

多変数の解析学

榎田洋介 *

* このノートはあくまで私的な勉強会の成果物を公開するものなので、私の所属などについては記しません。もし何らかの事情でそれらの情報をお探しの場合は <http://ehlfi.nifflh.net> をご覧ください。

まえがき

これは Michael Spivak “Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus” (CRC Press, 1971) および齋藤正彦による訳書「スピヴァック 多変数の解析学—古典理論への現代的アプローチ」(東京図書, 2007) に沿って、多変数解析学の勉強会をした(する)ときのノートである。大筋としては Spivak に沿った進行をしているが、詳細は異なる。記法や用語といった細かな点については勿論のこと、証明も私個人の嗜好を優先して書いたものが多い。特に積分は測度論に基づいており、加藤賢悟「測度論的確率論 講義ノート 2016 年版」を参考になっている。勉強会の進行に合わせて内容は随時加筆される。前提知識としては集合と写像の記号に関する初歩的な取り扱い、1 変数実関数の微積分、および行列と行列式の計算ができることを仮定しているが、実際のところは線型写像とその行列表示の概念までは理解していたほうが読みやすいと思う。

より具体的には、多変数関数の微積分の定義からはじめて、つぎの形の Stokes の定理およびその周辺の話題までをカバーすることを最終的な目標にしている。

定理. M を向きづけられた k 次元境界つきコンパクト微分可能多様体、 ω を M 上定義された微分 $k-1$ -形式とすると、

$$\int_M d\omega = \int_{\partial M} \omega$$

がなりたつ。ただし、 ∂M には M から誘導される向きを入れるものとする。特に、 M として \mathbb{R} の閉区間 $[a, b]$ 、 ω として $[a, b]$ 上の 0-形式 (すなわち C^∞ 級関数) F を取り、 F の導関数を f と書くことにすれば、

$$\int_a^b f(x) dx = F(b) - F(a)$$

がなりたつ。

定理の後半のステートメントは「微積分学の基本定理」と呼ばれているものにほかならない。いわゆる現代幾何学の入り口が、実際には初等的な解析で学ぶ概念と地続きになっていることをきちんと理解したいというのがこのノートを書くに至った動機のひとつである。

本文の議論とは関係ないかもしれないが、コメントしておくに値すると思った事柄を「余談」として設けてある。また、本文中のそこかしこに「問」が載っているが、読者に対して解くことを要求しているというよりは、私が証明や説明を省いた事柄がすべて「問」扱いされている、というほうが正確であろう。なお、省いた理由はだいたい「私が書くのをめんどくさがった」ないしは「私が答えを知らない、知りたい」のどちらかである。それゆえ、「～を示せ」という書き方はしていない。そのような立ち位置の「問」ではあるが、解いてもらっても構わないし、いい感じに解けたと思ったらぜひ教えてほしい。また、「問」に Spivak の章末問題を引用することはできるだけ避けてあり、私の手元で解けた問題についてはその解答を定理・命題などの形で本文に書くようにし

た^{*1}.

時に、このノートの趣旨から大幅に逸脱した「問」が掲載されていることがある。そのような「問」は星印 (*) を付した。逸脱度合いがひどいものは二つ星 (**) にした。これらは私が気になっている（が、ちゃんと勉強したことのない）事柄である。これらの「問」はつまみ食いが入らない限り勉強会で扱うことは稀であり、読み飛ばしても本文を読む上では問題ないようにしてある（はずである）。特に、私が解答のあらすじを知っているとは限らないこと、本ノートにおける前提知識などを全く踏まえていないことを注意してほしい。これらの「問」に解答をつけるつもりもとりあえずはないが、どこかで勉強したいと思っていることは事実である。

^{*1} のだが、どれが Spivak の章末問題を解いた結果の命題だったか、ちゃんと記録するのを怠ってしまい、書いた私自身もわからなくなっている。気が向いたら対応付けを調べて明記します。

謝辞

このノートは私が主催した勉強会を経て作られた（作られている）ノートである。勉強会の参加者である千葉直也氏，濱田将樹氏，今井雄毅氏，金子亮也氏，三井讓氏，森下真幸氏^{*2}からは，大小様々な質問を多数頂き，私の説明の不十分なところを補い，また話題を広げてくださった。勉強会の中で出た質問などに対する回答や議論のいくつかはこのノートに取り込まれた^{*3}。厚く感謝申し上げます。なお，本文中に存在する誤りや不適切な記述の全ては，筆者である私の責に帰する。

^{*2} 氏名掲載の同意を頂いた方について，名字のアルファベット順に氏名を掲載しています。

^{*3} 取り込むのを後回しにしていることがらも多々あるが，それらもいずれは取り込みたいと思っています。ほんとだよ。

目次

1	準備	6
1.1	\mathbb{R}^n の線型空間としての構造	6
1.2	\mathbb{R}^n の位相とコンパクト性	9
1.3	\mathbb{R}^n 上の連続関数	15
2	多変数関数の微分	18
2.1	微分の定義と基本的な性質	18
2.2	偏微分	25
2.3	C^∞ 級関数と解析関数	28
2.4	逆関数定理	33
2.5	陰関数定理と未定乗数法	37
3	測度論	44
3.1	測度空間と Lebesgue 測度	44
3.2	可測関数	51
3.3	積分と各種の収束定理	56
3.4	拡張定理の証明	65
4	多変数関数の積分	70
4.1	直積測度と Fubini の定理	70
4.2	1 の分割	74
4.3	変数変換公式	76

1 準備

多変数解析学を本格的に議論するにあたって必要な概念たちを手短に扱う。

1.1 \mathbb{R}^n の線型空間としての構造

定義 1.1. \mathbb{R} の n 組からなる集合のことを \mathbb{R}^n と書き, n 次元 Euclid 空間と呼ぶ. すなわち \mathbb{R}^n は, n 個の実数 x^1, x^2, \dots, x^n を用いて (x^1, x^2, \dots, x^n) と書かれるようなもの全体のことである. \mathbb{R}^n には次のような仕方で加法とスカラー倍が定まる;

$$\begin{aligned}(x^1, x^2, \dots, x^n) + (y^1, y^2, \dots, y^n) &:= (x^1 + y^1, x^2 + y^2, \dots, x^n + y^n) \\ a \cdot (x^1, x^2, \dots, x^n) &:= (ax^1, ax^2, \dots, ax^n)\end{aligned}$$

ただし, $a \in \mathbb{R}$.

通常 \mathbb{R}^n の元は (行列の計算との都合上) x^i たちを縦に並べたものとすることが多い (すなわち $n \times 1$ 行列と同一視することが多い) が, 縦ベクトルは紙幅を取るし, かといって逐一転置^Tを書くのも煩雑なので, 文脈から分かる場合および縦横の差異が問題にならない場合は横に並べて書いた上で転置の記号を省略する. 以下, $x \in \mathbb{R}^n$ に対して, その第 i 成分を x^i と表すことにする.

余談 1.2. ベクトルを細字で書いている理由はなんとなくシンプルでかっこよいのと, LATEX で書く量が減るからである. 数学の風習に合わせているというタテマエをつけられなくもないが, あくまでタテマエの域を出ない.

余談 1.3. ここで添え字を「上付き」にしていることには後々の議論を踏まえた意味がある. この時点で説明することもできるが, 多変数の微積分の範囲に限れば説明の必要はないので, 必要になってからやろうと思う. 言葉だけだしておくことにすると, 「座標変換に対して共変・反変のいずれに振る舞うかを意識して添字の上下を取り替えている». ちなみにべき乗と記号が衝突しているが, そういうものである. 混乱の可能性がある場合は都度注意するつもりである.

\mathbb{R}^n には次の仕方で内積と呼ばれる概念が定まっている;

定義 1.4. $x, y \in \mathbb{R}^n$ に対し,

$$\langle x, y \rangle := \sum_{i=1}^n x^i y^i$$

を対応させる写像 $\langle \cdot, \cdot \rangle: \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ を \mathbb{R}^n の標準内積と呼ぶ. $\|x\| := \sqrt{\langle x, x \rangle}$ を x のノルムと呼ぶ.

余談 1.5. わざわざ「標準」内積という呼び方をすることからわかるように、単に「内積」と呼んだ場合は「標準でない」ものも含めたもう少し広いクラスのことを指すことも多い。しかし、当面このノートでは標準内積以外の内積が登場しないので、誤解の恐れがない場合は標準内積のことを単に内積と呼ぶ。

多様体上の Riemann 計量について議論し始めると標準とは限らない内積がそこかしこに登場することになる。予定は未定だが、このあたりの話を書くことにした場合はその時に詳述しようと思う。

行列と行列式の計算ができることだけを前提に据えていることになっているので、線型空間・線型写像の定義もきちんと述べておくことにする。

定義 1.6. 集合 V が (実) 線型空間であるとは、 V にふたつの写像

- 加法 $V \times V \ni (x, y) \mapsto x + y \in V$
- スカラー倍 $\mathbb{R} \times V \ni (a, x) \mapsto ax \in V$

が定まっていて、以下の条件を充たすことをいう。

- (i) 任意の $x, y, z \in V$ に対して $(x + y) + z = x + (y + z)$
- (ii) 零ベクトルと呼ばれる特別な元 $0_V \in V$ があり、それは任意の $x \in V$ に対して $x + 0_V = x$ を充たす。
- (iii) 任意の $x \in V$ に対して、 $-x$ という元があり、それは $x + (-x) = 0_V$ を充たす。
- (iv) 任意の $x, y \in V$ に対して $x + y = y + x$ 。
- (v) 任意の $x, y \in V$ および $a \in \mathbb{R}$ に対して $a(x + y) = ax + ay$ 。
- (vi) 任意の $x \in V$ および $a, b \in \mathbb{R}$ に対して $(a + b)x = ax + bx$ 。
- (vii) 任意の $x \in V$ および $a, b \in \mathbb{R}$ に対して $a(bx) = (ab)x$ 。
- (viii) 任意の $x \in V$ に対して $1x = x$ 。

\mathbb{R}^n は線型空間の例である。実際には他にもいろいろなものが線型空間を成すが、このノートの主題から外れる例が多くなってしまうので、問にしておく；

問 1.7. 漸化式 $a_{n+2} = 3a_{n+1} - a_n$ を充たす数列の集合は線型空間をなす。

定義 1.8. V, W を線型空間とする。 $T: V \rightarrow W$ が線型写像であるとは、

- 任意の $x, y \in V$ に対して $T(x + y) = T(x) + T(y)$
- 任意の $x \in V$ および $a \in \mathbb{R}$ に対して $T(ax) = aT(x)$

が成り立つことをいう。 T が全単射のとき、 T は線型同型写像と言われる。

問 1.9. A を $m \times n$ 行列とする時, A 倍写像 $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^m$ は線型写像である. 逆に $T: \mathbb{R}^n \rightarrow \mathbb{R}^m$ を線型写像とすると, ある行列 A が存在して $T(x) = Ax$ が成り立つ.

定義 1.10. V を線型空間とする. 有限集合 $S = \{v_1, \dots, v_n\}$ が V の基底であるとは,

- $a_1, a_2, \dots, a_n \in \mathbb{R}$ に対して $\sum_i a_i v_i = 0$ ならば $a_1 = a_2 = \dots = a_n = 0$
- 任意の $v \in V$ に対して, 適当な $b_1, b_2, \dots, b_n \in \mathbb{R}$ が存在して $v = \sum_i b_i v_i$

が成り立つことをいう.

余談 1.11. 無限次元線型空間はこのノートの主題ではないので, 最初から基底は有限集合だとしてしまっている. なので厳密な読み方をすればこのノート上では「無限次元線型空間は基底を持たない」ことになってしまう (が, もちろん実際にはそんなことはなく, Zorn の補題によって任意の線型空間は基底を持つ).

問 1.12. S および T を V の基底とすると, S と T の濃度は等しい. この濃度のことを次元といい, $\dim V$ で表す.

問 1.13. $e_i \in \mathbb{R}^n$ を, 第 i 成分が 1, それ以外が 0 という元とすると, $\{e_1, \dots, e_n\}$ は \mathbb{R}^n の基底になる. これを \mathbb{R}^n の標準基底という.

命題 1.14. V, W を線型空間とし, $\{v_1, \dots, v_n\}$ を V の基底とする. $T: V \rightarrow W$ を線型写像とするとき, T は基底での値 $T(v_1), \dots, T(v_n)$ で特徴づけられる. すなわち, $w_1, \dots, w_n \in W$ とする時, $T(v_i) = w_i$ ($1 \leq i \leq n$) を充たす線型写像はひとつしかない.

証明. 任意の $v \in V$ は $\{v_1, \dots, v_n\}$ を用いて $v = \sum a_i v_i$ とただ一通りに表せる. したがって $T(v) = T(\sum a_i v_i) = \sum a_i T(v_i) = \sum a_i w_i$ となる.*4 ■

定義 1.15. $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$ を線型写像とする. このとき, 前命題より T は \mathbb{R}^m の標準基底での値で特徴づけられる. $T(e_i)$ は \mathbb{R}^n の元なので, これを \mathbb{R}^n の標準基底で展開することができる. すなわち mn 個の実数 $\{a_i^j\}$ があって,

$$T(e_i) = \sum a_i^j e_j$$

と表せる. (j, i) 成分に a_i^j を並べて得られる行列 (a_i^j) のことを, T の標準基底に関する行列表示と呼ぶ.

*4 本当はこの証明は少しいい加減ではあると思う. 「 T_1 および T_2 を, 命題にあるような条件を充たす線型写像とする. すると写像として $T_1 = T_2$ である」のようにやるほうが議論によどみがないと感じる.

問 1.16. 上記の定義は標準基底の性質を使っていないので、より一般化して「一般の基底に関する行列表示」を定義できる。

問 1.17. 「 (j, i) 成分に (a_i^j) を並べて」いる理由は、 A 倍写像 $\mathbb{R}^n \ni x \mapsto Ax \in \mathbb{R}^m$ の標準基底に関する行列表示が A になるようにするためであるので、実際にそうになっている。言い換えると、線型写像 $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$ の行列表示を A とするとき、 A の第 i 列は $T(e_i)$ に等しい。

余談 1.18. (a_{ji}) と書かずに (a_i^j) と書いていることにももちろん意味があるが、ベクトルの場合と同様の理由により、この節では説明しない。

最後に、線型写像の連続性・有界性について述べておく。

補題 1.19. $T: \mathbb{R}^m \rightarrow \mathbb{R}^n$ を線型写像とする。このとき、ある実数 M があって、任意の $h \in V$ に対して $\|T(h)\| \leq M\|h\|$ が成り立つ。

証明. 問 1.9 の結果より、ある行列 $A = (a_j^i)$ があって $T(h) = Ah$ であるとしてよい。 $M := \max \left\{ \|a_j^i\| \right\}$ とすればよい。 ■

問 1.20 (*). 上掲の問は、

$$\sup_{h \neq 0} \frac{\|T(h)\|}{\|h\|} \leq M$$

ということを述べている。この左辺に現れた量は T の作用素ノルムと呼ばれる。作用素ノルムは Euclid 空間上の線型写像のみならず、(無限次元かもしれない) ノルム空間の線型作用素に対しても定式化される。

1.2 \mathbb{R}^n の位相とコンパクト性

\mathbb{R} の閉区間

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

ならびに开区間

$$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$$

は容易に高次元に拡張できる。

定義 1.21. \mathbb{R}^n の閉方体とは、

$$[a^1, b^1] \times [a^2, b^2] \times \cdots \times [a^n, b^n] := \{(x^1, x^2, \dots, x^n) \in \mathbb{R}^n \mid a^i \leq x^i \leq b^i\}$$

なる形の集合のこととする。同様に、 \mathbb{R}^n の開方体とは

$$(a^1, b^1) \times (a^2, b^2) \times \cdots \times (a^n, b^n) := \{(x^1, x^2, \dots, x^n) \in \mathbb{R}^n \mid a^i < x^i < b^i\}$$

なる形の集合のこととする。

この定義を使って、より一般に \mathbb{R}^n の開集合・閉集合という概念が定義できる；

定義 1.22. U を \mathbb{R}^n の部分集合とする。 U の各点 x に対し、 x を含む開方体 O であって、 $O \subset U$ を満たすようなものが存在するとき、 U は \mathbb{R}^n の開集合である、あるいは U は \mathbb{R}^n において開である、という。

定義 1.23. U を \mathbb{R}^n の部分集合とする。 U の補集合 $\mathbb{R}^n \setminus U$ が \mathbb{R}^n において開であるとき、 U は \mathbb{R}^n の閉集合である、あるいは U は \mathbb{R}^n において閉である、という。

定義より直ちに分かる以下の性質は、しかしながら極めて重要である。

定理 1.24. U, V を \mathbb{R}^n の開集合とする。 $U \cap V$ は \mathbb{R}^n の開集合である。

証明. $x \in U \cap V$ を任意にひとつ取る。 U は開集合なので、 $x \in O_U \subset U$ を満たすような開方体 O_U がある。 V も開集合なので、 $x \in O_V \subset V$ を満たすような開方体 O_V がある。したがって $O_U \cap O_V$ が開方体であることが証明できればよい。 $n = 1$ のときは場合分けをすればあきらかである^{*5}し、 $n \geq 2$ の場合は各成分ごとに $n = 1$ の場合の議論が適用できることから従う。 ■

定理 1.25. U を \mathbb{R}^n の開集合とする。 U は開方体の和集合の形に表せる。

証明. 開集合の定義より、 U の各点 x に対して $x \in O_x$ を満たすような開方体 $O_x \subset U$ が存在する。 $U = \bigcup_{x \in U} O_x$ である。 ■

問 1.26. \mathbb{R}^n は \mathbb{R}^n において開である。

余談 1.27. 上記 3 性質は開集合の満たしている性質として「証明」された。しかしながら、この事実を逆手に取り、「上記 3 条件を満たすような集合の族のことを開集合系と呼び、開集合系が与えられた集合を一般的に考えてどこまで性質が復元できるか考えよう」というやりかたがあり得る。そのような「開集合系が与えられた集合」のことを位相空間と呼ぶ。以下で述べる命題の多くは位相空間でそのまま成り立つのだが、そのようなものはできるだけ「位相空間における命題・証明」としても読めるように書くよう努力したかったので、Spivak とは大きく流れを変えて上記の 3 点

^{*5} 訳：きちんと書くのがめんどくさい

を直ちに述べておく構成をとり、証明も Spivak のそれとは大きく変えてあるものがある。

以下、開集合や閉集合の例をいくつか挙げる。

問 1.28. 閉方体は閉集合である。これは頑張ればこのタイミングでも証明できそう。

問 1.29. \mathbb{R}^n の部分集合が有限ならばそれは閉集合である。ここで証明できないことはないがめんどくさい。後でやる Heine-Borel の定理（の逆）を使うのが一番早いと思う。

問 1.30. $x \in \mathbb{R}^n$, $r \in \mathbb{R}$ に対し, $B(x; r) := \{y \in \mathbb{R}^n \mid \|x - y\| < r\}$ と置いて, x を中心とする半径 r の開球という。開球は開集合であるが, これを示すのはこのタイミングでは相当めんどくさい。関数の連続性を開集合の言葉で言い換える方法を学んだあとの方が圧倒的に早く示せると思う。

閉集合の定義は「開集合でないこと」ではない。したがって論理的には「開かつ閉」な集合が存在し得ることになるし、実際存在する。

問 1.31. 例えば \emptyset や \mathbb{R}^n は \mathbb{R}^n において開かつ閉である。

余談 1.32. 「開かつ閉」を示す英単語は “clopen” である。

コンパクト性の定義に入る。

定義 1.33. $A \subset \mathbb{R}^n$ を部分集合とする。 \mathbb{R}^n の部分集合の族 $\mathcal{O} := \{U_\lambda\}_{\lambda \in \Lambda}$ が A を被覆する、あるいは A の被覆であるとは、 $A \subset \bigcup_{\lambda \in \Lambda} U_\lambda$ となることをいう。特に \mathcal{O} の元がすべて開集合のとき、 \mathcal{O} は A の開被覆であるという。

定義 1.34. $A \subset \mathbb{R}^n$ を部分集合とする。 A がコンパクトであるとは、任意の A の開被覆 $\{U_\lambda\}_{\lambda \in \Lambda}$ に対し、そこから適切な有限個 U_1, U_2, \dots, U_n を選ぶことで A の被覆にできることをいう。

問 1.35. 上記の定義で「任意の」がついていることは重要である。上記の定義を「ある A の開被覆 $\{U_\lambda\}_{\lambda \in \Lambda}$ に対し、 \sim 」とした場合、全く意味のない定義になる。

問 1.36. $(0, 1) \subset \mathbb{R}$ はコンパクトではない。

「コンパクト」という概念は初見だとかなりとっつきづらいのだが、 \mathbb{R}^n の場合に限って言えば有界閉集合と同値である。このことを順を追って証明していくが、実数 \mathbb{R} の性質を本質的に使うのはつぎの Heine-Borel の定理のみである。

問 1.37 (*). ということは、以下の議論の大部分は実数 \mathbb{R} に限らない一般の位相空間に対して拡張

できるから、どこまで拡張できるか考えると勉強になるし、位相空間を一通り学んでからこれらの定理群の証明を考えるのも勉強になると思う。私にとっては勉強になった。

定理 1.38 (Heine-Borel). \mathbb{R} の有界閉区間 $[a, b]$ はコンパクトである。

証明. $[a, b]$ の開被覆 \mathcal{O} を任意にひとつとって固定し、 $A = A_{\mathcal{O}} \subset [a, b]$ を、 $[a, x]$ が \mathcal{O} の有限個の元で覆えるような点 x の集合とする。あきらかに $a \in A$ であるので空ではなく、かつ A は上に有界である（ひとつの上界として b が取れる）。したがって実数の完備性より $\sup A$ が存在するから、これを α と置く。 $\alpha \in A$ かつ $\alpha = b$ を示せば良い。

- $\alpha \in A$ を示す。 $\alpha \in [a, b]$ なので、 \mathcal{O} の元 U であって $\alpha \in U$ であるようなものが存在する。このとき、 $y \leq \alpha$ なる数 y であって $y \in A$ かつ $y \in U$ なるものがある^{*6}。 A の定め方より $[a, y]$ は有限個の \mathcal{O} の元で覆うことができる。更に $[y, \alpha]$ は U だけで覆えるから、特に有限個の \mathcal{O} の元で覆うことができる。したがって $[a, \alpha]$ は有限個の \mathcal{O} の元で覆うことができるので、 $\alpha \in A$ である。
- $\alpha = b$ を示すべく、 $\alpha < b$ を仮定する。 $[a, \alpha]$ は有限個の \mathcal{O} の元たちで覆うことができる。更に $\alpha \in U$ となるような開集合 U をひとつ取ると、充分小さい $\varepsilon > 0$ に対して $[\alpha, \alpha + \varepsilon] \subset U$ が成り立つ。したがって $[a, \alpha + \varepsilon]$ が有限個の \mathcal{O} の元で覆えることになるが、これは α の定め方に反する。

■

この定理を高次元に拡張していくことにする。

補題 1.39. $B \subset \mathbb{R}^n$ がコンパクトであるとする。 $x \in \mathbb{R}^m$ に対し、 $\{x\} \times B \subset \mathbb{R}^{m+n}$ はコンパクトである。

証明. まず $\{x\} \times B \subset \mathbb{R}^{m+n}$ の開被覆は開方体のみからなっているとして一般性を失わないのでそうする^{*7}； $\{U_a \times V_a\}$ を $\{x\} \times B$ の開被覆であって、 U_a が \mathbb{R}^m の開方体、 V_a が \mathbb{R}^n の開方体であるようなものとする。 B がコンパクトなので、 $\{V_a\}$ の中から有限個を選べばそれが B の被覆になるから、有限部分被覆を選んで $\{V_{x_i}\}$ とする。 $\{U_{x_i} \times V_{x_i}\}$ は $\{x\} \times B$ の有限部分被覆

^{*6} U は開集合なので、ある $\varepsilon > 0$ が存在して、 $(\alpha - \varepsilon, \alpha) \subset U$ 。この開区間の中に A の元がなかったとすると $\alpha = \sup A \leq \alpha - \varepsilon$ となって $\sup A$ の定め方に反する。したがってこの開区間の中に A の元がひとつ以上あることがわかるから、そのようなものを任意にひとつ取って y とすればよい。この議論では \mathbb{R} の（完備性ではなく）稠密性を使っている。

^{*7} TODO: この議論はおかしいので手書きメモをもとにして直す!!! $A := \{x\}$ と置く。 $A \times B$ の開被覆 $\{X_i\}$ を取る。各 X_i は開集合なので、開方体の和集合で書ける。各開集合を開方体の和集合で表したとき、その表示に現れる開方体を（すべての開集合にわたって）あつめて得られる開被覆を $\{Y_\mu\}$ とする。 $\{Y_\mu\}$ の有限部分被覆があれば、それは $\{X_i\}$ の有限部分被覆にもなっている。この証明を検討すればわかるが、この議論は A が 1 点集合でなくても同様に通る。

である. ■

定理 1.40 (Tychonoff). $A \subset \mathbb{R}^m$ および $B \subset \mathbb{R}^n$ がそれぞれコンパクトであるならば, $A \times B \subset \mathbb{R}^{m+n}$ はコンパクトである.

証明. やはり開被覆は開方体のみからなっているとして一般性を失わないのでそうする; $\mathcal{O} := \{U_a \times V_a\}$ を $A \times B$ の開被覆であって, U_a が \mathbb{R}^m の開方体, V_a が \mathbb{R}^n の開方体であるようなものとする. $x \in A$ をひとつとって固定する. このとき $\{x\} \times B$ はコンパクトで, \mathcal{O} によって被覆されているので, この中から有限個 $\{U_{x_i} \times V_{x_i}\}$ を取り出すことで $\{x\} \times B$ を被覆できる. ところで $U_x := \bigcap U_{x_i}$ は x を含む開集合になるので, $\{U_x \times V_{x_i}\}$ は $\{x\} \times B$ の開被覆になっている. この操作を各 $x \in A$ に対し繰り返すことで, $x \in A$ に対して開集合 $U_x \subset A$ とそれに付随する有限個の開集合 $V_{x_i} \subset B$ を作ることができた. ところで A はコンパクトで, $\{U_x\}$ は A の開被覆なので, 有限部分被覆 $\{U_j\}$ が存在する. $\{U_j \times V_{j_i}\}$ は $A \times B$ の有限部分被覆である. ■

問 1.41 (*). 上記の定理は任意個数の直積まで拡張できる; $\{A_\lambda\}$ がすべてコンパクトの時, 積空間 $\prod A_\lambda$ もまたコンパクトである (こちらも Tychonoff の定理という). 証明には選択公理が必要.

系 1.42. \mathbb{R}^m の閉方体はコンパクトである.

証明. Tychonoff の定理と Heine-Borel の定理を組み合わせると帰納的に従う. ■

系 1.43. $A \subset \mathbb{R}^m$ が \mathbb{R}^m の有界閉集合であるならば, それはコンパクトである.

証明. A は有界なので, A を含む閉方体 B がある. A の開被覆 \mathcal{O} を任意にひとつ取る. A は閉集合なので, その補集合 $A^C := \mathbb{R}^m \setminus A$ は開集合である. したがって $\mathcal{O} \cup \{A^C\}$ は開集合の族であり, 更に \mathbb{R}^m を被覆している. 特に B の開被覆にもなっている. ゆえに $\mathcal{O} \cup \{A^C\}$ から有限被覆 \mathcal{O}_2 を選んで B を被覆できる. $A \subset B$ のだから, これは A の開被覆にもなっている. したがって $\mathcal{O}_2 \setminus \{A^C\}$ が \mathcal{O} から選んだ A の有限部分被覆である. ■

問 1.44 (*). 上記の証明では, 有界閉集合 A を含むようなより大きなコンパクト集合 B を取ってきて A のコンパクト性を示すということをした. この議論はより一般的な枠組みにおいて通る; K をコンパクト位相空間, $F \subset K$ を閉集合とすると, F はコンパクトである.

おしまいに, Heine-Borel の定理の逆を示しておこう. これで「 \mathbb{R}^m においてコンパクトであることと有界閉であることは同値」という主張の証明が完結する.

定理 1.45. \mathbb{R}^m のコンパクト集合は有界閉集合に限られる.

証明. この証明では「開球が開集合である」ことは認めて使う^{*8}. $A \subset \mathbb{R}^m$ をコンパクトとする. このとき, A の被覆であって開方体からなるものを取ると, A は有限個の開方体で被覆されるので, とくに有界である. A が閉であること, すなわち $A^C := \mathbb{R}^m \setminus A$ が開であることを示す. $x \in A^C$ を任意にひとつ取って固定する. $a \in A$ に対して $\varepsilon_a := \|x - a\|/4$ と定めると, $a \neq x$ なので $\varepsilon_a > 0$ である. $\{B(a; \varepsilon_a)\}$ は A の開被覆なので, A のコンパクト性より有限部分被覆 $\{B(i; \varepsilon_i)\}$ が取れる. ところで $V := B(x; \min_i \{\varepsilon_i\})$ は x をふくむ開集合であって, $V \cap A = \emptyset$ すなわち $V \subset A^C$ である. x は任意だったので, A^C は開集合である. ■

問 1.46 (*). 上掲した Heine-Borel の定理の逆の証明は, よくよく検討すれば \mathbb{R} の距離空間としての性質しか使っていないことがわかるので, 距離空間に一般化できる. また, 「コンパクトならば閉」という部分は距離特有の性質も必要なく, 本質的に効いているのは Hausdorff 性 (x と a を分離する開集合が取れること) である.

問 1.47 (*). Heine-Borel の定理は無限次元では成り立たない. 例えば有界数列全体の集合 $\ell^\infty(\mathbb{N})$ に対して $d(a, b) = \sup_{n \in \mathbb{N}} |a_n - b_n|$ で距離を入れると, この距離に関する位相のもとで単位閉球はコンパクトにならない. もっと強く, ノルム空間 V の単位閉球がノルム位相に関してコンパクトになることと, V が有限次元であることは同値である.

問 1.48 ().** V をノルム空間, V^* をその双対空間とすると, V^* の単位閉球は weak-* topology についてコンパクトである (Banach-Alaoglu の定理). 関数解析において複数の異なる位相を使い分けたい気持ちはこのあたりにあるのではないと思われる. ちなみにこの定理は Markov-角谷の定理の証明に使うことしか私は使いみちを知らない. Markov-角谷の不動点定理は角谷の不動点定理の相当な拡張と捉えられているらしく^{*9}, Banach-Tarski-Hausdorff の定理^{*10}の類似が S^1 では起こせないことの証明に使える.

コンパクト性は相当に便利な概念であることは事実である. 本題には関係ないものもあるが, ありがたみをいくつか列挙しておくことにする. 以下しばらく, $K \subset \mathbb{R}^n$ をコンパクト集合とする.

問 1.49. $f: K \rightarrow \mathbb{R}$ を連続関数とすると, f は最大値と最小値を持つ. したがって最適化が機能する. 非コンパクト集合上の関数の最適化に関して言えることは非常に少ない. ちなみにこの事実をここで証明するのはめんどくさい. 少なくともコンパクト性が連続関数によって遺伝することを認めたほうが良い…が, このノートでは閉集合の性質についてあまり詳しく議論していないので,

^{*8} 開球が開集合であることの証明は直接やってもよいし, 連続関数を用いた証明をする場合でもこの定理は使わないので, 循環論法にはならない.

^{*9} 私は角谷の不動点定理も, それがどのように経済学で応用されているのかもよく知らない.

^{*10} 一般に「Banach-Tarski のパラドックス」と呼ばれているもので, 大雑把には「 S^2 を 2 つに分解して再度組み合わせると S^2 が 2 つできる」という定理. ちなみにこの分解で得られる集合は Lebesgue 可測ではないので, 特に物理的に不可能な分解である.

