2010.04.13.

微分積分学　（医学部，S10 クラス）

担当：原　隆（数理学研究院）：伊都キャンパス数理研究教育棟 219 号室，tel: 092-802-4441,
e-mail: hara@math.kyushu-u.ac.jp,　http://www2.math.kyushu-u.ac.jp/~hara/lectures/lectures-j.html

Office hours: 月曜の午後 5 時～ 6 時半頃，僕のオフィスにて（ただし，その前のセミナーが長引いた場合には少し待って頂くことになります）．なお講義終了後にも質問を受け付けますし，これ以外でもお互いの都合の良い時間にお相手します．

概要： 以下のすべてをカバーすることは不可能なので，適当に取捨選択します．

- 1 変数関数と微分の話題
  - 合成関数の微分
  - 逆関数とは；逆三角関数の微分
  - 平均値の定理とテイラー展開
- 積分の話題
  - 積分とは何か（区分求積法による定義）
  - いろいろな積分法（置換積分，無理関数や有理関数の積分）
  - 広義積分
- 多変数の微分の話題
  - 多変数の微分（偏微分）
  - 連鎖律
  - 多変数の極値問題
- 多変数の積分（重積分）の話題
  - 重積分の定義
  - 累次積分への帰着
  - 変数変換

教科書： 磯崎，筧，木下，籠屋，砂川，竹山共著「微積分学入門」（裳華房）です．わかりやすく，材料を精選して書かれた良い本だと思います．

参考書： もっと突っ込んで勉強したい人には，斎藤正彦著「微分積分学」（東京図書）をお薦めします．

評価方法：

主に中間試験，レポートと期末試験の成績を総合して評価する．

> 詳しい評価方法を一旦，ここに書いたのだが，このクラスの雰囲気，指向性がよくわからないので，現時点で決めるのは危険だと気づいた．第２回目か３回目くらいまでには確定します．

「学習到達度再調査」について：

今学期，原の担当科目が異常に多いので，再調査を行う余裕がありません．再調査はないものと思って，しっかり勉強して下さい．

「例題」などについて：

教科書は，その副題にあるように，問題が充実しています．そこで，教科書の関連する問題はできるだけ自分で解き，理解するように心がけて下さい．場合によっては，簡単なレポートや「お奨めの宿題問題」を出す可能性もあります．この講義をこなす上では重要な意味があるので，是非，やって下さい．

「レポート」の作成はみんなで協力してやっても構わないし，むしろ協力することを奨励します．ただし、（友達と協力してレポート問題を解いた場合でも）各人のレポートは自分の言葉で記述し、かつ、「○○君と一緒に考えました」とぐらいは書きましょう．また，教えてもらった事はそのままにせず，自分でもう一回考えて納得しましょう．

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

- まず初めに，学生生活の最大の目的は勉強すること であると確認する．
- 講義中の私語，ケータイの使用はつつしむ．途中入室もできるだけ避ける（どうしても必要な場合は周囲の邪魔にならないように）．これらはいずれも講義に参加している 他の学生さんへの 最低限のエチケットです．
- 僕の方では時間通りに講義をはじめ、時間通りに終わるよう心がける．
- 重要な連絡・資料の配付は原則として講義を通して行う（補助として僕のホームページも使う —— アドレスは最初に載せた）．「講義に欠席したから知らなかった」などの苦情は一切，受け付けない．
- レポートを課した場合，その期限は厳密に取り扱う．
- E-mail による質問はいつでも受け付ける（hara@math.kyushu-u.ac.jp）ので積極的に利用するように．ただ，回答までには数日の余裕を見込んで下さい．

1 微分の話題

1.1 関数と逆関数

まずは高校の復習から始めよう．関数 f とは実数 x に対して実数 $f(x)$ を対応させる「対応関係」のことだ．例えば $f(x) = x^2$ というのは，実数 x に対して x^2 を対応させる．この際，注意すべきは（関数の定義域中の）x に対しては $f(x)$ が一意的に定まる（定まるようなもののみを関数という）ということ．

（念のため）ここで言ってるのは各 x に対してその行き先 $f(x)$ が一つに決まる，ということである．だから，$x_1 \neq x_2$ に対して $f(x_1) = f(x_2)$ となっていても構わない（異なる x の行き先が同じであっても良い）．x から $f(x)$ の方向で対応関係を見た場合,「多対 1」の対応は別に構わないが,「1 対多」はダメ，ということだ．

では逆関数を考えよう．関数

$$y = f(x) \tag{1.1}$$

が与えられた時，これを逆に解いて x を y で表す式

$$x = g(y) \qquad （ここで\ \ y = f(x)） \tag{1.2}$$

を作ってみる．この y から x への対応 g を「f の逆関数」というのだ（すぐ後の補足も参照）．要するに x の行き先が $y = f(x)$ である場合，y から x で戻ってくる対応関係を与えるものが逆関数なのである．

（注）一般に関数 f を考える時には，元の数を x，行き先を y と書くことが多く，$y = f(x)$ と書く．これに従うと，逆関数も $y = g(x)$ の流儀. で書きたい気がする．そこで (1.2) の x, y を交換して $y = g(x)$ と書くことが多く，教科書などでも高校でもこのように書いて習っただろう．もともと x, y は特定の数ではなく変数だからこのように入れ替えても一向に構わない．ただし，こうすると x, y の意味が元とは逆になっており，よく混乱するから注意しよう．

（補足）もちろん，上の (1.2) がいつも関数になるとは限らない．y から x に戻ろうとすると，戻り先が２つ以上ある場合もある．例えば，$y = x^2$ の時は $x = \pm\sqrt{y}$ であって,（$y > 0$ なら）もとの x は正負の２通りあるのだ．そもそも関数とは「行き先が一通りに定まるもの」だったから，これでは困る．そこで元の関数 $f(x)$ の定義域を制限して，x から y への対応が「1 対 1」になるようにし，この部分でのみ，逆関数を考える．$y = x^2$ の例なら，例えば $x \geq 0$ のみを考えると，y からもとの x へは $x = \sqrt{y}$ と一意に戻れるようになる．そこで $x \geq 0$ の範囲での $f(x) = x^2$ の逆関数 g は $g(y) = \sqrt{y}$ と決まる．

（例）指数関数 $y = e^x$ を逆に解くと $x = \log y$ となる．つまり，指数関数と対数関数は互いに逆関数の関係にある．

1.2 逆三角関数

以上は高校の復習．さて，この考え方を三角関数に適用しよう．三角関数は周期関数なので,（例えば）$y = \sin x$ の y を指定した場合，それを与える x は一杯ある．そこで上の（補足）のように，x の範囲を制限 して考える．具体的には $\sin x$ の逆関数を考える際には $|x| \leq \pi/2$ とすることが普通である．この範囲では $y = \sin x$ の x と y の対応は 1 対 1 だから，y から x への対応を一意に決めることができる．これを sin の逆関数と言い，

$$x = \arcsin y \tag{1.3}$$

と書く．この式の意味はあくまで

$$y = \sin x \tag{1.4}$$

を逆に解いただけ，でそれ以上でも以下でもない．

同様に cos の逆関数は $0 \le x \le \pi$ で考えて

$$y = \cos x \qquad \Longleftrightarrow \qquad x = \arccos y \tag{1.5}$$

となる関数 arccos と定義する．
tan の逆関数は $|x| < \pi/2$ で考えて

$$y = \tan x \qquad \Longleftrightarrow \qquad x = \arctan y \tag{1.6}$$

となる関数 arctan と定義する．

さて，このような逆三角関数の微分を考えてみよう．じつはこれは無意識のうちには「置換積分」で使っていることではあるのだが...

少し一般論を復習しよう．f の逆関数を g とすると，その定義から（$y = f(x)$ の場合）

$$g(f(x)) = g(y) = x \tag{1.7}$$

となっているはずである．そこでこの両辺を x で微分すると，左辺には合成関数の微分を用いて

$$g'(f(x))\, f'(x) = 1 \qquad \text{つまり} \qquad g'(f(x)) = \frac{1}{f'(x)} \tag{1.8}$$

が成り立つはずだ．$y = f(x)$ および $x = g(y)$ を用いて上のを y だけで書いてみると

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'\bigl(g(y)\bigr)} \tag{1.9}$$

が得られる．これが逆関数 $g(y)$ の導関数を f の導関数と $g(y)$ で表す式である．
これをまずは $g(y) = \arcsin y$ に適用してみよう．この場合，もとの f は $f(x) = \sin x$ であり，$f'(x) = \cos x$ だ．そこで

$$\frac{d}{dy} \arcsin y = \frac{1}{\cos\bigl(\arcsin y\bigr)} = \frac{1}{\cos x} \qquad \text{ただし，} x = \arcsin y \tag{1.10}$$

が得られるが，右辺の分母はもっと簡単になる．実際，$y = \sin x$ ならば $\cos x = \pm\sqrt{1 - x^2}$ のはずである（図を書いて納得しよう）．今 $|x| \le \pi/2$ だったからここでは cos は非負であり，$\cos x = \sqrt{1 - x^2}$ なのだ．従って結局，

$$\frac{d}{dy} \arcsin y = \frac{1}{\sqrt{1 - y^2}} \tag{1.11}$$

が得られる．$\arcsin y$ はよく訳がわからないのにその導関数は平方根を使って書けるのが面白いところだ．
同様に議論すると

$$\frac{d}{dy} \arccos y = \frac{1}{-\sin\bigl(\arccos y\bigr)} = -\frac{1}{\sqrt{1 - y^2}} \tag{1.12}$$

および

$$\frac{d}{dy} \arctan y = \frac{1}{\bigl\{\sec\bigl(\arctan y\bigr)\bigr\}^2} = \bigl\{\cos\bigl(\arctan y\bigr)\bigr\}^2 = \frac{1}{1 + y^2} \tag{1.13}$$

が得られる．

これを用いると置換積分がいろいろとできる．例は教科書を見ることにしよう．

４月２７日：今日から平均値の定理とテイラー展開に入ります．
成績評価のやり方について：大体，以下のように定めます．

- 成績の基準になる点としては，**中間試験と期末試験の点の平均**を用いる．
- 上の「最終素点」をよく見て，必要ならば全体に少し修正を加えたものをつくり，これをこの大学の基準と合わせて最終成績を出す．
- レポートを出題した場合，レポートの結果は原則としては最終素点には加えない．しかし，上の計算では合格基準に少し足りない人（百点満点で 10 点不足が限度）を助けるかどうかに使用する．また，チャレンジ問題などでずば抜けた解答をした人にも特例措置を講ずるかもしれない．
- 成績にはレポートはあまり関係ないのだが，これをやることで理解も深まるから，決しておろそかにしないで，ちゃんと取り組んで下さい．
- なお，上で「期末試験」と書きましたが，場合によっては，期末試験期間の前の週に試験をしてしまうかもしれません．

