

微分積分学・同演習 A 講義ノート（未完成版）*

原 隆
九大数理

hara@math.kyushu-u.ac.jp

Last updated: July 20, 2018

概要

これは上記科目のための講義ノート（講義メモ）です。教科書が少し簡単であるため、講義に関しての補足的な題材も含めてノートにしました。このノートは現在も改訂中ですが、大体、このような感じで進みます。ただし、このノート全般に渡って、授業では触れない（試験でも訊かない）少し高度な材料も載せてあります。なお、偏微分については時間の関係で、中途半端になってしまいました。混乱を避けるため、このノート（7/13 版）では、今学期にカバーする範囲のみを載せています。期末試験の後、来学期にカバーする材料の一部を直し直す予定です。

（受講生以外の方へのお断り）これはあくまで上記科目を受講した学生さんのためのもので、売り物になるくらいの品質で作っている訳ではありません。ところどころ、ミスもあるでしょう。もし、上記科目の受講生以外の方が奇特にも手に取ってくださった場合は、その点を十分了承した上でお使い頂こう、お願いします。

目次

1	極限と連続性	3
1.1	数列の極限： ϵ - N 論法 ¹	3
1.1.1	少しでも理解を助けるために	4
1.1.2	いろいろな例と定義の応用	6
1.2	関数の極限： ϵ - δ 論法 ² （ここは簡単に）	7
1.3	実数の連続性の公理 ³	8
1.4	単調な数列 ⁴	10
1.5	コーシー列 ⁵	11
1.6	連続関数とその性質 ⁶	15
1.7	連続関数の効用：指数関数と対数関数	17
2	微分	18
2.1	微分の定義 ⁷	18
2.2	オーダーの概念（ランダウの記号） ⁸	19
2.3	平均値の定理 ⁹	20
2.3.1	関数の増減	21
2.4	高階導関数 ¹⁰	22

*2018 年度前期、毎週火曜 3 限、基幹教育 1 年 S1-19 クラス（工学部物質科学工学科）用

¹教科書の 1.1 節

²教科書の 1.2 節前半

³教科書の 1.1 節に少しある

⁴教科書 1.1 節に少し

⁵教科書の定理 6.2 など

⁶教科書 1.2 節

⁷教科書 2.1 節前半

⁸教科書 p.31

⁹教科書 2.1 節の後半

¹⁰教科書 2.2 節前半

2.5	テイラーの定理とテイラー展開 ¹¹	23
2.5.1	テイラーの公式（有限項でとめた形）	23
2.5.2	テイラー展開（無限項まで）	25
2.5.3	テイラーの公式，テイラー展開の例	25
2.5.4	テイラーの公式の意味（函数の近似）	26
2.5.5	テイラー展開の効用	28
2.5.6	オイラーの公式	28
2.5.7	おまけ：剰余項が積分の形のテイラーの定理	29
2.6	函数の極値と凹凸 ¹²	30
2.6.1	函数の極大・極小	30
2.6.2	曲線の凹凸	31
3	多変数函数の微分	32
3.1	1変数の函数，多変数の函数 ¹³	32
3.2	偏微分 ¹⁴	34
3.2.1	偏導函数がゼロ，の函数は？	35
3.2.2	方向微分 ¹⁵	35
3.2.3	全微分可能性 ¹⁶	36
3.3	合成函数の微分（連鎖率，chain rule）	39
3.3.1	合成函数の微分（1変数の場合の復習，Case A）	40
3.3.2	合成函数の微分（1変数の場合に帰着，Case B）	40
3.3.3	合成函数の微分（本質的に多変数の場合，Cases C & D）	41
3.3.4	全微分可能性を仮定したときの chain rule	43
3.4	高階の偏導函数	44
3.4.1	2階の偏微分係数の幾何学的意味	46
3.4.2	高階偏導函数と連鎖律	47
3.4.3	（補足）偏導函数がゼロという函数は？ふたたび	47
3.5	平均値の定理	48
3.6	テイラーの定理とテイラー展開	49

¹¹教科書 2.2 節後半

¹²教科書の 2.3 節

¹³教科書の 3.1 節，3.2 節に相当

¹⁴教科書 3.3 節前半

¹⁵この小節の内容は，偏微分に関する理解を深めるための補助的なものである。

¹⁶この小節の内容は「進んだ話題」なので余裕のない人はとばしても良いし，講義でも触れない可能性が高い。

1 極限と連続性

理学・工学系（特に理論系）の人が将来、必要とする程度の、最低限の微積分の基礎、特に極限の概念についてまとめました。このくらいは一度は勉強しておいても悪くはないはず。

1.1 数列の極限： ϵ - N 論法¹⁷

まずは数列の極限を考える。数列の方が関数より簡単なはずだから、まずここで数列の極限（ ϵ - N 論法）に慣れようという狙いである。

皆さんは高校で $\lim_{n \rightarrow \infty} a_n = \alpha$ という式の意味を習ったはずだ。多分、

n が限りなく大きくなるとき、 a_n が限りなく α に近づく

などという「定義」を聞いたのではないか？この定義は特に間違っていないし、これで十分な場合はこれでやれば良い。しかし、この言い方は以下の理由で困ったものである。

- まず、「限りなく近づく」「限りなく大きく」には「限りなく」という感覚的な言葉が入っていて、あやふやだ。
- 次に、「近づく」「大きくなる」などの「動き」が何となく入っており、考えにくい。
- もっと困ったことに、この言い方には「どのくらい速く極限に収束するのか」の**収束の速さ**に関する言及が全くない。そのため、少しややこしい極限——特に2つ以上の変数が混ざった極限¹⁸——を考えだすと、お手上げになる。2つ以上の変数が現れないけど困ってしまう例としては、

(問) $\lim_{n \rightarrow \infty} a_n = 0$ のとき、 $\frac{1}{n} \sum_{k=1}^n a_k$ の極限を求めよ

がある。この答えは直感的には0だろうが、証明できますか？(この答えは後の命題 1.1.7 である)。

これらの欠点を克服すべく、極限への収束の速さまで含めた、定量的な定義が考えられた。これが ϵ - N 論法で、以下のように書かれる。

定義 1.1.1 数列 a_n と実数 α に対して、数列 a_n が $n \rightarrow \infty$ で α に収束する、つまり $\lim_{n \rightarrow \infty} a_n = \alpha$ というのは、以下の (ア) が成り立つことと定義する：

(ア) 任意の (どんなに小さい) 正の数 ϵ に対しても、適当な (大きい) 実数 $N(\epsilon)$ を見つけて、

$$\text{すべての } n > N(\epsilon) \text{ で、 } |a_n - \alpha| < \epsilon \text{ とできる。} \quad (1.1.1)$$

(ア) は以下のように言っても良い。

(アの言い換え) 任意の (どんなに小さい) 正の数 ϵ に対しても、

$$\text{すべての } n > N(\epsilon) \text{ で、 } |a_n - \alpha| < \epsilon \text{ が満たされる} \quad (1.1.2)$$

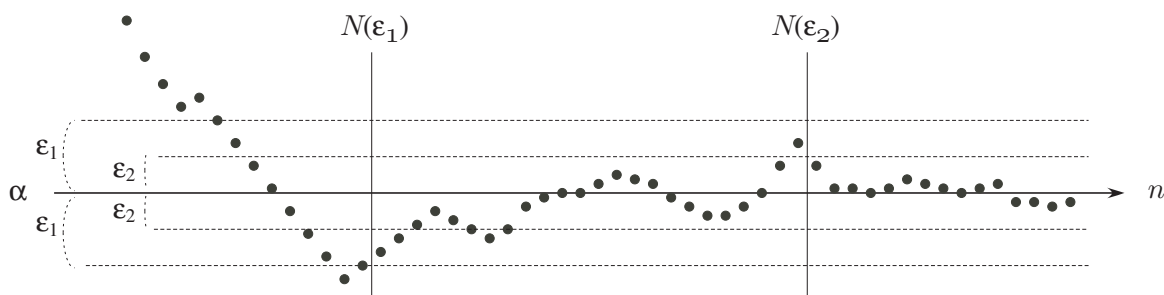
ような (十分に大きい) 実数 $N(\epsilon)$ が存在する。

(ア) は数式では以下のように書く (これは数学科の講義ではないので、この書き方は以下では使わない)：

$$\forall \epsilon > 0 \quad \exists N(\epsilon) \quad \left(n > N(\epsilon) \implies |a_n - \alpha| < \epsilon \right) \quad (1.1.3)$$

¹⁷教科書の 1.1 節

¹⁸俺はそんなもん考えたくないわ、と思った人は考えを改めよう。皆さんが高校でやってきたはずの「定積分」の存在を証明するだけでも、このような極限の問題が生じるので、この講義のメインテーマに直結してるのです。



少し補足説明：

- 上の定義の中で、括弧の中の(大きな)(小さな)はココロを述べたものである。これらは通常は省略されるが、慣れないうちは心の中で補うべきだ。
- $N(\epsilon)$ と書いたのは、「この N は ϵ によって決まる数なんだよ」と ϵ -依存性を強調するためである。
- (1.1.3) には2つの不等式 $n > N(\epsilon)$, $|a_n - \alpha| < \epsilon$ が現れている。ここはどちらも(または片方を) $n \geq N(\epsilon)$ や $|a_n - \alpha| \leq \epsilon$ (等号入り) に変えても、定義の意味する事は同じである(なぜ同じなのかは重要だから、各自で十分に納得せよ)。この講義では主に等号なしのバージョンを用いるが、等号入りのものを断りなく使うこともある。
- 通常は $N(\epsilon)$ を整数にとる事が多い。しかし、これは整数でなくても困らない上に、整数だとすると具体例の計算がややこしくなる。そこでこの講義では整数でない $N(\epsilon)$ を許すことにした。(気になる人は、後で十分に慣れてから、整数の $N(\epsilon)$ を使えば良い。)

この定義の最大の眼目は、極限という無限(ゼロ)の世界を扱っているのに、**ゼロでも無限でもない、有限の ϵ や N しか登場しない**点にある。有限のものなら(落ち着けば)我々は扱えるから、これは大きな利点だ。ただし、有限の ϵ や N を一つだけ考えても、これでは「極限」にならないのは明らかだ。そこで、上の定義では**その ϵ をいくらでも小さく選ぶ**ようにして、「どんどん大きくなる」「どんどん近づく」を表現している(以下で詳しく説明)。

細かい話に入る前に、 $\lim_{n \rightarrow \infty} a_n = +\infty$ など厳密に定義しておく：

定義 1.1.2 数列 a_n に対して、数列 a_n の $n \rightarrow \infty$ の極限がプラス無限大である、つまり $\lim_{n \rightarrow \infty} a_n = +\infty$ というのは、以下の(ア')が成り立つことと定義する：

(ア') 任意の(どんなに大きい)正の数 M に対しても、適当な(大きい)実数 $N(M)$ を見つけて、

$$\text{すべての } n > N(M) \text{ で、 } a_n > M \text{ とできる。} \quad (1.1.4)$$

(注) $\lim_{n \rightarrow \infty} a_n = +\infty$ や $\lim_{n \rightarrow \infty} a_n = -\infty$ の場合は、数列 (a_n) が収束するとは言わない。ただし、上のように「極限が無限大である」などとはいう。

1.1.1 少しでも理解を助けるために

上の定義 1.1.1 の意味するところは、自分でいろいろな例を作って納得するしかない。でも、理解を助けるために、少しだけ書いておこう。

1. 「いくらでも大きくなる」(無限大になる)の表現。 まず、「無限大」(一番大きい数)などは存在しない、ことを再確認しよう。なぜなら、一番大きい数があったとしたら、それに1を足したらもっと大きな数ができるから。だから、「 n が無限大」とは「 n がどんどん大きくなる状態」ととらえるしかない。これを有限の量のみを用いて表した結果が、「**適当に大きな N に対して、すべての $n > N$ では $\bigcirc\bigcirc$ が成り立つ**」という表現だ。

この表現には有限の N しか出てこない。けども、「 N より大きなすべての n 」を考えることで、実質、「無限大」に大きな n を考えてることを嘸み締めよう。

2. 「いくらでも近づく」の表現. 数列 $a_n = 1/n$ はいつでも正 (ゼロではない) だが, 極限はゼロになる. このように, 「その極限に ($n \rightarrow \infty$ で) いくらでも近づく」けれども「その極限には (有限の n では) 等しくなれない」ものの表現にも注意が必要だ. ここの「 n が無限大」と同様に, 有限の量のみを用いて表したい. それを実現するのが, 「**どんなに小さな $\epsilon > 0$ をとってきても, (n が大きくなっていくと, そのうちには) $|a_n - \alpha|$ が ϵ より小さくなる**」という表現だ.

ここにも有限, かつ正の ϵ しか登場しないが, この ϵ はこちらでいくらでも小さくとって行くのだ. $\epsilon = 10^{-6}$ より小さいか? $\epsilon = 10^{-14}$ よりも小さいか? $\epsilon = 10^{-200}$ なら? ... 「 N が無限大」と同じく, ここでも勝手にとってきた (どんなに小さくても良い) ϵ を考えることで, 実質的に「 $|a_n - \alpha|$ がいくらでも小さくなる」ことを表現していることを噛み締めてほしい.

3. N と ϵ のかけあい さて, 上の2つが非常にうまくむすびついて, いわば「掛け合い漫才」のように¹⁹ なっていることをよくよく理解しよう.

a_n が α に近づくか否かは, その距離 $|a_n - \alpha|$ で測っている. この距離は n を十分に大きくしない限りゼロに近づかない (ことが多い — 上の $a_n = 1/n$ の例を思い出せ). そこで, 本当にゼロに行くか否か判定するために,

「 $\epsilon = 0.0001$ になれるか?」「 $n > 100$ なら大丈夫」 (つまり, $n > 100$ なら $|a_n - \alpha| < 0.0001$)

「 $\epsilon = 10^{-6}$ になれるか?」「 $n > 20000$ としたら大丈夫」 ($n > 20000$ なら $|a_n - \alpha| < 10^{-6}$)

「 $\epsilon = 10^{-12}$ ならどや?」「 $n > 10^{20}$ で大丈夫」

「そしたら $\epsilon = 10^{-100}$ なら?」「それでも, $n > 10^{300}$ で大丈夫やで」

...

などといくらでも細かくしていけるかどうかを問うている訳だ. これがいくらでも小さい (つまり「任意の」) $\epsilon > 0$ でいけるのなら, $\lim_{n \rightarrow \infty} a_n = \alpha$ と言いましょう, というのが極限の定義.

逆に, 上の問答がどこかで切れてしまうなら, 例えば,

「 $\epsilon = 10^{-300}$ でどうや?」「ううん, N をいくら大きくしても今度はムリ!」

となってしまうたら, $\lim_{n \rightarrow \infty} a_n = \alpha$ とは言わないのだ.

4. N と ϵ の順序の問題 ϵ - N 論法で皆さんが戸惑う一つの理由は, N と ϵ の出てくる順番によると思われる. 高校までの言い方は「 n がどんどん大きくなると, a_n が α に近づく」または「 n を大きくすると, $a_n - \alpha$ がゼロに近づく」というものだ. ϵ が $a_n - \alpha$ を表していたつもりだから, これは「 $N \approx n$ が始めに出てきて, それから $\epsilon \approx |a_n - \alpha|$ が出る」構図である. ところが, ϵ - N 論法では順序が逆だ: 「どんなに小さな ϵ に対しても適当な $N(\epsilon)$ があつて」となっていて, ϵ が先, N が後.

この順序の逆転の理由は, 以下のような例を考えるとわかるかもしれない. 3つの数列を定義する ($n = 1, 2, 3, \dots$):

$$a_n = \frac{1}{n}, \quad b_n = \frac{1}{\log(2 + \log(2 + \log n))}, \quad c_n = \frac{1}{\log(2 + \log(2 + \log n))} + 10^{-8} \quad (1.1.5)$$

いくつかの n の値に対する, これらの数列の値を表にしてみると:

n	1	10	100	10^3	10^4	10^5	10^6	10^8	10^{16}
a_n	1	10^{-1}	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}	10^{-8}	10^{-16}
b_n	1.00938	0.80577	0.73645	0.69834	0.67321	0.65494	0.64084	0.62006	0.57692
c_n	1.00938	0.80577	0.73645	0.69834	0.67321	0.65494	0.64084	0.62006	0.57692

a_n の方は順調にゼロに行ってるが (アタリマエ!), b_n と c_n は動きが非常にノロい! また, b_n はゼロに行き, c_n はゼロに行かないはずだが, それもここまでの n では違いが全くわからない.

この例からわかるのは「同じ n の値で比べると, 数列によってはなかなかその極限の振る舞いが見えない」ということだ: a_n の方は $1/n$ だからまあ速くゼロに行くが, b_n は $\log$ が重なっている為に非常にゆっくりである. つまり, (アタリマエのことだが) 考える数列に応じて, 極限が見えやすいような大きな n をとってくる必要がある

¹⁹学習院大学物理学教室の田崎晴明氏の用語

わけだ。数列 c_n に至っては、初めは減っていくがそのうちに 10^{-8} に漸近して止まってしまう訳で、 n を大きくしたら収束が見えると思ってるとそのうちに裏切られる。

ここで困った理由は、 n の大きさを同じにして (n を先にとって) 3つの数列を比べようとしたことにある。これを避けるためには、順序を逆転させて、 N ではなくて ϵ を優先すれば良い。つまり、 $|a_n - \alpha|$ が (勝手にとってきた、非常に小さい) ϵ より小さくなるかどうかを知りたいわけだから、「 $|a_n - \alpha|$ の大きさの目安である ϵ を先に決めて、これに応じて n がどのくらい大きければ良いのか」を (またはいくら大きい n でも $|a_n - \alpha|$ が ϵ より小さくならないのかを) 考えるのが良い。これが ϵ - N 論法がこの順序で掛け合い漫才になっている理由である。

1.1.2 いろいろな例と定義の応用

この定式化の威力を知ってもらうには、下の命題 1.1.7 が良い例になってくれるだろう。しかしその前に、単純な例で具体計算をやって定式化に慣れる事が必要だ。以下の例をすべてやることを奨める。

問題 1.1.3 以下の数列が $n \rightarrow \infty$ で何に収束するのか (しないのか)、よくよく納得すること。その場合、 $N(\epsilon)$ がどのようにとれるのかを明示することが大切だ (いうまでもなく、 $n = 1, 2, 3, \dots$ である)。

$$a_n = 3, \quad b_n = \frac{1}{n}, \quad c_n = \frac{1}{\sqrt{n}}, \quad d_n = \frac{1}{n^2 + 1} \quad (1.1.6)$$

$$e_n = \begin{cases} 1 & (n \text{ が } 10, 10^2, 10^3, 10^4, 10^5, 10^6, \dots \text{ のとき}) \\ 0 & (\text{上以外のとき}) \end{cases} \quad (1.1.7)$$

(1.1.5) の 3つの数列も同様に考えてみよう。もう少し複雑な例も挙げておくから、考えてみよう ($n \rightarrow \infty$) :

$$f_n = \frac{n+3}{n}, \quad g_n = \frac{\sin n}{n}, \quad h_n = \sqrt{n+1} - \sqrt{n}, \quad p_n = \frac{2n+1}{n+1}, \quad q_n = \frac{1}{\log(n+1)} \quad (1.1.8)$$

具体的計算に少し慣れたら、以下のほとんどアタリマエに見える性質を ϵ - N を用いて証明しよう。

問題 1.1.4 極限に関する以下の性質を ϵ - N 論法を用いて厳密に証明せよ。

- $\lim_{n \rightarrow \infty} a_n = \alpha, \lim_{n \rightarrow \infty} b_n = \beta$ のとき, $\lim_{n \rightarrow \infty} (a_n + b_n) = \alpha + \beta$.
- $\lim_{n \rightarrow \infty} a_n = \alpha, \lim_{n \rightarrow \infty} b_n = \beta$ のとき, $\lim_{n \rightarrow \infty} a_n b_n = \alpha \beta$.
- $\lim_{n \rightarrow \infty} a_n = \alpha, \lim_{n \rightarrow \infty} b_n = \beta$ ($\beta \neq 0$) のとき, $\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\alpha}{\beta}$. この問題では分母の b_n がゼロになるかどうか、少し気になるところだ。実際、ある m では $b_m = 0$ となるような数列 $\{b_n\}$ もあるのだが、それでもこの性質が成り立つと言えるだろうか？

問題 1.1.5 (論理に弱い人にはキツイだろうから、できなくてもがっかりしないこと) 数列 $a_n = 1 + \frac{1}{n}$ はゼロには収束しない。このことを収束の定義に従って証明せよ。(「収束する」ことの定義は知っているから、その否定命題を考えればよい。) なお、以下の問題 1.1.6 を使って「この数列は 1 に収束するからゼロには収束しない」という証明も可能だが、これではなく、直接証明すること。

問題 1.1.6 (気がつけば簡単だが、これも慣れないと苦労するかも。) 数列 a_n が $n \rightarrow \infty$ で収束することがわかっている。収束先はただ一つであることを証明せよ。(収束先が 2つあるとすると、つまり、 $\lim_{n \rightarrow \infty} a_n = \alpha$ かつ $\lim_{n \rightarrow \infty} a_n = \beta$ であるとする、結局は $\alpha = \beta$ であることを証明せよ。) 証明すべき結論はアタリマエと思えるだろうが、そのアタリマエが証明できるかが問題だ。

少しは ϵ - N 論法に慣れたかな？ではこの辺りで、この論法の威力を示す命題を紹介しよう。この節の冒頭でも出したものである。

命題 1.1.7 数列 a_n から $b_n = \frac{1}{n} \sum_{k=1}^n a_k$ を定義する。 $\lim_{n \rightarrow \infty} a_n = \alpha$ ならば、 $\lim_{n \rightarrow \infty} b_n = \alpha$ である。

この命題の証明を、各自で高校までの定式化で試みると良い —— きちんと証明するのは大変だぞ (もし、高校までの定式化でもできたという人は僕のところまで来て下さい. 不可能とは言い切れないからね...). でも ϵ - N を用いると簡単にできてしまう. (まあ、簡単とは言ったけど、これが自力でできたら、それは大したものだ.) この定理の証明は、教科書の p.43 にある.

問題 1.1.8 (数列に関するチャレンジ問題) 命題 1.1.7 は

$$\lim_{n \rightarrow \infty} a_n = \alpha \quad \implies \quad \lim_{n \rightarrow \infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = \alpha$$

と主張している. そこで、右辺の「 a_1 から a_n の平均」をより一般の加重平均にして、同様の結果が成り立つかどうかを考えよう (より詳しくは以下に説明). まず、 $\rho_1, \rho_2, \rho_3, \dots$ を非負の数列として、

$$b_n := \left(\sum_{j=1}^n \rho_j a_j \right) / \left(\sum_{j=1}^n \rho_j \right)$$

を考える. 「 $\lim_{n \rightarrow \infty} a_n = \alpha$ ならば必ず $\lim_{n \rightarrow \infty} b_n = \alpha$ となる」ためには、 $\rho_1, \rho_2, \rho_3, \dots$ がどのような条件を満たしていれば良いか? できるだけ必要十分に近いものと考えてみよう. (命題 1.1.7 は $\rho_1 = \rho_2 = \rho_3 = \dots = 1$ に相当している.)

1.2 関数の極限: ϵ - δ 論法²⁰ (ここは簡単に)

前節では数列の極限、つまり、 n が無限大になったときに a_n がどうなるか、を見た. 今度は関数の極限、つまり、 x が連続変数で「 x が a に近づくとき $f(x)$ はどうなるか」を見たい. 考え方の基本は数列の場合と同じだから、少し簡単に行く.

