

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6月9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 数ベクトルに対する線型写像	47
2 線型空間	49

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。

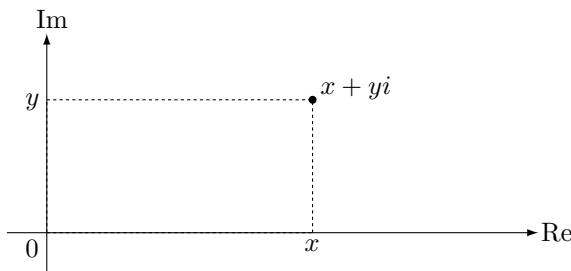


図 1.1 複素平面における数と点の対応

^{*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

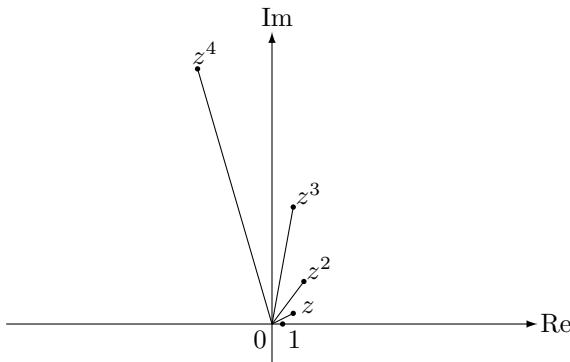
$$z^2 = 3 + 4i, \quad z^3 = 2 + 11i, \quad z^4 = -7 + 24i, \quad z^5 = -38 + 41i, \quad z^6 = -117 + 44i$$

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。

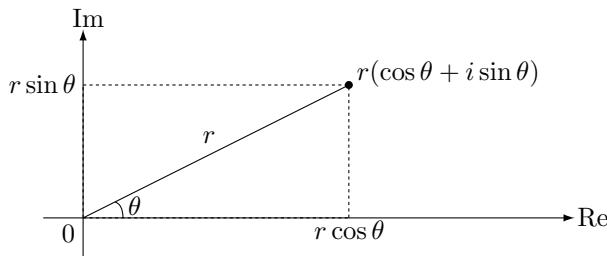


図 1.3 複素数の極形式表示

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、適當な複素数 $z_1 \in \mathbb{C}$ 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。

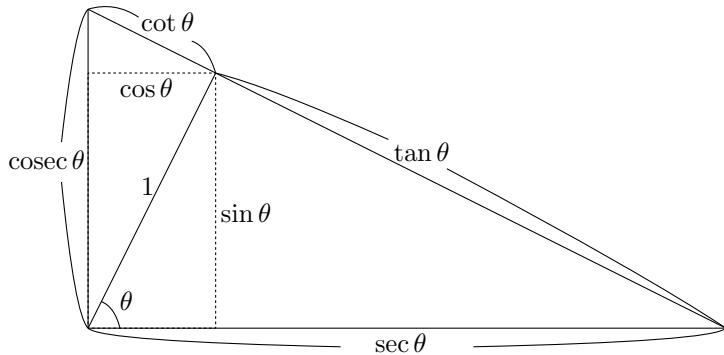
3 逆三角函数と円周率の近似公式

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。单位円周上の点 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。

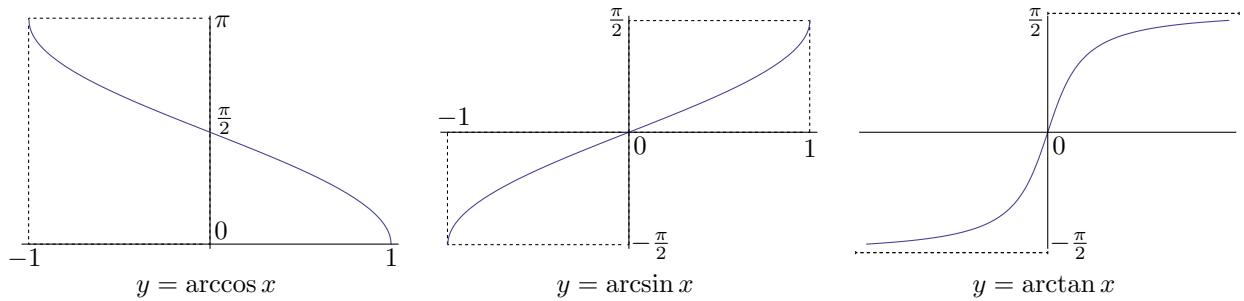
$\sec \theta := 1/\cos \theta$, $\cosec \theta := 1/\sin \theta$ と定めます。これを図に描くと次のようになります。



