数理科学基礎 (線形代数学) / 線形代数学
レポート解答解説

穂坂 秀昭

2015 年 10 月 12 日

目次

S1 ターム

数理科学基礎 (線形代数学)

数理科学基礎 (線形代数学) 第 1 回　　複素数と代数学の基本定理

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 4 月 15 日 1 限

1　はじめに

■ごあいさつ　みなさん、はじめまして。この授業の TA (ティーチング・アシスタント) をすることになりました、数理科学研究科博士課程の穂坂といいます。これから 1 セメスターの間、よろしくお願いします。

この授業では毎回レポート問題が課され、それを TA の穂坂が添削して返却します。また答案返却に合わせて、この文書のような解説プリントを配付していく予定です。何か疑問要望等があれば、提出するレポートの片隅にメッセージを書くなり、メールを hosaka@ms.u-tokyo.ac.jp に送るなり、授業後に聞くなりしてください。

■課題提出時のお願い　答案が消えたら困りますので、次の 2 点は必ず守ってください。

- **氏名、学生証番号の両方**を書いてください。
- 答案が複数枚に渡るときは、**左上をホチキス止め**してください。

■問題を解くにあたって　毎週出題される問題はレポート課題ですから、とにもかくにも期限までに提出しないといけません。とは言っても、どうせ解くなら 1 問 1 問からなるべく多くの教訓を引きずり出したいし、何よりなるべく楽しい問題を解きたいものです。レポートに取り組むときは、次のようなことを意識してください。

- 簡単な計算問題は、授業で扱った定理などを確かめるための具体例です。常に「どの定理を、どう使っているのか」を考えながら解きましょう。
- どんな問題であっても、ただ解くだけでなく、見通しの良い解法を探すべきです。計算問題ならなるべく手間を減らし、証明問題なら本質的な部分を捉えるよう努力しないといけません。

■このプリントの作り方について　せっかくなので、このプリントをどう作っているかについて説明しておきます。

ふつう「コンピュータで文書を作る」というと、大抵の人が Microsoft Word とか一太郎といったワープロソフトを連想すると思います。ところが残念なことに、市販のワープロソフトでは数式を入力するのに大変な苦労を強いられてしまいます。そこで数学が専門の人はどうするかというと、そういったワープロソフトの代わりに “LaTeX” というソフトウェアを使います。これは D. E. Knuth[*1]という非常に有名な数学者・計算機科学者が作った “TeX” というソフトウェアを、色々な人が改良してできあがったものです。

LaTeX はワープロソフトとはちょっと違い、文字のスタイルを変えたり見出しをつけたりするのに「コマンド」というものを使います。ですから LaTeX を使うにはコマンドの使い方を覚えなければいけません。加えてキーボードが打ち込んだものが、見た目通りに出てくるわけでもありません。一旦コマンドも含めて打ち込んだ文書を LaTeX というプログラムその他色々に処理させることによって、やっと整形された文書がでてきます。ですから使い始めるにはちょっとハードルが高いのですが、使い慣れるとワープロソフトよりも手際よく文書が書けるし、また数式を中心として文書の仕上がりが美しいというメリットもあります。

もしかしたら皆さんの中には LaTeX を既に知っている人がいるかもしれませんし、また将来 LaTeX を使う必要に迫られる人がいるかもしれません。そこで https://github.com/HideakiHosaka/2015_linear_algebra に、この文書の PDF ファイルと LaTeX ソースコードを置いておきます。もし LaTeX の方に興味がある人は、ページ右下の “Download ZIP” のボタンから一式をダウンロードしてください。また東京大学が持つ情報処理システムのオンライ

[*1] http://www-cs-faculty.stanford.edu/~uno/

ン自習教材「はいぱーワークブック」の第 27 章[*2]に、LaTeX の説明があります。LaTeX を使う人は、一度読んでおくと良いと思います。

ちなみにソースコードの公開には "GitHub" というサービス[*3]を利用しています。上に貼った URL を開くと、古いバージョンのプリントや、そうしたプリントがどう更新されていったかも見ることができます。授業自体の役には立たないと思いますが、興味があれば見てみてください。

2 複素 (数) 平面の幾何

2.1 複素数と複素平面

■複素数の定義 まず、最初に複素数の定義をおさらいしましょう。$i^2 = -1$ という規則で i という「数」[*4]を定めます。このとき、2 つの実数 $x, y \in \mathbb{R}$ を用いて $z = x + yi$ と表される数を複素数と言うのでした。また $z = x + yi$ の x を実部、y を虚部と言うことも知っているはずです。足し算と掛け算はそれぞれ、分配法則などが上手く成り立つよう

$$(x+yi)+(x'+y'i) := (x+x')+(y+y')i, \qquad (x+yi)(x'+y'i) := (xx'-yy')+(xy'+x'y)i$$

と決めていました[*5]。また、これらのルールがあれば、$x' + y'i \neq 0$ のとき

$$\frac{x+yi}{x'+y'i} = \frac{(x+yi)(x'-y'i)}{(x'+y'i)(x'-y'i)} = \frac{(xx'+yy')+(x'y-xy')i}{(x')^2+(y')^2} = \frac{xx'+yy'}{(x')^2+(y')^2} + \frac{x'y-xy'}{(x')^2+(y')^2}i$$

というように割り算もできます。こうして複素数では四則演算が全部できると分かりました。

■複素平面 さて、全ての複素数は $x + yi$ の形に、(x, y) という二つの実数 x, y のペアを用いて表せます。また複素数を二つの実数 x, y で $x + yi$ の形に表す方法がただ一通りなことも明らかでしょう。これらの事実[*6]から、**複素数 $z = x + yi$ と座標平面上の点 (x, y) とを 1 : 1 に対応させられる**ことが分かります。このように平面 $\mathbb{R}^2$ 上の点一つ一つを複素数と見なしたとき、平面 $\mathbb{R}^2$ のことを**複素 (数) 平面**[*7]と呼びます。また x 軸, y 軸をそれぞれ**実軸** (real axis)、**虚軸** (imaginary axis) と言います。

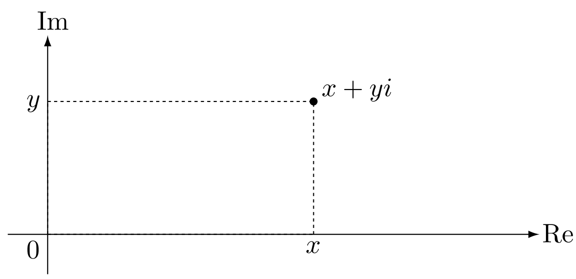


図 1.1 複素平面における数と点の対応

[*2] http://hwb.ecc.u-tokyo.ac.jp/current/applications/latex/

[*3] もしあなたが既に "GitHub" を知っているなら、きっと "pull request" の機能も知っているはずです。プリントに対して何か意見があれば、積極的に pull request を送ってください。:D

[*4] 誰もが一度は「-1 の平方根を数と呼んでいいのか」という疑問を抱いたことがあると思います。その疑問に対する答えをまだ書いていなかったので、ここでは括弧つきで「数」と書きました。でも一々こう書くと面倒なので、以下では括弧をつけないことにします。

[*5] ここで出てくる := は「左辺のものを右辺で定義する」という意味です。式変形をするときの = とは意味が違うので、定義の際には := が使われることがあります。また左右を入れ替えて =: とすると「右辺のものを左辺で定義する」という意味になります。

[*6] 集合と写像の言葉できちんと書くと、$(x, y) \in \mathbb{R}^2$ に $x + yi \in \mathbb{C}$ を対応させる写像 $\mathbb{R}^2 \to \mathbb{C}$ が全単射、ということです。

[*7] 複素平面と複素数平面は、どっちの言葉も同じ意味です。複素平面の方を使う人が多いですが、複素数平面と言っても誤解を招くことはないですし、昔は複素数平面という言葉も割と良く使われていたそうです。関西学院大学の示野信一先生のブログに詳しい事情が書いてありますので、気になる人は読んでみてください: http://mathsci.blog41.fc2.com/blog-entry-60.html

少し大げさに言うと、我々は「複素数という代数的なもの」と「平面という幾何的なもの」を対応付けました。このことには、非常に重要な意味があります。なぜなら**代数の観点と幾何の観点を行ったりきたりすることで、色々なことが分かるようになる**からです。たとえば長さなどの幾何的な情報を代数的な操作で捉えたり、逆に掛け算などの代数的な操作を図形的に捉えたりというように。これから、それをやってみましょう。

2.2 幾何を代数で捉える

■複素数の大きさと共役　幾何的な情報の最も典型的なものとして「2 点間の距離」が挙げられます。複素平面の場合、原点 0 と $z \in \mathbb{C}$ との距離を z の**絶対値**または**大きさ**と言い、$|z|$ で表します。三平方の定理から、すぐに $|x+yi| = \sqrt{x^2+y^2}$ が従います。またベクトルのときと同様、複素数 $z, z' \in \mathbb{C}$ の間の距離は $|z-z'|$ となります。

幾何的な観点からは、「線/点対称移動」といった操作を考えられるというメリットもあります。たとえば「実軸に対する線対称移動」で z が写る点を $\overline{z}$ と書くと、$\overline{x+yi} = x - yi$ です。この $\overline{z}$ を z の**共役**と言います[*8]。共役は $\overline{z+w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\,\overline{w}$ という性質を満たすことが、計算で確かめられます。また共役を用いて、複素数の大きさを表すことができます。

■複素数と複素数平面: 問 1 の解答　$z = x + iy$ のとき、$z\overline{z} = (x+yi)(x-yi) = x^2+y^2 = |z|^2$ となる。$z = 0$ であることは z と原点 0 との距離が 0 であることと同値なので、直ちに $z = 0 \Leftrightarrow |z| = 0$ を得る。 ■

■複素数と複素数平面: 問 4 の解答　(1) $z_1 = x_1 + y_1 i,\ z_2 = x_2 + y_2 i$ とおく。絶対値は 0 以上の実数だから、$|z_1+z_2| \leq |z_1| + |z_2|$ の代わりに $|z_1+z_2|^2 \leq (|z_1|+|z_2|)^2$ を示せばよい。実際計算すると

$$\begin{aligned}
(|z_1|+|z_2|)^2 - |z_1+z_2|^2 &= |z_1|^2 + 2|z_1||z_2| + |z_2|^2 - |z_1+z_2|^2 \\
&= x_1^2 + y_1^2 + 2\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2} + x_2^2 + y_2^2 - \left((x_1+x_2)^2 + (y_1+y_2)^2\right) \\
&= 2\Big(\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2} - (x_1x_2+y_1y_2)\Big)
\end{aligned}$$

となる。そして括弧の中は、平方根が非負であることと次の計算とを組み合わせれば、0 以上と分かる。

$$\left(\sqrt{x_1^2+y_1^2}\sqrt{x_2^2+y_2^2}\right)^2 - (x_1x_2+y_1y_2)^2 = x_1^2y_2^2 + y_1^2x_2^2 - 2x_1x_2y_1y_2 = (x_1y_2 - x_2y_1)^2 \geq 0$$

これで示すべきことが言えた。(2) は (1) を使えば $|z_1-z_2| = |z_1+(-z_2)| \leq |z_1| + |-z_2| = |z_1| + |z_2|$ となる。 ■

2.3 代数を幾何で捉える

続いて、四則演算という代数的な操作を複素平面で見てみましょう。複素数の足し算や引き算がベクトルの足し算や引き算と全く同じであることは、すぐに分かると思います。非自明なのは**平面上の点やベクトルには掛け算が定義されていないのに対し、複素数には掛け算がある**という点です[*9]。そこで「2 つの複素数を掛け算した結果は、複素平面上ではどのように見えるか」が問題となってきます。まずは問題を 1 つ解いてみましょう。

■複素数と複素数平面: 問 2 の解答　$z = 2+i$ とおくと

$$z^2 = 3+4i, \qquad z^3 = 2+11i, \qquad z^4 = -7+24i, \qquad z^5 = -38+41i, \qquad z^6 = -117+44i$$

である。これらをプロットした結果[*10]は次の図の通り。

[*8] この部分は嘘ではないですが、若干語弊があります。今は「実軸に関する線対称移動」という幾何学的な操作として共役を定義しましたが、本来「共役」とは、実数 $\mathbb{R}$ から複素数 $\mathbb{C}$ を作る「拡大」という代数的な操作に伴って定義されるものです。

[*9] 「ベクトルにも内積や外積があるじゃん」という声が聞こえてきそうですが、内積や外積は、いわゆる普通の「積」とは違う性質を持ちます。2 つのベクトルの内積は数になってしまい、ベクトルにはなりません。また 2 つのベクトルの外積はベクトルになりますが、ベクトルの外積は順序を入れ替わると結果が変わります。この辺が数の掛け算と全然違うところです。

[*10] z^5 と z^6 は図から激しくはみ出るので描いていません。

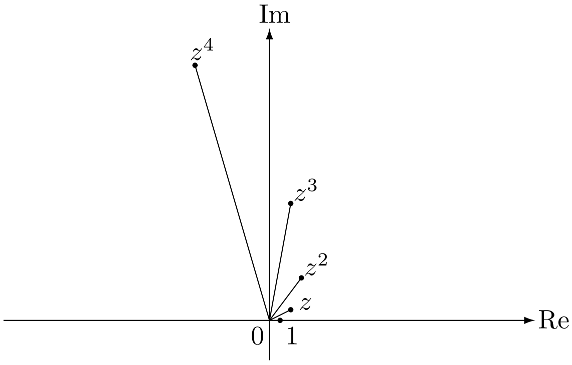


図 1.2　$z = 2 + i$ のべき乗のプロット

■極形式表示　いま複素数 $z \neq 0$ に対し $z^0 = 1$, $z^1 = 1$, z^2, z^3, z^4 を平面上にプロットし、これらの点を原点と結んだ結果を眺めると、隣り合う角が全て同じ大きさであるように見えます。それを実際に確かめてみましょう。

角度を計算したいので、極座標を使うのが筋がよさそうです。そこで $z = x + yi$ の表す点 (x, y) を、極座標 (r, θ) を用いて $x = r\cos\theta$, $y = r\sin\theta$ と表します。これを対応する複素数の方で表すと

$$z = x + yi = r\cos\theta + (r\sin\theta)i = r(\cos\theta + i\sin\theta)$$

となります。この書き方を、複素数の**極形式表示**と呼びます。極形式表示のもとで $r = |z|$ です。また θ は、複素平面の半直線 $0z$ と実軸の非負の部分がなす角を、反時計回りに測った角度となっています。この θ を複素数 z の**偏角**と言い、$\theta = \arg z$ と書きます。$\arg z$ は一通りでなく 2π の整数倍だけずらせますが、今は気にしないでおきます。

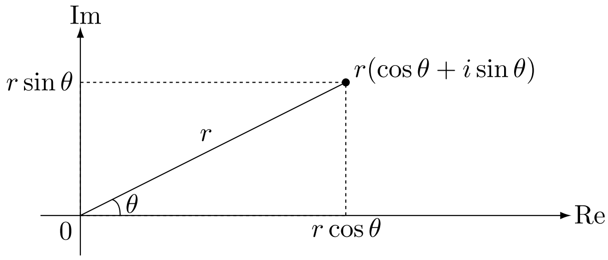


図 1.3　複素数の極形式表示

さて、極形式で表された 2 つの複素数を掛け算すると

$$\begin{aligned} r(\cos\theta + i\sin\theta) \times r'(\cos\theta' + i\sin\theta') &= rr'\bigl\{(\cos\theta\cos\theta' - \sin\theta\sin\theta') + i(\sin\theta\cos\theta' + \cos\theta\sin\theta')\bigr\} \\ &= rr'\bigl(\cos(\theta + \theta') + i\sin(\theta + \theta')\bigr) \end{aligned}$$

となります。最後の式変形は、もちろん三角函数[*11]の加法定理を使っています。この式は非常に重要なことを示唆しています。それは複素数 z, z' の積 zz' について

- 大きさは、$|zz'| = |z||z'|$ で与えられる
- 偏角は $\arg zz' = \arg z + \arg z'$ で与えられる

ということです。言い換えれば、複素数 z に対して別の複素数 z' を掛け算する操作は

- z の大きさを $|z'|$ 倍し
- z の偏角に $\arg z'$ を足し算する

[*11] 函数と関数は同じ意味です。少々古臭い言い回しですが、好みの問題でこちらを使います。

ということに他ならないからです。このように極形式を使うことによって、複素数の掛け算が「拡大縮小」と「回転」の組み合わせという図形的意味を持つことが読み取れるのです。

なお、一々 $\cos\theta + i\sin\theta$ と書いているのは長ったらしくて大変なので、以下では実数 θ に対して $e^{i\theta} = \cos\theta + i\sin\theta$ や $\exp i\theta = \cos\theta + i\sin\theta$ と表すことにします。たとえば $e^{i\pi} = \cos\pi + i\sin\pi = -1$ という感じです。指数函数 e^x と同じ記法を用いることには実は意味がある[*12]のですが、今は「単なる記号」だと思っていてください。この記号を使うと、極形式表示での掛け算は

$$re^{i\theta} \times r'e^{i\theta'} = rr'e^{i(\theta+\theta')}$$

と書けます。スッキリしてていいですね。

極形式表示は計算面でも、非常に威力を発揮することがあります。

■複素数と複素数平面: 問 3 の解答　$z = e^{2k\pi i/n}$ $(k = 0, 1, \dots, n-1)$ とおくと、

$$z^n = \left(e^{\frac{2k\pi i}{n}}\right)^n = \exp\left(\frac{2k\pi i}{n} + \frac{2k\pi i}{n} + \cdots + \frac{2k\pi i}{n}\right) = \exp\left(\frac{2k\pi i}{n} \times n\right) = e^{2k\pi i} = (e^{2\pi i})^k = 1$$

となる。ここで k が $k = 0, 1, \dots, n-1$ を動けば、n 個の異なる複素数が得られる[*13]。また多項式 $z^n - 1$ は n 次式だから、$n+1$ 個以上の根を持つことはない。ゆえに $z = e^{2k\pi i/n}$ $(k = 0, 1, \dots, n-1)$ が全ての根を与える。 ■

2.4　残りの問題

ここまでで紹介していなかった問題の解答を記します。

■複素数と複素数平面: 問 6 の解答　$x, y \in M$ とすると $x = a^2 + b^2$, $y = c^2 + d^2$ となる整数 $a, b, c, d \in \mathbb{Z}$ が取れる。このとき $x = |a+bi|^2$, $y = |c+di|^2$ なので $xy = |a+bi|^2|c+di|^2 = |(a+bi)(c+di)|^2 = |(ac-bd)+(ad+bc)i|^2 = (ac-bd)^2 + (ad+bc)^2$ となる。$a, b, c, d \in \mathbb{Z}$ より $ac-bd, ad+bc \in \mathbb{Z}$ である。よって $xy \in M$ である。 ■

■複素数と複素数平面: 問 7 の解答　次の式の t に好きな有理数を代入すればよい。

$$\left(\frac{2t}{1+t^2}\right)^2 + \left(\frac{1-t^2}{1+t^2}\right)^2 = 1$$

$t = \tan\theta$ とおくと、この式は $\sin^2 2\theta + \cos^2 2\theta = 1$ に化ける。したがって $0 \le t \le 1$ の範囲で $2t/(1+t^2)$ は短調増加する。これと 0 以上 1 以下の有理数が無限個存在することから、有理点も無限個存在することが分かる。 ■

3　多項式の性質

3.1　多項式の割り算

実数係数や複素数係数[*14]の多項式では、割り算と余りの計算ができます。$f(x)$ を m 次多項式、$g(x)$ を n 次多項式、$m \ge n$ とすると、$f(x) = g(x)q(x) + r(x)$ を満たす多項式 $g(x)$ と n 次未満の多項式 $r(x)$ がただ一つだけ存在します。実際 $f(x) = a_m x^m + a_{m-1}x^{m-1} + \cdots + a_0$, $g(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_0$ とおくと、$f(x) - b_n x^{m-n} g(x)/a_m$ の次数は $m-1$ 次以下になります。こうやって「$g(x)$ に上手い数と x のべき乗をかけ、$f(x)$ の項を次数が高い順に消していく」という操作をすれば、商と余りの計算ができます。実際の計算には筆算を使ったり、あるいは $g(x)$ が 1 次式のときは「組立除法」という技が使えたりしますが、その辺は割愛します。

[*12] そのうち微分積分学の授業で、函数の Taylor 展開というものを習うはずです。その後に巾 (べき) 級数で指数函数 e^x を定義し直すと、元々の「e のなんとか乗」という意味を越えて、e^x の x に複素数を代入できるようになります。そうして初めて $e^{i\theta} = \cos\theta + i\sin\theta$ という式に正しく意味を与えることができます。

[*13] 複素平面上にプロットすれば、異なることが直ちに分かります。

[*14] 有理数係数でも大丈夫です。より一般に、係数が「体」と呼ばれるものであれば、複素数係数と同様に割り算ができます。

特に $g(x)$ が 1 次式 $x-\alpha$ のとき、$f(x)=(x-\alpha)q(x)+r(x)$ に出てくる $r(x)$ は 0 次以下[*15]の式、つまり定数です。なので $r(x)$ の代わりに r と書くと、$f(x)=(x-\alpha)q(x)+r$ の両辺に $x=\alpha$ を代入して $r=f(\alpha)$ が得られます。よって、f が $x-\alpha$ で割り切れることと $f(\alpha)=0$ が同値になります。この事実を**因数定理**と呼ぶのでした。

因数定理などを用いて解ける問題を、まとめて片付けてしまいましょう。

■多項式: 問 1 の解答　$f(x)$ が $x-\alpha$ で割り切れるので、$f(x)=(x-\alpha)g(x)$ と書ける。また $f(x)$ は $x-\beta$ でも割り切れるので、$f(\beta)=0$ である。よって $0=f(\beta)=(\beta-\alpha)g(\beta)$ となるが、$\alpha\neq\beta$ より $g(\beta)=0$ でないといけない。よって $g(x)$ は $(x-\beta)$ で割り切れる。これより $f(x)$ は $(x-\alpha)(x-\beta)$ で割り切れる。 ■

■多項式: 問 2 の解答　$f(x)=x^3+10x^2+ax-2$ に $x=2,-3$ を代入した値が等しい。よって $f(2)=46+2a$ と $f(-3)=-3a+61$ が等しいのだから、$a=3$ が得られる。求める余りは 52 である。 ■

■多項式: 問 3 の解答　$f(x)=mx^3+nx^2-5$ とおく。$f(-\frac{1}{2})=0$ より $-\frac{m}{8}+\frac{n}{4}-5=0$ である。また $f(\frac{2}{3})=7$ より $\frac{8}{27}m+\frac{4}{9}n-5=7$ である。これより $m=6,\ n=23$ である。 ■

■多項式: 問 4 の解答　$F(x)$ を $2x^2+x-1$ で割った余りを $px+q$ と書くと、何か多項式 $P(x)$ を用いて

$$F(x)=(2x^2+x-1)P(x)+px+q$$

と書ける。$F(x)$ を $x+1$ で割った余りが 6 なので $F(-1)=6$ である、よって上式に $x=-1$ を代入して $6=-p+q$ を得る。同様に $F(x)$ を $2x-1$ で割った余りが 3 なので、$F(\frac{1}{2})=3$ である。これより $3=\frac{1}{2}p+q$ を得る。こうして p,q の連立 1 次方程式が得られたので、解くと $p=-2,\ q=4$ が得られる。よって余りは $-2x+4$ である。 ■

■多項式: 問 6 の解答　3 次多項式 $f(x)$ は、適当な多項式 $g(x)$ と $h(x)$ によって $f(x)=g(x)(x^2-1)+5x-8=h(x)(x^2-x-6)+17x+4$ と書ける。これに $x=1,-1,-2,3$ を代入すると、それぞれ $f(1)=-3,\ f(-1)=-13,\ f(-2)=-30,\ f(3)=55$ が得られる。そこで $f(x)=ax^3+bx^2+cx+d$ とおくと

$$\begin{cases} a+b+c+d & =-3 \\ -a+b-c+d & =-13 \\ -8a+4b-2c+d & =-30 \\ 27a+9b+3c+d & =55 \end{cases}$$

という連立一次方程式が得られる。これを a,b,c,d について解けば $f(x)=2x^3+3x-8$ と分かる。 ■

3.2　有名な多項式

今回の問題の中にはいくつか有名な多項式が出てくるので、問題を解きつつ紹介します。

■複素数と複素数平面: 問 5 の解答　$f(z)=z^4+z^3+z^2+z+1$ である。
(1) z を $f(z)=0$ の根とする。$z^5-1=(z-1)f(z)=0$ なので、$z\neq0$ である。よって $f(z)/z^2=0$ である。また $f(z)/z^2=z^2+z+1+z^{-1}+z^{-2}=(z+z^{-1})^2+(z+z^{-1})-1=t^2+t-1$ である。よって $t^2+t-1=0$ となるので、これを解いて $t=\frac{-1\pm\sqrt{5}}{2}$ を得る。
(2) $t=z+z^{-1}$ より $z^2-tz+1=0$ である。$t=\frac{-1+\sqrt{5}}{2}$ のとき、この方程式を解くと

$$z=\frac{t\pm\sqrt{t^2-4}}{2}=\frac{-1+\sqrt{5}\pm\sqrt{(1-\sqrt{5})^2-16}}{4}=\frac{-1+\sqrt{5}\pm\sqrt{-10-2\sqrt{5}}}{4}=\frac{-1+\sqrt{5}\pm\sqrt{10+2\sqrt{5}}\,i}{4}$$

[*15] 0 でない定数は 0 次式ですが、0 だけは次数を $-\infty$ と定めます。これは多項式の次数を deg で表すとき、$\deg fg=\deg f+\deg g$ が常に成り立つようにするためです。

となる。同様に $t=\frac{-1-\sqrt{5}}{2}$ のとき

$$z=\frac{t\pm\sqrt{t^2-4}}{2}=\frac{-1-\sqrt{5}\pm\sqrt{(-1-\sqrt{5})^2-16}}{4}=\frac{-1-\sqrt{5}\pm\sqrt{-10+2\sqrt{5}}}{4}=\frac{-1-\sqrt{5}\pm\sqrt{10-2\sqrt{5}}\,i}{4}$$

が得られる。これらが全ての解である。
(3) $z=e^{2k\pi i/5}$ $(k=0,1,2,3,4)$ が $z^5=1$ の全ての解である。複素平面にプロットすれば、(2) で求めた解のうち実部と虚部がともに正なものが $e^{2\pi i/5}$ だと分かる。これと $\sin(\frac{\pi}{2}-\theta)=\cos\theta$, $\cos(\frac{\pi}{2}-\theta)=\sin\theta$ より

$$\frac{-1+\sqrt{5}+\sqrt{10+2\sqrt{5}}\,i}{4}=e^{2\pi i/5}=\cos\frac{2\pi}{5}+i\sin\frac{2\pi}{5}=\sin\frac{\pi}{10}+i\cos\frac{\pi}{10}$$

となる。この式の実部と虚部を見ればよい。 ■

■多項式: 問 5 の解答　(1) $x^n-a^n=(x-a)(x^{n-1}+ax^{n-2}+a^2x^{n-3}+\cdots+a^{n-1})$
(2) もし x^n-a^n が $x+a$ で割り切れることは、x に $-a$ を代入した結果が 0 になることと同値である。すなわち $0=(-a)^n-a^n=\big((-1)^n-1\big)a^n$ より、$(-1)^n=1$ が必要十分条件である。これは n が偶数であることに他ならない。
(3) x^n+a^n に $x=-a$ を代入すると $(-a)^n+a^n=\big((-1)^n+1\big)a^n$ となる。これが 0 になることは n が奇数であることと同値である。よって n が奇数なら x^n+a^n は $x+a$ で割り切れる。 ■

■円周等分多項式　今の問題で出てきた因数分解 $x^n-a^n=(x-a)(x^{n-1}+ax^{n-2}+\cdots+a^{n-1})$ は非常に良く見かけます。特に $a=1$ と置いてできる多項式 x^n-1 の複素根は、極形式で考えれば直ちに $e^{2k\pi i/n}$ $(k=0,1,\ldots,n-1)$ と分かります。これらの解をプロットすると、原点を中心とする半径 1 の円周が n 等分されます。そういう理由で n が素数のとき[*16]、z^n-1 を $z-1$ で割ってできる多項式 $z^{n-1}+z^{n-2}+\cdots+1$ のことを円周等分多項式と呼びます。

■多項式: 問 7 の解答　答えのみ記す。

(1) $4(ab+cd)^2-(a^2+b^2-c^2-d^2)^2=(a+b+c-d)(a+b-c+d)(a-b+c+d)(-a+b+c+d)$
(2) $x^3-(a+b+c)x^2+(ab+bc+ca)x-abc=(x-a)(x-b)(x-c)$ ■

■基本対称式　今の問題の (2) で出てきた $a+b+c$, $ab+bc+ca$, abc はいずれも a, b, c について対称な多項式です。このような多項式を a, b, c の対称式と言う[*17]のでした。特に、ここに出てきた 3 つの対称式は**基本対称式**という名前がついています。これは「全ての対称式は基本対称式たちの定数倍の和と積で書ける」という重要な定理があるからです。遠くない将来にまた登場すると思いますので、その時に改めて詳しく解説します。

3.3　複素数について

“Was sind und was sollen die Zahlen?”
これは、ドイツの数学者 Richard Dedekind が記した本[*18]のタイトルです。『数とは何か、また何であるべきか？』この問題を、考えてみましょう。

■複素数は数なのか？　我々が日常生活の中で出会う「数」にはどんなものがあるでしょうか？ たとえば物の個数を数えるときは自然数を使いますし、お金の計算をするときは収入と支出を表すのに正の数と負の数を使います。また料理をすればレシピの中に分数が出てきますし、円周の長さを測ろうとしたら π のような無理数も現れます。これらに登場する数は、いずれも実数の範囲に収まっていますね。

[*16] n が素数でないときも円周等分多項式は定義されますが、諸々の事情で $z^{n-1}+z^{n-2}+\cdots+1$ そのものにはなりません。
[*17] 変数の数が増えても同様で、多項式 $f(x_1,\ldots,x_n)$ がどの 2 つの x_i, x_j を入れ替えても変わらないとき、f を n 変数の対称式と言います。
[*18] ちなみに、1893 年に出版されたドイツ語原著の第 2 刷が、東大駒場キャンパスの数理科学図書室にあります。

一方、複素数は実数の範囲を超えるものです。そのため $i^2 = -1$ となる数 i を我々の世界で目にすることはありません。たとえばおつかいに行った子供が「ママー！ おつりで $300 + 250\,i$ 円もらったよ！」なんて言うわけないですね。複素数を「気持ち悪い」と感じる主たる理由は、おそらくここにあるのではないでしょうか。英語では空想上の数 “imaginary number” と呼ばれますし、日本語ではさらにネガティブな含みを持つ「虚」数という呼び名もあります。

ですが「我々の身の回りに見当たらないから」というだけの理由で、複素数を数と呼ぶべきではないのでしょうか。ここで一度、我々の身近にある数について「何が数たらしめているのか」を考えてみましょう。数を考える上で何よりも大事なことは「計算」です。たとえば自然数だったら足し算と掛け算ができます。整数なら、引き算がいつでもできます。有理数や実数なら、0 以外の数による割り算もできます。また計算とは別に「大小の比較」ができることも、数の大きな特徴でしょう。

複素数は残念なことに「大小の比較」をすることはできません。ですが既に見てきたとおり、複素数では四則演算の全てを行うことができます。これをもって「数」と呼んでも良いのではないでしょうか。また「-1 の平方根」というと気味が悪いかもしれませんが、「形式的に i という数を付け加え、$i^2 = -1$ というルールで計算を行う」という風に思えば、i の存在も受け入れられる気がします。

実際、現代数学では今のような方法で複素数を捉えています。一般に実数や有理数のように「四則演算ができる数の範囲」を、数学では**体 (たい)** と呼びます。そしていま、実数係数の 1 変数多項式全体を $\mathbb{R}[x]$ と書きます。このとき多項式 $f(x)$ に対して「x^2 に -1 を代入する」という操作、より正確には「多項式 $f(x)$ を x^2+1 で割った余りを考える」という操作[*19]をすれば、多項式の変数だった x が -1 の平方根であるかのように振る舞います。これを「多項式 x^2+1 の根を添加して実数体 $\mathbb{R}$ を拡大する」と言います。このように体が与えられたとき、その中にない元を付け加えて大きい体を作る**拡大**という操作を定義することによって、きちんと複素数体 $\mathbb{C}$ を定義できるのです。

■複素数: **問 8 の解答**　大体の人が「実数の範疇を飛び出る」とか「大小がなくなる」と書いてくれました。もちろん、どちらも数の性質を踏まえた、真っ当な感じ方だと思います。 ∎

3.4 代数学の基本定理

今回最後の話題は、複素数係数の多項式の根です。多項式 $P(x)$ に対し、方程式 $P(x) = 0$ の解を根と言うのでした。まずは「係数が全て実数」という特別な場合に、複素数の根が「共役とペアで現れる」という事実を確認しましょう。

■代数学の基本定理: **問 1 の解答**　$f(z) = a_0 + a_1 z + \cdots + a_n z^n$ とおく。このとき $a_0, a_1, \ldots, a_n \in \mathbb{R}$ である。よって $\alpha \in \mathbb{C}$ が $f(z)$ の解であるとき、$f(\alpha) = 0$ の共役を取ると

$$0 = \overline{f(\alpha)} = \overline{a_0 + a_1\alpha + \cdots + a_n\alpha^n} = \overline{a_0} + \overline{a_1\alpha} + \cdots + \overline{a_n\alpha^n} = a_0 + a_1\overline{\alpha} + \cdots + a_n\overline{\alpha}^n = f(\overline{\alpha})$$

となる。よって $\overline{\alpha}$ も解である。 ∎

■代数学の基本定理　さて多項式が根を持つかどうかは、「考える数の範囲」によって変わってきます。たとえば $P(x) = x^2 + 1$ は実数の範囲で根を持ちませんが、複素数の範囲に広げると $x = \pm i$ という根を持ちます。このように「多項式がいつも根を持つか」は係数の範囲に依存するのですが、実はなんと、複素数で考える限り**定数でない多項式は必ず** 1 **個は根を持つ**のです[*20]。これが代数学の基本定理と呼ばれる内容です。

[*19] 代数学の用語では「多項式環 $\mathbb{R}[x]$ をイデアル (x^2+1) で割る」という操作に相当します。

[*20] ただし「解を持つ」ことと「解が計算で求まる」ことは全く別の問題です。たとえば f を連続函数とします。このとき $a < b$ とすると、f は閉区間 $[a, b]$ 上のどこかの点 c で最大値を取ります。この事実はグラフを描けばすぐ納得できますが、でも「最大値を取る c がどこにあるか」について、全く教えてくれません。代数学の基本定理も、これと同じ状況になっています。

　ちなみに 2 次方程式は解の公式を使えば常に解けますが、Galois 理論というものを使うと、5 次以上の多項式に対する「解の公式」が存在しないことまで証明できます。

そして「根が 1 個は存在する」ことが分かると、そこから「定数でない n 次多項式はいつも n 個の根を持つ」ことまで言えてしまいます。このことを問題で確かめてから、最後に代数学の基本定理の完全な証明を与えましょう[*21]。

■代数学の基本定理: 問 2 の解答 f の次数に関する帰納法で示す。まず $\deg f = 1$ のときは 1 次式 1 個の積である。次に、$n-1$ 次以下の全ての複素係数多項式が 1 次式の積に分解すると仮定する。このとき $f(z)$ を n 次多項式とすると、代数学の基本定理より $f(z)$ の根が存在する。それを α とすると $f(z) = (z-\alpha)g(z)$ と書け、$g(z)$ は $n-1$ 次多項式となる。帰納法の過程から $g(z)$ は 1 次式の積に分解するので、$f(z)$ 全体も 1 次式の積となる。

■代数学の基本定理: 問 3 の解答 $g(0) \neq 0$ より、$g(z)$ の定数項は 0 でない。定数項の次に低い次数の項が $a_k z^k$ だとして、$g(z) = a_0 + a_k z^k + a_{k+1}z^{k+1} + \cdots + a_n z^n$ とおく。うまく z の偏角を調節すれば a_0 と $a_k z^k$ の偏角が π だけずれ、原点から見て逆向きになるようにできる。その上で $|z|$ を十分小さくすれば、$|a_0 + a_k z^k|$ は $|a_0|$ より小さくなるはずである。さらに $|z|$ を小さくすれば、z の $(k+1)$ 乗以上の項の絶対値は、z^k の項の絶対値に比べていくらでも小さくなる。そうすれば $|g(z_0)| < |g(0)|$ となるはずである。この考え方に基づき、z の適切な偏角と絶対値を見出そう。

いま $a_0 = r_0 e^{i\theta_0}$, $a_k = r_k e^{i\theta_k}$ $(r_0, r_k > 0)$ と書ける。これらを用いて $z_0 := re^{i(\theta_0 - \theta_k + \pi)/k}$ と定めると

$$
\begin{aligned}
g(z_0) &= a_0 + a_k z_0^k\Bigl(1 + \frac{a_{k+1}}{a_k}z_0 + \cdots + \frac{a_n}{a_k}z_0^{n-k}\Bigr) = r_0 e^{i\theta_0} + r_k e^{i\theta_k}\Bigl(re^{i(\theta_0-\theta_k+\pi)/k}\Bigr)^k\Bigl(1 + \frac{a_{k+1}}{a_k}z_0 + \cdots + \frac{a_n}{a_k}z_0^{n-k}\Bigr) \\
&= r_0 e^{i\theta_0} + r_k r^k e^{i(\theta_0+\pi)}\Bigl(1 + \frac{a_{k+1}}{a_k}z_0 + \cdots + \frac{a_n}{a_k}z_0^{n-k}\Bigr) = (r_0 - r_k r^k)e^{i\theta_0} - r_k r^k e^{i\theta_0}\Bigl(\frac{a_{k+1}}{a_k}z_0 + \cdots + \frac{a_n}{a_k}z_0^{n-k}\Bigr)
\end{aligned}
$$

となる。よって $|e^{i\theta_0}| = 1$ と三角不等式より、$|f(z_0)|$ は

$$
\begin{aligned}
|g(z_0)| &\leq |r_0 - r_k r^k| + r_k r^k\Bigl|\frac{a_{k+1}}{a_k}z_0 + \cdots + \frac{a_n}{a_k}z_0^{n-k}\Bigr| = |r_0 - r_k r^k| + r_k r^k |z_0|\Bigl|\frac{a_{k+1}}{a_k} + \frac{a_{k+2}}{a_{k+1}}z_0 + \cdots + \frac{a_n}{a_k}z_0^{n-k-1}\Bigr| \\
&\leq |r_0 - r_k r^k| + r_k r^{k+1}\Bigl\{\Bigl|\frac{a_{k+1}}{a_k}\Bigr| + \Bigl|\frac{a_{k+2}}{a_k}z_0\Bigr| + \cdots + \Bigl|\frac{a_n}{a_k}z_0^{n-k-1}\Bigr|\Bigr\}
\end{aligned}
$$

となる。この式で

$$
M := \max\Bigl\{\Bigl|\frac{a_{k+1}}{a_k}\Bigr|, \Bigl|\frac{a_{k+2}}{a_k}z_0\Bigr|, \ldots, \Bigl|\frac{a_n}{a_k}z_0^{n-k-1}\Bigr|\Bigr\}
$$

とおくと、$|g(z_0)| \leq |r_0 - r_k r^k| + r_k r^{k+1}(n-k-1)M$ が得られる。そこで

$$
r := \min\Bigl\{\Bigl|\frac{r_0}{r_k}\Bigr|^{\frac{1}{k}}, \frac{1}{2(n-k-1)M}\Bigr\}
$$

と取れば、$r_0 - r_k r^k \geq 0$ かつ $r \leq \frac{1}{2(n-k-1)M}$ なので

$$
|g(z_0)| \leq r_0 - r_k r^k + r_k r^{k+1}(n-k-1)M \leq r_0 - r_k r^k\bigl(1 - r(n-k-1)M\bigr) \leq r_0 - \frac{r_k r^k}{2} < r_0
$$

である。これより $|g(z_0)| < r_0 = |g(0)|$ となることが分かった。 ■

■代数学の基本定理の証明 せっかく問 3 の解答を与えたので、代数学の基本定理の証明を与えておきます。以下、$f(z)$ を複素係数の多項式とします。証明のアイデアは次の通りです。

1. $f(z) = 0$ となる z が存在することを「z が複素平面全体を動く時の $|f(z)|$ の最小値が 0」と言い換える。
2. 原点を中心とする十分大きい半径 R の円板を考え、その外側には $|f(z)|$ の最小値が決して現れないことを言う。
3. 半径 R の円板内のどこかで、$|f(z)|$ が最小値を取ることを言う。
4. $|f(z)|$ の最小値が 0 でないと、矛盾が生じることを示す。

[*21] この証明には微分積分学の知識が必要なので、今は読み解くのが難しいかもしれません。連続函数の性質を習った後で読むと、手頃な勉強になるでしょう。

step 1. $|z| > R$ なる全ての複素数 z に対して $|f(z)| > |f(0)|$ が成り立つような正の実数 $R > 0$ が取れることを示す。

$f(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0$ とおく。このとき

$$|f(z)| = |a_n||z|^n \left|1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n}\right|$$

である。ここで三角不等式 $|z_1| - |z_2| \leq |z_1 - z_2|$ を使うと、$|z_1 + z_2| \geq |z_1| - |-z_2| = |z_1| - |z_2|$ なので

$$\left|1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n}\right| \geq \left|1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_1}{a_n z^{n-1}}\right| - \left|\frac{a_0}{a_n z^n}\right| \geq \cdots \geq 1 - \left|\frac{a_{n-1}}{a_n z}\right| - \cdots - \left|\frac{a_0}{a_n z^n}\right|$$

となる。ここで

$$R_0 := \max\left\{\frac{n|a_{n-1}|}{0.1|a_n|}, \left(\frac{n|a_{n-2}|}{0.1|a_n|}\right)^{\frac{1}{2}}, \ldots, \left(\frac{n|a_0|}{0.1|a_n|}\right)^{\frac{1}{n}}\right\}$$

とおく[*22]と、$|z| > |R_0|$ のとき、全ての $1 \leq k \leq n-1$ に対し

$$\left|\frac{a_{n-k}}{a_n z^k}\right| = \frac{|a_{n-k}|}{|a_n||z|^k} < \frac{|a_{n-k}|}{|a_n|R_0^k} \leq \frac{|a_{n-k}|}{|a_n|}\left(\frac{0.1|a_n|}{n|a_{n-k}|}\right)^{\frac{1}{k}\cdot k} = \frac{0.1}{n}$$

となる。よって

$$\left|1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n}\right| \geq 1 - \left|\frac{a_{n-1}}{a_n z}\right| - \cdots - \left|\frac{a_0}{a_n z^n}\right| > 1 - \frac{0.1}{n} - \cdots - \frac{0.1}{n} = 0.9$$

となる。さらに正の実数 R を

$$R := \max\left\{\left(\frac{|f(0)|}{0.9|a_n|}\right)^{\frac{1}{n}}, R_0\right\}$$

とおくと、$|z| > R$ のとき

$$|f(z)| = |a_n||z|^n \left|1 + \frac{a_{n-1}}{a_n z} + \cdots + \frac{a_0}{a_n z^n}\right| > |a_n|\left(\frac{|f(0)|}{0.9|a_n|}\right)^{\frac{1}{n}\cdot n} \cdot 0.9 = |f(0)|$$

となる。

step 2. 複素平面上の実数値函数 $|f(z)|$ が最小値を持つことを示す。

半径 R の閉円板 $\Delta_R := \{z \in \mathbb{C} \mid |z| \leq R\}$ を考える。$|f(z)|$ は $\mathbb{C} = \mathbb{R}^2$ 上の連続函数である。そして Δ_R は平面 $\mathbb{R}^2$ 内の有界閉集合だから、$|f(z)|$ は Δ_R 上で最小値を取る。その点を z_0 とする。いま $|z| > R$ とすると、R の取り方から $|f(z)| > |f(0)|$ が従う。一方 $0 \in \Delta_R$ より $|f(z_0)| \leq |f(0)|$ である。これらより $|z| > R$ のときも $|f(z)| > |f(z_0)|$ が従う。かくして $|f(z_0)|$ は、複素平面全体における $|f(z)|$ の最小値である。

step 3. 方程式 $f(z) = 0$ が解を持つことを示す。

もし $|f(z_0)| \neq 0$ だとすると、$f(z_0) \neq 0$ である。そこで新しい多項式 $g(z)$ を $g(z) := f(z + z_0)/f(z_0)$ で定めることができる。このとき $g(0) = 1$ なので、問 3 の結果より、適当な複素数 $z_1 \in \mathbb{C}$ を取ると $|g(z_1)| < 1$ とできる。ところが $|g(z_1)| = |f(z_1 + z_0)|/|f(z_0)|$ と合わせると $|f(z_1 + z_0)| < |f(z_0)|$ が導かれてしまう。これは $|f(z_0)|$ が複素平面全体で $|f(z)|$ の最小値になることに矛盾する。ゆえに $|f(z_0)| = 0$ である。つまり、z_0 は $f(z)$ の根である。 ∎

[*22] 0.1 という数に本質的な意味はありません。この 0.1 は、1 未満の任意の正の数 ε に置き換えて構いません。

数理科学基礎 (線形代数学) 第 2 回　　種々の函数

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 4 月 22 日 1 限

1　全体的な講評

今回のテーマは「高校生のうちに一応習いはするけど、でもあんまり深入りはしない函数たち」です。逆三角函数や双曲線函数について、名前自体は授業前から知っていた人が多かったのではないでしょうか。いずれの問題も、そうした函数の性質を知るためのものです。

問題は比較的易しいものが多かったので、ほとんどの人が良く解けていました。中には一つの問題に複数の方法で取り組んだり、あるいは関連する問題をまとめて解いたりした人がいました。大変良いことだと思います。普通に問題を解くのに加え、問題を解く過程から何かを見出せるよう頑張ってください。

また今回、答案の中に質問やリクエストを書いてくれた人もいました。そういうことが書いてあると解説プリントを作るときの指針になりますので、何か思うことがあれば、ぜひ答案内に書き込んでおいてください。

それでは、今回のテーマを一つずつ見ていきましょう。

2　逆函数について

2.1　逆函数の定義

函数とは、数に対して何かしらの数を対応させる規則のことです。たとえば $f(x) = x^2$ という函数は、実数 $x \in \mathbb{R}$ に $x^2 \in \mathbb{R}$ という数を対応させています。他にも三角函数、対数函数や指数函数といった有名な函数があり、そしてこれらの函数を合成するとさらに色々な函数を作ることができますが、どの函数も**数に数を対応させる**ことには間違いありません。別の言い方をすれば、僕たちが普段「函数」と呼んでいるものは、定義域と値域が実数全体 $\mathbb{R}$ (の部分集合) であるような写像です[*1]。

さて函数 $f(x)$ が与えられると、数 x が与えられた時に f の値 $f(x)$ が定まります。そしてしばしば「数 y が与えられたとき、$y = f(x)$ となる数 x を求めたい」という状況が生じます。対応 $x \mapsto f(x)$ ではなく、その逆の対応を求めたいわけです。このように函数 f が与えられたとき、「数 y に対し、$y = f(x)$ となる x を対応させる」という方法で定義される函数を f の**逆函数**と言い、f^{-1} と書きます。$x \mapsto f(x)$ の対応が逆になるので、逆函数のグラフは元の函数のグラフを直線 $y = x$ について線対称に折り返したものになります。

2.2　逆函数の主値

■逆函数を定義するための条件　残念なことに、与えられた函数に対していつでも逆函数が無条件で定義できるわけではありません。上で述べたように、函数は写像ですから「1 つの数に、1 つの数を対応させる」ことが大原則です。だから「1 つの数に 2 つの数を対応させる」という操作になってしまっては、函数とは呼べないのです。

いま、函数 f が相異なる実数 x_1, x_2 に対して $f(x_1) = f(x_2)$ を満たしたとしましょう。そして $y = f(x_1) = f(x_2)$ とおきます。すると x_1, x_2 のどちらに対しても y が対応するので、逆函数を作ろうとしたら「y に x_1, x_2 の両方が対応する」というマズい状況が実際に起きます。こうならないためには「異なる実数 x_1, x_2 に対して必ず $f(x_1) \neq f(x_2)$

[*1] ただし、この用語の使い方は慣用的なもので、厳密な取り決めがあるわけではありません。たとえば定義域や値域がベクトルの集合であっても、函数と呼ぶことはままあります。

が成り立つ」という条件が必要です。こういう性質を **単射性** というのでした。

また単射なら良いかというと、そうでもありません。たとえば指数函数 $\exp x := e^x$ を考えましょう。指数函数を $\exp\colon \mathbb{R} \to \mathbb{R}$ という写像と思えば、これは単射です。グラフを描けば狭義短調増加ですから、異なる値に対して同じ値が対応しようがありません。そして正の実数 $y > 0$ に対しては、$e^x = y$ となる x が $x = \log y$ で与えられます。ですが $y \leq 0$ のとき、$y = \exp x$ となる $x \in \mathbb{R}$ は存在しません。さっきは「1 つの数に 2 個 (以上) の数が対応してしまう」ことが問題でしたが、今度は「ある数に対して、対応させせられる数がなくなってしまった」という問題が起きています。これももちろん写像の定義を満たさないので、指数函数 exp を $\mathbb{R}$ から $\mathbb{R}$ への写像と思うと、逆函数は作れません。ですが exp の値域を正の整数全体の集合 $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$ だと思えば、どんな正の数 $y > 0$ に対しても $x := \log y$ とおくことで $y = e^x$ なる実数 x を見つけられます。このような、函数 f の「どんな値域の元 y に対しても、$y = f(x)$ となる x が存在する」という性質を、**全射性**と呼ぶのでした。

まとめると、f **の逆函数が存在するための必要条件は、**f **が全単射であること**です。また f の逆函数 g が存在すれば、f は全単射になると示せます。かくして f の逆函数が存在することと、f が全単射であることが同値だと分かります。

■逆函数の主値　ただし f が全単射でないときも、部分的には f の逆函数を作れることが多いです。さっきの指数函数の例では、値域を $\mathbb{R}$ から $\mathbb{R}_{>0}$ に取り換えれば逆函数が作れました。また f の定義域を制限してしまうのも一つの手です。たとえば $f(x) = x^2$ という函数を考えましょう。正の実数 $x > 0$ に対し $f(-x) = (-x)^2 = x^2 = f(x)$ という式が成り立ってしまうので、f は単射ではありません。ですが f の定義域を非負実数全体の集合 $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$ に制限してしまえば、f は単射になります。そしてどんな非負の実数 $y \in \mathbb{R}_{\geq 0}$ に対しても、$x := \sqrt{y}$ とおけば $y = x^2 = f(x)$ が成り立ちます。こうして f を制限して $\mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ という写像にすれば、f の逆函数 $y \mapsto \sqrt{y}$ が得られます。f の定義域を制限して逆写像 f^{-1} を作るとき、その値を f^{-1} の**主値**と言います。

ちなみに複素数の偏角 arg の「主値」という言葉遣いも、函数の主値と同じ意味です。0 以外の複素数全体の集合を $\mathbb{C}^\times := \{z \in \mathbb{C} \mid z \neq 0\}$ と書きます。すると $z \in \mathbb{C}^\times$ の偏角 $\arg z$ は 2π 単位でずらせるので、主値を決めないと普通の函数にはなりません。そこで主値を半開区間 $[0, 2\pi)$ とすれば、偏角は $\arg\colon \mathbb{C}^\times \to [0, 2\pi)$ という函数になります[*2]。

2.3　逆函数の微分

函数 f が微分可能で逆函数 f^{-1} が定義される場合、f^{-1} も大体は微分可能になります。いま $y_0 = f(x_0)$ だったとします。このとき $f'(x_0)$ は、$y = f(x)$ のグラフの点 (x_0, y_0) における接線 ℓ の傾きです。一方、$y = f(x)$ のグラフを直線 $y = x$ について折り返せば逆函数 $x = f^{-1}(y)$ のグラフが得られます。したがって $y = f(x)$ のグラフと一緒に接線 ℓ を折り返せば、$x = f^{-1}(y)$ の点 (y_0, x_0) における接線が得られます。その接線の傾きは ℓ の接線の傾きの逆数になります。したがって $(f^{-1})'(y_0) = 1/f'(x_0)$ となります。

逆函数の微分公式をちゃんと示すには少し神経を使いますから、ここでは証明はしません。でも、式自体は使いやすいと思います。また忘れても「グラフの接線を $y = x$ について折り返す」というイメージさえ覚えておけば、必要な場面で思い出せるはずです。

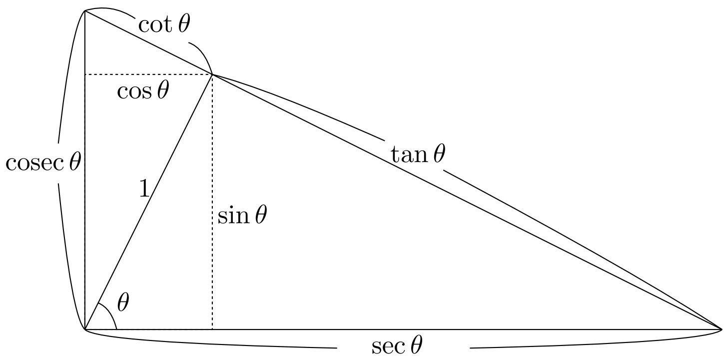
3　逆三角函数と円周率の近似公式

3.1　逆三角函数の定義

念の為に三角函数の定義を確認しておきます。単位円周上の点 P に対し、x 軸の非負の部分と半直線 OP のなす角が θ であるとき、P の座標を $(\cos\theta, \sin\theta)$ と表すのでした。そして残りは $\tan\theta := \sin\theta/\cos\theta$, $\cot\theta := 1/\tan\theta$,

[*2] ですが無理やり函数にした代償で、実軸正の部分での連続性が破れることに注意しましょう。偏角 0 付近で連続性を保ちたい場合は、主値を $(-\pi, \pi]$ などに取り直す必要があります。また一般に、どのような主値を選んでも arg を $\mathbb{C}^\times$ 全体で連続な函数にはできません。

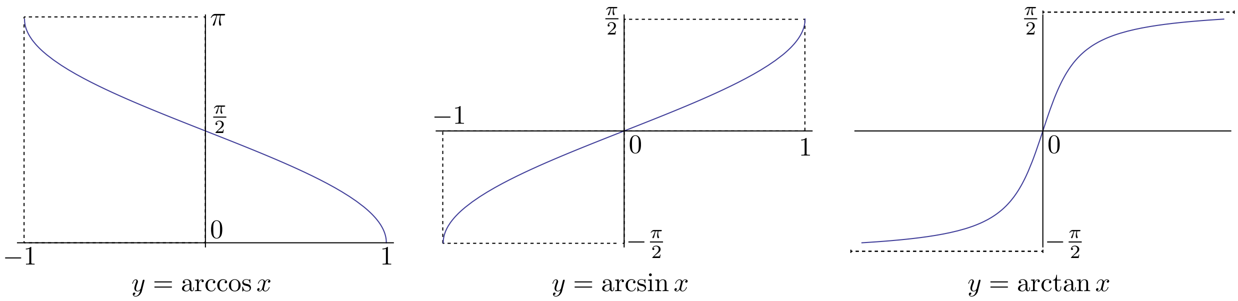
$\sec\theta := 1/\cos\theta$, $\operatorname{cosec}\theta := 1/\sin\theta$ と定めます。これを図に描くと次のようになります。



三角函数の逆函数が逆三角函数です。cos, sin, tan の逆函数をそれぞれ arccos, arcsin, arctan と書きます。ただし三角函数はすべて周期函数なので、ある意味「単射から最も遠い函数」です。たとえば $\sin x = 1/2$ となる実数 x は 2 個どころか無限個存在してしまいます。ですから先ほど説明したように、逆三角函数を考えるには、三角函数の定義域を制限して主値を選ぶ必要があります。通常、主値は次のように選びます。

函数	定義域	主値
arccos	$[-1,1]$	$[0,\pi]$
arcsin	$[-1,1]$	$[-\pi/2,\pi/2]$
arctan	$\mathbb{R}$	$(-\pi/2,\pi/2)$

この主値に関するグラフは次の通りです。直線 $y = x$ について折り返すと、元の三角函数のグラフになりますね。



この中で一番よく使うのは arctan だと思います。というのも直線 tan は角度から傾きを計算するのに使う函数ですから、arctan は「与えられた傾きを実現する角度は何度か？」を表しています。いかにも役立ちそうですよね。

3.2　$\arctan 1$ の公式たち

三角函数に色々な公式があるように、逆三角函数にも色々な公式があります。その中でも arctan の計算公式は比較的有名です。というのも後で見るように、arctan の計算ができると、$\pi/4 = \arctan 1$ を使って円周率を求めることができるからです。$\arctan 1$ を色々な方法で表してみましょう。

■問 2, 3 の解答　tan の加法定理

$$\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$$

に $u = \tan x$, $v = \tan y$ を代入すると

$$\arctan u + \arctan v = x + y = \arctan\tan(x+y) = \arctan\frac{\tan x + \tan y}{1 - \tan x \tan y} = \arctan\Big(\frac{u+v}{1-uv}\Big)$$

を得る。この式で $u = 1/2$, $v = 1/3$ とおくと

$$\arctan\frac{1}{2} + \arctan\frac{1}{3} = \arctan\Big(\frac{\frac{1}{2}+\frac{1}{3}}{1-\frac{1}{6}}\Big) = \arctan 1 = \frac{\pi}{4}$$

が得られる。これは Euler の公式と呼ばれる。

また $u = v = 1/5$ とおくと、同じように計算することで

$$2\arctan\frac{1}{5} = \arctan\frac{\frac{2}{5}}{1-\frac{1}{25}} = \arctan\frac{5}{12}$$

となる。さらに $u = v = 5/12$ に対しては

$$2\arctan\frac{5}{12} = \arctan\frac{\frac{10}{12}}{1-\frac{25}{144}} = \arctan\frac{120}{119}$$

である。そして $u = 120/119$ に対し $(u+v)/(1-uv) = 1$ となる v を求めると、$v = (1-u)/(1+u) = -1/239$ となる。これらをまとめると、Machin の公式

$$\frac{\pi}{4} = \arctan\frac{\frac{120}{119}+\frac{-1}{239}}{1-\frac{120}{119}\frac{-1}{239}} = \arctan\frac{120}{119} + \arctan\frac{-1}{239} = 2\arctan\frac{5}{12} - \arctan\frac{1}{239} = 4\arctan\frac{1}{5} - \arctan\frac{1}{239}$$

を得る[*3]。 ■

■問 4 の解答　計算自体は簡単である。

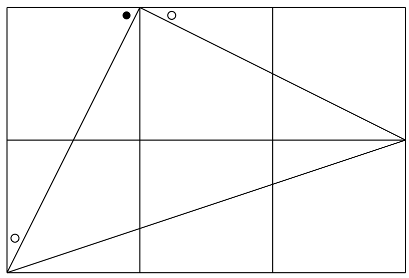
$$(1)\ (3+i)(7+i) = 20+10i, \qquad (2)\ (2+i)(3+i) = 5+5i, \qquad (3)\ (5+i)^4/(239+i) = 2+2i$$

ただ、この問題で重要なのは計算結果から arctan の公式が得られる点である。複素数 $z = x+yi$ の偏角が $\arctan(y/x)$ で与えられること、また偏角が $\arg zz' = \arg z + \arg z'$ という式を満たすことを思い出そう。両辺の偏角を取れば

$$\arctan\frac{1}{3} + \arctan\frac{1}{7} = \arctan\frac{1}{2}, \quad \arctan\frac{1}{2} + \arctan\frac{1}{3} = \arctan 1 = \frac{\pi}{4}, \quad 4\arctan\frac{1}{5} - \arctan\frac{1}{239} = \arctan 1 = \frac{\pi}{4}$$

が得られる。 ■

■問 5 の解答　次の図を見れば、求める角度は直角二等辺三角形の内角だから 45 度と分かる。



この図もまた、$\arctan(1/2) + \arctan(1/3) = \pi/4$ を表している。 ■

[*3] 途中で 1 回、主値 $(-\pi/2, \pi/2)$ のもとで arctan が奇函数であることを使います。

3.3 円周率の計算

arctan には、実は次のような近似公式があります:

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \quad (\text{if } -1 \le x \le 1)$$

この公式は、形式的には

$$\arctan x = \int_0^x (\arctan t)' dt = \int_0^x \frac{1}{1+t^2} dt = \int_0^x (1 - t^2 + t^4 - t^6 + \cdots) dt = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots$$

のようにして求められます[*4]。右辺は気合で頑張ればいくらでも精度よく計算できますから、この公式をガリガリ計算することで $\arctan x$ の値を求めることができます。たとえば $x = 1$ とすれば

$$\frac{\pi}{4} = \arctan 1 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots$$

という風にして円周率が求められます。これは Leibniz の公式と呼ばれているようです。

ただ、今の $\pi/4$ 公式は非常に効率が悪いです。というのも第 50 番目の項が $-1/99 = 0.010101\cdots$ ですから、50 番目の項を計算すると小数第 2 位の値が変わります。円周率を求める上では「ちょこっと計算しただけで、上の方の桁が正確に分かる」ような公式が望ましいわけで、こんな「50 項も計算してもまだ小数第 2 位が分からない」なんて公式は役に立ちません。そこで登場するのが Euler の公式や Machin の公式です。たとえば Euler の公式なら

$$\begin{aligned}\frac{\pi}{4} &= \arctan\frac{1}{2} + \arctan\frac{1}{3} = \left(\frac{1}{2} - \frac{1}{3\cdot 2^3} + \frac{1}{5\cdot 2^5} - \cdots\right) + \left(\frac{1}{3} - \frac{1}{3\cdot 3^3} + \frac{1}{5\cdot 3^5} - \cdots\right)\\ &= \left(\frac{1}{2} - \frac{1}{24} + \frac{1}{160} - \cdots\right) + \left(\frac{1}{3} - \frac{1}{81} + \frac{1}{1215} - \cdots\right)\end{aligned}$$

のように、分母にある x^n の形の項が効いてきて、足し引きされる項の値が急激に小さくなっていきます。今の場合、最初から 3 つずつの項を拾うだけで $\pi = 3.14\cdots$ が求まります。さっきより断然楽ですね。Machin の公式に至っては 1/239 という項がありますから、少ない手間でもっと精度よい値を求められます。William Jones という数学者が 1706 年に著した "Synopsis Palmariorum Mathesos" という本[*5]に、Machin が求めたとされる円周率が結果だけ 100 桁以上載っていますが、今の公式を使ったのでしょうか？

もちろん arctan に放り込む数が小さくなれば小さくなるほど、計算はどんどん楽になります。したがって「π が上手く出てくる arctan の公式を見つけて、君だけのオリジナルの円周率近似公式をつくろう！」…という話になりそうですが、実はこの公式は頑張ってもそんなに精度が出ません。いまの公式では分母に x^n が出てくることがキモだったので「1 つ先の項を計算したときに、何桁分だけ精度が良くなるか」は毎回同じです。もし x が 100 だったら、次の項を考えてもせいぜい 2 桁程度しか精度が良くならないわけです。ところが世の中には「計算を 1 手進めるだけで、**それまでの桁数の倍の桁数**だけ精度が良くなる」という、圧倒的に強い公式があります。Gauss–Legendre アルゴリズムといいますので、興味のある人は調べてみてください[*6]。

ちなみに、プリントをアップロードしている GitHub のページ[*7]に、Excel で円周率の近似計算の実験をした結果 `pi_approx.xlsx` を載せてみました。いくつかの公式に対し n 番目の項の値、n 番目の近似値と真の値からの誤差を表にしています。興味がある人は、このデータの誤差を対数プロットしてみると良いでしょう。ただし Excel は高精度計算には向かないので、途中で精度が打ち止めになっています。本気で円周率を求めるには、プログラムを書き、必要に応じて高精度計算のためのライブラリを援用しないといけません。

[*4] この「証明」は全然厳密ではありません。何も考えずに無限和と積分の順序を入れ替えてはいけないからです。Taylor 展開を習った後なら、こんなことをしなくても公式が導けるようになります。

[*5] ちなみに本郷キャンパスにある総合図書館の書庫内に、この本があるそうです。OPAC で検索すると出てきます。

[*6] たとえば四ツ谷 晶二・村井 実『楕円関数と仲良くなろう』(日本評論社) の第 2 章に詳しい説明があります。

[*7] `https://github.com/HideakiHosaka/2015_linear_algebra`

4　双曲線函数と逆双曲線函数

最後のテーマは、双曲線函数と呼ばれる函数たちです。これらは指数函数の四則演算で書けるという意味では、そこまで目新しいものではありません。ですが三角函数と良く似た性質を持ち、かつ置換積分などの際に役立つので、性質を知っておくと便利です。

4.1　双曲線函数

■定義　次に挙げる函数を、双曲線函数といいます。

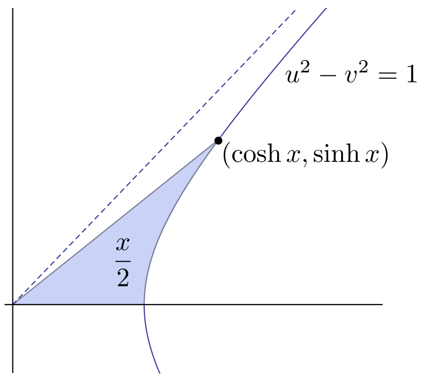
$$\cosh x := \frac{e^x + e^{-x}}{2}, \qquad \sinh x := \frac{e^x - e^{-x}}{2}, \qquad \tanh x := \frac{\sinh x}{\cosh x}$$

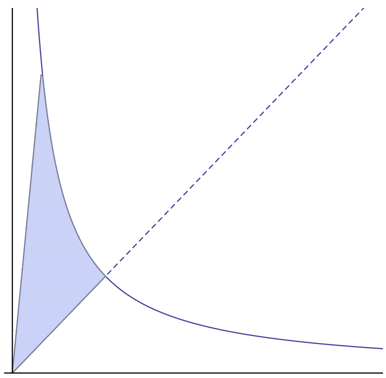
名前の付け方のルールは三角函数と同じです[*8]。

これらが双曲線函数と呼ばれる所以の一つは、双曲線のパラメータ表示を与えるからです。三角函数は $\cos^2 t + \sin^2 t = 1$ という恒等式を満たしましたが、双曲線函数では $\cosh^2 t - \sinh^2 t = 1$ という式が成り立ちます。したがって t が動けば、点 $(\cosh t, \sinh t)$ は双曲線 $x^2 - y^2 = 1$ の上を動きます。ただし $\cosh t > 0$ なので、このパラメータ表示は双曲線の右側だけしか与えていません。双曲線の左側のパラメータ表示は $(-\cosh t, \sinh t)$ で与えられます。

他にも、双曲線を使って cosh, sinh に図形的方法で定義を与えることもできます。

■問 6 の解答　双曲線 $u^2 - v^2 = 1$ 上に、下図で塗りつぶされた部分の面積が $x/2$ となるような点を取る。この点の座標を $(\cosh x, \sinh x)$ と書く。$\cosh x$ と $\sinh x$ の明示的な式を求めよう。面積の計算には積分が必要である。積分計算を楽にするため、この双曲線を正の向き[*9]に 45° 回転させる。そうすると双曲線 $u^2 - v^2 = 1$ は、直線 $v = u$ 上で原点からの距離が 1 である点、すなわち点 $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ を通る反比例のグラフになる。よって回転後の双曲線の式は $\eta = \frac{1}{2\xi}$ である[*10][*11]。





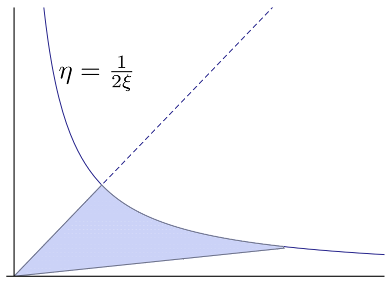
縦より横に積分した方が見やすいので、面積を求めるべき部分を直線 $\eta = \xi$ について折り返す。すると、次の二つの図で示された面積は等しいことが分かる。$\mathrm{AB} \cdot \mathrm{BC} = \mathrm{PQ} \cdot \mathrm{CQ} = \frac{1}{2}$ が成り立ち、$\triangle\mathrm{ABC}$ と $\triangle\mathrm{CQP}$ の面積が等しくなるからである。

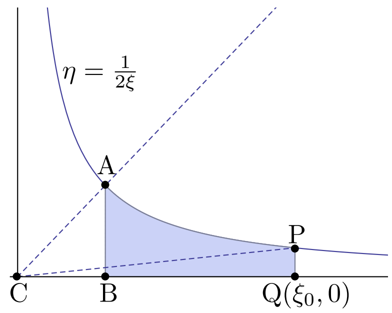
[*8] 三角函数の cot, sec, cosec に対応して双曲線函数でも coth, sech, cosech というのが定義されますが、使う場面はまず無いでしょう。

[*9] 数学でいう「正の向き」は、反時計回りです。

[*10] ξ, η はそれぞれギリシャ文字のグザイとエータです。

[*11] 軸を u, v から ξ, η に変えたことに、深い意味はありません。





さて点 Q の ξ 座標を ξ_0 とおく。B の ξ 座標は $\frac{1}{\sqrt{2}}$ なので、積分で面積を計算すると

$$\frac{x}{2} = \int_{\frac{1}{\sqrt{2}}}^{\xi_0} \frac{d\xi}{2\xi} = \frac{1}{2}\left(\log \xi_0 - \log \frac{1}{\sqrt{2}}\right) = \frac{1}{2} \log\left(\sqrt{2}\xi_0\right)$$

と求まる。一方で点 P は、点 $(\cosh x, \sinh x)$ を x 軸について折り返してから正の向きに 45° 度回転さることで得られるのであった。ゆえに、この平面を複素平面と同一視することで P の座標が

$$\overline{\cosh x + i \sinh x} \times e^{i\pi/4} = (\cosh x - i \sinh x)\left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}i\right) = \frac{\cosh x + \sinh x}{\sqrt{2}} + \frac{\cosh x - \sinh x}{\sqrt{2}} i$$

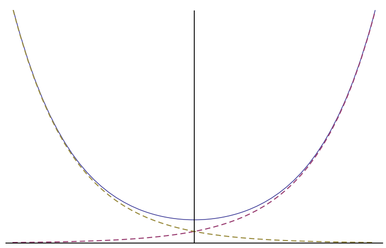
と分かる[*12]。この実部が ξ_0 なので、面積の計算と合わせて $x = \log(\sqrt{2}\xi_0) = \log(\cosh x + \sinh x)$ が従う。つまり $\cosh x + \sinh x = e^x$ である。

最後に、点 $(\cosh x, \sinh x)$ は双曲線 $u^2 - v^2 = 1$ の上にあったことを思い出す。これより

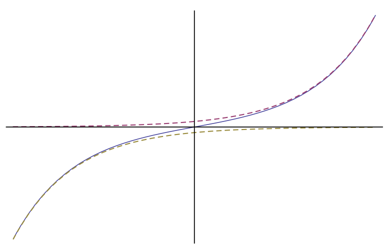
$$1 = \cosh^2 x - \sinh^2 x = (\cosh x - \sinh x)(\cosh x + \sinh x) = (\cosh x - \sinh x)e^x$$

という式が得られる。すなわち $\cosh x - \sinh x = e^{-x}$ である。かくして $\cosh x + \sinh x = e^x$ と合わせて $\cosh x$ と $\sinh x$ の連立 1 次方程式が得られたので、これを解いて $\cosh x$ と $\sinh x$ の求める式を得る。 ∎

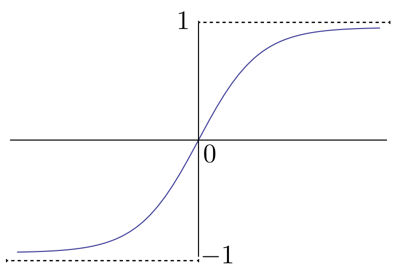
■グラフ　双曲線函数のグラフは次の通りです。$\cosh x$ と $\sinh x$ のグラフに点線で描き込まれているのは、函数 $e^x/2$ と $\pm e^{-x}/2$ です。



$y = \cosh x$



$y = \sinh x$



$y = \tanh x$

本体だけ見ると $\cosh x$, $\sinh x$ はそれぞれ 2 次、3 次函数のグラフに似てなくもないです。しかし中に指数函数がいるので、$x \to \pm\infty$ での絶対値の増大度が全然違います。気を付けましょう。また $\cosh x$ のグラフには「懸垂線」という名前が付いています。これは、糸の両端を持ってぶら下げた時にできる曲線が cosh で書けることに由来します。

■問 6 (d) の解答と導函数　計算すると一瞬で $(\cosh x)' = \sinh x$, $(\sinh x)' = \cosh x$ が分かります。これも符号の付き方が微妙に違うだけで、三角函数とよく似ています。また今の計算から、$y = \cosh x$ と $y = \sinh x$ は両方とも微分方程式 $y'' = y$ の解だとが分かります。これも、単振動の方程式 $y'' = -y$ の解が三角函数で得られることと似ています。

[*12] 回転行列のことを知っているなら、それを使ってももちろん同じ計算ができます。

■問 6 (b) の解答と加法定理　双曲線函数に関しても、三角函数と同様の加法定理が成り立ちます。ただし、符号の付き方が三角函数のときと少しだけ変わります。

$$\begin{aligned}\cosh x\cosh y+\sinh x\sinh y&=\frac{e^x+e^{-x}}{2}\frac{e^y+e^{-y}}{2}+\frac{e^x-e^{-x}}{2}\frac{e^y-e^{-y}}{2}\\&=\frac{e^{x+y}+e^{x-y}+e^{-x+y}+e^{-(x+y)}}{4}+\frac{e^{x+y}-e^{x-y}-e^{-x+y}+e^{-(x+y)}}{4}\\&=\cosh(x+y)\\\sinh x\cosh y+\cosh x\sinh y&=\frac{e^x-e^{-x}}{2}\frac{e^y+e^{-y}}{2}+\frac{e^x+e^{-x}}{2}\frac{e^y-e^{-y}}{2}\\&=\frac{e^{x+y}+e^{x-y}-e^{-x+y}-e^{-(x+y)}}{4}+\frac{e^{x+y}+e^{x-y}-e^{-x+y}-e^{-(x+y)}}{4}\\&=\sinh(x+y)\end{aligned}$$

加法定理が成り立ちますから、当然ながら n 倍角の公式や和積・積和公式の類も三角函数のやつを少し修正するだけで得られます。これらの公式は、たまに双曲線函数の掛け算を積分する際に役立ちます。大学入試のときみたいに頑張って公式を暗記する必要まではありませんが、「三角函数と同じ公式がある」ことは頭の片隅に置いておきましょう。

■双曲線函数と三角函数の関係　さて「三角函数と双曲線函数は似ている」ということを延々見てきたわけですが、なぜこんなにも似ているのでしょうか。それは定義域を複素数まで広げてみると分かります。前回 $e^{i\theta}=\cos\theta+i\sin\theta$ という式を紹介しました。そして双曲線函数は指数函数を使って定義されていますから、$e^{i\theta}$ の式を使えば、双曲線函数に複素数を放り込めます。「何で指数函数に複素数を入れていいのか」は全く説明していませんが、深いことは気にしないでおきましょう。実際にやってみると

$$\cosh i\theta=\frac{e^{i\theta}+e^{-i\theta}}{2}=\cos\theta,\qquad\qquad \sinh i\theta=\frac{e^{i\theta}-e^{-i\theta}}{2}=i\sin\theta$$

というように、なんと三角函数が出てきます。つまり複素数 $z\in\mathbb{C}$ に対し $(e^z\pm e^{-z})/2$ という函数を考えたとき[*13]、この函数は実軸上では双曲線函数に、虚軸上では三角函数 (の i 倍) に化けるのです。このように cosh と cos, sinh と sin の共通の親玉となる函数がいるので、似たような挙動を示していたのです。

4.2　逆双曲線函数

双曲線函数の逆函数を逆双曲線函数といい、三角函数の時と同様、頭に “arc” を付けて arccosh などと表します。これらの函数は、log と平方根で表せます。したがって微分も、今まで知っている知識だけでてできます。

■問 6 (c) の解答　$y=\cosh x=(e^x+e^{-x})/2$ とおくと、$(e^x)^2+1=2ye^x$ である。よって e^x の 2 次方程式 $(e^x)^2-2ye^x+1=0$ を解いて $e^x=y\pm\sqrt{y^2-1}$ を得る。よって $\operatorname{arccosh} y=x=\log\bigl(y\pm\sqrt{y^2-1}\bigr)$ である。$\pm$ の符号はそれぞれグラフの上の枝と下の枝に対応する。

同様に $y=\sinh x=(e^x-e^{-x})/2$ を e^x について解くと $e^x=y\pm\sqrt{y^2+1}$ を得る。ただし $e^x>0$ なので $-$ の符号は不適である。よって $x=\operatorname{arcsinh} x=\log\bigl(y+\sqrt{y^2+1}\bigr)$ が得られる。 ∎

■問 6 (d) の解答　地道に計算すると、次のようになる。

$$(\operatorname{arccosh} x)'=\frac{1+\frac{2x}{2\sqrt{x^2-1}}}{x+\sqrt{x^2-1}}=\frac{1}{\sqrt{x^2-1}},\qquad\qquad(\operatorname{arcsinh} x)'=\frac{1+\frac{2x}{2\sqrt{x^2+1}}}{x+\sqrt{x^2+1}}=\frac{1}{\sqrt{x^2+1}}$$

逆函数の微分法を使っても良い。$y=\cosh x$ のとき $(\operatorname{arccosh} y)'=1/(\cosh x)'=1/\sinh x=1/\sqrt{y^2-1}$ となる。 ∎

*13 一般の複素数 $z=a+i\theta$ に対しては、$e^z:=e^ae^{i\theta}$ と定めます。

数理科学基礎 (線形代数学) 第 3 回　　座標空間と数ベクトル

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 5 月 7 日 1 限

1　情報の調べ方

1.1　文献調査の一般論

数学的なことではありませんが、今回の「vector の語源を調べる」という問題に対して不適切な答案が多かったので、文献調査に関する最低限の基礎を述べておきます。きっと他の授業でも同じようなことを教わると信じていますが、念のため一読しておいてください。

まず**調査に使った文献を明記しないのは問答無用でアウト**です。今回は「調査すること」自体が課題ですから、どのような答案であってもどこかから引用されていることは分かります。ですが他のレポートで、出典を書かずにで他人の文章を取りこんだら、それは剽窃と呼ばれることになります。他人の文章を引っ張るときは「引用」という決まった作法に従い、自分の文章でないことを明らかにした上で、「誰がどこに書いたものなのか」を確実に分かるようにしてください*1。「たかが授業のレポートごときで」などと甘く見てはいけません。事が極限まで大きくなると、STAP の人がやらかしたように「学位論文の 1 つの章が丸ごとコピペ」という事態になるのですから*2。

次に、出典を明記するにしても**信頼できる文献をきちんと選んで**引用しましょう。特に Wikipedia は便利なもので、キーワードをインターネット上でググると頻繁にヒットします。ですが Wikipedia の記事は専門家が書いているとは限りません。大抵の場合 Wikipedia の記事には出典が付いていますから、引用する場合は元の文献にきちんと当たり、孫引き (引用の引用) にならないようにしましょう。引用でなくとも [要出典] タグが付いた場所など、出典がはっきりしない部分を鵜呑みにしてはいけません。

Wikipedia 以外の web サイトでも同じです。「公的機関の web サイトだから信頼できる」等の特別な事情がない限り、インターネット上の文書は「誰だか知らない人が書いたもの」に過ぎません。しばらく昔に Yahoo!知恵袋を使って京都大学の受験生がカンニングしたことが話題になりましたが、その際 Yahoo!知恵袋で「英文和訳問題の答」を教えた人は、翻訳ソフトを使っていたそうです*3。この程度の低品質な情報も平然と転がっているのですから、**インターネットで調べた情報を使うときは、その情報が信頼できるという裏付けを取ってください**。

実際 “vector” の語源を一つとっても、インターネット上には不正確な情報があふれていることが垣間見えます。せっかくの機会ですから、少し気合を入れて情報を調べてみましょう。

1.2　“vector” の語源

■ネット上の記事　最終的には信頼できる文献を探さねばいけないとはいえ、調査の最初の段階ではインターネットを使うのが楽です。手始めに「vector 語源」などでググってみましょう。見つかったサイトの記述から、関連個所をいくつか引用してみます*4。

- `http://www.nn.iij4u.or.jp/~hsat/techterm/vector.html`
  辞書を紐解くと, 数学用語の vector は 1843 年に数学者 William Rowen Hamilton (ハミルトン)(1805–1865)

---

*1 `http://www.bengo4.com/topics/1332/` に、著作権の視点からの分かりやすい説明があります。
*2 `http://stapcells.blogspot.jp/2014/02/blog-post_2064.html`
*3 `http://www.47news.jp/CN/201102/CN2011022801000578.html`
*4 以下での web サイトの引用は、いずれも 2015/5/14 に閲覧したときの情報に基づいています。

がラテン語 vectum から作ったとある。この vectum は「運ぶ」という意味であり, ギリシャ語の ‘ó χ o ζ が語源。 更に遡ると vehĕre に行き着くらしい。サンスクリットのヴァーハナは北インドで子供を守る女神の乗り物を意味するということである。

- `http://hooktail.sub.jp/vectoranalysis/restudyVector1/`
  vector という言葉は，1843 年に Hamilton がラテン語の vectum から作った造語ですが，vectum とは運ぶもの，という意味です．更に遡ると，サンスクリット語で女神の乗り物を意味するヴァーハナと同根のようです．
- `http://detail.chiebukuro.yahoo.co.jp/qa/question_detail/q1485274904`
  語源はラテン語です。ラテン語の意味は ‘carrier’ 「運び屋」です。
- `http://ja.wikipedia.org/wiki/`**ベクトル**
  ベクトルは ドイツ語: Vector[*5]に由来し、ベクターは 英語: vector に由来する。…(中略)… vector は「運ぶ」を意味するラテン語: vehere に由来し、18 世紀の天文学者によってはじめて使われた

これらの記事から、どうもラテン語の “vehere” とか “vectum” といった単語と関係があるようです。また Hamilton という数学者が関わっている気配も見てとれます。ところが vector という言葉が生まれた時期について、1843 年および 18 世紀という、2 種類の食い違う記述が見られます。どちらが正しいのでしょう？ もう少し Wikipedia を調べてみましょう。`http://ja.wikipedia.org/wiki/`**空間ベクトル** を見てみます。

> ハミルトンは 1846 年に四元数の複素数における実部と虚部に相当するものとしてスカラーとベクトルという用語を導入した (今日の用法とは異なることに注意されたい):

この記事は「ベクトルという用語を、Hamilton が 1846 年に導入した」と書いてあります。18 世紀でも 1843 年でもありません。ふと某名探偵の「真実はいつもひとつ！」という言葉が脳裏をよぎります。何が真実なのでしょうか？

■書籍や論文の調査　こういう時は、きちんとした文献の出番です。まず英単語としての語源を調べるなら、やはり使うべきは辞書でしょう。研究社から出ている『新英和大辞典 第 6 版』[*6]を引いてみると、p. 2727 に大体こんなことが書いてあります[*7]。この辞書は「18 世紀に vector が誕生」という説を支持するようです。

> vector は元々ラテン語で ‘carrier’ という意味の単語。そして vector は vehere の過去分詞形 vectus から派生している[*8]。初出は 1704 年[*9]。

一方 Wikipedia の「空間ベクトル」の記事には、Hamilton が 1846 年に導入したという事実がきちんと出典付で書いてあります。それによれば “London, Edinburgh & Dublin Philosophical Magazine 3rd series 29 27” が出典だそうです。皆さんはまだ見慣れないかもしれませんが、これは学術雑誌と呼ばれるものです。ですからこの文献を直接当たれば、1846 年に “vector” という言葉を Hamilton が使ったか確認できますね。

ここで皆さんの中に「1846 年の文献なんてどこで探すんだよ」と思った人がいるかもしれません。最近は便利なもので、なんと昔の学術雑誌がスキャンされてインターネット上に公開されていたりするのです。実際、今回のも雑誌のタイトルでググれば見つかります[*10]。その上、英語版 Wikipedia で Hamilton の記事[*11] を見てみると、References

---

[*5] (引用者注) これはスペルミスです。ドイツ語での「ベクトル」は “Vector” ではなく “Vektor” になります。

[*6] 辞書のチョイスは個人的な好みによるものです。ちゃんとした辞書なら、何でも OK です。受講生の皆さんの中には、大修館書店の『ジーニアス英和大辞典』を引用した人もいました。

[*7] まるっきりこのままではありません。辞書の方では、凡例記号で色々省略がされています。

[*8] (引用者注) 水谷智洋編『改訂版 羅和辞典』(研究社) の記述と一致することを確認しました。また vectus は「動詞 vehō の不定法現在形」だそうです。ラテン語を読めないので、この辺の事情を深く調べるのは諦めました。

[*9] (引用者注) ただし凡例によると、初出は「‘利用可能な文献による限りの初出例の年代’ ということで, 絶対的なものではない」とのこと。

[*10] `http://www.biodiversitylibrary.org/bibliography/58679#/summary` にある “ser.3:v.31 (1846:July-Dec)” です。

[*11] `http://en.wikipedia.org/wiki/William_Rowan_Hamilton`

の節に “David R. Wilkins's collection of Hamilton's Mathematical Papers”[*12] というリンクがあります。なんと Hamilton の論文を、全てまとめて (しかも TEX で打ち直して！) web サイトを作ってくれた人がいるのです。そこで http://www.maths.soton.ac.uk/EMIS/classics/Hamilton/OnQuat.pdf の 16 ページを見ると

> “... the algebraically imaginary part, being geometrically constructed by a straight line, or radius vector, which has, in general, for each determined quaternion, a determined length and determined direction in space, may be called the vector part ...”

という記述が見られます。Wikipedia の「空間ベクトル」の記事で引用されてたやつですね。これは 1846 年 7 月の記事らしいので、確かに 1846 年に Hamilton が vector という用語を使ったことは間違いありません[*13]。

■2 種類の問題　ここまでで、2 つ確実に分かったことがあります。

- 新英和大辞典によると、vector はラテン語の「運び屋」という意味の単語で、初出は 1704 年である
- Hamilton は 1846 年に vector part という用語を導入した

これらの事実は疑いようがありません。そこで浮かび上がってくるのは「vector という英単語が誕生したのは 1704 年で、それを数学の文脈で最初に用いたのが Hamilton だった」という可能性です。

このような「一般的に使われている単語が、数学用語として用いられるようになる」という事案は、現代でも日常的に起こっています。たとえば “cellular algebra” という数学の言葉がありますが、これは 1996 年、J. J. Graham と G. I. Lehrer という 2 人の数学者によって定義されたことがはっきりしています[*14]。ですが cellular という単語は、1996 年よりずっとずっと昔から存在していました[*15]。そこで vector についても、同じことが起きているのではないかと推測されるわけです。

インターネットで見つけた 2 つの web ページには、これらの記述と食い違う内容が書かれています。一番最初に挙げた web サイトを見直すと「辞書を紐解いたら」Hamilton が vector を作ったと書かれているそうです。ところが参考文献として挙げられている、三省堂の『カレッジクラウン英和辞典』[*16]および岩波書店の『岩波理化学辞典 第 5 版』を実際に紐解いても、ベクトルの語源とか意味だけが書いてあって、「Hamilton が作った」とはどこにも書いてありません。したがって、一番最初に挙げた web サイトの著者は vector について「言葉が作られたこと」と「言葉を数学に持ち込んだこと」を混同していると推測されます[*17]。また、1843 年に出版された Hamilton の論文全部を検索しても vector という単語は出ず、1844 年にやっと “radius vector” という単語が飛び出します。1843 年説も間違いでしたね。このようにインターネットには、著者の勘違い等に基づく不正確な情報が転がっていることがしばしばあるのです。

ついでに言えばページが作成された日時と内容を見るに、「物理のかぎしっぽ」の記事は 1 つ目の文献を剽窃した上、精査していないようにも見えます。「クロ」とまで言い切る証拠はないですが、皆さんはこのような疑わしい文章を書かないように心がけましょう。

■vector の語源と初出は？　これまでの調査で、vector の語源がラテン語の動詞 “vectus” であること、元々の意味が「運び屋」であることはほぼはっきりしました。ですが初出がまだはっきりしません。せっかくだから、もう少し頑張りましょう。こういう時に岩波書店の『数学辞典』が役立ちそうですが、残念ながら今回は欲しい情報が載っていませ

[*12] http://www.maths.soton.ac.uk/EMIS/classics/Hamilton/

[*13] さらに東京大学 OPAC で検索すると、本郷にある理学部の物理図書室に London, Edinburgh & Dublin Philosophical Magazine の現物が眠っていることが分かります。1846 年の文献でありながら、直接現物を確認することもできてしまうのです。

[*14] 正確な文献情報は次の通り: J. J. Graham, G. I. Lehrer, “Cellular algebras”, *Inventiones Mathematicae* **123**, 1–34.

[*15] 「Wikipedia によれば」1665 年に Robert Hooke が細胞を “cell” と名付け、1674 年に Leewenhoek が詳しく観察したそうです。気になる人はきちんと調べてみてください。

[*16] カレッジクラウン英和辞典第 2 版は 1986 年に出版された古い辞書ですが、頑張って現物を調達しました。

[*17] Hamilton が 1833 年に “Introductory Lecture on Astronomy” という本を出版していることを考えると、Hamilton が天文学の ‘vector’ という言葉を知らなかった可能性はほとんど無いでしょう。

んでした。そこで Wikipedia の「ベクトル」のページにある参考文献、Daniel Fleisch 著、河辺哲次訳 『物理のためのベクトルとテンソル』 (岩波書店、2013 年) の p. 1 を開いてみます。

> **ベクトル** (vector) という言葉は「運ぶ」を意味するラテン語 vehere に由来します．この言葉は，太陽の周りを「運ばれる」惑星の運動を研究していた 18 世紀の天文学者たちによって，初めて使われました．

新英和大辞典と合わせると、初出 1704 年で、かつ最初に天文学で使われたとみてよさようです。ですがこの本には「誰が使ったか」まで書いてないので、ここで Oxford English Dictionary (OED) の第 2 版を持ち出してみましょう。19 巻の p. 470 で “vector” を引くと、こんな風に書いてあります。

> †**1**. *Astr.* (See quot. 1704) Also *vector radius*, = *radius vector* RADIUS 3 **e**. *Obs.* **1704** J. HARRIS *Lex. Techn.* I. s.v., A Line supposed to be drawn from any Planet...

これでようやく、初出文献が分かりました。J. Harris という人の “Lex. Techn.” という本だそうです。略されているため正確なタイトルが分からないので、東京大学 OPAC で検索してみましょう。「著者 Harris, 出版年 1704」で検索をかけると何も出ないのですが、他大学の図書まで含めた検索をかけると、放送大学図書館の蔵書がヒットします。そのレコードで著者とタイトルが John Harris, “Lexicon technicum, or, An universal English dictionary of arts and sciences” (1704) だと確定します。これを見れば、1704 年に本当に vector が登場したか分かります。

あとはどうやってこの本を調べるか、という問題が残ります。東京大学内の図書館に所蔵がありません。一つには、東京大学の図書館を経由して放送大学付属図書館に文献の複写を依頼する手があります。また国立国会図書館の郵送複写サービス等も利用可能です。ところが今回「国立国会図書館サーチ」[*18]に書籍タイトルを突っ込むと、なんとこの本のスキャン画像が出てきます: `http://jairo.nii.ac.jp/0108/00002692` これの頭文字 “V” のページを見れば、“Vector” の項目に

> A Line supposed to be drawn from any Planet moving round a Center, or the Focus of an Ellipsis, to that Center or Focus, is by some Writers of the New Astronomy, called the Vector; because 'tis that Line by which the Planet seems to be carried round its Center, and with which it describes proportional Areas in proportional Times.