定義 1.2.1 関数 $f(x)$ と実数 a, b に対して、「 $f(x)$ が $x \rightarrow a$ で b に収束する、つまり $\lim_{x \rightarrow a} f(x) = b$ 」というのは、以下の (イ) が成り立つことと定義する:

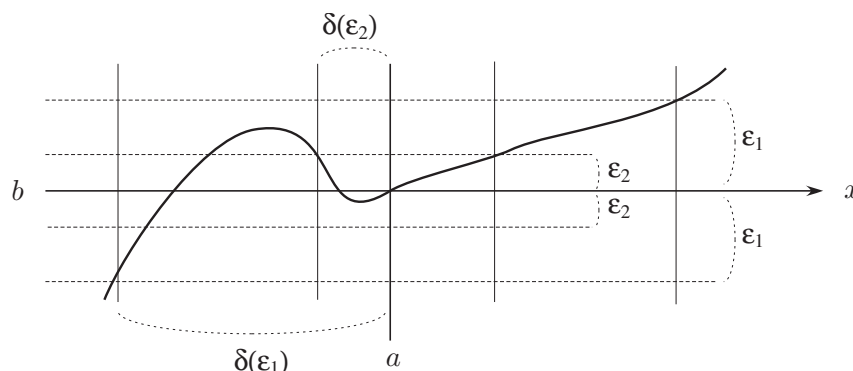
(イ) 任意の (どんなに小さい) 正の数 ϵ に対しても、適当な (小さな) 実数 $\delta(\epsilon)$ を見つけて、

$$0 < |x - a| < \delta(\epsilon) \text{ なるすべての } x \text{ で, } |f(x) - b| < \epsilon \text{ とできる.} \quad (1.2.1)$$

(イ) は数式では以下のように書かれる (以下では使わない. 将来の参考までに):

$$\forall \epsilon > 0 \quad \exists \delta(\epsilon) > 0 \quad \left(0 < |x - a| < \delta(\epsilon) \implies |f(x) - b| < \epsilon \right) \quad (1.2.2)$$

(注) 上の定義には $|x - a| > 0$ の条件がついている. つまり、 $x = a$ で何がおこっていようと、たとえ関数 $f(x)$ そのものが a で定義されなくとも、また $f(a) \neq b$ であっても、我々は気にしないのだ. (もちろん、 $f(a) = b$ でも文句はないが.) なぜ $x \neq a$ としているかの理由は、「関数の連続性」の定義を考えると理解できるのだが.



²⁰教科書の 1.2 節前半

注意: ϵ - N の時と同じく, 上の2つの不等式 $0 < |x-a| < \delta(\epsilon)$, $|f(x)-b| < \epsilon$ は, 等号入りの $0 < |x-a| \leq \delta(\epsilon)$, $|f(x)-b| \leq \epsilon$ に変えても同じである (ただし, $0 < |x-a|$ の方は等号入りにはしてはいけない, というのは上で注意した). この講義では主に等号なしバージョンを用いるが, 等号入りのものを断りなく使うこともあるので, また他の本では等号入りを用いていることもあるので, 注意されたい.

この定義にも ϵ - N 論法の時と同じ注意が当てはまる. 簡単に繰り返すと

- 極限を考えているのに, とともに正で有限の ϵ, δ しか定義に現れないところがミソである.
- ϵ, δ をどんなに小さくとっても良いという掛け合い漫才によって, 「 x が a に近づく」ときに「 $f(x)$ が b にいくらでも近づく」ことを表現しているのは, ϵ - N 論法と同じである.
- ϵ が先, δ が後になる理由も ϵ - N 論法と同じだ. 考えている函数によっては α への収束が非常に遅いこともあるから, そのような場合も扱うには「 $|f(x)-b| < \epsilon$ を実現するような $\delta(\epsilon)$ は何か (どのくらい小さい必要があるか)」を考える方が効率が良い.

ここの, いろいろな例をやることで感覚を身につけよう.

問題 1.2.2 以下の極限を, 定義に従って求めよ (極限は存在しないかもしれない). 極限が存在する場合は, $\delta(\epsilon)$ をどのようにとれば良いのか, 明記する事.

$$1) \lim_{x \rightarrow 0} x, \quad 2) \lim_{x \rightarrow 0} (x^2 - 2x + 3), \quad 3) \lim_{x \rightarrow 1} (x^2 - 2x + 3). \quad (1.2.3)$$

もうちょっとひねった例 ($a > 0$ は定数):

$$4) \lim_{x \rightarrow 0} \frac{1}{1+x}, \quad 5) \lim_{x \rightarrow 1} \frac{x^2-1}{x-1}, \quad 6) \lim_{x \rightarrow 0} \sin \frac{1}{x}, \quad (1.2.4)$$

$$7) \lim_{x \rightarrow a} \frac{x^3 - a^3}{x - a} \quad 8) \lim_{x \rightarrow 0} \frac{\sqrt{1+x} - \sqrt{1-x}}{x} \quad 9) \lim_{x \rightarrow 0} \sqrt{|x|} \quad (1.2.5)$$

問題 1.2.3 $f(x)$ を以下のように定めるとき, 極限 $\lim_{x \rightarrow 0} f(x)$ は存在するか? 存在するならその値と収束証明を, 存在しないならその理由 (収束しないことの証明) を ϵ - δ 論法の定義に基づいて述べよ.

$$f(x) := \begin{cases} 0.001 & (x = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, \dots) \\ x & (\text{上以外のとき}) \end{cases}$$

問題 1.2.4 $\lim_{x \rightarrow a} f(x) = \alpha$ かつ $\lim_{x \rightarrow a} g(x) = \beta$ の時, $\lim_{x \rightarrow a} \{f(x) + g(x)\} = \alpha + \beta$ と $\lim_{x \rightarrow a} \{f(x)g(x)\} = \alpha\beta$ が成り立つ. これらを ϵ - δ 論法によって証明せよ.

1.3 実数の連続性の公理²¹

「実数の連続性」は, その意義をつかみにくいと思われるので, 簡単にすませる. これでもまだわからない, という人は, 以下の1.4節に跳んでもまあ, 良い. 以下では断らない限り, 「数列」とは実数列 (実数でできた数列) の意味である.

実数と有理数との一番の違いは, 以下の公理 1.3.3 が満たされるか満たされないかにある. (その結果として, 「中間値の定理」が成り立つか成り立たないかなどの違いもおこる.) なお, 「公理」というと, こっちが勝手に仮定するような印象が出てくるが, 実際には, 「かくかくしかじかの定義²²によって実数を定義 (構成) すれば, 以下の公理 1.3.3 が満たされる」を証明することができる. この意味で, 公理 1.3.3 は仮定するものというよりは, (本講義では

²¹教科書の 1.1 節に少しある

²²実数の構成法には色々あるが, 原のお気に入りには, 「有理数のコーシー列の同値類」 (有理数の完備化) により定義する方法である (本講義では触れない)

省略する構成法によって) 証明されるべき、実数の重要な性質 (つまり公理 1.3.3 は実際には公理というよりは定理) と思ってもらった方が良いでしょう。

公理を述べるためにまず、補助概念を導入する。

定義 1.3.1 (部分列) 無限数列 $a_1, a_2, a_3, \dots$ が与えられた時、この数列から (順序を変えずに) 一部分を取り出して作った無限数列を数列 $\{a_n\}$ の部分列という。

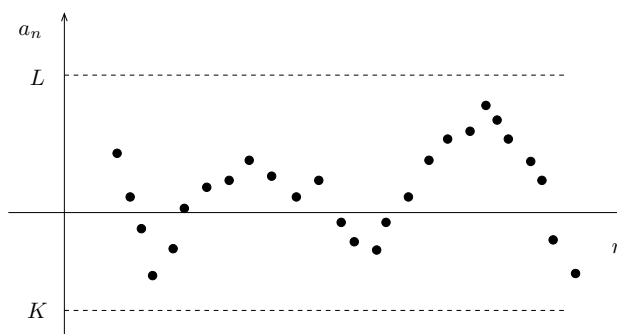
お約束として、 $\{a_n\}$ は $\{a_n\}$ それ自身の部分列とみなす (数列は、その数列自身の部分列)。

(例) 数列 $1, 2, 3, 4, 5, 6, \dots$ (自然数全体) の部分列の例としては $1, 3, 5, 7, 9, \dots$ とか、 $1, 4, 9, 16, 25, \dots$ とか $1, 2, 5, 10, 100, 10032, 2323445, \dots$ とか...

次に「有界な数列」の概念を定義する。

定義 1.3.2 (有界列) 数列 $\{a_n\}$ に対してある数 L が存在して、すべての n について $a_n < L$ が成り立っているとき、この数列は上に有界な数列という。また、ある数 K が存在してすべての n について $a_n > K$ が成り立っているとき、この数列は下に有界な数列という。上にも下にも有界な数列は単に有界な数列という。

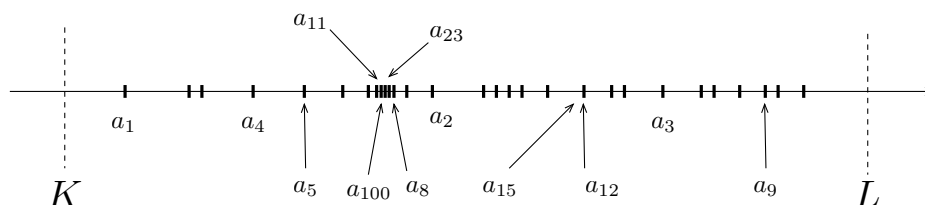
(注) K, L は一般に数列 $\{a_n\}$ に依存して決まるものであるが、もちろん、 n には依存してはいけない。



以上の下で、実数の連続性 (完備性) の公理を述べることができる。

公理 1.3.3 (実数の完備性または連続性) 有界な無限数列は必ず、収束する部分列を含む。つまり、有界な無限数列 $\{a_n\}$ が与えられれば、その部分列 $\{b_n\}$ をうまくとって、 $\{b_n\}$ が収束するようにできる。

この公理が何を言っているのかは、数直線上に $a_1, a_2, a_3, \dots$ の図を描いてみるのが良いだろう。図にすれば、かなりアタリマエに見えるものである。要するに、左を K 、右を L で区切られた数直線の区間に無限個の数を放り込むと、どこかにグチャッと集まるしかない、という主張である。(この、グチャッと集まった点を集積点 (accumulation point) という。集積点は 2 個以上存在するかもしれないが、少なくとも 1 個の集積点が存在するというのが公理の主張である。)



ただし、有理数の集合に対してはこの公理が成り立たないことは納得しておきたい。例えば、

$$a_n \text{ とは } \sqrt{2} \text{ の十進展開の小数点以下 } n \text{ 桁までとったやつ} \quad (1.3.1)$$

と定義してみる ($a_1 = 1.4, a_2 = 1.41, a_3 = 1.414, \dots$)。この数列の極限はもちろん $\sqrt{2}$ であって、上の公理を満たす数列の例になっている。(この場合、部分列をとるまでもなく収束している)。しかし、有理数の範囲でこの数列の極限を探しても極限は存在しない。

つまり、有理数の集合は、上の公理が成り立たない数の例になっている。有理数の集合と実数の集合、ほとんど同じように見えるのだが、極限などに関してはこれほどの違いがある。公理 1.3.3 は、上の図で「アタリマエ」に見えることが実際に成り立つように保証してくれるものともいえる。

数学的には重要な注

- この節の最初でも注意したことであるが、上ではさりげなく「実数の公理」と書いたけども、この公理を満たすような数の体系が本当に**存在するのか**(作れるのか)は大きな問題で、証明すべきである。これは「上の実数の公理系は無矛盾か」と言ってもよい。この講義ではこの問題には全く触れないが、結論だけ言うと、「上の公理を満たす実数の体系は存在する」となる。この辺りの詳しい話は原の過去の講義「数学 II」(2008 年度)の講義録にある。
- 「実数の公理」には互いに同値ないくつかの表現があり、以下に述べる「有界単調列は必ず収束する」「コーシー列は必ず収束する」などを公理とすることもある。特に個人的には「完備性」という場合には、「コーシー列は必ず収束する」を採用すべきだと思うが、これはちょっとわかりにくい。そこで、コーシー列は後でやることにして、現時点では直感的に分かりやすいと僕が思ったものを上の公理に採用した。

1.4 単調な数列²³

これまでに「行き先がわかっている極限」の定義はやった。 $\lim_{n \rightarrow \infty} a_n = \alpha$ とは、もちろん、数列 a_n の行き先が α だということであり、

$$\text{どんなに小さい } \epsilon > 0 \text{ に対しても } N(\epsilon) \text{ をうまくとると、 } (n > N(\epsilon) \text{ では } |a_n - \alpha| < \epsilon) \text{ となる} \quad (1.4.1)$$

という「定義」を行った。また、実際に数列の収束発散はこの定義に従って判定してきた。ところが、この定義は行き先 α がわかっていなければ使い物にならない。でも実際には、**行き先の値ははっきりわからなくても、それが収束するか否かを判定したい数列**はいくらでもある。

例えば、高校でも散々に出てきた非常に重要な数、 e の定義を考えよう。この数の定義 (のひとつ) は

$$e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n \quad (1.4.2)$$

という極限だが、この極限が実数として存在することを、今までの知識で証明できるだろうか? この数の存在が証明できなければ、物理で (多分) 最も重要な指数関数が定義できなくなるぞ...

これ以外にも、「行き先がきれいには書けないけども極限の存在を証明したい例」はいくらでもある。皆さんが知ってるはずの「定積分」も極限で定義されるから、その極限が存在することを示せなければ非常に困る。そもそも数学で扱う大抵の極限は「その値はきれいには書けないけど、その存在はわかっている」もので、実際にはその極限でその値を「定義」するものだから、極限の存在が言えなければ大変に困る!

さらにいうと、「極限が存在する」ことを知ると、**実際上も役にたつ**ことがある。というのは、「極限の値はわからないけど、極限が存在することがわかっている」場合には、その極限の値を x などとおいた上で、「この x の満たすべき方程式は何か?」などの問いを発してみると良い。 x の満たすべき方程式などが簡単にわかり、その方程式を解くことで極限の値もわかる、ことが往往にして起こるものである。(このテクは高校でも「漸化式で与えられた数列の極限」を求める時に使ったのでは?)

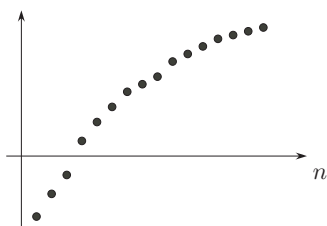
という訳で、**行き先の値がわからない数列でも、その数列が収束することだけは判定できる**ような定理が欲しい。これに応えようとして数学者が整備した概念が「単調増加 (減少) 列」「上極限と下極限」「コーシー列」などである。これらはそれほど簡単ではないものも含むので、この小節では一番簡単に直感的な単調列を考える。

定義 1.4.1 (単調列) $a_1 \leq a_2 \leq a_3 \leq \dots \leq a_n \leq \dots$ となっている数列 a_n を**広義の単調増加数列**、または**単調非減少数列**という (不等号にイコールが入ってないものは単調増加数列という)。不等号が逆向きになったのは「広義の単調減少」または「単調非増加」数列という。

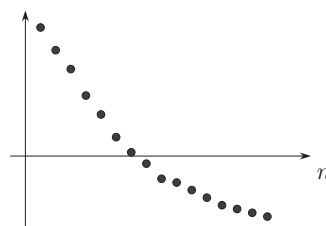
²³教科書 1.1 節に少し

(言葉に関する注)

- 英語では 単調増加 = (monotone) increasing, 単調減少 = (monotone) decreasing, 単調非減少 = (monotone) non-decreasing, 単調非増加 = (monotone) non-increasing.
- 上の定義中の「単調増加」を「狭義の単調増加」とか「真に単調増加」ということもある. 同様の用語は関数の増加・減少についても用いる.
- 「単調増加」を「広義の単調増加」の意味で使う事も時々あるので注意が必要である. 実際, 研究論文のレベルでは上の定義の意味での「広義の単調増加」を単に「単調増加」と言い, 上の定義の意味での「単調増加」は「真に単調増加 (strictly increasing)」という事が多い. この辺りの用語は人によって違うだから, どの意味で使ってるかは自分で確認すべし.



単調増加数列の例



単調減少数列の例

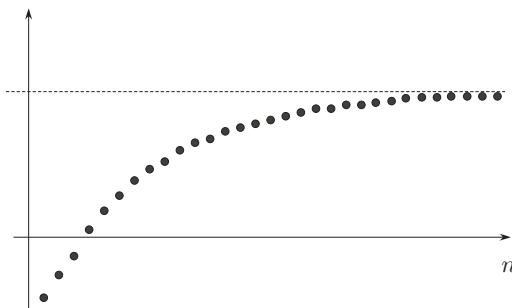
さて, 有界かつ単調な数列には, 以下の著しい性質がある. 直感的にはあたりまえに見えるだろう.

定理 1.4.2 (有界単調列の収束; 教科書の定理 1.2) 数列 $\{a_n\}$ が上に有界で広義単調増加のとき, $\lim_{n \rightarrow \infty} a_n$ は存在する. また, $\{a_n\}$ が下に有界で広義単調減少のときも, $\lim_{n \rightarrow \infty} a_n$ は存在する.

(注) $\{a_n\}$ が有界でない広義単調増加列の場合は $\lim_{n \rightarrow \infty} a_n = +\infty$ であるし, $\{a_n\}$ が有界でない広義単調減少列の場合は $\lim_{n \rightarrow \infty} a_n = -\infty$ である. このような場合には「極限が存在する」とは言わないのが数学のお約束だと前に注意したが, ここを敢えて「極限が $-\infty$ 」「極限が $+\infty$ 」という事にすれば, 上の定理は以下のようにも言える.

極限の値として $\pm\infty$ も許す事にすると, 単調な数列の極限 $\lim_{n \rightarrow \infty} a_n$ は常に存在する.

定理 1.4.2 はあたりまえには見えるが, 決してあたりまえではなく, 実数の連続性に強く依存している. それを示す簡単な例として, 数列 a_n を, 「 $\sqrt{2}$ を十進小数で書いたときの小数点以下 n 桁めまでの数」と定義してみる (この例はこれまでもよく使っている). a_n のそれぞれは有理数で, 単調増加, 更に有界でもある. しかしその極限は $\sqrt{2}$ という無理数であって有理数の中にはない. つまり, 極限を有理数の集合の中で探すと, この数列は (収束先が有理数ではないので) 収束しないことになってしまう. より広い実数全体の中で極限を探すと, (かつその実数が連続性を持っているおかげで), 極限の存在が保証され, 上の定理が成り立つ訳だ.



1.5 コーシー列²⁴

前小節の「有界単調列の収束」は, 『極限の値がわからなくても収束することは証明できる』例であった. ところが, 世の中にある数列のほとんどは単調列ではない. でもこれらについても, 収束の判定条件が欲しい. となると,

²⁴教科書の定理 6.2 など

単調列を超えた、もう少し難しいことが必要で、このような時に威力を発揮するのが**コーシー列**の概念である。

以下で述べるように、「この数列が収束する」ことと「この数列がコーシー列である」ことは全く**等価**(同値)である。つまり、ある数列が収束するか否かを判定するには、この数列がコーシー列であるか否かを判定すれば良いことになる。さらに(以下の通り)、**コーシー列か否かの判定は非常に簡単**であるから、数列が収束するか否かの判定が大変に簡単になる。

このようなわけで、コーシー列の概念は大変に有用であるので、(証明の部分は知らなくて良いから)、概念だけはしっかり理解して欲しい。

定義 1.5.1 (コーシー列; 教科書 p.158) 数列 a_n が以下の性質を満たすとき、これを**コーシー列** (Cauchy sequence) という。

任意の (小さい) $\epsilon > 0$ に対し、(十分大きな) 数 $N(\epsilon)$ がとれて、

$$\text{すべての } m, n \geq N(\epsilon) \text{ に対して } |a_m - a_n| < \epsilon \text{ とできる.} \quad (1.5.1)$$

上の条件は数式で書けば

$$\forall \epsilon > 0 \quad \exists N(\epsilon) > 0 \quad (m, n \geq N(\epsilon) \implies |a_m - a_n| < \epsilon) \quad (1.5.2)$$

である (数学科ではないので、この数式は今後使わない)。

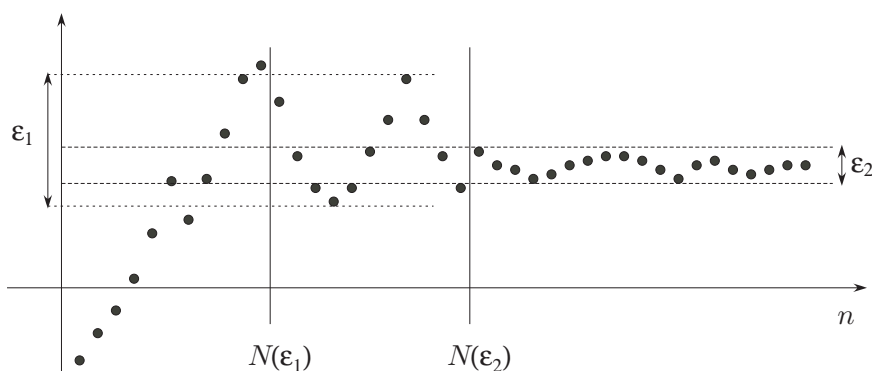
(ええ加減な書き方) ちょっと「ええ加減」な書き方では、上のコーシー列の条件 (1.5.1) は

$$\lim_{\substack{m \rightarrow \infty \\ n \rightarrow \infty}} |a_m - a_n| = 0 \quad (1.5.3)$$

と書ける。ただ、この書き方では「2重極限が実際にどのような意味であるか」が少し曖昧なため、そこを明らかにする意味で、(1.5.1) のように書くのが正しい。

(重要な注意) コーシー列の定義はなかなかとつきにくいもののようだ。最大の困難は、「 $N(\epsilon)$ 以上の任意の m, n 」(m, n 両方があるところが新しい) にあると思われる。わかりにくい人は直感的には (1.5.3) と思いつつ、その正確な意味を数式によって理解しようとするが良いだろう。

定義のキモは、「ある $N(\epsilon)$ より先なら、**この数列の変動の幅**が ϵ 以下になる」(それでももちろん、いつものように、 $\epsilon > 0$ はどんなに小さくても云々) という見方である。このイメージを描いたのが以下の図のつもりなのだが...



すぐには呑み込めないかもしれないが、この定義と次の定理の意味を各自で良く理解してほしい。収束先がわからないような数列を考えるのだから、収束先と a_n の差を計算する事はできない。それでも、 a_n と a_m の差 (の m, n が無限大になった極限; ただし、 m と n が独立に無限大に行ってよい、のがミソ) を見れば収束するかどうか判定できる、というのである。これは実用上、非常に重要だ。

命題 1.5.2 (教科書の定理 6.2 の一部)

- (1) 数列 a_n が (何かの値に) 収束するならば, a_n はコーシー列である.
- (2) コーシー列は有界である.

定理 1.5.3 (コーシーの収束条件; 教科書の定理 6.2, 非常に大事) a_n がコーシー列であれば, a_n は (何かの値に) 収束する. 上の命題と合わせると, 数列が収束することの**必要十分条件**は, その数列がコーシー列であることだ.

これらの証明は後回しにして, この定理の意味するところを述べておく.

