathbf{p}_2, \mathbf{p}_1)_{\mathbb{C}^2} = 0$ という式の方を書きました。

$$= -\frac{c}{\bar{b}} \cdot \begin{pmatrix} -\bar{b} \\ \bar{a} \end{pmatrix} \quad (155)$$

と表わせることが分かります. よって, (154) 式, (155) 式から, $a \neq 0$, あるいは, $b \neq 0$ のうちのいずれが成り立つとしても, 適当な複素数 $\zeta \in \mathbb{C}$ が存在して,

$$\begin{pmatrix} c \\ d \end{pmatrix} = \zeta \cdot \begin{pmatrix} -\bar{b} \\ \bar{a} \end{pmatrix} \quad (156)$$

というように表わせることが分かります. すなわち,

$$A = \begin{pmatrix} a & -\zeta\bar{b} \\ b & \zeta\bar{a} \end{pmatrix}$$

と表わせることが分かります. そこで, さらに, $\det A = 1$ という条件を考えてみると, (152) 式から,

$$\begin{aligned} 1 &= \det A \\ &= \begin{vmatrix} a & -\zeta\bar{b} \\ b & \zeta\bar{a} \end{vmatrix} \\ &= \zeta \cdot \begin{vmatrix} a & -\bar{b} \\ b & \bar{a} \end{vmatrix} \\ &= \zeta \cdot (|a|^2 + |b|^2) \\ &= \zeta \end{aligned} \quad ((152) \text{ 式から})$$

となることが分かりますから,

$$\zeta = 1$$

でなければならないことが分かります. よって,

$$A = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$$

というように表わせることが分かります.

逆に, (152) 式を満たすような複素数 $a, b \in \mathbb{C}$ に対して,

$$A = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix}$$

という形の行列を考えてみると,

$$\begin{aligned} {}^t\bar{A}A &= \begin{pmatrix} \bar{a} & \bar{b} \\ -b & a \end{pmatrix} \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \\ &= \begin{pmatrix} |a|^2 + |b|^2 & 0 \\ 0 & |a|^2 + |b|^2 \end{pmatrix} \end{aligned}$$

$$\begin{aligned}
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\
\det A &= \begin{vmatrix} a & -\bar{b} \\ b & \bar{a} \end{vmatrix} \\
&= |a|^2 + |b|^2 \\
&= 1
\end{aligned}$$

となることが分かりますから,

$$A \in \mathrm{SU}(2)$$

となることが分かります.

以上より, $\mathrm{SU}(2)$ という集合は,

$\mathrm{SU}(2)$ という集合の具体的な記述

$$\mathrm{SU}(2) = \left\{ A = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \in M_2(\mathbb{C}) \mid |a|^2 + |b|^2 = 1, a, b \in \mathbb{C} \right\}$$

というように表わせることが分かります. そこで,

$$\mathrm{SU}(2) \ni A = \begin{pmatrix} a & -\bar{b} \\ b & \bar{a} \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{C}^2$$

という対応によって, $\mathrm{SU}(2)$ を \mathbb{C}^2 に埋め込んで考えてみると,

$\mathrm{SU}(2)$ という集合は三次元の球面 S^3 と同一視することができる

$$\mathrm{SU}(2) \cong S^3 \subset \mathbb{C}^2 \cong \mathbb{R}^4$$

というように, $\mathrm{SU}(2)$ は三次元球面 S^3 と同一視できることが分かります. 特に, $\mathrm{SU}(2)$ は連結な集合であることが分かります.

そこで, いま, $A \in \mathrm{SU}(2)$ に対して,

$$f(A) = \det(T_A)$$

という式で与えられる写像 f を考えてみます. このとき, T や \det は連続写像であるということと, 3 節で見たように, 直交行列の行列式は ± 1 となることに注意すると, 写像 f は,

$$f : \mathrm{SU}(2) \rightarrow \{1, -1\}$$

という連続写像であることが分かります. ところが, 上で見たように, $\mathrm{SU}(2)$ は連結な集合で, $\mathrm{SU}(2)$ の単位元 $I_2 \in \mathrm{SU}(2)$ での写像 f の値は,

$$f(I_2) = \det(I_3) = 1$$

となりますから, 勝手な元 $A \in \mathrm{SU}(2)$ に対して,

$$f(A) = 1$$

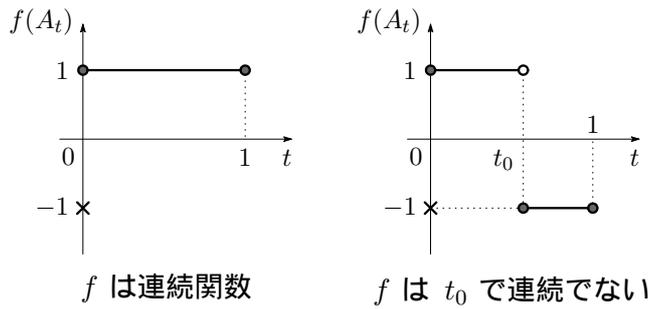


図 11: 離散集合 $\{1, -1\}$ に値をとる連結集合 $[0, 1]$ 上の連続関数は定数関数しか存在しない。

となることが分かります。⁶¹ したがって, (151) 式の写像 T は, 実は,

T は $SU(2)$ から $SO(3)$ への写像になる

$$T : SU(2) \rightarrow SO(3) \quad (157)$$

という写像であることが分かります。⁶²

そこで, 再び, (144) 式において, $\eta = -\theta$ としてみます. すると, (145) 式から, 勝手な実数 $\theta \in \mathbb{R}$ に対して,

$$A_\theta = \begin{pmatrix} e^{\sqrt{-1}\theta} & 0 \\ 0 & e^{-\sqrt{-1}\theta} \end{pmatrix} \in SU(2) \quad (158)$$

