

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7月14日

目次

S1 ターム 数理科学基礎 (線形代数学)	1
第 1 回 複素数と代数学の基本定理	3
1 はじめに	3
2 複素(数)平面の幾何	4
3 多項式の性質	7
第 2 回 種々の函数	13
1 全体的な講評	13
2 逆函数について	13
3 逆三角函数と円周率の近似公式	14
4 双曲線函数と逆双曲線函数	18
第 3 回 座標空間と数ベクトル	21
1 情報の調べ方	21
2 空間内における直線と平面の取り扱い	25
3 ベクトルの 1 次独立性と行列式	28
4 残りの問題	30
第 4 回 2 变数函数のグラフ	31
1 講評	31
2 2 变数函数のグラフの切断	31
3 座標変換とグラフの移動	34
4 曲面上の運動と接ベクトル	35
5 解答など	36
第 5 回 行列とその演算	37
1 総和記号 の使い方	37
2 行列とその演算	42
3 有名な公式たち	43
第 6 回 線型写像と行列	47
1 問題訂正のお詫びと雑談	47
2 数ベクトルに対する線型写像	48
3 線型空間	51

S2 ターム 線形代数学	55
第 1 回 ベクトル空間と線型写像	57
1 はじめに	57
2 線型空間と線型写像	57
3 重要な例	61
4 線型代数のさらなる一般論	68
第 2 回 ベクトル空間と線型写像の核と像	73
1 数学における議論のしかたについて	73
2 部分空間について	74
3 線型写像の核と像	77
第 3 回 掃き出し法	83
1 過去のプリントについて	83
2 問題の答えと計算のコツ	83
3 掃き出し法	85
4 行列の核、像と階数	91
第 4 回 線型空間の基底	95
1 証明のしかた	95
2 基底と次元	95
3 有名な線型空間における基底の例	102
4 次元と線型写像	104

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っています。そこで数学が専門の人はどうするかというと、そういう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できます。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ン

*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

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

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 に対応させることができます。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。

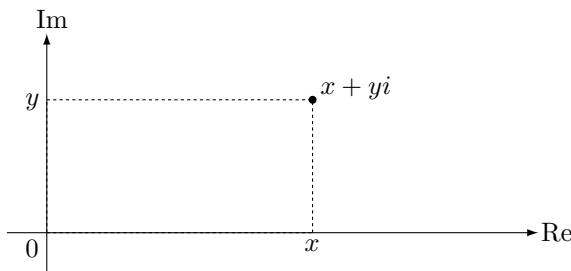


図 1.1 複素平面における数と点の対応

^{*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”的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いかったので、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 $\mathbb{R}^2 \rightarrow \mathbb{C}$ 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>

少し大きめに言うと、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に 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 $|z|$ 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 \bar{z} と書くと、 $\bar{x + yi} = x - yi$ です。この \bar{z} を z の共役と言います^{*8}。共役は $\bar{z+w} = \bar{z} + \bar{w}$, $\bar{zw} = \bar{z}\bar{w}$ 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、 $z\bar{z} = (x + yi)(x - yi) = x^2 + y^2 = |z|^2$ 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) $z_1 = x_1 + y_1i$, $z_2 = x_2 + y_2i$ 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

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

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

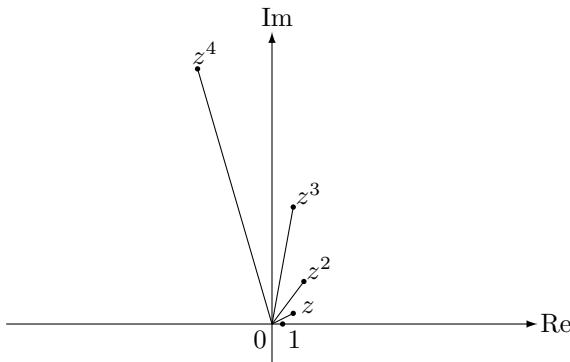
$$z^2 = 3 + 4i, \quad z^3 = 2 + 11i, \quad z^4 = -7 + 24i, \quad z^5 = -38 + 41i, \quad z^6 = -117 + 44i$$

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 の偏角といい、 $\theta = \arg z$ 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。

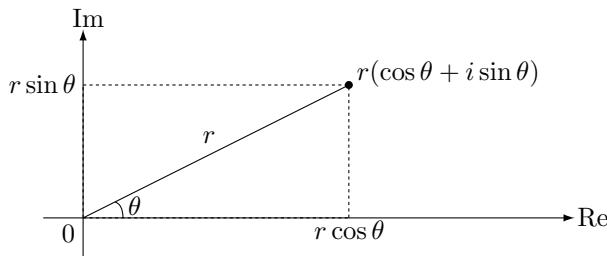


図 1.3 複素数の極形式表示

さて、極形式で表された 2 つの複素数を掛け算すると

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

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、

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

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 $0 \leq t \leq 1$ の範囲で $2t/(1+t^2)$ は単調増加する。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、 $m \geq n$ 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 $f(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_0, g(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_0$ 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、 $z \neq 0$ である。よって $f(z)/z^2 = 0$ である。また $f(z)/z^2 = z^2 + z + 1 + z^{-1} + z^{-2} = (z + z^{-1})^2 + (z + z^{-1}) - 1 = t^2 + t - 1 = 0$ 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。 $t = \frac{-1 \pm \sqrt{5}}{2}$ 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) $x^n - a^n = (x - a)(x^{n-1} + ax^{n-2} + a^2x^{n-3} + \dots + a^{n-1})$

(2) もし $x^n - a^n$ が $x + a$ で割り切ることは、 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 $0 = (-a)^n - a^n = ((-1)^n - 1)a^n$ 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 $(-a)^n + a^n = ((-1)^n + 1)a^n$ 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 $x^n - a^n = (x - a)(x^{n-1} + ax^{n-2} + \dots + a^{n-1})$ 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 $e^{2k\pi i/n}$ ($k = 0, 1, \dots, n-1$) 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 $z^{n-1} + z^{n-2} + \dots + 1$ 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}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 $z^{n-1} + z^{n-2} + \dots + 1$ 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 $f(x_1, \dots, x_n)$ 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 $300 + 250i$ 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”と呼ばれますが、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、「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、その中には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 $f(z) = a_0 + a_1z + \dots + a_nz^n$ とおく。このとき $a_0, a_1, \dots, a_n \in \mathbb{R}$ である。よって $\alpha \in \mathbb{C}$ が $f(z)$ の解であるとき、 $f(\bar{\alpha}) = 0$ の共役を取ると

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

となる。よって $\bar{\alpha}$ 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」とこと「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」とここ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、 $g(z) = a_0 + a_k z^k + a_{k+1} z^{k+1} + \dots + a_n z^n$ 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 $z_0 := r e^{i(\theta_0 - \theta_k + \pi)/k}$ と定めると

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

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

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

となる。この式で

$$M := \max \left\{ \left| \frac{a_{k+1}}{a_k} \right|, \left| \frac{a_{k+2}}{a_{k+1}} z_0 \right|, \dots, \left| \frac{a_n}{a_k} z_0^{n-k-1} \right| \right\}$$

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

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

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

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

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

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

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

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

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

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

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、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。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 $\exp: \mathbb{R} \rightarrow \mathbb{R}$ 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 $\mathbb{R}_{>0}$ に取り換えると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 $\mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ 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 $\arg: \mathbb{C}^\times \rightarrow [0, 2\pi)$ 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「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。

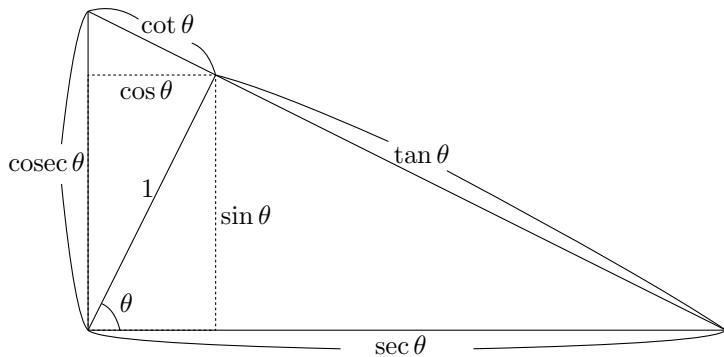
3 逆三角函数と円周率の近似公式

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。

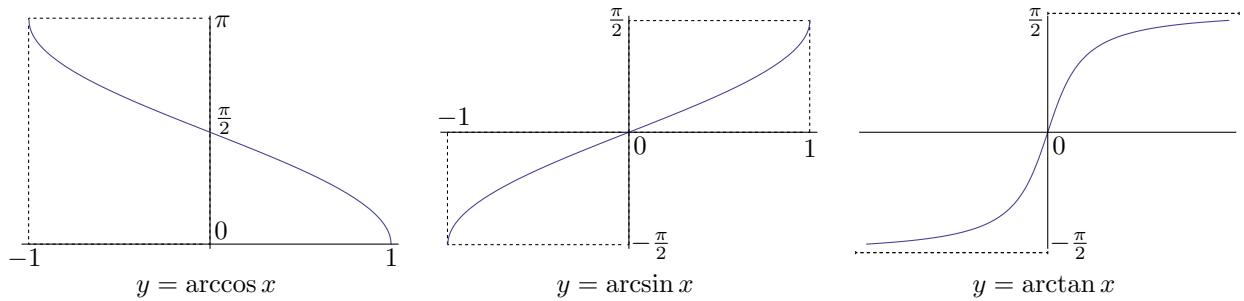
$\sec \theta := 1/\cos \theta$, $\cosec \theta := 1/\sin \theta$ と定めます。これを図に描くと次のようになります。



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「单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、 \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

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

を得る。この式で $u = 1/2, v = 1/3$ とおくと

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

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

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

を得る^{*3}。

問 4 の解答 計算自体は簡単である。

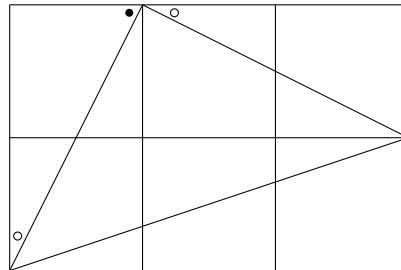
$$(1) (3+i)(7+i) = 20 + 10i, \quad (2) (2+i)(3+i) = 5 + 5i, \quad (3) (5+i)^4/(239+i) = 2 + 2i$$

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：

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

この公式は、形式的には

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

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

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

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 $-1/99 = 0.010101\dots$ 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「ちょっと計算しただけで、上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

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

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 $\pi = 3.14\dots$ が求まります。さっさと断然楽ですね。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 行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 術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での、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『橿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。

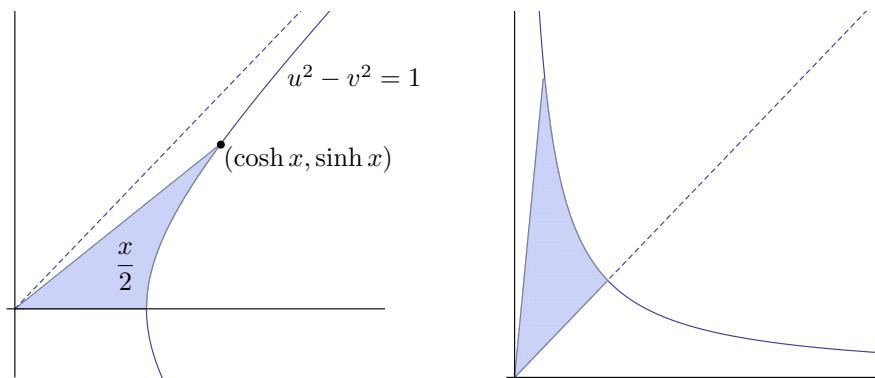
$$\cosh x := \frac{e^x + e^{-x}}{2}, \quad \sinh x := \frac{e^x - e^{-x}}{2}, \quad \tanh x := \frac{\sinh x}{\cosh x}$$

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}。



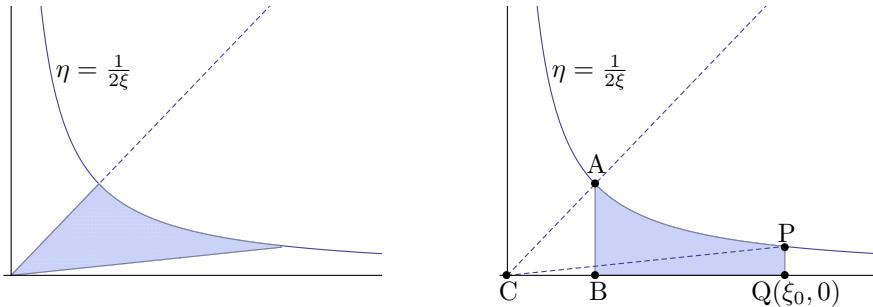
縦より横に積分した方が見やすいので、面積を求めるべき部分を直線 $\eta = \xi$ について折り返す。すると、次の二つの図で示された面積は等しいことが分かる。 $AB \cdot BC = PQ \cdot CQ = \frac{1}{2}$ が成り立ち、 $\triangle ABC$ と $\triangle CQP$ の面積が等しくなるからである。

^{*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

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

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° 度回転す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

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

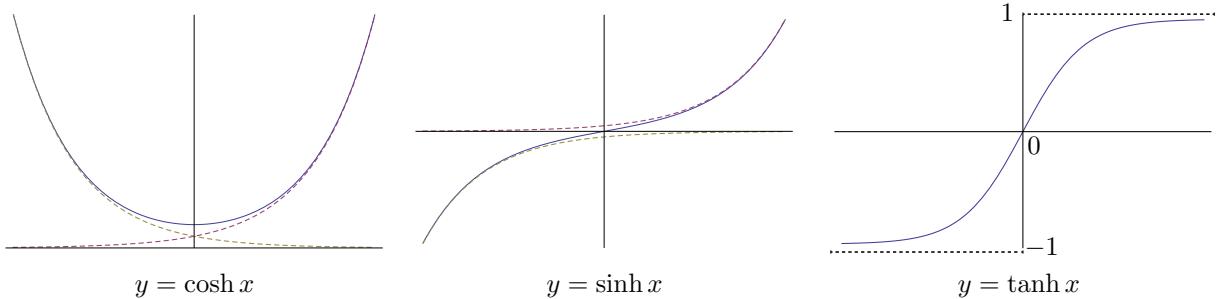
と分かる^{*12}。この実部が ξ_0 なので、面積の計算と合わせて $x = \log(\sqrt{2}\xi_0) = \log(\cosh x + \sinh x)$ が従う。つまり $\cosh x + \sinh x = e^x$ である。

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 $\sinh x$ のグラフに点線で書き込まれているのは、函数 $e^x/2$ と $\pm e^{-x}/2$ です。



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あるので、 $x \rightarrow \pm\infty$ 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

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

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 $\text{arccosh } y = x = \log(y \pm \sqrt{y^2 - 1})$ である。±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 $x = \text{arcsinh } y = \log(y + \sqrt{y^2 + 1})$ 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。

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

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

^{*13} 一般の複素数 $z = a + i\theta$ に対しては、 $e^z := e^a e^{i\theta}$ 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、出典を書かずして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.ij4u.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 ‘ό ον’ が語源。更に遡ると vehere 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か？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。皆さんまだ見慣れないかもしれません、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たりするのです。実際、今回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ない^{*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 年説も間違いましたね。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ん。

^{*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。略されていて正確なタイトルが分からぬので、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」とも書かれています。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、それを差し引いても

- 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”的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ります、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。そして皆さんのが後で学ぶ「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れます。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 $x_0 = (x_0, y_0, z_0)$ 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 x に対し $x - x_0$ は $t(a, b, c)$ と平行なので、適当な実数 $t \in \mathbb{R}$ を用いて $x - x_0 = t(a, b, c)$ と書けます。逆に全ての実数 t に対して $x = x_0 + t(a, b, c)$ は直線上の点です。こうして $x = x_0 + t(a, b, c)$ ($t \in \mathbb{R}$) 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。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 x_0 と平面を張るベクトル u, v が与えられると、 $x = x_0 + au + bv$ ($a, b \in \mathbb{R}$) がパラメータ表示です。

