解析概論 I ノート

桂田 祐史

2002 年 6 月

はじめに

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 (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。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[1]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。

「解析概論 I」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 年次の数学の科目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。これでは他の学問を学ぶ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、この文書である。少々脱線になるかもしれないが、世の中に微分積分学のテキストはたくさんあるのに、この講義でなぜ市販のテキストを教科書として採用しなかったかを説明しておきたい。

一番大きな理由は、1 年次に、数学科の専門科目として微分積分学を学ぶことが出来なかったからである (これは制度の問題である)。担当の諸先生方の努力にもかかわらず、日本の大学の標準 (数学科の学生は、数学を専門科目として学ぶ) からすると、やや不徹底な内容の講義になってしまっている。そのため、この講義では基本的には多変数関数を取り扱うが、1 変数関数についても必要に応じて詳しく説明する必要がある。ところが、大変残念なことに、この教え方に合う教科書があまりないのである[2]。

もう一つの理由は、私がテキストというものは世につれ人につれ、新しく作られるべきものだと信じるからである。 大学 1, 2 年次の講義の内容は基本的なものであり、少しくらい時間が経っても (特に数学の場合) 本質的には大きな変化はないのだが、その時代を生きる学生にどのように伝えるかは、その都度考える必要があると思う。 (例えば、数学科の学生といえども、最近では物理学の素養を持つ者は少なくなったこと[3]、コンピューターが普及したことなど、見過ごせない変化である。またこれは悲しいことだが、厚い本が受けなくなっている御時世である。) そのためかどうか、最近では、テキストもどんどん新しいものが作られている。その中には良い本も多いが、現時点で解析概論 I 用の教科書として納得行くものが見つからなかった。講義の内容を自分で取捨選択して作るのはやむをえない、と思った。

昔から大学の講義では「教科書なし」というものも珍しくない。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

[1]Newton の主著と言えるプリンキピア・マテマティカ (Principia Mathematica Philosophiae Naturalis, 1687) は Kepler の法則の証明ただ一つを目的に書かれているという意見がある。そうすると最初の本格的な応用は「力学の構築」というよりは「Kepler の法則の証明」と言うべきかもしれない。

[2]1 変数関数の微積分の記述と、多変数関数の微積分の記述の独立性が高く、数学的にきちんと書かれた本が何故かあまりない。

[3]ぴんと来ない学生が多くなったと思うので補足しておくと、以前は理科系の学生は、ほぼ例外なく高校で二年間みっちりと物理学を学んで来た。

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。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激減しているので、板書にはかなり注意を払わねばならなくなった。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、と考えたのが、この文書「解析概論 I ノート」の始まりであった。

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、

- 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。

という理由がある。

このノートを利用する場合の注意　要するに、講義が主であって、ノートは補助的なものである、ということを言いたい。

このノートは、講義の内容にかなり忠実に書かれているが、

- 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
- 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、かなり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多く、かなり読み難い**

というのは残念ながら真実なので、本文 100 ページ弱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、その時間に頭を回転さ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[4](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。

2002 年度版への注意　いつの間にか解析概論 I ノートも長くなってしまい、コピーしての配布に負担を感じるようになった。そこで今年度は基本的に本文のみ配布することにする (将来的には WWW オンリーにしようかと考えている)。付録に何があるか、目次には書いてあるので、読みたい人は

`http://www.math.meiji.ac.jp/~mk/lecture/kaisekigairon-1/`

をアクセスして下さい。

[4] 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

第1章　R^n の性質

この章では微分法に必要な $\mathbf{R}^n$ の性質をかけ足で説明する。

1.1　なぜ R^n を考える必要があるか

一つの理由は、

多変数関数を扱うための基礎とするためである

例 1.1.1 ある瞬間の温度を考える。場所によって異なるので、場所の関数である。

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

3 つの変数 x_1, x_2, x_3 についての関数、3 変数関数である。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。

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

(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 変数関数になる。)

集合、写像の言葉を使って書くと、n 変数関数とは、 $\mathbf{R}^n$ のある部分集合 A 上定義され、$\mathbf{R}$ に値を取る写像

$$f\colon \mathbf{R}^n \supset A \longrightarrow \mathbf{R}$$

のことである。

より一般には

ベクトル値関数が考えられる

例 1.1.2 ある瞬間の風の速度は, 位置 $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ を変数とするベクトル値の関数である:

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

つまり多変数ベクトル値関数 (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 A 上で定義され、 $\mathbf{R}^m$ に値を取る写像

$$\vec{g}\colon \mathbf{R}^n \supset A \longrightarrow \mathbf{R}^m$$

のことである。

つまり多変数関数やベクトル値関数を考えることにすると、必然的に舞台は数ベクトル空間 $\mathbf{R}^n$ になる。

1.2 n 次元ユークリッド空間 $\mathbf{R}^n$

定義 1.2.1 (内積空間としての $\mathbf{R}^n$ の定義) $n \in \mathbf{N}$ に対して

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

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。

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

注意 1.2.1 (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 1.2.2 (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、多くの本では $(\vec{x},\vec{y})$ 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。

$\mathbf{R}^n$ 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 $\vec{}$ 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。)

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る。

命題 1.2.1 (内積の公理) $\mathbf{R}^n$ の内積 $(\cdot,\cdot)$ は次の性質を満たす。

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

$$(x,x) = 0 \iff x = 0.$$

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

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

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

$$(x,y) = (y,x).$$

定理 1.2.1 (Schwarz (シュヴァルツ) の不等式) $\mathbf{R}^n$ の内積 $(\cdot,\cdot)$ は次の性質を満たす。$\forall x \in \mathbf{R}^n, \forall y \in \mathbf{R}^n$ に対して

$$(x,y)^2 \leq (x,x)(y,y). \tag{1.1}$$

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 x と y が線形従属である (片方がもう一方のスカラー倍である) ことである。

定理 1.2.1 の証明

(1) x, y が線形独立な場合。
$\forall t \in \mathbf{R}$ に対して $tx + y \neq 0$ であるから、

$$(tx+y, tx+y) > 0.$$

ゆえに

$$(x,x)t^2 + 2(x,y)t + (y,y) > 0.$$

t についての 2 次式の符号が変わらないことから

$$\frac{判別式}{4} = (x,y)^2 - (x,x)(y,y) < 0.$$

(2) x, y が線形従属な場合。
次のいずれかが成り立つ。

(1) $x = ty$ となる $t \in \mathbf{R}$ が存在する。

(2) $y = tx$ となる $t \in \mathbf{R}$ が存在する。

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

定義 1.2.2 ($\mathbf{R}^n$ のノルム) $x \in \mathbf{R}^n$ に対して、x の**ノルム**(**norm**, 長さ、大きさ) $\|x\|$ を

$$\|x\| \overset{\text{def.}}{=} \sqrt{(x,x)}$$

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

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}.$$

この記号を用いると、Schwarz の不等式 (1.2.1) は

$$|(x,y)| \le \|x\|\|y\|$$

とも表される。

問 上の不等式の絶対値を外すと

$$-\|x\|\|y\| \le (x,y) \le \|x\|\|y\|$$

という不等式が得られるが、等号が成立するのはどういう場合か調べよ。

命題 1.2.2 (ノルムの公理) $\mathbf{R}^n$ のノルム $\|\cdot\|$ は次の性質を満たす。

(i) $\forall x \in \mathbf{R}^n$ に対して

$$\|x\| \ge 0.$$

$\forall x \in \mathbf{R}^n$ に対して

$$\|x\| = 0 \iff x = 0.$$

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

$$\|\lambda x\| = |\lambda|\|x\|.$$

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

$$\|x+y\| \le \|x\| + \|y\|$$ (**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、x と 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。

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、2 乗した両辺を比較すればいい。

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

等号成立の条件については読者に任せる。■

例題 1.2.1 (逆三角不等式)

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

を示せ。

解答　(両辺共に正だから、自乗して比較しても良い[1]。) ここでは、三角不等式から導く。

$$\|x\|=\|x-y+y\|\leq\|x-y\|+\|y\|$$

から

(1.2) $$\|x\|-\|y\|\leq\|x-y\|.$$

x と y を入れ換えて

$$\|y\|-\|x\|\leq\|y-x\|,$$

すなわち

(1.3) $$-\|x-y\|\leq\|x\|-\|y\|.$$

(1.2) と (1.3) をまとめると

$$-\|x-y\|\leq\|x\|-\|y\|\leq\|x-y\|.$$

ゆえに

$$\Big|\|x\|-\|y\|\Big|\leq\|x-y\|. \quad ■$$

定義 1.2.3 (R^n のユークリッド距離) $x\in\mathbf{R}^n$, $y\in\mathbf{R}^n$ に対して、その間の**距離** (**distance**) $d(x,y)$ を

$$d(x,y)\overset{\text{def.}}{=}\|x-y\|=\sqrt{\sum_{i=1}^n(x_i-y_i)^2}$$

で定義する。

[1] この場合、内積の性質を使って証明することができる。

命題 1.2.3 (距離の公理) $\mathbf{R}^n$ の距離 $d = d(\cdot, \cdot)$ は次の性質を満たす。

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

$$d(x, y) \geq 0.$$

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

$$d(x, y) = 0 \Longleftrightarrow x = y.$$

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

$$d(x, y) = d(y, x).$$

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

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

注意 1.2.3 せっかく $d(\cdot, \cdot)$ という記号を定義したのだけれど、$d(x, y)$ よりも $\|x - y\|$ の方が簡潔だから、以後この講義では使わない。

1.3 行列のノルム

行列についてもノルムを考えることがある。
行列 $A = (a_{ij}) \in M(m, n; \mathbf{R})$ のノルム $\|A\|$ を次式で定める:

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

命題 1.3.1 $\|\cdot\|$ はノルムの公理を満たす。すなわち

(i) 任意の $A \in M(m, n; \mathbf{R})$ に対して

$$\|A\| \geq 0, \quad \|A\| = 0 \Leftrightarrow A = O.$$

(ii) 任意の $A \in M(m, n; \mathbf{R})$, $\lambda \in \mathbf{R}$ に対して

$$\|\lambda A\| = |\lambda| \|A\|.$$

(iii) 任意の $A \in M(m, n; \mathbf{R})$, $B \in M(m, n; \mathbf{R})$ に対して

$$\|A + B\| \leq \|A\| + \|B\|.$$

証明　$M(n,n;\mathbf{R})$ はノルムまで込めて、自然に $\mathbf{R}^{mn}$ と同一視出来ることから明らかである。■

定理 1.3.1 (1) 任意の $A \in M(m,n;\mathbf{R})$, $x \in \mathbf{R}^n$ に対して

$$\|Ax\| \le \|A\|\|x\|.$$

(2) 任意の $A \in M(m,n;\mathbf{R})$, $B \in M(n,\ell;\mathbf{R})$ に対して

$$\|AB\| \le \|A\|\|B\|.$$

証明　(1) Schwarz の不等式を利用して

$$\|Ax\|^2 = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}x_j\right)^2 \le \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 \sum_{j=1}^n x_j^2\right) = \left(\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2\right)\sum_{j=1}^n x_j^2 = \|A\|^2\|x\|^2.$$

(2) については、(1) と同様。■

例 1.3.1 (1 次写像の連続性) 上の定理から、1 次写像 (アファイン写像)

$$f\colon \mathbf{R}^n \ni x \longmapsto Ax + b \in \mathbf{R}^m$$

(ただし $A \in M(m,n,\mathbf{R})$, $b \in \mathbf{R}^m$) の連続性が見通しよく証明できる。実際

$$\|f(x) - f(a)\| = \|(Ax+b) - (Aa+b)\| = \|A(x-a)\| \le \|A\|\|x-a\|$$

であるから、$x \to a$ のとき $f(x) \to f(a)$ であることは明らかである。■

余談　行列 A のノルムを

$$\|A\| \overset{\text{def.}}{=} \sup_{x \in \mathbf{R}^n \setminus \{0\}} \frac{\|Ax\|}{\|x\|} = \sup_{\|x\|=1} \|Ax\| = \max_{\|x\| \le 1} \|Ax\|$$

と定義することもある (これを**作用素ノルム**と呼ぶ)[2]。作用素ノルムに対しても上の定理は成立するし、さらに

$$\|\ \text{単位行列}\ \| = 1$$

が成り立ち、便利である。

1.4　R^N の点列とその収束

これまで $\mathbf{R}^n$ と書いてきたが、この小節では数ベクトルの列 (点列とよぶ) を考え、番号として文字 n を使いたいので、空間の次元の方を大文字の N として、$\mathbf{R}^N$ を考えることにする。

また、列の n 番目の項 x_n を第 n 成分の記号と混乱させないため、ベクトルであることを強調する矢印 $\vec{}$ をつけて $\vec{x}_n$ と記すことにする (これはしばしば成分を用いて議論しなければならないこの節の中だけのこと)。

[2]むしろこの作用素ノルムを用いる方が普通かもしれない。

1.4.1 定義と簡単な性質

定義 1.4.1 ($\mathbf{R}^N$ の点列) $\mathbf{R}^N$ から無限個の点を取り出して、番号をつけたもの

$$\vec{x}_1, \vec{x}_2, \cdots, \vec{x}_n, \cdots$$

を $\mathbf{R}^N$ の**無限点列** (あるいは単に**点列** (**sequence**)) と呼び、$\{\vec{x}_n\}_{n=1}^{\infty}$ や $\{\vec{x}_n\}_{n\in\mathbf{N}}$ などの記号で表す。

注意 1.4.1 要するに $\mathbf{R}^N$ の点列とは $\mathbf{N}$ から $\mathbf{R}^N$ への写像である:

$$\mathbf{N} \ni n \longmapsto \vec{x}_n \in \mathbf{R}^N.$$

注意 1.4.2 点列を表すための括弧 $\{\quad\}$ は集合を表す括弧と形は同じだが、意味は違うものであることに注意しよう。$\mathbf{R}^1$ の点列としては

$$\{1,2,1,2,1,2,\cdots\} \neq \{2,1,2,1,2,1,\cdots\}$$

であるが、集合としては

$$\{1,2,1,2,1,2,\cdots\} = \{2,1,2,1,2,1,\cdots\} = \{1,2\}.$$

この紛らわしさを避けるために、点列を $(x_n)_{n\in\mathbf{N}}$ のような記号で表している本もある (筆者も割とこの流儀は好きなのだが、この講義では使わない)。

注意 1.4.3 $\mathbf{R}^N$ の点列があるということは N 個の実数列がある、ということである。つまり $\vec{x}_n \in \mathbf{R}^N$ の成分を

$$\vec{x}_n = \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_N^{(n)} \end{pmatrix}$$

と書くことにすると、

$$\{x_1^{(n)}\}_{n\in\mathbf{N}}, \{x_2^{(n)}\}_{n\in\mathbf{N}}, \cdots, \{x_N^{(n)}\}_{n\in\mathbf{N}}$$

という N 個の実数列が得られる。

定義 1.4.2 (点列の収束、極限) $\{\vec{x}_n\}_{n\in\mathbf{N}}$ を $\mathbf{R}^N$ の点列、$\vec{a} \in \mathbf{R}^N$ とするとき、

$$\{\vec{x}_n\}_{n\in\mathbf{N}} \text{ が } \vec{a} \text{ に}\textbf{収束}\text{する} \overset{\text{def.}}{\Leftrightarrow} \lim_{n\to\infty} \|\vec{x}_n - \vec{a}\| = 0.$$

このとき

$$\lim_{n\to\infty} \vec{x}_n = \vec{a}$$

あるいは

$$\vec{x}_n \to \vec{a} \quad (n \to \infty)$$

と書き、$\vec{a}$ を点列 $\{\vec{x}_n\}_{n\in\mathbf{N}}$ の**極限**(きょくげん) (**limit**) と呼ぶ。また「点列 $\{\vec{x}_n\}_{n\in\mathbf{N}}$ は**収束列** (**convergent sequence**) である」、「極限 $\lim_{n\to\infty} \vec{x}_n$ が存在する」、ともいう。

注意 1.4.4 (1) $N=1$ のときは「点列 = 数列」であり、収束、極限などという言葉が重なるが、内容が一致するので、混乱は起こらない。

(2) ε-N 論法を使えば、

$$\{\vec{x}_n\}_{n\in\mathbf{N}} \text{ が } \vec{a} \text{ に収束する} \overset{\text{def.}}{\Leftrightarrow} (\forall\varepsilon>0)\ (\exists\ell\in\mathbf{N})\ (\forall n\in\mathbf{N}\colon n\geq\ell)\ \|\vec{x}_n-\vec{a}\|\leq\varepsilon.$$

命題 1.4.1 (点列の収束は成分ごとに考えればよい) $\mathbf{R}^N$ の点列の収束は、各成分の作る数列の収束と同値である。つまり

$$\vec{x}_n=\begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_N^{(n)} \end{pmatrix},\quad \vec{a}=\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}$$

とおくと、

$$\lim_{n\to\infty}\vec{x}_n=\vec{a} \iff \text{すべての } i\ (1\leq i\leq N) \text{ について } \lim_{n\to\infty}x_i^{(n)}=a_i.$$

この命題の証明は、任意の $\vec{y}\in\mathbf{R}^N$ に対して成り立つ不等式

$$\max_{i=1,2,\cdots,N}|y_i|\leq\|\vec{y}\|\leq\sum_{i=1}^N|y_i|$$

を $\vec{y}=\vec{x}_n-\vec{a}$ に適用した

$$\max_{i=1,2,\cdots,N}|x_i^{(n)}-a_i|\leq\|\vec{x}_n-\vec{a}\|\leq|x_1^{(n)}-a_1|+|x_2^{(n)}-a_2|+\cdots+|x_N^{(n)}-a_N|$$

を用いれば簡単に証明できる。

要するに

$$\lim_{n\to\infty}\begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_N^{(n)} \end{pmatrix}=\begin{pmatrix} \lim_{n\to\infty}x_1^{(n)} \\ \lim_{n\to\infty}x_2^{(n)} \\ \vdots \\ \lim_{n\to\infty}x_N^{(n)} \end{pmatrix}.$$

ということである。

例 1.4.1

$$\vec{x}_n=\begin{pmatrix} 1+\frac{1}{n} \\ 1-\frac{1}{n^2} \end{pmatrix}$$

とすると

$$\lim_{n\to\infty}\vec{x}_n=\begin{pmatrix} \lim_{n\to\infty}\left(1+\dfrac{1}{n}\right) \\ \lim_{n\to\infty}\left(1-\dfrac{1}{n^2}\right) \end{pmatrix}=\begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

点列の収束については、数列の収束と同様の多くの命題が成り立つ。証明抜きでいくつか列挙しておく。

命題 1.4.2 (和、スカラー乗法、内積、ノルムの連続性) $\mathbf{R}^N$ の点列 $\{\vec{x}_n\}_{n\in\mathbf{N}}$, $\{\vec{y}_n\}_{n\in\mathbf{N}}$, と数列 $\{\lambda_n\}_{n\in\mathbf{N}}$ が、いずれも収束列であるとき、

(1) $\lim_{n\to\infty}(\vec{x}_n+\vec{y}_n)=\lim_{n\to\infty}\vec{x}_n+\lim_{n\to\infty}\vec{y}_n.$

(2) $\lim_{n\to\infty}(\lambda_n\vec{x}_n)=\lim_{n\to\infty}\lambda_n\lim_{n\to\infty}\vec{x}_n.$

(3) $\lim_{n\to\infty}(\vec{x}_n,\vec{y}_n)=(\lim_{n\to\infty}\vec{x}_n,\lim_{n\to\infty}\vec{y}_n).$

(4) $\lim_{n\to\infty}\|\vec{x}_n\|=\|\lim_{n\to\infty}\vec{x}_n\|.$

命題 1.4.3 (収束列の有界性) $\mathbf{R}^N$ の収束列は有界である。すなわち $\{\vec{x}_n\}_{n\in\mathbf{N}}$ が $\mathbf{R}^N$ の収束列ならば、

$$\exists M\in\mathbf{R}\quad s.t.\quad \forall n\in\mathbf{N}\quad \|\vec{x}_n\|\le M.$$

命題 1.4.4 (極限の一意性) $\mathbf{R}^N$ の収束列の極限は一意的である (収束先はただ一つしかない)。

命題 1.4.5 (収束列の部分列は収束列) $\mathbf{R}^N$ の収束列の任意の部分列は収束列である。

1.4.2 Cauchy 列と完備性

解析学では、様々なものを点列や関数の「極限」として見い出す (これが解析学の核心である、という意見が多数派を占めるようである)。解析学にとって、極限が存在することを保証する定理は非常に重要である。

この小節では「$\mathbf{R}^N$ 内の Cauchy 列は必ず極限を持つ」という定理 ($\mathbf{R}^N$ の完備性) を説明する。

定義 1.4.3 ($\mathbf{R}^N$ の Cauchy 列) $\mathbf{R}^N$ の点列 $\{\vec{x}_n\}_{n\in\mathbf{N}}$ が **Cauchy 列**(コーシー)[a]であるとは、

$$(\forall\varepsilon>0)(\exists\ell\in\mathbf{N})(\forall n\in\mathbf{N}: n\geq\ell)(\forall m\in\mathbf{N}: m\geq\ell)\quad \|\vec{x}_n-\vec{x}_m\|\leq\varepsilon$$

が成り立つことである。このことを

$$\lim_{n,m\to\infty}\|\vec{x}_n-\vec{x}_m\|=0$$

と書くこともある (あまり勧められない書き方である)。

[a] Cauchy 列のことを**基本列**と呼ぶこともある。

命題 1.4.6 (収束列は Cauchy 列) $\mathbf{R}^N$ 内の任意の収束列は Cauchy 列である。

証明 $\{x_n\}_{n\in\mathbf{N}}$ は収束列で、その極限を a とする。収束の定義から、$\varepsilon>0$ を任意に与えられた正数とすると、

$$(\exists\ell\in\mathbf{N})\ (\forall n\in\mathbf{N}\colon\ n\geq\ell)\quad \|x_n-a\|\leq\varepsilon/2.$$

が成り立つ。$m\in\mathbf{N}$ が $m\geq\ell$ を満たせば

$$\|x_m-a\|\leq\varepsilon/2$$

となるから、

$$\begin{aligned}\|x_n-x_m\| &= \|x_n-a+a-x_m\|\\ &\leq \|x_n-a\|+\|a-x_m\|\\ &\leq \varepsilon/2+\varepsilon/2=\varepsilon.\quad\blacksquare\end{aligned}$$

実は $N=1$ の場合、すなわち $\mathbf{R}$ については、この逆が成立することが知られていた。

補題 1.4.1 ($\mathbf{R}$ の完備性) $\mathbf{R}$ 内の任意の Cauchy 列は収束列である。

これはこの講義では証明抜きで認めることにする。任意の Cauchy 列が収束するような距離空間は**完備**(かんび)(**complete**) であると言われるので、上の補題は「$\mathbf{R}$ は完備である」と言い替えられる。

命題 1.4.7 (「有理数体 $\mathbf{Q}$ は完備でない」) $\mathbf{Q}$ 内の Cauchy 列で、$\mathbf{Q}$ 内では収束しないものがある。

証明 $\{x_n\}_{n\in\mathbf{N}}$ を

$$x_n=\sqrt{2}\text{ を 10 進小数に展開して小数 } n \text{ 位まで取ったもの}$$

として定義する。つまり

$$x_1 = 1.4, \quad x_2 = 1.41, \quad x_3 = 1.414, \quad x_4 = 1.4142, \quad x_5 = 1.41421, \cdots$$

もちろん $x_n \in \mathbf{Q}$ である。また $n \geq m$ なる $n \in \mathbf{N}$, $m \in \mathbf{N}$ に対して

$$0 \leq x_n - x_m = 0.\overbrace{00\cdots0}^{m\text{ 個}}***\cdots \leq 10^{-m}.$$

これから $\{x_n\}_{n\in\mathbf{N}}$ は $\mathbf{Q}$ の Cauchy 列である。しかし $\mathbf{Q}$ 内では収束しない。(もし $a \in \mathbf{Q}$ に収束したとすると、$\mathbf{R}$ でも a に収束する。極限の一意性から $a = \sqrt{2}$. これは $\sqrt{2} \notin \mathbf{Q}$ であることに反する。) ■

（「冷暖房完備」という言葉があるが、今ここで考えている「完備」は「収束先がすべて備わっている」というニュアンスがある。$\mathbf{Q}$ には、ぼこぼこ穴が空いているわけである。）

定理 1.4.1 ($\mathbf{R}^N$ の完備性) $\mathbf{R}^N$ の任意の Cauchy 列は収束列である。

証明 $\{\vec{x}_n\}_{n\in\mathbf{N}}$ を $\mathbf{R}^N$ の Cauchy 列とする。$\vec{x}_n = \begin{pmatrix} x_1^{(n)} \\ x_2^{(n)} \\ \vdots \\ x_N^{(n)} \end{pmatrix}$ とおこう。任意の i $(1 \leq i \leq N)$ に対して

$$\left| x_i^{(n)} - x_i^{(m)} \right| \leq \|\vec{x}_n - \vec{x}_m\| \quad (n, m \in \mathbf{N})$$

が成り立つから、$\{x_i^{(n)}\}_{n\in\mathbf{N}}$ は $\mathbf{R}$ の Cauchy 列である。ゆえに、補助定理 1.4.1 から収束列である。その極限を a_i として、$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_N \end{pmatrix}$ とおくと

$$\|\vec{x}_n - \vec{a}\| = \sqrt{\sum_{i=1}^{N} (x_i^{(n)} - a_i)^2} \to 0 \quad (n \to \infty).$$

すなわち $\lim_{n\to\infty} \vec{x}_n = \vec{a}$. ゆえに $\{\vec{x}_n\}_{n\in\mathbf{N}}$ は収束列である。 ■

よって、命題 1.4.6 と定理 1.4.1 をまとめて、

$\mathbf{R}^N$ において、収束列 $=$ Cauchy 列。

点列が収束列であることを示すよりも、Cauchy 列であることを示す方がずっと簡単なことが多いので、上の定理はとても役に立つ。

1.4.3 Bolzano-Weierstrass の定理

点列が「有界」というだけの比較的緩い条件のもとでの極限の存在 (ただし部分列を取るという操作は必要である) を保証する Bolzano-Weierstrass の定理を説明する。

定義 1.4.4 ($\mathbf{R}^N$ の有界集合, 有界点列) (1) $\mathbf{R}^N$ の部分集合 X が**有界**(ゆうかい) (**bounded**) であるとは、

$$\exists R \in \mathbf{R} \text{ s.t. } \forall x \in X \quad \|x\| \le R$$

が成り立つことである。

(2) $\mathbf{R}^N$ の点列 $\{\vec{x}_n\}_{n\in\mathbf{N}}$ が**有界**であるとは、

$$\exists M \in \mathbf{R} \text{ s.t. } \forall n \in \mathbf{N} \quad \|\vec{x}_n\| \le M$$

が成り立つことである。

次の定理も $N=1$ の時 (つまり $\mathbf{R}$ の場合) は知っている。

定理 1.4.2 (Bolzano-Weierstrass の定理) $\mathbf{R}^N$ 内の任意の有界点列は収束部分列を含む。

証明　記号が繁雑になるのを避けるため、$N=2$ として証明する (一般の次元でも本質は同じである)。$\{\vec{x}_n\}_{n=1}^{\infty}$ を $\mathbf{R}^N=\mathbf{R}^2$ 内の有界点列とする。成分を

$$\vec{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$$

のように書こう。

$$|x_n| \le \|\vec{x}_n\|, \quad |y_n| \le \|\vec{x}_n\| \quad (n \in \mathbf{N})$$

であることと、仮定から、$\{x_n\}_{n=1}^{\infty}$, $\{y_n\}_{n=1}^{\infty}$ は共に $\mathbf{R}$ 内の有界数列である。まず、 $\{x_n\}_{n=1}^{\infty}$ から収束部分列 $\{x_{n_k}\}_{k=1}^{\infty}$ を取り出し、$\lim_{k\to\infty} x_{n_k} = a$ とおく。次に $\{y_{n_k}\}_{k=1}^{\infty}$ を考える。これはやはり有界数列であるから、収束部分列 $\{y_{n_{k_j}}\}_{j=1}^{\infty}$ を取り出し、$\lim_{j\to\infty} y_{n_{k_j}} = b$ とおく。すると、当然 $\lim_{j\to\infty} x_{n_{k_j}} = a$ であるから、

$$\lim_{j\to\infty} \vec{x}_{n_{k_j}} = \begin{pmatrix} a \\ b \end{pmatrix}$$

となる。■

例題 1.4.1 $\vec{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$ を

$$x_n = (-1)^n + \frac{1}{n}, \quad y_n = \sin\frac{2n\pi}{3} + \frac{1}{n^2}$$

で定めるとき、$\{\vec{x}_n\}_{n\in\mathbf{N}}$ が $\mathbf{R}^2$ の有界な点列であることはすぐに分かるが、収束部分列を実際に取り出して見よ。

解答　$\{x_n\}_{n\in\mathbf{N}}$ は収束しないが、

$$x_2, x_4, x_6, \cdots, x_{2k}, \cdots$$

と偶数番目の項を取ると、

$$x_{2k} = 1 + \frac{1}{2k}$$

ゆえ、これは 1 に収束する。つまり $n_k = 2k$ としたわけ。次に $\{y_{n_k}\}_{k\in\mathbf{N}}$ を考えると、$y_{n_k} = \sin\frac{4k\pi}{3} + \frac{1}{4k^2}$ ゆえ、これも収束しないが、 3 項おきに取ると、 0 に収束することが分かる。つまり $n_{k_j} = 6j$ とすると、$\{\vec{x}_{n_{k_j}}\}_{j\in\mathbf{N}} = \begin{pmatrix} 1+\frac{1}{6j} \\ \frac{1}{36j^2} \end{pmatrix}$ ゆえ、これは $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ に収束する。■

1.5　$\mathbf{R}^n$ の位相

$\mathbf{R}^n$ の点列の極限の定義と基本的な性質については既に説明したが、一般に関数や点列の極限、ひいては関数の連続性を定義するために、**位相**（いそう） (**topology**) と呼ばれるものがある。位相の定義には色々な流儀があるが、現在では開集合を定義するやり方がよく使われている (もっとも、解析学では点列などの極限を用いて、位相を特徴づける方が便利なことも多い)。開集合の定義を理解し、有界閉集合の性質に慣れることが大事である。

1.5.1　開集合

まず $\mathbf{R}^n$ 内の「球」を定義する。

定義 1.5.1 ($\mathbf{R}^n$ の開球と閉球)　$a \in \mathbf{R}^n$, $r > 0$ とするとき、

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

を中心 a, 半径 r の**開球**（かいきゅう） (**open ball**) と呼ぶ。また、

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

を中心 a, 半径 r の**閉球**（へいきゅう） (**closed ball**) と呼ぶ。

例 1.5.1　$n=1$ の場合、開球は開区間であり、閉球は閉区間である: $B(a;r) = (a-r, a+r)$, $\overline{B}(a;r) = [a-r, a+r]$. また $n=2$ の場合は、開球は開円板、閉球は閉円板。$n=3$ の時は、ホントの球の内部と縁（ふち）付きの球。

定義 1.5.2 ($\mathbf{R}^n$ の開集合) A を $\mathbf{R}^n$ の部分集合とする。

(1) A が $\mathbf{R}^n$ の**開集合**(**開部分集合**, **open subset**) であるとは、$\forall x \in A$ に対して

$$\exists \varepsilon > 0 \text{ s.t. } B(x;\varepsilon) \subset A \tag{1.4}$$

が成り立つことである。

(2) 条件 (1.4) を満たす点 x を A の**内点** (**inner point**) と呼ぶ。

(3) A の内点全体を A° と書き、A の**内部** (**interior**) と呼ぶ。つまり

$$A^\circ = \{x \in \mathbf{R}^n; \exists \varepsilon > 0 \text{ s.t. } B(x;\varepsilon) \subset A\}.$$

例 1.5.2 (開球は開集合) 任意の開球 $A = B(a;r)$ は開集合である。実際、 $x \in A$ とするとき、$\|x-a\| < r$ であるから、

$$\varepsilon \overset{\text{def.}}{=} r - \|x-a\|$$

とおくと $\varepsilon > 0$ であり、$B(x;\varepsilon) \subset A$. なぜなら $y \in B(x;\varepsilon)$ とすると $\|x-y\| < \varepsilon$ だから、

$$\|y-a\| \le \|y-x\| + \|x-a\| < \varepsilon + \|x-a\| = r$$

となるので $y \in B(a;r)$. したがって、 $B(x;\varepsilon) \subset B(a;r) = A$. (以上の議論を、図示して納得すること。) ■

(この例の証明は有名で学ぶに価するものだが、付録 I で紹介する別証明も見ておくことを勧める。)

例 1.5.3 (閉球は開集合ではない) 閉球 $A = \overline{B}(a;r)$ は開集合ではない。実際 $x \overset{\text{def.}}{=} a + \begin{pmatrix} r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ とおくと $\|x-a\| = r$ ゆえ、 $x \in \overline{B}(a;r) = A$. ところが、$\varepsilon > 0$ をどんなに小さく取っても、$B(x;\varepsilon) \not\subset A$. (図を描いて説明する方が簡単だが: $y \overset{\text{def.}}{=} x + \begin{pmatrix} \varepsilon/2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ とおくと、$\|y-x\| = \varepsilon/2 < \varepsilon$ より $y \in B(x;\varepsilon)$. また $\|y-a\| = r + \varepsilon/2 > r$ より $y \notin \overline{B}(a;r) = A$.) ■

その他の例は、付録 I「開集合と閉集合に関するメモ」を見よ。

命題 1.5.1 (開集合系の公理) $\mathbf{R}^n$ の開集合全体の集合 $\mathcal{O}$ について、次の (1), (2), (3) が成立する。

(1) 空集合 $\emptyset \in \mathcal{O}$, 全空間 $\mathbf{R}^n \in \mathcal{O}$.

(2) $\mathbf{R}^n$ の部分集合の族 $\{U_\lambda\}_{\lambda\in\Lambda}$ に対して、$U_\lambda \in \mathcal{O}$ $(\forall\lambda \in \Lambda)$ ならば、合併 $\displaystyle\bigcup_{\lambda\in\Lambda} U_\lambda \in \mathcal{O}$.

(3) $U_1 \in \mathcal{O}, U_2 \in \mathcal{O}$ ならば、$U_1 \bigcap U_2 \in \mathcal{O}$.

証明 (1) 空集合は一つも要素を持たないので、確かに

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

は成立している (付録 E.5「空集合の論理」を参照せよ)。ゆえに空集合は開集合である。次に明らかに

$$\forall x \in \mathbf{R}^n \quad B(x;1) \subset \mathbf{R}^n$$

であるから ($\varepsilon = 1$ として条件が成り立ち) $\mathbf{R}^n$ は開集合である。
(2) $x \in \bigcup_{\lambda\in\Lambda} U_\lambda$ とすると、

$$\exists\lambda_0 \in \Lambda \quad \text{s.t.} \quad x \in U_{\lambda_0}.$$

U_{λ_0} は開集合であるから

$$\exists\varepsilon > 0 \quad \text{s.t.} \quad B(x;\varepsilon) \subset U_{\lambda_0}.$$

ゆえに

$$B(x;\varepsilon) \subset U_{\lambda_0} \subset \bigcup_{\lambda\in\Lambda} U_\lambda.$$

ゆえに $\bigcup_{\lambda\in\Lambda} U_\lambda$ は開集合である。
(3) $x \in U_1 \cap U_2$ とする。まず $x \in U_1$ で、U_1 は開集合であるから

$$\exists\varepsilon_1 > 0 \quad \text{s.t.} \quad B(x;\varepsilon_1) \subset U_1.$$

同様に $x \in U_2$ で、U_2 は開集合であるから

$$\exists\varepsilon_2 > 0 \quad \text{s.t.} \quad B(x;\varepsilon_2) \subset U_2.$$

そこで $\varepsilon \overset{\text{def.}}{=} \min(\varepsilon_1, \varepsilon_2)$ とおくと $\varepsilon > 0$ で

$$B(x;\varepsilon) \subset B(x;\varepsilon_1) \cap B(x;\varepsilon_2) \subset U_1 \cap U_2.$$

ゆえに $U_1 \cap U_2$ は開集合である。■

例題 1.5.1 $\mathbf{R}^n$ の部分集合の族 $\{U_\lambda\}$ について、$U_\lambda \in \mathcal{O}$ $(\forall\lambda \in \Lambda)$ とするとき、$\displaystyle\bigcap_{\lambda\in\Lambda} U_\lambda \in \mathcal{O}$ か?

答　Λ が有限集合ならばそうだが (これは数学的帰納法で簡単に証明できる)、無限集合の場合はそうとは限らない。例えば、$\Lambda = \mathbf{N}$, 各 $n \in \mathbf{N}$ に対して $U_n = B(0; 1/n)$ とおくと、$U_n \in \mathcal{O}$ だが、$\displaystyle\bigcap_{n \in \mathbf{N}} U_n = \{0\}$ は $\mathcal{O}$ に属さない。 ■

1.5.2　内部、外部、境界

さて、ここで、もう少し詳しく、$\mathbf{R}^n$ の部分集合の外部、境界について考えよう。例えば $A = [0,1] \times (0,1)$ とした場合、内部、外部、境界が何であるか、直観的には明らかであるが[3]、数学的にはどう定義すれば良いだろうか? 点の分類をするための準備として、次の命題を掲げる。

命題 1.5.2 (内点、外点、境界点の分類) $a \in \mathbf{R}^N$, $A \subset \mathbf{R}^N$ とするとき、次の 3 つの条件のうち、いずれか一つだけが必ず成立する。

(i) $\exists \varepsilon > 0$ s.t. $B(a; \varepsilon) \subset A$. (「$a$ を中心とする、十分小さな球は A に含まれる。」)

(ii) $\exists \varepsilon > 0$ s.t. $B(a; \varepsilon) \subset A^c$. (「$a$ を中心とする、十分小さな球は A の補集合に含まれる。」)

(iii) $\forall \varepsilon > 0$ $B(a; \varepsilon) \cap A \neq \emptyset$ かつ $B(a; \varepsilon) \cap A^c \neq \emptyset$. (「$a$ を中心とする、任意の球は A にも A の補集合にも含まれない。」)

注意 1.5.1 (補集合の記号) $\mathbf{R}^N$ の任意の部分集合 A に対して、A^c は**補集合**とする。つまり $A^c \overset{\text{def.}}{=} \mathbf{R}^N \setminus A$ と定義する。(この記号は、よく使われるが、それほど標準的というわけではないので、使う前には断りが必要である。)

命題 1.5.2 の証明　まず (i), (ii) が同時に成り立つことはないのは明らかである。

$$\begin{aligned}
\text{(i) でも (ii) でもない} \quad &\Leftrightarrow \forall \varepsilon > 0\ B(a; \varepsilon) \not\subset A \text{ かつ } B(a; \varepsilon) \not\subset A^c \\
&\Leftrightarrow \forall \varepsilon > 0\ B(a; \varepsilon) \cap A^c \neq \emptyset \text{ かつ } B(a; \varepsilon) \cap A \neq \emptyset. \\
&\Leftrightarrow \text{(iii) が成り立つ。}
\end{aligned}$$

より、(i), (ii), (iii) のうちいずれか一つだけが必ず成り立つ。 ■

注意 1.5.2 上の証明中、次の事実を利用した: 「集合 X の任意の部分集合 A, B に対して

$$A \subset B \iff A \cap B^c = \emptyset,$$

さらには

$$A \not\subset B \iff A \cap B^c \neq \emptyset$$

が成り立つ。」

[3]念のために書いておくと、内部は境界を含まない正方形 $(0,1) \times (0,1)$, 外部は全平面から境界を込めた正方形を抜いた集合 $\mathbf{R}^2 \setminus [0,1] \times [0,1]$, 境界は正方形の 4 つの辺の合併 $\{(x,0); 0 \le x \le 1\} \cup \{(x,1); 0 \le x \le 1\} \cup \{(0,y); 0 \le y \le 1\} \cup \{(1,y); 0 \le y \le 1\}$ である。

定義 1.5.3 (外点、外部、境界点、境界) $a \in \mathbf{R}^n$, $A \subset \mathbf{R}^n$ とする。上の (i) が成り立つような a のことを A の内点と呼ぶのだったが、 (ii) が成り立つような a のことを A の**外点**(がいてん)(**exterior point**)、(iii) が成り立つような a のことを A の**境界点** (**boundary point**) と呼ぶ。また、A の外点全体を A の**外部** (**exterior**)、A の境界点全体を A の**境界**(**boundary**) と呼ぶ。

この節では、A の外部を A^e, A の境界を A^b で表すことにする (一般に通じる記号ではない):

$$A^e \overset{\text{def.}}{=} A \text{ の外部}, \quad A^b \overset{\text{def.}}{=} A \text{ の境界}.$$

上の命題 1.5.2 からすぐに次の系が得られる。

系 1.5.1 (内部+境界+外部=全体) $\mathbf{R}^n$ の任意の部分集合 A に対して、$\mathbf{R}^n$ は A の内部、A の境界、A の外部の和 (共通部分のない合併) になる:

$$\mathbf{R}^n = A^\circ \cup A^b \cup A^e, \quad A^\circ \cap A^b = A^b \cap A^e = A^e \cap A^\circ = \emptyset.$$

また定義から簡単に次の二つの命題が得られる。

系 1.5.2 $\mathbf{R}^n$ の任意の部分集合 A に対して、A の内点は A の要素である。従って、$A^\circ \subset A.$

証明　a を A の内点とすると、定義から $\exists \varepsilon > 0$ s.t. $B(a;\varepsilon) \subset A.$ ところが $a \in B(a;\varepsilon)$ であるから $a \in A.$ ■

系 1.5.3 (外部は補集合の内部) A を $\mathbf{R}^n$ の部分集合とする時、 A の外点とは、A の補集合 $\mathbf{R}^n \setminus A$ の内点である。従って、A の外部は、A の補集合の内部である。

証明　命題 1.5.2 の条件 (i), (ii) をよく見れば、外点は補集合の内点であることは明らか。 ■

次の命題は、この講義では使わないので証明はしないが、事実だけでも知っていると便利かも知れない[4]。

命題 1.5.3 (内部と開集合) $\mathbf{R}^n$ の任意の部分集合 A について、次の (1), (2) が成り立つ。

(1) A の内部は、A に含まれる最大の開集合である。

(2) A が開集合であるための必要十分条件は、A が A の内部と一致することである:

$$A \text{ が開集合} \Leftrightarrow A = A^\circ.$$

[4]証明は特にむつかしくない。位相の説明をしてある本には大抵載っている。

1.5.3　閉集合

開集合と対になる概念である閉集合を定義しよう。

定義 1.5.4 ($\mathbf{R}^n$ の閉集合) $\mathbf{R}^n$ の部分集合 A が**閉集合**(**閉部分集合**, **closed subset**) であるとは、A の補集合 $A^c = \mathbf{R}^n \setminus A$ が開集合であることと定義する。

命題 1.5.4 (閉集合系の公理) $\mathbf{R}^n$ の閉集合全体の集合 $\mathcal{F}$ について、次の (1), (2), (3) が成立する。

(1) 空集合 $\emptyset \in \mathcal{F}$, 全空間 $\mathbf{R}^n \in \mathcal{F}$.

(2) $\mathbf{R}^n$ の部分集合の族 $\{F_\lambda\}_{\lambda\in\Lambda}$ に対して、$F_\lambda \in \mathcal{F}$ $(\forall\lambda \in \Lambda)$ ならば、共通部分 $\displaystyle\bigcap_{\lambda\in\Lambda} F_\lambda \in \mathcal{F}$.

(3) $F_1 \in \mathcal{F}$, $F_2 \in \mathcal{F}$ ならば、$F_1 \bigcup F_2 \in \mathcal{F}$.

証明　命題 1.5.1 と閉集合の定義、それと De Morgan (ド・モルガン) の法則

$$\left(\bigcap_{\lambda\in\Lambda} F_\lambda\right)^c = \bigcup_{\lambda\in\Lambda} F_\lambda^c, \quad \left(\bigcup_{\lambda\in\Lambda} F_\lambda\right)^c = \bigcap_{\lambda\in\Lambda} F_\lambda^c$$

からすぐ証明できる。■

定義 1.5.5 ($\mathbf{R}^n$ の部分集合の接触点、閉包、閉集合) $A \subset \mathbf{R}^n$, $a \in \mathbf{R}^n$ とする。

(1) a が A の**接触点**(**触点**) であるとは、

$$\forall\varepsilon > 0 \quad B(a;\varepsilon) \cap A \neq \emptyset.$$

が成り立つことである。

(2) A の接触点全体を A の**閉包**(へいほう) (**closure**) と呼び、$\overline{A}$ で表す。つまり

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

系 1.5.4 $\mathbf{R}^n$ の任意の部分集合 A に対して、A の点は A の接触点である。従って、$A \subset \overline{A}$.

証明　a を A の点とすると、$\forall\varepsilon > 0$ に対して $B(a;\varepsilon) \cap A \neq \emptyset$ $(\because a \in B(a;\varepsilon) \cap A)$ である。ゆえに a は A の接触点である。■

系 1.5.5 $\mathbf{R}^n$ の任意の部分集合 A に対して、A の境界点は A, A^c の接触点である。従って $A^b \subset \overline{A}$, $A^b \subset \overline{A^c}$.

証明　a を A の境界点とすると、定義から $\forall\varepsilon>0$ に対して $B(a;\varepsilon)\cap A\neq\emptyset$ かつ $B(a;\varepsilon)\cap A^c\neq\emptyset$. 前の方の条件は a が A の接触点であることを示し、後の方の条件は a が A^c の接触点であることを示す。■

閉包が何であるか、次の特徴づけを知っておくと良いだろう。

命題 1.5.5 (閉包=内部+境界) A を $\mathbf{R}^n$ の部分集合とすると、A の閉包は A の内部と A の境界の合併である。

証明　A の内点は A の接触点である (定義から明らか)。また系 1.5.5 で述べたように、A の境界点は A の接触点である。ゆえに

$$A^\circ\cup A^b\subset\overline{A}.$$

一方、 A の外点は A の接触点にはなりえない。ゆえに

$$A^e\cap\overline{A}=\emptyset.\quad\therefore\overline{A}\subset\mathbf{R}^n\setminus A^e.$$

系 1.5.1 から $\mathbf{R}^n\setminus A^e=A^\circ\cup A^b$ であるから、

$$\overline{A}\subset A^\circ\cup A^b.$$

以上から

$$\overline{A}=A^\circ\cup A^b.\quad\blacksquare$$

次の命題は、この講義では使わないので証明はしないが、事実だけでも知っていると便利かも知れない

命題 1.5.6 (閉集合と閉包) $\mathbf{R}^n$ の任意の部分集合 A について、次の (1), (2) が成り立つ。

(1) A の閉包は、A を含む最小の閉集合である。

(2) A が閉集合であるための必要十分条件は、A が A の閉包と一致することである:

$$A\text{ が閉集合}\Leftrightarrow A=\overline{A}.$$

$\mathbf{R}^n$ 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 1.5.7 (閉集合の点列による特徴づけ) $A\subset\mathbf{R}^n$ とするとき、

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

証明　いずれも背理法で証明する。

(⇒) $\{x_n\}_{n\in\mathbf{N}}$ 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。 ■

お勧め　次の表を念頭において、少し前まで戻って読み返してみると悟りが得られるかもしれない。

開集合	内部
閉集合	閉包

1.5.4　コンパクト集合

位相空間においてはコンパクト性という大変重要な概念がある。ここではまず $\mathbf{R}^n$ の部分集合がコンパクトであるための条件について述べる (証明はさぼる)。点列コンパクト性という「いかにも便利そうな」性質との同値性が得られるが、コンパクトという概念の重要性は、後の定理 2.1.1, 定理 2.1.2 で一層明らかになるであろう。

定理 1.5.1 (コンパクト集合の特徴づけ) $\mathbf{R}^n$ の空でない部分集合 A に対する次の三つの条件は同値である。

(i) A は有界閉集合である。

(ii) A の元からなる任意の点列 $\{\vec{x}_n\}_{n\in\mathbf{N}}$ は収束部分列を持ち、その極限は A に属する。(このことを A は**点列コンパクト** (**sequentially compact**) である、という。)

(iii) A は Heine-Borel(ハイネ ボレル) の条件「A の任意の開被覆は有限部分被覆を持つ」を満たす。(このことを A は**コンパクト**(**compact**) である、という。)

証明　(i), (ii) の同値性のみ証明する。
(i) $\Longrightarrow$ (ii) A は有界閉集合であると仮定する。 $\{x_n\}_{n\in\mathbf{N}}$ を A 内の任意の点列とすると、A が有界であるから、 $\{x_n\}_{n\in\mathbf{N}}$ も有界である。従って Bolzano-Weierstrass の定理 1.4.2 より、収束部分列 $\{x_{n_k}\}_{k\in\mathbf{N}}$ が存在する: $\displaystyle\lim_{k\to\infty} x_{n_k} = a$. A は閉集合と仮定したから、命題 1.5.7 より $a \in A$. ゆえに A は点列コンパクトである。
(ii) $\Longrightarrow$ (i)

A が有界であること　背理法を用いる。A が有界でないと仮定すると、$\|x_n\| > n$ を満たす A 内の点列 $\{x_n\}_{n\in\mathbf{N}}$ が取れる。これはいかなる収束部分列も含み得ないので、矛盾する。ゆえに A は有界である。

A が閉集合であること　A 内の点列 $\{x_n\}_{n\in\mathbf{N}}$ が収束すれば、その極限 a は、必ず A に含まれることを示す。点列コンパクト性の仮定より、部分列 $\{x_{n_k}\}_{k\in\mathbf{N}}$ と $a' \in A$ が存在して、$\lim_{k\to\infty} x_{n_k} = a'$. ところで、当然[5] $\lim_{k\to\infty} x_{n_k} = a'$ であり、極限の一意性から $a = a'$. よって、$a \in A$. ■

注意 1.5.3 (色々な言葉があるわけ) 色々な用語が出てきて、混乱しそうな人もいるだろう。特に、同じことを色々な表現で言い換えるのは馬鹿馬鹿しいようだが、実はこれらの概念は $\mathbf{R}^n$ の部分集合以外の様々な対象物に拡張され、一般には一致しないものなのである。これらの言葉をすぐに全部覚えられなくても構わないが、必要に応じて復習すると良い。

問.　上の定理の条件 (ii) は、Bolzano-Weierstrass の定理の結論とどこが違うか?

