数ＩＡシケプリ

作成者：pre

平成21 年8 月7 日

目 次

初めに

このシケプリは、演習で既に作ったシケプリに講義のみで扱ったtopic を加えるというコンセプトで作られています。演習で作った方のシケプリは LaTeX の恩恵を受けていないので読みにくいとは思いますが、勘弁願います。

1　実数の定義に関する補足

要点

ここでは**実数を定義し、実数の性質を整理する**ことを話題としています。ざっと内容を説明すると、有理数列の極限として実数を捉え、同値類**（グループ分け**[1]**したときのグループみたいなもの）**の概念を持ち出しつつ、正確な定義をしていきます。性質については単なる整理なので、後の方を御覧あれ。

尚、これから書くことは全ての section について言えることですが、これを一読してからで構わないので、**教科書の練習問題や、プリントの問は解いてください。**此処に在るのは「これは難しいかな？」と拙者が (独断と偏見により) 判断した topic の解説が殆どです。このシケプリが分かったら問題が全て解ける、ということは有り得ないのでご容赦を。

1.1　実数の定義

実数を定義する手掛かりとしてプリントで用いられている性質は、
「有理数列でも無理数に収束する場合がある」
「実数列は必ず実数に収束する」
というものです。これを利用して実数を、

定義 1.1. 実数の集合 $\mathbb{R}$（以下単に $\mathbb{R}$）とは、「有理数列の収束先」の全体である

のように定義することを考えるわけです。ただ、単純な収束を用いて定義してしまうと、**収束の定義**「$\forall\epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |x - \alpha| < \epsilon$」**において α が登場していることから、予め α の属する集合（即ち $\mathbb{R}$）を定義しておく必要に迫られます。**こうなってしまうと、実数を定義するためのプロセスで実数を定義しておかなくてはならなくなり、本末転倒ですね。そのような問題を回避するため、普通の収束の定義の代わりに Cauchy の収束条件を用いるわけです。ただし、Cauchy の収束条件を使う場合は収束先が明確に示されていないので、定義 1.1 をそのまま採用しても定義にはなりません。(「『収束先』って何？」と聞かれても厳密に答えられなくなりますからね。) そこで発想を変えて、

定義 1.2. $\mathbb{R}$ とは、各項が有理数である Cauchy 列そのものである

とするのが第 2 段階の定義になります。しかしまだまだこの定義には欠点があります。この定義によると、例えば $\sqrt{2}$ に収束する異なる数列は、(実際には同一のものとして扱いたいのですが、) 別々の異なるものであるという解釈をせざるを得なくなるのです。**「収束先が同じである有理数列は同じものとしてグループ分**

[1] ここでいうグループとは厳密な意味での group（群）とは別物です。念のため。

けをする」という考え方を持ち出す必要がありますね。
さてそのグループ分けですが、**0 への収束は有理数の範囲内で定義出来る**ことを利用します。詳しく言うと、「２つの有理 Cauchy 列の収束先が同じである」とは、「２つの有理 Cauchy 列の差が 0 に収束する」ということと同じであることに着目するのです。これによりグループ分けの方法は、**「差が 0 に収束する２つの有理 Cauchy 列は全て同じグループに属する」のように定めればよい**ことになります。
また細かい補足ですが、配られたプリントの命題 1.12 で言っていることはどういうことかというと、**「グループの代表としてどいつを選んでも、グループ同士の演算結果がブレない」**ということです。これを確認しないと、演算が定義できないことになりかねず、実数を構成するには不十分となってしまうのです。このことにより演算もめでたく定義出来たので、晴れて実数が定義出来たこととなります。長ったらしく書きますが、要はプリントの定義 1.13 です。やれやれ。

定義 1.3. $\mathbb{R}$ とは、F に属する Cauchy 列を 0 への収束という観点からグループ分けし、その各々のグループを１つの要素とするような集合のことである。[2]

1.2 実数の性質

実数の構成が終わったとはいえ、いちいち Cauchy 列を考えていたら疲れますね。なので、**成り立つ性質を Cauchy 列を用いて証明してしまって、それ以降実数について考える時はその性質を用いてやろう**、ということが次の目標です。ここでは証明そのものよりも、性質を纏めるだけにとどめておきます。

性質 1.4. 任意の有理 Cauchy 列は、「正の数」[3]に収束するもの、0 に収束するもの、「負の数」に収束するもの、の３つのうちどれか１つである

性質 1.5. $\mathbb{R}$ においては順序と積・和が両立する[4]

性質 1.6. $\mathbb{R} - \{0\}$（これは $\mathbb{R}$ から 0 を除いた集合、という意味）の要素には逆数が存在する[5]

性質 1.7. $\mathbb{R}$ はアルキメデス的である（アルキメデスの原理が成立する）

性質 1.8. 項が $\mathbb{R}$ に属する Cauchy 列は収束する。ここで「収束する」とは、収束の定義における収束先 α が、必ず $\mathbb{R}$ に属するという意味である。

とくに最後の１つが、有理数と実数を明確に区別する性質です。尚、これらを証明したい時には、逐一 Cauchy 列に立ち返って証明しようとすれば比較的楽に出来るはずです。なのでこれは省いて、もっと難しい (と個人的に感じた) 問題の解説を次節でしたいと思います。

---

[2] 本当はこの定義は厳密に書けていないのでもっと細かく書きたかったのですが、シケプリというものの性質を考慮して、テキトーにしました。注意。
[3] 「正の数」の正確な表現はプリント p3 の (1) を参照してください。
[4] 両立することの正確な意味はプリント p4 の問 2.5 を参照してください
[5] くどいようですが、正確な表現はプリント p4 の命題 2.6 を参照

