多変数の微分積分学1

http://www.math.meiji.ac.jp/~mk/lecture/tahensuu1-2013/

桂田 祐史

katurada@meiji.ac.jp

2013年4月15日, 2013年8月28日

序に代えて

一里塚というのでしょうか。

約 20 年、明治大学理工学部数学科の 2 年生を対象に、多変数の微分積分学についての講義を担当して来た。前期の微分法だけを担当した年度、後期の微分法だけを担当した年度、前期後期の演習抜きの講義だけ担当した年度、通年すべてを担当した年度、様々であった。この講義科目の到達目標はほぼ同じであったが、学生の準備状況は細かく変化してきた。最後の頃は、1 年次に数学科だけのクラスで、週 2 コマの (1 変数関数の微積分) の授業と、$\mathbf{R}^2$ までに限定して位相を学ぶ授業があった。さらにこの多変数の微積分の講義と並行して、集合・距離・位相に関する科目がおかれるようになった。位相に関する議論に長い時間を費やさずに、微積分に集中して取り組めるようになっている。これは非常に恵まれた状況にあると思われる。カリキュラム全体を少しずつ改良してきた賜物だと感じている。

今回、このお役ご免になった。清々した気がするが、色々なものが泡となって消えないうちに記録しておくことにした。ある時期、どういう内容の講義が行われていたか記録しておくことは重要だと考えるからである。

教科書をあきらめたことについて　当初は前任の H 先生にならって、あるテキストを教科書として採用していたが、以下の理由から止めることになった。

(1) 教科書に小さな誤植があったが、何度注意しても学生が間違えたまま覚えて使ってしまった。

(2) 積分のところの説明が難しめで学生の重荷になる。

色々な美点のある本でなかなか惜しい感じがしたが、すっぱりと諦めた。(1) については、今ならば授業中に教科書に直接書き込みをさせるなど、もっと工夫をしたと思う。(2) については、「こうすればより簡単」と考えて採用した説明が、後から細かい修正・補足を重ねることによって、結局はかなり重くなってしまい (あきらめたテキストの説明よりも重くなってしまったかもしれない)、結果的には「言いがかり」に近くなったような気がしている。

自分ならもっとうまくやれる、良いアイディアを持っている

という人は少なくないが、実際にやってみるとなかなか難しいものである。

講義ノートを配ったことについて　自分が学生の頃を振り返ると、教科書なしで、授業中の教師の板書・説明をノートに取ったものがすべて、という講義は珍しくなかった。どちらかというと、そういう講義の方が、他の教科書のある講義よりも充実している場合が多かったように思う (何人かの先生方の顔が浮かんで来る)。無謀にも、この多変数の微分積分学でも、しばらくはそういうやり方をしてみたが、学生にとって授業内容を正確にノートに取ることは結構難しいものだ、ということにすぐに気づいた。さらに、講義をしている自分も、板書自体をうっかり書き間違えてしまうことがないとは言えない (学生に間違いを指摘してもらったら、お礼を言ってきちんと学生達にノートを直してもらうように努めているのだけど、もれる可能性は否定できませんね)。そこで時々まとめのプリントのようなものを作って配布を始めたが (実

はそういうことに関しては身近に何人もお手本がいた)、それが一年続くとある程度まとまったノートが出来ていた。いっそのことそれを配ってしまえば良い、と考えて実行したわけである。

やってみるとすぐ分かることだが、一つにまとめてみると色々なアラ (説明し忘れたことがあった、整合的でない) が分るものである。以後毎年毎年修正・補足を重ねて、どんどんページ数が増えていく。結局、品質の問題はおいておいて、書いてある事項は、出版されているテキストと大きく変わらないようなものが出来上がる。何となく年を重ねるごとに見通しが悪くなっていくように感じるのは、気のせいだろうか…

講義ノート作成に使ったツールについて　知っている人が見れば、TEX を使っていることは一目瞭然であろう。TEX は素晴らしいツールであると思う。筆者は学生の頃から、自習した事項のノートの整備にはかなり時間を費やしていたが、なかなか思うように行かなかった。後から書いたことを差し挟みたくなった、同じことを複数のノートに書きたい、色々な欲求が生じるが、普通の筆記用具で作っていくノートで完成度をあげるのは実に難しい。電子的なツールは、ともすると「清書用」に思われかねないが、むしろ下書き段階から積極的に利用すべきものである。

図について述べておく必要があるであろう。いわゆる数式まじりの文章を作成することについて、TEX は十分使いやすいツールであるが、図は本来は守備範囲外と言ってよい。

作図法を考えることは重要である。これを書いている時点 (2013 年 8 月) で、数学全般の利用に満足できるツールは見つけられていない。幸い微分積分については、フリーハンドで描くような図が必要になることは少なく、具体的な関数のグラフ等が描ければ済むことが多い。

(脱線になるが、そうでないことのノートが TEX で済むかどうかは難しい問題だと思う。TEX で書くことによって、手書きだったら書けるはずのことが抜け落ちてしまう、その結果、いびつな作業に陥らないか、心配になることがある。)

この講義ノートに載せる図を描くための作図ツールは、最初のうちは描ければ何でも良し、としてきた。C のプログラムと独自のグラフィックス・ライブラリィで描いて一丁上がり、とかしたものである。しかし最近は学生が再現できることを考えて、できる限り Mathematica や gnuplot などの定番ツールを使って作画し、そのためのコードを添付するようにしている。自分で再現・追体験出来るということ以外に、止まっている絵では不十分ということもある。3 次元グラフィックスなどは視点を変えて眺める、特に自分がその図形をつまんで揺すってみることで劇的に分りやすくなる場合が多い。

スリム化を目指して — 位相の扱いを考える　講義時間はいつでも不足がちなものである。学習効果を考えると、数学としての完備性よりは、やることを精選して、重要なことに集中することが望ましい。

ランダウの記号は中途半端に使うのは時間がもったいないように思う。積極的に使うのか、あるいは切り捨てるのか。積極的に使うメリットが見出せなかったため、最近では使わずに講義している。

位相に関する事項は、やり始めるときりがない。以前は、距離空間・位相空間に関する授業がなかったこともあり、なるべくこの授業でまとまった内容を講義しよう、と考えていた。カリキュラムが整備されてその責任から解放されたため、位相に関する事項の見直しを考え始めた。

もともと集積点、近傍、内部、外部などはカットあるいは説明を短くするようになっていった。そこからもう一歩踏み出すことになったのは、2011 年度、震災の影響で授業回数が削減されたことによる。この授業で必要になることを見直した (例えば期末試験の問題を数学的に厳密に解くために、何が必要になるか、何はなくても済むか)。

参考までに、2010 年頃までの講義内容は、別の講義ノート [1] に全部説明してある。この文書は 2013 年度の講義内容に対応するノートである。

- 開集合はほとんどの定理に登場する基本語である。
- 閉包もないと困る (集積点とどちらかを選ぶものか？)。
- 重要かつ利用頻度の高い定理として、「$\mathbf{R}^n$ の有界閉集合上の実数値連続関数は最大値と最小値を持つ」がある。
- (そういうわけで) 有界集合も必要である。閉集合の重要性も非常に高い。
- 「$\mathbf{R}^n$ の有界閉集合上の連続関数は一様連続である」は、積分をするまでは必要がない。
- 関数の値域を明確にするためにも、中間値の定理はあった方が良い。

幸いなことに、「$\mathbf{R}^n$ の有界閉集合上の実数値連続関数は最大値と最小値を持つ」は、$\mathbf{R}^2$ の場合に学習しているので、「同様である」として証明を省略しても構わない。証明をする場合はそれなりの時間 (Bolzano-Weierstrass の定理と、閉集合の点列による特徴づけなど) が必要になるが、それはそれで教育的効果があると思う。

(なお、中間値の定理は、1 変数版を仮定し、弧連結な場合に限れば、ごく短時間で済ませられる。)

与えられた集合が開集合か、閉集合か、有界集合かを判定できる、もっと絞れば、そうである場合にそのことを厳密に確認できるようになること、を目標にすることにした (そうでない場合にそうでないことを証明することも出来るに越したことはないが、重要性は比較的低い、と割り切ることに決めた)。

距離空間・位相空間について学ぶときは、例えば開集合の定義に基づいて、与えられた集合が開集合であること、開集合でないことを確認出来るようになることが重要である。しかし実際問題、自分はそういうことをめったにやらない。

$$E := \left\{(x, y); x^2 + y^2/4 = 1\right\}$$

が閉集合であることは、連続関数 $(x, y) \mapsto x^2 + y^2/4$ による閉集合 $\{1\}$ の逆像だから、と考える。実際には、$f\colon \mathbf{R}^n \to \mathbf{R}$ が連続関数であるとき,

$$\{x \in \mathbf{R}^n; f(x) \ge a\}, \quad \{x \in \mathbf{R}^n; f(x) \le b\}, \quad \{x \in \mathbf{R}^n; a \le f(x) \le b\}, \quad \{x \in \mathbf{R}^n; f(x) = c\}$$

はいずれも閉集合である、という定理を紹介して、それに慣れてもらうことにした。

こうなると欲が出て、与えられた関数が連続である理由を厳密に述べる、ということが出来て欲しくなる。建前上は「簡単なので出来るはず、出来ないのはけしからん」であるが、実際に出来る学生は多くない。そこで

連続関数を組み立てて出来る関数は連続である

を標語として、処方箋を用意した。素材となる連続関数としては、

(a) 高校で習った、1 変数の、グラフまで明瞭に思い浮かべられる初等関数

(b) 多変数多項式で定義される関数 (「多項式関数」と呼ぶことにした)

を掲げた。

残念ながら、多項式とは何かきちんと説明できない学生は多く、有理式や有理式ですらないものを多項式と言ってのける困った人までいる。「平均的な学生はその程度のものだ」とも思うけれど、良い機会があるから、身につけてもらおう、と考えた。

配らなかった2011年度版講義ノートの「はじめに」

(震災の影響による時間短縮のため、この年度は講義内容をそれ以前のものから変更し、講義ノートを配らず、参考資料として WWW に載せておくにとどめた。)

関数の性質を解析する学問である微分積分学は、ニュートン (Sir Isaac Newton, 1642–1727), ライプニッツ (Leibniz, 1646–1716) 以来の長い歴史を持っている[1]。その最初の本格的な応用が、ニュートン力学の構築にあったという事実[2]を指摘するまでもなく、微分積分学は数学のみならず、現代の科学技術の一翼をになう、大変重要な学問である。

「多変数の微分積分学 1」は、明治大学数学科の 2 年生を対象に多変数関数の微分法を講義する目的で用意された科目である。多変数関数の微分法について基本的なことを、数学的にきちんと説明し、学生諸君にしっかりとしたイメージを持ってもらうことを目標にしている。既に高等学校や大学 1 年次の数学で微分積分学を学んできたはずだが、それらは (ほとんどすべて) 1 変数関数を対象とするものであったはずである。これでは他の学問で使うためにも十分であるとは言い難い。実際、多変数関数は、高等学校の理科にも普通に登場する概念である。1 変数関数に対して得られた結果の多くが多変数関数にまで拡張されるが、多変数であるがゆえの難しさ、面白さがあちこちに出て来る。もちろん、ここで学んだものは、基礎常識として、今後学ぶさまざまな数学で使われることになる。

この科目では、教科書を指定していない。それを補う目的で用意されたのが、この講義ノートである。昔から大学の講義では「教科書なし」というものが珍しくない。教師の板書あるいは喋ることそのものがテキストというわけである。ところが、最近の学生諸君を見ると、板書を正確かつ迅速に書き取る力がかなり落ちていると痛感する (この点は自覚を持って、実力向上に努めて欲しい)。また、教師側も板書の際に、書き間違いをしないとも限らない (私は常習犯です)。最近では、その種の教師のうっかりを指摘してくれる学生が減っているので、板書にはかなり注意を払う必要がある。そこでいわゆる「講義ノート」に相当するものを配布してしまおう、と考えるようになった。

書き進めていくうち、学部 4 年生の卒業研究や、大学院生の指導の際に気がついたことも含めたいと思うようになり、内容に多少反映されている。このように、すぐには必要にならないことまで一緒にまとめてしまうとページ数が増え、かえって読みにくくなる可能性があるが、それには目をつぶることにした。一応言い訳をしておくと、次のように考えるからである。

- 最初に読んだ本以外のものを使いこなすのは大変である。最初から、ある程度以上詳しい本を与える方が親切である。

- (基本中の基本である) 微分積分学とはいえ、生きた数学を研究するための道具であって、使う際の便利さを考えておくべきである。

このノートを利用する場合の注意

このノートは、講義の内容にかなり忠実に書いてあるつもりであるが、

- あくまでも 4 月の時点で出来ているものであり、講義の方はその後も可能な限り工夫をするので、どうしてもズレが出るのは仕方がない

[1]歴史的なことに言及した微分積分学のテキストとして、ハイラー・ヴァンナー [2] を推奨しておく。

[2]Newton は名高い『プリンキピア・マテマティカ』(Principia Mathematica Philosophiae Naturalis, 1687) において、運動の三法則と万有引力の法則を仮定すると、当時の課題であった Kepler の法則が導かれることを示し、Newton 力学を確立した。

- 本来、図を描いて説明すべきところを、手間の問題から省略しているところが多い (**だから、講義中の図には特に注意を払って欲しい**)
- 既に述べたように、細かい進んだ話題も、後で役に立ちそうに考えられる場合は、(講義で説明しない可能性が高くても) 書いてある

などの理由から、100% 一致しているとは言えない。
また、その点はクリアしたとしても、このノートを読めば、講義に出る必要はないとは考えない方が良い。

> **数学の本は、1 ページを読むのに要する時間が、そうでない本に比べて格段に長い (要するにかなり読み難い)**

というのは残念ながら真実なので、本文 130 ページ強の内容を自分一人で読破するのは、ほとんどの人にとって、困難なはずである。少しずつ時間をかけて理解して行くしかない。そうするためには、結局は授業に出席して、その時間に頭を働かせるのが近道である。
もう一つ注意しておきたいのは、「授業に出てみたが、分からないから、出ても意味がない。」という考え方をする人がときどきいるが、それは考え直した方が良い、ということである。数学も大学レベルになると、難しくなって来て、説明されてもすぐには良く理解できないのが普通である。勉強を続けていって、ある時点で、急に (不連続的に) 納得が出来るものである[3](数ヵ月のオーダーで、納得が勉強に遅れてついて来ることは良くある)。この種の我慢はどうしても必要であることを信じてほしい。

その他の注意

- (存在しなくなる WWW ページの URL の紹介なので省略)
- 集合・距離・位相に関して、1 年生向けの「数学演習 2」でもかなり詳しく説明されるようになったし、2 年生に「集合・距離・位相 1」,「集合・距離・位相 2」という講義科目も用意されている。そのため、この講義では、以前は説明していた事項のいくつかを省略することにしてある。しかしそれらの事項についても、この講義ノートには解説を残してある。

[3]このことを山登りに例えた人がいる。つまり見通しの良いところに出るまでは、自分が高いところまで登っている実感が得にくい、ということである。

目 次

記号表

開球と閉球　$a \in \mathbf{R}^n$, $r > 0$ に対して、

$$B(a;r) := \{x \in \mathbf{R}^n; \|x-a\| < r\},$$

$$\overline{B}(a;r) := \{x \in \mathbf{R}^n; \|x-a\| \leq r\}.$$

論理　標準的と思われる「かつ」 $\wedge$, 「または」 $\vee$, 「でない」 $\neg$, 「ならば」 $\Rightarrow$, 「必要十分条件である」 $\Leftrightarrow$, 「任意の」 $\forall$, 「存在する」 $\exists$ などの記号を用いる。

$\forall x\ P(x) \Rightarrow Q(x)$ を $(\forall x\colon P(x))\ Q(x)$ と書く。$\exists x\ P(x) \wedge Q(x)$ を $(\exists x\colon P(x))\ Q(x)$ と書く。

この記号の約束のもとに、

$$\neg\left((\forall x : P(x))\, Q(x)\right) \quad \Leftrightarrow \quad (\exists x : P(x))\neg Q(x),$$

$$\neg\left((\exists x : P(x))\, Q(x)\right) \quad \Leftrightarrow \quad (\forall x : P(x))\neg Q(x),$$

が成り立つ。

差集合、補集合　集合 A, B に対して、

$$A \setminus B := \{x \in A; x \notin B\}.$$

集合 A の補集合を A^c で表す。全体集合が $\mathbf{R}^n$ であるとき、$A \subset \mathbf{R}^n$ に対して、

$$A^c := \mathbf{R}^n \setminus A = \{x \in \mathbf{R}^n; x \notin A\}.$$

閉包　$\Omega \subset \mathbf{R}^n$ に対して、

$$\overline{\Omega} = \{x \in \mathbf{R}^n; \forall \varepsilon > 0 \quad B(x;\varepsilon) \cap \Omega \neq \emptyset\}.$$

はじめに

「多変数の微分積分学 1」では、多変数ベクトル値関数の微分法を学ぶ (積分法は後期の「多変数の微分積分学 2」で学ぶ)。

1 年次に次のようなことを学んだ。

- 基礎数学 1,2 は、$f\colon \mathbf{R}^n \to \mathbf{R}^m$, $f(x) = Ax$ (A は $m \times n$ 型行列) という関数についての数学であった。
  扱っている変数、関数の値ともにベクトルという意味では一般的だが、f は線形 (もしも $m = n = 1$ であれば f は中学校で学ぶ「正比例」の関数 $f(x) = ax$) で、そういう意味では簡単な場合を扱っている。

- 基礎数学 3,4 は、$f\colon I \to \mathbf{R}$ (ただし I は $\mathbf{R}$ の区間) という関数 (1 変数実数値関数) についての微積分であった。
  関数は「曲がっていても良く」て、そういう意味では一般的だが、変数も関数の値も実数、という意味で簡単な場合を扱っている。

これらを背景に，多変数の微分積分学 1,2 では、$f\colon \Omega \to \mathbf{R}^m$ (ただし Ω は $\mathbf{R}^n$ の部分集合) という関数 (n 変数 m 次元ベクトル値関数) の微分積分を学ぶわけである。

1 変数であればベクトル値になっても実数値の関数と違いはほとんどない。難しくなる原因は、関数が多変数になることである、と言える (多変数関数では、これまでになかったようなことが起こる)。そのため、講義名も「多変数の 微分積分学 1」となっている。

第1章　R^n の基本的な性質と 1 変数ベクトル値関数

1.1　多変数ベクトル値関数とは

例 1.1.1 ある瞬間の部屋の中の空気の温度を考える。場所によって異なるので、場所の関数である。適当な直交座標系を用意すると、つまり部屋の中の任意の点は、三つの実数の組 (x_1, x_2, x_3) で表される。その点での温度を

$$u(x_1, x_2, x_3)$$

とすると、u は 3 つの変数 x_1, x_2, x_3 についての関数である。ベクトル変数 $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ についての関数ともみなせる。

$$u(\vec{x}) = u(x_1, x_2, x_3).$$

部屋は $\mathbf{R}^3$ のある部分集合 Ω であると考えられ、u は $u\colon \Omega \to \mathbf{R}$ という写像となる。

(もし時間による変化を考えると、時刻 t の関数でもあることになり、4 変数関数 $u(x_1, x_2, x_3, t)$ になる。)

今度は、ある瞬間の部屋の中の風の速度を考える。位置 $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ における風の速度を

$$\vec{v}(\vec{x}) = \begin{pmatrix} v_1(\vec{x}) \\ v_2(\vec{x}) \\ v_3(\vec{x}) \end{pmatrix} = \begin{pmatrix} v_1(x_1, x_2, x_3) \\ v_2(x_1, x_2, x_3) \\ v_3(x_1, x_2, x_3) \end{pmatrix}$$

とすると、$\vec{v}$ は $\vec{v}\colon \Omega \to \mathbf{R}^3$ という写像となる。■

集合、写像の言葉を使って書くと、n 変数 m 次元ベクトル値関数とは、$\mathbf{R}^n$ のある部分集合 Ω 上定義され、 $\mathbf{R}^m$ に値を取る写像

$$\vec{f}\colon \Omega \longrightarrow \mathbf{R}^m$$

のことである。

$$\vec{f}(\vec{x}) = \begin{pmatrix} f_1(x_1, x_2, \ldots, x_n) \\ f_2(x_1, x_2, \ldots, x_n) \\ \vdots \\ f_m(x_1, x_2, \ldots, x_n) \end{pmatrix}$$

の形をしている。

上の u は 3 変数 1 次元ベクトル値 (実数値) 関数であり、$\vec{v}$ は 3 変数 3 次元ベクトル値関数である。

1.2　1変数ならばベクトル値でも簡単

まず証明は後回しにして、どういうことが成り立つか、説明する。

1 変数ならば、ベクトル値であっても、極限、連続性、微分可能性等は簡単である。
I を $\mathbf{R}$ の区間、$\vec{f}\colon I \to \mathbf{R}^m$ とする。各 $x \in I$ に対して、$\vec{f}(x) \in \mathbf{R}^m$ であるから、

$$\begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix} := \vec{f}(x)$$

とおくと、各 $i \in \{1, 2, \ldots, m\}$ に対して、$f_i\colon I \to \mathbf{R}$ である。
次のようにまとめておく (納得しやすいであろう)。

m 次元ベクトル値関数とは、実数値関数 m 個の組である。

微積分をするには、極限や連続性が問題となるが、1 変数ベクトル値関数 $\vec{f}$ については、極限や連続性、微分は「成分 f_i ごと」に考えれば良い。
例えば、極限については、

$$\lim_{x\to a} \vec{f}(x) = \vec{A} \quad \Leftrightarrow \quad \forall i \in \{1, 2, \cdots, m\} \quad \lim_{x\to a} \vec{f_i}(x) = \vec{A_i} \quad (\text{ただし } \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{pmatrix} := \vec{A}).$$

言い換えると、

$$\lim_{x\to a} \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix} = \begin{pmatrix} \lim_{x\to a} f_1(x) \\ \lim_{x\to a} f_2(x) \\ \vdots \\ \lim_{x\to a} f_m(x) \end{pmatrix}$$

ということであり、「ベクトル値関数の極限は実数値関数の極限に帰着される」。

例 1.2.1 $\vec{f}(x) = \begin{pmatrix} \dfrac{\sin x}{x} \\ x^3 - 3x + 2 \end{pmatrix}$ とするとき、

$$\lim_{x\to 0} \vec{f}(x) = \begin{pmatrix} \lim_{x\to 0} \dfrac{\sin x}{x} \\ \lim_{x\to 0} \left(x^2 - 3x + 2\right) \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}. \blacksquare$$

微分に関しても同様に、

$$\vec{f}'(x) = \begin{pmatrix} f_1'(x) \\ f_2'(x) \\ \vdots \\ f_m'(x) \end{pmatrix}$$

が成り立つ。高階の導関数についても同様である。

関数が実数値からベクトル値になっても、(成分ごとにやればよく) あまり変わらない

例 1.2.2 (質点の運動、特に等速円運動) 質点が時間の経過とともにその位置を変えるとき、時刻 $t \in I$ (I は $\mathbf{R}$ の区間) における位置ベクトルを $\vec{f}(t)$ で表すと，一つのベクトル値関数 $\vec{f}\colon I \to \mathbf{R}^m$ が得られる。

ここでは独立変数を t, 従属変数を $\vec{x}$ と書くことにしよう:

$$\vec{x} = \vec{f}(t).$$

このとき

$$\vec{v}(t) := \frac{d\vec{x}}{dt} = \vec{f}'(t)$$

を時刻 t における質点の**速度** (velocity) と呼ぶ。また速度のノルム (大きさ) $\|v(t)\|$ のことを**速さ** (speed) と呼ぶ。

速度の導関数

$$\vec{\alpha}(t) := \vec{v}'(t) = \frac{d^2\vec{x}}{dt^2} = \vec{f}''(t)$$

は質点の**加速度** (acceleration) と呼ばれる。

$m = 2$, $r > 0$, $\omega \in \mathbf{R}$ とするとき、$\vec{f}(t) := \begin{pmatrix} r\cos\omega t \\ r\sin\omega t \end{pmatrix}$ とおくと、原点を中心とする半径 r の円周上を一定の角速度 ω で移動する質点の運動とみなせる。

$$\vec{v}(t) = \vec{f}'(t) = \begin{pmatrix} -r\omega\sin\omega t \\ r\omega\cos\omega t \end{pmatrix}$$

であるから、

$$\|\vec{v}(t)\| = r|\omega|, \quad \vec{v}(t) \perp \vec{f}(t).$$

(後者は $\vec{v}(t)\cdot\vec{f}(t) = 0$ や、$\vec{v}(t) = \omega\begin{pmatrix} \cos 90^\circ & -\sin 90^\circ \\ \sin 90^\circ & \cos 90^\circ \end{pmatrix}\vec{f}(t)$ から分かる。) また

$$\vec{\alpha}(t) = \vec{f}''(t) = \begin{pmatrix} -r\omega^2\cos\omega t \\ -r\omega^2\sin\omega t \end{pmatrix}$$

であるから、

$$\|\vec{\alpha}(t)\| = r\omega^2, \quad \vec{\alpha}(t) = -\omega^2\vec{f}(t), \quad \vec{\alpha}(t) \perp \vec{v}(t).$$

(「加速度は中心を向く」、「速度と加速度は直交する」)。 ■

1.3 $\mathbf{R}^n$ の標準的な内積とノルム

$\mathbf{R}^n$ については、既に学んだはずであるが、記号の確認、復習を兼ねて説明しておく (授業では駆け足で通り過ぎるはずである)。後々、不等式が重要になる (極限の議論をする際に使うことが多い)。

定義 1.3.1 (内積空間としての R^n の定義) $n \in \mathbf{N}$ に対して

$$\begin{aligned}\mathbf{R}^n &:= \mathbf{R}\times\mathbf{R}\times\cdots\mathbf{R} \qquad (n \text{ 個の } \mathbf{R} \text{ の直積})\\ &= \left\{\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}; x_i \in \mathbf{R} \quad (i = 1, 2, \ldots, n)\right\}\end{aligned}$$

とおく。ただし n 次元**内積空間** (内積を持った n 次元線形空間) としての構造を入れておく。言い替えると

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbf{R}^n, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \in \mathbf{R}^n, \quad \lambda \in \mathbf{R}$$

に対して、**和** $\vec{x} + \vec{y}$, **スカラー倍** $\lambda\vec{x}$, **内積** (**inner product**) $(\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y}$ を以下のように定義する。

$$\vec{x} + \vec{y} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda\vec{x} := \begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \vdots \\ \lambda x_n \end{pmatrix}, \quad (\vec{x}, \vec{y}) = \vec{x} \cdot \vec{y} := \sum_{i=1}^{n} x_i y_i.$$

注意 1.3.2 (ベクトルの縦と横) この講義では、ベクトルが縦であるか横であるか、問題になるときは、断りのない限り縦ベクトルとする。行列 A とベクトル x のかけ算を Ax と書きたいからである。 ■

注意 1.3.3 (内積の呼び方、記号) 内積のことを**スカラー積** (**scalar product**), **ドット積** (**dot product**) とも呼ぶ。また、$(\vec{x}, \vec{y})$ という記号は、順序対 (要するに $\vec{x}$, $\vec{y}$ の組) と紛らわしいという理由で、内積を $\langle \vec{x}, \vec{y} \rangle$ と表している本も多い。

$(\vec{x}, \vec{y}) = \vec{y}^T \vec{x}$ ($\vec{y}^T$ は $\vec{y}$ を転置して出来る 1 行 n 列の行列) と書けることを知っておくと便利なことがある[1]。■

$\mathbf{R}^n$ のことを n 次元数ベクトル空間、あるいは n 次元 Euclid (ユークリッド) 空間と呼ぶ。$\mathbf{R}^n$ の要素のことを (その時の気分で) 点と呼んだり、ベクトルと呼んだりする。(ここまで、$\mathbf{R}^n$ の要素には矢印 $\vec{}$ をつけたが、面倒なので、以下は混同のおそれがない限り、省略することもある。)

以下 $\mathbf{R}^n$ と書いたとき、n が自然数を表すことは一々断らないことが多い。

次の命題が成り立つことは明らかであろう。

[1]例えば、2 次形式を扱う時に基本となる公式 $(A\vec{x}, \vec{y}) = (\vec{x}, A^T\vec{y})$ は、行列の積に関する結合法則で証明できる。

命題 1.3.4 (内積の公理) $\mathbf{R}^n$ の内積 $(\cdot,\cdot)$ は次の性質を満たす。

(1) $\forall \vec{x} \in \mathbf{R}^n$ に対して $(\vec{x},\vec{x}) \geq 0$. また $\forall \vec{x} \in \mathbf{R}^n$ に対して

$$(\vec{x},\vec{x}) = 0 \Longleftrightarrow \vec{x} = 0.$$

(2) $\forall \vec{x} \in \mathbf{R}^n$, $\forall \vec{y} \in \mathbf{R}^n$, $\forall \vec{z} \in \mathbf{R}^n$, $\forall \lambda \in \mathbf{R}$, $\forall \mu \in \mathbf{R}$ に対して

$$(\lambda\vec{x} + \mu\vec{y}, \vec{z}) = \lambda(\vec{x},\vec{z}) + \mu(\vec{y},\vec{z}).$$

(3) $\forall \vec{x} \in \mathbf{R}^n$, $\forall \vec{y} \in \mathbf{R}^n$ に対して

$$(\vec{x},\vec{y}) = (\vec{y},\vec{x}).$$

証明　簡単なので省略する。■

余談 1.3.1 (証明の手引き) 確かに簡単ではあるけれど、少しお手本を。例えば、(2) を簡単にした

$$(\bigstar) \qquad \forall \vec{x},\vec{y},\vec{z} \in \mathbf{R}^n \quad (\vec{x}+\vec{y},\vec{z}) = (\vec{x},\vec{z}) + (\vec{y},\vec{z})$$

をどのように証明するか、説明を補足しておこう。

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \vec{z} = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{pmatrix}$$

とおくと、

のように自分で書き出せることが、何でもないようでいて大事である。後は $(\bigstar)$ の式の部分を x_j, y_j, z_j で表せば良い。

$$\vec{x}+\vec{y} = \begin{pmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{pmatrix}$$

であるから

$$(\vec{x}+\vec{y},\vec{z}) = \sum_{j=1}^{n}(x_j+y_j)z_j = \sum_{j=1}^{n} x_j z_j + \sum_{j=1}^{n} y_j z_j = (\vec{x},\vec{z}) + (\vec{y},\vec{z})\,.$$

この真似をして下さい。■

定理 1.3.5 (Cauchy-Schwarz の不等式) $\mathbf{R}^n$ の内積 $(\cdot,\cdot)$ は次の性質を満たす。$\forall\vec x\in\mathbf{R}^n$, $\forall\vec y\in\mathbf{R}^n$ に対して

(1.1) $$(\vec x,\vec y)^2\le(\vec x,\vec x)(\vec y,\vec y).$$

この不等式において等号が成り立つための必要十分条件は、$\vec x$ と $\vec y$ が 1 次従属である (片方がもう一方のスカラー倍である) ことである。

(この後に書いてあるように $|(\vec x,\vec y)|\le\|\vec x\|\,\|\vec y\|$ や、$-\|\vec x\|\,\|\vec y\|\le(\vec x,\vec y)\le\|\vec x\|\,\|\vec y\|$ と書くことも出来る。そちらの形の方が分かりやすいかもしれない。

以下では "Cauchy" を略して単に「Schwarz の不等式」と呼ぶことにする。"Cauchy" を略してはいけないとおっしゃる先生もいるので、最初だけ Cauchy の顔を立てた。)

証明

(i) $\vec x$, $\vec y$ が 1 次独立な場合。
$\forall t\in\mathbf{R}$ に対して $t\vec x+\vec y\ne 0$ であるから、
$$(t\vec x+y,t\vec x+y)>0.$$
ゆえに
$$(\vec x,\vec x)t^2+2(\vec x,\vec y)t+(\vec y,\vec y)>0.$$
t についての 2 次式の符号が変わらないことから
$$\frac{\text{判別式}}{4}=(\vec x,\vec y)^2-(\vec x,\vec x)(\vec y,\vec y)<0.$$
移項して
$$(\vec x,\vec y)^2<(\vec x,\vec x)(\vec y,\vec y).$$

(ii) $\vec x$, $\vec y$ が 1 次属な場合。
次のいずれかが成り立つ。

(a) $\vec x=t\vec y$ となる $t\in\mathbf{R}$ が存在する。

(b) $\vec y=t\vec x$ となる $t\in\mathbf{R}$ が存在する。

(a) の場合、$(\vec x,\vec y)^2=[t(\vec y,\vec y)]^2=t^2(\vec y,\vec y)^2$, $(\vec x,\vec x)(\vec y,\vec y)=t^2(\vec y,\vec y)^2$ だから $(\vec x,\vec y)^2=(\vec x,\vec x)(\vec y,\vec y)$.
(b) の場合、$(\vec x,\vec y)^2=[t(\vec x,\vec x)]^2=t^2(\vec x,\vec x)^2$, $(\vec x,\vec x)(\vec y,\vec y)=t^2(\vec x,\vec x)^2$ だから $(\vec x,\vec y)^2=(\vec x,\vec x)(\vec y,\vec y)$.

(a), (b) どちらの場合も
$$(\vec x,\vec y)^2=(\vec x,\vec x)(\vec y,\vec y).$$

(i), (ii) をまとめると、一般に不等式 (1.1) が成り立ち、等号は $\vec x$, $\vec y$ が 1 次従属なとき、そのときに限り成立する。■

定義 1.3.6 ($\mathbf{R}^n$ のノルム) $\vec{x} \in \mathbf{R}^n$ に対して、$\vec{x}$ の**ノルム** (**norm**, 長さ、大きさ) $\|\vec{x}\|$ を

$$\|\vec{x}\| := \sqrt{(\vec{x}, \vec{x})}$$

で定める。すなわち $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ とするとき

$$\|\vec{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} = \sqrt{\sum_{j=1}^n x_j^2}.$$

($\vec{x}$ のノルムを $|\vec{x}|$ と表す流儀もある。)

この記号を用いると、Schwarz の不等式 (1.3.5) は

$$|(\vec{x}, \vec{y})| \le \|\vec{x}\| \|\vec{y}\| \tag{1.2}$$

とも表される。

任意の実数 a, b について、

$$|a| \le b \quad \Leftrightarrow \quad -b \le a \le b$$

であることから、Schwarz の不等式は

$$-\|\vec{x}\| \|\vec{y}\| \le (\vec{x}, \vec{y}) \le \|\vec{x}\| \|\vec{y}\|$$

と書くことも出来る。右の不等式で等号が成り立つためには、$\vec{x}$ と $\vec{y}$ が同じ方向＆同じ向きであること ($\exists t \ge 0$ s.t. $\vec{y} = t\vec{x} \vee \vec{x} = t\vec{y}$ である)、左の不等式で等号が成り立つためには、$\vec{x}$ と $\vec{y}$ が同じ方向＆反対の向きであること ($\exists t \le 0$ s.t. $\vec{y} = t\vec{x} \vee \vec{x} = t\vec{y}$ である) が必要十分である (そのことを証明するには、上の定理の証明を真似をすれば良い)。

Schwarz の不等式の別証明 (あらすじ) $\vec{y} = \vec{0}$ のときは明らかなので、$\vec{y} \ne \vec{0}$ とする。原点と $\vec{y}$ を通る直線に、$\vec{x}$ から垂線を下ろして、交点を $\vec{z}$ とする (昔風の表現をすると、$\vec{x}$ からその直線に下ろした垂線の足)。このとき直角三角形が出来るので

$$\|\vec{z}\|^2 + \|\vec{x} - \vec{z}\|^2 = \|\vec{x}\|^2$$

が成り立つ (ピタゴラスの定理)。ゆえに $\|\vec{z}\|^2 \le \|\vec{x}\|^2$. これから

$$\|\vec{z}\| \le \|\vec{x}\|.$$

$\vec{z}$ を $\vec{x}, \vec{y}$ で表した式をこの不等式に代入すると、Schwarz の不等式が得られる。つまり Schwarz の不等式は、直角三角形では、斜辺が一番長い、という意味を持っている。 ■

高等学校流の内積の幾何学的定義 高等学校の数学では、内積を図形的に定義することがある。$\vec{x}$ と $\vec{y}$ のなす角を θ ($\theta \in [0, \pi]$) とするとき、

$$(\vec{x}, \vec{y}) = \|\vec{x}\| \|\vec{y}\| \cos\theta$$

として内積を定義するわけである。一般に $-1 \le \cos\theta \le 1$ であるから、

$$-\|\vec{x}\| \|\vec{y}\| \le (\vec{x}, \vec{y}) \le \|\vec{x}\| \|\vec{y}\|$$

は明らかである。等号が成立する条件も分かりやすい。(例えば、$\vec{x}$ と $\vec{y}$ が同じ方向&同じ向きとは、$\theta=0$ ということで、それは $\cos\theta=1$ や $(\vec{x},\vec{y})=\|\vec{x}\|\,\|\vec{y}\|$ と同値である。) この授業では、図が描けるとは限らない一般の $\mathbf{R}^n$ のベクトル $\vec{x}$, $\vec{y}$ について

$$(\vec{x},\vec{y})=\sum_{j=1}^{n}x_jy_j$$

で内積を定義しているので、Schwarz の不等式は上でやったような証明をすることが必要になる。逆に $\vec{x},\vec{y}\neq\vec{0}$ とするとき、Schwarz の不等式から

$$\cos\theta=\frac{(\vec{x},\vec{y})}{\|\vec{x}\|\,\|\vec{y}\|}$$

を満たす $\theta\in[0,\pi]$ として、$\vec{x}$ と $\vec{y}$ のなす角 θ が定義できることになる。■

命題 1.3.7 (ノルムの公理) $\mathbf{R}^n$ のノルム $\|\cdot\|$ は次の性質を満たす。

(i) $\forall\vec{x}\in\mathbf{R}^n$ に対して

$$\|\vec{x}\|\geq 0.$$

$\forall\vec{x}\in\mathbf{R}^n$ に対して

$$\|\vec{x}\|=0\Longleftrightarrow\vec{x}=0.$$

(ii) $\forall\vec{x}\in\mathbf{R}^n$, $\forall\lambda\in\mathbf{R}$ に対して

$$\|\lambda\vec{x}\|=|\lambda|\|\vec{x}\|.$$

(iii) $\forall\vec{x}\in\mathbf{R}^n$, $\forall\vec{y}\in\mathbf{R}^n$ に対して

$$\|\vec{x}+\vec{y}\|\leq\|\vec{x}\|+\|\vec{y}\|$$ (**三角不等式**あるいは凸(とつ)不等式と呼ぶ).

この不等式で等号が成り立つための必要十分条件は、$\vec{x}$ と $\vec{y}$ が向きまで込めて同じ方向 (すなわち片方がもう一方の非負実数倍) であること。

三角不等式は、「三角形の任意の 2 辺の (長さの) 和は、残りの 1 辺より大きい」という初等幾何の定理の一般化であると考えられる。$\vec{x}$ と $\vec{y}$ が同じ方向&同じ向きであるとき、三角形がつぶれて、$\|\vec{x}+\vec{y}\|=\|\vec{x}\|+\|\vec{y}\|$ が成り立つ。

証明 (1), (2) は簡単だから、(3) のみ示す。証明すべき式の両辺は 0 以上だから、両辺の 2 乗を比較すればいい。

$$\begin{aligned}(\|\vec{x}\|+\|\vec{y}\|)^2-\|\vec{x}+\vec{y}\|^2&=(\|\vec{x}\|^2+2\,\|\vec{x}\|\,\|\vec{y}\|+\|\vec{y}\|^2)-(\vec{x}+\vec{y},\vec{x}+\vec{y})\\&=(\|\vec{x}\|^2+2\,\|\vec{x}\|\,\|\vec{y}\|+\|\vec{y}\|^2)-(\|\vec{x}\|^2+2(\vec{x},\vec{y})+\|\vec{y}\|^2)\\&=2(\|\vec{x}\|\,\|\vec{y}\|-(\vec{x},\vec{y}))\geq 0.\quad(\text{Schwarz の不等式による})\end{aligned}$$

等号成立の条件については読者に任せる。■

「三角形の任意の 2 辺の差は残りの 1 辺より小さい」という定理もある。これに対応する次の不等式がある。

例題 1.3.1 (逆三角不等式 (一般に通用する呼び方ではない))

$$\|\vec{x}-\vec{y}\|\geq\Big|\|\vec{x}\|-\|\vec{y}\|\Big|\quad(\vec{x},\,\vec{y}\in\mathbf{R}^n)$$

を示せ。

解答 (両辺共に正だから、自乗して比較しても良い[2]。) ここでは、三角不等式から導く。

$$\|\vec{x}\| = \|\vec{x}-\vec{y}+\vec{y}\| \le \|\vec{x}-\vec{y}\| + \|\vec{y}\|$$

から

$$\|\vec{x}\| - \|\vec{y}\| \le \|\vec{x}-\vec{y}\|. \tag{1.3}$$

$\vec{x}$ と $\vec{y}$ は任意であるから、入れ換えても成立する:

$$\|\vec{y}\| - \|\vec{x}\| \le \|\vec{y}-\vec{x}\|.$$

すなわち

$$-\|\vec{x}-\vec{y}\| \le \|\vec{x}\| - \|\vec{y}\|. \tag{1.4}$$

(1.3) と (1.4) をまとめると

$$-\|\vec{x}-\vec{y}\| \le \|\vec{x}\| - \|\vec{y}\| \le \|\vec{x}-\vec{y}\|.$$

ゆえに

$$\Big| \|\vec{x}\| - \|\vec{y}\| \Big| \le \|\vec{x}-\vec{y}\|. \quad \blacksquare$$

定義 1.3.8 ($\mathbf{R}^n$ のユークリッド距離) $\vec{x} \in \mathbf{R}^n$, $\vec{y} \in \mathbf{R}^n$ に対して、$d(\vec{x},\vec{y})$ を

$$d(\vec{x},\vec{y}) := \|\vec{x}-\vec{y}\| = \sqrt{\sum_{i=1}^{n}(x_i-y_i)^2}$$

で定義し、$\vec{x}$ と $\vec{y}$ の**距離** (**distance**) と呼ぶ。

命題 1.3.7 より、容易に次の命題が得られる。

命題 1.3.9 (距離の公理) $\mathbf{R}^n$ の距離 $d = d(\cdot,\cdot)$ は次の性質を満たす。

(d1) $\forall \vec{x} \in \mathbf{R}^n$, $\forall \vec{y} \in \mathbf{R}^n$ に対して

$$d(\vec{x},\vec{y}) \ge 0.$$

$\forall \vec{x} \in \mathbf{R}^n$, $\forall \vec{y} \in \mathbf{R}^n$ に対して

$$d(\vec{x},\vec{y}) = 0 \iff \vec{x} = \vec{y}.$$

(d2) $\forall \vec{x} \in \mathbf{R}^n$, $\forall \vec{y} \in \mathbf{R}^n$ に対して

$$d(\vec{x},\vec{y}) = d(\vec{y},\vec{x}).$$

(d3) $\forall \vec{x} \in \mathbf{R}^n$, $\forall \vec{y} \in \mathbf{R}^n$, $\forall \vec{z} \in \mathbf{R}^n$ に対して

$$d(\vec{x},\vec{y}) \le d(\vec{x},\vec{z}) + d(\vec{z},\vec{y}) \quad (\text{三角不等式}).$$

注意 1.3.10 せっかく $d(\cdot,\cdot)$ という記号を定義したのだけれど、$d(\vec{x},\vec{y})$ よりも $\|\vec{x}-\vec{y}\|$ の方が簡潔だから、以後この講義では使わない。(距離の公理 (d1), (d2), (d3) を見せるのが目的であった。) ■

[2] この場合、内積の性質を使って証明することができる。

1.4　1変数ベクトル値関数

1.4.1　1変数ベクトル値関数の極限の定義

ノルムの話をしたのは、極限の定義をするためである。ここではまず 1 変数ベクトル値関数の極限の定義をしよう。

極限を用いると、連続性、微分可能性、微分係数、導関数、k 回微分可能性、k 階微分係数 $\vec{f}^{(k)}(a)$、k 階導関数 $\vec{f}^{(k)}(x)$、C^k 級、C^∞ 級などの概念が定義できる。そのやり方は1変数の場合と (まったくと言って構わないほど) 同じである。

極限の定義　$\vec{f}(x) \to A$ とは、$\vec{f}(x)$ と $\vec{A}$ との距離が 0 に収束することと定義する。

定義 1.4.1 (ベクトル値関数の極限の定義) I は $\mathbf{R}$ の区間、$a \in \overline{I}$, $\vec{f}\colon I \to \mathbf{R}^m$, $\vec{A} \in \mathbf{R}^m$ とするとき、

$$\lim_{x\to a} \vec{f}(x) = \vec{A} \quad \overset{\text{def.}}{\Leftrightarrow} \quad \lim_{x\to a} \left\| \vec{f}(x) - \vec{A} \right\| = 0.$$

一見、lim の定義に lim を使っていて、循環論法をしているようだが、$\overset{\text{def.}}{\Leftrightarrow}$ の右辺は実数値関数の極限なので、すでにおなじみのものである。ε-δ 論法で書くと、

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in I \quad |x - a| < \delta \quad \Longrightarrow \quad \left\| \vec{f}(x) - \vec{A} \right\| < \varepsilon$$

ということになる。これは

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad (\forall x \in I : |x - a| < \delta) \qquad \left\| \vec{f}(x) - \vec{A} \right\| < \varepsilon$$

と書いても同じことである[3]。

閉包 (区間の場合)　$\overline{I}$ は I の閉包であるとする。閉包については、1年生のときに習ったと思うが、後で一般の場合 ($\mathbf{R}^n$ の部分集合) に正確な定義を与える (p. 26 を見よ)。ここでは、区間の端の点を加える、と覚えておけばよい。詳しく書くと、

- $I = (a,b)$, $[a,b)$, $(a,b]$, $[a,b]$ ($a, b \in \mathbf{R}$, $a < b$) の場合は、$\overline{I} = [a,b]$.
- $I = (a,\infty)$, $[a,\infty)$ ($a \in \mathbf{R}$) の場合は $\overline{I} = [a,\infty)$
- $I = (-\infty,b)$, $(-\infty,b]$ ($b \in \mathbf{R}$) の場合は $\overline{I} = (-\infty,b]$.

$a \in I$ でなく、$a \in \overline{I}$ としたのはなぜか？例で説明する。実変数の対数関数 $f(x) = \log x$ は (普通) $I := (0,\infty)$ を定義域とする。定義域 I に属する a, 例えば $a = 1$ に対して $\displaystyle\lim_{x\to a} f(x)$ を考えるのは自然である。

$$\lim_{x\to 1} \log x = \log 1 = 0.$$

$a = 0$ は定義域 I に属していないが ($\overline{I}$ には属していて)、$\displaystyle\lim_{x\to a} f(x)$ を考えることがある。

$$\lim_{x\to 0} \log x = -\infty.$$

しかし $\overline{I}$ に属さない $a = -1$ に対して、$\displaystyle\lim_{x\to a} f(x)$ を考えるのはナンセンスである。定義域 I に属する x によって近づくためには、a は $\overline{I}$ に属する必要がある。

[3] $\forall x\ P(x) \Rightarrow Q(x)$ を、$(\forall x\colon P(x))\ Q(x)$ と書くのである。同様に $\exists x\ P(x) \land Q(x)$ を $(\exists x\colon P(x))\ Q(x)$ と書く。

もう一つ例をあげると、$I=(0,\infty)$, $f\colon I\ni x\mapsto \dfrac{\sin x}{x}\in\mathbf{R}$ に対して、

$$\lim_{x\to 0} f(x)=1.$$

$0\notin I$, $0\in\overline{I}$ であることに注意しよう。

例年誤解する人が多いので、一つだけ書いておくと、$\vec{f}$ が I で C^k 級とは、$\vec{f}'$, $\vec{f}''$, $\dots$, $\vec{f}^{(k)}$ が I で存在して、$\vec{f}^{(k)}$ が I で連続なことをいう。

ベクトル値関数の極限は成分関数の極限を考えれば良い

次の定理は予告してあった。

定理 1.4.2 I は $\mathbf{R}$ の区間、$\vec{f}\colon I\to\mathbf{R}^m$, $a\in\overline{I}$, $\vec{A}\in\mathbf{R}^m$ とするとき、

$$\lim_{x\to a}\vec{f}(x)=\vec{A}\quad\Leftrightarrow\quad \forall i\in\{1,\dots,m\}\quad \lim_{x\to a} f_i(x)=A_i.$$

ただし $\begin{pmatrix} f_1\\ \vdots\\ f_m\end{pmatrix}:=\vec{f}$, $\begin{pmatrix} A_1\\ \vdots\\ A_m\end{pmatrix}:=\vec{A}$ とおいた。

この定理の証明には、次の補題が用いられる。

補題 1.4.3 任意の $\vec{x}=\begin{pmatrix} x_1\\ \vdots\\ x_m\end{pmatrix}\in\mathbf{R}^m$ に対して、

$$\max_{j=1,\dots,m}|x_j|\le\|\vec{x}\|\le\sum_{j=1}^m |x_j|\,.$$

特に $\forall j\in\{1,\dots,m\}$ に対して、$|x_j|\le\|\vec{x}\|$.

証明　任意の j に対して、

$$|x_j|=\sqrt{x_j^2}\le\sqrt{x_1^2+\cdots+x_m^2}=\|\vec{x}\|\,.$$

ゆえに

$$\max_{j=1,\dots,m}|x_j|\le\|\vec{x}\|\,.$$

一方 $\vec{e}_j$ を第 j 成分が 1 で、他の成分は 0 であるような単位ベクトルとすると、$\vec{x}=x_1\vec{e}_1+\cdots+x_m\vec{e}_m$ であるから、

$$\begin{aligned}\|\vec{x}\|&\le\|x_1\vec{e}_1\|+\cdots+\|x_m\vec{e}_m\|=|x_1|\,\|\vec{e}_1\|+\cdots+|x_m|\,\|\vec{e}_m\|\\ &=|x_1|\cdot 1+\cdots+|x_m|\cdot 1=\sum_{j=1}^m|x_j|\,.\blacksquare\end{aligned}$$

定理の証明　($\Rightarrow$) 任意の $j \in \{1,\dots,m\}$ に対して、$|f_j(x) - A_j| \le \left\| \vec{f}(x) - \vec{A} \right\|$ であるから、$\lim_{x\to a} \vec{f}(x) = A$ であれば、$\lim_{x\to a} f_j(x) = A_j$.

($\Leftarrow$) $\forall \varepsilon > 0$ に対して、$\exists \delta_1,\dots,\delta_m > 0$ s.t.

$$\forall j \in \{1,\cdots,m\} \quad |x-a| \le \delta_j \quad \Longrightarrow \quad |f_j(x) - A_j| < \frac{\varepsilon}{m}.$$

$\delta := \min\{\delta_1,\dots,\delta_m\}$ とおくと、$\delta > 0$ で、$|x-a| < \delta$ であれば、$|f_j(x) - A_j| < \dfrac{\varepsilon}{m}$ $(i = 1,\dots,m)$ であるから、

$$\left\| \vec{f}(x) - \vec{A} \right\| \le \sum_{j=1}^{m} |f_j(x) - A_j| < \sum_{j=1}^{m} \frac{\varepsilon}{m} = \varepsilon. \blacksquare$$

1.4.2　連続性

$\vec{f}\colon I \to \mathbf{R}^m$ が $a \in I$ で**連続** (continuous at a) であるとは、

$$\lim_{x\to a} \vec{f}(x) = \vec{f}(a)$$

が成り立つことをいう。

$\vec{f}\colon I \to \mathbf{R}^m$ が I で**連続** (continuous on I) であるとは、$\forall x \in I$ に対して、$\vec{f}$ が x で連続であることをいう。

上の定理から、以下のような系が得られる。

$\vec{f}$ が a (あるいは I) で連続であるためには、$\forall j \in \{1,\dots,m\}$ に対して、f_j が a (あるいは I) で連続なことが必要十分である。

1.4.3　微分可能性

$\vec{f}\colon I \to \mathbf{R}^m$ が $a \in I$ で**微分可能** (differentiable) であるとは、

$$\lim_{h\to 0} \frac{1}{h}\left(\vec{f}(a+h) - \vec{f}(a)\right) = \lim_{h\to 0} \begin{pmatrix} \frac{f_1(a+h)-f_1(a)}{h} \\ \vdots \\ \frac{f_m(a+h)-f_m(a)}{h} \end{pmatrix}$$

が存在することをいう。そのとき、その極限を $\vec{f}'(a)$ と表し、$\vec{f}$ の a における**微分係数** (the derivative of f at a) と呼ぶ。

$\vec{f}\colon I \to \mathbf{R}^m$ が I で**微分可能** (differentiable on I) であるとは、$\forall x \in I$ に対して、$\vec{f}$ が x で微分可能なことをいう。そのとき、関数 $I \ni x \mapsto \vec{f}'(x) \in \mathbf{R}^m$ を $\vec{f}$ の**導関数** (the derivative of f, the derivative function of f) と呼び、$\vec{f}'$ で表す。

$\vec{f}$ が a (あるいは I) で微分可能であるためには、$\forall i \in \{1,\dots,m\}$ に対して、f_i が a (あるいは I) で微分可能なことが必要十分条件である。

また $\vec{f}'(a) = \begin{pmatrix} f_1'(a) \\ \vdots \\ f_m'(a) \end{pmatrix}$ が成り立つ。

1.4.4 $\mathbf{R}^m$ 内の曲線

$\mathbf{R}$ の区間 I で定義された連続関数 $\vec{f}\colon I \to \mathbf{R}^m$ があるとき、$\vec{f}$ の値域 ($\vec{f}$ による I の像)

$$\vec{f}(I) = \left\{\vec{f}(t); t \in I\right\}$$

は $\mathbf{R}^m$ の部分集合であるが、直観的には曲線であると考えられる (図を描こう)。

数学では、$\mathbf{R}^m$ 内の**曲線** (curve) とは、$\mathbf{R}$ のある区間 I から $\mathbf{R}^m$ への連続写像のことであると定義する (場合が多い)。

上のように定義した連続曲線の中には、「曲線」らしくないものも含まれている。

例 1.4.4 (Peano 曲線 (the Peano curve)) 連続曲線 $\vec{f}\colon I = [0,1] \to \mathbf{R}^2$ で、像が正方形である、すなわち

$$\vec{f}(I) = [0,1] \times [0,1]$$

が成り立つようなものが存在する (平面や空間を「充填する」曲線 (space-filling curve) については、ザーガン [3] を見よ)。■

微分係数 $\vec{f}'(a)$ の意味

$\vec{f}'(a)$ が存在し、$\vec{f}'(a) \neq \vec{0}$ であれば、それは、$\vec{f}$ を曲線と考えたときの、$\vec{f}(a)$ における接線の方向を表すベクトルである。

例 1.4.5 $I := \mathbf{R}$, $\vec{f}(t) := \begin{pmatrix} t \\ t^2 \end{pmatrix}$, $a = 1$ とする。$\vec{f}'(a) = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ であるが、これは確かに $\vec{f}(a) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ における接線の方向ベクトルである。実際、この曲線は、関数 $F(x) := x^2$ のグラフ $y = F(x)$ であり、$x = 1$ における接線の傾きは $F'(1) = 2$ であり、確かに $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ は接線の方向を与えるベクトルである。■

$\vec{f}'(a) = \vec{0}$ である場合は、たとえ $\vec{f}$ が C^1 級であっても、接線が引けない (存在しない) 場合もある。

例 1.4.6 $\vec{f}(t) := \begin{pmatrix} t^3 \\ t^2 \end{pmatrix}$ は、$t = 0$ のところで、とがった曲線となっている (関数 $y = |x|^{2/3}$ のグラフ)。■

1.4.5 C^k 級

I は $\mathbf{R}$ の区間で、$\vec{f}\colon I \to \mathbf{R}^m$ とする。

- $k \in \mathbf{N}$ とする。$\vec{f}$ が C^k **級**とは、$\vec{f}$ が k 回微分可能で、$\vec{f}^{(k)}$ が連続であることをいう。
- $\forall k \in \mathbf{N}$ に対して $\vec{f}$ が C^k 級であるとき、$\vec{f}$ は C^∞ 級であるという。
- $\vec{f}$ が連続であるとき、$\vec{f}$ は C^0 **級**であるという。

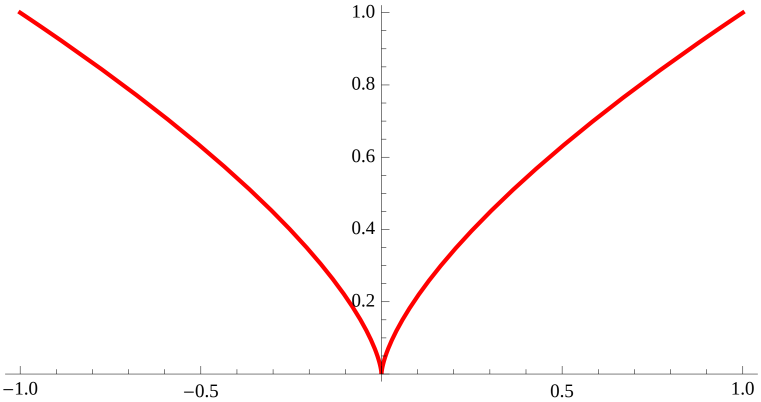


図 1.1: `ParametricPlot[{t^3,t^2},{t,-1,1},PlotStyle->{Thick,Red}]`

もちろん、$\vec{f}$ が C^k 級であることは、各 f_j が C^k 級であることと同値である。

切なる願い: 間違えるな

> 1 変数関数 $\vec{f}$ が C^1 級とは、$\vec{f}$ が 1 回微分可能かつ連続であることでは**ない！**正しくは、$\vec{f}$ が 1 回微分可能かつ $\vec{f}'$ が連続であることである。

後で、多変数関数について、C^1 級という言葉を定義するが、そのときは注意が必要になる。上のだけを丸暗記して済ませないように。

1.5 問題演習

問 1.5.1 I を $\mathbf{R}$ の区間、$\vec{f}\colon I \to \mathbf{R}^n$, $\vec{g}\colon I \to \mathbf{R}^n$ とする。
(1) $\vec{f}$ と $\vec{g}$ がともに微分可能であるならば

$$\frac{d}{dt}\left(\vec{f}(t), \vec{g}(t)\right) = \left(\vec{f}\,'(t), \vec{g}(t)\right) + \left(\vec{f}(t), \vec{g}\,'(t)\right) \quad (t \in I)$$

が成り立つことを示せ。
(2) 質点が等速運動するならば (つまり時刻 t における位置を $\vec{f}(t)$ と表すとき、$\left\|\vec{f}\,'(t)\right\|$ が定数関数となる)、速度と加速度はつねに直交することを示せ。

問題への取り組みかた　まず一般論として、

- 関係ありそうな定義、定理を思い出す。ノートを探す。思い出して書いてみる。
  面倒くさがらずに、使えそうなことを頭の中で並べて見るか (頭に自信がある場合)、一つの紙の上に書いて並べてみるのが良い。
- この講義に限り「成分で書いてみる」というのも有効である。
  (多次元化がテーマであるが、1 次元の場合から多次元の場合を証明、というのがよくあるパターンであるので。)
- 似ているものの証明を思い出す。

例えば、

$$(\bigstar) \qquad \forall \vec{x}, \vec{y}, \vec{z} \in \mathbf{R}^n \quad (\vec{x} + \vec{y}, \vec{z}) = (\vec{x}, \vec{z}) + (\vec{y}, \vec{z})$$

をどのように証明するか、説明を補足しておこう。

$$\vec{x}=\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix},\quad \vec{y}=\begin{pmatrix}y_1\\y_2\\\vdots\\y_n\end{pmatrix},\quad \vec{z}=\begin{pmatrix}z_1\\z_2\\\vdots\\z_n\end{pmatrix}$$

とおくと、

のように自分で書くことが、何でもないようでいて大事である。後は (★) の式の部分を x_j, y_j, z_j で表せば良い。

$$\vec{x}+\vec{y}=\begin{pmatrix}x_1+y_1\\x_2+y_2\\\vdots\\x_n+y_n\end{pmatrix}$$