コーシー列の応用 (重要性)

今までにも強調した通り, ある数列が「収束する」と「コーシー列である」ことは同値だ. だから, 「コーシー列」であるかどうかは, 収束するかどうかの最強の判定条件といえる. 実際, ある数列が収束するかどうかの判定のほとんどはコーシー列かどうかで行うと言ってもよい. この意味で, コーシー列は非常に非常に重要なのである. 以下の証明はわからなくても良いが, 「与えられた数列がコーシー列か否か」を判定するにはどうすれば良いかわかる程度には, コーシー列の定義を理解しておくことは必要だ.

(コーシー列よりも有界単調列かどうかの判定の方が簡単だが, 世の中それほど甘くはなく, 問題の数列が単調である事はそんなにないので, 単調列はあんまり役に立たない. これとは別に, この授業では扱わない $\limsup$ や $\liminf$ が役に立つ事はかなりある.)

問 1.5.4 「コーシー列」の定義を理解する問題. 以下の数列はすべて収束する数列であるから定理 1.5.3 によれば, コーシー列のはずである. そこで, コーシー列の定義に従って, 以下の数列のそれぞれがコーシー列であることを示せ. 特に, $N(\epsilon)$ をどのようにとれば十分か, できるだけギリギリの評価を与えよ.

$$a_n := \frac{1}{n} \quad b_n := \frac{1}{n^2} \quad c_n := \frac{(-1)^n}{n} \quad d_n := \frac{(-1)^n}{\sqrt{n}}$$

問 1.5.5 「コーシー列」または「有界単調列」の考えを用いて, 次の数列 $\{a_n\}, \{b_n\}, \{c_n\}$ が収束する事を証明せよ (α は正の定数). また, c_n の極限值を求めよ.

$$a_n := -\log n + \sum_{k=1}^n \frac{1}{k} \quad b_n := \sum_{k=1}^n \frac{(-1)^{k-1}}{k}$$

$$c_1 := 1, \quad n \geq 1 \text{ では } c_{n+1} := \frac{1}{2} \left(c_n + \frac{\alpha}{c_n} \right)$$

正直, c_n はそこそこ難しいと思う.

以下, 二つの定理の証明を述べるが, 難しい人は無視すれば良い. ただ, これらは $\epsilon-N$ 論法の良い練習になっているので, 参考までに載せておく.

命題 1.5.2 の証明

(1) これは簡単だ. 数列 a_n の収束先を α と書くと, 収束の定義から, 勝手な (小さな) $\epsilon > 0$ に対して $N(\epsilon)$ をとって, 全ての $n > N(\epsilon)$ では $|a_n - \alpha| < \epsilon/2$ とできる. つまり, $m, n > N(\epsilon/2)$ では

$$|a_m - \alpha| < \frac{\epsilon}{2}, \quad |a_n - \alpha| < \frac{\epsilon}{2} \tag{1.5.4}$$

となっている訳だ. でも三角不等式から, このような m, n では

$$|a_m - a_n| = |(a_m - \alpha) + (\alpha - a_n)| \leq |a_m - \alpha| + |\alpha - a_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{1.5.5}$$

が成り立つ. これはコーシー列の条件 (1.5.1) が成り立っていることを意味する.

(2) a_n がコーシー列である場合, $N(\epsilon)$ が存在して, $m, n \geq N(\epsilon)$ では $|a_m - a_n| < \epsilon$ となっている. よって, 特に $n = N(\epsilon)$ ととれば, $m \geq N(\epsilon)$ では $|a_m - a_{N(\epsilon)}| < \epsilon$, つまり

$$a_{N(\epsilon)} - \epsilon < a_m < a_{N(\epsilon)} + \epsilon \quad (1.5.6)$$

がなりたつ. つまり, $N(\epsilon)$ から先の a_m は有界なのだ. しかし, $N(\epsilon)$ 以下の a_m は有限個しかないから, これは絶対に有界. $N(\epsilon)$ 以上と $N(\epsilon)$ 以下の両方の場合を併せて, a_n は有界である. $\square$

定理 1.5.3 の証明

命題 1.5.2 の (2) から, コーシー列 $\{a_n\}$ は有界である. 従って, 実数の公理 1.3.3 から, $\{a_n\}$ は収束する部分列を少なくとも一つ含む. その部分列を b_k とし, その極限を β としよう. また, b_k と a_n の添字の対応を $b_k = a_{i_k}$ と書くことにする (a_n において, $n = i_k$ ととったものが b_k になっている, の意味. 当然, $i_k \geq k$ である.)

以下では a_n がこの β に収束することを証明する.

そのために, 任意の $\epsilon > 0$ を固定する. コーシー列の定義により, $N_1 = N(\epsilon/2)$ が存在して,

$$N_1 \text{ より大きい } m, n \text{ では } |a_m - a_n| < \frac{\epsilon}{2} \quad (1.5.7)$$

がなりたっている.

一方, $\{b_k\}$ の極限が β だから, 十分大きな N_2 をとると,

$$k \geq N_2 \text{ ならば } |b_k - \beta| < \frac{\epsilon}{2} \quad (1.5.8)$$

が成り立っているはずである.

そこで, $N_3 := i_{N_2}$ ととてやろう. つまり N_3 とは, b_{N_2} をあたえるような a_n の添字 n である. さらに, $N_4 = \max\{N_1, N_3\}$ と定義する. さて, $m, n > N_4$ なら, (1.5.7) から

$$|a_m - a_n| < \frac{\epsilon}{2} \quad (1.5.9)$$

がなりたつ. しかし一方で, $m > N_4$ である m の中には, 部分列 b_k を与えるようなものが絶対にある (つまり, $m = i_k$ と書けて, $a_m = b_k$ となるようなものがある). そのような a_m に対しては (1.5.8) から

$$|a_m - \beta| < \frac{\epsilon}{2} \quad (1.5.10)$$

が成り立っているはずである. この 2 つを組み合わせると, $n > N_4$ ならば

$$|a_n - \beta| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \quad (1.5.11)$$

が結論できる. これは $\lim_{n \rightarrow \infty} a_n = \beta$ を ϵ - N で書いたものに他ならない. $\square$

最後に, これまでの数列の収束 ($n \rightarrow \infty$) に関する収束条件を, 関数の収束 $x \rightarrow a$ に書き直した定理を挙げておこう.

定理 1.5.6 (コーシーの収束条件)

(1) $\lim_{x \rightarrow \infty} F(x)$ が存在するための必要充分条件は, $F(x)$ が以下のコーシーの条件を満たす事である:

(C_∞) 任意の $\epsilon > 0$ に対して (十分大きな) $L(\epsilon) > 0$ がとれて, $x, y > L(\epsilon)$ なる
任意の x, y に対して $|F(x) - F(y)| < \epsilon$ が成り立つ

(2) a を有限の実数とする. $\lim_{x \rightarrow a} F(x)$ が存在するための必要充分条件は, $F(x)$ が以下のコーシーの条件を満たす事である:

(C_a) 任意の $\epsilon > 0$ に対して $\delta(\epsilon) > 0$ がとれて, $0 < |x - a| < \delta(\epsilon)$ かつ $0 < |y - a| < \delta(\epsilon)$ なる
任意の x, y に対して $|F(x) - F(y)| < \epsilon$ が成り立つ

1.6 連続函数とその性質²⁵

連続函数については高校でも習ったと思うが、 ϵ - δ での定式化を行っておこう。

定義 1.6.1 点 a を含む区間で定義された $f(x)$ が「 a で連続」とは、 $\lim_{x \rightarrow a} f(x) = f(a)$ なることである。つまり、以下の (ウ) が成り立つことである：

(ウ) 任意の (どんなに小さい) 正の数 ϵ に対しても、適当な (小さな) 実数 $\delta(\epsilon)$ を見つけて、

$$|x - a| < \delta(\epsilon) \text{ なるすべての } x \text{ で, } |f(x) - f(a)| < \epsilon \text{ とできる.} \quad (1.6.1)$$

(ウ) は数式では

$$\forall \epsilon > 0, \exists \delta(\epsilon) > 0, \quad (|x - a| < \delta(\epsilon) \implies |f(x) - f(a)| < \epsilon) \quad (1.6.2)$$

となる。(この講義ではこの数式は使わない)。

函数の極限の定義 1.2.1 と比べると、 $0 < |x - a| < \delta(\epsilon)$ が $|x - a| < \delta(\epsilon)$ となっていて、 $0 <$ が無いのが不思議だ、と思った人もいるかもしれない。しかし、今の場合、 $\lim_{x \rightarrow a} f(x)$ が $f(a)$ そのものに等しくなって欲しいのだから、わざわざ $x \neq a$ のみを考える必要はない — というより、 $x = a$ が入っていることが本質的なのだ。このような理由で、 $0 <$ は書いていないのだ。

なお、片側連続を問題にすることもある。一応、定義を書くが、この講義ではあまり気にしなくて良い。

定義 1.6.2 函数 $f(x)$ が a で右連続とは、 $f(x)$ が a を左端とするある区間で定義されていて、かつ $\lim_{x \rightarrow a+0} f(x) = f(a)$ なることである。同様に、 a で左連続とは、 $f(x)$ が a を右端とするある区間で定義されていて、かつ $\lim_{x \rightarrow a-0} f(x) = f(a)$ なることである。

- 「右連続」を「右へ連続」, 「左連続」を「左へ連続」ということもある。英語ではそれぞれ right continuous, left continuous (または continuous to the right, continuous to the left)。
- $f(x)$ が閉区間 $[a, b]$ で連続とは、

$$c \in (a, b) \text{ では } \lim_{x \rightarrow c} f(x) = f(c), \quad \text{かつ} \quad \lim_{x \rightarrow a+0} f(x) = f(a), \quad \lim_{x \rightarrow b-0} f(x) = f(b) \quad (1.6.3)$$

となることである (区間の中では普通の連続, 区間の端点では右 (左) 連続)。

- 普通の連続にしても、片側連続にしても、比べるべきは $f(a)$ そのものと (右や左からの) 極限值 $\lim_{x \rightarrow a} f(x)$ だ。単に右側からと左側からの極限值が同じでも連続とは限らないから注意。(例を考えよ。)

問 1.6.3 函数 $f(x) = \sqrt{|x|}$ が、任意の x で連続であることを、定義に戻って示せ。

問 1.6.4 函数 $f(x)$ が

(あ) $x = a$ で連続である事と、

(い) $x = a$ で右連続かつ左連続でその値が等しい事

は同値であるか？

ϵ - δ を習ったので、以下の (一見、アタリマエの) 定理群を証明できる。

²⁵教科書 1.2 節

定理 1.6.5 点 a を含むある区間で定義された函数 $f(x)$ が $x = a$ で連続だとする.

- この時, a の近傍で f は有界である. 特に, 充分小さな $\delta > 0$ をとれば, $|x - a| < \delta$ なるすべての x で

$$|f(x)| < |f(a)| + 1 \quad (1.6.4)$$

がなりたつ.

- もし $f(a) > 0$ ならば, a の近傍では $f(x) > 0$ である. 特に, 充分小さな $\delta > 0$ をとると, $|x - a| < \delta$ なるすべての x で

$$f(x) > \frac{f(a)}{2} \quad (1.6.5)$$

がなりたつ. $f(a) < 0$ の時は不等号の向きをひっくり返せば同様の結論がなりたつ.

命題 1.6.6 函数 f が a で連続, 函数 g が $b = f(a)$ で連続なら, 合成函数 $h(x) = g(f(x))$ も a で連続である.

命題 1.6.7 函数 f, g が a で連続とする.

- (1) $f(x) + g(x)$, $f(x) - g(x)$ は共に a で連続である.
- (2) 積 $f(x)g(x)$ も a で連続である.
- (3) $g(a) \neq 0$ なら, $f(x)/g(x)$ も a で連続である.

さて, 実数の連続性を認めると, 連続函数の重要な性質 (2つ) を証明できる. その一つ目は, 高校でも習ったはずの中間値の定理である.

定理 1.6.8 (中間値の定理; 教科書の定理 1.13) 閉区間 $[a, b]$ で連続な函数 $f(x)$ を考える. $f(a)$ と $f(b)$ の間にある任意の数 F に対して, $f(c) = F$ なる $c \in [a, b]$ が少なくとも一つ存在する. つまり, x が a から b に動くとき, $f(x)$ は $f(a)$ と $f(b)$ の間のすべての値を (少なくとも一回は) とる.

これまでも強調してきたが, この定理は実数の連続性があって初めて成り立つものだ. 例えば函数 $f(x) = x^2 - 2$ が $f(x) = 0$ をとるような x の値を考えてみる. 無理数まで含めれば, もちろん, $x = \pm\sqrt{2}$ でゼロになる訳だが, 有理数の範囲ではそのような x は存在しない. つまり, 有理数 (連続性のない数の集合の例) の範囲で考えておれば, この定理の結論はなりたたないのだ.

この例では問題になる x の値が具体的にわかっているから「 $x = \sqrt{2}$ を数の集合に加える」ことで対症的に対処できるが, 一般の函数で同じことをやるのはまず, 不可能だ. その意味でも数の集合を「連続性」をもつ集合まで広げておく事は不可欠だったのだ.

さて, 連続函数の重要な性質その2は最大値, 最小値に関するものである. これもグラフを描けば直感的には明らかであるが, それがきちんと証明できるようになったこと (そしてその背後には実数の連続性があること) が重要である.

定理 1.6.9 (連続函数の最大値・最小値は常に存在; 教科書の定理 1.14) 閉区間で連続な函数は必ず, その区間で最大値, 最小値をとる. 従って特にこのような函数は有界である.

- 「閉区間で連続である」ことは重要な条件である. 例えば, 函数 $f(x) = 1/x$ を开区間 $(0, 1)$ で考えると, こいつは最大値を持たず, 有界でもない. また, $g(x) = x$ を同じく开区間 $(0, 1)$ で考えると, こいつは有界だが最大値も最小値も持たない.
- 実数の連続性が重要である事の傍証は以下のような例からもわかる. 有理数上だけで定義された函数 $g(x) = \sin x$ は, x が有理数に限定されている限り最大値を持たない. (ただし, $\sin x$ を有理数だけに対して定義することは不可能ではないが, 少し不自然である. この意味でこの例はちょっと「人工的」なものである.)

1.7 連続函数の効用：指数函数と対数函数

ここまでの準備を経て、 x^α (α は実数, $x > 0$), 指数函数, 対数函数などを自然に定義する事ができる. 例えば, x^α を定義するには, まず α に収束するような有理数の列 a_n を考え,

$$x^\alpha = \lim_{n \rightarrow \infty} x^{a_n} \quad (1.7.1)$$

とする. つまり, x^α が α について連続になるように定義するわけだ²⁶. このようにして, 有理数で定義された函数を非常に自然に実数に拡張することができる.

時間の関係上, この講義ではこの話題にはこれ以上ふれませんが, 教科書の p.19~p.24 の内容は, みなさんが後々使うことが多いでしょうから, 各自, 読んでおいてください (概念的に難しいものはないはず).

²⁶これで納得してしまったあなた, まだまだ甘いですね! このように定義するなら, 「上の極限が $\lim_{n \rightarrow \infty} a_n = \alpha$ となるすべての $\{a_n\}$ の取り方について同じである」ことを確かめる必要があります.

2 微分

これで漸く、「微分」に入ることができる。これまで延々と基礎の部分の準備をしてきたので、これを用いてまずは（高校でやったことになっている）微分の基礎付けを簡単に行う。そのあとで、高校ではやらなかった新しい題材も学習する（テイラー展開）。

微分を考える理由には大きく分けて2通りある。

- 微係数は函数の「変化率」を表すから、微分の値（正負）を知ることで、函数の増減を知ることができる。特に「微係数がゼロ」の点を探すことで極大・極小問題が奇麗に解ける。また、2階微分を考えるとグラフの凹凸も知ることができる。
- 微分を利用して函数を級数に展開できる（テイラー展開）。これを利用して、函数の近似値が計算できる。

このうち、第一の視点は受験などを通して散々やってきたものと思うので、この講義では簡単にすませる — これが形を変えて、「多変数函数の微分」（偏微分）として登場するが、これは後の話。ところが、第2の「テイラー展開」は、現在の高校のカリキュラムにはなく、案外、みなさんが苦勞するようだ。そこで、この重要なテーマをマスターするのが微分に関する大きな目標の一つになる。

2.1 微分の定義²⁷

微分については、かなり高校でやっている。大学で付け加えるべき事は、微分を定義している極限の定義が新しく厳密になった、ということくらいだ。だから、簡単に行きましょう。まずは高校の復習から。

定義 2.1.1 (微分係数) $x = a$ とその近傍で定義されている函数 $f(x)$ に対して、極限

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \quad (2.1.1)$$

が存在するとき、この極限を $f(x)$ の $x = a$ での微分係数 (derivative) とよび、 $f'(a)$ または $\frac{df}{dx}(a)$ と書く。またこのとき、 $f(x)$ は a で微分可能 (differentiable) という。なお、 f がある区間 I のすべての点で微分可能であるとき、 f は I で微分可能という。

色々な a に対する $f'(a)$ の全体は a に $f'(a)$ という値を対応させる函数だと考えられるので、これを f の導函数 (derived function, または derivative) とよぶ。微分係数は、考えている函数の「変化率」（増減の目安）であり、グラフの接線の傾きであったことを思い出しておこう。

(注) 極限のところで注意したように、 $x \rightarrow a$ というのは $|x - a| \rightarrow 0$ の事であったから、(2.1.1) の極限に於いても x は可能なすべての近づき方を考える。この極限の取り方を片側に制限すると以下の定義になる：

定義 2.1.2 (片側微分係数) 定義 2.1.1 の状況の下で、極限

$$f'_-(a) := \lim_{x \rightarrow a-0} \frac{f(x) - f(a)}{x - a} \quad (2.1.2)$$

が存在するとき、この極限を $f(x)$ の a での左微分係数 (left derivative) とよぶ。また、

$$f'_+(a) := \lim_{x \rightarrow a+0} \frac{f(x) - f(a)}{x - a} \quad (2.1.3)$$

が存在するとき、この極限を $f(x)$ の a での右微分係数 (right derivative) とよぶ。

f が a で微分可能なら、右微分係数も左微分係数も存在して、 $f'(a)$ に等しい事はすぐにわかる。実はその逆も成り立つ。つまり、右微分係数と左微分係数が両方存在して $f_-(a) = f_+(a)$ ならば、 f は a で微分可能で、 $f'(a) = f_-(a) = f_+(a)$ である。まあ、この辺りは片側連続と同じノリやわな。

²⁷教科書 2.1 節前半

教科書の定理 2.2, 定理 2.3, 定理 2.4 (いずれも微分の計算規則) などは高校でもやったと思うので, くり返さない. 各自で思い出して, 試験になったらできるようになっておくこと.

微分可能性と連続性の間には非常に重要な以下の関係がある (教科書の定理 2.1):

定理 2.1.3 函数 $f(x)$ が $x = a$ で微分可能であれば, f は a で連続である.

(証明) 微分可能性の定義を書き下せば簡単にるので略. ただし, 各自で一度はやっておく事. □

(注) 上の定理の逆はなりたたない. つまり, (1 点で) 連続だけれど微分不可能な函数の例はすぐに作れる (各自でやること!). なお, すべての点で連続だけれど, どの点でも微分不可能な函数も (なかなか想像しにくい) が存在する. 一つの例が田島本の p.129 に載っている (Weierstrass).

2.2 オーダーの概念 (ランダウの記号) ²⁸

いろいろな函数を扱って解析する場合, 「この項がああ項よりも大きい」「この項に比べるとあは項は無視できる」などとして近似的に議論を進めることが多い. この際に, 「大きい」「無視できる」などにどのような定量的な意味が付与されているのか, はっきりさせておくことは必要だ. その重要な概念と記号を学習しよう.

定義 2.2.1

- (1) x の函数 $f(x)$ が, $\lim_{x \rightarrow a} f(x) = 0$ を満たす時, 「 $x \rightarrow a$ において, $f(x)$ は**無限小**である」という.
 (2) $f(x)$ と $g(x)$ が共に $x \rightarrow a$ において無限小の時, 以下の 2 つの定義を導入する:

- もし $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$ であるなら, 「 $x \rightarrow a$ において, $f(x)$ は $g(x)$ よりも**高位の無限小**である」といい, 記号で $f(x) = o(g(x))$ と書く. (ここで, o は小文字のオーである.)
- もし, ある定数 K が存在して, $x = a$ の近くでは $\left| \frac{f(x)}{g(x)} \right| \leq K$ が成り立つなら (つまり, $x = a$ の近くで $\frac{f(x)}{g(x)}$ が有界なら), 「 $x \rightarrow a$ において, $f(x)$ は $g(x)$ で**押さえられている**」または「 $x \rightarrow a$ において, $f(x)$ は $g(x)$ の**オーダーである**」といい, 記号で $f(x) = O(g(x))$ と書く. (ここで, O は大文字のオーである.)

- (3) なお, 上の定義は, $x \rightarrow \infty$ や $x \rightarrow -\infty$ の場合にも流用する.

注意.

- これらの定義, 特に (2) は, 共にゼロに行く二つの函数の大きさを比較 (ゼロへの収束の速さを比較) するために導入されたものである.
- 以下の例でも示すように, $f(x) = O(g(x))$ の場合, f と g が同じくらいの大きさである必要はなく, $f(x) = o(g(x))$ でも良い.
- 等号 $=$ で両辺をつないでいるが, これはあくまで収束の速さを比べているだけのものであり, 通常の等号とは意味が異なることには注意. 例えば, $x \rightarrow 0$ に際しての $\sin x = O(3x) = O(\tan(5x))$ など ($3x, \sin x, \tan(5x)$ は $x \rightarrow 0$ での大きさもかなり異なるが, オーダーとしては同じなので, O でつないでる).

(例)

- $x \rightarrow 0$ において, $x^2 = o(x)$ かつ $x^2 = O(x)$ である.
- $x \rightarrow 0$ において, $\sin x = O(x)$ であるが, $\sin x = o(x)$ ではない.
- $x \rightarrow \infty$ において, $x = o(e^x)$, $\log x = o(x)$, $\sin x = O(1)$.

²⁸教科書 p.31

この記号を用いると, $f(x)$ が $x = a$ で微分可能なこと, つまり

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \quad (2.2.1)$$

は,

$$x \rightarrow a \text{ にて } f(x) = f(a) + f'(a) \times (x - a) + o(x - a) \quad (2.2.2)$$

と書ける. また逆に, ある定数 A が存在して

$$x \rightarrow a \text{ にて } f(x) = f(a) + A(x - a) + o(x - a) \quad (2.2.3)$$

が成り立つなら, $f(x)$ は $x = a$ で微分可能で, その微係数は A である. このような書き方はこれからも使うので慣れておいた方がよい.

2.3 平均値の定理²⁹

高校でもやったはずのロルの定理, 平均値の定理について述べよう. セッかく大学の内容なのだから, 定理の微妙な仮定 (閉区間で連続, 开区間では微分可能) に注目してほしい.

まずはロルの定理. 定理の下側の図を見れば, 直感的には明らかだろう.

定理 2.3.1 (ロル Rolle の定理) $f(x)$ が閉区間 $[a, b]$ で連続, 开区間 (a, b) で微分可能. 更に $f(a) = f(b)$ とする. このとき

$$f'(\xi) = 0 \quad (a < \xi < b) \quad (2.3.1)$$

となる ξ が存在する.

