微分積分学の講義参考資料*

原　隆

九大数理

hara@math.kyushu-u.ac.jp

Last updated: June 27, 2014

概 要

これは上記科目のための講義の参考資料（暫定版）です．使用している教科書が非常に「実践的」なため，数学的内容に物足りない人のために，ここにおきます．**これは参考資料なので，この内容を理解することは要求しません．実際，かなり程度は高いです．あくまで，個人の興味に応じて使ってください．**

なお，この参考資料はまだまだ作成中です —— 現在，偏微分くらいまでが出来て来ましたが，そのあとの積分，重積分は全く載っていません．載っていない部分は講義の進行とともに充足されて行く予定なので，当分は**あくまで暫定版である**との理解の下でご利用ください．

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

目 次

*2014 年度春学期，毎週金曜 4 限，基幹教育 1 年理系 12+13 クラス

0　記号のお約束

$a < b$ を２つの実数，n を非負（負でない）整数とする．

- 整数の全体は $\mathbb{Z}$，自然数（1 以上の整数）の全体を $\mathbb{N}$，有理数の全体を $\mathbb{Q}$，実数の全体を $\mathbb{R}$，複素数の全体を $\mathbb{C}$ と書く．例えば，$x \in \mathbb{R}$ と書けば，「x は実数」と同じことである．
- 集合 A の要素を大学では「元（げん）」ともいう．(例) 2 は $\mathbb{Z}$ の元である．$\sqrt{2}$ は $\mathbb{Q}$ の元ではない．
- 高校までと異なり，「$a < b$ または $a = b$」を $a \le b$ と書く（不等号の下が２本線ではなく，１本線）．同様に，「$a > b$ または $a = b$」を $a \ge b$ と書く．

[1] 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．
[2] この小節の内容は「進んだ話題」なので余裕のない人はとばしても良いし，講義でも触れない可能性が高い．

- $a < x < b$ なるすべての実数の集合を (a, b) と書き，開区間という．
  $a \le x \le b$ なるすべての実数の集合を $[a, b]$ と書き，閉区間という．
- 高校と同じく，$n! = n \cdot (n-1) \cdot (n-2) \cdots 2 \cdot 1$ は n の階乗である．ただし，$0! = 1$ と約束する．

「等号」には「両辺が等しい」の他に「片方をもう片方で定義する」の意味もある．この講義では両者を区別して

- $a = b$ は単に a と b が等しいことを表すために
- $a := b$ は左辺の a を右辺の b と定義するために

用いる．例えば，$f(x) := x \sin x$ とあれば，これは函数 $f(x)$ を $x \sin x$ という函数だと定義する，の意味である．

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

（漢字の注）昔は関数を函数と書いた．僕はこの癖が残っているので，所々で「函数」と書いてしまうかもしれないが，これは高校までの「関数」と同じ意味である．

1　1変数函数の微分

まずは（高校でやったことになっている）微分の基礎付けを簡単に行う．そのあとで，高校ではやらなかった新しい題材も学習する（テイラー展開）．

微分を考える理由には大きく分けて２通りある．

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

このうち，第一の視点は受験などを通して散々やってきたものと思うので，この講義では簡単にすませる．第２の「テイラー展開」は，現在の高校のカリキュラムにはないので，この重要なテーマをマスターするのが微分に関する大きな目標の一つになる．

1.1　微分の定義（教科書 3.1 節）

定義 1.1.1 (微分係数) $x=a$ とその近傍で定義されている関数 $f(x)$ に対して，極限

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a} \tag{1.1.1}$$

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

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

1.2　導関数（教科書 3.2 節）

教科書 3.2 節の内容．積の微分，商の微分など．すべて高校で習っているはずなので，ここでは略．

1.3　合成関数の微分（教科書 3.3 節）

$y=f(g(x))$ の時，$y'=f'(g(x))\,g'(x)$．高校でやったはずなのでここでは略．各自で復習すること．

1.4　逆関数の微分（教科書 3.4 節）

$y=f^{-1}(x)$ の微分は $y'=\dfrac{1}{f'\bigl(f^{-1}(x)\bigr)}$．教科書に詳しく載っているので，ここにはこれ以上書かない．

1.5　高階導関数（教科書 3.6 節）

高校でもやったと思うけど，高階の導関数についてまとめておく．

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

$$f^{(2)}(x)=\frac{d^2}{dx^2}f(x)=\frac{d}{dx}\Bigl\{\frac{d}{dx}f(x)\Bigr\},\qquad f^{(3)}(x)=\frac{d^3}{dx^3}f(x)=\frac{d}{dx}\Bigl[\frac{d}{dx}\Bigl\{\frac{d}{dx}f(x)\Bigr\}\Bigr],\quad \ldots \tag{1.5.1}$$

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

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

$$\frac{d}{dx}\{f(x)g(x)\} = f'(x)g(x) + f(x)g'(x), \qquad \frac{d^2}{dx^2}\{f(x)g(x)\} = f''(x)g(x) + 2f'(x)g'(x) + f(x)g''(x) \tag{1.5.2}$$

で，より一般には（n は自然数）

$$\frac{d^n}{dx^n}\{f(x)g(x)\} = \sum_{k=0}^{n}\binom{n}{k} f^{(k)}(x)\, g^{(n-k)}(x), \qquad \binom{n}{k} := {}_nC_k = \frac{n!}{k!\,(n-k)!} \tag{1.5.3}$$

となる [3]．この証明は数学的帰納法でできるから，一度は自力でやっておくこと．ただし，その途中で恒等式

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1} \tag{1.5.4}$$

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

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

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

1.6　平均値の定理（教科書にはないが重要なので加えた）

まずはロルの定理．定理の下の左側の図を見れば，直感的には明らかだろう．

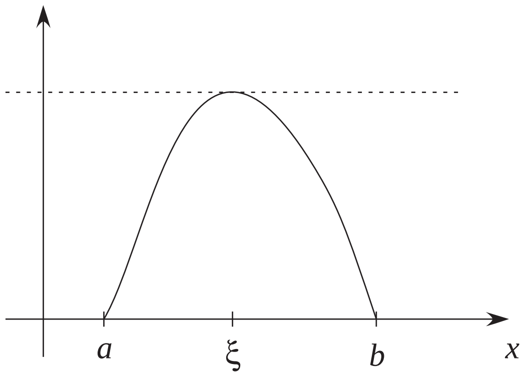
定理 1.6.1 (ロル Rolle の定理) $f(x)$ が閉区間 $[a, b]$ で連続，開区間 (a, b) で微分可能．更に $f(a) = f(b)$ とする．このとき

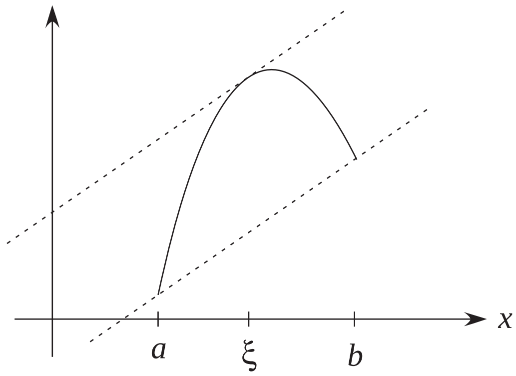
$$f'(\xi) = 0 \qquad (a < \xi < b) \tag{1.6.1}$$

となる ξ が存在する．

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

（証明の説明）連続関数の（当たり前ではない）性質を用いる必要があるので，ここでは略．図を見て直感的には明らかなことを理解していれば，ひとまずは十分である．





ロルの定理からすぐに次の（Lagrange による）平均値の定理が出る．上の図では右側の状況である．

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

定理 1.6.2 (平均値の定理) $f(x)$ が閉区間 $[a,b]$ で連続，開区間 (a,b) で微分可能と仮定する．このとき，

$$\frac{f(b)-f(a)}{b-a} = f'(\xi) \qquad (a<\xi<b) \tag{1.6.2}$$

となる ξ が存在する．

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

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

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

定理 1.6.3 (コーシーの平均値の定理) $f(x)$ と $g(x)$ が共に閉区間 $[a,b]$ で連続，開区間 (a,b) で微分可能とする．更に，(a,b) では $g'(x)\neq 0$ としよう．このとき，

$$\frac{f(b)-f(a)}{g(b)-g(a)} = \frac{f'(\xi)}{g'(\xi)} \qquad (a<\xi<b) \tag{1.6.3}$$

となる ξ が存在する．

（注）$g'(x)\neq 0$ なので，Rolle の定理から $g(a)\neq g(b)$ が保証されている．

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

1.7　テイラーの定理とテイラー展開（教科書 3.7 節）

これから暫く，微分の重要な応用.「テイラー展開」を扱う．これは案外，皆さん苦労するようだから，甘く見ないように.「テイラー展開」とは大雑把にいうと，$f(x)$ の値を $f(a)$ とその高階微係数で表す表式で，

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n \tag{1.7.1}$$

という形をしている（この表式の導出は後でじっくりやる）．皆さんの良く知っている関数の例では（上で $a=0$ としたものを書いた）

$$e^x = 1+x+\frac{x^2}{2}+\frac{x^3}{3!}+\frac{x^4}{4!}+\cdots = \sum_{n=0}^{\infty}\frac{1}{n!}\,x^n \tag{1.7.2}$$

$$\sin x = x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\cdots = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\,x^{2n+1} \tag{1.7.3}$$

$$\cos x = 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\cdots = \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\,x^{2n} \tag{1.7.4}$$

などとなる．

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

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

- まず，(1.7.2) などの式は，それ自身が数値計算にも適している —— $e^x, \sin x$ などの値を，右辺の級数（和）で計算できるのだ．もちろん，無限級数の値そのものを数値的に求める事はできないが，たくさんの項の和をとる事で，いくらでも精度良く計算できる [4]．
- (1.7.1) にはもう少し理論的な意味もある．つまり，$|x-a|$ が小さい場合に $f(x)$ を $f(a)$ で近似すると，誤差がどうなるかを表していると解釈できる．この誤差の評価は，もっと進んだ結果を得るのに不可欠である．

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

1.7.1　テイラーの公式（有限項でとめた形）

通常，テイラーの定理（テイラーの公式）というのは以下の形の定理をいう：

定理 1.7.1 ($f(x)$ の $x=a$ の周りのテイラーの公式) $f(x)$ がある開区間 I で n 回微分可能と仮定し，この区間内に $a \in I$ をとろう．このとき，勝手な $x \in I$ に対して，a と x の間の一点 ξ が存在して以下が成り立つ：

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(n)}(\xi)}{n!}(x-a)^n \tag{1.7.5}$$

なお，(1.7.5) の２つの項に名前をつけて

$$f(x) = S_n(x) + R_n(x), \tag{1.7.6}$$

$$S_n(x) := f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad R_n(x) := \frac{f^{(n)}(\xi)}{n!}(x-a)^n \tag{1.7.7}$$

と書く事もある．$S_n(x)$ をテイラー展開の n 次の <u>主要項</u>，$R_n(x)$ を n 次の <u>剰余項</u>（残項）という．

- $a=0$ とした場合の展開を特にマクローリン（Maclaurin）の公式（展開）ともいう．
- 実はマクローリンの公式とテイラーの公式は非常に近い親戚関係にあり，片方だけわかれば十分だ．理由は以下の通り：$y = x - a$ という変数変換によって，座標 x で見た時の点 $x = a$ は座標 y で見た時の $y = 0$ に移る．従って，座標 y でのマクローリンの公式は座標 x での $x = a$ の周りのテイラーの公式に対応している．
- テイラーの公式でも，平均値の定理でも，ξ は a と x（または b）の両方に依存しうることを再度強調しておく．同じ理由で，剰余項 $R_n(x)$ は x, a で決まるけども，$R_n(x)$ の ξ そのものが x, a に依存する事をお忘れなく．
- 細かいことであるが，定理 1.7.1 では $f^{(n)}(x)$ の存在は仮定するが，連続性は仮定しなくても良い．この点で，剰余項が積分形の定理 1.7.2（後出）より，こちらの方が少しだけ適用範囲はひろい（そのぶん，誤差評価は大抵，劣る —「ある ξ が存在して」とか言われても，どんな ξ かわからなければ細かい評価はできない）．

定理 1.7.1 の証明

$$F(x) := f(x) - \left[f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k \right], \qquad G(x) := (x-a)^n \tag{1.7.8}$$

とおく．$F(x)$ が (1.7.6) の $R_n(x)$ の表式で書けることを示せばよい．

そのために，コーシーの平均値の定理（定理 1.6.3）を F, G に適用する事を考えよう．$F(x)$ は $f(x)$ から $(x-a)^k$ の和を引いているだけなので，また $G(x)$ は多項式なので，共に n 階は微分できる．微分を具体的に計算すると

$$F(a) = F'(a) = F''(a) = \ldots = F^{(n-1)}(a) = 0, \qquad F^{(n)}(a) = f^{(n)}(a) \tag{1.7.9}$$

[4] 実際にコンピューターが $e^x, \sin x$ などを計算する場合には，上の (1.7.2) そのものではなく，これを更に効率よくしたものを用いる．しかし，計算の原理は（大体）同じである

$$G(a) = G'(a) = G''(a) = \ldots = G^{(n-1)}(a) = 0, \qquad G^{(n)}(a) = n! \tag{1.7.10}$$

となっている．この事実を用いて，以下のように進む．

（1）定理 1.6.3 そのもので

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(\xi_1)}{G'(\xi_1)} \tag{1.7.11}$$

を満たす ξ_1 の存在（ξ_1 は a と x の間にある）が言える．

（2）上の右辺の量は $F'(a) = G'(a) = 0$ を用いて強引に書き直すと，定理 1.6.3 が使える．その結果，

$$\frac{F'(\xi_1)}{G'(\xi_1)} = \frac{F'(\xi_1) - F'(a)}{G'(\xi_1) - G'(a)} = \frac{F''(\xi_2)}{G''(\xi_2)} \tag{1.7.12}$$

を満たす ξ_2 の存在（ξ_2 は a と ξ_1 の間にある）が言える．

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

$$\frac{F^{(k)}(\xi_k)}{G^{(k)}(\xi_k)} = \frac{F^{(k)}(\xi_k) - F^{(k)}(a)}{G^{(k)}(\xi_k) - G^{(k)}(a)} = \frac{F^{(k+1)}(\xi_{k+1})}{G^{(k+1)}(\xi_{k+1})} \tag{1.7.13}$$

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

（4）以上をまとめると，

$$\frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F^{(n)}(\xi_n)}{G^{(n)}(\xi_n)} \tag{1.7.14}$$

