1 積分（つづき）

1.1 広義積分

いままで，定積分としては有限区間 $[a,b]$ 上での関数 $f(x)$ の積分のみを考えてきた．この際，関数 $f(x)$ は区間 $[a,b]$ 上で定義できている（当然，その値は有限）ことが暗黙の前提であった．しかし，実際の応用では上の２条件が守られていない積分を考えたくなることは多い．例えば，

- $\displaystyle\int_{-1}^{1} \frac{1}{\sqrt{|x|}}dx$ これは積分区間は有限だが，被積分関数が（$x=0$ で）無限大になる例である．
- $\displaystyle\int_{0}^{\infty} e^{-x}dx$ これは被積分関数は有界だが，積分区間が無限大になっている例である．

もちろん，この２つが両方起こっているもの（積分区間も無限だし，被積分関数も有界でない；例えば $\int_{-\infty}^{\infty} e^{x}dx$）もありうる．これらの問題に共通しているのは，積分区間や被積分関数に無限大の芽が含まれており，春学期のリーマン積分の定義をそのままの形では適用できないということだ.（適用した場合，答えが「無限大」などになってしまうが，これは我々の欲しい積分の値としてはかえって不自然.）

この節では，このような問題を考えていく．解決法は単純だ：無限大の芽が隠れていそうな積分は，いつも「きちんと有限に定義できる積分」からの**極限として定義**する．その極限が存在すればよし，存在しない場合は「この積分は存在しない」と決めるのである．このように極限として定義するのが，物理や工学への応用上でも自然なのである．

（ことばについて）この節の内容で定義される積分を <u>広義積分</u>（improper integral）と呼ぶ．日本語の方はそのまま「積分の定義を拡張したもの」のつもりであろう．英語の方は正しい定義には含まれていない，というつもりだろうか.

1.1.1 有界区間上の積分だが，被積分関数が有界でない場合の広義積分

上で書いたように，ヤバいところをまず避けて積分を定義し，後でヤバいところまで積分区間を拡張する.

定義 1.1.1 $a<b$ とする.
(0) $f(x)$ が区間 $[a,b]$ において有界なら，今までのリーマン積分の定義により $\int_a^b f(x)dx$ を定義する.
(i) $f(x)$ が半開区間 $(a,b]$ で定義されていて

$$\lim_{c\to a+0}\int_c^b f(x)dx \tag{1.1.1}$$

が存在するとき（当然，各 c に対する $\int_c^b f(x)dx$ の存在は仮定している），$f(x)$ は $[a,b]$ で <u>広義積分可能</u>（または，広義積分が収束する）といい，その極限を広義積分 $\int_a^b f(x)dx$ の値と定める.
(ii) $f(x)$ が半開区間 $[a,b)$ で定義されていて

$$\lim_{d\to b-0}\int_a^d f(x)dx \tag{1.1.2}$$

が存在するときも $f(x)$ は $[a,b]$ で <u>広義積分可能</u> といい，その極限を広義積分 $\int_a^b f(x)dx$ の値と定める.
(iii) 最後に，$f(x)$ が開区間 (a,b) で定義されていて

$$\lim_{\substack{c\to a+0\\ d\to b-0}}\int_c^d f(x)dx \tag{1.1.3}$$

が存在するとき，$f(x)$ は $[a,b]$ で <u>広義積分可能</u> といい，その極限を広義積分 $\int_a^b f(x)dx$ の値と定める．ただしこ

こで c, d の極限は 互いに独立に a, b へ近づけるすべての近づけ方についてとる．
(iv) 最後に，上の (i), (ii), (iii) のそれぞれの極限が存在しない場合，その広義積分は存在しないという．

なお，上のようにして定義した広義積分は，特に断らずに $\int_a^b f(x)dx$ と書く事がある．つまり，$\int_a^b f(x)dx$ が通常のリーマン積分の定義で解釈できない時は，広義積分によって定義すると拡大解釈する場合があるので要注意.（きちんと「積分は広義積分の意味で考える」と書いてくれることもあるが，広義積分を考える事がほとんど自明な場合は省かれる事が多い.）

（注１）最後の (iii) の極限の取り方について注意しておこう．ここでは $c \to a+0$ と $d \to b-0$ を，互いの近づき方を気にせずに勝手バラバラに極限をとろう，と言っている．つまり，$c \to a+0$ よりも $d \to b-0$ を先にとったり，その逆に $d \to b-0$ よりも $c \to a+0$ を先にとったり，両方の極限を大体同じ速さでとったり，といろいろやってみて，どのような取り方をしても同じ一定値に近づく場合，かつその場合に限って，この極限が存在する，と言うのである．

（注２）通常のリーマン積分の定義によって $\int_a^b f(x)dx$ が定義できる場合に，敢えて上の (i) や (ii) の極限として $\int_a^b f(x)dx$ を定義すると，その結果は通常のリーマン積分による定義に一致する（各自，確かめよ）．この意味で，上の定義は，確かに通常の積分の定義の拡張になっている．

このようなものは変に覚えないで，具体例をやって自然に身につけるのが良い．ということで，レポート問題を出題の予定．

定義 1.1.1 では区間 $[a, b]$ の端に変態な（例えば f が有界でなくなる）点がある場合を考えた．もし $[a, b]$ の内部の点 c で f が変態である場合は，

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \tag{1.1.4}$$

の公式を使う．つまり，上の右辺の２つの積分のそれぞれが定義 1.1.1 によって広義積分として定義できるとき，上の式を使って $\int_a^b f(x)dx$ を定義する．具体的に書くと，

$$\int_a^b f(x)dx = \lim_{e \to c-0} \int_a^e f(x)dx + \lim_{d \to c+0} \int_d^b f(x)dx \tag{1.1.5}$$

ということだ．この場合も e, d の極限は互いに無関係にとることに注意しよう．

（例）次の積分 $\displaystyle\int_{-1}^1 \frac{1}{x} dx$ は，上の定義に従うと

$$\int_{-1}^1 \frac{1}{x} dx = \int_{-1}^0 \frac{1}{x} dx + \int_0^1 \frac{1}{x} dx \tag{1.1.6}$$

として定義したいが，右辺の積分は２つとも定義できない（定義 1.1.1 にしたがって極限を考えても $\pm\infty$ に発散してしまう）．従って $\int_{-1}^1 \frac{1}{x} dx$ 自身も定義できない．

（補足）この例でもし，右辺の極限を同じ速さでとると，つまり

$$\lim_{\epsilon \to +0} \left[\int_{-1}^{-\epsilon} \frac{1}{x} dx + \int_\epsilon^1 \frac{1}{x} dx \right] \tag{1.1.7}$$

を考えると，括弧の中は被積分関数が奇関数だからゼロになり，従って極限値もゼロである（というふうに極限値は存在してしまう）．正しい定義（極限は別々にとる）との違いをよく認識されたい．なお，この「補足」のようにそろえて極限をとったものには，「コーシーの主値（積分）」の名前がついている．これは将来，複素積分などで出てくると思うが，問題によっては，このようにちょっと「ずるい」定義[1]が役に立つ事もある．

[1]ずるいというのは，本来収束しないものを，うまく極限をとって収束するように見せかけているから

更にたくさんの特異点がある場合も同様に考える．例えば $f(x)$ が有界でない点が $[a,b]$ 中に c_1, c_2, c_3 と 3 点ある場合（$a < c_1 < c_2 < c_3 < b$）には，

$$\int_a^b f(x)dx = \int_a^{c_1} f(x)dx + \int_{c_1}^{c_2} f(x)dx + \int_{c_2}^{c_3} f(x)dx + \int_{c_3}^b f(x)dx \tag{1.1.8}$$

の公式を使うつもりになる．そして右辺のそれぞれの積分が定義 1.1.1 にしたがって定義できるかどうかを考える訳だ．

1.1.2　無限区間上の積分だが，被積分関数が有界な場合の広義積分

典型的な例は $\int_0^\infty e^{-x}dx$ である．まあ，この時はどう進むか，予想はつくだろう．実際，高校でも少しやった事があるかもしれない．

定義 1.1.2 $f(x)$ は有界な関数とする．
(i) 半無限区間 $[a,\infty)$ 上の有界な関数 $f(x)$ に対して，極限

$$\lim_{M\to\infty} \int_a^M f(x)dx \tag{1.1.9}$$

が存在するとき，$f(x)$ は $[a,\infty)$ で <u>広義積分可能</u> といい，その極限を $\int_a^\infty f(x)dx$ の値と定める．
(ii) 同様に半無限区間 $(-\infty,b]$ 上の有界な関数 $f(x)$ に対して，

$$\lim_{L\to-\infty} \int_L^b f(x)dx \tag{1.1.10}$$

が存在するとき，$f(x)$ は $(-\infty,b]$ で <u>広義積分可能</u> といい，その極限を $\int_{-\infty}^b f(x)dx$ の値と定める．
(iii) 最後に，無限区間 $(-\infty,\infty)$ 上の有界関数 $f(x)$ に対して，２重極限

$$\lim_{\substack{L\to-\infty \\ M\to\infty}} \int_L^M f(x)dx \tag{1.1.11}$$

が存在するとき， $f(x)$ は $(-\infty,\infty)$ で（または簡単に $\mathbb{R}$ で）<u>広義積分可能</u> といい，その極限を $\int_{-\infty}^\infty f(x)dx$ の値と定める．ここで L,M の極限は<u>互いに独立に</u> $-\infty,\infty$ へ近づけるすべての近づけ方についてとる．

最後の (iii) については定義 1.1.1(iii) と同じ注意が適用される．つまり，$L\to-\infty$ と $M\to\infty$ は別々に極限をとるのだ．なお，将来，$L=-M$ としてとった極限を考える場合もある（「フーリエ変換」などで出てくるはず）．

（例） $\displaystyle\int_0^\infty e^{-x}dx = \lim_{M\to\infty}\int_0^M e^{-x}dx = \lim_{M\to\infty}(1-e^{-M}) = 1$ であるので，この広義積分の値は 1．

1.1.3　（半）無限区間上の積分で，被積分関数も有界でない場合の広義積分

まあ，これは今まで考えてきた２つの場合の組み合わせであるから，どうやって進めるかは明らかだろう．区間が無限であるためにヤバい部分と，被積分関数が無限大になるのでヤバい部分を分離して，個々に片付ければ良い．例えば，

$$\int_0^\infty \frac{\sin x}{\sqrt{x}}dx = \int_0^1 \frac{\sin x}{\sqrt{x}}dx + \int_1^\infty \frac{\sin x}{\sqrt{x}}dx = \lim_{\epsilon\to+0}\int_\epsilon^1 \frac{\sin x}{\sqrt{x}}dx + \lim_{L\to\infty}\int_1^L \frac{\sin x}{\sqrt{x}}dx \tag{1.1.12}$$

のように分解するわけだ．なお，この例では $x=1$ で積分を分けたが，$x=1$ でなくても良い．ここはすきなように正の定数 c をとって，$x=c$ で分ければ良いのである．(もちろん，答えは c にはよらない．なぜよらないかは各自で確かめよ．)

このような場合をいろいろ書き下す事にあまり意味があるとは思えないので，後は演習にまかせる．

1.2 広義積分 II（積分が計算できないときの収束の判定条件その一）

概念としての広義積分は，前節で尽きている．しかし，実際問題として，与えられた広義積分が存在するか（収束するか）否かの判定には，これまでの話では不十分だ．

例えば，

$$I_1 := \int_0^\infty \frac{\sin x}{x} dx = \lim_{L\to\infty} \int_1^L \frac{\sin x}{x} dx \tag{1.2.1}$$

を考えてみる．右辺の積分はそう簡単に計算できないから，この極限が存在するかどうかは，すぐにはわからない．類似の問題として

$$I_2 := \int_1^\infty \frac{|\sin x|}{x} dx, \qquad I_3 := \int_1^\infty \frac{|\sin x|}{x^2} dx \tag{1.2.2}$$

なども挙げておこう．こたえを先に言ってしまうと，I_2 は発散するが，I_1, I_3 は収束する（広義積分が定義できる）．この節では，これらの判定条件（多くの場合は十分条件にすぎない）の一部を考える[2]．

1.2.1 非積分関数が一定符号の場合

まず，簡単な場合として，非積分関数が一定符号 —— いつも非負，またはいつも非正 —— の場合を考えよう．（正でも負でも一緒だから，以下では非負の場合のみ考える．）このときは簡単な（必要）十分条件がある．すこし読み進むと，春学期にやった「有界単調数列の収束」と同じノリであることがわかるだろう．

命題 1.2.1 (教科書では定義 3.5.1 と定義 3.5.8 の後のノート) この命題では $f(x) \ge 0$ とする．
(1) $f(x)$ が $x \ge a$ で有界の場合，広義積分 $\displaystyle\int_a^\infty f(x)dx$ の収束性は，b の関数として定義した $S(b) := \displaystyle\int_a^b f(x)dx$ の（$b \to \infty$ での）有界性と同等である．
(2) $f(x)$ が $x = a$ 以外では有界の場合，広義積分 $\displaystyle\int_a^b f(x)dx$ の収束性は，c の関数として定義した $S(c) := \displaystyle\int_c^b f(x)dx$ の（$c \to a+0$ での）有界性と同等である．

上の命題はより一般に，$\int_{-\infty}^b f(x)dx$ や b が特異点の場合の $\int_a^b f(x)dx$ に簡単に適用されるが，いちいち断らない．

証明：
(1) 数列 $S_n := S(n)$ を考える（$n > a$）と，$f(x) \ge 0$ ゆえ，これは広義単調増加である．また，$S(b)$ が有界なので，S_n も有界である．広義単調増加な有界数列は収束するから，極限 $S_\infty := \lim_{n\to\infty} S_n$ が存在する．

でもまだ証明は終わりではない．これまでのところでは，n を整数に限定した場合の $S(n)$ の極限の存在を言ったに過ぎぬ．本来は，整数に限定されない b を無限大にした場合でも極限が存在すること（それは当然，S_∞ に一致するはず）を示す必要がある．

しかし，これは $S(b)$ が b について広義単調増加であることからすぐにいえる．実際，任意の b に対して $n \le b < n+1$ となる整数 n を見つけられて，$S_n \le S(b) \le S_{n+1}$ が成り立っている．b を無限大にすれば S_n も S_{n+1} も S_∞ に行くから，挟まれた $S(b)$ も S_∞ に収束する．（ここのところ，ϵ-δ で仰々しくやることもできますが，必要ないでしょう．）

