2006.4.17.

微分積分学・同演習 A （理学部数学科）

担当：原　隆（数理学研究院）：六本松 3-312 号室，tel: 092-726-4774,　e-mail: hara@math.kyushu-u.ac.jp,
http://www.math.kyushu-u.ac.jp/~hara/lectures/lectures-j.html

Office hours: 月曜の午後４時半頃〜６時，僕のオフィスにて．講義終了後にも質問を受け付けますが，４限には別の講義が入っているので，講義の後ではあまり時間がとれません．

概要：

この講義は秋学期の「微分積分学 B」とあわせて完成する．春学期の「微分積分学 A」ではまず「偏微分」を高校と同じノリで（つまり，厳密性にはあまりこだわらず）学習する．そのあとで『本格的に厳密な大学の数学』に入り,「極限とは何か（その厳密な定義）」「１変数関数の微分とその応用」などを扱う．秋学期には「１変数関数の積分」や「多変数関数の微分」を厳密に扱う予定だが，春学期の積み残しも入るかも．なお，微積分では「級数論」も非常に大事だが，これは秋学期の箱崎日「数学概論 I」にほぼ丸投げする予定である．

春学期でキーとなる概念：偏微分，極限，ϵ-δ 論法，コーシー列，微分,

秋学期でキーとなる概念：極限，ϵ-δ 論法,（級数），積分，陰関数定理

特に講義を通して身につけて欲しいこと：

この講義で学んでほしい「能力」は以下の２つである．

- 微分や積分のいろいろな概念を習得し，実際に応用して使えるようになること
- その際，厳密に議論が展開でき，自分の議論に自信が持てるようになること．

「数学」の講義なんだから両者は切り離せないはずだが，高校までの数学では主に最初の面に力点が置かれていた．実際，極限や積分の定義には曖昧さが入り込む余地があるのだが，高校まではそのような曖昧さが入り込まないような例に限って,「応用して使える」ことを目指していた訳だ．

しかし，そのような限定された例だけでは話が閉じなくなってくる．これは数学に限らず，物理学や工学などで出てくる問題に関しても同じだ．そのような問題にまともに取り組むには，高校で学んだつもりの極限，微分，積分などの意味を問い直すことから始めなければならない．

そのため，この講義のかなりの部分は「厳密な理論を展開する」ことにあてられる．しかし，抽象的な「厳密理論」を振りかざすのは「畳の上の水練」と同じで意味がない.（そもそも，具体例を書いた「抽象論」にどれほどの意味があるのか？）そのため，具体的な微分積分の話題を用いて，厳密な数学理論の展開方法を学んで行く．

履修上の注意：

1. この講義は数学科向けのもので，他学科向けのものとは少し異なる．従って，他学科の学生さんの再履修にはお勧めしない.（それを承知でとってみたいという人は自己責任でどうぞ.）
2. 箱崎日に開講される「数学入門」で学習したことを前提にして，微分積分学 A の講義を進める．特に「数学入門」の最初の５回の内容は必須である．従って,「数学入門」をぜひ履修すること．

内容予定：（以下は大体の目安です．皆さんの理解度により，ある程度の変更や増減はあり得ます.）

1. 足慣らし：高校のノリでの偏微分（４回程度）
   1. ２変数の関数とは何か？
   2. 偏微分の定義とその意味
   3. 連鎖律
   4. 高階の偏微分
   5. 偏微分の応用：２変数関数の極値問題（もしかしたら秋学期に廻すかも）

2. ϵ-δ 論法と極限（５回程度）
   1. 極限の厳密な定義：ϵ-N 論法，ϵ-δ 論法
   2. 上極限，下極限,
   3. 単調増加列とコーシー列

3. 連続関数（２回程度）

1. 連続関数
2. 最大値最小値の定理，中間値の定理

4. 関数と微分（この学期の残りの時間；最後の方は秋に廻すかも）

1. 微分の厳密な定義
2. 平均値の定理
3. 関数の増減と凹凸
4. テイラーの定理とテイラー展開（これは秋学期か？）

教科書： 田島一郎「解析入門」（岩波書店）．ただし，この教科書には１変数関数の微積分しか載っていないので，初めの数回（偏微分）に関しては，配布するプリントを用いる．

参考書： 上の教科書は今時の学生さんにも読みやすいものとして選んだが，すこし物足りない部分もある．数学科の学生ならば，以下に掲げる参考書を一組は買ってほしい.（今すぐに読まなくても，後々で役に立つ時があろう.）

- 高木貞治「解析概論」（岩波）．今の学生さんには難しすぎる，との意見もあるが，不朽の名著だ．
- 小平邦彦「解析入門 I, II」（岩波）．上の解析概論を少しとっつきやすくした感じ．記述はおおむね平明かつ直感的で，名著といえよう．
- 杉浦光夫「解析入門 1, 2」（東大出版会）．最近の硬派の定番ともいえる．かなり分厚いけど，その分，記述は丁寧で読み応えはあるようだ．
- 溝畑茂「数学解析　上・下」（朝倉書店）．かなりユニークな本である．特に，微分と積分が渾然一体となって展開される点は非常に面白い．読み応えは非常にあるが，最初はかなり難しいと感じるだろう．

以上の教科書，参考書は「田島本」「小平本」などと引用する事がある．

評価方法：

中間試験（＋レポート）と 期末試験 の成績を総合して評価する．そのルールは以下の通り（ただし，例外あり．後の注意を参照）：

- 最終成績は一旦，１００点満点に換算してから，この大学の様式に従ってつける．
- その１００点満点（最終素点）は，以下のように計算する．
  - まず,「中間試験（＋レポート）の点」「期末試験の点」をそれぞれ 100 点満点で出す.
  - 次にこの２つを以下の式で「平均」し，一応の総合点を出す：

$$(\text{総合点 } A) = 0.50 \times (\text{中間（＋レポート）の点}) + 0.50 \times (\text{期末の点})$$

  - ただし，上の計算式の重みを若干変更する可能性はあることを承知されたい（例えば，総合点 A で，中間と期末の比を 4 : 6 にするなど）．
  - 最終素点は

$$(\text{最終素点}) = \max\{\,(\text{総合点 } A), (\text{期末の点})\,\}$$

  とする．つまり,（総合点 A）と（期末の点）を比べて，良い方をとる のだ．
- 上の「最終素点」をよく見て，必要ならば全体に少し修正を加えたものをつくり，これをこの大学の基準と合わせて最終成績を出す.
- レポートの点は原則として，総合点 A には加えない．ただし，上の計算では合格基準に少し足りない人（百点満点で 10 点不足が限度）を助けるかどうかに使用する．また，レポートがずば抜けて良い場合，この事実は最終成績に反映される事もある．

（期末一発逆転の可能性と例外について）

- この講義では（上位 10%の人だけがわかるような）進んだ話題はあまり扱わない．そのため,「できる」人が退屈することも考えられる．そのような人は自主的な学習を奨める意味で，講義などに出なくても「期末で一発逆転」も可能なようにした.

- ただし,「期末の一発勝負」がうまくいく人はそれほど多くないだろうと思われる（期末試験は中間試験やレポートよりは難しい）から，あくまで自己責任でやってくれ．期末の一発勝負に出て成績が悪くても，苦情は一切受け付けないからね！(できる人が少ないだろうと思いつつもこの形式をとるのは，僕の美学にこだわっているからである.）
- （重要な例外）上のルールは最大限，尊重する．ただし，最終成績が９０点以上になりそうな人については,他のことも加味して考慮することがある．つまり,「期末の点は確かに良いのだが，本当にわかって書いているのか，どうも疑問だ」などの場合には，形式的に９０点以上になっても最終成績を９０未満に引きずり落とす事がありうる.（今までにそんなことをした事はないが，可能性はゼロではない.）もちろん，これをやった場合には反論の余地は与えます.

「学習到達度再調査」について：

僕はこの大学は赴任して３年目だが,「学習到達度再調査」とかいう，変な制度があるらしい．この科目は必修科目でもあり，これに変に期待する人がいるかもしれないので，ここではっきり，宣言しておこう.

> 「再調査」は行わない可能性が高い．もし行うとしても，その権利を得るのはギリギリで不合格になった人だけである．再調査を行うか，誰を対象とするかは，こちらの一存で（もちろん公平に，しかし厳しく）決めさせていただく.

去年は「再調査」を積極的に利用したが，今年は僕が国際会議に行ってしまうかも知れないので，再調査をやりたくてもやれない可能性もあるから，再調査には期待しないように.（もちろん，再調査とは独立に，正規の理由があれば追試験は行うのでご安心を.）

更に付言するならば，再調査をする方が，こちらとしては厳しく点を付けやすい（厳しくつけておいて，誰を助けるかは再調査できちんと確かめれば良いから）.

だから，このようなものには頼らず，期末試験まででちゃんと合格できるよう，しっかり学習して下さい．期末試験までなら皆さんの学習を助ける努力は惜しまないつもりで，質問などにも忍耐強く相手することを保証する.

合格（最低）基準：

合格のための条件（A, B がとれる条件ではない！）は，講義中に出題する例題，レポート問題と同レベルの問題が解けることである.（ただし「時間がなくてレポートは出せないけど試験には出すぞ」などの指示を講義中に与えることもあり得る.）具体的に書くと，大体，以下のようになる（進度の都合で内容に若干の変更があるので，完全なリストを現時点で呈示する事はできない.）.

- １変数関数の微分とその応用について，厳密性を少し犠牲にしても良いから，計算ができること.
- ２変数関数の微分（偏微分）についてもその定義，計算法，応用がわかること.
- ϵ-δ 論法の基礎を理解しており，非常に簡単な（講義中に示すような）例題が解けること.

（これはあくまで最低基準であり，この講義のかなりの部分はより突っ込んだ極限の理論 — ϵ-δ 論法，コーシー列 — にあてられる.）

レポート，宿題について：

ほぼ毎回，簡単なレポートや「お奨めの宿題問題」を出す予定である．レポートの方は TA （teaching assistant）の方に採点してもらい，２週間後に返却の予定．これらの出題意図は「この程度できれば講義についていけるし，合格も可能だ」という目安を与えることと家庭学習の引き金にすること，である．成績評価に占めるレポートの比重は低いが，この講義をこなす上では重要な意味があるので，是非やること.

重要：レポートは友達と相談した結果を書いても良い．ただし，誰と相談したかは明記すること.（「俺は人に教えてやっただけで人からは全く教わってない」と思う人は書かなくても良いが.）相談した人の名前を書かせるのは,それで点数を左右するのが目的ではない.「お世話になった文献，人にはきちんと感謝する」という，科学上の最低ルールを守ってもらうためである.

プリントの使いかた：

必要に応じてプリントを配る予定である．これらのプリントは板書にアップアップしないでも講義が聴けるように，また，教科書の足りないところを補うために，配っているものである．なお，急いで作っているためにプリントにはタイプミスなどがかなりあると思うので，気づいたらできるだけ指摘してくれるとありがたい．

特に注意を要する題材：

１．この講義の一つの山場はやはり ϵ-δ 論法（とコーシー列）で，かなりの人が戸惑うでしょう．しかし，これはそんなに難しいものではなく，ゆっくり考えればさえすれば誰でも理解できます．この講義でも「数学入門」との連携を考えるなど，理解を助けるための努力をしています．ですから皆さんも諦めずに学習して下さい．(「わからない」という人は，初めから考える努力を放棄していることが多い．質問に来てくれたら，僕はかなりの忍耐をもってつきあいますし，そのうちにわかるようになる人も去年，何人かいました.）わからなくても，ここでこれからの４年間が決まると思って頑張って下さい．

２．もし万が一,「ϵ-δ 論法がわからない」場合も諦めないこと．ϵ-δ 論法がわからなくても，「偏微分」「微分の直感的意味」などの応用面の概要を理解することは十分に可能です．それらをやっているうちに ϵ-δ 論法がわかってくることもあるはず．

３．初めにやる「偏微分」は ϵ-δ に比べたら簡単に見えるため，馬鹿にしていて後で困る学生さんがかなり出る可能性があります．概念的には簡単だからと思っても，気は抜かないで下さい．

この科目に関するルール：

世相の移り変わりは激しく，僕が学生だったときには想像すらできなかったことが大学で行われるようになりました．そのうちのいくつかは良いことですが，悪いこともあります．オヤジだとの批判は覚悟の上で，互いの利益のために，以下のルールを定めます．

- まず初めに，学生生活の最大の目的は勉強することであると確認する．
- 講義中の私語，ケータイの使用はつつしむ．途中入室もできるだけ避ける（どうしても必要な場合は周囲の邪魔にならないように）．これらはいずれも講義に参加している他の学生さんへの最低限のエチケットです．
- 僕の方では時間通りに講義をはじめ、時間通りに終わるよう心がける．
- 重要な連絡・資料の配付は原則として講義を通して行う（補助として僕のホームページも使う —— アドレスは http://www.math.kyushu-u.ac.jp/~hara/lectures/lectures-j.html)．「講義に欠席したから知らなかった」などの苦情は一切，受け付けない．
- レポートを課した場合，その期限は厳密に取り扱う．
- E-mail による質問はいつでも受け付ける（hara@math.kyushu-u.ac.jp)．ただ，回答までには数日の余裕を見込んで下さい．なお，学生さんのメイルが往々にして spam mail に分類されてしまう事があります（多分，html mail で送られてくると自動的にスパムにされてしまうのだろう）．従って，僕にメイルしたのに，２，３日しても返事がない場合は返事を催促して下さい．たとえどんなに理不尽なメイルであっても，僕は返事はすることにしていますので，返事がないのはメイルが届いていない可能性が高いです．

講義を少し離れて一般的な勉強法など

> **（本論に入る前のお断り）以下では少し苦言を呈する．このような苦言を書くと自分が全面的に否定されたように感じる人がいるかもしれないが，それは僕の意図するところではない．自分が否定されたと思う前に，世の中には完璧な人間などいず，みんな長所と短所を持っていること，従って自分にも長所が一杯あるが短所も持っている（だろう）ことを，まず思い出して欲しい．皆さんの世代は我々の世代になかった良い面も多く持っているが，こと学問のやり方については未熟な面が我々の世代よりも多く見られる，というのが僕の指摘したい事であり，皆さんの世代を全否定するつもりは毛頭ない．この点，大丈夫とは思うが，決して誤解のないように念の為に強調しておく．**

もう皆さんは「学力低下」について耳にタコができるほど効かされているでしょうし,「そんなこと言われても」と反発を感じてもいるでしょう．確かにこんなことばかり聞かされたら反発するのもわかるけども，我々教官が実際に「学力低下」を感じているのも事実です．より正確に言うと，学力そのものよりも,「**学問に対する態度**」の面で昔と今ではかなりの差があるように思われます．態度（方法）が良くないといくら時間をかけても成果は望めません．ですから学問に対する正しい態度を身につけるつもりで，以下に述べる事に気を配って勉強することを強く奨めます.

- 講義の復習として，ともかく自分のノートや配られたプリント，教科書の対応する場所を読み直すこと.
- その際，教科書，プリントなどは少なくとも 3回は読む こと，１回目に読んだ時はわからなくても，３回目には何となくわかってくることもある．何となくわかったらもう一回読む.
- 第一原則として，自分の納得するまで考えて理解する ことを目指す.「考える」ことは「覚える」ことより百倍も大変だから，自分を無理にでも追い込んで考えるように努力しないと考える習慣は身に付かない.
- ただし，あまりに行き詰まったら，気分転換も兼ねて演習問題や演習書などをやる．特に教科書，プリントなどの 演習問題はともかく自力でやってみる 事．人間はみんな（もちろん僕も）アホだから，ある程度の訓練を通して慣れない限り，理解する事はほとんど不可能だ[1]．また具体的に手を動かすことで,「わかったつもりで全然わかってない」ことが見つかるかもしれない．だから問題を解く事が重要なのだ.
- 新しい概念などがわからない時は，その「定義」がそもそもわかってないことが非常に多い．重要な概念の定義が言えるか，自答しよう．定義が言えない時は定義を覚えられるまで，具体例を考えよう.（意味もわからずに定義を丸暗記するのは，たいていの場合は無駄だが，やらないよりはましだ．しかしもっと良いのは，具体例を考えているうちに自然と定義が覚えられてしまう事だ.）具体例さえ思い浮かばない時はかなりの重症です．友達や教官に質問しましょう.
- （上と関連するが）何かを定義したら，その 意味は何か，なぜそれを定義したいのか をいつも考えよう.
- 定義，定理などでは 反例 を常に思い浮かべるように.「定理のこの条件がなくなったらどこが困るのか」などを考えるとより身近に感じられるかもしれない.
- ここは大学で，これまでのように手取り足取りはしてくれない（少なくとも僕はしない）ことを思い出そう．皆さんが自分から動けば道は開けるけども，助けてくれるのを待っているだけでは何も解決しないよ.

（まとめ）学問に王道なし．地道な日々の努力が最後には実を結ぶのだ.

演習書の奨め：

教科書に載っている例題や節末問題，商末問題はできるだけやること．それでもわかった気がしなかったら，演習書（いわゆる問題集）をやることを勧めます．問題をやることによって，自分が曖昧にしかわかっていなかった部分がはっきりしてくることが多い．ただし，その際，**解答を鵜呑みにはせず，自分で納得するまで考えること．考えてもわからなかったら，友達や教官（僕を含む）に訊けばよい**．同じ理由で問題の解答を頭から覚える愚だけは避ける事.

演習書はどれでも良いが，一応，目についたものを列挙すると：

- 三村征雄編「大学演習 微分積分学」（裳華房）— 僕はこれを使った．ちょっとムズイかもね.

[1]昨今，受験勉強の分量すら減っているのではないか（高校での計算練習が絶対的に不足しているのでは？）と思わされる事がある

- 蟹江，桑垣，笠原「演習詳説 微分積分学」（培風館）— なかなか良いが，はじめは難しく感じるかも．
- 杉浦ほか「解析演習」（東大出版会） — これもまあ，大変ではありますが，良い本．
- 鶴丸ほか「微分積分 — 解説と演習」（内田老鶴圃） — 一番「普通」かも．
- 飯高茂監修「微積分と集合　そのまま使える答えの書き方」（講談社サイエンティフィック） — 題名は変だけど，馬鹿にはできない，なかなかの本．流石は飯高さん監修だけあるな．案外，おすすめ．

これ以外にもいくらでも出版されてるから，図書館や本屋さんで自分にあった（読みやすい，やる気になる）ものを選べば良い．ただしその際，解答や解説の詳しいものがよい．また，無理をして難しすぎるものを選ぶ必要はない．自分が簡単だと思うことでも,（人間はアホやから）わかってないことが一杯あり，むしろ簡単なところが盲点になって先に進めないのだ．簡単な演習書でもやれば，大きな効果があるはず．

なお，受験と違って死ぬほどの問題量をこなす必要はありません[2]— 自分が納得できるようにいくつか例題をやり，弱いところだけたくさんやれば大抵は十分です．

（ついでに気がついた本）「共立ワンポイント数学双書」というシリーズの中に「イプシロン–デルタ」とか「テイラー展開」のものがある．小さな本ではあるけど，トピックごとにわかりやすく書かれているから，ピンポイントでの勉強に適しています（特に５月半ば以降に）．

本論に入る前に記号のお約束．

$a<b$ を２つの実数，n を非負（負でない）整数とする．

- 整数の全体は $\mathbb{Z}$，自然数（1 以上の整数）の全体を $\mathbb{N}$，有理数の全体を $\mathbb{Q}$，実数の全体は $\mathbb{R}$ と書く．
- 高校までと異なり,「$a<b$ または $a=b$」を $a\le b$ と書く．同様に,「$a>b$ または $a=b$」を $a\ge b$ と書く．
- $a<x<b$ なるすべての実数の集合を (a,b) と書き，開区間という．
- $a\le x\le b$ なるすべての実数の集合を $[a,b]$ と書き，閉区間という．
- 高校と同じく，$n!=n\cdot(n-1)\cdot(n-2)\cdots 2\cdot 1$ は n の階乗である．ただし，$0!=1$ と約束する．

（用語の注）あるものがたった一通りに決まる（存在する）とき，業界用語では**○○が一意に決まる（存在する）**という．この表現（一意）は頻出するから覚えよう（英語の unique, uniquely の訳）．

「定義」と「定理」の意味はわかっているだろう．これの仲間として，以下のようなものがある.（使い分けはかなり，その時の気分による.）

- 定理：いうまでもなく，非常に重要で，一般的な結果．
- 命題：定理ほどは重要でない（または一般的ではない）が，そこそこ重要なもの．
- 補題：定理を導くのに用いられる，補助的な結果．
- 系：定理や命題から，ほんの少しだけ余分に頑張れば導かれる結果．

次ページから講義用のプリントの本体が始まる．このプリントの作成にはもちろん，いろいろな本，特に先に挙げた教科書と参考書を参考にした．

[2]いや，さっきも書いたけど，最近は受験勉強に於ける問題量が絶対的に不足しているようにも思えるが

1　多変数関数の微分

既に宣言した通り，この講義ではまず，偏微分（多変数関数の微分）を扱う．厳密性にはこだわらず，高校までの「ええ加減」なノリで[3]，ともかく概念を理解する事を目標にする．いくつかの重要な定理の証明は後回しになるだろう．その後で（５月の半ばから）本格的に厳密な大学の数学を始める．

なお，たいていの場合は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．

1.1　１変数の関数，多変数の関数

関数とは何か，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，

$$x^2, \quad x^4, \quad \sin x \quad \cos(x^2+2), \quad \ldots \tag{1.1.1}$$

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．１変数 x の全体は実軸全体を動くから，これは実軸（直線）から実軸への「写像」である．すべての実数値に対して値 $f(x)$ が定義されていなくても関数という．より正確に定義すると（高校でも少しはやったはず），

定義 1.1.1 D を実数の集合とする．定義域が D である実数値関数 f とは，D の各点 x に対して「関数の値」$f(x)$ を定める対応関係のことである．この事を

$$\begin{array}{cccc} f: & D & \to & \mathbb{R} \\ & \Cup & & \Cup \\ & x & \mapsto & f(x) \end{array} \tag{1.1.2}$$

と表現する．

後の方では定義域が大事になってくるが，当座はあまり気にしなくても良い．

（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か）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．この点についてはおいおい，５月以降にやりましょう．

