5 微分（続き）

5.4 テイラーの定理とテイラー展開

これから暫く，微分の重要な応用のもう一つ.「テイラー展開」を扱う．これは案外，皆さん苦労するようだから，少し時間をかけることにした．この節に対応する内容は教科書にはないので，プリントも詳しく作っている．

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

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

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

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

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

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

などとなる．

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

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

- まず，(5.4.2) などの式は，それ自身が数値計算にも適している —— $e^x, \sin x$ などの値を，右辺の級数（和）で計算できるのだ．もちろん，無限級数の値そのものを数値的に求める事はできないが，たくさんの項の和をとる事で，いくらでも精度良く計算できる[1].
- (5.4.1) にはもう少し理論的な意味もある．つまり，$|x-a|$ が小さい場合に $f(x)$ を $f(a)$ で近似すると，誤差がどうなるかを表していると解釈できる．この誤差の評価は，もっと進んだ結果を得るのに不可欠である．

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

5.4.1 テイラーの公式（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 5.4.1 (通常のテイラーの公式) $f(x)$ がある開区間 I で n 回微分可能と仮定し，この区間内に $a \in I$ をとろう．このとき，勝手な $x \in I$ に対して，a と x の間の一点 ξ が存在して以下が成り立つ：

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

なお，(5.4.5) の 2 つの項に名前をつけて

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

[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 (5.4.2) そのものではなく，これを更に効率よくしたものを用いる．しかし，計算の原理は（大体）同じである

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

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

- $a = 0$ とした場合の展開を特にマクローリン（Maclaurin）の公式（展開）ともいう．
- 実はマクローリンの公式とテイラーの公式は非常に近い親戚関係にあり，片方だけわかれば十分だ．理由は以下の通り：$y = x - a$ という変数変換によって，座標 x で見た時の点 $x = a$ は座標 y で見た時の $y = 0$ に移る．従って，座標 y でのマクローリンの公式は座標 x での $x = a$ の周りのテイラーの公式に対応している．
- テイラーの公式でも，平均値の定理でも，ξ は a と x（またはb）の両方に依存しうることを再度強調しておく．同じ理由で，剰余項 $R_n(x)$ は x, a で決まるけども，$R_n(x)$ の ξ そのものが x, a に依存する事をお忘れなく．
- 細かいことであるが，定理 5.4.1 では $f^{(n)}(x)$ の存在は仮定するが，連続性は仮定しなくても良い．この点で，剰余項が積分形の定理 5.4.7（後出）より，こちらの方が少しだけ適用範囲はひろい（そのぶん，誤差評価は大抵，劣る —「ある ξ が存在して」とか言われても，どんな ξ かわからなければ細かい評価はできない）．

定理 5.4.1 の証明[2]

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

とおく．$F(x)$ が (5.4.6) の $R_n(x)$ の表式で書けることを示せばよい．

そのために，コーシーの平均値の定理（定理 5.2.3）を F, G に適用する事を考えよう．$F(x)$ は $f(x)$ から $(x-a)^k$ の和を引いているだけなので，また $G(x)$ は多項式なので，共に n 階は微分できる．微分を具体的に計算すると

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

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

となっている．この事実を用いて，以下のように進む．

（1）定理 5.2.3 そのもので

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

を満たす ξ_1 の存在（ξ_1 は a と x の間にある）が言える．

（2）上の右辺の量は $F'(a) = G'(a) = 0$ を用いて強引に書き直すと，定理 5.2.3 が使える．その結果，

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

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

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

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

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

（4）以上をまとめると，

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

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

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

となっているので，分母を払うと定理が得られる． □

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

5.4.2 テイラー展開（無限項まで）

定理 5.4.1 において，公式 (5.4.6) がすべての $n \geq 1$ 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，

$$f(x) = \lim_{n\to\infty} S_n(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k \tag{5.4.16}$$

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，それは以下のように考えれば保証される：(5.4.6) 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．

5.4.3 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

$$e^x = \sum_{k=0}^{n-1} \frac{x^k}{k!} + R_n(x), \qquad R_n(x) := \frac{e^{\xi}}{n!}x^n \tag{5.4.17}$$

  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（レポート問題）．従って，すべての実数 x に対して

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \tag{5.4.18}$$

  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

$$\sin x = S_n(x) + R_n(x), \qquad S_n(x),\ R_n(x) \text{ の形はレポートでね} \tag{5.4.19}$$

  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（レポート問題）．従って，すべての実数 x に対して

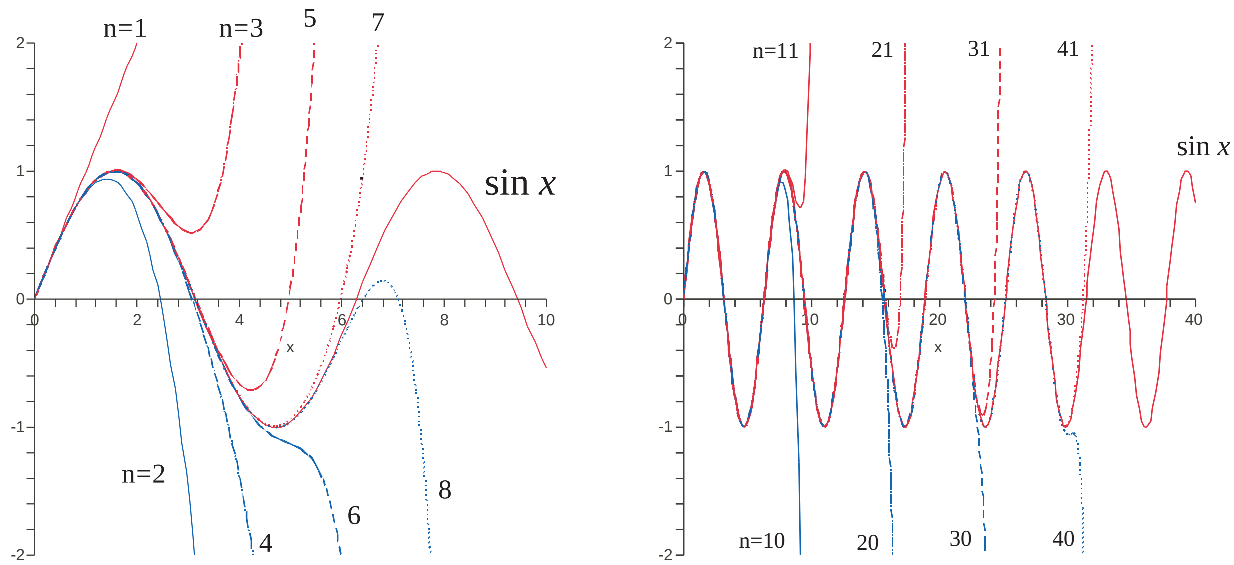
$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \qquad \text{また同様の考察により} \quad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \tag{5.4.20}$$

  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[3]．

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．

[3] 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



第１回レポート問題：テイラー展開に慣れるための問題です．テイラー展開を求めよ，というもののいくつかは既にプリントに書いてあるけど，自分で微分を実行して，テイラーの公式を使う事．

問１： 以下の関数の $x=0$ でのテイラー展開を求めよ（a は正の定数）．ただし (a), (b), (c), (d) は一般の n に対する主要項 $S_n(x)$ と剰余項 $R_n(x)$ を，(e), (f) は最初のゼロでない４項を求める事．

$$(a)\quad f(x)=\frac{1}{1-2x},\qquad (b)\quad g(x)=e^{ax},\qquad (c)\quad h(x)=\sin x,\qquad (d)\quad p(x)=\log(1+ax)$$

$$(e)\quad q(x)=(1+x)^{1/2}\qquad (f)\quad r(x)=\frac{(1+x)^{1/2}}{(1-x)^{1/2}}$$

問２： e^x と $\sin x$ のテイラーの公式における剰余項がなぜゼロになるのか，理由を述べよ（証明せよ）．

番外問題：（前学期と同じ．感想や改善の要望なども，あれば書いて下さい.）

レポート提出方法： 上の問に解答し，

10 月 12 日（木）17:00 （時刻は 24 時間制）までに，原の部屋（六本松３号館 3-312）の前の箱に

入れてください．整理の都合上，用紙は A4 を使ってください．また，２枚以上にわたる場合は何らかの方法で綴じてくだされ．

前回の補足（言葉）：テイラーの公式やテイラー展開で $a=0$ としたやつを「$f(x)$ の $x=0$ の周りのテイラーの公式」とか「テイラー展開」という，同様に，$a=3$ なら「$x=3$ の周りの」という．人によっては「周りの」の代わりに「において」と言う場合もある．つまり,「$x=2$ におけるテイラーの公式」など．

5.4.4 テイラーの公式の意味（関数の近似）

そもそも，テイラーの公式は

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

いう精神の下に生まれたものである．つまり，後述する条件の下では，(5.4.6) での $S_n(x)$ が $f(x)$ を良く近似し，$R_n(x)$ の方は小さな誤差項とみなせるのだ．

また，第１回のレポートをやった人ならわかるだろうが，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 5.4.2 (n 次より高く近似) $x=0$ の近くで定義された関数 $f(x), g(x)$ があり，

$$x \to 0 \quad \text{のときに} \qquad \lim_{x\to 0} \frac{f(x)-g(x)}{x^n} = 0 \qquad (n\text{ は正の整数}) \tag{5.4.21}$$

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

上の式では，$f(x)-g(x)$ はゼロに行くのだが，その行き方（ゼロへの収束の速さ）が，x^n よりも速い，と言っているのである．こ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[4]

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

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

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

$$\exists K>0 \quad \exists\delta>0 \qquad \Big(0<|x-a|<\delta \quad \Longrightarrow \quad \Big|\frac{f(x)}{h(x)}\Big| < K\Big) \tag{5.4.22}$$

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

（注）

- アとイは大文字と小文字だけの区別なので，特に手書きの際には注意が必要だ．
- 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，$f(x)=\Omega\big(g(x)\big)$ と書いたりすることはある．

[4] 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

この書き方によると，(5.4.21) は

$$f(x) - g(x) = o(x^n) \qquad \text{（n 次より高く近似．ここの o は小文字）} \tag{5.4.23}$$

と書ける．

この用語法に従うと，以下の命題が成り立つ．まあ，これは定義そのものであるが...

命題 5.4.4 (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 階のテイラーの公式

$$f(x) = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k, \quad R_n(x) := \frac{f^{(n)}(\theta x)}{n!} x^n \quad (0 < \theta < 1) \tag{5.4.24}$$

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

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + o(x^{n-1}) \tag{5.4.25}$$

となるための必要充分条件は以下の通り：

$$\lim_{x \to 0} \frac{R_n(x)}{x^{n-1}} = 0 \tag{5.4.26}$$

前の命題の (5.4.26) の十分条件として，以下がある．

命題 5.4.5 (多項式近似の十分条件)
1) 0 を内部に含むある区間で $f^{(n)}$ が有界，つまり $\delta > 0$ と $M > 0$ があって，

$$|x| < \delta \text{ ならば } \left| f^{(n)}(x) \right| < M \tag{5.4.27}$$

となっているとする．このとき，

$$f(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(0)}{k!} x^k + O(x^n) \tag{5.4.28}$$

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

これらはわざわざ命題とするほどのことではないかもしれないが，実用上大事だから載せた．特に，一年生で出てくる関数は C^∞-級（何回でも微分できる）のものが多く，これらに対しては上の十分条件が自動的に満たされており，命題 5.4.4 の結論も成り立つのである．

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 5.4.6 (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

$$g_1(x) = \sum_{j=0}^{n} a_j x^j, \qquad g_2(x) = \sum_{j=0}^{n} b_j x^j \tag{5.4.29}$$

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

$$\frac{|g_1(x) - g_2(x)|}{|x|^n} \le \frac{|g_1(x) - f(x)|}{|x|^n} + \frac{|f(x) - g_2(x)|}{|x|^n} \tag{5.4.30}$$

の両辺で $|x| \downarrow 0$ とすると，

$$\lim_{x\to 0} \frac{|g_1(x) - g_2(x)|}{|x|^k} = 0 \qquad 0 \le k \le n \tag{5.4.31}$$

がなりたつ．

$$g_1(x) - g_2(x) = \sum_{j=0}^{n} (a_j - b_j) x^j \tag{5.4.32}$$

である事に注目して $k = 0, 1, 2, \ldots$ に対して (5.4.31) を順次考えると，$a_k - b_k = 0$ しかあり得ないことがわかる． □

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．先週のレポート問題なら，$1/(1-2x)$ は等比級数で書ける事は高校からしってるから，これが答えになるしかないことがすぐにわかる．