それでも面倒かもしれない。

問 1.50 (*). $f: K \rightarrow \mathbb{R}$ を連続関数とすると、 f は一様連続である。もっと強く Lipschitz 連続である。さらに微分可能であると仮定すると L -平滑である。したがって勾配法による最適化速度に理論的な上界を与えられる。

問 1.51 (*). $C(K) := \{f: K \rightarrow \mathbb{R} \mid f \text{ は連続関数}\}$ と置く。 $f, g \in C(K)$ に対し、 $d(f, g) := \sup_{x \in K} |f(x) - g(x)|$ と置くと、 d に関して $C(K)$ は完備距離空間になる。更に、任意の $f \in C(K)$ と $\varepsilon > 0$ に対して、 $d(p, f) < \varepsilon$ を満たすような多項式 p が存在する (Weierstrass の多項式近似定理)。これを使えば、連続な周期関数が三角関数の有限和でいくらでも精度良く近似できることが示せる。もう少し議論すれば連続関数に Fourier 変換ができることを数学的に厳密に示せる気がするが、やったことがないし、もしかすると牛刀割鶏かもしれない。

問 1.52 ()**. 実際には K は \mathbb{R}^n の部分集合である必要はなく、コンパクト Hausdorff 空間であればよい。更に、 $C(K)$ の元を一様に任意精度で近似することができるような関数のクラスは多項式に限られない。このようなクラスを特徴づける定理として Stone-Weierstrass の定理というものがある； $\mathcal{A} \subset C(K)$ が K のあらゆる 2 点を分離するような部分代数であって、定数関数 1 を含むならば、 \mathcal{A} は距離空間 $(C(K), d)$ において稠密である。

問 1.53 ()**. 前々問の設定を引き継ぐ。 $M \subset C(K)$ に対して以下は同値である。

- M は $C(K)$ において (d の定める位相に関して) 相対コンパクト。
- M は一様有界かつ一様同程度連続。

この同値性は Ascoli-Arzelà の定理として知られており、常微分方程式の解の存在定理の証明や、連続な分布関数を持つ確率変数の族が弱収束する条件の特徴づけなどに用いられる。コンパクト性がなければこのような定理群は当然成り立たない。

1.3 \mathbb{R}^n 上の連続関数

多変数関数に対する連続性の定義は一変数の場合と同様である。一応きちんと定義しておくことにすると、

定義 1.54. $A \subset \mathbb{R}^m$ を部分集合とする。 $f: A \rightarrow \mathbb{R}^n$ が $a \in \mathbb{R}^m$ において連続であるとは、任意の $\varepsilon > 0$ に対し、適切な $\delta > 0$ を選べば、「 $\|x - a\| < \delta$ ならば $\|f(x) - f(a)\| < \varepsilon$ 」が成り立つこと、言い換えれば $\lim_{x \rightarrow a} f(x) = f(a)$ が成り立つことをいう。 f が任意の $a \in A$ において連続であるとき、 f は単に連続であるという。

上記のような、いわゆる ε - δ 論法は必要なときは必要である。特に誤差の評価をより具体的にする必要があるのである場合は上記の定義のほうがよい。他方、実際にはこの定義は（一見すると極限を用いない）別の形に述べ直すことができる。誤差に関する具体的な評価が必要ない場合、たとえば単に連続であることを証明したい場合などは以下に述べる特徴づけのほうが使いやすい。

定理 1.55. $A \subset \mathbb{R}^m$ を部分集合とする。 $f: A \rightarrow \mathbb{R}^n$ に対して以下は同値である。

- f は連続。
- 任意の開集合 $O \subset \mathbb{R}^n$ に対し、 $f^{-1}(O) = V \cap A$ を満たすような開集合 $V \subset \mathbb{R}^m$ が存在する。

証明. 前半から後半を示す。 $O \subset \mathbb{R}^n$ を開集合とする。 $a \in f^{-1}(O)$ を任意にとると、 $f(a) \in O$ である。 O は開集合なので、充分小さい $\varepsilon_{f(a)} > 0$ に対して^{*11} $B(f(a); \varepsilon_{f(a)}) \subset O$ である。 f は A 上定義された連続関数なので、適当な $\delta_{f(a)} > 0$ に対して $f(B(a; \delta_{f(a)}) \cap A) \subset B(f(a); \varepsilon_{f(a)})$ すなわち $B(a; \delta_{f(a)}) \cap A \subset f^{-1}(B(f(a); \varepsilon_{f(a)}))$ が成り立つ。これを $f^{-1}(O)$ の各点に対して一斉に考えることで

$$\left(\bigcup_{a \in f^{-1}(O)} B(a; \delta_{f(a)}) \right) \cap A \subset f^{-1} \left(\bigcup_{a \in f^{-1}(O)} B(f(a); \varepsilon_{f(a)}) \right)$$

を得る。右辺は実際には $f^{-1}(O)$ に等しいし、 $V := \bigcup_{a \in f^{-1}(O)} B(a; \delta_{f(a)})$ は開集合であるから、あとは左辺と右辺が等しいことを示せばよいが、これはあきらかである。

後半から前半を示す。 $a \in A$ を任意にひとつ取る。 $B(f(a); \varepsilon)$ は \mathbb{R}^n の開集合なので、仮定より $f^{-1}(B(f(a); \varepsilon)) = V \cap A$ を満たすような開集合 V がある。 V は a を含む開集合なので、適当な δ が存在して $B(a; \delta) \subset V$ である。したがって $\|x - a\| < \delta$ ならば $\|f(x) - f(a)\| < \varepsilon$ である。 a は任意であったから結論を得る。 ■

系 1.56. $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ に対して以下は同値である。

- f は連続。
- 任意の開集合 $O \subset \mathbb{R}^n$ に対し、 $f^{-1}(O) = V$ は開集合である。

証明. 先ほどの定理において $A = \mathbb{R}^m$ とせよ。 ■

問 1.57. 開球は開集合である。このタイミングで示すのが一番早い。

問 1.58 (*). $M_n(\mathbb{R})$ を $n \times n$ 実行列の全体とする。 $M_n(\mathbb{R})$ は「成分を縦一列に並び替えることで」 \mathbb{R}^{n^2} と同一視できるので、これを用いて $M_n(\mathbb{R})$ に対しても開集合および連続関数の概念が定義で

^{*11} ここは普通であれば添字の $f(a)$ を省いて記述するところであろうが、ここを省いてしまうと直下の改行した式が意味不明になってしまうので、重たくなるのを承知で添字をつけている。

きる.

$$\mathrm{GL}_n(\mathbb{R}) := \{ A \in M_n(\mathbb{R}) \mid \det A \neq 0 \}$$

$$\mathrm{SL}_n(\mathbb{R}) := \{ A \in M_n(\mathbb{R}) \mid \det A = 1 \}$$

をそれぞれ実一般線型群, 実特殊線型群と呼ぶ. $\mathrm{GL}_n(\mathbb{R})$ および $\mathrm{SL}_n(\mathbb{R})$ は $M_n(\mathbb{R})$ においてそれぞれ開集合, 閉集合である.

余談 1.59. 位相空間に対しても連続写像の定義はできるのだが, 位相空間では当然ながら連続性を定義するのに距離やノルムが使えない. その代わりにどうするかというと, 上記の「開集合の引き戻しが開集合」という (Euclid 空間上の関数に対しては「証明」した) 性質を, 逆に定義に据えてしまうのである.

余談 1.60. 上記の「位相空間に対する連続写像」の定義は, 測度論に現れる「可測写像」の定義と酷似している一方, 代数で現れる「準同型」の定義とは違う点がある. おそらくは「逆像を取る」という操作が共通部分を取る操作 \cap および和集合を取る操作 \cup を保存する一方, 「像を取る・写像で送る」という操作はこれらを保存しないからだろうと思っている.

余談 1.61. 中には像の定義よりも逆像の定義のほうが簡単だと主張する人もいる. ある面ではたしかにそうだと思う.

補題 1.62. $A \subset \mathbb{R}^m$ をコンパクト集合, $f: A \rightarrow \mathbb{R}^n$ を連続写像とする. このとき $f(A)$ はコンパクトである.

証明. $f(A)$ の開被覆 $\{U_\lambda\}$ をとる. f は連続なので, 各 $f^{-1}(U_\lambda)$ は \mathbb{R}^m の開集合である. $\{f^{-1}(U_\lambda)\}$ は A を被覆するから, A のコンパクト性よりその中の有限個 $\{f^{-1}(U_i)\}$ が A を覆う. この有限集合に現れた $\{U_i\}$ が $f(A)$ の有限被覆になっている. ■

2 多変数関数の微分

まず「微分とは線型写像によって関数を近似するおこないである」というテーゼに基づいて微分を高次元へと拡張する。この定義は微分可能多様体に対してもほぼそのまま拡張が通るという点で一般性の高い抽象的な定義であるが、他方で実際の計算においては使いづらい。実際には C^1 級という関数クラスにおいては「各変数に関する偏微分を計算することで微分の情報をすべて取り出せる」ということを見る。これゆえ大抵の場合で微分の計算は偏導関数の計算に帰着する。

続くふたつの話題はいずれも議論の詳細まで追いかけるには骨が折れるが、結果は知るに値すると思う。ひとつ目の話題は C^∞ 級関数と解析関数 (C^ω 級関数) の違いについてである。「無限回微分可能であるが Taylor 級数がもとの関数に一致しない」という一見病的な対象は、しかしながら後の理論構築で本質的な役割を果たす。ふたつ目の話題は陰関数定理と逆関数定理である。「行列が正則であることと、その行列が定める線型写像が可逆であることは同値」という線型代数で必ず学ぶ事柄を微積分の文脈に拡張して得られる定理であり、理論上の重要性は言うまでもなく、制限付き極値問題をはじめとする応用を持つという点で実用的な定理である。

2.1 微分の定義と基本的な性質

一変数関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ は、

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} = b$$

なる数 b が存在する時、 a で微分可能であるといった。この定義はもちろん f が多変数関数の場合には（たとえ左辺に対して適切にノルムをとったとしても）意味を持たない。

問 2.1. たとえば $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ に対する $a \in \mathbb{R}^2$ での微分可能性を

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a)\|}{\|h\|} = b$$

なる数 b が存在する時だというふうに定義したとしよう。この定義のもとでは、たとえば $f(x, y) = x + 2y$ がいかなる点でも微分可能にならない。

しかしながら上記の式を

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - bh}{h} = 0$$

と改めてみる。これは $f(a+h)$ を $f(a) + bh$ で近似していると見ることができる。更に進んで、 $\mathbb{R} \ni h \mapsto bh \in \mathbb{R}$ が線型写像であること、すなわち微分の定義は「関数を線型写像で近似している」とみなせることに注目すると高次元化へのひとつの道筋が見えてくる。すなわち、一変数関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ は、

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - \lambda(h)}{h} = 0$$

を充たす線型写像 $\lambda: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が存在する時、 a で微分可能である。この定義ならば高次元化は容易である。

定義 2.2. 関数 $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ は、次を充たす線型写像 $\lambda_a: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が存在する時、 $a \in \mathbb{R}^m$ で微分可能であるといわれる。

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - \lambda_a(h)\|}{\|h\|} = 0$$

命題 2.3. 関数 $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が $a \in \mathbb{R}^m$ で微分可能であるとする。このとき、

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - \lambda_a(h)\|}{\|h\|} = 0$$

を充たすような線型写像 λ_a は一意である。

証明. 上式を充たす線型写像をもう一つ取ってきて μ_a とする。 $h \in \mathbb{R}^m$ として、

$$\|\lambda_a(h) - \mu_a(h)\| = \|(f(a+h) - f(a) - \mu_a(h)) - (f(a+h) - f(a) - \lambda_a(h))\|$$

の両辺を $\|h\|$ で割った上で $h \rightarrow 0$ とする極限を取れば

$$\begin{aligned} \frac{\|\lambda_a(h) - \mu_a(h)\|}{\|h\|} &= \frac{\|(f(a+h) - f(a) - \mu_a(h)) - (f(a+h) - f(a) - \lambda_a(h))\|}{\|h\|} \\ &\leq \frac{\|(f(a+h) - f(a) - \mu_a(h))\|}{\|h\|} + \frac{\|(f(a+h) - f(a) - \lambda_a(h))\|}{\|h\|} \\ &\rightarrow 0 \end{aligned}$$

すなわち $\lim_{h \rightarrow 0} \frac{\|\lambda_a(h) - \mu_a(h)\|}{\|h\|} = 0$ を得る。ところで μ_a と λ_a は線型写像なので、これは

$$\lim_{h \rightarrow 0} \left\| \lambda_a \left(\frac{h}{\|h\|} \right) - \mu_a \left(\frac{h}{\|h\|} \right) \right\| = 0$$

を意味する。 h は任意だったので、実数 t と標準基底 e_i に対して $h = te_i$ として $t \rightarrow 0$ とする極限を取れば $\lambda_a(e_i) = \mu_a(e_i)$ を得る。基底での値が一致しているので、線型写像としても一致する。 ■

定義 2.4. $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が $a \in \mathbb{R}^m$ で微分可能なとき、前命題で言及された一意な線型写像 λ_a のことを df_a と書いて、 f の a における微分という。 df_a の標準基底に関する行列表示を f の a における Jacobi 行列といい、 $J_f(a)$ と書くことにする。ただし、 $m = n = 1$ の場合は 1×1 行列 $J_f(a)$ とスカラー倍写像 df_a を区別するのが煩雑なので、同一視してどちらも df_a で表すことがある。

定義 2.5. $A \subset \mathbb{R}^m$ を開集合とする。 $f: A \rightarrow \mathbb{R}^n$ が $a \in A$ において微分可能なとき、 f は A 上で微分可能であるという。 $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が \mathbb{R}^m 上で微分可能な時、 f は微分可能であるという。微分可能

な関数 $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ に対し、 a を df_a に対応させる写像 $a \mapsto df_a$ のことを df と書いて、 f の微分という。

余談 2.6. 微分の記号を Spivak における $Df(a)$ から df_a に変えたことには一定の意図があるが、このタイミングではきちんとした説明が難しいので大雑把なことだけを言っておく； df という記号は（正式な定義を知らないにしても）微分形式を思い起こさせるであろうが、実際この微分 df は「0-形式 f の外微分によって得られる 1-形式 df 」ともみなせる（ので、同じ記号を使うことに一定の正当性がある）ということを後で述べるつもりである。

他方、Jacobi 行列にあてる記号を $f'(a)$ から $J_f(a)$ に変えたことには前者ほどの意味はなく、行列に $f'(a)$ という記号を割り当てるのがあまり一般的だとは思わなかったから、というだけに過ぎない。文献によっては Jacobi 行列のほうに $Df(a)$ という記号を割り当てているものもあるが、この流儀を取るのは Spivak と併せて読む際に混乱を引き起こすと思ったのでやめた。

結局のところこのノートでは $J_f(a)$ 倍写像のことを df_a と書くことになった。しかし、とくに局所座標を取ってどうこうという議論をする必要のない今の時点では、無理して記号を変える必要はないかもしれない。

問 2.7. これは偏微分を見慣れた読者に向けての、やや先回り気味の注意である^{*12}；微分可能ならば偏微分可能であることは後で見るが、しかしながらあらゆる変数に関して偏微分可能であったとしても（もっと強く、あらゆる方向に対する方向微分が存在したとしても）、ここで定義した意味で微分可能であるとは限らない。次のような例がある； $S^1 := \{x \in \mathbb{R}^2 \mid \|x\| = 1\}$ を単位円周とし、 $g: S^1 \rightarrow \mathbb{R}$ を、 $g(0,1) = g(1,0) = 0$ および $g(-x) = -g(x)$ を満たすような連続関数とする。関数 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ を、

$$f(x) = \begin{cases} \|x\| \cdot g\left(\frac{x}{\|x\|}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

で定める。更に $x \in \mathbb{R}^2$ に対して $\varphi_x: \mathbb{R} \rightarrow \mathbb{R}$ を $\varphi_x(t) := f(tx)$ で定める。このとき、

- φ_x は微分可能である。
- $g \neq 0$ ならば f は $(0,0)$ において微分可能ではない。

基本的な性質を見ておくことにする。

命題 2.8 (合成関数の微分, 合成則). $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が $a \in \mathbb{R}^m$ において微分可能で、 $g: \mathbb{R}^n \rightarrow \mathbb{R}^l$ が $f(a) \in \mathbb{R}^n$ において微分可能であるとする。このとき合成写像 $g \circ f: \mathbb{R}^m \rightarrow \mathbb{R}^l$ は $a \in \mathbb{R}^m$ において

^{*12} TODO: ここで方向微分の定義を導入して、「微分の高次元化」の可能性は複数あり得るという話をしてもよいと思う。Gâteaux 微分と Fréchet 微分に関する余談を入れることも検討する。また、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ に対しては微分を考えるのと接平面を考えるのはだいたい一緒という話も（できれば微分の定義の前後に）差し込んでおきたい。もちろん、その「だいたい一緒」という話をより一般的な状況で述べていくのがこの節の主たる話題のひとつなのだが。

微分可能であり,

$$d(g \circ f)_a = dg_{f(a)} \circ df_a$$

が成り立つ. したがって, f と g がともに微分可能であるならば, $g \circ f$ は微分可能であり,

$$d(g \circ f) = dg \circ df$$

が成り立つ.

証明. 示すべきことは, $h \rightarrow 0$ のもとで

$$C := \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(df_a(h))\|}{\|h\|} \rightarrow 0$$

となることである.

$$\begin{aligned} C &= \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(df_a(h))\|}{\|h\|} \\ &= \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a)) + dg_{f(a)}(f(a+h) - f(a) - df_a(h))\|}{\|h\|} \\ &\leq \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\|}{\|h\|} + \frac{\|dg_{f(a)}(f(a+h) - f(a) - df_a(h))\|}{\|h\|} \\ &=: A + B. \end{aligned}$$

A と B が $h \rightarrow 0$ で 0 に収束することを示せばよい.

補題 1.19 より, ある実数 M があって $B \leq \frac{M\|f(a+h) - f(a) - df_a(h)\|}{\|h\|} \rightarrow 0$ なので B はよい. A を評価する. $\varepsilon > 0$ を任意に取る. ふたたび補題 1.19 より, 任意の $h \in \mathbb{R}^m$ に対して $\|df_a(h)\| \leq N\|h\|$ となるような実数 N がある. 以下, 3 つの評価を使う;

- g は $f(a)$ において微分可能なので,

$$\lim_{y \rightarrow f(a)} \frac{\|g(y) - g(f(a)) - dg_{f(a)}(y - f(a))\|}{\|y - f(a)\|} = 0$$

すなわちある $\delta > 0$ が存在して $\|y - f(a)\| < \delta$ ならば

$$\|g(y) - g(f(a)) - dg_{f(a)}(y - f(a))\| < \frac{\varepsilon}{2N} \|y - f(a)\|.$$

- f は a において微分可能, とくに連続なので,

$$\lim_{h \rightarrow 0} \|f(a+h) - f(a)\| = 0$$

すなわちある $\gamma_1 > 0$ が存在して $\|h\| < \gamma_1$ ならば

$$\|f(a+h) - f(a)\| < \delta.$$

- f は a において微分可能なので,

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - df_a(h)\|}{\|h\|} = 0$$

すなわちある $\gamma_2 > 0$ が存在して $\|h\| < \gamma_2$ ならば

$$\|f(a+h) - f(a) - df_a(h)\| < N\|h\|.$$

$\gamma := \min\{\gamma_1, \gamma_2\}$ とすると, $\|h\| < \gamma$ ならば $\|f(a+h) - f(a)\| < \delta$ なので,

$$\begin{aligned} \|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\| &< \frac{\varepsilon}{2N} \|f(a+h) - f(a)\| \\ &= \frac{\varepsilon}{2N} \|f(a+h) - f(a) - df_a(h) + df_a(h)\| \\ &\leq \frac{\varepsilon}{2N} \|f(a+h) - f(a) - df_a(h)\| + \frac{\varepsilon}{2N} \|df_a(h)\| \\ &< \frac{\varepsilon}{2N} \cdot N\|h\| + \frac{\varepsilon}{2N} \cdot N\|h\| \\ &= \varepsilon\|h\| \end{aligned}$$

すなわち $h \rightarrow 0$ で $A \rightarrow 0$ となる. ■

余談 2.9. 上記証明の A の評価は, 気分としては以下のようなことをしたいのである^{*13};

$$\begin{aligned} A &= \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\|}{\|h\|} \\ &= \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\|}{\|f(a+h) - f(a)\|} \cdot \frac{\|f(a+h) - f(a)\|}{\|h\|} \\ &= \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\|}{\|f(a+h) - f(a)\|} \cdot \frac{\|f(a+h) - f(a) - df_a(h) + df_a(h)\|}{\|h\|} \\ &\leq \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\|}{\|f(a+h) - f(a)\|} \cdot \frac{\|f(a+h) - f(a) - df_a(h)\| + \|df_a(h)\|}{\|h\|} \\ &\leq \frac{\|g(f(a+h)) - g(f(a)) - dg_{f(a)}(f(a+h) - f(a))\|}{\|f(a+h) - f(a)\|} \cdot \frac{\|f(a+h) - f(a) - df_a(h)\| + N\|h\|}{\|h\|} \\ &\rightarrow 0 \cdot (0 + N) = 0. \end{aligned}$$

のだが, この議論は $f(a+h) = f(a)$ となった場合に破綻する. 1次元であれば「 $f(a)$ の充分近くで常に $f(a+h) = f(a)$ となるかそうでないか」を場合分けすればできるかもしれないが, 2次元の場合にそういうことを考えるのはおそらくつらい. なので泣きながら前半のように ε - δ 論法に基づいてきちんと評価する必要があるし, 多分こちらのほうが最終的には近道な気がしている.

問 2.10. $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が a で微分可能であるならば a で連続であることをしれっと使ったが, 証明はしていない.

^{*13} もっと言えば, 最初はこのように証明を書いていて, 後で「あっ. 間違っている」と気づいて慌てて書き直した.

命題 2.11.

- (i) $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が定数関数であるならば, f は微分可能であり, 任意の $a \in \mathbb{R}^m$ に対して $df_a = 0$. ただし右辺の 0 は零写像 $0: \mathbb{R}^m \ni x \mapsto 0 \in \mathbb{R}^n$.
- (ii) $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が線型写像であるならば, f は微分可能であり, 任意の $a \in \mathbb{R}^m$ に対して $df_a = f$.
- (iii) $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が微分可能であることと, 各成分 $f^i: \mathbb{R}^m \rightarrow \mathbb{R}$ が微分可能であることは同値. このとき,

$$df_a(h) = (df_a^1(h), \dots, df_a^n(h))^T$$

が成り立つ. したがって $J_f(x)$ の第 i 行は $J_{f^i}(x)$ に等しい.

- (iv) 関数 $s: \mathbb{R}^2 \rightarrow \mathbb{R}$ を $s(a^1, a^2) := a^1 + a^2$ で定めると s は微分可能であり, 任意の $a \in \mathbb{R}^2$ に対して $ds_a = s$.
- (v) 関数 $p: \mathbb{R}^2 \rightarrow \mathbb{R}$ を $p(a^1, a^2) := a^1 a^2$ で定めると p は微分可能であり, $dp_{(a^1, a^2)}(h^1, h^2) = h^1 a^2 + h^2 a^1$.

証明.

- (i) $\lim_{h \rightarrow 0} \frac{\|f(x+h) - f(x) - 0\|}{\|h\|} = \lim_{h \rightarrow 0} \frac{\|0\|}{\|h\|} = 0$.
- (ii) $\lim_{h \rightarrow 0} \frac{\|f(x+h) - f(x) - f(h)\|}{\|h\|} = \lim_{h \rightarrow 0} \frac{\|f(x) + f(h) - f(x) - f(h)\|}{\|h\|} = \lim_{h \rightarrow 0} \frac{\|0\|}{\|h\|} = 0$.
- (iii) 各 f^i が微分可能であると仮定する. $\lambda_a(h) := (df_a^1(h), \dots, df_a^n(h))^T$ によって線型写像 $\lambda_a: \mathbb{R}^m \rightarrow \mathbb{R}^n$ を定めると,

$$f(a+h) - f(a) - \lambda_a(h) = (f^1(a+h) - f^1(a) - df_a^1(h), \dots, f^n(a+h) - f^n(a) - df_a^n(h))$$

となるので,

$$\lim_{h \rightarrow 0} \frac{\|f(a+h) - f(a) - \lambda_a(h)\|}{\|h\|} \leq \sum_{i=1}^n \lim_{h \rightarrow 0} \frac{\|f^i(a+h) - f^i(a) - df_a^i(h)\|}{\|h\|} = 0$$

より λ_a は f の a における微分である. 逆に f が微分可能であると仮定すると, 各 f^i は f と座標関数 $\pi_i: \mathbb{R}^n \ni x \mapsto x^i \in \mathbb{R}$ の合成 $f^i = \pi_i \circ f$ であり, π_i が線型写像であることから $df^i = d\pi_i \circ df = \pi_i \circ df$ である. したがって $df_a(h) = (df_a^1(h), \dots, df_a^n(h))^T$ を得る.

- (iv) s は線型写像であるからよい.

- (v) $(h^1, h^2) \mapsto h^1 a^2 + h^2 a^1$ が p の微分の定義を充たすことをしめせばよい^{*14}.

$$\frac{\|p(a+h) - p(a) - h^1 a^2 - h^2 a^1\|}{\|h\|} = \frac{|h^1 h^2|}{\sqrt{(h^1)^2 + (h^2)^2}}$$

^{*14} 上付き添字と冪乗の記号の相性が最悪なことがよくわかる証明だと思う. Spivak では h^1, h^2 ではなく h, k を使っていた.

であるが, $|h^1 h^2| \leq (\max\{h^1, h^2\})^2 \leq (h^1)^2 + (h^2)^2$ なので, $|h^1 h^2| / \sqrt{(h^1)^2 + (h^2)^2} \leq \sqrt{(h^1)^2 + (h^2)^2} \rightarrow 0$ となるからよい.

■

問 2.12 ().** ここまでの議論で以下のことがわかった;

- 微分可能な関数 f および g に対して $d(g \circ f) = dg \circ df$ が成り立つ.
- $\text{id}_{\mathbb{R}^m}: \mathbb{R}^m \rightarrow \mathbb{R}^m$ を恒等写像とすると, これは微分可能であって, $d(\text{id}_{\mathbb{R}^m}) = \text{id}_{\mathbb{R}^m}$ が成り立つ.

したがって d はおそらくどこかの圏からどこかの圏への共変関手である. このことをきちんと述べるにはベクトル束に関する議論が必要になる気がする. しかしこのノートでベクトル束の話まで書くのはやりすぎな気もしている. ちゃんと勉強したことがないので, やってみたいくはあるのだが.

系 2.13. $f, g: \mathbb{R}^n \rightarrow \mathbb{R}$ が x で微分可能であるならば, $f + g, fg$ もまた微分可能であって,

$$\begin{aligned} d(f + g)_x &= df_x + dg_x \\ d(fg)_x &= g(x)df_x + f(x)dg_x \end{aligned}$$

が成り立つ. 特に $g(x) \neq 0$ であるならば,

$$d(f/g)_x = \frac{g(x)df_x - f(x)dg_x}{(g(x))^2}.$$

証明. $f + g = s \circ (f, g)$ なので $f + g$ は微分可能であり, $d(f + g)_x = d(s \circ (f, g))_x = ds_{(f(x), g(x))} \circ d(f, g)_x = s \circ (df_x, dg_x) = df_x + dg_x$ となる.

$fg = p \circ (f, g)$ なので fg は微分可能であり, $d(fg)_x = d(p \circ (f, g))_x = dp_{(f(x), g(x))} \circ d(f, g)_x = dp_{(f(x), g(x))} \circ (df_x, dg_x) = g(x)df_x + f(x)dg_x$ となる. 最後の式は多項式の微分を知っていればできると思います. ■

まとめると, 各成分が微分可能な関数の加減乗除ないし合成で書ける関数は微分可能であるということがわかった. しかしながらここまでの知識だけで微分を計算するのは煩雑である. 煩雑であることを見るためだけに例をひとつだけ挙げておくことにする.

例 2.14. $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ を, $f(x, y) = x^y$ で定める. f の微分を計算しよう. 指数関数 $x \mapsto e^x$ を e と書くことにすれば, $f = e \circ (\pi^2 \log \pi^1)$ なので, $df_{(a,b)} = de_{b \log a} \circ d(\pi^2 \log \pi^1)_{(a,b)} = e^{b \log a} \cdot \left[(d\pi^2)_{(a,b)} \log a + b(d \log \pi^1)_{(a,b)} \right] = a^b \cdot \left[\pi^2 \log a + b \frac{d\pi^1_{(a,b)}}{a} \right] = a^b \cdot \left[\pi^2 \log a + \pi^1 \frac{b}{a} \right]$ を得る. Jacobi 行列の形で書けば $J_f(a, b) = (a^b \frac{b}{a}, a^b \log a)$ となる.

(偏微分の計算に慣れている人ならばよくご存知のように,) このやり方で微分を計算することは

まずあり得ない。もっと簡単な方法があるので、次項以降でそれを見ていくことにする。

2.2 偏微分

多変数関数に対して、ひとつの変数以外を固定してその変数に関して微分するという操作を考える。

定義 2.15. $f: \mathbb{R}^n \rightarrow \mathbb{R}$ と $a \in \mathbb{R}^n$ に対し、

$$\lim_{h \rightarrow 0} \frac{f(a^1, \dots, a^i + h, \dots, a^n) - f(a^1, \dots, a^i, \dots, a^n)}{h}$$

が存在する時、 f は第 i 変数に関して（偏）微分可能であるという。このとき、上記の極限の値を $\partial_i f(a)$ と書いて f の a における第 i 偏微分係数と呼ぶ。任意の $a \in \mathbb{R}^n$ に対して $\partial_i f(a)$ が定まる時、関数 $a \mapsto \partial_i f(a)$ が定まるが、これを $\partial_i f$ と書いて f の第 i 偏導関数という。

余談 2.16. Spivak では $D_i f(a)$ となっていた偏微分の記号を $\partial_i f(a)$ に変えたのはそこまで深い意味はない。 D_i だとあまり偏微分という感じがしなかったからという程度の理由である。ところで、微分に倣って $(\partial_i f)_a$ のように書かない理由もとくにない。書いたほうが良いかもしれない。あまりちゃんと検討してないからこうなっている、というだけである。

他方、 f の引数に充てる記号として (x, y) を使うとき、 $\partial f / \partial x$ という記号で $\partial_1 f$ のことを指すような書き方はよく採られるし、何なら私もよく使う。しかしこのノートではとりあえず断りなしに使うつもりはない。こちらには一定の理由がある。ひとつは引数に割り当てる記号をいちいち用意しないと厳密な記述ができないことである。他にもあるが、それらはすべて Spivak にそのまま述べてあり、それを書き写すのが面倒なので気が向いたら書き足します。

定義 2.17. f の第 i 偏導関数が a において第 j 変数に関して偏微分可能なとき、すなわち $\partial_j(\partial_i f)(a)$ が存在する時、これを $\partial_{i,j} f(a)$ と書く。このような関数たちを 2 階偏導関数という。高階偏導関数も同様に定義される。 r 階偏導関数がすべて存在してしかも連続であるような関数は C^r 級関数と呼ばれる。特に任意階数の偏導関数がすべて存在して連続であるような関数は C^∞ 級関数と呼ばれる。

一般には $\partial_{i,j} f = \partial_{j,i} f$ であるとは限らないが、この等式を保証する充分条件としては次のものがある。あとで Fubini の定理を使って証明するので、ここでは証明しない。

定理 2.18. $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が C^2 級関数であるならば、 $\partial_{i,j} f = \partial_{j,i} f$ である。

余談 2.19. この定理はいろいろな呼ばれ方をしているらしい。Wikipedia しらべでは Young の定理とも Clairaut の定理とも Schwarz の定理ともいうとのこと。

例 2.20. $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ を, $f(x, y) = x^y$ で定める. f の偏微分係数を計算してみよう. とはいえ 1 変数関数だと思って微分を計算すればよいので, $\partial_1 f(a, b) = a^b \frac{b}{a}$, $\partial_2 f(a, b) = a^b \log a$ である.