1.3 プリント問 1.7 と課題 3.2 の解答

1.3.1 問 1.7 の解答

ポイント:結局は**数列** $\{a_{n+1} - a_n\}$ **の漸化式を作れるかどうかが山場です。** a_n そのものの式をいじって示そうとしても、a_n そのものがわからないから Cauchy の収束条件を用いようとしている、という主旨に逆らうことになるのでうまくいきません。
以下解答。
漸化式より帰納的に $a_n > 1 \quad (n \geq 1$ のとき$) \quad \cdots(1)$ である。ここで、$|a_{n+2} - a_{n+1}|$ を計算してみると、$n \geq 1$ のとき

$$\begin{aligned} |a_{n+2} - a_{n+1}| &= \left| \frac{1}{1 + a_{n+1}} - \frac{1}{1 + a_n} \right| \\ &= \left| \frac{a_{n+1} - a_n}{(1 + a_{n+1})(1 + a_n)} \right| \\ &< \frac{1}{4} |a_{n+1} - a_n| \quad (\because (1)) \end{aligned}$$

であるとわかる。従って、漸化式より $a_1 = \frac{3}{2}$ に注意すると、

$$\begin{aligned} |a_{n+1} - a_n| &= \left(\frac{1}{4}\right)^n |a_1 - a_0| \\ &= \frac{1}{2} \left(\frac{1}{4}\right)^n \quad \cdots(2) \end{aligned}$$

がいえる。次に、任意の正数 ϵに対し或る正の整数 N が存在し、任意の $n \geq N$ に対し ((2) の最右辺) $< \epsilon$となること $\quad \cdots(3)$ を示す。そのためには、同値である

$$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, \frac{1}{\epsilon} < 2 \cdot 4^n \quad \cdots(3)'$$

を示せばよい。ここで、一般に正の整数 N に対し

$$\begin{aligned} 2 \cdot 4^N &= 2(1 + 3)^N \\ &> 2(1 + 3N) \quad (\text{二項定理}) \\ &> 6N \end{aligned}$$

である。一方、アルキメデスの原理 (教科書の公理　) より、任意の正数 ϵ に対し

$$\frac{1}{\epsilon} < 6N$$

なる正の整数 N が存在する。この 2 式より、任意の $\epsilon > 0$ に対し、或る $N \in \mathbb{N}$ が存在して、

$$\frac{1}{\epsilon} < 2 \cdot 4^N$$

が成立する。4^n は n に関して単調に増加するので、(3)' が言え、同値な (3) も成立することがわかった。よって (2),(3) より

$$\forall\epsilon>0, \exists N\in\mathbb{N}, \forall n\geq N, |a_{n+1}-a_n|<\epsilon \quad \cdots(4)$$

を得る。ここで、このような ϵ, N に対し、$m,n\geq N$ のとき

$$\begin{aligned}
|a_m-a_n| &= |(a_m-a_{m-1})+(a_{m-1}-a_{m-2})+\cdots+(a_{n+1}-a_n)| \\
&\leq |a_m-a_{m-1}|+|a_{m-1}-a_{m-2}|+\cdots+|a_{n+1}-a_n| \quad (\text{三角不等式}) \\
&< (m-n)\epsilon \quad (\because (4))
\end{aligned}$$

である。(ここでは $m>n$ の場合のみ書いたが、$m<n$ の場合も同様で、$m=n$ の場合は (上記の最左辺) $=0$ となるので、(最左辺) $<\epsilon$ は自明となる) これは、数列 $\{a_n\}$ が Cauchy 列であることを示している。

1.3.2 課題 3.2 の解答

ポイント:**この問題はポイントがてんこ盛り**なので解答を書いた、という経緯もあって、敢えてポイントをピックアップすることはしません。(というか無理です。) 一貫して「$\Rightarrow$」の証明なので、矢印の左側の仮定をどう使うか考えることがおおまかなコツです。
以下解答。

- 「(c)⇒(R3),(R4)」の証明
  (c) を仮定する。先ず (R3) を示す。数列 $\{a_n\}=\{\frac{c}{n}\}$ (c は任意の正の実数) を考えると、これは単調に減少する下に有界な数列なので、仮定 (c) より $\{a_n\}$ は収束する。その収束先を α とおくと、この条件は

  $$\forall\epsilon>0, \exists N\in\mathbb{N}, \forall n\geq N, \left|\frac{c}{n}-\alpha\right|<\epsilon \quad \cdots(1)$$

  と書ける。ここで $\alpha>0$ を仮定して矛盾を導く。
  (1) で ϵ は任意なので、例えば $\epsilon=\frac{\alpha}{2}$ の場合を考える。$n=N$ 及び $n=3N$ の場合をそれぞれ考えると、前者からは

  $$\frac{1}{2}\alpha<\frac{c}{N}<\frac{3}{2}\alpha$$

  が導かれ、後者からは

  $$\frac{3}{2}\alpha<\frac{c}{N}<\frac{9}{2}\alpha$$

  が導かれる。この 2 条件は矛盾しているので、仮定は誤りで $\alpha\leq 0$ が言える。ところが、$\alpha<0$ を仮定しても同様に、$\epsilon=-\frac{\alpha}{2}$ として $n=N$ の場合

と $n=3N$ の場合を考えると矛盾が導ける。結局、$\alpha=0$ である。
よって (1) は

$$\forall\epsilon>0,\exists N\in\mathbb{N},\forall n\geq N,\frac{c}{n}<\epsilon$$

と書き直せる。ここで $n=N$ の場合を考えて、両辺を N 倍すれば、

$$\forall\epsilon,c>0,\exists N\in\mathbb{N},c<\epsilon N$$

を得る。これはアルキメデスの原理 (即ち公理　或いは (R3) のこと) に他ならない。
次に (R4) を示す。任意の Cauchy 列を代表して $\{b_n\}$ と書く。すると $\{b_n\}$ は

$$\forall\epsilon>0,\exists N\in\mathbb{N},\forall m,n\geq N,|b_m-b_n|<\epsilon \quad \cdots(2)$$