なお，教科書のすべてをやる時間はないので,「逆関数」の後は「平均値の定理」「テイラーの公式」へ跳びます．

1.3 平均値の定理

（この節の内容は大体，教科書の 1.8 節）

直感的には簡単な定理（下図左を参照）．あまり厳密な事を言っても仕方ないので，ここは「こんな定理もある」ことと「その直感的な意味」さえわかれば十分だ．

定理 1.1 (平均値の定理；教科書の p.17) $f(x)$ 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，

$$\frac{f(b)-f(a)}{b-a} = f'(\xi) \qquad (a<\xi<b) \tag{1.14}$$

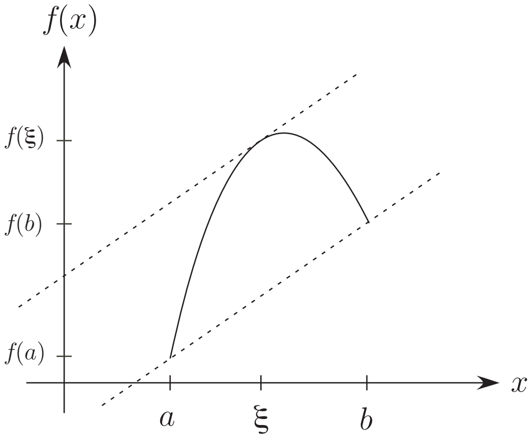
となる ξ が存在する．

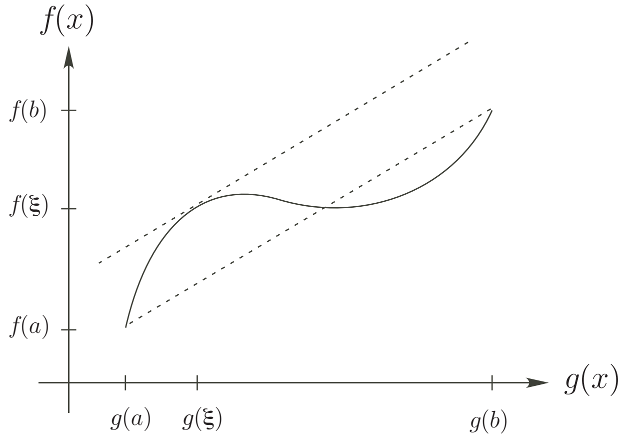
（注）ξ は一般には a,b の両方に依存して決まる．なお，上で (a,b) は「$a<x<b$ なる x の集合」，$[a,b]$ は「$a \le x \le b$ なる x の集合」のこと．

定理を応用する場合，分母を払って

$$f(b) = f(a) + f'(\xi)(b-a) \tag{1.15}$$

と書いた方が良いかもしれない．$f(b)$ の値を $f(a)$ から「計算したい」希望を表している式である（詳細は講義で）．この方向をもっと進めると「テイラーの定理」になる（次節）．





下の形の定理も役に立つことが多い（上図右を参照）．ただし，これは教科書にも載っていないので，深入りはしないことにする．

> **定理 1.2 (コーシーの平均値の定理；教科書にはない？)** $f(x)$ 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，
>
> $$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} \qquad (a < \xi < b) \tag{1.16}$$
>
> となる ξ が存在する．

（注）$g'(x) \neq 0$ から $g(a) \neq g(b)$ が保証されるので，左辺の分母はゼロでない．

なお，後ででてくる（積分形の）テイラーの公式に似せて書くならば，(1.15) を

$$f(b) - f(a) = \int_a^b f'(t)\, dt \tag{1.17}$$

と書いた方が良いかもしれない．これは微分と積分の関係を表す，アタリマエの式ではあるが，この類似物が後で非常に役に立つ（定理 1.3）ことを見るだろう．

1.4 テイラーの定理とテイラー展開

（大体，教科書の 1.9 節）

（大雑把な話）「テイラー展開」とは $f(x)$ の値を $f(a)$ とその高階微係数で表す表式で，

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{f^{(n)}(a)}{n!}(x - a)^n \tag{1.18}$$

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

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots = \sum_{n=0}^{\infty} \frac{1}{n!}\, x^n \tag{1.19}$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!}\, x^{2n+1} \tag{1.20}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!}\, x^{2n} \tag{1.21}$$

などとなる．

これはやはり**驚異的な式**である !!（と共感して欲しい...）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，(1.19) 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[1]．
- (1.18) 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（数学の）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．

1.4.1 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 1.3 (通常のテイラーの公式，教科書の p.19) $f(x)$ 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：

$$f(x) = f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)\,dt \tag{1.22}$$

がなりたつ．なお，(1.22) の 2 つの項に名前をつけて

$$f(x) = S_n(x) + R_n(x), \tag{1.23}$$

$$S_n(x) := f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad R_n(x) := \frac{1}{(n-1)!}\int_a^x (x-t)^{n-1} f^{(n)}(t)\,dt \tag{1.24}$$

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

この定理の証明は教科書の p.20 に載っているので略す．また，剰余項については，以下の表式もある：

定理 1.4 (通常のテイラーの公式その 2，教科書にはない？) 上の定理の剰余項は，また

$$R_n(x) := \frac{f^{(n)}(\xi)}{n!}(x-a)^n \tag{1.25}$$

とも書ける．ここで，ξ とは a と x の間にある数で，a と x に依存して決まる．

Remark.

- $a=0$ 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．
- 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．

[1] 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 (1.19) 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

- 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 1.4 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 1.3 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）．

1.4.2　テイラーの公式の例

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，既にテイラーの公式の形になっている．念のため，テイラーの公式を用いたら多項式が再現される事を各自で確かめてみよう．
- 指数関数．$f(x) = e^x$ は何回でも微分可能で，高階の導関数もすべて e^x である．従って，特に $a = 0$ としたテイラーの公式から

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x), \qquad R_n(x) = \frac{1}{(n-1)!} \int_0^x (x-t)^{n-1} e^t dt \tag{1.26}$$

が得られる．また剰余項は

$$R_n(x) = \frac{e^\xi}{n!} x^n \tag{1.27}$$

とも書ける（ξ は 0 と x の間の数）．

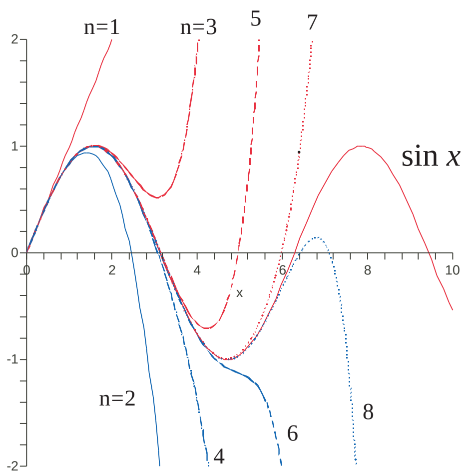
- 三角関数（$\sin, \cos$）も同様にして展開式を導くことができる．例えば

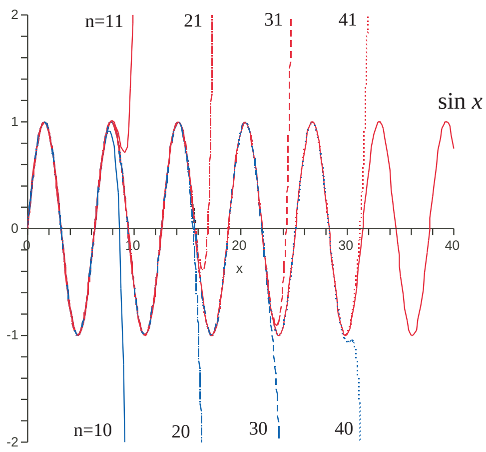
$$\sin x = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad R_n(x) = \frac{1}{(n-1)!} \int_0^x (x-t)^{n-1} \sin t\, dt \tag{1.28}$$

がなりたつ．

参考までに $\sin x$ のテイラーの公式の図を載せておく．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．





$\arctan x$ の例．$\tan x$ の逆関数を $\arctan x$ という．この導関数は

$$(\arctan x)' = \frac{1}{1+x^2}$$

である．また，$\arctan 0 = 0$ である．従って，導関数は

$$(\arctan x)' = \frac{1}{1+x^2}, \quad (\arctan x)'' = -\frac{2x}{(1+x^2)^2}, \quad (\arctan x)''' = \frac{2(3x^2-1)}{(1+x^2)^3}, \quad (\arctan x)'''' = -\frac{24x(x^2-1)}{(1+x^2)^4}, \ldots$$

となって行くので，

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \ldots$$

と見える．剰余項もちゃんと書くと，例えば $n=7$ では

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{-7\xi^6 + 35\xi^4 - 21\xi^2 + 1}{7(1+\xi^2)^7} \qquad (\xi \text{ は } 0 \text{ と } x \text{ の間の数})$$

が得られる．

さてここで $x=1$ としてみよう．$\tan$ が 1 になるのは $\pi/4$ のときなので，上の式は

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7}\frac{-7\xi^6 + 35\xi^4 - 21\xi^2 + 1}{(1+\xi^2)^7} \qquad (\xi \text{ は } 0 \text{ と } 1 \text{ の間の数})$$

となる．最後のの剰余項が $0 \le \xi \le 1$ でどの範囲にあるかはすぐにはわからないが，微分などしてみると，$-4/7$ と $2/7$ の間にあることはわかる．従って，以上から

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{4}{7} \le \frac{\pi}{4} \le 1 - \frac{1}{3} + \frac{1}{5} + \frac{2}{7}$$

が得られる．残念ながら，これはどうしようもないくらい役に立たない評価である（これなら「円周率は３」の方がマシ）．

1.4.3　テイラー展開（無限項まで）

（教科書の p.154）

定理 1.4 において，公式 (1.23) 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，

$$f(x) = \lim_{n\to\infty} S_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{1.29}$$

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，それは以下のように考えれば保証される：(1.23) 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．

- 指数関数．既にテイラーの公式が

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x), \qquad R_n(x) := \frac{e^\xi}{n!}x^n \tag{1.30}$$

となることは見た（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

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{1.31}$$

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，将来，みることもあるだろう．

- 三角関数（$\sin, \cos$）も同様に，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

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad \text{また同様の考察により} \quad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{1.32}$$

が成り立つことがわかる．このテイラー級数の形も大事である[2]．

- $\pi/4$ に対しても $\arctan x$ の展開から

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \dots \tag{1.33}$$

が成り立つ事がわかる．

1.4.4 テイラー展開の効用

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

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

少し進んだ話題．

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

$$e^{i\theta} = \cos\theta + i \sin\theta, \qquad \theta \in \mathbb{R} \tag{1.34}$$

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少し進んだ数学で確かめられる）．つまり，「複素数に対する指数関数」をテイラー展開を用いて定義するのである．すると，

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell+1}}{(2\ell+1)!} \tag{1.35}$$

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．

[2] 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

2 積分の話題

2.1 積分の定義（区分求積法）

高校での積分の「定義」は「微分の逆演算」というやつだった．つまり,「微分したら関数 $f(x)$ になる」ような関数を f の原始関数と呼び，$F(x) = \int f(x)dx$ などと書いたのだった．

しかし，これでは与えられた $f(x)$ に対して $F'(x) = f(x)$ となるような $F(x)$ が見つからない限り，積分が定義できない事になる．世の中には不定積分が（初等関数で）書き下せない関数はいくらでもあるから，これは困る.（例：微分して $f = e^{-x^2}$ になるような関数 F は何でしょう？じつはこの F は簡単に書けない —— というか，積分で定義するしかないのだが，高校の範囲ではこのような関数は定義できないし，その存在もわからない.）

この節では高校までの知識はいったん忘れて,「与えられた関数 $f(x)$ に対して，その積分を定義すること」を最初に行う．これから見ていくように，かなり広いクラスの関数に対してその積分（定積分）を定義することができる．定積分を通して不定積分も定義できるので，高校までの知識とのつながりがつくことになる[3]．そのあとで更に積分のいろいろな性質を見ていくことにしよう．

$f(x)$ を適当な（例えば連続な）関数とし，簡単のために $f(x) > 0$ とする．$a < b$ を定めたときの定積分 $\int_a^b f(x)dx$ とは，高校でやった通り，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．しかし,「面積とは何か」の定義自体が実はあやふやだ．そこで，この講義では，以下のようにして面積と定積分を同時に定義していく．このような考えは,「区分求積法」として見たことがあるかもしれない[4]．