と説明が書いてあります。確かに 1704 年に “vector” が登場したことがはっきりします[*19]。

ちなみに OED には「数学の “vector” は、Hamilton が 1846 年に使った」とも書かれていて、そこでは上で挙げた Hamilton の論文が引用されています。頑張れば、独自の調査で OED の初出年代にたどり着くこともできるんですね。

1.3 教訓

このように、普段「ネットでちょこっと調べればいいや」と思いがちな情報検索も、頑張ろうとすると結構時間と手間がかかるものなのです。ただ、ここまで読んだ皆さんは「何か無駄が多くなかったか？」と思うかもしれません。それは正しいです。実際 web サイトの記述を検証するために余分な手間がかかっていはいますが、それを差っ引いても

- 一番最初に OED と『新英和大辞典』をセットで調べてみる
- 出てきた参考文献の原本を、東京大学/国会図書館の OPAC や Hamilton の業績コレクションを使って検索する

という順で作業してしまえば、もっと手早く正確な情報にたどり着けたはずです。また今回、「昔の書物がオンラインで見られること」が非常に重要な役割を果たしました。ですから、今回の例を教訓に

[*18] `http://www.ndl.go.jp/index.html` の左上です。NDL-OPAC とは別なので注意してください。

[*19] ちなみに vector の初出については、Harris の本よりさらに遡ろうとしてみました。しかし Harris の本にある “the New Astronomy” がどこにあるのか、天文学専攻の友人に聞いても良く分からなかったので、諦めました。

- 情報検索に「使える」文献や手段 (今回は OED など) を知っておくこと
- オンラインで見られる文献を頑張って探すこと

の大事さを認識しておいてください。

最後に、これまでの “vector” の語源の話は全て検証可能です。ですからもし TA が嘘をついているとすれば、挙げた文献に当たればそれを見抜くことができます。今回そこまでするかどうかは皆さんにお任せしますが、先生や TA が言ったことも必要に応じて検証せねばならないし、もし検証不可能なことがあれば皆さんは考え、質問をしなければいけません。そのことを忘れないでください。

2　空間内における直線と平面の取り扱い

皆さんは既に、平面 $\mathbb{R}^2$ 内の直線の扱いについては十分習熟しているはずです。たとえば 2 点を通る直線の方程式を求めるとか、直線の法線ベクトルを求めるといった計算は、すぐにできるでしょう。同じようなことを 1 次元上げもできるようになろう、というのが今回のお題の一つです。

まず最初に、平面を決定するための条件を確認しておきます。平面は

- 3 つの相異なる点
- 1 つの点と、平面を張る 2 本の 1 次独立なベクトル
- 1 つの点と、1 本の法ベクトル

が与えられれば定まります。これは頭の中で想像すればすぐ分かります。また、平面の表し方には

- 方程式による表示
- パラメータ表示

という 2 通りのやり方がありました。ですから

- 与えられた平面の決定条件から、平面の方程式やパラメータ表示を求めること
- 平面の決定条件のうちどれか 1 つから、残りの 2 つの条件を導くこと
- 平面の方程式とパラメータ表示をいったりきたりすること

がスムーズにできなければいけません。

それから、今回の問題では「2 次元版と 3 次元版」が並列されていることに気づいた人が多いと思います。以下で見るように、問題の次元が 1 個変わっても、解き方はほとんど変わりません。そして皆さんが後で学ぶ「線型代数」とは、そのような**次元によらず同じように使えるテクニック**を身に着けることだと言えます。そのような「何で同じように問題が解けるのか」という原理に関しても、注意してみてください。

2.1　直線と平面のパラメータ表示

まず最初に、直線のパラメータ表示を復習しましょう。空間内の直線は、通る点 $\boldsymbol{x}_0 = (x_0, y_0, z_0)$ と方向ベクトル ${}^t(a, b, c)$ が[*20]与えられれば決まります。そのとき直線上の点 $\boldsymbol{x}$ に対し $\boldsymbol{x} - \boldsymbol{x_0}$ は ${}^t(a, b, c)$ と平行なので、適当な実数 $t \in \mathbb{R}$ を用いて $\boldsymbol{x} - \boldsymbol{x}_0 = t\,{}^t(a, b, c)$ と書けます。逆に全ての実数 t に対して $\boldsymbol{x} = \boldsymbol{x_0} + t\,{}^t(a, b, c)$ は直線上の点です。こうして $\boldsymbol{x} = \boldsymbol{x_0} + t\,{}^t(a, b, c)$ $(t \in \mathbb{R})$ が直線のパラメータ表示を与えることが分かりました。平面の場合も同様です。

[*20] 記号 ^{t} は、縦ベクトルを横ベクトルに直す記号です。縦ベクトルを使うと紙面上で面積を消費するので、印刷物ではこんな書き方をすることがしばしばあります。また「別に横ベクトルでもいいじゃん」と思うかもしれませんが、後で行列の話をするとき、縦ベクトルと横ベクトルを区別する必要が出てきます。なので以下、空間内のベクトルは転置記号をつけて縦ベクトルにします。

通る点 $\boldsymbol{x}_0$ と平面を張るベクトル $\boldsymbol{u}$, $\boldsymbol{v}$ が与えられると、$\boldsymbol{x} = \boldsymbol{x}_0 + a\boldsymbol{u} + b\boldsymbol{v}$ $(a, b \in \mathbb{R})$ がパラメータ表示です。

さて今度は、2 点 $\boldsymbol{p}_1$, $\boldsymbol{p}_2$ が与えられたとします。この 2 点を通る直線の方向ベクトルは $\boldsymbol{p}_2 - \boldsymbol{p}_1$ なので、直線のパラメータ表示は $\boldsymbol{x} = \boldsymbol{p}_1 + t(\boldsymbol{p}_2 - \boldsymbol{p}_1) = (1-t)\boldsymbol{p}_1 + t\boldsymbol{p}_2$ となります。したがってベクトル $\boldsymbol{p}_1$ と $\boldsymbol{p}_2$ を係数が 1 になるよう足し合わせることで、直線の方程式が得られます。書き方を変えると、$s\boldsymbol{p}_1 + t\boldsymbol{p}_2$ $(s+t=1)$ という表示もできます。

平面についても同じように、$\boldsymbol{p}_1$, $\boldsymbol{p}_2$, $\boldsymbol{p}_3$ を通る平面のパラメータ表示は $\boldsymbol{x} = \boldsymbol{p}_1 + s(\boldsymbol{p}_2 - \boldsymbol{p}_1) + t(\boldsymbol{p}_3 - \boldsymbol{p}_1) = (1-s-t)\boldsymbol{p}_1 + s\boldsymbol{p}_2 + t\boldsymbol{p}_3$ $(s, t \in \mathbb{R})$ あるいは $\boldsymbol{x} = s\boldsymbol{p}_1 + t\boldsymbol{p}_2 + u\boldsymbol{p}_3$ $(s+t+u=1)$ となります。

これらの式は結構利用価値が高いので、覚えておいてください。

2.2 平面の方程式と法ベクトル

次に、平面の法ベクトルについて考えてみましょう。平面は、通る点と法ベクトルがあれば一通りに決定されます。そこで $\boldsymbol{x}_0 = {}^t(x_0, y_0, z_0)$ を通り $\boldsymbol{n} = {}^t(a, b, c)$ を法ベクトルとする平面を考えます。すると平面上の任意の点 $\boldsymbol{x}$ に対し、$\boldsymbol{x} - \boldsymbol{x}_0$ は $\boldsymbol{n}$ と垂直になります。したがって

$$0 = \boldsymbol{n} \cdot (\boldsymbol{x} - \boldsymbol{x}_0) = a(x - x_0) + b(y - y_0) + c(z - z_0)$$

です。この式で $d := -ax_0 - by_0 - cz_0$ とおけば、平面上の点 $\boldsymbol{x}$ は方程式 $ax + by + cz + d = 0$ を満たします。

逆に方程式 $ax + by + cz + d = 0$ が与えられたとき、この方程式を満たす点全体の集合は平面になります。まず $(a, b, c) \neq (0, 0, 0)$ なので、$ax + by + cz + d = 0$ の解が一つは存在します。たとえば $a \neq 0$ なら、$\boldsymbol{x}_0 := {}^t(-d/a, 0, 0)$ とすれば良いです。そして $\boldsymbol{x} = {}^t(x, y, z)$ が方程式の解だとすれば、$ax + by + cz + d = 0$ なので

$$0 = ax + by + cz + d = {}^t(a, b, c) \cdot \boldsymbol{x} - {}^t(a, b, c) \cdot \boldsymbol{x}_0 = {}^t(a, b, c) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)$$

となります。こうして全ての解ベクトルは ${}^t(a, b, c)$ と垂直なので、方程式 $ax + by + cz + d = 0$ の定める集合は平面だと分かりました。

この事実を踏まえれば、問 6 と問 7 が解けます。

■問 6, 7 の解答　平面 Π: $ax + by + cz + d = 0$[*21]の法ベクトルは ${}^t(a, b, c)$ である。また点 (p, q, r) が与えられたとき、この点を通り方向ベクトルが ${}^t(a, b, c)$ であるような直線 ℓ のパラメータ表示は ${}^t(p, q, r) + \alpha\, {}^t(a, b, c)$ $(\alpha \in \mathbb{R})$ である。いま ${}^t(a, b, c)$ は平面 Π の法ベクトルだから、平面 Π と直線 ℓ は 1 点で交わる。すなわち平面 $ax + by + cz + d = 0$ にベクトル ${}^t(p, q, r) + \alpha\, {}^t(a, b, c)$ が乗るような $\alpha \in \mathbb{R}$ が唯一存在する。この α を用いて、(p, q, r) から平面までの距離は $\|\alpha\, {}^t(a, b, c)\| = |\alpha|\sqrt{a^2 + b^2 + c^2}$ と書ける。他方 ${}^t(p, q, r) + \alpha\, {}^t(a, b, c)$ が方程式 $ax + by + cz + d = 0$ を満たすので

$$a(a\alpha + p) + b(b\alpha + q) + c(c\alpha + r) + d = 0 \quad \text{すなわち} \quad \alpha(a^2 + b^2 + c^2) = -(ap + bq + cr + d)$$

よって求める距離は

$$|\alpha|\sqrt{a^2 + b^2 + c^2} = \frac{|ap + bq + cr + d|}{\sqrt{a^2 + b^2 + c^2}}$$

で与えられる。この式で $c = r = 0$ と置けば、平面内の場合の式が得られる。 ■

■与えられた点を通る平面　これまでの議論で、与えられた 3 点を通る平面を求める方法が分かりました。

- まずその方程式を $ax + by + cz + d = 0$ とおき
- この方程式に 3 点の座標を代入して、a, b, c の連立 1 次方程式を得る
- 得られた方程式を解く

[*21] Π は、ギリシャ文字 π（パイ）の大文字です。

とすれば OK です。ただ、ここで出てくる方程式は 3 変数の連立 1 次方程式なので、一般の場合に解くのは面倒です。問 2〜問 5 のように、通る点の座標に 0 がたくさん出てくる場合は解くのが楽ですが、そうでない場合は

- パラメータ表示を使う
- 後で説明するベクトルの外積を使い、法線ベクトルを求めてから方程式を得る

など別の手を使う方が賢いでしょう。

■問 2, 3 の解答　平面の方程式を $\alpha x + \beta y + \gamma z = \delta$ とおき、この式に $(a, 0, 0)$, $(0, b, 0)$, $(0, 0, c)$ を代入して、方程式

$$\frac{x}{a} + \frac{y}{b} + \frac{z}{c} = 1$$

が直ちに得られる。この式で z/c の項を落とせば、直線の式が得られる。■

■問 4, 5 の解答　原点を通るので、方程式の定数項は 0 である。あとは座標 $(1, 0, a)$ と $(1, 0, b)$ を代入すれば、求める式は $z = ax + by$ と分かる。平面上の直線の場合は、同様にして $y = ax$ が求まる。■

2.3　平面の方程式と決定条件

■問 8, 10 の解答　2 点 (p_1, q_1) と (p_2, q_2) を通る直線がただ一つに決まる条件は $(p_1, q_1) \neq (p_2, q_2)$ である。同様に 3 点 (p_1, q_1, r_1), (p_2, q_2, r_2), (p_3, q_3, r_3) を通る平面がただ一つに決まる条件は、3 点が同一直線上にないことである。■

■代数的な考察　この問題の結果自体は、直感的にはほとんど明らかです。そもそも Euclid の『原論』にも、公理として「相異なる 2 点を通る直線がただ 1 つ存在する」と書かれています。が、ここでちょっと視点を変え「直線がただ一つに決まる」という事実を、代数的に捉えてみましょう。

平面 $\mathbb{R}^2$ 内の直線 $ax + by + c = 0$ が点 (p_1, q_1) と (p_2, q_2) を両方通るとします。このとき $ap_1 + bq_1 + c = 0$ および $ap_2 + bq_2 + c = 0$ という式が成り立ちます。この 2 つの式の差を取れば

$$a(p_1 - p_2) + b(q_1 - q_2) = 0$$

という式が得られます。逆にこの式を a と b の満たすべき方程式だと思えば、この 0 でない解の一つ一つが (p_1, q_1) と (p_2, q_2) の両方を通る直線に対応することになります。ただし 0 でない任意の実数 $\alpha \in \mathbb{R}$ に対し、方程式 $ax+by+c = 0$ と $\alpha ax + \alpha by + \alpha c = 0$ は同じ直線を表します。ですから $(p_1 - p_2)a + (q_1 - q_2)b = 0$ の解 (a, b) に 0 でない定数 α をかけたとき、(a, b) と $(\alpha a, \alpha b)$ は同じ直線に対応します。

いま $(p_1, q_1) \neq (p_2, q_2)$ だとしたら、$(p_1 - p_2)a + (q_1 - q_2)b = 0$ は、原点を通る (a, b) 平面の直線を表します。したがって、$(0, 0)$ 以外のどの二つの解も定数倍でうつります。ゆえに直線はただ一つに定まります。ところが $(p_1, q_1) = (p_2, q_2)$ なら、全ての (a, b) が解となってしまいます。これらの解のうちには定数倍でうつり合わないものが無数にありますから、(p_1, q_1) と (p_2, q_2) の両方を通る直線は無数に存在します。

このように「相異なる 2 点 (p_1, q_1), (p_2, q_2) を通る直線がただ一つ存在する」という条件が「方程式 $(p_1 - p_2)a + (q_1 - q_2)b = 0$ の $(0, 0)$ でない解が、定数倍を除いてただ一つに決まる」という代数的な言葉に言い換えられたわけです。若干言い方がまどろっこしいですが、少しして線型代数の勉強をすると「方程式 $(p_1 - p_2)a + (q_1 - q_2)b = 0$ の解空間が 1 次元」という、もっとすっきりした表現が使えるようになります。

空間の場合も話は同じです。平面の方程式 $ax + by + cz + d = 0$ に通る 3 点の座標を代入すると、a, b, c, d に関する 3 つの連立 1 次方程式が得られます。変数が 4 つで式が 3 つですから、この連立方程式の解はただ 1 つに決まらず、解を動かす自由度が残ります。ただ 0 でない定数を (a, b, c, d) にかけても平面自体は変わらないので、「1 次元分の自由度」が「平面が 1 個に決まること」に相当します。

2.4 直線の方程式

ここまで使いませんでしたが、直線の方程式についても少しだけ触れておきます。パラメータ表示された直線 $\ell\colon {}^t(x_0, y_0, z_0) + t\,{}^t(a, b, c)$ が与えられたとします。このとき直線上の点 $\boldsymbol{x} = {}^t(x, y, z)$ は、ある実数 $t \in \mathbb{R}$ に対し

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases}$$

を満たします。この式を t について解くことで、直線上の点が満たす式

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}$$

が得られるのでした。ちなみにこの式は a, b, c の中に 0 が混ざっているとマズいですが、そういう時も使えるような式を得るなら、分母を払っておけば大丈夫です。a, b, c の中に、少なくとも 1 つは 0 でないものがあります。それが a だった場合は、今の方程式を

$$b(x - x_0) = a(y - y_0), c(x - x_0) = a(z - z_0)$$

と書き換えれば問題ありません。逆にこの方程式が与えられていれば、$t = (x - x_0)/a$ とおくことで、$a \neq 0$ と合わせて $y - y_0 = bt$, $z - z_0 = ct$ が得られます。こうして直線の方程式からパラメータ表示を復元することもできました。

ここで、直線の方程式は

$$\begin{cases} \Pi_1\colon bx - ay - (bx_0 - ay_0) = 0 \\ \Pi_2\colon cx - az - (cx_0 - az_0) = 0 \end{cases}$$

という 2 本の平面の方程式を連立して得られています。これらの平面は

- いずれも点 (x_0, y_0, z_0) を通り
- それぞれの法ベクトル $(b, -a, 0)$ と $(c, 0, -a)$ が共に ℓ の方向ベクトル (a, b, c) と垂直

なことに気を付けましょう。したがって平面 Π_1, Π_2 は共に直線 ℓ を含んでいることが分かります。すなわち直線 ℓ の方程式は、ℓ を **2 枚の平面 Π_1, Π_2 の共通部分**として表しているわけです。

このような手は、他の場合にも応用できます。空間内の図形をつかむには、その図形をいくつかの分かりやすい図形の共通部分として表しておいて、それぞれのパーツの方程式を作れば良いのです。そうすれば連立方程式の形で、求める図形を表す方程式が得られます。

3 ベクトルの 1 次独立性と行列式

3.1 ベクトルの 1 次独立性

■1 次独立性　空間内の n 本のベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ が 1 次独立であることの定義は、$\alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n$ となる $(\alpha_1, \ldots, \alpha_n)$ が $(0, \ldots, 0)$ 以外に存在しないことでした。1 次独立でなく、たとえばもし $\alpha_1 = 1$ だったなら、$\boldsymbol{v}_1 = -\alpha_2\boldsymbol{v}_2 - \cdots - \alpha_n\boldsymbol{v}_n$ というように、$\boldsymbol{v}_1$ が他のベクトルのスカラー倍と足し算で表されます。なので 1 次独立性は、大体「ベクトルたちがバラバラの方向を向いている」ということに相当します。また 1 次独立でないことを 1 次従属と呼びます。これを踏まえた上で、問 9, 11 の解答を見てみましょう。

■共線条件と共面条件 (問 9, 11 の解答)　2 点 (p_1, q_1), (p_2, q_2) を通る直線のパラメータ表示は、$s\,{}^t(p_1, q_1) + t\,{}^t(p_2, q_2)$ $(s + t = 1)$ で与えられます。したがって 3 点が同一直線上にあるためには、$(p_3, q_3) = s\,{}^t(p_1, q_1) + t\,{}^t(p_2, q_2)$, $s + t = 1$ という連立方程式が解を持つか、$(p_1, q_1) = (p_2, q_2)$ であることが必要十分です。

ただ、この方程式は「ベクトルのパラメータ表示」と「1 次方程式」がバラバラに出てきてちょっと扱いづらいです。そこで平面ベクトルの話ですがわざと次元を上げて、${}^t(1, p_1, q_1)$, ${}^t(1, p_2, q_2)$ というベクトルを考えてみます。そうすると $s\,{}^t(1, p_1, q_1) + t\,{}^t(1, p_2, q_2) = {}^t(s+t, sp_1, +tp_2, sq_1 + tq_2)$ となるので、求める条件が「$s\,{}^t(1, p_1, q_1) + t\,{}^t(1, p_2, q_2) = {}^t(1, p_3, q_3)$ となる s, t が存在すること」でまとまります*22。この式は移項すれば

$$s\,{}^t(1, p_1, q_1) + t\,{}^t(1, p_2, q_2) - {}^t(1, p_3, q_3) = \mathbf{0}$$

となります。すなわち「${}^t(1, p_1, q_1)$, ${}^t(1, p_2, q_2)$, ${}^t(1, p_3, q_3)$ が 1 次従属であること」が求める条件です。

空間の場合も、議論は全く同じです。求める共面条件は「ベクトル ${}^t(1, p_1, q_1, r_1)$, ${}^t(1, p_2, q_2, r_2)$, ${}^t(1, p_3, q_3, r_3)$, ${}^t(1, p_4, q_4, r_4)$ が 1 次従属であること」となります。

3.2 ベクトルの外積

■問 12 の解答 ${}^t(p_1, q_1)$ に垂直なベクトルは、明らかに $\alpha\,{}^t(-q_1, p_1)$ $(\alpha \neq 0)$ である。実際内積を取れば、${}^t(p_1, q_1) \cdot \alpha\,{}^t(-q_1, p_1) = \alpha(-p_1q_1 + q_1p_1) = 0$ となる。

■ベクトルの外積 (問 13 の解答) 平面 $\mathbb{R}^2$ の場合はこのように垂直なベクトルはすぐ求まるわけですが、空間の場合はどうなるでしょうか。すなわち空間の場合、2 本のベクトル $\boldsymbol{u}$, $\boldsymbol{v}$ が与えられたとき、両方のベクトルに垂直なベクトルを作るにはどうすれば良いでしょうか。実は「ベクトルの外積」というものを使うと、それを求めることができます。

$\boldsymbol{u} = {}^t(p_1, q_1, r_1)$, $\boldsymbol{v} = {}^t(p_2, q_2, r_2)$ に対し、$\boldsymbol{u}$ と $\boldsymbol{v}$ の外積は $\boldsymbol{u} \times \boldsymbol{v} := (q_1r_2 - q_2r_1, r_1p_2 - r_2p_1, p_1q_2 - p_2q_1)$ という式で与えられるベクトルです。そうすると

$$\begin{aligned}
\boldsymbol{u} \cdot (\boldsymbol{u} \times \boldsymbol{v}) &= p_1(q_1r_2 - q_2r_1) + q_1(r_1p_2 - r_2p_1) + r_1(p_1q_2 - p_2q_1) = 0 \\
\boldsymbol{v} \cdot (\boldsymbol{u} \times \boldsymbol{v}) &= p_2(q_1r_2 - q_2r_1) + q_2(r_1p_2 - r_2p_1) + r_2(p_1q_2 - p_2q_1) = 0
\end{aligned}$$

となるので、確かに $\boldsymbol{u} \times \boldsymbol{v}$ は $\boldsymbol{u}$ および $\boldsymbol{v}$ と垂直になることが分かります。今は天下りな定義を与えましたが、これから行列式を使い、もう少しマシな説明をします。

3.3 行列式

ベクトルの 1 次独立性 / 従属性や外積の話は「行列式」というものを知っていると良く理解できます。本格的に勉強するのは後ですから、ここではラフに話をしておきます。

色々な定義がありますが、$\mathbb{R}^2$ 内の 2 本の縦ベクトル $\boldsymbol{u}$, $\boldsymbol{v}$ の張る平行四辺形の符号付き面積*23を $\det(\boldsymbol{u}\ \boldsymbol{v})$ と書きます。また $\mathbb{R}^3$ 内の 3 本の縦ベクトル $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ の張る平行六面体の符号付き体積*24を $\det(\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w})$ と書きます。このとき、det をベクトルの成分で書き下すと

$$\det\begin{pmatrix} a & c \\ b & d \end{pmatrix} = ad - bc, \quad \det\begin{pmatrix} p_1 & p_2 & p_3 \\ q_1 & q_2 & q_3 \\ r_1 & r_2 & r_3 \end{pmatrix} = p_1 \det\begin{pmatrix} q_2 & q_3 \\ r_2 & r_3 \end{pmatrix} - q_1 \det\begin{pmatrix} p_2 & p_3 \\ r_2 & r_3 \end{pmatrix} + r_1 \det\begin{pmatrix} p_2 & p_3 \\ q_2 & q_3 \end{pmatrix}$$

となることが知られています。

さて平面内に 2 本のベクトルがあったとき、これらが 1 次独立だったら平行四辺形が張れます。逆に 1 次従属だったら、平行四辺形を張ろうとすると面積が 0 になります。したがって $\boldsymbol{u}$, $\boldsymbol{v}$ が 1 次独立であることと $\det(\boldsymbol{u}\ \boldsymbol{v}) \neq 0$ が同値になります。3 次元以上の場合でも同様に det を使えば、すごくややこしくなりますが「ベクトルが 1 次独立であるか」を成分計算する式が得られるのです。これを使えば共線・共面条件をたった 1 本の式で書き下せます。

*22 この「一度次元を上げてから、見やすいように書き直す」という手法は、他の場面でもしばしば用いられます。

*23 符号は、$\boldsymbol{u}$ から $\boldsymbol{v}$ に向かう方向が正の向きであるときに $+$ とし、負の向き $-$ とします。

*24 符号は、$\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ がこの順で右手系をなすときに $+$, 左手系をなすときに $-$ とします。

さらに 3 次元の det を注意深く観察すると、

$$\det\begin{pmatrix}p_1 & p_2 & p_3\\ q_1 & q_2 & q_3\\ r_1 & r_2 & r_3\end{pmatrix} = {}^t(p_1,q_1,r_1)\cdot{}^t\left(\det\begin{pmatrix}q_2 & q_3\\ r_2 & r_3\end{pmatrix}, -\det\begin{pmatrix}p_2 & p_3\\ r_2 & r_3\end{pmatrix}, \det\begin{pmatrix}p_2 & p_3\\ q_2 & q_3\end{pmatrix}\right)$$

のように、det が 2 本のベクトルの内積で書けることが分かります。ここで 2 次元のときの det を見ると、行の入れ替えで符号が (-1) 倍されると分かります。よって今の式で det の前の $-$ が消せて

$$\det\begin{pmatrix}p_1 & p_2 & p_3\\ q_1 & q_2 & q_3\\ r_1 & r_2 & r_3\end{pmatrix} = {}^t(p_1,q_1,r_1)\cdot{}^t\left(\det\begin{pmatrix}q_2 & q_3\\ r_2 & r_3\end{pmatrix}, \det\begin{pmatrix}r_2 & r_3\\ p_2 & p_3\end{pmatrix}, \det\begin{pmatrix}p_2 & p_3\\ q_2 & q_3\end{pmatrix}\right)$$

となります。この右辺に出てくる、成分に det が入ったベクトルが、さっき外積と呼んだものです。したがって、ここまでの議論で $\det(\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w}) = \boldsymbol{u}\cdot(\boldsymbol{v}\times\boldsymbol{w})$ が分かりました。この式を使えば、なぜ $\boldsymbol{v}\times\boldsymbol{w}$ が $\boldsymbol{v}$ と直交するのか一目瞭然ですね。$\boldsymbol{v}\cdot(\boldsymbol{v}\times\boldsymbol{w})$ はベクトル $\boldsymbol{v}$, $\boldsymbol{v}$, $\boldsymbol{w}$ の張る平行六面体の体積ですが、この平行六面体はぺっちゃんこなので、どう見ても体積 0 です。こうして内積の値が 0 となり、$\boldsymbol{v}$ と $\boldsymbol{v}\times\boldsymbol{w}$ が直交すると分かります。$\boldsymbol{w}$ についても同じです。

このようにベクトルの 1 次独立性を det と関連付けることによって、ベクトルが一段と便利な道具になります。また det を用いてベクトルの情報を得る過程には、det が持つ諸々の計算公式が重要な役割を果たします。そういう事も追々勉強しますので、楽しみにしていてください。

4　残りの問題

■問 14 の解答　3 つの点が同一直線上にないことが必要十分条件です。このとき 3 点は三角形をなすから、外接円が 3 点全てを通過します。続いて円の方程式を $(x-a)^2+(y-b)^2=r^2$ とおく。これに $(x,y)=(p_1,q_1),(p_2,q_2),(p_3,q_3)$ を代入すると、a,b,r に関する 3 変数の連立一次方程式が得られる…のですが、これを真面目に解いてもあまり徳しません。△ABC の 3 点の位置ベクトルを $\boldsymbol{a},\boldsymbol{b},\boldsymbol{c}$ とすると、外心が $\dfrac{\sin 2\angle A\boldsymbol{a}+\sin 2\angle B\boldsymbol{b}+\sin 2\angle C\boldsymbol{c}}{\sin 2\angle A+\sin 2\angle B+\sin 2\angle C}$ と表せます。こっちの式の方が使い勝手が良いと思います[*25]。

■問 15 の解答　存在しないことを背理法で示す。全ての頂点が格子点である正三角形があったとする。このとき

- 成分が整数であるようなベクトルに沿って格子点を平行移動したものは、再び格子点になる
- 図形を平行移動すると、元の図形と合同になる

という事実があるので、正三角形の頂点の 1 つが原点であると仮定して差支えない。そこで正三角形の頂点を $(0,0)$, (a,b), (c,d) $(a,b,c,d\in\mathbb{Z})$ とする。

いま正三角形の一辺の長さは $\sqrt{a^2+b^2}$ である。したがって正三角形の面積は

$$\frac{1}{2}\times\sqrt{a^2+b^2}\times\frac{\sqrt{3}}{2}\sqrt{a^2+b^2}=\frac{\sqrt{3}(a^2+b^2)}{4}$$

である。一方、座標を使って計算すると面積が $|ad-bc|/2$ と求まる。したがって

$$\frac{\sqrt{3}(a^2+b^2)}{4}=\frac{|ad-bc|}{2}\quad\text{より}\quad\sqrt{3}=\frac{2|ad-bc|}{a^2+b^2}\in\mathbb{Q}$$

となる。これは $\sqrt{3}$ が無理数である事実に反する。 ∎

[*25] 頂点の取り方を工夫していくらか計算してみましたが、座標ベースではあんまり綺麗な式や面白い式が得られませんでした。すいません。

数理科学基礎 (線形代数学) 第 4 回　　2 変数函数のグラフ

担当教員: 植野 義明 / TA: 穂坂 秀昭
講義日時: 2015 年 5 月 13 日 1 限

1　講評

最初に採点をしてて思ったことを、ちょっとだけ書いておきます。

- 全体的に正答率は高かったです。
- グラフを描く問題は、コンピュータで頑張った人と手書きで頑張った人が半々ぐらいでした。またコンピュータを使った人の中では Mathematica で描いた人が一番多く、次に gnuplot を使った人が多かったと思います。
- グラフ $z = xy$ を回転する問題や平面で切断したときの断面を考える問題は、正答率が極めて低かったです。2 変数函数のグラフの回転操作は、まだなじみがなかったのかもしれません。「グラフの平行移動のやり方」と「点の回転のしかた」が分かれば、グラフの回転もできるようになります。落ち着いて復習してみてください。

2　2 変数函数のグラフの切断

2 変数函数 $z = f(x, y)$ のグラフは 3 次元的形状を持つため、1 変数函数のグラフと比べて理解が難しくなります。それを何とかするための一つの手段として「平面で切断した様子を観察する」という手があります。

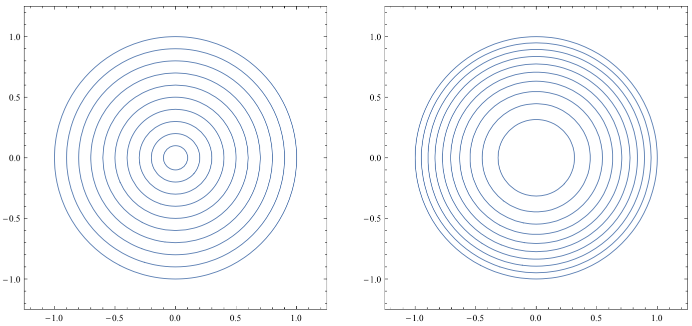
前回の内容を思い出しましょう。「x, y, z の 1 次方程式が定める図形は平面である」という事実を確認しました。ですからグラフの方程式 $z = f(x, y)$ と平面の方程式 $ax + by + cz + d = 0$ を連立すれば、グラフ $z = f(x, y)$ の平面 $ax + by + cz + d = 0$ による断面が得られます。特に平面の方程式として分かりやすいもの、たとえば

- xy 平面、yz 平面、zx 平面のいずれかに平行な平面
- z 軸を含む平面

を上手く取って来れば、グラフの形状をよく理解できる可能性が高まります。

2.1　座標軸の張る平面に平行な切断

■等高線　実数 c に対し、方程式 $z = c$ は z 座標の値が c である、xy 平面と平行な平面を表します。したがって $z = f(x, y)$ と連立して $c = f(x, y)$ とすれば、平面 $z = c$ による断面が見えます。この断面のことをグラフの**等高線**といいます。等高線が把握できれば、c の値を動かすことでグラフの形状が把握できます。たとえば $z = \sqrt{x^2 + y^2}$ と $z = x^2 + y^2$ のグラフを、xy 平面に平行な平面を $z = 0$ から $z = 1$ まで 0.1 刻みで動かしてみましょう。

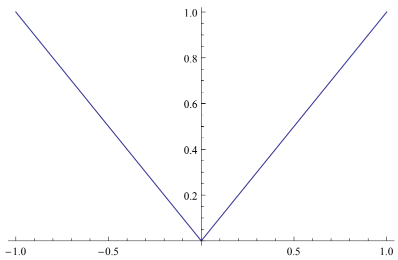


(a) $z = \sqrt{x^2 + y^2}$　　(b) $z = x^2 + y^2$

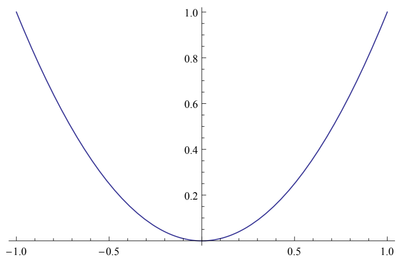
この図から明らかに、どちらのグラフも z 軸回りの回転対称性を持つと分かります。また等高線の密度から、$z=\sqrt{x^2+y^2}$ では高さが一定の間隔で増し、$z=x^2+y^2$ では急激に高さが増していく感じがします。ですがどの程度増えるのかは、この図だけでははっきりしません。詳しい情報を得るためには、他の平面での切断面などが必要です。

■x 軸や y 軸に垂直な平面での切断　全く同様に平面 $x=c$ あるいは $y=c$ ($c\in\mathbb{R}$ は定数) を使うことで、それぞれ x 軸と y 軸に垂直な平面でグラフを切断できます。

さっきの $z=\sqrt{x^2+y^2}$ と $z=x^2+y^2$ で試してみましょう。まずは $x=0$ とおくと、式はそれぞれ $z=\sqrt{y^2}=|y|$, $z=y^2$ となります。これらが yz 平面による断面です。そのグラフは次の通りです。

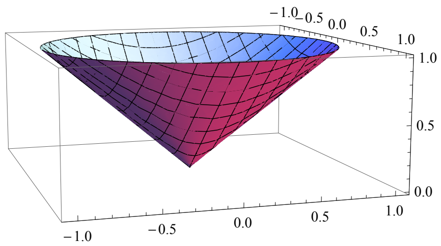


(c) $z=|y|$

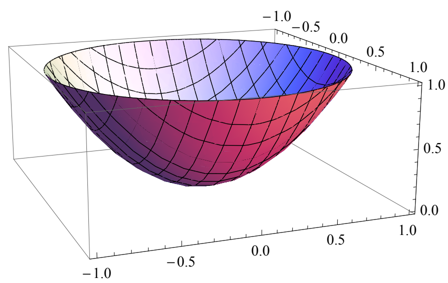


(d) $z=y^2$

あとは z 軸回りの回転対称性があるので、このグラフを z 軸回りにぐるっと一周回転させれば求めるグラフの正体が分かります。$z=\sqrt{x^2+y^2}$ は円錐、$z=x^2+y^2$ は回転放物面です。



(e) $z=\sqrt{x^2+y^2}$



(f) $z=x^2+y^2$

また x, y 軸に垂直な平面に関するグラフの切断は、2 変数函数 $f(x,y)$ を x,y の順に積分するときに役立ちます。特に体積の計算をするときは、まず真っ先にこれらの切断を考えるでしょう。ほとんどの人が大学入試の勉強でやったことがあるはずです。

2.2　z 軸を含む平面に関する切断

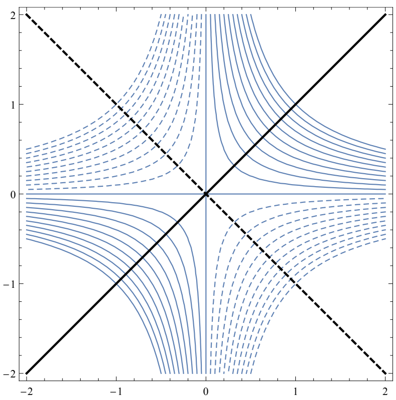
続いて平面の方程式のうち、$ax+by=0$ の形のものを考えます。これの法ベクトルは ${}^t(a,b,0)$ で、z 軸の方向ベクトル ${}^t(0,0,1)$ と直交します。さらに平面 $ax+by=0$ は原点を通るので、この平面は z 軸を含むことが分かりました。したがって方程式 $z=f(x,y)$ で $ax+by=0$ を連立して x ないし y を消去すれば、z 軸を含む平面でどういう形をしているか分かります。この式で $b=0$ や $a=0$ とした場合は、先ほどの x, y 軸に垂直な平面による切断と一致します。

たとえば方程式 $z=xy$ のグラフを考えましょう。次のように変形してみます。

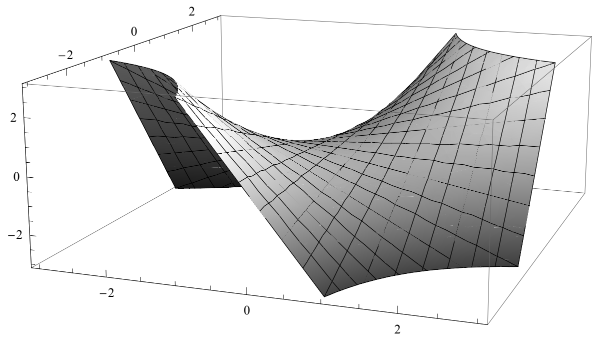
$$z=xy=\frac{(x+y)^2-(x-y)^2}{4}$$

この式に $x=y$ を代入すると $z=x^2$、$x=-y$ を代入すると $-x^2$ になります。したがって平面 $x=y$ での断面は上向きの放物線、それに直交する平面 $x=-y$ での断面は下向きの放物線となります。ただし、**この放物線がそのまま $y=\pm x^2$ にはならない**ことに気を付けてください。直線 $y=x$ 上で原点からの距離が 1 である点のうち、座標が正のものは $(1/\sqrt{2},1/\sqrt{2})$ です。ですから xy 平面の直線 $y=\pm x$ に x 軸や y 軸と同じ目盛りを刻むには、点 $(1/\sqrt{2},1/\sqrt{2})$ を基準に取らなければいけません。そうすると $(x,y)=(t/\sqrt{2},\pm t/\sqrt{2})$ のとき $xy=\pm t^2/2$ となりますから、$z=xy$ の平面 $x=\pm y$ による断面の放物線は $y=\pm x^2/2$ と合同になります。

一方で等高線は $xy=c$ という反比例の式で定まるので、双曲線または 2 直線になります。これを踏まえ、次の平面図を見てください。実線が $z\geq 0$ の部分、点線が $z<0$ の部分の等高線です。そして斜めの実線と点線がそれぞれ平面 $x=y$ と $x=-y$ に対応し、これらで切断した断面がそれぞれ上向きと下向きの放物線になるわけです。

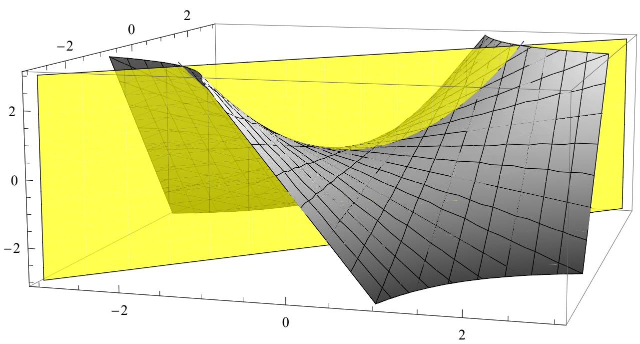


(g) $z=xy$ を上から見た図

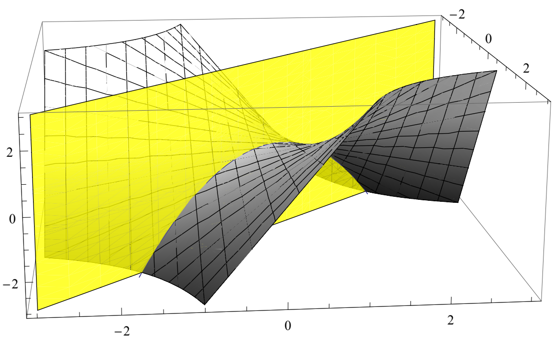


(h) $z=xy$ を俯瞰した図

そして右の図は、曲面 $z=xy$ を俯瞰した図です。このグラフは双曲放物面と呼ばれます[*1]。立体的な図を見て等高線が双曲線になること、それから断面に放物線が現れることを納得してください[*2]。



(i) $z=xy$ の平面 $x=y$ による断面



(j) $z=xy$ の平面 $x=-y$ による断面

[*1] テキストに「双曲 2 次形式」とも書いてありますが、これはどちらかというと $x^2-y^2-z=0$ という「式の形」を表す言葉です。岩波書店から出ている『数学辞典 第 4 版』を引くと 347 A の項に「双曲放物面」と書いてあるので、その呼び名を使いました。また、このグラフで原点は「接平面が xy 平面と平行であるが、極大でも極小でもない」という意味で**鞍点**と呼ばれます。ただしこれは一般名詞であって、グラフを固有名詞的に「馬の鞍」と呼ぶわけではありません。気を付けてください。

[*2] PDF ファイルの原本 `https://github.com/HideakiHosaka/2015_linear_algebra/raw/master/2015linear_algebra.pdf` も必要に応じて見てみてください。カラーな上に拡大可能なので、印刷版より図が綺麗なはずです。

3 座標変換とグラフの移動

2 変数函数のグラフを理解する別の手段として、グラフの移動を考えます。1 変数の場合、この手法の最も典型的な例は平方完成です。2 次函数を平方完成をすると、その 2 次函数が「放物線 $y = ax^2$ をどうやって動かして作られたか」が分かり、それによってグラフの形状がはっきり分かるのでした。そこで 2 変数函数についても、グラフを動かしましょう。2 変数の場合は平行移動だけでなく回転移動もできますから、回転移動の方法についても調べます。

また 2 変数以上の函数だと、グラフを移動する以外にも「座標を取り換える」という操作ができます。1 変数函数だと、せいぜい値域の目盛りを対数軸にして対数プロットをすることがあるくらいで、定義域の目盛りを変えることはありませんでした。ところが 2 変数函数だと、座標の取り換えが極めて有効なケースが出てきます。

3.1 円柱座標

平面上の極座標は、$(x, y) = (r\cos\theta, r\sin\theta)$ という式で定義されていました。そこで直交座標 (x, y, z) のうち x, y だけを極座標で置き換えることで (r, θ, z) という座標系が定まります。これを円柱座標といいます[*3]。

$z = f(x, y)$ の式の中にあからさまに $x^2 + y^2$ が出てくるなどする場合は、円柱座標が役立ちます。たとえばさっきの $z = \sqrt{x^2 + y^2}$ というグラフを円柱座標で考えてみると、$x^2 + y^2 = r^2$ より $z = \sqrt{r^2} = |r|$ となります。だから z の値は θ に依存せず、グラフが z 軸回りの任意の角度に対する回転対称性を持つことが分かります。そして z 軸を含む平面でグラフを切断すれば、$z = |r|$ という絶対値のグラフが現れます。これを回転させることで、元々のグラフは円錐を表していたことが再び分かります。このように z 軸回りの回転対称性を持つグラフは、円柱座標を使う事ですっきりと理解できます。

$z = x^2 + y^2$ のグラフでも状況は同じです。こっちでは $z = r^2$ となるのでやはり回転対称性を持ち、さらに z 軸を含む平面での断面は放物線になります。よって放物線を軸に回転させて、元のグラフが回転放物面だと分かります。

3.2 平行移動

函数 $y = f(x)$ のグラフを右に a だけ平行移動すると、式は $y = f(x - a)$ に変化します。右に a ずらすときに f の中に $x - a$ が入るのでたまに勘違いしそうになりますが、そういうときは $x = 0$ 等で検証してみましょう。平行移動後に $x = 0$ の位置**に写ってくる**のは、元々左に a だけ移動したところにある $x = -a$ の位置の点です。だから平行移動後のグラフで $x = 0$ における値は $f(-a)$ となります。

多変数函数でも全く事情は同じです。$z = f(x, y)$ のグラフを x 軸正の方向に a, y 軸正の方向に b だけ平行移動させるとき、平行移動で点 (x, y) **に写ってくる**点は $(x - a, y - b)$ です。よって平行移動後のグラフは $z = f(x - a, y - b)$ になります。

3.3 回転

1 変数のグラフは変数が x しかないので、グラフの移動といっても平行移動くらいしかすることがありません。ですが 2 変数のグラフになると、座標の回転ができるようになります。原点中心の回転を調べてみましょう。平行移動のときと同じく、見るべきは「回転移動でどの点がどこに写るか」ではなく「どの点が**どこから写ってくるか**」です[*4]。

複素平面上で原点中心の角度 θ 回転は、$e^{i\theta}$ の掛け算で与えられるのでした。これと $e^{i\theta}e^{-i\theta}(x + iy) = x + iy$ より、θ 回転で点 (x, y) **に写ってくる**点は $e^{-i\theta}(x + iy) = \big((\cos\theta)x + (\sin\theta)y\big) + i\big(-(\sin\theta)x + (\cos\theta)y\big)$ です。した

[*3] 3 次元の極座標 (r, θ, φ) とは別物です。

[*4] 後々線型代数の授業で座標変換を考えるときも、これと似たような状況が現れます。

がって $f(x,y)$ を原点中心に θ 回転させて得られる函数は、$f(x,y)$ の中に今の座標を代入して得られる $f\big((\cos\theta)x+(\sin\theta)y, -(\sin\theta)x+(\cos\theta)y\big)$ です。

さっきの $z=xy$ のグラフは、回転を使うとよく理解できます。xy 座標系を正の向きに $\pi/4$ だけ回転させて uv 座標系を作ります。すると導いた回転の公式で $\theta=\pi/4$ を代入することで、uv 座標系での点 (u,v) は xy 座標系で $\big((u+v)/\sqrt{2}, (u-v)/\sqrt{2}\big)$ に化けます。したがって $z=xy$ は uv 座標系で $z=(u^2-v^2)/2$ です。こうすれば曲面 $z=xy$ と $z=x^2-y^2$ が回転と縦方向の拡大・縮小で写り合うことが分かります。既に確認した通りですね。ちなみに $z=(u^2-v^2)/2$ に直してしまえば、どんな定数 $c\in\mathbb{R}$ に対しても、平面 $u=c$ での断面が放物線 $z=-v^2/2$ と同じ形だと分かります。実際 $z=-v^2/2+c^2/2$ は、放物線 $-v^2/2$ を上下方向に平行移動したものです。これを xy 座標系で見れば、平面 $y=-x+c$ $(c\in\mathbb{R})$ による断面となります。

また今は z 軸を中心とする回転を調べましたが、軸が z 軸と平行である限り、どこであっても回転の計算はできます。既に僕たちはグラフの平行移動の仕方を知っていますから、回転軸が一旦 z 軸と重なるように平行移動し、z 軸回りの回転をして、さらに最初の平行移動とは逆向きの平行移動をすれば OK です。

4　曲面上の運動と接ベクトル

4.1　曲面に沿う曲線と接ベクトル

2 変数函数 $z=f(x,y)$ と平面 $\mathbb{R}^2$ 内の曲線 γ を考えます。各 $t\in\mathbb{R}$ に対して平面 $\mathbb{R}^2$ 上の点 $\gamma(t)\in\mathbb{R}^2$ が定まるから、$\gamma(t)=\big(x(t),y(t)\big)$ と書けます。そこで $\gamma(t)$ の座標を $z=f(x,y)$ に代入する[*5]と、$f\big(x(t),y(t)\big)$ は点 $\gamma(t)$ における f の値を表します。こうしてしまえば $f\big(x(t),y(t)\big)$ は t の函数ですから、微分することができます。まだ証明していませんが、実は $f\big(x(t),y(t)\big)$ の t における微分は f と $\gamma'(t):=\big(x'(t),y'(t)\big)$ だけで決まります[*6]。そこで $f\big(x(t),y(t)\big)$ の微分を、$\gamma'(t)$ 方向の**方向微分**と言います。t を時刻、γ を点の運動だと思えば、$\gamma'(t)$ は時刻 t における速度ベクトルです。直感的に言うと $f\big(\gamma(t)\big)$ の微分は「$\gamma'(t)$ 方向に少し動くとどれだけ f が変化するか」を表す量ですから、方向微分という言葉がしっくりくるのではないでしょうか。

また $\tilde{\gamma}(t):=\big(x(t),y(t),f(x(t),y(t))\big)$ と置くと、$\tilde{\gamma}(t)$ は常に曲面 $z=f(x,y)$ 上の点を表します。したがって t を動かすことで、曲面 $z=f(x,y)$ に沿う曲線が得られます。そこで曲線 $\tilde{\gamma}$ の座標を t で微分すると、$\tilde{\gamma}$ の接ベクトルができ、それが曲面 $z=f(x,y)$ の接ベクトルにもなります。このようにして、曲面の接ベクトルが得られます。

4.2　偏導函数と接平面

今の方向微分の話で、特に $\gamma(t)=(a+t,b)$ あるいは $\gamma(t)=(a,b+t)$ の場合を考えます。つまり γ は x 軸や y 軸の方向を向いた直線の上を、速度 1 で進みます。このとき $f\big(\gamma(t)\big)$ を $t=0$ で微分した値を、それぞれ点 (a,b) における x 方向、y 方向の**偏微分係数**といいます。すなわち

$$\frac{\partial f}{\partial x}(a,b):=\frac{d}{dt}f(a+t,b)\Big|_{t=0},\quad \frac{\partial f}{\partial y}(a,b):=\frac{d}{dt}f(a,b+t)\Big|_{t=0}$$

です。新しい記号が出てきましたが、計算自体は難しくありません。x での偏微分は「y を定数と思って、x の函数として微分する」というだけです。また 1 変数の場合、微分係数は接線の傾きでした。だから x での偏微分係数は、点 (a,b) における曲面 $z=f(x,y)$ の x 軸方向の勾配を表します。

これを知っていると、2 変数函数のグラフ $z=f(x,y)$ の接平面が求められます。$\tilde{\gamma}$ の微分が曲面 $z=f(x,y)$ の接べ

[*5] 写像の言葉で言うなら、曲線を $\gamma\colon\mathbb{R}\to\mathbb{R}^2$ という写像と思い、2 変数函数 $f\colon\mathbb{R}^2\to\mathbb{R}$ の合成 $f\circ\gamma$ を考えています。
[*6] 合成函数の微分さえできれば良いのですが、それは多変数の場合「連鎖律」と呼ばれ、少しだけ計算が複雑になります。

クトルでした。その式に x 方向の偏微分と y 方向の偏微分を代入すると

$$\left(1, 0, \frac{\partial f}{\partial x}(a,b)\right), \quad \left(0, 1, \frac{\partial f}{\partial y}(a,b)\right)$$

が曲面 $z = f(x,y)$ の点 (a,b) における接ベクトルだと分かります。これらのベクトルは 1 次独立ですから、点 (a,b) における接平面はこれら 2 本のベクトルによって張られます。したがって外積を使って、接平面の法ベクトルが

$$\left(\frac{\partial f}{\partial x}(a,b), \frac{\partial f}{\partial y}(a,b), -1\right)$$

と分かります。これで接平面の方程式における x, y, z の係数が決まりました。あとは接平面が点 $\bigl(a, b, f(a,b)\bigr)$ を通るよう調整すれば、接平面の方程式

$$\frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) - \bigl(z - f(a,b)\bigr) = 0$$

が求まります。

5 解答など

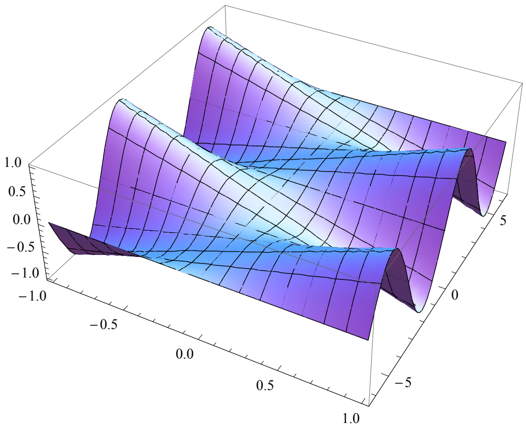
5.1 穴埋め問題の解答

既に一通りの問題を解説していますが、穴埋め問題の解答だけをもう一度まとめておきます。

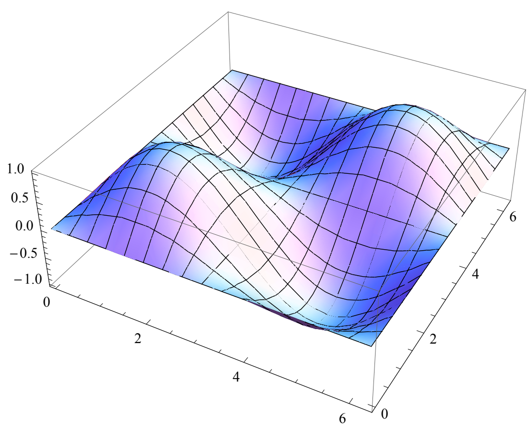
(1) 平面 (2) 双曲放物面 (3) z 軸 (4) 回転対称性 (5) 円錐 (6) 回転放物面 (7) a (8) b
(9) $\sqrt{c}$ (10) $2xy$ (11) 双曲放物面 (12) 双曲線 (13) a (14) b (15) $1/2$ (16) $-1/2$

5.2 グラフ

$z = xy$ のグラフは既に描きました。$z = x \sin y$ と $z = \sin x \sin y$ のグラフは次の通りです。グラフの雰囲気が分かるよう、問題より少し範囲を広げて描画しています。



(k) $z = x \sin y$



(l) $z = \sin x \sin y$

ちなみに、これらのグラフは Mathematica を使って描いています。Mathematica の使い方については、たとえば「はいぱーワークブック」の 29 章[*7]や、そこに紹介されている参考文献などを読んでください。

[*7] http://hwb.ecc.u-tokyo.ac.jp/current/applications/mathematica/

数理科学基礎 (線形代数学) 第 5 回　　行列とその演算

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 5 月 20 日 1 限

今回のテーマは行列の計算です。行列の計算を理解できていない人はいないようでしたが、**計算ミスをする人が少なくありませんでした**。行列の計算は今後多用しますから、この機会に十分慣れておいてください。

1　総和記号 $\sum$ の使い方

行列の積の計算やこの後でやる行列式の計算では、とにかく $\sum$ 記号が式中にたくさん現れます。そこで最初に、$\sum$ の色々な使い方をまとめておきましょう。必要に応じて読んでください。

1.1　最もありふれた使い方

総和記号 $\sum$ の使い方として、最もありふれたものは

$$\sum_{i=1}^{n} a_i := a_1 + a_2 + \cdots + a_n$$

というものでしょう。n 個の数 $a_1, a_2, \ldots, a_n$ が与えられたとき、それの全ての和を上に書いたように、$\sum$ 記号で表すのでした。$\sum$ 記号は数学でよく現れる「数列の和」を表せるという意味で便利な記号ですが、それ以上に「記号操作で複雑な計算を片付けられる」という点が強力です。そのことを、以下で見ていきましょう。

■線型性　$\sum$ の最も大事な性質は、次の**線型性**です。

$$\sum_{k=1}^{n}(a_k + b_k) = \sum_{k=1}^{n} a_k + \sum_{k=1}^{n} b_k, \qquad \sum_{k=1}^{n} \alpha a_k = \alpha \sum_{k=1}^{n} a_k$$

式自体の意味は単純で、1 つ目は足し算の順序の入れ替え、2 つ目はかっこでくくる操作を表しています。

$$(a_1 + b_1) + (a_2 + b_2) + \cdots + (a_n + b_n) = (a_1 + a_2 + \cdots + a_n) + (b_1 + b_2 + \cdots + b_n)$$
$$\alpha a_1 + \alpha a_2 + \cdots + \alpha a_n = \alpha(a_1 + a_2 + \cdots + a_n)$$

実際に式変形をするときは、線型性の公式を左から右、右から左の両方で使います。すなわち今の式を

- 変数が同じ範囲を走る $\sum$ が足されているときは、それをひとまとめにしてよい
- $\sum$ の中に足し算があれば、それを 2 つの $\sum$ に分割できる
- $\sum$ の変数と関係ない数は、いつでも $\sum$ を前後に飛び越えられる

と読むわけです。

■変数のシフトと和の順序の反転　$\sum$ の変数は、ずらすことが可能です。たとえば

$$\sum_{k=1}^{n} a_k = a_1 + a_2 + \cdots + a_n, \qquad \sum_{k=0}^{n} a_{k+1} = a_1 + a_2 + \cdots + a_n$$

ですね。項を書き出してしまえば当たり前ですが、k が走る範囲をずらしても、それに合わせて $\sum$ の中の式に出てくる k を適当にずらせば、表す式は全く同じになります。

同じようにして、足す順序を逆順にすることもできます。$a_1 + a_2 + \cdots + a_n = a_n + a_{n-1} + \cdots + a_1$ ですから

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n} a_{n+1-k}$$

という式が成り立ちます。

これらの変数の変換は、主に線型性と組み合わせて 2 つの $\sum$ をくっつけるときに使います。たとえば $\sum_{k=1}^{n} k$ の公式は、次のようにして導くことができます。

$$2\sum_{k=1}^{n} k = \sum_{k=1}^{n} k + \sum_{k=1}^{n} k = \sum_{k=1}^{n} k + \sum_{k=1}^{n}(n+1-k) = \sum_{k=1}^{n}\bigl\{k + (n+1-k)\bigr\} = \sum_{k=1}^{n}(n+1) = n(n+1)$$

もう一つ別の証明を挙げておきます。こっちの手法は、$\sum k^2$ や $\sum k^3$ の公式を導くのにも使えます。

$$\begin{aligned}
(n+1)^2 - 1 &= \sum_{k=1}^{n}\{(k+1)^2 - k^2\} = \sum_{k=1}^{n}(2k+1) = 2\sum_{k=1}^{n} k + \sum_{k=1}^{n} 1 = 2\sum_{k=1}^{n} k + n \\
\therefore \sum_{k=1}^{n} k &= \frac{(n+1)^2 - 1 - n}{2} = \frac{n^2 + n}{2}
\end{aligned}$$

どちらの証明でも、1 つ 1 つの式変形が「和の順序の反転」や「線型性で $\sum$ をまとめる / ばらす」といった操作だけしか行っていないことに気を付けてください。このように「$\sum$ の公式をそのまま当てはめる」という操作を繰り返すだけで証明を完了できることが、$\sum$ を使う大きなメリットなのです。

1.2 集合を走る変数に関する和

さて $\sum_{k=1}^{n} a_k$ は「k を順番に $1, 2, \ldots, n$ と動かし、出てくる項を全部足す」という意味でした。ですが「順番に足す」という意味は別に大事ではありません[*1]。どうせ足し算だから、順序を入れ替えたって結果は変わらない[*2]からです。そこで順番を考えず、「集合のそれぞれの元に対応する項を足す」という意味でも $\sum$ を使います。たとえば

$$\sum_{k \in \{1,3,5\}} a_k = a_1 + a_3 + a_5$$

という感じです。$\{1, 3, 5\}$ の元は 1 と 3 と 5 だから、k が $1, 3, 5$ を動くのに合わせて対応する a_k、つまり a_1 と a_3 と a_5 を足し合わせるというのが左辺の $\sum$ の意味です。新しい $\sum$ を使えば、これまで使っていた $\sum$ は

$$\sum_{k=1}^{n} a_k = \sum_{k \in \{1,2,\ldots,n\}} a_k$$

と書き表せます。だから新しい $\sum$ の使い方は、今までの $\sum$ が拡張されたものになっています。後々見ていくように、$\sum$ の変数に整数以外のものを使えるのは、実は非常に役立ちます。

さらに省略した書き方として、$\sum$ の下に「変数の満たすべき条件」を書く場合があります。たとえば

$$\sum_{\substack{1 \le k \le 5 \\ k\ \text{は奇数}}} a_k = a_1 + a_3 + a_5$$

[*1] ただし、たまに $n = 0$ のとき「$\sum_{k=1}^{n} a_k = 0$」という式が出ます。このような場合は $\sum$ を 1 から「順に足す」という意味で捉えないと、式を正しく解釈できません。

[*2] ここで考えているのは有限個の足し算だけです。無限級数などの代数的に扱えない場合は、全く考えていません。

と書きます。集合 $\{k \in \mathbb{N} \mid 1 \leq k \leq 5, k \text{ は奇数 }\}$ は $\{1,3,5\}$ と同じです。この集合を $\sum$ の下の狭いスペースに書くと非常に見苦しいので「縦棒の右側に書かれる条件だけを $\sum$ の下に書いてしまおう」という魂胆です。だから

$$\sum_{1\leq k\leq 5} a_k = \sum_{k\in\{1,2,3,4,5\}} a_k = \sum_{k=1}^{5} a_k = a_1 + a_2 + a_3 + a_4 + a_5$$

などという意味になります。このとき一番左の表記法では “$k \in \mathbb{N}$” という条件が地味に抜け落ちますが、その部分については「読者が空気を読んで察する」という暗黙のルールがあります[*3]。頑張って読み解いてください。

さらに、この $\sum$ 記法に慣れるとこんなこともできます。

$$\sum_{k\in A\cup B} a_k = \sum_{k\in A} a_k + \sum_{k\in B} a_k - \sum_{k\in A\cap B} a_k$$

$\sum$ の変数が集合を走る場合、集合の分割に応じて $\sum$ を分割できます。もし和にダブりがあればその部分を引いて補正しなければいけませんが、ダブりなくばらせば第 3 項は消えます。このような $\sum$ の分割は、$\sum$ を部分的に計算する場合に役立ちます。

1.3 変数について

ここで、$\sum$ の変数について 2 つほど注意をしておきます。

■ダミー変数　次の式を見てください。

$$\sum_{k=1}^{n} a_k = \sum_{l=1}^{n} a_l$$

当たり前の式ですが、ここで大事なのは $\sum$ **で走る文字は自由に取り換えても良い**という点です。左辺の $\sum$ では、k は「1 から n までの整数を動くこと」だけを示すのに使われているのであって、k という文字自体が式全体で意味を持っているわけではありません。k という文字は、$\sum$ の中だけで有効です。こういう変数を**ダミー変数**と言います。

$\sum$ を 1 個単独で使ってる場合は、ダミー変数の文字を変えるご利益はそんなにありません。せいぜい複素数を足し算する時、虚数単位とこんがらがらないよう変数を i から j に変えるくらいでしょうか。ですが $\sum$ が入った複数個の式をかけて変形するようになると、割と頻繁に変数の衝突が起きたりします。そういう場合に文字の取り換えは、地味ですが計算を遂行するのに大変役立ちます。

■変数のスコープ　今のダミー変数について、$\sum$ の変数は**式中の一部分でのみ意味がある**と言いました。この変数が有効な範囲を、プログラミングの言葉で**変数のスコープ**といいます。たとえば

$$\sum_{k=1}^{n} \underline{(2k-1)} = n^2$$

の左辺における k のスコープは、下線を引いた $2k-1$ の場所です。

当たり前のことですが、$\sum$ で使われる変数はスコープ外で意味を持ちません。ですから式を見直してみて「$\sum$ の外にダミー変数が飛び出していた」とか「$\sum$ の計算をし終わった後にダミー変数が残っていた」などという場合、スコープを考えるだけで明らかに計算ミスがあると分かります。今みたいな単純な式だと気になりませんが、$\sum$ の入った複雑な式では変数のスコープを意識すると計算がしやすくなります。心の片隅に置いといてください。

[*3] 勉強する人に向かって「空気読め」は割とヒドい言葉のような気がするのですが、ちょっとでも新しい $\sum$ 記号を使ってみると、逆に「一々 $\sum$ の下に集合を書くなんてかったるい」と思うようになってしまうんですよね……。

1.4 二重和

ここまで $\sum$ の使い方を色々説明してきましたが、実用面を考えると、さらに一歩踏み込んで多重の $\sum$ を取り扱う必要があります。

計算の定義自体は今までの $\sum$ と変わりません。たとえば

$$\sum_{k=1}^{m}\sum_{l=1}^{n} a_{k,l} = \sum_{k=1}^{m}\left\{\sum_{l=1}^{n} a_{k,l}\right\}$$

というように「内側の $\sum$ をまず足し、その結果を外側の $\sum$ でさらに足し合わせる」というだけです。$\sum$ を元々「1 列に並べた数を足し合わせる」という方法で使っていたことになぞらえれば、二重和は「長方形に並べた数を足し合わせる」という感じです。ただ二重和の場合

$$\sum_{k=1}^{m}\sum_{l=1}^{k} a_{k,l} = \sum_{k=1}^{m}\left\{\sum_{l=1}^{k} a_{k,l}\right\} = (a_{1,1}) + (a_{2,1} + a_{2,2}) + (a_{3,1} + a_{3,2} + a_{3,3}) + \cdots$$

のように、「外側の $\sum$ のダミー変数を使って、内側の $\sum$ における変数の範囲を指定する」という技ができたりします。上の例では、三角形に並べた数を全部足すという意味になります。

このような場合に、足し算の順序を上手くコントロールするための変数の扱い方を見ていきましょう。

■和の順序の交換　一番基本的な二重和

$$\sum_{k=1}^{m}\sum_{l=1}^{n} a_{k,l} = \sum_{k=1}^{m}(a_{k,1} + a_{k,2} + \cdots + a_{k,n}) = \begin{array}{l} \quad a_{1,1} \ + \ a_{1,2} + \cdots + a_{1,n} \\ + \, a_{2,1} \ + \ a_{2,2} + \cdots + a_{2,n} \\ + \ \cdots \ \ \cdots \\ + \, a_{m,1} \ + \ a_{m,2} + \cdots + a_{m,n} \end{array}$$

を考えてみましょう。この計算は、各行を横向きに足し合わせた結果を縦に足し合わせても、各列を縦向きに足し合わせた結果を横に足し合わせても全く同じ結果になります。これを $\sum$ で表現すると

$$\sum_{k=1}^{m}\sum_{l=1}^{n} a_{k,l} = \sum_{l=1}^{n}\sum_{k=1}^{m} a_{k,l}$$

となります。一方、

$$\sum_{k=1}^{n}\sum_{l=1}^{k} a_{k,l} = \sum_{k=1}^{n}(a_{k,1} + a_{k,2} + \cdots + a_{k,k}) = \begin{array}{l} \quad a_{1,1} \\ + \, a_{2,1} + a_{2,2} \\ + \ \cdots \\ + \, a_{n,1} + a_{n,2} + \cdots + a_{n,n} \end{array}$$

のような場合、外側と内側の $\sum$ を安直に入れ替えることはできません。この事実は「内側の $\sum$ の上限を与える k が、外側の $\sum$ のスコープの外に出られない」とも言えます。ですから「内側の $\sum$ の範囲指定に外側の $\sum$ の変数が使われていない限り、2 つの $\sum$ の順番を入れ替えることができる」というのが結論です。$\sum$ の変数が集合を走る場合でも、同じことが言えます。

■$\sum$ の結合と分割　次の式を見てください。

$$\sum_{k=1}^{m}\sum_{l=1}^{n} a_{k,l} = \sum_{1\le k\le m}\ \sum_{1\le l\le n} a_{k,l} = \sum_{\substack{1\le k\le m \\ 1\le l\le n}} a_{k,l}$$

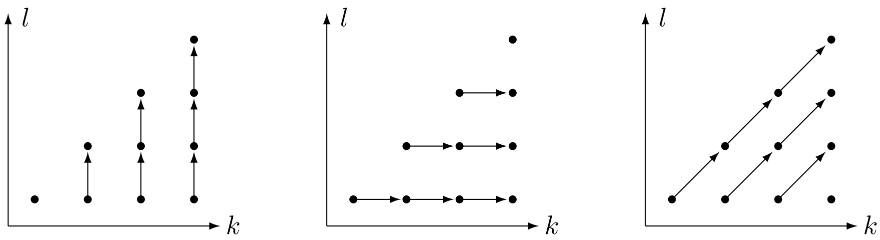
こんな感じで、2 つの $\sum$ をくっつけることができます。元々の二重和を「k が $\{1,2,\ldots,m\}$ を走り、l が $\{1,2,\ldots,n\}$ を走る」と書き換えた上で、さらに集合の直積を利用して「(k,l) が $\{1,2,\ldots,m\}\times\{1,2,\ldots,n\}$ を走る」と書き換えました。こういう風に集合の直積を使うと、2 つの $\sum$ を 1 つにくっつけられます。逆に $\sum$ の変数が直積集合を走るときは、$\sum$ を 2 つにばらすことができます。

あんまり大した意味が無いようにも見えますが、三重以上に $\sum$ が重なる状況では、$\sum$ を 1 つにまとめると案外式が見やすくなったりします。また「$\sum$ の変数は集合の上を走る」という意識を持っておくと、$\sum$ の変数を操作するときに、式が正しいかどうかを変数の走る集合が等しいかどうかという問題に帰着させて考えられます。

■変数変換 最後に、$\sum$ の変数変換を紹介します。次の二重和を見てください。

$$\sum_{k=1}^{n}\sum_{l=1}^{k} a_{k,l} = \begin{aligned} &\ a_{1,1} \\ &+ a_{2,1} + a_{2,2} \\ &+ \cdots \\ &+ a_{n,1} + a_{n,2} + \cdots + a_{n,n} \end{aligned}$$

この足し方を図示したのが、次の一番左の図です。内側の $\sum$ で足される項を、矢印で繋いでいます。



そして図中にも描いたように、この二重和は色々な足し方ができます。この図をじっくり見れば

$$\sum_{k=1}^{n}\sum_{l=1}^{k} a_{k,l} = \sum_{l=1}^{n}\sum_{k=l}^{n} a_{k,l} = \sum_{p=1}^{n}\sum_{k=1}^{n+1-p} a_{p-1+k,k}$$

が分かるはずです。集合で言えば、$\{(k,l)\in\{1,2,\ldots,n\}\mid l\le k\}$ を鉛直な直線や、あるいは軸と 45° の傾きをなす直線でスライスして分割しているわけです。このように二重和では「格子の各点に項が対応している」と捉えることで、外と中の $\sum$ をそのまま入れ替えることはできなくても、足し方を色々変えることができます。特に、斜めの足し算はしばしば式変形のキーポイントになります。

$\sum$ の変数変換は実際にやってみると案外間違えやすいのですが、そういう時はきちんと格子を描き「どの範囲の項が足されるか」を絵で表しましょう。そうすればぐっと間違いは減るはずです。

1.5 総積記号

ちなみに $\sum$ の掛け算バージョンもあって、それは $\prod$ (パイ) という記号で表されます。使い方は $\sum$ と全く同じで

$$\prod_{k=1}^{n} a_i = a_1a_2\cdots a_k$$

という意味です。たとえば $\prod_{k=1}^{n} k = n!$ など。$\sum$ のときと同様、変数は集合を走ることもあります。

この記号も色々使いどころがあります。特に

$$\sum_{k=1}^{n}\log a_k = \log\prod_{k=1}^{n} a_k, \qquad \prod_{k=1}^{n} p^{a_k} = p^{\sum_{k=1}^{n} a_k}$$

という形の式変形は知っておくと、計算がサクサク進むと思います。

2 行列とその演算

数を縦に m 個、横に n 個並べたものを $m \times n$ 型あるいは (m, n) 型の**行列**といいます。また、行列の中にある一つ一つの数を**成分**といいます。特に i 行 j 列[*4]にある成分を指して (i, j) 成分という呼び方をします。成分を表すときは

- A_{ij} のように右下に ij をつけることで、行列 A の (i, j) 成分を表す
- (i, j) 成分が a_{ij} である行列のことを、$(a_{ij})_{\substack{1 \le i \le m \\ 1 \le n \le n}}$ と表す

といった記法も使います。覚えておきましょう。

行列は数学の色々なところで使います。たとえば連立 1 次方程式を解いたり、微分方程式を解いたりといった用途に使えますし、また他の分野への応用もたくさんあります。じっくり勉強して、行列の扱いに慣れていってください。

2.1 行列の和と積

行列に対しては和と積が定義されます。まずはその定義を確認します。同じ型の行列に対しては「同じ位置にある成分を足す」、すなわち

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{pmatrix}$$

と定めます。そして積は、(m, n) 型行列 A と (n, l) 型行列 B との間でだけ定義され、その (i, j) 成分は

$$(AB)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

です。これだと何をやっているか分からないのですが、行列を

$$\left(\begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ \hline a_{21} & a_{22} & a_{23} \end{array}\right) \left(\begin{array}{c|c|c} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{array}\right)$$

と区切れば、AB の (i, j) 成分は「A の i 行目と B の j 列目との内積」だと分かります。慣れないうちは、こうやってきちんと行と列を区切ることをお勧めします。

こんなややこしい定義をするのには相応の事情があるのですが、そんなことよりも**行列の積は役立つ**という事実が大事です。数学で使うだけでなく、Google が検索クエリを処理するとき[*5]とか、生物の記憶のシミュレーションをするときとか[*6]、その他行列の使い道は山ほどあります。理論的な場面でも使うので、計算をコンピュータに丸投げできず、人間が手計算をしなければいけない場面もしばしばあります。ですので行列の掛け算の仕方は徹底的に練習して、体で覚えてください。

2.2 行列の積の性質

行列の積で非常に特徴的なのは**順序を変えると結果が変わる**という性質です。具体例を一つやってみます。

[*4] 横の並びが行 (row), 縦の並びが列 (column) です。自動車レースのフロント・ローとか、化学で使うカラムクロマトグラフィーといった言葉を思い出すと「row が横、column が縦」を間違えにくくなるかもしれません。

[*5] 学習院大学の田崎晴明先生が執筆中の教科書 `http://www.gakushuin.ac.jp/~881791/mathbook/index.html` 内の「グーグルのページランク」の節に、解説があります。あるいは原論文 `http://infolab.stanford.edu/pub/papers/google.pdf` を見ても良いでしょう。

[*6] たとえば池谷裕二『進化しすぎた脳』(講談社ブルーバックス) の付録を参照。

■問 8 の解答 $+\alpha$

$$\left(\begin{array}{ccc}1 & 2 & 4\\ \hline 3 & 1 & 1\end{array}\right)\left(\begin{array}{c|c}-2 & 2\\ 7 & 5\\ -1 & -3\end{array}\right)=\begin{pmatrix}8 & 0\\ 0 & 8\end{pmatrix},\ \left(\begin{array}{cc}-2 & 2\\ \hline 7 & 5\\ \hline -1 & -3\end{array}\right)\left(\begin{array}{c|c|c}1 & 2 & 4\\ 3 & 1 & 1\end{array}\right)=\begin{pmatrix}4 & -2 & -6\\ 22 & 19 & 33\\ -10 & -5 & -7\end{pmatrix}$$

■非可換性　今の例では、行列の積 AB と BA の結果がサイズ違いになりました。他にも「AB は掛け算できるが BA は掛け算ができない」とか「AB と BA は同じ形の行列になるが結果は違う」とか色々な場合がありますが、とにかく大事なのは**行列の積は、ほとんどの場合順序を入れ替えられない**ことです。だから多項式 $(x+y)^n$, $(x+y)(x-y)$ などの展開公式もそのままは使えません。気を付けてください。

■零因子　さらに行列の場合、**零行列でない行列同士の積が零行列になる**こともあります。たとえば $\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}^2=O$ です。このような「上手く行列をかけると O にできる」性質を持つ行列は**零因子**と呼ばれます。零因子の存在も、普通の数や多項式とは大きく違うことです。

3　有名な公式たち

今回の計算問題の中には、実は有名な公式がいくつも含まれています。計算の答え合わせをしながら、公式とその使い方を見ていきましょう。

3.1　Cayley–Hamilton の定理

問 6 (b) に現れる等式は **Cayley–Hamilton の定理**と呼ばれています。この等式が正しいことを確かめましょう。

■問 6 の解答　(a) $(A-aE)(A-dE)=A^2-(a+d)A+(ad-bc)E$ である。よって

$$\begin{aligned}(A-aE)(A-dE)&=\begin{pmatrix}a & b\\ c & d\end{pmatrix}^2-(a+d)\begin{pmatrix}a & b\\ c & d\end{pmatrix}+ad\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}\\ &=\begin{pmatrix}a^2+bc & ab+bd\\ ac+cd & bc+d^2\end{pmatrix}-\begin{pmatrix}a^2+ad & ab+bd\\ ac+cd & ad+d^2\end{pmatrix}+\begin{pmatrix}ad & 0\\ 0 & ad\end{pmatrix}\\ &=\begin{pmatrix}bc & 0\\ 0 & bc\end{pmatrix}\end{aligned}$$

(b) $A^2-(a+d)A+(ad-bc)E=(A-aE)(A-dE)-bcE=O$

■対角和と行列式　一般に 2 次正方行列に対し

$$\operatorname{tr}\begin{pmatrix}a & b\\ c & d\end{pmatrix}:=a+d,\qquad \det\begin{pmatrix}a & b\\ c & d\end{pmatrix}:=ad-bc$$

をそれぞれ行列の**対角和** (trace)、**行列式** (determinant) といいます。対角和は文字通り、行列の対角線上にある成分を全部足し合わせたものです。行列式については第 3 回の解答で「2 本のベクトルが張る平行四辺形の符号付き面積」という意味を説明しました。これらを用いて、Cayley–Hamilton の定理は $A^2-(\operatorname{tr}A)A+(\det A)E=O$ と書けます。

証明はしていませんが、対角和と行列式はそれぞれ適切な型の行列に対し $\operatorname{tr}(AB)=\operatorname{tr}(BA)$, $\det(AB)=\det(BA)$ という式を満たします。重要な公式なので練習問題がてら証明してみてください。2 次正方行列だったらすぐ示せるはずです。

3.2 余因子行列と逆行列

■余因子行列　問 4 (3) に現れる行列 $\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ を $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ の**余因子行列**といいます。余因子行列は「元の行列にかけたら単位行列の行列式倍になる」という大事な性質を持っています。次回以降 3 次以上の正方行列に対する余因子行列の定義もやりますが、2 次の余因子行列は 3 次以上のものと比べて格段に使用頻度が高いです。どんな行列が与えられても脊髄反射で余因子行列が即答できるよう、暗記しておいてください。

まず、余因子行列の掛け算をやってみましょう。

■問 4 (3) の解答

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = (ad-bc)E = (\det A)E$$

■Cayley–Hamilton の定理の帰結　ここで、先ほどの $A^2 - (\operatorname{tr} A)A + (\det A)E = O$ を思い出しましょう。この式を移項すると $(\det A)E = A\big((\operatorname{tr} A)E - A\big)$ が得られます。そして実際に計算してみても

$$A - (\operatorname{tr} A)E = \begin{pmatrix} a+d & 0 \\ 0 & a+d \end{pmatrix} - \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

なので、2 次正方行列 A の余因子行列が $(\operatorname{tr} A)E - A$ と表せることが分かりました。

この式からは、とても大事なことが分かります。一般に同じサイズの正方行列であっても、掛け算の順番を入れ替えると結果が変わることがあるのでした。しかし A と A、あるいは A と E の掛け算は、いつでも順番をひっくり返せます。したがって $A\big((\operatorname{tr} A)E - A\big) = \big((\operatorname{tr} A)E - A\big)A = (\det A)E$ が成り立ちます。余因子行列は**左右のどちら側からかけても、単位行列の** $(\det A)$ **倍**になるのです。

■逆行列　ここで、さらに $\det A \neq 0$ の場合を考えてみます。このとき

$$A^{-1} := \frac{(\operatorname{tr} A)E - A}{\det A} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

とおくと、余因子行列の計算から $AA^{-1} = E$ が分かります。単位行列 E は数における 1 みたいなものだから、A^{-1} は A の逆数みたいなものですね。そこで A^{-1} を A の**逆行列**といいます。

A^{-1} は余因子行列を $\det A$ で割っただけですから、A^{-1} と A の積も交換可能です。よって $A^{-1}A = AA^{-1} = E$ が成り立ちます。**逆行列は、左からかけても右からかけても単位行列になる**のです。行列の積を計算するときに順番がひっくり返せないのは少々面倒ですが、逆行列についてはそういう面倒なことは起きないので安心してください。

問 4 の残りと問 7 は、余因子行列の公式に全部押し付けられます。

■問 4 (1), (2) の解答

$$\begin{pmatrix} 5 & 6 \\ 4 & 5 \end{pmatrix}\begin{pmatrix} 5 & -6 \\ -4 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

■問 7 の解答

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

■逆行列の存在条件　0 でない数に対しては常に逆数を考えられますが、行列はいつも逆行列を持つとは限りません。余因子行列から逆行列を作るには $(\det A)$ で割る操作が必要なので、$\det A = 0$ の場合に破綻が生じます。そして、$\det A = 0$ の場合はどうやっても逆行列が作れないことが、次のように示せます。

成分計算によって、2 つの 2 次正方行列 A, B に対して $\det(AB) = (\det A)(\det B)$ が確かめられます。よって、もし行列 A の逆行列 A^{-1} が存在すれば、$AA^{-1} = E$ の両辺の行列式を取って $(\det A)(\det A^{-1}) = \det E = 1$ が得られます。これより $\det A \neq 0$ が従います。

この議論で、逆行列が存在するならば $\det A \neq 0$ だと分かりました。逆に $\det A \neq 0$ の場合は、余因子行列から逆行列を作れることを既に示しています。ですから 2 次正方行列の場合に、**逆行列が存在することと** $\det A \neq 0$ **は同値**な条件だと分かりました。この条件が一般の n 次正方行列でも正しいことを、追々証明します。

3.3　交換子

同じ大きさの 2 つの正方行列 A, B に対し、$[A, B] := AB - BA$ と書きます。この $[,]$ のことを**交換子**と言います。A と B が交換する、すなわち $AB = BA$ なら $[A, B] = 0$ ですから、交換子は「2 つの行列がどれくらい交換しないか」を測るものです。実際に、次の H, X, Y で計算してみましょう。

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad X = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad Y = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

■問 5 の解答

$$\begin{aligned}
[H, X] &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &&= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & -1 \\ 0 & 0 \end{pmatrix} &&= \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} &&= 2X \\
[H, Y] &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &&= \begin{pmatrix} 0 & 0 \\ -1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} &&= -\begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix} &&= -2Y \\
[X, Y] &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} &&= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} &&= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} &&= H
\end{aligned}$$

■Lie 環 $\mathfrak{sl}_2$　今の $[H, X] = 2X$, $[H, Y] = -2Y$, $[X, Y] = H$ という公式、ただ計算すればそれで終わりなのですが、実は「Lie 環 $\mathfrak{sl}_2$ の交換関係」という名前が付いています。また行列 X, Y, H の 3 つ組は $\mathfrak{sl}_2$-triple と呼ばれます。今は詳しい説明を省きますが、物理や数学を専攻にすると、将来きっとお世話になることでしょう。

3.4　特殊な形の行列の積

最後に問 9 を解きつつ、線型代数をやる上でよく見かける計算を紹介しましょう。まずは答えを載せておきます。

■問 9 の解答　(a) から (c) までは、地道な計算で示す。

(a) ${}^t\boldsymbol{v}M\boldsymbol{v} = 4xy + z^2$

(b), (c)

$$A^2 = B^2 = C^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad {}^tAMA = {}^tBMB = {}^tCMC = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = M$$

(d) ${}^t(A\boldsymbol{v})M(A\boldsymbol{v}) = {}^t\boldsymbol{v}\,{}^tAMA\boldsymbol{v} = {}^t\boldsymbol{v}M\boldsymbol{v}$　■

■転置と行列の積との関係　行列 A に対し、その縦と横をひっくり返した行列 ${}^tA$ を A の**転置行列**といいます。ベクトルの転置と記号の使い方は同じです。転置については、明らかに ${}^t({}^tA) = A$ という式が成り立ちます。

さて A, B の積 AB が定義されるとき、${}^t(AB) = {}^tB\,{}^tA$ という式が成り立ちます。実際 A が (m,n) 型、B が (n,l) 型のとき、AB は (m,l) 型なので ${}^t(AB)$ は (l,m) 型です。A と B の転置を組み合わせて (l,m) 型行列を作るには、(l,n) 型行列と (n,m) 型行列の積である ${}^tB\,{}^tA$ という組合せしかありません。そして実際に ${}^t(AB)$ と ${}^tB\,{}^tA$ の (i,j) 成分を比較すると

$$\left({}^t(AB)\right)_{ij} = (AB)_{ji} = \sum_{k=1}^{n} A_{jk}B_{ki} = \sum_{k=1}^{n} ({}^tA)_{kj}({}^tB)_{ik} = \sum_{k=1}^{n} ({}^tB)_{ik}({}^tA)_{kj} = ({}^tB\,{}^tA)_{ij}$$

となり、確かに一致しています。この公式はよく使うので、覚えておきましょう。そうすれば (d) は (c) を使ってすぐ解けます。

■内積　(a) の問題で計算結果の型を間違える人が多かったです。少し落ち着いて、計算をフォローしてみましょう。

n 次元の縦ベクトルは、$(n,1)$ 型の行列と同じです。その転置を取ったものは n 次元の横ベクトル、すなわち $(1,n)$ 型の行列です。よって n 次元の縦ベクトル $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ に対し、${}^t\boldsymbol{u}\boldsymbol{v}$ は $(1,n)$ 型行列と $(n,1)$ 型の行列の積だから、$(1,1)$ 型行列になります。たとえば $n=3$ で $\boldsymbol{u} = {}^t(u_1,u_2,u_3)$, $\boldsymbol{v} = {}^t(v_1,v_2,v_3)$ の場合に成分で書けば

$${}^t\boldsymbol{u}\boldsymbol{v} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1v_1 + u_2v_2 + u_3v_3 \end{pmatrix}$$

です。ここで計算結果の式は $(1,1)$ 型行列なので括弧をつけましたが、$(1,1)$ 型行列とスカラーとは自然に同一視することができます。ですから普通は計算結果にわざわざ括弧をつけず、単に ${}^t\boldsymbol{u}\boldsymbol{v} = u_1v_1 + u_2v_2 + u_3v_3$ と書きます。成分が実数なら、${}^t\boldsymbol{u}\boldsymbol{v}$ はベクトルの内積と一致します。

そしてこの問題の M に限らず、n 次正方行列 M と n 次元の縦ベクトル $\boldsymbol{u}$, $\boldsymbol{v}$ が与えられたとき、${}^t\boldsymbol{u}M\boldsymbol{v}$ はやはりスカラーになります。実際 $M\boldsymbol{v}$ が n 次元の縦ベクトルですから、この式を ${}^t\boldsymbol{u}(M\boldsymbol{v})$ と読めば、計算結果が $\boldsymbol{u}$ と $M\boldsymbol{v}$ の内積になります。また $M=E$ のときは ${}^t\boldsymbol{u}E\boldsymbol{v} = {}^t\boldsymbol{u}\boldsymbol{v}$ ですから、${}^t\boldsymbol{u}M\boldsymbol{v}$ は内積の一般化だと思えます。どんな M を持って来れば ${}^t\boldsymbol{u}M\boldsymbol{v}$ が内積と似た性質を示すかを調べるのが、1 年生の線型代数の後半で学ぶテーマの 1 つです。

■特別な行列の名前　問 9 の行列 A, B, C は 2 乗すると単位行列になるという、特別な性質を持っています。せっかくなのでこれに関連して、特別な行列のクラスを紹介しておきます。

まず 2 乗したら単位行列ということは、逆行列が自分自身ということと同じです。そして A, B は共に対称行列ですから、A, B は「自分自身の転置行列が逆行列」という性質を持ちます。このような行列を**直交行列**といいます。直交行列は 1 年生の終わりに学ぶ「対称行列の対角化」という話で重要な役割を果たします。

また一般に、何乗かしたら 1 になる行列を**巾単行列**と言います。A, B, C は全て巾単行列です。こちらも重要な行列ですが、1 年生の線型代数の範囲ではそこまで活躍しないかもしれません。

■随伴作用　問 9 (c) では ${}^tAMA$ を計算しました。ここで ${}^tA = A^{-1}$ ですから、${}^tAMA$ は $A^{-1}MA$ とも書けます。この「正方行列を左から A^{-1}、右から A で挟む操作」は後々非常に良く出てきて、A^{-1} による随伴作用[*7]という名前までついています。また随伴作用を表す $\mathrm{Ad}_A(M) := AMA^{-1}$ という記号も用意されていますが、ここまで覚える必要はたぶんありません。

[*7] 普通、随伴作用では M を $A^{-1}MA$ ではなく AMA^{-1} にうつすので、$A^{-1}MA = A^{-1}M(A^{-1})^{-1}$ は A^{-1} による随伴となります。

数理科学基礎 (線形代数学) 第 6 回　　線型写像と行列

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 5 月 27 日 1 限

1　問題訂正のお詫びと雑談

最初に問題訂正のお知らせです。今回の問題の 4 番と 5 番に間違いがあります。すいませんでした。どこが間違えていたのかは、それぞれの問題の解説で詳しく書きます。5 番に取り組んだ人については、多くの人が「何かがおかしい」ということに気づいていたようです。また何人か、問題が間違っている事実を的確に見抜いた人もいました。

ところで少々むちゃくちゃなことを言いますが、**先生や TA の言うことがいつも正しいなどと思ってはダメです**。もちろん (少なくとも、このプリントを書いている TA の穂坂は) 意図的に間違ったことを言おうと思っているわけではなく、なるべくミスをなくす努力をしています。とは言っても人間である以上、何かの拍子に間違いが入ることは避けられません。ですから最終的には、皆さん自身が物事が正しいかどうかを判断しなければいけません。

特に数学では命題の真偽が白黒はっきりします[*1]から、問題や解説が間違っているときは「反例」が見つかるはずです。授業で何か怪しいことを見つけたときは「先生や TA が言ってることなんだから、自分の方が間違っているんだろう」などと思わず「本当に正しいのか」を疑ってかかってください。そして間違いを見抜くことができたら、それを正した上で問題に取り組むなどしてください。

[*1] 残念ながら (?) ごく稀に白黒はっきりつかない問題もあります。それは「Gödel の不完全性定理」によって与えられます。せっかくなので、少し「証明不可能な命題」について、雰囲気を紹介してみたいと思います。この部分は読み飛ばしても、本編には全く影響しません。

Gödel の不完全性定理は「自然数の算術 (Peano 算術) を含むいかなる体系にも、真偽のいずれも証明できない命題が存在する」ことを主張します。そして不完全性定理には「真偽の判定が不可能な命題の存在」を保証する第 1 不完全性定理と、これより一段と強く、証明不可能な命題を具体的に与える第 2 不完全性定理とがあります。僕たちがやっている数学では、そのうち真偽が判定できない命題に突き当たるのです。

もう少し詳しいことを言うと、僕たちが普段「数学」と呼んでいるものは Zermelo–Fraenkel の公理系 (ZF) に選択公理 (Axiom of Choice; AC) という公理を付け加えた “ZFC” という公理系の上に構築されています。ZFC では自然数の算術を扱えますので、Gödel の不完全性定理により、この公理系で真偽の判定が不可能な命題 P が存在します。この場合 ZFC に証明不可能な命題 P を公理として付け加えても、その否定を付け加えても、元の ZFC より強い公理系ができます。僕たちは、このどちらに従って数学をしても良いのです。もちろん必要にならない限りは「証明不可能な命題に手を触れない」という立場で問題は起こりません。

このような「真偽のいずれも証明できない命題」の中で最も良く知られたものの一つが「連続体仮説 (Continuum Hypothesis; CH)」と呼ばれるものです。これを説明するには「無限集合の個数」に関する概念が必要なので、少しだけ説明しましょう。

2 つの集合 A と B が「同じ個数」であることは、どうやれば判定できるでしょうか？ A と B が共に有限集合であるときは、元の個数を数えれば同じ個数かどうかが分かります。ところが無限集合の場合には「数える」ことができません。そこで無限集合 A, B が「同じ個数」であることを「全単射 $f\colon A \to B$ が存在すること」で定義します。「集合の元の間に $1:1$ の対応が付けば、同じ個数と呼んでいいだろう」というわけです。この「同じ個数である」ということを、数学では「2 つの集合が同じ濃度である」と言います。また無限集合の「個数」に相当するものを、集合の「濃度」と言います。

さて、自然数全体の集合 $\mathbb{N}$ と実数全体の集合 $\mathbb{R}$ はどちらも無限集合です。また $\mathbb{N} \subset \mathbb{R}$ ですから、少なくとも $\mathbb{R}$ の濃度は $\mathbb{N}$ の濃度以上であることは間違いありません。そして実は「Cantor の対角線論法」と呼ばれる手法により、「$\mathbb{N}$ から $\mathbb{R}$ への全単射が存在しない」ことが証明できます。ですから実数全体の集合 $\mathbb{R}$ の濃度は、自然数全体の集合 $\mathbb{N}$ の濃度より真に大きくなります。ところが色々なものを探しても「自然数と実数の間の濃度を持つ集合」が見つからないのです。そこで「実数と自然数の間の濃度を持つ集合は存在しないのではないか」という仮説が登場しました。これが「連続体仮説」です。

そして我々の数学の体系 ZFC のもと、Gödel が 1940 年に連続体仮説の否定が証明不可能なこと、Cohen が 1963 年に連続体仮説が証明不可能なことをそれぞれ示しています。ですから、この「連続体仮説」が、まさに「白とも黒ともつかない数学的命題」にあたります。数学というと「証明すること」ばかりだと思われがちですが、その中には「証明不可能なことを証明する」という不思議な話もあるのです。

ちなみに、この説明を書くにあたり

- Gödel の不完全性定理に関しては、鹿島亮『数理論理学』(朝倉書店) を
- 連続体仮説に関しては、東大数理科学研究科の古田幹雄先生の解説 `http://www.ms.u-tokyo.ac.jp/~furuta/BernsteinZorn.pdf` を

それぞれ参考にしました。興味のある人は読んでみてください。

2　数ベクトルに対する線型写像

前回の授業で天下りに行列とその演算を定義しましたが、今回はいよいよ「その心」を説明します。一言で言ってしまうと、行列とは「$\mathbb{R}^n$ から $\mathbb{R}^m$ への線型写像」と呼ばれる特別な写像を実現する道具なのです。その意味を、これから確認していきましょう。

2.1　数ベクトルに対する線型写像の定義

$f\colon \mathbb{R}^n \to \mathbb{R}^m$ が線型写像であるとは、

- 任意の $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ に対し、$f(\boldsymbol{u}+\boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$
- 任意の $\boldsymbol{u} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ に対し、$f(\alpha\boldsymbol{u}) = \alpha f(\boldsymbol{u})$

が満たされることをいいます。難しいことを言っているような気もしますが、$m = n = 1$ とおけば f は普通の函数になります。そして u がただの数なので、$f(u) = uf(1)$ となります。つまり f は定数項のない 1 次函数 (正比例の式) です。だから線型写像は、1 **次函数** $y = ax$ **をベクトルに対して一般化したもの**と言えます。こう言えば、線型写像を考えるご利益がありそうだという気がしてくると思います。そして実際、線型写像は色々なところで出てきます。