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

$$\frac{F(x)}{(x-a)^n} = \frac{f^{(n)}(\xi_n)}{n!} \tag{1.7.15}$$

となっているので，分母を払うと定理が得られる．　□

1.7.2　テイラー展開（無限項まで）

定理 1.7.1 において，公式 (1.7.6) がすべての $n \ge 1$ で成り立ち，かつ剰余項 $R_n(x)$ が $n \to \infty$ でゼロになるならば，つまり，$\lim_{n\to\infty} R_n(x) = 0$ ならば，

$$f(x) = \lim_{n\to\infty} S_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{1.7.16}$$

が得られる．

> ここのところ，$\lim_{n\to\infty} S_n$ が存在するのかどうか気になる人がいるかもしれないが，それは**「R_n の極限がゼロ」の仮定の下では**以下のように保証される：(1.7.6) の左辺は n に依存せず，右辺では R_n がゼロに行く．従って，残りの S_n の $n \to \infty$ 極限が存在して，かつその極限は左辺の $f(x)$ に等しくなければならない．

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

1.7.3　テイラーの公式，テイラー展開の例

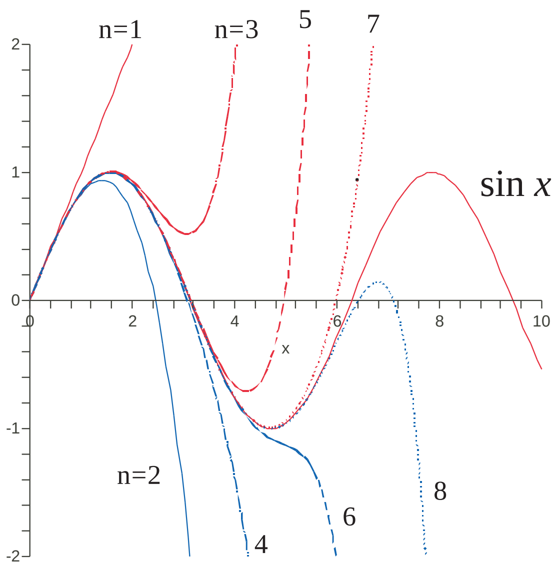
まずは具体例を見てみよう．もう少し「理論的」なことは後で詳しく見る．

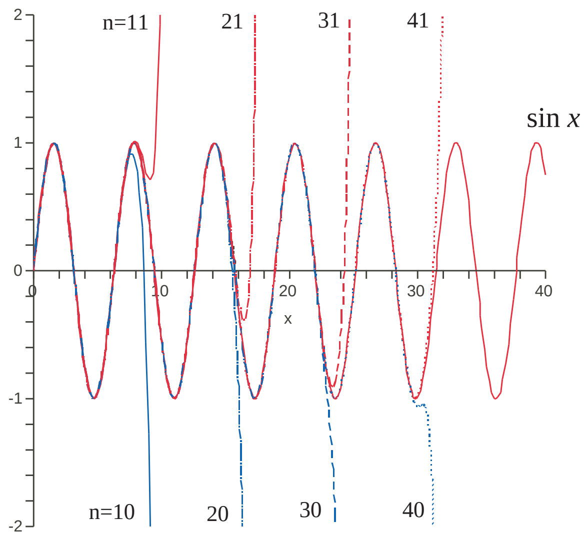
- まず，多項式．$f(x) = c_n(x-a)^n + c_{n-1}(x-a)^{n-1} + \ldots + c_1(x-a) + c_0$ は何回でも微分可能であり，既にテイラー展開の形になっている．念のため，テイラーの公式を用いたら多項式が再現される事を各自で確かめてみよう．

- 指数関数．$f(x) = e^x$ は何回でも微分可能で，高階の導関数もすべて e^x である．従って，特に $a = 0$ としたテイラーの公式から
$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x), \qquad R_n(x) := \frac{e^\xi}{n!} x^n \tag{1.7.17}$$
が得られる（ξ は 0 と x の間の数）．更に，少しややこしい計算を頑張ってやると，すべての実数 x に対して $\lim_{n\to\infty} R_n(x) = 0$ が証明できる．従って，すべての実数 x に対して
$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{1.7.18}$$
が成り立つ．このテイラー級数の形は非常に基本的だから，覚えておくことが望ましい．
- 三角関数（$\sin, \cos$）も同様にして展開式を導くことができる．例えば
$$\sin x = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad R_n(x) := \frac{(-1)^n \sin\xi}{(2n)!} x^{2n} \tag{1.7.19}$$
がなりたつ．指数関数と同様に，この場合もすべての実数 x に対して $\lim_{n\to\infty} R_n(x) = 0$ が証明できる．従って，すべての実数 x に対して
$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad \text{また同様の考察により} \qquad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{1.7.20}$$
が成り立つことがわかる．このテイラー級数の形も覚えてしまうくらいになろう [5]．

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

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





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

1.7.4　テイラー展開の効用

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

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

1.7.5　オイラーの公式

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

$$e^{i\theta} = \cos\theta + i\sin\theta, \qquad \theta \in \mathbb{R} \tag{1.7.21}$$

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

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell+1}}{(2\ell+1)!} \tag{1.7.22}$$

が得られる（２番目の等号は，単に k が偶数の場合と奇数の場合をわけて，i^k を計算しただけ）．ところがこの最右辺は $\cos\theta + i\sin\theta$ のテイラー展開に他ならない．従って，指数関数や三角関数はそのテイラー展開の式で定義し直すのだと思えば，オイラーの公式が証明されたことになる．テイラー展開によって関数を定義し直すというのは一見，奇妙に思えるかもしれないが，同値な命題がある場合にどれを仮定（公理）にしてどれを結論とするか，の一例と思えば良い．ただし，本当に定義し直す立場をとった場合は今まで知っていたはずの関数の性質（例：$\sin, \cos$ は周期 2π である，指数関数は $e^{a+b} = e^a e^b$ を満たす，等々）はすべて忘れて，テイラー展開だけからこれらを導き直す必要はある．この辺りは夏休みチャレンジ問題とする予定なので，我こそはと思う人は挑戦して欲しい．

1.7.6　おまけ：剰余項が積分の形のテイラーの定理

今までのものの他に，テイラーの公式には以下のようなバージョンもある．これは剰余項を積分で書くもので，剰余項の大きさを評価するには楽な事が多い．

定理 1.7.2 (剰余項が積分形のテイラーの公式) $f(x)$ がある開区間 I で C^n-級であると仮定する．この区間 I 内に $a \in I$ をとろう．このとき，勝手な $x \in I$ について，以下が成り立つ：

$$f(x) = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad R_n(x) := \int_a^x \frac{f^{(n)}(y)}{(n-1)!}(x-y)^{n-1}\,dy \tag{1.7.23}$$

（高校のノリでの証明；積分の基礎付けさえすれば，この証明は厳密に正しい）数学的帰納法で証明する．つまり $f(x)$ は C^N-級と仮定し，(1.7.23) をすべての $n \le N$ について証明することを目指す．それで n についての帰納法を用いる．

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

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

$$f(x)=f(a)+\int_a^x f'(y)dy \tag{1.7.24}$$

の第２項を，以下のように部分積分するとよい．

$$\begin{aligned}\int_a^x f'(y)dy &= \int_a^x \Bigl\{-\frac{d}{dy}(x-y)\Bigr\} f'(y)dy = \Bigl[-(x-y)f'(y)\Bigr]_a^x + \int_a^x (x-y)f''(y)dy \\ &= (x-a)f'(a)+\int_a^x (x-y)f''(y)dy \end{aligned}\tag{1.7.25}$$

II. n まで証明できたとして，$n+1$ をやってみよう（もちろん，$n \le N-1$ と仮定しておく）．n までできたと仮定したので，(1.7.23) が成り立っているが，最後の項を以下のように考えて部分積分する（分母の $(n-1)!$ は後で）：

$$\begin{aligned}\int_a^x f^{(n)}(y)(x-y)^{n-1}dy &= \int_a^x f^{(n)}(y)\Bigl\{-\frac{1}{n}\frac{d}{dy}(x-y)^n\Bigr\}dy \\ &= -\frac{1}{n}\Bigl[f^{(n)}(y)\,(x-y)^n\Bigr]_a^x + \frac{1}{n}\int_a^x f^{(n+1)}(y)\,(x-y)^n dy \\ &= \frac{1}{n}f^{(n)}(a)\,(x-a)^n + \frac{1}{n}\int_a^x f^{(n+1)}(y)\,(x-y)^n dy. \end{aligned}\tag{1.7.26}$$

これを (1.7.23) の最後の項に用いると（もちろん，分母の $(n-1)!$ を忘れない），(1.7.23) の $n+1$ のものが証明されてメデタシメデタシ．　□

1.7.7　おまけ：テイラーの公式の意味（関数の近似）

そもそも，テイラーの公式は

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

いう精神の下に生まれたものである．つまり，後述する条件の下では，(1.7.6) での $S_n(x)$ が $f(x)$ を良く近似し，$R_n(x)$ の方は小さな誤差項とみなせるのだ．

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

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

定義 1.7.3 (n 次より高く近似) $x=0$ の近くで定義された関数 $f(x), g(x)$ があり，

$$x\to 0 \ \text{のときに} \qquad \lim_{x\to 0}\frac{f(x)-g(x)}{x^n}=0 \qquad (n\ \text{は正の整数}) \tag{1.7.27}$$

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

上の式では，$f(x)-g(x)$ はゼロに行くのだが，その行き方（ゼロへの収束の速さ）が，x^n よりも速い，と言っているのである．このような事情をうまく表すため，以下のような書き方を導入する [6]

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

定義 1.7.4 (無限小の比較；オーダーまたはランダウの記号) $\lim_{x\to a} f(x) = \lim_{x\to a} h(x) = 0$ とする．

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

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

$$\exists K > 0 \quad \exists\delta > 0 \qquad \left(0 < |x-a| < \delta \quad\Longrightarrow\quad \left|\frac{f(x)}{h(x)}\right| < K\right) \tag{1.7.28}$$

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

(注)

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

この書き方によると，(1.7.27) は

$$f(x) - g(x) = o(x^n) \qquad (n \text{ 次より高く近似．ここの } o \text{ は小文字}) \tag{1.7.29}$$

と書ける．

この用語法に従うと，テイラーの定理を以下のように言い換えることができる．

命題 1.7.5 (テイラーの定理の言い換え) 関数 $f(x)$ の $x = 0$ を中心とした n 次のテイラーの公式

$$f(x) = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k, \quad R_n(x) := \frac{f^{(n)}(\theta x)}{n!} x^n \quad (0 < \theta < 1) \tag{1.7.30}$$

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

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + o(x^{n-1}) \tag{1.7.31}$$

となるための必要充分条件は以下の通り：

$$\lim_{x\to 0} \frac{R_n(x)}{x^{n-1}} = 0 \tag{1.7.32}$$

前の命題の (1.7.32) の十分条件として，以下がある．

命題 1.7.6 (多項式近似の十分条件)
1)　0 を内部に含むある区間で $f^{(n)}$ が有界，つまり $\delta > 0$ と $M > 0$ があって（これらは n に依存してもよいが，x に依存してはいけない），

$$|x| < \delta \text{ ならば } \left|f^{(n)}(x)\right| < M \tag{1.7.33}$$

となっているとする．このとき，

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + O(x^n) \tag{1.7.34}$$

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

これらはわざわざ命題とするほどのことではないかもしれないが，実用上大事だから載せた．特に，一年生で出てくる関数は C^∞-級（何回でも微分できる）のものが多く，これらに対しては上の十分条件が自動的に満たされており，命題 1.7.5 の結論も成り立つのである．

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

命題 1.7.7 (多項式近似の一意性) 原点の近くで定義された函数 $f(x)$ と多項式 $g(x) = \sum_{j=0}^{n} a_j x^j$ があり，$g(x)$ は $f(x)$ を n 次より高く近似しているものとする．このとき，$g(x)$ の係数 $a_0, a_1, \ldots, a_n$ は一意に決まる．

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

証明：
f を n 次より高く近似する g が２つあったとして，それらを

$$g_1(x) = \sum_{j=0}^{n} a_j x^j, \qquad g_2(x) = \sum_{j=0}^{n} b_j x^j \tag{1.7.35}$$

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

$$\frac{|g_1(x) - g_2(x)|}{|x|^n} \le \frac{|g_1(x) - f(x)|}{|x|^n} + \frac{|f(x) - g_2(x)|}{|x|^n} \tag{1.7.36}$$

の両辺で $|x| \downarrow 0$ とすると，

$$\lim_{x \to 0} \frac{|g_1(x) - g_2(x)|}{|x|^k} = 0 \qquad (0 \le k \le n) \tag{1.7.37}$$

がなりたつ．そこで

$$g_1(x) - g_2(x) = \sum_{j=0}^{n} (a_j - b_j) x^j \tag{1.7.38}$$

である事に注目して $k = 0, 1, 2, \ldots$ に対して (1.7.37) を順次考えると，$a_k - b_k = 0$ しかあり得ないことが，$k = 0, 1, 2, \ldots$ と順次わかる． □

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

また，複雑な関数のテイラー展開（の最初の何項か）を求めるには，いろいろな工夫が可能になる．例えば，$\tan x$ を $x = 0$ の周りで展開する場合，まともに微分して行くとなかなか大変だ．でも $\tan x = \dfrac{\sin x}{\cos x}$ を利用して，分母と分子を別々にテイラー展開した後で割り算（これにも工夫が必要だが）するのが良い．

1.8 関数の増減とグラフ

以下では高階導関数の応用を簡単に紹介する．どちらも高校でやったはずだ．

1.8.1 関数の極大・極小

定義 1.8.1 点 $x = a$ が関数 $f(x)$ の <u>極大点</u>（local maximum）であるとは，

$$\exists r > 0, \qquad 0 < |x - a| < r \quad \Longrightarrow \quad f(x) < f(a) \tag{1.8.1}$$

となることである．このとき，f は $x = a$ で極大，ともいう．同様に，点 $x = a$ が関数 $f(x)$ の <u>極小点</u>（local minimum）であるとは，

$$\exists r > 0, \qquad 0 < |x - a| < r \quad \Longrightarrow \quad f(x) > f(a) \tag{1.8.2}$$

であることをいう．なお，

$$\exists r > 0, \qquad |x - a| < r \quad \Longrightarrow \quad f(x) \le f(a) \tag{1.8.3}$$

となっている時（最後の不等号に等号を許す），f は a で <u>広義の極大</u> という．広義の極小も同様に定義する．

（注）高校でも強調されたかもしれないが，関数 $f(x)$ が $x = a$ で <u>最大</u>（maximum）とは，f の <u>定義域全体</u> を見渡した時に $f(a)$ が最大であることをいう．つまり，

$$f \text{ の定義域に入っているすべての } x \text{ に対して} \quad f(x) < f(a) \tag{1.8.4}$$

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

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

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

定理 1.8.2 $x = a$ 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.）

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

1.8.2　曲線の凹凸

これまた高校でもやったはずだが，２階導関数の幾何学的意味を復習しておこう．

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

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

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

2　多変数函数の微分

これから，偏微分（多変数函数の微分）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 [7]，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．