さて今度は、2点 p_1, 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 $p_2 - p_1$ なので、直線のパラメータ表示は $x = p_1 + t(p_2 - p_1) = (1-t)p_1 + tp_2$ となります。したがってベクトル p_1 と 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、 $sp_1 + tp_2$ ($s+t=1$) という表示もできます。

平面についても同じように、 p_1, p_2, p_3 を通る平面のパラメータ表示は $x = p_1 + s(p_2 - p_1) + t(p_3 - p_1) = (1-s-t)p_1 + sp_2 + tp_3$ ($s, t \in \mathbb{R}$) あるいは $x = sp_1 + tp_2 + up_3$ ($s+t+u=1$) 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 $x_0 = {}^t(x_0, y_0, z_0)$ を通り $n = {}^t(a, b, c)$ 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 x に対し、 $x - x_0$ は n と垂直になります。したがって

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

です。この式で $d := -ax_0 - by_0 - cz_0$ とおけば、平面上の点 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 $(a, b, c) \neq (0, 0, 0)$ 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、 $x_0 := {}^t(-d/a, 0, 0)$ とすれば良いです。そして $x = {}^t(x, y, z)$ が方程式の解だとすれば、 $ax + by + cz + d = 0$ ので

$$0 = ax + by + cz + d = {}^t(a, b, c) \cdot x - {}^t(a, b, c) \cdot x_0 = {}^t(a, b, c) \cdot (x - x_0)$$

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、一般的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 $\ell: {}^t(x_0, y_0, z_0) + t{}^t(a, b, c)$ が与えられたとします。このとき直線上の点 $x = {}^t(x, y, z)$ 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

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

と書き換えるべき問題ありません。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

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

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、それぞれのパート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 v_1, \dots, v_n が 1 次独立であることの定義は、 $\alpha_1 v_1 + \dots + \alpha_n v_n$ となる $(\alpha_1, \dots, \alpha_n)$ が $(0, \dots, 0)$ 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、 $v_1 = -\alpha_2 v_2 - \dots - \alpha_n v_n$ というように、 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

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

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、 ${}^t(p_1, q_1) \cdot \alpha {}^t(-q_1, p_1) = \alpha(-p_1 q_1 + q_1 p_1) = 0$ 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 u, 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。

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

$$\begin{aligned} u \cdot (u \times v) &= p_1(q_1 r_2 - q_2 r_1) + q_1(r_1 p_2 - r_2 p_1) + r_1(p_1 q_2 - p_2 q_1) = 0 \\ v \cdot (u \times v) &= p_2(q_1 r_2 - q_2 r_1) + q_2(r_1 p_2 - r_2 p_1) + r_2(p_1 q_2 - p_2 q_1) = 0 \end{aligned}$$

となるので、確かに $u \times v$ は u および 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 u, v の張る平行四辺形の符号付き面積^{*23}を $\det(u \ v)$ と書きます。また \mathbb{R}^3 内の3本の縦ベクトル u, v, w の張る平行六面体の符号付き体積^{*24}を $\det(u \ v \ w)$ 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 u, v が1次独立であることと $\det(u \ v) \neq 0$ 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、 u から 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、 u, v, 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、列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、ここまで議論で $\det(u v w) = u \cdot (v \times w)$ が分かりました。この式を使えば、なぜ $v \times w$ が v と直交するのか一目瞭然ですね。 $v \cdot (v \times w)$ はベクトル v, v, w の張る平行六面体の体積ですが、この平行六面体はペッちゃんこなので、どう見ても体積 0 です。こうして内積の値が 0 となり、 v と $v \times w$ が直交すると分かります。 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 a, b, c とすると、外心が $\frac{\sin 2\angle A a + \sin 2\angle B b + \sin 2\angle C c}{\sin 2\angle A + \sin 2\angle B + \sin 2\angle C}$ 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。

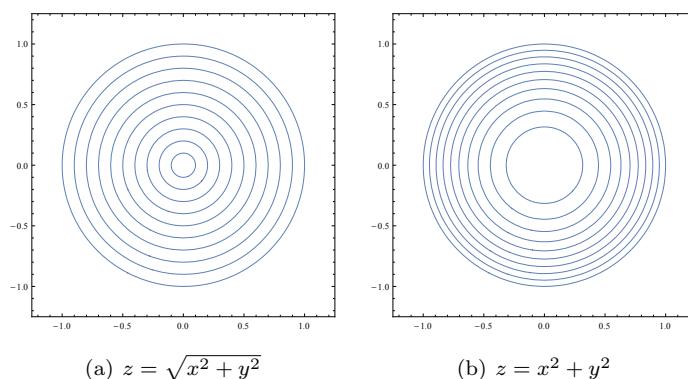
前回の内容を思い出しましょう。「 x, y, z の1次方程式が定める图形は平面である」という事実を確認しました。ですからグラフの方程式 $z = f(x, y)$ と平面の方程式 $ax + by + cz + d = 0$ を連立すれば、グラフ $z = f(x, y)$ の平面 $ax + by + cz + d = 0$ による断面が得られます。特に平面の方程式として分かりやすいもの、たとえば

- 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

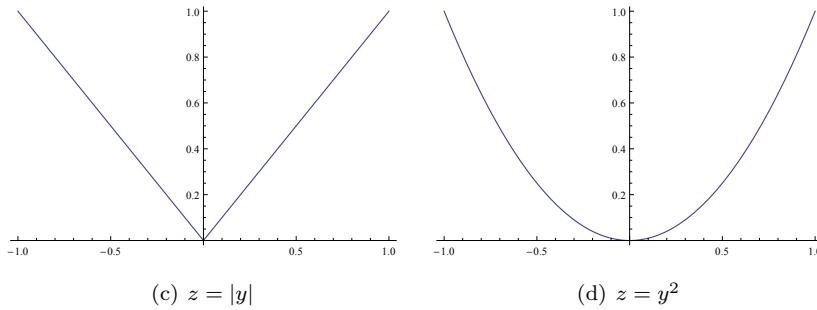
等高線 実数 c に対し、方程式 $z = c$ は z 座標の値が c である、 xy 平面と平行な平面を表します。したがって $z = f(x, y)$ と連立して $c = f(x, y)$ とすれば、平面 $z = c$ による断面が見えます。この断面のことをグラフの等高線といいます。等高線が把握できれば、 c の値を動かすことによってグラフの形状が把握できます。たとえば $z = \sqrt{x^2 + y^2}$ と $z = x^2 + y^2$ のグラフを、 xy 平面上に平行な平面を $z = 0$ から $z = 1$ まで 0.1 刻みで動かしてみましょう。



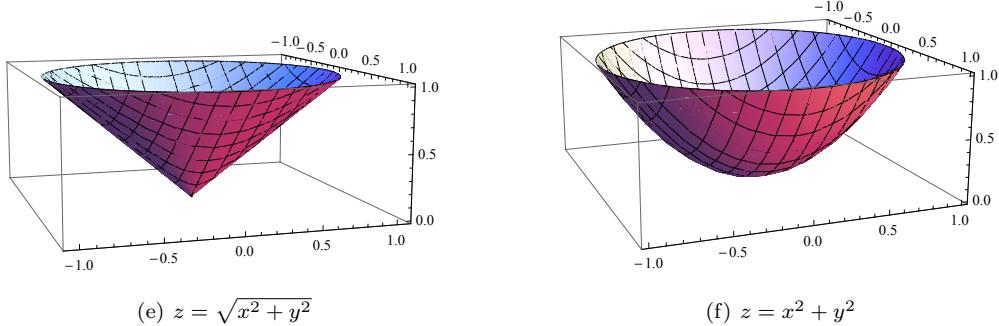
この図から明らかに、どちらのグラフも z 軸回りの回転対称性を持つと分かります。また等高線の密度から、 $z = \sqrt{x^2 + y^2}$ では高さが一定の間隔で増し、 $z = x^2 + y^2$ では急激に高さが増していく感じがします。ですがどの程度増えるのかは、この図だけでははっきりしません。詳しい情報を得るためにには、他の平面での切断面などが必要です。

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。

さつきの $z = \sqrt{x^2 + y^2}$ 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。



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。



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。

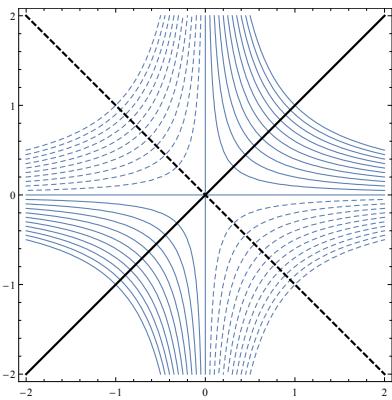
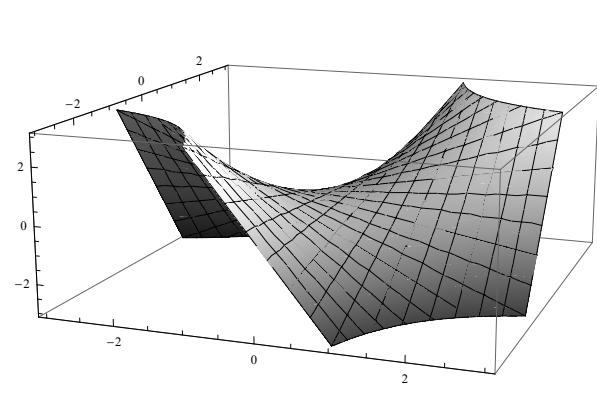
2.2 z 軸を含む平面に関する切断

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。

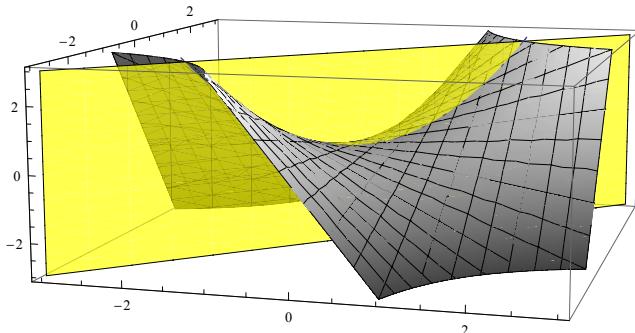
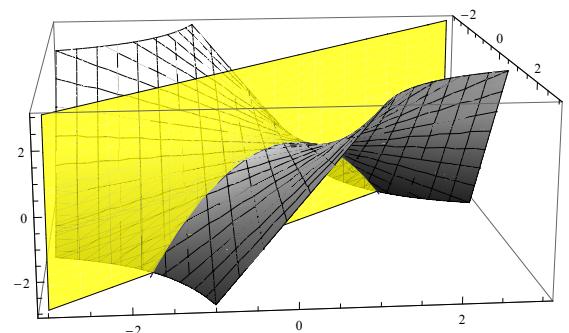
$$z = xy = \frac{(x+y)^2 - (x-y)^2}{4}$$

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、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、グラフの移動といつても平行移動くらいしかすることはありません。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 $e^{-i\theta}(x + iy) = ((\cos \theta)x + (\sin \theta)y) + i(-(\sin \theta)x + (\cos \theta)y)$ 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 $f((\cos \theta)x + (\sin \theta)y, -(\sin \theta)x + (\cos \theta)y)$ 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 $((u+v)/\sqrt{2}, (u-v)/\sqrt{2})$ 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、 $\gamma(t) = (x(t), y(t))$ 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、 $f(x(t), y(t))$ は点 $\gamma(t)$ における f の値を表します。こうしてしまえば $f(x(t), y(t))$ 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 $f(x(t), y(t))$ の t における微分は f と $\gamma'(t) := (x'(t), y'(t))$ だけで決まります^{*6}。そこで $f(x(t), y(t))$ 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 $f(\gamma(t))$ 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かでしょうか。

また $\tilde{\gamma}(t) := (x(t), y(t), f(x(t), y(t)))$ 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 を動かすことと、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

今の方角微分の話で、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 $f(\gamma(t))$ 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

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

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 变数の場合、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 变数函数のグラフ $z = f(x, y)$ の接平面が求められます。 $\tilde{\gamma}$ の微分が曲面 $z = f(x, y)$ の接ベ

^{*5} 写像の言葉で言うなら、曲線を $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$ という写像と思い、2 变数函数 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ 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、それは多变数の場合「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 $(a, b, f(a, b))$ を通るよう調整すれば、接平面の方程式

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

が求まります。

5 解答など

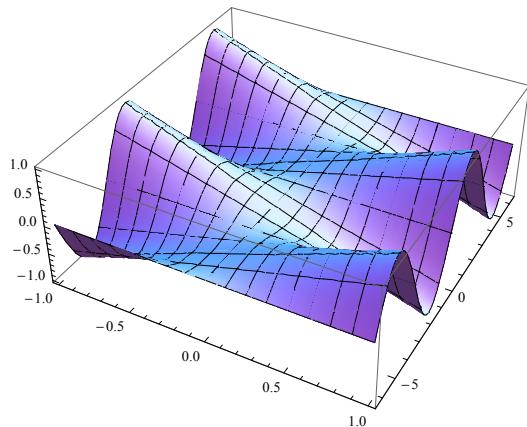
5.1 穴埋め問題の解答

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。

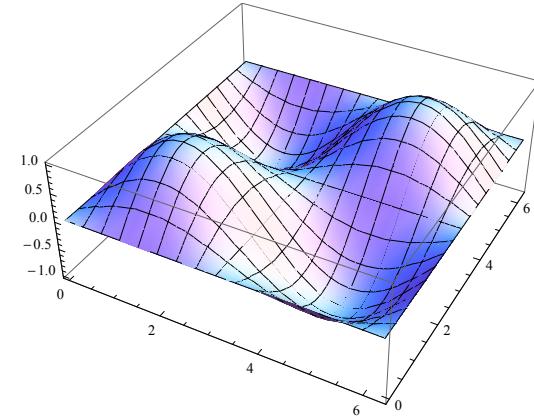
- | | | | | | | | | |
|----------------|------------|-----------|------------|----------|-----------|----------|------------|-------------|
| (1) 平面 | (2) 双曲放物面 | (3) z 軸 | (4) 回転対称性 | (5) 円錐 | (6) 回転放物面 | (7) a | (8) b | |
| (9) \sqrt{c} | (10) $2xy$ | | (11) 双曲放物面 | (12) 双曲線 | (13) a | (14) b | (15) $1/2$ | (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れるよう、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, \dots, a_n が与えられたとき、そ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。

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

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

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

ですね。項を書き出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。

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

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 $1, 2, \dots, n$ 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

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

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

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

^{*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、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 記法に慣れるところなこともできます。

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

\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。もし和にダブリ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、ダブリなくばらせば第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る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、虚数単位とこんな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

$$\sum_{k=1}^n \underbrace{(2k - 1)}_{\text{左辺における } k \text{ のスコープ}} = n^2$$

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ふんですよね……。

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、さらに一步踏み込んで多重の \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

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

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

$$\sum_{k=1}^m \sum_{l=1}^n a_{k,l} = \sum_{k=1}^m (a_{k,1} + a_{k,2} + \dots + a_{k,n}) = \begin{aligned} & a_{1,1} + a_{1,2} + \dots + a_{1,n} \\ & + a_{2,1} + a_{2,2} + \dots + a_{2,n} \\ & + \dots \dots \\ & + a_{m,1} + a_{m,2} + \dots + a_{m,n} \end{aligned}$$

を考えてみましょう。この計算は、各行を横向きに足し合せた結果を縦に足し合わせても、各列を縦向きに足し合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、

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

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。

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

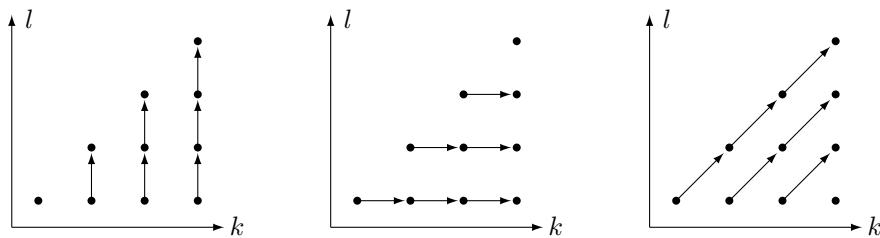
こんな感じで、2つの \sum をくっつけることができます。元々の二重和を「 k が $\{1, 2, \dots, m\}$ を走り、 l が $\{1, 2, \dots, n\}$ を走る」と書き換えた上で、さらに集合の直積を利用して「 (k, l) が $\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$ を走る」と書き換えました。こういう風に集合の直積を使うと、2つの \sum を1つにくっつけられます。逆に \sum の変数が直積集合を走るときは、 \sum を2つにばらすことができます。