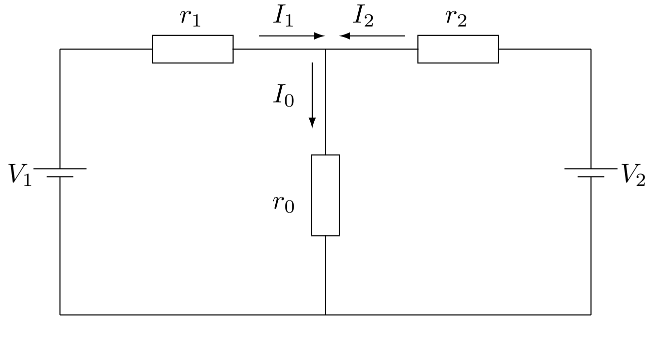
線型写像の一つの例は、行列の積です。A を (m, n) 型行列とします。このとき $\boldsymbol{v} \in \mathbb{R}^n$ に対し、$A\boldsymbol{v} \in \mathbb{R}^m$ です。したがって「A を $\mathbb{R}^n$ のベクトルに左からかける」という操作によって、A は写像 $\mathbb{R}^n \to \mathbb{R}^m$ を定めます。そして

- 任意の $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^n$ に対し、$A(\boldsymbol{u}+\boldsymbol{v}) = A\boldsymbol{u} + A\boldsymbol{v}$
- 任意の $\boldsymbol{u} \in \mathbb{R}^n$, $\alpha \in \mathbb{R}$ に対し、$A(\alpha\boldsymbol{u}) = \alpha A\boldsymbol{u}$

が成り立ちます。よって A 倍写像は線型写像であることが分かりました。

■問 2 の解答　抵抗や起電力の具体的な値は本質的でないので、最も一般的な状況で議論する。図のように[*2]電池の起電力 V_1, V_2、抵抗値 r_0, r_1, r_2 および電流 I_0, I_1, I_2 を定める。このとき

- 上側の分岐点における Kirchhoff の第 1 法則から $I_0 = I_1 + I_2$
- 左右それぞれのサイクルに対する Kirchhoff の第 2 法則から $V_1 = r_1 I_1 + r_0 I_0, V_2 = r_2 I_2 + r_0 I_0$



*2 問題中の図では抵抗がギザギザの折れ線で表されていました。この記号は今でもよく使われてはいますが古いですので、ここでは現行の規格 JIS C 0617-4 に従い、箱で抵抗を表しました。

が成り立つ*3。これらの式から I_0 を消去し、方程式を行列で書くと

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} r_0 + r_1 & r_0 \\ r_0 & r_0 + r_2 \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$$

が成り立つ*4。よって、逆行列を両辺に左からかけると

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \frac{1}{r_1 r_2 + r_0(r_1 + r_2)} \begin{pmatrix} r_0 + r_2 & -r_0 \\ -r_0 & r_0 + r_1 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$$

だから

$$\begin{aligned} I_0 = I_1 + I_2 &= \frac{1}{r_1 r_2 + r_0(r_1 + r_2)} \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} r_0 + r_2 & -r_0 \\ -r_0 & r_0 + r_1 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \frac{1}{r_1 r_2 + r_0(r_1 + r_2)} \begin{pmatrix} r_2 & r_1 \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \\ &= \frac{1}{r_1 r_2 + r_0(r_1 + r_2)} (r_2 V_1 + r_1 V_2) \end{aligned}$$

となる。確かにこれは、$V_1 = 0$ のときの解と $V_2 = 0$ のときの解を足し合わせたものになっている。数値の代入はやればできるので省略する。

■線型性　いまの答えの式を

$$I_0 = \frac{1}{r_1 r_2 + r_0(r_1 + r_2)} \begin{pmatrix} r_2 & r_1 \end{pmatrix} \begin{pmatrix} V_1 \\ 0 \end{pmatrix} + \frac{1}{r_1 r_2 + r_0(r_1 + r_2)} \begin{pmatrix} r_2 & r_1 \end{pmatrix} \begin{pmatrix} 0 \\ V_2 \end{pmatrix}$$

と書き換えてみます。こうすれば「$V_1 = 0$ のときの I_0 の値」「$V_2 = 0$ のときの I_0 の値」を足したものが最終的な I_0 の値になることがよりはっきり見えます。そして、この式の書き換えで使ったのは紛れもなく**行列をかける操作の線型性**です。だからこの問題は「I_0, I_1, I_2 が V_1, V_2 に対して線型に依存すること」を確認していたことになります。

もう少し別の線型写像の例も見てみましょう。

■問 1 の解答　(1) 任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ と実数 α に対し

$$\begin{aligned} \overline{\boldsymbol{x} + \boldsymbol{y}} &= \frac{1}{n} \sum_{i=1}^{n} (x_i + y_i) = \frac{1}{n} \sum_{i=1}^{n} x_i + \frac{1}{n} \sum_{i=1}^{n} y_i = \overline{\boldsymbol{x}} + \overline{\boldsymbol{y}} \\ \overline{\alpha \boldsymbol{x}} &= \frac{1}{n} \sum_{i=1}^{n} \alpha x_i = \alpha \frac{1}{n} \sum_{i=1}^{n} x_i = \alpha \overline{\boldsymbol{x}} \end{aligned}$$

が成り立つ。よって平均を取る写像は線型である。

(2) 元の問題では $n = 5$ であるが、一般の場合で示す。$p_1, \ldots, p_n$ を、$p_1 + \cdots + p_n = 1$ かつ $p_i > 0$ を満たす実数とする。また $\boldsymbol{x} = {}^t(x_1, \ldots, x_n) \in \mathbb{R}^n$ とする。このとき期待値は

$$E[\boldsymbol{x}] = p_1 x_1 + p_2 x_2 + \cdots + p_n x_n = \sum_{i=1}^{n} p_i x_i$$

で与えられる。よって任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ と $\alpha \in \mathbb{R}$ に対して

$$E[\boldsymbol{x} + \boldsymbol{y}] = \sum_{i=1}^{n} p_i (x_i + y_i) = \sum_{i=1}^{n} p_i x_i + \sum_{i=1}^{n} p_i y_i = E[\boldsymbol{x}] + E[\boldsymbol{y}]$$

*3 要は「導線の接点において、流れ込む電流と流れ出る電流は釣り合う」「回路を一周すると、抵抗での電圧低下と電池の起電力が釣り合う」と言っているだけです。物理を習っていなくても、こう書けば「それはそうなるだろう」という気がしてきませんか？ また、式さえ立ててしまえば、後は単なる連立一次方程式の問題です。

*4 I_0 を消去せずに 3 次正方行列を使って方程式を書き下すこともできますが、計算方法を知っていたとしても、3 次正方行列の逆行列を求めるのは若干面倒です。今の場合は簡単に変数 I_0 を消去できるので、2 次正方行列を使った計算に持ち込みました。

$$E[\alpha\boldsymbol{x}] = \sum_{i=1}^{n} p_i \alpha x_i = \alpha \sum_{i=1}^{n} p_i x_i = \alpha E[\boldsymbol{x}]$$

が成り立つ。すなわち期待値を取る写像は線型である。

■問 1 の補足　上の解答では線型性を定義に従って示しました。ですが (2) で $\boldsymbol{p} = {}^t(p_1, \ldots, p_n) \in \mathbb{R}^n$ とおくと、$E[\boldsymbol{x}] = {}^t\boldsymbol{p}\boldsymbol{x}$ という式が成り立ちます。つまり期待値を取る写像 E は横ベクトル ${}^t\boldsymbol{p}$ を左からかける写像と一致するので、ここから線型性が分かります。また (2) で $p_1 = \cdots = p_n := 1/n$ とおけば、(1) の場合になります。

2.2　線型写像の行列表示

行列の積は線型写像でしたが、実は逆に、**全ての線型写像は行列の積で書くことができます。**

いま $f\colon \mathbb{R}^n \to \mathbb{R}^m$ を線型写像とします。このとき $1 \leq i \leq n$ に対し、第 i 成分が 1 でそれ以外の成分が 0 である行列を $\boldsymbol{e}_i$ と書きます。これを用いて $\boldsymbol{v}_i := f(\boldsymbol{e}_i) \in \mathbb{R}^m$ とおき、さらにベクトル $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ たちを並べて $A = (\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots\ \boldsymbol{v}_n)$ という行列を作ります。この行列のサイズは (m, n) 型です。つまり、f による $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ の行先を並べて行列を作るわけです。

そして $\mathbb{R}^n$ のベクトル $\boldsymbol{v}$ を成分表示し $\boldsymbol{v} = {}^t(a_1, a_2, \ldots, a_n) = \sum_{i=1}^{n} a_i \boldsymbol{e}_i$ と書くと

$$f\left(\sum_{i=1}^{n} a_i \boldsymbol{e}_i\right) = \sum_{i=1}^{n} f(a_i \boldsymbol{e}_i) = \sum_{i=1}^{n} a_i f(\boldsymbol{e}_i) = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$$

が成り立ち、その一方で

$$A \sum_{i=1}^{n} a_i \boldsymbol{e}_i = \sum_{i=1}^{n} a_i A \boldsymbol{e}_i = \sum_{i=1}^{n} a_i \boldsymbol{v}_i$$

も成り立ちます。よって $f(\boldsymbol{v}) = A\boldsymbol{v}$ です。$\boldsymbol{v}$ の取り方は任意だったので、これで線型写像 f と A 倍写像が一致することが示せました。この A を**線型写像 f の行列表示**といいます。ナイーブには「線型写像 = 行列」と思ってよい、ということになります。

2.3　平面 $\mathbb{R}^2$ 上の線型変換

線型写像 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ において、特に $n = m$ が満たされる場合、f を**線型変換**と呼ぶことがあります。$m \neq n$ なら $\mathbb{R}^n$ と $\mathbb{R}^m$ は異なるので、f は空間 $\mathbb{R}^n$ の点を別の空間 $\mathbb{R}^m$ の点に対応させるだけです。ところが $m = n$ なら定義域と値域が同じ集合ですから、f は「空間 $\mathbb{R}^n$ の中で点を移動させる操作」と思うことができます。そこで $m = n$ の場合、線型写像 f は特別に「線型変換」と呼ばれます。

■等長写像と等角写像　線型写像や線型変換は、一般に図形の「形」を保ちません。線型性のおかげでそこまで大幅に形が崩れたりはしないのですが、たとえば「長方形をつぶして平行四辺形にする」くらいの変形が起きます。ところが運が良いと、線型写像が「任意の 2 点間の距離を保つ」とか「任意の 2 本のベクトルのなす角を保つ」といった良い性質を持つことがあります。これらの典型例は、回転行列です。

こうした写像はしばしば登場するので、名前が付いています。任意の 2 点間の距離を保つ写像を**等長写像**といい、任意の 2 本のベクトルのなす角を保つ写像を**等角写像**といいます。また等角写像であるような線型変換は、あまり「等角変換」とは呼ばれず、**共形変換**と呼ばれることが多いです。

線型変換の場合、等長変換ならば必ず共形変換になります。これは三角形の合同条件に「三辺相等」があるからです。どの 2 点間の距離も保たれるなら三角形の形がそのまま保たれるので、したがって角度も保たれるというわけです。この事実をベクトルを用いて証明するのが、問 4 の (1)⇒(2) です。ところが共形変換であっても、必ずしも等長

変換になるとは限りません。2 つの三角形の内角が揃っても、その 2 つの三角形は一般に合同にはならず、相似にしかならないからです。ですから線型な共形変換は相似変換ですが、等長性までは保証されません。

■問 4 の訂正　というわけで、問 4 の**問題が半分間違っていました**。ごめんなさい。(1)⇒(2) は正しいですが、(2)⇒(1) は正しくありません。反例はたとえば 2 倍写像 $f(\boldsymbol{u}) := 2\boldsymbol{u}$ です。

ベクトル $\boldsymbol{u}, \boldsymbol{v}$ のなす角は $\boldsymbol{u}\cdot\boldsymbol{v}/|\boldsymbol{u}||\boldsymbol{v}|$ で与えられます。いま f を 2 倍写像とすると、$f(\boldsymbol{u})\cdot f(\boldsymbol{v})/|f(\boldsymbol{u})||f(\boldsymbol{v})| = 2\boldsymbol{u}\cdot 2\boldsymbol{v}/|2\boldsymbol{u}||2\boldsymbol{v}| = \boldsymbol{u}\cdot\boldsymbol{v}/|\boldsymbol{u}||\boldsymbol{v}|$ なので、任意のベクトルのなす角は保存されます。しかし f は長さを保ちません。

■問 4 (1)⇒(2) の解答　任意のベクトル $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ に対し

$$|\boldsymbol{u}+\boldsymbol{v}|^2 - |\boldsymbol{u}-\boldsymbol{v}|^2 = (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v}) - (\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v}) = 4\boldsymbol{u}\cdot\boldsymbol{v}$$

が成り立つ。よって $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ が等長線型変換なら、$|F(\boldsymbol{u})+F(\boldsymbol{v})| = |\boldsymbol{u}+\boldsymbol{v}|$, $|F(\boldsymbol{u})-F(\boldsymbol{v})| = |\boldsymbol{u}-\boldsymbol{v}|$ なので

$$F(\boldsymbol{u})\cdot F(\boldsymbol{v}) = \frac{|F(\boldsymbol{u})+F(\boldsymbol{v})|^2 - |F(\boldsymbol{u})-F(\boldsymbol{v})|^2}{4} = \frac{|\boldsymbol{u}+\boldsymbol{v}|^2 - |\boldsymbol{u}-\boldsymbol{v}|^2}{4} = \boldsymbol{u}\cdot\boldsymbol{v}$$

となる。つまり F は内積を保っている。 そして F は等長写像だから、角度 $\boldsymbol{u}\cdot\boldsymbol{v}/|\boldsymbol{u}||\boldsymbol{v}|$ も保つ。 ■

■問 5 の解答　(1) $SR(\theta) = R(\theta)S$ とすると

$$SR(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ -\sin\theta & -\cos\theta \end{pmatrix}$$
$$R(\theta)S = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

これらが等しいので $\sin\theta = -\sin\theta$、つまり $\sin\theta = 0$ である。よって $\theta = n\pi$ $(n \in \mathbb{Z})$ である。
(2) 次の計算により、$SR(\theta)S$ は $(-\theta)$ 回転であると分かる。

$$SR(\theta)S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = R(-\theta)$$

■問 5 の訂正と線型空間の向き　いま見たように、どう見ても $SR(\theta)S$ は「折り返し」ではなく、回転を表します。これも問題が間違えていました。すいません。

ちなみに「向き」を考えると、$SR(\theta)S$ が折り返しでないという確証が得られます。平面上の 1 次独立な 2 本のベクトル $\boldsymbol{u}, \boldsymbol{v}$ について、始点を揃えたとき $\boldsymbol{u}$ から $\boldsymbol{v}$ へ向かう方向が正の向き (反時計回り) になることを「$(\boldsymbol{u}, \boldsymbol{v})$ は正の向きである」と呼びましょう。たとえば $(\boldsymbol{e}_1, \boldsymbol{e}_2)$ は正の向きです。このとき線型変換 A に対し「正の向きであるベクトルのペアを正の向きにうつす」ということと、$\det A > 0$ が同値になります。たとえば回転行列と折り返しは、それぞれ $\det R(\theta) = 1 > 0$, $\det S = -1 < 0$ を満たします。そして実際、正の向きのベクトルのペアを回転させたものは正の向きだし、折り返したものは負の向きになります。確かに「向きを保つかどうか」と det の符号に対応がついていますね。

一方、行列式は $\det AB = \det A \det B$ という性質を持つのでした。ですから $\det SR(\theta)S = \det S \det R(\theta) \det S = (-1)\cdot 1\cdot(-1) = 1$ となり、$SR(\theta)S$ は向きを保ちます。一方でどんな折り返し変換も向きを反転させますから、その行列式は負でないといけません。したがって、$SR(\theta)S$ は絶対に折り返しにならないと分かりました。

3　線型空間

これまで数ベクトル空間上の線型写像を調べました。ここでもう一歩、線型写像の考察を進めてみましょう。僕たちは数ベクトルの性質を使って線型写像を定義したわけですが、その発想を逆転させてみます。「線型写像が定義できる

には、定義域と値域にどんな性質があれば良いでしょうか？」[*5]

その答えが「線型空間」と呼ばれるものです。線型空間は数ベクトル空間の性質を一部だけ抜き出したもので、これを使うと、数ベクトル空間でなくても線型写像の概念を定義することができます。

3.1 線型空間の定義

後でちゃんと線型空間の定義を理解する必要がありますが、今は軽く眺め「数ベクトル空間だったら当たり前に成り立つ性質だよなあ」と納得してみてください。

■線型空間 次の条件を満たす集合 V を**線型空間**あるいは**ベクトル空間**と言います。

- V 上に加法が定義されており
  - 任意の $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ に対し $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} = \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ が成り立つ
  - 任意の $\boldsymbol{u} \in V$ に対し $\boldsymbol{u}+\boldsymbol{0} = \boldsymbol{0}+\boldsymbol{u} = \boldsymbol{0}$ となる $\boldsymbol{0} \in V$ が唯一存在する
  - 任意の $\boldsymbol{u} \in V$ に対し、$\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{0}$ となる $\boldsymbol{v} \in V$ が唯一存在する (この $\boldsymbol{v}$ を $-\boldsymbol{u}$ と書く)
  - 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ に対し $\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{u}$ が成り立つ
- V 上にスカラー倍が定義されている
  - 任意の $\alpha, \beta \in \mathbb{R}$ と $\boldsymbol{v} \in V$ に対し $\alpha(\beta\boldsymbol{v}) = (\alpha\beta)\boldsymbol{v}$ が成り立つ
  - 任意の $\boldsymbol{v} \in V$ に対し、$1\boldsymbol{v} = \boldsymbol{v}$ が成り立つ
- スカラー倍と和が分配法則を満たす
  - 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ と $\alpha \in \mathbb{R}$ に対し $(\alpha+\beta)\boldsymbol{v} = \alpha\boldsymbol{v}+\beta\boldsymbol{v}$
  - 任意の $\boldsymbol{u} \in V$ と $\alpha, \beta \in \mathbb{R}$ に対し、$\alpha(\boldsymbol{u}+\boldsymbol{v}) = \alpha\boldsymbol{u}+\alpha\boldsymbol{v}$

これらの性質から、$0\boldsymbol{v} = \boldsymbol{0}$, $(-1)\boldsymbol{v} = -\boldsymbol{v}$ が成り立つことが次のように示せます。

- $0\boldsymbol{v} = (0+0)\boldsymbol{v} = 0\boldsymbol{v}+0\boldsymbol{v}$ の両辺に $-(0\boldsymbol{v})$ を足して、$\boldsymbol{0} = 0\boldsymbol{v}+\boldsymbol{0} = 0\boldsymbol{v}$
- $\boldsymbol{0} = 0\boldsymbol{v} = \big(1+(-1)\big)\boldsymbol{v} = \boldsymbol{v}+(-1)\boldsymbol{v}$ より、$-\boldsymbol{v} = (-1)\boldsymbol{v}$

ちなみにスカラーが実数体 $\mathbb{R}$ のとき、実線型空間という言い方をします。スカラーを複素数体 $\mathbb{C}$ にすると、今と全く同じようにして複素線型空間が定義されます。

3.2 例: 微分方程式の解空間

線型空間の定義を抽象化したことによって、我々は「数ベクトル空間でない線型空間」を扱えるようになりました。その中で最も重要なものが、微分方程式の解空間と呼ばれるものです。どのような分野に進んでも、おそらく微分方程式と全く無縁な人生を過ごす人はいないでしょう。たとえば「時々刻々と変化する量」を扱えば、確実に時間を変数とする微分方程式が出ます。そして単振動の方程式や波動方程式、熱方程式などの基本的な微分方程式の多くが、これから説明する「線型微分方程式」と呼ばれるクラスに属します。また解くのが難しい微分方程式も、線型微分方程式で近似して解くことが良くあります。こうした線型微分方程式を解くとき、線型代数が極めて重要な役割を果たします。

まずは例を見てみましょう。

■問 3 の解答 微分方程式 $y''-y'-6y = 0$ を考える。
(1), (2) 一般に $y = e^{\alpha x}$ に対し、$y' = \alpha y$ を満たす。よって $y''-y'-6y = (\alpha^2-\alpha-6)y = (\alpha-3)(\alpha+2)y$ である。

[*5] ちなみに、この手の発想は数学ではよく見かけます。「○○は××の性質を持つ」ということが分かったら、逆に「××という性質が上手く定まるのはどういう場合か」と考え、抽象化を図るのです。たとえば函数の「連続性」の概念から位相空間の考え方に到達します。

これより $y=e^{\alpha x}$ が $y''-y'-6y=0$ を満たすことは、$\alpha=-2$ または $\alpha=3$ と同値である。
(3) y を微分方程式の解、$C\in\mathbb{R}$ とする。このとき $(Cy)'=Cy'$, $(Cy)''=Cy''$ であるから、$(Cy)''-(Cy)'-6(Cy)=C(y''-y'-6)=0$ である。よって Cy も解である。

また y_1,y_2 を微分方程式の解とする。このとき $(y_1+y_2)''-(y_1+y_2)'-6(y_1+y_2)=(y_1''-y_1'-6y_1)+(y_2''-y_2'-6y_2)=0$ であるから、y_1+y_2 もまた微分方程式の解である。
(4) は (1), (2), (3) から直ちに従う。
(5) 微分方程式の解 $y(x)=C_1e^{3x}+C_2e^{-2x}$ が初期条件 $y(0)=\alpha$, $y'(0)=\beta$ を満たすとする。このとき $\alpha=y(0)=C_1+C_2$, $\beta=y'(0)=3C_1-2C_2$ なので、$C_1=(2\alpha+\beta)/5, C_2=(3\alpha-\beta)/5$ である。

■線型微分方程式の解空間　今の問題の (3) で示したことを一言で言えば「$y''-y'-6y=0$ の解全体は線型空間をなす」ということです。この証明で本質的なのは「微分の線型性」と「$y,y',\ldots$ の項が斉次 1 次式[*6]」という点でした。

一般に $a_n(x)y^{(n)}+a_{n-1}(x)y^{(n-1)}+\cdots+a_1(x)y'+a_0(x)y=0$ のように、微分方程式であって $y,y',y'',\ldots$ に関する斉次 1 次式であるものを**線型微分方程式**といいます。問 3 と全く同様にして、この方程式の解全体は線型空間をなすことが示せます。そこで線型微分方程式の場合、解全体の集合を**解空間**といいます。

線型微分方程式の場合、2 つの解を足したら新しい解を作ることができます。すなわち**解の重ね合わせ**ができます。たとえば「波」を表す函数は波動方程式という線型偏微分方程式の解として与えられます。波については重ね合わせの原理が成り立つことはよく知られていますが、微分方程式のレベルで見ると、これは波の満たす方程式が線型だという事実に対応しています。

■線型空間の基底　線型空間 V のベクトルの組 $(\boldsymbol{u}_1,\ldots,\boldsymbol{u}_n)$ であって「任意の $\boldsymbol{v}\in V$ が $\boldsymbol{v}=a_1\boldsymbol{v}_1+a_2\boldsymbol{v}_2+\cdots+a_n\boldsymbol{u}_n$ の形に一意的に表せる」という性質を持つものを、V の**基底**と言います。たとえば第 i 成分だけが 1 で他の成分が 0 であるようなベクトルを $\boldsymbol{e}_i$ と書くと、$\mathbb{R}^n$ の基底として $(\boldsymbol{e}_1,\boldsymbol{e}_2,\ldots,\boldsymbol{e}_n)$ が取れます。

実は一般の場合でも、線型空間には常に基底が取れることが知られています。ですから基底を 1 つ取ってしまえば、どんなベクトルもその 1 次結合で表せます。特に線型微分方程式の解空間の場合、**基底をなす解を見つけてしまえば、他の全ての解は基底の重ね合わせで得られる**ということです。今の方程式 $y''-y'-6y=0$ では、(e^{3x},e^{-2x}) が解空間の基底であることが示せます。だから線型代数の議論によって、$C_1e^{3x}+C_2e^{-2x}$ $(C_1,C_2\in\mathbb{R})$ が**全ての解を尽くすこと**が保証されるのです。

また方程式 $y''-y'-6y=0$ については、初期条件として $y(0)$ と $y'(0)$ を決めるごとに解が 1 つ定まります。ですから解の自由度は 2 です。そして (5) の問題で見たように、初期条件 $y(0)=\alpha$, $y'(0)=\beta$ に対応する解を求めることは、${}^t(\alpha,\beta)=C_1{}^t(1,3)+C_2{}^t(1,-2)$ を満たすスカラー C_1,C_2 を求めることと同じでした。つまり初期条件を満たす解、e^{3x} と e^{-2x} とが、それぞれベクトル ${}^t(\alpha,\beta)$, ${}^t(1,3)$ と ${}^t(1,2)$ とにぴったり対応しているわけです[*7]。そして「解の自由度が 2 である」という事実は、解空間の次元が 2 であるという事実にちょうど対応しているのです。「自由度」という言葉は必ずしも厳密な定式化が易しくありませんが、線型微分方程式の場合はこのように「解空間」という概念を導入することで、その次元としてきちんと定式化ができるのです。このようにして、線型代数の議論がいかに強力であるかを垣間見ることができます。

[*6] ここで「斉次」というのは、定数項が 0 という意味です。
[*7] 線型写像のことをもう少しちゃんと学ぶと、この「微分方程式の解と数ベクトルとが対応する」という事実は「解空間と $\mathbb{R}^2$ の間の線型同型」という言葉できちんと定式化できます。

S2 ターム

線形代数学①

線形代数学① 第 7 回　　ベクトル空間と線型写像

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 6 月 10 日 1 限

1　はじめに

「数理科学基礎」の授業が終わり、いよいよ線型代数学[*1]の授業が始まりました。既に行列の計算法を学び、線型空間や線型写像出てはきましたが、この先は一段と重たい計算や抽象的な議論が出てきます。心してかかってください。

抽象的な議論をする理由は、ひとえに「適用範囲が広がるから」です。普通に行列を計算しているだけでもそれなりに得られるものはあるのですが、線型空間と線型写像の言葉によって、行列計算のテクニックを他の様々な場面に応用することができます。そしてこれから見ていくように「線型な場面」というのは多々あるので、線型代数を理解しているとこの先の人生で非常に役立ちます。

一方で「**線型空間と線型写像は、それぞれ数ベクトル空間と行列がモデルになっている**」という事実を忘れてはいけません。抽象論ばかりやっていると時々自分が何をしているのか分からなくなってきますが、そういう時は必ず「行列だったらどうだろう」と考えてみましょう。特に、最も手頃な 2 次正方行列で試してみると良いと思います。

■今回の内容について　今回配付するプリントの分量は、いつもの 2 倍くらいの厚さになってしまいました。おそらく**今回は、1 年間の授業全体を通して一番難易度の勾配が急**なところだと思います。最終的には全て理解することが望ましいですが、一度に理解するのは厳しいかもしれません。そういうときは、次のようにしてみてください。

- 線型空間と線型写像の定義は、$\mathbb{R}^2$ の部分空間や 1 次函数の例などを使い、確実に理解してください。これが分かっていないとにっちもさっちもいきません。
- 部分空間の定義も、分からないと後々困ります。また最低限、連立一次方程式の解空間が部分空間になっていることは理解してください。
- 多項式の空間、数列の空間と函数空間は「数ベクトル空間ではない線型空間の例」として典型的なものです。そういう意味でこれらの問題は重要なのですが、数ベクトル空間のことをよく理解していないと、勉強するのは大変だと思います。今取り組むのが大変だと思ったら、後で戻ってきてください。
- 多項式の空間は、数列の空間と函数空間に比べて扱いやすいと思います。また数列の空間と函数空間のどちらが分かりやすいかについては、人によって個人差がありそうです。好きな方から取り組みましょう。

また毎週のレポート問題に取り組むにあたっては、**自分が理解できる問題を着実に解くこと**が大事です。特にこれから先、色々な命題を**定義に基づいてきっちり示す**ことが問われます。ですから生半可な気持ちで雑に解いたり、まして他人の答案を (間違ったまま) 書き写したりすることは全く無意味です[*2]。それよりは「何が分からなかったか」をはっきりさせる方が大事でしょう。こういう点を意識して、きちんとレポートに取り組んでください。

2　線型空間と線型写像

これからの主役は言うまでもなく線型空間と線型写像です。まずそれらの定義を確認し、それから「線型空間の中に入っている、より小さな線型空間」に関する考察をしましょう。

[*1] すごく今更ですが、「線型代数」と「線形代数」は同じ意味で、どっちの漢字を使うかは気分の問題でしかありません。授業名は「線形代数学」になってますが、このプリントでは「線型代数」を使います。

[*2] 誰とは言いませんが、たとえばベクトル空間のことを「ベクトル場」と言い間違えている答案が 5 枚以上ありました。こういうミスがあると、採点してる側には「何も考えずに誰かのを書き写している」ということが一発で見えてしまいます。

2.1 線型空間の定義と例

■線型空間の定義 (再掲) 次の条件を満たす集合 V を**線型空間**あるいは**ベクトル空間**というのでした。

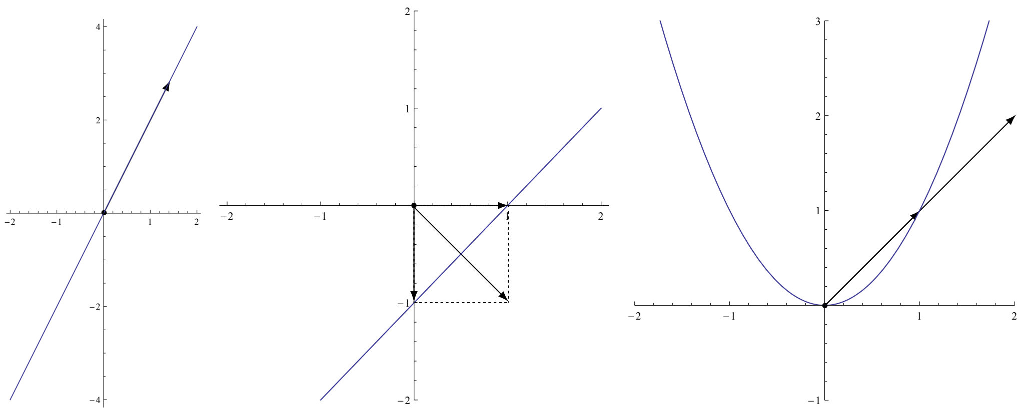
- V 上に加法が定義されており
  - 任意の $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in V$ に対し $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w} = \boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ が成り立つ*3
  - 任意の $\boldsymbol{u} \in V$ に対し $\boldsymbol{u}+\boldsymbol{0} = \boldsymbol{0}+\boldsymbol{u} = \boldsymbol{0}$ となる $\boldsymbol{0} \in V$ が唯一存在する
  - 任意の $\boldsymbol{u} \in V$ に対し、$\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{0}$ となる $\boldsymbol{v} \in V$ が唯一存在する (この $\boldsymbol{v}$ を $-\boldsymbol{u}$ と書く)
  - 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ に対し $\boldsymbol{u}+\boldsymbol{v} = \boldsymbol{v}+\boldsymbol{u}$ が成り立つ
- V 上にスカラー倍が定義されており
  - 任意の $\alpha, \beta \in \mathbb{R}$ と $\boldsymbol{v} \in V$ に対し $\alpha(\beta\boldsymbol{v}) = (\alpha\beta)\boldsymbol{v}$ が成り立つ
  - 任意の $\boldsymbol{v} \in V$ に対し、$1\boldsymbol{v} = \boldsymbol{v}$ が成り立つ
- スカラー倍と和が分配法則を満たす:
  - 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ と $\alpha \in \mathbb{R}$ に対し $(\alpha+\beta)\boldsymbol{v} = \alpha\boldsymbol{v}+\beta\boldsymbol{v}$
  - 任意の $\boldsymbol{u} \in V$ と $\alpha, \beta \in \mathbb{R}$ に対し、$\alpha(\boldsymbol{u}+\boldsymbol{v}) = \alpha\boldsymbol{u}+\alpha\boldsymbol{v}$

これらの性質から、$0\boldsymbol{v} = \boldsymbol{0}$, $(-1)\boldsymbol{v} = -\boldsymbol{v}$ が成り立つことに注意しておきます:

- $0\boldsymbol{v} = (0+0)\boldsymbol{v} = 0\boldsymbol{v}+0\boldsymbol{v}$ の両辺に $-(0\boldsymbol{v})$ を足して、$\boldsymbol{0} = 0\boldsymbol{v}+\boldsymbol{0} = 0\boldsymbol{v}$ を得る。
- $\boldsymbol{0} = 0\boldsymbol{v} = \big(1+(-1)\big)\boldsymbol{v} = \boldsymbol{v}+(-1)\boldsymbol{v}$ より、$-\boldsymbol{v} = (-1)\boldsymbol{v}$ を得る。

■$\mathbb{R}^2$ の部分空間の例 ここで、線型空間の何たるかを目で見て理解するため「平面 $\mathbb{R}^2$ の部分集合で、線型空間になっているものは何か」を考えてみましょう。次に挙げる 3 つの部分集合のうち、どれが線型空間でしょうか？*4

$$V_1 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = 2x \right\}, \qquad V_2 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = x-1 \right\}, \qquad V_3 := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid y = x^2 \right\}$$



(m) $y = 2x$ (n) $y = x-1$ (o) $y = x^2$

まず V_3 の放物線 $y = x^2$ は、見るからに線型じゃなさそうな格好をしています。実際、$\boldsymbol{v} := {}^t(1,1)$ は放物線の上に乗りますが、2 倍した $2\boldsymbol{v} = {}^t(2,2)$ は放物線からはみ出します。よってこれは線型空間の定義を満たしません。

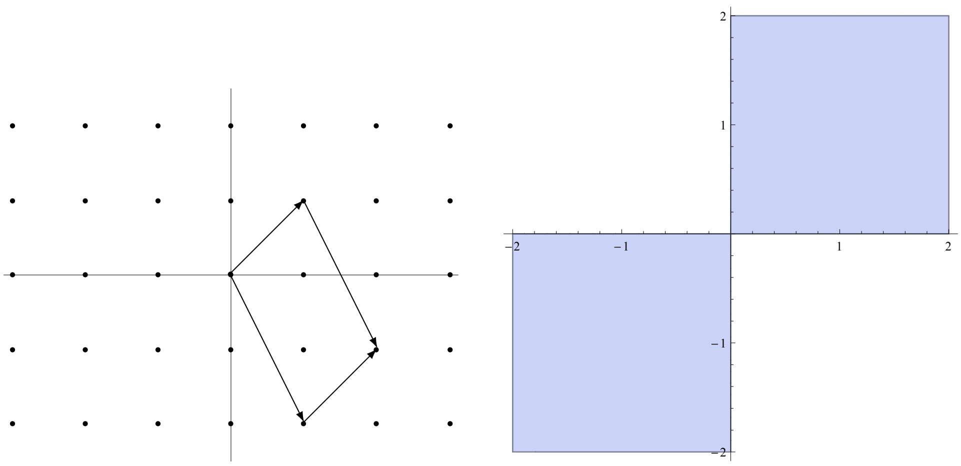
*3 このように「演算の結果が順番に依存しない」ということを指して、**結合法則**が成り立つといいます。
*4 すぐ後ろで議論はしますが、それを読む前に自分で定義に従って、どれが線型空間か判定してみると良いと思います。

次に V_1 の直線 $y = 2x$ は、いかにも線型と呼ぶにふさわしい図形です。そして実際、ちゃんと線型空間になります。この直線上の点は ${}^t(x, 2x)$ という格好で書けます。そこで直線上のベクトル ${}^t(a, 2a)$ と ${}^t(b, 2b)$ を取ってくると、これらの和 ${}^t\bigl(a + b, 2(a + b)\bigr)$ も再び直線 $y = 2x$ に乗ります。またスカラー倍も $\alpha^t(a, 2a) = (a\alpha, 2a\alpha)$ となり、ちゃんと直線 $y = 2x$ 上に乗ります。絵で描いてみると、どう見ても加法とスカラー倍で閉じていることが一段と良く分かりますね。確かに $y = 2x$ は線型空間でした。

ところが線型空間になるためには、単に**まっすぐならば良いというわけでもない**のです。実は直線 $y = x - 1$ は線型空間になっていません。たとえばベクトル ${}^t(1, 0)$, ${}^t(0, -1)$ はこの直線上に乗りますが、これらを足した ${}^t(1, -1)$ は直線 $y = x - 1$ からはみ出してしまいます。こうして $y = x - 1$ は線型空間でないことが分かりました。

以上をまとめると、線型空間は**原点を持つ、まっすぐな空間**ということになります。

ちなみに線型空間の定義は、大雑把には「ベクトルの足し算できること」「ベクトルのスカラー倍ができること」に分かれます。きちんと線型空間を定義するには、このどちらの条件も必要です。たとえば成分が全て整数であるベクトルの集合 $\mathbb{Z}^2 := \left\{{}^t(m, n) \in \mathbb{R}^2 \mid m, n \in \mathbb{Z}\right\}$ は、加法で閉じますがスカラー倍で閉じません。また x, y 軸と第 $1, 3$ 象限を合わせた $\left\{{}^t(x, y) \in \mathbb{R}^2 \mid xy \geq 0\right\}$ という集合は、スカラー倍で閉じますが加法では閉じません。



(p) $\mathbb{Z}^2$　　(q) $xy \geq 0$

2.2　線型写像

■線型写像の定義　V, W を共に線型空間とするとき、$f\colon V \to W$ が線型写像であるとは

- 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ に対し、$f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$
- 任意の $\boldsymbol{u} \in V$, $\alpha \in \mathbb{R}$ に対し、$f(\alpha\boldsymbol{u}) = \alpha f(\boldsymbol{u})$

が成り立つことでした。

この定義が、何となく「線型空間と相性が良い」ことを感じて欲しいのですが、どうでしょうか？ V と W は線型空間だから、どちらに対しても足し算とスカラー倍が定義されています。そして f が線型写像であることの条件は、**f を施してから足し算 / スカラー倍をしても、足し算 / スカラー倍をしてから f を施しても、結果が同じになる**ということを言っています[*5]。図式にすると次のようになります。これを見て、線型写像と線型空間の相性の良さを感じ取って

[*5] このことを称して「f は加法 / スカラー倍と整合的 (compatible) である」と言ったりもします。線型空間の構造と整合的な写像が線型写像、というわけです。

ください。

$$
\begin{tikzcd}
(\boldsymbol{v}, \boldsymbol{w}) \arrow[r, mapsto, "f\times f"] \arrow[d, mapsto, "+"] & \bigl(f(\boldsymbol{v}), f(\boldsymbol{w})\bigr) \arrow[d, mapsto, "+"] \\
\boldsymbol{v}+\boldsymbol{w} \arrow[r, mapsto, "f"] & f(\boldsymbol{v}+\boldsymbol{w}) = f(\boldsymbol{v}) + f(\boldsymbol{w})
\end{tikzcd}
\qquad
\begin{tikzcd}
(\alpha, \boldsymbol{v}) \arrow[r, mapsto, "f\times f"] \arrow[d, mapsto, "\cdot"] & \bigl(\alpha, f(\boldsymbol{w})\bigr) \arrow[d, mapsto, "\cdot"] \\
\alpha\boldsymbol{v} \arrow[r, mapsto, "f"] & f(\alpha\boldsymbol{v}) = \alpha f(\boldsymbol{v})
\end{tikzcd}
$$

また f が線型写像なら, $f(\mathbf{0}) = f(\mathbf{0}-\mathbf{0}) = f(\mathbf{0}) - f(\mathbf{0}) = \mathbf{0}$ が成り立つ[*6]ことにも注意しておきましょう．**線型写像は原点を原点にうつします。**

さて、前回のプリントで「行列の掛け算は線型写像である」ということを説明しました。その中で最も単純な場合、つまり f が $\mathbb{R}$ から $\mathbb{R}$ への線型写像である場合を考えましょう。前回のプリントでも少しだけ触れましたが、この場合 f は定数項が 0 であるような 1 次函数です。裏返して言えば、正比例の関係式をベクトルに一般化したものが線型写像となります。このことをチェックしましょう。

■問 1 の解答　$f(x) := ax + b$ で定まる写像 $f\colon \mathbb{R} \to \mathbb{R}$ が線型写像であるとする。このとき $f(x+y) = f(x) + f(y)$ が成り立つので、$0 = f(x+y) - f(x) - f(y) = \{a(x+y)+b\} - (ax+b) - (ay-b) = b$ となる。よって $b = 0$ でないといけない。

逆に $b = 0$ のとき、$f(x) = ax$ は線型である。実際、任意の $x, y \in \mathbb{R}$ と $\alpha \in \mathbb{R}$ に対し $f(x+y) = a(x+y) = ax + ay = f(x) + f(y)$, $f(\alpha x) = a(\alpha x) = \alpha(ax) = \alpha f(x)$ となっている。よって f が線型であることと $b = 0$ とが同値になる。 ■

2.3 部分空間

一般に、線型空間はその中にもっと小さい線型空間を含んでいます。一番分かりやすい例は、平面 $\mathbb{R}^2$ の中に含まれる x 軸や y 軸です。これらは共に、まっすぐで原点を通っていますね。また先ほど確認したように、$\mathbb{R}^2$ の原点を通る直線はいずれも線型空間です。そこで一般に、線型空間 V の部分集合 $W \subset V$ でそれ自身が線型空間になっているものを、V の**部分空間**といいます[*7]。

座標軸があるおかげで、僕たちは平面上で「x 方向」とか「y 方向」とかを考えることができます。それと同じように部分空間を使うと、親玉の線型空間を「こっちの空間の方向とあっちの空間の方向」というように、色々な方向に分けることができます。これから先、連立方程式や微分方程式の解空間を線型代数の手法で解析していくわけですが、その時に上手い部分空間を考えると問題を切り分けることができます。部分空間はそういう風に役に立つのです。

さて部分空間は色々なところに登場しますが、部分空間が本当に線型空間になっていることを一々定義に戻ってチェックするのは、無駄があります。実は W が線型空間 V の部分空間であることは

- 任意の $\boldsymbol{u}, \boldsymbol{v} \in W$ に対し、$\boldsymbol{u} + \boldsymbol{v} \in W$ が成り立つ
- 任意の $\boldsymbol{u} \in W$ と $\alpha \in \mathbb{R}$ に対し、$\alpha\boldsymbol{u} \in W$ が成り立つ

という 2 条件[*8]のチェックだけで事足りるのです。そのことを、問 3 で確かめましょう。そうすれば今後は、部分空間の確認で楽をすることができます。

■問 3 を解く前に　問題の解答に入る前に「そもそも、この問 3 は何がしたいのか？」を確認しておきましょう。線型空間の公理は 58 ページに書いた通りです。一方、問題では線型空間の部分集合が「加法とスカラー倍とで閉じること」

*6 雑な書き方をしましたが, f の中に入る $\mathbf{0}$ は V の元で, $f(\mathbf{0}) = \mathbf{0}$ の右辺は W の元です．厳密に言えば $\mathbf{0}_V$, $\mathbf{0}_W$ などと書いて区別するべきですが, 文脈から意味がはっきり取れる場合は略記して差し支えありません.

*7 部分線型空間 とか 線型部分空間 とか呼んだりもします。どれでも意味は一緒です。

*8 ちなみに 1 つ目の条件を「W が**加法で閉じている**」といい、2 つ目の条件を「W が**スカラー倍で閉じている**」と言います。「閉じている ＝ はみ出さない」という感じです。

だけを部分空間の定義としています。そして表面的には「部分空間の定義」の方が「線型空間の定義」より条件が少ないように見えますが、実はこれで十分なのだということが、問題になっているのです。ここをまず押さえてください。

そして、問題の意味を理解した上で「どうせ線型空間の部分集合なんだから条件は当然成り立つでしょ」と思った人がかなり多くいました。確かに「3 つのベクトル $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ に対して $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ が成り立つこと」などは、線型空間の部分集合であるから当たり前のように成り立ちます。そういう感じで大体のことは済むのですが「$\mathbf{0}$ ベクトルが部分空間に入ること」「逆向きのベクトルが部分空間に入ること」だけは自明ではありません。ここにどう部分空間の条件を使うかがポイントです。

細かいところですが、これらに注意して解答を読んでください。

■問 3 の解答　V を線型空間とし、その空でない部分集合 $W \subset V$ が V の加法とスカラー倍で閉じているとする。このとき W が線型空間の条件を満たすことを、定義に従ってチェックする[*9]。

- 加法について: 加法そのものは、V で定義されている。そして
  - 任意の $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in W$ に対し $(\boldsymbol{u}+\boldsymbol{v})+\boldsymbol{w}=\boldsymbol{u}+(\boldsymbol{v}+\boldsymbol{w})$ が成り立つことは、V が線型空間であることから保証される。
  - W は空集合ではないので、何か 1 つ元 $\boldsymbol{v} \in W$ が取れる。すると W がスカラー倍で閉じていることから、$\mathbf{0}=0\boldsymbol{v} \in W$ となる。
  - W がスカラー倍で閉じているので、$\boldsymbol{v} \in W$ のとき $-\boldsymbol{v}=(-1)\boldsymbol{v} \in W$ となる。また $\boldsymbol{v}+\boldsymbol{u}=\mathbf{0}$ となる $\boldsymbol{u}$ がただ一つ存在することは、V が線型空間であることから保証される。
  - 任意の $\boldsymbol{u}, \boldsymbol{v} \in W$ に対し $\boldsymbol{u}+\boldsymbol{v}=\boldsymbol{v}+\boldsymbol{u}$ が成り立つことは、V が線型空間であることから従う。
- スカラー倍について: スカラー倍そのものは、V の中で定義されている。そして
  - 任意の $\boldsymbol{v} \in W$ と $\alpha, \beta \in \mathbb{R}$ について、$\alpha(\beta\boldsymbol{v})=(\alpha\beta)\boldsymbol{v}$ が成り立つことは、V が線型空間であることから保証される。
  - 任意の $\boldsymbol{v} \in W$ について $1\boldsymbol{v}=\boldsymbol{v}$ が成り立つことは、V が線型空間であることから従う。
- 分配法則が成り立つことは、V が線型空間であることから保証される。

■

3　重要な例

線型空間と部分空間の定義が終わったので、具体的な例を確かめてみましょう。

3.1　連立一次方程式の解空間

連立 1 次方程式

$$
\begin{aligned}
a_{11}x_1+a_{12}x_2+a_{13}x_3&=b_1\\
a_{21}x_1+a_{22}x_2+a_{23}x_3&=b_2
\end{aligned}
$$

は、

$$
A:=\begin{pmatrix} a_{11} & a_{12} & a_{13}\\ a_{21} & a_{22} & a_{23}\end{pmatrix},\quad \boldsymbol{x}:=\begin{pmatrix} x_1\\ x_2\\ x_3\end{pmatrix},\quad \boldsymbol{b}:=\begin{pmatrix} b_1\\ b_2\end{pmatrix}
$$

[*9] ぶっちゃけた話、このチェックはかなりかったるいのですが、人生で一度はこのチェックをやっておかないといけません。最初のうちは大変だと思いますが、時間をかけてコツコツやってみてください。大体の場合はルーチンワークなので、何回かやれば慣れてきて、テキパキできるようになります。

とおくと、単に $A\boldsymbol{x}=\boldsymbol{b}$ と書けます。特に $\boldsymbol{b}=\boldsymbol{0}$ のとき、この連立方程式は**同次系**であるといい、その解全体の集合を**解空間**といいます。

この解全体の集合を「解空間」と呼ぶことは、次のように考えれば納得がいくのではないでしょうか。一般に空間 $\mathbb{R}^3$ 内で、1 次方程式は平面を定めるのでした。そして方程式を連立することは、図形の側で見れば「交わりを見ること」に相当しました。ですから 3 変数の 1 次方程式は大体の場合、1 本で平面を、2 本で直線を、そして 3 本で点を定めます[*10]。数学では直線とか平面とかも一緒くたに「空間」と呼んでしまう習わしがあるので、解集合のことを解空間と呼びます。また 1 次方程式の定数項が 0 でないと、それの定める平面は原点を通りません。線型空間は原点を持つまっすぐな図形でしたから、連立方程式の定める図形が線型空間になるためには、定数項が全て 0 でなければいけません[*11]。

事情は n 次元でも同じです。成分が実数の (m,n) 型行列全体の集合を $\mathrm{Mat}_{m,n}(\mathbb{R})$ で表します。$A\in\mathrm{Mat}_{m,n}(\mathbb{R})$ として方程式 $A\boldsymbol{x}=\boldsymbol{b}$ を立てると、各成分毎に $(n-1)$ 次元の超平面[*12]が定まります。そして連立する方程式が増えていくたびに超平面との交わりが取られ、解空間の次元がどんどん減っていくという仕組みになっています。

この解空間がちゃんと線型空間になっていることは、図形的な考察からすればほぼ明らかですが、もう一度定義通りに確かめてみましょう。

■問 4 の解答　$A\in\mathrm{Mat}_{m,n}(\mathbb{R})$ とし、$V:=\{\boldsymbol{x}\in\mathbb{R}^n\mid A\boldsymbol{x}=\boldsymbol{0}\}$ とおく。このとき $\boldsymbol{0}\in V$ より V は空でない。そして

(1) $\boldsymbol{u},\boldsymbol{v}\in V$ を任意に取る。このとき $A\boldsymbol{u}=A\boldsymbol{v}=\boldsymbol{0}$ なので、$A(\boldsymbol{u}+\boldsymbol{v})=A\boldsymbol{u}+A\boldsymbol{v}=\boldsymbol{0}+\boldsymbol{0}=\boldsymbol{0}$ である。よって $\boldsymbol{u}+\boldsymbol{v}\in V$ となる。

(2) $\alpha\in\mathbb{R},\ \boldsymbol{u}\in V$ を任意に取る。このとき $A(\alpha\boldsymbol{v})=\alpha(A\boldsymbol{v})=\boldsymbol{0}$ なので、$\alpha\boldsymbol{u}\in V$ である。

これで V が $\mathbb{R}^n$ の部分空間であることがチェックできた。 ■

3.2　多項式の空間

次に、僕たちが普段言う「空間」ではないような線型空間の典型的な例として、多項式の集合を考えてみましょう。$\mathbb{R}[x]$ で、x を変数とする実係数 1 変数多項式全体の集合を表します。これは線型空間になります。実際

- 多項式には加法が定義されていて
  - 任意の多項式 $f(x),g(x),h(x)\in\mathbb{R}[x]$ に対して、$\bigl(f(x)+g(x)\bigr)+h(x)=f(x)+\bigl(g(x)+h(x)\bigr)$ が成り立つ。
  - 0 は多項式で、任意の多項式 $f(x)$ に対して $f(x)+0=0+f(x)=0$ を満たす唯一のものである
  - 任意の多項式 $f(x)$ に対して、$f(x)+g(x)=0$ となる多項式は $g(x)=-f(x)$ に限る
- 多項式には実数倍が定義されていて
  - 多項式の掛け算はどのような順序で行ってもよい
  - 多項式の 1 倍は元の多項式と同じである
- 多項式の積は分配法則を満たすので、特に実数倍についても分配法則が成り立つ。

というように、線型空間の条件を満たしているからです[*13]。

[*10] たまに「3 枚の平面がの交わりが 1 直線になる」という場合が起きることがあります。これは裏返せば、連立している 3 本方程式に無駄があるという事実になります。次回以降きちんと定義しますが「連立している方程式にどれだけ無駄があるか」を測る量を行列の**階数** (rank) といいます。方程式に無駄があるとその分解空間の次元が上がりますから、rank の増減と解空間の次元の増減がぴったり対応します。

[*11] 「別に原点を通らなくたって、線型空間と大体同じじゃないか。それだけでのけ者にするのは可哀想だ」という意見は、全くその通りです。連立 1 次方程式 $A\boldsymbol{x}=\boldsymbol{b}$ の解空間は、$A\boldsymbol{x}=\boldsymbol{0}$ の解空間を平行移動したものに過ぎません。こういう空間を**アフィン空間**と呼んだりします。

[*12] 1 次方程式は平面 $\mathbb{R}^2$ 内では 1 次元の直線を、空間 $\mathbb{R}^3$ 内では 2 次元の平面を定めます。こんな感じに 1 次方程式は n 次元空間 $\mathbb{R}^n$ の中で $(n-1)$ 次元の図形を定めます。このとき「全体の空間に対して次元がどれくらい足りないか」を**余次元**といいます。そして n 次元空間 $\mathbb{R}^n$ の中で余次元が 1 のまっすぐな空間を、**超平面**といいます。

[*13] 実際には、多項式全体の集合 $\mathbb{R}[x]$ は**環**あるいは $\mathbb{R}$ **代数**と呼ばれるものになっており、線型空間の条件より大分強い条件を満たしていることが知られています。

■問 5 の解答　$V := \{f(x) \in \mathbb{R}[x] \mid \deg f(x) \leq 3\}$ とする。
(1) V は線型空間 $\mathbb{R}[x]$ の部分集合である。そして

- 3 次以下の任意の多項式 $f(x), g(x) \in V$ に対し、$f(x) + g(x)$ は再び 3 次以下になる
- 3 次以下の任意の多項式 $f(x) \in V$ と実数 $\alpha \in \mathbb{R}$ に対し、$\alpha f(x)$ は再び 3 次以下になる

ことから、V は $\mathbb{R}[x]$ の部分空間だと分かる。
(2) 多項式 $f(x) \in \mathbb{R}[x]$ に対し $\mathrm{ev}_\alpha\bigl(f(x)\bigr) := f(\alpha)$ と定める[*14]ことで、多項式に実数を対応させる写像 $\mathrm{ev}_\alpha \colon \mathbb{R}[x] \to \mathbb{R}$ を定める。このとき

- 任意の多項式 $f(x), g(x) \in \mathbb{R}[x]$ に対し $\mathrm{ev}_\alpha\bigl(f(x) + g(x)\bigr) = f(\alpha) + g(\alpha) = \mathrm{ev}_\alpha\bigl(f(x)\bigr) + \mathrm{ev}_\alpha\bigl(g(x)\bigr)$
- 任意の多項式 $f(x) \in \mathbb{R}[x]$ と実数 $\alpha \in \mathbb{R}$ に対し、$\mathrm{ev}_\alpha\bigl(af(x)\bigr) = af(\alpha) = a\,\mathrm{ev}_\alpha(f)$

である。故に ev_α は線型写像である。$V \subset \mathbb{R}[x]$ は部分空間なので、ev_α の定義域を V に制限してもやはり線型写像である。
(3) $W := \{f(x) \in V \mid \mathrm{ev}_\alpha(f) = 0\}$ とおく[*15]。このとき (2) より

- 任意の $f(x), g(x) \in W$ に対し、$\mathrm{ev}_\alpha\bigl(f(x) + g(x)\bigr) = f(\alpha) + g(\alpha) = 0 + 0 = 0$ となるので、$f(x) + g(x) \in W$ となる。
- 任意の $f(x) \in W$ と $a \in \mathbb{R}$ に対し、$\mathrm{ev}_\alpha\bigl(af(x)\bigr) = f(\alpha) = 0$ となるので、$af(x) \in W$ となる。

これより W は V の部分空間である。 ■

3.3　数列の空間

線型代数をやると「数列の空間」などというものも考えることができるようになります。数列とは言うまでもなく数が並んだものですが、今度はその数列を、つまり数列に並んでいる数全部をひっくるめて 1 つのものとして扱います。

■数列の記法　数列 $a_0, a_1, \ldots, a_n$ のことを、ひとまとめにして $(a_n)_{n\geq 0}$ と表します。ここで数列を表すとき、このプリントでは中括弧ではなく丸かっこを使うことにします。というのも中括弧で $\{a_0, a_1, a_2, \ldots\}$ と書くと集合のように見え、特に項の順序がどうでもいいような気がしてしまう[*16]からです。丸かっこで $(a_0, a_1, a_2, \ldots)$ とすれば一見して「順序を入れ替えたら違う数列になる」ことが分かり、このような問題は起きません[*17]。

■数列の演算　さて 2 つの数列 $(a_n)_{n\geq 0}$ と $(b_n)_{n\geq 0}$ があったとします。この 2 つの数列の「足し算」をどう定めるべきかと言ったら、それは各項を足して新しい数列を定めるのが自然でしょう。

$$\begin{array}{r@{\;}l}
(a_n)_{n\geq 0} &= (\quad a_0 \quad , \quad a_1 \quad , \quad a_2 \quad , \quad a_3 \quad , \ldots) \\
+\,(b_n)_{n\geq 0} &= (\quad b_0 \quad , \quad b_1 \quad , \quad b_2 \quad , \quad b_3 \quad , \ldots) \\ \hline
(a_n)_{n\geq 0} + (b_n)_{n\geq 0} &:= (a_0 + b_0, a_1 + b_1, a_2 + b_2, a_3 + b_3, \ldots)
\end{array}$$

$(a_n)_{n\geq 0}$ と $(b_n)_{n\geq 0}$ を足した数列の第 n 項は $a_n + b_n$ だから、$(a_n)_{n\geq 0} + (b_n)_{n\geq 0} := (a_n + b_n)_{n\geq 0}$ です[*18]。

[*14] 一般に、写像 f の α における値 α を見ることを「f の α での値を評価する」と言ったりします。この言葉を使うと、f に $f(\alpha)$ を対応させることは、α での値を評価することそのものです。そこでこの写像を α における**評価写像** (evaluation map) といいます。ev_α の記号はここから取りました。
[*15] つまり W は、$x = \alpha$ を解に持つ多項式の全体です。
[*16] たとえば集合だと $\{1, 2\} = \{2, 1\}$ ですよね。一方、丸かっこなら普通は $(1, 2) \neq (2, 1)$ と思うでしょう。
[*17] ですが岩波書店から出ている『数学辞典』を確かめると、数列を中括弧 $\{\}$ で表しています。なので括弧はどっちを使ってもいいです。
[*18] 慣れない人のため、式の読み方を説明しておきます。この右辺 $(a_n + b_n)_{n\geq 0}$ は、第 n 項が $a_n + b_n$ であるような数列を表します。一方、左辺は数列 $(a_n)_{n\geq 0}$ と数列 $(b_n)_{n\geq 0}$ の和です。そして、左辺の数列の和を、右辺で定義しています。たとえば $(a_n)_{n\geq 0} = (2n)_{n\geq 0} =$

また数列 $(a_n)_{n\geq 0}$ と実数 $\alpha \in \mathbb{R}$ が与えられたとき「数列 $(a_n)_{n\geq 0}$ の α 倍」はどう定義すれば良いでしょうか。これも全ての項を一斉に α 倍するのが自然でしょう。

$$\alpha(a_n)_{n\geq 0} = \alpha(a_0, a_1, a_2, \ldots) := (\alpha a_0, \alpha a_1, \alpha a_2, \ldots) = (\alpha a_n)_{n\geq 0}$$

いま、実数列全体の集合を $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ と書くことにします[*19]。これまでの議論で $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ 上には加法とスカラー倍とが定義されました。この $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ がきちんと線型空間になることを、チェックしましょう。58 ページの定義と照らし合わせながら、以下の証明を読んでください。

- 加法について
  - 数列 $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}, (c_n)_{n\geq 0}$ に対し、数列の和の定義を繰り返し使うと

    $$\begin{aligned} \bigl\{(a_n)_{n\geq 0} + (b_n)_{n\geq 0}\bigr\} + (c_n)_{n\geq 0} &= (a_n + b_n)_{n\geq 0} + (c_n)_{n\geq 0} = (a_n + b_n + c_n)_{n\geq 0} \\ (a_n)_{n\geq 0} + \bigl\{(b_n)_{n\geq 0} + (c_n)_{n\geq 0}\bigr\} &= (a_n)_{n\geq 0} + (b_n + c_n)_{n\geq 0} = (a_n + b_n + c_n)_{n\geq 0} \end{aligned}$$

    が得られる。確かに加法の結果は順番に依存していない。
  - 全ての項が 0 であるような数列 $(0)_{n\geq 0}$ は、いかなる数列 $(a_n)_{n\geq 0}$ に対しても $(a_n)_{n\geq 0} + (0)_{n\geq 0} = (a_n)_{n\geq 0} = (0)_{n\geq 0} + (a_n)_{n\geq 0}$ を満たす。このような性質を満たす数列が $(0)_{n\geq 0}$ に限ることは、全ての実数 $a \in \mathbb{R}$ に対して $a + b = a$ を満たす数 b が 0 しかないことから従う。
  - 数列 $(a_n)_{n\geq 0}$ に対し、その全ての項を (-1) 倍した数列 $(-a_n)_{n\geq 0}$ は、$(a_n)_{n\geq 0} + (-a_n)_{n\geq 0} = \bigl(a_n + (-a_n)\bigr)_{n\geq 0} = (0)_{n\geq 0}$ を満たす。また $(a_n)_{n\geq 0}$ に足したら $(0)_{n\geq 0}$ になるような数列がこれ以外に存在しないことは、$a + b = 0$ となる実数 b が $-a$ しかないことから従う。
- スカラー倍について
  - 任意の数列 $(a_n)_{n\geq 0}$ と実数 $\alpha, \beta \in \mathbb{R}$ に対して、数列のスカラー倍の定義を繰り返し使うと $\alpha\bigl\{\beta(a_n)_{n\geq 0}\bigr\} = \alpha(\beta a_n)_{n\geq 0} = (\alpha\beta a_n)_{n\geq 0} = (\alpha\beta)(a_n)_{n\geq 0}$ となる。
  - 任意の数列 $(a_n)_{n\geq 0}$ に対して、$1 \cdot (a_n)_{n\geq 0} = (1 \cdot a_n)_{n\geq 0} = (a_n)_{n\geq 0}$ である。
- 分配法則について
  - 任意の数列 $(a_n)_{n\geq 0}, (b_n)_{n\geq 0}$ と実数 $\alpha \in \mathbb{R}$ に対し、

    $$\begin{aligned} \alpha\bigl\{(a_n)_{n\geq 0} + (b_n)_{n\geq 0}\bigr\} &= \alpha(a_n + b_n)_{n\geq 0} && \text{(数列の和の定義)} \\ &= \bigl(\alpha(a_n + b_n)\bigr)_{n\geq 0} && \text{(数列のスカラー倍の定義)} \\ &= (\alpha a_n + \alpha b_n)_{n\geq 0} && \text{(各項ごとに括弧を展開)} \\ &= \alpha(a_n)_{n\geq 0} + \alpha(b_n)_{n\geq 0} && \text{(数列の和の定義)} \end{aligned}$$

    が成り立つ。
  - 任意の数列 $(a_n)_{n\geq 0}$ と実数 $\alpha, \beta \in \mathbb{R}$ に対し

    $$\begin{aligned} (\alpha + \beta)(a_n)_{n\geq 0} &= \bigl((\alpha + \beta)a_n\bigr)_{n\geq 0} && \text{(数列のスカラー倍の定義)} \\ &= (\alpha a_n + \beta a_n)_{n\geq 0} && \text{(各項ごとに括弧を展開)} \\ &= (\alpha a_n)_{n\geq 0} + (\beta a_n)_{n\geq 0} && \text{(数列の和の定義)} \\ &= \alpha(a_n)_{n\geq 0} + \beta(a_n)_{n\geq 0} && \text{(数列のスカラー倍の定義)} \end{aligned}$$

    が成り立つ。

$(0, 2, 4, 6, 8, \ldots)$, $(b_n)_{n\geq 0} = (3n)_{n\geq 0} = (0, 3, 6, 9, 12, 15, \ldots)$ なら、$(2n)_{n\geq 0} + (3n)_{n\geq 0} = (5n)_{n\geq 0} = (0, 5, 10, 15, 20, \ldots)$ といった感じです。見た目がややこしいだけで記号の意味さえ分かってしまえば難しくないですから、落ち着いて読んでください。

[*19] この記号の意味は、実は問 2 の $\mathrm{Map}(S, \mathbb{R})$ と整合的です。詳しいことは後で説明します。

■問 7 の解答　漸化式 $a_n + a_{n+1} = a_{n+2}$ を満たす数列 $(a_n)_{n\geq 0}$ の全体を V とする。

(1) 数列全体の空間 $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ が線型空間になることは既に確認した[*20]ので、V がその部分空間になっていることさえ示せばよい。まず、全ての項が 0 である数列 $(0)_{n\geq 0}$ は漸化式 $a_{n+2} = a_{n+1} + a_n$ を満たすので、$(0)_{n\geq 0} \in V$ である。よって V は空でない。そして

- $(a_n)_{n\geq 0}, (b_n)_{n\geq 0} \in V$ とする。このとき $(a_n)_{n\geq 0} + (b_n)_{n\geq 0} = (a_n + b_n)_{n\geq 0}$ である。そして $(a_n + b_n) + (a_{n+1} + b_{n+1}) = a_n + a_{n+1} + b_n + b_{n+1} = a_{n+2} + b_{n+2}$ なので、数列 $(a_n + b_n)_{n\geq 0}$ も同じ漸化式を満たす。つまり $(a_n)_{n\geq 0} + (b_n)_{n\geq 0} \in V$ である。
- $(a_n)_{n\geq 0} \in V$, $\alpha \in \mathbb{R}$ とする。$\alpha(a_n)_{n\geq 0} = (\alpha a_n)_{n\geq 0}$ である。そして $\alpha a_n + \alpha a_{n+1} = \alpha(a_n + a_{n+1}) = \alpha a_{n+2}$ なので、数列 $\alpha(a_n)_{n\geq 0}$ も同じ漸化式を満たす。つまり $\alpha(a_n)_{n\geq 0} \in V$ である。

よって、V は部分空間の条件を満たしている。

(2) $\varphi\colon V \to \mathbb{R}^2$ を、$\varphi\bigl((a_n)_{n\geq 0}\bigr) := {}^t(a_0, a_1)$ と定める[*21]。次のようにして、φ が線型写像の条件を満たすことが確かめられる。

$$\begin{aligned}
\varphi\bigl((a_n)_{n\geq 0} + (b_n)_{n\geq 0}\bigr) &= \varphi\bigl((a_n + b_n)_{n\geq 0}\bigr) && \text{(数列の和の定義)} \\
&= \begin{pmatrix} a_0 + b_0 \\ a_1 + b_1 \end{pmatrix} && \text{(φ の定義)} \\
&= \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} + \begin{pmatrix} b_0 \\ b_1 \end{pmatrix} && \text{($\mathbb{R}^2$ の和の定義)} \\
&= \varphi\bigl((a_n)_{n\geq 0}\bigr) + \varphi\bigl((b_n)_{n\geq 0}\bigr) && \text{(φ の定義)}
\end{aligned}$$

$$\begin{aligned}
\varphi\bigl(\alpha(a_n)_{n\geq 0}\bigr) &= \varphi\bigl((\alpha a_n)_{n\geq 0}\bigr) && \text{(数列のスカラー倍の定義)} \\
&= \begin{pmatrix} \alpha a_0 \\ \alpha a_1 \end{pmatrix} && \text{(φ の定義)} \\
&= \alpha \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} && \text{($\mathbb{R}^2$ のスカラー倍の定義)} \\
&= \alpha\varphi\bigl((a_n)_{n\geq 0}\bigr) && \text{(φ の定義)}
\end{aligned}$$

(3) V に属する $(a_n)_{n\geq 0}$ が等比数列で、かつ初項が $a_0 = 1$ を満たしたとする。このとき数列 $(a_n)_{n\geq 0}$ は等比数列だから、第 1 項を用いて $a_n = a_1^n$ と書ける。そして $(a_n)_{n\geq 0} \in V$ なので $a_2 = a_1 + a_0$ である。よって $a_1^2 = a_1 + 1$ でないといけない。この 2 次方程式を解いて、$a_1 = \frac{1\pm\sqrt{5}}{2}$ が従う。これで等比数列の候補が 2 つに絞られた。

逆に $a_1 = \frac{1\pm\sqrt{5}}{2}$ で定まる数列 $(a_n)_{n\geq 0} \in V$ が等比数列 $\left(\left(\frac{1\pm\sqrt{5}}{2}\right)^n\right)_{n\geq 0}$ になることを数学的帰納法で示せる。まず $a_0 = 1 = \left(\frac{1\pm\sqrt{5}}{2}\right)^0$, $a_1 = \frac{1\pm\sqrt{5}}{2}$ である。そして $a_n = \left(\frac{1\pm\sqrt{5}}{2}\right)^n$, $a_{n+1} = \left(\frac{1\pm\sqrt{5}}{2}\right)^{n+1}$ が分かっているとき

$$\begin{aligned}
a_{n+2} = a_{n+1} + a_n &= \left(\frac{1\pm\sqrt{5}}{2}\right)^{n+1} + \left(\frac{1\pm\sqrt{5}}{2}\right)^n = \left(\frac{1\pm\sqrt{5}}{2}\right)^n \left(1 + \frac{1\pm\sqrt{5}}{2}\right) \\
&= \left(\frac{1\pm\sqrt{5}}{2}\right)^n \left(\frac{3\pm\sqrt{5}}{2}\right) = \left(\frac{1\pm\sqrt{5}}{2}\right)^n \left(\frac{1\pm\sqrt{5}}{2}\right)^2 \\
&= \left(\frac{1\pm\sqrt{5}}{2}\right)^{n+2}
\end{aligned}$$

となる。

以上で、求める数列は $\left(\left(\frac{1\pm\sqrt{5}}{2}\right)^n\right)_{n\geq 0}$ の 2 つだと分かった。 ■

[*20] 本当なら解答にこの部分も含めるべきなのですが、プリントを作る都合上切り分けました。

[*21] 見れば分かりますが、φ は数列 $(a_n)_{n\geq 0}$ に対し、その初項 a_0 と第 1 項 a_1 のペアを対応させる写像です。

3.4 函数空間と微分方程式の解空間

今回扱う最後の例は、函数のなす空間です。既に多項式の全体 $\mathbb{R}[x]$ が線型空間になることを示しましたが、函数の全体を考えても線型空間ができるのです。

この先の議論では「函数それ自身」を 1 つのものとして扱うので、慣れないと戸惑うかもしれません。上で数列の空間を扱ったのと大体同じ状況ではありますが、もしかしたら数列の空間の方がとっつきやすいかもしれません。適宜参照してください。

■函数空間 $\mathrm{Map}(\mathbb{R},\mathbb{R})$ で、$\mathbb{R}$ 上の (連続性や微分可能性を仮定しない) 実数値函数全体の集合を表します。このとき 2 つの函数 f,g が与えられ「足し算を定義せよ」と言われたら、各 $x\in\mathbb{R}$ 毎に f と g の値を足し算するのが自然でしょう。つまり新しい函数 $f+g$ を、$(f+g)(x) := f(x)+g(x)$ で定義します。またスカラー倍についても、自然な定義のやり方は 1 通りです。$\alpha\in\mathbb{R}$ と函数 f に対し、新しい函数 αf の x における値は $(\alpha f)(x) := \alpha f(x)$ と定めます[*22]。

こう定めたときに、$\mathrm{Map}(\mathbb{R},\mathbb{R})$ が線型空間の公理を満たすことを確かめましょう。**2 つの函数 f と g が等しいとは、任意の $x\in\mathbb{R}$ に対し $f(x)=g(x)$ となること**でした。これを踏まえて、証明に取り掛かってください。

- 加法について
  - f,g,h を函数とする。このとき函数の和の定義を繰り返し使うと任意の $x\in\mathbb{R}$ に対して
$$\big((f+g)+h\big)(x) = (f+g)(x)+h(x) = f(x)+g(x)+h(x)$$
$$\big(f+(g+h)\big)(x) = f(x)+(g+h)(x) = f(x)+g(x)+h(x)$$
    となる。よって函数として $(f+g)+h = f+(g+h)$ である。
  - g を値が恒等的に 0 である函数とする。このとき任意に函数 f を取る。すると任意の x に対し $(f+g)(x) = f(x)+g(x) = f(x)+0 = f(x)$ となる。したがって函数として $f+g=f$ である。同様に $g+f=f$ が成り立つ。

    逆に函数 g が任意の函数 f に対し $f+g=f$ を満たすとする。このとき特に f として値が恒等的に 1 である函数を取ると、任意の x に対して $f(x)=1$, $(f+g)(x) = f(x)+g(x) = 1+g(x)$ となる。これらが等しいので、任意の x に対して $g(x)=0$ となる。
  - f を函数とする。このとき函数 g を $g(x) := -f(x)$ で定義すると、任意の x に対し $(f+g)(x) = f(x)+g(x) = f(x)-f(x) = 0$ となる。よって函数として $f+g=0$ である。逆に函数 g が $f+g=0$ を満たしたとすれば、任意の x に対して $f(x)+g(x)=0$ なので $g(x)=-f(x)$ となる。
  - f,g を函数とする。このとき任意の $x\in\mathbb{R}$ に対し $(f+g)(x) = f(x)+g(x) = g(x)+f(x) = (g+f)(x)$ である。よって函数として $f+g=g+f$ である。
- スカラー倍について
  - f を函数とし、$\alpha,\beta\in\mathbb{R}$ とする。このとき函数のスカラー倍の定義を繰り返し使うと、任意の $x\in\mathbb{R}$ に対し $\big(\alpha(\beta f)\big)(x) = \alpha\big((\beta f)(x)\big) = \alpha\big(\beta f(x)\big) = (\alpha\beta)f(x) = \big((\alpha\beta)f\big)(x)$ となることが分かる。よって函数として $\alpha(\beta f) = (\alpha\beta)f$ である。
  - f を函数とする。このとき任意の $x\in\mathbb{R}$ に対し $(1f)(x) = 1f(x) = f(x)$ なので、函数として $1f=f$ である。
- 分配法則について
  - f を函数とし、$\alpha,\beta\in\mathbb{R}$ とする。このとき任意の $x\in\mathbb{R}$ に対し、
$$\big((\alpha+\beta)f\big)(x) = (\alpha+\beta)f(x) \qquad \text{(函数のスカラー倍の定義)}$$

[*22] この右辺は「実数 $f(x)$ の α 倍」という意味です。$(\alpha f)(x)$ と $\alpha f(x)$ を混同しないでください。

$$
\begin{aligned}
&= \alpha f(x) + \beta f(x) && \text{(実数の分配法則)} \\
&= (\alpha f)(x) + (\beta f)(x) && \text{(函数のスカラー倍の定義)} \\
&= (\alpha f + \beta f)(x) && \text{(函数の和の定義)}
\end{aligned}
$$

となる。よって函数として $(\alpha+\beta)f = \alpha f + \beta f$ が成り立つ。

– f, g を函数とし、$\alpha \in \mathbb{R}$ とする。このとき任意の $x \in \mathbb{R}$ に対し、

$$
\begin{aligned}
\big(\alpha(f+g)\big)(x) &= \alpha\big((f+g)(x)\big) && \text{(函数のスカラー倍の定義)} \\
&= \alpha\big(f(x)+g(x)\big) && \text{(函数の和の定義)} \\
&= \alpha f(x) + \alpha g(x) && \text{(実数の分配法則)} \\
&= (\alpha f)(x) + (\alpha g)(x) && \text{(函数のスカラー倍の定義)} \\
&= (\alpha f + \alpha g)(x) && \text{(函数の和の定義)}
\end{aligned}
$$

となる。よって函数として $\alpha(f+g) = \alpha f + \alpha g$ が成り立つ。

■微分方程式の解空間　前回の問 3 を少し思い出しましょう。微分方程式 $y'' - y' - 6y = 0$ を考え

- y_1, y_2 が解なら、$y_1 + y_2$ も解
- y が解、$c \in \mathbb{R}$ なら、cy も解

という事をチェックしました。今にして思えば、この問題は「解全体の集合が函数空間の部分空間になること」を示す問題だったわけです。単振動の方程式でも全く同じようにして、解全体の集合が線型空間になることが言えます。

ちなみに単振動の方程式だったらすぐに解が分かってしまうのですが、**解全体の集合が線型空間になることを示すのに、微分方程式を解く必要はありません**。むしろ「線型空間になることが分かるから、解の全てが得られる」という話の流れになることが多いです。ですから以下の解答でも、方程式を解かないまま話を進めます。

■問 6 の解答　単振動の微分方程式 $y'' = -y$ を考える。この解全体のなす集合を V とする。

(1) V が函数全体の空間 $\mathrm{Map}(\mathbb{R}, \mathbb{R})$ の部分空間であることを示す。まず y が恒等的に 0 なら明らかに $y'' + y = 0$ が成り立つので、$0 \in V$ である。よって V は空ではない。そして

- $f_1, f_2 \in V$ とする。このとき $\big(f_1(x) + f_2(x)\big)'' = f_1''(x) + f_2''(x) = -f_1(x) - f_2(x) = -\big(f_1(x) + f_2(x)\big)$ より、$f_1 + f_2 \in V$ である。
- $f \in V$, $\alpha \in \mathbb{R}$ とする。このとき $\big(\alpha f(x)\big)'' = \alpha f''(x) = -\alpha f(x)$ より、$\alpha f \in V$ である。

これで V が部分空間だと言えた。

(2) 写像 $\varphi\colon V \to \mathbb{R}^2$ を $\varphi\big(f(x)\big) := {}^t\big(f(0), f'(0)\big)$ で定める[*23]。このとき

- 任意の $f, g \in V$ に対し
$$
\varphi(f+g) = \begin{pmatrix} (f+g)(0) \\ (f+g)'(0) \end{pmatrix} = \begin{pmatrix} f(0)+g(0) \\ f'(0)+g'(0) \end{pmatrix} = \begin{pmatrix} f(0) \\ f'(0) \end{pmatrix} + \begin{pmatrix} g(0) \\ g'(0) \end{pmatrix} = \varphi(f) + \varphi(g)
$$
が成り立つ。
- 任意の $f \in V$, $\alpha \in \mathbb{R}$ に対し
$$
\varphi(\alpha f) = \begin{pmatrix} (\alpha f)(0) \\ (\alpha f)'(0) \end{pmatrix} = \begin{pmatrix} \alpha f(0) \\ \alpha f'(0) \end{pmatrix} = \alpha \begin{pmatrix} f(0) \\ f'(0) \end{pmatrix} = \alpha\varphi(f)
$$

[*23] 集合と写像の記法に慣れないと目がチカチカするかもしれませんが、要は微分方程式の解に対し、その初期値 $f(0)$ と $f'(0)$ のペアを対応させているだけです。

が成り立つ。

よって φ は線型写像である。
(3) $a \in \mathbb{R}$ とする。$\mathbb{R}$ 上の函数 f に対し、$(\psi_a f)(x) := f(x-a)$ によって新しい函数 $\psi_a f$ を定める[*24]。このとき

- 任意の $x \in \mathbb{R}$ に対して

$$\begin{aligned}
\bigl(\psi_a(f+g)\bigr)(x) &= (f+g)(x-a) && (\psi_a \text{ の定義}) \\
&= f(x-a) + g(x-a) && (\text{函数の和の定義}) \\
&= (\psi_a f)(x) + (\psi_a g)(x) && (\psi_a \text{ の定義}) \\
&= \bigl((\psi_a f) + (\psi_a g)\bigr)(x) && (\text{函数の和の定義})
\end{aligned}$$

  が成り立つので、函数として $\psi_a(f+g) = \psi_a f + \psi_a g$ である。
- 任意の $x \in \mathbb{R}$ に対して

$$\begin{aligned}
\bigl(\psi_a(\alpha f)\bigr)(x) &= (\alpha f)(x-a) && (\psi_a \text{ の定義}) \\
&= \alpha f(x-a) && (\text{函数のスカラー倍の定義}) \\
&= \alpha\bigl((\psi_a f)(x)\bigr) && (\psi_a \text{ の定義}) \\
&= \bigl(\alpha(\psi_a f)\bigr)(x) && (\text{函数のスカラー倍の定義})
\end{aligned}$$

  が成り立つので、函数として $\psi_a(\alpha f) = \alpha(\psi_a f)$ である。

よって、ψ_a は線型写像である。また任意の $x \in \mathbb{R}$ に対し

$$\bigl((\psi_a f)(x)\bigr)' = \bigl(f(x-a)\bigr)' = f'(x-a) = (\psi_a f')(x)$$

が成り立つ。すなわち f を微分してから ψ_a を施しても、ψ_a を施してから微分しても結果は変わらない。これと ψ_a の線型性より $(\psi_a f)''(x) = (\psi_a f'')(x) = \bigl(\psi_a(-f)\bigr)(x) = (-\psi_a f)(x) = -(\psi_a f)(x)$ である。この式は $\psi_a f \in V$ を意味する。これで ψ_a が V から V への写像を定めることが言えた。 ∎

4 線型代数のさらなる一般論

さて、ここまでで既に大分色々なことを述べましたが、さらに一歩踏み込んで抽象的な議論をしましょう。「もう勘弁してほしい」と思っているかもしれませんが、線型代数が真の力を発揮するのはこの後です。

今回は「数列の空間」とか「写像の空間」とか「函数の空間」とか色々な空間が出てきましたが、実は適切なセッティングをすれば、全て「写像の空間」として扱うことができるのです。また「漸化式を満たす数列の空間」とか「微分方程式の解空間」とか「連立方程式の解空間」とかも出てきましたが、これらも個別に議論することはなく、全部まとめて扱うことができます。その一網打尽な感じを理解してください[*25]。

4.1 写像空間

まず手始めに、問題 2 で出てきた $\mathrm{Map}(S, \mathbb{R})$ を一般化し、ベクトル値の写像を考えます。

[*24] わざわざ言うまでもないかもしれませんが、函数を x 軸正の方向に a だけ平行移動させて、新しい函数を作っています。元の函数が f のとき、新しい函数を $\psi_a f$ と書いています

[*25] とは言っても無理は禁物です。今読めないようなら、暫くしてから振り返ってください。

■写像空間　V を線型空間、S を空でない集合とします。このとき S から V への写像全体の集合 $\mathrm{Map}(S, V)$ は

- $(f+g)(\boldsymbol{v}) := f(\boldsymbol{v}) + g(\boldsymbol{v})$
- $(\alpha f)(\boldsymbol{v}) := \alpha f(\boldsymbol{v})$

と定義することで、線型空間になります。この証明は、さっき 66 ページで函数空間が線型空間になることを示したのと全く同じです。ちなみに $\mathbb{R}$ も線型空間ですから、この問題で $V = \mathbb{R}$ とすると問 2 の解答になります。

■色々な例　さて $\mathrm{Map}(S, V)$ がベクトル空間になるためには、V がベクトル空間でさえあれば十分です。さらに S はベクトル空間である必要すらありません。そこで $\mathrm{Map}(S, V)$ の S と V を取り換えることで、色々なものがベクトル空間だと分かります。

たとえば $S = V = \mathbb{R}$ とすると、$\mathrm{Map}(\mathbb{R}, \mathbb{R})$ は (連続性や微分可能性を一切仮定しない) $\mathbb{R}$ 上の実数値函数全体の集合です。一般論から、これは自動的に線型空間になります。また $S = \mathbb{N}$, $V = \mathbb{R}$ とします。そうすると $a \in \mathrm{Map}(\mathbb{N}, \mathbb{R})$ は $a\colon \mathbb{N} \to \mathbb{R}$ という写像になります。定義域が $\mathbb{N}$ ですから、各自然数毎に $a(0), a(1), a(2), \ldots$ という値が定まっています。つまり a は数列に他なりません。$a(n)$ のことを a_n と書けば、より一層雰囲気が出るでしょう。このように**数列は、$\mathbb{N}$ から $\mathbb{R}$ への写像として捉えることができる**のです。こうすると一般論から、直ちに $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ は線型空間だと分かります。

このように、数列の空間や函数の空間に関する議論は一々個別に行う必要がなく、全て $\mathrm{Map}(S, V)$ の話に帰着させられるのです。これが一般論の力です。

4.2　線型写像の空間

ベクトル空間 V に値を取る写像の集合 $\mathrm{Map}(S, V)$ が線型空間になることが示せたわけですが、もし S も線型空間であれば、$\mathrm{Map}(S, V)$ の中に「線型写像の全体」という部分集合を考えることができます。実はこれは $\mathrm{Map}(S, V)$ の部分空間になります。

■Hom と End　V, W が共に線型空間のとき、V から W への線型写像の全体を $\mathrm{Hom}_{\mathbb{R}}(V, W)$ と表します。また、特に $\mathrm{Hom}_{\mathbb{R}}(V, V)$ のことを $\mathrm{End}_{\mathbb{R}}(V)$ とも表します[*26]。なんだか仰々しい記号ですが「数ベクトル空間の間の線型写像は行列だった」という事実を思い出せば、Hom は**行列全体の集合** $\mathrm{Mat}_{m,n}(\mathbb{R})$ **を一般化しただけ**だと分かります。End の方は正方行列に対応しています。また Hom や End の右下の $\mathbb{R}$ は、一々書かずに省略することも多いです。

数ベクトル空間 $\mathbb{R}^n$ が線型空間であることは今更言うまでもないですが、行列は数ベクトルを横に並べただけで、しかも加法のスカラー倍の定義はベクトルのときとまったく同じです。ですから $\mathrm{Mat}_{m,n}(\mathbb{R})$ は線型空間になります。よって一般化した $\mathrm{Hom}(V, W)$ や $\mathrm{End}(V)$ も線型空間になりそうです。実際に $\mathrm{Hom}(V, W)$ が $\mathrm{Map}(V, W)$ の部分空間であることは、次のようにチェックできます。

- $f, g \in \mathrm{Hom}(V, W)$ とする。このとき
  - 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ に対し

$$\begin{aligned}
(f+g)(\boldsymbol{u}+\boldsymbol{v}) &= f(\boldsymbol{u}+\boldsymbol{v}) + g(\boldsymbol{u}+\boldsymbol{v}) && \text{(写像の和の定義)} \\
&= f(\boldsymbol{u}) + f(\boldsymbol{v}) + g(\boldsymbol{u}) + g(\boldsymbol{v}) && \text{(f と g の線型性)} \\
&= (f+g)(\boldsymbol{u}) + (f+g)(\boldsymbol{v}) && \text{(写像の和の定義)}
\end{aligned}$$

が成り立つ。

[*26] 線型写像のことを、線型空間の間の準同型写像 (homomorphism) と呼ぶことがあります。また線型写像の定義域と値域が同じ線型空間のとき、それを自己準同型 (endomorphism) と呼ぶことがあります。Hom と End の名前はここに由来します。

– 任意の $\boldsymbol{u} \in V$ と $\alpha \in \mathbb{R}$ に対し

$$\begin{aligned}(f+g)(\alpha\boldsymbol{u}) &= f(\alpha\boldsymbol{u}) + g(\alpha\boldsymbol{u}) && (\text{写像の和の定義})\\ &= \alpha f(\boldsymbol{u}) + \alpha g(\boldsymbol{u}) && (f\ \text{と}\ g\ \text{の線型性})\\ &= \alpha(f+g)(\boldsymbol{u}) && (\text{写像の和の定義})\end{aligned}$$

が成り立つ。

よって $f+g$ は線型写像と分かり、$f+g \in \mathrm{Hom}(V,W)$ となる。

- $f \in \mathrm{Hom}(V,W)$, $\alpha \in \mathbb{R}$ とする。このとき

– 任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ に対し

$$\begin{aligned}(\alpha f)(\boldsymbol{u}+\boldsymbol{v}) &= \alpha\bigl(f(\boldsymbol{u}+\boldsymbol{v})\bigr) && (\text{写像のスカラー倍の定義})\\ &= \alpha\bigl(f(\boldsymbol{u}) + f(\boldsymbol{v})\bigr) && (f\ \text{の線型性})\\ &= \alpha f(\boldsymbol{u}) + \alpha f(\boldsymbol{v}) && (\text{分配法則})\\ &= (\alpha f)(\boldsymbol{u}) + (\alpha f)(\boldsymbol{v}) && (\text{写像のスカラー倍の定義})\end{aligned}$$

が成り立つ。

– 任意の $\boldsymbol{u} \in V$ と $\beta \in \mathbb{R}$ に対し

$$\begin{aligned}(\alpha f)(\beta\boldsymbol{u}) &= \alpha f(\beta\boldsymbol{u}) && (\text{写像のスカラー倍の定義})\\ &= \alpha\beta f(\boldsymbol{u}) && (f\ \text{の線型性})\\ &= \beta(\alpha f)(\boldsymbol{u}) && (\text{写像のスカラー倍の定義})\end{aligned}$$

が成り立つ。

よって (αf) は線型写像であり、$\alpha f \in \mathrm{Hom}(V,W)$ である。

■線型写像の合成　U, V, W を線型空間とし、$f\colon U \to V$ と $g\colon V \to W$ を共に線型写像とします。このとき合成写像 $g \circ f\colon U \to W$ は再び線型写像になります。線型写像が行列のときは「行列の掛け算がまた行列になる」という当たり前のことしか言っていないのですが、抽象的な線型空間の場合にも、次のようにして確かめられます。

- 任意の $\boldsymbol{u}, \boldsymbol{v} \in U$ に対し

$$\begin{aligned}(g\circ f)(\boldsymbol{u}+\boldsymbol{v}) &= g\bigl(f(\boldsymbol{u}+\boldsymbol{v})\bigr) && (\text{合成写像の定義})\\ &= g\bigl(f(\boldsymbol{u}) + f(\boldsymbol{v})\bigr) && (f\ \text{の線型性})\\ &= g\bigl(f(\boldsymbol{u})\bigr) + g\bigl(f(\boldsymbol{v})\bigr) && (g\ \text{の線型性})\\ &= (g\circ f)(\boldsymbol{u}) + (g\circ f)(\boldsymbol{v}) && (\text{合成写像の定義})\end{aligned}$$

が成り立つ。

- 任意の $\boldsymbol{u} \in U$ とスカラー倍 $\alpha \in \mathbb{R}$ に対し

$$\begin{aligned}(g\circ f)(\alpha)\boldsymbol{u} &= g\bigl(f(\alpha\boldsymbol{u})\bigr) && (\text{合成写像の定義})\\ &= g\bigl(\alpha f(\boldsymbol{u})\bigr) && (f\ \text{の線型性})\\ &= \alpha\bigl(g(f(\boldsymbol{u}))\bigr) && (g\ \text{の線型性})\\ &= \alpha(g\circ f)(\boldsymbol{u}) && (\text{合成写像の定義})\end{aligned}$$

が成り立つ。

これで $g \circ f$ が線型写像になることが言えました。

■End **環**　V を線型空間とします。既に示したように $\mathrm{End}(V) = \mathrm{Hom}(V, V)$ は線型空間になります。一方で $f, g \in \mathrm{End}(V)$ なら、$\mathrm{End}(V)$ の中で写像の合成 $f \circ g$ を考えることができます。すると「写像の合成と線型空間の構造との間には、どんな関係があるのか」という疑問が沸き起こってきます。

定義を考えれば分かるのですが、$f, g, h \in \mathrm{End}(V)$ のとき $(f + g) \circ h = f \circ h + g \circ h$, $f \circ (g + h) = f \circ g + f \circ h$ が成り立ちます。ここで写像の合成を表す $\circ$ を省略してみると、今の式は $(f + g)h = fh + gh$, $f(g + h) = fg + fh$ となります。つまり $\mathrm{End}(V)$ では、**写像の合成 $\circ$ を掛け算だと思ったとき、分配法則が成り立っている**のです。また $a \in \mathbb{R}$ で $f, g \in \mathrm{End}(V)$ のとき $(af) \circ g = f \circ (ag) = a(f \circ g)$ が成り立ちます。つまり**写像の合成とスカラー倍とは順序が入れ替えられます**。

このように $\mathrm{End}(V)$ は

- 「足し算と掛け算が上手くできる」という意味で整数のような性質を持っており[27]
- 線型空間の構造を持っており
- 写像の合成とスカラー倍との順序が入れ替えられる

という非常に良い性質を持っているのです。こういうこととその他いくつかの条件を合わせて、$\mathrm{End}(V)$ は**環構造**あるいは $\mathbb{R}$ **代数の構造**を持つといいます。

特に $\mathrm{End}(V)$ の環構造からは「$\mathrm{End}(V)$ に属する線型写像を、好き勝手に足したりかけたりしてできるものが、再び線型写像になる」という事実が分かります[28]。これに注意しておきましょう。

4.3　線型写像の核

$f\colon V \to W$ が線型写像であるとき、$\mathrm{Ker}\, f := \{\boldsymbol{v} \in V \mid f(\boldsymbol{v}) = \boldsymbol{0}\}$ と定めます。このとき $\mathrm{Ker}\, f \subset W$ は W の部分空間になっています。実際 $\boldsymbol{0} \in \mathrm{Ker}\, f$ より $\mathrm{Ker}\, f$ は空でありません。そして

- $\boldsymbol{u}, \boldsymbol{v} \in \mathrm{Ker}\, f$ のとき、$f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v}) = \boldsymbol{0} + \boldsymbol{0} = \boldsymbol{0}$
- $\boldsymbol{u} \in \mathrm{Ker}\, f$, $\alpha \in \mathbb{R}$ のとき、$f(\alpha\boldsymbol{u}) = \alpha f(\boldsymbol{u}) = \alpha\boldsymbol{0} = \boldsymbol{0}$

が成り立ちます。これで $\mathrm{Ker}\, f$ は部分空間だと分かりました。

さて、今の証明を問題 4、あるいは問題 5 の (3) と見比べてください。出てくる文字や写像の名前が違うだけで、やっていることは全て同じですよね。というか「線型写像の核が部分空間になる」という事実を知っていれば、

- 問題 4 は $V = \mathrm{Ker}\, A$
- 問題 5 (3) は $W = \mathrm{Ker}\, \mathrm{ev}_\alpha$