この例と, さきほどの微分の計算例とを比較すれば, 「Jacobi 行列というのは偏導関数を縦横にならべたものに等しいのではないか?」という予想をすることは難しくないだろう. この予想は実際正しい. それゆえ微分可能であることがわかっている場合, 微分の計算は偏微分係数の計算に帰着する.

定理 2.21. $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が a において微分可能であるとする. このとき, f は各変数に関して偏微分可能であり, f の a における Jacobi 行列 $J_f(a)$ の (i, j) 成分は $\partial_j f^i(a)$ である.

証明. 命題 2.11 より, 結局は各成分関数ごとの考察に帰着するため, $n = 1$ の場合を証明すればよい. $h: \mathbb{R} \rightarrow \mathbb{R}^m$ を, $h(x) = (a^1, \dots, a^{j-1}, x, a^{j+1}, \dots, a^m)$ で定めると $f \circ h: \mathbb{R} \rightarrow \mathbb{R}$ は微分可能であるが, 導関数 $d(f \circ h)$ とは $\partial_j f$ のことである. 更に $J_h(x) = e_j$ であるから, $J_{f \circ h}(a^j) = J_f(h(a^j)) \cdot J_h(a^j) = J_f(a) \cdot e_j$ である. つまり $1 \times m$ 行列 $J_f(a)$ の j 列目は $\partial_j f(a)$ に等しい. ■

問 2.22. 上記の証明では合成則を用いて証明したが, 合成則なしでも証明できる. $n = 1$ としてよいのは同様に, f の微分の存在を仮定してよいので, $h = (0, \dots, 0, h^i, 0, \dots, 0)$ として

$$\lim_{h^i \rightarrow 0} \frac{|f(a+h) - f(a) - df_a(h)|}{|h^i|} = 0$$

が成り立つから, ということは

$$\lim_{h^i \rightarrow 0} \frac{f(a+h) - f(a) - df_a(h)}{|h^i|} = 0$$

すなわち

$$\lim_{h^i \rightarrow 0} \frac{f(a+h) - f(a)}{|h^i|} = \lim_{h^i \rightarrow 0} \frac{df_a(h)}{|h^i|} = \lim_{h^i \rightarrow 0} df_a(e_i) = df_a(e_i)$$

が成り立つ. ここまで書いたところで, 埋める行間がなくなって問の体裁すら保てなくなったことに気づいたのですが, 行間を見つけたら埋めてください.

この逆は成り立つとは限らない; 例えば偏微分可能だからといって微分可能とは限らない例をすでに見た. しかしながら大抵の場合において逆も成り立つので, 大抵の場合において微分の計算は(微分可能性の判定まで含めて) 偏導関数の計算に帰着すると思ってよい.

定理 2.23. $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が C^1 級であるならば f は微分可能である.

証明. 命題 2.11 より, 各成分関数の微分可能性を示せば充分であるから, やはり $n = 1$ の場合を証明すれば良い. $f(a+h) - f(a)$ を,

$$\begin{aligned} f(a+h) - f(a) &= f(a^1 + h^1, a^2, \dots, a^n) - f(a^1, a^2, \dots, a^n) \\ &\quad + f(a^1 + h^1, a^2 + h^2, \dots, a^n) - f(a^1 + h^1, a^2, \dots, a^n) \\ &\quad + \dots \\ &\quad + f(a^1 + h^1, a^2 + h^2, \dots, a^{n-1} + h^{n-1}, a^n + h^n) - f(a^1 + h^1, a^2 + h^2, \dots, a^{n-1} + h^{n-1}, a^n) \end{aligned}$$

のように表示しておく. f は各変数に関して偏微分可能なので, 平均値の定理により

$$f(a^1 + h^1, a^2, \dots, a^n) - f(a^1, a^2, \dots, a^n) = h^1 \partial_1 f(b^1, a^2, \dots, a^n)$$

を充たすような $b^1 \in [a^1, a^1 + h^1]$ がある. 同様の仕方で $b^i \in [a^i, a^i + h^i]$ が取れるので取っておくことで,

$$\begin{aligned} f(a+h) - f(a) &= h^1 \partial_1 f(b^1, a^2, \dots, a^n) \\ &\quad + h^2 \partial_2 f(a^1 + h^1, b^2, \dots, a^n) \\ &\quad + \dots \\ &\quad + h^n \partial_n f(a^1 + h^1, a^2 + h^2, \dots, a^{n-1} + h^{n-1}, b^n) \end{aligned}$$

のように表示できる. さて, 線型写像 $\lambda_a: \mathbb{R}^m \rightarrow \mathbb{R}$ を,

$$\lambda_a(h^1, \dots, h^n) = \sum_{i=1}^n h^i \partial_i f(a)$$

で定める. λ_a が f の微分になっていることを示す^{*15}.

$$\begin{aligned} \frac{\|f(a+h) - f(a) - \lambda(h)\|}{\|h\|} &= \frac{\|\sum h^i \partial_i f(a^1 + h^1, \dots, b^i, a^{i+1}, \dots, a^n) - \sum h^i \partial_i f(a)\|}{\|h\|} \\ &\leq \sum_{i=1}^n \frac{|h^i|}{\|h\|} \|\partial_i f(a^1 + h^1, \dots, a^{i-1} + h^{i-1}, b^i, a^{i+1}, \dots, a^n) - \partial_i f(a)\| \end{aligned}$$

なので, $\partial_i f$ の連続性より右辺は $h \rightarrow 0$ で 0 に収束する. ■

問 2.24 (*). 前定理における C^1 級という仮定は, 微分可能であるための必要条件とは限らない.

- 実際 $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が微分可能でも各成分関数 $f^i: \mathbb{R}^m \rightarrow \mathbb{R}$ が C^1 級であるとは限らないらしい. $m = n = 1$ の場合はそれはそうなのだが, そうでない例をまだ検討していない.

^{*15} 突然天から降ってきた式に見えるかもしれないが, 「仮に微分可能だったとすると前定理より微分はこの形になるしかない」というのがわかるので, そこから逆算している.

さて、微分可能性から偏微分係数の存在が合成則なしで示せて、かつ偏微分係数の連続性から微分可能性を言うのに合成則を使っていないので、ここまでの議論で合成則はなくてもよかった。しかしながら合成則は次の形の（特に実用上の計算において）重要な帰結を持つ；

命題 2.25. 各 $1 \leq i \leq n$ に対して $g^i: \mathbb{R}^m \rightarrow \mathbb{R}$ が $a \in \mathbb{R}^m$ において C^1 級であり、 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が $(g^1(a), g^2(a), \dots, g^n(a))$ において微分可能であったと仮定する。このとき

$$F: \mathbb{R}^m \ni x \mapsto f(g^1(x), g^2(x), \dots, g^n(x)) \in \mathbb{R}$$

は $a \in \mathbb{R}^m$ において微分可能であって、

$$\partial_i F(a) = \sum_{j=1}^m \partial_j f(g^1(a), g^2(a), \dots, g^n(a)) \cdot \partial_i g^j(a)$$

が成り立つ。

証明. $g: \mathbb{R}^m \rightarrow \mathbb{R}^n$ を $g(x) := (g^1(x), \dots, g^n(x))$ で定めると、定理 2.23 より g は微分可能であるので、 $F = f \circ g$ も微分可能である。定理 2.21 より F の a における Jacobi 行列 $J_F(a)$ は $1 \times m$ 行列であって $(1, i)$ 成分が $\partial_i f(a)$ に等しいことがわかる。したがって F の Jacobi 行列を計算すればよいが、合成則より $J_F(a) = J_f(g(a)) \cdot J_g(a)$ であって、 $J_f(g(a))$ は

$$(\partial_1 f(g(a)), \partial_2 f(g(a)), \dots, \partial_n f(g(a)))$$

という $1 \times n$ 行列、 $J_g(a)$ は (j, i) 成分が $\partial_i g^j(a)$ であるような $n \times m$ 行列であるから、あとは行列の掛け算をすることで結論を得る。 ■

2.3 C^∞ 級関数と解析関数

C^∞ 級関数の概念はすでに述べたが、これと近い概念として次のものがある；

定義 2.26. C^∞ 級関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ が $a \in \mathbb{R}$ において解析的であるとは、ある a を含む開集合 D が存在して、 D の各点 x で Taylor 級数

$$\sum_{i=0}^{\infty} \frac{f^{(i)}(a)}{i!} (x-a)^i := \lim_{N \rightarrow \infty} \sum_{i=0}^N \frac{f^{(i)}(a)}{i!} (x-a)^i$$

が $f(x)$ に収束することをいう。 f が任意の $a \in \mathbb{R}$ において解析的なとき、 f は解析的である、ないしは C^ω 級関数、解析（的）関数であると言われる。

問 2.27. たとえば多項式や指数関数は解析的であるし、複素解析的な関数の実軸への制限もそうである。

問 2.28. 実際には多変数関数に対しても解析性の概念は定義できるのだが、多重指数の説明を書くのがいやだったのでやめた。

この項の目的は、 C^∞ であるが解析的でない関数を構成することである。この関数はこの時点では「変な例」以上の感想が湧かないと思われるが、実際にはこの関数を 1 の分割の構成で使う（らしい）。

定理 2.29 (隆起関数の存在). \mathbb{R} のコンパクト集合 K と開集合 A を、 $K \subset A \subset \mathbb{R}$ を満たすように取る。このとき、 C^∞ 級関数 $\varphi: \mathbb{R} \rightarrow [0, 1]$ であって、 K 上では $\varphi(x) = 1$ 、 A に含まれるあるコンパクト集合の外では $\varphi(x) = 0$ を満たすようなものが存在する。

問 2.30 (*). この項では隆起関数を具体的に構成する。構成された関数が解析的でないことも見るが、では作り方を工夫すれば解析的になり得るだろうか、という疑問が生じる。実はどのように作ったとしても φ は解析的にはなり得ない。それは大まかには次のようにしてわかる； φ が解析的であったとして、 K を含むような \mathbb{C} 上の適当な開集合まで φ を解析接続する。すると一致の定理から φ は定数関数 1 になって矛盾する。

問 2.31 ()**. ところで 1 の分割は積分の定義とか Riemann 計量の存在を示すのに使うのだが、作り方から解析的であることが許されない。ということは複素多様体上で類似の概念を作ろうと思ってもうまく行かなそうであるということがわかる。事実、ある複素多様体を与えられたとして、それにどのような計量が入るかというのはそんなに自明な問題ではないらしいし、Kähler 多様体の定義も（「Kähler 計量の入った複素多様体」ではなく）「複素多様体であって Kähler 計量を入れられるもの」として構わないらしい。

では隆起関数の構成に入る。大まかな道筋としては、

- まず開集合 $(-1, 1)$ 上で正で、 ± 1 に近づくにつれ 0 に減衰するような関数を作り、これを各軸ごとに組み合わせることで「ある点を含む開方体の内部だけで値を持ち、境界に近づくにつれ減衰するような関数」を作る。
- 上記の関数を「 K 上の適切な有限箇所に配置することで」 K 上で正、 A の境界の「十分手前で」0 に減衰し切るような関数を作る。この「適切な有限箇所」の選定に K のコンパクト性を用いる。
- 上記で作った関数の値をいい感じに調整する。

補題 2.32. 任意の $k \in \mathbb{R}$ に対して $\lim_{h \rightarrow 0} h^{-k} \exp(-h^{-1}) = 0$.

証明.

$$\log(h^{-k} \exp(-h^{-1})) = -k \log h - h^{-1} = -\frac{1 + kh \log h}{h}$$

であり, $\lim_{h \rightarrow 0} h \log h = \lim_{h \rightarrow 0} \log h / (h^{-1}) = \lim_{h \rightarrow 0} (1/h) / (-h^{-2}) = \lim_{h \rightarrow 0} -h = 0$ なので結局 $\lim_{h \rightarrow 0} \log(h^{-k} \exp(-h^{-1})) = -\infty$ すなわち $\lim_{h \rightarrow 0} h^{-k} \exp(-h^{-1}) = 0$ を得る. ■

命題 2.33. 関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ を,

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

で定めると, f は C^∞ 級である. 更に任意の非負整数 i に対して $f^{(i)}(0) = 0$ が成り立つので, f は解析的ではない. ただし $f^{(i)}$ は f の i 階導関数であり, $f^{(0)} = f$ と理解する.

証明. f は原点以外で C^∞ 級であるのは明らかなので, 非自明な点は任意階数の導関数が原点で 0 になり, かつ連続になることである. これを i に関する帰納法によって示す. まず

$$f^{(1)}(0) = \lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h} = \lim_{h \rightarrow 0} h^{-1} \exp(-h^{-2}) = 0$$

なので $f^{(1)}$ は \mathbb{R} 上定義されているとわかる. 更に $x \neq 0$ において $f^{(1)}(x) = 2x^{-3}e^{-x^{-2}}$ なので, 補題 2.32 によって $\lim_{x \rightarrow 0} f^{(1)}(x) = 0$ であるから $f^{(1)}$ は連続関数である. $i \geq 0$ に対して帰納法の仮定より

$$f^{(i+1)}(0) := \lim_{h \rightarrow 0} \frac{f^{(i)}(h) - f^{(i)}(0)}{h} = \lim_{h \rightarrow 0} \frac{f^{(i)}(h)}{h}$$

を得る. $f^{(i)}$ は帰納法の仮定より微分可能であり, $\lim_{h \rightarrow 0} f^{(i)}(h) = \lim_{h \rightarrow 0} h = 0$ であるから l'Hôpital の定理が使えて,

$$\lim_{h \rightarrow 0} \frac{f^{(i)}(h)}{h} = \lim_{h \rightarrow 0} \frac{f^{(i+1)}(h)}{1}$$

を得るので, この時点で $f^{(i+1)}$ が原点で定義されていればそれは連続であるとわかる. 更に帰納的に議論を繰り返すことで,

$$f^{(i+1)}(0) := \lim_{h \rightarrow 0} \frac{f^{(i)}(h)}{h} = \lim_{h \rightarrow 0} \frac{f^{(i-1)}(h)}{h^2} = \cdots = \lim_{h \rightarrow 0} \frac{f^{(1)}(h)}{h^i}$$

がわかるが, ふたたび補題 2.32 によってこれは 0 である. ■

命題 2.34. 関数 $\psi: \mathbb{R} \rightarrow \mathbb{R}$ を,

$$\psi(x) = \begin{cases} \exp\left(-\frac{1}{(x-1)^2}\right) \cdot \exp\left(-\frac{1}{(x+1)^2}\right) & \text{if } x \in (-1, 1) \\ 0 & \text{if } x \notin (-1, 1) \end{cases}$$

で定めると, ψ は C^∞ 級であり, $(-1, 1)$ においては正であってそれ以外では 0 である. $a \in \mathbb{R}^n$ および $\varepsilon > 0$ に対して関数 $g_{a,\varepsilon}: \mathbb{R}^n \rightarrow \mathbb{R}$ を,

$$g_{a,\varepsilon}(x) := \psi\left(\frac{x^1 - a^1}{\varepsilon}\right) \cdot \psi\left(\frac{x^2 - a^2}{\varepsilon}\right) \cdots \psi\left(\frac{x^n - a^n}{\varepsilon}\right)$$

で定めると、 $g_{a,\varepsilon}$ は開方体 $(a^1 - \varepsilon, a^1 + \varepsilon) \times (a^2 - \varepsilon, a^2 + \varepsilon) \times \cdots \times (a^n - \varepsilon, a^n + \varepsilon)$ においては正であってそれ以外では 0 である。

証明. 前半の主張は明らか。後半は $(x^i - a^i)/\varepsilon \in (-1, 1) \Leftrightarrow x^i \in (a^i - \varepsilon, a^i + \varepsilon)$ なのでよい。 ■

ここまでで部品として使う関数の構成が終わった。ここからは上記の関数を配置する「 K 上の適切な有限箇所」の選定・構成に移る。

ところでここまで書いたところで、このノートでは \mathbb{R}^n の集合の境界などについてろくに議論していなかったことに思い至ったので、一旦議論しておくことにする。

定義 2.35. $A \subset \mathbb{R}^n$ を部分集合とする。 $x \in \mathbb{R}^n$ は以下の 3 条件のうちいずれか 1 つのみを充たす。

- (i) ある $\varepsilon > 0$ に対して $B(x; \varepsilon) \subset A$.
- (ii) ある $\varepsilon > 0$ に対して $B(x; \varepsilon) \subset A^C$.
- (iii) 任意の $\varepsilon > 0$ に対して $B(x; \varepsilon) \cap A \neq \emptyset$ および $B(x; \varepsilon) \cap A^C \neq \emptyset$ の両方を充たす。

1. を充たす点の集合を A の内部といい、 A° で表す。 2. を充たす点の集合を A の外部という。 3. を充たす点の集合を A の境界といい、 ∂A で表す。 $\bar{A} := A \cup \partial A$ を A の閉包という。

定義より $A^\circ \subset A$ および $(A^C)^\circ \subset A^C$ は直ちに分かるし、 A の外部とは $(A^C)^\circ$ のことにほかならない。もっと強く次のことがわかる。

問 2.36. A° は A を含むような開集合のうちで最大のものである。 ∂A は閉集合である。 \bar{A} は A を含むような閉集合の中で最小のものである。

問 2.37 (*). このあたりの話は「当たり前の話をくどくどとやっているようにしか見えず、実際にはもっとかんたんに議論が済むのではないか？」と思うかもしれない。その疑念をある意味で強める、ある意味ではこの疑念に割り切りをつける助けとなる定理として Jordan の閉曲線定理というものがある； $C \subset \mathbb{R}^2$ を単純閉曲線とする。このとき $\mathbb{R}^2 \setminus C$ は 2 つの連結成分からなり、ひとつは有界でひとつは非有界である。

この定理は「平面に輪っかを描くと平面はその内側と外側に分かれる」という極めて当たり前のことを述べただけの定理に思えるが、その主張の素朴さに反して、証明は大変であるらしい。少なくとも私は証明を読んだことがなく、調べるたびに「読む気しないな…」と思ってそっ閉じしている。

定義 2.38. 部分集合 $A, B \subset \mathbb{R}^n$ に対し、 $d(A, B) := \inf \{ \|x - y\| \mid x \in A, y \in B \}$ と定める。ただし $A = \{x\}$ (1 点集合) の場合は (いちいち波括弧を書くのが煩雑なので) $d(A, B)$ のことを $d(x, B)$ とも書く。 B が 1 点集合の場合も同様の記法を用いる。

命題 2.39. $C \subset \mathbb{R}^n$ を閉集合, $K \subset \mathbb{R}^n$ をコンパクト集合とする. このとき, $C \cap K = \emptyset$ であるならば $d(C, K) > 0$.

証明. $C \cap K = \emptyset$ なので, K の各点 x に対して $d(x, C) =: \varepsilon_x > 0$ である. $\{B(x; \varepsilon_x/2)\}_{x \in K}$ は K の開被覆であるから, K のコンパクト性によりこの開被覆から有限個を選んだ $\{B(x_i; \varepsilon_{x_i}/2)\}_{i=1}^n$ がすでに K を被覆する. ε_x の定め方より, $y \in B(x_i; \varepsilon_{x_i}/2)$ であるならば, $d(y, C) \geq \varepsilon_{x_i}/2 > 0$ である. 従って, 任意の $y \in K$ に対して $d(y, C) \geq \min_i \{\varepsilon_{x_i}/2\} > 0$ となる. ■

問 2.40. 命題 2.39 において, K のコンパクト性は本質的である. すなわち $C \cap K = \emptyset$ を充たす閉集合 $C, K \subset \mathbb{R}^n$ であって $d(C, K) = 0$ となるようなものがある.

命題 2.41. $K \subset A \subset \mathbb{R}^n$ とし, 更に K をコンパクト, A を開集合とする. このとき, $K \subset L \subset A$ を充たすコンパクト集合 L であって, L の内部に K を含むようなものがある. 更に C^∞ 級関数 $h: \mathbb{R}^n \rightarrow \mathbb{R}$ であって, K 上では正, L の外部では 0 であるようなものがある.

証明. この証明において, $a \in \mathbb{R}^n$ を中心とする幅 2ε の開方体 $\prod_{i=1}^n (a^i - \varepsilon, a^i + \varepsilon)$ のことを $R(a; \varepsilon)$ と書くことにする.

$x \in K$ を任意にとると, $\{x\}$ はコンパクトで, $A^c := \mathbb{R}^n \setminus A$ は閉集合なので, 命題 2.39 によって $\varepsilon_x > 0$ があり $d(x, A^c) = 2\varepsilon_x$ となる. このとき $R(x; \varepsilon_x) \subset A$ であり, 特に $d(R(x; \varepsilon_x), A^c) = \varepsilon_x$ である. これをすべての $x \in K$ に対して考えることで, K の開被覆 $\{R(x, \varepsilon_x)\}_{x \in K}$ を得る. K のコンパクト性により, ここから K の有限部分被覆 $\{R(x_i, \varepsilon_i)\}_{i=1}^n$ を取ることができるから取る. $M := \bigcup_{i=1}^n R(x_i; \varepsilon_i)$ と置き, $L := \overline{M}$ とする. L は有界閉集合なのでコンパクトで, 明らかにその内部 M が K を含み, かつ $L \subset A$ である.

命題 2.34 に即して関数 g_{x_i, ε_i} を作り, $h := \sum_{i=1}^n g_{x_i, \varepsilon_i}$ と定めると, これは M 上で正なので K 上でも正であり, L の外部で 0 である. ■

ここまでで関数の概形の構成が済んだので, あとは値を調整すれば良い.

補題 2.42. C^∞ 級関数 $k_\varepsilon: \mathbb{R} \rightarrow \mathbb{R}$ であって, $x \geq \varepsilon$ ならば $k_\varepsilon(x) = 1$, $x \leq 0$ ならば $k_\varepsilon(x) = 0$ を充たすものがある.

証明.

$$k_\varepsilon(x) := \frac{\int_0^x g_{\varepsilon/2, \varepsilon/2}(t) dt}{\int_0^\infty g_{\varepsilon/2, \varepsilon/2}(t) dt}$$

とせよ^{*16}. ■

^{*16} 要するに $g_{\varepsilon/2, \varepsilon/2}$ を確率密度関数に持つような確率分布関数を作るのを真似ればよい, と言っている.

ここまで来れば定理 2.29 を示すのは難しくない。

証明 (定理 2.29). $h: \mathbb{R}^n \rightarrow \mathbb{R}$ を、命題 2.41 で構成した関数とすると、 h の K への制限は正值であり、かつ K はコンパクトなので h は K 上最小値を持つからそれを $\varepsilon > 0$ とする。 $k_\varepsilon \circ h$ がそのような関数である。 ■

2.4 逆関数定理

我々は一変数の解析学において次のことを学んでいる；

命題 2.43. C^1 級関数 $f: \mathbb{R} \rightarrow \mathbb{R}$ が $a \in \mathbb{R}$ において $f'(a) \neq 0$ を満たしているならば、ある a を含む開集合 U があり、 f は U 上で単射であって、 $f(U)$ 上定義される f の逆写像 $g: f(U) \rightarrow U$ は C^1 級である。

証明. f は C^1 級すなわち f' が連続であり、 $f'(a) \neq 0$ なので、 a を含む開集合 U が存在して、 $x \in U$ ならば $f'(x) \neq 0$ である。一般性を失わず $f'(a) > 0$ と仮定してよい ($f'(a) < 0$ ならば $-f$ に対して同様の考察をすれば良い) ので、特に $x \in U$ ならば $f'(x) > 0$ であるとしてよい。微分の定義より f は U 上で狭義単調増加であるとわかるから、 f の V への制限 $f|_V: V \rightarrow f(V)$ は単射であるので、逆写像 $g: f(U) \rightarrow U$ がある。

g が $f(U)$ の各点で微分可能であることを示すため、 $f(U)$ の二点 $w, w+h \in f(U)$ をとる。するとある $v, v+k \in U$ があって $w = f(v)$, $w+h = f(v+k)$ である。したがって

$$\lim_{h \rightarrow 0} \frac{g(w+h) - g(w)}{h} = \lim_{h \rightarrow 0} \frac{g(f(v+k)) - g(f(v))}{h} = \lim_{h \rightarrow 0} \frac{v+k-v}{f(v+k)-f(v)} = \lim_{h \rightarrow 0} \frac{k}{f(v+k)-f(v)}$$

を得る。 f の連続性より $h \rightarrow 0$ と $k \rightarrow 0$ は同値なので、この極限は結局 $1/f'(v)$ に等しい。 g が C^1 級であることはこのことから直ちに従う。 ■

定義 2.44. $f: A \rightarrow B$ が C^r 級微分同相であるとは、 f が C^r 級であり、 f の逆写像 $g: B \rightarrow A$ が存在して更に C^r 級であることをいう。この言葉を用いれば、上述の定理の主張は「 C^1 級関数 f の微分が a において 0 でないならば、 f は a の充分近くで C^1 級微分同相となる」とまとめることができる。

この証明のあらましは、 C^1 級関数の微分がある点で 0 でないならば、その点の充分近くでもやはり 0 ではなく、したがって値が常に変化し続けるから、局所的に逆写像を持つ、というように述べることができよう。この「値が常に変化し続ける」とは、「微分が非零の微小変動を零につぶさない」と述べることができる。これはすなわち微分を線型写像として見た時に、それが可逆であるということである。この観察を高次元の場合に拡張して得られるのが、逆関数定理である。

定理 2.45 (逆関数定理). $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ が $a \in \mathbb{R}^n$ において C^1 級であるとし、更に $\det J_f(a) \neq 0$ で

あると仮定する．このとき， a を含む開集合 $U \subset \mathbb{R}^n$ があり， $f|_U: U \rightarrow f(U)$ は C^1 級微分同相である．

証明にはいくつかの方法がある．ここでは Spivak とはやり方を変えて，Banach の縮小写像の原理に依る証明を与えることにする．やり方を変える強い理由があるわけではないが，単に本と違ったことをやってみたいというだけである．まず Banach の縮小写像の原理を述べる．

定義 2.46 (縮小写像). $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ が縮小的である，あるいは縮小写像であるとは，ある $0 \leq \mu < 1$ が存在して，任意の $x, y \in \mathbb{R}^n$ に対して $\|f(x) - f(y)\| \leq \mu \|x - y\|$ が成り立つことをいう．

問 2.47. f は連続写像である．

定理 2.48 (Banach の縮小写像の原理). $D \subset \mathbb{R}^n$ を閉集合とし， $f: D \rightarrow D$ を縮小写像とする．このとき， D に f の不動点がただひとつ存在する．すなわち， $f(y) = y$ を満たす $x \in D$ がただひとつ存在する．更に，任意の $x \in D$ に対し， $\lim_{k \rightarrow \infty} f^k(x) = y$ が成り立つ．ただし自然数 k に対して， $f^1(x) := f(x)$ ， $f^k(x) := f(f^{k-1}(x))$ によって帰納的に定める．

証明. f は縮小的なので，ある $0 \leq \mu < 1$ が存在して，任意の $x, y \in \mathbb{R}^n$ に対して $\|f(x) - f(y)\| \leq \mu \|x - y\|$ が成り立つ．

$y_0 \in D$ を任意に取り，点列 $\{y_i\}_{i=1}^\infty$ を $y_i := f^i(y_0)$ によって定める．すると任意の $k \in \mathbb{N}$ に対して $\|y_{k+1} - y_k\| \leq \mu^k \|y_1 - y_0\|$ が成り立つから， $m > n$ を満たすような $m, n \in \mathbb{N}$ に対して

$$\begin{aligned} \|y_m - y_n\| &\leq \|y_m - y_{m-1}\| + \|y_{m-1} - y_{m-2}\| + \cdots + \|y_{n+1} - y_n\| \\ &\leq (\mu^{m-1} + \mu^{m-2} + \cdots + \mu^n) \|y_1 - y_0\| = \mu^n \frac{1 - \mu^{m-n}}{1 - \mu} \|y_1 - y_0\|. \end{aligned}$$

最右辺は $m, n \rightarrow \infty$ で 0 に収束するので， $\{y_i\}$ は \mathbb{R}^n の Cauchy 列である．したがって収束先 $y := \lim_{i \rightarrow \infty} y_i$ が \mathbb{R}^n に存在するが， D は閉であって $\{y_i\}$ は D の点列なので， $y \in D$ である．更に f の連続性より $y = \lim_{i \rightarrow \infty} y_i = \lim_{i \rightarrow \infty} f(y_{i-1}) = f(\lim_{i \rightarrow \infty} y_i) = f(y)$ であるから y は f の不動点になっている．更に z も f の不動点であったとすると， $\|y - z\| = \|f(y) - f(z)\| \leq \mu \|y - z\|$ を得るが， $\mu < 1$ なので $\|y - z\| = 0$ すなわち $y = z$ である． ■

問 2.49. \mathbb{R} の Cauchy 列が収束列であることは実数の完備性（の同値な言い換えのひとつ）である． \mathbb{R}^n の Cauchy 列が収束列であることは，実数の完備性および「 \mathbb{R}^n の点列が収束することと，各成分が \mathbb{R} の点列として収束することは同値」という事実から従う．

問 2.50. $D \subset \mathbb{R}^n$ が閉で， D 内の点列 $\{y_i\}$ が収束するならばその収束先は D に属する．

次いで証明に入る前にひとつ補題を用意しておく．微分は関数の線型近似であることはすでに述

べたが, その近似についてより定量的な評価を与えておく必要がある.

補題 2.51. $D \subset \mathbb{R}^n$ を開集合とし, $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ が D で微分可能であると仮定する. 更に任意の $a \in D$ に対して $J_f(a)$ の各成分の絶対値は M を超えないと仮定する. このとき, 任意の $a, a+h \in D$ に対して $\|f(a+h) - f(a)\| \leq n^2 M \|h\|$ が成り立つ.

証明. 各成分関数 f^i に対して, $f^i(a+h) - f^i(a)$ を,

$$\begin{aligned} f^i(a+h) - f^i(a) &= f^i(a^1 + h^1, a^2, \dots, a^n) - f^i(a^1, a^2, \dots, a^n) \\ &\quad + f^i(a^1 + h^1, a^2 + h^2, \dots, a^n) - f^i(a^1 + h^1, a^2, \dots, a^n) \\ &\quad + \dots \\ &\quad + f^i(a^1 + h^1, a^2 + h^2, \dots, a^{n-1} + h^{n-1}, a^n + h^n) - f^i(a^1 + h^1, a^2 + h^2, \dots, a^{n-1} + h^{n-1}, a^n) \end{aligned}$$

のように表示しておく. f^i は各変数に関して偏微分可能なので, 平均値の定理により

$$|f^i(a^1 + h^1, a^2, \dots, a^n) - f^i(a^1, a^2, \dots, a^n)| \leq h^1 M$$

が成り立つ. したがって

$$|f^i(a+h) - f^i(a)| \leq (h^1 + h^2 + \dots + h^n)M \leq n\|h\|M$$

である. ふたたびノルムの三角不等式より $\|f(a+h) - f(a)\| \leq \sum |f^i(a+h) - f^i(a)|$ が得られるから結論が従う. ■

以上で逆関数定理を示す準備が整った.

証明 (定理 2.45). まず座標軸を平行移動することで $a = 0$ および $f(a) = 0$ と仮定して構わない. 更に $\lambda: \mathbb{R}^n \rightarrow \mathbb{R}^n$ を可逆線型写像とすると, $\lambda^{-1} \circ f$ に関して逆関数定理が成り立てばあきらかに f に対しても成り立つとわかる上, 仮定より df_a は可逆なので, 必要ならば $(df_a)^{-1}$ 倍した写像を考えることで $df_a = \text{id}_{\mathbb{R}^n}$ すなわち $J_f(a) = I_n$ と仮定してもやはり一般性を失わない. ただし I_n は $n \times n$ 単位行列. なのでこれらの仮定を置いて証明を進める.

f は C^1 級なので, ある半径 δ の閉球 $D = \overline{B(0; \delta)}$ が存在して, 任意の $x \in D$ および $1 \leq i, j \leq n$ に対して $\|\partial_j f^i(x) - \partial_j f^i(0)\| \leq 1/2n^2$ となるようにできる. このとき, 関数 $w: D \rightarrow D$ を $w(x) := f(x) - x$ によって定めると, $J_w(x) = J_f(x) - I_n = (\partial_j f^i(x) - \partial_j f^i(0))$ なので, 補題 2.51 によって $\|w(x+h) - w(x)\| \leq \|h\|/2$ すなわち $\|f(x+h) - f(x) - h\| \leq \|h\|/2$ を得る.