であるから

$$(\vec{x}+\vec{y},\vec{z})=\sum_{j=1}^{n}(x_j+y_j)z_j=\sum_{j=1}^{n}(x_jz_j+y_jz_j)=\sum_{j=1}^{n}x_jz_j+\sum_{j=1}^{n}y_jz_j=(\vec{x},\vec{z})+(\vec{y},\vec{z})\,.$$

1.6 問題の解答

問 1.5.1

(1) $\begin{pmatrix}f_1\\\vdots\\f_n\end{pmatrix}:=\vec{f},\ \begin{pmatrix}g_1\\\vdots\\g_n\end{pmatrix}:=\vec{g}$ とするとき、

$$\left(\vec{f}(t),\vec{g}(t)\right)=\sum_{j=1}^{n}f_j(t)g_j(t)$$

であるから、(1 変数実数値関数の) 積の微分法を用いて、

$$\begin{aligned}\frac{d}{dt}\left(\vec{f}(t),\vec{g}(t)\right)&=\frac{d}{dt}\sum_{j=1}^{n}f_j(t)g_j(t)=\sum_{j=1}^{n}\frac{d}{dt}\left(f_j(t)g_j(t)\right)=\sum_{j=1}^{n}\left(f_j'(t)g_j(t)+f_j(t)g_j'(t)\right)\\&=\sum_{j=1}^{n}f_j'(t)g_j(t)+\sum_{j=1}^{n}f_j(t)g_j'(t)=\left(\vec{f}\,'(t),\vec{g}(t)\right)+\left(\vec{f}(t),\vec{g}\,'(t)\right).\end{aligned}$$

(2) 仮定から、$\exists C\in\mathbf{R}$ s.t. $\forall\in I\ \left\|\vec{f}'(t)\right\|=C$. ゆえに

$$\left(\vec{f}'(t),\vec{f}'(t)\right)=\left\|\vec{f}'(t)\right\|^2=C^2.$$

両辺を t で微分すると、(1) を用いて

$$\left(\vec{f}\,''(t),\vec{f}\,'(t)\right)+\left(\vec{f}\,'(t),\vec{f}\,''(t)\right)=0.$$

左辺は $2\left(\vec{f}''(t), \vec{f}'(t)\right)$ であるから、

$$\left(\vec{f}''(t), \vec{f}'(t)\right) = 0.$$

これは $\vec{f}''(t)$ と $\vec{f}'(t)$ が直交することを示す。■

(1) の別解　積の微分法の証明を思い出して、それをベクトル値関数化する (この方法は無限次元でも通用する)。

$$\begin{aligned}
&\frac{1}{h}\left[\left(\vec{f}(t+h), \vec{g}(t+h)\right) - \left(\vec{f}(t), \vec{g}(t)\right)\right] \\
&\quad = \frac{1}{h}\left[\left(\vec{f}(t+h), \vec{g}(t+h)\right) - \left(\vec{f}(t), \vec{g}(t+h)\right) + \left(\vec{f}(t), \vec{g}(t+h)\right) - \left(\vec{f}(t), \vec{g}(t)\right)\right] \\
&\quad = \frac{1}{h}\left[\left(\vec{f}(t+h) - \vec{f}(t), \vec{g}(t+h)\right) + \left(\vec{f}(t), \vec{g}(t+h) - \vec{g}(t)\right)\right].
\end{aligned}$$

ゆえに

$$\begin{aligned}
&\frac{1}{h}\left[\left(\vec{f}(t+h), \vec{g}(t+h)\right) - \left(\vec{f}(t), \vec{g}(t)\right)\right] - \left[\left(\vec{f}'(t), \vec{g}(t)\right) + \left(\vec{f}(t), \vec{g}'(t)\right)\right] \\
&\qquad = \left(\frac{\vec{f}(t+h) - \vec{f}(t)}{h} - \vec{f}'(t), \vec{g}(t+h)\right) + \left(\vec{f}'(t), \vec{g}(t+h) - \vec{g}(t)\right) \\
&\qquad + \left(\vec{f}(t), \frac{\vec{g}(t+h) - \vec{g}(t)}{h} - \vec{g}'(t)\right).
\end{aligned}$$

$$\begin{aligned}
&\left|\frac{1}{h}\left[\left(\vec{f}(t+h), \vec{g}(t+h)\right) - \left(\vec{f}(t), \vec{g}(t)\right)\right] - \left[\left(\vec{f}'(t), \vec{g}(t)\right) + \left(\vec{f}(t), \vec{g}'(t)\right)\right]\right| \\
&\qquad \leq \left\|\frac{\vec{f}(t+h) - \vec{f}(t)}{h} - \vec{f}'(t)\right\| \|\vec{g}(t+h)\| + \left\|\vec{f}'(t)\right\| \|\vec{g}(t+h) - \vec{g}(t)\| \\
&\qquad + \left\|\vec{f}(t)\right\| \left\|\frac{\vec{g}(t+h) - \vec{g}(t)}{h} - \vec{g}'(t)\right\| \\
&\qquad \to 0 \cdot \|\vec{g}(t)\| + \left\|\vec{f}'(t)\right\| \cdot 0 + \left\|\vec{f}(t)\right\| \cdot 0 = 0.
\end{aligned}$$

ただし

- $h \to 0$ のとき $\vec{F}(t+h) \to \vec{A}$ ならば $\left\|\vec{F}(t+h)\right\| \to \left\|\vec{A}\right\|$
- $\vec{g}$ が t で微分可能ならば $\vec{g}(t+h) \to \vec{g}(t)$ $(h \to 0)$

ということを使った (その証明も 1 変数の場合と同様に出来る)。■

第2章　多変数関数

忘れないうちに言っておく　これまでベクトルは $\vec{x}$ のように矢印 $\vec{}$ をつけてきた。$\boldsymbol{x}$ のように太字で表す、という流儀もある。これからは少しサボって、単に x のように書くことにする。(ベクトルとその成分を混同して欲しくないときは、また $\vec{}$ をつけるかも知れない。)

2.1　多変数関数の極限

最初に記号から。$\vec{a} \in \mathbf{R}^n$, $r > 0$ に対して、

$$B(\vec{a}; r) := \{\vec{x} \in \mathbf{R}^n; \|\vec{x} - \vec{a}\| < r\}.$$

これを $\vec{a}$ を中心とする半径 r の**開球** (open ball) と呼ぶ。また、

$$\overline{B}(\vec{a}; r) := \{\vec{x} \in \mathbf{R}^n; \|\vec{x} - \vec{a}\| \le r\}.$$

これを $\vec{a}$ を中心とする半径 r の**閉球** (closed ball) と呼ぶ。

問　$n = 1, 2, 3$ の場合にどういう集合か図を描いて説明せよ。

定義 2.1.1 $\Omega \subset \mathbf{R}^n$ に対して、

$$\overline{\Omega} := \{\vec{x} \in \mathbf{R}^n; \forall \varepsilon > 0 \quad \Omega \cap B(\vec{x}; \varepsilon) \ne \emptyset\}$$

とおき、Ω の**閉包** (the closure of Ω) と呼ぶ。

直観的には、Ω に Ω の縁を加えたものである (後でもう少し詳しく説明する)。

例 2.1.2 $\Omega = (1, 2)$ のとき、$\overline{\Omega} = [1, 2]$.　$\Omega = (1, 2) \times (3, 4)$ のとき、$\overline{\Omega} = [1, 2] \times [3, 4]$. $\Omega = B(\vec{a}; r)$ のとき、$\overline{\Omega} = \overline{B}(\vec{a}; r)$. ■

問　例で述べたことを確かめよ。

$\displaystyle\lim_{\vec{x} \to \vec{a}} \vec{f}(x) = \vec{A}$ とはどういう意味だろうか? $\vec{x} \to \vec{a}$ の意味が問題であるが、結論から先に言うと、$\vec{x}$ と $\vec{a}$ との距離 $\|\vec{x} - \vec{a}\|$ が $\to 0$ となること、と約束する。

定義 2.1.3 (多変数関数の極限) $\Omega \subset \mathbf{R}^n$, $f \colon \Omega \to \mathbf{R}^m$, $\vec{a} \in \overline{\Omega}$, $\vec{A} \in \mathbf{R}^m$ とするとき、

$$\lim_{\vec{x} \to \vec{a}} \vec{f}(\vec{x}) = \vec{A} \quad \overset{\text{def.}}{\Leftrightarrow} \quad (\forall \varepsilon > 0) \quad (\exists \delta > 0) \quad (\forall \vec{x} \in \Omega : \|\vec{x} - \vec{a}\| < \delta) \quad \left\|\vec{f}(\vec{x}) - \vec{A}\right\| < \varepsilon.$$

「**収束する**」(convergent, converges),「**極限**」(limit) などの言葉は 1 変数関数と同様に用いる。

言い換えをいくつか。

- $\forall\varepsilon>0,\ \exists\delta>0,\ \forall\vec{x}\in\Omega\cap B(\vec{a};\delta)\ \left\|\vec{f}(\vec{x})-\vec{A}\right\|<\varepsilon.$
- $\forall\varepsilon>0,\ \exists\delta>0,\ \forall\vec{x}\in\Omega\quad \|\vec{x}-\vec{a}\|<\delta\Longrightarrow\left\|\vec{f}(\vec{x})-\vec{A}\right\|<\varepsilon.$

ここで図を描いて説明する。$\vec{x}$ が $\vec{a}$ に近づくというのは、1 変数の場合とは大きく様子が異なる。1 次元では、方向は 1 つしかなかったが、2 次元以上では、直線に沿った場合だけを考えても、無限に多くの方向が存在するし、曲線に沿って接近したりする場合もある。

記号の約束: A と B の差集合 $A\setminus B:=\{x\in A; x\notin B\}$.

例 2.1.4 (極限の存在する例) $\Omega:=\mathbf{R}^2\setminus\{(0,0)\}$, $f\colon\Omega\to\mathbf{R}$ を

$$f(x,y):=\frac{x^2y}{x^2+y^2}$$

で定める。実は

$$\lim_{(x,y)\to(0,0)}f(x,y)=0$$

である。実際、$(x,y)\in\Omega$, $(x,y)\to(0,0)$ とするとき

$$|f(x,y)-0|=\frac{x^2}{x^2+y^2}\cdot|y|\le\frac{x^2+y^2}{x^2+y^2}\cdot|y|=|y|\le\sqrt{x^2+y^2}=\|(x,y)-(0,0)\|\to 0.$$

(別解: 極座標を使うと、$f(x,y)=\dfrac{r^2\cos^2\theta\cdot r\sin\theta}{r^2}=r\cos^2\theta\sin\theta$ なので $|f(x,y)|\le|r|\to 0$ と出来て見通しが良い。) ■

問 $\displaystyle\lim_{(x,y)\to(0,0)}f(x,y)=0$ であることを、ε-δ 論法を使ってきちんと書け。

極限が存在しない例をあげる前に、一つ準備をしておくと便利である。

問 2.1.1 (関数の極限が存在すれば制限した関数の極限もそれに等しい) $\Omega\subset\mathbf{R}^n$, $f\colon\Omega\to\mathbf{R}^m$, $\Omega'\subset\Omega$, $a\in\overline{\Omega'}$, $A\in\mathbf{R}^m$ とするとき、

$$\lim_{x\to a}f(x)=A$$

であれば、

$$\lim_{\substack{x\to a\\ x\in\Omega'}}f(x)=A$$

であることを示せ。ただし、

$$\lim_{\substack{x\to a\\ x\in\Omega'}}f(x):=\lim_{x\to a}f|_{\Omega'}(x)\quad(f|_{\Omega'}\text{ は }f\text{ の }\Omega'\text{ への制限})$$

とする。

例 2.1.5 (極限の存在しない例) $\Omega:=\mathbf{R}^2\setminus\{(0,0)\}$, $f\colon\Omega\to\mathbf{R}$ を

$$f(x,y):=\frac{xy}{x^2+y^2}$$

で定める。

(i) 点 (x,y) を、x 軸に沿って $(0,0)$ に近づけると、

$$\lim_{\substack{(x,y)\to(0,0)\\ y=0}} f(x,y) = \lim_{x\to 0} f(x,0) = \lim_{x\to 0}\frac{x\cdot 0}{x^2+0^2} = \lim_{x\to 0} 0 = 0.$$

(ii) 点 (x,y) を、y 軸に沿って $(0,0)$ に近づけると、

$$\lim_{\substack{(x,y)\to(0,0)\\ x=0}} f(x,y) = \lim_{y\to 0} f(0,y) = \lim_{y\to 0}\frac{0\cdot y}{0^2+y^2} = \lim_{y\to 0} 0 = 0.$$

(iii) 点 (x,y) を、直線 $y=kx$ (ここで k はある実定数) に沿って $(0,0)$ に近づけると、

$$\lim_{\substack{(x,y)\to(0,0)\\ y=kx}} f(x,y) = \lim_{x\to 0} f(x,kx) = \lim_{x\to 0}\frac{x\cdot kx}{x^2+(kx)^2} = \lim_{x\to 0}\frac{k}{1+k^2} = \frac{k}{1+k^2}.$$

これは $k=0$ でない限り、0 ではない。

以上より、$\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ は存在しない。実際、もしも $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)=A$ となる A が存在すれば、

$$\lim_{x\to 0} f(x,0) = \lim_{y\to 0} f(0,y) = \lim_{x\to 0} f(x,kx) = A$$

となるはずだが、問 2.1.1 によって、$0=0=\dfrac{k}{1+k^2}=A$ となって矛盾が生じる。 ■

このような不定形の極限が重要かつ難しいが、その演習は後日にまわす。

問 2.1.2 $\Omega := \{(x,y)\in\mathbf{R}^2; x>0\wedge y>0\}$, $f\colon \Omega\ni(x,y)\mapsto x^y\in\mathbf{R}$ とするとき、

$$\lim_{x\to 0}\lim_{y\to 0} f(x,y),\quad \lim_{y\to 0}\lim_{x\to 0} f(x,y),\quad \lim_{(x,y)\to(0,0)} f(x,y)$$

を求めよ。

極限に関する色々な命題、多くはこれまで (1 変数関数の場合) と同様のものが成り立ち、同様に証明出来る。例えば $\displaystyle\lim_{x\to a} f(x)$, $\displaystyle\lim_{x\to a} g(x)$ が存在するならば (仮定を細かく書くことは省略するが)

$$\begin{aligned}
&\lim_{x\to a}(f(x)+g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x),\\
&\lim_{x\to a}(f(x)-g(x)) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x),\\
&\lim_{x\to a}(f(x)g(x)) = \lim_{x\to a} f(x)\lim_{x\to a} g(x),\\
&\lim_{x\to a}\frac{g(x)}{f(x)} = \frac{\lim_{x\to a} g(x)}{\lim_{x\to a} f(x)}\quad (\text{ただし } \lim_{x\to a} f(x)\neq 0 \text{ とする}).
\end{aligned}$$

証明は、$\vec{\ }$ や $\|\cdot\|$ をつけるだけ、というのが多い。

分母の極限が 0 である場合をクローズアップしてみよう。

問 2.1.3 $\displaystyle\lim_{x\to a} f(x)=0$ で、

$$\lim_{x\to a}\frac{g(x)}{f(x)}$$

が (有限な) 極限を持つならば、$\displaystyle\lim_{x\to a} g(x)=0$ であることを示せ。

命題 2.1.6 $\Omega \subset \mathbf{R}^n$, $f\colon \Omega \to \mathbf{R}$, $a \in \overline{\Omega}$ とする。

(1) $\displaystyle\lim_{x\to a} f(x) = \infty$ ならば $\displaystyle\lim_{x\to a} \frac{1}{f(x)} = 0.$

(2) $\forall x \in \Omega\ f(x) > 0$, $\displaystyle\lim_{x\to a} f(x) = 0$ ならば $\displaystyle\lim_{x\to a} \frac{1}{f(x)} = \infty.$

証明

(1) $\forall \varepsilon > 0$ に対して、$R := \dfrac{1}{\varepsilon}$ とおくと、$R > 0$. $\displaystyle\lim_{x\to a} f(x) = \infty$ より、

$$\exists \delta > 0 \quad \forall x \in B(a;\delta) \cap \Omega \quad f(x) > R(> 0).$$

このとき、

$$0 < \frac{1}{f(x)} < \frac{1}{R} = \varepsilon.$$

ゆえに $\displaystyle\lim_{x\to a} \frac{1}{f(x)} = 0.$

(2) $\forall R \in \mathbf{R}$ に対して、$\varepsilon := \dfrac{1}{|R|+1}$ とおくと、$\varepsilon > 0$. $\displaystyle\lim_{x\to a} f(x) = 0$ であるから、

$$\exists \delta > 0 \quad \forall x \in B(a;\delta) \cap \Omega \quad |f(x)| < \varepsilon.$$

このとき、($f(x) > 0$ に注意して)

$$\frac{1}{f(x)} = \frac{1}{|f(x)|} > \frac{1}{\varepsilon} = |R| + 1 > R.$$

ゆえに $\displaystyle\lim_{x\to a} \frac{1}{f(x)} = 0.$ ■

2.2 連続関数の定義と簡単な性質

定義 2.2.1 $\Omega \subset \mathbf{R}^n$, $\vec{f}\colon \Omega \to \mathbf{R}^m$ とする。
$\vec{f}$ が $\vec{a}$ で**連続** (continuous at a) であるとは、

$$\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a})$$

が成り立つをいう。

問 ε-δ 論法で書いてみよう。
(解答)

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall \vec{x} \in \Omega \quad \|\vec{x} - \vec{a}\| < \delta \quad \Longrightarrow \quad \left\|\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right\| < \varepsilon.$$

あるいは

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad (\forall \vec{x} \in \Omega : \|\vec{x} - \vec{a}\| < \delta) \quad \left\|\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right\| < \varepsilon.$$

あるいは

$$\forall \varepsilon > 0, \quad \exists \delta > 0, \quad \forall \vec{x} \in \Omega \cap B(\vec{a};\delta) \quad \left\|\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right\| < \varepsilon. ■$$

微積分で扱う多くの関数は、定義域全体で連続である。そのことを経済的に証明する方法を学ぼう。

連続関数を組み立てたものは連続である

連続関数を “組み立てたもの” は連続関数 (実は微分可能な関数を組み立てたものは微分可能な関数、のように他での「応用」がある考え方)

- $f\colon \Omega \to \mathbf{R}$, $g\colon \Omega \to \mathbf{R}$ がともに連続ならば、$f+g$, $f-g$, fg はいずれも Ω から $\mathbf{R}$ への連続関数。$g \neq 0$ (on Ω) ならば f/g も Ω から $\mathbf{R}$ への連続関数。
- 連続関数の和、差、内積、ノルム、実数値関数倍、実数値関数による商 $\vec f + \vec g$, $\vec f - \vec g$, $\left(\vec f, \vec g\right)$, $\left\|\vec f\right\|$, $k\vec f$, $\dfrac{1}{k}\vec f$ (ただし $k \neq 0$) も連続である。
- $f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ について、f が連続 $\iff$ すべての $i \in \{1, \ldots, m\}$ について f_i が連続。
- 連続関数の合成関数 $\vec g \circ \vec f$ は連続関数。
- n 変数実係数多項式は $\mathbf{R}^n$ 上の連続関数を定める: $f(x_1, \ldots, x_n) \in \mathbf{R}[x_1, \ldots, x_n]$ ならば、$\mathbf{R}^n \ni (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n) \in \mathbf{R}$ は連続である。
  簡潔のため、多項式 $f(x_1, \cdots, x_n)$ が定める関数を f と書くことを約束し、多項式の定める関数を「多項式関数」と呼ぶことにする。
- 高校生以来知っている (切れていないグラフが思い浮かべられる) 指数関数 $e^x = \exp x$, a^x (ただし $a > 0$, $a \neq 1$)、対数関数 $\log x$ ($x > 0$)、三角関数 $\cos x$, $\sin x$、$\sqrt{x}$ ($x \geq 0$), 冪乗関数 $x \mapsto x^\alpha$ ($x > 0$), n 乗根 $\sqrt[n]{x}$ ($x \in \mathbf{R}$ または $x \in [0, \infty)$) 絶対値 $|x|$ は、は、それらの定義域上で連続である。
- 大学に入ってから教わった逆三角函数 $\tan^{-1} x = \arctan x$, $\sin^{-1} x = \arcsin x$ ($x \in [-1, 1]$), $\cos^{-1} x = \arccos x$ ($x \in [-1, 1]$) もそれらの定義域上で連続である。

これらの関数の多くは C^∞ 級であることが分かり、証明も同様である (ただし $\sqrt{x}$ が C^∞ 級であるのは、$x > 0$ の範囲で、$x = 0$ を含めると成り立たなくなる、などの注意は必要である)。

例 2.2.2 $f(x, y) = x^2 + 2xy + 3y^2 + 4x + 5y + 6$ は 2 変数の多項式関数であるから、$\mathbf{R}^2$ 上の関数として連続である。

$\varphi(x, y) = \sin(x^2 + 2xy + 3y^2 + 4x + 5y + 6)$ は、$g(z) = \sin z$ とすると、$\varphi = g \circ f$. f も g も連続関数であるから、合成関数 φ は連続である。■

問 2.2.1 次の各関数が $\mathbf{R}^2$ で連続であることを示せ (理由を述べよ)。
(1) $f(x, y) = x^2 + \sqrt{2}xy + (\log 3)y^2 + \frac{\pi}{4}x + e^5 y + 6$ (2) $g(x, y) = \exp(3x + 2y + 1)$
(3) $h(x, y) = \dfrac{x^2 + 2x + 3}{x^2 + y^2 + 1}$ (4) $\varphi(x, y) = \log\left(1 + \sqrt{x^2 + y^2}\right)$ (5) $F(x, y) = \begin{pmatrix} x^3 - 3xy^2 \\ 3x^2 y - y^3 \end{pmatrix}$

上で連続関数を組み立てたものは連続と言い、いくつか定理を述べた。それらの証明をどうするか簡単に説明しよう。

ほとんどは、1 変数の場合の定理の証明を少し修正するだけで良い。
(絶対値 $|\cdot|$ をノルム $\|\cdot\|$ に変える程度である)

例えば、「 f と g が a で連続であれば、$f+g$ も a で連続である。」の証明をやってみよう。$f+g$ は

$$(f+g)(x)=f(x)+g(x)$$

で定まる関数であることを思い出しておく。

$$\lim_{x\to a}(f+g)(x)=\lim_{x\to a}\left(f(x)+g(x)\right)=\lim_{x\to a}f(x)+\lim_{x\to a}g(x)=f(a)+g(a)=(f+g)(a)$$

となるので、$f+g$ は a で連続である。

ここで、和の極限が極限の和になるという

$$\lim_{x\to a}\left(f(x)+g(x)\right)=\lim_{x\to a}f(x)+\lim_{x\to a}g(x)$$

という定理を用いた。これについて、前回さらっと言ったが、証明を見ておこう。

$$\lim_{x\to a}f(x)=A,\quad \lim_{x\to a}g(x)=B$$

とする。$\forall\varepsilon>0$ に対して、$\exists\delta_1>0$ s.t.

$$\forall x\in\Omega\cap B(a;\delta_1)\quad \|f(x)-A\|<\varepsilon/2.$$

同様に $\exists\delta_2>0$ s.t.

$$\forall x\in\Omega\cap B(a;\delta_2)\quad \|g(x)-B\|<\varepsilon/2.$$

$\delta:=\min\{\delta_1,\delta_2\}$ とおくと、$\delta>0$ であり、$\forall x\in\Omega\cap B(a;\delta)$ に対して、$x\in B(a;\delta_1)$ かつ $x\in B(a;\delta_2)$ であるので、$\|f(x)-A\|<\varepsilon/2$, $\|g(x)-B\|<\varepsilon/2$. ゆえに

$$\begin{aligned}\|f(x)+g(x)-(A+B)\|&=\|(f(x)-A)+(g(x)-B)\|\\&\le\|f(x)-A\|+\|g(x)-B\|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.\end{aligned}$$

多変数の多項式関数が連続というのはどう証明するのだろうか? 多項式は、“座標関数” と定数関数から (和と積で) 組み立てられることに気がつけば、次の補題に帰着される。

補題 2.2.3 (座標関数と定数関数の連続性) $n\in\mathbf{N}$ とする。

(1) $\forall j\in\{1,\cdots,n\}$ に対して、$\varphi_j\colon\mathbf{R}^n\to\mathbf{R}$ を $\varphi(x):=x_j$ $(x=(x_1,\cdots,x_n)^T\in\mathbf{R}^n)$ で定めるとき、φ_j は連続である。

(2) $\forall c\in\mathbf{R}$ に対して、$\psi\colon\mathbf{R}^n\to\mathbf{R}$ を $\psi(x):=c$ $(x\in\mathbf{R}^n)$ で定めるとき、ψ は連続である。

証明

(1) $\forall a\in\mathbf{R}^n$ に対して、

$$|\varphi_j(x)-\varphi_j(a)|=|x_j-a_j|\le\|x-a\|$$

であるから、$x\to a$ のとき $\varphi_j(x)\to\varphi_j(a)$.

(2) $\forall a\in\mathbf{R}^n$ に対して、

$$|\varphi_j(x)-\varphi_j(a)|=|c-c|=0$$

であるから、$x\to a$ のとき $\varphi_j(x)\to\varphi_j(a)$. ■

2.2.1　極限を求める問題

連続ならば⋯

試験の問題に出されることはあまり多くはないのだが、実際、微積に現れる多くの関数は連続である。連続関数の定義

$$「f \text{ が } a \text{ で連続} \overset{\text{def.}}{\Leftrightarrow} \lim_{x\to a} f(x) = f(a)」$$

より、明らかに「f が a で連続 $\Longrightarrow \lim_{x\to a} f(x) = f(a)$」が成り立つ。つまり連続関数については、極限の計算は単に関数値の計算 ($x = a$ を代入?) をするだけである。

微分の定義を考えると、分数関数の極限 $\displaystyle\lim_{x\to a} \frac{q(x)}{p(x)}$ が重要、ということが分かる。難しいのは $\displaystyle\lim_{x\to a} p(x) = 0$ の場合である。

分母が 0 に近づく、しかし分子はそうでないならば⋯

$f(x) = \dfrac{q(x)}{p(x)}$ で、$\displaystyle\lim_{x\to a} p(x) = 0$, $\displaystyle\lim_{x\to a} q(x) \neq 0$ (収束しないか、収束して極限が 0 でない) ならば、$\displaystyle\lim_{x\to a} f(x)$ は存在しない。証明は背理法による。つまり $\displaystyle\lim_{x\to a} f(x) = A$ であれば、

$$\lim_{x\to a} q(x) = \lim_{x\to a} \left(\frac{q(x)}{p(x)} \cdot p(x)\right) = \lim_{x\to a} (f(x) \cdot p(x)) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} p(x) = A \cdot 0 = 0$$

であり、仮定に反する。

というわけで、問題はいわゆる不定形 $\dfrac{0}{0}$ の場合、ということになる。

不定形

不定形 $\dfrac{0}{0}$ は難しい (でも大事)。これについては、必ずうまく行く、というやり方はない。

$\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ の場合に少し練習してみよう。(i) $y = kx$ 作戦, (ii) 極座標作戦などが有効である。

$y = kx$ **作戦**　簡単だけど役に立つ命題「$f\colon \Omega \to \mathbf{R}^m$, $a \in \overline{\Omega}$, $A \in \mathbf{R}^m$, $\displaystyle\lim_{x\to a} f(x) = A$ ならば、$\Omega' \subset \Omega$, $a \in \overline{\Omega'}$ となる Ω' に対して、$\displaystyle\lim_{\substack{x\in\Omega' \\ x\to a}} f(x) = A$.」に注目しよう (問 2.1.1)。

2 次元で $a = (0,0)$ の場合に、$\Omega' = \{(x,y) \in \Omega; y = kx\}$ として利用すると、

$$「\lim_{(x,y)\to(0,0)} f(x,y) = A \text{ ならば、} \forall k \in \mathbf{R}\ \lim_{\substack{y=kx \\ x\to 0}} f(x,y) = A.」$$

これから $\displaystyle\lim_{\substack{y=kx \\ x\to 0}} f(x,y)$ が存在しなかったり、存在しても k に依存したりすると、$\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ は存在しない (対偶というか、背理法で証明できる)。

「逆は必ずしも真ならず」で、$\displaystyle\lim_{\substack{y=kx \\ x\to 0}} f(x,y)$ が k によらない極限を持っていても、$\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ が存在するとは限らない。

(反例: $f(x,y) = \dfrac{x^2+y^2}{x+y}$ について、$\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ は存在しないが、直線 $y = kx$ にそった極限は 0 となる。)

しかし $\displaystyle\lim_{\substack{y=kx\\x\to 0}} f(x,y)$ が k によらない極限 A を持つことが示せたとして、何も言えないかというと、そうではない。もしも $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ が存在するならば、それは A でしかあり得ない、ことが分かる。

$\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = A$ が成り立つかどうか。$|f(x,y)-A|$ を考え、これを抑えるような量で、$(x,y)\to(0,0)$ のとき 0 に収束するようなものが見つかるか調べる、という作業になる。

極座標作戦　(工事中)

問 2.2.2 つぎの lim が存在するかどうか調べ、存在する場合はそれを求めよ。発散する場合も ∞ または $-\infty$ であるときはそれを指摘せよ。出来る限り根拠を書くこと。

(1) $\displaystyle\lim_{(x,y)\to(1,2)}(3x^2+4xy+5y^2)$ (2) $\displaystyle\lim_{(x,y)\to(0,1)}\frac{2+3xy}{4x^2+5y^2}$ (3) $\displaystyle\lim_{(x,y)\to(0,0)}\frac{1}{x^2+y^2}$ (4) $\displaystyle\lim_{(x,y)\to(0,0)}\frac{x+y}{\log(x^2+y^2)}$

(5) $\displaystyle\lim_{(x,y)\to(0,0)}\frac{x-y}{x+y}$ (6) $\displaystyle\lim_{(x,y)\to(0,0)}\frac{|x|}{\sqrt{x^2+y^2}}$ (7) $\displaystyle\lim_{(x,y)\to(0,0)}\frac{x^2y^2}{x^2+y^2}$ (8) $\displaystyle\lim_{(x,y)\to(0,0)}\frac{\sin(xy)}{xy}$.

問 2.2.3 つぎの関数が原点 $(0,0)$ で連続かどうか調べよ。

(1) $f(x,y)=\begin{cases}\dfrac{xy^2}{x^2+y^2} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$　(2) $f(x,y)=\begin{cases}\dfrac{x^2y^2}{x^2+y^2} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$

(3) $f(x,y)=\begin{cases}\dfrac{x^2-y^2}{x^2+y^2} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$　(4) $f(x,y)=\begin{cases}\dfrac{x+y}{\log\left(x^2+y^2\right)} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$

(5) $f(x,y)=\begin{cases}\dfrac{xy}{x+y} & (x+y\neq 0)\\ 0 & (x+y=0).\end{cases}$

2.3　連続関数と開集合、閉集合

1 変数関数の微積分を扱うのにも開集合、閉集合という概念を用いた。多変数関数を扱うには (さらにはこの先の数学では)、1 変数のとき以上に重要となる。それに焦点を合わせた、「集合・距離・位相 x」 $(x=1,2)$ という講義が用意されている。

ここでは多変数関数の微積分に必要な、最低限のことを用意する。

定義 2.3.1 (開集合, 閉集合) $A\subset\mathbf{R}^n$ とする。

(i) A が $\mathbf{R}^n$ の開集合であるとは、

$$\forall x\in A\quad \exists\varepsilon>0\quad \text{s.t.}\quad B(x;\varepsilon)\subset A$$

が成り立つことをいう

(ii) A が $\mathbf{R}^n$ の閉集合であるとは、$A^c:=\mathbf{R}^n\setminus A$ が $\mathbf{R}^n$ の開集合であることをいう。

次の 2 つの命題は既に習ったはずである (「集合・距離・位相 1 」でも出て来るはず)。

命題 2.3.2 (開集合系の公理) 次の (i), (ii), (iii) が成り立つ。

(i) $\emptyset$, $\mathbf{R}^n$ は $\mathbf{R}^n$ の開集合である。

(ii) 集合族 $\{U_\lambda\}_{\lambda\in\Lambda}$ において、$\forall\lambda\in\Lambda$ に対して U_λ が $\mathbf{R}^n$ の開集合ならば、$\displaystyle\bigcup_{\lambda\in\Lambda} U_\lambda$ は $\mathbf{R}^n$ の開集合である。

(iii) U_1, U_2 は $\mathbf{R}^n$ の開集合ならば、$U_1\cap U_2$ は $\mathbf{R}^n$ の開集合である。

証明 念のため、(iii) だけでも証明しておく。$x\in U_1\cap U_2$ とすると、$x\in U_1$ で、U_1 は開集合だから、$\exists\varepsilon_1>0$ s.t. $B(x;\varepsilon_1)\subset U_1$. 同様に $\exists\varepsilon_2>0$ s.t. $B(x;\varepsilon_2)\subset U_2$. $\varepsilon:=\min\{\varepsilon_1,\varepsilon_2\}$ とおくと、$\varepsilon>0$ で、

$$B(x;\varepsilon)\subset B(x;\varepsilon_1)\subset U_1,\quad B(x;\varepsilon)\subset B(x;\varepsilon_2)\subset U_2$$

であるから、$B(x;\varepsilon)\subset U_1\cap U_2$. ゆえに $U_1\cap U_2$ は $\mathbf{R}^n$ の開集合である。■

命題 2.3.3 (閉集合系の公理) (1) $\emptyset$, $\mathbf{R}^n$ は $\mathbf{R}^n$ の閉集合である。(2) $\{U_\lambda\}_{\lambda\in\Lambda}$ において、$\forall\lambda\in\Lambda$ U_λ が $\mathbf{R}^n$ の閉集合ならば、$\displaystyle\bigcap_{\lambda\in\Lambda} U_\lambda$ は $\mathbf{R}^n$ の閉集合である。(3) U_1, U_2 は $\mathbf{R}^n$ の閉集合ならば、$U_1\cup U_2$ は $\mathbf{R}^n$ の閉集合である。

証明 命題 2.3.2 の系である。■

次の命題は非常に便利である。

命題 2.3.4 (とても便利: 連続関数による逆像の開集合、閉集合の判定) $f\colon\mathbf{R}^n\to\mathbf{R}$ が連続関数、$a,b,c\in\mathbf{R}$ とするとき、以下が成立する。

(1) $\{x\in\mathbf{R}^n; f(x)>a\}$, $\{x\in\mathbf{R}^n; f(x)<b\}$, $\{x\in\mathbf{R}^n; a<f(x)<b\}$, $\{x\in\mathbf{R}^n; f(x)\neq c\}$ は $\mathbf{R}^n$ の開集合である。

(2) $\{x\in\mathbf{R}^n; f(x)\geq a\}$, $\{x\in\mathbf{R}^n; f(x)\leq b\}$, $\{x\in\mathbf{R}^n; a\leq f(x)\leq b\}$, $\{x\in\mathbf{R}^n; f(x)=c\}$ は $\mathbf{R}^n$ の閉集合である。

証明

(1) $A:=\{x\in\mathbf{R}^n; f(x)>a\}$ とする。$x_0\in A$ とすると、$f(x_0)>a$. $\varepsilon:=f(x_0)-a$ とおくと、$\varepsilon>0$. f は x_0 で連続だから、

$$\exists\delta>0\quad(\forall x\in\mathbf{R}^n:\|x-x_0\|<\delta)\quad|f(x)-f(x_0)|<\varepsilon.$$

ゆえに $\forall x\in B(x_0;\delta)$ に対して、$|f(x)-f(x_0)|<\varepsilon$ であるから、$-\varepsilon<f(x)-f(x_0)<\varepsilon$. ゆえに $f(x)>f(x_0)-\varepsilon=f(x_0)-(f(x_0)-a)=a$. ゆえに $x\in A$. これは $B(x_0;\delta)\subset A$ を意味している。ゆえに A は $\mathbf{R}^n$ の開集合である。

$\{x\in\mathbf{R}^n; f(x)<b\}$ も同様に証明できる。あるいは $F(x):=b-f(x)$ とおくと、$\{x\in\mathbf{R}^n; F(x)>0\}$ であることから上で証明したことに帰着される。

$\{x\in\mathbf{R}^n; a<f(x)<b\}$ と $\{x\in\mathbf{R}^n; f(x)\neq c\}$ については、

$$\{x\in\mathbf{R}^n; a<f(x)<b\}=\{x\in\mathbf{R}^n; f(a)>a\}\cap\{x\in\mathbf{R}^n; f(x)<b\},$$
$$\{x\in\mathbf{R}^n; f(x)\neq c\}=\{x\in\mathbf{R}^n; f(x)<c\}\cup\{x\in\mathbf{R}^n; f(x)>c\}$$

と、命題2.3.2 の (iii), (ii) による。(後者は、あるいは $\{x \in \mathbf{R}^n; f(x) \neq c\} = \{x \in \mathbf{R}^n; |f(x) - c| > 0\}$ の方が拡張性があって良いかも。)

(2) $A := \{x \in \mathbf{R}^n; f(x) \geq a\}$ とするとき、

$$A^c = \mathbf{R}^n \setminus A = \{x \in \mathbf{R}^n; f(x) < a\}$$

は、(1) の 2 番目より $\mathbf{R}^n$ の開集合である。ゆえに A は $\mathbf{R}^n$ の閉集合である。$\{x \in \mathbf{R}^n; f(x) \leq b\}$ が閉集合であることも同様に証明できる。

また $\{x \in \mathbf{R}^n; a \leq f(x) \leq b\}$ については、

$$\{x \in \mathbf{R}^n; a \leq f(x) \leq b\} = \{x \in \mathbf{R}^n; f(a) \geq a\} \cap \{x \in \mathbf{R}^n; f(x) \leq b\}$$

と、二つの閉集合の共通部分になっているから、命題 2.3.3 の (ii) により、閉集合である。

$B := \{x \in \mathbf{R}^n; f(x) = c\}$ の補集合

$$B^c = \{x \in \mathbf{R}^n; f(x) = c\}^c = \{x \in \mathbf{R}^n; f(x) \neq c\}$$

は、(1) より $\mathbf{R}^n$ の開集合であるから、B は閉集合である。 ■

「開球は開集合である」、「閉球は閉集合である」というのは有名で、直接的な計算に基づく証明も良く知られているが、上の命題と系を使うと、次のように見通し良く証明できる。

命題 2.3.5 (1) $\mathbf{R}^n$ の開球は開集合である。(2) $\mathbf{R}^n$ の閉球は閉集合である。

証明

(1) $a \in \mathbf{R}^n$, $r > 0$ とするとき、

$$B(a; r) = \{x \in \mathbf{R}^n; \|x - a\| < r\}$$

は、$f \colon \mathbf{R}^2 \ni (x, y) \mapsto \|x - a\|^2 = \sum_{j=1}^{n} (x_j - a_j)^2 \in \mathbf{R}$ (これは多項式関数だから連続) と $b := r^2$ を用いて、

$$B(a; r) = \{x \in \mathbf{R}^n; f(x) < b\}$$

と表されるので、開球 $B(a; r)$ は $\mathbf{R}^n$ の開集合である。

(2) 同様に

$$\overline{B}(a; r) = \{x \in \mathbf{R}^n; \|x - a\| \leq r\} = \{x \in \mathbf{R}^n; f(x) \leq b\}$$

であるから、閉球 $B(a; r)$ は $\mathbf{R}^n$ の閉集合である。 ■

問 2.3.1 $a \in \mathbf{R}^n$ とするとき、$A := \{a\}$ は $\mathbf{R}^n$ の閉集合である ($\mathbf{R}^n$ のシングルトン (単元集合) は閉集合である) ことを示せ。

(ヒント: この問題をここにおいたのには、理由がある。「開球 $B(a; r)$ は開集合、閉球 $\overline{B}(a; r)$ は閉集合」という定理を紹介した。そこでは球であるために $r > 0$ という仮定をおいたが、$r = 0$ の場合も、証明は成立する。$r = 0$ のときに $B(a; r)$, $\overline{B}(a; r)$ が何になるか、考えてみよう。)

例 2.3.6 (開区間が開集合であることの別証明) $\mathbf{R}$ の開区間というのは、(a,b), $(-\infty,b)$, (a,∞), $(-\infty,\infty)$ (ここで a, b は $a<b$ を満たす実数) のいずれかのことで、これらが開集合であることは証明済みであるが、次のようにも理解できる。$f\colon \mathbf{R}^1 \to \mathbf{R}$ を $f(x)=x$ で定めると、これは多項式関数なので連続で、

$$(a,b)=\left\{x\in\mathbf{R}^1; a<f(x)<b\right\},\quad (-\infty,b)=\left\{x\in\mathbf{R}^1; f(x)<b\right\},\quad (a,\infty)=\left\{x\in\mathbf{R}^1; f(x)>a\right\}$$

は命題 2.3.4 $\mathbf{R}^1$ の開集合である。

同様に

$$[a,b]=\left\{x\in\mathbf{R}^1; a\le f(x)\le b\right\},\quad (-\infty,b]=\left\{x\in\mathbf{R}^1; f(x)\le b\right\},\quad [a,\infty)=\left\{x\in\mathbf{R}^1; f(x)\ge a\right\}$$

は命題 2.3.4 $\mathbf{R}^1$ の閉集合である。■

例 2.3.7 第 1 象限 $A=\{(x,y)\in\mathbf{R}^2; x>0, y>0\}$ は、$f,g\colon\mathbf{R}^2\to\mathbf{R}$ を $f(x,y)=x$, $g(x,y)=y$ で定めると、

$$A=A_1\cap A_2,\quad A_1=\left\{(x,y)\in\mathbf{R}^2; f(x,y)>0\right\},\quad A_2=\left\{(x,y)\in\mathbf{R}^2; g(x,y)>0\right\}$$

と書ける。命題 2.3.4 から A_1, A_2 は $\mathbf{R}^1$ の開集合であり、命題 2.3.2 から $A_1\cap A_2$ も $\mathbf{R}^1$ の開集合である。ゆえに A は $\mathbf{R}^1$ の開集合である。■

問 2.3.2 $\mathbf{R}^2$ における次の各集合について、(a) 図示できる場合は図示せよ, (b) 開集合である場合は証明せよ, (c) 閉集合である場合は証明せよ[1]。
(1) $\emptyset$ (2) $\mathbf{R}^2$ (3) $\{(0,0)\}$ (4) $\{(0,0),(1,1)\}$ (5) $(1,2)\times(3,4)$ (6) $[1,2]\times(3,4)$
(7) $[1,2]\times[3,4]$ (8) $\{(x,y); 5<x^2+y^2<6\}$ (9) $(0,\infty)\times(0,\infty)$ (10) $\{(x,y); x^3\le y\le x^2\}$
(11) $\mathbf{R}^2\setminus\{(0,0)\}$.

問 2.3.3 (1) 前問の集合のうち、開集合でないものについて、それが開集合でないことを証明せよ。(2) 前問の集合のうち、閉集合でないものについて、それが閉集合でないことを証明せよ。

問 2.3.4 U, V をそれぞれ $\mathbf{R}^n$, $\mathbf{R}^m$ の開集合、$f\colon U\to V$ を連続関数とする。このとき $W\subset V$ なる任意の開集合 W に対して、$f^{-1}(W):=\{x\in U; f(x)\in W\}$ は $\mathbf{R}^n$ の開集合となることを証明するため、以下の空欄を埋めよ。「任意の $a\in$ $\boxed{\text{ア}}$ をとると、$a\in U$ かつ $f(a)\in$ $\boxed{\text{イ}}$. $\boxed{\text{イ}}$ は $\boxed{\text{ウ}}$ であるから、$\exists\varepsilon>0$ s.t. $B(f(a);\varepsilon)\subset$ $\boxed{\text{イ}}$ (ここで $B(\alpha\backslash r)$ は中心 α, 半径 r の開球を表す記号). f の連続性から $\boxed{\text{エ}}$ $\delta>0$ s.t. $\|x-a\|<\delta \Longrightarrow x\in U$ かつ $\|f(x)-f(a)\|<\varepsilon$. ゆえに $f(B(a;\delta))\subset B(f(a);\varepsilon)\subset W$ となるが、これから $B(a;\delta)\subset$ $\boxed{\text{オ}}$. ゆえに $f^{-1}(W)$ は開集合である。」

問 2.3.5 $F\subset\mathbf{R}^n$ で $f\colon F\to\mathbf{R}$ が連続とする。(1) $\{x\in F; f(x)\ge 0\}$ は $\mathbf{R}^n$ の閉集合とは限らないことを示せ。(2) F が $\mathbf{R}^n$ の閉集合であるとき、$\{x\in F; f(x)\ge 0\}$ は $\mathbf{R}^n$ の閉集合であることを示せ。

[1]開集合、または閉集合である場合、上で説明したやり方を使って証明できる。そうでない場合はその証明をするため、定義に戻ったりする必要があるが、それは今回要求しない。

2.4 連続関数の重要な性質

次の 3 つの定理は重要である。

(a) 中間値の定理

(b) 「$\mathbf{R}^n$ の有界閉集合 K 上の実数値連続関数 $f\colon K \to \mathbf{R}$ は、必ず最大値を持つ」 (Weierstrass) — この講義では「Weierstrass の最大値定理」と呼ぶことにする[2]。

(c) 「$\mathbf{R}^n$ の有界閉集合 K 上の連続関数 $f\colon K \to \mathbf{R}^m$ は、K で一様連続である」 (Weierstrass, Heine)

(c) は積分論で重要な役割を果たすが、この『多変数の微分積分学 1 』では必要ないので、とりあえず後回しにする。

2.4.1 中間値の定理

1 変数の中間値の定理を思い出そう。

定理 2.4.1 (1 変数実数値連続関数に関する中間値の定理) $f\colon [a,b] \to \mathbf{R}$ が連続で、$k \in \mathbf{R}$ は $f(a)$, $f(b)$ の間にある (i.e. $f(a) < f(b)$ の場合 $f(a) < k < f(b)$, $f(a) > f(b)$ の場合 $f(a) < k < f(b)$) ならば、

$$\exists c \in (a,b) \text{ s.t. } f(c) = k.$$

多変数関数では、区間 $[a,b]$ をどのように一般化するかが問題である。結論を先に言うと、ある意味で区間を一般化した「連結集合」を用いる。

定義 2.4.2 (連結集合) $\Omega \subset \mathbf{R}^n$ とするとき、Ω が**連結 (connected)** であるとは、Ω 内の任意の 2 点 x, y に対して、Ω 内の連続曲線で x と y を結ぶものが存在する、すなわち

$$\forall x, y \in \Omega \quad \exists \varphi\colon [0,1] \to \Omega \text{ 連続 s.t. } \varphi(0) = x \text{ かつ } \varphi(1) = y$$

が成り立つことをいう。

注意 2.4.3 (連結性の定義について) 実は、一般の位相空間論においては、連結性は、上とは違った (あまり直観的でない) やり方で定義される。上の定義の条件を満足する集合は、**弧状連結**(こじょうれんけつ) **(arcwise connected)** と呼ばれるのが普通である。しかし、$\mathbf{R}^n$ の開集合においては、連結=弧状連結なので、ここでは簡単で直観的な定義法を採用した。 ■

$\Omega \subset \mathbf{R}$ とするとき、Ω が連結であるためには、Ω が区間であることが必要十分である。

問 そのことを証明せよ。

(区間が (弧) 連結であるのは簡単に分かる。逆に A が弧連結であるとすると、$\forall x, y \in A$ に対して、$[x,y] \subset A$ であることがすぐ導かれる。…といういい加減な話をしました。)

[2] この定理はとても重要なため、一言で呼べると便利であるので、定着した呼び方ではないけれども、あえて名前をつけることにする。

定理 2.4.4 ((多変数関数に関する) 中間値の定理) Ω は $\mathbf{R}^n$ の連結な部分集合、$f\colon \Omega \to \mathbf{R}$ は連続ならば、次のことが成り立つ。

$$a \in \Omega,\ b \in \Omega,\ f(a) < k < f(b) \Longrightarrow \exists c \in \Omega \text{ s.t. } f(c) = k.$$

(板書の図が大事なのです…手書きのを取り込むかな？)

証明 Ω が連結であるという仮定から、a と b を結ぶ Ω 内の曲線 φ が取れる:

$$\varphi\colon [\alpha, \beta] \to \Omega \text{ 連続}, \quad \varphi(\alpha) = a, \quad \varphi(\beta) = b.$$

このとき $f \circ \varphi\colon [\alpha, \beta] \to \mathbf{R}$ は、連続関数の合成なので連続関数であり、

$$f \circ \varphi(\alpha) = f(\varphi(\alpha)) = f(a) < k < f(b) = f(\varphi(\beta)) = f \circ \varphi(\beta).$$

ゆえに 1 変数関数の中間値の定理から

$$\exists \gamma \in (\alpha, \beta) \text{ s.t. } f \circ \varphi(\gamma) = k.$$

そこで $c := \varphi(\gamma)$ とおけば、$f(c) = k$ となる。■

2.4.2 Weierstrass の最大値定理

楽屋裏で考えたこと

悩ましいな。聴いている学生の理解を無視して、とりあえず最速で証明してしまう、というシナリオもありうる。

1. $\mathbf{R}$ での Bolzano-Weierstrass の定理。習ったはず。一つの証明のあらすじは、区間縮小法により Cauchy 列であるような部分列が取れる。$\mathbf{R}$ の完備性により、$\mathbf{R}$ で収束する。

2. $\mathbf{R}^n$ での Bolzano-Weierstrass の定理。一つの証明のあらすじ：第 1 成分の作る数列は有界数列なので、それが収束するような部分列を取れる。さらに部分列を取って、第 2 成分の作る数列も収束するようにする。以下続けて第 n 成分も収束するような部分列を取る。終了。

3. 閉集合の点列による特徴づけ (後の命題 2.4.8)。これは書いてもノート半ページ。

4. $\mathbf{R}^n$ の有界閉集合 K の点列コンパクト性。これは有界だから、Bolzano-Weierstrass で収束部分列が取れて、閉集合だから、その極限が K に含まれる、ということ。逆「点列コンパクトならば有界閉集合」も言える。

5. $\mathbf{R}^n$ の有界閉集合の連続関数像は有界閉集合。ここまで準備してあれば 3 行で証明できる。

6. Weierstrass の最大値定理。$\mathbf{R}$ の有界閉集合は最大値を持つから。

授業では、結局証明は余所の授業に任せることにして、コンパクト集合の連続関数像はコンパクト、という定理の系という見方だけを紹介した。変かな？

次の定理が重要である (中間値の定理と、この定理さえ証明できれば、多変数の微分積分学 1 としては、ほぼ完璧という気もする)。

定理 2.4.5 (有界閉集合上の実数値連続関数は最大値と最小値を持つ, Weierstrass) K は $\mathbf{R}^n$ の空でない有界閉集合、$f\colon K \to \mathbf{R}$ は連続とするならば、f は K 上で最大値、最小値を持つ。

1 年次に (数学演習 2 で)、$\mathbf{R}^2$ の場合を学んだはずである。$\mathbf{R}^n$ でも本質的な違いはない。

ここでは、この定理 2.4.5 が以下の命題 (非常に一般的に成り立つ定理の $\mathbf{R}^n$ バージョン) の系とみなせることを注意しておく。

命題 2.4.6 (コンパクト集合の連続関数による像はコンパクト、の $\mathbf{R}^n$ バージョン) K を $\mathbf{R}^n$ の有界閉集合、$f\colon K \to \mathbf{R}^m$ は連続とするとき、$f(K)$ は $\mathbf{R}^m$ の有界閉集合である。

定理 2.4.5 の証明 $f\colon K \to \mathbf{R}$ が連続とすると、$f(K)$ は $\mathbf{R}$ の有界閉集合である。次の補題から $\max f(K)$, $\min f(K)$ が存在する。■

補題 2.4.7 $\mathbf{R}$ の空でない上に有界な閉集合 K は最大値を持つ。

証明 Weierstrass の定理より、K の上限 U が存在する。$\exists \{x_n\}_{n\in\mathbf{N}}$ s.t. $\forall n \in \mathbf{N}$ $x_n \in K$, $\lim_{n\to\infty} x_n = U$. 次の命題を認めれば、$U \in K$. ゆえに $U = \max K$. ■

$\mathbf{R}^n$ における閉集合は、次のように点列の言葉を用いて特徴づけられる (解析学にとっては大変便利である):

命題 2.4.8 (閉集合の点列による特徴づけ) $A \subset \mathbf{R}^n$ とするとき、

$$A \text{ が閉集合} \Leftrightarrow A \text{ 内の点列がもし } \mathbf{R}^n \text{ 内で収束するならば、極限は } A \text{ に属する。}$$

証明 いずれも背理法で証明する。

($\Rightarrow$) $\{x_n\}_{n\in\mathbf{N}}$ は A 内の点列で、$\mathbf{R}^n$ 内で a に収束しているとする。$a \in A$ を証明するために、$a \notin A$ と仮定する。これは $a \in A^c$ ということだから、A^c が開集合であることから

$$\exists \varepsilon > 0 \text{ s.t. } B(a;\varepsilon) \subset A^c.$$

ゆえに $B(a;\varepsilon) \cap A = \emptyset$. ところが、$x_n \to a$ $(n \to \infty)$ であるから、十分大きな $n \in \mathbf{N}$ に対して $x_n \in B(a;\varepsilon)$. $x_n \in A$ であるから、これは矛盾である。ゆえに $a \in A$ でなければならない。

($\Leftarrow$) A が閉集合でないと仮定する。すると A^c は開集合でないので、

$$\exists a \in A^c \text{ s.t. } \forall \varepsilon > 0\ B(a;\varepsilon) \not\subset A^c.$$

これは $B(a;\varepsilon) \cap A \neq \emptyset$ ということである。これから各 $n \in \mathbf{N}$ に対して $x_n \in B(a;1/n) \cap A$ なる x_n を取ることが出来る。$x_n \to a$ $(n \to \infty)$ で、$x_n \in A$ であるから、仮定から $a \in A$. これは矛盾である。ゆえに A は閉集合である。■

2.5　微分についての短いガイダンス

多変数になると変数の増分 h がベクトルになるので、1 変数の場合の微分係数の定義

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$$

は (ベクトル h で割るというナンセンスな式になってしまうので) そのままの形では使えない。

多変数関数については、2 つの微分がある。

(1) 全微分 $f'(a)$

(2) 偏微分 $\dfrac{\partial f}{\partial x_j}(a)$

1 変数関数の微分 $f'(a)$ と良く対応するのは、全微分の方である (だから同じ記号を使うことにしたし、最近は「全微分」と言わずに単に「微分」と呼ぶ人も増えている[3])。例えば、1 変数実数値関数 f のグラフ $y=f(x)$ 上の点 $(a, f(a))$ における接線の方程式は

$$y = f'(a)(x-a)+f(a)$$

であるが、多変数実数値関数 f のグラフ $y=f(\vec{x})$ 上の点 $(\vec{a}, f(\vec{a}))$ における接平面の方程式は

$$y = f'(\vec{a})(\vec{x}-\vec{a})+f(\vec{a})$$

である。形式上はまったく違いがなく、覚える苦労がない。

なお、f がベクトル値 $f=\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ である場合、$\dfrac{\partial f}{\partial x_j}(a) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_j}(a) \\ \dfrac{\partial f_2}{\partial x_j}(a) \\ \vdots \\ \dfrac{\partial f_m}{\partial x_j}(a) \end{pmatrix}$ となるのは、これまで (例えば 1.4.3) と同様である。

全微分と偏微分の関係はある意味簡単で、

$$\begin{aligned} f'(a) &= \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) & \dfrac{\partial f}{\partial x_2}(a) & \cdots & \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix} \\ &= \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \cdots & \dfrac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \dfrac{\partial f_m}{\partial x_2}(a) & \cdots & \dfrac{\partial f_m}{\partial x_n}(a) \end{pmatrix} = \left(\dfrac{\partial f_i}{\partial x_j}(a) \right). \end{aligned}$$

つまり

全微分係数は、偏微分係数を成分とする行列である。

[3]数学の本に書かれている内容はすぐには変化せず、微分積分については、30〜40 年前に書かれた教科書が現在も十分現役として使うことが出来る。しかし、全微分を単に「微分」と呼んだり、それを $f'(a)$ という記号で表す習慣は、比較的新しいと思われる。古い本には見られない。

2.6 偏微分

2.6.1 定義

数学のテキスト、講義では定義から始めるのが普通だが、まずは実例を見せよう。

例 2.6.1 (2 変数 2 次関数) 実定数 a, b, c, d, p, q, r に対して、

$$f(x,y) := ax^2 + bxy + cy^2 + px + qy + r \quad ((x,y) \in \mathbf{R}^2)$$

とおいて $f\colon \mathbf{R}^2 \to \mathbf{R}$ を定めるとき、

$$\frac{\partial f}{\partial x} = 2ax + by + p.$$

x で偏微分するときは、他の変数 (ここでは y) を定数と見なして微分する。以下、同様に

$$\frac{\partial f}{\partial y} = bx + 2cy + q.$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = 2a,$$
$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = b,$$
$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = b,$$
$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = 2c.$$

ちなみに f の (全) 微分 $f'(x,y)$ は

$$f'(x,y) = \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 2ax + by + p & bc + 2cy + q \end{pmatrix}.$$

これは $\begin{pmatrix} 2ax + by + p \\ bc + 2cy + q \end{pmatrix}$ ではない。これは $\nabla f(x,y)$ という記号で表される。■

記号 ∂ は、多変数関数の 1 つの変数に関する微分 (偏微分) であることを強調するためのもので、partial ‘d’, round ‘d’, または単に ‘d’ と読まれる (Jacobi に始まるものだそうである)。

定義 2.6.2 (1 点における偏微分係数) Ω は $\mathbf{R}^n$ の開集合, $f\colon \Omega \to \mathbf{R}^m$, $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \Omega$, $j \in \{1, \dots, n\}$ とする。f が点 a で変数 x_j について**偏微分可能**であるとは、極限

$$\lim_{h \to 0} \frac{f(a + he_j) - f(a)}{h}$$

が存在することをいう。ここで e_j は、第 j 成分が 1 で、それ以外の成分がすべて 0 であるような、$\mathbf{R}^n$ のベクトルである。このとき、この極限値 ($\in \mathbf{R}^m$) を、f の点 a での変数 x_j についての**偏微分係数** (the partial derivative of f in the direction x_j at a) と呼び、

$$\frac{\partial f}{\partial x_j}(a), \quad \frac{\partial}{\partial x_j} f(a), \quad f_{x_j}(a)$$

などの記号で表す。

ベクトル記法を使わずに、成分を用いて表すと

$$\frac{f(a+he_j)-f(a)}{h}=\frac{f(a_1,\ldots,a_{j-1},a_j+h,a_{j+1},\ldots,a_n)-f(a_1,\ldots,a_{j-1},a_j,a_{j+1},\ldots,a_n)}{h}$$

である。

偏導関数、高階偏微分、C^k 級 $(0\le k\le\infty)$ について述べる。

定義 2.6.3 (偏導関数、高階微分、C^k 級) Ω は $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}^m$ とする。

(1) $j\in\{1,\ldots,n\}$ とする。f が Ω で x_j について**偏微分可能**であるとは、$\forall x\in\Omega$ に対して、f は x で変数 x_j について偏微分可能であることをいう。このとき、写像

$$\Omega\ni x\mapsto\frac{\partial f}{\partial x_j}(x)\in\mathbf{R}^m$$

を f の変数 x_j に関する**偏導関数**と呼び、

$$\frac{\partial f}{\partial x_j},\quad\frac{\partial}{\partial x_j}f,\quad f_{x_j}$$

などの記号で表す。

(2) $\dfrac{\partial f}{\partial x_1},\ldots,\dfrac{\partial f}{\partial x_n}$ を f の**1 階偏導関数**と呼ぶ。