の一言で片付きます。これが一般論の力です。

さらに言ってしまえば、問題 7 の (1) も線型写像の核として記述できます。数列の空間 $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ 上の「左シフト写像」$L\colon \mathrm{Map}(\mathbb{N}, \mathbb{R}) \to \mathrm{Map}(\mathbb{N}, \mathbb{R})$ を、

$$L\bigl((a_n)_{n \geq 0}\bigr) := (a_{n+1})_{n \geq 0} = (a_1, a_2, \ldots)$$

で定めます。つまり L は、数列の左端を切り落として新しい数列を作る写像です。この S は線型写像です。実際

[27] ただし、掛け算の順序が入れ替えられないことには注意する必要があります。

[28] なんか難しいことを言ってる気がしますが、V が「有限次元」という条件を満たせば、正方行列を足したりかけたりして正方行列ができると言っているだけです。

- 任意の数列 $(a_n)_{n\geq 0}, (b_n)_{n\geq 0} \in \mathrm{Map}(\mathbb{N}, \mathbb{R})$ に対し

$$
\begin{aligned}
L\big((a_n)_{n\geq 0} + (b_n)_{n\geq 0}\big) &= L\big((a_n + b_n)_{n\geq 0}\big) && \text{(数列の和の定義)} \\
&= (a_{n+1} + b_{n+1})_{n\geq 0} && \text{(左シフト写像の定義)} \\
&= (a_{n+1})_{n\geq 0} + (b_{n+1})_{n\geq 0} && \text{(数列の和の定義)} \\
&= L\big((a_n)_{n\geq 0}\big) + L\big((b_n)_{n\geq 0}\big) && \text{(左シフト写像の定義)}
\end{aligned}
$$

- 任意の数列 $(a_n)_{n\geq 0} \in \mathrm{Map}(\mathbb{N}, \mathbb{R})$ と実数 $\alpha \in \mathbb{R}$ に対し

$$
\begin{aligned}
L\big(\alpha(a_n)_{n\geq 0}\big) &= L\big((\alpha a_n)_{n\geq 0}\big) && \text{(数列のスカラー倍の定義)} \\
&= (\alpha a_{n+1})_{n\geq 0} && \text{(左シフト写像の定義)} \\
&= \alpha(a_{n+1})_{n\geq 0} && \text{(数列のスカラー倍の定義)} \\
&= \alpha L\big((a_n)_{n\geq 0}\big) && \text{(左シフト写像の定義)}
\end{aligned}
$$

となっています。これより L の 2 回合成 L^2 も線型写像です。さらに線型写像をスカラー倍したり足したりしても線型写像なので、$L^2 - L - \mathrm{id}$ が線型写像だと分かりました[*29]。そして数列 $(a_n)_{n\geq 0}$ に対し、$(L^2 - L - \mathrm{id})\big((a_n)_{n\geq 0}\big)$ の第 n 項目は、$a_{n+2} - a_{n+1} - a_n$ です。ですから漸化式 $a_n + a_{n+1} = a_{n+2}$ を満たす数列の全体は、線型写像 $L^2 - L - \mathrm{id}$ の核 $\mathrm{Ker}(L^2 - L - \mathrm{id})$ で与えられるというわけです。かくして問題 7 の (1) も、核が部分空間という一般論から直ちに従うことが分かりました。

問題 6 の (1) も似たようにして解けます。まず問題の仮定では $y'' + y = 0$ を満たす「2 回微分可能な函数」だけを考えていましたが、$y'' = -y$ の解は何回でも微分可能なことを確かめましょう。実際 y が 2 回微分可能で $y'' = -y$ とすると、$-y$ はさらにもう 2 回微分できます。その結果 $y^{(4)} = (y'')'' = (-y)'' = y$ と、y は 4 回微分で元に戻ります。ですから y は 4 の倍数回微分可能で、これを繰り返し使えば 8 回微分可能とか 16 回微分可能とか 3864 回微分可能とかがたちどころに分かります。かくして y は無限回微分可能[*30]です。というわけで V は、無限回微分可能な函数全体のなす線型空間 $C^\infty(\mathbb{R})$ の部分空間になります。

そして微分 $\frac{d}{dx}$ は $C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$ という写像を定めます。よく知られているように $(y_1 + y_2)' = y_1' + y_2'$, $(cy)' = cy'$ だから、微分は線型写像です。したがって微分を合成して得られる 2 回微分 $\frac{d^2}{dx^2}$ も線型写像です。線型写像全体の空間は線型空間でしたから、$\frac{d^2}{dx^2} - \mathrm{id}$ も線型写像です。そうすると微分方程式 $y'' + y = 0$ の解空間は $\left(\frac{d}{dx}\right)^2 - \mathrm{id}$ という線型写像の核として表せます。これで解の集合が線型空間になると言えました。

4.4 まとめ

いかがでしょうか？ ここまで非常に長々と文章を書き連ねてきましたが、結局線型代数の一般論が良く分かってさえいれば、今回の一連の問題における線型代数的な議論は

- 写像空間 $\mathrm{Map}(S, V)$ が線型空間になることを示す
- $\mathrm{End}(V)$ が環構造を持つことを示す
- 線型写像の核 Ker が部分空間になることを示す

という**たった 3 つに集約できる**のです。単に問題の解説を並べるだけだったら、数ページで収まることでしょう。何度も言いますが、これが線型代数の力です。皆さんも 1 年生が終わる頃には、自由自在に線型代数を使いこなせるようになってください。

[*29] ここで id は恒等写像、すなわち全ての x に対し $\mathrm{id}(x) = x$ を満たす写像です。

[*30] 「無限回微分可能」とは「任意の自然数 $n \in \mathbb{N}$ に対し n 回微分可能」という意味です。「無限回の微分」という操作を定義しているわけではありません。念のため。

線形代数学① 第 8 回　　ベクトル空間と線型写像の核と像

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 6 月 17 日 1 限

1　数学における議論のしかたについて

今回の答案を採点していて、数学での抽象的な議論の仕方に慣れていない人が一定数いるような気がしたので、少々補足をしようと思います。

■論理的であること　まず、数学では**論理的であること**が非常に大事です。たとえば

- 「P かつ Q を示す」というタイプの証明 (たとえば、部分空間の条件を全て確かめる問題) では、条件 P と条件 Q が両方とも示されているかどうか
- 「すべての x に対し $P(x)$ が成り立つことを示す」というタイプの証明では、本当に x は何でも大丈夫か
- 「$P(x)$ となる x が存在する」というタイプの証明では、x を具体的に構成できているか、あるいは x の存在を保証する定理を何か使っているか
- 「P かつ $P \Rightarrow Q$ が成り立つので、Q が成り立つ」というタイプの証明では、何が P で何が Q なのか、なぜ $P \Rightarrow Q$ が成り立つのか

に気を付けてください。これらが不明瞭だと、あっという間に議論がよく分からなくなってしまいます。もし記号論理学などの授業を取っている場合は、自分の証明が述語論理の形式に乗っかっているかを確かめると良いと思います。また簡単な判定法として「他人がこれを読んだら、意味が分かるかどうか」を想像してみることをお勧めします。

■「イメージ」への頼り方　数学は厳密だとは言いますが、一方でイメージを持つことは大変重要です。数学を専門にする人は何かの定理の証明を読むにしても、本当に字面だけ追いかけることはまずありません。定理の証明がなぜ上手くいくのかを、適切な具体例で確かめたり図を描いたりしながら、理解を深めていきます。特に「非自明な中で最も簡単な例」を見ることは、非常に重要です。

かたや数学は厳密ですから、数学的なものに対する「イメージ」が何を表しているのかには、十分な注意を払う必要があります。たとえば今回、線型空間 W_1 と W_2 の共通部分や和集合が再び線型空間になるかを考える問題で、Venn 図[*1]を描いた人がいました。このイメージでは、線型空間が「集合」であるという面は捉えられています。ですが線型空間が「まっすぐ」であるということは、全く反映されていません。そういう意味で、線型空間を表すのに「丸」を描くのは適当ではないと言えるでしょう。

こんな感じで、数学的対象の「イメージ」は真実を反映する部分とそうでない部分とがあります。自分が何を捉えているのか見失わないよう、気を付けてイメージを使ってください。

■判定問題　今回、いくつか「〜であるかどうか？」という問題があります。この手の問題を見たら、何も言われずとも自分で後ろに「**正しければ証明し、誤りならば反例を挙げよ**」と補ってください。また誤っている場合について、多くの人が「一般には〜とは言えない」という書き方をしていました。これでは正しくないのかはっきりしませんから、正しくないなら決定的な証拠である反例を 1 つ、必ず**具体的に**挙げましょう。

[*1] 和集合や共通部分を議論するときによく使う、交わる丸 2 個からなるあの図です。

2　部分空間について

2.1　ベクトルの張る部分空間

前回、部分空間が「原点を持つまっすぐな空間」だということを確認しました。今回はその作り方や、部分空間相互の関係について見ていきましょう。

まず、ベクトルが何本か与えられたら、それが「張る」部分空間というものを定義できます。直感的には明らかなのですが、まずは証明をしましょう。

■問 1 の解答　V を線型空間とする。
(1) $W \subset V$ を部分空間とする。このとき W は空でないから、何か $\boldsymbol{v} \in W$ が取れる。すると $0\boldsymbol{v} = \boldsymbol{0}$ である。W はスカラー倍で閉じているので、$\boldsymbol{0} \in W$ が従う。
(2) $\boldsymbol{a} \in V$ として、$\mathbb{R}\boldsymbol{a} := \{C\boldsymbol{a} \mid C \in \mathbb{R}\}$ とおく。すると

- $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}\boldsymbol{a}$ とする。このとき $\boldsymbol{u} = \alpha\boldsymbol{a}$, $\boldsymbol{v} = \beta\boldsymbol{a}$ となる実数 $\alpha, \beta \in \mathbb{R}$ が取れる。よって $\boldsymbol{u} + \boldsymbol{v} = (\alpha + \beta)\boldsymbol{a} \in \mathbb{R}\boldsymbol{a}$ となる。
- $\boldsymbol{u} \in \mathbb{R}\boldsymbol{a}$, $\alpha \in \mathbb{R}$ とする。このとき $\boldsymbol{u} = \beta\boldsymbol{a}$ となる実数 $\beta \in \mathbb{R}$ が取れる。すると $\alpha\boldsymbol{u} = \alpha\beta\boldsymbol{a} \in \mathbb{R}\boldsymbol{a}$ となる。

よって $\mathbb{R}\boldsymbol{a}$ は部分空間である。
(3) $\boldsymbol{a}, \boldsymbol{b} \in V$ として、$\mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b} := \{\alpha\boldsymbol{a} + \beta\boldsymbol{b} \in V \mid \alpha, \beta \in \mathbb{R}\}$ とおく。

- $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b}$ とする。このとき $\boldsymbol{u} = \alpha_1\boldsymbol{a} + \beta_1\boldsymbol{b}$, $\boldsymbol{v} = \alpha_2\boldsymbol{a} + \beta_2\boldsymbol{b}$ と書ける。よって $\boldsymbol{u} + \boldsymbol{v} = (\alpha_1 + \alpha_2)\boldsymbol{a} + (\beta_1 + \beta_2)\boldsymbol{b} \in \mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b}$ となる。
- $\boldsymbol{u} \in \mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b}$, $\gamma \in \mathbb{R}$ とする。このとき $\boldsymbol{u} = \alpha\boldsymbol{a} + \beta\boldsymbol{b}$ と書けるので $\gamma\boldsymbol{u} = \gamma(\alpha\boldsymbol{a} + \beta\boldsymbol{b}) = \gamma\alpha\boldsymbol{a} + \gamma\beta\boldsymbol{b} \in \mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b}$ を得る。

よって $\mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b}$ は部分空間である。 ■

■ベクトルの張る部分空間　今の問題で出てきた $\mathbb{R}\boldsymbol{a}$ を、ベクトル $\boldsymbol{a}$ が張る部分空間といいます。上では一応証明をつけましたが、絵で見れば「原点を通る $\boldsymbol{a}$ 方向の直線」が $\mathbb{R}\boldsymbol{a}$ です。線型空間が「原点を持つまっすぐな空間」ということを思い出せば、これが部分空間になることは明らかでしょう。

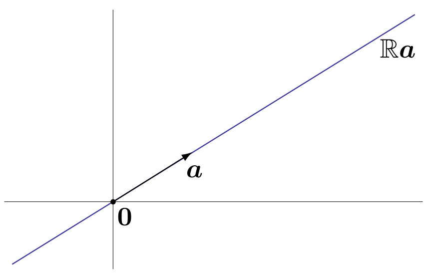


図 8.1　ベクトルの張る部分空間

$\mathbb{R}\boldsymbol{a} + \mathbb{R}\boldsymbol{b}$ も状況は同じです。今度はベクトルが 2 本あって、$\boldsymbol{a}$ と $\boldsymbol{b}$ の 1 次結合で表せるベクトル全体を考えているわけです。そうすれば原点を通る平面[*2]が張れることはほとんど明らかでしょう。

[*2] $\boldsymbol{a}$ と $\boldsymbol{b}$ が 1 次独立にならない場合でも、証明が破綻していないことに気を付けましょう。この場合はもちろん、張る部分空間は平面ではなく直線になります。

■「張る空間」の記号について　ここで、この $\mathbb{R}\boldsymbol{a}+\mathbb{R}\boldsymbol{b}$ という記号の心積もりを一応説明しておきます。スカラー α とベクトル $\boldsymbol{a}$ に対し、$\boldsymbol{a}$ の α 倍は $\alpha\boldsymbol{a}$ と書きます。また実数全体の集合は $\mathbb{R}$ で表されます。そこで「$\alpha\boldsymbol{a}$ の α を全てのスカラーの範囲で動かして、得られるベクトルをかき集める」という意味で $\mathbb{R}\boldsymbol{a}$ という表記が用いられます。

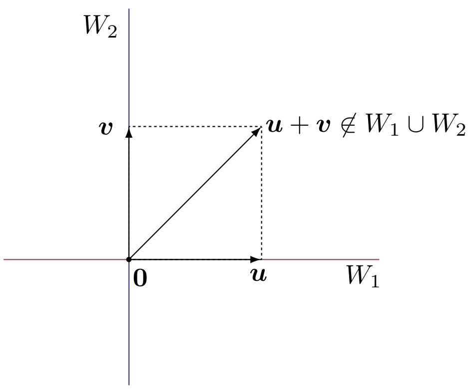
$\mathbb{R}\boldsymbol{a}+\mathbb{R}\boldsymbol{b}$ も同じです。$\mathbb{R}\boldsymbol{a}$ と $\mathbb{R}\boldsymbol{b}$ はどちらもベクトルの集合で、$\boldsymbol{u}\in\mathbb{R}\boldsymbol{a}$ と $\boldsymbol{v}\in\mathbb{R}\boldsymbol{b}$ に対して和 $\boldsymbol{u}+\boldsymbol{v}$ が考えられます。そして $\boldsymbol{u}\in\mathbb{R}\boldsymbol{a}$ と $\boldsymbol{v}\in\mathbb{R}\boldsymbol{b}$ とを全て動かしたときにできるベクトル $\boldsymbol{u}+\boldsymbol{v}$ をかき集めるというのが、$\mathbb{R}\boldsymbol{a}+\mathbb{R}\boldsymbol{b}$ という記号の意味です。

この他にも一般に、集合と元、あるいは集合と集合とが足し算やスカラー倍の記号を介して繋げられることがあります。その場合は「それぞれの集合から元を取ってきて演算をした結果を、ありとあらゆる元の組合せについてかき集める」と思ってください。後で問 2 に現れる部分空間 W_1, W_2 に対する記号 W_1+W_2 も、同じ意味です。

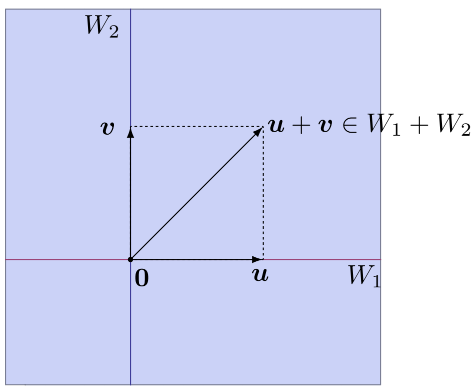
■2 つの部分空間　さて、集合の部分集合が 2 つ与えられたらそれらの共通部分や和集合を考えることができます。それらが線型空間の部分空間だった場合、どのような性質を示すでしょうか？

V を線型空間、$W_1, W_2\subset V$ をその部分空間とします。このとき**共通部分** $W_1\cap W_2$ **は再び部分空間になります**。ここでも「原点を持つまっすぐな空間」が線型空間だったことを思い出しましょう。W_1 と W_2 は両方とも部分空間だから原点を持ち、したがって $W_1\cap W_2$ も原点を持ちます。また W_1 と W_2 はまっすぐなので、その交わり $W_1\cap W_2$ もまっすぐです。ですから $W_1\cap W_2$ は部分空間になりそうです。絵としては「3 次元空間 $\mathbb{R}^3$ の中で、原点を通る 2 枚の平面の交わりが直線になっている様子」を想像できれば OK です。

一方、**和集合** $W_1\cup W_2$ **はほとんどの場合、部分空間になりません**。たとえば W_1 と W_2 が 1 次元だったら、その和集合 $W_1\cup W_2$ は交わる 2 本の直線でしかなく、その「間」がスカスカです。これでは真っ直ぐとは言えないでしょう。一般の場合も同じで、W_1 の方向のベクトルと W_2 の方向のベクトルを足し合わせた結果があさっての方向に飛んで行ってしまっては、$W_1\cup W_2$ には入りません。加えて今の議論から「$W_1\cup W_2$ が和集合になるには、$W_1\subset W_2$ または $W_2\subset W_1$ が必要」ということが推測できます。



(a) 和集合が部分空間にならない例



(b) 2 つの部分空間が張る部分空間

でも $W_1\cup W_2$ が部分空間にならないとはいえ、その「間」をきちんと埋めれば、W_1 と W_2 を両方とも含む部分空間が得られそうです。つまり「ベクトルの張る部分空間」を一般化し「2 **つの部分空間が張る部分空間**」を考えれば良いのです。それが W_1+W_2 です。

こうしたイメージを踏まえた上で、きちんと証明をつけましょう。

■問 2 の解答　V を線型空間、W_1, W_2 をその部分空間とする。

(1) $W_1\cap W_2$ が部分空間になることを、次のように確かめられる。

- $\boldsymbol{u}, \boldsymbol{v} \in W_1 \cap W_2$ とする。このとき $\boldsymbol{u}, \boldsymbol{v} \in W_1$ なので $\boldsymbol{u}+\boldsymbol{v} \in W_1$ である。また $\boldsymbol{u}, \boldsymbol{v} \in W_2$ なので $\boldsymbol{u}+\boldsymbol{v} \in W_2$ である。よって $\boldsymbol{u}+\boldsymbol{v} \in W_1 \cap W_2$ である。
- $\boldsymbol{u} \in W_1 \cap W_2$, $\alpha \in \mathbb{R}$ とする。このとき $\boldsymbol{u} \in W_1$ より $\alpha\boldsymbol{u} \in W_1$ である。また $\boldsymbol{u} \in W_2$ より $\alpha\boldsymbol{u} \in W_2$ である。よって $\alpha\boldsymbol{u} \in W_1 \cap W_2$ である。

(2) $W_1 \cup W_2$ は必ずしも線型空間ではない。反例を挙げる。

$V = \mathbb{R}^2$ とする。$W_1 := \mathbb{R}\boldsymbol{e}_1$, $W_2 := \mathbb{R}\boldsymbol{e}_2$*3とおくと、これはいずれも $\mathbb{R}^2$ の部分空間である。しかし $W_1 \cup W_2$ は部分空間でない。実際 $\boldsymbol{e}_1 \in W_1 \subset W_1 \cup W_2$, $\boldsymbol{e}_2 \in W_2 \subset W_1 \cup W_2$ だが $\boldsymbol{e}_1 + \boldsymbol{e}_2 \notin W_1 \cup W_2$ である。よって $W_1 \cup W_2$ は線型空間でない。

(3) $W_1 + W_2$ が部分空間になることを、次のように確かめられる。

- $\boldsymbol{u}, \boldsymbol{v} \in W_1 + W_2$ とする。このとき $\boldsymbol{u} = \boldsymbol{a}_1 + \boldsymbol{b}_1$, $\boldsymbol{v} = \boldsymbol{a}_2 + \boldsymbol{b}_2$ $(\boldsymbol{a}_1, \boldsymbol{a}_2 \in W_1,\ \boldsymbol{b}_1, \boldsymbol{b}_2 \in W_2)$ と書ける。よって $\boldsymbol{u}+\boldsymbol{v} = (\boldsymbol{a}_1+\boldsymbol{a}_2) + (\boldsymbol{b}_1+\boldsymbol{b}_2)$ となる。ここで W_1, W_2 は共に部分空間だから、$\boldsymbol{a}_1+\boldsymbol{a}_2 \in W_1$, $\boldsymbol{b}_1+\boldsymbol{b}_2 \in W_2$ となる。よって $\boldsymbol{u}+\boldsymbol{v} \in W_1 + W_2$ となる。
- $\boldsymbol{u} \in W_1 + W_2$, $\alpha \in \mathbb{R}$ とする。このとき $\boldsymbol{u} = \boldsymbol{a} + \boldsymbol{b}$ $(\boldsymbol{a} \in W_1,\ \boldsymbol{b} \in W_2)$ と書ける。W_1, W_2 は部分空間なので $\alpha\boldsymbol{a} \in W_1$, $\alpha\boldsymbol{b} \in W_2$ となる。よって $\alpha\boldsymbol{u} = \alpha(\boldsymbol{a}+\boldsymbol{b}) = \alpha\boldsymbol{a} + \alpha\boldsymbol{b} \in W_1 + W_2$ を得る。■

2.2 部分空間の例

部分空間の議論に慣れるため、目に見えない例を見てみます。問 3 で挙がっている空間が全て数列の空間 $\mathrm{Map}(\mathbb{N}, \mathbb{R})$ の部分集合であるこをと、定義通り確かめましょう。

■問 3 の解答　実数列全体のなす線型空間を $V = \mathrm{Map}(\mathbb{N}, \mathbb{R})$ と書く。

(1) 0 に収束する数列の全体を W と書くと、W は部分空間である。実際

- $(a_n)_{n\geq 0}, (b_n)_{n\geq 0} \in W$ とする。$n \to \infty$ のとき $a_n \to 0$, $b_n \to 0$ なので、$a_n + b_n \to 0 + 0 = 0$ となる。よって $(a_n)_{n\geq 0} + (b_n)_{n\geq 0} = (a_n + b_n)_{n\geq 0} \in W$ である。
- $(a_n)_{n\geq 0} \in W$, $\alpha \in \mathbb{R}$ とする。$n \to \infty$ のとき $a_n \to 0$ なので、$\alpha a_n \to \alpha \cdot 0 = 0$ である。よって $\alpha(a_n)_{n\geq 0} = (\alpha a_n)_{n\geq 0} \in W$ となる。

これより W は加法とスカラー倍で閉じている。

(2) 有界な数列の全体を W は部分空間であることを示す。数列 $(a_n)_{n\geq 0}$ が有界であることは「ある正の実数 $M > 0$ が存在し、任意の $n \in \mathbb{N}$ に対して $|a_n| < M$ が成り立つ」と言える*4。

- $(a_n)_{n\geq 0}, (b_n)_{n\geq 0} \in W$ とする。このとき正の実数 $M_1, M_2 > 0$ が存在して、任意の自然数 $n \in \mathbb{N}$ に対し $|a_n| < M_1$, $|b_n| < M_2$ が成り立つようにできる。すると任意の $n \in \mathbb{N}$ に対し、三角不等式より $|a_n + b_n| < |a_n| + |b_n| < M_1 + M_2$ となる。よって数列 $(a_n)_{n\geq 0} + (b_n)_{n\geq 0} = (a_n + b_n)_{n\geq 0}$ は有界で、W の元となる。
- $(a_n)_{n\geq 0} \in W$, $\alpha \in \mathbb{R}$ とする。このとき正の実数 $M > 0$ が存在して、任意の自然数 $n \in \mathbb{N}$ に対して $|a_n| < M$ となる。すると $|\alpha a_n| = |\alpha||a_n| < |\alpha|M$ が成り立つ。よって $\alpha(a_n)_{n\geq 0} = (\alpha a_n)_{n\geq 0}$ も有界で、W の元となる。

---

*3 $\boldsymbol{e}_1 := {}^t(1, 0)$, $\boldsymbol{e}_2 := {}^t(0, 1)$ です。

*4 「有界」の意味を「上に有界または下に有界」と取り違えていた人がいましたが、ふつう有界と言う時は「上に有界かつ下に有界」です。間違えないよう気を付けてください。

この議論をするときは、もちろん「上に有界だから $a_n < M_1$ となる実数 $M_1 \in \mathbb{R}$ が存在し、下に有界だから $-M_2 < a_n$ となる実数 $M_2 \in \mathbb{R}$ が存在する」と書いても間違いではありません。が、一々上と下とを分けて議論するのは面倒なので、$M := \max\{|M_1|, |M_2|\}$ を考えて $|a_n| < M$ とおいた方が楽です。$|a_n| < M$ なら $-M < a_n < M$ だから、上と下とをいっぺんに抑えることができます。

よって W は加法とスカラー倍で閉じている。

(3) 有限個を除く全ての項が 0 である数列の全体を W と書くと、W は部分空間である。これを示すにあたり、数列 $(a_n)_{n\geq 0}$ について「有限個を除いて項が 0」という条件は、「十分大きい自然数 $N\in\mathbb{N}$ が存在して、$n>N$ ならば $a_n=0$ が成り立つ」と言い換えられることに気を付けよう[*5]。

- $(a_n)_{n\geq 0},(b_n)_{n\geq 0}\in W$ とする。このときある自然数 $N_1,N_2\in\mathbb{N}$ を取って、$n>N_1$ ならば $a_n=0$, $n>N_2$ ならば $b_n=0$ となるようにできる。ここで $N:=\max\{N_1,N_2\}$ とおくと、$n>N$ なら $n>N_1$, $n>N_2$ の両方が同時に満たされるので、$a_n=b_n=0$ となる。これより $n>N$ ならば $a_n+b_n=0$ で、$(a_n)_{n\geq 0}+(b_n)_{n\geq 0}=(a_n+b_n)_{n\geq 0}\in W$ と言えた。
- $(a_n)_{n\geq 0}\in W$, $\alpha\in\mathbb{R}$ とする。このときある自然数 $N\in\mathbb{N}$ を取ると、$n>N$ ならば $a_n=0$ となるようにできる。すると $n>N$ ならば $\alpha a_n=0$ である。よって $\alpha(a_n)_{n\geq 0}=(\alpha a_n)_{n\geq 0}\in W$ である。

よって、W は加法とスカラー倍で閉じている。 ■

3 線型写像の核と像

線型写像には、もれなく「核」という定義域の部分空間と「像」という値域の部分空間が付いてきます。線型写像の性質を調べるときに、これらの部分空間が大きな役割を果たすことを、順番に見ていきます。

3.1 線型写像の核

■Ker の定義 V と W を線型空間とし、$f\colon V\to W$ を線型写像とします。このとき線型写像 f の核 (kernel) は

$$\operatorname{Ker} f := \{\boldsymbol{v}\in V \mid f(\boldsymbol{v})=\boldsymbol{0}\}$$

と定義されます。つまり f によって $\boldsymbol{0}$ ベクトルに潰される[*6] V の元全体が f の核です。核が部分空間になることは、前回のプリント 71 ページでチェックしました。問 6 を解いて、核の性質を確かめてみましょう。

■問 6 の解答 V,W を線型空間、$\Psi\colon V\to W$ を線型写像とする。

(1) $\Psi(\boldsymbol{0})=\Psi(\boldsymbol{0}-\boldsymbol{0})=\Psi(\boldsymbol{0})-\Psi(\boldsymbol{0})=\boldsymbol{0}'$ である。

(2) $\Psi(\boldsymbol{u})=\Psi(\boldsymbol{v})$ ならば、Ψ の線型性を使って $\boldsymbol{0}=\Psi(\boldsymbol{u})-\Psi(\boldsymbol{v})=\Psi(\boldsymbol{u}-\boldsymbol{v})$ を得る。よって $\boldsymbol{u}-\boldsymbol{v}\in\operatorname{Ker}\Psi$ である。

(3) $\boldsymbol{a}\in V$, $\boldsymbol{b}=\Psi(\boldsymbol{a})\in W$ とする。このとき $\boldsymbol{v}\in V$ が $\Psi(\boldsymbol{v})=\boldsymbol{b}$ を満たすとすると、$\Psi(\boldsymbol{v}-\boldsymbol{a})=\Psi(\boldsymbol{v})-\Psi(\boldsymbol{a})=\boldsymbol{b}-\boldsymbol{b}=\boldsymbol{0}'$ となる。よって $\boldsymbol{v}-\boldsymbol{a}\in\operatorname{Ker}\Psi$ となるから、$\boldsymbol{v}=\boldsymbol{a}+(\boldsymbol{v}-\boldsymbol{a})$ が求める分解を与える。

(4) $\Psi(\boldsymbol{0})=\boldsymbol{0}'$ である。よって Ψ が単射なら $\Psi(\boldsymbol{v})=\boldsymbol{0}'$ なる $\boldsymbol{v}$ は $\boldsymbol{0}$ しかない。これより $\operatorname{Ker}\Psi=\{\boldsymbol{0}\}$ である。

逆に $\operatorname{Ker}\Psi=\{\boldsymbol{0}\}$ とする。このとき $\boldsymbol{u},\boldsymbol{v}\in V$ が $\Psi(\boldsymbol{u})=\Psi(\boldsymbol{v})$ を満たしたとすると、(2) より $\boldsymbol{u}-\boldsymbol{v}\in\operatorname{Ker}\Psi=\{\boldsymbol{0}\}$ となる。よって $\boldsymbol{u}-\boldsymbol{v}=\boldsymbol{0}$ だから $\boldsymbol{u}=\boldsymbol{v}$ となり、Ψ の単射性が従う。 ■

[*5] 直接「a_n が 0 にならないような n の個数」を考えても OK です。数列 $(a_n)_{n\geq 0}$ に対し、$\operatorname{supp}(a_n)_{n\geq 0}:=\{n\in\mathbb{N}\mid a_n\neq 0\}$ と定めます。つまり $\operatorname{supp}(a_n)_{n\geq 0}$ は、a_n が 0 にならないような自然数 $n\in\mathbb{N}$ 全体の集合です。このとき 2 つの数列 $(a_n)_{n\geq 0}$, $(b_n)_{n\geq 0}$ に対し、$a_n+b_n\neq 0$ が成り立つには $a_n\neq 0$ または $b_n\neq 0$ が必要です。これより $\operatorname{supp}(a_n+b_n)_{n\geq 0}\subset\operatorname{supp}(a_n)_{n\geq 0}\cup\operatorname{supp}(b_n)_{n\geq 0}$ が成り立ちます。よって集合 X に属する元の個数を $\#X$ と書くことにすると

$$\#\operatorname{supp}(a_n+b_n)_{n\geq 0}\leq\#\bigl(\operatorname{supp}(a_n)_{n\geq 0}\cup\operatorname{supp}(b_n)_{n\geq 0}\bigr)\leq\#\operatorname{supp}(a_n)_{n\geq 0}+\#\operatorname{supp}(b_n)_{n\geq 0}<\infty$$

となり、数列 $(a_n+b_n)_{n\geq 0}$ は有限項を除いて 0 だと分かります。

ちなみに、ここで使った supp という記号は “support” の略です。一般に $\mathbb{R}$ 上の函数 f に対し、f の値が 0 でない点を集めた集合の閉包 $\overline{\{x\in\mathbb{R}\mid f(x)\neq 0\}}$ を f の台 (support) といい、$\operatorname{supp} f$ と書きます。数列は $\mathbb{N}$ 上の函数と思えるので、記号を合わせました。

[*6] 「潰れる」という言葉は数学用語ではないですが、こう書くと雰囲気が分かりやすいと思います。また $\boldsymbol{v}\in\operatorname{Ker} f$ のことを “f kills $\boldsymbol{v}$” と表すこともよくあります。ちょっと物騒な言い方ですが、これもこれで言いたいことが伝わると思います。

■線型写像の核と単射性　今の問 6 (4) は非常に重要なことを言っています。そもそも一般に、全ての写像に対して「単射性」という性質が定義されていました。集合 X から集合 Y への写像 $f\colon X \to Y$ が単射であるとは「任意の $x, y \in X$ について、$f(x) = f(y)$ ならば $x = y$ が成り立つ」ことでした。一般には、単射性を示すには定義通りに議論するしかありません。ところが問 6 (4) は、**線型写像の場合は** Ker **を調べれば単射性が分かる**と教えてくれるのです。さらに、この後の授業で「線型空間の次元」や、線型写像の核と像の次元を結びつける「次元公式」を学ぶと、特別な場合に全射性と単射性が連動すること、あるいは単射性 / 全射性が成り立たないことを一瞬で見抜ける状況などが出てきます。そういう意味で Ker は非常に役立つのです。

さらに問 6 (3) は、もっと強力なことを言っています。ふつう「写像 f が単射でない」と言われても「どのように単射でないか」などは一言で言えません。たとえば函数 $y = e^{-x^2} \sin 30x$ を x 軸に平行な直線 $y = c$ で切断したら、断面に現れる点が何個になるかなんて、わけが分かりません。

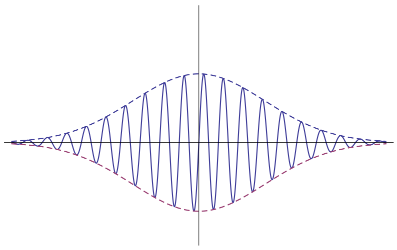


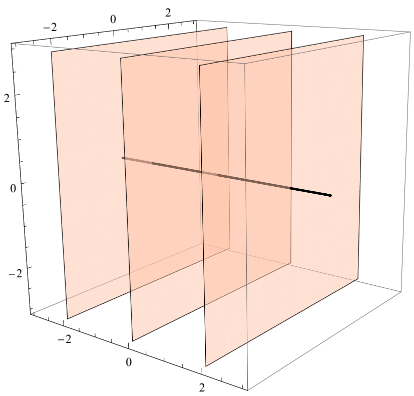
図 8.2　$y = e^{-x^2} \sin 30x$ のグラフ

ところが線型写像の場合、問 6 (2), (3) で示したように

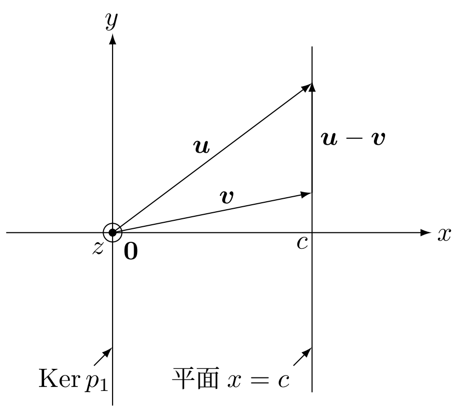
- $\boldsymbol{u}$ と $\boldsymbol{v}$ のズレ $\boldsymbol{u} - \boldsymbol{v}$ が $\operatorname{Ker} f$ に入っていたら、f による値は同じ
- $f(\boldsymbol{u}) = f(\boldsymbol{v})$ のように f による値が一致したら、$\boldsymbol{u}$ と $\boldsymbol{v}$ のズレ $\boldsymbol{u} - \boldsymbol{v}$ は必ず $\operatorname{Ker} f$ に含まれる

と分かります。つまり**線型写像の単射性が破れる具合は、核で完全にコントロールされる**というわけです。

たとえば線型写像 $p_1\colon \mathbb{R}^3 \to \mathbb{R}$ を第 1 成分への射影、つまり $p_1\big({}^t(x, y, z)\big) := x$ と定めます[*7]。すると p_1 はベクトルの x 座標の値だけを見るので、$\operatorname{Ker} p_1$ は平面 $x = 0$、つまり yz 平面だと分かります。そして yz 平面に平行な平面 $x = c$ 上では p_1 の値が一定です。だから $c \in \mathbb{R}$ を 1 つ決めるごとに、p_1 の等位面 $x = c$ が定まります。



(a) x 軸と垂直な p_1 の等位面 $x = c$ が $\operatorname{Ker} p_1$ と平行に並ぶ様子



(b) $\operatorname{Ker} p_1$ を真上から見た図

[*7] p_1 は横ベクトル $(1, 0, 0)$ を左から掛け算する写像なので、ここから直ちに線型写像であることが分かります。

さっき「単射性の破れが Ker でコントロールされる」といったのは、今の場合

- $\boldsymbol{u}-\boldsymbol{v}$ が yz 平面と平行だったら、$\boldsymbol{u}$ と $\boldsymbol{v}$ の x 座標が同じ
- $\boldsymbol{u}$ と $\boldsymbol{v}$ の x 座標が同じだったら、$\boldsymbol{u}-\boldsymbol{v}$ は yz 平面と平行

という意味になります。図で考えれば当たり前ですよね。でも、**この当たり前の状況が線型写像ではいつでも成り立っている**ということが大事です。

さらに言ってしまえば、定義域の線型空間を「Ker の方向」と「それ以外の方向」に分けて考えることができます。そうすると Ker の方向には単射性が破れるわけですから、それ以外の方向を見れば単射になることが推察されます。そうした「空間をいくつかの方向に分ける」という話も、おいおい授業で行います。

3.2 線型写像の像

線型写像を扱うにあたり、核と対をなして重要な役割を果たすのが像と呼ばれるものです。核が線型写像の単射性を司ったのに対し、像の方は線型写像の全射性を司ります。

■Im の定義　一般に写像 $f\colon X\to Y$ に対し、「X の元の f による行き先を全て集めた集合」を f の**像** (image) といい、$\operatorname{Im} f$ と書きます。記号で書けば $\operatorname{Im} f := \{y\in Y \mid \exists x\in X \text{ s.t. } y=f(x)\}$ です。線型写像も写像ですから、一般の場合と全く同じように像が定義されます。ただ線型写像の場合は

- 定義域と値域が共に線型空間
- 線型空間の和やスカラー倍に対し、線型写像は整合的に振る舞う

という良い条件があります。このことを反映し、**線型写像の像は値域の部分空間になります**。それを証明しておきましょう。V, W を線型空間とし、$f\colon V\to W$ を線型写像とします。

- $\boldsymbol{w}_1,\boldsymbol{w}_2\in\operatorname{Im} f$ とする。このとき像の定義から $f(\boldsymbol{v}_1)=\boldsymbol{w}_1,\ f(\boldsymbol{v}_2)=\boldsymbol{w}_2$ となる $\boldsymbol{v}_1,\boldsymbol{v}_2\in V$ が存在する。したがって f の線型性から $\boldsymbol{w}_1+\boldsymbol{w}_2=f(\boldsymbol{v}_1)+f(\boldsymbol{v}_2)=f(\boldsymbol{v}_1+\boldsymbol{v}_2)\in\operatorname{Im} f$ となる。
- $\boldsymbol{w}\in\operatorname{Im} f,\ \alpha\in\mathbb{R}$ とする。このとき像の定義から、$f(\boldsymbol{v})=\boldsymbol{w}$ となる $\boldsymbol{v}\in V$ が存在する。したがって f の線型性から $\alpha\boldsymbol{w}=\alpha f(\boldsymbol{v})=f(\alpha\boldsymbol{v})\in\operatorname{Im} f$ となる。

これで $\operatorname{Im} f$ が W の部分空間になることが示せました。

f が行列 A で表される場合には、$\operatorname{Im} f$ は「A の各列のベクトルが張る空間」になります。これを考えると少し $\operatorname{Im} f$ のイメージが分かり易くなります。ですが、このことは連立 1 次方程式の解の存在問題と関連させた方が話しやすいので、後回しにしたいと思います。今回は少し抽象的な議論に徹します。

■集合の包含と等号　問 4 と問 5 で「集合の等号」の話が出てくるので、念のために補足です。集合 A と B について、$A=B$ の定義は「$A\subset B$ かつ $B\subset A$ であること」です。したがって 2 つの集合が等しいことを示すには、$\subset$ 方向および $\supset$ 方向の包含関係という、**異なる 2 つの命題を示す必要があります**。これに気を付けておいてください。

■問 4 の解答　写像 $\Psi\colon\mathbb{R}[x]\to\mathbb{R}[x]$ を $\Psi\bigl(f(x)\bigr):=(x-\alpha)f(x)$ と定める。また、写像 $\mathrm{ev}_\alpha\colon\mathbb{R}[x]\to\mathbb{R}$ を $\mathrm{ev}_\alpha\bigl(f(x)\bigr):=f(\alpha)$ で定める[*8]。

(1) 次のようにして Ψ が線型写像であることを確かめられる。

- 任意の $f(x),g(x)\in\mathbb{R}[x]$ に対し $\Psi\bigl(f(x)+g(x)\bigr)=(x-\alpha)\bigl(f(x)+g(x)\bigr)=(x-\alpha)f(x)+(x-\alpha)g(x)=\Psi\bigl(f(x)\bigr)+\Psi\bigl(g(x)\bigr)$ である。

[*8] この ev_α は問題文中の Φ です。ev_α という記号の由来は、前回 63 ページの脚注で説明しました。

- 任意の $f(x), g(x) \in \mathbb{R}[x]$ に対し、$\Psi\bigl(f(x)g(x)\bigr) = (x-\alpha)f(x)g(x) = g(x)(x-\alpha)f(x) = g(x)\Psi\bigl(f(x)\bigr)$ となる[*9]。特に $g(x)$ を定数にとって $g(x) = \alpha$ とすれば、$\Psi\bigl(\alpha f(x)\bigr) = \alpha\Psi\bigl(f(x)\bigr)$ となる。

(2) $f(x) \in V$ を任意に取る。このとき $\mathrm{ev}_\alpha\bigl(\Psi(f(x))\bigr) = \mathrm{ev}_\alpha\bigl((x-\alpha)f(x)\bigr) = 0f(\alpha) = 0$ である。

(3) 今の (2) で、$\mathrm{Im}\,\Psi \subset \mathrm{Ker}\,\mathrm{ev}_\alpha$ が示せた。そこで逆向きの包含関係 $\mathrm{Ker}\,\mathrm{ev}_\alpha \subset \mathrm{Im}\,\Psi$ を示せばよい。

$f(x) \in \mathrm{Ker}\,\mathrm{ev}_\alpha$ を任意に取る。このとき $f(\alpha) = 0$ なので、因数定理から $f(x) = (x-\alpha)g(x)$ となる多項式 $g(x) \in \mathbb{R}[x]$ が存在する。よって $f(x) = \Psi\bigl(g(x)\bigr) \in \mathrm{Im}\,\Psi$ である。 ■

■問 5 の解答　写像 $\Phi\colon \mathbb{R}[x] \to \mathbb{R}^2$ を $\Phi\bigl(f(x)\bigr) := {}^t\bigl(f(\alpha), f'(\alpha)\bigr)$ で定める。

(1) Φ の線型性が、次のように確かめられる。

- 任意の多項式 $f(x), g(x) \in \mathbb{R}[x]$ に対し

$$\Phi\bigl(f(x)+g(x)\bigr) = \begin{pmatrix} f(\alpha)+g(\alpha) \\ f'(\alpha)+g'(\alpha) \end{pmatrix} = \begin{pmatrix} f(\alpha) \\ f'(\alpha) \end{pmatrix} + \begin{pmatrix} g(\alpha) \\ g'(\alpha) \end{pmatrix} = \Phi\bigl(f(x)\bigr) + \Phi\bigl(g(x)\bigr)$$

が成り立つ。

- 任意の多項式 $f(x) \in \mathbb{R}[x]$ と実数 $\beta \in \mathbb{R}$ に対し、

$$\Phi\bigl(\beta f(x)\bigr) = \begin{pmatrix} \beta f(\alpha) \\ \beta f'(\alpha) \end{pmatrix} = \beta \begin{pmatrix} f(\alpha) \\ f'(\alpha) \end{pmatrix} = \beta\Phi\bigl(f(x)\bigr)$$

が成り立つ。

(2) 線型写像 $\Psi\colon \mathbb{R}[x] \to \mathbb{R}[x]$ を $\Psi\bigl(f(x)\bigr) := (x-\alpha)^2 f(x)$ で定める[*10]。すると任意の多項式 $f(x) \in \mathbb{R}[x]$ に対し、$\Psi\bigl(f(x)\bigr) = (x-\alpha)^2 f(x)$ だから $\Psi\bigl(f(x)\bigr)' = 2(x-\alpha)f(x) + (x-\alpha)^2 f'(x)$ である。よって $\Phi\bigl(\Psi(f(x))\bigr) = 0$ となるから、$\mathrm{Im}\,\Psi \subset \mathrm{Ker}\,\Phi$ である。

逆に $f(x) \in \mathrm{Ker}\,\Phi$ とする。このとき $f(\alpha) = f'(\alpha) = 0$ である。$f(\alpha) = 0$ より、因数定理から $f(x) = (x-\alpha)g(x)$ となる多項式 $g(x) \in \mathbb{R}[x]$ が取れる。そして $f'(x) = g(x) + (x-\alpha)g'(x)$ の両辺に $x = \alpha$ を代入することで、$0 = g(\alpha)$ を得る。ゆえに再び因数定理により、$g(x) = (x-\alpha)h(x)$ となる多項式 $h(x) \in \mathbb{R}[x]$ が取れる。これを用いると $f(x) = (x-\alpha)^2 h(x) \in \mathrm{Im}\,\Psi$ が分かる。ゆえに $\mathrm{Ker}\,\Phi \subset \mathrm{Im}\,\Psi$ である。以上より、$\mathrm{Ker}\,\Phi = \mathrm{Im}\,\Psi$ である。 ■

■多項式の Taylor 展開　一般に実数 $\alpha \in \mathbb{R}$ を含む開区間で定義された「良い」函数[*11] f に対し、$x = \alpha$ の近くで

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = f(\alpha) + f'(\alpha)(x-\alpha) + \frac{f''(\alpha)}{2!}(x-\alpha)^2 + \cdots$$

という式が成り立ちます。これを函数 f の $x = \alpha$ における **Taylor 展開**といいます。気分としては、与えられた函数を 0 次式、1 次式、2 次式、... で順番に近似しているという感じです。たとえば右辺の第 1 項だけ見ると $y = f(\alpha)$ という定数函数が得られます。もし「$x = \alpha$ で f を定数で近似しろ」と言われたら、$f(a)$ を選ぶ他ないでしょう。次の項まで見て $y = f(\alpha) + f'(\alpha)(x-\alpha)$ という 1 次式を考えると、これは $y = f(x)$ の点 $x = \alpha$ における接線に他なりません。$y = f(x)$ を $x = \alpha$ で 1 次式で近似するとしたら、もちろん接線を考えるのが自然です。こんな感じで右辺を第 n 項まで見ると、f を $x = \alpha$ の周りで最もよく近似する多項式が得られます。そして、この近似を延々と高次の項まで繰り返すことで、いくらでも f を精度よく近似することができます。これが Taylor 展開と呼ばれるものです。

[*9] 線型写像の条件は「<u>実数</u> $\alpha \in \mathbb{R}$ と多項式 $f(x) \in \mathbb{R}[x]$ に対し $\Psi\bigl(\alpha f(x)\bigr) = \alpha\Psi\bigl(f(x)\bigr)$」ですが、今は一段と強く「<u>多項式</u> $g(x) \in \mathbb{R}[x]$ と多項式 $f(x) \in \mathbb{R}[x]$ に対し、$\Psi\bigl(g(x)f(x)\bigr) = g(x)\Psi\bigl(f(x)\bigr)$」が成り立っています。このように Ψ は「多項式に対してもスカラー倍のように振る舞う」ので、これを称して「Ψ は $\mathbb{R}[x]$ **線型**」と言ったりします。

[*10] この問題では直接使いませんが、Ψ は線型写像です。このことは問 4 (1) と全く同様に示せます。

[*11] 細かい条件を書くのが嫌だったので誤魔化しました。多項式、三角函数や指数・対数函数、それにこれらの四則演算や合成で得られる函数は大体「良い」函数になります。ただ、「良い」函数になるための条件は「無限回微分可能」よりも真に強力です。

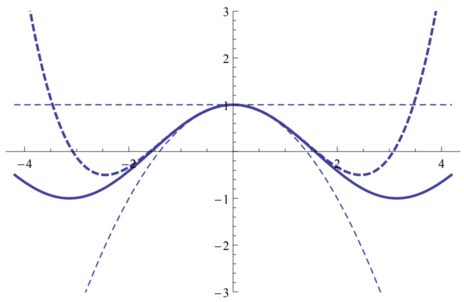
試しに有名な例で計算してみましょう。$f(x) := \cos x$ とおくと、$f'(x) = -\sin x$, $f''(x) = -\cos x$, $f'''(x) = \sin x$, $f''''(x) = \cos x$ なので、f は 4 回微分したら元に戻ります。したがって f の k 回微分 $f^{(k)}(x)$ に $x = 0$ を代入すると、生き残るのは k が偶数の項だけです。そして ± 1 が交互に登場するので、$k = 2l$ とおくと $f^{(2l)}(0) = (-1)^l$ が分かります。これをさっきの Taylor 展開の式に代入すれば、$\cos x$ の $x = 0$ における Taylor 展開

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^l}{(2l)!} x^{2l} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots$$

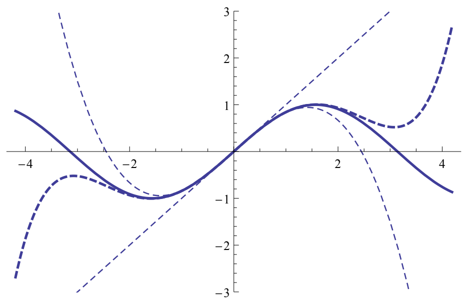
が得られます。同様にして、$\sin x$ の $x = 0$ における Taylor 展開が

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^l}{(2l+1)!} x^{2l+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

と求まります。最初の 3 項までを順番に加えて行ってグラフを描いたのが、次の図です。真のグラフを実線、近似グラフを点線で描き、その上で最も高次の近似グラフを太い点線で表しています。多項式のグラフが三角函数のグラフにどんどん近づいていく様子を、目で見て納得してください。



(c) $\cos x$ の Taylor 展開



(d) $\sin x$ の Taylor 展開

さて Taylor 展開は「函数を多項式で近似する」という話でしたが、そもそも元々の函数が多項式だったらどうなるのでしょうか？ 多項式 $f \in \mathbb{R}[x]$ が n 次式だったら、$f(x)$ は $(n+1)$ 回微分したら 0 になります。ですから今の式は

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(\alpha)}{k!}(x-\alpha)^k = f(\alpha) + f'(\alpha)(x-\alpha) + \frac{f''(\alpha)}{2!}(x-\alpha)^2 + \cdots + \frac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n$$

となり、有限項で打ち止めになります。こうして x の多項式を、$x - \alpha$ の多項式として表す式が得られました[*12]。

これを見れば、問 5 の (2) は自然に解けると思います。というのも $f(x) \in \operatorname{Ker}\Phi$ になるための条件は $f(\alpha) = f'(\alpha) = 0$ と同じです。そうすると Taylor 展開の右辺は $(x-\alpha)^2$ の項から始まります。かたや $f(x)$ が $(x-\alpha)^2$ で割り切れるなら、$f(x) = (x-\alpha)^2 g(x)$ と表せます。そして $g(x)$ の $x = \alpha$ における Taylor 展開に $(x-\alpha)^2$ をかければ $f(x)$ の Taylor 展開が得られるので、$f(x)$ の Taylor 展開はやはり $(x-\alpha)^2$ の項から始まることが分かります。これで $f(\alpha) = f'(\alpha) = 0$ となることが見えます。このように Taylor 展開を通すと、f が $(x-\alpha)^2$ で割り切れることと $f(\alpha) = f'(\alpha) = 0$ とを綺麗に対応づけられます。これで写像 Ψ に当たりがつくわけです。

[*12] ちなみに多項式であれば、Taylor 展開の証明は簡単です。$f(x) \in \mathbb{R}[x]$ を n 次式として、$f(x) = \sum_{k=0}^{n} a_k (x-\alpha)^k$ とおきます。この両辺を m 回微分すれば、$\left\{(x-\alpha)^m\right\}^{(m)} = m!$ より $f^{(m)}(x) = m! a_m + (x-\alpha) \times (\text{多項式})$ という格好の式が得られます。これに $x = \alpha$ を代入すれば $f^{(m)}(\alpha) = m! a_m$ となり、a_m が求まります。

■おまけ: 完全列　微妙に紙面が余ったので、ちょこっとだけ余分なことを書きます。この話は本当におまけなので、知らなくても 1 年生の線型代数の授業では何も困りません。

さて問 4 では、次のような図式が得られています

$$\mathbb{R}[x] \xrightarrow{\Psi} \mathbb{R}[x] \xrightarrow{\Phi} \mathbb{R}$$

そして、この図式の真ん中で $\operatorname{Ker}\Phi = \operatorname{Im}\Psi$ が成り立っているのでした。このことを称して「図式が真ん中の $\mathbb{R}[x]$ のところで**完全**である」といいます。

さらに「0 ベクトルだけを元に持つ線型空間」を 0 と書くと、線型空間 0 から線型空間 $\mathbb{R}[x]$ への線型写像は、0 を 0 へ移すものしかありません。これを用いると $\operatorname{Ker}\Psi$ が 0 ということは、線型写像 $0 \to \mathbb{R}[x]$ の像と $\operatorname{Ker}\Psi$ とが一致することと同じです。つまり $0 \to \mathbb{R}[x] \xrightarrow{\Phi} \mathbb{R}[x]$ が真ん中で完全なことが、単射性と同じ意味になります。

また $\mathbb{R}$ から 0 への線型写像は、全ての実数を 0 に送るものしかありません。この写像 $\mathbb{R} \to 0$ の核は $\mathbb{R}$ 全体ですから、ev_α の全射性[*13]は、$\mathrm{Im}\ \mathrm{ev}_\alpha$ と $\mathbb{R} \to 0$ の核が一致することと言えます。つまり $\mathbb{R}[x] \xrightarrow{\mathrm{ev}_\alpha} \mathbb{R} \to 0$ は完全です。まとめると、**線型写像の単射性と全射性は、完全性を用いて表せます**。

かくして、

$$0 \longrightarrow \mathbb{R}[x] \xrightarrow{\Psi} \mathbb{R}[x] \xrightarrow{\mathrm{ev}_\alpha} \mathbb{R} \longrightarrow 0$$

という、全ての場所が完全であるような図式が得られます。このような図式を**完全列**といいます。特に両端が 0 で 5 項からなる完全列を**短完全列**といいます。後で空間を 2 方向に分解する話をするとき、この短完全列が登場するかもしれません。

[*13] 定数函数 $a \in \mathbb{R}$ に対し $\mathrm{ev}_\alpha(a) = a$ なので、ev_α は全射です。

線形代数学① 第 9 回　　掃き出し法

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 6 月 24 日 1 限

1　過去のプリントについて

前回何人かの人から「過去のプリントはどこか」と尋ねられました。過去のプリントは全て 1 つのファイルにまとめて

`https://github.com/HideakiHosaka/2015_linear_algebra/raw/master/2015linear_algebra.pdf`

に置いてあります。授業を休むなどしてしまった人は、ここからダウンロードしてください。ただファイルサイズが 30MB ほどあるので、容量には気を付けてください。

ちなみに過去に配ったプリントに間違いがあった場合は、オンラインの方を随時修正しています。ですから配られたプリントを読んで「なんかおかしいぞ？」と思ったときは、オンラインで修正されていないか適宜確認してください。そして、オンラインでも修正されていない間違いを見つけたときは、指摘していただけると助かります。ご協力よろしくお願いします。

2　問題の答えと計算のコツ

2.1　解答一覧

最初に答えの一覧をまとめておきます。計算方法さえ分かってしまえば解ける問題ですし、掃き出しの順序は色々あり得るので、解くステップは一々載せません。後で掃き出し法に関連する色々なことを説明する際に、一部の問題を詳しい手順付きで解いてみせます。

なお解が複数ある場合、解の表示を 1 つだけ紹介しています。ここに書かれていない表記法が間違いだというわけではありません。

■問 1　すべて $(x, y) = (-1, -2)$

■問 2　(1) $(x, y) = (0, 4/3)$　(2) $(x, y) = (3, 2)$　(3) $(x, y) = (1, -1)$　(4) $(x, y) = (-3, 4)$

■問 3　(1) $(x, y, z) = (2, -3, -1)$　(2) $(x, y, z) = (-10, 1, 1)$　(3) $(x, y, z) = (0, 1, -2)$

■問 4　(1) $(x, y, z) = (1, 1, -1)$　(2) $(x, y, z) = (2, 3, -1)$　(3) $(x, y, z) = (-11, 0, 1)$

■問 5　(1) z は任意で $(x, y) = (z - 1, -8z + 2)$　(2) 解なし　(3) z は任意で $(x, y) = \big((5z - 6)/7, (11z - 2)/7\big)$

■問 6　(1) $(x, y, z) = (0, 0, 0)$　(2) z, w は任意で $(x, y) = \big((3z - 7w)/5, (11z - 4w)/5\big)$

■問 7　タコとイカの数をそれぞれ x, y とおく。タコは 8 本足、イカは 10 本足の生き物[*1]なので、

$$\begin{cases} x + y &= 8 \\ 8x + 10y &= 70 \end{cases}$$

[*1] 僕たちがイカの「足」だと思っているものは本当は「腕」らしいのですが、そこは見なかったことにします。また生きているイカは 8 本しか足がないように見えますが、残り 2 本は体の中に収まっているんだそうです。「マリンワールド 海の中道」という水族館のページに情報がありました: `http://www.marine-world.co.jp/er/topics/2010_03.html`

という方程式が立つ。これを解いて $x = 5,\ y = 3$ を得る。

■問 8　狐狸と鵬鳥の数をそれぞれ x, y とおく[*2]。狐狸は 1 頭あたり 9 本の尾を、鵬鳥は 1 尾あたり 9 つの頭を持つので

$$\begin{cases} x + 9y &= 72 \\ 9x + y &= 88 \end{cases}$$

という方程式が立つ。これを解いて $x = 9,\ y = 7$ を得る。

■問 9　絹織物の反数を x, 盗人の人数を y とおく[*3]。盗人たちの会話から

$$\begin{cases} x - 8y &= -7 \\ x - 7y &= 8 \end{cases}$$

という方程式が立つ。これを解いて $x = 113,\ y = 15$ を得る。

■問 11　1 番目の方程式から $7y - x = 6$ が得られる。これに $x = ny$ を代入すると $(7 - n)y = 6$ となる。王子の人数と 1 人あたりのダイヤの個数は共に正の整数だから、これで $(7 - n, y)$ の組合せが $(1, 6), (2, 3), (3, 2), (6, 1)$ に限られる。一方 k 番目の方程式は $y = k + \bigl(x - (k-1)y - k\bigr)/7$ と書けるので、この両辺を $\sum_{k=1}^{n}$ すると

$$ny = \sum_{k=1}^{n}\Bigl(k + \frac{x - (k-1)y - k}{7}\Bigr) = \sum_{k=1}^{n}\Bigl(\frac{x+y}{7} + \frac{6-y}{7}k\Bigr) = \frac{n(x+y)}{7} + \frac{6-y}{7}\frac{n(n+1)}{2}$$

となる。この両辺を 7 倍して $x = ny$ を代入すると $7ny = n(n+1)(3 + y/2)$ を得る。さっきの $(7 - n, y)$ の組合せでこの式を満たすものは、$(n, y) = (1, 1), (6, 6)$ だけである。前者の場合 $x = 1$ で、後者の場合 $x = 36$ である。

■問 12　(1) z は任意で $(x, y) = (z + 3, z + 2)$　(2) $(x, y, z) = (1/2, 1/2, 1/2)$　(3) x, y は任意で $z = 3 - x - y$　(4) $(x, y, z) = (1, 2, 4)$　(5) z は任意で $(x, y) = (z + 1, z)$　(6) y は任意で $x = 2y$

■問 13　(1) $(x, y, z) = (0, 0, 0)$　(2) z は任意で $(x, y) = (z, z)$　(3) $(x, y) = (0, 0)$

2.2　問題を解くコツについて

今回の問題の解き方は、基本的にどれも同じです。単に拡大係数行列に対して掃き出しを行えば良いだけの話ですので、答え以上の添削はほとんどしていません。

■検算　ややこしい計算だから当たり前ですが、結構計算ミスをする人がいました。しかしミス自体は一定の確率で起きてしまうものの、**解を元の方程式に代入すれば、正しいかどうかが判断できます**。これで答えを確かめるとよいでしょう。

■間違えたときは　この手の問題で計算ミスをすると、後から修正するのはかなり困難です。頑張って掃き出しを間違えた場所を探すより、**潔く諦めて最初から計算し直す**方が間違えずに済むと思います。ですので答えが間違っていた人は、最初から注意深くやり直してみましょう。

[*2] この問題の出典となった程大位『算法統宗』第 4 巻の 90 ページに、狐狸と鵬鳥の問題が乗っています。写真がオンラインで公開されていますので、気になる人は見てください: `http://edb.kulib.kyoto-u.ac.jp/exhibit/wa2/wa2cont.html`

[*3] この問題の出典となった吉田光由『塵劫記』の初版は 1627 年に出版されています。しかし佐藤健一『『塵劫記』初版本 —影印、現代文字、そして現代語訳—』(研成社) 237〜238 ページによれば、何度かの改訂の後、1641 年の版で「絹盗人を知る算」が初登場するそうです。また塵劫記は、1645 年版を底本として「岩波文庫 青 24-1」で再版されています。これには 222 ページに「きぬぬす人を知る事」と書かれています。

3　掃き出し法

今回のテーマは、連立 1 次方程式です。先週までと比べて急に易しくなった気がしますが、でも「変数の個数と方程式の本数が一般の場合」を考えるので「こどものお遊び」などと侮ってはいけません。何より、ここでの具体的な計算が、後々抽象論を展開するにも役立つのです。

3.1　連立 1 次方程式の解法再考

まずは連立 1 次方程式の中でも最も易しい、2 変数の 2 本の連立 1 次方程式を解いてみましょう。繰り返しになりますが、「こどものお遊び」感がしても侮ってはいけません。問題が易しい分、観察の方をじっくりと読んでください。

■問 1 (3) の解答　上の式を「式 1」、下の式を「式 2」と書くことにする。

$$\begin{cases} x+2y=3 \\ 2x+3y=4 \end{cases} \xrightarrow[(-2)\text{ 倍を加える}]{\text{式 2 に式 1 の}} \begin{cases} x+2y=\ \ 3 \\ \quad -\ y=-2 \end{cases} \xrightarrow[\text{2 倍を加える}]{\text{式 1 に式 2 の}} \begin{cases} x \qquad =-1 \\ \quad -y=-2 \end{cases} \xrightarrow[(-1)\text{ 倍する}]{\text{式 2 を}} \begin{cases} x \qquad =-1 \\ \qquad y=\ \ 2 \end{cases}$$

よって解は $x=-1,\ y=2$ である。 ■

■係数行列の考え方　ここで、今の方程式の解き方を落ち着いて眺めると「実は x とか y とか書く必要なくて、単に係数だけ書き出せばいいんじゃないの？」という気がしてきます。実際その通りで、連立 1 次方程式を解くのに式全体を一々書く必要はありません。また、いまの方程式は行列を用いて

$$\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$$

と書けます。この式を $A\boldsymbol{x}=\boldsymbol{b}$ と表せば、「係数だけ書き出す」という操作は、A と $\boldsymbol{b}$ とを横に並べてできる行列 $\tilde{A} := (A \mid \boldsymbol{b})$ を取り出す操作に相当します。

一般の場合も、連立 1 次方程式は必ず $A\boldsymbol{x}=\boldsymbol{b}$ と書けます。この方程式 $A\boldsymbol{x}=\boldsymbol{b}$ に対し、A を**係数行列**といいます。また A と $\boldsymbol{b}$ とを横に並べてできる行列 $\tilde{A} := (A\ \boldsymbol{b})$ を**拡大係数行列**と言います。ちょっとくどいですが、方程式を解く操作と拡大係数行列の変形とを並べて書き、「拡大係数行列だけ見ればよい」という事実を確認しましょう。もう 1 つ問題を解きます。

■問 1 (2) の解答　次のようにして、$x=-1,\ y=2$ が解と分かる[*4]。

$$\begin{cases} x+2y=3 \\ 3x+4y=5 \end{cases} \xrightarrow[(-3)\text{ 倍を加える}]{\text{式 2 に式 1 の}} \begin{cases} x+2y=\ \ 3 \\ \ \ -2y=-4 \end{cases} \xrightarrow[\text{1 倍を加える}]{\text{式 1 に式 2 の}} \begin{cases} x+ \quad =-1 \\ \ \ -2y=-4 \end{cases} \xrightarrow[(-1/2)\text{ 倍する}]{\text{式 2 を}} \begin{cases} x \qquad =-1 \\ \qquad y=\ \ 2 \end{cases}$$

$$\left(\begin{array}{cc|c} 1 & 2 & 3 \\ 3 & 4 & 5 \end{array}\right) \xrightarrow[(-3)\text{ 倍を加える}]{\text{2 行目に 1 行目の}} \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & -2 & -4 \end{array}\right) \xrightarrow[\text{1 倍を加える}]{\text{1 行目に 2 行目の}} \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & -2 & -4 \end{array}\right) \xrightarrow[(-1/2)\text{ 倍する}]{\text{2 行目を}} \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 2 \end{array}\right)$$

■

■掃き出しと枢軸選択　今の方程式を解く操作は、どちらも

- 式 1 の定数倍を式 2 に加えて、式 2 から x を消去する
- 式 2 の定数倍を式 1 に加えて、式 1 から y を消去する

という順番で解いていました。もっと変数が多い場合でも、方程式が解を持つときは大体同じようにして解けます。つまり係数行列の側で見た時に

[*4] 拡大係数行列に縦線を入れたのは、見易さのためだけです。別に書かなくても構いません。

- 1 行目を適当に定数倍したものを他の行に加え、1 行目以外の第 1 列の成分を全て 0 にする
- 2 行目を適当に定数倍したものを他の行に加え、2 行目以外の第 2 列の成分を全て 0 にする
- ……

といった操作を繰り返すわけです。こうして拡大係数行列の右端以外の行に 1 つだけ 1 が残れば、1 つの変数について、解が求まったことになります。この「第 i 行の定数倍を加え、i 行目以外の j 列目の成分を全て 0 にする」という操作を、「(i, j) **成分を要 (pivot) にした他の行の掃き出し**」といいます。連立 1 次方程式を解くには、1 行目から順番に掃き出しをしていけば OK です。

ただし、たまに「$(1, 1)$ 成分を要にして他の行を掃き出したら、$(2, 2)$ 成分も消えてしまった」なんて現象が起きます。要とする成分が 0 では掃き出しができないので、こういうときは

- 行の入れ替えをして、0 でない成分を次の行に持ってくる[*5]
- 0 でない成分のある行を、次の要に選択する

といういずれかの操作をする必要があります。この操作を**枢軸選択 (pivotting)** といいます。枢軸選択の絡む方程式も、1 つ解いてみましょう。

■問 4 (1) の解答　次のようにして、解が $(x, y, z) = (1, 1, -1)$ と分かる。

$$\left(\begin{array}{ccc|c} 1 & -2 & -4 & 3 \\ 2 & -4 & -7 & 5 \\ 4 & -7 & -13 & 10 \end{array}\right) \xrightarrow[2,3\ \text{行目を掃き出す}]{(1,1)\ \text{成分を要に}} \left(\begin{array}{ccc|c} 1 & -2 & -4 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 3 & -2 \end{array}\right) \xrightarrow[\text{入れ替える}]{2\ \text{行目と}\ 3\ \text{行目を}} \left(\begin{array}{ccc|c} 1 & -2 & -4 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

$$\xrightarrow[1\ \text{行目を掃き出す}]{(2,2)\ \text{成分を要に}} \left(\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 \end{array}\right) \xrightarrow[\text{他の行を掃き出す}]{(3,3)\ \text{成分を要に}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right)$$

■

■行基本変形　ここまでで、連立 1 次方程式を解くのに必要な操作が全て出揃いました。係数行列だけを見て解くのにあたり、使っているのは

- 2 つの行を入れ替える
- ある行に、別の行の定数倍を加える
- ある行に 0 でない定数をかける

という操作の 3 つです。後で示すように、連立 1 次方程式が解けるときは、必ずこれらの組合せで事足ります。そこで、これらの 3 つを**行基本変形**といいます。

3.2　行基本変形の行列による実現

さて、実は**行基本変形は実は左からの行列の掛け算で実現できます**。それを問題で確認しましょう。ちょっとくどいですが「掃き出す操作」を行基本変形 1 つ 1 つに分解して、どのような行列の掛け算をすれば良いのか記してみます[*6]。

[*5] 連立 1 次方程式の側で見れば、方程式の並ぶ順番を入れ替えるだけです。この操作が解に影響しないことは明らかでしょう。

[*6] ちなみに、行基本変形で登場する行列を右からかけると、列基本変形をすることができます。今回は使いませんが、後で行列の階数の性質を詳しく調べるときに使います。

■問 5 (3) **の解答**　まずは掃き出しを行う。

$$
\begin{pmatrix} 2 & 1 & -3 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{pmatrix}
\xrightarrow[\text{2 倍を加える}]{\text{1 行目に 3 行目の}}
\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 2 & 1 & -3 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{pmatrix}
= \begin{pmatrix} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{pmatrix}
$$

$$
\xrightarrow[\text{(−1) 倍する}]{\text{3 行目を}}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}
\begin{pmatrix} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{pmatrix}
= \begin{pmatrix} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ 1 & -3 & 4 & 0 \end{pmatrix}
$$

$$
\xrightarrow[\text{3 倍を加える}]{\text{2 行目に 3 行目の}}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ 1 & -3 & 4 & 0 \end{pmatrix}
= \begin{pmatrix} 0 & 7 & -11 & -2 \\ 0 & -7 & 11 & 2 \\ 1 & -3 & 4 & 0 \end{pmatrix}
$$

$$
\xrightarrow[\text{入れ替える}]{\text{1 行目と 3 行目を}}
\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}
\begin{pmatrix} 0 & 7 & -11 & -2 \\ 0 & -7 & 11 & 2 \\ 1 & -3 & 4 & 0 \end{pmatrix}
= \begin{pmatrix} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 7 & -11 & -2 \end{pmatrix}
$$

$$
\xrightarrow[\text{1 倍を加える}]{\text{3 行目に 2 行目の}}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}
\begin{pmatrix} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 7 & -11 & -2 \end{pmatrix}
= \begin{pmatrix} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
\xrightarrow[\text{(−1/7) 倍する}]{\text{2 行目を}}
\begin{pmatrix} 1 & 0 & 0 \\ 0 & -\frac{1}{7} & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 0 & 0 & 0 \end{pmatrix}
= \begin{pmatrix} 1 & -3 & 4 & 0 \\ 0 & 1 & -\frac{11}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

$$
\xrightarrow[\text{3 倍を加える}]{\text{1 行目に 2 行目の}}
\begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}
\begin{pmatrix} 1 & -3 & 4 & 0 \\ 0 & 1 & -\frac{11}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & 0 \end{pmatrix}
= \begin{pmatrix} 1 & 0 & -\frac{5}{7} & -\frac{6}{7} \\ 0 & 1 & -\frac{11}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & 0 \end{pmatrix}
$$

変形後の拡大係数行列を方程式で書き直すと

$$
\begin{aligned} x - \frac{5}{7}z &= -\frac{6}{7} \\ y - \frac{11}{7}z &= -\frac{2}{7} \end{aligned}
$$

となる。よって答えは、z をパラメータとして $x = (5z-6)/7,\ y = (11z-2)/7$ である。 ■

■基本行列　いまの変形を見れば分かると思いますが、一般に、行列の i 行目と j 行目を入れ替えるには

$$
P_{i,j} := \begin{array}{cc}
 & \begin{array}{ccccccccccc} & & & i^{\mathrm{th}} & & & & j^{\mathrm{th}} & & & \end{array} \\
\begin{array}{c} \\ \\ \\ i^{\mathrm{th}} \\ \\ \\ \\ j^{\mathrm{th}} \\ \\ \\ \\ \end{array} &
\begin{pmatrix}
1 & & & & & & & & & & \\
 & \ddots & & & & & & & & & \\
 & & 1 & & & & & & & & \\
 & & & 0 & & & & 1 & & & \\
 & & & & 1 & & & & & & \\
 & & & & & \ddots & & & & & \\
 & & & & & & 1 & & & & \\
 & & & 1 & & & & 0 & & & \\
 & & & & & & & & 1 & & \\
 & & & & & & & & & \ddots & \\
 & & & & & & & & & & 1
\end{pmatrix}
\end{array}
$$

を左からかけます[*7]。また i 行目の c 倍、i 行目に j 行目の c 倍を加える操作は、それぞれ

$$M_i(c) := \begin{array}{cc} & \begin{array}{ccccccc} & & & i^{\text{th}} & & & \end{array} \\ \begin{array}{c} \\ \\ \\ i^{\text{th}} \\ \\ \\ \\ \end{array} & \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 &&&& \\ &&& c &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \end{array}, \quad Q_{i,j}(c) := \begin{array}{cc} & \begin{array}{ccccccc} & & i^{\text{th}} & & j^{\text{th}} & & \end{array} \\ \begin{array}{c} \\ \\ i^{\text{th}} \\ \\ j^{\text{th}} \\ \\ \\ \end{array} & \begin{pmatrix} 1 &&&&&& \\ & \ddots &&&&& \\ && 1 && c && \\ &&& \ddots &&& \\ &&&& 1 && \\ &&&&& \ddots & \\ &&&&&& 1 \end{pmatrix} \end{array}$$

という行列を左からかけることで実現できます。これら 3 種類の行列を**基本行列**[*8]といいます。

3.3　基本行列と逆行列

連立 1 次方程式の中でも特別な場合として、係数行列が正方行列である場合を考えましょう。つまり変数の個数とちょうど同じ本数の方程式がある場合を考えます。このとき方程式を解く操作は「係数行列に行基本変形を施して、単位行列にする」ことに他なりません。そして行基本変形は左からの行列の積で表せますから、これは「係数行列の逆行列を考えていること」と同じになります。これを応用して、行基本変形を用いて逆行列を求めることができます。

■行列のブロック分解　少し今更感があり脇道に逸れますが、後で必要になるので行列のブロック分解の話をしておきます。たとえば 4 つの 2 次正方行列 $A, B, C, D \in \mathrm{Mat}_2(\mathbb{R})$ が与えられたとき、A と B、C と D を対角線上に並べると、大きい 4 次正方行列が作れます。そして

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix} \begin{pmatrix} C & O \\ O & D \end{pmatrix} = \begin{pmatrix} AC & O \\ O & BD \end{pmatrix}$$

という式が成り立ちます。もっと一般に、行列 A と B を $A = (A_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$, $B = (B_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le l}}$ という「ブロック (小さい行列)」に上手く分ける[*9]と、AB の「(i, j) ブロック」が

$$\sum_{k=1}^{n} A_{ik} B_{kj}$$

という形で表せます。**行列の積と全く同じ公式が、ブロック分解された行列の積についても成り立つのです**。典型的な例は「対角線に沿って正方行列が並ぶ場合」ですが、他にも知っておくと、計算の手間が省けることがあるでしょう。実際、すぐ後で使います。

■逆行列の計算法　正方行列 A に対し、その逆行列とは $AA^{-1} = A^{-1}A = E$ となる行列のことでした。逆行列はいつでも存在するとは限らず、逆行列が存在するような正方行列 A は**正則**であるといいます。A が正則な場合、基本変形を用いると、次のようにして逆行列を求めることができます。

いま、A に基本行列 $X_1, X_2, \ldots, X_k$ を順番に左からかけて単位行列になったとします。つまり $X_k X_{k-1} \cdots X_1 A = E$ です。すると $A^{-1} = X_k X_{k-1} \cdots X_1$ だと分かります。ここでこの式の一番右に単位行列を補って $A^{-1} = X_k X_{k-1} \cdots X_1 E$ とします。「各 X_i が行基本変形に対応していた」ことを思い出せば、右辺は「単位行列に対す

[*7] 何もないところには 0 が詰まっています。また左上から右下に並ぶ $1 \cdots 1$ の個数は、時と場合によって変わります。点を並べてるくせに 1 個しかないこともありますし、0 個のこともあります。

[*8] 基本行列を表す記号は、必ずしも P, Q, M が使われるわけではありません。本によって記号はまちまちだと思います。

[*9] ここでいう「上手く分ける」の意味は、「すぐ下の式に出てくるブロックごとの行列の積が定義されるように」という意味です。

る基本変形」を行っているように見えます。つまり「**A を単位行列にしたのと同じ手順の行基本変形を単位行列に施すと、A の逆行列が得られる**」と言っているわけです。こうして逆行列が求まります。

さらに基本変形で逆行列を求めるときに、一々「まず A を行基本変形で単位行列に変形して、その手順通りに単位行列にもう一度行基本変形を施す」なんてする必要はありません。**最初から A の右側に単位行列を並べ、一緒に行基本変形をしてしまえば良いのです**。実際、A と同じサイズの単位行列を並べた行列 $(A \mid E)$ に対して[*10]、A を単位行列にするような基本変形をすれば

$$X_k X_{k-1} \cdots X_1 (A \mid E) = (X_k X_{k-1} \cdots X_1 A \mid X_k X_{k-1} \cdots X_1 E) = (E \mid X_k X_{k-1} \cdots X_1)$$

となります。**A が単位行列に変形されたのと同時に、元々右側にいた単位行列が A^{-1} に変形されるのです**。

この方法が機能することを確かめるため、1 つ問題を解いてみましょう。

■問 4 (2) の解答　係数行列と単位行列を横に並べ、同時に基本変形する。

$$\left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ -2 & 4 & 19 & 0 & 1 & 0 \\ -2 & 5 & 25 & 0 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\text{2, 3 行目を掃き出す}]{\text{(1, 1) 成分を要に}} \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ -2 & 4 & 19 & 0 & 1 & 0 \\ -2 & 5 & 25 & 0 & 0 & 1 \end{array}\right) = \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \end{array}\right)$$

$$\xrightarrow[\text{入れ替える}]{\text{1 行目と 2 行目を}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \end{array}\right) = \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow[\text{1 行目を掃き出す}]{\text{(2, 2) 成分を要に}} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array}\right) = \left(\begin{array}{ccc|ccc} 1 & 0 & 5 & 5 & 0 & 2 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array}\right)$$

$$\xrightarrow[\text{1, 2 行目を掃き出す}]{\text{(3, 3) 成分を要に}} \begin{pmatrix} 1 & 0 & -5 \\ 0 & 1 & -7 \\ 0 & 0 & 1 \end{pmatrix} \left(\begin{array}{ccc|ccc} 1 & 0 & 5 & 5 & 0 & 2 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array}\right) = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & -5 & 2 \\ 0 & 1 & 0 & -12 & -7 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array}\right)$$

こうして係数行列の逆行列が求まるので、これを拡大係数行列にかけて

$$\begin{pmatrix} -5 & -5 & 2 \\ -12 & -7 & 1 \\ 2 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -2 & -9 & 5 \\ -2 & 4 & 19 & -11 \\ -2 & 5 & 25 & -14 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

を得る[*11]。答えは $x = 2,\ y = 3,\ z = -1$ である。 ■

■基本行列の逆行列　2 次正方行列の場合には、正則なことと $\det A \neq 0$ とが同値であることを (大分昔になりますが)「数理科学基礎 (線形代数学) 第 5 回」の配布プリント 45 ページにて直接証明しました。一般の場合も同様に行列式を用いて正則性の判定ができますが、高次の行列式を定義しないと、その手法は使えません。

が、基本行列に関してはこんな一般論をゴタゴタ使わずとも、直に逆行列が計算できてしまいます。基本行列は「行基本変形」と対応します。そして**「行基本変形を元に戻す操作」もまた行基本変形です**。そこで行基本変形を「どう戻せばいいか」を考えれば、基本行列の逆行列に当たりが付きます。

[*10] ここでも縦線は、見易さのために入れているだけです。

[*11] 逆行列が得られたことは理論的には分かっているのですが、これが本当に逆行列になっていることは、一度は直接計算で確かめるべきです。

まず「i 行目と j 行目」の入れ替えは、同じ操作を 2 度繰り返せば元に戻ります。実際、基本行列でも

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

となっています[*12]。次に、i 行目に j 行目の c 倍を加える基本変形の逆は、i 行目に j 行目の $-c$ 倍を加える基本変形です。i 行目に j 行目の c 倍を加えた後 $-c$ 倍を加えたら、元に戻りますよね。それを反映して

$$\begin{pmatrix} 1 & & c \\ & \ddots & \\ & & 1 \end{pmatrix} \begin{pmatrix} 1 & & -c \\ & \ddots & \\ & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix}$$

となります。そして最後に、i 行目を c 倍する基本変形の逆は、i 行目を $1/c$ 倍する操作です。行列でこれを表すと

$$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & \frac{1}{c} & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 1 & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

となります。まとめると

$$P_{k,l}^2 = E, Q_{k,l}(c)Q_{k,l}(-c) = E, M_k(c)M_k(1/c) = E$$

という関係式が得られます。これで全ての基本行列について逆行列が得られました。

3.4 LU 分解

さて、ここで連立 1 次方程式のうち、枢軸選択が要らない問題を考えます。この問題も当然掃き出しでは解けるのですが、掃き出す順序を

1. 1 行目から順に対角線に沿って、下の方だけを先に掃き出し
2. その後でまた 2 行目から順に対角線に沿って、対角線より上の方を掃き出す

という順番で行ってみます。

■問 3 (3) の解答

$$\begin{pmatrix} 1 & 3 & 2 & -1 \\ -3 & -8 & -4 & 0 \\ 2 & 5 & 3 & -1 \end{pmatrix} \xrightarrow[\text{2,3 行目を掃き出す}]{(1,1) \text{ 成分を要に}} \begin{pmatrix} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & -1 & -1 & 1 \end{pmatrix} \xrightarrow[\text{3 行目だけを掃き出す}]{(2,2) \text{ 成分を要に}} \begin{pmatrix} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

これで下半分の掃き出しが済んだ。続けて掃き出しを行うと

$$\begin{pmatrix} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix} \xrightarrow[\text{1 行目を掃き出す}]{(2,2) \text{ 成分を要に}} \begin{pmatrix} 1 & 0 & -4 & 8 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix} \xrightarrow[\text{1,2 行目を掃き出す}]{(3,3) \text{ 成分を要に}} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{pmatrix}$$

となる。 ■

[*12] 元々の基本行列は、いまの行列の左上と右下に単位行列がありました。しかし行列の積をブロック分割して考えれば、左下と右下のブロックでの計算は「単位行列を 2 個かけたら単位行列になる」というだけの結果になります。だから中央のブロックだけ取ってくれば十分です。

■LU **分解**　今の問題を解く基本変形を「下半分の掃き出し」のところで止めてみます。これを基本行列の積で書くと

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 \\ -3 & -8 & -4 \\ 2 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

となります。この両辺に左から、基本変形の逆行列をかけると*13

$$\begin{pmatrix} 1 & 3 & 2 \\ -3 & -8 & -4 \\ 2 & 5 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 3 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$$

というように、元の方程式の係数行列を**下三角行列と上三角行列の積に分解**できます*14。

今の手順を振り返って、何でこのような分解ができたのかを考えてみましょう。右側に上三角行列が出てきた理由は言うまでもなく、わざと行列の下半分だけ掃き出したからです。一方、左側に出てきた行列が下三角行列になるのは

- 「対角線の下側を掃き出す」という操作に登場する基本行列が、下三角行列で表されること
- その時に登場する基本行列の逆行列が、また下三角行列であること
- 下三角行列同士の積が再び下三角行列になること

という理由によるものです。だから一般の正則行列の場合でも、枢軸選択が不要な限りにおいて、「下半分だけを掃き出す」という操作で下三角 (<u>L</u>ower triangular) 行列と上三角 (<u>U</u>pper triangular) 行列への分解が得られます。これを **LU 分解**といいます。また、係数行列の下半分が掃き出せた段階で、方程式は

$$\begin{pmatrix} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix} \leftrightarrow \begin{cases} x + 3y + 2z = -1 \\ y + 2z = -3 \\ z = -2 \end{cases}$$

という格好をしています。ここまで来れば、あとは掃き出すまでもありません。一番下を解けば z が求まり、その z を真ん中に代入すれば y が求まり、y, z を一番上に代入すれば x が求まります。このように LU 分解の後、下から順に方程式を解く手法を **Gauss の消去法**といいます。この手法はコンピュータで連立 1 次方程式を解く際、よく使います。

4　行列の核、像と階数

線型写像に対しては核と像が定義され、核は定義域の、像は値域の部分空間となるのでした。行列の掛け算は線型写像ですから、行列に対しても当然その核や像が定義でき、これらは線型空間になります。

前の章では計算手順等を説明するため、意図的に上手く解ける問題ばかりを選んで解説しました。ですが世の中の連立 1 次方程式は、解をとてもたくさん持つ場合や、逆に解を全く持たない場合もあります。今度は線型写像の核と像を用いて、解の存在状況を解析しましょう。

4.1　行列の像と連立 1 次方程式の解

連立 1 次方程式*15

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned}$$

*13 行列の積は非可換なので、AB の逆行列は $(AB)^{-1} = B^{-1}A^{-1}$ となります。順番に気を付けましょう。
*14 この式は作り方からして正しい式なのですが、直接計算でも正しさを確かめてみてください。
*15 話を簡単にするため変数と方程式の本数を両方とも 3 としていますが、以下の議論は、一般の場合でも全く同様に成り立ちます。

$$a_{31}x_1 + a_{32}x_2 + a_{33}x_3 = b_3$$

を、いつものように行列で $A\boldsymbol{x} = \boldsymbol{b}$ と表しましょう。そしてここで、A の列ベクトルを 1 列目から順に $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ とし、行列 A を $A = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3)$ と表します。そうすると

$$A\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3$$

となります。ここで一旦方程式のことを忘れ、$\boldsymbol{x}$ を 3 次元空間 $\mathbb{R}^3$ の中で動かします。そうしてできる $A\boldsymbol{x}$ 全部を集めてできる集合は、A の像 $\operatorname{Im} A$ に他なりません。一方 $x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + x_3\boldsymbol{a}_3$ で x_1, x_2, x_3 を動かしてできるのは、ベクトル $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ の張る空間です。ですから**行列 A の像は、A の列ベクトルたちが張る空間と一致します**。式で書けば $\operatorname{Im} A = \mathbb{R}\boldsymbol{a}_1 + \mathbb{R}\boldsymbol{a}_2 + \mathbb{R}\boldsymbol{a}_3$ です。

そして方程式 $A\boldsymbol{x} = \boldsymbol{b}$ に話を戻すと、この方程式が解を持つことは、$\boldsymbol{b}$ が A の像に入ることに他なりません。一方 $\operatorname{Im} A = \mathbb{R}\boldsymbol{a}_1 + \mathbb{R}\boldsymbol{a}_2 + \mathbb{R}\boldsymbol{a}_3$ ですから、これらを合わせて考えると「像 $\operatorname{Im} A$ が大きいほど、方程式 $A\boldsymbol{x} = \boldsymbol{b}$ は解を持ちやすくなる」と想像ができます。

また $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ が $\operatorname{Im} A$ を張るときに「無駄」がある場合もあります。もし $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ が 1 次独立でなければ、$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \alpha_3\boldsymbol{a}_3 = \boldsymbol{0}$ となる、全てが 0 ではない実数 $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$ が存在します。これは言い換えれば ${}^t(\alpha_1, \alpha_2, \alpha_3) \in \operatorname{Ker} A$ となります。核の方向にベクトルを動かしても像は変わりませんから、核の方向にパラメータが生じることが想像できるわけです。

4.2 (拡大) 係数行列の階数と解の存在

■行列の階数　行列 A に対し、$\operatorname{Im} A$ がどれだけ大きいかを測るのが「階数」と呼ばれる量です。行列 A を列ベクトルの並びで $A = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_n)$ と表します。このとき列ベクトル $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ のうち、1 次独立な組合せの最大本数を行列 A の**階数** (rank) といい、$\operatorname{rank} A$ で表します。

後で線型空間の「次元」を定義すると、$\operatorname{rank} A$ は線型空間 $\operatorname{Im} A$ の次元と一致します。まだきちんと説明していないものの、「空間内における 1 次独立なベクトルの最大本数が次元である」という文章は、何となく分かるのではないでしょうか。たとえば僕たちの住んでいる空間 $\mathbb{R}^3$ は「3 次元」と言われます。このことが「$\mathbb{R}^3$ 内で 1 次独立なベクトルは 3 本までしか取れない」という事実と対応するのは、直感的には明らかでしょう。ですから $\operatorname{rank} A$ の大小は $\operatorname{Im} A$ の次元の大小と同じで、したがって $\operatorname{Im} A$ がどれだけ広いかを反映しています。

また**ある行列に正則行列を左からかけても、階数は変わりません**。このことを確かめましょう。一般に X を正則な n 次正方行列、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k \in \mathbb{R}^n$ をベクトルとするとき、X は逆行列を持つので、

$$\begin{aligned} \alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k &= \boldsymbol{0} \\ X(\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_k\boldsymbol{v}_k) &= \boldsymbol{0} \end{aligned}$$

という 2 式が同値になります。一方、2 つ目の式は $\alpha_1(X\boldsymbol{v}_1) + \alpha_2(X\boldsymbol{v}_2) + \cdots + \alpha_k(X\boldsymbol{v}_k) = \boldsymbol{0}$ と書き直せます。だから $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_k$ の 1 次独立性と $X\boldsymbol{v}_1, X\boldsymbol{v}_2, \ldots, X\boldsymbol{v}_k$ の 1 次独立性は同値です。そして行列 XA は、A の各列ベクトルをそれぞれ X 倍して得られます。だから「A の列ベクトルの組合せが 1 次独立であるかどうか」は、「同じ位置にある XA の列ベクトルの組合せが 1 次独立であるかどうか」で判定できます。

■連立 1 次方程式の解の存在判定　いま、連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ を考えましょう。拡大係数行列を $\tilde{A} = (A\ \boldsymbol{b})$ とおくと、$\tilde{A}$ は A に列ベクトルを 1 本付け加えただけなので、$\operatorname{rank} \tilde{A} = \operatorname{rank} A$ または $\operatorname{rank} \tilde{A} = \operatorname{rank} A + 1$ のいずれかが成り立ちます。一方 $\boldsymbol{b} \in \operatorname{Im} A$ かそうでないかが、A の列ベクトルたちに $\boldsymbol{b}$ を加えた時、1 次独立なベクトルの最大本数が増えるかどうかと対応しています。したがって $\operatorname{rank} \tilde{A} = \operatorname{rank} A + 1$ なら $\boldsymbol{b} \notin \operatorname{Im} A$ であり、方程式 $A\boldsymbol{x} = \boldsymbol{b}$ が解を持たないと結論付けることができます。逆に $\operatorname{rank} \tilde{A} = \operatorname{rank} A$ なら $\boldsymbol{b} \in \operatorname{Im} A$ で、方程式 $A\boldsymbol{x} = \boldsymbol{b}$ は解を持ちます。

そして連立 1 次方程式を掃き出し法で解くとき、拡大係数行列 $\tilde{A}$ は基本変形されて姿を変えていきます。ですが基本変形は左からの基本行列の積で書けること、そして基本行列が正則なことを合わせると、基本変形の前後で (拡大) 係数行列の階数は変わりないと分かります。だから連立 1 次方程式の解の存在を判定するには、**拡大係数行列を適当に見やすく基本変形し、一番右の列の有無で階数が変化するかを調べれば良い**のです。

この方針で、1 つ問題を解いてみましょう。

■問 5 (2) の解答　拡大係数行列を基本変形すると

$$\begin{pmatrix} 3 & 1 & 5 & -1 \\ 2 & 1 & 6 & 0 \\ 5 & 2 & 11 & 3 \end{pmatrix} \to \begin{pmatrix} 3 & 1 & 5 & -1 \\ 2 & 1 & 6 & 0 \\ 2 & 1 & 6 & 3 \end{pmatrix} \to \begin{pmatrix} 3 & 1 & 5 & -1 \\ 2 & 1 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

となる。ここで係数行列の部分

$$\begin{pmatrix} 3 & 1 & 5 \\ 2 & 1 & 6 \\ 0 & 0 & 0 \end{pmatrix}$$

だけ取り出すと、全ての列ベクトルが空間 $\mathbb{R}^3$ 内の xy 平面に含まれている。そして係数行列の左上 2×2 ブロックの行列式を計算すると 1 になるから、第 1 列と第 2 列の列ベクトルが 1 次独立だと分かる。したがって係数行列の階数は 2 である。

一方、拡大係数行列の一番右の列ベクトル ${}^t(-1, 0, 3)$ は第 3 成分が 0 でない。よって拡大係数行列の第 1 列、第 2 列、第 4 列の列ベクトル

$$\begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 3 \end{pmatrix}$$

たちは 1 次独立である。また空間 $\mathbb{R}^3$ 内で 4 本の 1 次独立なベクトルは取れないから、これで拡大係数行列の階数が 3 と決まる。

以上で、係数行列と拡大係数行列の階数が異なることが分かった。これは $\boldsymbol{b} \notin \operatorname{Im} A$ を意味するので、方程式 $A\boldsymbol{x} = \boldsymbol{b}$ は解を持たない。 ■

4.3　行列の核と解の任意性

今度は解がパラメータを持つ場合を考えましょう。連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ が複数の解を持つとき、そのうち異なる 2 個を $\boldsymbol{x}_1, \boldsymbol{x}_2$ とおくと、$A\boldsymbol{x}_1 = A\boldsymbol{x}_2 = \boldsymbol{b}$ です。したがってベクトルを A 倍する写像は単射でありません。そして A 倍写像の単射でない具合は $\operatorname{Ker} A$ を調べれば分かるのでした。特に $\boldsymbol{b} = \boldsymbol{0}$ の場合、方程式を解くことは $\operatorname{Ker} A$ をことそのものです。この方針で、問題を 1 つ解いてみましょう。