１変数関数と同じノリで多変数の関数も考えるが，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[4]．

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．従って２変数の関数は平面から実軸への写像と考えられる．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．

[3]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

[4]いつの時代からか,「絶対値はとも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

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[5]．

- まず，n 次元空間の点を，$\boldsymbol{x}$ のように太字で書く：$\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$．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）では太字で書くのだ．（もっ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：$\mathbb{R}^n := \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \,\big|\, x_1, x_2, \ldots, x_n \in \mathbb{R}\}$．

そして

定義 1.1.2 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 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ に対して「関数の値」$f(\boldsymbol{x}) = f(x_1, x_2, \ldots, x_n)$ を定める対応関係のことである．この事を

$$\begin{array}{cccc} f: & D & \to & \mathbb{R} \\ & \Cup & & \Cup \\ & \boldsymbol{x} & \mapsto & f(\boldsymbol{x}) \end{array} \tag{1.1.3}$$

と表現する．

さらに１次元の時に倣って

- 点 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ と $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)$ の間の <u>距離</u> を

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \Bigl(\sum_{j=1}^{n} (x_j - y_j)^2\Bigr)^{1/2} \tag{1.1.4}$$

  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 $\boldsymbol{x} = (x, y)$，$\boldsymbol{z} = (u, v)$ などとも書く．(だいたい,「空間内の点」のような幾何学的視点を強調するときには $f(\boldsymbol{x})$ と書く．それに比べて，$\boldsymbol{x}$ 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 <u>開集合でかつ連結</u> なものを <u>領域</u>（domain, region）という．(これが何かは簡単に説明する．今は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高校数学と同じノリで考える（厳密理論は５月半ばから）．まずは１変数の場合を思い出そう．

定義 1.1.3 (１変数関数の極限) x の関数 $f(x)$ に対して $\lim_{x \to a} f(x) = \alpha$ とは，以下が成り立つことをいう．

$$|x - a| \to 0 \quad \text{ならば} \quad \bigl|f(\boldsymbol{x}) - \alpha\bigr| \to 0 \tag{1.1.5}$$

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（高校数学のノリで）定義すると以下のようになる．

定義 1.1.4 (多変数関数の極限) n 変数関数 $f(\boldsymbol{x})$ に対して $\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = \alpha$ とは，以下が成り立つことをいう．

$$\|\boldsymbol{x} - \boldsymbol{a}\| \to 0 \quad \text{ならば} \quad \bigl|f(\boldsymbol{x}) - \alpha\bigr| \to 0 \tag{1.1.6}$$

１変数の時の $|x - a| \to 0$ の条件が，$\|\boldsymbol{x} - \boldsymbol{a}\| \to 0$ に変わっただけで，どちらも「２点の距離がゼロに行く」極限を考えている．

注意：２変数以上が１変数と違うところ： $\|\boldsymbol{x} - \boldsymbol{a}\|$ というのは２点 $\boldsymbol{a}$ と $\boldsymbol{x}$ 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，$x \to a$ 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

[5]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 が出ているのだが，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

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，$\boldsymbol{x}$ が $\boldsymbol{a}$ へどのような近づき方をしても $f(\boldsymbol{x})$ 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 1.1.5 (多変数関数の連続性の定義) n 変数関数 $f(\boldsymbol{x})$ が $\boldsymbol{x}=\boldsymbol{a}$ で連続 とは，$\lim_{\boldsymbol{x}\to\boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a})$ 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 $\boldsymbol{x}\to\boldsymbol{a}$ 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 $\boldsymbol{x}, \boldsymbol{a}$ 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，慣れてもらうつもりで敢えて書いてみた.）

1.2 偏微分

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 1.2.1 (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

$$\lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h} = \lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \tag{1.2.1}$$

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 $\dfrac{\partial f}{\partial x}(a,b)$，$f_1(a,b)$，$f_x(a,b)$，$D_1 f(a,b)$ などと書く．同様に，第２変数に関する偏微分係数とは

$$\lim_{h\to 0}\frac{f(a,b+h)-f(a,b)}{h} = \lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b} \tag{1.2.2}$$

のことであって，$\dfrac{\partial f}{\partial y}(a,b)$，$f_2(a,b)$, $f_y(a,b)$，$D_2 f(a,b)$ 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）偏導関数と呼ぶ.

（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，$\dfrac{\partial f}{\partial x}(a,b)$ 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うが，式がどうしても複雑になって混乱するといけないので，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 1.2.2 (C^1-級) 多変数関数 $f(x_1,x_2,\ldots,x_n)$ がその定義域（の一部）D で

- f は各変数 $x_1,x_2,\ldots,x_n$ のそれぞれについて偏微分可能で
- かつ，その n-この偏導関数が $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ の連続関数である

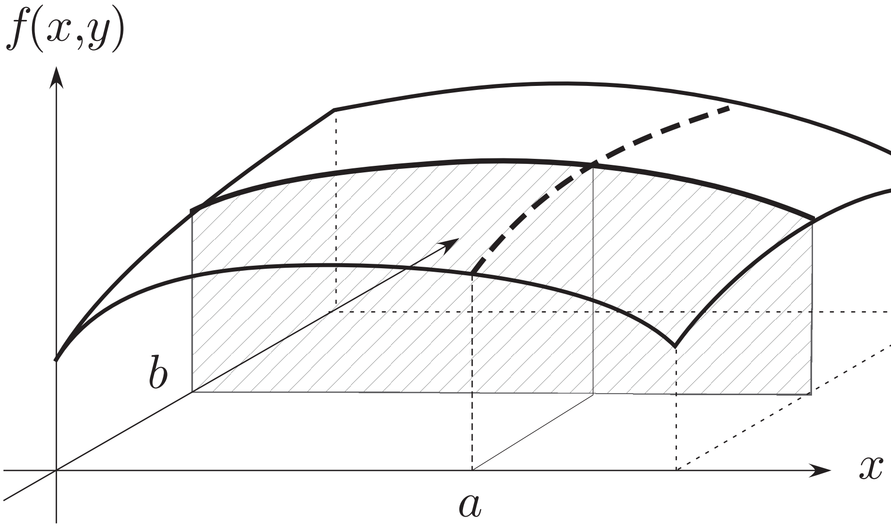
であるとき，f は D で C^1-級 であるという.

（大体想像がつくと思うが）この後で「高階の偏導関数」を学ぶ．そうすると n-階までの偏導関数がすべて存在してかつ連続，な関数を C^n-級という．これらの定義では（考えている階数までの）すべての偏導関数の存在と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．



問 1.2.1. 次の関数をそれぞれの独立変数で偏微分せよ．

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

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

1.2.1 偏導関数がゼロ，の関数は？

１変数の関数 f の場合，導関数 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，多変数の関数では事情が異なる．例えば，２変数関数 $f(x,y)$ 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

$$f(x,y) = g(y) \qquad g \text{ は任意の関数} \tag{1.2.3}$$

と書ける．これは一般には定数関数ではない！

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

（記号についての注意）

$f(x,y)$ の偏導関数 $\frac{\partial f}{\partial x}$ の記号であるが，前回のプリントに書いたように，$\frac{\partial f}{\partial x}$ $D_x f$ $D_1 f$ $\partial_x f$ $\partial_1 f$ 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（溝畑本が一例），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' とだけ書いて $\frac{\partial f}{\partial x}$ のつもりである人がいるから，念のために注意しておく．

1.2.2 方向微分[6]

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変数の関数 $f(x,y)$ 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，この関数の 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 の関数として，もっと他の方向での変化率を考えたくなることもあるだろう．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）$x=y$ の直線にそって x, y が動いた時にどうなるか，なども見たい．

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 1.2.3 (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）$\boldsymbol{v}=(v_x,v_y)$ が与えられたとせよ．極限

$$f_{\boldsymbol{v}}(a,b) := \lim_{h\to 0} \frac{f(a+hv_x, b+hv_y)-f(a,b)}{h} \tag{1.2.4}$$

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

$$f_{\boldsymbol{v}}(\boldsymbol{a}) := \lim_{h\to 0} \frac{f(\boldsymbol{a}+h\boldsymbol{v})-f(\boldsymbol{a})}{h} \tag{1.2.5}$$

が存在するなら，これを $f(\boldsymbol{x})$ の点 $\boldsymbol{a}$ における $\boldsymbol{v}$ 方向の <u>方向微係数</u> という．この方向微分は $D_{\boldsymbol{v}}f(\boldsymbol{a})$ とも書く．

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

さて，関数 $f(\boldsymbol{a})$ 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

$$f(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad f(x,y)=\frac{2xy}{x^2+y^2} \tag{1.2.6}$$

$$g(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad g(x,y)=\frac{xy}{\sqrt{x^2+y^2}} \tag{1.2.7}$$

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，

$$f_x(0,0)=f_y(0,0)=0, \qquad (x,y)\neq(0,0) \quad \text{では} \quad f_x(x,y)=\frac{2y(y^2-x^2)}{(x^2+y^2)^2}, \quad f_y(x,y)=\frac{2x(x^2-y^2)}{(x^2+y^2)^2} \tag{1.2.8}$$

$$g_x(0,0)=g_y(0,0)=0, \qquad (x,y)\neq(0,0) \quad \text{では} \quad g_x(x,y)=\frac{y^3}{(x^2+y^2)^{3/2}}, \qquad g_y(x,y)=\frac{x^3}{(x^2+y^2)^{3/2}} \tag{1.2.9}$$

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）．

---

[6] 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．もちろん，数学科の学生ならちゃんと理解して欲しい事ではあるが，入学後まもないから，ある程度の理解で良しとする．ただし，初めから諦めずに理解しようとする努力はすること．同じ注意はこの次の小節 — 全微分可能性 — にも当てはまる．

1.2.3 全微分可能性[7]

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：

$$f(x)\approx f(a)+f'(a)\times(x-a). \tag{1.2.10}$$

我々は当然，偏微分にも同じ役割を担ってほしい．つまり，$\boldsymbol{x}=\boldsymbol{a}$ の点の近傍での $f(\boldsymbol{x})$ 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のだ．例えば先の (1.2.6)，(1.2.7) 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[8]，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」を導入したのだった．

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，$\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$，および $\|\boldsymbol{x}-\boldsymbol{a}\|=\Big(\sum_{j=1}^{n}(x_j-a_j)^2\Big)^{1/2}$ である．

定義 1.2.4 (全微分可能性) ある領域 D で定義された n 変数関数 $f(\boldsymbol{x})$ と，D 内の1点 $\boldsymbol{a}$ がある．定数 $A_1,A_2,\ldots,A_n$ が存在して（「定数」という意味は $\boldsymbol{x}$ に依存しないということ．もちろん，$\boldsymbol{a}$ には依存してよい），

$$f(\boldsymbol{x})=f(\boldsymbol{a})+\sum_{j=1}^{n}A_j(x_j-a_j)+\tilde{f}(\boldsymbol{x}),\qquad \text{with}\quad \lim_{\boldsymbol{x}\to\boldsymbol{a}}\frac{\tilde{f}(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|}=0 \tag{1.2.11}$$

が成り立つとき，f は $\boldsymbol{x}=\boldsymbol{a}$ 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，(1.2.6) や (1.2.7) 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 (1.2.11) が $\boldsymbol{a}$ に近いすべての $\boldsymbol{x}$，つまり $\boldsymbol{a}$ 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）のだ．まとめると，以下の命題になる．

命題 1.2.5 ある領域 D で定義された n 変数関数 $f(\boldsymbol{x})$ と，D 内の1点 $\boldsymbol{a}$ があって，$f(\boldsymbol{x})$ が $\boldsymbol{x}=\boldsymbol{a}$ で全微分可能だとする．このとき，$\boldsymbol{x}=\boldsymbol{a}$ において

1. $f(\boldsymbol{x})$ は 連続 であり，
2. $f(\boldsymbol{x})$ はすべての x_j について 偏微分可能 で，

$$\Big(\text{(1.2.11) の } A_j\Big)=\frac{\partial f}{\partial x_j}(\boldsymbol{a})\qquad (j=1,2,\ldots,n), \tag{1.2.12}$$

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

$$D_{\boldsymbol{v}}f(\boldsymbol{a})=\sum_{j=1}^{n}A_jv_j=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}v_j \tag{1.2.13}$$

[7] 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，講義でもあまり触れない．また，この小節の内容を良く理解するには後からやる「オーダー」の概念が役立つ．ただ，偏微分の話をして全微分の話をしないのは筋が通らないので，簡単に触れる．

[8] この事実がアタリマエに見えない人が多いと思うが，心配はいらない．このようなことは5月以降，じっくりやります

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 (1.2.11) は $f(\boldsymbol{x})-f(\boldsymbol{a})$ 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 $x_2,x_3,...$ は $a_2,a_3,...$ に固定して考えるので，(1.2.11) から

$$\frac{f(x_1,a_2,a_3,\ldots,a_n)-f(a_1,a_2,a_3,\ldots,a_n)}{x_1-a_1}=A_1+\frac{\tilde{f}(\boldsymbol{x})}{x_1-a_1} \tag{1.2.14}$$

の $x_1\to a_1$ の極限が $\frac{\partial f}{\partial x_1}$ を与えることになる．ところが，$x_2,x_3,...$ を $a_2,a_3,...$ に固定した場合は $|x_1-a_1|=\|\boldsymbol{x}-\boldsymbol{a}\|$ であるので，(1.2.11) から $\dfrac{\tilde{f}(\boldsymbol{x})}{x_1-a_1}$ 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 (1.2.11) から

$$\frac{f(\boldsymbol{a}+h\boldsymbol{v})-f(\boldsymbol{a})}{h}=\frac{\displaystyle\sum_{j=1}^{n}A_jhv_j+\tilde{f}(\boldsymbol{a}+h\boldsymbol{v})}{h}=\sum_{j=1}^{n}A_jv_j+\frac{\tilde{f}(\boldsymbol{a}+h\boldsymbol{v})}{h} \tag{1.2.15}$$

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

$$\frac{\tilde{f}(\boldsymbol{a}+h\boldsymbol{v})}{h}=\frac{f(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|} \tag{1.2.16}$$

であるため，(1.2.11) から，(1.2.15) の最後の項は $h\to 0$ でゼロに行く．よって，(1.2.13) が証明される． □

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 (1.2.11) は図形的には以下のように解釈できる．まず，(1.2.11) の最後の項がない場合を考えると，定数 A,B があって

$$f(x,y)=f(a,b)+A(x-a)+B(y-b) \tag{1.2.17}$$

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 $z-c=A(x-a)+B(y-b)$（$c=f(a,b)$ 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（線形代数でやるはずなのだが現時点ではちょっと苦しいかな）．この平面を S としよう．

実際には (1.2.11) 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，$\boldsymbol{x}-\boldsymbol{a}$ 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 <u>接平面</u> になっているわけだ．

つまり，全微分可能の条件 (1.2.11) 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 <u>接平面で良く近似できる</u> 条件とも解釈できるのである．(1.2.6) や (1.2.7) 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 1.2.6 (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 $f(\boldsymbol{x})$ 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，$f(\boldsymbol{x})$ 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，今の段階では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，$\boldsymbol{a}:=(a,b),\boldsymbol{x}=(x,y)$ と書く．全微分可能であることをいうためには，差

$$f(x,y)-f(a,b)=\bigl\{f(x,y)-f(a,y)\bigr\}+\bigl\{f(a,y)-f(a,b)\bigr\} \tag{1.2.18}$$

が $\boldsymbol{x}\to\boldsymbol{a}$ でどのように振る舞うか — (1.2.11) を満たすか — を調べなければならない．

さて，(1.2.18) 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）

$$f(a,y)-f(a,b)=f_y(a,\tilde{y})\times(y-b) \tag{1.2.19}$$

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）

$$f(x,y)-f(a,y)=f_x(\tilde{x},y)\times(x-a) \tag{1.2.20}$$

となる（$\tilde{x}$ は a と x の間の適当な数）．これを (1.2.18) に代入して

$$f(x,y)-f(a,b)=f_x(\tilde{x},y)\times(x-a)+f_y(a,\tilde{y})\times(y-b) \tag{1.2.21}$$

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

$$f_x(u,v)=f_x(a,b)+g(u,v), \qquad \text{with} \qquad \lim_{(u,v)\to(a,b)} g(u,v)=0, \tag{1.2.22}$$

$$f_y(u,v)=f_y(a,b)+h(u,v), \qquad \text{with} \qquad \lim_{(u,v)\to(a,b)} h(u,v)=0 \tag{1.2.23}$$

が成り立っている．ここで (1.2.22) を $u=\tilde{x}, v=y$ として用いると，$(x,y)\to(a,b)$ の時には $(\tilde{x},y)\to(a,b)$ でもあるから，

$$f_x(\tilde{x},y)=f_x(a,b)+g(\tilde{x},y), \qquad \text{with} \qquad \lim_{(x,y)\to(a,b)} g(\tilde{x},y)=0 \tag{1.2.24}$$

が結論できる．同様に，

$$f_y(a,\tilde{y})=f_y(a,b)+h(a,\tilde{y}), \qquad \text{with} \qquad \lim_{(x,y)\to(a,b)} h(a,\tilde{y})=0 \tag{1.2.25}$$

も結論できる．これを (1.2.21) に代入すると

$$f(x,y)-f(a,b)=f_x(a,b)\times(x-a)+f_y(a,b)\times(y-b)+g(\tilde{x},y)\times(x-a)+h(a,\tilde{y})\times(y-b) \tag{1.2.26}$$

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，

$$\lim_{(x,y)\to(a,b)} \frac{\tilde{f}(x,y)}{\|\boldsymbol{x}-\boldsymbol{a}\|}=0 \qquad \text{（ここで}\quad \boldsymbol{x}=(x,y),\ \boldsymbol{a}=(a,b) \quad \text{と書いた）} \tag{1.2.27}$$

が結論できる．結果として，このような $\tilde{f}$ を用いて

$$f(x,y)=f(a,b)+A(x-a)+B(y-b)+\tilde{f}(x,y), \qquad \text{with} \qquad A=f_x(a,b),\ B=f_y(a,b) \tag{1.2.28}$$

と書ける事がわかった．これは全微分可能の定義式 (1.2.11) 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. □

1.3 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 1.3.3 節である[9]．

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，

$$h(x)=f(g(x)) \tag{1.3.1}$$

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.「そんな言葉は知らない」という人も $\sin(x^3)$ は $f(x)=\sin x$ と $g(x)=x^3$ の合成関数だといえば，高校の時から知っているものと納得できるはずだ．

このとき，関数 $h(x)$ の導関数については，高校以来，

$$h'(x)=f'\bigl(g(x)\bigr)\,g'(x) \tag{1.3.2}$$

が成り立つことは知っている.（このように書くと良くわからない，という人も $\sin(x^3)$ を x で微分する事はできるはずだから，受験数学でやってるんだよ.）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 $h(x)=f\bigl(z(x)\bigr)$ 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 $h(x,y)=f\bigl(z(x,y)\bigr)$ 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 $h(t)=f\bigl(x(t),y(t)\bigr)$ 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 $h(u,v)=f\bigl(x(u,v),y(u,v)\bigr)$ の，u,v による偏微分.