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、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。

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

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、 $\{(k, l) \in \{1, 2, \dots, n\} \mid l \leq k\}$ 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

$$\prod_{k=1}^n a_i = a_1 a_2 \cdots a_k$$

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

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

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、 $(a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ 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くください。

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

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

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。慣れない方は、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『進化しそうな脳』(講談社ブルーバックス) の付録を参照。

問 8 の解答 $+ \alpha$

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

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

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

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 $A^2 - (\text{tr } A)A + (\det A)E = O$ 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 $\text{tr}(AB) = \text{tr}(BA)$, $\det(AB) = \det(BA)$ 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 $A^2 - (\text{tr } A)A + (\det A)E = O$ を思い出しましょう。この式を移項すると $(\det A)E = A((\text{tr } A)E - A)$ が得られます。そして実際に計算してみても

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

なので、2 次正方行列 A の余因子行列が $(\text{tr } A)E - A$ 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 $A((\text{tr } A)E - A) = ((\text{tr } A)E - A)A = (\det A)E$ 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

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

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。この条件が一般的の 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。

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

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) ${}^t \mathbf{v} M \mathbf{v} = 4xy + z^2$

(b), (c)

$$A^2 = B^2 = C^2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad {}^t A M A = {}^t B M B = {}^t C M C = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = M$$

(d) ${}^t(A\mathbf{v})M(A\mathbf{v}) = {}^t \mathbf{v} {}^t A M A \mathbf{v} = {}^t \mathbf{v} M \mathbf{v}$ ■

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 ${}^t A$ 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 ${}^t({}^t A) = A$ 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

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

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 $u, v \in \mathbb{R}^n$ に対し、 tuv 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 $u = {}^t(u_1, u_2, u_3)$, $v = {}^t(v_1, v_2, v_3)$ の場合に成分で書けば

$${}^tuv = (u_1 \quad u_2 \quad u_3) \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = (u_1v_1 + u_2v_2 + u_3v_3)$$

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 ${}^tuv = u_1v_1 + u_2v_2 + u_3v_3$ と書きます。成分が実数なら、 tuv 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 u, v が与えられたとき、 tuMv はやはりスカラーになります。実際 Mv が n 次元の縦ベクトルですから、この式を ${}^tu(Mv)$ と読めば、計算結果が u と Mv の内積になります。また $M = E$ のときは ${}^tuEv = {}^tuv$ ですから、 tuMv は内積の一般化だと思えます。どんな M を持って来れば tuMv 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になる行列を巾單行列と言います。 A, B, C は全て巾單行列です。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 $\text{Ad}_A(M) := AMA^{-1}$ 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。どこが間違っていたのかは、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が言っていることなんだから、自分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(?)ごく稀に白黒はつ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、このどちらかに従って数学をして良いのです。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 $f: A \rightarrow B$ 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索しても「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

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

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

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。

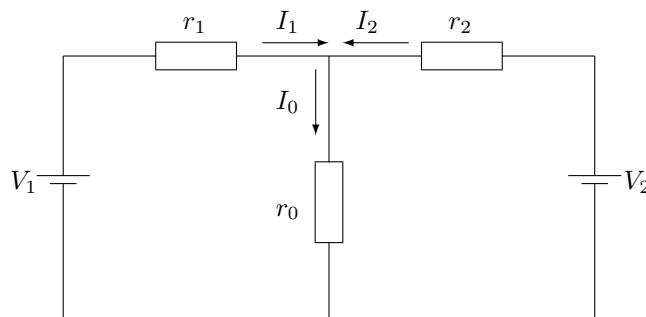
線型写像の一つの例は、行列の積です。 A を (m, n) 型行列とします。このとき $v \in \mathbb{R}^n$ に対し、 $Av \in \mathbb{R}^m$ です。したがって「 A を \mathbb{R}^n のベクトルに左からかける」という操作によって、 A は写像 $\mathbb{R}^n \rightarrow \mathbb{R}^m$ を定めます。そして

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

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

- 上側の分歧点における 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

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

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

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

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」を確認していましたことになります。

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 $x, y \in \mathbb{R}^n$ と実数 α に対し

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

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, \dots, p_n を、 $p_1 + \dots + p_n = 1$ かつ $p_i > 0$ を満たす実数とする。また $x = {}^t(x_1, \dots, x_n) \in \mathbb{R}^n$ とする。このとき期待値は

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

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

$$\begin{aligned} E[x+y] &= \sum_{i=1}^n p_i (x_i + y_i) = \sum_{i=1}^n p_i x_i + \sum_{i=1}^n p_i y_i = E[x] + E[y] \\ E[\alpha x] &= \sum_{i=1}^n p_i \alpha x_i = \alpha \sum_{i=1}^n p_i x_i = \alpha E[x] \end{aligned}$$

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。

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 $p = {}^t(p_1, \dots, p_n) \in \mathbb{R}^n$ とおくと、 $E[x] = {}^t p x$ 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 ${}^t p$ 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 $p_1 = \dots = p_n := 1/n$ 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 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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 e_i と書きます。これを用いて $v_i := f(e_i) \in \mathbb{R}^m$ とおき、さらにベクトル v_1, \dots, v_n たちを並べて $A = (v_1 \ v_2 \ \dots \ v_n)$ 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 e_1, \dots, 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 v を成分表示し $v = {}^t(a_1, a_2, \dots, a_n) = \sum_{i=1}^n a_i e_i$ と書くと

$$f\left(\sum_{i=1}^n a_i e_i\right) = \sum_{i=1}^n f(a_i e_i) = \sum_{i=1}^n a_i f(e_i) = \sum_{i=1}^n a_i v_i$$

が成り立ち、その一方で

$$A \sum_{i=1}^n a_i e_i = \sum_{i=1}^n a_i A e_i = \sum_{i=1}^n a_i v_i$$

も成り立ちます。よって $f(v) = Av$ です。 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 $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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 $(1) \Rightarrow (2)$ 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) \Rightarrow (2)は正しいですが、(2) \Rightarrow (1)は正しくありません。反例はたとえば2倍写像 $f(u) := 2u$ です。

ベクトル u, v のなす角は $u \cdot v / |u||v|$ で与えられます。いま f を2倍写像とすると、 $f(u) \cdot f(v) / |f(u)||f(v)| = 2u \cdot 2v / |2u||2v| = u \cdot v / |u||v|$ 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) \Rightarrow (2) の解答 任意のベクトル $u, v \in \mathbb{R}^2$ に対し

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

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

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

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 $u \cdot v / |u||v|$ も保つ。 ■

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

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

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 u, v について、始点を揃えたとき u から v へ向かう方向が正の向き（反時計回り）になることを「 (u, v) は正の向きである」と呼びましょう。たとえば (e_1, 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。ここでもう一步、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

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

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 $u, v, w \in V$ に対し $(u + v) + w = u + (v + w)$ が成り立つ
 - 任意の $u \in V$ に対し $u + 0 = 0 + u = u$ となる $0 \in V$ が唯一存在する
 - 任意の $u \in V$ に対し、 $u + v = 0$ となる $v \in V$ が唯一存在する（この v を $-u$ と書く）
 - 任意の $u, v \in V$ に対し $u + v = v + u$ 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 $v \in V$ に対し $\alpha(\beta v) = (\alpha\beta)v$ が成り立つ
 - 任意の $v \in V$ に対し、 $1v = v$ が成り立つ
- スカラー倍と和が分配法則を満たす
 - 任意の $u, v \in V$ と $\alpha \in \mathbb{R}$ に対し $(\alpha + \beta)v = \alpha v + \beta v$
 - 任意の $u \in V$ と $\alpha, \beta \in \mathbb{R}$ に対し、 $\alpha(u + v) = \alpha u + \alpha v$

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

- $0v = (0 + 0)v = 0v + 0v$ の両辺に $-(0v)$ を足して、 $0 = 0v + 0 = 0v$
- $0 = 0v = (1 + (-1))v = v + (-1)v$ より、 $-v = (-1)v$

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。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', \dots の項が齊次 1 次式^{*6}」という点でした。

一般に $a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = 0$ のように、微分方程式であって y, y', y'', \dots に関する齊次 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 (u_1, \dots, u_n) であって「任意の $v \in V$ が $v = a_1u_1 + a_2u_2 + \dots + a_nu_n$ 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 e_i と書くと、 \mathbb{R}^n の基底として (e_1, e_2, \dots, 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には、 ${}^t(\alpha, \beta) = C_1 {}^t(1, 3) + C_2 {}^t(1, -2)$ 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「齊次」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

線形代数学 第1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く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しまいます。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。

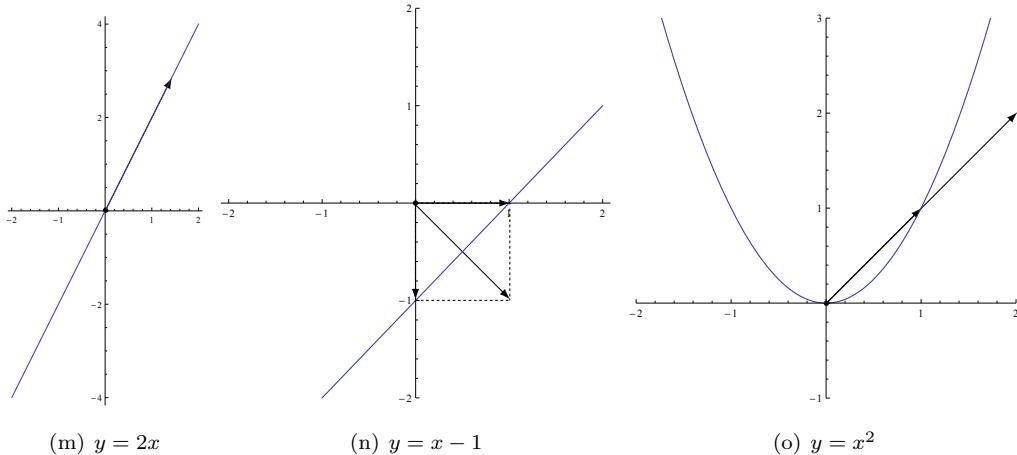
- V 上に加法が定義されており
 - 任意の $u, v, w \in V$ に対し $(u + v) + w = u + (v + w)$ が成り立つ^{*3}
 - 任意の $u \in V$ に対し $u + \mathbf{0} = \mathbf{0} + u = u$ となる $\mathbf{0} \in V$ が唯一存在する
 - 任意の $u \in V$ に対し、 $u + v = \mathbf{0}$ となる $v \in V$ が唯一存在する（この v を $-u$ と書く）
 - 任意の $u, v \in V$ に対し $u + v = v + u$ が成り立つ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

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

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

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



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、 $v := {}^t(1, 1)$ は放物線の上に乗りますが、2 倍した $2v = {}^t(2, 2)$ 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。

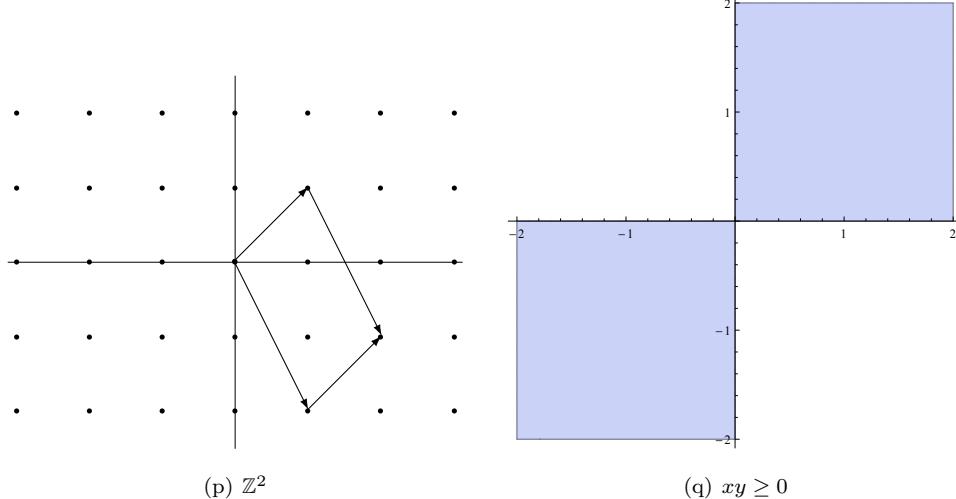
^{*4} すぐ後ろで議論はしますが、それを読む前に自分で定義に従って、どれが線型空間か判定してみると良いと思います。

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 ${}^t(b, 2b)$ を取ってみると、これらの和 ${}^t(a+b, 2(a+b))$ も再び直線 $y = 2x$ に乘ります。またスカラー倍も $\alpha {}^t(a, 2a) = (\alpha a, 2\alpha a)$ となり、ちゃんと直線 $y = 2x$ 上に乘ります。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 $y = x - 1$ は線型空間にならないません。たとえばベクトル ${}^t(1, 0), {}^t(0, -1)$ はこの直線上に乘りますが、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 $\mathbb{Z}^2 := \{{}^t(m, n) \in \mathbb{R}^2 \mid m, n \in \mathbb{Z}\}$ 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 $\{{}^t(x, y) \in \mathbb{R}^2 \mid xy \geq 0\}$ 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。



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、 $f: V \rightarrow W$ が線型写像であるとは

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

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。

$$\begin{array}{ccc} (\mathbf{v}, \mathbf{w}) & \xrightarrow{f \times f} & (f(\mathbf{v}), f(\mathbf{w})) \\ \downarrow + & & \downarrow + \\ \mathbf{v} + \mathbf{w} & \xrightarrow{f} & f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w}) \end{array} \quad \begin{array}{ccc} (\alpha, \mathbf{v}) & \xrightarrow{f \times f} & (\alpha, f(\mathbf{w})) \\ \downarrow \cdot & & \downarrow \cdot \\ \alpha \mathbf{v} & \xrightarrow{f} & f(\alpha \mathbf{v}) = \alpha f(\mathbf{v}) \end{array}$$

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 $f: \mathbb{R} \rightarrow \mathbb{R}$ 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 $u, v \in W$ に対し、 $u + v \in W$ が成り立つ
- 任意の $u \in W$ と $\alpha \in \mathbb{R}$ に対し、 $\alpha u \in W$ 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あります。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} 雜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。ここをまず押させてください。

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 u, v, w に対して $(u + v) + w = u + (v + w)$ 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「 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 $u, v, w \in W$ に対し $(u + v) + w = u + (v + w)$ 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 $v \in W$ が取れる。すると W がスカラー倍で閉じていることから、 $\mathbf{0} = 0v \in W$ となる。
 - W がスカラー倍で閉じているので、 $v \in W$ のとき $-v = (-1)v \in W$ となる。また $v + u = \mathbf{0}$ となる 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 $u, v \in W$ に対し $u + v = v + u$ 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 の中に定義されている。そして
 - 任意の $v \in W$ と $\alpha, \beta \in \mathbb{R}$ について、 $\alpha(\beta v) = (\alpha\beta)v$ 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 $v \in W$ について $1v = v$ 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、

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

^{*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 $Ax = b$ と書けます。特に $b = \mathbf{0}$ 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かでしょうか。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 变数の 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ともに「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图形でしたから、連立方程式の定める图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 $\text{Mat}_{m,n}(\mathbb{R})$ で表します。 $A \in \text{Mat}_{m,n}(\mathbb{R})$ として方程式 $Ax = b$ 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、图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 $A \in \text{Mat}_{m,n}(\mathbb{R})$ とし、 $V := \{x \in \mathbb{R}^n \mid Ax = \mathbf{0}\}$ とおく。このとき $\mathbf{0} \in V$ より V は空でない。そして