(2) これは (1) とほとんど同じ．今度は $S_n := S\Big(a + \dfrac{1}{n}\Big)$ を考えれば良い． □

これをオーダーの概念を用いて言い換えると，以下のようになる．春学期よりも少し拡張して用いるので，定義から復習しておく．

定義 1.2.2 (教科書では定義 3.5.4 の前半)
(1) $f(x), g(x)$ が半開区間 $[a, b)$ で定義されているとする．ある数 K をとると b の近くで $|f(x)| \le K|g(x)|$ が成り立つとき，「$x = b$ の近くで f は g のオーダー」であるといい，

$$f(x) = O\big(g(x)\big) \qquad (x \to b - 0) \tag{1.2.3}$$

[2] 完全な議論のためには「コーシー列」の概念が必要だが，これは後で「級数」を論じる時に勉強する

と書く．
(2) $f(x), g(x)$ が半無限区間 $[a,\infty)$ で定義されているとする．ある数 K をとると十分大きな x で $|f(x)| \le K|g(x)|$ が成り立つとき,「$x\to\infty$ で f は g のオーダー」であるといい,

$$f(x) = O\bigl(g(x)\bigr) \qquad (x\to\infty) \tag{1.2.4}$$

と書く．

（注意）

- 上では半開区間 $[a,b)$ について述べたが，$(a,b]$ などでも同様の定義を行う．
- 「f は g のオーダー」とは言っても，$|f(x)$ が $|g(x)|$ よりも格段に小さい場合も含むことに注意．
- 教科書ではわかりやすいように，$f,g\ge 0$ の場合に限定しているが，実際には f,g の正負に関わらずこの表現を使うことが多いので，上ではそうした．
- 教科書の定義 3.5.4 の後半には $f(x)\sim g(x)$ の記号が導入されているが，この教科所の用法は（1 年生には良いかもしれないが）数学の大勢の使い方とは異なり，非常によろしくないと考える．よって，この講義ではこの記号法は用いない．
- 興味のある人のために書いておくと，数学の大勢を占める使い方では,「$x\to a$ の時に $f(x)\sim g(x)$」とは,

$$\lim_{x\to a}\frac{f(x)}{g(x)} = 1 \tag{1.2.5}$$

となることを指す．つまり，$f(x)$ と $g(x)$ が同じオーダーだけでなく，その大きさまでほとんど同じ（比をとって 1）ことを意味する．教科書の定義の後半にある「$f(x)$ と $g(x)$ が同じオーダー」という状況は

$$f(x)\asymp g(x) \qquad (x\to a) \tag{1.2.6}$$

と書かれることが多いが，ひとによっては $f(x)\approx g(x)$ と書く場合もある．

上の定義を用いると，以下の十分条件を得る．

命題 1.2.3 (教科書の命題 3.5.5 と命題 3.5.11) この命題では $f(x), g(x)\ge 0$ とする．
(1) $x\ge a$ で $f(x)$ が有界，かつ $x\to\infty$ で $f(x) = O(g(x))$ であるとする．このとき，広義積分 $\displaystyle\int_a^\infty g(x)dx$ が収束すれば，広義積分 $\displaystyle\int_a^\infty f(x)dx$ も収束する．
(2) $x=a$ 以外では $f(x)$ が有界，かつ $x\to a+0$ では $f(x) = O(g(x))$ であるとする．このとき，広義積分 $\displaystyle\int_a^b g(x)dx$ が収束すれば，広義積分 $\displaystyle\int_a^b f(x)dx$ も収束する．

証明：
命題 1.2.1 を用いる．(1), (2) とも，$\int g(x)dx$ の収束性は，命題 1.2.1 の $S(b)$ または $S(c)$ の有界性を保証する．従って，$\int f(x)dx$ の存在が直ちに証明される． □

この定理から直ちに，始めの I_3 の収束性を結論できる．実際,

$$0\le\frac{|\sin x|}{x^2}\le\frac{1}{x^2} \tag{1.2.7}$$

である上に $\displaystyle\int_1^\infty\frac{1}{x^2}dx$ は収束するから，上の命題から直ちに，I_3 の収束性が結論できるのだ．

もう少し典型例を書いておこう．上の命題を用いることにより，以下に挙げた例以外にも判定できるものがあることには注意のこと．

- $\displaystyle\int_1^\infty \frac{dx}{x^\alpha}$ は $\alpha > 1$ ならば収束し，$\alpha \le 1$ ならば発散する．
- $\displaystyle\int_2^\infty \frac{dx}{x^\alpha(\log x)^\beta}$ は，$\alpha > 1$ ならば収束し，$\alpha < 1$ ならば発散する．$\alpha = 1$ の時は，$\beta > 1$ なら収束し，$\beta \le 1$ なら発散する．
- $\displaystyle\int_0^1 \frac{dx}{x^\alpha}$ は $\alpha < 1$ ならば収束し，$\alpha \ge 1$ ならば発散する．

以上よりも強力な「コーシー列」の方法については，この後の「級数論」で勉強する．

1.3 積分の応用

ここは特に目新しいことはない．教科書の 3.7 節に目を通して頂ければ良いだろう．

2 微分方程式

微分方程式について，その基本を簡単に説明する．教科書には対応する部分はない．

2.1 微分方程式とは

まず例から始めよう．x の関数 $y(x)$ が

$$y'(x) = \{y(x)\}^2 + x \tag{2.1.1}$$

を満たしているとする（$y'(x)$ はもちろん，$y(x)$ の導関数）．この式は，y の導関数の値が x と y で決まる形になっており,「微分方程式」と呼ばれる．

より一般に $f(x, y)$ は何か決まった x と y の関数だとして

$$y'(x) = f(x, y(x)) \tag{2.1.2}$$

の形の方程式を **x を独立変数，$y(x)$ を未知関数とする常微分方程式**と呼ぶ．(最初の例では $f(x, y) = y^2 + x$ である.) この方程式を満たす $y(x)$ を求めるのが，常微分方程式の典型的な問題である．

常微分方程式は上の形のものに限らない．一般化の方向としては 2 つある：(1) 出てくる微分を高階にする，(2) 未知関数の数を増やす．

(1) の例としては

$$y''(x) = y'(x) + \sin y(x) + x^3 \tag{2.1.3}$$

などが挙げられる．また，(2) の例としては y, z を未知関数として

$$\begin{cases} y'(x) = -z(x) \\ z'(x) = y(x) \end{cases} \tag{2.1.4}$$

などが挙げられる.（一般に (2) の場合は未知関数の数と同じだけの方程式が連立されることが多い — もし方程式の数が未知関数の数より少なければ，未知関数が一意に決まらないことが多い）．このように一般の常微分方程式を考えることができるが，**この講義では簡単のため，未知関数が一つだけ，出てくる微分も一階だけ，のものを主に考える．**

なお，常微分方程式は英語では ordinary differential equation であるので，この講義では常微分方程式のことを ODE と略記することが多い．

x を独立変数，$y(x)$ を未知関数とする常微分方程式

$$y'(x) = f(x, y(x)) \tag{2.1.5}$$

を考えよう（f は決まった関数である）．この方程式を満たす $y(x)$ のことをこの常微分方程式の**解**（または一般解）という.（実は常微分方程式の解は一般には一つに決まらない．この点については次の段落を見よ.）

実は常微分方程式 (2.1.5) を考える場合，ある x_0 における y の値 $y(x_0)$ が，ある特定の値であるような解に注目することが多い．すなわち，常微分方程式のたくさんある解（一般解）のうち，ある決まった y_0 に対して

$$y_0 = y(x_0) \tag{2.1.6}$$

となるような解を考えるのである．この条件を**初期条件**と呼び，上の解を「初期条件 (2.1.6) を満たす解」という．

> この言葉の由来は以下の通り．そもそも，常微分方程式という概念は，物事が時間とともにどう変化するかを記述するために導入された．つまり，独立変数 x というのは時間（時刻）であって，常微分方程式 (2.1.5) は我々の注目する量 $y(x)$ が時間 x とともにどう変わるかを記述するものなのである．このような問題では，現在の時刻を x_0 と書いた場合，現状での y の値 $y_0 = y(x_0)$ を決めて，この量が未来（$x > x_0$）においてどう変化するかを知りたい．
>
> このように考えれば，$x = x_0$ での状態を決める条件 (2.1.6) は y の初期の状態を決める条件と解釈できる．よって，これを初期条件と呼ぶのである．

これらの用語を用いると，我々の問題は「常微分方程式 (2.1.5) の解（一般解）を求めること」および「初期条件 (2.1.6) を満たす常微分方程式 (2.1.5) の解を求めること」となる．以下，この問題を考えて行く．

2.2 解の存在と一意性

具体的に常微分方程式を解き始める前に，そもそも，常微分方程式には解はあるのか，あるとしたらどのくらいの数の解があるのか，を考えておこう．非常に一般的な定理として，以下のものが挙げられる．この定理は少々長いので，まず問題設定から説明しよう．

設定：

- 正の定数 a, b と実数の組 (x_0, y_0) に対して，x, y 平面上の領域
$$D := \{(x,y) \,\big|\, |x - x_0| \le a,\ |y - y_0| \le b\} \tag{2.2.1}$$
を定義する．
- x, y の関数 $f(x,y)$ があり，これは D 内では x, y のそれぞれに関して**連続**とする．
- 閉領域 D において f が連続なので，$|f|$ の最大値が存在する．それを M と書こう．
- 更に，正の定数 K が存在して，D 内の 2 点 (x, y_1) と (x, y_2) に対して
$$|f(x,y_1) - f(x,y_2)| \le K|y_1 - y_2| \tag{2.2.2}$$
がなりたつ．（このとき，f は D において **Lipschitz 条件**を満たすという．）

定理 2.2.1 上の設定を満たす $f(x,y)$ を用いて定義された微分方程式
$$y'(x) = f\bigl(x, y(x)\bigr) \tag{2.2.3}$$
の解で，初期条件
$$y(x_0) = y_0 \tag{2.2.4}$$
を満たすものは
$$\bigl|x - x_0\bigr| \le \min\Bigl\{a, \frac{b}{M}\Bigr\} \tag{2.2.5}$$
において存在する．しかもこの解は一意に定まる．

この定理は非常に重要である．(2.2.5) を満たす x の範囲では初期条件 (2.2.4) を満たす解が存在し，しかもそれが一つに決まってしまうことを保証してくれるのである．

一般解との関係：特定の初期条件を満たすとは限らない解はどのくらいあるのか？これも上の定理が答えてくれる．

この定理の初期条件にでてくる y_0 は（Lipschitz 条件が満たされる限り）任意の定数である．つまりこの定理は「初期条件 y_0 を変えることにより無数の解が作れる」「初期条件の数だけ解がある」ことも保証してくれる．この y_0 が不定積分の積分定数に相当している．

上記定理の効用：上の定理は一見，単なる衒学的な意味以上のものを持たないように見えるだろうが，これは大きな間違いである．もちろん，求めようとしている解が存在しないならそもそも微分方程式を習う意味もないのだから，解の存在を保証してくれる上の定理は大事だ．しかしそれ以上に，上の定理は（定理の成立範囲においては）**どんな汚い手を使ってでも解をひとつだけ見つければこっちの勝ちだ**，ということを保証してくれる意味で，実用上も非常に役に立つ．

例えば，微分方程式
$$y'(x) = \{y(x)\}^3 + \{y(x)\}^2 + y(x) - e^{3x} - e^{2x} \tag{2.2.6}$$

を，初期条件 $y(0) = 1$ の下で解くことを考えよう．ぱっと見たところ，この微分方程式は解けそうにない．しかし，$y(x) = e^x$ が実際に解になっていることは，元の微分方程式に代入すればすぐに確かめられる．上の定理から，これ以外の解がないことが保証されるので，これで解が完全に見つけられたことになる．

将来，非常に複雑な形の微分方程式が出て来て，しかもその解はこの関数○○である，と言われることがあるかもしれない．このような場合，どのようにしてその解を求めるかはそれほど重要ではない．むしろ，○○を微分方程式に代入して「○○が実際に解であること」および上記定理を援用して「○○以外には解がないこと」を納得する方がよほど大切である．

ともかく，こういう訳で，この節の冒頭の問いの答えが出た．おおざっぱに言うと，Lipschitz 条件を満たす微分方程式に対しては，解が存在し，しかもただ一つに定まる，ということである．

この講義ではほとんど触れる余裕がないだろうが，未知関数が 2 つ以上の連立常微分方程式についても，同様の存在と一意性の定理が成り立つ．以下では未知関数が n 個の場合を参考までに書いておく．かなりややこしいので，無理に読まなくても良い — 講義では簡単にどういうことかを説明する．

設定：

- 正の定数 a, b と実数の組 $(x^{(0)}, y_1^{(0)}, y_2^{(0)}, \ldots, y_n^{(0)})$ に対して，$n+1$ 次元空間の領域
$$D := \{(x, y_1, y_2, \ldots, y_n) \,\big|\, |x - x^{(0)}| \le a,\ |y - y_1^{(0)}| \le b,\ |y - y_2^{(0)}| \le b,\ \ldots, |y - y_n^{(0)}| \le b,\ \} \tag{2.2.7}$$
を定義する．
- $x, y_1, y_2, \ldots, y_n$ の関数が n 個あり，これを $f_i(x, y_1, y_2, \ldots, y_n)$ とする（$i = 1, 2, \ldots, n$）．各 f_i は D 内では $x, y_1, y_2, \ldots, y_n$ のそれぞれに関して**連続**である．
- 閉領域 D において f_i が連続なので，各 $|f_i|$ の最大値が存在する．それらの $1 \le i \le n$ に関しての最大値を M と書こう．
- 更に，正の定数 K が存在して，D 内の 2 点 $(x, y_1, y_2, \ldots, y_n)$ と $(x, z_1, z_2, \ldots, z_n)$ に対して
$$|f_i(x, y_1, y_2, \ldots, y_n) - f_i(x, z_1, z_2, \ldots, z_n)| \le K \sum_{j=1}^{n} |y_j - z_j| \qquad (1 \le i \le n) \tag{2.2.8}$$
がなりたつ.（このとき，f は D において **Lipschitz 条件**を満たすという.）

定理 2.2.2　上の設定を満たす $f_i(x, y_1, y_2, \ldots, y_n)$ を用いて定義された連立微分方程式
$$y_i'(x) = f_i\bigl(x, y_1(x), y_2(x), \ldots, y_n(x)\bigr) \qquad (1 \le i \le n) \tag{2.2.9}$$
を考える．この常微分方程式の解で，初期条件
$$y_j(x^{(0)}) = y_j^{(0)} \qquad (1 \le j \le n) \tag{2.2.10}$$
を満たすものは
$$\bigl|x - x^{(0)}\bigr| \le \min\Bigl\{a, \frac{b}{M}\Bigr\} \tag{2.2.11}$$
において存在する．しかもこの解は一意に定まる．

2.3　変数分離形の微分方程式

では，微分方程式の解法その 1 として,「変数分離形の微分方程式の解法」を説明しよう.「変数分離形の微分方程式」とは以下の形のものをいう：
$$y'(x) = \frac{X(x)}{Y(y(x))} \tag{2.3.1}$$

ここで $Y(y)$ は y だけの，$X(x)$ は x だけの関数である．これは左辺を $\dfrac{dy}{dx}$ とかくと

$$\frac{dy}{dx} = \frac{X}{Y} \tag{2.3.2}$$

となって，分母と分子に x だけ，y だけの関数が分離されてでたように見える．このために変数分離形という．

（例）

$$y'(x) = y, \qquad y'(x) = y^2\, x^3, \qquad y'(x) = \frac{x^2+1}{y} \tag{2.3.3}$$

このとき，この方程式の，初期条件 $y(x_0) = y_0$ （ただし，$y_0 \neq 0$ と仮定する）に相当する解 $y(x)$ は

$$\int_{y_0}^{y} Y(s)\, ds = \int_{x_0}^{x} X(t)\, dt \tag{2.3.4}$$

で与えられる，というのが変数分離法である.（解であるところの $y = y(x)$ における x, y が (2.3.4) を満たす，つまり，(2.3.4) を書き下せば x と y の関係がわかる，という主張である.）[3]