定義 2.1 (定積分) $a < b$ と．区間 $[a, b]$ で定義された（連続な）関数 $f(x)$ に対して，定積分 $\displaystyle\int_a^b f(x)dx$ を以下のように定義する（下図を参照）．

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

$$R(f; P, \vec{\zeta}) = \sum_{i=1}^{n} f(\zeta_i)(x_i - x_{i-1}) \tag{2.1}$$

  を定義する．
- さて，$|P| \to 0$（区間の幅がゼロ）を満たすような任意の P と，P に対して上のようにとった任意の $\vec{\zeta}$ を考える．$|P| \to 0$ の極限で $R(f; P, \vec{\zeta})$ の値が（$P, \vec{\zeta}$ の取り方によらず）一定の値に <u>近づくならば</u>, <u>$f(x)$ は $[a, b]$ 上で**積分可能**</u>（または <u>可積分</u>）といい，その極限値を定積分 $\displaystyle\int_a^b f(x)dx$ の値と定める．模式的に数式で書けば

$$\int_a^b f(x)dx \equiv \text{“}\lim_{|P|\to 0}\text{”}\, R(f; P, \vec{\zeta}) \tag{2.2}$$

  とするのである（上の極限はかなり複雑なので “ ” を付けた）．

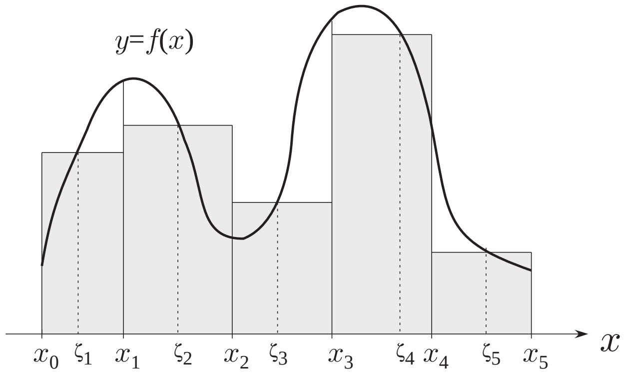
なお，$a = b$ の場合は $\displaystyle\int_a^a f(x)dx = 0$ と定義する．

また，$a > b$ の場合は $\displaystyle\int_a^b f(x)dx = -\int_b^a f(x)dx$ と定義する．($a > b$ の時の定義はもちろん，$\displaystyle\int_b^a f(x)dx$ が定義できる時のみ有効である.）このようにして定義した積分をリーマン式積分，または**リーマン積分**という．

[3]定積分より先に不定積分を考えようとすると,「微分したら $f(x)$ になるような関数 $F(x)$ は何？という問いに応える必要がある．これは一般に非常に難しい．しかしこれからやる定積分の定義なら，このような場合にも使えるのだ

[4]厳密には,「区分求積法」とは区間を等間隔に区切った場合をいうようだ．以下では等間隔でない場合も扱う

$f(x) > 0$ 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．



図を見ればわかるように，この定義は大体において，面積の近似値を作るだろうと予想される．少なくとも，上の極限が存在する場合にこの値を面積とすることに異論はないだろう．非常に大きな問題はこの極限がいつ存在するのか（面積がいつ定義できるのか），そもそもこのような極限が存在する関数（可積分な関数）は存在するのか，であるが，これは後の節で少し考察する．

この節ではまず，定積分とは，グラフの下の図形の面積を細い短冊の和で近似する（近似したい）ものである，ということをはっきりと認識してほしい[5]．

（注）繰り返しになるが，ここで学んでいる定積分の定義から出発して高校でやった「原始関数」につなげていくことはこの後で行う．従って,「微分の逆演算は積分」ということは一旦，忘れて頂きたい．この意味で，これからやることは高校での積分の導入に厳密な根拠を与える作業である．

2.1.1 定積分はいつ定義できるのか？

先に注意したように，定義 2.1 の極限値 (2.2) はいつも存在するとは限らない．有名な例（Dirichlet）だが

$$f(x) = \begin{cases} 0 & (x\text{ が有理数の時}) \\ 1 & (x\text{ が無理数の時}) \end{cases} \qquad \text{に対して} \qquad \int_0^1 f(x)dx \tag{2.3}$$

を考えると，これは定義 2.1 では定義できない（なぜ定義できないのか，各自で納得するまで考えること）．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 Lebesgue が彼の博士論文で提唱した「ルベーグ積分」である．いろいろな意味で，ルベーグ積分の方がリーマン積分より自然な積分だと僕は考えるが，その厳密な理論はそれなりに大変なので，この講義ではルベーグ積分は扱わない．

どのような関数が積分可能かを考えるのはこの講義の程度を超えているので，ここでは積分可能性の簡単な十分条件を挙げるにとどめる：

定理 2.2 (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．

理解を深める問題：

高校の時にもやったかもしれないが，良く知っている関数に対して，上積分，下積分を計算しよう．例えば，積分区間は $[-1,1]$ にして，$f(x) = x^2, x^3$ の場合など，いくつかやってみることを強く奨める．

まあ，定義に従って定積分を求めるのは大変だ（上の問題をやった人は同意するだろう）．でも，高校で習ったように（それでこれから見るように）定積分は微分の逆演算なのだ．この事実により，積分の計算は非常に簡単になるのだ．

[5]煎じ詰めれば「積分は和のお化け」である．ついでに「微分は差のお化け」である.「お化け」は別名，極限ともいう

2.2 積分の性質

ほとんど当たり前ではあるが，定積分の基本的な性質をまとめて述べておこう．まず，$a \geq b$ の場合の定積分の定義を思い出しておこう．

- $\displaystyle\int_a^a f(x) = 0$ と定める．
- $a < b$ のとき，$\displaystyle\int_a^b f(x)dx$ が定義できるならば，$\displaystyle\int_b^a f(x)dx = -\int_a^b f(x)dx$ と定める．

さて，定積分の定義から 以下の諸性質が簡単に導かれる．

定理 2.3 (区間に関する加法性)
(i) $a < c < b$ のとき，

$$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx \tag{2.4}$$

である（もちろん，３つの積分が定義できることは仮定する）．
(ii) 実は上の (2.4) は任意の実数 a, b, c について成り立つ（やはり３つの積分が定義できることは仮定する）．

これは区間を合わせた（足した）場合に対応する積分も足し算になることを主張しているので，積分の加法性と呼ばれる．これも（厳密な証明はともかく）高校の時から知ってるはずだ．

証明：
(i) $f(x) \geq 0$ の場合，定積分をグラフの下の面積だと思えば，これはほとんどアタリマエであるが，定積分の定義からもすぐに導かれる．その際，区間 $[a, b]$ の分割として点 c を分点に持つようなものを考えて，$[a, c]$ 上，および $[c, b]$ 上の積分との関連をつけるとよい．
(ii) これは簡単で，(i) の結果と積分の上端が下端より小さい場合の定積分の定義を組み合わせるとすぐに出る． □

定理 2.4 (積分の線形性)
(i) $\int_a^b f(x)dx$, $\int_a^b g(x)dx$ がともに定義できるとき，

$$\int_a^b \bigl\{f(x) \pm g(x)\bigr\}dx = \int_a^b f(x)dx \pm \int_a^b g(x)dx \tag{2.5}$$

である．
(ii) $\int_a^b f(x)dx$ が定義できるとき，任意の実数 α に対して

$$\int_a^b \bigl\{\alpha f(x)\bigr\}dx = \alpha \int_a^b f(x)dx \tag{2.6}$$

である．

いうまでもなく，上の性質は定積分で定義される関数から実数への写像

$$f \qquad \mapsto \quad \int_a^b f(x)dx \tag{2.7}$$

が線形写像であることを主張している．線形代数で注意されたかもしれないが，線形写像の一番基本的なものは普通の微分演算や積分演算なのだ．それはともかく，これは高校の時から親しんできた性質であろう．

証明：
(i), (ii) ともに非常に簡単である．定積分はリーマン和の極限として定義されたが，そのリーマン和に対して (i), (ii) に相当する線形写像の関係式が成り立っている．そのため，極限をとった後の定積分でも同じ関係式が成り立つ． □

定理 2.5 $a < b$ のとき，以下が成り立つ（もちろん，登場する積分は定義できているものとする）．
(i) （正値性）

$$[a,b] \text{ で } f(x) \geq 0 \qquad \Longrightarrow \qquad \int_a^b f(x)dx \geq 0 \tag{2.8}$$

(ii) （単調性）

$$[a,b] \text{ で } f(x) \leq g(x) \qquad \Longrightarrow \qquad \int_a^b f(x)dx \leq \int_a^b g(x)dx \tag{2.9}$$

(iii) $[a,b]$ で $f(x)$ は連続かつ非負とする．もし $f(x)$ がこの区間で恒等的にゼロでなければ，$\int_a^b f(x)dx > 0$（ゼロではなく，完全に正）である．

証明：

(i) $f(x) \geq 0$ ならば定積分の定義のリーマン和がそもそも非負である．従って，極限として定義される定積分も非負である．
(ii) $h(x) = g(x) - f(x)$ に (i) を適用すればよい．
(iii) 仮定から $a \leq c \leq b$ なる c があって，$f(c) > 0$ となっているはずである．今，$f(x)$ が連続と仮定したので，$f(x)$ の値は $x = c$ の十分近くでは正である．特に十分小さな $\delta > 0$ をとれば

$$|x - c| < \delta \text{ かつ } a \leq x \leq b \quad \text{では} \quad f(x) \geq \frac{f(c)}{2} \tag{2.10}$$

となっているはずだ．

ここで簡単のため，$c - \delta \geq a$，$c + \delta \leq b$ だったとする（そうでない場合にどのように証明を修正すべきかは以下から明らかだろう）．積分の区間に関する加法性を用いて積分区間を $c \pm \delta$ でわけると，

$$\int_a^b f(x)dx = \int_a^{c-\delta} f(x)dx + \int_{c-\delta}^{c+\delta} f(x)dx + \int_{c+\delta}^b f(x)dx \tag{2.11}$$

となるが，始めと終わりの積分は $f(x) \geq 0$ のために非負である．また，真ん中の積分は (2.10) をもちいると，

$$\int_{c-\delta}^{c+\delta} f(x)dx \geq \int_{c-\delta}^{c+\delta} \frac{f(c)}{2} dx = \frac{f(c)}{2} \times 2\delta = f(c)\,\delta > 0 \tag{2.12}$$

である．従って，(2.11) 自身も正である． □

系 2.6 $a < b$ のとき，両辺の積分が定義できているなら，

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b \bigl| f(x) \bigr| \, dx. \tag{2.13}$$

証明 簡単だ．

$$-|f(x)| \leq f(x) \leq |f(x)| \tag{2.14}$$

がいつでも成り立っているので，この不等式の 3 辺をそれぞれ a から b まで積分すれば良い．定理 2.5 の (ii) から，積分結果に対しても不等号が成り立つ：

$$-\int_a^b |f(x)|\, dx \leq \int_a^b f(x)dx \leq \int_a^b |f(x)|\, dx \tag{2.15}$$

これは (2.13) に他ならない． □

定理 2.7 (積分の平均値の定理) $a < b$ とし，区間 $[a, b]$ 上では $f(x)$ が連続，かつ $g(x) \geq 0$ と仮定する．以下の両辺の積分が定義できるとき，区間 $[a, b]$ 内の一点 ξ が存在して，

$$\int_a^b f(x)g(x)dx = f(\xi) \int_a^b g(x)dx \tag{2.16}$$

特に $g(x) \equiv 1$ とおくと，

$$\int_a^b f(x)dx = f(\xi)\,(b-a) \tag{2.17}$$

となるような ξ の存在が証明される．

証明 簡単だ．閉区間 $[a, b]$ 上の連続関数は最大値と最小値をもつから，それらを M, m と書こう．すると，$g(x) \geq 0$ なので，区間 $[a, b]$ では $mg(x) \leq f(x)g(x) \leq Mg(x)$ がなりたつ．この不等式のそれぞれの辺を積分すると，積分の単調性から，

$$m \int_a^b g(x)dx \leq \int_a^b mg(x)dx \leq \int_a^b f(x)g(x)\,dx \leq \int_a^b Mg(x)dx = M \int_a^b g(x)dx \tag{2.18}$$