$f(a) = 0$ を含む開集合として $V = B(0; \delta/2)$ を取る. ここで, 任意の $y \in V$ に対して $f(x) = y$ を充たす $x \in D$ が一意に存在すること, すなわち単射 $g: V \rightarrow D$ であって $f \circ g = \text{id}$ となるものが存在することを示したい. これが示せれば f の $g(V)$ への制限が可逆であるとわかる. g は次のようにして構成する; 関数 $u_y: D \rightarrow \mathbb{R}^n$ を $u_y(x) = x + (y - f(x))$ によって定め, u_y の値域が実際には D であること, および $u_y: D \rightarrow D$ が縮小的であることを示せば, Banach の縮小写像の原理によって u_y は唯一の不動点を持つとわかるので, それを $g(y)$ とすればよい.

u_y の値域が D であることは、先程得た評価 $\|f(x+h) - f(x) - h\| \leq \|h\|/2$ において $x = 0$ とすることで $\|f(h) - h\| \leq \|h\|/2 \leq \delta/2$ が従うので、 $\|u_y(x)\| \leq \|y\| + \|x - f(x)\| < \delta/2 + \delta/2 = \delta$ となるからよい。 $\|u_y(x+h) - u_y(x)\| = \|h - f(x+h) - f(x)\| \leq \|h\|/2$ となるから縮小的であることもわかる。更にこの時点で u_y の値域は実際には D° であるとわかったので、開集合 $U := g(V) = D^\circ$ へ f を制限したものは可逆であることが従う。

g の連続性を示すため、相異なる $y, y+k \in V$ を取ると、 $f(x) = y$ および $f(x+h) = y+k$ を充たすような $x, x+h \in D^\circ$ が一意的に存在する。したがって $\|k - h\| \leq \|y+k - y - h\| \leq \|h\|/2$ を得るが、 $\|h\| - \|k\| \leq \|k - h\|$ であるから結局 $\|h\| \leq 2\|k\|$ を得る。したがって g は連続である。

g の微分可能性を示す。 $x, x+h \in D^\circ$ および $y, y+k \in V$ を連続性の議論において用いた点たちとし、 $f(x+h) - f(x) - J_f(x)h$ を考える。この量は $y+k - y - J_f(x)(g(y+k) - g(y)) = k - J_f(x)(g(y+k) - g(y))$ と書き直せる。線型写像 $h \mapsto (J_f(x) - I_n)h$ に対して補題 2.51 を適用することで $\|J_f(x)h - h\| \leq \|h\|/2$ を得るので、特に $\|h\| - \|J_f(x)h\| \leq \|h\|/2$ すなわち $\|J_f(x)h\| \geq \|h\|/2$ を得るから $J_f(x)$ は正則行列である。したがって適当な実数 $M > 0$ があって

$$\begin{aligned} \frac{\|g(y+k) - g(y) - J_f(x)^{-1}k\|}{\|k\|} &= \frac{\|g(y+k) - g(y) - J_f(x)^{-1}k\|}{\|h\|} \frac{\|h\|}{\|k\|} \\ &\leq 2 \frac{\|g(y+k) - g(y) - J_f(x)^{-1}k\|}{\|h\|} \\ &= 2 \frac{\|h - J_f(x)^{-1}(f(x+h) - f(x))\|}{\|h\|} \\ &\leq 2M \frac{\|J_f(x)h - f(x+h) + f(x)\|}{\|h\|} \end{aligned}$$

を得るが、最右辺は $h \rightarrow 0$ で 0 に収束するし、 g の連続性より $h \rightarrow 0$ と $k \rightarrow 0$ は同値であるから g は微分可能である。

ここまでの議論で $J_g(f(x)) = J_f(x)^{-1}$ がわかっているので、Cramer の公式より g は C^1 級である。 ■

問 2.52 ().** 逆関数定理の主張は Banach 空間までほぼそのまま拡張できるらしく、証明もここに挙げたように縮小写像の原理に依ればよいらしい。無限次元に一般化する場合、逆作用素の存在がデリケートな問題になるらしい。

問 2.53. Spivak では逆写像の構成にあたってまず $v_y(x) := \|y - f(x)\|^2$ という関数を考えている。これはコンパクト集合 D 上で最小値を持つ上、実際には D の内部で最小値をとることがわかる。それゆえ最小値を実現する点で微分が消えているはずなので $y = f(x)$ を充たすような x の存在が言える、という議論をしている。一意性は g の連続性において現れた議論をすこし変形することで示している。これはこれで味があってよいと思う反面、 D がコンパクトという性質、すなわち有限次元であることをクリティカルに使っている。

問 2.54 (**). それゆえ、Spivak の方針そのままでは Banach 空間への拡張はつらいと思われる。Banach 空間での逆関数定理の定式化には Fréchet 微分が現れる以上、ノルム位相を避けて通ることはできないが、ノルム位相に関して閉球は一般にコンパクトにならない。しかしながら閉球がコンパクトになるような弱い位相を併せて考えることで示せたりしないだろうか？

問 2.55 (*). 証明にあらわれた $u_y: D \rightarrow D$ が不動点を持つこと自体は Brouwer の不動点定理からもう。Brouwer の不動点定理では不動点の一意性までは示せないが、Spivak での一意性の証明と同じようなやり方が通ると思う（のだが、検証はしていない）。

問 2.56 (**). ちなみに Brouwer の不動点定理の証明はいくつか方法があるらしいが、多分縮小写像の原理を示すよりは手間だと思うのでやめた。すくなくともこのためだけにホモロジー群を導入するのは大変なのでやる気がしない。組合せ論的に示す方法もあるらしく、しかも比較的初等的に示せるという話を聞いたことはあるが、私はちゃんと勉強したことはない。他にも不動点定理とよばれるものはたくさんあるし、応用もまたたくさんあるらしい。

問 2.57. ところで逆関数定理の応用はあるのかということが気になる。ので、探してみたのだが、この時点で書けるようなよい例が見つからなかった。いい応用を見つけたらその話について勉強して書きたいのでおしえてください。

問 2.58 (**). この証明においては f が C^1 級であることをクリティカルに使った。ところが、Terrence Tao 氏がブログに書いている^{*17}ところによると、この仮定は落とせる（すなわち、 f がある点 a で微分可能かつ微分が可逆であるならば、 $f(a)$ のまわりで定義された微分可能な逆写像が存在する）らしい。

2.5 陰関数定理と未定乗数法

逆関数定理の兄弟分とでも言うべき定理が陰関数定理である。言葉による大雑把な説明をすると、いくつかの変数によって陰的に表される束縛条件があるときに、その束縛条件を書き表す陽的な方程式ないしは関数の存在を主張する定理である。

例 2.59. $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ を $f(x, y) := x^2 + y^2$ で定め、単位円周 $S^1 := \{(x, y) \in \mathbb{R}^2 \mid f(x, y) = 1\}$ を考える。このとき、任意の $(x, y) \in S^1$ に対して $y = g(x)$ を満たすような関数 $g: \mathbb{R} \rightarrow \mathbb{R}$ は存在しない。しかしながら「ある点の充分近くで考えれば」このような関数を見つけることができる場合がある。例えば $y > 0$ を満たすような $(x, y) \in S^1$ に対しては $y = \sqrt{1 - x^2}$ と書ける。

^{*17} <https://terrytao.wordpress.com/2011/09/12/the-inverse-function-theorem-for-everywhere-differentiable-maps/>

この現象を高次元の場合も含めて記述することが、本項の目的である。

定理 2.60 (陰関数定理). m, n を自然数とし, $f: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ を C^1 級関数とする. ある $a \in \mathbb{R}^m$ および $b \in \mathbb{R}^n$ に対して $f(a, b) = 0$ であり, $n \times n$ 行列 $M := (\partial_{m+j} f^i(a, b))$ が正則であると仮定する. このとき, a を含む開集合 $A \subset \mathbb{R}^m$ および b を含む開集合 $B \subset \mathbb{R}^n$ ならびに C^1 級関数 $g: A \rightarrow B$ があり, A 上で $f(x, g(x)) = 0$ を満たす. 更に g は次の意味で一意的である; 任意の $x \in A$ に対して $f(x, h(x)) = 0$ を満たす関数 $h: A \rightarrow B$ があれば, A 上で $g = h$.

証明. 関数 $F: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^{m+n}$ を, $x \in \mathbb{R}^m$ および $y \in \mathbb{R}^n$ に対して $F(x, y) := (x, f(x, y))$ によって定めるとこれは C^1 級である. 更に F の $(a, b) \in \mathbb{R}^{m+n}$ における Jacobi 行列は (左上の $m \times m$ ブロックが単位行列に, 右上の $m \times n$ ブロックがゼロ行列に, 右下の $n \times n$ ブロックが M に一致するので) 正則である. したがって逆関数定理より $(a, f(a, b))$ を含む開集合 $A' \times B \subset \mathbb{R}^m \times \mathbb{R}^n$ 上定義された関数 $G: A' \times B \rightarrow \mathbb{R}^{m+n}$ があり, $F \circ G = \text{id}$ である. F の関数の形から, 関数 $k: \mathbb{R}^{m+n} \rightarrow \mathbb{R}^n$ があって $G(x, y) = (x, k(x, y))$ と書ける. 関数 π を $\pi: \mathbb{R}^m \times \mathbb{R}^n \ni (x, y) \mapsto y \in \mathbb{R}^n$ で定めると, $f = \pi \circ F$ が成り立つ. したがって, $x \in A'$ かつ $k(x, y) \in B$ ならば

$$f(x, k(x, y)) = (f \circ G)(x, y) = (\pi \circ F \circ G)(x, y) = \pi(x, y) = y$$

である. 関数 $g: A' \rightarrow \mathbb{R}^n$ を $g(x) := k(x, 0)$ で定めれば, $(a, b) = G(a, 0) = (a, k(a, 0)) = (a, g(a))$ すなわち $g(a) = b$ なので, 充分小さな開集合 $A \subset A'$ をとることで $g(A) \subset B$ とできる. このとき $x \in A$ ならば $g(x) \in B$ なので, 特に $f(x, g(x)) = 0$ である. $k = \pi \circ G$ から k が C^1 級であるとわかるので g もそうである.

$h: A \rightarrow B$ が任意の $x \in A$ に対して $f(x, h(x)) = 0$ を満たすと仮定する. このとき, $F(x, h(x)) = (x, f(x, h(x))) = (x, 0) = (x, f(x, g(x))) = F(x, g(x))$ となるが, F は $A \times B$ において単射なので $g(x) = h(x)$ である. ■

陰関数定理は次のように述べ直すこともできる.

定理 2.61. n, p を $p \leq n$ を満たすような自然数とし, $a \in \mathbb{R}^n$ とする. $f: \mathbb{R}^n \rightarrow \mathbb{R}^p$ が C^1 級であって, 更に $f(a) = 0$ かつ f の $a \in \mathbb{R}^n$ における Jacobi 行列の階数は p であったと仮定する. このとき, a を含む開集合 U 上の C^1 級微分同相 $F: U \rightarrow \mathbb{R}^n$ であって, $f \circ F^{-1}(x^1, \dots, x^n) = (x^{n-p+1}, \dots, x^n)$ を満たすものがある.

証明. $f: \mathbb{R}^n \rightarrow \mathbb{R}^p$ を $f: \mathbb{R}^{n-p} \times \mathbb{R}^p \rightarrow \mathbb{R}^p$ と考える. $p \times p$ 行列 $(\partial_{n-p+j} f^i(a))$ の階数が p ならば, 逆関数定理より $F: \mathbb{R}^{n-p} \times \mathbb{R}^p \ni (x, y) \mapsto (x, f(x, y)) \in \mathbb{R}^n$ は a を含む開集合 U 上で C^1 級微分同相である. 更に F の逆写像を $G: F(U) \rightarrow U$ と書けば, $f \circ G(x^1, x^2, \dots, x^{n-p}, x^{n-p+1}, \dots, x^n) = (x^{n-p+1}, \dots, x^n)$ である.

$p \times p$ 行列 $(\partial_{n-p+j} f^i(a))$ の階数が p 未満であるときも, $1 \leq j_1 < j_2 < \dots < j_p \leq n$ を適切に選べば $(\partial_{j_i} f^i(a))$ の階数は p になる. $g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ を, 「 x^{j_i} たちが右側にまとまるように並べ替

える」写像とすると、 $f \circ g$ に対しては前段の議論が使える、 a を含む開集合 U 上の C^1 級微分同相 $k: U \rightarrow \mathbb{R}^n$ があって $f \circ g \circ k^{-1}(x^1, x^2, \dots, x^{n-p}, x^{n-p+1}, \dots, x^n) = (x^{n-p+1}, \dots, x^n)$ となる。 $F := k \circ g^{-1}$ とすればよい。 ■

この述べかえは、のちのち微分可能多様体の判定法として使えることが明らかになるが、直ちに分かる応用をひとつ紹介しておこう；制限付き極値問題である。

命題 2.62 (Lagrange の未定乗数法). n, p を $p \leq n$ を満たすような自然数、 $A \subset \mathbb{R}^n$ を開集合、 $g: A \rightarrow \mathbb{R}^p$ を C^1 級関数とし、任意の $a \in A$ に対して $J_g(a)$ の階数は p であったとする。更に $f: g^{-1}(0) \rightarrow \mathbb{R}$ を微分可能な関数とし、 $a \in \mathbb{R}^n$ において f は極値を取ると仮定する。このとき、 p 個の実数 $\lambda_1, \dots, \lambda_p$ が存在して、 $1 \leq j \leq p$ に対して

$$\partial_j f(a) = \sum_{i=1}^p \lambda_i \partial_j g^i(a)$$

を満たす。

証明. 証明を 2 段階に分割する。まず $g(x) = (x^{n-p+1}, \dots, x^n)$ のときを考える。このとき、 $i, j = 1, 2, \dots, p$ および $j = 1, 2, \dots, n$ に対して

$$\partial_j g^i(a) = \begin{cases} 1 & \text{if } j = n - p + i \\ 0 & \text{otherwise} \end{cases}$$

である。 f を $g^{-1}(0)$ に制限するというのは $x^{n-p+1} = x^{n-p+2} = \dots = x^n = 0$ に固定して $n - p$ 変数の関数だと見なすことと同値である。したがって f が $a \in \mathbb{R}^n$ において極値を取るならば、 $\partial_1 f(a) = \partial_2 f(a) = \dots = \partial_{n-p} f(a) = 0$ である。したがって $\lambda_i = \partial_{n-p+i} f(a)$ とすれば命題の主張が成り立つ。一般の g に対しては、陰関数定理によって a を含む開集合 U で定義された C^1 級微分同相 F があって $g \circ F^{-1} = (x^{n-p+1}, \dots, x^n)$ となるから、 F の逆写像を G と書けば、前段の議論により実数 $\lambda_1, \dots, \lambda_p$ があって

$$\partial_j(f \circ G)(F(a)) = \sum_{i=1}^p \lambda_i \partial_j(g \circ G)^i(F(a)) = \sum_{i=1}^p \lambda_i \partial_j(g^i \circ G)(F(a))$$

となる。命題 2.25 より

$$\begin{aligned} \partial_j(f \circ G)(F(a)) &= \sum_{k=1}^n \partial_k f(G(F(a))) \cdot \partial_j G^k(F(a)) = \sum_{k=1}^n \partial_k f(a) \cdot \partial_j G^k(F(a)), \\ \sum_{i=1}^p \lambda_i \partial_j(g^i \circ G)(F(a)) &= \sum_{i=1}^p \lambda_i \sum_{k=1}^n \partial_k g^i(G(F(a))) \cdot \partial_j G^k(F(a)) = \sum_{i=1}^p \lambda_i \sum_{k=1}^n \partial_k g^i(a) \cdot \partial_j G^k(F(a)) \end{aligned}$$

を得るから、

$$\sum_{k=1}^n \partial_k f(a) \cdot \partial_j G^k(F(a)) = \sum_{i=1}^p \lambda_i \sum_{k=1}^n \partial_k g^i(a) \cdot \partial_j G^k(F(a))$$

が従う。この式を行列演算の形で書いた上で $\partial_j G^k(F(a))$ の逆行列を掛ければ結論を得る (G が微分同相なので, $\partial_j G^k(F(a))$ は正則行列である)。 ■

以下しばらく $p = 1$ とし, $\nabla f := (\partial_1 f, \dots, \partial_n f)$ と置く。Lagrange の未定乗数法の主張は, f を $g = 0$ に制限した関数の臨界点は $\nabla(f - \lambda g) = 0$ を充たす, ということを主張している。したがって $\partial_i(f - \lambda g) = 0$ および $g = 0$ という $n+1$ 本の連立方程式が立つので, $x^1, x^2, \dots, x^n, \lambda$ に関して解くことができるから, 実際にやればよい。ということでいくつか例を見てみることにしよう。

例 2.63. ここでは \mathbb{R}^n の座標として p^1, p^2, \dots, p^n を使うことにする。 $\Delta_n := \{p \in \mathbb{R}^n \mid \sum_{i=1}^n p^i = 1, p^1, p^2, \dots, p^n \geq 0\}$ を標準 n -単体という。 $p \in \Delta_n$ に対して $H(p) := -\sum_{i=1}^n p^i \log p^i$ を p のエントロピーと呼ぶ。ただし $0 \log 0 = 0$ と定める。いくつかの状況について, エントロピーを最大化する $p \in \mathbb{R}^n$ を求めてみよう。エントロピーは凹関数なので, 極値を達成する点があればその点で最大値を達成することに注意しておく。

- $p \in \Delta_n$ 以外に特に制約のない場合。 $L(p, \lambda) := H(p) - \lambda(\sum_{i=1}^n p^i - 1)$ を p^i で微分して 0 と置いた式 $\partial_i L(p, \lambda) = 0$ を考えることで, $p^i = \exp(-\lambda + 1)$ を得る。すなわち $p^1 = p^2 = \dots = p^n = 1/n$ である。
- $k = 1, 2, \dots, m$ に対し, $f_k: \{1, \dots, n\} \rightarrow \mathbb{R}$ を関数とする。 $\sum_{i=1}^n f_k(i)p^i$ を, f_k の $p \in \Delta_n$ における期待値という。ここでは, f_k たちの期待値に制約がついている状況を考える。すなわち, U_1, U_2, \dots, U_m を実数として, $\sum_{i=1}^n f_k(i)p^i = U_k$ という制約のもとでエントロピーを最大化する $p \in \Delta_n$ を求めることを考える。もちろん f_k たちの具体的な形がわからないので, 解析解の計算はできないが, f_k たちの具体形に依らないところまでは計算を進めてみることにしよう。

$$L(p, \lambda^0, \lambda^1, \dots, \lambda^m) := H(p) - (\lambda^0 - 1) \left(\sum_{i=1}^n p^i - 1 \right) - \sum_{k=1}^m \lambda^k \left(\sum_{i=1}^n f_k(i)p^i - U_k \right)$$

と置く (λ^0 を定数シフトして定義しているのは, 後のノーターションをきれいにするためだけの便宜である)。これを p^i で微分して 0 と置いた式は

$$\partial_i L(p, \lambda) = -\log p_i - 1 - (\lambda^0 - 1) - \sum_{k=1}^m \lambda^k f_k(i) = 0$$

すなわち

$$p_i = \exp \left(-\lambda^0 - \sum_{k=1}^m \lambda^k f_k(i) \right)$$

を得る。他の制約式にこの表式を代入してみる^{*18}。まず

$$1 = \sum_{i=1}^n p^i = e^{-\lambda^0} \sum_{i=1}^n \exp\left(-\sum_{k=1}^m \lambda^k f_k(i)\right)$$

を得る。ここに現れた

$$Z(\lambda) = Z(\lambda^1, \lambda^2, \dots, \lambda^m) := \sum_{i=1}^n \exp\left(-\sum_{k=1}^m \lambda^k f_k(i)\right)$$

を分配関数と呼ぶ。 $\lambda^0 = \log Z(\lambda)$ なので、 λ^0 の計算は分配計算の計算に帰着する。分配関数を計算するために期待値に対する制約式を検討すると、

$$U_k = \sum_{i=1}^n f_k(i) p^i = \sum_{i=1}^n f_k(i) p^i = \sum_{i=1}^n f_k(i) \frac{e^{-\sum_k \lambda^k f_k(i)}}{Z(\lambda)} = -\frac{\partial_i Z(\lambda)}{Z(\lambda)} = -\partial_i \log Z(\lambda)$$

を得る。これは λ^k たちに関する m 本の方程式になるので、(実際の計算の面倒具合はともかくとして、また f_k たちが互いに相異なるなど、いくつか付加的な条件は必要であると思われるが) λ^k たちを決定するに十分な本数の方程式である。

余談 2.64. $p \in \Delta_n$ は、実際には有限標本空間上の確率分布というほうが普通であるが、そのことを知らないふりをして書いてある。次節は測度論に基づいて積分を定義するが、気が向いたら、例として確率分布や確率測度の話を入れてもいいかもしれない。

例 2.65. やや脱線となるが、前例で現れたエントロピーに対する考察を続けよう。前例の議論を反芻すると、関数 f_k たちの期待値が取る値を U_1, \dots, U_m に制約すれば、Lagrange の未定乗数法によって制約下でのエントロピーの最大値が求められた。これは、

$$S(U) = S(U_1, U_2, \dots, U_m) := \max_p H(p) = \max_{p, \lambda^0, \lambda} \left[H(p) - (\lambda^0 - 1) \left(\sum_{i=1}^n p^i - 1 \right) - \sum_{k=1}^m \lambda^k \left(\sum_{i=1}^n f_k(i) p^i - U_k \right) \right]$$

なる関数 S がある、と言い換えることができる。ただし \max は $p \in \Delta_n$ ならびに期待値に関する制約を満たすような範囲で取る。最大値を達成する点では $p_i = \exp\left(-\lambda^0 - \sum_{k=1}^m \lambda^k f_k(i)\right)$ であること

^{*18} 未定乗数法でなにか問題を解く時に一番難しい、ないし時間がかかるのはここから先ではないだろうか？未定乗数ならびにももとの引数に関する何本かの方程式を具体的に与えられて、それに対してどのように変形すればよいかがよくわからない、ということが私はよくある。このような状況に対して一般論ができることは殆どなく、ここから先で物を言うのは高校数学や算数の計算をちゃんと遂行できる腕力だと思う。「理論は落ち着いて考えればわかることしか書いてない」という言葉を聞いたことがあるが、言い得て妙だと思う（と同時に、学部生のときの私にこの言葉を聞かせても聞く耳を持たなかっただろう）。

がわかっているの、これを代入して計算を続けると

$$\begin{aligned}
S(U) &= \max_{p, \lambda^0, \lambda} H(p) \\
&= \max_{p, \lambda^0, \lambda} \left[- \sum_{i=1}^n p^i \left(-\lambda^0 - \sum_{k=1}^m \lambda^k f_k(i) \right) \right] \\
&= \max_{p, \lambda^0, \lambda} \left[\lambda^0 + \sum_{k=1}^m \lambda^k \sum_{i=1}^n f_k(i) p^i \right] \\
&= \max_{\lambda} \left[\log Z(\lambda) + \sum_{k=1}^m \lambda^k U_k \right]
\end{aligned}$$

を得る。一般に、凸関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ に対して $f^*(p) := \max_x [\langle x, p \rangle - f(x)]$ を f の Legendre 変換という。この言葉を用いれば、「 S は分配関数 $-\log Z(\lambda)$ の Legendre 変換である」と言い換えることができる。この S は熱力学的エントロピーと呼ばれる。

さて、 S の表式には λ^k たちが現れているが、実際には S は U_k の関数なので、 $H(p)$ の制約付き極値における λ^k は U_k の関数として書ける^{*19}。ここでは更に λ^k たちが U_k に関して微分可能であると仮定してしまうことにする。すると命題 2.25 によって

$$\partial_i S(U) = \sum_{k=1}^m \left[\partial_k \log Z(\lambda(U)) \cdot \partial_i \lambda^k(U) \right] + \sum_{k=1}^m \left[U_k \partial_i \lambda^k(U) \right] + \lambda^k(U) = \lambda^k(U)$$

を得る。したがって未定乗数は熱力学的エントロピーの導関数として現れることがわかった。いくつかの物理学的設定のもとでは、 λ^k は「示量変数 U_k に共役な示強変数」と呼ばれる。また物理学においては熱力学的エントロピー S の引数に「エネルギー」と呼ばれる引数 U を含むように要請することが多い。この U に共役な示強変数には「逆温度」という名前がついている；もちろん、適切な設定ならびに考察のもとで、この量と我々が見知っている温度との間に深い関わりを見いだせるだろうが、このノートで書くのはここまでにしておく。

問 2.66 ().** ここでは λ^k たちが U_k に関して微分可能であると仮定してしまったが、この仮定はどれくらい本質的なのだろうか。本質的であるとして、それを正当化する物理的考察はどのようになされるか。正当化を許さない物理的状況は（もし存在するならば）どのようなものか。このいずれについても、私は答えを知らない。

問 2.67 ().** Legendre 変換が顔を出す物理学的な分野といえば解析力学が挙げられるだろう。解析力学を多変数の微積分の延長上に捉えることはできるだろうか。シンプレクティック幾何というものがあるということだけを聞きかじっているが、勉強したことはない。

問 2.68 ().** ここまでは、制約が p^i たちの線型結合で表されるものだけを考えてきた。しかし Lagrange の未定乗数法は、線型結合に限らない C^1 級関数 g について、 $g(p) = 0$ に制限したもと

^{*19} 本当は記号を分けて $\lambda^k_*(U_1, \dots, U_n)$ などと書くべきだろうが、記号を分けないほうが一般的であるように思う。

での f の極値を求める方法を与えている．ということはここまでの議論がある程度は同様に通り，より一般化された状況での Legendre 変換 $\max [\lambda g - f]$ が見いだされるはずである．この拡張された Legendre 変換はどこまでよく振る舞い，拡張に伴ってどのように振る舞いを変えるのだろうか．微分可能多様体にも通るような定義に拡張できるだろうか．拡張することではじめて得られるありがたみは何で，どのような応用があるのだろうか．ちなみに Legendre 変換のひとつの一般化として，接触変換というものがあるらしく，これがひとつの答えになっているのではないかと期待しているが，一方この概念はどのような文脈で現れるのだろうか．

3 測度論

積分の扱いに入る．このノートでは Spivak から大きく構成を変えて、測度論に基づいて積分を定義する．理由は大小含めていくつかあるが、主要な理由の第一としては積分と極限の順序交換ができる条件を見通しよく述べたかったからである^{*20}．

多変数の積分を本格的に取り扱うのは次節になる．また、あくまで本稿は多変数の解析学のノートなので、測度論の一般論に深く立ち入ることはしない．

余談 3.1. ちなみに Spivak では測度 0 の定義はしているが、測度の定義はしていない．数学においては赤ニシンは単なるニシンでも単なる赤でもないという原則を反映していると言える．

問 3.2. ところで、可算無限と非可算無限の違いだけは把握していないとこの節は読めないかもしれない．

定義 3.3. 単調増加な集合の族 $A_1 \subset A_2 \subset \cdots$ が $A = \bigcup_{i=1}^{\infty} A_i$ を満たすことを $A_i \uparrow A$ と書く．単調減少な集合の族 $B_1 \supset B_2 \supset \cdots$ が $B = \bigcap_{i=1}^{\infty} B_i$ を満たすことを $B_i \downarrow B$ と書く．

3.1 測度空間と Lebesgue 測度

本項の目的は、Lebesgue 測度を定義すること、およびそのために必要な測度論の結果を最低限カバーすることである．我々は高校までの数学で面積や体積といった概念に慣れ親しんできた．これらの概念を厳密に捉え、積分に関する理論構成を見通しよくしつつ、一般の集合に対して拡張できるようにしたものが測度である．実際のところ、面積や体積に対して我々がまず疑念を抱かないような点のみを抽象化したと言っても過言ではない；測度とは、大雑把に言えば次の条件を満たす写像として定式化される．

- 空集合は測度 0.
- 排反な集合たちの測度の和は、それらの和集合の測度にひとしい．

大雑把にはこれだけなのだが、実際には注意を必要とする点がある．

問 3.4. たとえば、「長さ」に関するわれわれの素朴な直感を定式化して \mathbb{R} 上の測度 μ を作ることを考えよう． $\{x\} \subset \mathbb{R}$ を 1 点集合とするとき、 $\mu(\{x\}) = 0$ と定めるのが自然に思える．他方で、 $\mu([0, 1]) = 1$ と定めたいとも思える．ところで、先程の「排反な集合の測度は、それらの和集合

^{*20} いくらか個人的な理由たちとしては、Spivak 自身が測度 0 集合の性質に関してそれなりに丁寧に議論しており、測度論の枠組みに片足を突っ込んでしまっていること、私が測度論で出てくる「骨の折れる証明」に対して手を動かしてみたことがなかったのも、手を動かして理解したいと思っていたこと、などがある．

の測度にひとしい」という言明を何も考えずに数式に落とし込むと、 $\mu([0, 1]) = \sum_{x \in [0, 1]} \mu(\{x\}) = \sum_{x \in [0, 1]} \mu(\{x\}) = \sum_{x \in [0, 1]} 0$ を得る。最右辺の値が定義されるならば、それは 0 であろうが、そうすると結局は矛盾 $1 = 0$ を得てしまうので嬉しくない。細かいことを言えば、非可算個の項に対する総和の定義はそんなに自明ではない。可算個の総和というのはいわゆる無限級数なのでがんばれば扱えるのだが、非可算個の項からなる総和を考えることはあまりない。原因はおそらく次の命題にある； $\{a_\lambda\}_{\lambda \in \Lambda}$ を添字付けられた正の実数の集合とする。このとき、適切な定義のもとでその総和 $\sum a_\lambda$ が定義できるが、その総和が有限の値を取るならば、 Λ の濃度は高々可算である。

上記の観察を踏まえると、種々の操作ができるだけ可算無限程度で閉じていることを求めたくなる。その結果、ある集合 X が与えられた時、 X の任意の部分集合に対して（要求したい性質と両立するように）測度の値を定めることは難しくなる。のだが、殆どの状況において「任意の」部分集合に対する測度の値が欲しくなることはなく、例えば開集合や閉集合といった「ある程度常識的な」集合の測度を測れば充分であることも多い。このような実用上・理論上の要請をすり合わせることが Lebesgue 測度を含む測度の定式化には必要となる。

以下しばらく、 X を非空集合とする。

定義 3.5. $\mathcal{A} \subset 2^X$ が $(X$ 上の) σ -加法族であるとは、次の 3 条件を全て満たすことをいう。

- $\emptyset \in \mathcal{A}$.
- $A \in \mathcal{A}$ ならば、 $X \setminus A \in \mathcal{A}$.
- 任意の自然数 i に対して $A_i \in \mathcal{A}$ ならば、 $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

X とその上の σ -加法族の組 (X, \mathcal{A}) のことを可測空間といい、 \mathcal{A} の元のことを (\mathcal{A}) -可測集合という。

定義 3.6. あまり無限大や拡張実数について構成を厳密にやっても得るものがなさそうなのでここはいい加減にやる； $[0, \infty] := \{x \in \mathbb{R} \mid x \geq 0\} \cup \{\infty\}$ を片側拡張実数、 $\overline{\mathbb{R}} := [-\infty, \infty] := \mathbb{R} \cup \{\infty\} \cup \{-\infty\}$ を（両側）拡張実数という。任意の実数 $a \in \mathbb{R}$ に対して $-\infty \leq a \leq \infty$ が成り立つものとする。 $\overline{\mathbb{R}}$ 上の演算は、

- 任意の $x \in \mathbb{R}$ に対して $x \pm \infty = \pm\infty + x = \pm\infty$
- 任意の正の実数 x に対して $x \cdot \pm\infty = \pm\infty \cdot x = \pm\infty$
- 任意の負の実数 x に対して $x \cdot \pm\infty = \pm\infty \cdot x = \mp\infty$
- $\infty + \infty = \infty$, $-\infty - \infty = -\infty$, $0 \cdot \infty = \infty \cdot 0 = 0$

を充たすように \mathbb{R} の演算を拡張することで定める。ただし $\infty - \infty$ は定義しない。

余談 3.7. $0 \cdot \infty = 0$ と定めるのは積分を定義する際の便宜のためである。たぶんアルゴリズムの分野などでこの定義をするのは目くじらを立てられる可能性があるだろう。

定義 3.8. (X, \mathcal{A}) を可測空間とする. $\mu: \mathcal{A} \rightarrow [0, \infty]$ が測度であるとは, 次の 2 条件を充たすことをいう.