5.4.5 テイラー展開の効用

テイラーの公式とテイラー級数の効用については既に述べたが，重要なの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．

少し進んだ話題． 少し先走るが，２番目の効用の例として（多分，どこかで見ただろう）オイラー（Euler）の公式

$$e^{i\theta} = \cos\theta + i \sin\theta, \qquad \theta \in \mathbb{R} \tag{5.4.33}$$

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

$$e^{i\theta} = \sum_{k=0}^{\infty} \frac{(ix)^k}{k!} = \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell}}{(2\ell)!} + i \sum_{\ell=0}^{\infty} (-1)^\ell \frac{x^{2\ell+1}}{(2\ell+1)!} \tag{5.4.34}$$

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した．

5.4.6 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.（大体，微分よりは積分の方が評価しやすいのである —— これは皆さんが４年生くらいになるとわかってくるだろう）．ただ，これは積分を使っているから（そして，我々は積分の厳密な理論をまだやっていないから）現時点ではこの定理の完全な証明を与えるわけにはいかない．

定理 5.4.7 (剰余項が積分形のテイラー（Taylor）の公式) $f(x)$ がある開区間 I で C^n-級であると仮定する．この区間 I 内に $a \in I$ をとろう．このとき，勝手な $x \in I$ について，以下が成り立つ：

$$f(x) = S_n(x) + R_n(x), \qquad S_n(x) := \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k, \quad R_n(x) := \int_a^x \frac{f^{(n)}(y)}{(n-1)!}(x-y)^{n-1}\,dy \tag{5.4.35}$$

（高校のノリでの証明；ただし積分の基礎付けさえすれば，この証明は厳密に正しい）数学的帰納法で証明する．つまり $f(x)$ は C^N-級と仮定し，(5.4.35) をすべての $n \le N$ について証明することを目指す．それで n についての帰納法を用いる．

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

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

$$f(x) = f(a) + \int_a^x f'(y)dy \tag{5.4.36}$$

の第 2 項を，以下のように部分積分するとよい．

$$\begin{aligned}\int_a^x f'(y)dy &= \int_a^x \left\{-\frac{d}{dy}(x-y)\right\} f'(y)dy = \Big[-(x-y)f'(y)\Big]_a^x + \int_a^x (x-y)f''(y)dy \\ &= (x-a)f'(a) + \int_a^x (x-y)f''(y)dy \end{aligned}\tag{5.4.37}$$

II. n まで証明できたとして，$n+1$ をやってみよう（もちろん，$n \le N-1$ と仮定しておく）．n までできたと仮定したので，(5.4.35) が成り立っているが，最後の項を以下のように考えて部分積分する（分母の $(n-1)!$ は後で）：

$$\begin{aligned}\int_a^x f^{(n)}(y)(x-y)^{n-1}dy &= \int_a^x f^{(n)}(y)\left\{-\frac{1}{n}\frac{d}{dy}(x-y)^n\right\}dy \\ &= -\frac{1}{n}\Big[f^{(n)}(y)\,(x-y)^n\Big]_a^x + \frac{1}{n}\int_a^x f^{(n+1)}(y)\,(x-y)^n dy \\ &= \frac{1}{n}f^{(n)}(a)\,(x-a)^n + \frac{1}{n}\int_a^x f^{(n+1)}(y)\,(x-y)^n dy. \end{aligned}\tag{5.4.38}$$

これを (5.4.35) の最後の項に用いると（もちろん，分母の $(n-1)!$ を忘れない），(5.4.35) の $n+1$ のものが証明されてメデタシメデタシ． □

6 積分

積分については高校でも習ってはいるが，その基礎を突き詰めていくといろいろと困ったことがでてくる．特に「積分は微分の逆演算」として定義すると,「ある関数 f の積分を求めよ」という問題や「この関数の積分は定義できるか？」という問題でハタと困ってしまう.（微分して f になるような関数がわからない場合，高校までの知識ではお手上げだ.）この節では高校までの知識はいったん忘れて,「積分とは何か」「積分をどのように定義すべきか」から話を始める．その後で高校で習ったこととの関連をつけ，更に積分のいろいろな性質を見ていくことにしよう．

6.1 積分（定積分）の定義

ということで，まずやるべきは「与えられた関数 $f(x)$ に対して，その積分を定義すること」である．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[5]．

$f(x)$ 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」自体が定義を要する問題である．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．

定義 6.1.1 (定積分) $a < b$ と．区間 $[a, b]$ で定義された関数 $f(x)$ 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：$|P| = \max\limits_{1\le i\le n}(x_i - x_{i-1})$.
- 各小区間 $[x_{i-1}, x_i]$ に勝手に点 ζ_i をとる（$i = 1, 2, \ldots, n$）．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 $P, \vec{\zeta}$ に対して，<u>リーマン和</u>

$$R(f; P, \vec{\zeta}) = \sum_{i=1}^{n} f(\zeta_i)\,(x_i - x_{i-1}) \tag{6.1.1}$$

  を計算する．
- さて，$|P| \to 0$ 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

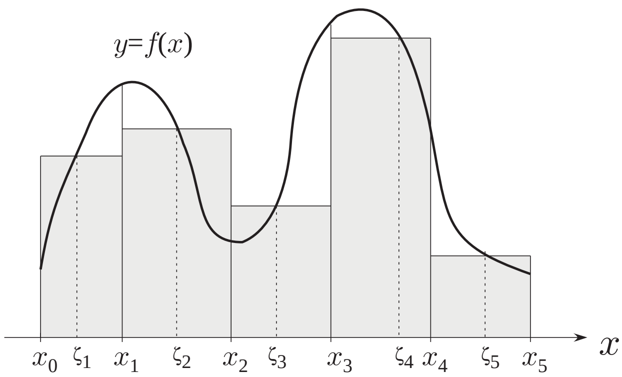
$$\int_a^b f(x)dx \equiv \text{“}\lim_{|P|\to 0}\text{”}\, R(f; P, \vec{\zeta}) \tag{6.1.2}$$

  とするのである（上の極限はかなり複雑なので“ ”を付けた）．

最後に，$a = b$ 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．

$f(x) > 0$ 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．

[5]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



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，これは次の節で詳しく考察する．

ここではまず，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[6]．

（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．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．

6.2 定積分はいつ定義できるのか？

先に注意したように，定義 6.1.1 の極限値 (6.1.2) はいつも存在するとは限らない．有名な例（Dirichlet）だが

$$f(x) = \begin{cases} 0 & (x \text{ が有理数の時}) \\ 1 & (x \text{ が無理数の時}) \end{cases} \qquad \text{に対して} \qquad \int_0^1 f(x)dx \tag{6.2.1}$$

を考えると，これは定義 6.1.1 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（３年までお預けです）．

これから積分の厳密な構築に入る．ちょっと理論的でうるさいところではあるが，大事なところだから，大筋だけでも理解するように心がけてほしい．その際にキーになるのは

- 定積分は定義できなくても,「上積分」「下積分」はいつでも定義できること（Darboux の定理，以下の定理 6.2.1）
- 定積分が定義できる必要十分条件は上積分と下積分の値が等しいこと（定理 6.2.2）
- 定積分が定義できる十分条件の一つは f が連続関数であること（定理 6.2.3）

である．特に３番目の「連続関数は可積分である」は非常に重要だから，結果だけでも頭に叩き込んでおくように！

まず,「上積分」などの定義から始めよう．ここでは区間 $[a,b]$ で定義された有界な関数 $f(x)$ に話を限る．$f(x)$ が有界でない場合や $[a,b]$ が有限の区間でない場合は，後（6.4 節）で「広義積分」として取り扱う．

- 分割 P に対して以下のように定義する：区間 $[x_{i-1}, x_i]$ における $f(x)$ の下限と上限を $m_i(f;P), M_i(f;P)$ と書く．そして

$$s(f;P) \equiv \sum_{i=1}^{n} m_i(f;P) \times (x_i - x_{i-1}), \qquad S(f;P) \equiv \sum_{i=1}^{n} M_i(f;P) \times (x_i - x_{i-1}) \tag{6.2.2}$$

を定義する．$s(f;P)$ を下限和，$S(f;P)$ を上限和という．

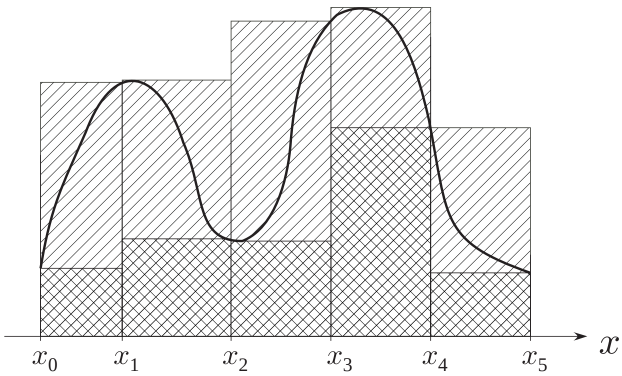
[6] 煎じ詰めれば「積分は和のお化け」である．ついでに「微分は差のお化け」である

- 更に，様々な細かさの P を考え，

$$s(f) = \sup\{s(f;P) \,\big|\, P \text{ は } [a,b] \text{ の分割 }\}, \qquad S(f) = \inf\{S(f;P) \,\big|\, P \text{ は } [a,b] \text{ の分割 }\} \tag{6.2.3}$$

  も定義する．$s(f)$ を下積分，$S(f)$ を上積分という．

$n = 5$ の場合の例を以下に示した．右上から左下への斜め斜線のところの面積が上限和，左上から右下への斜め斜線のところの面積が下限和である．ただし，図では下限和に相当する部分は両方の斜め線が入って十文字の模様になっている．



上の定義から，分割内の分点 $\vec{\zeta}$ の取り方にかかわらず，

$$s(f;P) \le R(f;P,\vec{\zeta}) \le S(f;P) \tag{6.2.4}$$

であることに注意しておこう（上の２つの図を比べてみよ）．

次の定理は，$s(f;P)$ や $S(f;P)$ は，それぞれが極限を持つことを保証する．

定理 6.2.1 (Darboux の定理，教科書の定理 33) 分割 P を限りなく細かくする（$|P| \to 0$）とき，下限和と上限和はそれぞれ一定の値に収束し，その行き先は (6.2.3) で定義された s と S である．つまり，

$$\lim_{|P|\to 0} s(f;P) = s(f), \qquad \lim_{|P|\to 0} S(f;P) = S(f) \tag{6.2.5}$$

がなりたつ.（ただし，$s(f) = S(f)$ とは限らない.）

では，上積分・下積分と積分可能性の関係はどうか？それぞれの P に対しては $s(f;P) \le S(f;P)$ だったから，

$$s(f) \le S(f) \tag{6.2.6}$$

であることはわかる．問題は上積分と下積分がいつ等しいかだ — 関数 f や区間 $[a,b]$ の取り方によってはこの２つは等しくないこともある．しかし，この２つが等しいことは定義 6.1.1 の積分可能性と同値だ，というのが次の定理である．

定理 6.2.2 (積分可能性の必要十分条件，教科書の定理 34) f が区間 $[a,b]$ 上で積分可能である必要十分条件は，上積分と下積分が一致することである．つまり

$$s(f) = S(f) \qquad \Longleftrightarrow \qquad f \text{ は可積分で，} \int_a^b f(x)dx = s(f) = S(f) \tag{6.2.7}$$

これで積分可能性の一般論はおしまいである．しかしこのままでは，与えられた関数に対して上積分，下積分を計算しないと積分可能かどうかがわからない．これは不便だから，積分可能性の簡単な十分条件を挙げておく：

定理 6.2.3 (連続関数は積分可能，教科書の定理 35) 関数 $f(x,y)$ 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．

講義ではこれらの定理の証明（説明）を行う．少し難解かもしれないが，大事なところだし，$\epsilon-\delta$ の非常に良い練習問題にもなっているから，ちょっと辛抱して欲しい．なお，残念ながら証明がチンプンカンプンな人も，諦める必要はない．次回からの積分の応用を勉強すれば，証明がわからなくても単位を取る事は十分に可能だ．