が得られる．以下，これから (2.16) を示す．

まず，$\int_a^b f(x)dx = 0$ ならば，(2.16) の両辺が共にゼロとなり，(2.16) はアタリマエに正しい．

次に，$\int_a^b g(x)dx \neq 0$ ならば，$g(x) \geq 0$ から $\int_a^b g(x)dx > 0$ である．よって，上の不等式 (2.18) の両辺を $\int_a^b g(x)dx$ で割って

$$m \leq \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \leq M \tag{2.19}$$

が結論できる．m, M は区間 $[a, b]$ における $f(x)$ の最大値と最小値であったので，$f(x)$ が連続なら，x が a から b まで動くとき，$f(x)$ は m と M の間すべての値をかならず一度はとる（中間値の定理）．従って，特に，

$$\frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} = f(\xi) \tag{2.20}$$

となる $\xi \in [a, b]$ が存在する．この式の分母をはらえば (2.16) になる． □

（注）上の定理は以下のように書く方が「平均値」の定理という感じがするのだが，どうだろうか？

$$f(\xi) = \frac{1}{b-a} \int_a^b f(x)dx, \qquad f(\xi) = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \tag{2.21}$$

左側は $f(x)$ を区間 $[a, b]$ で普通に平均したつもりだし，右側のは $f(x)$ を $g(x)$ という重みで加重平均した感じになっている．

最後に，積分の性質の中では最も重要とも言えるものを証明しよう．

定理 2.8 (微分積分学の基本定理) I 上では $f(x)$ が連続とする．区間 I 内の一点を a として

$$F(y) := \int_a^y f(x)dx \tag{2.22}$$

を定義する．このとき，F は I 内の各点 y で微分可能で，

$$\frac{d}{dy} F(y) = f(y). \tag{2.23}$$

（注）$F(y)$ がちゃんと定義できていることは定理 2.2 で保証されている．

証明 ここでは積分の平均値の定理を使った証明を与えておく．$F(y)$ の微分を定義どおり計算しようとすると

$$\frac{d}{dy}F(y) = \lim_{h\to 0}\frac{F(y+h)-F(y)}{h} = \lim_{h\to 0}\frac{1}{h}\int_y^{y+h} f(x)dx \tag{2.24}$$

の極限が問題になる．(以下，簡単のため，$h \to +0$ の極限を考えるが，$h \to -0$ も全く同様にできる．) 積分形の平均値の定理，特に系?? によると，右辺の積分は $[y, y+h]$ 内の適当な ξ を用いて $hf(\xi)$ と書ける．つまり，問題の極限は

$$\frac{d}{dy}F(y) = \lim_{h\to 0}\frac{F(y+h)-F(y)}{h} = \lim_{h\to 0} f(\xi) \tag{2.25}$$

というものだ（ここで ξ は y と $y+h$ の間の適当な数）．$h \to 0$ の極限では，ξ は y に収束する．更にこのとき，f が連続なので，$f(\xi)$ は $f(y)$ に収束する．従って，問題の極限は $f(y)$ に等しく，定理は証明された． □

高校でも既にやったように，「微分したら $f(x)$ になる関数」を $f(x)$ の 原始関数 と呼ぶ．上の定理で定義した $F(x)$ は原始関数の一つである．すると当然，$f(x)$ の原始関数はどのくらいあるのか，が問題になるが，これには以下の命題が答えてくれる．

系 2.9 $a < b$ とし，区間 $[a, b]$ 上では $f(x)$ が連続とする．このとき，f の原始関数 $F(x)$ は，付加定数を除いて一意に定まる．すなわち，

(i) $F_1(x)$ が $f(x)$ の原始関数である場合，任意の定数 C を用いて

$$F_2(x) := F_1(x) + C \tag{2.26}$$

を定義すると，$F_2(x)$ も $f(x)$ の原始関数である．

(ii) $F_1(x)$ と $F_2(x)$ が $f(x)$ の原始関数である場合，x に依存しない定数 C がとれて

$$F_2(x) - F_1(x) = C \qquad (a \le x \le b) \tag{2.27}$$

と書ける．

証明 (i) のほうは，$F_1' = f$ ならば $F_2' = f$ でもあることから，明らか．
(ii) の方は，F_1, F_2 が f の原始関数ならば $\frac{d}{dx}\{F_2(x) - F_1(x)\} = f(x) - f(x) = 0$ であるべきだから，この両辺を積分すればすぐに出る． □

なお，この付加定数の自由度は (2.22) での c の選び方が全く任意であったことに対応していることに注意しよう．

これで漸く，高校で習った積分のお話を閉じることができる．特に，置換積分と部分積分は高校で習った通りに成り立つ．

命題 2.10 (置換積分) （置換積分）有限閉区間 $[\alpha, \beta]$ の C^1-級関数 $\varphi(t)$ があり，$\varphi(\alpha) = a, \varphi(\beta) = b$ $(a \neq b)$ であって，$\alpha < t < \beta$ では $\varphi(t)$ は a, b の間にあるとする．このとき，区間 $[a, b]$ または $[b, a]$ 上の連続関数 $f(x)$ に対して

$$\int_a^b f(x)dx = \int_\alpha^\beta f\bigl(\varphi(t)\bigr)\,\varphi'(t)\,dt \tag{2.28}$$

命題 2.11 (部分積分) （置換積分）有限閉区間 $[a, b]$ の C^1-級関数 $f(x), g(x)$ に対して

$$\int_a^b f(x)\,g'(x)\,dx = \Bigl[f(x)g(x)\Bigr]_a^b - \int_a^b f'(x)\,g(x)\,dx \tag{2.29}$$

2.3 広義積分

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

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

もちろん，この２つが両方起こっているもの（積分区間も無限だし，被積分関数も有界でない；例えば $\int_{-\infty}^{\infty} e^{x}dx$）もありうる．これらの問題に共通しているのは，積分区間や被積分関数に無限大の芽が含まれており，定義 2.1 をそのままの形では適用できないということだ.（適用した場合，答えが「無限大」などになってしまうが，これは我々の欲しい積分の値としてはかえって不自然.）

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

（ことばについて）この節の内容で定義される積分を 広義積分（improper integral）と呼ぶ．日本語の方はそのまま「積分の定義を拡張したもの」のつもりであろう．英語の方は正しい定義 2.1 には含まれていない，というつもりだろうか.

2.3.1 有界区間上の積分だが，被積分関数が有界でない場合の広義積分

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

定義 2.12 $a<b$ とする．
(0) $f(x)$ が区間 $[a,b]$ において有界なら，今までのリーマン積分の定義により $\int_a^b f(x)dx$ を定義する．
(i) $f(x)$ が半開区間 $(a,b]$ で定義されていて

$$\lim_{c\to a+0}\int_c^b f(x)dx \tag{2.30}$$

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

$$\lim_{d\to b-0}\int_a^d f(x)dx \tag{2.31}$$

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

$$\lim_{\substack{c\to a+0\\ d\to b-0}}\int_c^d f(x)dx \tag{2.32}$$

が存在するとき， $f(x)$ は $[a,b]$ で 広義積分可能 といい，その極限を広義積分 $\int_a^b f(x)dx$ の値と定める．ただしここで c,d の極限は 互いに独立に a,b へ近づけるすべての近づけ方についてとる．
(iv) 最後に，上の (i), (ii), (iii) のそれぞれの極限が存在しない場合，その広義積分は存在しないという．

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

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

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

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

定義 2.12 では区間 $[a, b]$ の端に変態な（例えば f が有界でなくなる）点がある場合を考えた．もし $[a, b]$ の内部の点 c で f が変態である場合は，

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \tag{2.33}$$

の公式を使う．つまり，上の右辺の２つの積分のそれぞれが定義 2.12 によって広義積分として定義できるとき，上の式を使って $\int_a^b f(x)dx$ を定義する．具体的に書くと，

$$\int_a^b f(x)dx = \lim_{e \to c-0} \int_a^e f(x)dx + \lim_{d \to c+0} \int_d^b f(x)dx \tag{2.34}$$

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

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

$$\int_{-1}^1 \frac{1}{x}dx = \int_{-1}^0 \frac{1}{x}dx + \int_0^1 \frac{1}{x}dx \tag{2.35}$$

として定義したいが，右辺の積分は２つとも定義できない（定義 2.12 にしたがって極限を考えても $\pm\infty$ に発散してしまう）．従って $\int_{-1}^1 \frac{1}{x}dx$ 自身も定義できない.

（補足）この例でもし，右辺の極限を同じ速さでとると，つまり

$$\lim_{\epsilon \to +0} \left[\int_{-1}^{-\epsilon} \frac{1}{x}dx + \int_\epsilon^1 \frac{1}{x}dx \right] \tag{2.36}$$

を考えると，括弧の中は被積分関数が奇関数だからゼロになり，従って極限値もゼロである（というふうに極限値は存在してしまう）．正しい定義（極限は別々にとる）との違いをよく認識されたい．なお，この「補足」のようにそろえて極限をとったものには,「コーシーの主値（積分)」の名前がついている．

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

$$\int_a^b f(x)dx = \int_a^{c_1} f(x)dx + \int_{c_1}^{c_2} f(x)dx + \int_{c_2}^{c_3} f(x)dx + \int_{c_3}^b f(x)dx \tag{2.37}$$

の公式を使うつもりになる．そして右辺のそれぞれの積分が定義 2.12 にしたがって定義できるかどうかを考える訳だ.