- $\mu(\emptyset) = 0$.
- (可算加法性) 高々可算個の \mathcal{A} の元からなる族 $\{A_i\}_{i=1}^{\infty}$ が互いに排反であるならば,
 $\mu(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mu(A_i)$.

3 つ組 (X, \mathcal{A}, μ) を測度空間という. $\mu(X) < \infty$ のとき, μ は有限測度であるといわれる. $\mu(A_i) < \infty$ ならびに $A_i \uparrow X$ を充たす集合列 $\{A_i\}$ が存在する時, μ は σ -有限測度であるといわれる.

命題 3.9. (X, \mathcal{A}, μ) を測度空間とする.

- (i) (単調性) $A, B \in \mathcal{A}$ が $A \subset B$ を充たすならば, $\mu(A) \leq \mu(B)$.
- (ii) (可算劣加法性) $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$ に対して, $\mu(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$.
- (iii) (連続性) $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$ に対して $A_i \uparrow A$ ならば $\lim_{i \rightarrow \infty} \mu(A_i) = \mu(A)$.
- (iv) (連続性) $\{A_i\}_{i=1}^{\infty} \subset \mathcal{A}$ に対して $A_i \downarrow A$ かつ $\mu(A_1) < \infty$ ならば $\lim_{i \rightarrow \infty} \mu(A_i) = \mu(A)$.

証明.

- (i) $B \setminus A \in \mathcal{A}$ なので測度の可算加法性より $\mu(B) = \mu(A) + \mu(B \setminus A) \geq \mu(A)$.
- (ii) $B_1 := A_1, B_i := (\bigcup_{k=1}^i A_k) \setminus B_{i-1} = A_i \setminus B_{i-1}$ によって帰納的に B_i を定めれば, $\{B_i\}_{i=1}^{\infty}$ は互いに排反で, $\bigcup_{i=1}^{\infty} B_i = \bigcup_{i=1}^{\infty} A_i$ ならびに $B_i \subset A_i$ を充たす. したがって $\mu(\bigcup_{i=1}^{\infty} A_i) = \mu(\bigcup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i) \leq \sum_{i=1}^{\infty} \mu(A_i)$.
- (iii) $B_1 := A_1, B_i := A_i \setminus A_{i-1}$ によって帰納的に B_i を定めれば, $\{B_i\}_{i=1}^{\infty}$ は互いに排反で, $A_i = \bigcup_{k=1}^i B_k$ ならびに $A = \bigcup_{k=1}^{\infty} B_k$ を充たす. 従って $\mu(A_i) = \mu(\bigcup_{k=1}^i B_k) = \sum_{k=1}^i \mu(B_k)$ であるから, $\mu(A) = \mu(\bigcup_{k=1}^{\infty} B_k) = \sum_{k=1}^{\infty} \mu(B_k) = \lim_{i \rightarrow \infty} \mu(A_i)$.
- (iv) $A_1 \setminus A_i \uparrow A_1 \setminus A$ なので, (iii) の結果より $\lim_{i \rightarrow \infty} \mu(A_1 \setminus A_i) = \mu(A_1 \setminus A)$. $\mu(A_1) < \infty$ なので $\mu(A_1) - \lim_{i \rightarrow \infty} \mu(A_i) = \mu(A_1) - \mu(A)$ すなわち $\lim_{i \rightarrow \infty} \mu(A_i) = \mu(A)$.

■

例 3.10. 簡単な測度の例をいくつか挙げておくことにしよう;

- $x \in X$ をひとつ固定する. $\mathcal{A} = 2^X$ として, $\mu: \mathcal{A} \rightarrow [0, 1]$ を

$$\mu(A) := \delta_x(A) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

と定めると, μ は \mathcal{A} 上の測度になる. このような測度を Dirac 測度という.

- X の点列 $\{x_n\}_{n=1}^{\infty}$ に対して $\mu := \sum_{n=1}^{\infty} \delta_{x_n}$ で定めると, μ は 2^X 上の測度になる. この測度を $\{x_n\}$ 上の数え上げ測度, あるいは計数測度という. 特に $X = \mathbb{N}$ として, $\mu(A) := \sum_{i \in A} 1 = \#A$

というのは「 A に含まれる元の数」を返す写像のことである。

- $\mathcal{A} = 2^X$ として, $\mu: \mathcal{A} \rightarrow [0, \infty]$ を

$$\mu(A) := \begin{cases} |A| & \text{if } |A| < \infty \\ \infty & \text{otherwise} \end{cases}$$

と定めると, μ は \mathcal{A} 上の測度になる. X が高々可算である場合を除いて, この測度は σ -有限ではない.

これらの測度は直接定義を与えることが比較的簡単にできるが, これは例外的な状況である. たいていの場合に σ -加法族は非常に大きな集合となり, それらの各元に対して測度の値を直接割り当てることは非常に困難である. それゆえ, 測度を定義する際は, より小さな集合族の上で測度の概形を構成しておいて, 「拡張定理」と呼ばれる定理によってそれを測度に拡張する, という戦略を取ることが多い. Lebesgue 測度もまたそのような戦略に則って構成される.

定義 3.11. $\mathcal{A} \subset 2^X$ が (X 上の) 有限加法族であるとは, 次の 3 条件を全て満たすことをいう.

- $\emptyset \in \mathcal{A}$.
- $A \in \mathcal{A}$ ならば, $X \setminus A \in \mathcal{A}$.
- $A_1, A_2, \dots, A_n \in \mathcal{A}$ ならば, $\bigcup_{i=1}^n A_i \in \mathcal{A}$.

定義 3.12. \mathcal{A} を X 上の有限加法族とする. $\mu: \mathcal{A} \rightarrow [0, \infty]$ が有限加法的測度であるとは, 次の 2 条件を満たすことをいう.

- $\mu(\emptyset) = 0$.
- 高々有限個の \mathcal{A} の元からなる族 $\{A_i\}_{i=1}^N$ が互いに排反であるならば, $\mu\left(\bigcup_{i=1}^N A_i\right) = \sum_{i=1}^N \mu(A_i)$.

$\mu(A_i) < \infty$ ならびに $A_i \uparrow X$ を満たす集合列 $\{A_i\}$ が存在する時, μ は \mathcal{A} 上で σ -有限であるといわれる.

問 3.13. 有限加法的測度は単調性と劣加法性を満たす. 証明は (ちゃんと書き下してないけど, 多分) 測度の場合と同様 (だと思う).

余談 3.14. 後に見るが, 左半開区間 $(a, b]$ に対して $\mu((a, b]) := b - a$ と定め, これを $\overline{\mathcal{H}}$ 上の有限加法的測度に拡張すれば Lebesgue 測度の「概形」が得られる. この有限加法的測度のことを Jordan 測度ということもある. きちんと確かめていないが, Spivak における「容量 0」概念は Jordan 測度 0 のことを指している可能性が高い.

命題 3.15. $C \subset 2^X$ を部分集合の族とする. この時, C を含む σ -加法族の中で (集合の包含関係に関して) 最小のものが一意に存在する. これを C の生成する σ -加法族といい, $\sigma(C)$ と書く.

証明. σ -加法族は集合の共通部分を取る操作について閉じている；すなわち $\{\mathcal{A}_\lambda\}_{\lambda \in \Lambda}$ が σ -加法族からなる族であるならば、 $\bigcap_{\lambda \in \Lambda} \mathcal{A}_\lambda$ もまた σ -加法族である。したがって C を含む σ -加法族全体の集合を $D := \{\mathcal{A} \mid \mathcal{A} \text{ は } C \text{ を含む } \sigma\text{-加法族}\}$ として、 $\sigma(C) := \bigcap_{\mathcal{A} \in D} \mathcal{A}$ とおけばよい。 ■

問 3.16. ちゃんと確かめていないが、 C を含む σ -加法族全体はおそらく集合になるので上記の証明で問題ないと思う。たぶん濃度を $|2^{2^X}|$ で抑えられる気がする。「この性質を充たすものぜんぶとってきて！」という無邪気なことをするとたまに集合でないものができあがるので気をつけないといけない。

定理 3.17 (Carathéodory の拡張定理). \mathcal{A} を有限加法族、 $\mu: \mathcal{A} \rightarrow [0, \infty]$ を有限加法的測度とする。更に μ は次の意味で \mathcal{A} 上可算加法的であると仮定する；高々可算個の \mathcal{A} の元からなる族 $\{A_i\}_{i=1}^\infty$ が互いに排反であり、かつ $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$ であるならば、 $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$ 。このとき、 μ は $\sigma(\mathcal{A})$ 上の測度に拡張できる。更に、 μ が \mathcal{A} 上で σ -有限ならば、この拡張は一意的である。

拡張定理の証明はこの時点でできるのだが、Lebesgue 測度をできるだけはやく定義したいので、証明は後の節に回す。ただし、拡張の一意性の証明に用いる π - λ 定理については、拡張定理以外にも使い所があるので、この時点でステートメントだけは紹介しておくことにする。

定義 3.18. $\mathcal{P} \subset 2^X$ が π -システムであるとは、次の 2 条件を充たすことをいう。(i) $\mathcal{P} \neq \emptyset$, (ii) $A, B \in \mathcal{P}$ ならば $A \cap B \in \mathcal{P}$ 。

$\mathcal{L} \subset 2^X$ が λ -システムであるとは、次の 3 条件を充たすことをいう。(i) $X \in \mathcal{L}$, (ii) $A, B \in \mathcal{L}$ かつ $A \subset B$ ならば $B \setminus A \in \mathcal{L}$, (iii) $A_i \uparrow A$ かつ任意の $i \in \mathbb{N}$ に対して $A_i \in \mathcal{L}$ であるならば、 $A \in \mathcal{L}$ 。

問 3.19. 有限加法族は π -システムである。 σ -加法族は π -システムであり λ -システムでもある。

定理 3.20 (Dynkin の π - λ 定理). $\mathcal{P} \subset 2^X$ を π -システム、 $\mathcal{L} \subset 2^X$ を λ -システムとし、更に $\mathcal{P} \subset \mathcal{L}$ であると仮定する。このとき $\sigma(\mathcal{P}) \subset \mathcal{L}$ 。すなわち $\sigma(\mathcal{P})$ は \mathcal{P} を含む λ -システムの中で最小のものである。

命題 3.21. \mathbb{R} の左半开区間をすべて集めた集合 $\mathcal{H} := \{(a, b) \mid -\infty \leq a < b < \infty\} \cup \{(a, \infty) \mid -\infty \leq a < \infty\} \cup \{\emptyset\}$ は有限加法族ではないが、

$$\overline{\mathcal{H}} := \left\{ \bigcup_{i=1}^N A_i \mid A_i \in \mathcal{H}, A_i \text{ は排反} \right\}$$

は有限加法族である。更に $\sigma(\mathcal{H}) = \sigma(\overline{\mathcal{H}})$ である。

証明. 左半開区間 2 つの和集合は排反な左半開区間 2 つになるか、ひとつの左半開区間になる。また左半開区間の補集合は（高々 2 つの）左半開区間の和で表せる。従って $\overline{\mathcal{H}}$ は有限加法族である。 $\sigma(\mathcal{H}) \subset \sigma(\overline{\mathcal{H}})$ はよく、 σ -加法族の定義から $\overline{\mathcal{H}} \subset \sigma(\mathcal{H})$ なので π - λ 定理によって $\sigma(\mathcal{H}) \supset \sigma(\overline{\mathcal{H}})$ を得る。 ■

命題 3.22. 測度 $\lambda: \sigma(\mathcal{H}) \rightarrow [0, \infty]$ であって、以下の 3 条件を充たすものが一意的に存在する。この測度 λ を \mathbb{R} 上の Lebesgue 測度と呼ぶ。

- 任意の $(a, b] \in \mathcal{H}$ に対して $\lambda((a, b]) = b - a$.
- 任意の $(a, \infty) \in \mathcal{H}$ に対して $\lambda((a, \infty)) = \infty$.
- $\lambda(\emptyset) = 0$.

証明. $\mu: \mathcal{H} \rightarrow [0, \infty]$ を、 $(a, b] \in \mathcal{H}$ に対して $\mu((a, b]) := b - a$, $(a, \infty) \in \mathcal{H}$ に対して $\mu((a, \infty)) := \infty$, $\mu(\emptyset) := 0$ と定める。更に、排反な $A_1, A_2, \dots, A_n \in \mathcal{H}$ に対して $\mu\left(\bigcup_{i=1}^n A_i\right) := \sum_{i=1}^n \mu(A_i)$ と定めることで μ を $\overline{\mathcal{H}}$ 上の有限加法的測度に拡張できたとする。その上で $\overline{\mathcal{H}}$ 上で可算加法的であることが示せれば、 μ は $\overline{\mathcal{H}}$ 上 σ -有限なので Carathéodory の拡張定理によって結論を得る。以下、これらのことを示していく。

まず μ が $\overline{\mathcal{H}}$ 上の有限加法的測度に拡張できること、すなわち上記の拡張が well-defined であることを示す。 $A \in \overline{\mathcal{H}}$ をとると、 A は排反な集合族 $A_1, A_2, \dots, A_m \in \mathcal{H}$ を用いて $A = \bigcup_{i=1}^m A_i$ と表されている。ここで、別の排反な集合族 $B_1, B_2, \dots, B_n \in \mathcal{H}$ を用いて $A = \bigcup_{j=1}^n B_j$ と表されていたとする。このとき $1 \leq i \leq m$ ならびに $1 \leq j \leq n$ に対して $C_{ij} := A_i \cap B_j$ と置くと、 $\{C_{ij}\}$ は排反な $\overline{\mathcal{H}}$ の集合族であって、 $\bigcup_{i=1}^m C_{ij} = B_j$ および $\bigcup_{j=1}^n C_{ij} = A_i$ を充たす。したがって $\mu\left(\bigcup_{i=1}^n A_i\right) = \mu\left(\bigcup_{i=1}^n \bigcup_{j=1}^m C_{ij}\right) = \mu\left(\bigcup_{j=1}^n B_j\right)$ を得るから μ は well-defined である。

続いて μ の \mathcal{H} 上での可算加法性、すなわち $A \in \mathcal{H}$ が排反な区間の族 $\{(a_i, b_i]\}_{i=1}^\infty$ を用いて $A = \bigcup_{i=1}^\infty (a_i, b_i]$ と表されていたときに $\mu(A) = \sum_{i=1}^\infty \mu((a_i, b_i])$ となることを示す。

- $-\infty < a < b < \infty$ を用いて $A = (a, b]$ と表せる場合。まず任意の $n > 0$ に対して $\bigcup_{i=1}^n (a_i, b_i] \subset (a, b]$ なので $\mu\left(\bigcup_{i=1}^n (a_i, b_i]\right) \leq \mu((a, b])$ である。更に $(a_i, b_i]$ たちの排反性より $\mu\left(\bigcup_{i=1}^n (a_i, b_i]\right) = \sum_{i=1}^n \mu((a_i, b_i])$ である。結局 $\sum_{i=1}^n \mu((a_i, b_i]) \leq \mu((a, b])$ を得るので、 $n \rightarrow \infty$ として $\sum_{i=1}^\infty \mu((a_i, b_i]) \leq \mu((a, b])$ である。逆向きの不等式を示す。 $\varepsilon > 0$ を任意に固定し、 $\eta_m := \varepsilon/2^m$ と置く。このとき $\{(a_i, b_i + \eta_i)\}$ は $[a + \varepsilon, b]$ の開被覆なので、Heine-Borel の定理よりある $n \in \mathbb{N}$ が存在して、 $(a + \varepsilon, b] \subset [a + \varepsilon, b] \subset \bigcup_{i=1}^n (a_i, b_i + \eta_i) \subset \bigcup_{i=1}^n (a_i, b_i + \eta_i]$

である。以上の結果より,

$$\begin{aligned}
\mu((a, b]) &\leq \mu\left((a, a + \varepsilon] \cup \bigcup_{i=1}^n (a_i, b_i + \eta_i]\right) \\
&\leq \varepsilon + \mu\left(\bigcup_{i=1}^n (a_i, b_i + \eta_i]\right) \\
&\leq \varepsilon + \sum_{i=1}^n (\mu((a_i, b_i]) + \eta_i) \\
&\leq 2\varepsilon + \sum_{i=1}^n (\mu((a_i, b_i])) \\
&\leq 2\varepsilon + \sum_{i=1}^{\infty} (\mu((a_i, b_i]))
\end{aligned}$$

を得る。 $\varepsilon > 0$ は任意だったので $\mu((a, b]) \leq \sum_{i=1}^{\infty} \mu((a_i, b_i])$ である。

- $b \in \mathbb{R}$ を用いて $A = (-\infty, b]$ と書ける場合。 $\mu(A) = \infty$ なので $\mu(A) \geq \sum_{i=1}^{\infty} \mu((a_i, b_i])$ はよい。 $c \in \mathbb{R}$ に対して $(c, b] = \bigcup_{i=1}^{\infty} ((c, b] \cap (a_i, b_i])$ なので、前段の結果と併せて $\mu((c, b]) = \bigcup_{i=1}^{\infty} \mu((c, b] \cap (a_i, b_i]) \leq \sum_{i=1}^{\infty} \mu((a_i, b_i])$ を得るから $c \rightarrow \infty$ として $\mu(A) \leq \sum_{i=1}^{\infty} \mu((a_i, b_i])$ を得る。 $A = (a, \infty)$ ならびに $A = \mathbb{R}$ の場合も同様である。

最後に、 μ が $\overline{\mathcal{H}}$ 上で可算加法的であることを示す。排反な $\overline{\mathcal{H}}$ の元の族 $\{A_i\}_{i=1}^{\infty}$ に対して $A := \bigcup_{i=1}^{\infty} A_i \in \overline{\mathcal{H}}$ であったと仮定する。 $A \in \overline{\mathcal{H}}$ なので、 $A = \bigcup_{j=1}^m B_j$ を充たす排反な $B_1, B_2, \dots, B_m \in \mathcal{H}$ がある。同様に $A_i \in \overline{\mathcal{H}}$ なので、 $A_i = \bigcup_{k=1}^{n_i} C_{ik}$ を充たす排反な $C_{i1}, C_{i2}, \dots, C_{in_i} \in \mathcal{H}$ がある。 $B_j \cap C_{ik}$ が互いに排反であることに気をつけて、

$$\begin{aligned}
\mu(A) &= \sum_{j=1}^m \mu(B_j) \\
&= \sum_{j=1}^m \mu\left(\bigcup_{i=1}^{\infty} \bigcup_{k=1}^{n_i} B_j \cap C_{ik}\right) \\
&= \sum_{j=1}^m \sum_{i=1}^{\infty} \sum_{k=1}^{n_i} \mu(B_j \cap C_{ik}) \\
&= \sum_{i=1}^{\infty} \sum_{j=1}^m \sum_{k=1}^{n_i} \mu(B_j \cap C_{ik}) \\
&= \sum_{i=1}^{\infty} \mu(A_i)
\end{aligned}$$

となる。 ■

問 3.23 (*). ここでは $\lambda((a, b]) = b - a$ と定めたが、実際には非減少かつ右連続な関数 $F: \mathbb{R} \rightarrow \mathbb{R}$ に対して $\lambda((a, b]) := F(b) - F(a)$ と定めた場合でも上記の証明がほぼ同様に通る、 $\sigma(\mathcal{H})$ 上の測度が

得られる。これは F から定まる Lebesgue-Stieltjes 測度と呼ばれる。

ここまで、 $\overline{\mathcal{H}}$ という左半開区間の和集合からなる集合を考えてきたが、これは $\overline{\mathcal{H}}$ が補集合を取る操作に関して閉じているという技術的事情による。例えば開集合や閉集合が全て含まれているのか、ということが気になると思うが、実際にはその点を心配する必要はなく、 $\sigma(\overline{\mathcal{H}})$ は多くの集合を含む；

定義 3.24. \mathbb{R} の開集合をすべて含む最小の σ -加法族を \mathbb{R} の Borel σ -加法族といい、 \mathcal{B} であらわす。すなわち、 \mathbb{R} の開集合をすべて集めた集合を \mathcal{O} と書くとき、 $\mathcal{B} := \sigma(\mathcal{O})$ と定める。

問 3.25. \mathbb{R} の開集合は \mathbb{R} の開区間の可算合併で書けるので、実際には $\mathcal{B} = \sigma(\{(a, b) \mid -\infty < a < b < \infty\})$ である。

余談 3.26. 実際には位相空間に対して Borel σ -加法族という言葉が定義できる；位相空間 X に対し、その開集合全体が生成する σ -加法族のことを X の Borel σ -加法族という。

命題 3.27. $\mathcal{B} = \sigma(\mathcal{H})$ 。

証明. $(a, b] \in \mathcal{H}$ を任意にとると、 $(a, b] = \bigcap_{n=1}^{\infty} (a, b + 1/n)$ と表せるので、 $\mathcal{H} \subset \mathcal{B}$ すなわち $\sigma(\mathcal{H}) \subset \mathcal{B}$ である。逆に $(a, b) \in \mathcal{B}$ を任意にとると $(a, b) = \bigcup_{n=1}^{\infty} (a, b - 1/n]$ と表せるので $\mathcal{B} \subset \sigma(\mathcal{H})$ であるから $\sigma(\mathcal{B}) \subset \sigma(\mathcal{H})$ を得る。 ■

3.2 可測関数

1 変数関数の Riemann 積分がどのような関数に対して定義できたわけではないのと同様に、測度に関する積分もまた、任意の関数に定義できるわけではない。測度に関する積分が定義できる関数のクラスが可測写像と呼ばれるものである。本節では可測写像の定義ならびに基本的な性質を見ていく。それと並行して、拡張実数や \mathbb{R}^n の Borel σ -加法族の構造についても調べる。とりあえず抑えておくべきことがあるとするならば、(Euclid 空間における) 可測写像は連続写像よりもより広いクラスになるということ、可測性は（連続性や微分可能性とは異なって）極限操作に関して保存されるということの 2 点であろう。

以下しばらく、 (X, \mathcal{A}) ならびに (X_i, \mathcal{A}_i) ($i = 1, 2, 3, \dots$) を可測空間とする。

定義 3.28. 写像 $f: X_1 \rightarrow X_2$ が $(\mathcal{A}_1/\mathcal{A}_2)$ -可測であるとは、任意の $A \in \mathcal{A}_2$ に対して $f^{-1}(A) \in \mathcal{A}_1$ となることをいう。

余談 3.29. この定義は「開集合の逆像が開集合」という連続写像の特徴づけと併置して考えるのが

よいと思っている。開集合という位相的な構造を壊さない写像のことが連続写像であったのを真似るように、可測空間の構造を壊さない写像のことを可測写像として定義している。

σ -加法族が一般に複雑であるという話をした手前、各可測集合の逆像がどうなっているかを確かめるのは煩雑に思えるかも知れない。が、実際にはそのような手間を踏む必要はないことのほうが多い；

補題 3.30. \mathcal{A}_2 が $C \subset 2^{X_2}$ によって生成されている時、すなわち $\mathcal{A}_2 = \sigma(C)$ のとき、 $f: X_1 \rightarrow X_2$ が可測であることと、任意の $C \in C$ に対して $f^{-1}(C) \in \mathcal{A}_1$ であることは同値である。

証明. 充分性は明らか。 $\mathcal{D} := \{A \in \mathcal{A}_2 \mid f^{-1}(A) \in \mathcal{A}_1\}$ と置くと、逆像は補集合・共通部分・和集合を取る操作を保存することから \mathcal{D} は σ -加法族であり、かつ $C \subset \mathcal{D}$ 。したがって $\sigma(C) \subset \mathcal{D}$ なのでよい。 ■

補題 3.31. $f: X \rightarrow \mathbb{R}$ に対して以下はすべて同値である。

- (i) f は可測。
- (ii) 任意の $a \in \mathbb{R}$ に対して $f^{-1}((-\infty, a)) \in \mathcal{A}$ 。
- (iii) 任意の $a \in \mathbb{R}$ に対して $f^{-1}((-\infty, a]) \in \mathcal{A}$ 。
- (iv) 任意の $a \in \mathbb{R}$ に対して $f^{-1}((a, \infty)) \in \mathcal{A}$ 。
- (v) 任意の $a \in \mathbb{R}$ に対して $f^{-1}([a, \infty)) \in \mathcal{A}$ 。

証明. 議論はほぼ同様なので、(i) と (ii) の同値性のみ示す。 $\mathcal{B} = \sigma(\{(-\infty, a) \mid a \in \mathbb{R}\})$ を示せばよい。任意の $a \in \mathbb{R}$ に対して $(-\infty, a) \in \mathcal{B}$ なので $\mathcal{B} \supset \sigma(\{(-\infty, a) \mid a \in \mathbb{R}\})$ はよい。任意の $n \in \mathbb{N}$ ならびに开区間 (a, b) に対して $(a, b) = \bigcap_{n=1}^{\infty} ((-\infty, b) \cap (-\infty, a + 1/n)^c)$ なので $(a, b) \in \sigma(\{(-\infty, a) \mid a \in \mathbb{R}\})$ となるから $\mathcal{B} \subset \sigma(\{(-\infty, a) \mid a \in \mathbb{R}\})$ である。 ■

実数に値を持つ関数が主たる関心である場合でも、積分や極限に関する種々の操作をする中で拡張実数値関数を考えたほうがいい場合がある。たとえば実数値関数の族 $\{f_n\}_{n=1}^{\infty}$ があるときに $\sup_n f_n(x)$ が有限であるとは限らないし、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ に対して $y \mapsto \int f(x, y) dx$ という関数は（適当な意味で積分が定義できるとはして） f の形状いかんでは有限にならないことがある。

余談 3.32. このノートの範囲に限れば、前者が問題になることはないと思う。なので拡張実数値関数のことを（測度の定義を除いて）陽に取り扱う必要はなさそうに見えていた。のだが、後者の問題を捨て置くわけにはいかないと後で気づいた。というのも、この形の関数が Fubini の定理の主張に直接的に現れるのである。

というわけで拡張実数 $\overline{\mathbb{R}}$ の Borel σ -加法族について考えることになるのだが、そもそも $\overline{\mathbb{R}}$ の開

集合の話をしていなかったで、そこから話をしよう。とはいえこのノートでは位相の一般論をわざと表に出さないで、以下の議論はややアドホックではある。

定義 3.33. $\overline{\mathbb{R}}$ の開集合は以下のルールによって帰納的に定める。

- \mathbb{R} の開集合は $\overline{\mathbb{R}}$ の開集合である。
- $a \in \mathbb{R}$ に対して、 $[-\infty, a)$ および $(a, \infty]$ は $\overline{\mathbb{R}}$ において開集合である。
- U, V が $\overline{\mathbb{R}}$ において開集合であるならば、 $U \cap V$ も $\overline{\mathbb{R}}$ において開集合である。
- $\{U_\lambda\}_{\lambda \in \Lambda}$ が $\overline{\mathbb{R}}$ における開集合の族であるならば、 $\bigcup_{\lambda \in \Lambda} U_\lambda$ は $\overline{\mathbb{R}}$ において開集合である。
- $\overline{\mathbb{R}}$ の開集合は上記で定まるものに限る。

$\overline{\mathbb{R}}$ の開集合を全て含む最小の σ -加法族を $\overline{\mathbb{R}}$ の Borel σ -加法族といい、 \mathcal{B}^* と書く。

問 3.34 (*). すでに白状したとおり、このように位相を指定するのは一般的ではないとも思う（その位相の定め方は恣意的ではないか？という疑問に答えられない）。実際には次のようにしたほうが多少は行儀がよいだろう；逆正接関数 $\arctan: \mathbb{R} \rightarrow (-1, 1)$ を、 $\arctan(-\infty) = -\pi/2$ および $\arctan(\infty) = \pi/2$ と定めることで $\arctan: \overline{\mathbb{R}} \rightarrow [-\pi/2, \pi/2]$ に拡張し、この写像による誘導位相を $\overline{\mathbb{R}}$ に入れる。ということで $\overline{\mathbb{R}}$ は $[-\pi/2, \pi/2]$ と同相なので、距離化可能なコンパクト位相空間である。

問 3.35 ()**. とはいえこれ以外に $\overline{\mathbb{R}}$ に定める位相構造はありえないのか？というのはよくわからない（なんとなく自然なものはない気がするが、証明が思いついているわけではない）。少なくとも $\overline{\mathbb{R}}$ が \mathbb{R} に定める相対位相と、 \mathbb{R} の通常の位相が一致してほしいと思う（そうでなければ、 \mathbb{R} -値関数を $\overline{\mathbb{R}}$ -値関数だと思う際に位相ならびに可測性に関する議論が煩雑になる）が、この点は充たしているし、 \arctan を連続拡張する位相になっているので自然ではあると思う（個人の感想です）。 $\overline{\mathbb{R}}$ に完備距離化可能な位相を入れる限り可測構造はいやでも一意に決まるので（後述する Kuratowski の定理による）、可測構造だけを気にするのであれば特に何も考えなくてもよいのかもしれないが…

拡張実数値関数に対しても、実数値関数の場合と同様に次の事実がある。証明は実数値関数の場合とほとんど同じなので省きます。

問 3.36. $f: X \rightarrow \overline{\mathbb{R}}$ に対して以下はすべて同値である。

- f は可測。
- 任意の $a \in \mathbb{R}$ に対して $f^{-1}([-\infty, a)) \in \mathcal{A}$ 。
- 任意の $a \in \mathbb{R}$ に対して $f^{-1}([-\infty, a]) \in \mathcal{A}$ 。
- 任意の $a \in \mathbb{R}$ に対して $f^{-1}((a, \infty]) \in \mathcal{A}$ 。
- 任意の $a \in \mathbb{R}$ に対して $f^{-1}([a, \infty]) \in \mathcal{A}$ 。

拡張実数に次いで、 \mathbb{R}^n の可測空間としての構造を考える。 \mathbb{R} を真似て Borel σ -加法族を与えることも勿論できる。が、一般的に、ふたつの可測空間の直積に対して可測空間の構造を作ることができるので、まずその話から始めよう；

定義 3.37. (X_1, \mathcal{A}_1) ならびに (X_2, \mathcal{A}_2) を可測空間とする。 $X_1 \times X_2$ 上の σ -加法族

$$\mathcal{A}_1 \otimes \mathcal{A}_2 := \sigma(\{A_1 \times A_2 \mid A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\})$$

を \mathcal{A}_1 と \mathcal{A}_2 の直積 σ -加法族という。 n 個の可測空間に対する直積 σ -加法族 $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ も同様に定める。特に $\mathcal{A}_1 = \mathcal{A}_2 = \cdots = \mathcal{A}_n =: \mathcal{A}$ のとき、 $\mathcal{A}_1 \otimes \cdots \otimes \mathcal{A}_n$ を \mathcal{A}^n と書く。

余談 3.38. 直積 σ -加法族のことを $\mathcal{A}_1 \times \mathcal{A}_2$ と書く文献も多い。これは直積集合と紛らわしい気もするが、 σ -加法族の直積集合を考えることはあまりないからか、一般的な記法である。他方、 $\mathcal{A}_1 \otimes \mathcal{A}_2$ は（後に現れる）テンソル積と記号が衝突しているけれども、この記号衝突が混乱を引き起こすこともないと思う。個人的な所感としてはどっちもどっちなので、気分でお \otimes を使うことにしたが、 \mathcal{A}^n を $\mathcal{A}^{\otimes n}$ とは書いていないので、お行儀が悪いと感じる人もいるかもしれない。

この時点で、 \mathbb{R}^n には 2 つの σ -加法族の入れ方があり得るとわかる。すなわち、直積 σ -加法族 \mathcal{B}^n ならびに \mathbb{R}^n の開集合から生成される σ -加法族 $\mathcal{B}(\mathbb{R}^n)$ である。あり得る構造が 2 つもあると話が煩雑になりそうな予感がするが、実際にはこの 2 つは同じものなのでやはり心配する必要はない；

命題 3.39. $\mathcal{B}^n = \mathcal{B}(\mathbb{R}^n)$.