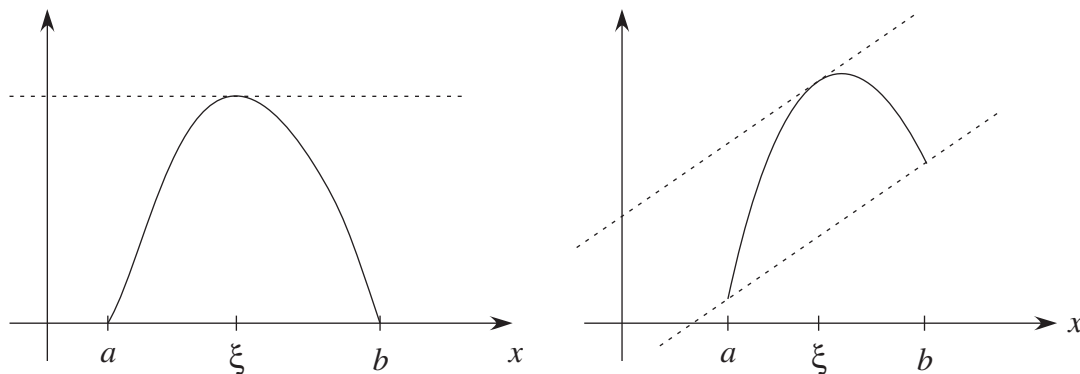
(注) 定理の ξ は一般には a, b の両方に依存して決まる. アタリマエだが, 注意の事.

(証明) $f(x)$ が定数であればいつでも $f'(x) = 0$ だから, 証明は終わっている. そこで, $f(x)$ が定数でない場合を考える. 定数でない $f(x)$ は (a, b) で正または負の値をとる³⁰ので, ある点では正をとったと仮定しよう. (負の場合は $-f(x)$ は正だから, 同じように議論できる).

ここで, **閉区間で定義された連続関数は必ず最大値, 最小値を持つ**ことを思い出そう (定理 1.6.9). その最大値をとる点 (の一つ) を ξ と書くと, ここでは f が正だから $\xi \in (a, b)$. また, ξ で最大値なんだから, ξ の周りでは $f(\xi) \geq f(x)$ である. 従って, ξ での微分係数の定義

$$f'(\xi) = \lim_{h \rightarrow 0} \frac{f(\xi + h) - f(\xi)}{h} \quad (2.3.2)$$

において, 分子はいつも非正であり, 分母は h の正負に応じて正負になっている. 従って, この極限に出ている分数は, $h > 0$ なら非正, $h < 0$ なら非負である. しかし, $h \rightarrow 0$ ということは h を正負両方の方向からゼロにする訳だから, 定理の仮定 (微分可能) にあるように極限が存在するなら, それは非負でも非正でもある. この両方を満たすのは極限がゼロの時だけだ. □



²⁹教科書 2.1 節の後半

³⁰数学用語の注: 高校でも散々聞かされたと思うが, 「正または負」というときは「正だけ」「負だけ」「正も負も」の3通りをすべて含む. この点, 日常用語とズレているので注意

ロルの定理からすぐに次の (Lagrange による) 平均値の定理が出る. これが本節の主要な結果である. 上の図では右側の状況である.

定理 2.3.2 (平均値の定理; 教科書の定理 2.5) $f(x)$ が閉区間 $[a, b]$ で連続, 开区間 (a, b) で微分可能と仮定する. このとき,

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) \quad (a < \xi < b) \quad (2.3.3)$$

となる ξ が存在する.

(注) ロルの定理と同様, 平均値の定理の ξ も一般には a, b の両方に依存して決まる.

(証明) ロルの定理を認めれば簡単だ. $g(x) = f(x) - f(a) - \frac{x-a}{b-a}\{f(b) - f(a)\}$ を作ると, ロルの定理の条件をみたす. よって, $0 = g'(\xi) = f'(\xi) - \frac{1}{b-a}\{f(b) - f(a)\}$ がなりたつ $a < \xi < b$ が存在する. $\square$

以上で平均値の定理の主要な部分はおしまいだ, 下の形の定理も有効である. 実際, 後で「テイラーの定理」の証明に用いるであろう.

定理 2.3.3 (コーシーの平均値の定理; 教科書の定理 2.6) $f(x)$ と $g(x)$ が共に閉区間 $[a, b]$ で連続, 开区間 (a, b) で微分可能とする. 更に, (a, b) では $g'(x) \neq 0$ としよう. このとき,

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} \quad (a < \xi < b) \quad (2.3.4)$$

となる ξ が存在する.

(注) $g'(x) \neq 0$ から $g(a) \neq g(b)$ は保証されている.

(証明) $k := \frac{f(b) - f(a)}{g(b) - g(a)}$ において, $F(x) := f(x) - f(a) - k\{g(x) - g(a)\}$ を考える. すると, $F(a) = F(b) = 0$, かつ F の微分可能性なども f, g の微分可能性と同じだから大丈夫なので, ロルの定理から $F'(\xi) = 0$ なる ξ が存在するといえる. これは $f'(\xi) - kg'(\xi) = 0$ を意味するので, これを書き直すと定理を得る. $\square$

以下では平均値の定理の応用を考える. これらは大まかには高校でやっていると思うので, 簡単にすませる. 平均値の定理の応用として非常に大事な (かつ, 高校ではやってない) 「テイラー展開」については後の節で考える.

2.3.1 函数の増減

微分の応用として最重要なものの一つは, 函数の増減や極大・極小との関連である. 類似の結果は高校から散々やってきているだろうから, 講義でも簡単に触れるにとどめる. ただ, 以下のように (また教科書にも強調されているように) 仮定の微妙な入り方が面白いところではある. まずは言葉の定義から始める.

定義 2.3.4 (単調な函数) 区間 (开区間でも閉区間でも) I で定義された函数 f に対して

- $x, y \in I$ かつ $x < y$ ならば常に $f(x) < f(y)$ であるなら, f は I で狭義の**単調増加**であるという.
- $x, y \in I$ かつ $x < y$ ならば常に $f(x) \leq f(y)$ であるなら, f は I で**広義の単調増加** (または, 単調非減少) であるという.
- $x, y \in I$ かつ $x < y$ ならば常に $f(x) > f(y)$ であるなら, f は I で狭義の**単調減少**であるという.
- $x, y \in I$ かつ $x < y$ ならば常に $f(x) \geq f(y)$ であるなら, f は I で**広義の単調減少** (または, 単調非増加) であるという.

なお, 単調増加な函数を単に「増加函数」, 単調減少な函数を「減少函数ともいう. また単調増加と単調減少の両方をまとめて, 「単調な」函数という.

数列のところでも注意したが、単に「単調増加」と言った場合に広義の単調増加を指すのか狭義の単調増加を指すのかは分野やレベルによる。この講義では教科書に従い「狭義の単調増加」を単に「単調増加」という事が多いだろう。

定理 2.3.5 (導函数の符号と函数の増減; 教科書の p.36 付近) $f(x)$ が开区間 $I = (a, b)$ で微分可能と仮定する。このとき、

- I で常に $f'(x) \geq 0$ $\implies$ I で $f(x)$ は広義単調増加.
- I で常に $f'(x) > 0$ $\implies$ I で $f(x)$ は狭義単調増加. (逆はなりたない.)
- I で常に $f'(x) = 0$ $\iff$ I で $f(x)$ は定数函数.

(注) 上の定理の仮定では「区間 I 全体で $f'(x) > 0$ 」などを仮定しているが、これはほとんど必要である。つまり、ある一点 a で $f'(a) > 0$ だとしても、これだけでは $x = a$ で増加しているとはいえない (例は田島本の p.135)。

(注) 狭義単調増加や狭義単調減少だからと言って、 $f'(x) > 0$ や $f'(x) < 0$ とは言い切れない (反例は $f(x) = x^3$)。

2.4 高階導函数³¹

高校でもやったと思うけど、高階の導函数についてまとめておく。

函数 $f(x)$ に対して、それを n -回微分してできる函数を n -階の導函数 (n^{th} derivative) といい、 $f^{(n)}(x)$ と書く。ただし、1 階、2 階、3 階くらいはそれぞれ $f'(x), f''(x), f'''(x)$ と書く。具体的には

$$f^{(2)}(x) = \frac{d^2}{dx^2} f(x) = \frac{d}{dx} \left\{ \frac{d}{dx} f(x) \right\}, \quad f^{(3)}(x) = \frac{d^3}{dx^3} f(x) = \frac{d}{dx} \left[\frac{d}{dx} \left\{ \frac{d}{dx} f(x) \right\} \right], \quad \dots \quad (2.4.1)$$

というわけ。なお、 $f^{(0)}(x)$ は $f(x)$ そのものを表すものと理解する (これは今後、断りなく多用する)。

高階の導函数については ライプニッツ (Leibniz) の公式 が成り立つ。つまり

$$\frac{d}{dx} \{f(x)g(x)\} = f'(x)g(x) + f(x)g'(x), \quad \frac{d^2}{dx^2} \{f(x)g(x)\} = f''(x)g(x) + 2f'(x)g'(x) + f(x)g''(x) \quad (2.4.2)$$

で、より一般には (n は自然数)

$$\frac{d^n}{dx^n} \{f(x)g(x)\} = \sum_{k=0}^n \binom{n}{k} f^{(k)}(x) g^{(n-k)}(x), \quad \binom{n}{k} := {}_n C_k = \frac{n!}{k!(n-k)!} \quad (2.4.3)$$

となる³²。この証明は数学的帰納法のできるから、一度は自力でやっておくこと。ただし、その途中で恒等式

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \quad (2.4.4)$$

を用いることは注意しておく。(この恒等式の意味は何だろう? 順列組み合わせで考えてみよう。)

(用語) ある开区間 I で定義された函数 $f(x)$ が n 回微分可能で、更に $f^{(n)}(x)$ が連続のとき、この函数は开区間 I で C^n -級である、という。いうまでもなく、 $m < n$ ならば、 C^n -級の函数は C^m -級でもある。

(注) 「**連続性は遺伝しない**」とは高木貞治の名言である。つまり、連続な函数の導函数は連続とは限らない。このような例はいくらでも作れるから、各自で作って納得しておくこと。

³¹教科書 2.2 節前半

³²この公式は 2 項展開の公式 $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$ に良く似ている。その導出法を思い出すと、同じ二項係数 $\binom{n}{k}$ が出る理由がわかるだろう

2.5 テイラーの定理とテイラー展開³³

これから暫く、微分の重要な応用のもう一つ、「テイラー展開」を扱う。これは案外、皆さん苦勞するようだから、甘く見ないように。

「テイラー展開」とは大雑把にいうと、 $f(x)$ の値を $f(a)$ とその高階微係数で表す表式で、

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n \quad (2.5.1)$$

という形をしている（この表式の成立条件は後でじっくりやる）。皆さんの良く知っている函数の例では（上で $a=0$ としたものを書いた）

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!} x^n \quad (2.5.2)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} \quad (2.5.3)$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \quad (2.5.4)$$

などとなる。

これはある種、驚異的な式である。高校から知ってたはずの函数が、上のような変な級数（和）で書けるというのだ。物事を深く考える人ほど、初めはこの式に違和感を持つものと思う。特に変なのは $\sin x$ と $\cos x$ であって、上の表式からは $\sin x$ と $\cos x$ が周期 2π の周期函数である事が全く自明ではない！（ $\sin \pi = 0$ が上の式から見えますか？）

しかし、後で証明するように、上の3つの式はもちろんすべて正しい。 $\sin x$ や $\cos x$ の周期性は暫く各自で考えてもらうことにして、テイラー展開の持ちうる意味（意義）について簡単に述べておこう。

- まず、(2.5.2) などの式は、それ自身が数値計算にも適している —— $e^x, \sin x$ などの値を、右辺の級数（和）で計算できるのだ。もちろん、無限級数の値そのものを数値的に求める事はできないが、たくさんの項の和をとる事で、いくらでも精度良く計算できる³⁴。
- (2.5.1) にはもう少し理論的な意味もある。つまり、 $|x-a|$ が小さい場合に $f(x)$ を $f(a)$ で近似すると、誤差がどうなるかを表していると解釈できる。この誤差の評価は、もっと進んだ結果を得るのに不可欠である。

以下、このテイラー展開について詳しく述べる。まずはおおとの「テイラーの定理」から始めよう。

2.5.1 テイラーの公式（有限項でとめた形）

通常、テイラーの定理（テイラーの公式）というのは以下の形の定理をいう：

定理 2.5.1 (通常 of テイラーの公式) $f(x)$ がある開区間 I で n 回微分可能と仮定し、この区間内に $a \in I$ をとろう。このとき、勝手な $x \in I$ に対して、 a と x の間の一点 ξ が存在して以下が成り立つ：

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k + \frac{f^{(n)}(\xi)}{n!} (x-a)^n \quad (2.5.5)$$

なお、(2.5.5) の2つの項に名前をつけて

$$f(x) = S_n(x) + R_n(x), \quad (2.5.6)$$

³³教科書 2.2 節後半

³⁴実際にコンピューターが $e^x, \sin x$ などを計算する場合には、上の (2.5.2) そのものではなく、これを更に効率よくしたものを用いる。しかし、計算の原理は（大体）同じである

$$S_n(x) := f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k, \quad R_n(x) := \frac{f^{(n)}(\xi)}{n!} (x-a)^n \quad (2.5.7)$$

と書く事もある。\$S_n(x)\$ をテイラー展開 (テイラーの公式) の \$n\$ 次の主要項, \$R_n(x)\$ を \$n\$ 次の剰余項 (残項) という。

- 上のテイラーの公式を「\$f(x)\$ の \$x=a\$ の周りのテイラーの公式」という, 例えば, \$a=3\$ なら「\$f(x)\$ の \$x=3\$ の周りのテイラーの公式」という。場合によっては「\$x=\dots\$ の周りの」の代わりに「\$x=\dots\$ において」とかと「\$x=\dots\$ での」と言うこともある。つまり, 「\$x=2\$ におけるテイラーの公式」や「\$x=2\$ でのテイラーの公式」などともいう。
- \$a=0\$ とした場合の展開を特にマクローリン (Maclaurin) の公式 (展開) ともいう。
- 実はマクローリンの公式とテイラーの公式は非常に近い親戚関係にあり, 片方だけわかれば十分だ。理由は以下の通り: \$y=x-a\$ という変数変換によって, 座標 \$x\$ で見た時の点 \$x=a\$ は座標 \$y\$ で見た時の \$y=0\$ に移る。従って, 座標 \$y\$ でのマクローリンの公式は座標 \$x\$ での \$x=a\$ の周りのテイラーの公式に対応している。
- テイラーの公式でも, 平均値の定理でも, \$\xi\$ は \$a\$ と \$x\$ (または \$b\$) の両方に依存しうすることを再度強調しておく。同じ理由で, 剰余項 \$R_n(x)\$ は \$x, a\$ で決まるけども, \$R_n(x)\$ の \$\xi\$ そのものが \$x, a\$ に依存する事をお忘れなく。
- 細かいことであるが, 定理 2.5.1 では \$f^{(n)}(x)\$ の存在は仮定するが, 連続性は仮定しなくても良い。この点で, 剰余項が積分形の定理 2.5.6 (後出) より, こちらの方が少しだけ適用範囲はひろい (そのぶん, 誤差評価は大抵, 劣る — 「ある \$\xi\$ が存在して」とか言われても, どんな \$\xi\$ かわからなければ細かい評価はできない)。

定理 2.5.1 の証明³⁵

$$F(x) := f(x) - \left[f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k \right], \quad G(x) := (x-a)^n \quad (2.5.8)$$

とおく。\$F(x)\$ が (2.5.6) の \$R_n(x)\$ の表式で書けることを示せばよい。

そのために, コーシーの平均値の定理 (定理 2.3.3) を \$F, G\$ に適用する事を考えよう。\$F(x)\$ は \$f(x)\$ から \$(x-a)^k\$ の和を引いているだけなので, また \$G(x)\$ は多項式なので, 共に \$n\$ 階は微分できる。微分を具体的に計算すると

$$F(a) = F'(a) = F''(a) = \dots = F^{(n-1)}(a) = 0, \quad F^{(n)}(a) = f^{(n)}(a) \quad (2.5.9)$$

$$G(a) = G'(a) = G''(a) = \dots = G^{(n-1)}(a) = 0, \quad G^{(n)}(a) = n! \quad (2.5.10)$$

となっている。この事実を用いて, 以下のように進む。

(1) 定理 2.3.3 そのもので

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(\xi_1)}{G'(\xi_1)} \quad (2.5.11)$$

を満たす \$\xi_1\$ の存在 (\$\xi_1\$ は \$a\$ と \$x\$ の間にある) が言える。

(2) 上の右辺の量は \$F'(a) = G'(a) = 0\$ を用いて強引に書き直すと, 定理 2.3.3 が使える。その結果,

$$\frac{F'(\xi_1)}{G'(\xi_1)} = \frac{F'(\xi_1) - F'(a)}{G'(\xi_1) - G'(a)} = \frac{F''(\xi_2)}{G''(\xi_2)} \quad (2.5.12)$$

を満たす \$\xi_2\$ の存在 (\$\xi_2\$ は \$a\$ と \$\xi_1\$ の間にある) が言える。

(3) この議論は, \$F^{(k)}(a) = G^{(k)}(a) = 0\$ である限り, つまり \$k \leq n-1\$ である限りくりかえす事ができて,

$$\frac{F^{(k)}(\xi_k)}{G^{(k)}(\xi_k)} = \frac{F^{(k)}(\xi_k) - F^{(k)}(a)}{G^{(k)}(\xi_k) - G^{(k)}(a)} = \frac{F^{(k+1)}(\xi_{k+1})}{G^{(k+1)}(\xi_{k+1})} \quad (2.5.13)$$

³⁵正直, 僕は高校の頃からこの定理の証明がどうもすんなりできないままである。典型的な証明は以下に述べる「コーシーの平均値の定理」を使うもので, それは理解できるものの, どうも回りくどい気がして仕方ない。そこで, 微積の講義を受け持った際に「コーシーの平均値の定理」を使わない証明を何度か試みるのだが, いつもうまくいかないのだ。仕方ないので, 「コーシーの平均値の定理」を用いるバージョンを載せておく (高木本からのカンニング)

を満たす ξ_{k+1} の存在 (ξ_{k+1} は a と ξ_k の間にある) が, $k \leq n-1$ で順次, 証明される.

(4) 以上をまとめると,

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F^{(n)}(\xi_n)}{G^{(n)}(\xi_n)} \quad (2.5.14)$$

を満たす ξ_n の存在 (ξ_n は a と x の間にある) が, 証明された. この両辺を具体的に計算すると

$$\frac{F(x)}{(x-a)^n} = \frac{f^{(n)}(\xi_n)}{n!} \quad (2.5.15)$$

となっているので, 分母を払うと定理が得られる. □

2.5.2 テイラー展開 (無限項まで)

定理 2.5.1 において, 公式 (2.5.6) がすべての $n \geq 1$ で成り立ち, かつ剰余項 $R_n(x)$ が $n \rightarrow \infty$ でゼロになるならば, つまり, $\lim_{n \rightarrow \infty} R_n(x) = 0$ ならば,

$$f(x) = \lim_{n \rightarrow \infty} S_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k \quad (2.5.16)$$

が得られる.

このところ, $\lim_{n \rightarrow \infty} S_n$ が存在するのかどうか気になる人がいるかもしれないが, それは「 R_n の極限がゼロ」の仮定の下では以下のように保証される: (2.5.6) の左辺は n に依存せず, 右辺では R_n がゼロに行く. 従って, 残りの S_n の $n \rightarrow \infty$ 極限が存在して, かつその極限は左辺の $f(x)$ に等しくなければならない.

このように無限級数の形になったものをテイラー展開またはテイラー級数とよび, 有限項の「テイラーの公式」と区別する. なお, 剰余項 $R_n(x)$ が $n \rightarrow \infty$ でゼロになるか否かは展開される函数 f と考えている区間 I に依存するので, 個別に考察する必要がある. この問題は個々の例で見て行こう.

2.5.3 テイラーの公式, テイラー展開の例

まずは具体例を見てみよう. もう少し「理論的」なことは後で詳しく見る.

- まず, 多項式. $f(x) = c_n(x-a)^n + c_{n-1}(x-a)^{n-1} + \dots + c_1(x-a) + c_0$ は何回でも微分可能であり, 既にテイラー展開の形になっている. 念のため, テイラーの公式を用いたら多項式が再現される事を各自で確かめてみよう.
- 指数函数. $f(x) = e^x$ は何回でも微分可能で, 高階の導函数もすべて e^x である. 従って, 特に $a=0$ としたテイラーの公式から

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x), \quad R_n(x) := \frac{e^\xi}{n!} x^n \quad (2.5.17)$$

が得られる (ξ は 0 と x の間の数). 更に, 少しややこしい計算を頑張ってやると, すべての実数 x に対して $\lim_{n \rightarrow \infty} R_n(x) = 0$ が証明できる (レポート問題). 従って, すべての実数 x に対して

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{x^k}{k!} \quad (2.5.18)$$

が成り立つ. このテイラー級数の形は非常に基本的だから, 覚えておくことが望ましい.

- 三角関数 ($\sin, \cos$) も同様にして展開式を導くことができる。例えば

$$\sin x = S_n(x) + R_n(x), \quad S_n(x), R_n(x) \text{ の形はレポートでね} \quad (2.5.19)$$

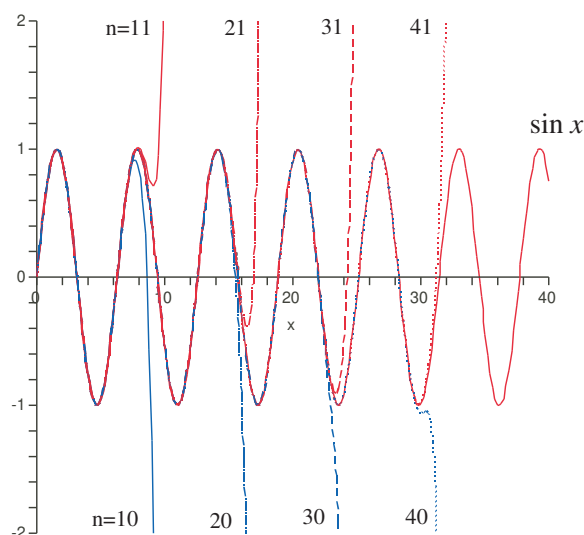
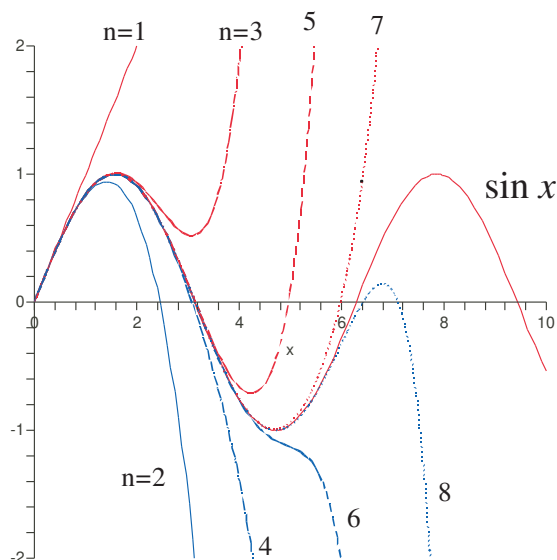
がなりたつ。指数関数と同様に、この場合もすべての実数 x に対して $\lim_{n \rightarrow \infty} R_n(x) = 0$ が証明できる (レポート問題)。従って、すべての実数 x に対して

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \quad \text{また同様の考察により} \quad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \quad (2.5.20)$$

が成り立つことがわかる。このテイラー級数の形も覚えてしまうくらいになろう³⁶。

参考までに $\sin x$ のテイラー展開の図を載せておく。下の左図は、 $n = 1, 2, \dots, 8$ の $y = S_n(x)$ の様子を、 $y = \sin x$ のグラフ (実線) とともに書いたもの。 n が奇数のものはいつも正の方に大きくなって視界から消えている。一方、 n が偶数のものは負の方に大きくなって視界から消えていく。

右図は $n = 11, 21, 31, 41$ と $n = 10, 20, 30, 40$ の様子を、 $y = \sin x$ とともに書いたもの。 n が増えるにつれて、近似はどんどん良くなっていくが、ある x から先では急速にダメになって上下に離れてしまう様子が見て取れる。



2.5.4 テイラーの公式の意味 (関数の近似)

そもそも、テイラーの公式は

よく訳のわからない関数 $f(x)$ を、訳のわかっている関数 $(x-a)^k$ の和 (多項式) $S_n(x)$ で書く

いう精神の下に生まれたものである。つまり、後述する条件の下では、(2.5.6) での $S_n(x)$ が $f(x)$ を良く近似し、 $R_n(x)$ の方は小さな誤差項とみなせるのだ。

また、関数の種類によってはテイラーの公式を杓子定規に使うよりも簡単な方法もある [例: $f(x) = 1/(1-x)$]。しかしそのように「ずるい」方法がテイラーの公式を杓子定規に使ったものと同じかどうかは現時点ではまだわからない。

このような事情を明確にするため、以下の考察を行う。まずは「関数を近似する」とはどういう事かをはっきりさせよう。

³⁶このような公式は無理に丸暗記してもダメだ。自分で導出した、実際に使ってみるうちに自然に覚えるようになるのが望ましい

定義 2.5.2 (n 次より高く近似) $x = 0$ の近くで定義された函数 $f(x), g(x)$ と正の整数 n があり,

$$x \rightarrow 0 \text{ のときに } \lim_{x \rightarrow 0} \frac{f(x) - g(x)}{x^n} = 0 \quad \text{つまり, } f(x) - g(x) = o(x^n) \quad (2.5.21)$$

となるとき, 0 の近くで $g(x)$ は $f(x)$ を n 次より高く (n 次よりも良く) 近似する という.