2.3.2 無限区間上の積分だが，被積分関数が有界な場合の広義積分

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

定義 2.13 $f(x)$ は有界な関数とする．
(i) 半無限区間 $[a,\infty)$ 上の有界な関数 $f(x)$ に対して，極限

$$\lim_{M\to\infty}\int_a^M f(x)dx \tag{2.38}$$

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

$$\lim_{L\to-\infty}\int_L^b f(x)dx \tag{2.39}$$

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

$$\lim_{\substack{L\to-\infty\\ M\to\infty}}\int_L^M f(x)dx \tag{2.40}$$

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

最後の (iii) については定義 2.12(iii) と同じ注意が適用される．つまり，$L\to-\infty$ と $M\to\infty$ は別々に極限をとるのだ．なお，将来，$L=-M$ としてとった極限を考える場合もある（「フーリエ変換」などで出てくるはず）．

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

2.3.3 （半）無限区間上の積分で，被積分関数も有界でない場合の広義積分

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

$$\int_0^\infty\frac{\sin x}{\sqrt{x}}dx=\int_0^1\frac{\sin x}{\sqrt{x}}dx+\int_1^\infty\frac{\sin x}{\sqrt{x}}dx=\lim_{\epsilon\to+0}\int_\epsilon^1\frac{\sin x}{\sqrt{x}}dx+\lim_{L\to\infty}\int_1^L\frac{\sin x}{\sqrt{x}}dx \tag{2.41}$$

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

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

2.3.4 広義積分が計算できないときの収束の判定条件）

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

例えば，

$$I_1:=\int_0^\infty\frac{\sin x}{x}dx=\lim_{L\to\infty}\int_1^L\frac{\sin x}{x}dx \tag{2.42}$$

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

$$I_2 := \int_1^\infty \frac{|\sin x|}{x} dx, \qquad I_3 := \int_1^\infty \frac{|\sin x|}{x^2} dx \tag{2.43}$$

なども挙げておこう．こたえを先に言ってしまうと，I_2 は発散するが，I_1, I_3 は収束する（広義積分が定義できる）．この節では，これらの判定条件（多くの場合は十分条件にすぎない）を考える．ただし，この講義で扱えるのは，上の I_2, I_3 のように，非積分関数が一定符号の場合のみである．

以下，非積分関数が一定符号 —— いつも非負，またはいつも非正 —— の場合に話を限る．(正でも負でも一緒だから，以下では非負の場合のみ考える．) このときは簡単な（必要）十分条件がある．

命題 2.14 この命題では $f(x) \geq 0$ とする．
(1) $f(x)$ が $x \geq a$ で有界の場合，広義積分 $\displaystyle\int_a^\infty f(x)dx$ の収束性は，b の関数として定義した $S(b) := \displaystyle\int_a^b f(x)dx$ の（$b \to \infty$ での）有界性と同等である．
(2) $f(x)$ が $x = a$ 以外では有界の場合，広義積分 $\displaystyle\int_a^b f(x)dx$ の収束性は，c の関数として定義した $S(c) := \displaystyle\int_c^b f(x)dx$ の（$c \to a+0$ での）有界性と同等である．

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

証明：

（参考までに証明を載せたが，無視して良い．）

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

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

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

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

この定理から直ちに，始めの I_3 の収束性を結論できる．実際，

$$0 \leq \frac{|\sin x|}{x^2} \leq \frac{1}{x^2} \tag{2.44}$$

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

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

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

3 偏微分

これから残りの時間で「偏微分」を簡単にやります．これは将来，どこかで見る可能性が高いから，頑張ってやりましょう．主な話題は（順不同）

- 偏微分の定義
- 連鎖律（合成関数の微分）
- テイラー展開
- 高階の偏微分；偏微分は順序によらない
- 極大極小問題

です．

3.0 多変数の関数

偏微分が出てくるのは,「多変数の関数」に関してである．今まで皆さんが知っていたのは主に「1 変数」の関数だった．これは「変数」が一つだけで，それを決めたら関数の値が変わるようなの．例えば

$$x^2 \qquad \sin x \qquad \frac{x^2+\cos(x^3)}{\sqrt{1+x^2}} \tag{3.1}$$

など．でも世の中には２つ以上の変数に依存する関数はいくらでもある．n 個の変数に依存するのを「n 変数の関数」という．例えば（変数を x, y, z などと書いて）

$$f(x,y)=\sqrt{x^2+y^2} \qquad g(x,y)=x^2\cos y \qquad h(x,y,z)=\frac{x+z^3}{y^2+1} \tag{3.2}$$

など．

こういうと仰々しいけど，そんなに難しい話ではない．上の $f(x,y)$ なら「直角を挟む２辺の長さが x, y の時の斜辺の長さ」を表しているわけで，このような例は数学でも，日常生活でも直感的には扱ったことがあるはず．

ということで難しく考える必要は全くないのだが，一つだけ注意：多変数関数のグラフはちょっと大変．例えば２変数の関数 $f(x,y)$ なら，変数は x, y の２つあるから，こいつは平面（の一部）を動く．平面上の各点 (x,y) に対して関数の値 $f(x,y)$ が決まるので，このグラフは２次元では描けず，第３の軸（z 軸）をとって $z=f(x,y)$ を xyz の３次元空間で図にしなければならない.（f の性質が良ければ，結果はこの３次元空間での曲面になる.）以下では主に２変数の関数を扱うが，いつもこのような立体的なグラフのイメージを持つことにしよう．

(多変数関数の極限)

多変数関数の極限について：変数が２個以上ある場合，極限の取り方には注意が必要だ．つまり，(x,y) が (a,b) に近づく，といっても，その近づき方にはいろいろな方向がありうる．つまり，(1) $y=b$ として，x-軸に平行に近づくのか，(2) $x=a$ として y-軸に平行に近づくのか，(3) 斜めに近づくのか，(4) 螺旋を描くように近づくのか... といくらでも可能性がある．

そこで，上のようなどのような近づき方をしても行き先が同じ数である場合，かつそのときに限り，(行き先の数を A と書いて)

$$\lim_{(x,y)\to(a,b)} f(x,y)=A \tag{3.3}$$

と書く．

3.1 偏微分

さて，いよいよ偏微分を考えよう．以下の話のほとんどは一般の n 変数のときにも成り立つが，式がいたずらに複雑になるので，主に２変数の関数を考える．

定義 3.1 (偏微分係数) ２変数関数 $f(x,y)$ の点 (a,b) における 第１変数に関する偏微分係数 とは極限

$$\lim_{h\to 0}\frac{f(a+h,b)-f(a,b)}{h}=\lim_{x\to a}\frac{f(x,b)-f(a,b)}{x-a} \tag{3.4}$$

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

$$\lim_{h\to 0}\frac{f(a,b+h)-f(a,b)}{h}=\lim_{y\to b}\frac{f(a,y)-f(a,b)}{y-b} \tag{3.5}$$

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

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

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

以下の定義はよく使うので，ここで与えておく．

定義 3.2 (C^1-級) 多変数関数 $f(x_1,x_2,\ldots,x_n)$ がその定義域（の一部）D で

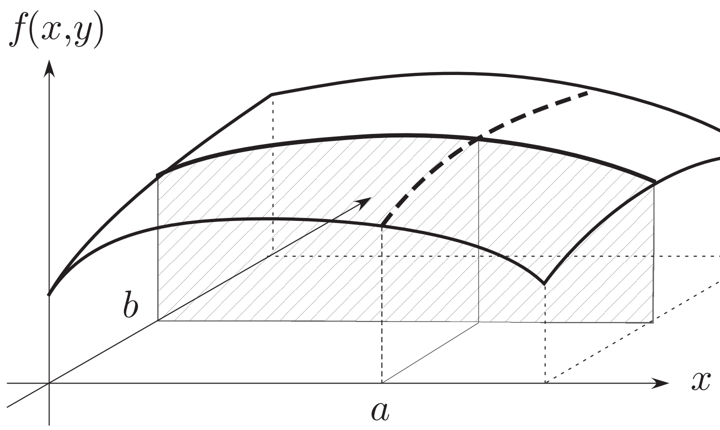
- f は各変数 $x_1,x_2,\ldots,x_n$ のそれぞれについて偏微分可能で
- かつ，その n-この偏導関数が $\boldsymbol{x}=(x_1,x_2,\ldots,x_n)$ の連続関数である

であるとき，f は D で C^1-級 であるという．

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

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

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



（記号についての注意）

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

問 3.2.1. 次の関数をそれぞれの独立変数で偏微分せよ．

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

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

3.1.1 偏導関数がゼロ，の関数は？

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

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

$$f(x,y)=g(y)\qquad g\text{ は任意の関数}\tag{3.6}$$

と書ける．これは一般には定数関数ではない！

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

3.1.2 方向微分[6]

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

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

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

そこで，このような変化率をみるために，以下の定義を行う．

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

定義 3.3 (方向微分) 2 変数関数 $f(x,y)$ と 2 次元の単位ベクトル（長さ 1 のベクトル）$\boldsymbol{v}=(v_x,v_y)$ が与えられたとせよ．極限

$$f_{\boldsymbol{v}}(a,b) := \lim_{h\to 0} \frac{f(a+hv_x, b+hv_y)-f(a,b)}{h} \tag{3.7}$$

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

$$f_{\boldsymbol{v}}(\boldsymbol{a}) := \lim_{h\to 0} \frac{f(\boldsymbol{a}+h\boldsymbol{v})-f(\boldsymbol{a})}{h} \tag{3.8}$$

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．

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

その前に少し例を挙げておこう．以下の関数 f,g

$$f(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad f(x,y)=\frac{2xy}{x^2+y^2} \tag{3.9}$$

$$g(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad g(x,y)=\frac{xy}{\sqrt{x^2+y^2}} \tag{3.10}$$

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

$$f_x(0,0)=f_y(0,0)=0, \qquad (x,y)\neq(0,0) \quad \text{では} \quad f_x(x,y)=\frac{2y(y^2-x^2)}{(x^2+y^2)^2}, \quad f_y(x,y)=\frac{2x(x^2-y^2)}{(x^2+y^2)^2} \tag{3.11}$$

$$g_x(0,0)=g_y(0,0)=0, \qquad (x,y)\neq(0,0) \quad \text{では} \quad g_x(x,y)=\frac{y^3}{(x^2+y^2)^{3/2}}, \qquad g_y(x,y)=\frac{x^3}{(x^2+y^2)^{3/2}} \tag{3.12}$$

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

3.2 合成関数の微分（連鎖率，chain rule）

ここでは偏微分での最初の山場,「連鎖率」（合成関数の微分）を学ぶ．この題材は簡単に見えて，案外たいへんなことがあるから，注意する事．特に，この後でやる「高階の導関数」を計算する時にひっかかる人が多いはずだ．なお，この節の山場は後の 3.2.3 節である[7]．

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

$$h(x)=f(g(x)) \tag{3.13}$$

で定義される関数 h を f と g の 合成関数 といい，$f\circ g$ などと書いたのだった.「そんな言葉は知らない」という人も $\sin(x^3)$ は $f(x)=\sin x$ と $g(x)=x^3$ の合成関数だといえば，高校の時から知っているものと納得できるはずだ．

このとき，関数 $h(x)$ の導関数については，高校以来，

$$h'(x)=f'\bigl(g(x)\bigr)\,g'(x) \tag{3.14}$$

[7]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

が成り立つことは知っている．(このように書くと良くわからない，という人も $\sin(x^3)$ を x で微分する事はできるはずだから，受験数学でやってるんだよ.）この節の主題は，これの多変数関数版を考える事である．

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

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

このうち，A は高校以来知っていることだ（この後でも改めて証明する)．

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

$$\frac{\partial h(x,y)}{\partial x} = f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial x}, \qquad \frac{\partial h(x,y)}{\partial y} = f'\bigl(z(x,y)\bigr)\,\frac{\partial z(x,y)}{\partial y} \tag{3.15}$$

となる．

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

3.2.1 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．

定理 3.4（１変数の合成関数の微分） $g(x)$ が区間 I 内の点 $x = a$ にて x について微分可能，かつ $f(z)$ が点 $b = f(a)$ にて z について微分可能のとき，合成関数 $h(x) = f\bigl(g(x)\bigr)$ は点 $x = a$ で微分可能であり，

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

がなりたつ．

（ちょっとマニアックな注）１変数の関数に関するこの定理では，「$g(x)$ が $x = a$ で微分可能，かつ $f(z)$ が $z = f(a)$ で微分可能」であれば十分で，導関数の連続性などは必要ない．後出の多変数の場合の定理 3.6 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．(が，講義では触れなかった.）

定理 3.4 の少しだけええかげんな証明. この定理はほとんど当たり前だ．$h(x)$ の $x = a$ での微分を定義するニュートン商を

$$\frac{h(a+\epsilon) - h(a)}{\epsilon} = \frac{f\bigl(g(a+\epsilon)\bigr) - f\bigl(g(a)\bigr)}{\epsilon} = \frac{f\bigl(g(a+\epsilon)\bigr) - f\bigl(g(a)\bigr)}{g(a+\epsilon) - g(a)} \times \frac{g(a+\epsilon) - g(a)}{\epsilon} \tag{3.16}$$