(3) $i,j\in\{1,\ldots,n\}$ とする。f が Ω で変数 x_j について偏微分可能で、偏導関数 $\dfrac{\partial f}{\partial x_j}$ が Ω で変数 x_i について偏微分可能であるとき、$\dfrac{\partial}{\partial x_i}\left(\dfrac{\partial f}{\partial x_j}\right)$ を

$$\frac{\partial^2 f}{\partial x_i\partial x_j},\quad f_{x_jx_i}$$

などの記号で表す。$i=j$ である場合、つまり $\dfrac{\partial^2 f}{\partial x_j\partial x_j}$ を $\dfrac{\partial^2 f}{\partial x_j^2}$ とも書く。

(4) $\dfrac{\partial^2 f}{\partial x_i\partial x_j}$ $(i,j=1,\ldots,n)$ を f の **2 階偏導関数**と呼ぶ。

(5) 同様に任意の k $(k\in\mathbf{N},\ k\ge3)$ に対して、f の $\boldsymbol{k}$ **階偏導関数**が定義される。

(6) $k\in\mathbf{N}$ とする。f が Ω で C^k **級**であるとは、f が Ω で k 階のすべての偏導関数を持ち、それらすべてと f 自身が Ω で連続であることをいう。

(7) f が Ω で C^∞ **級**であるとは、$\forall k\in\mathbf{N}$ に対して、f が Ω で C^k 級であることをいう。

(8) f が Ω で C^0 **級**であるとは、f が Ω で連続であることをいう。f 自身を f の 0 階偏導関数ともいう。

注意 2.6.4 (C^k 級の定義) C^k 級の定義は、上とは別の定義をしている本もある。よくあるのは、f が C^k 級であるためには、

($\heartsuit$) f が k 階までのすべての偏導関数を持ち、それら偏導関数 (f 自身も含む) が連続

とするものである。後述の「C^1 級 $\Longrightarrow$ 全微分可能」の証明を精査すれば分かるように、k 階のすべての偏導関数の存在とそれらの連続性から、$k-1$ 以下の階数の偏導関数 (f 自身を含む) の連続性が導かれる。■

注意 2.6.5 (定義域が開集合である理由) 定義域を開集合としてあるので、$\forall a \in \Omega$ に対して、$\exists \varepsilon > 0$ s.t. $B(a;\varepsilon) \subset \Omega$ となるので、$|h| < \varepsilon$ なる任意の h に対して $a + he_j \in \Omega$ が保証される。この保証がないと議論が著しく面倒になる。

開集合でない Ω で定義された関数が微分可能とは、Ω を含む開集合 $\widetilde{\Omega}$ が存在して、f が $\widetilde{\Omega}$ 上で微分可能な関数 $\widetilde{f}$ に拡張可能なことと定義する場合が多い。■

問 2.6.1 $f(x,y,z) := \dfrac{1}{\sqrt{x^2+y^2+z^2}}$ とするとき、次式を示せ (まず f_x, f_{xx} を計算せよ)。

$$f_{xx} + f_{yy} + f_{zz} = 0.$$

2.6.2 偏微分の順序交換

高階の微分係数の値は、多くの場合、偏微分の順序にはよらない、という重要な定理 (複数) がある。

次に掲げる定理が基本的である。

定理 2.6.6 Ω は $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}^m$ は C^2 級とするとき、任意の $i, j \in \{1, \ldots, n\}$, $a \in \Omega$ に対して、

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

$i = j$ のときは明らかに成立するので、$i \neq j$ の場合の証明が問題となる。

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_i \partial x_j}(a) &= \lim_{h\to 0} \frac{1}{h}\left(\frac{\partial f}{\partial x_j}(a+he_i) - \frac{\partial f}{\partial x_j}(a)\right) \\
&= \lim_{h\to 0} \frac{1}{h}\left(\lim_{k\to 0} \frac{f(a+he_i+ke_j) - f(a+he_i)}{k} - \lim_{k\to 0} \frac{f(a+ke_j) - f(a)}{k}\right) \\
&= \lim_{h\to 0}\lim_{k\to 0} \frac{1}{hk}\left(f(a+he_i+ke_j) - f(a+he_i) - f(a+ke_j) + f(a)\right)
\end{aligned}$$

であり、

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(a) = \lim_{k\to 0}\lim_{h\to 0} \frac{1}{hk}\left(f(a+he_i+ke_j) - f(a+he_i) - f(a+ke_j) + f(a)\right)$$

である。極限の順序の問題であることが分かる。実は

$$\lim_{(h,k)\to(0,0)} \frac{1}{hk}\left(f(a+he_i+ke_j) - f(a+he_i) - f(a+ke_j) + f(a)\right)$$

が存在するので、両者は一致する、というのがあらすじである。

証明 $i = j$ の場合に成り立つことは明らかなので、$i \neq j$ の場合を証明する。x_i と x_j 以外の変数 x_k ($k \neq i, j$) は、$x_k = a_k$ と固定しているので、本質的に 2 変数関数の話である。$x_i = x$, $x_j = y$, $a_i = a$, $a_j = b$, $f(a_1, \cdots, a_{i-1}, x, a_{i+1}, \ldots, a_{j-1}, y, a_{j+1}, \ldots, a_n) = f(x,y)$ として、

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$$

を示せば良い。

$$\Delta(h,k) := f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b)$$

とおく[4]。$\phi(x) := f(x,b+k) - f(x,b)$ とおくと、十分小さな正数 ε を取ると、$0 < |h| < \varepsilon$, $0 < |k| < \varepsilon$ を満たす任意の h,k に対して、$\exists\theta,\theta' \in (0,1)$ s.t.

$$\begin{aligned}\Delta(h,k) &= \phi(a+h) - \phi(a)\\ &= \phi'(a+\theta h)h = [f_x(a+\theta h,b+k) - f_x(a+\theta h,b)]\,h\\ &= f_{xy}(a+\theta h,b+\theta' k)hk.\end{aligned}$$

ゆえに

$$\lim_{(h,k)\to(0,0)} \frac{\Delta(h,k)}{hk} = \lim_{(h,k)\to(0,0)} f_{xy}(a+\theta h,b+\theta' k) = f_{xy}(a,b).$$

同様に $\psi(y) := f(a+h,y) - f(a,y)$ とおくと、$0 < |h| < \varepsilon$, $0 < |k| < \varepsilon$ を満たす任意の h,k に対して、$\exists\theta'',\theta''' \in (0,1)$ s.t.

$$\Delta(h,k) = \psi(b+k) - \psi(b) = f_{yx}(a+\theta''' h,b+\theta'' k)hk.$$

これから

$$\lim_{(h,k)\to(0,0)} \frac{\Delta(h,k)}{hk} = \lim_{(h,k)\to(0,0)} f_{yx}(a+\theta''' h,b+\theta'' k) = f_{yx}(a,b).$$

ゆえに

$$f_{xy}(a,b) = f_{yx}(a,b).\blacksquare$$

C^2 級でない場合には、偏微分の順序交換は成立しないこともある。

反例をあげる前に、一つの例 (1 年生で学んだはず) を復習しておこう。

例 2.6.7 (1 回微分可能だが、C^1 級でない関数) 次の関数は 1 回微分可能であるが、C^1 級ではない。

$$f(x) := \begin{cases} x^2 \sin\dfrac{1}{x} & (x \neq 0)\\ 0 & (x = 0).\end{cases}$$

f は $\mathbf{R}\setminus\{0\}$ で C^∞ 級である。0 で微分可能であることを示す。$h \neq 0$ とするとき

$$\frac{f(0+h)-f(0)}{h} = \frac{h^2\sin\frac{1}{h} - 0}{h} = h\sin\frac{1}{h} \to 0 \quad (h\to 0)$$

実際

$$\left|h\sin\frac{1}{h} - 0\right| = |h|\left|\sin\frac{1}{h}\right| \le |h|\cdot 1 = |h| \to 0 \quad (h \to 0)$$

であるから。ゆえに $f'(0) = 0$. ところで $x \neq 0$ に対して、

$$f'(x) = 2x\cdot\sin\frac{1}{x} + x^2\cdot\left(-\frac{1}{x^2}\right)\cos\frac{1}{x} = 2x\sin\frac{1}{x} - \cos\frac{1}{x}$$

[4] これは二つの差分を施したものであり、それらは順序交換可能である:

$$\Delta(h,k) = \Big[[f(x,y)]_{x=a}^{x=a+h}\Big]_{y=b}^{y=b+k} = \Big[[f(x,y)]_{y=b}^{y=b+k}\Big]_{x=a}^{x=a+h}.$$

であり、これは $x \to 0$ のとき、$0 = f'(0)$ に収束しない (出来ますか?)。ゆえに f' は 0 では連続でない。ゆえに f は C^1 級でない。

たまに「$x=0$ のとき $f(x)=0$ だから $f'(0)=0$」というトンでもないことをする人がいる。もしそれが正しければ、$g(x)=x$ に対して、$x=0$ のとき $g(x)=0$ だから $g'(0)=0$ となるはずだが、$g'(0)=1$ であるから矛盾する。$f(h)$ は $f(x)=x^2\sin\dfrac{1}{x}$ $(x\neq 0)$ で計算することを納得すること。■

しかし、つねに (2.1) が成立するわけではない。反例をあげておく。

例 2.6.8 (偏微分の順序交換が成立しない関数, Peano の例)

$$f(x,y) := \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & ((x,y)\neq(0,0)) \\ 0 & ((x,y)=(0,0)) \end{cases}$$

について $f_{xy}(0,0)=-1$, $f_{yx}(0,0)=1$ を示せ。

(解) f_{xy} は、f_x を y について偏微分したものであるから、

$$f_{xy}(0,0)=\lim_{h\to 0}\frac{f_x(0,h)-f_x(0,0)}{h}.$$

$$f_x(0,0)=\lim_{k\to 0}\frac{f(0+k,0)-f(0,0)}{k}=\lim_{k\to 0}\frac{k\cdot 0\cdot\frac{k^2-0^2}{k^2+0^2}-0}{k}=\lim_{k\to 0}0=0,$$

$h\neq 0$ のとき、

$$\begin{aligned} f_x(0,h) &= \lim_{k\to 0}\frac{f(0+k,h)-f(0,h)}{k}=\lim_{k\to 0}\frac{kh\frac{k^2-h^2}{k^2+h^2}-0\cdot h\cdot\frac{0^2-h^2}{0^2+h^2}}{k} \\ &= \lim_{k\to 0}\frac{h(k^2-h^2)}{k^2+h^2}=\frac{h(0-h^2)}{0^2+h^2}=-h. \end{aligned}$$

ゆえに

$$f_{xy}(0,0)=\lim_{h\to 0}\frac{-h-0}{h}=\lim_{h\to 0}(-1)=-1.$$

同様に

$$f_{yx}(0,0)=1$$

が得られる。■

問 2.6.2 $f\colon \mathbf{R}^2\to\mathbf{R}$ を次式で定めるとき、$f_{xy}(0,0)$ と $f_{yx}(0,0)$ を求めよ。

$$f(x,y)=\begin{cases} \dfrac{x^3y}{x^2+y^2} & ((x,y)\neq(0,0)) \\ 0 & ((x,y)=(0,0)).\blacksquare \end{cases}$$

例えば f が C^3 級ならば

$$\frac{\partial^3 f}{\partial x_2\partial x_1\partial x_1}=\frac{\partial^3 f}{\partial x_1\partial x_2\partial x_1}=\frac{\partial^3 f}{\partial x_1\partial x_1\partial x_2}=\frac{\partial^3 f}{\partial x_1^2\partial x_2}.$$

(各等号の理由をちゃんと言えるか？考えてみよう。)

結局 f が C^k 級ならば、k 階の偏導関数は、各変数 x_i $(i=1,\dots,n)$ について何回偏微分したかで定まり、

$$\frac{\partial^k f}{\partial x_1^{i_1}\partial x_2^{i_2}\cdots\partial x_n^{i_n}},\quad i_1+i_2+\cdots+i_n=k$$

という形に表すことが出来る。

余談 2.6.1 (定理の名前) ところで、上の定理 2.6.6 を「Schwarz の定理」と呼ぶ本が少なくないのだが、それが正しいのか、確信が持てない。高木 [4] には、上の定理以外に以下の定理が紹介されている。

命題 2.6.9 (Schwarz) ある領域で f_x, f_y, f_{xy} が存在して、領域内の 1 点において f_{xy} が連続ならば、その点において f_{yx} も存在し、かつ $f_{xy} = f_{yx}$.

命題 2.6.10 (Young) ある領域で f_x, f_y が存在して、それらが領域内の 1 点において全微分可能ならば、その点において $f_{xy} = f_{yx}$.

命題 2.6.9 と混同しているのかな…という気がしないでもない。とりあえず解析概論に従っておく。■

余談 2.6.2 (C^2 級でなくても 2 回全微分可能ならば OK) 次の定理は、上記の Young の定理に含まれるようではあるが、仮定が自然なので、重要であると筆者は考える。

シュヴァルツ [5], p.

f が 2 回全微分可能であれば、

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a). \tag{2.1}$$

ところが「2 回全微分可能」という概念の定義を省略するテキストが多い。f が Ω で 2 回全微分可能であるとは、f が Ω で全微分可能で、すべての偏導関数 $\frac{\partial f}{\partial x_j}$ が Ω で全微分可能なことをいう。この定義は極めて自然なのだが、それを説明するのに案外と手間がかかるので (扱う関数の範囲を少し広げると (例えばシュヴァルツ [5], かいつまんで言うと、関数の終域を $\mathbf{R}^m$ でなく、$\mathbf{R}$ 上の有限次元ベクトル空間、と少し一般化する)、導関数 $x \mapsto f'(x)$ の導関数が定義できるので、2 回微分可能の概念が自然に確定する。それは全微分可能であり、かつすべての 1 階偏導関数が全微分可能であることと同値であるのはすぐ分かる。)、ここでは省略する。

応用上の観点からは、2 回全微分可能であるが、C^2 級ではないような関数が登場することは稀なので、ここに述べたことを気にする必要はあまりない。それが現在の微分積分学のテキストの多くで、高階の全微分の概念を省略して、C^k 級という概念だけで済まされる理由であろう。■

一方で、超関数[5]の世界では、つねに (2.1) が成立する。結局、応用上現れるほとんどの場合に、偏微分の順序交換が成立すると思ってよい。

問 2.6.3 2 変数関数で、1 回全微分可能であるが、C^1 級ではない関数の例をあげよ。(p. ?? を見よ。)

問 2.6.4 2 回微分可能であるが、C^2 級でない 1 変数関数の例をあげよ。(ヒント: $x^p \sin \dfrac{1}{x}$)

2.7 全微分

やや天下りに定義をするが、少しでも天下り感をなくそう、ということで、それがもっともに感じられる理由を説明していく。

[5]超関数とは、微積分が自由に出来るように関数概念を拡張したもの。

2.7.1 定義

まず復習から。1 変数関数 $f\colon I \to \mathbf{R}^m$ (I は $\mathbf{R}$ の区間), $a \in I$ の場合、

$$
\begin{aligned}
f \text{ が } a \text{ で微分可能} &\overset{\text{def.}}{\Leftrightarrow} \lim_{h\to 0}\frac{f(a+h)-f(a)}{h} \quad \text{が存在} \\
&\Leftrightarrow \exists A \in \mathbf{R}^m \quad \lim_{h\to 0}\frac{f(a+h)-f(a)}{h} = A \\
&\Leftrightarrow \exists A \in \mathbf{R}^m \quad \lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{|h|} = 0.
\end{aligned}
$$

実数を成分とする m 行 n 列の行列全体の集合を $M(m,n;\mathbf{R})$ で表す。

定義 2.7.1 (全微分可能性, 全微分係数, 導関数) Ω は $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}^m$ とする。

(1) $a \in \Omega$ とする。f が a で **(全) 微分可能** ((totally) differentiable at a) であるとは、

$$\exists A \in M(m,n;\mathbf{R}) \quad \text{s.t.} \quad \lim_{h\to 0}\frac{1}{\|h\|}\left(f(a+h)-f(a)-Ah\right) = 0$$

が成り立つことをいう。(この A は実は f と a から一意的に定まる — 証明は後述。) このとき A を f の a における **(全) 微分係数** (the (total) derivative of f at a, the (total) differential of f at a) と呼び、$f'(a)$ で表す。

(2) f が Ω で **(全) 微分可能**であるとは、$\forall x \in \Omega$ で f が全微分可能であることをいう。このとき、

$$\Omega \ni x \mapsto f'(x) \in M(m,n;\mathbf{R})$$

を f の**導関数**と呼び、f' で表す。

問 f の a における全微分係数は、もし存在するならば一意的であることを示せ (すぐ後の定理の証明を見れば分かることだが、直接証明も可能である)。

2.7.2 全微分可能性, 連続性, 偏微分可能性、C^1 級

定理 2.7.2 (全微分可能ならば連続かつすべての変数につき偏微分可能で全微分係数はヤコビ行列) Ω は $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}^m$, $a \in \Omega$, f は a で全微分可能 (i.e., $\exists A \in M(m,n;\mathbf{R})$ s.t. $\displaystyle\lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{\|h\|} = 0$) とするとき、次の (1), (2) が成り立つ。

(1) f は a で連続である。

(2) f は a で、すべての変数 x_j について偏微分可能で、

$$A = \left(\frac{\partial f_i}{\partial x_j}(a)\right).$$

(これで全微分係数の一意性の一つの証明が得られる。)

証明のために、行列のノルムを導入し、その性質を一つ紹介する[6]。$A=(a_{ij})\in M(m,n;\mathbf{R})$ とするとき、

$$\|A\| := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$$

とおき、これを A の**ノルム**と呼ぶ。

$M(m,n;\mathbf{R})$ の元 A を、自然に (強引に？) $\mathbf{R}^{mn}$ の元に対応させたとき、その ($\mathbf{R}^{mn}$ における) ノルムと $\|A\|$ は一致することから、

$$\|A\| \geq 0, \quad \text{等号成立} \Leftrightarrow A=0,$$

$$\|A+B\| \leq \|A\| + \|B\|,$$
$$\|\lambda A\| = |\lambda|\,\|A\|$$

などが成立することは明らかである。

命題 2.7.3 $\forall A\in M(m,n;\mathbf{R})$, $\forall x\in\mathbf{R}^n$ に対して、

$$\|Ax\| \leq \|A\|\ \|x\|$$

が成り立つ。

証明　A の第 i 行ベクトルを転置したものを $\boldsymbol{a}_i\in\mathbf{R}^n$ とする。

$$\boldsymbol{a}_i = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}$$

であり、

$$A = \begin{pmatrix} \boldsymbol{a}_1^T \\ \vdots \\ \boldsymbol{a}_m^T \end{pmatrix}, \quad \|A\| = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} = \sqrt{\sum_{i=1}^{m}\|\boldsymbol{a}_i\|^2},$$

$$Ax = \begin{pmatrix} \boldsymbol{a}_1^T x \\ \vdots \\ \boldsymbol{a}_m^T x \end{pmatrix} = \begin{pmatrix} (\boldsymbol{a}_1, x) \\ \vdots \\ (\boldsymbol{a}_m, x) \end{pmatrix}.$$

Schwarz の不等式から、

$$\|Ax\|^2 = \sum_{i=1}^{m}(\boldsymbol{a}_i, x)^2 \leq \sum_{i=1}^{m}\|\boldsymbol{a}_i\|^2\,\|x\|^2 = \left(\sum_{i=1}^{m}\|\boldsymbol{a}_i\|^2\right)\|x\|^2 = \|A\|^2\,\|x\|^2.$$

ゆえに

$$\|Ax\| \leq \|A\|\ \|x\|.\blacksquare$$

[6]実は、$\lim_{h\to 0} Ah = 0$ を証明しているだけなので、「成分を考えて 1 次多項式は関数として連続」と言えば済む。しかし、線形写像 $\mathbf{R}^n \ni x \mapsto Ax\in\mathbf{R}^m$ の連続性を見通し良く示すことは有意義であると思うので、少々大げさではあるが、行列のノルムを導入してみた。

定理 2.7.2 の証明　f が a で全微分可能であるから、$\exists A \in M(m,n;\mathbf{R})$ s.t.

$$(\sharp)\qquad \lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{\|h\|}=0.$$

(1) $h\to 0$ のとき、$\|Ah\|\le\|A\|\ \|h\|\to 0$ であるから、$\|Ah\|\to 0$ であることに注意しておく。

$$f(a+h)-f(a)=f(a+h)-f(a)-Ah+Ah=\|h\|\frac{f(a+h)-f(a)-Ah}{\|h\|}+Ah$$

であるから、$h\to 0$ のとき、

$$\begin{aligned}\|f(a+h)-f(a)\|&=\left\|\|h\|\frac{f(a+h)-f(a)-Ah}{\|h\|}+Ah\right\|\\&\le\|h\|\frac{\|f(a+h)-f(a)-Ah\|}{\|h\|}+\|Ah\|\to 0\cdot 0+0=0.\end{aligned}$$

ゆえに

$$\lim_{h\to 0}f(a+h)=f(a).$$

すなわち f は a で連続である。

(2) 突然だが、ベクトルを表す文字の上に $\vec{}$ を書くことにする。$\vec{f}=\begin{pmatrix}f_1\\ \vdots\\ f_m\end{pmatrix}$, $A=(a_{ij})$, $\vec{h}=\begin{pmatrix}h_1\\ \vdots\\ h_n\end{pmatrix}$ とおくと

$$(\text{再掲 }\sharp)\qquad \lim_{\vec{h}\to\vec{0}}\frac{\vec{f}(\vec{a}+\vec{h})-\vec{f}(\vec{a})-A\vec{h}}{\left\|\vec{h}\right\|}=\vec{0}$$

において、第 i 成分は、

$$\lim_{\vec{h}\to\vec{0}}\frac{f_i\left(\vec{a}+\vec{h}\right)-f_i(\vec{a})-\displaystyle\sum_{k=1}^n a_{ik}h_k}{\left\|\vec{h}\right\|}=0.$$

$j\in\{1,\dots,n\}$ に対して、$\vec{h}=h\vec{e}_j$ とすると、$\displaystyle\sum_{k=1}^n a_{ik}h_k=\sum_{k=1}^n a_{ik}h\delta_{kj}=a_{ij}h$ であるから

$$\lim_{h\to 0}\frac{f_i(\vec{a}+h\vec{e}_j)-f_i(\vec{a})-a_{ij}h}{|h|}=0.$$

これから

$$\lim_{h\to 0}\frac{f_i(\vec{a}+h\vec{e}_j)-f_i(\vec{a})}{h}=a_{ij}.$$

ゆえに f_i は $\vec{a}$ で変数 x_j につき偏微分可能で、

$$\frac{\partial f_i}{\partial x_j}(\vec{a})=a_{ij}.$$

ゆえに

$$A=(a_{ij})=\left(\frac{\partial f_i}{\partial x_j}(\vec{a})\right).\blacksquare$$

この定理によって、f が全微分可能であるとき、その全微分係数は偏微分することで求められることが分かった。偏微分は本質的に 1 変数関数の世界の話であるから、簡単に実行できることが多い。

余談 2.7.1 (有限次元空間の間の線形写像の連続性) 線形写像 $f\colon \mathbf{R}^n \ni x \mapsto Ax \in \mathbf{R}^m$ の連続性を示すのに、上では不等式

$$\|Ax\| \le \|A\| \, \|x\|$$

を用いた。これは行列のノルムを重要であると考えたため、紹介する良い機会と考えたためであるが、単に f の連続性を示すだけであれば、これまでと同様に次のようにすれば良い。

$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix}$ とおくと、任意の i に対して

$$f_i(x) = \sum_{j=1}^{n} a_{ij} x_j \quad (x = (x_1, \cdots, x_n)^T \in \mathbf{R}^n).$$

これは $x_1, \cdots, x_n$ の実係数 (1 次) 多項式であるから、$f_i\colon \mathbf{R}^n \to \mathbf{R}$ は連続である。ゆえに $f\colon \mathbf{R}^n \to \mathbf{R}^m$ は連続である。

線形写像はつねに連続というわけではない。無限次元線形空間には、線形演算を連続とするような位相が複数存在し、線形写像はつねに連続というわけではなくなる。そういうものを扱うのに、成分に分解して考えるやり方はほとんど役に立たない。一方、線形写像のノルムという考え方は無限次元空間においても、しばしば有効である。

乱暴にまとめると

> 無限次元空間における線形写像については、成分に分解して考えるという方法は無力で、ノルムを用いる方法が重要になってくる

となる。■

余談 2.7.2 (全微分係数の一意性の別証明) $A, B \in M(m, n; \mathbf{R})$ が

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - Ah}{\|h\|} = 0, \quad \lim_{h \to 0} \frac{f(a+h) - f(a) - Bh}{\|h\|} = 0$$

を満たすならば、実は $A = B$ であることを証明する。

まず

$$\begin{aligned}
\lim_{h \to 0} \frac{(A-B)h}{\|h\|} &= \lim_{h \to 0} \frac{(f(a+h) - f(a) - Bh) - (f(a+h) - f(a) - Ah)}{\|h\|} \\
&= \lim_{h \to 0} \frac{f(a+h) - f(a) - Bh}{\|h\|} - \lim_{h \to 0} \frac{f(a+h) - f(a) - Ah}{\|h\|} \\
&= 0 - 0 = 0.
\end{aligned}$$

もし $A = B$ でなければ、$B - A \ne 0$. ゆえに $\exists \widetilde{h} \in \mathbf{R}^n$ s.t. $(B-A)\widetilde{h} \ne 0$. このとき、$h = \varepsilon \widetilde{h}$ を考えると、$\varepsilon \to 0$ のとき $h \to 0$ であるから、

$$\lim_{\varepsilon \to 0} \frac{(A-B)(\varepsilon \widetilde{h})}{\left\| \varepsilon \widetilde{h} \right\|} = 0.$$

ゆえに

$$\lim_{\varepsilon\to 0}\frac{\left\|(A-B)(\varepsilon\widetilde{h})\right\|}{\left\|\varepsilon\widetilde{h}\right\|}=0.$$

ところが

$$\frac{\left\|(A-B)(\varepsilon\widetilde{h})\right\|}{\left\|\varepsilon\widetilde{h}\right\|}=\frac{|\varepsilon|\left\|(A-B)\widetilde{h}\right\|}{|\varepsilon|\left\|\widetilde{h}\right\|}=\frac{\left\|(A-B)\widetilde{h}\right\|}{\left\|\widetilde{h}\right\|}\neq 0 \quad (\varepsilon\text{ によらない }0\text{ でない定数})$$

であるから、矛盾である。■

定理 2.7.4 (C^1 級ならば全微分可能) Ω は $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}^m$ はすべての変数 x_j につき Ω で偏微分可能で、偏導関数 $\dfrac{\partial f}{\partial x_j}\colon\Omega\to\mathbf{R}^m$ は連続とする。このとき、f は Ω で全微分可能である。

証明　Ω が開集合であるから、$\exists\varepsilon>0$ s.t. $B(a;\varepsilon)\subset\Omega$. 任意の $i\in\{1,\cdots,n\}$, $h\in\mathbf{R}^n$, $0<\|h\|<\varepsilon$ に対して、

$$\begin{aligned}
f_i(a+h)-f_i(a)&=f_i(a_1+h_1,a_2+h_2,\dots,a_n+h_n)-f_i(a_1,a_2,\dots,a_n)\\
&=f_i(a_1+h_1,a_2+h_2,\dots,a_n+h_n)-f_i(a_1,a_2+h_2,\dots,a_n+h_n)\\
&\quad+f_i(a_1,a_2+h_2,a_3+h_3\dots,a_n+h_n)-f_i(a_1,a_2,a_3+h_3,\dots,a_n+h_n)\\
&\quad+f_i(a_1,a_2,a_3+h_3,\dots,a_n+h_n)-f_i(a_1,a_2,a_3,a_4+h_4,\dots,a_n+h_n)\\
&\qquad\qquad\vdots\qquad\qquad\qquad\qquad\vdots\\
&\quad+f_i(a_1,a_2,\dots,a_{n-1},a_n+h_n)-f_i(a_1,a_2,\dots,a_{n-1},a_n).
\end{aligned}$$

平均値の定理より、$\exists\theta_1,\theta_2,\dots,\theta_n\in(0,1)$ s.t.

$$f_i(a+h)-f_i(a)=\sum_{j=1}^n\frac{\partial f_i}{\partial x_j}(a_1,\dots,a_{j-1},a_j+\theta_jh_j,a_{j+1}+h_{j+1},\dots,a_n+h_n)h_j.$$

ゆえに

$$f_i(a+h)-f_i(a)-\sum_{j=1}^n\frac{\partial f_i}{\partial x_j}(a)h_j=\sum_{j=1}^n\varepsilon_{ij}(h)h_j.$$

ただし、

$$\varepsilon_{ij}(h):=\frac{\partial f_i}{\partial x_j}(a_1,\dots,a_{j-1},a_j+\theta_jh_j,a_{j+1}+h_{j+1},\dots,a_n+h_n)-\frac{\partial f_i}{\partial x_j}(a)\quad(j=1,\dots,n).$$

$h\to 0$ のとき

$$\|(a_1,\dots,a_{j-1},a_j+\theta_jh_j,a_{j+1}+h_{j+1},\dots,a_n+h_n)-a\|^2\le\sum_{j=1}^n h_j^2=\|h\|^2\to 0$$

であることと、f は C^1 級であると仮定したので $\dfrac{\partial f_i}{\partial x_j}$ は連続であることから、

$$\varepsilon_{ij}(h)\to 0.$$

ゆえに、三角不等式と、$|h_j| \le \|h\|$ $(j=1,\dots,n)$ であることを用いると、

$$\frac{\left|\sum_{j=1}^n \varepsilon_{ij}(h)h_j\right|}{\|h\|} \le \frac{\sum_{j=1}^n |\varepsilon_{ij}(h)|\,|h_j|}{\|h\|} \le \sum_{j=1}^n |\varepsilon_{ij}(h)| \to 0 \quad (h \to 0).$$

ゆえに

$$\frac{f_i(a+h) - f_i(a) - \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(a)h_j}{\|h\|} = \frac{\sum_{j=1}^n \varepsilon_{ij}(h)h_j}{\|h\|} \to 0 \quad (h \to 0).$$

これは f_i が a で全微分可能であることを示している。ゆえに $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$ も a で全微分可能である。■

これから、関数の 1 階偏導関数をすべて求めて (これは容易な場合が多い — ほとんど高校数学)、それらが連続関数であることが確められれば (やり方は演習しているはず)、その関数が全微分可能であることが分かる。これは与えられた関数が全微分可能であることの強力な確認手段である。

この定理の証明では、f がすべての 1 階偏導関数を持ち、それらが連続であることしか用いていない。つまり f の連続性は仮定していない。それで全微分可能であることが示されたので、「全微分可能ならば連続」という定理によって、f は連続である。ゆえに次のことが分かった。

系 2.7.5 f が C^1 級であるためには、f がすべての変数 x_j に関して偏微分可能で、$\dfrac{\partial f}{\partial x_j}$ と f が連続であることが必要十分である。

本によっては、後者の条件で C^1 級を定義している。

2.7.3 諸条件の間の関係を振り返る

多変数関数の全微分可能性、偏微分可能性、C^1 級等の条件の間の関係を振り返ってみよう。

まず、1 変数関数の場合は非常に簡単である (要点は「微分可能ならば連続」くらいで、証明も高校数学)。

1 変数関数の場合

$$C^1 \text{級} \quad \Longrightarrow \quad \text{微分可能} \quad \Longrightarrow \quad \text{連続}$$

(C^1 級とは、微分可能かつ導関数が連続なことであるから、左の $\Longrightarrow$ は明らかである。右の $\Longrightarrow$ は高校数学である。)

多変数関数の場合は、微分に (大きく分けて) 二つの概念があり、C^1 級の概念もやや覚えにくい (実際、勘違いして覚えている人がかなり多い)。

多変数関数の場合

$$
\begin{array}{ccccc}
C^1 \text{ 級} & \Longrightarrow & \text{微分可能} & \Longrightarrow & \text{連続} \\
 & & & \searrow & \text{各変数につき偏微分可能}
\end{array}
$$

(C^1 級とは、各変数につき偏微分可能かつすべての 1 階偏導関数が連続ということである。一番左の $\Longrightarrow$ も明らかではない。右の $\Longrightarrow$ は 1 変数の場合と本質的に同じ。$\searrow$ は 1 変数関数にはなかったもので重要。)

関数が全微分可能であることをどう示すか、考えてみよう。

一般的戦略　まず偏微分を試みる。偏微分可能でなければ全微分可能ではないことが分かる。

すべての変数について偏微分可能であることを確認した後、偏導関数が連続であるかどうかを調べる。連続であれば、もとの関数は C^1 級であるから、全微分可能である。

連続でない場合は、各 i について

$$\lim_{h\to 0}\frac{1}{\|h\|}\left(f_i(a+h)-f_i(a)-\sum_{j=1}^{n}\frac{\partial f_i}{\partial x_j}(a)h_j\right)=0$$

かどうかを調べる。これが成り立つことが f が全微分可能であるための必要十分条件である。

実際的には…　連続関数でやったのと同様に (合成関数の微分はまだ説明していないので、フライングになるが)、C^∞ 級の関数から組み立てた関数は C^∞ 級である。多項式関数や、高校で学んだ多くの関数は定義域全体で C^∞ 級である。(例外は、$f(x)=|x|$ の $x=0$, $f(x)=\sqrt{x}$ の $x=0$ など、ごく少数しかない。)

問 2.7.1 次の各関数は、$(x,y)\neq(0,0)$ の範囲で C^∞ 級であることは明らかであるが、$\mathbf{R}^2$ 全体で (a) 連続である, (b) 各変数につき偏微分可能である, (c) 全微分可能である, (d) C^1 級である, の各条件を満たすかどうか調べよ。

(1) $f(x,y)=\sqrt{x^2+y^2}$

(2) $f(x,y)=\begin{cases}\dfrac{xy}{x^2+y^2} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$

(3) $f(x,y)=\begin{cases}\dfrac{xy}{\sqrt{x^2+y^2}} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$

(4) $f(x,y)=\begin{cases}(x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$

(5) $f(x,y)=\begin{cases}xy\dfrac{x^2-y^2}{x^2+y^2} & ((x,y)\neq(0,0))\\ 0 & ((x,y)=(0,0))\end{cases}$

2.7.4 微分の例

Jacobi 行列, gradient ベクトル

Ω は $\mathbf{R}^n$ の開集合、$a \in \Omega$, $f\colon \Omega \to \mathbf{R}^m$ が a で微分可能とするとき、$f'(a)$ は行列であった。具体的には、$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ とおくとき、

$$f'(a) = \left(\frac{\partial f_i}{\partial x_j}(a)\right) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \cdots & \dfrac{\partial f_1}{\partial x_n}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \cdots & \dfrac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x_1}(a) & \dfrac{\partial f_m}{\partial x_2}(a) & \cdots & \dfrac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

この行列を f の a における**ヤコビ行列** (the Jacobian matrix of f at a) と呼ぶ。

さて、$m = 1$ の場合を考えよう。このとき、

$$f'(a) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) & \dfrac{\partial f}{\partial x_2}(a) & \cdots & \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}$$

と、ヤコビ行列は 1 行 n 列の行列、すなわち n 次元横ベクトルになる。この転置である n 次元縦ベクトルを $\operatorname{grad} f(a)$ または $\nabla f(a)$ で表し、f の a における **gradient** (**勾配ベクトル**) と呼ぶ:

$$\operatorname{grad} f(a) = \nabla f(a) := f'(a)^T = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

記号 ∇ は単独でも **nabla** (ナブラ) と呼ばれ、

$$\nabla = \begin{pmatrix} \dfrac{\partial}{\partial x_1} \\ \vdots \\ \dfrac{\partial}{\partial x_n} \end{pmatrix}$$

という意味で用いられる。いわゆる**ベクトル解析**では多用される。

余談 2.7.3 (全微分という言葉とヤコビ行列の記号) 全微分に関する用語と記法はかなりの揺れがある。「全微分とは、写像 $h \mapsto f'(a)h$ のことである」とか、それを $df = \sum_{j=1}^{n} \dfrac{\partial f}{\partial x_j}(a) dx_j$ と「微分形式」で書くとか。ヤコビ行列のことは $J_f(a)$ と書くとか。私は良く使われる記号や名称はなるべく尊重しようと考えているが (言葉遣いに関しては、保守的と言われても、慎重になるべきだと考えている)、この全微分に関しては、もうそういう過去のしがらみは忘れて、1 変数のときと同じ記号 $f'(a)$ を使って、「全微分係数」でなく単に「微分係数」と言うことにしよう、と思っている。

$$\text{微分係数 } f'(a) = \text{ヤコビ行列 } \left(\frac{\partial f_i}{\partial x_j}(a)\right)$$

と頭に入れましょう。■

1 次関数, 2 次関数

例 2.7.6 (1 次関数の微分) $A=(a_{ij})\in M(m,n;\mathbf{R})$, $b=(b_i)\in\mathbf{R}^m$, $c\in\mathbf{R}$ とするとき、

$$f(x):=Ax+b\quad(x\in\mathbf{R}^n)$$

により $f\colon\mathbf{R}^n\to\mathbf{R}^m$ を定めると、f は $\mathbf{R}^n$ で全微分可能で、

$$f'(x)=A.$$

(証明 1) $f=\begin{pmatrix}f_1\\ \vdots\\ f_m\end{pmatrix}$ とおくと、

$$f_i(x)=\sum_{i=1}^m a_{ij}x_j+b_i.$$

ゆえに

$$\frac{\partial f_i}{\partial x_j}(x)=\frac{\partial}{\partial x_j}\left(\sum_{k=1}^n a_{ik}x_k+b_i\right)=\sum_{k=1}^n a_{ik}\frac{\partial}{\partial x_j}x_k+\frac{\partial}{\partial x_j}b_i=\sum_{k=1}^n a_{ik}\delta_{jk}+0=a_{ij}.$$

ここで **Kronecker のデルタ** $\delta_{jk}=\begin{cases}1 & (j=k)\\ 0 & (j\neq k)\end{cases}$ と、任意の数列 $\{A_k\}_{1\le k\le n}$ について成り立つ公式 $\displaystyle\sum_{k=1}^n A_k\delta_{kj}=A_j$ を用いた。

$\dfrac{\partial f_i}{\partial x_j}$ は定数関数であるから (あるいは多項式関数であるから)、$\mathbf{R}^n$ で連続である。ゆえに f は $\mathbf{R}^n$ で C^1 級であるから、全微分可能で、

$$f'(x)=\left(\frac{\partial f_i}{\partial x_j}(x)\right)=(a_{ij})=A.$$

証明の途中ですが…もっと具体的な例。

$$\begin{pmatrix}y_1\\ y_2\end{pmatrix}=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}+\begin{pmatrix}7\\8\end{pmatrix}.$$

$y_1=x_1+2x_2+3x_3+7$, $y_2=4x_1+5x_2+6x_3+8$ であるから、

$$\frac{\partial y_1}{\partial x_1}=1,\quad\frac{\partial y_1}{\partial x_2}=2,\quad\frac{\partial y_1}{\partial x_3}=3,$$
$$\frac{\partial y_2}{\partial x_1}=4,\quad\frac{\partial y_2}{\partial x_2}=5,\quad\frac{\partial y_2}{\partial x_3}=6$$

であり、ヤコビ行列は

$$\begin{pmatrix}\dfrac{\partial y_1}{\partial x_1}&\dfrac{\partial y_1}{\partial x_2}&\dfrac{\partial y_1}{\partial x_3}\\[2mm]\dfrac{\partial y_2}{\partial x_1}&\dfrac{\partial y_2}{\partial x_2}&\dfrac{\partial y_2}{\partial x_3}\end{pmatrix}=\begin{pmatrix}1&2&3\\4&5&6\end{pmatrix}.$$

(証明 2) $\forall h\in\mathbf{R}^n$ に対して、$f(a+h)-f(a)=A(a+h)+b-(Aa+b)=Ah$ であるから、$h\neq 0$ のとき

$$\frac{f(a+h)-f(a)-Ah}{\|h\|}=0.$$

ゆえに

$$\lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{\|h\|}=\lim_{h\to 0}0=0.$$

もちろん、$A\in M(m,n;\mathbf{R})$. ゆえに f は a で全微分可能で、$f'(a)=A$. ■

例 2.7.7 (2 変数 2 次関数) a, b, c, p, q, r を実定数とするとき、

$$f(x,y)=ax^2+2bxy+cy^2+px+qy+r$$

で定まる $f\colon \mathbf{R}^2\to\mathbf{R}$ に対して、 $f'(x,y)$ を求めよ。

解 $\dfrac{\partial f}{\partial x}$ と $\dfrac{\partial f}{\partial y}$ は、ともに x と y の多項式関数で、連続であるから、f が C^1 級であることが分かる[7]。ゆえに全微分可能であり、導関数は、偏導関数を並べた

$$f'(x,y)=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right)=(2ax+2by+p,2cy+2bx+q).$$

ちなみに

$$\nabla f(x,y)=f'(x,y)^T=\begin{pmatrix}2ax+2by+p\\2cy+2bx+q\end{pmatrix}.\quad ■$$

例 2.7.8 (2 次関数の微分) n 変数 x_1, $\cdots$, x_n に関する 2 次関数とは、

$$f(x)=\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j+\sum_{i=1}^n b_ix_i+c\quad (x\in\mathbf{R}^n)$$

という式で表される関数だということには誰でも賛成してくれるであろう (ここで a_{ij}, b_i, $c\in\mathbf{R}$)。

$A:=(a_{ij})\in M(n,n;\mathbf{R})$, $b:=(b_i)\in\mathbf{R}^n$, $x:=(x_i)\in\mathbf{R}^n$ とおくと、

$$f(x)=\frac{1}{2}(Ax,x)+(b,x)+c.$$

実際、$\displaystyle\sum_{j=1}^n a_{ij}x_j$ は Ax の第 i 成分であるから、$\displaystyle\sum_{i=1}^n\sum_{j=1}^n a_{ij}x_ix_j=\sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}x_j\right)x_i=(Ax,x)$.

このままでは、関数の表示に一意性がないためもあって、通常

$$a_{ij}=a_{ji}\quad(\forall i,j)$$

を仮定する (例えば $3x_1x_2+x_2x_1$, $x_1x_2+3x_2x_1$, $2x_1x_2+2x_1x_2$ のいずれも同じ関数を表すので、係数を等しく割り振った $2x_1x_2+2x_2x_1$ のみ使うことにする、ということである)。つまり行列 A に対称性を仮定する (線形代数で習う 2 次形式と同じ)。

[7]一般に、多項式関数は何度偏微分しても多項式関数で、それは連続であるから、多項式関数は C^∞ 級であることが分かる。

$$
\begin{aligned}
\frac{\partial f}{\partial x_j} &= \frac{\partial}{\partial x_j}\left(\frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}x_i x_k + \sum_{i=1}^{n} b_i x_i + c\right) \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}\frac{\partial}{\partial x_j}(x_i x_k) + \sum_{i=1}^{n} b_i \frac{\partial}{\partial x_j} x_i + \frac{\partial}{\partial x_j} c \\
&= \frac{1}{2}\sum_{i=1}^{n}\sum_{k=1}^{n} a_{ik}(\delta_{ji}x_k + x_i\delta_{jk}) + \sum_{i=1}^{n} b_i\delta_{ji} + 0 \\
&= \frac{1}{2}\left(\sum_{k=1}^{n}\left(\sum_{i=1}^{n} a_{ik}\delta_{ji}\right)x_k + \sum_{i=1}^{n}\left(\sum_{k=1}^{n} a_{ik}\delta_{jk}\right)x_i\right) + b_j \\
&= \frac{1}{2}\left(\sum_{k=1}^{n} a_{jk}x_k + \sum_{i=1}^{n} a_{ij}x_i\right) + b_j.
\end{aligned}
$$

ここで $\sum_{k=1}^{n} a_{jk}x_k$ は Ax の第 j 成分である。また対称性より $\sum_{i=1}^{n} a_{ij}x_i = \sum_{i=1}^{n} a_{ji}x_i$ も Ax の第 j 成分である。ゆえに

$$
\frac{\partial f}{\partial x_j}(x) = \text{「}Ax \text{ の第 } j \text{ 成分」} + b_j = \text{「}Ax + b \text{ の第 } j \text{ 成分」}.
$$

ゆえに

$$
\begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix} = Ax + b.
$$

ゆえに

$$
\nabla f(x) = f'(x)^T = Ax + b.
$$

偏導関数 $\dfrac{\partial f}{\partial x_j}$ は $x_1, \ldots, x_n$ についての多項式関数なので、$\mathbf{R}^n$ で連続である。ゆえに f は $\mathbf{R}^n$ で C^1 級であるので、全微分可能である。 ■

2.7.5 微分の意味

線形化写像は元の関数を近似する $f(x) \fallingdotseq f(a) + f'(a)(x-a)$

まず「全微分」の定義を復習しよう。f が a で全微分可能であるとは、

$$
\exists A \in M(m, n; \mathbf{R}) \quad \text{s.t.} \quad \lim_{h \to 0} \frac{1}{\|h\|}(f(a+h) - f(a) - Ah) = 0
$$

が成り立つことをいい (Ah は行列 A とベクトル h の積を意味していることに注意する)、一意的に定まる行列 A のことを f の a における全微分係数と呼び、$f'(a)$ で表す。

これから、f が a で全微分可能ならば

$$
\lim_{h \to 0} \frac{1}{\|h\|}(f(a+h) - f(a) - f'(a)h) = 0
$$

が成り立つ。$h \to 0$ のとき、分母は 0 に収束するので、分子はそれよりも速く 0 に収束する、ということである。このことを

$$
f(a+h) - f(a) - f'(a)h = o(\|h\|) \quad (h \to 0)
$$

とも書く (**Landau の little "o" notation**)。

不正確な書き方になるが[8]、$\|h\|$ が十分小さいとき、

$$f(a+h) \fallingdotseq f(a) + f'(a)h$$

が成り立つ。言い方を変えると、x が a に十分近いとき、

$$f(x) \fallingdotseq f(a) + f'(a)(x-a)$$

が成り立つ。この右辺の式で表される写像、すなわち

$$\Omega \ni x \mapsto f(a) + f'(a)(x-a) \in \mathbf{R}^m$$

を f の (a における) **線形化写像** (**1 次近似**) と呼ぶ (これはちゃんとした定義である)。

$\mathrm{grad}\, F$ はレベルセットの法線ベクトル

以下、Ω は $\mathbf{R}^n$ の開集合、$F\colon \Omega \to \mathbf{R}$ は C^1 級とする。$c \in \mathbf{R}$ に対して、

$$L_c := \{x \in \Omega; F(x) = c\}$$

とおき、F のレベル c の**レベル・セット** (level set) と呼ぶ。特に 3 次元の場合は**等値面** (**iso-surface**), 2 次元の場合は**等値線** (**isoline**) とも呼ぶ。

$L_c = \emptyset$ となることもありうる。$a \in \Omega$ のとき $c := F(a)$ とおくと、$a \in L_c$ が成り立つので、$L_c \neq \emptyset$ である。

例 2.7.9 (円はレベル・セットとして扱える) $\Omega = \mathbf{R}^2$, $F(x,y) := x^2 + y^2$ とするとき、L_1 は原点中心半径 1 の円周, L_2 は原点中心半径 $\sqrt{2}$ の円周, L_0 は原点 1 点からなる集合 $\{(0,0)\}$, $L_{-1} = \emptyset$. ■

例 2.7.10 (等高線とレベル・セット) Ω をある地図、$(x,y) \in \Omega$ に対して、$F(x,y)$ は (x,y) という地点の標高とするとき、L_c は高さ c の**等高線** (**contour line**) である。 ■

実は、**$\nabla F(a)$ は、L_c の点 a における法線ベクトル**と呼ぶにふさわしいものである。実際、

$$\lim_{x \to a} \frac{1}{\|x-a\|} \left(F(x) - F(a) - F'(a)(x-a)\right) = 0$$

が成り立つが、$c := F(a)$ とするとき、$\forall x \in L_c$ に対して $F(x) = c$ であることと、$F'(a)(x-a) = (\nabla F(a), x-a)$ であることから、

$$\begin{aligned}
\frac{1}{\|x-a\|} \left(F(x) - F(a) - F'(a)(x-a)\right) &= \frac{1}{\|x-a\|} \left(c - c - (\nabla F(a), x-a)\right) \\
&= -\frac{1}{\|x-a\|} (\nabla F(a), x-a) \\
&= -\left(\nabla F(a), \frac{x-a}{\|x-a\|}\right)
\end{aligned}$$

となるので、

$$\lim_{\substack{x \in L_c \\ x \to a}} \left(\nabla F(a), \frac{x-a}{\|x-a\|}\right) = 0.$$

[8] $\fallingdotseq$ というのは、数学語ではない。

(この辺のの議論は「発見的考察」であり、「証明」ではない。盲目的に信じ込まれても困るので、少し疵を指摘しておくと、L_c がどういう集合であるか、一般には良く分からない。そもそも L_c をたどって、a に近づくことが出来るのかも分からない。後で「陰関数定理」という定理を学ぶとこのあたりのことが解決される。)

$\nabla F(a) \neq 0$, $x \neq a$ であるとき、$\nabla F(a)$ と $x - a$ のなす角 $\theta(x) \in [0, \pi]$ が

$$(\nabla F(a), x - a) = \|\nabla F(a)\| \, \|x - a\| \cos\theta(x)$$

により定義できる。このとき

$$\lim_{\substack{x \in L_c \\ x \to a}} \cos\theta(x) = 0$$

であるから

$$\lim_{\substack{x \in L_c \\ x \to a}} \theta(x) = \frac{\pi}{2}. \tag{$\angle R$}$$

(注意: $x \to a$ のとき $\dfrac{x-a}{\|x-a\|}$ 自身が何かあるベクトルに収束するというわけではなく、収束が保証出来るのは角度だけである。)

以上の考察を背景に、次のように定義する。

定義 2.7.11 Ω が $\mathbf{R}^n$ の開集合、$F\colon \Omega \to \mathbf{R}$ が C^1 級、$a \in \Omega$, $\nabla F(a) \neq 0$ とする。このとき、

$$\{x \in \mathbf{R}^n; (\nabla F(a), x - a) = 0\}$$

を、L_c の a における**接超平面** ($n = 3$ のときは単に**接平面**, $n = 2$ のときは**接線**), $\nabla F(a)$ と同じ方向を持つベクトル ($\neq 0$) を、L_c の a における**法線ベクトル** (normal vector) と呼ぶ。

また点 a を通り $\nabla F(a)$ を方向ベクトルとする直線を、L_c の a における**法線** (normal line) と呼ぶ。

現時点で ($\angle R$) は、根拠にあいまいなところがあるので、「∇F が法線ベクトル」というのは**定義したこと**である。この点、何かずるいと感じられるかも知れないが[9]、後で陰関数定理を学ぶと、($\angle R$) が根拠を持った主張として浮上して来る。

$\nabla F(a)$ は点 a において、F の変化が最も急な方向を表す (地図解釈では、傾斜がもっとも急な方向ということになる)。

証明もどき

$$F(a + h) - F(a) = (\nabla F(a), h) + o(\|h\|) \quad (h \to 0).$$

右辺の第 2 項は h より高位の無限小だから、右辺第 1 項が F の増分の主部といえる。h と $\nabla F(a)$ のなす角を $\theta(x) \in [0, \pi]$ とすると、

$$|(\nabla F(a), h)| = \|\nabla F(a)\| \, \|h\| \cos\theta(x)$$

が成り立ち、

$$-\|\nabla F(a)\| \, \|h\| \leq (\nabla F(a), h) \leq \|\nabla F(a)\| \, \|h\| \, .$$

[9] この辺の事情は、曲線の弧長の定義のそれと少し似ている。

左の不等号の等号は、$\theta(x)=\pi$, すなわち $\exists\lambda\le 0$ s.t. $h=\lambda\nabla F(a)$ のとき、右の不等号の等号は、$\theta(x)=0$, すなわち $\exists\lambda\ge 0$ s.t. $h=\lambda\nabla F(a)$ のとき、成立する。つまり、$\nabla F(a)$ の方向に移動すると、高さの増加が最も大きい。 ■

例 2.7.12 (平面内の直線の接線) $A,B\in\mathbf{R}^2$, $(A,B)\ne(0,0)$ とするとき、

$$F(x,y):=Ax+By$$

は C^∞ 級の $F\colon\mathbf{R}^2\to\mathbf{R}$ を定める。$\forall x\in\mathbf{R}$ に対して、F の高さ c のレベルセット

$$L_c=\left\{(x,y)\in\mathbf{R}^2;F(x,y)=c\right\}=\left\{(x,y)\in\mathbf{R}^2;Ax+By=c\right\}$$

はベクトル $\begin{pmatrix}A\\B\end{pmatrix}$ に垂直な直線である、ということは良く知られている。$\nabla F(x,y)=\begin{pmatrix}F_x\\F_y\end{pmatrix}=\begin{pmatrix}A\\B\end{pmatrix}$ であり、確かに ∇F は L_c の法線ベクトルである。

$(a,b)\in L_c$ とする。このとき、L_c の (a,b) における接超平面は

$$A(x-a)+B(y-b)=0.$$

仮定より $Aa+Bb=c$ であるから、この方程式を整理すると

$$Ax+By=c.$$

つまり接超平面は接線で、L_c 自身に一致する。 ■

余談 2.7.4 (直線 $Ax+By=c$ がベクトル (A,B) に垂直なこと) 昔は常識だったはずだけど、こういうのも説明しないといけないのかな。

(a) 直線 $Ax+By=c$ は、$\boldsymbol{n}=(A,B)^T$ に垂直である。

(b) 点 $\mathrm{C}(x_0,y_0)$ を通り、$\boldsymbol{n}=(A,B)^T$ に垂直な直線の方程式は、$A(x-x_0)+B(y-y_0)=0$.

証明

(a) 直線上の点 $\mathrm{C}(x_0,y_0)$ を一つ取る ($Ax_0+By_0=c$ が成り立つ)。

直線上の任意の点 $\mathrm{P}(x,y)$ に対して、$Ax+By=c$ が成り立つが、辺々引き算して $A(x-x_0)+B(y-y_0)=0$. これは $\boldsymbol{n}\cdot\overrightarrow{\mathrm{CP}}=0$ と書き直せるので、$\boldsymbol{n}\perp\overrightarrow{\mathrm{CP}}$ を示している。

(b) $\mathrm{P}(x,y)$ が直線上にあるためには、$\boldsymbol{n}\perp\overrightarrow{\mathrm{CP}}$ が必要十分である。書き直すと $A(x-x_0)+B(y-y_0)=0$. ■

以上は2次元の話であるが、何次元でも通用する話で、3次元の場合も重要である。方程式 $Ax+By+Cz=d$ は $(A,B,C)^T$ に垂直な平面を表す。 ■

例 2.7.13 (平面内の円の法線ベクトルと接線) $F(x,y):=x^2+y^2$ とすると、$\nabla F(x,y)=\begin{pmatrix}F_x\\F_y\end{pmatrix}=\begin{pmatrix}2x\\2y\end{pmatrix}$. $R>0$ とするとき、$L_{R^2}=\{(x,y)\in\mathbf{R}^2;F(x,y)=R^2\}$ は、原点中心、半径 R の円周である。$(a,b)\in L_{R^2}$ とするとき、

$$\nabla F(a,b)=\begin{pmatrix}2a\\2b\end{pmatrix}$$

であるが、これは確かに L_{R^2} の (a,b) における法線ベクトルである (円周上の点 (a,b) を通る円の半径は、円の接線と直交するので、円の中心から (a,b) に向うベクトル $\begin{pmatrix} a \\ b \end{pmatrix}$ は L_c の法線ベクトルである)。

(a,b) における接超平面 (これも 2 次元なので接線) は

$$\nabla F(a,b)\cdot\begin{pmatrix} x-a \\ y-b \end{pmatrix}=0.$$

すなわち

$$2a(x-a)+2b(y-b)=0.$$

$a^2+b^2=R^2$ を用いて整理すると $ax+by=R^2$. ■

例 2.7.14 (2 変数関数のグラフの接線) Ω は $\mathbf{R}^2$ の開集合、$f\colon\Omega\to\mathbf{R}$ とするとき、f のグラフ $\operatorname{graph} f=\{(x,y,z);(x,y)\in\Omega,\ z=f(x,y)\}$ は、

$$F(x,y,z):=f(x,y)-z\quad((x,y,z)\in\Omega\times\mathbf{R})$$

のレベル 0 のレベルセットである。(a,b,c) (ただし $c=f(a,b)$) における接平面は

$$\begin{pmatrix} f_x(a,b) \\ f_y(a,b) \\ -1 \end{pmatrix}\cdot\begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix}=0.$$

$$f_x(a,b)(x-a)+f_y(a,b)(y-b)+(-1)(z-c)=0.$$

$c=f(a,b)$ であるから、

$$z=f_x(a,b)(x-a)+f_y(a,b)(y-b)+f(a,b).$$

これは

$$z=f'(\boldsymbol{a})\,(\boldsymbol{x}-\boldsymbol{a})+f(\boldsymbol{a})$$

と書ける。つまり、f の線形化写像のグラフである。■

例題 2.7.1 楕円 $x^2/a^2+y^2/b^2=1$ (a, b は正定数) 上の点 (x_0,y_0) における接線の方程式を求めよ。

解 $F(x,y):=x^2/a^2+y^2/b^2-1$ とおくと、楕円は $F(x,y)=0$ となり、

$$\nabla F\,(x_0,y_0)=\begin{pmatrix} \dfrac{\partial F}{\partial x}(x_0,y_0) \\ \dfrac{\partial F}{\partial y}(x_0,y_0) \end{pmatrix}=\begin{pmatrix} 2x_0/a^2 \\ 2y_0/b^2 \end{pmatrix}.$$

これが 0 にならないから ($\because (x_0,y_0)\neq(0,0)$)、$(x_0,y_0)$ における楕円の接線の法線ベクトルとなる。したがって、接線の方程式は

$$\frac{2x_0}{a^2}\,(x-x_0)+\frac{2y_0}{b^2}\,(y-y_0)=0.$$

これを整理すると

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=\frac{x_0^2}{a^2}+\frac{y_0^2}{b^2}.$$

点 (x_0, y_0) が楕円上の点であるから、$\dfrac{x_0^2}{a^2}+\dfrac{y_0^2}{b^2}=1$ が成り立つので、

$$\frac{x_0 x}{a^2}+\frac{y_0 y}{b^2}=1. \quad \blacksquare$$

問 2.7.2 (1) 関数 $f(x,y):=\sqrt{17-5x^2-4y^2}$ のグラフの $(-1,1,2\sqrt{2}\,)$ における接平面の方程式を求めよ。
(2) 関数 $F(x,y,z):=\dfrac{(x+1)^2}{1}+\dfrac{y^2}{4}+\dfrac{(z-1)^2}{9}$ のレベルセット $L_1=\{(x,y,z)\in\mathbf{R}^3; F(x,y,z)=1\}$ について、次の問に答えよ。(a) $\left(-\dfrac{2}{3},\dfrac{4}{3},3\right)\in L_1$ における、L_1 の接平面の方程式と法線を求めよ。(b) 平面 $x+y+z=k$ (k は実定数) が L_1 の接平面になるように、k の値を求めよ。

問 (2-b) を解くための例題　楕円 $F(x,y):=x^2/4+y^2/3=1$ に直線 $x+y=k$ が接するような k を求めよ。もちろん高校数学的に判別式を考えても解ける。(a,b) が $F(x,y)=1$ 上にあり、$\nabla F(a,b)\parallel(1,1)$ となるように

$$F(a,b)=1 \quad \wedge \quad \exists t\in\mathbf{R} \quad \nabla F(a,b)=t\begin{pmatrix}1\\1\end{pmatrix}$$

から $(a,b)=\pm\left(\frac{4}{\sqrt{7}},\frac{3}{\sqrt{7}}\right)$. これから $k=a+b=\pm\sqrt{7}$. $\blacksquare$

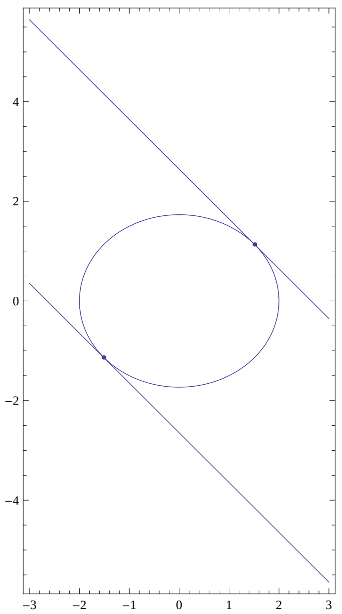


図 2.1: 楕円と接線

2.7.6　高校数学の補足

直線 $Ax+By=c$ は、$\boldsymbol{n}=\begin{pmatrix}A\\B\end{pmatrix}$ に垂直である。

証明　直線上の点 $\mathrm{C}(x_0,y_0)$ を一つ取る ($Ax_0+By_0=c$ が成り立つ)。
直線上の任意の点 $\mathrm{P}(x,y)$ に対して、$Ax+By=c$ が成り立つが、辺々引き算して $A(x-x_0)+B(y-y_0)=0$. これは $\boldsymbol{n}\cdot\overrightarrow{\mathrm{CP}}=0$ と書き直せるので、$\boldsymbol{n}\perp\overrightarrow{\mathrm{CP}}$ を示している。$\blacksquare$

点 $\mathrm{C}(x_0,y_0)$ を通り、$\boldsymbol{n}=\begin{pmatrix}A\\B\end{pmatrix}$ に垂直な直線の方程式は、$A(x-x_0)+B(y-y_0)=0$.