- (1) $u, v \in V$ を任意に取る。このとき $Au = Av = \mathbf{0}$ なので、 $A(u + v) = Au + Av = \mathbf{0} + \mathbf{0} = \mathbf{0}$ である。よって $u + v \in V$ となる。
- (2) $\alpha \in \mathbb{R}, u \in V$ を任意に取る。このとき $A(\alpha u) = \alpha(Au) = \alpha(\mathbf{0}) = \mathbf{0}$ なので、 $\alpha u \in V$ 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 变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、 $(f(x) + g(x)) + h(x) = f(x) + (g(x) + h(x))$ が成り立つ。
 - 0 は多項式で、任意の多項式 $f(x)$ に対して $f(x) + 0 = 0 + f(x) = f(x)$ 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が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 $Ax = b$ の解空間は、 $Ax = \mathbf{0}$ 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 $(n - 1)$ 次元の图形を定めます。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 $\text{ev}_\alpha(f(x)) := f(\alpha)$ と定める^{*14}ことで、多項式に実数を対応させる写像 $\text{ev}_\alpha: \mathbb{R}[x] \rightarrow \mathbb{R}$ 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 $\text{ev}_\alpha(f(x) + g(x)) = f(\alpha) + g(\alpha) = \text{ev}_\alpha(f(x)) + \text{ev}_\alpha(g(x))$
- 任意の多項式 $f(x) \in \mathbb{R}[x]$ と実数 $\alpha \in \mathbb{R}$ に対し、 $\text{ev}_\alpha(\alpha f(x)) = \alpha f(\alpha) = \alpha \text{ev}_\alpha(f)$

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) $W := \{f(x) \in V \mid \text{ev}_\alpha(f) = 0\}$ とおく^{*15}。このとき (2) より

- 任意の $f(x), g(x) \in W$ に対し、 $\text{ev}_\alpha(f(x) + g(x)) = f(\alpha) + g(\alpha) = 0 + 0 = 0$ 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、 $\text{ev}_\alpha(af(x)) = f(\alpha) = 0$ 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」などというのも考えることができます。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, \dots, 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 $\{a_0, a_1, a_2, \dots\}$ 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, \dots) 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。

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

$(a_n)_{n \geq 0}$ 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 の記号はこれから取りました。

*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。

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

いま、実数列全体の集合を $\text{Map}(\mathbb{N}, \mathbb{R})$ と書くことにします^{*19}。これまでの議論で $\text{Map}(\mathbb{N}, \mathbb{R})$ 上には加法とスカラーベ倍とが定義されました。この $\text{Map}(\mathbb{N}, \mathbb{R})$ 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

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

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、 $(a_n)_{n \geq 0} + (-a_n)_{n \geq 0} = (a_n + (-a_n))_{n \geq 0} = (0)_{n \geq 0}$ 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 $\alpha\{\beta(a_n)_{n \geq 0}\} = \alpha(\beta a_n)_{n \geq 0} = (\alpha\beta a_n)_{n \geq 0} = (\alpha\beta)(a_n)_{n \geq 0}$ 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、

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

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

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

が成り立つ。

$(0, 2, 4, 6, 8, \dots), (b_n)_{n \geq 0} = (3n)_{n \geq 0} = (0, 3, 6, 9, 12, 15, \dots)$ なら、 $(2n)_{n \geq 0} + (3n)_{n \geq 0} = (5n)_{n \geq 0} = (0, 5, 10, 15, 20, \dots)$ 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 $\text{Map}(S, \mathbb{R})$ 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 $\text{Map}(\mathbb{N}, \mathbb{R})$ 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) $\varphi: V \rightarrow \mathbb{R}^2$ を、 $\varphi((a_n)_{n \geq 0}) := {}^t(a_0, a_1)$ 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。

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

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

(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。適宜参考してください。

函数空間 $\text{Map}(\mathbb{R}, \mathbb{R})$ で、 \mathbb{R} 上の（連續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 $x \in \mathbb{R}$ 每に 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、 $\text{Map}(\mathbb{R}, \mathbb{R})$ 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

$$\begin{aligned} ((f + g) + h)(x) &= (f + g)(x) + h(x) = f(x) + g(x) + h(x) \\ (f + (g + h))(x) &= f(x) + (g + h)(x) = f(x) + g(x) + h(x) \end{aligned}$$

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 $(\alpha(\beta f))(x) = \alpha((\beta f)(x)) = \alpha(\beta f(x)) = (\alpha\beta)f(x) = ((\alpha\beta)f)(x)$ 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

^{*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。

– 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、

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

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、

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

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。单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。

ちなみに单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 单振動の微分方程式 $y'' = -y$ 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 $\text{Map}(\mathbb{R}, \mathbb{R})$ 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 $(f_1(x) + f_2(x))'' = f_1''(x) + f_2''(x) = -f_1(x) - f_2(x) = -(f_1(x) + f_2(x))$ 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 $(\alpha f(x))'' = \alpha f''(x) = -\alpha f(x)$ より、 $\alpha f \in V$ である。

これで V が部分空間だとえた。

(2) 写像 $\varphi: V \rightarrow \mathbb{R}^2$ を $\varphi(f(x)) := {}^t(f(0), f'(0))$ 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。

^{*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、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。

- 任意の $f \in V, \alpha \in \mathbb{R}$ に対し

$$\varphi(\alpha f) = \begin{pmatrix} (\alpha f)(0) \\ (\alpha f)'(0) \end{pmatrix} = \begin{pmatrix} \alpha f(0) \\ \alpha f'(0) \end{pmatrix} = \alpha \begin{pmatrix} f(0) \\ f'(0) \end{pmatrix} = \alpha \varphi(f)$$

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

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

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

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

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

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

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 $(\psi_a f)''(x) = (\psi_a f'')(x) = (\psi_a(-f))(x) = (-\psi_a f)(x) = -(\psi_a f)(x)$ 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、さらに一步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、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 $\text{Map}(S, \mathbb{R})$ 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、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 $\text{Map}(S, V)$ は

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

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 $\text{Map}(S, V)$ 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はありません。そこで $\text{Map}(S, V)$ 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、 $\text{Map}(\mathbb{R}, \mathbb{R})$ 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 $a \in \text{Map}(\mathbb{N}, \mathbb{R})$ は $a: \mathbb{N} \rightarrow \mathbb{R}$ 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 $a(0), a(1), a(2), \dots$ 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ます。こうすると一般論から、直ちに $\text{Map}(\mathbb{N}, \mathbb{R})$ 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 $\text{Map}(S, V)$ 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 $\text{Map}(S, V)$ 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、 $\text{Map}(S, V)$ 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 $\text{Map}(S, V)$ 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 $\text{Hom}_{\mathbb{R}}(V, W)$ と表します。また、特に $\text{Hom}_{\mathbb{R}}(V, V)$ のことを $\text{End}_{\mathbb{R}}(V)$ 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 $\text{Mat}_{m,n}(\mathbb{R})$ 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 $\text{Mat}_{m,n}(\mathbb{R})$ は線型空間になります。よって一般化した $\text{Hom}(V, W)$ や $\text{End}(V)$ も線型空間になりそうです。実際に $\text{Hom}(V, W)$ が $\text{Map}(V, W)$ の部分空間であることは、次のようにチェックできます。

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

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

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 $\mathbf{u} \in V$ と $\alpha \in \mathbb{R}$ に対し

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

が成り立つ。

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

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

- 任意の $\mathbf{u}, \mathbf{v} \in V$ に対し

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

が成り立つ。

- 任意の $\mathbf{u} \in V$ と $\beta \in \mathbb{R}$ に対し

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

が成り立つ。

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

線型写像の合成 U, V, W を線型空間とし、 $f: U \rightarrow V$ と $g: V \rightarrow W$ を共に線型写像とします。このとき合成写像 $g \circ f: U \rightarrow W$ 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 $\mathbf{u}, \mathbf{v} \in U$ に対し

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

が成り立つ。

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

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

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。

$\text{End 環 } V$ を線型空間とします。既に示したように $\text{End}(V) = \text{Hom}(V, V)$ は線型空間になります。一方で $f, g \in \text{End}(V)$ なら、 $\text{End}(V)$ 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」という疑問が湧き起きます。

定義を考えれば分かるのですが、 $f, g, h \in \text{End}(V)$ 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。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 $\text{End}(V)$ では、写像の合成。を掛け算だと思ったとき、分配法則が成り立っています。また $a \in \mathbb{R}$ で $f, g \in \text{End}(V)$ 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 $\text{End}(V)$ 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、 $\text{End}(V)$ 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 $\text{End}(V)$ の環構造からは「 $\text{End}(V)$ 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

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

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

が成り立ちます。これで $\text{Ker } f$ 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」という事実を知つていれば、

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

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 $\text{Map}(\mathbb{N}, \mathbb{R})$ 上の「左シフト写像」 $L: \text{Map}(\mathbb{N}, \mathbb{R}) \rightarrow \text{Map}(\mathbb{N}, \mathbb{R})$ を、

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

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 L 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 $(a_n)_{n \geq 0}, (b_n)_{n \geq 0} \in \text{Map}(\mathbb{N}, \mathbb{R})$ に対し

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

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

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

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、 $L^2 - L - \text{id}$ 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、 $(L^2 - L - \text{id})(a_n)_{n \geq 0}$ 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 $L^2 - L - \text{id}$ の核 $\text{Ker}(L^2 - L - \text{id})$ 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 $C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ 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、 $\frac{d^2}{dx^2} - \text{id}$ 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 $(\frac{d}{dx})^2 - \text{id}$ 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 $\text{Map}(S, V)$ が線型空間になることを示す
- $\text{End}(V)$ 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 $\text{id}(x) = x$ 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 第2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

今回の答案を探点していて、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 $v \in W$ が取れる。すると $0v = 0$ である。 W はスカラー倍で閉じているので、 $0 \in W$ が従う。

(2) $a \in V$ として、 $\mathbb{R}a := \{Ca \mid C \in \mathbb{R}\}$ とおく。すると

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

よって $\mathbb{R}a$ は部分空間である。

(3) $a, b \in V$ として、 $\mathbb{R}a + \mathbb{R}b := \{\alpha a + \beta b \mid \alpha, \beta \in \mathbb{R}\}$ とおく。

- $u, v \in \mathbb{R}a + \mathbb{R}b$ とする。このとき $u = \alpha_1 a + \beta_1 b, v = \alpha_2 a + \beta_2 b$ と書ける。よって $u + v = (\alpha_1 + \alpha_2)a + (\beta_1 + \beta_2)b \in \mathbb{R}a + \mathbb{R}b$ となる。
- $u \in \mathbb{R}a + \mathbb{R}b, \gamma \in \mathbb{R}$ とする。このとき $u = \alpha a + \beta b$ と書けるので $\gamma u = \gamma(\alpha a + \beta b) = \gamma \alpha a + \gamma \beta b \in \mathbb{R}a + \mathbb{R}b$ を得る。

よって $\mathbb{R}a + \mathbb{R}b$ 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 $\mathbb{R}a$ を、ベクトル 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 a 方向の直線」が $\mathbb{R}a$ 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。

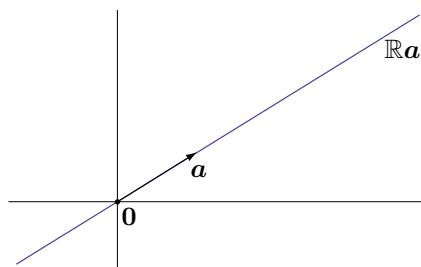


図 2.1 ベクトルの張る部分空間

$\mathbb{R}a + \mathbb{R}b$ も状況は同じです。今度はベクトルが 2 本あって、 a と 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 a と 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つけましょう。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 $\mathbb{R}a + \mathbb{R}b$ 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 a に対し、 a の α 倍は αa 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「 αa 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 $\mathbb{R}a$ という表記が用いられます。

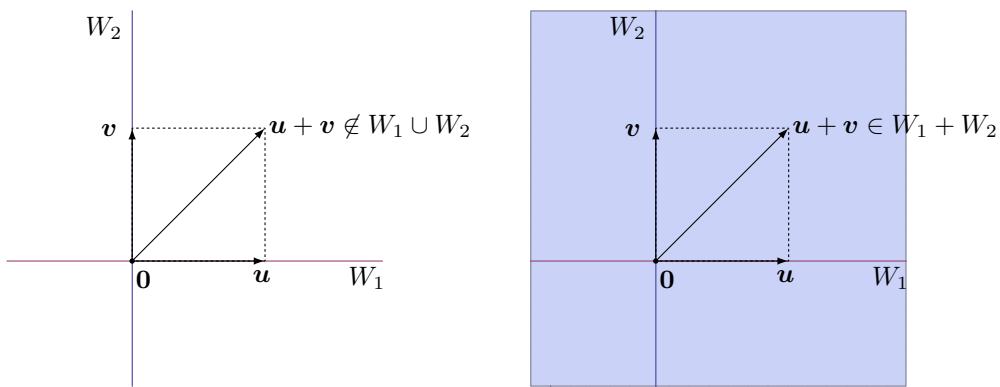
$\mathbb{R}a + \mathbb{R}b$ も同じです。 $\mathbb{R}a$ と $\mathbb{R}b$ はどちらもベクトルの集合で、 $u \in \mathbb{R}a$ と $v \in \mathbb{R}b$ に対して和 $u + v$ が考えられます。そして $u \in \mathbb{R}a$ と $v \in \mathbb{R}b$ とを全て動かしたときにできるベクトル $u + v$ をかき集めるというのが、 $\mathbb{R}a + \mathbb{R}b$ という記号の意味です。

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、あと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。一般的の場合も同じで、 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って、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ます。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。

- $u, v \in W_1 \cap W_2$ とする。このとき $u, v \in W_1$ ので $u + v \in W_1$ である。また $u, v \in W_2$ ので $u + v \in W_2$ である。よって $u + v \in W_1 \cap W_2$ である。
- $u \in W_1 \cap W_2, \alpha \in \mathbb{R}$ とする。このとき $u \in W_1$ より $\alpha u \in W_1$ である。また $u \in W_2$ より $\alpha u \in W_2$ である。よって $\alpha u \in W_1 \cap W_2$ 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。 $W_1 := \mathbb{R}e_1, W_2 := \mathbb{R}e_2$ ^{*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 $e_1 \in W_1 \subset W_1 \cup W_2, e_2 \in W_2 \subset W_1 \cup W_2$ だが $e_1 + e_2 \notin W_1 \cup W_2$ 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。

- $u, v \in W_1 + W_2$ とする。このとき $u = a_1 + b_1, v = a_2 + b_2$ ($a_1, a_2 \in W_1, b_1, b_2 \in W_2$) と書ける。よって $u + v = (a_1 + a_2) + (b_1 + b_2)$ となる。ここで W_1, W_2 は共に部分空間だから、 $a_1 + a_2 \in W_1, b_1 + b_2 \in W_2$ となる。よって $u + v \in W_1 + W_2$ となる。
- $u \in W_1 + W_2, \alpha \in \mathbb{R}$ とする。このとき $u = a + b$ ($a \in W_1, b \in W_2$) と書ける。 W_1, W_2 は部分空間なので $\alpha a \in W_1, \alpha b \in W_2$ となる。よって $\alpha u = \alpha(a + b) = \alpha a + \alpha b \in W_1 + W_2$ 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 $\text{Map}(\mathbb{N}, \mathbb{R})$ の部分集合であることをと、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 $V = \text{Map}(\mathbb{N}, \mathbb{R})$ 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。 $n \rightarrow \infty$ のとき $a_n \rightarrow 0, b_n \rightarrow 0$ ので、 $a_n + b_n \rightarrow 0 + 0 = 0$ 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。 $n \rightarrow \infty$ のとき $a_n \rightarrow 0$ ので、 $\alpha a_n \rightarrow \alpha \cdot 0 = 0$ 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 $|a_n + b_n| \leq |a_n| + |b_n| < M_1 + M_2$ 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} $e_1 := {}^t(1, 0), e_2 := {}^t(0, 1)$ 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、順番に見ていきましょう。

3.1 線型写像の核

Ker の定義 V と W を線型空間とし、 $f: V \rightarrow W$ を線型写像とします。このとき線型写像 f の核 (kernel) は

$$\text{Ker } f := \{v \in V \mid f(v) = \mathbf{0}\}$$

と定義されます。つまり f によって 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、 $\Psi: V \rightarrow W$ を線型写像とする。