答.　極限が A に含まれること。(「閉集合である」という仮定が増えたから、結論も増えたわけである。)

注意 1.5.4 (この節を終えるにあたっての言い訳) 位相の話は、詳しくやると非常に長いものになってしまう。この節の記述は決して簡潔とは言えないが、これでも煩わしくならないようにかなり簡略にしてある。例えば、「集積点」、「近傍」、「相対位相」などの重要な概念の説明はさぼっている (どれもあれば便利なのだが、それを言い出すとキリがないので涙を飲んで省略した)。一方で、初めて学ぶ人にはそれでもゴチャゴチャしているように感じられるかもしれない。

1.6　$\mathbf{R}^n$ 内の曲線

連続な 1 変数のベクトル値関数を曲線と呼ぶ ⋯ そのためにまずベクトル値関数の極限を定義することから始めよう。

[5]真面目に理由を述べると「収束列の部分列は収束列である」から。

定義 1.6.1 (1 変数ベクトル値関数の極限) I を $\mathbf{R}$ の区間, $t_0 \in \overline{I}$, $\vec{\varphi}\colon I \to \mathbf{R}^n$, $\vec{a} \in \mathbf{R}^n$ とするとき、

$$t \to t_0 \text{ のとき、} \vec{\varphi}(t) \text{ が } \vec{a} \text{ に収束する} \overset{\text{def.}}{\Leftrightarrow} \lim_{t \to t_0} \|\vec{\varphi}(t) - \vec{a}\| = 0.$$

このとき、

$$\lim_{t \to t_0} \vec{\varphi}(t) = \vec{a}$$

あるいは

$$\vec{\varphi}(t) \to \vec{a} \quad (t \to t_0)$$

と書き、a を、$\vec{\varphi}(t)$ の $t \to t_0$ のときの**極限**と呼ぶ。また「極限 $\lim_{t \to t_0} \vec{\varphi}(t)$ が存在する」ともいう。

命題 1.6.1 (ベクトル値関数の収束は成分ごとに考えればよい) ベクトル値関数の収束は、各成分の作る数列の収束と同値である。つまり

$$\vec{\varphi}(t) = \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix}, \quad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

とおくと、

$$\lim_{t \to t_0} \vec{\varphi}(t) = \vec{a} \iff \text{すべての } i \ (1 \le i \le n) \text{ について } \lim_{t \to t_0} \varphi_i(t) = a_i.$$

要するに $\lim_{t \to t_0} \begin{pmatrix} \varphi_1(t) \\ \varphi_2(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix} = \begin{pmatrix} \lim_{t \to t_0} \varphi_1(t) \\ \lim_{t \to t_0} \varphi_2(t) \\ \vdots \\ \lim_{t \to t_0} \varphi_n(t) \end{pmatrix}$ ということであり、命題 1.4.1 のベクトル値関数版である (証明も同じである)。この命題から、次の命題は明らかであろう。

命題 1.6.2 (和、スカラー乗法、内積、ノルムの連続性) $\mathbf{R}$ の区間 I 上定義されたベクトル値関数 $\vec{\varphi}: I \to \mathbf{R}^n$, $\vec{\psi}: I \to \mathbf{R}^n$, 実数値関数 $\lambda: I \to \mathbf{R}$ が、$t \to t_0$ (ただし $t_0 \in \overline{I}$) のとき収束するならば、

(1) $\lim_{t\to t_0}(\vec{\varphi}(t) + \vec{\psi}(t)) = \lim_{t\to t_0}\vec{\varphi}(t) + \lim_{t\to t_0}\vec{\psi}(t)$.

(2) $\lim_{t\to t_0}(\lambda(t)\vec{\varphi}(t)) = \lim_{t\to t_0}\lambda(t)\lim_{t\to t_0}\vec{\varphi}(t)$.

(3) $\lim_{t\to t_0}(\vec{\varphi}(t), \vec{\psi}(t)) = (\lim_{t\to t_0}\vec{\varphi}(t), \lim_{t\to t_0}\vec{\psi}(t))$.

(4) $\lim_{t\to t_0}\|\vec{\varphi}(t)\| = \|\lim_{t\to t_0}\vec{\varphi}(t)\|$.

定義 1.6.2 (曲線) I を $\mathbf{R}$ の区間、$t_0 \in I$, $\varphi: I \to \mathbf{R}^n$ とする。

(1) φ が t_0 で**連続** $\overset{\text{def.}}{\Leftrightarrow}$ $\lim_{I\ni t\to t_0}\varphi(t) = \varphi(t_0)$ すなわち $\lim_{I\ni t\to t_0}\|\varphi(t) - \varphi(t_0)\| = 0$.

(2) φ が I で連続 (φ が I 上の連続なベクトル値関数、φ が $\mathbf{R}^n$ における曲線) $\overset{\text{def.}}{\Leftrightarrow}$ $\forall t \in I$ φ は t で連続。(つまり、**曲線とは、数直線上の区間で定義された連続関数のこと**である。)

(3) φ が t_0 で**微分可能** (differentiable) $\overset{\text{def.}}{\Leftrightarrow}$ $\lim_{I\ni t\to t_0}\dfrac{1}{t-t_0}(\varphi(t) - \varphi(t_0))$ が存在
この極限を $\varphi'(t_0)$ あるいは $\dfrac{d\varphi}{dt}(t_0)$ と書き、 φ の t_0 における**微分係数** (**differential coefficient**)、あるいは φ の点 $\varphi(t_0)$ における**接ベクトル** (**tangent vector**) という。

(4) φ が I で微分可能 $\overset{\text{def.}}{\Leftrightarrow}$ $\forall t \in I$ に対して φ は t で微分可能。
このとき、写像 $\varphi': I \ni t \mapsto \varphi'(t) \in \mathbf{R}^n$ を φ の 1 **階導関数** (**first derivative**, derived function) と呼ぶ。

(5) φ' も曲線だから、連続、微分可能、導関数 $(\varphi')'$ が考えられる。φ' が I 上微分できるとき、φ は I 上 2 **回微分可能** (**two-times differentiable**) という。このとき、$(\varphi')'$ を φ'' あるいは $\varphi^{(2)}$ と書き、 φ の 2 **階導関数**(**second derivative**, **second derived function**) と呼ぶ。以下、任意の自然数 k について帰納的に k **回微分可能**、k **階導関数** (**k-th derivative**, **k-th derived function**) $\varphi^{(k)}$ が定義できる。

(6) $k \geq 1$ とするとき、 φ が I で C^k**-級** (k **回連続的微分可能**, **k-times continuously differentiable**, **function of class** C^k) $\overset{\text{def.}}{\Leftrightarrow}$ φ が I で k 階微分可能、かつ k 階導関数 $\varphi^{(k)}$ が I で連続。
φ が I で連続である時、 C^0**-級**であると言う。φ の 0 **階導関数**は φ 自身のことであるとする: $\varphi^{(0)} = \varphi$.

(7) φ が I で C^∞**-級** (**無限回微分可能、無限回連続的微分可能**, **infinitely differentiable**) $\overset{\text{def.}}{\Leftrightarrow}$ 任意の自然数 k に対し φ は I で C^k-級。

問　上で定義した色々な条件相互の関係を図示せよ。

注意 1.6.1 日常用語の感覚からすると、φ の像 $\varphi(I) = \{\varphi(t); t \in I\}$ のことを曲線と呼びたいが、ここでは写像 φ そのものを曲線と呼んでいる (現代の数学としては普通の流儀である)。また、この曲線の定義では、(曲がっていない) 直線も曲線に入ってしまうことに注意しておこう。

注意 1.6.2 (よくある勘違いについて) 連続的微分可能とは、連続かつ微分可能なことでは**ない** (微分可能ならば連続なので、連続でかつ微分可能というような冗長なことは言うはずがない)。微分可能で、導関数が連続なことである。

注意 1.6.3 (1 変数ベクトル値関数は早い話が) 成分毎(ごと)に考えればよい。例えば、

$$\varphi(t) = \begin{pmatrix} \varphi_1(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix}$$

とすると、

$$\varphi \text{ が } t_0 \text{ で連続} \Leftrightarrow \text{各 } i\ (1 \le i \le n) \text{ に対して } \varphi_i \text{ が } t_0 \text{ で連続。}$$

$$\varphi \text{ が } t_0 \text{ で微分可能} \Leftrightarrow \text{各 } i\ (1 \le i \le n) \text{ に対して } \varphi_i \text{ が } t_0 \text{ で微分可能。}$$

例 1.6.1 (等速円運動) $n = 2$, r は正定数、ω は実定数とするとき、$\varphi(t) = \begin{pmatrix} r\cos\omega t \\ r\sin\omega t \end{pmatrix}$ とおくと、

$$\varphi'(t) = \begin{pmatrix} -r\omega\sin\omega t \\ r\omega\cos\omega t \end{pmatrix}, \quad \varphi''(t) = \begin{pmatrix} -r\omega^2\cos\omega t \\ -r\omega^2\sin\omega t \end{pmatrix},$$

$$\|\varphi'(t)\| = r|\omega|.$$

t を時刻、$\varphi(t)$ を時刻 t における質点の位置を表すものと解釈すると、$\varphi'(t)$ は速度、$\|\varphi'(t)\|$ は速さ、$\varphi''(t)$ は加速度となる。■

　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 1.6.2 (Peano(ペアノ) 曲線 (Peano curve)) 連続曲線 $\varphi\colon I = [0,1] \to \mathbf{R}^2$ で、像が正方形である:

$$\varphi(I) = [0,1] \times [0,1]$$

ようなものが存在する。

注意 1.6.4 (滑らかな曲線 — 正則曲線) 本によっては、C^1-級の関数のことを「滑らか」と言うことがあるが、C^1-級の曲線の像 (図形としての集合、形) は滑らかであるとは限らず、角(かど)を持つ (尖(とが)っている) こともありうる (例えば $x = t^3$, $y = t^2$, $t \in [-1,1]$)。こういう角は特異点 ($\varphi'(t_0) = 0$ となる点のこと) になっているので、特異点のない C^1-級曲線と言えば、ホントに像も滑らかになる。特異点のない C^1-級曲線のことを、**正則曲線**(せいそくきょくせん) (**regular curve**) ということがある。

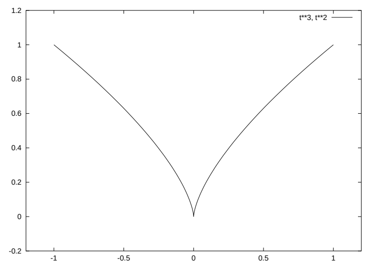


図 1.1: C^1-級だが滑らかでない曲線

命題 1.6.3 (曲線と微分の公式) I, J を $\mathbf{R}$ の区間、$\varphi\colon I \to \mathbf{R}^n$, $\psi\colon I \to \mathbf{R}^n$ を微分可能な曲線、 $f\colon I \to \mathbf{R}$, $g\colon J \to I$ を微分可能な関数とするとき、以下の (1) 〜 (5) が成り立つ。

(1) $\dfrac{d}{dt}(\varphi(t) \pm \psi(t)) = \dfrac{d\varphi}{dt}(t) \pm \dfrac{d\psi}{dt}(t)$ (複号同順).

(2) $\dfrac{d}{dt}(f(t)\varphi(t)) = \dfrac{df}{dt}(t)\varphi(t) + f(t)\dfrac{d\varphi}{dt}(t)$.

(3) $\dfrac{d}{dt}(\varphi(t), \psi(t)) = \left(\dfrac{d\varphi}{dt}(t), \psi(t)\right) + \left(\varphi(t), \dfrac{d\psi}{dt}(t)\right)$.

(4) $\dfrac{d}{dt}\varphi(g(t)) = g'(t)\varphi'(g(t))$.

例題 1.6.1 速さ一定の運動では、速度と加速度は互いに直交することを示せ。

解答 φ が速さ一定とする。すなわち

$$\|\varphi'(t)\| \equiv C \quad (C \text{ は定数}).$$

これから

$$(\varphi'(t), \varphi'(t)) \equiv C^2.$$

両辺を t で微分すると、

$$(\varphi''(t), \varphi'(t)) + (\varphi'(t), \varphi''(t)) = 0$$

すなわち

$$(\varphi''(t), \varphi'(t)) = 0.$$

これは速度 φ' と加速度 φ'' が直交することを示す。 ■

C^1-級曲線の長さは、簡単に定義できる。

定義 1.6.3 (C^1-級曲線の長さ) C^1-級曲線 $\varphi\colon I \overset{\text{def.}}{=} [a,b] \to \mathbf{R}^n$ の長さを

$$\int_a^b \|\varphi'(t)\| dt$$

であると定義する。$\varphi(t) = \begin{pmatrix} \varphi_1(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix}$ とすれば、$\displaystyle\int_a^b \sqrt{\varphi_1'(t)^2 + \cdots + \varphi_n'(t)^2}\, dt.$

曲線を質点の運動と解釈すれば、速さを時間で積分したものが、曲線の長さということになる。

「曲線の長さはパラメーターの取り方によらない」 $\varphi\colon I = [a,b] \to \mathbf{R}^n$ を C^1-級の曲線、$f\colon [\alpha,\beta] \to [a,b]$ を C^1-級の関数で、$f' > 0$, $f(a) = \alpha$, $f(b) = \beta$ を満足するものとする。このとき、合成写像

$$\Phi \overset{\text{def.}}{=} \varphi \circ f\colon [\alpha,\beta] \ni s \mapsto \varphi(f(s)) \in \mathbf{R}^n$$

も C^1-級の曲線である。このとき、

$$\begin{aligned} \Phi \text{ の長さ} &= \int_\alpha^\beta \|\Phi'(s)\| ds \\ &= \int_\alpha^\beta \|\varphi'(f(s)) f'(s)\| ds = \int_\alpha^\beta \|\varphi'(f(s))\| f'(s) ds \\ &= \int_a^b \|\varphi'(t)\| dt = \varphi \text{ の長さ}. \end{aligned}$$

φ の像と Φ の像は、点集合 (図形) としては等しいわけで、心情的には同じ曲線であるが、現在の数学の流儀では、曲線とは写像のことで、そのセンで行くと、Φ と φ は一応は別物である。でも、その長さはめでたく等しい、と言うわけである。この事実を

曲線の長さは、パラメーターの取り方によらない

と言う。

さて、φ を正則曲線とするとき、

$$\ell(t) \overset{\text{def.}}{=} \int_a^t \|\varphi'(\eta)\| d\eta \quad (t \in [a,b])$$

とおくと、

$$\ell\colon [a,b] \longrightarrow [0,L] \quad C^1\text{-級},$$
$$\ell(a) = 0, \quad \ell(b) = L,$$
$$\ell'(t) = \|\varphi'(t)\| > 0 \quad (t \in [a,b]),$$

が成り立つ (ここで L は曲線 φ の長さである)。この ℓ の逆関数を ψ としよう:

$$\psi\colon [0,L] \ni s \mapsto \psi(s) \in [a,b]$$

こうして作った $\Phi \overset{\text{def.}}{=} \varphi \circ \psi\colon [0,L] \to \mathbf{R}^n$ は、$\|\Phi'(s)\| \equiv 1$ という性質をもつ。Φ は弧長 s によりパラメーターづけられた曲線であるという。

曲線の長さのより一般的な定義、長さを持つ曲線、長さを持たない曲線　C^1-級でない曲線の長さはどう考えたらよいだろうか？ここでは詳しく述べないが、曲線上に節点を持つ、折れ線で曲線を近似して、その折れ線の長さの上限として定義する。注意すべきは、すべての曲線に対して、この意味の長さが定義できるとは限らないことである (折れ線の長さの (有限な) 上限が存在せず、いくらでも長さの大きい折れ線が取れることがある)。そこで、この意味の長さが定義できる曲線のことを、「**長さを持つ曲線** (**rectifiable curve**)」と呼び、そうでないものと区別する。長さを持たない曲線の存在は早くから知られていたが、最近では**フラクタル** (**fractal**) の典型例として現われるため、詳しく研究されている[6]。次に示すのは、長さを持たない曲線として有名な **Koch 曲線** (Koch curve) である。この図形 Γ は、

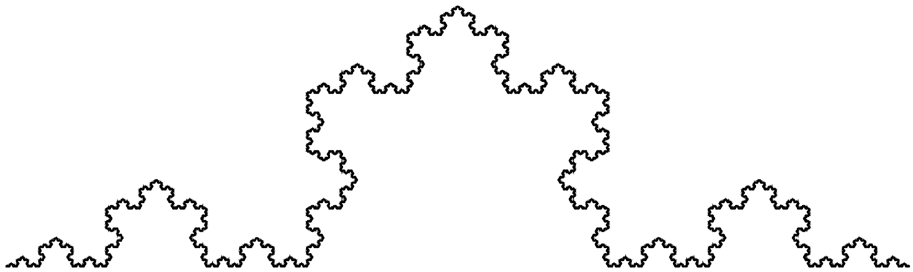


図 1.2: Koch 曲線

$$\Gamma = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4, \quad \Gamma_1 \equiv \Gamma_2 \equiv \Gamma_3 \equiv \Gamma_4$$

と分解できるが、実は各小部分 Γ_i は、全体 Γ と相似で、相似比は

$$\Gamma_i : \Gamma = 1 : 3 \quad (i = 1, 2, 3, 4)$$

となっていることが図から読み取れる。これからもし Γ が長さを持てば、矛盾が導かれることが分かる。■

[6] フラクタルとは、大ざっぱに言って、分数次元を持つ図形のことである。

第2章　多変数関数

(前章で述べたことの繰り返しになるが) Ω を $\mathbf{R}^n$ の部分集合とするとき、写像 $f\colon\Omega\longrightarrow\mathbf{R}^m$ は、

- $n=1$ のとき、 1 変数 (関数)
- $n>1$ のとき、 多変数 (関数)
- $m=1$ のとき、 実数値 (関数)
- $m>1$ のとき、 ベクトル値 (関数)

と呼ばれる。

どちらかと言うと、1 変数ベクトル値関数 $(m>1)$ よりも、多変数実数値関数 $(n>1)$ の方が難しい。1 変数ベクトル値関数 (曲線) は、ほとんど各成分関数を別々に考えるだけで、大抵の議論がすんでしまうが、多変数関数の場合、例えば $x\to a$ と言っても、色々な近付き方を考えに入れる必要があり、独特の注意が必要になる。

2.1　多変数連続関数

2.1.1　定義と簡単な性質

定義 2.1.1 (多変数関数の極限) Ω を $\mathbf{R}^n$ の部分集合、$f\colon\Omega\to\mathbf{R}^m$, $a\in\overline{\Omega}$, $A\in\mathbf{R}^m$ とするとき、

$$\lim_{x\to a}f(x)=A\overset{\text{def.}}{\Leftrightarrow}\forall\varepsilon>0,\ \exists\delta>0,\ \forall x\in\Omega\cap B(a;\delta)\quad\|f(x)-A\|<\varepsilon.$$

このことを「$x\to a$ のときの $f(x)$ の **極限**(きょくげん) (**limit**) は A である」、「$x\to a$ のとき $f(x)$ は A に**収束**する」などと言い、$f(x)\to A$ $(x\to a)$ とも書く。

注意 2.1.1 (連続の条件を $\Longrightarrow$ を使って表す) 上の条件の書き方は、次のように書いた方が分かりやすいかも知れない。

$$\forall\varepsilon>0,\exists\delta>0,\forall x\in\Omega,\|x-a\|<\delta\Longrightarrow\|f(x)-A\|<\varepsilon.$$

しかし、条件の否定を作るなど、定義の中の書き方が便利なこともある。

$$\lim_{x\to a}f(x)=A\ \text{ではない}\Leftrightarrow\exists\varepsilon>0,\ \forall\delta>0,\ \exists x\in\Omega\cap B(a;\delta)\quad\|f(x)-A\|\geq\varepsilon.$$

注意 2.1.2 特に $x \in \Omega$ としていることを強調したいときは、

$$\lim_{x \in \Omega, x \to a} f(x) = A \text{ あるいは } \lim_{\Omega \ni x \to a} f(x) = A$$

と書くことにする。

命題 2.1.1 (和、差、スカラー倍、積、商の極限) Ω は $\mathbf{R}^n$ の部分集合、$a \in \Omega$ で、 f, $g \colon \Omega \to \mathbf{R}$ は $\lim_{x \to a} f(x) = A$, $\lim_{x \to a} g(x) = B$ を満たし、$\lambda \in \mathbf{R}$ とするならば、次の (1), (2) が成り立つ。

(1) $\lim_{x \to a} (f(x) + g(x)) = A + B$, $\lim_{x \to a} \lambda f(x) = \lambda A$, $\lim_{x \to a} f(x)g(x) = AB$, $\lim_{x \to a} |f(x)| = |A|$.

(2) $B \neq 0$ ならば $\lim_{x \to a} f(x)/g(x) = A/B$.

証明は 1 変数関数のときと同様なので省略する。次の命題も同様である。

命題 2.1.2 (和、差、スカラー倍、内積の極限) Ω は $\mathbf{R}^n$ の部分集合、$a \in \Omega$ で、 f, $g \colon \Omega \to \mathbf{R}^m$ は $\lim_{x \to a} f(x) = A$, $\lim_{x \to a} g(x) = B$ を満たし、$\lambda \in \mathbf{R}$ とするならば、$x \to a$ のとき、$f(x) + g(x) \to A + B$, $\lambda f(x) \to \lambda A$, $(f(x), g(x)) \to (A, B)$, $\|f(x)\| \to \|A\|$.

命題 2.1.3 (合成関数の極限) U は $\mathbf{R}^n$ の部分集合、V は $\mathbf{R}^m$ の部分集合で、$f \colon U \to \mathbf{R}^m$, $f(U) \subset V$, $g \colon V \to \mathbf{R}^\ell$ とする。$\lim_{x \to a} f(x) = b$, $\lim_{y \to b} g(y) = c$ ならば、$\lim_{x \to a} g(f(x)) = c$.

注意 2.1.3 (重要: 多変数関数の極限は「成分変数ごとに極限を取ったもの」ではない) Ω を $\mathbf{R}^2$ から原点 $(0,0)$ を除いたものとし ($\Omega \overset{\text{def.}}{=} \mathbf{R}^2 \setminus \{(0,0)\}$)、$f \colon \Omega \to \mathbf{R}$ を

$$f(x_1, x_2) = \frac{2x_1 x_2}{x_1^2 + x_2^2} \quad ((x_1, x_2) \in \Omega)$$

で定める。$(x_1, x_2) \to (0,0)$ としたときの極限について考えよう。

$$\lim_{x_1 \to 0} f(x_1, 0) = \lim_{x_1 \to 0} \frac{2x_1 \cdot 0}{x_1^2 + 0^2} = \lim_{x_1 \to 0} 0 = 0,$$

$$\lim_{x_2 \to 0} f(0, x_2) = \lim_{x_2 \to 0} \frac{20 \cdot x_2}{0^2 + x_1^2} = \lim_{x_2 \to 0} 0 = 0$$

である。しかし、直線 $L_k \overset{\text{def.}}{=} \{(x_1, x_2) \in \mathbf{R}^2; x_2 = kx_1\}$ を考えると、

$$f(x_1, x_2) = \frac{2k}{1 + k^2} \qquad ((x_1, x_2) \in L_k \setminus \cap \Omega)$$

であるから、L_k に沿って、(x_1, x_2) が $(0,0)$ に近付くとき、

$$\lim_{L_k \ni (x_1, x_2) \to (0,0)} f(x_1, x_2) = \frac{2k}{1 + k^2}.$$

この値は k に依存するので、$\lim_{(x_1, x_2) \to (0,0)} f(x_1, x_2)$ は存在しない。■

多変数関数の連続性を定義しよう。

定義 2.1.2 (多変数連続関数) Ω を $\mathbf{R}^n$ の部分集合、$f\colon \Omega \to \mathbf{R}^m$ とする。

(1) $a \in \Omega$ とするとき、

$$
\begin{aligned}
f \text{ が } a \text{ で}\textbf{連続}\ (\textbf{continuous}) &\overset{\text{def.}}{\Leftrightarrow} \lim_{x \to a} f(x) = f(a) \\
&\Leftrightarrow \forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in B(a;\delta) \cap \Omega \\
&\qquad \|f(x) - f(a)\| < \varepsilon.
\end{aligned}
$$

(2) f が Ω で連続 $\overset{\text{def.}}{\Leftrightarrow}$ $\forall x \in \Omega$ で f が連続。

注意 2.1.4 $n = 1$ の場合は、すでに知られていること (1 変数関数の連続性) と同じである (矛盾しない)。

注意 2.1.5 (ベクトル値関数の連続性は、各成分関数の連続性である) $f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix}$ とするとき、一般に

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

であるから (命題 1.6.1 と同様)、

$$
\begin{aligned}
f \text{ が } a \text{ で連続} &\Leftrightarrow \lim_{x \to a} f(x) = f(a) \\
&\Leftrightarrow \lim_{x \to a} f_i(x) = f_i(a) \quad (1 \le i \le m) \\
&\Leftrightarrow f_i \text{ が } a \text{ で連続 } (1 \le i \le m).
\end{aligned}
$$

注意 2.1.6 (重要: 多変数関数の連続性は「成分変数ごとの連続性」ではない) $f\colon \mathbf{R}^2 \to \mathbf{R}$ を

$$
f(x_1, x_2) = \begin{cases} \dfrac{2x_1 x_2}{x_1^2 + x_2^2} & ((x_1, x_2) \ne (0,0) \text{ のとき}) \\ 0 & ((x_1, x_2) = (0,0) \text{ のとき}) \end{cases}
$$

で定める。既に見たように

$$
\lim_{(x_1, x_2) \to (0,0)} f(x_1, x_2)
$$

は存在しない。ゆえに f は $(0,0)$ で連続ではない。

$$
\lim_{x_1 \to 0} f(x_1, 0) = 0 = f(0,0),
$$

$$
\lim_{x_2 \to 0} f(0, x_2) = 0 = f(0,0)
$$

となるから、「各変数に関して $(0,0)$ で連続」であるが、連続とは限らない関数があるわけである。■

例 2.1.1 (定数関数は連続) $\mathbf{R}^m$ の要素 c を任意に選んで、$f(x)=c$ $(x\in\mathbf{R}^n)$ で $f\colon\mathbf{R}^n\to\mathbf{R}^m$ を定めたとき、f は $\mathbf{R}^n$ で連続である。実際、任意の $a\in\mathbf{R}^n$ に対して明らかに

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

であるから。■

例 2.1.2 (1 次関数は連続) $A\in M(m,n;\mathbf{R})$, $b\in\mathbf{R}^m$ とするとき、

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

で定まる $f\colon\mathbf{R}^n\to\mathbf{R}^m$ を 1 次関数と呼ぶが、これは連続である。実際、a を任意の $\mathbf{R}^n$ の要素とするとき、

$$\|f(x)-f(a)\|=\|(Ax+b)-(Aa+b)\|=\|A(x-a)\|\le\|A\|\|x-a\|$$

であるから、

$$f(x)\to f(a) \quad (x\to a \text{ のとき}).$$

すなわち、f は a で連続である。■

連続関数はたくさんあることを示すために、いくつか命題を用意しよう。

命題 2.1.4 (多変数実数値関数について連続性は遺伝する) Ω は $\mathbf{R}^n$ の部分集合、$a\in\Omega$ で、 f, $g\colon\Omega\to\mathbf{R}$ は a で連続、 $\lambda\in\mathbf{R}$ とするとき、次の (1), (2) が成り立つ。

(1) $f+g$, λf, fg, $|f|$ も a で連続。

(2) $g(a)\ne 0$ ならば f/g も a で連続。

証明 命題 2.1.1 より明らかである。■

命題 2.1.5 (多変数ベクトル値関数について連続性は遺伝する) Ω は $\mathbf{R}^n$ の部分集合、$a\in\Omega$ で、 f, $g\colon\Omega\to\mathbf{R}^m$ は a で連続、$\lambda\in\mathbf{R}$ とするとき、$f+g$, λf, (f,g), $\|f\|$ も a で連続。

証明 命題 2.1.2 より明らかである。■

系 2.1.1 (多項式関数は連続) 多項式関数は連続である。すなわち $P(x_1,x_2,\cdots,x_n)$ を実係数の多項式とするとき、

$$\mathbf{R}^n\ni\begin{pmatrix}x_1\\x_2\\\vdots\\x_n\end{pmatrix}\mapsto P(x_1,x_2,\cdots,x_n)\in\mathbf{R}$$

は連続関数である。

証明　各 i $(1 \le i \le n)$ に対して、関数 $\mathbf{R}^n \ni {}^t(x_1, x_2, \cdots, x_n) \mapsto x_i \in \mathbf{R}$ の連続性が不等式

$$|x_i - y_i| \le \|\vec{x} - \vec{y}\|$$

から導かれる。後は例 2.1.1 と命題 2.1.4 から明らか。 ■

同様にして、有理式 (多項式/多項式 の形をしている式) は分母が 0 にならない点全体の集合で関数を定義し、それは連続であることが分かる。

命題 2.1.6 (連続関数の合成は連続) $U \subset \mathbf{R}^n$, $V \subset \mathbf{R}^m$, $f\colon U \to \mathbf{R}^m$, $g\colon V \to \mathbf{R}^\ell$, $a \in U$, $f(U) \subset V$, f は a で連続、 g は $b \overset{\text{def.}}{=} f(a)$ で連続 $\Longrightarrow$ $g \circ f\colon U \to \mathbf{R}^\ell$ は a で連続。

証明は演習問題にする。

ここまで来れば連続関数の例はいくらでも簡単に作ることができる。

例 2.1.3 (連続関数の具体例) 以下の各関数は $\mathbf{R}^2$ で連続である。

$$f(x, y) = 1 + x + 2y + 3x^2 + 4xy + 5y^2,$$

$$g(x, y) = e^{-(x^2+y^2)},$$

$$h(x, y) = \frac{\sin(xy)}{x^2 + y^2 + 1},$$

$$\varphi(x, y) = \log(1 + x^2 + y^2).$$

分母が 0 になりうる分数関数については、いくつか演習問題を用意してある。

2.1.2　3 つの重要な定理

この小節では連続関数に関する 3 つの非常に重要な定理を紹介する (解析概論 I の範囲でその重要性が明らかにならない定理もあるが)。そのうちの二つがコンパクト性と関係した命題である。

まず準備として、それ自身も重要な次の命題から始めよう。

命題 2.1.7 (連続性の点列による表現) Ω は $\mathbf{R}^N$ の部分集合、$a \in \Omega$ で、 $f\colon \Omega \to \mathbf{R}^m$ とするとき、次の (i), (ii) は互いに同値である。

(i) f は a で連続である。

(ii) a に収束する, Ω 内の任意の点列 $\{x_n\}_{n \in \mathbf{N}}$ に対して、$\displaystyle\lim_{n\to\infty} f(x_n) = f(a)$.

特に、f が連続ならば、$\displaystyle\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n)$ が成り立つ。

証明　(i) $\Longrightarrow$ (ii) は簡単である。(ii) $\Longrightarrow$ (i) を示す。(ii) が成り立つが、(i) は成り立たないと仮定すると、

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in \Omega \bigcap B(a; \delta)\ \|f(x) - f(a)\| \ge \varepsilon.$$

各 $n\in\mathbf{N}$ に対して $\delta=\dfrac{1}{n}$ とおいて、上の事実を適用すると、点列 $\{x_n\}_{n\in\mathbf{N}}\subset\Omega$ で

$$\|x_n-a\|<\frac{1}{n},\quad \|f(x_n)-f(a)\|\ge\varepsilon$$

なるものが取れる。すると、

$$x_n\to a,\quad f(x_n)\not\to f(a)\quad (n\to\infty)$$

となるが、これは矛盾である。ゆえに (ii)$\Longrightarrow$ (i) が成り立つ。■

定理 2.1.1 (有界閉集合上の実数値連続関数は最大値、最小値を持つ) 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。

証明 最大値の存在を証明する (最小値も同様)。 $\sup_{x\in K} f(x)=M$ とおく[1]。sup の定義 (付録 F 「$\mathbf{R}^n$ の有界集合と上限、下限」を参照せよ) から、

$$\exists\{x_n\}_{n\in\mathbf{N}}\colon K \text{ 内の点列 s.t. } \lim_{n\to\infty} f(x_n)=M.$$

$\mathbf{R}^n$ の有界閉集合の点列コンパクト性 (定理 1.5.1) より、

$$\exists\{x_{n_j}\}_{j\in\mathbf{N}}\colon \{x_n\}_{n\in\mathbf{N}} \text{ の部分列}, \exists a\in K \text{ s.t. } \lim_{j\to\infty} x_{n_j}=a.$$

f は連続であると仮定したから、命題 2.1.7 の (i) $\Longrightarrow$ (ii) が適用できて

$$\lim_{j\to\infty} f(x_{n_j})=f(a).$$

一方 $\lim_{j\to\infty} f(x_{n_j})=M$ であるから、極限の一意性より、 $M=f(a)$. ゆえに $M=\max_{x\in K} f(x)$. ■

注意 2.1.7 (最大値を持たない連続関数) K が有界閉集合でない場合には、

- $\sup_{x\in K} f(x)=\infty$ もありうる (当然、最大値は存在しない)。例えば $K=(0,1)$, $f(x)=1/x$ とすると $\sup_{x\in K} f(x)=\infty$.
- $\sup_{x\in K} f(x)<\infty$ であっても、最大値がないことがありうる。例えば $K=(0,1)$, $f(x)=(x-1/2)^2$ とすると $\sup_{x\in K} f(x)=1/4$. しかし $f(a)=1/4$ となる $a\in(0,1)$ はない。

注意 2.1.8 (細かい注意 — 最初読むときは無視) この節にある 3 つの定理のうち、上の定理だけは、 (ベクトル値ではない) 実数値関数に関する定理である。 最大値・最小値は、実数体の順序構造に関係しているので、この定理をベクトル値関数に直接に拡張することは出来ない。(多くの教科書で行なわれているように) 上の定理を、一般に成り立つ重要な定理

「コンパクト集合の連続関数による像はコンパクト集合である」

[1] f が上に有界でない場合は $M=\infty$ とみなす。

の系と見なすのが良いのかも知れない。($\mathbf{R}$ のコンパクト集合 (有界閉集合) は、空でなければ最大値・最小値を持つ、というのは簡単な命題だから。)

定理 2.1.2 (有界閉集合上の多変数連続関数は一様連続である) K を $\mathbf{R}^n$ の有界閉集合、$f\colon K \to \mathbf{R}^m$ は連続とするならば、f は K 上**一様連続** (**uniformly continuous**) である。すなわち

$$\forall \varepsilon > 0,\ \exists \delta > 0,\ \forall x \in K,\ \forall x' \in K \cap B(x;\delta) \quad \|f(x) - f(x')\| < \varepsilon.$$

定理 1.5.1 より、K はコンパクトかつ点列コンパクトである。K のコンパクト性を用いると、明快な証明が出来るが、ここでは点列コンパクト性を用いた証明を紹介しよう[2]。

証明　背理法で証明する。f が一様連続でないと仮定すると、

$$\exists \varepsilon > 0,\ \forall \delta > 0,\ \exists x \in K,\ \exists x' \in K \cap B(x;\delta) \quad \|f(x) - f(x')\| \ge \varepsilon.$$

そこで δ として $1/n$ ($n \in \mathbf{N}$) を取ることによって、次の条件を満たす二つの点列 $\{x_n\}_{n\in\mathbf{N}}$, $\{y_n\}_{n\in\mathbf{N}}$ が存在することが分かる。

$$x_n \in K, \quad y_n \in K, \quad \|x_n - y_n\| < \frac{1}{n}, \quad \|f(x_n) - f(y_n)\| \ge \varepsilon.$$

K の点列コンパクト性から、適当な部分列 $\{x_{n_k}\}_{k\in\mathbf{N}}$ を取ると、

$$\exists a \in K \quad \text{s.t.} \quad x_{n_k} \to a \quad (k \to \infty).$$

ところで $\|x_n - y_n\| \le 1/n$ であるから、

$$\|y_{n_k} - a\| \le \|y_{n_k} - x_{n_k}\| + \|x_{n_k} - a\| \le \frac{1}{n_k} + \|x_{n_k} - a\| \to 0 \quad (k \to \infty).$$

すなわち $y_{n_k} \to a$ ($k \to \infty$). f の連続性より $k \to \infty$ のとき $f(x_{n_k}) \to f(a)$, $f(y_{n_k}) \to f(a)$. ゆえに

$$\|f(x_{n_k}) - f(y_{n_k})\| \ge \varepsilon$$

で $k \to \infty$ とすると、

$$0 = \|f(a) - f(a)\| \ge \varepsilon > 0.$$

これは矛盾である。ゆえに f は K 上で一様連続でなければならない。 ■

注意 2.1.9 (一様連続でない連続関数) K が有界閉集合でない場合は、連続であって、一様連続でない関数が存在し得る。$K = (0,1)$, $f(x) = 1/x$ とするとき、 $\forall \delta > 0$ に対して、

$$\delta' \overset{\text{def.}}{=} \min\{\delta, 1/2\}, \quad x \overset{\text{def.}}{=} \frac{\delta'}{2}, \quad x' \overset{\text{def.}}{=} \frac{\delta'}{4}$$

とおくと、

$$|x - x'| = \frac{\delta'}{4} < \delta, \quad |f(x) - f(x')| = \left|\frac{2}{\delta'} - \frac{4}{\delta'}\right| = \frac{2}{\delta'} \ge 1.$$

要するに

[2] このノートでは、Heine-Borel の定理の (i) $\Longrightarrow$ (iii) の部分の証明を省略しているので、それを必要としない証明を与えることにした。

$$\exists\varepsilon > 0,\, \forall\delta > 0,\, \exists x \in K,\, \exists x' \in K \cap B(x;\delta) \quad \|f(x) - f(x')\| \geq \varepsilon$$

が $\varepsilon = 1$ として示されたので、一様連続でないことが分かった。■

定義 2.1.3 (連結集合) $\mathbf{R}^n$ の部分集合 Ω が**連結**(れんけつ) **(connected)** であるとは、Ω 内の任意の 2 点が、Ω 内の曲線で結ばれることである:

$$\forall x \in \Omega,\, \forall y \in \Omega,\, \exists\varphi\colon [\alpha,\beta] \to \Omega \text{ 連続曲線 s.t. } \varphi(\alpha) = x \text{ かつ } \varphi(\beta) = y.$$

注意 2.1.10 (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 (あまり直感的でない) 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、 $\mathbf{R}^n$ の部分集合においては、連結＝弧状連結なので、ここでは簡単で直感的な定義法を採用した。

$\mathbf{R}$ の部分集合 A について、A が連結であるための必要十分条件は、A が区間であることである。これは、次の定理 (知っているはずだから証明略) から分かる。

補題 2.1.1 (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、

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

系 2.1.2 (連続関数による区間の像は区間) I を $\mathbf{R}$ の区間、 $f\colon I \to \mathbf{R}$ を連続関数とすると、その像 $f(I) \overset{\text{def.}}{=} \{f(x); x \in I\}$ は $\mathbf{R}$ の区間である。

この結果は「連続関数による連結集合の像は連結である」という形で一般化される。その証明は簡単であるが (理解を深めたければ試してみよ)、ここでは次の形の命題にまとめておこう。

定理 2.1.3 (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。

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

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 $c \overset{\text{def.}}{=} \varphi(\gamma)$ とおけば、 $f(c) = k$ となる。■

2.2 偏微分

2.2.1 偏微分の定義

まず f が多変数関数の場合、差分商

$$\frac{f(x+h)-f(x)}{h}$$

が意味を持たないことに注意する (ベクトルで割ることは出来ないから)。

定義 2.2.1 (偏微分可能、偏微分係数) Ω を $\mathbf{R}^n$ の開集合、$a = {}^t(a_1, \cdots, a_n) \in \Omega$, $f \colon \Omega \to \mathbf{R}^m$ とするとき、f が a で変数 x_i について**偏微分可能**(へんびぶんかのう) (**partially differentiable**) とは、極限

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

が存在することであると定義する。ここで、 e_i は、第 i 成分のみ 1 で、他の成分は 0 であるような $\mathbf{R}^n$ のベクトルである。つまり、

$$\lim_{h\to 0}\frac{f(a+he_i)-f(a)}{h} = \lim_{h\to 0}\frac{f(a_1,\cdots,a_{i-1},a_i+h,a_{i+1},\cdots,a_n)-f(a_1,\cdots,a_{i-1},a_i,a_{i+1},\cdots,a_n)}{h}$$

である。この値のことを、f の a における、変数 x_i についての **偏微分係数** (**partial differential coefficient**) と呼び、

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

などの記号で表す。∂ は「パーシャル・ディー (partial “D”)」、「ラウンド・ディー (round “D”)」、あるいは単に「ディー」と読む。

(偏微分を表す記号として $\partial_{x_i} f$ というのもあり、便利なこともあるが、この講義では使わない。)

例 2.2.1 2 変数関数 $f \colon (x,y) \mapsto f(x,y)$ の場合は、

$$\begin{aligned}
\frac{\partial f}{\partial x}(a,b) &= f_x(a,b) = \lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}, \\
\frac{\partial f}{\partial y}(a,b) &= f_y(a,b) = \lim_{h\to 0}\frac{f(a,b+h)-f(a,b)}{h}.
\end{aligned}$$

定義 2.2.2 (偏導関数、高階の偏導関数、C^k-級) Ω を $\mathbf{R}^n$ の開集合、$f\colon\Omega \to \mathbf{R}^m$, $a = {}^t(a_1, \cdots, a_n) \in \Omega$ とする。

(1) f が Ω で、変数 x_i について**偏微分可能** (**partially differentiable**) $\overset{\text{def.}}{\Leftrightarrow}$ $\forall x \in \Omega$ に対して、f は x で、変数 x_i について偏微分可能。
このとき、写像

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

を、f の変数 x_i に関する**偏導関数** (**partial derivative**, **partial derived function**) と呼び、

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

などの記号で表す。
f の変数 x_i に関する偏導関数 $\dfrac{\partial f}{\partial x_i}$ $(1 \le i \le n)$ を f の 1 **階偏導関数** (**first partial derivative**, **first partial derived function**) と総称する。

(2) f が Ω で変数 x_i について偏微分可能で、その偏導関数 $\dfrac{\partial f}{\partial x_i}$ が Ω で、変数 x_j について偏微分可能であるとき、

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)$$

を

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

などの記号で表し、f の 2 **階偏導関数** (**second partial derivative**, **second partial derived function**) と呼ぶ。

(3) (2) と同様にして、$\forall k \in \mathbf{N}$ に対して、k **階偏導関数**(k**-th partial derivative**, k**-th partial derived function**) が定義できる。

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_2} \partial x_{i_1}} \overset{\text{def.}}{=} \frac{\partial}{\partial x_{i_k}}\left(\cdots\left(\frac{\partial}{\partial x_{i_2}}\left(\frac{\partial}{\partial x_{i_1}}f\right)\right)\cdots\right).$$

(4) $m \in \mathbf{N}$ とするとき、f が Ω で C^m-級 (m **回連続的微分可能**) とは、f が Ω において、m 階以下のすべての偏導関数を持ち、それらすべてと、f 自身が Ω 上で連続であることと定義する。1 回連続的微分可能であることを単に連続的微分可能であるという。

(5) 関数 $f\colon\Omega \to \mathbf{R}$ が連続であることを、f は C^0-級である、と言うことがある。f 自身のことを f の 0 階 (偏) 導関数と言うことがある。

(6) f が C^∞-級 (**無限回微分可能**) であるとは、$\forall k \in \mathbf{N}$ に対して、f が C^k-級であることと定義する。

注意 2.2.1 (記号について) (1) 同じ変数に関する偏微分については、巾の記号を用いる。例えば $i=j$ のとき、$\dfrac{\partial^2 f}{\partial x_j \partial x_i}=\dfrac{\partial^2 f}{\partial x_i \partial x_i}$ のことを $\dfrac{\partial^2 f}{\partial x_i^2}$ と書く。

(2) 人によっては、$\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ のことを $f_{x_j x_i}$ と書く (上の定義とは反対)。

2.2.2 偏微分の順序交換

定理 2.2.1 (偏微分の順序交換) Ω を $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$ を C^2-級の関数とするとき

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

証明 $i=j$ のときは証明する必要がない。$i\neq j$ とする。$x_i=x$, $x_j=y$ とおき、f を 2 変数関数

$$(x,y)\longmapsto f(x,y)$$

として証明すればよい (他の変数は一切変化させないわけだから)。$(a,b)\in\Omega$ において考える。

$$\Delta(h,k)\overset{\text{def.}}{=}f(a+h,b+k)-f(a+h,b)-f(a,b+k)+f(a,b)$$

とおく ($|h|$, $|k|$ は十分小さいとする)。

1. 主張 $\displaystyle\lim_{h,k\to 0}\frac{\Delta(h,k)}{hk}=\frac{\partial^2 f}{\partial y\partial x}(a,b)$.
   実際、 $\phi(x)\overset{\text{def.}}{=}f(x,b+k)-f(x,b)$ とおくと、$\Delta(h,k)=\phi(a+h)-\phi(a)$ であるが、平均値の定理から $\exists\theta_1\in(0,1)$ s.t.
   $$\Delta(h,k)=\phi_x(a+\theta_1 h)h=\left\{\frac{\partial f}{\partial x}(a+\theta_1 h,b+k)-\frac{\partial f}{\partial x}(a+\theta_1 h,b)\right\}h.$$
   もう一回、今度は y について平均値の定理を使って、$\exists\theta_2\in(0,1)$ s.t.
   $$\Delta(h,k)=\frac{\partial^2 f}{\partial y\partial x}(a+\theta_1 h,b+\theta_2 k)hk.$$
   仮定より、$\dfrac{\partial^2 f}{\partial y\partial x}$ は連続ゆえ、
   $$\lim_{h,k\to 0}\frac{\Delta(h,k)}{hk}=\frac{\partial^2 f}{\partial y\partial x}(a,b).$$
2. 上と全く同様に $\displaystyle\lim_{h,k\to 0}\frac{\Delta(h,k)}{hk}=\frac{\partial^2 f}{\partial x\partial y}(a,b)$ が示せる。
3. ゆえに $\dfrac{\partial^2 f}{\partial y\partial x}(a,b)=\dfrac{\partial^2 f}{\partial x\partial y}(a,b)$. ■

寄り道: 上の証明は一体？　上の証明は、正しいことは納得できるが、ピンと来ないという感想が多かったので、少し背景説明を補足する。高等学校で微分係数の定義

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

を学んだ際に、

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

や

$$f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a-h)}{2h}$$

などの式を証明せよ、という練習問題があったであろう。右辺に現れる $(f(a+h)-f(a))/h$ のような式を「差分商」と呼ぶのであるが (正確な定義は省略する)、

微分係数は差分商の極限である

と標語的に説明できる。実は 1 階の微分係数だけでなく、例えば

$$f''(a) = \lim_{h\to 0} \frac{f(a+h)-2f(a)+f(a-h)}{h^2}$$

のように、高階の微分係数も差分商の極限として表すことができる。以上は、1 変数関数についての話であったが、同じことが多変数関数の微分係数についても成り立つ。例えば C^2-級の 2 変数関数 $f\colon (x,y)\mapsto f(x,y)$ の場合

$$\begin{aligned}
f_x(a,b) &= \lim_{h\to 0} \frac{f(a+h,b)-f(a,b)}{h},\\
f_y(a,b) &= \lim_{k\to 0} \frac{f(a,b+k)-f(a,b)}{k},\\
f_{xy}(a,b) &= \lim_{h,k\to 0} \frac{f(a+h,b+k)-f(a+h,b)-f(a,b+k)+f(a,b)}{hk},\\
\frac{\partial^2 f}{\partial x^2}(a,b) &= \lim_{h\to 0} \frac{f(a+h,b)-2f(a,b)+f(a-h,b)}{h^2},
\end{aligned}$$

等々が成り立つ。このような事実を知っていると、上の定理の証明は極めて自然に見えてくる。つまり差分商を作る操作の順序の可換性が偏微分の順序の可換性に遺伝する、ということである。■

系 2.2.1 (C^k-級なら、偏微分の順序交換が出来る) f が $\mathbf{R}^n$ の開集合の上で定義された C^k-級の関数とするならば、f の k 階までの偏微分の順序は、交換可能である。

証明　明らかか。■

以下は、高木 [7] による。

偏微分の順序交換

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

が成立しない「病的な[3]」例もある。

[3]何をもって病的と言うかは主観の問題であるが、超関数においては常に偏微分の順序交換が出来ることを考えると、順序交換が出来ない関数は超関数ともみなせない関数であるということで、かなりタチが悪い、と筆者 (桂田) は思う。

例 2.2.2 (偏微分の順序交換ができない関数)

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

について

$$f_{xy}(0,0)=-1,\quad f_{yx}(0,0)=1.$$

また、偏微分の順序交換が成立するための条件も、上の定理以外に色々知られている。しかし、そういうものに深入りすることはやめておくことにする。証明抜きで二つほどあげておくにとどめる。

命題 2.2.1 (Schwarz) ある領域で f_x, f_y, f_{xy} が存在して、領域内の一点において 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 2.2.2 (Young) ある領域で f_x, f_y が存在して、それらが領域内の一点において全微分可能ならば、その点において $f_{xy}=f_{yx}$.