理解を深める問題：

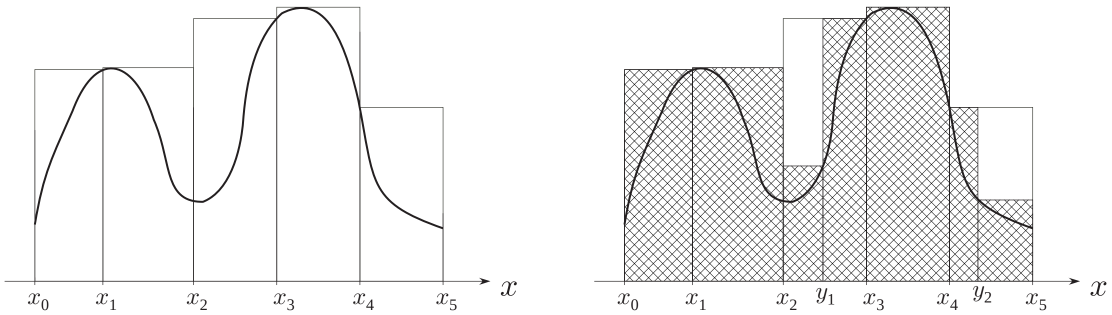
高校の時にもやったかもしれないが，良く知っている関数に対して，上積分，可積分を計算しよう．例えば，積分区間は $[-1,1]$ にして，$f(x)=x^2, x^3$ など，いくつかやってみることを強く奨める．

6.2.1 定理 6.2.1 と定理 6.2.2 の証明

定理 6.2.1 の証明の基本になるのは，以下の性質である．定理 6.2.2 の方は定理 6.2.1 からすぐに出る．

補題 6.2.4 $S(f;P)$ は，分割を細かくすると減少する．より正確にいうと，区間 $[a,b]$ の勝手な分割 P_1, P_2 をとってきて，これを合わせた（つまり，両方の分割の分点を全部集めた）分割を $P_{12}=P_1\cup P_2$ と書くと，$S(f;P_{12})\le S(f;P_1)$ および $S(f;P_{12})\le S(f;P_2)$ である．
同様に，$s(f;P)$ は分割を細かくすると増加する．

この補題は，$S(f;P)$ の定義からほとんどあたりまえである．以下にこの事情を図で例示した．



P_1 $\qquad\qquad$ $P_1 \bigcup P_2$

左側の図（の長方形の下の面積）が $P_1=(x_0,x_1,x_2,x_3,x_4,x_5)$ のみの場合の $S(f;P_1)$ である．一方，$P_2=(y_0,y_1,y_2,y_3)$ を考えると（$y_0=a, y_3=b$），右側の図の陰をつけた部分の面積が $P_1\cup P_2$ の場合の $S(f;P_1\cup P_2)$ である．図に示すように，白い 2 つの長方形の部分だけ，$S(f;P_1\cup P_2)$ が小さくなっている． □

（注）上ではわかりやすいようにわざと不正確な書き方をしたが，本来は「$S(f;P)$ は，分割を細かくすると増加しない」と書くべきであった．同様に，「$s(f;P)$ は，分割を細かくすると 減少しない」が正しい．

以下ではこの補題を用いて定理 6.2.1 と定理 6.2.2 を証明する．

定理 6.2.1 の証明 S の方のみ，証明する．s のほうも，いくつかの不等号の向きが逆になるだけで同じだ．

ちょっと考えると，定理 6.2.1 は当たり前に思える．なぜなら，補題 6.2.4 より，$S(f;P)$ は単調減少っぽく見えて，「有界な単調減少列は極限を持つ」から．しかし，これは早とちりだ．というのは，補題 6.2.4 は「P をより細かくしたら $S(f;P)$ は非増加」と言っているだけで，他の分割から出発して 細かくした行き先が，この P から出発した行き先と 等しいかどうか は保証の限りではない．この問題を解決するため，以下のように進む．

まず inf としての $S(f)$ の定義から，どんな分割 P に対しても $S(f)\le S(f;P)$ であることに注意しておこう：

$$\forall P,\quad S(f)\le S(f;P). \tag{6.2.8}$$

また，$S(f)$ は $S(f;P)$ の inf であるから，$S(f)$ と $S(f;P)$ の差がいくらでも小さくなるような分割 P もある：

$$\forall\epsilon>0,\ \exists P,\quad S(f;P)\le S(f)+\epsilon. \tag{6.2.9}$$

問題は，(6.2.9) が $|P'| \to 0$ なる任意の P' に対して成り立つか，つまり

$$(??) \qquad \forall\epsilon > 0,\ \exists\delta > 0, \quad |P'| < \delta \implies S(f;P') \le S(f) + \epsilon \qquad (??) \tag{6.2.10}$$

となっているか，ということである．

そこでまず，(6.2.8) の P を固定し，十分細かい分割 P' を,「P' の各ブロック内に P の分点が高々一つしかない」ようにとる．これは $|P'|$ を $|P|$ より小さくとれば絶対に実現できる．次に，P と P' を合わせた分割を考えると，これは P, P' よりも細かいので，細かい方の S の値が小さくなる：

$$S(f;P \cup P') \le S(f;P). \tag{6.2.11}$$

一方，n を P の分点の数，M, m は $[a,b]$ 内での f の上限と下限とすると，

$$S(f;P') - S(f;P \cup P') \le n(M-m)|P'| \tag{6.2.12}$$

が成り立つ．なぜなら，左辺の差への寄与は P' の分割ブロック中に P の分点が入っているときのみゼロでないが，このような分点の数は最大で n 個しかなく，そのような一つのブロックからの寄与は $(M-m)|P'|$ で押さえられるからだ（ここのところは図で納得するのがよい）．(6.2.11) と (6.2.12) から

$$S(f;P') \le S(f;P \cup P') + n(M-m)|P'| \le S(f;P) + n(M-m)|P'| \tag{6.2.13}$$

が結論できた．これと (6.2.9) を組み合わせると

$$S(f;P') \le S(f;P) + n(M-m)|P'| \le S(f) + \epsilon + n(M-m)|P'| \tag{6.2.14}$$

が得られる．さてここで P' を十分細かく，$n(M-m)|P'| < \epsilon$ となるようにとると（ここで，n は P のみで決まり，P' には関係ないことが効いている），

$$\forall\epsilon > 0,\ \exists\delta > 0, \quad |P'| < \delta \implies S(f;P') \le S(f) + 2\epsilon \tag{6.2.15}$$

が言える．よって,（ϵ は任意だから 2ϵ を ϵ と思い直して）(6.2.10) が結論できる． □

定理 6.2.2 の証明

（十分であること）Darboux の定理の証明中，任意の $\epsilon > 0$ に対して，$\delta > 0$ がとれて，

$$|P| < \delta \text{ ならば} \quad S(f;P) < S(f) + \epsilon \quad \text{かつ} \quad s(f;P) > s(f) - \epsilon \tag{6.2.16}$$

であることを見た — (6.2.15) 式．ところで，その定義から，リーマン和は

$$s(f;P) \le R(f;P,\vec{\zeta}) \le S(f;P) \tag{6.2.17}$$

を満たす．従って，$|P| < \delta$ である限り，どんな分割でも，どんな分点 $\vec{\zeta}$ の取り方に対しても，

$$s(f) - \epsilon \le s(f;P) \le R(f;P,\vec{\zeta}) \le S(f;P) \le S(f) + \epsilon \tag{6.2.18}$$

が成り立つことがわかる．ここでもし，定理の仮定のように $s(f) = S(f)$ であれば，$\delta \downarrow 0$ として（このとき，もちろん $\epsilon \downarrow 0$）

$$\lim_{|P| \to 0} R(f;P,\vec{\zeta}) = s(f) = S(f) \tag{6.2.19}$$

が結論できる．リーマン和の極限が確定するから，積分可能である．

（必要であること）ほとんど自明である．というのも，$S(f) - s(f) = c > 0$ と仮定すると，sup, inf としての定義から，

$$s(f;P) \le s(f) = S(f) - c \le S(f;P') - c \tag{6.2.20}$$

が勝手な P, P' に関して成り立つ．つまり，いくら頑張っても $S(f;P)$ と $s(f;P')$ のギャップを埋めることはできず，リーマン和の極限が存在しない（そのような分点をいくらでもとれる）．従って積分不可能である． □

6.2.2 一様連続性

定理 6.2.3 の証明のキーになるのは，以下の「一様連続性」と呼ばれる性質である．これは非常に大事な概念なので，少し詳しく述べておこう．

関数 $f(x)$ が $x=a$ で連続とは，$\lim_{x\to a} f(x) = f(a)$ であることだった．また，関数 $f(x)$ が区間 $[a.b]$ の各点で連続とは，その字のごとく，$[a,b]$ の中の任意の点 c にて $\lim_{x\to c} f(x) = f(c)$ となることであった．これを $\epsilon-\delta$ で書いてみると，

$$\forall c \in [a,b] \quad \forall \epsilon > 0 \quad \exists \delta(\epsilon, c) > 0 \qquad |x-c| < \delta(\epsilon, c) \implies |f(x)-f(c)| < \epsilon \tag{6.2.21}$$

ということになる．δ は ϵ に依存するのはもちろんであるが，一般には c にも依存する．特に $c\to\infty$ や $c\to 0$ で $\delta(\epsilon, c)$ がゼロになってしまうことも良くある．実はこのような例は期末テストや中間テストでも出題していた（例：$\lim_{x\to c} \frac{1}{x}$ を $\epsilon-\delta$ で求め，δ の取り方をも示せ，など）．ともかく，このような連続性は単に「連続」または「各点連続」という．

ところが，関数 $f(x)$ と考えている区間 $[a,b]$ の取り方によっては，上の $\delta(\epsilon, c)$ を c によらずにとれる，つまり $[a,b]$ 内のすべての c に共通の $\delta(\epsilon)$ をとれる，場合がある．このような場合，$f(x)$ は区間 $[a,b]$ で <u>一様連続</u> であるという．数式で書けば，

$$\forall \epsilon > 0 \quad \exists \delta(\epsilon) \quad \forall c \in [a,b] \qquad |x-c| < \delta(\epsilon) \implies |f(x)-f(c)| < \epsilon \tag{6.2.22}$$

となる場合，$f(x)$ は一様連続というのである.（c と δ の順序に注意！）

Remark. 「一様」という概念は $\epsilon-\delta$ の次に待ち受けている，大きな難関のようだ．一般に「一様」が問題になるのは以下のような状況である：

- ある種の極限が問題である．極限をとる変数を x とする．
- 極限をとる変数とは別の変数（パラメーター） y も存在する．
- 問題の極限が「別の変数」y をさほど気にしなくてもとれる場合，つまりいろいろな y の値に関しても同じように極限がとれ，<u>同じような収束の速さ</u> である場合，極限は「一様」であるという．
- y を気にする必要がある（特に y によって収束の速さが非常に異なる）場合，一様とは言わない．

今考えている「一様連続」の場合，極限をとるのは x について（$x\to c$）であり，別の変数（パラメーター）とは c である．連続になるための δ の取り方（x の極限を規定する）がパラメーター c に依存しない（すべての c に共通）ようにとれる，というのはパラメーター c の値をそんなに気にしなくてもよいと言うことだ．この事情を指して「一様」と言っているのである．

ついでに「一様」の例をもう一つ挙げておこう．比較のために一様連続も書いておく．

定義 6.2.5
(i) 区間 $[a,b]$ で定義された関数 $f(x)$ が一様連続とは

$$\forall \epsilon > 0 \quad \exists \delta(\epsilon) \quad \forall c \in [a,b] \qquad |x-c| < \delta(\epsilon) \implies |f(x)-f(c)| < \epsilon \tag{6.2.23}$$

が成り立つ場合をいう．
(ii) 区間 $[a,b]$ で定義された関数の列 $f_n(x)$ がある（$n=1,2,3,\ldots$）．このとき，$\lim_{n\to\infty} f_n(x) = g(x)$ が一様収束であるとは，

$$\forall \epsilon > 0 \quad \exists N(\epsilon) \quad \forall x \in [a,b] \qquad n > N(\epsilon) \implies |f_n(x)-g(x)| < \epsilon \tag{6.2.24}$$

となることをいう．

2 番目の例では $n\to\infty$ の極限を考えているのだが，その際の $\epsilon-N$ の N が x によらずにとれる（収束の速さが x にほとんどよらない）ことを指して「一様」と言っている．一様収束とは限らない収束のことを「各点収束」という．式で書けば

$$\forall x \in [a,b] \quad \forall \epsilon > 0 \quad \exists N(\epsilon, x) \qquad n > N(\epsilon, x) \implies |f_n(x)-g(x)| < \epsilon \tag{6.2.25}$$

ということで，N が一般には x にもよる．一様収束については数学概論 I でじっくりと，この講義でも後の方で少し，やるだろう．

さて，一様連続性については，以下の非常に重要な定理がある．

定理 6.2.6 (連続関数は閉区間で一様連続) $a < b$ を任意の実数とするとき，閉区間 $[a, b]$ 上の連続関数は<u>一様連続</u>である．つまり，任意の $\epsilon > 0$ に対して適当な $\delta(\epsilon) > 0$ がとれて，

$$\forall x, y \in [a, b] \qquad |x - y| < \delta(\epsilon) \quad \Longrightarrow \quad |f(x) - f(y)| < \epsilon \tag{6.2.26}$$