を満たす。ここで、ϵをある値に固定し、$n=N$ の場合を考えると、N 以上の整数 m に対して

$$-\epsilon+b_N<b_m<\epsilon+b_N$$

が成立する。これと $m=1,2,\ldots,N-1$ の場合も加味し、数列 $\{b_n\}$ が有界であること　$\cdots(3)$　が分かる。
次に、n により変化する集合 S_nを、

$$S_n=\{b_n,b_{n+1},\ldots\}$$

により定義し、さらに数列 $\{A_n\},\{B_n\}$ を、

$$A_n=\inf S_n,B_n=\sup S_n$$

により定義する。このとき定義より $A_n\leq b_n\leq B_n \quad \cdots(4)$　であり、A_nは単調に増加し、B_nは単調に減少する。また (3) より A_n,B_nはともに有界である。よって A_n,B_nには仮定(c) を用いることが出来て、A_n,B_nはともに収束する。A_n,B_nの収束先をそれぞれ A,B とおく。
ここで、$\{B_n-A_n\}$ について考えよう。(2) における ϵ,N をとったとき、$n\geq N$ のときは (2) より、

$$|B_n-A_n|<\epsilon$$

が成立する。(B_n,A_nはいずれも $\{b_n\}$ のうち番号が N 以上のもののうちどれかだから。) これは数列 $\{B_n-A_n\}$ が 0 に収束することを示している。一方この数列の収束先は $B-A$ で表されるので、$A=B$　(これを b とおく) が言えた。即ち、$|A_n-b|$ 及び $|B_n-b|$ はいずれも (2) のϵ 未満に押さえられる。この事実及び (4) より、

$$\forall\epsilon>0,\exists N\in\mathbb{N},\forall n\geq N,|b_n-b|<\epsilon$$

がわかり、$\{b_n\}$ は収束する。よって (R4) の成立が言えた。

- 「(R3),(R4)$\Rightarrow$ 教科書の公理　,　」の証明
  (R3) は教科書の公理　のことなので、公理　を示せばよい。(R3),(R4) を仮定する。
  数列 $\{a_n\}, \{b_n\}$ は以下の (あ),(い) を満たす任意の数列とする。

  (あ)　$a_0 \leq a_1 \leq \ldots \leq a_n \leq \ldots \leq b_n \leq \ldots \leq b_1 \leq b_0$
  (い)　$\forall \epsilon > 0, \exists N \in \mathbb{N}, \forall n \geq N, |b_n - a_n| < \epsilon$

  このとき、このような ϵ 及び整数 m, n(但し $m > n \geq N$) に対し、

$$
\begin{aligned}
|b_m - b_n| &= |(b_m - a_m) + (a_m - b_n)| \\
&\leq |b_m - a_m| + |b_n - a_m| \quad (\text{三角不等式}) \\
&= |b_m - a_m| + (b_n - a_m) \quad (\because (\text{あ})) \\
&< |b_m - a_m| + (b_n - a_n) \quad (\because (\text{あ})) \\
&= |b_m - a_m| + |b_n - a_n| \quad (\because (\text{あ})) \\
&< \epsilon + \epsilon \quad (\because (\text{い})) \\
&= 2\epsilon
\end{aligned}
$$

  であるから、数列 $\{b_n\}$ は Cauchy 列である。よって仮定 (R4) より、数列 $\{b_n\}$ は収束する。その収束先を c とおく。　$\cdots(1)$　すると、任意の正の数 ϵ に対し或る N が定まり、$n \geq N$ であれば、

$$
\begin{aligned}
|a_n - c| &= |(a_n - b_n) + (b_n - c)| \\
&\leq |b_n - a_n| + |b_n - c| \quad (\text{三角不等式}) \\
&< \epsilon + \epsilon \quad (\because (\text{い}) \text{ 及び } (1)) \\
&= 2\epsilon \quad \cdots(2)
\end{aligned}
$$

  なので、$\{a_n\}$ も c に収束する。
  ここで、$a_{n_0} > c$ なる n_0 が存在すると仮定すると、(2) でとくに $\epsilon = \frac{1}{4}(a_{n_0} - c)$ とおいた場合に反する。
  実際そのようにおいてみると、(2) より $n \geq \max\{N, n_0\}$ において、

$$
\frac{1}{2}(3c - a_{n_0}) < a_n < \frac{1}{2}(c + a_{n_0})
$$

  であることを導くが、(最右辺)$< a_{n_0}$ より、$\{a_n\}$ が 単調に増加すること (条件 (あ) より) と矛盾することがわかる (から)。
  したがって全ての $n \in \mathbb{N}$ に対して $a_n \leq c$ であると分かる。同様にして全ての $n \in \mathbb{N}$ に対して $c \leq b_n$ であることが導けるので、教科書の公理　が証明された。

- 「教科書の公理　,　⇒(c)」の流れ
  これは教科書の流れそのままなので、解答そのものは略します。教科書見てね。そのかわりにおおよそ教科書の何処を見れば良いのかを押さえます: まずは Weierstraß(英綴り Weierstrass、ドイツ人) の定理 (p.6) を示します。続いて教科書の定理 1.8(p.8) を示せばお終いです。
  Weierstraßの定理の証明におけるポイントは、上界である数と上界でない数の中点をとり、それが上界であるかないかによって数列の作り方が違うところです。次に作った数列に対して公理　を使えることに気がつけば良い。その後の、得られた極限値が上限であることの証明は、ひたすら背理法を用いると上手くいきます。意識すべきは、上限とは最小の上界であること。「最小」も勿論、「上界」も立派な条件ですので、しっかり背理法に組み込みましょう。
  教科書の定理 1.8 を示すポイントはやはり背理法で、決定打は単調性です。証明のエッセンスは大凡上での公理　の証明と同じであると気がつければ、数列の極限については大丈夫だと思います。(但し、このシケプリは独断と偏見に満ちていることを忘れずに！)

2　一変数関数と二変数 (多変数) 関数の違いについての補足

要点

一変数関数に対応する概念があっても、その関連が見えてこない部分もなかなか多いのが厄介なところです。如何にして一変数との対応を見出すかがポイントだと言えるので、**その手助けになるような「解釈」を加えておくのが此処の狙いです。**

2.1　連続性、微分可能性

連続性、微分可能性で嫌というほど出てくる「限りなく近づく」という概念について考えます。一変数のときには「x が a に限りなく近づく」は大凡 $|x-a|<\delta$ で表されました。これは単に「x と a の誤差がδ未満 である」と解釈しても良いのですが、さらに進んだ解釈として、「x と a の**距離**がδ未満である」というものを導入すると、変数が増えてからの概念に対応出来ます。つまり、例えば「$(x,y)\to(a,b)$」とは、大凡「(x,y) と (a,b) の距離が δ未満である」という程の意味になります。