2.1　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，

$$x^2, \quad x^4, \quad \sin x \quad \cos(x^2+2), \quad \ldots \tag{2.1.1}$$

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 [8]．

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 [9]．

- 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 2.1.1** 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

[7] 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

[8] 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

[9] 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
$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\| = \Bigl(\sum_{j=1}^{n} (x_j - y_j)^2\Bigr)^{1/2} \tag{2.1.2}$$
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 2.1.2 (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．
$$|x - a| \to 0 \quad ならば \quad \bigl|f(\boldsymbol{x}) - \alpha\bigr| \to 0 \tag{2.1.3}$$

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 2.1.3 (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．
$$\|\boldsymbol{x} - \boldsymbol{a}\| \to 0 \quad ならば \quad \bigl|f(\boldsymbol{x}) - \alpha\bigr| \to 0 \tag{2.1.4}$$

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 2.1.4 (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.）

2.2 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 2.2.1 (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

$$\lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}=\lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \tag{2.2.1}$$

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

$$\lim_{h\to 0}\frac{f(a,b+h)-f(a,b)}{h}=\lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b} \tag{2.2.2}$$

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 2.2.2 (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

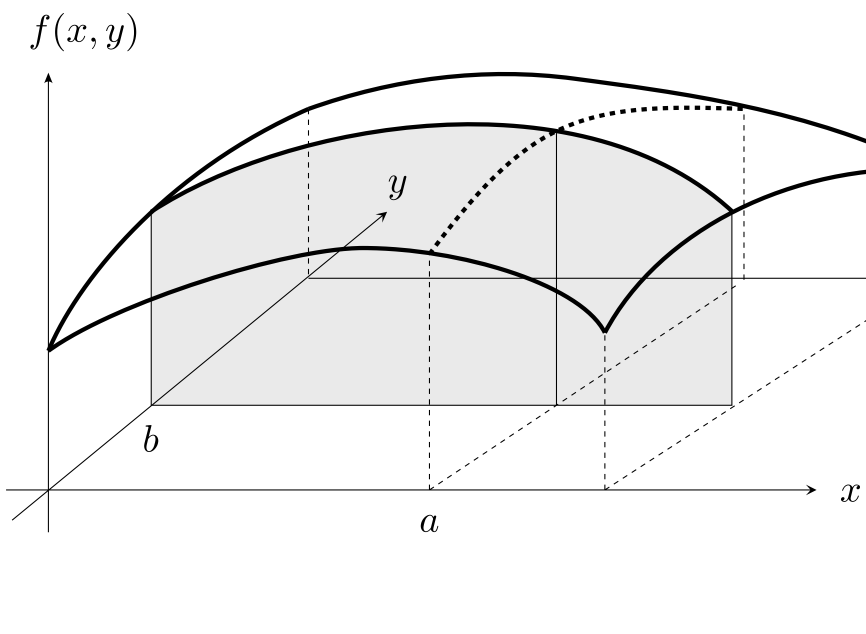
- f は各変数 $x_1,x_2,\ldots,x_n$ のそれぞれについて偏微分可能で
- かつ，その n-この偏導関数が $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ の連続関数である

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．

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

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

2.2.1　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

$$f(x,y)=g(y)\qquad g\ \text{は任意の関数} \tag{2.2.3}$$

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．

2.2.2　方向微分 [10]

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 2.2.3 (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

$$f_{\boldsymbol{v}}(a,b):=\lim_{h\to 0}\frac{f(a+hv_x,b+hv_y)-f(a,b)}{h} \tag{2.2.4}$$

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 方向微係数（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

$$f_{\boldsymbol{v}}(\boldsymbol{a}):=\lim_{h\to 0}\frac{f(\boldsymbol{a}+h\boldsymbol{v})-f(\boldsymbol{a})}{h} \tag{2.2.5}$$

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 方向微係数 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 x_1-軸の向きを向いた単位ベクトル方向の方向微分，ということになる．

[10] 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

$$f(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad f(x,y)=\frac{2xy}{x^2+y^2} \tag{2.2.6}$$

$$g(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad g(x,y)=\frac{xy}{\sqrt{x^2+y^2}} \tag{2.2.7}$$

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，

$$f_x(0,0)=f_y(0,0)=0, \qquad (x,y)\neq(0,0) \quad \text{では} \quad f_x(x,y)=\frac{2y(y^2-x^2)}{(x^2+y^2)^2}, \quad f_y(x,y)=\frac{2x(x^2-y^2)}{(x^2+y^2)^2} \tag{2.2.8}$$

$$g_x(0,0)=g_y(0,0)=0, \qquad (x,y)\neq(0,0) \quad \text{では} \quad g_x(x,y)=\frac{y^3}{(x^2+y^2)^{3/2}}, \qquad g_y(x,y)=\frac{x^3}{(x^2+y^2)^{3/2}} \tag{2.2.9}$$

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）．

2.2.3　全微分可能性 [11]

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：

$$f(x)\approx f(a)+f'(a)\times(x-a). \tag{2.2.10}$$

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 (2.2.6)，(2.2.7) の例を考えてみるとよい．$x=y=0$（原点）では f の偏微分係数はともにゼロであるが，$f(x,y)$ は原点付近でゼロではない．たとえば $x=y$ では $f(x,x)=1$（$x\neq 0$）であって，原点で連続ですらない！1 変数関数の場合は「微分可能ならば連続」であるのに，２変数関数ではこのような変態もありうるわけだ．

しかし，これは実は驚くにはあたらない．x での偏微分というのは y を固定して x を動かした時の振る舞いしか見ないから，x-軸に平行に動いたときの振る舞いは偏微分からわかるけども，$x=y$ のように x-軸に平行でない動きは 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 2.2.4 (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），

$$f(\boldsymbol{x})=f(\boldsymbol{a})+\sum_{j=1}^{n}A_j(x_j-a_j)+o(\|\boldsymbol{x}-\boldsymbol{a}\|) \tag{2.2.11}$$

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，(2.2.6) や (2.2.7) の f, g は原点 $(0,0)$ では全微分可能でない．

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

上の定義のミソは (2.2.11) 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 2.2.5 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，
$$\Bigl(\ (2.2.11)\ \text{の}\ \ A_j\ \Bigr) = \frac{\partial f}{\partial x_j}(\boldsymbol{a}) \qquad (j=1,2,\ldots,n), \tag{2.2.12}$$
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
$$D_{\boldsymbol{v}} f(\boldsymbol{a}) = \sum_{j=1}^{n} A_j v_j = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j} v_j \tag{2.2.13}$$

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 (2.2.11) 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，(2.2.11) から
$$\frac{f(x_1,a_2,a_3,\ldots,a_n)-f(a_1,a_2,a_3,\ldots,a_n)}{x_1-a_1} = A_1 + \frac{\tilde f(\boldsymbol{x})}{x_1-a_1} \tag{2.2.14}$$
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，(2.2.11) 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 (2.2.11) から
$$\frac{f(\boldsymbol{a}+h\boldsymbol{v})-f(\boldsymbol{a})}{h} = \frac{\displaystyle\sum_{j=1}^{n} A_j h v_j + \tilde f(\boldsymbol{a}+h\boldsymbol{v})}{h} = \sum_{j=1}^{n} A_j v_j + \frac{\tilde f(\boldsymbol{a}+h\boldsymbol{v})}{h} \tag{2.2.15}$$
が得られる．ここで $\boldsymbol{x}=\boldsymbol{a}+h\boldsymbol{v}$ と書くと
$$\frac{\tilde f(\boldsymbol{a}+h\boldsymbol{v})}{h} = \frac{f(\boldsymbol{x})}{\|\boldsymbol{x}-\boldsymbol{a}\|} \tag{2.2.16}$$
であるため，(2.2.11) から，(2.2.15) の最後の項は $h\to 0$ でゼロに行く．よって，(2.2.13) 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 (2.2.11) は図形的には以下のように解釈できる．まず，(2.2.11) の最後の項がない場合を考えると，定数 A,B があって
$$f(x,y) = f(a,b) + A(x-a) + B(y-b) \tag{2.2.17}$$
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 (2.2.11) 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 (2.2.11) 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．(2.2.6) や (2.2.7) 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 2.2.6 (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

$$f(x, y) - f(a, b) = \bigl\{f(x, y) - f(a, y)\bigr\} + \bigl\{f(a, y) - f(a, b)\bigr\} \tag{2.2.18}$$

が $\boldsymbol{x} \to \boldsymbol{a}$ でどのように振る舞うか — (2.2.11) を満たすか — を調べなければならない．

さて，(2.2.18) 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）

$$f(a, y) - f(a, b) = f_y(a, \tilde{y}) \times (y - b) \tag{2.2.19}$$

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）

$$f(x, y) - f(a, y) = f_x(\tilde{x}, y) \times (x - a) \tag{2.2.20}$$

となる（$\tilde{x}$ は a と x の間の適当な数）．これを (2.2.18) に代入して

$$f(x, y) - f(a, b) = f_x(\tilde{x}, y) \times (x - a) + f_y(a, \tilde{y}) \times (y - b) \tag{2.2.21}$$

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

$$f_x(u, v) = f_x(a, b) + g(u, v), \qquad \text{with} \qquad \lim_{(u,v)\to(a,b)} g(u, v) = 0, \tag{2.2.22}$$

$$f_y(u, v) = f_y(a, b) + h(u, v), \qquad \text{with} \qquad \lim_{(u,v)\to(a,b)} h(u, v) = 0 \tag{2.2.23}$$

が成り立っている．ここで (2.2.22) 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，

$$f_x(\tilde{x}, y) = f_x(a, b) + g(\tilde{x}, y), \qquad \text{with} \qquad \lim_{(x,y)\to(a,b)} g(\tilde{x}, y) = 0 \tag{2.2.24}$$

が結論できる．同様に，

$$f_y(a, \tilde{y}) = f_y(a, b) + h(a, \tilde{y}), \qquad \text{with} \qquad \lim_{(x,y)\to(a,b)} h(a, \tilde{y}) = 0 \tag{2.2.25}$$

も結論できる．これを (2.2.21) に代入すると

$$f(x, y) - f(a, b) = f_x(a, b) \times (x - a) + f_y(a, b) \times (y - b) + g(\tilde{x}, y) \times (x - a) + h(a, \tilde{y}) \times (y - b) \tag{2.2.26}$$

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，

$$\lim_{(x,y)\to(a,b)} \frac{\tilde{f}(x, y)}{\|\boldsymbol{x} - \boldsymbol{a}\|} = 0 \qquad \text{（ここで}\quad \boldsymbol{x} = (x, y),\ \ \boldsymbol{a} = (a, b)\quad\text{と書いた）} \tag{2.2.27}$$

が結論できる．結果として，このような $\tilde{f}$ を用いて

$$f(x,y) = f(a,b) + A(x-a) + B(y-b) + \tilde{f}(x,y), \qquad \text{with} \qquad A = f_x(a,b),\ B = f_y(a,b) \tag{2.2.28}$$

と書ける事がわかった．これは全微分可能の定義式 (2.2.11) 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．□

2.3　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 2.3.3 節である [12]．

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，

$$h(x) = f(g(x)) \tag{2.3.1}$$

で定義される関数 h を f と g の <u>合成関数</u> 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，

$$h'(x) = f'\bigl(g(x)\bigr)\, g'(x) \tag{2.3.2}$$

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

$$\frac{\partial h(x,y)}{\partial x} = f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial x}, \qquad \frac{\partial h(x,y)}{\partial y} = f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial y} \tag{2.3.3}$$

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．

2.3.1　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．

[12] 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 2.3.1（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 2.3.3 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 2.3.1 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

$$\frac{h(a+\epsilon)-h(a)}{\epsilon}=\frac{f\bigl(g(a+\epsilon)\bigr)-f\bigl(g(a)\bigr)}{\epsilon}=\frac{f\bigl(g(a+\epsilon)\bigr)-f\bigl(g(a)\bigr)}{g(a+\epsilon)-g(a)}\times\frac{g(a+\epsilon)-g(a)}{\epsilon}\tag{2.3.4}$$

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 (2.3.4) 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．

2.3.2　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 2.3.2 $g(x,y)$ 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，

$$h_x(a,b)=f'\bigl(g(a,b)\bigr)\,g_x(a,b),\qquad \text{つまり}\ z=g(x,y),\ w=f(z)\ \text{とおくと}\quad \frac{\partial w}{\partial x}=\frac{dw}{dz}\,\frac{\partial z}{\partial x}\tag{2.3.5}$$

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 2.3.1 が使える． □

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

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

２変数の関数 $f(x,y)$ と $x(t), y(t)$ から合成関数 $h(t) = f(x(t), y(t))$ を作る．この t での微分を考えると，ニュートン商の極限として［記号を見やすくするため．$x_0 = x(t)$, $y_0 = y(t), x_1 = x(t+\epsilon), y_1 = y(t+\epsilon)$ と書く [13]］

$$\begin{aligned}\lim_{\epsilon\to 0} \frac{h(t+\epsilon) - h(t)}{\epsilon} &= \lim_{\epsilon\to 0} \frac{f\bigl(x(t+\epsilon), y(t+\epsilon)\bigr) - f\bigl(x(t), y(t)\bigr)}{\epsilon} = \lim_{\epsilon\to 0} \frac{f(x_1,y_1) - f(x_0,y_0)}{\epsilon} \\ &= \lim_{\epsilon\to 0} \frac{f(x_1,y_1) - f(x_1,y_0)}{\epsilon} + \lim_{\epsilon\to 0} \frac{f(x_1,y_0) - f(x_0,y_0)}{\epsilon} \qquad (2.3.6)\end{aligned}$$

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）

$$\lim_{\epsilon\to 0} \frac{f(x_2,y_1) - f(x_2,y_0)}{\epsilon} \qquad \text{with} \qquad y_0 = y(t), \quad y_1 = y(t+\epsilon) \qquad (2.3.7)$$

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，(2.3.6) から

$$h'(t) = \frac{\partial f}{\partial x}(x_0,y_0)\, x'(t) + \frac{\partial f}{\partial y}(x_0,y_0)\, y'(t) \qquad (2.3.8)$$

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 (2.3.8) はかなり一般に成り立つ．ただし，上でも明記したように，(2.3.6) 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：

$$g(0,0) = 0, \qquad (x,y) \neq (0,0) \ \text{では} \quad g(x,y) = \frac{xy}{\sqrt{x^2+y^2}}. \qquad (2.3.9)$$