- (1) $\Psi(\mathbf{0}) = \Psi(\mathbf{0} - \mathbf{0}) = \Psi(\mathbf{0}) - \Psi(\mathbf{0}) = \mathbf{0}'$ である。
- (2) $\Psi(u) = \Psi(v)$ ならば、 Ψ の線型性を使って $\mathbf{0} = \Psi(u) - \Psi(v) = \Psi(u - v)$ を得る。よって $u - v \in \text{Ker } \Psi$ である。
- (3) $a \in V, b = \Psi(a) \in W$ とする。このとき $v \in V$ が $\Psi(v) = b$ を満たすとすると、 $\Psi(v - a) = \Psi(v) - \Psi(a) = b - b = \mathbf{0}'$ となる。よって $v - a \in \text{Ker } \Psi$ となるから、 $v = a + (v - a)$ が求める分解を与える。
- (4) $\Psi(\mathbf{0}) = \mathbf{0}'$ である。よって Ψ が単射なら $\Psi(v) = \mathbf{0}'$ なる v は 0 しかない。これより $\text{Ker } \Psi = \{\mathbf{0}\}$ である。

逆に $\text{Ker } \Psi = \{\mathbf{0}\}$ とする。このとき $u, v \in V$ が $\Psi(u) = \Psi(v)$ を満たしたとすると、(2) より $u - v \in \text{Ker } \Psi = \{\mathbf{0}\}$ となる。よって $u - v = \mathbf{0}$ だから $u = v$ 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、 $\text{supp}(a_n)_{n \geq 0} := \{n \in \mathbb{N} \mid a_n \neq 0\}$ と定めます。つまり $\text{supp}(a_n)_{n \geq 0}$ 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 $\text{supp}(a_n + b_n)_{n \geq 0} \subset \text{supp}(a_n)_{n \geq 0} \cup \text{supp}(b_n)_{n \geq 0}$ 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

$$\#\text{supp}(a_n + b_n)_{n \geq 0} \leq \#(\text{supp}(a_n)_{n \geq 0} \cup \text{supp}(b_n)_{n \geq 0}) \leq \#\text{supp}(a_n)_{n \geq 0} + \#\text{supp}(b_n)_{n \geq 0} < \infty$$

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 $\{x \in \mathbb{R} \mid f(x) \neq 0\}$ を f の台 (support) といい、 $\text{supp } f$ 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 $v \in \text{Ker } f$ のことを “ f kills v ” と表すこともあります。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 $f: X \rightarrow Y$ 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 $\text{Ker } f$ を調べれば単射性が分かることを教えてくれるので。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でてきます。そういう意味で $\text{Ker } f$ 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」と言わっても「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。

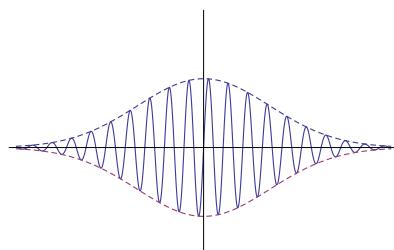


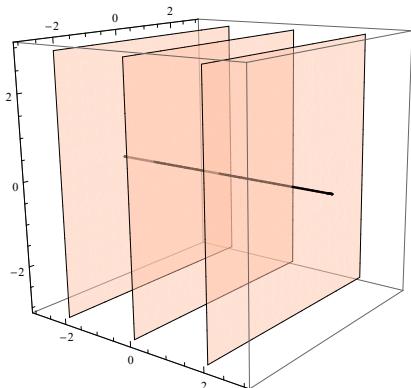
図 2.2 $y = e^{-x^2} \sin 30x$ のグラフ

ところが線型写像の場合、問 6 (2), (3) で示したように

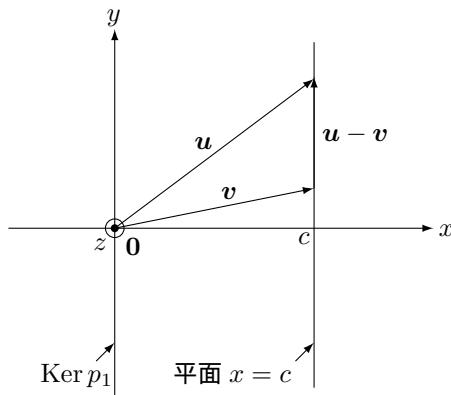
- u と v のズレ $u - v$ が $\text{Ker } f$ に入っていたら、 f による値は同じ
- $f(u) = f(v)$ のように f による値が一致したら、 u と v のズレ $u - v$ は必ず $\text{Ker } f$ に含まれる

と分かります。つまり線型写像の単射性が破れる場合は、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 $p_1: \mathbb{R}^3 \rightarrow \mathbb{R}$ を第 1 成分への射影、つまり $p_1((x, y, z)) := x$ 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、 $\text{Ker } p_1$ は平面 $x = 0$ 、つまり yz 平面だと分かります。そして yz 平面上に平行な平面 $x = c$ 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 $\text{Ker } p_1$ と平行に並ぶ様子



(b) $\text{Ker } p_1$ 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「单射性の破れが Ker でコントロールされる」といったのは、今の場合

- $u - v$ が yz 平面と平行だったら、 u と v の x 座標が同じ
- u と v の x 座標が同じだったら、 $u - v$ 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 の方向には单射性が破れるわけですから、それ以外の方向を見れば单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。核が線型写像の单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 $f: X \rightarrow Y$ 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、 $\text{Im } f$ と書きます。記号で書けば $\text{Im } f := \{y \in Y \mid \exists x \in X \text{ s.t. } y = f(x)\}$ です。線型写像も写像ですから、一般的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、線型写像の像是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、 $f: V \rightarrow W$ を線型写像とします。

- $w_1, w_2 \in \text{Im } f$ とする。このとき像の定義から $f(v_1) = w_1, f(v_2) = w_2$ となる $v_1, v_2 \in V$ が存在する。したがって f の線型性から $w_1 + w_2 = f(v_1) + f(v_2) = f(v_1 + v_2) \in \text{Im } f$ となる。
- $w \in \text{Im } f, \alpha \in \mathbb{R}$ とする。このとき像の定義から、 $f(v) = w$ となる $v \in V$ が存在する。したがって f の線型性から $\alpha w = \alpha f(v) = f(\alpha v) \in \text{Im } f$ となる。

これで $\text{Im } f$ 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、 $\text{Im } f$ は「 A の各列のベクトルが張る空間」になります。これを考慮すると少し $\text{Im } f$ 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 $\Psi: \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ を $\Psi(f(x)) := (x - \alpha)f(x)$ と定める。また、写像 $\text{ev}_\alpha: \mathbb{R}[x] \rightarrow \mathbb{R}$ を $\text{ev}_\alpha(f(x)) := f(\alpha)$ 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 $\Psi(f(x) + g(x)) = (x - \alpha)(f(x) + g(x)) = (x - \alpha)f(x) + (x - \alpha)g(x) = \Psi(f(x)) + \Psi(g(x))$ である。

^{*8} この ev_α は問題文中的 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、 $\Psi(f(x)g(x)) = (x - \alpha)f(x)g(x) = g(x)(x - \alpha)f(x) = g(x)\Psi(f(x))$ となる^{*9}。特に $g(x)$ を定数にとって $g(x) = \alpha$ とすれば、 $\Psi(\alpha f(x)) = \alpha\Psi(f(x))$ となる。

(2) $f(x) \in V$ を任意に取る。このとき $\text{ev}_\alpha(\Psi(f(x))) = \text{ev}_\alpha((x - \alpha)f(x)) = 0f(\alpha) = 0$ である。

(3) 今の(2)で、 $\text{Im } \Psi \subset \text{Ker } \text{ev}_\alpha$ が示せた。そこで逆向きの包含関係 $\text{Ker } \text{ev}_\alpha \subset \text{Im } \Psi$ を示せばよい。

$f(x) \in \text{Ker } \text{ev}_\alpha$ 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 $f(x) = \Psi(g(x)) \in \text{Im } \Psi$ である。■

問5の解答 写像 $\Phi: \mathbb{R}[x] \rightarrow \mathbb{R}^2$ を $\Phi(f(x)) := {}^t(f(\alpha), f'(\alpha))$ 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

$$\Phi(f(x) + g(x)) = \begin{pmatrix} f(\alpha) + g(\alpha) \\ f'(\alpha) + g'(\alpha) \end{pmatrix} = \begin{pmatrix} f(\alpha) \\ f'(\alpha) \end{pmatrix} + \begin{pmatrix} g(\alpha) \\ g'(\alpha) \end{pmatrix} = \Phi(f(x)) + \Phi(g(x))$$

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、

$$\Phi(\beta f(x)) = \begin{pmatrix} \beta f(\alpha) \\ \beta f'(\alpha) \end{pmatrix} = \beta \begin{pmatrix} f(\alpha) \\ f'(\alpha) \end{pmatrix} = \beta \Phi(f(x))$$

が成り立つ。

(2) 線型写像 $\Psi: \mathbb{R}[x] \rightarrow \mathbb{R}[x]$ を $\Psi(f(x)) := (x - \alpha)^2 f(x)$ で定める^{*10}。すると任意の多項式 $f(x) \in \mathbb{R}[x]$ に対し、 $\Psi(f(x)) = (x - \alpha)^2 f(x)$ だから $\Psi(f(x))' = 2(x - \alpha)f(x) + (x - \alpha)^2 f'(x)$ である。よって $\Phi(\Psi(f(x))) = 0$ となるから、 $\text{Im } \Psi \subset \text{Ker } \Phi$ である。

逆に $f(x) \in \text{Ker } \Phi$ 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 $f(x) = (x - \alpha)^2 h(x) \in \text{Im } \Psi$ が分かる。ゆえに $\text{Ker } \Phi \subset \text{Im } \Psi$ である。以上より、 $\text{Ker } \Phi = \text{Im } \Psi$ 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

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(\alpha)}{k!} (x - \alpha)^k = f(\alpha) + f'(\alpha)(x - \alpha) + \frac{f''(\alpha)}{2!} (x - \alpha)^2 + \dots$$

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」と言わされたら、 $f(a)$ 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「実数 $\alpha \in \mathbb{R}$ と多項式 $f(x) \in \mathbb{R}[x]$ に対し $\Psi(\alpha f(x)) = \alpha\Psi(f(x))$ 」ですが、今は一段と強く「多項式 $g(x) \in \mathbb{R}[x]$ と多項式 $f(x) \in \mathbb{R}[x]$ に対し、 $\Psi(g(x)f(x)) = g(x)\Psi(f(x))$ 」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。

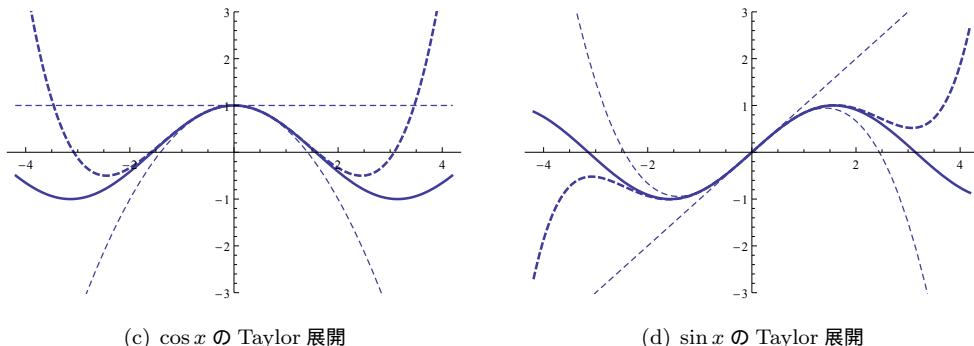
試しに有名な例で計算してみましょう。 $f(x) := \cos x$ とおくと、 $f'(x) = -\sin x$, $f''(x) = -\cos x$, $f'''(x) = \sin x$, $f''''(x) = \cos x$ なので、 f は 4 回微分したら元に戻ります。したがって f の k 回微分 $f^{(k)}(x)$ に $x = 0$ を代入すると、生き残るのは k が偶数の項だけです。そして ± 1 が交互に登場するので、 $k = 2l$ とおくと $f^{(2l)}(0) = (-1)^l$ が分かります。これをさっきの Taylor 展開の式に代入すれば、 $\cos x$ の $x = 0$ における Taylor 展開

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^l}{(2l)!} x^{2l} = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots$$

が得られます。同様にして、 $\sin x$ の $x = 0$ における Taylor 展開が

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^l}{(2l+1)!} x^{2l+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots$$

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。



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

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(\alpha)}{k!} (x-\alpha)^k = f(\alpha) + f'(\alpha)(x-\alpha) + \frac{f''(\alpha)}{2!}(x-\alpha)^2 + \dots + \frac{f^{(n)}(\alpha)}{n!}(x-\alpha)^n$$

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 $f(x) \in \text{Ker } \Phi$ 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 $(x-\alpha)^2$ で割り切れるごとに $f(\alpha) = f'(\alpha) = 0$ 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、 $\{(x-\alpha)^m\}^{(m)} = m!$ 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 $\text{Ker } \Phi = \text{Im } \Psi$ 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 $\text{Ker } \Psi$ が 0 ということは、線型写像 $0 \rightarrow \mathbb{R}[x]$ の像と $\text{Ker } \Psi$ とが一致することと同じです。つまり $0 \rightarrow \mathbb{R}[x] \xrightarrow{\Phi} \mathbb{R}$ 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 $\mathbb{R} \rightarrow 0$ の核は \mathbb{R} 全体ですから、 ev_α の全射性^{*13}は、 $\text{Im } \text{ev}_\alpha$ と $\mathbb{R} \rightarrow 0$ の核が一致することと言えます。つまり $\mathbb{R}[x] \xrightarrow{\text{ev}_\alpha} \mathbb{R} \rightarrow 0$ 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、

$$0 \longrightarrow \mathbb{R}[x] \xrightarrow{\Psi} \mathbb{R}[x] \xrightarrow{\text{ev}_\alpha} \mathbb{R} \longrightarrow 0$$

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 $\text{ev}_\alpha(a) = a$ なので、 ev_α は全射です。

線形代数学 第3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かったときは、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 $(x, y) = ((5z - 6)/7, (11z - 2)/7)$

問6 (1) $(x, y, z) = (0, 0, 0)$ (2) z, w は任意で $(x, y) = ((3z - 7w)/5, (11z - 4w)/5)$

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 , 盜人の人数を 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 $y = k + (x - (k - 1)y - k)/7$ と書けるので、この両辺を $\sum_{k=1}^n$ すると

$$ny = \sum_{k=1}^n \left(k + \frac{x - (k - 1)y - k}{7} \right) = \sum_{k=1}^n \left(\frac{x + y}{7} + \frac{6 - y}{7} k \right) = \frac{n(x + y)}{7} + \frac{6 - y}{7} \frac{n(n + 1)}{2}$$

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ってしまうものの、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 卷の 90 ページに、狐狸と鵬鳥の問題が載っています。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「絹盜人を知る算」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「きぬねす人を知る事」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的の場合」を考えるので「こどもの遊び」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 变数の 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。

$$\begin{cases} x + 2y = 3 & \xrightarrow{\text{式 2 に式 1 の}} \\ 2x + 3y = 4 & \xrightarrow{\text{(-2) 倍を加える}} \end{cases} \begin{cases} x + 2y = 3 & \xrightarrow{\text{式 1 に式 2 の}} \\ -y = -2 & \xrightarrow{\text{2 倍を加える}} \end{cases} \begin{cases} x = -1 & \xrightarrow{\text{式 2 を}} \\ -y = -2 & \xrightarrow{\text{(-1) 倍する}} \end{cases} \begin{cases} x = -1 & \\ y = 2 & \end{cases}$$

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 $Ax = b$ と表せば、「係数だけ書き出す」という操作は、 A と b とを横に並べてできる行列 $\tilde{A} := (A | b)$ を取り出す操作に相当します。