次に、微分に限った話です。但し、偏微分そのものについては簡単な方ですので略します。全微分については、変数が増えると「微分」の意味が掴みづらくなる人もいるかと思います。「二変数関数の『傾き』って何？」という具合に。そのために、一変数の解釈からまず考え直します。(傾き、という解釈も悪くはないけれど。) 一変数関数で、微分可能とは「一次関数 (直線) で近似可能であること」という解釈を習ったはずです。(演習用のシケプリにも書きましたし、演習、講義どちらの授業でも習った。) だから、二変数では**「平面で近似可能であること」**が全微分可能を意味します。その数学的な記述は以下のようになるはずです。

定義 2.1. 関数 $f(x,y)$ が点 (a,b) で微分可能であるとは、

$$f(x,y) = f(a,b) + A(x-a) + B(y-b) + o(\rho(x,y))$$

(但し $\rho(x,y) = \sqrt{(x-a)^2 + (y-b)^2}$(= (点 (x,y) と (a,b) の距離)) であり、A, B は或る定数 (実は一通りに定まる) である。また、o は Landau(英綴り同一、ドイツ人) の記号である。)

の形に書けることである。

この定義で、Landau の記号を書き直せば教科書にある形と全く同一のものになります。教科書の定義を見ると結局、**連続性の判定が出来なければ微分可能性は判定できない**ということが分かります。しっかり練習しましょう。因みに、この定義が実はいかに厳しい条件か、ということは教科書の p.148 にある練習問題を眺めるだけでも分かると思います。どんな方向でも微分可能なのに、全微分は不可能、なんて例もあるくらいですからね。(6 番)

2.2 Chain rule(連鎖律)

プリントではベクトル値関数[6]を用いて記述していますが、その言葉で連鎖律を記述するのは次節にて。今回は教科書の形で書きます。

先ずは定理の形から。いろいろありますが、まずは一番基本的なものを書き、次いで一番一般的なものを書きます。

定理 2.2. $f(x,y)$ は x, y を変数とする関数で、x, y はそれぞれ $x = \phi(s,t), y = \psi(s,t)$ という形で変数 s, t の関数になっているものとする。このとき、合成関数 $f(\phi(s,t), \psi(s,t))$ の s, t による偏微分はそれぞれ次式で与えられる。

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial s}$$

[6]ベクトル値関数とは、関数の出力先が数 1 つではなく、複数の数となっているような関数のことです。要は値域がベクトル空間に属するのだということ。次節で詳説

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial t}$$

定理 2.3. f は変数 $x_1, x_2, \ldots, x_n$ の関数で、これらはそれぞれ変数 $t_1, t_2, \ldots, t_m$ の関数になっているものとする。(但し $m.n$ は正の整数。) このとき、f の t_i による偏微分は次式で与えられる。

$$\frac{\partial f}{\partial t_i} = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}\frac{\partial x_k}{\partial t_i}$$

これを行列表示すると、次のように (m,n) 行列を使って表せる。
(これは上式の理解が出来ていれば暗記せずに書けるので覚えなくて良い。)

$$\begin{pmatrix} \frac{\partial f}{\partial t_1} \\ \frac{\partial f}{\partial t_2} \\ \vdots \\ \frac{\partial f}{\partial t_m} \end{pmatrix} = \begin{pmatrix} \frac{\partial x_1}{\partial t_1} & \frac{\partial x_2}{\partial t_1} & \cdots & \frac{\partial x_n}{\partial t_1} \\ \frac{\partial x_1}{\partial t_2} & \frac{\partial x_2}{\partial t_2} & \cdots & \frac{\partial x_n}{\partial t_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial x_1}{\partial t_m} & \frac{\partial x_2}{\partial t_m} & \cdots & \frac{\partial x_n}{\partial t_m} \end{pmatrix} \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

この定理は次のように見ます：例えば二変数の方に於いて、変数 s を僅かに変化させると、**変化するのは x だけではなくて、y も変化します。**(これが「連鎖」という所以かと思われる。) そこで、大雑把には $\frac{\partial x}{\partial s}$ と $\frac{\partial y}{\partial s}$ が公式の中に含まれることが予想できます。ただ、**実際に変化するのはさらに x,y の変化の影響を受ける f** ですので、$\frac{\partial x}{\partial s}$ には $\frac{\partial f}{\partial x}$ が掛けられている必要があります。y についても同じです。x, y 相互に、直接関数の関係がある訳ではないので、最後にこれら 2 つを足し合わせれば出来上がりです。変数が増えても同じ見方を試みてください。敢えて和の表現や行列の表現を導入したので、その表現に慣れる練習としても。

2.3 ベクトルを定義域、値域とする関数

この節は単なる整理の節であって、不要かも知れません。しかし、連鎖律がプリントではベクトル値関数の言葉で記述されているので、軽く説明をしておきます。

関数の定義域、値域の要素はベクトルでも構いません。例えば、定義域の要素がベクトルになっている関数は、単なる多変数関数です。(変数 n 個を入力することは、ベクトル 1 つを入力する、と解釈可能なので。) これに対し、値域の要素がベクトルになっている関数はどうなのか。これについては、**1 つの値を出力する関数が n 個セットになっている、そのワンセットが関数そのものである**、と解釈すると良いです。(このように、値域の要素がベクトルとなっているような関数をベクトル値関数と呼びます。)

すると、微分係数に関する解釈もそれぞれで異なってきます。2 変数関数では、平面での近似という解釈に基づき、比例定数を 2 つ指定しました。同様に多変数関

数、即ちベクトルを定義域の要素とする関数では、そのベクトルの次元だけの比例係数を指定しなければ、微分という観念が意味を持たなくなるのです。ベクトル値関数では関数が幾つかセットになっているのですから、当然それだけ必要な比例係数の数が増えることは想像に難くないでしょう。例えば、プリントで述べられているような $\mathbb{R}^n$から $\mathbb{R}^m$への 写像 F の微分を考えたとき、ベクトル $(\mathbf{x}-\mathbf{a})$ に (左から) 掛けられる比例定数は mn 個の要素を含む行列であるはずです。このような行列をヤコビ行列と言います。正確な形で書くと下のようになりますね。