証明. $n = 2$ の場合を示す。 $\{A \mid A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)\}$ は \mathbb{R} の開集合をすべて含む σ -加法族なので、 \mathcal{B} を含む。従って任意の $A \in \mathcal{B}$ に対して $A \times \mathbb{R} \in \mathcal{B}(\mathbb{R}^2)$ が成り立つ。同様に任意の $B \in \mathcal{B}$ に対して $\mathbb{R} \times B \in \mathcal{B}(\mathbb{R}^2)$ が成り立つ。ゆえに任意の $A, B \in \mathcal{B}$ に対して $A \times B = (A \times \mathbb{R}) \cap (\mathbb{R} \times B) \in \mathcal{B}(\mathbb{R}^2)$ を得るので、 $\mathcal{B}^2 \subset \mathcal{B}(\mathbb{R}^2)$ 。逆の包含を示すには $\mathcal{B} = \sigma(\{(a_1, b_1) \times (a_2, b_2) \mid -\infty < a_i < b_i < \infty, i = 1, 2\})$ を示せばよいが、これは \mathbb{R}^2 の開集合が開方体の可算和で書けることから従う。 $n > 2$ の場合の証明もほぼ同様である。 ■

問 3.40 (*). \mathbb{R}^2 の開集合は開方体の可算和で書ける。これは位相の言葉を陽に使うと、 \mathbb{R}^2 は可算開基を持つ、ということにほかならない。可算開基を持つ位相空間は第二可算空間と呼ばれる。距離空間に限れば、第二可算性と可分性（可算な稠密部分集合を持つこと）は同値である。

問 3.41 ().** この議論は可分位相空間にほとんどそのまま一般化できる。可分でない位相空間においては、ここで考えた 2 つの σ -加法族は一致するとは限らない（一般には直積位相に関する σ -加法族のほうが大きくなる）。これが原因で、可分でない空間に対しては「連続ならば（Borel σ -加法族に関して）可測」という、 \mathbb{R}^n の場合に成り立ってくれる性質が成り立たない可能性が現れる。可

分でない空間を考えなければよさそうにも見えるし、個人的にはそうしたいのだが、実際にはノンパラメトリック推定などの文脈で関数空間に値を取る確率変数を考えはじめるとそういうわけにもいかない。

ここまでの議論をもとに、写像が可測になる状況のいくつかを挙げることができる。主張を見ればわかるように、極限操作を含む様々な操作に対して可測性は頑健である。

命題 3.42.

- (i) $f: \mathbb{R}^m \rightarrow \mathbb{R}^n$ が連続であるならば、それは $\mathcal{B}^m/\mathcal{B}^n$ -可測である。
- (ii) $f^1, f^2, \dots, f^n: X_1 \rightarrow \mathbb{R}$ が可測であることと、 $f = (f^1, f^2, \dots, f^n): X_1 \rightarrow \mathbb{R}^n$ が可測であることは同値である。
- (iii) $f: X_1 \rightarrow X_2$ ならびに $g: X_2 \rightarrow X_3$ がそれぞれ可測であるならば、合成写像 $g \circ f: X_1 \rightarrow X_3$ も可測である。
- (iv) $g: \mathbb{R}^n \rightarrow \mathbb{R}$ ならびに $f^1, f^2, \dots, f^n: X \rightarrow \mathbb{R}$ がすべて可測であるならば、 $X_1 \ni x \mapsto g(f^1(x), f^2(x), \dots, f^n(x)) \in \mathbb{R}$ は可測である。従って $f, g: X \rightarrow \mathbb{R}$ が可測であれば、 $f + g, fg$ などとは全て可測である。
- (v) $f_1, f_2, \dots: X \rightarrow \overline{\mathbb{R}}$ が可測であるとする。 $\sup_n f_n, \inf_n f_n, \limsup_n f_n, \liminf_n f_n$ は可測である。したがって任意の $x \in X$ に対して $\lim_n f_n(x)$ が $\pm\infty$ を含めて存在するならば、 $f: X \ni x \mapsto \lim_n f_n(x) \in \mathbb{R}$ は可測である。

証明.

- (i) 任意の開集合 $O \subset \mathbb{R}^n$ に対して $f^{-1}(O) \in \mathcal{B}^m$ なので、開集合全体が \mathcal{B}^n を生成することと前補題から結論が従う。
- (ii) $n = 2$ の場合のみ示す。任意の $A, B \in \mathcal{B}$ に対して、 $f^{-1}(A \times B) \in \mathcal{A}_1$ は $(f^1)^{-1}(A) \cap (f^2)^{-1}(B) \in \mathcal{A}_1$ と同値であることに注意する。このことから f^1, f^2 が可測ならば f が可測であるとわかる。逆に f が可測であると仮定すると、特に $B = \mathbb{R}$ とすることで $(f^1)^{-1}(A) \in \mathcal{A}_1$ を得るから f^1 は可測である。 f^2 の可測性も同様である。
- (iii) $A \in \mathcal{A}_3$ を任意にとってくると、 $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$ であり、 g の可測性から $g^{-1}(A) \in \mathcal{A}_2$ を得るから、 f の可測性より $(g \circ f)^{-1}(A) \in \mathcal{A}_1$ 。
- (iv) (ii) ならびに (iii) から出る。
- (v) $(\inf_n f_n)^{-1}((-\infty, a)) = \bigcap_n (f_n)^{-1}((-\infty, a)) \in \mathcal{A}$ なので $\inf_n f_n$ は可測である。 $\sup_n f_n$ も同様。 $\liminf_n f_n = \sup_n \inf_{m \geq n} f_m$ なので $\liminf_n f_n$ の可測性もよい。ほかも同様である。

■

問 3.43 (*). ということでたいていの場合に写像は可測になるという話をしてきたが、だからと

いて任意の関数が可測になるわけではない。このノートの範囲で可測性が問題になることはおそくないので、証明まで含めて説明することはしないが、以下のような例がある；

- (i) $x, y \in [0, 1)$ 上の二項関係 \sim を $x \sim y \Leftrightarrow x - y \in \mathbb{Q}$ で定めると、これは同値関係である。この同値関係に関する代表元をみつめた集合 V を Vitali 集合という。 $V \notin \mathcal{B}$ である。したがって、その指示関数 $1_V: \mathbb{R} \rightarrow \mathbb{R}$ は \mathcal{B}/\mathcal{B} -可測ではない。
- (ii) $\{f_\lambda: X \rightarrow \mathbb{R}\}_{\lambda \in \Lambda}$ を \mathbb{R} -値可測関数の族とすると、 Λ が非可算であるならば、 $\sup_{\lambda \in \Lambda} f_\lambda(x)$ は可測とは限らない。例えば Λ として V を取り、 $\{1_x\}_{x \in V}$ を考えればよい。

ところで、ここまでで \mathbb{R}^n 上の σ -加法族に関する理解は深まった気がする一方、それ以外の σ -加法族については相変わらずわからないままである。そもそも可測構造というのはどれくらいいろいろなものがあるのだろうか、というのは自然な疑問の一つだと思う。このノートでは証明まで踏み込むことはしないものの、かなり面白い（と個人的には思っている）一般論があるので述べておくことにしよう。ちなみにこの話は私は聞きかじっただけでちゃんと勉強したことがあるわけではない。

定義 3.44. (X_1, \mathcal{A}_1) と (X_2, \mathcal{A}_2) が（可測空間として）同型であるとは、全単射な $\mathcal{A}_1/\mathcal{A}_2$ -可測写像 $f: X_1 \rightarrow X_2$ があり、その逆写像が $\mathcal{A}_2/\mathcal{A}_1$ -可測であることをいう。可測空間 (X, \mathcal{A}) が標準 Borel 空間であるとは、ある完備距離空間ならびにその Borel σ -加法族によって定まる可測空間と同型であることをいう。

問 3.45 (Kuratowski, **). (X, \mathcal{A}) を標準 Borel 空間とする。

- (i) $|X| = n$ であれば (X, \mathcal{A}) は $(\{1, 2, \dots, n\}, 2^{\{1, 2, \dots, n\}})$ と同型である。
- (ii) X が可算無限であれば (X, \mathcal{A}) は $(\mathbb{N}, 2^{\mathbb{N}})$ と同型である。
- (iii) (i), (ii) のいずれも当てはまらないならば、 (X, \mathcal{A}) は $(\mathbb{R}, \mathcal{B})$ と同型である。

というわけで、少なくとも完備距離空間くらいに限れば、可測構造というのは随分種類が少ないらしい。

3.3 積分と各種の収束定理

(X, \mathcal{A}, μ) を測度空間とする。可測写像 $f: X \rightarrow \overline{\mathbb{R}}$ に対し、その積分を定義する。一旦全て定義を述べてしまい、論理の飛躍は後で埋める。

定義 3.46. $A \in \mathcal{A}$ に対して、

$$1_A(x) := \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases}$$

で定まる関数 $1_A: X \rightarrow \mathbb{R}$ を A の指示関数という。 $f: X \rightarrow \mathbb{R}$ が単関数であるとは、 $a_1, a_2, \dots, a_n \in$

\mathbb{R} ならびに $A_1, A_2, \dots, A_n \in \mathcal{A}$ を用いて $f = \sum_{i=1}^n a_i 1_{A_i}$ と表せることをいう。

定義 3.47 (積分の定義).

- (非負単関数の場合) 非負単関数 $f = \sum_{i=1}^n a_i 1_{A_i}$ に対して $\int f d\mu := \sum_{i=1}^n a_i \mu(A_i)$ と定める. well-defined であることは非自明だが, 後で証明する.
- (非負可測関数の場合) 可測関数 $f: X \rightarrow [0, \infty]$ に対して,

$$\int f d\mu := \sup \left\{ \int g d\mu \mid 0 \leq g \leq f, g \text{ は単関数} \right\}$$

で定める.

- (一般の場合) 可測関数 $f: X \rightarrow \overline{\mathbb{R}}$ に対して, $f^+ := \max\{f, 0\}$, $f^- := -\min\{f, 0\}$ と定めると, この2つは非負可測関数であって, $f = f^+ - f^-$ である. $\int f^+ d\mu$ ならびに $\int f^- d\mu$ の少なくとも片方が有限であるとき, f の積分が定義できるといい,

$$\int f d\mu := \int f^+ d\mu - \int f^- d\mu$$

で定める. 両方が無限大になった場合の積分は定義しない.

- 積分が定義できる関数 $f: X \rightarrow \mathbb{R}$ ならびに $A \in \mathcal{A}$ に対して, f の A 上での積分を

$$\int_A f d\mu := \int f 1_A d\mu$$

で定める. f の積分が定義できるならば $\int_A f d\mu$ も定まるが, このことは後で示す.

関数 $f: X \ni x \mapsto f(x) \in \overline{\mathbb{R}}$ の積分について取り扱う際, f の引数の記号として x を用いていることを明示したい場合は, $\int f d\mu$ のことを $\int f(x) d\mu(x)$ や $\int f(x) \mu(dx)$ のように書く.

定義 3.48. 可測関数 $f: X \rightarrow \mathbb{R}$ は, $|f|$ の積分が有限の値を取るとき, すなわち $\int f^+ d\mu + \int f^- d\mu < \infty$ のとき, (絶対) 可積分であるという. 絶対可積分ならば f の積分が定義でき, それは有限の値になる. 自然数 p に対し $|f|^p$ の積分が有限の値を取るような関数は p 乗可積分であると言われる.

(X, \mathcal{A}, μ) 上の p 乗可積分関数全体の集合を $L^p(X, \mathcal{A}, \mu)$ と書く. 前後の文脈から明らかな場合や誤解の恐れがない場合は $L^p(\mu)$ あるいは L^p と書く場合もある.

例 3.49. 具体的な測度に関する積分をいくつか述べる.

- $\mu = \delta_x$ ($x \in X$ 上の Dirac 測度) の場合, $\int f d\mu = f(x)$.
- $\mu = \sum_{n=1}^{\infty} \delta_{x_n}$ (計数測度) の場合, $\int d\mu = \sum_{n=1}^{\infty} f(x_n)$. 特に無限級数 $\sum_{n=1}^{\infty} a_n$ とは, 数列 $a: \mathbb{N} \rightarrow \mathbb{R}$ を自然数上の数え上げ測度 $\sum_{n \in \mathbb{N}} \delta_n$ に関して積分したものにはかならない.
- $\mathcal{B}/\mathcal{B}^*$ -可測関数 $f: \mathbb{R} \rightarrow \overline{\mathbb{R}}$ が閉区間 $[a, b]$ 上で Riemann 積分可能であるならば, $\int_{[a,b]} f d\lambda = \int_a^b f(x) dx$ が成り立つ (左辺は Lebesgue 測度に関する積分, 右辺は Riemann 積分). それを

踏まえて今後は Riemann 積分可能な関数の Lebesgue 測度に関する積分 $\int f(x)d\lambda(x)$ のことも $\int f(x)dx$ と書いてよいとわかる. なお, 実際にはもっと強いことが言えるので, (積分が定義できる前提で) $\int f(x)d\lambda(x)$ のことを $\int f(x)dx$ と書いてよいからそうする; 次の問を参照せよ^{*21}.

問 3.50. (X, \mathcal{A}) を可測空間とする.

- (i) 測度 $\mu: \mathcal{A} \rightarrow [0, \infty]$ が完備であるとは, $\mu(N) = 0$ かつ $M \subset N$ ならば $M \in \mathcal{A}$ が成り立つことをいう. 言葉で述べれば, 測度が 0 となるような可測集合の部分集合が必ず可測となるとき, その測度は完備であると言われる. このノートにおける Lebesgue 測度は完備ではない.
- (ii) 一般に, 完備とは限らない測度 $\mu: \mathcal{A} \rightarrow [0, \infty]$ が与えられたとき, $\mathcal{A} \subset \mathcal{A}^*$ を満たす σ -加法族ならびにその上の完備な測度 $\mu^*: \mathcal{A}^* \rightarrow [0, \infty]$ であって, \mathcal{A} 上で μ に一致するものが存在する. このような μ^* の作り方の一つに「完備化」というものがある. λ の完備化によって得られる σ -加法族を \mathcal{L} と書き, \mathcal{L} の元を Lebesgue 可測集合という. また, 文献によってはこのノートにおける Lebesgue 測度の完備化を単に Lebesgue 測度と呼ぶものもある.
- (iii) Riemann 積分可能な関数 $f: [a, b] \rightarrow \mathbb{R}$ は, 一般には $\mathcal{B}/\mathcal{B}^*$ -可測であるとは限らない. が, $\mathcal{L}/\mathcal{B}^*$ -可測にはなる. 更に $\int_{[a,b]} f d\lambda = \int_a^b f(x)dx$ が成り立つ (左辺は Lebesgue 測度の完備化に関する積分, 右辺は Riemann 積分). 完備化まで考えれば, Lebesgue 測度に関する積分は Riemann 積分を包括していると言える.
- (iv) 他方, 広義 Riemann 積分まで含めて考えると事情はやや複雑である. $f: [0, \infty) \rightarrow \overline{\mathbb{R}}$ が非負であるか $f \in L^1$ であるならば, $\int_{(0,\infty)} f d\lambda = \int_0^\infty f dx$ が成り立つ (右辺は広義 Riemann 積分として $\int_0^\infty f dx := \lim_{R \rightarrow \infty} \int_0^R f dx$ を考えているが, 他の形の極限であったとしても本質的には同様に議論できると思う). 一方, $f(x) = \frac{\sin x}{x}$ は広義 Riemann 積分としては有限確定値 $\int_0^\infty f dx = \frac{\pi}{2}$ を持つが, f は絶対可積分ではないので $\int_{(0,\infty)} f d\lambda = \infty$ である. この例は優収束定理が使えない例としても捉えることができると思う.

余談 3.51. 実際には Riemann 積分可能とは限らない関数も積分できる場合がある (たとえば $1_{\mathbb{Q}}$ がそう). なので Lebesgue 測度に関する積分は Riemann 積分の一般化と言える…が, 個人的には $1_{\mathbb{Q}}$ が積分できるようになったからといって特にありがたい気はしない. 測度論の主たるありがたいところはそこではないと思っているので, Riemann 積分の限界だの一般化だのという文脈に沿って測度論を展開するのは好きになれない (勿論, 測度に関する積分の具体例の中で Riemann 積分は致命的に重要であるとも思うが)^{*22}. 少なくとも計算機に載せようとするときは Riemann 積分の定義に

^{*21} TODO: この問は特に, できればちゃんと証明をつけて本文に組み込んでしまいたい.

^{*22} 昔の私は「測度論というのは Riemann 積分の拡張としての Lebesgue 積分論のことで, それをキチンと理解していないと測度論的確率論はわからないのではないか」と思っていたのだが, 実際にはこの認識は誤りであった. また測度論に対しても苦手意識があったのだが, 確率論を勉強する際に抽象的な形で測度論を勉強したところわかりやすく, むしろ Riemann 積分の拡張という歴史的経緯を重視した論理展開に理解を阻害されていた (と, あとになって気づ

沿ってやったほうが簡単なのではないかと思っているが、そうでもないのだろうか？

命題 3.52. 上記の単関数に対する積分の定義は well-defined である。

証明. 単関数 f が $a_i, b_j \in \mathbb{R}$ ならびに $A_i, B_j \in \mathcal{A}$ (ただし $1 \leq i \leq n$ および $1 \leq j \leq m$) を用いて $f = \sum_{i=1}^n a_i 1_{A_i} = \sum_{j=1}^m b_j 1_{B_j}$ と 2 通りに表示できていたとする。このとき $C_{ij} := A_i \cap B_j$ と置けば $C_{ij} \in \mathcal{A}$ であり、かつ C_{ij} 上で $a_i = b_j$ なので、

$$\sum_{i=1}^n a_i \mu(A_i) = \sum_{i=1}^n \sum_{j=1}^m a_i \mu(C_{ij}) = \sum_{i=1}^n \sum_{j=1}^m b_j \mu(C_{ij}) = \sum_{j=1}^m b_j \mu(B_j)$$

となる。 ■

さて、実際の積分の計算では、積分の評価しやすい関数列 f_n であって被積分関数 f に何らかの意味で収束するものを取ってきて、 $\lim_{n \rightarrow \infty} \int f_n d\mu$ を計算する戦略を取ることが多々ある。が、これが $\int f d\mu$ に一致するには、言い換えれば積分と極限の順序交換ができるには一定の条件が必要である。この問題に対する（個人的には Riemann 積分の場合よりも遥かに単純だと思う）充分条件を以下に述べていく。

定義 3.53. 関数 $f: X \rightarrow \mathbb{R}$ と $g: X \rightarrow \mathbb{R}$ が任意の $x \in X$ に対して $f(x) \leq g(x)$ を充たすとき、 $f \leq g$ と書く。 $f_1 \leq f_2 \leq \dots$ を充たす関数の族 $\{f_n\}_{n=1}^\infty$ が、任意の $x \in X$ に対して $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ を充たすとき $f_n \uparrow f$ と書く。

補題 3.54. 単関数に対する積分は単調性ならびに線型性を充たす。すなわち、 $f, g: X \rightarrow [0, \infty]$ を非負単関数、 $a, b \in \mathbb{R}$ とするとき、

- $f \leq g$ を充たすならば $\int f d\mu \leq \int g d\mu$.
- $\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu$.

したがって特に $A \in \mathcal{A}$ とするとき、 $f^+ 1_A \leq f^+$ なので、 f の積分が定義できるならば $\int_A f d\mu$ も定義される。

証明. $f = \sum_{i=1}^m a_i 1_{A_i}$, $g = \sum_{j=1}^n b_j 1_{B_j}$ と表示し、更に $C_{ij} := A_i \cap B_j$ と置く。このとき

いた)。もちろん歴史的には測度論の走りは Riemann 積分の拡張にあったのだろうとは思いますが、論理構成の上では歴史的経緯を踏まえることは必ずしも必要ない（どころか、場合によっては遠回りにすらなり得る）と思う。更に個人的な事情としては、私はとくに多変数の Riemann 積分を真面目に勉強したことがなく、他の概念を理解するための中継地として Riemann 積分が機能しづらかったというのものもあるかもしれない。

$f + g = \sum_{i=1}^m \sum_{j=1}^n (a_i + b_j) 1_{C_{ij}}$ が成り立つので、

$$\begin{aligned} \int (f + g) d\mu &= \sum_{i=1}^m \sum_{j=1}^n (a_i + b_j) \mu(C_{ij}) \\ &= \sum_{i=1}^m \sum_{j=1}^n a_i \mu(C_{ij}) + \sum_{i=1}^m \sum_{j=1}^n b_j \mu(C_{ij}) \\ &= \sum_{i=1}^m a_i \mu(A_i) + \sum_{j=1}^n b_j \mu(B_j) \\ &= \int f d\mu + \int g d\mu \end{aligned}$$

を得る。定数倍が積分記号と交換できることは積分の定義から直ちに従う。

f, g が $f \leq g$ を満たす非負単関数ならば $g - f \geq 0$ すなわち $\int (g - f) d\mu \geq 0$ なので、積分の線型性より $\int g d\mu \geq \int f d\mu$ である（単関数は決して $\pm\infty$ を取らないので、この議論の過程で $\infty - \infty$ は起こらない）。 ■

余談 3.55. 積分の線型性は「何を当たり前な…」と思うかも知れないが、非負可測関数に対する積分の線型性を証明するには単調収束定理（の特別な場合）が必要になるので、あまり自明ではないと思う。ところで Riemann 積分に対する積分の線型性をどうやって示すのか、私は忘れしました。

補題 3.56. $f: X \rightarrow \mathbb{R}$ を非負可測関数とすると、非負単関数からなる列 $\{f_n\}_{n=1}^\infty$ であって、 $f_n \uparrow f$ を満たすようなものがある。

証明. $n \in \mathbb{N}$ ならびに $i = 1, 2, \dots, n2^n$ に対して $B_n := f^{-1}([n, \infty))$, $A_{n,i} := f^{-1}\left(\left[\frac{i-1}{2^n}, \frac{i}{2^n}\right)\right)$ と定めた上で、

$$f_n := n 1_{B_n} + \sum_{i=1}^{n2^n} \frac{i-1}{2^n} 1_{A_{n,i}}$$

とすれば、 $0 \leq f_n \leq f$ である。自然数 N が充分大きければ $|f(x) - f_N(x)| < 2^{-N}$ なので、 $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ もよい。 n に関する単調性は以下の考察から分かる^{*23}。

- $x \in B_{n+1}$ の場合. $n = f_n(x) < f_{n+1}(x) = n + 1$.
- $x \in B_n \setminus B_{n+1}$ の場合. $f_n(x) = n$ はよい. $f_{n+1}(x) < n$ であつたとすると、 $i/2^{n+1} \leq n$ を満たすある i について $x \in [(i-1)/2^{n+1}, i/2^{n+1})$ であることになるが、これは $x \notin B_n$ を含意するので矛盾である。
- $x \notin B_n$ の場合. ある i について $x \in A_{n,i}$ であるが、 $A_{n,i} = A_{n+1,2i-1} \cup A_{n+1,2i}$ と分解でき、 $f_n|_{A_{n,i}} = (i-1)/2^n = f_{n+1}|_{A_{n+1,2i-1}} \leq f_{n+1}|_{A_{n+1,2i}}$ なのでよい。

^{*23} ここ、もっと簡単にならないか？

■

余談 3.57. 上記の補題は証明とその気持ちがかなり離れた証明のひとつだと思う。気分としては f_n は「 f の値が n 未満の部分はちゃんと近似する気があるが、それ以上は n につぶす」という関数である。その上で f_n が f に各点収束してほしいので、近似する範囲が $O(n)$ で増えていくことを踏まえてそれ以上に細かい分割（たとえば $O(2^n)$ で増える分割）を取ることで、近似する範囲の近似精度を高めようとしている。実際には $O(2^n)$ で増える分割を採らずとも、 $O(n^2)$ で増える分割でも間に合うのかも知れない。収束は遅くなりそうだが、収束さえしてくれればとりあえずの使い勝手は困らない、ような気がする。

定理 3.58 (Beppo-Levi の単調収束定理). $f: X \rightarrow [0, \infty]$ ならびに $f_n: X \rightarrow [0, \infty]$ ($n \in \mathbb{N}$) を可測関数とする. $f_n \uparrow f$ ならば $\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu$.

証明. 証明を二段階に分割する. まず f_n が非負単関数の列であった場合を考えると, $\int f_n d\mu$ は n に関して非減少なので, $c := \lim_{n \rightarrow \infty} \int f_n d\mu$ が $[0, \infty]$ の範囲に存在する. 更に非負可測関数に対する積分の定義より, $c \leq \int f d\mu$ が成り立つ. $g: X \rightarrow [0, \infty]$ を, $g \leq f$ を満たすような任意の非負単関数とすると, $\int g d\mu \leq c$ を示せばよい. $g = \sum_{i=1}^m a_i 1_{A_i}$ と表示する. 非負単関数に対する積分の線型性より, 各 $1 \leq i \leq m$ に対して $a_i \mu(A_i) \leq \lim_{n \rightarrow \infty} \int f_n 1_{A_i} d\mu$ を示せば充分なのでそうする. $\varepsilon \in (0, 1)$ を任意に固定し, $B_{n,i} := \{x \in X \mid f_n(x) 1_{A_i} > (1 - \varepsilon) a_i\}$ と置くと, $B_{n,i} \uparrow A_i$. 更に $\int f_n 1_{A_i} d\mu \geq \int f_n 1_{B_{n,i}} d\mu \geq (1 - \varepsilon) \mu(B_{n,i})$ であるから, $n \rightarrow \infty$ の極限をとって $\lim_{n \rightarrow \infty} \int f_n 1_{A_i} d\mu \geq (1 - \varepsilon) a_i \mu(A_i)$ であり, $\varepsilon \rightarrow 0$ の極限を取ることで結論を得る.

次に f_n が非負可測関数の列であった場合を考える. 補題 3.56 より非負単関数の列 $\{g_{n,i}\}_{i=1}^\infty$ であって $g_{n,i} \uparrow f_n$ であるようなものがある. $h_m := \max\{g_{m,1}, g_{m,2}, \dots, g_{m,m}\}$ と置けば $\{h_m\}_{m=1}^\infty$ は単関数であって, $h_m \leq f_m$ ならびに $h_m \uparrow f$ を満たす. 積分の単調性より $\int h_m d\mu \leq \int f_m d\mu \leq \int f d\mu$ であり, 前段の結果より $\lim_{m \rightarrow \infty} \int h_m d\mu = \int f d\mu$ なので, はさみうちの原理より $\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu$ である. ■

定理 3.59 (Fatou の補題). $f_n: X \rightarrow [0, \infty]$ ($n \in \mathbb{N}$) を可測関数の族とする. このとき $\int \liminf_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$.

証明. $g_n := \inf_{m \geq n} f_m$ と置くと, $g_n \uparrow \sup_n g_n = \liminf_{n \rightarrow \infty} f_n$ なので, 単調収束定理より $\lim_{n \rightarrow \infty} \int g_n d\mu = \int \liminf_{n \rightarrow \infty} f_n d\mu$. また $g_n \leq f_n$ なので $\int g_n d\mu \leq \int f_n d\mu$ すなわち $\liminf_{n \rightarrow \infty} \int g_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$ である. 以上より $\int \liminf_{n \rightarrow \infty} f_n d\mu = \lim_{n \rightarrow \infty} \int g_n d\mu = \liminf_{n \rightarrow \infty} \int g_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$ を得る. ■

定理 3.60 (Lebesgue の優収束定理). $f_n \in L^1$ ($n \in \mathbb{N}$) がある関数 $f: X \rightarrow \mathbb{R}$ に各点収束していた,

すなわち任意の $x \in X$ に対して $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ であったと仮定する. このとき, ある $g \in L^1$ に対して $|f_n| \leq g$ であるならば, $f \in L^1$ であり, $\lim_{n \rightarrow \infty} \int f_n d\mu = \int f d\mu$.

証明. $|f| \leq |g|$ なので $f \in L^1$ はよい. $g + f_n \geq 0$ なので Fatou の補題が使えて, $\int \lim_{n \rightarrow \infty} (g + f_n) d\mu \leq \liminf_{n \rightarrow \infty} \int (g + f_n) d\mu$ すなわち $\int \lim_{n \rightarrow \infty} f_n d\mu \leq \liminf_{n \rightarrow \infty} \int f_n d\mu$ である. 同様に $g - f_n \geq 0$ なので再び Fatou の補題により $\int \lim_{n \rightarrow \infty} (g - f_n) d\mu \leq \liminf_{n \rightarrow \infty} \int (g - f_n) d\mu$ つまり $\limsup_{n \rightarrow \infty} \int f_n d\mu \leq \int \lim_{n \rightarrow \infty} f_n d\mu$ である. あとははさみうちの原理によればよい. ■

収束定理たちの応用や, 積分にまつわる具体例たちを見ていくことにしよう.

系 3.61 (積分の線型性). 可測関数 $f, g: X \rightarrow \overline{\mathbb{R}}$ が非負であるか, または $f, g \in L^1$ であるならば, 任意の $a, b \in \mathbb{R}$ に対して $\int (af + bg) d\mu = a \int f d\mu + b \int g d\mu$.

証明. f, g が非負ならば, 非負単関数の列 $\{f_n\}$ ならびに $\{g_n\}$ であって $f_n \uparrow f, g_n \uparrow g$ を充たすものをつとて上で単調収束定理を繰り返し使うと $a \int f d\mu + b \int g d\mu = a \int \lim_{n \rightarrow \infty} f_n d\mu + b \int \lim_{n \rightarrow \infty} g_n d\mu = \lim_{n \rightarrow \infty} a \int f_n d\mu + b \int g_n d\mu = \lim_{n \rightarrow \infty} \int (af_n + bg_n) d\mu = \int (f + g) d\mu$. $f, g \in L^1$ の場合は正部分と負部分に分けて非負可測関数に対する積分の線型性を適用すればよい. ■

定理 3.62 (積分と級数の交換, 項別積分). 非負可測関数 $f_n: X \rightarrow [0, \infty]$ ($n \in \mathbb{N}$) に対して, $\sum_{n=1}^{\infty} \left(\int f_n d\mu \right) = \int \left(\sum_{n=1}^{\infty} f_n \right) d\mu$. すなわち非負可測関数に関する限り, 積分と級数は自由に交換できる.

証明. $g_N := \sum_{n=1}^N f_n$ と置けば $g_N \uparrow \sum_{n=1}^{\infty} f_n$ なので, 積分の線型性と単調収束定理より $\lim_{N \rightarrow \infty} \sum_{n=1}^N \left(\int f_n d\mu \right) = \lim_{N \rightarrow \infty} \int g_N d\mu = \int \lim_{N \rightarrow \infty} g_N d\mu = \int \left(\sum_{n=1}^{\infty} f_n \right) d\mu$. ■

余談 3.63. (書くのを忘れていなければ) 後に見るが, 上記は Fubini の定理を用いても証明できる (はず).