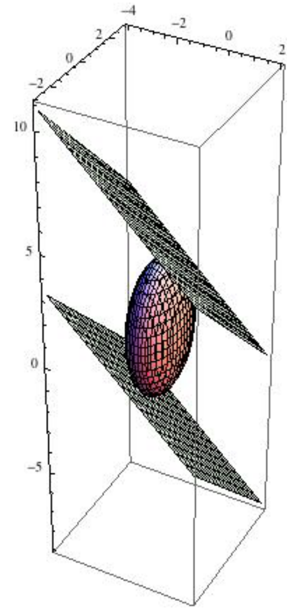


図 2.2: 楕円体 $\dfrac{(x+1)^2}{1}+\dfrac{y^2}{4}+\dfrac{(z-1)^2}{9}=1$ と接平面 $x+y+z=k$

```
g1=ContourPlot3D[(x+1)^2/1 +y^2/4+(z-1)^2/9==1,{x,-4,2}, {y,-3,3},{z,-3,5}]
g2=Plot3D[{-x,-y,-Sqrt[14],-x-y+Sqrt[14]},{x,-4,2},{y,-3,3}]
Show[g1,g2,BoxRatios->Automatic,PlotRange->All]
```

証明 $\mathrm{P}(x,y)$ が直線上にあるための条件は、$\boldsymbol{n}\perp\overrightarrow{\mathrm{CP}}$ である。書き直すと $A(x-x_0)+B(y-y_0)=0$. ■

2.8 合成関数の微分法

非常に重要であるが、一つの難所である。簡単のようでいて難しいところがある。

- 公式そのものは、1 変数の場合の素直な拡張とみなせるので、覚えるのは簡単である。
- 具体的な合成関数を微分するのは簡単、というか、ひょっとすると不要なのかもしれない。使わずに解ける問題が実は多い。(要は、多変数関数でも、偏微分を計算すれば良いので、結局は 1 変数関数の問題になってしまう場合が多い。)
- 具体的な関数でない場合にとても重要なものが多く (例えば微分方程式を変数変換する)、しかも計算が非常に面倒というのが少なくない。
- 具体的な計算問題のドリルをこなしてマスターする、という学校数学の勉強法は通用しない。

2.8.1 定理と証明

1 変数の場合

$y = f(x)$, $z = g(y)$ がいずれも微分可能であれば、合成関数 $z = g(f(x))$ も微分可能で、

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

つまり

$$(g \circ f)'(a) = g'(b)f'(a) \quad (\text{ただし } b = f(a))$$

である。

これは多変数の場合にも、自然に拡張できる。右辺をヤコビ行列の積とみなせばよい。

定理 2.8.1 (合成関数の微分法, chain rule (連鎖律)) Ω, D はそれぞれ $\mathbf{R}^n$, $\mathbf{R}^m$ の開集合で、$a \in \Omega$, $f\colon \Omega \to \mathbf{R}^m$, $g\colon D \to \mathbf{R}^\ell$, $f(\Omega) \subset D$, $b = f(a)$, f は a で微分可能、g は b で微分可能ならば、$g \circ f$ は a で微分可能で、

$$(g \circ f)'(a) = g'(b)f'(a).$$

$y = f(x)$, $z = g(y)$ と書けば、上式の (i, j) 成分は

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k}\frac{\partial y_k}{\partial x_j} \quad (1 \le i \le \ell,\ 1 \le j \le n).$$

証明の前に、1 次関数だったら当たり前、ということを見よう。$f(x) = Ax+b$, $g(y) = Cy+d$ とすると、

$$g \circ f(x) = g\left(f(x)\right) = g(Ax+b) = C(Ax+b)+d = CAx+(Cb+d).$$

これから、

$$(g \circ f)'(x) = CA = g'(y)f'(x).$$

余談 2.8.1 (証明もどき) 高校数学の某教科書には、合成関数の微分法に次のような「証明」がついていた。

「証明」

$$\frac{dz}{dx} = \lim_{\Delta x \to 0} \frac{\Delta z}{\Delta x} = \lim_{\Delta x \to 0} \left(\frac{\Delta z}{\Delta y}\frac{\Delta y}{\Delta x}\right).$$

z は y の関数として微分可能なので、連続であるから、$\Delta x \to 0$ のとき、$\Delta y \to 0$ である。ゆえに

$$\frac{dz}{dx} = \lim_{\Delta y \to 0} \frac{\Delta z}{\Delta y} \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \frac{dz}{dy}\frac{dy}{dx}.$$

しかし $\Delta x \ne 0$ であっても $\Delta y \ne 0$ とは限らないので、

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y}\frac{\Delta y}{\Delta x}$$

は一般に成り立つ式ではない。つまり上の「証明」には穴がある (証明ではない)。こういう杜撰なことを許せば、上の定理の証明は簡単に出来るのだが…以下に紹介する証明は、補助関数の導入によって、その問題をていねいに回避するものであると言える。 ■

証明 f が a で全微分可能であるから、

$$\varepsilon_1(x) := f(x) - f(a) - f'(a)(x-a) \quad (x \in \Omega) \tag{2.2}$$

とおくと

$$\lim_{x\to a} \frac{\varepsilon_1(x)}{\|x-a\|} = 0. \tag{2.3}$$

また g が $b = f(a)$ で全微分可能であるから、

$$\varepsilon_2(y) := g(y) - g(b) - g'(b)(y-b) \quad (y \in D) \tag{2.4}$$

とおくと

$$\lim_{y\to b} \frac{\varepsilon_2(y)}{\|y-b\|} = 0. \tag{2.5}$$

(2.4) から得られる

$$g(y) - g(b) = g'(b)(y-b) + \varepsilon_2(y) \quad (y \in D)$$

に $y = f(x)$ を代入して ($b = f(a)$ に注意して)

$$\begin{aligned}(g\circ f)(x) - (g\circ f)(a) &= g(f(x)) - g(f(a)) \\ &= g'(b)(f(x) - f(a)) + \varepsilon_2(f(x)) \quad (x \in \Omega).\end{aligned}$$

この式の右辺に、(2.2) から得られる

$$f(x) - f(a) = f'(a)(x-a) + \varepsilon_1(x)$$

を代入して

$$\begin{aligned}(g\circ f)(x) - (g\circ f)(a) &= g'(b)\left[f'(a)(x-a) + \varepsilon_1(x)\right] + \varepsilon_2(f(x)) \\ &= g'(b)f'(a)(x-a) + g'(b)\varepsilon_1(x) + \varepsilon_2(f(x)).\end{aligned}$$

ゆえに

$$\lim_{x\to a} \frac{g'(b)\varepsilon_1(x)}{\|x-a\|} = 0 \tag{2.6}$$

と

$$\lim_{x\to a} \frac{\varepsilon_2(f(x))}{\|x-a\|} = 0 \tag{2.7}$$

を証明できれば、

$$\lim_{x\to a} \frac{g'(b)\varepsilon_1(x) + \varepsilon_2(f(x))}{\|x-a\|} = 0$$

となるので、$g\circ f$ が a で全微分可能で $(g\circ f)'(a) = g'(b)f'(a)$ であることが結論できる。

(2.6) は、(2.3) から

$$\left\|\frac{g'(b)\varepsilon_1(x)}{\|x-a\|}\right\| \le \|g'(b)\| \frac{\|\varepsilon_1(x)\|}{\|x-a\|} \to 0 \quad (x \to a)$$

が成り立つので明らかである。

(2.7) を証明するために、補助関数 M を導入する。

$$M(y) := \begin{cases} \dfrac{\varepsilon_2(y)}{\|y-b\|} & (y \in D \setminus \{b\}) \\ 0 & (y = b) \end{cases}$$

とおくと ($\varepsilon_2(b) = 0$ に注意して)

$$M\colon D \to \mathbf{R}^m, \quad \lim_{y\to b} M(y) = 0, \quad \varepsilon_2(y) = \|y-b\|\, M(y) \quad (y \in D).$$

特に $\varepsilon_2\left(f(x)\right) = \|f(x) - b\|\, M(f(x))$ である。さて、

$$f(x) - b = f(x) - f(a) = f'(a)(x-a) + \varepsilon_1(x)$$

より

$$\|f(x) - b\| = \|f'(a)(x-a) + \varepsilon_1(x)\| \le \|f'(a)\|\,\|x-a\| + \|\varepsilon_1(x)\|\,.$$

ゆえに

$$\begin{aligned}\frac{\|\varepsilon_2(f(x))\|}{\|x-a\|} &= \frac{\|f(x)-b\|\,\|M\left(f(x)\right)\|}{\|x-a\|} \le \left(\|f'(a)\| + \frac{\|\varepsilon_1(x)\|}{\|x-a\|}\right)\|M(f(x))\| \\ &\to (\|f'(a)\| + 0)\cdot 0 = 0 \quad (x \to a).\end{aligned}$$

ゆえに (2.7) が示された。■

2 点ほど指摘しておく。

- 行列のノルムを導入しておいたのが、見通しを良くすることに役立っている。
- 上で補助関数 M を導入する理由は、

$$\frac{\|\varepsilon_2(f(x))\|}{\|x-a\|} = \frac{\|f(x)-b\|}{\|x-a\|}\cdot\frac{\|\varepsilon_2(f(x))\|}{\|f(x)-b\|} \le C\frac{\|\varepsilon_2(f(x))\|}{\|f(x)-b\|} \to C\cdot 0 = 0 \quad (x \to a)$$

  という素朴な「証明」の穴をふさぐためである。$x \ne a$ であっても、$f(x) \ne b$ であるとは限らないので、分母 $\|f(x) - b\|$ が 0 となりうることに注意する。

2.8.2 例

曲線に沿った微分

例 2.8.2 (chain rule の簡単な例: 曲線に沿った微分) (1) $z = xy$, $x = \varphi(t)$, $y = \psi(t)$ とするとき、

$$\frac{dz}{dt} = \frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial t} = y\frac{dx}{dt} + x\frac{dy}{dt} = \psi(t)\varphi'(t) + \varphi(t)\psi'(t).$$

別解として

$$\frac{dz}{dt} = \frac{d}{dt}\left(\varphi(t)\psi(t)\right) = \varphi'(t)\psi(t) + \varphi(t)\psi'(t).$$

(こうやってしまうと、1 変数関数についての定理だけで済む。)

(2) $f(x,y)=\sqrt{x^2+y^2}$, $x=\varphi(t)$, $y=\psi(t)$ とするとき、

$$\frac{d}{dt}f\left(\varphi(t),\psi(t)\right)=f_x x_t+f_y y_t=\frac{x}{\sqrt{x^2+y^2}}\varphi'(t)+\frac{y}{\sqrt{x^2+y^2}}\psi'(t)=\frac{\varphi(t)\varphi'(t)+\psi(t)\psi'(t)}{\sqrt{\varphi(t)^2+\psi(t)^2}}.$$

これも別解として

$$\frac{d}{dt}f\left(\varphi(t),\psi(t)\right)=\frac{d}{dt}\sqrt{\varphi(t)^2+\psi(t)^2}=\frac{\frac{d}{dt}\left(\varphi(t)^2+\psi(t)^2\right)}{2\sqrt{\varphi(t)^2+\psi(t)^2}}=\frac{\varphi'(t)\varphi(t)+\psi'(t)\psi(t)}{\sqrt{\varphi(t)^2+\psi(t)^2}}.$$

(こうやってしまうと、1 変数関数についての定理だけで済む。) ■

学生との Q&A から　上の例 ($z=xy$, $x=\varphi(t)$, $y=\psi(t)$) で

$$\frac{\partial z}{\partial t}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial t}$$

という chain rule を使ったが、これはどうやって出せるか (定理の「あてはめ」ですね)、という基本的 (ということは重要) な質問をされた。

z が 2 つの変数 x と y の関数である、つまり $z=z(x,y)$ であることを見て取るのが第 1 ステップである。それが分かれば

$$\frac{\partial z}{\partial \boxed{}}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial \boxed{}}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial \boxed{}}$$

と書ける。この $\boxed{}$ に t を入れて出来上がり。

もし u が 3 変数 x, y, z の関数、つまり $u=u(x,y,z)$ であれば、

$$\frac{\partial u}{\partial \boxed{}}=\frac{\partial u}{\partial x}\frac{\partial x}{\partial \boxed{}}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial \boxed{}}+\frac{\partial u}{\partial z}\frac{\partial z}{\partial \boxed{}}$$

となる。

z_i が $y_1,\dots,y_m$ の関数、つまり $z_i=z_i(y_1,\dots,y_m)$ であれば、

$$\frac{\partial z_i}{\partial \boxed{}}=\sum_{k=1}^{m}\frac{\partial z_i}{\partial y_k}\frac{\partial y_k}{\partial \boxed{}}.$$

空間極座標

重要な変数変換に、極座標変換がある。平面極座標は良く知っているだろうから、ここでは 3 次元空間の極座標を紹介する。古い講義ノート (http://www.math.meiji.ac.jp/~mk/tahensuu1/tahensuu1-2011.pdf) の付録に n 次元極座標の紹介がある。

空間に、互いに直交する座標軸 x 軸, y 軸, z 軸を取って座標を入れた xyz 座標系で、(x,y,z) という座標を持つ点を P とする。

- P の原点からの距離を r $(r\ge 0)$
- $\overrightarrow{\mathrm{OP}}$ が z 軸の正方向となす角を θ $(0\le\theta\le\pi)$
- P を xy 平面に正射影した点を P$'$ として、動径 $\overrightarrow{\mathrm{OP}'}$ を x 軸の正の部分から反時計回りに測った角を ϕ $(0\le\phi<2\pi)$

とすると

$$\begin{cases} x = r\sin\theta\cos\phi \\ y = r\sin\theta\sin\phi \\ z = r\cos\theta \end{cases}$$

が成り立つ。このとき、r, θ, ϕ を P の**空間極座標** (**3 次元極座標**あるいは**球 (面) 座標**, **spherical coordinate**) と呼ぶ。

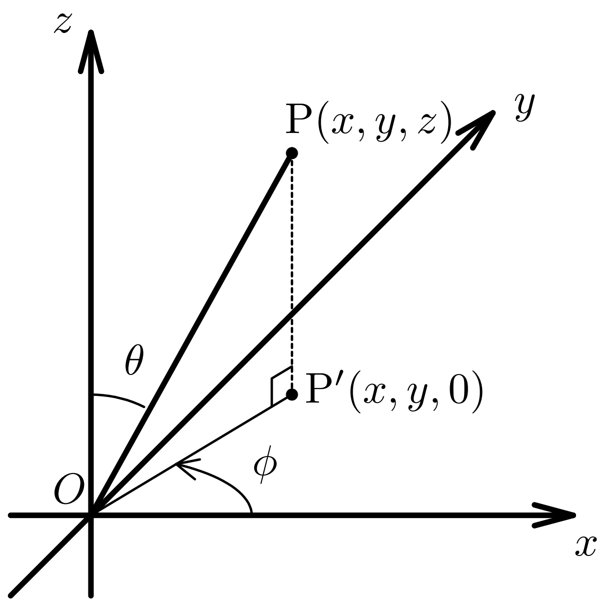


図 2.3: 球座標の θ, ϕ

θ と ϕ の測り方の違いに注意

注意 2.8.3 (極座標とのつきあい方) 3 次元以上の極座標の式には、色々なバリエーションがある。x 軸の正方向から測った角度を θ としたり ($x = r\cos\theta$ となる)、地球儀のように緯度・経度形式にしたり。ここに紹介した式を丸暗記するよりは、導出の仕方を理解して、自分で導出が出来るように練習しておくのが望ましい。 ■

問 2.8.1 地球の表面上にある 2 点の緯度経度が分かっているときに、その 2 点間の (表面に沿っての最短の) 道のりの長さの求め方を説明せよ (ただし地球は球であると考える)。具体的な問題が欲しければ、アレクサンドリアの図書館 (北緯 $31°12' = 31.20°$, 東経 $29°55' = 29.91°$) と生田キャンパス (北緯 $35°37' = 35.61°$, 東経 $139°33' = 139.55°$) の地球表面に沿った道のりの長さを求めよ。(答: 昔、某学生が計算して、9555km くらい、とのことでした。正しいかどうか保証しません。)

例 2.8.4 (変数の極座標変換、2 変数 1 階の場合) 関数の変数をデカルト座標から極座標に「変換」する、つまり関数 $f = f(x, y)$ が与えられたとき、

$$g(r, \theta) := f(x, y), \quad x = r\cos\theta, \quad y = r\sin\theta$$

で定義される関数 g を考える — と分かりやすくなることが (非常にしばしば) ある。

$$g(r, \theta) = f\left(r\cos\theta, r\sin\theta\right)$$

であるから、g は $\varphi(r,\theta) := \begin{pmatrix} r\cos\theta \\ r\sin\theta \end{pmatrix}$ と f との合成関数である。

$$g_r = f_x x_r + f_y y_r, \quad g_\theta = f_x x_\theta + f_y y_\theta$$

であり (これは $(g_r\ g_\theta) = (f_x\ f_y)\begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}$ と書けるので、確かに $(f\circ\varphi)' = f'\varphi'$ が成り立っている)、

$$x_r = \cos\theta, \quad y_r = \sin\theta, \quad x_\theta = -r\sin\theta, \quad y_\theta = r\cos\theta$$

であるから、

$$g_r = f_x\cos\theta + f_y\sin\theta, \quad g_\theta = -f_x r\sin\theta + f_y r\cos\theta.$$

実は g の 2 階偏導関数を f の偏導関数で表したり、逆に f の 2 階偏導関数を g の偏導関数で表すことが重要であるが、それはまた後で紹介する。■

余談 2.8.2 上の例は簡単だが、f を具体的に与えていないため、本質的に 2 変数の例となっている (つまり 1 変数関数の合成関数の微分法だけでは得られない)。

1 変数関数から「組み立てられない[10]」という意味で本質的に多変数の関数の具体例はほとんど知られていない (皆無というわけではないらしいが、1 変数関数が豊富な具体例を持っているのとは対照的である)。従って、具体的な計算問題で、多変数関数の合成関数の微分法の定理が必要となるものを出題することは難しい。これは、本質的に多変数でない関数は応用上現れないということとは違う。応用上現れる関数のほとんどは具体的な式では書けない、ということである。■

2.8.3 高階の導関数

要するに偏導関数を計算すれば良いので、既に述べたことと積の微分法くらいで計算はどんどん出来る。

試しに 1 変数では、

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$$

より

$$\begin{aligned}\frac{d^2z}{dx^2} &= \frac{d}{dx}\left(\frac{dz}{dy}\frac{dy}{dx}\right) = \left(\frac{d}{dx}\frac{dz}{dy}\right)\frac{dy}{dx} + \frac{dz}{dy}\left(\frac{d}{dx}\frac{dy}{dx}\right) \\ &= \frac{d^2z}{dy^2}\frac{dy}{dx}\cdot\frac{dy}{dx} + \frac{dz}{dy}\frac{d^2y}{dx^2} = \frac{d^2z}{dy^2}\left(\frac{dy}{dx}\right)^2 + \frac{dz}{dy}\frac{d^2y}{dx^2}.\end{aligned}$$

偏微分方程式論からの有名な例を 2 つ紹介する。変数変換 (独立変数の変換は、要するに合成関数である！) をして「見方を変える」ことが重要なテクニックである。具体的に分からない関数 (なにしろ未知関数だから！) の合成関数の、高階の偏導関数の計算が必要になるのは仕方がない。

[10] 大学 1 年生が知っているような 1 変数関数と、和・差・積・商、合成、テンソル積を取ることで得られる関数の意味。

例 2.8.5 (変数の極座標変換、2 変数 2 階の場合 (偏微分方程式で重要な Laplacian の極座標表示)) C^2 級の関数 $f\colon (x,y)\mapsto f(x,y)$ があるとき、

$$x=r\cos\theta,\quad y=r\sin\theta,\quad g(r,\theta):=f(x,y),$$

すなわち、

$$g(r,\theta):=f(r\cos\theta,r\sin\theta).$$

このとき

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}=\frac{\partial^2 g}{\partial r^2}+\frac{1}{r}\frac{\partial g}{\partial r}+\frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2}$$

が成り立つ。chain rule と積の微分法により、

$$\begin{aligned}
g_r &= f_x x_r + f_y y_r,\\
g_{rr} &= (f_{xx}x_r+f_{xy}y_r)x_r + f_x x_{rr} + (f_{yx}x_r+f_{yy}y_r)y_r + f_y y_{rr}\\
&= f_{xx}x_r^2 + (f_{xy}+f_{yx})\,x_r y_r + f_{yy}y_r^2 + f_x x_{rr} + f_y y_{rr},\\
g_{\theta\theta} &= f_{xx}x_\theta^2 + (f_{xy}+f_{yx})\,x_\theta y_\theta + f_{yy}y_\theta^2 + f_x x_{\theta\theta} + f_y y_{\theta\theta}.
\end{aligned}$$

$x_r=\cos\theta$, $y_r=\sin\theta$, $x_{rr}=0$, $y_{rr}=0$, $x_\theta=-r\sin\theta$, $y_\theta=r\cos\theta$, $x_{\theta\theta}=-r\cos\theta$, $y_{\theta\theta}=-r\sin\theta$ であるから、

$$\begin{aligned}
g_{rr} &= f_{xx}\cos^2\theta + (f_{xy}+f_{yx})\cos\theta\sin\theta + f_{yy}\sin^2\theta,\\
\frac{1}{r}g_r &= \frac{f_x\cos\theta}{r}+\frac{f_y\sin\theta}{r},\\
\frac{1}{r^2}f_{\theta\theta} &= \frac{1}{r^2}\left(f_{xx}r^2\cos^2\theta-(f_{xy}+f_{yx})\,r^2\sin\theta\cos\theta+f_{yy}r^2\cos^2\theta-f_x r\cos\theta-f_y r\sin\theta\right)\\
&= f_{xx}\cos^2\theta-(f_{xy}+f_{yx})\cos\theta\sin\theta+f_{yy}\sin^2\theta-\frac{f_x\cos\theta}{r}-\frac{f_y\sin\theta}{r}.
\end{aligned}$$

ゆえに

$$g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta}=f_{xx}+f_{yy}.\blacksquare$$

ちなみに 3 変数バージョンは

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}+\frac{\partial^2 f}{\partial z^2}=\frac{\partial^2 g}{\partial r^2}+\frac{2}{r}\frac{\partial g}{\partial r}+\frac{1}{r^2}\left(\frac{\partial^2 g}{\partial \theta^2}+\frac{1}{\tan\theta}\frac{\partial g}{\partial \theta}+\frac{1}{\sin^2\theta}\frac{\partial^2 g}{\partial \phi^2}\right)$$

となり、工夫なしに馬鹿正直に計算すると、1 時間半以上かかる (桂田先生調査)。

余談 2.8.3 (ギリシャ文字に親しもう) 上の問のようにギリシャ文字を用いると苦戦する学生が一定の数いるので、あるとき、ギリシャ文字を避けて授業したことがある。そうすると確かにその部分はスムーズに進行するのだが、ともすると学生がギリシャ文字に慣れる機会を奪ってしまいかねないことに気づいた。今ではネットで「ギリシャ文字 書き方」のように検索すると、ギリシャ文字の筆順まで調べられるご時世なので、ギリシャ文字を避けずに授業をする気になった。■

2.8.4 逆関数の微分法

1 変数関数の場合の

$$\frac{dx}{dy}=\frac{1}{\dfrac{dy}{dx}}$$

に相当する定理がある。

定理 2.8.6 (逆関数の微分法) U と V は $\mathbf{R}^n$ の開集合で、$\varphi\colon U\to V$ は全単射、$a\in U$, $b=\varphi(a)$, φ は a で、φ^{-1} は b で全微分可能であるならば、

$$\left(\varphi^{-1}\right)'(b)=\left(\varphi'(a)\right)^{-1}.$$

(左辺の $^{-1}$ は逆関数を表し、右辺の $^{-1}$ は逆行列を表す。)

注意: 上の定理では微分可能な逆関数の存在を仮定しているが、実は $\det\varphi'(a)\neq 0$ という条件が成り立てば、局所的に微分可能な逆関数の存在が導かれるという、**逆関数定理**がある (非常に重要!)。それはこの講義科目の終盤に解説する予定である。

証明 逆関数の定義により、

$$\varphi^{-1}\left(\varphi(x)\right)=x\quad(x\in U).$$

φ は a で、φ^{-1} は $b=f(a)$ で全微分可能であるから、合成関数の微分法より、

$$\left(\varphi^{-1}\right)'(b)\varphi'(a)=I\quad(I\ \text{は}\ n\ \text{次の単位行列})$$

が成り立つ。ゆえに $\left(\varphi^{-1}\right)'(b)=\varphi'(a)^{-1}$. ■

問 $\varphi(x)=x$ $(x\in\mathbf{R}^n)$ とするとき、$\varphi'(x)=I$ であることを示せ。
(解法 1) 一般に「$f(x)=Ax+b$ ならば $f'(x)=A$」であるから、$\varphi(x)=x=Ix+0$ より、$\varphi'(x)=I$.
(解法 2) $\varphi=\begin{pmatrix}\varphi_1\\ \vdots\\ \varphi_n\end{pmatrix}$ とおくと、$\varphi_i(x)=x_i$ $(i=1,\dots,n)$.

$$\frac{\partial\varphi_i}{\partial x_j}(x)=\frac{\partial}{\partial x_j}x_i=\begin{cases}1 & (i=j)\\ 0 & (i\neq j)\end{cases}$$

であるから、$\varphi'(x)=(\delta_{ij})=I$. ■

次の例で、逆関数の微分法を使って $\left(\varphi^{-1}\right)'$ を計算しているが、微分可能性を証明するには、本当は逆関数の定理のお世話になる必要があるだろう。

例 2.8.7 (極座標変換の逆変換のヤコビ行列 (とても重要)) $\varphi\colon(0,\infty)\times(0,2\pi)\ni\begin{pmatrix}r\\ \theta\end{pmatrix}\mapsto\begin{pmatrix}x\\ y\end{pmatrix}=\begin{pmatrix}r\cos\theta\\ r\sin\theta\end{pmatrix}\in\mathbf{R}^2\setminus\{(x,0);x\in[0,\infty)\}$ は明らかに C^1 級であるから (実は C^∞ 級である)、全微分可能である。

$$\varphi'(r,\theta)=\begin{pmatrix}x_r & x_\theta\\ y_r & y_\theta\end{pmatrix}=\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}.$$

逆関数の微分法によって、$\begin{pmatrix} r \\ \theta \end{pmatrix} = \varphi^{-1}(x,y)$ の全微分係数は、

$$\begin{aligned}
\begin{pmatrix} r_x & r_y \\ \theta_x & \theta_y \end{pmatrix} = \left(\varphi^{-1}\right)'(x,y) = \varphi'(r,\theta)^{-1} &= \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}^{-1} \\
&= \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}^{-1} \\
&= \frac{1}{\cos\theta\cdot r\cos\theta - (-r\sin\theta)\sin\theta}\begin{pmatrix} r\cos\theta & +r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{pmatrix}.
\end{aligned}$$

ゆえに[11]

$$r_x = \cos\theta, \quad r_y = \sin\theta, \quad \theta_x = -\frac{\sin\theta}{r}, \quad \theta_y = \frac{\cos\theta}{r}. \blacksquare$$

(**以下はやや脱線** — 無視してよろしい) この結果を逆関数の微分法を使わずに求めてみよう。$r = \sqrt{x^2+y^2}$ であるから、r_x と r_y は比較的簡単に得られる。

$$\begin{aligned}
r_x &= \frac{\partial}{\partial x}\left(x^2+y^2\right)^{1/2} = \frac{1}{2}\left(x^2+y^2\right)^{-1/2}\cdot\frac{\partial}{\partial x}(x^2+y^2) = \frac{x}{\sqrt{x^2+y^2}}, \\
r_y &= \frac{\partial}{\partial y}\left(x^2+y^2\right)^{1/2} = \frac{1}{2}\left(x^2+y^2\right)^{-1/2}\cdot\frac{\partial}{\partial y}(x^2+y^2) = \frac{y}{\sqrt{x^2+y^2}}.
\end{aligned}$$

これらがそれぞれ $\cos\theta, \sin\theta$ に等しいことは容易に分かる[12]。θ の導関数の方は少し難しい。割と多くのテキストに

$$\theta = \tan^{-1}\frac{y}{x}$$

と書かれているが[13]、これは $\pmod{\pi}$ でしか正しくない式である。本当は、$\tan^{-1}\dfrac{y}{x}$ が主値 ($(-\pi/2,\pi/2)$ 内の値) を意味するとして、

$$\theta = \begin{cases}
\tan^{-1}\dfrac{y}{x} & ((x,y)\ \text{が第 1 象限内あるいは}\ x\ \text{軸の正の部分上の点}) \\
\tan^{-1}\dfrac{y}{x} + \pi & (x<0,\ \text{言い換えると}\ (x,y)\ \text{が第 2,3 象限内あるいは}\ x\ \text{軸の負の部分上の点}) \\
\tan^{-1}\dfrac{y}{x} + 2\pi & ((x,y)\ \text{が第 4 象限内の点}) \\
\dfrac{\pi}{2} & (x=0\ \text{かつ}\ y>0) \\
\dfrac{3\pi}{2} & (x=0\ \text{かつ}\ y<0)
\end{cases}$$

---

[11] こういう計算をするとき、ヤコビ行列の成分の並べ方を間違えると、とんでもない結果になってしまうことに注意しよう。

[12] 例えば $\dfrac{x}{\sqrt{x^2+y^2}} = \dfrac{r\cos\theta}{\sqrt{(r\cos\theta)^2+(r\sin\theta)^2}} = \dfrac{r\cos\theta}{r} = \cos\theta.$

[13] 本当に困ったことである。C 言語のプログラムで、デカルト座標を極座標に直すには、`r=sqrt(x*x+y*y); theta=atan2(y,x);` のようにする。`theta=atan(y/x);` ではマズイ — というのは常識的なことなのだが、数学書の方が旧態依然のままなのは情けない。

となるはずである。

$$\frac{\partial}{\partial x}\tan^{-1}\frac{y}{x}=\frac{1}{1+(y/x)^2}\frac{\partial}{\partial x}\left(\frac{y}{x}\right)=\frac{1}{1+(y/x)^2}\cdot\left(-\frac{y}{x^2}\right)=-\frac{y}{x^2+y^2},$$
$$\frac{\partial}{\partial y}\tan^{-1}\frac{y}{x}=\frac{1}{1+(y/x)^2}\frac{\partial}{\partial y}\left(\frac{y}{x}\right)=\frac{1}{1+(y/x)^2}\cdot\frac{1}{x}=\frac{x}{x^2+y^2}$$

であることから、

$$\theta_x=-\frac{y}{x^2+y^2},\quad \theta_y=\frac{x}{x^2+y^2}$$

が導けるが、少々面倒である (例えば y 軸上でどうすれば良いか分かりますか？[14])。これがそれぞれ $-\dfrac{\sin\theta}{r}$, $\dfrac{\cos\theta}{r}$ に等しいことは容易に確められる。 ■

例 2.8.8 (Laplacian の極座標表示 (再び)) C^2 級の関数 $f\colon (x,y)\mapsto f(x,y)$ があるとき、

$$x=r\cos\theta,\quad y=r\sin\theta,\quad g(r,\theta):=f(x,y)$$

で g を定める。すなわち、

$$g(r,\theta):=f(r\cos\theta,r\sin\theta).$$

このとき

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}=\frac{\partial^2 g}{\partial r^2}+\frac{1}{r}\frac{\partial g}{\partial r}+\frac{1}{r^2}\frac{\partial^2 g}{\partial\theta^2}$$

が成り立つ。上では、右辺を変形していって左辺になることで示した。

応用上は、 $\triangle f=f_{xx}+f_{yy}$ が先にあって、これを g とその偏導関数で表したいので、以下のように左辺を変形していって右辺になることを示す方が望ましい。まず x, y についての 1 階偏導関数

$$f_x=g_r r_x+g_\theta\theta_x=g_r\cos\theta-g_\theta\frac{\sin\theta}{r},$$
$$f_y=g_r r_y+g_\theta\theta_y=g_r\sin\theta+g_\theta\frac{\cos\theta}{r}$$

から、

$$\frac{\partial}{\partial x}=\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta},\quad \frac{\partial}{\partial y}=\sin\theta\frac{\partial}{\partial r}+\frac{\cos\theta}{r}\frac{\partial}{\partial\theta}.$$

以下は (面倒ではあるが、機械的計算で)

$$\begin{aligned}
f_{xx}&=\frac{\partial}{\partial x}f_x=\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(g_r\cos\theta-g_\theta\frac{\sin\theta}{r}\right)\\
&=\cos\theta\frac{\partial}{\partial r}\left(g_r\cos\theta-g_\theta\frac{\sin\theta}{r}\right)-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\left(g_r\cos\theta-g_\theta\frac{\sin\theta}{r}\right)\\
&=\cos\theta\left(g_{rr}\cos\theta-g_{\theta r}\frac{\sin\theta}{r}-g_\theta\frac{-\sin\theta}{r^2}\right)-\frac{\sin\theta}{r}\left(g_{r\theta}\cos\theta+f_r(-\sin\theta)-g_{\theta\theta}\frac{\sin\theta}{r}-g_\theta\frac{\cos\theta}{r}\right)\\
&=g_{rr}\cos^2\theta-\frac{2g_{r\theta}\cos\theta\sin\theta}{r}+\frac{g_{\theta\theta}\sin^2\theta}{r^2}+\frac{g_r\sin^2\theta}{r}+\frac{2g_\theta\cos\theta\sin\theta}{r^2}.
\end{aligned}$$

---

[14] 例えば、$x=0$, $y>0$ の範囲では、偏微分の定義から $\theta_y=0=\dfrac{x}{x^2+y^2}$ となることは容易に分かるが、$\theta_x=-\dfrac{y}{x^2+y^2}$ となることは、偏導関数の右側からの極限と、左側からの極限がともに $-\dfrac{y}{x^2+y^2}$ であることを確め、「$f\colon (a,b)\to\mathbf{R}$ が連続で、$c\in(a,b)$ 以外では微分可能で、$\lim\limits_{\substack{x\neq c\\ x\to c}}f'(x)=D$ であれば、f は c で微分可能で、$f'(c)=D$.」という定理を用いる。

同様に

$$f_{yy}=g_{rr}\sin^2\theta+\frac{2g_{r\theta}\cos\theta\sin\theta}{r}+\frac{g_{\theta\theta}\cos^2\theta}{r^2}+\frac{g_r\cos^2\theta}{r}-\frac{2g_\theta\cos\theta\sin\theta}{r^2}.$$

ゆえに

$$f_{xx}+f_{yy}=g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta}.\blacksquare$$

2.8.5 演習問題

問 2.8.2 $z=g(y)$, $y=f(x)$ がともに 2 回微分可能であれば、

$$\frac{d^2z}{dx^2}=\frac{d^2z}{dy^2}\left(\frac{dy}{dx}\right)^2+\frac{dz}{dy}\frac{d^2y}{dx^2}$$

であることを示せ。

問 2.8.3 $f,g\colon\mathbf{R}\to\mathbf{R}$ が C^2 級ならば、定数 $c>0$ に対して

$$u(t,x):=f(x-ct)+g(x+ct)$$

で定義される関数 u は、次の **1 次元波動方程式**を満たすことを示せ。

$$\frac{\partial^2u}{\partial t^2}=c^2\frac{\partial^2u}{\partial x^2}.$$

問 2.8.4 c を正定数、$F\colon\mathbf{R}\to\mathbf{R}$ を C^2 級の関数とするとき、u を

$$u(x,y,z,t)=\frac{F(r-ct)}{r},\quad r=\sqrt{x^2+y^2+z^2}$$

で定義すると、u は定義される範囲で次の **3 次元波動方程式**を満たすことを示せ。

$$u_{tt}=c^2\left(u_{xx}+u_{yy}+u_{zz}\right).$$

問 2.8.5 関数

$$u(t,x)=\frac{1}{(4\pi t)^{n/2}}\exp\left(-\frac{\|x\|^2}{4t}\right)\quad(t>0,\ x\in\mathbf{R}^n)$$

は次の n 次元**熱伝導方程式**を満たすことを示せ。

$$\frac{\partial u}{\partial t}=\triangle u.$$

ただし

$$\triangle:=\sum_{j=1}^n\frac{\partial^2}{\partial x_j^2}=\frac{\partial^2}{\partial x_1^2}+\cdots\frac{\partial^2}{\partial x_n^2}.$$

問 2.8.6 $\nu\in\mathbf{R}^n$ は $\|\nu\|=1$ を満たし、$c>0$, $U\colon\mathbf{R}\to\mathbf{R}$ は C^2 級の関数とするとき、

$$u(x,t):=U(\nu\cdot x-ct)\quad((x,t)\in\mathbf{R}^n\times\mathbf{R})$$

で定義される $u\colon\mathbf{R}^n\times\mathbf{R}\to\mathbf{R}$ は、

$$\frac{1}{c^2}\frac{\partial^2u}{\partial t^2}(x,t)=\triangle u(x,t)$$

を満たすことを示せ。ただし、

$$\nu\cdot x:=\sum_{j=1}^n\nu_jx_j,\quad\triangle:=\sum_{j=1}^n\frac{\partial^2}{\partial x_j^2}.$$

問 2.8.7 C^2 級の関数 $u\colon \mathbf{R}^2 \ni (x,t) \mapsto u(x,t) \in \mathbf{R}$ と正定数 c があるとき、

$$\xi = x - ct, \quad \eta = x + ct, \quad v(\xi,\eta) = u(x,t), \quad \text{すなわち} \quad v(\xi,\eta) := u\left(\frac{\xi+\eta}{2}, \frac{\eta-\xi}{2c}\right)$$

とおく。

(1) u_x, u_{xx}, u_t, u_{tt} を v の偏導関数で表せ。 (2) $v_\eta, v_{\eta\xi}$ を u の偏導関数で表し，$\dfrac{1}{c^2}u_{tt} - u_{xx} = -4v_{\eta\xi}$ を示せ。

問 2.8.8 $x = r\cos\theta$, $y = r\sin\theta$ ($r \geq 0$, $\theta \in [0, 2\pi)$) とするとき、以下の問に答えよ。(1) 写像 $\varphi\colon \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$ のヤコビ行列を求めよ。(2) $f = f(x,y)$ を C^2 級の関数として、$g := f \circ \varphi$, すなわち $g(r,\theta) = f(x,y) = f(r\cos\theta, r\sin\theta)$ で関数 g を定める。このとき g_r, g_{rr}, $g_{\theta\theta}$ を f の偏微分係数 (f_x, f_y, f_{xx}, f_{xy}, f_{yy} 等) を用いて表し、$f_{xx} + f_{yy} = g_{rr} + \dfrac{1}{r}g_r + \dfrac{1}{r^2}g_{\theta\theta}$ が成り立つことを示せ。

問 2.8.9 $x = r\cos\theta$, $y = r\sin\theta$ ($r \geq 0$, $\theta \in [0, 2\pi)$) とするとき、以下の問に答えよ。(1) 写像 $\varphi\colon \begin{pmatrix} r \\ \theta \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix}$ のヤコビ行列を求めよ。(2) $r \neq 0$ であれば、φ は逆関数定理の仮定を満たすことを示せ。(3) 逆関数 $\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} r \\ \theta \end{pmatrix}$ (定義域は適当に定める) の偏微分係数 r_x, r_y, θ_x, θ_y を求めよ。(4) $f = f(x,y)$ を C^2 級の関数として、$g := f \circ \varphi$, すなわち $g(r,\theta) = f(x,y)$ で関数 g を定める。このとき f_x, f_y, f_{xx}, f_{yy} を g の偏微分係数 (g_r, g_θ, g_{rr}, $g_{r\theta}$, $g_{\theta\theta}$ 等) を用いて表し、$f_{xx} + f_{yy} = g_{rr} + \dfrac{1}{r}g_r + \dfrac{1}{r^2}g_{\theta\theta}$ が成り立つことを示せ。

2.9 平均値の定理、Taylor の定理

合成関数の微分法を用いると、多変数関数版平均値の定理・Taylor の定理が得られる。これらは例えば極値問題の解析に利用できる。

問題意識 Ω が $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$, $a \in \Omega$ とするとき、($\|h\|$ が十分小さい h に対して $a + h \in \Omega$ となるわけだが) $f(a+h) - f(a)$ はどうなるか？以前 $f(a+h) - f(a) \fallingdotseq f'(a)h = \nabla f(a) \cdot h$ ということを書いたが、$\fallingdotseq$ を用いた式は数学的な主張とは見なせない。そこをちゃんとしたい。

注意 f の値域の次元 (これまで m と書いてきたもの) は、1 とする。その理由は後述する。

アイディア

解くためには、次のアイディア一発で十分である。

$$F(t) := f(a + th) \quad (t \in [0,1])$$

とおくと、$F(0) = f(a)$, $F(1) = f(a+h)$ であるから、

$$f(a+h) - f(a) = F(1) - F(0).$$

この F は 1 変数関数であるから、1 変数関数の平均値定理、Taylor の定理が使える。

F の導関数 F' はどうなるか？$\varphi(t) := a + th$ とおくと、$F(t) = f(\varphi(t))$ であるから、F は f と φ の合成関数である: $F = f \circ \varphi$. また、$\varphi'(t) = h$. 実際、

$$\varphi(t) = a + th = \begin{pmatrix} a_1 + th_1 \\ a_2 + th_2 \\ \vdots \\ a_n + th_n \end{pmatrix}$$

であるから、

$$\varphi'(t) = \begin{pmatrix} (a_1 + th_1)' \\ (a_2 + th_2)' \\ \vdots \\ (a_n + th_n)' \end{pmatrix} = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = h.$$

ゆえに合成関数の微分法から、

$$F'(t) = \frac{d}{dt} f(a + th) = f'(\varphi(t))\varphi'(t) = f'(a + th)h = \nabla f(a + th) \cdot h.$$

特に

$$F'(0) = \left. \frac{d}{dt} f(a + th) \right|_{t=0} = f'(a)h = \nabla f(a) \cdot h.$$

この量を、a における、f の h 方向の**方向微分係数**と呼び、$\dfrac{\partial f}{\partial h}(a)$ で表す (h が変数の名前でなく、ベクトルであることに注意する)。

記号の約束: 方向微分係数

$$\frac{\partial f}{\partial h}(a) := \left. \frac{d}{dt} f(a + th) \right|_{t=0} = f'(a)h = \nabla f(a) \cdot h.$$

2.9.1 平均値の定理

定理 2.9.1 (多変数関数の平均値の定理) Ω は $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$ は全微分可能、$a, b \in \Omega$, $a \neq b$, $[a, b] \subset \Omega$ とするとき、

$$\exists c \in (a, b) \quad \text{s.t.} \quad f(b) - f(a) = f'(c)(b - a).$$

ただし

$$[a, b] := \{(1 - t)a + tb; t \in [0, 1]\},$$
$$(a, b) := \{(1 - t)a + tb; t \in (0, 1)\}.$$

注意 2.9.2 $n \geq 2$ のとき、1 変数の場合に良く登場する式

$$\frac{f(b) - f(a)}{b - a} = f'(c)$$

はナンセンスである (ベクトル $b - a$ で割れない)。■

証明 $h := b - a$, $F(t) := f(a + th)$ $(t \in \mathbf{R})$ とおくと、

$$f(b) - f(a) = F(1) - F(0), \quad F'(t) = f'(a + th)h.$$

1 変数関数の平均値の定理から、

$$\exists \theta \in (0,1) \quad \text{s.t.} \quad F(1) - F(0) = F'(\theta) \cdot 1.$$

ゆえに

$$f(b) - f(a) = f'(a + \theta h)h.$$

$c := a + \theta h$ とおくと、$c \in (a,b)$ で、$f(b) - f(a) = f'(c)(b - a)$. ■

平均値の定理をベクトル値関数に拡張することは出来ない。実数値関数の場合の定理の証明を振り返り、さかのぼると、実数値関数の最大値の存在まで行き着く。この証明を一般化するのが難しいことは理解できるであろう。ここでは一つの反例を示す。

例 2.9.3 (ベクトル値関数では平均値の定理が成り立たない) $f\colon \mathbf{R}^2 \ni x \mapsto \begin{pmatrix} \cos x \\ \sin x \end{pmatrix} \in \mathbf{R}$ は明らかに C^∞ 級であり、全微分可能である。$a = 0$, $b = 2\pi$ とするとき、

$$f(b) - f(a) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

一方

$$f'(x) = \begin{pmatrix} (\cos x)' \\ (\sin x)' \end{pmatrix} = \begin{pmatrix} -\sin x \\ \cos x \end{pmatrix}$$

で、このノルムは $\sqrt{\sin^2 x + \cos^2 x} = 1$ であるから、$f'(x) \neq 0$. 特に $\forall c \in (a,b)$ に対して、$f'(c)(b - a) \neq 0$. ゆえに $f(b) - f(a)$ と $f'(c)(b - a)$ は等しくなり得ない。 ■

2.9.2 Taylor の定理の準備

2 変数関数の高階偏導関数

問 2.9.1 C^∞ 級の 2 変数関数 $f(x,y)$ $((x,y) \in \mathbf{R}^2)$ と、$(a,b) \in \mathbf{R}^2$, $(h,k) \in \mathbf{R}^2$ があるとき、

$$F(t) := f\left(a + th, b + tk\right) \quad (t \in \mathbf{R})$$

とおくとき、次の (1), (2) に答えよ[15]。 ——— 合成関数の微分法で、Taylor の定理の準備

(1) $F'(t)$, $F''(t)$, $\ldots$ を (いくつか) 計算せよ。

(2) $F^{(m)}(t)$ $(m \in \mathbf{N})$ の公式を推測し、数学的帰納法で証明せよ。

前半は計算であるが、要点は次の二つ。

(a) chain rule

[15] (授業でやるけれど) n 変数関数 $f(x_1, \ldots, x_n)$ について、$F(t) := f(a_1 + th_1, a_2 + th_2, \cdots, a_n + th_n)$ とおき、上と同様のことを行なえ。

(b) f が C^2 級ならば $f_{xy}=f_{yx}$, f が C^3 級ならば

$$f_{xxy}=f_{xyx}=f_{yxx},\quad f_{yyx}=f_{yxy}=f_{xyy}.$$

f が C^∞ 級ならば、f の偏導関数は、x で何回、y で何回偏微分したかで決まり、$\dfrac{\partial^{p+q}f}{\partial x^p\partial y^q}$ だけで表すことが出来る。

多くの人が (1) で 3 階微分 $F'''(t)$ を計算できるくらいまで時間を取って、

$$\begin{aligned}F^{(m)}(t)&=\sum_{r=0}^{m}\binom{m}{r}\frac{\partial^m f}{\partial x^r\partial y^{n-r}}(a+th,b+tk)h^r k^{m-r}\\&=\left(h\frac{\partial}{\partial x}+k\frac{\partial}{\partial y}\right)^m f(a+th,b+tk)\qquad (\text{乱暴であるが}\cdots)\end{aligned}$$

を提示する。「$(a+b)^m$ を表す二項定理を思い出して下さい、その数学的帰納法による証明と同様にして証明できます。」

$(a_1+\cdots+a_n)^m$ の展開 — 多項定理

この節の目標は次を証明することである。

定理 2.9.4 (多項定理 (the multinomial theorem)) $n,m\in\mathbf{N}$, $n\ge 2$ とするとき

(2.8)
$$\begin{aligned}(a_1+a_2+\cdots+a_n)^m&=\sum_{1\le i_1,i_2,\cdots,i_m\le n}a_{i_1}a_{i_2}\cdots a_{i_m}\\&=\sum_{\substack{\alpha_1,\ \alpha_2,\ \cdots,\ \alpha_n\ \text{は非負整数}\\ \alpha_1+\alpha_2+\cdots+\alpha_n=m}}\frac{m!}{\alpha_1!\alpha_2!\cdots\alpha_n!}a_1^{\alpha_1}a_2^{\alpha_2}\cdots a_n^{\alpha_n}.\end{aligned}$$

以下、インデックスが非負整数であることは常に仮定するので、それを書くのは省略する。

まず (2.8) の前半部分

(2.9)
$$(a_1+a_2+\cdots+a_n)^m=\sum_{1\le i_1,i_2,\cdots,i_m\le n}a_{i_1}a_{i_2}\cdots a_{i_m}$$

を証明しよう。分配法則でとにかくバラすのが方針である。

分配法則は良く知っているであろう。

$$(a_1+\cdots+a_n)A=a_1A+\cdots+a_nA,\quad \text{つまり}\quad \left(\sum_{i=1}^n a_i\right)A=\sum_{i=1}^n(a_iA).$$

($\sum$ の外の数を中に入れる、ということ。)

これを 2 回使うと、

$$(a_1+\cdots+a_n)^2=\left(\sum_{i=1}^n a_i\right)\left(\sum_{j=1}^n a_j\right)=\sum_{i=1}^n\left(a_i\sum_{j=1}^n a_j\right)=\sum_{i=1}^n\sum_{j=1}^n(a_ia_j)$$

同様にして、m 個の積の場合、$m-1$ 回分配法則を用いて

$$(a_1+\cdots+a_n)^m=\sum_{i=1}^n\sum_{i_2=1}^n\cdots\sum_{i_m=1}^n(a_{i_1}a_{i_2}\cdots a_{i_m}).$$

この右辺を

$$\sum_{1\le i_1,i_2,\cdots,i_m\le n}a_{i_1}a_{i_2}\cdots a_{i_m}$$

と書く、というわけである。

(2.8) の後半部分の証明の参考のため、二項定理を復習しよう。

二項定理 (the binomial theorem)　授業では時間がないので、「二項定理は証明まで込めて自分で復習しておいて下さい」というのだろう。

定理 2.9.5 (二項定理 (the binomial theorem))

$$(a+b)^m = \sum_{k=0}^{m} \binom{m}{k} a^{m-k} b^k = \sum_{k=0}^{m} \binom{m}{k} a^k b^{m-k}.$$

高校数学では帰納法による証明が標準的と思われる[16]。

二項係数に関する次の公式は、Pascal の三角形を作るときにも使われるので、良く知っているはずである (知らずに Pascal の三角形を書いていたらダメよ)。

補題 2.9.6

$$\binom{m}{k-1} + \binom{m}{k} = \binom{m+1}{k}.$$

証明

$$\begin{aligned}
\binom{m}{k-1} + \binom{m}{k} &= \frac{m!}{(k-1)!(m-(k-1))!} + \frac{m!}{k!(m-k)!} \\
&= \frac{m!}{(k-1)!(m-k+1))!} + \frac{m!}{k!(m-k)!} \\
&= \frac{k}{k(k-1)!} \cdot \frac{m!}{(m-k+1))!} + \frac{m!}{k!} \cdot \frac{m-k+1}{(m-k+1)(m-k)!} \\
&= \frac{m!}{k!(m-k+1)!} (k + (m-k+1)) = \frac{m!(m+1)}{k!(m-k+1)!} \\
&= \frac{(m+1)!}{k!((m+1)-k)!} = \binom{m+1}{k}. \blacksquare
\end{aligned}$$

定理 2.9.5 の証明　帰納法による。$m=1$ のとき明らかに成り立つ。$m=k$ のとき成り立つと仮定すると、

$$\begin{aligned}
(a+b)^{k+1} = (a+b)(a+b)^k &= (a+b) \sum_{r=0}^{n} \binom{k}{r} a^{k-r} b^r \\
&= \sum_{r=0}^{k} \binom{k}{r} a^{k+1-r} b^r + \sum_{r=0}^{k} \binom{k}{r} a^{k-r} b^{r+1}.
\end{aligned}$$

右辺第 1 項は

$$\sum_{r=0}^{k} \binom{k}{r} a^{k+1-r} b^r = a^{k+1} + \sum_{r=1}^{k} \binom{k}{r} a^{k+1-r} b^r.$$

右辺第 2 項は、途中で $r+1$ を r' と置き換えて

$$\begin{aligned}
\sum_{r=0}^{k} \binom{k}{r} a^{k-r} b^{r+1} = \sum_{r=0}^{k} \binom{k}{r} a^{(k+1)-(r+1)} b^{r+1} &= \sum_{r'=1}^{k+1} \binom{k}{r'-1} a^{k+1-r'} b^{r'} \\
&= \sum_{r=1}^{k} \binom{k}{r-1} a^{k+1-r} b^r + b^{k+1}.
\end{aligned}$$

[16] $(a+b)^m = (a+b)(a+b)\cdots(a+b)$ の右辺を展開して $a^k b^{m-k}$ が出て来るには、m 個の因子のうち、a を k 個選ぶということで、全部で $\binom{m}{k}$ 通りある、と数え上げることも出来る。

ゆえに

$$\begin{aligned}
(a+b)^{k+1} &= a^{k+1} + \sum_{r=1}^{k} \left(\binom{k}{r} + \binom{k}{r-1} \right) a^{k+1-r} b^r + b^{k+1} \\
&= a^{k+1} + \sum_{r=1}^{k} \binom{k+1}{r} a^{k+1-r} b^r + b^{k+1} \\
&= \sum_{r=0}^{k+1} \binom{k+1}{r} a^{k+1-r} b^r + b^{k+1}.
\end{aligned}$$

これは $m = k+1$ のときも成り立つことを示している。帰納法により、任意の自然数 m について成り立つ。■

この帰納法による証明は、問 10 の (2) の証明部分と本質的に同じである。

二項定理が

$$(a+b)^n = \sum_{r=0}^{n} \binom{n}{r} a^r b^{n-r} = \sum_{\alpha+\beta=n} \frac{n!}{\alpha!\beta!} a^\alpha b^\beta$$

と書けることを注意しておく。

定理 2.9.4 の後半部分の証明 n に関する帰納法による。$n=2$ のとき (2.8) が成り立つのは明らかである。$n=k$ のとき (2.8) が成り立つと仮定する。二項定理に続いて帰納法の仮定を用いて、

$$\begin{aligned}
(a_1 + a_2 + \cdots + a_k + a_{k+1})^m &= ((a_1 + a_2 + \cdots + a_k) + a_{k+1})^m \\
&= \sum_{\alpha+\beta=m} \frac{m!}{\alpha!\beta!} (a_1 + a_2 + \cdots + a_k)^\alpha a_{k+1}^\beta \\
&= \sum_{\alpha+\beta=m} \frac{m!}{\alpha!\beta!} \sum_{\alpha_1+\alpha_2+\cdots+\alpha_k=\alpha} \frac{\alpha!}{\alpha_1!\alpha_2!\cdots\alpha_k!} a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_k^{\alpha_k} a_{k+1}^\beta \\
&= \sum_{\alpha_1+\alpha_2+\cdots+\alpha_k+\beta=m} \frac{m!}{\alpha_1!\alpha_2!\cdots\alpha_k!\beta!} a_1^{\alpha_1} a_2^{\alpha_2} \cdots a_k^{\alpha_k} a_{k+1}^\beta.
\end{aligned}$$

これは $n = k+1$ のときに (2.8) が成り立つことを示している。ゆえに (2.8) は、2 以上の任意の自然数に対して成立する。■

2.9.3 多変数版 Taylor の定理

記述を簡単にするため、

$$(d^m f)_x(h) := \sum_{1 \le i_1, i_2, \cdots, i_m \le n} \frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m}}(x) h_{i_1} h_{i_2} \cdots h_{i_m} \tag{2.10}$$

とおく[17]。これを、f の x における m **次微分**と呼ぶ。これは h に関する m 次同次多項式 (m **次形式**) である。

(少し先走っておくと) $m = 1, 2$ のときが重要である。$m=1$ のとき、$(d^1 f)_a(h) = f'(a)h$. $m=2$ のとき、$(d^2 f)_a(h)$ は後で定義する f の a における Hesse 行列 $H(a) = \left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) \right)$ を係数とする 2 次形式 $(H(a)h, h) = \displaystyle\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j$ である。

[17] このあたりは標準的な記号がない。ここで紹介した記号はいくつかの教科書に載っているものではあるが、誰でも分かるとは限らない。この講義だけの記号と思っておいた方が良い。

これは桂田が発明した記号というわけではなくて、某テキストにあった記号だが、あまりメジャーな記号ではない (誰でも分かる記号ではない)。

補題 2.9.7 Ω が $\mathbf{R}^n$ の開集合, $k \in \mathbf{N}$, $f\colon \Omega \to \mathbf{R}$ は C^k 級, $a \in \Omega$, $h \in \mathbf{R}^n$, $[a, a+h] \subset \Omega$ とするとき、

$$F(t) := f\,(a+th) \quad (t \in [0,1])$$

とおくと、$F\colon [0,1] \to \mathbf{R}$ は C^k 級で、$\forall m \in \{1,\dots,k\}$ に対して、

$$F^{(m)}(t) = (d^m f)_{a+th}(h).$$

証明 m に関する帰納法を用いる。$m=1$ に対しては、chain rule から

$$\begin{aligned} F^{(1)}(t) = F'(t) &= \frac{d}{dt} f(a+th) = f'(a+th)\frac{d}{dt}(a+th) = f'(a+th)h \\ &= \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+th)h_i. \end{aligned}$$

これは $\displaystyle\sum_{1 \le i_1 \le n} \frac{\partial f}{\partial x_{i_1}}(a+th)h_{i_1}$ と書き直せる。ゆえに $m=1$ のとき成立する。

m のとき成立する、すなわち

$$F^{(m)}(t) = \sum_{1 \le i_1, i_2, \dots, i_m \le n} \frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m}}(a+th)h_{i_1}h_{i_2}\cdots h_{i_m}$$

と仮定すると、

$$\begin{aligned} F^{(m+1)}(t) = \frac{d}{dt}F^{(m)}(t) &= \sum_{1 \le i_1, i_2, \dots, i_m \le n} \frac{d}{dt} \frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m}}(a+th)h_{i_1}h_{i_2}\cdots h_{i_m} \\ &= \sum_{1 \le i_1, i_2, \dots, i_m \le n} \left(\sum_{i=1}^{n} \frac{\partial^{m+1} f}{\partial x_i \partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m}}(a+th)h_i \right) h_{i_1}h_{i_2}\cdots h_{i_m} \\ &= \sum_{1 \le i_1, i_2, \dots, i_m, i_{m+1} \le n} \frac{\partial^{m+1} f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m} \partial x_{i_{m+1}}}(a+th)h_{i_1}h_{i_2}\cdots h_{i_m}h_{i_{m+1}}. \end{aligned}$$

これは $m+1$ のときも成立することを示す。ゆえにすべての $m(\le k)$ について成立する。■

定理 2.9.8 (多変数の Taylor の定理) $n, k \in \mathbf{N}$, Ω を $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$ を C^k 級の関数、$a \in \Omega$, $h \in \mathbf{R}^n$, 線分 $[a, a+h] \subset \Omega$ とするとき、次の式を満たすような $0 < \theta < 1$ が存在する:

$$f(a+h) = \sum_{m=0}^{k-1} \frac{1}{m!} \left(d^m f\right)_a (h) + \frac{1}{k!} \left(d^k f\right)_{a+\theta h} (h).$$