定義 2.4. $F(x_1, x_2, \ldots, x_n)$ を、ベクトル値関数とする。(見て分かるように、定義域もまた、ベクトル空間である。) (但し F の第 i 成分を f_iとし、これは値域の要素を実数とする関数である。) このとき、F が定義域内の或る点 $\mathbf{a}$ で微分可能であるということは、以下のように F が記述出来ることである、と定義する。

$$F(\mathbf{x}) = F(\mathbf{a}) + J_F(\mathbf{a})(\mathbf{x}-\mathbf{a})$$

$$\text{但し、}J_F(\mathbf{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \cdots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}\text{であり、}$$

これを F の $\mathbf{a}$ における ヤコビ行列と言う。

最後に連鎖律との関係：合成関数のヤコビ行列は、合成前のヤコビ行列の積で与えられる、というだけの事です。つまり、ベクトル値関数の微分を記述する比例定数の値は、行列積で計算して構わない、ということを述べているのが連鎖律です。連鎖律は 2 つの方法で書きましたが、先ずは「連鎖」のイメージが掴みやすい方で覚えて、残り 1 つの理解に役立てると良いでしょう。

2.4 多変数の Taylor の定理

先ずは変数の形を増やした一般形を書いておきます。教科書にもプリントにも載っていませんが、これで合っている筈。只、基本的には二変数の場合のみ理解しておけば事は足りるので、これは覚えなくて良いです。(Σ を二つ使うのはやりたくなかったのですが、スペースの都合上御免なさい。)

定理 2.5. f は $x_1, x_2, \ldots, x_n$を変数とする関数である。このとき、正の整数 n 及び f の定義域内の点 $(a_1, a_2, \ldots, a_n)$ に対して、次式を成立させるような θ(但し $0 < \theta < 1$) が存在する。

$$f(x_1,x_2,\ldots,x_n)=\left\{\sum_{k=0}^{n-1}\frac{1}{k!}\left(\sum_{i=1}^{n}(x_i-a_i)\frac{\partial}{\partial x_i}\right)^k f(a_1,a_2,\ldots,a_n)\right\}$$
$$+\frac{1}{n!}\left(\sum_{i=0}^{n}(x_i-a_i)\frac{\partial}{\partial x_i}\right)^n f(a_1+\theta(x_1-a_1),a_2+\theta(x_2-a_2),\ldots,a_n+\theta(x_n-a_n))$$

証明そのもののポイント：一変数 t が、$x_i=b_it+a_i\quad(i=1,2,\ldots,n)$ の形で変数 $x_1,x_2,\ldots,x_n$を動かし、最終的に多変数関数 $f(x_1,x_2,\ldots,x_n)$ を動かす、という特殊な場合をまず考えて、それに一変数の Taylor の定理を使います。(Taylor はイギリス人。) この際 a_i,b_iは任意に取れ、また t も任意に動かせることから、**この特殊な場合が全ての場合を代表している**ことを断れば証明が終わりです。

次いで使い方のポイント：この定理は変数の数が増えた分だけ、使い勝手がさらに良くなります。ある点の近傍[7]における関数の挙動を捉える Taylor の定理は、**関数の極大・極小を判断するためには丁度良い道具となっている**からです。具体的には次節 (というか教科書の極値判定に関する部分) を参照。

2.5 極値判定条件

ここでは二変数に限って話を進めます。n 変数の場合はおまけを参照。停留点 ($\frac{\partial f}{\partial x}=\frac{\partial f}{\partial y}=0$ となるような点) が極値かどうかを調べる際には、極値判定条件を満たすかチェックする、という一変数の時とは異なった手法により調べます。一変数のときと異なり、**「前後」の「導関数」の「符合」を調べるという概念が成立しない**からです。ここでは極値の定義に沿って考えなければなりません。(極値の定義に限って一般形で書いておく。)

定義 2.6. 領域で定義されている関数 $f(x_1,x_2,\ldots,x_n)$ が定義域内の点 $A(a_1,a_2,\ldots,a_n)$ で極大値 (又は極小値) をとるとは、あるδが存在して、$0<d(P,A)<\delta$を満たす任意の点 P に対して、$f(P)<f(A)$(又は $f(P)>f(A)$) となることである、と定義する。

この定義を見て分かるように、極大・極小とは本来 1 点のみに注目して判断するものではなく、**周囲の点との比較により定めるものである**ことが大前提です。しかし、この作業を或る程度省略させてくれるのが極値判定条件です：(教科書とは少し異なる形で書きます。)

定理 2.7. $f(x,y)$ は C^2級関数とする。この関数は停留点として点 (a,b) をもつものとする。(即ち、$\frac{\partial f}{\partial x}=\frac{\partial f}{\partial y}=0$ を満たす。) このとき、行列 A を、

$$A=\begin{pmatrix} f_{xx}(a,b) & f_{yx}(a,b)\\ f_{xy}(a,b) & f_{yy}(a,b)\end{pmatrix}$$

[7]正確な意味はさておき、ひとまず近い点のことと思ってくだされば良いです。

で定義すると、以下が分かる。
(あ) $\det(A) > 0$ かつ $\mathrm{tr}(A) > 0$ [8]ならば $f(a,b)$ は極小値である。
(い) $\det(A) > 0$ かつ $\mathrm{tr}(A) > 0$ ならば $f(a,b)$ は極大値である。
(う) $\det(A) < 0$ ならば $f(a,b)$ は極値ではない。