この関数の偏導関数は (2.2.9) 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

$$h(t) = g(x(t), y(t)) = \frac{t}{\sqrt{2}}, \qquad h'(t) = \frac{1}{\sqrt{2}} \qquad (2.3.10)$$

であって，特に $h'(0) \neq 0$ だ．ところが (2.3.8) 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 (2.3.8) 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 2.3.4 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 2.3.3 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

$$h'(t) = f_x\, x'(t) + f_y\, y'(t), \qquad \text{つまり} \ \ z = f(x,y) \ \text{とおくと} \quad \frac{dz}{dt} = \frac{\partial z}{\partial x}\frac{dx}{dt} + \frac{\partial z}{\partial y}\frac{dy}{dt} \qquad (2.3.11)$$

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：

$$\frac{d}{dt} f\bigl(x_1(t), x_2(t), \ldots, x_n(t)\bigr) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{dx_j}{dt}. \qquad (2.3.12)$$

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

言うまでもなく，(2.3.9) 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　(2.3.6) 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）

$$f(x_1,y_1)-f(x_1,y_0)=f_y(x_1,y_3)\times(y_1-y_0) \tag{2.3.13}$$

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，

$$\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}=\lim_{\epsilon\to 0}\left[f_y(x_1,y_3)\times\frac{y_1-y_0}{\epsilon}\right] \tag{2.3.14}$$

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，

$$\lim_{\epsilon\to 0}f_y(x_1,y_3)=f_y(x_0,y_0) \tag{2.3.15}$$

である．また，(2.3.14) の後ろの方は y_1,y_0 の定義から

$$\lim_{\epsilon\to 0}\frac{y_1-y_0}{\epsilon}=\lim_{\epsilon\to 0}\frac{y(t+\epsilon)-y(t)}{\epsilon}=y'(t) \tag{2.3.16}$$

である．以上をまとめると

$$\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}=f_y(x_0,y_0)\times y'(t) \tag{2.3.17}$$

が厳密に証明されたので，(2.3.6) 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 2.3.4（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，

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u},\qquad \text{かつ}\qquad \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v}. \tag{2.3.18}$$

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 2.3.3 からすぐに証明される．　□

問題 2.3.5　（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 2.3.6　(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）

$$u=\alpha x+\beta y,\;\; v=\gamma x+\delta y \tag{2.3.19}$$

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 2.3.7　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）．

$$1)\;\;\frac{\partial f}{\partial y}\equiv 0\qquad 2)\;\;\frac{\partial f}{\partial x}+\frac{\partial f}{\partial y}\equiv 0\qquad 3)\;\;a\frac{\partial f}{\partial x}+b\frac{\partial f}{\partial y}\equiv 0 \tag{2.3.20}$$

1) は既にやった（2.2.1 節）．問題は 2) と 3) だが，問題 2.3.6 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.)

2.3.4　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 2.3.8 (定理 2.3.3 の一般形)　2 変数関数 $f(x,y)$ が全微分可能で，$x(t),y(t)$ が t について微分可能なら，$h(t)=f\bigl(x(t),y(t)\bigr)$ は t で微分可能で，

$$h'(t)=f_x\,x'(t)+f_y\,y'(t),\qquad \text{つまり}\ \ z=f(x,y)\ \text{とおくと}\quad \frac{dz}{dt}=\frac{\partial z}{\partial x}\frac{dx}{dt}+\frac{\partial z}{\partial y}\frac{dy}{dt} \tag{2.3.21}$$

が成り立つ．n-変数の場合も同様であるが，略.

定理 2.3.9 (定理 2.3.4 の一般形)　2 変数関数 $f(x,y)$ と $x(u,v),y(u,v)$ がそれぞれその引数について全微分可能なら，合成関数 $z=h(u,v)=f\bigl(x(u,v),y(u,v)\bigr)$ も (u,v) について全微分可能で，

$$\frac{\partial z}{\partial u}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u},\qquad \text{かつ}\qquad \frac{\partial z}{\partial v}=\frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v} \tag{2.3.22}$$

である.

定理 2.3.4 が定理 2.3.3 から出るのと同じようにして，定理 2.3.9 は定理 2.3.8 から証明できる．従って，以下では定理 2.3.8 を簡単に証明する.

定理 2.3.8 の証明

以下では 5 月初めに導入した「オーダー」の記号を用いている．今までと同じく，式を見やすくするために $x_0=x(t),x_1=x(t+\epsilon),y_0=y(t),y_1=y(t+\epsilon)$ と書く．全微分可能性の定義と命題 2.2.5 から，

$$h(t+\epsilon)-h(t)=f(x_1,y_1)-f(x_0,y_0)=A(x_1-x_0)+B(y_1-y_0)+o\Bigl(\sqrt{|x_1-x_0|^2+|y_1-y_0|^2}\Bigr) \tag{2.3.23}$$

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），

$$x_1-x_0=x'(t)\epsilon+o(\epsilon)=p\epsilon+o(\epsilon),\qquad y_1-y_0=y'(t)\epsilon+o(\epsilon)=q\epsilon+o(\epsilon) \tag{2.3.24}$$

も成り立っている．従って，

$$\begin{aligned} h(t+\epsilon)&=h(t)+A\{p\epsilon+o(\epsilon)\}+B\{q\epsilon+o(\epsilon)\}+o\Bigl(\sqrt{\{p\epsilon+o(\epsilon)\}^2+\{q\epsilon+o(\epsilon)\}^2}\Bigr)\\ &=h(t)+(Ap+Bq)\epsilon+o(\epsilon)+o\Bigl(\sqrt{(p^2+q^2)\epsilon^2+o(\epsilon^2)}\Bigr) \end{aligned} \tag{2.3.25}$$

となるが，

$$o\Bigl(\sqrt{(p^2+q^2)\epsilon^2+o(\epsilon^2)}\Bigr)=o\Bigl(\sqrt{p^2+q^2}\,\epsilon+o(\epsilon)\Bigr)=o(\epsilon) \tag{2.3.26}$$

であるので，結局 (2.3.25) は

$$h(t+\epsilon)=h(t)+(Ap+Bq)\epsilon+o(\epsilon) \tag{2.3.27}$$

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 (2.2.10) そのものだ. □

2.4　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

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \qquad \text{同様に} \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{2.4.1}$$

なども定義する（もちろん，これらの微分が定義できる場合）．まあ，この定義は非常に自然だから問題ないでしょう [14]．**ただし，微分の順序には注意：すぐに「順序はあんまり重要でない」とわかるが，$\dfrac{\partial^2 f}{\partial x \partial y} = \dfrac{\partial}{\partial x}\left(\dfrac{\partial f}{\partial y}\right)$ は先に y で偏微分してから x で偏微分，ということである．**

問題 2.4.1 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．

$$f(x,y) = x^2 + y^2, \qquad g(x,y) = x^2 y^3, \qquad h(x,y) = \cos(x^2 y) \tag{2.4.2}$$

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 2.4.1 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 2.4.2 ２変数関数 $f(x,y)$ が C^2-級の場合，

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \tag{2.4.3}$$

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 2.4.2 での「C^2-級」は十分条件であって，もう少しだけ条件を緩めることも可能である．例えば，以下のようなものがあるようだ [15]（D は領域，A は D 内の一点．また $f_x = \frac{\partial f}{\partial x}, f_{yx} = \frac{\partial^2 f}{\partial x \partial y}$ などと略記する）．

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

定理 2.4.2 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．

$$\Delta(h,k) := f(a+h,b+k) - f(a+h,b) - f(a,b+k) + f(a,b) \tag{2.4.4}$$

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

$$\begin{aligned}\lim_{h\to 0}\Bigl(\lim_{k\to 0}\frac{\Delta(h,k)}{hk}\Bigr) &= \lim_{h\to 0}\frac{1}{h}\Bigl[\lim_{k\to 0}\Bigl(\frac{f(a+h,b+k)-f(a+h,b)}{k} - \frac{f(a,b+k)-f(a,b)}{k}\Bigr)\Bigr] \\ &= \lim_{h\to 0}\frac{1}{h}\Bigl[\frac{\partial f}{\partial y}(a+h,b) - \frac{\partial f}{\partial y}(a,b)\Bigr] = \frac{\partial}{\partial x}\Bigl(\frac{\partial f}{\partial y}\Bigr) = \frac{\partial^2 f}{\partial x \partial y}\end{aligned} \tag{2.4.5}$$

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

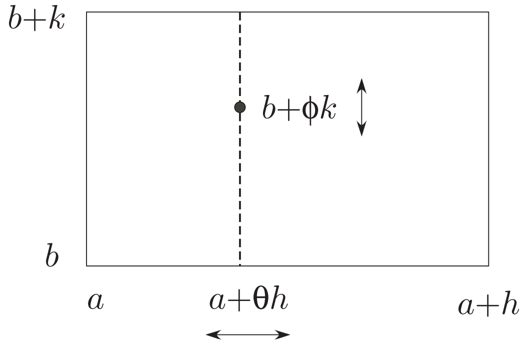
[15] 小平本の 6.2 節の d) を参考にした

であるし，同様に考えると

$$\lim_{k\to 0}\Bigl(\lim_{h\to 0}\frac{\Delta(h,k)}{hk}\Bigr)=\frac{\partial^2 f}{\partial y\partial x} \tag{2.4.6}$$

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，

$$\varphi(a+h)-\varphi(a)=\varphi'(a+\theta h)h \tag{2.4.7}$$

つまり

$$\Delta(h,k)=\varphi(a+h)-\varphi(a)=\Bigl(f_x(a+\theta h,b+k)-f_x(a+\theta h,b)\Bigr)h \tag{2.4.8}$$

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

次に a,h,θ を固定して $\psi(y) := f_x(a+\theta h,y)$ を y の関数と考えると [16] またもや 1 変数の場合の平均値の定理から

$$\psi(b+k)-\psi(b)=\psi'(b+\phi k)k \tag{2.4.9}$$

つまり

$$f_x(a+\theta h,b+k)-f_x(a+\theta h,b)=f_{xy}(a+\theta h,b+\phi k)\,k \tag{2.4.10}$$

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 (2.4.7) と (2.4.10) から

$$\frac{\Delta(h,k)}{hk}=\frac{f_{xy}(a+\theta h,b+\phi k)\,hk}{hk}=f_{xy}(a+\theta h,b+\phi k) \tag{2.4.11}$$

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：

$$\lim_{(h,k)\to(0,0)}\frac{\Delta(h,k)}{hk}=f_{xy}(a,b) \tag{2.4.12}$$

これで既に $h\to 0$ と $k\to 0$ の極限が順序によらない事は示せたから (2.4.5) と (2.4.6) 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 (2.4.12) 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．)　□

---

[16] (注) ここでは少しひっかかるかもしれない．(2.4.7) 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 (2.4.10) は (2.4.7) がなりたっていた $b+k,b$ についてのものであるので，(2.4.7) と (2.4.10) 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

2.4.1　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

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{2.4.13}$$

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

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial f}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial f}{\partial v} = \frac{1}{\sqrt{2}}\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right] \tag{2.4.14}$$

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

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial}{\partial v} = \frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial u} - \frac{\partial}{\partial v}\right] \tag{2.4.15}$$

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

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &= \frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right]\left\{\frac{1}{\sqrt{2}}\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right]\right\} = \frac{1}{2}\left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right]\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right] \\ &= \frac{1}{2}\left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v \partial u} - \frac{\partial^2 f}{\partial v^2}\right] = \frac{1}{2}\left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial v^2}\right] \end{aligned} \tag{2.4.16}$$

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．

2.4.2　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

$$\frac{\partial f}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial f}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial f}{\partial \theta} = \frac{x}{\sqrt{x^2 + y^2}}\frac{\partial f}{\partial r} - \frac{y}{x^2 + y^2}\frac{\partial f}{\partial \theta} = \frac{x}{r}\frac{\partial f}{\partial r} - \frac{y}{r^2}\frac{\partial f}{\partial \theta} \tag{2.4.17}$$

となるので，微分演算子としては

$$\frac{\partial}{\partial x} = \frac{x}{r}\frac{\partial}{\partial r} - \frac{y}{r^2}\frac{\partial}{\partial \theta} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta} \tag{2.4.18}$$

という作用をもっている．さて，もう一回やるときには

$$\frac{\partial^2 f}{\partial x^2} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) \tag{2.4.19}$$

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

$$= \cos\theta \frac{\partial}{\partial r}\Bigl(\cos\theta \frac{\partial f}{\partial r}\Bigr) - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(\cos\theta\frac{\partial f}{\partial r}\Bigr) + \cos\theta\frac{\partial}{\partial r}\Bigl(-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\Bigr) - \frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(-\frac{\sin\theta}{r}\frac{\partial f}{\partial\theta}\Bigr) \tag{2.4.20}$$

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，

$$\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(\cos\theta\frac{\partial f}{\partial r}\Bigr) = \frac{\sin\theta}{r}\frac{\partial\cos\theta}{\partial\theta}\times\frac{\partial f}{\partial r} + \frac{\sin\theta}{r}\cos\theta\frac{\partial}{\partial r}\Bigl(\frac{\partial f}{\partial\theta}\Bigr) = -\frac{\sin^2\theta}{r}\frac{\partial f}{\partial r} + \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{2.4.21}$$

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

$$\frac{\partial^2 f}{\partial x^2} = \cos^2\theta\frac{\partial^2 f}{\partial r^2} + \frac{\sin^2\theta}{r}\frac{\partial f}{\partial r} + \frac{\sin 2\theta}{r^2}\frac{\partial f}{\partial\theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 f}{\partial\theta^2} - \frac{\sin 2\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{2.4.22}$$

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 2.4.3 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，

$$\Delta f(x,y) := \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{2.4.23}$$

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．

2.4.3 （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

$$\frac{\partial}{\partial y}\bigl(f_x\bigr) = 0 \tag{2.4.24}$$

と見ると,「f_x は y には依存しません」ということなので,

$$f_x(x,y) = g(x) \tag{2.4.25}$$

と書けるはずだ（g は任意の関数，2.2.1 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，(2.4.25) から

$$\frac{\partial}{\partial x}\bigl\{f(x,y) - G(x)\bigr\} = 0 \tag{2.4.26}$$

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，

$$f(x,y) = G(x) + h(y) \tag{2.4.27}$$

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，例えば物理で「波動方程式」を考える時にでてくる．

2.5　平均値の定理

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

定理 2.5.1 (多変数の平均値の定理)　2 変数関数 $f(x, y)$ が C^1-級の場合，

$$f(a+h, b+k) = f(a, b) + \frac{\partial f}{\partial x}(a+\theta h, b+\theta k)h + \frac{\partial f}{\partial y}(a+\theta h, b+\theta k)k \tag{2.5.1}$$