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

（補足）上の「証明」でごまかしたのは，「$\epsilon \neq 0$ であっても $g(a+\epsilon) - g(a) = 0$ かもしれない」という可能性を見て見ぬふりをしたことだ — この可能性の例としては，$g(x) \equiv 1$（恒等的に 1）を考えよ．もし $g(a+\epsilon) - g(a) = 0$

ならば (3.16) 右辺の書き換え（分母がゼロ！）に意味がつけられなくなり，右辺の積の極限を別々に考えることができなくなる．

でも，これは大した問題ではない．実際，$g(a+\epsilon)-g(a)=0$ ならば $h(g(a+\epsilon))-h(g(a))=0$ でもあるはずだから，もともとのニュートン商の値もゼロ，よって困ることは何もないはずだ．実際，ここのところはちょっと書き方を工夫すれば厳密に議論できる．上の「証明」は不完全だが，まずは「このような感じだな」と大体の筋道を理解することが一番大切である．

3.2.2 合成関数の微分（１変数の場合に帰着, Case B）

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

定理 3.5 $g(x,y)$ が R 内の一点 (a,b) にて x について偏微分可能，かつ $f(z)$ が $c=g(a,b)$ で微分可能とする．このとき，$h(x,y)=f\bigl(g(x,y)\bigr)$ は (a,b) にて x について偏微分可能で，

$$h_x(a,b)=f'\bigl(g(a,b)\bigr)\,g_x(a,b), \qquad \text{つまり}\ \ z=g(x,y),\ w=f(z)\ \text{とおくと}\quad \frac{\partial w}{\partial x}=\frac{dw}{dz}\frac{\partial z}{\partial x} \tag{3.17}$$

がなりたつ．

証明 x についての偏微分のみを問題にしているから，変数 y は単なる定数と思っても同じだ．だから，定理 3.4 が使える． □

3.2.3 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[8]］

$$\begin{aligned}\lim_{\epsilon\to 0}\frac{h(t+\epsilon)-h(t)}{\epsilon}&=\lim_{\epsilon\to 0}\frac{f\bigl(x(t+\epsilon),y(t+\epsilon)\bigr)-f\bigl(x(t),y(t)\bigr)}{\epsilon}=\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_0,y_0)}{\epsilon}\\&=\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}+\lim_{\epsilon\to 0}\frac{f(x_1,y_0)-f(x_0,y_0)}{\epsilon}\end{aligned} \tag{3.18}$$

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

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

$$\lim_{\epsilon\to 0}\frac{f(x_2,y_1)-f(x_2,y_0)}{\epsilon}\qquad \text{with}\qquad y_0=y(t),\quad y_1=y(t+\epsilon) \tag{3.19}$$

であれば，$x=x_2$ は定数で y だけが ϵ に依存するから，極限は合成関数の微分（case B）により $\frac{\partial f}{\partial y}(x_2,y)\,y'(t)$ になる．さらに x_2 も x_0 に近づくと 思えば，これは多分，$\frac{\partial f}{\partial y}(x_0,y_0)\,y'(t)$ になるだろう（**ここでゴマカシ終わり**）．よってこのゴマカシによると，(3.18) から

$$h'(t)=\frac{\partial f}{\partial x}(x_0,y_0)\,x'(t)+\frac{\partial f}{\partial y}(x_0,y_0)\,y'(t) \tag{3.20}$$

が得られると 予想 される．

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

答えを言ってしまうと，以上の結論 (3.20) はかなり一般に成り立つ．ただし，上でも明記したように，(3.18) 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：

$$g(0,0)=0, \qquad (x,y)\neq(0,0) \text{ では } \quad g(x,y)=\frac{xy}{\sqrt{x^2+y^2}}. \tag{3.21}$$

この関数の偏導関数は (3.12) で計算した．特に，$g_x(0,0)=g_y(0,0)=0$ である．さて，地道に計算するとすべての t で

$$h(t)=g(x(t),y(t))=\frac{t}{\sqrt{2}}, \qquad h'(t)=\frac{1}{\sqrt{2}} \tag{3.22}$$

であって，特に $h'(0)\neq 0$ だ．ところが (3.20) を闇雲に使うと（$x=y=0$ での偏微分の値を使って計算するから）$h'(0)=0$ が得られてしまう．つまり，この $g(x,y)$ に対しては (3.20) 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 1 つ，定理の形で述べる．なお，あまり細かいことを書くと肝心のところが見えなくなりそうだから，関数の定義域と値域は，合成関数が定義できるようになっていると適当に仮定する．

定理 3.6 2 変数関数 $f(x,y)$ が $\underline{C^1\text{-級}}$ で，$x(t),y(t)$ が t について微分可能なら，$h(t)=f\bigl(x(t),y(t)\bigr)$ は t で微分可能である．更にその導関数について

$$h'(t)=f_x\,x'(t)+f_y\,y'(t), \qquad \text{つまり } z=f(x,y) \text{ とおくと } \quad \frac{dz}{dt}=\frac{\partial z}{\partial x}\frac{dx}{dt}+\frac{\partial z}{\partial y}\frac{dy}{dt} \tag{3.23}$$

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

$$\frac{d}{dt}f\bigl(x_1(t),x_2(t),\ldots,x_n(t)\bigr)=\sum_{j=1}^{n}\frac{\partial f}{\partial x_j}\frac{dx_j}{dt}. \tag{3.24}$$

言うまでもなく，(3.21) 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　（証明は参考のために載せた．この講義では無視しても全く問題ない．）

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

$$f(x_1,y_1)-f(x_1,y_0)=f_y(x_1,y_3)\times(y_1-y_0) \tag{3.25}$$

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

$$\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}=\lim_{\epsilon\to 0}\Bigl[f_y(x_1,y_3)\times\frac{y_1-y_0}{\epsilon}\Bigr] \tag{3.26}$$

となる．ところで，$\underline{f \text{ が } C^1 \text{ 級と仮定しているから，} f_y(x,y) \text{ は } x,y \text{ の連続関数}}$である（ここがキーでした）．従って $\epsilon\to 0$ では $x_1\to x_0$ かつ $y_3\to y_0$ であることも使うと，

$$\lim_{\epsilon\to 0}f_y(x_1,y_3)=f_y(x_0,y_0) \tag{3.27}$$

である．また，(3.26) の後ろの方は y_1,y_0 の定義から

$$\lim_{\epsilon\to 0}\frac{y_1-y_0}{\epsilon}=\lim_{\epsilon\to 0}\frac{y(t+\epsilon)-y(t)}{\epsilon}=y'(t) \tag{3.28}$$

である．以上をまとめると

$$\lim_{\epsilon\to 0}\frac{f(x_1,y_1)-f(x_1,y_0)}{\epsilon}=f_y(x_0,y_0)\times y'(t) \tag{3.29}$$

が厳密に証明されたので，(3.18) に戻ると定理が証明できた．n-変数の場合も（式が汚くなるだけで）同様．□

D の場合についてもだめ押しで述べておこう．

定理 3.7 (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 $z = h(u,v) = f\bigl(x(u,v), y(u,v)\bigr)$ も C^1-級で，

$$\frac{\partial z}{\partial u} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial u}, \qquad \text{かつ} \qquad \frac{\partial z}{\partial v} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial v}. \tag{3.30}$$

証明 u で偏微分する場合には v は動かさない（定数）から先の定理 3.6 からすぐに証明される． □

問題 3.8 （連鎖率）２変数の関数 $f(x,y) = \dfrac{xy}{\sqrt{x^2+y^2}}$ を考える．また，変数 x,y と変数 r,θ は平面の曲座標の関係，つまり $x = r\cos\theta, y = r\sin\theta$ を満たしているものとする．このとき，合成関数 $h(r,\theta) = f\bigl(x(r,\theta), y(r,\theta)\bigr)$ の，変数 r,θ に関する偏導関数を計算せよ．ただしその場合,（１）合成関数の微分法（連鎖率）を用いて計算する,（２）h を r,θ の関数として具体的に書き下してから偏微分する，の２通りで行い，結果を比べること．なお，場合によっては微分できない点があるかもしれないが，今はそれは無視して良い．つまり，微分できる点で偏導関数を求めればよい．

問題 3.9 (x,y) と (u,v) が 1 次変換（$\alpha,\beta,\gamma,\delta$ は $\alpha\delta \neq \beta\gamma$ を満たす定数）

$$u = \alpha x + \beta y,\ \ v = \gamma x + \delta y \tag{3.31}$$

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

問題 3.10 以下の方程式を満たす C^1-級の関数 $f(x,y)$ は一般にどんな形か，求めよ（a,b はゼロでない定数である）．

$$1)\ \ \frac{\partial f}{\partial y} \equiv 0 \qquad 2)\ \ \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \equiv 0 \qquad 3)\ \ a\frac{\partial f}{\partial x} + b\frac{\partial f}{\partial y} \equiv 0 \tag{3.32}$$

1) は既にやった（3.1.1 節）．問題は 2) と 3) だが，問題 3.9 をヒントにせよ．つまり，新しい変数 (u,v) をうまく見つけて，$\dfrac{\partial f}{\partial u} \equiv 0$ が成り立つようにしてみると良い.（もちろん，このようなやり方でうまく行く保証はないが，ある程度の経験を積めば，この形の方程式ならこのような一時変換で行ける，ことが予想できるものである.）

3.3 高階の偏導関数

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

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

$$\frac{\partial^2 f}{\partial y\partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \qquad \text{同様に} \qquad \frac{\partial^2 f}{\partial x\partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \tag{3.33}$$

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[9]．

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

問題 3.11 以下の関数 f, g, h について，２階の偏導関数（４通り）をすべて計算せよ．実は以下の定理で $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ だというのだが，この定理には頼らず，実際に計算して確かめる事．

$$f(x,y) = x^2 + y^2, \qquad g(x,y) = x^2 y^3, \qquad h(x,y) = \cos(x^2 y) \tag{3.34}$$

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

さて，これまでの定義によるだけでは，$\frac{\partial^2 f}{\partial x \partial y}$ と $\frac{\partial^2 f}{\partial y \partial x}$ とは全く別ものであり（微分の順序が逆），何の関係もないように見える．実際，問題 3.11 の微分を実際にやった人には，$\frac{\partial g}{\partial x}$ と $\frac{\partial g}{\partial y}$ が全然別ものだから，$\frac{\partial^2 g}{\partial x \partial y}$ と $\frac{\partial^2 g}{\partial y \partial x}$ に関係がつく事の方がかえって奇跡に見えるかもしれない．しかし，(我々が扱うような「普通の」関数では）この２つは等しい．

定理 3.12 ２変数関数 $f(x,y)$ が C^2-級の場合，

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} \tag{3.35}$$

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

定理 3.12 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[10]（D は領域，A は D 内の一点．また $f_x = \frac{\partial f}{\partial x}, f_{yx} = \frac{\partial^2 f}{\partial x \partial y}$ などと略記する）．

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

定理 3.12 の証明 (a,b) を f の定義域中の一点として固定し，ここでの微分を考える．

$$\Delta(h,k) := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a,b) \tag{3.36}$$

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

$$\begin{aligned}\lim_{h\to 0}\Bigl(\lim_{k\to 0}\frac{\Delta(h,k)}{hk}\Bigr) &= \lim_{h\to 0}\frac{1}{h}\Bigl[\lim_{k\to 0}\Bigl(\frac{f(a+h,b+k)-f(a+h,b)}{k} - \frac{f(a,b+k)-f(a,b)}{k}\Bigr)\Bigr] \\ &= \lim_{h\to 0}\frac{1}{h}\Bigl[\frac{\partial f}{\partial y}(a+h,b) - \frac{\partial f}{\partial y}(a,b)\Bigr] = \frac{\partial}{\partial x}\Bigl(\frac{\partial f}{\partial y}\Bigr) = \frac{\partial^2 f}{\partial x \partial y}\end{aligned} \tag{3.37}$$