上の式では, $f(x) - g(x)$ はゼロに行くのだが, その行き方 (ゼロへの収束の速さ) が, x^n よりも速い, と言っている.

この用語法に従うと, テイラーの定理を以下のように言い換えることができる.

命題 2.5.3 (テイラーの定理の言い換え) 函数 $f(x)$ の $x = 0$ を中心とした n 次のテイラーの公式

$$f(x) = S_n(x) + R_n(x), \quad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k, \quad R_n(x) := \frac{f^{(n)}(\theta x)}{n!} x^n \quad (0 < \theta < 1) \quad (2.5.22)$$

において, $S_n(x)$ が $f(x)$ を $(n-1)$ 次より高く近似する, つまり

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + o(x^{n-1}) \quad (2.5.23)$$

となるための必要充分条件は以下の通り:

$$\lim_{x \rightarrow 0} \frac{R_n(x)}{x^{n-1}} = 0 \quad (2.5.24)$$

前の命題の (2.5.24) の十分条件として, 以下がある.

命題 2.5.4 (多項式近似の十分条件)

1) 0 を内部に含むある区間で $f^{(n)}$ が有界, つまり $\delta > 0$ と $M > 0$ があって (これらは n に依存してもよいが, x に依存してはいけない),

$$|x| < \delta \text{ ならば } |f^{(n)}(x)| < M \quad (2.5.25)$$

となっているとする. このとき,

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + O(x^n) \quad (2.5.26)$$

である.

2) 0 を内部に含むある区間で $f^{(n)}$ が連続, つまりこの区間で $f(x)$ が C^n -級なら, 1) のためには十分である.

これらはわざわざ命題とするほどのことではないかもしれないが, 実用上大事だから載せた. 特に, 一年生で出てくる函数は C^∞ -級 (何回でも微分できる) のものが多く, これらに対しては上の十分条件が自動的に満たされており, 命題 2.5.3 の結論も成り立つのである.

さて, テイラー展開 (より一般に函数を級数で近似すること) については, 以下の非常に重要な性質がある. これはほとんどアタリマエだが, テイラーの公式を直接使わずに S_n を求める方法の基礎を与えてくれる.

命題 2.5.5 (多項式近似の一意性) 原点の近くで定義された函数 $f(x)$ と多項式 $g(x) = \sum_{j=0}^n a_j x^j$ があり, $g(x)$ は $f(x)$ を n 次より高く近似しているものとする. このとき, $g(x)$ の係数 $a_0, a_1, \dots, a_n$ は一意に決まる.

テイラーの公式があることを考えると, 要するに a_j はテイラーの公式にでてくる係数と一致しなければならない事がわかる.

証明:

f を n 次より高く近似する g が 2 つあったとして、それらを

$$g_1(x) = \sum_{j=0}^n a_j x^j, \quad g_2(x) = \sum_{j=0}^n b_j x^j \quad (2.5.27)$$

とする. $a_j = b_j$ ($0 \leq j \leq n$) を示したい. さて ($x \neq 0$),

$$\frac{|g_1(x) - g_2(x)|}{|x|^n} \leq \frac{|g_1(x) - f(x)|}{|x|^n} + \frac{|f(x) - g_2(x)|}{|x|^n} \quad (2.5.28)$$

の両辺で $|x| \downarrow 0$ とすると,

$$\lim_{x \rightarrow 0} \frac{|g_1(x) - g_2(x)|}{|x|^k} = 0 \quad (0 \leq k \leq n) \quad (2.5.29)$$

がなりたつ. そこで

$$g_1(x) - g_2(x) = \sum_{j=0}^n (a_j - b_j) x^j \quad (2.5.30)$$

である事に注目して $k = 0, 1, 2, \dots$ に対して (2.5.29) を順次考えると, $a_k - b_k = 0$ しかあり得ないことが, $k = 0, 1, 2, \dots$ と順次わかる. $\square$

この命題から, 与えられた函数 $f(x)$ のテイラーの公式 (S_n の方のみ考える) を求めるには, どのようなやり方でも良いから $f(x)$ を $(n-1)$ 次よりも高く近似するものを見つければよいことがわかる.

また, 複雑な函数のテイラー展開 (の最初の何項か) を求めるには, いろいろな工夫が可能になる. 例えば, $\tan x$ を $x = 0$ の周りで展開する場合, まともに微分して行くとなかなか大変だ. でも $\tan x = \frac{\sin x}{\cos x}$ を利用して, 分母と分子を別々にテイラー展開した後で割り算 (これにも工夫が必要だが) するのが良い. これは教科書の p.50 問 13 などに載ってるから見ておくこと (レポートでも出題予定).

2.5.5 テイラー展開の効用

テイラーの公式とテイラー級数の効用については既に述べたが, 重要なのもう一度繰り返す.

1. テイラーの公式では, 剰余項以外は単なる級数 ($(x-a)^n$ の和) で, 四則演算で計算できる. 剰余項を何らかの工夫で押さえれば, 問題の函数の値の近似値を計算できる. その例をレポート問題に与える予定なので, やってみてほしい.
2. テイラー展開 (無限級数の形) が成立するならば, テイラー展開によって函数を定義するのだと考え直すこともできる. そうすれば, その級数をより広い x に拡張して適用することにより, 函数の定義域を一気に広げることにも可能である. これは特に, 「いままで実数だと思ってきた x を複素数に拡張する」場合に非常に有効である. この一つの例 (オイラーの公式) を下に示した.

2.5.6 オイラーの公式

2 番目の効用の例として (多分, どこかで見ただろう) オイラー (Euler) の公式

$$e^{i\theta} = \cos \theta + i \sin \theta, \quad \theta \in \mathbb{R} \quad (2.5.31)$$

を挙げておこう. 指数函数のテイラー展開において, $x = i\theta$ とおいてしまおう (このようににおいてもテイラー展開が収束することは, 実部と虚部を別々に考察すると確かめられる). すると,

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell+1}}{(2\ell+1)!} \quad (2.5.32)$$

が得られる (2 番目の等号は, 単に k が偶数の場合と奇数の場合をわけて, i^k を計算しただけ). ところがこの最右辺は $\cos \theta + i \sin \theta$ のテイラー展開に他ならない. 従って, **指数函数や三角函数はそのテイラー展開の式で定義し直**

すのだと思えば、オイラーの公式が証明されたことになる。テイラー展開によって関数を定義し直すというのは一見、奇妙に思えるかもしれないが、同値な命題がある場合にどれを仮定（公理）にしてどれを結論とするか、の一例と思えば良い。ただし、本当に定義し直す立場をとった場合は今まで知っていたはずの関数の性質（例： $\sin, \cos$ は周期 2π である、指数関数は $e^{a+b} = e^a e^b$ を満たす、等々）はすべて忘れて、テイラー展開だけからこれらを導き直す必要はある。この辺りは夏休みチャレンジ問題とする予定なので、我こそはと思う人は挑戦して欲しい。

2.5.7 おまけ：剰余項が積分の形のテイラーの定理

今までのものの他に、テイラーの公式には以下のようなバージョンもある。これは剰余項を積分で書くもので、剰余項の大きさを評価するには楽な事が多い。（大体、微分よりは積分の方が評価しやすいのである——これは皆さんがもう少し勉強するとわかってくるだろう）。ただ、これは積分を使っているから（そして、我々は積分の厳密な理論をまだやっていないから）現時点ではこの定理の完全な証明を与えるわけにはいかない。

定理 2.5.6 (剰余項が積分形のテイラーの公式) $f(x)$ がある开区間 I で C^n -級であると仮定する。この区間 I 内に $a \in I$ をとろう。このとき、勝手な $x \in I$ について、以下が成り立つ：

$$f(x) = S_n(x) + R_n(x), \quad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!} (x-a)^k, \quad R_n(x) := \int_a^x \frac{f^{(n)}(y)}{(n-1)!} (x-y)^{n-1} dy \quad (2.5.33)$$

（高校のノリでの証明；ただし積分の基礎付けさえすれば、この証明は厳密に正しい）数学的帰納法で証明する。つまり $f(x)$ は C^N -級と仮定し、(2.5.33) をすべての $n \leq N$ について証明することを目指す。それで n についての帰納法を用いる。

I. $n=1$ では、 $\int_a^x f'(y) dy = f(x) - f(a)$ であるから、 $f(a)$ を移行すれば証明できる—— $f^{(0)}(x) := f(x)$ の記号法を思い出せ。

II. $n=2$ の場合（これは証明には必要ないが、ウォームアップとしてやる）。 $n=1$ の

$$f(x) = f(a) + \int_a^x f'(y) dy \quad (2.5.34)$$

の第2項を、以下のように部分積分するとよい。

$$\begin{aligned} \int_a^x f'(y) dy &= \int_a^x \left\{ -\frac{d}{dy} (x-y) \right\} f'(y) dy = \left[-(x-y) f'(y) \right]_a^x + \int_a^x (x-y) f''(y) dy \\ &= (x-a) f'(a) + \int_a^x (x-y) f''(y) dy \end{aligned} \quad (2.5.35)$$

II. n まで証明できたとして、 $n+1$ をやってみよう（もちろん、 $n \leq N-1$ と仮定しておく）。 n までできたと仮定したので、(2.5.33) が成り立っているが、最後の項を以下のように考えて部分積分する（分母の $(n-1)!$ は後で）：

$$\begin{aligned} \int_a^x f^{(n)}(y) (x-y)^{n-1} dy &= \int_a^x f^{(n)}(y) \left\{ -\frac{1}{n} \frac{d}{dy} (x-y)^n \right\} dy \\ &= -\frac{1}{n} \left[f^{(n)}(y) (x-y)^n \right]_a^x + \frac{1}{n} \int_a^x f^{(n+1)}(y) (x-y)^n dy \\ &= \frac{1}{n} f^{(n)}(a) (x-a)^n + \frac{1}{n} \int_a^x f^{(n+1)}(y) (x-y)^n dy. \end{aligned} \quad (2.5.36)$$

これを (2.5.33) の最後の項に用いると（もちろん、分母の $(n-1)!$ を忘れない）、(2.5.33) の $n+1$ のものが証明されてメデタシメデタシ。□

2.6 関数の極値と凹凸³⁷

関数の極地や凹凸について、簡単にまとめておく。ただ、高校で十分にやってるはずなので、この講義ではほとんど触れない。

2.6.1 関数の極大・極小

定義 2.6.1 点 $x = a$ が函数 $f(x)$ の極大点 (local maximum) であるとは、

$$\exists r > 0, \quad 0 < |x - a| < r \implies f(x) < f(a) \quad (2.6.1)$$

となることである。このとき、 f は $x = a$ で極大、ともいう。同様に、点 $x = a$ が函数 $f(x)$ の極小点 (local minimum) であるとは、

$$\exists r > 0, \quad 0 < |x - a| < r \implies f(x) > f(a) \quad (2.6.2)$$

であることをいう。なお、

$$\exists r > 0, \quad |x - a| < r \implies f(x) \leq f(a) \quad (2.6.3)$$

となっている時 (最後の不等号に等号を許す)、 f は a で広義の極大という。広義の極小も同様に定義する。

(注) 高校でも強調されたかもしれないが、函数 $f(x)$ が $x = a$ で最大 (maximum) とは、 f の定義域全体を見渡した時に $f(a)$ が最大であることをいう。つまり、

$$f \text{ の定義域に入っているすべての } x \text{ に対して } f(x) \leq f(a) \quad (2.6.4)$$

であることをいう (上の極大の定義のように x の範囲を我々が勝手に設定してはいけない)。最小 (minimum) についても同様である。要するに極大・極小とは local な性質、最大、最小とは (全体を見渡した時の) global な性質である。この点は英語の方が良く表現されている。

実際問題として、極大や極小を求めるのは (みんなが高校で習ったように) 割合簡単なことが多い。それに引き換え、最大や最小を求めるのはなかなか大変なことが多く、すべての極大点や極小点を探し出した上でそれらの中で最大や最小のものを求める、という 2 段階が必要になる。(場合によっては、境界での値も考えに入れなければならない。) この節では最大最小は避けて、極大・極小問題に重点をおきたい。

さて、1 変数の場合の極大、極小問題は以下のようにになっている。この結果そのものは高校でやったはずだが、今では厳密に証明できるようになったから、再録する。

定理 2.6.2 $x = a$ の近傍で定義された 1 変数の函数 $f(x)$ について、以下が成り立つ。

- (i) $f(x)$ が $x = a$ で微分可能、かつ $x = a$ で $f(x)$ が極大または極小の場合、 $f'(a) = 0$ である。逆は必ずしもなりたたない。
- (ii) $f(x)$ が $x = a$ で 2 階微分可能で $f'(a) = 0$ の場合には、以下が成り立つ：
 - a. $f''(a) > 0$ の場合、 $f(x)$ は $x = a$ で極小である。
 - b. $f''(a) < 0$ の場合、 $f(x)$ は $x = a$ で極大である。
 - c. $f''(a) = 0$ の場合、 $f(x)$ の $x = a$ での極大極小については何も言えない (極大の場合、極小の場合、どちらでもない場合もある。これについては、教科書の定理 2.12 を参照のこと)。

(上の定理の (ii)-c は「定理」の中に入れるほどのことではないが、わかりやすさを考えて入れておいた。)

講義ノートにはこれ以上書かないが、各自でいくつかの計算問題はやっておくこと (受験数学の復習みたいなものだが)。

³⁷教科書の 2.3 節

2.6.2 曲線の凹凸

これまた高校でもやったはずだが、2階導関数の幾何学的意味を復習しておこう。

1階導関数 $f'(x)$ は x での $f(x)$ の変化率 (増減) を表すので、 $y = f(x)$ のグラフの傾きを表す。

それに対して、2階導関数 $f''(x)$ は $f'(x)$ の増減を表し、これは $y = f(x)$ のグラフの曲がり具合に対応している。つまり、 $f''(x) > 0$ ならば x でのグラフは下に尖っている (これを下に凸という)。 $f''(x) < 0$ ならば x でのグラフは上に尖っている (これを上に凸または凹という)。 f' と f'' の正負を調べてグラフを書くことは高校のときに散々やっただろうから、詳細は省く。

用語についての注意： 英語では下に凸の函数を単に convex function (直訳：凸函数) といい、上に凸の函数を concave function (直訳：凹函数) とよぶ。日本人にとっては不幸なことに、函数の凹凸に関する用語が、漢字から受ける印象と逆になってしまっている。

3 多変数関数の微分

これから、偏微分（多変数関数の微分）を扱う。厳密性にはあまりこだわらず、高校までの「ええ加減」なノリで³⁸、とにかく概念を理解する事を目標にしよう。

なお、たいていの場合には2変数の関数を扱う。最後に扱う「極大・極小問題」を除いては、3変数以上への拡張は容易かつ自明である。

3.1 1変数の関数, 多変数の関数³⁹

関数とは何か、の復習から始めよう。高校でもいろんな「関数」をやったはずだ。例えば、

$$x^2, \quad x^4, \quad \sin x \quad \cos(x^2 + 2), \quad \dots \quad (3.1.1)$$

要するに1変数の関数 $f(x)$ とは、実数の変数 x に対して $f(x)$ という実数値が決まるもの（実数値を決める規則）であった。

（余談）この定義通り、「関数」とは何でも良く、高校までの常識からは関数に見えないようなもの——例えば、そのグラフがガザガザすぎて描けないようなもの——も入る。ただ、あまり一般的すぎると病的な関数も入ってくるので、どのような関数なら扱えるか（どのような関数を扱いたい）を見極める事が重要になってくる。近代の微分積分学（より一般に解析学）の大きなテーマは「一般の関数とは何か？その関数に対して有効な微分や積分の概念は何か？」を見極める事であった。

1変数関数と同じノリで多変数の関数を考えるが、その前に1次元での記号を整理しておこう。

- 実軸上の点は x や y のように書く。すべての実数からなる集合を $\mathbb{R}$ と書く。
- 1変数関数 f の x での値は $f(x)$ と書く。
- 点 x と y の間の距離を $\rho(x, y) = |x - y|$ （普通の絶対値）により定義する。

とする。ここまでは高校と同じだが、強いて言えば、 $|x - y|$ という絶対値を2点 x, y の間の距離と解釈することが目新しいかもしれない。「差の絶対値は距離」という見方はこれからも頻出する、非常に重要なものである⁴⁰。

では、**2変数の関数**にうつる。2変数 x, y の関数とは、2つの変数の値 (x, y) に対して $f(x, y)$ という値を定めるもののことをいう。2つの変数 x, y が勝手に動くと、 (x, y) は2次元の xy -平面全体を動く。この意味で2変数の関数は変数の空間が1次元から2次元になった拡張である。

n 変数の関数は以下のように定義される ($n \geq 2$)。一般の n で考えにくい人は $n = 2$ (平面)、 $n = 3$ (空間) を思い浮かべれば十分だ⁴¹。

- まず、 n 次元空間の点を、 $\mathbf{x}$ のように太字で書く： $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 。高校まではベクトルは矢印で書いたと思うが、大学（初年度）では太字で書く。（もっとと学年が進むと太字ですら書かず、普通の細字で書く）。
- n 次元空間は $\mathbb{R}^n$ と表す： $\mathbb{R}^n := \{\mathbf{x} = (x_1, x_2, \dots, x_n) \mid x_1, x_2, \dots, x_n \in \mathbb{R}\}$ 。

³⁸ここで「ええ加減」と書いたが、高校での数学を馬鹿にしているのではない。特に昨今の厳しい状況の中でも数学の神髄を伝えようと努力されている高校の先生方には深い尊敬の念を抱いている。また、物事は最初は大抵「ええ加減」であるが、この「ええ加減」な時代の精神は後々まで重要である——大学の数学が難しく感じられる理由は、当初の精神を忘れて形式的にだけ厳密になろうとするからかもしれない。従って「ええ加減」というのは決して悪い意味ではないことを強調しておく

³⁹教科書の3.1節、3.2節に相当

⁴⁰いつの時代からか、「絶対値はとにかく場合分けして外せ」と受験数学では指導するようになったようだ。場合分けして外せば良い場合も多いが、これでは「差の絶対値は距離」という見方が育ってくれないだろう。大学生になったらむやみに絶対値を外すのではなく、まずは「差の絶対値は距離」という見方をしてみよう

⁴¹大学の数学では $n = 2, 3$ などの例を省いて、いきなり一般の n の式が出てくる事がある。これは本来は $n = 2, 3$ を考えた結果として一般の n が出てくるのだが、そのすべてを書くのが面倒なので一般の n のみを書いていることが多い。もし一般の n に困難を覚えた場合はためらわずに $n = 1, 2, 3$ くらいを具体的に書き下してみるべきである。この注意（一般の n の式は具体的に書き下す）は以下では繰り返さないが、大学におけるすべての数学の講義において有効なはずだから労力を惜しまない事

そして

定義 3.1.1 D を n 次元空間 $\mathbb{R}^n$ の部分集合とする: $D \subset \mathbb{R}^n$. 定義域が D である n 変数の実数値関数 f とは, D の各点 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ に対して「関数の値」 $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$ を定める対応関係のことである.

さらに 1 次元の時に倣って

- 点 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ と $\mathbf{y} = (y_1, y_2, \dots, y_n)$ の間の距離を

$$\rho(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \left(\sum_{j=1}^n (x_j - y_j)^2 \right)^{1/2} \quad (3.1.2)$$

により定義する. これは $n=2$ の時には普通の平面での距離, $n=3$ の時には 3 次元空間での距離である.

- ただし, いつでも上のように x_1, x_2, x_3 などとしているとかえって書きにくいこともあるので, 適宜 $\mathbf{x} = (x, y)$, $\mathbf{z} = (u, v)$ などとも書く. (だいたい, 『空間内の点』における関数の値』のような幾何学的視点を強調するときには $f(\mathbf{x})$ と書く. それに比べて, $\mathbf{x}$ の個々の成分の関数であることを強調したいときには $f(x, y)$ などと書く.)
- なお, $\mathbb{R}^n$ の部分集合で開集合でかつ連結なものを領域 (domain, region) という. (あんまり気にしないで良い.)

以上の準備の下に, これから関数の極限を考える. まずは 1 変数の場合を思い出そう.

定義 3.1.2 (1 変数関数の極限) x の関数 $f(x)$ に対して $\lim_{x \rightarrow a} f(x) = \alpha$ とは, 以下が成り立つことをいう.

$$|x - a| \rightarrow 0 \quad \text{ならば} \quad |f(x) - \alpha| \rightarrow 0 \quad (3.1.3)$$

これは「 x と a の距離がゼロになる極限では, $f(x)$ と $f(a)$ の距離もゼロになる」ということだ. これを素直に拡張して, 多変数関数の極限を定義すると以下ようになる.

定義 3.1.3 (多変数関数の極限) n 変数関数 $f(\mathbf{x})$ に対して $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = \alpha$ とは, 以下が成り立つことをいう.

$$\|\mathbf{x} - \mathbf{a}\| \rightarrow 0 \quad \text{ならば} \quad |f(\mathbf{x}) - \alpha| \rightarrow 0 \quad (3.1.4)$$

1 変数の時の $|x - a| \rightarrow 0$ の条件が, $\|\mathbf{x} - \mathbf{a}\| \rightarrow 0$ に変わっただけで, どちらも「2 点の距離がゼロに行く」極限で「関数の値が α に行く」という状況を考えている.

注意: 2 変数以上が 1 変数と違うところ: $\|\mathbf{x} - \mathbf{a}\|$ というのは 2 点 $\mathbf{a}$ と $\mathbf{x}$ の距離であるから, これがゼロに行く行き方は非常に多様である. 1 変数のときですら, $x \rightarrow a$ とは x が a の大きい方から近づくか, 小さい方から近づくか, または a をまたぐ様にして振動しながら近づくか, などの自由度があったが, 2 変数以上では比べ物にならないほど大きな自由度を持ってしまったことには注意しておこう. (上の定義に従えば, $\mathbf{x}$ が $\mathbf{a}$ へどのような近づき方をしても $f(\mathbf{x})$ が同じ α という値に近づくときのみ, 極限が存在するという.)