（証明） (2.3.4) の両辺を x で微分すると（左辺の y は x の関数 $y(x)$ であることを忘れない）

$$Y(y(x))\, y'(x) = X(x)$$

となるから，両辺を $Y(y)$ で割ればもともとの微分方程式 (2.3.1) になる．つまり，(2.3.4) をみたす $y(x)$ が (2.3.1) の解であることが示された．更に，存在と一意性の定理 2.2.1 を用いると，これ以外に解がないこともわかる．つまり，天下りに与えた (2.3.4) が，実際に我々の問題の解を与えてくれることがわかった． □

練習問題：次の微分方程式を，与えられた初期条件の下で解け．

$$y' = 3y, \qquad y(0) = 2$$

$$y' = y^2 - 1, \qquad y(0) = 2$$

次の微分方程式を解け（一般解を求めよ）．

$$y' = -2xy$$

$$y' = y \sin x$$

$$y' = 2y(3-y)$$

重要な注意：通常の方程式の場合と同じく，「微分方程式を解け」の問題では多くの場合，検算が可能である — 単に，得られた解が本当に初期条件を満たしているのか，またもともとの常微分方程式を満たしているのか，代入して計算してみればよい．

[3] 厳密なことをいうと，(2.3.4) が解であることが保証されるのは，x_0 周りにある程度限定された x においてのみであるが，この講義では立ち入らない

3 多変数関数の微分

これから，偏微分（多変数関数の微分）を扱う．厳密性にはあまりこだわらず，高校までの「ええ加減」なノリで[4]，ともかく概念を理解する事を目標にしよう．

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

3.1 １変数の関数，多変数の関数

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

$$x^2, \quad x^4, \quad \sin x \quad \cos(x^2+2), \quad \ldots \tag{3.1.1}$$

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

（余談）この定義通り,「関数」とは何でも良く，高校までの常識からは関数に見えないようなもの —— 例えば，そのグラフが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．

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

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

とする．ここまでは高校と同じだが，強いて言えば，$|x-y|$ という絶対値を２点 x, y の間の距離と解釈することが目新しいかもしれない.「差の絶対値は距離」という見方はこれからも頻出する，非常に重要なものである[5]．

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

n 変数の関数は以下のように定義される（$n \geq 2$）．一般の n で考えにくい人は $n=2$（平面），$n=3$（空間）を思い浮かべれば十分だ[6]．

- 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：$\mathbb{R}^n := \{\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}$．

そして

> **定義 3.1.1** D を n 次元空間 $\mathbb{R}^n$ の部分集合とする：$D \subset \mathbb{R}^n$．定義域が D である n 変数の実数値関数 f とは，D の各点 $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ 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．

さらに１次元の時に倣って

[4]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

[5]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

[6]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

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

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \Bigl(\sum_{j=1}^{n} (x_j - y_j)^2\Bigr)^{1/2} \tag{3.1.2}$$

  により定義する．これは $n = 2$ の時には普通の平面での距離，$n = 3$ の時には 3 次元空間での距離である．
- ただし，いつでも上のように x_1, x_2, x_3 などとしているとかえって書きにくいこともあるので，適宜 $\boldsymbol{x} = (x, y)$, $\boldsymbol{z} = (u, v)$ 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 開集合でかつ連結 なものを 領域（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考える．まずは 1 変数の場合を思い出そう．

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

$$|x - a| \to 0 \quad ならば \quad \bigl|f(\boldsymbol{x}) - \alpha\bigr| \to 0 \tag{3.1.3}$$

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

定義 3.1.3 (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．

$$\|\boldsymbol{x} - \boldsymbol{a}\| \to 0 \quad ならば \quad \bigl|f(\boldsymbol{x}) - \alpha\bigr| \to 0 \tag{3.1.4}$$

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 3.1.4 (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.）

3.2 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 3.2.1 (偏微分係数) 2 変数関数 $f(x,y)$ の点 (a,b) における 第 1 変数に関する偏微分係数 とは極限

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

のことである（もちろん，この極限が存在する場合のみ，この定義は有効）．これは記号で $\dfrac{\partial f}{\partial x}(a,b)$，$f_1(a,b)$，$f_x(a,b)$，$D_1f(a,b)$ などと書く．同様に，第 2 変数に関する偏微分係数とは

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

のことであって，$\dfrac{\partial f}{\partial y}(a,b)$，$f_2(a,b)$, $f_y(a,b)$，$D_2f(a,b)$ 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 3.2.2 (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

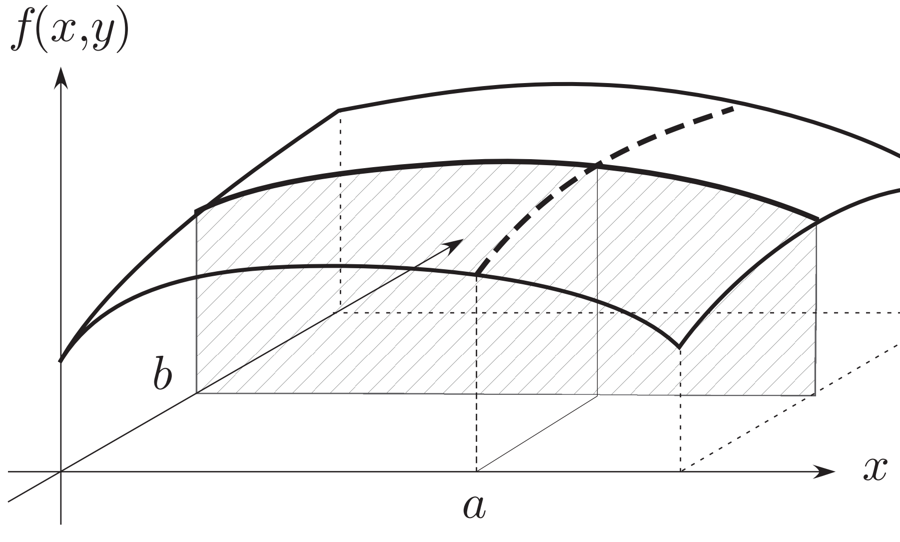
- f は各変数 $x_1,x_2,\ldots,x_n$ のそれぞれについて偏微分可能で
- かつ，その n-この偏導関数が $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ の連続関数である

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

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



(記号についての注意)

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

問 3.2.1. 次の関数をそれぞれの独立変数で偏微分せよ．

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

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

3.2.1 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

$$f(x,y)=g(y)\qquad g\text{ は任意の関数}\tag{3.2.3}$$

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．

3.2.2 方向微分[7]

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

$$f_{\boldsymbol{v}}(a,b):=\lim_{h\to 0}\frac{f(a+hv_x,b+hv_y)-f(a,b)}{h}\tag{3.2.4}$$

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

$$f_{\boldsymbol{v}}(\boldsymbol{a}):=\lim_{h\to 0}\frac{f(\boldsymbol{a}+h\boldsymbol{v})-f(\boldsymbol{a})}{h}\tag{3.2.5}$$

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．

[7]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．

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

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

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

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

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

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

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

3.2.3 全微分可能性[8]

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：

$$f(x)\approx f(a)+f'(a)\times(x-a). \tag{3.2.10}$$

我々は当然，偏微分にも同じ役割を担ってほしい．つまり，$\boldsymbol{x}=\boldsymbol{a}$ 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 (3.2.6)，(3.2.7) の例を考えてみるとよい．$x=y=0$（原点）では f の偏微分係数はともにゼロであるが，$f(x,y)$ は原点付近でゼロではない．たとえば $x=y$ では $f(x,x)=1$（$x\neq0$）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[9]，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 $\|\boldsymbol{x}-\boldsymbol{a}\|=\left(\sum_{j=1}^{n}(x_j-a_j)^2\right)^{1/2}$ である．

定義 3.2.4 (全微分可能性) ある領域 D で定義された n 変数関数 $f(\boldsymbol{x})$ 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），

$$f(\boldsymbol{x})=f(\boldsymbol{a})+\sum_{j=1}^{n}A_j(x_j-a_j)+\tilde{f}(\boldsymbol{x}), \qquad \text{with} \quad \lim_{\boldsymbol{x}\to\boldsymbol{a}}\frac{\tilde{f}(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|}=0 \tag{3.2.11}$$

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，(3.2.6) や (3.2.7) の f,g は原点 $(0,0)$ では全微分可能でない．

[8] 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，
[9] 春学期に少しやった

上の定義のミソは (3.2.11) 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 3.2.5 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，

$$\Bigl(\ (3.2.11)\ \text{の}\ \ A_j\ \Bigr) = \frac{\partial f}{\partial x_j}(\boldsymbol{a}) \qquad (j=1,2,\ldots,n), \tag{3.2.12}$$

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

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

証明.

1. f が連続なのはほとんど自明だ．というのも，全微分可能の条件 (3.2.11) は $f(\boldsymbol{x})-f(\boldsymbol{a})$ 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 $x_2,x_3,\ldots$ は $a_2,a_3,\ldots$ に固定して考えるので，(3.2.11) から

$$\frac{f(x_1,a_2,a_3,\ldots,a_n)-f(a_1,a_2,a_3,\ldots,a_n)}{x_1-a_1} = A_1 + \frac{\tilde{f}(\boldsymbol{x})}{x_1-a_1} \tag{3.2.14}$$

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，$x_2,x_3,\ldots$ を $a_2,a_3,\ldots$ に固定した場合は $|x_1-a_1| = \|\boldsymbol{x}-\boldsymbol{a}\|$ であるので，(3.2.11) 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 (3.2.11) から

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

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

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

であるため，(3.2.11) から，(3.2.15) の最後の項は $h \to 0$ でゼロに行く．よって，(3.2.13) 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 (3.2.11) は図形的には以下のように解釈できる．まず，(3.2.11) の最後の項がない場合を考えると，定数 A,B があって

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

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

実際には (3.2.11) には余分な項がついているわけで，$z=f(x,y)$ のグラフは簡単な平面ではない．しかし，$\boldsymbol{x}-\boldsymbol{a}$ 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 接平面 になっている．

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

全微分可能の十分条件

最後に，全微分可能の十分条件を一つ，与えておこう．

定理 3.2.6 (C^1 級なら全微分可能) ある領域 D で定義された n 変数関数 $f(\boldsymbol{x})$ 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，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

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

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

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

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

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

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

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

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

を得る．

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

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

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

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

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

が結論できる．同様に，

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

も結論できる．これを (3.2.21) に代入すると

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

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

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

が結論できる．結果として，このような $\tilde{f}$ を用いて

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

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

3.3 合成関数の微分（連鎖率，chain rule）

ここでは偏微分での最初の山場,「連鎖率」（合成関数の微分）を学ぶ．この題材は簡単に見えて，案外たいへんなことがあるから，注意する事．特に，この後でやる「高階の導関数」を計算する時にひっかかる人が多いはずだ．なお，この節の山場は後の 3.3.3 節である[10]．

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，

$$h(x) = f(g(x)) \tag{3.3.1}$$

で定義される関数 h を f と g の 合成関数 といい，$f \circ g$ などと書いたのだった.「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，

$$h'(x) = f'\bigl(g(x)\bigr)\, g'(x) \tag{3.3.2}$$

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

$$\frac{\partial h(x,y)}{\partial x} = f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial x}, \qquad \frac{\partial h(x,y)}{\partial y} = f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial y} \tag{3.3.3}$$

となる．

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

3.3.1 合成関数の微分（１変数の場合の復習，Case A）

$g(x)$ は区間 I で，$f(y)$ は区間 J で，それぞれ定義されており，かつ，g の値域 $g(I) = \{g(x) \,|\, x \in I\}$ が J の部分集合であるとする．このとき，合成関数 $h(x) = f\bigl(g(x)\bigr)$ 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．

[10] 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

定理 3.3.1（1 変数の合成関数の微分） $g(x)$ が区間 I 内の点 $x=a$ にて x について微分可能，かつ $f(z)$ が点 $b=f(a)$ にて z について微分可能のとき，合成関数 $h(x)=f\bigl(g(x)\bigr)$ 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 3.3.3 では事情が異なり，導関数の連続性（またはそれに類する条件）が必要になってくる．この事情は「全微分可能性」と関連している．

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

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

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 (3.3.4) 右辺の書き換え（分母がゼロ！）に意味がつけられなくなり，右辺の積の極限を別々に考えることができなくなる．

でも，これは大した問題ではない．実際，$g(a+\epsilon)-g(a)=0$ ならば $h(g(a+\epsilon))-h(g(a))=0$ でもあるはずだから，もともとのニュートン商の値もゼロ，よって困ることは何もないはずだ．実際，こ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．

3.3.2 合成関数の微分（1 変数の場合に帰着, Case B）

既に注意したように，B の場合は上からすぐに出る．つまり，区間 J で定義された関数 $f(z)$ と，ある領域 D で定義された関数 $g(x,y)$ があって，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 $h(x,y)=f\bigl(g(x,y)\bigr)$ を D で定義することができるが...

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

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

がなりたつ．

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

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

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

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[11]］

$$\lim_{\epsilon\to 0} \frac{h(t+\epsilon) - h(t)}{\epsilon} = \lim_{\epsilon\to 0} \frac{f\bigl(x(t+\epsilon), y(t+\epsilon)\bigr) - f\bigl(x(t), y(t)\bigr)}{\epsilon} = \lim_{\epsilon\to 0} \frac{f(x_1,y_1) - f(x_0,y_0)}{\epsilon}$$
$$= \lim_{\epsilon\to 0} \frac{f(x_1,y_1) - f(x_1,y_0)}{\epsilon} + \lim_{\epsilon\to 0} \frac{f(x_1,y_0) - f(x_0,y_0)}{\epsilon} \tag{3.3.6}$$

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**）．もし これが（x_2 は ϵ に無関係な数で）

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

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

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

が得られると 予想 される．

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

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

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

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

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

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

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 3.3.4 節で述べる.) なお，あまり細かいことを書くと肝心のところが見えなくなりそうだから，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 3.3.3 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

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

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

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

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

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

証明　(3.3.6) の第一項がゴマカシだったので，きちんとやりなおそう．$f(x,y)$ を，y だけの関数と見て 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）

$$f(x_1,y_1)-f(x_1,y_0)=f_y(x_1,y_3)\times(y_1-y_0) \tag{3.3.13}$$

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

$$\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}=\lim_{\epsilon\to 0}\Bigl[f_y(x_1,y_3)\times\frac{y_1-y_0}{\epsilon}\Bigr] \tag{3.3.14}$$

となる．ところで，f が C^1 級と仮定しているから，$f_y(x,y)$ は x,y の連続関数である（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，

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

である．また，(3.3.14) の後ろの方は y_1,y_0 の定義から

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

である．以上をまとめると

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

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

D の場合についてもだめ押しで述べておこう．

定理 3.3.4（２変数関数の連鎖律） 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，

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u},\qquad\text{かつ}\qquad\frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v}. \tag{3.3.18}$$

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

問題 3.3.5　（連鎖率） 2 変数の関数 $f(x,y)=\dfrac{xy}{\sqrt{x^2+y^2}}$ 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 3.3.6 (x,y) と (u,v) が 1 次変換（$\alpha,\beta,\gamma,\delta$ は $\alpha\delta\neq\beta\gamma$ を満たす定数）

$$u=\alpha x+\beta y,\ \ v=\gamma x+\delta y \tag{3.3.19}$$

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

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

1) は既にやった（3.2.1 節）．問題は 2) と 3) だが，問題 3.3.6 をヒントにせよ．つまり，新しい変数 (u,v) をうまく見つけて，$\dfrac{\partial f}{\partial u}\equiv 0$ 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．）

3.3.4 全微分可能性を仮定したときの chain rule

（この小節の内容は「おまけ」であり，余裕のある人だけが読めばよい.）

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

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

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

が成り立つ．n-変数の場合も同様であるが，略.