■問 6 (2) の解答　問題の方程式を行列 $A\boldsymbol{x} = \boldsymbol{0}$ と表す。右辺は $\boldsymbol{0}$ ベクトルだから拡大係数行列を考える意味はなく、係数行列だけを変形すれば良い。

$$\begin{pmatrix} 1 & 2 & -5 & 3 \\ 2 & -1 & 1 & 2 \\ 3 & 1 & -4 & 5 \\ 2 & 4 & -10 & 6 \end{pmatrix} \xrightarrow[\text{他の行を掃き出し}]{(1,1)\text{ 成分を要に}} \begin{pmatrix} 1 & 2 & -5 & 3 \\ 0 & -5 & 11 & -4 \\ 0 & -5 & 11 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[2\text{ 行目を引く}]{3\text{ 行目から}} \begin{pmatrix} 1 & 2 & -5 & 3 \\ 0 & -5 & 11 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[(-1/5)\text{ 倍する}]{2\text{ 行目を}} \begin{pmatrix} 1 & 2 & -5 & 3 \\ 0 & 1 & -\frac{11}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[1\text{ 行目を掃き出す}]{(2,2)\text{ 成分を要に}} \begin{pmatrix} 1 & 0 & -\frac{3}{5} & \frac{7}{5} \\ 0 & 1 & -\frac{11}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

変形後の係数行列の第 1 列から第 4 列を、それぞれ $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$, $\boldsymbol{a}_4$ と表す。そうすると

$$\boldsymbol{a}_3 = -\frac{3}{5}\boldsymbol{a}_1 - \frac{11}{5}\boldsymbol{a}_2, \boldsymbol{a}_4 = \frac{7}{5}\boldsymbol{a}_1 + \frac{4}{5}\boldsymbol{a}_2$$

である。よって

$$\boldsymbol{v}_1 := \begin{pmatrix} 3/5 \\ 11/5 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{v}_2 := \begin{pmatrix} -7/5 \\ -4/5 \\ 0 \\ 1 \end{pmatrix} \in \operatorname{Ker} A$$

が従う。これらのベクトルは、第 3, 4 成分を見ることで 1 次独立と分かる。

一方背理法で、$\operatorname{Ker} A$ の中からこれ以上 1 次独立なベクトルを取れないことが示せる。仮にそのようなベクトル $\boldsymbol{v}_3 = {}^t(u_1, u_2, u_3, u_4)$ があったとする。$\boldsymbol{v}_1$ と $\boldsymbol{v}_2$ の 1 次結合を $\boldsymbol{v}_3$ に足しても、$\boldsymbol{v}_1$, $\boldsymbol{v}_2$ と $\boldsymbol{v}_3$ が 1 次独立であることに変わりはない。したがって

$$\boldsymbol{v}_3 - u_3\boldsymbol{v}_1 - u_4\boldsymbol{v}_2 = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} - \begin{pmatrix} 3/5 \\ 11/5 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -7/5 \\ -4/5 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}$$

と $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ も一次独立なベクトルの組である[*16]。しかし、これが $\operatorname{Ker} A$ に属するということは

$$\boldsymbol{0} = A \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix} = *\boldsymbol{a}_1 + *\boldsymbol{a}_2$$

を意味する。$\boldsymbol{a}_1$, $\boldsymbol{a}_2$ の係数の少なくとも一方は 0 でないから、$\boldsymbol{a}_1$ と $\boldsymbol{a}_2$ は 1 次独立ではない。しかし $\boldsymbol{a}_1 = {}^t(1, 0, 0, 0)$, $\boldsymbol{a}_2 = {}^t(0, 1, 0, 0)$ は明らかに 1 次独立なので、これは矛盾している。したがって $\operatorname{Ker} A$ の元は、全て $\boldsymbol{v}_1$ と $\boldsymbol{v}_2$ の 1 次結合で書ける。これで解が

$$z\boldsymbol{v}_1 + w\boldsymbol{v}_2 = \begin{pmatrix} (3z - 7w)/5 \\ (11z - 4w)/5 \\ z \\ w \end{pmatrix} \quad (z, w \in \mathbb{R})$$

と求まった。 ■

■解のパラメータと核の基底の対応　今の問題をよくよく観察すると

- 連立 1 次方程式がパラメータを持つのは、係数行列の列ベクトルたちが 1 次独立でない場合である
- 列ベクトルの 1 次従属な組合せと、$\operatorname{Ker} A$ のベクトルとが $1:1$ に対応する
- $\operatorname{Ker} A$ を無駄なく張るベクトルを集めると、それらの 1 次結合を作るときの係数が解のパラメータになる

ことが分かります。この考え方を推し進めると、核と像がどのように関係するかを詳しく知ることができます。

ただ、そろそろ「基底」という言葉を使わずに説明を書くのが苦しい状況になってきました。次回の演習問題で基底の扱いを学びますので、今回はここで話をやめ、次元に関する話は先送りにしたいと思います。

[*16] 第 1, 2 行目の値は計算できますが、後々の議論において「具体的な値が何になるか」は重要でありません。ただ「どちらかの成分が 0 ではないこと」だけが重要なので「具体的な値はどうでもいい」という意味を込めて、$*$ という記号を書いています。

線形代数学① 第 10 回　　線型空間の基底

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 7 月 1 日 1 限

1　講評

■問題の訂正　問 5 の (2) に間違いがありました。すいません。「この列から $\boldsymbol{f}_i$ を除いてできる列も生成系であるような i」が正しい文章です。

この問題については、問題を正しく訂正した上で取り組んだ人が少しいました。自分で適切に問題を訂正できるのは、大変良いことだと思います。今後も「問題がおかしい」と感じたら、問題を疑う姿勢を忘れないでください。

■証明のしかた　今回は証明問題が多かったせいか、解くのに苦しんでいる人がそこそこいたような気がします。そこで、2 回くらい前のプリントで書いた内容と重複しますが、何点か役立つかもしれないアドバイスをお送りします。

- 大抵の場合、問題はいくつかの「示すべきこと」に分割できるはずです。たとえば今回の「基底であることを示せ」なら「1 次独立であること」と「生成系であること」の両方を示せばよいわけです。また「P と Q が同値であることは「P ならば Q」と「Q ならば P」の 2 つに分割できます。こんな感じに**「何を示すべきか」をまずはっきりさせましょう**。これを間違えると、問題とは全くあさっての方向を向いた議論をする羽目になります。
- **定義にはきちんと従いましょう**。たとえば「与えられたベクトルの組が生成系であること」なら、生成系の定義に従って「全てのベクトルが、与えられたベクトルの組の 1 次結合で表せること」を示せば OK です。逆に定義に従わなかったら、何を言っているのかさっぱり分からなくなります。
- これは証明の仕方ではないですが、**書き上げた証明は自分で一度読んでみましょう**。読んで「あれ？」と思うところがあれば修正をしてください。自分で読んでも分からなければ、おそらく他の誰が読んでも理解できません。もし可能なら近くにいるお友達に読ませてみると、一段と「自分の証明の分かりにくい部分」がはっきりすると思います。

2　基底と次元

今回はいよいよ「線型空間の基底」を定義します。基底を導入すると抽象的な線型空間と数ベクトル空間とを同一視できるようになり、行列に対する諸々の演算の意味が明らかになります。

2.1　ベクトルの 1 次独立性と線型空間の生成系

すぐ後で線型空間の基底は「1 次独立な生成系」と定義されます。この定義を理解するためには「1 次独立性」「生成系」という言葉を理科うする必要があるので、まずはその説明をします。ただ、みなさんの中には「それより何で基底が役立つのか先に知りたい」という人もいると思います。次の節を読んで後でここに戻るのも良いでしょう。

■1 次独立性　既にベクトルの 1 次独立性は何度も使っていますが、一応、復習をしておきましょう。

V を線型空間、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n \in V$ とします。$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ が 1 **次独立**であるとは $\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_n\boldsymbol{v}_n = \boldsymbol{0}$ となる $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ の 1 次結合[*1]が $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ 以外に存在しないことを言います。また 1 次独立でないこと

[*1] $\alpha_1\boldsymbol{v}_1 + \alpha_2\boldsymbol{v}_2 + \cdots + \alpha_n\boldsymbol{v}_n$ の形の式を「$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ の 1 **次結合**」といいます。既に知っていると思いますが、念のため補足します。

を 1 **次従属**といいます。

直感的には 1 次独立性は「それぞれのベクトルが違う向きをしており、空間の張り方に無駄がないこと」と言えます。「張る空間」の記号を使えば、1 次独立性は、任意の $1 \leq i \leq n$ に対し $\boldsymbol{v}_i \notin \mathbb{R}\boldsymbol{v}_1 + \cdots + \widehat{\mathbb{R}\boldsymbol{v}_i} + \cdots + \mathbb{R}\boldsymbol{v}_n$ が成り立つ[*2]ことと同値です。

実際にいくつかの場合で、1 次独立性を確認しましょう。ただし問 10, 問 11 は連立 1 次方程式を解くだけですから、細かい解説は省略します。

■問 10 の解答　(1) 1 次従属 (2) 1 次独立

■問 11 の解答　(1) 1 次従属 (2) 1 次独立 (3) 1 次独立 (4) 1 次独立 (5) 1 次独立 (6) 1 次従属 (7) 1 次従属

■問 11 (6) の補足　後で説明する「線型空間の次元」の議論を使うと、(6) の答えに当たりをつけることができます。

ここに出てくる $(1,1,2)$, $(3,5,8)$, $(13,21,34)$ は、有名な Fibonacci 数列の一部分です。漸化式 $a_{n+2} = a_{n+1} + a_n$ を満たす数列で、初項と第 1 項[*3]を共に $a_0 = a_1 = 1$ とすると、a_2 より先が $(2,3,5,8,13,21,34)$ となります。

ここで、Fibonacci 数列と同じ漸化式を満たす数列全体の集合 $V = \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid a_{n+2} = a_{n+1} + a_n\}$ が線型空間になったことを思い出しましょう。この隣接 2 項間漸化式を満たす数列は最初の 2 項と 1 対 1 に対応するので、$\dim V = 2$ です。一方で $(1,1,2)$, $(3,5,8)$, $(13,21,34)$ はいずれも「第 1 成分 + 第 2 成分 = 第 3 成分」となっているので、後ろの項を付け加えて V の元を作れます。2 次元線型空間 V の中に 3 本の 1 次独立なベクトルを取ることはできないので、これで計算をしないでも、問題の答えが「1 次従属」だと分かります。

■問 17 の解答　(1) $f(x) := a_0 + a_1x + a_2x^2 + a_3x^3$ とおき、f が閉区間 $[0, 2\pi]$ 上で恒等的に 0 だったとする。このとき $f(0) = 0$ より、$a_0 = 0$ が従う。次に開区間 $(0, 2\pi)$ 上で $f'(x) = a_1 + 2a_2x + 3a_3x^2 \equiv 0$ なので、$x \to +0$ の極限を取って $a_1 = 0$ を得る。同様に $f''(x)$, $f'''(x)$ の $x \to +0$ における極限を順番に考えれば、$a_2 = 0$, $a_3 = 0$ を得る。よって $1, x, x^2, x^3$ は 1 次独立である。

(2) $a + b\sin x + c\cos x \equiv 0$ とおく。この両辺を x で微分すると $b\cos x - c\sin x \equiv 0$ が得られる。これに $x = 0, 3\pi/2$ を代入すれば、それぞれ $b = 0$ と $c = 0$ が得られる。これを元の $a + b\sin x + c\cos x \equiv 0$ に代入すれば $a = 0$ を得る。故に $\{1, \sin x, \cos x\}$ は 1 次独立である。

(3) $f(x) = a + be^x + ce^{2x} \equiv 0$ とおく。このとき f とその 2 階までの微分に 0 を代入すると、

$$0 = f(0) = a + b + c, \quad 0 = f'(0) = b + 2c, \quad 0 = f''(0) = b + 4c$$

となる。後ろ 2 つの方程式を解くと $b = c = 0$ が分かる。よって残りも $a = 0$ となる。

(4) $0 < \alpha < \pi$ を固定する。$a\sin x + b\sin(x + \alpha) \equiv 0$ とおく。このとき $x = 0$ を両辺に代入すると $b\sin\alpha = 0$ となる。α の取り方から $\sin\alpha \neq 0$ なので、$b = 0$ となる。よって $a\sin x \equiv 0$ となるので、両辺に $x = \pi/2$ を代入して $a = 0$ を得る。

(5) $f(x) = a + b\sin x + c\cos x + d\sin 2x + e\cos 2x \equiv 0$ とおく[*4]。微分を計算すると $f'(x) = b\cos x - c\sin x + 2d\cos 2x - 2e\sin 2x \equiv 0$ である。よって $0 = f'(0) = b + 2d$, $0 = f'(\pi) = -b + 2d$ である。これを解いて $b = d = 0$ を得るので、$f'(x) = -c\sin x - 2e\sin 2x$ となる。これに $x = 3\pi/2$ を代入すると $0 = f'(3\pi/2) = c$ となる。続けて $x = \pi/4$ を代入すれば $0 = f'(\pi/4) = -2e$ となる。これで b, c, d, e が全て 0 だと分かったので、最後に元の $f(x)$ の式に代入して $a = 0$ を得る。

[*2] ここで $\widehat{\mathbb{R}\boldsymbol{v}_i}$ と書いたのは「$\mathbb{R}\boldsymbol{v}_i$ を除く」という意味です。一々 $\mathbb{R}\boldsymbol{v}_1 + \cdots + \mathbb{R}\boldsymbol{v}_{i-1} + \mathbb{R}\boldsymbol{v}_{i+1} + \cdots + \mathbb{R}\boldsymbol{v}_n$ と書くと長くて大変なので、しばしばこういう記法が用いられます。

[*3] ここでは数列を「0 番目」から数えました。

[*4] e が自然対数の底を表す記号と重複しますが、誤解の余地は無いでしょう。

■生成系　V を線型空間、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n \in V$ とします。V の全ての元が $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ の 1 次結合で表せるとき、$\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ は V を**生成する**といいます。記号で書けば、$V = \mathbb{R}\boldsymbol{v}_1 + \mathbb{R}\boldsymbol{v}_2 + \cdots + \mathbb{R}\boldsymbol{v}_n$ ということですね。

これと 1 次独立性を合わせて、ようやく基底を定義する準備ができました。おまけ的になりますが、この節の最後に問 4 の解答を記しておきましょう。大体当たり前のことなのですが、証明の中で地味に使われます。

■問 4 の解答　V を線型空間、$(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ をその要素の列とする。
(1) $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ が 1 次独立とする。このとき $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ とし、部分列 $(\boldsymbol{f}_{i_1}, \boldsymbol{f}_{i_2}, \ldots, \boldsymbol{f}_{i_k})$ を考える[*5]と、これも 1 次独立となる。実際 $\alpha_1 \boldsymbol{f}_{i_1} + \alpha_2 \boldsymbol{f}_{i_2} + \cdots + \alpha_k \boldsymbol{f}_{i_k} = \boldsymbol{0}$ とおくと、この両辺に $i_1, i_2, \ldots, i_k$ に現れない α_j たちを全て 0 倍して足しても、0 である。そうすると元々の $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ の 1 次独立性から、$\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$ が従う。
(2) $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ が V を生成したとする。このとき勝手な $\boldsymbol{v} \in V$ を足した列 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n, \boldsymbol{v})$ も V を生成する。実際、全てのベクトルは $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ の 1 次結合で表せる。その表式の両辺に $0\boldsymbol{v} = \boldsymbol{0}$ を足した式を考えれば、$(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n, \boldsymbol{v})$ が V を生成することになる。 ■

■問 4 の別解　やっていること自体は同じなのですが、「ベクトルの張る線型空間」や「線型空間の和」の記号を使った証明もしてみます。
(1) $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ が 1 次独立とする。このとき任意の $1 \leq j \leq n$ に対し、$\boldsymbol{f}_j \notin \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\hat{\boldsymbol{f}}_j + \cdots + \mathbb{R}\boldsymbol{f}_j$ である。ここで $1 \leq i_1 < i_2 < \cdots < i_k \leq n$ とし、部分列 $(\boldsymbol{f}_{i_1}, \boldsymbol{f}_{i_2}, \ldots, \boldsymbol{f}_{i_k})$ を考える。すると任意の $1 \leq l \leq k$ に対し $\mathbb{R}\boldsymbol{f}_{i_1} + \cdots + \mathbb{R}\hat{\boldsymbol{f}}_{i_l} + \cdots + \mathbb{R}\boldsymbol{f}_{i_k} \subset \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\hat{\boldsymbol{f}}_{i_l} + \cdots + \mathbb{R}_n$ となるから、$\boldsymbol{f}_{i_l} \notin \mathbb{R}\boldsymbol{f}_{i_1} + \cdots + \mathbb{R}\hat{\boldsymbol{f}}_{i_l} + \cdots + \mathbb{R}\boldsymbol{f}_{i_k}$ でないといけない。これで $(\boldsymbol{f}_{i_1}, \boldsymbol{f}_{i_2}, \ldots, \boldsymbol{f}_{i_k})$ の 1 次独立性が言えた。
(2) $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ が V を生成したとする。このとき $\mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n = V$ である。いま勝手に $\boldsymbol{v} \in V$ を取ると、$\mathbb{R}\boldsymbol{v} \subset V$ である。よって $V = \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n \subset \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n + \mathbb{R}\boldsymbol{v} \subset V$ となり、$V = \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n + \mathbb{R}\boldsymbol{v}$ が従う。よって $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n, \boldsymbol{v})$ も V を生成する。 ■

2.2　基底の定義

■基底の定義　V を線型空間とし、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ を V の元とします。このとき $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ たちが 1 次独立で、かつ V を生成するとき、V の元の列 $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ を V の**基底**といいます。また基底をなすベクトルの本数を V の**次元**といい、$\dim V$ で表します。次元が正しく定まること、つまりどんな基底を持ってきても同じ本数のベクトルから構成されるという事実は、決して自明ではありません。ですがその証明は後回しにして、まずは次元の存在を認めた上で、基底の性質を調べましょう[*6]。

僕たちにとって最も身近な基底の例は、数ベクトル空間 $\mathbb{R}^n$ における「座標軸」の正の方向を向いた単位ベクトル

$$\boldsymbol{e}_1 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \boldsymbol{e}_n := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

の組 $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ です。この基底は $\mathbb{R}^n$ の**標準基底**と呼ばれます。最初に、これがきちんと基底の条件を満たしていることを確認しましょう。

[*5] 「部分列」というときは、元の列からどんな間引き方をすることも許します。だから「最初の n 個だけを取る」という議論では不十分です。この辺の言葉遣いは、微積分をするときに「収束する部分列」と言ったりするのと同じです。

[*6] 念のため補足すると、議論が循環論法に陥らないようプリントを作っています。気になる人は先に 105 ページに行き、次元の一意性の証明を読んできてください。

■問 12 の解答　勝手な $\mathbb{R}^n$ の元は

$$\begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix} = \alpha_1\boldsymbol{e}_1 + \alpha_2\boldsymbol{e}_2 + \cdots + \alpha_n\boldsymbol{e}_n$$

と書けるので、$\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ たちは $\mathbb{R}^n$ を生成する。また

$$\boldsymbol{0} = \alpha_1\boldsymbol{e}_1 + \alpha_2\boldsymbol{e}_2 + \cdots + \alpha_n\boldsymbol{e}_n = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{pmatrix}$$

とおくと、各成分が 0 であることから $\alpha_1 = \cdots = \alpha_n = 0$ が従う。これより $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ たちは 1 次独立である。以上で $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ が $\mathbb{R}^n$ の基底であり、$\dim \mathbb{R}^n = n$ が言えた。 ■

■基底と座標　$\mathbb{R}^n$ の標準基底の例では、$\mathbb{R}^n$ のベクトルを $\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ で表したときの $\boldsymbol{e}_i$ が、i 番目の座標と全く同じでした。一般の場合でも全く状況は同じで、**基底を定めることと座標を入れることは同じ**です。そして**座標はどうあるべきか**を考えることによって、基底の定義に込められた意味が見えてきます。

V を一般の線型空間とし、$(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ をその基底とします。数ベクトル空間の場合を真似して、僕たちは $\boldsymbol{x} \in V$ が

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n$$

を満たすときに「$\boldsymbol{x}$ の座標は ${}^t(x_1, x_2, \ldots, x_n)$ である」と言いたいのです。このように言うためには、まず「全ての V の元に対して、きちんと座標が定まること」が不可欠です。ここで $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ は基底だったから、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ たちは V を張ります。だから全ての $\boldsymbol{x} \in V$ を、$\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ たちの 1 次結合で表せます。これでどんな V の元も、座標で表示できることが分かりました。

ですが、これで十分ではありません。座標と呼ぶからには「異なる座標が同じ点を表してしまう」という事態があってはいけません。こちらの条件はどうでしょうか？ いま $\boldsymbol{x} \in V$ が

$$\boldsymbol{x} = x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 + \cdots + x_n\boldsymbol{v}_n = x'_1\boldsymbol{v}_1 + x'_2\boldsymbol{v}_2 + \cdots + x'_n\boldsymbol{v}_n$$

と、基底 $(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ を用いて 2 通りに表示されたとしましょう。すると第 2 式と第 3 式から

$$(x_1 - x'_1)\boldsymbol{v}_1 + (x_2 - x'_2)\boldsymbol{v}_2 + \cdots + (x_n - x'_n)\boldsymbol{v}_n = \boldsymbol{0}$$

を得ます。ここで $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ は 1 次独立でしたから、$x_1 - x'_1 = x_2 - x'_2 = \cdots = x_n - x'_n = 0$ を得ます。つまり全ての $\boldsymbol{x} \in V$ に対し「$\boldsymbol{x}$ を $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ の 1 次結合で表す式を持ってきたら、それらは全く同じ表式である」と言っているわけです。これは座標の一意性に他なりません。

かくして基底を決めることは、座標を入れることと全く同じであり

- 基底が V を生成することは、全ての V の元に座標を対応させるための条件
- 基底が 1 次独立であることは、座標がただ一通りに定まるための条件

だったのです。このことを用いると、$(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ が基底であるための条件が「全てのベクトルが $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ の 1 次結合で、ただ 1 通りに表せること」とまとめられます。

さらに「座標を入れる」という話は、線型写像の言葉を使うとすっきり書くことができます。

■問 1 の解答　V を線型空間とし、$(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ をその要素の列とする。このとき写像 $\varphi\colon \mathbb{R}^n \to V$ を、$\boldsymbol{x} = {}^t(x_1, x_2, \dots, x_n)$ に対し $\varphi(\boldsymbol{x}) := x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \cdots + x_n\boldsymbol{f}_n$ で定める。
(0) まず φ が線型写像であることを示しておく。

- $\boldsymbol{x} = {}^t(x_1, x_2, \dots, x_n)$, $\boldsymbol{y} = {}^t(y_1, y_2, \dots, y_n)$ に対し

$$\begin{aligned}\varphi(\boldsymbol{x}+\boldsymbol{y}) &= \varphi\big({}^t(x_1+y_1, x_2+y_2, \dots, x_n+y_n)\big) = (x_1+y_1)\boldsymbol{f}_1 + (x_2+y_2)\boldsymbol{f}_2 + \cdots + (x_n+y_n)\boldsymbol{f}_n \\ &= (x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \cdots + x_n\boldsymbol{f}_n) + (y_1\boldsymbol{f}_1 + y_2\boldsymbol{f}_2 + \cdots + y_n\boldsymbol{f}_n) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})\end{aligned}$$

- $\boldsymbol{x} = {}^t(x_1, x_2, \dots, x_n)$, $\alpha \in \mathbb{R}$ に対し

$$\begin{aligned}\varphi(\alpha\boldsymbol{x}) &= \varphi\big((\alpha x_1, \alpha x_2, \dots, \alpha x_n)\big) = \alpha x_1\boldsymbol{f}_1 + \alpha x_2\boldsymbol{f}_2 + \cdots + \alpha x_n\boldsymbol{f}_n \\ &= \alpha(x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \cdots + x_n\boldsymbol{f}_n) = \alpha\varphi(\boldsymbol{x})\end{aligned}$$

(1) $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ が 1 次独立であることは、$\varphi(\boldsymbol{x}) = \boldsymbol{0}$ を満たす $\boldsymbol{x} = {}^t(x_1, x_2, \dots, x_n)$ が $\boldsymbol{x} = \boldsymbol{0}$ 以外に存在しないことと同値である。よって $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ の 1 次独立性は $\operatorname{Ker}\varphi = \{\boldsymbol{0}\}$ と同値である。そして $\operatorname{Ker}\varphi = \{\boldsymbol{0}\}$ と φ の単射性が同値だったので、$(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ の 1 次独立性と φ の単射性が同値になる。
(2) $(\boldsymbol{f}_1, \boldsymbol{f}_2, \dots, \boldsymbol{f}_n)$ は $\operatorname{Im}\varphi$ を生成する。一方 φ の全射性は、$\operatorname{Im}\varphi = V$ と同値である。よって $(\boldsymbol{f}_1, \boldsymbol{f}_2, \dots, \boldsymbol{f}_n)$ が V を生成することと全射性が同値になる。 ■

■座標による同一視　この問 1 から、特に「φ が全単射であることと、$(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ が基底をなすことが同値」と分かります。つまり V **の基底を取って来れば、それに応じて全単射な線型写像** $\varphi\colon \mathbb{R}^n \to V$ **が付いてくる**というわけです。そして φ が全単射ということは $\mathbb{R}^n$ の元と V の元との間に 1 : 1 の対応が付くということです。すなわち φ によって、$\mathbb{R}^n$ の元と V の元とを同一視できます。しかも φ は線型写像ですから、任意の $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ と $\alpha \in \mathbb{R}$ に対し $\varphi(\boldsymbol{x}+\boldsymbol{y}) = \varphi(\boldsymbol{x}) + \varphi(\boldsymbol{y})$, $\varphi(\alpha\boldsymbol{x}) = \alpha\varphi(\boldsymbol{x})$ が成り立ちます。つまり「φ で V に送ってから足し算/スカラー倍をしても、先に $\mathbb{R}^n$ の方で足し算/スカラー倍をして φ に送っても結果が同じ」です。だから今の「φ による $\mathbb{R}^n$ と V の同一視」は、単に 1 : 1 の対応が付くというだけでなく、**加法とスカラー倍の構造まで込めて** $\mathbb{R}^n$ **と** V **とを同一視できる**ということを言っているのです。基底を取ることによって、線型空間としての $\mathbb{R}^n$ と V の構造は全く同じだと分かります。このことを称して「$\mathbb{R}^n$ と V は**同型**である」とか「φ は線型空間の間の**同型写像**である」とかいいます。また $\mathbb{R}^n \simeq V$ とか $\varphi\colon \mathbb{R}^n \xrightarrow{\sim} V$ などと書いたりします[*7]。

■線型写像の行列表示　今度は、2 つの線型空間 V, W とその間の線型写像 $f\colon V \to W$ を考えましょう。そして $\dim V = n$, $\dim W = m$ とします。基底を取ることで、V は $\mathbb{R}^n$ と、W は $\mathbb{R}^m$ とそれぞれ同型になります。その同型を与える写像を $\varphi\colon \mathbb{R}^n \to V$, $\psi\colon \mathbb{R}^m \to W$ とします。そうすると、次のような図式ができます。

$$\begin{tikzcd} V \arrow[r, "f"] & W \\ \mathbb{R}^n \arrow[u, "\varphi"] & \mathbb{R}^m \arrow[u, "\psi"'] \end{tikzcd}$$

ここで ψ は全単射ですから、逆写像 ψ^{-1} が定義できます。そして ψ が線型写像であることを使うと、ψ^{-1} も線型写像であることが示せます。そこで右側にある縦の矢印の向きを反転させると

$$\begin{tikzcd} V \arrow[r, "f"] & W \arrow[d, "\psi^{-1}"] \\ \mathbb{R}^n \arrow[u, "\varphi"] \arrow[r, dashed] & \mathbb{R}^m \end{tikzcd}$$

*7 この辺は記号の亜種が色々あるのですが、大体 $=$ とか $\to$ に $\sim$ がくっついてたら「同型」と思って差支えありません。

という図式が得られます。元々あったのは $f\colon V \to W$ という線型写像でしたが、基底を用いて V, W を数ベクトル空間と同一視すると、$\psi^{-1} \circ f \circ \varphi$ という、$\mathbb{R}^n$ から $\mathbb{R}^m$ への写像が得られます*8。線型写像の合成は再び線型写像になりますから、$\psi^{-1} \circ f \circ \varphi\colon \mathbb{R}^n \to \mathbb{R}^m$ は数ベクトル空間の間の線型写像です。つまりこれは、行列そのものです。そして**線型写像の合成と表現行列の積がぴったり対応します**。最初のあの「よく分からない行列の積の定義」は、線型写像の合成に合うようきめられていたのです。

こんな感じで線型写像 $f\colon V \to W$ が与えられると、定義域と値域の両方を数ベクトル空間と同一視することで、f を行列と同一視することができます。この f のことを線型写像の**行列表示**といいます。「行列が線型写像」という話は既にしましたが、逆に「線型写像は行列で表せる」のです。ですから僕たちは行列に関する諸々の性質を調べましたが、それは「行列相手にしか適用できない話」ではなく「一般の線型写像に対しても適用できる話」をしていたのです*9。

2.3 基底の構成と判定条件

さて、一般に線型空間の基底を作るにはどうすれば良いでしょうか？ この問題に対する処方箋が、問 5, 6, 7 で与えられます。結果を信じるならば、問 5 は

- 最初に $\mathbf{0}$ でないベクトルを取ると、そこに「1 次独立性が保たれるように別のベクトルを適当に足す」という操作は、基底が出来上がるまで続けられる
- 最初に適当な生成系を取ると、「生成系であるという条件を保ったままベクトルを間引いていく」という操作は、基底が出来上がるまで続けられる

と言っているわけです。そして問 6 と問 7 を合わせると

- 1 次独立なベクトルの組について「生成系であること」と「これ以上 1 次独立なベクトルを足せないこと」は同値
- 生成系をなすベクトルの組について「1 次独立であること」と「これ以上間引くと生成系にならないこと」は同値

だと分かります。ですから「1 次独立なベクトルを集める」あるいは「生成系を構成するベクトルを減らす」という操作を「これ以上は無理」というところまで行うことで、基底が得られると分かります。特に線型空間 V の部分空間 $W \subset V$ が与えられたとき「W の基底をまず作って、そこに V のベクトルを継ぎ足して V 全体の基底を作る」という方法で V の基底が作れます。この手法は**基底の延長**と呼ばれ、しばしば活躍します。たとえば基底の延長を考えることで、部分空間 $W \subset V$ がいつも $\dim W \leq \dim V$ を満たすことが分かります。

それでは、問題を解きましょう。

■問 5 の解答　V を線型空間とし、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ をその要素の列とする。

(1) $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が 1 次独立だが、V を生成しないとする。このとき $\mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n$ は V の真部分集合なので、$\boldsymbol{f}_{n+1} \in V$ で $\boldsymbol{f}_{n+1} \notin \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n$ を満たすものが取れる。このとき $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n, \boldsymbol{f}_{n+1})$ は 1 次独立である。実際 $\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n + \alpha_{n+1}\boldsymbol{f}_{n+1} = 0$ とおくと、$\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n = -\alpha_{n+1}\boldsymbol{f}_{n+1}$ である。もし $\alpha_{n+1} \neq 0$ なら $\boldsymbol{f}_{n+1} = (\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n)/(-\alpha_{n+1})$ となってしまい、$\boldsymbol{f}_{n+1}$ の取り方に矛盾する。よって $\alpha_{n+1} = 0$ で $\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n = \mathbf{0}$ が得られる。$\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ は元々 1 次独立だったので、結局 $\alpha_1 = \cdots = \alpha_n = 0$ となる。

(2) $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が V を生成するが、1 次独立でないとする。このとき $\alpha_1 \boldsymbol{f}_1 + \ldots + \alpha_n \boldsymbol{f}_n = \mathbf{0}$ となる実数の組 $(\alpha_1, \ldots, \alpha_n)$ であって、少なくとも 1 つは 0 でないようなものが存在する。ここで、$\alpha_i \neq 0$ となる $1 \leq i$

---

*8 図式の側では「既にある 3 本の矢印を合成して、点線部分の矢印が作れる」ということです。

*9 だから線型代数の話は抽象的な議論で進めることもできるし、基底の定義を済ませてから行列計算によるゴリ押しで進めることもできます。どっちか一方が大事なのではなく両者を対応付けて理解することが大事なのですが、話の順番やウェイトの置き方には色々バリエーションがあります。この辺が線型代数の教科書が巷にあふれる 1 つの原因のような気がします。好みは人によってまちまちだと思うので、皆さんも「自分にとって馴染みやすい理解のしかた」を探してみてください。

を 1 つ選ぶと、$\boldsymbol{f}_i$ を取り除いた列 $(\boldsymbol{f}_1, \ldots, \hat{\boldsymbol{f}}_i, \ldots, \boldsymbol{f}_n)$ も V を生成する[*10]。なぜなら $\alpha_i \neq 0$ より $\boldsymbol{f}_i$ は $\boldsymbol{f}_i = -(\alpha_1 \boldsymbol{f}_1 + \cdots + \widehat{\alpha_i \boldsymbol{f}_i} + \cdots + \alpha_n \boldsymbol{f}_n)/\alpha_i$ を満たす。よって $\boldsymbol{f}_i$ は $\boldsymbol{f}_1, \ldots, \hat{\boldsymbol{f}}_i, \ldots, \boldsymbol{f}_n$ の 1 次結合で書けてしまうから、$\boldsymbol{f}_1, \ldots, \boldsymbol{f}_i, \ldots, \boldsymbol{f}_n$ の 1 次結合は $\boldsymbol{f}_i$ を使わない形で書ける。 ■

■問 6 の解答 V を線型空間とし、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ をその要素の列とする。

(1) $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が 1 次独立で、ここに V のどんな元を付け加えても 1 次独立にならないとする。このとき任意に $\boldsymbol{v} \in V$ を取ると、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n, \boldsymbol{v})$ は 1 次従属だから、$\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n + \alpha_{n+1} \boldsymbol{f}_{n+1} = \boldsymbol{0}$ となる実数の組 $(\alpha_1, \ldots, \alpha_{n+1})$ で、少なくとも 1 つは 0 でないものが存在する。ここでもし $\alpha_{n+1} = 0$ だと $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ の 1 次独立性に反するので、$\alpha_{n+1} \neq 0$ である。よって $\boldsymbol{v} = -(\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n)/\alpha_{n+1}$ となる。これは $\boldsymbol{v} \in \mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n$ に他ならない。$\boldsymbol{v}$ は任意だったので、$\mathbb{R}\boldsymbol{f}_1 + \cdots + \mathbb{R}\boldsymbol{f}_n = V$ である。つまり $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ は V を生成する。

(2) $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ は生成系で、ここからどの元を取り去っても生成系にならないとする。このとき $\alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n = \boldsymbol{0}$ とおく。もし $\alpha_i \neq 0$ となる i が存在すれば、$\boldsymbol{f}_i = -(\alpha_1 \boldsymbol{f}_1 + \cdots + \widehat{\alpha_i \boldsymbol{f}_i} + \cdots + \alpha_n \boldsymbol{f})/\alpha_i$ となり、$\boldsymbol{f}_i$ が他のベクトルの 1 次結合で表せてしまう。したがって $\boldsymbol{f}_i$ を取り除いた列 $(\boldsymbol{f}_1, \ldots, \hat{\boldsymbol{f}}_i, \ldots, \boldsymbol{f}_n)$ が V の生成系となり、矛盾する。よって全ての i について $\alpha_i = 0$ が従うから、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ たちは 1 次独立である。 ■

■問 7 の解答 V を線型空間とし、その要素の列 $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が 1 次独立な生成系とする。

(1) $\boldsymbol{v} \in V$ を $\boldsymbol{0}$ でないベクトルとする。このとき $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ は V の生成系だから、$\boldsymbol{v} = \alpha_1 \boldsymbol{f}_1 + \cdots + \alpha_n \boldsymbol{f}_n$ と表せる。よって $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n, \boldsymbol{v})$ は 1 次独立でない。

(2) $1 \leq i \leq n$ とする。このとき $\boldsymbol{f}_i = \alpha_1 \boldsymbol{f}_1 + \cdots + \widehat{\alpha_i \boldsymbol{f}_i} + \cdots + \alpha_n \boldsymbol{f}_n$ と表せたとすると、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が 1 次独立であることに反する。よって $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ の中からどの $\boldsymbol{f}_i$ を引き抜いても、残りのベクトルの 1 次結合で $\boldsymbol{f}_i$ を表すことはできない。したがって $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ からベクトルを取り除いた部分列は、V を生成しえない。 ■

2.4 基底の順序と空間の向き

線型空間 V に属するベクトルが何本か与えられたとき、それらが 1 次独立であるかどうかとか、あるいは V を張るかどうかとかは、ベクトルたちを並べる順番に全く依存しません。ほとんど当たり前ですが、一応証明しておきます。

■問 2 の解答 V を線型空間とし、$\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n \in V$ とする。また $\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n$ を $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$ の並べ替えとし[*11]、添字の対応を $\boldsymbol{g}_i = \boldsymbol{f}_{\sigma(i)}$ および $\boldsymbol{f}_i = \boldsymbol{g}_{\tau(i)}$ と表す。

いま $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$ たちが 1 次独立だったとする。このとき $a_1 \boldsymbol{g}_1 + \cdots + a_n \boldsymbol{g}_n = \boldsymbol{0}$ とおくと

$$\boldsymbol{0} = a_1 \boldsymbol{g}_1 + \cdots + a_n \boldsymbol{g}_n = a_1 \boldsymbol{f}_{\sigma(1)} + \cdots + a_n \boldsymbol{f}_{\sigma(n)}$$

である。ここで $\sigma(1), \ldots, \sigma(n)$ は $1, \ldots, n$ を並べ替えに過ぎないから、第 3 式は項の順番を並び替えれば $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ の 1 次結合になる。$\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ は 1 次独立だから、その係数についている a_i $(1 \leq i \leq n)$ たちは全て 0 である。これで $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n$ も 1 次独立だと言えた。

また $\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ が V を生成したとする。このとき勝手な $\boldsymbol{x} \in V$ を取ると、$\boldsymbol{x} = x_1 \boldsymbol{f}_1 + \cdots + x_n \boldsymbol{f}_n$ と表せる。すると $\boldsymbol{f}_i$ たちは $\boldsymbol{g}_j$ たちのどれかなので、$\boldsymbol{x} = x_1 \boldsymbol{g}_{\tau(1)} + \cdots + x_n \boldsymbol{g}_{\tau(n)}$ と書ける。$\boldsymbol{x}$ は任意なので、結局全ての $\boldsymbol{x} \in V$ が $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n$ の 1 次結合で表せると言えた。 ■

[*10] ここで $\boldsymbol{f}_i$ の上に ^ をつけて $\hat{\boldsymbol{f}}_i$ と書いたのは「$\boldsymbol{f}_i$ を取る」という意味です。さっきは足し算の中でこの記号を使いましたが、「特定の 1 つを取り除く」という言葉が意味をなす文脈では、所構わず使われます。

[*11] この問題で考えられているのは「ありとあらゆる並べ替え方」です。ですから「順番を正反対にした場合」とか「2 つだけ入れ替えた場合」だけを考えるのでは、議論が不十分です。事実としては「2 つの入れ替れかえで大丈夫なら、どんな入れ替えでも大丈夫」ではあるのですが、このことは証明する必要があります。

■基底の順序　ここで基底の定義を思い出しましょう。基底は「1 次独立」な「生成系」であることでした。ベクトルの順番を並べ替えてもこれらの性質は成り立ちますので、結局のところ「基底は並べ替えても基底である」と分かります。では基底が与えられたとき、その順序は考慮すべきなのでしょうか？

結論から先に言うと、基底は「同じベクトルから構成されていても、並び順が違えば別物である」と考えるべきです。というのも、基底には「座標を定める」という役割がありました。そして基底の順番を並び替えると、それに対応して座標の順番が入れ替わります。ですから「基底の並び順を区別しない」ということは「座標の並び順が入れ替わっても気にしない」という意味になってしまいます。これはさすがに変ですよね。ですから基底を考えるときは、並び順まで込めること、言い換えればベクトルの「組」ではなく「列」を考えるのだということを忘れないでください。

■空間の向き　実ベクトル空間の場合、基底は「空間の向き」を定めます。そして基底を並べる順序を変えると、向きが変わります。

たとえば 2 次元平面 $\mathbb{R}^2$ では、普通「反時計回りが正の向き」という約束をします。これは $\mathbb{R}^2$ の標準基底 $\boldsymbol{e}_1, \boldsymbol{e}_2$ に対し「$\boldsymbol{e}_1$ の方向から $\boldsymbol{e}_2$ の方向を眺める回り方」が正の向きだと言っているわけです。また 3 次元だと、座標系が「右手系か左手系か」という言い回しを良く使います。座標と基底は対応していますから、$\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ をこの順に並べたものが右手系、どれか 2 つの順序を入れ替えたものが左手系だと言っています。

いま紹介したのは「$\mathbb{R}^2, \mathbb{R}^3$ の標準基底」という特別な例なので、これだけで「一般の線型空間に対しても、基底の取り方で向きが定まる」という事実を理解するのは難しいと思います。厳密に定義するには行列式を使い「全ての基底を集めてできる集合を 2 つに分割する」という議論をしなければいけません。こうした話もおいおい紹介しますが、今は何となく「基底の並べ方が向きと関係する」と身近な具体例で感じておいてください。

■問 14 の解答　$\mathbb{C}$ の $\mathbb{R}$ 上の基底として $(1, i)$ が取れる。実際、任意の複素数 $z \in \mathbb{C}$ は $z = a + bi$ と書けるので、$(1, i)$ は $\mathbb{C}$ を生成する。また $a + bi = 0$ なら $a = b = 0$ なので、$(1, i)$ は $\mathbb{R}$ 上 1 次独立である。 ■

■複素平面の向き　基底の並べる順序を決めると、線型空間の「向き」が定まるという話をしました。単に抽象的な線型空間を相手にするだけなら、ここで話は終わってしまいます。ですが相手が複素平面 $\mathbb{C}$ だと話は別で、$\mathbb{C}$ の中には $1, i$ という自然な基底が最初からついてきています。そして 1 と i を「どちらを先に並べるか」といったら、1 を先に並べる方が自然でしょう。こんなわけで、$\mathbb{C}$ には「$\mathbb{R}$ 線型空間としての自然な向き」が最初から定まっています。この $\mathbb{C}$ の向きを使うことで、複素数の数ベクトル空間 $\mathbb{C}^n$ にも最初から実線型空間としての自然な向きが入ります。

3　有名な線型空間における基底の例

さて、基底の抽象論から一旦離れ、具体的な基底の例を見てみましょう。特に線型空間の中でも、行列のなす線型空間や多項式のなす線型空間は、非常によく使われるものです。そこでこれらの空間を例にして、基底を見てみます。

3.1　行列のなす線型空間

■問 13 の解答　まず (m, n) 型行列全体の集合 $\mathrm{Mat}_{m,n}(\mathbb{R})$ が線型空間であることを示す。

- $\mathrm{Mat}_{m,n}(\mathbb{R})$ には、加法が定義されている
  - 加法は各成分毎に定義されており、実数の和は交換可能なので、行列の和も交換可能である
  - ゼロ行列 $O \in \mathrm{Mat}_{m,n}(\mathbb{R})$ がある
  - 任意の行列 $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ に対し、その成分を全て (-1) 倍した行列 $-A$ が、$A + (-A) = (-A) + A = O$ を満たす唯一の元である。
- $\mathrm{Mat}_{m,n}(\mathbb{R})$ には、スカラー倍が定義されている

  - 任意の $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ と $\alpha, \beta \in \mathbb{R}$ に対し、$\alpha(\beta A) = (\alpha\beta)A$ が成り立つ
  - 任意の $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ に対し、$1A = A$ が成り立つ
- 加法とスカラー倍について、分配法則が成り立つ。
  - 任意の $A, B \in \mathrm{Mat}_{m,n}(\mathbb{R})$ と $\alpha \in \mathbb{R}$ に対し、$\alpha(A + B) = \alpha A + \alpha B$ が成り立つ
  - 任意の $A \in \mathrm{Mat}_{m,n}(\mathbb{R})$ と $\alpha, \beta \in \mathbb{R}$ に対し、$(\alpha + \beta)A = \alpha A + \beta A$ が成り立つ

次に、$\mathrm{Mat}_{m,n}(\mathbb{R})$ が有限次元であることを示す。(i, j) 成分だけが 1 で他の成分が全て 0 であるような行列を、E_{ij} と表す。このとき E_{ij} は 1 次独立である。実際

$$O = \sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} a_{ij} E_{ij} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

とおくと、全ての a_{ij} が 0 となる。また、勝手な行列 $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ を取ると

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \sum_{\substack{1 \le i \le m \\ 1 \le j \le n}} a_{ij} E_{ij}$$

は E_{ij} たちの 1 次結合で表されている。よって $\{E_{ij} \mid 1 \le i \le m, 1 \le j \le n\}$ の元を適当な順序で並べたものは、$\mathrm{Mat}_{m,n}(\mathbb{R})$ を生成する。これで $\mathrm{Mat}_{m,n}(\mathbb{R})$ が mn 次元の線型空間であると分かった[*12]。 ■

■問 16 の解答　(1) n 次実対称行列全体の集合を $\mathrm{Sym}_n(\mathbb{R})$ と書く[*13]と、これは $\mathrm{Mat}_n(\mathbb{R}) := \mathrm{Mat}_{n,n}(\mathbb{R})$ の部分空間である[*14]。実際

- $A, B \in \mathrm{Sym}_n(\mathbb{R})$ のとき、${}^t(A + B) = {}^tA + {}^tB = A + B$ より $A + B \in \mathrm{Sym}_n(\mathbb{R})$
- $A \in \mathrm{Sym}_n(\mathbb{R})$, $\alpha \in \mathbb{R}$ のとき、${}^t(\alpha A) = \alpha {}^tA = \alpha A$ より $\alpha A \in \mathrm{Sym}_n(\mathbb{R})$

となっている。そして $\mathrm{Mat}_n(\mathbb{R})$ は n^2 次元の有限次元線型空間だから、その部分空間 $\mathrm{Sym}_n(\mathbb{R})$ も有限次元である。また一般の対称行列は

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} = \sum_{i=1}^{n} a_{ii} E_{ii} + \sum_{1 \le i < j \le n} (E_{ij} + E_{ji})$$

の形に 1 通りに表せる。よって基底として $\{E_{ii} \mid 1 \le i \le n\} \cup \{E_{ij} + E_{ji} \mid 1 \le i < j \le n\}$ を適当な順に並べたものが取れる。これより $\mathrm{Sym}_n(\mathbb{R})$ は $n + n(n-1)/2 = n(n+1)/2$ 次元だと分かる。

(2) n 次実交代行列全体の集合を $\mathrm{Alt}_n(\mathbb{R})$ と書く[*15]と、これは $\mathrm{Mat}_n(\mathbb{R})$ の部分空間である。実際

- $A, B \in \mathrm{Alt}_n(\mathbb{R})$ のとき、${}^t(A + B) = {}^tA + {}^tB = -A - B = -(A + B)$ より $A + B \in \mathrm{Sym}_n(\mathbb{R})$
- $A \in \mathrm{Alt}_n(\mathbb{R})$, $\alpha \in \mathbb{R}$ のとき、${}^t(\alpha A) = \alpha {}^tA = \alpha A$ より $\alpha A \in \mathrm{Alt}_n(\mathbb{R})$

[*12] この証明は、$\mathbb{R}^n$ が n 次元線型空間であることの証明と全く同じです。行列は成分が長方形の形に並んでいるものの、線型空間としての構造は、所詮 mn 個の数を成分毎に足したりスカラー倍するだけに過ぎません。だから線型空間としては $\mathrm{Mat}_{m,n}(\mathbb{R}) \simeq \mathbb{R}^{mn}$ です。

[*13] 対称行列のことを "symmetric matrix" といいます。$\mathrm{Sym}_n(\mathbb{R})$ の記号はここから取りました。

[*14] 正方行列のサイズを指定するには 1 つの自然数で事足りるので、しばしば $\mathrm{Mat}_{n,n}(\mathbb{R})$ のことを $\mathrm{Mat}_n(\mathbb{R})$ と略記します。

[*15] 交代行列のことを "alternating matrix" といいます。$\mathrm{Alt}_n(\mathbb{R})$ の記号はここから取りました。

となっている。そして $\mathrm{Mat}_n(\mathbb{R})$ は n^2 次元の有限次元線型空間だから、その部分空間 $\mathrm{Alt}_n(\mathbb{R})$ も有限次元である。また、全ての交代行列は

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix} = \sum_{1 \le i < j \le n} (E_{ij} - E_{ji})$$

の形に 1 通りに表せる。よって具体的な基底として $\{E_{ij} - E_{ji} \mid 1 \le i < j \le n\}$ を適当な順に並べたものが取れる。よって $\mathrm{Alt}_n(\mathbb{R})$ は $n(n-1)/2$ 次元である。 ∎

■直和分解　一般に、全ての n 次正方行列は対称行列と交代行列の和に分解できます。実際

$$X = \frac{X + {}^tX}{2} + \frac{X - {}^tX}{2}$$

という式がいつでも成り立ち、この第 1 項は対称行列に、第 2 項は交代行列になっています。そしてこのような分解は一意的です。実際 $X = S + A = S' + A'$ を、どちらも X の対称行列と交代行列への分解とします。そうすると $S - S' = A' - A$ が成り立ちます。この左辺は対称行列、右辺は交代行列なので、$A' - A = S - S' = {}^t(S - S') = {}^t(A' - A) = -A' - A$ となります。これより $A' - A = O$ となり、$A' = A$ が従います。このことから $S - S' = O$ となるので、$S = S'$ となります。

つまり今の状況では、線型空間 $\mathrm{Mat}_n(\mathbb{R})$ の元を、その部分空間 $\mathrm{Sym}_n(\mathbb{R})$ と $\mathrm{Alt}_n(\mathbb{R})$ の和に一意的に分解できるようになっています。次元を勘定しても、$n^2 = n(n+1)/2 + n(n-1)/2$ なのでつじつまが合います。この状況を指して「$\mathrm{Mat}_n(\mathbb{R})$ は、$\mathrm{Sym}_n(\mathbb{R})$ と $\mathrm{Alt}_n(\mathbb{R})$ の直和である」といい、$\mathrm{Mat}_n(\mathbb{R}) = \mathrm{Sym}_n(\mathbb{R}) \oplus \mathrm{Alt}_n(\mathbb{R})$ と書きます。

基底を取ることで線型空間はいくつかの方向に分解できますが、この直和の話は「基底を取ることなく、空間を色々な方向に分けることを記述する」という雰囲気で使われます。必要になった段階で、改めて詳しく紹介します。

3.2　多項式の空間

多項式の空間 $\mathbb{R}[x]$ も、僕たちにとってなじみのある線型空間です。これの基底も調べてみましょう。

■問 15 の解答　自然数 $n \in \mathbb{N}$ に対し、n 次以下の多項式全体のなす集合を $\mathbb{R}[x]_{\le n}$ とする。全ての多項式の集合 $\mathbb{R}[x]$ は線型空間である。そして

- $f(x), g(x) \in \mathbb{R}[x]$ が n 次以下なら、$f(x) + g(x)$ も n 次以下[*16]
- $f(x) \in \mathbb{R}[x]$, $\alpha \in \mathbb{R}$ なら、$\alpha f(x)$ も n 次以下

となっているので、$\mathbb{R}[x]_{\le n}$ は和とスカラー倍で閉じている。よって $\mathbb{R}[x]_{\le n}$ は $\mathbb{R}[x]$ の部分空間である。

$\mathbb{R}[x]_{\le n}$ の基底としては $(1, x, x^2, \ldots, x^n)$ が取れる。実際多項式の定義から、全ての多項式は $1, x, x^2, \ldots, x^n$ の 1 次結合でただ一通りに書ける[*17]。よって $\mathbb{R}[x]_{\le n}$ は $n + 1$ 次元である。 ∎

■全ての多項式のなす線型空間　今の問 15 では次数を区切り、$\mathbb{R}[x]_{\le n}$ が $(n+1)$ 次元線型空間であることを示しました。が、次数を区切らずとも、全ての多項式のなす線型空間 $\mathbb{R}[x]$ の基底として $(1, x, x^2, \ldots)$ が取れることは同じです。実際、全ての多項式は $1, x, x^2, \ldots$ の線型結合で、ただ 1 通りに表せます。

ですが多項式の次数はいくらでも高くすることができますから、$1, x, x^2, \ldots$ のうち有限個を持ってきても $\mathbb{R}[x]$ を生成できないことは明らかでしょう。つまり $\mathbb{R}[x]$ は、無限次元の線型空間になっています。他にも無限次元の線型空間

[*16] ただし最高次の項が打ち消し合って、$f + g$ の次数が下がることはあり得ます。ですから「n 次多項式だけを集めた集合」だと、線型空間になりません。

[*17] 「ただ一通り」の部分が、1 次独立性と対応しています。

は色々ありますが、$\mathbb{R}[x]$ はその中でも最もなじみがあるものの 1 つです。

4　次元と線型写像

これまで余り深く考えずに「次元」という言葉を使ってきましたが、実はこんな問題が残っています。「基底の取り方を変えたら、実は本数が変わったりしないだろうか？」もし基底の取り方で本数が変わってしまうなら「線型空間の次元」という言葉は全く意味を失います。

もちろん実際にはそんなことはないのですが、この事実はきちんと証明しなければいけません。

4.1　次元の一意性

「線型空間の基底の本数」を次元と呼びたいのですから、次元の一意性を議論するには「2 つの異なる基底がいつも同じ本数であるか」を調べる必要があります。2 つの基底を比較して、その上で次元の一意性を証明しましょう。

■基底の変換行列　V を線型空間とし、$(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_m)$ と $(\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n)$ が共に基底だったとします[*18]。このとき各 $\boldsymbol{f}_i$ $(1 \leq i \leq m)$ は基底 $(\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_n)$ を用いて

$$\boldsymbol{g}_i = a_{1i}\boldsymbol{f}_1 + a_{2i}\boldsymbol{f}_2 + \cdots + a_{ni}\boldsymbol{f}_n \quad (a_{1i}, a_{2i} \ldots, a_{ni} \in \mathbb{R})$$

と一意的に表せます。これを $i = 1, 2, \ldots, n$ と並べてみると

$$\begin{aligned} \boldsymbol{g}_1 &= a_{11}\boldsymbol{f}_1 + a_{21}\boldsymbol{f}_2 + \cdots + a_{m1}\boldsymbol{f}_m \\ \boldsymbol{g}_2 &= a_{12}\boldsymbol{f}_1 + a_{22}\boldsymbol{f}_2 + \cdots + a_{m2}\boldsymbol{f}_m \\ &\vdots \\ \boldsymbol{g}_n &= a_{1m}\boldsymbol{f}_1 + a_{2m}\boldsymbol{f}_2 + \cdots + a_{mn}\boldsymbol{f}_m \end{aligned}$$

となります[*19]。連立 1 次方程式を行列で表したときのことを思い出すと、今の式も行列を使えばまとめて書けそうですね。実際、記号の濫用ですが「ベクトルを並べたベクトル」に対しても行列の掛け算を同様に定義すると

$$\begin{pmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_m \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

となっています。かくして、2 つの基底の関係は $(\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n) = (\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m)A$ と、行列 A を用いて表せることがあります。この A を**基底の変換行列**といいます。

■次元の一意性　ここで、$(\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m)$ から $(\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n)$ への基底の変換行列と、その逆向きの $(\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n)$ から $(\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m)$ への基底の変換行列両方を考えます。つまり

$$(\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n) = (\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m)A, \quad (\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m) = (\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n)B$$

とおきます。そうすると、基底を 2 度変換することで

$$(\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m) = (\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n)B = (\boldsymbol{f}_1\ \boldsymbol{f}_2\ \cdots\ \boldsymbol{f}_m)AB$$

[*18] まだ $m = n$ かどうかは分かりません。注意しましょう。

[*19] 係数の添字が、いつもの「行列を書くときの順番」と異なるので注意してください。普段僕たちは座標 (を並べた列ベクトル) に行列を当てていますが、ここでは座標ではなく基底の側に行列を当てるため、係数の並び方が変化します。

が従います。これは「基底 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ を基底 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ で表す式」です。が基底の定義から明らかに、その表し方は $\boldsymbol{f}_i = \boldsymbol{f}_i$ しかありません。ですから AB は m 次の単位行列 E_m でないといけません。同様に、今度は基底 $(\boldsymbol{g}_1\ \boldsymbol{g}_2\ \cdots\ \boldsymbol{g}_n)$ から 2 度基底の変換を行うと、今度は BA が n 次の単位行列 E_n だと分かります。ここで AB, BA の対角和は等しいので、$m = \mathrm{tr}\, E_m = \mathrm{tr}\, AB = \mathrm{tr}\, BA = \mathrm{tr}\, E_n = n$ が従います。これで 2 つの基底の本数が等しいと言えました。

4.2 次元と単射性/全射性

線型写像の場合、単射性と全射性がそれぞれ核と像の次元で判定できるという著しい性質があります。

$\varphi\colon U \to V$ を線型写像とします。まず φ の単射性と $\mathrm{Ker}\,\varphi = \{\boldsymbol{0}\}$ とが同値でした。そして $\mathrm{Ker}\,\varphi = \{\boldsymbol{0}\}$ と $\dim \mathrm{Ker}\,\varphi = 0$ とが同値なので、結局 φ の単射性は $\dim \mathrm{Ker}\,\varphi = 0$ と言い換えられます。また φ の全射性の定義は $\mathrm{Im}\,\varphi = V$ です。一般に $\mathrm{Im}\,\varphi \subset V$ が成り立つから $\mathrm{Im}\,\varphi$ の基底を延長して V の基底が作れるので、$\mathrm{Im}\,\varphi = V$ が成り立つことは $\mathrm{Im}\,\varphi$ の基底が V の基底にもなること、つまり $\dim \mathrm{Im}\,\varphi = \dim V$ と同値です。

そして実は「線型写像の核と像の次元」には関係が付くので、それを使うことで、単射性や全射性に関する色々な情報を引き出すことができます。実務上も非常に役立つので、ぜひメカニズムを理解してください。まずは問 3 の解答から見てみます。

■問 3 の解答　U, V を線型空間、$\varphi\colon U \to V$ を線型写像、$\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n \in U$ とする。

(1) $\bigl(\varphi(\boldsymbol{f}_1), \varphi(\boldsymbol{f}_2), \ldots, \varphi(\boldsymbol{f}_n)\bigr)$ が 1 次独立とする。このとき $\alpha_1 \boldsymbol{f}_1 + \alpha_2 \boldsymbol{f}_2 + \cdots + \alpha_n \boldsymbol{f}_n = \boldsymbol{0}$ とおくと、両辺に φ を施して

$$\boldsymbol{0} = \varphi(\alpha_1 \boldsymbol{f}_1 + \alpha_2 \boldsymbol{f}_2 + \cdots + \alpha_n \boldsymbol{f}_n) = \alpha_1 \varphi(\boldsymbol{f}_1) + \alpha_2 \varphi(\boldsymbol{f}_2) + \cdots + \alpha_n \varphi(\boldsymbol{f}_n)$$

を得る。$\bigl(\varphi(\boldsymbol{f}_1), \varphi(\boldsymbol{f}_2), \ldots, \varphi(\boldsymbol{f}_n)\bigr)$ の 1 次独立性から $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$ が従う。よって $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ も 1 次独立である。

(2) $\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$ が U を生成したとする。任意に $\boldsymbol{v} \in \mathrm{Im}\,\varphi$ を取ると、$\boldsymbol{v} = \varphi(\boldsymbol{x})$ となる $\boldsymbol{x} \in U$ が取れる。$\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n$ は U を生成するので、$\boldsymbol{x} = \alpha_1 \boldsymbol{f}_1 + \alpha_2 \boldsymbol{f}_2 + \cdots + \alpha_n \boldsymbol{f}_n$ と表せる。すると

$$\boldsymbol{v} = \varphi(\boldsymbol{x}) = \varphi(\alpha_1 \boldsymbol{f}_1 + \alpha_2 \boldsymbol{f}_2 + \cdots + \alpha_n \boldsymbol{f}_n) = \alpha_1 \varphi(\boldsymbol{f}_1) + \alpha_2 \varphi(\boldsymbol{f}_2) + \cdots + \alpha_n \varphi(\boldsymbol{f}_n)$$

となり、$\boldsymbol{v}$ は $\varphi(\boldsymbol{f}_1), \varphi(\boldsymbol{f}_2), \ldots, \varphi(\boldsymbol{f}_n)$ の 1 次結合で表せる。$\boldsymbol{v} \in \mathrm{Im}\,\varphi$ は任意だったので、$\varphi(\boldsymbol{f}_1), \varphi(\boldsymbol{f}_2), \ldots, \varphi(\boldsymbol{f}_n)$ は $\mathrm{Im}\,\varphi$ を生成する。 ■

■問 3 の帰結　問 3 の結果を使うと、線型写像の単射性/全射性を次元で判定する簡単な方法が得られます。いま、$\varphi\colon U \to V$ を線型写像とします。そして $\dim U = n$, $\dim V = m$ とし、$(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ を U の基底とします。

まず (1) を見ましょう。(1) は「$\mathrm{Im}\,\varphi$ に 1 次独立なベクトルがあれば、少なくともそれと同じ数だけの 1 次独立なベクトルが U にある」と言っています。つまり $\dim \mathrm{Im}\,\varphi \leq \dim V$ が成り立ちます。

(2) も同様です。$(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ は U の基底なので、特に生成系です。したがって (2) の結果から、$\bigl(\varphi(\boldsymbol{f}_1), \varphi(\boldsymbol{f}_2), \ldots, \varphi(\boldsymbol{f}_n)\bigr)$ は $\mathrm{Im}\,\varphi$ を生成します。つまり $\mathrm{Im}\,\varphi$ は高々 n 本のベクトルがあれば張れます。一方、基底は「最小本数の生成系」でしたから、ここまでの議論でやはり $\dim \mathrm{Im}\,\varphi \leq n = \dim U$ が従います。つまり**線型写像の像の次元は、定義域の次元より大きくなりません**。特に $\dim U < \dim V$ なら $\dim \mathrm{Im}\,\varphi < \dim V$ となりますから、φ は全射にはなりえません。

■次元定理　次元については、さらに詳しい情報を得ることができます。いま $\varphi\colon U \to V$ を線型写像とします。このとき常に $\dim \mathrm{Ker}\,\varphi + \dim \mathrm{Im}\,\varphi = \dim U$ という式が成り立ちます。これを**次元定理**といいます。

証明の雰囲気はこんな感じです。基底を上手く取ることで U を「$\mathrm{Ker}\,\varphi$ の方向」と「それ以外の方向」とに分けることができます。すると線型写像 φ の値が潰れる方向は必ず $\mathrm{Ker}\,\varphi$ 方向なので、それ以外の方向には φ は潰れません。

したがって「それ以外の方向」を張る基底を φ で移すと、それが $\operatorname{Im}\varphi$ の基底になります。

この証明を細部まで詰めましょう。まず「基底の延長」を使います。$\operatorname{Ker}\varphi$ は U の部分空間なので、$\operatorname{Ker}\varphi$ の基底 $(\boldsymbol{f}_1,\ldots,\boldsymbol{f}_k)$ を延長して U 全体の基底 $(\boldsymbol{f}_1,\ldots,\boldsymbol{f}_k,\boldsymbol{f}_{k+1},\ldots,\boldsymbol{f}_n)$ を作れます。そうすると $\boldsymbol{f}_{k+1},\ldots,\boldsymbol{f}_n$ は $\operatorname{Ker}\varphi=\mathbb{R}\boldsymbol{f}_1+\cdots+\mathbb{R}\boldsymbol{f}_k$ に入らないので、$\varphi(\boldsymbol{f}_{k+1}),\ldots,\varphi(\boldsymbol{f}_n)$ はいずれも $\mathbf{0}$ にはなりません。

次に $\varphi(\boldsymbol{f}_{k+1}),\ldots,\varphi(\boldsymbol{f}_n)$ が 1 次独立なことを示します。$\alpha_1\varphi(\boldsymbol{f}_{k+1})+\alpha_2\varphi(\boldsymbol{f}_{k+2})+\cdots+\alpha_{n-k}\varphi(\boldsymbol{f}_n)=\mathbf{0}$ とおくと $\varphi(\alpha_1\boldsymbol{f}_{k+1}+\alpha_2\boldsymbol{f}_{k+2}+\cdots+\alpha_{n-k}\boldsymbol{f}_n)=\mathbf{0}$ となります。よって $\alpha_1\boldsymbol{f}_{k+1}+\alpha_2\boldsymbol{f}_{k+2}+\cdots+\alpha_{n-k}\boldsymbol{f}_n\in\operatorname{Ker}\varphi$ ですが、$(\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n)$ が $\operatorname{Ker}\varphi$ の基底だったので、$\alpha_1\boldsymbol{f}_{k+1}+\alpha_2\boldsymbol{f}_{k+2}+\cdots+\alpha_{n-k}\boldsymbol{f}_n=\mathbf{0}$ が従います。よって $\boldsymbol{f}_{k+1},\ldots,\boldsymbol{f}_n$ の 1 次独立性から $\alpha_1=\cdots=\alpha_{n-k}=0$ となります。

最後に $\varphi(\boldsymbol{f}_{k+1}),\ldots,\varphi(\boldsymbol{f}_n)$ が $\operatorname{Im}\varphi$ を生成することを示します。任意に $\boldsymbol{v}\in\operatorname{Im}\varphi$ を取ると、$\boldsymbol{v}=\varphi(\boldsymbol{u})$ $(\boldsymbol{u}\in U)$ と書けます。そして $(\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n)$ が U の基底なので、$\boldsymbol{u}=\alpha_1\boldsymbol{f}_1+\cdots+\alpha_n\boldsymbol{f}_n$ と書けます。ところが $\boldsymbol{f}_1,\ldots,\boldsymbol{f}_k\in\operatorname{Ker}\varphi$ だから $\boldsymbol{v}=\varphi(\alpha_1\boldsymbol{f}_1+\cdots+\alpha_n\boldsymbol{f}_n)=\alpha_{k+1}\varphi(\boldsymbol{f}_{k+1})+\cdots+\alpha_n\varphi(\boldsymbol{f}_n)$ となります。よって $\boldsymbol{u}$ は $\varphi(\boldsymbol{f}_{k+1}),\ldots,\varphi(\boldsymbol{f}_n)$ の 1 次結合で表せます。$\boldsymbol{u}$ は任意だったので、これで $\varphi(\boldsymbol{f}_{k+1}),\ldots,\varphi(\boldsymbol{f}_n)$ が $\operatorname{Im}U$ を生成することが言えました。

これらをまとめると、$(\boldsymbol{f}_1,\ldots,\boldsymbol{f}_n)$ の最初の k 本が $\operatorname{Ker}\varphi$ の基底で、残りの $k+1$ 本に φ を施すと $\operatorname{Im}\varphi$ の基底が得られることが分かります。よって次元の定義から、$\dim\operatorname{Ker}\varphi+\dim\operatorname{Im}\varphi=\dim U$ となります。

■次元定理の帰結　この次元定理を使うと、写像の全射性や単射性について更なる情報が得られます。

$\varphi\colon U\to V$ を線型写像とします。次元定理から $\dim\operatorname{Im}\varphi=\dim U-\dim\operatorname{Ker}\varphi\leq\dim U$ が従います。これはさっき、問 3 の帰結として述べたことです。また $\dim\operatorname{Ker}\varphi=\dim U-\dim\operatorname{Im}\varphi\geq\dim U-\dim V$ なので、もし $\dim U>\dim V$ であれば $\dim\operatorname{Ker}\varphi\geq 1$ となります。つまり定義域の方が値域より大きければ、それだけで φ は単射ではあり得ないと分かってしまうのです。

さらに極端なのが $\dim U=\dim V$ の場合です。このとき $\dim\operatorname{Ker}\varphi+\dim\operatorname{Im}\varphi=\dim V$ が成り立つので、$\dim\operatorname{Ker}\varphi=0$ と $\dim\operatorname{Im}\varphi=\dim V$ とが同値になります。つまり φ が単射であることと全射であることは同値で、**単射性または全射性の一方が成り立つだけで、自動的に φ が全単射になります**。こんな感じで、次元定理は非常に使いでがあるのです[*20]。

[*20] でも証明でみたとおり、この公式は「次元の足し算がたまたま良い式を満たす」という程度の話ではありません。重要なのは f の定義域が「$\operatorname{Ker} f$ の方向」と「それ以外の方向」に分解でき、かつ「それ以外の方向が潰れずに $\operatorname{Im} f$ に行く」という点です。つまり次元公式は「定義域の空間 V を 2 つの部分空間に分解する公式がまずあって、その式で次元だけに着目することによって次元公式が得られる」と理解すべきものです (ちなみにこの例のように、式などを「空間の (正確には、空間の “圏” というものの) レベル」まで持ち上げて理解することを “categorification” と呼びます)。

線形代数学① 第 11 回　　線型写像と行列

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 7 月 8 日 1 限

1　授業の締めくくりに向けて

レポートの解説プリントは次回分・次々回分が残っていますが、授業自体は今回が最後です。そこで一度、ここまでに勉強したことをまとめておきましょう。

1.1　簡単なまとめ

■線型代数の二本柱　だいたい S1 タームの最後 2 回くらいからずっと、授業では「線型代数」と呼ばれる内容を扱ってきました。その内容は、ざっくりと

- 連立 1 次方程式の解法と行列の計算
- 線型空間と線型写像の理論

の 2 つに分かれます。で、これらの理論と計算はもちろん無関係ではありません。線型空間や線型写像の理論は、行列計算を抽象化したものです。たとえば線型空間の理論の中で出てきた諸々の概念は、全て行列とか連立 1 次方程式の話の中に対応物があります。抽象化する 1 つの理由は、適用範囲を広げるためです。

抽象的な理論での概念	行列の話で対応するもの
線型空間	数ベクトル空間 $\mathbb{R}^n$
線型空間の次元	$\mathbb{R}^n$ の n
線型写像	行列
線型写像 A の核	$A\boldsymbol{x} = \boldsymbol{0}$ の解集合
A の像の次元	行列 A の階数
A の核の次元	$A\boldsymbol{x} = \boldsymbol{b}$ の解のパラメータの個数

■理論と計算の関わり　線型代数の理論の中に「基底」というものがありました。これは非常に強力なツールです。というのも、基底を取るということは、線型空間に座標を入れることそのものでした。ですから基底を取ってしまえば、線型空間は数ベクトル空間 $\mathbb{R}^n$ と同一視できます。そして線型写像 $f\colon U \to V$ の定義域 U と値域 V のそれぞれで基底を取ってしまえば、f は数ベクトル空間の間の線型写像と同一視できます。これは行列そのものです。

このように「基底」を使えば、線型空間や線型写像の話は、出自とは無関係に数ベクトル空間や行列の話に落とし込めます。世の中には行列とは無関係に線型写像が飛び出てくることがままあります。たとえば線型微分方程式を解くときや、特定の漸化式を満たす数列を求めるときなどです。でもそれらのややこしい話は、基底を使って行列の話にすり替えてしまう[*1]ことで、連立 1 次方程式の話に帰着させられます。こうすれば、問題が解けますね。線型代数は、実際に生じる問題を解くために非常に有用なツールなのです。こうして、線型代数をよく理解しておくと

- 抽象論を使って、色々な問題に適応できるようにしておいて
- その後で具体的な行列の問題に直し、問題を解く

[*1] 行列の話が使えるのは、基底を取った後です。基底を取るところまでは、個々の問題に応じて別途頑張る必要があります。

という方法で、色々な問題を解けるようになります。

また、線型代数の抽象論をよく理解していると、具体的な問題にも取り組みやすくなります。たとえば今回、再び連立1次方程式の問題に取り組みます。でも前回と全く同じ扱いでは面白くないですから、今回は次元の理論を使い、いかに手際よく解を求めるかという話を中心にします。具体的な計算をする上でも、抽象的な理論は計算の見通しを綺麗にする、非常に便利な道具なのです。このことも、抽象論を展開する一つのモチベーションです。

具体的な計算と抽象的な理論の両方に習熟して、線型代数を使いこなせるようになってください。

1.2 なぜ線型代数を学ぶのか？

さて、ここまで「線型代数が数学の問題を解く時に便利だ」という話をしてきましたが、皆さんの中には「本当に線型代数使う場面なんてあるんかいな？」と思っている人がいるかもしれません。そこで線型代数が現実に必要そうな場面を、2つほど紹介します[*2]。

■統計 おそらく、ほとんど全ての人が今後何らかの「データ」を扱うことになるでしょう。それは実験データかもしれませんし、社会調査の結果かもしれませんが、何にせよデータを基にした推論が必要になります。その際変数が複数個あると、しばしば行列計算が必要になります。たとえば主成分分析や重回帰分析といった手法を使うと、そもそも問題の記述に行列が不可欠になります。そして統計の問題では、最終的には計算を行わなければいけないものの、抽象的な線型代数の議論を援用することが多々あります。

■微分方程式 もう1つの重要な例は、微分方程式です。

何を勉強するにしても、大抵の場合は微分方程式と格闘する必要性に迫られます。たとえば物理を勉強すると、基本法則の多くが微分方程式で書かれているはずです。化学をするなら、電子の軌道を支配する Schrödinger 方程式という微分方程式のことを知らないといけません。また化学反応の速度は、反応機構に応じた微分方程式に従います。感染症の流行モデルを記述する手法の1つに SIR モデルというものがありますが、これも微分方程式で記述されます。その他色々例があると思いますが、何をするにしても、とにかく微分方程式は必要なのです。そして線型微分方程式が相手なら、線型代数を使うことができます。この辺の事情は、解空間のことを過去にちょこっとだけ紹介しました。

ただ世の中の微分方程式は、解けることの方が稀です。大体の場合は解けない微分方程式が出てくるので、近似で頑張る必要などが出てきます。その際に良く使う手法の一つが「線型化」と呼ばれるものです。これは「平衡状態」に対応する微分方程式の解があるとき、「平衡状態からのズレが従う微分方程式」を調べるというものです。こうするとズレの主要部分だけに着目した、線型微分方程式が得られます。こちらを解けば、元の方程式が完全に解けなくてもある程度の情報が得られます。たとえば「物をちょっと動かしたときに元の場所に戻るかどうか」とかくらいなら、確かめられます。だから線型微分方程式は、そのまま解ける場合でも近似方程式として出てくる場合でも、重要なのです。

こういう感じで、線型代数というのは非常に重要です。実際の問題に出くわしてみないとあまり実感が分からないかもしれませんが、後々使うということだけは、知っておいてください。

2 いくつかの便利な補題

問題を解く前に、脇道に逸れますが便利な話を紹介します。

[*2] なお、ここまでの授業では扱われていない中で、非常に重要なテーマが2つあります。それは正方行列の「行列式」と「固有値」です。これらを知っていると、線型代数の重要性が一段と身に染みて理解できるようになります。

2.1　1 次独立性の判定法

1 次独立性には便利な判定法があるので紹介します[*3]。それは「一部の成分だけを抜き出して 1 次独立だったら、元のベクトルの組も 1 次独立」ということです。たとえば

$$\boldsymbol{v}_1 := \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad \boldsymbol{v}_2 := \begin{pmatrix}0\\1\\1\end{pmatrix}$$

というベクトルを考えます。このとき $\boldsymbol{v}_1$ の $1,2$ 行目は ${}^t(1,0)$、$\boldsymbol{v}_2$ の $1,2$ 行目は ${}^t(0,1)$ なので、ここを見るだけで直ちに「$\alpha_1\boldsymbol{v}_1 + \alpha\boldsymbol{v}_2 = \boldsymbol{0}$ ならば $\alpha_1 = \alpha_2 = 0$」が分かります。今は一番分かり易い形を例として示しましたが、ベクトルがもう少しややこしい形でも、あるいは本数が増えても、1 次独立性を示すには「一部の成分さえ見て示せれば十分」というわけです。たとえば $(2,2)$ 行列の場合、2 本の列ベクトルが 1 次独立なことと行列式が 0 でないことは同値です。こういうのを組み合わせると、ラクに 1 次独立性が示せます。

証明はもちろん愚直にやってもできますが、ここでは「射影」というテクニックを使って示してみます。正整数の増加列 $1 \leq j_1 < j_2 < \cdots < j_p \leq n$ に対し、(p,n) 型行列 $\mathrm{Proj}_{j_1,j_2,\ldots,j_p} \in \mathrm{Mat}_{p,n}(\mathbb{R})$ を

$$\mathrm{Proj}_{j_1,j_2,\ldots,j_p} := \begin{array}{c} \begin{array}{cccc} {j_1}^{\mathrm{th}} & {j_2}^{\mathrm{th}} & & {j_p}^{\mathrm{th}} \end{array} \\ \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \end{array}$$

で定めます。つまり各 $1 \leq i \leq p$ に対し「i 行目は j_i 列目だけ 1 で、他は 0」という条件で定まる行列が $\mathrm{Proj}_{j_1,j_2,\ldots,j_p}$ です。これを $\mathbb{R}^n$ のベクトルに当てると、$j_1, j_2, \ldots, j_p$ 番目の成分だけをちょうど抜き出せます。たとえば

$$\mathrm{Proj}_{2,3,5}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}0&1&0&0&0\\0&0&1&0&0\\0&0&0&0&1\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\\x_5\end{pmatrix} = \begin{pmatrix}x_2\\x_3\\x_5\end{pmatrix}$$

となっていますね。

これを使えば、証明はすぐ終わります。前回の問 3 で、線型写像 φ に対し「$\varphi(\boldsymbol{f}_1), \ldots, \varphi(\boldsymbol{f}_n)$ が 1 次独立なら、$\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n$ も 1 次独立である」という事実を示しています。これを $\mathrm{Proj}_{j_1,j_2,\ldots,j_p}$ に適用するだけです。

2.2　正則行列と 1 次独立性/生成系

逆行列を持つ正方行列のことを正則行列と呼ぶのでした。行列に左右から正方行列をかけたとき、核や像がどう変わるかを見てみます。A を (m,n) 型行列とします。僕たちはこれから、m 次の正則行列 X と n 次の正則行列 Y に対し、XAY の核や像が A の核や像と比べてどうズレるかを調べます。特に念頭にあるのは、X や Y が基本行列の場合です。つまり「基本変形で核や像がどうズレるか」を調べます。

証明に入る前に、この問題を「線型写像」の立場からざっくり眺めてみましょう。行列 A はベクトル $\boldsymbol{x} \in \mathbb{R}^n$ に $A\boldsymbol{x} \in \mathbb{R}^m$ を対応させる写像なので、$A\colon \mathbb{R}^n \to \mathbb{R}^m$ と表せます。同じようにして m 次正方行列 X は $X\colon \mathbb{R}^m \to \mathbb{R}^m$

[*3] 今までも暗黙のうちに使っていますし、みなさんもきっと当然に感じていることなので、今更言う必要がないかもしれませんが……

という線型写像を、n 次正方行列 Y は $Y\colon \mathbb{R}^n \to \mathbb{R}^n$ という写像を定めます。こう思えば、$\boldsymbol{x}$ に $XAY\boldsymbol{x}$ を対応させる写像は $\boldsymbol{x} \mapsto Y\boldsymbol{x} \mapsto AY\boldsymbol{x} \mapsto XAY\boldsymbol{x}$ という合成写像だと分かります。これを図式で書くと

$$\mathbb{R}^n \xrightarrow{Y} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m \xrightarrow{X} \mathbb{R}^m$$

となります。元々の写像 $A\colon \mathbb{R}^n \to \mathbb{R}^m$ に、写像 Y を右から、写像 X を左からかませているわけです[*4]。

そして X や Y は正則行列なので、全単射を定めています。特に Y は全射なので、$\boldsymbol{x}$ が $\mathbb{R}^n$ の全ての元を動き回るとき、$Y\boldsymbol{x}$ も $\mathbb{R}^n$ の全ての元を動きます。こう思えば、$\operatorname{Im} AY$ と $\operatorname{Im} A$ は同じになりそうですね。また X は単射なので、X が作用して $\mathbf{0}$ になるベクトルは $\mathbf{0}$ 以外にありません。だから $\operatorname{Ker} XA$ と $\operatorname{Ker} A$ は同じになりそうです。

また X が全射だとはいえ、一般には $\operatorname{Im} A$ のベクトルたちを X で移して得られる $\operatorname{Im} XA$ は、当然元の $\operatorname{Im} XA$ とは異なります。でも X で潰れるベクトルがないので、$\dim \operatorname{Im} XA = \dim \operatorname{Im} A$ は期待してよさそうです。また $\operatorname{Ker} A$ のベクトルを Y^{-1} で動かすと $\operatorname{Ker} AY$ になるので、一般には $\operatorname{Ker} AY \neq \operatorname{Ker} A$ です。でも Y で潰れるベクトルはないので、$\dim \operatorname{Ker} AY = \dim \operatorname{Ker} A$ も期待してよさそうです。

こんなイメージを頭の中に置きながら、ちゃんと証明をしてみましょう。

■左からの正則行列の作用　まず X が m 次の正則行列のとき、$\operatorname{Ker} XA = \operatorname{Ker} A$ を示します。実際 $A\boldsymbol{x} = \mathbf{0}$ ならば $XA\boldsymbol{x} = \mathbf{0}$ は当たり前なので $\operatorname{Ker} A \subset \operatorname{Ker} XA$ です。そして X^{-1} が存在するので、$XA\boldsymbol{x} = \mathbf{0}$ が成り立てば両辺に左から X^{-1} をかけて、$A\boldsymbol{x} = X^{-1}\mathbf{0} = \mathbf{0}$ を得ます。よって $\operatorname{Ker} XA \subset \operatorname{Ker} A$ です。こうして $\operatorname{Ker} XA = \operatorname{Ker} A$ が分かりました。

■右からの正則行列の作用　次に Y が n 次の正則行列のとき、$\operatorname{Im} AY = \operatorname{Im} A$ を示します。勝手に $\boldsymbol{y} \in \operatorname{Im} AY$ を取ると $\boldsymbol{y} = AY\boldsymbol{x}$ なる $\boldsymbol{x} \in \mathbb{R}^n$ が存在します。この式を $\boldsymbol{y} = A(Y\boldsymbol{x})$ と読めば、$\boldsymbol{y} \in \operatorname{Im} A$ が従います。故に $\operatorname{Im} AY \subset \operatorname{Im} A$ です。逆に $\boldsymbol{y} \in \operatorname{Im} A$ とすると、$\boldsymbol{y} = A\boldsymbol{x}$ となる $\boldsymbol{x} \in \mathbb{R}^n$ が取れます。そうすると Y^{-1} の存在から $\boldsymbol{y} = AY(Y^{-1}\boldsymbol{x})$ と書けるので $\boldsymbol{y} \in \operatorname{Im} AY$ です。よって $\operatorname{Im} A \subset \operatorname{Im} AY$ です。これで $\operatorname{Im} A = \operatorname{Im} AY$ が示せました。

■次元について　今の状況で、$\operatorname{Ker} AY = \operatorname{Ker} A$ や $\operatorname{Im} XA = \operatorname{Im} A$ は成り立ちません。ですが $\dim \operatorname{Ker} AY = \dim \operatorname{Ker} A$ や $\dim \operatorname{Im} XA = \dim \operatorname{Im} A$ は成り立ちます。さらにまとめると X, Y が正則行列のとき、$\dim \operatorname{Ker} XAY = \dim \operatorname{Ker} A, \dim \operatorname{Im} XAY = \dim \operatorname{Im} A$ です。なぜなら $\dim \operatorname{Ker} XAY = \dim \operatorname{Ker} AY = n - \dim \operatorname{Im} AY = n - \dim A = \dim \operatorname{Ker} A$ です。これより $\dim \operatorname{Im} XAY = n - \dim \operatorname{Ker} XAY = n - \dim \operatorname{Ker} A = \dim \operatorname{Im} A$ となります。

2.3　基本変形で変わらないもの

今の話を基本行列に当てはめてみましょう。基本行列には 3 つの系列がありましたが、どれも正則なことに変わりありません。そして基本行列を左からかけると行基本変形に、右からかけると列基本変形[*5]になるのでした。ですから

- A を行基本変形しても、Ker は変化しない
- A を列基本変形しても、Im は変化しない

と分かります。「行基本変形で Ker は変化しない」というのは単に「行基本変形による掃き出しで方程式が解ける」というだけですから、既に知っていたことです。でも列基本変形で Im が求まるというのは、新しい情報です[*6]。

[*4] 写像の合成を書く順序は、作用する順序と逆向きなので気を付けましょう。写像 XAY は、ベクトルに「Y, A, X の順」で作用します。

[*5] 列基本変形はまだちゃんと扱っていませんが、次の節で 1 つ例を出します。

[*6] 問題 2 のヒントに「与えられた列ベクトルを転置で行ベクトルにして、それを並べた行列を行基本変形すれば答えが求まる」と書いてあるのは、今述べた原理によるものです。${}^t(AX) = {}^tX\,{}^tA$ なので「A に列基本変形をすること」は「A の転置に行基本変形をすること」と同じです。行基本変形だけで話を済ませるために、転置を取ったのだと思います。
　ですが個人的には、列基本変形を使った方が「像が保たれる」という事実が見やすくて良いように思いました。この**問 2 は間違える人が非常に多かった**ですので、計算の原理を今一度復習しておいてください。

また $\mathrm{Ker}\, A$ や $\mathrm{Im}\, A$ を完全に求めることはできなくても、次元の情報を組み合わせて上手く解を作れることもあります。たとえば一般に $\mathrm{Im}\, XA \neq \mathrm{Im}\, A$ なので、A を行基本変形したら $\mathrm{Im}\, A$ は分からなくなってしまいます。でも $\dim XA = \dim A$ までは分かるので、方程式 $A\boldsymbol{x} = \boldsymbol{0}$ を解いて $\mathrm{Ker}\, A$ を求めれば、次元定理 $\dim \mathrm{Im}\, A = n - \dim \mathrm{Ker}\, A$ から $\mathrm{Im}\, A$ の基底の本数が分かります。あとはたとえば A の見た目を利用するとかして、A の列ベクトルの中から $\dim \mathrm{Im}\, A$ 本の 1 次独立なベクトルを持ってくれば、それで $\mathrm{Im}\, A$ の基底が得られたことになります。

こんな風に「次元の情報」を使うことで、$\mathrm{Ker}\, A$ や $\mathrm{Im}\, A$ の基底が何本あるかを知ることができます。だから「$\mathrm{Ker}\, A$ と $\mathrm{Im}\, A$ を両方求める」という問題の場合、大体の場合「行基本変形で $\mathrm{Ker}\, A$ を求め、その後列基本変形で $\mathrm{Im}\, A$ を求める」という二度手間をする必要はあまりありません。さらに言えば、行基本変形と列基本変形を組み合わせても良いのです。どちらを併用しても Ker と Im の次元は保たれますから、単に基底の本数を求めるだけなら、行/列のやりやすい基本変形を両方使って大丈夫です。たとえば「行基本変形をしてたら、2 つ同じ列ベクトルが出てきた」なんて場合は、列基本変形で重複する列ベクトルを消してから、行基本変形を進めたりできます。

どういう手順で計算を進めるのが一番良いかは、問題によってまちまちですし、一発で最善策を見抜くことは難しいでしょう。でも最善策にたどり着けないとしても、手持ちの手法を上手く組み合わせることで、行き当たりばったりよりは大分マシな解き方ができるはずです。

■問 5 の解答　行列の階数 (Im の次元) や Ker の次元は基本変形で変わらない。

2.4　列基本変形の定義と例

それでは早速、列基本変形を使ってみましょう。行基本変形を定義する文章中で、「行」を「列」に置き換えると、そのまま列基本変形の定義になります。すなわち

- 2 つの列を入れ替えること
- ある列に別の列の定数倍を加えること
- ある列に 0 でない定数をかけること

が、列基本変形です。そして行基本変形が「基本行列の左からの積」で実現できたのと同様に、列基本変形は「基本行列の右からの積」で実現できます。

この辺の事情は、実際に手を動かして理解する方がよいでしょう。「列基本変形が像を保つ」という事実を使い、問 2 を列基本変形で解いてみます。計算を追いかけてみてください。繰り返し言いますが、**問 2 は間違えた人が非常に多い**です。必ず復習してください。

■問 2 の解答

$$A := \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{pmatrix} = \begin{pmatrix} -1 & 3 & 5 \\ 1 & -1 & 1 \\ 2 & 0 & 8 \end{pmatrix}$$

とおくと $W = \mathrm{Im}\, A$ である。そこで

$$\begin{aligned}
\begin{pmatrix} -1 & 3 & 5 \\ 1 & -1 & 1 \\ 2 & 0 & 8 \end{pmatrix} &\xrightarrow[\text{2, 3 列目の掃き出し}]{\text{(1, 1) 成分を要に}} \begin{pmatrix} -1 & 0 & 0 \\ 1 & 2 & 6 \\ 2 & 6 & 18 \end{pmatrix} = \begin{pmatrix} -1 & 3 & 5 \\ 1 & -1 & 1 \\ 2 & 0 & 8 \end{pmatrix} \begin{pmatrix} 1 & 3 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&\xrightarrow[\text{2 列目を 1/2 倍}]{\text{1 列目を (−1) 倍}} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 6 \\ -2 & 3 & 18 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 0 \\ 1 & 2 & 6 \\ 2 & 6 & 18 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \\
&\xrightarrow[\text{1, 3 列目を掃き出し}]{\text{(2, 2) 成分を要に}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 6 \\ -2 & 3 & 18 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & -6 \\ 0 & 0 & 1 \end{pmatrix}
\end{aligned}$$

となる。これで $W = \operatorname{Im} A$ の基底が ${}^t(1,0,1)$ と ${}^t(0,1,3)$ だと分かる。また、これらのベクトルの $\mathbf{0}$ でない 1 次結合は、必ず第 1 成分または第 2 成分のいずれかが 0 でない。よって ${}^t(0,0,1)$ を加えると、$\mathbb{R}^3$ の基底になる。 ∎

3 連立 1 次方程式再訪

さて今回の問題の多くは、再び連立 1 次方程式です。「またかよ」と思った人がいるかもしれませんが、前とは一つ違うことがあります。それは「線型空間の次元」という道具があることです。前回は掃き出し法の様々な使い方を知ることがテーマでしたが、今度は次元定理[*7]という強力な道具が加わるので、これを使っていっそうサクサクと解を求めにいくことができます。

■問 1 の解答 まず問 1 の解答をまとめて載せておきます。例によって、解の表示方法は 1 つだけ与えています。
(1) y は任意で $x = 1+2y$ (2) y は任意で $x = 2y$ (3) z は任意で $(x,y) = (2+3z, 1+2z)$ (4) x は任意で $(y,z) = (2x, x)$ (5) $(x,y,z) = (1,2,4)$ (6) $(x,y,z) = (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ (7) x は任意で $(y,z) = (x-1, x-3)$ (8) x, y は任意で $z = 3 - x - y$ (9) x は任意で $y = z = -1 + x$ (10) z は任意で $(x,y) = (-2+5z, -3-3z)$ (11) z は任意で $(x,y) = (1-2z, 1+2z)$

3.1 線型代数の言葉による連立 1 次方程式の解釈

連立 1 次方程式

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

を、いつものように $A\boldsymbol{x} = \boldsymbol{b}$ と表します。ここで $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{b} \in \mathbb{R}^m$ です。

■係数行列の表す線型写像 係数行列 A は (m,n) 型行列なので、$A\colon \mathbb{R}^n \to \mathbb{R}^m$ という線型写像を定めます[*8]。この見方に立つと、$A\boldsymbol{x} = \boldsymbol{b}$ を満たす $\boldsymbol{x}$ を見つけるということは、A による $\boldsymbol{b}$ の逆像 $\{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} = \boldsymbol{b}\}$ を全て求めることに他なりません。特に $\boldsymbol{b} = \mathbf{0}$ の場合、方程式を解くことは $\operatorname{Ker} A$ を求めることに他なりません。

■解のパラメータ 線型写像の場合、単射性の破れ方は核で支配されるのでした。もう一度復習しておくと、$\boldsymbol{x}$ と $\boldsymbol{x}'$ が共に解であれば、$\boldsymbol{x} - \boldsymbol{x}' \in \operatorname{Ker} A$ です。実際 $A\boldsymbol{x} = A\boldsymbol{x}' = \boldsymbol{b}$ なら $A(\boldsymbol{x} - \boldsymbol{x}') = \boldsymbol{b} - \boldsymbol{b} = \mathbf{0}$ です。また $\boldsymbol{x}$ が解で $\boldsymbol{y} \in \operatorname{Ker} A$ なら、$A(\boldsymbol{x} + \boldsymbol{y}) = \boldsymbol{b} + \mathbf{0} = \boldsymbol{b}$ が成り立ちます。だから $A\boldsymbol{x} = \boldsymbol{b}$ の解 $\boldsymbol{x}_0 \in \mathbb{R}^n$ を 1 つ見つけてしまえば、残りの解は全て $\boldsymbol{x}_0 + \boldsymbol{y}$ ($\boldsymbol{y} \in \operatorname{Ker} A$) と表せるし、逆にこの形以外の解は存在しません。