(全微分可能性の定義は、次の節に初めて出て来るので、順番が前後しているが、大目に見て下さい。)

これらの定理の証明は、いずれも上記の本にある。

2.3 (全) 微分

この節では、もう一つの微分の概念である**全微分**(ぜんびぶん)を説明する。1 変数の場合の微分によく対応するのは (偏微分よりはむしろ) 全微分の方で、こちらがより自然な概念と考えられる。それで最近では「全」を省略して、単に「微分」と呼ぶことも多い。

高等学校の微分法では、接線が大活躍したが、

関数のグラフの接線は、関数を 1 次関数で近似したもの (のグラフ) である。

そして実は

関数を微分するとは、関数を 1 次関数で近似することである。

このことをしみじみ納得することが、現代の微分法を理解する際の急所である[4]。

2.3.1 微分の定義

1 変数関数 $f\colon \mathbf{R}\supset\Omega\to\mathbf{R}^m$ の場合には

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

こう書くと多変数への拡張が見えて来る。

[4]このように 1 階微分係数を 1 次関数による近似ととらえるのは、 Fréchet (1878–1973) に始まるそうで、特に無限次元空間における微分法では、今でも Fréchet 微分という語はよく使われている。

- A を $m \times n$ 行列 (i.e. $A \in M(m,n;\mathbf{R})$)
- h を $\mathbf{R}^n$ のベクトル
- $|h|$ を $\|h\|$

に置き換えればよい。これがふさわしいことは段々に納得できるであろう。

定義 2.3.1 (全微分の定義) Ω を $\mathbf{R}^n$ の開集合, $a \in \Omega$, $f\colon \Omega \to \mathbf{R}^m$ とする。

(1) f が a で (**全**) **微分可能 (totally differentiable)** $\overset{\text{def.}}{\Leftrightarrow}$ $\exists A \in M(m,n;\mathbf{R})$ s.t.

$$\lim_{h\to 0} \frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} = 0 \tag{2.1}$$

この A のことを f の a における**微分係数**といい、 $f'(a)$ あるいは $Df(a)$ などの記号で表す。

(2) f が Ω で (全) 微分可能 $\overset{\text{def.}}{\Leftrightarrow}$ $\forall x \in \Omega$ に対して f が x で微分可能。
このとき

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

を f の**導関数**と呼び f' あるいは Df と記す。

注意 2.3.1 (1) f の a における微分係数は (もし存在するならば) 一意的に定まる。

(2) これは 1 変数の場合の拡張になっている。

(3) (Landau の記号による表現) 上の極限についての式 (2.1) は、$f(a+h) - f(a) - Ah$ が h よりも速く 0 に収束するということで、Landau の記号 (付録 G 参照) を用いると

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

と書ける。要するに $o(\|h\|)$ $(h \to 0)$ とは、$h \to 0$ のとき h よりも速く 0 に収束する量を表す。 Landau の記号に慣れると、極限に関する計算の多くを、あたかも代数計算であるかのように遂行できて便利であるが、慣れるのにある程度の時間がかかりそうなので、この講義では、なるべく lim で表現して済ませるようにした。

(4) 古い本では、この微分係数に **Stolz の微分係数**という名をつけている。

(5) $m = 1$ の場合、すなわち

$$f\colon \Omega \longrightarrow \mathbf{R}$$

の場合は $A \in M(1,n;\mathbf{R})$, つまり A は n 次元横ベクトルである。そこで $A = (a_1 a_2 \cdots a_n)$,

$$h = \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}$$ とおくと

$$Ah = (a_1 a_2 \cdots a_n) \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix} = \sum_{j=1}^{n} a_j h_j.$$

そこで A を転置した縦ベクトルを u とすると、すなわち

$$u = {}^t A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

とおくと

$$Ah = (u, h).$$

この u のことを $\operatorname{grad} f(a)$ や $\nabla f(a)$ などの記号で表し、f の**勾配**(こうばい) (**gradient**) と呼ぶ。この記号を用いると

$$f(a+h) - f(a) - (\operatorname{grad} f(a), h) = o(\|h\|) \quad (h \to 0).$$

定理 2.3.1 (微分可能 $\Longrightarrow$ 連続、各変数に関して偏微分可能で微分係数はヤコビ行列) Ω を $\mathbf{R}^n$ の開集合、 $a \in \Omega$, $f\colon \Omega \to \mathbf{R}^m$ とするとき以下の (1), (2) が成り立つ。

(1) f が a で微分可能ならば、f は a で連続。

(2) f が a で微分可能ならば、f はすべての変数 x_i $(i = 1, 2, \ldots, n)$ について偏微分可能で

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

このように f の 1 階偏微分係数を並べた行列を f の**Jacobi**(ヤコビ) **行列** (**Jacobian matrix**) と呼ぶ。

注意 2.3.2 (ヤコビ行列式) (解析概論 I の範囲では使わないのだが)

$$\frac{\partial(f_1, f_2, \cdots, f_n)}{\partial(x_1, x_2, \cdots, x_n)}(a) \overset{\text{def.}}{=} \det f'(a)$$

とおき[5]、f の a における**ヤコビ行列式** (**ヤコビアン**, **Jacobian**) あるいは**関数行列式** (**functional determinant**) と呼ぶ。すなわち、f のヤコビ行列式とは、f のヤコビ行列の行列式のことである。

系 2.3.1 (特に $m = 1$ の場合) $f\colon \Omega \to \mathbf{R}$ が $a \in \Omega$ で微分可能ならば

$$\operatorname{grad} f(a) = \nabla f(a) = {}^t\left(f'(a)\right) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1}(a) \\ \dfrac{\partial f}{\partial x_2}(a) \\ \vdots \\ \dfrac{\partial f}{\partial x_n}(a) \end{pmatrix}.$$

覚えて欲しいことを一行で書けば

多変数ベクトル値関数の微分係数は、ヤコビ行列である: $f'(a) = \left(\dfrac{\partial f_i}{\partial x_j}(a)\right)$.

定理 2.3.1 の証明

(i) f が微分可能であることから $\exists A = (a_{ij}) \in M(m,n;\mathbf{R})$ s.t.

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

ゆえに

$$\begin{aligned} \|f(a+h) - f(a)\| &= \|f(a+h) - f(a) - Ah + Ah\| \\ &\le \|f(a+h) - f(a) - Ah\| + \|A\|\,\|h\| \\ &= \|h\| \frac{\|f(a+h) - f(a) - Ah\|}{\|h\|} + \|A\|\,\|h\| \to 0 \quad (h \to 0). \end{aligned}$$

すなわち f は a で連続である。

(ii) f が a で微分可能であることから $\exists A = (a_{ij}) \in M(m,n;\mathbf{R})$ s.t.

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

この式の右辺の分子の $\|\cdot\|$ 内のベクトルの第 i 成分を取り出すと

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

[5]実は $\dfrac{\partial(f_1, f_2, \cdots, f_n)}{\partial(x_1, x_2, \cdots, x_n)}(a)$ で (行列式ではなく) ヤコビ行列 $f'(a)$ そのものを表すという流儀もある (微分法の記号の統一性の無さはかなり困ったものだ)。

ここで $h_k = \delta_{jk}$ (Kronecker のデルタ)、すなわち

$$h = \varepsilon e_j = \varepsilon \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow j \text{ 番目} \qquad (\varepsilon \text{ は } |\varepsilon| \text{ が十分小さな実数})$$

とすると

$$\lim_{\varepsilon \to 0} \frac{|f_i(a + \varepsilon e_j) - f_i(a) - a_{ij}\varepsilon|}{|\varepsilon|} = 0.$$

ゆえに

$$\lim_{\varepsilon \to 0} \frac{f_i(a + \varepsilon e_j) - f_i(a)}{\varepsilon} = a_{ij}.$$

これは $\dfrac{\partial f_i}{\partial x_j}(a)$ が存在して a_{ij} に等しいことを示している。■

例題 2.3.1 (1 次関数の微分係数) $A \in M(m, n; \mathbf{R})$, $b \in \mathbf{R}^m$ とするとき、

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

で定義される $f\colon \mathbf{R}^n \to \mathbf{R}^m$ について、$f'(x)$ は何か？

答 証明は後回しにして、結論を先に述べると、任意の $x \in \mathbf{R}^n$ に対して

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

つまり f の x における微分係数 (ヤコビ行列) は A である (これは 1 変数実数値関数の世界で良く知られている「$f(x) = ax + b$ ならば $f'(x) = a$」という事実の一般化である)。このノートでは二つの証明を与えるが、まずは微分係数の定義に基づく証明を示しておく。

$$f(x+h) - f(x) - Ah = A(x+h) + b - (Ax + b) - Ah = Ax + Ah + b - Ax - b - Ah = 0.$$

よって

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

となり、定義によって $f'(x) = A$. ■

上の定理で見たように

微分可能 $\Longrightarrow$ 各変数に関して偏微分可能

しかし逆は成立しない。

例 2.3.1 (各変数につき偏微分可能だが、微分可能でない関数)

$$f(x,y) \overset{\text{def.}}{=} \begin{cases} \dfrac{2xy}{x^2+y^2} & ((x,y)\neq(0,0) \text{ のとき}) \\ 0 & ((x,y)=(0,0) \text{ のとき}) \end{cases}$$

で定義される $f\colon \mathbf{R}^2 \to \mathbf{R}$ は、0 で変数 x, y の双方に関して偏微分可能である。実際

$$\begin{aligned} f_x(0,0) &= \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0, \\ f_y(0,0) &= \lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0}\frac{0-0}{h} = 0. \end{aligned}$$

しかし f は 0 で微分可能ではない (f は 0 で連続でないことは既に注意 2.1.6 の中で示してあるから)。 ■

すべての偏導関数の連続性を仮定すると、微分可能性が出て来る。

定理 2.3.2 (C^1-級ならば全微分可能) Ω を $\mathbf{R}^n$ の開集合、 $f\colon \Omega \to \mathbf{R}^m$ を C^1-級の写像とするならば f は Ω 上微分可能。

証明

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \text{ が微分可能} \iff \text{各 } f_i\ (i=1,\dots,m) \text{ が微分可能.}$$

$$f = \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} \text{ が } C^1\text{-級} \iff \text{各 } f_i\ (i=1,\dots,m) \text{ が } C^1\text{-級.}$$

であるから (どうしてか各自考えよ)、$m=1$ として証明すれば十分である。

$$a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \Omega, \quad h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}, \quad a+h\in\Omega$$

とすると、平均値の定理から $\exists \theta_j \in (0,1)$ $(j=1,2,\cdots,n)$ s.t.

$$\begin{aligned} f(a+h)-f(a) &= f(a_1+h_1,a_2+h_2,\cdots,a_n+h_n) - f(a_1,a_2,\cdots,a_n) \\ &= f(a_1+h_1,a_2+h_2,\cdots,a_n+h_n) - f(a_1,a_2+h_2,\cdots,a_n+h_n) \\ &\quad +f(a_1,a_2+h_2,\cdots,a_n+h_n) - f(a_1,a_2,a_3+h_3,\cdots,a_n+h_n) \\ &\quad +\cdots \\ &\quad +f(a_1,\cdots,a_{j-1},a_j+h_j,a_{j+1}+h_{j+1},\cdots,a_n+h_n) - f(a_1,\cdots,a_{j-1},a_j,a_{j+1}+h_{j+1},\cdot \\ &\quad +\cdots \end{aligned}$$

$$
\begin{aligned}
&\quad +f(a_1,a_2,\cdots,a_{n-1},a_n+h_n)-f(a_1,a_2,a_3,\cdots,a_{n-1},a_n) \\
&= \sum_{j=1}^{n} f_{x_j}(a_1,a_2,\cdots,a_{j-1},a_j+\theta_j h_j,a_{j+1}+h_{j+1},\cdots,a_n+h_n)h_j \\
&= \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)h_j+\sum_{j=1}^{n}\varepsilon_j(h)h_j.
\end{aligned}
$$

ただし

$$
\varepsilon_j(h) \overset{\text{def.}}{=} \frac{\partial f}{\partial x_j}(a_1,\cdots,a_{j-1},a_j+\theta_j h_j,a_{j+1}+h_{j+1},\cdots,a_n+h_n)-\frac{\partial f}{\partial x_j}(a).
$$

仮定より $\dfrac{\partial f}{\partial x_j}$ は連続ゆえ $\lim_{h\to 0}\varepsilon_j(h)=0$. よって $h\to 0$ のとき、

$$
\frac{\left\|\sum_{j=1}^{n}\varepsilon_j(h)h_j\right\|}{\|h\|} \le \sum_{j=1}^{n}|\varepsilon_j(h)|\frac{|h_j|}{\|h\|} \le \sum_{j=1}^{n}|\varepsilon_j(h)|\cdot 1=\sum_{j=1}^{n}|\varepsilon_j(h)|\to 0.
$$

ゆえに

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

これは f が a で全微分可能であることを示している。 ■

全微分の定義をやや天下りに感じた人も多いと思うが、上の定理は、その定義の正当性の一つの裏付けになると言えよう。また、これから多くの関数の全微分可能性が簡単に証明できる(これで一安心)。

例題 2.3.2 次の 4 つの条件の間の関係について、自分なりにまとめよ。
(i) C^1-級 (連続的微分可能)、(ii) 全微分可能、 (iii) 各変数について偏微分可能、(iv) 連続.

解答　定理 2.3.2 によって「(i) $\Longrightarrow$ (ii)」, また定理 2.3.1 によって「(ii) $\Longrightarrow$ (iii)」 と「(ii) $\Longrightarrow$ (iv)」が示されている。これから明らかに「(i)$\Longrightarrow$ (iii)」, 「(i)$\Longrightarrow$ (iv)」が成り立つ。これ以外には、一般に成り立つことはない。

- 「(ii) $\Longrightarrow$ (i)」は一般には成り立たない。すなわち微分可能であるが、連続的微分可能でないような関数がある。例えば 1 変数で次のような例がある。

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

- 「(iii) $\Longrightarrow$ (ii)」は一般には成り立たない。すなわち連続であるが、微分可能でないような関数がある。例えば 1 変数で次のような例がある。$f(x)=|x|$ は $x=0$ で連続であるが、微分可能ではない。

- 「(iv) $\Longrightarrow$ (ii)」は一般には成り立たない。すなわち偏微分可能であるが、微分可能でないような関数がある。これは例 2.3.1 で見た。
- 「(iii) $\Longrightarrow$ (iv)」は一般には成り立たない。すなわち連続であるが、偏微分可能でないような関数がある。例えば $f(x,y)=|x|$ は連続だが、原点 $(0,0)$ で変数 x について偏微分可能ではない。
- 「(iv) $\Longrightarrow$ (iii)」は一般には成り立たない。すなわち偏微分可能であるが、連続でないような関数がある。これは例 2.3.1 で見た。■

2.3.2 いくつかの例

定理 2.3.2 によって、多くの場合に与えられた関数が微分可能であることが簡単に調べられる (とにかく偏微分してみて、それが連続であるかどうか調べればよい)。

手始めに 1 次関数の微分係数の問題を再度取り上げてみよう。

例題 2.3.3 (1 次関数の微分係数 (計算に基づく方法))

$$f(x)=Ax+b=\left(\sum_{k=1}^{n}a_{ik}x_k+b_i\right)$$

より

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

この結果が連続であることは明らかであるから (「定数関数は連続」)、f は全微分可能であることが分かり、

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

例 2.3.2 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

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

ちなみに

$$\nabla f(x,y)=\begin{pmatrix}2ax+2by+p\\2cy+2bx+q\end{pmatrix}. \quad ■$$

例 2.3.3

$$f(x,y)=\begin{pmatrix}x^2-y^2\\2xy\end{pmatrix}$$

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

解

$$f'(x,y)=\begin{pmatrix}\dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y}\\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y}\end{pmatrix}=\begin{pmatrix}2x & -2y\\ 2y & 2x\end{pmatrix}. \quad \blacksquare$$

例 2.3.4 (2 次関数)

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

で定義される n 変数の実数値 2 次関数 $f\colon \mathbf{R}^n\to\mathbf{R}$ について考えよう。表現の一意性のため

$$a_{ij}=a_{ji}\quad (1\le i,j\le n)$$

を仮定する。

$$\sum_{i,j=1}^{n}a_{ij}x_ix_j=\sum_{i=1}^{n}\left(\sum_{j=1}^{n}a_{ij}x_j\right)x_i=\sum_{i=1}^{n}(Ax\text{ の第 }i\text{ 成分})\,x_i=(Ax,x)$$

と書ける (最後の括弧は内積を表す)。ただし $A=(a_{ij})$ とおいた。さらに $b=(b_i)$ とおくと、

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

と表せる。このとき

$$\nabla f(x)=Ax+b$$

が成り立つことが分かる (確かめるのは良い演習問題である)。これは 1 変数実数値関数における公式

$$\left(\frac{1}{2}ax^2+bx+c\right)'=ax+b$$

の一般化である。■

空間極座標　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 P の

- 原点からの距離を r
- 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、$\overrightarrow{OP'}$ が x 軸の正方向となす角を反時計回りに計った角度を φ $(0\le\varphi<2\pi)$

とすると

$$\begin{cases} x = r\sin\theta\cos\varphi\\ y = r\sin\theta\sin\varphi\\ z = r\cos\theta\end{cases}$$

が成り立つ。このとき (r,θ,φ) を点 P の**極座標**(きょくざひょう) (**polar coordinates**) と呼ぶ。

計算練習

$$I \overset{\text{def.}}{=} [0,\infty) \times [0,\pi] \times [0,2\pi) \subset \mathbf{R}^3$$

とおき、

$$f\colon I \ni \begin{pmatrix} r \\ \theta \\ \varphi \end{pmatrix} \longmapsto \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbf{R}^3$$

を

$$\begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \qquad (0 \le r < \infty,\ 0 \le \theta \le \pi,\ 0 \le \varphi < 2\pi) \\ z &= r\cos\theta \end{cases}$$

で定義したとき、微分係数 f' とヤコビアン $\det f'$ を求めよ。

2.3.3 $\operatorname{grad} f$ の幾何学的意味

この小節の要点は一行で書ける:

> $\operatorname{grad} f = \nabla f$ **は、レベル・セットの法線ベクトルである。**

$\mathbf{R}^n$ の開集合 Ω 上で定義された C^1-級の関数 $f\colon \Omega \to \mathbf{R}$ があるとする。$a \in \Omega$, $f(a) = c$ として、f の高さ c の**等高面** (レベル・セット, **contour**, **level set**)

$$L_c \equiv \{x \in \Omega; f(x) = c\}$$

を考える。素朴に考えると L_c は $\mathbf{R}^n$ の中で余次元 1 の曲面[6] (超曲面) を定めるが、厳密には以下の仮定をおく必要がある:

$$\forall x \in L_c \quad \nabla f(x) \neq 0.$$

(曲面のきちんとした話は、もう少し準備が整った段階でないとできないので、今はフィーリングで読んで欲しい。以下の記述は、うるさいことを言い出すと問題だらけで…)

f は a で微分可能であることから、

$$f(x) - f(a) - (\nabla f(a), x - a) = o(\|x - a\|) \quad (x \to a)$$

が成り立つが、ここで $x \in L_c$ とすると $f(x) = f(a) = c$ ゆえ

$$(\nabla f(a), x - a) = o(\|x - a\|) \quad (x \to a).$$

すなわち

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

[6] $n = 2$ の場合は、曲線 (1 次元的存在), $n = 3$ の場合は普通の曲面 (2 次元的存在) というように、属している空間の次元よりも 1 だけ小さい次元を持ったものになっている。このことを余次元は 1 である、という。余次元 1 の図形を超曲面 (hypersurface) と呼ぶ。ちなみに 1 次式 $= 0$ で定義される図形を超平面 (hyperplane) と呼ぶ (つまり hyperplane は特別な hypersurface である)。$\mathbf{R}^2$ の超平面とは普通の直線のこと、$\mathbf{R}^3$ の超平面とは普通の平面のことである。$\mathbf{R}^2$ の超曲面とは普通の曲線のこと、$\mathbf{R}^3$ の超曲面とは普通の曲面のことである。

ここに現われるベクトル $\dfrac{x-a}{\|x-a\|}$ は、a から x に向かう方向の単位ベクトルであるから、この式は、a から x に向かう方向が、$x \to a$ での極限では、$\nabla f(a)$ と直交することを意味している[7]。これから

$\nabla f(a)$ は L_c の「**接超平面**の **法線ベクトル (normal vector)**」というべき性質を持つ

ことがわかる。実際、L_c の a における接超平面は普通

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

で定義される。

さて、c を変化させて、それぞれの値に対して L_c を描いてみよう。こうして出来た図を地図とみなすと、

$\nabla f(a)$ は点 a において傾斜が最も急な方向を表す。

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。Schwarz の不等式から

$$|(\nabla f(a), h)| \le \|\nabla f(a)\| \cdot \|h\|, \quad \text{等号} \Leftrightarrow \exists \lambda \in \mathbf{R} \text{ s.t. } h = \lambda \nabla f(a)$$

より詳しくは

$$-\|\nabla f(a)\| \cdot \|h\| \le (\nabla f(a), h) \le \|\nabla f(a)\| \cdot \|h\|.$$

左の不等号の等号は $h = \lambda \nabla f(a)$ $(\lambda \le 0)$ のとき、右の不等号の等号は $h = \lambda \nabla f(a)$ $(\lambda \ge 0)$ のとき、成立する。つまり、$\nabla f(a)$ の方向に移動すると、高さの変化が最も大きい。 ■

例題 2.3.4 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、

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

これが 0 にならないから ($\because (x_0, y_0) \ne (0,0)$)、$(x_0, y_0)$ 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

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

[7] $\dfrac{x-a}{\|x-a\|}$ が $x \to a$ のとき収束するわけではないことに注意して読もう。

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、

$$\frac{x_0x}{a^2}+\frac{y_0y}{b^2}=1. \quad \blacksquare$$

2.3.4 線形化写像とグラフの接超平面

f が a で微分可能であるとき

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

であるから、 $\|x-a\|$ が十分小さいとき

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

右辺は 1 次関数である。つまり、これは f を a の近くで、1 次関数で近似していることになる。この右辺の関数に名前をつけておこう。

定義 2.3.2 (線形化写像) Ω を $\mathbf{R}^n$ の開集合, $f\colon\Omega\to\mathbf{R}^m$ が $a\in\Omega$ で微分可能とするとき、写像

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

を f の a における**線形化写像** (1 **次近似**) と呼ぶ。

上の式の形には見覚えがあるはずである (高校の数学で出て来た接線の式)。つまり

1 **変数実数値関数の線形化写像のグラフは、その関数のグラフの接線に他ならない**

多変数ベクトル値関数の場合も、線形化写像にそのような意味づけをすることが可能である。
次の「問題」は、そのままの形では試験には出しにくいが、本当はとても重要である。

例題 2.3.5 2 変数関数 $f(x,y)$ について、

$$f(0,0)=1,\quad \frac{\partial f}{\partial x}(0,0)=2,\quad \frac{\partial f}{\partial y}(0,0)=3$$

であることが分かっている。$f(0.1,0.2)$ の近似値を求めよ。

解 (他に情報がなければ、線形化写像の値を近似値に採用すべきだろう)

$$\begin{aligned}f(0.1,0.2) &\fallingdotseq f(0,0)+\frac{\partial f}{\partial x}(0,0)\cdot 0.1+\frac{\partial f}{\partial y}(0,0)\cdot 0.2\\ &= 1+2\cdot 0.1+3\cdot 0.2\\ &= 1+0.2+0.6=1.8. \quad \blacksquare\end{aligned}$$

さて C^1-級の実数値関数 $f\colon \mathbf{R}^n \supset \Omega \to \mathbf{R}$ のグラフ

$$\operatorname{graph} f \overset{\text{def.}}{=} \{(x,y) \in \Omega \times \mathbf{R}; y = f(x)\}$$

上の点 $(a, f(a))$ における接超平面を考えよう。

$$F(x,y) \overset{\text{def.}}{=} f(x) - y$$

とおくと、

$$\operatorname{graph} f = \{(x,y) \in \Omega \times \mathbf{R}; F(x,y) = 0\}.$$

よって、前小節の議論で接超平面の方程式が求まる。

$$\nabla F(x,y) = \begin{pmatrix} \frac{\partial F}{\partial x_1} \\ \vdots \\ \frac{\partial F}{\partial x_n} \\ \frac{\partial F}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \\ -1 \end{pmatrix}$$

であるから、 $(a, f(a)) = (a_1, \cdots, a_n, f(a))$ における接超平面の方程式は

$$\frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n - a_n) + (-1)(y - f(a)) = 0.$$

すなわち

$$y - f(a) = \nabla f(a) \cdot (x - a),$$

あるいは

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

かくして

高等学校の数学の「グラフの接線の方程式」は、
多変数関数のグラフの接超平面の方程式として一般化される
(表面上はまったく違いがない式で表される)

2.4 合成関数の微分法

「合成関数の導関数の公式 (chain rule) は微分法で最も重要な公式である」

という人もいるくらい、合成関数の導関数は良く現れる。このことは、微積分の基本的なテクニックの一つである変数変換は、実は合成関数を作ることだと気がつけば、納得しやすいであろう。

2.4.1 定理の陳述

1 変数の場合の合成関数の微分に関しては、次の定理があった:「$y = f(x)$, $z = g(y)$ がいずれも微分可能であれば $z = g(f(x))$ も微分可能で

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

つまり

$$(g \circ f)'(a) = g'(b)f'(a) \quad (\text{ただし } b = f(a).) \tag{2.2}$$

である。」

　この公式は多変数の場合にも自然に拡張できる。(2.2) において、右辺の積を行列の積とみなせばよい。

定理 2.4.1 (合成関数の微分法、連鎖律(れんさりつ) (chain rule(チェイン・ルール))) Ω, D をそれぞれ $\mathbf{R}^n$, $\mathbf{R}^m$ の開集合で、$a \in \Omega$, $f\colon \Omega \to \mathbf{R}^m$ は a で微分可能、 $g\colon D \to \mathbf{R}^\ell$ は $b \overset{\text{def.}}{=} f(a)$ で微分可能, $f(\Omega) \subset D$ とするならば、$g \circ f$ は a において微分可能で

$$(g \circ f)'(a) = g'(b)f'(a).$$

$y = f(x)$, $z = g(y)$ と書けば、上式の (i, j) 成分は次のように表される。

$$\frac{\partial z_i}{\partial x_j} = \sum_{k=1}^m \frac{\partial z_i}{\partial y_k}\frac{\partial y_k}{\partial x_j} \quad (1 \le i \le \ell,\ 1 \le j \le n).$$

乱暴な説明

$$\begin{aligned} z - z_0 &\fallingdotseq g'(y_0)(y - y_0), \\ y - y_0 &\fallingdotseq f'(x_0)(x - x_0) \end{aligned}$$

であるから、

$$z - z_0 \fallingdotseq g'(y_0)f'(x_0)(x - x_0).$$

ゆえに $(g \circ f)'(x_0) = g'(y_0)f'(x_0)$ である。 ■

　この議論を精密化したものが、以下の証明である。

証明　関数 f が a で微分可能であることから、

$$f(x) - f(a) - f'(a)(x - a) = \varepsilon_1(x) \quad (x \in \Omega)$$

で $\varepsilon_1(x)$ を定めると

$$\lim_{x \to a} \frac{\varepsilon_1(x)}{\|x - a\|} = 0.$$

関数 g が b で微分可能であることから、

$$g(y) - g(b) - g'(b)(y - b) = \varepsilon_2(y) \quad (y \in D)$$

で $\varepsilon_2(y)$ を定めると

$$\lim_{y \to b} \frac{\varepsilon_2(y)}{\|y - b\|} = 0. \tag{2.3}$$

$y = f(x)$ とすると第一式より

$$y - b = f'(a)(x - a) + \varepsilon_1(x)$$

となるので、

$$(g\circ f)(x) = g(f(x)) = g(y) = g(b)+g'(b)(y-b)+\varepsilon_2(y) = g(f(a))+g'(b)\left[f'(a)(x-a)+\varepsilon_1(x)\right]+\varepsilon_2(y) = g(f(a))+g'(b)f'(a)(x-a)+g'(b)\varepsilon_1(x)+\varepsilon_2(y). \tag{2.4}$$

この右辺第 3 項については

$$\frac{\|g'(b)\varepsilon_1(x)\|}{\|x-a\|}\le \|g'(b)\|\frac{\|\varepsilon_1(x)\|}{\|x-a\|}\to 0 \quad (x\to a).$$

以下は (2.4) 右辺第 4 項について考える。まず

$$\begin{aligned}\|y-b\| &= \|f'(a)(x-a)+\varepsilon_1(x)\|\le\|f'(a)(x-a)\|+\|\varepsilon_1(x)\|\\ &\le \|f'(a)\|\|x-a\|+\|\varepsilon_1(x)\|\end{aligned}$$

であり、x が a に十分近いとき $\|\varepsilon_1(x)\|\le\|x-a\|$ であるから

$$\|y-b\|\le(\|f'(a)\|+1)\|x-a\| \quad (x \text{ が } a \text{ に十分近いとき}). \tag{2.5}$$

そこで

$$M(y)\overset{\text{def.}}{=}\begin{cases}\dfrac{\|e_2(y)\|}{\|y-b\|} & (y\ne b)\\ 0 & (y=b)\end{cases}$$

とおくと

$$\frac{\|\varepsilon_2(y)\|}{\|x-a\|}=\frac{\|y-b\|}{\|x-a\|}M(y)\le(\|f'(a)\|+1)M(y)\to 0 \quad (x\to a).$$

(最初の不等号は $e_2(b)=0$ に注意する。第 2 の不等号は (2.5) そのもの。また (2.3) より、$y\to b$ のとき $M(y)\to 0$.) よって、

$$\varepsilon(x)\overset{\text{def.}}{=}g'(b)\varepsilon_1(x)+\varepsilon_2(y)$$

とおくとき

$$\lim_{x\to a}\frac{\|\varepsilon(x)\|}{\|x-a\|}=0,$$

$$(g\circ f)(x)=(g\circ f)(a)+g'(b)f'(a)(x-a)+\varepsilon(x).$$

これは $g\circ f$ が a で微分可能で、その微分係数が $g'(b)f'(a)$ であることを示している。■

系 2.4.1 $I = (\alpha, \beta)$ を $\mathbf{R}$ の区間,

$$\varphi: I \ni t \longmapsto \varphi(t) = \begin{pmatrix} \varphi_1(t) \\ \vdots \\ \varphi_n(t) \end{pmatrix} \in \mathbf{R}^n$$

は $t_0 \in I$ で微分可能、Ω は $\varphi(I) \subset \Omega$ なる $\mathbf{R}^n$ の開集合、$f: \Omega \to \mathbf{R}$ は $\varphi(t_0)$ で微分可能とするならば、$f \circ \varphi(t) = f(\varphi(t))$ は $t = t_0$ で微分可能で

$$\begin{aligned} \left.\frac{d}{dt} f(\varphi(t))\right|_{t=t_0} &= (f \circ \varphi)'(t_0) = f'(\varphi(t_0))\varphi'(t_0) \\ &= \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\varphi(t_0)) \frac{d\varphi_k}{dt}(t_0) = \nabla f(\varphi(t_0)) \cdot \frac{d\varphi}{dt}(t_0) \end{aligned}$$

が成り立つ。

(上で $t \mapsto \varphi(t)$ は t についての 1 変数関数なので、$\dfrac{\partial \varphi_k}{\partial t}$ は常微分 $\dfrac{d\varphi_k}{dt}$ となる。)

この系は、特に $\varphi(t) = a + tu$, $t_0 = 0$ のときが大事である。$f: \Omega \to \mathbf{R}$ が $a \in \Omega$ で微分可能であるとき、 $\forall u \in \mathbf{R}^n$ に対して

$$\left.\frac{d}{dt} f(a + tu)\right|_{t=0} = f'(a)u = \nabla f(a) \cdot u \qquad (\nabla f(a) \text{ と } u \text{ の内積}). \tag{2.6}$$

上の式の値のことを、a における f の u 方向の **(方向) 微分係数**と呼ぶ。

- 偏微分係数は $\dfrac{\partial f}{\partial x_i}$ は $\vec{e}_i = (\delta_{ij})$ 方向の微分係数である。実際

$$\frac{\partial f}{\partial x_i}(a) = \nabla f(a) \cdot \vec{e}_i.$$

- (2.6) は 「f が微分可能ならば、すべての方向 u に関する方向微分係数が存在して、それは勾配と u の内積に等しい」と読める。

- u が、ある曲面 S 上の点 a における S の外向き単位法線ベクトル n である場合が応用上よく現れる。これはしばしば

$$\frac{\partial f}{\partial n}(a)$$

  と書かれる。もちろん

$$\frac{\partial f}{\partial n}(a) = \nabla f(a) \cdot n.$$

2.4.2 簡単な例と注意

例 2.4.1 (1 次関数) $f(x) = Ax + p$ ($A \in M(m, n; \mathbf{R})$, $p \in \mathbf{R}^m$) で定義される $f: \mathbf{R}^n \to \mathbf{R}^m$ を 1 次関数 (アフィン写像、あるいはアファイン写像とも) と呼ぶ。既に見たように

$$f'(x) = A \quad (x \in \mathbf{R}^n).$$

さて、二つの 1 次関数

$$\begin{aligned} y &= f(x) = Ax + p, \\ z &= g(y) = By + q \end{aligned}$$

の合成は

$$z = (g \circ f)(x) = g(f(x)) = B(Ax + p) + q = B(Ax) + Bp + q = (BA)x + (Bp + q)$$

となりやはり 1 次関数で、

$$(g \circ f)'(x) = BA \quad (x \in \mathbf{R}^n).$$

ゆえに chain rule

$$(g \circ f)'(x) = g'(f(x))f'(x) \quad (x \in \mathbf{R}^n)$$

は ($BA = B \cdot A$ として) 確かに成り立っている。

例 2.4.2 二つの関数 $g = g(u, v, w)$, $h = h(x, y)$ が与えられたとき、

$$\varphi(x, y) \stackrel{\text{def.}}{=} g(x, y, h(x, y))$$

で定義される関数 φ の導関数は？

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} = \begin{pmatrix} x \\ y \\ h(x, y) \end{pmatrix} \stackrel{\text{def.}}{=} f(x, y)$$

とすると $\varphi = g \circ f$ となることに注意する。

$$\begin{aligned} \frac{\partial \varphi}{\partial x} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial x} = \frac{\partial g}{\partial u} \cdot 1 + \frac{\partial g}{\partial v} \cdot 0 + \frac{\partial g}{\partial w}\frac{\partial h}{\partial x} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial w}\frac{\partial h}{\partial x}, \\ \frac{\partial \varphi}{\partial y} &= \frac{\partial g}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial g}{\partial v}\frac{\partial v}{\partial y} + \frac{\partial g}{\partial w}\frac{\partial w}{\partial y} = \frac{\partial g}{\partial u} \cdot 0 + \frac{\partial g}{\partial v} \cdot 1 + \frac{\partial g}{\partial w}\frac{\partial h}{\partial y} = \frac{\partial g}{\partial v} + \frac{\partial g}{\partial w}\frac{\partial h}{\partial y}. \end{aligned}$$

注意 2.4.1 (ある不合理な習慣について) 上の例で、u, v はそれぞれ x, y に等しいのだから、わざわざ別の記号を持ち出すまでもないと考え、(記号をけちって) しばしば $g = g(x, y, z)$, $z = h(x, y)$ として

$$\frac{\partial \varphi}{\partial x} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial z}\frac{\partial h}{\partial x}, \quad \frac{\partial \varphi}{\partial y} = \frac{\partial g}{\partial y} + \frac{\partial g}{\partial z}\frac{\partial h}{\partial y}$$

と書くことがある。この場合 $\dfrac{\partial}{\partial x}$, $\dfrac{\partial}{\partial y}$ の意味が場所ごとに違うことに注意する必要がある。

一方、 φ という新しい文字を導入せずに

$$\frac{\partial g}{\partial x} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial w}\frac{\partial h}{\partial x}, \quad \frac{\partial g}{\partial y} = \frac{\partial g}{\partial v} + \frac{\partial g}{\partial w}\frac{\partial h}{\partial y}.$$

と書くことも多い。この場合は g は登場する位置 (左辺と右辺で) によって異なる意味を持つ。これも注意しないと混乱する。例えば、$u = x$ であるが、

$$\frac{\partial g}{\partial x} \neq \frac{\partial g}{\partial u}.$$

このように、微分法のこのあたりの記法は少々問題含みである。しかし古くから普及している記号法なので、今から別の書き方に換えることは難しい。使う側の心得としては、普段は適当にさぼったり、けちった書き方につき合うにしても、自分ではいざとなったら、現代的な流儀 (写像の定義域や終域をきちんと述べたりすること) で書けるようにしておくべきである。

例 2.4.3 (積の微分法) $\mathbf{R}^n$ の開集合 Ω で定義された微分可能な実数値関数 $f,\ g\colon \Omega \to \mathbf{R}$ に対して、

$$(fg)(x) = f(x)g(x) \quad (x \in \Omega)$$

で積 $fg\colon \Omega \to \mathbf{R}$ が定義できるが、

$$(fg)'(x) = g(x)f'(x) + f(x)g'(x).$$

(微分の定義に戻っても簡単に証明できるが、ここでは合成関数の微分法を利用した証明を示そう。次に示す二つの証明は、少し見栄えが違うが、本質的には同じものである。)

証明 1. fg は

$$G(u,v) = uv, \quad \begin{pmatrix} u \\ v \end{pmatrix} = F(x) = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$$

で定まる写像 $G\colon \mathbf{R}^2 \to \mathbf{R}$, $F\colon \Omega \to \mathbf{R}^2$ の合成 $G \circ F$ である。そして

$$G'(u,v) = \begin{pmatrix} \dfrac{\partial G}{\partial u} & \dfrac{\partial G}{\partial v} \end{pmatrix} = (v\ u)$$

より

$$G'(F(x)) = G'(f(x), g(x)) = (g(x)\ \ f(x)),$$

さらに

$$F'(x) = \begin{pmatrix} f'(x) \\ g'(x) \end{pmatrix}$$

であるから、chain rule より

$$(G \circ F)'(x) = G'(F(x))F'(x) = (g(x)\ \ f(x)) \begin{pmatrix} f'(x) \\ g'(x) \end{pmatrix} = g(x)f'(x) + f(x)g'(x).$$

証明 2. $y = uv$, $\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} f(x) \\ g(x) \end{pmatrix}$ の合成関数であるから、

$$\frac{\partial y}{\partial x_i} = \frac{\partial y}{\partial u}\frac{\partial u}{\partial x_i} + \frac{\partial y}{\partial v}\frac{\partial v}{\partial x_i} = v \cdot \frac{\partial f}{\partial x_i} + u \cdot \frac{\partial g}{\partial x_i} = g(x)\frac{\partial f}{\partial x_i} + f(x)\frac{\partial g}{\partial x_i}.$$

ゆえに

$$(fg)'(x) = \left(\frac{\partial y}{\partial x_1} \cdots \frac{\partial y}{\partial x_n} \right) = g(x) \left(\frac{\partial f}{\partial x_1} \cdots \frac{\partial f}{\partial x_n} \right) + f(x) \left(\frac{\partial g}{\partial x_1} \cdots \frac{\partial g}{\partial x_n} \right) = g(x)f'(x) + f(x)g'(x). \quad \blacksquare$$

例 2.4.4 (2 変数関数と曲線の合成) (1) $z=xy$, $x=\varphi(t)$, $y=\psi(t)$ とすれば、

$$\frac{dz}{dt}=\frac{\partial z}{\partial x}\frac{dx}{dt}+\frac{\partial z}{\partial y}\frac{dy}{dt}=z_x\varphi'+z_y\psi'=y\varphi'+x\psi'=\varphi'\psi+\varphi\psi'.$$

(2) $f(x,y)=\sqrt{x^2+y^2}$, $x=\varphi(t)$, $y=\psi(t)$ とすれば、

$$\frac{d}{dt}f\left(\varphi(t),\psi(t)\right)=\frac{\partial f}{\partial x}\frac{dx}{dt}+\frac{\partial f}{\partial y}\frac{dy}{dt}=\frac{x}{\sqrt{x^2+y^2}}\varphi'(t)+\frac{y}{\sqrt{x^2+y^2}}\psi'(t)=\frac{\varphi'(t)\varphi(t)+\psi(t)\psi'(t)}{\sqrt{\varphi(t)^2+\psi(t)^2}}.$$

例 2.4.5 (平面極座標と変数変換) 平面極座標

$$x=r\cos\theta,\quad y=\sin\theta\quad (r\ge 0,\ \theta\in\mathbf{R})$$

により $\varphi\colon(r,\theta)\to(x,y)$ を定義し、微分可能な関数 $f\colon\mathbf{R}^2\supset\Omega\ni(x,y)\mapsto f(x,y)\in\mathbf{R}$ に対して、$g\overset{\text{def.}}{=}f\circ\varphi$ とおく。つまり

$$g(r,\theta)=(f\circ\varphi)(r,\theta)=f(r\cos\theta,r\sin\theta).$$

これから

$$\begin{aligned}g_r &= f_xx_r+f_yy_r=f_x\cos\theta+f_y\sin\theta,\\ g_\theta &= f_xx_\theta+f_yy_\theta=-f_xr\sin\theta+f_yr\cos\theta.\end{aligned}$$

これを行列の形で表現しておくと

$$\begin{pmatrix}g_r & g_\theta\end{pmatrix}=\begin{pmatrix}f_x & f_y\end{pmatrix}\begin{pmatrix}x_r & x_\theta\\ y_r & y_\theta\end{pmatrix}=\begin{pmatrix}f_x & f_y\end{pmatrix}\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}.$$

応用上は、逆に f_x, f_y を g_r, g_θ で表す必要が生じることが多いが、これには上の式を f_x, f_y に関する連立 1 次方程式とみなして解けばよい。

$$\begin{aligned}\begin{pmatrix}f_x & f_y\end{pmatrix} &= \begin{pmatrix}g_r & g_\theta\end{pmatrix}\begin{pmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{pmatrix}^{-1}=\begin{pmatrix}g_r & g_\theta\end{pmatrix}\frac{1}{r}\begin{pmatrix}r\cos\theta & r\sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}\\ &= \begin{pmatrix}g_r & g_\theta\end{pmatrix}\begin{pmatrix}\cos\theta & \sin\theta\\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r}\end{pmatrix}.\end{aligned}$$

すなわち

$$\begin{aligned}f_x &= g_r\cos\theta-\frac{g_\theta}{r}\sin\theta,\\ f_y &= g_r\sin\theta+\frac{g_\theta}{r}\cos\theta.\end{aligned}$$

すぐ後で、別の方法 (本質的にはここで示した方法と同じ) でこの結果を導き直す[8]。■

[8] この例で示した計算をそのまま試験の答案に書く人が結構いて、それは決して間違いではないのだが、賢い (=勧められる) やり方ではないので注意したつもり。

例 2.4.6 (空間極座標と変数変換)

$$\begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{cases}$$

で、 $f\colon \mathbf{R}^3 \supset \Omega \ni (x,y,z) \mapsto f(x,y,z) \in \mathbf{R}$ を微分可能な関数としよう。このとき $g(r,\theta,\varphi) \overset{\text{def.}}{=} f(x,y,z)$ とおくと、

$$\begin{aligned} g_r &= f_x x_r + f_y y_r + f_z z_r = f_x \sin\theta\cos\varphi + f_y \sin\theta\sin\varphi + f_z\cos\theta, \\ g_\theta &= f_x x_\theta + f_y y_\theta + f_z z_\theta = f_x r\cos\theta\cos\varphi + f_y r\cos\theta\sin\varphi - f_z r\sin\theta, \\ g_\varphi &= f_x x_\varphi + f_y y_\varphi + f_z z_\varphi = -f_x r\sin\theta\sin\varphi + f_y r\sin\theta\cos\varphi. \end{aligned}$$

これを f_x, f_y, f_z について解くとどうなるか？ (生真面目にやると大変です。)

2.4.3 逆関数の微分法

上の二つの例で、最後に方程式を解いているのは、実は逆写像 (逆関数) の導関数を計算していることに相当する。高校数学の逆関数の微分法

$$\frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}}.$$

も多変数の場合に自然に拡張できる (逆数を逆行列と読み換えればよい)。

命題 2.4.1 (逆関数の微分法) U, V を $\mathbf{R}^n$ の開集合、 $f\colon U \to V$ は全単射で、 f と f^{-1} の両方とも微分可能であるとする。このとき

$$\left(f^{-1}\right)'(y) = (f'(x))^{-1}, \quad \text{ただし } y = f(x)$$

が成り立つ。

証明

$$f^{-1}(f(x)) = x, \quad f(f^{-1}(y)) = y$$

の両辺を微分して

$$(f^{-1})'(y) f'(x) = I, \quad f'(x)(f^{-1})'(y) = I$$

(ここで I は単位行列を表す) であるから $(f^{-1})'(y) = (f'(x))^{-1}$. ■

例 2.4.7 (平面極座標と変数変換ふたたび) 例 2.4.5 では、

$$(g_r\ g_\theta) = (f_x\ f_y) \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}$$

より

$$(f_x \ f_y) = (g_r \ g_\theta) \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}^{-1}.$$

として f_x, f_y を計算したが、上で述べた定理から、実は右辺の逆行列は

$$\begin{pmatrix} r_x & r_y \\ \theta_x & \theta_y \end{pmatrix}$$

に他ならない、ということである。これから、$r_x, r_y, \theta_x, \theta_y$ が求められる。これは応用上非常に重要なテクニックである。具体的に実行してみよう。

$$\begin{pmatrix} r_x & r_y \\ \theta_x & \theta_y \end{pmatrix} = \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.$$

ゆえに

$$r_x = \cos\theta, \quad r_y = \sin\theta, \quad \theta_x = \frac{-\sin\theta}{r}, \quad \theta_y = \frac{\cos\theta}{r}.$$

先にこれらを求めておけば、f_x, f_y は chain rule を使うだけで次のように計算できる。

$$f_x = g_r r_x + g_\theta \theta_x = g_r \cos\theta - \frac{g_\theta \sin\theta}{r}, \quad f_y = g_r r_y + g_\theta \theta_y = g_r \sin\theta + \frac{g_\theta \cos\theta}{r}. \quad \blacksquare$$

2.4.4 高階導関数について

合成関数の高階導関数、例えば 2 階導関数はどう計算するか?