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（$\boldsymbol{a} = (a_1, a_2, \ldots, a_n)$，$\boldsymbol{h} = (h_1, h_2, \ldots, h_n)$）

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{j=1} \frac{\partial f}{\partial x_j}(\boldsymbol{a}+\theta\boldsymbol{h})\, h_j \tag{2.5.2}$$

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

$$f(a+h, b+k) - f(a, b) = g(1) - g(0) = g'(\theta) \tag{2.5.3}$$

である．ところが，g' については,「連鎖律」定理 2.3.3 を $x(t) = a+th, y(t) = b+yk$ として用いると

$$g'(t) = \frac{\partial f}{\partial x}x'(t) + \frac{\partial f}{\partial y}y'(t) = \frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k \tag{2.5.4}$$

であるから，定理 2.5.1 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 2.5.2 (平均値の定理)　2 変数関数 $f(x, y)$ が全微分可能なら，

$$f(a+h, b+k) = f(a, b) + \frac{\partial f}{\partial x}(a+\theta h, b+\theta k)h + \frac{\partial f}{\partial y}(a+\theta h, b+\theta k)k \tag{2.5.5}$$

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 2.5.1 が定理 2.3.3 から出るのと同じようにして，定理 2.5.2 は定理 2.3.8 から証明される．　□

2.6　テイラーの定理とテイラー展開

先に，1 変数については「テイラーの定理」「テイラー展開」を学習した．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）

$$g(t) = f(a+th, b+tk) - f(a, b) \tag{2.6.1}$$

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，

$$f(a+h, b+k) = g(1) = \sum_{k=0}^{n-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(n)}(\theta)}{n!} \qquad (0 < \theta < 1) \tag{2.6.2}$$

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）

$$\frac{d}{dt} = \frac{dx(t)}{dt}\frac{\partial}{\partial x} + \frac{dy(t)}{dt}\frac{\partial}{\partial y} = h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} \tag{2.6.3}$$

であるから，例えば，f が C^2-級ならば

$$\begin{aligned} g^{(2)}(\theta) &= \frac{d}{dt}\frac{dg}{dt} = \Big(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\Big)\Big(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\Big) \\ &= h^2\frac{\partial^2 f}{\partial x^2} + hk\frac{\partial^2 f}{\partial x \partial y} + kh\frac{\partial^2 f}{\partial y \partial x} + k^2\frac{\partial^2 f}{\partial y^2} \end{aligned} \tag{2.6.4}$$

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，

$$\big|g^{(2)}(\theta)\big| \le 2M(h^2 + k^2) = O(\|\boldsymbol{a} - \boldsymbol{c}\|^2) = o(\|\boldsymbol{a} - \boldsymbol{c}\|) \tag{2.6.5}$$

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，(2.6.2) の $n = 2$ を考えると，

$$f(a + h, b + k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + o(\|\boldsymbol{a} - \boldsymbol{c}\|) \tag{2.6.6}$$

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，

$$f(a+h, b+k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + \frac{1}{2}\Big[f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2\Big] + o(\|\boldsymbol{a} - \boldsymbol{c}\|^2) \tag{2.6.7}$$

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 2.6.1　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

$$\begin{aligned} f(a + h, b + k) = f(a, b) &+ \sum_{m=1}^{r-1}\frac{1}{m!}\sum_{\ell=0}^{m}\binom{m}{\ell}\frac{\partial^m f}{\partial x^\ell\,\partial y^{m-\ell}}(a, b)\,h^\ell\,k^{m-\ell} \\ &+ \frac{1}{r!}\sum_{\ell=0}^{r}\binom{r}{\ell}\frac{\partial^r f}{\partial x^\ell\,\partial y^{r-\ell}}(a + \theta h, b + \theta k)\,h^\ell\,k^{r-\ell} \end{aligned} \tag{2.6.8}$$

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，(2.6.7) 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

2.7　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．

2.7.1　問題の定義

定義 2.7.1 n-成分ベクトルの空間において，上の記号のもとで，

$$B_r(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{a}\| < r\} \tag{2.7.1}$$

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 2.7.2 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，

$$\text{ある正の } r \text{ が存在して，} \qquad 0 < \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \text{では} \quad f(\boldsymbol{x}) < f(\boldsymbol{a}) \tag{2.7.2}$$

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，

$$\text{ある正の } r \text{ が存在して，} \qquad 0 < \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \text{では} \quad f(\boldsymbol{x}) > f(\boldsymbol{a}) \tag{2.7.3}$$

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 (2.7.2) と (2.7.3) の代わりに

$$\exists r > 0, \qquad \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{2.7.4}$$

$$\exists r > 0, \qquad \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \ge f(\boldsymbol{a}) \tag{2.7.5}$$

  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，

$$f \text{ の定義域に入っているすべての } \boldsymbol{x} \text{ に対して} \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{2.7.6}$$

  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，(2.7.6) 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．

2.7.2　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 2.7.3 $x = a$ 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

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + o(|x-a|^2) \tag{2.7.7}$$

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

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \tag{2.7.8}$$

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

$$f(x) = f(a) + \frac{f''(a)}{2}(x-a)^2 + o(|x-a|^2) \tag{2.7.9}$$

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］，

$$f(x) = f(a) + \frac{f^{(3)}(a)}{6}(x-a)^3 + \frac{f^{(4)}(a)}{24}(x-a)^4 + \frac{f^{(5)}(a)}{120}(x-a)^5 + o(|x-a|^5) \tag{2.7.10}$$

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．

2.7.3　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 2.7.4 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

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y)=(a,b) \text{ での値} \tag{2.7.11}$$

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 Hesse 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 2.7.5 $(x,y)=(a,b)$ 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），

$$f(x,y) = f + f_x(x-a) + f_y(y-b) + \frac{1}{2}\left[f_{xx}(x-a)^2 + 2f_{xy}(x-a)(y-b) + f_{yy}(y-b)^2\right] + o(\|\boldsymbol{x}-\boldsymbol{a}\|^2) \tag{2.7.12}$$

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

$$g(x,y) = A(x-a)^2 + 2B(x-a)(y-b) + C(y-b)^2 \tag{2.7.13}$$

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，

$$g(x,y) = A\left[\left\{(x-a) + \frac{B}{A}(y-b)\right\}^2 + \frac{CA-B^2}{A^2}(y-b)^2\right] \tag{2.7.14}$$

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 2.7.5 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

$$f_x(a,b) = f_y(a,b) = 0 \tag{2.7.15}$$

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 2.7.4 と定理 2.7.5）．上を満たす (a, b) の一つ一つについて，ヘッセ行列

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y) = (a,b) \text{ での値} \tag{2.7.16}$$

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）

2.7.4　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，(多分，この講義では扱わない)「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

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \sum_{j=1}^{n} (x_j - a_j) \frac{\partial f}{\partial x_j}(\boldsymbol{a}) + \frac{1}{2} \sum_{i,j=1}^{n} (x_i - a_i)(x_j - a_j) \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) + o(\|\boldsymbol{x} - \boldsymbol{a}\|^2) \tag{2.7.17}$$

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，

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = 0 \qquad (j = 1, 2, \ldots, n) \tag{2.7.18}$$

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

$$\sum_{i,j=1}^{n} h_i h_j a_{ij} = {}^t\boldsymbol{h} A \boldsymbol{h} \qquad \text{ここで} \quad h_i = x_i - a_i, \quad a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}), \tag{2.7.19}$$

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 [17] 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（${}^tPP = P\,{}^tP = E$），

$$B = {}^tP\,A\,P \qquad A = P\,B\,{}^tP \tag{2.7.20}$$

[17] 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

$$(\boldsymbol{h}, A\boldsymbol{h}) = (\boldsymbol{h}, P\,B\,{}^tP\,\boldsymbol{h}) = ({}^tP\,\boldsymbol{h}, B\,{}^tP\,\boldsymbol{h}) = (\boldsymbol{g}, B\boldsymbol{g}) = \sum_{j=1}^{n} \lambda_j (g_j)^2 \tag{2.7.21}$$

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 2.7.6 $\boldsymbol{x} = \boldsymbol{a}$ 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 2.7.7 $n \times n$ 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．

2.8　陰関数定理（完全なおまけ：この講義では扱わない）

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．

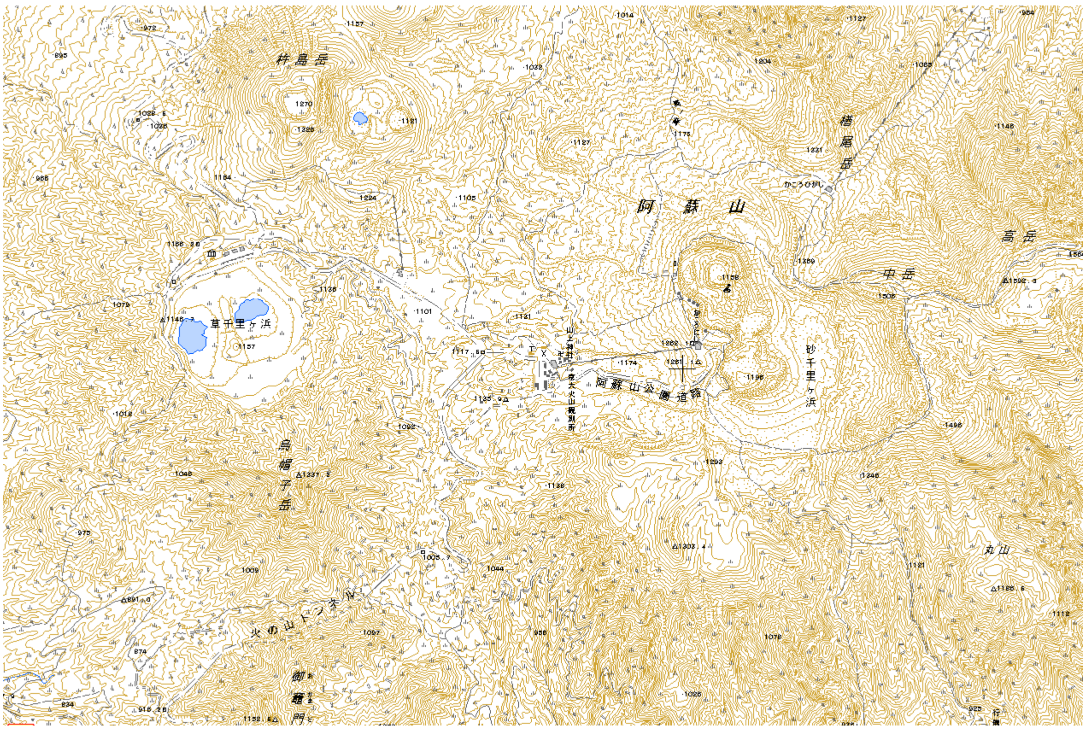
問題 2.8.1 xy-平面全体で定義された C^1-級の関数 $f(x,y)$ がある．$f(x,y)=0$ を y について解いて y を x の関数として表せ．別の言い方をすると，$f(x,y)$ の**零点**，つまり $f(x,y)=0$ となる点の集合を求めよ．

$f(x,y)$ 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[18] 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．



[18] 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 2.8.2 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 2.8.3（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，

$$b=\varphi(a)\quad \text{かつ}\qquad (a,b)\ \text{の近傍で}\ \ f(x,\varphi(x))=0 \tag{2.8.1}$$

がなりたつ．更に (a,b) の近傍では

$$\frac{d}{dx}\varphi(x)=-\left.\frac{f_x(x,y)}{f_y(x,y)}\right|_{y=\varphi(x)} \tag{2.8.2}$$

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）

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 (2.8.1) 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. (2.8.1) 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．

2.9　条件付き極値問題：ラグランジュの未定乗数法（多分，講義では扱わない）

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

$$F(x,y,\lambda)=f(x,y)-\lambda g(x,y) \tag{2.9.1}$$

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．

$$0=\frac{\partial F}{\partial x}=\frac{\partial f}{\partial x}(x,y)-\lambda\frac{\partial g}{\partial x}(x,y),\qquad 0=\frac{\partial F}{\partial y}=\frac{\partial f}{\partial y}(x,y)-\lambda\frac{\partial g}{\partial y}(x,y) \tag{2.9.2}$$

$$0=\frac{\partial F}{\partial\lambda}=g(x,y) \tag{2.9.3}$$

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 (2.9.2) と (2.9.3)（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

$$0=4x^3+2\lambda x, \qquad 0=4y^3+2\lambda y, \qquad 0=x^2+y^2-1 \tag{2.9.4}$$

の３つである．これを解くと，

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \pm 1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm 1 \\ 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \pm\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \mp\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \text{（ベクトルの中では複合同順）} \tag{2.9.5}$$

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

$$0=h'(x)=f_x+f_y\,\varphi'(x) \tag{2.9.6}$$

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

$$0=\frac{d}{dx}g(x,\varphi(x))=g_x+g_y\,\varphi'(x) \tag{2.9.7}$$

この２つから，

$$\varphi'(x)=-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \tag{2.9.8}$$

が導かれるが，これは見方を変えれば

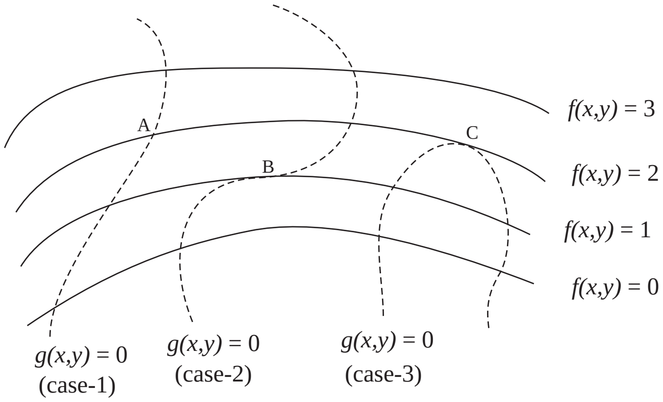
$$\frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{2.9.9}$$

ということであり，この値を λ と書けば，これは (2.9.2) 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 2.8.3 の (2.8.2) で与えられている．つまり

$$f(x,y)=c \text{ の接線の傾きは } -\frac{f_x}{f_y}, \qquad g(x,y)=0 \text{ の接線の傾きは } -\frac{g_x}{g_y} \tag{2.9.10}$$

なのだ．従って，両者が接する条件は

$$-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \qquad \text{つまり} \qquad \frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{2.9.11}$$

であるが，これは (2.9.2) 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

$$F(\boldsymbol{x},\lambda_1,\lambda_2,\ldots,\lambda_m)=f(\boldsymbol{x})-\Bigl\{\lambda_1 g_1(\boldsymbol{x})+\lambda_2 g_2(\boldsymbol{x})+\cdots+\lambda_m g_m(\boldsymbol{x})\Bigr\} \tag{2.9.12}$$