が成立する．数式の書き方からもわかる通り，$\delta(\epsilon)$ はすべての $x, y \in [a, b]$ に共通にとれる．

定理 6.2.6 の証明

ある種の背理法で証明してみよう．考えている区間 $[a, b]$ を 2^n こに等分割してできる小区間を $I_j^{(n)} = [a + \frac{j-1}{2^n}, a + \frac{j}{2^n}]$ と書く（$j = 1, 2, \ldots, 2^n$）．そして，$I_j^{(n)}$ 中の $f(x)$ の最大値を $M_j^{(n)}$，最小値を $m_j^{(n)}$ と書き，$d_j^{(n)} = M_j^{(n)} - m_j^{(n)}$ としよう（f が連続で，小区間は閉区間だから，ここでの最大値，最小値は必ず存在する—春学期の定理 4.2.3—ので，この定義は意味を持つ）．そして，それらの最大値を $d^{(n)} = \max_j d_j^{(n)}$ と書くことにする．

その定義から $d^{(n)}$ は n の単調非増加数列であり，非負（つまり下に有界）である．従って，$\lim_{n\to\infty} d^{(n)}$ が存在すること（更に極限値が非負であること）は保証されている．問題はこの極限の値であるが，定理を証明するにはこの極限がゼロ，つまり

$$\lim_{n\to\infty} d^{(n)} = 0 \tag{6.2.27}$$

を示せば十分である．理由は以下の通りである．(6.2.27) は

$$\forall \epsilon > 0 \quad \exists N \qquad n \geq N \quad \Longrightarrow \quad d^{(n)} < \epsilon \tag{6.2.28}$$

と同値であり，$d^{(n)}$ の定義から，これは

$$\forall \epsilon > 0 \quad \exists N \qquad n \geq N \quad \Longrightarrow \quad \Bigl(1 \leq j \leq 2^n \quad \Longrightarrow \quad d_j^{(n)} < \epsilon\Bigr) \tag{6.2.29}$$

とも同値である．

ところが，$d_j^{(n)}$ の定義を思い出すと，（1）(6.2.29) は同じ小区間 $I_j^{(n)}$ 内の任意の 2 点 x, y に対して $|f(x) - f(y)| < \epsilon$ を保証することがすぐにわかる．（2）また，x, y が隣り合った区間 $I_j^{(n)}$ と $I_{j+1}^{(n)}$ にそれぞれ入っている場合は（式を見やすくするため，両区間の境目の点を $c = a + \frac{j}{2^n}$ と書く）

$$|f(x) - f(y)| = \bigl|f(x) - f(c) + f(c) - f(y)\bigr| \leq \bigl|f(x) - f(c)\bigr| + \bigl|f(c) - f(y)\bigr| < \epsilon + \epsilon = 2\epsilon \tag{6.2.30}$$

となる．

互いの距離が $\frac{b-a}{2^n}$ 以下になっている $[a, b]$ 内の任意の 2 点 x, y はかならず上の（1）または（2）でカバーできるから，結果的に

$$\forall \epsilon > 0, \quad \exists N \qquad n \geq N \quad \Longrightarrow \quad \Bigl(|x - y| < \frac{b-a}{2^n} \quad \Longrightarrow \quad |f(x) - f(y)| < 2\epsilon\Bigr), \tag{6.2.31}$$

いやもっと簡単に（$\delta = \frac{b-a}{2^N}$ ととるつもりで）

$$\forall \epsilon > 0, \quad \exists \delta \qquad |x - y| < \delta > 0 \quad \Longrightarrow \quad |f(x) - f(y)| < 2\epsilon \tag{6.2.32}$$

が結論できる．これは一様連続性の定義に他ならないから，(6.2.27) が定理の証明には十分であることが示された．

さて，(6.2.27) 自身を証明するには，ある種の背理法を用いる．つまり，(6.2.27) が成り立たなかったと仮定してみる．この極限値が存在して非負であることが単調性から従うことは既に注意したから，(6.2.27) が成り立たないなら，残された可能性は極限値が正，つまり $d > 0$ が存在して

$$\lim_{n\to\infty} d^{(n)} = d \tag{6.2.33}$$

となることである．以下，これがあり得ない（f の連続性に反する）ことを示そう．

さて，もし仮に (6.2.33) となっているとすると，($d^{(n)}$ は単調非増加だったから）すべての n に対して

$$d^{(n)} \geq d \qquad \text{つまり} \qquad \text{各 } n \text{ に対して } j_n \text{ が存在して } d_{j_n}^{(n)} \geq d \tag{6.2.34}$$

となっているはずだ．上の j はもちろん n に依存するが，うまく j_n を選んで区間の減少列

$$I_{j_1}^{(1)} \supset I_{j_2}^{(2)} \supset \ldots \supset I_{j_n}^{(n)} \supset I_{j_{n+1}}^{(n+1)} \supset \ldots, \qquad \text{かつ　すべての } n \text{ で } d_{j_n}^{(n)} \geq d \tag{6.2.35}$$

となるものを選び出すことができる[7]．区間 $I_{j_n}^{(n)}$ の幅は $\frac{b-a}{2^n}$ であって，これは $n \to \infty$ でゼロに行く．従って「区間縮小法」によって，$\displaystyle\bigcap_{n=1}^{\infty} I_{j_n}^{(n)}$ に属する実数 c がただ一つ存在する[8]．ここでもちろん，$a \leq c \leq b$ である．

ところが，この c においては関数 $f(x)$ は連続ではあり得ない[9]．なぜなら，もし c にて連続であれば $\delta > 0$ が存在して，$|x-c| < \delta$ ならば $|f(x) - f(c)| < \frac{d}{3}$ となっているはずである．(連続性の定義の ϵ を $\frac{d}{3}$ とした.）しかし，もしこうであれば三角不等式から

$$|x-c| < \delta \text{ かつ } |y-c| < \delta \quad \text{ならば} \quad |f(x) - f(y)| \leq |f(x) - f(c)| + |f(c) - f(y)| < \frac{d}{3} + \frac{d}{3} = \frac{2}{3}d \tag{6.2.36}$$

となるが，$\frac{b-a}{2^n} < \delta$ となるような大きな n では，(6.2.36) は $d_{j_n}^{(n)} \geq d$ に矛盾する．つまり，c では $f(x)$ が連続になれないのだ．

というわけで，仮に (6.2.33) であれば，$f(x)$ は c にて連続でない，ことが結論されてしまった．これは定理の仮定（f の連続性）に矛盾するから容認できない．つまり，(6.2.33) 自身があり得ないわけで，背理法が完成した． □

6.2.3　定理 6.2.3 の証明

一様連続性がわかれば，証明は簡単だ．

定理 6.2.3 の証明

定理 6.2.2 を考えに入れると，$s(f) = S(f)$ が言えれば定理 6.2.3 の証明には十分だ．そのためには，

$$(\,?\,)\ \text{任意の}\epsilon > 0 \ \text{ に対して，} 0 \leq S(f;P) - s(f;P) < \epsilon \ \text{ なる分割 } P \text{ がとれる} \tag{6.2.37}$$

となることを証明すればよい．以下ではこれよりも強い，

$$(\,?\,)\ \forall\epsilon > 0 \quad \exists\delta > 0 \qquad |P| < \delta \implies 0 \leq S(f;P) - s(f;P) < \epsilon \tag{6.2.38}$$

を証明しよう．さて，積分領域の長さは $b-a$ なので，$\epsilon' = \dfrac{\epsilon}{b-a}$ とおく．f の一様連続性から $\delta > 0$ が存在して，

$$|x-y| < \delta \quad \Longrightarrow \quad |f(x) - f(y)| < \epsilon' \tag{6.2.39}$$

とすることができる．そこで，分割 P を，$|P| < \delta$ となるようなものにとろう．この分割は十分小さいので，同じ小区間 $I_i = [x_{i-1}, x_i]$ に属する x, y は $|x-y| < \delta$ を満たしており，したがって $|f(x) - f(y)| < \epsilon'$ も満たされる．よって，I_i 上での f の上限 M_i と下限 m_i は $0 \leq M_i - m_i \leq \epsilon'$ を満たす．これを i について和をとると，

$$0 \leq S(f;P) - s(f;P) = \sum_i (M_i - m_i)(x_i - x_{i-1}) \leq \sum_i \epsilon'(x_i - x_{i-1}) = \epsilon' \, |b-a| = \epsilon \tag{6.2.40}$$

が得られる．つまり，(6.2.38) が証明できた．メデタシメデタシ． □

---

[7]なぜかというと，小区間 $I_j^{(n)}$ が小区間 $I_k^{(n+1)}$ と $I_{k+1}^{(n+1)}$ にわかれたとすると，$d_j^{(n)} \geq \max\left\{d_k^{(n+1)}, d_{k+1}^{(n+1)}\right\}$ が成り立っている（つまり，ある小区間で $d_k \geq d$ であれば，この小区間の「親」区間でも $d_j \geq d$）からである

[8]証明を進めるためにはこのような c が少なくとも一つ存在すれば十分だから，区間縮小法を持ち出すまでもない

[9]以下，$a < c < b$ となっている場合に説明する．$a = c$ または $c = b$ の場合もほとんど同じである

理解を深めるための問題（レポートで出題済み）：

定積分の定義に従って（上の定理を使っても良い），積分 $\displaystyle\int_a^b f(x)dx$ を求めてみよ．ここで

$$a=0, \quad b=1, \qquad f(x)=x \tag{6.2.41}$$

とする．余力のある人は他の $f(x)$ の場合も（例：$f(x)=x^2$）やってみようか．

この問題の計算は大変である．しかし，一回はやっておいた方が，後々のためになる．

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．

6.3 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 以下の性質が簡単に導かれる．

定理 6.3.1 (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，

$$\int_a^b \bigl\{f(x)+g(x)\bigr\}dx = \int_a^b f(x)dx + \int_a^b g(x)dx \tag{6.3.1}$$

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

$$\int_a^b \bigl\{\alpha f(x)\bigr\}dx = \alpha\int_a^b f(x)dx \tag{6.3.2}$$

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

$$f \qquad \mapsto \quad \int_a^b f(x)dx \tag{6.3.3}$$

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 6.3.2 (区間に関する加法性)
(i) $a<c<b$ のとき，

$$\int_a^c f(x)dx + \int_c^b f(x)dx = \int_a^b f(x)dx \tag{6.3.4}$$

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 (6.3.4) 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 6.3.3 $a < b$ 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)

$$[a,b] \text{ で } f(x) \geq 0 \qquad \Longrightarrow \qquad \int_a^b f(x)dx \geq 0 \tag{6.3.5}$$

(ii) （単調性）

$$[a,b] \text{ で } f(x) \leq g(x) \qquad \Longrightarrow \qquad \int_a^b f(x)dx \leq \int_a^b g(x)dx \tag{6.3.6}$$

(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

$$|x - c| < \delta \text{ かつ } a \leq x \leq b \quad \text{では} \quad f(x) \geq \frac{f(c)}{2} \tag{6.3.7}$$

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，

$$\int_a^b f(x)dx = \int_a^{c-\delta} f(x)dx + \int_{c-\delta}^{c+\delta} f(x)dx + \int_{c+\delta}^b f(x)dx \tag{6.3.8}$$

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 (6.3.7) をもちいると，

$$\int_{c-\delta}^{c+\delta} f(x)dx \geq \int_{c-\delta}^{c+\delta} \frac{f(c)}{2}dx = \frac{f(c)}{2} \times 2\delta = f(c)\,\delta > 0 \tag{6.3.9}$$

である．従って，(6.3.8) 自身も正である． □

系 6.3.4 $a < b$ のとき，両辺の積分が定義できているなら，

$$\left| \int_a^b f(x)dx \right| \leq \int_a^b \left| f(x) \right| dx. \tag{6.3.10}$$

証明 簡単だ．

$$-|f(x)| \leq f(x) \leq |f(x)| \tag{6.3.11}$$

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 6.3.3 の (ii) から，積分結果に対しても不等号が成り立つ：

$$-\int_a^b |f(x)|\,dx \leq \int_a^b f(x)dx \leq \int_a^b |f(x)|\,dx \tag{6.3.12}$$