(念のため枠内に再度書いておく) ここで $\left(d^m f\right)_x (h)$ は、f の x における m 次微分と呼ばれる、h についての m 次形式で、次の式で定められる。

$$\left(d^m f\right)_x (h) := \sum_{1 \le i_1, i_2, \cdots, i_m \le n} \frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m}}(x)h_{i_1}h_{i_2}\cdots h_{i_m}.$$

証明 補題 2.9.7 より、$F(t) := f(a+th)$ は $[0,1]$ で C^k 級で、

$$F^{(m)}(t) = (d^m f)_{a+th}(h) \quad (0 \le m \le k).$$

1 変数関数についての Taylor の定理から、$\exists\theta\in(0,1)$ s.t.

$$F(1)=\sum_{m=0}^{k-1}\frac{F^{(m)}(0)}{m!}\cdot 1^m+\frac{1}{k!}F^{(k)}(\theta)\cdot 1^k$$

後は代入するだけで結論を得る。 ■

m **次微分の同類項の整理**　偏微分係数は偏微分の順序によらないのだから、$(d^m f)_x(h)$ の $\sum$ には同類項が含まれている。それをまとめると

$$\begin{aligned}(d^m f)_x(h) &= \sum_{1\le i_1,i_2,\cdots,i_m\le n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(x)h_{i_1}h_{i_2}\cdots h_{i_m}\\ &= \left(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}+\cdots+h_n\frac{\partial}{\partial x_n}\right)^m f(x)\\ &= \sum_{\substack{\alpha_1,\ldots,\alpha_n\ \text{は非負整数}\\ \alpha_1+\cdots+\alpha_n=m}}\frac{m!}{\alpha_1!\alpha_2!\cdots\alpha_n!}\frac{\partial^m f}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n}}(x)\,h_1^{\alpha_1}h_2^{\alpha_2}\cdots h_n^{\alpha_n}.\end{aligned}$$

となる。ただし、ここでは二項定理

$$(a+b)^m=\sum_{r=0}^{m}\binom{m}{r}a^r b^{m-r}=\sum_{r=0}^{m}\frac{m!}{r!(m-r)!}a^r b^{m-r}$$

を一般化した**多項定理**[18]

$$(a_1+a_2+\cdots+a_n)^m=\sum_{1\le i_1,i_2,\cdots,i_m\le n}a_{i_1}a_{i_2}\cdots a_{i_m}=\sum_{\substack{\alpha_1+\alpha_2+\cdots+\alpha_n=m\\ \alpha_j\ \text{は非負整数}}}\frac{m!}{\alpha_1!\alpha_2!\cdots\alpha_n!}{a_1}^{\alpha_1}{a_2}^{\alpha_2}\cdots{a_n}^{\alpha_n}$$

を用いて、多少形式的な[19]計算を行なった (証明したとは言えません)。

例えば 2 変数関数 f の 2 次微分については、既に示したように、

$$\begin{aligned}(d^2 f)_a(h) &= \sum_{i,j=1}^{2}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)h_i h_j\\ &= \frac{\partial^2 f}{\partial x_1\partial x_1}(a)h_1h_1+\frac{\partial^2 f}{\partial x_1\partial x_2}(a)h_1h_2+\frac{\partial^2 f}{\partial x_2\partial x_1}(a)h_2h_1+\frac{\partial^2 f}{\partial x_2\partial x_2}(a)h_2h_2\\ &= \frac{\partial^2 f}{\partial x_1^2}(a)h_1^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}(a)h_1h_2+\frac{\partial^2 f}{\partial x_2^2}(a)h_2^2.\end{aligned}$$

例 2.9.9 (荷見 [6] から) $f(x,y)$ は C^2 級で、$f(0,0)=1$, $f_x(0,0)=0.5$, $f_y(0,0)=0.1$ であり、さらに原点と点 $P=(0.1,0.2)$ を結ぶ線分上で

$$|f_{xx}(x,y)|\le 0.02,\quad |f_{xy}(x,y)|\le 0.05,\quad |f_{yy}(x,y)|\le 0.05$$

が成り立つとき、$f(P)$ を評価せよ (1 次近似で値を求め、Taylor の定理で誤差を評価せよ)。
(解答) $\boldsymbol{a}$, $\boldsymbol{a}+\boldsymbol{h}$ (ただし $\boldsymbol{a}=(a,b)$, $\boldsymbol{h}=(h,k)$) を端点とする線分が f の定義域に含まれるならば、Taylor の定理から、$\exists\theta\in(0,1)$ s.t.

$$\begin{aligned}f(\boldsymbol{a}+\boldsymbol{h}) &= f(\boldsymbol{a})+f_x(\boldsymbol{a})h+f_y(\boldsymbol{a})k\\ &\quad+\frac{1}{2}\left(f_{xx}(\boldsymbol{a}+\theta\boldsymbol{h})h^2+2f_{xy}(\boldsymbol{a}+\theta\boldsymbol{h})hk+f_{xx}(\boldsymbol{a}+\theta\boldsymbol{h})k^2\right).\end{aligned}$$

[18] 二項定理を認めれば、後は n に関する帰納法で簡単に証明できる。

[19] これを正当化することは可能である。$\dfrac{\partial}{\partial x_i}$ は数ではないが、多項定理の証明に使うような数の性質は満足していることを示すのは難しくはない。

ゆえに

$$
\begin{aligned}
&|f(\boldsymbol{a}+\boldsymbol{h})-(f(\boldsymbol{a})+f_x(\boldsymbol{a})h+f_y(\boldsymbol{a})k)|\\
&\quad\le \frac{1}{2}\left(\max_{t\in[0,1]}|f_{xx}(\boldsymbol{a}+\theta\boldsymbol{h})|\,h^2+2\max_{t\in[0,1]}|f_{xy}(\boldsymbol{a}+\theta\boldsymbol{h})|\,hk+\max_{t\in[0,1]}|f_{yy}(\boldsymbol{a}+\theta\boldsymbol{h})|\,k^2\right).
\end{aligned}
$$

$\boldsymbol{a}=(0,0)$, $\boldsymbol{h}=(0.1,0.2)$ として用いると、

$$f(\boldsymbol{a})+f_x(\boldsymbol{a})h+f_y(\boldsymbol{a})k=f(0,0)+f_x(0,0)0.2+f_y(0,0)0.2=1+0.5\times 0.1+0.1\times 0.2=1.07,$$

$$
\begin{aligned}
&\frac{1}{2}\left(\max_{t\in[0,1]}|f_{xx}(\boldsymbol{a}+t\boldsymbol{h})|\,h^2+2\max_{t\in[0,1]}|f_{xy}(\boldsymbol{a}+t\boldsymbol{h})|\,hk+\max_{t\in[0,1]}|f_{yy}(\boldsymbol{a}+t\boldsymbol{h})|\,k^2\right)\\
&\qquad\le \frac{1}{2}(0.02\times 0.1^2+2\times 0.05\times 0.1\times 0.2+0.05\times 0.2^2)=0.0021.
\end{aligned}
$$

ゆえに

$$|f(0.1,0.2)-1.07|\le 0.0021.$$

これから $1.07-0.0021\le f(0.1,0.2)\le 1.07+0.0021$ であるから、

$$1.0679\le f(0.1,0.2)\le 1.0721.\blacksquare$$

2.10 極値問題

2.10.1 まずは問題から

K を $\mathbf{R}^2$ 内の 3 点 $(0,0)$, $(1,0)$, $(0,1)$ を頂点とする三角形とする: $K=\{(x,y)\in\mathbf{R}^2\ x\ge 0,\ y\ge 0,\ x+y\le 1\}$. $f\colon K\to\mathbf{R}$ を

$$f(x,y)=3x^2+2y^2+2xy-2x-2y+1$$

で定めるとき、f の最大値、f の最小値を求めよ。

とりあえず微分してみましょう。

$$f'(x,y)=\begin{pmatrix}\dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y}\end{pmatrix}=\begin{pmatrix}6x+2y-2 & 4y+2x-2\end{pmatrix}.$$

これから

$$f'(x,y)=0\quad\Leftrightarrow\quad (x,y)=\left(\frac{1}{5},\frac{2}{5}\right).$$

さて、ここからどうしたら良いか？1 変数関数の場合のような増減表は書けない。そもそも f' の符号を調べるというのが、どう多変数関数に拡張したら良いか分からない ($f'(x,y)$ はベクトルなので)。

しかし、すぐ後で証明するように

f が定義域の内点 a で極大 (or 極小) ならば $f'(a)=0$

は多変数関数でも成立する。

また

$f'(a)=0$, a の十分近くで $f''>0\Longrightarrow f$ は a で極小

は多変数関数への拡張が出来る (今回、定理を紹介する)。

2.10.2 内点 a で極値を取れば $f'(a)=0$

まず極値を定義しないと話が始まりません。

定義 2.10.1 (極大, 極小) $A\subset\mathbf{R}^n$, $f\colon A\to\mathbf{R}$, $a\in A$ とするとき、f が a で極大とは、

$$\exists\varepsilon>0\quad\text{s.t.}\quad f(a)=\max_{x\in A\cap B(a;\varepsilon)} f(x)=\max_{\substack{x\in A\\ \|x-a\|<\varepsilon}} f(x)$$

が成り立つことをいう。また f が a で狭義の極大とは、

$$\exists\varepsilon>0\quad\text{s.t.}\quad(\forall x\in A: 0<\|x-a\|<\varepsilon)\quad f(a)>f(x)$$

が成り立つことをいう。「極小」「狭義の極小」も同様である。

注意 2.10.2 内点でしか極値を考えないという立場もある。後の条件付き極値問題とのからみで上のような定義を採用した。上の定理は、どちらの流儀でも考えても良い (開集合に属する点は、すべて内点なので)。 ■

定理 2.10.3 (微分可能な関数が内点で極値を取れば微分は 0) Ω が $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$ は全微分可能で、$a\in\Omega$, f は a で極大 (or 極小) ならば、$f'(a)=0$ (これは $\nabla f(a)=0$ とも書ける).

証明 Ω が開集合であるから、$\exists r>0$ s.t. $B(a;r)\subset\Omega$. 各 $i\in\{1,\dots,n\}$ に対して、

$$\varphi_i\colon(a_i-r,a_i+r)\ni x_i\mapsto f\left(a_1,\dots,a_{i-1},x_i,a_{i+1},\dots,a_n\right)\in\mathbf{R}$$

を考えると、これは $x_i=a_i$ で極大値を取る。ゆえに

$$0=\varphi_i'(a_i)=\frac{\partial f}{\partial x_i}(a)=0.$$

これが任意の i について成り立つから、

$$f'(a)=\left(\frac{\partial f}{\partial x_1}(a)\cdots\frac{\partial f}{\partial x_n}(a)\right)=(0\cdots0)=0.\blacksquare$$

幾何学的考察 この定理を図形的に考えてみよう。一般に関数 f のグラフ $z=f(\vec{x})$ の $\vec{x}=\vec{a}$ における接超平面は、

$$z=f'(\vec{a})(\vec{x}-\vec{a})+f(\vec{a})$$

であった。$n=2$ の場合、

$$z=f'(a,b)\begin{pmatrix}x-a\\ y-b\end{pmatrix}+f(a,b)$$

となるが、$f'(a,b)=0$ であれば、

$$z=f(a,b).$$

これは xy 平面に水平な平面である。 ■

上の定理の逆は成り立たない。すなわち $f'(a)=0$ であっても、f が a で極値を取らないということがありうる。これは 1 変数関数でもそうである。(反例: $f(x)=x^3$, $a=0$ とすると、$f'(a)=0$ であるが、f は a で極大でも極小でもない。)

2.10.3 Hesse 行列の符号による極値の判定定理

Taylor の定理を $k=2$ で用いる。f が C^2 級ならば、十分小さい $\forall h \neq 0$ に対して、

$$\begin{aligned}
f(a+h) &= f(a) + \left(d^2 f\right)_a(h) + \frac{1}{2!}\left(d^2 f\right)_{a+\theta h}(h) \\
&= f(a) + f'(a)h + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h) h_i h_j.
\end{aligned}$$

$f'(a)=0$ とすると、$f'(a)h=0$ なので

$$f(a+h) = f(a) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h) h_i h_j.$$

$\|h\|$ が小さいとき、右辺第 2 項は「大体」h の 2 次式なので (かなり乱暴な議論だけれど)、符号が一定になる場合がある。

$$\exists \varepsilon > 0 \quad (\forall h : 0 < \|h\| < \varepsilon) \quad \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h) h_i h_j > 0 \quad \Longrightarrow \quad f \text{ は } a \text{ で極小},$$

$$\exists \varepsilon > 0 \quad (\forall h : 0 < \|h\| < \varepsilon) \quad \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h) h_i h_j < 0 \quad \Longrightarrow \quad f \text{ は } a \text{ で極大}.$$

もちろん以下では厳密な議論を行なう。

定義 2.10.4 (Hesse 行列) C^2 級の関数 f に対して、

$$H(x) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)$$

とおき、これを f の x における **Hesse 行列**と呼ぶ。

Hesse 行列は実対称行列である。これを使うと、上の式は

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2}\left(H(a+\theta h)h, h\right) = f(a) + \frac{1}{2}\left(H(a+\theta h)h, h\right)$$

と書ける。ポイントは

$$\left(d^2 f\right)_x(h) = \left(H(x)h, h\right),$$

すなわち 2 次の微分が Hesse 行列を係数とする 2 次形式になることである。

次の定理がこの節のメインである。

定理 2.10.5 (Hesse 行列の符号による極値の判定) Ω が $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}$ が C^2 級, $a \in \Omega$, $f'(a)=0$, $H(a) := f$ の a における Hesse 行列 とするとき、次の (i), (ii), (iii) が成り立つ。

(i) $H(a)$ が正値 $\Longrightarrow$ f は a で狭義の極小となる。

(ii) $H(a)$ が負値 $\Longrightarrow$ f は a で狭義の極大となる。

(iii) $H(a)$ が不定符号 $\Longrightarrow$ f は a で極値を取らない。

注意 2.10.6 正値でも負値でも不定符号でもない場合がある (次項で述べる)。そういう場合は、もっと詳しく調べないと判定できない。■

2.10.4 実対称行列の正値性、負値性

定義 2.10.7 $A=(a_{ij})$ を n 次実対称行列とする。

(i) A が正値 $\overset{\text{def.}}{\Leftrightarrow}$ A の固有値がすべて正。

(ii) A が負値 $\overset{\text{def.}}{\Leftrightarrow}$ A の固有値がすべて負。

(iii) A が不定符号 $\overset{\text{def.}}{\Leftrightarrow}$ A の固有値に正のもの、負のものがある。

例 2.10.8 $A=\begin{pmatrix}2 & 0\\ 0 & 3\end{pmatrix}$ のとき、固有値は $2,3$ で、すべて正であるから、A は正値。

$A=\begin{pmatrix}-1 & 0\\ 0 & -2\end{pmatrix}$ のとき、固有値は $-1,-2$ で、すべて負であるから、A は負値。

$A=\begin{pmatrix}5 & 0\\ 0 & -2\end{pmatrix}$ のとき、固有値は $5,-2$ で、正のもの、負のもの両方あるので、A は不定符号。

$A=\begin{pmatrix}3 & 0\\ 0 & 0\end{pmatrix}$ のとき、固有値は $3,0$ で、すべて正でもないし、すべて負でもないし、不定符号でもない (負の固有値がない) ので、A は正値でも、負値でも、不定符号でもない。 ■

2.10.5 2次形式の符号=係数行列の符号

定理 2.10.9 (線形代数の復習: 実対称行列の符号と2次形式の符号) $A=(a_{ij})$ が n 次実対称行列とするとき、次の (i), (ii), (iii) が成り立つ。

(i) A が正値 $\Leftrightarrow \forall h\in\mathbf{R}^n\setminus\{0\}\quad (Ah,h)>0.$

(ii) A が負値 $\Leftrightarrow \forall h\in\mathbf{R}^n\setminus\{0\}\quad (Ah,h)<0.$

(iii) A が不定符号 $\Longrightarrow \exists h,h'\in\mathbf{R}^n$ s.t. $(Ah,h)>0,\ (Ah',h')<0.$

(証明のあらすじ) 適当な実直交行列 U が存在して、

$$U^TAU=\begin{pmatrix}\lambda_1 & & \text{\huge 0}\\ & \ddots & \\ \text{\huge 0} & & \lambda_n\end{pmatrix}.$$

このとき $x=Uy$ とおくと ($y=U^Tx$ とおくと)、

$$(Ax,x)=(AUy,Uy)=\left(U^TAUy,y\right)=\left(\begin{pmatrix}\lambda_1 & & \text{\huge 0}\\ & \ddots & \\ \text{\huge 0} & & \lambda_n\end{pmatrix}y,y\right)=\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2.$$

後は $x=0\Leftrightarrow y=0$ に注意して符号を考えれば良い。 ■

2.10.6 実対称行列の符号の判定 (オーソドックス版)

与えられた実対称行列が正値か、負値か、不定符号か、あるいはそのいずれでもないか、判定する必要が生じることがある。

「そんなのは線形代数の話だから、微積分の授業ではカットする」と言いたいのをぐっとこらえて。

実対称行列 A が正値であるとは、定義によれば A の固有値がすべて 0 であることだから、A の特性方程式 $\det(\lambda I - A) = 0$ の解がすべて正であることと同値である。残念ながら特性方程式は解くのが大変な場合が多く、3 次方程式の場合で既に難しい。

定理 2.10.10 (首座小行列式の符号による正値性、負値性の判定) n 次実対称行列 $A = (a_{ij})$ に対して、A_k を A の k 次首座行列とするとき、以下の (1), (2) が成り立つ。

(1) A が正値 $\iff \forall k \in \{1, \cdots, n\} \quad \det A_k > 0.$

(2) A が負値 $\iff \forall k \in \{1, \cdots, n\} \quad (-1)^k \det A_k > 0.$ ($\det A_1$, $\det A_2$, $\ldots$, $\det A_n$ の符号が交互に負, 正, 負, 正,...)

補題 2.10.11 対称行列の逆行列は対称行列である:

$$A^T = A, \quad \det A \neq 0 \quad \Rightarrow \quad (A^{-1})^T = A^{-1}.$$

証明 A が対称かつ正則であるとする。$AA^{-1} = A^{-1}A = I$ の転置を取ると、$(A^{-1})^T A^T = A^T (A^{-1})^T = I^T$. A が対称であることから、$(A^{-1})^T A = A(A^{-1})^T = I$. これから、$(A^{-1})^T = A^{-1}$. ■

定理 2.10.10 の証明 以下の証明は齋藤 [7] pp. 156–157 を書き直したものである。

(必要性) A は正値であると仮定する。$k \in \{1, \cdots, n\}$, $x \in \mathbf{R}^k$ とするとき、$x' := (x, \overbrace{0, \cdots, 0}^{n-k\text{ 個}}) \in \mathbf{R}^n$ とおくと、$(A_k x, x) = (Ax', x')$. $x \neq 0$ とすると、$x' \neq 0$ であるから、A の正値性の仮定から

$$(A_k x, x) = (Ax', x') > 0.$$

ゆえに A_k は正値であり、その固有値 $\lambda_j^{(k)}$ $(j = 1, \cdots, k)$ はすべて正であるから、

$$\det A_k = \prod_{j=1}^{k} \lambda_j^{(k)} > 0.$$

(十分性) 「任意の n 次実対称行列 A に対して、$\det A_k > 0$ $(k = 1, \cdots, n)$ ならば A は正値である」を、n に関する数学的帰納法によって証明する。$n = 1$ のとき、$A = a_{11}$ について $\det A_1 = a_{11}$ であるから、$\det A_1 > 0$ であれば A は正値である。$n - 1$ のときに成り立つと仮定する。A は n 次実対称行列で、$\det A_k > 0$ $(k = 1, \cdots, n)$ を満たすと仮定すると、A_{n-1} は正値である。

$$A = \begin{pmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & c \end{pmatrix}$$

とブロック分けする。

$$P := \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad B := \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \quad (I_{n-1} \text{ は } n-1 \text{ 次単位行列})$$

とおくと、

$$P^T BP = A. \tag{2.11}$$

(これは良くある変換のようだが、分かりやすい説明はないものか…)

実際

$$\begin{aligned}
P^T BP &= \begin{pmatrix} I_{n-1} & \mathbf{0} \\ (A_{n-1}^{-1}\boldsymbol{b})^T & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \\
&= \begin{pmatrix} I_{n-1} & \mathbf{0} \\ \boldsymbol{b}^T A_{n-1}^{-1} & 1 \end{pmatrix} \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \mathbf{0}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & \mathbf{0} \\ \boldsymbol{b}^T & c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \begin{pmatrix} I_{n-1} & A_{n-1}^{-1}\boldsymbol{b} \\ \mathbf{0}^T & 1 \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} + c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} \end{pmatrix} \\
&= \begin{pmatrix} A_{n-1} & \boldsymbol{b} \\ \boldsymbol{b}^T & c \end{pmatrix} = A.
\end{aligned}$$

A が正値であることを示すには、B が正値であることを示せば十分である。

$\det P = 1$ に注意すると、(2.11) から

$$\det A = \det A_{n-1} \left(c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b}\right).$$

$\det A > 0$, $\det A_{n-1} > 0$ より $c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b} > 0$.

$x = (x_1, \cdots, x_n)^T \in \mathbf{R}^n$ に対して、$x' := (x_1, \cdots, x_{n-1})^T$ とおくと、

$$(Bx, x) = (A_{n-1}x', x') + (c - \boldsymbol{b}^T A_{n-1}^{-1}\boldsymbol{b})x_n^2.$$

$x \neq 0$ より $x' \neq 0$ または $x_n \neq 0$ であることに注意すると、$(Bx, x) > 0$. ゆえに B は正値である。■

- 与えられた行列が対角行列ならば、対角成分が固有値なので、判定は簡単である。

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

  は対角行列で、固有値は 1, 2, 3 でいずれも正なので A は正値。

- 与えられた行列が 2 行 2 列ならば、特性方程式を解いて固有値を求めるのもアリ。

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

  の特定多項式は $\det(\lambda I - A) = \lambda^2 - (2+1)\lambda + (2\cdot 1 - 1\cdot 1) = \lambda^2 - 3\lambda + 1$. 固有値は $\lambda = \dfrac{3 \pm \sqrt{3^2 - 4\cdot 1\cdot 1}}{2} = \dfrac{3 \pm \sqrt{5}}{2} > 0$ であるから、いずれも正である。ゆえに A は正値である。

- 与えられた行列の首座小行列式 $\det A_k$ の符号を調べる。すべて正ならば A は正値、$\det A_k$ $(k=1,2,\dots,n)$ が順に $-+-+\cdots$ であれば A は負値。例えば

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

の首座小行列式は

$$\begin{aligned}&\det A_1=2>0,\quad \det A_2=2\cdot 2-1\cdot 1=3>0,\\&\det A_3=2\cdot 2\cdot 2+1\cdot 1\cdot 1+1\cdot 1\cdot 1-1\cdot 2\cdot 1-2\cdot 1\cdot 1-2\cdot 1\cdot 1\\&\qquad\ =8+1+1-2-2-2=4>0\end{aligned}$$

であるから正値である。

- A が 2 次実対称行列で $\det A<0$ ならば不定符号であり、$\det A=0$ ならば正値でも負値でも不定符号でもない。
(A の固有値を λ_1, λ_2 とするとき、$\lambda_1\lambda_2=\det A$ であるから。)

問 2.10.1 1 変数の 2 次形式について論ぜよ。(それはどういうものか。正値、負値、不定符号などの判定法は？)

問 2.10.2 次の行列が正値であるか、負値であるか、不定符号であるか、そのどれでもないか、判定せよ。

(1) $\begin{pmatrix}2&0\\0&1\end{pmatrix}$ (2) $\begin{pmatrix}-2&0\\0&-1\end{pmatrix}$ (3) $\begin{pmatrix}2&0\\0&-1\end{pmatrix}$ (4) $\begin{pmatrix}4&3\\3&1\end{pmatrix}$

(5) $\begin{pmatrix}4&3\\3&4\end{pmatrix}$ (6) $\begin{pmatrix}-4&-3\\-3&-4\end{pmatrix}$ (7) $\begin{pmatrix}4&2\\2&1\end{pmatrix}$ (8) $\begin{pmatrix}0&0\\0&0\end{pmatrix}$

(9) $\begin{pmatrix}2&0\\0&0\end{pmatrix}$ (10) $\begin{pmatrix}2&0&0\\0&1&0\\0&0&3\end{pmatrix}$ (11) $\begin{pmatrix}-2&0&0\\0&-1&0\\0&0&-3\end{pmatrix}$ (12) $\begin{pmatrix}1&0&0\\0&-2&0\\0&0&0\end{pmatrix}$

(13) $\begin{pmatrix}4&3&0\\3&4&0\\0&0&1\end{pmatrix}$ (14) $\begin{pmatrix}-4&3&0\\3&-4&0\\0&0&-2\end{pmatrix}$ (15) $\begin{pmatrix}4&1&0\\1&4&1\\0&1&4\end{pmatrix}$ (16) $\begin{pmatrix}1&2&3\\2&2&0\\3&0&3\end{pmatrix}$ (固有値求めづらい)

問 2.10.3 2 変数の 2 次形式について論ぜよ。

2.10.7 Hesse 行列の符号による極値の判定定理の証明

定理 2.10.12 (Hesse 行列の符号による極値の判定定理) Ω が $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$ が C^2 級, $a\in\Omega$, $f'(a)=0$, $H(a):=f$ の a における Hesse 行列 とするとき、次の (i), (ii), (iii) が成り立つ。

(i) $H(a)$ が正値 $\Longrightarrow$ f は a で狭義の極小となる。

(ii) $H(a)$ が負値 $\Longrightarrow$ f は a で狭義の極大となる。

(iii) $H(a)$ が不定符号 $\Longrightarrow$ f は a で極値を取らない。

証明 Ω が開集合であるから、$\exists\varepsilon > 0$ s.t. $B(a;\varepsilon) \subset \Omega$. Taylor の定理から、

$$\begin{aligned}(\forall h \in \mathbf{R}^2 : \|h\| < \varepsilon)(\exists\theta \in (0,1)) \quad f(a+h) &= f(a) + f'(a)h + \frac{1}{2}\left(H(a+\theta h)h, h\right) \\ &= f(a) + \frac{1}{2}\left(H(a+\theta h)h, h\right).\end{aligned}$$

(i) $H(a)$ が正値ならば、$\forall k \in \{1, \cdots, n\}$ $\det H_k(a) > 0$. f が C^2 級で、$H(x)$ は連続だから、$\exists\varepsilon' \in (0,\varepsilon)$, $(\forall h \in \mathbf{R}^n\colon \|h\| < \varepsilon')(\forall\theta \in (0,1))$ $\det H_k(a+\theta h) > 0$. このとき $H(a+\theta h)$ は正値である。ゆえに $h \neq 0$ ならば $(H(a+\theta h)h, h) > 0$ であるから、$f(a+h) > f(a)$. ゆえに f は a で狭義の極小である。

(ii) (i) と同様に証明できるので省略する。

(iii) $H(a)$ が不定符号とする。$\exists\lambda, \mu$: $H(a)$ の固有値で、$\lambda > 0$ かつ $\mu < 0$. $\exists u, v \in \mathbf{R}^n \setminus \{0\}$ s.t. $H(a)u = \lambda u$, $H(a)v = \mu v$. $0 < \|u\| < \varepsilon$, $0 < \|v\| < \varepsilon$ として良い (固有ベクトルは実数倍しても固有ベクトルであるから)。

$$g(t) := f(a+tu), \quad h(t) := f(a+tv)$$

とおくと、

$$\begin{aligned} g'(t) &= f'(a+tu)u, \quad g''(t) = (H(a+tu)u, u), \\ h'(t) &= f'(a+tv)v, \quad h''(t) = (H(a+tv)v, v), \\ g'(0) &= f'(a)u = 0, \quad g''(0) = (H(a)u, u) = \lambda(u, u) > 0, \\ h'(0) &= f'(a)v = 0, \quad h''(0) = (H(a)v, v) = \mu(v, v) < 0 \end{aligned}$$

である。$g(t)$ ($|t| < 1$) は $t = 0$ で狭義の極小、$h(t)$ ($|t| < 1$) は $t = 0$ で狭義の極大である。ゆえに f は a で極値を取らない (変化する方向によって極小となったり極大となったりする**峠点** (**鞍点**, **saddle point**) であるから)。 ■

2.10.8 典型例

例題 2.10.1 $f(x,y) = x^3 + y^3 - 3xy$ の極値を求めよ。

解 まず導関数を計算しよう:

$$\nabla f(x) = \begin{pmatrix} f_x \\ f_y \end{pmatrix} = \begin{pmatrix} 3x^2 - y \\ 3y^2 - x \end{pmatrix},$$

$$f \text{ の Hesse 行列 } H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

f の停留点を求める。

$$\begin{aligned}\nabla f(x,y) = 0 \quad &\Leftrightarrow \quad \begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 - y = 0 \\ y^2 - x = 0 \end{cases} \\ &\Leftrightarrow \quad \begin{cases} y = x^2 \\ x^4 - x = 0 \end{cases} \\ &\Leftrightarrow \quad x = y = 0 \quad \text{or} \quad x = y = 1.\end{aligned}$$

(1) 点 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ においては

$$H = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}, \quad \therefore \det H = -9 < 0.$$

H は不定符号であるから、この点は極値点ではない。

(2) 点 $\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ においては

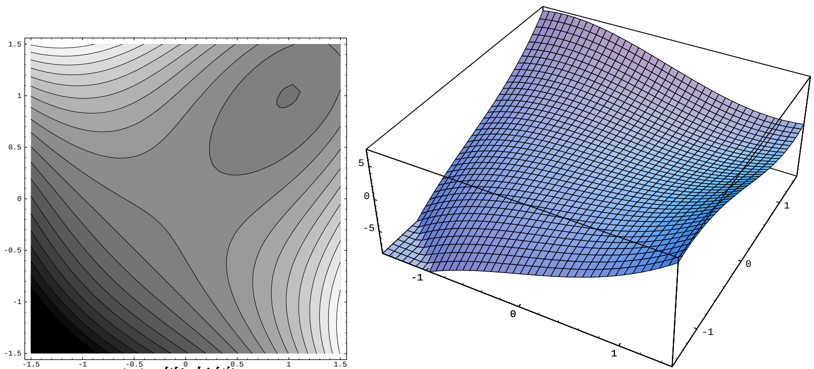
$$H = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}, \quad \therefore \det D_1 = 6 > 0, \quad \det D_2 = \det H = 6 \cdot 6 - (-3)(-3) = 27 > 0.$$

H は正値であるから、この点は極小点である。値は

$$f(1,1) = 1^3 + 1^3 - 3 \cdot 1 \cdot 1 = -1.$$

以上をまとめると、$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ のとき極小値 -1. ■

図 2.4 は $[-1.5, 1.5] \times [-1.5, 1.5]$ で関数の様子を調べたものである。



(a) 等高線　(b) 鳥瞰図（ちょうかんず）

図 2.4: $f(x,y) = x^3 + y^3 - 3xy$.

問 2.10.4 $f(x,y) := 2x^3 + 6xy^2 - 2x$ について、以下の問に答えよ。

(1) $f'(x,y)$ を求めよ。(2) f の Hesse 行列 $H(x,y)$ を求めよ。(3) f の極値を求めよ。(4) f のグラフ $z = f(x,y)$ の $(x,y) = (1,1)$ での接平面の方程式を求めよ。

2.10.9 実対称行列の符号の判定 — 少し真剣にアルゴリズムの追求

固有値計算作戦

固有多項式 $\varphi(\lambda) := \det(\lambda I - A)$ を計算して、その根を求める。$n = 2$ のときは 2 次方程式の解の公式で計算可能である。実際、$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ のとき、

$$\det(\lambda I - A) = \lambda^2 - t\lambda + D, \quad t := a + b, \quad D := ad - bc$$

であるから、

$$\lambda = \frac{t \pm \sqrt{t^2 - 4D}}{2}.$$

しかし $n \geq 3$ になると困難が生じる。$n = 3$ であっても、いわゆる不還元の場合には、解が虚数の 3 乗根を含む形で表されることになる (もちろん、微積分を使えば何とか処理できるが)。

首座小行列式作戦

$k = 1, 2, \cdots$ に対して、$\det A_k$ (A_k は A の k 次首座小行列) を計算して符号を調べる。

(i) すべて正である ($\forall k \in \{1, \cdots, n\}$ $\det A_k > 0$) ことは、A が正値であるための必要十分条件である。

(ii) 負から始まり、負と正が交互に現れる ($\forall k \in \{1, \cdots, n\}$ $(-1)^k \det A_k > 0$) ことは、A が負値であるための必要十分条件である。

(iii) 上の (i), (ii) のいずれでもない場合、$\det A$ を計算する。もし $\det A \neq 0$ であるならば、A は不定符号である。$\det A = 0$ のときは、一般には面倒だが、

   (a) $n = 2$ の場合は、正値、負値、不定符号のいずれでもないと結論できる ($\det A < 0 \iff A$ は不定符号)。

   (b) また $n = 3$ の場合は、固有多項式が容易に因数分解可能で、符号の判定は容易である。結論だけ書いておくと

$$A = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

   の固有多項式は

$$\det(\lambda I - A) = \lambda^3 - (a + d + f)\lambda^2 + (-b^2 - c^2 - e - 2 + ad + af + df)\lambda - \det A$$

   であるから、$\det A = 0$ を満たすとき、

$$\lambda_1 \lambda_2 = -b^2 - c^2 - e - 2 + ad + af + df$$

   が負ならば不定符号、そうでないならば正値でも負値でも不定符号でもない。

   (c) $n \geq 4$ の場合は研究課題であろう。

Gauss の消去法作戦

A の対角線から下を掃き出す。A が正値であれば、対角線に非零要素は現れず、行交換なしに最後まで上三角化が進められて、対角線に正数が並ぶはずである。一方、A が負値であれば、対角線に非零要素は現れず、行交換なしに最後まで上三角化が進められて、対角線に負数が並ぶはずである。そのいずれでもない場合、必要ならば行交換を施して計算を進めて A の行列式を計算する (行交換を全部で r 回した場合、最終的には対角成分の積 $\times(-1)^r = \det A$ である)。$\det A \neq 0$ ならば、A は 0 を固有値に持たず、正値でも負値でもないので、A は実は不定符号であることが分かる。$\det A = 0$ の場合は少々難しいが、シフトしてみるなどして (つまり A の代りに、$A + \sigma I$ (σ は適当に選ぶ実数) を調べる)、「何とかなる」場合が多いであろう。

$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix}$ に対する Gauss の消去法は、行交換なしに

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -6 \\ 0 & -6 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -6 \\ 0 & 0 & 15 \end{pmatrix}$$

となるので、いわゆる符号は $(2,1)$ (正の固有値は 2 個、負の固有値が 1 個) で、不定符号である。

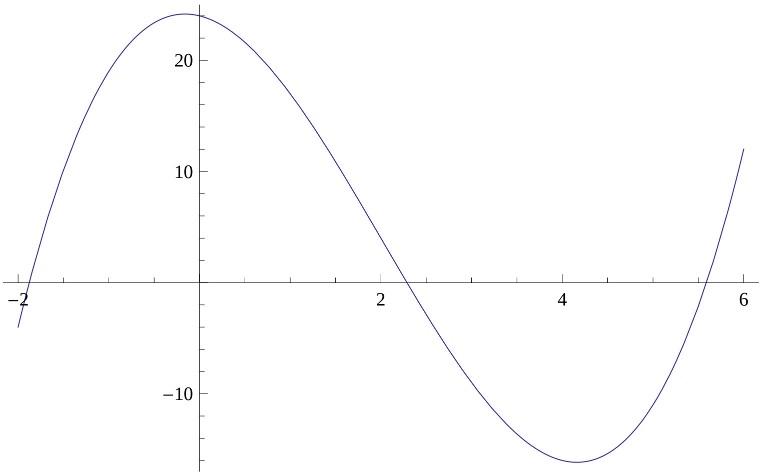
$p(\lambda) := \det(\lambda I - A)$ の根は

$$\lambda = \frac{1}{3}\left(6 + \frac{7\,6^{2/3}}{\sqrt[3]{-9+i\sqrt{1977}}} + \sqrt[3]{6i\left(\sqrt{1977}+9i\right)}\right),$$
$$2 - \frac{\left(1+i\sqrt{3}\right)\sqrt[3]{-9+i\sqrt{1977}}}{6^{2/3}} + \frac{-7+7i\sqrt{3}}{\sqrt[3]{6i\left(\sqrt{1977}+9i\right)}},$$
$$2 + \frac{i\left(\sqrt{3}+i\right)\sqrt[3]{-9+i\sqrt{1977}}}{6^{2/3}} + \frac{-7-7i\sqrt{3}}{\sqrt[3]{6i\left(\sqrt{1977}+9i\right)}}.$$

分かりづらいけれど、これは (もちろん) いずれも実数で

$$\lambda \fallingdotseq 5.580664\cdots, 2.2874\cdots, -1.877074\cdots.$$

p のグラフは次のようになる。



処世術

Gauss の消去法で行列の符号が計算できることは、知らない人も結構いると思われる。ペーパーテストで Gauss の消去法で符号を調べるのは少し危険があるかもしれない。

2.10.10 多変数関数の最大最小問題やりかけの問題

極値問題の最初に、

$$f(x,y) := 3x^2 + 2y^2 + 2xy - 2x - 2y + 1 \tag{2.12}$$

の、

$$K := \{(x,y)\in\mathbf{R}^2; x\ge 0,\ y\ge 0,\ x+y\le 1\} \tag{2.13}$$

における最大値、最小値を求めよ、という問題を提示した。

f の $\mathbf{R}^2$ での極値を求めてみよう。$\nabla f(x,y)=0 \iff (x,y)=\left(\dfrac{1}{5},\dfrac{2}{5}\right)$. この点で Hesse 行列は正値であるので、$f$ は狭義の極小になることが分かる。極小値 $f\left(\dfrac{1}{5},\dfrac{2}{5}\right)=\dfrac{2}{5}$.

定理 2.10.13 (Weierstrass の最大値定理) コンパクト集合 K で定義された連続関数 $f\colon K\to\mathbf{R}$ は、必ず最大値と最小値を持つ。

系 2.10.14 K が $\mathbf{R}^n$ の空でない有界閉集合とするとき、連続関数 $f\colon K\to\mathbf{R}$ は必ず最大値と最小値を持つ。

(2.13) の K は明らかに閉集合である (多項式関数 (それは連続！) と $\ge$, $\le$ で定義されている)。

「有界」については定義を思い出そう。

定義 2.10.15 (有界集合) $A\subset\mathbf{R}^n$ とする。A が有界であるとは、

$$\exists R\in\mathbf{R}\quad \text{s.t.}\quad \forall x\in A\quad \|x\|\le R$$

が成り立つことをいう。

言い換えると、A が有界であるとは、原点を中心とする十分大きな半径 R の閉球 $\overline{B}(0;R)$ に A が含まれることである:

$$A\text{ が有界}\quad\Leftrightarrow\quad \exists R\in\mathbf{R}\quad\text{s.t.}\quad A\subset\overline{B}(0;R).$$

(2.13) の K は、

$$K\subset\overline{B}(0;1)$$

を満たすので、有界性の条件が $R=1$ として成り立つ。ゆえに (2.13) の K は有界である。

従って Weierstrass の最大値定理によって、(2.12) で定義された f は K で最大値、最小値を持つ。

K を内部

$$K^\circ=\{(x,y)\in\mathbf{R}^2; x>0,\ y>0,\ x+y<1\}$$

と境界

$$\begin{aligned}
&K^b=\overline{K}\setminus K^\circ=E_1\cup E_2\cup E_2=(\text{三角形の周}),\\
&E_1:=\{(x,0);x\in[0,1]\},\quad E_2:=\{(0,y);y\in[0,1]\},\quad E_2:=\{(x,1-x);x\in[0,1]\}
\end{aligned}$$

に分けて最大値、最小値を考える。

$$\max_{K^b} f=\max\left\{\max_{E_1} f,\max_{E_2} f,\max_{E_3} f\right\}$$

であるが、

$$\begin{aligned}
\max_{E_1} f &= \max_{x\in[0,1]} f(x,0) = 2 \quad (= f(1,0)),\\
\max_{E_2} f &= \max_{y\in[0,1]} f(0,y) = 1 \quad (= f(0,0) = f(0,1)),\\
\max_{E_3} f &= \max_{x\in[0,1]} f(x,1-x) = 2 \quad (= f(1,0))
\end{aligned}$$

であるから、

$$\max_{K^b} f = 2 \quad (= f(1,0)).$$

一方、

$$\min_{K^b} f = \min\{\min_{E_1} f, \min_{E_2} f, \min_{E_3} f\}$$

において

$$\begin{aligned}
\min_{E_1} f &= \min_{x\in[0,1]} f(x,0) = \frac{2}{3} \quad (= f\left(\frac{1}{3},0\right)),\\
\min_{E_2} f &= \min_{y\in[0,1]} f(0,y) = \frac{1}{2} \quad (= f\left(0,\frac{1}{2}\right),\\
\min_{E_3} f &= \min_{x\in[0,1]} f(x,1-x) = \frac{2}{3} \quad (= f\left(\frac{1}{3},\frac{2}{3}\right))
\end{aligned}$$

であるから、

$$\min_{K^b} f = \frac{1}{2} \quad \left(f(\left(0,\frac{1}{2}\right))\right).$$

f が $a \in K$ で最大になったとする ($f(a) = \max_K f$)。$a \notin K^\circ$ である。実際、$a \in K^\circ$ であれば、$\nabla f(a) = 0$, f は a で極大となるはずであるが、K° で $\nabla f(x) = 0$ となる点は $x = \left(\frac{1}{5},\frac{2}{5}\right)$ のみで、そこで f は狭義の極小となるので、極大とはなりえない。ゆえに $a \in K^b$. ゆえに

$$\max_K f = \max_{K^b} f = 2 \quad (= f(1,0)).$$

一方、$f\left(\frac{1}{5},\frac{2}{5}\right) = \frac{2}{5} < \frac{1}{2} = \max_{K^b} f$ であるから、f は K^b で最小値を取らず、K° で最小値を取る。それは極小値であるから、$f\left(\frac{1}{5},\frac{2}{5}\right)$ でなければならない。

まとめると、f は $\left(\frac{1}{5},\frac{2}{5}\right)$ で最小値 $\frac{2}{5}$ を取り、$(1,0)$ で最大値 2 を取る。■

2.10.11 補足: グラフを見る

2変数の場合に典型的な2次形式 (行列が対角行列) の場合を見てみる。プリントを配布する。

鞍点 (峠点, saddle point) の場合を納得すること。

$f_3(x,y) = x^2 - y^2$ と $f_4(x,y) = 2xy$ は座標系を回転させると重なる (つまりグラフは合同である)。

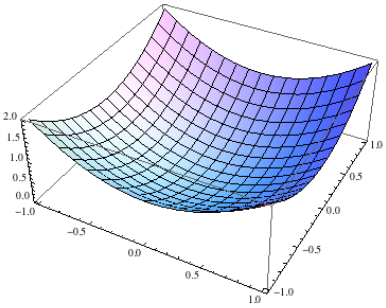


図 2.5: $f_1(x, y) = x^2 + y^2$

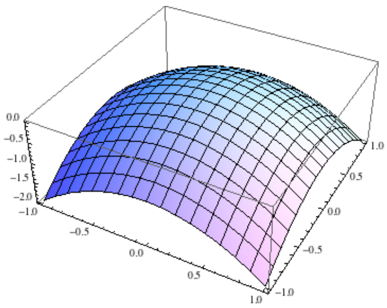


図 2.6: $f_2(x, y) = -x^2 - y^2$

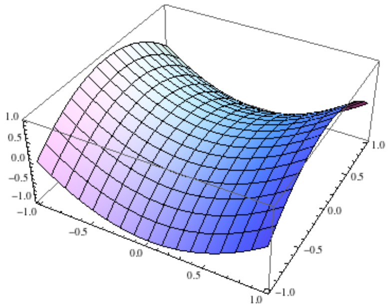


図 2.7: $f_3(x, y) = x^2 - y^2$

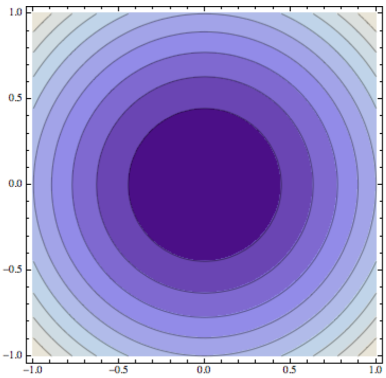


図 2.8: $f_1(x, y) = x^2 + y^2$

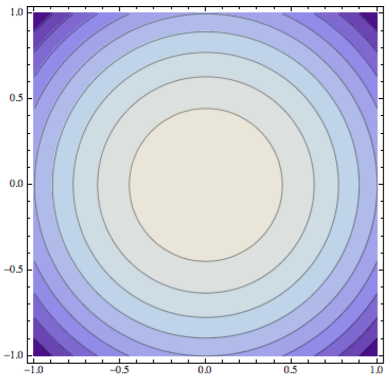


図 2.9: $f_2(x, y) = -x^2 - y^2$

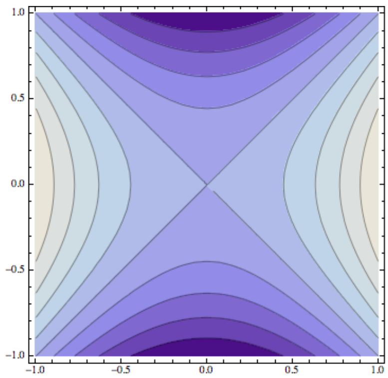


図 2.10: $f_3(x, y) = x^2 - y^2$

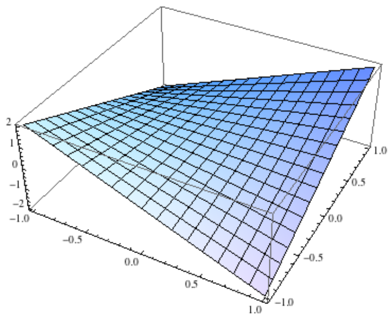


図 2.11: $f_4(x, y) = 2xy$

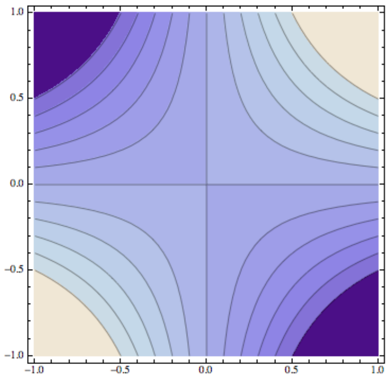


図 2.12: $f_4(x, y) = 2xy$

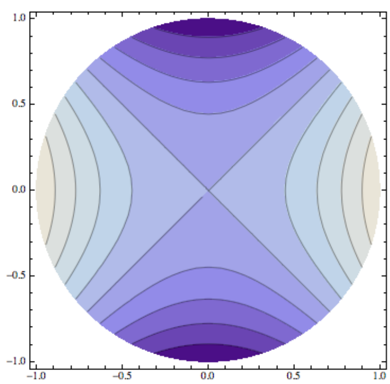


図 2.13: $f_3(x, y) = x^2 - y^2$

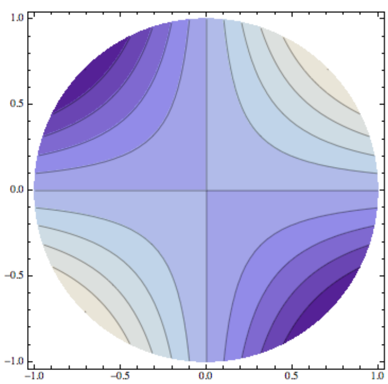


図 2.14: $f_4(x, y) = 2xy$

2.10.12 Hesse 行列が “どれでもない” 場合

$f_5(x,y)=x^4+y^4$, $f_6(x,y)=-x^4-y^4$, $f_7(x,y)=x^4-y^4$ は、いずれも $f'(0,0)=0$, $H(0,0)=\begin{pmatrix}0&0\\0&0\end{pmatrix}$ を満たすが、これは正値でも負値でも不定符号でもない。それぞれ狭義の極小 ($(x,y)\neq(0,0)$ ならば $f_5(x,y)=x^4+y^4>0=f_5(0,0)$)、狭義の極大 ($(x,y)\neq(0,0)$ ならば $f_6(x,y)=-(x^4+y^4)<0=f_6(0,0)$)、極値でない ($x,y\neq 0$ ならば $f_7(x,0)=x^4>0>-y^4=f_7(0,y)$)。2 階までの微分をしらべても分からないこともある、ということ。

2.10.13 連立方程式の解き方

極値問題の定理は、多変数関数を Taylor 展開を基礎として取り扱う、という意味で教育的であると考えているが、期末試験のために問題を作成するのは、結構教師にとって悩ましい作業である。連立方程式を解く方法を論じる授業ではないので、$f'(x)=0$ は簡単に解けるのが望ましいのだが、そのような例を探すのは結構面倒なのである。

時々失敗して、連立方程式を解けない人が大勢出てしまうことがある。(気づき次第ヒントを板書したりするのだが…)

いずれにせよ、少し問題を解いて練習して下さい。

「未知数の消去」が基本だけれど、対称性に目をつけて (そういう問題を出します)、辺々加えたり、引いてみたりするとうまく行く場合がある。

例として、問 11 に現れた

$$\begin{gathered}x^3+3xy^2-x=0,\\ y^3+3x^2y-y=0\end{gathered}$$

でやってみよう。辺々引き算して $(x-y)(x^2-2xy+y^2-1)=0$. さらに

$$(x-y)(x-y+1)(x-y-1)=0.$$

これから $x-y=0,1,-1$. ゆえに $y=x,x-1,x+1$. それぞれを第 2 式に代入すれば、y が消去できて解ける。

一方、最初に因数分解して

$$\begin{cases}x(x^2+3y^2-1)=0\\ y(y^2+3x^2-1)=0\end{cases}$$
$$\Leftrightarrow \bigl(x=0\vee x^2+3y^2-1=0\bigr)\wedge(y=0\vee y^2+3x^2-1=0).$$

これを $(P\vee Q)\wedge(R\vee S)$ と見て、$(P\wedge R)\vee(P\wedge S)\vee(Q\wedge R)\vee(Q\wedge S)$ と展開する (論理の分配法則)。具体的には、

$$\begin{cases}x=0\\ y=0\end{cases}\vee\begin{cases}x=0\\ y^2+3x^2-1=0\end{cases}\vee\begin{cases}x^2+3y^2-1=0\\ y=0\end{cases}\vee\begin{cases}x^2+3y^2-1=0\\ y^2+3x^2-1=0.\end{cases}$$

後は省略する。答合わせ用に結果のみ書いておくと、

$$(x,y)=(0,0),(0,\pm1),(\pm1,0),(1/2,\pm1/2),(-1/2,\pm1/2)\quad(\text{全部で 9 個}).$$

問 2.10.5 (1) A が 2 次の実対称行列で $\det A<0$ を満たすならば、A は不定符号であることを示せ。(2) A が実対称行列で $\det A<0$ であっても、A は不定符号とは限らない。反例を示せ。(3) $a\in\mathbf{R}$, $A=\begin{pmatrix}0&a\\a&0\end{pmatrix}$ とするとき、A が不定符号であるための必要十分条件を求めよ。

2.11 陰関数定理と逆関数定理 — 存在定理

兄弟の関係にある「陰関数定理」と「逆関数定理」を駆け足で説明する。
どちらも

「指定された点の近くで (局所的に) 関数 (それぞれ陰関数、逆関数) が存在する」

という**存在定理**である。
存在定理というと、2 年生にはなじみが薄いかも知れないが、まったくの初めてというわけではなくて、

1. 中間値の定理「連続関数 $f\colon [a,b]\to\mathbf{R}$ が $f(a)f(b)<0$ を満たすならば、$f(c)=0$ を満たす $c\in(a,b)$ が存在する」
2. 代数学の基本定理「複素係数の n 次多項式 $a_0z^n+\cdots+a_{n-1}z+a_n$ は複素数の範囲に少なくとも一つの根を持つ」
3. Weierstrass の最大値定理「コンパクト集合 K 上の実数値連続関数は、最大値を持つ」

という例 (どれも非常に重要) がある。
陰関数定理も逆関数定理も「関数の存在」を主張している。証明においては、x が与えられたときに $F(x,y)=0$ を y について解く、y が与えられたときに $f(x)=y$ を x について解く、と方程式の解の存在をするのが関門である。

2.11.1 逆関数定理を超特急で説明

逆関数については、

「写像 f が逆写像 f^{-1} を持つためには、f が全単射であることが必要十分である」

というのが基本中の基本である。
与えられた関数 f そのものが全単射でなくても、それを適当に制限したものが全単射になり、その逆写像 (逆関数) が便利、というのが良くある話である。

例 2.11.1 (高校数学からの例) $f\colon \mathbf{R}\ni x\mapsto x^2\in\mathbf{R}$ は全射でも単射でもないが、その制限

$$\widetilde{f}\colon [0,\infty)\ni x\mapsto x^2\in[0,\infty)$$

は全単射であり、その逆関数 $\widetilde{f}^{-1}$ はいわゆるルートである。

$$\widetilde{f}^{-1}(y)=\sqrt{y}\quad (y\in[0,\infty)).$$

似たようなことは、exp と log, 各種逆三角関数であった。■

1 変数関数に対する逆関数の定理は簡単であるので、概略を述べてみよう。

例 2.11.2 (1 変数の逆関数の定理) I が $\mathbf{R}$ の開区間、$f\colon I\to\mathbf{R}$ は C^1 級、$a\in I$, $f'(a)\neq 0$ ならば、

$$(\exists U\subset I : a\ \text{を含む開区間})\ (\exists V : b:=f(a)\ \text{を含む開区間})$$
$$\widetilde{f}\colon U\ni x\longmapsto f(x)\in V\quad \text{は全単射で逆関数も}\ C^1\ \text{級}$$

が成り立つ。実際 $f'(a) \neq 0$ であるから、$f'(a) > 0$ or $f'(a) < 0$. $f'(a) > 0$ の場合、f' の連続性から、$\exists \varepsilon > 0$ s.t. $f' > 0$ on $[a-\varepsilon, a+\varepsilon]$. このとき f は $[a-\varepsilon, a+\varepsilon]$ で狭義単調増加である。このとき、$U := (a-\varepsilon, a+\varepsilon)$, $V = (f(a-\varepsilon), f(a+\varepsilon))$ とすると、$\widetilde{f}$ は明らかに定義できて、単射である。また中間値の定理を用いて全射であることが分かる。ゆえに $\widetilde{f}$ は全単射であるから逆関数が存在する。少し頑張ると $\widetilde{f}^{-1}$ の連続性と、微分可能性、$\widetilde{f}^{-1}$ の連続性が証明できる (詳しくは桂田 [1] の付録 H.2「1 変数の逆関数の定理」)。■

逆関数の定理は、この例の素直な多次元化である。しかしそれを説明する前に、線形代数の復習をしておく。

例 2.11.3 (線型写像が全単射となる条件) 有限次元線型空間の間の線型写像を考える。一般形は、$n, m \in \mathbf{N}$, $A \in M(m,n;\mathbf{R})$ として、

$$f\colon \mathbf{R}^n \ni x \mapsto Ax \in \mathbf{R}^m$$

である。このとき、有名な次元定理

$$\operatorname{rank} f = n - \dim\ker f$$

が成り立つ。これはやや高級な定理であるが、これを認めれば後の議論は簡単である。まず、

- f が全射 $\iff \operatorname{rank} f = m$
- f が単射 $\iff \dim\ker f = 0$

であるから、

$$\begin{aligned} f \text{ が全単射} &\Longrightarrow (\operatorname{rank} f = m \quad \text{and} \quad \dim\ker f = 0) \\ &\Longrightarrow m = n. \end{aligned}$$

ゆえに全単射であるためには、空間の次元に関する条件 $m = n$ が必要である。そこで以下 $m = n$ を前提条件とする。このとき

$$\begin{aligned} f \text{ が全射} &\iff \operatorname{rank} f = m (= n) \\ &\iff \dim\ker f = 0 \\ &\iff f \text{ は単射} \\ &\iff f \text{ は全単射} \\ &\iff f^{-1} \text{ が存在} \\ &\iff A^{-1} \text{ が存在} \\ &\iff \det A \neq 0. \end{aligned}$$

後のために次のように覚えておこう。「全単射な線型写像が存在するために、定義域と終域の空間次元が等しいことが必要で、それが成り立つという前提のもとで、与えられた線型写像が全単射であるためには、その行列式が 0 でないことが必要十分である」 ■

いよいよ逆関数定理の登場である。

定理 2.11.4 (逆関数定理) Ω は $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}^n$ は C^1 級、$a \in \Omega$, $\det f'(a) \neq 0$ とするとき、($\exists U$: a を含む開集合) ($\exists V$: $b = f(a)$ を含む開集合) s.t. $\widetilde{f} := f|_U : U \ni x \mapsto f(x) \in V$ は全単射で、逆関数 $\widetilde{f}^{-1}\colon V \to U$ も C^1 級である。

時間の関係で、証明は涙を飲んで省略するが (苦笑)、逆関数の導関数については、既に学んだ逆関数の微分法 (定理 2.8.6) が成立することを注意 (「思い出せ！」) しておく。

2.11.2 陰関数についてのイントロ (2 変数関数版)

陰関数定理については、簡単な具体例を並べるところから始めよう。

直観的には、方程式 $F(x,y)=0$ は、(例外的な状況を除けば) 平面曲線を定め、適当に範囲を限定すると、変数 x の関数 $y=\varphi(x)$ を定めることがある (このとき、その関数 $y=\varphi(x)$ を $F(x,y)=0$ の定める陰関数と呼ぶ)。

いくつか実例を並べてみよう。

(1) $F(x,y)=y-\varphi(x)$ のとき、$y=\varphi(x)$. $F(x,y)=0$ の定める曲線は、関数 φ のグラフである。

(2) $F(x,y)=ax+by+c$ $(a,b,c\in\mathbf{R},\ (a,b)\neq(0,0))$ のとき、$F(x,y)=0$ の定める曲線は直線である。$b\neq 0$ であれば、$y=-\dfrac{a}{b}x-\dfrac{c}{b}$ と解ける。

(3) $F(x,y)=x^2+y^2-1$ のとき、$y=\pm\sqrt{1-x^2}$. $F(x,y)=0$ の定める曲線は、原点を中心とする半径 1 の円周である。(一般に、$F(x,y)$ が x と y の 2 次多項式であるならば、$F(x,y)=0$ の定める曲線は、いわゆる 2 次曲線で、具体的には、空集合、1 点、2 直線、楕円、放物線、双曲線である — 線形代数のテキストを見よ)。

(4) $F(x,y)=y^2-x^2(x-a)$ (a は実定数) のとき、$F(x,y)=0$ は、$y=0$ ($x=0$ のとき), または $y=\pm x\sqrt{x-a}$ ($x\geq a$ のとき) と解ける[20]。$F(x,y)=0$ の定める曲線は、

(a) $a<0$ のときは原点で自己交差する曲線 (原点を結節点と呼ぶ)

(b) $a=0$ のときは原点で尖っている曲線 (原点を尖点と呼ぶ)

(c) $a>0$ のときは原点と、$x\geq a$ の範囲にある曲線 (原点を孤立点と呼ぶ)

```
g0=ListPlot[{{0,0}}]
myg[a_]:=ContourPlot[y^2-(x-a)x^2==0,{x,-2,2},{y,-2,2},ContourStyle->Red]
g=Show[myg[1],g0]
```

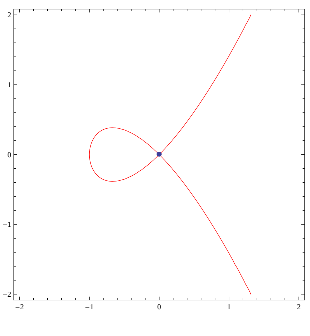


図 2.15: $a=-1$

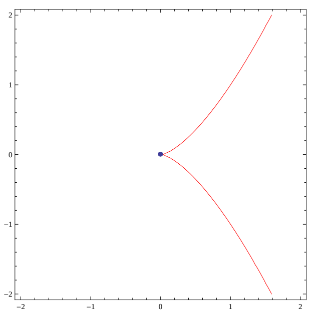


図 2.16: $a=0$

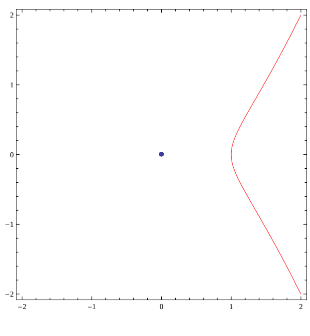


図 2.17: $a=1$

(5) (ヤコブ・ベルヌーイのレムニスケート, 1694 年) $F(x,y)=(x^2+y^2)^2-2(x^2-y^2)$ のとき、いわゆるヤコブ・ベルヌーイのレムニスケート (連珠形) $F(x,y)=0$ は y についての 4 次方程式であるが、2 次方程式を解くことを 2 回行って、y について解ける。

[20] $y^2=x^2(x-a)$ としたとき、実数の範囲で解ける $\Leftrightarrow x^2(x-a)\geq 0 \Leftrightarrow [x=0$ または $x\geq a]$ であることに注意せよ。$x=0$ のときは $y=0$, $x\geq a$ のときは、$y=\pm\sqrt{x^2(x-a)}=\pm|x|\sqrt{x-a}=\pm x\sqrt{x-a}$.