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

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u},\qquad \text{かつ}\qquad \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v} \tag{3.3.22}$$

である.

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

定理 3.3.8 の証明

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 3.2.5 から，

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

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

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

も成り立っている．従って，

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

となるが，

$$o\Bigl(\sqrt{(p^2+q^2)\epsilon^2+o(\epsilon^2)}\Bigr)=o\Bigl(\sqrt{p^2+q^2}\,\epsilon+o(\epsilon)\Bigr)=o(\epsilon) \tag{3.3.26}$$

であるので，結局 (3.3.25) は

$$h(t+\epsilon)=h(t)+(Ap+Bq)\epsilon+o(\epsilon) \tag{3.3.27}$$

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

3.4 高階の偏導関数

（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

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

なども定義する（もちろん，これらの微分が定義できる場合）．まあ，この定義は非常に自然だから問題ないでしょう[12]．

問題 3.4.1 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．

$$f(x, y) = x^2 + y^2, \qquad g(x, y) = x^2 y^3, \qquad h(x, y) = \cos(x^2 y) \tag{3.4.2}$$

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 3.4.1 の微分を実際にやった人には，$\frac{\partial g}{\partial x}$ と $\frac{\partial g}{\partial y}$ が全然別ものだから，$\frac{\partial^2 g}{\partial x \partial y}$ と $\frac{\partial^2 g}{\partial y \partial x}$ に関係がつく事の方がかえって奇跡に見えるかもしれない．しかし，(我々が扱うような「普通の」関数では）この２つは等しい．

定理 3.4.2 ２変数関数 $f(x, y)$ が C^2-級の場合，

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \tag{3.4.3}$$

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 3.4.2 での「C^2-級」は十分条件であって，もう少しだけ条件を緩めることも可能である．例えば，以下のようなものがあるようだ[13]（D は領域，A は D 内の一点．また $f_x = \frac{\partial f}{\partial x}, f_{yx} = \frac{\partial^2 f}{\partial x \partial y}$ などと略記する）．

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

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

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

[12] 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

[13] 小平本の 6.2 節の d) を参考にした

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，

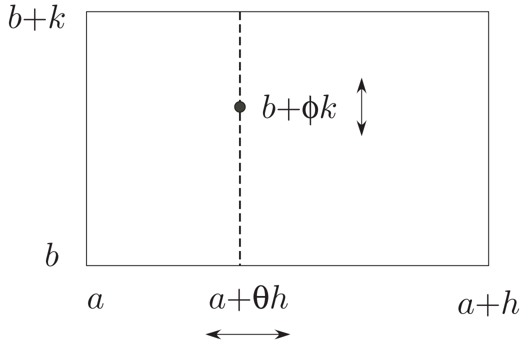
$$\begin{aligned}\lim_{h\to 0}\Bigl(\lim_{k\to 0}\frac{\Delta(h,k)}{hk}\Bigr) &= \lim_{h\to 0}\frac{1}{h}\Bigl[\lim_{k\to 0}\Bigl(\frac{f(a+h,b+k)-f(a+h,b)}{k}-\frac{f(a,b+k)-f(a,b)}{k}\Bigr)\Bigr] \\ &= \lim_{h\to 0}\frac{1}{h}\Bigl[\frac{\partial f}{\partial y}(a+h,b)-\frac{\partial f}{\partial y}(a,b)\Bigr]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

であるし，同様に考えると

$$\lim_{k\to 0}\Bigl(\lim_{h\to 0}\frac{\Delta(h,k)}{hk}\Bigr) = \frac{\partial^2 f}{\partial y \partial x} \tag{3.4.6}$$

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

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）．



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，

$$\varphi(a+h)-\varphi(a) = \varphi'(a+\theta h)h \tag{3.4.7}$$

つまり

$$\Delta(h,k) = \varphi(a+h)-\varphi(a) = \Bigl(f_x(a+\theta h,b+k)-f_x(a+\theta h,b)\Bigr)h \tag{3.4.8}$$

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

次に a,h,θ を固定して $\psi(y) := f_x(a+\theta h,y)$ を y の関数と考えると[14]またもや 1 変数の場合の平均値の定理から

$$\psi(b+k)-\psi(b) = \psi'(b+\phi k)k \tag{3.4.9}$$

つまり

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

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 (3.4.7) と (3.4.10) から

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

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

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

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

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 (3.4.5) と (3.4.6) を思い出すと証明は完成した.（余分ではあるが，上の議論を順序を変えて行うと (3.4.12) の極限が $f_{yx}(a,b)$ に等しい事もいえて，証明はより明確になる.） □

3.4.1 ２階の偏微分係数の幾何学的意味

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

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

問題は $f_{xy} = f_{yx}$ だ．x で微分してから y で微分と言われても，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

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

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

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

となる.（余談だが，上の関係は微分演算子の部分だけを取り出して

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

とも書ける．このような書き方は後々，便利だ.）これを y で偏微分して

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

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

3.4.2 高階偏導関数と連鎖律

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

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

となるので，微分演算子としては

$$\frac{\partial}{\partial x} = \frac{x}{r}\frac{\partial}{\partial r} - \frac{y}{r^2}\frac{\partial}{\partial \theta} = \cos\theta\,\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta} \tag{3.4.18}$$

という作用をもっている．さて，もう一回やるときには

$$\frac{\partial^2 f}{\partial x^2}=\Bigl(\cos\theta\,\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigr)\Bigl(\cos\theta\frac{\partial f}{\partial r}-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\Bigr) \tag{3.4.19}$$

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

$$=\cos\theta\frac{\partial}{\partial r}\Bigl(\cos\theta\frac{\partial f}{\partial r}\Bigr)-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(\cos\theta\frac{\partial f}{\partial r}\Bigr)+\cos\theta\frac{\partial}{\partial r}\Bigl(-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\Bigr)-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\Bigr) \tag{3.4.20}$$

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

$$\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(\cos\theta\frac{\partial f}{\partial r}\Bigr)=\frac{\sin\theta}{r}\frac{\partial\cos\theta}{\partial\theta}\times\frac{\partial f}{\partial r}+\frac{\sin\theta}{r}\cos\theta\frac{\partial}{\partial r}\Bigl(\frac{\partial f}{\partial\theta}\Bigr)=-\frac{\sin^2\theta}{r}\frac{\partial f}{\partial r}+\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{3.4.21}$$

などとなる訳だ．このように計算していくと，最終的な答えは

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

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

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

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

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

3.4.3 （補足）偏導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

$$\frac{\partial}{\partial y}\bigl(f_x\bigr)=0 \tag{3.4.24}$$

と見ると,「f_x は y には依存しません」ということなので,

$$f_x(x,y)=g(x) \tag{3.4.25}$$

と書けるはずだ（g は任意の関数，3.2.1 節を思い出そう）．この両辺を x で積分すると，左辺は $f(x,y)$ になり，右辺は g の原始関数になるが，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，(3.4.25) から

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

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，

$$f(x,y)=G(x)+h(y) \tag{3.4.27}$$

となることがわかった（h がその「積分定数」としてでてくる y の任意関数）．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．

3.5 平均値の定理

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

定理 3.5.1 (多変数の平均値の定理；教科書の定理 5.1.11 と 5.1.13) 2 変数関数 $f(x, y)$ が C^1-級の場合,

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

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

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{j=1} \frac{\partial f}{\partial x_j}(\boldsymbol{a}+\theta\boldsymbol{h})\, h_j \tag{3.5.2}$$

が成り立つ（$0 < \theta < 1$）.

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

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

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

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

であるから，定理 3.5.1 を得る．n 変数の場合も同様である． □

平均値の定理が成立するには，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．この辺りは数学としては興味のあるところだが，余裕のない人はあまりこだわる必要はない．上の定理だけ理解すれば（一年生の間は）十分だ．

定理 3.5.2 (平均値の定理) 2 変数関数 $f(x, y)$ が全微分可能なら，

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

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

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

3.6 テイラーの定理とテイラー展開

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

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

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

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

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

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

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

であるから，例えば，f が C^2-級ならば

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

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

$$\big|g^{(2)}(\theta)\big| \le 2M(h^2 + k^2) = O(\|\boldsymbol{a} - \boldsymbol{c}\|^2) = o(\|\boldsymbol{a} - \boldsymbol{c}\|) \tag{3.6.5}$$

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

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

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

この先もどんどんやれる．f が C^3-級だと仮定すると，

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

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

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

定理 3.6.1 (教科書の定理 5.2.4 などの一般形) ２変数の関数 $f(x, y)$ が C^r-級であるとき，適当な $0 < \theta < 1$ に対して

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

が成り立つ．

（注）上の θ が a, b, h, k に依存するのは，１変数の場合と同じである．

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

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

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

3.7 極大・極小問題

高校で習った微分の応用は，ほとんど最大・最小の問題につきるだろう．実際，微分の意義は最大・最小問題が簡単にわかることにあると言ってよい．となれば当然，偏微分を用いれば多変数関数の最大・最小問題が解けると期待したくなる．実際，その通りなのだが，1 変数の場合よりは少し複雑だ．この節の主な目的は，その事情を良く理解することにある．

3.7.1 問題の定義

定義 3.7.1 n-成分ベクトルの空間において，上の記号のもとで，

$$B_r(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathbb{R}^n \,\big|\, \|\boldsymbol{x}-\boldsymbol{a}\| < r\} \tag{3.7.1}$$

なる集合 $B_r(\boldsymbol{a})$ を **$\boldsymbol{a}$ の r-開近傍**という．$\boldsymbol{a}$ を中心とした半径 r の球（の内部）ということである．

なお，適当に $r > 0$ をとったら a の r-開近傍で性質○○が成り立つ場合，単に「性質○○が $x = a$ の近傍で成り立つ」ということがある．

定義 3.7.2 n-変数の関数 $f(\boldsymbol{x})$ が $\boldsymbol{x} = \boldsymbol{a}$ で**極大**であるとは，適当な $r > 0$ に対して $\boldsymbol{a}$ の r-開近傍 $B_r(\boldsymbol{a})$ があって，その中では $f(\boldsymbol{a})$ の値が最大であることをいう（r は我々が勝手に設定してよい）．つまり，

$$\text{ある正の } r \text{ が存在して，} \qquad 0 < \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \text{では} \quad f(\boldsymbol{x}) < f(\boldsymbol{a}) \tag{3.7.2}$$

となることである．同様に，$f(\boldsymbol{x})$ が $\boldsymbol{x} = \boldsymbol{a}$ で**極小**であるとは，

$$\text{ある正の } r \text{ が存在して，} \qquad 0 < \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \text{では} \quad f(\boldsymbol{x}) > f(\boldsymbol{a}) \tag{3.7.3}$$

であることをいう．

- この代わりに等号も含めたもの，つまり (3.7.2) と (3.7.3) の代わりに

$$\exists r > 0, \qquad \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{3.7.4}$$

$$\exists r > 0, \qquad \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \ge f(\boldsymbol{a}) \tag{3.7.5}$$

  としたものを「広義の極大」「広義の極小」とよぶ．
- 高校でも強調されたかもしれないが，関数 $f(x)$ が $x = a$ で**最大**とは，f の**定義域全体**を見渡した時に $f(a)$ が最大であることをいう．つまり，

$$f \text{ の定義域に入っているすべての } \boldsymbol{x} \text{ に対して} \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{3.7.6}$$

  であることをいう（上の極大の定義のように $\boldsymbol{x}$ の範囲を我々が勝手に設定してはいけない）．最小についても同様である．なお，(3.7.6) で等号を入れるか入れないかはまた，悩ましい定義の問題だが，ここでは一応，等号も許す事にする．

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，またこの節でやるように）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．) この節では最大・最小問題にはほとんど触れず，極大・極小問題に話を限る．

3.7.2 １変数の場合の復習

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た（高校でやったはず）．

定理 3.7.3 $x = a$ 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.）
念のために定理のそれぞれの場合に相当する例を挙げておこう（すべて $a = 0$ の例）．

- $f(x) = x^2$ は (ii)-a，$f(x) = -x^2$ は (ii)-b の典型的な例である．
- $f(x) = x^3$ は (i) で「逆が成り立たない」例である．($x = 0$ で微係数がゼロでも極大でも極小でもない.)
- $f(x) = x^4$ や $f(x) = -x^4$ は (ii)-c の，極大や極小になる例である．
- $f(x) = x^3$ や $f(x) = x^5$ は (ii)-c で極大でも極小でもない例である．

この定理の厳密な証明は平均値の定理を用いるが，定理のような振る舞いは（少なくともええ加減には）テイラーの定理（テイラー展開）から理解できる．すなわち，$x = a$ の周りのテイラーの公式を

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + o(|x-a|^2) \tag{3.7.7}$$

と書いてみよう．もし $f'(a) \neq 0$ なら $x \to a$ では

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \tag{3.7.8}$$

となるから極大・極小にはなれないはずだ（この対偶をとると定理の (i)）．次に，$f'(a) = 0$ の場合は

$$f(x) = f(a) + \frac{f''(a)}{2}(x-a)^2 + o(|x-a|^2) \tag{3.7.9}$$

となるから，$f''(a) > 0$ なら $x \neq a$ では第２項が正になって，$f(x) > f(a)$ となるだろう．$f''(a) < 0$ の場合も同様である．最後に，$f''(a) = 0$ の場合はテイラーの公式をここまで書いたのではわからない．もっと高階の微係数も存在すると仮定して書いてみると［$f'(a) = f''(a) = 0$ の場合］，

$$f(x) = f(a) + \frac{f^{(3)}(a)}{6}(x-a)^3 + \frac{f^{(4)}(a)}{24}(x-a)^4 + \frac{f^{(5)}(a)}{120}(x-a)^5 + o(|x-a|^5) \tag{3.7.10}$$

となる．$x \to a$ では $(x-a)$ の次数の低い項が一番効く．従って，$f^{(3)}(a) \neq 0$ ならば $x = a$ は極大でも極小でもない［$(x-a)^3$ と同じような振る舞いになる］．一方，$f^{(3)}(a) = 0, f^{(4)}(a) > 0$ ならばこの $(x-a)^4$ の項が一番効いて，$x = a$ は極小になる．次に $f^{(3)}(a) = f^{(4)}(a) = 0$ で $f^{(5)}(a) \neq 0$ なら $(x-a)^5$ と同じような振る舞いで，極大でも極小でもない．以下同様で，テイラー展開の始めの数項がどうなっているかから考えていくと良い．