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ます。 $\arctan 1$ 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。

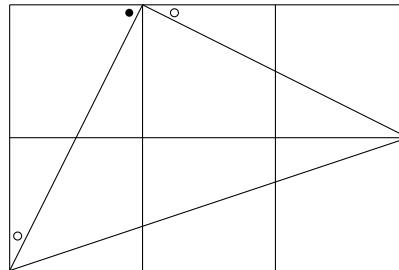
$$(1) (3+i)(7+i) = 20 + 10i, \quad (2) (2+i)(3+i) = 5 + 5i, \quad (3) (5+i)^4/(239+i) = 2 + 2i$$

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。

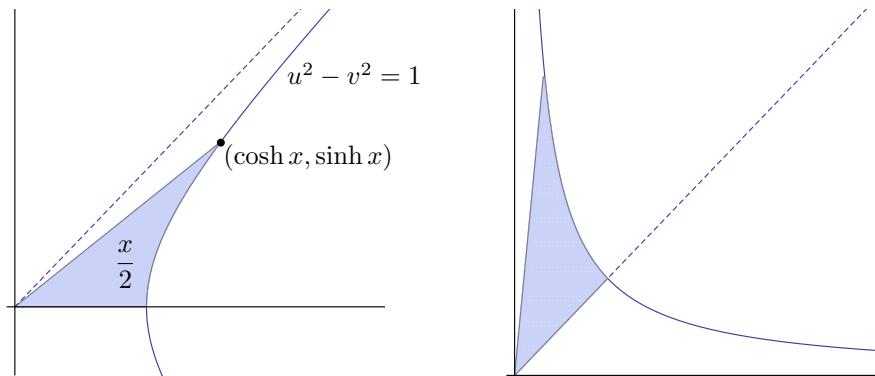
$$\cosh x := \frac{e^x + e^{-x}}{2}, \quad \sinh x := \frac{e^x - e^{-x}}{2}, \quad \tanh x := \frac{\sinh x}{\cosh x}$$

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}。



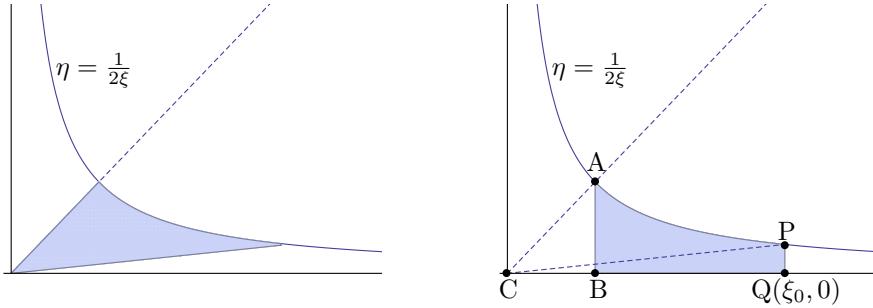
縦より横に積分した方が見やすいので、面積を求めるべき部分を直線 $\eta = \xi$ について折り返す。すると、次の二つの図で示された面積は等しいことが分かる。 $AB \cdot BC = PQ \cdot CQ = \frac{1}{2}$ が成り立ち、 $\triangle ABC$ と $\triangle CQP$ の面積が等しくなるからである。