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．

$$\begin{aligned} 0&=\frac{\partial F}{\partial x_j}=\frac{\partial f}{\partial x_j}(\boldsymbol{x})-\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial x_j}(\boldsymbol{x}), \qquad (j=1,2,\ldots,n)\\ 0&=\frac{\partial F}{\partial \lambda_k}=g_k(\boldsymbol{x}), \qquad (k=1,2,\ldots,m) \end{aligned} \tag{2.9.13}$$

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．

3　重積分

時間の関係で，1 変数函数の積分についてはすべて省略し（高校の知識で大体，良しとする），重積分を残った時間でやります．

3.1　1 変数関数の積分

まず，高校の復習をも兼ねて,「1 変数関数の積分」を復習しておこう．このところがはっきりしていれば，この講義で出てくるいろんな積分は簡単なはず．

$f(x)$ を適当な（例えば連続な）関数とし，簡単のために $f(x) > 0$ としておこう．$a < b$ を定めたときの定積分 $\displaystyle\int_a^b f(x)dx$ とは直感的には区間 $[a, b]$ 上での $y = f(x)$ の**グラフと** x**-軸との間の図形の面積**である（$f(x)$ が負のときはもちろん，面積の符号を変えたもの）．この積分の数学での定義は以下のようなものである．

- まず，区間 $[a, b]$ を n 個（n は大きな整数）の小区間に分ける：$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$. できる小区間は (x_{i-1}, x_i) である（$i = 1, 2, \ldots, n$）．これを区間 $[a, b]$ の**分割**といい，Δ で表す．
- 各小区間の幅を $\Delta_i = x_i - x_{i-1}$ と書く．また，これらの小区間の長さの最大値を $|\Delta|$ と書く：$|\Delta| = \max_{1\le i\le n} (x_i - x_{i-1}) = \max_{1\le i\le n} \Delta_i$.
- 各小区間 $[x_{i-1}, x_i]$ に勝手に点 ζ_i をとる（$i = 1, 2, \ldots, n$）．簡単のために $\zeta_1, \zeta_2, \ldots, \zeta_n$ をまとめて $\vec{\zeta}$ と書く．
- 上のように決めた $\Delta, \vec{\zeta}$ に対して，**リーマン和**

$$S(f; \Delta, \vec{\zeta}) = \sum_{i=1}^{n} f(\zeta_i)\,(x_i - x_{i-1}) = \sum_{i=1}^{n} f(\zeta_i)\,\Delta_i \tag{3.1.1}$$

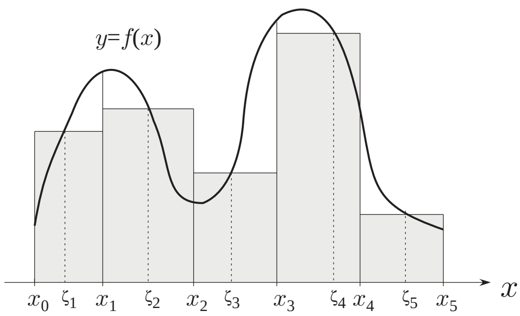
を計算する．

- さて，$|\Delta| \to 0$ を満たすような任意の分割 Δ とそれに対する任意の $\vec{\zeta}$ を考える．$|\Delta| \to 0$ の極限で $S(f; \Delta, \vec{\zeta})$ の値が（$\Delta, \vec{\zeta}$ の取り方によらず）一定の値に近づくならばその値を $\int_a^b f(x)dx$ の値と定める．模式的に数式で書けば

$$\int_a^b f(x)dx \equiv \lim_{|\Delta|\to 0} S(f; \Delta, \vec{\zeta}) \tag{3.1.2}$$

とするのである．

$f(x) > 0$ の場合の模式図（$n = 5$）を以下に示した．図で陰をつけた部分の面積がこの場合の $R(f; P, \vec{\zeta})$ である．



上の極限値はいつもあるとは限らない．例えば,

$$f(x) = \begin{cases} 0 & (x \text{ が有理数の時}) \\ 1 & (x \text{ が無理数の時}) \end{cases} \qquad \text{に対して} \qquad \int_0^1 f(x)dx \tag{3.1.3}$$

を考えても，これは定義できない.（このような関数に対しても「積分」を定義しよう，というのが「ルベーグ積分」なのであるが，この講義ではルベーグ積分は扱わない.）定積分が定義できるかどうかなどについては

定理 3.1.1 (連続関数は積分可能) 関数 $f(x)$ が区間 $[a,b]$ 上で連続なら，f は $[a,b]$ 上で積分可能である．また，有限個の点を除くと連続な場合も積分可能である．

ともかく，定積分の基本は，グラフの下の図形を細い短冊の和で近似する，ということなのだった．これからいろいろな積分が出てくるが，これらはすべて，上のような意味での「うまく近似した和の極限」として理解すべきものである．

3.2　２重積分の定義とその意味

２重積分は単に「重積分」ということも多い．まず，重積分で何をやりたいのか，考えてみよう．

2 変数 x,y の関数 $f(x,y)$ が与えられている（例：$f(x,y)=xy$）．また，xy-平面上の長方形の領域 $A=\{(x,y)\,\big|\,a\le x\le b, c\le y\le d\}$ も与えられている．そのとき，関数 f の領域 A 上での積分 $\displaystyle\iint_A f(x,y)dxdy$ を定義したい．もともと積分は「グラフの下の図形の面積」を表すものだったから，そのノリを保って，以下のように考える．(実際，この定義が自然で役に立つことはこれから見ていく．) —— ここのところはまず黒板で概念を説明するつもり．

関数 $z=f(x,y)$ のグラフは (x,y,z)-空間での曲面になる．簡単のために $f(x,y)\ge 0$ とする．この曲面と xy 平面の間にあり，底面が A である立体（A を底面とする柱のようなもの）を考え，この体積を表すものが重積分 $\displaystyle\iint_A f(x,y)dxdy$ であるように，定義を考えたい．1 変数の時に倣って，問題の体積を小さな部分の和で近似するつもりで定義する．

- まず，A を小さな長方形に分割する．つまり，x-軸方向には $a=x_0<x_1<x_2<\ldots\quad<x_{n-1}<x_n=b$，$y$-軸方向には $c=y_0<y_1<y_2<\ldots<y_{m-1}<y_m=d$，と分ける（$m,n$ は大きな整数；これで A は mn 個の小さな長方形に分割された）．この分割を Δ で表す．この時にできる小さな長方形を $I_{ij}=[x_{i-1},x_i]\times[y_{j-1},y_j]$ と書く（$1\le i\le n, 1\le j\le m$）．また，これらの小長方形の辺の長さのの最大値を $|\Delta|$ と書く：$|\Delta|=\max_{i,j}\{(x_i-x_{i-1}),(y_j-y_{j-1})\}$．
- 次に，それぞれの小長方形 I_{ij} の中に勝手に点 $\zeta_{ij}=(\xi_{ij},\eta_{ij})$ をとる．mn 個の ζ_{ij} をまとめて $\vec{\zeta}$ と書く．
- このように決めた $\Delta,\vec{\zeta}$ に対してリーマン和を

$$S(\Delta,\vec{\zeta})=\sum_{i=1}^{n}\sum_{j=1}^{m}f(\xi_{ij},\eta_{ij})\,(x_i-x_{i-1})\,(y_j-y_{j-1}) \tag{3.2.1}$$

  として定義する．
- 最後に，$|\Delta|\to 0$ となるようないろいろな Δ と，その Δ に対するいろいろな $\vec{\zeta}$ の取り方を考える．$|\Delta|\to 0$ の極限でリーマン和が一定値に近づくならば，その値を「A の上での f の重積分」の値と定義する．つまり，

$$\iint_A f(x,y)dxdy=\lim_{|\Delta|\to 0}S(\Delta,\vec{\zeta}) \tag{3.2.2}$$

これが重積分の定義である（考えやすいように $f(x,y)\ge 0$ の制限を始めにつけたが，勝手な f で上の定義を用いる）．概念図は黒板で説明．もちろん，これは定義の大筋を述べただけで，以下のような問題点（取り扱わなかった点）が残されている．

- 1 変数の積分と同様，$|\Delta|\to 0$ の極限値が存在するか（すなわち重積分が定義できるか）どうかは全く自明ではない．実際に f の性質が悪いと極限値が存在しないことも多い．極限値が存在する（重積分が定義できる）十分条件など，この辺りの詳しいことは来週．
- 今は長方形上の重積分を考えているが，本当はもっと一般に，xy-平面上の勝手な図形 A の上での重積分を考えたい．この場合も定義のアイディアは同じである（底面 A を小長方形に分けて，小さな柱の体積の和の極限で定義）．ただし，A が性質の良くない図形であれば，またもや積分が定義できないことがおこる．このような点については後に触れる．

一つ目の疑問に対する答えは，１次元の積分と同じだ：

定理 3.2.1 (連続関数は積分可能) 関数 $f(x,y)$ が長方形 A 上で連続なら，f は A 上で積分可能である．また，有限個の点を除くと連続な場合も積分可能である．

3.3 一般の領域での重積分

3.2 節への補足として，一般の領域での重積分の定義を簡単に述べておく．この節の内容は，まあ常識的なものだから,「油断すると変なこともある」点以外はそれほど気にしなくてよい.（ただし「縦線図形上の積分」は後で一杯出てくる.）

今までは xy-平面の長方形 $A = [a,b] \times [c,d]$ 上の積分を考えてきた．xy-平面の一般の図形 B での積分はどう定義したら良いだろうか？まず天下りに定義を与え，その後で意味を説明する．

B を xy 平面内の有界な図形とする．B の 特性関数（定義関数）と呼ばれる関数 $\chi_B(x,y)$ を，

$$\chi_B(x,y) = \begin{cases} 1 & (x,y) \in B \text{ のとき} \\ 0 & (x,y) \notin B \text{ のとき} \end{cases} \tag{3.3.1}$$

と定義する．また，B をその内部におさめられるような，十分大きな長方形 B^* をとる．

この準備の下で，まず図形 B の面積を定義しよう．

定義 3.3.1 (一般の図形の面積) 図形 B の面積は，重積分

$$\iint_{B^*} \chi_B(x,y)\,dxdy \tag{3.3.2}$$

によって定義する（B^* とは上で述べた，B を含む長方形）．この積分が存在しない場合は，B の面積は定義できないと考える．この積分が存在する場合，B は**面積確定**であるという．

上の定義の右辺は，B^* が長方形であるから，前節までの定義によって解釈できる．χ_B の定義を見ればわかるように，この右辺では積分に実質的に効いてくるのは B の中だけである．この意味で上の定義は直感的に「正しい」ものと考えて良い．

図形 B の「面積」とは何か？は決してアタリマエの事ではない．我々が日常見かけるような図形は，その境界が**滑らかな曲線**であるから，その面積は直感的にも定義できる．しかし，その境界が連続な曲線まで拡げると，既に上の定義では面積が確定できない図形も多々ある．物理や工学に出てくる曲線でも滑らかでないものもある（例：ブラウン運動の軌跡．もう少し講義で）．

この意味で,「図形の面積」は**我々が**苦労して，意識的に**定義すべき**ものである．以下に，面積が確定するための十分条件を少し挙げる．

定理 3.3.2 (B の面積確定のための十分条件) 図形 B の面積が定義できるための十分条件の２例は：

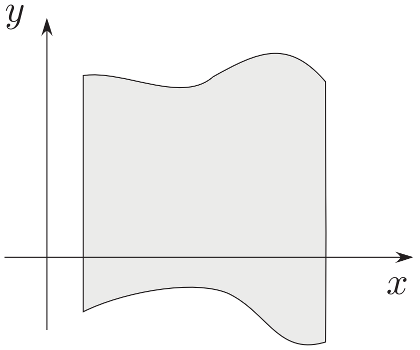
（１）B の境界が滑らかな曲線であること．つまり，$x = x(t), y = y(t)$ という閉曲線（$0 \le t \le 1$ かつ $x(0) = x(1), y(0) = y(1)$）があって，その内部が B であり，かつ，$x(t), y(t)$ が t の関数として C^1-級（一階微分可能で，導関数が連続）であること．

（２）$x = a, x = b$ の２つの直線と $y = \varphi(x), y = \psi(x)$ （ただし $a \le x \le b$ で $\varphi(x) \le \psi(x)$，かつ $\varphi(x)$ と $\psi(x)$ は x の連続関数）なる２つの曲線で囲まれた部分が B であること．

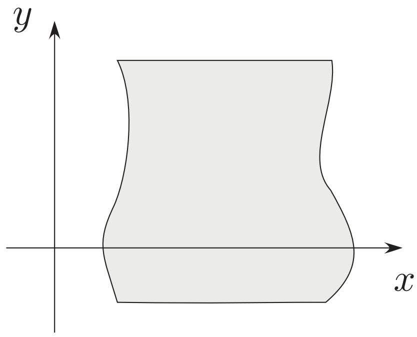
定理の（２）に出ているような図形を**縦線図形**という（下図参照）．上では y-方向の「縦線」でできた図形の例を示したが，x-方向の「横線」でできた図形，つまり

$$y = c, \quad y = d, \qquad x = \varphi(y), \quad x = \psi(y) \quad \text{で囲まれた図形} \tag{3.3.3}$$

に対しても定理は成り立つ．これは本来,「横線図形」と呼ぶべきだろうが．このような図形も「縦線図形」という．



縦線図形



横線図形

この準備の下で,

定義 3.3.3 (面積確定の図形の上の重積分) 面積が確定する図形 B が与えられたとする．B 上で定義された関数 $f(x,y)$ が与えられたとき，f の B での積分は，重積分

$$\iint_B f(x,y)dxdy \equiv \iint_{B^*} f(x,y)\chi_B(x,y)\,dxdy \tag{3.3.4}$$

によって定義する（B^* とは先に述べた，B を含む長方形）．右辺の積分が定義できる場合は f は B で可積分（積分できる），定義できない場合は f は B で積分できないという．

最後に，積分可能の十分条件を挙げておく．図形 B が面積確定との条件をつければ，だいたい，長方形上での積分と同じ事になる．

定理 3.3.4 (積分可能の条件) 図形 B の面積が定義できる時，その上の関数 f が B で積分可能なための十分条件は，f **が** B **上で連続**な事である．

3.4 重積分と累次積分

重積分をその定義から（分割を使って）求めるのは大変だ．ところが，n 重積分は（大抵の場合）n 個の 1 次元積分のくり返しで求められる．これは非常な省力化であり，実用上も非常に有り難い．この節ではこの重要な性質を学ぶ．

まずは簡単のために，$A=[a,b]\times[c,d]$（長方形） の上での $f(x,y)$ の重積分を考える．図形 A が前小節の「一般の図形」の場合については後で少しだけ触れる．

定理 3.4.1 (累次積分への帰着) 関数 $f(x,y)$ が A 上で積分可能とする．このとき，すべての $x\in[a,b]$ に対して

$$F(x)=\int_c^d f(x,y)dy \tag{3.4.1}$$

を定義できるならば,

$$\iint_A f(x,y)dxdy=\int_a^b F(x)dx=\int_a^b\left[\int_c^d f(x,y)dy\right]dx \tag{3.4.2}$$

が成り立つ．x,y を入れ替えた形の定理ももちろん，なりたつ．すなわち，すべての $y\in[c,d]$ に対して

$$G(y)=\int_a^b f(x,y)dx \tag{3.4.3}$$

を定義できるならば，

$$\iint_A f(x,y)dxdy = \int_c^d G(y)dy = \int_c^d \left[\int_a^b f(x,y)dx\right] dy \tag{3.4.4}$$

である．

講義で詳しく説明するが，$z = f(x,y)$ のグラフの下の体積を求めると思えば，この定理の主張を高校までの積分で理解できる．高校では，このような立体を x-軸に垂直な面で切り，その**断面の面積を積分**することで体積を求めた．(3.4.2) は，正にこれになっている［$F(x)$ が断面の面積に相当］．

記号：(3.4.2) の積分は，通常はカッコを省略して

$$\int_a^b \int_c^d f(x,y)\,dy\,dx \tag{3.4.5}$$

と書かれる（a,b,c,d と x,y の順番に注意）．ただ，これでは x,y どっちの変数がどこまで動くのかが混乱しがちなので，積分範囲を明確にするために

$$\int_a^b dx \int_c^d dy\, f(x,y) \tag{3.4.6}$$

と書くことも多い．(物理や工学では後者の書き方が一般的である.）後者では「積分記号の直後にある積分変数がその範囲を動く」点がわかりやすいが，その反面，この後に更に数式が続いた場合など，「どこまでが非積分関数か」わかりにくい面もある...