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

$$\frac{\partial h(x,y)}{\partial x}=f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial x},\qquad \frac{\partial h(x,y)}{\partial y}=f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial y} \tag{1.3.3}$$

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．

1.3.1 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 $g(I)=\{g(x)\,|\,x\in I\}$ が[10] J の部分集合であるとする．このとき，合成関数 $h(x)=f\bigl(g(x)\bigr)$ 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．

---

[9] 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）に書く事が多い．しかしこの節のように，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ので，注意されたい

[10] 細かいけどやっぱり大事な注：関数 g とは区間 I 内の実数 x を実数 $g(x)$ に対応させるもので，g の括弧の中には x（実数）が入るべきである．ところがここでは g の括弧の中に実数の集合（区間）I が入っている．ここが気になったあなたは非常に注意深いのでほめてあげましょう．確かに g は最初，実数を実数に移す規則として定義したのだが，集合 $\{g(x)\,|\,x\in I\}$ を $g(I)$ と書いたのだ．厳密な事をいうと，g に関係したある関数（これは集合を集合に移す）G を $G(I)=\{g(x)\,|\,x\in I\}$ の作用を持つようなもの，と定義すべきである．でもこれは煩雑だから g を二重の意味につかって $g(I)$ と書いたのだ．このような二重の使い方はこれからも頻繁に出るだろう.（この脚注自身が何を言ってるか理解できない人も多いと思う．実は僕も大学入学時点ではこのような注はウザイだけだったから，わからないなら今はそれほど気にしなくても良い —— ２年生くらいまでにははっきりわかってくれないと困るけど.）

定理 1.3.1（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 $h(x)=f\bigl(g(x)\bigr)$ は点 $x=a$ で微分可能であり，

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

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 1.3.3 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 1.3.1 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

$$\frac{h(a+\epsilon)-h(a)}{\epsilon}=\frac{f\bigl(g(a+\epsilon)\bigr)-f\bigl(g(a)\bigr)}{\epsilon}=\frac{f\bigl(g(a+\epsilon)\bigr)-f\bigl(g(a)\bigr)}{g(a+\epsilon)-g(a)}\times\frac{g(a+\epsilon)-g(a)}{\epsilon} \tag{1.3.4}$$

と書いて，$\epsilon\to 0$ としてやれば良い．後ろはモロに $g'(a)$ に行くし，$g(a+\epsilon)\to g(a)$ 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,「$\epsilon\neq 0$ であっても $g(a+\epsilon)-g(a)=0$ かもしれない」という可能性を見て見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 (1.3.4) 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ることが一番大切である．大学の数学なので，以下に「完全な証明」を与えるが，これは興味のある人だけ理解すれば良いのであって，このような細部にはこだわるべきではない．

定理 1.3.1 の正しい証明． 微分係数の定義から $g'(a)$ が存在するということは

$$g'(a)=\lim_{\epsilon\to 0}\frac{g(a+\epsilon)-g(a)}{\epsilon}\qquad\Longleftrightarrow\qquad \lim_{\epsilon\to 0}\frac{g(a+h)-g(a)-g'(a)\epsilon}{\epsilon}=0 \tag{1.3.5}$$

であるので，右側の式の分子を $\epsilon\times\tilde g(\epsilon)$ と書くと，つまり新しい関数 $\tilde g(\epsilon)$ を

$$\tilde g(\epsilon):=\frac{g(a+h)-g(a)-g'(a)\epsilon}{\epsilon} \tag{1.3.6}$$

と定義すると[11]，(1.3.5) は

$$g(a+\epsilon)=g(a)+\epsilon g'(a)+\epsilon\tilde g(\epsilon),\qquad \text{with}\quad \lim_{\epsilon\to 0}\tilde g(\epsilon)=0\ \ \text{かつ}\ \ \tilde g(0)=0 \tag{1.3.7}$$

と書ける．$g(a)=b$ とすると，f に対する同様の考察によって

$$f(b+\eta)=f(b)+\eta f'(b)+\eta\tilde f(\eta),\qquad \text{with}\quad \lim_{\eta\to 0}\tilde f(\eta)=0\ \text{かつ}\ \ \tilde f(0)=0 \tag{1.3.8}$$

である．これらを使って問題の微分係数

$$\lim_{\epsilon\to 0}\frac{h(a+\epsilon)-h(a)}{\epsilon}=\lim_{\epsilon\to 0}\frac{f\bigl(g(a+\epsilon)\bigr)-f\bigl(g(a)\bigr)}{\epsilon} \tag{1.3.9}$$

を考えよう．この分子は (1.3.7) と (1.3.8) から

$$\begin{aligned} f\bigl(g(a+\epsilon)\bigr)-f\bigl(g(a)\bigr)&=f\bigl(g(a)+\epsilon g'(a)+\epsilon\tilde g(\epsilon)\bigr)-f\bigl(g(a)\bigr)=f\bigl(b+\epsilon g'(a)+\epsilon\tilde g(h)\bigr)-f(b)\\ &=f(b)+\eta f'(b)+\eta\tilde f(\eta)-f(b)=\eta f'(b)+\eta\tilde f(\eta)\qquad\text{with}\qquad \eta:=\epsilon g'(a)+\epsilon\tilde g(\epsilon) \end{aligned} \tag{1.3.10}$$

[11] $\tilde g$ と書いたのは単に「関数 g と関係した別の関数」の意味である．$\tilde g(\epsilon)$ の替わりに $p(\epsilon)$ などでも良かったのだが．なお，今は a をとめて考えているので，この関数は ϵ の関数とみなす

と計算できるので，問題の極限は

$$\lim_{\epsilon\to 0}\frac{h(a+\epsilon)-h(a)}{\epsilon}=\lim_{\epsilon\to 0}\frac{\eta f'(b)+\eta\tilde{f}(\eta)}{\epsilon}=\lim_{\epsilon\to 0}\Big[\{g'(a)+\tilde{g}(h)\}\{f'(b)+\tilde{f}(\eta)\}\Big]$$
$$=f'(b)g'(a)+\lim_{\epsilon\to 0}\Big[f'(b)\tilde{g}(\epsilon)+g'(a)\tilde{f}(\eta)+\tilde{g}(\epsilon)\tilde{f}(\eta)\Big] \tag{1.3.11}$$

となる．第一項だけなら我々の欲しい結果になっているから，第２項の極限がゼロであることを示せば証明は完結する．しかし，これは簡単だ．まず，(1.3.7) から $\tilde{g}(\epsilon)$ はゼロに行く．また，(1.3.8) と (1.3.10) での η の定義から

$$\lim_{\epsilon\to 0}\eta=\lim_{\epsilon\to 0}\Big[\epsilon g'(a)+\epsilon\tilde{g}(h)\Big]=0\times g'(a)+0\times 0=0 \tag{1.3.12}$$

であるとわかる[12]．従って (1.3.8) から $\tilde{f}(\eta)$ もゼロに行くことがわかる．というわけで (1.3.11) の第２項はゼロになり，結果として $h'(a)=f'(b)g'(a)=f'(g(a))g'(a)$ を得る． □

1.3.2 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 $g(x,y)$ があって，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 $h(x,y)=f\big(g(x,y)\big)$ を D で定義することができるが...

定理 1.3.2 $g(x,y)$ 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，$h(x,y)=f\big(g(x,y)\big)$ は (a,b) にて x について偏微分可能で，

$$h_x(a,b)=f'\big(g(a,b)\big)\,g_x(a,b),\qquad \text{つまり}\ \ z=g(x,y),\ w=f(z)\ \text{とおくと}\quad \frac{\partial w}{\partial x}=\frac{dw}{dz}\frac{\partial z}{\partial x} \tag{1.3.13}$$

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 1.3.1 が使える． □

1.3.3 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．まずは発見法的に考えてみる．

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[13]］

$$\lim_{\epsilon\to 0}\frac{h(t+\epsilon)-h(t)}{\epsilon}=\lim_{\epsilon\to 0}\frac{f\big(x(t+\epsilon),y(t+\epsilon)\big)-f\big(x(t),y(t)\big)}{\epsilon}=\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_0,y_0)}{\epsilon}$$
$$=\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}+\lim_{\epsilon\to 0}\frac{f(x_1,y_0)-f(x_0,y_0)}{\epsilon} \tag{1.3.14}$$

が出てくる．さて，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 $\epsilon\to 0$ 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**）．<u>もし</u> これが（x_2 は ϵ に無関係な数で）

$$\lim_{\epsilon\to 0}\frac{f(x_2,y_1)-f(x_2,y_0)}{\epsilon}\qquad\text{with}\qquad y_0=y(t),\quad y_1=y(t+\epsilon) \tag{1.3.15}$$

---

[12]実はここで,「 $\lim_{\epsilon\to 0}f(\epsilon)=\lim_{\epsilon\to 0}g(\epsilon)=0$ ならば $\lim_{\epsilon\to 0}f(\epsilon)g(\epsilon)=0$ 」などの事実を用いた．これは証明すべき事であり，５月以降に行う

[13]$x(t+\epsilon)$ 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

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，(1.3.14) から

$$h'(t) = \frac{\partial f}{\partial x}(x_0, y_0)\, x'(t) + \frac{\partial f}{\partial y}(x_0, y_0)\, y'(t) \tag{1.3.16}$$

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 (1.3.16) はかなり一般に成り立つ．ただし，上でも明記したように，(1.3.14) 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：

$$g(0,0) = 0, \qquad (x, y) \neq (0,0) \text{ では } \quad g(x, y) = \frac{xy}{\sqrt{x^2 + y^2}}. \tag{1.3.17}$$

この関数の偏導関数は (1.2.9) 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

$$h(t) = g(x(t), y(t)) = \frac{t}{\sqrt{2}}, \qquad h'(t) = \frac{1}{\sqrt{2}} \tag{1.3.18}$$

であって，特に $h'(0) \neq 0$ だ．ところが (1.3.16) 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 (1.3.16) は適用できない！　（反例終わり）

以下ではこ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 1.3.4 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，関数の定義域と値域は，合成関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 1.3.3 2 変数関数 $f(x, y)$ が C^1-級 で，$x(t), y(t)$ が t について微分可能なら，$h(t) = f\bigl(x(t), y(t)\bigr)$ は t で微分可能である．更にその導関数について

$$h'(t) = f_x\, x'(t) + f_y\, y'(t), \qquad \text{つまり } z = f(x, y) \text{ とおくと } \quad \frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \tag{1.3.19}$$

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 変数の関数の場合は：

$$\frac{d}{dt} f\bigl(x_1(t), x_2(t), \ldots, x_n(t)\bigr) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{dx_j}{dt}. \tag{1.3.20}$$

言うまでもなく，(1.3.17) 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　(1.3.14) 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 だけの関数と見て 1 変数関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で，平均値の定理は使える）

$$f(x_1, y_1) - f(x_1, y_0) = f_y(x_1, y_3) \times (y_1 - y_0) \tag{1.3.21}$$

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 $\epsilon \to 0$ とすると，

$$\lim_{\epsilon \to 0} \frac{f(x_1, y_1) - f(x_1, y_0)}{\epsilon} = \lim_{\epsilon \to 0} \left[f_y(x_1, y_3) \times \frac{y_1 - y_0}{\epsilon} \right] \tag{1.3.22}$$

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 の連続関数 である（ここがキーでした）．従って $\epsilon \to 0$ では $x_1 \to x_0$ かつ $y_3 \to y_0$ であることも使うと，

$$\lim_{\epsilon \to 0} f_y(x_1, y_3) = f_y(x_0, y_0) \tag{1.3.23}$$

である．また，(1.3.22) の後ろの方は y_1, y_0 の定義から

$$\lim_{\epsilon\to 0}\frac{y_1-y_0}{\epsilon}=\lim_{\epsilon\to 0}\frac{y(t+\epsilon)-y(t)}{\epsilon}=y'(t) \tag{1.3.24}$$

である．以上をまとめると

$$\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}=f_y(x_0,y_0)\times y'(t) \tag{1.3.25}$$

が厳密に証明されたので，(1.3.14) 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 1.3.4 (２変数関数の連鎖律) 2 変数関数 $f(x,y)$ が C^1-級で，$x(u,v), y(u,v)$ が u,v について C^1-級なら，合成関数 $\mathrm{z}=h(u,v)=f\bigl(x(u,v),y(u,v)\bigr)$ も C^1-級で，

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u},\qquad \text{かつ}\qquad \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v}. \tag{1.3.26}$$

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 1.3.3 からすぐに証明される． □

問題 1.3.5 （連鎖率）2 変数の関数 $f(x,y)=\dfrac{xy}{\sqrt{x^2+y^2}}$ 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，合成関数 $h(r,\theta)=f\bigl(x(r,\theta),y(r,\theta)\bigr)$ の，変数 r,θ に関する偏導関数を計算せよ．ただしその場合,（1）合成関数の微分法（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 1.3.6 (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）

$$u=\alpha x+\beta y,\quad v=\gamma x+\delta y \tag{1.3.27}$$

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

問題 1.3.7 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）．

$$1)\ \ \frac{\partial f}{\partial y}\equiv 0\qquad 2)\ \ \frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\equiv 0\qquad 3)\ \ a\frac{\partial f}{\partial x}+b\frac{\partial f}{\partial y}\equiv 0 \tag{1.3.28}$$

1) は既にやった（1.2.1 節）．問題は 2) と 3) だが，問題 1.3.6 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，$\dfrac{\partial f}{\partial u}\equiv 0$ 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.）

1.3.4 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 1.3.8 (定理 1.3.3 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，$h(t)=f\bigl(x(t),y(t)\bigr)$ は t で微分可能で，

$$h'(t)=f_x\,x'(t)+f_y\,y'(t),\qquad \text{つまり}\ \ z=f(x,y)\ \text{とおくと}\quad \frac{dz}{dt}=\frac{\partial z}{\partial x}\frac{dx}{dt}+\frac{\partial z}{\partial y}\frac{dy}{dt} \tag{1.3.29}$$

が成り立つ．n-変数の場合も同様であるが，略．

定理 1.3.9 (定理 1.3.4 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 $z = h(u,v) = f\bigl(x(u,v), y(u,v)\bigr)$ も (u,v) について全微分可能で，

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \qquad \text{かつ} \qquad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v} \tag{1.3.30}$$

である．

定理 1.3.4 が定理 1.3.3 から出るのと同じようにして，定理 1.3.9 は定理 1.3.8 から証明できる．従って，以下では定理 1.3.8 を簡単に証明する．

定理 1.3.8 の証明

この証明では後でやる「オーダー」の概念を使って説明した．「オーダー」を使わないと式が煩雑になるし，肝心のところが良くわからないだろうと思うからである．この証明は後で「オーダー」を理解してから読んでくれれば良く，現時点での理解は求めていない．

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 1.2.5 から，

$$h(t+\epsilon) - h(t) = f(x_1,y_1) - f(x_0,y_0) = A(x_1 - x_0) + B(y_1 - y_0) + o\Bigl(\sqrt{|x_1 - x_0|^2 + |y_1 - y_0|^2}\Bigr) \tag{1.3.31}$$

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），

$$x_1 - x_0 = x'(t)\epsilon + o(\epsilon) = p\epsilon + o(\epsilon), \qquad y_1 - y_0 = y'(t)\epsilon + o(\epsilon) = q\epsilon + o(\epsilon) \tag{1.3.32}$$

も成り立っている．従って，

$$\begin{aligned} h(t+\epsilon) &= h(t) + A\{p\epsilon + o(\epsilon)\} + B\{q\epsilon + o(\epsilon)\} + o\Bigl(\sqrt{\{p\epsilon + o(\epsilon)\}^2 + \{q\epsilon + o(\epsilon)\}^2}\Bigr) \\ &= h(t) + (Ap + Bq)\epsilon + o(\epsilon) + o\Bigl(\sqrt{(p^2+q^2)\epsilon^2 + o(\epsilon^2)}\Bigr) \end{aligned} \tag{1.3.33}$$

となるが，

$$o\Bigl(\sqrt{(p^2+q^2)\epsilon^2 + o(\epsilon^2)}\Bigr) = o\Bigl(\sqrt{p^2+q^2}\,\epsilon + o(\epsilon)\Bigr) = o(\epsilon) \tag{1.3.34}$$

であるので，結局 (1.3.33) は

$$h(t+\epsilon) = h(t) + (Ap + Bq)\epsilon + o(\epsilon) \tag{1.3.35}$$

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 (1.2.10) そのものだ． □

1.4 高階の偏導関数

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

2変数以上の関数についても，同様のものを考えたい．すなわち，$f(x,y)$ の1階導関数 $\frac{\partial f}{\partial x}(x,y)$ は x, y の関数であるが，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階導関数 $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial x}\right)$ を定義する．同様に，$\frac{\partial f}{\partial x}$ を y で微分して

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \qquad \text{同様に} \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{1.4.1}$$

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[14]．

[14]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

問題 1.4.1 以下の関数 f, g, h について，2階の偏導関数（4通り）をすべて計算せよ．実は以下の定理で $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ だというのだが，この定理には頼らず，実際に計算して確かめる事．

$$f(x,y) = x^2 + y^2, \qquad g(x,y) = x^2 y^3, \qquad h(x,y) = \cos(x^2 y) \tag{1.4.2}$$

k-階の導関数も同様に定義する．つまり，$f(x,y)$ の k-階の導関数とは $f(x,y)$ を x または y で合計 k 回，微分してできる関数のことである．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-階までの偏導関数がすべて存在してかつ連続のとき，その関数は 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 1.4.1 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」関数では）この2つは等しい．

定理 1.4.2 2変数関数 $f(x,y)$ が C^2-級の場合，

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \tag{1.4.3}$$

である．つまり，偏導関数は偏微分の順序によらない．n-変数の関数の場合も同様で，関数が C^2-級ならば偏導関数は偏微分の順序によらない．さらに，n-変数の関数が C^k-級ならば，その k-階までの偏導関数は偏微分の順序によらない．

定理 1.4.2 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[15]（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 1.4.2 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．

$$\Delta(h,k) := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a,b) \tag{1.4.4}$$

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，

$$\begin{aligned}\lim_{h\to 0}\Bigl(\lim_{k\to 0}\frac{\Delta(h,k)}{hk}\Bigr) &= \lim_{h\to 0}\frac{1}{h}\Bigl[\lim_{k\to 0}\Bigl(\frac{f(a+h,b+k)-f(a+h,b)}{k} - \frac{f(a,b+k)-f(a,b)}{k}\Bigr)\Bigr] \\ &= \lim_{h\to 0}\frac{1}{h}\Bigl[\frac{\partial f}{\partial y}(a+h,b) - \frac{\partial f}{\partial y}(a,b)\Bigr] = \frac{\partial}{\partial x}\Bigl(\frac{\partial f}{\partial y}\Bigr) = \frac{\partial^2 f}{\partial x \partial y}\end{aligned} \tag{1.4.5}$$

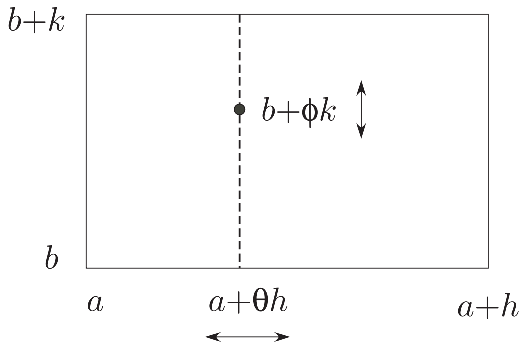
であるし，同様に考えると

$$\lim_{k\to 0}\Bigl(\lim_{h\to 0}\frac{\Delta(h,k)}{hk}\Bigr) = \frac{\partial^2 f}{\partial y \partial x} \tag{1.4.6}$$

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がポイントになる．

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）．

[15] 小平本の 6.2 節の d) を参考にした



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変数の平均値の定理から，

$$\varphi(a+h) - \varphi(a) = \varphi'(a+\theta h)h \tag{1.4.7}$$

つまり

$$\Delta(h,k) = \varphi(a+h) - \varphi(a) = \Big(f_x(a+\theta h, b+k) - f_x(a+\theta h, b)\Big)h \tag{1.4.8}$$

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[16]またもや１変数の場合の平均値の定理から

$$\psi(b+k) - \psi(b) = \psi'(b+\phi k)k \tag{1.4.9}$$

つまり

$$f_x(a+\theta h, b+k) - f_x(a+\theta h, b) = f_{xy}(a+\theta h, b+\phi k)\, k \tag{1.4.10}$$

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 (1.4.7) と (1.4.10) から

$$\frac{\Delta(h,k)}{hk} = \frac{f_{xy}(a+\theta h, b+\phi k)\, hk}{hk} = f_{xy}(a+\theta h, b+\phi k) \tag{1.4.11}$$

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：

$$\lim_{(h,k)\to(0,0)} \frac{\Delta(h,k)}{hk} = f_{xy}(a,b) \tag{1.4.12}$$

これで既に $h \to 0$ と $k \to 0$ の極限が順序によらない事は示せたから (1.4.5) と (1.4.6) 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うと (1.4.12) 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.） □

1.4.1　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

[16]（注）ここでは少しひっかかるかもしれない．(1.4.7) 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 (1.4.10) は (1.4.7) がなりたっていた $b + k, b$ についてのものであるので，(1.4.7) と (1.4.10) 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

$z = f(x, y)$ 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，ううむ，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

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{1.4.13}$$

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

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial f}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial f}{\partial v} = \frac{1}{\sqrt{2}}\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right] \tag{1.4.14}$$

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

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial}{\partial v} = \frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial u} - \frac{\partial}{\partial v}\right] \tag{1.4.15}$$

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

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &= \frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right]\left\{\frac{1}{\sqrt{2}}\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right]\right\} = \frac{1}{2}\left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right]\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right] \\ &= \frac{1}{2}\left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v \partial u} - \frac{\partial^2 f}{\partial v^2}\right] = \frac{1}{2}\left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial v^2}\right] \end{aligned} \tag{1.4.16}$$

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．

1.4.2 高階偏導関数と連鎖律

上で，高階偏導関数の出てく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

$$\frac{\partial f}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial f}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial f}{\partial \theta} = \frac{x}{\sqrt{x^2+y^2}}\frac{\partial f}{\partial r} - \frac{y}{x^2+y^2}\frac{\partial f}{\partial \theta} = \frac{x}{r}\frac{\partial f}{\partial r} - \frac{y}{r^2}\frac{\partial f}{\partial \theta} \tag{1.4.17}$$

となるので，微分演算子としては

$$\frac{\partial}{\partial x} = \frac{x}{r}\frac{\partial}{\partial r} - \frac{y}{r^2}\frac{\partial}{\partial \theta} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta} \tag{1.4.18}$$

という作用をもっている．さて，もう一回やるときには

$$\frac{\partial^2 f}{\partial x^2} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) \tag{1.4.19}$$

となるのだが，左側の括弧の中の微分はその <u>右側</u> 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

$$= \cos\theta\frac{\partial}{\partial r}\left(\cos\theta\frac{\partial f}{\partial r}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\cos\theta\frac{\partial f}{\partial r}\right) + \cos\theta\frac{\partial}{\partial r}\left(-\frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(-\frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) \tag{1.4.20}$$

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り，例えば，

$$\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Big(\cos\theta\frac{\partial f}{\partial r}\Big) = \frac{\sin\theta}{r}\frac{\partial\cos\theta}{\partial\theta}\times\frac{\partial f}{\partial r} + \frac{\sin\theta}{r}\cos\theta\frac{\partial}{\partial r}\Big(\frac{\partial f}{\partial\theta}\Big) = -\frac{\sin^2\theta}{r}\frac{\partial f}{\partial r} + \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{1.4.21}$$

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

$$\frac{\partial^2 f}{\partial x^2} = \cos^2\theta\frac{\partial^2 f}{\partial r^2} + \frac{\sin^2\theta}{r}\frac{\partial f}{\partial r} + \frac{\sin 2\theta}{r^2}\frac{\partial f}{\partial\theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 f}{\partial\theta^2} - \frac{\sin 2\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{1.4.22}$$

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 1.4.3 2 変数の関数 $f(x,y)$ と座標変換 $x = r\cos\theta, y = r\sin\theta$ を考える．このとき，

$$\Delta f(x,y) := \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{1.4.23}$$

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．

1.4.3 （補足）偏導関数がゼロという関数は？ふたたび

以前に「1 階偏微分がゼロ」の関数はどんなものか考えたが，今度は 2 階導関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

$$\frac{\partial}{\partial y}\big(f_x\big) = 0 \tag{1.4.24}$$

と見ると,「f_x は y には依存しません」ということなので，

$$f_x(x,y) = g(x) \tag{1.4.25}$$

と書けるはずだ（g は任意の関数，1.2.1 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 の原始関数になるが，x で積分したときの積分定数は y の任意の関数になれる．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.（ここのところがわかりにくい人は $g(x)$ の原始関数を $G(x)$ とすると，(1.4.25) から

$$\frac{\partial}{\partial x}\big\{f(x,y) - G(x)\big\} = 0 \tag{1.4.26}$$

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

$$f(x,y) = G(x) + h(y) \tag{1.4.27}$$

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 の任意関数）．結局，f は x と y の関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，このような考察は物理で「波動方程式」を考える時にでてくるのだが，皆さんは接する機会はないかな...