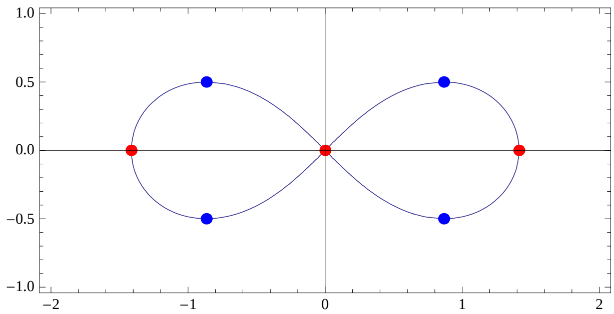


図 2.18: レムニスケート $(x^2+y^2)^2-2(x^2-y^2)=0$

(6) (デカルトの葉線, folium cartesii, 1638 年) $F(x,y)=x^3+y^3-3xy$ のとき、$F(x,y)=0$ の定める曲線は、いわゆるデカルトの葉線で、原点において自分自身と交差する曲線である (例 2.11.8, p.104)。$F(x,y)=0$ は y についての 3 次方程式である。これは y について簡単に解くことは…？出来ないと思ったら、Mathematica は答を返して来た。あ、そうか。でも使いにくそう。

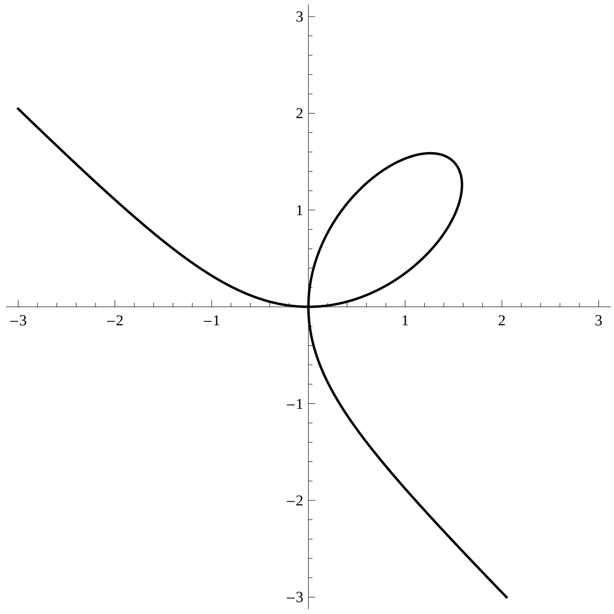


図 2.19: Decartes の葉線 $x^3+y^3-3xy=0$

余談 2.11.1 (Decartes の葉線の伝統的な描き方) 極座標を使うと、

$$r=\frac{3\cos\theta\sin\theta}{\cos^3\theta+\sin^3\theta}$$

という極方程式がすぐに得られる。あるいは $y=tx$ として、

$$x=\frac{3t}{1+t^3},\quad y=\frac{3t^2}{1+t^3}$$

という有理パラメーター表示も得られる。$x+y=-1$ が漸近線になっている。 ■

次のことが分かる。

- F によっては、$F(x,y)=0$ を具体的な式変形で y について解くことは不可能である。
  → 抽象的な「存在定理[21]」が望み得るゴールとなる。

[21]アナロジーとして、中間値の定理を思い出させる。

- 1つの x に 2 つ以上の y が対応したり、逆に 1 つも y がなかったりする。
  → 最初に $F(a,b)=0$ を満たす点 (a,b) があったとして、その点の「近傍」で考えることにする。とっかかりは要求することにする。(a,b) によってうまく行ったり行かなかったりする。

- 1つの x に複数の y が対応する場合も、注目している点を中心とした十分小さい範囲に限れば、1 つの x に 1 つの y が対応するようになることもある。
  → a を含む開集合 U, b を含む開集合 V をとり、$U\times V$ (イメージとしては窓枠、ウィンドウ) に考察を限定する、という方針が良さそう。

  もっとも、どんなに小さい範囲にしぼってもダメなこともある (その点で曲線が自己交差していたり、x について片側にしか対応する y がない)。うまく行くための十分条件はないか？
  → 実は $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ という条件が満たされればOK, と後で分かる。

- 陰関数の導関数は (そもそも存在するかはすぐには分からないことであるが、存在するならば)、合成関数の微分法で計算するのは簡単である。

  例: $x^2+y^2=1$ より、$2x+2y\cdot\dfrac{dy}{dx}=0$ だから、$\dfrac{dy}{dx}=-\dfrac{x}{y}$.
  一般には、$F(x,\varphi(x))=0$ より、

$$\frac{\partial F}{\partial x}(x,\varphi(x))+\frac{\partial F}{\partial y}(x,\varphi(x))\varphi'(x)=0 \quad \text{より} \quad \varphi'(x)=-\frac{\partial F}{\partial y}(x,\varphi(x))^{-1}\frac{\partial F}{\partial x}(x,\varphi(x)).$$

2.11.3 定理の陳述

以下 $\mathbf{R}^m\times\mathbf{R}^n$ が登場する。これはもちろん

$$\mathbf{R}^m\times\mathbf{R}^n=\left\{(x,y);x=\begin{pmatrix}x_1\\ \vdots\\ x_m\end{pmatrix}\in\mathbf{R}^m, y=\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}\in\mathbf{R}^n\right\}$$

であるから、

$$z=\begin{pmatrix}z_1\\ \vdots\\ z_m\\ z_{m+1}\\ \vdots\\ z_{m+n}\end{pmatrix}\quad (z_j\in\mathbf{R},\ j=1,2,\cdots,m+n)$$

全体の集合である $\mathbf{R}^{m+n}$ と同一視できる。そこで例えば $\Omega\subset\mathbf{R}^m\times\mathbf{R}^n$ が開集合と言った場合はこの同一視によって Ω が $\mathbf{R}^{m+n}$ の開集合であることを意味する。単に (x,y) が $\mathbf{R}^m\times\mathbf{R}^n$ の要素であると言った場合は、特に断りがなければ $x\in\mathbf{R}^m$, $y\in\mathbf{R}^n$ であるとする。

さて、$F\colon\Omega\to\mathbf{R}^n$ があるとき、

$$F=\begin{pmatrix}F_1\\ \vdots\\ F_n\end{pmatrix}$$

と書けば、

$$F'(x,y) = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_m} & \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_n} \\ \vdots & & \vdots & \vdots & & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \cdots & \dfrac{\partial F_n}{\partial x_m} & \dfrac{\partial F_n}{\partial y_1} & \cdots & \dfrac{\partial F_n}{\partial y_n} \end{pmatrix}$$

となるわけだが、m 列, n 列とブロックわけして、それぞれ $\dfrac{\partial F}{\partial x}$, $\dfrac{\partial F}{\partial y}$ と書く。すなわち

$$\frac{\partial F}{\partial x} = \begin{pmatrix} \dfrac{\partial F_1}{\partial x_1} & \cdots & \dfrac{\partial F_1}{\partial x_m} \\ \vdots & & \vdots \\ \dfrac{\partial F_n}{\partial x_1} & \cdots & \dfrac{\partial F_n}{\partial x_m} \end{pmatrix}, \quad \frac{\partial F}{\partial y} = \begin{pmatrix} \dfrac{\partial F_1}{\partial y_1} & \cdots & \dfrac{\partial F_1}{\partial y_n} \\ \vdots & & \vdots \\ \dfrac{\partial F_n}{\partial y_1} & \cdots & \dfrac{\partial F_n}{\partial y_n} \end{pmatrix}.$$

以下しばらくこの記号を使おう。

定理 2.11.5 (陰関数定理, implicit function theorem) Ω は $\mathbf{R}^m \times \mathbf{R}^n$ の開集合、$F\colon \Omega \ni (x,y) \mapsto F(x,y) \in \mathbf{R}^n$ は C^1 級、$(a,b) \in \Omega$, $F(a,b) = 0$, $\det \dfrac{\partial F}{\partial y}(a,b) \neq 0$ が成り立つとする。このとき、a を含む $\mathbf{R}^m$ の開集合 U, b を含む $\mathbf{R}^n$ の開集合 V, C^1 級の関数 $\varphi\colon U \to V$ で、以下の (0), (i), (ii), (iii) を満たすものが存在する。

(0) $U \times V \subset \Omega$.

(i) $\varphi(a) = b$.

(ii) $\forall (x,y) \in U \times V$ について、$F(x,y) = 0 \iff y = \varphi(x)$.

(iii) $\forall x \in U$ について、$\varphi'(x) = -\left(\dfrac{\partial F}{\partial y}(x,\varphi(x))\right)^{-1} \dfrac{\partial F}{\partial x}(x,\varphi(x))$.

注意 2.11.6 (覚え方のヒント) 上の定理は、大事なことをひとまとめにしたものだが、最低限必要なことと、それから導かれることに分けた方が覚えやすいかも知れない。

短縮版陰関数定理

Ω は $\mathbf{R}^m \times \mathbf{R}^n$ の開集合、$F\colon \Omega \ni (x,y) \mapsto F(x,y) \in \mathbf{R}^n$ は C^1 級、$(a,b) \in \Omega$, $F(a,b) = 0$, $\det \dfrac{\partial F}{\partial y}(a,b) \neq 0$ が成り立つならば、$\exists U$, $\exists V$, $\exists \varphi \in C^1(U;V)$ s.t.

(a) U は a の開近傍、V は b の開近傍で、$U \times V \subset \Omega$.

(b) $\forall (x,y) \in U \times V$ について、$F(x,y) = 0 \iff y = \varphi(x)$.

上の定理 2.11.5 に書いてあって、この短縮版に書いてないことを導こう。まず $F(a,b) = 0$ と (b) から $\varphi(a) = b$ が導かれる。また (b) から $F(x,\varphi(x)) = 0$ が得られるが、F と φ が C^1 級であるから、$F_x(x,\varphi(x)) + F_y(x,\varphi(x))\varphi'(x) = 0$. $\det F_y(a,b) \neq 0$ であるから、(a,b) の十分小さな近傍で $F_y(x,y)^{-1}$ が存在するので、$\varphi'(x) = -\left(F_y(x,\varphi(x))\right)^{-1} F_x(x,\varphi(x))$. ■

注意 2.11.7 (陰関数定理の条件 (ii) の言い換え「零点集合がグラフになる」) 定理 2.11.5 の (ii) は、「方程式が解ける」といういわば解析的な表現であるが、幾何学的な表現である次の (ii)' で置き換えることも出来る。

(ii)' $U \times V$ において、F の零点集合は φ のグラフに一致する: $N_F \cap (U \times V) = \operatorname{graph}\varphi$.

ここで N_F, $\operatorname{graph}\varphi$ はこれまでも登場した記号で、

$$N_F := \{(x,y) \in \Omega; F(x,y) = 0\}, \quad \operatorname{graph}\varphi := \{(x,\varphi(x)); x \in U\}. \blacksquare$$

2.11.4 単純な例

既に述べたように、陰関数定理は広範な応用を持つが、ここではなるべく単純な例を紹介する。

例 2.11.8 [デカルトの葉線 (folium of Descartes, folium cartesii, 1694)] $a > 0$ とするとき、$F(x,y) := x^3 + y^3 - 3axy$, $P = \left(\dfrac{3a}{2}, \dfrac{3a}{2}\right)$ とおく。点 P の十分小さな開近傍において、$F(x,y) = 0$ の陰関数 $y = \varphi(x)$ が存在することを示し、その点における微分係数を求めよ。

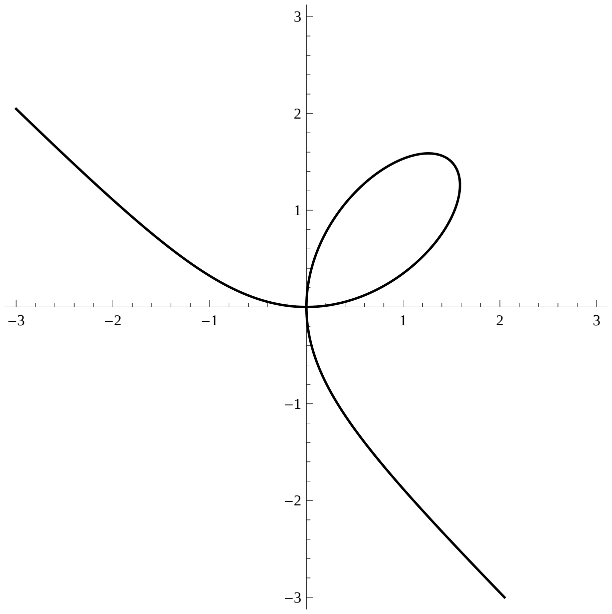


図 2.20: Mathematica による $F(x,y) := x^3 + y^3 - 3axy$ の零点集合 ($a = 2/3$ の場合)

```
g = ContourPlot[x^3 + y^3 - 3 x y==0, {x, -3, 3}, {y, -3, 3},
  ContourStyle->Thick, Axes->True, Frame->None]
```

解答 $F\colon \mathbf{R}^2 \to \mathbf{R}$ は C^1 級で、

$$F\left(\frac{3a}{2}, \frac{3a}{2}\right) = \left(\frac{3a}{2}\right)^3 + \left(\frac{3a}{2}\right)^3 - 3a\left(\frac{3a}{2}\right)^2 = \left(\frac{27}{8} + \frac{27}{8} - \frac{27}{4}\right)a^3 = 0,$$

$$F_y(x,y) = 3y^2 - 3ax, \quad F_y\left(\frac{3a}{2}, \frac{3a}{2}\right) = 3\left(\frac{3a}{2}\right)^2 - 3a\frac{3a}{2} = \frac{9a^2}{4} \neq 0$$

であるから、$\dfrac{3a}{2}$ の十分小さな開近傍 U と V が存在して、$U \times V$ で $F(x,y) = 0$ は $y = \varphi(x)$ と解けて、$\varphi\colon U \to V$ は C^1 級となる。$F(x,\varphi(x)) = 0$ より、$F_x(x,\varphi(x)) + F_y(x,\varphi(x))\varphi'(x) = 0$ となるので、$\varphi'(x) = -\dfrac{F_x(x,\varphi(x))}{F_y(x,\varphi(x))}$. $F_x(x,y) = 3x^2 - 3ay$, $F_x\left(\dfrac{3a}{2}, \dfrac{3a}{2}\right) = \dfrac{9a^2}{4}$ であるから、

$$\varphi'\left(\frac{3a}{2}\right) = -\frac{F_x\left(3a/2, 3a/2\right)}{F_y\left(3a/2, 3a/2\right)} = -\frac{9a^2/4}{9a^2/4} = -1. \blacksquare$$

注意 2.11.9 (陰関数の存在しない点) $F_y(x_0,y_0)\neq 0$ であれば、(x_0,y_0) の近傍で、$y=\varphi(x)$ の形の陰関数が存在することが保証されるので、その形の陰関数の存在しない可能性がある点は、連立方程式 $F(x,y)=0$, $F_y(x,y)=0$ の解として得られる。実際に解くと、$(x,y)=(0,0), \left(2^{2/3}a, 2^{1/3}a\right)$. この後者は、円 $x^2+y^2=a^2$ の場合の $(\pm a,0)$ のような点であるが、原点 $(0,0)$ の方は、少し様子が違って、どんなに小さな開近傍を取っても、1 つの x に対して $F(x,y)=0$ を満たす y が 3 つ存在したりする。いずれにせよ、$(0,0)$, $\left({}^{2/3}a, 2^{1/3}a\right)$ とも、そのいかなる近傍でも、$y=\varphi(x)$ の形の ($F(x,y)=0$ の) 陰関数は存在しない。 ■

問 2.11.1 $F(x,y):=(x^2+y^2)^2-2(x^2-y^2)$ とおく。

(1) 点 $\left(\dfrac{\sqrt{3}}{2},\dfrac{1}{2}\right)$ の十分小さな開近傍において、$F(x,y)=0$ の陰関数 $y=\varphi(x)$ が存在することを陰関数定理を用いて示せ。(本当は、定理を使わないでも、2 次方程式を解けば陰関数が具体的に求まる。そういう単純な場合で、定理を使う練習をしましょう、ということである。)

(2) $F(a,b)=0$ を満たす点 (a,b) のうちで、陰関数定理の仮定の成立しない点[22]を求めよ。

(3) 曲線 $F(x,y)=0$ 上の点で、その点における接線の傾きが 0 となる点を求めよ。

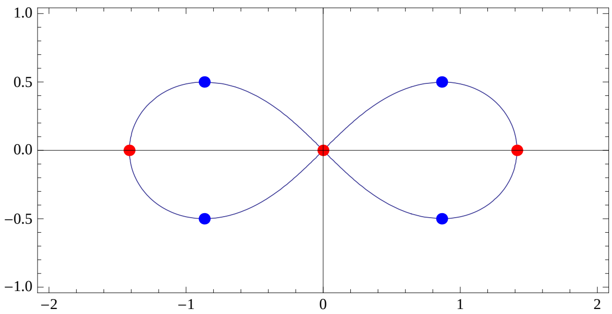


図 2.21: ヤコブ・ベルヌーイのレムニスケート, $(x^2+y^2)^2-2(x^2-y^2)=0$

(p.??を見よ。)

例 2.11.10 連立方程式 $x+y+z+w=0$, $e^x+e^{2y}+e^z+e^w=4$ は、0 の十分小さな開近傍で x, y について解けることを証明せよ。

解答

$$X:=\begin{pmatrix} z \\ w \end{pmatrix},\quad Y:=\begin{pmatrix} x \\ y \end{pmatrix},$$
$$F_1(X,Y):=x+y+z+w,\quad F_2(X,Y):=e^x+e^{2y}+e^z+e^w-4,$$
$$F(X,Y):=\begin{pmatrix} F_1(X,Y) \\ F_2(X,Y) \end{pmatrix}$$

とおくと、$F\colon \mathbf{R}^2\times\mathbf{R}^2 \ni (X,Y)\longmapsto F(X,Y)\in\mathbf{R}^2$ は C^1 級で、

$$\frac{\partial F}{\partial Y}(X,Y)=\begin{pmatrix} \dfrac{\partial F_1}{\partial x} & \dfrac{\partial F_1}{\partial y} \\ \dfrac{\partial F_2}{\partial x} & \dfrac{\partial F_2}{\partial y} \end{pmatrix}=\begin{pmatrix} 1 & 1 \\ e^x & 2e^y \end{pmatrix}.$$

[22]ただし、陰関数としては $y=\varphi(x)$ の形のものを考える ($x=\psi(y)$ の形のものは考えない)。

これから

$$\det \frac{\partial F}{\partial Y}(0,0) = \det \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = 1\cdot 2 - 1\cdot 1 = 1 \neq 0.$$

ゆえに $F(X,Y)=0$ は 0 の近傍で Y について解ける。いいかえると (x,y) について解ける。ついでに

$$\varphi'(X) = -\left(\frac{\partial F}{\partial Y}\right)^{-1}\frac{\partial F}{\partial X} = -\frac{1}{2e^{2y}-e^x}\begin{pmatrix} 2e^{2y} & -1 \\ -e^x & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ e^z & e^w \end{pmatrix}$$

が得られる。■

2.11.5 陰関数、逆関数の高階数導関数

陰関数、逆関数の高階導関数については、次の命題が成り立つ。

命題 2.11.11 (陰関数、逆関数の微分可能性) (1) 陰関数定理で F が C^k 級 $(k\geq 2)$ であれば φ も C^k 級。

(2) 逆関数定理で f が C^k 級 $(k\geq 2)$ であれば f^{-1} も C^k 級。

証明 陰関数定理、逆関数定理における (1 階の) 導関数の公式を眺めると明らかである (以下の例を見よ)。■

高階導関数を実際に計算するには、合成関数の微分法を用いれば良い。陰関数の場合に $k=2$ に対して調べてみよう。まず陰関数定理から、陰関数 φ は C^1 級で

$$\varphi'(x) = -\left(\frac{\partial F}{\partial x}(x,\varphi(x))\right)^{-1}\frac{\partial F}{\partial y}(x,\varphi(x)). \tag{2.14}$$

ここで F が C^2 級という仮定から $\dfrac{\partial F}{\partial x}, \dfrac{\partial F}{\partial y}$ は C^1 級である。また φ は C^1 級であるから、(2.14) の右辺は C^1 級関数の合成関数として C^1 級である。ゆえに φ' が C^1 級となるから φ は C^2 級である。一般の場合もこれと同じことである。

その気になれば、合成関数の微分法に関する定理を用いて、実際に (2.14) の右辺を微分して、φ の 2 階導関数を表す公式を具体的に求められる。$m=n=1$ の場合に実行してみよう。

$$\begin{aligned}
\varphi''(x) &= -\frac{F_y(x,\varphi(x))\dfrac{d}{dx}F_x(x,\varphi(x)) - \dfrac{d}{dx}F_y(x,\varphi(x))F_x(x,\varphi(x))}{F_y(x,\varphi(x))^2} \\
&= -\Big\{F_y(x,\varphi(x))\left[F_{xx}(x,\varphi(x)) + F_{xy}(x,\varphi(x))\varphi'(x)\right] \\
&\qquad - \left[F_{yx}(x,\varphi(x)) + F_{yy}(x,\varphi(x))\varphi'(x)\right]F_x(x,\varphi(x))\Big\} \times \frac{1}{F_y(x,\varphi(x))^2} \\
&= -\frac{F_yF_{xx} + F_yF_{xy}(-F_x/F_y) - F_{yx}F_x - F_{yy}(-F_x/F_y)F_x}{F_y^2} \\
&= -\frac{F_y^2F_{xx} - 2F_{xy}F_xF_y + F_{yy}F_x^2}{F_y^3}.
\end{aligned}$$

なかなか面倒なようだが、例えば極値の判定をするときは $\varphi'(x)=0$, すなわち $F_x(x,\varphi(x))=0$ となる点 x における値のみ興味があるわけで、そういう点では

$$\varphi''(x) = -\frac{F_y^2F_{xx}}{F_y^3} = -\frac{F_{xx}}{F_y} \tag{2.15}$$

とかなりシンプルになる。

例題 2.11.1 方程式
$$xy^2 - x^2y - 2 = 0$$
によって定められる陰関数 y の極値を求めよ。(古い演習書を見ると、この手の問題が載っていたりします。以下のようにして解いても、それだけではどういう陰関数なのか全然分からない、というのが変な感じがしますが…参考まで。)

解 まず与式を微分して

$$y^2 + 2xyy' - 2xy - x^2y' = 0. \tag{2.16}$$

これから

$$\begin{aligned} y' = 0 &\Leftrightarrow y(y-2x) = 0 \\ &\Leftrightarrow y = 2x \quad (y = 0 \text{ は元の式を満たさない}) \\ &\Leftrightarrow x = 1, \quad y = 2. \end{aligned}$$

ところで (2.16) から

$$2yy' + 2yy' + 2x(y')^2 + 2xyy'' - 2y - 2xy' - 2xy' - x^2y'' = 0.$$

よって
$$y'(4y + 2xy' - 4x) + (2xy - x^2)y'' - 2y = 0.$$
ここで $x = 1$, $y = 2$, $y' = 0$ を代入すると $3y'' - 4 = 0$ となるので、

$$y'' = 4/3 > 0.$$

よって極小値である。■

2.11.6 陰関数定理の応用について

陰関数定理は、初めて学ぶ人にとっては、きちんと述べるだけでも大変な定理である。その本質は、いわゆる存在定理であって、ご利益(りやく)が分かりづらいところがある。しかし陰関数定理は多くの重要な応用を持つ。ここでは、多様体、条件付き極値問題、分岐理論[23]を紹介する。

多様体 幾何学の諸理論を展開する場である**多様体**(たようたい) (**manifold**) は (狭い見方をすれば) 曲線や曲面の概念を一般化したものであるが、現代の数学にとって基本的な言語である。その理論の基礎固めをするときに陰関数定理が必要になる。(例えば、局所的に $F = 0$ という方程式の解集合として定義されるものと、 $\operatorname{graph} \varphi$ として定義されるものが同等であることを保証するために使われる。この種の応用のごく簡単な場合を、次項「関数のレベル・セット」で説明する。)

条件つき極値問題 次の 2.12 節で詳しく説明する。

[23]非線形数学の重要なテーマである。

分岐理論 パラメーター λ を含む方程式

$$F(x,\lambda)=0$$

の解 $x=x(\lambda)$ のパラメーター依存性 (特に解の一意性がなくなる場合) を研究するのが**分岐理論**(ぶんきりろん) (**bifurcation theory**) である。陰関数定理が適用できる場合であれば、解の一意性が成立するので、分岐が起るためには、陰関数定理の条件が成立しないことが必要と分かる。

2.11.7 関数のレベル・セット

$$\text{内点 } a \text{ が } f \text{ の極値点} \Longrightarrow a \text{ は } f \text{ の停留点 i.e. } \nabla f(a)=0.$$

という定理の図形的な解釈を、既に ?? で与えておいたが、ここでは、f のレベル・セットとからめた意味付けを補足しておく。

簡単のため、Ω を $\mathbf{R}^2$ の開集合とし、$f\colon\Omega\to\mathbf{R}$ を C^1 級の関数とする。$c\in\mathbf{R}$ に対して

$$L_c:=\{(x,y)\in\Omega; f(x,y)=c\}$$

を f の高さ c の**レベル・セット** (**level set**) あるいは**等高線** (**contour**) という。特に $c=0$ の場合、L_c を f の**零点集合**とも呼び、N_f という記号で表したこともあった。

今 $(a,b)\in\Omega$ を任意に取って、$c:=f(a,b)$ とおく ($(a,b)\in L_c$ なので $L_c\neq\emptyset$ が成り立つ)。既に

> (a,b) から $\nabla f(a,b)$ の方向に移動すると標高が高くなり、$-\nabla f(a,b)$ の方向に移動すると標高が低くなる

ということは分かっている。

「∇f が 0 でなければ、レベル・セット L_c は曲線」 $F(x,y):=f(x,y)-c$ とおき、F について陰関数定理を適用することによって、$\nabla f(a,b)\neq\begin{pmatrix}0\\0\end{pmatrix}$ ならば、(a,b) の十分小さな開近傍 $U\times V$ で、$f(x,y)=c$ は、以下示すように、1 つの変数について解くことができる。

(1) $f_y(a,b)\neq 0$ の場合。y について解ける。すなわち $\mathbf{R}$ の開集合 U, V と C^1 級の関数 $\varphi\colon U\to V$ が存在して、$b=\varphi(a)$,

$$N_F\cap(U\times V)=\operatorname{graph}\varphi:=\{(x,\varphi(x)); x\in U\}.$$

(2) $f_x(a,b)\neq 0$ の場合。x について解ける。すなわち $\mathbf{R}$ の開集合 U, V と C^1 級の関数 $\psi\colon V\to U$ が存在して、$a=\psi(b)$,

$$N_F\cap(U\times V)=\operatorname{graph}\psi\equiv\{(\psi(y),y); y\in V\}.$$

$N_F=L_c$ であることに注意すると、レベル・セット L_c は、(a,b) の十分小さな開近傍で 1 変数関数のグラフ、従って曲線になることが分かる。■

「$\nabla f = 0$ の場合は…」 狭義の極値点 (山や谷) の近傍におけるレベル・セット L_c は「点」である。ちなみに峠点の近傍におけるレベル・セットは、峠点で交わる 2 曲線である[24]。■

同様にして、f が $\mathbf{R}^3$ の開集合 Ω で定義された C^1 級の関数で、$\nabla f \neq 0$ を満たす場合は、f のレベル・セット L_c は、局所的に 2 変数関数のグラフとして表され、特に曲面であることが分かる。

2.11.8 陰関数定理と逆関数定理の証明

ここでは逆関数の定理を証明し、それを利用して陰関数の定理を証明することにする。
後者は簡単なので、先に片付けよう。

逆関数定理を認めた上での陰関数定理の証明 $f\colon \mathbf{R}^{m+n} \supset \Omega \to \mathbf{R}^{m+n}$ を、$f(x,y) := \begin{pmatrix} x \\ F(x,y) \end{pmatrix}$ で定義すると、これは C^1 級で、$f(a,b) = \begin{pmatrix} a \\ F(a,b) \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}$,

$$f'(a,b) = \begin{pmatrix} I & 0 \\ \dfrac{\partial F}{\partial x}(a,b) & \dfrac{\partial F}{\partial y}(a,b) \end{pmatrix}.$$

これから

$$\det f'(a,b) = \det I \cdot \det \frac{\partial F}{\partial y}(a,b) = \det \frac{\partial F}{\partial y}(a,b) \neq 0.$$

ゆえに逆関数定理が適用できて、点 (a,b) を含む開集合 $\widetilde{\Omega} \subset \Omega$ と、点 $f(a,b) = (a,0)$ を含む開集合 W が存在して、 $f|_{\widetilde{\Omega}}\colon \widetilde{\Omega} \to W$ は C^1 級の逆関数 g を持つ。

$\forall (x,y) \in \widetilde{\Omega}$ に対して $g(x,y) = (x, \psi(x,y))$ と書ける。(実際、$(\eta(x,y), \psi(x,y)) := g(x,y)$ とおくと、$(x,y) = f(\eta(x,y), \psi(x,y)) = (\eta(x,y), F(\eta(x,y), \psi(x,y)))$. ゆえに $x = \eta(x,y)$ であるから、$g(x,y) = (x, \psi(x,y))$.)

射影 $\pi\colon \mathbf{R}^m \times \mathbf{R}^n \to \mathbf{R}^n$ を $\pi(x,y) = y$ で定めると、$\psi = \pi \circ g$ と表現できる。これから ψ は C^1 級であることが分かる。

一方 $\pi \circ f = F$ ゆえ、$\forall (x,y) \in W$ に対して

$$F(x, \psi(x,y)) = F(g(x,y)) = (\pi \circ f) \circ g(x,y) = \pi \circ (f \circ g)(x,y) = \pi(x,y) = y.$$

さて a を含む開集合 $\tilde{U}$, b を含む開集合 V を十分小さく取って

$$\tilde{U} \times V \subset \tilde{\Omega}, \quad \tilde{U} \times \{0\} \subset W$$

が成り立つようにする。そして $\tilde{\varphi}\colon \tilde{U} \to \mathbf{R}^m$ を $\tilde{\varphi}(x) = \psi(x,0)$ で定める ($x \in \tilde{U}$ の時 $(x,0) \in \tilde{U} \times \{0\} \subset W = \psi$の定義域 であることに注意)。$\psi$ が C^1 級ゆえ $\tilde{\varphi}$ も C^1 級である。そして $\tilde{\varphi}(a) = b$. 実際 $\varphi(a) = \psi(a,0) = \pi \circ g(a,0) = \pi(a,b) = b$.

$U = \tilde{U} \cap \tilde{\varphi}^{-1}(V)$ とおくと U は a を含む開集合で $\varphi(U) \subset V$.

そして $x \in U$ とすると $\varphi(x) \in \varphi(U) \subset V$. よって $(x, \varphi(x)) \in U \times V \subset \tilde{U} \times V \subset \tilde{\Omega}$. ゆえに $F(x, \varphi(x)) = F(x, \psi(x,0)) = 0$.

---

[24] この事実は、Morse の補題という定理から簡単に証明できる。Morse の補題については、例えば服部晶夫、「いろいろな幾何 II」、岩波書店 (1993) の命題 3.1 や横田一郎、「多様体とモース理論」、現代数学社 (1991) を参照するとよい。

逆に $F(x,y_1)=0$ となったとすると、

$$f(x,y_1)=(x,F(x,y_1))=(x,0)=(x,F(x,\varphi(x)))=f(x,\varphi(x)).$$

$f|_{\tilde\Omega}$ は 1 対 1 ゆえ、$y_1=\varphi(x)$. ■

この逆に、陰関数定理から逆関数を導く論法も紹介しておく (我々の話の筋「逆関数定理を証明し、それから陰関数定理を導く」には必要がないわけだが)。

陰関数定理を認めた上での逆関数定理の証明　講義の話の流れからは必要ないので、アイディアだけ。$F(x,y):=f(x)-y$ により F を定義すると、$\dfrac{\partial F}{\partial x}(x,y)=f'(x)$ であるから、

$$\det\frac{\partial F}{\partial x}(a,b)=\det f'(a)\neq 0.$$

これから F について陰関数定理が適用できて、(a,b) の近傍で $F(x,y)=0$ が x について解けることが分かる。■

それでは逆関数の定理の証明を始めよう。証明には色々な方法があり、解析学の常套手段である「**逐次近似法**」を使う証明は捨てがたいが、準備に手間がかかるので、ここでは「コンパクト集合上の連続関数は最小値を持つ」という定理に持ち込む方法を採用する。

逆関数の定理の証明

1° $A:=f'(a)$, $\tilde f:=A^{-1}\circ f$ とおくと、$(\tilde f)'(a)=I$ (I は単位行列) となる。$\tilde f$ について定理を証明すれば $f=A\circ\tilde f$ について示せたことになる。そこで以下 $f'(a)=I$ と仮定する。

2° **主張 A:** $\exists U$: a を内点として含む閉区間 $\subset \mathbf{R}^n$ s.t.

(2.17) $$\forall x\in U\setminus\{a\}\quad f(x)\neq f(a).$$

(2.18) $$\forall x\in U\quad \det f'(x)\neq 0.$$

(2.19) $$\forall x\in U\quad \|f'(x)-f'(a)\|<\frac{1}{2}.$$

主張 A の証明　f' の連続性により、U を十分小さく取れば (2.19) は成り立つ。同様に $\det f'(a)=\det I=1\neq 0$ に注意すれば、U を十分小さく取れば (2.18) も成り立つ。(2.17) については、まず f が a で微分可能であることから

$$\lim_{x\to a}\frac{\|f(x)-f(a)-f'(a)(x-a)\|}{\|x-a\|}=0.$$

特に $\exists\varepsilon>0$ s.t.

$$0<\|x-a\|<\varepsilon\quad\Longrightarrow\quad\frac{\|f(x)-f(a)-f'(a)(x-a)\|}{\|x-a\|}<\frac{1}{2}.$$

ところが $f(x)=f(a)$ とすると

$$\frac{\|f(x)-f(a)-f'(a)(x-a)\|}{\|x-a\|}=\frac{\|0-I(x-a)\|}{\|x-a\|}=\frac{\|x-a\|}{\|x-a\|}=1.$$

ゆえに $0<\|x-a\|<\varepsilon$ ならば $f(x)\neq f(a)$ が成り立つ。■

3° **主張 B:**

(2.20) $$\forall x_1,x_2\in U\quad \|x_1-x_2\|\le 2\,\|f(x_1)-f(x_2)\|.$$

(これから $f|_U$ の単射性はすぐ分かるし、後述の逆写像が連続であることの証明の鍵となる。)

主張 B の証明　$g(x) := f(x) - x$ とおくと

$$g'(x) = f'(x) - I = f'(x) - f'(a)$$

であるから

$$\|g(x_1) - g(x_2)\| \le \sup_{\xi\in U} \|g'(\xi)\| \, \|x_1 - x_2\| = \sup_{\xi\in U} \|f'(\xi) - f'(a)\| \, \|x_1 - x_2\| \le \frac{\|x_1 - x_2\|}{2}.$$

すなわち

$$\|f(x_1) - f(x_2) - (x_1 - x_2)\| \le \frac{1}{2} \|x_1 - x_2\| .$$

ゆえに

$$\|x_1 - x_2\| - \|f(x_1) - f(x_2)\| \le \frac{1}{2} \|x_1 - x_2\| .$$

移項して両辺を 2 倍すれば、(2.20) を得る。■

4° $B := U^b$ (U の境界), $d := \inf_{y\in f(B)} \|y - f(a)\|$ とおくと $d > 0$. 実際

- (2.17) より $f(a) \notin f(B)$.
- B は $\mathbf{R}^n$ の有界閉集合で、コンパクトであるから、連続写像 f による像 $f(B)$ もコンパクトで、特に $f(B)$ は閉集合である。
- 「閉集合とそれに属さない点との距離は正である」

であるから[25]。さて $W := B(f(a); d/2)$ とおくと

$$y \in W, \quad x \in B \Longrightarrow \|y - f(a)\| < \|y - f(x)\| . \tag{2.21}$$

(図を描くことを勧める) 実際、まず W の定義から

$$\|y - f(a)\| < \frac{d}{2},$$

一方 $x \in B$ より

$$\|f(x) - f(a)\| \ge \inf_{y\in f(B)} \|y - f(a)\| = d$$

であるから

$$\begin{aligned}\|f(x) - y\| &= \|f(x) - f(a) + f(a) - y\| \ge \|f(x) - f(a)\| - \|f(a) - y\| \\ &> d - \frac{d}{2} = \frac{d}{2} > \|y - f(a)\| .\end{aligned}$$

5° **主張 C:**

$$\forall y \in W \quad \exists! \overline{x} \in U \setminus B \quad \text{s.t. } f(\overline{x}) = y.$$

[25] (初等的な証明) $d = 0$ とすると、$\exists \{y_n\}_{n\in\mathbf{N}}$ s.t. (i) $\forall n \in \mathbf{N}$ $y_n \in f(B)$, (ii) $\|y_n - f(a)\| \to 0$. $f(B)$ が閉集合であるから、$f(a) \in f(B)$ だが、これは $f(a) \notin f(B)$ に矛盾する。

主張の C 証明　関数 $h\colon U \to \mathbf{R}$ を

$$h(x) := \|y - f(x)\|^2 \equiv (y - f(x), y - f(x))$$

で定義する。これはコンパクト集合 U 上の連続関数であるから、最小値 $h(\overline{x}), \overline{x} \in U$ を取る。ところで (2.21) より

$$x \in B \Longrightarrow h(a) < h(x).$$

ゆえに $\overline{x} \not\in B$ i.e. $\overline{x} \in U^\circ$. ゆえに h は内点 $\overline{x}$ で最小値を取ることになり、$\nabla h(\overline{x}) = 0$. $\nabla h(\overline{x}) = f'(\overline{x})^T(f(\overline{x}) - y)$ であり、(2.18) より $f'(\overline{x})$ は正則ゆえ $f(\overline{x}) - y = 0$. すなわち $f(\overline{x}) = y$. $\overline{x}$ の一意性は (2.20) から分かる。

6° ここで

$$V := (U \setminus B) \cap f^{-1}(W)$$

とおくと V は a の開近傍である。(実際 W は開球であるから開集合であり、連続写像 f による逆像 $f^{-1}(W)$ は開集合である。$U \setminus B$ は U の内部であるから、もちろん開集合である。2 つの開集合の交わりであるから、W は開集合である。

一方、$a \in U$, $a \neq B$ は明らかで、$f(a) \in W = B(f(a); d/2)$ より $a \in f^{-1}(W)$ であるから、$a \in V$.) 前項から

$$f|_V\colon V \longrightarrow W$$

は逆関数 $f^{-1}\colon W \to V$ を持つ (本当は ${f_V}^{-1}$ と書くべきであるが、繁雑になるので、以下この証明中では単に f^{-1} と書く)。

7° f^{-1} は連続である。実際 (2.20) より $y_1, y_2 \in W$ とするとき

$$\left\|f^{-1}(y_1) - f^{-1}(y_2)\right\| \le 2\left\|y_1 - y_2\right\|$$

であるから。

8° **主張 D:** $\forall x \in V$ に対して、f^{-1} は $y := f(x)$ で微分可能で

$$(f^{-1})'(y) = \left(f'(x)\right)^{-1}.$$

主張 D の証明　$x_0 \in V$ に対して、$A := f'(x_0)$ とおく。微分可能性の定義から

$$f(x) - f(x_0) = A(x - x_0) + \varepsilon(x) \tag{2.22}$$

とおくと

$$\lim_{x \to x_0} \frac{\|\varepsilon(x)\|}{\|x - x_0\|} = 0.$$

さて $\forall y \in W$ に対して $x := f^{-1}(y)$ とおくと $x \in V$ で $f(x) = y$. それで (2.22) の両辺に A^{-1} をかけ、y_0, y で書き直すと

$$A^{-1}(y - y_0) = f^{-1}(y) - f^{-1}(y_0) + A^{-1}\varepsilon(f^{-1}(y)).$$

ゆえに

$$f^{-1}(y) - f^{-1}(y_0) = A^{-1}(y - y_0) - A^{-1}\varepsilon(f^{-1}(y)).$$

そこで次のことを示せばよい。

$$\lim_{y \to y_0} \frac{\|A^{-1}\varepsilon(f^{-1}(y))\|}{\|y - y_0\|} = 0.$$

これを示すには

$$\lim_{y\to y_0}\frac{\|\varepsilon(f^{-1}(y))\|}{\|y-y_0\|}=0$$

を示せばよい。

$$\frac{\|\varepsilon(f^{-1}(y))\|}{\|y-y_0\|}=\frac{\|\varepsilon(f^{-1}(y))\|}{\|f^{-1}(y)-f^{-1}(y_0)\|}\cdot\frac{\|f^{-1}(y)-f^{-1}(y_0))\|}{\|y-y_0\|}.$$

f^{-1} の連続性より、$y\to y_0$ のとき $f^{-1}(y)\to f^{-1}(y_0)$. よって右辺の第 1 因子 $\to 0$. 一方第 2 因子は、第 6° より 2 で押さえられる。

9° f^{-1} が C^1 級であること。f^{-1} のヤコビ行列 $(f^{-1})'(y)$ は $f'(x)$ の逆行列であり、成分は Cramer の公式から、分母が $\det f'(x)$, 分子は $\dfrac{\partial f_i}{\partial x_j}(x)$ の多項式として表現できる。これは y の関数として見て連続である。ゆえに f^{-1} は C^1 級である。■

2.12 条件付き極値問題 (Lagrange の未定乗数法)

2.12.1 2 変数の場合

まず 2 変数関数の場合に説明する。

これまで扱った極値問題では、定義域が基本的には開集合であった。すると

$$\boxed{\text{内点 } a \text{ が } f \text{ の極値点} \Longrightarrow a \text{ は } f \text{ の停留点}, \quad \text{i.e. } \nabla f(a)\equiv\begin{pmatrix}\dfrac{\partial f}{\partial x}(a)\\ \dfrac{\partial f}{\partial y}(a)\end{pmatrix}=0.}$$

という命題が成り立ち、極値点を探すことは比較的簡単であった。以下では、関数 f の極値を条件

$$g(x,y)=0 \tag{2.23}$$

の下で求めることを考える。すなわち g の零点集合

$$N_g:=\{(x,y);g(x,y)=0\}$$

に f を制限して考える。これは普通、開集合にはならない。よって、f の停留点 ($f'(a)=0$ となる点 a のこと) を探しても意味がない。

例 2.12.1 条件 $g(x,y):=x^2+y^2-1=0$ の下で

$$f(x,y):=x^2+3xy+y^2$$

の最大値、最小値を求めよ。

解答の方針 条件 $g(x,y)=0$ を $y=\pm\sqrt{1-x^2}$ と y について解いて代入して、x のみの関数についての最大最小問題に直せば解ける。ここで陰関数が出て来ていることに気がつくだろうか？(余談: これは2次形式なので、2次形式の取り扱いに詳しければ、微積分を一切使わないで解くことも可能である。また、$g=0$ は単位円周で、三角関数を使ってパラメーター付けできるから、それを使って1変数関数の最大最小問題にすることも出来る。) ■

例 2.12.2 平面 $\mathbf{R}^2$ 内の曲線 $y=x^2$ 上の点と点 $(0,1)$ との距離の極値を求めよ。これは

$$\begin{cases} f(x,y) := \sqrt{x^2+(y-1)^2} \\ g(x,y) := x^2 - y \end{cases}$$

とする条件付き極値問題である。

この場合、$g(x,y)=0$ は $y=x^2$ と解ける。これを f に代入して出来る

$$h(x) := f(x,x^2) = \sqrt{x^2+(x^2-1)^2}$$

についての普通の極値問題に帰着できる。■

陰関数定理の節で述べたように、すべての方程式 $g(x,y)=0$ から陰関数 $y=\varphi(x)$ が具体的に求まるとは限らないので、上の 2 つの例のような解き方は、運が良くない限り出来ないわけである。そこで、次の定理の出番となる。

定理 2.12.3 (条件つき極値問題に対する Lagrange の未定乗数法, Lagrange (1788 年)) Ω を $\mathbf{R}^2$ の開集合、f と g を Ω で定義され $\mathbf{R}$ に値を持つ C^1 級の関数として、

$$N_g := \{(x,y)\in\Omega;\ g(x,y)=0\}$$

とおいたとき

$$\nabla g \neq 0 \quad \text{on } N_g$$

が成り立つとする。また条件 $g(x,y)=0$ の下で、f は $a=\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\in N_g$ で極値を取るとする (すなわち、$\exists\varepsilon>0$ s.t. $f(a)=\max\{f(x); x\in B(a;\varepsilon)\cap N_g\}$ または $f(a)=\min\{f(x); x\in B(a;\varepsilon)\cap N_g\}$)。このとき $\exists\lambda\in\mathbf{R}$ s.t.

$$\nabla f(a) = \lambda\nabla g(a)$$

i.e.

$$\frac{\partial f}{\partial x}(a) = \lambda\frac{\partial g}{\partial x}(a),$$
$$\frac{\partial f}{\partial y}(a) = \lambda\frac{\partial g}{\partial y}(a).$$

注意 2.12.4 (1) λ のことを **Lagrange の未定乗数** (**Lagrange multiplier**) という。

(2) 極値点の座標 α, β, 未定乗数 λ は、連立方程式

$$\begin{cases} g(\alpha,\beta)=0 \\ \dfrac{\partial f}{\partial x}(\alpha,\beta) = \lambda\dfrac{\partial g}{\partial x}(\alpha,\beta) \\ \dfrac{\partial f}{\partial y}(\alpha,\beta) = \lambda\dfrac{\partial g}{\partial y}(\alpha,\beta) \end{cases}$$

の解であり、これから具体的に求まる場合が多い (未知数 3 個、方程式 3 個)。この場合は条件付き極値問題が解けるわけである。この方法を **Lagrange の未定乗数法**と呼ぶ。この条件は

$$F(x,y,\lambda) := f(x,y) - \lambda g(x,y)$$

とおくと

$$F_\lambda = F_x = F_y = 0 \quad \text{i.e. } \nabla F = 0$$

と書ける。ここで ∇ は (x, y, λ) に関する勾配 (gradient) を表す。この形で定理を述べている本も多い。■

証明 仮定

$$\nabla g(a) \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

より

$$\frac{\partial g}{\partial x}(a) \neq 0 \quad \text{or} \quad \frac{\partial g}{\partial y}(a) \neq 0.$$

(i) $\dfrac{\partial g}{\partial y}(a) \neq 0$ の場合. 陰関数の定理から、点 a の近傍で

$$g(x, y) = 0 \Longleftrightarrow y = \varphi(x)$$

と y について解けて

$$\varphi'(x) = -\frac{\dfrac{\partial g}{\partial x}(x, \varphi(x))}{\dfrac{\partial g}{\partial y}(x, \varphi(x))}$$

となる。そこで

$$h(x) := f(x, \varphi(x))$$

とおくと

$$\begin{aligned} h'(x) &= \frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x))\varphi'(x) \\ &= \frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x)) \left(-\frac{\dfrac{\partial g}{\partial x}(x, \varphi(x))}{\dfrac{\partial g}{\partial y}(x, \varphi(x))} \right). \end{aligned}$$

h は α で極値となるから $h'(\alpha) = 0$. $(\alpha, \varphi(\alpha)) = (\alpha, \beta) = a$ に注意して

$$\frac{\partial f}{\partial x}(a) - \frac{\dfrac{\partial f}{\partial y}(a)}{\dfrac{\partial g}{\partial y}(a)} \cdot \frac{\partial g}{\partial x}(a) = 0.$$

ここで λ を

$$\lambda = \frac{\frac{\partial f}{\partial y}(a)}{\frac{\partial g}{\partial y}(a)} \quad \text{i.e.} \quad \frac{\partial f}{\partial y}(a) - \lambda \frac{\partial g}{\partial y}(a) = 0$$

とおくと

$$\frac{\partial f}{\partial x}(a) - \lambda \frac{\partial g}{\partial x}(a) = 0.$$

まとめると

$$\nabla f(a) = \lambda \nabla g(a).$$

(ii) $\dfrac{\partial g}{\partial x}(a) \neq 0$ の場合. 今度は $g(x, y) = 0$ を x について解けばよい。後は同様である。■

問 2.12.1 上の定理は 2 変数関数に関するものだが、n 変数関数に一般化して、証明せよ。

極値の条件の図形的な解釈　$c=f(a)$ に対して、f のレベル・セット $L_c=\{(x,y)\in\Omega; f(x,y)=c\}$ を考える。条件

$$\exists\lambda\in\mathbf{R}\quad \text{s.t.}\quad \nabla f(a)=\lambda\nabla g(a)$$

は、$N_g=\{(x,y); g(x,y)=0\}$ と L_c が接することを意味する ($\nabla f(a)$ は L_c の法線ベクトル、$\nabla g(a)$ は N_g の法線ベクトルだから)。こうなることは次のように考えても納得できる。

点 (x,y) が N_g に沿って、$a=(\alpha,\beta)$ から動くとき、$g(x,y)=g(\alpha,\beta)=0$ であるから

$$0=g(x,y)-g(\alpha,\beta)\fallingdotseq\left(\nabla g(a),\begin{pmatrix}x-\alpha\\ y-\beta\end{pmatrix}\right).$$

f は $a=(\alpha,\beta)$ で極値を取るので、$(x,y)\in N_g$ が (α,β) に近いとき、$f(x,y)$ は $f(\alpha,\beta)$ に非常に近いと考えられる。ゆえに

$$0\fallingdotseq f(x,y)-f(\alpha,\beta)\fallingdotseq(\nabla f(a),\begin{pmatrix}x-\alpha\\ y-\beta\end{pmatrix}).$$

よって $\nabla f(\alpha,\beta)$ と $\nabla g(\alpha,\beta)$ は、ともに $\begin{pmatrix}x-\alpha\\ y-\beta\end{pmatrix}$ に直交する。これから

$$\nabla g(\alpha,\beta)\parallel\nabla f(\alpha,\beta).$$

ゆえに f の等高線 L_c と N_g は接している。■

2.12.2　n 変数, d 個の制約条件の場合

一般の n 変数関数で、制約条件も複数の場合に拡張した定理を (証明抜きで) 掲げておこう。

定理 2.12.5 (Lagrange の未定乗数法, n 変数で制約条件 d 個の場合) Ω を $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$ を C^1 級の関数、d を $1\le d<n$ なる自然数、

$$\boldsymbol{g}=\begin{pmatrix}g_1\\ \vdots\\ g_d\end{pmatrix}:\Omega\longrightarrow\mathbf{R}^d$$

を C^1 級の関数で

$$\operatorname{rank}\boldsymbol{g}'(x)=d\quad(\forall x\in\Omega)$$

を満たすもの、$a\in\Omega$ で $\boldsymbol{g}(a)=0$, f は条件 $\boldsymbol{g}=0$ の下で a で極値を取る、とすると
$\exists\boldsymbol{\lambda}=\begin{pmatrix}\lambda_1\\ \vdots\\ \lambda_d\end{pmatrix}\in\mathbf{R}^d$ s.t.

$$\nabla f(a)=\sum_{j=1}^d\lambda_j\nabla g_j(a),\quad \text{また (もちろん)}\quad \boldsymbol{g}(a)=0. \tag{2.24}$$

注意 2.12.6 やはり

$$F(x,\boldsymbol{\lambda}):=f(x)-\sum_{j=1}^d\lambda_j g_j(x)$$

とおくと、方程式は (2.24) は次と同値になる:

$$\begin{cases} \dfrac{\partial F}{\partial x_i}(a, \boldsymbol{\lambda}) = 0 & (i = 1, \cdots, n) \\ \dfrac{\partial F}{\partial \lambda_j}(a, \boldsymbol{\lambda}) = 0 & (j = 1, \cdots, d). \end{cases}$$

(これは、F のすべての変数についての gradient が 0 と書けることに注意。) ■

2.12.3 例題

Lagrange の未定乗数法の例を二つほどあげる。いずれも意味が明らかな (高校数学でも答が出る) 問題である。

例題 2.12.1 方程式 $ax + by + c = 0$ ($(a, b) \in \mathbf{R}^2 \setminus \{0\}$, $c \in \mathbf{R}$) で表される平面内の曲線を L とする。点 (x, y) が直線 L 上を動くときの、関数 $f(x, y) = x^2 + y^2$ の最小値を求めよ。

解答 (求めるものは原点と直線 L との距離の平方になることは (直観的にすぐ) 分かるだろうから、微分法を用いなくても「解ける」問題であるが、Lagrange の未定乗数法で求めてみる。)

1. 最小値が存在することの証明 (この問題の場合は、図形的な意味が分かるので「明らか」であるが、そうでない場合もあるので、きちんと書くとどうなるか、紹介する意味で以下に示す。実は良く出て来る論法である。) L 上の点 (x_0, y_0) を一つ取り (存在することは自明)、正数 R を $R^2 = x_0^2 + y_0^2$ で定め、$D := \{(x, y) \in \mathbf{R}^2; x^2 + y^2 \le R^2\}$ とおく。L を

$$L = L \cap \mathbf{R}^2 = L \cap (D \cup D^c) = (L \cap D) \cup (L \cap D^c)$$

と分解すると、 $L \cap D$ は $\mathbf{R}^2$ の空でない有界閉集合であるから、関数 f は $L \cap D$ において最小値 $m = f(\alpha, \beta)$ を持つ。ところで $(\alpha, \beta) \in D$ であるから、$m = f(\alpha, \beta) = \alpha^2 + \beta^2 \le R^2$. 一方、$L \cap D^c$ においては、$f(x, y) = x^2 + y^2 > R^2$ であるから、m は f の L 全体における最小値であることが分かる。

2. 唯一の極値は最小値である 前段で最小値が存在することが分かったが、最小値は極値であるから、もしも極値が一つしか無いことが分かれば、それが最小値である。

3. f の条件付き極値を求める 関数 $g \colon \mathbf{R}^2 \to \mathbf{R}$ を $g(x, y) := ax + by + c$ で定義すると、

$$\nabla g(x, y) = \begin{pmatrix} a \\ b \end{pmatrix} \ne 0.$$

したがって、条件 $g(x, y) = 0$ の下での f の極値点は (もし存在するならば) Lagrange の未定乗数法で求まる。未定乗数を λ とおくと、方程式は、

$$\begin{aligned} 0 &= f_x(x, y) - \lambda g_x(x, y), \\ 0 &= f_y(x, y) - \lambda g_y(x, y), \\ 0 &= g(x, y). \end{aligned}$$

これは

$$
\begin{aligned}
0 &= 2x - \lambda a, \\
0 &= 2y - \lambda b, \\
0 &= ax + by + c
\end{aligned}
$$

となるから、解は

$$\lambda = -\frac{2c}{a^2+b^2}, \quad x = -\frac{ac}{a^2+b^2}, \quad y = -\frac{bc}{a^2+b^2}$$

ただ一つだけである。

このように Lagrange の未定乗数法で求められた点が極値点であるかどうかは、一般にはすぐには分からないが、この場合は前段の議論から、これは極値点であり、さらには最小点に他ならないことが分かる[26]。すなわち

$$(x, y) = \left(\frac{-ac}{a^2+b^2}, \frac{-bc}{a^2+b^2}\right)$$

のとき、f は最小値

$$f\left(\frac{-ac}{a^2+b^2}, \frac{-bc}{a^2+b^2}\right) = \frac{c^2}{a^2+b^2}$$

を取る。 ■

問 2.12.2 直線 $ax + by + c = 0$ $((a, b) \neq (0, 0))$ と点 (x_0, y_0) との距離は、$\dfrac{|ax_0 + by_0 + c|}{\sqrt{a^2+b^2}}$ であることを示せ。

問 2.12.3 平面 $ax + by + cz + d = 0$ $((a, b, c) \neq (0, 0, 0))$ と点 (x_0, y_0, z_0) との距離は、$\dfrac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2+b^2+c^2}}$ であることを示せ。

例題 2.12.2 方程式 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$ (a, b は正の定数) 表される平面内の楕円を E とする。点 (x, y) が直線 E 上を動くときの、関数 $f(x, y) = x + y$ の最大値、最小値を求めよ。

解答 これも図形的に考えると意味は明瞭で、Lagrange の未定乗数法を講義しなかった年度にこの問題の 3 次元版を期末試験に出したことがある (接平面をきちんと求めて、使いこなせるかというのが、出題のねらい)。

まず E は有界閉集合であるから、連続関数 f は E 上で最大値、最小値を持つことが分かる。また

$$g(x, y) := \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$$

とおくと、

$$\nabla g(x, y) = \begin{pmatrix} 2x/a^2 \\ 2y/b^2 \end{pmatrix}$$

であり、$g(x, y) = 0$ を満たす任意の (x, y) に対して

$$\nabla g(x, y) \neq 0$$

[26]「犯人は確かに存在し、この部屋の中にいる」、「(もし存在するならば) 犯人は男性である」、「この部屋の中に男性は一人だけいる」ならば、この部屋にいる唯一の男性が犯人である。

であることが分かる。ゆえに条件 $g(x,y)=0$ の下での関数 f の極値は、Lagrange の未定乗数法で求まる。方程式は

$$\begin{aligned}0&=f_x(x,y)-\lambda g_x(x,y),\\0&=f_y(x,y)-\lambda g_y(x,y),\\0&=g(x,y).\end{aligned}$$

これは

$$\begin{aligned}0&=1-\lambda\frac{2x}{a^2},\\0&=1-\lambda\frac{2y}{b^2},\\0&=\frac{x^2}{a^2}+\frac{y^2}{b^2}-1\end{aligned}$$

であり、

$$(x,y,\lambda)=\pm\left(\frac{a^2}{\sqrt{a^2+b^2}},\frac{b^2}{\sqrt{a^2+b^2}},\frac{\sqrt{a^2+b^2}}{2}\right)\quad(\text{複号同順}).$$

$$f\left(\pm\frac{a^2}{\sqrt{a^2+b^2}},\pm\frac{b^2}{\sqrt{a^2+b^2}}\right)=\pm\sqrt{a^2+b^2}\quad(\text{複号同順}).$$

f が最大値、最小値を持つことは既に分かっているから、これらがその最大値、最小値に他ならない。すなわち、$(x,y)=(a^2/\sqrt{a^2+b^2},b^2/\sqrt{a^2+b^2})$ のとき最大値 $\sqrt{a^2+b^2}$, $(x,y)=(-a^2/\sqrt{a^2+b^2},-b^2/\sqrt{a^2+b^2})$ のとき最小値 $-\sqrt{a^2+b^2}$. ■

問 2.12.4 $f(x,y):=x+y$, $g(x,y):=\dfrac{x^2}{4}-y^2-1$, $N_g:=\{(x,y)\in\mathbf{R}^2;g(x,y)=0\}$ とする。

(1) N_g の概形を描け。

(2) N_g 上の点 $(2\sqrt{2},1)$ における接線の方程式を求めよ。

(3) Lagrange の未定乗数法により、N_g 上での f の極値の候補をすべて求めよ。

(4) N_g 上での f の値の範囲を求めよ。

問 2.12.5 (最短距離は垂線で実現される) $\mathbf{R}^3$ の開集合 Ω で定義された C^1 級関数 $g\colon\Omega\to\mathbf{R}$ が、$\nabla g(x,y,z)\neq 0$ ($(x,y,z)\in\Omega$) を満たすとする。また $P(a,b,c)$ は $\mathbf{R}^3$ 内の定点とする。このとき $N_g:=\{(x,y,z)\in(x,y,z)\in\Omega;g(x,y,z)=0\}$ は曲面となるが、N_g 上の点 $Q(x_0,y_0,z_0)$ で、P からの距離が最小となるものが存在するならば、それは P から N_g に下ろした垂線の足であることを示せ。

(**注意**　多くの場合に「最短距離＝垂線の長さ」が成り立つことを知っていると思うが、これは上に示すような形で (かなり一般に) 成り立ち、証明も出来る、ということである。難しいことを問われているようだが、$\overrightarrow{PQ}=(x_0-a,y_0-b,z_0-c)$ が、法線ベクトル $\nabla g(x_0,y_0,z_0)$ と平行ということで、やってみるとすごく簡単である。) ■

問 2.12.6 (相加平均と相乗平均) n を任意の自然数とする。n 個の任意の正数 x_1, x_2, $\dots$, x_n に対して、不等式

$$\frac{x_1+x_2+\cdots+x_n}{n}\geq\sqrt[n]{x_1x_2\dots x_n}$$

が成り立ち、等号が成立するためには

$$x_1 = x_2 = \cdots = x_n$$

が必要十分であることを示せ (相加平均 $\geq$ 相乗平均)。

(**注意**　これは凸関数の性質を用いて証明するのが簡単であるが、Lagrange の未定乗数法によって証明することも出来る。) ■

問 2.12.7 (対称行列の対角化) スピヴァックの有名な教科書 [8] に載っている次の問題 (**5-17**) の (a) を解け。(普通、固有値は特性多項式の根として特徴づけられて存在証明されるが、ここでは固有値と固有ベクトルを、条件付きの最大値問題の解として得ようということである。いわゆる Rayleigh の原理の基礎となる事実。)

$T\colon \mathbf{R}^n \to \mathbf{R}^n$ を対称線型変換, $A = (a_{ij})$ を T の行列とする ($a_{ij} = a_{ji}$).