3.7.3 ２変数の極大極小問題

さて，本題の n-変数の場合にもどろう．まずは２変数関数の場合を考える．１変数の場合の経験から，f の２階微分が大事であろうことは想像できるだろうが，その通りである．まず，用語の定義：

定義 3.7.4 2変数の関数 $f(x,y)$ の，点 (a,b) における**ヘッセ行列**とは，以下の形の行列

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y)=(a,b) \text{ での値} \tag{3.7.11}$$

のことである．同様に，C^2-級の n-変数の関数 $f(x_1, x_2, \ldots, x_n)$ の点 $\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ におけるヘシアンとは，その ij 成分が $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a})$ となっているような $n \times n$ 行列のことである．ヘッセ行列の行列式を**ヘシアン**という．

（注）少し用語の混乱があるようで，ヘッセ行列そのものも「ヘシアン」ということもある（特に英語の文献では Hessian matrix の代わりに Hessian という事も多い）．多分，僕自身もヘッセ行列をヘシアンと言ってしまうことがあるでしょう．

すると，

定理 3.7.5 $(x,y) = (a,b)$ の近傍で定義された2変数の関数 $f(x,y)$ について，以下が成り立つ．（簡単のため，$\boldsymbol{x} = (x,y), \boldsymbol{a} = (a,b)$ とかく．）

(i) $f(\boldsymbol{x})$ が $\boldsymbol{x} = \boldsymbol{a}$ で微分可能，かつ $\boldsymbol{x} = \boldsymbol{a}$ で $f(\boldsymbol{x})$ が極大または極小の場合，$f_x(\boldsymbol{a}) = f_y(\boldsymbol{a}) = 0$ である．逆は必ずしもなりたたない．

(ii) $f(\boldsymbol{x})$ が $\boldsymbol{x} = \boldsymbol{a}$ で2階微分可能，$f_x(\boldsymbol{a}) = f_y(\boldsymbol{a}) = 0$ の場合，以下が成り立つ（微係数はすべて $\boldsymbol{a} = (a,b)$ における値を表す）．

a. $f_{xx}f_{yy} - f_{xy}f_{yx} > 0$（ヘシアンが正）の場合，$f(\boldsymbol{x})$ は $\boldsymbol{x} = \boldsymbol{a}$ で極小または極小である．詳しくは，
  - $f_{xx} > 0$ ならば f は (a,b) にて極小,
  - $f_{xx} < 0$ ならば f は (a,b) にて極大

  である．

b. $f_{xx}f_{yy} - f_{xy}f_{yx} < 0$（ヘシアンが負）の場合，$f(\boldsymbol{x})$ は $\boldsymbol{x} = \boldsymbol{a}$ で極大にも極小にもなれない（鞍点）．

c. $f_{xx}f_{yy} - f_{xy}f_{yx} = 0$ の場合，$f(\boldsymbol{x})$ の $\boldsymbol{x} = \boldsymbol{a}$ における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）．もっと詳しく調べる必要がある．

（注）上の b のような場合を「鞍点」と呼ぶ．

この定理のきちんとした証明は平均値の定理を用いて行えるが，それは教科書にも書いてあるからここには再現しない．もちろん，その証明が良くわかる人はそれで十分だが，その証明がわかりにくい人は，「なぜこうなのか」を大体でも理解することがまず大切だ（厳密にちゃんとやるのはその後でも良い）．そのために，テイラーの公式を使う理解の仕方を紹介しておこう．

関数が3階くらいまで微分可能だと思って2変数のテイラーの公式を書いてみると（f や f_x, f_{xy} などの引数はすべて (a,b) であるが，式がややこしくなるので省略した），

$$f(x,y) = f + f_x(x-a) + f_y(y-b) + \frac{1}{2}\Big[f_{xx}(x-a)^2 + 2f_{xy}(x-a)(y-b) + f_{yy}(y-b)^2\Big] + o(\|\boldsymbol{x}-\boldsymbol{a}\|^2) \tag{3.7.12}$$

となっていたことをまず，思い出そう．

(i) 1階微分の少なくとも1つがゼロでない場合．

さて，$f_x \neq 0$ や $f_y \neq 0$ の場合は点 (a,b) のごくごく近傍では $(x-a)$ や $(y-b)$ の1次の項が一番効く（2次以上の項は1次の項より凄く小さい）から，$f(x,y)$ は (a,b) では極大にも極小にもなれない（各自，確かめよ）．この対偶をとれば定理の (i) になる．

(ii) 1階微分が2つともゼロで，3つの2階微分の少なくとも一つがゼロでない場合．

次に，$f_x = f_y = 0$ の時には上の2次以上の項が重要になる．まずは2次の項のどれかがゼロでない場合を考えよう．この時は $o(\|\boldsymbol{x}-\boldsymbol{a}\|^2)$ の項が2次の項に比べて無視できる．

さて，１変数の時と異なって厄介なのは，真ん中の $2f_{xy}(x-a)(y-b)$ の項だ．他の２つの項では $(x-a)^2, (y-b)^2$ は共に正であるが，この真ん中の項では $(x-a)(y-b)$ は正にも負にもなるから，困ってしまう．これをちゃんと理解するには「行列の対角化」（線形代数でやりましたね）をやる必要がある．ここでは今考えている２変数に限って簡単に理解できる方法を説明しよう．

問題は（$A = f_{xx}, B = f_{xy} = f_{yx}, C = f_{yy}$）

$$g(x,y) = A(x-a)^2 + 2B(x-a)(y-b) + C(y-b)^2 \tag{3.7.13}$$

が $x = a, y = b$ の近傍で正か負かということだが，これは受験数学でやった平方完成の問題だ．

$A \neq 0$ の場合をまず考えると，

$$g(x,y) = A\left[\left\{(x-a) + \frac{B}{A}(y-b)\right\}^2 + \frac{CA-B^2}{A^2}(y-b)^2\right] \tag{3.7.14}$$

である．よって場合分けすると

- $A > 0$ かつ $CA - B^2 > 0$ ならば（$(x-a)^2 + (y-b)^2 > 0$ の時）これはいつも正
- $A < 0$ かつ $CA - B^2 > 0$ ならば（$(x-a)^2 + (y-b)^2 > 0$ の時）これはいつも負
- A の符号にかかわらず $CA - B^2 < 0$ ならばこれは正にも負にもなる
- $CA - B^2 = 0$ なら $x - a = B(y-b)/A$ の時にこれはゼロ $\Longrightarrow$ もっと高次の項まで考えないとわからない

となって，定理の a, b, c の場合がでてくる．

$C \neq 0$ の場合は x, y の役割を取り替えれば同様．

最後に $A = C = 0$ の場合は $g(x,y) = 2B(x-a)(y-b)$ であって，$B \neq 0$ ならこれは正にも負にもなりうるので，極大や極小にはなれない．$A = B = C = 0$ ならば $g(x,y) \equiv 0$ だから，高次の項を考えないと何も言えない．

(iii) １階微分も２階微分もすべてゼロの場合：

この時は $o(\|\boldsymbol{x} - \boldsymbol{a}\|^2)$ についてもっとたくさんの情報が得られない限りは，どうしようもない．この場合は定理では (ii) の c の場合に分類されてしまっているが．

ともかく，２変数の関数の場合に定理 3.7.5 を理解するのは，このように地道に考えれば可能である．なお，同様の議論を「行列の対角化」の話を用いて，この後で定式化しなおす． □

以上をまとめると，２変数の関数の極値問題の解き方は以下のようになる．

（１）極値を取る点の候補を求める．点 (a,b) で極値をとるとすると，そこでは

$$f_x(a,b) = f_y(a,b) = 0 \tag{3.7.15}$$

である**必要**がある．従って，上の連立方程式を解けば，極値を取る点の候補はわかる．

（２）実際に極値になっているかを調べる（講義ノートの定義 3.7.4 と定理 3.7.5）．上を満たす (a,b) の一つ一つについて，ヘッセ行列

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y) = (a,b) \text{ での値} \tag{3.7.16}$$

を定義すると，

- $\det H(a,b) > 0$ かつ $f_{xx}(a,b) > 0$ なら，$f(x,y)$ は (a,b) にて極小
- $\det H(a,b) > 0$ かつ $f_{xx}(a,b) < 0$ なら，$f(x,y)$ は (a,b) にて極大
- $\det H(a,b) < 0$ なら $f(x,y)$ は (a,b) にて極大でも極小でもない
- $\det H(a,b) = 0$ なら極大とも極小とも判定できない（もっと詳しく調べるべし）

3.7.4　３変数以上の極大極小

３変数以上の場合に同様の考察を行うのは，原理的には簡単だが，実際には計算が大変だ．教科書にも載ってないけども，やはり触れない訳には行かない．この場合は線形代数で習いつつあるはずの「行列の対角化と２次形式の標準形」を用いるのが良い．この節の内容はこれまでに述べた２変数の場合もカバーしているので，前節の内容はなくても良い訳だが，n-変数の一般論はそれなりにわかりにくいだろうと考えて，前節を設けた．

（余談） 行列の対角化を習う大きな理由の一つは正にこの極大極小問題にある．つまり，今まで見てきたように，$f_x = f_y = 0$ となるような点の近傍では，テイラー展開の最初の数項だけみておれば大体の振る舞いがわかる．そして，特にテイラー展開の２次の項がゼロでない場合はテイラー展開の２次の項の振る舞いを「行列の対角化と２次形式」の理論で奇麗に理解することができるのだ．

対角化が非常に有用なもう一つの例は，後で習う「陰関数定理」である．この場合，考えている非線形の関数をそのテイラー展開の第１項で近似して考えれば大体良い，という主張がなされる．

この世の中には「線形」の現象は数少ないけども，**線形で近似**することにより本質が理解できる非線形現象も非常に多い．(他の具体例としては，微分方程式の理論，力学系の理論などいくらでもある．) いやむしろ，我々の思考は線形のものとは非常に相性が良いので，**非線形現象の中から線形で理解できる部分を抜き出して**いると言った方が良いかもしれない．ともかく，このような訳で，線形代数は（それ自身も美しい理論ではあるが）応用上も非常に重要なのである．(余談終わり)

定理を述べるのは簡単だが，考え方の方がより大事なので，発見法的にすすむ．いま，C^2-級の n-変数の関数 $f(x_1, x_2, \ldots, x_n)$ を考える．(いつも通り，$\boldsymbol{x} = (x_1, x_2, \ldots, x_n), \boldsymbol{a} = (a_1, a_2, \ldots, a_n)$ である）．これについてテイラーの公式を書くと

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \sum_{j=1}^{n} (x_j - a_j) \frac{\partial f}{\partial x_j}(\boldsymbol{a}) + \frac{1}{2} \sum_{i,j=1}^{n} (x_i - a_i)(x_j - a_j) \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) + o(\|\boldsymbol{x} - \boldsymbol{a}\|^2) \tag{3.7.17}$$

となる．

（１）極値の候補：２変数の場合と全く同じで，$x_j - a_j$ の項は正にも負にもなりうるから，これらの項が残っていては極値にはなり得ない．従って，

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = 0 \qquad (j = 1, 2, \ldots, n) \tag{3.7.18}$$

が $\boldsymbol{x} = \boldsymbol{a}$ であるための**必要条件**である．

（２）上の条件が満たされているとき，$\boldsymbol{x} - \boldsymbol{a}$ の２次の項（＋高次の項）が残る．２次の項は

$$\sum_{i,j=1}^{n} h_i h_j a_{ij} = {}^t\boldsymbol{h} A \boldsymbol{h} \qquad ここで \quad h_i = x_i - a_i, \quad a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}), \tag{3.7.19}$$

の形にかける（$\boldsymbol{h}$ は h_j を集めたベクトル，A は a_{ij} を成分に持つ行列；つまりヘッセ行列そのもの）．２変数の場合を思い出すと，この２次形式 $(\boldsymbol{h}, A\boldsymbol{h})$ が一定の符号を持てば[15]極大や極小，一定の符号を持たなければ極大でも極小でもない，一定の符号を持つか持たないかが判定できないならば情報不足（もっと調べるべし）となる．

という訳で，問題は線形代数の２次形式の問題に帰着された．線形代数の方でもお話があった（ある）はずだが，２次形式の問題は，要するに行列の対角化の応用である．特に今の場合，f が C^2-級だから $a_{ij} = a_{ji}$ となっていて A は実対称行列である．よって A を対角化する直交行列を P と書くと（${}^tPP = P\,{}^tP = E$），

$$B = {}^tP\,A\,P \qquad A = P\,B\,{}^tP \tag{3.7.20}$$

[15] 線形代数で講義されると思うが，２次形式の符号が一定の場合,「定符号の２次形式」という．特にいつでも正（$\boldsymbol{h} \neq \boldsymbol{0}$ ならば $(\boldsymbol{h}, A\boldsymbol{h}) > 0$）の２次形式を**正定値**（positive definite）の２次形式，いつでも負（$\boldsymbol{h} \neq \boldsymbol{0}$ ならば $(\boldsymbol{h}, A\boldsymbol{h}) < 0$）の２次形式を**負定値**（negative definite）の２次形式，という．また，いつでも正とは言い切れないけど負にはならない（すべての $\boldsymbol{h}$ で $(\boldsymbol{h}, A\boldsymbol{h}) \geq 0$）場合，**半正定値**（positive semi-definite）の２次形式という．「２次形式 $(\boldsymbol{h}, A\boldsymbol{h})$ が正定値」というのは,「行列 A の固有値がすべて正」と同値である．また,「２次形式 $(\boldsymbol{h}, A\boldsymbol{h})$ が半正定値」というのは,「行列 A の固有値がすべて非負」と同値である．

を満たす B が対角行列になる．これを用いると

$$(\boldsymbol{h}, A\boldsymbol{h}) = (\boldsymbol{h}, P\,B\,{}^tP\,\boldsymbol{h}) = ({}^tP\,\boldsymbol{h}, B\,{}^tP\,\boldsymbol{h}) = (\boldsymbol{g}, B\boldsymbol{g}) = \sum_{j=1}^{n} \lambda_j (g_j)^2 \tag{3.7.21}$$

と書ける（λ_j は A の固有値，$\boldsymbol{g} = {}^tP\,\boldsymbol{h}$．また B の対角成分は A の固有値 λ_j であることを用いた）．

ここまでくれば，この２次形式の正負は判定できる．

- λ_j がすべて正なら上の和は正であり，$\boldsymbol{x}$ が $\boldsymbol{a}$ に十分近ければ高次の項はこの２次形式よりも小さいので，$f(\boldsymbol{x}) - f(\boldsymbol{a})$ の符号はこの２次形式で決まる．従ってこの場合，$\boldsymbol{x} = \boldsymbol{a}$ が極小である．
- λ_j がすべて負なら上の和は負である．従って上と同様の議論により，$\boldsymbol{x} = \boldsymbol{a}$ が極大である．
- λ_j の中にかつ正のものと負のものが混じっている場合はどうか？わかりやすいように $\lambda_1 > 0$ かつ，$\lambda_n < 0$ の場合を考えよう（他の $\lambda_i > 0$，$\lambda_j < 0$ の場合も同様である）．g_1 のみがゼロでない場合（そのような $\boldsymbol{g}$ を与えるような $\boldsymbol{h}$ は，いつでも $\boldsymbol{h} = P\boldsymbol{g}$ から作れる）はこの２次形式は正であるが，g_n のみがゼロでない場合はこの２次形式は負である．つまり，この２次形式の符号は一定ではない．繰り返し述べたように高次の項はこの２次形式よりも（絶対値が）小さくなるから，２次形式の符号が定まらない今のケースでは極大にも極小にもなり得ない．
- 上のいずれでもない場合，つまり，λ_j は「ゼロまたは正」のみ，または「ゼロまたは負」のみの場合．$\lambda_1 = 0$ だと仮定しよう（他の固有値がゼロなら添字を付け替える）．g_1 のみゼロでない場合，２次形式は丁度ゼロであって，高次の項がどうかがわからない限り $f(\boldsymbol{x}) - f(\boldsymbol{a})$ の符号について結論することができない．つまり，この場合はもっと詳しく調べないとなんとも言えない．