そしていま $\boldsymbol{x}_0 \in \mathbb{R}^n$ を $A\boldsymbol{x} = \boldsymbol{b}$ の 1 つの解とし、さらに $(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_l)$ を $\operatorname{Ker} A$ の基底とします。このとき全ての $\operatorname{Ker} A$ の元は $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_l$ の 1 次結合でただ 1 通りに表せます。したがって全ての解は、$\boldsymbol{x}_0 + t_1\boldsymbol{x}_1 + \cdots + t_l\boldsymbol{x}_l$ ($t_1, \ldots, t_l \in \mathbb{R}$) の形に一意的に表せます。逆に解が $\boldsymbol{x}_0 + t_1\boldsymbol{x}_1 + \cdots + t_l\boldsymbol{x}_l$ の形にパラメータ表示されるとき、$\boldsymbol{x}_1, \ldots, \boldsymbol{x}_l$ は $\operatorname{Ker} A$ の元です。このパラメータ表示が全ての解を尽くし、そして一意的であることが、$(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_l)$ が $\operatorname{Ker} A$ の基底であることを導きます。

このように解のパラメータは、$\operatorname{Ker} A$ の元を基底の 1 次結合で表すときの係数だったのです。だから解に現れるパラメータは、ちょうど $\dim \operatorname{Ker} A$ 個だと分かります。

[*7] A が (m,n) 型行列のとき、$n = \dim \operatorname{Ker} A + \dim \operatorname{Im} A$ が成り立つというやつです。証明は前回のプリントの 106 ページに書いてありますが、まずは式を認めて使ってみるというのでも良いでしょう。とにかく使い道のある公式なので、使い方は絶対に覚えてください。
[*8] 問題プリントでは f_A と書いていました。

■像の次元　A の列ベクトルのうち、1 次独立なものの最大本数を $\operatorname{rank} A$ というのでした。これは $\dim \operatorname{Im} A$ と同じです。というのも $A = (\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_n)$ とおくと、

$$A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_n \end{pmatrix}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x_1\boldsymbol{a}_1 + x_2\boldsymbol{a}_2 + \cdots + x_n\boldsymbol{a}_n$$

より $\operatorname{Im} A = \mathbb{R}\boldsymbol{a}_1 + \mathbb{R}\boldsymbol{a}_2 + \cdots + \mathbb{R}\boldsymbol{a}_n$ でした。ここで前回の結果を思い出しましょう。もし $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ が 1 次独立でなければ、このうち上手く 1 本を削っても、$\operatorname{Im} A$ の生成系が得られるのでした。そうして 1 次独立になるまで上手くベクトルを削り続けると、基底ができます。ここで $\operatorname{rank} A$ の定義は「1 次独立な A の列ベクトルの組の最大本数」であり、A の列ベクトルを上手く選ぶと $\operatorname{Im} A$ の基底ができるのだから、$\operatorname{rank} A \geq \operatorname{Im} A$ です。でも次元の性質から、$\operatorname{Im} A$ の中に $\dim \operatorname{Im} A$ 本以上の 1 次独立なベクトルを取ってくることはできません。だから $(\dim \operatorname{Im} A + 1)$ 本以上からなる A の列ベクトルの組は、必ず 1 次従属になります。これで $\operatorname{rank} A = \operatorname{Im} A$ が成り立つことが分かりました。

3.2　次元定理の利用例

実際に次元定理を使って、問 3 と問 4 を解いてみましょう。問 3 の (2) は「次元を考えるだけで解が分かる」という問題の典型例です。また問 4 では行基本変形と列基本変形のどちらを用いても解くことができます。なんとなく (1) では列基本変形、(2) では行基本変形を使うと解きやすい気がする[*9]ので、両方使ってみます。

■問 3 の解答　(2) $A := (1\quad 1\quad 1\quad 1)$ とおくと、A は線型写像 $\mathbb{R}^4 \to \mathbb{R}$ を定める。このとき $W = \operatorname{Ker} A$ である。$A \neq O$ より $\dim \operatorname{Im} A = \operatorname{rank} A \geq 1$ である。かたや $\operatorname{Im} A \subset \mathbb{R}$ より $\dim \operatorname{Im} A \leq 1$ なので、これで $\dim \operatorname{Im} A = 1$ と分かる。よって次元定理から $\dim \operatorname{Ker} A = 4 - 1 = 3$ が従う。

これより $\operatorname{Ker} A$ の基底を求めるには、3 本の 1 次独立なベクトルを見つければ良い。A の形から、1 次独立なベクトル 3 本が次のように取れると分かる:

$$\boldsymbol{v}_1 := \begin{pmatrix} -1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 := \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_3 := \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

(1)

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

とおくと、A は線型写像 $\mathbb{R}^4 \to \mathbb{R}^3$ を定める。このとき $W = \operatorname{Ker} A$ である。行基本変形で $\operatorname{Ker} A$ と $\operatorname{rank} A$ は不変なので、掃き出しを行う。すると

$$\begin{pmatrix} 1 & 1 & -1 & 1 \\ -2 & -1 & 0 & 1 \\ 3 & 1 & 1 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & -1 & 1 \\ 0 & 1 & -2 & 3 \\ 0 & -2 & 4 & -6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & -2 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

となる。これより $\operatorname{rank} A = 2$ が分かる。よって $\dim \operatorname{Ker} A = 4 - \dim \operatorname{Im} A = 2$ である。あとは $\operatorname{Ker} A$ の中に 2 本の 1 次独立なベクトルを探せばよいが、掃き出し後の表式から解は

$$\begin{pmatrix} -z + 2w \\ 2z - 3w \\ z \\ w \end{pmatrix} = z\begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix} + w\begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

[*9] (1) は 1 行目の 2 列目だけに 1 があり、(2) は 1 列目にたくさん 1 があるから、こう思いました。これ以上の深い理由はありません。

と表せる。よって

$$\boldsymbol{v}_1 := \begin{pmatrix} -1 \\ 2 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 := \begin{pmatrix} 2 \\ -3 \\ 0 \\ 1 \end{pmatrix}$$

が基底である。 ∎

■問 4 の解答　(1) 与えられた行列を列基本変形すると

$$\begin{pmatrix} 2 & 1 & 4 \\ 4 & 3 & 10 \\ 2 & 3 & 8 \end{pmatrix} \xrightarrow[\text{入れ替え}]{\text{1 列目と 2 列目を}} \begin{pmatrix} 1 & 2 & 4 \\ 3 & 4 & 10 \\ 3 & 2 & 8 \end{pmatrix} \xrightarrow[\text{2, 3 列目を掃き出し}]{(1,1)\text{ 成分を要に}} \begin{pmatrix} 1 & 0 & 0 \\ 3 & -2 & -2 \\ 3 & -4 & -4 \end{pmatrix}$$

$$\xrightarrow[(-1/2)\text{ 倍}]{\text{2 列目を}} \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & -2 \\ 3 & 2 & -4 \end{pmatrix} \xrightarrow[\text{1, 3 列目を掃き出し}]{(2,2)\text{ 成分を要に}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -3 & 2 & 0 \end{pmatrix}$$

となる。これで $\operatorname{Im} A$ の基底は $\left({}^t(1,0,-3), {}^t(0,1,2)\right)$ で、$\dim \operatorname{Im} A = 2$ と分かる。これより $\dim \operatorname{Ker} A = 3 - 2 = 1$ と分かる。だから $\operatorname{Ker} A$ に属する $\mathbf{0}$ でないベクトルを 1 本見つければ、それが基底になる。たとえば A を眺めると ${}^t(1,2,-1) \in \operatorname{Ker} A$ が分かる[*10]。

(2) 与えられた行列を行基本変形すると

$$\begin{pmatrix} 1 & 1 & 3 & 0 & 0 \\ 1 & 2 & 5 & -2 & -5 \\ 1 & 1 & 3 & 1 & 2 \\ 2 & 0 & 2 & 4 & 10 \end{pmatrix} \xrightarrow[\text{2, 3, 4 行目を掃き出し}]{(1,1)\text{ 成分を要に}} \begin{pmatrix} 1 & 1 & 3 & 0 & 0 \\ 0 & 1 & 2 & -2 & -5 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & -2 & -4 & 4 & 10 \end{pmatrix}$$

$$\xrightarrow[\text{1, 4 行目を掃き出し}]{(2,2)\text{ 成分を要に}} \begin{pmatrix} 1 & 0 & 1 & 2 & 5 \\ 0 & 1 & 2 & -2 & -5 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[(1,2)\text{ 行目を掃き出し}]{(3,4)\text{ 成分を要に}} \begin{pmatrix} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 2 & 0 & 1 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

となる。これで $\operatorname{Ker} A$ の基底が求まる:

$$\boldsymbol{v}_1 := \begin{pmatrix} -1 \\ -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 := \begin{pmatrix} -1 \\ -1 \\ 0 \\ -2 \\ 1 \end{pmatrix}$$

$\dim \operatorname{Ker} A = 2$ だから $\dim \operatorname{Im} A = 5 - 2 = 3$ である。そこで $\operatorname{Im} A$ を張る 3 本の列ベクトル、列基本変形で探す。

$$\begin{pmatrix} 1 & 1 & 3 & 0 & 0 \\ 1 & 2 & 5 & -2 & -5 \\ 1 & 1 & 3 & 1 & 2 \\ 2 & 0 & 2 & 4 & 10 \end{pmatrix} \xrightarrow[\text{5 列目に足す}]{\text{4 列目の }(-3)\text{ 倍を}} \begin{pmatrix} 1 & 1 & 3 & 0 & 0 \\ 1 & 2 & 5 & -2 & 1 \\ 1 & 1 & 3 & 1 & -1 \\ 2 & 0 & 2 & 4 & -2 \end{pmatrix} \xrightarrow[\text{4 列目に足す}]{\text{5 列目の 2 倍を}} \begin{pmatrix} 1 & 1 & 3 & 0 & 0 \\ 1 & 2 & 5 & 0 & 1 \\ 1 & 1 & 3 & -1 & -1 \\ 2 & 0 & 2 & 0 & -2 \end{pmatrix}$$

とすると、上 3 行を見ることで 1, 4, 5 列目のベクトルが 1 次独立だと分かる[*11]。$\dim \operatorname{Im} A = 3$ なので、これらが $\operatorname{Im} A$ の基底をなす。

[*10] これは、1 行目を眺めて何となく「1 列目と 2 列目の 2 倍を足したら 3 列目になりそうだなあ」と思ってたら答えが見つかりました。今回はうまく答えを見つけましたが、一般の場合は行基本変形による掃き出しで攻めるのが正攻法です。$\dim \operatorname{Ker} A = 1$ と分かってるので、掃き出し法を途中まで行えば答えに当たりがつくはずです。

[*11] もうちょっとだけ補足すると、1 列目、5 列目、4 列目の順番にベクトルを並べると下三角行列ができます。ここまでくれば 1 次独立なことはほとんど明らかです。

線形代数学① 第 12 回　　基底の取り換えと行列表示

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 7 月 15 日 1 限

1　連絡事項など

■レポート解答の返却返却　7/22 に回収した皆さんのレポートは、添削してアドミニストレーション棟のレポートボックスに入れています。各自受け取っていってください。なお答案は学生証番号順に並べているので、**自分のを持っていく際に順番を乱さないでください**。

なお、過去に提出された答案も一緒に返却しています。受け取りそびれた人は、持って行ってください。

■記法について　「基底の変換行列」の話を書いていて気づいたのですが、もしかしたら、このプリントと授業と、あるいは他の本とで記法が違う箇所があるかもしれません。時と場合によってどういう記法が便利かが違ってくるので、このような現象が起きえます。もし試験で解答する際に気になることがあったら「どういう記法を使っているか」が分かる答案を書きましょう。そうすれば誤解の余地はありません。

■その他、試験に向けて　質問がある人は、先生にメールを送っても、TA の穂坂 `hosaka@ms.u-tokyo.ac.jp` にメールを送っても構いません。可能な限り応対します。

授業範囲を理解するのに必要なことは、これまでのプリントで大体書いたつもりです。試験まで残り 2 日しかないので、あまり色々な事はできないと思いますが、ノートとプリントを参考にして勉強しましょう。がんばってください。

2　線型写像の行列表示

これは非常に大事な話なのですが、あまりできている人が多くなかったので、なるべく丁寧めに解説をします。

まずは、話の筋を言ってしまいます。$f\colon U \to V$ を線型写像とします。このとき

- U と V は共に基底を入れることで、数ベクトル空間と同一視される
- この同一視のもと、$f\colon U \to V$ は数ベクトル空間の間の線型写像になるから、行列で表示される

というわけです。とにかく**線型写像は行列だと思える**という点が大事です。行列にさえしてしまえば、後は掃き出し法とか色々な道具で計算ができるのですから。

2.1　基底の復習

V を線型空間とします。V の元の列 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \dots, \boldsymbol{f}_n)$ が V **の基底**であるとは

- $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \dots, \boldsymbol{f}_n\}$ は V を生成する
- $\{\boldsymbol{f}_1, \boldsymbol{f}_2, \dots, \boldsymbol{f}_n\}$ は 1 次独立である

という 2 つの条件が満たされることでした。そして $\boldsymbol{v} \in V$ に対し

$$\boldsymbol{v} = x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \cdots + x_n\boldsymbol{f}_n$$

と表したとき、基底の係数 $x_1, \dots, x_n$ を並べてできる数ベクトル ${}^t(x_1, \dots, x_n)$ を (基底 $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ に関する) $\boldsymbol{v}$ の**座標**あるいは**成分表示**といいます。

基底を特徴づける 2 つの条件が、それぞれ「全てのベクトルに座標が割り当てられること」「座標の割り当て方がただ 1 通りであること」を保証することは、既に 97 ページで説明しました。

■線型空間の同型　いま $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ を V の基底として、線型写像 $\varphi\colon \mathbb{R}^n \to V$ を

$$\varphi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) := x_1\boldsymbol{f}_1 + \cdots + x_n\boldsymbol{f}_n$$

と定めます。すると基底の定義から

- $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ は V を生成するので、φ は全射
- $\boldsymbol{f}_1, \dots, \boldsymbol{f}_n$ は 1 次独立なので $\operatorname{Ker}\varphi = \{\boldsymbol{0}\}$ となり、φ は単射

です。これより φ は全単射となり、φ によって $\mathbb{R}^n$ と V の元とが $1:1$ に対応します。さらに φ は線型写像だから、$\mathbb{R}^n$ での加法/スカラー倍の結果を φ で送ったものは、先にベクトルを φ で送ってから V の中で加法/スカラー倍した結果と一致します。だから**線型空間の構造まで込めて、φ で $\mathbb{R}^n$ と V が同一視されます**。

そして見れば分かりますが、φ で ${}^t(x_1, \dots, x_n) \in \mathbb{R}^n$ を写した先にあるのは、基底 $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ のもとで座標 ${}^t(x_1, \dots, x_n)$ に対応する点です。ですから φ は、**座標とベクトルとの対応を与える写像**だといえます。逆に線型空間の同型 $\varphi\colon \mathbb{R}^n \to V$ が与えられれば、$\boldsymbol{f}_1 := \varphi(\boldsymbol{e}_1), \dots, \boldsymbol{f}_n := \varphi(\boldsymbol{e}_n)$ は V の基底をなします。したがって**座標を入れることは、$\varphi\colon \mathbb{R}^n \to V$ で線型空間を同一視すること**に他なりません。

もう一度同じ式を書きますが

$$\varphi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) := x_1\boldsymbol{f}_1 + \cdots + x_n\boldsymbol{f}_n$$

によって、V のベクトルとその座標とが対応します。だからこの φ は「座標をベクトルに読み替えるときのおまじない」みたいなものです。逆に φ^{-1} がついてたら「ベクトルを基底で展開した係数を拾い、対応する座標に読み替える」という意味になります。一度理解してしまえば大したことはないので、記号に慣れてください。

いっそ「線型空間に基底を入れちゃったら、もう $x_1\boldsymbol{f}_1 + \cdots + x_n\boldsymbol{f}_n$ のことを ${}^t(x_1, \dots, x_n)$ と書いていいじゃん」と思わなくもないし、実際に慣れてくるとそうすることも多いのですが、一方で「同一視のしかた」に気を配るのも数学の習わしです。今のうちは座標を考えるとき、きちんと $\varphi\bigl({}^t(x_1, \dots, x_n)\bigr)$ と書きましょう。

2.2　線型写像の行列表示

ここで U, V を線型空間とし、$f\colon U \to V$ を線型写像とします。$(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ を U の基底、$(\boldsymbol{g}_1, \dots, \boldsymbol{g}_m)$ を V の基底とします。そうすると座標による同一視 $\varphi\colon \mathbb{R}^n \to U$, $\psi\colon \mathbb{R}^m \to V$ が得られます。図式では

$$\begin{tikzcd} U \arrow[r, "f"] & V \arrow[d, "\psi^{-1}"] \\ \mathbb{R}^n \arrow[u, "\varphi"] \arrow[r, dashed] & \mathbb{R}^m \end{tikzcd}$$

という感じです。ですから写像の合成 $\psi^{-1} \circ f \circ \varphi$ によって、**元々は U から V への線型写像だった f を、座標の入った $\mathbb{R}^n$ から $\mathbb{R}^m$ への線型写像に読み替える**ことができます。これは行列に他なりません。基底を取ることで、線型写像は行列表示できるのです[*1]。

[*1] なお $U = V$ で $f\colon V \to V$ のときは、ふつう定義域と値域で同じ基底を取ります。線型変換を考えるときに、わざわざ違う基底を持ってくる理由はないですので。

非常に易しい例として、問 4 を解いてみましょう。皆さんは複素平面のことを知っていますから、$\mathbb{R}^2$ のベクトルと $\mathbb{C}$ を同一視できるはずです。この同一視のもと、$\mathbb{C}$ の元に対する操作をベクトルに対する操作と読み替えて、行列で表示してみます。

■問 4 の解答　実線型空間 $\mathbb{C}$ の基底として $\{1, i\}$ を取り、$\mathbb{C}$ と $\mathbb{R}^2$ を同一視する。このとき同一視を与える写像 $\varphi\colon \mathbb{R}^2 \to \mathbb{C}$ は

$$\begin{pmatrix} a \\ b \end{pmatrix} \underset{\varphi}{\overset{\varphi^{-1}}{\leftrightarrows}} a \cdot 1 + b \cdot i$$

である[*2]。

(1) 複素数 $\alpha = a + bi \in \mathbb{C}$ について、α 倍写像 $f_\alpha\colon \mathbb{C} \to \mathbb{C}$ を考える。このとき $z = x + yi$ に対し $f_\alpha(z) = \alpha z = (ax - by) + (bx + ay)i$ である。この式を同一視 $\varphi\colon \mathbb{R}^2 \to \mathbb{C}$ を使って読み替えると

$$\begin{pmatrix} x \\ y \end{pmatrix} \xmapsto{\varphi} x \cdot 1 + y \cdot i \xmapsto{f_\alpha} (ax - by) \cdot 1 + (bx + ay) \cdot i \xmapsto{\varphi^{-1}} \begin{pmatrix} ax - by \\ bx + ay \end{pmatrix} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

となっている。よって f_α の行列表示は

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

である。

(2) $\iota\colon \mathbb{C} \to \mathbb{C}$ を複素共役を取る写像とする。このとき $\iota(x + yi) = x - yi$ なので、同一視 $\varphi\colon \mathbb{R}^2 \to \mathbb{C}$ で読み替えると

$$\begin{pmatrix} x \\ y \end{pmatrix} \xmapsto{\varphi} x \cdot 1 + y \cdot i \xmapsto{\iota} x \cdot 1 - y \cdot i \xmapsto{\varphi^{-1}} \begin{pmatrix} x \\ -y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

となっている。よって ι の行列表示は

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

である。 ■

■行列表示の計算　では一般の場合 $f\colon U \to V$ は、どのような行列で表現されるのでしょうか？ それを追いかけるために、「行列を線型写像とみたとき、成分の意味は何か」を考えてみましょう。(m, n) 型行列 $A = (a_{ij}) \in \mathrm{Mat}_{m,n}(\mathbb{R})$ を $\mathbb{R}^n$ の標準基底に作用させると

$$A\boldsymbol{e}_j = \begin{array}{c} \begin{array}{ccccccc} \phantom{a_{11}} & \phantom{a_{12}} & & j^{\mathrm{th}} & & \phantom{a_{1n}} \end{array} \\ \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mj} & \cdots & a_{mn} \end{pmatrix} \end{array} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

となっています。つまり行列 A の j 列目は「標準基底の j 番目のベクトル $\boldsymbol{e}_j$ を A で移した先の $A\boldsymbol{e}_j$ を、成分で表示したもの」になっています。

一般の場合もこれを真似て、U の基底を f で写した後に、V の基底で表示してみましょう。U の基底 $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ のそれぞれに対し、$f(\boldsymbol{f}_i)$ は V の基底 $(\boldsymbol{g}_1, \ldots, \boldsymbol{g}_m)$ の 1 次結合で表されるはずです。その式を

$$f(\boldsymbol{f}_1) = a_{11}\boldsymbol{g}_1 + a_{21}\boldsymbol{g}_2 + \cdots + a_{m1}\boldsymbol{g}_m$$

[*2] 右辺は単に $a + bi$ と書いてよいのですが、基底を明示するため、ここでは敢えて $a \cdot 1 + b \cdot i$ と書いています。

$$
\begin{aligned}
f(\boldsymbol{f}_2) &= a_{12}\boldsymbol{g}_1 + a_{22}\boldsymbol{g}_2 + \cdots + a_{m2}\boldsymbol{g}_m \\
&\vdots \\
f(\boldsymbol{f}_n) &= a_{1n}\boldsymbol{g}_1 + a_{2n}\boldsymbol{g}_2 + \cdots + a_{mn}\boldsymbol{g}_m
\end{aligned}
$$

とおきます。これを使って f の行列表示を求めます。まず

$$
\varphi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) = x_1\boldsymbol{f}_1 + \cdots + x_n\boldsymbol{f}_n
$$

で、座標を U のベクトルに読み替えます。これに f を施すと

$$
\begin{aligned}
f\circ\varphi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) &= f(x_1\boldsymbol{f}_1 + \cdots + x_n\boldsymbol{f}_n) = x_1 f(\boldsymbol{f}_1) + \cdots + x_n f(\boldsymbol{f}_n) \\
&= x_1(a_{11}\boldsymbol{g}_1 + a_{21}\boldsymbol{g}_2 + \cdots + a_{m1}\boldsymbol{g}_m) + \cdots + x_n(a_{1n}\boldsymbol{g}_1 + a_{2n}\boldsymbol{g}_2 + \cdots + a_{mn}\boldsymbol{g}_m) \\
&= (a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)\boldsymbol{g}_1 + \cdots + (a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n)\boldsymbol{g}_m
\end{aligned}
$$

となります。そして、このベクトルに ψ^{-1} を当てると座標が取り出せるので

$$
\begin{aligned}
\psi^{-1}\circ f\circ\varphi\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) &= \psi^{-1}\big((a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n)\boldsymbol{g}_1 + \cdots + (a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n)\boldsymbol{g}_m\big) \\
&= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}
\end{aligned}
$$

となります。ですから f は、行列

$$
\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}
$$

で表現されると分かります。

この式をもう一度よく見ると、j 列目に並ぶ $a_{1j}, a_{2j}, \ldots, a_{mj}$ の出所は $f(\boldsymbol{f}_j) = a_{1j}\boldsymbol{g}_1 + a_{2j}\boldsymbol{g}_2 + \cdots + a_{mj}\boldsymbol{g}_m$ という式です。ですから f の行列表示では、**j 番目の基底に f を施してできるベクトルの成分表示が j 列目にくる**ということになります。さっき、行列の積を線型写像と思ったときの話と全く同じですね。

これを踏まえた上で、問 2, 3 も解いてみましょう。

■問 3, 2 の解答　n 次以下の多項式全体のなす線型空間 $\mathbb{R}[x]_{<n}$ を

$$
\varphi\colon \mathbb{R}^{n+1} \to \mathbb{R}[x]_{<n};\quad \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} \mapsto b_0 + b_1 x + \cdots + b_n x^n
$$

によって $\mathbb{R}^{n+1}$ と同一視する[*3]。

[*3] ここで行っているのは「多項式の各項の係数を拾ってベクトルを作る」操作であって、決して「項そのものを並べてベクトルを作る操作」ではありません。勘違いしている人が非常に多かったので、注意しておきます。

全ての自然数 $k \in \mathbb{N}$ に対し $Dx^k = kx^{k-1}$ である。よって同一視 φ を使うと、D は

$$\begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix} \xmapsto{\varphi} b_0 + b_1 x + \cdots + b_n x^n \xmapsto{D} b_1 + 2b_2 x + \cdots + nb_n x^{n-1} \xmapsto{\varphi^{-1}} \begin{pmatrix} b_1 \\ 2b_2 \\ \vdots \\ nb_n \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & 0 & n \\ & & & & 0 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_{n-1} \\ b_n \end{pmatrix}$$

と見えるので、D の行列表示は

$$\varphi^{-1} \circ D \circ \varphi = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \ddots & \ddots & \\ & & & 0 & n \\ & & & & 0 \end{pmatrix}$$

となる。また二項定理より、全ての自然数 $k \in \mathbb{N}$ に対し

$$T_a x^k = (x-a)^k = (-a)^k + \binom{k}{1}(-a)^{k-1}x + \cdots + \binom{k}{k-1}(-a)x^{k-1} + x^k$$

が成り立つ[*4]。これで基底の行先が分かったので、x^k と $\boldsymbol{e}_{k+1} \in \mathbb{R}^{n+1}$ が対応することを踏まえて、T_a の行列表示が

$$\varphi^{-1} \circ T_a \circ \varphi = \begin{pmatrix} 1 & -a & a^2 & -a^3 & \cdots & (-a)^n \\ & 1 & -2a & 3a^2 & \cdots & \binom{n}{1}(-a)^{n-1} \\ & & 1 & -3a & & \vdots \\ & & & 1 & \ddots & \vdots \\ & & & & \ddots & \binom{n}{n-1}(-a) \\ & & & & & 1 \end{pmatrix}$$

と分かる。 ∎

■関係式の「表現」 合成函数の微分法を知っていると、多項式 $f(x) \in \mathbb{R}[x]$ に対し $\bigl(f(x-\alpha)\bigr)' = f'(x-\alpha)$ が成り立つと分かります。この式をいまの T_a と D を使って書くと

$$D\bigl(T_a f(x)\bigr) = T_a\bigl(Df(x)\bigr)$$

に他なりません。さらに少し書き換えると $(DT_a - T_a D)\bigl(f(x)\bigr) = 0$ が全ての多項式 $f(x) \in \mathbb{R}[x]$ に対して成り立ちます。だから $\mathbb{R}[x]$ 上の線型写像として、$DT_a - T_a D = 0$ が成り立ちます。

この事実を反映して、さっき僕たちが求めた行列表示も $DT_a - T_a D = 0$ を満たすはずです。たとえば $n = 3$ のとき、実際に計算をすると確かに

$$\begin{pmatrix} 1 & -a & a^2 & -a^3 \\ & 1 & -2a & 3a^2 \\ & & 1 & -3a \\ & & & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & & \\ & 0 & 2 & \\ & & 0 & 3 \\ & & & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & -2a & 3a^2 \\ 0 & 0 & 2 & -6a \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & & \\ & 0 & 2 & \\ & & 0 & 3 \\ & & & 0 \end{pmatrix} \begin{pmatrix} 1 & -a & a^2 & -a^3 \\ & 1 & -2a & 3a^2 \\ & & 1 & -3a \\ & & & 1 \end{pmatrix}$$

となっています。n は何でも良いですから、n に応じて $DT_a - T_a D = 0$ という式を様々なサイズの行列で表現できる、というわけです。その他 $T_a T_b = T_{a+b} = T_b T_a$ などの関係式も、同様に成り立っています。興味のある人は確かめてみてください。

そして今は「平行移動」と「微分」という操作が先に与えられていましたが、逆に $DT_a - T_a D$ のような「関係式」だけを最初に取り出して、後から「どのように行列で表現できるか」を考えることもしばしばあります。このような分野は**表現論**と呼ばれ、現代数学の主たる分野の 1 つを形成しています。

[*4] ここで $\binom{k}{l}$ は、二項係数を表します。${}_kC_l$ と全く同じ意味ですが、大学の教科書や洋書ではほとんど $\binom{k}{l}$ という記号が使われます。

3 基底の変換

線型空間にはいくつもの基底があるので、それに応じていくつもの座標があります。それらの座標がどう関係するのかを、調べてみましょう。

3.1 基底の変換行列

■基底の変換 線型空間 V に 2 つの基底が与えられたとします。このとき 2 つの基底に応じて 2 つの座標が定まります。すると当然「2 つの座標はどのように対応するのか」が問題になります。すなわち、座標による同一視 $\varphi, \psi\colon \mathbb{R}^n \to V$ が与えられたとき、図式

$$\begin{tikzcd} & V & \\ \mathbb{R}^n \arrow[ur, "\varphi"] \arrow[rr, dashed] & & \mathbb{R}^n \arrow[ul, "\psi"'] \end{tikzcd}$$

における点線の矢印 $\psi^{-1} \circ \varphi$ が座標変換を表していて、これを具体的に行列で表示したいのです。

恒等写像 $\mathrm{id}\colon \mathbb{R}^n \to \mathbb{R}^n$ を使って今の図式に手を加えると

$$\begin{tikzcd} V \arrow[r, "\psi^{-1}"] & \mathbb{R}^n \arrow[d, "\mathrm{id}"] \\ \mathbb{R}^n \arrow[u, "\varphi"] \arrow[r, dashed] & \mathbb{R}^n \end{tikzcd}$$

とも書けます。こうすると、さっき線型写像を行列表示したときの図式と同じになります。だから座標変換の話は、座標を与える写像 ψ の逆写像 $\psi^{-1}\colon V \to \mathbb{R}^n$ を、φ の定める基底および $\mathbb{R}^n$ に関して行列表示することだと言えます。そうすれば「基底の行先を見ればよい」という話が自ずと見えてきます[*5]。

いま φ が定める基底を $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ とし、ψ が定める基底を $(\boldsymbol{g}_1, \dots, \boldsymbol{g}_n)$ とします。このどちらもが基底だから、それぞれの $\boldsymbol{g}_i$ は

$$\begin{aligned} \boldsymbol{g}_1 &= a_{11}\boldsymbol{f}_1 + a_{21}\boldsymbol{f}_2 + \cdots + a_{n1}\boldsymbol{f}_n \\ \boldsymbol{g}_2 &= a_{12}\boldsymbol{f}_1 + a_{22}\boldsymbol{f}_2 + \cdots + a_{n2}\boldsymbol{f}_n \\ &\vdots \\ \boldsymbol{g}_n &= a_{1n}\boldsymbol{f}_1 + a_{2n}\boldsymbol{f}_2 + \cdots + a_{nn}\boldsymbol{f}_n \end{aligned}$$

とただ 1 通りに表せます。これを行列っぽい記号で書くと

$$\begin{pmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

となります。ここに出てくる行列を**基底 $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ から基底 $(\boldsymbol{g}_1, \dots, \boldsymbol{g}_n)$ への変換行列**といいます。

■座標変換と基底変換 少々こんがらがる話になるのですが、**基底の変換と座標の変換を混同してはいけません**。今から見るように、基底をある正則行列 A で別の基底に変換すると、対応する座標は A^{-1} で変換されます。

これまで通り $(\boldsymbol{f}_1, \dots, \boldsymbol{f}_n)$ と $(\boldsymbol{g}_1, \dots, \boldsymbol{g}_n)$ を共に V の基底とし、その基底の変換行列を A とします。すなわち

$$\begin{pmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n \end{pmatrix} A$$

[*5] ただ、座標を変換する話と基底を変換する話を区別しないと混乱します。この点については、すぐ後ろで述べます。

です。そして $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ に関するベクトル $\boldsymbol{v}$ の座標表示が ${}^t(x_1, x_2, \ldots, x_n)$ だったとしましょう。このとき $\boldsymbol{v} = x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \cdots + x_n\boldsymbol{f}_n$ となるわけですが、この式を行列の積っぽい記法で

$$\boldsymbol{v} = x_1\boldsymbol{f}_1 + x_2\boldsymbol{f}_2 + \cdots + x_n\boldsymbol{f}_n = \begin{pmatrix}\boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

と書き直してみます。こうすると**基底と座標のペアによって、ベクトルが定まること**が見やすい。そして、この表示だと基底と座標の間に行列を挟み込めるので

$$\boldsymbol{v} = \begin{pmatrix}\boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n\end{pmatrix}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n\end{pmatrix}A\cdot A^{-1}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n\end{pmatrix}\cdot A^{-1}\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}$$

という式変形ができます。これで座標の変換規則が分かります。基底 $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ から $(\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n)$ に対する変換行列が A で与えられるとき、基底 $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ での座標に A^{-1} をかけたものが基底 $(\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n)$ での座標になります。**基底の変換の仕方と座標の変換の仕方は逆行列の関係にある**ので、混同しないよう気を付けてください。

■問 1 の解答　(1)

$$(x-a)^k = (-a)^k + \binom{k}{1}(-a)^{k-1}x + \cdots + \binom{k}{k-1}(-a)x^{k-1} + x^k$$

なので

$$\begin{pmatrix}1 & x-a & (x-a)^2 & \cdots & (x-a)^n\end{pmatrix} = \begin{pmatrix}1 & x & x^2 & \cdots & x^n\end{pmatrix}\begin{pmatrix}1 & -a & a^2 & -a^3 & \cdots & (-a)^n\\ & 1 & -2a & 3a^2 & \cdots & \binom{n}{1}(-a)^{n-1}\\ & & 1 & -3a & & \vdots\\ & & & 1 & \ddots & \vdots\\ & & & & \ddots & \binom{n}{n-1}(-a)\\ & & & & & 1\end{pmatrix}$$

で基底の変換行列が求まる。

(2) $x^k = \bigl((x-a)+a\bigr)^k$ より、(1) の変換行列で a を $-a$ に置き換えたものが求める行列である。すなわち

$$\begin{pmatrix}1 & x & x^2 & \cdots & x^n\end{pmatrix} = \begin{pmatrix}1 & x-a & (x-a)^2 & \cdots & (x-a)^n\end{pmatrix}\begin{pmatrix}1 & a & a^2 & a^3 & \cdots & a^n\\ & 1 & 2a & 3a^2 & \cdots & \binom{n}{1}a^{n-1}\\ & & 1 & 3a & & \vdots\\ & & & 1 & \ddots & \vdots\\ & & & & \ddots & \binom{n}{n-1}a\\ & & & & & 1\end{pmatrix}$$

である。 ∎

■基底変換の合成　基底の変換行列の話が分かってしまえば、基底の変換を 2 度行うのは簡単です。いま V を線型空間、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$, $(\boldsymbol{g}_1, \ldots, \boldsymbol{g}_n)$ と $(\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n)$ とをそれぞれ V の基底とします。そして基底の変換行列が

$$\begin{pmatrix}\boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n\end{pmatrix}A, \quad \begin{pmatrix}\boldsymbol{h}_1 & \boldsymbol{h}_2 & \cdots & \boldsymbol{h}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n\end{pmatrix}B$$

で与えられていたとしましょう。このとき基底 $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ から基底 $(\boldsymbol{h}_1, \ldots, \boldsymbol{h}_n)$ への変換行列は、上に書いた式を代入すれば

$$\begin{pmatrix} \boldsymbol{h}_1 & \boldsymbol{h}_2 & \cdots & \boldsymbol{h}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{g}_1 & \boldsymbol{g}_2 & \cdots & \boldsymbol{g}_n \end{pmatrix} B = \begin{pmatrix} \boldsymbol{f}_1 & \boldsymbol{f}_2 & \cdots & \boldsymbol{f}_n \end{pmatrix} AB$$

より AB と求まります。

■基底変換による行列表示の変化　ここまでの話のまとめとして、「線型写像の行列表示が、基底の変換でどう影響を受けるか」を調べます。ちょっとセッティングがややこしいので、箇条書きでまとめておきます。

- U, V は線型空間で、$f \colon U \to V$ は線型写像
- $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$, $(\boldsymbol{f}'_1, \boldsymbol{f}'_2, \ldots, \boldsymbol{f}'_n)$ は共に U の基底で、それぞれが線型空間の同一視 $\varphi, \varphi' \colon \mathbb{R}^n \to U$ を与える
- $(\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m)$, $(\boldsymbol{g}'_1, \boldsymbol{g}'_2, \ldots, \boldsymbol{g}'_m)$ は共に V の基底で、それぞれが線型空間の同一視 $\psi, \psi' \colon \mathbb{R}^m \to V$ を与える

このとき f について

- U の基底 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ と V の基底 $(\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m)$ に関する行列表示 A
- U の基底 $(\boldsymbol{f}'_1, \boldsymbol{f}'_2, \ldots, \boldsymbol{f}'_n)$ と V の基底 $(\boldsymbol{g}'_1, \boldsymbol{g}'_2, \ldots, \boldsymbol{g}'_m)$ に関する行列表示 B

を考えます。すると、2 つの図式

$$\begin{tikzcd} U \arrow[r, "f"] & V \arrow[d, "\psi^{-1}"] \\ \mathbb{R}^n \arrow[u, "\varphi"] \arrow[r, dashed, "A"] & \mathbb{R}^m \end{tikzcd}, \qquad \begin{tikzcd} U \arrow[r, "f"] & V \arrow[d, "(\psi')^{-1}"] \\ \mathbb{R}^n \arrow[u, "\varphi'"] \arrow[r, dashed, "B"] & \mathbb{R}^m \end{tikzcd}$$

ができます。続いて

- 基底 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ から基底 $(\boldsymbol{f}'_1, \boldsymbol{f}'_2, \ldots, \boldsymbol{f}'_n)$ への変換行列 P
- 基底 $(\boldsymbol{g}_1, \boldsymbol{g}_2, \ldots, \boldsymbol{g}_m)$ から $(\boldsymbol{g}'_1, \boldsymbol{g}'_2, \ldots, \boldsymbol{g}'_m)$ への変換行列 Q

を考えます。するとさらに 2 つ

$$\begin{tikzcd} & V & \\ \mathbb{R}^n \arrow[ur, "\varphi"] \arrow[rr, "P^{-1}"] & & \mathbb{R}^n \arrow[ul, "\varphi'"'] \end{tikzcd}, \qquad \begin{tikzcd} & V & \\ \mathbb{R}^m \arrow[ur, "\psi"] \arrow[rr, "Q^{-1}"] & & \mathbb{R}^m \arrow[ul, "\psi'"'] \end{tikzcd}$$

という図式ができます[*6]。こうしてできた 4 つの図式を合体させると、こうなります:

$$\begin{tikzcd} \mathbb{R}^n \arrow[rrr, "A"] \arrow[dr, "\varphi"] \arrow[dd, "P^{-1}"'] & & & \mathbb{R}^m \arrow[dd, "Q^{-1}"] \\ & U \arrow[r, "f"] & V \arrow[ur, "\psi^{-1}"] \arrow[dr, "(\psi')^{-1}"'] & \\ \mathbb{R}^n \arrow[ur, "\varphi'"] \arrow[rrr, "B"'] & & & \mathbb{R}^m \end{tikzcd}$$

複数の回り方があるところは、どう回っても同じです。だから左下の $\mathbb{R}^n$ から右下の $\mathbb{R}^m$ への写像 $B \colon \mathbb{R}^n \to \mathbb{R}^m$ は、一旦 $P \colon \mathbb{R}^n \to \mathbb{R}^n$ で上に行き、そこから A で右上の $\mathbb{R}^m$ に行き、さらに Q^{-1} で右下に移動するのと同じです。もうちょっとちゃんと書くと

$$B = (\psi')^{-1} \circ f \circ \varphi' = (Q^{-1} \circ \psi^{-1}) \circ f \circ (\varphi \circ P) = Q^{-1} \circ A \circ P$$

[*6] この図式は「2 通りに回れるところは、どっちを回っても同じ結果になる」という意味です。このような図式を**可換図式**といいます。

という変形をやっています。上にある図式とにらめっこしながら、どういう経路をたどっているのか、一度追いかけてみてください。これで基底の変換に応じて、行列表示が $B = Q^{-1}AP$ という公式で置き換わることが分かりました[*7]。

ちなみにこの図式は「座標変換で行列表示は色々変化するが、その中心にいるのは $f\colon U \to V$ である」と言っているように見えます。特に P や Q として基本行列を持ってくると、B は A に行/列基本変形をした結果になります。僕たちは上手い基本変形で行列の像や核を求める計算をしましたが、それは「真ん中にいる f を色々な角度から眺めて、その性質を探ろうとしていた」と言うことができるのです。

3.2 一般化された内積

問 5 は、区間 $[0, 2\pi]$ の適当な値を代入したり、微分したりとかで解くことができます。ですが今回はちょっとやり方を変えて、積分を使って解いてみましょう。

■問 5 の解答　W は $\{1, \cos x, \sin x, \cos 2x, \sin 2x\}$ が張る空間である。この集合が 1 次独立であることを示して、$\dim W = 5$ をいう。

step 1. まず、正の整数 $m, n \in \mathbb{Z}_{>0}$ に対し

$$\frac{1}{\pi}\int_0^{2\pi} \cos mx \cos nx \, dx = \frac{1}{\pi}\int_0^{2\pi} \sin mx \sin nx \, dx = \delta_{mn} := \begin{cases} 1 & (m = n) \\ 0 & (m \neq n) \end{cases}$$

$$\frac{1}{\pi}\int_0^{2\pi} \cos mx \sin nx \, dx = \frac{1}{\pi}\int_0^{2\pi} \cos mx \, dx = \frac{1}{\pi}\int_0^{2\pi} \sin mx \, dx = 0$$

が成り立つ[*8]ことを確認しておく[*9]。

step 2. 次に $C^0\big([0, 2\pi]\big)$ を、閉区間 $[0, 2\pi]$ 上の連続函数全体の集合とする[*10]。$f, g \in C^0\big([0, 2\pi]\big)$ に対し

$$(f, g) := \frac{1}{\pi}\int_0^{2\pi} f(x)g(x)dx$$

と定める。すると全ての $f_1, f_2, g \in C^0\big([0, 2\pi]\big)$ と $\alpha, \beta \in \mathbb{R}$ に対し

$$\begin{aligned}(\alpha f_1 + \beta f_2, g) &= \frac{1}{\pi}\int_0^{2\pi} \big(\alpha f_1(x) + \beta f_2(x)\big)g(x)dx = \frac{\alpha}{\pi}\int_0^{2\pi} f_1(x)g(x)dx + \frac{\beta}{\pi}\int_0^{2\pi} f_2(x)g(x)dx \\ &= \alpha(f_1, g) + \beta(f_2, g)\end{aligned}$$

が成り立つ。つまり g を止めたとき、写像 $f \mapsto (f, g)$ は線型である。さらにこの記号を使うと、さっき示した三角函数の公式は

$$(\cos mx, \cos nx) = (\sin mx, \sin nx) = \delta_{mn}, \quad (\cos mx, \sin nx) = 0$$

と表せる。

step 3. これで準備ができたので、$\{1, \cos x, \sin x, \cos 2x, \sin 2x\}$ の 1 次独立性を示す。$f(x) := a + b\cos x + c\sin x +$

---

*7 この公式は重要ですが、式を丸覚えするのはやめた方がいいです。というのも「変換行列はどっち向きだっけ？」とかを思い出そうとすると、大体どこかでミスするからです。上に書いた図式を再現できるようになっておきましょう。

*8 ここで δ_{nm} は、$n = m$ なら 1、そうでないなら 0 という意味です。この記号は「Kronecker の δ」と呼ばれ、たとえば単位行列を $I = (\delta_{ij})$ と表したりするときに使います。

*9 この式は大学入試でも割とよく見かける式なので、証明は略します。三角函数の積和公式を使って変形した後、$m = n$ かそうでないかに応じて愚直に積分すれば求まります。

*10 しばしば $\mathbb{R}$ 上の連続函数のことを「C^0 級函数」というので、その記法を使いました。

$d\cos 2x + e\sin 2x \equiv 0$ とおく。このとき任意の函数 $g \in C^0\big([0,2\pi]\big)$ に対し $(f,g) = (0,g) = 0$ が成り立つので

$$\begin{aligned}
0 &= \big(f(x),1\big) &&= a(1,1) &&+ b(\cos x,1) &&+ c(\sin x,1) &&+ d(\cos 2x,1) &&+ e(\sin 2x,1) &&= a\\
0 &= \big(f(x),\cos x\big) &&= a(1,\cos x) &&+ b(\cos x,\cos x) &&+ c(\sin x,\cos x) &&+ d(\cos 2x,\cos x) &&+ e(\sin 2x,\cos x) &&= b\\
0 &= \big(f(x),\sin x\big) &&= a(1,\sin x) &&+ b(\cos x,\sin x) &&+ c(\sin x,\sin x) &&+ d(\cos 2x,\sin x) &&+ e(\sin 2x,\sin x) &&= c\\
0 &= \big(f(x),\cos 2x\big) &&= a(1,\cos 2x) &&+ b(\cos x,\cos 2x) &&+ c(\sin x,\cos 2x) &&+ d(\cos 2x,\cos 2x) &&+ e(\sin 2x,\cos 2x) &&= d\\
0 &= \big(f(x),\sin 2x\big) &&= a(1,\sin 2x) &&+ b(\cos x,\sin 2x) &&+ c(\sin x,\sin 2x) &&+ d(\cos 2x,\sin 2x) &&+ e(\sin 2x,\sin 2x) &&= e
\end{aligned}$$

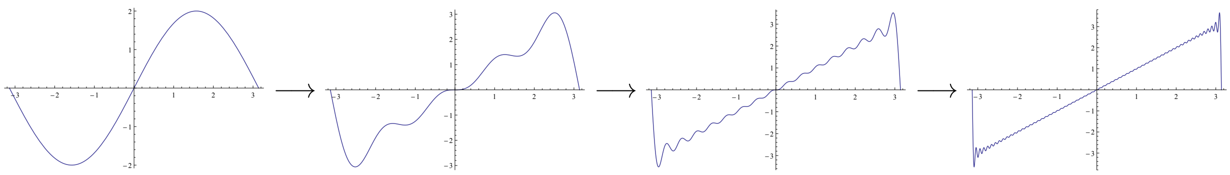
となる。これで 1 次独立性が言えた。 ■

■Fourier 級数展開　いまの証明からほとんど明らかですが、cos や sin の中に入るのは別に x や $2x$ でなくとも構いません。積分を使って定義された (f,g) を使えば、$\{1,\cos x,\sin x,\cos 2x,\sin 2x,\cos 3x,\sin 3x,\ldots\}$ は 1 次独立だと示せます[*11]。では、$C^0\big([0,2\pi]\big)$ の基底にはなるでしょうか？

この答えは、普通の基底の定義に照らせば "No" です。三角函数ではない函数はいくらでもあります。たとえば多項式がその典型的な例です。ですから $\{1,\cos x,\sin x,\cos 2x,\sin 2x,\cos 3x,\sin 3x,\ldots\}$ の元を単純に足し算するだけでは、$C^0\big([0,2\pi]\big)$ の全ての元を作ることはとてもできません。ですが単純な足し算ではなく「無限和」、つまり函数の和の極限を考えれば、$C^0\big([0,2\pi]\big)$ の函数たちをいくらでも良く近似できることが知られています。たとえば直線 $y=x$ を区間 $[-\pi,\pi]$ で[*12]近似する式は

$$x = \sum_{m=1}^{\infty} \sin mx \cdot \frac{1}{\pi}\int_{-\pi}^{\pi} t\sin mt\,dt = \sum_{m=1}^{\infty}(-1)^{m+1}\frac{2}{m}\sin mx$$

です。問 5 を解くときに「$\big(f(x),\cos x\big)$ がちょうど $f(x)$ における $\cos x$ の係数になる」などの事実を使いましたが、それと同じようにして、x を $\{1,\cos x,\sin x,\cos 2x,\sin 2x,\ldots\}$ の「線型結合」で表したときの係数を、積分で取り出しています。途中までの和をグラフにしてみると、確かに近似になっていることが分かりますね。



(e) $m=1$ まで　(f) $m=4$ まで　(g) $m=16$ まで　(h) $m=64$ まで

図 12.1　閉区間 $[-\pi,\pi]$ における $y=x$ の Fourier 級数展開の様子

このように、函数を $\{1,\cos mx,\sin mx \mid m\in\mathbb{Z}_{>0}\}$ の無限級数で表すことを **Fourier 級数展開**といいます。音波や交流など「波」を扱う際、Fourier 級数展開が基本的な道具になります。詳細な計算法は別として、名前だけは知っておいてください。

■内積の一般化　さて、今の話がどことなく「内積」に似ていることに気が付きましたか？ 平面のベクトルに対しては

$$\left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\boldsymbol{e}_1\right)=x_1,\quad \left(\begin{pmatrix}x_1\\x_2\end{pmatrix},\boldsymbol{e}_2\right)=x_2$$

---

*11 より正確には「この集合に含まれるどんな有限部分集合を持ってきても、その元たちが 1 次独立になる」という意味です。

*12 計算の都合上、区間 $[0,2\pi]$ を $-\pi$ だけずらして $[-\pi,\pi]$ 上で考えます。元の区間 $[0,2\pi]$ でも計算できるのですが、今回は $[-\pi,\pi]$ の方が計算しやすかったのです。

というように、標準基底との内積を取るとそれぞれの成分が飛び出してきます。今の函数の場合も

$$f(x)=a_0+\sum_{m=1}^{\infty}a_m\sin 2mx+\sum_{m=1}^{\infty}b_m\cos 2mx$$

と表したとき、その「$\sin 2mx$ 成分」である a_m や「$\cos 2mx$ 成分」である b_m が

$$a_m=\bigl(f(x),\sin 2mx\bigr),b_m=\bigl(f(x),\cos 2mx\bigr)$$

という式で取り出せています。

いまはいきなり函数空間の例でやってしまいましたが、もっと手頃な有限次元線型空間の場合でも、同じようなことができます。線型空間に「内積」と呼ぶべきものを定義したり、さらに「基底との内積で成分を取り出す」といった操作を考えたりします。このような「内積」を付け加えた線型空間を**計量線型空間**といいます。また、その内積に関して「長さが 1」で「異なるベクトルがお互いに直交する」という条件を満たす基底を、**正規直交基底**といいます。

これまでに出てきた函数のなす線型空間や多項式のなす線型空間では、その上にいろいろな内積を入れることができ、内積に応じて様々な面白いことが起こります。そうした内積の詳しい取扱いを、来学期に行います。

4　最終回の解答

もう試験が来てしまうので、最終回の授業で配布された問題についても一部解答を載せておきます。ただしレポートとして後から回収するので、丸写しすれば済むような書き方をするわけにもいきません[*13]。そこで、

- 計算問題については、答えだけ
- 証明問題については、理論的に綺麗な解き方

を書いてみます。これがすらすら読めるなら、きっと試験も何とかなると思います。チャレンジしてみてください。

■線型空間の直和分解　U を線型空間とし、$V,W\subset U$ をその部分空間とします。そして「任意の $\boldsymbol{u}\in U$ が、$\boldsymbol{u}=\boldsymbol{v}+\boldsymbol{w}$ $(\boldsymbol{v}\in V,\boldsymbol{w}\in W)$ の形にただ一通りに書ける」という条件が成り立つとき、U は V と W の**直和**であるといい、$U=V\oplus W$ と書きます。

直感的に言えば、これは「U を V 方向と W 方向に分ける」ことを表しています。ですから直和の条件は、次のようにも言い換えられます。

- V の基底と W の基底を連結すると、U 全体の基底が得られる。
- $U=V+W$ かつ $V\cap W=\{\boldsymbol{0}\}$ が成り立つ。

3 つ以上の空間の直和についても定義は同様です。

■問 1　$\operatorname{Ker} f=\mathbb{R}\,{}^t(1,-2,1)$ ∎

■問 2　(1) $\operatorname{Ker} f=\mathbb{R}\,{}^t(1,-3,1)$ (2) $\operatorname{Im} f=\mathbb{R}\,{}^t(1,0,1)\oplus\mathbb{R}\,{}^t(0,1,2)$ ∎

■問 3　(1) $\operatorname{Ker} f=\mathbb{R}\,{}^t(2,2,-3,-3)$ (2) $\operatorname{Im} f=\mathbb{R}\,{}^t(1,0,0,1)\oplus\mathbb{R}\,{}^t(0,1,0,0)\oplus\mathbb{R}\,{}^t(0,0,1,2)$ ∎

■問 4 と問 5　$f,g\colon \mathrm{Mat}_2(\mathbb{R})\to\mathrm{Mat}_2(\mathbb{R})$ を、$f(X):=X+{}^tX,\ g(X):=X-{}^tX$ と定める。このとき直和分解 $\mathrm{Mat}_2(\mathbb{R})=\mathrm{Sym}_2(\mathbb{R})\oplus\mathrm{Alt}_2(\mathbb{R})$ を考え[*14]、それぞれの空間に $f,\ g$ を制限すると

[*13] この略解のギャップをきちんと埋めてレポートにするのは構いませんが、そのまま書き写すのはやめてくださいね。
[*14] $\mathrm{Sym}_2(\mathbb{R})$ と $\mathrm{Alt}_2(\mathbb{R})$ は、それぞれ対称行列全体のなす集合と交代行列全体のなす集合です。103 ページで一度出てきました。

- $f|_{\mathrm{Sym}_2(\mathbb{R})} = 2\,\mathrm{id}\colon \mathrm{Sym}_2(\mathbb{R}) \to \mathrm{Sym}_2(\mathbb{R}),\ f|_{\mathrm{Alt}_2(\mathbb{R})} = 0$
- $g|_{\mathrm{Sym}_2(\mathbb{R})} = 0,\ f|_{\mathrm{Alt}_2(\mathbb{R})} = 2\,\mathrm{id}\colon \mathrm{Alt}_2(\mathbb{R}) \to \mathrm{Alt}_2(\mathbb{R})$

となっている。これで $\mathrm{Ker}\, f = \mathrm{Im}\, g = \mathrm{Alt}_2(\mathbb{R}),\ \mathrm{Im}\, f = \mathrm{Ker}\, g = \mathrm{Sym}_2(\mathbb{R})$ が分かる。 ■

■問 6 と問 7 線型写像 $\mathrm{tr}\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathbb{R}$ について、$\mathrm{tr}(E/2) = 1 \in \mathbb{R}$ なので、$\mathrm{Im\ tr} = \mathbb{R}$ である。よって $\mathfrak{sl}_2(\mathbb{R}) := \mathrm{Ker\, tr}$ とおく[*15]と、$\dim \mathfrak{sl}_2(\mathbb{R}) = \dim \mathrm{Mat}_2(\mathbb{R}) - \dim \mathbb{R} = 4 - 1 = 3$ である。そして

$$E := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad F := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

という 1 次独立な元が $\mathfrak{sl}_2(\mathbb{R})$ の中に取れる。これが $\mathfrak{sl}_2(\mathbb{R})$ の基底を与える。

さて $X \in \mathrm{Mat}_2(\mathbb{R})$ に対し、線型写像 $\mathrm{ad}_X\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathrm{Mat}_2(\mathbb{R})$ を $\mathrm{ad}_X(Y) := [X, Y] = XY - YX$ と定める。いかなる X を取ってきても、この ad_X は全射にならない。なぜなら $\mathrm{tr}\,\mathrm{ad}_X(Y) = \mathrm{tr}(XY - YX) = \mathrm{tr}\, XY - \mathrm{tr}\, YX = 0$ なので、ad_X の像は $\mathfrak{sl}_2(\mathbb{R}) = \mathrm{Ker\, tr} \subsetneq \mathrm{Mat}_2(\mathbb{R})$ に含まれるからである。

特に $X = H$ の場合を考える。このとき $\mathrm{ad}_H(E) = [H, E] = 2E,\ \mathrm{ad}_H(F) = [H, F] = -2F \in \mathrm{Im\ ad}_H$ は 1 次独立である。また 2 次単位行列 I と H とは $\mathrm{Mat}_2(\mathbb{R})$ の中で 1 次独立であり、ともに $\mathrm{ad}_H(H) = [H, H] = O$, $\mathrm{ad}_H(I) = [H, I] = O$ を満たす。よって $\mathrm{Ker\, ad}_H$ の中に I, H という 2 つの 1 次独立な元が取れた。

ここまでで $\dim \mathrm{Im\ ad}_H \geq 2,\ \dim \mathrm{Ker\, ad}_H \geq 2$ が分かったが、一方次元定理より $\dim \mathrm{Ker\, ad}_H + \dim \mathrm{Im\ ad}_H = \dim \mathrm{Mat}_2(\mathbb{R}) = 4$ である。よって $\dim \mathrm{Im\ ad}_H = \dim \mathrm{Ker\, ad}_H = 2$ であり、基底が求まったことになる。 ■

■問 10 $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ を線型空間 V の基底とする。このとき $\boldsymbol{f}_i^\vee \in V^* := \mathrm{Hom}_{\mathbb{R}}(V, \mathbb{R})$ を

$$\boldsymbol{f}_i^\vee(\boldsymbol{f}_j) := \begin{cases} 1 & (i = j) \\ 0 & (i \neq j) \end{cases}$$

と定める[*16]と、$\left(\boldsymbol{f}_1^\vee, \ldots, \boldsymbol{f}_n^\vee\right)$ が V^* の基底となる[*17]。 ■

[*15] n 次正方行列の場合でも、$\mathrm{tr}\colon \mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R}$ の核を $\mathfrak{sl}_n(\mathbb{R})$ と書き、「n 次特殊線型群 $SL_n(\mathbb{R})$ の Lie 環」といいます。

[*16] V^* のことを**双対（そうつい）空間**といいます。「そうたい」ではありませんので、読み方に気を付けてください。

[*17] この基底は $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ の**双対基底**といいます。標語的にいうと「V の座標に応じて、V^* の自然な座標が定まる」という感じです。ちょっと大事な話なので、来学期の頭に配るプリントで詳しく解説するかもしれません。

線形代数学① 第 13 回　　おまけ

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 7 月 22 日 1 限

1　最初に

みなさんこんにちは。夏休み、いかがでしたか？

今回のテーマは S2 タームの補足ということで、まずは行列の核と像について簡単に復習をします。夏休みを挟んで色々忘れてしまった人は、ちょっとリハビリをしてください。A セメスターの頭に行列式を勉強する際にも、後々で基本変形を使います。また、S2 ターム最後の授業で出された課題には、実は数学的に面白いネタが潜んでいました。そこで夏休み明け初回のこのプリントでは、ちょっと発展的なお話をいくつか紹介したいと思います。その後で最後に、この先の授業の展望を紹介します。

なお、今回は過去のプリントを参照する箇所が何ヶ所かあります。繰り返しになりますが、プリントをなくしてしまった人は

`https://github.com/HideakiHosaka/2015_linear_algebra/raw/master/2015linear_algebra.pdf`

からダウンロードしてください。ちょっと重たいので、PC からダウンロードすることをお勧めします。

2　復習: 行列の核と像の求め方

最初に、問 1 から問 3 までの計算問題について復習しておきましょう。これらの問題は、いずれも「与えられた行列 A に対し、A の核と像の基底を求める」という問題です。そして夏学期に勉強したのは

- 行列 A を<u>行</u>基本変形すると<u>核</u>の基底が求まる
- 行列 A を<u>列</u>基本変形すると<u>像</u>の基底が求まる

という事実でした。列基本変形をすべき場面で行基本変形をしてしまっていた人が少なからずいたので、気を付けてください。行基本変形と列基本変形のやり方については、昔配った「掃き出し法」の回のプリント p. 85 と「線型写像と行列」の回のプリント p. 113 を復習してください。

「行と列どっちだっけ」とこんがらがったときは、

- 核を求めること = 連立一次方程式を解くこと
- 連立一次方程式に対する消去法 = (拡大) 係数行列の行基本変形

を思い出しましょう。これを覚えていれば、「核を求めるときは行基本変形だ」と確信をもって言えるようになります。

また核と像両方の基底を求める問 2, 3 では、次元定理を使って計算を避けている人がいました。良い作戦だと思います。ただ次元定理は、次元に関する情報しか与えてくれません。たとえば問 3 なら、(1) で $\operatorname{Ker} A$ を求めれば $\dim \operatorname{Im} A = 3$ は分かります。$\operatorname{Im} A$ は A の 4 本の列ベクトルで生成されますから、そのうち 1 次独立な 3 本を見つければ、それが $\operatorname{Im} A$ の基底になります。でも「どの 3 本が当たりか」は次元を見ていても分かりません。きちんと 1 次独立性を確かめる必要があります。

■問 1–問 3 の解答

問 1 $(1, -2, 1)$
問 2 (1) $(1, -3, 1)$ (2) $(1, 0, 1), (0, 1, 2)$
問 3 (1) $(2, 2, -3, -3)$ (2) $(1, 0, 0, 1), (0, 1, 0, 0), (0, 0, 1, -2)$ ■

3 線型写像のなす線型空間

今回はいくつかの問題で、行列のなす線型空間や、それらを定義域とする線型写像が登場します。「線型写像のなす線型空間で定義された線型写像」とか書いていると訳が分からなくなりそうですが、よく出てくるので、この機会にぜひ扱いに慣れてください。

3.1 行列のなす線型空間

線型空間の中で一番典型的なものは、列ベクトルの空間 $\mathbb{R}^n$ です。これが線型空間であることは既に知っているはずです。一方 (m, n) 型行列全体のなす集合を $\mathrm{Mat}_{m,n}(\mathbb{R})$ と書いていました。これもまた線型空間になります。というのも、$\mathrm{Mat}_{m,n}(\mathbb{R})$ には和とスカラー倍が定義されています。そして $\mathrm{Mat}_{m,n}(\mathbb{R})$ と $\mathbb{R}^{mn}$ とを見比べたとき、和とスカラー倍に関してだけいえば、成分が長方形に並んでいるか縦一列に並んでいるかの違いしかありません。ですから $\mathrm{Mat}_{m,n}(\mathbb{R})$ が線型空間の公理を満たすことは明らかでしょう。

そして $\mathbb{R}^n$ には**標準基底**と呼ばれる基底がありました。i 番目の成分だけが 1 で他の成分が全て 0 であるようなベクトルを $\boldsymbol{e}_i$ と書くとき、$(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ が標準基底でした。各軸の方向に「単位ベクトル」を取ってきて、それを集めたものですね。同様にして $\mathrm{Mat}_{m,n}(\mathbb{R})$ の場合も「1 つの成分だけが 1 で他は 0」という行列[*1]を全部集めてきて 1 列に並べれば、基底がます。そのことを確認しましょう。

■問 8 の解答 (i, j) 成分だけが 1 で他の成分が全て 0 であるような行列を $E_{ij} \in \mathrm{Mat}_{m,n}(\mathbb{R})$ と書く[*2]。一般に行列 $A = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathrm{Mat}_{m,n}(\mathbb{R})$ に対し

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{n2} & \cdots & a_{mn} \end{pmatrix} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} E_{ij}$$

が成り立つ。この式から、全ての行列は E_{ij} $(1 \le i \le m, 1 \le j \le n)$ の 1 次結合で書けることが分かる。つまり E_{ij} たちは $\mathrm{Mat}_{m,n}(\mathbb{R})$ を生成する。また $A = O$ であれば、全ての a_{ij} が 0 になる。よって E_{ij} たちは 1 次独立である。よって E_{ij} たちを 1 列に並べた $(E_{11}, \ldots, E_{1n}, E_{21}, \ldots, E_{2n}, \ldots, E_{m1}, \ldots, E_{mn})$ は $\mathrm{Mat}_{m,n}(\mathbb{R})$ の基底である。 ■

3.2 線型写像のなす線型空間

次に、行列と線型写像の関係を復習しておきましょう。線型空間 V, W の間の写像 $f\colon V \to W$ が線型写像であるとは、任意の $\boldsymbol{u}, \boldsymbol{v} \in V$ と $\alpha \in \mathbb{R}$ に対し $f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$, $f(\alpha \boldsymbol{u}) = \alpha f(\boldsymbol{u})$ が成り立つことでした[*3]。そして

- 行列のかけ算は数ベクトル空間の間の線型写像である
- 線型空間を基底によって数ベクトル空間と同一視すると、線型写像は行列として表される

[*1] こういう行列を**行列単位**といいます。
[*2] 元の問 8 は n 次正方行列が問題になっていますが、別に縦横のサイズが違っても問題の解き方は全く変わらないので、一般化して解きます。
[*3] S2 ターム初回のプリント p. 59 にて、図式つきで線型写像の定義を説明しています。どうして「線型写像が、線型空間の間の良い写像と言えるのか」を忘れてしまった人は、参考にしてください。

という事実がありました。基底によって、線型写像は行列とぴったり対応しています。だから V から W への線型写像全体の集合を $\mathrm{Hom}(V, W)$ と書くことにすれば、行列の集合 $\mathrm{Mat}_{m,n}(\mathbb{R})$ が線型空間だったのと全く同様に、$\mathrm{Hom}(V, W)$ も線型空間になります。

この線型空間の構造についても、復習しましょう。行列の和とスカラー倍は、A, B を行列、$\boldsymbol{u}$ を数ベクトル、α をスカラーとして

$$(A+B)\boldsymbol{u} = A\boldsymbol{u} + B\boldsymbol{u}, (\alpha A)(\boldsymbol{u}) = \alpha(A\boldsymbol{u})$$

という式を満たしていました。ですから一般の線型写像 $f, g \in \mathrm{Hom}(V, W)$ の場合にも、

$$(f+g)(\boldsymbol{u}) := f(\boldsymbol{u}) + g(\boldsymbol{u}), (\alpha f)(\boldsymbol{u}) := \alpha f(\boldsymbol{u})$$

によって写像 $f+g, \alpha f\colon V \to W$ を定めれば、これらは線型写像になります。こうして $\mathrm{Hom}(V, W)$ に和とスカラー倍が定まります。線型写像と行列が大体同じだと思えば、$\mathrm{Hom}(V, W)$ が線型空間になることは納得いくでしょう[*4]。

3.3 行列の線型空間を定義域とする線型写像

ここでもう一度確認しますが、(m, n) 型行列全体の集合 $\mathrm{Mat}_{m,n}(\mathbb{R})$ は線型空間でした。また線型写像は、線型空間を定義域として、和とスカラー倍について良い振る舞いをする写像のことでした。したがって「$\mathrm{Mat}_{m,n}(\mathbb{R})$ を定義域とする」線型写像を考えることができます。その例をいくつか見てみましょう。

なお、いつもは「線型写像が分からなくなったら行列に戻ることが大事だ」と言っていましたが、ここでは**一旦線型写像を行列と切り離して考える**方がいいです。線型写像の定義だけを使って、話を進めます。

■問 6 (1) の解答　行列のトレースは

$$\mathrm{tr}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} := x_{11} + x_{22}$$

で定まっていた。これは 2 次正方行列に対して数を対応させているので、$\mathrm{tr}\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathbb{R}$ という写像が定まっていることになる[*5]。そして

$$\begin{aligned}
\mathrm{tr}\left\{\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} + \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix}\right\} &= \mathrm{tr}\begin{pmatrix} x_{11}+y_{11} & x_{12}+y_{12} \\ x_{21}+y_{22} & x_{22}+y_{22} \end{pmatrix} = x_{11}+y_{11}+x_{22}+y_{22} \\
&= (x_{11}+x_{22}) + (y_{11}+y_{22}) = \mathrm{tr}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} + \mathrm{tr}\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \\
\mathrm{tr}\left\{\alpha\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}\right\} &= \mathrm{tr}\begin{pmatrix} \alpha x_{11} & \alpha x_{12} \\ \alpha x_{21} & \alpha x_{22} \end{pmatrix} = \alpha x_{11} + \alpha x_{22} = \alpha(x_{11}+x_{22}) = \alpha\,\mathrm{tr}\begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}
\end{aligned}$$

となっているので、$\mathrm{tr}\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathbb{R}$ は線型写像である。 ■

■問 9 の解答　n 次正方行列 A を左からかける写像 $m_A\colon \mathrm{Mat}_n(\mathbb{R}) \to \mathrm{Mat}_n(\mathbb{R})$ は線型写像である。実際、任意の $X, Y \in \mathrm{Mat}_n(\mathbb{R})$ と $\alpha \in \mathbb{R}$ に対し

$$m_A(X+Y) = A(X+Y) = AX + AY = m_A(X) + m_A(Y), m_A(\alpha X) = A(\alpha X) = \alpha AX = \alpha m_A(X)$$

が成り立っている。またトレース $\mathrm{tr}\colon \mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R}$ も線型写像である。t_A は合成写像として $t_A := \mathrm{tr} \circ m_A$ のように書けるので、やはり線型である。 ■

[*4] S2 タームの一番最初に配ったプリントの p. 69 には $\mathrm{Hom}(V, W)$ が線型空間になることのきちんとした証明もつけていますので、合わせて参考にしてください。

[*5] 見ればすぐ分かりますが、以下での議論は n 次元でも全く同様に成り立ちます。

■問 7 の解答　(1) 2 次正方行列 $H \in \mathrm{Mat}_2(\mathbb{R})$ を

$$H := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

で定め、写像 h を $h(X) = [H, X] := HX - XH$ で定める。X が 2 次正方行列なら HX, XH は共に 2 次正方行列だから $HX - XH$ もそうであり、結果として h は $h\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathrm{Mat}_2(\mathbb{R})$ という写像を定めている。そして任意の $X \in \mathrm{Mat}_2(\mathbb{R})$ と $\alpha \in \mathbb{R}$ に対し

$$\begin{aligned} h(X+Y) &= H(X+Y) - (X+Y)H = HX - XH + HY - YH = h(X) + h(Y) \\ h(\alpha X) &= H(\alpha X) - (\alpha X)H = \alpha(HX - XH) = \alpha h(X) \end{aligned}$$

が成り立つので、h は線型写像である。
(3) 2 次の単位行列を I と書く。このとき $h(I) = HI - IH = H - H = O$ である。また $h(H) = H^2 - H^2 = O$ である。これより $I, H \in \mathrm{Ker}\, h$ が従う。一方

$$E := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

と定めると、

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = bE + \frac{a+d}{2}I + \frac{a-d}{2}H + cF$$

と書ける。これと $h(E) = [H, E] = 2E$, $h(F) = [H, F] = -2F$ より、h の線型性を使って

$$h\left(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\right) = 2bE - 2cF$$

と計算できる。よって $\mathrm{Im}\, h$ は E, F の張る $\mathrm{Mat}_2(\mathbb{R})$ の部分空間に一致している。また E, F は 1 次独立なので、これらが $\mathrm{Im}\, h$ の基底になっていることが分かる。
(2) 次元定理と (3) から $\dim \mathrm{Ker}\, h = \dim \mathrm{Mat}_2(\mathbb{R}) - \dim \mathrm{Im}\, h = 4 - 2 = 2$ と分かる。一方 $I, H \in \mathrm{Ker}\, h$ は既に示しており、これらは 1 次独立である。故に I, H が $\mathrm{Ker}\, h$ の基底をなす。 ■

■問 7 の h の行列表示　さて、今の問 7 を解く途中で出てきた

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = bE + \frac{a-d}{2}H + cF + \frac{a+d}{2}I$$

を思い出しましょう。この式より (E, H, F, I) は $\mathrm{Mat}_2(\mathbb{R})$ の生成系です[*6]。一方 $\mathrm{Mat}_2(\mathbb{R})$ は 4 次元ですから、この (E, H, F, I) は $\mathrm{Mat}_2(\mathbb{R})$ の基底になります。この基底 (E, H, F, I) について、h を行列表示してみましょう。

　線型写像の行列表示とは、基底を用いて定義域と値域の線型空間を数ベクトル空間と同一視したとき、h が行列としてどう表されるかを見たものです。今の場合 $\mathrm{Mat}_2(\mathbb{R})$ の基底として (E, H, F, I) を取っていますから

$$E \longleftrightarrow \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, H \longleftrightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, F \longleftrightarrow \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, I \longleftrightarrow \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

[*6] 基底として (E, H, F, I) を取ったのは、TA の個人的な数学的趣味によるものです。上手い基底を取ったおかげで、行列表示したときに対角行列でています。こういうとき、線型写像が**対角化されている**といいます。さらに言うと Lie 環 $\mathfrak{sl}_2$ の「表現の既約分解」「最高ウェイト表現」というものに対応した基底の並べ方をしているのですが、今はここまで気にしなくていいです。

という読み替えで $\mathrm{Mat}_2(\mathbb{R})$ と $\mathbb{R}^4$ を対応させます。これに合わせて $h(E)=2E$, $h(F)=-2F$, $h(H)=h(I)=O$ を読み替えると、h は

$$\begin{pmatrix}1\\0\\0\\0\end{pmatrix}\longleftrightarrow E\overset{h}{\longmapsto}2E\longleftrightarrow\begin{pmatrix}2\\0\\0\\0\end{pmatrix},\qquad \begin{pmatrix}0\\1\\0\\0\end{pmatrix}\longleftrightarrow H\overset{h}{\longmapsto}O\longleftrightarrow\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

$$\begin{pmatrix}0\\0\\1\\0\end{pmatrix}\longleftrightarrow F\overset{h}{\longmapsto}-2F\longleftrightarrow\begin{pmatrix}0\\0\\-2\\0\end{pmatrix},\qquad \begin{pmatrix}0\\0\\0\\1\end{pmatrix}\longleftrightarrow I\overset{h}{\longmapsto}O\longleftrightarrow\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$$

という対応を与えることが分かります。この対応を実現する行列は、h の行先となる列ベクトルを並べた

$$\begin{pmatrix}2&0&0&0\\0&0&0&0\\0&0&-2&0\\0&0&0&0\end{pmatrix}$$

となります。

ここで h は、元々 $\mathrm{Mat}_2(\mathbb{R})$ の

$$\begin{pmatrix}1&0\\0&-1\end{pmatrix}$$

という行列を使い、$h(X)=HX-XH$ という式で定義された写像でした。ところが上で導いた h の行列表示は、定義に使った行列 H や、あるいは基底の行列 E,F,I と特に関係を見出せません。これは偶然というわけではなく、一般に行列の空間 $\mathrm{Mat}_{m,n}(\mathbb{R})$ を定義域とする線型写像 f について、**「f の中に放り込まれる行列」と「f の行列表示」との間に関係はありません**。

さっき「線型写像と行列を切り離した方が良い」といったのは、関係ない 2 つの行列がごっちゃになるのを防ぐためでした。一度切り分け方が分かってしまえば、こんがらがることは無くなると思います。もし途中でつっかえたら、落ち着いて問 7 を最初から追いかけてください。

また今は h の行列表示を求めてみましたが、一般の場合に行列表示を求める必要があるかというと、そうでもありません。たとえば n 次行列のなす線型空間の次元は $\dim\mathrm{Mat}_n(\mathbb{R})=n^2$ なので、n が 3 とか 4 になるだけで空間の次元はすごく大きくなります。そうした空間を定義域とする線型写像を考えると、行列表示をしたら行列の 1 辺が 9 列とか 16 列になってしまい、書くだけで大変です。その上そもそも上に書いた通り、行列表示をしたからといって意味のある情報が出てくるわけでもありません[*7]。ですので行列のなす線型空間が絡む残りの問題についても、行列表示は使わずに問題を解いてみます。

■問 6 (2), (3) の解答　2 次正方行列 $X=(x_{ij})$ について

$$\mathrm{tr}\begin{pmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{pmatrix}=x_{11}+x_{22}$$

なので、トレースが 0 であることと $x_{22}=-x_{11}$ は同値である。よって全ての $X\in\mathrm{Ker}\,\mathrm{tr}$ は

$$\begin{pmatrix}x_{11}&x_{12}\\x_{21}&x_{22}\end{pmatrix}=x_{11}H+x_{12}E+x_{21}F$$

と表せる[*8]。また $x_{11}H+x_{12}E+x_{21}F=O$ ならば $x_{11}=x_{12}=x_{21}=0$ となる。よって (H,E,F) は 1 次独立でもある。これより $\mathrm{Ker}\,\mathrm{tr}$ の基底として (H,E,F) が取れる。

[*7] 行列表示の例として問 7 の h を選んだのは、「対角化」という意味のある例だったからです。

[*8] H,E,F は問 7 の解答に出てきた行列です。

また、任意の $a \in \mathbb{R}$ に対して

$$\operatorname{tr}\begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} = a$$

が成り立つので、$\operatorname{tr}\colon \operatorname{Mat}_2(\mathbb{R}) \to \mathbb{R}$ は全射である。特に $\operatorname{Im} \operatorname{tr} = \mathbb{R}$ は 1 次元で、その基底として 1 が取れる。 ■

4　線型空間の直和分解

今回の問 4, 5 には、線型空間の直和分解というものが見えています。これは一言で言えば「空間を 2 つの部分空間の方向に分ける」という操作に相当します。たとえば線型空間の基底を取ってくれれば、線型空間を各基底の方向にバラすことができます。でも時と場合によっては、基底を取ってまで細かくしなくても、ほどほどに空間をバラしたいことがあります。そこで登場するのが、直和分解という考え方です。

4.1　直和の定義

前回のプリントの再掲になりますが、直和の定義と性質を確認しておきましょう。U を線型空間とし、$V, W \subset U$ をその部分空間とします。そして「任意の $\boldsymbol{u} \in U$ が、$\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{w}$ $(\boldsymbol{v} \in V, \boldsymbol{w} \in W)$ の形にただ一通りに書ける」という条件が成り立つとき、U は V と W の**直和**である[*9]といい、$U = V \oplus W$ と書きます。

直感的に言えば、これは「U を V 方向と W 方向に分ける」ことを表しています。ですから直和の条件は、次のようにも言い換えられます。

- V の基底と W の基底を連結すると、U 全体の基底が得られる。
- $U = V + W$ かつ $V \cap W = \{\boldsymbol{0}\}$ が成り立つ。
- $V \cap W = \{\boldsymbol{0}\}$ かつ $\dim V + \dim W = \dim U$ が成り立つ。

3 つ以上の空間の直和についても定義は同様です。

この中でも最も典型的な例は基底でしょう。線型空間 V について、$(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が基底であることは、全ての $\boldsymbol{v} \in V$ が $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ の一次結合でただ一通りに書けることでした。一方、$\boldsymbol{f}_i$ の張る V の部分空間を $\mathbb{R}\boldsymbol{f}_i$ と書いていました。ですから $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ が基底であるという事実は、直和の記号を使って表せば $V = \mathbb{R}\boldsymbol{f}_1 \oplus \cdots \oplus \mathbb{R}\boldsymbol{f}_n$ となります。こんな感じで、問 1 から問 3 の答えを書き直してみましょう。

■問 1–問 3 の解答

問 1 $\operatorname{Ker} f = \mathbb{R}\,{}^t(1, -2, 1)$
問 2 (1) $\operatorname{Ker} f = \mathbb{R}\,{}^t(1, -3, 1)$ (2) $\operatorname{Im} f = \mathbb{R}\,{}^t(1, 0, 1) \oplus \mathbb{R}\,{}^t(0, 1, 2)$
問 3 (1) $\operatorname{Ker} f = \mathbb{R}\,{}^t(2, 2, -3, -3)$ (2) $\operatorname{Im} f = \mathbb{R}\,{}^t(1, 0, 0, 1) \oplus \mathbb{R}\,{}^t(0, 1, 0, 0) \oplus \mathbb{R}\,{}^t(0, 0, 1, -2)$ ■

4.2　直和分解の例

さて既に述べた通り、直和分解が真に役立つのは「基底よりもう少し大きいブロックに空間を分ける」という場合です。基底を持ち出して話が進むなら、わざわざ直和を持ち出す理由がないですからね。そこで行列のなす線型空間 $\operatorname{Mat}_n(\mathbb{R})$ で、典型的な直和分解の例を確認します。

[*9] 正確には、ここで定義した直和は「内部直和」といいます。この他に「外部直和」というものがあって、慣れてくると内部直和と外部直和をどっちも「直和」と略すようになります。事実としては内部直和と外部直和は自然に同型になるので、そんな深く気にすることはありません。でも他の場所で「この場所は、僕の知ってる直和と違うな？」と思ったときは、どっちの直和なのかを考えてみてください。

■問 4, 問 5 の解答　$f, g\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathrm{Mat}_2(\mathbb{R})$ を、$f(X) := X + {}^tX$, $g(X) := X - {}^tX$ と定める。このとき

- $f(X) = O$ と ${}^tX = -X$ は同値なので、$\mathrm{Ker}\, f = \mathrm{Alt}_2(\mathbb{R})$
- $g(X) = O$ と ${}^tX = X$ は同値なので、$\mathrm{Ker}\, g = \mathrm{Sym}_2(\mathbb{R})$

である*10。一方

- 任意の $X \in \mathrm{Mat}_2(\mathbb{R})$ に対し ${}^tf(X) = {}^t(X + {}^tX) = {}^tX + X = f(X)$ だから、$\mathrm{Im}\, f \subset \mathrm{Sym}_2(\mathbb{R})$
- $X \in \mathrm{Sym}_2(\mathbb{R})$ に対し、$f(X/2) = (X + {}^tX)/2 = X$ となるので、$\mathrm{Sym}_2(\mathbb{R}) \subset \mathrm{Im}\, f$

だから、$\mathrm{Im}\, f = \mathrm{Sym}_2(\mathbb{R})$ である。同様にして

- 任意の $X \in \mathrm{Mat}_2(\mathbb{R})$ に対し ${}^tg(X) = {}^t(X - {}^tX) = {}^tX - X = g(X)$ だから、$\mathrm{Im}\, f \subset \mathrm{Alt}_2(\mathbb{R})$
- $X \in \mathrm{Alt}_2(\mathbb{R})$ に対し、$f(X/2) = (X - {}^tX)/2 = X$ となるので、$\mathrm{Alt}_2(\mathbb{R}) \subset \mathrm{Im}\, f$

だから、$\mathrm{Im}\, f = \mathrm{Alt}_2(\mathbb{R})$ である。

これより、残りの問題は $\mathrm{Sym}_2(\mathbb{R})$, $\mathrm{Alt}_2(\mathbb{R})$ の基底を求めることに帰着された。一般に行列

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

が対称行列であることは $b = c$ と、交代行列であることは「$b = -c$ かつ $a = d = 0$」とそれぞれ同値である。よって全ての対称行列は

$$\begin{pmatrix} a & b \\ b & d \end{pmatrix} = a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

の形に表せ、全ての交代行列は

$$\begin{pmatrix} 0 & -c \\ c & 0 \end{pmatrix} = c\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

の形に表せる。これらの表式は一意的なので、基底が求まったことになる。 ■

■対称行列と交代行列による分解　今の問 4, 5 に出てきた $f, g\colon \mathrm{Mat}_2(\mathbb{R}) \to \mathrm{Mat}_2(\mathbb{R})$ は、全ての n 次正方行列 X に対し $f(X) + g(X) = 2X$ を満たしています。この両辺を 2 で割れば

$$X = \frac{1}{2}f(X) + \frac{1}{2}g(X) = \frac{X + {}^tX}{2} + \frac{X - {}^tX}{2}$$

という式が得られます。注目すべきは、右辺で第 1 項が $\mathrm{Im}\, f = \mathrm{Sym}_2(\mathbb{R})$ に、第 2 項が $\mathrm{Im}\, g = \mathrm{Alt}_2(\mathbb{R})$ に入っていることです。つまり全ての行列は、対称行列と交代行列の和に書けます。これより $\mathrm{Mat}_2(\mathbb{R}) = \mathrm{Sym}_2(\mathbb{R}) + \mathrm{Alt}_2(\mathbb{R})$ だと分かります。

さらに $X \in \mathrm{Sym}_2(\mathbb{R}) \cap \mathrm{Alt}_2(\mathbb{R})$ だとしたら、${}^tX = X$ かつ ${}^tX = -X$ が成り立つので、$X = O$ です。よって $\mathrm{Sym}_2(\mathbb{R}) \cap \mathrm{Alt}_2(\mathbb{R}) = \{O\}$ なので、$\mathrm{Mat}_2(\mathbb{R})$ は $\mathrm{Sym}_2(\mathbb{R})$ と $\mathrm{Alt}_2(\mathbb{R})$ の直和で $\mathrm{Mat}_2(\mathbb{R}) = \mathrm{Sym}_2(\mathbb{R}) \oplus \mathrm{Alt}_2(\mathbb{R})$ が成り立っています。つまり**全ての行列は、対称行列と交代行列の和にただ一通りに書ける**というわけです。この分解は知っておいて損はありませんので、心の片隅に留めておいてください。

5　行列のなす Lie 環

今回の問 7 には「交換子」と呼ばれるものが登場します。問 7 では行列の交換子だけを計算しますが、実は行列に限らず、交換子は世の中の色々な場面で登場することが知られています。その性質を、少し突っ込んで調べましょう。

*10 $\mathrm{Sym}_n(\mathbb{R})$ は n 次対称行列、$\mathrm{Alt}_n(\mathbb{R})$ は n 次交代行列のなす線型空間です。

なお、以下の話では具体例としてちょくちょく量子力学が顔を出します。別に「量子力学を知らなければ線型代数が理解できない」というわけではないですが、これを機に量子力学と線型代数を関連付けて勉強するのも良いでしょう。たとえば、駒場キャンパスにいらっしゃる総合文化研究科の清水明先生が書かれた『新版 量子論の基礎』(サイエンス社) は手頃で良い本だと思います。このプリントに出てくる量子力学の話は、ほとんどこの本に書いてあります。

5.1 交換子

V を線型空間とし、$X, Y \in \mathrm{End}(V)$ を V 上の線型変換とします。このとき $[X, Y] := XY - YX$ を X と Y の**交換子**といいます。

最も典型的な例は、行列の交換子です。たとえば S1 タームの終わりの方に、一度演習問題で

$$H := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad F := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

の交換子を調べる問題が登場しました[*11]。また X や Y は行列で書かれている必要もありません。たとえば実数値の無限回微分可能な函数全体の集合 $V = C^\infty(\mathbb{R})$ の上で、X は微分 $\frac{d}{dx}$、Y は x をかける線型写像だとします。このとき $f \in C^\infty(\mathbb{R})$ に対し

$$[X, Y]f(x) = XYf(x) - YXf(x) = \frac{d}{dx}\bigl(xf(x)\bigr) - x\Bigl(\frac{d}{dx}f(x)\Bigr) = f(x) + xf'(x) - xf'(x) = f(x)$$

が成り立ちます。

■交換子の性質　交換子は次の性質を満たします。

- 双線型性: $[\alpha X_1 + \beta X_2, Y] = \alpha[X_1, Y] + \beta[X_2, Y]$, $[X, \alpha Y_1 + \beta Y_2] = \alpha[X, Y_1] + \beta[X, Y_2]$
- 交代性: $[Y, X] = -[X, Y]$
- Jacobi 恒等式: $\bigl[X, [Y, Z]\bigr] + \bigl[Y, [Z, X]\bigr] + \bigl[Z, [X, Y]\bigr] = 0$

これは定義式 $[X, Y] = XY - YX$ を逐一代入して計算するだけで証明ができます。試してみてください。

5.2 Lie 環

さて、いまは行列などの掛け算を使って交換子 $[,]$ を定義しました。ですが色々計算をしてみると、時と場合によっては**交換子の性質だけを使って議論を進められたりします**。そこで一般に線型空間 $\mathfrak{g}$[*12]の上で Jacobi 恒等式を満たす交代双線型写像 $[,]\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ が定義されているとき、$\mathfrak{g}$ を **Lie 環**といいます。n 次正方行列の集合 $\mathrm{Mat}_n(\mathbb{R})$ は Lie 環の一例です。

■なぜ Lie 環を考えるのか？　僕たちはなぜ、わざわざ Lie 環などという対象を定義し、その性質を調べるのでしょうか？

答え方には色々ありますが、一つの理由としては **Lie 環に現れる関係式が、色々なところで登場する**という事実が挙げられます。たとえば複素数を成分とする 2 次正方行列[*13]

$$\sigma_x := \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

[*11] 45 ページにあるので、思い出してみてください。

[*12] '$\mathfrak{g}$' は、アルファベット 'g' のフラクトゥール体です。Lie 環の研究の創始者である数学者 Sophus Lie が Lie 環を表すのにフラクトゥール体を使ったことが、今では数学者の間で標準的な習わしになっています。

[*13] 授業中に複素数成分の行列を表だって扱ったことはないですが、計算の仕方は実数成分のときと全く同じです。

は $[\sigma_x, \sigma_y] = i\sigma_z, [\sigma_y, \sigma_z] = i\sigma_x, [\sigma_z, \sigma_x] = i\sigma_y$ という式を満たします。一方で 3 変数の函数に対する微分作用素 L_x, L_y, L_z を

$$L_x := y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}, L_y := z\frac{\partial}{\partial x} - x\frac{\partial}{\partial x}, L_z := x\frac{\partial}{\partial y} - y\frac{\partial}{\partial z}$$

で定める[*14]と、これらもまた $[L_x, L_y] = L_z, [L_y, L_z] = L_z, [L_z, L_x] = L_y$ という関係式を満たします。これは $\sigma_x, \sigma_y, \sigma_z$ の交換関係と大体同じです[*15]。

また量子力学の勉強をしていると、必ず $[\hat{x}, \hat{p}] = 1$ という式が登場します[*16]。たとえば交換子の例として微分作用素 $X = \frac{d}{dx}$ と x 倍作用素 Y を挙げましたが、これらは $[X, Y] = 1$ を満たしていました。この X と Y は、正準交換関係を満たす演算子の **Schrödinger 表現**といいます。他にも正準交換関係を満たす演算子は色々ありますが、とにかく $[\hat{x}, \hat{p}] = 1$ を満たすことが大事なのです。

このように「同じ交換関係式が、時と場所を変えて色々な状況で出現する」という実態を知ると「一々個別に調べるのは面倒だから、交換関係だけに注目して、その性質を調べ上げてしまえ」という発想に至るのは自然なことでしょう。こうして「特定の交換関係だけ」を抽出したものが Lie 環で、その交換関係を満たす行列や演算子のことを「Lie 環の表現」といいます。

■Lie 環の由緒　さらに言うと、Lie 環の役割を一段と良く理解するには「なぜ Lie 環の交換関係が色々なところに登場するか」を知らないといけません。そのキーワードが「Lie 群」と「対称性」です。

たとえば水素原子を考えると、原子核は陽子 1 個だけからなります。陽子の作る電場は、原点を通るあらゆる軸に対する回転対称性を持ちます。回転対称性以外にも対称性には色々な種類がありますが、数学では対称性を表すのに「群」という言葉を使います。Lie 群はその中でも特別な群であり、Lie 群の対称性を「無限小」のレベルで見ようとすると Lie 環が登場します。いよいよ Lie 環が重要そうな気がしてきますね[*17]。

ですが Lie 群と Lie 環の対応を述べるには、行列の指数函数[*18]の性質を詳しく調べなければいけません。事実としては、全ての行列 $X, Y \in \mathrm{Mat}_n(\mathbb{R})$ に対し

$$e^{tX}e^{tY}e^{-tX}e^{-tY} = \exp\Big(t^2[X, Y] + (t^3 \text{ より高次の項})\Big)$$

が成り立つことが、Lie 環の交換子 $[,]$ が重要な理由です。でもこの書き方では「何で $e^{tX}e^{tY}e^{-tX}e^{-tY}$ という式が出てくるのか」を説明できていません。これを理解するには少し「群論」というものを勉強する必要があります。行列式の話をする時に少し群論が出てくると思いますので、Lie 群と Lie 環の対応を述べるのはその機会に譲りたいと思います。

5.3　Lie 環 $\mathfrak{sl}_n$

ここまで出てきた具体的な Lie 環は $\mathrm{Mat}_n(\mathbb{R})$ しかありませんが、これ以外にも $\mathfrak{sl}_n$ という重要な Lie 環があるので、それを紹介します。中でも $\mathfrak{sl}_2$ は、Lie 環の理論を展開する上でも、あるいは量子力学等への応用を考えるにしても、非常に重要な役割を果たします。

■定義　一般に $\mathrm{tr}\colon \mathrm{Mat}_n(\mathbb{R}) \to \mathbb{R}$ の核を $\mathfrak{sl}_n(\mathbb{R})$ と書きます[*19]。任意の $X, Y \in \mathrm{Mat}_n(\mathbb{R})$ に対して $\mathrm{tr}\, XY = \mathrm{tr}\, YX$ なので、特に $X, Y \in \mathfrak{sl}_n(\mathbb{R})$ に対しても $[X, Y] \in \mathfrak{sl}_n(\mathbb{R})$ が成り立ちます。よって $\mathfrak{sl}_n(\mathbb{R})$ は $[,]$ で閉じており、Lie 環

[*14] つまり「3 変数函数 $f(x, y, z)$ に対し $L_x f := yf_z - zf_y$ などと定める」という意味です。

[*15] $\sigma'_x := -i\sigma_x$, $\sigma'_y := -i\sigma_y$, $\sigma'_z := -i\sigma_z$ とおくと、本当に L_x, L_y, L_z の交換関係と同じ式を満たすようになります。

[*16] この式を**正準交換関係**といいます。量子力学の授業だと普通 $[\hat{x}, \hat{p}] = i\hbar$ という形で出てきますが、右辺が 1 でも $i\hbar$ でも Lie 環を考えるにあたっては大差ないので、数学では $[x, p] = 1$ としてしまうことが多いです。

[*17] なお Lie 環の対称性は、いつも Lie 群に付随して現れるわけではありません。Lie 環の対称性が単独で現れる状況もあります。

[*18] 指数函数 $e^x = \exp x$ を Taylor 展開した式には、正方行列を代入することができます。その式で行列の指数函数を定義します。

[*19] この $\mathfrak{sl}$ も、"sl" のフラクトゥール体です。

になります。問 7 の議論を使えば、$\dim \mathfrak{sl}_n(\mathbb{R}) = n-1$ と分かります。

■$\mathfrak{sl}_2(\mathbb{C})$ と Pauli 行列　ここで数の範囲を広げて、複素係数の行列で $\mathfrak{sl}_2(\mathbb{C})$ を考えます。既に E, H, F という基底が $\mathfrak{sl}_2(\mathbb{C})$ に取れることを知っていますが、複素係数の場合は特に

$$\sigma_x := \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y := \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z := \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