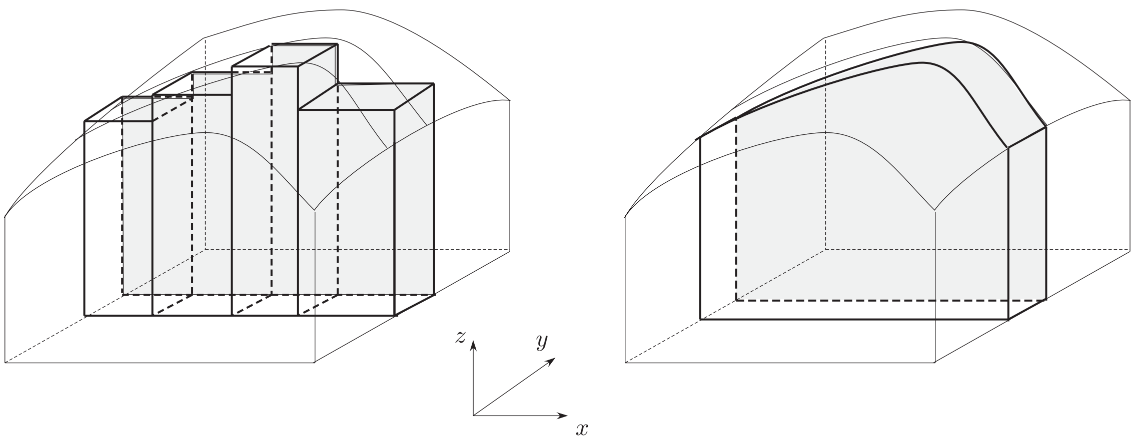
系 3.4.2 (Riemann 積分に対する Fubini の定理) 関数 f が A 上で積分可能のとき（両辺に意味がつく限り）

$$\int_a^b \left[\int_c^d f(x,y)dy\right] dx = \int_c^d \left[\int_a^b f(x,y)dx\right] dy \tag{3.4.7}$$

が成り立つ．つまり，累次積分の順序を交換できる．

定理 3.4.1 の証明について：立体の体積だと思えばアタリマエのようなものだが，一応，講義では大体の感じを説明する．要点は以下の通り：

重積分の定義において，x-方向の分割だけを先に細かくして，x-方向の分割の幅がゼロに行った極限を考える．y が一定の面で切った切り口を考えると，このとき，切り口に見えている短冊は素直に細くなって行くので，この切り口の面積が出てくる．この後に y-方向の分割を細かくすると，この切り口の面積を y-方向に積分したものが得られて，累次積分の公式に到達する． □



この定理は，その証明は大まかなアイディアがわかれば良い．それよりは実際に計算できる事が至上命題だ．いくつか例題を掲げておくからやっておくように．

問 3.4.1　A 上で積分可能であるが，上の $F(x)$ が（ある $x \in [a,b]$ に対しては）定義できないような $f(x,y)$ の例を作れ．

問 3.4.2　以下の重積分を計算せよ．$A = [0,1] \times [0,1]$．

$$\iint_A (x^2 + y^2) dxdy, \qquad \iint_A xy\, dxdy, \tag{3.4.8}$$

問 3.4.3　以下の場合に，重積分 $\displaystyle\iint_A f(x,y)\, dxdy$ を計算せよ..

a) $A = [1,3] \times [0,2]$，$f(x,y) = xy$.

b) $A = [0,1] \times [0,1]$，$f(x,y) = \dfrac{1}{3x+y+1}$.

c) A は 3 直線 $x = 0, y = 0, x + y = 1$ で囲まれた図形，$f(x,y) = \dfrac{1}{3x+y+1}$.

d) A は直線 $y = x$ と放物線 $y = x^2$ で囲まれた図形，$f(x,y) = (y - x^2)^2$.

（少し進んだ話題）

１．定理 3.4.1 は「$f(x,y)$ が A で積分可能ならば，重積分は累次積分で書ける」ことを主張している．この逆は成り立つかを考えたい．つまり，

$$\int_a^b \left[\int_c^d f(x,y) dy \right] dx = \int_c^d \left[\int_a^b f(x,y) dx \right] dy \qquad \text{（累次積分が定義できて積分順序によらない）} \tag{3.4.9}$$

ならば，f は A で可積分だろうか？

残念ながら，そうとは限らない．次のやや人工的な f は (3.4.9) が成り立つにもかかわらず，A で積分できない例である．このような意味で，累次積分の性質からもともとの重積分が定義できるかどうかを判定することはできない．

例： 正方形 $[0,1] \times [0,1]$ を S と書く．また，k-番目に大きい素数を p_k と書く（$k = 1, 2, \ldots$）．更に，適当な 1 以上の整数 k と整数 m, n を用いて $\left(\frac{m}{p_k}, \frac{n}{p_k}\right)$ の形に書ける S の内点の全体を T と書く：

$$T = \bigcup_{k=1}^{\infty} \left\{ \left(\frac{m}{p_k}, \frac{n}{p_k} \right) \middle| \, 0 < m < p_k,\ 0 < n < p_k \right\} \tag{3.4.10}$$

更に，

$$f(x,y) = \begin{cases} 0 & \big((x,y) \in T\big) \\ 1 & \big((x,y) \in S \backslash T\big) \end{cases} \tag{3.4.11}$$

と定義する．このとき，

$$\int_0^1 \left[\int_0^1 f(x,y) dy \right] dx = \int_0^1 \left[\int_0^1 f(x,y) dx \right] dy = 1 \tag{3.4.12}$$

であるが，$\displaystyle\iint_S f(x,y)\, dxdy$ は<u>定義できない</u>．

２．このように，Riemann 積分では何かと話がヤヤコシイのだが，この点は Lebesgue 積分を考えると大幅に改善される．粗っぽくいうと，上の反例などは大体が非常に些細なところから出ているので，その部分を「無視」するような定義を用いれば反例が消滅する．Lebesgue 積分というのはどの部分が「些細」で無視して良いかを合理的に決めたものとも考えられる．このような点から Lebesgue 積分は非常に自然なので，現在の数学では Lebesgue 積分が（Riemann 積分に替わって）用いられるようになった．

一般の図形上での積分に対する累次積分

長方形でない図形 B 上での積分を累次積分に帰着することも，同様に考えていける．たとえば，B が

$$x = a, \quad x = b, \qquad y = \varphi(x), \quad y = \psi(x) \tag{3.4.13}$$

で囲まれた縦線図形の場合（$a \le x \le b$ で $\varphi(x) < \psi(x)$），

$$\iint_B f(x,y)\,dxdy = \int_a^b \left[\int_{\varphi(x)}^{\psi(x)} f(x,y)dy \right] dx \tag{3.4.14}$$

が成り立つ．

重積分の順序交換

系 3.4.2（Fubini の定理）で非積分関数が「良い」性質を持っていれば重積分の順序は交換できる，ことをみた．これは重積分を累次積分に帰着することから出てきたもので，考え方は非常に単純である —— 要するに，与えられた積分範囲をどんな順番でも良いから尽くすように覆えばよいわけだ．でもこれは応用上，非常に大事なものである．

例えば，問 3.4.3 c) では $x = 0, y = 0, 2x + y = 1$ で囲まれた三角形の領域での積分を考えた．上の解答にも載せたように，x, y どちらの積分を先にやっても構わない．また，例えば x での積分を先にやる形の累次積分で与えられた積分の順序を変えて，y から積分しても良い．問題によっては，このように順序を変えることで簡単に計算できる場合があるので，これは実際問題としては大事である．(以下の具体例をみよ．)

問 3.4.4　積分領域 B を，$(x-1)^2 + y^2 \le 1$ と $y \ge 0$ で囲まれた図形として，積分 $\iint_B x^2 y\,dxdy$ を計算せよ．（この問題は x, y のどちらか一方を先にやると簡単にできる．順序を逆にすると大変だよ．）

順序交換は実用上，非常に大事だから，いくつか例題を挙げておく：

問 3.4.5　以下の積分の順序交換を納得せよ（f は適当な関数，$a, b > 0$ は定数）

$$\int_0^a dx \int_{-bx}^{bx} dy\, f(x,y) = \int_0^{ab} dy \int_{y/b}^a dx\, f(x,y) + \int_{-ab}^0 dy \int_{-y/b}^a dx\, f(x,y),$$

$$\int_0^{1/2} dx \int_{x^2}^x dy\, f(x,y) = \int_0^{1/4} dy \int_y^{\sqrt{y}} dx\, f(x,y) + \int_{1/4}^{1/2} dy \int_y^{1/2} dx\, f(x,y)$$

このような問題では，積分領域の図を描いて，間違わないように書き換えるのが良い．

発展問題：

重積分の順序交換の応用として，1 変数関数のテイラーの定理（剰余項の表式つき）を導いてみよう．簡単のため，関数 $f(x)$ はすべての x で無限階微分可能だとする．

1. 微分積分学の基本定理から以下が成り立つことに注意しよう（ここで $f'(t)$ は $f(t)$ の，t に関する導関数）

$$f(x) = f(a) + \int_a^x f'(t)dt, \qquad \text{同様に，}\quad f'(t) = f'(a) + \int_a^t f''(s)ds \tag{3.4.15}$$

2. 右の式を左の式に代入して，$f(x)$ の表式を作れ．そこに出てくる 2 重積分の順序を交換して

$$f(x) = f(a) + f'(a)(x-a) + \int_a^x (x-s)f''(s)ds \tag{3.4.16}$$

が成り立つことを示せ．

3. これを帰納法的にくり返して，$f(x)$ の n 次のテイラー展開の式を求めよ．

この結果として得られる表式は剰余項を積分で与えてくれるものなので，かなり使いやすい．通常は「区間 $[a, x]$ 中の一点 x_1 があって，$f^{(n+1)}(x_1)(x-a)^{n+1}/(n+1)!$ が剰余項」などとするが，これでは x_1 がどこにあるのかわかりにくいので，困ることがある．

3.5　重積分の変数変換

1 変数関数の積分では変数変換（置換積分）の公式が存在した，多重積分においても同様の公式が存在し，かつ実際上，非常に有用である．

１次元の場合を思い出そう．この場合，$x = x(t)$ と変数変換すると，

$$\int_{x_1}^{x_2} f(x)\,dx = \int_{t_1}^{t_2} f(x(t))\,x'(t)\,dt \tag{3.5.1}$$

であった（t_1, t_2 は $x(t)$ がそれぞれ x_1, x_2 になる t の値）．x と t の間で，座標が伸び縮みした分を考慮に入れるために，$x'(t)$ が必要になったのである．

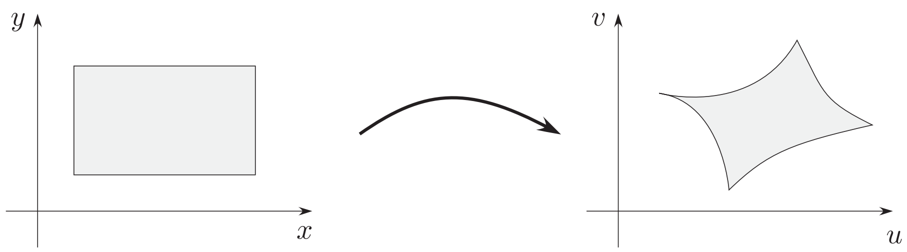
２重積分の時に，これに相当するものは何だろうか？今，(x, y) から新しい座標 (u, v) に移ることを考える．ここで新しい座標系が (u, v) だが，旧い座標を新しい座標で表して，(x, y) と (u, v) の関係を

$$x = x(u, v), \qquad y = y(u, v) \tag{3.5.2}$$

と書くことにする．例としては，$x = u + v, y = u - v$ などを想定して欲しい．このとき，(x, y) でみた時の積分領域 A が，(u, v) では B になるとしよう．また，上の変数変換をして f を表したものを $g(u, v)$ と書こう：

$$g(u, v) \equiv f(x(u, v), y(u, v)). \tag{3.5.3}$$

さて，問題：重積分 $\displaystyle\iint_A f(x, y)dxdy$ は，u, v での重積分として，どのように書けるだろうか？下図の左が領域 A，右が変数変換後の領域 B のつもりである．



単純に考えて，積分領域 A が B になるのだから，

$$\text{（間違い !! ）} \qquad \iint_A f(x, y)\,dxdy = \iint_B g(u, v)\,dudv \qquad \text{（間違い !! ）} \tag{3.5.4}$$

となると思ったら，一般には間違いである．これが間違いであることは，１次元の時を思いだせば，ある程度は理解できる．１次元の場合，区間 $[x_1, x_2]$ が区間 $[t_1, t_2]$ に変わったからと言って，

$$\text{（間違い !! ）} \qquad \int_{x_1}^{x_2} f(x)\,dx = \int_{t_1}^{t_2} f(x(t))\,dt \qquad \text{（間違い !! ）} \tag{3.5.5}$$