2 極限の厳密な定義

お待たせしました！これから今学期の終わりまで，極限を厳密に定義することから始めて，微分，関数の連続性などについて学んでいきます．ここからの講義には，原則として「ゴマカシ」はありません.（時間の都合で証明を省いたりする事はありますが.）

（お断り）極限を扱う場合には,「実数の連続性」を学び（定義し），それから極限の定義におもむろに入っていくのが一般的です.「実数とは何か」がはっきりしない限り，実数の極限なども議論しにくいからです．しかし,「実数の連続性」はかなり抽象的で，それがどこで極限と結びつくのかがそれほど明らかではありませんし，かなりとっつきにくい題材でもあります．また，皆さんの鬼門であろう「ϵ-δ 論法」と実数の連続性には直接の関係はありません.

そこで，この講義では「実数の連続性」は後回し（多分，６月上旬，3 章）にし，皆さんが日頃持っている実数の感覚に基づいて，ともかく「ϵ-δ 論法」をわかることを目指します．このようにしても，論理の飛躍はありません.

2.1 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 $\lim_{n\to\infty} a_n = \alpha$ 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はいないし，これで十分な場合はこれでやれば良い．しかし，この言い方はやはり困ったものである.

まず,「限りなく近づく」「限りなく大きくなる」の言い方には「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ていないけど困ってしまう例としては,

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

を考えてみると良い．この答えは直感的には 0 だろうという気はするだろうが，証明できますか？(この答えは後の命題 2.1.7 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 論法というもので，以下のようなものである.

定義 2.1.1 数列 a_n と実数 α に対して，数列 a_n が $n\to\infty$ で α に収束する，つまり $\lim_{n\to\infty} a_n = \alpha$ 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,

$$\text{すべての } n > N(\epsilon) \text{ で，} \quad |a_n - \alpha| < \epsilon \quad \text{とできる．} \tag{2.1.1}$$

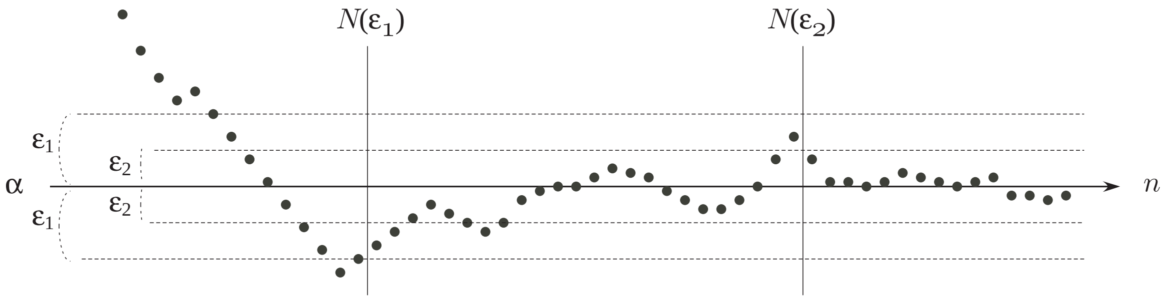
（ア）を数式で書くと

$$\forall\epsilon > 0 \quad \exists N(\epsilon) \qquad n > N(\epsilon) \implies |a_n - \alpha| < \epsilon \tag{2.1.2}$$

となる.

[17] この節の内容は教科書の 1.1 節と 1.2 節に対応しているから，良く読んで復習すべし.

[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．詳しくは秋学期にガンガンやります



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．
- (2.1.2) には 2 つの不等式 $n > N(\epsilon)$, $\left|a_n - \alpha\right| < \epsilon$ が現れている．ここはどちらも（または片方を）$n \geq N(\epsilon)$ や $\left|a_n - \alpha\right| \leq \epsilon$（等号入り）に変えても，定義の意味する事は同じである（ここは重要だから，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，証明の流れによっては等号入りのものを断りなく使うこともあるので，注意されたい．
- 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，後で充分に慣れてから，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 をこちらで勝手に選べる 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（以下の小節で詳しく説明する）.

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

定義 2.1.2 数列 a_n に対して，数列 a_n の $n \to \infty$ の極限がプラス無限大である，つまり $\lim_{n\to\infty} a_n = +\infty$ 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,

$$\text{すべての } n > N(M) \text{ で，}\quad a_n > M \quad \text{とできる．} \tag{2.1.3}$$

（ア′）を数式で書くと

$$\forall M > 0 \quad \exists N(M) \qquad n > N(M) \implies a_n > M \tag{2.1.4}$$

となる．
同様に，$\lim_{n\to\infty} a_n = -\infty$ というのは，以下が成り立つことと定義する：

$$\forall M < 0 \quad \exists N(M) \qquad n > N(M) \implies a_n < M \tag{2.1.5}$$

（注）$\lim_{n\to\infty} a_n = +\infty$ や $\lim_{n\to\infty} a_n = -\infty$ の場合は $\{a_n\}$ 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.

2.1.1 少しでも理解を助けるために

上の定義 2.1.1 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ても，それに1を足したらもっと大きくなるから．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,「どんなに大きな N をとってきても，そのうちに n が N より大きくなる」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，この N は好きなように大きなものを持ってこれる．$N = 10^4$ ならどうだ？ $N = 10^{10}$ ならどうだ？ $N = 10^{100}$ なら？ $\ldots$ いくらでも大きな N を許すことで実質的に「n がいくらでも大きくなる」ことを表現していることを噛み締めよう．

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（$n \to \infty$ 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．ここも「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）$|a_n - \alpha|$ 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 はこちらで勝手にとれるのだ．$\epsilon = 10^{-6}$ より小さいか？ $\epsilon = 10^{-14}$ よりも小さいか? $\epsilon = 10^{-200}$ なら? $\ldots$ 「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 を許すことで，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 に近づくかどうかは，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どうか判定するために，

「$\epsilon = 0.0001$ 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」

$\cdots$

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，$\lim_{n\to\infty} a_n = \alpha$ と言いましょう，というわけ．
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アカン！」
となってしまったら，$\lim_{n\to\infty} a_n = \alpha$ 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 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 $N(\epsilon)$ があって」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（$n = 1, 2, 3, \ldots$）：

$$a_n = \frac{1}{n}, \qquad b_n = \frac{1}{\log(2 + \log(2 + \log n))}, \qquad c_n = \frac{1}{\log(2 + \log(2 + \log n))} + 10^{-8} \tag{2.1.6}$$

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：

[19]学習院大学物理学教室の田崎晴明氏の用語

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 n をとってくる必要があるわけだ．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，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，ϵ を先に決めて，これに応じて n がどのくらい大きい必要があるかを（またはいくら大きい n でも $|a_n-\alpha|$ が ϵ より小さくなれないのかを）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．

2.1.2 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 2.1.7 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ってみること．

問題 2.1.3 以下の数列が $n\to\infty$ 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，$n=1,2,3,\ldots$ である）．

$$a_n=3,\qquad b_n=\frac{1}{n},\qquad c_n=\frac{1}{\sqrt{n}},\qquad d_n=\frac{1}{n^2+1} \tag{2.1.7}$$

$$e_n=\begin{cases}1 & （n \text{ が } 10,10^2,10^3,10^4,10^5,10^6,\ldots \text{ のとき}）\\ 0 & （\text{上以外のとき}）\end{cases} \tag{2.1.8}$$

(2.1.6) 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（$n\to\infty$）：

$$f_n=\frac{n+3}{n},\qquad g_n=\frac{\sin n}{n},\qquad h_n=\sqrt{n+1}-\sqrt{n},\qquad p_n=\frac{2n+1}{n+1},\qquad q_n=\frac{1}{\log(n+1)} \tag{2.1.9}$$

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 2.1.4 極限に関する以下の性質を証明せよ．

- $\lim_{n\to\infty}a_n=\alpha$, $\lim_{n\to\infty}b_n=\beta$ のとき，$\lim_{n\to\infty}(a_n+b_n)=\alpha+\beta$.
- $\lim_{n\to\infty}a_n=\alpha$, $\lim_{n\to\infty}b_n=\beta$ のとき，$\lim_{n\to\infty}a_nb_n=\alpha\beta$.
- $\lim_{n\to\infty}a_n=\alpha$, $\lim_{n\to\infty}b_n=\beta$ （$\beta\neq0$）のとき，$\lim_{n\to\infty}\dfrac{a_n}{b_n}=\dfrac{\alpha}{\beta}$ ．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 2.1.5 （論理に弱い人にはキツいだろうが，頑張ろう）数列 $a_n=1+\dfrac{1}{n}$ 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 2.1.6 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 2.1.6 （気がつけば簡単だが...）数列 a_n が $n \to \infty$ 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，$\lim_{n\to\infty} a_n = \alpha$ かつ $\lim_{n\to\infty} a_n = \beta$ であるとすると，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 2.1.7 数列 a_n から $b_n = \dfrac{1}{n}\sum_{k=1}^{n} a_k$ を定義する．$\lim_{n\to\infty} a_n = \alpha$ ならば，$\lim_{n\to\infty} b_n = \alpha$ 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（各自やってみること）.

問題 2.1.8 （数列に関するチャレンジ問題）命題 2.1.7 は

$$\lim_{n\to\infty} a_n = \alpha \qquad \Longrightarrow \qquad \lim_{n\to\infty} \frac{a_1 + a_2 + \cdots + a_n}{n} = \alpha$$

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，$\rho_1, \rho_2, \rho_3, \ldots$ を非負の数列として，

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

を考える.「$\lim_{n\to\infty} a_n = \alpha$ ならば必ず $\lim_{n\to\infty} b_n = \alpha$ となる」ためには，$\rho_1, \rho_2, \rho_3, \ldots$ 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を考えてみよう.（命題 2.1.7 は $\rho_1 = \rho_2 = \rho_3 = \ldots = 1$ に相当している.）

2.2 関数の極限：ϵ-δ 論法[20]

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，少し簡単に書く.

定義 2.2.1 関数 $f(x)$ と実数 a, b に対して,「$f(x)$ が $x \to a$ で b に収束する，つまり $\lim_{x\to a} f(x) = b$」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，

$$0 < |x - a| < \delta(\epsilon) \text{ なるすべての } x \text{ で，} \quad \bigl|f(x) - b\bigr| < \epsilon \quad \text{とできる．} \tag{2.2.1}$$

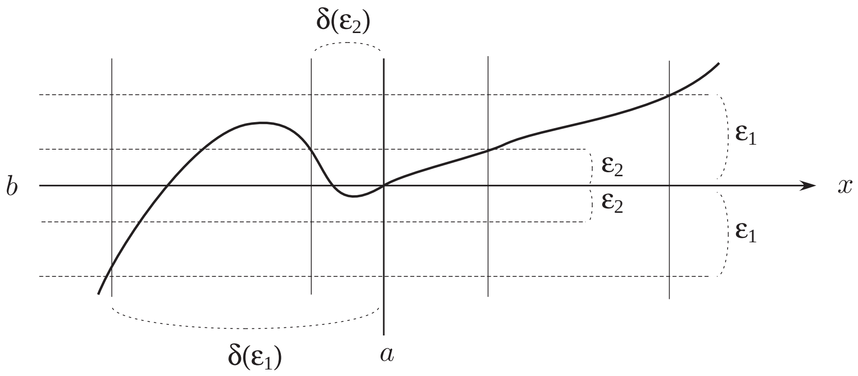
（イ）は数式では

$$\forall\epsilon > 0 \quad \exists\delta(\epsilon) > 0 \qquad 0 < |x - a| < \delta(\epsilon) \implies \bigl|f(x) - b\bigr| < \epsilon \tag{2.2.2}$$

となる.

（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，すぐ後で「関数の連続性」を考えると理解できるだろう.

[20] この節の内容は教科書の 1.1 節と 1.3 節



注意： ϵ-N の時と同じく，上の 2 つの不等式 $0 < |x-a| < \delta(\epsilon)$，$\bigl|f(x)-b\bigr| < \epsilon$ は，等号入りの $0 < |x-a| \le \delta(\epsilon)$，$\bigl|f(x)-b\bigr| \le \epsilon$ 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，ともに <u>正で有限</u> 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 が後になってる理由も ϵ-N 論法と同じだ．考えている関数によっては 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（どのくらい小さいのか)」を考える方が効率が良い．

ここも，いろいろな例をやることで感覚を身につけよう．

問題 2.2.2 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よ）．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.

$$1)\ \lim_{x\to 0} x, \qquad 2)\ \lim_{x\to 0}\Bigl(x^2-2x+3\Bigr), \qquad 3)\ \lim_{x\to 1}\Bigl(x^2-2x+3\Bigr). \tag{2.2.3}$$

もうちょっとひねった例（$a>0$ は定数)：

$$4)\ \lim_{x\to 0}\frac{1}{1+x}, \qquad 5)\ \lim_{x\to 1}\frac{x^2-1}{x-1}, \qquad 6)\ \lim_{x\to 0}\sin\frac{1}{x}, \tag{2.2.4}$$

$$7)\ \lim_{x\to a}\frac{x^3-a^3}{x-a} \qquad 8)\ \lim_{x\to 0}\frac{\sqrt{1+x}-\sqrt{1-x}}{x} \qquad 9)\ \lim_{x\to 0}\sqrt{|x|} \tag{2.2.5}$$

問題 2.2.3 $f(x)$ を以下のように定めるとき，極限 $\lim_{x\to 0} f(x)$ 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 2.2.4 $\lim_{x\to a} f(x) = \alpha$ かつ $\lim_{x\to a} g(x) = \beta$ の時，$\lim_{x\to a}\bigl\{f(x)+g(x)\bigr\} = \alpha+\beta$ と $\lim_{x\to a}\bigl\{f(x)g(x)\bigr\} = \alpha\beta$ が成り立つ．これらを ϵ-δ 論法によって証明せよ．

2.2.1 いろいろな極限の定義と例

以上で極限の定義に関する基本はおしまい．ただし，これからは $\lim_{x\to a} f(x) = b$ 以外のいろいろな極限の定義もでてくるから，ここでまとめて述べておく．まず，無限大が絡んでくるものについては，

- 数列の時と同じく，$\lim_{x\to\infty} f(x) = b$ というのは,「どんな（小さな）$\epsilon > 0$ に対しても，うまく（大きな）$L(\epsilon)$ をとってやると，$x > L(\epsilon)$ なるすべての x にて $|f(x) - b| < \epsilon$ が成り立つ」ということだ．
- $\lim_{x\to-\infty} f(x) = b$ というのは,「どんな（小さな）$\epsilon > 0$ に対しても，うまく（大きな）$L(\epsilon)$ をとってやると，$x < -L(\epsilon)$ なるすべての x にて $|f(x) - b| < \epsilon$ が成り立つ」ということだ．
- $\lim_{x\to a} f(x) = +\infty$ というのは,「どんな (大きな) $M > 0$ に対しても，うまく $\delta(M)$ をとってやると，$|x-a| < \delta(M)$ なるすべての x にて $f(x) > M$ が成り立つ」ということだ．
- $\lim_{x\to a} f(x) = -\infty$ というのは,「どんな (大きな) $M > 0$ に対しても，うまく $\delta(M)$ をとってやると，$|x-a| < \delta(M)$ なるすべての x にて $f(x) < -M$ が成り立つ」ということだ．
- これらを組み合わせると，$\lim_{x\to\infty} f(x) = \infty$ や $\lim_{x\to-\infty} f(x) = \infty$ などの定義も書けるが，明らかだろうから詳細は略する．

> 以上，ゴチャゴチャといろいろな定義をしたけども，混乱しそうな人は (初めのうちは) 最も基本的な $\lim_{n\to\infty} a_n = \alpha$ と $\lim_{x\to a} f(x) = b$ のみに限って理解すればよい．この二つがわかってれば何とかなる．

ここで更にもう一つ，片側からの極限も定義しておく．いうまでもなく，単に $\lim_{x\to a}$ と書いた時は a の正負両側からの極限（今までやってきたもの）のことである．

> **定義 2.2.5** 関数 $f(x)$ と実数 a, b に対して
>
> - $\lim_{x\to a+0} f(x) = b$ というのは，a の右側から，つまり $x > a$ を保って，x を a に近づける，ということだ．具体的に書くと,「どんな（小さな）$\epsilon > 0$ に対しても，うまく $\delta(\epsilon) > 0$ をとってやると，$0 < x - a < \delta(\epsilon)$ なるすべての x にて $|f(x) - b| < \epsilon$ が成り立つ」ということで，数式では
>
> $$\forall\epsilon > 0\ \exists\delta(\epsilon) > 0 \qquad 0 < x - a < \delta(\epsilon) \quad\Longrightarrow\quad |f(x) - b| < \epsilon \tag{2.2.6}$$
>
> - 同様に，$\lim_{x\to a-0} f(x) = b$ というのは a の左側から，つまり $x < a$ を保って，x を a に近づける，ということだ．具体的に書くと，
>
> $$\forall\epsilon > 0\ \exists\delta(\epsilon) > 0 \qquad -\delta(\epsilon) < x - a < 0 \quad\Longrightarrow\quad |f(x) - b| < \epsilon \tag{2.2.7}$$

なお，欧米では $\lim_{x\to a+0}$ の代わりに $\lim_{x\to a^+}$ と書き，$\lim_{x\to a-0}$ の代わりに $\lim_{x\to a^-}$ と書くこともある.（$\pm$ 記号は a の右肩についている.）

問題 2.2.6 関数 $f(x)$ と実数 a, b について，通常の極限の存在 $\lim_{x\to a} f(x) = b$ と，左右の片側極限が存在して一致する事 $\lim_{x\to a+0} f(x) = \lim_{x\to a+0} f(x) = b$ が同値である事を証明せよ.（これがアタリマエやろって思った人は多いでしょうが，ちょっとだけ微妙なことはあるのよ．その点がわかってアタリマエと思ってるのなら良いんですがね.）

問題 2.2.7 以下の式を，極限の定義に戻って証明せよ．

$$\lim_{x\to\infty} e^x = \infty, \qquad \lim_{x\to-\infty} e^x = 0, \qquad \lim_{x\to\infty} \log x = \infty, \tag{2.2.8}$$

$$\lim_{x\to+0} \log x = -\infty, \qquad \lim_{x\to 0} \frac{e^x - 1}{x} = 1, \qquad \lim_{x\to 0} \frac{\sin x}{x} = 1, \tag{2.2.9}$$

$$\lim_{x\to\infty} x^{\alpha} = \infty, \qquad \lim_{x\to+0} x^{\alpha} = 0, \qquad (\alpha > 0) \tag{2.2.10}$$

$$\lim_{x\to\infty} x^{\alpha} = 0, \qquad \lim_{x\to+0} x^{\alpha} = \infty, \qquad (\alpha < 0) \tag{2.2.11}$$

上では $x^{\alpha} = e^{\alpha \log x}$ と定義したつもりである．

実を言うと，指数関数 e^x や対数関数 $\log x$ を任意の実数 x に対してどう定義するか，は自明ではない（高校の数学ではここを少しごまかしている）けども，いまのところは感覚的にとらえておれば良いとしよう．これらの関数の厳密な定義も今学期の後半（か来学期）で行う．

2.3 数列の極限と関数の極限の関係[21]

ここまでで，数列の極限，関数の極限をそれぞれ定義した．これらの間の関係を考えるべく，以下の 2 つの命題を考えたい：

（あ）$\displaystyle\lim_{x\to a} f(x) = b$ である．

（い）$\displaystyle\lim_{n\to\infty} a_n = a$ （でもすべての n について $a_n \neq a$）となるすべての数列 $\{a_n\}$ に対して $\displaystyle\lim_{n\to\infty} f(a_n) = b$ である．

両者にはどんな関係があるのだろうか？（あ）ならば（い）であることはすぐにわかる[22]．問題はその逆だ．（い）から（あ）が言えるだろうか？答えは「言える」であって，まとめると以下の定理になる：

定理 2.3.1 上の命題（あ）と（い）は同値である．

関数の極限（$0 < |x-a| < \delta$ なるすべての実数）は考えにくい事がままあるが，数列の収束なら n が一つずつ増えて行くのだからそんなに大変ではない．その意味で，この定理は関数の収束を数列の収束の問題に置き換える事を可能にする，非常に重要なものである．

（注）上の命題の（い）は $a_n \to a$ となる すべての 数列に関する命題である事には注意を要する（特定の数列に対してのみではダメなことを例を作って納得せよ）．

定理 2.3.1 の証明 （あ）から（い）を出すのは簡単だから，各自の演習に任せる．（い）から（あ）を出すには対偶をとって[23]考えるのがよいだろう．まず（あ）と（い）の否定命題をそれぞれ作ってみると

$\overline{（あ）}$ $\exists\epsilon > 0\ \forall\delta > 0\ \exists x (0 < |x-a| < \delta)\ |f(x) - b| \geq \epsilon$.