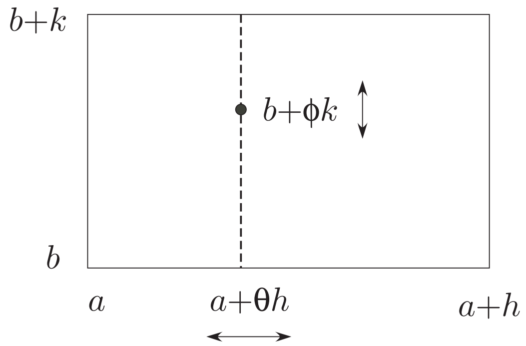
であるし，同様に考えると

$$\lim_{k\to 0}\Bigl(\lim_{h\to 0}\frac{\Delta(h,k)}{hk}\Bigr) = \frac{\partial^2 f}{\partial y \partial x} \tag{3.38}$$

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）．

[10] 小平本の 6.2 節の d) を参考にした



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

$$\varphi(a+h) - \varphi(a) = \varphi'(a+\theta h)h \tag{3.39}$$

つまり

$$\Delta(h,k) = \varphi(a+h) - \varphi(a) = \Big(f_x(a+\theta h, b+k) - f_x(a+\theta h, b)\Big)h \tag{3.40}$$

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[11]またもや 1 変数の場合の平均値の定理から

$$\psi(b+k) - \psi(b) = \psi'(b+\phi k)k \tag{3.41}$$

つまり

$$f_x(a+\theta h, b+k) - f_x(a+\theta h, b) = f_{xy}(a+\theta h, b+\phi k)\,k \tag{3.42}$$

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

以上 (3.39) と (3.42) から

$$\frac{\Delta(h,k)}{hk} = \frac{f_{xy}(a+\theta h, b+\phi k)\,hk}{hk} = f_{xy}(a+\theta h, b+\phi k) \tag{3.43}$$

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

$$\lim_{(h,k)\to(0,0)} \frac{\Delta(h,k)}{hk} = f_{xy}(a,b) \tag{3.44}$$

これで既に $h \to 0$ と $k \to 0$ の極限が順序によらない事は示せたから (3.37) と (3.38) を思い出すと証明は完成した．(余分ではあるが，上の議論を順序を変えて行うと (3.44) の極限が $f_{yx}(a,b)$ に等しい事もいえて，証明はより明確になる．) □

3.3.1 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．

---

[11] (注) ここでは少しひっかかるかもしれない．(3.39) では a, b, h, k を固定していたので，それに応じて θ が決まった．ところがここではその θ を固定した上で $\psi(y)$ を考え，b や $b+k$ の方を動かしているように見え，何となく気持ちが悪い．だけど，ここの我々の結論 (3.42) は (3.39) がなりたっていた $b+k, b$ についてのものであるので，(3.39) と (3.42) 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

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

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

$$\begin{pmatrix} u \\ v \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}, \qquad \begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \tag{3.45}$$

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

$$\frac{\partial f}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial f}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial f}{\partial v} = \frac{1}{\sqrt{2}}\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right] \tag{3.46}$$

となる.（余談だが，上の関係は微分演算子の部分だけを取り出して

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x}\frac{\partial}{\partial u} + \frac{\partial v}{\partial x}\frac{\partial}{\partial v} = \frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial u} - \frac{\partial}{\partial v}\right] \tag{3.47}$$

とも書ける．このような書き方は後々，便利だ.）これを y で偏微分して

$$\begin{aligned}\frac{\partial^2 f}{\partial y \partial x} &= \frac{1}{\sqrt{2}}\left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right]\left\{\frac{1}{\sqrt{2}}\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right]\right\} = \frac{1}{2}\left[\frac{\partial}{\partial u} + \frac{\partial}{\partial v}\right]\left[\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right] \\ &= \frac{1}{2}\left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v \partial u} - \frac{\partial^2 f}{\partial v^2}\right] = \frac{1}{2}\left[\frac{\partial^2 f}{\partial u^2} - \frac{\partial^2 f}{\partial v^2}\right]\end{aligned} \tag{3.48}$$

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

3.3.2 高階偏導関数と連鎖律

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

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

$$\frac{\partial f}{\partial x} = \frac{\partial r}{\partial x}\frac{\partial f}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial f}{\partial \theta} = \frac{x}{\sqrt{x^2+y^2}}\frac{\partial f}{\partial r} - \frac{y}{x^2+y^2}\frac{\partial f}{\partial \theta} = \frac{x}{r}\frac{\partial f}{\partial r} - \frac{y}{r^2}\frac{\partial f}{\partial \theta} \tag{3.49}$$

となるので，微分演算子としては

$$\frac{\partial}{\partial x} = \frac{x}{r}\frac{\partial}{\partial r} - \frac{y}{r^2}\frac{\partial}{\partial \theta} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta} \tag{3.50}$$

という作用をもっている．さて，もう一回やるときには

$$\frac{\partial^2 f}{\partial x^2} = \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta\frac{\partial f}{\partial r} - \frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) \tag{3.51}$$

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

$$= \cos\theta\frac{\partial}{\partial r}\left(\cos\theta\frac{\partial f}{\partial r}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\cos\theta\frac{\partial f}{\partial r}\right) + \cos\theta\frac{\partial}{\partial r}\left(-\frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(-\frac{\sin\theta}{r}\frac{\partial f}{\partial \theta}\right) \tag{3.52}$$

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

$$\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\Bigl(\cos\theta\frac{\partial f}{\partial r}\Bigr)=\frac{\sin\theta}{r}\frac{\partial\cos\theta}{\partial\theta}\times\frac{\partial f}{\partial r}+\frac{\sin\theta}{r}\cos\theta\frac{\partial}{\partial r}\Bigl(\frac{\partial f}{\partial\theta}\Bigr)=-\frac{\sin^2\theta}{r}\frac{\partial f}{\partial r}+\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{3.53}$$

などとなる訳だ．このように計算していくと，最終的な答えは

$$\frac{\partial^2 f}{\partial x^2}=\cos^2\theta\frac{\partial^2 f}{\partial r^2}+\frac{\sin^2\theta}{r}\frac{\partial f}{\partial r}+\frac{\sin 2\theta}{r^2}\frac{\partial f}{\partial\theta}+\frac{\sin^2\theta}{r^2}\frac{\partial^2 f}{\partial\theta^2}-\frac{\sin 2\theta}{r}\frac{\partial^2 f}{\partial r\partial\theta} \tag{3.54}$$

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

問題 3.13 2 変数の関数 $f(x,y)$ と座標変換 $x=r\cos\theta, y=r\sin\theta$ を考える．このとき，

$$\Delta f(x,y):=\frac{\partial^2 f}{\partial x^2}+\frac{\partial^2 f}{\partial y^2} \tag{3.55}$$

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

3.3.3 （補足）偏導関数がゼロという関数は？ふたたび

以前に「1 階偏微分がゼロ」の関数はどんなものか考えたが，今度は 2 階導関数がゼロのものを考えてみよう．例えば，$\frac{\partial^2 f}{\partial y\partial x}=f_{xy}\equiv 0$ というのを考えてみる．これは

$$\frac{\partial}{\partial y}\bigl(f_x\bigr)=0 \tag{3.56}$$

と見ると，「f_x は y には依存しません」ということなので，

$$f_x(x,y)=g(x) \tag{3.57}$$

と書けるはずだ（g は任意の関数，3.1.1 節を思い出そう）．この両辺を x で積分すると，左辺は $f(x,y)$ になり，右辺は g の原始関数になるが，x で積分したときの積分定数は y の任意の関数になれる．なぜなら，積分定数は積分している変数に依存していなければなんでもよく，いまは x で積分しているので，y に依存するのは勝手である．(ここのところがわかりにくい人は $g(x)$ の原始関数を $G(x)$ とすると，(3.57) から

$$\frac{\partial}{\partial x}\bigl\{f(x,y)-G(x)\bigr\}=0 \tag{3.58}$$

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

$$f(x,y)=G(x)+h(y) \tag{3.59}$$

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

なお，このような考察は物理で「波動方程式」を考える時にでてくる．

3.4 平均値の定理，テイラー展開

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

定理 3.14 (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,

$$f(a+h,b+k) = f(a,b) + \frac{\partial f}{\partial x}(a+\theta h, b+\theta k)h + \frac{\partial f}{\partial y}(a+\theta h, b+\theta k)k \tag{3.60}$$

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

$$f(\boldsymbol{a}+\boldsymbol{h}) = f(\boldsymbol{a}) + \sum_{j=1} \frac{\partial f}{\partial x_j}(\boldsymbol{a}+\theta\boldsymbol{h})\, h_j \tag{3.61}$$

が成り立つ（$0<\theta<1$）.

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

$$f(a+h,b+k) - f(a,b) = g(1) - g(0) = g'(\theta) \tag{3.62}$$

である．ところが，g' については,「連鎖律」定理 3.6 を $x(t) = a+th, y(t) = b+yk$ として用いると

$$g'(t) = \frac{\partial f}{\partial x}x'(t) + \frac{\partial f}{\partial y}y'(t) = \frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k \tag{3.63}$$

であるから，定理 3.14 を得る．n 変数の場合も同様である． □

平均値の定理が成立するには，関数が C^1 級である必要はない．全微分可能性を仮定すると，以下の定理になる．この辺りは数学としては興味のあるところだが，余裕のない人はあまりこだわる必要はない．上の定理だけ理解すれば（一年生の間は）十分だ.

定理 3.15 (平均値の定理) 2 変数関数 $f(x,y)$ が全微分可能なら,

$$f(a+h,b+k) = f(a,b) + \frac{\partial f}{\partial x}(a+\theta h, b+\theta k)h + \frac{\partial f}{\partial y}(a+\theta h, b+\theta k)k \tag{3.64}$$

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

定理 3.14 が定理 3.6 から出るのと同じようにして，定理 3.15 は定理**??**から証明される. □

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

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

$$g(t) = f(a+th, b+tk) - f(a,b) \tag{3.65}$$

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

$$f(a+h,b+k) = g(1) = \sum_{k=0}^{n-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(n)}(\theta)}{n!} \qquad (0<\theta<1) \tag{3.66}$$

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

$$\frac{d}{dt} = \frac{dx(t)}{dt}\frac{\partial}{\partial x} + \frac{dy(t)}{dt}\frac{\partial}{\partial y} = h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} \tag{3.67}$$

であるから，例えば，f が C^2-級ならば

$$g^{(2)}(\theta) = \frac{d}{dt}\frac{dg}{dt} = \Bigl(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\Bigr)\Bigl(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\Bigr)$$
$$= h^2\frac{\partial^2 f}{\partial x^2} + hk\frac{\partial^2 f}{\partial x\partial y} + kh\frac{\partial^2 f}{\partial y\partial x} + k^2\frac{\partial^2 f}{\partial y^2} \tag{3.68}$$

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，

$$\bigl|g^{(2)}(\theta)\bigr| \le 2M(h^2+k^2) = O(\|\boldsymbol{a}-\boldsymbol{c}\|^2) = o(\|\boldsymbol{a}-\boldsymbol{c}\|) \tag{3.69}$$

が成り立つ（記号を簡単にするため，$\boldsymbol{a} = (a,b), \boldsymbol{c} = (a+h, b+k)$ とおいた）．つまり，(3.66) の $n=2$ を考えると，

$$f(a+h, b+k) = f(a,b) + f_x(a,b)h + f_y(a,b)k + o(\|\boldsymbol{a}-\boldsymbol{c}\|) \tag{3.70}$$