(a) $f(x) = \langle Tx, x\rangle = \sum a_{ij}x^i x^j$ に対し，$D_k f(x) = 2\sum_{j=1}^{n} a_{kj}x^j$ であることを示せ．$S^{n-1} = \{x \in \mathbf{R}^n; |x| = 1\}$ 上での f の最大値を考えることにより，$Tx = \lambda x$ となる $x \in S^{n-1}$ および $\lambda \in \mathbf{R}$ が存在することを示せ．

(b) この x に対し，$V = \{x \in \mathbf{R}^n; \langle x, y\rangle = 0\}$ と置くとき，$T(V) \subset V$ および $T\colon V \to V$ が対称線型変換であることを示せ．

(c) T の固有ベクトルから成る $\mathbf{R}^n$ の基底が存在することを示せ．

(b), (c) は多くの線形代数のテキストに書いてあるので探せば見つかると思う。自力で (b), (c) を解いてみたくなった人のために: 対称線型変換の定義については、やはり問題 (**4-11**) 中で定義されている。

T を V 上の内積，$f\colon V \to V$ を線型変換とする．$x, y \in V$ に対して $T(x, f(y)) = T(f(x), y)$ が成り立つとき，f を T に関する**対称変換**と言う．$v_1, \cdots, v_n$ が T に関する正交基底 (正規直交基底のこと)，この基底に関する f の行列 $A = (a_{ij})$ が対称行列 (すなわち $a_{ij} = a_{ji}$) であることを示せ．

2.13 演習問題解答

問 2.1.1 (p. 27) の解答　仮定を ε-δ 論法で表現する。$(\forall x \in \Omega)$ を $(\forall x \in \Omega')$ で置き換えることが出来る。■

問 2.1.2 (p. 28) の解答　順に 1, 0, 極限は存在しない。

問 2.1.3 (p. 28) の解答　$h(x) := \dfrac{g(x)}{f(x)}$, $A := \lim_{x\to a} h(x)$ とおく。$\dfrac{g(x)}{f(x)} = h(x)$ より $g(x) = f(x)h(x)$ であるから、

$$\lim_{x\to a} g(x) = \lim_{x\to a} f(x)h(x) = \lim_{x\to a} f(x) \lim_{x\to a} h(x) = 0 \cdot A = 0. ■$$

問 2.2.1 (p. 30) の解答

(1) $1, \sqrt{2}, \log 3, \pi/4, e^5, 6 \in \mathbf{R}$ であるので $f(x,y) \in \mathbf{R}[x,y]$ である。ゆえに $f\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。

(2) $F(x,y) := 3x+2y+1$, $G(z) := \exp z$ とおく。$F(x,y) \in \mathbf{R}[x,y]$ であるから、$F\colon \mathbf{R}^2 \ni (x,y) \mapsto F(x,y) \in \mathbf{R}$ は $\mathbf{R}^2$ 全体で連続である。また $G\colon \mathbf{R} \ni z \mapsto G(z) \in \mathbf{R}$ は連続である。ゆえにそれらの合成である $g = G \circ F\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。

(3) $Q(x,y) := x^2+2x+3$, $P(x,y) := x^2+y^2+1$ とおくと、$P(x,y), Q(x,y) \in \mathbf{R}[x,y]$ である。ゆえに $Q\colon \mathbf{R}^2 \to \mathbf{R}$ と $P\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。また $\forall (x,y) \in \mathbf{R}^2$ に対して $P(x,y) \geq 1$ であるから、$P(x,y) \neq 0$. ゆえに $h = \dfrac{Q}{P}\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。

$Q(x) := x^2+2x+3$ として、1 変数多項式とするのではないことに注意する。

(4) $f(x,y) := x^2+y^2$ とおくと、$f(x,y) \in \mathbf{R}[x,y]$ であるから、$f\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。一方 $f(\mathbf{R}^2) = [0,\infty)$ である。$g\colon [0,\infty) \to \mathbf{R}$ を $g(z) := \sqrt{z}$ で定めると、g は連続である。ゆえに合成関数 $g \circ f\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。また $h\colon \mathbf{R}^2 \ni (x,y) \mapsto 1 \in \mathbf{R}$ は定数関数だから連続である。ゆえに $F = h + g \circ f\colon \mathbf{R}^2 \ni (x,y) \mapsto 1 + \sqrt{x^2+y^2} \in \mathbf{R}$ は連続である。そして $F(\mathbf{R}^2) = [1,\infty)$. 対数関数 $G\colon (0,\infty) \ni z \mapsto \log z \in \mathbf{R}$ は連続である。$F(\mathbf{R}^2) \subset (0,\infty)$ であるから、G と F は合成可能で、$\varphi = G \circ F\colon \mathbf{R}^2 \ni (x,y) \mapsto \log(1+\sqrt{x^2+y^2}) \in \mathbf{R}$ は連続である

(5) $F_1(x,y) := x^3 - 3xy^2 \in \mathbf{R}$, $F_2(x,y) := 3x^2y - y^3 \in \mathbf{R}$ とおくと、$F_1(x,y), F_2(x,y) \in \mathbf{R}[x,y]$ であるから、関数 $F_1\colon \mathbf{R}^2 \to \mathbf{R}$ と $F_2\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。ゆえに $F = \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}\colon \mathbf{R}^2 \ni (x,y) \mapsto \begin{pmatrix} x^3-3xy^2 \\ 3x^2y-y^3 \end{pmatrix} \in \mathbf{R}^2$ は連続である。■

問 2.2.2 (p. 33) の解答

(1) $f(x,y) := 3x^2+4xy+5y^2$ は多項式関数であるから、$f\colon \mathbf{R}^2 \to \mathbf{R}$ は連続である。特に $(1,2)$ で連続であるから、$(x,y) \to (1,2)$ のときの極限は、$f(1,2)$ に等しい:

$$\lim_{(x,y)\to(1,2)} f(x,y) = f(1,2) = 3\cdot 1^2 + 4\cdot 1\cdot 2 + 5\cdot 2^2 = 3+8+20 = 31.$$

(2) $f(x,y) := \dfrac{2+3xy}{4x^2+5y^2}$ は有理式で、その分母が 0 にならない範囲 $\Omega := \mathbf{R}^2 \setminus \{(0,0)\}$ に対して、$f\colon \Omega \to \mathbf{R}$ が定義されて連続である。特に $(0,1) \in \Omega$ で連続であるから、$(x,y) \to (0,1)$ のときの極限は、$f(0,1)$ に等しい:

$$\lim_{(x,y)\to(0,1)} f(x,y) = f(0,1) = \frac{2+3\cdot 0\cdot 1}{4\cdot 0^2 + 5\cdot 1^2} = \frac{2}{5}.$$

(3) f は有理関数で、$(0,0)$ で分母が 0 になることに注意する。$(x,y) \to (0,0)$ のとき 分母 $= x^2+y^2 \to +0$, 分子 $= 1 \not\to 0$ であるから、極限は存在しない。

(別解) $\Omega := \mathbf{R}^2 \setminus \{(0,0)\}$, $\widetilde{f}\colon \Omega \to \mathbf{R}$, $\widetilde{f}(x,y) = x^2+y^2$ ($(x,y) \in \Omega$), $a = (0,0)$ とおくと、$\widetilde{f}\colon \Omega \to \mathbf{R}$, $a \in \overline{\Omega}$, $\widetilde{f} > 0$ on Ω, $\displaystyle\lim_{(x,y)\to a} \widetilde{f}(x,y) = 0$ であるから、命題 2.1.6 (2) によって、

$$\lim_{(x,y)\to(0,0)} f(x) = \lim_{(x,y)\to a} \frac{1}{\widetilde{f}(x,y)} = \infty.$$

(4) $(x,y)\to(0,0)$ のとき、分子 $=x+y\to 0$, $x^2+y^2\to +0$, 分母 $=\log(x^2+y^2)\to-\infty$ であるから、$\dfrac{0}{-\infty}$ で、

$$\lim_{(x,y)\to(0,0)}\frac{x+y}{\log(x^2+y^2)}=0.$$

もう少し丁寧にやると、

$$\lim_{(x,y)\to(0,0)}\frac{x+y}{\log\left(x^2+y^2\right)}=\lim_{(x,y)\to(0,0)}(x+y)\times\lim_{(x,y)\to(0,0)}\frac{1}{\log\left(x^2+y^2\right)}=0\cdot 0=0.$$

第 2 因子が 0 であることは、命題 2.1.6 (1) や、合成関数の極限を用いて (例えば $r:=\sqrt{x^2+y^2}$ とおき、$\dfrac{1}{\log r^2}\to 0$ $(r\to+0)$ を示すのは簡単) 証明することが出来るが、有界であることを言うのも簡単だから、次のような解答もある。

$x^2+y^2\le\frac{1}{2}$ のとき、$\log(x^2+y^2)\le\log\frac{1}{2}=-\log 2<0$ より $|\log(x^2+y^2)|\ge\log 2$ に注意すると、

$$\left|\frac{x+y}{\log(x^2+y^2)}\right|\le\frac{|x+y|}{\log 2}\to 0\quad((x,y)\to(0,0)).$$

(5) いわゆる不定形 $\dfrac{0}{0}$ である。近づく方向を限定して考えてみると何か分かることがある。x 軸に沿って近づけた場合

$$\lim_{\substack{(x,y)\to(0,0)\\ y=0}}\frac{x-y}{x+y}=\lim_{x\to0}\frac{x}{x}=\lim_{x\to0}1=1.$$

y 軸に沿って近づけた場合

$$\lim_{\substack{(x,y)\to(0,0)\\ x=0}}\frac{x-y}{x+y}=\lim_{y\to0}\frac{-y}{y}=\lim_{y\to0}(-1)=-1.$$

これら 2 つの極限が一致しないので、$\displaystyle\lim_{(x,y)\to(0,0)}\frac{x-y}{x+y}$ は存在しない。

(6) これも不定形 $\dfrac{0}{0}$ である。x 軸に沿って近づけた場合

$$\lim_{\substack{y=0\\ (x,y)\to(0,0)}}\frac{|x|}{\sqrt{x^2+y^2}}=\lim_{x\to0}\frac{x}{\sqrt{x^2}}=\lim_{x\to0}1=1.$$

y 軸に沿って近づけた場合

$$\lim_{\substack{x=0\\ (x,y)\to(0,0)}}\frac{|x|}{\sqrt{x^2+y^2}}=\lim_{y\to0}0=0.$$

これら 2 つの極限が一致しないので、$\displaystyle\lim_{(x,y)\to(0,0)}\frac{|x|}{\sqrt{x^2+y^2}}$ は存在しない。

(7) これも不定形 $\dfrac{0}{0}$ である。x 軸, y 軸や、$y=kx$ (k は定数) にそっての極限は、すべて 0 であることが分かる。実際例えば

$$\lim_{\substack{(x,y)\to(0,0)\\ y=kx}}\frac{x^2y^2}{x^2+y^2}=\lim_{x\to0}\frac{x^2\cdot(kx)^2}{x^2+(kx)^2}=\lim_{x\to0}\frac{k^2x^2}{1+k^2}=0.$$

これから 0 に収束しそうだと見当をつけて証明を考える。

$$\left|\frac{x^2y^2}{x^2+y^2}-0\right| = \left|\frac{x^2y^2}{x^2+y^2}\right| = x^2\frac{y^2}{x^2+y^2} \le x^2\frac{x^2+y^2}{x^2+y^2} = x^2 \to 0^2 = 0 \quad ((x,y)\to(0,0)).$$

はさみうちの原理から、

$$\lim_{(x,y)\to(0,0)}\frac{x^2y^2}{x^2+y^2} = 0.$$

(8) これも不定形 $\dfrac{0}{0}$ である。$f(x,y) := xy$, $g(z) := \dfrac{\sin z}{z}$, $a=(0,0)$, $b=0$, $c=1$ とおくと、

$$\lim_{(x,y)\to a} f(x,y) = b, \quad \lim_{z\to b} g(z) = c$$

であるから、

$$\lim_{(x,y)\to(0,0)}\frac{\sin(xy)}{xy} = \lim_{(x,y)\to a} g(f(x,y)) = \lim_{z\to b} g(z) = c = 1.$$

もう少しきちんと書くと: $A := \{(x,y)\in\mathbf{R}^2; xy\ne 0\}$, そして

$$f\colon A\ni(x,y)\mapsto xy\in\mathbf{R}, \quad g\colon \mathbf{R}\setminus\{0\}\ni z\mapsto\frac{\sin z}{z}\in\mathbf{R}$$

とおく。$f(x,y)$ は x と y の多項式であるから、いたるところ連続である。ゆえに $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = f(0,0) = 0$. 一方 (高校で学んだように) $\displaystyle\lim_{z\to 0} g(z) = 1$ である。$f(A)\subset\mathbf{R}\setminus\{0\}$ であるから、g と f は合成できて、

$$\lim_{(x,y)\to(0,0)}\frac{\sin(xy)}{xy} = \lim_{(x,y)\to(0,0)} g(f(x,y)) = 1.$$

ϵ-δ 論法でやると、以下のようになる。

$\displaystyle\lim_{\theta\to 0}\frac{\sin\theta}{\theta} = 1$ であるから、$\forall\varepsilon>0$, $\exists\delta'>0$ s.t.

$$|\theta|\le\delta' \quad\Rightarrow\quad \left|\frac{\sin\theta}{\theta}-1\right|\le\varepsilon.$$

$\delta := \sqrt{2\delta'}$ とおくと、$\delta>0$ で、$\|(x,y)-(0,0)\|<\delta$ のとき、

$$|xy| \le \frac{1}{2}\left(x^2+y^2\right) < \frac{\delta^2}{2} = \delta'.$$

ゆえにこのとき、

$$\left|\frac{\sin(xy)}{xy}-1\right|\le\varepsilon.$$

ゆえに $\displaystyle\lim_{(x,y)\to(0,0)}\frac{\sin(xy)}{xy} = 1.$ ■

問 2.2.3 (p. 33) の解答 (1) 連続である (2) 連続である (3) 連続でない (4) 連続である (5) 連続でない

問 2.3.1 (p. 35) の解答 $\{a\} = \overline{B}(a;0)$ であるから、$\{a\}$ は $\mathbf{R}^n$ の閉集合である。

問 2.3.2(p. 36) の解答　すみませんが、図を描くのは省略させてもらいます (だれか描いてくれないかなあ)。

(1) $\emptyset$ は $\mathbf{R}^2$ の開集合であり、$\mathbf{R}^2$ の閉集合でもある。これは 命題2.3.2, 2.3.3 で済んでいる。

(2) $\mathbf{R}^2$ は $\mathbf{R}^2$ の開集合であり、$\mathbf{R}^2$ の閉集合でもある。これも 命題2.3.2, 2.3.3 で済んでいる。

(3) $\{(0,0)\}$ は $\mathbf{R}^2$ の閉集合である。一般に $\forall a \in \mathbf{R}^n$ に対して、$A = \{a\}$ は $\mathbf{R}^n$ の閉集合である。実際、$f\colon \mathbf{R}^n \ni x \mapsto \|x-a\|^2 = \sum_{j=1}^{n}(x_j - a_j)^2 \in \mathbf{R}$ は、多項式関数であるから、$\mathbf{R}^n$ 上の連続関数で、$A = \{x \in \mathbf{R}^n; f(x) = 0\}$ は命題 2.3.4(2) により $\mathbf{R}^n$ の閉集合である。あるいは、

$$A = \bigcap_{j=1}^{n} F_j, \quad F_j := \{x \in \mathbf{R}^n; x_j = a_j\}$$

と書き直して、各 F_j が命題2.3.4 (2) により $\mathbf{R}^n$ の閉集合であること、それと命題 2.3.3 (2) を使う、ということも出来る。
あるいは、閉球 $\overline{B}(a;r)$ で $a = (0,0)$, $r = 0$ としても良い。

(4) 前問から、$A_1 = \{(0,0)\}$, $A_2 = \{(1,1)\}$ は $\mathbf{R}^2$ の閉集合である。命題 2.3.3 (3) を使えば $A = A_1 \cup A_2$ も $\mathbf{R}^2$ の閉集合である。

(5) 閉集合ではないが、開集合である。$A := (1,2) \times (3,4) = U_1 \cap U_2$, $U_1 := \{(x,y) \in \mathbf{R}^2; 1 < x < 2\}$, $U_2 := \{(x,y) \in \mathbf{R}^2; 3 < y < 4\}$. 命題 2.3.4 (1) を使えば、$U_1$ と U_2 が $\mathbf{R}^2$ の開集合であることが分かり、命題 2.3.2 (3) を使えば A が $\mathbf{R}^2$ の開集合であることが分かる。

(6) $[1,2] \times (3,4)$ は $\mathbf{R}^2$ の開集合でもないし、$\mathbf{R}^2$ の閉集合でもない。

(7) $A = [1,2] \times [3,4]$ は $\mathbf{R}^2$ の開集合ではないが、$\mathbf{R}^2$ の閉集合である。実際 $F_1 := \{(x,y) \in \mathbf{R}^2; 1 \le x \le 2\}$, $F_2 := \{(x,y) \in \mathbf{R}^2; 3 \le y \le 4\}$ とおくと、 $A = F_1 \cap F_2$ で、命題 2.3.3 (2) を使えば F_1 と F_2 が $\mathbf{R}^2$ の閉集合であることが分かるので、命題 2.3.3 (2) を使えば A が $\mathbf{R}^2$ の閉集合であることが分かる。

(8) $A = \{(x,y) \in \mathbf{R}^2; 5 < x^2+y^2 < 6\}$ は $\mathbf{R}^2$ の開集合である。$f(x,y) := x^2+y^2$ $((x,y) \in \mathbf{R}^2)$, $a = 5$, $b = 6$ とおくと、$f(x,y)$ は x,y の多項式で、$f\colon \mathbf{R}^2 \to \mathbf{R}$ は連続関数であり、$A = \{(x,y) \in \mathbf{R}^2; a < f(x,y) < b\}$ と書けるので、命題 2.3.4 (1) を使えば、A が $\mathbf{R}^2$ の開集合であることが分かる。

(9) $A = (0,\infty) \times (0,\infty)$ は $\mathbf{R}^2$ の開集合である。$U_1 := \{(x,y) \in \mathbf{R}^2; x > 0\}$, $U_2 := \{(x,y) \in \mathbf{R}^2; y > 0\}$ とおくと、U_1 と U_2 は命題 2.3.4 (1) より、$\mathbf{R}^2$ の開集合である。そして $A = U_1 \cap U_2$ であるから、命題 2.3.3 (3) より A は $\mathbf{R}^2$ の開集合である。

(10) $A = \{(x,y) \in \mathbf{R}^2; x^3 \le y \le x^2\}$ は $\mathbf{R}^2$ の閉集合である。$f_1(x,y) := y - x^3$, $f_2(x,y) := y - x^2$, $F_1 := \{(x,y) \in \mathbf{R}^2; f_1(x,y) \ge 0\}$, $F_2 := \{(x,y) \in \mathbf{R}^2; f_2(x,y) \le 0\}$ とおくと、F_1 と F_2 は命題 **jprop:連続関数による逆像の開・閉** (2) より $\mathbf{R}^2$ の閉集合である。また $A = F_1 \cap F_2$ であるから、命題 2.3.3 (2) より A は $\mathbf{R}^2$ の閉集合である。

(11) $A = \mathbf{R}^2 \setminus \{(0,0)\}$ は $\mathbf{R}^2$ の開集合である。実際、$f\colon \mathbf{R}^2 \ni (x,y) \mapsto x^2+y^2 \in \mathbf{R}$ は連続関数で、$A = \{(x,y) \in \mathbf{R}^2; f(x,y) > 0\}$ であるから、命題 2.3.4 (1) より A は $\mathbf{R}^2$ の開集合である。 ■

問 2.3.4(p. 36) の解答　「連続関数による開集合の逆像は開集合である」という一般的になりたつ命題の証明である。何も見ずに証明せよと言われたら簡単ではないかもしれないが、この種の証明を見慣れていればいくつかの部分は（極論すれば考えないでも）分かってしまうであろう。(ア) $f^{-1}(W)$　(イ) W　(ウ) 開集合　(エ) $\exists$　(オ) $f^{-1}(W)$. ■

$m=1$, $U=\mathbf{R}^n$, $V=\mathbf{R}^m=\mathbf{R}$, $W=(a,\infty)$ とすると、$f^{-1}(W)=\{x\in\mathbf{R}^n; f(x)>a\}$ となり、何度も使った定理の別証明となる。

問 2.6.1 (p. 33) の解答

$$\begin{aligned} f_x &= \frac{\partial}{\partial x}\left(x^2+y^2+z^2\right)^{-1/2} = -\frac{1}{2}\left(x^2+y^2+z^2\right)^{-3/2}\cdot\frac{\partial}{\partial x}\left(x^2+y^2+z^2\right) \\ &= -\left(x^2+y^2+z^2\right)^{-3/2}x = -\frac{x}{(x^2+y^2+z^2)^{3/2}}. \end{aligned}$$

$$\begin{aligned} f_{xx} &= \frac{\partial}{\partial x}f_x = \frac{\partial}{\partial x}\left(-\left(x^2+y^2+z^2\right)^{-3/2}x\right) \\ &= -\left\{-\frac{3}{2}\left(x^2+y^2+z^2\right)^{-5/2}\cdot(2x)\cdot x + \left(x^2+y^2+z^2\right)^{-3/2}\cdot 1\right\} \\ &= \frac{3x^2}{(x^2+y^2+z^2)^{5/2}} - \frac{1}{(x^2+y^2+z^2)^{3/2}}. \end{aligned}$$

同様に

$$f_{yy} = \frac{3y^2}{(x^2+y^2+z^2)^{5/2}} - \frac{1}{(x^2+y^2+z^2)^{3/2}},\quad f_{zz} = \frac{3z^2}{(x^2+y^2+z^2)^{5/2}} - \frac{1}{(x^2+y^2+z^2)^{3/2}}.$$

ゆえに

$$\begin{aligned} f_{xx}+f_{yy}+f_{zz} &= \frac{3\left(x^2+y^2+z^2\right)}{(x^2+y^2+z^2)^{5/2}} - \frac{3}{(x^2+y^2+z^2)^{3/2}} \\ &= \frac{3}{(x^2+y^2+z^2)^{3/2}} - \frac{3}{(x^2+y^2+z^2)^{3/2}} = 0. \blacksquare \end{aligned}$$

問 2.7.1　どの条件を満たすか、結果だけ先に書いておこう。(1) (a)　(2) (b)　(3) (a), (b)　(4) (a), (b), (c)　(5) (a),(b),(c),(d) となる。

(1) $g(t)=\sqrt{t}$ $(t\ge 0)$, $\varphi(x,y)=x^2+y^2$ $((x,y)\in\mathbf{R}^2)$ は連続であり、$f=g\circ\varphi$. f は連続関数の合成関数なので連続である。しかし極限

$$\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{\sqrt{h^2}-0}{h} = \lim_{h\to 0}\frac{|h|}{h}$$

は存在しないので、f は $(0,0)$ で x について偏微分可能でない。「全微分可能ならば各変数について偏微分可能」であるから、全微分可能ではない。「C^1 とは、各変数につき偏微分可能で、偏導関数がすべて連続のこと」であるから、C^1 級でない。

(2) $\displaystyle\lim_{\substack{(x,y)\to(0,0)\\(x,y)\ne(0,0)}} f(x,y)$ は存在しないので (これは講義でやった。$y=kx$ に沿った極限が食い違うから、が理由。別のやり方として、$|f(x,y)-f(0,0)|$ が 0 に収束しないことを確かめる方法がある)、f は $(0,0)$ で連続ではない。

$$f_x(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{\dfrac{h\cdot 0}{h^2+0^2}-0}{h} = \lim_{h\to 0}0 = 0,$$

$$f_y(0,0)=\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{\dfrac{0\cdot h}{0^2+h^2}-0}{h}=\lim_{h\to 0}0=0$$

であるから、f は $(0,0)$ で x, y のそれぞれについて偏微分可能。「全微分可能ならば連続」であるから、全微分可能ではない。「C^1 級ならば全微分可能」であるから、C^1 級でない。

(3) $(x,y)\neq(0,0)$ のとき、

$$|f(x,y)-f(0,0)|=\left|\frac{xy}{\sqrt{x^2+y^2}}\right|=|x|\sqrt{\frac{y^2}{x^2+y^2}}\leq|x|\sqrt{\frac{x^2+y^2}{x^2+y^2}}=|x|$$

で、$(x,y)\to(0,0)$ のとき、$|x|\to 0$ であるから、

$$|f(x,y)-f(0,0)|\to 0,\quad \text{i.e.}\quad \lim_{(x,y)\to(0,0)}f(x,y)=f(0,0).$$

ゆえに f は $(0,0)$ で連続である。

$$f_x(0,0)=\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{\dfrac{h\cdot 0}{\sqrt{h^2+0^2}}-0}{h}=\lim_{h\to 0}0=0,$$

$$f_y(0,0)=\lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to 0}\frac{\dfrac{0\cdot h}{\sqrt{0^2+h^2}}-0}{h}=\lim_{h\to 0}0=0$$

であるから、f は $(0,0)$ で、x と y のそれぞれについて偏微分可能である。$(x,y)\neq(0,0)$ とするとき、

$$\frac{f(x,y)-f(0,0)-f_x(0,0)x-f_y(0,0)y}{\sqrt{x^2+y^2}}=\frac{\dfrac{xy}{\sqrt{x^2+y^2}}-0-0\cdot x-0\cdot y}{\sqrt{x^2+y^2}}=\frac{xy}{x^2+y^2}$$

であるが、これは $(x,y)\to(0,0)$ のとき 0 に収束しない。ゆえに f は $(0,0)$ で全微分可能ではない。「C^1 級ならば全微分可能」なので C^1 級ではない。

(4) $\left|h\sin\dfrac{1}{|h|}\right|\leq|h|\to 0\ (h\to 0)$ に注意すると、

$$f_x(0,0)=\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{(h^2+0^2)\sin\dfrac{1}{\sqrt{h^2+0^2}}-0}{h}=\lim_{h\to 0}h\sin\frac{1}{|h|}=0,$$

$$f_y(0,0)=\lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to 0}\frac{(0^2+h^2)\sin\dfrac{1}{\sqrt{0^2+h^2}}-0}{h}=\lim_{h\to 0}h\sin\frac{1}{|h|}=0$$

であるから、f は $(0,0)$ で、x と y のそれぞれについて偏微分可能である。$(x,y)\neq(0,0)$ とするとき、

$$\begin{aligned}\frac{f(x,y)-f(0,0)-f_x(0,0)x-f_y(0,0)y}{\sqrt{x^2+y^2}}&=\frac{(x^2+y^2)\sin\dfrac{1}{\sqrt{x^2+y^2}}-0-0\cdot x-0\cdot y}{\sqrt{x^2+y^2}}\\&=\sqrt{x^2+y^2}\sin\frac{1}{\sqrt{x^2+y^2}}.\end{aligned}$$

$(x,y)\to(0,0)$ のとき、$\left|\sqrt{x^2+y^2}\sin\dfrac{1}{\sqrt{x^2+y^2}}\right| \le \sqrt{x^2+y^2}\to 0$ であるから、(はさみうちの原理によって)

$$\lim_{(x,y)\to(0,0)}\frac{f(x,y)-f(0,0)-f_x(0,0)x-f_y(0,0)y}{\sqrt{x^2+y^2}}=0.$$

ゆえに f は $(0,0)$ で全微分可能である。一方、$(x,y)\ne(0,0)$ とするとき、

$$\begin{aligned}
f_x(x,y)&=2x\cdot\sin\frac{1}{\sqrt{x^2+y^2}}+(x^2+y^2)\cdot\cos\frac{1}{\sqrt{x^2+y^2}}\frac{\partial}{\partial x}\left[\left(x^2+y^2\right)^{-1/2}\right]\\
&=2x\sin\frac{1}{\sqrt{x^2+y^2}}-\frac{x}{\sqrt{x^2+y^2}}\cos\frac{1}{\sqrt{x^2+y^2}}
\end{aligned}$$

であるから、

$$f_x(x,0)=2x\sin\frac{1}{|x|}-\frac{x}{|x|}\cos\frac{1}{|x|}.$$

ゆえに

$$\lim_{x\to+0}f_x(x,0)=\lim_{x\to+0}\left(2x\sin\frac{1}{x}-\cos\frac{1}{x}\right)$$

これは極限が存在しない。従って、$f_x(x,y)$ は $(x,y)\to(0,0)$ のとき $f_x(0,0)$ には収束しない。ゆえに f は C^1 級ではない。

(5)

$$\begin{aligned}
f_x(0,0)&=\lim_{h\to0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to0}\frac{0-0}{h}=\lim_{h\to0}0=0,\\
f_y(0,0)&=\lim_{h\to0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to0}\frac{0-0}{h}=\lim_{h\to0}0=0
\end{aligned}$$

であるから、f は $(0,0)$ で x と y のそれぞれについて偏微分可能である。$(x,y)\ne(0,0)$ とするとき、

$$f_x(x,y)=\frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2},\quad f_y(x,y)=\frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2}$$

であるが、

$$|f_x(x,y)-f_x(0,0)|\le|y|\frac{x^4+4x^2y^2+y^4}{(x^2+y^2)^2}\le2|y|\frac{x^4+2x^2y^2+y^4}{(x^2+y^2)^2}=2|y|\to0,$$

$$|f_y(x,y)-f_y(0,0)|\le|x|\frac{x^4+4x^2y^2+y^4}{(x^2+y^2)^2}\le2|x|\frac{x^4+2x^2y^2+y^4}{(x^2+y^2)^2}=2|x|\to0$$

であるから、C^1 級である。一般に「C^1 級ならば全微分可能」であるから、全微分可能である。

解説＆お説教 (？)

- (a),(b),(c),(d) の定義と、(d) $\Longrightarrow$ (c), (c) $\Longrightarrow$ (a), (c) $\Longrightarrow$ (b) が成り立つことはしっかりマスターしておくこと。
- 極限の計算がきちんと出来ること。はさみうちの原理を使うために不等式にも慣れる必要がある。(これらは、例や例題の解答を熟読して、分からないことがあれば、それを解消しておく努力が必要である。)

- 諸君の先輩達の期末試験の答案を見ると、不等式の扱いがめちゃくちゃであるものがかなり多い。
  - 平然と $x^2 \le x^3$ とか $y^2 \le y^4$ とする人がいるが (指数が大きい方が大きいと信じている？)、無条件では成立しない。x が正とは限らない ($x<0$ ならば $x^3<0<x^2$ である)。またそもそも x, y は 0 に近づけることが多いので、どちらかというと逆であろう。例えば $0<x<1$ ならば

$$0<a<b \quad \Rightarrow x^a > x^b.$$

    この場合、正しくは $x^2 \ge |x^3|$, $y^2 \ge y^4$ (x, y が十分 0 に近いとき)
  - はさみうちの原理を誤解しているケース。$(x,y)\to(a,b)$ のとき

$$|f(x,y)-A| \le g(x,y), \quad g(x,y)\to 0$$

    が示されれば、$-g(x,y)\le f(x,y)-A\le g(x,y)$ の両辺が 0 に収束することから、$\displaystyle\lim_{(x,y)\to(a,b)} f(x,y)=A$ が言える。絶対値抜きの $f(x,y)-A\le g(x,y)\to 0$ では、はさめない！また $g(x,y)\not\to 0$ が分かったからと言って、$\displaystyle\lim_{(x,y)\to(a,b)} f(x,y)=A$ でないとは結論できない (大きい方が 0 に収束しなくても、小さい方は 0 に収束するかも知れないでしょう？単に $g(x,y)$ の取り方がまずくて、$|f(x,y)-A|\le g(x,y)$ の評価が甘すぎるのかもしれない。)。■

- f が a で全微分可能であるとは、ある行列 A が存在して、

$$\lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{\|h\|}=0$$

  が成り立つこと、と講義で定義してあるが、このままではチェックしづらい。f が a で全微分可能であるためには、f が a で各変数について偏微分可能であり、かつヤコビ行列 $A := \left(\frac{\partial f_i}{\partial x_j}(a)\right)$ に対して、

$$(\bigstar) \qquad \lim_{h\to 0}\frac{f(a+h)-f(a)-Ah}{\|h\|}=0$$

  が成り立つことが必要十分である。こちらの方がチェックしやすい。f が 2 変数の実数値関数であれば、($\bigstar$) は

$$\lim_{(h,k)\to(0,0)}\frac{f(a+h,b+k)-f(a,b)-f_x(a,b)h-f_y(a,b)k}{\sqrt{h^2+k^2}}=0$$

  と書き換えられる。

問 2.6.2 (p. 33) の解答　まず

$$f_{xy}(0,0)=\lim_{h\to 0}\frac{f_x(0,0+h)-f_x(0,0)}{h}=\lim_{h\to 0}\frac{f_x(0,h)-f_x(0,0)}{h},$$

$$f_{yx}(0,0)=\lim_{h\to 0}\frac{f_y(0+h,0)-f_y(0,0)}{h}=\lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h}.$$

$$f_x(0,0)=\lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{\dfrac{h^3\cdot 0}{h^2+0^2}-0}{h}=\lim_{h\to 0}\frac{0}{h}=\lim_{h\to 0}0=0,$$

$$f_y(0,0)=\lim_{h\to0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to0}\frac{\dfrac{0^3\cdot h}{0^2+h^2}-0}{h}=\lim_{h\to0}\frac{0}{h}=\lim_{h\to0}0=0.$$

$h\neq0$ とするとき、

$$f_x(0,h)=\lim_{k\to0}\frac{f(0+k,h)-f(0,h)}{k}=\lim_{k\to0}\frac{\frac{k^3h}{k^2+h^2}-\frac{0^3\cdot h}{0^2+h^2}}{k}=\lim_{k\to0}\frac{k^2h}{k^2+h^2}=\frac{0}{h^2}=0.$$

$$f_y(h,0)=\lim_{k\to0}\frac{f(h,0+k)-f(h,0)}{k}=\lim_{k\to0}\frac{\frac{h^3k}{h^2+k^2}-\frac{h^3\cdot0}{h^2+0^2}}{k}=\lim_{k\to0}\frac{h^3}{h^2+k^2}=\frac{h^3}{h^2}=h.$$

ゆえに

$$f_{xy}(0,0)=\lim_{k\to0}\frac{f_x(0,h)-f_x(0,0)}{h}=\lim_{h\to0}\frac{0-0}{h}==\lim_{h\to0}0=0.$$

$$f_{yx}(0,0)=\lim_{h\to0}\frac{f_y(h,0)-f_y(0,0)}{h}=\lim_{h\to0}\frac{h-0}{h}=\lim_{h\to0}1=1.\blacksquare$$

別解　$f_x(0,h)$, $f_y(h,0)$ を求めるのに、商の微分法から得られる

$$f_x(x,y)=\frac{x^2y(x^2+3y^2)}{(x^2+y^2)^2}\quad((x,y)\neq(0,0)),\quad f_x(0,y)=0\quad(y\neq0),$$

$$f_y(x,y)=\frac{x^3(x^2-y^2)}{(x^2+y^2)^2}\quad((x,y)\neq(0,0)),\quad f_y(x,0)=x\quad(x\neq0).$$

を用いることも出来る。■

注意事項

- 類似の問題でそうなることが多いので、$f_x(0,0)$ や $f_y(0,0)$ は途中経過なしに 0 と書く人が多いけれど、いつもそうとは限らないので、省略せずに書くべきである。

- $$f_x(0,h)-f_x(0,0)=\lim_{k\to0}\frac{f(0+k,h)-f(0,h)-f(0+k,0)+f(0,0)}{k}$$
  は $f_x(0,h)$, $f_x(0,0)$ の存在が分かっていれば正しい式だが、それらの存在を示していない段階では根拠不十分である。

問 2.6.3(p. 46) の解答

問 2.6.4(p. 46) の解答

問 2.7.2 (p. 62) の解答

(1) 授業中に説明したことを確認しておく。

> 2 変数関数 f のグラフ $z=f(x,y)$ の $(x,y)=(a,b)$ における接平面は
>
> $$z=f(a,b)+f_x(a,b)(x-a)+f_y(a,b)(y-b).$$

$f(-1,1) = \sqrt{17-5x^2-4y^2}\Big|_{x=-1,y=1} = \sqrt{17-5-4} = \sqrt{8} = 2\sqrt{2}$ である。

$$f_x(x,y) = \frac{-10x}{2\sqrt{17-5x^2-4y^2}} = \frac{-5x}{\sqrt{17-5x^2-4y^2}},$$
$$f_y(x,y) = \frac{-8y}{2\sqrt{17-5x^2-4y^2}} = \frac{-4y}{\sqrt{17-5x^2-4y^2}}$$

であるから

$$f_x(-1,1) = \frac{-5\cdot(-1)}{2\sqrt{2}} = \frac{5\sqrt{2}}{4}, \quad f_y(-1,1) = \frac{-4\cdot 1}{2\sqrt{2}} = -\sqrt{2}.$$

ゆえに $x = f(x,y)$ の $(x,y) = (-1,1)$ における接平面の方程式は

$$z = f(-1,1) + f_x(-1,1)(x-(-1)) + f_y(-1,1)(y-1) = 2\sqrt{2} + \frac{5\sqrt{2}}{4}(x+1) - \sqrt{2}(y-1).$$

整理して

$$z = \frac{17+5x-4y}{2\sqrt{2}}.$$

(2) 授業中に「F のレベルセット $F(x) = c$ 上の点 a における接 (超) 平面の方程式は、$\nabla F(a)\cdot(x-a) = 0$」と説明した。3 次元バージョンは次のようになる。

$F(x,y,z) = d$ 上の点 (a,b,c) における接平面の方程式は、

$$\nabla F(a,b,c)\cdot\begin{pmatrix} x-a \\ y-b \\ z-c \end{pmatrix} = 0 \quad \text{i.e.,} \quad F_x(a,b,c)(x-a)+F_y(a,b,c)(y-b)+F_z(a,b,c)(z-c) = 0.$$

$F\left(-\frac{2}{3},\frac{4}{3},3\right) = \left(\frac{1}{3}\right)^2 + \left(\frac{4}{3}\right)^2/4 + 2^2/9 = 1$ であるから、確かに $\left(-\frac{2}{3},\frac{4}{3},3\right)$ は $F(x,y,z) = 1$ 上の点である。$\left(-\frac{2}{3},\frac{4}{3},3\right)$ における接平面は

$$(\sharp) \qquad \nabla F\left(-\frac{2}{3},\frac{4}{3},3\right)\cdot\begin{pmatrix} x-\left(-\frac{2}{3}\right) \\ y-\frac{4}{3} \\ z-2 \end{pmatrix} = 0.$$

ところで

$$\nabla F(x,y,z) = \begin{pmatrix} 2(x+1) \\ \dfrac{y}{2} \\ \dfrac{2(z-1)}{9} \end{pmatrix}, \quad \nabla F\left(-\frac{2}{3},\frac{4}{3},3\right) = \begin{pmatrix} \dfrac{2}{3} \\ \dfrac{2}{3} \\ \dfrac{4}{9} \end{pmatrix} = \frac{2}{9}\begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}.$$

ゆえに $(\sharp)$ は

$$\begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix}\cdot\begin{pmatrix} x+\frac{2}{3} \\ y-\frac{4}{3} \\ z-2 \end{pmatrix} = 0.$$

整理して

$$3x+3y+2z = 6.$$

一方、$(-\frac{2}{3}, \frac{4}{3}, 3)$ における法線は

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\frac{2}{3} \\ \frac{4}{3} \\ 3 \end{pmatrix} + t \begin{pmatrix} 3 \\ 3 \\ 2 \end{pmatrix} \quad (t \in \mathbf{R}).$$

(3) $F(x, y, z) = 1$ 上の点 (x_0, y_0, z_0) における法線ベクトルは

$$\nabla F(x_0, y_0, z_0) = \begin{pmatrix} 2(x_0+1) \\ \dfrac{y_0}{2} \\ \dfrac{2(z_0-1)}{9} \end{pmatrix}.$$

これが $(1, 1, 1)$ と平行であるには、

$$\exists t \in \mathbf{R} \quad \text{s.t.} \quad \begin{pmatrix} 2(x_0+1) \\ \dfrac{y_0}{2} \\ \dfrac{2(z_0-1)}{9} \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

このとき

$$x_0 = -1 + \frac{t}{2}, \quad y_0 = 2t, \quad z_0 = 1 + \frac{9t}{2}.$$

(x_0, y_0, z_0) が $F(x, y, z) = 1$ 上にあるので

$$\left(\frac{t}{2}\right)^2 + \frac{1}{4}(2t)^2 + \frac{1}{9}\left(\frac{9t}{2}\right)^2 = 1.$$

$$\frac{1}{4}t^2 + t^2 + \frac{9}{4}t^2 = 4 \quad \Leftrightarrow \quad \frac{14}{4}t^2 = 1 \quad \Leftrightarrow \quad t = \pm\frac{2}{\sqrt{14}}.$$

であるから

$$\begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} = \begin{pmatrix} -1 \pm \frac{1}{\sqrt{14}} \\ \pm\frac{4}{\sqrt{14}} \\ 1 \pm \frac{9}{\sqrt{14}} \end{pmatrix}.$$

接平面は $x + y + z = k$ の形になることから (複号同順[27]で計算して)

$$k = -1 \pm \frac{1}{\sqrt{14}} \pm \frac{4}{\sqrt{14}} + 1 \pm \frac{9}{\sqrt{14}} = \pm\sqrt{14}. \blacksquare$$

問 2.8.1(p. 68) の解答

問 2.8.2(p. 74) の解答　2.8.3 を見よ。

[27] こういう古めかしい言葉は教科書からは駆逐されているはずで、もう使うべきではないかもしれない。「垂線の足」、「題意」なんてのもそのお仲間か。

問 2.8.3(p. 74) の解答　高校数学で、f が微分可能な関数で、a と b が実定数であるとき、

$$(\text{☆}) \qquad (f(ax+b))' = af'(ax+b)$$

という公式を学んだ。証明は 1 変数関数の合成関数の微分法で，$u=ax+b$, $y=f(u)$ とおくと、

$$(f(ax+b))' = \frac{d}{dx}y = \frac{dy}{du}\frac{du}{dx} = f'(u)\cdot a = af'(ax+b)$$

となる、というものであった。この問題は多変数関数の微分の問題であるが、偏微分に関しては (☆) が使えて、

$$\begin{aligned}
\frac{\partial u}{\partial x} &= 1\cdot f'(x-ct)+1\cdot g'(x+ct) = f'(x-ct)+g'(x+ct),\\
\frac{\partial^2 u}{\partial x^2} &= 1\cdot f''(x-ct)+1\cdot g''(x+ct) = f''(x-ct)+g''(x+ct),\\
\frac{\partial u}{\partial x} &= (-c)\cdot f'(x-ct)+c\cdot g'(x+ct) = -cf'(x-ct)+cg'(x+ct),\\
\frac{\partial^2 u}{\partial t^2} &= (-c)^2\cdot f''(x-ct)+c^2\cdot g''(x+ct) = c^2(f''(x-ct)+g''(x+ct)).
\end{aligned}$$

ゆえに

$$\frac{\partial^2 u}{\partial x^2} = f''(x-ct)+g''(x+ct) = \frac{1}{c^2}\frac{\partial^2 u}{\partial t^2}.$$

(別解) $\xi := x-ct$, $\eta := x+ct$, $F(\xi,\eta) := f(\xi)+g(\eta)$ とおくと、

$$u(x,t) = F(\xi,\eta)$$

となるので、

$$\begin{aligned}
u_x &= F_\xi \xi_x + F_\eta \eta_x = f'(\xi)\cdot 1 + g'(\eta)\cdot 1\\
&= f'(\xi)+g'(\eta),\\
u_{xx} &= \left(f'(\xi)+g'(\eta)\right)_\xi \xi_x + \left(f'(\xi)+g'(\eta)\right)_\eta \eta_x = f''(\xi)\cdot 1 + g''(\eta)\cdot 1\\
&= f''(\xi)+g''(\eta),\\
u_t &= F_\xi \xi_t + F_\eta \eta_t = f'(\xi)(-c)+g'(\eta)\cdot c\\
&= c(-f'(\xi)+g'(\eta)),\\
u_{tt} &= c\left(-f'(\xi)+g'(\eta)\right)_\xi \xi_t + c\left(-f'(\xi)+g'(\eta)\right)_\eta \eta_t = c(-f''(\xi))\cdot(-c) + cg''(\eta)\cdot c\\
&= c^2(f''(\xi)+g''(\eta)).
\end{aligned}$$

ゆえに

$$u_{xx} = f''(\xi)+g''(\eta) = \frac{1}{c^2}u_{tt}.\blacksquare$$

問 2.8.4 (p. 74) の解答

問 2.8.5 (p. 74) の解答

$$u(t,x) = (4\pi)^{-n/2}t^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)$$

と書ける。

$$\frac{\partial}{\partial t}\left(-\frac{\|x\|^2}{4t}\right) = \frac{\partial}{\partial t}\left(-\frac{\|x\|^2}{4}t^{-1}\right) = \frac{\|x\|^2}{4}t^{-2} = \frac{\|x\|^2}{4t^2}$$

であるから、

$$\begin{aligned}
\frac{\partial u}{\partial t} &= (4\pi)^{-n/2}\left[\frac{\partial}{\partial t}t^{-n/2}\cdot\exp\left(-\frac{\|x\|^2}{4t}\right)+t^{-n/2}\frac{\partial}{\partial t}\exp\left(-\frac{\|x\|^2}{4t}\right)\right]\\
&= (4\pi)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)\left[-\frac{n}{2}t^{-n/2-1}+t^{-n/2}\frac{\|x\|^2}{4t^2}\right]\\
&= (4\pi)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)t^{-n/2-2}\left(\frac{\|x\|^2}{4}-\frac{nt}{2}\right)\\
&= (4\pi t)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)\left(\frac{\|x\|^2}{4t^2}-\frac{n}{2t}\right).
\end{aligned}$$

一方、

$$\frac{\partial}{\partial x_j}\left(-\frac{\|x\|^2}{4t}\right)=\frac{\partial}{\partial x_j}\left(-\frac{1}{4t}\left(x_1^2+\cdots+x_n^2\right)\right)=-\frac{x_j}{2t}$$

であるから、

$$\begin{aligned}
\frac{\partial u}{\partial x_j} &= (4\pi t)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)\frac{\partial}{\partial x_j}\left(-\frac{\|x\|^2}{4t}\right)=(4\pi t)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)\left(-\frac{2x_j}{4t}\right)\\
&= -\frac{x_j}{2t}(4\pi t)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right).
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 u}{\partial x_j^2} &= -(4\pi t)^{-n/2}\frac{1}{2t}\frac{\partial}{\partial x_j}\left[x_j\exp\left(-\frac{\|x\|^2}{4t}\right)\right]=-(4\pi t)^{-n/2}\frac{1}{2t}\exp\left(-\frac{\|x\|^2}{4t}\right)\left(1-x_j\cdot\frac{x_j}{2t}\right)\\
&= -(4\pi t)^{-n/2}\frac{1}{2t}\exp\left(-\frac{\|x\|^2}{4t}\right)\left(1-\frac{x_j^2}{2t}\right)\\
&= (4\pi t)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)\left(\frac{x_j^2}{4t^2}-\frac{1}{2t}\right).
\end{aligned}$$

$\displaystyle\sum_{j=1}^n x_j^2=\|x\|^2$, $\displaystyle\sum_{j=1}^n 1=n$ に注意すると、

$$\triangle u=\sum_{j=1}^n\frac{\partial^2 u}{\partial x_j^2}=(4\pi t)^{-n/2}\exp\left(-\frac{\|x\|^2}{4t}\right)\left(\frac{\|x\|^2}{4t^2}-\frac{n}{2t}\right)=\frac{\partial u}{\partial t}.\blacksquare$$

問 2.8.6 (p. 74) の解答

$$u_t=-cU'(\nu\cdot x-ct),\quad u_{tt}=(-c)^2U''(\nu\cdot x-ct)=c^2U''(\nu\cdot x-ct),$$

$$u_{x_j}=\nu_jU'(\nu\cdot x-ct),\quad u_{x_jx_j}=\nu_j^2U''(\nu\cdot x-ct)$$

であるから、

$$\triangle u(x,t)=\sum_{j=1}^n u_{x_jx_j}=\sum_{j=1}^n\nu_j^2U''(\nu\cdot x-ct)=\|\nu\|^2U''(\nu\cdot x-ct)=U''(\nu\cdot x-ct)=\frac{1}{c^2}u_{tt}.\blacksquare$$

問 2.8.7(p. 75) の解答　(右辺から左辺)

$$x=\frac{1}{2}(\xi+\eta),\quad t=\frac{1}{2c}(\eta-\xi)$$

であるから、

$$x_\xi = \frac{1}{2}, \quad x_\eta = \frac{1}{2}, \quad t_\xi = -\frac{1}{2c}, \quad t_\eta = \frac{1}{2c}.$$

chain rule によって

$$\begin{aligned}
v_\eta &= u_x x_\eta + u_t t_\eta = \frac{1}{2}u_x + \frac{1}{2c}u_t,\\
v_{\eta\xi} &= \frac{1}{2}\left(u_{xx}x_\xi + u_{xt}t_\xi\right) + \frac{1}{2c}\left(u_{tx}x_\xi + u_{tt}t_\xi\right) = \frac{1}{4}u_{xx} - \frac{1}{4c}u_{xt} + \frac{1}{4c}u_{tx} - \frac{1}{4c^2}u_{tt}\\
&= \frac{1}{4}\left(u_{xx} - \frac{1}{c^2}u_{tt}\right).
\end{aligned}$$

ただし u が C^2 級であるから $u_{xt} = u_{tx}$ が成り立つことを用いた。ゆえに

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} = -4\frac{\partial^2 v}{\partial \xi \partial \eta}.$$

あるいは

$$\begin{aligned}
v_{\eta\xi} &= \frac{\partial}{\partial \xi}u_x \cdot x_\eta + u_x \frac{\partial}{\partial \xi}x_\eta + \frac{\partial}{\partial \xi}u_t \cdot t_\eta + u_t\frac{\partial}{\partial \xi}t_\eta\\
&= (u_{xx}x_\xi + u_{xt}t_\xi)x_\eta + u_x \cdot x_{\eta\xi} + (u_{tx}x_\xi + u_{tt}t_\xi)t_\eta + u_t \cdot t_{\eta\xi}\\
&= u_{xx}x_\xi x_\eta + u_{xt}t_\xi x_\eta + u_{tx}x_\xi t_\eta + u_{tt}t_\xi t_\eta + u_x x_{\eta\xi} + u_t t_{\eta\xi}\\
&= \frac{1}{4}u_{xx} - \frac{1}{4c}u_{xt} + \frac{1}{4c}u_{tx} + \left(-\frac{1}{4c^2}\right)u_{tt}.
\end{aligned}$$

(左辺から右辺)

$$\xi_x = 1, \quad \xi_t = -c, \quad \eta_x = 1, \quad \eta_t = c.$$

chain rule によって

$$\begin{aligned}
u_t &= v_\xi \xi_t + v_\eta \eta_t = -cv_\xi + cv_\eta,\\
u_{tt} &= -c(v_{\xi\xi}\xi_t + v_{\xi\eta}\eta_t) + c\left(v_{\eta\xi}\xi_t + v_{\eta\eta}\eta_t\right) = c^2 v_{\xi\xi} - c^2 v_{\xi\eta} - c^2 v_{\eta\xi} + c^2 v_{\eta\eta}\\
&= c^2\left(v_{\xi\xi} - 2v_{\eta\xi} + v_{\eta\eta}\right),\\
u_x &= v_\xi \xi_x + v_\eta \eta_x = v_\xi + v_\eta,\\
u_{xx} &= v_{\xi\xi} + v_{\eta\xi} + v_{\eta\xi} + v_{\eta\eta} = v_{\xi\xi} + 2v_{\xi\eta} + v_{\eta\eta}.
\end{aligned}$$

ただし v が C^2 級であるから、$v_{\xi\eta} = v_{\eta\xi}$ であることを用いた。ゆえに

$$\frac{1}{c^2}\frac{\partial^2 u}{\partial t^2} = \frac{\partial^2 u}{\partial x^2} = (v_{\xi\xi} - 2v_{\eta\xi} + v_{\eta\eta}) - (v_{\xi\xi} + 2v_{\eta\xi} + v_{\eta\eta}) = -4\frac{\partial^2 v}{\partial \xi \partial \eta}.\blacksquare$$

問 2.8.8 (p. 75) の解答

(1) $\varphi'(r,\theta) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial r} & \frac{\partial \varphi_1}{\partial \theta} \\ \frac{\partial \varphi_2}{\partial r} & \frac{\partial \varphi_2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$

(2) chain rule から

$$g_r = f_x x_r + f_y y_r = f_x \cos\theta + f_y \sin\theta.$$

$x_{rr} = y_{rr} = 0$ であるので、

$$\begin{aligned}
g_{rr} &= (f_{xx}x_r + f_{xy}y_r)x_r + f_x x_{rr} + (f_{yx}x_r + f_{yy}y_r)y_r + f_y y_{rr}\\
&= f_{xx}(x_r)^2 + 2f_{xy}x_r y_r + f_{yy}(y_r)^2 + f_x x_{rr} + f_y y_{rr}\\
&= f_{xx}\cos^2\theta + 2f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta.
\end{aligned}$$

同様に

$$g_\theta = f_x x_\theta + f_y y_\theta.$$

$x_{\theta\theta} = -r\cos\theta$, $y_{\theta\theta} = -r\sin\theta$ であるから

$$\begin{aligned}
g_{\theta\theta} &= (f_{xx}x_\theta + f_{xy}y_\theta)\,x_\theta + f_x x_{\theta\theta} + (f_{yx}x_\theta + f_{yy}y_\theta)\,y_\theta + f_y y_{\theta\theta} \\
&= f_{xx}(x_\theta)^2 + 2f_{xy}x_\theta y_\theta + f_{yy}(y_\theta)^2 + f_x x_{\theta\theta} + f_y y_{\theta\theta} \\
&= f_{xx}r^2\sin^2\theta + 2f_{xy}(-r^2\cos\theta\sin\theta) + f_{yy}r^2\cos^2\theta + f_x(-r\cos\theta) + f_y(-r\sin\theta) \\
&= r^2\left(f_{xx}\sin^2\theta - 2f_{xy}\cos\theta\sin\theta + f_{yy}\cos^2\theta\right) - r(f_x\cos\theta + f_y\sin\theta).
\end{aligned}$$

ゆえに

$$\begin{aligned}
g_{rr} + \frac{1}{r}g_r + \frac{1}{r^2}g_{\theta\theta} &= f_{xx}\cos^2\theta + 2f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta + \frac{1}{r}(f_x\cos\theta + f_y\sin\theta) \\
&\quad + \left(f_{xx}\sin^2\theta - 2f_{xy}\cos\theta\sin\theta + f_{yy}\cos^2\theta\right) - \frac{1}{r}(f_x\cos\theta + f_y\sin\theta) \\
&= f_{xx} + f_{yy}.\blacksquare
\end{aligned}$$

問 2.8.9 (p. 75) の解答　(1) は上と同じ。(2) $\det\varphi'(r,\theta) = r\cos^2\theta - (-r\sin^2\theta) = r$ であるから、$r \neq 0$ であれば $\det\varphi'(r,\theta) \neq 0$. ゆえに逆関数定理の仮定が成り立ち、考えている点 (r,θ) の近傍で、C^1 級の逆関数が存在する。(3) 逆関数の微分法より

$$\begin{aligned}
\begin{pmatrix} r_x & r_y \\ \theta_x & \theta_y \end{pmatrix} &= \left(\varphi^{-1}\right)'(x,y) = \left(\varphi'(r,\theta)\right)^{-1} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{pmatrix}.
\end{aligned}$$

(4)

$$\begin{aligned}
f_x &= g_r r_x + g_\theta\theta_x = g_r\cos\theta - \frac{g_\theta\sin\theta}{r}, \\
f_y &= g_r r_y + g_\theta\theta_y = g_r\sin\theta + \frac{g_\theta\cos\theta}{r}.
\end{aligned}$$

これから

$$\frac{\partial}{\partial x} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}, \quad \frac{\partial}{\partial y} = \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}$$

が成り立つ。

$$\begin{aligned}
f_{xx} + f_{yy} &= \frac{\partial}{\partial x}f_x + \frac{\partial}{\partial y}f_y \\
&= \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(g_r\cos\theta - \frac{g_\theta\sin\theta}{r}\right) \\
&\quad + \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\right)\left(g_r\sin\theta + \frac{g_\theta\cos\theta}{r}\right) \\
&= \cdots\text{中略}\cdots \\
&= g_{rr}\cos^2\theta - \frac{2g_{r\theta}\sin\theta\cos\theta}{r} + \frac{g_{\theta\theta}\sin^2\theta}{r^2} + \frac{g_r\sin^2\theta}{r} + \frac{2g_\theta\sin\theta\cos\theta}{r^2} \\
&\quad + g_{rr}\sin^2\theta + \frac{2g_{r\theta}\sin\theta\cos\theta}{r} + \frac{g_{\theta\theta}\cos^2\theta}{r^2} + \frac{g_r\cos^2\theta}{r} - \frac{2g_\theta\sin\theta\cos\theta}{r^2} \\
&= g_{rr} + \frac{1}{r}g_r + \frac{1}{r^2}g_{\theta\theta}.\blacksquare
\end{aligned}$$