これは (6.3.10) に他ならない． □

定理 6.3.5 (積分の平均値の定理，教科書の定理 36 の拡張) $a < b$ 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，

$$\int_a^b f(x)g(x)dx = f(\xi)\int_a^b g(x)dx \tag{6.3.13}$$

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，

$$m\int_a^b g(x)dx \leq \int_a^b mg(x)dx \leq \int_a^b f(x)g(x)\,dx \leq \int_a^b Mg(x)dx = M\int_a^b g(x)dx \tag{6.3.14}$$

が得られる．以下，これから (6.3.13) を示す．

まず，$\int_a^b f(x)dx = 0$ ならば，(6.3.13) の両辺が共にゼロとなり，(6.3.13) 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 (6.3.14) の両辺を $\int_a^b g(x)dx$ で割って

$$m \leq \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \leq M \tag{6.3.15}$$

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，

$$\frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} = f(\xi) \tag{6.3.16}$$

となる $\xi \in [a, b]$ が存在する．この式の分母をはらえば (6.3.13) になる． □

上の定理で $g(x) \equiv 1$ とおくと以下の系になる：

系 6.3.6 (積分の平均値の定理，教科書の定理 36) $a < b$ のとき，区間 $[a, b]$ 上で $f(x)$ が連続なら，区間 $[a, b]$ 内の一点 ξ が存在して，

$$\int_a^b f(x)dx = f(\xi)\,(b-a) \tag{6.3.17}$$

（注）上の定理と系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？

$$f(\xi) = \frac{1}{b-a}\int_a^b f(x)dx, \qquad f(\xi) = \frac{\int_a^b f(x)g(x)dx}{\int_a^b g(x)dx} \tag{6.3.18}$$

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 6.3.7 (微分積分学の基本定理，教科書の定理 37) $a < b$ とし，区間 $[a, b]$ 上では $f(x)$ が連続とする．区間 $[a, b]$ 内の一点を c として

$$F(y) := \int_c^y f(x)dx \tag{6.3.19}$$

を定義する．このとき，F は $[a, b]$ 内の各点 y で微分可能で，

$$\frac{d}{dy}F(y) = f(y) \tag{6.3.20}$$

である．

（注）$F(y)$ がちゃんと定義できていることは定理 6.2.3 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

$$\frac{d}{dy}F(y) = \lim_{h\to 0}\frac{F(y+h)-F(y)}{h} = \lim_{h\to 0}\frac{1}{h}\int_y^{y+h} f(x)dx \tag{6.3.21}$$

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 6.3.6 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

$$\frac{d}{dy}F(y) = \lim_{h\to 0}\frac{F(y+h)-F(y)}{h} = \lim_{h\to 0} f(\xi) \tag{6.3.22}$$

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 6.3.8 $a < b$ 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

$$F_2(x) := F_1(x) + C \tag{6.3.23}$$

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

$$F_2(x) - F_1(x) = C \qquad (a \le x \le b) \tag{6.3.24}$$

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 (6.3.19) 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．

6.4 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 6.1.1 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もより自然なのである．

（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 6.1.1 には含まれていない，というつもりだろうか.

6.4.1 関数の極限についての補足

以下で広義積分をある種の極限として定義するが，そのような極限が存在するか否かの判定には，以下の「コーシーの判定法」の関数バージョンが便利である．これは春学期のノートに定理 3.4.6 として書いてあるが，念のために再録する．

定理 6.4.1 $a<c<b$ とし，$F(x)$ を区間 (a,c) と区間 (c,b) で定義された関数とする．極限 $\lim_{x\to c} F(x)$ が存在するための必要十分条件は

$$\forall\epsilon>0,\ \exists\delta(\epsilon)>0,\quad \Bigl(|x-c|<\delta(\epsilon) \text{ かつ } |y-c|<\delta(\epsilon) \quad\Longrightarrow\quad |F(x)-F(y)|<\epsilon\Bigr) \tag{6.4.1}$$

がなりたつ事である．

6.4.2 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 6.4.2 $a<b$ とする．
(0) 今までのリーマン積分の定義による $\int_a^b f(x)dx$ が存在するなら，これまで通りでよい．今までの通りには $\int_a^b f(x)dx$ が定義できない場合は以下のように定義していく．
(i) 半開区間 $(a,b]$ 上の関数 $f(x)$ に対して，

$$\lim_{c\to a+0}\int_c^b f(x)dx \tag{6.4.2}$$

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) 同様に半開区間 $[a, b)$ 上の関数 $f(x)$ に対して，

$$\lim_{d\to b-0}\int_a^d f(x)dx \tag{6.4.3}$$

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，上の (i)(ii) が適用できない（つまり $\int_c^b f(x)dx$ も $\int_a^d f(x)dx$ もうまく定義できない）が

$$\lim_{\substack{c\to a+0\\ d\to b-0}}\int_c^d f(x)dx \tag{6.4.4}$$

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，上のようにしても広義積分が定義できないとき，この広義積分は存在しない（または発散する）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 6.4.2 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，

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx \tag{6.4.5}$$

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 6.4.2 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，

$$\int_a^b f(x)dx = \lim_{e\to c-0}\int_a^e f(x)dx + \lim_{d\to c+0}\int_d^b f(x)dx \tag{6.4.6}$$

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

$$\int_{-1}^1 \frac{1}{x}dx = \int_{-1}^0 \frac{1}{x}dx + \int_0^1 \frac{1}{x}dx \tag{6.4.7}$$

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 6.4.2 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

$$\lim_{\epsilon\to+0}\left[\int_{-1}^{-\epsilon}\frac{1}{x}dx+\int_{\epsilon}^{1}\frac{1}{x}dx\right] \tag{6.4.8}$$

を考えると，括弧の中は被積分関数が奇関数だからゼロになり，従って極限値もゼロである（というふうに極限値は存在してしまう）．正しい定義（極限は別々にとる）との違いをよく認識されたい．なお，この「補足」のようにそろえて極限をとったものには,「コーシーの主値（積分）」の名前がついている．これは将来，複素積分などで出てくると思うが，問題によっては，このようにちょっと「ずるい」定義[10]が役に立つ事もある．

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,

$$\int_a^b f(x)dx=\int_a^{c_1} f(x)dx+\int_{c_1}^{c_2} f(x)dx+\int_{c_2}^{c_3} f(x)dx+\int_{c_3}^{b} f(x)dx \tag{6.4.9}$$

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 6.4.2 にしたがって定義できるかどうかを考える訳だ.

6.4.3 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 6.4.3 $f(x)$ 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

$$\lim_{M\to\infty}\int_a^M f(x)dx \tag{6.4.10}$$

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

$$\lim_{L\to-\infty}\int_L^b f(x)dx \tag{6.4.11}$$

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

$$\lim_{\substack{L\to-\infty\\ M\to\infty}}\int_L^M f(x)dx \tag{6.4.12}$$

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

最後の (iii) については定義 6.4.2(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.

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

6.4.4 （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，

$$\int_0^\infty \frac{\sin x}{\sqrt{x}} dx = \int_0^1 \frac{\sin x}{\sqrt{x}} dx + \int_1^\infty \frac{\sin x}{\sqrt{x}} dx = \lim_{\epsilon \to +0} \int_\epsilon^1 \frac{\sin x}{\sqrt{x}} dx + \lim_{L \to \infty} \int_1^L \frac{\sin x}{\sqrt{x}} dx \tag{6.4.13}$$

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．

最後に補足として，広義積分の存在に関する定理を結果だけ述べておく．類似の定理としては，教科書の定理 40 と定理 40′ がある．

定理 6.4.4 広義積分の収束について，以下がなりたつ（共に十分条件であることには注意が必要）．
(i) 半開区間 $[a, b)$ 上で連続な関数 $f(x)$ に対して定数 $C > 0, \alpha < 1$ が存在して

$$|f(x)| \le C(b - x)^{-\alpha} \tag{6.4.14}$$

がなりたっているなら，広義積分 $\displaystyle\int_a^b f(x)dx$ は（絶対）収束する．
(ii) 半無限区間 $[a, \infty)$ 上の連続な関数 $f(x)$ に対して，定数 $C > 0, \alpha > 1$ が存在して

$$|f(x)| \le Cx^{-\alpha} \tag{6.4.15}$$

がなりたっているなら，広義積分 $\displaystyle\int_a^\infty f(x)dx$ は（絶対）収束する．

（注）広義積分 $\displaystyle\int_a^b |f(x)|dx$ が存在するとき，広義積分 $\displaystyle\int_a^b f(x)dx$ は<u>絶対収束</u>するという．級数の場合と同じく，絶対収束する広義積分は普通の収束もする（証明はコーシーの条件をチェック）．

6.5 一様収束と極限の順序交換（この節は半分おまけ）

この節の内容は「数学概論 I」とも重複する部分が多いと思われるので，基礎的な部分に絞って述べることにする．いろいろな応用例は多分，数学概論にお任せすることになるだろう（教科書では６章の 3,4 節）．

まず，一様収束の定義を書いておこう．比較のために，普通の収束も書くと，以下のようになる．

定義 6.5.1 (一様収束) 区間 $[a,b]$ で定義された関数の列 $f_n(x)$ がある（$n=1,2,3,\ldots$）．この列について：
(i) 関数列 $\{f_n(x)\}$ が区間 $[a,b]$ で関数 $f(x)$ に <u>各点収束</u> するとは各点 x で $\lim_{n\to\infty} f_n(x) = f(x)$ となること，つまり

$$\forall\epsilon>0 \quad \forall x\in[a,b] \quad \exists N(\epsilon,x) \qquad n>N(\epsilon,x) \quad\Longrightarrow\quad |f_n(x)-f(x)|<\epsilon \tag{6.5.1}$$

が成り立つ場合をいう．
(ii) 関数列 $\{f_n(x)\}$ が区間 $[a,b]$ で関数 $f(x)$ に <u>一様収束</u> するとは

$$\forall\epsilon>0 \quad \exists N(\epsilon) \quad \forall x\in[a,b] \qquad n>N(\epsilon) \quad\Longrightarrow\quad |f_n(x)-f(x)|<\epsilon \tag{6.5.2}$$

となることをいう.（この状況を「$\lim_{n\to\infty} f_n(x) = f(x)$ の収束が一様である」ということもある.）

各点収束と一様収束の違いは <u>N が x に依存するかしないか</u> である．より正確に言うと，x に依存しないように N をとることができれば一様収束，いくら頑張っても N が x に依存してしまう場合が（一様収束でない）各点収束，である．なお，定義をよく見ればわかるように，一様収束であれば各点収束の条件も満たされている．この意味で，一様収束は各点収束よりも強い（より強い性質を要求する）概念である．

以下では，この一様収束の概念が，如何に自然に現れるかを，いくつかの「２つの極限の問題」を通してみていく事にしよう．以下では特に断らない限り，ある有限な区間 $I=[a,b]$ で定義された関数の列 $f_n(x)$（$n=1,2,3,\ldots$）を考える．

6.5.1 一様収束，極限と積分の順序交換

まずは積分つながりで,「積分と極限の交換」から行ってみよう．積分自身がリーマン和の極限で定義されているから，これはれっきとした「極限の順序交換」の問題である．新居さんの数学入門でもプロジェクト問題の一つとして考えた．

関数列 $f_n(x)$ を

$$f_n(x) = \begin{cases} n & (0<x<1/n) \\ 0 & (\text{それ以外}) \end{cases} \tag{6.5.3}$$

と定義する．このとき，

$$(\text{？？}) \qquad \lim_{n\to\infty}\left[\int_0^1 f_n(x)dx\right] = \int_0^1 \left(\lim_{n\to\infty} f_n(x)\right)dx \qquad (\text{？？}) \tag{6.5.4}$$

が成り立つだろうか？

答えは「成り立たない」である[11]．つまり，この関数列については，極限 $\lim_{n\to\infty}$ と積分 $\int_0^1$ を交換することはできないのだ．しかし一方で，極限と積分が交換できるような例もある．例えば,

$$g_n(x) = \begin{cases} 1 & (0<x<1/n) \\ 0 & (\text{それ以外}) \end{cases} \tag{6.5.5}$$

に対しては

$$\lim_{n\to\infty}\left[\int_0^1 g_n(x)dx\right] = \int_0^1 \left(\lim_{n\to\infty} g_n(x)\right)dx \tag{6.5.6}$$

[11] なぜ成り立たないのか，各自で納得すること．少なくとも「数学入門」ではここのところが怪しかった人が多かったと聞いている

が成り立つ（両辺ともにゼロ）．この２つのケースの違いは何だろうか？

もう少し問題を整理したい．$f(x) = \lim_{n\to\infty} f_n(x)$ と書くと (6.5.4) は

$$（？？）\qquad \lim_{n\to\infty}\left[\int_0^1 f_n(x)dx\right] = \int_0^1 f(x)dx \qquad （？？） \tag{6.5.7}$$