^{*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

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

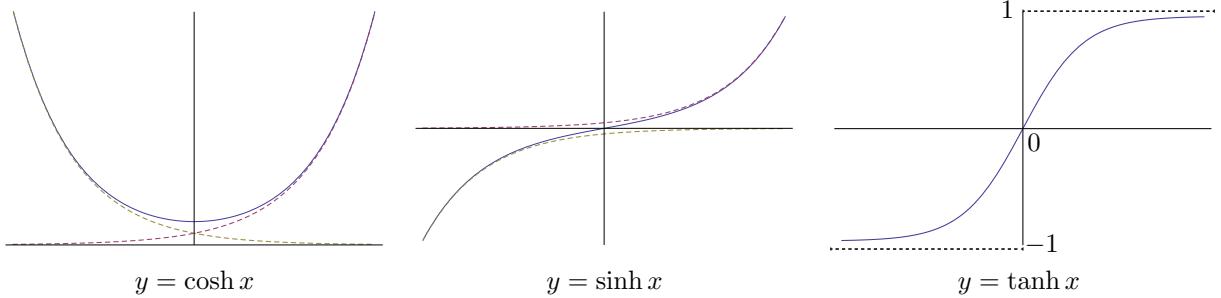
と分かる^{*12}。この実部が ξ_0 なので、面積の計算と合わせて $x = \log(\sqrt{2}\xi_0) = \log(\cosh x + \sinh x)$ が従う。つまり $\cosh x + \sinh x = e^x$ である。

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には「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

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

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、列の入れ替えて符号が (-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。

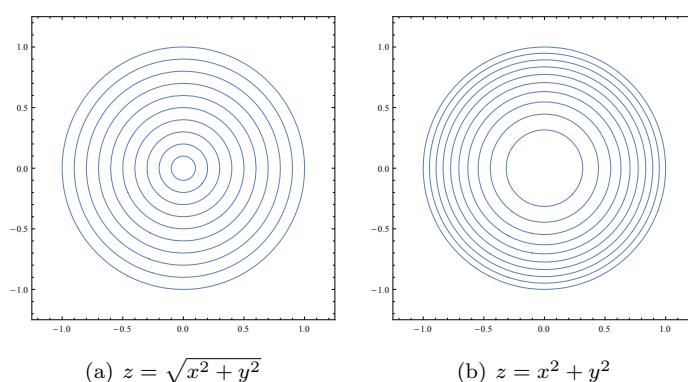
前回の内容を思い出しましょう。「 x, y, z の1次方程式が定める图形は平面である」という事実を確認しました。ですからグラフの方程式 $z = f(x, y)$ と平面の方程式 $ax + by + cz + d = 0$ を連立すれば、グラフ $z = f(x, y)$ の平面 $ax + by + cz + d = 0$ による断面が得られます。特に平面の方程式として分かりやすいもの、たとえば

- 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

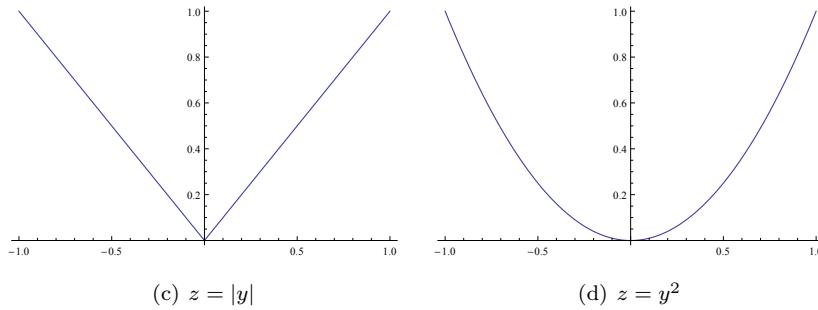
等高線 実数 c に対し、方程式 $z = c$ は z 座標の値が c である、 xy 平面と平行な平面を表します。したがって $z = f(x, y)$ と連立して $c = f(x, y)$ とすれば、平面 $z = c$ による断面が見えます。この断面のことをグラフの等高線といいます。等高線が把握できれば、 c の値を動かすことによってグラフの形状が把握できます。たとえば $z = \sqrt{x^2 + y^2}$ と $z = x^2 + y^2$ のグラフを、 xy 平面上に平行な平面を $z = 0$ から $z = 1$ まで 0.1 刻みで動かしてみましょう。