例 3.64 (Fatou の補題で等号が達成されない例). $n \in \mathbb{N}$ に対して $f_n := 1_{[n, n+1)}$ と置くと, 任意の $n \in \mathbb{N}$ に対して $\int f_n d\mu = 1$ であるが, $f_n \rightarrow 0$ なので $\int \lim_{n \rightarrow \infty} f_n d\mu = 0$. 従って特に $\int \lim_{n \rightarrow \infty} f_n d\mu < \lim_{n \rightarrow \infty} \int f_n d\mu$ である.

定理 3.65 (微分と積分の交換). $f: X \times (a, b) \rightarrow \mathbb{R}$ を, 次の 3 条件を充たす関数とする.

- 任意の $t \in (a, b)$ に対して $(x \mapsto f(x, t)) \in L^1$.
- 任意の $x \in X$ に対して, $g_x: (a, b) \ni t \mapsto f(x, t) \in \mathbb{R}$ は (a, b) において微分可能.
- ある $h \in L^1$ が存在して, 任意の $x \in X$ ならびに $t_0 \in (a, b)$ に対して $|dg_x/dt(t_0)| \leq h(x)$.

このとき, $g: (a, b) \ni t \mapsto \int f(x, t) d\mu(x) \in \mathbb{R}$ は (a, b) において微分可能であって, 任意の $t_0 \in (a, b)$ に対して

$$\frac{dg}{dt}(t_0) = \int \frac{dg_x}{dt}(t_0) d\mu(x)$$

が成り立つ. 人口に膾炙した記号を用いてより象徴的に書けば,

$$\frac{d}{dt} \Big|_{t=t_0} \int f(x, t) d\mu(x) = \int \frac{\partial f}{\partial t}(t_0, x) d\mu(x)$$

が任意の $t_0 \in (a, b)$ に対して成り立つ.

証明. $t_0 \in (a, b)$ を任意にひとつ固定する. \mathbb{R} の点列 $\{\delta_n\}_{n=1}^\infty$ を, $\delta_1 < b - t_0$ ならびに $\delta_n \downarrow 0$ を充たすように取る. 平均値の定理より, 任意の $x \in X$ ならびに $n \in \mathbb{N}$ に対してある $\theta_{t_0, x, n} \in [0, 1]$ が存在して

$$\frac{f(x, t_0 + \delta_n) - f(x, t_0)}{\delta_n} = \frac{dg_x}{dt}(x, t_0 + \theta_{t_0, x, n} \delta_n)$$

が成り立つ. 従って任意の $t_0 \in (a, b)$, $x \in X$ ならびに $n \in \mathbb{N}$ に対して $\left| \frac{f(x, t_0 + \delta_n) - f(x, t_0)}{\delta_n} \right| \leq h(x)$ が成り立つから, 優収束定理より

$$\lim_{n \rightarrow \infty} \frac{\int f(x, t_0 + \delta_n) d\mu(x) - \int f(x, t_0) d\mu}{\delta_n} = \int \lim_{n \rightarrow \infty} \frac{f(x, t_0 + \delta_n) - f(x, t_0)}{\delta_n} d\mu(x)$$

を得るので結論が従う. ■

系 3.66 (項別微分). $f_n: (a, b) \ni t \mapsto f_n(t) \in \mathbb{R}$ を微分可能な関数とし, 任意の $t \in (a, b)$ に対して $\sum_{n=1}^\infty |f_n(t)| < \infty$ であったと仮定する. このとき, 任意の $t \in (a, b)$ に対して $\sum_{n=1}^\infty \frac{d}{dt} f_n(t) < \infty$ であるならば, 任意の $t_0 \in (a, b)$ に対して

$$\frac{d}{dt} \sum_{n=1}^\infty f_n(t_0) = \sum_{n=1}^\infty \frac{d}{dt} f_n(t_0).$$

証明. 前定理において μ を \mathbb{N} 上の計数測度にとれ. ■

命題 3.67 (Leibniz の積分公式). 以下, $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ を連続関数とし, 更に $\partial_2 f$ が存在して連続であるとする.

- (i) $F(x, y) := \int_a^x f(t, y) dt$ と定めると, $\partial_1 F = f$ および $\partial_2 F(x, y) = \int_a^x \partial_2 f(t, y) dt$ が成り立つ.
- (ii) $g: \mathbb{R} \rightarrow (a, b)$ を微分可能な関数とし, $G(x) := \int_a^{g(x)} f(t, x) dt$ と置くと, $G'(x) = f(g(x), x) \cdot g'(x) + \int_a^{g(x)} \partial_2 f(t, x) dt$ が成り立つ. 特に $g(x) = x$ のとき, $G'(x) = f(x, x) + \int_a^x \partial_2 f(t, x) dt$ である.

証明.

(i) $\partial_1 F = f$ は微積分学の基本定理である. f は定理 3.65 の前件をすべて満たすので, 積分と微分を入れ替えてよいから後半も従う.

(ii) 合成関数 $x \mapsto (g(x), x) \mapsto F(g(x), x)$ に合成則と (i) の結果を使え.

■

問 3.68. 定理 3.65 の前件を検討すれば, Leibniz の積分公式において f に課した条件はより弱められることがわかる. ところでこの公式は名前がついているくらいだし, いろいろなところで使われると思うのだが, 私は使いみちをよく知りません.

以下, やや寄り道となるが, L^p たちの包含関係について議論する. $\|f\|_p := \left(\int |f|^p d\mu\right)^{1/p}$ と定める.

問 3.69. $f \in L^p$ と $\|f\|_p < \infty$ は同値である. $\|f\|_p$ のことを L^p ノルムと呼びたいのだが, 実際には $f = g$ μ -a.e でさえあれば ($f = g$ ではなかったとしても) $\int f d\mu = \int g d\mu$ となるので, ノルムの同一律を充たさない. なので「測度零集合での違いを除いて一致する」事に関する同値関係を入れ, その同値類の集合を考えないとノルムにはならない. 文献によっては, 同値関係で割る前の集合のことを \mathcal{L}^p と書き, 割ったあとの集合のことを L^p と書いているものもある.

定理 3.70 (Young の不等式). a, b を非負実数, p, q を $p^{-1} + q^{-1} = 1$ を充たす自然数とするととき, $ab \leq (a^p/p) + (b^q/q)$.

証明. $a = 0$ または $b = 0$ の場合はあきらか. そうでないとする, 対数関数の凸性より $\log((a^p/p) + (b^q/q)) \geq \log a + \log b$ を得るので, 両辺の指数を取って結論を得る. ■

定理 3.71 (Hölder の不等式). 可測関数 $f, g: X \rightarrow \overline{\mathbb{R}}$ および $1/p + 1/q = 1$ を充たす自然数 p, q に対して $\int |fg| d\mu \leq \|f\|_p \|g\|_q$ が成り立つ. $p = q = 2$ の場合を Cauchy-Schwarz の不等式という.

証明. $\|f\|_p = 0$ ないしは $\|g\|_q = 0$ であるならば μ -a.e に $f = 0$ または μ -a.e に $g = 0$ なので, 不等式の両辺とも 0 である. 従って $\|f\|_p > 0$ かつ $\|g\|_q > 0$ と仮定してよい. 更に積分の線型性より f を $f/\|f\|_p$ に, g を $g/\|g\|_q$ に置き換えて証明すれば良いことがわかるので, 最初から $\|f\|_p = 1$ かつ $\|g\|_q = 1$ であるとして示して良い. Young の不等式より各 $x \in X$ に対して $|f(x)g(x)| \leq (|f(x)|^p/p) + (|f(x)|^q/q)$ を得るので, 両辺を積分して $\int |fg| d\mu \leq (\|f\|_p^p/p) + (\|f\|_q^q/q) = (1/p) + (1/q) = 1$ となる. ■

命題 3.72. $\mu(X) < \infty$, すなわち μ が有限測度であるとする. このとき $p \geq q$ ならば $L^p(\mu) \subset L^q(\mu)$ である.

証明. $f \in L^p$ を任意に取る. $r := p/q$, $s = p/(p-q)$ として Hölder の不等式を使えば

$$\int |f|^q d\mu \leq \| |f|^q \|_r \|1\|_s = \mu(X)^{(p-q)/p} \left(\int |f|^p d\mu \right)^{q/p}$$

を得る. $\int |f|^p d\mu < \infty$ ならびに $\mu(X) < \infty$ なので $\|f\|_q < \infty$, すなわち $f \in L^q$ である. ■

上記の包含関係が成り立つ上で μ の有限性は本質的である.

例 3.73. $f: \mathbb{R} \rightarrow \mathbb{R}$ を, $f(x) := \frac{1}{x} 1_{[1, \infty)}(x)$ で定める. f の Lebesgue 測度に関する積分を考えると, $\int f d\lambda = \int_1^\infty \frac{dx}{x} = [\log x]_{x=1}^{x=\infty} = \infty$ なので $f \notin L^1$. ところが任意の $p \geq 2$ に対して $\int f^p d\lambda = \int_1^\infty \frac{dx}{x^p} = \left[\frac{1}{-p+1} x^{-p+1} \right]_{x=1}^{x=\infty} = \frac{1}{-p+1} < \infty$ なので $f \in L^p$ である. 他方 $g: \mathbb{R} \rightarrow \mathbb{R}$ を $g(x) := 1_{(0,1]}(x) \frac{1}{\sqrt{x}}$ で定めると, $\int g d\lambda = \int_0^1 \frac{dx}{\sqrt{x}} = [2\sqrt{x}]_{x=0}^{x=1} = 2$ なので $g \in L^1$ である. 他方, $\int g^2 d\lambda = \int_0^1 \frac{dx}{x} = [\log x]_{x=0}^{x=1} = -\infty$ なので $g \notin L^2$. したがって, この場合は $L^1 \subset L^2$ でも $L^2 \subset L^1$ でもない.

例 3.74. ここでは $(\mathbb{N}, 2^{\mathbb{N}})$ 上の計数測度を考える. この測度に関する L^p 空間のことを ℓ^p と書くことにする; すなわち, $\ell^p := \{ \{x_i\} \in \mathbb{R}^{\mathbb{N}} \mid \sum_{i=1}^\infty |x_i|^p < \infty \}$. この設定のもとでは, $p \leq q$ ならば $\ell^p \subseteq \ell^q$ が成り立つ. それを見るために $x \in \ell^p$ を取ると, ある適当な $N \in \mathbb{N}$ が存在して, $i \geq N$ ならば $x_i < 1$ である. 従って $\sum_{i=N}^\infty |x_i|^q \leq \sum_{i=N}^\infty |x_i|^p < \infty$ だから $f \in \ell^q$. また, $y_i := 1/i$ によって数列 $\{y_i\}$ を定めると, 例えば $\sum_{i=1}^\infty (y_i)^2 = \pi^2/6 < \infty$ であるが, $\sum_{i=1}^\infty y_i = \infty$ なので $\ell^1 \not\subseteq \ell^2$ である.

3.4 拡張定理の証明

後回しにしていた, 拡張定理 (定理 3.17) の証明を与える. 存在と一意性のそれぞれに証明を分割する.

定理 3.17 (Carathéodory の拡張定理). \mathcal{A} を有限加法族, $\mu: \mathcal{A} \rightarrow [0, \infty]$ を有限加法的測度とする. 更に μ は次の意味で \mathcal{A} 上可算加法的であると仮定する; 高々可算個の \mathcal{A} の元からなる族 $\{A_i\}_{i=1}^\infty$ が互いに排反であり, かつ $\bigcup_{i=1}^\infty A_i \in \mathcal{A}$ であるならば, $\mu(\bigcup_{i=1}^\infty A_i) = \sum_{i=1}^\infty \mu(A_i)$. このとき, μ は $\sigma(\mathcal{A})$ 上の測度に拡張できる. 更に, μ が \mathcal{A} 上で σ -有限ならば, この拡張は一意的である.

証明 (定理 3.17 における拡張の存在). $\mu^*: 2^X \rightarrow [0, \infty]$ を,

$$\mu^*(E) := \inf \left\{ \sum_{i=1}^\infty \mu(A_i) \mid E \subset \bigcup_{i=1}^\infty A_i \text{ and } A_1, A_2, \dots \in \mathcal{A} \right\}$$

によって定める. $E \in \mathcal{A}$ であるならば, 集合列 A_i を $A_1 = E, A_2 = \emptyset, A_3 = \emptyset, \dots$ のように取って考えることで $\mu^*(E) \leq \mu(E)$ がわかる. 逆向きの不等式を示せば μ^* は写像として μ の拡張になっていることがわかるが, このノートではまだサボっている (ここに μ の \mathcal{A} 上での可算加法

性を使う). したがって μ^* が $\sigma(\mathcal{A})$ 上の測度であることを示せばよい. 以下, 一般に実数の部分集合が $A \subset B$ を満たすならば $\inf B \leq \inf A$ であることに注意する.

まず μ^* の単調性, すなわち $E, F \in 2^X$ が $E \subset F$ を満たすならば $\mu^*(E) \leq \mu^*(F)$ となることを示す.

$$\left\{ \{A_i\} \subset \mathcal{A} \mid F \subset \bigcup_{i=1}^{\infty} A_i \right\} \subset \left\{ \{A_i\} \subset \mathcal{A} \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\}$$

であるから,

$$\left\{ \sum_{i=1}^{\infty} \mu(A_i) \mid F \subset \bigcup_{i=1}^{\infty} A_i \text{ and } A_1, A_2, \dots \in \mathcal{A} \right\} \subset \left\{ \sum_{i=1}^{\infty} \mu(A_i) \mid E \subset \bigcup_{i=1}^{\infty} A_i \text{ and } A_1, A_2, \dots \in \mathcal{A} \right\}$$

すなわち $\mu^*(E) \leq \mu^*(F)$ を得る.

つづいて μ^* の可算劣加法性, すなわち $E_1, E_2, \dots \in 2^X$ に対して $\mu^*(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu^*(E_n)$ となることを示す. $\varepsilon > 0$ を任意にひとつ取って固定する. 各 $n = 1, 2, \dots$ に対して, \mathcal{A} の元の族 $\{A_{n,i}\}_{i=1}^{\infty}$ を $E_n \subset \bigcup_{i=1}^{\infty} A_{n,i}$ ならびに $\sum_{i=1}^{\infty} \mu(A_{n,i}) < \mu^*(E_n) + \frac{\varepsilon}{2^n}$ を満たすように取れるので取る. このとき $\bigcup_{n=1}^{\infty} E_n \subset \bigcup_{i,n=1}^{\infty} A_{n,i}$ なので,

$$\mu^*\left(\bigcup_{n=1}^{\infty} E_n\right) \leq \sum_{i,n=1}^{\infty} \mu(A_{n,i}) < \varepsilon + \sum_{n=1}^{\infty} \mu^*(E_n)$$

を得る. $\varepsilon > 0$ は任意だったので可算劣加法性が従う.

更に, 次の性質を満たすような集合族 \mathcal{C} が \mathcal{A} を含む σ -加法族であることを示す; $A \in \mathcal{C}$ であるならば, 任意の $E \in 2^X$ に対して, $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A)$. まず $A \in \mathcal{A}$ を任意に取る. すると $E \subset \bigcup_{i=1}^{\infty} A_i$ ならば $E \cap A \subset \bigcup_{i=1}^{\infty} (A_i \cap A)$ なので,

$$\left\{ \{A_i \cap A\} \subset \mathcal{A} \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\} \subset \left\{ \{B_i\} \subset \mathcal{A} \mid E \cap A \subset \bigcup_{i=1}^{\infty} B_i \right\}$$

すなわち

$$\mu^*(E \cap A) \leq \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i \cap A) \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\}$$

を得る. 同様にして

$$\mu^*(E \setminus A) \leq \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i \setminus A) \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\}$$

も従う. ゆえに

$$\begin{aligned} \mu^*(E \cap A) + \mu^*(E \setminus A) &\leq \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i \setminus A) \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\} + \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i \cap A) \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\} \\ &\leq \inf \left\{ \sum_{i=1}^{\infty} \mu(A_i \setminus A) + \sum_{i=1}^{\infty} \mu(A_i \cap A) \mid E \subset \bigcup_{i=1}^{\infty} A_i \right\} \\ &= \mu^*(E) \end{aligned}$$

を得る（最後の等式は測度の非負性，ならびに正項級数は項の順番を並べ替えても値が変わらないことから出る）．逆向きの不等式は可算劣加法性から従う．したがって $\mathcal{A} \subset \mathcal{C}$. $X \in \mathcal{C}$ ならびに $A \in \mathcal{C} \Rightarrow X \setminus A \in \mathcal{C}$ はよいので， $A_1, A_2, \dots \in \mathcal{C}$ のときに $\bigcup_{i=1}^{\infty} A_i \in \mathcal{C}$ となること，すなわち $\mu^*(E) \geq \mu^*(E \cap \bigcup_{i=1}^{\infty} A_i) + \mu^*(E \setminus \bigcup_{i=1}^{\infty} A_i)$ となることを示す（逆向きの不等式はやはり可算劣加法性から従うのでよい）． $A_1, A_2 \in \mathcal{C}$ に対して

$$\begin{aligned}\mu^*(E) &= \mu^*(E \cap A_1) + \mu^*(E \setminus A_1) \\ &= \mu^*(E \cap A_1) + \mu^*((E \setminus A_1) \cap A_2) + \mu^*((E \setminus A_1) \setminus A_2) \\ &= \mu^*(E \cap A_1) + \mu^*((E \setminus A_1) \cap A_2) + \mu^*((E \setminus A_1) \setminus A_2) \\ &= \mu^*\left(E \cap \bigcup_{i=1}^2 A_i \cap A_1\right) + \mu^*\left(\left(E \cap \bigcup_{i=1}^2 A_i\right) \setminus A_1\right) + \mu^*\left(E \setminus \bigcup_{i=1}^2 A_i\right)\end{aligned}$$

を得るので， $B_1 = A_1, B_i = A_i \setminus B_{i-1}$ と定めて帰納的に議論を繰り返すことで

$$\mu^*(E) = \mu^*\left(E \setminus \bigcup_{i=1}^n A_i\right) + \sum_{i=1}^n \mu^*(E \cap B_i)$$

を得る．なので

$$\mu^*(E) \geq \mu^*\left(E \setminus \bigcup_{i=1}^{\infty} A_i\right) + \sum_{i=1}^n \mu^*(E \cap B_i)$$

となるから，第二項に対して $n \rightarrow \infty$ とした上で可算劣加法性を使うことで結論が従う．以上より特に任意の $A \in \sigma(\mathcal{A})$ ならびに $E \in 2^X$ に対して， $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A)$ であることがわかった．

最後に， μ^* が $\sigma(\mathcal{A})$ 上で可算加法的であることを示す． $\{A_i\}_{i=1}^{\infty}$ を排反な $\sigma(\mathcal{A})$ の元の族とする．やはり可算劣加法性により $\mu^*(\bigcup_{i=1}^{\infty} A_i) \geq \sum_{i=1}^{\infty} \mu^*(A_i)$ を示せば充分なので，そうする． $\mu^*(\bigcup_{i=1}^{\infty} A_i) = \infty$ ならばやることはないのので，以下 $\mu^*(\bigcup_{i=1}^{\infty} A_i) < \infty$ とする．先程示した性質によって $\mu^*(\bigcup_{i=1}^{\infty} A_i) = \mu^*(\bigcup_{i=1}^n A_i) + \mu^*(\bigcup_{i=n+1}^{\infty} A_i)$ を得るので，特に $\mu^*(\bigcup_{i=1}^{\infty} A_i) - \mu^*(\bigcup_{i=1}^n A_i) = \mu^*(\bigcup_{i=n+1}^{\infty} A_i) \geq 0$ すなわち $\mu^*(\bigcup_{i=1}^{\infty} A_i) \geq \sum_{i=1}^n \mu^*(A_i)$ となるから， $n \rightarrow \infty$ とすることで定理の主張を得る． ■

余談 3.75. 証明の中に現れた μ^* は μ の外測度と呼ばれる．

ところで，「任意の $E \in 2^X$ に対して， $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A)$ 」という性質を充たす A に対して，「 A は Carathéodory 可測」という言い方もする．証明をよくよく検討すればわかるように，Carathéodory 可測な集合の全体 \mathcal{C} は $\sigma(\mathcal{A})$ よりも大きくなることもあり，かつ \mathcal{C} まで測度を拡張することも（証明はほぼそのままに）できる．そうしなかった理由は，あくまで拡張定理を便利な道具扱いにとどめて使い倒そうという観点からすれば， \mathbb{R} の測度は $\sigma(\mathcal{H})$ まで拡張できれば実用上そんなに問題なく，それ以上に測度を作り込んでも仕方がないと判断したからである*24．

*24 数学に YAGNI 原則を持ち込んでもしょうがないような気がするが，個人の好みということで大目に見てください．

問 3.76 (*). 内測度という概念が（外測度と同様に）定義でき、外測度と内測度が一致する時に可測、とすることもある。そのようにして測度を構成した場合、その測度は自動的に完備になる。その際の可測集合の全体は C に一致するような気がするのだが、確かめてみたことはない。

一意性の証明にうつる。 π - λ 定理の証明もここで行う。

定理 3.20 (Dynkin の π - λ 定理). $\mathcal{P} \subset 2^X$ を π -システム、 $\mathcal{L} \subset 2^X$ を λ -システムとし、更に $\mathcal{P} \subset \mathcal{L}$ であると仮定する。このとき $\sigma(\mathcal{P}) \subset \mathcal{L}$ 。すなわち $\sigma(\mathcal{P})$ は \mathcal{P} を含む λ -システムの中で最小のものである。

証明. λ -システムは σ -加法族と同様に共通部分をとる操作について閉じている； $\{\mathcal{L}_\mu\}_{\mu \in M}$ が λ -システムの族であるならば、 $\bigcap_{\mu \in M} \mathcal{L}_\mu$ もまた λ -システムである。そこで、 $\mathcal{L} := \{\mathcal{L} \mid \mathcal{L} \text{ は } \mathcal{P} \text{ を含む } \lambda\text{-システム}\}$ として、 $\mathcal{L}^* := \bigcap_{\mathcal{L} \in \mathcal{L}} \mathcal{L}$ が σ -加法族であることを示せば証明が完結するからそうする。

「 $\emptyset \in \mathcal{L}^*$ 」および「 $A \in \mathcal{L}^*$ ならば $X \setminus A \in \mathcal{L}^*$ 」は \mathcal{L}^* の定め方からわかる。 \mathcal{L}^* が π -システムであることを示せば、 $A, B \in \mathcal{L}^*$ に対して $A \cup B = X \setminus (A^C \cap B^C) \in \mathcal{L}^*$ がわかる。このことから \mathcal{L}^* の族 $\{A_i\}_{i=1}^\infty$ に対して $\bigcup_{i=1}^n A_i \in \mathcal{L}^*$ が従うので、 \mathcal{L}^* が λ -システムであることと併せて $\bigcup_{i=1}^n A_i \uparrow \bigcup_{i=1}^\infty A_i \in \mathcal{L}^*$ を得る。

以下、 \mathcal{L}^* が π -システムであることを示す。 $A \in \mathcal{P}$ を任意に固定し、 $\mathcal{D}_A := \{B \in \mathcal{L}^* \mid A \cap B \in \mathcal{L}^*\}$ と置く。 \mathcal{D}_A が \mathcal{P} を含む λ -システムであることを示せば、 $\mathcal{L}^* = \mathcal{D}_A$ がわかるので、 A の任意性と併せて次のことが言える；任意の $A \in \mathcal{P}$ および $B \in \mathcal{L}^*$ に対して $A \cap B \in \mathcal{L}^*$ 。その上で今度は $B \in \mathcal{L}^*$ を任意に固定し、 $\mathcal{E}_B := \{A \in \mathcal{L}^* \mid A \cap B \in \mathcal{L}^*\}$ と置いて、 \mathcal{E}_B が \mathcal{P} を含む λ -システムであることを示せば、 $\mathcal{L}^* = \mathcal{E}_B$ がわかるので、 B の任意性と併せて \mathcal{L}^* が π -システムであることが従う。

議論は同様なので \mathcal{D}_A が \mathcal{P} を含む λ -システムであることだけを示す。 \mathcal{P} は π -システムなので $\mathcal{P} \subset \mathcal{D}_A$ である。 $A \in \mathcal{L}^*$ だから $X \in \mathcal{D}_A$ もよい。 $B, C \in \mathcal{D}_A$ を $B \subset C$ を満たすようにとれば、 $A \cap (C \setminus B) = (A \cap C) \setminus (A \cap B) \in \mathcal{L}^*$ となるので $C \setminus B \in \mathcal{D}_A$ である。集合の増大列 $\{B_i\}$, $B_i \in \mathcal{D}_A$ を取ると、 $A \cap \bigcup_{i=1}^n B_i \in \mathcal{L}^*$ なので $A \cap \bigcup_{i=1}^n B_i \uparrow A \cap \bigcup_{i=1}^\infty B_i \in \mathcal{L}^*$ であるから $\bigcup_{i=1}^\infty B_i \in \mathcal{D}_A$ である。 ■

証明（定理 3.17 における拡張の一意性）。 ν_1, ν_2 を μ の拡張とし、 $A \in \mathcal{A}$ を $\mu(A) < \infty$ を満たすように任意に 1 つ取って固定する。その上で、

$$\mathcal{L}_A := \{B \in \sigma(\mathcal{A}) \mid \nu_1(A \cap B) = \nu_2(A \cap B)\}$$

と置く。 \mathcal{L}_A が \mathcal{A} を含む λ -システムであることが示せたとなると、 π - λ 定理によって $\sigma(\mathcal{A}) = \mathcal{L}_A$ を得る。 A は $\mu(A) < \infty$ を満たせば任意だったので、 $\mu(A) < \infty$ を満たす任意の $A \in \mathcal{A}$ ならびに任意の $B \in \sigma(\mathcal{A})$ に対して $\nu_1(A \cap B) = \nu_2(A \cap B)$ を得る。 μ は \mathcal{A} 上 σ -有限なので、 $\mu(A_i) < \infty$

を充たす集合の増大列 A_i があり $\bigcup_{i=1}^{\infty} A_i = X$ である. したがって任意の $B \in \sigma(\mathcal{A})$ に対して $B \cap A_i \uparrow B$ であるから, $\nu_1(B) = \lim_{i \rightarrow \infty} \nu_1(B \cap A_i) = \lim_{i \rightarrow \infty} \nu_2(B \cap A_i) = \nu_2(B)$ すなわち $\nu_1 = \nu_2$ を得る.

あとは \mathcal{L}_A が \mathcal{A} を含む λ -システムであることを示せば良い. \mathcal{A} 上 $\nu_1 = \nu_2$ で, かつ \mathcal{A} は π -システムなので $\mathcal{A} \subset \mathcal{L}_A$ である. したがって $X \in \mathcal{L}_A$ もよい. $B_1, B_2 \in \mathcal{L}_A$ が $B_1 \subset B_2$ を充たすならば, $\nu_1(A \cap (B_2 \setminus B_1)) = \nu_1((A \cap B_2) \setminus (A \cap B_1)) = \nu_1(A \cap B_2) - \nu_1(A \cap B_1) = \nu_2(A \cap B_2) - \nu_2(A \cap B_1) = \nu_2((A \cap B_2) \setminus (A \cap B_1)) = \nu_2(A \cap (B_2 \setminus B_1))$ なので $B_2 \setminus B_1 \in \mathcal{L}_A$ である (ここで $\nu_1(A) = \nu_2(A) < \infty$ を使っている). $B_i \uparrow B$ なる $\{B_i\} \subset \mathcal{L}_A$ に対して $\nu_1(A \cap B) = \lim_{i \rightarrow \infty} \nu_1(A \cap B_i) = \lim_{i \rightarrow \infty} \nu_2(A \cap B_i) = \nu_2(A \cap B)$ であるので, $B \in \mathcal{L}_A$ である. ■

4 多変数関数の積分

前節に引き続き、積分の議論を続ける。第一の話題である Fubini の定理は、いわゆる多変数関数の重積分を各変数に関する逐次積分で置き換えてよい充分条件を与える。この定理を Riemann 積分の範囲内で厳密に記述するとやや複雑となる。この複雑さを排し、よりシンプルな形で述べたい、というのが測度論を導入した所以のひとつである。

この章の目標のもう一つは重積分の変数変換公式である。その証明の過程で「1 の分割」と呼ばれる概念を導入する。1 の分割は、局所的な考察を張り合わせて大域的な結果を得るために使う。後の章でも、例えば Riemann 計量の存在証明に用いられる。

4.1 直積測度と Fubini の定理

重積分を定義するために Lebesgue 測度を \mathbb{R}^n に拡張する。より一般に、2 つ以上の測度空間が与えられているときに、その可測空間の直積に直積測度と呼ばれる測度をつくることができる。 \mathbb{R}^n 上の Lebesgue 測度は直積測度の一例として捉えられる。

以下暫く、 $(X_1, \mathcal{A}_1, \mu_1)$ および $(X_2, \mathcal{A}_2, \mu_2)$ を σ -有限な測度空間とする。 $\mathcal{A} := \{A_1 \times A_2 \mid A_1 \in \mathcal{A}_1, A_2 \in \mathcal{A}_2\}$ と置く。すでに定義したように、 $\mathcal{A}_1 \otimes \mathcal{A}_2 = \sigma(\mathcal{A})$ である。

定理 4.1. $\mu: \mathcal{A} \rightarrow [0, \infty]$ を $\mu(A_1 \times A_2) := \mu_1(A_1)\mu_2(A_2)$ で定める。このとき、 μ は $\mathcal{A}_1 \otimes \mathcal{A}_2$ 上の測度に一意的に拡張できる。

証明. μ が \mathcal{A} 上で可算加法的であることが示せたとすると、あとの議論は Lebesgue 測度の構成と同様にして進められる。すなわち \mathcal{A} を

$$\overline{\mathcal{A}} := \left\{ \bigcup_{i=1}^N A_i \mid A_i \in \mathcal{A}, A_i \text{ は排反} \right\}$$

と有限加法族に拡張する。排反な $A_1, A_2, \dots, A_n \in \mathcal{A}$ に対して $\mu\left(\bigcup_{i=1}^n A_i\right) := \sum_{i=1}^n \mu(A_i)$ と定めることで μ を $\overline{\mathcal{A}}$ 上の有限加法的測度に拡張すると、これは well-defined であって $\overline{\mathcal{A}}$ 上可算加法的であることが Lebesgue 測度の構成のときと同様にして示せるから、 $\sigma(\mathcal{A}) = \sigma(\overline{\mathcal{A}})$ と併せて Carathéodory の拡張定理によって主張が従う。

以下、 μ が \mathcal{A} 上で可算加法的であることを示す。すなわち排反な \mathcal{A} の元の族 $\{A_i\}_{i=1}^{\infty}$ に対して $\bigcup_{i=1}^{\infty} A_i =: A \in \mathcal{A}$ であったと仮定して $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$ を示す。 $A_i \in \mathcal{A}$ なので、 $B_i \in \mathcal{A}_1$ ならびに $C_i \in \mathcal{A}_2$ を用いて $A_i = B_i \times C_i$ と書ける。 $A \in \mathcal{A}$ についても同様に $A = B \times C$ を充たす $B \in \mathcal{A}_1$ ならびに $C \in \mathcal{A}_2$ がある。 A_i たちの排反性から、 $1_B 1_C = 1_{B \times C} = \sum_{i=1}^{\infty} 1_{B_i \times C_i} = \sum_{i=1}^{\infty} 1_{B_i} 1_{C_i}$ である。両辺を μ_1 に関して積分すると $\mu_1(B) 1_C = \sum_{i=1}^{\infty} \mu_1(B_i) 1_{C_i}$ である（ここで級数と積分が交換できることを使った）。同様に μ_2 に関して積分することで $\mu_1(B) \mu_2(C) = \sum_{i=1}^{\infty} \mu_1(B_i) \mu_2(C_i)$ すなわち $\mu(A) = \sum_{i=1}^{\infty} \mu(A_i)$ を得る。 ■

定義 4.2. 上記で得られた $\mathcal{A}_1 \otimes \mathcal{A}_2$ 上の測度を μ_1 と μ_2 の直積測度といい、 $\mu_1 \times \mu_2$ と書く。また、上記の構成を帰納的に繰り返すことで、 n 個の σ -有限測度空間 $(X_1, \mathcal{A}_1, \mu_1), (X_2, \mathcal{A}_2, \mu_2), \dots, (X_n, \mathcal{A}_n, \mu_n)$ に対して μ_i たちの直積測度 $\mu_1 \times \dots \times \mu_n: \mathcal{A}_1 \otimes \dots \otimes \mathcal{A}_n \rightarrow [0, \infty]$ を、 $\mu_1 \times \dots \times \mu_n(A_1 \times \dots \times A_n) = \prod_{i=1}^n \mu_i(A_i)$ を満たすように一意に定めることができる。特に $(X_1, \mathcal{A}_1, \mu_1) = (X_2, \mathcal{A}_2, \mu_2) = (X_n, \mathcal{A}_n, \mu_n) = (\mathbb{R}, \mathcal{B}, \lambda)$ のとき、 $(\mathbb{R}^n, \mathcal{B}^n)$ 上の直積測度 $\lambda^n := \lambda \times \dots \times \lambda$ のことを n 次元 Lebesgue 測度という。

Fubini の定理に話題を移そう。Fubini の定理を厳密に述べるには「a.e. に定義された関数」の概念を定式化した上で、積分の定義を拡張する必要があるので、まずその話からはじめる。いつものように (X, \mathcal{A}, μ) を測度空間とする。

定義 4.3. $\mu(N) = 0$ を満たすような $N \in \mathcal{A}$ のことを μ -零集合という。 μ が文脈から明らかな場合は単に零集合とか測度零集合という場合もある。

$x \in X$ に関する命題が (μ) -a.e. に成り立つとは、 μ -零集合 N が存在して、その命題がすべての $x \in X \setminus N$ に対して成り立つことをいう。

余談 4.4. 「零集合」のことを「0 集合」と綴る例は見たことがないが、「測度零集合」は「測度 0 集合」と綴ってもそんなに違和感を感じない。

定義 4.5. N を μ -零集合とし、 $f: X \setminus N \rightarrow \overline{\mathbb{R}}$ を写像とする。任意の $B \in \mathcal{B}^*$ に対して $f^{-1}(B) \in \mathcal{A}$ となるとき、 f は μ -a.e. に定義された可測関数であると言われる。

μ -a.e. に定義された可測関数 $f: X \setminus N \rightarrow \overline{\mathbb{R}}$ に対して、関数 $\tilde{f}: X \rightarrow \overline{\mathbb{R}}$ を

$$\tilde{f}(x) := \begin{cases} f(x) & \text{if } x \in X \setminus N \\ 0 & \text{if } x \in N \end{cases}$$

と定めれば、 $\tilde{f}: X \rightarrow \overline{\mathbb{R}}$ は可測である。このもとで、 $\int f d\mu := \int \tilde{f} d\mu$ と定める。

定理 4.6 (Fubini).