証明は教科書 (p158〜p161) 又は授業プリントに任せます。もしくは一般の次元での証明 (おまけでやっている) を見てくださってもいいかと。この条件では、**或る一点の情報のみに基づいて極大又は極小が分かる**ことに注意してください。さっきの大前提からするとかなり使い勝手の良い定理になっていることが分かります。しかし、自ずと限界はあって、**(あ)〜(う) で等号が成立する場合については何もわからない**ことに注意する必要があります。等号が成立する場合は、この定理を使うことはキッパリ諦めて、極値の定義に立ち戻って考えましょう。ただ、そう言われても何をすべきかは一概には分かりません。例えば、停留点を含む平面で切断して一変数で考えてみるとよいかもしれません。(それで解決出来ないこともあるが、そのような難しい問題にはテストでお目にかからないことを祈りましょう。)

3　級数、巾級数についての補足

要点

「積分」という言葉や「一様収束」という言葉をまだ授業で定義していないので、教科書に載っている正確な証明を問われる可能性は低い、と**個人的には**思います。したがって、特に巾級数 (べき級数、と読みます) については授業プリントを中心に復習するとよいと思われます。

3.1　級数の扱い

級数に関しては細かい話が沢山あるので、ちょびちょび触れていきます。この節では勉強内容を確認する程度のことしかしません。

ゼータ関数 (教科書 p238) について：収束か発散か、を聞かれたら、方針は「積分を用いて有界性を示す、ないしは発散することを示す」です。(要点で触れたように出る可能性は低いと思いますが…)

級数の収束判定法について：Cauchy の判定法と d'Alembert の判定法があります。(どちらもフランス人。多分綴りは英仏で同じ。) これの証明に類似したことを実は演習の授業中にやっています。ので、教科書にも問題の解答として証明が載っていることですし、証明は割愛します。

[8] ここで、行列式、トレースを用いて記述することが一般の次元に於いて本質的かどうかは私には分かりかねます、申し訳ない。誰か分かる方は御一報下さい。

絶対収束の重要性について：絶対収束する級数に関しては教科書に様々な定理があります。比較的結論がスッキリしているので、各自整理しましょう。但し、特に教科書の定理 6.9(p.241) は重要だと思います。

級数の証明で心がけること：こういうの書かないとシケプリになりませんよね。級数は Σ 記号が付いているのでそれだけで見通しが悪くなることもあります。$S_n = \sum_{k=0}^{n} a_k$ などと置いて数列として扱う側面と、そのまま級数として扱う側面の二つを区別しましょう。

3.2 巾級数の収束半径

要点でも述べたように、基本的にプリントベースで学習します。ここでは巾級数のうち特に重要な収束半径について考えます。(テストで出せそう、と私が思ったのが此処くらいなので…) 収束半径とは、大凡「この巾級数は変数がどの範囲にあれば収束するか」を指し示すものです。収束半径について重要な点は、一意に定まること、及び**収束半径を求める Cauchy と d'Alembert の方法があること**です。なぜ重要か：収束半径そのものよりも、これを用いればすぐさま収束する限界が分かる上、その限界のもとでいろいろな近似が行えるからです。また、Euler の公式が導けるのも収束半径という概念が確立しているからです。(Euler はスイス人、つまりこれドイツ語読み。) (Euler の公式はどちらかというと絶対収束に関する定理により成立するのですが、それは気にしない。)

勉強に際しては、どちらの判定法に於いても 収束半径が $\frac{1}{A}$ で与えられることに注意してくださいね。

4 おまけ

要点

この節は数ＩＡには関係ない事柄を多く含むので、おまけ、としてあります。従って発展的な内容であり、太字で強調することもしていません。余力のある人のみどうぞ。変分法に至っては全く数ＩＡには関係がなく、物理でやった数学っぽい topic なので載せただけです。

4.1 教科書の定理 4.17(多変数関数一般に於ける極値判定条件) の略証

証明ではなく略証です。しかし無いよりは在った方が良いと思い執筆した次第。

まずは補題を。証明は後者に限り力学の授業でチラッと述べられたくらいの略証を書きます。数学的には厳密性に欠けますが…[9]

補題 4.1. 実対称行列の固有値は全て実数である。

補題 4.2. 実対称行列の固有ベクトルはどの 2 つを取っても直交する

略証：実対称行列を A とし、任意に取った 2 つの異なる固有ベクトルをそれぞれ縦ベクトルで $\mathbf{x_i}, \mathbf{x_j}$ とし、対応する固有値をそれぞれ λ_i, λ_j とする。このとき、

$$A\mathbf{x_i} = \lambda_i \mathbf{x_i} \quad \cdots(1)$$

および添え字を j に変えて transpose した、

$$\mathbf{x_j}^t A^t = \lambda_j \mathbf{x_j}^t \quad \cdots(2)$$

が成立する。(1) の左辺に左から $\mathbf{x_j}^t$ を掛けたものと、(2) の左辺に右から $\mathbf{x_i}$ を掛けたものは同一なので、($\because A$ は実対称行列より $A^t = A$)

$$(\lambda_i - \lambda_j)\mathbf{x_j}^t \mathbf{x_i} = 0$$

を得る。この式より $\lambda_i \neq \lambda_j$ の時は $\mathbf{x_j}$ と $\mathbf{x_i}$が直交することが分かる。(本当は$\lambda_i = \lambda_j$ の時でも直交するが、その場合の証明は 略します。[10])

ここからが教科書の定理の略証になります。

$f(x_1, x_2, \ldots, x_n)$ は停留点 $(a_1, a_2, \ldots, a_n)$ を持つものとする。(以下、縦ベクトルで $\mathbf{x} = (x_1, x_2 \ldots, x_n)$, $\mathbf{a} = (a_1, a_2, \ldots, a_n)$ と おく。) 多変数の Taylor の定理より、

$$f(\mathbf{x}) - f(\mathbf{a}) = \frac{1}{2}\left(\sum_{k=1}^{n}(x_k - a_k)\frac{\partial}{\partial x_k}\right)^2 f(\mathbf{a} + \theta(\mathbf{x} - \mathbf{a}))$$