一般的の場合も、連立 1 次方程式は必ず $Ax = b$ と書けます。この方程式 $Ax = b$ に対し、 A を係数行列といいます。また A と b とを横に並べてできる行列 $\tilde{A} := (A | b)$ 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}。

$$\begin{array}{l} \begin{cases} x + 2y = 3 & \xrightarrow{\text{式 2 に式 1 の}} \\ 3x + 4y = 5 & \xrightarrow{\text{(-3) 倍を加える}} \end{cases} \begin{cases} x + 2y = 3 & \xrightarrow{\text{式 1 に式 2 の}} \\ -2y = -4 & \xrightarrow{\text{1 倍を加える}} \end{cases} \begin{cases} x = -1 & \xrightarrow{\text{式 2 を}} \\ -2y = -4 & \xrightarrow{\text{(-1/2) 倍する}} \end{cases} \begin{cases} x = -1 & \\ y = 2 & \end{cases} \\ \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 3 & 4 & 5 \end{array} \right) \xrightarrow{\substack{\text{2 行目に 1 行目の} \\ \text{(-3) 倍を加える}}} \left(\begin{array}{cc|c} 1 & 2 & 3 \\ 0 & -2 & -4 \end{array} \right) \xrightarrow{\substack{\text{1 行目に 2 行目の} \\ \text{1 倍を加える}}} \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & -2 & -4 \end{array} \right) \xrightarrow{\substack{\text{2 行目を} \\ \text{(-1/2) 倍する}}} \left(\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 2 \end{array} \right) \end{array} \quad \blacksquare$$

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。もっと变数が多い場合でも、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、解が求まることになります。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

といいう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 (pivoting) 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。

$$\left(\begin{array}{ccc|c} 1 & -2 & -4 & 3 \\ 2 & -4 & -7 & 5 \\ 4 & -7 & -13 & 10 \end{array} \right) \xrightarrow{\substack{(1, 1) \text{ 成分を要に} \\ 2, 3 \text{ 行目を掃き出す}}} \left(\begin{array}{ccc|c} 1 & -2 & -4 & 3 \\ 0 & 0 & 1 & -1 \\ 0 & 1 & 3 & -2 \end{array} \right) \xrightarrow{\substack{2 \text{ 行目と } 3 \text{ 行目を入れ替える}} \substack{}} \left(\begin{array}{ccc|c} 1 & -2 & -4 & 3 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 \end{array} \right)$$

$$\xrightarrow{\substack{(2, 2) \text{ 成分を要に} \\ 1 \text{ 行目を掃き出す}}} \left(\begin{array}{ccc|c} 1 & 0 & 2 & -1 \\ 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 \end{array} \right) \xrightarrow{\substack{(3, 3) \text{ 成分を要に} \\ \text{他の行を掃き出す}}} \left(\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array} \right)$$

■

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。

$$\left(\begin{array}{cccc} 2 & 1 & -3 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{array} \right) \xrightarrow[\text{2倍を加える}]{\text{1行目に3行目の}} \left(\begin{array}{cccc} 1 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \left(\begin{array}{cccc} 2 & 1 & -3 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{array} \right) = \left(\begin{array}{cccc} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{array} \right)$$

$$\xrightarrow[\text{(-1)倍する}]{\text{3行目を}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{array} \right) \left(\begin{array}{cccc} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ -1 & 3 & -4 & 0 \end{array} \right) = \left(\begin{array}{cccc} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ 1 & -3 & 4 & 0 \end{array} \right)$$

$$\xrightarrow[\text{3倍を加える}]{\text{2行目に3行目の}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \left(\begin{array}{cccc} 0 & 7 & -11 & -2 \\ -3 & 2 & -1 & 2 \\ 1 & -3 & 4 & 0 \end{array} \right) = \left(\begin{array}{cccc} 0 & 7 & -11 & -2 \\ 0 & -7 & 11 & 2 \\ 1 & -3 & 4 & 0 \end{array} \right)$$

$$\xrightarrow[\text{入れ替える}]{\text{1行目と3行目を}} \left(\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array} \right) \left(\begin{array}{cccc} 0 & 7 & -11 & -2 \\ 0 & -7 & 11 & 2 \\ 1 & -3 & 4 & 0 \end{array} \right) = \left(\begin{array}{cccc} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 7 & -11 & -2 \end{array} \right)$$

$$\xrightarrow[\text{1倍を加える}]{\text{3行目に2行目の}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array} \right) \left(\begin{array}{cccc} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 7 & -11 & -2 \end{array} \right) = \left(\begin{array}{cccc} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\xrightarrow[\text{(-1/7)倍する}]{\text{2行目を}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{7} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \left(\begin{array}{cccc} 1 & -3 & 4 & 0 \\ 0 & -7 & 11 & 2 \\ 0 & 0 & 0 & 0 \end{array} \right) = \left(\begin{array}{cccc} 1 & -3 & 4 & 0 \\ 0 & 1 & -\frac{11}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & 0 \end{array} \right)$$

$$\xrightarrow[\text{3倍を加える}]{\text{1行目に2行目の}} \left(\begin{array}{cccc} 1 & 3 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \left(\begin{array}{cccc} 1 & -3 & 4 & 0 \\ 0 & 1 & -\frac{11}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & 0 \end{array} \right) = \left(\begin{array}{cccc} 1 & 0 & -\frac{5}{7} & -\frac{6}{7} \\ 0 & 1 & -\frac{11}{7} & -\frac{2}{7} \\ 0 & 0 & 0 & 0 \end{array} \right)$$

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

$$P_{i,j} := \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & 1 \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \\ j^{\text{th}} & & & 1 & & 0 & \\ & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix}$$

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

$$M_i(c) := \underset{i^{\text{th}}}{\left(\begin{array}{cccc} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & c \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{array} \right)}, \quad Q_{i,j}(c) := \underset{i^{\text{th}}}{\left(\begin{array}{ccccc} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & c & \\ & & & & \ddots \\ & & & & & 1 \\ & & & & & & \ddots \\ & & & & & & & 1 \end{array} \right)}$$

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、係数行列が正方形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 次正方形 $A, B, C, D \in \text{Mat}_2(\mathbb{R})$ 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 次正方形が作れます。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 $A = (A_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$, $B = (B_{ij})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq l}}$ 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形が並ぶ場合」ですが、他にも知つておくと、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 正方形 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形 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, \dots, 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 … 1 の個数は、時と場合によって変わります。点を並べてくせに 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。本によって授業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 $(A | E)$ 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

$$X_k X_{k-1} \cdots X_1 (A | E) = (X_k X_{k-1} \cdots X_1 A | X_k X_{k-1} \cdots X_1 E) = (E | X_k X_{k-1} \cdots X_1)$$

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。

$$\begin{array}{c} \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ -2 & 4 & 19 & 0 & 1 & 0 \\ -2 & 5 & 25 & 0 & 0 & 1 \end{array} \right) \\ \xrightarrow{(1,1) \text{ 成分を要に}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & -9 \\ 2 & 1 & 0 & -2 & 4 & 19 \\ 2 & 0 & 1 & -2 & 5 & 25 \end{array} \right) \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2 & 1 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \end{array} \right) \\ \xrightarrow[入れ替える]{1 \text{ 行目と } 2 \text{ 行目を}} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & -2 & -9 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 7 \end{array} \right) \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right) \\ \xrightarrow[1 \text{ 行目を掃き出す}]{(2,2) \text{ 成分を要に}} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 1 & -2 & -9 \\ 0 & 1 & 0 & 0 & 1 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 1 & 0 & 0 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & 0 & 5 & 5 & 0 & 2 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right) \\ \xrightarrow[1, 2 \text{ 行目を掃き出す}]{(3,3) \text{ 成分を要に}} \left(\begin{array}{ccc|ccc} 1 & 0 & -5 & 1 & 0 & 5 \\ 0 & 1 & -7 & 0 & 1 & 7 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array} \right) \left(\begin{array}{ccc|ccc} 1 & -2 & -9 & 5 & 0 & 2 \\ 0 & 1 & 7 & 2 & 0 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right) = \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -5 & -5 & 2 \\ 0 & 1 & 0 & -12 & -7 & 1 \\ 0 & 0 & 1 & 2 & 1 & 0 \end{array} \right) \end{array}$$

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

$$\left(\begin{array}{ccc|cc} -5 & -5 & 2 & 1 & -2 \\ -12 & -7 & 1 & -2 & 4 \\ 2 & 1 & 0 & -2 & 5 \end{array} \right) \left(\begin{array}{ccccc} 1 & -2 & -9 & 5 & 0 \\ -2 & 4 & 19 & -11 & 0 \\ -2 & 5 & 25 & -14 & 0 \end{array} \right) = \left(\begin{array}{ccccc} 1 & 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{array} \right)$$

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。一般的な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

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & c \\ & & & & 1 \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \frac{1}{c} \\ & & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 1 \\ & & & & 1 \end{pmatrix}$$

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

$$\begin{pmatrix} 1 & 3 & 2 & -1 \\ -3 & -8 & -4 & 0 \\ 2 & 5 & 3 & -1 \end{pmatrix} \xrightarrow{\substack{(1,1) \text{ 成分を要に} \\ 2,3 \text{ 行目を掃き出す}} \left(\begin{array}{cccc} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & -1 & -1 & 1 \end{array} \right)} \xrightarrow{\substack{(2,2) \text{ 成分を要に} \\ 3 \text{ 行目だけを掃き出す}} \left(\begin{array}{cccc} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{array} \right)}$$

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

$$\begin{pmatrix} 1 & 3 & 2 & -1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{pmatrix} \xrightarrow{\substack{(2,2) \text{ 成分を要に} \\ 1 \text{ 行目を掃き出す}} \left(\begin{array}{cccc} 1 & 0 & -4 & 8 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 1 & -2 \end{array} \right)} \xrightarrow{\substack{(3,3) \text{ 成分を要に} \\ 1,2 \text{ 行目を掃き出す}} \left(\begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -2 \end{array} \right)}$$

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 (Lower triangular) 行列と上三角 (Upper 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}

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 &= b_3 \end{aligned}$$

^{*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。

を、いつものように行列で $Ax = b$ 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 a_1, a_2, a_3 とし、行列 A を $A = (a_1 \ a_2 \ a_3)$ と表します。そうすると

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix} = x_1a_1 + x_2a_2 + x_3a_3$$

となります。ここで一旦方程式のことを忘れ、 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 Ax 全部を集めてできる集合は、 A の像 $\text{Im } A$ に他なりません。一方 $x_1a_1 + x_2a_2 + x_3a_3$ で x_1, x_2, x_3 を動かしてできるのは、ベクトル a_1, a_2, a_3 の張る空間です。ですから行列 A の像是、 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 $\text{Im } A = \mathbb{R}a_1 + \mathbb{R}a_2 + \mathbb{R}a_3$ です。

そして方程式 $Ax = b$ に話を戻すと、この方程式が解を持つことは、 b が A の像に入ることに他なりません。一方 $\text{Im } A = \mathbb{R}a_1 + \mathbb{R}a_2 + \mathbb{R}a_3$ ですから、これらを合わせて考えると「像 $\text{Im } A$ が大きいほど、方程式 $Ax = b$ は解を持ちやすくなる」と想像ができます。

また a_1, a_2, a_3 が $\text{Im } A$ を張るときに「無駄」がある場合もあります。もし a_1, a_2, a_3 が 1 次独立でなければ、 $a_1a_1 + a_2a_2 + a_3a_3 = \mathbf{0}$ 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 $\alpha_1a_1 + \alpha_2a_2 + \alpha_3a_3 = \mathbf{0}$ 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、 $\text{Im } A$ 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 $A = (a_1 \ a_2 \ \cdots \ a_n)$ と表します。このとき列ベクトル a_1, a_2, \dots, 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、 $\text{rank } A$ で表します。

後で線型空間の「次元」を定義すると、 $\text{rank } A$ は線型空間 $\text{Im } A$ 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。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 $\text{rank } A$ の大小は $\text{Im } A$ の次元の大小と同じで、したがって $\text{Im } A$ 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、 $v_1, \dots, v_k \in \mathbb{R}^n$ をベクトルとするとき、 X は逆行列を持つので、

$$\begin{aligned} \alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_kv_k &= \mathbf{0} \\ X(\alpha_1v_1 + \alpha_2v_2 + \cdots + \alpha_kv_k) &= \mathbf{0} \end{aligned}$$

という 2 式が同値になります。一方、2 つ目の式は $\alpha_1(Xv_1) + \alpha_2(Xv_2) + \cdots + \alpha_k(Xv_k) = \mathbf{0}$ と書き直せます。だから v_1, v_2, \dots, v_k の 1 次独立性と Xv_1, Xv_2, \dots, Xv_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 $Ax = b$ を考えましょう。拡大係数行列を $\tilde{A} = (A \ b)$ 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、 $\text{rank } \tilde{A} = \text{rank } A$ または $\text{rank } \tilde{A} = \text{rank } A + 1$ のいずれかが成り立ちます。一方 $b \in \text{Im } A$ かそうでないかが、 A の列ベクトルたちに 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 $\text{rank } \tilde{A} = \text{rank } A + 1$ なら $b \notin \text{Im } A$ であり、方程式 $Ax = b$ が解を持たないと結論付けることができます。逆に $\text{rank } \tilde{A} = \text{rank } A$ なら $b \in \text{Im } A$ で、方程式 $Ax = b$ 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

$$\begin{pmatrix} 3 & 1 & 5 & -1 \\ 2 & 1 & 6 & 0 \\ 5 & 2 & 11 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 1 & 5 & -1 \\ 2 & 1 & 6 & 0 \\ 2 & 1 & 6 & 3 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 1 & 5 & -1 \\ 2 & 1 & 6 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

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 $b \notin \text{Im } A$ を意味するので、方程式 $Ax = b$ 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 $Ax = b$ が複数の解を持つとき、そのうち異なる 2 個を x_1, x_2 とおくと、 $Ax_1 = Ax_2 = b$ です。したがってベクトルを A 倍する写像は単射ではありません。そして A 倍写像の単射でない場合は $\text{Ker } A$ を調べれば分かるのでした。特に $b = 0$ の場合、方程式を解くことは $\text{Ker } A$ 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 $Ax = 0$ と表す。右辺は 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。

$$\begin{pmatrix} 1 & 2 & -5 & 3 \\ 2 & -1 & 1 & 2 \\ 3 & 1 & -4 & 5 \\ 2 & 4 & -10 & 6 \end{pmatrix} \xrightarrow[\text{他の行を掃き出し}]{(1, 1) \text{ 成分を要に}} \begin{pmatrix} 1 & 2 & -5 & 3 \\ 0 & -5 & 11 & -4 \\ 0 & -5 & 11 & -4 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{2 行目を引く}]{3 \text{ 行目から}} \begin{pmatrix} 1 & 2 & -5 & 3 \\ 0 & -5 & 11 & -4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\xrightarrow[\text{(-1/5) 倍する}]{2 \text{ 行目を}} \begin{pmatrix} 1 & 2 & -5 & 3 \\ 0 & 1 & -\frac{11}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{1 行目を掃き出す}]{(2, 2) \text{ 成分を要に}} \begin{pmatrix} 1 & 0 & -\frac{3}{5} & \frac{7}{5} \\ 0 & 1 & -\frac{11}{5} & \frac{4}{5} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

変形後の係数行列の第1列から第4列を、それぞれ a_1, a_2, a_3, a_4 と表す。そうすると

$$a_3 = -\frac{3}{5}a_1 - \frac{11}{5}a_2, a_4 = \frac{7}{5}a_1 + \frac{4}{5}a_2$$

である。よって

$$v_1 := \begin{pmatrix} 3/5 \\ 11/5 \\ 1 \\ 0 \end{pmatrix}, v_2 := \begin{pmatrix} -7/5 \\ -4/5 \\ 0 \\ 1 \end{pmatrix} \in \text{Ker } A$$

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、 $\text{Ker } A$ 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 $v_3 = {}^t(u_1, u_2, u_3, u_4)$ があったとする。 v_1 と v_2 の1次結合を v_3 に足しても、 v_1, v_2 と v_3 が1次独立であることに変わりはない。したがって

$$v_3 - u_3 v_1 - u_4 v_2 = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} - \begin{pmatrix} 3/5 \\ 11/5 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} -7/5 \\ -4/5 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix}$$

と v_1, v_2 も一次独立なベクトルの組である^{*16}。しかし、これが $\text{Ker } A$ に属するということは

$$\mathbf{0} = A \begin{pmatrix} * \\ * \\ 0 \\ 0 \end{pmatrix} = *a_1 + *a_2$$

を意味する。 a_1, a_2 の係数の少なくとも一方は0でないから、 a_1 と a_2 は1次独立ではない。しかし $a_1 = {}^t(1, 0, 0, 0)$, $a_2 = {}^t(0, 1, 0, 0)$ 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 $\text{Ker } A$ の元は、全て v_1 と v_2 の1次結合で書ける。これで解が

$$zv_1 + wv_2 = \begin{pmatrix} (3z - 7w)/5 \\ (11z - 4w)/5 \\ z \\ w \end{pmatrix} \quad (z, w \in \mathbb{R})$$