1 変数実数値関数の場合

まず 1 変数の場合から考えよう。$y = g(u)$, $u = f(x)$ となっているとき、$y = g(f(x))$ の導関数は

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}$$

であるから、積の微分法によって

$$\frac{d^2y}{dx^2} = \frac{d}{dx}\left(\frac{dy}{du}\frac{du}{dx}\right) = \frac{d}{dx}\left(\frac{dy}{du}\right)\cdot\frac{du}{dx} + \frac{dy}{du}\cdot\frac{d}{dx}\left(\frac{du}{dx}\right).$$

右辺第 1 項については、1 階導関数のときと同様に

$$\frac{d}{dx}\left(\frac{dy}{du}\right)\cdot\frac{du}{dx} = \left[\frac{d}{du}\left(\frac{dy}{du}\right)\cdot\frac{du}{dx}\right]\frac{du}{dx} = \frac{d^2y}{du^2}\left(\frac{du}{dx}\right)^2$$

と計算できるから、

$$\frac{d^2y}{dx^2} = \frac{d^2y}{du^2}\left(\frac{du}{dx}\right)^2 + \frac{dy}{du}\frac{d^2u}{dx^2}.$$

多変数関数の場合

一般に書いても面倒なだけだから、2 変数関数の 2 階偏導関数のみ考える。原理は変数の個数が増えても、階数が高くなっても同じである。

$z=f(x,y)$ と $\begin{pmatrix} x \\ y \end{pmatrix}=\varphi(\xi,\eta)$ を合成した $z=f(\varphi(\xi,\eta))$ について、

$$\frac{\partial z}{\partial \xi}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial \xi}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial \xi},\quad \frac{\partial z}{\partial \eta}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial \eta}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial \eta}$$

が成り立つが、2 階偏導関数は次のようになる。

$$\begin{aligned}
\frac{\partial^2 z}{\partial \xi^2} &= \frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial \xi}\right)^2+2\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi}+\frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial \xi}\right)^2+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial \xi^2}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial \xi^2},\\
\frac{\partial^2 z}{\partial \eta^2} &= \frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial \eta}\right)^2+2\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \eta}+\frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial \eta}\right)^2+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial \eta^2}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial \eta^2},\\
\frac{\partial^2 z}{\partial \xi\partial \eta} &= \frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial \xi}\frac{\partial x}{\partial \eta}+\frac{\partial^2 z}{\partial x\partial y}\left(\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial x}{\partial \eta}\frac{\partial y}{\partial \xi}\right)+\frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial \xi}\frac{\partial y}{\partial \eta}+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial \xi\partial \eta}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial \xi\partial \eta}.
\end{aligned}$$

どれでも同じことなので、最初の式のみ示す。

$$\begin{aligned}
\frac{\partial^2 z}{\partial \xi^2} &= \frac{\partial}{\partial \xi}\frac{\partial z}{\partial \xi}=\frac{\partial}{\partial \xi}\left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial \xi}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial \xi}\right)\\
&= \left(\frac{\partial}{\partial \xi}\frac{\partial z}{\partial x}\right)\frac{\partial x}{\partial \xi}+\frac{\partial z}{\partial x}\cdot\frac{\partial}{\partial \xi}\frac{\partial x}{\partial \xi}+\left(\frac{\partial}{\partial \xi}\frac{\partial z}{\partial y}\right)\frac{\partial y}{\partial \xi}+\frac{\partial z}{\partial y}\cdot\frac{\partial}{\partial \xi}\frac{\partial y}{\partial \xi}\\
&= \left(\frac{\partial^2 z}{\partial x^2}\frac{\partial x}{\partial \xi}+\frac{\partial^2 z}{\partial y\partial x}\frac{\partial y}{\partial \xi}\right)\frac{\partial x}{\partial \xi}+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial \xi^2}+\left(\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial \xi}+\frac{\partial^2 z}{\partial y^2}\frac{\partial y}{\partial \xi}\right)\frac{\partial y}{\partial \xi}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial \xi^2}\\
&= \frac{\partial^2 z}{\partial x^2}\left(\frac{\partial x}{\partial \xi}\right)^2+2\frac{\partial^2 z}{\partial x\partial y}\frac{\partial x}{\partial \xi}\frac{\partial y}{\partial \xi}+\frac{\partial^2 z}{\partial y^2}\left(\frac{\partial y}{\partial \xi}\right)^2+\frac{\partial z}{\partial x}\frac{\partial^2 x}{\partial \xi^2}+\frac{\partial z}{\partial y}\frac{\partial^2 y}{\partial \xi^2}.
\end{aligned}$$

このやり方で、原理的には、どんな高階の偏導関数も計算できることが分かってもらいたい。

■

2.4.5 理解度判定問題

次の問は、結果も非常に重要であるし、結果を得る過程も理解度チェックに最適だ、と思われる。(1), (2) は既に例として説明済みだが、ノートを見ないでも解答できるようにしておこう。(3) だけが新しいが、これは前小節で学んだことの応用問題である。

問 (Laplacian の平面極座標による表示) 平面の極座標を考える。つまり $x=r\cos\theta$, $y=r\sin\theta$ とするとき、次の問に答えよ。

(1) (a) x_r, x_θ, y_r, y_θ を求めよ。

(b) r_x, r_y, θ_x, θ_y を求めよ。

(2) 関数 $f=f(x,y)$ が与えられているとき、$g(r,\theta)=f(x,y)$ で関数 g を定める。このとき

(a) g_r, g_θ を f_x, f_y で表せ。

(b) f_x, f_y を g_r, g_θ で表せ。

(3) 次の式が成り立つことを確かめよ。

$$\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2}=\frac{\partial^2 g}{\partial r^2}+\frac{1}{r}\frac{\partial g}{\partial r}+\frac{1}{r^2}\frac{\partial^2 g}{\partial \theta^2}.$$

(注意: 結果が分かっていれば、右辺を計算・整理して左辺に等しいことを示す方が簡単である。)

2.5　多変数の平均値の定理、Taylor の定理

この節の内容

f が点 a で微分可能であるとき、 $f(a+h)-f(a)$ は $h\to 0$ の極限においては $f'(a)h$ で近似されるわけだが、 h が有限[a]の大きさのときはどうなるであろうか？ 1 変数実数値関数の場合にこれに応える定理として、

1. 平均値の定理
2. Taylor の定理

を学んだ。この節では、これらの多次元化を目標にする。実は多変数実数値関数への拡張は自然に出来るが、ベクトル値関数への拡張は適当に定理を修正しないといけない。
多変数版定理の証明の要点は

$$F(t)\overset{\text{def.}}{=}f(a+th)\quad (t\in[0,1])$$

という 1 変数関数を導入して、 f の変化量を

$$f(a+h)-f(a)=F(1)-F(0)$$

と表して、問題を F に関する平均値の定理、 Taylor の定理に帰着させることにある。

[a]有限という言葉には二つの使い方がある。一つは無限大に対する有限で、単に $\mathbf{R}^n$ の要素であること。もう一つは無限小 or 0 に対する有限で、0 でない $\mathbf{R}^n$ の要素であること。今の場合は後者に相当する。

2.5.1　平均値の定理の多次元への拡張

平均値の定理の多変数への拡張は簡単　「多変数でも、(多次元ベクトル値でなく) 実数値関数ならば、平均値の定理が成り立つ。」

定理 2.5.1 (多変数の平均値の定理) Ω が $\mathbf{R}^n$ の開集合で、$f\colon\Omega\to\mathbf{R}$ が微分可能な関数、$a,\ b\in\Omega$ を端点とする線分が Ω に含まれている、つまり

$$[a,b]\overset{\text{def.}}{\equiv}\{(1-t)a+tb;t\in[0,1]\}\subset\Omega$$

とするならば、

$$\exists c\in[a,b]\ s.t.\ f(b)-f(a)=f'(c)(b-a).$$

証明

$$F(t)\overset{\text{def.}}{\equiv}f((1-t)a+tb)=f(a+t(b-a))\quad(t\in[0,1])$$

とおくと、合成関数の微分法より、$F\colon[0,1]\to\mathbf{R}$ は微分可能で

$$F'(t)=f'((1-t)a+tb)(b-a)$$

となる。1 変数の場合の平均値の定理を用いて $\exists\theta\in(0,1)$ s.t.

$$F(1)-F(0)=F'(\theta)1=F'(\theta).$$

すなわち

$$f(b)-f(a)=f'((1-\theta)a+\theta b)(b-a).$$

よって $c=(1-\theta)a+\theta b$ とすればよい。 ■

平均値の定理はベクトル値では駄目 多次元ベクトル値関数に対しては平均値の定理は成立しない。実際 $f\colon[0,2\pi]\to\mathbf{R}^2$ を $f(x)=\begin{pmatrix}\cos x\\ \sin x\end{pmatrix}$ とすると、もし平均値の定理が成り立つのならば、

$$\exists c\in(0,2\pi)\ \text{s.t.}\ f(2\pi)-f(0)=f'(c)(2\pi-0)=2\pi f'(c).$$

ところが $f(2\pi)-f(0)=\begin{pmatrix}0\\0\end{pmatrix}$, $f'(c)=\begin{pmatrix}-\sin c\\ \cos c\end{pmatrix}\neq\begin{pmatrix}0\\0\end{pmatrix}$ (なぜなら $\|f'(c)\|=1$) であるから、そういう c は存在し得ない。しかし、平均値の定理を用いる多くの場合は $\|f(a+h)-f(a)\|$ を上から評価する[9]目的に使うので、そういう場合は次の有限増分の公式があれば大丈夫である。

命題 2.5.1 (有限増分の公式) $f\colon[a,a+h]\to\mathbf{R}^m$ は連続、$(a,a+h)$ で微分可能、

$$\sup_{\theta\in(0,1)}\|f'(a+\theta h)\|\overset{\text{def.}}{\equiv}M<+\infty$$

とするならば

$$\|f(a+h)-f(a)\|\leq Mh.$$

[9]与えられた A に対して、適当な B を探して $A\leq B$ なる形の不等式を得ることを「A を上から評価する」という。

この公式の証明は難しくはなく、今すぐにでも出来るが、f に C^1-級くらいの滑らかさを仮定して、積分を使う方がずっと簡単なので後回しにする[10]。

あるいは、平均値の定理を以下述べるように修正すれば

$$\|f(a+h)-f(a)\| \le C\|h\| \quad (C \text{ は } h \text{ に関係ない定数})$$

の形の評価式を証明するのは難しくない。

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_m \end{pmatrix} : \Omega \to \mathbf{R}^m$$

において、各成分関数 f_i は実数値関数であるから

$$f_i(a+h) = f_i(a) + f_i'(a+\theta_i h)h \quad (1 \le i \le m)$$

となる $\theta_i \in (0,1)$ $(1 \le i \le m)$ は存在する。つまり、各成分ごとに異なる θ_i を必要とすることを我慢する。

2.5.2 Taylor の定理の多変数への拡張

ここでは f は実数値の多変数関数とする (ベクトル値とすると、前小節のような修正が必要になってやや面倒だから — 有限増分の公式だとそういうことはないのだが…)。

2 変数関数でウォーミング・アップ　簡単のため、一般的に考える前に、2 変数関数 $f = f(x,y)$ を 2 階の項まで展開してみよう。

$$f(a+h, b+k) - f(a,b)$$

を扱うために、

$$F(t) \overset{\text{def.}}{=} f(a+th, b+tk) \quad (t \in [0,1])$$

という 1 変数関数を導入する。合成関数の微分法によって

$$F'(t) = \frac{\partial f}{\partial x}(a+th, b+tk)h + \frac{\partial f}{\partial y}(a+th, b+tk)k$$

である。特に

$$F'(0) = \frac{\partial f}{\partial x}(a,b)h + \frac{\partial f}{\partial y}(a,b)k.$$

もう一度合成関数の微分法を用いて微分すると

$$F''(t) = \frac{\partial}{\partial x}\frac{\partial f}{\partial x}(a+th, b+tk)h \cdot h + \frac{\partial}{\partial y}\frac{\partial f}{\partial x}(a+th, b+tk)h \cdot k$$

[10] 昔々、偉い先生が「微分のことは (積分などの手は借りずに) 微分でしなくちゃ」とか言ったそうであるが (駄洒落のつもり)、はじめて学習するときは、明快な話運びが出来ることを尊重してもよいだろう。

$$
\begin{aligned}
&\quad + \frac{\partial}{\partial x}\frac{\partial f}{\partial y}(a+th,b+tk)k\cdot h + \frac{\partial}{\partial y}\frac{\partial f}{\partial y}(a+th,b+tk)k\cdot k \\
&= \frac{\partial^2 f}{\partial x^2}(a+th,b+tk)h^2 + \frac{\partial^2 f}{\partial y\partial x}(a+th,b+tk)hk \\
&\quad + \frac{\partial^2 f}{\partial x\partial y}(a+th,b+tk)kh + \frac{\partial^2 f}{\partial y^2}(a+th,b+tk)k^2 \\
&= \frac{\partial^2 f}{\partial x^2}(a+th,b+tk)h^2 + 2\frac{\partial^2 f}{\partial x\partial y}(a+th,b+tk)hk + \frac{\partial^2 f}{\partial y^2}(a+th,b+tk)k^2.
\end{aligned}
$$

1 変数関数 F に対して平均値の定理を適用すると、$\exists\theta \in (0,1)$ s.t.

$$
\begin{aligned}
f(a+h)-f(a) &= F(1)-F(0) = F'(0)\cdot 1 + \frac{1}{2}F''(\theta)\cdot 1^2 \\
&= F'(0) + \frac{1}{2}F''(\theta) \\
&= \frac{\partial f}{\partial x}(a,b)h + \frac{\partial f}{\partial y}(a,b)k \\
&+ \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(a+\theta h,b+\theta k)h^2 + 2\frac{\partial^2 f}{\partial x\partial y}(a+\theta h,b+\theta k)hk + \frac{\partial^2 f}{\partial y^2}(a+\theta h,b+\theta k)k^2\right).
\end{aligned}
$$

これが 2 変数関数の 2 階までの Taylor 展開の公式である。以下で導く一般の場合の公式はかなり複雑であるが、導出の原理はまったく同様である。

n 変数関数の m 次微分　さて、それでは一般の f に対して考察を始めよう。関数 $F(t) = f(a+tu)$ の 1 階導関数については

$$
F'(t) = \frac{d}{dt}\left[f(a+tu)\right] = f'(a+tu)u
$$

という結果があり、これが前小節の議論の基礎となったわけだが、まずこれを高階の導関数まで拡張しよう。

補題 2.5.1 Ω は $\mathbf{R}^n$ の開集合、$f\colon\Omega\to\mathbf{R}$ は C^k- 級 (k は自然数), $a\in\Omega$, $h = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \in \mathbf{R}^n$,

$$
[a,a+h] \overset{\text{def.}}{\equiv} \{a+th; t\in[0,1]\} \subset \Omega
$$

とするとき、

$$
F(t) \overset{\text{def.}}{=} f(a+th) \quad (t\in[0,1])
$$

とおくならば $F\colon[0,1]\to\mathbf{R}$ は C^k-級で

$$
F^{(m)}(t) = \sum_{1\le i_1,i_2,\cdots,i_m\le n} \frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a+th)h_{i_1}h_{i_2}\cdots h_{i_m} \quad (1\le m\le k).
$$

証明　帰納法による。$m=1$ の場合は系 2.4.1 で済んでいる。m のとき成立すると仮定しよう:

$$F^{(m)}(t)=\sum_{1\le i_1,i_2,\cdots,i_m\le n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a+th)h_{i_1}h_{i_2}\cdots h_{i_m}.$$

すると $\dfrac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a+th)$ に関する合成関数の微分法により、

$$\begin{aligned}F^{(m+1)}(t)=\frac{d}{dt}F^{(m)}(t) &= \sum_{1\le i_1,i_2,\cdots,i_m\le n}\frac{d}{dt}\left[\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a+th)\right]h_{i_1}h_{i_2}\cdots h_{i_m}\\ &= \sum_{1\le i_1,i_2,\cdots,i_m\le n}\left[\sum_{i=1}^n\frac{\partial}{\partial x_i}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a+th)\,h_i\right]h_{i_1}h_{i_2}\cdots h_{i_m}\\ &= \sum_{1\le i_1,i_2,\cdots,i_m,i_{m+1}\le n}\frac{\partial^{m+1} f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}\partial x_{i_{m+1}}}(a+th)h_{i_1}h_{i_2}\cdots h_{i_m}h_{i_{m+1}}.\end{aligned}$$

($(i,i_1,i_2,\cdots,i_m)$ を $(i_1,i_2,\cdots,i_{m+1})$ と書き換えた。)

これは $m+1$ でも成立することを示している。■

特に

$$F^{(m)}(0)=\sum_{1\le i_1,i_2,\cdots,i_m\le n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a)h_{i_1}h_{i_2}\cdots h_{i_m}$$

であるが、これを f の a における m **次微分**と呼び、$(d^m f)_a(h)$ と書く[11]。 $(d^m f)_a(h)$ は h に関する m 次同次多項式 (m **次形式**) である。

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、上式の $\sum$ には同類項が含まれている。まとめるとどうなるか？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、

$$\begin{aligned}(d^2 f)_a(h) &= \sum_{i,j=1}^2\frac{\partial^2 f}{\partial x_i\partial x_j}(a)h_ih_j\\ &= \frac{\partial^2 f}{\partial x_1\partial x_1}(a)h_1h_1+\frac{\partial^2 f}{\partial x_1\partial x_2}(a)h_1h_2+\frac{\partial^2 f}{\partial x_2\partial x_1}(a)h_2h_1+\frac{\partial^2 f}{\partial x_2\partial x_2}(a)h_2h_2\\ &= \frac{\partial^2 f}{\partial x_1^2}(a)h_1^2+2\frac{\partial^2 f}{\partial x_1\partial x_2}(a)h_1h_2+\frac{\partial^2 f}{\partial x_2^2}(a)h_2^2.\end{aligned}$$

一般には

$$\begin{aligned}(d^m f)_a(h) &= \sum_{1\le i_1,i_2,\cdots,i_m\le n}\frac{\partial^m f}{\partial x_{i_1}\partial x_{i_2}\cdots\partial x_{i_m}}(a)h_{i_1}h_{i_2}\cdots h_{i_m}\\ &= \left(h_1\frac{\partial}{\partial x_1}+h_2\frac{\partial}{\partial x_2}+\cdots+h_n\frac{\partial}{\partial x_n}\right)^m f(a)\\ &= \sum_{\alpha_1+\alpha_2+\cdots+\alpha_n=m}\frac{m!}{\alpha_1!\alpha_2!\cdots\alpha_n!}\frac{\partial^m f}{\partial x_1^{\alpha_1}\partial x_2^{\alpha_2}\cdots\partial x_n^{\alpha_n}}(a)h_1^{\alpha_1}h_2^{\alpha_2}\cdots h_n^{\alpha_n}.\end{aligned}$$

[11] 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。

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**[12]

$$\begin{aligned}(a_1 + a_2 + \cdots + a_n)^m &= \sum_{1 \le i_1, i_2, \cdots, i_m \le n} a_{i_1} a_{i_2} \cdots a_{i_m} \\ &= \sum_{\substack{\alpha_1 + \alpha_2 + \cdots + \alpha_n = m \\ \alpha_j \text{ は非負整数}}} \frac{m!}{\alpha_1! \alpha_2! \cdots \alpha_n!} {a_1}^{\alpha_1} {a_2}^{\alpha_2} \cdots {a_n}^{\alpha_n}\end{aligned}$$

を用いて、多少形式的な[13]計算を行なった。

以上をまとめておこう。

定理 2.5.2 [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、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:

$$f(a+h) = \sum_{m=0}^{k-1} \frac{1}{m!} \left(d^m f\right)_a (h) + \frac{1}{k!} \left(d^k f\right)_{a+\theta h} (h).$$

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。

$$\begin{aligned}\left(d^m f\right)_x (h) &= \sum_{1 \le i_1, i_2, \cdots, i_m \le n} \frac{\partial^m f}{\partial x_{i_1} \partial x_{i_2} \cdots \partial x_{i_m}} (x) h_{i_1} h_{i_2} \cdots h_{i_m} \\ &= \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^m f(a) \\ &= \sum_{\alpha_1 + \alpha_2 + \cdots + \alpha_n = m} \frac{m!}{\alpha_1! \alpha_2! \cdots \alpha_n!} \frac{\partial^m f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}} (x)\, h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_n^{\alpha_n}.\end{aligned}$$

証明 $F(t) \overset{\text{def.}}{\equiv} f(a+th)$ は $[0,1]$ で C^k-級で、1 変数関数についての Taylor の定理から $\exists \theta \in (0,1)$ s.t.

$$F(1) = \sum_{m=0}^{k-1} F^{(m)}(0) + \frac{1}{k!} F^{(k)}(\theta)$$

後は補助定理 2.5.1 より

$$F^{(m)}(0) = (d^m f)_a(h), \quad F^{(m)}(\theta) = (d^m f)_{a+\theta h}(h)$$

であるから、結論を得る。■

---

[12] 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。

[13] 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。

次の小節で説明する **Schwartz の多重指数** (**multi-index**) の記法を用いると、

$$(d^m f)_x (h) = m! \sum_{|\alpha|=m} \frac{f^{(\alpha)}(x)}{\alpha!} h^\alpha.$$

よって

$$f(a+h) = \sum_{0\le|\alpha|\le k-1} \frac{f^{(\alpha)}(a)}{\alpha!} h^\alpha + R_k, \quad R_k = \sum_{|\alpha|=k} \frac{f^{(\alpha)}(a+\theta h)}{\alpha!} h^\alpha$$

と 1 変数の場合と非常に良く似た形をした式が得られる[14]。

剰余項 (**remainder**) R_k を Landau の記号を用いて書くと、次のようになる[15]。

系 2.5.1 定理 2.5.2 と同じ仮定の下で、

$$f(a+h) = \sum_{0\le|\alpha|\le k-1} \frac{f^{(\alpha)}(a)}{\alpha!} h^\alpha + O(\|h\|^k) = \sum_{0\le|\alpha|\le k-1} \frac{f^{(\alpha)}(a)}{\alpha!} h^\alpha + o(\|h\|^{k-1}) \quad (h \to 0).$$

2.5.3 余談あれこれ

この小節に書いてあることはいずれもかなり役に立つものであるが、はじめて勉強するときは省略しても構わない。

Taylor の定理の逆

Taylor の定理の逆に相当する次の命題 (証明略) は、覚えておくと便利である。

命題 2.5.2 定理 2.5.2 と同じ仮定の下で、

$$f(a+h) = \sum_{0\le|\alpha|\le k-1} C_\alpha h^\alpha + o(\|h\|^{k-1}) \quad (h \to 0) \tag{2.7}$$

が成り立っているならば、

$$C_\alpha = \frac{f^{(\alpha)}(a)}{\alpha!} \quad (|\alpha| \le k-1).$$

これから、とにかく (2.7) の形に書ければ、主要部が Taylor の定理のそれと一致することが分かる。

例 2.5.1 $f(x_1, x_2) = \exp(x_1^2 + x_2^2)$ を原点のまわりで 4 階の項まで展開してみよう。$e^t = 1 + t + \frac{1}{2}t^2 + o(|t|^2)$ より、

$$e^{x_1^2+x_2^2} = 1 + (x_1^2+x_2^2) + \frac{(x_1^2+x_2^2)^2}{2} + o((x_1^2+x_2^2)^2) = 1 + (x_1^2+x_2^2) + \frac{(x_1^2+x_2^2)^2}{2} + o(\|x\|^4)$$

[14]筆者自身はこちらの書き方の方が覚えやすいと感じていて、こちらだけ記憶していた時期があった。もっともこれは偏微分方程式の勉強をして、多重指数の記法に慣れたせいかもしれない。

[15]実はこの形にしておけば、ベクトル値関数でも成立する。一番暗記向きな公式かも知れない。

$$= \quad 1 + x_1^2 + x_2^2 + \frac{1}{2}x_1^4 + x_1^2 x_2^2 + \frac{1}{2}x_2^4 + o(\|x\|^4).$$

であるが、上の命題から Taylor の定理の展開に他ならないことが分かる。ゆえに

$$\frac{\partial f}{\partial x_1}(0,0) = \frac{\partial f}{\partial x_2}(0,0) = 0,$$

$$\frac{\partial^2 f}{\partial x_1^2}(0,0) = \frac{\partial^2 f}{\partial x_2^2}(0,0) = 2! \cdot 1 = 2, \quad \frac{\partial^2 f}{\partial x_1 \partial x_2}(0,0) = 0,$$

$$f^{(\alpha)}(0,0) = 0 \quad (|\alpha| = 3 \text{ なる任意の } \alpha \in (\mathbf{N} \cup \{0\})^2),$$

$$\begin{aligned}
\frac{\partial^4 f}{\partial x_1^4}(0,0) &= \frac{\partial^4 f}{\partial x_2^4}(0,0) = 4! \cdot \frac{1}{2} = 12, \quad \frac{\partial^4 f}{\partial x_1^2 \partial x_2^2}(0,0) = 2!2! \cdot 1 = 4, \\
f^{(\alpha)}(0,0) &= 0 \quad (|\alpha| = 4 \text{ かつ } \alpha \neq (4,0), (0,4), (2,2) \text{ なる任意の } \alpha). \quad \blacksquare
\end{aligned}$$

多重指数の記法

(色々な公式が、1 次元のときと良く似た公式で書けることを面白く感じてくれれば幸いだが、こういう記号は肌にあわない、と感じたら無理に覚える必要はない。)

以下 $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, $\beta = (\beta_1, \beta_2, \cdots, \beta_n)$ で、各 α_j, β_j は 0 以上の整数であるとする。

$$\begin{aligned}
|\alpha| &\overset{\text{def.}}{=} \alpha_1 + \alpha_2 + \cdots + \alpha_n, \\
\alpha! &\overset{\text{def.}}{=} \alpha_1! \alpha_2! \cdots \alpha_n!, \\
h^\alpha &\overset{\text{def.}}{=} h_1^{\alpha_1} h_2^{\alpha_2} \cdots h_n^{\alpha_n}, \\
\left(\frac{\partial}{\partial x}\right)^\alpha &\overset{\text{def.}}{=} \left(\frac{\partial}{\partial x_1}\right)^{\alpha_1} \left(\frac{\partial}{\partial x_2}\right)^{\alpha_2} \cdots \left(\frac{\partial}{\partial x_n}\right)^{\alpha_n}, \quad f^{(\alpha)} \overset{\text{def.}}{=} \left(\frac{\partial}{\partial x}\right)^\alpha f.
\end{aligned}$$

さらに

$$\begin{aligned}
\alpha \geq \beta &\overset{\text{def.}}{\Leftrightarrow} \alpha_1 \geq \beta_1,\ \alpha_2 \geq \beta_2,\ \cdots,\ \alpha_n \geq \beta_n, \\
\alpha > \beta &\overset{\text{def.}}{\Leftrightarrow} \alpha \geq \beta,\ \alpha \neq \beta
\end{aligned}$$

と定義し、$\alpha \geq \beta$ のとき

$$\binom{\alpha}{\beta} \overset{\text{def.}}{=} \frac{\alpha!}{\beta!\,(\alpha - \beta)!}$$

とおく。

この記法は偏微分方程式論などではバンバン使われる。微積分の段階でも、Taylor の定理だけでなく色々使い道がある。例えば、積の微分法 $(fg)' = f'g + fg'$ の一般化である **Leibniz**(ライプニッツ) **の公式**[16]は

$$\left(\frac{\partial}{\partial x}\right)^\alpha (fg) = \sum_{\beta \leq \alpha} \binom{\alpha}{\beta} f^{(\beta)} g^{(\alpha - \beta)}$$

のように表される。

[16] 1 変数実数値関数の場合は $(fg)^{(k)} = \sum_{r=0}^{k} \binom{k}{r} f^{(r)} g^{(k-r)}$ であった。

積分を利用した評価

関数の増分 $f(a+h)-f(a)$ を評価するのに、積分を用いるのが有効なことも多い。$f\in C^1$ など、少し強い仮定が必要になるが、簡単で強力である。

$$f(a+h)-f(a)=[f(a+th)]_{t=0}^{t=1}=\int_0^1\frac{d}{dt}\left(f(a+th)\right)\,dt=\int_0^1 f'(a+th)h\,dt.$$

この式から有限増分の公式はすぐに導かれる。実際

$$\|f(a+h)-f(a)\|\le\int_0^1\|f'(a+th)h\|dt\le\max_{t\in[0,1]}\|f'(a+th)h\|\int_0^1 dt\le M\|h\|,$$

ただし

$$M\overset{\text{def.}}{=}\max_{t\in[0,1]}\|f'(a+th)\|.$$

(f は C^1-級であるから、$t\mapsto\|f'(x+th)\|$ は連続関数であり、コンパクト集合 $[0,1]$ 上最大値を持つ。)

後は、少しマイナーな話題になるが、同様の積分表示式から、部分積分を繰り返すことにより

$$f(b)=\sum_{j=0}^{k}\frac{f^{(j)}(a)}{j!}(b-a)^j+R_{k+1},\quad R_{k+1}=\frac{1}{k!}\int_a^b f^{(k+1)}(t)(b-t)^k dt.$$

が導かれる。この形の剰余項を **Schlömlich の剰余項**という。

$$R_{k+1}=\frac{f^{(k+1)}(c)}{k!}(b-c)^k(b-a)\quad(c\in[a,b])$$

を **Cauchy の剰余項**とよぶ。

2.5.4 Taylor の定理記憶のススメ

多変数の Taylor の定理の公式はなかなか複雑である。筆者自身、正確な式を覚えるようになったのは、微分積分学を学んで大部後になってからである。自分が良く覚えていられないものを記憶することを他人に強要するつもりはない。しかし、知らないよりは知っている方が有利なことがあるのも真実なので、覚えたいという人はチャレンジすることを勧める[17]。定理 2.5.2 とその後の多重指数の記法による書換えを眺めて、どちらか一方、気に入った方だけでも覚えるとよいだろう。

定理そのものを正確に暗記出来ない (or その気になれない) という場合は、次のようなことだけ身につけておけばよい。

$$\begin{aligned}
f(a+h) &= f(a)+f'(a)h+\frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i\partial x_j}(a)h_ih_j\\
&\quad+\sum_{m=3}^{k-1}[(f\text{ の }m\text{ 階微分係数})\times(h\text{ の }m\text{ 次単項式})\text{ の和}]\\
&\quad+k\text{ 次剰余項 }R_k,\\
k\text{ 次剰余項 }R_k &= (f\text{ の }k\text{ 階微分係数})\times(h\text{ の }k\text{ 次単項式})\text{ の和}.
\end{aligned}$$

[17] この手の話になると、「丸暗記は馬鹿馬鹿しい。そういう場合は公式の導き方を覚えることにして、必要ならばその場で導出できるようにしておくのが良い。」という意見が出ることがある。今の場合も、大変もっともなところがあると感じられるが、一方で公式そのものが教えてくれる「知見」のようなものもあり、公式を真剣に見つめる時間を持たないのももったいないことだと思う。公式を結果として覚えてしまうくらい眺める時間を持つのは決して無駄ではあるまい。

要点は次のようなものである。

- Taylor の定理とは、つまりは多項式による近似の話である。
- 一般項は、f の m 階微分係数と h の m 次単項式の積である。
- 1 次の項は (もう大部慣れて来たはずの) $f'(a)h$ である。
- 2 次の項は、まだ慣れていないかもしれないが、重要なのでこの機会に覚えよう! 次の節で詳しく説明することになる、f の Hesse 行列 $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)$ を係数とする、 h の 2 次形式である。
- R_k を構成する項は、主要部の一般項とよく似ている。ただし a における値ではなくて、 $a+\theta h$ における値 ($\theta \in (0,1)$) であり、大きさは $O(\|h\|^k)$ である。

2.6 極値問題への応用

この節では、1 変数関数の微分法ではよく知られた

$$f'(a)=0,\ f''(a)>0 \Longrightarrow f \text{ は } a \text{ で極小}$$

という定理を多変数関数に拡張する (かなり長いが、やっているのはそれだけ)。

2.6.1 極値問題の概説

この節では、微分法の応用のうちでも最もポピュラーな極値問題を扱う。結果を得る過程 (主に Taylor の定理を利用することになる) も重要である。値の大小の話であるから、ベクトル値関数を考えるのはナンセンスで、実数値関数が考察の対象になる。

1 変数実数値関数についての「$f\colon A \to \mathbf{R}$ の最大値を求めよ!」という類の問題に対しては以下のようなことを調べるのであった。

(1) 定義域 A の境界点を調べる。

(2) $f'(a)=0$ となる点 (このような点 a を f の**停留点**(ていりゅうてん)と呼ぶ) を調べる。

(3) f が微分可能でない点を何らかの方法で調べる[18]。

多変数関数については (1) が重要になる問題も多い[19]。これについてある程度は、後の**条件付き極値問題** (Lagrange の未定乗数法) という節で考察する。

(3) については (基本的には) 微積分の範囲外ということで省略する。

この節では (2) についてのみ考察する。まず、極値点の候補を探すために使える定理

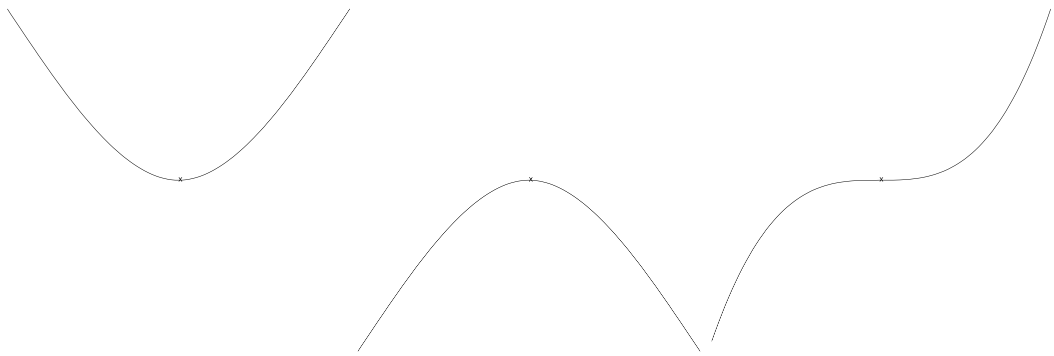
$$f \text{ が定義域の内点 } a \text{ で極値を取るならば } f'(a)=0$$

[18] たとえば $f(x)=1-\sqrt{|x|}$ の場合、f が微分可能でない点 0 で f は最大となる。

[19] 1 変数関数で、定義域が区間であった場合など、境界点は高々 2 つの点だけであるから、考えやすい (一つずつ調べても手間は大したことがない)。これに対して、多変数関数の場合は、境界は大抵は無限個の点を含む。

は多変数関数に対してもそのまま成り立つ (すぐ後で定理 2.6.1 としてきちんと証明する)。

ところで、上の定理を背景に、$f'(a) = 0$ を満たす点 a を探し出したとして、その点で本当に f が極値になっているかどうかを調べるにはどうしたらよいか?上の定理の逆は成立しない (すなわち、$f'(a) = 0$ が成り立つからと言って、f が a で極値を取るとは限らない — 例えば $f(x) = x^3$, $a = 0$ など) し、極値と言っても極大・極小のどちらなのかも含めて、何か別の手段を使って調べなければならない。



(a) 極小　　(b) 極大　　(c) 極値ではない

図 2.1: 停留点と言っても色々ある。

1 変数関数に対しては、関数の増減を調べるために、**増減表**というとても便利なものがあったが、多変数関数に対して増減表を書くことは不可能に近い。多変数関数に対して拡張できるのは、f の 2 階導関数の符号を用いてチェックする、次の定理である[20]。

1 変数関数の場合の極値の判定

$f'(a) = 0$ かつ $f''(a) > 0$ ならば f は a で狭義の極小.

実際、次の定理が成り立つ。

多変数関数の場合の極値の判定

$f'(a) = 0$ かつ H_a が**正値**ならば f は a で狭義の極小.

(ここで H_a は f の a における**Hesse**(ヘッセ)**行列**)

初めて学ぶ人は、ここで次の二つのことが当然の疑問として浮かんで来るであろう。

(1) Hesse 行列とは何か?

(2) 行列が正値であるとはどういうことか?

(1) に対する答: f の a における **Hesse 行列**とは、a における f の 2 階の偏微分係数を並

[20] $f''(a)$ の正・負と極大・極小のどっちがどっちに対応しているか分からなくなったら、2 次関数 $f(x) = px^2$ を考えるといい。$f''(0) = 2p > 0$ ならばグラフは下に凸な放物線なので当然極小で、$f''(0) = 2p < 0$ ならばグラフは上に凸な放物線なので当然極大である。

べて作った行列である:

$$H_a \overset{\text{def.}}{=} \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_2 \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \cdots & \vdots \\ \dfrac{\partial^2 f}{\partial x_n \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{pmatrix}.$$

これを見ると、f の a における Hesse 行列のことを $f''(a)$ と書いてしまいたい!、そう思うかもしれない。残念ながらこう書く人はいないようだが (いたとしてもごく小数派 — 真似しないこと)、直感的には Hesse 行列とは 2 階の微分係数と考えて良い。

(2) に対する**とりあえずの**答: 行列が正値であるとは、その行列の固有値がすべて正の数であることである。

次節以降は以下の点に注意して学んで欲しい。

- 何で H_a が出てくるのか (証明に Taylor の定理がどのように使われるのか)。
- 行列が正値であることのホントの意味 (固有値が正というだけではよく分からないでしょ)。
- 行列の固有値の符号の判定はどうするのか。

注意 2.6.1 (「狭義の極小」とは?) $a \in \Omega$, $f \colon \Omega \to \mathbf{R}$ に対して、f が a で**極小**とは、$f(a)$ が a の十分近くでの f の最小値になっていることであった:

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

ところで、この定義では、$f(a)$ が唯一の最小値であることは要求していない。唯一の最小であるとき、すなわち

$$\exists \varepsilon > 0\ \forall x \in (B(a;\varepsilon) \setminus \{a\}) \cap \Omega \quad f(x) > f(a).$$

であるとき、f は a で**狭義の極小**である、ということにする。

注意 2.6.2 (他の本を読むための注意) 極大、極小の定義にも色々ある。このテキストの「狭義の極小」を「極小」の定義に採用している本もある。また内点以外では極大・極小を定義しない流儀もある。おかしいと思ったら、その著者がその言葉をどういう意味で使っているか注意深く調べること。

さて、上で約束した定理とその証明を書いておこう。

定理 2.6.1 (内点で極値を取れば、微分係数は 0) Ω は $\mathbf{R}^n$ の開集合、$f \colon \Omega \to \mathbf{R}$ は微分可能、$a \in \Omega$ で f は極値を取るならば、$f'(a) = 0$.

証明　各 $i \in \{1, \cdots, n\}$ を固定するごとに、1 変数関数

$$\varphi\colon (a_i - \varepsilon, a_i + \varepsilon) \ni t \mapsto f(a_1, \cdots, a_{i-1}, t, a_{i+1}, \cdots, a_n) \in \mathbf{R}$$

について考える (ε は十分小さな正の数) と、これは $t = a_i$ で極値を取る。従って、1 変数関数についての定理から $\varphi'(a_i) = 0$ であるが、これは $\dfrac{\partial f}{\partial x_i}(a) = 0$ を意味する。ゆえに $f'(a) = 0$. ■

以下では、簡単のため Ω が $\mathbf{R}^2$ の開集合で、f が Ω 上の C^1-級関数である場合に、上の定理の主張

$$\text{内点 } a \text{ が } f \text{ の極値点} \Longrightarrow a \text{ は } f \text{ の停留点 i.e. } \nabla f(a) = 0.$$

の図形的な意味を考えよう。

図形的な解釈 (1)

f のグラフ

$$\operatorname{graph} f \overset{\text{def.}}{=} \{(x, y, z); (x, y) \in \Omega, z = f(x, y)\}$$

は関数

$$F(x, y, z) \overset{\text{def.}}{=} f(x, y) - z$$

の零点集合である。この曲面の法線ベクトルは

$$\nabla F = \begin{pmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial z} \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ -1 \end{pmatrix}.$$

よって

$$\begin{aligned} \nabla f(a) = 0 \quad &\Leftrightarrow \quad \nabla F(a) \text{ が } \vec{e}_3 \text{ と平行} \\ &\Leftrightarrow \quad \text{グラフ上の点 } (a, f(a)) \text{ における法線が } \vec{e}_3 \text{ と平行.} \end{aligned}$$

図形的な解釈 (2)

もしも $\nabla f(a) \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ ならば a は極値点ではないことを、もう少し具体的に見てみよう。Ω を地図と考え、f の値を「標高」と考えよう。

$$F(t) \overset{\text{def.}}{=} f(a + t\nabla f(a)) \quad (|t| \text{ は十分小})$$

とおくと、

$$F'(0) = f'(a)\nabla f(a) = \|\nabla f(a)\|^2 > 0$$

したがって $\exists \varepsilon > 0$ s.t.

$$F(t) < F(0) \quad (t \in (-\varepsilon, 0)),$$

$$F(t) > F(0) \quad (t \in (0, \varepsilon)).$$

つまり a から $\nabla f(a)$ の方向に移動すると標高が高くなり、$-\nabla f(a)$ の方向に移動すると標高が低くなる (だから a は山でも谷でもありえず、坂の途中である)。これから a は f の極値点ではないことが分かる。対偶を取ると

$$a \text{ が } f \text{ の極値点ならば } \nabla f(a) = 0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

が得られる。

2.6.2　1 変数関数の場合 — 凸関数と 2 階導関数

(詳しくは付録 K を参照。)

定義 2.6.1 (凸関数) I を $\mathbf{R}$ の区間、$f\colon I \to \mathbf{R}$ とするとき、

$$f \text{ が}\overset{\text{とつかんすう}}{\textbf{凸関数}}\ (\textbf{convex function}) \overset{\text{def.}}{\Leftrightarrow} \begin{array}{l} \forall a \in I,\ \forall b \in I,\ \forall t \in (0,1) \\ f(ta+(1-t)b) \le tf(a)+(1-t)f(b). \end{array}$$

命題 2.6.1 (2 階導関数の値と凸性、極値) $\mathbf{R}$ の区間 I で f'' が存在して、$f'' \ge 0$ on I であるとき、以下のことが成立する。

(1) 任意の $a, x \in I$ に対して, $f(x) \ge f(a) + (x-a)f'(a)$.

(2) f は I で凸。

(3) $f'(\alpha) = 0$ となる $\alpha \in I$ が存在すれば、α は f の最少点である。

証明　(すべて付録に書いてあるから) 特に重要な (3) のみ示す。Taylor の定理 K.2.1 を $k = 2$ として使うと、α, x の間に c が取れて、

$$\begin{aligned} f(x) &= f(\alpha) + f'(\alpha)(x-\alpha) + \frac{f''(c)}{2}(x-\alpha)^2 \\ &= f(\alpha) + \frac{f''(c)}{2}(x-\alpha)^2 \ge f(\alpha) \quad (\because f''(c) \ge 0,\ (x-\alpha)^2 \ge 0). \end{aligned}$$

ゆえに $f(x)$ は $x = \alpha$ で最小となる。■

2.6.3　線形代数から — 2 次形式

$\mathbf{R}^n$ 上の m **次形式**とは、n 個の実変数 x_1, x_2, $\cdots$, x_n の m 次同次多項式 $F(x_1, x_2, \cdots, x_n)$ のことをいう。以下では $\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x$ と書き、$F(x_1, x_2, \cdots, x_n) = F(x)$ のように表す。

(関数 F が m 次**同次** (**斉次**とも言う) とは、$F(\lambda x) = \lambda^m F(x) \quad (\forall \lambda \in \mathbf{R}, \forall x \in \mathbf{R}^n)$ が成り立つことを言う。m 次形式は、ちょうど m 次の単項式の和として表される多項式である。)

任意の 1 次形式 $L(x)$ は適当な n 次元ベクトル $a = {}^t(a_1, a_2, \cdots, a_n) \in \mathbf{R}^n$ によって $L(x) = \sum_{j=1}^{n} a_j x_j$ の形に表すことが出来る。この式の右辺は内積を使うと (a, x) と書ける。写像 $x \mapsto L(x)$ は $\mathbf{R}^n$ から $\mathbf{R}$ への線形写像になっている。

一方 $\mathbf{R}^n$ 上の任意の 2 次形式 $Q(x)$ は適当な n 次正方行列 $H = (a_{ij})$ を用いて、$Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j$ の形に表すことが出来る。この式は内積を使うと (Hx, x) と書ける。以下では H が対称行列、すなわち $a_{ij} = a_{ji}$ $(i, j = 1, 2, \cdots, n)$ であることを仮定する (こうすると写像としての 2 次形式 $Q(x)$ と対称行列 H が 1 対 1 に対応することになるから)。この場合 H を 2 次形式 $Q(x)$ の係数行列、$Q(x)$ を行列 H で定まる 2 次形式と呼ぶ。$Q(x)$ のことをしばしば $H[x]$ などの記号で表す。

さて H は実対称行列であるから、H のすべての固有値 $\lambda_1, \lambda_2, \cdots, \lambda_n$ は実数であって、ある直交行列 T が存在して

$$
{}^tTHT = \begin{pmatrix} \lambda_1 & & & \text{\Large 0} \\ & \lambda_2 & & \\ & & \ddots & \\ \text{\Large 0} & & & \lambda_n \end{pmatrix}
$$

と対角化できる。$x = Ty$ とおくと、

$$
Q(x) = (Hx, x) = (HTy, Ty) = ({}^tTHTy, y) = \lambda_1 {y_1}^2 + \lambda_2 {y_2}^2 \cdots + \lambda_n {y_n}^2 \tag{2.8}
$$

と変形出来る。このことから、$Q(x)$ の符号と、H の固有値の符号の間に密接な関係があることが分かる。

定義 2.6.2 2 次形式 $Q(x)$ について、

(i) $Q(x)$ が**正値** (**正定値**、**正定符号**, **positive definite**) $\overset{\text{def.}}{\Leftrightarrow}$ $\forall x \neq 0\ Q(x) > 0$.

(ii) $Q(x)$ が**負値** (**負定値**、**負定符号**, **negative definite**) $\overset{\text{def.}}{\Leftrightarrow}$ $\forall x \neq 0\ Q(x) < 0$.

(iii) $Q(x)$ が**不定符号** $\overset{\text{def.}}{\Leftrightarrow}$ $\exists x, \exists x'$ s.t. $Q(x) > 0$, $Q(x') < 0$.

定義 2.6.3 実対称行列 H について、

(i) H が**正値** (**正定値**、**正定符号**) $\overset{\text{def.}}{\Leftrightarrow}$ H の固有値がすべて正。

(ii) H が**負値** (**負定値**、**負定符号**) $\overset{\text{def.}}{\Leftrightarrow}$ H の固有値がすべて負。

(iii) H が**不定符号** $\overset{\text{def.}}{\Leftrightarrow}$ H の固有値の中に正のものと負のものがある。

H が正則な ($\Leftrightarrow$ 0 は H の固有値でない) 場合は、上の 3 つ (i), (ii), (iii) で、すべての場合がつくされることになる。

式 (2.8) から明らかなように

命題 2.6.2 実対称行列 H の定める 2 次形式 $Q(x)$ について、

(i) $Q(x)$ が正値 $\Leftrightarrow$ H が正値.

(ii) $Q(x)$ が負値 $\Leftrightarrow$ H が負値.

(iii) $Q(x)$ が不定符号 $\Leftrightarrow$ H が不定符号.

従って関数 $Q(x)$ の符号を調べるには、行列 H の固有値の符号を調べればいいことが分かる。固有値の計算は通常は結構大変[21]だが、固有値の符号を調べるだけなら簡単な判定法がある[22]。

命題 2.6.3 (対称行列の正値性を主座行列式で判定する方法) $H = (h_{ij})$ を n 次実対称行列で、D_r をその第 r 主座行列、すなわち H の最初の r 行 r 列を取って作った r 次の正方行列であるとする:

$$D_r = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1r} \\ h_{21} & h_{22} & \cdots & h_{2r} \\ \vdots & \vdots & \cdots & \vdots \\ h_{r1} & h_{r2} & \cdots & h_{rr} \end{pmatrix}.$$

このとき、以下の (1), (2) が成り立つ。

(1) H が正値 $\Leftrightarrow$ $\det D_r > 0$ $(r = 1, 2, \cdots, n)$.

(2) H が負値 $\Leftrightarrow$ $-H$ が正値 $\Leftrightarrow$ $(-1)^r \det D_r > 0$ $(r = 1, 2, \cdots, n)$.

この命題の証明は省略する (線形代数の教科書に載っている)。四則演算だけで固有値の符号が判定出来ることを納得してくれれば良い。

特に $n = 2$ の場合、すなわち 2 変数 x_1, x_2 の 2 次形式を調べてみよう。添字を書くのは面倒なので $x_1 = x$, $x_2 = y$, $\begin{pmatrix} x \\ y \end{pmatrix} = \vec{x}$ と書く。 $H = \begin{pmatrix} p & q \\ q & r \end{pmatrix}$ とすると

$$Q(\vec{x}) = (H\vec{x}, \vec{x}) = px^2 + qxy + qyx + ry^2 = px^2 + 2qxy + ry^2.$$

特に H が正則である ($\Leftrightarrow$ $\det H = pr - q^2 \neq 0$ $\Leftrightarrow$ H の固有値が両方とも 0 でない) 場合は、完全に分類出来て、

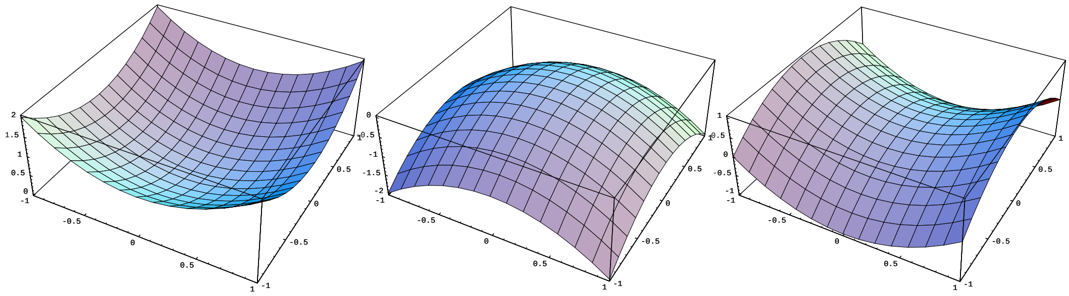
(1) H が正値であるには $\det D_1 = p > 0$, $\det D_2 = pr - q^2 > 0$ であることが必要十分条件である。 この場合 $\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ は関数 $\vec{x} \mapsto Q(\vec{x})$ の極小点になる (実は最小点)。正値な H の例として $H = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $H[\vec{x}] = x^2 + y^2$.

[21] n が大きいとき、n 次代数方程式を解くのは大変である。

[22] ここにあげる方法以外にも、**符号数** (正の固有値の個数と負の固有値の個数の組のこと) を計算する一般的な手段として **Lagrange の方法**と呼ばれる方法がある。これは各変数について順番に平方完成をして行くというもので、この方が強力ではあるが、ここでは名前をあげるにとどめておく。線形代数の教科書 (例えば齋藤 [2] p.157 など) に書かれていることが多い。

(2) H が負値であるには $(-1)^1 \det D_1 = -\det D_1 = -p > 0$, $(-1)^2 \det D_2 = \det D_2 = pr - q^2 > 0$ であることが必要十分である。この場合 $\vec{0}$ は関数 $\vec{x} \mapsto Q(\vec{x})$ の極大点になる (実は最大点)。負値な H の例として $H = (-1 \quad 00 \quad -1)$, $H[\vec{x}] = -x^2 - y^2$.

(3) H が不定符号であるには、H が正値でも負値でもないことが必要十分条件で、$pr - q^2 < 0$ と書ける[23]。この場合 $\vec{0}$ は関数 $\vec{x} \mapsto Q(\vec{x})$ の**鞍点**(あんてん) (**saddle point**, **峠点**(とうげてん)) である。不定符号な H の例として $H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, $H[\vec{x}] = x^2 - y^2$.



(a) $z = x^2 + y^2$. (b) $z = -x^2 - y^2$. (c) $z = x^2 - y^2$.

図 2.2: 代表的な 2 変数 2 次関数.

2.6.4 n 変数関数の極値の判定

Taylor の定理の 2 次剰余項 R_2 は Hesse 行列を使って書けるのがミソ。$f'(a) = 0$ となる点 a においては

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{1}{2}(H_{a+\theta h}h, h) \\ &= f(a) + \frac{1}{2}(H_{a+\theta h}h, h) \\ &\fallingdotseq f(a) + \frac{1}{2}(H_a h, h). \end{aligned}$$

以下では $Q(h) = (H_a h, h)$ とおく。

定理 2.6.2 (多変数関数の極値の Hesse 行列による判定) f は $\mathbf{R}^n$ の開集合 Ω 上定義された C^2-級の実数値関数で $f'(a) = 0$ を満たすものとする。H_a を f の a における Hesse 行列とするとき、

(i-a) H_a が正値 $\Longrightarrow$ f は a で狭義の極小.

(i-b) H_a が負値 $\Longrightarrow$ f は a で狭義の極大.

(ii) H_a が不定符号 $\Longrightarrow$ f は a で極値を取らない.

[23] あるいは $\det H$ は H のすべての固有値の積であることを思い出せれば、今の場合固有値は 2 個しかないので、正負の固有値がある ⇔ 二つの固有値の積が負 ⇔ $\det H < 0$、と考えてもよい。

証明　Ω は開集合であるから、$\exists\varepsilon > 0$ s.t. $B(a;\varepsilon) \subset \Omega$. Taylor の定理より、$\|h\| < \varepsilon$ なる任意の h に対して、$\exists\theta \in (0,1)$ s.t.