という形のユニタリ行列に対応する直交行列 T_{A_θ} は,

$$T_{A_\theta} = \begin{pmatrix} \cos(2\theta) & -\sin(2\theta) & 0 \\ \sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 1 \end{pmatrix} \in SO(3) \quad (159)$$

という形で与えられることが分かります. そこで, いま, \mathbb{R}^3 のベクトル $\mathbf{x} \in \mathbb{R}^3$ を,

$$\mathbf{x}_\theta = T_{A_\theta} \mathbf{x} \in \mathbb{R}^3$$

というように, T_{A_θ} という直交行列を用いて動かしてみます. すると, $\theta = \pi$ となったときに,

$$\begin{aligned} \mathbf{x}_\pi &= T_{A_\pi} \mathbf{x} \\ &= I_3 \mathbf{x} \\ &= \mathbf{x} \end{aligned}$$

⁶¹いま, $A_t \in SU(2)$, ($t \in [0, 1]$) という変数 t で連続的にパラメータ付けられた $SU(2)$ の元が, t とともに, $A_0 = I_2$ から $A_1 = A$ まで連続的に動いていくとします. このとき, f が連続写像であることから, $f(A_t)$ という値も t とともに, $\{1, -1\}$ という集合の中を連続的に動いていくこととなりますが, $\{1, -1\}$ は離散的な集合なので, 連続的な変化において, 値が 1 から, 突然, -1 に「ジャンプする」ということは不可能であることが分かります (図 11 も参照). したがって, $f(A) = f(A_1) = f(A_0) = f(I_2) = 1$ となることが分かります.

⁶²ここで, 3 節で説明した行列式が 1 となる n 行 n 列の直交行列全体のなす集合を表わす $SO(n)$ という記号を用いました.

というように元の状態に戻ってくることが分かります。同様に、 \mathbb{C}^2 のベクトル $\mathbf{u} \in \mathbb{C}^2$ を、

$$\mathbf{u}_\theta = A_\theta \mathbf{u} \in \mathbb{C}^2$$

というように、 A_θ というユニタリ行列を用いて動かしてみます。すると、今度は、 $\theta = \pi$ では、

$$\begin{aligned} \mathbf{u}_\pi &= A_\pi \mathbf{u} \\ &= (-I_2) \mathbf{u} \\ &= -\mathbf{u} \end{aligned}$$

というように元の状態には戻らずに、 (-1) 倍された状態に移り、 $\theta = 2\pi$ となったときに、

$$\begin{aligned} \mathbf{u}_{2\pi} &= A_{2\pi} \mathbf{u} \\ &= I_2 \mathbf{u} \\ &= \mathbf{u} \end{aligned}$$

というように、初めて元の状態に戻ってくることが分かります。このように、 \mathbb{R}^3 上の回転という視点から眺めてみると、 \mathbb{C}^2 のベクトル $\mathbf{u} \in \mathbb{C}^2$ は二回転して始めて元の状態に戻ってくるという不思議な変換性を持ったベクトルとして現われてくるということが分かります。一般に、空間の座標軸を回転させる操作に対して、 \mathbb{C}^2 のベクトルのような不思議な変換性を示すようなベクトルはスピノールと呼ばれていて、物理学において、物質の素であるフェルミオンと呼ばれる粒子を記述するために基本的な道具となっています。上では、(158) 式や (159) 式で与えられるような特別な形をした元に対して T という写像を調べましたが、実は、 $SU(3)$ の元と $SO(3)$ の元は、

写像 T によって $SU(2)$ の元と $SO(3)$ の元が 2 対 1 にピッタリ対応する

$$T : SU(2) \rightarrow SO(3)$$

という写像によって、「ほぼピッタリと対応する」ということ、より正確には、それぞれの直交行列 $P \in SO(3)$ に対して、

$$T_A = P$$

となるような行列式が 1 のユニタリ行列 $A \in SU(2)$ がちょうど二つ存在するということが分かります。⁶³ すなわち、 $SU(2)$ の元は $SO(3)$ の元より「二倍多く存在している」ことが分かります。

こうしたことは、行列の集合の間の単なる数学的な関係に過ぎないように思われますが、上で注意したように、このような事実がスピノールと呼ばれる空間の回転に対して不思議な変換性を示すようなベクトルの存在を可能にしています。そして、このようなスピノールの存在が、現実の世界で、我々の肉体の構成要素でもあり、「すべての物質の素」でもあるフェルミオンと呼ばれる粒子が存在できる数学的な根拠となっています。すなわち、この世の中で我々が肉体を持って存在しているということと、 $SU(2)$ と $SO(3)$ との間の対応

⁶³そのうちの一方を $A_0 \in SU(2)$ と表わすことにすると、もう一方は $(-A_0) \in SU(2)$ で与えられることも分かります。

が一對一ではなく「二対一であった」ということが、実は深く関係しているわけです。そう考えると、世の中というものは、実に不思議なものであるというような気がします。皆さんも、あれこれと想像をたくましくしながら、いろいろなことを勉強してみることで、感動とともに「世の中の不思議さ」に対する理解を深めていかれたら面白いのではないかと思います。