なる θ(但し $0 < \theta < 1$) が存在する。(停留点周りの展開なので 1 階の偏微分は全て 0 になることに注意。) (ここからの数行はかなり略をしますが、) この右辺のうち $f(\mathbf{a} + \theta(\mathbf{x} - \mathbf{a}))$ の部分を $f(\mathbf{a})$ に置き換えた式を考えると、両者の符合は $\mathbf{a}$ の近傍で一致する。
(より正確には、教科書と同様に、両者の差は $\rho^2 G(\mathbf{x} - \mathbf{a})$ の形で書けるから、任意の$\epsilon > 0$ に対しある$\delta > 0$ が存在し、$0 < |\mathbf{x} - \mathbf{a}| < \delta$を満たす任意の $\mathbf{x}$ に対し $|G(\mathbf{x} - \mathbf{a})| < \epsilon$となる。従って適当に$\epsilon$を定めれば、それに対し定まる$\delta$が定める 範

[9] 偏に私の力不足の所為です、申し訳ない。
[10] というか此処が分かりませんでした…。

囲内において両者の符合が一致する訳です。)
したがって、係数 $\frac{1}{2}$ も無視して、

$$\left(\sum_{k=1}^{n}(x_k-a_k)\frac{\partial}{\partial x_k}\right)^2 f(\mathbf{a})$$

の符合が正ならば f はこの点で極小値をとり、符合が負ならば f はこの点で極大値をとり、正負どちらにも変わりうるならば $f(\mathbf{a})$ は極値にならない ことが分かった。以下ではこれを調べます。この式は、(1,1) 行列を実数と同一視すれば、

$$(\mathbf{x}-\mathbf{a})^t \begin{pmatrix} f_{x_1x_1}(\mathbf{a}) & f_{x_1x_2}(\mathbf{a}) & \cdots & f_{x_1x_n}(\mathbf{a}) \\ f_{x_2x_1}(\mathbf{a}) & f_{x_2x_2}(\mathbf{a}) & \cdots & f_{x_2x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{a}) & f_{x_nx_2}(\mathbf{a}) & \cdots & f_{x_nx_n}(\mathbf{a}) \end{pmatrix} (\mathbf{x}-\mathbf{a})$$

に等しい。各自計算して確かめて欲しいのですが、確かめる際のヒントを少し書いておきます。作用素の和で記述された式は

$$\left(\sum_{k=1}^{n}(x_k-a_k)\frac{\partial}{\partial x_k}\right)\left(\sum_{k=1}^{n}(x_k-a_k)\frac{\partial}{\partial x_k}\right) f(\mathbf{a})$$

とも書ける。一番目の括弧の i 番目と二番目の括弧の j 番目を掛けた項を計算することを考え、その後 i と j に関して和を取ったものが元の式に等しいことが分かる。つまり、初めの式は

$$\sum_{i=0}^{n}\sum_{j=0}^{n}(x_i-a_i)(x_j-a_j)\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})$$

に等しい。行列で記述された式を真面目にかけ算すると、この式が現れる筈です。

符合を確かめる話に戻ります。実は行列で記述された式に含まれる真ん中の (n,n) 行列はヘッセ行列になっています。このヘッセ行列は偏微分の順序が交換できる ($\because$ f は C^2級) ことから、対称行列です。以下では補題 4.1 より、ヘッセ行列の固有値が全て実数であることは保証されていることに注意しましょう。補題 4.2 より、固有ベクトルは長さを自由に取れることに注意すると、n 個の固有ベクトルは正規直交基底をなすように取れる。[11] したがってベクトル $(\mathbf{x}-\mathbf{a})$ は、この基底を用いて新たに $(y_1, y_2, \ldots, y_n)$ のように成分表示出来る。(やはり縦ベクトルを用いる。)

[11] 正規直交基底をなすベクトルの組を取ると、三次元の任意のベクトルが xyz 空間で成分表示できるのと同じ感覚で n 次元のベクトルを扱えるようになる。正確な話を書くとなると線形代数のベクトル空間から学ばなければならない為、流石に割愛させて頂きます。

これより、基底を取り替える前の式

$$(\mathbf{x}-\mathbf{a})^t \begin{pmatrix} f_{x_1x_1}(\mathbf{a}) & f_{x_1x_2}(\mathbf{a}) & \cdots & f_{x_1x_n}(\mathbf{a}) \\ f_{x_2x_1}(\mathbf{a}) & f_{x_2x_2}(\mathbf{a}) & \cdots & f_{x_2x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{a}) & f_{x_nx_2}(\mathbf{a}) & \cdots & f_{x_nx_n}(\mathbf{a}) \end{pmatrix} (\mathbf{x}-\mathbf{a})$$

は、新たな基底を取れば

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} \begin{pmatrix} f_{x_1x_1}(\mathbf{a}) & f_{x_1x_2}(\mathbf{a}) & \cdots & f_{x_1x_n}(\mathbf{a}) \\ f_{x_2x_1}(\mathbf{a}) & f_{x_2x_2}(\mathbf{a}) & \cdots & f_{x_2x_n}(\mathbf{a}) \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1}(\mathbf{a}) & f_{x_nx_2}(\mathbf{a}) & \cdots & f_{x_nx_n}(\mathbf{a}) \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

と書き表され、固有ベクトルの意味を考えればこの式は

$$\begin{pmatrix} y_1 & y_2 & \cdots & y_n \end{pmatrix} \begin{pmatrix} \lambda_1 y_1 \\ \lambda_2 y_2 \\ \vdots \\ \lambda_n y_n \end{pmatrix} = \sum_{k=1}^{n} \lambda_k y_k^2$$

に等しくなる。
(念のためこの部分の解説：右二つの行列を掛け合わせることに関して、ベクトルを一次変換するという解釈が出来る。この行列の積を計算する際、縦ベクトルの上から k 番目の項を一次変換した結果は、積の結果の上から k 番目に書かれる、という一次変換の性質を思い出しましょう。縦ベクトルの上から k 番目の項は、そもそも固有ベクトル $\mathbf{x}_k$ の y_k 倍を 意味しました。これを一次変換するのですから、上から k 番目の項は $\lambda_k y_k$ になる筈ですね。)