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{1}{2}(d^2 f)_{a+\theta h}(h) \\ &= f(a) + \frac{1}{2}(H_{a+\theta h}h, h). \end{aligned}$$

(i-a) H_a が正値ならば、主座小行列 $D_r(a)$ に対して

$$\det D_r(a) > 0 \quad (1 \le r \le n).$$

f は C^2-級と仮定してあるから、f の 2 階偏微分係数を成分とする行列 $D_r(x)$ の行列式は x の関数として連続である。ゆえに $\varepsilon > 0$ を十分小さく取ると、

$$\det D_r(a+\theta h) > 0 \quad (1 \le r \le n,\ \|h\| < \varepsilon,\ 0 < \theta < 1).$$

ゆえに $\|h\| < \varepsilon$, $0 < \theta < 1$ に対して $H_{a+\theta h}$ は正値である。したがって $0 < \|h\| < \varepsilon$ に対して

$$f(a+h) - f(a) = \frac{1}{2}(H_{a+\theta h}h, h) > 0.$$

これは a が f の狭義の極小点であることを意味する。

(i-b) 上と同様であるので省略。

(ii) (以下の証明を理解するには $n=2$, $a=0$, $f(x) = {x_1}^2 - {x_2}^2$, $u = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ の場合を頭に思い浮かべるとよい。) H_a が不定符号ならば、

$$(H_a u, u) > 0 > (H_a v, v) \tag{2.9}$$

を満たす $u \in \mathbf{R}^n$, $v \in \mathbf{R}^n$ が存在する。u, v の代わりに λu, λv $(\lambda \in \mathbf{R})$ で置き換えても (2.9) は成立するので、$\|u\| < \varepsilon$, $\|v\| < \varepsilon$ としてよい。このとき、

$$g(t) \overset{\text{def.}}{=} f(a+tu), \quad h(t) \overset{\text{def.}}{=} f(a+tv) \quad (|t| < 1)$$

とおくと、 g, h は C^2-級で、

$$g^{(k)}(0) = (d^k f)_a(u), \quad h^{(k)}(0) = (d^k f)_a(v) \qquad (k = 1, 2).$$

仮定より $(df)_a = 0$ ゆえ

$$g'(0) = f'(a)u = 0, \quad h'(0) = f'(a)v = 0.$$

また

$$g''(0) = (H_a u, u) > 0, \quad h''(0) = (H_a v, v) < 0.$$

ゆえに $\left\{ \begin{array}{l} t=0 \text{ は } g \text{ の狭義の極小点} \\ t=0 \text{ は } h \text{ の狭義の極大点} \end{array} \right\}$. すなわち a は f の峠点である。特に a は f の極値点ではない。 ■

H_a の符号の判定法については、前小節で述べたが、 Hesse 行列に即して、繰り返して述べておこう。 D_r を H_a の第 r 主座行列とする. すなわち

$$D_r \overset{\text{def.}}{=} \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_1 \partial x_r}(a) \\ \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_2 \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_2 \partial x_r}(a) \\ \vdots & \vdots & \cdots & \cdots \\ \dfrac{\partial^2 f}{\partial x_r \partial x_1}(a) & \dfrac{\partial^2 f}{\partial x_r \partial x_2}(a) & \cdots & \dfrac{\partial^2 f}{\partial x_r \partial x_r}(a) \end{pmatrix}.$$

とおくと、

(i-a) (正値性の判定)

$$\begin{aligned} & Q(h)\text{ が正値} \overset{\text{def.}}{\Leftrightarrow} \forall h \neq 0 \quad Q(h) > 0 \\ \Leftrightarrow \quad & H_a\text{が正値} \overset{\text{def.}}{\Leftrightarrow} H_a \text{ のすべての固有値が正} \Leftrightarrow \det D_r > 0 \quad (r = 1, 2, \cdots, n) \end{aligned}$$

(i-b) (負値性の判定)

$$\begin{aligned} & Q(h)\text{ が負値} \overset{\text{def.}}{\Leftrightarrow} \forall h \neq 0 \quad Q(h) < 0 \\ \Leftrightarrow \quad & H_a\text{が負値} \overset{\text{def.}}{\Leftrightarrow} H_a \text{ のすべての固有値が負} \Leftrightarrow \det(-D_r) = (-1)^r \det D_r > 0 \quad (r = 1, 2, \cdots, n) \end{aligned}$$

(ii) (不定符号性の判定)

$$\begin{aligned} & Q(h)\text{ が不定符号} \overset{\text{def.}}{\Leftrightarrow} \exists h, \exists h' \quad \text{s.t.} \quad Q(h) > 0,\ Q(h') < 0 \\ \Leftrightarrow \quad & H_a\text{が不定符号} \overset{\text{def.}}{\Leftrightarrow} H_a \text{ の固有値の中に正のものと負のものがある} \end{aligned}$$

2.6.5 例題

例題 2.6.1 $f(x, y) = x^3 + y^3 - 3xy$ の極値を求めよ。

解 まず導関数を計算しよう:

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix} = \begin{pmatrix} 3x^2 - y \\ 3y^2 - x \end{pmatrix},$$

$$f \text{ の Hesse 行列 } H = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

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

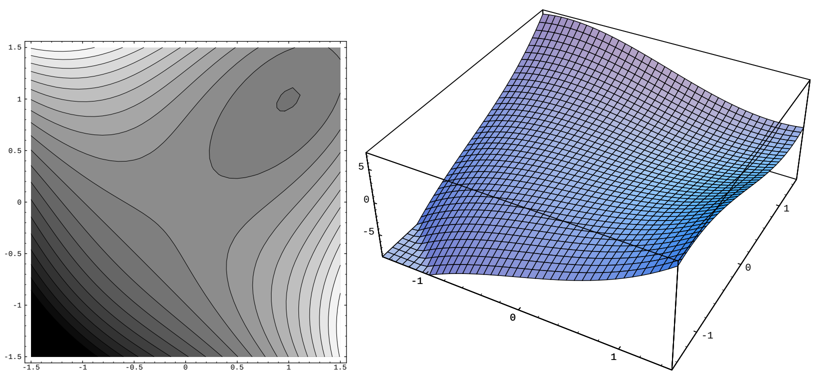
$$H = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}, \quad \therefore \det D_1 = 6 > 0, \quad \det D_2 = \det H = 6 \cdot 6 - (-3)(-3) = 27 > 0.$$

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 2.3 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 2.3: $f(x,y) = x^3 + y^3 - 3xy$.

注意 2.6.3 (教師のつぶやき) この手の極値問題は「話が分かりやすい」ので、期末試験でも定番の問題なのだが、微分法自体の理解よりも、(停留点を求めるための) 連立方程式を解くことに苦戦する (間違える、そこから先に進めなくなる) 人が後を絶たない。大抵は代数方程式なのだが、単純に未知数を消去していくと人間の手には終えなくなることが多い。しばしば、問題の対称性を生かした式変形を使うと簡単になるようである。この点は少し練習で慣れることが必要であろう。一方で、この手の工夫は「試験向きのテクニック」に過ぎないのかもしれないとも感じている。コンピューターを用いて連立代数方程式を解く研究が近年かなり進展したのだが、その成果が大学学部レベルに降りてくるにはまだ時間がかかるのだろう…

2.6.6 細かい話

ところで、定理 2.6.2 の逆は成立しないが、それに近いものはある。

命題 2.6.4 f は $\mathbf{R}^N$ の開集合 Ω 上定義された C^2- 級の実数値関数で $f'(a)=0$ を満たすものとする。H_a を f の a における Hesse 行列とするとき、

$$f \text{ は } a \text{ で極小} \Longrightarrow H_a \text{ が半正値},$$
$$f \text{ は } a \text{ で極大} \Longrightarrow H_a \text{ が半負値}$$

が成り立つ。

ただし、

$$\begin{aligned} Q(h) \text{ が半正値 (positive semi-definite}^{24}) &\overset{\text{def.}}{\Leftrightarrow} \forall h \neq 0 \quad Q(h) \geq 0 \\ \Leftrightarrow \quad H_a \text{が半正値} &\overset{\text{def.}}{\Leftrightarrow} H_a \text{ のすべての固有値が 0 以上} \\ &\Leftrightarrow \det D_r \geq 0 \quad (r=1,2,\cdots,n) \end{aligned}$$

$$\begin{aligned} Q(h) \text{ が半負値 (negative semi-definite}^{25}) &\overset{\text{def.}}{\Leftrightarrow} \forall h \neq 0 \quad Q(h) \leq 0 \\ \Leftrightarrow \quad H_a \text{が半負値} &\overset{\text{def.}}{\Leftrightarrow} H_a \text{ のすべての固有値が 0 以下} \\ &\Leftrightarrow (-1)^r \det D_r \geq 0 \quad (r=1,2,\cdots,n) \end{aligned}$$

である。付加条件として「H_a は正則」とすると、半正値=正値、半負値=負値であるから、次のように簡単な命題が得られる。

系 2.6.1 Hesse 行列が正則な場合は、極大・極小の完全な判定が出来る。すなわち

$$f \text{ は } a \text{ で極小} \Leftrightarrow H_a \text{ が正値},$$
$$f \text{ は } a \text{ で極大} \Leftrightarrow H_a \text{ が負値},$$
$$f \text{ は } a \text{ で極値を取らない} \Leftrightarrow H_a \text{ が不定符号}$$

が成り立つ。

一方、Hesse 行列が特異である場合は、大ざっぱに言って、

(i) (正値でない) 半正値

(ii) (負値でない) 半負値

(iii) 不定符号

の 3 つに分類出来る ($H_a=O$ など (i) かつ (ii) ということもあるので、「分類」は厳密に言えば不適当な表現であるが大目に見て下さい)。このうち不定符号である場合は極値ではないと断定出来るが、半正値、半負値だけでは極大、極小は判定できない。例えば $f(x,y)=x^4+y^4$ とするとき、$f'(0)=0$, $H_a=O$ で、f は 0 で極小だが、$f(x,y)=-x^4-y^4$ とするとき、$f'(0)=0$, $H_a=O$ で、f は 0 で極大、さらに $f(x,y)=x^4-y^4$ とするとき、$f'(0)=0$, $H_a=O$ で、f は 0 で極大でも極小でもない。これらの例からも示唆されるように、半正値、半負値の場合は、より高次の導関数の項まで調べないと分からない。

2.7 陰関数定理と逆関数定理

2.7.1 逆関数に関する序

要点は少なくて以下のように簡単にまとめられる (特に最後の一つだけが本質的に新しい)。

- 写像 $f\colon A \to B$ が全単射であるとき、逆写像 (逆関数) $f^{-1}\colon B \to A$ が存在する。
- 逆関数の存在、非存在を考えるときは、定義域、終域が大事である。両方とも適当に狭めることで逆関数が存在するようになることが多い。
- 線形写像 $f\colon \mathbf{R}^n \ni x \mapsto Ax \in \mathbf{R}^n$ については、f の逆写像 f^{-1} が存在するための必要十分条件は $\det A \neq 0$.
- 逆関数の導関数については、簡単な公式がある: 対応する点における元の関数の微分係数の逆行列である。つまり $(f^{-1})'(y) = (f'(x))^{-1}$ (ただし $y = f(x)$).
- より一般の C^1-級関数 $f\colon \mathbf{R}^n \supset \Omega \to \mathbf{R}^n$ についても、$\det f'(a) \neq 0$ ならば、a の近傍で f の逆関数が存在する。— これがこの節で学ぶ重要な「**逆関数定理**」である。

1 変数関数の逆関数については既に実例を多く知っている。まずその復習から始めよう。

指数関数 $\exp$ **と対数関数** $\log$　指数関数

$$\tilde{f}\colon \mathbf{R} \ni x \longmapsto \exp x = e^x \in \mathbf{R}$$

は単射 (1 対 1 の写像) である。すなわち

$$x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2).$$

(このことは f が狭義の単調増加関数であることから明らか。) f の値域 $f(\mathbf{R})$ は正の実数全体の集合 $(0,\infty)$ であるから、終域を $\mathbf{R}$ から $(0,\infty)$ に入れ替えた関数

$$f\colon \mathbf{R} \ni x \longmapsto \exp x = e^x \in (0,\infty)$$

は単射であるのみならず全射 (上への写像) である。すなわち

$$\forall y \in (0,\infty) \quad \exists x \in \mathbf{R} \text{ s.t.} \quad f(x) = y$$

が成り立つ。よって f は逆関数 g を持つ。つまり

$$g\colon (0,\infty) \longrightarrow \mathbf{R}$$

で

$$f \circ g = id_{(0,\infty)}, \quad g \circ f = id_{\mathbf{R}}$$

を満たす[26]。$g(y)$ のことを $\log y$ と記すのであった。逆関数の微分法より

$$\frac{d}{dy}\log y = \frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}} = \frac{1}{e^x} = \frac{1}{y}.$$

[26]集合 A に対して、id_A は $id_A(x) = x$ $(\forall x \in A)$ で定められる写像 $id_A\colon A \to A$ で、A 上の**恒等写像** (**identity mapping**) と呼ばれる。

ルート $\sqrt{\ }$ 関数

$$\tilde{f}\colon \mathbf{R} \ni x \longmapsto x^2 \in \mathbf{R}$$

は単射でも全射でもない。しかし定義域を $[0,\infty)$ に制限して、終域も $[0,\infty)$ と替えた

$$f\colon [0,\infty) \ni x \longmapsto x^2 \in [0,\infty)$$

は全単射である。よって f は逆関数 g を持つ。つまり

$$g\colon [0,\infty) \longrightarrow [0,\infty)$$

で

$$f \circ g = id_{[0,\infty)}, \quad g \circ f = id_{[0,\infty)}$$

を満たす。$g(y)$ のことを $\sqrt{y}$ と記すのであった。

$$\frac{d}{dy}\sqrt{y} = \frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}} = \frac{1}{2x} = \frac{1}{2\sqrt{y}} \quad (\text{ただし } y \neq 0).$$

逆三角関数

$$\begin{aligned}
f_1 &: \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \ni x \longmapsto \sin x \in [-1,1], \\
f_2 &: [0,\pi] \ni x \longmapsto \cos x \in [-1,1], \\
f_3 &: \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \ni x \longmapsto \tan x \in \mathbf{R}
\end{aligned}$$

はいずれも全単射であるから、逆関数が存在する。それぞれ g_1, g_2, g_3 とすると

$$\begin{aligned}
g_1 &: [-1,1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \\
g_2 &: [-1,1] \longrightarrow [0,\pi], \\
g_3 &: \mathbf{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)
\end{aligned}$$

となる。$g_1(y)$. $g_2(y)$, $g_3(y)$ はそれぞれ $\operatorname{Arcsin} y$, $\operatorname{Arccos} y$, $\operatorname{Arctan} y$ と書かれる。

$$\begin{aligned}
\frac{d}{dy}\operatorname{Arcsin} y &= \frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}} \quad (\text{ただし } y \neq 1), \\
\frac{d}{dy}\operatorname{Arccos} y &= \frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}} = \frac{1}{-\sin x} = -\frac{1}{\sqrt{1-\cos^2 x}} = -\frac{1}{\sqrt{1-y^2}} \quad (\text{ただし } y \neq 1), \\
\frac{d}{dy}\operatorname{Arctan} y &= \frac{dx}{dy} = \frac{1}{\dfrac{dy}{dx}} = \frac{1}{\frac{1}{\cos^2 x}} = \frac{1}{1+\tan^2 x} = \frac{1}{1+y^2}.
\end{aligned}$$

一般の 1 変数実数値関数　一般の $\tilde{f}\colon \mathbf{R} \supset \Omega \to \mathbf{R}$, $a \in \Omega$ について $\tilde{f}'(a) \neq 0$ ならば a の十分近くでは逆関数が存在する。実際 $\tilde{f}'(a) > 0$ または $\tilde{f}'(a) < 0$ であり、$\tilde{f}'$ は連続であるから $\exists \varepsilon > 0$ s.t. $\tilde{f}' > 0$ in $(a-\varepsilon, a+\varepsilon)$ または $\tilde{f}' < 0$ in $(a-\varepsilon, a+\varepsilon)$. 例えば $\tilde{f}' > 0$ としよう。すると $\tilde{f}$ は狭義の単調増加関数だから、$\tilde{f}$ による 開区間 $(a-\varepsilon, a+\varepsilon)$ の像は $(\tilde{f}(a-\varepsilon), \tilde{f}(a+\varepsilon))$ で、制限写像

$$f\colon (a-\varepsilon, a+\varepsilon) \ni x \longmapsto \tilde{f}(x) \in (\tilde{f}(a-\varepsilon), \tilde{f}(a+\varepsilon))$$

は全単射である。ゆえに逆関数

$$g\colon (\tilde{f}(a-\varepsilon), \tilde{f}(a+\varepsilon)) \to (a-\varepsilon, a+\varepsilon)$$

が存在する。

一般の n 変数 n 次元ベクトル値関数　一般の $\tilde{f}\colon \mathbf{R}^n \supset \Omega \to \mathbf{R}^n$, $a \in \Omega$ についても同様である。すなわち、

$$\tilde{f}'(a) \text{ の逆行列 } \tilde{f}'(a)^{-1} \text{ が存在 } (\Leftrightarrow \det \tilde{f}'(a) \neq 0)$$

ならば、a を含む十分小さな開集合 U が存在して、

$$f\colon U \ni x \mapsto \tilde{f}(x) \in \tilde{f}(U)$$

は全単射になり、その逆関数が存在する。これが後で詳述する**逆関数定理**である。

2.7.2　陰関数についての序

まずは 1 次元の例から。$F\colon \mathbf{R}^2 \to \mathbf{R}$ を

$$F(x,y) = x^2 + y^2 - 1$$

で定義する。$(a,b) \in \mathbf{R}^2$ を

$$F(a,b) = 0, \quad (a,b) \neq (\pm 1, 0)$$

を満たす点とする。F の零点集合 $N_F \overset{\text{def.}}{=} \{(x,y); F(x,y) = 0\}$ は、もちろん原点を中心とする半径 1 の円周である。このとき a を含む開区間 U と b を含む開区間 V とで

$$\forall x \in U \quad \exists! y \in V \text{ s.t. } F(x,y) = 0$$

を満たすものが存在する[27]。$x \in U$ に対応する y を $\varphi(x)$ と書くと、$\varphi\colon U \to V$ であって、

$$\forall x \in U \quad F(x, \varphi(x)) = 0$$

がなりたつ。こういう φ を方程式 $F = 0$ から定まる**陰関数** (**implicit function**) と呼ぶ。(陰 (implicit) $\leftrightarrow$ 陽(よう) (explicit), y が $y = \varphi(x)$ と陽(あらわ)になっていなくて、隠されていた、という意味で陰関数).

いくつか注意しておこう。

[27] $\exists!$ は一意的に存在することを表す記号である。

(1) 最初から $y = \varphi(x)$ と普通の関数 (陽関数) となっているものは、取り扱いに便利なことが多いが、中には

$$F(x, y) = 0$$

の形にしておいた方が望ましく、好まれることもある。そういう場合に、いざというときは $y = \varphi(x)$ と y について解けることが保証されると、とても助かる。そのための定理が陰関数定理である。

(2) この例ではもちろん

$$\begin{cases} b > 0 \implies \varphi(x) = \varphi_1(x) \overset{\text{def.}}{=} \sqrt{1 - x^2} \\ b < 0 \implies \varphi(x) = \varphi_2(x) \overset{\text{def.}}{=} -\sqrt{1 - x^2} \end{cases}$$

である。$F(x, \varphi(x)) = 0$ を満たす φ としては他に

$$\varphi_3(x) \overset{\text{def.}}{=} \begin{cases} \sqrt{1 - x^2} & (x \in \mathbf{Q} \cap (-1, 1)) \\ -\sqrt{1 - x^2} & (x \in (-1, 1) \setminus \mathbf{Q}) \end{cases}$$

なんてのもあるが、これは微分可能でも連続でもないし、面白くもない (役に立たない)。

(3) $a = \pm 1$ のときは、こういう φ はない。これはよーっく納得するように。陰関数定理では、こういう点を除外するための条件がつくわけだが、それは

$$\det \frac{\partial F}{\partial y}(a, b) \neq 0 \qquad (\text{解こうとしている変数 } y \text{ に関する微分が } 0 \text{ でない})$$

というものである。

(4) $F(x, y) = 0$ が簡単な式変形では解けないような場合も多い。例えば

$$F(x, y) \overset{\text{def.}}{=} y + \sin y + x^2 - ex + 1 = 0.$$

陰関数定理はこういう場合にも適用できる。それは陰関数 φ についての、**抽象的な存在定理**であって、φ の具体的な表現などは与えていないが、その代わり広い範囲の F に適用が出来る。

(5) ところで φ は C^1- 級である (実は陰関数定理が成り立つときは φ は常に C^1- 級である)。すると

$$F(x, \varphi(x)) = 0$$

が成り立つことから、

$$\frac{\partial F}{\partial x}(x, \varphi(x)) + \frac{\partial F}{\partial y}(x, \varphi(x))\varphi'(x) = 0,$$

よって

$$\varphi'(x) = -\frac{\dfrac{\partial F}{\partial x}(x, \varphi(x))}{\dfrac{\partial F}{\partial y}(x, \varphi(x))}.$$

つまり陰関数定理は φ の表現については何も言及しないけれども、その導関数については合成関数の微分法から導かれる簡単な公式がある。上の例では

$$x^2 + y^2 - 1 = 0$$

の両辺を x で微分して

$$2x + 2y\frac{dy}{dx} = 0,$$

これから

$$\frac{dy}{dx} = -\frac{x}{y},$$

というもので、これは高校数学で習った人もいるであろう。

多次元の場合の例　上の例では x, y が 1 次元の量であったが、多次元の場合はどうであろうか。簡単な (線形の) 例で考えてみよう。例えば

$$A \in M(n, m; \mathbf{R}), \quad B \in M(n; \mathbf{R}), \quad \vec{c} \in \mathbf{R}^n$$

とするとき

$$\vec{F}(\vec{x}, \vec{y}) \overset{\text{def.}}{=} A\vec{x} + B\vec{y} + \vec{c}$$

で定められる

$$\vec{F}\colon \mathbf{R}^m \times \mathbf{R}^n \ni (\vec{x}, \vec{y}) \longmapsto \vec{F}(\vec{x}, \vec{y}) \in \mathbf{R}^n$$

を取り上げる。

$$\vec{F}(\vec{x}, \vec{y}) = \vec{0}$$

は $\vec{y}$ について解けるか？この答は簡単で、$\det B \neq 0$ ならば、

$$\vec{F}(\vec{x}, \vec{y}) \equiv A\vec{x} + B\vec{y} + \vec{c} = \vec{0}$$

を移項して出来る

$$B\vec{y} = -(A\vec{x} + \vec{c})$$

の両辺から B^{-1} をかけて

$$\vec{y} = -B^{-1}(A\vec{x} + \vec{c})$$

となる。$\det B = 0$ であるときは case by case (A, B について詳しい情報がないと分からない).

2.7.3 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 $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ 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 2.7.1 (陰関数定理 (implicit function theorem)) Ω を $\mathbf{R}^m \times \mathbf{R}^n$ の開集合、$F\colon \Omega \to \mathbf{R}^n$ を C^1- 級の関数とする。さらに点 $(a, b) \in \Omega$ において $F(a, b) = 0$, $\dfrac{\partial F}{\partial y}(a, b)$ は逆行列を持つとする。

$\Longrightarrow$ $\exists U$: a を含む開集合 ($\subset \mathbf{R}^m$), $\exists V$: b を含む開集合 ($\subset \mathbf{R}^n$), $\exists \varphi\colon U \to V$: C^1- 級の関数 s.t.

(1) $\varphi(a) = b$.

(2) 任意の $(x, y) \in U \times V$ について

$$F(x, y) = 0 \Longleftrightarrow y = \varphi(x).$$

(3) $\varphi'(x) = -\left(\dfrac{\partial F}{\partial y}(x, \varphi(x))\right)^{-1} \dfrac{\partial F}{\partial x}(x, \varphi(x))$.

注意 2.7.1 (陰関数定理の条件 (2) の言い換え) (2) 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 (2)' で置き換えることも出来る。
(2)' $U \times V$ において、F の零点集合は φ のグラフに一致する:

$$(U \times V) \cap N_F = \operatorname{graph} \varphi,$$

ただし

$$N_F \overset{\text{def.}}{=} \{(x, y) \in \Omega; F(x, y) = 0\}, \quad \operatorname{graph} \varphi \overset{\text{def.}}{=} \{(x, \varphi(x)); x \in U\}.$$

定理 2.7.2 (逆関数定理 (inverse function theorem)) Ω を $\mathbf{R}^n$ の開集合、$f\colon \Omega \to \mathbf{R}^n$ を C^1- 級の関数とする。さらに $f'(a)$ は逆行列を持つとする。
$\Longrightarrow$ $\exists U$: a を含む開集合 ($\subset \Omega \subset \mathbf{R}^n$), $\exists V$: $b = f(a)$ を含む開集合 ($\subset \mathbf{R}^n$) s.t. $f|_U\colon U \to V$ は全単射で、逆写像 f^{-1} は V で C^1- 級, ${f^{-1}}'(y) = (f'(x))^{-1}$, $x = f^{-1}(y)$.

注意 2.7.2 $f|_U$ は f の U への制限写像を表わす。

陰関数定理と逆関数定理は見かけは異なるが、どちらか一方を認めると、もう一方は比較的簡単に証明できるという意味で、ほとんど同等の定理であると言える。

2.7.4 簡単な例

例 2.7.1 $F(x, y) = x^3 + y^3 - 2xy$, $P = (1, 1)$ とする。$F(x, y) = 0$ の点 P の近くにおける陰関数 $y = \varphi(x)$ の存在を示し、その点における微分係数を求めよ。

解答 $F(x, y) = x^3 + y^3 - 2xy$ とおくと $F\colon \mathbf{R}^2 \to \mathbf{R}$ は C^1-級で、$F_y(x, y) = 3y^2 - 2x$. ゆえに $F_y(1, 1) = 3 \cdot 1^1 - 2 \cdot 1 = 1 \neq 0$ であるから $(1, 1)$ の近傍で $F(x, y) = 0$ は $y = \varphi(x)$ と解ける。$F_x(x, y) = 3x^2 - 2y$ より $F_x(1, 1) = 3 \cdot 1^1 - 2 \cdot 1 = 1$ であるから

$$\varphi'(1) = -\frac{F_x(1, 1)}{F_y(1, 1)} = -\frac{1}{1} = -1. \quad \blacksquare$$

例 2.7.2 連立方程式 $x + y + z + w = 0$, $e^x + e^{2y} + e^z + e^w = 4$ は、 0 の近傍で x, y について解けることを証明せよ。

解答

$$\begin{aligned} X &= \begin{pmatrix} z \\ w \end{pmatrix}, \quad Y = \begin{pmatrix} x \\ y \end{pmatrix}, \\ F_1(X, Y) &= x + y + z + w, \quad F_2(X, Y) = e^x + e^{2y} + e^z + e^w - 4, \\ F(X, Y) &= \begin{pmatrix} F_1(X, Y) \\ F_2(X, Y) \end{pmatrix} \end{aligned}$$

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、

$$\frac{\partial F}{\partial Y}(X,Y)=\begin{pmatrix} \frac{\partial F_1}{\partial x} & \frac{\partial F_1}{\partial y} \\ \frac{\partial F_2}{\partial x} & \frac{\partial F_2}{\partial y}\end{pmatrix}=\begin{pmatrix} 1 & 1 \\ e^x & 2e^y\end{pmatrix}.$$

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。■

2.7.5 陰関数、逆関数の高階数導関数

命題 2.7.1 (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。

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

$$\varphi'(x)=-\left(\frac{\partial F}{\partial x}(x,\varphi(x))\right)^{-1}\frac{\partial F}{\partial y}(x,\varphi(x)). \tag{2.10}$$

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、(2.10) 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だから、証明は略する。

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 (2.10) 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

$$\varphi''(x)=-\frac{F_y^2F_{xx}}{F_y^3}=-\frac{F_{xx}}{F_y}.$$

とかなりシンプルになる。

例題 2.7.1 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。

解　まず与式を微分して

$$y^2+2xyy'-2xy-x^2y'=0. \tag{2.11}$$

これから

$$\begin{aligned} y'=0 \quad &\Leftrightarrow \quad y(y-2x)=0 \\ &\Leftrightarrow \quad y=2x \quad (y=0 \text{ は元の式を満たさない}) \\ &\Leftrightarrow \quad x=1, \quad y=2. \end{aligned}$$

ところで *(2.11)* 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

$$3y''-4=0, \quad \therefore y''=4/3>0.$$

よって極小値である。■

2.7.6　陰関数定理の応用について

陰関数定理はきちんと述べるのも大変な定理であるが、一体何の役に立つのか (苦労に見合うだけの実入りがあるか)？微積分を他の諸科学に直接的に応用するという立場からは、導関数の公式さえマスターしておけばいいだろう (こう書くと、みもふたもないが)。しかし数学自身の事情から言えば、陰関数定理は色々な重要な応用を持つ[28]。ここでは、多様体[29], 条件付き極値問題, 分岐理論[30]を紹介する。

[28]最近は数学以外の科学に数学を応用する際に、かなり深いところが使われる。そういう場合は陰関数定理程度のことは常識であろう。

[29]3 年次の幾何学 (=幾何学入門) のメイン・テーマである。

[30]非線形数学の重要なテーマである。

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、次小節「関数のレベル・セット」で説明する。)

条件つき極値問題 以下で節を改めて詳しく説明する。

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。

2.7.7 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 2.6.1 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

$$L_c \overset{\text{def.}}{=} \{(x,y)\in\Omega; f(x,y)=c\}$$

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。

今 $a\in\Omega$ を任意に取って、$c\overset{\text{def.}}{=}f(a)$ とおく。既に

> a から $\nabla f(a)$ の方向に移動すると標高が高くなり、$-\nabla f(a)$ 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」 陰関数定理によれば、$\nabla f(a)\neq\begin{pmatrix}0\\0\end{pmatrix}$ ならば、a の近傍 W で $f(x,y)=c$ は一つの変数について解くことができる。

(1) $f_y(a)\neq 0$ の場合。y について解ける。すなわち開集合 U と $\varphi\colon U\to\mathbf{R}$ が存在して、

$$N_{f-c}\cap W=\operatorname{graph}\varphi\equiv\{(x,\varphi(x)); x\in U\}.$$

(2) $f_x(a)\neq 0$ の場合。x について解ける。すなわち開集合 V と $\psi\colon V\to\mathbf{R}$ が存在して、

$$N_{f-c}\cap W=\operatorname{graph}\psi\equiv\{(\psi(y),y); y\in V\}.$$

$L_c=N_{f-c}$ であることに注意すると、レベル・セット L_c は a の近傍で 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) におけるレベル・セット L_c は「点」である。ちなみに峠におけるレベル・セットは、峠点で交わる 2 曲線である[31]。

2.7.8　陰関数定理と逆関数定理の証明

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 $f(x,y) = (x, F(x,y))$ で定義すると、これは C^1- 級で、$f(a,b) = (a, F(a,b)) = (a,0)$,

$$f'(a,b) = \begin{pmatrix} I & 0 \\ \dfrac{\partial F}{\partial x}(a,b) & \dfrac{\partial F}{\partial y}(a,b) \end{pmatrix}.$$

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 $\tilde{\Omega} \subset \Omega$ と、点 $f(a,b) = (a,0)$ を含む開集合 W が存在して、 $f|_{\tilde{\Omega}}\colon \tilde{\Omega} \to W$ は C^1- 級の逆関数 g を持つ。

$\forall (x,y) \in \tilde{\Omega}$ に対して $g(x,y) = (x, \psi(x,y))$ と書ける。($\because g(x,y) = (\eta(x,y), \psi(x,y))$ とすると、$(x,y) = f(\eta(x,y), \psi(x,y)) = (\eta(x,y), F(\eta(x,y), \psi(x,y)))$. $\therefore x = \eta(x,y)$)

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 $F(x, \varphi(x)) = F(x, \psi(x,0)) = 0$.

逆に $F(x, y_1) = 0$ となったとすると、

$$f(x, y_1) = (x, F(x, y_1)) = (x, 0) = (x, F(x, \varphi(x))) = f(x, \varphi(x)).$$

$f|_{\tilde{\Omega}}$ は一対一ゆえ $y_1 = \varphi(x)$. ■

[31] 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。

陰関数定理を認めた上での逆関数定理の証明　今回の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、

$$\det \frac{\partial F}{\partial x}(a,b)=\det f'(a)\neq 0.$$

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。

逆関数の定理の証明

1° $A \overset{\text{def.}}{=} f'(a)$, $\tilde{f} \overset{\text{def.}}{=} A^{-1}\circ f$ 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.

$$\forall x\in U\setminus\{a\}\quad f(x)\neq f(a). \tag{2.12}$$

$$\forall x\in U\quad \det f'(x)\neq 0. \tag{2.13}$$

$$\forall x\in U\quad \|f'(x)-f'(a)\|<\frac{1}{2}. \tag{2.14}$$

主張 A の証明　f' の連続性により、U を十分小さく取れば (2.14) は成り立つ。同様に $f'(a)=I$ に注意すれば、U を十分小さく取れば (2.13) も成り立つ。(2.12) については、まず

$$\begin{aligned} f(x) &= f(a)+f'(a)(x-a)+o(\|x-a\|)\quad (x\to a)\\ &= f(a)+(x-a)+o(\|x-a\|)\quad (x\to a)\end{aligned}$$

から

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

ところが $f(x)=f(a)$ とすると

$$\frac{\|f(x)-f(a)-(x-a)\|}{\|x-a\|}=\frac{\|x-a\|}{\|x-a\|}=1.$$

よって a に十分近い x $(\neq a)$ では $f(x)=f(a)$ とはなりえない。■

3° **主張 B:**

$$\forall x_1\in U\quad \forall x_2\in U\quad \|x_1-x_2\|\le 2\|f(x_1)-f(x_2)\|. \tag{2.15}$$

主張 B の証明　$g(x)\overset{\text{def.}}{=} f(x)-x$ 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
$$\|x_1-x_2\|-\|f(x_1)-f(x_2)\| \le \frac{1}{2}\|x_1-x_2\|.$$
■

4° $W \overset{\text{def.}}{=} \partial U$, $d=\mathrm{dist}(f(a),f(B))$ とおくと $d>0$. 実際

- (2.12) より $f(a)\notin f(B)$.
- B はコンパクトだから連続写像による像 $f(B)$ もコンパクトで、特に $f(B)$ は閉集合。
- 「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。さて $W \overset{\text{def.}}{=} B(f(a);d/2)$ とおくと

(2.16) $$y\in W,\quad x\in B \Longrightarrow \|y-f(a)\|<\|y-f(x)\|.$$

実際まず W の定義から
$$\|y-f(a)\|<\frac{d}{2},$$
一方 $x\in B$ より
$$\|f(x)-f(a)\|\ge \mathrm{dist}(f(a),f(B))\equiv d$$
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:**
$$\forall y\in W\quad \exists!\overline{x}\in U\setminus B \text{ s.t. } f(\overline{x})=y$$

主張の C 証明　関数 $g\colon\overline{U}\to\mathbf{R}$ を
$$g(x)\overset{\text{def.}}{=}\|y-f(x)\|^2\equiv(y-f(x),y-f(x))$$
で定義する。これはコンパクト集合上の連続関数ゆえ最小値 $g(\overline{x})$ を取る。ところで (2.16) より
$$x\in B\Longrightarrow g(a)<g(x).$$
ゆえに $\overline{x}\notin B$ i.e. $\overline{x}\in U$. ゆえに g は内点 $\overline{x}$ で最小値を取ることになり $g'(\overline{x})=0$ i.e. $f'(\overline{x})(y-f(\overline{x}))=0$. (2.13) より $f'(\overline{x})$ は正則ゆえ $y-f(\overline{x})=0$. すなわち $f(\overline{x})=y$. $\overline{x}$ の一意性は (2.15) から分かる。

6° ここで

$$V \overset{\text{def.}}{=} (U \setminus B) \cap f^{-1}(W)$$

とおくと V は a の開近傍である。(実際 W は開集合であり、連続写像による逆像 $f^{-1}(W)$ は開集合。それと開集合 $U \setminus B$ の交わりであるから W は開集合。$a \in U$, $a \neq B$ は明らかで、$f(a) \in W$ より $a \in f^{-1}(W)$ であるから $a \in V$.) 前項から

$$f|_V : V \longrightarrow W$$

は逆関数 $f^{-1}: W \to V$ を持つ。

7° f^{-1} は連続である。実際 (4) より $y_1 \in W$, $y_2 \in W$ とするとき

$$\|f^{-1}(y_1) - f^{-1}(y_2)\| \leq 2\|y_1 - y_2\|$$

であるから。

8° **主張 D:** $\forall x \in V$ f^{-1} は $y \overset{\text{def.}}{=} f(x)$ で微分可能で

$$(f^{-1})'(y) = (f'(x))^{-1}.$$

主張 D の証明 $A \overset{\text{def.}}{=} f'(x)$ とおく。微分可能性の定義から

(2.17) $$f(X) - f(x) = A(X - x) + \varepsilon(x, X)$$

とおくと

$$\lim_{X \to x} \frac{\|\varepsilon(x, X)\|}{\|X - x\|} = 0.$$

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 (2.17) の両辺に A^{-1} をかけ、 y, Y で書き直すと

$$A^{-1}(Y - y) = f^{-1}(Y) - f^{-1}(y) + A^{-1}\varepsilon(f^{-1}(y), f^{-1}(Y)).$$

ゆえに

$$f^{-1}(Y) - f^{-1}(y) = A^{-1}(Y - y) - A^{-1}\varepsilon(f^{-1}(y), f^{-1}(Y)).$$

そこで次のことを示せばよい。

$$\lim_{Y \to y} \frac{\|A^{-1}\varepsilon(f^{-1}(y), f^{-1}(Y))\|}{\|Y - y\|} = 0.$$

これを示すには

$$\lim_{Y \to y} \frac{\|\varepsilon(f^{-1}(y), f^{-1}(Y))\|}{\|Y - y\|} = 0$$

を示せばよい。

$$\frac{\|\varepsilon(f^{-1}(y), f^{-1}(Y))\|}{\|Y - y\|} = \frac{\|\varepsilon(f^{-1}(y), f^{-1}(Y))\|}{\|f^{-1}(Y) - f^{-1}(y)\|} \cdot \frac{\|f^{-1}(y) - f^{-1}(Y))\|}{\|Y - y\|}.$$

f^{-1} の連続性より $Y \to y$ のとき $f^{-1}(Y) \to f^{-1}(y)$. よって右辺第 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 の関数としてみて連続である。ゆえに f^{-1} は C^1- 級。■

2.8　条件付き極値問題 (Lagrange の未定乗数法)

2.8.1　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は定義域が基本的には開集合であった。すると

$$a \text{ が } f \text{ の極値点} \Longrightarrow a \text{ は } f \text{ の停留点}, \quad \text{i.e. } \nabla f(a) \equiv \begin{pmatrix} \dfrac{\partial f}{\partial x}(a) \\ \dfrac{\partial f}{\partial y}(a) \end{pmatrix} = 0.$$

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

$$g(x,y) = 0 \tag{2.18}$$

の下で求めることを考える。すなわち g の零点集合

$$N_g \overset{\text{def.}}{=} \{(x,y); g(x,y) = 0\}$$

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。

例題 2.8.1 条件 $g(x,y) \overset{\text{def.}}{=} x^2 + y^2 - 1 = 0$ 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？

例 2.8.1 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

$$h(x) \overset{\text{def.}}{=} f(x,x^2) = \sqrt{x^2 + (x^2-1)^2}$$

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 $y = \varphi(x) = 0$ が具体的に求まるとは限らないので、上の二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 2.8.1 (条件つき極値問題の極値点の探し方 (Lagrange の未定乗数法) 2 変数版)
Ω を $\mathbf{R}^2$ の開集合、f, 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 (i.e. $\exists \varepsilon > 0$ s.t. $f(a) = \max\{f(x); x \in B(a;\varepsilon) \cap N_g\}$ または $f(a) = \min\{f(x); x \in B(a;\varepsilon) \cap N_g\}$) $\Longrightarrow \exists \lambda \in \mathbf{R}$ s.t.

$$\nabla f(a) = \lambda \nabla g(a)$$

i.e.

$$\begin{aligned} \frac{\partial f}{\partial x}(a) &= \lambda \frac{\partial g}{\partial x}(a), \\ \frac{\partial f}{\partial y}(a) &= \lambda \frac{\partial g}{\partial y}(a). \end{aligned}$$

注意 2.8.1 (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

$$\begin{cases} g(\alpha,\beta) = 0 \\ \dfrac{\partial f}{\partial x}(\alpha,\beta) = \lambda \dfrac{\partial g}{\partial x}(\alpha,\beta) \\ \dfrac{\partial f}{\partial y}(\alpha,\beta) = \lambda \dfrac{\partial g}{\partial y}(\alpha,\beta) \end{cases}$$

の解として求まることが多い (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

$$F(x,y,\lambda) \overset{\text{def.}}{=} f(x,y) - \lambda g(x,y)$$

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

$$h(x) \overset{\text{def.}}{=} f(x, \varphi(x))$$

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。■

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 $L_c = \{x \in \Omega; f(x) = c\}$ を考える。条件

$$\exists \lambda \in \mathbf{R} \quad \text{s.t.} \quad \nabla f(a) = \lambda \nabla g(a)$$

は、$N_g = \{(x, y); g(x) = 0\}$ 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)。

点 x が N_g に沿って、a から動くとき

$$(x - a, \nabla g(a)) \fallingdotseq 0.$$

すると
$$0 \fallingdotseq f(x) - f(a) \fallingdotseq (\nabla f(a), x-a).$$
よって $\nabla f(a)$, $\nabla g(a)$ ともに $x-a$ に直交する。これから
$$\nabla g(a) \parallel \nabla f(a).$$
ゆえに f の等高線と N_g は接している。 ■

2.8.2 n 変数の場合

一般の n 変数関数の場合も定理だけは掲げておこう。

定理 2.8.2 (条件つき極値問題の極値点の探し方 (Lagrange の未定乗数法) n 変数版)
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.
$$\nabla f(a) = \sum_{j=1}^{d} \lambda_j \nabla g_j(a), \quad \text{また (もちろん)} \quad \boldsymbol{g}(a) = 0. \tag{2.19}$$

注意 2.8.2 やはり
$$F(x,\boldsymbol{\lambda}) \overset{\text{def.}}{=} f(x) - \sum_{j=1}^{d} \lambda_j g_j(x)$$
とおくと方程式は (2.19) は次と同値になる:
$$\begin{cases} \dfrac{\partial F}{\partial x_i}(a,\boldsymbol{\lambda}) = 0 & (i = 1, \cdots, n) \\ \dfrac{\partial F}{\partial \lambda_j}(a,\boldsymbol{\lambda}) = 0 & (j = 1, \cdots, d). \end{cases}$$

2.8.3 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 2.8.2 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 (直感的にすぐ) 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、$D\overset{\text{def.}}{=}\{(x,y)\in\mathbf{R}^2;x^2+y^2\le R^2\}$ とおく。L を

$$L=(L\cap D)\cup(L\cap D^c)$$

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 $g(x,y)\overset{\text{def.}}{=}ax+by+c$ 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[32]。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。■

例題 2.8.3 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

$$g(x,y)\overset{\text{def.}}{=}\frac{x^2}{a^2}+\frac{y^2}{b^2}-1$$

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

$$\nabla g(x,y)\neq 0$$

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

$$\begin{aligned}0&=1-\lambda\frac{2x}{a^2},\\0&=1-\lambda\frac{2y}{b^2},\\0&=\frac{x^2}{a^2}+\frac{y^2}{b^2}-1\end{aligned}$$

[32]「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 $-\sqrt{a^2+b^2}$. ■

付 録A　期末試験の採点から — 教師の憂鬱な時間

「今年もあれほど言ったのに、また同じことをやっている…」採点の時間はため息をつきながら過ぎていく。

A.1　定義を書こう

例題　(1) $\mathbf{R}^n$ の開集合の定義を記せ。(2) $\mathbf{R}^n$ の開集合の例をあげ、定義に従ってそれが開集合であることを証明せよ。(3) $\mathbf{R}^n$ の部分集合で、開集合でないものの例をあげ、それが開集合でないことを証明せよ。

この (1) について。

満点の取れない解答例　「$\forall a \in A, \exists \varepsilon > 0 \text{ s.t. } B(a;\varepsilon) \subset A.$」

上記の一行「解答」はこれまでの期末試験の解答でかなり多いものであるが、まずい点が二つある。そのうちで大事なこと (減点の対象となる) は、「何が開集合であるか」書いていないことである。

修正した解答例 1　「$A \subset \mathbf{R}^n$ が $\mathbf{R}^n$ の開集合であるとは、条件

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

が成り立つことである。」

修正した解答例 2　「$A \subset \mathbf{R}^n$ が $\mathbf{R}^n$ の開集合 $\Leftrightarrow \forall a \in A, \exists \varepsilon > 0 \text{ s.t. } B(a;\varepsilon) \subset A.$」

次に些細な点を指摘すると、それは $B(a;\varepsilon)$ という記号である。これは解答者は a を中心とする半径 ε の $\mathbf{R}^n$ の開球、つまり

$$B(a;\varepsilon) = \{x \in \mathbf{R}^n; \|x-a\| < \varepsilon\}, \quad (\text{ただし } \|\cdot\| \text{ は } \mathbf{R}^n \text{ のノルムを表す})$$

の意味で使っていると思われるが、実はそれほど一般的な記号ではない (例えば、同じ意味で $U(a;\varepsilon)$ のような記号を使うテキストも多い)。解析概論の講義では、最初に「この講義ではこの記号をこの意味で使う」と断って、以後一貫してこの記号を使い続けたので、期末試験においては「了解事項」として使っていいかもしれないが、他の試験では一言断り書きをする必要がある。その意味で、解析概論の試験の答案に断りなしに $U(a;\varepsilon)$ という記号を使うのは本来は減点の対象とすべきなのだろう (減点したことはないが … われながら甘いなあ)。

仮想の問答　「自分が学んだ記号が一般的なものかどうか、どうやったら分かるのでしょう？」「一口に言えば『常識』で、それは色々な機会を通して、例えば複数の講義を受けたり、複数の本を読んだりして、自分で学んでいくものです。」「それは初めて勉強する場合大変なのでは？」「個人的には、大学の学部程度の講義では使う記号がどれだけ一般的であるか、教師が説明しておくのが親切だとは思いますが…まあ、そんなに神経質になる必要はないでしょう。この記号はどういう意味かと尋ねられたら、そのとき答えればいいのですから。」

A.2　連続性、偏微分、全微分

「次式で定義される関数 $f\colon \mathbf{R}^2 \to \mathbf{R}$ について、(1) 原点で連続であるか、(2) 原点で偏微分可能であるか、(3) 原点で全微分可能であるか、答えよ。」という内容の出題文の問題を出すことが多い。

A.2.1　連続性のチェック

概説　一般的には、連続であることがすぐ分かる関数というものも結構多い。

(1) $f(x,y) = 1 + x - 2y + 3xy + 4y^2$. 多項式関数だから連続。

(2) $f(x,y) = \dfrac{1 + x - 2y + 3xy + 4y^2}{2x + 3y - 1}$. 分母が考えている点 (原点) で 0 にならない有理関数だから連続。

(3) $f(x,y) = \log(1 + 2xy)$. $h(z) = \log z$, $g(x,y) = 1 + 2xy$ という関数の合成関数 $h \circ g$ に他ならない。g は多項式だから連続。h は $g(0,0) = 1$ $(\neq 0)$ で連続だから、f は $(0,0)$ で連続。

残念ながら (?)、そういう関数をこのタイプの問題に出したことはあまりない。多くは原点とそれ以外の点で場合わけされるような定義をしてあり、原点以外の点では $(x,y) \to (0,0)$ での極限が 0 になるような分母を持つような問題である。こういう問題の場合は、連続であるならば

$$|f(x,y) - f(0,0)|$$

を不等式で評価して行って、$(x,y) \to (0,0)$ のとき 0 に収束することを示すのが普通である。連続でない場合は、極限が存在しないか、存在しても $f(0,0)$ に等しくないことを示すわけである。後者の場合を問題に出したことはない (その場合、原点での定義さえ修正すれば連続になる。だからもともとの定義が不自然ということで変な問題と思うから)。前者の場合は例えば $y = kx$ という直線に沿って原点に近づけた場合の極限を調べて解決する場合があるが、残念ながら万能というわけではない。

連続な場合の例　まず連続な関数の例として

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

を取り上げる。任意の $(x,y)\neq(0,0)$ に対して

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

ここで

$$\frac{y^2}{x^2+y^2}\le\frac{x^2+y^2}{x^2+y^2}=1$$

であるから

$$|f(x,y)-f(0,0)|\le|x|\cdot\sqrt{1}=|x|.$$

$(x,y)\to(0,0)$ のとき $x\to 0$ であるから、右辺 $\to 0$. ゆえに

$$|f(x,y)-f(0,0)|\to 0\quad((x,y)\to(0,0)).$$

ゆえに

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

であり、f は $(0,0)$ で連続であることがわかる。■

不連続な場合の例 1　上の例と良く似ているが不連続な関数の例として

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

これは直線 $y=kx$ 上で考えると、原点以外では

$$f(x,kx)=\frac{x\cdot kx}{x^2+(kx)^2}=\frac{k}{1+k^2}.$$

これは k によって値が異なる。ゆえに極限 $\displaystyle\lim_{(x,y)\to(0,0)}f(x,y)$ は存在しない。ゆえに f は $(0,0)$ で連続ではない。■

不連続な場合の例 2　最後に少し難しい例をあげておく (間違えて連続であると解説してある本があった)。

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

結論から先に言うと、実はこの f は $(0,0)$ で連続ではない。そこで $y=kx$ にそって $(0,0)$ に近づけたときの極限を調べる方法をやってみよう。$k\neq 1$ とすると $x\neq 0$ であるとき $x+kx\neq 0$ となるので、

$$f(x,kx)=\frac{x^2+(kx)^2}{x+kx}=x\frac{1+k^2}{1+k}.$$

これは $x\to 0$ とすると 0 という共通の極限に収束する。これを見て、

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

と結論した人が多かった。しかしこれは間違いである。k が任意であっても、それを一度固定してしまってから、$y = kx$ にそって $(0, 0)$ に近づけるというのは「特殊な」近づけ方に過ぎないのである。

それでは解答。(x, y) が半径 r の円周上にある場合を考える。$x = r\cos\theta$, $y = r\sin\theta$ $(\theta \in [0, 2\pi])$ と書けるので、$x + y \neq 0$ のとき

$$\begin{aligned} |f(x,y) - f(0,0)| &= |f(r\cos\theta, r\sin\theta) - 0| \\ &= \left|\frac{(r\cos\theta)^2 + (r\sin\theta)^2}{r\cos\theta + r\sin\theta}\right| \\ &= \left|\frac{r}{\cos\theta + \sin\theta}\right| = \frac{r}{\left|\sqrt{2}\sin(\theta + \pi/4)\right|}. \end{aligned}$$