$\overline{（い）}$ $\displaystyle\lim_{n\to\infty} a_n = a$ $(\forall n\ a_n \neq a)$ となる数列 $\{a_n\}$ で，更に「$\displaystyle\lim_{n\to\infty} f(a_n) = b$ ではない」ものが存在する．

となる．（新居さんの数学入門でやったはずだけど，アヤシイ人が多いのではないかな？ちゃんと復習しておくように．）

$\overline{（あ）}$ から $\overline{（い）}$ を導くためには，$\overline{（い）}$ にでてくる数列 $\{a_n\}$ を具体的に作ってやれば良い．そのために，$\epsilon > 0$ をまず固定し，次に $\delta = 1/n$（n は正の整数）ととってみる．この δ に対して $|f(x) - b| \geq \epsilon$ となる x（$0 < |x-a| < \delta = 1/n$ を満たす）が存在することは，$\overline{（あ）}$ が保証する．そこで，$\delta = 1/n$ に対する上のような x を a_n と定義する（$n = 1, 2, 3, \ldots$）．

この数列 $\{a_n\}$ に対して $\overline{（い）}$ が成り立っている事は，すぐにわかる．実際，上の作り方から $0 < |a_n - a| < 1/n$ なので $\displaystyle\lim_{n\to\infty} a_n = \alpha$ かつ $a_n \neq a$ である．また，すべての a_n に対して $|f(a_n) - b| \geq \epsilon$ であるから，この数列に対しては「$\displaystyle\lim_{n\to\infty} f(a_n) = b$」ではありえない． □

[21] 教科書の 1.4 節．この内容は少し「余談」に近いが，後で非常に重要になってくる．また，厳密な思考に慣れる良い練習問題でもある

[22] 「すぐにわかる」と書いたが，きちんと証明できる人は案外すくないのではないか？考えておくように

[23] 「対偶をとって証明」というのは，「（い）ならば（あ）」が「$\overline{（あ）}$ ならば $\overline{（い）}$」と同値であることを用いる証明法である．念のため

3 実数の連続性と極限

今までで，数列と関数の極限を厳密に定義した．また，その厳密な定義の効用（例えば命題 2.1.7）も紹介した．しかし，実はここまでの話は，考えている数の集合が実数でなくても成り立つものなのだ（少なくとも考えている数をすべて有理数だとしても，2 章の内容はすべて正しい）．

では実数を考える意味はどこにあるのだろう？それは「数を有理数に限っていると，存在してほしい極限が存在しない例が山ほどでてくる」ことにある[24]．それのみならず，実数をやらないままだと，高校でもやったはずの中間値の定理，連続関数の最大値・最小値の定理，平均値の定理，など軒並みアウトで，受験数学のかなりの部分の基礎がなくなってしまうのだ！

というわけで，実数の基礎をちゃんとやるのは本当は非常に大事だ（少なくとも数学科では）．決して衒学趣味でやってるのではない．しかし，今の時点で実数の構成に深入りして講義ノートをいたずらに長くするのは得策とは思われないので（長くなればなるほど読まない人が増えそう），重要な実数の性質は「公理」として天下りに与える事にする．実際には有理数を基にして実数を構成し（定義し），以下の公理を証明する事も可能だが，これは今は触れない[25]．

3.1 上限と下限[26]

まず，上界（下界）と上限（下限）の概念から話を始める．

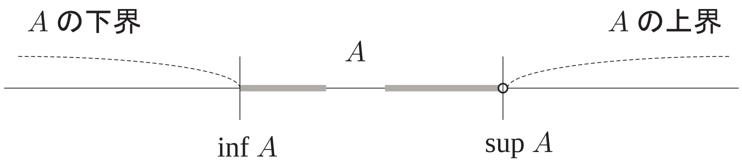
定義 3.1.1 (上界と下界) A を実数の集合とする．ある数 N があって，A の任意の元 a が $a \le N$ を満たすとき，A は上に有界（bounded from above）といい，N を A の 上界（upper bound）という．同様に，ある数 M があって，A の任意の元 a が $a \ge M$ を満たすとき，A は下に有界（bounded from below）といい，M を A の 下界（lower bound）という．A が上にも下にも有界な場合は単に 有界（bounded）という．

定義からわかるように，上界や下界はギリギリの数でなくても良い．例えば，A を区間 $[0,1]$ とした場合には，-1 や -10 や -2345 はすべて A の下界である．同様に 1 や 123 や 33556 は A の上界である．でもこの定義では A がどこまで広がっているのかがわからない．そこで，A の端と端を決める（ギリギリの数にする）つもりで,「上限」と「下限」を定義する．でもその前に：

以下の定義を導入する：

定義 3.1.2 (上限と下限) A を実数の集合とする．A が上に有界のとき，A の上界の最小値 を A の上限（supremum）と定義し，$\sup A$ と書く．同様に A が下に有界のとき，A の下界の最大値 を A の下限（infimum）と定義し，$\inf A$ と書く．

（注）上限と上界は間違いやすいから，注意する事.（正直，僕は日本語だとどっちがどっちだったかすぐにわからなくなる.）



（注意！）上では「A の上界の最小値」や「A の下界の最大値」があたかも存在するかのような書き方をしたが，これは以下の公理 3.1.3 で要請する．だから，論理の順序を重んじるなら，まず下の公理を書いてから，上の定義

[24] もちろんそれ以外にも,「代数方程式の解が存在するように」などの理由もあるが，解析学の立場からは極限の存在を保証するためと言って良い（と思うんだが）

[25] 実際にどのように構成するかについては，この講義ノートへの付録の形で述べることにした．まだ書きかけではあるが，その原稿がこの講義の web page にあるので，興味のある人は自己責任で（間違いがある可能性も考慮の上）どうぞ

[26] 教科書の 2.2 節前半．ただし，この講義では実数の「切断」には触れていないので，定理 4 の証明は跳ばして良い

で上限や下限を定義すべきなのだ（教科書ではちゃんとそう書いている）．しかしその順序ではかえってわかりにくいと思ったので、敢えて上の順序で書いた．下では［$\cdots$］の中はそれぞれ置き換えて読むべし．

公理 3.1.3 (上限と下限の存在) 実数の集合 S が上に［下に］有界ならば S の上界 [下界] の最小値 [最大値] が存在する．上の定義の用語を使うと，S が上に［下に］有界ならば S の上限 [下限] が存在する．

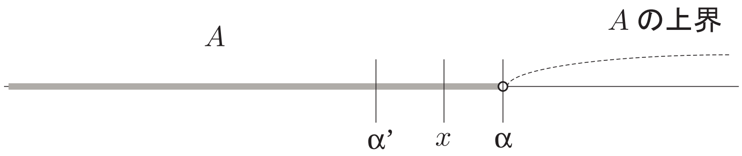
この公理から，我々に不可欠の以下の定理が導かれる．

定理 3.1.4 (上限と下限の特徴付け；教科書の定理 6) A を実数の集合とする．$\alpha = \sup A$ となる必要充分条件は次の 2 つが成り立つ事である：

(i) A に属するすべての x に対して，$x \le \alpha$

(ii) $\alpha' < \alpha$ なる任意の α' をとると，必ず，$\alpha' < x$ なる $x \in A$ が存在する

$\beta = \inf A$ の必要充分条件も同様である（不等式の向きがいろいろと変わるから注意．各自書き下して，教科書でチェックする事）．



定理 3.1.4 の証明　両方の向きを示せば良い．上の図を見て，直感を養おう．

I. $\alpha = \sup A$ のとき，定理の (i), (ii) が成り立つ事を示す．

まず，α は A の上界の最小値だから，A の上界ではある．だから，上界の定義から $x \le \alpha$（$\forall x \in A$）はアタリマエになりたつ．よって (i) は O.K.

次に，(ii) を示すには，背理法を使うのが良いだろう．もし，(ii) がなりたたないとすると，ある α' では $\alpha' < x$ となるような $x \in A$ が存在しない，ということだ．これはこの α' では $\alpha' \ge x$（$\forall x \in A$）という事だが，これでは α' が A の上界になってしまうぞ．でもこれは α が A の上界の最小値であった（上限の定義）に反する．ので背理法から，(ii) が成り立つ必要がある．

II. (i), (ii) がなりたつとき，$\alpha = \sup A$，つまり α は A の上界の最小値だということを示す．

まず，(i) は α が A の上界であることを保証している．問題は，α より小さい上界があるかどうかだけども，やはり背理法を使ってみる．

つまり，α より小さい上界が存在したとして，それを α' と書いてみると，$\alpha' \ge x$（$\forall x \in A$）が成り立っているはずである．ところが，(ii) によると，これは許されない！というわけで，α よりも小さい上界は存在せず，α は A の上界の最小値（つまり，上限）なのである． □

3.2 単調な数列[27]

これまで実数の連続性や上限・下限を延々とやった．これで本題に戻って極限の話を続けられる．これまでにも「行き先がわかっている極限」の定義は散々やってきた．$\lim_{n\to\infty} a_n = \alpha$ 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，

$$\forall \epsilon > 0, \quad \exists N(\epsilon), \qquad n > N(\epsilon) \implies |a_n - \alpha| < \epsilon \tag{3.2.1}$$

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，その収束を判定したい数列はいくらでもある．

[27] 教科書の 2.2 節後半

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

$$e = \lim_{n\to\infty}\Bigl(1+\frac{1}{n}\Bigr)^n \tag{3.2.2}$$

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？(高校まででではこの極限が存在する事をアタリマエのようにごまかして来たのだが，大学の数学科ではそれは許されない！）この数の存在が証明できなければ，皆さんが受験数学で一生懸命やった指数関数が存在しないことになってしまう...

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．というより，数学で扱う大抵の極限は「その値はきれいに書けないけど，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」したりするのだ.

（例）後で「テイラー展開」というものをやる．これは

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \lim_{N\to\infty}\sum_{n=0}^{N}\frac{x^n}{n!} \tag{3.2.3}$$

のような形の級数（上の級数がどうして出てくるかはあとのお楽しみ）だが，右辺の級数の値が一般の x でどうなるかなんて，さっぱりわからんでしょ？実は上の右辺を e^x の定義としてしまうことさえある．こうしたいのなら，右辺の極限の存在を証明できなければ非常に困る！

更に付け加えるなら，e^x については裏のズルイ手を使って，上の級数が存在することを証明できるからまあ良いのだ[28]．しかし，上の級数を少し変えて

$$\lim_{N\to\infty}\sum_{n=0}^{N}\frac{x^n}{n\,n!} \qquad \lim_{N\to\infty}\sum_{n=0}^{N}\frac{x^n}{\sqrt{n}\,n!} \tag{3.2.4}$$

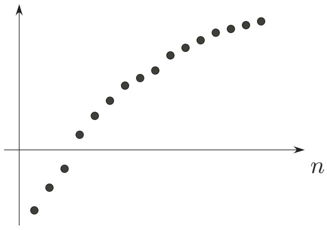
などを考えだすともうお手上げだ...

という訳で，<u>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言える</u>ような定理が欲しい．これに応えてくれるのが,「単調増加（減少）列」（この節）,「上極限と下極限」（3.3 節）,「コーシー列」（3.4 節）である．この小節では一番簡単で直感的な単調列を考える.

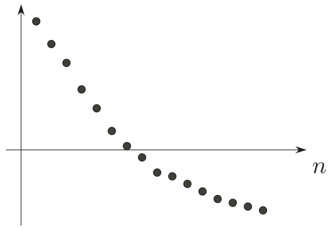
定義 3.2.1 (単調列) $a_1 \le a_2 \le a_3 \le \ldots \le a_n \le \ldots$ となっている数列 a_n を<u>広義の単調増加</u>数列，または<u>単調非減少</u>数列という（不等号にイコールが入ってないものは単調増加数列という）．不等号が逆向きになったのは「広義の単調減少」または「単調非増加」数列という.

（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（関数については後の定義 5.3.1 参照）.
- 「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.



単調増加数列の例



単調減少数列の例

[28]ただし，e^x をいう関数そのものの定義には関数の連続性など，結局は実数の連続性に関連する事をどこかで使う必要がある．というわけで，ケッキョクのところ，実数の連続性（とその帰結）抜きには指数関数は扱えないから,「まあ良い」というのはちょっと言い過ぎ

さて，単調列と並んで大事な概念を定義しておこう

定義 3.2.2 (有界列) 数列 $\{a_n\}$ に対してある数 M が存在して，すべての n について $|a_n| < M$ が成り立っているとき，この数列は 有界 な数列だという．これは要するに集合 $\{a_n \,|\, n \geq 1\}$ が有界な集合（定義 3.1.1）だということだ．

（注）M 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．

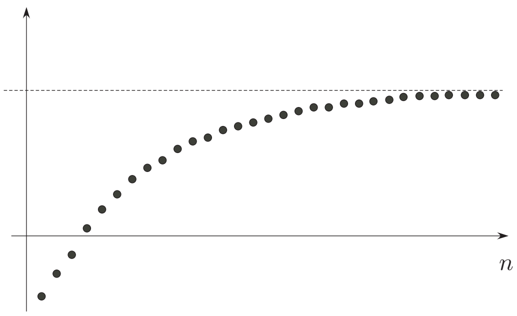
さて，有界かつ単調な数列には，以下の著しい性質がある．直感的にはあたりまえに見えるだろう．

定理 3.2.3 (有界単調列の収束；教科書の定理 7) 数列 $\{a_n\}$ が上に有界で広義単調増加のとき，$\lim_{n\to\infty} a_n$ 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，$\lim_{n\to\infty} a_n$ 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 $\lim_{n\to\infty} a_n = +\infty$ であるし，$\{a_n\}$ が有界でない広義単調減少列の場合は $\lim_{n\to\infty} a_n = -\infty$ 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，単調な数列では $\lim_{n\to\infty} a_n$ は常に存在する．

定理 3.2.3 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．a_n のそれぞれは有理数で，単調増加でもあり，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事で,（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．



定理 3.2.3 の証明　a_n が有界かつ広義単調増加であるとする（広義単調減少の場合は不等号の向きをひっくり返せば同じだから略）．集合 $\{a_1, a_2, a_3, \ldots\}$ の上限を α としよう（上限の存在は定理 3.1.3 で保証されている；ここに実数の連続性が効いている）．$\alpha = \lim_{n\to\infty} a_n$ であることを証明すれば十分である．

まず，上限の定義から，$a_n \leq \alpha$ が成り立っている事に注意する．逆向きの不等式がほしいので，a_n を下から $\alpha - \epsilon$（$\epsilon > 0$ は n に依存するけど，n を大きくとるといくらでも小さくできる）のような形で押さえることを試みよう．

そこで，まず任意の $\epsilon > 0$ を決めよう．次に，$\alpha' = \alpha - \epsilon$ を考える．すると α が集合 $\{a_n\}$ の上限だから，定理 3.1.4 によれば，$\alpha' < a_m \leq \alpha$ となるような自然数 m が存在する．そこで，この m を使って $N(\epsilon) := m$ と定義すると，$n \geq N(\epsilon)$ では

$$\alpha' < a_m \leq a_n \leq \alpha \quad \text{つまり} \quad |a_n - \alpha| < \epsilon \tag{3.2.5}$$

が成り立つ（真ん中の不等号は a_n の広義単調性から）．これは ϵ-N で書いた収束の定義そのものなので，$\lim_{n\to\infty} a_n = \alpha$ が証明された． □

3.3 上極限と下極限

数列の収束判定に関わる重要な概念の２番目は「上極限」「下極限」である．この内容は教科書にはないが，重要かつ便利なのでここで導入しておく.（単調列の考えがうまく使われるのを実感して欲しい.）

定義 3.3.1 (上極限，下極限) 数列 $\{a_n\}$ に対して，

$$\liminf_{n\to\infty} a_n := \lim_{N\to\infty}\Bigl(\inf_{n\ge N} a_n\Bigr), \qquad \limsup_{n\to\infty} a_n := \lim_{N\to\infty}\Bigl(\sup_{n\ge N} a_n\Bigr) \tag{3.3.1}$$

の２つの極限を考え（この極限が存在するか否かは下の定理参照），前者を下極限，後者を上極限という．また，別の記法として $\liminf_{n\to\infty} a_n = \underline{\lim}_{n\to\infty}\, a_n$， $\limsup_{n\to\infty} a_n = \overline{\lim}_{n\to\infty}\, a_n$ とも書く（下線と上線に注意）.

(3.3.1) の主張は，新しく数列

$$b_N = \inf_{n\ge N} a_n = \inf\{a_n \,|\, n \ge N\}, \qquad c_N = \sup_{n\ge N} a_n = \sup\{a_n \,|\, n \ge N\} \tag{3.3.2}$$

を定義した上で，上極限と下極限を

$$\liminf_{n\to\infty} a_n := \lim_{N\to\infty} b_N, \qquad \limsup_{n\to\infty} a_n := \lim_{N\to\infty} c_N \tag{3.3.3}$$

と定義する，という事である．上の定義を簡略化して書いたものだと思えば，この記号は理解しやすいのではないか.

なお，$\limsup$，$\liminf$ の定義は sup，inf の定義と関連はしているが，もちろん同じものではない．違いをいろんな数列の例で体感すること.

通常の $\lim_{n\to\infty} a_n$ とは異なり，$\limsup$ と $\liminf$ には以下の著しい性質がある.

定理 3.3.2 (上極限，下極限はいつでも存在)
(i)（極限の値として $+\infty, -\infty$ も許せば）数列 a_n の上極限，下極限は いつでも存在 し，以下を満たす：

$$\liminf_{n\to\infty} a_n \le \limsup_{n\to\infty} a_n \tag{3.3.4}$$

(ii) 特に数列 $\{a_n\}$ が有界（$\exists M,\ \forall n,\ |a_n| \le M$）ならば上極限と下極限は有限の値をとる.
(iii) 「数列が収束すること」と「上極限と下極限が一致する事」は同値である．つまり，

$$\lim_{n\to\infty} a_n = \alpha \qquad \Longleftrightarrow \qquad \limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = \alpha \tag{3.3.5}$$

（上極限・下極限の効用）極限がわからない数列を扱うのは通常，なかなか大変だ．特に，極限がそもそも存在するかどうかわからないのは非常に困る．ところが上のように，有界な数列に対しては（たとえ普通の極限が存在しなくても）上極限と下極限はいつでも存在するので，非常にありがたい．特に上の定理の (iii) によれば，上極限の値を β，下極限の値を γ と仮定して，この２つが等しくなるかどうかを吟味して数列の収束を判定することができる.

定理 3.3.2 の略証

(i) と (ii). 数列 $\{a_n\}$ が有界の場合のみを考える（有界でない場合はむしろ簡単だから練習問題）．記号と考える対象を明確にするため，(3.3.2) で定義された数列 b_N, c_N を考える．まず上限，下限の定義から，

$$b_n \le b_{n+1} \le b_{n+2} \le \ \cdots\ \le c_{n+2} \le c_{n+1} \le c_n \tag{3.3.6}$$

が成り立つことに注意する．つまり，b_n は広義の単調増加，c_n は広義の単調減少で，更に $b_m \le c_n$ がいつでも成り立つ．また，上限，下限としての b_n, c_n の定義から

$$m \ge n \quad \text{ならば} \quad b_n \le a_m \le c_n \tag{3.3.7}$$

も成り立っている（後で使う）．

さて，$\{a_n\}$ が有界なら，$\{b_N\}$ も有界である（why?）．従って定理 3.2.3 から $\lim_{N\to\infty} b_N$ の存在が保証される．(3.3.3) を思い出せば，これは $\liminf a_n$ の存在に他ならない．上極限 $\limsup a_n$ についても同様である．

(iii). 同値関係を示すのだから，両方の向きの矢印を示せば良い．まず，右向きの矢印を示す．$\lim_{n\to\infty} a_n = \alpha$ ということは

$$\forall\epsilon \quad \exists N \qquad n \geq N \quad \Longrightarrow \quad \alpha - \epsilon < a_n < \alpha + \epsilon \tag{3.3.8}$$

という事であった．ここで N を固定して最後の不等式部分で $n \geq N$ について inf と sup をとると，

$$\forall\epsilon \quad \exists N \qquad \Bigl(\alpha - \epsilon \leq b_N \leq \alpha + \epsilon \ \text{かつ}\ \alpha - \epsilon \leq c_N \leq \alpha + \epsilon\Bigr) \tag{3.3.9}$$

が得られる．ところが $\{b_N\}$ は広義の単調増加だから，$\alpha - \epsilon \leq b_N$ ならば，すべての $n \geq N$ で $\alpha - \epsilon \leq b_n$ がなりたつ．同様に，c_N が広義の単調減少なので，$c_N \leq \alpha + \epsilon$ ならば，すべての $n \geq N$ で $c_n \leq \alpha + \epsilon$ がなりたつ．従って，

$$\forall\epsilon \quad \exists N \qquad n > N \quad \Bigl(\alpha - \epsilon \leq b_n \ \text{かつ}\ c_n \leq \alpha + \epsilon\Bigr) \tag{3.3.10}$$

が結論できる．ところが (3.3.7) で注意したように，すべての m, n に対して $b_n \leq c_m$ である．よって上のは

$$\forall\epsilon \quad \exists N \qquad n > N \quad \Bigl(\alpha - \epsilon \leq b_n \leq c_n \leq \alpha + \epsilon\Bigr) \tag{3.3.11}$$

とまとめられる．これから特に

$$\forall\epsilon \quad \exists N \qquad n > N \quad \Bigl(|b_n - \alpha| \leq \epsilon \ \text{かつ}\ |c_n - \alpha| \leq \epsilon\Bigr) \tag{3.3.12}$$

が結論できるが，これは $\lim_{n\to\infty} b_n = \alpha$，$\lim_{n\to\infty} c_n = \alpha$ を ϵ-N で書いたものに他ならない．

次に左向きの矢印を示す．$\lim_{n\to\infty} b_n = \alpha$，$\lim_{n\to\infty} c_n = \alpha$ をまとめて ϵ-N で書くと

$$\forall\epsilon > 0 \quad \exists N_1, N_2, \qquad \Bigl\{\bigl(\, n > N_1 \implies |b_n - \alpha| < \epsilon \,\bigr) \cap \bigl(\, m > N_2 \implies |c_m - \alpha| < \epsilon \,\bigr)\Bigr\} \tag{3.3.13}$$

となる．ここで N_1 と N_2 の大きい方を改めて N と書くと，上から

$$\forall\epsilon > 0 \quad \exists N, \qquad n > N \quad \Longrightarrow \quad \Bigl(|b_n - \alpha| < \epsilon \quad \text{かつ} \quad |c_n - \alpha| < \epsilon\Bigr) \tag{3.3.14}$$

が得られる．ここで更に $b_n \leq c_n$ を考えに入れると

$$\forall\epsilon > 0 \quad \exists N, \qquad n > N \quad \Longrightarrow \quad \alpha - \epsilon < b_n \leq c_n < \alpha + \epsilon \tag{3.3.15}$$

が得られる．ところが，(3.3.7) の特殊な場合として $b_n \leq a_n \leq c_n$ であるから，上から直ちに

$$\forall\epsilon > 0 \quad \exists N, \qquad n > N \quad \Longrightarrow \quad \alpha - \epsilon < a_n < \alpha + \epsilon \tag{3.3.16}$$

が得られる．これは $\lim_{n\to\infty} a_n = \alpha$ を ϵ-N で書いたものに他ならない． □

ついでに，知っておくと便利な定理をあげておこう．

定理 3.3.3 (有界な数列は収束する部分列を含む；教科書の定理 9) 有界な数列 $a_1, a_2, a_3, \ldots$ から収束する部分列 $b_1, b_2, \ldots$ を取り出すことができる．

この定理も図を書いてみるとほとんど当たり前ではある．数列が有界なんだから，有界なところに無限個の数列の値を放り込むと，どこかで無限個が溜まった形にならざるを得ない．以下の証明は区間を分割する事で「無限個が溜まる」ところを捕まえている．もちろん，このような収束部分列の収束先は二つ以上ある事もある．