ではなく，正しくは

$$\int_{x_1}^{x_2} f(x)\,dx = \int_{t_1}^{t_2} f(x(t))\,x'(t)\,dt \tag{3.5.6}$$

だった．しつこいけど，**変数変換によって区間が伸び縮みする効果**を考えに入れるために，$x'(t)$ が必要だったわけね．

重積分でも事情は同じで，変数変換によって座標が伸び縮みした効果を表すものが必要である．ただし，考えている座標の変換が２次元的だから，伸び縮みだけでなく，「ひねり」の要素も加わるので，話がややこしい．

答えを言ってしまうと，以下のようになる．

定理 3.5.1 変数変換 (3.5.2) に対応して，ヤコビアン と呼ばれる関数 $J(u, v)$ を，以下の行列式で定義する：

$$J(u, v) \equiv \frac{\partial(x, y)}{\partial(u, v)} \equiv \det \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix}. \tag{3.5.7}$$

また，変数変換 (3.5.2) は十分に性質の良いもの，つまり

- 領域 A と B の点が 1 対 1 に対応し，
- $x = x(u,v)$ と $y = y(u,v)$ が偏微分可能で導関数が連続，
- B 内でヤコビアン $J(u,v)$ がゼロでない

とする．このとき，

$$\iint_A f(x,y)\,dxdy = \iint_B g(u,v)\,|J(u,v)|\,dudv \tag{3.5.8}$$

である．

- A の点と B の点が 1 対 1 に対応することは非常に重要だ．もし例えば 1 ： 2 であったら，A を B でみると 2 回数えてしまってる感じになって，答えが 2 倍ずれてくる．
- 上の定理では，ヤコビアンの**絶対値**をとったものが現れていることにも注意．1 次元の積分では $x'(t)$ （絶対値ではない）が出ていたことと違う．この理由は，重積分では本質的に「積分の向き」がないことに関係している．

非常に重要な例：平面の極座標

直交座標 (x,y) から極座標 (r,θ) への変換を考えよう．座標変換の式は

$$x = r\cos\theta, \qquad y = r\sin\theta \tag{3.5.9}$$

であるから，ヤコビアンは

$$J(r,\theta) = \det\begin{bmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{bmatrix} = r\cos^2\theta + r\sin^2\theta = r \tag{3.5.10}$$

というわけで，皆さんのよく知っている（はずの）$dxdy$ を $rdrd\theta$ に変換するのが出てきた..

言うまでもなく，このような変数変換は，それをやることによって初めて積分できる場合が多いから重要なのだ．例えば積分 $\displaystyle\iint_{x^2+y^2\le 1} e^{-(x^2+y^2)}dxdy$ はこのままでは積分が非常に難しい．しかし，極座標に変換すると

$$\iint_{x^2+y^2\le 1} e^{-(x^2+y^2)}dxdy = \int_0^1 dr\,r\int_0^{2\pi} d\theta\,e^{-r^2} = 2\pi\int_0^1 e^{-r^2}\,r\,dr = 2\pi\left[-\frac{e^{-r^2}}{2}\right]_0^1 = \pi(1-e^{-1}) \tag{3.5.11}$$

と計算できる．

このような応用例としては

$$\int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi} \tag{3.5.12}$$

の証明がある．答えを知ってないととても出来そうにないが，これは

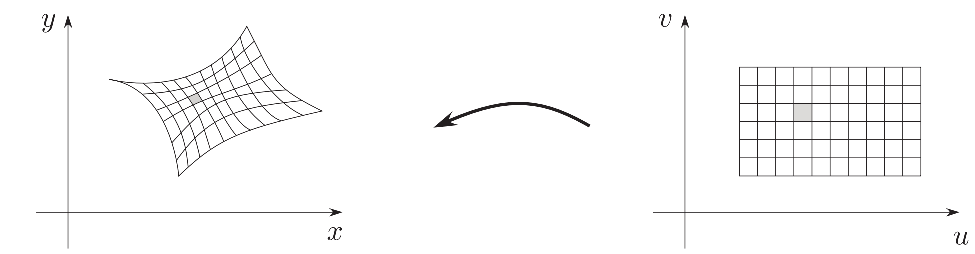
$$\left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)^2 = \int_{-\infty}^{\infty} e^{-x^2}dx\int_{-\infty}^{\infty} e^{-y^2}dy = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}dxdy \tag{3.5.13}$$

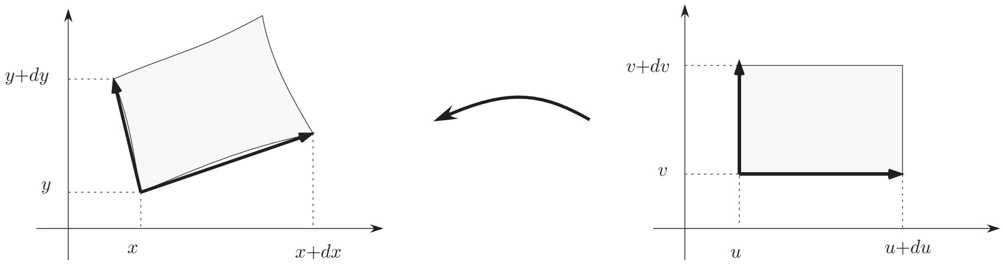
と考えて極座標に変換すると計算できるのだ.（これも箱崎日に散々聞かされたかな？）

ヤコビアンの意味（変数変換の公式の略証）

上の変数変換の式（ヤコビアン）が出てくる理由を述べよう．そのためには重積分の定義に戻って考えるのが良い．

何度も強調したように，$\iint_A f(x,y)dxdy$ とは，xy 座標系を細かく四角に区切って，その四角の面積と $f(x,y)$ の値をかけたものを足し併せ（たものの極限をと）る，ことだった．同様に，$\iint_B h(u,v)dudv$ は，uv 座標系での四角の面積と h の値をかけて和をとるわけ．





さて，uv-平面での小さな四角 $[u, u+du] \times [v, v+dv]$ を考えよう．これがもとの xy-平面で囲む図形は，その４つの頂点が

$$(x(u,v), y(u,v)),\ \ (x(u+du,v), y(u+du,v)),\ \ (x(u,v+dv), y(u,v+dv)),\ \ (x(u+du,v+dv), y(u+du,v+dv)) \tag{3.5.14}$$

で与えられる，近似的な平行四辺形になる．この平行四辺形を作る２つのベクトルは du, dv が非常に小さいとすると，

$$\begin{bmatrix} x(u+du,v)-x(u,v) \\ y(u+du,v)-y(u,v) \end{bmatrix} \approx \begin{bmatrix} \dfrac{\partial x}{\partial u}du \\ \\ \dfrac{\partial y}{\partial u}du \end{bmatrix} = \begin{bmatrix} \dfrac{\partial x}{\partial u} \\ \\ \dfrac{\partial y}{\partial u} \end{bmatrix} du \quad と \quad \begin{bmatrix} x(u,v+dv)-x(u,v) \\ y(u,v+dv)-y(u,v) \end{bmatrix} \approx \begin{bmatrix} \dfrac{\partial x}{\partial v}dv \\ \\ \dfrac{\partial y}{\partial v}dv \end{bmatrix} = \begin{bmatrix} \dfrac{\partial x}{\partial v} \\ \\ \dfrac{\partial y}{\partial v} \end{bmatrix} dv \tag{3.5.15}$$

である．ところで，ベクトル (a, b) と (c, d) の作る平行四辺形の面積は，行列 $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ の行列式，つまり $ad - bc$（の絶対値）で与えられた（線形代数の講義を思い出そう；忘れていても，初等的にも導けるよ）．これを用いると，考えている近似的な平行四辺形の面積は以下（の絶対値）になるはずだ．

$$\det \begin{bmatrix} \dfrac{\partial x}{\partial u}du & \dfrac{\partial x}{\partial v}dv \\ \\ \dfrac{\partial y}{\partial u}du & \dfrac{\partial y}{\partial v}dv \end{bmatrix} = \det \begin{bmatrix} \dfrac{\partial x}{\partial u} & \dfrac{\partial x}{\partial v} \\ \\ \dfrac{\partial y}{\partial u} & \dfrac{\partial y}{\partial v} \end{bmatrix} dudv = J(u,v)\, du\, dv \tag{3.5.16}$$

このようにしてヤコビアンが登場するのである．　□

3.6　３次元以上の重積分

（時間の関係で，この節はおまけ気味である．）

いままで，平面上の領域での重積分を考えてきた．空間内での積分（３重積分）は平面の場合と全く同様のアイディアで定義される．ただし，平面の場合には考える積分領域（２次元）を細かい長方形のメッシュで覆ったが，３次元の場合には積分領域（３次元）を細かい直方体で覆うところが異なる．（これ以外は全く同じだからくり返さない．でもよく考えると，１次元と２次元の差も，覆う対象が１次元の領域か２次元の領域か，それに応じて分け方を変えただけだった．）４次元以上での積分（４重積分，５重積分）も同様に定義できる．

これらの n 重積分も，２重積分と同様の性質を持っている．すなわち，

1. n 重積分は，非積分関数と積分領域の性質が良ければ，n 個の累次積分で表せる．実際に累次積分に直すには，n-次元空間での積分領域を図示して（4 次元以上では不可能だが，少なくともできるだけ思い浮かべて），丁寧に直していけばよい．
2. n 重積分においても，性質の良い変数変換に対しては，ヤコビアンを用いた変数変換の公式が成り立つ．勿論，この場合のヤコビアンは $n \times n$ 行列の行列式である．

ヤコビアンについて補足しておく．n 次元での元々の座標が $(x_1, x_2, \ldots, x_n)$, 新しい座標が $(u_1, u_2, \ldots, u_n)$ で，各 x_i は u_1 から u_n の関数として書けているとする．このとき，

$$\iint \cdots \int_A f(x_1, x_2, \ldots, x_n)\, dx_1 dx_2 \cdots dx_n = \iint \cdots \int_B g(u_1, u_2, \ldots, u_n) \left| \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} \right| du_1 du_2 \cdots du_n \tag{3.6.1}$$

となる．ここで B は新しい座標 $(u_1, u_2, \ldots, u_n)$ で見た積分領域 A のことであり，g は対応する点での f の値を表す．また，ヤコビアンは

$$\frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(u_1, u_2, \ldots, u_n)} = \det \begin{bmatrix} \frac{\partial x_1}{\partial u_1} & \frac{\partial x_1}{\partial u_2} & \cdots & \frac{\partial x_1}{\partial u_n} \\ \frac{\partial x_2}{\partial u_1} & \frac{\partial x_2}{\partial u_2} & \cdots & \frac{\partial x_2}{\partial u_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial x_n}{\partial u_1} & \frac{\partial x_n}{\partial u_2} & \cdots & \frac{\partial x_n}{\partial u_n} \end{bmatrix} \tag{3.6.2}$$

という $n \times n$ 行列の行列式である．

座標変換で最も重要なのは極座標への変換であろう．３次元の場合，これは

$$x = r\,\sin\theta\,\cos\phi, \quad y = r\,\sin\theta\,\sin\phi, \quad z = r\cos\theta \tag{3.6.3}$$

で，(r, θ, ϕ) の動く範囲はそれぞれ $r \geq 0, 0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$ である．この場合のヤコビアンは（今までに散々出てきただろうが，各自確かめる）

$$\frac{\partial(x, y, z)}{\partial(r, \theta, \phi)} = \det \begin{bmatrix} \sin\theta\,\cos\phi & r\cos\theta\,\cos\phi & -r\sin\theta\,\sin\phi \\ \sin\theta\,\sin\phi & r\cos\theta\,\sin\phi & r\sin\theta\,\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{bmatrix} = r^2 \sin\theta \tag{3.6.4}$$

となる．以下に極座標関連の発展問題を２つ載せる．どちらもかなり大変だから，無理にできるようになる必要はない．今できることよりも，将来，何らかの役に立つだろうと思って載せている．

発展問題その１：

４次元以上の極座標と，そのヤコビアンがどうなるか，一度やってみると良い．ただし，一般次元では計算が非常に大変であるから，相当の覚悟が必要．

発展問題その２：

n を３以上の整数とする．原点を中心とする半径 r の n-次元球とは，$x_1^2 + x_2^2 + \cdots + x_n^2 \leq r^2$ を満たす点の集合である．こいつの体積 $V_n(r)$ を求めるには，いくつかの方法がある．

- 上の発展問題で求まったはずのヤコビアンを用いて，n-次元極座標へ変換した積分を行う．非積分関数には，この球の定義関数をとればよい．この方法は一番の基本だが，ヤコビアンの計算が大変で死ぬことが多い...
- この球を x_n が一定の面で切ると，その切り口は半径が $\sqrt{r^2 - x_n^2}$ の，$(n-1)$ 次元球になる．この切り口の面積（体積）を x_n で積分することで，n 次元球の体積が求められるはずだ：

$$V_n(r) = \int_{-r}^{r} dx_n\, V_{n-1}(\sqrt{r^2 - x_n^2}) \tag{3.6.5}$$

一方，相似な図形の体積を考えると，半径 r と半径 1 の球の体積の比は丁度 r^n のはずである（a_n は係数）：

$$V_n(r) = r^n V_n(1) = r^n \times a_n \tag{3.6.6}$$

この２つの式を組み合わせると，a_n と a_{n-1} の間の漸化式が得られる．２次元球（円）や３次元球では $a_2 = \pi, a_3 = 4\pi/3$ を知っているから，漸化式を解くことで a_n と $V_n(r)$ がわかる．

- n 次元のガウス積分 $\displaystyle\iint\cdots\int_{\mathbb{R}^n} e^{-(x_1^2+x_2^2+\ldots+x_n^2)}\,dx_1dx_2\ldots dx_n$ を以下の別々の方法で計算して，結果を比較する．
  - 指数関数が積に分かれることから，各成分でバラバラに積分する．結果は $\pi^{n/2}$ のはず．
  - n 次元の極座標に変数変換するつもりになる．しかし，今は非積分関数が球対称だから，角度成分は積分できてしまって，結局 $\int_0^\infty drr^{n-1}c_n\,e^{-r^2}$ の形になるはずだ．ここで c_n は n–次元の立体角（半径 1 の n-次元球の表面積）であり，上で出てきた a_n とは，$c_n = na_n$ の関係にある．

  両者を等置して，

$$\pi^{n/2} = n\,a_n \int_0^\infty dr\,r^{n-1}\,e^{-r^2} \tag{3.6.7}$$

  となるはずで，これから a_n と $V_n(r)$ が求まる．(右辺の積分は Γ-関数で表される．)