と等価であり，これは

$$（？？）\qquad \lim_{n\to\infty}\left[\int_0^1 \Big\{f_n(x) - f(x)\Big\}dx\right] = 0 \qquad （？？） \tag{6.5.8}$$

とも等価である．そこで $g_n(x) = f_n(x) - f(x)$ と書けば，問題は次のように定式化される．

> 問題：区間 $[a,b]$ で定義された関数列 $g_n(x)$ がすべての x で $\lim_{n\to\infty} g_n(x) = 0$ をみたす場合，$\lim_{n\to\infty}\int_a^b g_n(x) = 0$ と言えるだろうか？一般にこうとは言えないならば，言えるための十分条件は何だろうか？

少し発見法的に考えてみよう．$\lim_{n\to\infty} g_n(x) = 0$ ということは

$$\forall\epsilon > 0 \quad \exists N(\epsilon, x) \qquad n > N(\epsilon, x) \quad \Longrightarrow \quad |g_n(x)| < \epsilon \tag{6.5.9}$$

ということだ．一見，これで十分のように見える．なぜなら，もしすべての x に対して $|g_n(x)| < \epsilon$ と なっているなら，

$$\left|\int_a^b g_n(x)dx\right| \le \int_a^b |g_n(x)|dx \le (b-a)\epsilon \tag{6.5.10}$$

となるからだ．上の「もし」以下は完全に正しい．問題はむしろ,「もし」以下の条件がなりたつとは限らない点にある．というのは，$\lim_{n\to\infty} g_n(x) = 0$ というだけでは，(6.5.9) の N は一般には x にも依存するからだ．つまり，すべての x に対して $|g_n(x)| < \epsilon$ となるような n がとれないかもしれないのである．実際，(6.5.3) の $f_n(x)$ に対して $g_n(x) = f_n(x) - 0$ （この例では $f_n(x)$ の極限は恒等的にゼロだから）を考えると，上のような n がとれないことがわかる．

逆にいうと，もし適当な n に対して，すべての x で $|g_n(x)| < \epsilon$ が成り立つならば何も問題なく，(6.5.10) が結論できる．つまり，普通の収束よりつよい，新たな収束の概念が必要とされている訳だ．定義を思い出すと，これが「一様収束」に他ならない．

以上の発見法的な議論から直ちに，極限と積分の順序交換に関する以下の定理が証明できる．この定理を見れば，「一様収束」の概念は割合自然に見えるであろう．

> **定理 6.5.2 (積分と極限の交換；教科書の定理 52)** 区間 $[a,b]$ で定義された関数の列 $f_n(x)$（$n = 1, 2, 3, \ldots$）がこの区間で $f(x)$ に 一様収束 するなら，
>
> $$\lim_{n\to\infty}\int_a^b f_n(x)dx = \int_a^b \Big\{\lim_{n\to\infty} f_n(x)\Big\}dx = \int_a^b f(x)dx \tag{6.5.11}$$
>
> が成立する．つまり，極限と積分を交換できる．

（注）一様収束は (6.5.11) の順序交換ができるための 十分条件 にすぎないことは強調しておく．一様収束していなくても (6.5.11) ができる例はいくらでもある．

証明：

上に書いた事でほとんどつきているが，非常に重要だから書いておく．一様収束の定義から

$$\forall\epsilon > 0 \quad \exists N(\epsilon) \quad \forall x \in [a,b] \qquad n > N(\epsilon) \quad \Longrightarrow \quad |f_n(x) - f(x)| < \epsilon \tag{6.5.12}$$

である．上の ϵ を固定して積分の差を計算すると

$$\int_a^b f_n(x)dx - \int_a^b f(x)dx = \int_a^b \{f_n(x) - f(x)\}dx \tag{6.5.13}$$

なので，両辺の絶対値をとって

$$\left|\int_a^b f_n(x)dx - \int_a^b f(x)dx\right| \le \int_a^b |f_n(x) - f(x)|dx \le \int_a^b \epsilon\, dx = \epsilon(b-a) \tag{6.5.14}$$

が得られる．ところが，ϵ は（n を大きくとる事で）いくらでも小さくできる．これはつまり，上の左辺の差が（n を十分に大きくとると）いくらでも小さくできる事を意味する．つまり左辺の $n \uparrow \infty$ での極限はゼロである． □

上の $f_n(x)$ が級数の形の場合を特に書いておくと，以下のようになる.（これは「数学概論」の範囲だが，参考のために載せた.）

系 6.5.3 (教科書の定理 52′)
(i) 区間 $I = [a, b]$ で $\sum_{n=0}^{\infty} f_n(x)$ が $F(x)$ に一様収束し，かつ各 $f_n(x)$ が連続である時，

$$\sum_{n=0}^{\infty}\left[\int_a^b f_n(x)dx\right] = \int_a^b \left[\sum_{n=0}^{\infty} f_n(x)\right]dx = \int_a^b F(x)dx.$$

(iii) べき級数 $\sum_{n=0}^{\infty} a_n x^n$ の収束半径を R とすると，$f(x) \equiv \sum_{n=0}^{\infty} a_n x^n$ は $(-R, R)$ 内の任意の閉区間 $[a, b]$ において

$$\int_a^b f(x)dx = \sum_{n=0}^{\infty}\left[\int_a^b a_n x^n dx\right]$$

を満たす．特に，$a = 0, b = x$ として $|x| < R$ では

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

（補足１）積分と極限の順序交換については，一様収束を仮定しない，より一般の形の定理が成り立つ．例えば，

定理 6.5.4 (Arzelà の定理，小平の本の定理 5.10, 5.11)
区間 $I = (a, b)$ で定義された連続関数の列 $\{f_n(x)\}$ が

- n について一様有界，つまり n, x によらない定数 M があってすべての $n \ge 0$ と $x \in I$ に対して $|f(x)| \le M$
- 極限 $f(x) := \lim_{n\to\infty} f_n(x)$ が存在して区間 I で連続

を満たしているとする．このとき，$\int_a^b dx$ と極限 $n \to \infty$ は交換できる．つまり

$$\lim_{n\to\infty}\int_a^b f(x)dx = \int_a^b \Bigl(\lim_{n\to\infty} f_n(x)\Bigr)dx = \int_a^b f(x)dx \tag{6.5.15}$$

が成り立つ．

（補足２）考えている区間が無限の場合（広義積分）の積分と極限の順序交換はもう少し大変だ．単に一様収束しているだけでは足りない．その場合の典型的な定理は以下のようになる

定理 6.5.5 (ルベーグの優越収束の定理もどき；小平の本の定理 5.12)
区間 $I = (a, \infty)$ で定義された連続関数の列 $\{f_n(x)\}$ と関数 $g(x)$ が

- g は f の優関数，つまりすべての $n \ge 0$ と $x \in I$ に対して $|f(x)| \le g(x)$
- $\displaystyle\int_a^{\infty} g(x)dx < \infty$ （広義積分として）
- 極限 $f(x) := \lim_{n\to\infty} f_n(x)$ が存在して区間 I で連続

を満たしているとする．このとき，$\int_a^\infty dx$ と極限 $n \to \infty$ は交換できる．つまり（積分は広義積分として解釈して）

$$\lim_{n\to\infty} \int_a^\infty f(x)dx = \int_a^\infty \Big(\lim_{n\to\infty} f_n(x) \Big) dx = \int_a^\infty f(x)dx \tag{6.5.16}$$

が成り立つ．

この定理で要求されている条件は単なる一様収束よりは強い事に注意しよう．単なる一様収束では足りない例を考えてみると，理解が深まるだろう．また，小平の本の 5.4 節にはこのようなことが一杯載っているから，興味のある人は一読をお勧めする．

（補足３） 何回か言ったように,「リーマン積分」はその定義が少しきつすぎる（条件が厳しくて，リーマン積分が定義できない関数が多すぎ）．それを改良した，もっと自然な「ルベーグ積分」というものがあり，現在の解析学ではこのルベーグ積分を使う事が普通になっているし，それが自然である．そのような訳で，リーマン積分ではややこしい条件（一様収束）つきの定理もルベーグ積分で書けば簡単になる事は多い．実際，上の定理 6.5.5 は，実はルベーグ積分で成り立つ定理を翻訳したものである．

ただし，ルベーグ積分を理解するには,「測度論」をかなり一生懸命やる必要があり，一年のこの時期では少し無理があると思われるので，この講義では触れない．

6.5.2 極限と連続性

今度は「連続性と極限の交換」を考える．と言っても，これでは何の事かわからんかもしれないが，要するに以下の問題を考えるわけ．

> 区間 $I = [a, b]$ で定義された関数の列 $f_n(x)$（$n = 1, 2, 3, \ldots$）があって，各点 $x \in I$ で $\lim_{n\to\infty} f_n(x) = f(x)$ が存在する．更に，各 n では $f_n(x)$ は x について連続である．このとき，極限の $f(x)$ は x について連続だろうか？

極限をとる前の関数が連続なら，極限の後も連続か，ということで，形式的には「極限と連続性の交換」という感じの問いかけである．

まあ，もう予想がついているだろうが，上の問いに対する答えも，一般には「なりたたない」である．例えば，区間 $[-1, 1]$ で定義された関数列を

$$f_n(x) = \begin{cases} 0 & (-1 \le x \le 0) \\ nx & (0 < x \le 1/n) \\ 1 & (1/n < x \le 1) \end{cases} \tag{6.5.17}$$

と定義すると，これは連続である．しかし $n \to \infty$ の極限は

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & (-1 \le x \le 0) \\ 1 & (0 < x \le 1) \end{cases} \tag{6.5.18}$$

となって，$x = 0$ で連続ではない！

そこで上の問いの結論が「成り立つ」ための十分条件として，またもや一様連続が登場するのである：

定理 6.5.6 (極限と連続性；教科書の定理 51) 区間 I で $\{f_n\}$ が $f(x)$ に一様収束し，かつ各 $f_n(x)$ が連続である時， $f(x)$ も連続である．

（証明らしきもの）ちゃんとした証明はどの本にも書いてあるから，ここでは発見法的に理解する事を試みる．やりたいのは $f(x)$ の連続性の証明だから，$I = [a, b]$ 内の一点を c として，

$$\lim_{x\to c} f(x) = f(c), \quad \text{つまり} \quad \forall\epsilon > 0,\ \exists\delta > 0 \quad |x - c| < \delta \Longrightarrow |f(x) - f(c)| < \epsilon \tag{6.5.19}$$

という事だ．これが証明できるとしたら「$f_n(x)$ が連続であること」しか手がかりが無いだろうから，こいつを使うつもりで書き直していく．つまり，$f(x)-f(c)$ を $f_n(x)-f_n(c)$ で近似しようと思って書き直すと，恒等式：

$$f(x)-f(c)=f(x)-f_n(x)+f_n(x)-f_n(c)+f_n(c)-f(c) \tag{6.5.20}$$

に三角不等式を使って

$$|f(x)-f(c)|\le|f(x)-f_n(x)|+|f_n(x)-f_n(c)|+|f_n(c)-f(c)| \tag{6.5.21}$$

が得られる．ここで右辺の 3 つの項はそれぞれゼロに行くように見える：$f_n(x)-f(x)$ と $f_n(c)-f(c)$ は，共に f_n が f に収束するから $n\to\infty$ でゼロに行く．$f_n(x)-f_n(c)$ は f_n が連続だから，$x\to c$ でゼロに行く．

さて問題は，我々はここで 2 つの極限（$x\to c$ と $n\to\infty$）をとる必要があることだ．(6.5.21) の第 3 項は（c が固定されているから）$n\to\infty$ だけ考えれば良くって，何も問題ない．しかし，第 1 項と第 2 項は x と n の関係によっては，うまく行かないかもしれない．つまり，x を固定した上で $n\to\infty$ とするなら第 1 項はゼロになるけども，$x\to c$ と動きつつ $n\to\infty$ でなら，どうなるかわからない．第 2 項も，n を固定して $x\to c$ ならゼロになるけど，n が無限大にいくのと同時進行されると，良くわからない．困った事に，第一項と第二項がうまく行くための極限の順序が逆のようなのだ．そのため，単に「各点収束」だけでは困った事が起こりうる．

実際，(6.5.17) の関数に対して $c=0$ として，(6.5.21) を考えてみよう．n を固定して $x>0$ を 0 に近づけると，第 2 項はゼロに近づくが第一項は $1-0=1$ に近づく．逆に $x>0$ を固定して $n\to\infty$ とすると，第 1 項はゼロに行くけど，第 2 項は $1-0=1$ に近づく．どっちにせよ，うまく行かない．