以上をまとめると，以下の定理になる：

> **定理 3.7.6** $\boldsymbol{x} = \boldsymbol{a}$ の近傍で定義された C^2-級の n 変数の関数 $f(\boldsymbol{x})$ について，以下が成り立つ．
> (i) $f(\boldsymbol{x})$ が $\boldsymbol{x} = \boldsymbol{a}$ で極大または極小の場合，$\dfrac{\partial f}{\partial x_j}(\boldsymbol{a}) = 0$（$j = 1, 2, \ldots, n$）である．逆は必ずしもなりたたない（必要条件）．
> (ii) $\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = 0$（$j = 1, 2, \ldots, n$）の場合，$f$ の $\boldsymbol{a}$ におけるヘッセ行列を H と書き，H の固有値を（重複も含めて）$\lambda_1, \lambda_2, \ldots, \lambda_n$ と書く．すると，
>
> a. $\lambda_j > 0$（$j = 1, 2, \ldots, n$）の場合，$f(\boldsymbol{x})$ は $\boldsymbol{x} = \boldsymbol{a}$ で極小である．
> b. $\lambda_j < 0$（$j = 1, 2, \ldots, n$）の場合，$f(\boldsymbol{x})$ は $\boldsymbol{x} = \boldsymbol{a}$ で極大である．
> c. $\lambda_1, \lambda_2, \ldots, \lambda_n$ の中に正のものと負のものが混在している場合（他にゼロがあっても可），f は $\boldsymbol{x} = \boldsymbol{a}$ にて極大でも極小でもあり得ない．
> d. $\lambda_j \geq 0$（$j = 1, 2, \ldots, n$）または $\lambda_j \leq 0$（$j = 1, 2, \ldots, n$）ではあるが，$\lambda_1, \lambda_2, \ldots, \lambda_n$ の中にゼロがある場合，$f(\boldsymbol{x})$ の $\boldsymbol{x} = \boldsymbol{a}$ における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）．もっと調べなければならない．

なお，行列の正定値，負定値を判定するための条件として，以下がある（参考までに載せる；斎藤正彦「線形代数入門」の定理 4.3 と系 4.4 などを参照）．

> **定理 3.7.7** $n \times n$ 行列 A が与えられたとき，$1 \leq k \leq n$ に対して，行列 A の第１行から第 k 行と第１列から第 k 列までを使って $k \times k$ 行列を作り，これを A_k と書く．このとき，行列 A が
>
> a. 正定値であるための必要十分条件はすべての $1 \leq k \leq n$ に対して $\det A_k > 0$ となることである．
> b. 負定値であるための必要十分条件はすべての $1 \leq k \leq n$ に対して $(-1)^k \det A_k > 0$ となることである．

3.8 陰関数定理

さてと，いよいよ「陰関数定理」に入ります．正直，僕はこの項目が大嫌いだ．重要な定理である事は認めるものの，微積の他の題材と異なり，最初は「何が言いたいのかわからない定理」と思い，一旦わかってしまえば今度は「そんなアタリマエの事をやる必要があるのか」と思うだろう（僕自身，一年の時はそう思った）から．と言っていても仕方ないので，やりましょう．すぐの応用としては，この後でやる「ラグランジュの未定定数法」があります．

まずは，何を問題にしているかを規定しよう．３変数以上は極端に大変なので，まずは２変数で考える．

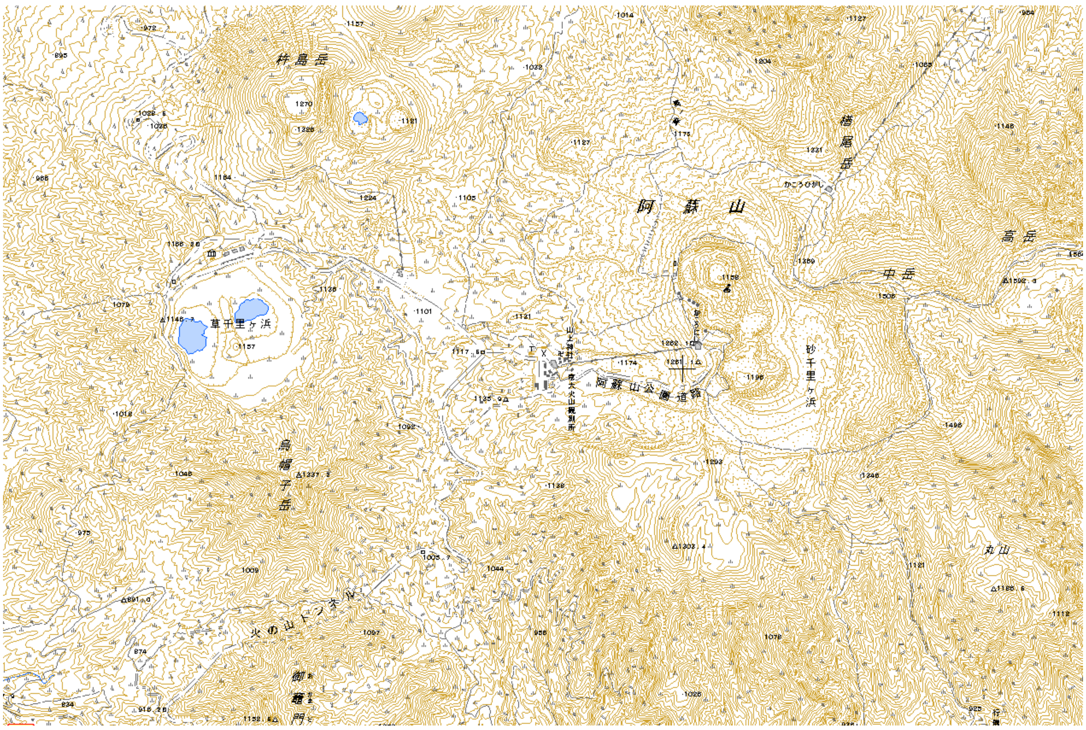
> **問題 3.8.1** xy-平面全体で定義された C^1-級の関数 $f(x,y)$ がある．$f(x,y)=0$ を y について解いて y を x の関数として表せ．別の言い方をすると，$f(x,y)$ の**零点**，つまり $f(x,y)=0$ となる点の集合を求めよ．

$f(x,y)$ が簡単な場合には，これは高校までの知識で解ける．

- $f(x,y)=2x+y-1$ の時は，f の零点は直線 $y=1-2x$ である．
- $f(x,y)=xy$ のとき：零点は $x=0$ または $y=0$，つまり x 軸と y 軸だ．
- $f(x,y)=x^2+y^2-1$ の時：零点は $x^2+y^2=1$ で，単位円だね．無理に書けば $y=\pm\sqrt{1-x^2}$
- $f(x,y)=x^2-y^2+1$ の時：零点は $y^2-x^2=1$ で，双曲線だ．$y=\pm\sqrt{x^2+1}$

上の例では f の零点は何らかの曲線（またはその集まり；直線も曲線の一種と考える）になっていて，そのお陰で $y=y(x)$ の形に表せた．これは「次元」を考えればある程度は自然なことで，もともとの２次元平面（x,y）に条件が一つ（$f=0$）ついたので，その解は次元が一つ下がって「１次元[16]のようなもの」（＝曲線）になるのだ（ろう）．でも，このようなことはより一般の f でも成り立つのだろうか？どのような f なら成り立つのだろうか？実際の問題では $f(x,y)$ が具体的には書き下せない場合も多いから，そのような時にも判定できる条件が欲しい．これに答えるのが陰関数定理である．

定理そのものに入る前に，少し直感的な話をしておく．$z=f(x,y)$ が地点 (x,y) でその土地の標高を表していると思えば，$f(x,y)=C$ （C は定数）というのは標高が C のところの**等高線**である．我々は特に $C=0$（海岸線）を知りたい訳だが，$f(x,y)-C$ を改めて $f(x,y)$ だと思えば（標高を測る原点をずらせば），同じ事である．ともかく，「どのような土地の形ならきれいに等高線が描けるか」が問題になっている訳だ．



[16] このところの「次元」の定義は線形代数でやっている厳密なものからはほど遠く，今の段階ではかなりええ加減な話だ．ただしもちろん現代数学ではこのような「曲がった」ものの「次元」も定義できる

さて，地図を見た事がある人ならわかるように，大抵の場所（なだらかな山の斜面など）にはきれいに等高線が描けている．等高線が描けない（描きにくい）可能性があるのは大体，以下の２つだ：

a. 土地がものすごく平らで，標高 C メートルの**平坦な台地**みたいになっているところ

b. **垂直な崖**が，$C-10$ メートルから $C+10$ メートルまで続いているところ

１つ目の例では $f(x,y)=C$ を満たすところが平面的に広がってしまって,「線」にならない．２つ目の例では標高が $C-10$ から $C+10$ にジャンプしてしまって，丁度 C のところがない．

このような事を（等高線が描ける）十分条件の形にすると，以下の定理になる．この定理では，上の b（崖）の可能性は，f が C^1-級である事を仮定して，始めから排除してある．その上で a の可能性もなければ等高線が描ける，というのが定理の主張であり，直感的には上でやった議論を出ていない（数学的に厳密にできるということはもちろん，凄いことだが）．定理を述べるためにまず，用語を定義する．

定義 3.8.2 xy-平面全体で定義された C^1-級の関数 $f(x,y)$ がある．$f(a,b)=0$ かつ，$f_x(a,b)=f_y(a,b)=0$ となる場合，(a,b) を f の**特異点**という．特異点でない $f(a,b)=0$ となる点は**通常点**という．

すると，

定理 3.8.3（２変数の陰関数定理） xy-平面全体で定義された C^1-級の関数 $f(x,y)$ がある．$f(a,b)=0$ かつ (a,b) が通常点ならば，$f(x,y)=0$ は (a,b) の近傍で一つの曲線を表す．例えば $f_y(a,b)\neq 0$ ならば，$y=\varphi(x)$ が求める曲線になるような C^1-級の関数 $\varphi(x)$ が一意に存在する．すなわち，

$$b=\varphi(a) \quad \text{かつ} \qquad (a,b) \text{ の近傍で } f(x,\varphi(x))=0 \tag{3.8.1}$$

がなりたつ．更に (a,b) の近傍では

$$\frac{d}{dx}\varphi(x)=-\left.\frac{f_x(x,y)}{f_y(x,y)}\right|_{y=\varphi(x)} \tag{3.8.2}$$

もなりたつ．なお，f が C^r-級（$r\geq 1$）なら，$\varphi(x)$ も C^r-級である．($\varphi(x)$ の r-階導関数を f の偏導関数を使って書く事もできるが，ちょっと大変なので略)．

（注意）他の大抵の定理と同様に，この定理も十分条件しか与えていない．(つまり，特異点の周りでも曲線 $y=\varphi(x)$ が定まる事もある.)

定理の形にすれば厳めしいが，要するにみんなの知っている等高線の問題だと思って乗り切る事にしよう．証明は易しくはないが，これも等高線を実際に描くつもりになればわかるのではないかな．

（証明の概略）詳しくは教科書の pp.188-190 を参照．

Step 1. $\varphi(x)$ を実際につくる．$f(x,y)=0$ をみたすような y が存在する事，つまり (3.8.1) をみたすような $\varphi(x)$ が存在する事を，中間値の定理から示せば良い．

Step 2. $\varphi(x)$ が連続である事をいう．連続でなかったとして矛盾を導く．

Step 3. (3.8.1) をみたす $\varphi(x)$ が一意に決まる事をいう．とは言っても，大半は Step 1 で言ってあるのだが...

Step 4. $\varphi(x)$ が C^1-級である事をいって，導関数を計算する．$f(x,y)$ のテイラー展開を用いる．ここは簡単な計算だから，変に覚えようとせずに，各自で再現してみるのが良いだろう． □

３変数以上の，また条件が２つ以上ある場合の陰関数定理については教科書の定理 5.4.3，定理 5.4.4 を参考にして下さい．講義で宣言したように，この講義ではこの題材は深くは扱いません．

3.9 条件付き極値問題：ラグランジュの未定乗数法

実用上は大事な項目ですが，計算はなかなか大変なので，ある程度簡単に済ませます．わかりやすいように２変数の場合をまず考え，一般の場合は後で簡単に触れるにとどめます．

以下の問いを考えたい．

（問１）関数 $f(x,y)$ を，条件 $g(x,y)=0$ の下で最大・最小（極大・極小）にするような (x,y) と，その時の $f(x,y)$ の値を求めよ．

ここで「条件 $g(x,y)=0$ の下に (a,b) で極小」の意味は以下の２つが成り立つ事である．

- $g(a,b)=0$ である．
- $g(x,y)=0$ かつ $(x,y)\neq(a,b)$ であるような，(a,b) に十分近い (x,y) に対しては $f(x,y)>f(a,b)$ である．

このような問題を**条件付き極値（最大最小）問題**という．

（注）以前にも注意したが，最大・最小の問題は極大・極小の問題よりも難しい —— 極大・極小点をすべて求めた上で，考えている領域の境界での値とも比べる必要があるからである．ここでは極大・極小問題に注力する．

このような問題がいままでの極大・極小問題と異なるのは，$g(x,y)=0$ などの条件（拘束条件，constraint）がついていることだ．この条件のため，x,y は独立に動く事ができない．従って,「２変数関数の極値問題」のように単純に偏微分してやる訳にはいかない．

少し気をつければ，今までの知識だけでも「愚直に」解く事は大体，可能だ．つまり $g(x,y)=0$ を y について解いて y を x の関数として表し，それを $f(x,y)$ に代入して $f(x,y)$ を x だけの関数として表す．こうすれば x は自由に動けるから，問題は（高校でやった）１変数関数の極値問題になる．従って，普通に x で微分してやればよい．

（例１）$f(x,y)=x^4+y^4$ の極値を，条件 $x^2+y^2=1$ の下で求めよ．

これなら $y=\pm\sqrt{1-x^2}$ と解いて $f=x^4+(1-x^2)^2=2x^4-2x^2+1=2(x^2-\frac{1}{2})^2+\frac{1}{2}$ となるから，$x=\pm1/\sqrt{2}$ で極小（この場合は最小）になる．極小値は $\frac{1}{2}$．極値をとる (x,y) は $(\pm\frac{1}{\sqrt{2}},\pm\frac{1}{\sqrt{2}})$（複合任意）．