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，

$$f(a+h,b+k) = f(a,b)+f_x(a,b)h+f_y(a,b)k+\frac{1}{2}\bigl[f_{xx}(a,b)h^2+2f_{xy}(a,b)hk+f_{yy}(a,b)k^2\bigr]+o(\|\boldsymbol{a}-\boldsymbol{c}\|^2) \tag{3.71}$$

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

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

定理 3.16 ２変数の関数 $f(x,y)$ が C^r-級であるとき，適当な $0<\theta<1$ に対して

$$f(a+h,b+k) = f(a,b)+\sum_{m=1}^{r-1}\frac{1}{m!}\sum_{\ell=0}^{m}\binom{m}{\ell}\frac{\partial^m f}{\partial x^\ell\,\partial y^{m-\ell}}(a,b)\,h^\ell\,k^{m-\ell}+\frac{1}{r!}\sum_{\ell=0}^{r}\binom{r}{\ell}\frac{\partial^r f}{\partial x^\ell\,\partial y^{r-\ell}}(a+\theta h,b+\theta k)\,h^\ell\,k^{r-\ell} \tag{3.72}$$

が成り立つ．

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

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

（問題）次の関数を，あたえられた点 (a,b) の周りで，２次までテイラー展開せよ．つまり，(3.71) に相当する式を,（具体的に偏微分を計算して）書き下せ．

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

3.5 極大・極小問題

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

3.5.1 問題の定義

定義 3.17 n-成分ベクトルの空間において，上の記号のもとで，

$$B_r(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathbb{R}^n \,\big|\, \|\boldsymbol{x}-\boldsymbol{a}\| < r\} \tag{3.73}$$

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

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

定義 3.18 n-変数の関数 $f(\boldsymbol{x})$ が $\boldsymbol{x} = \boldsymbol{a}$ で**極大**であるとは，適当な $r > 0$ に対して $\boldsymbol{a}$ の r-開近傍 $B_r(\boldsymbol{a})$ があって，その中では $f(\boldsymbol{a})$ の値が最大であることをいう（r は我々が勝手に設定してよい）．つまり，

$$\text{ある正の } r \text{ が存在して，} \qquad 0 < \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \text{では} \quad f(\boldsymbol{x}) < f(\boldsymbol{a}) \tag{3.74}$$

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

$$\text{ある正の } r \text{ が存在して，} \qquad 0 < \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \text{では} \quad f(\boldsymbol{x}) > f(\boldsymbol{a}) \tag{3.75}$$

であることをいう．

- この代わりに等号も含めたもの，つまり (3.74) と (3.75) の代わりに

$$\exists r > 0, \qquad \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{3.76}$$

$$\exists r > 0, \qquad \|\boldsymbol{x}-\boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \ge f(\boldsymbol{a}) \tag{3.77}$$

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

$$f \text{ の定義域に入っているすべての } \boldsymbol{x} \text{ に対して} \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{3.78}$$

  であることをいう（上の極大の定義のように $\boldsymbol{x}$ の範囲を我々が勝手に設定してはいけない）．最小についても同様である．なお，(3.78) で等号を入れるか入れないかはまた，悩ましい定義の問題だが，ここでは一応，等号も許す事にする．

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

3.5.2 1変数の場合の復習

さて，1変数の場合の極大，極小問題は以下のようになっていた（高校でやったはず）．

定理 3.19 $x=a$ の近傍で定義された1変数の関数 $f(x)$ について，以下が成り立つ．
(i) $f(x)$ が $x=a$ で微分可能，かつ $x=a$ で $f(x)$ が極大または極小の場合，$f'(a)=0$ である．逆は必ずしもなりたたない．
(ii) $f(x)$ が $x=a$ で2階微分可能で $f'(a)=0$ の場合には，以下が成り立つ：

a. $f''(a)>0$ の場合，$f(x)$ は $x=a$ で極小である．
b. $f''(a)<0$ の場合，$f(x)$ は $x=a$ で極大である．
c. $f''(a)=0$ の場合，$f(x)$ の $x=a$ での極大極小については何も言えない（極大の場合，極小の場合，どちらでもない場合もある）．

（上の定理の (ii)-c は「定理」の中に入れるほどのことではないが，わかりやすさを考えて入れておいた.）
念のために定理のそれぞれの場合に相当する例を挙げておこう（すべて $a=0$ の例）．

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

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

$$f(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2}(x-a)^2+o(|x-a|^2) \tag{3.79}$$

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

$$f(x)=f(a)+f'(a)(x-a)+o(x-a) \tag{3.80}$$

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

$$f(x)=f(a)+\frac{f''(a)}{2}(x-a)^2+o(|x-a|^2) \tag{3.81}$$

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

$$f(x)=f(a)+\frac{f^{(3)}(a)}{6}(x-a)^3+\frac{f^{(4)}(a)}{24}(x-a)^4+\frac{f^{(5)}(a)}{120}(x-a)^5+o(|x-a|^5) \tag{3.82}$$

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

3.5.3 2変数の極大極小問題

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

定義 3.20 2変数の関数 $f(x,y)$ の，点 (a,b) における**ヘッセ行列**とは，以下の形の行列

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y)=(a,b) \text{ での値} \tag{3.83}$$

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

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

すると，

定理 3.21 $(x,y)=(a,b)$ 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），

$$f(x,y) = f + f_x(x-a) + f_y(y-b) + \frac{1}{2}\Bigl[f_{xx}(x-a)^2 + 2f_{xy}(x-a)(y-b) + f_{yy}(y-b)^2\Bigr] + o(\|\boldsymbol{x}-\boldsymbol{a}\|^2) \tag{3.84}$$

となっていたことをまず，思い出そう．

(i) 1階微分の少なくとも1つがゼロでない場合．

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

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

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

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

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

$$g(x,y) = A(x-a)^2 + 2B(x-a)(y-b) + C(y-b)^2 \tag{3.85}$$

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

$A \neq 0$ の場合をまず考えると，

$$g(x,y) = A\left[\left\{(x-a) + \frac{B}{A}(y-b)\right\}^2 + \frac{CA-B^2}{A^2}\,(y-b)^2\right] \tag{3.86}$$

である．よって場合分けすると

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

となって，定理の a, b, c の場合がでてくる．

$C \neq 0$ の場合は x, y の役割を取り替えれば同様．

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

(iii) １階微分も２階微分もすべてゼロの場合：

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

ともかく，２変数の関数の場合に定理 3.21 を理解するのは，このように地道に考えれば可能である．なお，同様の議論を「行列の対角化」の話を用いて，この後で定式化しなおす． □

以上をまとめると，２変数の関数の極値問題の解き方は以下のようになる．

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

$$f_x(a,b) = f_y(a,b) = 0 \tag{3.87}$$

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

（２）実際に極値になっているかを調べる（講義ノートの定義 3.20 と定理 3.21）．上を満たす (a,b) の一つ一つについて，ヘッセ行列

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y) = (a,b) \text{ での値} \tag{3.88}$$

を定義すると，

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

3.5.4　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．この場合は線形代数で習うはずの「行列の対角化と２次形式の標準形」を用いるのが良いのだが，この講義では省略する．

3.6 条件付き極値問題：ラグランジュの未定乗数法

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

以下の問いを考えたい．

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

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

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

このような問題を**条件付き極値（最大最小）問題**という．

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

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

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

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

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

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

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

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

$$F(x,y,\lambda)=f(x,y)-\lambda g(x,y) \tag{3.89}$$

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

(i) $g(x,y)=0$ の特異点，

(ii) 未知変数を x,y,λ とする以下の連立方程式の解．

$$0=\frac{\partial F}{\partial x}=\frac{\partial f}{\partial x}(x,y)-\lambda\frac{\partial g}{\partial x}(x,y),\qquad 0=\frac{\partial F}{\partial y}=\frac{\partial f}{\partial y}(x,y)-\lambda\frac{\partial g}{\partial y}(x,y) \tag{3.90}$$

$$0=\frac{\partial F}{\partial\lambda}=g(x,y) \tag{3.91}$$

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

考案者の名前をとって λ を Lagrange の未定乗数（Lagrange multiplier）という．なお，この方法では**極値をとる**
(x,y) **の候補が見つかるだけ**であって，それらが実際の極値を与えるか否かを決める一般論は存在しない．(より正確には，そのような一般論がない訳ではないが，実用的なものはほとんどない．）ただし，極値点の候補が見つかれば，その点の周りでのテイラー展開などを用いて，実際に極値になっているかどうかの判定は可能な事が多いから，これは実用上は大した問題ではない（少なくとも計算機の助けを借りれば何とかなる）．また，方程式 (3.90) と (3.91)（やその多変数の場合の該当物）を解くのは大変だと強調している本が多いが，これも計算機の助けを借りればそんなに大した問題ではない（事も多い）．というわけで，未定乗数法はやはり偉大なのである．

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

$$0=4x^3+2\lambda x, \qquad 0=4y^3+2\lambda y, \qquad 0=x^2+y^2-1 \tag{3.92}$$

の３つである．これを解くと，

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ \pm1 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm1 \\ 0 \\ -2 \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \pm\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \begin{bmatrix} \pm\frac{1}{\sqrt{2}} \\ \mp\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \ \text{（ベクトルの中では複合同順）} \tag{3.93}$$

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

（未定乗数法がうまく行く理由１）

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

この $h(x)$ は x のみの関数だから極値の条件は

$$0=h'(x)=f_x+f_y\,\varphi'(x) \tag{3.94}$$

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

$$0=\frac{d}{dx}g(x,\varphi(x))=g_x+g_y\,\varphi'(x) \tag{3.95}$$

この２つから，

$$\varphi'(x)=-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \tag{3.96}$$

が導かれるが，これは見方を変えれば

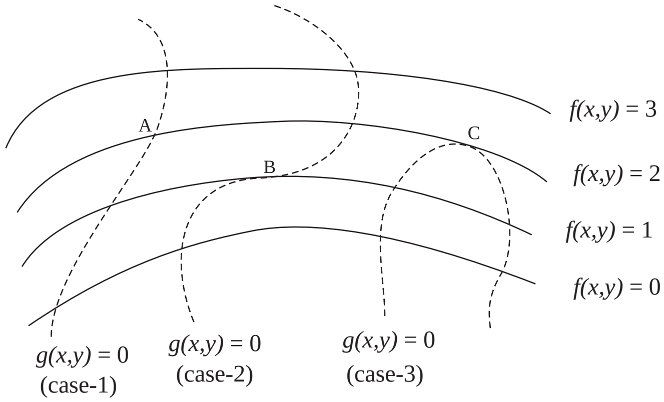
$$\frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{3.97}$$

ということであり，この値を λ と書けば，これは (3.90) 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 **??**の (??) で与えられている．つまり

$$f(x,y)=c \text{ の接線の傾きは } -\frac{f_x}{f_y}, \qquad g(x,y)=0 \text{ の接線の傾きは } -\frac{g_x}{g_y} \tag{3.98}$$

なのだ．従って，両者が接する条件は

$$-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \qquad \text{つまり} \qquad \frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{3.99}$$

であるが，これは (3.90) 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

$$F(\boldsymbol{x},\lambda_1,\lambda_2,\ldots,\lambda_m)=f(\boldsymbol{x})-\left\{\,\lambda_1 g_1(\boldsymbol{x})+\lambda_2 g_2(\boldsymbol{x})+\cdots+\lambda_m g_m(\boldsymbol{x})\,\right\} \tag{3.100}$$

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．

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

$$0=\frac{\partial F}{\partial \lambda_k}=g_k(\boldsymbol{x}), \qquad (k=1,2,\ldots,m) \tag{3.101}$$

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．