この極限の定義を使うと, n 変数関数の連続性は以下のように定義される.

定義 3.1.4 (多変数関数の連続性の定義) n 変数関数 $f(\mathbf{x})$ が $\mathbf{x} = \mathbf{a}$ で連続 とは, $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$ となることである.

要するに 1 変数の場合と形式的にはまったく同じだが, 上で注意したように $\mathbf{x} \rightarrow \mathbf{a}$ の中身 (近づき方の自由度) が非常に大きい事に注意しよう.

(慣れないうちは n この変数をまとめて $\mathbf{x}, \mathbf{a}$ のように書かれるとわかりにくいかもしれない. しかし, このような幾何的な見方が後々重要になってくるので, 慣れてもらおうつもりで敢えて書いてみた.)

3.2 偏微分⁴²

さて、いよいよ偏微分を考えよう。これからは n 変数のそれぞれをあらわに書いた方が楽なので、 $f(x, y)$ のような書き方に戻る。また、一般の n 変数のときには式がいたずらに複雑になるので、主に 2 変数の場合を考える。

定義 3.2.1 (偏微分係数) 2 変数関数 $f(x, y)$ の点 (a, b) における第 1 変数に関する偏微分係数とは極限

$$\lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h} = \lim_{x \rightarrow a} \frac{f(x, b) - f(a, b)}{x - a} \quad (3.2.1)$$

のことである（もちろん、この極限が存在する場合のみ、この定義は有効）。これは記号で $\frac{\partial f}{\partial x}(a, b)$, $f_1(a, b)$, $f_x(a, b)$, $D_1 f(a, b)$ などと書く。同様に、第 2 変数に関する偏微分係数とは

$$\lim_{h \rightarrow 0} \frac{f(a, b+h) - f(a, b)}{h} = \lim_{y \rightarrow b} \frac{f(a, y) - f(a, b)}{y - b} \quad (3.2.2)$$

のことであって、 $\frac{\partial f}{\partial y}(a, b)$, $f_2(a, b)$, $f_y(a, b)$, $D_2 f(a, b)$ などと書く。

上のように各点で偏微分係数を計算すると、 (x, y) の関数として $\frac{\partial f}{\partial x}(x, y)$, $\frac{\partial f}{\partial y}(x, y)$ が定まる。これを f の (x, y) に関する **偏導函数** と呼ぶ。

(記号の注意) 括弧に 2 重の意味があるためになかなか避けにくいのだが、 $\frac{\partial f}{\partial x}(a, b)$ などというのは、点 (a, b) における $\frac{\partial f}{\partial x}$ の値のつもりであって、 $\frac{\partial f}{\partial x}$ に (a, b) をかけたものではない。これは文脈から明らかとは思いますが、式がどうしても複雑になって混乱するといけなないので、念のため。

以下の定義はよく使うので、ここで与えておく。

定義 3.2.2 (C^1 -級) 多変数関数 $f(x_1, x_2, \dots, x_n)$ がその定義域（の一部） D で

- f は各変数 $x_1, x_2, \dots, x_n$ のそれぞれについて偏微分可能で
- かつ、その n -この偏導函数が $\mathbf{x} = (x_1, x_2, \dots, x_n)$ の連続函数である

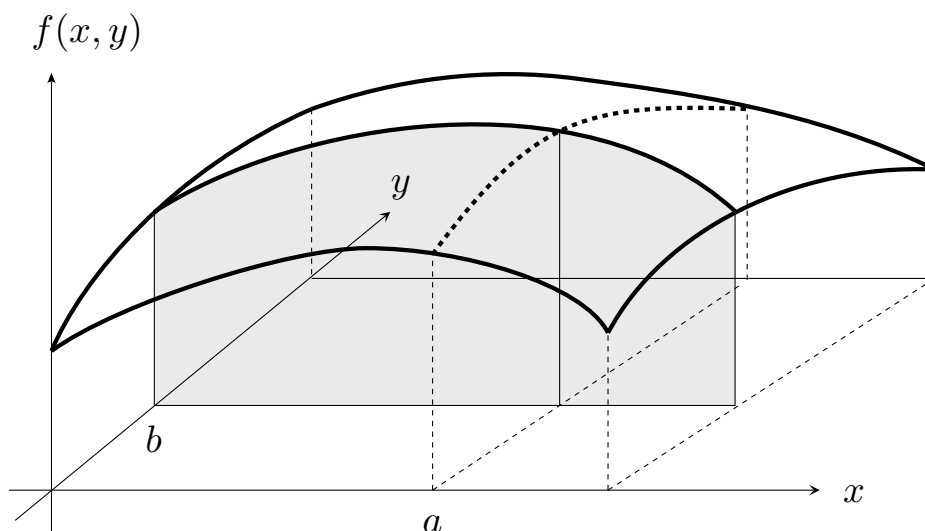
であるとき、 f は D で C^1 -級であるという。

この後で「高階の偏導函数」を学ぶ。そこでは「 n -階までの偏導函数がすべて存在してかつ連続」な函数を C^n -級の函数という（一変数函数の時と同様の定義）。これらの定義では（考えている階数までの）すべての偏導函数の存在と偏導函数の連続性を仮定していることに注意せよ。

偏微分の図形的な意味について、簡単に述べておこう。その定義からわかるように、 x での偏微分というのは $y = b$ を一定にして x だけを動かして微分、という事だ。これは $z = f(x, y)$ のグラフを $y = b$ の面で切った切り口を見て、この切り口のグラフの変化率を考えていることになる。下図では太い実線がそれにあたる。一方、 y での偏微分は $x = a$ の面での断面を問題にしている。下図では太い点線のグラフを見ていることになる。

このようなイメージは非常に役に立つものだから、できるだけ持つように心がけよう。

⁴²教科書 3.3 節前半



(記号についての注意)

$f(x, y)$ の偏導函数 $\frac{\partial f}{\partial x}$ の記号としては, $\frac{\partial f}{\partial x}$, $D_x f$, $D_1 f$, $\partial_x f$, $\partial_1 f$, f_x , f_1 などが一般的である. 時たまに f'_x というのも見かけるが, それほど一般的ではない. いずれにせよ, **どの変数で微分するのかがわかるように何らかの明記を行うことが不可欠である**. 時々, f' と書いて $\frac{\partial f}{\partial x}$ のつもりである人がいるから, 念のために注意しておく.

問 3.2.1. 次の函数をそれぞれの独立変数で偏微分せよ.

$$a) \quad x^2 + y^3, \quad b) \quad 2x^2y \quad c) \quad \sin(xy^2) \quad d) \quad (x^2 + y + z^3)^2$$

$$e) \quad f(x, y) = \begin{cases} 0 & (x, y) = (0, 0) \text{ の時} \\ \frac{2xy}{x^2 + y^2} & (x, y) \neq (0, 0) \text{ の時} \end{cases}$$

3.2.1 偏導函数がゼロ, の函数は?

1 変数の函数 f の場合, 導函数 f' が恒等的にゼロというのは簡単だった — f は定数しかない.

ところが, 多変数の函数では事情が異なる. 例えば, 2 変数函数 $f(x, y)$ が $f_x(x, y) \equiv 0$ を満たしていると, これは f が x には依存しないと言ってるにすぎない. (1 変数の時も「 x に依存しない」ことは同じだけど, あの場合は x しか変数がなかったから, x に依存しないなら定数だった.) いまは y にはいくら依存してもよいのだから, このような f は

$$f(x, y) = g(y) \quad g \text{ は任意の函数} \quad (3.2.3)$$

と書ける. これは一般には定数函数ではない!

1 変数に慣れすぎたあまり, 「導函数がゼロなら定数」と思い込みがちだが, 偏導函数に関してはこれは正しくないから, 注意しよう.

3.2.2 方向微分⁴³

偏微分の持つ意味を明らかにするため, 偏微分よりも広い, 「方向微分」という概念を導入しよう.

2 変数の函数 $f(x, y)$ を考える. その定義から, 偏微分 $\frac{\partial f}{\partial x}$ や偏微分 $\frac{\partial f}{\partial y}$ とは, この函数の x -方向, y -方向での変化率を表すと考えられる (各自, 理由を納得せよ).

しかし, x, y の函数として, もっと他の方向での変化率を考えたいこともあるだろう. 例えば, 点 (a, b) でのまわりで $f(x, y)$ がどのように変化しているかを見たい場合, x -方向, y -方向だけでは不十分で, (例えば) $x = y$ の斜め 45° の直線にそって x, y が動いた時にどうなるか, などを見たい.

⁴³ この小節の内容は, 偏微分に関する理解を深めるための補助的なものである.

そこで、このような変化率をみるために、以下の定義を行う。

定義 3.2.3 (方向微分) 2変数関数 $f(x, y)$ と2次元の単位ベクトル (長さ1のベクトル) $\mathbf{v} = (v_x, v_y)$ が与えられたとせよ。極限

$$f_{\mathbf{v}}(a, b) := \lim_{h \rightarrow 0} \frac{f(a + hv_x, b + hv_y) - f(a, b)}{h} \quad (3.2.4)$$

が存在するとき、これを、 $f(x, y)$ の点 (a, b) における $\mathbf{v}$ 方向の方向微係数 (方向微分) という。同様に、 n 変数関数 $f(\mathbf{x})$ と n 次元の単位ベクトル $\mathbf{v}$ が与えられたとき、極限

$$f_{\mathbf{v}}(\mathbf{a}) := \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h} \quad (3.2.5)$$

が存在するなら、これを $f(\mathbf{x})$ の点 $\mathbf{a}$ における $\mathbf{v}$ 方向の方向微係数という。この方向微分は $D_{\mathbf{v}}f(\mathbf{a})$ とも書く。

いうまでもなく、 $f(\mathbf{a})$ の $\mathbf{v}$ の方向での変化率を表すのがこの方向微分 $D_{\mathbf{v}}f(\mathbf{a})$ なのである。またこの定義に従うと、 x_1 による偏微分 $f_1(\mathbf{x})$ は正に x_1 -軸の向きを向いた単位ベクトル方向の方向微分、ということになる。

さて、関数 $f(\mathbf{a})$ の各座標軸方向の偏微分が存在しても、それだけではいろいろな方向微分が存在するとは限らない。これを保証するのが次の小節で述べる「全微分可能性」である。

その前に少し例を挙げておこう。以下の関数 f, g

$$f(0, 0) = 0, \quad (x, y) \neq (0, 0) \text{ では } f(x, y) = \frac{2xy}{x^2 + y^2} \quad (3.2.6)$$

$$g(0, 0) = 0, \quad (x, y) \neq (0, 0) \text{ では } g(x, y) = \frac{xy}{\sqrt{x^2 + y^2}} \quad (3.2.7)$$

を考える。定義通り計算すると、これらの関数はすべての (x, y) で偏微分できて、

$$f_x(0, 0) = f_y(0, 0) = 0, \quad (x, y) \neq (0, 0) \text{ では } f_x(x, y) = \frac{2y(y^2 - x^2)}{(x^2 + y^2)^2}, \quad f_y(x, y) = \frac{2x(x^2 - y^2)}{(x^2 + y^2)^2} \quad (3.2.8)$$

$$g_x(0, 0) = g_y(0, 0) = 0, \quad (x, y) \neq (0, 0) \text{ では } g_x(x, y) = \frac{y^3}{(x^2 + y^2)^{3/2}}, \quad g_y(x, y) = \frac{x^3}{(x^2 + y^2)^{3/2}} \quad (3.2.9)$$

である (各自、確かめるんだよ! 特に $(0, 0)$ での微係数の計算に注意)。しかし、単位ベクトル $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ 方向の方向微分は、原点では存在しない (これも確かめる事)。

3.2.3 全微分可能性⁴⁴

偏微分のもつ意味について、もう少し考える。1変数関数 $f(x)$ の場合、 $x = a$ での微係数 $f'(a)$ は $y = f(x)$ のグラフの接線の傾きだった。でもグラフから (また平均値の定理やテイラーの定理から) 明らかなように、これはまた $x \approx a$ での $f(x)$ の近似値をも与えてくれた:

$$f(x) \approx f(a) + f'(a) \times (x - a). \quad (3.2.10)$$

我々は当然、偏微分にも同じ役割を担ってほしい。つまり、 $\mathbf{x} = \mathbf{a}$ の点の近傍での $f(\mathbf{x})$ のふるまいを、偏微分を使って近似したい。

ところが (!) 多変数関数ではこれは全く自明ではない。例えば先の (3.2.6), (3.2.7) の例を考えてみるとよい。 $x = y = 0$ (原点) では f の偏微分係数はともにゼロであるが、 $f(x, y)$ は原点付近でゼロではない。たとえば $x = y$ では $f(x, x) = 1$ ($x \neq 0$) であって、**原点で連続ですらない!** 1変数関数の場合は「微分可能ならば連続」であるのに、2変数関数ではこのような変態もありうるわけだ。

しかし、これは実は驚くにはあたらない。 x での偏微分というのは y を固定して x を動かした時の振る舞いしか見ないから、 x -軸に平行に動いたときの振る舞いは偏微分からわかるけども、 $x = y$ のように x -軸に平行

⁴⁴この小節の内容は「進んだ話題」なので余裕のない人はとばしても良いし、講義でも触れない可能性が高い。

でない動きは x での偏微分だけでは見えないのだ。 y での偏微分も y -軸に平行な動きしか教えてくれないから、座標軸に平行でない動きは偏微分だけでは予測不可能、ということになる。そしてこのような動きを反映する概念として「方向微分」を導入したのである。

この方向微分と密接に関連するのが以下に定義する「全微分可能性」という概念である。以下の定義などの中ではお約束通り、 $\mathbf{x} = (x_1, x_2, \dots, x_n)$ 、および $\|\mathbf{x} - \mathbf{a}\| = \left(\sum_{j=1}^n (x_j - a_j)^2\right)^{1/2}$ である。

定義 3.2.4 (全微分可能性) ある領域 D で定義された n 変数関数 $f(\mathbf{x})$ と、 D 内の 1 点 $\mathbf{a}$ がある。定数 $A_1, A_2, \dots, A_n$ が存在して（「定数」という意味は $\mathbf{x}$ に依存しないということ。もちろん、 $\mathbf{a}$ には依存してよい）、

$$f(\mathbf{x}) = f(\mathbf{a}) + \sum_{j=1}^n A_j(x_j - a_j) + o(\|\mathbf{x} - \mathbf{a}\|) \quad (3.2.11)$$

が成り立つとき、 f は $\mathbf{x} = \mathbf{a}$ で全微分可能という。「全微分可能」を単に「微分可能」と言うこともある。

言うまでもなく、(3.2.6) や (3.2.7) の f, g は原点 $(0, 0)$ では全微分可能でない。

上の定義のミソは (3.2.11) が $\mathbf{a}$ に近いすべての $\mathbf{x}$ 、つまり $\mathbf{a}$ へのあらゆる近づき方について要求されていることである。繰り返しになるが、偏微分では x -軸、 y -軸などの特定の方向からの近づきかたしか考えていない。このため、あらゆる近づき方を考えている全微分可能性は、偏微分可能性よりも偉い（条件がきつい）。まとめると、以下の命題になる。

命題 3.2.5 ある領域 D で定義された n 変数関数 $f(\mathbf{x})$ と、 D 内の 1 点 $\mathbf{a}$ があって、 $f(\mathbf{x})$ が $\mathbf{x} = \mathbf{a}$ で全微分可能だとする。このとき、 $\mathbf{x} = \mathbf{a}$ において

1. $f(\mathbf{x})$ は連続であり、
2. $f(\mathbf{x})$ はすべての x_j について偏微分可能で、

$$\left((3.2.11) \text{ の } A_j \right) = \frac{\partial f}{\partial x_j}(\mathbf{a}) \quad (j = 1, 2, \dots, n), \quad (3.2.12)$$

3. 更に、任意の n 次元単位ベクトル $\mathbf{v} = (v_1, v_2, \dots, v_n)$ 方向の方向微分が存在して

$$D_{\mathbf{v}}f(\mathbf{a}) = \sum_{j=1}^n A_j v_j = \sum_{j=1}^n \frac{\partial f}{\partial x_j} v_j \quad (3.2.13)$$

証明.

1. f が連続なのはほとんど自明だ。というのも、全微分可能の条件 (3.2.11) は $f(\mathbf{x}) - f(\mathbf{a})$ がゼロに行くことを保証しているから。

2. 偏微分についても簡単だ。なぜなら、 x_1 で偏微分するときには $x_2, x_3, \dots$ は $a_2, a_3, \dots$ に固定して考えるので、(3.2.11) から

$$\frac{f(x_1, a_2, a_3, \dots, a_n) - f(a_1, a_2, a_3, \dots, a_n)}{x_1 - a_1} = A_1 + \frac{\tilde{f}(\mathbf{x})}{x_1 - a_1} \quad (3.2.14)$$

の $x_1 \rightarrow a_1$ の極限が $\frac{\partial f}{\partial x_1}$ を与えることになる。ところが、 $x_2, x_3, \dots$ を $a_2, a_3, \dots$ に固定した場合は $|x_1 - a_1| = \|\mathbf{x} - \mathbf{a}\|$ であるので、(3.2.11) から $\frac{\tilde{f}(\mathbf{x})}{x_1 - a_1}$ がゼロに行く事が保証される。これは f の x_1 での偏微分係数が存在して A_1 である、と言っているのと同値である。 x_2 以下での偏微分も同様である。

3. 方向微分についても、同様に議論する。つまり (3.2.11) から

$$\frac{f(\mathbf{a} + h\mathbf{v}) - f(\mathbf{a})}{h} = \frac{\sum_{j=1}^n A_j h v_j + \tilde{f}(\mathbf{a} + h\mathbf{v})}{h} = \sum_{j=1}^n A_j v_j + \frac{\tilde{f}(\mathbf{a} + h\mathbf{v})}{h} \quad (3.2.15)$$

が得られる. ここで $\mathbf{x} = \mathbf{a} + h\mathbf{v}$ と書くと

$$\frac{\tilde{f}(\mathbf{a} + h\mathbf{v})}{h} = \frac{f(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|} \quad (3.2.16)$$

であるため, (3.2.11) から, (3.2.15) の最後の項は $h \rightarrow 0$ でゼロに行く. よって, (3.2.13) が証明される. $\square$

全微分可能の図形的意味

2変数の関数 $f(x, y)$ の全微分可能性 (3.2.11) は図形的には以下のように解釈できる. まず, (3.2.11) の最後の項がない場合を考えると, 定数 A, B があって

$$f(x, y) = f(a, b) + A(x - a) + B(y - b) \quad (3.2.17)$$

となっている. このとき, $z = f(x, y)$ のグラフを考えると, これは $z - c = A(x - a) + B(y - b)$ (ここで $c = f(a, b)$ は定数) となって, 空間内の点 (a, b, c) を通る平面になっている (線型代数でやるはず). この平面を S としよう.

実際には (3.2.11) には余分な項がついているわけで, $z = f(x, y)$ のグラフは簡単な平面ではない. しかし, $\mathbf{x} - \mathbf{a}$ が小さい場合にはこの項はほとんど無視できるから, $z = f(x, y)$ のグラフは, 上の平面 S とほとんど同じと思ってよい. 上の平面 S は $z = f(x, y)$ のグラフの, (a, b) における接平面になっている.

つまり, 全微分可能の条件 (3.2.11) は, $z = f(x, y)$ のグラフが接平面を持つ, または $z = f(x, y)$ のグラフがその接平面で良く近似できる条件とも解釈できるのである. (3.2.6) や (3.2.7) の f, g では, $(0, 0)$ でのグラフの接平面が存在しないことを直感的に理解しよう.

全微分可能の十分条件

最後に, 全微分可能の十分条件を一つ, 与えておこう.

定理 3.2.6 (C^1 級なら全微分可能) ある領域 D で定義された n 変数関数 $f(\mathbf{x})$ が C^1 -級なら, つまり D 内の各点で 1 階の偏導関数がすべて存在して連続なら, $f(\mathbf{x})$ は D の各点で全微分可能である.

証明. (この証明は高校までの知識で大体は理解できるが, 跳ばしても構わない.)

D 内の点 (a, b) で全微分可能であることを証明する. 式を見やすくするため, $\mathbf{a} := (a, b), \mathbf{x} = (x, y)$ と書く. 全微分可能であることをいうためには, 差

$$f(x, y) - f(a, b) = \{f(x, y) - f(a, y)\} + \{f(a, y) - f(a, b)\} \quad (3.2.18)$$

が $\mathbf{x} \rightarrow \mathbf{a}$ でどのように振る舞うか — (3.2.11) を満たすか — を調べなければならない.

さて, (3.2.18) の 2 つめの差では (x 座標は a で共通だから) 変数 y についての 1 変数の平均値の定理をつかうと (f が C^1 級だと仮定しているので, 平均値の定理は使える)

$$f(a, y) - f(a, b) = f_y(a, \tilde{y}) \times (y - b) \quad (3.2.19)$$

が得られる — ここで $\tilde{y}$ は y と b の間の適当な数である. また, 言うまでもなく, $f_y = \frac{\partial f}{\partial y}$ である.) 一方, 一つ目の差は (y が共通だから x についての平均値の定理から)

$$f(x, y) - f(a, y) = f_x(\tilde{x}, y) \times (x - a) \quad (3.2.20)$$

となる ($\tilde{x}$ は a と x の間の適当な数). これを (3.2.18) に代入して

$$f(x, y) - f(a, b) = f_x(\tilde{x}, y) \times (x - a) + f_y(a, \tilde{y}) \times (y - b) \quad (3.2.21)$$

を得る.