ここまでの結果を纏めると、$f(\mathbf{x})-f(\mathbf{a})$ の符合は $\sum_{k=1}^{n} \lambda_k y_k^2$ の符合に等しくなります。従って
(あ) ヘッセ行列の固有値が全て正のとき、この符合は正なので、$f(\mathbf{a})$ は極小値となる。
(い) ヘッセ行列の固有値が全て負のとき、この符合は負なので、$f(\mathbf{a})$ は極大値となる。
(う) ヘッセ行列の固有値が正負いずれも含むとき、(あ) より正の固有値を持つ全ての固有ベクトルが貼る空間に沿って f を切断すれば、f は $\mathbf{a}$ に於いて極小となるが、反対に (い) より負の固有値を持つ全ての固有ベクトルが貼る空間に沿って f を切断すれば、f は $\mathbf{a}$ に於いて極大となる。これは f が $\mathbf{a}$ に於いて $(n+1)$ 次元空間全体では極値をとらないことを示す。

以上より定理が示された。

4.2 変分法

以下のような量 I を想定してみましょう：ある量 I は、x ではなくて $f(x)$ そのものによって変化する 量です。例えば、

$$I = \int_0^1 f(x)dx$$

とおけば、これは x の関数ではありません。(定積分ですから。) ところが、$f(x)$ の形を変えれば I の値が変わるのも分かります。定積分に限らず、関数 $f(x)$ の形により値が変化する量 I のことを 汎関数と呼びます。

これから調べたいのは、$f(x)$ の変化が具体的に汎関数 I にどのような変動を与えるかです。例えば、下のような量を I としましょう。

$$I = \int_a^b g(x, f_1(x), f_1'(x), f_2(x), f_2'(x))dx$$

(但し、g は x の他に関数 f_1, f_2及びその導関数を含む関数である。) さらに、定積分の始点、終点での関数 f_1, f_2 の値は不変であるものとします。このような「端を固定する」条件の下で、定積分を最小にしてみましょう。(力学で描いたグラフを思い出してください。力学では x に相当するのが時間 t であり、関数 f_1, f_2は時間をパラメータ として変化する量に相当しました。)

変化を調べるには、何事に於いても僅かな変化を考えます。ここでは関数 f_1, f_2 の形を調べたいので、関数 f_1, f_2をそれぞれ僅かな量 $\delta f_1, \delta f_2$ だけ変化させることを考えます。そのとき、汎関数 I は $I + \delta I$ だけ変化した、とします。ここで注意すべきは、固定端で考えているので、

$$\delta f_1(a) = \delta f_2(a) = \delta f_1(b) = \delta f_2(b) = 0 \quad \cdots(1)$$

が成立する点です。ここで、次は δg を $\delta f_1, \delta f_2$ を用いて表すことを考えます。これは難しいのでここでは結論のみ書くと、Chain rule と似た形で、

$$\delta g = \frac{\partial g}{\partial f_1}\delta f_1 + \frac{\partial g}{\partial f_1'}\delta f_1' + \frac{\partial g}{\partial f_2}\delta f_2 + \frac{\partial g}{\partial f_2'}\delta f_2'$$

が成立します。また、これも証明なしで認めてしまうのですが、

$$\delta f_1' = \frac{d}{dx}\delta f_1$$

$$\delta f_2' = \frac{d}{dx}\delta f_2$$

が成立します。これらの結果より、

$$\begin{aligned}\delta I &= (I+\delta I)-I \\ &= \int_a^b (g+\delta g)dx - \int_a^b g dx \\ &= \int_a^b \delta g dx \\ &= \int_a^b \left(\frac{\partial g}{\partial f_1}\delta f_1 + \frac{\partial g}{\partial f_1'}\frac{d\delta f_1}{dx} + \frac{\partial g}{\partial f_2}\delta f_2 + \frac{\partial g}{\partial f_2'}\frac{d\delta f_2}{dx} \right) dx\end{aligned}$$

が成立する、と分かります。積分の第二項、第四項について考えると、部分積分より第二項は、

$$\begin{aligned}(\text{第二項}) &= \left[\frac{\partial g}{\partial f_1'}\delta f_1\right]_a^b - \int_a^b \frac{d}{dx}\left(\frac{\partial g}{\partial f_1'}\right)\delta f_1 dx \\ &= -\int_a^b \frac{d}{dx}\left(\frac{\partial g}{\partial f_1'}\right)\delta f_1 dx\end{aligned}$$

となる。((1) より、部分積分の第一項は 0 となることに注意しよう。) 第四項は添え字を変えればよいので、結局

$$\delta I = \int_a^b \left\{ \frac{\partial g}{\partial f_1} - \frac{d}{dx}\left(\frac{\partial g}{\partial f_1'}\right) \right\}\delta f_1 + \left\{ \frac{\partial g}{\partial f_2} - \frac{d}{dx}\left(\frac{\partial g}{\partial f_2'}\right) \right\}\delta f_2 dx$$

が成立することになる。ところで、δf_1及びδf_2 がどのような形であっても δI が 0 になる、という条件が満たされれば I が最小になることに注意しましょう。
(例えば、ある δf_1 をとると $\delta I > 0$ が成立する場合、関数 f_1を $f_1 - \delta f_1$ に取り替えた方が、I の値は小さくなる。) つまり、以下の式が I を最小にする条件となります。

$$\frac{\partial g}{\partial f_1} - \frac{d}{dx}\left(\frac{\partial g}{\partial f_1'}\right) = 0, \frac{\partial g}{\partial f_2} - \frac{d}{dx}\left(\frac{\partial g}{\partial f_2'}\right) = 0$$

力学に於いては此処までのプロセスは省略し、この条件を用いて汎関数の最小値を求めているのです。尚、今回は変化する関数の数を 2 つにしましたが、数が増えても同様に議論すればよいことは直ぐ分かります。

終わりに

最初の方にも書きましたが、何よりも問題を解きましょう。その際、必ず「何をやっていいのか分からない問題」が出てきます。そのような問題にあたった時、教科書を読んで分かればいいのですが、「教科書の定理が何を言っているか今ひとつ分からない」というケースも生じうると思います。そのようなケースに当たった時、このプリントが (少しでも) 役に立つように、と作りました。上手に使ってやってください。(多少の趣味が入っているのは赦して下さい。)