ところが，このようなやり方は往々にして非常に面倒になる．上の例では $g(x,y)$ が簡単だから助かったけど，例えば，$g(x,y)=x^6+3xy-y^2$ だったらどうだろう？$g(x,y)$ が多項式でなく，$\sin,\cos,\log$ などで書かれていたらほとんどお手上げだ．

と言うわけで，応用上，もっと簡便な方法がないとやってられない．これを与えてくれるのが「Lagrange の未定乗数法」である．そのやり方をまず説明しよう（理由はあとで）．

（Lagrange の未定乗数法）上の（問１）の条件付き極値問題を考える．まず，天下りではあるが，新しい変数 λ を導入して

$$F(x,y,\lambda)=f(x,y)-\lambda g(x,y) \tag{3.9.1}$$

を定義する．すると，この条件付き極値問題において，極値を取る点の候補 (x,y) は，以下の (i), (ii) のいずれかである．

(i) $g(x,y)=0$ の特異点,

(ii) 未知変数を x,y,λ とする以下の連立方程式の解．

$$0=\frac{\partial F}{\partial x}=\frac{\partial f}{\partial x}(x,y)-\lambda\frac{\partial g}{\partial x}(x,y),\qquad 0=\frac{\partial F}{\partial y}=\frac{\partial f}{\partial y}(x,y)-\lambda\frac{\partial g}{\partial y}(x,y) \tag{3.9.2}$$

$$0=\frac{\partial F}{\partial \lambda}=g(x,y) \tag{3.9.3}$$

つまり，($g(x,y)=0$ の特異点を除けば）形式的には，この条件付き極値問題は新しく定義した関数 $F(x,y,\lambda)$ の普通の極値問題 —— x,y と λ が自由に動く —— のように見える．

考案者の名前をとって λ を Lagrange の未定乗数（Lagrange multiplier）という．なお，この方法では**極値をとる (x,y) の候補が見つかるだけ**であって，それらが実際の極値を与えるか否かを決める一般論は存在しない．(より正

確には，そのような一般論がない訳ではないが，実用的なものはほとんどない.）ただし，極値点の候補が見つかれば，その点の周りでのテイラー展開などを用いて，実際に極値になっているかどうかの判定は可能な事が多いから，これは実用上は大した問題ではない（少なくとも計算機の助けを借りれば何とかなる）．また，方程式 (3.9.2) と (3.9.3)（やその多変数の場合の該当物）を解くのは大変だと強調している本が多いが，これも計算機の助けを借りればそんなに大した問題ではない（事も多い）．というわけで，未定乗数法はやはり偉大なのである．

具体例：上の（例 1 ）なら，$g(x,y)=0$ の特異点はないので，解くべきは $F(x,y,\lambda)=x^4+y^4+\lambda(x^2+y^2-1)$ を考えて

$$0=4x^3+2\lambda x, \qquad 0=4y^3+2\lambda y, \qquad 0=x^2+y^2-1 \tag{3.9.4}$$

の 3 つである．これを解くと，

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \pm 1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm 1 \\ 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \pm\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \mp\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \text{（ベクトルの中では複合同順）} \tag{3.9.5}$$

となる．後ろの 2 つは変数を消去して解いたものと同じでメデタシメデタシ.（前の 2 つは極値の「候補」ではあったけど，やってみたら極値にはなってなかった，ということ.）

（未定乗数法がうまく行く理由 1 ）

条件 $g(x,y)=0$ が嫌らしいわけだから,「愚直」な方法で解くつもりになって，y を x で表してやろう．これを $y=\varphi(x)$ と書く（実際にこのように表せるかどうかは自明ではないが,「陰関数定理」によって，$g(x,y)=0$ の特異点以外では可能である — 場合によっては $x=\psi(y)$ の形にしか解けない事もあるが）．これを元の f に代入して $h(x)=f(x,\varphi(x))$ を作る．

この $h(x)$ は x のみの関数だから極値の条件は

$$0=h'(x)=f_x+f_y\,\varphi'(x) \tag{3.9.6}$$

となっている（偏微分は $(x,\varphi(x))$ での値）．ところが，$g(x,\varphi(x))=0$ であるから，この両辺を x で微分すると

$$0=\frac{d}{dx}g(x,\varphi(x))=g_x+g_y\,\varphi'(x) \tag{3.9.7}$$

この 2 つから，

$$\varphi'(x)=-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \tag{3.9.8}$$

が導かれるが，これは見方を変えれば

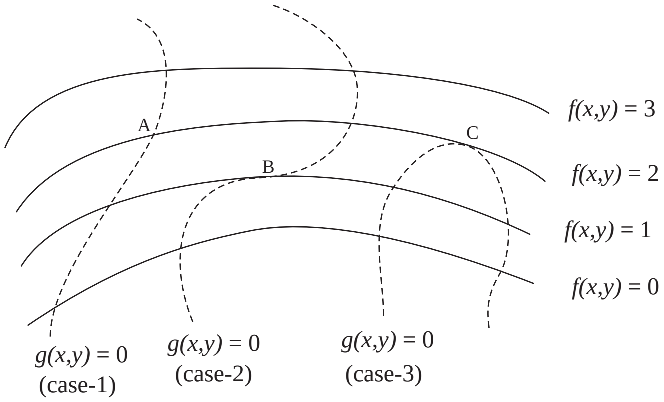
$$\frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{3.9.9}$$

ということであり，この値を λ と書けば，これは (3.9.2) に他ならない.（以上では g_y や f_y などがゼロでないと仮定して分数の形に書いたが，これらがゼロの場合は個別に扱えば大丈夫である事はわかる）． □

（未定乗数法がうまく行く理由 2 — 直感的意味）上の「証明」は愚直な方法で計算してみたらこうなった，というもので，どうも直感的ではない．ここではその直感的な説明を試みる.（以下は「解析概論」などを参考にした.）

陰関数定理を扱ったとき，$g(x,y)=0$ は $g(x,y)=0$ の「等高線」を表していることを指摘した．同様に c を定数として，$f(x,y)=c$ は $f=c$ の等高線を表している．我々の問題は，$g(x,y)=0$ の等高線上で $f(x,y)$ の値を極大（極小）にすること，言い換えれば $g(x,y)=0$ の等高線と $f(x,y)=c$ の等高線の交わりが存在するような c の値を探し，その極大や極小を探すことである．

以下に $f(x,y)=c$ の等高線と $g(x,y)=0$ の等高線の様子を模式的に描いてみた．$f(x,y)=0,1,2,3$ の 4 本の等高線が図の実線，$g(x,y)=0$ の等高線が図の点線である（ただし，3 つの典型的な場合を同じ図の中に描きこんだ）．



通常，$f(x,y)=c$ の等高線と $g(x,y)=0$ の等高線は（接しないで）交わり，図の case-1 のようになっている．この場合，$g(x,y)=0$ の等高線（点線）に沿って進むと，$f(x,y)$ の値は $0,1,2,3$ と増えてくるので，極値はない．

しかし，case-3 の場合には $g(x,y)=0$ に沿って進むと，始めは $f(x,y)=0,1$ と増えて行くが，$f(x,y)=2$ になったのを最高にして，f の値が減少してしまう．つまり，この場合には $f=2$ が極大になっているわけだ．この場合，図でも明らかなように，$f(x,y)=2$ と $g(x,y)=0$ の曲線が点 C で**接している**．

一方，case-2 の場合にも２つの曲線が点 B で接してはいるが，点 B では極値にはなっていない．つまり，接する事は必要条件ではあるが，十分条件ではない．

以上から，点 (a,b) で**極値になるための必要条件は，**$f(x,y)=c$ **と** $g(x,y)=0$ **の曲線が** (a,b) **で接する**事だと予想できる．(もちろん，接線がひけないような曲線の場合には話は別だが．) そこで，２つの曲線が接する条件を具体的に書き下してみよう．そのためには，$f(x,y)=c$ の接線の傾きを知る必要があるが，その答えは既に陰関数定理 3.8.3 の (3.8.2) で与えられている．つまり

$$f(x,y)=c \text{ の接線の傾きは } -\frac{f_x}{f_y}, \qquad g(x,y)=0 \text{ の接線の傾きは } -\frac{g_x}{g_y} \tag{3.9.10}$$

なのだ．従って，両者が接する条件は

$$-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \qquad \text{つまり} \qquad \frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{3.9.11}$$

であるが，これは (3.9.2) に他ならない． □

より一般の条件付き極値問題は以下のようになる（今までのように $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ と書く）：

（問２）n-変数の関数 $f(\boldsymbol{x})$, $g_1(\boldsymbol{x}),g_2(\boldsymbol{x}),\ldots,g_m(\boldsymbol{x})$ がある．$m<n$ として，m この条件 $g_i(\boldsymbol{x})=0$（$i=1,2,\ldots,m$）の下で $f(\boldsymbol{x})$ を最大・最小（極大・極小）にする $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ と，その時の $f(\boldsymbol{x})$ の値を求めよ．

（Lagrange の未定乗数法）上の（問２）の条件付き極値問題を考える．ただし，f,g_i は C^1-級の関数とする．このとき，新しい変数 $\lambda_1,\lambda_2,\ldots,\lambda_m$ を導入して

$$F(\boldsymbol{x},\lambda_1,\lambda_2,\ldots,\lambda_m)=f(\boldsymbol{x})-\Big\{\lambda_1 g_1(\boldsymbol{x})+\lambda_2 g_2(\boldsymbol{x})+\cdots+\lambda_m g_m(\boldsymbol{x})\Big\} \tag{3.9.12}$$

を定義する．すると，この条件付き極値問題において，極値を取る点の候補 $\boldsymbol{x}$ は，以下の (i), (ii) のどちらかを満たす．

(i) $\boldsymbol{x}$ でのヤコビ行列 $\frac{D\boldsymbol{g}}{D\boldsymbol{x}}$ の階数が m より小さい．

(ii) $\boldsymbol{x}$ は未知変数を $\boldsymbol{x}$ および $\lambda_1,\lambda_2,\ldots,\lambda_m$ とする以下の連立方程式を満たす．

$$\begin{aligned} 0&=\frac{\partial F}{\partial x_j}=\frac{\partial f}{\partial x_j}(\boldsymbol{x})-\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial x_j}(\boldsymbol{x}), \qquad (j=1,2,\ldots,n)\\ 0&=\frac{\partial F}{\partial \lambda_k}=g_k(\boldsymbol{x}), \qquad (k=1,2,\ldots,m) \end{aligned} \tag{3.9.13}$$

大雑把に言えば，m 個の条件があった場合には，m 個の未定乗数を導入して，条件が 1 個のときと同じように解けば良いのである．ただし，条件が 1 個の時と同様に，このようにして求めたものはあくまで「極値を取る点の候補」である．これらの候補で実際に極値になっているかどうかの簡単な判定条件はない．

4 級数

これまで，数列 $a_1, a_2, a_3, \ldots$ について，いくつかのこと（上に有界な単調増加数列は収束する，など）を学んだ．一方，高校では以下のような無限和（r は実数）

$$\sum_{n=1}^{\infty} r^n = \frac{1}{1-r} \qquad (|r|<1 \text{ のとき}) \tag{4.0.14}$$

を考えた．また,「テイラーの定理」などに関連して，

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \tag{4.0.15}$$

などの無限和による表現も少しだけ顔をだした．

この講義最後のこの章では，このような無限の和について考える．特に (1) どのような場合にこのような無限和が収束するのか，(2) 無限和の中でも特に重要な「ベキ級数」とはどんなものか，の 2 点を中心に考える．

4.1 問題設定と一般論のまとめ

定義 4.1.1 (教科書の定義 4.1.6) 数列，級数について以下のように定義する．

(1) 数列 $a_1, a_2, a_3, \ldots$ に対して，$\displaystyle\sum_{n=1}^{\infty} a_n$ を**級数**という.（この段階ではこの和の値が何か，などは言ってない.）

(2) $S_k := \sum_{n=1}^{k} a_n$ を第 k 部分和という．

(3) $\displaystyle\lim_{k\to\infty} S_k := \lim_{k\to\infty} \sum_{n=1}^{k} a_n$ が存在するとき，級数 $\displaystyle\sum_{n=1}^{\infty} a_n$ は**収束する**という．また，$S := \displaystyle\lim_{k\to\infty} S_k$ を級数 $\displaystyle\sum_{n=1}^{\infty} a_n$ の**和**といい，$\displaystyle\sum_{n=1}^{\infty} a_n = S$ と書く．

(4) 収束しない級数は**発散する**という．

（注）英語では数列は sequence，級数は series である．

（注）級数を考える場合，項を足して行く順序が重要である．実際，足して行く順序を変えると．級数の値が変わったり収束しなくなったり，などすることが多い.（ただし，絶対収束している級数は足して行く順序によらない．後の定理 4.1.13 などを参照.）

（記号）このノートではところどころで $\displaystyle\sum_{n=1}^{\infty} a_n$ を $\displaystyle\sum_{n} a_n$ と略記する．

定義から直ちにわかること 2 つ：

命題 4.1.2 (教科書の命題 4.1.7) 級数 $\displaystyle\sum_{n=1}^{\infty} a_n$ が収束するならば，$\displaystyle\lim_{n\to\infty} a_n = 0$ である．逆は必ずしも成り立たない．

この命題の証明は簡単だから，各自でやってみてほしい．

命題 4.1.3 (教科書の命題 4.1.8) (1) $\displaystyle\sum_{n=1}^{\infty} a_n = S$ および $\displaystyle\sum_{n=1}^{\infty} b_n = T$ のとき，$\displaystyle\sum_{n=1}^{\infty} (a_n \pm b_n) = S \pm T$ である（複合同順）．

(2) c を実数の定数とする．$\displaystyle\sum_{n=1}^{\infty} a_n = S$ のとき，$\displaystyle\sum_{n=1}^{\infty} (c a_n) = cS$ である．

4.1.1 正項級数

春学期に「単調増加数列」を考え，その性質を学んだ．この小節では単調増加数列の性質をもちいてすぐに理解できる級数を考える．級数 $\sum_n a_n$ の部分和が単調増加数列になるためには，和の中身 a_n が正であることが必要十分である．というわけで，以下の定義をする．

定義 4.1.4 (教科書の定義 4.1.10) 各 n について $a_n \geq 0$ である級数 $\sum_n a_n$ を**正項級数**という．

春学期に単調増加数列について習ったことから，直ちに次が言える．これらの性質は無理に覚えようとしないで，できるだけ証明を理解しようとするのが良い（幸いなことに，証明は高校に毛の生えたレベルだ）．

命題 4.1.5 (教科書の命題 4.1.11, 4.1.12, 4.1.13) この命題では $\sum_n a_n$，$\sum_n b_n$ は正項級数とする．

a. 正項級数 $\sum_n a_n$ が収束 $\iff$ 部分和の作る数列 $S_1, S_2, S_3, \ldots$ が有界

b. 正項級数 $\sum_n a_n$ が収束するとする．また級数 $\sum_n b_n$ があって，a_n と b_n の間には，$b_n \leq c a_n$ $(n = 1, 2, 3, \ldots)$ が成り立っているものとする（c は正の定数）．このとき，$\sum_n b_n$ も収束し，$\sum_n b_n \leq c \sum_n a_n$

c. 正項級数 $\sum_n a_n$ が収束し，$\dfrac{b_{n+1}}{b_n} \leq \dfrac{a_{n+1}}{a_n}$ ならば，$\sum_n b_n$ も収束する．

d. （ダランベールの判定条件; ratio test）$c := \lim_{n\to\infty} \dfrac{a_{n+1}}{a_n}$ が存在する場合，
   - $c < 1$ ならば，$\sum_n a_n$ は収束する．
   - $c > 1$ ならば，$\sum_n a_n$ は発散する．

e. （コーシーの判定条件; root test）$c := \lim_{n\to\infty} \sqrt[n]{a_n}$ が存在する場合，
   - $c < 1$ ならば，$\sum_n a_n$ は収束する．
   - $c > 1$ ならば，$\sum_n a_n$ は発散する．