（証明の概略）数列が有界なので，$a < b$ をうまくとれば，すべての n で $a \leq a_n \leq b$ とできる．そこで閉区間 $[a, b]$ を 2 等分して区間 $[a, \frac{a+b}{2}]$ と $[\frac{a+b}{2}, b]$ を作ろう（$\frac{a+b}{2}$ は両方の区間に入ってるけど問題ない）．

$\{a_n\}$ は全体で無限個あるから，$[a, b]$ を 2 等分してできた区間の少なくとも片方は無限個の a_n を含んでいるはずだ．そこで例えば左側の区間（これを I_1 とする）に無限個の a_n があったとして，I_1 をまた 2 等分してみる．I_1 に

は無限個の a_n が入っているから，I_1 を 2 等分した区間の少なくとも一方にはやはり無限個の a_n が入っているはずである．そこで無限個入ってるのを I_2 としよう．

以下，こんな調子で $[a, b]$ を分割していくと，区間の列 $I_1 \supset I_2 \supset I_3 \supset \ldots$ ができ，I_ℓ の幅は $\dfrac{b-a}{2^\ell}$ になっている．

これを用いて，$\{a_n\}$ が初めて I_ℓ に入ったときの a_n の値を b_ℓ と定義することで部分列 $\{b_\ell\}$ を作ろう．さて，$\limsup_{\ell\to\infty} b_\ell$ と $\liminf_{\ell\to\infty} b_\ell$ は常に存在するが，この両者は等しい（なぜなら $m \geq \ell$ ではすべての b_m が区間 I_ℓ に入っているから；各自きちんと納得すること）．従って定理 3.3.2 によってこの部分列 $\{b_\ell\}$ の収束が保証される． □

3.4 コーシー列[29]

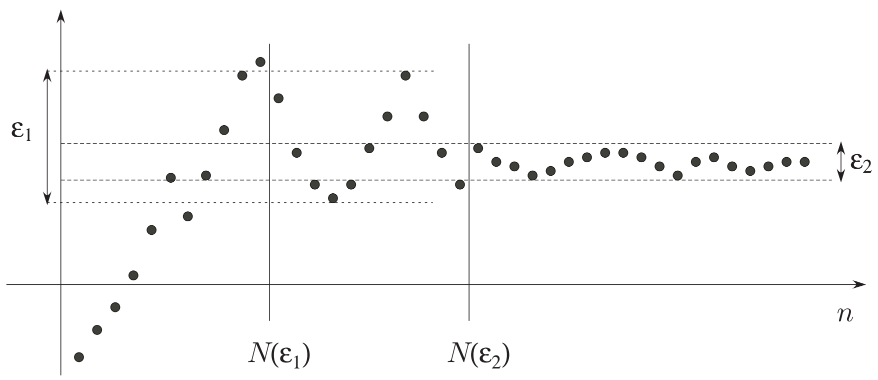
さて，世の中の収束する数列の中には，単調列でないものもたくさんある．そのようなものが本当に収束するかを判断するには，前節の単調列の定理だけでは足りない．そもそも，ある数列が収束する事の必要十分条件は何なのだろう？この答えは以下の「コーシー列」で与えられる（コーシー偉い！）

定義 3.4.1 (コーシー列) 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）整数 $N(\epsilon)$ がとれて，

$$\text{すべての } m, n \geq N(\epsilon) \text{ に対して} \quad \bigl|a_m - a_n\bigr| < \epsilon \quad \text{とできる．} \tag{3.4.1}$$

（注）この定義そのものがなかなか理解しにくいようだ．「次の数列がコーシー列かどうか判定せよ」というような，定義そのものを訊く問題でさえ，テストになると珍解答が続出するので要注意である．



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）を見れば収束するかどうかが判定できる，というのである．これは実用上，非常に重要だ．

定理 3.4.2 (コーシーの収束条件；教科書の定理 11，非常に大事) 数列 a_n が（何かの値に）収束することと，a_n がコーシー列であることは同値である．つまり，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．

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」ことと「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．（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．前節でやった lim sup と lim inf が役に立つ事はかなりある．）

[29]教科書の 2.4 節に相当．ぼんやり聞いているとコーシー列の概念はなかなかわからないだろう．ϵ-δ に次ぐこの科目の鬼門だから，心するように．教科書もちゃんと読むこと

問 3.4.3 「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 3.4.2 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．

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

問 3.4.4 「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 の極限値を求めよ．

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

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

正直，c_n はそこそこ難しいと思うが，ともかく a_n, b_n はできるようになろう．

問 3.4.5 以下の（例）のそれぞれが収束する事を実際に証明せよ．(3.2.4) の例にも挑戦してみよう．コツがわかれば，そんなに難しいものではないですよ．

（例）コーシー列の考えを使うと収束が証明できるものの典型例（コーシー列を使わなければ証明できないという訳ではないが）を挙げておこう．

- 既に言ったけど，e^x のテイラー展開 $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ は[30]すべての実数 x で収束する．$x > 0$ なら有界単調列の性質を用いても証明できるが，コーシー列になっていることを確かめた方がすべての x ができて簡単だ．とはいえ，実際にコーシー列になっていることを示すには，ある程度の計算力が必要だ．腕に覚えのある人は挑戦してみるとよい．
- $\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$ もすべての実数 x で収束する．この場合もコーシー列になっていることを確かめるのが簡単だろう．
- $0 < r < 1$ を定数とする．数列 $\{a_n\}$ が，$|a_{n+2} - a_{n+1}| \leq r|a_{n+1} - a_n|$ （$n = 1, 2, 3, ...$）を満たすとき，この数列はコーシー列であって，従って収束する．(この例をより一般の空間に拡張したものは「縮小写像の原理」とよばれ，関数解析の強力な手法の一つになっている.)

収束列ならコーシー列，の証明

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

$$|a_m - \alpha| < \frac{\epsilon}{2}, \quad |a_n - \alpha| < \frac{\epsilon}{2} \tag{3.4.2}$$

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

$$|a_m - a_n| = \bigl|(a_m - \alpha) + (\alpha - a_n)\bigr| \leq |a_m - \alpha| + |\alpha - a_n| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \tag{3.4.3}$$

が成り立つ．これはコーシー列の条件 (3.4.1) が成り立っていることを意味する． □

[30]この段階では，これがなぜ e^x かは理解していないことになっている．だから，左辺の e^x は単なる記号だと思ってください

コーシー列なら収束列，の証明

こちらの証明が大変だ．実数の連続性，上極限・下極限を延々とやってきたのは，この証明をやりたかったからである．証明の概略を述べるが，上極限，下極限がうまく使われているところを噛み締めて欲しい．

1．まず，コーシー列は有界な数列である．これは簡単だから，各自で証明すべし．

2．有界な数列は有限の上極限，下極限をもつ（定理 3.3.2）から，これを

$$\beta := \liminf_{n\to\infty} a_n, \qquad \gamma := \limsup_{n\to\infty} a_n \tag{3.4.4}$$

と書こう．もちろん，$\beta \le \gamma$ である．以下では記号を簡単にするため，(3.3.2) で定義された数列も考える，つまり，$b_N := \inf_{m\ge N} a_m$, $c_N := \sup_{n\ge N} a_n$.

3．定理 3.3.2 によれば，$\beta = \gamma$ を言えば十分であるから，これを目指そう．そのために $\{a_n\}$ がコーシー列であるとの条件を使う．コーシー列だから，

$$\forall\epsilon>0, \quad \exists N \qquad \Bigl(\ell, m \ge N \quad\Longrightarrow\quad |a_\ell - a_m| < \epsilon\Bigr) \tag{3.4.5}$$

が成り立っているが，最後の絶対値を外した

$$\forall\epsilon>0, \quad \exists N \qquad \Bigl(\ell, m \ge N \quad\Longrightarrow\quad a_\ell - a_m < \epsilon\Bigr) \tag{3.4.6}$$

も当然なりたつ．

4．最後の不等式の両辺で N, m を固定したまま $\ell \ge N$ に関する sup をとると，$\sup_{\ell\ge N} a_\ell = c_N$ であるから，

$$\forall\epsilon>0, \quad \exists N \qquad \Bigl(m \ge N \quad\Longrightarrow\quad c_N - a_m \le \epsilon\Bigr) \tag{3.4.7}$$

が得られる．

4′．今度はここで $m \ge N$ に関する inf をとってみると $\inf_{m\ge N} a_\ell = b_N$ であるから，

$$\forall\epsilon>0, \quad \exists N \qquad c_N - b_N \le \epsilon \tag{3.4.8}$$

が得られる．

5．さて，常に $c_n \ge b_n$ であることは何回も注意した．また，$\{b_n\}$ は広義の単調増加，$\{c_n\}$ は広義の単調減少であるから，$\{c_n - b_n\}$ は広義の単調減少である．従ってある N にて $c_N - b_N \le \epsilon$ であれば，すべての $n \ge N$ でも $c_n - b_n \le \epsilon$ である．つまり，(3.4.8) から

$$\forall\epsilon>0, \quad \exists N \qquad \Bigl(n \ge N \Longrightarrow \quad |c_n - b_n| \le \epsilon\Bigr) \tag{3.4.9}$$

が結論できる．これは $\lim_{n\to\infty}(c_n - b_n) = 0$ を ϵ-N で書いたものに他ならず，$\beta = \gamma$ が結論できる． □

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

定理 3.4.6 (コーシーの収束条件；教科書の定理 11，非常に大事) $\lim_{x\to a} f(x)$ 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)　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

証明：

$\lim_{x\to a} f(x)$ が存在するならばコーシーの条件が成り立つのは，数列の場合の定理 3.4.2 の証明と同様にできるから，略（各人でちゃんと確かめる事！）．

コーシーの条件 (C) が成り立っておれば $\lim_{x\to a} f(x)$ が存在することの証明をこれから行おう．そのために，以前に簡単に触れた定理 2.3.1 を用いる．定理 2.3.1 によると，$\lim_{x\to a} f(x) = b$ ということは，

$$a_n \neq a \text{ かつ } \lim_{n\to\infty} a_n = a \text{ であるすべての数列 } \{a_n\} \text{ について } \lim_{n\to\infty} f(a_n) = b \text{ である}$$

と同値であった．そこで，上が成り立っている事を証明しよう．

さて，コーシーの条件 (C) は，$a_n \neq a$ かつ $\lim_{n\to\infty} a_n = a$ である数列 $\{a_n\}$ をもってくると，数列 $\{f(a_n)\}$ がコーシー列になっている事を保証する．よって，$\lim_{n\to\infty} f(a_n)$ は存在する．

ところがこれで終わりではない．上では数列 $\{a_n\}$ を決める毎に $\lim_{n\to\infty} f(a_n)$ が存在することが保証されたけども，この極限が $\{a_n\}$ の取り方に依らないことも証明しなければ，定理 2.3.1 をつかうわけに行かないからだ．しかし，極限が $\{a_n\}$ の取り方に依らないことは，またもやコーシーの条件 (C) から簡単に証明できる．(例えば，極限が 2 つ以上あったとして，これが (C) に矛盾する事を示せば良い．) □

4 連続関数

さて，実数の連続性やコーシー列をやったので，連続関数をちゃんと扱えるようになった．まずは高校でもやったはずの定義を思い出そう．

4.1 連続関数の定義

定義 4.1.1 点 a を含む区間で定義された $f(x)$ が「a で連続」とは，$\lim_{x\to a} f(x) = f(a)$ 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，

$$0 < |x-a| < \delta(\epsilon) \text{ なるすべての } x \text{ で，} \quad \bigl|f(x)-f(a)\bigr| < \epsilon \quad \text{とできる．} \tag{4.1.1}$$

（ウ）は数式では

$$\forall\epsilon > 0, \quad \exists\delta(\epsilon) > 0, \qquad 0 < |x-a| < \delta(\epsilon) \implies \bigl|f(x)-f(a)\bigr| < \epsilon \tag{4.1.2}$$

となる．

なお，片側連続を問題にすることもある．

定義 4.1.2 関数 $f(x)$ 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 $\lim_{x\to a+0} f(x) = f(a)$ 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 $\lim_{x\to a-0} f(x) = f(a)$ 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，

$$c \in (a,b) \text{ では } \lim_{x\to c} f(x) = f(c), \qquad \text{かつ} \qquad \lim_{x\to a+0} f(x) = f(a), \quad \lim_{x\to b-0} f(x) = f(b) \tag{4.1.3}$$

  となることである（区間の中では普通の連続，区間の端点では左（右）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）極限値 $\lim_{x\to a} f(x)$ だ．単に右側からと左側からの極限値が同じでも連続ではないから注意.（例を考えよ.）

問 4.1.3 関数 $f(x) = \sqrt{|x|}$ が，任意の x で連続であることを，定義に戻って示せ．

問 4.1.4 関数 $f(x)$ 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？

4.2 連続関数の性質

連続関数について特に重要な，２つの定理を学ぶ．その前に，（実数の連続性とは関係なく成り立つ）連続関数の簡単な性質を述べておこう．

定理 4.2.1 (教科書の p.93，例 2) 点 a を含むある区間で定義された関数 $f(x)$ が $x = a$ で連続で，かつ $f(a) > 0$ ならば，a の十分近くでは $f(x) > 0$ である（x の範囲をもっと絞れば，$f(x) > \frac{1}{2}f(a)$ とも言える）．$f(a) < 0$ 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．

（証明）$f(x)$ が $x=a$ で連続であることの定義を ϵ-δ で書けばすぐに出る．各自必ずやってみること． □

さて，いよいよ本論．連続関数の重要な性質の一つ目は，高校でも結果だけはやったはずの中間値の定理である．

定理 4.2.2 (中間値の定理；教科書の定理 19) 閉区間 $[a,b]$ で連続な関数 $f(x)$ 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．

これまでにも強調してきたが，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．例えば関数 $f(x)=x^2-2$ 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」ことで対症療法的に対処できるが，一般の関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」をもつ集合まで拡げておく事は不可欠だったのだ．

定理 4.2.2 の証明： $f(a)<F<f(b)$ の場合を考えれば十分だ．$f(a)>f(b)$ の場合に以下の証明をどう変えるかは明らかだろうし，$f(a)=f(b)$ または $f(a)=F$ などの場合は $c=a$ とでもすれば良い．

以下では定理 4.2.1 を使うので，$g(x):=f(x)-F$ を考える．我々は $g(c)=0$ となるような c を探したい．

さて，$x=a$ から出発して x を段々と大きくしていく事を考える．出発点では $g(a)<0$ で，かつ $g(x)$ は連続だから，定理 4.2.1 によって $x=a$ のものすごく近くでは $g(x)<0$ のはずである．つまり，a に十分近い y では，

$a\le x<y$ なるすべての x にて $g(x)<0$ が成り立つ

ようにできるわけだ．このような y の全体を Y と書こう．$g(b)>0$ だから，上のような y は b を超える事はない．つまり，Y は有界である．従って特に，Y の上限 α が存在する（ここで実数の連続性から出る定理 3.1.3 を使った；実数の連続性が必要になるのはここだけだ）．以下では $g(\alpha)=0$ であることを示して，証明を完結する．そのためには，$g(\alpha)<0$ と $g(\alpha)>0$ が共に許されない事を示せば良い．

まず，$g(\alpha)<0$ と仮定すると，上の定理 4.2.1 から，α の近くの x でも $g(x)<0$ が結論される．これは Y が α よりも右にのびている事を意味し，Y の上限としての α の定義に矛盾する．

次に，$g(\alpha)>0$ と仮定すると，やはり定理 4.2.1 から，α の近くの x でも $g(x)>0$ が結論される．つまり，$\delta>0$ が存在して，$\alpha-\delta<x\le\alpha$ なる x でも $g(x)<0$ となってしまう．これでは Y の上限は α よりも左にあることになり，やはり Y の上限としての α の定義に矛盾する．

以上から $g(\alpha)=0$ しかあり得ないことがわかった．つまり，定理の結論は $c=\alpha$ で成り立つ． □

さて，連続関数の重要な性質その 2 は以下の最大値，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 4.2.3 (連続関数の最大値・最小値；教科書の定理 20) 閉区間で連続な関数は必ず，その区間内で最大値，最小値をとる．従って特にこのような関数は有界である．

- 「閉 区間で連続である」ことは重要な条件である．例えば，関数 $f(x)=1/x$ を開区間 $(0,1)$ で考えると，こいつは最大値を持たず，有界でもない．また，$g(x)=x$ を同じく開区間 $(0,1)$ 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関数 $g(x)=2-x^2$ は，x が有理数に限定されている限り最大値を持たない．

定理 4.2.3 の証明： 問題の閉区間を $[a,b]$ とし，最大値の存在を示す（最小値も同様にできる）．

連続な変数 x についての最大は考えにくいので，区間 $[a,b]$ を n 等分した点をまず考えて，この $(n+1)$ 個の点の中での最大値をまず探すことにしよう．あとで $n\to\infty$ としてやることで，本当の最大値を探す作戦だ．

そこで，この $(n+1)$ 個のなかで $f(x)$ を最大にするもの（これが複数ある場合はそのうちのどれでも良い）を x_n と書くことにする．$n = 1, 2, 3, \ldots$ とすると，数列 $x_1, x_2, x_3, \ldots$ ができる．この数列は有界な数列であるから，定理 3.3.3 によれば収束する部分列を含む．その部分列の要素を $y_\ell = x_{i_\ell}$ と書こう．(この意味は，y_ℓ はもともとの $x_1, x_2, \ldots$ の添字が i_ℓ となったもので与えられている，つまり y_ℓ は i_ℓ 等分した時の分点である，ということ.)

さて，$\alpha := \lim_{\ell\to\infty} y_\ell$ と書いたとき，この α で $f(x)$ が最大になる事を証明しよう．そのためには，任意の x を一つとり（$a \le x \le b$），この x に対して $f(\alpha) \ge f(x)$ であることを示せば良い．そこで，y_ℓ を定義した時の i_ℓ 分点の中から，x に一番近いものを選んで，それを z_ℓ と書く事にする．i_ℓ 分点の中で $f(x)$ が最大のものを y_ℓ としたから，

$$f(z_\ell) \le f(y_\ell) \tag{4.2.1}$$

がなりたっている．

この両辺で ℓ を無限大にしてみよう．右辺の y_ℓ は α に収束する（α の定義）．一方，$\ell \to \infty$ では分点の数も無限大に（従って，分割は限りなく細かく）なるから，左辺の z_ℓ は x に収束するはずだ．従って，f が連続であることも考えに入れると，左辺は $f(x)$ に，右辺は $f(\alpha)$ に収束するはずである．つまり，

$$f(x) \le f(\alpha) \tag{4.2.2}$$

が言え，x は任意だったから，これから $f(\alpha)$ が最大値であるとわかる． □

4.3 連続関数の効用：指数関数と対数関数

ここまでの準備を経て，x^α（α は実数，$x > 0$），指数関数，対数関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，

$$x^\alpha = \lim_{n\to\infty} x^{a_n} \tag{4.3.1}$$

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[31]．時間の関係上，この講義ではこれ以上触れることができないだろうから，教科書の 3.3 節を各自，読んで下さい．

[31] 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 $\lim_{n\to\infty} a_n = \alpha$ 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．教科書の p.116 も参照

5 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（高校でやったことになっている）微分の基礎付けを行って行こう．もちろん，高校ではやらなかった（はずの）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．

- 微係数は関数の「変化率」を表すから，微分の値（正負）を知ることで，関数の 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た．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．
- 微分を利用して関数を級数に展開できる（テイラー展開）．これを利用して，関数の近似値 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 — ただしこれが形を変えて,「多変数関数の微分」（偏微分）として登場する．

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陰も形もなくなっている．こんな重要な話題を削るのは怪しからん！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．ただし，その前に，高校で「ええ加減」にやった微分の基礎付けから始めよう —— とは言っても，かなりの部分は極限の基礎付けをやったことで終わっているのだが...