問題は $f_x(\tilde{x}, y), f_y(a, \tilde{y})$ がどのような量かということであるが, 今 f が C^1 -級 (つまり, f_x, f_y がともに連続関数) だと仮定しているので, 一般の (u, v) に対して

$$f_x(u, v) = f_x(a, b) + g(u, v), \quad \text{with} \quad \lim_{(u, v) \rightarrow (a, b)} g(u, v) = 0, \quad (3.2.22)$$

$$f_y(u, v) = f_y(a, b) + h(u, v), \quad \text{with} \quad \lim_{(u, v) \rightarrow (a, b)} h(u, v) = 0 \quad (3.2.23)$$

が成り立っている。ここで (3.2.22) を $u = \tilde{x}, v = y$ として用いると, $(x, y) \rightarrow (a, b)$ の時には $(\tilde{x}, y) \rightarrow (a, b)$ でもあるから,

$$f_x(\tilde{x}, y) = f_x(a, b) + g(\tilde{x}, y), \quad \text{with} \quad \lim_{(x, y) \rightarrow (a, b)} g(\tilde{x}, y) = 0 \quad (3.2.24)$$

が結論できる。同様に,

$$f_y(a, \tilde{y}) = f_y(a, b) + h(a, \tilde{y}), \quad \text{with} \quad \lim_{(x, y) \rightarrow (a, b)} h(a, \tilde{y}) = 0 \quad (3.2.25)$$

も結論できる。これを (3.2.21) に代入すると

$$f(x, y) - f(a, b) = f_x(a, b) \times (x - a) + f_y(a, b) \times (y - b) + g(\tilde{x}, y) \times (x - a) + h(a, \tilde{y}) \times (y - b) \quad (3.2.26)$$

が得られる。 g, h は両方ともゼロに行くから, 後ろの 2 つを $\tilde{f}(x, y)$ とまとめると,

$$\lim_{(x, y) \rightarrow (a, b)} \frac{\tilde{f}(x, y)}{\|x - a\|} = 0 \quad (\text{ここで } x = (x, y), a = (a, b) \text{ と書いた}) \quad (3.2.27)$$

が結論できる。結果として, このような $\tilde{f}$ を用いて

$$f(x, y) = f(a, b) + A(x - a) + B(y - b) + \tilde{f}(x, y), \quad \text{with} \quad A = f_x(a, b), B = f_y(a, b) \quad (3.2.28)$$

と書ける事がわかった。これは全微分可能の定義式 (3.2.11) そのものであって, 定理が証明された。 n -変数の場合も (式が汚くなるだけで) 同様。 $\square$

3.3 合成関数の微分 (連鎖率, chain rule)

ここでは偏微分での最初の山場, 「連鎖率」 (合成関数の微分) を学ぶ。この題材は簡単に見えて, 案外たいへんなことがあるから, 注意する事。特に, この後でやる「高階の偏導関数」を計算する時にひっかかる人が多いはずだ。なお, この節の山場は後の 3.3.3 節である⁴⁵。

まず 1 変数の場合を思い出そう。実数値関数 $f(x)$ と $g(y)$ が与えられたとき,

$$h(x) = f(g(x)) \quad (3.3.1)$$

で定義される関数 h を f と g の 合成関数 といい, $f \circ g$ などと書いたのだった。例えば $\sin(x^3)$ は $f(x) = \sin x$ と $g(x) = x^3$ の合成関数である。

このとき, 関数 $h(x)$ の導関数については, 高校以来,

$$h'(x) = f'(g(x)) g'(x) \quad (3.3.2)$$

が成り立つことは知っている。この節の主題は, これの多変数関数版を考える事である。

f と g のどちらが多変数かによって 4 通りあるから, 場合分けして考えよう (ただし, 以下では一般の n 変数をやると大変だから, 2 変数までを主に考える):

- A. 1 変数の関数 $f(z)$, $z(x)$ があるとき, 合成関数 $h(x) = f(z(x))$ の, x による微分。
- B. 1 変数の関数 $f(z)$ と 2 変数の関数 $z(x, y)$ があるとき, 合成関数 $h(x, y) = f(z(x, y))$ の, x, y による偏微分。
- C. 2 変数の関数 $f(x, y)$ と 1 変数の関数 $x(t), y(t)$ があるとき, 合成関数 $h(t) = f(x(t), y(t))$ の, t による微分。
- D. 2 変数の関数 $f(x, y), x(u, v), y(u, v)$ があるとき, 合成関数 $h(u, v) = f(x(u, v), y(u, v))$ の, u, v による偏微分。

⁴⁵ この講義ノートでは, 重要なことは小節 (1.2.3 節など) ではなく節 (1.2 節など) に書くようにしている。しかしこの小節のように, どうしても話の流れ上, 大事な事が小節に入ってしまうことがある。これはできるだけ指摘するようにするが, 注意されたい

このうち、A は高校以来知っていることだ (この後でも改めて証明する)。

また、B も見かけ倒しである。既に学んだように、 $h(x, y)$ を x で偏微分する場合には、 y をとめて偏微分する。つまり微分操作をやる限りでは y は定数と思って、 $f(x, y)$ は x のみの関数と思って微分すればよい。これなら B は高校までの A と全く同じことである。従って

$$\frac{\partial h(x, y)}{\partial x} = f'(z(x, y)) \frac{\partial z(x, y)}{\partial x}, \quad \frac{\partial h(x, y)}{\partial y} = f'(z(x, y)) \frac{\partial z(x, y)}{\partial y} \quad (3.3.3)$$

となる。

問題は C だ。(D は C ができればすぐにわかる —— この事情は B が A からすぐにわかるのと同じ。) これは多変数特有の現象なので、注意が必要である。いずれにせよ、1 変数の場合が (証明のアイディアも含めて) わからないと話にならないので、まずは高校以来の 1 変数の場合を復習しよう。

3.3.1 合成関数の微分 (1 変数の場合の復習, Case A)

$g(x)$ は区間 I で、 $f(y)$ は区間 J で、それぞれ定義されており、かつ、 g の値域 $g(I) = \{g(x) \mid x \in I\}$ が J の部分集合であるとする。このとき、合成関数 $h(x) = f(g(x))$ を区間 I で定義することができるが、その微分係数に関しては以下が成り立つ。

定理 3.3.1 (1 変数の合成関数の微分) $g(x)$ が区間 I 内の点 $x = a$ にて x について微分可能、かつ $f(z)$ が点 $b = f(a)$ にて z について微分可能のとき、合成関数 $h(x) = f(g(x))$ は点 $x = a$ で微分可能であり、

$$h'(x) = f'(g(x)) g'(x), \quad \text{つまり } z = g(x), w = h(z) \text{ とおくと } \frac{dw}{dx} = \frac{dw}{dz} \frac{dz}{dx}$$

がなりたつ。

(ちょっとマニアックな注) 1 変数の関数に関するこの定理では、「 $g(x)$ が $x = a$ で微分可能、かつ $f(z)$ が $z = f(a)$ で微分可能」であれば十分で、導関数の連続性などは必要ない。後出の多変数の場合の定理 3.3.3 では事情が異なり、導関数の連続性 (またはそれに類する条件) が必要になってくる。この事情は「全微分可能性」と関連している。

定理 3.3.1 の少しだけええかげんな証明。 この定理はほとんど当たり前だ。 $h(x)$ の $x = a$ での微分を定義するニュートン商を

$$\frac{h(a+\epsilon) - h(a)}{\epsilon} = \frac{f(g(a+\epsilon)) - f(g(a))}{\epsilon} = \frac{f(g(a+\epsilon)) - f(g(a))}{g(a+\epsilon) - g(a)} \times \frac{g(a+\epsilon) - g(a)}{\epsilon} \quad (3.3.4)$$

と書いて、 $\epsilon \rightarrow 0$ としてやれば良い。後ろはモロに $g'(a)$ に行くし、 $g(a+\epsilon) \rightarrow g(a)$ であるから (以下の注参照) 前の項は f' の $g(a)$ での値に行く。□

(補足) 上の「証明」でごまかしたのは、「 $\epsilon \neq 0$ であつても $g(a+\epsilon) - g(a) = 0$ かもしれない」という可能性を見えぬふりをしたことだ — この可能性の例としては、 $g(x) \equiv 1$ (恒等的に 1) を考えよ。もし $g(a+\epsilon) - g(a) = 0$ ならば (3.3.4) 右辺の書き換え (分母がゼロ!) に意味がつけられなくなり、右辺の積の極限を別々に考えることができなくなる。

でも、これは大した問題ではない。実際、 $g(a+\epsilon) - g(a) = 0$ ならば $h(g(a+\epsilon)) - h(g(a)) = 0$ でもあるはずだから、もともとのニュートン商の値もゼロ、よって困ることは何もないはずだ。実際、このところはちょっと書き方を工夫すれば厳密に議論できる。上の「証明」は不完全だが、まずは「このような感じだな」と大体の筋道を理解すればこの講義には十分だ。

3.3.2 合成関数の微分 (1 変数の場合に帰着, Case B)

既に注意したように、B の場合は上からすぐに出る。つまり、区間 J で定義された関数 $f(z)$ と、ある領域 D で定義された関数 $g(x, y)$ があつて、 g の値域が f の定義域に含まれているとする。このとき合成関数 $h(x, y) = f(g(x, y))$ を D で定義することができるが...

定理 3.3.2 $g(x, y)$ が R 内の一点 (a, b) にて x について偏微分可能, かつ $f(z)$ が $c = g(a, b)$ で微分可能とする. このとき, $h(x, y) = f(g(x, y))$ は (a, b) にて x について偏微分可能で,

$$h_x(a, b) = f'(g(a, b)) g_x(a, b), \quad \text{つまり } z = g(x, y), w = f(z) \text{ とおくと } \frac{\partial w}{\partial x} = \frac{dw}{dz} \frac{\partial z}{\partial x} \quad (3.3.5)$$

がなりたつ.

証明 x についての偏微分のみを問題にしているから, 変数 y は単なる定数と思っても同じだ. だから, 定理 3.3.1 が使える. $\square$

3.3.3 合成関数の微分 (本質的に多変数の場合, Cases C & D)

この小節の内容がこの節のメインである. いよいよ, C の場合に進もう. ここに至って, 本質的に新しい問題が生じる. まずは発見的に考えてみる.

2変数の関数 $f(x, y)$ と $x(t), y(t)$ から合成関数 $h(t) = f(x(t), y(t))$ を作る. この t での微分を考えると, ニュートン商の極限として [記号を見やすくするため. $x_0 = x(t), y_0 = y(t), x_1 = x(t + \epsilon), y_1 = y(t + \epsilon)$ と書く⁴⁶]

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} \frac{h(t + \epsilon) - h(t)}{\epsilon} &= \lim_{\epsilon \rightarrow 0} \frac{f(x(t + \epsilon), y(t + \epsilon)) - f(x(t), y(t))}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{f(x_1, y_1) - f(x_0, y_0)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{f(x_1, y_1) - f(x_1, y_0)}{\epsilon} + \lim_{\epsilon \rightarrow 0} \frac{f(x_1, y_0) - f(x_0, y_0)}{\epsilon} \end{aligned} \quad (3.3.6)$$

が出てくる. この第2項の極限は簡単だ. ϵ に依存した項は $x_1 = x(t + \epsilon)$ しかないから, $y_0 = y(t)$ の方は $\epsilon \rightarrow 0$ の極限をとる際に定数と思っても良い. これは (y_0 を定数と思って) 合成関数の微分の公式そのものだから $\frac{\partial f}{\partial x}(x_0, y_0) x'(t)$ になる. (ここまではゴマカシなし.)

第1項はもっとややこしい (**ここからゴマカシ**). もしこれが (x_2 は ϵ に無関係な数で)

$$\lim_{\epsilon \rightarrow 0} \frac{f(x_2, y_1) - f(x_2, y_0)}{\epsilon} \quad \text{with} \quad y_0 = y(t), \quad y_1 = y(t + \epsilon) \quad (3.3.7)$$

であれば, $x = x_2$ は定数で y だけが ϵ に依存するから, 極限は合成関数の微分 (case B) により $\frac{\partial f}{\partial y}(x_2, y) y'(t)$ になる. さらに x_2 も x_0 に近づくと思えば, これは多分, $\frac{\partial f}{\partial y}(x_0, y_0) y'(t)$ になるだろう (**ここでゴマカシ終わり**). よってこのゴマカシによると, (3.3.6) から

$$h'(t) = \frac{\partial f}{\partial x}(x_0, y_0) x'(t) + \frac{\partial f}{\partial y}(x_0, y_0) y'(t) \quad (3.3.8)$$

が得られると予想される.

答えを言ってしまうと, 以上の結論 (3.3.8) はかなり一般に成り立つ. ただし, 上でも明記したように, (3.3.6) の第一項の極限を求めるときにごまかしてしまったのが問題だ. 実際, 以下の反例が示すように, ここはもう少し仮定が必要である.

(反例) 以前に出た例だが, 以下の関数 $g(x, y)$ と $x(t) = y(t) = t$ ($t \geq 0$) を考える:

$$g(0, 0) = 0, \quad (x, y) \neq (0, 0) \text{ では } g(x, y) = \frac{xy}{\sqrt{x^2 + y^2}}. \quad (3.3.9)$$

この関数の偏導関数は (3.2.9) で計算した. 特に, $g_x(0, 0) = g_y(0, 0) = 0$ である. さて, 地道に計算するとすべての t で

$$h(t) = g(x(t), y(t)) = \frac{t}{\sqrt{2}}, \quad h'(t) = \frac{1}{\sqrt{2}} \quad (3.3.10)$$

であって, 特に $h'(0) \neq 0$ だ. ところが (3.3.8) を闇雲に使うと ($x = y = 0$ での偏微分の値を使って計算するから) $h'(0) = 0$ が得られてしまう. つまり, この $g(x, y)$ に対しては (3.3.8) は適用できない! (反例終わり)

⁴⁶ $x(t + \epsilon)$ などの括弧は関数の依存性を示すもので, 掛け算ではない

以下ではこのところを厳密にやれるような十分条件を2つ、定理の形で述べる。まず、覚えやすい形としては C^1 級を仮定するものがあるので、それを述べよう。(もう一つの十分条件はもっとマニアックなので3.3.4節で述べる。) なお、あまり細かいことを書くと肝心のところが見えなくなりそうだから、函数の定義域と値域は、合成函数が定義できるようになっていると適当に仮定する。

定理 3.3.3 2変数函数 $f(x, y)$ が C^1 -級で、 $x(t), y(t)$ が t について微分可能なら、 $h(t) = f(x(t), y(t))$ は t で微分可能である。更にその導函数について

$$h'(t) = f_x x'(t) + f_y y'(t), \quad \text{つまり } z = f(x, y) \text{ とおくと } \frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \quad (3.3.11)$$

がなりたつ (f の偏微分はもちろん、 $(x(t), y(t))$ での値)。更に一般に f が n 変数の函数の場合は：

$$\frac{d}{dt} f(x_1(t), x_2(t), \dots, x_n(t)) = \sum_{j=1}^n \frac{\partial f}{\partial x_j} \frac{dx_j}{dt}. \quad (3.3.12)$$

言うまでもなく、(3.3.9) の例は C^1 級ではないから、上の定理が適用できなくても仕方がない。

証明 (3.3.6) の第一項がゴマカシだったので、きちんとやりなおそう。 $f(x, y)$ を、 y だけの函数と見て1変数函数の平均値の定理を用いると (f が C^1 級だと仮定しているの、1変数函数の平均値の定理は使える)

$$f(x_1, y_1) - f(x_1, y_0) = f_y(x_1, y_3) \times (y_1 - y_0) \quad (3.3.13)$$

が得られる — ここで y_3 は y_0 と y_1 の間の適当な数である。この両辺を ϵ で割って $\epsilon \rightarrow 0$ とすると、

$$\lim_{\epsilon \rightarrow 0} \frac{f(x_1, y_1) - f(x_1, y_0)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \left[f_y(x_1, y_3) \times \frac{y_1 - y_0}{\epsilon} \right] \quad (3.3.14)$$

となる。ところで、 f が C^1 級と仮定しているから、 $f_y(x, y)$ は x, y の連続函数である (ここがキー)。従って $\epsilon \rightarrow 0$ では $x_1 \rightarrow x_0$ かつ $y_3 \rightarrow y_0$ であることも使うと、

$$\lim_{\epsilon \rightarrow 0} f_y(x_1, y_3) = f_y(x_0, y_0) \quad (3.3.15)$$

である。また、(3.3.14) の後ろの方は y_1, y_0 の定義から

$$\lim_{\epsilon \rightarrow 0} \frac{y_1 - y_0}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{y(t+\epsilon) - y(t)}{\epsilon} = y'(t) \quad (3.3.16)$$

である。以上をまとめると

$$\lim_{\epsilon \rightarrow 0} \frac{f(x_1, y_1) - f(x_1, y_0)}{\epsilon} = f_y(x_0, y_0) \times y'(t) \quad (3.3.17)$$

が厳密に証明されたので、(3.3.6) に戻ると定理が証明できた。 n -変数の場合も (式が汚くなるだけで) 同様。□

D の場合についてもだめ押しで述べておこう。

定理 3.3.4 (2変数函数の連鎖律) 2変数函数 $f(x, y)$ が C^1 -級で、 $x(u, v), y(u, v)$ が u, v について C^1 -級なら、合成函数 $z = h(u, v) = f(x(u, v), y(u, v))$ も C^1 -級で、

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}, \quad \text{かつ} \quad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v}. \quad (3.3.18)$$

証明 u で偏微分する場合には v は動かさない (定数) から先の定理 3.3.3 からすぐに証明される。□

問題 3.3.5 (連鎖率) 2変数の函数 $f(x, y) = \frac{xy}{\sqrt{x^2 + y^2}}$ を考える。また、変数 x, y と変数 r, θ は平面の曲座標の関係、つまり $x = r \cos \theta, y = r \sin \theta$ を満たしているものとする。このとき、合成函数 $h(r, \theta) = f(x(r, \theta), y(r, \theta))$ の、変数 r, θ に関する偏導函数を計算せよ。ただしその場合、(1) 合成函数の微分法 (連鎖率) を用いて計算する、

(2) h を r, θ の関数として具体的に書き下してから偏微分する, の2通りで行い, 結果を比べること. なお, 場合によっては微分できない点があるかもしれないが, 今はそれは無視して良い. つまり, 微分できる点で偏導関数を求めればよい.

問題 3.3.6 (x, y) と (u, v) が 1 次変換 ($\alpha, \beta, \gamma, \delta$ は $\alpha\delta \neq \beta\gamma$ を満たす定数)

$$u = \alpha x + \beta y, \quad v = \gamma x + \delta y \quad (3.3.19)$$

の関係にある時, $\frac{\partial f}{\partial x}$ と $\frac{\partial f}{\partial y}$ を, $\frac{\partial f}{\partial u}$ と $\frac{\partial f}{\partial v}$ および $\alpha, \beta, \gamma, \delta$ を用いて表せ.

問題 3.3.7 以下の方程式を満たす C^1 -級の関数 $f(x, y)$ は一般にどんな形か, 求めよ (a, b はゼロでない定数である).

$$1) \frac{\partial f}{\partial y} \equiv 0 \quad 2) \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \equiv 0 \quad 3) a \frac{\partial f}{\partial x} + b \frac{\partial f}{\partial y} \equiv 0 \quad (3.3.20)$$

1) は既にやった (3.2.1 節). 問題は 2) と 3) だが, 問題 3.3.6 をヒントにせよ. つまり, 新しい変数 (u, v) をうまく見つけて, $\frac{\partial f}{\partial u} \equiv 0$ が成り立つようにしてみると良い. (もちろん, このようなやり方でうまく行く保証はないが, ある程度の経験を積めば, この形の方程式ならこのような一時変換で行ける, ことが予想できるものである.)

3.3.4 全微分可能性を仮定したときの chain rule

(この小節の内容は「おまけ」であり, 余裕のある人だけが読めばよい.)

上では f が C^1 級であることを仮定して連鎖律に関する定理を導いた. でもこれらの定理は「 f が全微分可能」と仮定するだけでも証明できる. 実際, 全微分可能の概念は, 上のような定理を自然に成り立たせるための条件として発見されてきたものなので, これは自然である. 以下ではこれらをまとめて述べる.

定理 3.3.8 (定理 3.3.3 の一般形) 2 変数関数 $f(x, y)$ が全微分可能で, $x(t), y(t)$ が t について微分可能なら, $h(t) = f(x(t), y(t))$ は t で微分可能で,

$$h'(t) = f_x x'(t) + f_y y'(t), \quad \text{つまり } z = f(x, y) \text{ とおくと } \frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt} \quad (3.3.21)$$

が成り立つ. n -変数の場合も同様であるが, 略.

定理 3.3.9 (定理 3.3.4 の一般形) 2 変数関数 $f(x, y)$ と $x(u, v), y(u, v)$ がそれぞれその引数について全微分可能なら, 合成関数 $z = h(u, v) = f(x(u, v), y(u, v))$ も (u, v) について全微分可能で,

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial u}, \quad \text{かつ} \quad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial z}{\partial y} \frac{\partial y}{\partial v} \quad (3.3.22)$$

である.

定理 3.3.4 が定理 3.3.3 から出ると同じようにして, 定理 3.3.9 は定理 3.3.8 から証明できる. 従って, 以下では定理 3.3.8 を簡単に証明する.

定理 3.3.8 の証明

以下では 5 月ごろに導入した「オーダー」の記号を用いている. 今までと同じく, 式を見やすくするために $x_0 = x(t), x_1 = x(t + \epsilon), y_0 = y(t), y_1 = y(t + \epsilon)$ と書く. 全微分可能性の定義と命題 3.2.5 から,

$$h(t + \epsilon) - h(t) = f(x_1, y_1) - f(x_0, y_0) = A(x_1 - x_0) + B(y_1 - y_0) + o(\sqrt{|x_1 - x_0|^2 + |y_1 - y_0|^2}) \quad (3.3.23)$$

が成り立っており, $A = \frac{\partial f}{\partial x}(x_0, y_0), B = \frac{\partial f}{\partial y}(x_0, y_0)$ であることが既にわかっている. $x(t), y(t)$ が微分可能であるから (ここでも式を簡単にするため, $p = x'(t), q = y'(t)$ と書く),

$$x_1 - x_0 = x'(t)\epsilon + o(\epsilon) = p\epsilon + o(\epsilon), \quad y_1 - y_0 = y'(t)\epsilon + o(\epsilon) = q\epsilon + o(\epsilon) \quad (3.3.24)$$

も成り立っている。従って、

$$\begin{aligned} h(t+\epsilon) &= h(t) + A\{p\epsilon + o(\epsilon)\} + B\{q\epsilon + o(\epsilon)\} + o\left(\sqrt{\{p\epsilon + o(\epsilon)\}^2 + \{q\epsilon + o(\epsilon)\}^2}\right) \\ &= h(t) + (Ap + Bq)\epsilon + o(\epsilon) + o\left(\sqrt{(p^2 + q^2)\epsilon^2 + o(\epsilon^2)}\right) \end{aligned} \quad (3.3.25)$$

となるが、

$$o\left(\sqrt{(p^2 + q^2)\epsilon^2 + o(\epsilon^2)}\right) = o\left(\sqrt{p^2 + q^2}\epsilon + o(\epsilon)\right) = o(\epsilon) \quad (3.3.26)$$

であるので、結局 (3.3.25) は

$$h(t+\epsilon) = h(t) + (Ap + Bq)\epsilon + o(\epsilon) \quad (3.3.27)$$

となる。でもこれは $h(t)$ の t での微係数が $Ap + Bq$ である、という式 (3.2.10) そのものだ。□

3.4 高階の偏導函数

(高校でやったこと) 1 変数の函数 $f(x)$ を x で微分したものを 1 階の導函数 $f'(x)$ といった。また $f'(x)$ を x でもう一回微分したものを 2 階の導函数といい、 $f''(x) = f^{(2)}(x)$ と書いた。同様に n -階の導函数 $f^{(n)}(x)$ も定義した。(高校、終わり)

2 変数以上の函数についても、同様のものを考えたい。すなわち、 $f(x, y)$ の 1 階導函数 $\frac{\partial f}{\partial x}(x, y)$ は x, y の函数であるが、これが x についてもう一回偏微分可能のとき、 f の x についての 2 階導函数 $\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)$ を定義する。同様に、 $\frac{\partial f}{\partial x}$ を y で微分して

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \text{同様に} \quad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \quad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \quad (3.4.1)$$

なども定義する (もちろん、これらの微分が定義できる場合)。まあ、この定義は非常に自然だから問題ないでしょう⁴⁷。ただし、微分の順序には注意：すぐに「順序はあんまり重要でない」とわかるが、 $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)$ は先に y で偏微分した後で x で偏微分、ということである。

問題 3.4.1 以下の函数 f, g, h について、2 階の偏導函数 (4 通り) をすべて計算せよ。実は以下の定理で $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ だというのが、この定理には頼らず、実際に計算して確かめる事。

$$f(x, y) = x^2 + y^2, \quad g(x, y) = x^2 y^3, \quad h(x, y) = \cos(x^2 y) \quad (3.4.2)$$

k -階の導函数も同様に定義する。つまり、 $f(x, y)$ の k -階の導函数とは $f(x, y)$ を x または y で合計 k 回、微分してできる函数のことである。 n -変数 ($n \geq 3$) の場合も同様に定義するが、自明だろうからここには書かない。また予告したように、 k -階までの偏導函数がすべて存在してかつ連続のとき、その函数は C^k -級という。