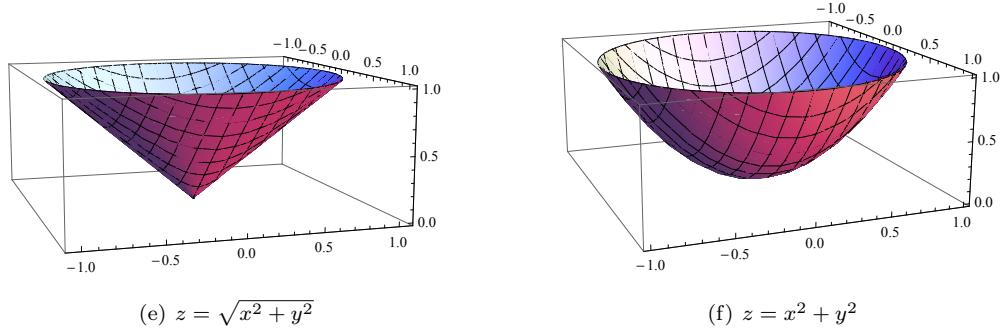
この図から明らかに、どちらのグラフも z 軸回りの回転対称性を持つと分かります。また等高線の密度から、 $z = \sqrt{x^2 + y^2}$ では高さが一定の間隔で増し、 $z = x^2 + y^2$ では急激に高さが増していく感じがします。ですがどの程度増えるのかは、この図だけでははっきりしません。詳しい情報を得るためにには、他の平面での切断面などが必要です。

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。

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



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。

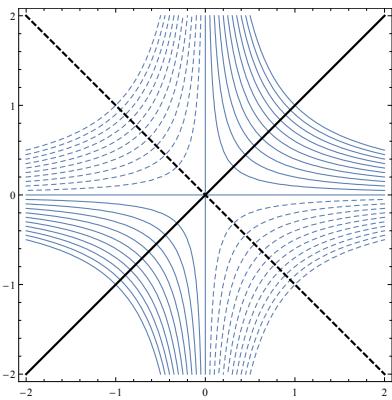
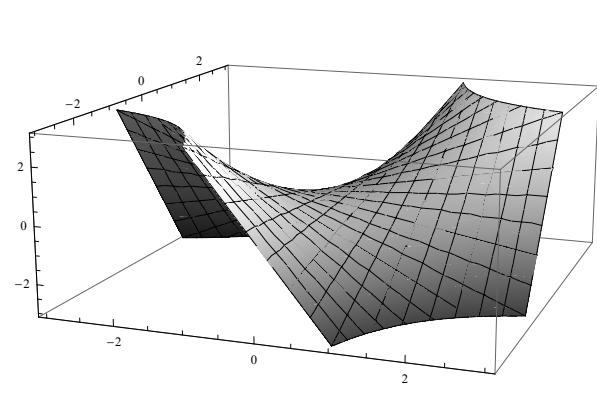
2.2 z 軸を含む平面に関する切断

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。

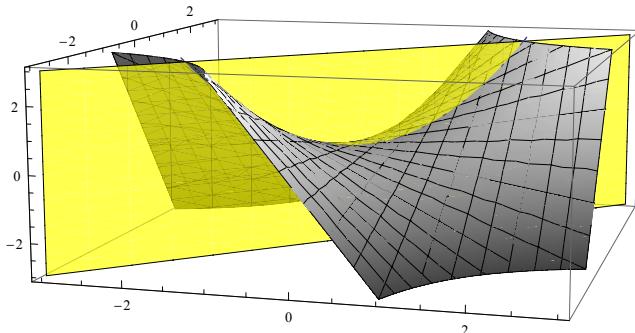
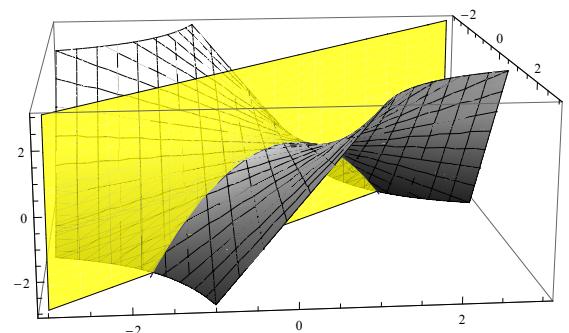
$$z = xy = \frac{(x+y)^2 - (x-y)^2}{4}$$