右辺の分母は θ を $3\pi/4$ あるいは $7\pi/4$ に近づけるといくらでも 0 に近くなる。そこで r がどんなに 0 に近かったとしても、θ を $3\pi/4$ の十分近くに取れば[1]、$|f(x, y) - f(0, 0)|$ の値はいくらでも大きくなる — 不連続性を証明するには 1 よりも大きいことを言えば十分。念のため論理式で書いておくと

$$\forall r > 0, \quad \exists\theta \in [0, 2\pi] \setminus \left\{\frac{3\pi}{4}, \frac{7\pi}{4}\right\} \quad \text{s.t.} \quad |f(r\cos\theta, r\sin\theta)| \geq 1.$$

ゆえに

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

は 0 でない (本当は極限すら存在しない)。ゆえに f は $(0, 0)$ で連続ではない。■

仮想の問答　「上の解説は、問題の解き方に関しては、『色々な場合がある』と言っているだけで、どうやればいいのか方針を説明できていないと思うんですが。」「うーん (『だけ』と言うことはないだろうに)、分母と分子が 0 に近づく速さを比較するわけなのだけど、あまり一般的な方針はないですね。比較的多くの場合に使える目安としては、分母と分子の『次数』を考えるというのがあります。

$$\frac{xy}{x^2 + y^2} = \frac{2\,次}{2\,次}$$

なんてのは、次数が同じだから、分子が勝って 0 に収束することにはならない。

$$\frac{xy}{\sqrt{x^2 + y^2}} = \frac{2\,次}{1\,次}$$

は分子の次数が高いので、勝ちそうだ (0 に収束しそうだ) という見通しをつけるわけです。ただこれで行くと、最後の例も

$$\frac{x^2 + y^2}{x + y} = \frac{2\,次}{1\,次}$$

となって、0 に収束しように見えるのが難点ですね。まあ、この分母は原点以外でも 0 になりうるというのが少し特殊なんだけど。」「完璧でなくてもいいけれど、何か便利そうな手は他

[1] 具体的に θ をどう取ればいいか書けば完璧であるが、そこまでする必要はないだろう。そんなに難しいことではないが。

にないですか？」「これは試験テクニック的な感じがして、こればかりを覚えてもらっても困るのですが、このように分母、分子ともに x と y のバランスの取れた式である場合は、最後の例でもやったように、極座標を使って式を表現するというのがあります。例えば

$$\frac{xy}{x^2+y^2} = \frac{r\cos\theta\cdot r\sin\theta}{r^2\cos^2\theta + r^2\sin^2\theta} = \cos\theta\sin\theta = \frac{1}{2}\sin 2\theta.$$

これは $r\to 0$ としても 0 に収束しそうでないのはすぐ分かる。一方、

$$\frac{xy}{\sqrt{x^2+y^2}} = \frac{r\cos\theta\cdot r\sin\theta}{\sqrt{r^2\cos^2\theta + r^2\sin^2\theta}} = r\cos\theta\sin\theta = \frac{r}{2}\sin 2\theta$$

だから

$$\left|\frac{xy}{\sqrt{x^2+y^2}}\right| \le \frac{r}{2}$$

で、これは $r\to 0$ のとき 0 に近づく。最後の例は、繰返しになるけど

$$f(x,y) - f(0,0) = \frac{r}{\sqrt{2}\sin(\theta+\pi/4)}$$

となって、ちょっと見た目には r が残っているので 0 に収束しそうな気もするけれど、0 になりうる分母が残っているので、そうは行かないと。二番目の例では

$$|\sin 2\theta| \le 1$$

と θ に関係した因子が有界になるのに、三番目の例では

$$\left|\frac{1}{\sin(\theta+\pi/4)}\right|$$

は有界でないというのが違いと言えます。」「うーん、何か複雑ですね。」「説明している方も簡単な説明になっているとは思っていません (ごめん)。」

A.2.2 偏微分可能性のチェック

例題

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

は原点で偏微分可能かどうか調べよ。

よくある 間違い は

$$f_x(x,y) = \frac{(x^2+y^2)\cdot y - 2x\cdot xy}{(x^2+y^2)^2} = \frac{x^2y+y^3-2x^2y}{x^2+y^2} = \frac{y^3-x^2y}{(x^2+y^2)^2},$$

$$f_y(x,y) = \frac{(x^2+y^2)\cdot x - 2y\cdot xy}{(x^2+y^2)^2} = \frac{x^3+x^2y-2xy^2}{x^2+y^2} = \frac{x^3-xy^2}{(x^2+y^2)^2}$$

としてから $(x,y)\to(0,0)$ の極限を取ろうとして、それが存在しないから、偏微分可能でない、と答えるという解答。これは根本的な勘違いである。そもそもこれでは $f(0,0)=0$ という情報を使う機会がないので、それだけでも変だと感じないとおかしい。

尋ねられているのは、極限

$$f_x(0,0)=\lim_{h\to0}\frac{f(0+h,0)-f(0,0)}{h},\quad f_y(0,0)=\lim_{h\to0}\frac{f(0,0+h)-f(0,0)}{h}$$

が存在するかということである。この式で $f(h,0)$, $f(0,h)$ については、$(h,0)\neq(0,0)$, $(0,h)\neq(0,0)$ なので、

$$\frac{xy}{x^2+y^2}$$

という式に代入すればいいが、$f(0,0)$ については場合わけの片割れである 0 という情報を使う必要がある。

色々文句を書いたが、正解は次のようにごくあっさりしたものになる。

$$f_x(0,0)=\lim_{h\to0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to0}\frac{1}{h}\left(\frac{h\cdot0}{h^2+0^2}-0\right)=\lim_{h\to0}\frac{0}{h}=\lim_{h\to0}0=0.$$

$$f_y(0,0)=\lim_{h\to0}\frac{f(0,0+h)-f(0,0)}{h}=\lim_{h\to0}\frac{1}{h}\left(\frac{0\cdot h}{0^2+h^2}-0\right)=\lim_{h\to0}\frac{0}{h}=\lim_{h\to0}0=0.$$

ゆえに f は $(0,0)$ で変数 x, y のいずれについても偏微分可能である。■

仮想問答　「これは覚えてしまえば簡単ですね。」「そうだと思うんだけど案外と出来ないんです。」「え、一体どこを間違えるんですか？」「間違えている解答は、ほとんど途中経過が書いていないので、採点側としては想像するしかないのだけれど、二つほど理由が考えられます。一つは、なまじっか 0 が多いので、『正解』を見ても、それがどうしてそう計算するのか理解できず、例題と違った問題になると対応できないというもの (要するに分かっていない)。もう一つは省略することによる計算間違いですね。そういえば

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

について $f_x(0,0)$, $f_y(0,0)$ を求めよ、という問題を出したとき、あっさりと

$$f_x(0,0)=\lim_{h\to0}\frac{f(0+h,0)-f(0,0)}{h}=0,\quad f_y(0,0)=\lim_{h\to0}\frac{f(0,0+h)-f(0,0)}{h}=0$$

と書いて間違った人がたくさん出て、とても驚いたことがあります。こういう問題の答は 0 になると思い込んでいるのでしょうか。」「えーと、正解がすぐには分からないのですが、教えてください。」「『すぐ分からない』のが当たり前です。こういうのは計算しないと私でも分かりません。自分で計算してみてください。」「ええと、

$$f_x(0,0)=\lim_{h\to0}\frac{f(0+h,0)-f(0,0)}{h}=\lim_{h\to0}\frac{1}{h}\left(\frac{h^2+0^2}{h+0}-0\right)=\lim_{h\to0}\frac{h^2}{h^2}=\lim_{h\to0}1=1.$$

ああ、本当に 0 ではないんですね。$f_y(0,0)$ もそうなんでしょうか。」「そうそう。…そういえば話は飛ぶけれど、この手の問題は、高校数学の範囲でもあったはずです。例えば

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

が $x=0$ で微分可能であるかどうか調べよ、とかね。」「そうだったかも…」「そういえば、大学院入試の面接で

$$f(x)=\begin{cases} |x|^{4/3}\sin\left(\dfrac{1}{x}\right) & (x\neq 0)\\ 0 & (x=0)\end{cases}$$

の微分可能性を尋ねていた先生もいました。受験生は結構あたふたしていましたね[2]。」「はあ。」

A.2.3 全微分可能性をチェックする

すぐに全微分可能であることが分かる関数も多い。

(1) 多項式

(2) 微分可能な 1 変数関数 (log, exp, sin, cos, $\sqrt{x}$ 等、ただし定義域や特異点に注意)

(3) 微分可能な関数の合成

(4) 分母、分子とも微分可能な関数で、考えている点で分母が 0 にならない。

不連続関数は全微分不可能　「全微分可能ならば連続」という定理の対偶「連続でなければ全微分可能でない」はピンと来てほしい。例えばすでに何回も出てきた

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

は、$(0,0)$ で不連続なことが分かるから、$(0,0)$ で全微分可能ではない。■

連続だが全微分可能でない例

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

は連続であることは既に見た。全微分可能性はどうか？定義に戻って考えると、

$$\lim_{(x,y)\to(0,0)}\frac{f(x,y)-f(0,0)-ax-by}{\sqrt{x^2+y^2}}=0$$

[2] ここまで来ると脱線でしょうか。でもこのような問題がきちんと解けることは大事だと考える先生がいるということで、なるほどなあ、と感じました。

となる定数 a, b が存在するならば全微分可能、そうでないならば全微分可能でないとなる (自分で納得するまで考えること — ここらへんをうろ覚えで沈没する人が多い)。このままでは少し難しいが、「全微分可能な場合、$a = f_x(0,0)$, $b = f_y(0,0)$ である」という定理があった[3]。そこでまず $f_x(0,0)$, $f_y(0,0)$ を求めてみよう。結果だけ書くと

$$f_x(0,0) = 0, \quad f_y(0,0) = 0.$$

さらに $f(0,0) = 0$ であるから、調べるべき式は

$$\begin{aligned}\lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - f_x(0,0)\cdot x - f_y(0,0)\cdot y}{\sqrt{x^2+y^2}} &= \lim_{(x,y)\to(0,0)} \frac{f(x,y)}{\sqrt{x^2+y^2}} \\ &= \lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}.\end{aligned}$$

これは既に出てきていて (すぐ上！)、結論は「極限なし」であった。ゆえに f は $(0,0)$ で全微分可能ではない。■

全微分可能な例　最後は少し凝った例を。

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

の全微分可能性はどうか？まず f は $(0,0)$ で連続で、$f_x(0,0) = f_y(0,0) = 0$ である (各自確かめよ)。

すると、残る問題は

$$\begin{aligned}\frac{f(x,y) - f(0,0) - f_x(0,0)x - f_y(0,0)y}{\sqrt{x^2+y^2}} &= \frac{f(x,y) - 1}{\sqrt{x^2+y^2}} \\ &= \frac{1}{\sqrt{x^2+y^2}}\left(\frac{\sin x^2 + \sin y^2}{x^2+y^2} - 1\right) \\ &= \left(\frac{\sin x^2 - x^2}{(x^2+y^2)^{3/2}}\right) + \left(\frac{\sin y^2 - y^2}{(x^2+y^2)^{3/2}}\right)\end{aligned}$$

の極限が 0 になるかどうか、である。

$\sin x$ の Taylor 展開

$$\sin x = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!} x^{2n-1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots$$

から、$\sin x^2$ の展開

$$\sin x^2 = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{(2n-1)!} x^{2(2n-1)} = x^2 - \frac{x^6}{3!} + \frac{x^{10}}{5!} - \cdots$$

[3]もう少し正確に書いておくと「関数が全微分可能ならば、各変数に関して偏微分可能で、微分係数は偏微分係数を並べたヤコビ行列に等しい。」

が得られる。これから

$$x^2 - \sin x^2 = \sum_{n=2}^{\infty} \frac{(-1)^n}{(2n-1)!} x^{2(2n-1)} = \frac{x^6}{3!} - \frac{x^{10}}{5!} + \cdots$$

であるが、右辺が交代級数であることに注意すると

$$0 \le x^2 - \sin x^2 \le \frac{x^6}{3!}.$$

これから

$$\begin{aligned}
\left| \frac{f(x,y) - f(0,0) - f_x(0,0)\cdot x - f_y(0,0)\cdot y}{\sqrt{x^2+y^2}} \right| &\le \left| \frac{\sin x^2 - x^2}{(x^2+y^2)^{3/2}} \right| + \left| \frac{\sin y^2 - y^2}{(x^2+y^2)^{3/2}} \right| \\
&\le \frac{|\sin x^2 - x^2|}{(x^2+0)^{3/2}} + \frac{|\sin y^2 - y^2|}{(0+y^2)^{3/2}} \\
&= \frac{x^2 - \sin x^2}{|x|^3} + \frac{y^2 - \sin y^2}{|y|^3} \\
&\le \frac{x^6}{3!}\cdot\frac{1}{|x|^3} + \frac{y^6}{3!}\cdot\frac{1}{|y|^3} \\
&= \frac{|x|^3}{6} + \frac{|y|^3}{6} \to 0 \quad ((x,y)\to(0,0)).
\end{aligned}$$

ゆえに f は $(0,0)$ で全微分可能である。■

A.3 極座標で合成関数の微分法を学ぶ

A.3.1 イントロ

$$x = r\cos\theta, \quad y = r\sin\theta.$$

これで $\varphi\colon [0,\infty)\times[0,2\pi) \ni (r,\theta) \mapsto (x,y) \in \mathbf{R}^2$ という写像ができる。一階の偏導関数は簡単に求まり、

(A.1) $$x_r = \cos\theta, \quad x_\theta = -r\sin\theta, \quad y_r = \sin\theta, \quad y_\theta = r\cos\theta.$$

$f\colon \mathbf{R}^2 \to \mathbf{R}$ という関数があったとき、

$$g(r,\theta) = f(x,y) = f(r\cos\theta, r\sin\theta)$$

で $g\colon [0,\infty)\times[0,2\pi) \to \mathbf{R}$ という関数を作ることができる。要するに $g = f\circ\varphi$ である。f が C^1-級であれば g もそうなるのだが、g の偏導関数は f の偏導関数を使ってどのように表されるだろうか？これは合成関数の微分法 (chain rule) の簡単な練習問題で、(A.1) を用いて

(A.2) $$g_r = f_x x_r + f_y y_r = f_x\cos\theta + f_y\sin\theta, \quad g_\theta = f_x x_\theta + f_y y_\theta = -f_x r\sin\theta + f_y r\cos\theta$$

となる。

それでは逆に f_x, f_y を g_r, g_θ で表すにはどうしたら良いか？そのためには (A.2) を f_x, f_y に関する連立方程式とみなして解いてもよいが、ここでは $f = g \circ \varphi^{-1}$ と考えて計算してみよう。やはり合成関数の微分法から

$$\text{(A.3)} \qquad f_x = g_r r_x + g_\theta \theta_x, \quad f_y = g_r r_y + g_\theta \theta_y.$$

この式をさらに具体的に書くには r_x, θ_x, r_y, θ_y を求める必要がある。逆関数のヤコビ行列は、もとの関数のヤコビ行列の逆行列であるから、

$$\begin{aligned}\begin{pmatrix} r_x & r_y \\ \theta_x & \theta_y \end{pmatrix} &= (\varphi^{-1})'(x,y) = (\varphi'(r,\theta))^{-1} = \begin{pmatrix} x_r & x_\theta \\ y_r & y_\theta \end{pmatrix}^{-1} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}^{-1} \\ &= \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}.\end{aligned}$$

ゆえに

$$r_x = \cos\theta, \quad r_y = \sin\theta, \quad \theta_x = \frac{-\sin\theta}{r}, \quad \theta_y = \frac{\cos\theta}{r}.$$

これを (A.3) に代入して

$$f_x = g_r\cos\theta - \frac{g_\theta\sin\theta}{r}, \quad f_y = g_r\sin\theta + \frac{g_\theta\cos\theta}{r}.$$

よくある間違い　さて、それで以上の計算を試験に出したときの出来であるが、r_x, r_y, θ_x, θ_y の計算あたりからつまずいてしまう。間違い方には大きく分けて二通りあって、まず

$$r_x = \frac{\partial r}{\partial x} = \frac{1}{\dfrac{\partial x}{\partial r}} = \frac{1}{\cos\theta} \quad \text{（この式は間違い！）}$$

のような「一変数の場合の逆関数の微分の公式の “機械的な” 適用」をしたのがある (もちろん間違いである)。もう一つは $x = r\cos\theta$ から

$$r = \frac{x}{\cos\theta}$$

という式を導き、これから

$$r_x = \frac{\partial}{\partial x}\left(\frac{x}{\cos\theta}\right) = \frac{1}{\cos\theta}\frac{\partial}{\partial x}x = \frac{1}{\cos\theta} \quad \text{（この式は間違い！）}$$

とするものである。これはどこがおかしいかと言うと、$\partial/\partial x$ を計算するときは、x 以外のすべての変数、この場合は y を固定しておいて変化率を考えるわけで、θ を定数と考えてはいけない、ということである。後を引き続き正しく計算すると

$$\begin{aligned} r_x &= \frac{\partial}{\partial x}\left(\frac{x}{\cos\theta}\right) = \frac{1}{\cos\theta}\cdot\frac{\partial}{\partial x}x + \frac{\partial}{\partial x}\left(\frac{1}{\cos\theta}\right)\cdot x = \frac{1}{\cos\theta} + \frac{\sin\theta}{\cos^2\theta}\theta_x x \\ &= \frac{1}{\cos\theta} + \frac{\sin\theta}{\cos^2\theta}\cdot\frac{-\sin\theta}{r}\cdot r\cos\theta = \frac{1}{\cos\theta} - \frac{\sin^2\theta}{\cos\theta} \\ &= \frac{1-\sin^2\theta}{\cos\theta} = \frac{\cos^2\theta}{\cos\theta} = \cos\theta. \end{aligned}$$

(説明のためにこういう泥臭い計算をしてあるだけで、こんなやり方を勧めているわけではない。念のため。)

A.3.2 二階導関数

記号は上と同じものを使う。

$$f_{xx}+f_{yy}=g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta} \tag{A.4}$$

を証明せよ、という問題をよく出題する。以下の方法 1 と方法 3 をマスターしてほしい。

方法 1　右辺を計算していって左辺に等しいことを示す、という方針で計算してみよう。まず $g_r=f_xx_r+f_yy_r$ より

$$\begin{aligned}
g_r &= f_x\cos\theta+f_y\sin\theta,\\
g_{rr} &= \frac{\partial}{\partial r}g_r=\frac{\partial}{\partial r}(f_xx_r+f_yy_r)=\frac{\partial}{\partial r}f_x\cdot x_r+f_x\frac{\partial}{\partial r}x_r+\frac{\partial}{\partial r}f_y\cdot y_r+f_y\frac{\partial}{\partial r}y_r\\
&= (f_{xx}x_r+f_{xy}y_r)x_r+f_xx_{rr}+(f_{yx}x_r+f_{yy}y_r)y_r+f_yy_{rr}\\
&= f_{xx}(x_r)^2+(f_{xy}+f_{yx})x_ry_r+f_{yy}(y_r)^2+f_xx_{rr}+f_yy_{rr}.
\end{aligned}$$

$x_r=\cos\theta$, $y_r=\sin\theta$ より $x_{rr}=y_{rr}=0$ が導かれること、また f が C^2-級であれば $f_{xy}=f_{yx}$ であることに注意すると、

$$g_{rr}=f_{xx}\cos^2\theta+2f_{xy}\cos\theta\sin\theta+f_{yy}\sin^2\theta.$$

次に $g_\theta=f_xx_\theta+f_yy_\theta$ より

$$\begin{aligned}
g_{\theta\theta} &= \frac{\partial}{\partial\theta}g_\theta=\frac{\partial}{\partial\theta}(f_xx_\theta+f_yy_\theta)=\frac{\partial}{\partial\theta}f_x\cdot x_\theta+f_x\frac{\partial}{\partial\theta}x_\theta+\frac{\partial}{\partial\theta}f_y\cdot y_\theta+f_y\frac{\partial}{\partial\theta}y_\theta\\
&= (f_{xx}x_\theta+f_{xy}y_\theta)x_\theta+f_xx_{\theta\theta}+(f_{yx}x_\theta+f_{yy}y_\theta)y_\theta+f_yy_{\theta\theta}\\
&= f_{xx}(x_\theta)^2+(f_{xy}+f_{yx})x_\theta y_\theta+f_{yy}(y_\theta)^2+f_xx_{\theta\theta}+f_yy_{\theta\theta}.
\end{aligned}$$

$x_\theta=-r\sin\theta$, $y_\theta=r\cos\theta$ より

$$x_{\theta\theta}=-r\cos\theta,\quad y_{\theta\theta}=-r\sin\theta$$

が導かれること、また f が C^2-級であれば $f_{xy}=f_{yx}$ であることに注意すると、

$$\begin{aligned}
g_{\theta\theta} &= f_{xx}(-r\sin\theta)^2+2f_{xy}(-r\sin\theta\cdot r\cos\theta)+f_{yy}(r\cos\theta)^2+f_x(-r\cos\theta)+f_y(-r\sin\theta)\\
&= f_{xx}r^2\sin^2\theta-2f_{xy}r^2\sin\theta\cos\theta+f_{yy}r^2\cos^2\theta-(f_xr\cos\theta+f_yr\sin\theta)
\end{aligned}$$

ゆえに

$$\begin{aligned}
g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta} &= f_{xx}\cos^2\theta+2f_{xy}\cos\theta\sin\theta+f_{yy}\sin^2\theta+\frac{1}{r}(f_x\cos\theta+f_y\sin\theta)\\
&\quad+\frac{1}{r^2}\left[f_{xx}r^2\sin^2\theta-2f_{xy}r^2\sin\theta\cos\theta+f_{yy}r^2\cos^2\theta-(f_xr\cos\theta+f_yr\sin\theta)\right]\\
&= f_{xx}(\cos^2\theta+\sin^2\theta)+f_{yy}(\sin^2\theta+\cos^2\theta)=f_{xx}+f_{yy}.
\end{aligned}$$

方法 2　(A.4) の右辺を計算する。方法 1 と同様に計算すると、

$$
\begin{aligned}
f_{xx} &= g_{rr}(r_x)^2+(g_{r\theta}+g_{\theta r})r_x\theta_x+g_{\theta\theta}(\theta_x)^2+g_r r_{xx}+g_\theta\theta_{xx},\\
f_{yy} &= g_{rr}(r_y)^2+(g_{r\theta}+g_{\theta r})r_y\theta_y+g_{\theta\theta}(\theta_y)^2+g_r r_{yy}+g_\theta\theta_{yy}
\end{aligned}
$$

となる。r_{xx}, θ_{xx}, r_{yy}, θ_{yy} は何か？[4]

$$
r_x=\cos\theta,\quad r_y=\sin\theta,\quad \theta_x=\frac{-\sin\theta}{r},\quad \theta_y=\frac{\cos\theta}{r}
$$

から一目で計算というわけには行かない。また合成関数の微分法を使って

$$
\begin{aligned}
r_{xx} &= \frac{\partial}{\partial x}r_x=\frac{\partial}{\partial x}\cos\theta=\frac{\partial}{\partial r}\cos\theta\cdot r_x+\frac{\partial}{\partial\theta}\cos\theta\cdot\theta_x\\
&= 0+(-\sin\theta)\frac{-\sin\theta}{r}=\frac{\sin^2\theta}{r},\\
r_{yy} &= \frac{\partial}{\partial y}r_y=\frac{\partial}{\partial y}\sin\theta=\frac{\partial}{\partial r}\sin\theta\cdot r_y+\frac{\partial}{\partial\theta}\sin\theta\cdot\theta_y\\
&= 0+(\cos\theta)\frac{\cos\theta}{r}=\frac{\cos^2\theta}{r},\\
\theta_{xx} &= \frac{\partial}{\partial x}\theta_x=\frac{\partial}{\partial x}\frac{-\sin\theta}{r}=\frac{\partial}{\partial r}\frac{-\sin\theta}{r}\cdot r_x+\frac{\partial}{\partial\theta}\frac{-\sin\theta}{r}\cdot\theta_x\\
&= \frac{\sin\theta}{r^2}\cos\theta+\frac{-\cos\theta}{r}\cdot\frac{-\sin\theta}{r}=\frac{2\sin\theta\cos\theta}{r^2},\\
\theta_{yy} &= \frac{\partial}{\partial y}\theta_y=\frac{\partial}{\partial y}\frac{\cos\theta}{r}=\frac{\partial}{\partial r}\frac{\cos\theta}{r}\cdot r_y+\frac{\partial}{\partial\theta}\frac{\cos\theta}{r}\cdot\theta_y\\
&= \frac{-\cos\theta}{r^2}\sin\theta+\frac{-\sin\theta}{r}\cdot\frac{\cos\theta}{r}=\frac{-2\sin\theta\cos\theta}{r^2}.
\end{aligned}
$$

これから

$$
(r_x)^2+(r_y)^2=1,\quad r_x\theta_x+r_y\theta_y=0,\quad (\theta_x)^2+(\theta_y)^2=\frac{1}{r^2},\quad r_{xx}+r_{yy}=\frac{1}{r},\quad \theta_{xx}+\theta_{yy}=0
$$

であるから、

$$
\begin{aligned}
f_{xx}+f_{yy} &= g_{rr}[(r_x)^2+(r_y)^2]+(g_{r\theta}+g_{\theta r})(r_x\theta_x+r_y\theta_y)+g_{\theta\theta}[(\theta_x)^2+(\theta_y)^2]\\
&\quad +g_r(r_{xx}+r_{yy})+g_\theta(\theta_{xx}+\theta_{yy})\\
&= g_{rr}\cdot 1+(g_{r\theta}+g_{\theta r})0+g_{\theta\theta}\frac{1}{r^2}+g_r\frac{1}{r}+g_\theta\cdot 0\\
&= g_{rr}+\frac{1}{r}g_r+\frac{1}{r^2}g_{\theta\theta}.
\end{aligned}
$$

(これは大変、ということで少しだけ工夫したのが次の方法である。)

[4]特に根拠なく、これらを 0 として計算する人が試験で後を絶たないのだけれど、一体なぜだろう？x_{rr}, y_{rr} が 0 になったのを機械的に真似している？「分からなくても何か書いておけ」という受験指導を受けてきた結果なのかとも思うが、こちらが理解できない毎年よく現われる誤答がいくつかある。デタラメでも 0 点で、どんなに悪くてもマイナスの点をつけないから (そうかね？)、何か書いておいた方が得である、という考え方なのだろうけど、何か大事なものを確実に損なっていると思う。この病気にかかっている自覚のある人は自分で治療するよう心掛けた方がよい (直そうと思っても、きっと急には直らない)。

方法 3　やはり (A.4) の右辺を変形していく。以下説明する方法は本質的には方法 2 と同じであるが、こちらの方が間違いにくいと思われる。

既に見た

$$f_x = g_r \cos\theta - \frac{g_\theta \sin\theta}{r}, \quad f_y = g_r \sin\theta + \frac{g_\theta \cos\theta}{r}$$

という公式から、関数の x, y に関する偏微分を r, θ に関する偏微分で表現する公式

$$\frac{\partial}{\partial x} = \cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}, \quad \frac{\partial}{\partial y} = \sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}$$

を得る。これを使って、

$$\begin{aligned}
f_{xx} &= \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)^2 g = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\cos\theta g_r - \frac{\sin\theta}{r}g_\theta\right) \\
&= \cos\theta\frac{\partial}{\partial r}\left(\cos\theta g_r - \frac{\sin\theta}{r}g_\theta\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\left(\cos\theta g_r - \frac{\sin\theta}{r}g_\theta\right) \\
&= \cos\theta\left(\cos\theta g_{rr} + \frac{\sin\theta}{r^2}g_\theta - \frac{\sin\theta}{r}g_{\theta r}\right) - \frac{\sin\theta}{r}\left(-\sin\theta g_r + \cos\theta g_{r\theta} - \frac{\cos\theta}{r}g_\theta - \frac{\sin\theta}{r}g_{\theta\theta}\right) \\
&= g_{rr}\cos^2\theta - \frac{(g_{r\theta}+g_{\theta r})\sin\theta\cos\theta}{r} + \frac{g_{\theta\theta}\sin^2\theta}{r^2} + \frac{g_r\sin^2\theta}{r} + \frac{2g_\theta\sin\theta\cos\theta}{r^2},
\end{aligned}$$

$$\begin{aligned}
f_{yy} &= \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\right)^2 g = \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\sin\theta g_r + \frac{\cos\theta}{r}g_\theta\right) \\
&= \sin\theta\frac{\partial}{\partial r}\left(\sin\theta g_r + \frac{\cos\theta}{r}g_\theta\right) + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\left(\sin\theta g_r + \frac{\cos\theta}{r}g_\theta\right) \\
&= \sin\theta\left(\sin\theta g_{rr} - \frac{\cos\theta}{r^2}g_\theta + \frac{\cos\theta}{r}g_{\theta r}\right) + \frac{\cos\theta}{r}\left(\cos\theta g_r + \sin\theta g_{r\theta} - \frac{\sin\theta}{r}g_\theta + \frac{\cos\theta}{r}g_{\theta\theta}\right) \\
&= g_{rr}\sin^2\theta + \frac{(g_{r\theta}+g_{\theta r})\sin\theta\cos\theta}{r} + \frac{g_{\theta\theta}\cos^2\theta}{r^2} + \frac{g_r\cos^2\theta}{r} - \frac{2g_\theta\sin\theta\cos\theta}{r^2}.
\end{aligned}$$

(式は長いけれど、途中で迷うところがないと思われる。こういうのが「工夫」であろう。)

ゆえに

$$f_{xx} + f_{yy} = g_{rr}(\cos^2\theta + \sin^2\theta) + g_{\theta\theta}\frac{\sin^2\theta+\cos^2\theta}{r^2} + g_r\frac{\sin^2\theta+\cos^2\theta}{r} = g_{rr} + \frac{1}{r}g_r + \frac{1}{r^2}g_{\theta\theta}.$$

付 録B　参考文献案内

私が参考にしたもの

このノートの内容を作るにあたっては、たくさんの本を参考にした。代表的なものをあげておく (読むことを学生諸君に勧めているわけではない)。

中尾 [9] はこの講義の前身である「微分積分学 I・同演習」の教科書に指定されたことのある本である。誤植が多い、「微分積分学 II・同演習」に相当する範囲がかなり難しい、などの理由から、私は教科書に指定することをあきらめたが、特に微分法の部分 (解析概論 I の範囲) は内容的に多くを負っている。

杉浦 [4], [5] は内容が多く (辞書的と言ってもよい)、非常にきちんと書かれた良い本である。このノートを作る際に参考にしたところは多い (というか、自力で一所懸命に内容を作った後で、この本を見ると、ずっと要領良く説明してあって、結局そちらに乗り換えたことが、一度や二度ではなかった)。この本を教科書に指定する誘惑が強かったが、まず厚さを見ただけで give up する人が出てきそうなので、あきらめることにした。

スピヴァック [6] は幾何学の先生が書いたコンパクトな (要領のいい) 本である。初めて学ぶ人が読む本としては、簡潔過ぎてかえって分かり難く[1]、適当でない気もするが、一通り勉強が済んだ後で、頭の中を整理するにはいい本だと思う。(もっとも、幾何学で多様体論を学ぶのならば、わざわざこの本を読む必要はない、という考え方もありうる。)

荷見・堀内 [10] は自習に向いていると思う。記述は素直で (あまり大げさでなく)、それでいて数学科向けの内容になっていて、好感が持てる。

高木 [7] は、明治以降における、日本の最初の世界的数学者、と言われる高木貞治 (1875–1960) による、有名な微積分の教科書であるが、この解析概論 I の範囲に関しては、あまり勧められない。この本の良いところは、積分、複素関数論や Fourier 級数の章であり (これらを勉強する場合は、現在でも大いにお勧めである)、多変数の微分法については、記述がやや古めかしくて、あえて選ぶ理由はないと思う。

参考書が欲しい人に

多変数の微分法については、必要なことはこのノートに書いたつもりである。(足りないと思うことがあったら、どうぞ指摘してください。次から書き足します。)

とはいえ、何か一冊の本を通読するのは良いことである。何か自分で微積分についての本を買って通読したい場合は、最近は特色ある本が色々と出版されているので (例えば新しく出版された岩波講座など)、自分の目で見て選ぶことを勧めたい。

[1] 私の学生当時の率直な感想である。

やる気のある人に

ある程度の微分積分学の勉強が出来たら、複素関数論をきちんと学ぶことを勧めたい。高木 [7] に魅力的な解説があるが、それ以外では高橋 [8] も面白く読める。

解析学者としての本音を言うと、解析学の楽しいテーマである、微分方程式に早くから親しんでもらいたい。常微分方程式の標準的な教科書としては、笠原 [1] などが勧められるが、読み物としては藤田 [11] が楽しめると思う。

関連図書

[1] 笠原 晧司(こうじ), 微分方程式の基礎, 朝倉書店 (1982).

[2] 齋藤 正彦; 線型代数入門, 東京大学出版会 (1966).

[3] 佐竹 一郎; 線型代数学, 裳華房 (1958).

[4] 杉浦 光夫, 解析入門 I, 東京大学出版会 (1980).

[5] 杉浦 光夫; 解析入門 II, 東京大学出版会 (1985).

[6] M. スピヴァック, 多変数解析学, 東京図書 (1972).

[7] 高木 貞治(ていじ), 解析概論 改訂第 3 版, 岩波書店 (1961).

[8] 高橋 礼司(れいじ), 複素解析, 東京大学出版会 (1990).

[9] 中尾 愼宏(みつひろ), 微分積分学, 近代科学社 (1987).

[10] 荷見(はすみ) 守助・堀内 利郎, 現代解析の基礎, 内田老鶴舗.

[11] 藤田 宏, 改訂版 応用数学, 放送大学出版協会 (1995).

[12] 一松(ひとつまつ) 信(しん), 微分積分学入門第一〜四課, 現代数学社 (1989〜).

[13] 一松(ひとつまつ) 信(しん), 代数学入門第一〜三課, 現代数学社 (1992, 1992, 1994).

[14] 一松(ひとつまつ) 信(しん), 数学とコンピュータ, 共立出版 (1995).

付 録C　ギリシャ文字、記号、注意すべき言い回し

C.1　ギリシャ文字

ギリシャ語の文字は 24 文字からなっている。大文字、小文字、対応するローマ字、読み (英語、仮名、発音記号) を以下に示す。

A	α	a	alpha	アルファ	ǽlfə
B	β	b	beta	ベータ	bíːtə, béitə
Γ	γ	g	gamma	ガンマ	gǽmə
Δ	δ	d	delta	デルタ	déltə
E	ϵ, ε	e	epsilon	イプシロン	épsilən/-lan, epsáilən
Z	ζ	z	zeta	ゼータ	zíːtə
H	η	e	eta	エータ	íːtə, éitə
Θ	θ, ϑ	t	theta	シータ	θíːtə, θéitə
I	ι	i	iota	イオタ	íoutə, aióutə
K	κ	k	kappa	カッパ	kǽpə
Λ	λ	l	lambda	ラムダ	lǽmdə
M	μ	m	mu	ミュー	mjuː, muː
N	ν	n	nu	ニュー	njuː, nuː
Ξ	ξ	x	xi	クシー	gzai, ksiː/-sai
O	o	o	omicron	オミクロン	óumikrən, oumái-
Π	π, ϖ	p	pi	パイ	pai
P	ρ, ϱ	r	rho	ロー	rou
Σ	σ, ς	s	sigma	シグマ	sígmə
T	τ	t	tau	タウ	tau, tɔː
Υ	υ	u	upsilon	ウプシロン	júːpsilən, juːpsáilən
Φ	ϕ, φ	p	phi	ファイ	fiː, fai
X	χ	c	chi	カイ	kai
Ψ	ψ	p	psi	プサイ	*p*sai, psiː/-sai
Ω	ω	o	omega	オメガ	óumigə, oumégə/-míː-

C.2　よく使われる記号

ここで説明してある記号以外にも、「集合」、「論理」の項にある記号は必見である。

i.e.	読み方は "that is," で意味は「すなわち」、「いいかえると」。
s.t.	読み方は "such that" で意味は「〜のような」。
Q.E.D.	証明の終りを表す。
$P \Longrightarrow Q$	「P ならば Q である」,「P は Q であるための十分条件」, 「Q は P であるための必要条件」。
$Q \Longleftarrow P$	(上と同じ)
$P \Longleftrightarrow Q$	「P は Q であるための必要十分条件」,「P と Q は同値」
P iff Q	"if and only if P, Q" 古くは「P のとき、またその時に限り Q」 と訳された。要するに $P \Leftrightarrow Q$ ということである。
$a \stackrel{\text{def.}}{=} b$	a を b で定義する。 (ただし b は既に意味の定まった式で、a はまだ未定義の記号とする。)
$a \equiv b$	(上と同じ) a を b で定義する。a は定義により b である。
$a \equiv b$	a と b は恒等的に等しい。
$a \equiv b$	a と b は合同である。
$a \leq b$	$a < b$ または $a = b$. ($a \leqq b$ と同じ。)
$a \geq b$	$a > b$ または $a = b$. ($a \geqq b$ と同じ。)
$\mathbf{C}$, $\mathbb{C}$	複素数全体の集合 (the set of all complex numbers).
$\mathbf{N}$, $\mathbb{N}$	自然数全体の集合 (the set of all natural numbers). (この講義では、自然数は 1 以上の整数のこととする。)
$\mathbf{Q}$, $\mathbb{Q}$	有理数全体の集合 (the set of all rational numbers).
$\mathbf{R}$, $\mathbb{R}$	実数全体の集合 (the set of all real numbers).
$\mathbf{Z}$, $\mathbb{Z}$	整数全体の集合 (the set of all integers).
(a, b)	開区間 (open interval) $\{x \in \mathbf{R}; a < x < b\}$.
$[a, b]$	閉区間 (closed interval) $\{x \in \mathbf{R}; a \leq x \leq b\}$.
$(a, b]$	$\{x \in \mathbf{R}; a < x \leq b\}$.
$[a, b)$	$\{x \in \mathbf{R}; a \leq x < b\}$.
$\displaystyle\sum_{i=n}^{m} a_i$	$a_n + a_{n+1} + \cdots + a_m$ (ただしこれは $m \geq n$ の場合で、$m < n$ のときは 0 であると約束する。)
$\displaystyle\prod_{i=n}^{m} a_i$	$a_n \times a_{n+1} \times \cdots \times a_m$ (ただしこれは $m \geq n$ の場合で、$m < n$ のときは 1 であると約束する。)
$\displaystyle\binom{n}{r}$	二項係数 ${}_nC_r$.
e	自然対数の底、ネイピアの数 $(= 2.7182818284590\cdots)$ (注: 最近の工学系の本では立体 e で表すことが多い。)
π	円周率 $(= 3.14159265358979323846\cdots)$
$x \uparrow a$	x を小さい方から a に近づける。高等学校流なら $x \to a - 0$ と書くところ。
$x \downarrow a$	x を大きい方から a に近づける。高等学校流なら $x \to a + 0$ と書くところ。

δ_{ij}	Kronecker のデルタ。 $i = j$ のとき 1, $i \neq j$ のとき 0 を表す。
$\vec{a}$	$\vec{}$ はベクトルであることを強調するための表現。
	高等学校の数学ではベクトルは必ず矢印をつけたが、大学では
	$\boldsymbol{a}$ と太字にしたり、あるいは単に a ですませる。
$[a]$	a を越えない最大の整数。いわゆる Gauss の括弧 (これは日本方言だそうです)。
$\max A$	集合 A に含まれる要素の最大値。
$\max\limits_{x \in A} f(x)$	集合 A 上の関数 f の最大値。
	言い換えると集合 $f(A) = \{f(x); x \in A\}$ に含まれる要素の最大値。
$\min A$	集合 A の最小値。
$\min\limits_{x \in A} f(x)$	集合 A 上の関数 f の最小値。
$\sup A$	集合 A が上に有界な場合には A の上限, そうでないとき ∞.
$\sup\limits_{x \in A} f(x)$	集合 A 上の関数 f の値の集合 $f(A) = \{f(x); x \in A\}$ の sup.
$\inf A$	集合 A が下に有界な場合には A の下限, そうでないとき $-\infty$.
$\inf\limits_{x \in A} f(x)$	集合 A 上の関数 f の値の集合 $f(A) = \{f(x); x \in A\}$ の inf.
$\log x$	x の自然対数 $\log_e x$.
	(工学系では $\log x = \log_{10} x$ (常用対数), $\ln x = \log_e x$ である。)
$\exp x$	x の指数 (exponential) e^x のこと。
$\sqrt[n]{a}$	$a > 0$ の場合は a の n 乗根のうち正のもの。
	$a \leq 0$ の場合は n 乗根のうち実数であるもの。
sin, cos, tan	三角関数。引数の単位はラジアン。
cot, sec, cosec	これも三角関数。それぞれ tan, cos, sin の逆数を表す。
$B(p, q)$	ベータ関数の (p, q) における値。
$\Gamma(x)$	ガンマ関数の x における値。
Arcsin, Sin^{-1}	sin の逆関数の主値。
Arccos, Cos^{-1}	cos の逆関数の主値。
Arctan, Tan^{-1}	tan の逆関数の主値。
$\sinh x$	hyperbolic sine $(= (e^x - e^{-x})/2)$.
$\cosh x$	hyperbolic cosine $(= (e^x + e^{-x})/2)$.
$\tanh x$	hyperbolic tangent $(= \sinh x / \cosh x)$.
$\lvert x \rvert$ (x は実数)	x の絶対値。
$\lvert z \rvert$ (z は複素数)	z の絶対値。
i	虚数単位 $(= \sqrt{-1})$.
	工学系では立体 i で表すことも。また電気系では j を使うことが多い。
$\Re z$, $\mathrm{Re}\, z$	複素数 z の実部。
$\Im z$, $\mathrm{Im}\, z$	複素数 z の虚部。
$\overline{z}$	複素数 z の共役複素数。
${}^tA$, A^T	行列 A の転置行列。工学系は後者の書き方が多い。
$M(m, n; \mathbf{R})$	実数を成分とする、m 行 n 列の行列全体の集合。
$M(n; \mathbf{R})$	実数を成分とする、n 次正方行列全体の集合 $(= M(n, n; \mathbf{R}))$.
$M(m, n; \mathbf{C})$	複素数を成分とする、m 行 n 列の行列全体の集合。
$M(n; \mathbf{C})$	複素数を成分とする、n 次正方行列全体の集合 $(= M(n, n; \mathbf{C}))$.

$(\vec{x}, \vec{y})$	ベクトル $\vec{x}, \vec{y}$ の内積。
$\vec{x} \cdot \vec{y}$	ベクトル $\vec{x}, \vec{y}$ の内積
$\vec{x} \times \vec{y}$	3 次元ベクトル $\vec{x}, \vec{y}$ のベクトル積。
Δx	変数 x の増分 (変化量)。
f'	関数 f の (1 階) 導関数。
f''	関数 f の 2 階導関数。
$f^{(n)}$	関数 f の n 階導関数。 ただし n は非負整数。$n = 0$ のときは f 自身を表す。

このテキストでは、なるべく標準的な記法や言い回しを採用するように努めたが、中には標準的な記法と言えるものがないものも多い。以下に掲げる記号は、かなり多くの本に載っているもので、あまり突飛なものではないが、使う場合は、最初に注意しておいた方がよいであろう。

$\neg P$	「P でない」
$P \vee Q$	「P または Q である」
$P \wedge Q$	「P かつ Q である」
$\exists! a$	「a は一意的に存在する」
$B(a; r)$	考えている空間での a を中心とする半径 r の開球。
$\overline{B}(a; r)$	考えている空間での a を中心とする半径 r の閉球。
$\vec{e}_j$	第 j 成分が 1 で、他のすべての成分が 0 であるベクトル。

C.3 その他

C.3.1 ラテン語由来の略語

etc.	et cetera の略で and so on という意味。
et al.	et alii の略で and others という意味。
i.e.	id est の略で that is という意味。
e.g.	exempli gratia の略で for example という意味。
viz.	videlicet の略で namely という意味。
q.e.d.	quod erat demonstrandum の略で which was to be deomonstrated という意味。

C.3.2 言葉遣いあれこれ

「すなわち」　「すなわち」は英語の “, that is,” に対応するもので、「言い換えると」くらいの意味である。つまり

むにゃむにゃ、すなわち、かくかくしかじか

というのは、「むにゃむにゃ」の内容を表現を変えて言い直したものが「かくかくしかじか」になるということである。

「実際」　英語で言うと、以下の二つの意味に分類できる。

“indeed,” に相当 直前に述べたことの根拠を以下に述べること (理由の説明) を意味する。“∵” くらいに考えれば良い。

“in fact,” に相当 直前に述べた以上のことが以下に言えることを意味する。

C.3.3 関数と関数値

高等学校の数学の教科書や参考書では、「関数 $f(x)$ が」という書き方が普通だったと想像する。大学で使っている数学の本にも、同じような書き方を使っているものは少なくないのだが、“(x)” のついていない「関数 f が」という書き方が多い。この微妙な相違点について説明する。

後者の流儀を一言で説明すると、

> $f(x)$ は関数 f の x での値のことで、関数そのものは f と書かなければいけない

となる。この流儀に従うと、「関数 f が」と書くのが正しく、「関数 $f(x)$ が」と書くのは厳密に言えば間違い、ということになる。決して「関数 f が」と “(x)” を省略するのは、単に面倒だから、あるいは簡潔だからという理由で省略したのではないことに注意しよう。

高等学校では「関数 $f(x)$ が」のように、関数はその変数を添えて表すのが普通であった。「y が x の関数であるとき」という表現にも現れているように、関数は二つの「ともなって変わる変数」のことだったわけである (余談だが、x を独立変数、y を従属変数と言って区別したりする)。つまり、高等学校流の数学では、変数の名前に特別の意味を与えている。

ところが、大学で学ぶ数学では、現代の数学における標準的な解釈「関数とは写像のことである」を採用している。つまり

> 集合 X の任意の要素それぞれに対して、集合 Y の要素がただ一つ決まる対応があるとき、その対応を集合 X から集合 Y への写像と呼ぶ。

この立場では、関数の変数を表す文字に何を使うかはあまり問題ではない。そもそも、

$$f\colon \mathbf{R} \to \mathbf{R}, \quad f(x) = x^2 \quad (x \in \mathbf{R})$$

というのと、

$$f\colon \mathbf{R} \to \mathbf{R}, \quad f(y) = y^2 \quad (y \in \mathbf{R})$$

は同じことを表しているわけである。似たことは高等学校の数学でも時々出て来たはずである。例えば定積分

$$\int_0^1 f(x)\,dx$$

と

$$\int_0^1 f(t)\,dt$$

は同じものを表す。

注意 C.3.1 (関数は規則？) よく「集合 X の任意の要素に集合 Y の要素をただ一つだけ対応させる規則のことを、集合 X から集合 Y への写像と呼ぶ」と言う人がいるが (このテキストでも使ったかもしれないし、授業でもうっかり言ってしまうかもしれない)、これは誤解を生みやすい表現である。規則と言うと、何か実際的な式とか、計算手順 (アルゴリズム) があるような印象を与えてしまうが、そういうものは必ずしも必要ではない。単に対応がある、だけで良い。

付 録D　集合 — 数学の言葉としての集合

現代の数学は集合論の上に記述されている。ここでは、解析概論 I の説明に必要な程度の、記号を紹介する目的で、素朴な集合論を説明する。きちんと勉強したい人は例えば

1. 井関 清志; 集合と論理, 新曜社 (1990).
2. 河田 敬義, 三村 征雄; 現代数学概説 II, 岩波書店 (1965).
3. 彌永 昌吉、彌永 健一; 集合と位相I, 岩波講座 基礎数学 (1976).
4. 島内 剛一; 数学の基礎, 日本評論社 (1971).

などを参考にすると良い (手近にあった中から選んだだけであり、他にも色々あるはずである)。

D.1　定義

思考の対象のうちで一定の範囲にあるものを一つの全体として考えた時、それを**集合**(**set**) と呼び、その範囲内の個々の対象をその集合の**元**または**要素** (**element**) と呼ぶ。

a が集合 A の要素である (a が A に含まれる、A は a を含む、とも言う) ことを $a \in A$ あるいは $A \ni a$ と書く。そうでないことを $a \notin A$ あるいは $A \not\ni a$ と書く。

A, B を集合とする。A の任意の要素 a が B の要素になっているとき、A は B の**部分集合** (**subset**) である (A は B に含まれる、B は A を含む) といい、$A \subset B$ と書く。(もちろん同じことを $B \supset A$ とも書き、その否定を $A \not\subset B$ あるいは $B \not\supset A$ と書く。以下この種のことは一々断らない。)

$A \subset B$ かつ $B \subset A$ のとき、 A と B は等しいといい、$A = B$ と書く。(つまり、集合が等しいというのは、同じ要素だけからなることを意味する。)

$A \subset B$ かつ $A \neq B$ のとき、 A は B の**真部分集合**であると言い、$A \subsetneqq B$ と書く。

注意 D.1.1 我々の流儀では、$A \subset B$ は $A = B$ の場合を含んでいることに注意しよう。別の流儀では、「$A \subset B$」という表現を A が B の真部分集合であるという意味に用い、$A = B$ も含み得る場合は $A \subseteqq B$ あるいは $A \subseteq B$ と表す。

要素を一つも含まない集合を**空集合** (**empty set**) と呼び、$\emptyset$ あるいは ϕ で表す。空集合は任意の集合に含まれる。

D.2　集合の表し方、良く使う記号

集合の表し方として、次の二つの方法が基本的である。

(1) $A = \{1, 2, 3\}$ のように元を表示して $\{,\}$ で括る方法。

(2) $A = \{x; x$ は整数かつ $1 \leq x \leq 3\}$ のように “$\{x; x$ についての条件 $\}$” と表す方法。

例 D.2.1 次の数の集合の記号は良く使われる。

(1) $\mathbf{N} = \{1, 2, 3, \cdots\}$: 自然数 (natural numbers) 全体の集合。

(2) $\mathbf{Z} = \{\cdots, -2, -1, 0, 1, 2, \cdots\}$: 整数 (integers) 全体の集合。

(3) $\mathbf{Q} = \left\{\dfrac{m}{n}; m, n \in \mathbf{Z}, n \neq 0\right\}$: 有理数 (rational numbers) 全体の集合。

(4) $\mathbf{R} = \{x; x$ は実数 $\}$: 実数 (real numbers) 全体の集合。

(5) $\mathbf{C} = \{z; z$ は複素数 $\}$: 複素数 (complex numbers) 全体の集合。

例 D.2.2 以下 a, b は $a \leq b$ なる実数とする。

(1) $[a, b] = \{x \in \mathbf{R}; a \leq x \leq b\}$:　閉区間。

(2) $(a, b) = \{x \in \mathbf{R}; a < x < b\}$:　開区間。

(3) $(a, b] = \{x \in \mathbf{R}; a < x \leq b\}$,　$[a, b) = \{x \in \mathbf{R}; a \leq x < b\}$:　半開区間。

端点として $+\infty$, $-\infty$ を用いることもある。例えば

$$(-\infty, b] = \{x \in \mathbf{R}; x \leq b\}, \quad (a, +\infty) = \{x \in \mathbf{R}; a < x\}, \quad (-\infty, +\infty) = \mathbf{R}.$$

端点が $\pm\infty$ である場合、そこは開いているはずで $[-\infty, b]$ や $(a, \infty]$ と書くのは変である。

D.3　集合算

A, B を集合とする。

(1) $A \bigcup B \stackrel{\text{def.}}{=} \{x; x \in A$ または $x \in B\}$ で定義される集合 $A \bigcup B$ を A と B の**合併**（がっぺい） (**union**) と呼ぶ。

(2) $A \bigcap B \stackrel{\text{def.}}{=} \{x; x \in A$ かつ $x \in B\}$ で定義される集合 $A \bigcap B$ を A と B の**共通部分** (**交わり**, **intersection**) と呼ぶ。

(3) $A \setminus B \stackrel{\text{def.}}{=} \{x; x \in A$ かつ $x \notin B\}$.