という別の基底が取れます。これらの行列を **Pauli 行列**といいます。

Pauli 行列の特徴的な点は、Hermite 行列であること[*20]です。Hermite 行列はいつも「固有値」と呼ばれるものが実数になります。そして量子力学では

- 物理系の状態は、複素線型空間のベクトルで表される
- 物理量は Hermite 行列で表される
- ある状態における物理量の観測値は、物理量を表す Hermite 行列の固有値のどれかから確率的に選択される

という事情があります。物理量の観測値が実数になることを、Hermite 性が保証してくれるのです。なので $\mathfrak{sl}_2(\mathbb{C})$ の基底としては (E, H, F) の方が便利でも、物理量を考えるときは $(\sigma_x, \sigma_y, \sigma_z)$ が好まれたりします。たとえば電子の持つスピンの観測値は、そのまま $\sigma_x, \sigma_y, \sigma_z$ の固有値になることが知られています。

5.4 随伴表現

最後に、問 7 に出てきた線型写像 h がどうして大事なのか、軽く紹介します。

$\mathfrak{g}$ を Lie 環とします[*21]。交換子 $[,]$ は双線型なので、片側については線型になっています。そこで $X \in \mathfrak{g}$ に対して $\operatorname{ad} X(Y) := [X, Y]$ と定義すれば、$\operatorname{ad} X \colon \mathfrak{g} \to \mathfrak{g}$ という線型写像、つまり $\operatorname{ad} X \in \operatorname{End}(\mathfrak{g})$ が得られます。これには**随伴表現**という名前がついています。

交換子が Jacobi 恒等式を満たすことから、$\operatorname{ad} X$ について次が成り立ちます。

- $\operatorname{ad} X\big([Y, Z]\big) = [\operatorname{ad} X(Y), Z] + [X, \operatorname{ad} Y(Z)]$
- $\operatorname{ad}[X, Y](Z) = [\operatorname{ad} X, \operatorname{ad} Y](Z)$

そして $\operatorname{ad} X \in \operatorname{End}(\mathfrak{g})$ の $X \in \mathfrak{g}$ を動かすことで、$\operatorname{ad} \colon \mathfrak{g} \to \operatorname{End}(\mathfrak{g})$ という写像が得られます。$\mathfrak{g}$ は線型空間なので、$\operatorname{End}(\mathfrak{g})$ もまた Lie 環です。そして今の 2 つ目の式 $\operatorname{ad}[X, Y] = [\operatorname{ad} X, \operatorname{ad} Y]$ は、写像 ad が Lie 環の構造と整合的であることを表しています。

$$\begin{tikzcd}
\mathfrak{g} \times \mathfrak{g} \arrow[r, "\operatorname{ad} \times \operatorname{ad}"] \arrow[d, "{[,]}"'] & \operatorname{End}(\mathfrak{g}) \times \operatorname{End}(\mathfrak{g}) \arrow[d, "{[,]}"] \\
\mathfrak{g} \arrow[r, "\operatorname{ad}"] & \operatorname{End}(\mathfrak{g})
\end{tikzcd}$$

ここで線型写像が「線型空間の構造と整合的な写像」だったことを思い出せば、Lie 環の間の写像を考えるときも、整合的な写像を考えるのは自然だと思えるはずです。特に ad は、どんな Lie 環 $\mathfrak{g}$ にももれなくついてくるという点で、とりわけ重要なものです。

Lie 環論と呼ばれる分野では、Lie 環の構造に整合的な線型写像をうまく使うことで、Lie 環の構造を調べます。その中でも $\operatorname{ad} - \alpha \operatorname{id}$ $(\alpha \in \mathbb{R})$ の形をした写像の核は ad の固有空間と呼ばれ、極めて重要な役割を果たします。線型代数を使いこなすことの良い具体例だったので、今回ちょっと詳しく取り上げてみました。

[*20] 複素数を成分とする行列 X が $\overline{{}^t X} = X$ を満たすとき、X を **Hermite 行列**といいます。

[*21] 慣れなければ $\mathfrak{g} = \operatorname{Mat}_n(\mathbb{R})$ と思っていいです。

6　双対空間

一般に線型空間 V があると、もれなく「V の双対空間」と呼ばれる線型空間 V^* が定義できます。V^* のことを知っていると線型代数の理解が一段と深まりますし、また相対論を勉強する時に出てくる「共変/反変テンソル」というものを理解するのにも役立ちます。今回のプリントの最初に説明した「線型写像の線型空間」の例だから、慣れないと多少こんがらがるかもしれません。でも双対空間を考えてみることは、抽象論の良い訓練になるでしょう。

6.1　双対空間の定義

線型空間 V, W があったとき、V から W への線型写像全体の集合 $\mathrm{Hom}(V,W)$ も再び線型空間になるのでした。ここで特に $W=\mathbb{R}$ のとき、$V^* := \mathrm{Hom}(V,\mathbb{R})$ と書き、V の**双対 (そうつい) 空間**といいます。

■$\mathbb{R}^n$ の双対空間　一番簡単な $\mathbb{R}^2$ の場合に、双対空間 $(\mathbb{R}^2)^*$ は何かを考えてみましょう。今の場合、$\mathbb{R}^2$ から $\mathbb{R}$ への線型写像は $(1,2)$ 型行列、つまり行ベクトルと同じです。よって $\mathrm{Hom}(\mathbb{R}^2,\mathbb{R}) = \mathrm{Mat}_{1,2}(\mathbb{R})$ です。$\mathbb{R}^n$ の場合でも全く話は同じで、列ベクトル全体のなす数ベクトル空間 $\mathbb{R}^n$ の双対空間 $(\mathbb{R}^n)^*$ は、同じ長さの行ベクトル全体がなす空間です。

6.2　双対基底

線型空間 V に基底 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ が与えられると、それに応じて自然に V^* の基底 $(\boldsymbol{f}_1^\vee, \boldsymbol{f}_2^\vee, \ldots, \boldsymbol{f}_n^\vee)$ が定義されます。この作り方を今から説明します。

まず V を定義域とする線型写像は、基底の行先を決めればただ一通りに決まります。そこで $\boldsymbol{f}_i^\vee : V \to \mathbb{R}$ を

$$\boldsymbol{f}_i^\vee(\boldsymbol{f}_j) := \delta_{ij} = \begin{cases} 1 & (i=j) \\ 0 & (i \neq j) \end{cases}$$

と定めます。これから示すように、$(\boldsymbol{f}_1^\vee, \ldots, \boldsymbol{f}_n^\vee)$ は V^* の基底になります。これを $(\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$ の**双対基底**といいます。まず具体例を見てみましょう。

■$\mathbb{R}^n$ の場合　$\mathbb{R}^n$ の標準基底が相手の場合、双対基底も非常に分かり易いものになります。$\boldsymbol{e}_i$ を転置してできる行ベクトルを $\boldsymbol{e}_i^\vee$ と書くと

$$\boldsymbol{e}_i^\vee(\boldsymbol{e}_j) = \begin{array}{c} \begin{array}{ccccccc} & & & i^{\mathrm{th}} & & & \end{array} \\ \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix} \end{array} \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \delta_{ij}$$

が成り立つので、$(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n)$ の双対基底 $(\boldsymbol{e}_1^\vee, \ldots, \boldsymbol{e}_n^\vee)$ は、各 $\boldsymbol{e}_i$ を行列だと思って転置することで得られます。この例では、双対基底がちゃんと双対空間の基底をなしていることがほとんど明らかです 。

■双対基底が基底になること　例を確認したところで、一般の場合に話を戻しましょう。V を線型空間とし、その基底 $(\boldsymbol{f}_1, \boldsymbol{f}_2, \ldots, \boldsymbol{f}_n)$ を取ります。

さっき定義した $\boldsymbol{f}_1^\vee, \boldsymbol{f}_2^\vee, \ldots, \boldsymbol{f}_n^\vee$ は 1 次独立です。実際 $\alpha_1 \boldsymbol{f}_1^\vee + \alpha_2 \boldsymbol{f}_2^\vee + \cdots + \alpha_n \boldsymbol{f}_n^\vee = 0 \in V^*$ とおく[*22]と、この写

[*22] 全てのベクトルに 0 を対応させる写像は、線型写像です。これを**零写像**といい、0 で表します。

像に何を代入しても値は 0 です。そこで $\boldsymbol{f}_j$ $(j=1,2,\ldots,n)$ を代入すると

$$0=\sum_{i=1}^{n}\alpha_i\boldsymbol{f}_i^{\vee}(\boldsymbol{f}_j)=\sum_{i=1}^{n}\alpha_i\delta_{ij}=\alpha_j$$

となるので、$\alpha_1=\alpha_2=\cdots=\alpha_n=0$ となります。これで 1 次独立性がいえました。

また、全ての $f\in V^*$ は $\boldsymbol{f}_1^{\vee},\boldsymbol{f}_2^{\vee},\ldots,\boldsymbol{f}_n^{\vee}$ の 1 次結合で書けます。実際 $f\in V^*$ に対し、$f':=f(\boldsymbol{f}_1)\boldsymbol{f}_1^{\vee}+f(\boldsymbol{f}_2)\boldsymbol{f}_2^{\vee}+\cdots+f(\boldsymbol{f}_n)\boldsymbol{f}_n^{\vee}$ とおくと、各 $\boldsymbol{f}_j$ $(j=1,2,\ldots,n)$ を f' に代入して

$$f'(\boldsymbol{f}_j)=\sum_{i=1}^{n}f(\boldsymbol{f}_i)\boldsymbol{f}_i^{\vee}(\boldsymbol{f}_j)=\sum_{i=1}^{n}f(\boldsymbol{f}_i)\delta_{ij}=f(\boldsymbol{f}_j)$$

が得られます。すなわち基底 $(\boldsymbol{f}_1,\boldsymbol{f}_2,\ldots,\boldsymbol{f}_n)$ の全ての行先が f と f' とで同じなので、写像として $f=f'$ が従います。これで V^* が $f\in V^*$ は $\boldsymbol{f}_1^{\vee},\boldsymbol{f}_2^{\vee},\ldots,\boldsymbol{f}_n^{\vee}$ で生成されることが示せました。以上より $(\boldsymbol{f}_1^{\vee},\boldsymbol{f}_2^{\vee},\ldots,\boldsymbol{f}_n^{\vee})$ は V の基底です。

このことより、特に $\dim V^*=\dim V$ が従います。また一般の線型空間 V は「標準的」と呼ぶべき基底を持つとは限りませんが、V の基底を決めたなら、それに応じて V^* に双対基底という自然な基底が定まります。言い方を変えれば「V に座標を入れたら、それに応じて V^* に自然な座標の入れ方が決まる」と言えます。

■問 10 の解答　双対基底の本数を数えれば $\dim V^*=\dim V$ が従う。 ■

6.3　2 回双対

線型空間 V から双対空間 V^* を作ったのと同じようにして、今度は V^* の双対空間 $V^{**}:=(V^*)^*$ を考えることができます。$\mathbb{R}^n$ の場合を思い出すと、双対 $*$ を取ることは行列としての転置と同じでした。ですから $*$ を 2 回取ったら、元に戻りそうな気がします。そこで一般の場合に、この V^{**} の構造を調べましょう。

■評価写像　V^{**} の元は「V^* の元に対して、実数 $\mathbb{R}$ を対応させる」という写像です。そして V^* は $V\to\mathbb{R}$ という写像です。だから $\boldsymbol{v}\in V$ を 1 つ固定すると、$\varphi\mapsto\varphi(\boldsymbol{v})$ という方法で $\mathrm{ev}_{\boldsymbol{v}}:V^*\to\mathbb{R}$ という写像が作れます。これを**評価写像**というのでした。63 ページでは多項式に対する評価写像を考えましたが、V^* の元に対する評価写像についても全く同様に、$\mathrm{ev}_{\boldsymbol{v}}$ の線型性が示せます。よって $\mathrm{ev}_{\boldsymbol{v}}\in V^{**}$ です。

そして $\mathrm{ev}_{\boldsymbol{v}}$ の $\boldsymbol{v}$ を動かすことで、$\mathrm{ev}\colon V\to V^{**};\boldsymbol{v}\mapsto\mathrm{ev}_{\boldsymbol{v}}$ という写像ができます。これも線型写像になっています。実際、任意の $\varphi\in V^*$ に対し

- $\mathrm{ev}_{\boldsymbol{v}+\boldsymbol{w}}(\varphi)=\varphi(\boldsymbol{v}+\boldsymbol{w})=\varphi(\boldsymbol{v})+\varphi(\boldsymbol{w})=(\mathrm{ev}_{\boldsymbol{v}}+\mathrm{ev}_{\boldsymbol{w}})(\varphi)$
- $\mathrm{ev}_{\alpha\boldsymbol{v}}(\varphi)=\varphi(\alpha\boldsymbol{v})=\alpha\varphi(\boldsymbol{v})=\alpha\,\mathrm{ev}_{\boldsymbol{v}}(\varphi)$

なので、$\mathrm{ev}_{\boldsymbol{v}+\boldsymbol{w}}=\mathrm{ev}_{\boldsymbol{v}}+\mathrm{ev}_{\boldsymbol{w}}$ かつ $\mathrm{ev}_{\alpha\boldsymbol{v}}=\alpha\,\mathrm{ev}_{\boldsymbol{v}}$ です。

■V と V^{**} の同型　この $\mathrm{ev}\colon V\to V^{**}$ が同型になることを示しましょう。$\dim V^{**}=\dim V^*=\dim V$ なので、単射性だけ示せば自動的に全射性が従います。

いま $\boldsymbol{v}\neq\boldsymbol{0}\in V$ とすると、$\boldsymbol{v}$ を延長することで V の基底 $(\boldsymbol{f}_1:=\boldsymbol{v},\boldsymbol{f}_2,\ldots,\boldsymbol{f}_n)$ が作れます。この基底の双対基底 $(\boldsymbol{f}_1^{\vee},\boldsymbol{f}_2^{\vee},\ldots,\boldsymbol{f}_n^{\vee})$ を取ると、$\boldsymbol{f}_1^{\vee}(\boldsymbol{v})=1\neq0$ となります。よって $\varphi\in V^*$ で $\varphi(\boldsymbol{v})\neq0$ となるものが存在します。

この対偶を取ると「任意の $\varphi\in V^*$ に対して $\varphi(\boldsymbol{v})=0$ なら、$\boldsymbol{v}=\boldsymbol{0}$」となります。よって $\boldsymbol{v}\in V$ が $\mathrm{ev}_{\boldsymbol{v}}=0\in V^{**}$ を満たせば、任意の $\varphi\in V^*$ に対し $\varphi(\boldsymbol{v})=\mathrm{ev}_{\boldsymbol{v}}(\varphi)=0$ となるから、$\boldsymbol{v}=\boldsymbol{0}$ です。よって ev は単射です[*23]。

[*23] ここの証明は $\dim V=\infty$ でも大体有効です。無限次元の場合に破綻するのは「単射ならば全射」の部分だけであって、$\mathrm{ev}\colon V\to V^{**}$ が単射なことはいつでも正しいです。

だから双対空間を考えると $V \to V^* \to V^{**} \to V^{***} \to \cdots$ というように、いくらでも線型空間を作ることはできるのですが、結局は 1 つおきに同型になってしまうのです。また今は V から V^* を作りましたが、V^* から V と同型な V^{**} を作れることを考えると「V と V^* の立場は対等なのではないか」という気がしてきます。そこで $\boldsymbol{v} \in V$ と $f \in V^*$ に対し、$f(\boldsymbol{v})$ のことを $\langle \boldsymbol{v}, f \rangle$ と内積のように書くことがあります。また「函数」という言葉を使うと対等性が崩れるので、$\langle \boldsymbol{v}, f \rangle$ のことを「標準的なペアリング」と言ったりもします。

6.4 Hom の分解

V, W が線型空間のとき、V^* の元と W の元を組み合わせることで V から W への線型写像を作ることができます。

いきなり抽象論をやる前に、数ベクトル空間の例を見てみましょう。既に見たとおり、$(\mathbb{R}^n)^*$ は n 次元の横ベクトルのなす線型空間でした。また $\mathbb{R}^m$ は m 次元の縦ベクトルの掛け算でした。すると $\mathbb{R}^m$ の縦ベクトルは $(m, 1)$ 型行列、$(\mathbb{R}^n)^*$ の横ベクトルは $(1, n)$ 型行列と同一視できるので、これらを掛け算[*24]して (m, n) 型行列が作れます:

$$\boldsymbol{e}_{n,i}\boldsymbol{e}_{m,j}^{\vee} = \begin{array}{c} \\ i^{\text{th}} \\ \\ \end{array}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{array}{c} \cdots \, j^{\text{th}} \\ \begin{pmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{pmatrix} \end{array} = \begin{array}{cc} & j^{\text{th}} \\ i^{\text{th}} & \begin{pmatrix} & & \\ & 1 & \\ & & \end{pmatrix} \end{array} = E_{ij}$$

そして E_{ij} に $\boldsymbol{e}_{n,k}$ を当てると $E_{ij}\boldsymbol{e}_{n,k} = \delta_{jk}\boldsymbol{e}_{m,i}$ となりますが、この式は $\boldsymbol{e}_{m,i}\boldsymbol{e}_{n,j}^{\vee}(\boldsymbol{e}_{n,k})$ と同じです。図式にすると

$$\begin{array}{ccccc} \mathbb{R}^m \times (\mathbb{R}^n)^* & \longrightarrow & \mathrm{Hom}(\mathbb{R}^n, \mathbb{R}^m) & = & \mathrm{Mat}_{m,n}(\mathbb{R}) \\ \cup\!\shortmid & & \cup\!\shortmid & & \cup\!\shortmid \\ (\boldsymbol{e}_{m,i}, \boldsymbol{e}_{n,j}^{\vee}) & \longmapsto & \boldsymbol{e}_{m,i}\boldsymbol{e}_{n,j}^{\vee} & = & E_{ij} \end{array}$$

という感じです。一般の場合にも同じことができます。V, W が線型空間で $\boldsymbol{w} \in W$, $\varphi \in V^*$ とします。このとき写像 $\boldsymbol{w} \otimes \varphi \colon V \to W$ を[*25]

$$(\boldsymbol{w} \otimes \varphi)(\boldsymbol{v}) := \varphi(\boldsymbol{v})\boldsymbol{w}$$

で定義できます。これも図式にすると

$$\begin{array}{ccc} W \times (V)^* & \longrightarrow & \mathrm{Hom}(V, W) \\ \cup\!\shortmid & & \cup\!\shortmid \\ (\boldsymbol{w}, \varphi) & \longmapsto & \boldsymbol{w} \otimes \varphi \end{array}$$

となります。

そして $\mathrm{Hom}(V, W)$ の元は、$\boldsymbol{w} \otimes \varphi$ の形の元の 1 次結合で表すことができます。たとえば $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ の場合、$\boldsymbol{e}_{m,i} \otimes \boldsymbol{e}_{n,j}^{\vee} = E_{ij}$ の 1 次結合によって、どんな行列も作れますね。これと同じことが抽象的な線型空間についても成り立つのです。

■トレースとの関係　最後に、$\mathrm{Hom}(V, W)$ の元が $\boldsymbol{w} \otimes \varphi$ の形の写像の和で書けることのご利益を一つ紹介します。それは**行列表示によらないトレースの定義ができる**という事実です。

$\boldsymbol{v} \in V$, $\varphi \in V^*$ があると、線型写像 $\boldsymbol{v} \otimes \varphi \colon V \to V$ が定義できるのでした。一方 $\varphi(\boldsymbol{v}) \in \mathbb{R}$ でもあります。実はこのとき、$\mathrm{tr}\, \boldsymbol{v} \otimes \varphi = \varphi(\boldsymbol{v})$ という式が成り立っています。

[*24] サイズの違うベクトルが出てくるので、以下では「$\mathbb{R}^n$ における i 番目の標準基底」を表すのに $\boldsymbol{e}_{n,i}$ と書くことにします。

[*25] 本当は、記号 $\otimes$ は「テンソル積」と呼ばれるものを表すのですが、今は $\boldsymbol{w} \otimes \varphi$ を「まとめて一つの記号」だと思って差支えありません。

$V=\mathbb{R}^n$ の場合に確かめてみましょう。$\boldsymbol{e}_i \otimes \boldsymbol{e}_j^\vee = E_{ij}$ だったので、$\operatorname{tr} \boldsymbol{e}_i \otimes \boldsymbol{e}_j^\vee = \delta_{ij}$ です。かたや双対基底の定義から $\boldsymbol{e}_j^\vee(\boldsymbol{e}_i) = \delta_{ij}$ なので、両者は確かに一致しています。そして tr は線型写像だったので、$\operatorname{tr} \boldsymbol{e}_i \otimes \boldsymbol{e}_j^\vee = \boldsymbol{e}_j^\vee(\boldsymbol{e}_i)$ から、全ての $\boldsymbol{v} \in \mathbb{R}^n$, $\varphi \in (\mathbb{R}^n)^*$ に対し $\operatorname{tr} \boldsymbol{v} \otimes \varphi = \varphi(\boldsymbol{v})$ が成り立つことが従います。

まだ「(正方行列とは限らない、一般の) 線型変換のトレース」は定義していませんでしたが、普通は「線型変換を行列表示したときのトレース」を考えます。実は線型変換をどのように行列表示しても、トレースが変わらないことが証明できるのです。ですが行列は、線型変換を基底で表示したときに見えてくる一つの姿でしかありません。その表示によらず、線型変換の言葉だけでトレースを捉えようとすると、今のような議論が必要になるのです。

7　A セメスターの展望

最後に、A セメスターに何をやるはずかについて、簡単に述べておきます。

■行列式　まず最初に扱うのは行列式です。これは n 次元空間における「立体の体積」にあたるもので、n 本のベクトルが一次独立性かどうかを判定することに使ったり、連立 1 次方程式を解いたりするのに非常に役立ちます。理論上も実務上も極めて大事です。ですから

- 行列式の理論的な意味を知ること
- 行列式を手際よく計算できるようになること

を心がけてください。

■計量線型空間　これまでの線型代数の授業では「線型性」について突っ込んだ議論を行ってきましたが、「内積」にあたるものは登場していません。たとえば我々の空間 $\mathbb{R}^3$ では、$\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^3$ のなす角を θ とすれば $\boldsymbol{u} \cdot \boldsymbol{v} = \|\boldsymbol{u}\|\|\boldsymbol{v}\| \cos\theta$ が成り立ちます。行列の積を定義するときに $\mathbb{R}^n$ の内積っぽい計算をしていても、「角度」や「長さ」といった幾何学的な話は、線型空間の中で出てきませんでした。こういう「角度」や「長さ」にあたるものを**計量**といいます。この計量の扱いを、勉強することになります。

その際僕たちは、「座標」と「計量」を切り離して考えます。これはたとえば「$\mathbb{R}^3$ のベクトル $\boldsymbol{e}_i$ の長さを 1 ではなくす」ということに相当します。なぜこんなことをするのか分からないかもしれませんが、「長さの目盛りを取り換える」と思えば、そんな不自然ではないはずです。たとえば僕たちは長さを測るのに僕たちは 1cm とか 1mm とか、いくつか単位を使い分けます。そうすると $\boldsymbol{e}_1$ の長さが 1cm だとして、mm 単位で測れば $\|\boldsymbol{e}_1\| = 10\mathrm{mm}$ になったりするわけです。ですから、座標と長さは必ずしも連動する必要はありません。僕たちは後で、単なる拡大 / 縮小にとどまらず、もっと広い範疇で「長さ」に相当する概念を定式化します。

■行列の固有値と対角化　f を線型空間 V 上の線型変換としましょう。このとき V に基底を取れば f は行列によって表示されました。基底を色々取り換えると f の行列表示も色々変わるわけですが、運が良いと f が対角行列で表示されることがあります。このとき、f の行列表示の対角成分に並ぶ数を f の**固有値**といいます。

後々証明しますが、f の対角化のやり方は (存在すれば) 本質的に 1 通りしかありません。違う方法で対角化できたとしても、対角成分に数が並ぶ順番が変わるだけなのです。ここで、抽象的に線型写像を扱う上で大事なのは「基底による見かけの表示の向こう側にある、本質的なもの」を取り出すことでした。f の固有値も行列表示に依存しないので、f にとって本質的な量だと言えます。実際、微分方程式など色々な場面に線型代数を応用する時に、固有値の計算がカギになる場面は山ほどあります。そうした固有値の計算方法や利用法を、1 年生の最後で学ぶことになるでしょう。

A セメスター

線形代数学②

線形代数学② 第 1 回　　行列式

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 9 月 16 日 1 限

1　行列式とは

今回は待ちに待った (?) 行列式を扱います。行列式とは正方行列に対して定義される数で

- 行列が正則かどうかの判定
- 行列をなす行ベクトル / 列ベクトルたちが 1 次独立かどうかの判定

に使えるという、大変ありがたい性質を持っています。さらに正方行列 A を用いて記述される連立一次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解も、Cramer の公式と呼ばれる公式を使うと、行列式で書き下すことができます。こんな感じで、行列式は非常に便利なのです。

今回の演習では、行列式に関する計算問題と、行列式の定義をするのに使う置換に関する問題が出ました。まだ 4 次以上の行列式は授業でやっていなかったようなので、そこができなかった人は復習してください。また 3 次までの行列式は、何が何でもできるようになっておいてください。

2　2 次と 3 次の行列式

さて、n 次正方行列 $A = (a_{ij})_{1\leq i,j\leq n} \in \mathrm{Mat}_n(\mathbb{R})$ の行列式の定義は、一応

$$\det A = \sum_{\sigma\in\mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

という式で与えられます[*1]。$\det A$ のことを $|A|$ と書くこともあります。最終的にはこの式を理解して、n 次の行列式の計算ができるようにならないといけません。ですがこの式には初見の記号が出てきており、いきなり理解するのは大変です。なのでこの式は一旦脇に置いといて、まず $n = 2, 3$ の簡単な場合だけ調べてみましょう。

なお $n = 2, 3$ の場合だけを先に扱うのは、決して単なる「逃げ」ではありません。実務上、2 次や 3 次の行列式の計算は非常にたくさん出てきます。というのも行列式には

- n 次の行列式の計算は、$(n-1)$ 次の行列式の計算に帰着できる
- 行列式の計算の手間は大体 $n!$ に比例してややこしくなるので、高次の行列式をそのまま定義通りに計算することはまずない[*2]

という性質があるからです。高次の行列式を計算する場合であっても、実際には 2 次や 3 次の行列式をたくさん計算することになるのです。ですから上に書いた行列式の定義式を理解するのとは別個に、2 次と 3 次の行列式は暗記して、すらすら使えるようになっておく必要があります。

[*1] 行列式は英語で<u>det</u>erminant なので、その頭 3 文字を取っています。

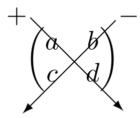
[*2] これは手計算に限った話ではありません。コンピュータであっても、定義通りに計算できるのはせいぜい 10 次くらいまでだと思います。また「同価格帯のコンピュータは 1 年半で性能が大体 2 倍になる」という Moore の法則によれば、コンピュータは時間を追うごとに指数函数的に性能が上がります。かたや行列式はサイズが大きくなると階乗の勢いで計算の手間が増えるので、どんな明るい未来がやってきても、今の汎用コンピュータ (より正確には、Turing 機械と呼ばれるタイプに属する機械) があらゆる行列式を定義通りに計算できる見込みはありません。基本変形と組み合わせたり近似公式を使ったりして、なんとか計算しているのが実情です。

2.1　2 次の行列式

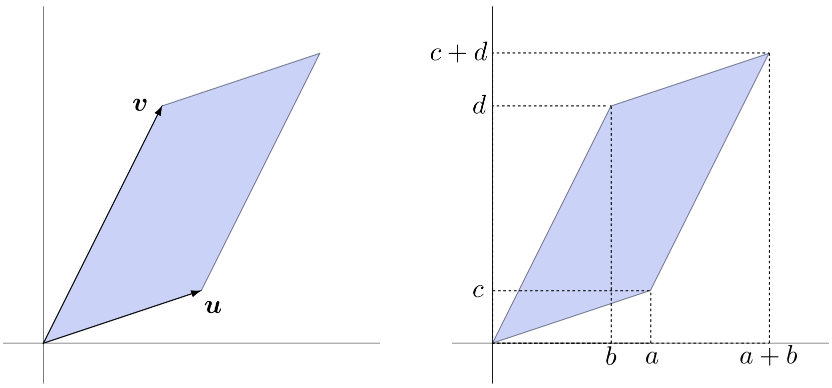
2 次の行列式は

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc$$

で定義されます。既に知っている人も、そこそこいるかと思います。知らなかった人は、この式を今すぐ暗記して、脊髄反射で書けるようにしておいてください。たとえば次の図のように「たすきがけ」っぽく覚える方法があります。同じ矢印の上にある数をかけ、$+$ の矢印の値は足し、$-$ の矢印の値は引くのです。



■面積との関係　2 次正方行列の行列式には、2 **本の列ベクトルが張る平行四辺形の面積**という重要な意味があります。それを確認しましょう。平面上に $\boldsymbol{u} = {}^t(a, c), \boldsymbol{v} = {}^t(b, d)$ という 2 本のベクトルを描いてみます。簡単のため、図のような状況で考えます。



右の図を見れば、算数を使うことで面積が

$$(a+b)(c+d) - \frac{ac}{2} - \frac{bd}{2} - \frac{\bigl(b + (a+b)\bigr)c}{2} - \frac{\bigl(c + (c+d)\bigr)b}{2} = ad - bc$$

と求まります。確かに行列式 $\det(\boldsymbol{u}\ \boldsymbol{v})$ になっていますね。

次に行列式の符号を考えてみます。図のように $a, b, c, d > 0$ の状況なら

$$ad - bc > 0 \Longleftrightarrow \frac{c}{a} < \frac{d}{b}$$

です。つまりベクトル ${}^t(a, c)$ の傾きが ${}^t(b, d)$ の傾きより小さいときに $ad - bc > 0$ というわけです。これは「第 1 象限内で ${}^t(a, c)$ を反時計回りに回すと ${}^t(b, d)$ の方向になる」ということと同じです。

これ以外の場合も地道にチェックしなければいけませんが、平面内の 2 本のベクトル $\boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^2$ を並べてできる行列の行列式 $\det(\boldsymbol{u}\ \boldsymbol{v})$ は、$\boldsymbol{u}, \boldsymbol{v}$ の位置関係がどんな場合であっても

- 絶対値が $\boldsymbol{u}$ と $\boldsymbol{v}$ の張る平行四辺形の面積
- 符号は、平行四辺形において $\boldsymbol{u}$ から $\boldsymbol{v}$ を見る方向が右回りなら $+$、左回りなら $-$

という性質を持っています。

このことを考えれば、$\boldsymbol{u}, \boldsymbol{v}$ が一次従属であることと $\det(\boldsymbol{u}\ \boldsymbol{v}) = 0$ が同値なのは一目瞭然でしょう。$\det(\boldsymbol{u}\ \boldsymbol{v}) = 0$ は「$\boldsymbol{u}$ と $\boldsymbol{v}$ の張る平行四辺形の面積が 0 になること」と同値です。この条件は「$\boldsymbol{u}$ と $\boldsymbol{v}$ が同一直線上に乗ること」つまり $\boldsymbol{u}$ と $\boldsymbol{v}$ が 1 次従属なことに他なりません。僕たちは後に一般の次数で、$\det A \neq 0$ と A の正則性が同値なことを証明します。その際に大事なのは、やはり「符号つき体積」というイメージです。また多重積分で置換積分をする際に「Jacobi 行列」というものの行列式が登場します。この証明をするときにも、det が体積を表すという事実は本質的です。今の話をきちんと理解してください。

■問 5 の解答　連立 1 次方程式

$$\begin{cases} ax + by &= e \\ cx + dy &= f \end{cases}$$

に対し、$d \times$ (第 1 式) $- b \times$ (第 2 式), $a \times$ (第 2 式) $- c \times$ (第 1 式) を計算すると

$$(ad - bc)x = ed - cf, (ad - bc)y = af - eb$$

を得る。これを行列式で書きなおすと

$$x = \det\begin{pmatrix} e & b \\ f & d \end{pmatrix} / \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}, y = \det\begin{pmatrix} a & e \\ c & f \end{pmatrix} / \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

となる。

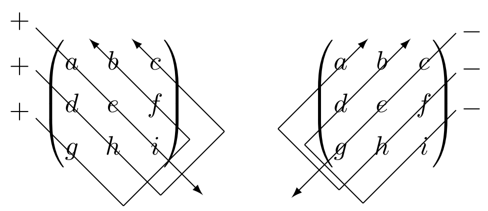
■Cramer の公式　こんな感じで、2 本の方程式からなる 2 変数の 1 次方程式の解が行列式で表せました。より一般に A が n 次正方行列で、連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解が唯一存在するとき[*3]、それを行列式で書き下す公式が存在します。それが **Cramer の公式**と呼ばれるものです。まだ n 次の行列式を定義していないので、それが終わった後、来週一般の場合を紹介します。

2.2　3 次の行列式

3 次の行列式も、まず定義式を書いてみます。

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} := aei + bfg + chd - ceg - bdi - ahf$$

この式には 2 つの覚え方があります。1 つ目は **Sarrus の方法**と呼ばれるものです。次の図のように、行列式を取る行列と曲がった矢印を重ねます。そして同じ矢印に乗った 3 つの数をかけて、最初右下に向かう矢印は $+$ の符号、左下に向かう矢印には $-$ の符号をつけて、足し合わせるのです。この方法で上の det の式が得られることを確認してください。



もう 1 つの覚え方は、**余因子展開**というものです。さっきの行列式をよく見ると

$$\det\begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + chd - ceg - bdi - ahf = a(ei - hf) - b(di - fg) + c(hd - eg)$$

[*3] つまり $\det A \neq 0$ のときです。

$$= a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \det \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \det \begin{pmatrix} d & e \\ g & h \end{pmatrix}$$

となっています。ルールが分かるような書き方をすると、こうなります:

- $\underline{1}$ 行目と $\underline{1}$ 列目を取り去ってできる小行列の行列式を、$\underline{(1,1)}$ 成分の a にかける
- $\underline{1}$ 行目と $\underline{2}$ 列目を取り去ってできる小行列の行列式を、$\underline{(1,2)}$ 成分の b にかける
- $\underline{1}$ 行目と $\underline{3}$ 列目を取り去ってできる小行列の行列式を、$\underline{(1,3)}$ 成分の c にかける

ということをやって、$\pm$ **の符号を交互にして**足しているのです[*4]。要となる成分が 1 行目を左から順に動くので、これを 1 **行目に関する余因子展開**といいます。

今と同じことを、列に関しても考えられます。行列式の定義式は

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = aei + bfg + chd - ceg - bdi - ahf = a(ei - hf) - d(bi - ch) + g(bf - ce)$$
$$= a \det \begin{pmatrix} e & f \\ h & i \end{pmatrix} - d \det \begin{pmatrix} b & c \\ h & i \end{pmatrix} + g \det \begin{pmatrix} b & c \\ e & f \end{pmatrix}$$

と書き直せます。この式は

- $\underline{1}$ 行目と $\underline{1}$ 列目を取り去ってできる小行列の行列式を、$\underline{(1,1)}$ 成分の a にかける
- $\underline{2}$ 行目と $\underline{1}$ 列目を取り去ってできる小行列の行列式を、$\underline{(2,1)}$ 成分の b にかける
- $\underline{3}$ 行目と $\underline{1}$ 列目を取り去ってできる小行列の行列式を、$\underline{(3,1)}$ 成分の c にかける

ということをやって、出てくる数に $\pm$ **を交互につけて**足しています。こっちは要となる成分が 1 列目を縦に動くので、1 **列目に関する余因子展開**といいます。同じようにして i 行目あるいは j 列目に関する余因子展開をすることもできますが、まずは 1 行目 / 1 列目での余因子展開を確実にできるようになりましょう。これさえできれば、行 / 列を動かしてもやることは同じです。

■問 4 (1) の解答　定義通りに計算すれば、次を得る。

$$D_2 = \det \begin{pmatrix} t & 1 \\ 1 & t \end{pmatrix} = t^2 - 1, \quad D_3 = \det \begin{pmatrix} t & 1 & 0 \\ 1 & t & 1 \\ 0 & 1 & t \end{pmatrix} = t^3 - 2t$$

■

■問 2 (1), (4) の解答　定義通りに計算すれば、次を得る。

$$\det \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix} = 0, \quad \begin{pmatrix} 1 & a & a^2 \\ 1 & b & b^2 \\ 1 & c & c^2 \end{pmatrix} = (b-a)(c-a)(c-b)$$

■

■Vandermonde 行列式　ここで出てきた (4) の行列式は (3 次の) **Vandermonde 行列式**と呼ばれます[*5]。また、Vandermonde 行列式の値 $(b-a)(c-a)(c-b)$ は a, b, c の**差積**と呼ばれます。a, b, c から 2 つを選ぶ全ての組合せについて差を計算し、その積を取っているから、差積という名前がついています。

この「Vandermonde 行列式 = 差積」という式は非常に重要です。というのも差積は

[*4] こういう「$\pm$ を交互にして足す」和のことを**交代和**と言ったりします。

[*5] 今はこの名前で定着していますが、実のところは Vandermonde 自身とそこまで深い関係にはないようです。この名前がついた詳しい歴史が `http://arxiv.org/abs/1204.4716` に載っています。

- 差積が 0 にならないことと a, b, c が相異なることが同値
- a, b, c の交代式は必ず差積で割り切れる

という性質を持っているからです。一方で Vandermonde 行列式の中身を係数行列に持つような連立一次方程式を解きたい場面が、この後たまに出てきます。そういう時に計算結果を知っていれば、方程式の解が存在するかどうか、直ちに判定することができます。

3　4 次以上の行列式

まだ授業ではやっていなかったようですが、問題の中に 4 次以上の行列式が出ているので、解き方を眺めてみましょう。今回解けなかった人も、後で復習してください。

3.1　高次の行列式の計算法

4 次より大きい行列式に対しても、やることは同じです。高次の行列式は、余因子展開で計算することができます。たとえば 4 次の行列式の場合

$$
\begin{aligned}
&\det\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} \\
&= a_{11}\det\begin{pmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{pmatrix} - a_{12}\det\begin{pmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{pmatrix} + a_{13}\det\begin{pmatrix} a_{21} & a_{22} & a_{24} \\ a_{31} & a_{32} & a_{34} \\ a_{41} & a_{42} & a_{44} \end{pmatrix} - a_{14}\det\begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}
\end{aligned}
$$

という式が成り立ちます。一般に n 次の行列式は余因子展開によって $n-1$ 次の行列式の計算に帰着できます。これを使えば、再帰的に行列式の計算を 2 次まで落としこめるというわけです[*6]。後で余因子展開の式が正しいことを証明しますが、証明は脇に置いといて、今回は使い方を見てみましょう。

■問 2 (2), (3), (5) の解答　(3)

$$
\begin{aligned}
\det\begin{pmatrix} 0 & 0 & 4 & 0 & 3 \\ 0 & 0 & 0 & 6 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 & 5 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix} &= 4\det\begin{pmatrix} 0 & 0 & 6 & 0 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 5 \\ 0 & 1 & 0 & 0 \end{pmatrix} + 3\det\begin{pmatrix} 0 & 0 & 0 & 6 \\ 2 & 0 & 0 & 0 \\ 0 & 0 & 7 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \\
&= 24\det\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 5 \\ 0 & 1 & 0 \end{pmatrix} - 18\det\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 7 \\ 0 & 1 & 0 \end{pmatrix} = 12
\end{aligned}
$$

(2)

$$
\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} = +0\det\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} - 1\det\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} + 1\det\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix} - 1\det\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = -3
$$

---

*6 行列式には

- このプリントに書いた、置換に関する $\sum$ で定義する式
- 「交代性」および「多重線型性」というものを用いる定義

という 2 種類の定義があり、どちらを採用しても同じ結果が得られることが知られています。どちらの定義も一長一短あるのですが、今回僕たちは前者の定義を採用して話を進めるので、余因子展開は最終的に「定義」ではなく「定理」として得られます。

(5)

$$
\det\begin{pmatrix} a & 0 & b & 0 \\ 0 & a & 0 & b \\ c & 0 & d & 0 \\ 0 & c & 0 & d \end{pmatrix} = a\det\begin{pmatrix} a & 0 & b \\ 0 & d & 0 \\ c & 0 & d \end{pmatrix} + b\det\begin{pmatrix} 0 & a & b \\ c & 0 & 0 \\ 0 & c & d \end{pmatrix}
$$
$$
= a(ad^2 - bdc) - b(c^2b - dca) = (ad - bc)^2
$$

■

■問 4 (2) **の解答**　D_n の det の中身を

$$
A_n := \begin{pmatrix} t & 1 & 0 & \cdots & 0 \\ 1 & t & 1 & \ddots & \vdots \\ 0 & 1 & t & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & 1 & t \end{pmatrix}
$$

とおく。この式を 1 行目について余因子展開すると

$$
D_n = \det A_n = t\det A_{n-1} + 1\det\begin{pmatrix} 1 & * \\ \mathbf{0} & A_{n-2} \end{pmatrix}
$$

となる[*7]。第 2 項は 1 列目に関して余因子展開すると $D_{n-2} = \det A_{n-2}$ となるから、結局 $D_n = tD_{n-1} + D_{n-2}$ という漸化式が得られる。これによって

$$
D_t = tD_4 - D_3 = t(tD_3 - D_2) - D_3 = (t^2 - 1) - tD_2
$$
$$
= (t^2 - 1)(t^3 - 2t) - t(t^2 - 1) = (t^2 - 1)(t^3 - 3t)
$$

を得る。 ■

3.2　特別な形の行列式

■問 3 (1) **の解答**　1 列目に関する余因子展開を繰り返すことで

$$
\det\begin{pmatrix} 1 & 5 & 8 & 10 \\ 0 & 2 & 6 & 9 \\ 0 & 0 & 3 & 7 \\ 0 & 0 & 0 & 4 \end{pmatrix} = 1\det\begin{pmatrix} 2 & 6 & 9 \\ 0 & 3 & 7 \\ 0 & 0 & 4 \end{pmatrix} = 2\det\begin{pmatrix} 3 & 7 \\ 0 & 4 \end{pmatrix} = 24
$$

を得る。 ■

■三角行列の行列式　今の式変形を見れば分かると思いますが、上三角行列の行列式は、対角成分を全て掛け算すれば得られます。理由は簡単で、

- 1 行目に関する余因子展開をすると、0 でない項が 1 個しか現れない
- 生き残る項に現れる行列式も、また上三角行列の行列式になっている

という事情があるからです。下三角行列の場合も、1 列目に関する余因子展開を繰り返せば同じ結果が得られます。

[*7] 行列式の計算に寄与しないどうでもいい部分を $*$ で表しました。

■問 3 (2) の解答　1 列目について余因子展開すると

$$\det\begin{pmatrix}1 & 5 & 8 & 10\\ a & 2 & 6 & 9\\ 0 & 0 & 3 & 7\\ 0 & 0 & 0 & 4\end{pmatrix} = 1\det\begin{pmatrix}2 & 6 & 9\\ 0 & 3 & 7\\ 0 & 0 & 4\end{pmatrix} - a\begin{pmatrix}5 & 8 & 10\\ 0 & 3 & 7\\ 0 & 0 & 4\end{pmatrix} = 24 - 60a$$

を得る。 ■

■多項式の行列式　今回の問 3 (2) や問 4 では、多項式を成分とする正方行列が登場します。このような場合、行列式の値も多項式になります。さらに**行列式の値を取る前と後のどっちで変数に値を代入しても、出てくる結果は同じ**です。というのも行列式は定義式が多少ややこしいものの、所詮は成分を適当に掛け算して足し合わせるだけに過ぎません。一方で任意の多項式 $f(x), g(x) \in \mathbb{R}[x]$ に対し、$h(x) := f(x)g(x)$ とおけば、全ての $a \in \mathbb{R}$ について $h(a) = f(a)g(a)$ が成り立ちます。多項式について掛け算と代入の順序が入れ替えられることから、行列式についても掛け算と数の代入の順序が入れ替えられます。

これを知っていれば、たとえば問 3 (2) は「$a = 0$ を代入したときに (1) と同じ値になるかどうか」で検算ができます。あるいは「最初に (2) を解いて、その結果に $a = 0$ を代入したものが (1) の答えになる」という手抜きもできます。ここで求めたのは上三角行列の行列式だから普通に解いても簡単ですが、このテクニック自体は覚えておいて損はしません。

4　置換と n 次対称群

行列式の定義式をもう一度見直してみましょう。n 次正方行列 $A = (a_{ij})_{1 \leq i,j \leq n}$ の行列式は

$$\det A = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

という式で定義されています。この式の $\sum_{\sigma \in \mathfrak{S}_n}$ は「集合 $\mathfrak{S}_n$ の全ての元 σ にわたって和を取る」という意味なので、行列式を理解するには “$\mathfrak{S}_n$” が何だか分からないと、話が進みません。また $\mathrm{sgn}(\sigma)$ という記号も、まだ定義していません。この式が読めるようになるよう、「n 次対称群 $\mathfrak{S}_n$」について調べてみましょう。

4.1　置換の定義

■定義　天下りですが、まず定義を与えてしまいましょう。$1, 2, \ldots, n$ を並べ替える写像を、$\{1, 2, \ldots, n\}$ の**置換**といいます。たとえば $n = 3$ なら

$$\sigma(1) = 3, \quad \sigma(2) = 1, \quad \sigma(3) = 2$$

で定まる写像が $\{1, 2, 3\}$ の置換です。σ に $1, 2, 3$ を放り込むと、それぞれの値として $3, 1, 2$ が出てきます。これは $1, 2, 3$ の並び替えになっていますね。他にもたとえば

$$\tau(2) = 1, \quad \tau(2) = 3, \quad \tau(3) = 2$$

も $\{1, 2, 3\}$ の置換です。$1, 2, 3$ を並べ替える方法は全部で $3! = 6$ 通りですから、上に挙げた他にも 4 つの置換があります。

もう少し写像としての見方を強調するなら、$1, 2, 3$ の並べ替えとは「値として $1, 2, 3$ がちょうど 1 度ずつ現れるような写像 $\{1, 2, 3\} \to \{1, 2, 3\}$」に他なりません。つまり $1, 2, 3$ **の置換とは、集合** $\{1, 2, 3\}$ **上の全単射**です。

■n **次対称群**　いま定義した $\{1,2,\ldots,n\}$ の置換全てを集めてできる集合を n **次対称群**[*8]といい、$\mathfrak{S}_n$ と書きます。

この見慣れない記号 ‘$\mathfrak{S}$’ は、フラクトゥール体 (いわゆるドイツ文字) でのアルファベット ‘S’ です。やたら飾りがでかいので、初見では「これのどこが S だよ！」と感じることでしょう。でも手書きにしてみると、中に S が入っているのが分かると思います[*9]。

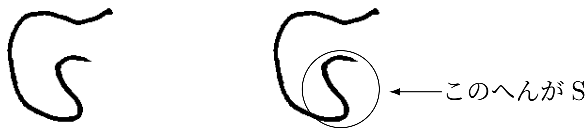
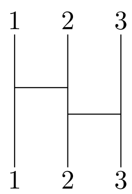


図 1.1　手書きの ‘$\mathfrak{S}$’

本によっては $\mathfrak{S}_n$ の代わりに S_n という記号を使っていることもあります。ですから皆さんが「n 次対称群の記号には S_n を使うんだ」と心の中で決心するなら、別に S_n と書いても何も問題ありません。ただ $\mathfrak{S}$ という記号を使うことにしておくと、$\mathfrak{S}$ という文字を見た瞬間に「対称群だ！」と気づくことができます。そういうささやかなメリットがあるからか、対称群を表すのに $\mathfrak{S}$ を使う伝統があるので、このプリントでも $\mathfrak{S}$ を使うことにします。

ちなみに英語に筆記体があるように、ドイツ語にも (英語とは異なる) 筆記体があります。かつて東大にいらした数学者の岩堀長慶先生[*10]は「‘$\mathfrak{S}$’ はあくまで印刷用の字体であるから、手書きで S を書く時はドイツ語の筆記体の S を使うべきだ」とおっしゃっていたそうです[*11]。こだわりたい人は、ドイツ語の筆記体を使ってみても良いでしょう[*12]。ただ残念ながらドイツ語の筆記体が読める人は (少なくとも、日本人の数学者には) あまりいないので、レポートや試験で使う時は断りを入れてください。

■**あみだくじ**　n 次対称群を表すのには、あみだくじを使うと便利です。たとえば $n=3$ のとき、$\sigma(1)=3, \sigma(2)=1, \sigma(3)=2$ で定まる置換を考えます。



このあみだくじの上にある 1, 2, 3 の行き先は、それぞれ 3, 1, 2 になっていますね。ですから σ があみだくじによって表されたと言って良いでしょう。

4.2　置換の表記

さて、n 次の置換 σ は $\sigma\colon \{1,2,\ldots,n\} \to \{1,2,\ldots,n\}$ という写像を与えています。したがって「σ がどんな写像であるか」を記述するには、$1,2,\ldots,n$ の行き先 $\sigma(1), \sigma(2), \ldots, \sigma(n)$ を全て書く必要があります。この表記法にはいくつかよく使われるものがあるので、それをまとめておきます。

[*8] 「群」という言葉の意味は、後で説明します。今は気にしないでいいです。

[*9] と言ってはみたものの、苦しい言い訳にしか聞こえないですよね……

[*10] 植野先生の師匠です。残念ながら、2011 年にご逝去されました。

[*11] このお話は、東大数理科学研究科の寺田至先生に伺いました。寺田先生の師匠も岩堀先生です。

[*12] ドイツ語筆記体の書き方については、たとえば林 メーナー エルケ『アルファベットの正しい書き方 –ドイツ語を例にとって』(ぎょうせい) に書いてあります。

■行列記法　集合 $\{1,2,\ldots,n\}$ の置換 σ を表すのに、「上の行に入力、下の行に出力」を並べた $2\times n$ の行列を使うことができます。たとえば $n=3$ で $\sigma(1)=3,\sigma(2)=1,\sigma(3)=2$ のときは

$$\sigma=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$$

といった感じです。より一般の場合も

$$\sigma=\begin{pmatrix}1&2&\cdots&n\\\sigma(1)&\sigma(2)&\cdots&\sigma(n)\end{pmatrix}$$

というルールにのっとって σ の行先を並べれば、σ がどんな写像なのかが分かります。最もよく使われる記法なので、意味をちゃんと抑えておいてください。

　また「上の行き先が下」というルールがあるので、1 行目が $1,2,\ldots,n$ の順番に並ぶ必要はありません。たとえば

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

といった書き方もできます。積極的に順番を乱す理由はないですが、後で逆置換を考えるときなどは、この記法が役に立ちます。

■互換　さて行列記法はどんな置換でも表せるものの、しばしば無駄が多いです。その際たる例は「2 つ数だけを入れ替え、他はそのままにする」という置換です。ほとんどの数がそのままにされるなら、省略したいと思うのが自然でしょう。そこで $i\neq j$ のとき、「i と j を入れ替え、他は何もしない」というルールで定まる置換を $(i\ j)$ と書きます。たとえば $n=5$ のとき

$$(2\ 5)=\begin{pmatrix}1&2&3&4&5\\1&5&3&4&2\end{pmatrix}$$

といった具合です。このように $(i\ j)$ の形で書ける置換を**互換**といいます。特に $j=i+1$ のとき、$(i\ i+1)$ を**隣接互換**といいます。

■巡回置換　互換は「2 つの数を入れ替える」という置換でしたが、これをもう少し拡張した「n 個の数を順繰りに入れ替える置換」というのも良く使います。このような置換を**巡回置換**といいます。

　巡回置換は、順繰りに入れ替わる数を 1 列に並べて表示します。たとえば $n=4$ で

$$\sigma=\begin{pmatrix}1&2&3&4\\4&3&1&2\end{pmatrix}$$

の場合、1 に σ を繰り返し施すと $1\to4\to2\to3\to1$ という順番で数が動きます。そこでこの σ を $\sigma=(1423)$ と書き表します。また σ を当てると数の列が巡回するので、1 から書き始める必要もありません。$(1423)=(4231)=(2314)=(3142)$ です。巡回置換に現れる数が 2 個のときは、巡回置換は互換になります。

「巡回置換に現れない数字はそのままにしておく」というルールも、互換のときと同じです。たとえば $n=4$ のとき、(243) は 2 を 4 に、4 を 3 に、そして 3 を 2 にうつす置換です。(243) は 1 を動かしません。つまり

$$(243)=\begin{pmatrix}1&2&3&4\\1&4&2&3\end{pmatrix}$$

です。

　なお、全ての置換が巡回置換の形で書けるわけではありません。たとえば

$$\sigma=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}$$

を何回か $1,2,3,4$ に施すと、$1\to2\to1$, $3\to4\to3$ という 2 つのサイクルがあることに気づきます。後で見るように、一般の場合でも置換を何個かの巡回置換に分解することができます。

4.3 置換の積

さて、置換が 2 つあったときに、それらの「積」を定義することができます。それを説明します。

$\sigma, \tau \in \mathfrak{S}_n$ とします。このとき σ, τ は共に $\{1, \ldots, n\} \to \{1, \ldots, n\}$ という写像なので、合成写像 $\sigma \circ \tau \colon \{1, \ldots, n\} \to \{1, \ldots, n\}$ を考えることができます。そこで $\sigma\tau := \sigma \circ \tau$ を σ と τ の**積**と定めます[*13]。

一つ例を見てみましょう。たとえば $n = 4$ で

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}, \quad \tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$$

とします。このとき

$$\sigma\bigl(\tau(1)\bigr) = \sigma(3) = 3 \qquad \sigma\bigl(\tau(2)\bigr) = \sigma(1) = 4 \qquad \sigma\bigl(\tau(3)\bigr) = \sigma(2) = 2, \qquad \sigma\bigl(\tau(4)\bigr) = \sigma(4) = 1$$

だから

$$\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 3 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$$

という具合です[*14]。

積を定義するにあたっては「$\mathfrak{S}_n$ の元が全単射」という事実が大事です。σ, τ は共に全単射なので、合成 $\sigma \circ \tau$ も全単射で、したがって $\mathfrak{S}_n$ の元となります。具体例だと気づきにくいですが、どんな $\sigma, \tau \in \mathfrak{S}_n$ を取ってきても $\sigma\tau \in \mathfrak{S}_n$ となることは、全単射という性質が保証してくれるのです。

ちなみにあみだくじで描くと、置換の積はあみだくじの連結に対応します。$\sigma, \tau \in \mathfrak{S}_n$ のとき、τ を表すあみだくじの下に σ を表すあみだくじをくっつけたものが、$\sigma\tau$ を表します。たとえば前のページに書いたあみだくじの図は $(132) = (23)(12)$ を表しています。確認してください。

■互換による表示　置換の中である意味最も簡単なのが、2 個の数を入れ替える互換です。あらゆる置換は、互換の積で表すことができます。たとえば

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 4 \end{pmatrix}\begin{pmatrix} 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 4 \end{pmatrix}$$

といった具合です。今の式変形を見れば分かるように、一般に $\sigma \in \mathfrak{S}_n$ に対して互換 $(i\ j)$ を右からかけると、$\sigma(i)$ と $\sigma(j)$ の値を入れ替えることができます。したがって、σ を行列表示したときに

- 右からうまく互換をかけて、σ の下の行で n を一番右に持っていく
- 右からうまく互換をかけて、σ の下の行で $n-1$ を右から 2 番目に持っていく
- ……

という操作を繰り返せば、最終的に σ を恒等置換に変形することができます。この式を見れば σ を互換の積に表す式が得られている、というわけです。上に書いた例も同じ手順で式変形をしています。確認してください。

また $i < j$ のとき

$$\begin{pmatrix} i & j \end{pmatrix} = \begin{pmatrix} i & i+1 \end{pmatrix}\begin{pmatrix} i+1 & i+2 \end{pmatrix} \cdots \begin{pmatrix} j-2 & j-1 \end{pmatrix}\begin{pmatrix} j-1 & j \end{pmatrix}\begin{pmatrix} j-2 & j-1 \end{pmatrix}\begin{pmatrix} j-3 & j-2 \end{pmatrix} \cdots \begin{pmatrix} i & i+1 \end{pmatrix}$$

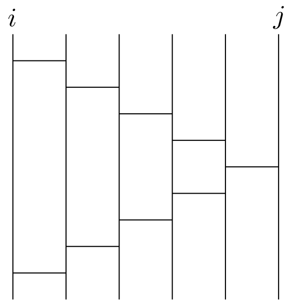
という式が成り立ちます。何をやっているのかと思うかもしれませんが、あみだくじで書いてみれば一発で分かります。次のあみだくじを辿って、

[*13] なんでこれが「積」と呼ぶにふさわしいかは、次回説明します。

[*14] 見た目は行列と同じですが、この文脈では置換を表しています。置換の積は**行列の積とは別物**ですので、ごっちゃにしないでください。

- 左端から出発したら右端に
- 右端から出発したら左端に
- 他の場所から出発したら元の位置の真下に

辿り着くことを確認してください。



これで結局

- 全ての置換が互換の積で表せる
- 全ての互換は隣接互換の積で表せる

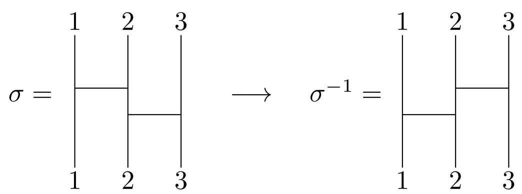
ことが分かったので、**全ての置換は隣接互換の積で表せる**ことが言えました。平たく言えば、「隣り合う縦線の間に横線を引く」という操作を使うだけで、どんな並びのあみだくじも作れるということです。

4.4 逆置換

$\{1, 2, \ldots, n\}$ の置換 $\sigma \in \mathfrak{S}_n$ が与えられると、$1, 2, \ldots, n$ のそれぞれに $\sigma(1), \sigma(2), \ldots, \sigma(n)$ という数が対応します。この $\sigma(1), \sigma(2), \ldots, \sigma(n)$ という数の並びには $1, 2, \ldots, n$ がちょうど 1 回ずつ登場するのでした。そこで「各 $\sigma(i)$ に対して i を対応させる」という方法を考えることで、別の置換を作ることができます[*15]。これを σ の**逆置換**といい、σ^{-1} で表します。

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \cdots & \sigma(n) \\ 1 & 2 & \cdots & n \end{pmatrix}$$

あみだくじで見れば、σ^{-1} を表すのは簡単です。あみだくじはいつも上から下に辿りますが、逆に下から上に辿ることもできますよね。なので σ が表すあみだくじの上下をひっくり返したものが、σ^{-1} に他なりません。



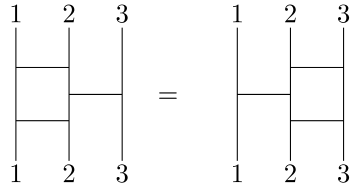
4.5 置換の符号

置換は互換の積で表せると言いましたが、その表し方は一通りではありません。たとえば

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

[*15] 要は、全単射 $\sigma\colon \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$ の逆写像を考えるということです。

が成り立ちます。事実、次のあみだくじは同じ結果を与えます。



ところが置換を互換の積で表す方法が何通りかあっても、「表すのに必要な互換の個数の偶奇」は常に一致するのです。上の例も確かにそうなっています。この事実を使って、$\sigma \in \mathfrak{S}_n$ が偶数個の互換の積で表されるとき**偶置換**、奇数個の互換の積で表されるとき**奇置換**といいます。そして σ が偶置換のとき $\mathrm{sgn}(\sigma) = 1$、奇置換のとき $\mathrm{sgn}(\sigma) = -1$ と定めます[*16]。この置換の符号がきちんと定まることを証明しましょう[*17]。

■置換の多項式への作用　突然ですが、一般に多項式 $f(x_1, \ldots, x_n)$ と $\{1, 2, \ldots, n\}$ の置換 $\sigma \in \mathfrak{S}_n$ が与えられたとき「$f(x_1, \ldots, x_n)$ の中で x_i を $x_{\sigma(i)}$ へと入れ替えて得られる多項式」を $\sigma f(x_1, \ldots, x_n)$ と表すことにします[*18]。すると、多項式を掛け算してから変数を入れ替えても変数を入れ替えてから掛け算しても結果が同じです。よって

$$\sigma\big(f(x_1, \ldots, x_n)g(x_1, \ldots, x_n)\big) = \big(\sigma f(x_1, \ldots, x_n)\big)\big(\sigma g(x_1, \ldots, x_n)\big)$$

が全ての多項式 $f(x_1, \ldots, x_n), g(x_1, \ldots, x_n)$ について成り立ちます。

■差積への作用　なぜいきなり多項式を持ち出したのかというと、もちろん「置換の符号」を引っ張り出すのに使うからです。差積と呼ばれる多項式

$$\Delta := \prod_{1 \leq i < j \leq n} (x_j - x_i)$$

を考えます。実はこのとき、任意の互換 $(p\ q)$ $(p < q)$ に対して $(p\ q)\Delta = -\Delta$ が成り立ちます[*19]。

この理由を調べるため、Δ に出てくる項を次のように三角形に並べてみます。これを全部かけたものが差積です[*20]。

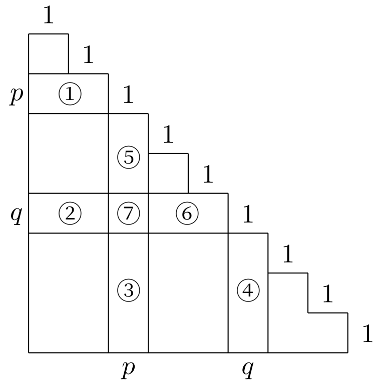
$$\begin{array}{cccccc}
1 & & & & & \\
x_2 - x_1 & 1 & & & & \\
x_3 - x_1 & x_3 - x_2 & 1 & & & \\
\vdots & \vdots & \ddots & \ddots & & \\
\vdots & \vdots & & \ddots & \ddots & \\
x_n - x_1 & x_n - x_2 & \cdots & \cdots & x_n - x_{n-1} & 1
\end{array}
\qquad
\begin{array}{cccccc}
1 & & & & & \\
x_2 - x_1 & 1 & & & & \\
x_3 - x_1 & x_3 - x_2 & 1 & & & \\
x_4 - x_1 & x_4 - x_2 & x_4 - x_3 & 1 & & \\
x_5 - x_1 & x_5 - x_2 & x_5 - x_3 & x_5 - x_4 & 1 & \\
x_6 - x_1 & x_6 - x_2 & x_6 - x_3 & x_6 - x_4 & x_6 - x_5 & 1
\end{array}$$

左側が一般の場合、右側が $n = 6$ の例です。さて、これに互換 $(p\ q)$ を当てた $(p\ q)\Delta$ を考えてみます。$(p\ q)$ を当てた時に影響を受ける項は

- $x_p - x_k$ $(1 \leq k < p)$ の形の項
- $x_k - x_p$ $(p \leq k < n)$ の形の項
- $x_q - x_k$ $(1 \leq k < q)$ の形の項
- $x_k - x_q$ $(q \leq k < n)$ の形の項

[*16] ふつう $\mathrm{sgn}(\sigma)$ は $\pm$ だけではなく、± 1 という数で表します。
[*17] ちょっとテクニカルな議論になるので、最初のうちは読み飛ばしても構いません。
[*18] このことを「置換 σ を多項式 $f(x_1, \ldots, x_n)$ に作用させる」などといいます。
[*19] 「こんなのどうして思いつくんだよ」と言う声が聞こえてきそうですが、対称式や交代式のことを詳しく知っていると、そこそこ思いつける式です。更なる後知恵ですが、この式は「n 次対称群 $\mathfrak{S}_n$ の符号表現に対応する Specht 加群」というものになっており、それを知っていると自然とこういう式を使えます。先人の知恵をありがたく拝借しましょう。
[*20] わざと 1 を並べたのには理由があります。後で互換の作用を調べるときに行番号をこのように調節しておくと、話を進めやすいのです。

の 4 種類です。これらの項の中で x_p が x_q に、x_q が x_p へと変化します。さっきの三角形に並べた式を見て、どこの項が $(p\ q)$ で影響を受けるのかを図にすると、こんな感じになります[*21]。①から⑦までの番号を振ったブロックが、影響を受ける場所です。



この式で x_p と x_q を入れ替えると、項は次のように変化します。

- ①と②が入れ変わる: $x_p - x_k \longleftrightarrow x_q - x_k\ (1 \le k < p)$
- ③と④が入れ変わる: $x_k - x_p \longleftrightarrow x_k - x_q\ (q < k \le n)$
- ⑤と⑥が (-1) 倍されて入れ替わる: $x_k - x_p \to -(x_q - x_k),\ x_q - x_k \to -(x_k - x_p)\ (p < k < q)$
- ⑦の $x_p - x_q$ は $x_q - x_p = -(x_p - x_q)$ に変化する

さらに⑤および⑥にある項の数はどちらも $q - p$ 個で同じですから、全部の項をかけてしまえば、この部分の -1 倍は打ち消し合います。よって結局、$x_p - x_q$ のところから出てくる -1 倍だけが残り、$(p\ q)\Delta = -\Delta$ が従います。

これが分かると、一般の置換 $\sigma \in \mathfrak{S}_n$ に対しても $\sigma\Delta$ が $\pm\Delta$ のどちらになるかが計算できます。互換を 1 個当てるごとに Δ が -1 倍されるのだから、σ を互換の積で表しておいて、その個数だけ -1 をかければよいのです。

さて $\sigma \in \mathfrak{S}_n$ を k 個の互換で表す方法と l 個の互換で表す方法があったとしましょう。Δ に互換 1 個を当てると -1 倍されるので、k 個の互換を当てれば $(-1)^k\Delta$ に、l 個の互換を当てれば $(-1)^l\Delta$ になります。ところがこれらが一致するので、k と l の偶奇は等しくないといけません。これで示すべきことが言えました。

■sgn の乗法性　さっきの $\mathrm{sgn}(\sigma)$ の定義は、(-1) を「σ を表すのに必要な互換の個数」乗したものと言っても同じです。したがって任意の $\sigma, \tau \in \mathfrak{S}_n$ に対して $\mathrm{sgn}(\sigma\tau) = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau)$ が成り立ちます。なぜなら

$$\sigma = \begin{pmatrix} a_1 & b_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \end{pmatrix}\cdots\begin{pmatrix} a_k & b_k \end{pmatrix}, \tau = \begin{pmatrix} p_1 & q_1 \end{pmatrix}\begin{pmatrix} p_2 & q_2 \end{pmatrix}\cdots\begin{pmatrix} p_l & q_l \end{pmatrix}$$

と表されるとき、

$$\sigma\tau = \begin{pmatrix} a_1 & b_1 \end{pmatrix}\begin{pmatrix} a_2 & b_2 \end{pmatrix}\cdots\begin{pmatrix} a_k & b_k \end{pmatrix}\begin{pmatrix} p_1 & q_1 \end{pmatrix}\begin{pmatrix} p_2 & q_2 \end{pmatrix}\cdots\begin{pmatrix} p_l & q_l \end{pmatrix}$$

より、$\mathrm{sgn}(\sigma\tau) = (-1)^{k+l} = (-1)^k(-1)^l = \mathrm{sgn}(\sigma)\,\mathrm{sgn}(\tau)$ となるからです。この公式はぜひ覚えておきましょう。

■問 1 の解答

$$\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 3 & 1 & 2 \end{pmatrix}, \quad \sigma\tau\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix}$$

である。また $\sigma = (14)(34)(24)$ より $\mathrm{sgn}(\sigma) = -1$ なので、$\mathrm{sgn}(\sigma^9) = \mathrm{sgn}(\sigma)^9 = -1$ である。 ■

[*21] $n = 6, p = 3, q = 5$ の例で確認してみてください。

4.6 行列式の定義

ここまでの準備を経て、ようやく det の定義をきちんと読むことができます。n 次正方行列 $A=(a_{ij})_{1\leq i,j\leq n}$ の行列式の定義

$$\det A := \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$$

は「各 $\sigma\in\mathfrak{S}_n$ に対し $\operatorname{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}\cdots a_{n\sigma(n)}$ を計算し、それを全部足し合わせる」という意味の式です[*22]。n が小さい場合に、この式に現れる項を書き下してみましょう。

まずは $n=2$ の場合です。$\{1,2\}$ の置換は恒等置換 id[*23]と互換 $(1\ 2)$ の 2 つしかありません。$s=(1\ 2)$ と書くことにすると、$\mathfrak{S}_2=\{\mathrm{id},s\}$ で、かつ $\operatorname{sgn}(\mathrm{id})=1$, $\operatorname{sgn}(s)=-1$ です。したがって

$$\sum_{\sigma\in\mathfrak{S}_2} \operatorname{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)} = \operatorname{sgn}(\mathrm{id})a_{1\,\mathrm{id}(1)}a_{2\,\mathrm{id}(2)} + \operatorname{sgn}(s)a_{1s(1)}a_{2s(2)} = a_{11}a_{22}-a_{12}a_{21}$$

となります。これは紛れもなく 2 次の行列式ですね。

次に $n=3$ でやってみます。$\{1,2,3\}$ を並べ替えるやり方は全部で 6 通りだから、$\mathfrak{S}_3$ は 6 個の元からなります。そこで各 $\sigma\in\mathfrak{S}_3$ について $\operatorname{sgn}(\sigma)$ と $\operatorname{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}$ の表を作ってみます。

表 1.1　$\mathfrak{S}_3$ の元の一覧

置換	$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$	$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$	$\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$
符号	1	-1	-1	1	1	-1
対応する項	$a_{11}a_{22}a_{33}$	$-a_{12}a_{21}a_{33}$	$-a_{11}a_{23}a_{32}$	$a_{12}a_{23}a_{31}$	$a_{13}a_{21}a_{32}$	$-a_{13}a_{22}a_{31}$

この表に基づいて det の定義式を書き下してみると

$$\sum_{\sigma\in\mathfrak{S}_3} \operatorname{sgn}(\sigma)a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)} = a_{11}a_{22}a_{33}-a_{12}a_{21}a_{33}-a_{11}a_{23}a_{32}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32}-a_{13}a_{22}a_{31}$$

となります。Sarrus の方法で計算する行列式と、確かに一致していますね。n が 4 以上の場合も同じようにやれば、計算ができるはずです。ただし $n!$ 個の項が出てきて大変なので、わざわざ書き下してみる必要はありません。「この式で定義ができる」ということが分かれば十分です。

こんな感じで、行列式はきちんと定義するのにも一苦労するわけですが、僕たちはさらに行列式の計算公式を証明しないといけません。次の週は、この定義式をもとに高次の行列式を計算するための様々な公式を導きましょう。

[*22] S1 タームの「数理科学基礎 (線形代数学)」の 5 回目 p. 37 に、この $\sum$ の使い方を述べています。必要に応じて復習してください。
[*23] 「何も並べ替えない」という置換のことです。恒等写像が表す置換なので、id という記号を使います。

線形代数学② 第 2 回　　行列式の性質

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 9 月 23 日 1 限

今回の演習問題は「行列式の計算」がメインです。愚直に計算すればどの問題も解けはしますが、単に解くだけではなく「手際のよい解き方」を考えることがポイントです。少し時間をかけて、公式の証明と計算例を眺めてみましょう。

1　置換について

前回、置換に関する基礎的な事柄を述べました。しかしその際一つ大事なことを言い忘れたので、まずはその補足をします。その後で、置換の符号を計算するのに役立つ公式を 2 つ紹介します。

1.1　置換の積の記法について

前回話し忘れた大事なことは「**置換の積の順序には流儀がある**」ということです。

前回のプリントの中で「2 つの置換 $\sigma, \tau \in \mathfrak{S}_n$ の積を $\sigma\tau := \sigma \circ \tau$ で定義する」と書きましたが、授業では「$\sigma\tau := \tau \circ \sigma$」で定義していました。見れば分かる通り、掛け算の順序が逆になっています。

どうしてこんなことになったのかというと、$\sigma\tau$ を「$\tau \circ \sigma$ で定義する流儀」と「$\sigma \circ \tau$ で定義する流儀」の 2 つが世の中に存在するからです。そして流儀が 2 つあることの理由は、一方の表記が絶対的に優れているわけではなく、一長一短だからです。たとえば写像の合成を考えるだけなら、$\sigma\tau := \sigma \circ \tau$ としておいた方が明らかに見やすいです。でも $\sigma\tau := \tau \circ \sigma$ とした方が式が見やすくなる場面も後々出てくるのです。どちらが良いという単純な話ではありません。本によっても流儀はまちまちです。たとえば佐武一郎『線型代数学』(裳華房) では $\sigma\tau := \sigma \circ \tau$ が、足助太郎『線型代数学』(東京大学出版会) では $\sigma\tau := \tau \circ \sigma$ が採用されています。

先生に確認したところ「きちんと定義がされていれば、どちらの流儀を使っても良い」とおっしゃっていたので、もし授業と違う流儀を採用する場合は、そのことを明示してください。また、このプリントでは授業と同じ流儀に合わせることにします[*1]。

1.2　サイクル分解

巡回置換に対しては、符号が簡単に計算できます。なぜなら集合 $\{1, 2, \ldots, n\}$ から相異なる数 $a_1, a_2, \ldots, a_l$ を取るとき

$$(a_1\ a_2\ \cdots\ a_l) = (a_{l-1}\ a_l)(a_{l-2}\ a_{l-1}) \cdots (a_1\ a_2)$$

という関係式があるからです。l 個の数を入れ替える巡回置換は $l-1$ 個の互換の積で書けるので、その符号は $(-1)^{l-1}$ となります。特に**巡回置換の符号は、長さだけで決まり**、中に入っている数に依存しません。

この事実は、一般の $\sigma \in \mathfrak{S}_n$ の符号の求めるのにも使えます。というのも、**全ての置換は巡回置換の積に表せる**からです。いま $\sigma \in \mathfrak{S}_n$ を 1 に何回も施すことを考えます。すると $1 \mapsto \sigma(1) \mapsto \sigma^2(1) := \sigma\bigl(\sigma(1)\bigr) \mapsto \cdots$ というようにどんどん数が変わりますが、所詮は $\{1, 2, \ldots, n\}$ という有限集合の中を動き回るだけなので、いつか 1 に戻ってきます。$\sigma^k(1) = 1$ となる最小の正の整数 k を取ると、$1, \sigma(1), \sigma^2(1), \ldots$ に対する σ の振る舞いは、巡回置換 $(1\ \sigma(1)\ \sigma^2(1)\ \cdots)$ と変わらないわけです。次に、$1, \sigma(1), \sigma^2(1), \ldots$ に現れない $\{1, 2, \ldots, n\}$ の元があったら、それを a_2 として、再び $a_2, \sigma(a_2), \sigma^2(a_2), \ldots$ という列を考えます。この列もやはり有限回で元に戻るので、巡回置換が得られます。しかもこ

[*1] といっても、日頃の癖で書き間違えてしまう可能性は十分ありえます。間違いに気づいたら教えてください。

の列の中には $1, \sigma(1), \sigma^2(1), \ldots$ は一度も登場しません。このように

- ある数から出発し、σ を施して得られる列からなる巡回置換を作る
- 巡回置換を作った後に余っている数があったら、そのうちどれかを起点に取り直す

という操作を繰り返すことによって、σ は巡回置換の積に書き表せるのです。たとえば

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 4 & 1 & 2 \end{pmatrix} = (1\ 3\ 4)(2\ 5)$$

という感じです。そして一つ一つの巡回置換の符号は既に計算したので、あとは全てをかけることで全体の符号が得られます。

1.3 転倒数と符号

$\sigma \in \mathfrak{S}_n$ に対し、「$1 \leq i < j \leq n$ かつ $\sigma(i) > \sigma(j)$」を満たす (i, j) の組み合わせの個数を σ の**転倒数**といい、$\mathrm{inv}(\sigma)$ で表します。たとえば $n = 4$ で

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 1 & 3 \end{pmatrix}$$

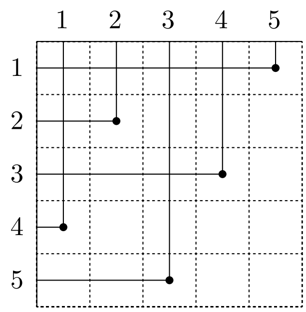
なら、$i < j$ かつ $\sigma(i) > \sigma(j)$ となる組合せは、$(i, j) = (1, 2), (1, 3), (1, 4), (1, 5), (2, 4), (3, 4), (3, 5)$ の 7 つです。よって $\mathrm{inv}(\sigma) = 7$ という感じです。

実は符号と転倒数に関して、$\mathrm{sgn}(\sigma) = (-1)^{\mathrm{inv}(\sigma)}$ という公式が成り立ちます。言い方を変えると、置換の符号が転倒数の偶奇に対応すると言っているのです。この事実を確かめましょう。

■Rothe 図　置換の転倒数を見るときに便利な表現を紹介します。それは **Rothe 図**と呼ばれるもの (を少しいじったもの) です[*2]。定義をいきなり書くより絵で見た方が早いので、今の

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 2 & 4 & 1 & 3 \end{pmatrix}$$

を例にして図を描いてみます。



この図を見れば、大体やっていることは分かるはずです。

1. サイズが $n \times n$ の正方形の格子を用意する
2. 各 $1 \leq i \leq n$ に対し、$\bigl(i, \sigma(i)\bigr)$ の位置にある正方形の中心に点を打つ
3. それぞれの点から、左と上の方向にまっすぐな線を引く

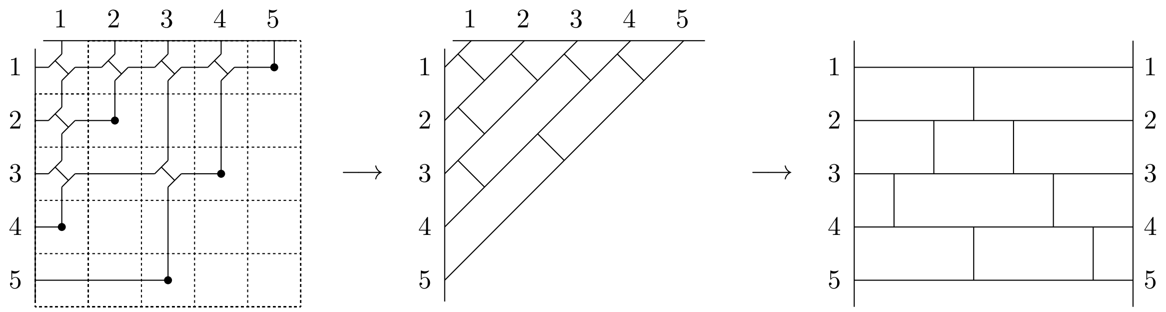
---

*2 ここに描いた図では、それぞれの点から真っ直ぐな線が左および上に出ています。ふつうの Rothe 図ではこの向きを逆にし、線を右および下に生やします。どっちでも情報量は同じですが、後でこの図を変形する都合上、ここでは左と上にしました。

という方法で図を作っています。これが Rothe 図です。

Rothe 図が便利なのは、転倒数を「線の交点の個数」として読み取れるからです。Rothe 図では $i<j$ のとき、「$\sigma(i)>\sigma(j)$」と「縦の i から出た線が縦の j から出た線とクロスする」ことが同値になるように線を引いています。だから $i<j$ かつ $\sigma(i)>\sigma(j)$ を満たす (i,j) のペアと Rothe 図の交点が 1 : 1 に対応します。

■転倒数と符号の関係　そして Rothe 図を変形することによって、どうして転倒数が符号と関係するのかが分かります。Rothe 図の中で、線の交点を「あみだくじの横棒」に置き換えます。Rothe 図をこう変形して「右または上の方向へ進むあみだくじ」と読むのです。



ここに現れたあみだくじをたどると、元々 Rothe 図の交点だったところでは「左から来たら右へ行く」「下から来たら上へ行く」ようになっています。そして元々の Rothe 図で、左の i から出た線は上の $\sigma(i)$ のところに辿りつくのでした。ですからこのあみだくじは、σ を実現するあみだくじに他なりません。

この対応によって σ は、Rothe 図における交点の個数と同じ個数の隣接互換の積で表せることが分かりました。これが $\mathrm{sgn}(\sigma)=(-1)^{\mathrm{inv}(\sigma)}$ の理由です。ちなみに Rothe 図を使うこの方法では「置換をできるだけ少ない個数の隣接互換の積で表すやり方」が得られることが知られています。

実際に今の公式を使って問題を解いてみましょう。転倒数を使うと、置換の符号判定が大分楽になります。

■「偶順列と奇順列」の言葉遣い　紛らわしい話で申し訳ないのですが、「順列」というときと「置換」というときとでは指すものが違っています。たとえば順列 (321) と書いたときは

$$\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}$$

を考えていますが、巡回置換 (321) と書いたときは

$$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}$$

を表しています。見た目が同じ記号でも文脈で表すものが違うので、気を付けてください。

■問 1 の解答　転倒数を数えれば (1) 奇 (2) 偶 (3) 奇 (4) 偶 と分かる。 ■

■問 5 の解答　転倒数を数えれば (1) 奇 (2) 偶 (3) 偶 (4) 奇 と分かる。 ■

■問 7 の解答　(1) $a_{12}a_{23}a_{31}a_{44}$ は巡回置換 $\sigma=(123)$ に対応する項なので、符号は $\mathrm{sgn}(\sigma)=+1$ (2) $a_{14}a_{23}a_{32}a_{41}$ は $\sigma=(14)(23)$ に対応する項なので、符号は $\mathrm{sgn}(\sigma)=+1$ ■

2 行列式の交代性と多重線型性

行列式には様々な公式があります。その中でも、行列式を特徴づける最も重要な性質は「交代性」および「多重線型性」と呼ばれる性質です。最初に「転置をとっても行列式が変わらないこと」を示してから、まずはこの 2 つの性質の証明を目標にします。

証明にあたっては「$\sigma \in \mathfrak{S}_n$ にわたる和」に対する式変形を繰り返すことになります。慣れるまでは多少戸惑うかもしれないので、最初のうちは丁寧めに証明をつけておきます。1 ステップずつじっくり追いかけてください。

2.1 行と列の入れ替えに関する対称性

これから先に出てくる公式は、行と列の両方について成り立つものばかりです。行と列のことを一々別個に証明をしていては面倒なので、先に**行列式は転置しても変わらない**という事実を示しておきましょう。これさえ知っておけば、例えば行に関する公式を転置によって列に関する公式に読み替えることができるようになります。なぜなら転置によって行は列に、列は行に変わるからです。

n 次正方行列 $A \in \mathrm{Mat}_n(\mathbb{R})$ を成分で $A = (a_{ij})_{1 \leq i,j \leq n}$ と表示しておきましょう。このとき A を転置した行列 ${}^tA$ の (i,j) 成分は a_{ji} です。よって

$$\det {}^tA = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$$

が成り立ちます。ここで $\sigma(1), \sigma(2), \ldots, \sigma(n)$ が $1, 2, \ldots n$ の並び替えであることを思い出すと、$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n}$ の掛け算の順番を入れ替えて $a_{1*} a_{2*} \cdots a_{n*}$ の恰好にすることができます[*3]。さらに元々の式に出てくる $a_{\sigma(i)i}$ という成分は「2 つ目の添字に σ を施すと 1 つ目の添字になる」という性質を持っています。このことから

$$a_{\sigma(1)1} a_{\sigma(2)2} \cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}$$

であると分かります。この議論によって

$$\det {}^tA = \sum_{\sigma \in \mathfrak{S}_n} \mathrm{sgn}(\sigma) a_{1\sigma^{-1}(1)} a_{2\sigma^{-1}(2)} \cdots a_{n\sigma^{-1}(n)}$$

が分かりました。

この式は「$\sum$ の添字が σ、成分の右下に現れるのは σ^{-1}」なので、det の定義式とは微妙に違っています。でも心配は要りません。σ が $\mathfrak{S}_n$ の元を全て動くとき、σ^{-1} もまた $\mathfrak{S}_n$ の元を動くからです。また σ^{-1} と σ の符号は同じです。そこで σ^{-1} を τ と表すと

$$\det {}^tA = \sum_{\tau \in \mathfrak{S}_n} \mathrm{sgn}(\tau) a_{1\tau(1)} a_{2\tau(2)} \cdots a_{n\tau(n)} = \det A$$

となります。

■問 8　証明そのものは上の議論で終わっていますが、3 次の行列式なら直接計算して確かめることもできます。やってみてください。

[*3] たまに行列の成分として普通の数ではなく、非可換な数、つまり $ab \neq ba$ を満たす数を考えることがあります。たとえば四元数 (クォータニオン) と呼ばれる数がこれに当たります。このような場合はもちろん、掛け算の順番を入れ替えられません。ですが今ここで考えている数は実数ないし複素数ですから、掛け算の順序を気にする必要はありません。

2.2 交代性

行列式には「2 つの行 / 列を入れ替えると (-1) 倍される」という性質があります。たとえば 2 次だと

$$\det\begin{pmatrix} b & a \\ d & c \end{pmatrix} = bc - ad = -(ad - bc) = -\det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

となっていますね。この性質を行列式の**交代性**といいます。2 次に限らず一般の次数でも行列式の交代性が成り立つことを示します。

行列 A の k 列目と l 列目を入れ替えた行列を A' とし、A を成分と列ベクトルで表しましょう。すなわち

$$\begin{aligned} A &= (a_{ij})_{1\leq i,j\leq n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_n \end{pmatrix} \\ A' &= (a'_{ij})_{1\leq i,j\leq n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_n \end{pmatrix} \end{aligned}$$

とおきます。このとき $A = (a_{ij}), A' = (a'_{ij})$ と書くことにすると

$$a'_{ij} = \begin{cases} a_{ij} & (j \neq k, l) \\ a_{il} & (j = k) \\ a_{ik} & (j = l) \end{cases}$$

なので

$$\det A' = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a'_{\sigma(1)1} \cdots a'_{\sigma(k)k} \cdots a'_{\sigma(l)l} \cdots a'_{\sigma(n)n} = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(k)l} \cdots a_{\sigma(l)k} \cdots a_{\sigma(n)n}$$

となります。k 番目と l 番目の項だけが規則から外れているので気を付けてください。この 2 つの項を並び替えると

$$\det A' = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(l)k} \cdots a_{\sigma(k)l} \cdots a_{\sigma(n)n}$$

となります。k 番目と l 番目の因子の添字を見ると、2 文字目は他と揃いましたが、1 文字目がまだ規則から外れています。ここで σ に左から互換 $(k\ l)$ をかけた置換 $(k\ l)\sigma$ を考えると、$i \neq k, l$ なら $(k\ l)(i) = i$ となるので[*4]

$$\big((k\ l)\sigma\big)(i) = \begin{cases} \sigma(i) & (i \neq k, l) \\ \sigma(l) & (i = k) \\ \sigma(k) & (i = l) \end{cases}$$

が成り立ちます。これを使うと、さっきの $\det A'$ の式を

$$\det A' = \sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{((k\ l)\sigma)(1)1} \cdots a'_{((k\ l)\sigma)(k)k} \cdots a'_{((k\ l)\sigma)(l)l} \cdots a'_{((k\ l)\sigma)(n)n}$$

と変形することができます。k 番目と l 番目まで込めて**添字が全て規則通りに並んだ**ことに注目してください。もう少しで行列式の定義式がひねり出せそうです。ここでさらに $\operatorname{sgn}(\sigma) = -\operatorname{sgn}\big((k\ l)\sigma\big)$ に気を付けると

$$\det A' = -\sum_{\sigma\in\mathfrak{S}_n} \operatorname{sgn}\big((k\ l)\sigma\big) a_{((k\ l)\sigma)(1)1} \cdots a'_{((k\ l)\sigma)(k)k} \cdots a'_{((k\ l)\sigma)(l)l} \cdots a'_{((k\ l)\sigma)(n)n}$$

が得られます。これで行列式の定義式とほとんど同じになりました。あと一息です。

最後に σ は $\mathfrak{S}_n$ の全ての元を動くとき、$(k\ l)\sigma$ もまた $\mathfrak{S}_n$ の元を全て動くことに気を付けましょう。したがって $(k\ l)\sigma = \tau$ と書くことにすれば

$$\det A' = -\sum_{\tau\in\mathfrak{S}_n} \operatorname{sgn}(\tau) a_{\tau(1)1} \cdots a'_{\tau(k)k} \cdots a'_{\tau(l)l} \cdots a'_{\tau(n)n} = -\det A$$

となります。

[*4] 丸括弧が並んで微妙に見辛いですが、1 つ目の $(k\ l)$ は互換で、2 つ目の (i) は「写像に i を代入する」という意味です。

■同じ列をもつ行列式　交代性からすぐに従う帰結の一つに、**行列式に同じ列があったら値は** 0 という事実があります。実際 A の i 列目と j 列目が同じベクトル $\boldsymbol{a}$ だったとしましょう。

$$A = \begin{array}{ccccc} & i^{\mathrm{th}} & & j^{\mathrm{th}} & \\ \big(\cdots & \boldsymbol{a} & \cdots & \boldsymbol{a} & \cdots\big) \end{array}$$

このとき

- A の i 列目と j 列目を入れ替えると、行列式の値は -1 倍される
- A の i 列目と j 列目は元々同じベクトルなので、入れ替えても A のまま

なので、$\det A = -\det A$ が従います。これより $\det A = 0$ です。交代性がこの形で役立つことはよくあるので、覚えておきましょう。もちろん、2 つの行が同じである行列についても、行列式は 0 になります。

2.3　多重線型性

　次に示すのは、多重線型性と呼ばれる性質です。n 次正方行列は n 本の列ベクトルの集まりと同じなので、det は n 本の列ベクトルに数を対応させる写像 $\det\colon \mathbb{R}^n \times \mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ と思うことができます。このとき det **は** 1 **つ** 1 **つのベクトルについて線型**です。すなわち、全ての $1 \leq j \leq n$ に対して

$$\det \begin{array}{c} \begin{array}{ccc} & j^{\mathrm{th}} & \\ \end{array} \\ \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_j + \boldsymbol{b} & \cdots \boldsymbol{a}_n \end{pmatrix} \end{array} = \det \begin{array}{c} \begin{array}{ccc} & j^{\mathrm{th}} & \\ \end{array} \\ \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_j & \cdots & \boldsymbol{a}_n \end{pmatrix} \end{array} + \det \begin{array}{c} \begin{array}{ccc} & j^{\mathrm{th}} & \\ \end{array} \\ \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{b} & \cdots & \boldsymbol{a}_n \end{pmatrix} \end{array}$$

$$\det \begin{array}{c} \begin{array}{ccc} & j^{\mathrm{th}} & \\ \end{array} \\ \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \alpha\boldsymbol{a}_j & \cdots & \boldsymbol{a}_n \end{pmatrix} \end{array} = \alpha \det \begin{array}{c} \begin{array}{ccc} & j^{\mathrm{th}} & \\ \end{array} \\ \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_j & \cdots & \boldsymbol{a}_n \end{pmatrix} \end{array}$$

が成り立つのです。この「一つ一つについて線型」という性質を**多重線型性**といいます。これも証明しましょう。

■列ベクトルの和　まず和に関する振る舞いを確認します。n 次正方行列 A を成分および列ベクトルで

$$\begin{aligned} A &= (a_{ij})_{1 \leq i,j \leq n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_k \end{pmatrix} \\ A' &= (a'_{ij})_{1 \leq i,j \leq n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_l + \boldsymbol{b} & \cdots & \boldsymbol{a}_k \end{pmatrix} \end{aligned}$$

と表しておきます。A の l 列目を $\boldsymbol{a}_l$ から $\boldsymbol{a}_l + \boldsymbol{b}$ に置き換えた行列が A' です。このとき det が行と列の入れ替えで対称なことより

$$\det A' = \det {}^t\!A' = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a'_{\sigma(1)1} \cdots a'_{\sigma(i)i} \cdots a'_{\sigma(n)n}$$

が得られます。そして A と A' は l 列目に $\boldsymbol{b} = {}^t(b_1, b_2, \ldots, b_n)$ の差があるだけなので

$$a'_{ij} = \begin{cases} a_{ij} & (j \neq l) \\ a_{il} + b_i & (j = l) \end{cases}$$

が成り立ちます。これを今の式に代入すると

$$\begin{aligned} \det A' &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots (a_{\sigma(l)l} + b_{\sigma(l)}) \cdots a_{\sigma(n)n} \\ &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(l)l} \cdots a_{\sigma(n)n} + \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots b_{\sigma(l)} \cdots a_{\sigma(n)n} \\ &= \det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_k \end{pmatrix} + \det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{b} & \cdots & \boldsymbol{a}_k \end{pmatrix} \end{aligned}$$

が得られます。同じことが行ベクトルでも成り立ちます。

■列ベクトルの定数倍　次に、列ベクトルを定数倍したらどうなるか調べてみましょう。

$$A = (a_{ij})_{1 \le i,j \le n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_k \end{pmatrix}$$
$$A' = (a'_{ij})_{1 \le i,j \le n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \alpha\boldsymbol{a}_l & \cdots & \boldsymbol{a}_k \end{pmatrix}$$