5.1 微分の定義

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 5.1.1 (微分係数) $x = a$ とその近傍で定義されている関数 $f(x)$ に対して，極限

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} \tag{5.1.1}$$

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 $\dfrac{df}{dx}(a)$ 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．

また，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関数だと考えられるので，これを f の導関数（derived function，または derivative）とよぶ．微分係数は，考えている関数の「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，$x \to a$ というのは $|x - a| \to 0$ の事であったから，(5.1.1) 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 5.1.2 (片側微分係数) 定義 5.1.1 の状況の下で，極限

$$f_-(a) := \lim_{x\to a-0} \frac{f(x) - f(a)}{x - a} \tag{5.1.2}$$

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，

$$f_+(a) := \lim_{x\to a+0} \frac{f(x) - f(a)}{x - a} \tag{5.1.3}$$

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（証明できますか？）．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．

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：

定理 5.1.3 (教科書の定理 25) 関数 $f(x)$ 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出るので略．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関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ど，どの点でも微分不可能な関数も（なかなか想像しにくいが）存在する．一つの例が教科書の p.129 に載っている（Weierstrass）．

5.1.1 高階導関数の定義

高校でもやったと思うけど，高階の導関数についてまとめておく．

関数 $f(x)$ に対して．それを n-回微分してできる関数を <u>n-階の導関数</u>（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 $f'(x), f''(x), f'''(x)$ とも書く．具体的には

$$f^{(2)}(x)=\frac{d^2}{dx^2}f(x)=\frac{d}{dx}\Big\{\frac{d}{dx}f(x)\Big\}, \qquad f^{(3)}(x)=\frac{d^3}{dx^3}f(x)=\frac{d}{dx}\Big[\frac{d}{dx}\Big\{\frac{d}{dx}f(x)\Big\}\Big], \quad \ldots \tag{5.1.4}$$

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）．

高階の導関数については <u>ライプニッツ（Leibniz）の公式</u> が成り立つ．つまり

$$\frac{d}{dx}\big\{f(x)g(x)\big\}=f'(x)g(x)+f(x)g'(x), \qquad \frac{d^2}{dx^2}\big\{f(x)g(x)\big\}=f''(x)g(x)+2f'(x)g'(x)+f(x)g''(x) \tag{5.1.5}$$

で，より一般には（n は自然数）

$$\frac{d^n}{dx^n}\big\{f(x)g(x)\big\}=\sum_{k=0}^{n}\binom{n}{k}f^{(k)}(x)\,g^{(n-k)}(x), \qquad \binom{n}{k}\equiv {}_nC_k=\frac{n!}{k!\,(n-k)!} \tag{5.1.6}$$

となる[32]．この証明は数学的帰納法でできるから，一度は自力でやっておくこと．ただし，その途中で恒等式

$$\binom{n}{k}=\binom{n-1}{k}+\binom{n-1}{k-1} \tag{5.1.7}$$

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）ある開区間 I で定義された関数 $f(x)$ が n 回微分可能で，<u>更に $f^{(n)}(x)$ が連続</u> のとき，この関数は開区間 I で <u>C^n-級</u> である，という．いうまでもなく，$m<n$ ならば，C^n-級の関数は 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，連続な関数の導関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．

5.2 平均値の定理

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．せっ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，開区間では微分可能）に注目してほしい．

まずはロルの定理．定理の下の左側の図を見れば，直感的には明らかだろう．

定理 5.2.1 (ロル Rolle の定理；教科書の定理 28) $f(x)$ が閉区間 $[a,b]$ で連続，開区間 (a,b) で微分可能．更に $f(a)=f(b)=0$ とする．このとき

$$f'(\xi)=0 \qquad (a<\xi<b) \tag{5.2.1}$$

となる ξ が存在する．

[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．その導出法を比べると，同じ二項係数 $\binom{n}{k}$ が出る理由がわかるだろう

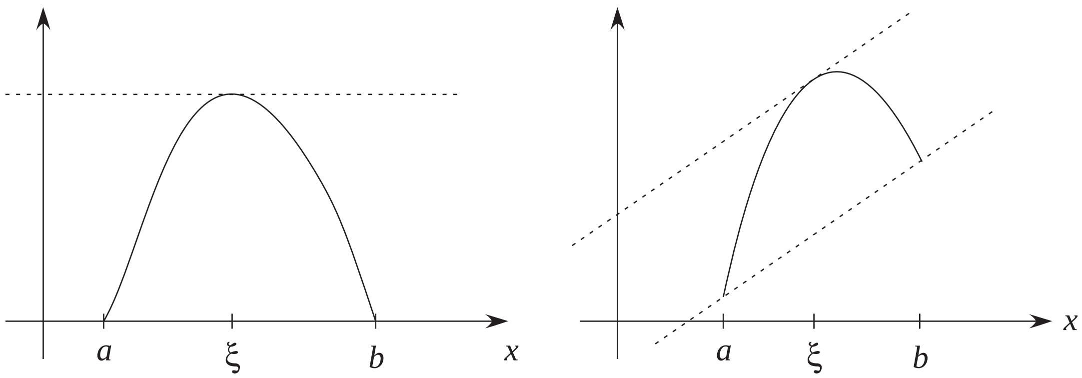
（注）定理の ξ は一般には a, b の両方に依存して決まる．アタリマエだが，注意の事．

（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[33]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，同じことである)．

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 4.2.3）．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

$$f'(\xi) = \lim_{h \to 0} \frac{f(\xi + h) - f(\xi)}{h} \tag{5.2.2}$$

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，$h \to 0$ 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． □



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 5.2.2 (平均値の定理；教科書の定理 27) $f(x)$ が閉区間 $[a, b]$ で連続，開区間 (a, b) で微分可能と仮定する．このとき，

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) \qquad (a < \xi < b) \tag{5.2.3}$$

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． □

以上で平均値の定理の主要な部分はおしまいだが，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 5.2.3 (コーシーの平均値の定理；教科書の定理 30) $f(x)$ と $g(x)$ が共に閉区間 $[a, b]$ で連続，開区間 (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，

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} \qquad (a < \xi < b) \tag{5.2.4}$$

となる ξ が存在する．

[33]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

（注）教科書にも注意してあるように，$g'(x) \neq 0$ から $g(a) \neq g(b)$ は保証されている．

（証明）$k := \dfrac{f(b)-f(a)}{g(b)-g(a)}$ とおいて，$F(x) := f(x) - f(a) - k\{g(x) - g(a)\}$ 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，定理を得る． □

5.3 平均値の定理の応用

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」については次節（来学期）で考える．

5.3.1 関数の増減

微分の応用として最重要なものの一つは，関数の増減や極大・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ある．まずは言葉の定義から始める．

定義 5.3.1 (単調な関数) 区間（開区間でも閉区間でも）I で定義された関数 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，単調増加な関数を単に「増加関数」．単調減少な関数を「減少関数ともいう．また単調増加と単調減少の両方をまとめて,「単調な」関数という．

数列のところ（定義 3.2.1）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 5.3.2 (導関数の符号と関数の増減；教科書の p.131, 定理 26) $f(x)$ が開区間 $I = (a, b)$ で微分可能と仮定する．このとき，

- I で常に $f'(x) > 0$ $\Longrightarrow$ I で $f(x)$ は狭義単調増加.
- I で常に $f'(x) < 0$ $\Longrightarrow$ I で $f(x)$ は狭義単調減少.
- I で常に $f'(x) = 0$ $\Longleftrightarrow$ I で $f(x)$ は定数関数.

（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（例は教科書の p.135）．

定理 5.3.2 は非常にわかりやすいものだが，実用上は不便なことがある —— 開区間の端点でどうなっているかが，この定理だけからはわからない.（例は教科書の pp.141–143）．この点を改良すると以下のようになる．

系 5.3.3 (導関数の符号と関数の増減；教科書の p.142，定理 26′) $f(x)$ が閉区間 $[a,b]$ で連続，かつ開区間 (a,b) で微分可能と仮定する．このとき，

- 開区間 (a,b) で常に $f'(x) > 0$ $\Longrightarrow$ 閉区間 $[a,b]$ で $f(x)$ は狭義単調増加．
- 開区間 (a,b) で常に $f'(x) < 0$ $\Longrightarrow$ 閉区間 $[a,b]$ で $f(x)$ は狭義単調減少．
- 開区間 (a,b) で常に $f'(x) = 0$ $\Longleftrightarrow$ 閉区間 $[a,b]$ で $f(x)$ は定数関数．

この系の仮定の微妙さに注目して欲しい：f は閉区間の端点では微分可能とは仮定していない —— これは平均値の定理と同じノリである．端点は f の連続性だけで十分だ，というのが上の定理のミソだ．

なお，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 $f(x) = x^3$）．しかし，これを広義単調にすると以下が成り立つ．

定理 5.3.4 (広義単調関数の増減と導関数の符号；教科書の定理 29) $f(x)$ が開区間 $I = (a,b)$ で微分可能と仮定する．このとき，

- I で常に $f'(x) \geq 0$ $\Longleftrightarrow$ I で $f(x)$ は広義単調増加．
- I で常に $f'(x) \geq 0$ $\Longleftrightarrow$ I で $f(x)$ は広義単調減少．

5.3.2 関数の極大・極小

では，極大と極小の問題に進む．これも高校で大略はやったはずだから，簡単に．

定義 5.3.5 点 $x = a$ が関数 $f(x)$ の<u>極大点</u>（local maximum）であるとは，

$$\exists r > 0, \qquad 0 < |x-a| < r \quad \Longrightarrow \quad f(x) < f(a) \tag{5.3.1}$$

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 $x = a$ が関数 $f(x)$ の<u>極小点</u>（local minimum）であるとは，

$$\exists r > 0, \qquad 0 < |x-a| < r \quad \Longrightarrow \quad f(x) > f(a) \tag{5.3.2}$$

であることをいう．なお，

$$\exists r > 0, \qquad |x-a| < r \quad \Longrightarrow \quad f(x) \leq f(a) \tag{5.3.3}$$

となっている時（最後の不等号に等号を許す），f は a で<u>広義の極大</u>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，関数 $f(x)$ が $x = a$ で <u>最大</u>（maximum）とは，f の <u>定義域全体</u> を見渡した時に $f(a)$ が最大であることをいう．つまり，

$$f \text{ の定義域に入っているすべての } x \text{ に対して} \quad f(x) < f(a) \tag{5.3.4}$$

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に大変なことが多く，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いといけない.）この節では極大・極小問題に話を限る．より複雑な最大・最小問題については，時間があれば秋学期に（多変数関数の場合として）触れる．

さて，１変数の場合の極大，極小問題は以下のようになっている．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 5.3.6 $x=a$ の近傍で定義された１変数の関数 $f(x)$ 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）．

（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.）

5.3.3 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．より詳しくは教科書の pp.150-154 を参照.

用語についての注意： 英語では下に凸の関数を単に convex function（直訳：凸関数）といい，上に凸の関数を concave function（直訳：凹関数）とよぶ．日本人にとっては不幸なことに，関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.

5.4 テイラーの定理とテイラー展開

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，皆さん苦労するようだから，少し時間をかけることにした.

「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,

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{5.4.1}$$

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関数の例では（上で $a = 0$ としたものを書いた）

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}\, x^n \tag{5.4.2}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\, x^{2n+1} \tag{5.4.3}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\, x^{2n} \tag{5.4.4}$$

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関数が，上のような変な級数（和）で書けるというのだ．物事を深く考えるひとほど，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 の周期関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，(5.4.2) 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].
- (5.4.1) 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．まずはおおもとの「テイラーの定理」から始めよう.

5.4.1 テイラーの公式

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：

[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 (5.4.2) 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，計算の原理は同じである

定理 5.4.1 (通常のテイラーの公式) $f(x)$ 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：

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}(\xi)}{n!}(x-a)^n \tag{5.4.5}$$

なお，(5.4.5) の 2 つの項に名前をつけて

$$f(x) = S_n(x) + R_n(x), \tag{5.4.6}$$

$$S_n(x) := f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad R_n(x) := \frac{f^{(n)}(\xi)}{n!}(x-a)^n \tag{5.4.7}$$

と書く事もある．この際，$S_n(x)$ 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 次の剰余項という．

- $a = 0$ 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，ξ は a と x（またはb）の両方に依存しう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 ξ 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 5.4.1 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 5.4.6（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 ξ が存在して」とか言われても，どんな ξ かわからなければ細かい評価はできない）．

定理 5.4.1 の証明[35]

$$F(x) := f(x) - \left[f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k \right], \qquad G(x) := (x-a)^n \tag{5.4.8}$$

とおく．$F(x)$ が (5.4.6) 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 5.2.3）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

$$F(a) = F'(a) = F''(a) = \ldots = F^{(n-1)}(a) = 0, \qquad F^{(n)}(a) = f^{(n)}(a) \tag{5.4.9}$$

$$G(a) = G'(a) = G''(a) = \ldots = G^{(n-1)}(a) = 0, \qquad G^{(n)}(a) = n! \tag{5.4.10}$$

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 5.2.3 そのもので

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(\xi_1)}{G'(\xi_1)} \tag{5.4.11}$$

を満たす ξ_1 の存在（ξ_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 5.2.3 が使える．その結果,

$$\frac{F'(\xi_1)}{G'(\xi_1)} = \frac{F'(\xi_1) - F'(a)}{G'(\xi_1) - G'(a)} = \frac{F''(\xi_2)}{G''(\xi_2)} \tag{5.4.12}$$

[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つたびに「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．今年も考えたけど，やっぱりダメだった．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_2 の存在（ξ_2 は a と ξ_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 $k \le n-1$ である限りくりかえす事ができて，

$$\frac{F^{(k)}(\xi_k)}{G^{(k)}(\xi_k)} = \frac{F^{(k)}(\xi_k) - F^{(k)}(a)}{G^{(k)}(\xi_k) - G^{(k)}(a)} = \frac{F^{(k+1)}(\xi_{k+1})}{G^{(k+1)}(\xi_{k+1})} \tag{5.4.13}$$

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

（4）以上をまとめると，

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F^{(n)}(\xi_n)}{G^{(n)}(\xi_n)} \tag{5.4.14}$$

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

$$\frac{F(x)}{(x-a)^n} = \frac{f^{(n)}(\xi_n)}{n!} \tag{5.4.15}$$

となっているので，分母を払うと定理が得られる． □

5.4.2 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 $S_n(x)$ 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，(5.4.6) 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．

この事情を明確にするため,「関数を近似する」とはどういう事かをはっきりさせよう．

定義 5.4.2 (n 次より高く近似) $x = 0$ の近くで定義された関数 $f(x), g(x)$ があり，

$$x \to 0 \text{ のときに} \qquad \lim_{x \to 0} \frac{f(x) - g(x)}{x^n} = 0 \qquad (n \text{ は正の整数}) \tag{5.4.16}$$

となるとき，0 の近くで <u>$g(x)$ は $f(x)$ を n 次より高く（n 次よりも良く）近似する</u> 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，と言っているのである．

ここで少し，後々便利な書き方を導入しておこう[36]

定義 5.4.3 (無限小の比較；オーダー) $\lim_{x \to a} f(x) = \lim_{x \to a} h(x) = 0$ とする．

ア. $\lim_{x \to a} \dfrac{f(x)}{h(x)} = 0$ の時，<u>$f(x)$ は $h(x)$ より高位の無限小</u> であると言い，$f(x) = o\big(h(x)\big)$ と書く（ここの o は小文字）．

イ. 上よりもう少し弱く，$\dfrac{f(x)}{h(x)}$ が $x \to a$ で有界であるとき，つまり，

$$\exists K > 0 \quad \exists \delta > 0 \qquad 0 < |x - a| < \delta \quad \Longrightarrow \quad \left| \frac{f(x)}{h(x)} \right| < K \tag{5.4.17}$$

のとき，<u>$f(x)$ は $h(x)$ のオーダーである</u> といい，$f(x) = O\big(h(x)\big)$ と書く（ここの O は大文字）．

（注）

- アとイは大文字と小文字だけの区別なので，特に手書きの際には注意が必要だ．

[36] この内容は別に小節を設けても良いくらいなのだが，話の流れを切らないために，必要最小限だけを書くことにした

- また，これらのオーダー比較はどのような極限を考えているのか（x がどこに近づいた時のものか）に当然，依存する．通常は文脈でわかるけども，どんな極限を考えているかはいつも意識すること．
- 上のイは当然アの場合を含み，実際には $f(x)$ が $g(x)$ よりずっと速くゼロに行く場合でも，$f(x)$ は $g(x)$ のオーダーである，という．この点，極限を計算する場合に注意を要する．
- では $f(x)$ は少なくとも $g(x)$ と同じくらいか大きい，という場合に使う記号はないのだろうか？ない訳ではないのだが，それほどポピュラーではない．分野によっては $f(x) \approx g(x)$ と書いたり，$f(x) = \Omega\big(g(x)\big)$ と書いたりすることはある．

この書き方によると，(5.4.16) は

$$f(x) - g(x) = o(x^n) \qquad \text{（ここの } o \text{ は小文字）} \tag{5.4.18}$$

と書ける．

この用語法に従うと，テイラーの定理は以下のように書き直せる． $\boxed{\text{テイラーの定理そのものは書いたか？}}$

命題 5.4.4 (テイラーの定理の言い換え) 関数 $f(x)$ の $x = 0$ と中心とした n 階のテイラーの公式

$$f(x) = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k, \quad R_n(x) := \frac{f^{(n)}(\theta x)}{n!} x^n \quad (0 < \theta < 1) \tag{5.4.19}$$

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

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + o(x^{n-1}) \tag{5.4.20}$$

となるための必要充分条件は以下の通り：

$$\lim_{x \to 0} \frac{R_n(x)}{x^{n-1}} = 0 \tag{5.4.21}$$

前の命題の (5.4.21) の十分条件として，以下がある．

命題 5.4.5 (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，

$$|x| < \delta \text{ ならば } \quad \left|f^{(n)}(x)\right| < M \tag{5.4.22}$$

となっているとする．このとき，

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + O(x^n) \tag{5.4.23}$$

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，一年生で出てくる関数は 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 5.4.4 の結論も成り立つのである．

5.4.3 テイラーの公式，テイラー展開の例

ともかく具体例をやらないと意味がないだろうから，やってみよう．

- まず，多項式．$f(x) = c_n(x-a)^n + c_{n-1}(x-a)^{n-1} + \ldots + c_1(x-a) + c_0$ 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．

- 指数関数．$f(x) = e^x$ は何回でも微分可能で，高階の導関数もすべて e^x である．従って，特に $a = 0$ としたテイラーの公式から

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x), \qquad R_n(x) := \frac{1}{(n-1)!} \int_0^x e^y\,(x-y)^{n-1} dy \tag{5.4.24}$$

が得られる．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 $\lim_{n\to\infty} R_n(x) = 0$ が証明できる．従って，対数関数の時と同じ議論により，すべての実数 x に対して

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{5.4.25}$$

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，

$$\sin x = \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \frac{(-1)^n}{(2n-1)!} \int_0^x (\sin y)\,(x-y)^{2n-1}\,dy \tag{5.4.26}$$

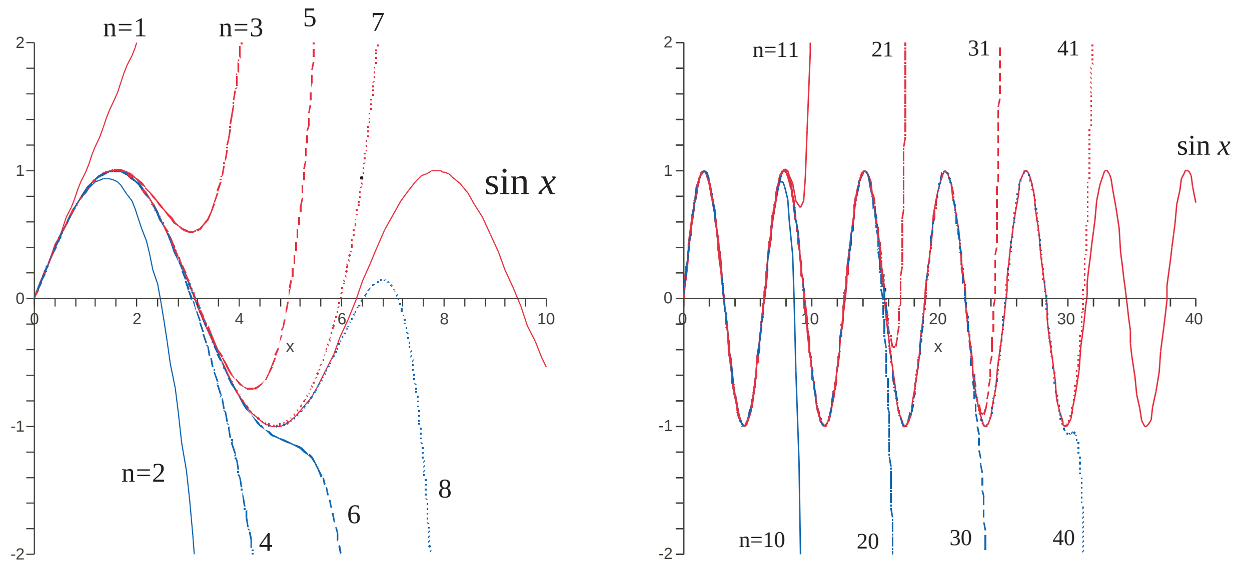
指数関数と同様に，この場合もすべての実数 x に対して $\lim_{n\to\infty} R_n(x) = 0$ が証明できる．従って，すべての実数 x に対して

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad \text{また同様の考察により} \qquad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{5.4.27}$$

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おくくらいになろう[37]．

$\sin x$ のテイラー展開の図を載せておく．下の左図は，$n = 1, 2, \ldots, 8$ 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．



まず,「オーダー」の威力を理解する為に：

[37] ただし，このような公式は無理に丸暗記してもダメだ．自分で導出したり，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

例：$x \to 0$ を考える．$f(x) = 1 + x + x^2 + o(x^2)$，$g(x) = 3 + x + o(x)$ の時，その積 $f(x)g(x)$ と商 $f(x)/g(x)$ を出来るだけ高次まで多項式で近似せよ.（今は $x = 0$ の近傍のみ考えるから，商を考える際にも $g(x) \neq 0$ と思ってよい.）

このような例をいくつか出す．その後，教科書の $\tan x$ の展開に触れる．

補足問題：以下の関数の $x = 0$ におけるテイラー展開（マクローリン展開）の最初の４項を求めよ（a は定数）．

$$a)\ \ \log(1+3x) \qquad b)\ \ e^{ax} \qquad c)\ \ \log(1+x^2) \qquad d)\ \ \tan x \qquad e)\ \ \cosh x$$

このうち，a), b), e) は定義通りにやればできる．c), d) については，少し工夫した方が簡単だが，定義通りに地道にやっても可能だ.

前回の補足から始めよう.（前回，強調すべきだったのだが）級数展開の話では，以下の２つを区別する必要がある.

（ア）級数展開を有限項で止めた，余りの項（剰余項という）つきの表式（n が有限）

（イ）n を無限大にとばした，無限級数（無限項の和）の形で書かれた表式

対数関数の例でいうと，

$$\log(1+x) = \sum_{k=0}^{n} \frac{(-1)^k x^{k+1}}{k+1} + \int_0^x \frac{(-y)^{n+1}}{1+y} dy \tag{5.4.28}$$

というのが（ア）であり，ここで形式的に $n \to \infty$ とやって（かつ剰余項がゼロに行くことを確かめて）得た

$$\log(1+x) = \lim_{n\to\infty} \left[\sum_{k=0}^{n} \frac{(-1)^k x^{k+1}}{k+1} \right] = \sum_{k=0}^{\infty} \frac{(-1)^k x^{k+1}}{k+1} \tag{5.4.29}$$

が（イ）である（最右辺は単に真ん中の極限の略記である）．これらについて，少し補足しておく.

（ア）について

対数関数の，有限項までで止めた展開（ア）の導き方を整理する：

1. まず，有限の n に対して，
$$\frac{1}{1+y} = \sum_{k=0}^{n} (-y)^k + \frac{(-y)^{n+1}}{1+y} \tag{5.4.30}$$
を出した．これは単なる多項式の展開計算（の後で $1+y$ で割った）だった．ここの y は（割り算ができるように）$y \neq -1$ ならば何でもよい.

2. 次に，この両辺を 0 から x まで積分した．上の右辺はあくまで n が有限のままだから，和と積分を交換してもよい．結果は上に書いた (5.4.28) である．この式での x には（非積分関数 $\frac{1}{1+y}$ の分母がゼロにならない条件から）$x > -1$ の制限がつく.