考察の全範囲となる集合 X があり、$A \subset X$ となっているとき、

$$A^c \overset{\text{def.}}{=} X \setminus A$$

とおき、A の**補集合** (**complement** of A) と呼ぶ。

A, B を集合とするとき、$x \in A, y \in B$ の順序を考慮した対[1](x, y) 全体の集合 $\{(x, y); x \in A,\ y \in B\}$ を A と B の**直積** (**product**) と呼び、$A \times B$ で表す。

D.4 写像

X, Y を空でない集合とする。X の各元に Y の元を一つずつ対応させる規則 f が与えられた時、f を X から Y への**写像** (**mapping**) といい、$f : X \to Y$ あるいは $X \xrightarrow{f} Y$ と書く。

写像 f により、$x \in X$ に $y \in Y$ が対応する時、y を x による x の**像** (**image**) と呼び (x における f の**値**ともいう)、$f(x)$ で表す。また f により x が y に対応していることを、しばしば $f\colon x \longmapsto y$ と表す。

X を写像 f の**定義域**という。A を X の部分集合とするとき、A の元の f のよる像の全体 $\{f(x); x \in A\}$ を f による A の**像**と呼び、$f(A)$ で表す。特に f による定義域 X の像 $f(X)$ を f の**値域** (**range**) と呼ぶ。Y には定着した名前がないが、ここでは終域と呼んでおく。

逆像　B を Y の部分集合とするとき、f によって B の要素に写される $x \in X$ の全体 $\{x \in X; f(x) \in B\}$ を f による B の**逆像** (**inverse image**, 引き戻し (**pull back**)) と呼び、$f^{-1}(B)$ で表す。

全射　f が X から Y の**上への**写像 (**全射**, **surjection**) であるとは、$f(X) = Y$ であること、言い換えると、Y の任意の要素 y に対して $y = f(x)$ となるような $x \in X$ が存在することである。

単射　f が X から Y への **1 対 1 の**写像 (**単射**, **injection**) であるとは、X の任意の要素 x_1, x_2 について、$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$ が成り立つこと。(対偶を取ると $f(x_1) = f(x_2) \Rightarrow x_1 = x_2$ ということでもある。)

全単射、逆写像　f が X から Y への**全単射** (**双射**, **bijection**) であるとは、f が X から Y への全射かつ単射であることを言う。このとき f によって X の元と Y の元は一対一に対応する。それゆえ逆向きの対応が定義される。それを f の**逆写像** (**inverse mapping**) と呼び、

[1]つまり $(x_1, y_1) = (x_2, y_2) \Leftrightarrow (x_1 = x_2 \text{ and } y_1 = y_2)$ ということ。**順序対**と言う。

f^{-1} と書く。$f^{-1}: Y \to X$ である。$x \in X, y \in Y$ に対して

$$f(x) = y \Longleftrightarrow x = f^{-1}(y)$$

である。

f の逆写像が存在するとき、Y の部分集合 B に対し、$f^{-1}(B)$ という記号は、f による B の逆像、f^{-1} による B の像という二つの意味を持ち得るが、集合として両者は一致するので、混乱はない。

合成写像　二つの写像 $f\colon X \to Y$, $g\colon Z \to W$ があり、$f(X) \subset Z$ のとき、合成可能であるといい、その場合

$$h(x) = g(f(x)) \quad (x \in X)$$

によって定義される $h\colon X \to W$ を f と g の**合成写像**(ごうせいしゃぞう)といい、$g \circ f$ で表す。

グラフ　X と Y の直積集合 $X \times Y$ の部分集合 $\{(x, f(x)); x \in X\}$ のことを f の**グラフ** (**graph** of f) と呼び、しばしば $\operatorname{graph} f$ と書く。

添字づけられた集合族の合併、共通部分　Λ, X を集合とし、$P(X)$ を X の部分集合全体のなす集合 (**冪集合**(べきしゅうごう)と呼ばれる) とする。このとき写像 $f\colon \Lambda \to P(X)$ を、X の部分集合の**添字**(そえじ)**づけられた族**と言う。このとき $f(\lambda) = X_\lambda$ のように書き、$f = (X_\lambda)_{\lambda \in \Lambda}$ のように表す。このとき Λ を族 f の**添字集合**と呼ぶ。X の部分集合の族 $\{X_\lambda; \lambda \in \Lambda\}$ に対し、その**合併**を

$$\bigcup_{\lambda \in \Lambda} X_\lambda \overset{\text{def.}}{=} \{x \in X; \exists \lambda \in \Lambda \text{ s.t. } x \in X_\lambda\}$$

で、**共通部分**を

$$\bigcap_{\lambda \in \Lambda} X_\lambda \overset{\text{def.}}{=} \{x \in X; \forall \lambda \in \Lambda \text{ に対して } x \in X_\lambda\}$$

で定義する。

制限　A を X の部分集合とするとき、

$$g(x) = f(x) \quad (x \in A)$$

により定まる関数 $g\colon A \to Y$ を f の A への**制限**(せいげん) (**restriction**) と呼び、$f|_A$ で表す。しばしば終域の方も $f(A) \subset B \subset Y$ なる B (大抵は $B = f(A)$ とする) に置き換えることもある。(f と $f|_A$, 両者の値を考えることの出来る集合 A の任意の要素 x に対して $f(x) = f|_A(x)$ であるから、わざわざ区別して考えるまでもない、と感じる人がいるかも知れないが、逆関数が存在するかどうか、のような問題の場合、全射性、単射性は大事なので、関数の定義域や終域の違いは重要である。)

D.5　公式あれこれ

写像による集合の像、逆像について。

$$f(\emptyset) = \emptyset.$$

$$A \subset B \Longrightarrow f(A) \subset f(B).$$

$$f(A \cup B) = f(A) \cup f(B).$$

$$f(A \cap B) \subset f(A) \cap f(B).$$

$$f^{-1}(\emptyset) = \emptyset.$$

$$A \subset B \Longrightarrow f^{-1}(A) \subset f^{-1}(B).$$

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B).$$

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B).$$

$$f^{-1}(A^c) = f^{-1}(A)^c.$$

$$f^{-1}(f(A)) \supset A \quad (f \text{ が単射なら等号成立}).$$

$$f(f^{-1}(B)) \subset B \quad (f \text{ が全射なら等号成立}).$$

付 録E　論理についてのメモ

「量称記号」$\forall$, $\exists$ を含む論理式の扱いについて、説明する (決して系統的ではない)。

E.1　量称の読み方

以下では、x は集合 X の上を動く変数で、$P(x)$, $Q(x)$ は x に関する条件 (述語, 命題関数) とする。

量称記号で述語中の変数を「束縛」して命題が得られる。

$\forall x P(x)$　すべて (任意) の $x \in X$ に対して $P(x)$ である。

$\exists x P(x)$　$P(x)$ であるような $x \in X$ が存在する (ある x が存在して $P(x)$)。

以下の公式が成り立つ ($\neg$ は否定を表す)。

(1) $\neg(\forall x P(x)) = (\exists x)(\neg P(x))$.

(2) $\neg(\exists x P(x)) = (\forall x)(\neg P(x))$.

(3) $\forall x(P(x) \wedge Q(x)) = (\forall x P(x)) \wedge (\forall x Q(x))$.

(4) $\exists x(P(x) \vee Q(x)) = (\exists x P(x)) \vee (\exists x Q(x))$.

条件の中に 2 つの変数がある場合は、量称記号も 2 つ現れうるが、これについては以下の公式が大切である (x は集合 X 上を、y は集合 Y 上を動くものとし、$P(x,y)$ は x, y に関する条件であるとする)。

(1) $\forall x \forall y P(x,y) = \forall y \forall x P(x,y)$.

(2) $\exists x \exists y P(x,y) = \exists y \exists x P(x,y)$.

(3) $\exists y \forall x P(x,y) \Rightarrow \forall x \exists y P(x,y)$.

注意 E.1.1 *(3) だけ等号でなく、$\Rightarrow$ であることに注意しよう。一言でまとめると $\forall$ 同士、$\exists$ 同士は順序を入れ換えてもよいが、$\forall$ と $\exists$ を入れ換えることはできない (意味が変わってしまう)。これについては、次節で例をあげて調べることにする。*

E.2 「普通の数学向き」の量称

第 1 節のような書き方で理論上は十分なのだが、通常の数学の議論に用いるには不便なところがある (記述が繁雑になりやすい)。普通はもう少し工夫した表現が使われる。後のためにきちんとまとめておこう。

$(\exists x)\ (A(x) \wedge P(x))$ は、「$A(x)$ かつ $P(x)$ であるような x が存在する」という意味だが、これを、「$A(x)$ という条件の下で、$P(x)$ を満たす x が存在する」とみなすのが便利であることが多い。このことを、$(\exists x : A(x))\ P(x)$ と書こう。 (この “$: A(x)$” のところは、場合に応じて「柔軟な」書き方をする。以下の例を参照せよ。)

$(\forall x)(A(x) \Rightarrow P(x))$ は、「すべての x に対し、$A(x)$ ならば $P(x)$ が成り立つ」という意味だが、これを、「条件 $A(x)$ を満たすすべての x に対して $P(x)$ が成り立つ」とみなすのが便利であることが多い。このことを、$(\forall x : A(x))\ P(x)$ と書こう。

例 E.2.1

$$\forall x \quad (x > 0 \Rightarrow x + \frac{1}{x} \geq 2)$$

を

$$(\forall x > 0) \quad x + \frac{1}{x} \geq 2$$

のように書く。

例 E.2.2

$$\exists x \quad ((x > 0) \wedge (x^2 = 2))$$

を

$$(\exists x > 0) \quad x^2 = 2$$

のように書く。

例 E.2.3 (ピタゴラス数 (Pythagorean numbers))

$$\exists x \exists y \exists z \quad (x \in \mathbf{N} \wedge y \in \mathbf{N} \wedge z \in \mathbf{N} \wedge x^2 + y^2 = z^2)$$

を

$$(\exists x \in \mathbf{N})(\exists y \in \mathbf{N})(\exists z \in \mathbf{N}) \quad x^2 + y^2 = z^2$$

のように書く。

例 E.2.4 (連続関数の定義) 関数 f の点 a における連続性の定義は「任意の正数 ε に対して、十分小さな正数 δ が存在して、$|x-a| < \delta$ を満たすすべての x に対して、$|f(x) - f(a)| < \varepsilon$」というものであった。これは、この節で導入した記法によれば

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x : |x - a| < \delta) \quad |f(x) - f(a)| < \varepsilon$$

と書けるが、前節の流儀で書こうとするとかなり面倒になる (興味ある人は試してみるとよい)。

前節のようなタダの $\forall x, \exists x$ に対する公式と同様のものが、$(\forall x : A(x)), (\exists x : A(x))$ に対しても成り立つ。

(1) $\neg(\exists x : A(x))P(x) = (\forall x : A(x))(\neg P(x))$.

(2) $\neg(\forall x : A(x))P(x) = (\exists x : A(x))(\neg P(x))$.

(3) $(\forall x : A(x))(P(x) \wedge Q(x)) = (\forall x : A(x))P(x) \wedge (\forall x : A(x))Q(x)$.

(4) $(\exists x : A(x))(P(x) \vee Q(x)) = (\exists x : A(x))P(x) \vee (\exists x : A(x))Q(x)$.

(5) $(\exists x : A(x))(\exists y : B(y))P(x, y) = (\exists y : B(y))(\exists x : A(x))P(x, y)$.

(6) $(\forall x : A(x))(\forall y : B(y))P(x, y) = (\forall y : B(y))(\forall x : A(x))P(x, y)$.

(7) $(\exists y : B(y))(\forall x : A(x))P(x, y) \Rightarrow (\forall x : A(x))(\exists y : B(y))P(x, y)$.

E.3 量称記号 $\forall, \exists$ の順序について

量称記号の順序について、実例で考えてみよう。

例 E.3.1 (連続性と一様連続性) I を $\mathbf{R}$ の区間、$f : I \to \mathbf{R}$ を写像とする。f が I で連続であるということを論理記号で書くと

$$(\forall x \in I)(\forall \varepsilon > 0)(\exists \delta > 0)(\forall y \in I : |x - y| < \delta) \quad |f(x) - f(y)| < \varepsilon$$

となる。また、f が I で一様連続であるということを、論理記号で書くと

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in I)(\forall y \in I : |x - y| < \delta) \quad |f(x) - f(y)| < \varepsilon$$

となる。
連続性と一様連続性は異なる概念であるが、それが $\forall, \exists$ の順番の違いだけで表されている。

$\forall$ 同士、あるいは $\exists$ 同士は順番を入れ換えても式の内容は変わらないのだが、$\forall$ と $\exists$ の順番は入れ換えると意味が変わってしまうのである。$\exists$ が先に現れる方（上の例では一様連続性）が強い条件である。

例 E.3.2 (Archimedes（アルキメデス） の公理) 2 つの正の実数 a, b を取ったとき (b がどんなに大きくても、あるいは a がどんなに小さくても)、a を十分たくさん足してやれば b を追い抜く。

$$(\forall a > 0)(\forall b > 0)(\exists n \in \mathbf{N}) \quad na > b$$

これは公理というくらいで、真な命題であるが、

$$(\forall a > 0)(\exists n \in \mathbf{N})(\forall b > 0) \quad na > b$$

は偽である。

例 E.3.3 (じゃんけん) 3 点からなる集合 $J = \{\, ぐう, ちょき, ぱあ \,\}$ に、次のような 2 項関係 $\succ$ を導入する。

$$ぐう \succ ちょき, \quad ちょき \succ ぱあ, \quad ぱあ \succ ぐう$$

これ以外の場合は、$\succ$ は不成立とする (要するに「じゃんけん」の勝ち負けの判定)。例えば、ぱあ $\not\succ$ ちょき. こうすると、

$$(\forall t \in J)(\exists k \in J) \quad k \succ t$$

(どんな手 t に対しても、それに勝つ手 k がある) は真であるが、

$$(\exists k \in J)(\forall t \in J) \quad k \succ t$$

(どんな手 t に対しても勝ってしまう "必勝手" k がある) は偽である。

E.4 複数の量称記号を含む式の読み方についての注意

複数の量称がある場合、左から右に順に読み進むのが基本である。

(1) $\forall x \forall y P(x,y)$ 　すべて (任意) の x, すべて (任意) の y について $P(x,y)$.

(2) $\exists x \exists y P(x,y)$ 　ある x, ある y が存在して $P(x,y)$.

(3) $\forall x \exists y P(x,y)$ 　すべて (任意) の x に対して、ある y が存在して $P(x,y)$.

(4) $\exists y \forall y P(x,y)$ 　ある y が存在して、すべて (任意) の x に対して $P(x,y)$.

世の中には、このような読み方があまり好きでなくて、

(3) $\forall x \exists y P(x,y)$ を「すべての x に対して、$P(x,y)$ であるような y が存在する」.

(4) $\exists y \forall x P(x,y)$ を「すべての x に対して $P(x,y)$ であるような y が存在する」.

と「読む」人もいる (読点「、」に注意)。日本語としては、あるいはこちらの表現の方が普通かもしれないが、間違えやすいので注意した方がよいであろう。もっと表現を工夫するか、あるいは上に示したように、機械的に左から右に読んだ方が誤解が生じにくい。機械的な読み方は、しばしば「日本語とは思えない」と非難されるが仕方がない。

E.5 空集合の論理

空集合 $\emptyset$ は、任意の集合 A の部分集合である、すなわち

$$\emptyset \subset A \tag{E.1}$$

が成り立つ。このことを「知っている」人は多いだろうが、証明を考えたことはあるだろうか?

集合 X の部分集合 A, B について、

$$A \subset B \overset{\text{def.}}{\Leftrightarrow} (\forall x \in A) \quad x \in B$$

であった。つまり A のメンバー全員に「B のメンバーであるか」というテストを受けさせ、全員合格ならば晴れて「$A \subset B$」であると言えるわけである。

すると (E.1) を証明するには、

$$(\forall x \in \emptyset) \quad x \in A$$

を示さねばならない。これは真なのであるが、納得できるだろうか？

空集合はその定義から、要素を一つも持たない、すなわち $x \in \emptyset$ となる x は存在しない。テストの例え話を続けると、受験生がいないテストは「全員合格」なのだろうか、そうでないのだろうか、ということである。初めて出くわした人は戸惑うかも知れないが、数学ではこういう場合「全員合格」であると考える (言い換えると、全称記号 $\forall$ はそういう意味である、と約束する)。

似たようなことはあちこちで出て来る。

- 空集合は有界である (めったに使わないが)。
- 空集合は開集合である。

数学で「p ならば q」と推論するときに、条件 p を満たす場合は本当に存在するかどうか、直接は問題にならないことが多いことに注意しよう。

付 録F　R の有界集合と上限、下限

F.1　有界集合

定義 F.1.1 (上に有界、上界、下に有界、下界、有界) A を $\mathbf{R}$ の部分集合とする。

(1) A が**上(うえ)に有界(ゆうかい)** (bounded from above) $\overset{\text{def.}}{\Leftrightarrow}$ $(\exists U \in \mathbf{R})\ (\forall x \in A) \quad x \leq U.$
(この U を A の**上界(じょうかい)** (**upper bound**) と呼ぶ。)

(2) A が**下(した)に有界** (bounded from below) $\overset{\text{def.}}{\Leftrightarrow}$ $(\exists L \in \mathbf{R})\ (\forall x \in A) \quad L \leq x.$
(この U を A の**下界(かかい)** (**lower bound**) と呼ぶ。)

(3) A が**有界** $\overset{\text{def.}}{\Leftrightarrow}$ $\left\{ \begin{array}{l} A \text{ が上に有界かつ下に有界、すなわち} \\ (\exists R > 0)\ (\forall x \in A) \quad |x| \leq R. \end{array} \right.$

数列について有界性を考えることが多い。上の定義から明らかなように思う人もいるだろうが、厳密に言うと、数列 $\{x_n\}_{n \in \mathbf{N}}$ は集合ではなく、$\mathbf{N}$ から $\mathbf{R}$ への写像である。例えば $x_n = (-1)^n$ で定義される数列

$$\{1, -1, 1, -1, 1, -1, \cdots\}$$

は

$$\{1, -1\}$$

のように簡略してはいけない。従って、数列の有界性は別に定義する必要がある。もちろん、写像と考えた場合の値域の有界性として定義するわけである。

定義 F.1.2 (数列の有界性) $\mathbf{R}$ 内の数列 $\{x_n\}_{n \in \mathbf{N}}$ が有界であるとは、値の集合

$$\{x_n; n \in \mathbf{N}\}$$

が有界であること、言い替えると、

$$(\exists R > 0)\ (\forall n \in \mathbf{N}) \quad |x_n| \leq R$$

が成り立つことであると定義する。

F.2 上限、Weierstrass の定理

定義 F.2.1 (上限、下限) A を $\mathbf{R}$ の部分集合とする。

(1) A の上界のうちで最小のものが存在するとき、それを A の**上限**（じょうげん）(**supremum**) と呼ぶ。

(2) A の下界のうちで最大のものが存在するとき、それを A の**下限**（かげん）(**infimum**) と呼ぶ。

命題 F.2.1 (上限、下限の論理式による表現) A を $\mathbf{R}$ の空でない部分集合、$S, I \in \mathbf{R}$ とする。

(1) S が A の上限であるための必要十分条件は次の (a), (b) が成り立つことである。

(1) (S は A の上界である)

$$(\forall x \in A) \quad x \le S.$$

(2) (S より小さな数は A の上界ではない)

$$(\forall S' < S)(\exists x \in A) \quad S' < x.$$

この (b) を "$(\forall \varepsilon > 0)$ $(\exists x \in A)$ $S - \varepsilon < x$" と書いてある本も多い (もちろん同じことである)。

(2) I が A の下限であるための必要十分条件は次の (a), (b) が成り立つことである。

(1) (I は A の下界である)

$$(\forall x \in A) \quad I \le x.$$

(2) (I より大きな数は A の下界ではない)

$$(\forall I' > I)(\exists x \in A) \quad x < I'.$$

この (b) を "$(\forall \varepsilon > 0)$ $(\exists x \in A)$ $x < I + \varepsilon$" と書いてある本も多い。

命題 F.2.2 (最大値は上限、最小値は下限) A を $\mathbf{R}$ の部分集合とする。

(1) A の最大値が存在するならば、それは上限である。

(2) A の最小値が存在するならば、それは下限である。

証明 (1) を証明する。A の最大値 M が存在すると仮定する。M が A の最大値であるとは、次の (a), (b) が成り立つことである。

(a) $(\forall x \in A) \quad x \le M.$

(b) $M \in A$.

このうち (a) は M が A の上界であることを主張している。次に S' を $S' < M$ なる任意の実数としよう。このとき、

$$(\exists x \in A) \quad S' < x$$

は確かに成り立つ。実際 $x = M$ とすると、(b) より $x \in A$ で $S' < x$. ゆえに (前命題を使って) M は A の上限である。■

注意 F.2.1 (上限は最大値よりも一般的) 集合に最大値がなくても上限は存在する場合がある。例えば

$$A = (0, 1) \text{ について、} \sup A = 1 \text{ だが } \max A \text{ は存在しない。}$$

まず

$$\sup A = 1$$

であることを確かめるのはやさしい。実際、$\forall x \in (0, 1)$ に対して $x < 1$ であるから当然 $x \le 1$ である。また、1 より小さな任意の S' に対して $S' < x$ なる $x \in A$ が存在するのは、次のようにして明らか。

- $S' < 0$ の場合、 $x = 1/2$ とすればよい。
- $0 \le S' < 1$ の場合、$x = (1 + S')/2$ とすると、$x \in A$ かつ $S' < x$ となる。

一方、A の最大値はもしも存在したと仮定すると、それは $\sup A = 1$ に等しいはずだが、$1 \notin A$ であるから、1 は A の最大値ではない。

分かったことをまとめておくと、

上限は最大値を一般化した概念である (最大値は上限だが、上限は最大値とは限らない)。

次の定理は重要である。

定理 F.2.1 (Weierstrass) (1) $\mathbf{R}$ の空でない上に有界な部分集合は上限を持つ。

(2) $\mathbf{R}$ の空でない下に有界な部分集合は下限を持つ。

証明に換えてお話　上の定理の証明は実数の構成法に依存する。実数を構成する議論をするのはなかなか大変である (微妙な話が長く続く)。ここではこれを実数の公理のように考えることにする (要するに認めてしまう)。$\mathbf{R}$ の代わりに有理数全体の集合 $\mathbf{Q}$ を考えると、上の定理は成立しなくなる。「$\mathbf{R}$ の上に有界な部分集合には必ず上限が存在する」というのは、**実数の連続性**と呼ばれる $\mathbf{R}$ の重要な性質の一つの表現である。実数の連続性の別の表現の仕方としては、「Cauchy 列は必ず収束する」という**完備性**などがある。

注意 F.2.2 (Weierstrass の定理の守備範囲外のケースについて) もちろん、

- 上に有界でない集合は上限を持たない。
- 下に有界でない集合は下限を持たない。

ことはすぐ分かる。ところで、空集合は有界であるが、任意の実数が空集合の上界にも下界にもなる (考えてみよう)。当然、最小の上界、最大の下界は存在しない。つまり

- 空集合は上限、下限を持たない。

次の命題は是非とも知っておくべき重要なものである。

命題 F.2.3 (有界単調数列の収束定理) $\{x_n\}_{n\in\mathbf{N}}$ を $\mathbf{R}$ 内の数列とする。

(1) $\{x_n\}_{n\in\mathbf{N}}$ が上に有界で、単調増加

$$i \le j \Longrightarrow x_i \le x_j$$

ならば、収束する。

(2) $\{x_n\}_{n\in\mathbf{N}}$ が下に有界で、単調減少

$$i \le j \Longrightarrow x_i \ge x_j$$

ならば、収束する。

証明　どちらでも同じことだから (1) を証明する。数列 $\{x_n\}_{n\in\mathbf{N}}$ の値の集合 $\{x_n; n\in\mathbf{N}\}$ を A とする。仮定から A は上に有界であり、また明らかに空でないから上限が存在する。それを S と書こう。まず、S は A の上界であるから、

$$(\forall n\in\mathbf{N})\quad x_n \le S$$

は明らかである。さて、ε を任意の正数とすると、$S-\varepsilon$ は S より小さいから、

$$(\exists N\in\mathbf{N})\quad S-\varepsilon < x_N.$$

数列の単調増加性より $(\forall n\in N : n\ge N)$ に対して

$$x_N \le x_n.$$

ゆえに

$$S-\varepsilon \le x_n \le S,$$

従って、

$$|S-x_n| \le \varepsilon.$$

これは $\lim\limits_{n\to\infty} x_n = S$ を示している。■

上限、下限について、記号を準備しておこう。

定義 F.2.2 (集合の sup, inf**)** (1) A を空でない $\mathbf{R}$ の部分集合とする時、$\sup A$ を

$$\sup A \overset{\text{def.}}{=} \begin{cases} A \text{ の上限} & (A \text{ が上に有界なとき } (=\text{上限が存在するとき})) \\ \infty & (A \text{ が上に有界でないとき } (=\text{上限が存在しないとき})) \end{cases}$$

で定める。

(2) A を空でない $\mathbf{R}$ の部分集合とする時、$\inf A$ を

$$\inf A \overset{\text{def.}}{=} \begin{cases} A \text{ の下限} & (A \text{ が下に有界なとき } (=\text{下限が存在するとき})) \\ -\infty & (A \text{ が下に有界でないとき } (=\text{下限が存在しないとき})) \end{cases}$$

で定める。

定義 F.2.3 (関数の sup, inf**)** $f\colon X \to \mathbf{R}$ とするとき、$f(X) \overset{\text{def.}}{=} \{f(x); x \in X\}$ とおく。

(1) f が上に有界 $\overset{\text{def.}}{\Leftrightarrow}$ $f(X)$ が $\mathbf{R}$ の上に有界な部分集合。
このとき、

$$\sup_{x\in X} f(x) \overset{\text{def.}}{=} \sup f(X)$$

により $\sup\limits_{x\in X} f(x)$ を定める。

(2) f が下に有界 $\overset{\text{def.}}{\Leftrightarrow}$ $f(X)$ が $\mathbf{R}$ の下に有界な部分集合。
このとき、

$$\inf_{x\in X} f(x) \overset{\text{def.}}{=} \inf f(X)$$

により $\inf\limits_{x\in X} f(x)$ を定める。

次の命題は便利で、しばしば使われる。

命題 F.2.4 (sup **に収束する数列の存在)** A を $\mathbf{R}$ の空でない部分集合とすると、A 内の数列 $\{x_n\}_{n\in\mathbf{N}}$, $\{y_n\}_{n\in\mathbf{N}}$ で $\lim\limits_{n\to\infty} x_n = \sup A$, $\lim\limits_{n\to\infty} y_n = \inf A$ なるものが存在する。

問 上の命題を証明せよ。

F.3 R の有界集合にまつわる有名な定理

$\mathbf{R}$ の有界部分集合、有界数列に関係する重要な定理がたくさんあるが、ここでは証明抜きで列挙しておく。

Weierstrass の定理 F.2.1 以外に、まず次が重要である。

定理 F.3.1 (Bolzano-Weierstrass) $\mathbf{R}$ の有界な数列は収束部分列を含む。

有界な閉集合はとりわけ重要である。まず、

定理 F.3.2 (Heine-Borel) $\mathbf{R}$ の部分集合 A がコンパクトであるための必要十分条件は、A が有界閉集合であることである。言い替えると、$\mathbf{R}$ の任意の部分集合 A について、次の二つの条件は互いに同値である。

(i) A は有界閉部分集合である。

(ii) A の任意の開被覆は有限部分被覆を持つ。

これから重要な結果がたくさん導かれる。まず「距離空間において、コンパクト $\Leftrightarrow$ 点列コンパクト」という一般的な原理から、

定理 F.3.3 (R の点列コンパクト性) $\mathbf{R}$ の有界閉集合 A 内の数列は、A 内の数に収束する部分列を含む。

さらには、「コンパクト集合の連続関数による像はコンパクト」という一般的な原理の $\mathbf{R}$ 版

定理 F.3.4 連続関数による $\mathbf{R}$ の有界閉集合の像は $\mathbf{R}$ の有界閉集合である。

と、その系

定理 F.3.5 $\mathbf{R}$ の有界閉集合上の連続関数は、最大値、最小値を持つ。

が導かれる。

さらに

定理 F.3.6 (Heine の定理) $\mathbf{R}$ の有界閉集合上の連続関数は、一様連続である。

という定理も成り立つ。

付 録G　Landau の記号

微積分では、記述を簡略化するために、$O(h)$ $(h \to 0)$ とか $o(n^{-k})$ $(n \to \infty)$ のような記号が頻繁(ひんぱん)に用いられる。これは **Landau の記号**と呼ばれるものであるが、ごく簡単に解説しておこう。

この記法は、数列に対しても関数に対しても使われるが、どれでも大した違いはないから、ここでは 1 変数実数値関数について説明する。

以下では、関数 f, g がともに α の除外近傍[1]

$$B(\alpha;\varepsilon) \setminus \{\alpha\} = \{x; 0 < |x-\alpha| < \varepsilon\}$$

で定義されているとして、$x \to \alpha$ の場合を考える ($x \to \infty$ 等も同様である)。

定義 G.0.1 (小さな $o(\cdot)$) 条件

$$\lim_{x\to\alpha} \frac{|f(x)|}{|g(x)|} = 0$$

が成り立つことを記号

$$f(x) = o(g(x)) \quad (x \to \alpha)$$

で表す。

この記号は、f, g がともに α で無限小である、すなわち

$$\lim_{x\to\alpha} f(x) = \lim_{x\to\alpha} g(x) = 0$$

であるとき使われることが多い。その場合「α において f が g より**高位の無限小**である」と言う (直感的には、$x \to \alpha$ のとき、$g(x)$ よりも $f(x)$ の方が速く 0 に近付く、ということである)。もっとも、この記号 $o(\cdot)$ は、そうでない場合に使われることも少なくない。

例 G.0.1 $\lim_{x\to\alpha} f(x) = 0$ ということを表すため、$f(x) = o(1)$ $(x \to \alpha)$ と書いたりする。$\lim_{x\to\alpha} \frac{|f(x)|}{1} = 0$ ということで、$g(x) \equiv 1$ とみなしているわけである。もちろん、この $g(x)$ は $x = \alpha$ で無限小ではない。

[1] あまり一般的な用語ではない。

定義 G.0.2 (大きな $O(\cdot)$) α において f が g で押さえられるとは

　十分大きい正数 C を取ると、α に十分近い任意の x に対して $|f(x)| \le C|g(x)|$

がなりたつことと定義し、記号

$$f(x) = O(g(x)) \quad (x \to \alpha)$$

で表す。すなわち

$$f(x) = O(g(x)) \quad (x \to a) \quad \overset{\text{def.}}{\Leftrightarrow} \quad \exists C > 0, \exists \varepsilon > 0, \forall x \in B(\alpha; \varepsilon) \setminus \{\alpha\} \\ |f(x)| \le C|g(x)|.$$

この記号も f, g が α で無限小のときに使われることが多い。その場合「f は (少なくとも) g と**同位の無限小**である」などと言われる。
微積分に現れる典型的な例をあげよう。

例 G.0.2 関数 $f\colon \mathbf{R} \supset I \to \mathbf{R}$ が $\alpha \in I$ において微分可能で、微分係数が A であるとは、

$$\lim_{h \to 0} \frac{f(\alpha + h) - f(\alpha)}{h} = A$$

が成り立つことであるが、この条件は

$$f(\alpha + h) - f(\alpha) - Ah = o(h) \quad (h \to 0)$$

のように書き直すことが出来る。また関数 f が C^n- 級であるとき、Taylor の定理から

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(x)}{(n-1)!}h^{n-1} + \frac{f^{(n)}(x+\theta h)}{n!}h^n$$

となる θ $(0 < \theta < 1)$ が存在することが示されるが、この式から導かれる

$$f(x+h) = f(x) + f'(x)h + \frac{f''(x)}{2!}h^2 + \cdots + \frac{f^{(n-1)}(x)}{(n-1)!}h^{n-1} + O(h^n) \quad (h \to 0)$$

はしばしば使われる (剰余項の面倒な形が必要ないことも多く、その場合は簡単に書けて便利だから)。

付 録H　極座標

H.1　平面極座標

平面に直交する座標軸 x 軸、y 軸を取って座標を入れる xy 座標系で (x, y) という座標を持つ点 P の原点からの距離を r, x 軸の正方向となす角を φ $(0 \leq \varphi < 2\pi)$ とすると、

$$\begin{cases} x = r\cos\varphi \\ y = r\sin\varphi \end{cases}$$

がなりたつ。

写像

$$f \colon [0, \infty) \times [0, 2\pi) \ni (r, \varphi) \longmapsto (x, y) \in \mathbf{R}^2$$

は C^∞- 級で、定義域を $r \neq 0$ の範囲、すなわち $(0, \infty) \times [0, 2\pi)$ に制限すれば 1 対 1 である。特に

$$f|_{(0,\infty)\times[0,2\pi)} \colon (0, \infty) \times [0, 2\pi) \ni (r, \varphi) \longmapsto f(r, \varphi) \in \mathbf{R}^2 \setminus \{(0, 0)\}$$

は全単射である。

逆の計算、つまり (x, y) から (r, φ) を求めるには、r の方は

$$r = \sqrt{x^2 + y^2}$$

と x, y の式として簡単に表されるが、φ の方は標準的な記法がない。強いて書けば

$$(x, y) \neq (0, 0) \text{ のとき} \quad \varphi = \arg(x, y) \in [0, 2\pi)$$

であろうか。多くの本に

$$\varphi = \tan^{-1}\left(\frac{y}{x}\right)$$

とあるが、これは色々と問題を含んでいる式である (要するにマズイ)。実際、$\tan^{-1}$ は $\tan$ の逆関数であるが、これが主値を表すと解釈すると値の範囲が $(-\pi/2, \pi/2)$ と幅 π に制限されてしまう。この式だけでは角度 π の差は無視されることになる。そもそも $(x', y') = -(x, y)$ として定義した (x', y') は (x, y) とは角度 φ が π 異なるはずであるが、$y/x = y'/x'$ であるから $\tan^{-1}$ を施す以前に角度 π の違いが消えてしまう。それ以外の情報 (x, y の符号など) から再生する手続きが必要になる。

$$\varphi \equiv \tan^{-1}\left(\frac{y}{x}\right) \pmod{\pi}$$

は正しい式であるのだが、これだけでは不十分であろう。

蛇足的注意　プログラミング言語の C や Fortran には角度を計算するための関数 `atan2()` がある[1]。`atan2(y,x)` とすると、点 (x, y) の角度を $[-\pi, \pi]$ の範囲で返す。そこで C プログラムで (x, y) から (r, φ) を計算するには

```
r = sqrt(x * x + y * y);
phi = atan2(y, x);
if (phi < 0.0) phi += 2 * pi;
```

とするとよい (もちろん `pi` には π の値が入っているとしている)。関数 `atan2()` を使わずに `atan()` のみで角度 φ を求めようとすると、分かりにくいプログラムになる。■

なお、

平面極座標のヤコビアン

$x = r\cos\varphi$, $y = r\sin\varphi$ により $f\colon (r, \varphi) \mapsto (x, y)$ を定義すると

$$\det f = \frac{\partial(x, y)}{\partial(r, \varphi)} = r.$$

H.2　空間極座標

空間に直交する座標軸 x 軸、y 軸, z 軸を取って座標を入れる xyz 座標系で (x, y, z) という座標を持つ点 P の原点からの距離を r, 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、$\overrightarrow{OP'}$ が x 軸の正方向となす角を反時計回りに計った角度を φ $(0 \le \varphi < 2\pi)$ とすると

$$\begin{cases} x = r\sin\theta\cos\varphi \\ y = r\sin\theta\sin\varphi \\ z = r\cos\theta \end{cases}$$

が成り立つ。

写像

$$f\colon [0, \infty) \times [0, \pi] \times [0, 2\pi) \ni (r, \theta, \varphi) \longmapsto (x, y, z) \in \mathbf{R}^3$$

は C^∞- 級で、定義域を $\Omega \overset{\text{def.}}{=} (0, \infty) \times (0, \pi) \times [0, 2\pi)$ に制限すれば 1 対 1 である。特に

$$f|_\Omega\colon (0, \infty) \times (0, \pi) \times [0, 2\pi) \ni (r, \theta, \varphi) \longmapsto f(r, \theta, \varphi) \in \mathbf{R}^3 \setminus \{(0, 0, z); z \in \mathbf{R}\}$$

は全単射である。

逆の計算、つまり (x, y, z) から (r, θ, φ) を求めるには、r, θ は

$$\begin{aligned} r &= \sqrt{x^2 + y^2 + z^2}, \\ \theta &= \operatorname{Arccos}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) \quad (\text{ただし } (x, y, z) \ne (0, 0, 0)) \end{aligned}$$

と x, y, z の式として簡単に表される。φ の方は前と同様、強いて書けば

$$(x, y) \ne (0, 0) \text{ のとき}\quad \varphi = \arg(x, y) \in [0, 2\pi).$$

[1] `atan2()` は、Fortran なら組み込み関数、C ならばライブラリィ関数である。

なお、

空間極座標のヤコビアン

$x = r\sin\theta\cos\varphi$, $y = r\sin\theta\sin\varphi$, $z = r\cos\theta$ により $f\colon (r,\theta,\varphi) \mapsto (x,y,z)$ を定義すると

$$\det f = \frac{\partial(x,y,z)}{\partial(r,\theta,\varphi)} = r^2\sin\theta.$$

H.3 一般の R^n における極座標

$$\text{(H.1)}\qquad \begin{cases} x_1 &= r\cos\theta_1 \\ x_2 &= r\sin\theta_1\cos\theta_2 \\ & \cdots \\ x_i &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{i-1}\cos\theta_i \quad (2\le i\le n-1) \\ & \cdots \\ x_{n-1} &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\cos\theta_{n-1} \\ x_n &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{n-2}\sin\theta_{n-1} \end{cases}$$

ただし $(r,\theta_1,\theta_2,\cdots,\theta_{n-1})$ は次の条件式で定義される $\mathbf{R}^n$ の部分集合 D を動くとする:

$$\text{(H.2)}\qquad \begin{array}{l} r \ge 0, \\ \theta_1,\theta_2,\cdots,\theta_{n-2} \in [0,\pi], \\ \theta_{n-1} \in [0,2\pi) \end{array}$$

注意 H.3.1 $n=3$ のとき、既に紹介した空間極座標

$$\text{(H.3)}\qquad \begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \qquad (r\ge 0,\ \theta\in[0,\pi],\ \varphi\in[0,2\pi)) \\ z &= r\cos\theta \end{cases}$$

とは異なる (座標の順番が異なるだけで、本質的には同じものであるが)。混乱しないこと。

写像

$$\varphi_n\colon D \ni (r,\theta_1,\theta_2,\cdots,\theta_{n-1}) \longmapsto (x_1,x_2,\cdots,x_n) \in \mathbf{R}^n$$

のヤコビアンについては、次の定理が成り立つ。

定理 H.3.1 (極座標のヤコビアン)

$$\text{(H.4)}\qquad \det\varphi_n' = \frac{\partial(x_1,x_2,\cdots,x_n)}{\partial(r,\theta_1,\cdots\theta_{n-1})} = r^{n-1}\sin^{n-2}\theta_1\sin^{n-3}\theta_2\cdots\sin\theta_{n-2}.$$

この証明を与えるかわりに、そのヒントとなる問題を掲げておく。

問題 $\begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} \in \mathbf{R}^4$ に対して、次の式を満たす $\begin{pmatrix} r \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix}$ を 4 次元極座標と呼ぶ:

$$\begin{cases} x &= r\cos\theta_1 \\ y &= r\sin\theta_1\cos\theta_2 \\ z &= r\sin\theta_1\sin\theta_2\cos\theta_3 \\ w &= r\sin\theta_1\sin\theta_2\sin\theta_3 \end{cases}, \quad \begin{pmatrix} r \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \in I \overset{\text{def.}}{=} [0,\infty)\times[0,\pi]\times[0,\pi]\times[0,2\pi).$$

(1) $\begin{cases} X &= r\cos\theta_1 \\ Y &= r\sin\theta_1\cos\theta_2 \\ u &= r\sin\theta_1\sin\theta_2 \\ v &= \theta_3 \end{cases}$ とすると、写像 $\begin{pmatrix} r \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \mapsto \begin{pmatrix} X \\ Y \\ u \\ v \end{pmatrix}$ のヤコビアンは $r^2\sin\theta_1$ であることを示せ。

(2) $\begin{cases} x &= X \\ y &= Y \\ z &= u\cos v \\ w &= u\sin v \end{cases}$ とするとき、写像 $\begin{pmatrix} X \\ Y \\ u \\ v \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ のヤコビアンを求め、$\begin{pmatrix} r \\ \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}$ のヤコビアンが $r^3\sin^2\theta_1\sin\theta_2$ であることを示せ。

(3) 4 次元単位球 $\{(x,y,z,w); x^2+y^2+z^2+w^2<1\}$ の 4 次元 Jordan 測度を求めよ。

H.4 Laplacian の極座標表示

$\mathbf{R}^2$, $\mathbf{R}^3$ における Laplacian $\triangle$ の極座標表示を書いておく。

例 H.4.1 (平面極座標の Laplacian) 平面の極座標

$$\begin{cases} x &= r\cos\varphi \\ y &= r\sin\varphi \end{cases}$$

によって、平面の Laplacian

$$\triangle f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

を表示すると

$$\triangle f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \varphi^2}$$

となる。証明は本文中に問題として出しておいたので、ここでは略する。■

例 H.4.2 (空間極座標の Laplacian) 空間の極座標

$$\begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{cases}$$

によって、空間の Laplacian

$$\triangle f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

を表示すると

$$\begin{aligned} \triangle f &= \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}\right] \\ &= \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\left(\frac{\partial^2 f}{\partial\theta^2} + \frac{1}{\tan\theta}\frac{\partial f}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2 f}{\partial\varphi^2}\right) \end{aligned}$$

となる。

証明　すべて連鎖律で計算していくのは、とても大変なので、少し工夫をする (解析概論 I の範囲を出てしまうが)。まず

$\mathbf{R}^3$ の極座標による grad の内積の公式

$\mathbf{R}^3$ の領域上定義された u, v に対して、

(H.5)
$$\nabla u \cdot \nabla v = u_r v_r + \frac{1}{r^2}u_\theta v_\theta + \frac{1}{r^2\sin^2\theta}u_\varphi v_\varphi$$

を示す。連鎖律より

$$(u_r, u_\theta, u_\varphi) = (u_x, u_y, u_z)\begin{pmatrix} x_r & x_\theta & x_\varphi \\ y_r & y_\theta & y_\varphi \\ z_r & z_\theta & z_\varphi \end{pmatrix}.$$

ここに現れるヤコビ行列を U とおく。具体的には

$$U = \begin{pmatrix} x_r & x_\theta & x_\varphi \\ y_r & y_\theta & y_\varphi \\ z_r & z_\theta & z_\varphi \end{pmatrix} = \begin{pmatrix} \sin\theta\cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

となるが、これを (e_1, e_2, e_3) とおくと、e_i $(i = 1, 2, 3)$ は直交系になる:

$$e_i \cdot e_j = 0 \quad (i \neq j).$$

それゆえ、

$$U^T U = \begin{pmatrix} e_1^T \\ e_2^T \\ e_3^T \end{pmatrix}(e_1 e_2 e_3) = \begin{pmatrix} e_1\cdot e_1 & & \\ & e_2\cdot e_2 & \\ & & e_3\cdot e_3 \end{pmatrix} = \begin{pmatrix} 1 & & \\ & r^2 & \\ & & r^2\sin^2\theta \end{pmatrix}.$$

$$\nabla u = (U^{-1})^T\begin{pmatrix} u_r \\ u_\theta \\ u_\varphi \end{pmatrix}, \quad \nabla v = (U^{-1})^T\begin{pmatrix} v_r \\ v_\theta \\ v_\varphi \end{pmatrix}$$

であるから、

$$\nabla u\cdot\nabla v=(\nabla v)^T\nabla u=\begin{pmatrix}v_r\\ v_\theta\\ v_\varphi\end{pmatrix}^T(U^{-1})^{TT}(U^{-1})^T\begin{pmatrix}u_r\\ u_\theta\\ u_\varphi\end{pmatrix}.$$

ここで

$$(U^{-1})^{TT}(U^{-1})^T=U^{-1}(U^T)^{-1}=(U^TU)^{-1}=\begin{pmatrix}1&&\\&\dfrac{1}{r^2}&\\&&\dfrac{1}{r^2\sin^2\theta}\end{pmatrix}.$$

ゆえに

$$\begin{aligned}\nabla u\cdot\nabla v&=\begin{pmatrix}v_r\\ v_\theta\\ v_\varphi\end{pmatrix}^T\begin{pmatrix}1&&\\&\dfrac{1}{r^2}&\\&&\dfrac{1}{r^2\sin^2\theta}\end{pmatrix}\begin{pmatrix}u_r\\ u_\theta\\ u_\varphi\end{pmatrix}\\&=u_rv_r+\frac{1}{r^2}u_\theta v_\theta+\frac{1}{r^2\sin^2\theta}u_\varphi v_\varphi.\quad\blacksquare\end{aligned}$$

さて、v を $v=0$ (on $\partial\Omega$) を満たす関数とすると、Green の定理から

$$\iiint_\Omega v\,\triangle u\,dxdydz=-\iiint_\Omega\nabla u\cdot\nabla v\,dxdydz.$$

$\nabla u\cdot\nabla v$ を極座標表示した式 (H.5) を代入して

$$\begin{aligned}\iiint_\Omega v\,\triangle u\,dxdydz&=-\iiint_{\widetilde\Omega}(u_rv_r+\frac{1}{r^2}u_\theta v_\theta+\frac{1}{r^2\sin^2\theta}u_\varphi v_\varphi)r^2\sin\theta\,drd\theta d\varphi\\&=-\iiint_{\widetilde\Omega}(r^2\sin\theta u_rv_r+\sin\theta u_\theta v_\theta+\frac{1}{\sin\theta}u_\varphi v_\varphi)\,drd\theta d\varphi\end{aligned}$$

(ただし $\widetilde\Omega$ は、極座標で Ω に対応する領域)。

部分積分を施し、積分変数を元に戻すと

$$\begin{aligned}\iiint_\Omega v\,\triangle u\,dxdydz&=\iiint_{\widetilde\Omega}\left[\frac{\partial}{\partial r}(r^2\sin\theta u_r)+\frac{\partial}{\partial\theta}(\sin\theta u_\theta)+\frac{\partial}{\partial\varphi}(\frac{1}{\sin\theta}u_\varphi)\right]v\,drd\theta d\varphi\\&=\iiint_\Omega\frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}(r^2\sin\theta u_r)+\frac{\partial}{\partial\theta}(\sin\theta u_\theta)+\frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}u_\varphi\right)\right]v\,dxdydz.\end{aligned}$$

v の任意性から (変分法の基本補題から、というべきか)

$$\triangle u=\frac{1}{r^2\sin\theta}\left[\frac{\partial}{\partial r}(r^2\sin\theta u_r)+\frac{\partial}{\partial\theta}(\sin\theta u_\theta)+\frac{\partial}{\partial\varphi}\left(\frac{1}{\sin\theta}u_\varphi\right)\right].$$

整理して

$$\triangle u=\frac{\partial^2u}{\partial r^2}+\frac{2}{r}\frac{\partial u}{\partial r}+\frac{1}{r^2}\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial u}{\partial\theta}\right)+\frac{1}{\sin^2\theta}\frac{\partial^2u}{\partial\varphi^2}\right].\quad\blacksquare$$

$\mathbf{R}^n$ における $\triangle$ の極座標表示については、

島倉紀夫, 楕円型偏微分作用素, 紀伊國屋書店 (1978)

を見ると良い。

付 録I 開集合、閉集合についてのメモ

(昔は、開集合、閉集合にまつわる話 (いわゆる位相空間論) は、単独の講義科目として用意されることも多かったのだが、最近は「必要に応じて (何かのついでに)」説明されることが多い。それも仕方がないと思うが、もし勉強したくなった場合、参考書は豊富にある。古い本だが、

河田 敬義, 三村 征雄; 現代数学概説 II, 岩波書店 (1965).