この問題を解決してくれるのが「一様収束」である．すなわち，一様収束を仮定すれば，$x\to c$ と $n\to\infty$ は以下のようにとっていくと良い．

- まず勝手な $\epsilon>0$ を決める．以下では (6.5.21) の各項が ϵ より小さくなる事を示そう．
- 一様収束の定義から，$N(\epsilon)$ がとれて（これは x に依存しない事にくれぐれも注意！），$n\ge N(\epsilon)$ ならば第一項と第 3 項は ϵ より小さくできる．そこで，このような n を一つ固定する．
- 次に，上で決めた $N(\epsilon)$ に対して，(6.5.21) の第 2 項が ϵ より小さくなるような x の範囲を考える．$f_{N(\epsilon)}(x)$ は連続だから，ある $\delta(\epsilon,N(\epsilon))>0$ がとれて，$|x-c|<\delta(\epsilon,N(\epsilon))$ ならば (6.5.21) の第 2 項が ϵ より小さくなる．

以上から $\epsilon, N(\epsilon), \delta(\epsilon,N(\epsilon))$ の順番に決めていくことができ，この時に (6.5.21) の各項が ϵ より小さくなる事がわかった．つまり，このとき，(6.5.21) は 3ϵ よりも小さい．任意の ϵ に対して，$|x-c|$ を十分小さくすると $|f(x)-f(c)|<3\epsilon$ とできるのだから，これは $f(x)$ が $x=c$ で連続である事を意味する． □

6.5.3 極限と微分の順序交換

次に，微分と極限をみてみよう．残念ながら，積分と極限の時ほど条件は簡単ではない．特に，$f_n(x)\to f(x)$ の収束が一様収束だけでは足りない．

定理 6.5.7 (極限と微分；教科書の定理 53, 53′) (i) $f_n(x)$ は $I=[a,b]$ で連続的微分可能（C^1-級），$\{\frac{d}{dx}f_n(x)\}$ がある関数に I で一様収束し，かつ $\{f_n(x)\}$ は一点 $x_0\in I$ で収束しているとする．この時，$\{f_n(x)\}$ は I のすべての点で収束し，連続的微分可能（C^1-級）で

$$\lim_{n\to\infty}\left[\frac{d}{dx}f_n(x)\right]=\frac{d}{dx}\left[\lim_{n\to\infty}f_n(x)\right]=\frac{d}{dx}f(x).$$

(ii) $f_n(x)$ は I で連続的微分可能（C^1-級），$\sum_n\frac{d}{dx}f_n(x)$ がある関数に I で一様収束し，かつ $\sum_n f_n(x)$ は一点 $x_0\in I$ で収束しているとする．この時，$\sum_n f_n(x)$ は I のすべての点で収束し，連続的微分可能（C^1-級）で

$$\sum_{n=0}^{\infty}\left[\frac{d}{dx}f_n(x)\right]=\frac{d}{dx}\left[\sum_{n=0}^{\infty}f_n(x)\right].$$

(iii) べき級数 $\sum_n a_n x^n$ の収束半径を R とすると，$f(x) \equiv \sum_n a_n x^n$ は $(-R, R)$ で微分可能で，その導関数は

$$\frac{d}{dx} f(x) = \sum_{n=1}^{\infty} n\, a_n x^{n-1}$$

を満たす.

6.5.4 積分と微分の順序交換

最後に，積分と微分を考えよう．とは言っても，同じ変数で微分・積分をするのは互いに逆演算なだけであるので，領域 $R \equiv \{(x,y) | x \in [a,b], y \in [c,d]\}$ で定義された 2 変数の関数 $f(x,y)$ を問題にしよう．この $f(x,y)$ を x だけで積分すると，結果は y の関数になる：

$$I(y) \equiv \int_a^b f(x,y) dx.$$

この $I(y)$ を y で微分するとどうなるか？

定理 6.5.8 (積分下の微分) 関数 $f(x,y)$ が領域 R で定義されていて，各 $y \in [c,d]$ に対して上の積分 $I(y)$ が存在するものとする．更に，この領域 R で $\dfrac{\partial}{\partial y} f(x,y)$ が連続であるとする．この時，$y \in [c,d]$ に対して $I(y)$ は微分可能で，その導関数は

$$\frac{d}{dy} I(y) = \frac{d}{dy}\left[\int_a^b f(x,y) dx\right] = \int_a^b \left[\frac{\partial}{\partial y} f(x,y)\right] dx$$

つまり，積分記号の下で y で微分して良い.

6.5.5 最後に：このような順序交換はなぜ大事なのか？

「数学概論」でも強調されていると思うが，我々が扱わなければならない関数は非常に多種多様であり，大抵のものは何らかの級数としてしか表せないことが多い．そのような訳のわからない関数に対しては，当然，その微分や積分なども良くわからない.

良くわからないけども，級数の形で書けている関数に対しては，級数の各項を微分・積分する事で形式的に微分や積分を行う事が可能だ．つまり，

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \qquad \text{ならば} \qquad f'(x) = \sum_{n=1}^{\infty} a_n\, n x^{n-1} \qquad (??) \tag{6.5.22}$$

でも，これが本当に正しいかどうかはわからない．左辺は級数の和をとった後で微分，右辺は微分してから和をとっているので，正に「微分と極限（和）」の交換をしているからだ.

定理 6.5.2，系 6.5.3 と定理 6.5.7 から，べき級数に関しては，その収束半径の中では極限，微分，積分の交換を勝手にやって良いことがわかる．このおかげで，冪級数の方法は非常に強力なものとなる.（例：ある関数の積分を求めたい時，非積分関数を冪級数に展開してから，項別に積分すればよい.）この考えが形をなしてきたのは，Newton, Leibnitz の頃からである．彼らが解析学の祖と言われる所以である.

（ついでに）今のところ，関数の引数 x は（暗黙のうちに）実数と仮定している．しかし，これを複素数に拡張し，その際に自然な「微分可能性」の定義を考えていくと，非常に面白い事が見えてくる．特に，一階微分できれば何回でも微分できる，とか,（その結果として）微分可能な関数はベキ級数に展開できる，とか,（そのベキ級数にこの節の定理を用いて）項別微分，項別積分などがやり放題になるとか... このように大変に奇麗な理論が存在するのだが，それは「複素関数論」でじっくりと習う事になるだろう.（なお，この段落に関しては細かい条件を落として書いているので，少し数学的に不正確なところがあるので，ご注意.）

7 偏微分の続き

偏微分については最初に少しやった．これからその応用（極値問題）を中心にして学習する．教科書には該当部分がないので，春学期の最初に配ったプリントを参照されたい．(僕のプリントもかなり詳しくするが.)

記号を思い出そう： 春学期と同じく，ここでは n-変数の関数を考える（ただし，大抵は $n=2$ に限定する）．関数の引数は n-成分あるから，これを n-成分のベクトルだと思って $\boldsymbol{x}=(x_1,x_2,x_3,\ldots,x_n)$ と書き，n-成分のベクトルの空間を $\mathbb{R}^n$ と書く．また，$\boldsymbol{x}$ の長さを $\|\boldsymbol{x}\|=\left(x_1^2+x_2^2\ +x_3^2+\ldots+x_n^2\right)^{1/2}$ と定義し，２つの点（ベクトル）$\boldsymbol{x},\boldsymbol{y}$ の距離は $\|\boldsymbol{x}-\boldsymbol{y}\|$ と決める．いうまでもなく，$n=1$（１変数関数）の場合は $\boldsymbol{x}=x$ であり，$\|\boldsymbol{x}\|=|x|$（普通の絶対値），$\|\boldsymbol{x}-\boldsymbol{y}\|=|x-y|$ である．一般の n が考えにくい人は，$n=2$ を主に考えるとよい．

7.1 平均値の定理

偏微分を学習した際,「連鎖律」を学んだ．その応用として，多変数の場合の平均値の定理が導かれる．

定理 7.1.1 (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,

$$f(a+h,b+k)=f(a,b)+\frac{\partial f}{\partial x}(a+\theta h,b+\theta k)h+\frac{\partial f}{\partial y}(a+\theta h,b+\theta k)k \tag{7.1.1}$$

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

$$f(\boldsymbol{a}+\boldsymbol{h})=f(\boldsymbol{a})+\sum_{j=1}\frac{\partial f}{\partial x_j}(\boldsymbol{a}+\theta\boldsymbol{h})\,h_j \tag{7.1.2}$$

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

$$f(a+h,b+k)-f(a,b)=g(1)-g(0)=g'(\theta) \tag{7.1.3}$$

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

$$g'(t)=\frac{\partial f}{\partial x}x'(t)+\frac{\partial f}{\partial y}y'(t)=\frac{\partial f}{\partial x}h+\frac{\partial f}{\partial y}k \tag{7.1.4}$$

であるから，定理 7.1.1 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 7.1.2 (平均値の定理) ２変数関数 $f(x,y)$ が全微分可能なら,

$$f(a+h,b+k)=f(a,b)+\frac{\partial f}{\partial x}(a+\theta h,b+\theta k)h+\frac{\partial f}{\partial y}(a+\theta h,b+\theta k)k \tag{7.1.5}$$

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 7.1.1 が定理 1.3.3 から出るのと同じようにして，定理 7.1.2 は定理 1.3.8 から証明される． □

7.2 テイラー展開

春学期は高階の偏導関数を定義した辺りで終わってしまったが，そのあと，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．

テイラー展開． テイラー展開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）

$$g(t) = f(a+th, b+tk) - f(a, b) \tag{7.2.1}$$

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，

$$f(a+h, b+k) = g(1) = \sum_{k=0}^{n-1} \frac{g^{(k)}(0)}{k!} + \frac{g^{(n)}(\theta)}{n!} \qquad (0 < \theta < 1) \tag{7.2.2}$$

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）

$$\frac{d}{dt} = \frac{dx(t)}{dt}\frac{\partial}{\partial x} + \frac{dy(t)}{dt}\frac{\partial}{\partial y} = h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y} \tag{7.2.3}$$

であるから，例えば，f が C^2-級ならば

$$\begin{aligned} g^{(2)}(\theta) &= \frac{d}{dt}\frac{dg}{dt} = \Bigl(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\Bigr)\Bigl(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\Bigr) \\ &= h^2\frac{\partial^2 f}{\partial x^2} + hk\frac{\partial^2 f}{\partial x \partial y} + kh\frac{\partial^2 f}{\partial y \partial x} + k^2\frac{\partial^2 f}{\partial y^2} \end{aligned} \tag{7.2.4}$$

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，

$$\bigl|g^{(2)}(\theta)\bigr| \le 2M(h^2 + k^2) = O(\|\boldsymbol{a} - \boldsymbol{c}\|^2) = o(\|\boldsymbol{a} - \boldsymbol{c}\|) \tag{7.2.5}$$

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，(7.2.2) の $n = 2$ を考えると，

$$f(a+h, b+k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + o(\|\boldsymbol{a} - \boldsymbol{c}\|) \tag{7.2.6}$$

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，

$$f(a+h, b+k) = f(a, b) + f_x(a, b)h + f_y(a, b)k + \frac{1}{2}\bigl[f_{xx}(a, b)h^2 + 2f_{xy}(a, b)hk + f_{yy}(a, b)k^2\bigr] + o(\|\boldsymbol{a} - \boldsymbol{c}\|^2) \tag{7.2.7}$$

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 7.2.1 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

$$f(a+h, b+k) = f(a, b) + \sum_{m=1}^{r-1} \frac{1}{m!} \sum_{\ell=0}^{m} \binom{m}{\ell} \frac{\partial^m f}{\partial x^\ell \, \partial y^{m-\ell}}(a, b)\, h^\ell\, k^{m-\ell} + \frac{1}{r!} \sum_{\ell=0}^{r} \binom{r}{\ell} \frac{\partial^r f}{\partial x^\ell \, \partial y^{r-\ell}}(a+\theta h, b+\theta k)\, h^\ell\, k^{r-\ell} \tag{7.2.8}$$

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．

（問題）演習問題としても出題するつもりだが，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，(7.2.7) 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

7.3 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．

7.3.1 問題の定義

定義 7.3.1 n-成分ベクトルの空間において，上の記号のもとで，

$$B_r(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathbb{R}^n \mid \|\boldsymbol{x} - \boldsymbol{a}\| < r\} \tag{7.3.1}$$

なる集合 $B_r(\boldsymbol{a})$ を $\boldsymbol{a}$ の r-開近傍 という．図では，$\boldsymbol{a}$ 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 7.3.2 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，

$$\exists r > 0, \qquad 0 < \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) < f(\boldsymbol{a}) \tag{7.3.2}$$

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，

$$\exists r > 0, \qquad 0 < \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) > f(\boldsymbol{a}) \tag{7.3.3}$$

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 (7.3.2) と (7.3.3) の代わりに

$$\exists r > 0, \qquad \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{7.3.4}$$

$$\exists r > 0, \qquad \|\boldsymbol{x} - \boldsymbol{a}\| < r \quad \Longrightarrow \quad f(\boldsymbol{x}) \ge f(\boldsymbol{a}) \tag{7.3.5}$$

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，

$$f \text{ の定義域に入っているすべての } \boldsymbol{x} \text{ に対して} \quad f(\boldsymbol{x}) \le f(\boldsymbol{a}) \tag{7.3.6}$$

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，(7.3.6) 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.）この節では極大・極小問題に話を限る.