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、 $\text{Ker } A$ のベクトルとが1:1に対応する
- $\text{Ker } A$ 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 第4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 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、 $v_1, \dots, v_n \in V$ とします。 v_1, \dots, v_n が1次独立であるとは $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n = \mathbf{0}$ となる v_1, \dots, v_n の1次結合^{*1}が $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ 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 $v_i \notin \mathbb{R}v_1 + \dots + \hat{\mathbb{R}v_i} + \dots + \mathbb{R}v_n$ が成立^{*2}ことと同値です。

^{*1} $\alpha_1 v_1 + \alpha_2 v_2 + \dots + \alpha_n v_n$ の形の式を「 v_1, \dots, 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。

^{*2} ここで $\hat{\mathbb{R}v_i}$ と書いたのは「 $\mathbb{R}v_i$ を除く」という意味です。一々 $\mathbb{R}v_1 + \dots + \mathbb{R}v_{i-1} + \mathbb{R}v_{i+1} + \dots + \mathbb{R}v_n$ と書くと長くて大変なので、し

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、 $x \rightarrow +0$ の極限を取って $a_1 = 0$ を得る。同様に $f''(x), f'''(x)$ の $x \rightarrow +0$ 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。ここで 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。

生成系 V を線型空間、 $v_1, v_2, \dots, v_n \in V$ とします。 V の全ての元が v_1, v_2, \dots, v_n の1次結合で表せるとき、 v_1, v_2, \dots, v_n は V を生成するといいます。記号で書けば、 $V = \mathbb{R}v_1 + \mathbb{R}v_2 + \dots + \mathbb{R}v_n$ 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、 (f_1, f_2, \dots, f_n) をその要素の列とする。

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。

(1) (f_1, f_2, \dots, f_n) が 1 次独立とする。このとき $1 \leq i_1 < i_2 < \dots < i_k \leq n$ とし、部分列 $(f_{i_1}, f_{i_2}, \dots, f_{i_k})$ を考える^{*5}と、これも 1 次独立となる。実際 $\alpha_1 f_{i_1} + \alpha_2 f_{i_2} + \dots + \alpha_k f_{i_k} = \mathbf{0}$ とおくと、この両辺に i_1, i_2, \dots, 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 (f_1, f_2, \dots, f_n) の 1 次独立性から、 $\alpha_1 = \alpha_2 = \dots = \alpha_k = 0$ が従う。

(2) (f_1, f_2, \dots, f_n) が V を生成したとする。このとき勝手な $v \in V$ を足した列 $(f_1, f_2, \dots, f_n, v)$ も V を生成する。実際、全てのベクトルは (f_1, f_2, \dots, f_n) の 1 次結合で表せる。その表式の両辺に $0v = \mathbf{0}$ を足した式を考えれば、 $(f_1, f_2, \dots, f_n, v)$ 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) (f_1, f_2, \dots, 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、 $f_j \notin \mathbb{R}f_1 + \dots + \mathbb{R}\hat{f}_j + \dots + \mathbb{R}f_j$ である。ここで $1 \leq i_1 < i_2 < \dots < i_k \leq n$ とし、部分列 $(f_{i_1}, f_{i_2}, \dots, f_{i_k})$ を考える。すると任意の $1 \leq l \leq k$ に対し $\mathbb{R}f_{i_1} + \dots + \mathbb{R}\hat{f}_{i_l} + \dots + \mathbb{R}f_{i_k} \subset \mathbb{R}f_1 + \dots + \mathbb{R}\hat{f}_{i_l} + \dots + \mathbb{R}_n$ となるから、 $f_{i_l} \notin \mathbb{R}f_{i_1} + \dots + \mathbb{R}\hat{f}_{i_l} + \dots + \mathbb{R}f_{i_k}$ でないといけない。これで $(f_{i_1}, f_{i_2}, \dots, f_{i_k})$ の 1 次独立性が言えた。

(2) (f_1, f_2, \dots, f_n) が V を生成したとする。このとき $\mathbb{R}f_1 + \dots + \mathbb{R}f_n = V$ である。いま勝手に $v \in V$ を取ると、 $\mathbb{R}v \subset V$ である。よって $V = \mathbb{R}f_1 + \dots + \mathbb{R}f_n \subset \mathbb{R}f_1 + \dots + \mathbb{R}f_n + \mathbb{R}v \subset V$ となり、 $V = \mathbb{R}f_1 + \dots + \mathbb{R}f_n + \mathbb{R}v$ が従う。よって $(f_1, f_2, \dots, f_n, v)$ 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、 v_1, \dots, v_n を V の元とします。このとき v_1, \dots, 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 (v_1, \dots, 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

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

の組 (e_1, \dots, 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。

問 12 の解答 勝手な \mathbb{R}^n の元は

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n$$

^{*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ください。

と書けるので、 e_1, \dots, e_n たちは \mathbb{R}^n を生成する。また

$$\mathbf{0} = \alpha_1 e_1 + \alpha_2 e_2 + \cdots + \alpha_n e_n = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

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 e_1, \dots, e_n たちは 1 次独立である。以上で (e_1, \dots, 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 e_1, \dots, e_n で表したときの e_i が、 i 番目の座標と全く同じでした。一般的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、 (v_1, \dots, 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 $x \in V$ が

$$x = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n$$

を満たすときに「 x の座標は ${}^t(x_1, x_2, \dots, x_n)$ 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 (v_1, \dots, v_n) は基底だったから、 v_1, \dots, v_n たちは V を張ります。だから全ての $x \in V$ を、 v_1, \dots, 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」という事態があつてはいけません。こちらの条件はどうでしょうか？ いま $x \in V$ が

$$x = x_1 v_1 + x_2 v_2 + \cdots + x_n v_n = x'_1 v_1 + x'_2 v_2 + \cdots + x'_n v_n$$

と、基底 (v_1, \dots, 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

$$(x_1 - x'_1)v_1 + (x_2 - x'_2)v_2 + \cdots + (x_n - x'_n)v_n = \mathbf{0}$$

を得ます。ここで v_1, \dots, 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 $x \in V$ に対し「 x を v_1, \dots, 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、 (v_1, \dots, v_n) が基底であるための条件が「全てのベクトルが v_1, \dots, 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、 (f_1, \dots, f_n) をその要素の列とする。このとき写像 $\varphi: \mathbb{R}^n \rightarrow V$ を、 $x = {}^t(x_1, x_2, \dots, x_n)$ に対し $\varphi(x) := x_1 f_1 + x_2 f_2 + \cdots + x_n f_n$ で定める。

(0) まず φ が線型写像であることを示しておく。

- $x = {}^t(x_1, x_2, \dots, x_n)$, $y = {}^t(y_1, y_2, \dots, y_n)$ に対し

$$\begin{aligned} \varphi(x+y) &= \varphi({}^t(x_1+y_1, x_2+y_2, \dots, x_n+y_n)) = (x_1+y_1)f_1 + (x_2+y_2)f_2 + \cdots + (x_n+y_n)f_n \\ &= (x_1 f_1 + x_2 f_2 + \cdots + x_n f_n) + (y_1 f_1 + y_2 f_2 + \cdots + y_n f_n) = \varphi(x) + \varphi(y) \end{aligned}$$

- $x = {}^t(x_1, x_2, \dots, x_n)$, $\alpha \in \mathbb{R}$ に対し

$$\begin{aligned}\varphi(\alpha x) &= \varphi((\alpha x_1, \alpha x_2, \dots, \alpha x_n)) = \alpha x_1 f_1 + \alpha x_2 f_2 + \dots + \alpha x_n f_n \\ &= \alpha(x_1 f_1 + x_2 f_2 + \dots + x_n f_n) = \alpha \varphi(x)\end{aligned}$$

(1) (f_1, \dots, f_n) が 1 次独立であることは、 $\varphi(x) = \mathbf{0}$ を満たす $x = {}^t(x_1, x_2, \dots, x_n)$ が $x = \mathbf{0}$ 以外に存在しないことと同値である。よって (f_1, \dots, f_n) の 1 次独立性は $\text{Ker } \varphi = \{\mathbf{0}\}$ と同値である。そして $\text{Ker } \varphi = \{\mathbf{0}\}$ と φ の単射性が同値だったので、 (f_1, \dots, f_n) の 1 次独立性と φ の単射性が同値になる。

(2) (f_1, f_2, \dots, f_n) は $\text{Im } \varphi$ を生成する。一方 φ の全射性は、 $\text{Im } \varphi = V$ と同値である。よって (f_1, f_2, \dots, 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、 (f_1, \dots, f_n) が基底をなすことが同値」と分かれます。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 $\varphi: \mathbb{R}^n \rightarrow V$ 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 $x, y \in \mathbb{R}^n$ と $\alpha \in \mathbb{R}$ に対し $\varphi(x+y) = \varphi(x) + \varphi(y)$, $\varphi(\alpha x) = \alpha \varphi(x)$ が成立立ちます。つまり「 φ で V に送ってから足し算/スカラー倍をして、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れます。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 $\varphi: \mathbb{R}^n \xrightarrow{\sim} V$ 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 $f: V \rightarrow W$ 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 $\varphi: \mathbb{R}^n \rightarrow V$, $\psi: \mathbb{R}^m \rightarrow W$ とします。そうすると、次のような図式ができます。

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \varphi \uparrow & & \downarrow \psi \\ \mathbb{R}^n & & \mathbb{R}^m \end{array}$$

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

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \varphi \uparrow & & \downarrow \psi^{-1} \\ \mathbb{R}^n & \dashrightarrow & \mathbb{R}^m \end{array}$$

という図式が得られます。元々あったのは $f: V \rightarrow W$ 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、 $\psi^{-1} \circ f \circ \varphi: \mathbb{R}^n \rightarrow \mathbb{R}^m$ 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 $f: V \rightarrow W$ 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

⁷ この辺は記号の亜種が色々あるのですが、大体 $=$ とか \rightarrow に \sim がくっついてたら「同型」と思って差支えありません。

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。

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」をしていましたの^{*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繰り返される
- 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繰り返される

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、基底が得られることがあります。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、 (f_1, \dots, f_n) をその要素の列とする。

- (f_1, \dots, f_n) が1次独立だが、 V を生成しないとする。このとき $\mathbb{R}f_1 + \dots + \mathbb{R}f_n$ は V の真部分集合なので、 $f_{n+1} \in V$ で $f_{n+1} \notin \mathbb{R}f_1 + \dots + \mathbb{R}f_n$ を満たすものが取れる。このとき $(f_1, \dots, f_n, f_{n+1})$ は1次独立である。実際 $\alpha_1 f_1 + \dots + \alpha_n f_n + \alpha_{n+1} f_{n+1} = 0$ とおくと、 $\alpha_1 f_1 + \dots + \alpha_n f_n = -\alpha_{n+1} f_{n+1}$ である。もし $\alpha_{n+1} \neq 0$ なら $f_{n+1} = (\alpha_1 f_1 + \dots + \alpha_n f_n) / (-\alpha_{n+1})$ となってしまい、 f_{n+1} の取り方に矛盾する。よって $\alpha_{n+1} = 0$ で $\alpha_1 f_1 + \dots + \alpha_n f_n = 0$ が得られる。 f_1, \dots, f_n は元々1次独立だったので、結局 $\alpha_1 = \dots = \alpha_n = 0$ となる。
- (f_1, \dots, f_n) が V を生成するが、1次独立でないとする。このとき $\alpha_1 f_1 + \dots + \alpha_n f_n = 0$ となる実数の組 $(\alpha_1, \dots, \alpha_n)$ であって、少なくとも1つは0でないようなもののが存在する。ここで、 $\alpha_i \neq 0$ となる $1 \leq i \leq n$ を1つ選ぶと、 f_i を取り除いた列 $(f_1, \dots, \hat{f}_i, \dots, f_n)$ も V を生成する^{*10}。なぜなら $\alpha_i \neq 0$ より f_i は $f_i = -(\alpha_1 f_1 + \dots + \alpha_{i-1} f_{i-1} + \dots + \alpha_n f_n) / \alpha_i$ を満たす。よって f_i は $f_1, \dots, \hat{f}_i, \dots, f_n$ の1次結合で書けてしまうから、 $f_1, \dots, f_i, \dots, f_n$ の1次結合は 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、 (f_1, \dots, f_n) をその要素の列とする。

- (f_1, \dots, 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 $v \in V$ を取ると、 (f_1, \dots, f_n, v) は1次従属だから、 $\alpha_1 f_1 + \dots + \alpha_n f_n + \alpha_{n+1} f_{n+1} = 0$ となる実数の組

^{*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。どちら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。

^{*10} ここで f_i の上に¹⁰をつけて \hat{f}_i と書いたのは「 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。

$(\alpha_1, \dots, \alpha_{n+1})$ 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 (f_1, \dots, 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 $v = -(\alpha_1 f_1 + \dots + \alpha_n f_n)/\alpha_{n+1}$ となる。これは $v \in \mathbb{R}f_1 + \dots + \mathbb{R}f_n$ に他ならない。 v は任意だったので、 $\mathbb{R}f_1 + \dots + \mathbb{R}f_n = V$ である。つまり (f_1, \dots, f_n) は V を生成する。

(2) (f_1, \dots, 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 $\alpha_1 f_1 + \dots + \alpha_n f_n = 0$ とおく。もし $\alpha_i \neq 0$ となる i が存在すれば、 $f_i = -(\alpha_1 f_1 + \dots + \hat{\alpha_i f_i} + \dots + \alpha_n f_n)/\alpha_i$ となり、 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 f_i を取り除いた列 $(f_1, \dots, \hat{f}_i, \dots, f_n)$ 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、 (f_1, \dots, 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 (f_1, \dots, f_n) が 1 次独立な生成系とする。

(1) $v \in V$ を 0 でないベクトルとする。このとき (f_1, \dots, f_n) は V の生成系だから、 $v = \alpha_1 f_1 + \dots + \alpha_n f_n$ と表せる。よって (f_1, \dots, f_n, v) は 1 次独立でない。

(2) $1 \leq i \leq n$ とする。このとき $f_i = \alpha_1 f_1 + \dots + \hat{\alpha_i f_i} + \dots + \alpha_n f_n$ と表せたとすると、 (f_1, \dots, f_n) が 1 次独立であることに反する。よって (f_1, \dots, f_n) の中からどの f_i を引き抜いても、残りのベクトルの 1 次結合で f_i を表すことはできない。したがって (f_1, \dots, 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、 $f_1, f_2, \dots, f_n \in V$ とする。また g_1, g_2, \dots, g_n を f_1, f_2, \dots, f_n の並べ替えとし^{*11}、添字の対応を $g_i = f_{\sigma(i)}$ および $f_i = g_{\tau(i)}$ と表す。

いま f_1, f_2, \dots, f_n たちが 1 次独立だったとする。このとき $a_1 g_1 + \dots + a_n g_n = 0$ とおくと

$$0 = a_1 g_1 + \dots + a_n g_n = a_1 f_{\sigma(1)} + \dots + a_n f_{\sigma(n)}$$

である。ここで $\sigma(1), \dots, \sigma(n)$ は $1, \dots, n$ 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 f_1, \dots, f_n の 1 次結合になる。 f_1, \dots, 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 g_1, \dots, g_n も 1 次独立だと言えた。

また f_1, \dots, f_n が V を生成したとする。このとき勝手な $x \in V$ を取ると、 $x = x_1 f_1 + \dots + x_n f_n$ と表せる。すると f_i たちは g_j たちのどれかなので、 $x = x_1 g_{\tau(1)} + \dots + x_n g_{\tau(n)}$ と書ける。 x は任意なので、結局全ての $x \in V$ が g_1, \dots, g_n の 1 次結合で表せると言えた。■

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これら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。

*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 つの入れ替われば大丈夫なら、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。

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 e_1, e_2 に対し「 e_1 の方向から 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、 e_1, e_2, 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 $\text{Mat}_{m,n}(\mathbb{R})$ が線型空間であることを示す。