問 2.9.1(p. 77) の解答　(1)

$$F'(t) = f_x x_t + f_y y_t = f_x(a+th, b+tk)h + f_y(a+th, b+tk)k,$$

簡単のため、$c := (a+th, b+tk)$ とおく。f が C^2 級であることから、$f_{xy} = f_{yx}$ であることに注意すると、

$$\begin{aligned} F''(t) &= \left(f_{xx}(c)h + f_{xy}(c)k\right)h + \left(f_{yx}(c)h + f_{yy}(c)k\right)k \\ &= f_{xx}(c)h^2 + 2f_{xy}(c)hk + f_{yy}(c)k^2, \end{aligned}$$

f が C^3 級であることから、$f_{xxy} = f_{xyx}$, $f_{xyy} = f_{yyx}$ であることに注意すると、

$$\begin{aligned} F''(t) &= \left(f_{xxx}(c)h + f_{xxy}(c)k\right)h^2 + 2\left(f_{xyx}(c)h + f_{xyy}(c)k\right)hk + \left(f_{yyx}(c)h + f_{yyy}(c)k\right)y^2 \\ &= f_{xxx}(c)h^3 + 3f_{xxy}(c)h^2k + 3f_{xyy}(c)hk^2 + f_{yyy}(c)k^3. \end{aligned}$$

(2) $m = 1, 2, 3$ での結果から

$$(\heartsuit) \qquad F^{(m)}(t) = \sum_{r=0}^{m} \binom{m}{r} \frac{\partial^m f}{\partial x^r \partial y^{m-r}}(a+th, b+tk)h^r k^{m-r}$$

と推測される。$m = n$ のとき $(\heartsuit)$ が成立したと仮定すると、chain rule と $\binom{n}{r} + \binom{n}{r-1} = \binom{n+1}{r}$ から、

$$\begin{aligned} F^{(n+1)}(t) &= \frac{d}{dt}F^{(n)}(t) = \frac{d}{dt}\sum_{r=0}^{n}\binom{n}{r}\frac{\partial^n f}{\partial x^r \partial y^{n-r}}(a+th, b+tk)h^r k^{n-r} \\ &= \sum_{r=0}^{n}\binom{n}{r}\left(\frac{\partial^{n+1} f}{\partial x^{r+1}\partial y^{n-r}}(c)h + \frac{\partial^{n+1} f}{\partial x^r \partial y^{n-r+1}}(c)k\right)h^r k^{n-r} \\ &= \sum_{r=0}^{n}\binom{n}{r}\frac{\partial^{n+1} f}{\partial x^{r+1}\partial y^{n+1-(r+1)}}(c)h^{r+1}k^{n+1-(r+1)} + \sum_{r=0}^{n}\binom{n}{r}\frac{\partial^{n+1} f}{\partial x^r \partial y^{n+1-r}}(c)h^r k^{n+1-r} \\ &= \sum_{r'=1}^{n+1}\binom{n}{r'-1}\frac{\partial^{n+1} f}{\partial x^{r'}\partial y^{n+1-r'}}(c)h^{r'}k^{n+1-r'} + \sum_{r=0}^{n}\binom{n}{r}\frac{\partial^{n+1} f}{\partial x^r \partial y^{n+1-r}}(c)h^r k^{n+1-r} \\ &= \frac{\partial^{n+1} f}{\partial x^{n+1}}(c)h^{n+1} + \sum_{r=1}^{n}\left(\binom{n}{r-1} + \binom{n}{r}\right)\frac{\partial^{n+1} f}{\partial x^r \partial y^{n+1-r}}(c)h^r k^{n+1-r} + \frac{\partial^{n+1} f}{\partial y^{n+1}}(c)k^{n+1} \\ &= \frac{\partial^{n+1} f}{\partial x^{n+1}}(c)h^{n+1} + \sum_{r=1}^{n}\binom{n+1}{r}\frac{\partial^{n+1} f}{\partial x^r \partial y^{n+1-r}}(c)h^r k^{n+1-r} + \frac{\partial^{n+1} f}{\partial y^{n+1}}(c)k^{n+1} \\ &= \sum_{r=0}^{n+1}\binom{n+1}{r}\frac{\partial^{n+1} f}{\partial x^r \partial y^{n+1-r}}(a+th, b+tk)h^r k^{n+1-r}. \end{aligned}$$

これは $m = n+1$ のときも $(\heartsuit)$ が成り立つことを示している。ゆえに帰納法により、$(\heartsuit)$ は任意の $m \in \mathbf{N}$ に対して成り立つ。 ■

問 2.10.1 (p. 89) の解答　1 変数の 2 次形式とは、中学校で習う関数 $f(x) = ax^2$ ($x \in \mathbf{R}$; a は実定数) のことである。それが正値であるためには $a > 0$ が必要十分である。負値であるためには $a < 0$ が必要十分である。不定符号となることはない。$a = 0$ のとき、正値でも負値でもない。

問 2.10.2(p. 89) の解答　先に結果だけ書くと (1) 正値, (2) 負値, (3) 不定符号, (4) 不定符号, (5) 正値, (6) 負値, (7) どれでもない, (8) どれもない, (9) どれでもない, (10) 正値, (11) 負値, (12) 不定符号, (13) 正値, (14) 負値, (15) 正値, (16) 不定符号
((7), (8), (9) を「どちらでもない」と書く人が少なくないけれど、3 つあるのだから「どれでもない」と書くべきであろう。)

以下、行列 A の k 次首座小行列を A_k と書くことにする。

(1) 対角行列だから、固有値は対角成分の 2, 1. ともに正だから正値である。

(2) 対角行列だから、固有値は対角成分の -2, -1. ともに負だから負値である。

(3) 対角行列だから、固有値は対角成分の 2, -1. 正と負だから、不定符号である。

(4) 問題の行列を A とおくと、

$$\det A_1 = A_1 = 4 > 0, \quad \det A_2 = \det A = \begin{vmatrix} 4 & 3 \\ 3 & 1 \end{vmatrix} = 4 \cdot 1 - 3^2 = -5 < 0.$$

ゆえに、A は負値でなく、正値でもない。A の固有値を λ_1, λ_2 とおくと、これらは実数で $\lambda_1 \lambda_2 = \det A < 0$. ゆえに固有値は異符号であり、$A$ は不定符号である。

(別解) A の特性多項式は $\det(\lambda I - A) = \begin{vmatrix} \lambda - 4 & -3 \\ -3 & \lambda - 1 \end{vmatrix} = (\lambda - 4)(\lambda - 1) - (-3)(-3) = \lambda^2 - 5\lambda - 5$ であり、固有値は $\dfrac{5 \pm \sqrt{45}}{2} = \dfrac{5 \pm 3\sqrt{5}}{2}$ である。正と負だから不定符号である。

(5) 問題の行列を A とおくと、

$$\det A_1 = A_1 = 4 > 0, \quad \det A_2 = \det A = 4 \cdot 4 - 3 \cdot 3 = 16 - 9 = 7 > 0.$$

任意の k に対して $\det A_k > 0$ が成り立っているので、A は正値である。

(別解) A の特性多項式は $\det(\lambda I - A) = \begin{vmatrix} \lambda - 4 & -3 \\ -3 & \lambda - 4 \end{vmatrix} = (\lambda - 4)^2 - 3^2 = \lambda^2 - 8\lambda + 7 = (\lambda - 1)(\lambda - 7)$ であり、固有値は 1, 7 である。共に正であるから正値である。

(6) (これは前問の行列の -1 倍だから、負値である。)

問題の行列を A とおくと、

$$\det A_1 = A_1 = -4 < 0, \quad \det A_2 = \det A = (-4)^2 - (-3)^2 = 7 > 0.$$

$(-1)^k \det A_k > 0$ $(\forall k)$ を満たしているので、A は負値である。

(別解) A の特性多項式は $\det(\lambda I - A) = \begin{vmatrix} \lambda + 4 & 3 \\ 3 & \lambda + 4 \end{vmatrix} = (\lambda + 4)^2 - 3^2 = \lambda^2 + 8\lambda + 7 = (\lambda + 1)(\lambda + 7)$ であり、固有値は -1, -7 である。共に負であるから負値である。

(7) 問題の行列を A とおくと、

$$\det A_1 = 4 > 0, \quad \det A_2 = 4 \cdot 1 - 2 \cdot 2 = 0.$$

ゆえに A は負値でなく、正値でもない。A の固有値を λ_1, λ_2 とおくと、$\lambda_1 \lambda_2 = \det A = 0$ であるから、2 つの固有値のうち少なくとも一つは 0 である。ゆえに A は不定符号ではない (2 次の正方行列 A が不定符号であれば、$\det A$ は負である)。

(別解) A の特性多項式は $\det(\lambda I - A) = \begin{vmatrix} \lambda - 4 & -2 \\ -2 & \lambda - 1 \end{vmatrix} = (\lambda-4)(\lambda-1) - 2^2 = \lambda^2 - 5\lambda = \lambda(\lambda - 5)$ であり、固有値は $0, 5$ である。固有値に 0 があるので、正値でも負値でもない。また正の固有値はあるが、負の固有値はないので、不定符号でもない。

(8) これは対角行列なので、固有値は対角成分で、$0, 0$. 正でない固有値 0 があるので正値ではなく、負でない固有値 2 があるので負値ではなく、正と負両方の固有値があるわけでない (負の固有値はない) ので不定符号ではない。

(9) これは対角行列なので、固有値は対角成分で、$2, 0$. 正でない固有値があるので正値ではなく、負でない固有値があるので負値ではなく、正と負両方の固有値があるわけでないので不定符号ではない。

(10) 対角行列だから、固有値は対角成分の $2, 1, 3$. みな正だから正値である。

(11) 対角行列だから、固有値は対角成分の $-2, -1, -3$. みな負だから負値である。

(12) 対角行列だから、固有値は対角成分の $1, -2, 0$. 正の固有値と負の固有値があるので不定符号である。

(13) 問題の行列を A とおくと、

$$\det A_1 = 4 > 0, \quad \det A_2 = \begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} = 4^2 - 3^2 = 16 - 9 = 7 > 0,$$
$$\det A_3 = \begin{vmatrix} 4 & 3 & 0 \\ 3 & 4 & 0 \\ 0 & 0 & 1 \end{vmatrix} = \begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} \cdot 1 = 7 > 0$$

であるから、$\det A_k > 0\ (\forall k)$ が成り立っていて、正値であることが分かる。

(別解) A の 特性多項式は $\det(\lambda I - A) = \begin{vmatrix} \lambda - 4 & -3 & 0 \\ -3 & \lambda - 4 & 0 \\ 0 & 0 & \lambda - 1 \end{vmatrix} = (\lambda - 1) \begin{vmatrix} \lambda - 4 & -3 \\ -3 & \lambda - 4 \end{vmatrix} = (\lambda-1)((\lambda-4)^2 - (-3)^2) = (\lambda-1)(\lambda^2 - 8\lambda + 7) = (\lambda-1)(\lambda-1)(\lambda-7) = (\lambda-1)^2(\lambda-7)$ であり、固有値は $1, 1, 7$ である。すべて正であるから正値である。

(別解) $A = \left(\begin{array}{cc|c} 4 & 3 & 0 \\ 3 & 4 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$ とブロック分けでき、対角線上にあるブロック以外はすべて 0 である。ゆえに対角線上にあるブロックの固有値を調べればよい。$\begin{pmatrix} 4 & 3 \\ 3 & 4 \end{pmatrix}$ は既に見たように正値である。右下のブロックの固有値は 1 でこれも正である。ゆえに正値である。

(14) 問題の行列を A とすると、

$$\det A_1 = -4 < 0, \quad \det A_2 = \begin{vmatrix} -4 & 3 \\ 3 & -4 \end{vmatrix} = (-4)^2 - 3^2 = 16 - 9 = 7 > 0,$$
$$\det A_3 = \begin{vmatrix} -4 & 3 & 0 \\ 3 & -4 & 0 \\ 0 & 0 & -2 \end{vmatrix} = \begin{vmatrix} -4 & 3 \\ 3 & -4 \end{vmatrix} \cdot (-2) = 7 \cdot (-2) = -14 < 0$$

であり、$(-1)^k \det A_k > 0\ (\forall k)$ が成り立っているので、A は負値であることが分かる。

(別解) A の特性多項式は $\det(\lambda I - A) = \begin{vmatrix} \lambda+4 & -3 & 0 \\ -3 & \lambda+4 & 0 \\ 0 & 0 & \lambda+2 \end{vmatrix} = (\lambda+2)\begin{vmatrix} \lambda+4 & -3 \\ -3 & \lambda+4 \end{vmatrix} =$ $(\lambda+2)((\lambda+4)^2-(-3)^2) = (\lambda+2)(\lambda^2+8\lambda+7) = (\lambda+3)(\lambda+1)(\lambda+7) = (\lambda+1)(\lambda+3)(\lambda+7)$ であり、固有値は $-1, -3, -7$ である。すべて負であるから負値である。

(別解) これも $A = \left(\begin{array}{cc|c} -4 & 3 & 0 \\ 3 & -4 & 0 \\ \hline 0 & 0 & -2 \end{array}\right)$ とブロック分けすると、固有値は、$A_2 = \begin{pmatrix} -4 & 3 \\ 3 & -4 \end{pmatrix}$ の固有値と、-2 を合わせたものだと分かる。A_2 は負値であるので (省略)、問題の行列の固有値はすべて負であることが分かり、負値である。

(15) 問題の行列を A とすると、

$$\det A_1 = 4 > 0, \quad \det A_2 = \begin{vmatrix} 4 & 1 \\ 1 & 4 \end{vmatrix} = 4^2 - 1^2 = 15 > 0,$$
$$\det A_3 = \begin{vmatrix} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 4 \end{vmatrix} = 4 \cdot 4 \cdot 4 - 4 \cdot 1 \cdot 1 - 4 \cdot 1 \cdot 1 = 64 - 4 - 4 = 56 > 0$$

であり、$\det A_k > 0\ (\forall k)$ が成り立っているので、正値であることが分かる。

(別解) A の特性多項式は

$$\det(\lambda I - A) = \begin{vmatrix} \lambda-4 & -1 & 0 \\ -1 & \lambda-4 & -1 \\ 0 & -1 & \lambda-4 \end{vmatrix} = (\lambda-4)\begin{vmatrix} \lambda-4 & -1 \\ -1 & \lambda-4 \end{vmatrix} + (-1)^{(2+1)}(-1)\begin{vmatrix} -1 & 0 \\ -1 & \lambda-4 \end{vmatrix}$$
$$= (\lambda-4)\left((\lambda-4)^2 - (-1)^2\right) - (\lambda-4) = (\lambda-4)\left(\lambda^2 - 8\lambda + 14\right)$$

であり、固有値は $4, 4 \pm \sqrt{2}$ である。すべて正であるから正値である。

(16) 問題の行列を A とすると、$\det A_1 = 1 > 0,\ \det A_2 = -2 < 0,\ \det A_3 = -24 < 0.$ $\det A_1 < 0$ でないので負値でなく、$\det A_2 > 0$ でないので正値でない。一方で $\det A_3 = 0$ でないので 0 は固有値でない (従って、すべての固有値は正または負の実数である)。以上より、A は正負両方の固有値を持つことが分かるので、不定符号である。

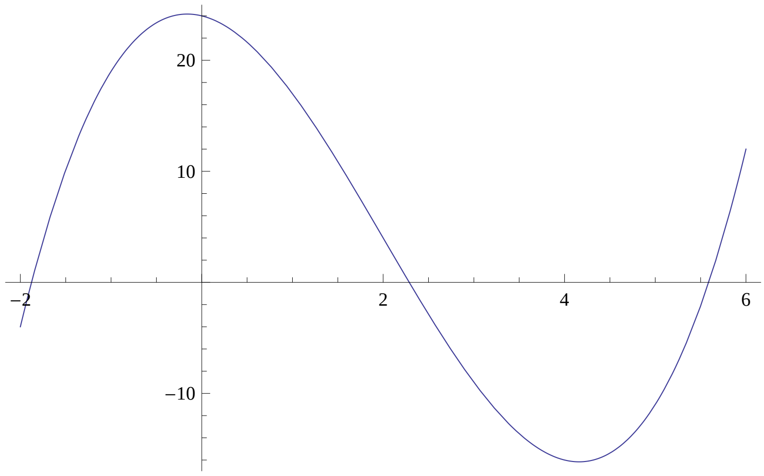
(注意) A の特性多項式は $\det(\lambda I - A) = \lambda^3 - 6\lambda^2 - 2\lambda + 24 =: p(\lambda)$. これは不還元の場合で、Cardano の公式でも虚数の 3 乗根が現れ、扱いづらい。実際、この多項式の根は

$$\lambda = \frac{1}{3}\left(6 + \frac{7\, 6^{2/3}}{\sqrt[3]{-9+i\sqrt{1977}}} + \sqrt[3]{6i\left(\sqrt{1977}+9i\right)}\right),$$
$$2 - \frac{\left(1+i\sqrt{3}\right)\sqrt[3]{-9+i\sqrt{1977}}}{6^{2/3}} + \frac{-7+7i\sqrt{3}}{\sqrt[3]{6i\left(\sqrt{1977}+9i\right)}},$$
$$2 + \frac{i\left(\sqrt{3}+i\right)\sqrt[3]{-9+i\sqrt{1977}}}{6^{2/3}} + \frac{-7-7i\sqrt{3}}{\sqrt[3]{6i\left(\sqrt{1977}+9i\right)}}.$$

これは (もちろん) いずれも実数で

$$\lambda \fallingdotseq 5.580664\cdots, 2.2874\cdots, -1.877074\cdots.$$

関数としての p のグラフは次のようになる ($p'(\lambda)=0$ の根 $\lambda=\frac{6\mp\sqrt{42}}{3}$ について、$\frac{6-\sqrt{42}}{3}<0<\frac{6+\sqrt{42}}{3}$ であること、p の極値 $p\left(\frac{6\mp\sqrt{42}}{3}\right)=\frac{4\left(9\pm 7\sqrt{42}\right)}{9}$ (正と負), $p(0)=24>0$, であることを使うと、$p(\lambda)$ の根の符号が分かる程度の p のグラフの概形は分かる)。



概形が分かれば、$p(-2)<0$, $p(0)>0$, $p(4)<0$, $p(6)>0$ をチェックして、$(-2,0)$, $(0,4)$, $(4,6)$ に根があることが分かる。3 次方程式なので、他に根はない。正と負の根があるので、不定符号である。■

問 2.10.3 (p. 89) の解答　a, b, c を実定数として、$f(x,y)=ax^2+2bxy+cy^2$ の形をしている関数 $f\colon \mathbf{R}^2\to\mathbf{R}$ のことである。正値である $\Leftrightarrow$ $a>0$ かつ $ac-b^2>0$. 負値である $\Leftrightarrow$ $a<0$ かつ $ac-b^2>0$. 不定符号である $\Leftrightarrow$ $ac-b^2<0$. $ac-b^2=0$ のときは正値でも負値でも不定符号でもない。

余談: アルゴリズムの追求

固有値計算作戦　固有多項式 $\varphi(\lambda):=\det(\lambda I-A)$ を計算して、その根を求める。$n=2$ のときは 2 次方程式の解の公式で計算可能であるが、$n\geq 3$ になると困難になる。$n=3$ であっても、いわゆる不還元の場合には、解が虚数の 3 乗根を含む形で表されることになる (微積分を使えば何とか処理できるが面倒である)。

首座小行列式作戦　$k=1,2,\cdots$ に対して、$\det A_k$ (A_k は A の k 次首座小行列) を計算して符号を調べる。

(i) すべて正である ($\forall k\in\{1,\cdots,n\}$ $\det A_k>0$) ことは、A が正値であるための必要十分条件である。

(ii) 負から始まり、負と正が交互に現れる ($\forall k\in\{1,\cdots,n\}$ $(-1)^k\det A_k>0$) ことは、A が負値であるための必要十分条件である。

(iii) 上の (i), (ii) のいずれでもない場合、$\det A$ を計算する。もし $\det A\neq 0$ であるならば、A は不定符号である。$\det A=0$ のときは、一般には面倒だが、

  (a) $n=2$ の場合は、正値、負値、不定符号のいずれでもないと結論できる ($\det A<0$ $\Longleftrightarrow$ A は不定符号)。

(b) また $n=3$ の場合は、固有多項式が容易に因数分解可能で、符号の判定は容易である。結論だけ書いておくと

$$A=\begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix}$$

が $\det A=0$ を満たすとき、

$$\lambda_1\lambda_2=-b^2-c^2-e-2+ad+af+df$$

が負ならば不定符号、そうでないならば正値でも負値でも不定符号でもない。

(c) $n\ge 4$ の場合は研究課題であろう。

Gauss の消去法作戦 (「ベクトル空間論」で、2 次形式を平方完成することで標準形に変形して (2 次形式の対角化とも呼ばれる式変形である) 符号数を調べる方法を学んだようですが、それと本質的に同じやり方です。こちらの方が解答はあっさり書けます。) A の対角線から下を掃き出す。A が正値であるためには、対角線に非零要素は現れず、行交換なしに最後まで上三角化が進められて、対角線に正数が並ぶことが必要十分。一方、A が負値であるためには、対角線に非零要素は現れず、行交換なしに最後まで上三角化が進められて、対角線に負数が並ぶことが必要十分。そのいずれでもない場合、必要ならば行交換を施して計算を進めて A の行列式を計算する (行交換を全部で r 回した場合、最終的には対角成分の積 $\times(-1)^r=\det A$ である)。$\det A\ne 0$ ならば、A は 0 を固有値に持たず、正値でも負値でもないので、A は実は不定符号であることが分かる。$\det A=0$ の場合は少々難しいが、シフトしてみるなどして、「何とかなる」場合が多いであろう。

$A=\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix}$ の場合、行交換なしに

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{pmatrix}\to\begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -6 \\ 0 & -6 & -3 \end{pmatrix}\to\begin{pmatrix} 1 & 2 & 3 \\ 0 & -2 & -6 \\ 0 & 0 & 15 \end{pmatrix}$$

となるので、いわゆる符号は $(2,1)$ (正の固有値が 2 個、負の固有値が 1 個) で、不定符号である。

首座小行列式を用いた方法、Gauss の消去法による方法をマスターすべきでしょう。■

問 2.10.4(p. 91) の解答

(1) $f'(x,y)=(f_x\ f_y)=\begin{pmatrix}6x^2+6y^2-2 & 12xy\end{pmatrix}$.

(2) $H(x,y)=\begin{pmatrix} f_{xx} & f_{yx} \\ f_{xy} & f_{yy}\end{pmatrix}=12\begin{pmatrix} x & y \\ y & x\end{pmatrix}$.

(3) $f\colon \mathbf{R}^2\to\mathbf{R}$ は C^∞ 級である。極値を取る点 (x,y) では $f'(x,y)=0$ が成り立つ。

$$\begin{aligned} f'(x,y)=0 &\Leftrightarrow 6x^2+6y^2-2=0\wedge 12xy=0 \\ &\Leftrightarrow 3x^2+3y^2=1\wedge(x=0\vee y=0) \\ &\Leftrightarrow (x=0\wedge 3x^2+3y^2=1)\vee(y=0\wedge 3x^2+3y^2=1) \\ &\Leftrightarrow (x,y)=\left(0,\pm\frac{1}{\sqrt{3}}\right)\vee(x,y)=\left(\pm\frac{1}{\sqrt{3}},0\right). \end{aligned}$$

$f'(x,y)=0$ となる点 (x,y) は、$\left(0,\dfrac{1}{\sqrt{3}}\right)$, $\left(0,-\dfrac{1}{\sqrt{3}}\right)$, $\left(\dfrac{1}{\sqrt{3}},0\right)$, $\left(-\dfrac{1}{\sqrt{3}},0\right)$.

$H\left(0,\pm\dfrac{1}{\sqrt{3}}\right)=\pm\begin{pmatrix}0 & 4\sqrt{3}\\ 4\sqrt{3} & 0\end{pmatrix}$ でこれは不定符号なので極値を取らない。

$H\left(\dfrac{1}{\sqrt{3}},0\right)=\begin{pmatrix}4\sqrt{3} & 0\\ 0 & 4\sqrt{3}\end{pmatrix}$ は正値である。ゆえに f は狭義の極小値を取る。極小値 $f\left(\dfrac{1}{\sqrt{3}},0\right)=-\dfrac{4\sqrt{3}}{9}$.

$H\left(-\dfrac{1}{\sqrt{3}},0\right)=-\begin{pmatrix}4\sqrt{3} & 0\\ 0 & 4\sqrt{3}\end{pmatrix}$ は負値である。ゆえに f は狭義の極大値を取る。極大値 $f\left(-\dfrac{1}{\sqrt{3}},0\right)=\dfrac{4\sqrt{3}}{9}$.

(4) $z=f(x,y)$ の $(-1,1)$ における接平面の方程式は

$$z=f'(-1,1)\begin{pmatrix}x+1\\ y-1\end{pmatrix}+f(-1,1).$$

$f'(-1,1)=(6+6-2 \quad -12)=(10 \quad -12)$ であるから、右辺は

$$10(x+1)-12(y-1)-6=10x-12y+16.$$

ゆえに接平面の方程式は $z=10x-12y+16$.

$z=f(x,y)$ の (a,b) における接平面の方程式は

$$z=f_x(a,b)(x-a)+f_y(a,b)(y-b)+f(a,b).$$

$$f_x(-1,1)=6+6-2=10,\quad f_y(-1,1)=-12,\quad f(-1,1)=-2-6+2=-6$$

であるから、

$$z=10(x+1)-12(y-1)-6.$$

整理して $z=10x-12y+16$. ■

問 2.10.5 (p. 97) の解答

(1) A の固有値を λ_1, λ_2 とすると、

$$\lambda_1\lambda_2=\det A<0.$$

A が実対称であることから、λ_1 と λ_2 は実数であるから、$\lambda_1\lambda_2<0$ は、λ_1 と λ_2 が異符号であること示している。ゆえに A は正負両方の固有値を持つので, A は不定符号である。

(2) $A=\begin{pmatrix}-1 & 0 & 0\\ 0 & -2 & 0\\ 0 & 0 & -3\end{pmatrix}$ とすると、A の固有値は -1, -2, -3 で、$\det A=(-1)(-2)(-3)<0$. A の固有値がすべて負であるから、A は負値であり、不定符号ではない。

(3) $a\neq 0$ は A が不定符号であるための必要十分条件である。実際 $\det A=-a^2$ に注意すると、

$$a\neq 0 \quad\Leftrightarrow \det A<0 \quad\Leftrightarrow\quad \text{固有値は異符号} \quad\Leftrightarrow\quad A\ \text{は不定符号}.\blacksquare$$

問 2.11.1(p. 105 の解答　(2) $(a,b)=(0,0),(\sqrt{2},0),(-\sqrt{2},0)$. (3) $(x,y)=\left(\pm\frac{\sqrt{3}}{2},\pm\frac{1}{2}\right)$ (複号任意) ■

問 2.12.1(p. 115) の解答　簡単です。自分でやってみましょう。■

問 2.12.4 の答　(1) 高校数学でお馴染の双曲線 ($y=\pm\frac{x}{2}$ が双曲線)。(2) $\sqrt{2}x-2y=2$　(3) $\pm\sqrt{3}$　(4) $\left(-\infty,-\sqrt{3}\right]\cup\left[\sqrt{3},\infty\right))$

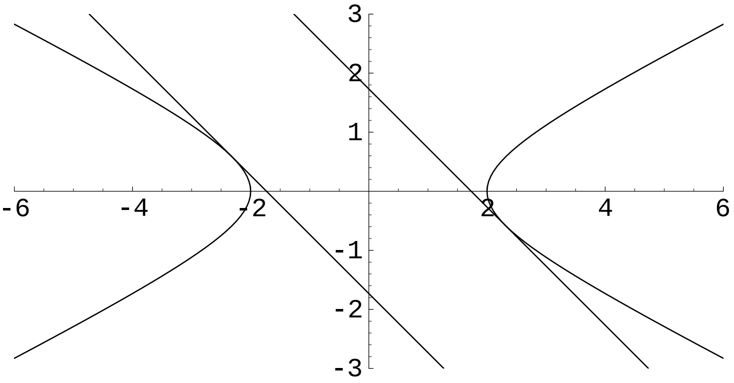


図 2.22: $x^2/4-y^2=1$ と傾きが -1 の接線

Mathematica による図 2.22 の描画

```
Needs["Graphics‘ImplicitPlot‘"]
g = ImplicitPlot[x^2/4 - y^2 == 1, {x, -6, 6},
                  PlotRange -> {{-6, 6}, {-3, 3}},
                  AspectRatio -> 1/2]
t = Plot[{-x + Sqrt[3], -x - Sqrt[3]}, {x, -6, 6}]
Show[g,t]
```

■

問 2.12.6(p. 119) の解答

付 録A　misc

A.1　準備のためのミニ問題集

A.1.1　高校数学から

問 A.1.1 $f(x) = \dfrac{x}{x^2+1}$ とおくとき、$y = f(x)$ のグラフの概形を描きなさい。(増減表を作成すること。また第 2 次導関数 (2 階の導関数) を求め、曲線の凹凸も調べること。$\lim_{x\to\infty} f(x)$ と $\lim_{x\to-\infty} f(x)$ も求めること。)

A.1.2　論理

問 A.1.2 次の論理式の否定を作れ。

(1) $(\forall\varepsilon > 0)\ (\exists N \in \mathbf{N})\ (\forall n \in \mathbf{N},\ n \ge N)\ \|x_n - a\| \le \varepsilon.$

(2) $(\forall x \in A)\ (\exists\varepsilon > 0)$ s.t. $B(x;\varepsilon) \subset A.$

(3) $(\forall\varepsilon > 0)\ (\exists\delta > 0)\ (\forall x \in B(a;\delta))\ \|f(x) - f(a)\| \le \varepsilon.$

(4) $(\exists R \in \mathbf{R})\ (\forall x \in A)\ \|x\| \le R.$

問 A.1.3 連立方程式 $\begin{cases} x(x^2+4y+y^2) = 0 \\ (y+2)(y+x^2) = 0 \end{cases}$ を解け。

A.1.3　集合

問 A.1.4 $A \subset \mathbf{R}^n$, $A \ne \mathbf{R}^n$, $a \in A$ に対して、$\varepsilon := \inf_{y \in A^c} \|a - y\|$, $B_\varepsilon := \{x \in \mathbf{R}^n; \|x - a\| < \varepsilon\}$ とおくと、$B_\varepsilon \subset A$ が成り立つことを示せ。

(B_ε のことを $B(a;\varepsilon)$ と書いても良いかもしれないが、これは普通 a を中心とする半径 ε の開球を表す記号で，$\varepsilon = 0$ の場合に使わない (そのとき $B_\varepsilon = \emptyset$ であり、それを球とは言いづらい)、と考える人が多いかも知れないので、ここでは直接定義を書いた。)

問 A.1.5 次のことを証明せよ。(1) $\displaystyle\bigcap_{n=1}^{\infty} B(0;1/n) = \{0\}$　(2) $\displaystyle\bigcup_{n=1}^{\infty} B(0;n) = \mathbf{R}^n$

A.1.4 解答

問 A.1.1　この問題は高校数学の範疇に属するが、基本の確認ということで入れてある。いわゆる増減の様子以外にも、外して欲しくない常識がある。

(i) 極値を取る変数の値や、座標軸との交点など重要な値はグラフと一緒に書き込む。
→折角計算して調べて増減表に書き込みもしたのに、グラフに反映されていなくて (何のために調べたのだろう？)、変なことになっている人がタマにいます。

(ii) 連続関数のグラフは「切れない」 (正確に言うと連結集合)
→笑い話みたいですが、この問題の場合にも、グラフを切ってしまう人がいたりします。分数関数 (正確には有理関数) は、分母が 0 になるところで、不連続になりうるので、グラフが切れることもありますが、この問題の関数ではそうなりません (当たり前すぎるかもしれないけれど、連続かそうでないか、間違えないように)。

(iii) C^1 級関数のグラフは「滑らか」 (接線の傾きが連続的に変化するので、とがらない)
→こうしてしまう人は結構多い。そうなってしまうのは何か間違えているせいなので、要再チェック、です。(たとえ話をすると、面積を計算して負の値が出たり、確率が 1 より大きくなったりする類いの非常識な結果です。)

(iv) グラフを描く範囲が指定されていない場合は自分で判断する。例えば $\mathbf{R}$ 全体で定義されている場合、$x \to \pm\infty$ でどうなるか調べる。勝手に $x \geq 0$ の範囲に限定したりしない。周期関数の場合は、そのことを指摘して、最低 1 周期分は描く。

(v) 入試の採点をしていて、最近増減表が書けない (途中まで書いてギブアップすることもある)、書いてもそれをグラフに生かせない受験生が多くなったと感じています (凝ったことをしようとして失敗しているケースがあるみたい)。添削して返してあげたいくらいで…

それから、増減表を書かない人も多くなっています。増減表を書かずに、言葉で説明するのは結構大変 (難しい) です。結局は、増減表を書いて、それを使って考えてグラフも描けるようにするのが、一番楽だと思います。高校の先生になる人も多いと思いますが、生徒にしっかり練習させてあげてください。

略解　商の微分法 $\left(\dfrac{F}{G}\right)' = \dfrac{GF' - G'F}{G^2}$ は既知とする (公式は自力で導出できるように)。
まず高校の数学 II までの知識でどこまで出来るかやってみよう。

$$f'(x) = \frac{(2x^2+1)\cdot 1 - 4x\cdot x}{(2x^2+1)^2} = \frac{1-2x^2}{(2x^2+1)^2} = \frac{-2\left(x-\dfrac{1}{\sqrt{2}}\right)\left(x+\dfrac{1}{\sqrt{2}}\right)}{(2x^2+1)^2}$$

であるから、増減表は

x		$-1/\sqrt{2}$		$1/\sqrt{2}$	
$f'(x)$	$-$	0	$+$	0	$-$
$f(x)$	$\searrow$	$-\dfrac{1}{2\sqrt{2}}$	$\nearrow$	$\dfrac{1}{2\sqrt{2}}$	$\searrow$

であり、極小値と極大値はそれぞれ $f\left(-\dfrac{1}{\sqrt{2}}\right) = -\dfrac{1}{2\sqrt{2}}$, $f\left(\dfrac{1}{\sqrt{2}}\right) = \dfrac{1}{2\sqrt{2}}$.
$x \to \infty$, $x \to -\infty$ の時の極限が気になるので調べると、

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} \frac{1/x}{2+1/x^2} = \frac{0}{2+0} = 0, \quad \lim_{x\to-\infty} f(x) = \lim_{x\to-\infty} \frac{1/x}{2+1/x^2} = \frac{0}{2+0} = 0.$$

数学３の知識を用いると、２階導関数 (２次の導関数) を調べて、グラフの凹凸が分かる。

$$\begin{aligned} f''(x) &= \frac{(2x^2+1)^2\cdot(-4x)-2(2x^2+1)4x\cdot(1-2x^2)}{(2x^2+1)^4} = \frac{-4x(2x^2+1)-8x(1-2x^2)}{(2x^2+1)^3} \\ &= \frac{4x(2x^2-3)}{(2x^2+1)^3} = \frac{8x\left(x-\sqrt{\frac{3}{2}}\right)\left(x+\sqrt{\frac{3}{2}}\right)}{(2x^2+1)^3} \end{aligned}$$

であり、増減表は

x	$-\infty$		$-\sqrt{3/2}$		$-1/\sqrt{2}$		0		$1/\sqrt{2}$		$\sqrt{3/2}$		∞
$f'(x)$		$-$	$-$	$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	$-$
$f''(x)$		$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$	
$f(x)$	0	↘	変曲点	↘	$-\frac{1}{2\sqrt{2}}$ 極小	↗	変曲点	↗	$\frac{1}{2\sqrt{2}}$ 極大	↘	変曲点	↘	0

となる[1]。f'' の符号から、$x<-\sqrt{\frac{3}{2}}$ では上に凸、$-\sqrt{\frac{3}{2}}<x<0$ では下に凸、$0<x<\sqrt{\frac{3}{2}}$ では上に凸、$\sqrt{\frac{3}{2}}<x$ では下に凸。

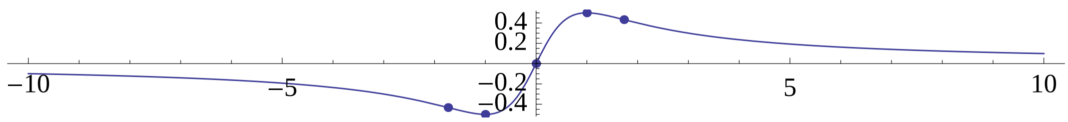


図 A.1: $f(x)=\dfrac{x}{x^2+1}$ のグラフ

問 A.1.2

(1) $(\exists\varepsilon>0)\ (\forall N\in\mathbf{N})\ (\exists n\in\mathbf{N},\, n\geq N)\ \|x_n-a\|>\varepsilon.$

(2) $(\exists x\in A)\ (\forall\varepsilon>0)$ s.t. $B(x;\varepsilon)\not\subset A.$

(3) $(\exists\varepsilon>0)\ (\forall\delta>0)\ (\exists x\in B(a;\delta))\ \|f(x)-f(a)\|>\varepsilon.$

(4) $(\forall R\in\mathbf{R})\ (\exists x\in A)\ \|x\|>R.$

問 A.1.3 $(x,y)=(0,0),(0,-2),(\pm 2,-2),(\pm\sqrt{3},-3).$

問 A.1.4 $x\in B_\varepsilon$ とすると、$\|x-a\|<\varepsilon$. もし $x\in A^c$ であれば、$\|x-a\|\geq\inf_{y\in A^c}\|a-y\|=\varepsilon$ となり矛盾するので、$x\notin A^c$. すなわち $x\in A$. ゆえに $B_\varepsilon\subset A$. ■

与えられた集合 A が開集合であることを定義に基づいて証明するとき、

$$\forall a\in A\quad \exists\varepsilon>0\quad B(a;\varepsilon)\subset A$$

を示すわけであるが、それには普通 ε を具体的に指示することになる。$\varepsilon:=\inf_{y\in A^c}\|a-y\|$ とおくのは、一つの処方箋である。こうすれば (上で証明したことから) $B_\varepsilon=\{x\in\mathbf{R}^n;\|x-a\|<\varepsilon\}\subset A$ である。(なお、ε をこれより大きく取ると $B_\varepsilon\not\subset A$ となることも証明できる。) 後は $\varepsilon>0$ であるかどうかだけが問題である。

[1] ↘, ↘, ↗, ↗ は便利だが、間違える人も多い。慣れていなければ無理して使わないのが無難かもしれない。

A.2 実対称行列の正値性の効率的判定法 (Gaussの消去法作戦)

与えられた実対称行列が正値であるか、負値であるか、判定するための条件として、定理 2.10.10 (次の (a)) を紹介したが、これだけに頼って、素朴に首座小行列式を全部計算してその符号を調べるのは、あまり効率の良い方法ではない。実は、非常に簡単かつ効率的な方法がある。

要点は次の 2 つ。

(a) (大抵の線形代数のテキストに書いてある) n 次実対称行列 A が正値であるためには、$\delta_k := \det A_k > 0$ $(k = 1, \ldots, n)$ が必要十分 (A_k は A の最初の k 行、k 列を取って作った、k 次首座小行列。MATLAB 風に表記すると $A_k = A(1:k, 1:k)$.)

(b) (数値線形代数のテキストには書いてある[2]) 行や列の交換を一切行わない Gauss の消去法で現れる、第 k ピボットを $a_{kk}^{(k)}$ とするとき、

$$a_{11}^{(1)} = a_{11} = \delta_1, \quad a_{kk}^{(k)} = \frac{\delta_k}{\delta_{k-1}} \quad (k = 2, 3, \ldots, n).$$

Gauss の消去法とは

「Gauss の消去法」とは、連立 1 次方程式の解法の一つである。数学の授業では習わないかも知れないが、数値線形代数では常識的な事項であり、コンピューター関係の授業ではしばしば紹介される (A.3 節を見よ)。この方法の前半部分は行列の変形を行うが、それを対称行列の符号の判定に利用できる、ということである。

定理 2.10.10 より、

$$\begin{aligned}
A \text{ が正値} &\Leftrightarrow \delta_k = \det A_k > 0 \quad (k = 1, 2, \cdots, n) \\
&\Leftrightarrow a_{kk}^{(k)} > 0 \quad (k = 1, 2, \cdots, n), \\
A \text{ が負値} &\Leftrightarrow (-1)^k \delta_k = (-1)^k \det A_k > 0 \quad (k = 1, 2, \cdots, n) \\
&\Leftrightarrow a_{kk}^{(k)} < 0 \quad (k = 1, 2, \cdots, n). \blacksquare
\end{aligned}$$

定理としてまとめておく。

定理 A.2.1 (Gauss の消去法による正値性、負値性の判定法) n 次実対称行列 $A = (a_{ij})$ に対して、以下の (1), (2) が成り立つ。

(1) A が正値 $\Longleftrightarrow$ A が行や列の交換なしの Gauss の消去法で上三角行列に変形され、その対角成分がすべて正となる。

(2) A が負値 $\Longleftrightarrow$ A が行や列の交換なしの Gauss の消去法で上三角行列に変形され、その対角成分がすべて負となる。

なお、正則な実対称行列 A に対して、行や列の交換を一切行わないで Gauss の消去法を最後まで実行できたとするとき (これは任意の k に対して $a_{kk}^{(k)} \neq 0$ であることと同値で、任意の k に対して $\delta_k \neq 0$ であることとも同値)、対角成分 $a_{kk}^{(k)}$ の中の正数の個数、負数の個数の組 (p, m) は、A の符号 (正の固有値、負の固有値の個数の組) と一致する。しかし、つねに Gauss の消去法を最後まで実行できるとは限らないので、実対称行列の符号の計算法として

[2] 余分なことがたくさん書いてあって分かり辛いので勧めないが、桂田 [9] にもとりあえず書いてある。

は、必ず答えが得られる方法ではない (実対称行列が正値であったり負値であったりする場合は、Gauss の消去法が最後まで実行できることが保証されることに注意する)。

Gauss の消去法が現れることに唐突な印象を持つかも知れないが、実は 2 次式に対しての基本操作である「平方完成」は Gauss の消去法と関係がある。実対称行列 A を係数とする 2 次形式を、(係数) $\times$ (多項式)^2 の和の形に変形したとき、出て来た係数は Gauss の消去法で得られた上三角行列の対角成分となっている。

このことをまず例で見てみよう。

例 A.2.2

$$A=\begin{pmatrix}1&0&2\\0&1&2\\2&2&-1\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&1&2\\0&2&-5\end{pmatrix}\to\begin{pmatrix}1&0&2\\0&1&2\\0&0&-9\end{pmatrix}.$$

定理 A.2.1 より A は正値でも負値でもない。また行列式は 0 でない (-9 に等しい)。ゆえに不定符号である (実際固有値は $\lambda=1,1,-1$ である)。A を係数とする 2 次形式を平方完成してみよう。

$$\begin{aligned}x^2+y^2-z^2+4xz+4yz&=(x+2z)^2-4z^2+y^2-z^2+4yz\\&=(x+2z)^2+y^2+4yz-5z^2\\&=(x+2z)^2+(y+2z)^2-9z^2.\end{aligned}$$

係数の 1, 1, -9 に注目。

もう一つ、

$$A=\begin{pmatrix}2&1&1\\1&2&1\\1&1&2\end{pmatrix}\to\begin{pmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&\frac{1}{2}&\frac{3}{2}\end{pmatrix}\to\begin{pmatrix}2&1&1\\0&\frac{3}{2}&\frac{1}{2}\\0&0&\frac{4}{3}\end{pmatrix}.$$

やはり定理 A.2.1 より A は正値である (実際、固有値は 4, 1, 1 である)。(念のため、$\det A_1=2>0$, $\det A_2=2^2-1^2=3>0$, $\det A_3=\det A=4>0$ であるから、定理 2.10.10 からも A は正値である。) A を係数とする 2 次形式の平方完成は

$$\begin{aligned}2x^2+2y^2+2z^2+2xy+2yz+2zx&=2(x^2+(y+z)x)+2y^2+2z^2+2yz\\&=2\left(x+\frac{y+z}{2}\right)^2-\frac{(y+z)^2}{2}+2y^2+2z^2-2yz\\&=2\left(x+\frac{y+z}{2}\right)^2+\frac{3y^2}{2}+yz+\frac{3z^2}{2}\\&=2\left(x+\frac{y+z}{2}\right)^2+\frac{3}{2}\left(y^2+\frac{2}{3}yz\right)+\frac{3z^2}{2}\\&=2\left(x+\frac{y+z}{2}\right)^2+\frac{3}{2}\left(y+\frac{z}{3}\right)^2-\frac{z^2}{6}+\frac{3z^2}{2}\\&=2\left(x+\frac{y+z}{2}\right)^2+\frac{3}{2}\left(y+\frac{z}{3}\right)^2+\frac{4}{3}z^2.\end{aligned}$$

係数 2, 3/2, 4/3 が、やはり上の Gauss の消去法の結果の対角成分として現れている。■

$A = (a_{ij})$ が実対称行列であるとする。$a_{11} \neq 0$ であれば

$$a_{11}\left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}} x_j\right)^2 = a_{11}x_1^2 + 2x_1 \sum_{j=2}^n a_{1j}x_j + \frac{1}{a_{11}}\left(\sum_{j=2}^n a_{1j}x_j\right)^2$$
$$= a_{11}x_1^2 + x_1 \sum_{j=2}^n a_{1j}x_j + x_1 \sum_{i=2}^n a_{i1}x_i + \frac{1}{a_{11}}\left(\sum_{i=2}^n a_{i1}x_i\right)\left(\sum_{j=2}^n a_{1j}x_j\right)$$

であるから、

$$a_{11}x_1^2 + \sum_{j=2}^n a_{1j}x_1x_j + \sum_{i=2}^n a_{i1}x_ix_1 = a_{11}\left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}}x_j\right)^2 - \sum_{i,j=2}^n \frac{a_{i1}a_{1j}}{a_{11}}x_ix_j.$$

ゆえに

$$\sum_{i,j=1}^n a_{ij}x_ix_j = a_{11}\left(x_1 + \sum_{j=2}^n \frac{a_{1j}}{a_{11}}x_j\right)^2 + \sum_{i,j=2}^n \left(a_{ij} - \frac{a_{i1}a_{1j}}{a_{11}}\right)x_ix_j.$$

これは Gauss の消去法の前進消去過程と同じである。

前進消去の途中でピボット $(a_{kk}^{(k)})$ に 0 が現れなければ、

$$LA = \begin{pmatrix} a_{11} & & & & * \\ 0 & a_{22}^{(2)} & & & \\ 0 & 0 & a_{33}^{(3)} & & \\ 0 & 0 & \ddots & \ddots & \\ 0 & 0 & & 0 & a_{nn}^{(n)} \end{pmatrix}$$

と上三角化される (L は対角成分が 1 に等しい下三角行列である)。これから

$$\sum_{i,j=1}^n a_{ij}x_ix_j = \sum_{k=1}^n a_{kk}^{(k)} y_k^2$$

が分かる。あるいは

$$LAL^T = \operatorname{diag}[a_{11}, a_{22}^{(2)}, \dots, a_{nn}^{(n)}].$$

例 A.2.3 (吉田・加藤 [10] に載っている例)

$$\begin{aligned} & x_1^2 + \frac{1}{3}x_2^2 + \frac{1}{5}x_3^2 + x_1x_2 + \frac{2}{3}x_1x_2 + \frac{1}{2}x_2x_3 \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3\right)^2 + \left(\frac{1}{3} - \frac{1}{4}\right)x_2^2 + \left(\frac{1}{2} - \frac{1}{3}\right)x_2x_3 + \left(\frac{1}{5} - \frac{1}{9}\right)x_3^2 \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3\right)^2 + \frac{1}{12}x_2^2 + \frac{1}{6}x_2x_3 + \frac{4}{25}x_3^3 \\ &= \left(x_1 + \frac{1}{2}x_2 + \frac{1}{3}x_3\right)^2 + \frac{1}{12}(x_2 + x_3)^2 + \frac{1}{180}x_3^2 \end{aligned}$$

の計算は、次のような Gauss 消去法計算と対応している。

$$\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} & \frac{1}{4} \\ \frac{1}{3} & \frac{1}{4} & \frac{1}{5} \end{pmatrix} \to \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{12} & \frac{1}{12} \\ 0 & \frac{1}{12} & \frac{4}{45} \end{pmatrix} \to \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3} \\ 0 & \frac{1}{12} & \frac{1}{12} \\ 0 & 0 & \frac{1}{180} \end{pmatrix}. \blacksquare$$

A.3 ガウス (Gauss) の消去法のアルゴリズム

連立 1 次方程式の解法として、線形代数の教科書には **クラーメル (Cramer) の公式**や**掃き出し法** (Jordan(ヨルダン) の消去法ともいう) が説明されていることが多いが、ガウスの消去法は、掃き出し法を改良したものである[3]。

例として次の方程式を取りあげて説明しよう。

$$
\left\{
\begin{array}{rcrcrl}
2x_1 & + & 3x_2 & - & x_3 & = 5 \\
4x_1 & + & 4x_2 & - & 3x_3 & = 3 \\
-2x_1 & + & 3x_2 & - & x_3 & = 1.
\end{array}
\right.
\tag{A.1}
$$

掃き出し法では係数行列と右辺のベクトルを並べた行列を作り、それに

1. ある行に 0 でない定数をかける。
2. 2 つの行を入れ換える。
3. ある行に別の行の定数倍を加える。

のような操作 — **行に関する基本変形**と呼ぶ — をほどこして、連立方程式の係数行列に相当する部分を単位行列にするのであった。

$$
\begin{pmatrix} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix}
\to
\begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix}
\to
\begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & -2 & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix}
\to
$$

$$
\to
\begin{pmatrix} 1 & \frac{3}{2} & -\frac{1}{2} & \frac{5}{2} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 6 & -2 & 6 \end{pmatrix}
\to
\begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & -\frac{5}{2} & -\frac{15}{2} \end{pmatrix}
\to
\begin{pmatrix} 1 & 0 & -\frac{5}{4} & -\frac{11}{4} \\ 0 & 1 & \frac{1}{2} & \frac{7}{2} \\ 0 & 0 & 1 & 3 \end{pmatrix}
\to
$$

$$
\to
\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix},
\quad \text{ゆえに} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.
$$

ガウスの消去法も、前半の段階はこの方法に似ていて、同様の変形を用いて掃き出しを行なうのだが、以下のように対角線の下側だけを 0 にする。

$$
\begin{pmatrix} 2 & 3 & -1 & 5 \\ 4 & 4 & -3 & 3 \\ -2 & 3 & -1 & 1 \end{pmatrix}
\to
\begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 6 & -2 & 6 \end{pmatrix}
\to
\begin{pmatrix} 2 & 3 & -1 & 5 \\ 0 & -2 & -1 & -7 \\ 0 & 0 & -5 & -15 \end{pmatrix}.
$$

最後の行列は

$$
2x_1 + 3x_2 - x_3 = 5, \quad -2x_2 - x_3 = -7, \quad -5x_3 = -15
$$

ということを表しているので、後の方から順に

$$
x_3 = \frac{-15}{-5} = 3, \quad x_2 = \frac{-7 + x_3}{-2} = 2, \quad x_1 = \frac{5 - 3x_2 + x_3}{2} = \frac{5 - 3 \times 2 + 3}{2} = 1
$$

と解くことが出来る。前半の対角線の下側を 0 にする掃き出しの操作を**前進消去** (forward elimination)、後半の代入により解の値を求める操作を**後退代入** (backward substitution) と呼ぶ。

[3] 見方によっては、ガウスの消去法は中学校で習う加減法 (初めて習う解法！) そのものであり、大学の線形代数で習う解法は、実用性では退化していると言えなくもない (？)。

A.4 陰関数定理を覚える

結構長いから、段階的に詳細化するのが一つの手である。これを説明してみよう。 陰関数定理とは、授業でも言ったのだが、2 変数の

$$F(x,y)=0$$

を 1 つの変数 (ここで y とする) について

$$y=\varphi(x)$$

のように解くための定理であり、そのために一番重要な仮定が

$$\det\frac{\partial F}{\partial y}(a,b)\neq 0$$

である[4]。つまり、もの凄く乱暴に言うと

第 1 近似

$\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ ならば、$F(x,y)=0$ は $y=\varphi(x)$ と解ける。

これを見ると、(a,b) って一体なんだろう？φ って一体なんだろう？「解ける」とはどういうことか？と疑問が湧いて来る (そうでないといけない)。例えば、まず (a,b) について少し書き足すことにしよう。$F(x,y)=0$ が「全体で」解けることは一般には望めなくて、注目している点の近くだけで解けることくらいしか期待できない。その注目している点が (a,b) ということだ。それは $F(x,y)=0$ の上にある。そこで次のようにする。

第 2 近似

$F(a,b)=0$, $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ ならば、$F(x,y)=0$ は、(a,b) のある近傍で $y=\varphi(x)$ と解ける。

φ というのは、陰関数で、この存在を主張しているのが大事なところ、という話もした。そこでそれをはっきり言ってみよう。

第 3 近似

$F(a,b)=0$, $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ ならば、ある C^1 級の関数 φ が存在して、(a,b) のある近傍で、$F(x,y)=0\Leftrightarrow y=\varphi(x)$ が成り立つ。

大部よくなって来た。採点基準は実は結構甘いのであまり言いたくないが、それによるとこの状態の答案には (満点はやらないが) 結構イイ点がつく、とだけ言っておこう。次がちょっと大変だ。ある近傍と言うのが $U\times V$ である。その U,V というのが、φ については定義域と終域、つまり $\varphi\colon U\to V$ で、U は a の開近傍、V は b の開近傍ということである。これらは一部だけ書いて全部は書かないというのは変なので、次は一気に書くことが増える (と言っても分量で 1 行未満の増加)。

[4] この仮定がもしも覚えにくければ、F が 1 次関数、つまり $F(x,y)=Ax+By+c$ の場合を考えると良いかもしれない。つまり $Ax+By+c=0$ から、$By=-Ax-c$ としておいて、次にやりたいのは B^{-1} を左からかけること。そのためには $\det B\neq 0$ という仮定をおきたい。そして $B=\dfrac{\partial F}{\partial y}$ である。ということで、仮定が $\det\dfrac{\partial F}{\partial t}\neq 0$ であるのはもっともらしい。

第 4 近似

$F(a,b)=0$, $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ ならば、a のある開近傍 U, b のある開近傍 V, ある C^1 級の関数 $\varphi\colon U\to V$ が存在して、$\forall(x,y)\in U\times V$ について、$F(x,y)=0\Leftrightarrow y=\varphi(x)$ が成り立つ。

お好みならば、黒板文体もあるな。

第 4 近似′

$F(a,b)=0$, $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0\Longrightarrow(\exists U\colon a$ のある開近傍$)$ $(\exists V\colon b$ のある開近傍$)$ $(\exists\varphi\colon U\to V$ C^1 級$)$ s.t. $\forall(x,y)\in U\times V\quad F(x,y)=0\Leftrightarrow y=\varphi(x)$.

そろそろ F についても、ちゃんと書かないとまずいだろう。

Ω は $\mathbf{R}^m\times\mathbf{R}^n$ の開集合で、$F\colon\Omega\to\mathbf{R}^n$ は C^1 級の写像とし、$(a,b)\in\Omega$ とする。

というのを書き足す。C^1 というのは、そんなに難しくないであろう。$\dfrac{\partial F}{\partial y}$ や $\dfrac{\partial F}{\partial x}$ が出て来るのだから。x が m 次元ベクトル、y が n 次元ベクトルとするとき、F の値も n 次元ベクトルというのが押えておきたいところである。これも授業中にしゃべったが、そうしておくことで、$\dfrac{\partial F}{\partial y}$ が正方行列になって (そうでないと det も考えられない)、まともな逆が存在する可能性か生じるのである (ここら辺は線形代数がちゃんと身についているかだな)。

第 5 近似

Ω は $\mathbf{R}^m\times\mathbf{R}^n$ の開集合で、$F\colon\Omega\to\mathbf{R}^n$ は C^1 級の写像とし、$(a,b)\in\Omega$ とする。$F(a,b)=0$, $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ ならば、a のある開近傍 U, b のある開近傍 V, ある C^1 級の関数 $\varphi\colon U\to V$ が存在して、$\forall(x,y)\in U\times V$ について、$F(x,y)=0\Leftrightarrow y=\varphi(x)$ が成り立つ。

後は $U\times V\subset\Omega$ を入れるくらいか。導関数の公式 $\varphi'(x)=-\left(\dfrac{\partial F}{\partial y}(x,\varphi(x))\right)^{-1}\dfrac{\partial F}{\partial x}(x,\varphi(x))$ は書いておかなくても、φ が C^1 と分かっていれば後から自前で出せる (出せないといけない)。

第 6 近似

Ω は $\mathbf{R}^m\times\mathbf{R}^n$ の開集合で、$F\colon\Omega\to\mathbf{R}^n$ は C^1 級の写像とし、$(a,b)\in\Omega$ とする。$F(a,b)=0$, $\det\dfrac{\partial F}{\partial y}(a,b)\neq 0$ ならば、a のある開近傍 U, b のある開近傍 V, ある C^1 級の関数 $\varphi\colon U\to V$ が存在して、$\forall(x,y)\in U\times V$ について、$F(x,y)=0\Leftrightarrow y=\varphi(x)$, $U\times V\subset\Omega$ が成り立つ。

これで一応の出来上がり。

A.4.1 気になること、やり残したこと

微分法は積分法やベクトル解析と比べると難しいところが少ないと思われるが、気になるところがないわけではない。

(a) $k\geq 2$ なる k に対して、多変数関数が k 回 (全) 微分可能であることの定義をして、その場合の Taylor の定理の証明をすること。同時に微分係数をテンソルとして捉える必要が出て来るのかもしれない。

(b) 逆関数定理の見通しの良い証明を与えること。(2.11.8 はあちこちで見掛ける出来合いの証明であるが、個人的にはあまりしっくり来ない。解析系の人間なので、反復法で方程式を解く話としてまとめたい。)

(c) 対称行列の符号を判定する一般のアルゴリズムをきちんと紹介すること。

(a) については、シュヴァルツ [5] や、稲葉 [11] などが参考になる。
(c) については、伊理 [12], [13] などが参考になるはずだが、自分流儀で書き直すのに一仕事が必要そうである。

関連図書

[1] 桂田祐史：多変数の微分積分学 1 講義ノート, http://www.math.meiji.ac.jp/~mk/lecture/tahensuu1-2013/tahensuu1-2011.pdf (2011).

[2] E. ハイラー, G. ヴァンナー：解析教程 上，下, シュプリンガー・ジャパン (2006).

[3] H. ザーガン著, 鎌田清一郎訳：空間充填曲線とフラクタル, シュプリンガーフェアラーク東京 (1998), H. Sagan, Space-Filling Curves, Springer (1994) の翻訳.

[4] 高木貞治：解析概論 改訂第 3 版, 岩波書店 (1961).

[5] L. シュヴァルツ：シュヴァルツ解析学 2 微分法, 東京図書 (1970).

[6] 荷見守助：現代解析の基礎演習, 内田老鶴圃 (1993).

[7] 齋藤正彦：線型代数入門, 東京大学出版会 (1966).

[8] M. スピヴァック：スピヴァック多変数の解析学 — 古典理論への現代的アプローチ, 東京図書 (2007), 齋藤正彦訳. 1972 年に出版されたものをお色直しして復刊.

[9] 桂田祐史：発展系の数値解析の続き, http://www.math.meiji.ac.jp/~mk/labo/text/heat-fdm-0-add.pdf (1997 年〜).

[10] 吉田 耕作・加藤敏夫：大学演習 応用数学 I, 裳華房 (1961).

[11] 稲葉三男：新装版 微積分の根底をさぐる, 現代数学社 (2008), 1977 年に出版された。

[12] 伊理正夫：一般線形代数, 岩波書店 (2003), 伊理正夫, 線形代数 I, II, 岩波講座応用数学, 岩波書店 (1993, 1994) の単行本化.

[13] 伊理正夫：線形代数汎論, 朝倉書店 (2009), 「一般線形代数」のリニューアル.

索 引