とおきます。A の l 列目を α 倍した行列が A' です。このとき l 列目だけ $a'_{il} = \alpha a_{il}$ で、他の列については $a'_{ij} = a_{ij}$ が成り立っています。よって

$$\begin{aligned} \det A' &= \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a'_{\sigma(1)1} \cdots a'_{\sigma(l)l} \cdots a'_{\sigma(n)n} = \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots \bigl(\alpha a_{\sigma(l)l}\bigr) \cdots a_{\sigma(n)n} \\ &= \alpha \sum_{\sigma \in \mathfrak{S}_n} \operatorname{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(l)l} \cdots a_{\sigma(n)n} = \alpha \det A \end{aligned}$$

となります。行ベクトルについても同じです。

3　行列式の計算公式とその応用

ここまでで行列式が転置に関して変わらないことと、交代性・多重線型性という性質を示しました。これでどんな行列も、ある程度計算できるようになります。ただ、これ以外にも

- ある行に別の行の定数倍を足しても、det の値は変わらないこと
- ブロック三角行列の行列式公式
- 余因子展開公式
- $\det AB = \det A \det B$

という便利な計算公式があります。行列式の計算にあたっては、これらの公式を的確に組み合わせることがキモです。また行列式を使うと、正則性の判定ができたり、逆行列の計算ができるようになったりもします。ちょっと長くなりますが、そうした便利な公式たちを片っ端から示しておきましょう。

3.1　基本変形に対する振る舞い

行列に対する「基本変形」という操作を思い出しましょう。連立 1 次方程式を解く時に使う、係数行列に対する

- ある行を定数倍する
- 2 つの行を入れ替える
- ある行の定数倍を別の行に加える

という 3 つの操作を行基本変形というのでした。また「行」を「列」で置き換えると、列基本変形になります。

列基本変形によって行列式の値がどう変わるのかを調べましょう。といっても最初の 2 つは交代性と多重線型性そのものです。最後の 1 つだけ示しましょう。

行列 $A,\ A'$ を

$$A = (a_{ij})_{1 \le i,j \le n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_k \end{pmatrix}$$
$$A' = (a'_{ij})_{1 \le i,j \le n} = \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_l + \alpha\boldsymbol{a}_k & \cdots & \boldsymbol{a}_k \end{pmatrix}$$

で定めます。A の k 列目を α 倍して l 列目に加えた行列が A' です。このとき多重線型性より

$$\det A' = \det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_l & \cdots & \boldsymbol{a}_k \end{pmatrix} + \alpha \det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_k & \cdots & \boldsymbol{a}_k \end{pmatrix}$$

が従います。そして右辺第 2 項の行列式には同じ列があるので、交代性から値は 0 です。これで $\det A' = \det A$ が言えました。行についても同様です。

この公式は非常に強力です。というのも「ある行を別の行に足しても行列式が変わらない」ということは、ある成分を要に**行を掃き出しても行列式が変わらない**と言っているわけです。掃き出しをすると行列の中に 0 がたくさん出てくるので、余因子展開をするときの計算が非常に楽になります。

3.2 ブロック三角行列の行列式

正方行列 A が、ブロック分けによって

$$A = \begin{pmatrix} B & O \\ C & D \end{pmatrix}$$

と分けられ、かつ B, D が共に正方行列となっているとき、A を**ブロック下三角行列**といいます。B, D は正方行列でさえあればよく、サイズが違っても構いません。このとき $\det A = \det B \det D$ という公式が成り立ちます。これを示します。

$B = (b_{ij})_{1 \le i,j \le p}$ が p 次、$D = (d_{ij})_{1 \le i,j \le q}$ が q 次の正方行列としましょう。そして $A = (a_{ij})_{1 \le i,j \le p+q}$ とおきます。このとき $1 \le i, j \le p$ なら $a_{ij} = b_{ij}$ で、$1 \le i, j \le q$ なら $a_{p+i,p+j} = d_{ij}$ です。さて定義に従えば

$$\det A = \sum_{\sigma \in \mathfrak{S}_{p+q}} a_{1\sigma(1)} \cdots a_{p\sigma(p)} a_{p+1,\sigma(p+1)} \cdots a_{p+q,\sigma(p+q)}$$

です。ところが A の右上 $p \times q$ ブロックの成分は全て 0 です。したがって $\sigma(1), \ldots, \sigma(p)$ の中に 1 個でも $p+1$ 以上の数が出てきたら、その σ に対応する項は 0 です。よって $\det A$ に寄与する項の σ は、$\sigma(1), \ldots, \sigma(p)$ の値が全て p 以下です。元々 σ は $1, 2, \ldots, p+q$ の並び替えですが、ここで効いてくる σ は $1, 2, \ldots, p$ を $1, 2, \ldots, p$ のどれかに並び替えないといけないのです。このことから、σ は $p+1, \ldots, p+q$ を $p+1, \ldots, p+q$ のどれかに並び替えることも従います。そこで $\mathfrak{S}_n$ の元のうち、$1, \ldots, p$ を $1, \ldots, p$ の中で、$p+1, \ldots, p+q$ を $p+1, \ldots, p+q$ の中で並べ替える置換を集めた部分集合

$$\mathfrak{S}_p \times \mathfrak{S}_q := \left\{ \sigma \in \mathfrak{S}_n \mid \sigma(i) \in \begin{cases} \{1, 2, \ldots, p\} & (1 \le i \le p) \\ \{p+1, p+2, \ldots, p+q\} & (p+1 \le i \le p+q) \end{cases} \right\}$$

を考えます。このような置換は $\{1, 2, \ldots, p\}$ の置換と $\{p+1, p+2, \ldots, p+q\}$ の置換のペアで一意的に表せるので、$\mathfrak{S}_p \times \mathfrak{S}_q$ という記号を使っています。

さて、ここまでの考察で

$$\det A = \sum_{\sigma \in \mathfrak{S}_p \times \mathfrak{S}_q} a_{1\sigma(1)} \cdots a_{p\sigma(p)} a_{p+1,\sigma(p+1)} \cdots a_{p+q,\sigma(p+q)}$$

が言えました。ここで $\mathfrak{S}_p \times \mathfrak{S}_q$ の元は、$\{1, 2, \ldots, p\}$ の置換と $\{p+1, p+2, \ldots, p+q\}$ の置換のペアで一意的に表せます。$\tau \in \mathfrak{S}_p$ と $\rho \in \mathfrak{S}_q$ のペアに対応する $\mathfrak{S}_{p+q}$ の元を $\tau \times \rho$ と書くこと[*5]にすれば

$$\det A = \sum_{\tau \in \mathfrak{S}_p} \sum_{\rho \in \mathfrak{S}_q} a_{1(\tau\times\rho)(1)} \cdots a_{p(\tau\times\rho)(p)} a_{p+1,(\tau\times\rho)(p+1)} \cdots a_{p+q,(\tau\times\rho)(p+q)}$$

と書けます。そして $(\tau \times \rho)(i) = \tau(i)$ と $(\tau \times \rho)(p+i) = p + \rho(i)$ に注目すれば

$$\det A = \sum_{\tau \in \mathfrak{S}_p} \sum_{\rho \in \mathfrak{S}_q} a_{1\tau(1)} \cdots a_{p\tau(p)} a_{p+1,p+\rho(1)} \cdots a_{p+q,p+\rho(q)} = \sum_{\tau \in \mathfrak{S}_p} a_{1\tau(1)} \cdots a_{p\tau(p)} \sum_{\rho \in \mathfrak{S}_q} a_{p+1,p+\rho(1)} \cdots a_{p+q,p+\rho(q)}$$

[*5] ここで $\rho \in \mathfrak{S}_q$ は $\{1, 2, \ldots, q\}$ の置換ですが、$\tau \in \mathfrak{S}_p$ とペアにして $\tau \times \rho \in \mathfrak{S}_{p+q}$ を考えるときは、ρ を「$p+k$ を $p+\rho(k)$ にずらす置換」と読み替えることにします。あみだくじで言えば、「τ を表すあみだくじ」と「ρ を表すあみだくじ」を横に並べたものが $\tau \times \rho$ です。

となるが、$1 \le i,j \le p$ のとき $a_{ij} = b_{ij}$、$1 \le i,j \le q$ のとき $a_{p+i,p+j} = d_{ij}$ なので、

$$\det A = \sum_{\tau \in \mathfrak{S}_p} b_{1\tau(1)} \cdots b_{p\tau(p)} \sum_{\rho \in \mathfrak{S}_q} d_{1,\rho(1)} \cdots d_{q,\rho(q)} = \det B \det D$$

が得られます。

ブロック上三角行列の場合は、転置と組み合わせて

$$\det \begin{pmatrix} B & C \\ O & D \end{pmatrix} = \det {}^t\begin{pmatrix} B & C \\ O & D \end{pmatrix} = \det \begin{pmatrix} {}^tB & {}^tO \\ {}^tC & {}^tD \end{pmatrix} = \det {}^tB \det {}^tD = \det B \det D$$

が分かります。また、ブロックが何段かある場合でも

$$\det \begin{pmatrix} A_1 & * & * \\ O & A_2 & * \\ O & O & A_3 \end{pmatrix} = \det A_1 \det A_2 \det A_3$$

が成り立ちます。実際、今示した 2 段の場合の公式を繰り返し使えば

$$\det \left(\begin{array}{c|cc} A_1 & * & * \\ \hline O & A_2 & * \\ O & O & A_3 \end{array}\right) = \det A_1 \det \begin{pmatrix} A_2 & * \\ O & A_3 \end{pmatrix} = \det A_1 \det A_2 \det A_3$$

となるからです。4 段以上でも同様なのは明らかでしょう。

■問 3 の解答　次のようにブロック分解すれば分かる。

$$\det \left(\begin{array}{cc|c} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ \hline 0 & 0 & a_{33} \end{array}\right) = a_{33} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \det \left(\begin{array}{cc|cc} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ \hline 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & a_{43} & a_{44} \end{array}\right) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \det \begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix}$$

■

3.3　余因子展開

■余因子とは　A を n 次正方行列、$1 \le i,j \le n$ とします。このとき A から i 行目と j 列目を取り去ってできる行列 $\tilde{A}_{ij}$ の行列式 $\det \tilde{A}_{ij}$ に $(-1)^{i+j}$ をかけた $\Delta_{ij} := (-1)^{i+j} \det \tilde{A}_{ij}$ を A の (i,j) **余因子**といいます。余因子を使うことで、高次の行列式の計算に帰着させることができます。

余因子展開を考えるにあたり、一番基本的なのは

$$\det \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2} & \cdots & a_{nn} \end{pmatrix} = a_{11} \det \begin{pmatrix} a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots \\ a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

という事実です。ブロック三角行列の展開公式から直ちに従うこの公式を、これから幾度となく使います。

■1 列目に対する余因子展開公式　まずは 1 列目に関する余因子展開公式を示しましょう。それは n 次正方行列 A に対する

$$\det A = \sum_{i=1}^{n} a_{i1} \Delta_{i1}$$

という公式です。たとえば 3 次なら

$$\det A = a_{11}a_{22}a_{33} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

$$= a_{11}\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{21}\det\begin{pmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{pmatrix} + a_{31}\det\begin{pmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{pmatrix}$$

といった感じです。これが一般の n で正しいことを証明しましょう。

いつものように、A を成分と列ベクトルとで $A = (a_{ij})_{1\leq i,j\leq n} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix}$ と表します。このとき $\boldsymbol{a}_1 = a_{11}\boldsymbol{e}_1 + a_{12}\boldsymbol{e}_2 + \cdots + a_{1n}\boldsymbol{e}_n$ だから、多重線型性より

$$\begin{aligned} \det A &= a_{11}\det\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} + a_{21}\det\begin{pmatrix} \boldsymbol{e}_2 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} + \cdots + a_{n1}\det\begin{pmatrix} \boldsymbol{e}_n & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} \\ &= \sum_{i=1}^{n} a_{i1}\det\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} \end{aligned}$$

と表せます。ここで出てくる行列 $\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix}$ の 1 列目を見ると、i 行目のみに 1 があり、他の行の成分は全て 0 です。よって $(i,1)$ 成分を要にして 2 列目以降を掃き出すことで

$$\det\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} = \det\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,2} & \cdots & a_{i-1,n} \\ 1 & a_{i,2} & \cdots & a_{i,n} \\ 0 & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = \det\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,2} & \cdots & a_{i-1,n} \\ 1 & 0 & \cdots & 0 \\ 0 & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix}$$

となります。この右辺の行列で、i 行目を $i-1$ 行目と入れ替え、$i-1$ 行目を $i-2$ 行目と入れ替え……という操作を順番に行うと、行ベクトル $(1,0,\ldots,0)$ を一番上の行まで持ってくることができます。この操作は $i-1$ 回行われ、その都度 (-1) がかかるので、結局

$$\begin{aligned} \det\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} &= (-1)^{i-1}\det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,2} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} = (-1)^{i+1}\det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{i-1,2} & \cdots & a_{i-1,n} \\ 0 & a_{i+1,2} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n,2} & \cdots & a_{n,n} \end{pmatrix} \\ &= \Delta_{i1} \end{aligned}$$

が得られました。よって

$$\det A = \sum_{i=1}^{n} a_{i1}\det\begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix} = \sum_{i=1}^{n} a_{i1}\Delta_{i1}$$

となります。

■一般の余因子展開公式　今の議論が分かってしまえば、後の話は簡単です。j 列目に関して余因子展開をするときは

1. $\boldsymbol{a}_j$ を $\boldsymbol{a}_j = a_{1j}\boldsymbol{e}_1 + a_{2j}\boldsymbol{e}_2 + \cdots + a_{nj}\boldsymbol{e}_n$ と書き直した上で、行列式を多重線型性でバラし
2. $\boldsymbol{e}_i$ を含む行列式の i 行目を、(i,j) 成分を要にして掃き出し
3. i 行目を 1 行目に、j 列目を 1 列目に動かす

というステップで、j 列目に関する余因子展開公式

$$\det A = \sum_{i=1}^{n} a_{ij}\Delta_{ij}$$

の証明ができます。全く同様に、i 行目に関する余因子展開公式

$$\det A = \sum_{j=1}^{n} a_{ij}\Delta_{ij}$$

も得られます。証明で唯一気を付けるべきは、3 番目のステップで i 行目を 1 行目に、j 列目を 1 列目に持っていくとき「(-1) が何個くっつくか」を正しく数えることです。$\pm$ の符号が交互に並ぶことは簡単に分かりますが、**最初の項が $\pm$ のどっちから始まるのか慎重に数えないと全体の符号を間違えます**。余因子展開は a_{ij} に「A の i 行目と j 列目を取り除いた行列の行列式と $(-1)^{i+j}$ をかけたもの」を足し合わせる公式です。符号を正しく覚えてください。

3.4 正則性の判定条件と乗法性

次に $\det AB = \det A \det B$ という公式を示します。この公式には色々な証明がありますが、ここでは正則性の判定条件と絡め

1. 非正則な行列 A が $\det A = 0$ を満たすこと
2. A, B の少なくとも一方が非正則なときに $\det AB = \det A \det B$ が正しいこと
3. A, B が正則で、かつ一方が基本行列のときに $\det AB = \det A \det B$ が正しいこと
4. 正則行列 A は $\det A \neq 0$ を満たすこと

という順番で証明をしていきます。

■正則性の言い換え　n 次正方行列 A については、単射性と全射性が同値だったことを思い出しましょう。したがって n 次正方行列 A が正則であることは、n 本の列ベクトルたちが 1 次独立なことと同値です。

■正則でない行列の行列式　まず、非正則行列の行列式が 0 になることを示します。いま $A = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_n \end{pmatrix}$ が正則でないとしましょう。このとき $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ は 1 次従属なので、$\alpha_1\boldsymbol{a}_1 + \alpha_2\boldsymbol{a}_2 + \cdots + \alpha_n\boldsymbol{a}_n = \boldsymbol{0}$ を見たすスカラーの組 $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ で、$(0, 0, \ldots, 0)$ 以外のものが存在します。このとき $\alpha_i \neq 0$ なる i を 1 つ取ると

$$\boldsymbol{a}_i + \sum_{j \neq i} \frac{\alpha_j}{\alpha_i}\boldsymbol{a}_j = \boldsymbol{0}$$

が成り立ちます。したがって全ての $j \neq i$ について j 列目の $\frac{\alpha_j}{\alpha_i}\boldsymbol{a}_j$ 倍を i 列目に足せば

$$\det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_i & \cdots & \boldsymbol{a}_n \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{a}_i + \sum_{j \neq i} \frac{\alpha_j}{\alpha_i}\boldsymbol{a}_j & \cdots & \boldsymbol{a}_n \end{pmatrix} = \det \begin{pmatrix} \boldsymbol{a}_1 & \cdots & \boldsymbol{0} & \cdots & \boldsymbol{a}_n \end{pmatrix} = 0$$

となります。

また、正方行列 A が正則でなければ、どんな正方行列 B を右からかけても AB は正則になりません。もし AB が正則なら $I = AB(AB)^{-1} = A\bigl(B(AB)^{-1}\bigr)$ となり、A が非正則であることに矛盾するからです。ですから A が正則でなければ、$\det AB = \det A \det B$ という式は両辺とも 0 で成り立ちます。B が非正則な場合も同様です。

■正則行列が基本行列の積で書けること　次に A, B が同じサイズの正則行列であるときに、$\det AB = \det A \det B$ が成り立つことを示しましょう。それには、正則行列が基本行列の積で書けるという事実を思い出す必要があります。

いま A が正則行列だったとします。このとき逆行列 A^{-1} が存在するので、連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ の解は $\boldsymbol{x} = \boldsymbol{0}$ ただ一つです。一方、連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{0}$ は掃き出し法で解くことができます。解がただ一つ存在する方程式なので、係数行列 A に適切な基本変形を施すことで、A を単位行列まで変形できます。これは A^{-1} が基本行列の積で書けることに他なりません。そして基本行列の逆行列もまた基本行列ですから、A 自身も基本行列の積で表されます。

■乗法性の証明　今の事実を使って、乗法性を示しましょう。全ての正則行列は基本行列の積で書けるから、我々は A が基本行列だと仮定して $\det AB = \det A \det B$ を示せば十分です。

まず A が「B の i 行目を c 倍する」基本変形だったとしましょう。このとき

$$B = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

なので、$\det B = c$ です。一方行列式の多重線型性から、B の i 行目を α 倍した行列 AB の行列式は $\det AB = c \det B$ を満たしています。これで $\det AB = \det A \det B$ が確認できました。

次に A が「2 つの行を入れ替える」基本変形だったとしましょう。このとき A は

$$A = \begin{pmatrix} 1 & & & & & & & & & & \\ & \ddots & & & & & & & & & \\ & & 1 & & & & & & & & \\ & & & 0 & & & & 1 & & & \\ & & & & 1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & 1 & & & & \\ & & & 1 & & & & 0 & & & \\ & & & & & & & & 1 & & \\ & & & & & & & & & \ddots & \\ & & & & & & & & & & 1 \end{pmatrix}$$

という格好です。A の 2 つの行を入れ替えると単位行列になるので、$\det A = -\det I = -1$ と分かります。一方 AB は、B の 2 つの行を入れ替えた行列ですから、やはり $\det AB = -\det B$ を満たします。よってここでも $\det AB = \det A \det B$ が確かめられました。

最後に「ある行の c 倍を別の行に足す」という基本変形を考えます。このとき A は

$$A = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & c & & \\ & & & \ddots & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

という行列です。c と同じ列にある 1 を要に掃き出しをすれば A を単位行列に変形できるので、$\det A = \det I = 1$ です。また AB は B のある行に別の行の定数倍を足した行列なので、$\det AB = \det B$ です。この場合も確かに $\det AB = \det A \det B$ です。

以上、場合分けによっていずれの場合も $\det AB = \det A \det B$ が正しいことを示せました。

■正則な行列の行列式　乗法性を使うことで、正則行列の行列式が決して 0 にならないことが示せます。A が仮に n 次の正則行列だったとすると、逆行列 A^{-1} が存在し、$I = AA^{-1}$ を満たします。この両辺で det を取ると、乗法性か

ら $1 = \det I = \det AA^{-1} = \det A \det A^{-1}$ となります。もし $\det A$ が 0 なら、何をかけても 1 にならないので矛盾が起きます。よって $\det A \neq 0$ です。

また、今の計算から $\det A^{-1} = 1/\det A$ という式も得られました。この式もよく使います。

■問 11 の解答　三角行列なので $\det A = 1$ と $\det B = 3$ はすぐ求まる。あとは乗法性から $\det AB = \det A \det B = 3$, $\det B^{-1} = 1/\det B = 1/3$ と求まる。 ■

■問 12 の解答　$\det AB = \det A \det B$ と $\det A^{-1} = 1/\det A$ より、直ちに $\det AB = \det BA$, $\det(ABA^{-1}B^{-1}) = 1$ が従う。 ■

3.5 余因子行列と Cramer の公式

■余因子行列と逆行列　もう一度余因子展開の公式を眺めてみましょう。n 次正方行列 $A = (a_{ij})_{1\leq i,j\leq n}$ の行列式 $\det A$を i 行目に関して余因子展開すると

$$\det A = \sum_{j=1}^{n} a_{ij}\Delta_{ij}$$

となるのでした。この式の右辺をよーく眺めると、行列の積の形をしています:

$$\det A = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \begin{pmatrix} \Delta_{i1} \\ \Delta_{i2} \\ \vdots \\ \Delta_{in} \end{pmatrix}$$

これは全ての i について正しい式です。そこで行列 $\operatorname{adj} A$ を

$$\operatorname{adj} A := \begin{pmatrix} \Delta_{11} & \Delta_{21} & \cdots & \Delta_{n1} \\ \Delta_{12} & \Delta_{22} & \cdots & \Delta_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{1n} & \Delta_{2n} & \cdots & \Delta_{nn} \end{pmatrix}$$

で定めると[*6]、積 $A \operatorname{adj} A$ の i 番目の対角成分は

$$(A \operatorname{adj} A)_{ii} = \sum_{j=1}^{n} a_{ij}\Delta_{ij} = \det A$$

となります。そして $i \neq j$ のとき、$A \operatorname{adj} A$ の (i,j) 成分は

$$(A \operatorname{adj} A)_{ij} = \sum_{k=1}^{n} a_{ik}\Delta_{jk}$$

となります。ここで余因子展開公式を逆向きに使うと、右辺の値は「A の j 行目を i 行目で置き換えた行列 A' の行列式」だと分かります。ところが A' は i 行目と j 行目が一致しているので、$\det A' = 0$ です。かくして $A \operatorname{adj} A$ の非対角成分は全て 0 だと分かります。

以上より

$$A \operatorname{adj} A = \begin{pmatrix} \det A & & & \\ & \det A & & \\ & & \ddots & \\ & & & \det A \end{pmatrix} = (\det A)I$$

[*6] 添字の番号付けに気を付けてください。$\operatorname{adj} A$ は (i,j) 成分が Δ_{ji} です。

と分かりました。この「(i,j) 成分に A の (j,i) 余因子 Δ_{ji} を置いてできる行列」を A の**余因子行列**といい、$\operatorname{adj} A$ で表します[*7]。既に A が正則なことと $\det A \neq 0$ が同値なことを示していますから、この式から

$$A \frac{1}{\det A} \operatorname{adj} A = I$$

が分かります。$(1/\det A)\operatorname{adj} A$ は右から A にかけると単位行列になります。同様にして $(1/\det A)\operatorname{adj} A$ を左から A にかけると、列に関する余因子展開を用いることで、$(\operatorname{adj} A)A = (\det A)I$ を示せます。よって $(1/\det A)\operatorname{adj} A$ は左右どちらから A にかけても単位行列になるので、A の逆行列です。これで A^{-1} を具体的に書き下す公式が得られました。

■Cramer の公式　余因子行列を使うと、A が正則行列の場合に、連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解を書き下す公式が得られます。この場合 A の逆行列が存在するので $\boldsymbol{x} = A^{-1}\boldsymbol{b}$ が唯一の解となります。しかも僕たちは A^{-1} を求める公式を知っているので、これを代入すれば解の表式が得られるという算段です。

$A^{-1} = (1/\det A)\operatorname{adj} A$ なので、$(\operatorname{adj} A)\boldsymbol{b}$ を計算しましょう。$\boldsymbol{b} = {}^t(b_1, b_2, \ldots, b_n)$ とすれば、ベクトル $(\operatorname{adj} A)\boldsymbol{b}$ の第 j 成分は

$$\bigl((\operatorname{adj} A)\boldsymbol{b}\bigr)_j = \sum_{k=1}^{n} \Delta_{kj} b_k$$

です。ここで列に関する余因子展開公式を逆に使えば

$$\bigl((\operatorname{adj} A)\boldsymbol{b}\bigr)_j = \sum_{k=1}^{n} \Delta_{kj} b_k = \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2,j-1} & b_2 & a_{2,j+1} & \cdots & a_{2n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{pmatrix}$$

となります。A の j 列目を $\boldsymbol{b}$ で置き換えて得られる行列を A'_j と表すと、この式は $\det A'_j$ に他なりません。したがって連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の解は

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = \frac{1}{\det A}(\operatorname{adj} A)\boldsymbol{b} = \frac{1}{\det A}\begin{pmatrix} \det A'_1 \\ \det A'_2 \\ \vdots \\ \det A'_n \end{pmatrix}$$

となります。これを **Cramer の公式**といいます。行列式が 0 でないときは、いつでもこの方法で解を求めることができます[*8]。

4　行列式の計算例

めでたく公式を証明し終わったので、色々な計算をしてみましょう。公式の使いどころを感じてください。

[*7] この名前には注意すべき点があり、東大数理科学研究科にいらっしゃる足助太郎先生の web サイトに解説が載っています: `http://www.ms.u-tokyo.ac.jp/~asuke/linear_algebra/corrections.html` これによれば、(i,j) 成分に A の (j,i) 余因子 Δ_{ji} を置いてできる行列を

- 英語ではadjugate matrix
- 日本語では余因子行列

と呼ぶことが多いです。しかし

- 余因子は英語で “cofactor”
- 英語では “cofactor matrix” は “adjugate matrix” の転置を表す

ので、日本語の「余因子行列」は英語の直訳になっていません。ここではよく使われる記号に合わせて $\operatorname{adj} A$ を採用しましたが、英語の文献を読むときは気を付けてください。

[*8] ただし、Cramer の公式はいつでも使えるものの、計算にはあまり向きません。見ての通り式中に det が $n+1$ 個現れるので、愚直に計算すると $n+1$ 次の行列式を $n+1$ 個計算しなければいけません。普通の人が計算するのは $n=3$ くらいが限界でしょう。行列が特別な形をしている等の事情がない限り、実際の計算に使うのはお勧めしません。

4.1 Vandermonde 行列式と Schur 多項式

■Vandermonde 行列式　いきなりですが

$$\det\begin{pmatrix}1 & x_1 & x_1^2 & \cdots & x_1^{n-1}\\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1}\end{pmatrix} = \prod_{1\le i<j\le n}(x_j - x_i)$$

という公式が知られています。この左辺の行列式を **Vandermonde 行列式**というのでした。前回 $n=3$ の場合に Vandermonde 行列式が差積と一致することを示しましたが、今回は一般の n で証明します。

Vandermonde 行列式の値を $V(x_1,x_2,\dots,x_n)$ と書くことにします。この行列の (i,j) 成分は x_i^{j-1} なので、

$$V(x_1,x_2,\dots,x_n) = \sum_{\sigma\in\mathfrak{S}_n}\mathrm{sgn}(\sigma)x_{\sigma(1)}^0 x_{\sigma(2)}^1\cdots x_{\sigma(n)}^{n-1}$$

です。このことから、$V(x_1,x_2,\dots,x_n)$ は $x_1,x_2,\dots,x_n$ の多項式だと分かります。さらに $\sum$ で足される各項は、σ が何であるかに関わらず、次数は $0+1+\cdots+n-1=\frac{n(n-1)}{2}$ です。よって $V(x_1,x_2,\dots,x_n)$ は斉次 $\frac{n(n-1)}{2}$ 式だと分かりました。

さて行列式には「同じ行があったら値が 0」という公式がありました。よって $i\neq j$ のとき $x_i=x_j$ おけば、$V(x_1,x_2,\dots,x_n)=0$ となります。したがって $i\neq j$ となる全ての (i,j) の組合せについて、多項式 $V(x_1,x_2,\dots,x_n)$ は x_i-x_j で割り切れます。このことから結局

$$V(x_1,x_2,\dots,x_n) = c(x_1,x_2,\dots,x_n)\prod_{1\le i<j\le n}(x_j-x_i)$$

という形に書けると分かります。ここで $V(x_1,x_2,\dots,x_n)$ と $\prod_{1\le i<j\le n}(x_j-x_i)$ が共に斉次多項式なので、商 $c(x_1,x_2,\dots,x_n)$ も斉次多項式だと分かります。そして両辺の次数を比較すると

$$\frac{n(n-1)}{2} = \deg V(x_1,x_2,\dots,x_n) = \deg c(x_1,x_2,\dots,x_n) + \deg\prod_{1\le i<j\le n}(x_j-x_i) = \deg c(x_1,x_2,\dots,x_n) + \frac{n(n-1)}{2}$$

なので、$\deg c(x_1,x_2,\dots,x_n)=0$ です。つまり c は定数でなければいけないと分かります。

あとは $x_2x_3^2\cdots x_n^{n-1}$ の係数を比較すれば c が求まります。両辺とも $x_2x_3^2\cdots x_n^{n-1}$ の係数が $+1$ なので、$c=1$ が得られます。これで Vandermonde 行列式が差積そのものであると示せました。

この議論は実はすごく役立つことが知られています。応用例として、問題を解いてみましょう。

■問 15 (5), (6) の解答　(5) は Vandermonde 行列式そのもので $(b-a)(c-a)(c-b)$ である。(6) も Vandermonde 行列式と同様に、行列式の定義を考えれば、a,b,c の 4 次斉次多項式になることが分かる。そして $a=b$, $a=c$, $b=c$ とおけば値が 0 になるから、det の値は $(b-a)(c-a)(c-b)$ で割り切れて

$$\det\begin{pmatrix}1 & a & a^3\\ 1 & b & b^3\\ 1 & c & c^3\end{pmatrix} = (b-a)(c-a)(c-b)f(a,b,c)$$

と書ける。

この式の両辺で a と b を入れ替えると、左辺の det は 1 行目と 2 行目が入れ替わるので -1 倍される。一方 $(b-a)(c-a)(c-b)$ もまた -1 倍される。よって $f(a,b,c)=f(b,a,c)$ でないといけない。同様に b と c、a と c を

入れ替えても、f の値は変化しない。これより f は a,b,c の対称多項式だと分かる。f の次数は $4-3=1$ だから、$f(a,b,c)=\alpha(a+b+c)$ と書ける。

$$\det\begin{pmatrix}1 & a & a^3\\ 1 & b & b^3\\ 1 & c & c^3\end{pmatrix}=\alpha(b-a)(c-a)(c-b)(a+b+c)$$

残る α を決定するには、係数比較をすれば良い。左辺からは bc^3 という項が出る。右辺から bc^3 という項が出てくるには、$\alpha=+1$ でないといけない。これで答えが求まった。 ∎

■問 9 (2), (3) と問 10 (1) の解答　Vandermonde 行列式を使うと、問 9 (2) は $(2-1)(3-1)(3-2)=2$、問 10 (1) は $(2-1)(3-1)(3-2)(4-1)(4-2)(4-3)=12$ と求まる。また問 9 (3) は問 15 (6) の結果を使えば $(2-1)(3-1)(3-2)(1+2+3)=12$ と分かる。 ∎

■問 15 (3), (4) の解答　どちらの行列式も、結果は a,b,c の多項式である。そしてどちらの行列式も、$a=b$ とおくと 2 行目と 3 行目が一致するので値は 0 になる。したがって $(b-a)$ で割り切れる。全く同様にして $(c-a)$ と $(c-b)$ でも割り切れることが分かるので

$$\det\begin{pmatrix}1 & b+c & a^2\\ 1 & c+a & b^2\\ 1 & a+b & c^2\end{pmatrix}=\alpha(b-a)(c-a)(c-b),\det\begin{pmatrix}1 & a+b & ab\\ 1 & b+c & bc\\ 1 & c+a & ca\end{pmatrix}=\beta(b-a)(c-a)(c-b)$$

と書ける。どちらの行列式も結果は a,b,c の 3 次斉次式なので、α,β はただの数である。あとは行列式を定義通り展開することを考えると、$\sigma=\mathrm{id}$ に対応する項が $+ac^2$ という項を含むことが分かる。これと係数が合うためには、$\alpha=\beta=-1$ でないといけない。すなわち

$$\det\begin{pmatrix}1 & b+c & a^2\\ 1 & c+a & b^2\\ 1 & a+b & c^2\end{pmatrix}=\det\begin{pmatrix}1 & a+b & ab\\ 1 & b+c & bc\\ 1 & c+a & ca\end{pmatrix}=-(b-a)(c-a)(c-b)=(a-b)(a-c)(b-c)$$

である。 ∎

■Schur 多項式　Vandermonde 行列式をちょっといじって、長さ n の自然数の列 $\alpha=(\alpha_1,\alpha_2,\ldots,\alpha_n)$ に対し

$$a_\alpha(x_1,x_2,\ldots,x_n):=\det\begin{pmatrix}x_1^{\alpha_1} & x_1^{\alpha_2} & x_1^{\alpha_3} & \cdots & x_1^{\alpha_n}\\ x_2^{\alpha_1} & x_2^{\alpha_2} & x_2^{\alpha_3} & \cdots & x_2^{\alpha_n}\\ \vdots & \vdots & \vdots & \ddots & \vdots\\ x_n^{\alpha_1} & x_n^{\alpha_2} & x_n^{\alpha_3} & \cdots & x_n^{\alpha_n}\end{pmatrix}$$

という形の行列式を考えてみましょう。$\delta_n:=(n-1,n-2,\ldots,1,0)$ とおくと、$\alpha=\delta_n$ のときが、さっき計算した Vandermonde 行列式 (の ±1 倍) です。$\alpha_1,\alpha_2,\ldots,\alpha_n$ の中に同じ数があると det の値が 0 になってしまうので、α_i たちは全て相異なるとしておきます。また列の並び替えをしても、(-1) 倍が何個かかかるだけなので、最初から $\alpha_1>\alpha_2>\cdots>\alpha_n$ と仮定してしまいます。

さて Vandermonde 行列式が差積であることを示すとき、行列式の交代性から「$x_i=x_j$ なら det の値が 0 になる」ことを利用しました。いまの場合でも、全く同じことが言えます。$a_\alpha(x_1,x_2,\ldots,x_n)$ は差積で割り切れ、a_α/a_{δ_n} は斉次多項式になるのです。さらに a_α の中で x_i と x_j を入れ替えたら、2 つの行が入れ替わるので符号が -1 倍されます。よって任意の置換 $\sigma\in\mathfrak{S}_n$ に対し $\sigma a_\alpha(x_1,x_2,\ldots,x_n)=\mathrm{sgn}(\sigma)a_\alpha(x_1,x_2,\ldots,x_n)$ が成り立ちます。かくして

$$\sigma\left(\frac{a_\alpha(x_1,x_2,\ldots,x_n)}{a_{\delta_n}(x_1,x_2,\ldots,x_n)}\right)=\frac{\sigma a_\alpha(x_1,x_2,\ldots,x_n)}{\sigma a_{\delta_n}(x_1,x_2,\ldots,x_n)}=\frac{\mathrm{sgn}(\sigma)a_\alpha(x_1,x_2,\ldots,x_n)}{\mathrm{sgn}(\sigma)a_{\delta_n}(x_1,x_2,\ldots,x_n)}=\frac{a_\alpha(x_1,x_2,\ldots,x_n)}{a_{\delta_n}(x_1,x_2,\ldots,x_n)}$$

となるので、多項式 a_α/a_{δ_n} はどんな置換 $\sigma\in\mathfrak{S}_n$ で $x_1,x_2,\ldots,x_n$ を入れ替えても不変、すなわち対称式だとが分かります。こうして僕たちは、対称多項式を系統的に作り出す方法を手に入れました。

ちなみにこの方法で多項式を作る場合、諸々の都合で $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ のことを $(\lambda_1 + n-1, \lambda_2 + n-2, \ldots, \lambda_n)$ と書くことが多いです。最初に $\alpha_n > \alpha_{n-1} > \cdots > \alpha_1$ と仮定したので、$\alpha_k \geq k-1$ が常に成り立ちます。そのため $\lambda_k := \alpha_k - k + 1$ とおけば、$\lambda_k \geq 0$ となります。そして $\alpha_{k+1} > \alpha_k$ より、$\lambda_{k+1} \geq \lambda_k$ です。こうして自然数の狭義単調減少列 α と、自然数の広義単調減少列 λ を対応付けることができました。そして λ に対し

$$s_\lambda(x_1, x_2, \ldots, x_n) := \frac{a_{\lambda+\delta_n}(x_1, x_2, \ldots, x_n)}{a_{\delta_n}(x_1, x_2, \ldots, x_n)}$$

と書き、これを **Schur 多項式**といいます。Schur 多項式は非常に面白い性質や計算公式をたくさん持っており、対称多項式の理論で中心的な役割を果たしています。今でも Schur 多項式から派生した様々な多項式や函数が、研究対象となっています。興味がある人は

- 岡田聡一『古典群の表現論と組合せ論 (下)』(培風館)
- I. G. Macdonald, "Symmetric functions and Hall polynomials, 2nd ed.," Oxford University Press

などを読んでみてください。

4.2 複素数の行列による実現

問 13, 14 はそのまま定義通り計算すればできる問題です。が、普通に計算してみた人は「どうも複素数と似てるっぽい」ことに気づいたのではないでしょうか？ それは正しいです。

2 次正方行列 J を

$$J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

とおきます。このとき $J^2 = -I$ が成り立つので、

- $(aI + bJ) + (cI + dJ) = (a+c)I + (b+d)J$
- $(aI + bJ)(cI + dJ) = (ac - bd)I + (ad + bd)J$

という式が従います。これは**複素数の計算公式と全く同じ**です。すなわち $aI + bJ \leftrightarrow a + bi$ という対応によって、集合

$$\left\{ \begin{pmatrix} x & -y \\ y & x \end{pmatrix} \mid x, y \in \mathbb{R} \right\}$$

と複素数全体の集合 $\mathbb{C}$ とが $1:1$ に対応し、しかもこの対応の下では足し算・掛け算という演算までもがぴったり対応づくのです。また、行列式を計算すると $\det(aI + bJ) = a^2 + b^2$ と分かります。これは複素数の側で見れば、絶対値の 2 乗 $|a + bi|^2$ ですね。この対応さえ見破ってしまえば、計算はサクサク進むはずです。

ちなみに、行列の側で出てくる成分は実数だけです。ですから今の対応は、**行列を使って複素数を構成する方法**を与えることにもなっています。

■問 14 直接計算すれば $\det XY = (ac - bd)^2 + (ad + bc)^2$ と $\det X \det Y = (a^2 + b^2)(c^2 + d^2)$ が求まる。展開すれば、これらが等しいことが分かる。 ∎

■問 13 (1) は問 14 で得られる $(ac - bd)^2 + (ad + bd)^2 = (a^2 + b^2)(c^2 + d^2)$ という式に $c = a$, $b = d$ を代入すれば得られる。(2) も手間がかかるが

$$A^3 = (x^3 - 3xy^2)^2 I + (3x^2 y - y^3) J$$

を頑張って計算した上で、$\det A^3 = (\det A)^3$ の式に代入すればよい。 ∎

4.3 計算問題の解答

残った計算問題の答えを、まとめて書いておきます。今回問題となったほとんどの行列式は、工夫をすると簡単に計算できるので、どう公式を使ったのかも合わせて記します。参考にしてください。また、ここに書いてあるものが唯一の答えではないので、もっと手際のよい計算方法があるかもしれません。みなさん各自で考えてみてください。

■問 2 の解答　(1) 2 行目を 3 行目から引き、その後 1 行目を 2 行目から引いて、交代性を使う。

$$\det\begin{pmatrix}7&8&9\\8&9&10\\9&10&11\end{pmatrix}=\det\begin{pmatrix}7&8&9\\8&9&10\\1&1&1\end{pmatrix}=\det\begin{pmatrix}7&8&9\\1&1&1\\1&1&1\end{pmatrix}=0$$

(2) 2 行目と 3 行目を 1 行目に足すと、多重線型性で $a+b+c$ がくくり出せる。

$$\det\begin{pmatrix}a&c&b\\b&a&c\\c&b&a\end{pmatrix}=\det\begin{pmatrix}a+b+c&a+b+c&a+b+c\\b&a&c\\c&b&b\end{pmatrix}=(a+b+c)\det\begin{pmatrix}1&1&1\\b&a&c\\c&b&a\end{pmatrix}$$

さらに $(1,1)$ 成分を要に 1 行目を掃き出すと

$$\begin{aligned}\det\begin{pmatrix}a&c&b\\b&a&c\\c&b&a\end{pmatrix}&=(a+b+c)\det\begin{pmatrix}1&1&1\\b&a&c\\c&b&a\end{pmatrix}=(a+b+c)\det\begin{pmatrix}1&0&0\\b&a-b&c-b\\c&b-c&a-c\end{pmatrix}\\&=(a+b+c)(a^2+b^2+c^2-ab-bc-ca)\end{aligned}$$

となる。　■

■問 4 の解答　いずれも左上の 1×1 ブロックと右下の 2×2 ブロックに分けて計算すればすぐ解ける。答えは (1) 18 (2) 82 (3) 8 となる。　■

■問 6 の解答　(2) 0 と (3) $\cos^2\theta$ はそのまま計算すれば良い。(4) は 1 行目で余因子展開すれば 112 と求まる。(1) と (5) は掃き出しを使えば

$$\det\begin{pmatrix}98&99\\99&100\end{pmatrix}=\det\begin{pmatrix}98&99\\1&1\end{pmatrix}=-1,\det\begin{pmatrix}2&1&1\\1&2&1\\1&1&1\end{pmatrix}=\det\begin{pmatrix}1&0&0\\0&1&0\\1&1&1\end{pmatrix}=1$$

と求まる。(6) は、まず多重線型性で $2,3$ 列目の r と 3 列目の $\sin\theta$ をくくりだすとよい。

$$\det\begin{pmatrix}\sin\theta\cos\varphi&r\cos\theta\cos\varphi&-r\sin\theta\sin\varphi\\\sin\theta\sin\varphi&r\cos\theta\sin\varphi&r\sin\theta\cos\varphi\\\cos\theta&-r\sin\theta&0\end{pmatrix}=r^2\sin\theta\det\begin{pmatrix}\sin\theta\cos\varphi&\cos\theta\cos\varphi&-\sin\varphi\\\sin\theta\sin\varphi&\cos\theta\sin\varphi&\cos\varphi\\\cos\theta&-\sin\theta&0\end{pmatrix}$$

これを 3 列目について余因子展開すると

$$\begin{aligned}&=r^2\sin\theta\left\{(-\sin\varphi)\det\begin{pmatrix}\sin\theta\sin\varphi&\cos\theta\sin\varphi\\\cos\theta&-\sin\theta\end{pmatrix}-\cos\varphi\det\begin{pmatrix}\sin\theta\cos\varphi&\cos\theta\cos\varphi\\\cos\theta&-\sin\theta\end{pmatrix}\right\}\\&=r^2\sin\theta(\sin^2\varphi+\cos^2\varphi)=r^2\sin\theta\end{aligned}$$

となる。　■

■極座標変換の Jacobi 行列式　この問 6 (6) に出てきた行列は、いかにも意味ありげな行列です。実際この行列は、3 次元の極座標

$$\begin{cases} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{cases}$$

の **Jacobi 行列**という名前がついており、空間内での 3 重積分を極座標に変数変換して解く時に使います。今は名前だけ知っておいてください。

■問 9　(5) の答えは 360。先に掃き出しても何も考えずに余因子展開しても、手間は大差ないかもしれない。(2), (3) は Vandermonde 行列式の応用で片付ける。それ以外の計算は次の通り。
(1) $(1,1)$ 成分を要にして 1 列目を掃き出す。

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 1 \\ 3 & 2 & 1 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 3 \\ 0 & -1 & -2 \\ 0 & -4 & -8 \end{pmatrix} = 0$$

(4) 1 行目を 3 行目から引いてから、3 行目に関する余因子展開を使う。

$$\det\begin{pmatrix} -1 & 3 & -7 \\ 3 & -6 & 6 \\ -1 & 3 & -5 \end{pmatrix} = \det\begin{pmatrix} -1 & 3 & -7 \\ 3 & -6 & 6 \\ 0 & 0 & 2 \end{pmatrix} = (-1)^{3+3}\,2\det\begin{pmatrix} -1 & 3 \\ 3 & -6 \end{pmatrix} = -6$$

■

■問 10 の解答　(2) 1 列目を 2 列目から、3 列目を 4 列目から引き、交代性を使う。

$$\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix} = \det\begin{pmatrix} 1 & 1 & 3 & 1 \\ 5 & 1 & 7 & 1 \\ 9 & 1 & 11 & 1 \\ 13 & 1 & 15 & 1 \end{pmatrix} = 0$$

(3) $2,3,4$ 行目から 1 行目を引き、4 列目に関する余因子展開を使う。

$$\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 2 & 3 & 4 \\ 3 & 3 & 3 & 4 \\ 4 & 4 & 3 & 4 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 \end{pmatrix} = (-1)^{4+1}4\det\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 2 & 1 \end{pmatrix} = -4$$

(4) $(1,1)$ 成分を要に 1 列目を掃き出す。

$$\det\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 2 \\ 1 & 2 & 3 & 3 \\ 1 & 2 & 3 & 4 \end{pmatrix} = \det\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & 2 & 3 \end{pmatrix} = \det\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix} = \det\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix} = 1$$

(5) 1 行目と 4 行目、2 行目と 3 行目を入れ替えれば 1 と分かる。
(6) このまま余因子展開するのはしんどいので、$(2,2)$ 成分の 1 を要にして 2 行目を掃き出してみる。その後、多重線型性で 3 列目から 3 を、3 行目から 12 をくくりだすと

$$\det\begin{pmatrix} 6 & 12 & 7 & 9 \\ 15 & 1 & 14 & 4 \\ 10 & 8 & 11 & 5 \\ 3 & 13 & 2 & 16 \end{pmatrix} = \det\begin{pmatrix} -174 & 12 & -161 & -39 \\ 0 & 1 & 0 & 0 \\ -110 & 8 & -101 & -27 \\ -192 & 13 & -180 & -36 \end{pmatrix} = (-1)^{2+2}\det\begin{pmatrix} -174 & -161 & -39 \\ -110 & -101 & -27 \\ -192 & -180 & -36 \end{pmatrix}$$

$$
= -\det\begin{pmatrix} 174 & 161 & 39 \\ 110 & 101 & 27 \\ 192 & 180 & 36 \end{pmatrix} = -3 \times 12 \det\begin{pmatrix} 174 & 161 & 13 \\ 110 & 101 & 9 \\ 16 & 15 & 1 \end{pmatrix}
$$

となる。ここまで来れば 1 列目と 2 列目の差がぴったり 3 列目と一致することが見える。1 列目から 2 列目を引いて

$$
\det\begin{pmatrix} 174 & 161 & 13 \\ 110 & 101 & 9 \\ 16 & 15 & 1 \end{pmatrix} = \det\begin{pmatrix} 13 & 161 & 13 \\ 9 & 101 & 9 \\ 1 & 15 & 1 \end{pmatrix} = 0
$$

を得る。 ∎

■問 15 の解答　(3) から (6) は全て Vandermonde 行列式と類似の議論で片付く。

(1) 掃き出しと多重線型性を使う。

$$
\det\begin{pmatrix} 1 & 1 & 1 \\ 1 & a & a \\ 1 & a & b \end{pmatrix} = \det\begin{pmatrix} 1 & 1 & 1 \\ 0 & a-1 & a-1 \\ 0 & a-1 & b-1 \end{pmatrix} = (a-1)\det\begin{pmatrix} 1 & 1 \\ a-1 & b-1 \end{pmatrix} = (a-1)(b-a)
$$

(2) 掃き出すと三角行列の行列式に帰着される。

$$
\det\begin{pmatrix} a & b & c \\ -a & 1 & c \\ -a & -b & 1 \end{pmatrix} = \det\begin{pmatrix} a & b & c \\ 0 & 1+b & 2c \\ 0 & 0 & 1+c \end{pmatrix} = a(1+b)(1+c)
$$

∎

線形代数学② 第 3 回　　行列式の余因子展開

担当教員: 植野 義明 / TA: 穂坂 秀昭

講義日時: 2015 年 9 月 30 日 1 限

今回のお題は余因子です。もう主要な命題の証明は大体済んでしまったので計算方法を軽く確認してから、その後行列式の重要な応用例として、3 次元空間 $\mathbb{R}^3$ におけるベクトルの外積を説明します。また、おまけに期末試験の解説を載せました。

1　余因子と余因子展開

1.1　よくある間違い

採点をしていたら、目立つ間違いがいくつかありました。当てはまる人は修正してください。

- n 次正方行列 $A = (a_{ij})_{1 \leq i,j \leq n} \in \mathrm{Mat}_n(\mathbb{R})$ の (i,j) 余因子は

$$\Delta_{ij} := (-1)^{i+j} \det \begin{pmatrix} a_{11} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

  で定義されます。「A から i 行目と j 列目を取り去った行列」ではなく、その行列式に $(-1)^{i+j}$ をかけたものです。**行列のほうを余因子だと勘違いしていた人が少なからずいました**ので、今のうちに修正してください。
- 余因子を計算するときに、行列式の前にくっつく符号 $(-1)^{i+j}$ を忘れている人がいました。**符号まで込めて余因子です**ので、間違えないでください。
- 2 次以上の正方行列の**行列式は 1 個の数であって、行列とは別物です**。したがって

$$\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}, \quad \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$$

  は別のものです。何故かこれらを混同している人がいましたが、区別してください。
- 余因子行列の (i,j) 成分は Δ_{ji} です。(i,j) 成分に$\underline{(j,i)\ 余因子}$が来るので、間違えないでください。
- 逆行列の計算を間違える人がそこそこいました。**逆行列の計算が正しいかどうかは検算できる**ので、検算を怠らないでください。

1.2　余因子の計算

余因子の定義は上に記した通りです。計算例を見てみましょう。

■問 1 の解答

$$A_{12} = (-1)^{1+2} \det \begin{pmatrix} 1 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = -3, A_{23} = (-1)^{2+3} \det \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{pmatrix} = 4, A_{33} = (-1)^{3+3} \det \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = 6$$

■

■問 3 の解答　(1) $A_{12}=-3,\ A_{22}=1$
(2)

$$A_{12}=(-1)^{1+2}\begin{pmatrix}2&6\\3&9\end{pmatrix}=0, A_{22}=(-1)^{2+2}\begin{pmatrix}1&3\\3&9\end{pmatrix}=0$$

■

続いて、余因子展開の公式を復習します。n 次正方行列 $A=(a_{ij})_{1\le i,j\le n}\in \mathrm{Mat}_n(\mathbb{R})$ の (i,j) 余因子を Δ_{ij} と書くことにすると、任意の $1\le k,l\le n$ に対して

$$\det A=\sum_{j=1}^{n}a_{kj}\Delta_{kj}=\sum_{i=1}^{n}a_{il}\Delta_{il}$$

という公式が成り立つのでした。これらが余因子展開と呼ばれるものです。余因子展開することによって、高次の行列式の計算を低次の行列式の計算に帰着させることができます。

問題を解いて、例をみてみましょう。

■問 2 の解答

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

■

■問 4 の解答　(1) 12 (2) -30 (3) 0 ■

1.3　逆行列の計算

余因子の計算を使うと、逆行列の計算もできます。

n 次正方行列 A に対し、A の (i,j) 余因子を Δ_{ij} 書きます。このとき $\mathrm{adj}\,A:=(\Delta_{ji})_{1\le i,j\le n}$ という式で A の余因子行列を定めると、A が正則な場合には $A^{-1}=\mathrm{adj}\,A/\det A$ が成り立つのでした。

試しに 2 次正方行列でやってみましょう。

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

の余因子を計算すると、

$$\Delta_{11}=(-1)^{1+1}\det(d)=d,\qquad \Delta_{12}=(-1)^{1+2}\det(c)=-c,$$
$$\Delta_{21}=(-1)^{2+1}\det(b)=-b,\qquad \Delta_{11}=(-1)^{1+1}\det(a)=a$$

です[*1]。よって $\det A=ad-bc\neq 0$ なら

$$A^{-1}=\frac{1}{\det A}\begin{pmatrix}\Delta_{11}&\Delta_{21}\\\Delta_{12}&\Delta_{22}\end{pmatrix}=\frac{1}{ad-bc}\begin{pmatrix}d&-b\\-c&a\end{pmatrix}$$

です。これは僕たちが知っている逆行列の公式に他なりません。高次の正方行列でも同様に、余因子行列を使って逆行列を計算できます[*2]。問題を解いてみましょう。

[*1] 念の為補足しておくと 1 次正方行列 (a_{11}) の行列式は a_{11} そのものです。実際、1 個の数 1 を並び替えるやり方は 1 通りだけですので、$\mathfrak{S}_1=\{\mathrm{id}\}$ です。よって $\det(a_{11})=a_{1\,\mathrm{id}(1)}=a_{11}$ となります。

[*2] ただし次数が上がると行列式の計算は面倒です。「逆行列を求めたい行列 A と単位行列とを横に並べ、行基本変形によって A を単位行列にする」という方法の方が、楽なことが多いでしょう。

■問 6 の解答　答えのみ記す。

$$\begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad \frac{1}{abc}\begin{pmatrix} bc & c & 1 \\ 0 & ac & a \\ 0 & 0 & ab \end{pmatrix}$$

■

1.4　Cramer の公式

Cramer の公式は、連立 1 次方程式 $A\boldsymbol{x} = \boldsymbol{b}$ の係数行列が正則行列であるときに、その唯一の解を書き下す公式です。A の j 列目を $\boldsymbol{b}$ で置き換えた行列を A_j と書くと、

$$\boldsymbol{x} = A^{-1}\boldsymbol{b} = \frac{1}{\det A}{}^t\begin{pmatrix} \det A_1 & \det A_2 & \cdots & \det A_n \end{pmatrix}$$

が成り立ちます。証明自体は前回やってしまったので、今回は使い方を見てみましょう[*3]。

■問 7 の解答　(1) $x + y = 13$, $x - y = 5$ のとき

$$\det\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2$$

なので

$$x = \frac{1}{-2}\det\begin{pmatrix} 13 & 1 \\ 5 & -1 \end{pmatrix} = \frac{-18}{-2} = 9, y = \frac{1}{-2}\det\begin{pmatrix} 1 & 13 \\ 1 & 5 \end{pmatrix} = \frac{-8}{-2} = 4$$

(2) $x + y = 6$, $2x - y = 3$ のとき

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

なので

$$x = \frac{1}{-3}\det\begin{pmatrix} 6 & 1 \\ 3 & -1 \end{pmatrix} = 3, y = \frac{1}{-3}\det\begin{pmatrix} 1 & 6 \\ 2 & 3 \end{pmatrix} = 3$$

(3) $99x + 100y = 98$, $100x - 101y = 99$ のとき

$$\det\begin{pmatrix} 99 & 100 \\ 100 & -101 \end{pmatrix} = -19999$$

なので

$$x = \frac{1}{-19999}\det\begin{pmatrix} 98 & 100 \\ 99 & -101 \end{pmatrix} = \frac{19798}{19999}, y = \frac{1}{-19999}\det\begin{pmatrix} 99 & 98 \\ 100 & 99 \end{pmatrix} = -\frac{1}{19999}$$

■

■問 8 の解答　計算のやり方は問 7 と同じなので、答えのみ記す。(1) $(x, y, z) = (1, 2, 3)$ (2) $(x, y, z) = (1, 2, -1)$ (3) $(x, y, z) = (3, 2, 1)$ (4) $(x, y, z) = (14/3, -13/3, 29/3)$ ■

2　行列式の幾何学的応用

これまで行列式の色々な公式を示しはしたものの、行列式が持つ図形的な意味は述べていませんでした。そこで今回、行列式が「体積」という図形的意味を持つことを調べてみたいと思います。これによって、3 次元空間 $\mathbb{R}^3$ におけるベクトルの外積の意味も、的確に捉えられるようになります。

[*3] 逆行列の計算公式と同様、次数が高くなると Cramer の公式は実用的でなくなります。使う前に一度、掃き出し法とどっちが楽かを考えてみるべきでしょう。

2.1 行列式の特徴づけ

前回、行列式の重要な性質は「交代性と多重線型性である」と述べました。なぜこれらが重要なのかというと、**交代性と多重線型性を共に満たす写像は、**$\det$ **の定数倍しかない**ことが証明できるからです。

n 本の n 次元ベクトルに数を対応させる写像 $F\colon \mathbb{R}^n \times \mathbb{R}^n \times \cdots \mathbb{R}^n \to \mathbb{R}$ が交代性と多重線型性を満たすとしましょう。このとき多重線型性から

$$\begin{aligned}
F(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \ldots, \boldsymbol{a}_n) &= F\left(\sum_{i_1=1}^n a_{i_1 1}\boldsymbol{e}_{i_1}, \sum_{i_2=1}^n a_{i_2 2}\boldsymbol{e}_{i_2}, \sum_{i_3=1}^n a_{i_3 3}\boldsymbol{e}_{i_3}, \ldots, \sum_{i_n=1}^n a_{i_n n}\boldsymbol{e}_{i_n}\right) \\
&= \sum_{i_1=1}^n a_{i_1 1} F\left(\boldsymbol{e}_{i_1}, \sum_{i_2=1}^n a_{i_2 2}\boldsymbol{e}_{i_2}, \sum_{i_3=1}^n a_{i_3 3}\boldsymbol{e}_{i_3}, \ldots, \sum_{i_n=1}^n a_{i_n n}\boldsymbol{e}_{i_n}\right) \\
&= \sum_{i_1=1}^n a_{i_1 1} \sum_{i_2=1}^n a_{i_2 2} F\left(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \sum_{i_3=1}^n a_{i_3 3}\boldsymbol{e}_{i_3}, \ldots, \sum_{i_n=1}^n a_{i_n n}\boldsymbol{e}_{i_n}\right) \\
&= \cdots \\
&= \sum_{i_1=1}^n a_{i_1 1} \sum_{i_2=1}^n a_{i_2 2} \cdots \sum_{i_n=1}^n a_{i_n n} F(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \ldots, \boldsymbol{e}_{i_n}) \\
&= \sum_{i_1=1}^n \sum_{i_2=1}^n \cdots \sum_{i_n=1}^n a_{i_1 1} a_{i_2 2} \cdots a_{i_n n} F(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \ldots, \boldsymbol{e}_{i_n})
\end{aligned}$$

と変形できます。次に F の交代性より、$i_1, i_2, \ldots, i_n$ の中に同じものがあれば、$F(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \ldots, \boldsymbol{e}_{i_n}) = 0$ となります。これより、上の $\sum$ で足される項の多くは 0 で、生き残るのは

$$F(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3, \ldots, \boldsymbol{a}_n) = \sum_{\substack{1 \le i_1, \ldots, i_n \le n \\ i_k \ne i_l \ (k \ne l)}} a_{i_1 1} \cdots a_{i_n n} F(\boldsymbol{e}_{i_1}, \boldsymbol{e}_{i_2}, \cdots, \boldsymbol{e}_{i_n})$$

となります。そしてこの和は「$1, 2, \ldots, n$ の並べ替え $i_1, i_2, \ldots, i_n$ にわたる和」ですから、$\sigma(k) := i_k$ で定まる置換 $\sigma \in \mathfrak{S}_n$ を用いて

$$F(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \ldots, \boldsymbol{a}_n) = \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} F(\boldsymbol{e}_{\sigma(1)}, \boldsymbol{e}_{\sigma(n)}, \cdots, \boldsymbol{e}_{\sigma(n)})$$

と書き直せます。この $\boldsymbol{e}_{\sigma(1)}, \boldsymbol{e}_{\sigma(n)}, \cdots, \boldsymbol{e}_{\sigma(n)}$ は $\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ を並び替えたものです。そこで σ を互換の積で $\sigma = (p_1\ q_1)(p_2\ q_2)\cdots(p_l\ q_l)$ と書いておけば、$\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n$ という列の p_1 番目と q_1 番目、p_2 番目と q_2 番目、$\ldots$ を順番に並び替えれば $\boldsymbol{e}_{\sigma(1)}, \boldsymbol{e}_{\sigma(n)}, \cdots, \boldsymbol{e}_{\sigma(n)}$ になります。したがって F の交代性から

$$F(\boldsymbol{e}_{\sigma(1)}, \boldsymbol{e}_{\sigma(n)}, \cdots, \boldsymbol{e}_{\sigma(n)}) = \operatorname{sgn}(\sigma) F(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$$

となり、結局

$$F(\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3 \ldots, \boldsymbol{a}_n) = \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \operatorname{sgn}(\sigma) F(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = (\det A) F(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$$

を得ます。確かに交代性と多重線型性がある写像は、$\det$ の定数倍です。そしてその比例定数は $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ における写像の値と一致します。

■$\det$ **の乗法性の別証明**　これを使うと、実は $\det$ の乗法性が簡単に示します。$A \in \mathrm{Mat}_n(\mathbb{R})$ を n 次正方行列とします。このとき列ベクトルの並びで表された n 次正方行列 $B = (\boldsymbol{b}_1\ \boldsymbol{b}_2\ \cdots\ \boldsymbol{b}_n)$ に対し

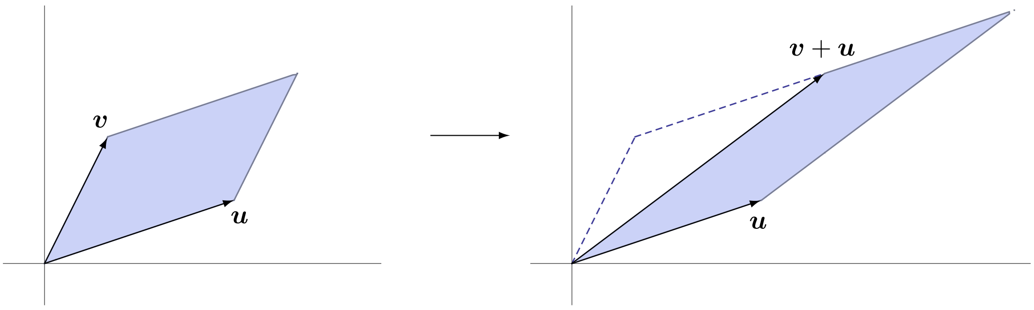
$$F(B) = F(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_n) := \det AB = \det A \begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_n \end{pmatrix} = \det \begin{pmatrix} A\boldsymbol{b}_1 & A\boldsymbol{b}_2 & \cdots & A\boldsymbol{b}_n \end{pmatrix}$$

という写像を考えてみます。行列の積が線型なことと det の多重線型性を合わせると、F が多重線型性を持つと分かります。また $\boldsymbol{b}_i$ と $\boldsymbol{b}_j$ を入れ替えると det の中で $A\boldsymbol{b}_i$ と $A\boldsymbol{b}_j$ が入れ替わるので、F は交代性も持ちます。故に $F(B) = (\det B)F(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n)$ と分かります。そして $F(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n) = \det A$ なので、$\det AB = \det A \det B$ が言えました。

2.2 行列式と面積・体積の関係

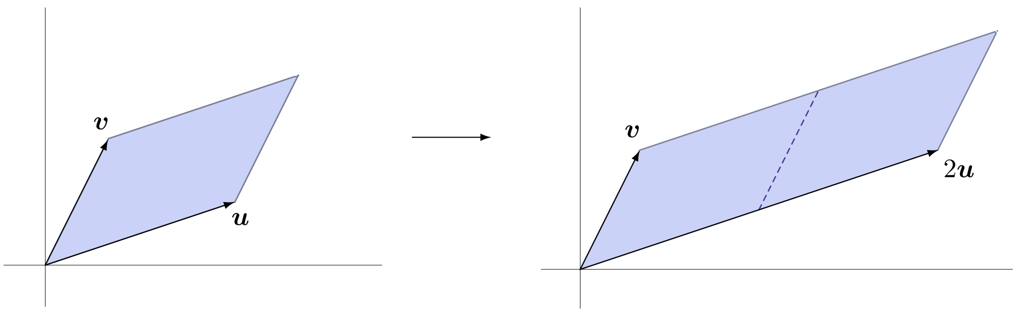
前々回、2 次の正方行列について**行列式の値は 2 本の列ベクトルが張る平行四辺形の符号付き面積と等しい**ことを示しました。同じことが 3 次元の場合にも成り立つことを調べましょう。

■基本変形に対する振る舞い　最初に行列式の性質が、面積とどう関係しているのかを見ておきましょう。行列式には、ある列を何倍かして別の列に足しても値が変化しないという性質がありました。たとえば $\det(\boldsymbol{u}\ \boldsymbol{v}) = \det(\boldsymbol{u}\ \boldsymbol{v}+\boldsymbol{u})$ が成り立ちます。この式を「2 本の列ベクトルの張る平行四辺形の面積」として見てみましょう。



絵で見れば、両方の面積が同じことは明らかですよね。平行四辺形の 1 つの辺を固定して、高さが変形しないように変形しているわけですから。α が入った $\det(\boldsymbol{u}\ \boldsymbol{v}+\alpha\boldsymbol{u})$ という式でも状況は同じです。α の値によって平行四辺形がどれだけひしゃげるかが変わるものの、底辺と高さが変わらないようことに変わりはありません。

■交代性・多重線型性の面積的意味　続いて、多重線型性を見てみましょう。まず $\det(\alpha\boldsymbol{u}\ \boldsymbol{v}) = \alpha\det(\boldsymbol{u}\ \boldsymbol{v})$ という性質がありました。これは $\alpha\boldsymbol{u}, \boldsymbol{v}$ の張る平行四辺形が $\boldsymbol{u}, \boldsymbol{v}$ の張る平行四辺形の α 倍の面積を持つという意味です。図にしてみれば、明らかでしょう。

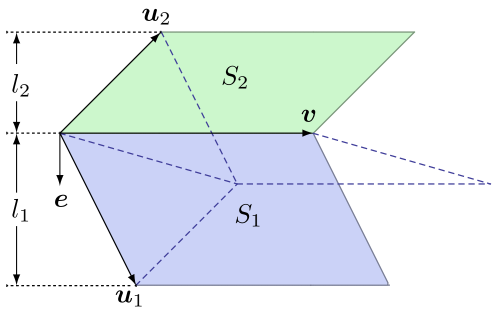


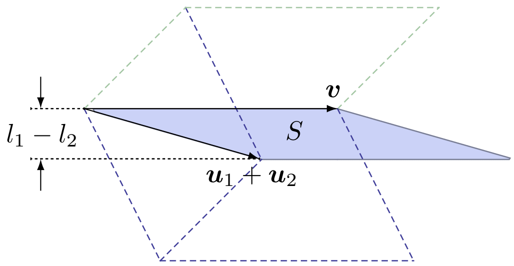
次に、$\det(\boldsymbol{u}_1+\boldsymbol{u}_2\ \boldsymbol{v}) = \det(\boldsymbol{u}_1\ \boldsymbol{v}) + \det(\boldsymbol{u}_2\ \boldsymbol{v})$ という式を見てみましょう。

- $\boldsymbol{u}_1$ と $\boldsymbol{v}$ が張る平行四辺形を S_1
- $\boldsymbol{u}_2$ と $\boldsymbol{v}$ が張る平行四辺形を S_2
- $\boldsymbol{u}_1+\boldsymbol{u}_2$ と $\boldsymbol{v}$ が張る平行四辺形を S

とおきます。そして $\boldsymbol{v}$ と垂直な方向の単位ベクトル $\boldsymbol{e}$ を取っておきます。このとき $\boldsymbol{v}$ から測った S_1, S_2 の高さをそれ

ぞれ l_1, l_2 とおくと、$l_1 = |\boldsymbol{u}_1 \cdot \boldsymbol{e}|, l_2 = |\boldsymbol{u}_2 \cdot \boldsymbol{e}|$ です。そして次の図で、S の高さ $l_1 - l_2$ は $(\boldsymbol{u}_1 + \boldsymbol{u}_2) \cdot \boldsymbol{e}$ です。

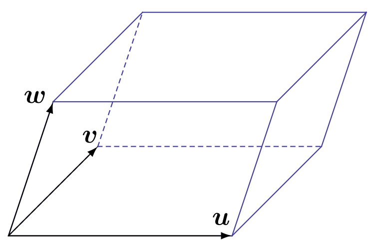




一般の場合でも同じです。$(\boldsymbol{u}_1 + \boldsymbol{u}_2) \cdot \boldsymbol{e} = \boldsymbol{u}_1 \cdot \boldsymbol{e} + \boldsymbol{u}_2 \cdot \boldsymbol{e}$ なので、向きまで込めて考えれば S_1, S_2 の高さを足したものが S の高さになります。これで S の面積が、S_1, S_2 の面積の和であることが示せました。det の言葉で言い換えると、$\det(\boldsymbol{u}_1 + \boldsymbol{u}_2\ \boldsymbol{v}) = \det(\boldsymbol{u}_1\ \boldsymbol{v}) + \det(\boldsymbol{u}_2\ \boldsymbol{v})$ に他なりません。また $\boldsymbol{u}, \boldsymbol{v}$ を入れ替えると、$\boldsymbol{u}$ から $\boldsymbol{v}$ へと回る向きが反転します。この事実を反映して $\det(\boldsymbol{u}\ \boldsymbol{v}) = -\det(\boldsymbol{v}\ \boldsymbol{u})$ という式が成り立ちます。

■体積の持つ交代性・多重線型性　今の話を、3 次元でも考えてみましょう。3 本のベクトル $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ を引数に取る関数 $\mathrm{vol}\colon \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}$ を次で定めます。

- 絶対値は、$\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ の張る平行 6 面体の体積
- 符号は、$(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ が右手系のとき $+1$、左手系のとき -1



この vol は、作り方から交代性を満たします。右手系をなすベクトルのうち 2 本を入れ替えると左手系になり、またその逆も成り立つからです。また、$\mathrm{vol}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ は $\boldsymbol{u}$ について線型です。実際 $\boldsymbol{v}, \boldsymbol{w}$ の両方に垂直な単位ベクトル $\boldsymbol{e}$ を取り、$\boldsymbol{v}$, $\boldsymbol{w}$ の張る平行四辺形の面積を A とすれば、$\mathrm{vol}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{u} \cdot \boldsymbol{e})A$ です。この式は $\boldsymbol{u}$ について線型です。そして交代性から $\mathrm{vol}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = -\mathrm{vol}(\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{w}) = -\mathrm{vol}(\boldsymbol{w}, \boldsymbol{v}, \boldsymbol{u})$ だから、$\boldsymbol{v}$ や $\boldsymbol{w}$ に対する線型性も従います。

これで $\mathrm{vol}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w})$ が交代性と多重線型性の両方を満たすことが言えました。さっき見た通り、交代多重線型性を持つ関数は $\mathrm{vol}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \det(\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w})\,\mathrm{vol}(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3)$ と書けます。そして $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ の張る平行 6 面体 (1 辺の長さ 1 の立方体) の体積は 1 ですから、$\mathrm{vol}(\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}) = \det(\boldsymbol{u}\ \boldsymbol{v}\ \boldsymbol{w})$ です。より一般に、n 次元であっても $\mathrm{vol} = \det$ という式が正しいことを示せます。**体積は交代性と多重線型性を満たすべき**という考察だけから、$\mathrm{vol} = \det$ が導かれるのです。

2.3　3 次元ベクトルの外積

3 次元空間 $\mathbb{R}^3$ の中の 3 本のベクトル $\boldsymbol{a} = {}^t(a_1, a_2, a_3), \boldsymbol{b} = {}^t(b_1, b_2, b_3), \boldsymbol{c} = {}^t(c_1, c_2, c_3)$ を並べた行列の行列式は、3 列目に関する余因子展開をすることで

$$\det\begin{pmatrix}\boldsymbol{a} & \boldsymbol{b} & \boldsymbol{c}\end{pmatrix} = \det\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} = c_1 \det\begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} - c_2 \det\begin{pmatrix} a_1 & b_1 \\ a_3 & b_3 \end{pmatrix} + c_3 \det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

$$= c_1 \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix} + c_2 \det \begin{pmatrix} a_3 & b_3 \\ a_1 & b_1 \end{pmatrix} + c_3 \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$$

となります。そこでベクトル $\boldsymbol{a} \times \boldsymbol{b}$ を

$$\boldsymbol{a} \times \boldsymbol{b} := {}^t\!\left(\det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix}, \det \begin{pmatrix} a_3 & b_3 \\ a_1 & b_1 \end{pmatrix}, \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \right)$$

と定めると、$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})$ が成り立ちます[*4]。このベクトル $\boldsymbol{a} \times \boldsymbol{b}$ のことを、$\boldsymbol{a}$ と $\boldsymbol{b}$ の**外積**といいます。行列式の交代性を使うと $\det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c}) = -\det(\boldsymbol{b}\ \boldsymbol{a}\ \boldsymbol{c})$ となるので、$\boldsymbol{a} \times \boldsymbol{b} = -\boldsymbol{b} \times \boldsymbol{a}$ が分かります。**外積は順番をひっくり返せない**ので、注意してください。

■外積の特徴づけ　$\mathbb{R}^3$ のベクトルは向きと大きさで決定されます。そこで外積の向きと大きさを調べてみましょう。

過去に一度「$\boldsymbol{a}$ と $\boldsymbol{b}$ の両方に垂直なベクトル」を作る方法として $\boldsymbol{a} \times \boldsymbol{b}$ を紹介したことがありますが、むしろ外積の定義で本質的なのはスカラー三重積の公式 $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})$ です。行列式の交代性を使うと、$\boldsymbol{c} = \boldsymbol{a}$, $\boldsymbol{c} = \boldsymbol{b}$ とおいたとき、ただちに

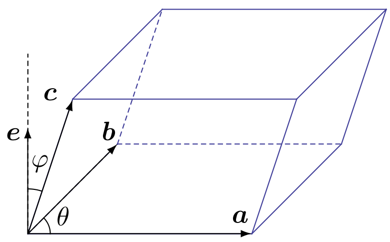
$$(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{a} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{a}) = 0, (\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{b} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{b}) = 0$$

が分かります。これが $\boldsymbol{a} \times \boldsymbol{b}$ の計算で、$\boldsymbol{a}$ と $\boldsymbol{b}$ の両方と垂直なベクトルを作れる理由です。

次に大きさ $|\boldsymbol{a} \times \boldsymbol{b}|$ を考えてみます。$\boldsymbol{a}$ と $\boldsymbol{b}$ が平行なら $\boldsymbol{a} \times \boldsymbol{b} = \boldsymbol{0}$ なので、そうではないと仮定します。$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ が 1 次独立になるような $\boldsymbol{c} \in \mathbb{R}^3$ を取ります。このとき $|\boldsymbol{c}| \neq 0$ で、$|\det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})| = |(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}|$ は $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ の張る平行 6 面体の体積でした。かたや $\boldsymbol{a}, \boldsymbol{b}$ の張る面を S、その面の法線方向と $\boldsymbol{c}$ のなす角を $[0, \pi/2)$ の範囲で測ったものを φ とすると、S から測った平行 6 面体の高さは $|\boldsymbol{c}| \cos\varphi$ です。$\boldsymbol{c}$ の取り方から $\cos\varphi \neq 0$ にもなります。よって

$$|\boldsymbol{a} \times \boldsymbol{b}| = \frac{|(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c}|}{|\boldsymbol{c}| \cos\varphi} = \frac{|\det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})|}{|\boldsymbol{c}| \cos\varphi} = \frac{|\operatorname{vol}(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})|}{|\boldsymbol{c}| \cos\varphi}$$

が従います。これは体積を高さで割ったものなので、外積の大きさは $\boldsymbol{a}, \boldsymbol{b}$ が張る平行四辺形の面積に一致します。$\boldsymbol{a}$ と $\boldsymbol{b}$ のなす角を θ とすれば、$|\boldsymbol{a} \times \boldsymbol{b}| = |\boldsymbol{a}||\boldsymbol{b}| \sin\theta$ です[*5]。



最後に行列式の多重線型性を使うと、任意の $\boldsymbol{c} \in \mathbb{R}^3$ に対し

$$\det\big(\boldsymbol{a} + \boldsymbol{a}'\ \boldsymbol{b}\ \boldsymbol{c}\big) = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c}) + \det(\boldsymbol{a}'\ \boldsymbol{b}\ \boldsymbol{c})$$
$$\det(\alpha\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c}) = \alpha \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})$$

となるから、$\big((\boldsymbol{a} + \boldsymbol{a}') \times \boldsymbol{b}\big) \cdot \boldsymbol{c} = (\boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a}' \times \boldsymbol{b}) \cdot \boldsymbol{c}$ および $\big((\alpha\boldsymbol{a}) \times \boldsymbol{b}\big) \cdot \boldsymbol{c} = \big(\alpha(\boldsymbol{a} \times \boldsymbol{b})\big) \cdot \boldsymbol{c}$ が成り立ちます。特に $\boldsymbol{c}$ として標準基底をなす $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ を取れば、$(\boldsymbol{a} + \boldsymbol{a}') \times \boldsymbol{b}$ と $\boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a}' \times \boldsymbol{b}$、$(\alpha\boldsymbol{a}) \times \boldsymbol{b}$ と $\alpha(\boldsymbol{a} \times \boldsymbol{b})$ のそれぞれについ

---

*4 $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})$ という式は、よく**スカラー三重積の公式**と呼ばれます。たとえば電磁気学やベクトル解析の授業などでは、必ず登場すると思います。外積を成分で定義すると $(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{c} = \det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c})$ が証明できるので「公式」という名前がついています。

*5 $\boldsymbol{a} \cdot \boldsymbol{b} = |a||b| \cos\theta$ と $\sin\theta = \sqrt{1 - \cos^2\theta}$ を使ってゴリゴリ計算しても $|\boldsymbol{a} \times \boldsymbol{b}|$ は求まります。ですが「体積は行列式である」という事実とスカラー三重積の公式を組み合わせれば、大した計算をせずとも綺麗に答えが求まるので、こちらの方法で説明してみました。

て、x, y, z 成分が一致することが確認できます。これより $(\boldsymbol{a}+\boldsymbol{a}')\times\boldsymbol{b}=\boldsymbol{a}\times\boldsymbol{b}+\boldsymbol{a}'\times\boldsymbol{b}$ と $(\alpha\boldsymbol{a})\times\boldsymbol{b}=\alpha\boldsymbol{a}\times\boldsymbol{b}$ が従います。また $\boldsymbol{a}, \boldsymbol{b}$ が 1 次独立なら $|\boldsymbol{a}|, |\boldsymbol{b}| \neq 0$ かつ $\theta \neq 0, \pi$ なので、$\boldsymbol{a}\times\boldsymbol{b}\neq\boldsymbol{0}$ だから

$$\det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{a}\times\boldsymbol{b}) = (\boldsymbol{a}\times\boldsymbol{b})\cdot(\boldsymbol{a}\times\boldsymbol{b}) = |\boldsymbol{a}\times\boldsymbol{b}|^2 = |\boldsymbol{a}|^2|\boldsymbol{b}|^2\sin^2\theta > 0$$

が従います。よって $\boldsymbol{a}, \boldsymbol{b}$ が 1 次独立なとき、$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}\times\boldsymbol{b}$ をこの順に並べると右手系[*6]をなします。これで $\boldsymbol{a}\times\boldsymbol{b}$ の向きが完全に決定されました。

かくして外積 $\boldsymbol{a}, \boldsymbol{b}$ は

- 大きさが $\boldsymbol{a}, \boldsymbol{b}$ の張る平行四辺形の面積
- $\boldsymbol{a}, \boldsymbol{b}$ の両方と垂直で、$\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{a}\times\boldsymbol{b}$ が右手系となる向き

という条件で特徴づけられるベクトルだということも分かりました。

■n 次元空間 $\mathbb{R}^n$ における外積　今の話を一般化すれば、$n \geq 4$ のときも $\mathbb{R}^n$ においてもベクトルの外積を考えることができます。ただし 3 次元での外積の作り方を見れば分かるように、外積は

$$\det(\boldsymbol{a}_1\ \boldsymbol{a}_2\ \cdots\ \boldsymbol{a}_{n-1}\ \boldsymbol{b}) = (\boldsymbol{a}_1\times\boldsymbol{a}_2\times\cdots\times\boldsymbol{a}_{n-1})\cdot\boldsymbol{b}$$

という式で定義されるべきものです。$\mathbb{R}^3$ の 2 本のベクトルに対して外積が定義できたのは $\dim\mathbb{R}^3 - 1 = 2$ だからであって、n 次元の場合は $n-1$ 本のベクトルの組に対して外積が計算されます。

3　おまけ: S2 ターム期末試験の解説

「S2 タームの期末試験の解説がほしい」というリクエストをいただいたので、解答と解説を作りました。復習に役立ててください。なお配点がどうなっているのかはそもそも知らないので、書いていません。

3.1　問 1

(2) は中々難しいと思います。$|z|=1$ となる複素数は $\cos\theta + i\sin\theta$ と表せますが、θ をパラメータに取った状態では $\cos\theta$ と $\sin\theta$ の両方が同時に有理数になるタイミングは見えてきません。ここで三角函数の入った積分を置換積分で解くときに、たまに使う公式を思い出します。$t := \tan(\theta/2)$ とおくと

$$\cos\theta = \frac{1-t^2}{1+t^2}, \quad \sin\theta = \frac{2t}{1+t^2}$$

が成り立つのを、覚えていますか？ この式は大変都合の良いことに、t の有理式です。したがって t が有理数であれば、$\cos\theta$ と $\sin\theta$ が両方とも有理数になることが分かります。そうすれば後は、t が有理数になるような θ が無限個存在することを示すだけです。

■解答　(1) $(a+bi)\overline{(a+bi)} = a^2+b^2$ を使うと (a) 2, (b) 5, (c) 13 と分かる。
(2) $(1-t^2)^2+(2t)^2=(1+t^2)^2$ の両辺を $1+t^2$ で割ると

$$\left(\frac{1-t^2}{1+t^2}\right)^2 + \left(\frac{2t}{1+t^2}\right)^2 = 1$$

[*6] この部分では「$\det > 0$ だから右手系」という書き方をしていますが、実は逆です。1 次独立な 3 本のベクトル $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ が右手系であることを、$\det(\boldsymbol{a}\ \boldsymbol{b}\ \boldsymbol{c}) > 0$ で定義します。こうすると「右手の法則にしたがう」という曖昧な言い方を回避してきちんと定義ができ、さらに「右手系と左手系がきちんと区別されること」「右手系と左手系以外の○○系がないこと」などが示せます。将来「内積」の話をする際、一緒に「空間の向き」のことを紹介します。

となる。よって実数 t に対し

$$z(t) := \frac{1-t^2}{1+t^2} + \frac{2t}{1+t^2}i$$

と定めると、t の値が何であっても $|z(t)| = 1$ である。式の形から、t が有理数なら $z(t)$ の実部、虚部は共に有理数である。そして

$$\operatorname{Re} z(t) = \frac{2}{1+t^2} - 1$$

は t について単調減少で、$\operatorname{Re} z(0) = 1$, $\operatorname{Re} z(1) = 0$ である。したがって閉区間 $[0,1]$ の点 t に $z(t)$ を対応させる写像は単射である。これと閉区間 $[0,1]$ に有理数が無数に存在することから、求める結果を得る。 ∎

3.2 問 2

この問題については、特に注釈は要らないでしょう。地道に計算するだけです。

■解答 (1) $\overrightarrow{\mathrm{AB}} = {}^t(\sqrt{2}-1, \sqrt{3}-\sqrt{2}, 1-\sqrt{3})$, $\overrightarrow{\mathrm{AC}} = {}^t(\sqrt{3}-1, 1-\sqrt{2}, \sqrt{2}-\sqrt{3})$ なので

$$\begin{aligned}\cos\angle\mathrm{BAC} = \frac{\overrightarrow{\mathrm{AB}}\cdot\overrightarrow{\mathrm{AC}}}{|\overrightarrow{\mathrm{AB}}||\overrightarrow{\mathrm{AC}}|} &= \frac{(\sqrt{2}-1)(\sqrt{3}-1)+(\sqrt{3}-\sqrt{2})(1-\sqrt{2})+(1-\sqrt{3})(\sqrt{2}-\sqrt{3})}{(\sqrt{2}-1)^2+(\sqrt{3}-\sqrt{2})^2+(1-\sqrt{3})^2} \\ &= \frac{(\sqrt{2}-1)^2+(1-\sqrt{3})(\sqrt{2}-\sqrt{3})}{12-2(\sqrt{2}+\sqrt{3}+\sqrt{6})} = \frac{6-\sqrt{2}-\sqrt{3}-\sqrt{6}}{12-2(\sqrt{2}+\sqrt{3}+\sqrt{6})} \\ &= \frac{1}{2}\end{aligned}$$

となる。よって $\angle\mathrm{BAC} = \pi/3$ である。

(2) $\alpha\boldsymbol{a} + \beta(\boldsymbol{a}+\boldsymbol{b}) = \boldsymbol{0}$ とおく。このとき $(\alpha+\beta)\boldsymbol{a} + \beta\boldsymbol{b} = \boldsymbol{0}$ なので、$\boldsymbol{a}, \boldsymbol{b}$ の 1 次独立性から $\alpha+\beta = 0, \beta = 0$ が従う。これを解くと $\alpha = \beta = 0$ となる。よって $\boldsymbol{a}, \boldsymbol{a}+\boldsymbol{b}$ は 1 次独立である。

(3) $2\boldsymbol{a} + t(\boldsymbol{b}-2\boldsymbol{a}) = \boldsymbol{a} + s(\boldsymbol{a}+3\boldsymbol{b})$ のとき、$(1-2t-s)\boldsymbol{a} + (t-3s)\boldsymbol{b} = \boldsymbol{0}$ となる。よって $\boldsymbol{a}, \boldsymbol{b}$ の 1 次独立性から $2t+s = 1$, $t-3s = 0$ という連立方程式を得る。これを解いて $s = 1/7, t = 3/7$ を得る。 ∎

3.3 問 3

ただの計算問題です。下三角同士の積が再び下三角行列になることを覚えていると、(2) では対角線より上の計算を端折れます。

■解答

(1)

$$\frac{1}{2}\begin{pmatrix}1&4&7\\2&5&8\\3&6&9\end{pmatrix} - \frac{1}{2}\begin{pmatrix}1&2&3\\4&5&6\\7&8&9\end{pmatrix} = \begin{pmatrix}0&1&2\\-1&0&1\\-2&-1&0\end{pmatrix}$$

(2)

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

(3)

$$\begin{pmatrix}\sqrt{2}&-1&1\\\sqrt{2}&1&-1\\0&\sqrt{2}&\sqrt{2}\end{pmatrix}\begin{pmatrix}\sqrt{2}&\sqrt{2}&0\\-1&1&\sqrt{2}\\1&-1&\sqrt{2}\end{pmatrix} = \begin{pmatrix}4&0&0\\0&4&0\\0&0&4\end{pmatrix}$$

∎

3.4　問 4

拡大係数行列に未知数 a が入り込んでいるのがややこしいところです。a の値によって、係数行列や拡大係数行列の rank が変わることが予想されます。

なるべく手間をかけたくないので、式中に a を残したまま、できる限り行基本変形を進めてみましょう。その際 **a の入った式で割り算をしないこと**が大事です。たとえば $a-1$ が 0 だろうがそれ以外だろうが、「ある行の $a-1$ 倍を別の行に足す」という操作はできます。しかし $a-1=0$ のときは「$a-1$ で割る」という操作をすると破綻が生じてしまいます。

■解答　まず、拡大係数行列を次のように行基本変形する

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ a & 1 & 1 & 1 & 1 \\ 1 & a & 3-a & 1 & 3 \\ 2 & 2 & a & 2 & 0 \end{array}\right) \xrightarrow[\text{1 列目を掃き出す}]{(1,1)\text{ 成分を要に}} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1-a & 1-a & 1-a & 1-a \\ 0 & a-1 & 2-a & 0 & 2 \\ 0 & 0 & a-2 & 0 & -2 \end{array}\right)$$

$$\xrightarrow[\text{加える}]{\text{2 行目を 3 行目に}} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1-a & 1-a & 1-a & 1-a \\ 0 & 0 & 3-2a & 1-a & 3-a \\ 0 & 0 & a-2 & 0 & -2 \end{array}\right)$$

$$\xrightarrow[\text{2 倍を加える}]{\text{3 行目に 4 行目の}} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1-a & 1-a & 1-a & 1-a \\ 0 & 0 & -1 & 1-a & -1-a \\ 0 & 0 & a-2 & 0 & -2 \end{array}\right)$$

$$\xrightarrow[a-2\text{ 倍を加える}]{\text{4 行目に 3 行目の}} \left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1-a & 1-a & 1-a & 1-a \\ 0 & 0 & -1 & 1-a & -1-a \\ 0 & 0 & 0 & (1-a)(a-2) & -a(a-1) \end{array}\right)$$

この式より、係数行列の rank が $a=1,2$ とそれ以外で変化しそうだと分かる。

- $a=1$ のとき、拡大係数行列は
$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$
となる。よって係数行列と拡大係数行列の rank は共に 2 である。
- $a=2$ のとき、拡大係数行列は
$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & -1 & -3 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right)$$
となる。$(2,2)$ 成分、$(3,3)$ 成分と $(4,5)$ 成分を要にして 2 列目、3 列目と 5 列目とをそれぞれ掃き出してから、符号を調整すると
$$\left(\begin{array}{cccc|c} 1 & 0 & 0 & * & 0 \\ 0 & 1 & 0 & * & 0 \\ 0 & 0 & 1 & * & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right)$$
の形に変形できる。よって係数行列の rank は 3 で、拡大係数行列の rank は 4 である。

- それ以外の場合 $a-1, a-2 \neq 0$ なので、2 行目を $1-a$ で、4 行目を $(1-a)(a-2)$ で割ってよい。

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & a-1 & 1+a \\ 0 & 0 & 0 & 1 & a/(2-a) \end{array}\right)$$

この後 $(1,1)$ 成分から $(4,4)$ 成分までを順番に要として 1 列目から 4 列目を掃き出せば

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & a-1 & 1+a \\ 0 & 0 & 0 & 1 & a/(a-2) \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2-a & -a \\ 0 & 0 & 1 & a-1 & 1+a \\ 0 & 0 & 0 & 1 & a/(a-2) \end{array}\right) \to \left(\begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -2/(a-2) \\ 0 & 0 & 0 & 1 & a/(a-2) \end{array}\right)$$

となる。よって係数行列の rank は 4 である。拡大係数行列の rank は係数行列の rank 以上で、かつ行数以下である。よって拡大係数行列の rank もまた 4 になる。

(2), (3) 行列が解を持つための必要十分条件は、係数行列と拡大係数行列の rank が一致することであった。したがって解が存在するための条件は $a \neq 2$ である。また解がただ 1 つ存在するための必要十分条件は、係数行列が正方行列で、かつ係数行列の rank が行列のサイズと一致することである。よって $a \neq 1, 2$ のとき、解はただ 1 つに定まる。

$a \neq 1, 2$ のとき、解は上で計算した通り

$$x = y = 0, z = -\frac{2}{a-2}, w = \frac{a}{a-2}$$

である。また $a = 1$ のときは、行基本変形後の拡大係数行列を見れば、解は x, y をパラメータとして

$$z = 2, w = -1 - x - y$$

と書ける。 ■

3.5 問 5

表現行列の意味を確かめる問題です。落ち着いて定義通りに計算すれば、解けるはずです。

■解答 (1)

$$\begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}$$

の逆行列を求めるには、右側に単位行列を並べて、左側の行列が単位行列になるよう行基本変形を行えば良い。

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -2 & 1 & 2 \\ 0 & 0 & 1 & 2 & -1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & -1 \\ 0 & 1 & 0 & 2 & -1 & 2 \\ 0 & 0 & 1 & 2 & -1 & 1 \end{array}\right)$$

(2) 基底 $(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3)$ に関する T の表現行列を A と書くと、$T(\boldsymbol{u}_1) = \boldsymbol{u}_1 + \boldsymbol{u}_2$, $T(\boldsymbol{u}_2) = \boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3$ と $T(\boldsymbol{u}_3) = \boldsymbol{u}_1 - \boldsymbol{u}_3$ より

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

である。

(3) A を行基本変形すると

$$\begin{pmatrix}1 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & -1\end{pmatrix} \to \begin{pmatrix}1 & 1 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & -1\end{pmatrix} \to \begin{pmatrix}1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 0\end{pmatrix}$$

となる。よって $A\boldsymbol{x} = 0$ の解は

$$\boldsymbol{x} = \begin{pmatrix}-2z \\ z \\ z\end{pmatrix} = z\begin{pmatrix}-2 \\ 1 \\ 1\end{pmatrix}$$

と書ける。T の言葉で書きなおすと $T(-2\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3) = 0$ だから、$\mathrm{Ker}\, T$ の $\mathrm{Im}\, T$ の基底として $-2\boldsymbol{u}_1 + \boldsymbol{u}_2 + \boldsymbol{u}_3$ が取れる。

(4) $\dim \mathrm{Im}\, T = 3 - \dim \mathrm{Ker}\, T = 2$ である。よって A の列ベクトルのうち一次独立な 2 本を取り、それに対応する $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ の 1 次結合を作れば、$\mathrm{Im}\, T$ の基底が得られる。

$$\det \begin{pmatrix}1 & 1 \\ 1 & 2\end{pmatrix} = 2 - 1 = 1 \neq 0$$

より、最初の 2 本が 1 次独立である。よって基底として $\boldsymbol{u}_1 + \boldsymbol{u}_2$, $\boldsymbol{u}_1 + 2\boldsymbol{u}_2 + \boldsymbol{u}_3$ が取れる。 ∎