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

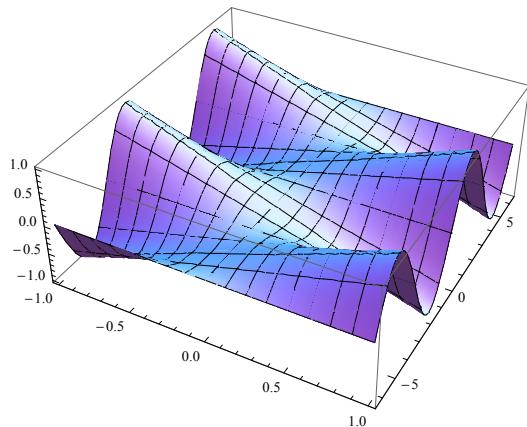
5.1 穴埋め問題の解答

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。

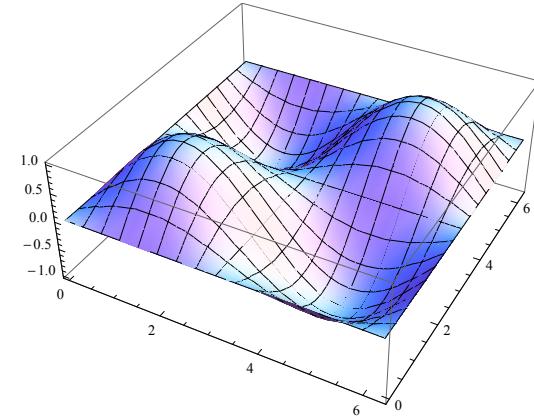
- | | | | | | | | | |
|----------------|------------|-----------|------------|----------|-----------|----------|------------|-------------|
| (1) 平面 | (2) 双曲放物面 | (3) z 軸 | (4) 回転対称性 | (5) 円錐 | (6) 回転放物面 | (7) a | (8) b | |
| (9) \sqrt{c} | (10) $2xy$ | | (11) 双曲放物面 | (12) 双曲線 | (13) a | (14) b | (15)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

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

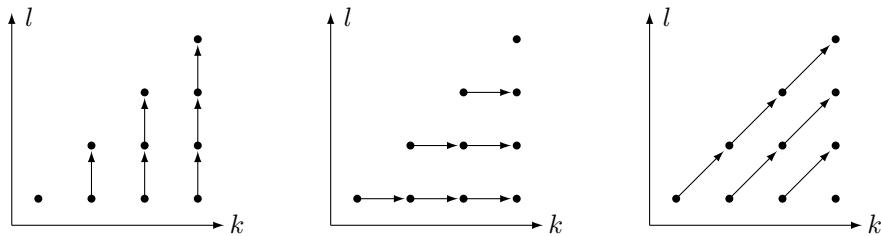
こんな感じで、2つの \sum をくっつけることができます。元々の二重和を「 k が $\{1, 2, \dots, m\}$ を走り、 l が $\{1, 2, \dots, n\}$ を走る」と書き換えた上で、さらに集合の直積を利用して「 (k, l) が $\{1, 2, \dots, m\} \times \{1, 2, \dots, n\}$ を走る」と書き換えました。こういう風に集合の直積を使うと、2つの \sum を1つにくっつけられます。逆に \sum の変数が直積集合を走るときは、 \sum を2つにばらすことができます。

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 次正方形行列に対し

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

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 次正方形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 数ベクトルに対する線型写像

1.1 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。

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}^m 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 $\mathbb{R}^m \rightarrow \mathbb{R}^n$ 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。