さて、これまでの定義によるだけでは、 $\frac{\partial^2 f}{\partial x \partial y}$ と $\frac{\partial^2 f}{\partial y \partial x}$ とは全く別ものであり (微分の順序が逆)、何の関係もないように見える。実際、問題 3.4.1 の微分を実際にやった人には、 $\frac{\partial g}{\partial x}$ と $\frac{\partial g}{\partial y}$ が全然別ものだから、 $\frac{\partial^2 g}{\partial x \partial y}$ と $\frac{\partial^2 g}{\partial y \partial x}$ に関係がつく事の方がかえって奇跡に見えるかもしれない。しかし、(我々が扱うような「普通の」函数では) この 2 つは等しい。

定理 3.4.2 2 変数函数 $f(x, y)$ が C^2 -級の場合、

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \quad (3.4.3)$$

である。つまり、**偏導函数は偏微分の順序によらない**。 n -変数の函数の場合も同様で、函数が C^2 -級ならば偏導函数は偏微分の順序によらない。さらに、 n -変数の函数が C^k -級ならば、その k -階までの偏導函数は偏微分の順序によらない。

⁴⁷ただし、ときどき、「 x で偏微分の時は y は定数」と変な覚え方をしてる人が「 x で微分した時に y が定数になったのにその定数の y で微分するんですか?」と混乱する事はあるようだ。これについては偏微分の意味 (x で偏微分というのは y が一定の面での変化率をみること) を思い出せば何の問題もないはずだ

定理 3.4.2 での「 C^2 -級」は十分条件であって、もう少しだけ条件を緩めることも可能である。例えば、以下のようなものがあるようだ⁴⁸ (D は領域, A は D 内の一点, また $f_x = \frac{\partial f}{\partial x}$, $f_{yx} = \frac{\partial^2 f}{\partial x \partial y}$ などと略記する)。

- D にて f_{yx} , f_{xy} が連続なら, D の各点で $f_{xy} = f_{yx}$
- D にて f_x, f_y, f_{yx} が存在し, A にて f_{yx} が連続ならば, f_{xy} も A で存在してかつ $f_{xy} = f_{yx}$
- D にて f_x, f_y が存在し, これらが A にて全微分可能ならば, A にて $f_{xy} = f_{yx}$

定理 3.4.2 の証明 (a, b) を f の定義域中の一点として固定し, ここでの微分を考える。

$$\Delta(h, k) := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b) \quad (3.4.4)$$

を考えよう。これを hk で割ってから h や k を適当な順序でゼロに持って行くと, 考えたい偏微分が出てくる。実際,

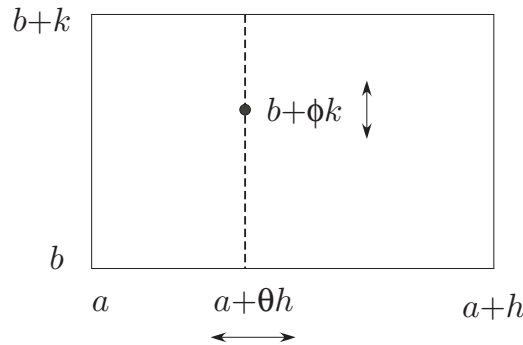
$$\begin{aligned} \lim_{h \rightarrow 0} \left(\lim_{k \rightarrow 0} \frac{\Delta(h, k)}{hk} \right) &= \lim_{h \rightarrow 0} \frac{1}{h} \left[\lim_{k \rightarrow 0} \left(\frac{f(a+h, b+k) - f(a+h, b)}{k} - \frac{f(a, b+k) - f(a, b)}{k} \right) \right] \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left[\frac{\partial f}{\partial y}(a+h, b) - \frac{\partial f}{\partial y}(a, b) \right] = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = \frac{\partial^2 f}{\partial x \partial y} \end{aligned} \quad (3.4.5)$$

であるし, 同様に考えると

$$\lim_{k \rightarrow 0} \left(\lim_{h \rightarrow 0} \frac{\Delta(h, k)}{hk} \right) = \frac{\partial^2 f}{\partial y \partial x} \quad (3.4.6)$$

でもある。つまり, 上の 2 つの極限の順序が交換できるかどうかのポイントになる。

極限の順序が交換できるかどうかを論じるには, 極限をとる前, つまり h, k がゼロでないところの表式をうまく書き直すしかない。それをやってみよう (下図も参照)。



まず, $\varphi(x) := f(x, b+k) - f(x, b)$ という関数を b, k を固定して考えると, これは 1 変数 x の関数とみなせる。従って, 1 変数関数の平均値の定理から,

$$\varphi(a+h) - \varphi(a) = \varphi'(a+\theta h)h \quad (3.4.7)$$

つまり

$$\Delta(h, k) = \varphi(a+h) - \varphi(a) = \left(f_x(a+\theta h, b+k) - f_x(a+\theta h, b) \right) h \quad (3.4.8)$$

となる $0 < \theta < 1$ が存在するといえる。これは上図の点線を左右に動かして, ちょうど良い $a+\theta h$ を見つけた事にあたる。

次に a, h, θ を固定して $\psi(y) := f_x(a+\theta h, y)$ を y の関数と考えると⁴⁹ またもや 1 変数の場合の平均値の定理から

$$\psi(b+k) - \psi(b) = \psi'(b+\phi k)k \quad (3.4.9)$$

つまり

$$f_x(a+\theta h, b+k) - f_x(a+\theta h, b) = f_{xy}(a+\theta h, b+\phi k)k \quad (3.4.10)$$

⁴⁸ 小平本の 6.2 節の d) を参考にした

⁴⁹ (注) ここでは少しひっかかるかもしれない。(3.4.7) では a, b, h, k を固定していたので, それに応じて θ が決まった。ところがここではその θ を固定した上で $\psi(y)$ を考え, b や $b+k$ の方を動かしているように見え, 何となく気持ちが悪い。だけど, ここの我々の結論 (3.4.10) は (3.4.7) がなりたっていた $b+k, b$ についてのものであるのだから, (3.4.7) と (3.4.10) は両立する。これは図の点線上を動かしているのだと思えば納得できるだろう

となる $0 < \phi < 1$ の存在がいえる (ここで平均値の定理が使えるための条件として $\psi(y)$ が C^1 -級である事を使うが、これは f が C^2 -級なので保証されている)。これは上の図では $x = a + \theta h$ での点線上を動かして、適切な $b + \phi k$ を探し当てたことに相当する。

以上 (3.4.7) と (3.4.10) から

$$\frac{\Delta(h, k)}{hk} = \frac{f_{xy}(a + \theta h, b + \phi k) hk}{hk} = f_{xy}(a + \theta h, b + \phi k) \quad (3.4.11)$$

となるような $0 < \theta < 1, 0 < \phi < 1$ の存在がいえた。 θ, ϕ が 0 と 1 の間にあり、更に f_{xy} が連続であることを用いると、上の極限は h, k をどのようにゼロに持って行っても $f_{xy}(a, b)$ に行く事がいえる：

$$\lim_{(h, k) \rightarrow (0, 0)} \frac{\Delta(h, k)}{hk} = f_{xy}(a, b) \quad (3.4.12)$$

これで既に $h \rightarrow 0$ と $k \rightarrow 0$ の極限が順序によらない事は示せたから (3.4.5) と (3.4.6) を思い出すと証明は完成した。(余分ではあるが、上の議論を順序を変えて行くと (3.4.12) の極限が $f_{yx}(a, b)$ に等しい事もいえて、証明はより明確になる。) $\square$

3.4.1 2 階の偏微分係数の幾何学的意味

2 階の偏導関数の意味は以下の幾何学的考察から少しはわかる (かなあ)。もう一つの意味付けは後で習うテイラー展開で与えられる。

$f_{xx} = \frac{\partial^2 f}{\partial x^2}$ の意味付けははっきりしている。 $z = f(x, y)$ のグラフを y が一定の面で切った切り口で、 x の 2 階微分を考えている訳だ。高校の時から知ってるように、2 階微分はグラフの凹凸 (曲がり方) を表す。従って、 f_{xx} は「 $z = f(x, y)$ のグラフを y が一定の面で切った切り口」での、 x -方向でのグラフの曲がり方を表している。同様に、 f_{yy} は x が一定の面で切った場合の、 y -方向でのグラフの曲がり方を表している。

問題は $f_{xy} = f_{yx}$ だ。 x で微分してから y で微分と言われても、あんまりよくわからないよね。そこで多少天下りだが、変数変換

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \quad \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \quad (3.4.13)$$

を考えてみよう。これは、もともとの変数 (x, y) から座標軸を 45° 回転した新しい変数 (u, v) へ移る変換である。新しい変数 u, v で偏微分 f_{yx} を表してみるとどうなるだろうか？連鎖律を使って素直に計算してみると

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial f}{\partial u} + \frac{\partial v}{\partial x} \frac{\partial f}{\partial v} = \frac{1}{\sqrt{2}} \left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} \right] \quad (3.4.14)$$

となる。(余談だが、上の関係は微分演算子の部分だけを取り出して

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial}{\partial u} + \frac{\partial v}{\partial x} \frac{\partial}{\partial v} = \frac{1}{\sqrt{2}} \left[\frac{\partial}{\partial u} - \frac{\partial}{\partial v} \right] \quad (3.4.15)$$

とも書ける。このような書き方は後々、便利だ。) これを y で偏微分して

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &= \frac{1}{\sqrt{2}} \left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v} \right] \left\{ \frac{1}{\sqrt{2}} \left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} \right] \right\} = \frac{1}{2} \left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v} \right] \left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v} \right] \\ &= \frac{1}{2} \left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v \partial u} - \frac{\partial^2 f}{\partial v^2} \right] = \frac{1}{2} \left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial v^2} \right] \end{aligned} \quad (3.4.16)$$

を得る (上では $f_{uv} = f_{vu}$ を仮定した)。何となく変な量ではあるが、新しい座標系での u -方向の曲がり方と v -方向の曲がり方の差が $f_{xy} = f_{yx}$ なのである。

3.4.2 高階偏導函数と連鎖律

上で、高階偏導函数の出る場合の連鎖律の応用例を扱った。これは落ち着いて意味を考えながらやれば何の問題もないが、案外間違いやすいので注意が必要である。一回 x で偏微分したあとの $\frac{\partial f}{\partial u}$ 自身が u, v を通して x に依存しているから $\frac{\partial f}{\partial u}$ を x で偏微分する際にはまた、連鎖律が必要だ、でも慣れないうちはここを良く間違えてしまう。

例えば、 $x = r \cos \theta, y = r \sin \theta$ のときに、 $\frac{\partial^2 f}{\partial x^2}$ を r, θ の微分で表す問題を考えてみる。一回目の微分は簡単だ。連鎖律で

$$\frac{\partial f}{\partial x} = \frac{\partial r}{\partial x} \frac{\partial f}{\partial r} + \frac{\partial \theta}{\partial x} \frac{\partial f}{\partial \theta} = \frac{x}{\sqrt{x^2 + y^2}} \frac{\partial f}{\partial r} - \frac{y}{x^2 + y^2} \frac{\partial f}{\partial \theta} = \frac{x}{r} \frac{\partial f}{\partial r} - \frac{y}{r^2} \frac{\partial f}{\partial \theta} \quad (3.4.17)$$

となるので、微分演算子としては

$$\frac{\partial}{\partial x} = \frac{x}{r} \frac{\partial}{\partial r} - \frac{y}{r^2} \frac{\partial}{\partial \theta} = \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \quad (3.4.18)$$

という作用をもっている。さて、もう一回やるときには

$$\frac{\partial^2 f}{\partial x^2} = \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) \left(\cos \theta \frac{\partial f}{\partial r} - \frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) \quad (3.4.19)$$

となるのだが、**左側の括弧の中の微分はその右側にあるものすべてにかかる**。(右側と言っても、右の括弧内のものだけで、左の括弧内のものにはかからない。念のため。) つまり、しつこく書くと

$$= \cos \theta \frac{\partial}{\partial r} \left(\cos \theta \frac{\partial f}{\partial r} \right) - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \left(\cos \theta \frac{\partial f}{\partial r} \right) + \cos \theta \frac{\partial}{\partial r} \left(-\frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \left(-\frac{\sin \theta}{r} \frac{\partial f}{\partial \theta} \right) \quad (3.4.20)$$

となり、微分演算子の左にあるものは微分されない。また、それぞれの項には「積の微分」を適用する必要がある、例えば、

$$\frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \left(\cos \theta \frac{\partial f}{\partial r} \right) = \frac{\sin \theta}{r} \frac{\partial \cos \theta}{\partial \theta} \times \frac{\partial f}{\partial r} + \frac{\sin \theta}{r} \cos \theta \frac{\partial}{\partial \theta} \left(\frac{\partial f}{\partial r} \right) = -\frac{\sin^2 \theta}{r} \frac{\partial f}{\partial r} + \frac{\sin \theta \cos \theta}{r} \frac{\partial^2 f}{\partial r \partial \theta} \quad (3.4.21)$$

などとなる訳だ。このように計算していくと、最終的な答えは

$$\frac{\partial^2 f}{\partial x^2} = \cos^2 \theta \frac{\partial^2 f}{\partial r^2} + \frac{\sin^2 \theta}{r} \frac{\partial f}{\partial r} + \frac{\sin 2\theta}{r^2} \frac{\partial f}{\partial \theta} + \frac{\sin^2 \theta}{r^2} \frac{\partial^2 f}{\partial \theta^2} - \frac{\sin 2\theta}{r} \frac{\partial^2 f}{\partial r \partial \theta} \quad (3.4.22)$$

となるはずだ。この辺りは落ち着いて、意味を考えながらやれば何ともないはずだが、慣れないうちは非常に間違いやすいから、注意されたし。

問題 3.4.3 2変数の函数 $f(x, y)$ と座標変換 $x = r \cos \theta, y = r \sin \theta$ を考える。このとき、

$$\Delta f(x, y) := \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (3.4.23)$$

を f の r, θ に関する適当な偏微分を用いて表せ。上の Δ を2次元のラプラシアンといい、物理で頻出するだろう。

3.4.3 (補足) 偏導函数がゼロという函数は？ふたたび

以前に「1階偏導函数がゼロ」の函数はどんなものか考えたが、今度は2階偏導函数がゼロのものを考えてみよう。例えば、 $\frac{\partial^2 f}{\partial y \partial x} = f_{xy} \equiv 0$ というのを考えてみる。これは

$$\frac{\partial}{\partial y}(f_x) = 0 \quad (3.4.24)$$

と見ると、「 f_x は y には依存しません」ということなので、

$$f_x(x, y) = g(x) \quad (3.4.25)$$

と書けるはずだ (g は任意の函数, 3.2.1 節を思い出そう). この両辺を x で積分すると, 左辺は $f(x, y)$ になり, 右辺は g の原始函数になるが, x で積分したときの積分定数は y の任意の函数になれる. なぜなら, 積分定数は積分している変数に依存していなければなんでもよく, いまは x で積分しているので, y に依存するのは勝手である. (このところがわかりにくい人は $g(x)$ の原始函数を $G(x)$ とすると, (3.4.25) から

$$\frac{\partial}{\partial x} \{f(x, y) - G(x)\} = 0 \quad (3.4.26)$$

が成り立つこと, 従って, $f(x, y) - G(x)$ は y の任意の函数になれることに注目するとよい.) よって, $g(x)$ の原始函数を $G(x)$ と書いて,

$$f(x, y) = G(x) + h(y) \quad (3.4.27)$$

となることがわかった (h がその「積分定数」としてでてくる y の任意函数). 結局, f は x と y の函数の和であれば何でも良い, という驚愕の (というほどでもないかいな) 事実が得られたのである!

なお, このような考察は, 例えば物理で「波動方程式」を考える時などに良くでてくる.

3.5 平均値の定理

「連鎖律」の応用として, 多変数の場合の平均値の定理が導かれる. これはこの後のテイラー展開と極大極小問題の考察に使える.

定理 3.5.1 (多変数の平均値の定理) 2 変数函数 $f(x, y)$ が C^1 -級の場合,

$$f(a + h, b + k) = f(a, b) + \frac{\partial f}{\partial x}(a + \theta h, b + \theta k)h + \frac{\partial f}{\partial y}(a + \theta h, b + \theta k)k \quad (3.5.1)$$

がなりたつ. (ここで θ は $0 < \theta < 1$ なる適当な数で, 一般に θ は a, b, h, k に依存する.) 同様に, C^1 級の n 変数函数に対しては ($\mathbf{a} = (a_1, a_2, \dots, a_n)$, $\mathbf{h} = (h_1, h_2, \dots, h_n)$)

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\mathbf{a} + \theta \mathbf{h}) h_j \quad (3.5.2)$$

が成り立つ ($0 < \theta < 1$).

証明 簡単だ. $g(t) = f(a + th, b + tk)$ を t の函数と見て, 1 変数函数の平均値の定理を使うと

$$f(a + h, b + k) - f(a, b) = g(1) - g(0) = g'(\theta) \quad (3.5.3)$$

である. ところが, g' については, 「連鎖律」定理 3.3.3 を $x(t) = a + th, y(t) = b + tk$ として用いると

$$g'(t) = \frac{\partial f}{\partial x}x'(t) + \frac{\partial f}{\partial y}y'(t) = \frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k \quad (3.5.4)$$

であるから, 定理 3.5.1 を得る. n 変数の場合も同様である. □

平均値の定理が成立するには, 函数が C^1 級である必要はない. 全微分可能性を仮定すると, 以下の定理になる. この辺りは数学としては興味のあるところだが, 余裕のない人はあまりこだわる必要はない.

定理 3.5.2 (多変数の平均値の定理) 2 変数函数 $f(x, y)$ が全微分可能なら,

$$f(a + h, b + k) = f(a, b) + \frac{\partial f}{\partial x}(a + \theta h, b + \theta k)h + \frac{\partial f}{\partial y}(a + \theta h, b + \theta k)k \quad (3.5.5)$$

がなりたつ. (ここで θ は $0 < \theta < 1$ なる適当な数で, 一般に θ は a, b, h, k に依存する.)

定理 3.5.1 が定理 3.3.3 から出ると同じようにして, 定理 3.5.2 は定理 3.3.8 から証明される. □

3.6 テイラーの定理とテイラー展開

先に、1変数については「テイラーの定理」「テイラー展開」を学習した。この節ではこれを多変数に拡張する。これらの話題はそれ自身でも非常に重要であるが、2階の偏導関数の意味付けも与えてくれる。

簡単のため、2変数の場合を考える。 h, k が小さいとき、 $f(a+h, b+k)$ を $f(a, b)$ で近似するものとして平均値の定理がある。その導き方は（前節でやったように）

$$g(t) = f(a+th, b+tk) - f(a, b) \quad (3.6.1)$$

を考えて、1変数 t に対する平均値の定理を使うものであった。この $g(t)$ は1変数 t の関数なんだから、平均値の定理で止まらずに、 t についてのテイラーの公式やテイラー展開を考えてみるのは自然である。実際、もし $g(t)$ が C^n -級だとすると、

$$f(a+h, b+k) = g(1) = \sum_{k=0}^{n-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(n)}(\theta)}{n!} \quad (0 < \theta < 1) \quad (3.6.2)$$

が成立する。さらに右辺の導関数がいづつ存在してそれは何なのか、については、連鎖律（を何回もつかうこと）が答えてくれる。つまり、一回の微分ごとに $(x(t) = a+th, y(t) = b+tk)$ のつもりで

$$\frac{d}{dt} = \frac{dx(t)}{dt} \frac{\partial}{\partial x} + \frac{dy(t)}{dt} \frac{\partial}{\partial y} = h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \quad (3.6.3)$$

であるから、例えば、 f が C^2 -級ならば

$$\begin{aligned} g^{(2)}(\theta) &= \frac{d}{dt} \frac{dg}{dt} = \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right) \left(h \frac{\partial f}{\partial x} + k \frac{\partial f}{\partial y} \right) \\ &= h^2 \frac{\partial^2 f}{\partial x^2} + hk \frac{\partial^2 f}{\partial x \partial y} + kh \frac{\partial^2 f}{\partial y \partial x} + k^2 \frac{\partial^2 f}{\partial y^2} \end{aligned} \quad (3.6.4)$$

と計算できる（偏微分はすべて $(a+\theta h, b+\theta k)$ での値；またもし f が C^2 -級なら、上の真ん中の2つの項はもちろん、等しい）。上に出てくる偏微分の絶対値は f が C^2 -級なら有界（ $\leq M$ ）であるから、

$$|g^{(2)}(\theta)| \leq 2M(h^2 + k^2) = O(\|\mathbf{a} - \mathbf{c}\|^2) = o(\|\mathbf{a} - \mathbf{c}\|) \quad (3.6.5)$$

が成り立つ（記号を簡単にするため、 $\mathbf{a} = (a, b), \mathbf{c} = (a+h, b+k)$ とおいた）。つまり、(3.6.2) の $n=2$ を考えると、

$$f(a+h, b+k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + o(\|\mathbf{a} - \mathbf{c}\|) \quad (3.6.6)$$

が得られた訳である。期待通り、 $f(a+h, b+k)$ の h, k の1次での近似になっている。

この先もどんどんやれる。 f が C^3 -級だと仮定すると、

$$f(a+h, b+k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + \frac{1}{2} [f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2] + o(\|\mathbf{a} - \mathbf{c}\|^2) \quad (3.6.7)$$

が得られる。今度は h, k の2次式（の3つの可能性）が出ているが、これも当然であろう。 h^2 の係数が f_{xx} 、 hk と kh の係数が f_{xy} と f_{yx} である（これらは f が C^2 -級であることを仮定すれば等しいから、上ではまとめてしまったが）ことにも注意しよう。

以上のような計算を一般化すれば、以下の定理になる。（記号がうるさいから、2変数の函数に限定した。）

定理 3.6.1 2変数の函数 $f(x, y)$ が C^r -級であるとき、適当な $0 < \theta < 1$ に対して

$$\begin{aligned} f(a+h, b+k) &= f(a, b) + \sum_{m=1}^{r-1} \frac{1}{m!} \sum_{\ell=0}^m \binom{m}{\ell} \frac{\partial^m f}{\partial x^\ell \partial y^{m-\ell}}(a, b) h^\ell k^{m-\ell} \\ &\quad + \frac{1}{r!} \sum_{\ell=0}^r \binom{r}{\ell} \frac{\partial^r f}{\partial x^\ell \partial y^{r-\ell}}(a+\theta h, b+\theta k) h^\ell k^{r-\ell} \end{aligned} \quad (3.6.8)$$

が成り立つ。（ θ は、 a, b, h, k, r に依存しうる。）

ともかく、このようにして多変数でもテイラーの公式が成り立つのである。当然、上の公式で $n \rightarrow \infty$ とできる場合には2変数関数のテイラー展開（級数）が成り立つことになるが、概念的には1変数の場合と全く同じだから、これ以上は省略する。むしろ、以下に掲げるような具体例を計算して感覚を身につけることが大事である。

(問題) 次の関数を、あたえられた点 (a, b) の周りで、2次までテイラー展開せよ。つまり、(3.6.7) に相当する式を、(具体的に偏微分を計算して) 書き下せ。

- $f(x, y) = y \sin(x^2 y)$ を $(0, 0)$ の周りで.
- $f(x, y) = x e^{x+y^2}$ を $(0, 0)$ の周りで
- $f(x, y) = \cos(x \sqrt{y})$ を $(\pi, 1)$ の周りで