3. ここでわかりやすいように，
$$S_{n+1}(x) := \sum_{k=0}^{n} \frac{(-1)^k x^{k+1}}{k+1}, \qquad R_{n+1}(x) := \int_0^x \frac{(-y)^{n+1}}{1+y} dy \tag{5.4.31}$$
とおいてみよう．対数関数の展開 (5.4.28) は
$$\log(1+x) = S_{n+1}(x) + R_{n+1}(x) \qquad (n = 0, 1, 2, \ldots) \tag{5.4.32}$$
となる.

ここまではおかしなことは何もしてないし，高校の数学で十分に理解できる．極限の概念も（「積分」は本来，ある種の極限で定義されるべきものであるということ以外は）入っていない．ただし，若干の注意が適当だろう.

- 些細なことではあるが，教科書に合わせて級数展開の k の入り方を，前回のプリントと変えた（前回は和が $k=1$ から始まったが，今回は $k=0$ から）．
- S_{n+1}, R_{n+1} の添字が $n+1$ になっているのは，S_{n+1} が $n+1$ 項の和になっていることを表すつもりである．R_{n+1} の添字は，S_{n+1} におつきあいした．これは教科書の p.1 とは異なる —— この項は教科書では R_n になっている —— が，教科書 1.3 節の「テイラー展開」とは一致した書き方である．(要するに，教科書の p.1 と 1.3 節の書き方が合ってない；それでこれからはより一般の 1.3 節の書き方に合わせる，というわけ）

（イ）について

これから，(5.4.32) において，$n \to \infty$ を考える．ちゃんと考えると，ここはちょっと微妙なのだ．以下に説明しよう．(以下では x を一つ固定 して，n だけを無限大にすることを考える.）

ここのところ，普通は以下の（直感的にもわかるし，後でちゃんとやる定理）

数列 a_n, b_n について，a_n, b_n 両方の $n \to \infty$ 極限が存在するならば $\displaystyle\lim_{n\to\infty}\bigl[a_n + b_n\bigr] = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$

を，$a_n = S_{n+1}(x), b_n = R_{n+1}(x)$ として用いたくなるだろう．ところが，今の場合，$-1 < x \le 1$ ならば $\lim_{n\to\infty} R_{n+1}(x) = 0$ であるが（先週の講義と教科書 p.1 の中程を参照），$\lim_{n\to\infty} S_{n+1}(x)$ が存在するかどうかについては我々は何の判定基準も持っていない．実際，この極限が存在することを示すには，5〜6 月に学ぶ「極限の厳密化」,「実数概念の精密化」,「コーシー列の概念」が不可欠だ．

そこで 4 月の時点では困った，となるのであるが，もちろん，本当には困らない．というのは，今は（$C := \log(1+x)$ は n によらない定数）

$$C = a_n + b_n, \quad つまり, \quad a_n = C - b_n \tag{5.4.33}$$

がすべての $n \ge 1$ で成立している．従って，この場合は a_n または b_n のどちらかの極限が存在することだけ確かめれば，もう片方の極限が存在することは自動的に出るわけだ．特にどちらかの極限が存在すれば，

$$C = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n \tag{5.4.34}$$

が成立する．対数関数の展開では $-1 < x \le 1$ に対して，$b_n = R_{n+1}(x)$ がゼロに行くことは確かめた．従って，

$$\log(1+x) = \lim_{n\to\infty} S_{n+1}(x) + \lim_{n\to\infty} R_{n+1}(x) = \lim_{n\to\infty} S_{n+1}(x) + 0 = \lim_{n\to\infty}\left[\sum_{k=0}^{n} \frac{(-1)^k x^{k+1}}{k+1}\right] \tag{5.4.35}$$

を得る．

以上，長々と説明したが，これが教科書 p.1 の (1) 式の正しい導出である．これからしばらくの間は，$n \to \infty$ の極限はあまりとらないが，もしとった場合はこのような議論を繰り返しているのだと思ってほしい．

（注）$n \to \infty$ の極限をとるところで，$-1 < x \le 1$ の制限がついた．これは $\lim_{n\to\infty} R_{n+1}(x) = 0$ のためには必要だったからである．それに対して，有限の n で止めた式 (5.4.32) は $x > 1$ でも成り立っている．各自，たとえば $x = 2$ の時に，$S_{n+1}(x), R_{n+1}(x)$ がどのようにふるまうか，$n = 1, 2, 3, 4$ くらいで確かめて実感してほしい．

5.4.4 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，重要なのでもう一度繰り返す．

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．対数関数のときに少しやったように，??? 剰余項を何らかの工夫で押さえれば，問題の関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たので，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関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，関数の定義域を一気に拡げることも可能である．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．この視点は秋以降（また 2 年時の「複素関数論」で）たくさんやるだろう．

少し進んだ話題． 少し先走るが，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

$$e^{i\theta} = \cos\theta + i\sin\theta, \qquad \theta \in \mathbb{R} \tag{5.4.36}$$

を挙げておこう．指数関数のテイラー展開において，$x = i\theta$ とおいてしまおう（このようにおいてもテイラー展開が収束することは確かめられる）．すると，

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell+1}}{(2\ell+1)!} \tag{5.4.37}$$

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，指数関数や三角関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．これについては６月頃にまた触れるつもりだが，根性のある人は「$\sin x$ のテイラー展開から $\sin x$ が周期関数と言えるだろうか」「指数関数のテイラー展開から指数法則は出せるだろうか」などを考えてみて欲しい．

5.4.5 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４年生くらいにな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 5.4.6 (剰余項が積分形のテイラー（Taylor）の公式) $f(x)$ がある開区間 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：

$$f(x) = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad R_n(x) := \int_a^x \frac{f^{(n)}(y)}{(n-1)!}(x-y)^{n-1}\,dy \tag{5.4.38}$$

（注）上の公式での 剰余項 $R_n(x)$ が $n \to \infty$ でゼロになるならば，つまり，$\lim_{n\to\infty} R_n(x) = 0$ ならば，対数関数の級数展開と同じ注意により，

$$f(x) = \lim_{n\to\infty} S_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{5.4.39}$$

が得られる．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，本当に $n \to \infty$ として良いかどうかは，展開される関数 f と考えている区間 I によるので，個別に考察する必要がある．実際，対数関数の場合，$-1 < x \le 1$ ならば $n \to \infty$ として良いが，$|x| > 1$ ではダメだった．(そもそも，級数の各項が発散してしまう！)

（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，(5.4.38) をすべての $n \le N$ 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．

I′. $n = 2$ 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

$$f(x) = f(a) + \int_a^x f'(y)dy \tag{5.4.40}$$

の第２項を，以下のように部分積分するとよい．

$$\int_a^x f'(y)dy = \int_a^x \Bigl\{-\frac{d}{dy}(x-y)\Bigr\} f'(y)dy = \Bigl[-(x-y)f'(y)\Bigr]_a^x + \int_a^x (x-y)f''(y)dy$$
$$= (x-a)f'(a) + \int_a^x (x-y)f''(y)dy \tag{5.4.41}$$

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，$n \le N-1$ と仮定しておく）．n までできたと仮定したので，(5.4.38) 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）：

$$\int_a^x f^{(n)}(y)(x-y)^{n-1}dy = \int_a^x f^{(n)}(y)\Bigl\{-\frac{1}{n}\frac{d}{dy}(x-y)^n\Bigr\}dy$$
$$= -\frac{1}{n}\Bigl[f^{(n)}(y)\,(x-y)^n\Bigr]_a^x + \frac{1}{n}\int_a^x f^{(n+1)}(y)\,(x-y)^n dy$$
$$= \frac{1}{n}f^{(n)}(a)\,(x-a)^n + \frac{1}{n}\int_a^x f^{(n+1)}(y)\,(x-y)^n dy. \tag{5.4.42}$$

これを (5.4.38) の最後の項に用いると（もちろん，分母の $(n-1)!$ を忘れない），(5.4.38) の $n+1$ のものが証明されてメデタシメデタシ． □

5.5 対数関数の級数展開

この節の内容は教科書の 1.1 節そのものなので，プリントでは簡単にすませる．結果だけ書くと，$-1 < x \le 1$ の時に

$$\log(1+x) = \lim_{n\to\infty}\Biggl[\sum_{k=1}^{n}\frac{(-1)^{k-1}x^k}{k}\Biggr] = \sum_{k=1}^{\infty}\frac{(-1)^{k-1}x^k}{k} \tag{5.5.1}$$

である．これは「テイラー展開」の一種だが，テイラー展開の詳しい定義は来週にやるので，それまでは単に「展開」と呼ぶ．

ここで大事なのはこのような展開式だけでなく，これが何の役に立つのかということだ．答えは２通りある．

- log などの「よくわかった」関数に対しては，これを用いていろいろな関数の値を計算できる．対数関数に対する例が教科書の 1.1.2 節に載っている．昨今，コンピューターの発達により，log の値は瞬時に計算できるようになった．そのため,「どのようにしてその値を出したのか」が盲点になっているように思う．我々（とコンピューター）にできることは主に四則演算に限られていることに注意しよう．その限られた四則演算から関数の値を求めるには，このような展開（とその仲間）が必須なのだ．実際，コンピューターの中では，教科書に説明されているようなこと（を更に改良して）四則演算から log（や他の特殊関数）の値を計算している．
- 更に立場を逆転して，上のような級数を関数の定義にしてしまうことも可能だ．物理や工学の問題では，対象としている関数が級数の形でしか与えられないことも往々にしてあるので，この発想の転換は重要である．

２番目の考え方や例にはしばらくは出会わないだろうが，後期に行う「級数論」でその一端を見ることになるだろう．

この級数の収束の様子を図に示す．左側は $\log(1+x)$ そのものと (5.5.1) の第 n 項までの部分和のグラフである．n を上げていくと近似が良くなっているのが見える．右はこの級数を教科書の２ページのように改良したもの．つまり，

$$1+x = \frac{1+y}{1-y}, \quad \text{つまり} \quad y = \frac{x}{2+x} \tag{5.5.2}$$

とおいた上で，

$$\log(1+x) = \log\Bigl(\frac{1+y}{1-y}\Bigr) = 2\bigl(y + \frac{y^3}{3} + \frac{y^5}{5} + \dots\bigr) \tag{5.5.3}$$

としたものである．図からも，こちらの級数は非常に良い近似になっていることがわかる．

———夏休みチャレンジ問題———

問１３で出題した問題のより詳しいバージョンです.（今まで高校で習った sin, cos の定義や性質はいったん忘れて）$\sin x$ と $\cos x$ をそのテイラー級数で定義した場合，我々の知っている sin, cos と同じになるかどうかを考える問題．

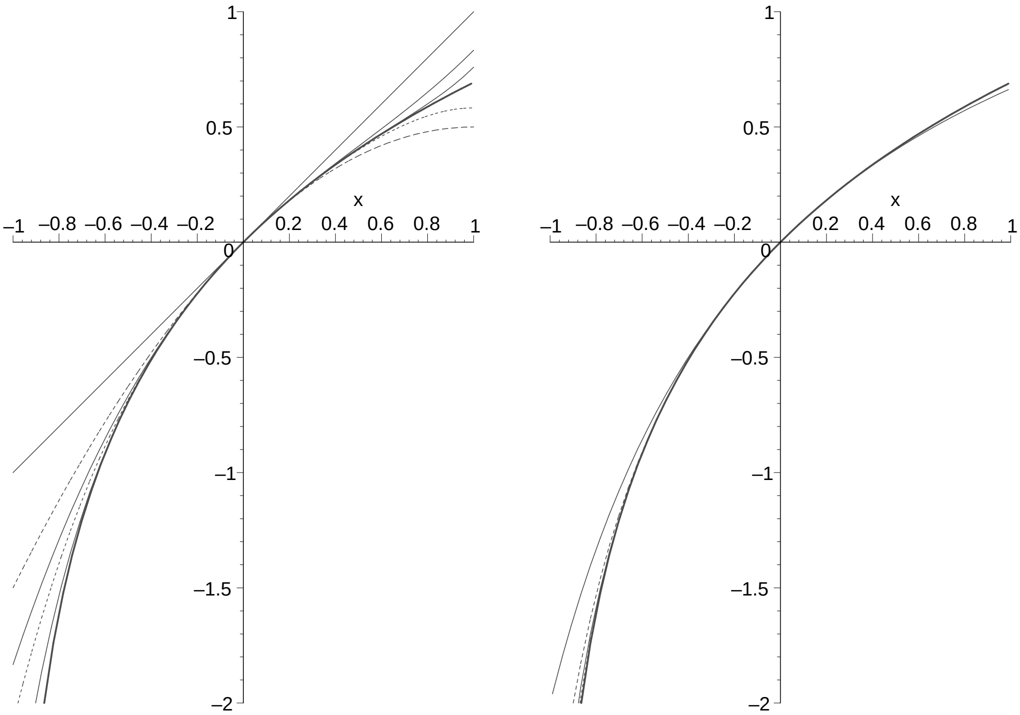


図 1: （左）$\log(1+x)$ の展開を第 n 項までとったものと $\log(1+x)$ の比較．太い実線が $\log(1+x)$，細い実線は $n=1,3,7$，点線が $n=2,4$．$\log(1+x)$ から遠いものほど n が大きい．
（右）(5.5.3) の級数を第 n 項までとったもの（$n=1,2,3,4$）と $\log(1+x)$ の比較．$n \geq 3$ では $\log(1+x)$ とほとんど重なって区別がつかない（近似が良い）．

いろいろなアプローチがあろうが，この前には一人しか解いてくれなかったので,（でもその一人は非常に頑張ってました）以下のようにある程度の道筋を示しておく（順番は割合とええ加減）．我こそはと思う者は，是非，チャレンジしてほしい．

- ともかく，$\sin x, \cos x$ をそのテイラー級数で定義する．
- （お約束）このテイラー級数はすべての実数 x で収束していることを確かめよう．
- このように定義したものが $(\sin x)' = \cos x, (\cos x)' = -\sin x$ をみたすこと（$'$ は x での微分）を示せ．この場合，$\sin, \cos$ ともに無限級数で定義されているから，級数の無限和と x による微分の順序交換ができるかどうかが問題だ．一般論は後期にやるが，今までの知識でも（微分の定義から出発して）できるから，頑張れ！
- $\sin^2 x + \cos^2 x = 1$ を示せ．
- 加法定理が成り立つことを示せ.（無限級数の積を考える必要があるから，極限の取り方などに注意して厳密に議論しよう.）
- $\sin\alpha = 0, \cos\beta = 0$ となるようなゼロでない α, β が存在することを何とかして示せ．
- $\sin, \cos$ が周期関数であることを導け．
- 上の α, β と，円周率 π の関係をつけよう．π の定義としては「半径 1 の円の円周の長さが 2π」とする．これまで考えてきた $\sin, \cos$ からその円周の長さを計算することで，π との関係をつけよ．
- その他，わかることがあったら何でもやってみよう．

以上，時間不足で思いつくままに書いたから，もしかしたらミスが混じってるかもしれない．その時は申し訳ないけど，自力で修正してくだされ．また，上のはあくまで一つのアプローチであって，これにとらわれる必要はない．

5.6 平均値の定理

多変数の続き，秋かな？ 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．

定理 5.6.1 (多変数の平均値の定理) 2変数関数 $f(x,y)$ が C^1-級の場合，

$$f(a+h,b+k)=f(a,b)+\frac{\partial f}{\partial x}(a+\theta h,b+\theta k)h+\frac{\partial f}{\partial y}(a+\theta h,b+\theta k)k \tag{5.6.1}$$

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 変数関数に対しては（$\boldsymbol{a}=(a_1,a_2,\ldots,a_n)$，$\boldsymbol{h}=(h_1,h_2,\ldots,h_n)$）

$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+\sum_{j=1}\frac{\partial f}{\partial x_j}(\boldsymbol{a}+\theta\boldsymbol{h})\,h_j \tag{5.6.2}$$

が成り立つ（$0<\theta<1$）．

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

$$f(a+h,b+k)-f(a,b)=g(1)-g(0)=g'(\theta) \tag{5.6.3}$$

である．ところが，g' については，定理 1.3.3 を $x(t)=a+th, y(t)=b+yk$ として用いると

$$g'(t)=\frac{\partial f}{\partial x}x'(t)+\frac{\partial f}{\partial y}y'(t)=\frac{\partial f}{\partial x}h+\frac{\partial f}{\partial y}k \tag{5.6.4}$$

であるから，定理 5.6.1 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，関数が 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．

定理 5.6.2 (平均値の定理) 2変数関数 $f(x,y)$ が全微分可能なら，

$$f(a+h,b+k)=f(a,b)+\frac{\partial f}{\partial x}(a+\theta h,b+\theta k)h+\frac{\partial f}{\partial y}(a+\theta h,b+\theta k)k \tag{5.6.5}$$

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 5.6.1 が定理 1.3.3 から出るのと同じようにして，定理 5.6.2 は定理 1.3.8 から証明される． □

2006.07.31

期末テストなどに関する注意：

- この期末試験はできるだけ早く採点の後，8/3（木曜）の午後５時から返却の予定です．また，メイルでの成績の問い合わせには，そのメイルに返信する形で応じます．自分の名前と学生番号を書いた上で「成績を知りたい」旨のメイルを送って下さい（hara@math.kyushu-u.ac.jp）．
- 学力再調査は一週間後の 8/7 の 3 限に行う予定です（場所は未定）．学力再調査を受験する資格のある人のリストは８月３日（木曜）午後５時までに僕の部屋（六本松３号館３階，3-312 室）の前に張り出します．再調査の場所もそのときに掲示します.

秋学期の最初「新学期の準備は万全かな？」テストについて：

- 今学期の皆さんの出来（特に期末の結果）を基にして，皆さんの弱点を突く形での小テストを秋学期の最初の頃にやる可能性があります.
- 夏休み，大いに遊んで下さって結構ですが，今学期の弱点はきちんと押さえておいて下さい.
- どの辺りが弱点か，については 8/3 に返却予定の期末テストを参考にして下さい.
- ただし，皆さんに共通のヒドイ弱点が発見された場合には，その弱点を 8/3 に掲示して，注意を促すかもしれません.

実数論に関する補足：

新居さんのプロジェクトでもあったようだけど，実数論に関する僕の講義ノート（暫定版）を web page においてあります（このページの上にある URL のところに行き，そこからこの講義（微積 A）のページに行くとわかると思う）．このノートでは切断による構成の他にコーシー列の同値類としての構成などにも触れています．実のところ，まだ書き足す予定で，最終版は 8/7 頃にならないと出来ない予定ですが，興味のある人はその頃にでも見て下さい.

夏休みチャレンジ問題

これは去年の一年生にも出した問題です．テイラー展開はまだやってはいませんが,「そのような変なものがある」また「それを基にして関数が定義できる」ことを知るには良い例になるでしょう.（見通しを良くするために「テイラー展開」という用語は出しましたが，これをまったく知らなくても今学期に習った事だけでできます.）

（問題の趣旨）今まで高校で習った $\sin, \cos$ の定義や性質はいったん忘れ，$\sin x$ と $\cos x$ をそのテイラー級数で定義した場合，我々の知っている $\sin, \cos$ と同じになるかどうかを考える問題です．何をもって「同じ」というかはなかなか難しいのですが，一見異なる性質のものがどうも同じらしい，と実感してもらえれば良いですね.

我こそはと思う者は，是非，チャレンジしてほしい.

- まず，関数 $S(x)$ と $C(x)$ を
$$S(x) := \sum_{n=0}^{\infty} \frac{(-1)^n\, x^{2n+1}}{(2n+1)!}, \qquad C(x) := \sum_{n=0}^{\infty} \frac{(-1)^n\, x^{2n}}{(2n)!}$$
と定義する．右辺の意味はもちろん，$S(x)$ なら $a_N := \sum_{n=0}^{N} \frac{(-1)^n\, x^{2n+1}}{(2n+1)!}$ という数列を考え，その $N \to \infty$ の極限を $S(x)$ と定義する，ということ．先走っていうと，$S(x)$ は $\sin x$，$C(x)$ は $\cos x$ になるはずのものである.
- （お約束）この級数はすべての実数 x で収束していることを確かめよう.
- このように定義したものが $S'(x)' = C(x), C'(x) = -S(x)$ をみたすこと（$'$ は x での微分）を示せ．この場合，$\sin, \cos$ ともに無限級数で定義されているから，級数の無限和と x による微分の順序交換ができるかどうかが問題だ．一般論は後期にやるが，今までの知識でも（微分の定義から出発して）できるから，頑張れ！
- $\{S(x)\}^2 + \{C(x)\}^2 = 1$ を示せ.
- $\sin, \cos$ の加法定理に相当するものが成り立つことを示せ．例えば，$S(x+y) = S(x)C(y) + C(x)S(y)$ などということである.（無限級数の積を考える必要があるから，極限の取り方などに注意して厳密に議論しよう.）
- $S(\alpha) = 0, C(\beta) = 0$ となるようなゼロでない α, β が存在することを何とかして示せ．$S(x), C(x)$ の満たす不等式を作ってみるのが良いかもしれない.
- $S(x), C(x)$ が周期関数であることを導け．つまり，正の数 α があって，$S(x+\alpha) = S(x)$, $C(x+\alpha) = C(x)$ がすべての x について成り立つことを示せ.
- 上の α, β と，円周率 π の関係をつけよう．π の定義としては「半径 1 の円の円周の長さが 2π」とする．これまで考えてきた $S(x), C(x)$ からその円周の長さを計算することで，π との関係をつけよ.
- その他，わかることがあったら何でもやってみて，上の $S(x), C(x)$ と皆さんの知っているはずの $\sin x, \cos x$ が「同じ」である状況証拠を積み重ねてみよう．あと，どのようなことを言えば，本当に「同じ」だといえるだろうか？そもそも，皆さんの知っているはずの $\sin x, \cos x$ の定義は何だったんだろう？

上に書いたものはこの順序でやる必要はない．また，上のはあくまで一つのアプローチであって，これにとらわれる必要もない．ともかく上で定義した級数 $S(x), C(x)$ と皆さんの知っているはずの $\sin x, \cos x$ との関係をいろいろとつけてくれれば良い.