7.3.2 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 7.3.3 $x = a$ 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

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + o(|x-a|^2) \tag{7.3.7}$$

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

$$f(x) = f(a) + f'(a)(x-a) + o(x-a) \tag{7.3.8}$$

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

$$f(x) = f(a) + \frac{f''(a)}{2}(x-a)^2 + o(|x-a|^2) \tag{7.3.9}$$

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］,

$$f(x) = f(a) + \frac{f^{(3)}(a)}{6}(x-a)^3 + \frac{f^{(4)}(a)}{24}(x-a)^4 + \frac{f^{(5)}(a)}{120}(x-a)^5 + o(|x-a|^5) \tag{7.3.10}$$

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.

7.3.3　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 7.3.4　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

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y)=(a,b) \text{ での値} \tag{7.3.11}$$

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

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

すると，

定理 7.3.5　$(x,y) = (a,b)$ 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）．

（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．それは春に配った「教科書」にも書いてあるからここには再現しない．もちろん，その証明が良くわかる人はそれで十分だが，その証明がわかりにくい人は,「なぜこうなのか」を大体でも理解することがまず大切だ（厳密にちゃんとやるのはその後でも良い）．そのために，テイラーの公式を使う理解の仕方を紹介しておこう．

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

$$f(x,y) = f + f_x(x-a) + f_y(y-b) + \frac{1}{2}\Big[f_{xx}(x-a)^2 + 2f_{xy}(x-a)(y-b) + f_{yy}(y-b)^2\Big] + o(\|\boldsymbol{x}-\boldsymbol{a}\|^2) \tag{7.3.12}$$

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」（線形代数でやってる頃かな）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（$A = f_{xx}, B = f_{xy} = f_{yx}, C = f_{yy}$）

$$g(x,y) = A(x-a)^2 + 2B(x-a)(y-b) + C(y-b)^2 \tag{7.3.13}$$

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，

$$g(x,y) = A\Big\{(x-a) + \frac{B}{A}(y-b)\Big\}^2 + \frac{CA - B^2}{A}\,(y-b)^2 \tag{7.3.14}$$

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 7.3.5 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

$$f_x(a,b) = f_y(a,b) = 0 \tag{7.3.15}$$

である <u>必要</u> 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 7.3.4 と定理 7.3.5）．上を満たす (a,b) の一つ一つについて，ヘシアン

$$H(a,b) = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \qquad \text{偏微分は } (x,y) = (a,b) \text{ での値} \tag{7.3.16}$$

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）

7.3.4 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，なかなか難しい．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ならう「行列の対角化」を用いるのが良い．もちろん，以下のやり方は２変数の場合もカバーしている．

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

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

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

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

$$f(\boldsymbol{x}) = f(\boldsymbol{a}) + \sum_{j=1}^{n}(x_j - a_j)\frac{\partial f}{\partial x_j}(\boldsymbol{a}) + \frac{1}{2}\sum_{i,j=1}^{n}(x_i - a_i)(x_j - a_j)\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) + o(\|\boldsymbol{x} - \boldsymbol{a}\|^2) \tag{7.3.17}$$

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，

$$\frac{\partial f}{\partial x_j}(\boldsymbol{a}) = 0 \qquad (j = 1, 2, \ldots, n) \tag{7.3.18}$$

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

$$\sum_{i,j=1}^{n} h_i h_j a_{ij} = {}^t\boldsymbol{h}A\boldsymbol{h} \qquad \text{ここで} \quad h_i = x_i - a_i, \quad a_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}), \tag{7.3.19}$$

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[12]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．廣島さん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，対角化可能である．そこで A を対角化する行列を P と書くと（P は直交行列にとれるので ${}^tPP = P\,{}^tP = E$），

$$B = {}^tP\,A\,P \qquad A = P\,B\,{}^tP \tag{7.3.20}$$

を満たす B が対角行列になる．これを用いると

$$(\boldsymbol{h}, A\boldsymbol{h}) = (\boldsymbol{h}, P\,B\,{}^tP\,\boldsymbol{h}) = ({}^tP\,\boldsymbol{h}, B\,{}^tP\,\boldsymbol{h}) = (\boldsymbol{g}, B\boldsymbol{g}) = \sum_{j=1}^{n}\lambda_j (g_j)^2 \tag{7.3.21}$$

[12] 廣島さんが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．

と書ける（λ_j は A の固有値，$\boldsymbol{g} = {}^t\!P\,\boldsymbol{h}$．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．わかりやすいように $\lambda_1 > 0$ かつ，$\lambda_n < 0$ とする（そうでないなら添字を付け替えれば同じ）．g_1 のみがゼロでない場合（そのような $\boldsymbol{h}$ 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からである．

以上をまとめると，以下の定理になる：

定理 7.3.6 $\boldsymbol{x} = \boldsymbol{a}$ 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）．

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；皆さんの線形代数の教科書の定理 4.3 と系 4.4）．

定理 7.3.7 $n \times n$ 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．

7.4 陰関数定理

さてと，いよいよ「陰関数定理」に入ります．正直，僕はこの項目が大嫌いだ．重要な定理である事は認めるものの，微積の他の題材と異なり，最初は「何が言いたいのかわからない定理」と思い，一旦わかってしまえば今度は「そんなアタリマエの事をやる必要があるのか」と思うだろう（僕自身，一年の時はそう思った）から．と言っていても仕方ないので，やりましょう．すぐの応用としては，この後でやる「ラグランジュの未定乗数法」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 7.4.1 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 <u>零点</u>，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 $y=0$ で 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[13]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 <u>等高線</u> 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．

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．実のところ，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，ということである．定理を述べるためにまず，用語を定義する．

定義 7.4.2 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 <u>特異点</u> という．特異点でない $f(a,b)=0$ となる点は <u>通常点</u> という．

すると，

定理 7.4.3（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，

$$b=\varphi(a) \quad \text{かつ} \qquad (a,b) \ \text{の近傍で} \ \ f(x,\varphi(x))=0 \tag{7.4.1}$$

[13] 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．ただし将来，このような「曲がった」ものなどにも「次元」を定義する事を行うだろう

がなりたつ．更に (a,b) の近傍では

$$\frac{d}{dx}\varphi(x) = -\frac{f_x(x,y)}{f_y(x,y)}\bigg|_{y=\varphi(x)} \tag{7.4.2}$$

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

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

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

（証明の概略）詳しくは教科書の pp.22-26 を参照．Step 1. $\varphi(x)$ を実際につくる．$f(x,y) = 0$ をみたすような y が存在する事，つまり (7.4.1) 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. (7.4.1) 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． □

（ここには阿蘇山の 1:25,000 の地図を貼付けました．予算がないので，スキャナーが買えず，ここには載せられません...）

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 1.5, 1.6 節など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．

7.5 条件付き極値問題：ラグランジュの未定乗数法

今学期（そしてこの科目）最後の主な話題です．
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る．

（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）今までにも意識的に避けてきたのだが，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．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？
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

$$F(x,y,\lambda)=f(x,y)-\lambda g(x,y) \tag{7.5.1}$$

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.

$$0=\frac{\partial F}{\partial x}=\frac{\partial f}{\partial x}(x,y)-\lambda\frac{\partial g}{\partial x}(x,y), \qquad 0=\frac{\partial F}{\partial y}=\frac{\partial f}{\partial y}(x,y)-\lambda\frac{\partial g}{\partial y}(x,y) \tag{7.5.2}$$

$$0=\frac{\partial F}{\partial \lambda}=g(x,y) \tag{7.5.3}$$

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 (7.5.2) と (7.5.3)（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

$$0=4x^3+2\lambda x, \qquad 0=4y^3+2\lambda y, \qquad 0=x^2+y^2-1 \tag{7.5.4}$$

の３つである．これを解くと，

$$\begin{bmatrix}x\\y\\\lambda\end{bmatrix}=\begin{bmatrix}0\\\pm1\\-2\end{bmatrix},\quad\begin{bmatrix}\pm1\\0\\-2\end{bmatrix},\quad\begin{bmatrix}\pm\frac{1}{\sqrt2}\\\pm\frac{1}{\sqrt2}\\-\frac{1}{\sqrt2}\end{bmatrix},\quad\begin{bmatrix}\pm\frac{1}{\sqrt2}\\\mp\frac{1}{\sqrt2}\\-\frac{1}{\sqrt2}\end{bmatrix},\ \text{（ベクトルの中では複合同順）} \tag{7.5.5}$$

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

$$0=h'(x)=f_x+f_y\,\varphi'(x) \tag{7.5.6}$$

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

$$0=\frac{d}{dx}g(x,\varphi(x))=g_x+g_y\,\varphi'(x) \tag{7.5.7}$$

この２つから，

$$\varphi'(x)=-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \tag{7.5.8}$$

が導かれるが，これは見方を変えれば

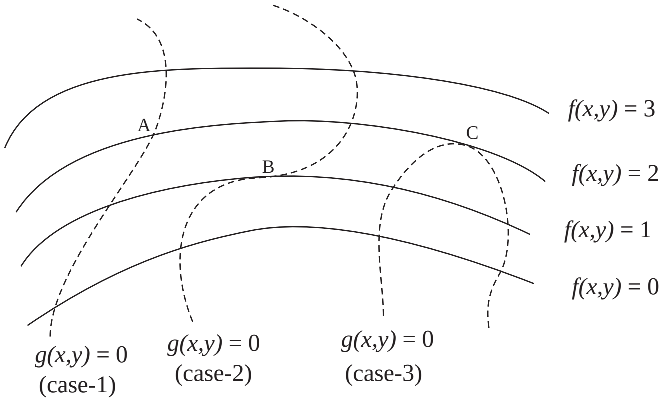
$$\frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{7.5.9}$$

ということであり，この値を λ と書けば，これは (7.5.2) 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 の値を探す事である．

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 事だとわかった．(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 7.4.3 の (7.4.2) で与えられている．つまり

$$f(x,y)=c \text{ の接線の傾きは } -\frac{f_x}{f_y}, \qquad g(x,y)=0 \text{ の接線の傾きは } -\frac{g_x}{g_y} \tag{7.5.10}$$

なのだ．従って，両者が接する条件は

$$-\frac{f_x}{f_y}=-\frac{g_x}{g_y} \qquad \text{つまり} \qquad \frac{g_x}{f_x}=\frac{g_y}{f_y} \tag{7.5.11}$$

であるが，これは (7.5.2) 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

$$F(\boldsymbol{x},\lambda_1,\lambda_2,\ldots,\lambda_m)=f(\boldsymbol{x})-\left\{\lambda_1 g_1(\boldsymbol{x})+\lambda_2 g_2(\boldsymbol{x})+\cdots+\lambda_m g_m(\boldsymbol{x})\right\} \tag{7.5.12}$$

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．

$$\begin{aligned} 0&=\frac{\partial F}{\partial x_j}=\frac{\partial f}{\partial x_j}(\boldsymbol{x})-\sum_{k=1}^{m}\lambda_k\frac{\partial g_k}{\partial x_j}(\boldsymbol{x}), \qquad (j=1,2,\ldots,n)\\ 0&=\frac{\partial F}{\partial \lambda_k}=g_k(\boldsymbol{x}), \qquad (k=1,2,\ldots,m) \end{aligned} \tag{7.5.13}$$

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．

（最後に：この講義を終えるにあたって）

この講義では平均的な学生さんを主な対象としたため，かなりの題材が扱えませんでした．微積の基本的なところは何とか入れようとしたものの，以下のような題材が手薄，または全くやらないままになっています．

- 陰関数定理と逆関数定理．特に逆関数と逆関数定理はやってませんね．
- 曲線，曲面に関すること．接線，接平面など．
- 包絡線.
- 積分の応用としての曲線の長さなど.（これは「微積続論」で少しやるはず.）

これらについては適宜，これから補って下さい．

（講義を終えた感想を書こうかと思ったけど，時間切れになってしまった.）