もう少し具体的な例を見てみましょう。

問1の解答 (1) 任意の x, y 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

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

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、 $E[x] = {}^t px$ 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。

1.2 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、 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。

1.3 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。

問 4 の訂正 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。ここで 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。しかし明らかに 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

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

である。よって $F: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ が等長線型変換なら

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

となり、 F が内積を保つことが分かる。 ■

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) $SR(\theta) = R(\theta)S$ の両辺に右から S をかけると、 $S^2 = E$ より $SR(\theta)S = R(\theta)S^2 = R(\theta)$ である。

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

2 線型空間

これまで数ベクトル空間上の線型写像を調べました。ここでもう一歩、線型写像の考察を進めてみましょう。僕たちは数ベクトルの性質を使って線型写像を定義したわけですが、その発想を逆転させてみます。「線型写像が定義できるには、定義域と値域にどんな性質があれば良いでしょうか？」¹

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。

2.1 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 $u + \mathbf{0} = \mathbf{0} + u = \mathbf{0}$ 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、 $0v = \mathbf{0}$, $(-1)v = -v$ 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、 $\mathbf{0} = 0v + \mathbf{0} = 0v$
- $\mathbf{0} = 0v = (1 + (-1))v = v + (-1)v$ より、 $-v = (-1)v$

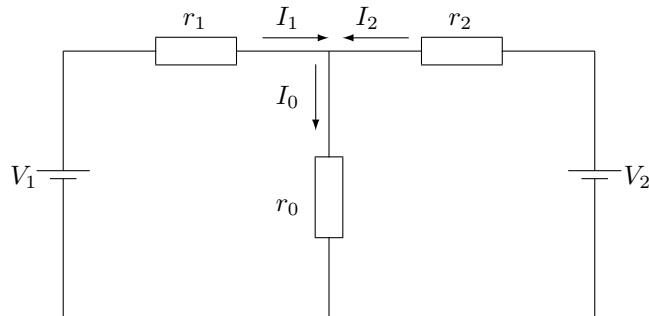
¹ ちなみに、この手の発想は数学ではよく見かけます。「 \square は $\times \times$ の性質を持つ」ということが分かったら、逆に「 $\times \times$ という性質が上手く定まるのはどういう場合か」と考え、抽象化を図るのであります。たとえば函数の「連続性」の概念から位相空間の考え方へ到達します。

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。

数ベクトル空間は当然線型空間なのですが、その部分集合が線型空間になることがあります。

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 $I_0 = I_1 + I_2$
- 左右それぞれのループに対する Kirchhoff の第 2 法則から $V_1 = r_1 I_1 + r_0 I_0, V_2 = r_2 I_2 + r_0 I_0$



が成り立つ。これらの式から I_0 を消去し、方程式を行行列で書くと

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

が成り立つ^{*3}。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

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

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。

連立 1 次方程式の解空間

2.2 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

^{*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。

^{*3} 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。

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、 $(Cy)'' - (Cy)' - 6(Cy) = C(y'' - y' - 6y) = 0$ 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 $y(x) = C_1 e^{3x} + C_2 e^{-2x}$ 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、 $C_1 = (2\alpha + 3\beta)/5$, $C_2 = (3\alpha - \beta)/5$ 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「微分は線型写像である」とことと「 y, y', \dots の項が齊次1次式」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 に関する項が全て1次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 (u_1, \dots, u_n) であって「任意の $v \in V$ が $v = a_1 u_1 + a_2 u_2 + \dots + a_n u_n$ の形に一意的に表せる」という性質を持つものを、 V の基底と言います。たとえば \mathbb{R}^n の基底は (e_1, e_2, \dots, e_n) です。

実は一般的な場合でも、線型空間には常に基底が取れることが知られています。ですから基底を1つ取ってしまえば、どんなベクトルもその1次結合で表せます。特に線型微分方程式の解空間の場合、基底をなす解を見つければ、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、 $C_1 e^{3x} + C_2 e^{-2x}$ ($C_1, C_2 \in \mathbb{R}$) が全ての解を尽くすことが保証されるのです。

S2 ターム

線形代数学