は辞書的に使える便利な本である[1]。)

I.1 直観的な話 — まとめ

早い話、フチなしが開集合 (自分自身の境界をまったく含んでいないのが開集合)、フチつきが閉集合 (自分自身の境界をすべて含んでいるのが閉集合)。境界というのは「図示したときに境目となるところ」。

もう少していねいに言うと、まず $\mathbf{R}^n$ の任意の部分集合 A があったとき、$\mathbf{R}^n$ は次の 3 つの (交わりのない) 部分に分割される：$\mathbf{R}^n = A^\circ \cup A^b \cup A^e$, ここで

(i) A の内部 $A^\circ = \{x \in \mathbf{R}^n; \exists\varepsilon > 0 \text{ s.t. } B(x;\varepsilon) \subset A\}$.

(ii) A の境界 $A^b = \{x \in \mathbf{R}^n; \forall\varepsilon > 0\ B(x;\varepsilon) \cap A \neq \emptyset \text{ and } B(x;\varepsilon) \cap A^c \neq \emptyset\}$

(iii) A の外部 $A^e = \{x \in \mathbf{R}^n; \exists\varepsilon > 0 \text{ s.t. } B(x;\varepsilon) \subset A^c\}$.

(ただし A^c で A の補集合 $\mathbf{R}^n \setminus A$ を表している。)

以上とは別に、A の閉包があるが、それは A の内部と A の境界を合併したものに等しい:

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

そして

- A が開集合であるとは、A が自分自身の内部に一致すること。
- A が閉集合であるとは、A が自分自身の閉包に一致すること。

一方、境界 A^b の計算は簡単なことが多い。そこから考えると、

- $A^\circ = A \setminus A^b$, $\overline{A} = A \cup A^b$, $A^e = \mathbf{R}^n \setminus \overline{A} = \mathbf{R}^n \setminus (A \cup A^b)$.
- A が開集合 $\Leftrightarrow$ $A^b \cap A = \emptyset$.　(フチなしが開集合)
- A が閉集合 $\Leftrightarrow$ $A^b \subset A$.　(フチつきが開集合)

[1]一方で位相空間論もそれなりに進歩しているという話を聞いたことがある。日本語で読める解説はある？

I.2 開集合

開集合であることを示す

このテキストにおける開集合の定義は、

$$\Omega \text{ が開集合} \overset{\text{def.}}{\Leftrightarrow} \forall x \in \Omega \ \exists \varepsilon > 0 \text{ s.t. } B(x;\varepsilon) \subset \Omega$$

というものだった ($B(x;\varepsilon)$ は x を中心とする半径 ε の開球を表す)。これは比較的素直なものだから、それを直接チェックするという方針で解けることが多いと思われる。それ以外に以下のことに注意するとよい。

- 開集合系の公理
  (1) 空集合 $\emptyset$, 全空間 $\mathbf{R}^n$ は開集合
  (2) 任意の開集合族[2]の合併は開集合
  (3) 有限個の開集合の共通部分は開集合
- 開区間は開集合 (以下で証明する)
- 開球は開集合 (本文で既に証明済)
- A が開集合 $\Leftrightarrow A = A^\circ$ (A° の計算は、たとえば $A = A \setminus A^b$ による).
  あるいは、A が開集合 $\Leftrightarrow A \cap A^b = \emptyset$).

例 I.2.1 (有界開区間は開集合) $a < b$ なる $a \in \mathbf{R}$, $b \in \mathbf{R}$ に対して、 $\Omega = (a,b)$ は $\mathbf{R}$ の開集合である。

証明 $\left((a,b) = B\left(\dfrac{a+b}{2}; \dfrac{b-a}{2}\right)\right.$ であるから、(a,b) は実は $\mathbf{R}$ の開球であり、「開球は開集合」であることは別に証明済み、という論法でも良いが、ここでは直接証明してみよう。) $\forall x \in \Omega$ を取る。

$$\varepsilon \overset{\text{def.}}{=} \min\{b-x, x-a\}$$

とおくと

- $\varepsilon > 0$ ($\because a < x < b$ より、$b - x > 0$ かつ $x - a > 0$ だから)
- $B(x;\varepsilon) \subset \Omega$. 実際、

  $$a - (x-\varepsilon) = a - x + \varepsilon < a - x + (x-a) = 0 \quad \therefore a < x - \varepsilon,$$

  かつ

  $$b - (x+\varepsilon) = b - x - \varepsilon > b - x - (b-x) = 0 \quad \therefore x + \varepsilon < b$$

  であるから、$(x-\varepsilon, x+\varepsilon) \subset (a,b)$. すなわち $B(x;\varepsilon) \subset \Omega$.

ゆえに Ω は開集合である。 ■

[2]集合、族、類、これらは同じ意味であることが多い。つまり「集合族」=「集合の集合」。しかし「集合の集合」と言うのは、あまり気分が良くないので「集合族」と言う。なお、集合については、付録 D を参照せよ。

問 $c \in \mathbf{R}$ とするとき、$(-\infty, c)$, (c, ∞) はともに $\mathbf{R}$ の開集合である。これを証明せよ。

もちろん $(-\infty, \infty) = \mathbf{R}$ も $\mathbf{R}$ の開集合である (全空間は開集合!)。以上まとめると、

R の任意の開区間は開集合である。

例 I.2.2 (一点の補集合は開集合) $A = \{x \in \mathbf{R}; x \neq 0\}$ は $\mathbf{R}$ の開集合である。

証明 $A = (-\infty, 0) \cup (0, \infty)$ と、A は二つの開集合 ($\because$ 開区間は開集合) の合併になるから、開集合である。■

多次元の例を一つあげておこう。

例 I.2.3 $A = \{(x, y); x > 0,\ y > 0\}$ は $\mathbf{R}^2$ の開集合である。

証明 任意の $(x, y) \in A$ に対して、$\varepsilon \overset{\text{def.}}{=} \min\{x, y\}$ とおくと、$\varepsilon > 0$ かつ $B((x, y); \varepsilon) \subset A$ となる。ゆえに A は開集合である。■

少し抽象的な例もあげておこう。

例 I.2.4 (真不等式で定義される集合は開集合) $f \colon \mathbf{R}^n \to \mathbf{R}$ が連続関数であるとき、

$$\Omega \overset{\text{def.}}{=} \{x \in \mathbf{R}^n; f(x) > 0\}$$

は $\mathbf{R}^n$ の開集合である。

証明 $\forall a \in \Omega$ を取る。Ω の定義から $f(a) > 0$ である。連続性から

$$\exists \delta > 0\ \forall x \in B(a; \delta)\ |f(x) - f(a)| < \frac{f(a)}{2}.$$

これから

$$-\frac{f(a)}{2} < f(x) - f(a) < \frac{f(a)}{2}, \quad \therefore f(x) > \frac{f(a)}{2} > 0.$$

よって $x \in \Omega$. これは $B(a; \delta) \subset \Omega$ を示している。ゆえに Ω は開集合である。■

同様に $\{x \in \mathbf{R}^n; f(x) > a\}$, $\{x \in \mathbf{R}^n; f(x) < b\}$, $\{x \in \mathbf{R}^n; a < f(x) < b\}$ なども $\mathbf{R}^n$ の開集合であることが分かる (最初は $f - a$ を新たに f とおき、次は $b - f$ を新たに f とおく。最後はそれらの共通部分だから開集合。)。これから、開区間、開球、一点の補集合などが開集合であることの別証明が簡単に得られる。さらに、

例 I.2.5 $A = (0, 1) \times (3, 4)$ は $\mathbf{R}^2$ の開集合である。

証明 $A = B \cap C$, $B = \{(x, y); 0 < x < 1\}$, $C = \{(x, y); 3 < y < 4\}$ で、B, C は開集合だから A も開集合。■

開集合でないことを示す

「Ω が開集合でないことを示せ」という問題をどう解くか？定義の条件を否定すると、

$$\Omega \text{ が開集合でない} \overset{\text{def.}}{\Leftrightarrow} \exists x \in \Omega\ \forall \varepsilon > 0\ B(x;\varepsilon) \not\subset \Omega.$$

が得られる。

例 I.2.6 半開区間 $\Omega \overset{\text{def.}}{=} [0,1)$ は $\mathbf{R}$ の開集合ではないことを示せ。

証明　$0 \in \Omega$ であるが、

$$\forall \varepsilon > 0 \quad B(0;\varepsilon) \not\subset \Omega$$

である。($\because$ $y = -\varepsilon/2$ とおくと、

$$-\varepsilon < y < +\varepsilon, \quad \therefore y \in (-\varepsilon, \varepsilon) = B(0;\varepsilon),$$

ところが、 $y \notin [0,1) = \Omega$. ゆえに $B(0;\varepsilon) \not\subset \Omega$.) よって Ω は開集合ではない。 ■

I.3　閉集合

閉集合であることを示す

この授業での閉集合の定義は、

$$A \text{ が閉集合} \overset{\text{def.}}{\Leftrightarrow} A^c \text{ が開集合}$$

というもので、少し間接的である。A の補集合 $A^c = \mathbf{R}^n \setminus A$ が複雑な分かりにくい集合である場合など、これを直接チェックするという方針では、面倒になるかもしれない。そこで、閉集合の点列による特徴づけ:

$$A \text{ が閉集合} \Leftrightarrow A \text{ 内の任意の収束点列の極限が } A \text{ に属する}$$

などに慣れておくことを勧める。もちろん、

- 閉集合系の公理
  (1) 空集合 $\emptyset$, 全空間 $\mathbf{R}^n$ は閉集合
  (2) 任意の閉集合族の共通部分は閉集合
  (3) 有限個の閉集合の合併は閉集合
- 閉区間は閉集合
- 閉球は閉集合
- 1 点 a だけからなる集合 $\{a\}$ は閉集合

なども適宜使う。

例 I.3.1 $a \in \mathbf{R}$ とするとき、$\{a\}$ は $\mathbf{R}$ の閉集合である。

証明　$\mathbf{R}\setminus\{a\}=(-\infty,a)\cup(a,\infty)$ で、 $(-\infty,a)$, (a,∞) は開区間ゆえ、開集合。ゆえにその合併である $\mathbf{R}\setminus\{a\}$ は開集合。ゆえに $\{a\}$ は $\mathbf{R}$ の閉集合。■

例 I.3.2 $a<b$ なる $a\in\mathbf{R}, b\in\mathbf{R}$ に対して、$[a,b]$ は $\mathbf{R}$ の閉集合である。

証明　上と同様に、補集合を考えても簡単。ここでは「閉集合の点列による特徴づけ」を用いて証明しよう。$\{x_n\}_{n\in\mathbf{N}}$ を A 内の点列で、$\mathbf{R}$ で $\lim_{n\to\infty} x_n = x_\infty$ と収束するものとする。$a\le x_n\le b$ から $a\le x_\infty\le b$ すなわち $x_\infty\in A$ が導かれる。ゆえに A は $\mathbf{R}$ の閉集合である。■

問　$a\in\mathbf{R}^n$ とするとき、$\{a\}$ は $\mathbf{R}^n$ の閉集合である。これを証明せよ。

例 I.3.3 $f\colon\mathbf{R}^n\to\mathbf{R}$ が連続関数であるとき、

$$A \overset{\text{def.}}{=} \{x\in\mathbf{R}^n; f(x)\le 0\}$$

は $\mathbf{R}^n$ の閉集合である。

証明 1　$\mathbf{R}^n\setminus A=\{x\in\mathbf{R}^n; f(x)>0\}$ は前節で示したように $\mathbf{R}^n$ の開集合である。ゆえに A は閉集合である。■

証明 2　$\{x_n\}_{n\in\mathbf{N}}$ を A 内の任意の収束列とする。すなわち、

$$x_n\in A \quad (n\in\mathbf{N}), \qquad \lim_{n\to\infty} x_n = \exists x_\infty \quad (\text{in } \mathbf{R}^n).$$

さて、定義から

$$f(x_n)\ge 0 \quad (n\in\mathbf{N}),$$

これと f の連続性から

$$f(x_\infty)=f(\lim_{n\to\infty} x_n)=\lim_{n\to\infty} f(x_n)\ge 0.$$

よって $x_\infty\in A.$ 従って A は閉集合である。■

この例は覚えておくと便利である。例えば

例 I.3.4 $\mathbf{R}^n$ の閉球 $\overline{B}(a;r)$ は閉集合である。

証明　$f\colon\mathbf{R}^n\to\mathbf{R}$ を $f(x)=\|x-a\|^2-R^2=(x_1-a_1)^2+(x_2-a_2)^2+\cdots+(x_n-a_n)^2-R^2$ で定義すると、これは連続関数である。そして $\overline{B}(a;R)=\{x\in\mathbf{R}^n; f(x)\le 0\}$ であるから、$\overline{B}(a;R)$ は閉集合。■

例 I.3.5 $A=\{0\}\bigcup\{1/n; n\in\mathbf{N}\}$ は $\mathbf{R}$ の閉集合である。

証明

$$A^c = (-\infty, 0) \cup \left(\bigcup_{n=1}^{\infty} \left(\frac{1}{n+1}, \frac{1}{n} \right) \right) \cup (1, \infty)$$

のように A^c は開集合の和集合として表されるから開集合である。ゆえに A は閉集合である。

■

閉集合でないことを示す

A の補集合が開集合でないことを示すか、前小節の閉集合の点列による特徴づけを使う。

例 I.3.6 $A = \{1/n; n \in \mathbf{N}\}$ は閉集合でないことを示せ。

証明 $x_n = 1/n$ とおくと、$\{x_n\}_{n \in \mathbf{N}}$ は A 内の点列で、$\mathbf{R}$ で

$$\lim_{n \to \infty} x_n = 0$$

と収束する。ところがその極限 0 は A に属さない。ゆえに A は閉集合ではない。 ■

I.4 開集合の連続関数による逆像は開集合

一般の位相空間論を学ぶと「開集合の連続関数による逆像は開集合」という定理がある。上で紹介した、「連続関数 $f\colon \mathbf{R}^n \to \mathbf{R}$ に対して $\{x \in \mathbf{R}^n; f(x) > 0\}$ は $\mathbf{R}^n$ の開集合」というのも、実はこの定理の系とみなせる。この定理を本文で説明したかったのであるが、**部分位相** (**相対位相**) という概念を学ばないと、使い方を間違える危険が大きいので、あきらめた。ここにその一端を紹介しておく (使うときは、下線部に注意して使うこと)。

> <u>定義域が開集合である</u> 連続関数による開集合の逆像は開集合である。

次に示すのは、ある年度の期末試験の問題と、その解説文である。

例題 I.4.1 U を $\mathbf{R}^n$ の開集合、$f\colon U \to \mathbf{R}^m$ を連続関数とする。このとき $\mathbf{R}^m$ の任意の開集合 W に対して $f^{-1}(W) \overset{\text{def.}}{=} \{x \in U; f(x) \in W\}$ 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 $B(\alpha; r)$ 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。」

解 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) $f^{-1}(W)$. ■

付 録J　数列の収束についての補足

数列については、「基礎数学」で十分学んだはずなので、すべてを繰り返し説明することはしない。分からないことがあったら、当時のテキストやノートを復習すること。

J.1　アルキメデスの公理

高等学校以来 $\lim_{n\to\infty}\frac{1}{n}=0$ は当り前のように使っているが、これをきちんと証明しようとするとどうなるか?、を考えて見よう。ここに書いてあることは、数学を使う立場からは「当り前」のことであるが、実数を定義するところから話を始めると、ここは要所である。

公理 J.1.1 (アルキメデスの公理) $a,\ b$ を任意の正数とするとき、適当な自然数 N をとれば、

$$Na>b$$

となる。 — 「どんな小さな正数でも、十分たくさん集めれば、大きな数を追い抜ける」

例題 J.1.1 数列 $\left\{\frac{1}{n}\right\}_{n\in\mathbf{N}}$ が 0 に収束することを示せ。

解答. ε を任意の正数とする。アルキメデスの公理から、

$$\exists N\in\mathbf{N}\quad s.t.\quad N\varepsilon>1.$$

すると $n\ge N$ なる任意の $n\in\mathbf{N}$ に対して $1/n\le 1/N<\varepsilon.$ これから

$$0<\frac{1}{n}<\varepsilon\quad ゆえに\quad \left|\frac{1}{n}-0\right|<\varepsilon.$$

これは $\lim_{n\to\infty}\frac{1}{n}=0$ を示している。 ■

J.2　はさみうちの原理

命題 J.2.1 (はさみうちの原理) 数列 $\{a_n\}_{n\in\mathbf{N}}$, $\{b_n\}_{n\in\mathbf{N}}$, $\{c_n\}_{n\in\mathbf{N}}$ について、二条件

(i) $a_n \le b_n \le c_n$ $(n\in\mathbf{N})$.

(ii) ある $\alpha\in\mathbf{R}$ が存在して $\displaystyle\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n = \alpha$.

が成り立つならば、$\{b_n\}_{n\in\mathbf{N}}$ も収束して、その極限は α である。

例題 J.2.1 数列 $\left\{\dfrac{1}{n}\sin n\right\}_{n\in\mathbf{N}}$ が 0 に収束することを示せ。

解答　$-1\le \sin n\le 1$ であるから、

$$-\frac{1}{n}\le\frac{1}{n}\sin n\le\frac{1}{n}\quad(n\in\mathbf{N}).$$

ところで $\displaystyle\lim_{n\to\infty}\left(-\frac{1}{n}\right)=\lim_{n\to\infty}\frac{1}{n}=0$ であるから、$\left\{\dfrac{1}{n}\sin n\right\}_{n\in\mathbf{N}}$ も収束し、極限は 0. ■

上の議論を次のように書いてすますことが多い。

$$\left|\frac{1}{n}\sin n\right|\le\frac{1}{n}\to 0\quad(n\to\infty)\quad\text{より}\quad\frac{1}{n}\sin n\to 0\quad(n\to\infty).$$

J.3　有界単調列の収束

次の定理 (実は定理 F.2.3) は、極限がすぐには分からない数列に対しても有効なことが多く、役に立つ。

定理 J.3.1 上に有界な単調増加数列は収束する。

例題 J.3.1 次の式で定義される数列 $\{a_n\}_{n\in\mathbf{N}}$ が収束することを示せ。

$$a_n = 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!}=\sum_{k=0}^{n}\frac{1}{k!}.$$

解答　まず $\{a_n\}_{n\in\mathbf{N}}$ が単調増加であることは明らかである。一方、

$$\begin{aligned}
\frac{1}{2!} &= \frac{1}{2^1},\\
\frac{1}{3!} &= \frac{1}{3}\cdot\frac{1}{2!}<\frac{1}{2}\cdot\frac{1}{2^1}=\frac{1}{2^2},\\
\frac{1}{4!} &= \frac{1}{4}\cdot\frac{1}{3!}<\frac{1}{2}\cdot\frac{1}{2^2}=\frac{1}{2^3},\\
\cdots &\ \cdots\ \cdots\\
\frac{1}{n!} &< \frac{1}{2^{n-1}},
\end{aligned}$$

より

$$\begin{aligned} a_n &= 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{n!} \\ &< 1+1+\frac{1}{2^1}+\frac{1}{2^2}+\cdot+\frac{1}{2^{n-1}} \ = 1+\frac{1-\left(\dfrac{1}{2}\right)^n}{1-\dfrac{1}{2}} < 1+\frac{1}{1/2} = 3. \end{aligned}$$

であるから、$\{a_n\}_{n\in\mathbf{N}}$ は有界でもある。ゆえに $\{a_n\}_{n\in\mathbf{N}}$ は収束する。 ■

もちろん、よく知られているように $\{a_n\}_{n\in\mathbf{N}}$ の極限は自然対数の底である:

$$\lim_{n\to\infty} a_n = e = 2.7182818284590452\cdots$$

付 録K　1 変数の平均値の定理、Taylor の定理

1 変数実数値関数に関する平均値の定理、Taylor の定理を復習しよう。特に平均値の定理はどういうことに使われるかを学んでもらいたい。

K.1　平均値の定理の復習

高校の数学の微分法で、イメージで納得していたような「よく知られた事実」をきちんと証明しようとすると、平均値の定理のお世話になることが多い。

微分係数の定義より、直感的に明らかに、十分小さな $|h|$ に対して

$$f'(x) \fallingdotseq \frac{f(x+h)-f(x)}{h}$$

であるが、大体等しい $\fallingdotseq$ では厳密な論証に使えない。

定義 K.1.1 (極大値、極小値、極値) I を $\mathbf{R}$ の区間、 $c \in I$, $f\colon I \to \mathbf{R}$ とする。$f(c)$ が、x が c に十分近い範囲での f の最大値になっているとき[a]、$f(c)$ は f の**極大値**(きょくだいち)である、f は c で**極大**である、という。同様に **極小値**(きょくしょうち), **極小**が定義される。極大値、極小値をあわせて **極値**と呼ぶ。f が c で極値を取るとき、 c を**極値点**と呼ぶ。

[a]十分小さな $\varepsilon > 0$ を取ると、f の $I \cap (c-\varepsilon, c+\varepsilon)$ における最大値が $f(c)$ である、ということ。

定理 K.1.1 (内点で極値を取れば、微分係数 $= 0$) I を $\mathbf{R}$ の区間、$f\colon I \to \mathbf{R}$, c は I の内点, $f(c)$ は f の極値、f は c で微分可能 $\Longrightarrow$ $f'(c) = 0$.

証明　f が c で極大になる場合を考える (極小になる場合も同様)。極大の定義から、$|h|$ が十分小さいならば、

$$f(c+h) - f(c) \le 0.$$

まず $h > 0$ の場合を考えると、

$$\frac{f(c+h)-f(c)}{h} \le 0.$$

ここで $h \downarrow 0$ として[1] $f'(c) \le 0$ が得られる。一方 $h < 0$ の場合を考えると、

$$\frac{f(c+h)-f(c)}{h} \ge 0.$$

[1]h を正の方から 0 に近付けることを $h \downarrow 0$ と書く。$h \to +0$ と同じこと。同様に $h \to -0$ を $h \uparrow 0$ とも書く。

ここで $h \uparrow 0$ として $f'(c) \geq 0$ が得られる。ゆえに $f'(c) = 0$. ■

定理 K.1.2 (Rolle の定理) $f\colon [a,b] \to \mathbf{R}$ は連続で、(a,b) で微分可能, $f(a) = f(b) \Longrightarrow \exists c \in (a,b)\ s.t.\ f'(c) = 0$.

証明 $I = [a,b]$ は $\mathbf{R}$ の有界閉集合であるから、I 上の連続関数である f は最大値、最小値を持つ。最大値 $=$ 最小値の場合、f は定数関数であるから、$c = \dfrac{a+b}{2}$ とおけば $f'(c) = 0$, $a < c < b$. 最大値 $>$ 最小値の場合、少なくとも一方は $f(a) = f(b)$ に等しくない。すると、ある内点 c で、f は最大値かまたは最小値を取ることになる。定理 K.1.1 によって、$f'(c) = 0$. ■

定理 K.1.3 (平均値の定理 (Mean value theorem)) $f\colon [a,b] \to \mathbf{R}$ は連続で、(a,b) で微分可能[a] $\Longrightarrow \exists c \in (a,b)\ s.t.$

$$\frac{f(b)-f(a)}{b-a} = f'(c).$$

[a]区間の端点 a, b での微分可能性は不要であるから、「$[a,b]$ で連続、 (a,b) で微分可能」とした。もちろん f が区間 $[a,b]$ で微分可能ならば $[a,b]$ で連続となるから、「f が $[a,b]$ で微分可能ならば $\exists c \in (a,b)$ s.t. $(f(b)-f(a))/(b-a) = f'(c)$.」は成り立つ。

証明

$$g(x) \overset{\text{def.}}{=} f(x) - \frac{f(b)-f(a)}{b-a}(x-a)$$

とおくと

$$\begin{cases} g\colon [a,b] \to \mathbf{R} \quad \text{連続}, \\ g \text{ は } (a,b) \text{ で微分可能}, \\ g(a) = g(b) \end{cases}$$

となり, g は Rolle の定理の仮定を満たす。よって、$\exists c \in (a,b)$ s.t. $g'(c) = 0$. ところで、

$$g'(c) = f'(c) - \frac{f(b)-f(a)}{b-a}$$

であるから、

$$\frac{f(b)-f(a)}{b-a} = f'(c).$$

■

注意 K.1.1 *(1)* この定理は、いろいろな表し方がある。例えば、

$$\exists \theta \in (0,1)\ s.t.\ \frac{f(b)-f(a)}{b-a} = f'(a + \theta(b-a))$$

のように c のかわりに θ で表現したり、

$$\exists c \in (a,b)\ s.t.\ f(b) - f(a) = f'(c)(b-a)$$

のように分母を払った形にしたりする。また $b-a=h$ として、

$$\exists\theta\in(0,1)\ s.t.\ f(a+h)=f(a)+f'(a+\theta h)h$$

とした形は、$h<0$ としても成り立ち、よく使われる。

(2) 平均値の定理はいわゆる「**存在定理**」であって、c の存在は主張するが、c の値については、a と b の間にあるという以外に何の情報も与えていない。

定理 K.1.4 (微分がいたるところ正ならば狭義単調増加) $f\colon[a,b]\to\mathbf{R}$ が連続で、f は (a,b) で微分可能、 $f'>0\ in\ (a,b)$ とするとき f は $[a,b]$ で狭義の単調増加である。すなわち

$$a\le x_1<x_2\le b\Longrightarrow f(x_1)<f(x_2).$$

証明 平均値の定理から $f(x_2)-f(x_1)=f'(c)(x_2-x_1)$ を満たす $c\in(x_1,x_2)$ が存在する。 仮定より $f'(c)>0,\ x_2-x_1>0$ であるから、 $f(x_2)>f(x_1)$. ■

定理 K.1.5 (微分がいたるところ非負ならば単調増加) $f'\ge 0$ の場合は広義の単調増加となる。すなわち

$$a\le x_1<x_2\le b\Longrightarrow f(x_1)\le f(x_2).$$

次の「微分法を習った人なら誰でも知っている (が、案外証明を知っている人は少ない)」定理も平均値の定理で証明できる。

定理 K.1.6 (微分がいたるところ 0 ならば定数) $f\colon[a,b]\to\mathbf{R}$ が連続で、f は (a,b) で微分可能、$f'\equiv 0\ in\ (a,b)$ とするとき f は $[a,b]$ で定数となる。

次の命題は、知っていると便利である。

命題 K.1.1 $\varepsilon>0,\ I=(a-\varepsilon,a+\varepsilon),\ f\colon I\to\mathbf{R}$ 連続、a 以外で微分可能、

$$\lim_{\substack{x\to a\\ x\ne a}}f'(x)=\ell$$

$\Longrightarrow f$ は a で微分可能で $f'(a)=\ell$.

証明 $x\in I\setminus\{a\}$ に対して、a と x の間にある数 c で

$$\frac{f(x)-f(a)}{x-a}=f'(c)$$

となるものが存在する。$x\to a$ とすると、$c\to a$ となり、上式の右辺 $\to\ell$ となるから、

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}=\ell.$$

すなわち $f'(a)$ が存在して ℓ に等しい。 ■

注意 K.1.2 平均値の定理はベクトル値関数では成立しない。例えば $f(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ $(t \in [0, 2\pi])$ とすると、$f(0) = f(2\pi)$, $f'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ であるから、 $f(2\pi) - f(0) = f'(c)(2\pi - 0)$ を満たす c は存在しない (左辺は 0 だが、右辺は 0 にならない)。しかし、次の定理は多次元でも成立するので、あまり困ることはない。

定理 K.1.7 (有限増分の公式) $f\colon [a,b] \to \mathbf{R}$ が連続で、f は (a,b) で微分可能、$\sup_{x\in(a,b)} |f'(x)| = M$ とするとき、 $|f(b) - f(a)| \leq M|b-a|$.

証明は平均値の定理からただちに導かれる。

系 K.1.1 $f\colon [a,b] \to \mathbf{R}$ が連続で、f は (a,b) で微分可能、$L \in \mathbf{R}$ とするとき、 $|f(b) - f(a) - L(b-a)| \leq \sup_{x\in(a,b)} |f'(x) - L|\,|b-a|$.

K.2 Taylor の定理

定理 K.2.1 (Taylor の定理) $k \in \mathbf{N}$, I を $\mathbf{R}$ の区間、 $f\colon I \to \mathbf{R}$ を k 回微分可能な関数, $a \in I$, $x \in I \Longrightarrow$

(K.1) $f(x) = f(a) + \dfrac{f'(a)}{1!}(x-a) + \dfrac{f''(a)}{2!}(x-a)^2 + \cdots + \dfrac{f^{(k-1)}(a)}{(k-1)!}(x-a)^{k-1} + R_k$

によって R_k を定義するとき、

$$R_k = \frac{f^{(k)}(c)}{k!}(x-a)^k \tag{K.2}$$

を満たす c が a と x の間に[a]存在する。

[a] つまり $a < x$ ならば $c \in (a,x)$, $x < a$ ならば $c \in (x,a)$, $a = x$ ならば $c = a$.

証明 $a = x$ ならば明らかだから、$a \neq x$ とする。

$$g(t) = -f(x) + \left[\sum_{j=0}^{k-1} \frac{(x-t)^j}{j!} f^{(j)}(t) + \frac{(x-t)^k}{k!}\omega\right]$$

とおく。ただし、定数 ω は $g(a) = 0$ となるように、すなわち

$$f(x) = \sum_{j=0}^{k-1} \frac{(x-a)^j}{j!} f^{(j)}(a) + \frac{(x-a)^k}{k!}\omega$$

が成り立つように定める。$g(x)=0$ であるから、$g(a)=g(x)$ であり、g について Rolle の定理を適用すると、$g'(c)=0$ (c は a と x の間) を満たす c の存在が分かる。ところで (根性で) 計算すると

$$g'(t)=\frac{(x-t)^{k-1}}{(k-1)!}\left[f^{(k)}(t)-\omega\right]$$

となることがわかる。$t=c$ を代入して、$f^{(k)}(c)=\omega$ を得る。 ■

注意 K.2.1 (1) (C.1) の右辺は、$f^{(0)}=f$, $0!=1$ であることに注意すると、

$$\sum_{j=0}^{k-1}\frac{f^{(j)}(a)}{j!}(x-a)^j+R_k$$

と書ける。ただし $h^0\equiv 1$ であるとした[2]。

(2) R_k を k **次剰余項**(じょうよこう)と呼ぶが、特に (K.2) の形で表されたものを **Lagrange の剰余項**という。**気分的には**剰余項は小さな項である。$\dfrac{1}{k!}$ は小さいし、多くの場合は $|x-a|$ が小さいので $|(x-a)^k|$ はとても小さいから (本当は、$|x-a|$ が小さくないときにもこの定理を用いるし、$|f^{(k)}(c)|$ がものすごく大きくなることもありうるわけで、R_k が本当に小さいかどうかはケース・バイ・ケースである)。

(3) $k=1$ ならば平均値の定理に相当する。すなわち Taylor の定理は平均値の定理の一般化である。

(4) $x-a=h$ とおくと (K.1), (K.2) は以下のようになる:

$$f(a+h)=f(a)+f'(a)h+\frac{1}{2}f''(a)h^2+\cdots+\frac{f^{(k-1)}(a)}{(k-1)!}h^{k-1}+R_k,$$

$$R_k=\frac{f^{(k)}(a+\theta h)}{k!}h^k.$$

ここで θ は $0<\theta<1$ を満たす、ある数である。

(5) 応用上重要な多くの場合に、a に十分近い任意の x に対して $\lim_{k\to\infty}R_k=0$ が成り立つ (このとき、f は a の近傍で実解析的であるという)。すなわち

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(a)}{n!}(x-a)^n$$

が成立する。これを f の a を中心とする **Taylor 展開** (**Taylor expansion**) と呼ぶ。特に $a=0$ の場合、すなわち

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n$$

を **Maclaurin**(マクローリン) **展開**と呼ぶ[3]。

[2]正数でない h に対して h^0 は一般には定義されない。例えば 0^0 が定義されないのは有名な話であるが、ここでは 1 であると考える。

[3]Taylor 展開の中心が 0 であるものを Maclaurin 展開と呼ぶ、というのはすっかり普及している用語であるが、歴史的には正しくないのだそうである (このことも良く知られているが、今さら変えられないらしい)。

K.3 凸関数と 2 階導関数

定義 K.3.1 (凸関数) I を $\mathbf{R}$ の区間、$f\colon I \to \mathbf{R}$ とするとき、

$$f \text{ が}\overset{\text{とつかんすう}}{\textbf{凸関数}}\ (\textbf{convex function}) \overset{\text{def.}}{\Leftrightarrow} \begin{cases} \forall a \in I,\ \forall b \in I,\ \forall t \in (0,1) \\ f(ta+(1-t)b) \leq tf(a)+(1-t)f(b). \end{cases}$$

直観的には、グラフが下に凸であるような関数のことを凸関数というわけである。

定理 K.3.1 (2 階導関数の符号と凸性) $\mathbf{R}$ の区間 I で f'' が存在して、$f'' \geq 0$ on I であるとき、以下のことが成立する。

(1) 任意の $a, x \in I$ に対して, $f(x) \geq f(a)+(x-a)f'(a)$.

(2) f は I で凸。

(3) $f'(\alpha)=0$ となる $\alpha \in I$ が存在すれば、α は f の最少点である。

証明

(1) Taylor の定理を $n=2$ として用いると、$f'' \geq 0$ から

$$\begin{aligned} f(x) &= f(a)+f'(a)(x-a)+\frac{f''(c)}{2}(x-a)^2 \\ &\geq f(a)+f'(a)(x-a). \end{aligned}$$

(2) $x=ta+(1-t)b$ とすると、前項より

$$\begin{aligned} f(a) &\geq f(x)+(a-x)f'(x), \\ f(b) &\geq f(x)+(b-x)f'(x) \end{aligned}$$

となる。両式にそれぞれ $t,\, 1-t\ (\geq 0)$ を乗じて、辺々加えると、

$$tf(a)+(1-t)f(b) \geq [t+(1-t)]f(x)+[t(a-x)+(1-t)(b-x)]f'(x).$$

整理すると[4]

$$tf(a)+(1-t)f(b) \geq f(x) = f(ta+(1-t)b).$$

(3) Taylor の定理を $n=2$ として使うと、$\alpha,\, x$ の間に c が取れて、

$$\begin{aligned} f(x) &= f(\alpha)+f'(\alpha)(x-\alpha)+\frac{f''(c)}{2}(x-\alpha)^2 \\ &= f(\alpha)+\frac{f''(c)}{2}(x-\alpha)^2 \geq f(\alpha) \quad (f'(c) \geq 0,\ (x-\alpha)^2 \geq 0) \end{aligned}$$

ゆえに $f(x)$ は $x=\alpha$ で最小となる。 ■

[4] $t(a-x)+(1-t)(b-x) = ta+(1-t)b-[t+(1-t)]x = ta+(1-t)b-x = 0.$

K.4　おまけ: 2 変数関数の極値

(2 変数の場合は、その気になれば線形代数を使わず、高校数学だけで極値の判定定理が理解可能である。こことは独立に一般の n 変数の場合の説明ができるので、論理的にはこの小節は蛇足となってしまうが…)

一般の場合に考える前に 2 変数でやってみよう。$h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ とすると、Taylor の定理から $\exists\theta \in (0,1)$ s.t.

$$\begin{aligned} f(a+h) &= f(a) + f'(a)h + \frac{1}{2}(d^2f)_{a+\theta h}(h) \\ &= f(a) + (f_x(a)h_1 + f_y(a)h_2) + \frac{1}{2}\left(f_{xx}(a+\theta h)h_1^2 + 2f_{xy}(a+\theta h)h_1h_2 + f_{yy}(a+\theta h)h_2^2\right) \end{aligned}$$

となる。$f'(a) = (f_x(a), f_y(a)) = (0,0)$ であれば

$$\begin{aligned} f(a+h) &= f(a) + \frac{1}{2}\left(f_{xx}(a+\theta h)h_1^2 + 2f_{xy}(a+\theta h)h_1h_2 + f_{yy}(a+\theta h)h_2^2\right) \\ &\fallingdotseq f(a) + \frac{1}{2}\left(f_{xx}(a)h_1^2 + 2f_{xy}(a)h_1h_2 + f_{yy}(a)h_2^2\right). \end{aligned}$$

これから

$$Q(h) = Q(h_1,h_2) = ph_1^2 + 2qh_1h_2 + rh_2^2, \quad (p = f_{xx}(a),\, q = f_{xy}(a),\, r = f_{yy}(a))$$

という 2 次形式 (2 次同次多項式) の性質を調べることが鍵であることは想像がつくであろう。例えば

(1) $p=1$, $q=0$, $r=1$ ならば $\forall h \neq 0$ に対して $Q(h) = h_1^2 + h_2^2 > 0$. よって $f(a)$ は極小値。

(2) $p=-1$, $q=0$, $r=-1$ ならば $\forall h \neq 0$ に対して $Q(h) = -h_1^2 - h_2^2 < 0$. よって $f(a)$ は極大値。

(3) $p=1$, $q=0$, $r=-1$ ならば $Q(h) = h_1^2 - h_2^2$ は h によって正にも負にもなりうる。よって $f(a)$ は極値ではない。

$q=0$ でない場合はどうなるか？実は $pr - q^2 \neq 0$ ならば、本質的には上の 3 つの場合のいずれかと同じであり、そのいずれであるかの判定には、「判別式を使えばよい」ことが分かる。きちんと述べると

命題 K.4.1 (2 **変数関数の極大極小の判定**) C^2-級の関数 $f\colon \mathbf{R}^2 \supset \Omega \to \mathbf{R}$ が内点 $a \in \Omega$ において $f_x(a) = f_y(a) = 0$ を満たすとする。 $p = f_{xx}(a)$, $q = f_{xy}(a)$, $r = f_{yy}(a)$, $Q(h) = ph_1^2 + 2qh_1h_2 + rh_2^2$ (ただし $h = (h_1, h_2)$) とおく。

(i) $pr - q^2 > 0$ ならば h についての 2 次形式 $Q(h)$ は $\forall h \neq 0$ に対していつも同じ符号を取り、したがって f は a で極値を取る。より詳しくは

(a) $p > 0$ ならば $Q(h)$ は常に正となるので $f(a)$ は極小値。

(b) $p < 0$ ならば $Q(h)$ は常に負となるので $f(a)$ は極大値。

(ii) $pr - q^2 < 0$ ならば 2 次形式 $Q(h)$ は正にも負にもなりうるので、$f(a)$ は極値ではない。

(iii) $pr - q^2 = 0$ ならば、もっと詳しく調べないと分からない。

このことを証明するのは難しくはない[5]が、一般の場合への拡張を考えて、Hesse 行列の性質で分類した形で書いておこう ($\det H = pr - q^2$ は固有値の積になることに注意しよう)。

命題 K.4.2 (2 **変数関数の極大極小の判定 — Hesse 行列版**) C^2- 級の関数 $f\colon \mathbf{R}^2 \supset \Omega \to \mathbf{R}$ が内点 $a \in \Omega$ において $f_x(a) = f_y(a) = 0$ を満たすとする。 $p = f_{xx}(a)$, $q = f_{xy}(a)$, $r = f_{yy}(a)$, $H = \begin{pmatrix} p & q \\ q & r \end{pmatrix}$ とおくとき

(i-a) H が正定値 (固有値が両方とも正) ならば $\forall h \neq 0$ に対して $Q(h) > 0$. f は a で極小である。

(i-b) H が負定値 (固有値が両方とも負) ならば $\forall h \neq 0$ に対して $Q(h) < 0$. f は a で極大である。

(ii) H が不定符号 (固有値に正のものと負のものが存在する) ならば、h によって $Q(h)$ は正にも負にもなりうる。従って $f(a)$ は極値ではない。

(iii) H が特異 (H の固有値に 0 が存在する) ならば、より詳しく調べないと分からない。

さて、証明 (もどき) をやってみよう。$Q(h) = (Hh, h)$ と書けることがミソである。実際

$$Hh = \begin{pmatrix} ph_1 + qh_2 \\ qh_1 + rh_2 \end{pmatrix}, \quad h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

であるから

$$(Hh, h) = (ph_1 + qh_2)h_1 + (qh_1 + rh_2)h_2 = ph_1^2 + 2qh_1h_2 + rh_2^2 = Q(h).$$

[5] 証明の出出(でだし)の部分を書いておこう。 $h_2 \neq 0$ ならば $Q(h) = {h_2}^2(p\xi^2 + 2q\xi + r)$, $\xi = h_1/h_2$ と変形できる。括弧内を $f(\xi)$ とおくと、f は 1 変数の 2 次関数であり、変数 ξ は $\mathbf{R}$ 全体を動くことになる。判別式を D とすると、 $D/4 = q^2 - pr$. $D < 0$ ならば $f(\xi)$ はいつも同じ符号になる。以下略。 — 確かに高校数学で間に合う。

H は対称行列であるから、適当な直交行列 T によって

$$ {}^tTHT = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} $$

と対角化できる。ここで λ_1, λ_2 は H の固有値である (対称行列の固有値だから実数である!)。$k = {}^tTh$ とおくと $h = Tk$ であるから

$$ Q(h) = (Hh, h) = (HTk, Tk) = ({}^tTHTk, k) = \lambda_1 k_1^2 + \lambda_2 k_2^2. $$

ここで $h \mapsto k$ という対応が 1 対 1 対応で、「$h = 0 \Leftrightarrow k = 0$」であることに注意すると

(i-a) $\lambda_1, \lambda_2 > 0$ ならば $\forall h \neq 0$ に対して $Q(h) > 0$ となる。

(i-b) $\lambda_1, \lambda_2 < 0$ ならば $\forall h \neq 0$ に対して $Q(h) < 0$ となる。

(ii) λ_1 と λ_2 が異符号ならば $Q(h)$ は正にも負にもなる。

(iii) λ_1, λ_2 の少なくとも一方が 0 ならば $Q(h) = 0$ となる $h \neq 0$ が存在する。

ということが分かる。以上大ざっぱではあるが、2 変数関数の場合には一通りの様子が分かった。一般の n 変数関数を扱うための準備として、2 次形式の話をしよう。

付 録L　点と閉集合、閉集合と閉集合の距離

以下は、

L.Schwartz, シュヴァルツ解析学 I, 東京図書 (1970)

による。

命題 L.0.3 (点と閉集合との距離) (X, d) を距離空間、A を X の閉集合、$x \in X$ とするとき、x と A の距離 $d(x, A)$ を

$$d(x, A) \overset{\text{def.}}{=} \inf_{a \in A} d(x, a)$$

で定義するとき、以下の (1) ～ (4) が成り立つ。

(1) $d(x, A) \geq 0$.

(2) $d(x, A) = 0 \Leftrightarrow x \in A$. したがって $d(x, A) > 0 \Leftrightarrow x \notin A$.

(3) $X \ni x \mapsto d(x, A)$ は連続写像。

(4) X の閉球がすべてコンパクトならば、任意の $x \in X$, 任意の閉集合 A に対して、$d(x, A) = d(x, a)$ となる $a \in A$ が存在する。

証明

(1) $\forall a \in A$ に対して $d(x, a) \geq 0$ であるから。

(2)
$$\begin{aligned} d(x, A) = 0 \quad &\Leftrightarrow \forall \varepsilon > 0\ \exists a \in A \text{ s.t. } d(x, a) < \varepsilon. \\ &\Leftrightarrow \forall \varepsilon > 0\ B(x; \varepsilon) \cap A \neq \emptyset. \\ &\Leftrightarrow x \in \overline{A}. \\ &\Leftrightarrow x \in A. \end{aligned}$$

(3) x, y を X の二点とするとき、

$$|d(x, A) - d(y, A)| \leq d(x, y)$$

が成り立つことを示す。$d(x, A)$ の定義から、任意の $\varepsilon > 0$ に対して $\exists a \in A$ s.t.

$$d(x, a) \leq d(x, A) + \varepsilon$$

ゆえに
$$d(y,A) \le d(y,a) \le d(y,x) + d(x,a) \le d(y,x) + d(x,A) + \varepsilon$$
ε の任意性から
$$d(y,A) \le d(y,x) + d(x,A).$$
x と y を入れ替えて同じ議論をすると
$$d(x,A) \le d(y,x) + d(y,A).$$
これから
$$-d(y,x) \le d(y,A) - d(x,A) \le d(y,x),$$
すなわち
$$|d(y,A) - d(x,A)| \le d(y,x).$$
これから $x \mapsto d(x,A)$ は連続であることが分かる。

(4) $d \overset{\text{def.}}{=} d(x,A)$ とおき、x を中心とする半径 $d+1$ の閉球を B とする。仮定によって、B はコンパクトである。$B \cap A$ も (コンパクト距離空間の閉集合なので) コンパクトであるから、連続関数 $a \mapsto d(x,a)$ は $B \cap A$ 上最小値を持つ。最小値を与える点を a^* としよう:
$$a^* \in B \cap A, \quad d(x,a^*) = \inf_{a \in B \cap A} d(x,a).$$
このとき実は a^* は A における最小値を与える (i.e. $d = d(a,a^*)$)。実際、

- $x \in B \cap A$ ならば $d(a,a^*) \le d(a,x)$.
- $x \in A \setminus B$ ならば、$d(a,x) \ge d+1$ であるから、$d(a,a^*) \le d+1 \le d(a,x)$.

よって、すべての $x \in A$ に対して $d(a,a^*) \le d(a,x)$ となり、$d(a,a^*)$ は A における最小値 d である。■

命題 L.0.4 (閉集合と閉集合の距離) (X,d) を距離空間、A, B を X の閉集合とするとき、A, B の距離 $d(A,B)$ を
$$d(A,B) \overset{\text{def.}}{=} \inf_{a \in A, b \in B} d(a,b)$$
で定義する。

(1) 一般には、$A \cap B = \emptyset$ であっても $d(A,B) = 0$ とは限らない。

(2) A, B の少なくとも一方がコンパクトで、$A \cap B = \emptyset$ のときは、$d(A,B) > 0$. さらに X の閉球がすべてコンパクトならば、この下限は最小値である、すなわち $\exists (a,b) \in A \times B$ s.t. $d(A,B) = d(a,b)$.

(1) 平面で、双曲線とその漸近線を考えればよい。共通部分はないが、距離は 0 である。

(2) A がコンパクトであると仮定しよう (B がコンパクトな場合も同様)。

$$d(A,B)=\inf_{a\in A,b\in B}d(a,b)=\inf_{a\in A}\inf_{b\in B}d(a,b)=\inf_{a\in A}d(a,B).$$

$a\mapsto d(a,B)$ は連続関数であり、A がコンパクトであると仮定したから、最右辺の下限は実は最小値である。つまり $\exists a^*\in A$ s.t. $d(A,B)=d(a^*,B)$. $A\cap B=\emptyset$ と仮定したから、$a^*\notin B$ であり、B は閉集合であるから $d(a^*,B)>0$. ゆえに $d(A,B)>0$. X の閉球がすべてコンパクトな場合、命題 L.0.3 の (4) によって、$\exists b^*\in B$ s.t. $d(a^*,B)=d(a^*,b^*)$. ゆえに $d(A,B)=d(a^*,b^*)$. ■

例 L.0.1 X を距離空間、Ω を X の開集合、 K を Ω に含まれるコンパクト集合[1]とすると、$d(K,\Omega^c)=d(K,\Omega^b)>0$. (ただし $\Omega^c=X\setminus\Omega$, Ω^b は Ω の境界を表す。) — この事実は非常にしばしば使われるにもかかわらず、証明が書いてある本が案外少ない。

注意 L.0.1 集合の距離には、ここで紹介したものの他に、Hausdorff 距離と呼ばれるものがある。それについては、例えば

山口昌哉(まさや)・畑政義(はたまさよし)・木上淳(きがみじゅん), フラクタルの数理, 岩波講座 応用数学, 岩波書店 (1993)

を参照せよ。

[1]応用上は関数の台 (support) であることが多い。

付 録M　最近気になっていること

コンピューターの数学への影響が無視できない。既に研究者達は自分の都合に応じて積極的にコンピューターを利用しているが、教育をどのように変える必要があるのか、真剣に考えるべき時期である (もう何年も前から…)。

忙しい中、考えていることは色々あるのだが、まとまった構想はおいそれとは出せないと感じている。以下雑感。

- 高校までの数学では、手計算の範疇という限界の中で、それなりに算法を重視しているが、大学の数学では「算法にこだわらない数学」もあることを学ぶ。それは大学入学者の数学のレベルを考えると自然な段階であると思うが、やり方を間違えて、ともすると算法の軽視をする数学感を植え込んでいないだろうか？
- 現実的に計算で処理できる問題の規模が人手とコンピューターとでは大違いであるが、多くの教科書では人手で扱えるような小規模の問題にしか使えない算法しか説明していない。原理を理解する (証明を理解する) ためには、その種の算法でも十分であるから構わない面もあるが、原理を理解するにも、大規模な計算にも使えるような算法があるのならば、そちらを説明するように乗り換えた方が良くはないか？
- 大学初年級の数学というと、微分積分学と線形代数が二本の柱となっているが、その内容を見ると、数学としては若い線形代数の方に、考えるべき問題があると感じる。脱線になるけれども…
  - 「線形」でない代数も必要ではないか。例えば多項式代数とか[1]。最近の連立代数方程式の解法の説明は無理にしても、何も伝えないで良いのだろうか？
  - 空間図形の幾何学 (直線、平面、球) のような内容が、高校数学から大きく削られた[2]のを埋め合わせる場所は「線型代数」が適当と思われるが、教員・著者の方にその自覚があるか？ (少なくとも書いてないテキストが多くなっているように感じる。)
  - 名前に「代数」とついているせいか、解析的なことは一切省くつもりか、ノルムも何も出てこない。しかし、例えば固有値問題ともなると、実際の計算には反復計算が欠かせないので、解析的な視点は本質的に重要だと思われる。重要性・有用性から考えて、Gerschgorin の円板定理くらいはテキストに載せるのが本当だと思うが、どういう形で載せるか、結構むつかしい… でも、それが難しくなっていることが、逆に現在の線形代数の不健全性を表わしていると思う。

[1]むしろ、古く書かれた本には、ちゃんと書かれているのに、新しいテキストでは省略されていることが多い。

[2]線形計算みたいのをカリキュラムに入れることでコンピューターの時代に合わせたつもりなのだろうか…(高次元) 空間のイメージを思い浮かべ、計算で処理することが出来るという、面白くかつ役に立つ内容を放り投げてどうするのだろう。

– 一松先生の本 ([13], [14] など) を読むと、なるほどと思うことが多いのだが、では [13] を教科書に採用できるかというと、考え込んでしまうのであった。

索 引