- $\text{Mat}_{m,n}(\mathbb{R})$ 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 $O \in \text{Mat}_{m,n}(\mathbb{R})$ がある
 - 任意の行列 $A \in \text{Mat}_{m,n}(\mathbb{R})$ 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。
- $\text{Mat}_{m,n}(\mathbb{R})$ には、スカラー倍が定義されている
 - 任意の $A \in \text{Mat}_{m,n}(\mathbb{R})$ と $\alpha, \beta \in \mathbb{R}$ に対し、 $\alpha(\beta A) = (\alpha\beta)A$ が成り立つ
 - 任意の $A \in \text{Mat}_{m,n}(\mathbb{R})$ 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 $A, B \in \text{Mat}_{m,n}(\mathbb{R})$ と $\alpha \in \mathbb{R}$ に対し、 $\alpha(A + B) = \alpha A + \alpha B$ が成り立つ
 - 任意の $A \in \text{Mat}_{m,n}(\mathbb{R})$ と $\alpha, \beta \in \mathbb{R}$ に対し、 $(\alpha + \beta)A = \alpha A + \beta A$ が成り立つ

次に、 $\text{Mat}_{m,n}(\mathbb{R})$ 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

$$O = \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} a_{ij} E_{ij} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

とおくと、全ての a_{ij} が 0 となる。また、勝手な行列 $A = (a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$ を取ると

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} = \sum_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} a_{ij} E_{ij}$$

は E_{ij} たちの 1 次結合で表されている。よって $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ の元を適当な順序で並べたものは、 $\text{Mat}_{m,n}(\mathbb{R})$ を生成する。これで $\text{Mat}_{m,n}(\mathbb{R})$ 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 $\text{Sym}_n(\mathbb{R})$ と書く^{*13}と、これは $\text{Mat}_n(\mathbb{R}) := \text{Mat}_{n,n}(\mathbb{R})$ の部分空間である^{*14}。実際

- $A, B \in \text{Sym}_n(\mathbb{R})$ のとき、 ${}^t(A + B) = {}^tA + {}^tB = A + B$ より $A + B \in \text{Sym}_n(\mathbb{R})$
- $A \in \text{Sym}_n(\mathbb{R}), \alpha \in \mathbb{R}$ のとき、 ${}^t(\alpha A) = \alpha {}^tA = \alpha A$ より $\alpha A \in \text{Sym}_n(\mathbb{R})$

となっている。そして $\text{Mat}_n(\mathbb{R})$ は n^2 次元の有限次元線型空間だから、その部分空間 $\text{Sym}_n(\mathbb{R})$ も有限次元である。また一般的の対称行列は

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{12} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \cdots & a_{nn} \end{pmatrix} = \sum_{i=1}^n a_{ii} E_{ii} + \sum_{1 \leq i < j \leq n} (E_{ij} + E_{ji})$$

の形に 1 通りに表せる。よって基底として $\{E_{ii} \mid 1 \leq i \leq n\} \cup \{E_{ij} + E_{ji} \mid 1 \leq i < j \leq n\}$ を適当な順に並べたものが取れる。これより $\text{Sym}_n(\mathbb{R})$ 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 $\text{Alt}_n(\mathbb{R})$ と書く^{*15}と、これは $\text{Mat}_n(\mathbb{R})$ の部分空間である。実際

- $A, B \in \text{Alt}_n(\mathbb{R})$ のとき、 ${}^t(A + B) = {}^tA + {}^tB = -A - B = -(A + B)$ より $A + B \in \text{Sym}_n(\mathbb{R})$
- $A \in \text{Alt}_n(\mathbb{R}), \alpha \in \mathbb{R}$ のとき、 ${}^t(\alpha A) = \alpha {}^tA = \alpha A$ より $\alpha A \in \text{Alt}_n(\mathbb{R})$

となっている。そして $\text{Mat}_n(\mathbb{R})$ は n^2 次元の有限次元線型空間だから、その部分空間 $\text{Alt}_n(\mathbb{R})$ も有限次元である。また、全ての交代行列は

$$\begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ -a_{12} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{1n} & -a_{2n} & \cdots & 0 \end{pmatrix} = \sum_{1 \leq i < j \leq n} (E_{ij} - E_{ji})$$

^{*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 $\text{Mat}_{m,n}(\mathbb{R}) \simeq \mathbb{R}^{mn}$ です。

^{*13} 対称行列のことを “symmetric matrix” といいます。 $\text{Sym}_n(\mathbb{R})$ 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 つの自然数で事足るので、しばしば $\text{Mat}_{n,n}(\mathbb{R})$ のことを $\text{Mat}_n(\mathbb{R})$ と略記します。

^{*15} 交代行列のことを “alternating matrix” といいます。 $\text{Alt}_n(\mathbb{R})$ の記号はここから取りました。

の形に 1 通りに表せる。よって具体的な基底として $\{E_{ij} - E_{ji} \mid 1 \leq i < j \leq n\}$ を適当な順に並べたものが取れる。よって $\text{Alt}_n(\mathbb{R})$ 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

$$X = \frac{X + {}^t X}{2} + \frac{X - {}^t X}{2}$$

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 $\text{Mat}_n(\mathbb{R})$ の元を、その部分空間 $\text{Sym}_n(\mathbb{R})$ と $\text{Alt}_n(\mathbb{R})$ 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「 $\text{Mat}_n(\mathbb{R})$ は、 $\text{Sym}_n(\mathbb{R})$ と $\text{Alt}_n(\mathbb{R})$ の直和である」といい、 $\text{Mat}_n(\mathbb{R}) = \text{Sym}_n(\mathbb{R}) \oplus \text{Alt}_n(\mathbb{R})$ 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 $\mathbb{R}[x]_{\leq n}$ 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、 $\mathbb{R}[x]_{\leq n}$ は和とスカラー倍で閉じている。よって $\mathbb{R}[x]_{\leq n}$ は $\mathbb{R}[x]$ の部分空間である。

$\mathbb{R}[x]_{\leq n}$ の基底としては $(1, x, x^2, \dots, x^n)$ が取れる。実際多項式の定義から、全ての多項式は $1, x, x^2, \dots, x^n$ の 1 次結合でただ一通りに書ける^{*17}。よって $\mathbb{R}[x]_{\leq n}$ 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、 $\mathbb{R}[x]_{\leq n}$ 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 $(1, x, x^2, \dots)$ が取れることは同じです。実際、全ての多項式は $1, x, x^2, \dots$ 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、 $1, x, x^2, \dots$ 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

*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。

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、 (f_1, f_2, \dots, f_m) と (g_1, g_2, \dots, g_n) が共に基底だったとします^{*18}。このとき各 f_i ($1 \leq i \leq m$) は基底 (g_1, g_2, \dots, g_n) を用いて

$$g_i = a_{1i}f_1 + a_{2i}f_2 + \cdots + a_{ni}f_n \quad (a_{1i}, a_{2i}, \dots, a_{ni} \in \mathbb{R})$$

と一意的に表せます。これを $i = 1, 2, \dots, n$ と並べてみると

$$g_1 = a_{11}f_1 + a_{21}f_2 + \cdots + a_{m1}f_m$$

$$g_2 = a_{12}f_1 + a_{22}f_2 + \cdots + a_{m2}f_m$$

⋮

$$g_n = a_{1n}f_1 + a_{2n}f_2 + \cdots + a_{mn}f_m$$

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

$$(g_1 \ g_2 \ \cdots \ g_n) = (f_1 \ f_2 \ \cdots \ f_m) \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

となっています。かくして、2つの基底の関係は $(g_1 \ g_2 \ \cdots \ g_n) = (f_1 \ f_2 \ \cdots \ f_m)A$ 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、 $(f_1 \ f_2 \ \cdots \ f_m)$ から $(g_1 \ g_2 \ \cdots \ g_n)$ への基底の変換行列と、その逆向きの $(g_1 \ g_2 \ \cdots \ g_n)$ から $(f_1 \ f_2 \ \cdots \ f_m)$ への基底の変換行列両方を考えます。つまり

$$(g_1 \ g_2 \ \cdots \ g_n) = (f_1 \ f_2 \ \cdots \ f_m)A, \quad (f_1 \ f_2 \ \cdots \ f_m) = (g_1 \ g_2 \ \cdots \ g_n)B$$

とおきます。そうすると、基底を2度変換することで

$$(f_1 \ f_2 \ \cdots \ f_m) = (g_1 \ g_2 \ \cdots \ g_n)B = (f_1 \ f_2 \ \cdots \ f_m)AB$$

が従います。これは「基底 (f_1, f_2, \dots, f_n) を基底 (f_1, f_2, \dots, 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 $f_i = f_i$ 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 $(g_1 \ g_2 \ \cdots \ g_n)$ 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、 $m = \text{tr } E_m = \text{tr } AB = \text{tr } BA = \text{tr } E_n = n$ 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。

^{*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。

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。

$\varphi: U \rightarrow V$ を線型写像とします。まず φ の単射性と $\text{Ker } \varphi = \{0\}$ とが同値でした。そして $\text{Ker } \varphi = \{0\}$ と $\dim \text{Ker } \varphi = 0$ とが同値なので、結局 φ の単射性は $\dim \text{Ker } \varphi = 0$ と言い換えられます。また φ の全射性の定義は $\text{Im } \varphi = V$ です。一般に $\text{Im } \varphi \subset V$ が成り立つから $\text{Im } \varphi$ の基底を延長して V の基底が作れるので、 $\text{Im } \varphi = V$ が成り立つことは $\text{Im } \varphi$ の基底が V の基底にもなること、つまり $\dim \text{Im } \varphi = \dim V$ 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、 $\varphi: U \rightarrow V$ を線型写像、 $f_1, f_2, \dots, f_n \in U$ とする。

(1) $(\varphi(f_1), \varphi(f_2), \dots, \varphi(f_n))$ が1次独立とする。このとき $\alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_n f_n = 0$ とおくと、両辺に φ を施して

$$0 = \varphi(\alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_n f_n) = \alpha_1 \varphi(f_1) + \alpha_2 \varphi(f_2) + \dots + \alpha_n \varphi(f_n)$$

を得る。 $(\varphi(f_1), \varphi(f_2), \dots, \varphi(f_n))$ の1次独立性から $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$ が従う。よって (f_1, f_2, \dots, f_n) も1次独立である。

(2) f_1, f_2, \dots, f_n が U を生成したとする。任意に $v \in \text{Im } \varphi$ を取ると、 $v = \varphi(x)$ となる $x \in U$ が取れる。 f_1, f_2, \dots, f_n は U を生成するので、 $x = \alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_n f_n$ と表せる。すると

$$v = \varphi(x) = \varphi(\alpha_1 f_1 + \alpha_2 f_2 + \dots + \alpha_n f_n) = \alpha_1 \varphi(f_1) + \alpha_2 \varphi(f_2) + \dots + \alpha_n \varphi(f_n)$$

となり、 v は $\varphi(f_1), \varphi(f_2), \dots, \varphi(f_n)$ の1次結合で表せる。 $v \in \text{Im } \varphi$ は任意だったので、 $\varphi(f_1), \varphi(f_2), \dots, \varphi(f_n)$ は $\text{Im } \varphi$ 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、 $\varphi: U \rightarrow V$ を線型写像とします。そして $\dim U = n$, $\dim V = m$ とし、 (f_1, f_2, \dots, f_n) を U の基底とします。

まず(1)を見ましょう。(1)は「 $\text{Im } \varphi$ 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 $\dim \text{Im } \varphi \leq \dim V$ が成り立ちます。

(2)も同様です。 (f_1, f_2, \dots, 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、 $(\varphi(f_1), \varphi(f_2), \dots, \varphi(f_n))$ は $\text{Im } \varphi$ を生成します。つまり $\text{Im } \varphi$ 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、ここまで議論でやはり $\dim \text{Im } \varphi \leq n = \dim U$ 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 $\dim \text{Im } \varphi < \dim V$ となりますから、 φ は全射にはなりません。

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 $\varphi: U \rightarrow V$ を線型写像とします。このとき常に $\dim \text{Ker } \varphi + \dim \text{Im } \varphi = \dim U$ 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「 $\text{Ker } \varphi$ 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 $\text{Ker } \varphi$ 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 $\text{Im } \varphi$ 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。 $\text{Ker } \varphi$ は U の部分空間なので、 $\text{Ker } \varphi$ の基底 (f_1, \dots, f_k) を延長して U 全体の基底 $(f_1, \dots, f_k, f_{k+1}, \dots, f_n)$ を作れます。そうすると f_{k+1}, \dots, f_n は $\text{Ker } \varphi = \mathbb{R}f_1 + \dots + \mathbb{R}f_k$ に入らないので、 $\varphi(f_{k+1}), \dots, \varphi(f_n)$ はいずれも 0 にはなりません。

次に $\varphi(\mathbf{f}_{k+1}), \dots, \varphi(\mathbf{f}_n)$ が 1 次独立なことを示します。 $\alpha_1\varphi(\mathbf{f}_{k+1}) + \alpha_2\varphi(\mathbf{f}_{k+2}) + \dots + \alpha_{n-k}\varphi(\mathbf{f}_n) = \mathbf{0}$ とおくと $\varphi(\alpha_1\mathbf{f}_{k+1} + \alpha_2\mathbf{f}_{k+2} + \dots + \alpha_{n-k}\mathbf{f}_n) = \mathbf{0}$ となります。よって $\alpha_1\mathbf{f}_{k+1} + \alpha_2\mathbf{f}_{k+2} + \dots + \alpha_{n-k}\mathbf{f}_n \in \text{Ker } \varphi$ ですが、 $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ が $\text{Ker } \varphi$ の基底だったので、 $\alpha_1\mathbf{f}_{k+1} + \alpha_2\mathbf{f}_{k+2} + \dots + \alpha_{n-k}\mathbf{f}_n = \mathbf{0}$ が従います。よって $\mathbf{f}_{k+1}, \dots, \mathbf{f}_n$ の 1 次独立性から $\alpha_1 = \dots = \alpha_{n-k} = 0$ となります。

最後に $\varphi(\mathbf{f}_{k+1}), \dots, \varphi(\mathbf{f}_n)$ が $\text{Im } \varphi$ を生成することを示します。任意に $\mathbf{v} \in \text{Im } \varphi$ を取ると、 $\mathbf{v} = \varphi(\mathbf{u})$ ($\mathbf{u} \in U$) と書けます。そして $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ が U の基底なので、 $\mathbf{u} = \alpha_1\mathbf{f}_1 + \dots + \alpha_n\mathbf{f}_n$ と書けます。ところが $\mathbf{f}_1, \dots, \mathbf{f}_k \in \text{Ker } \varphi$ だから $\mathbf{v} = \varphi(\alpha_1\mathbf{f}_1 + \dots + \alpha_n\mathbf{f}_n) = \alpha_{k+1}\varphi(\mathbf{f}_{k+1}) + \dots + \alpha_n\varphi(\mathbf{f}_n)$ となります。よって \mathbf{u} は $\varphi(\mathbf{f}_{k+1}), \dots, \varphi(\mathbf{f}_n)$ の 1 次結合で表せます。 \mathbf{u} は任意だったので、これで $\varphi(\mathbf{f}_{k+1}), \dots, \varphi(\mathbf{f}_n)$ が $\text{Im } U$ を生成することが言えました。

これらをまとめると、 $(\mathbf{f}_1, \dots, \mathbf{f}_n)$ の最初の k 本が $\text{Ker } \varphi$ の基底で、残りの $k+1$ 本に φ を施すと $\text{Im } \varphi$ 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、 $\dim \text{Ker } \varphi + \dim \text{Im } \varphi = \dim U$ 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。

$\varphi: U \rightarrow V$ を線型写像とします。次元定理から $\dim \text{Im } \varphi = \dim U - \dim \text{Ker } \varphi \leq \dim U$ 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 $\dim \text{Ker } \varphi = \dim U - \dim \text{Im } \varphi \geq \dim U - \dim V$ なので、もし $\dim U > \dim V$ であれば $\dim \text{Ker } \varphi \geq 1$ 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 $\dim \text{Ker } \varphi + \dim \text{Im } \varphi = \dim V$ が成り立つので、 $\dim \text{Ker } \varphi = 0$ と $\dim \text{Im } \varphi = \dim V$ 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度があるのです^{*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「 $\text{Ker } f$ 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 $\text{Im } f$ 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）。