このように，正項級数については，春学期にやったことの簡単な応用で，すべてわかってしまう．

（重要な例；教科書の 4.1.16）級数 $\sum_{n=1}^{\infty} n^{-s}$ は

- $s > 1$ なら収束する．
- $s \leq 1$ なら発散する．

4.1.2 交代級数（交項級数）

正項級数はまあ簡単だった．次に簡単なものとして，以下の定義をする：

定義 4.1.6 (教科書の定義 4.1.18) n を一つ増やす毎に a_n の符号が変わる場合，$\sum_n a_n$ を**交代級数（交項級数）**という．

交代級数の例としては $\sum_{n=1}^{\infty} \dfrac{(-1)^n}{n}$ などがあり，春学期のレポート問題でも少し考えた．交代級数については，実は以下の一般論（ライプニッツによる）がある：

命題 4.1.7 (教科書の定理 4.1.19) 交代級数 $\sum_n a_n$ を考える．もし a_n が

$$|a_n| \geq |a_{n+1}| \qquad \text{かつ} \qquad \lim_{n\to\infty} a_n = 0 \tag{4.1.1}$$

を満たすならば，交代級数 $\sum_n a_n$ は収束する．

4.1.3 コーシーの判定条件

さて，春学期にやったことで簡単にわかることは以上でおしまい．これでは，(正項や交代でない）一般の級数が収束するのか発散するのかの判定条件が釈然としない．振り返れば春学期にも，(単調増加などでない）一般の数列の収束条件を追いつめないまま，ここまで来てしまっている．ここで本腰を入れて，この条件（コーシーの判定条件）を扱う．ただし，時間の関係もあって，できるだけ簡単にすませる．級数に行く前に，まずは数列に対する「コーシーの条件」について述べよう．

定義 4.1.8 (コーシー列；教科書の定義 4.1.1) 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．

勝手に選んだ（小さい）$\epsilon > 0$ に対し，(十分大きな）整数 $N(\epsilon)$ がとれて，

$$\text{すべての } m, n \geq N(\epsilon) \text{ に対して} \quad \big|a_m - a_n\big| < \epsilon \quad \text{とできる} \tag{4.1.2}$$

(注）上の定義の条件は

$$\lim_{m,n\to\infty} |a_m - a_n| = 0 \tag{4.1.3}$$

と同じことである．つまり，(4.1.3) を満たすような数列がコーシー列ということである．

(注）実はコーシー列という概念は ϵ-δ の次に待ち構えている，大学数学の鬼門である．ただし，この講義では時間の関係もあって深入りはしない．

コーシー列というものをわざわざ定義したのは，ひとえに次の定理のためだ．

定理 4.1.9 (コーシーの収束条件；教科書の命題 4.1.2 と定理 4.1.3 に対応) 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．

この定理が威力を発揮するのは，収束先がわかっていない数列の場合である．この場合，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）を見て，この極限がゼロになること」が収束と同値だ，というのである．収束の必要十分条件を与えてくれるのだから，この定理は非常に強力，かつ重要なものである．

上の定理の証明は教科書を参照されたい．

では，以上のコーシー列の知識を，級数に応用しよう．定義により，級数が収束するとは，その部分和の作る列 $S_k := \sum_{n=1}^{k} a_n$ が収束することだった．この数列 $S_1, S_2, S_3, \ldots$ に上の定理を用いると，以下の定理がすぐに得られる．数列に対する場合と同じく，この定理も級数の収束先の値がわからない時に真価を発揮する．

定理 4.1.10 級数 $\sum_n a_n$ が収束するための必要十分条件は，以下である：

勝手に選んだ（小さい）$\epsilon > 0$ に対し，（十分大きな）整数 $N(\epsilon)$ がとれて，

$$\text{すべての } m > n \geq N(\epsilon) \text{ に対して} \quad \left|\sum_{k=n}^{m} a_k\right| < \epsilon \quad \text{とできる} \tag{4.1.4}$$

（注）上の条件は

$$\lim_{\substack{m,n\to\infty \\ m>n}} \sum_{k=n}^{m} a_k = 0 \tag{4.1.5}$$

と同値である．

4.1.4 絶対収束と条件収束

定義 4.1.11 (絶対収束) 級数 $\sum_n a_n$ に対応して，和の中身を絶対値でおきかえた級数 $\sum_n |a_n|$ を考える．

- 級数 $\sum_n |a_n|$ が収束するとき，元の級数 $\sum_n a_n$ は**絶対収束する**という．
- 級数 $\sum_n |a_n|$ は発散するが，元の級数 $\sum_n a_n$ は収束するとき，元の級数 $\sum_n a_n$ は**条件収束する**という．

絶対収束と言った場合，あくまで元の数列 $\sum_n a_n$ に対して言っているのである．ただし，絶対収束かどうかは，(定義に従って）絶対値をとった方の級数 $\sum_n |a_n|$ で行う．ここのところ，ちょっと混乱しやすいかもしれないので注意．

わざわざこのような概念を定義したのは，以下の定理のためである：

定理 4.1.12 (教科書の定理 4.1.22) 級数 $\sum_n a_n$ が絶対収束するなら，すなわち級数 $\sum_n |a_n|$ が収束するなら，もとの級数 $\sum_n a_n$ も収束する（収束の十分条件）．

上の定理は単なる十分条件ではあるが，符号が一定しない数を足している場合，その級数の収束判定に役立つことが多い．

また，絶対収束する級数については，和が無限ということを忘れて，あたかも普通の有限和のように扱って良い．特に重要な性質をまとめておく：

定理 4.1.13 (教科書の定理 4.1.23) 以下が成り立つ．

- 級数 $\sum_n a_n$ が絶対収束する，つまり，級数 $\sum_n |a_n|$ が収束すると仮定する．この場合，$\{a_n\}$ の順番を好き勝手に並び替えた数列を $\{b_n\}$ とすると，$\sum_n a_n = \sum_n b_n$．つまり，級数の中で，**和をとって行く順序をいろいろと変えても答えは同じ**．
- 級数 $\sum_n a_n$ と $\sum_n b_n$ が両方とも絶対収束すると仮定する．このとき，

$$\sum_{m,n} a_m b_n = \left(\sum_n a_n\right) \times \left(\sum_n b_n\right) \tag{4.1.6}$$

が成り立つ．ここで左辺の m, n の和はどんな順序でとっても構わない．

上の定理の結論はアタリマエに見えるかもしれないが，絶対収束しない級数に対しては一般にはなりたたない．

4.2 べき級数（整級数）

さて，いろいろな級数のうち，微積分にとって非常に重要なものに,「ベキ級数」がある（教科書では「整級数」と呼ばれている）．まず，ベキ級数のついての基本的な事項を列挙しよう．

定義 4.2.1 (ベキ級数または整級数；教科書の定義 4.2.1) 変数 x を含む級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ （x は一応，実数とする．また，a_n は x とは無関係の数列である）を，x の**ベキ級数**（または整級数）という．この級数が収束するような x の全体をこの級数の**収束域**という．

実はベキ級数の理論は x を複素数と思っても，x が実数の場合とほとんど同じように展開できる．しかし，この講義では x が実数と思って話を進める．

命題 4.2.2 ベキ級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ が $x = \alpha$ で収束するならば，この級数は $|x| < |\alpha|$ なるすべての x でも収束する．しかも，その収束は絶対収束である．

定理 4.2.3 (教科書の定理 4.2.4) ベキ級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ について，以下の 3 つのどれか一つが成り立つ．

(a) この級数は全ての実数 x について収束する．

(b) この級数は 0 以外の全ての実数 x について発散する．

(c) ある正の実数 r が存在し，$|x| < r$ では絶対収束，$|x| > r$ では発散する．

上の (c) の場合，$x = \pm r$ については何も主張していないことに注意．実際，$x = r$ または $x = -r$ で収束する級数もあれば，発散する級数もある．

定義 4.2.4 (ベキ級数の収束半径；教科書の定義 4.2.5) 上の (c) の r をベキ級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ の**収束半径**という．なお，上の (a) の場合には収束半径は無限大，(b) の場合には収束半径はゼロ，と約束する．

（注）収束「半径」という理由は，x を複素数に拡大した場合を想定しているからである．x が複素数の場合でも，定理 4.2.3 は成り立ち，r を収束半径とすると，$|x| < r$ を満たす全ての複素数では級数が収束，$|x| > r$ では級数が発散，となる．複素平面で $|x| < r$ なる x を表すと，これは半径 r の円の内部になる．つまり，r は収束域を表す円の半径になっているのである．

さて，一般論の命題 4.1.5 を用いると，直ちに以下が得られる．

命題 4.2.5 ベキ級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ の収束半径 r は以下を満たす．

(1) $\displaystyle b := \lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}$ が存在するならば，$r = \dfrac{1}{b}$

(2) $\displaystyle b := \lim_{n\to\infty} \sqrt[n]{|a_n|}$ が存在するならば，$r = \dfrac{1}{b}$

いずれの場合も，$b = 0$ なら $r = +\infty$，$b = +\infty$ なら $r = 0$ と解釈する．

この命題はベキ級数の収束半径を求める際に非常に有効であるから知っていて損はないだろう．

さて，このようなベキ級数が有効なのは，(1) これがいろいろな関数を表すのに利用できること，(2) さらに，ベキ級数は（少なくとも形式的には）いつでも微分や積分ができること，にある．

4.2.1 テイラー級数

既に春学期に「テイラーの定理」をやった：0 を内部に含む開区間で関数 f が何回でも微分可能とすると，

$$f(x)=\sum_{k=0}^{n-1}\frac{f^{(k)}(0)}{k!}x^k+R_n(x),\qquad R_n(x)=\frac{f^{(n)}(\theta x)}{n!}x^n \tag{4.2.1}$$

ここで θ は x によって決まる，0 と 1 の間の数であった．

もし，上の $R_n(x)$ が $n\to\infty$ でゼロに行くなら，

$$f(x)=\sum_{k=0}^{\infty}\frac{f^{(k)}(0)}{k!}x^k \tag{4.2.2}$$

となる．つまり，関数 f をベキ級数の形に表せるのである．逆に，(4.2.2) が成り立つ時には，もちろん，$\lim_{n\to\infty}R_n(x)=0$ である．つまり，f が (4.2.2) の形に表せることの必要十分条件が $\lim_{n\to\infty}R_n(x)=0$ なのである．

（用語の注）f が (4.2.2) の形に表せる場合，f はそのテイラー級数に展開できる，という．また，(4.2.2) 右辺の級数を f の（$x=0$ の周りの）テイラー級数という．なお，$x=0$ の周りのテイラー級数はマクローリン級数ともいう．

（注意）(4.2.2) 右辺の級数が収束する場合でも，これが元の関数 $f(x)$ に等しいとは限らない．例えば，

$$f(x):=\begin{cases}\exp(-x^{-2}) & (x\neq 0)\\ 0 & (x=0)\end{cases} \tag{4.2.3}$$

を考えてみよ．

f がそのテイラー級数に展開できる十分条件の例には，教科書の定理 4.2.12, 4.2.13 などがある．これらは今は覚える必要は全くない．将来，必要になったら教科書を見直してほしい．

4.2.2 項別微分と項別積分

以下の有限個の和で表された $f(x)$ を考えよう（n は正の整数)：

$$f(x)=\sum_{k=0}^{n}a_k\,x^k \tag{4.2.4}$$

有限個の和であれば，自由に微分，積分ができる．つまり，

$$f'(x)=\sum_{k=0}^{n}ka_k\,x^{k-1}\qquad \text{（ただし，}k=0\text{ の項はもちろん，ゼロ）} \tag{4.2.5}$$

$$\int_0^y f(x)dx=\sum_{k=0}^{n}a_k\,\frac{y^{k+1}}{k+1}\qquad \text{（}y\text{ は任意の実数）} \tag{4.2.6}$$

$f(x)$ が無限個の和で表されている場合（特にそのテイラー級数に展開できる場合)，同様のことは成り立つだろうか？もし成り立つとすれば，そのような関数の微積分は（級数の形で）いつでも簡単にできるから，非常に嬉しい！

上のような問題については，以下の定理群が答えを与えてくれる．まず，級数 $\sum_{n=0}^{\infty}a_nx^n$ とそれを形式的に微分積分して得られる級数の収束半径について

定理 4.2.6 (教科書の定理 4.2.16) ベキ級数 $\sum_{n=0}^{\infty}a_nx^n$ の収束半径が r のとき，2 つの級数

$$\sum_{n=1}^{\infty}na_nx^{n-1},\qquad \sum_{n=0}^{\infty}\frac{a_n}{n+1}x^{n+1} \tag{4.2.7}$$

の収束半径も r である．

定理が扱っている級数は，元の級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ の各項をそれぞれ微分，積分した形をしている．上の定理は，この新しい級数も，元の級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ と同じ収束半径を持つことを主張している．

もちろん，このように作った新しい級数が，もとの級数の微分や積分に等しいかどうかは自明ではない．これについては，以下の定理がある．

定理 4.2.7 (項別微積分；教科書の定理 4.2.17) ベキ級数 $\displaystyle\sum_{n=0}^{\infty} a_n x^n$ の収束半径が r のとき，$|x| < r$ なる x に対して

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \tag{4.2.8}$$

とおく．すると，

(1) $|x| < r$ の範囲では，f は x の連続関数である．

(2) $|x| < r$ の範囲では，f は x について積分可能で，

$$\int_0^x f(t)dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} \tag{4.2.9}$$

(3) $|x| < r$ の範囲では，f は x に関して微分可能で，

$$f'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \tag{4.2.10}$$

つまり，級数が絶対収束している $|x| < r$ の範囲では，級数の各項を微分・積分することで，元の級数の微分・積分がわかるのである．これを項別微分・項別積分という．通常，関数を積分することは大変で，結果を初等関数で表せないことも多い．ところが，項別積分はいつでもできるから，これは実用上，非常に有効なのだ．

なお，$f(x)$ が (4.2.8) の形に表される場合の右辺は f のテイラー級数になっている可能性が高いが，実際にそうである：

定理 4.2.8 (項別微積分；教科書の定理 4.2.18 の一部) 上の定理においては，

$$a_n = \frac{f^{(n)}(0)}{n!} \tag{4.2.11}$$

である．つまりこの場合，

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \tag{4.2.12}$$

であり，(4.2.8) は f のテイラー級数に他ならないことがわかる．

以上で級数についての重要な事実はすべて網羅した．教科書の 4.2 節には更に,「収束域の端点での様子」（教科書 pp.159–161）が載っているが，これは少しマニアックなのでこの講義では触れない．将来，必要があったら教科書を見直して下さい．