- (i) 非負可測関数 $f: X_1 \times X_2 \rightarrow [0, \infty]$ に対して、以下が成り立つ。
 - (i-a) 任意の $x^2 \in X_2$ に対し $x^1 \mapsto f(x^1, x^2)$ は $\mathcal{A}_1/\mathcal{B}$ 可測であり、
 - (i-b) $x^2 \mapsto \int f(x^1, x^2) d\mu_1(x^1)$ は $\mathcal{A}_2/\mathcal{B}$ 可測である。従って $\iint f(x^1, x^2) d\mu_1(x^1) d\mu_2(x^2)$ が定義でき、
 - (i-c) $\int f d(\mu_1 \times \mu_2) = \iint f(x^1, x^2) d\mu_1(x^1) d\mu_2(x^2)$.
- (ii) $f \in L^1(X_1 \times X_2, \mathcal{A}_1 \otimes \mathcal{A}_2, \mu_1 \times \mu_2)$ に対して、以下が成り立つ。
 - (ii-a) 任意の $x^2 \in X_2$ に対し $x^1 \mapsto f(x^1, x^2)$ は $\mathcal{A}_1/\mathcal{B}$ 可測であり、
 - (ii-b) $x^2 \mapsto \int f(x^1, x^2) d\mu_1(x^1)$ は μ_2 -a.e. に定義された $\mathcal{A}_2/\mathcal{B}$ 可測関数である。従って $\iint f(x^1, x^2) d\mu_1(x^1) d\mu_2(x^2)$ が定義でき、

$$(ii-c) \int f d(\mu_1 \times \mu_2) = \iint f(x^1, x^2) d\mu_1(x^1) d\mu_2(x^2).$$

証明.

(i) $\mathcal{A}_1 \otimes \mathcal{A}_2$ の元の指示関数に対して示せたならば、単調収束定理と補題 3.56 を用いることで一般の非負可測関数に対しても結論が従うので、以下では f が指示関数の場合を示す。

$$C := \{A \in \mathcal{A}_1 \otimes \mathcal{A}_2 \mid 1_A \text{ に対して (i-a), (i-b), (i-c) が成り立つ}\}$$

と置いて、 $C \supset \mathcal{A}_1 \otimes \mathcal{A}_2$ を示す。 π - λ 定理を踏まえれば、次の 3 点を示せば充分である。(α) $\mathcal{A} \subset C$. (β) \mathcal{A} は π -システムである。(γ) C は λ -システムである。このうち (α), (β) はすぐわかるので (γ) を確認する。

- まず $X_1 \times X_2 \in \mathcal{A}$ なので $X_1 \times X_2 \in C$ である。
 - $E, F \in C$ かつ $E \subset F$ ならば $F \setminus E \in C$ となることを確認する。まず $\mu(X), \mu(Y) < \infty$ となる場合を見よう。 $1_{F \setminus E} = 1_F - 1_E$ なので (i-a) はよい。任意の $x^2 \in X_2$ に対して $\int 1_{F \setminus E}(x^1, x^2) d\mu_1(x^1) = \int 1_F(x^1, x^2) d\mu_1(x^1) - \int 1_E(x^1, x^2) d\mu_1(x^1)$ なので (ii-b) もよい (ここで $\mu(X) < \infty$ なので右辺は必ず定義される)。最後に $\int 1_{F \setminus E} d(\mu_1 \times \mu_2) = \int 1_F d(\mu_1 \times \mu_2) - \int 1_E d(\mu_1 \times \mu_2) = \int 1_F d\mu_1 d\mu_2 - \int 1_E d\mu_1 d\mu_2 = \int 1_{F \setminus E} d\mu_1 d\mu_2$ なので (i-c) も従う (ここで $\mu(Y) < \infty$ なのでやはりこの過程で未定義な値は現れない)。
 $\mu(X) = \infty$ ないしは $\mu(Y) = \infty$ の場合は、 X と Y の σ -有限性より $A_i \uparrow X$ および $B_i \uparrow Y$ となる測度有限な集合の可算列 $\{A_i\}$ および $\{B_i\}$ がとれる。 E, F をそれぞれ $E \cap A_i \times B_i, F \cap A_i \times B_i$ に置き換えれば上記の考察がそのまま適用できるので、適用した上で単調収束定理を用いればよい。
 - $E_n \in C$ かつ $E_n \uparrow E$ のときに $E \in C$ を示す。 $\lim_{n \rightarrow \infty} 1_{E_n} = 1_E$ なので (i-a) はよい。また単調収束定理より任意の $x^2 \in X_2$ に対して $\int 1_{E_n}(x^1, x^2) d\mu_1(x^1) \uparrow \int 1_E(x^1, x^2) d\mu_1(x^1)$ なので (i-b) もよい。 $\int 1_{E_n} d\mu_1 d\mu_2 = \int 1_{E_n} d(\mu_1 \times \mu_2)$ の両辺に単調収束定理を適用することで $\int 1_E d\mu_1 d\mu_2 = \int 1_E d(\mu_1 \times \mu_2)$ すなわち (i-c) を得る。
- (ii) $f = f^+ - f^-$ と分解すれば、 f^+ および f^- は非負可測なので (i-a), (i-b), (i-c) が成り立つ。従って (ii-a) はよい。続いて、 $x^2 \mapsto \int f^+ d\mu_1(x^1)$ は μ_2 -a.e. に有限な関数である (そうでないとすると、 $f \in L^1$ 、とくに $f^+ \in L^1$ であることに矛盾する)。同様に $x^2 \mapsto \int f^- d\mu_1(x^1)$ も μ_2 -a.e. に有限である。従って $x^2 \mapsto \int f d\mu_1 = \int f^+ - f^- d\mu_1$ は μ_2 -a.e. に定義された $\mathcal{A}_2/\mathcal{B}$ -可測関数である*25。なので (ii-c) もよい。

■

*25 主に私自身の備忘のため、この部分についてもう少し述べておこう。 $x^2 \mapsto \int f^+ d\mu_1(x^1)$ が有限値を取らない x^2 の集合を N_+ 、 $x^2 \mapsto \int f^- d\mu_1(x^1)$ が有限値を取らない x^2 の集合を N_- と置く。 $x^2 \in N_+ \cap N_-$ である場合 $x^2 \mapsto \int f(x^1, x^2) d\mu_1(x^1)$ は定義できないが、 $\mu_1(N_+) = \mu_1(N_-) = 0$ であるので、特に $\mu(N_+ \cap N_-) = 0$ である。従って $x \mapsto \int f(x^1, x^2) d\mu_1(x^1)$ は μ_2 -a.e. に定義されている。

余談 4.7. σ -有限性を用いて有限測度の場合に考察を帰着する部分について、より簡潔に「 μ_1, μ_2 の σ -有限性より、 $\mu_1(X_1) < \infty, \mu_2(X_2) < \infty$ の場合を証明すれば充分である」とだけ書いてある文献もある。手の内を一通り見たあとであれば、このように書いてしまうほうが簡潔でよいだろう。

余談 4.8. Spivak を見てもらうとわかるが、被積分関数の連続性まで課してしまえば Riemann 積分の場合でも主張はすっきりするけれども、そうでない場合までカバーしようとする途端に主張がごたついて見える。積分の順序を入れ替えてよい充分条件をシンプルに記述するというのは実用上の重要性を持つと思う。一方、可測関数というものがずいぶんいろいろな関数を含んでいて、そこに難しさを押し付けているだけでは？といわれると返す言葉もないですが…

Fubini の定理の応用は数多いと思われる。ここではそのごく一部を紹介する。手始めに、以前証明を保留した Clairaut の定理（定理 2.18）に証明をつけることにする。

定理 2.18. $f: \mathbb{R}^n \rightarrow \mathbb{R}$ が C^2 級関数であるならば、 $\partial_{i,j}f = \partial_{j,i}f$ である。

証明. 必要ならば変数を並べ替え、更に他の変数をすべて止めて考えることで、最初から $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $i = 1, j = 2$ であるとして証明して構わないのでそうする。対称性より、ある点 $a \in \mathbb{R}^2$ において $\partial_{1,2}f(a) - \partial_{2,1}f(a) > 0$ であると仮定して矛盾を導けばよい。 f が C^2 級なので、 $\partial_{1,2}f$ ならびに $\partial_{2,1}f$ はいずれも連続関数である。従って a を含む適当な閉方体 $A = [a, b] \times [c, d]$ が存在して、 $x \in A$ ならば $\partial_{1,2}f(x) - \partial_{2,1}f(x) > 0$ である。このことより $\int_A (\partial_{1,2}f - \partial_{2,1}f) d\lambda^2 > 0$ を得る。 $\partial_{1,2}f - \partial_{2,1}f$ は連続関数なので、コンパクト集合 A 上で最大値と最小値を持つから特に絶対可積分である。従って Fubini の定理より $\int_A (\partial_{1,2}f - \partial_{2,1}f) d\lambda^2 = \int_{[c,d]} \int_{[a,b]} (\partial_{1,2}f(x,y) - \partial_{2,1}f(x,y)) d\lambda(x) d\lambda(y) = \int_{[c,d]} \int_{[a,b]} \partial_{1,2}f(x,y) d\lambda(x) d\lambda(y) - \int_{[a,b]} \int_{[c,d]} \partial_{2,1}f(x,y) d\lambda(y) d\lambda(x)$ である。微積分学の基本定理を 2 回使うと $\int_{[c,d]} \int_{[a,b]} \partial_{1,2}f(x,y) d\lambda(x) d\lambda(y) = \int_{[a,b]} \int_{[c,d]} \partial_{2,1}f(x,y) d\lambda(y) d\lambda(x)$ を得るが、これは $\int_A (\partial_{1,2}f - \partial_{2,1}f) d\lambda^2 = 0$ を含意するので矛盾である。 ■

命題 4.9 (Cavalieri の原理). $A, B \in \mathcal{B}^3$ とする。 $c \in \mathbb{R}$ に対して $A_c := \{(x, y) \in \mathbb{R}^2 \mid (x, y, c) \in A\}$, $B_c := \{(x, y) \in \mathbb{R}^2 \mid (x, y, c) \in B\}$ と定める。任意の $c \in \mathbb{R}$ に対して $\lambda^2(A_c) = \lambda^2(B_c)$ であるならば、 $\lambda^3(A) = \lambda^3(B)$ である。

証明. 指示関数の非負性と Fubini より $\int 1_A d\lambda^3 = \int \left(\int 1_{A_c} d\lambda^2 \right) d\lambda(c) = \int \left(\int 1_{B_c} d\lambda^2 \right) d\lambda(c) = \int 1_B d\lambda^3$. ■

4.2 1 の分割

この項の目標は次の定理の証明である。

定理 4.10 (1 の分割). $A \subset \mathbb{R}^n$ を部分集合, \mathcal{O} を A の開被覆とする. このとき, 次の 4 条件を充たすような C^∞ 級関数の可算族 Φ が存在する.

- (i) 任意の $\varphi \in \Phi$ と $x \in A$ に対して $0 \leq \varphi(x) \leq 1$.
- (ii) 任意の $x \in A$ に対し, 次の条件を充たすような開集合 V が存在する: $x \in V$ であり, 「 V 上で恒等的に 0」でないような $\varphi \in \Phi$ は有限個.
- (iii) 任意の $x \in A$ に対して $\sum_{\varphi \in \Phi} \varphi(x) = 1$. ここで, 条件 (ii) より, x で非零の値を持つような $\varphi \in \Phi$ は有限個しかないから, 左辺が意味を持つことに注意する.
- (iv) 任意の $\varphi \in \Phi$ に対し, 次の条件を充たす $U_\varphi \in \mathcal{O}$ が存在する: U_φ に含まれるコンパクト集合 K_φ があり, φ は K_φ の外で 0 である. 特に, $\text{supp } \varphi := \overline{\{x \mid \varphi(x) \neq 0\}} \subset U_\varphi$ が成り立っている.

条件 (i)-(iii) を充たすような Φ を, A 上の 1 の (C^∞ 級) 分割という. A 上の 1 の分割が更に (iv) を充たすとき, Φ は \mathcal{O} に従属している, といわれる.

証明には隆起関数の存在を含めたいいくつかの命題を使う. この証明で使う命題たちを再掲する.

命題 2.41. $K \subset A \subset \mathbb{R}^n$ とし, 更に K をコンパクト, A を開集合とする. このとき, $K \subset L \subset A$ を充たすコンパクト集合 L であって, L の内部に K を含むようなものがある. 更に C^∞ 級関数 $h: \mathbb{R}^n \rightarrow \mathbb{R}$ であって, K 上では正, L の外部では 0 であるようなものがある.

定理 2.29 (隆起関数の存在). \mathbb{R} のコンパクト集合 K と開集合 A を, $K \subset A \subset \mathbb{R}$ を充たすように取る. このとき, C^∞ 級関数 $\varphi: \mathbb{R} \rightarrow [0, 1]$ であって, K 上では $\varphi(x) = 1$, A に含まれるあるコンパクト集合の外では $\varphi(x) = 0$ を充たすようなものが存在する.

証明. 証明を 4 段階に分割する.

- (i) A がコンパクトの場合. \mathcal{O} から有限個を選んだ $\{U_1, \dots, U_m\}$ が A の被覆になっている. $\{U_1, \dots, U_m\}$ に従属する A 上の 1 の分割を構成すればよい. 以下, $K_i \subset U_i$ を充たすようなコンパクト集合 K_i であって, K_i たちの内部が A の被覆になるものが存在した, すなわち $\{K_1^\circ, \dots, K_m^\circ\}$ が A の被覆になっているようなものが取れたとする. このとき命題 2.41 によって, $K_i \subset L_i \subset U_i$ を充たすコンパクト集合 L_i ならびに C^∞ 級関数 $\psi_i: \mathbb{R}^n \rightarrow \mathbb{R}$ であって, K_i 上で正で, L_i の外部で 0 になっているようなものが存在する. このとき, ある開集

合 $U \supset \bigcup_i K_i^\circ$ が存在して、 U 上で $\sum_{j=1}^m \psi_j > 0$ である。

$$\varphi_i(x) := \begin{cases} \frac{\psi_i(x)}{\sum_{j=1}^m \psi_j(x)} & \text{if } x \in U, \\ 0 & \text{otherwise,} \end{cases}$$

によって関数 φ_i を定める。また、定理 2.29 より、 C^∞ 級関数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ であって、 A 上で 1, U に含まれるコンパクト集合の外で 0 であるようなものが存在する。 $\Phi := \{f \cdot \varphi_i\}_{i=1}^m$ とすればよい^{*26}。

あとは $K_i \subset U_i$ を充たすようなコンパクト集合 K_i であって、 K_i たちの内部が A の被覆になるものが存在することを示せばよい。 m に関する帰納法により、 $\{K_1^\circ, \dots, K_i^\circ, U_{i+1}, \dots, U_m\}$ が A の被覆になっていたとする。 $M_{i+1} := A \setminus (K_1^\circ \cup K_2^\circ \cup \dots \cup K_i^\circ \cup U_{i+2} \cup \dots \cup U_m)$ と置くと、これは有界閉集合なのでコンパクトであり、作り方より $M_{i+1} \subset U_{i+1}$ を充たす。したがって命題 2.41 によって $M_{i+1} \subset K_{i+1} \subset U_{i+1}$ を充たすコンパクト集合 K_{i+1} があり、 $M_{i+1} \subset K_{i+1}^\circ$ を充たす。従って $\{K_1^\circ, \dots, K_{i+1}^\circ, U_{i+2}, \dots, U_m\}$ が A の被覆になるから帰納法が進む。

- (ii) A がコンパクト集合による充填をもつ、即ちコンパクト集合の列 $K_1 \subset K_2 \subset \dots$ であって、 $K_i \subset K_{i+1}^\circ$ ならびに $K_i \uparrow A$ を充たすようなものがある場合。 $L_i := K_i \setminus K_{i-1}^\circ$ と置くと、 L_i はコンパクトで、 $\mathcal{O}_i := \{O \cap (K_{i+1}^\circ \setminus K_{i-2}^\circ) \mid O \in \mathcal{O}\}$ は L_i の開被覆である。従って \mathcal{O}_i に従属した（すなわち \mathcal{O} にも従属した） L_i 上の 1 の分割 Φ_i が取れる。 $\sigma(x) := \sum_{i=1}^\infty \sum_{\varphi \in \Phi_i} \varphi(x)$ と定めると、 σ は各点で有限和なので実数値 C^∞ 級関数になり、 \bar{A} を含む開集合 U 上で正である。定理 2.29 を用いて $f: \mathbb{R}^n \rightarrow \mathbb{R}$ であって A 上で 1, U に含まれるコンパクト集合の外で 0 であるようなものを取った上で、 $\Phi := \{f \cdot (\varphi/\sigma) \mid \varphi \in \bigcup_i \Phi_i\}$ と置けば、これが求める 1 の分割である。
- (iii) A が開集合の場合。 $K_i := \{x \in A \mid \|x\| \leq i, d(x, \partial A) \leq 1/i\}$ とすれば $\{K_i\}$ が A のコンパクト集合による充填になるので、前段に帰着する。
- (iv) A が一般の集合の場合。 \mathcal{O} に属する開集合すべての和集合を U とすると、 U は開集合なので、前段により U 上の 1 の分割 Φ であって \mathcal{O} に従属するものが取れる。 Φ は A 上の 1 の分割でもある。

■

余談 4.11. このノートでは 1 の分割を変数変換公式の証明で使いたないので、このタイミングで証明した。Spivak でも変数変換公式の証明は 1 の分割に拠っているが、それ以外にも Darboux 積分を非有界関数に拡張するのに 1 の分割を用いている。このノートでは測度論に基づいて積分を定義したので、積分の定義には 1 の分割は特に必要なくなっている。

^{*26} 「 $f \cdot \varphi_i: \mathbb{R}^n \rightarrow \mathbb{R}$ が C^∞ 級である」ことは非自明だと思うので、この点について補っておく； $f \cdot \varphi_i$ が U において C^∞ 級であることはよい。 $f \cdot \varphi_i$ は U に含まれるコンパクト集合の外で 0 なので、 ∂U の近傍および U の外部において $f \cdot \varphi_i$ は恒等的に 0 である。

問 4.12 ().** 1 の分割の (このノートの主目標からは外れる) 応用としては Riesz-Markov-Kakutani の表現定理がある. S を局所コンパクト Hausdorff 空間とし, $C(S)$ を S 上の実数値連続関数からなる線型空間とする. 線型形式 $L: C(S) \rightarrow \mathbb{R}$ が次の意味で非負であったとする; $f \geq 0$ ならば $L(f) \geq 0$. このとき, S 上の有限 Borel 測度 μ が一意的に存在して $L(f) = \int f d\mu$ が成り立つ. この表現定理は Prohorov の定理の証明に用いる; 完備可分距離空間 S 上の確率測度の集合を $P(S)$ とするとき, $\Pi \subset P(S)$ が一様緊密であることと (弱位相, 即ち分布収束を距離化する位相に関して) 相対コンパクトであることは同値である.

4.3 変数変換公式

いわゆる重積分の変数変換公式を証明する. 証明の本質的な流れは Spivak にあるそれとほぼ変わらないが, 測度論の議論で片付けられる部分は (折角なので) 最大限そのようにして進めることにする.

手始めに, 測度に関する積分の変数変換公式から始める. (X_1, \mathcal{A}_1) ならびに (X_2, \mathcal{A}_2) をそれぞれ可測空間とする.

定義 4.13. μ を \mathcal{A}_1 上の測度, $\varphi: X_1 \rightarrow X_2$ を可測写像とする. このとき, $\varphi_*\mu: \mathcal{A}_2 \rightarrow [0, \infty]$ を $\varphi_*\mu(A) := \mu(\varphi^{-1}(A))$ によって定めれば, $\varphi_*\mu$ は \mathcal{A}_2 上の測度になる. $\varphi_*\mu$ を μ の φ による像測度, ないしは押し出しという.

定理 4.14 (変数変換公式). $\varphi: X_1 \rightarrow X_2$ ならびに $f: X_2 \rightarrow \mathbb{R}$ を可測関数とする. $\int f d(\varphi_*\mu)$ と $\int f \circ \varphi d\mu$ のどちらかが定義されるならば, もう片方も定義され, 両者は同じ値を取る.

証明. まず $A \in \mathcal{A}_2$ に対して $f = 1_A$ である場合を考える. このときは $\int f d(\varphi_*\mu) = \varphi_*\mu(A) = \mu(\varphi^{-1}(A)) = \int 1_{\varphi^{-1}(A)} d\mu = \int 1_A \circ \varphi d\mu$ なのでよい. 積分の線型性より任意の単関数に対しても主張が成り立つから, 非負可測関数についても正しい. 一般の可積分関数に対しては正部分と負部分に分けてここまでの議論を適用すればよい. ■

余談 4.15. $A \in \mathcal{A}_2$ に対して $f = 1_A$ である場合を除いては, 積分の構成をなぞれば成り立つことが示せた. このことを以て, 「積分の構成の仕方より, $A \in \mathcal{A}_2$ に対して $f = 1_A$ である場合を示せば充分である」と初手で言ってしまうても証明としては正しいとわかる. とはいえ, いきなりそう言われると「ほんとか?」と思ってしまうので, やや冗長に感じられるのを承知の上で書き下してみることにした.

定理 4.16 (\mathbb{R}^n における変数変換公式). $A \subset \mathbb{R}^n$ を開集合, $f: \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ を非負可測または $f \in$

$L^1(\mathbb{R}^n, \mathcal{B}^n, \lambda^n)$ とし, $\varphi: A \rightarrow \varphi(A)$ を C^1 級微分同相とする. このとき,

$$\int_{\varphi(A)} f(y) d\lambda^n(y) = \int_A f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x).$$

証明. A 全体ではなく, 各 $a \in A$ に対して a を含む開集合 $U_a \subset A$ 上で定理が成り立つことを証明すればよいことが次のようにわかる; 各 U_a 上で証明できたとして, $\varphi(A)$ の開被覆 $\mathcal{O} := \{\varphi(U_a) \mid x \in A\}$ に従属した $\varphi(A)$ 上の 1 の分割 Ψ を取る. 1 の分割の定義より $\psi \in \Psi$ はある $\varphi(U) \subset \varphi(A)$ の外で 0 なので,

$$\begin{aligned} \int_{\varphi(A)} (\psi \cdot f)(y) d\lambda^n(y) &= \int_{\varphi(U)} (\psi \cdot f)(y) d\lambda^n(y) \\ &= \int_U (\psi \cdot f) \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x) \\ &= \int_A (\psi \cdot f) \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x) \end{aligned}$$

が成り立つ. 双方を ψ に関して和をとれば

$$\sum_{\psi \in \Psi} \int_{\varphi(A)} (\psi \cdot f)(y) d\lambda^n(y) = \sum_{\psi \in \Psi} \int_A (\psi \cdot f) \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x)$$

となり, 単調収束定理より総和と極限を入れ替えて定理の結論を得る.

また, 写像 $x \mapsto |\det J_\varphi(x)|$ は可測なので, $\mu_\varphi(U) := \int_U |\det J_\varphi(x)| d\lambda^n(x)$ は A 上の測度を定め, 更に

$$\int_U f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x) = \int_U f \circ \varphi(x) d\mu_\varphi(x)$$

が成り立つ. 測度の変数変換公式より

$$\int_U f \circ \varphi(x) d\mu_\varphi(x) = \int_{\varphi(U)} f(y) d(\varphi_* \mu_\varphi)(y)$$

であるから, 定理の主張は $\int_{\varphi(U)} f(y) d\lambda^n(y) = \int_{\varphi(U)} f(y) d(\varphi_* \mu_\varphi)(y)$ と同値である (すなわち, この等式を示せばよい). 積分の構成より, $f = 1$ の場合を証明すればよく, 更に $\int_{\varphi(U)} d(\varphi_* \mu_\varphi) = \mu_\varphi(\varphi^{-1}(\varphi(U))) = \mu_\varphi(U) = \int_U |\det J_\varphi(x)| d\lambda^n(x)$ である.

結局, 各 $x \in A$ に対して x を含む開集合 $U \subset A$ が存在し, $\int_U |\det J_\varphi(x)| d\lambda^n(x) = \int_{\varphi(U)} d\lambda^n$ が成り立つことをいえばよいとわかった. 証明をいくつかのステップに分割する.

- 主張が合成に関して閉じていることを言う. 即ち, ふたつの微分同相 $\varphi_1: A_1 \rightarrow \varphi_1(A_1)$, $\varphi_2: A_2 \rightarrow \varphi_2(A_2)$ が $\varphi_1(A_1) \subset A_2$ を充たしており, 更に $(\varphi_1)_* \mu_{\varphi_1} = \lambda^n$ かつ $(\varphi_2)_* \mu_{\varphi_2} = \lambda^n$ で

あるならば, $(\varphi_2 \circ \varphi_1)_* \mu_{\varphi_2 \circ \varphi_1} = \lambda^n$ が成り立つことを示す.

$$\begin{aligned}
(\varphi_2 \circ \varphi_1)_* \mu_{\varphi_2 \circ \varphi_1}(U) &= \mu_{\varphi_2 \circ \varphi_1}(\varphi_1^{-1}(\varphi_2^{-1}(U))) \\
&= \int_{\varphi_1^{-1}\varphi_2^{-1}(U)} |\det J_{\varphi_2 \circ \varphi_1}(x)| d\lambda^n(x) \\
&= \int_{\varphi_1^{-1}\varphi_2^{-1}(U)} |\det J_{\varphi_2}(\varphi_1(x))| \cdot |\det J_{\varphi_1}(x)| d\lambda^n(x) \\
&= \int_{\varphi_1^{-1}\varphi_2^{-1}(U)} |\det J_{\varphi_2}(\varphi_1(x))| d\mu_{\varphi_1}(x) \\
&= \int_{\varphi_2^{-1}(U)} |\det J_{\varphi_2}(y)| d(\varphi_1)_* \mu_{\varphi_1}(y) \\
&= \int_{\varphi_2^{-1}(U)} |\det J_{\varphi_2}(y)| d\lambda^n(y) = \lambda^n(U)
\end{aligned}$$

なのでよい.

- φ が正則な線型変換の場合.
 - 実数 m と, ある i に対して $\varphi(e_i) = me_i$, $\varphi(e_j) = e_j$ ($j \neq i$) の場合. Fubini の定理と 1 変数関数に対する置換積分よりよい.
 - ある i, j に対して $\varphi(e_i) = e_i + e_j$, $\varphi(e_k) = e_k$ ($k \neq i, j$) の場合. Fubini の定理と 1 次元 Lebesgue 測度の並進不変性から従う.
 - ある i, j に対して $\varphi(e_i) = e_j$, $\varphi(e_j) = e_i$, $\varphi(e_k) = e_k$ ($k \neq i, j$) の場合. Fubini の定理で示せる.
 - 一般の線型変換は上 3 つの線型変換の合成で書けるので, 並びに前段で示した「主張が合成に関して閉じていること」と併せて結果を得る.
- 一般の φ の場合. $J_\varphi(a)^{-1} \circ \varphi$ に対して主張を示せば φ に対しても結果が従うので, 最初から $J_\varphi(a) = I_n$ (単位行列) だと仮定して証明してよい. 次元 n に関する帰納法に依る. $\varphi_1(x) := (\varphi^1(x), \varphi^2(x), \dots, \varphi^{n-1}(x), x^n)$ と定めれば $J_{\varphi_1}(a) = I_n$ なので, 逆関数定理より φ_1 は a の近傍 V' で定義された微分同相である. したがって $\varphi_2(x) := (x^1, x^2, \dots, x^{n-1}, \varphi^n(\varphi_1^{-1}(x)))$ が定義でき, $\varphi = \varphi_2 \circ \varphi_1$ が成り立つ. したがって φ_2 は $\varphi_1(a)$ の近傍で定義された微分同相である. 主張が合成に関して閉じていることから, φ_1 および φ_2 の双方に対して主張を示せばよい. 議論は同様なので φ_1 に対してのみ示す. a を含む開方体 $V \subset V'$ を取る. V は $W \subset \mathbb{R}^{n-1}$ 並びに $a^n, b^n \in \mathbb{R}$ を用いて $V = W \times (a^n, b^n)$ と書ける. Fubini の定理より

$$\int_V |\det J_{\varphi_1}(x)| d\lambda^n(x) = \int_{(a^n, b^n)} \int_W |\det J_{\varphi_1}(x)| d\lambda^{n-1}(x^1, \dots, x^{n-1}) dx^n$$

である. 関数 $\xi_{x^n}: W \rightarrow \mathbb{R}^{n-1}$ を $\xi_{x^n}(x^1, \dots, x^{n-1}) := ((\varphi_1)^1(x), (\varphi_1)^2(x), \dots, (\varphi_1)^{n-1}(x))$ で定めれば, $\det J_{\varphi_1}(x) = \det J_{\xi_{x^n}}(x^1, \dots, x^{n-1})$ が成り立つ. したがって

$$\int_{(a^n, b^n)} \int_W |\det J_{\varphi_1}(x)| d\lambda^{n-1}(x^1, \dots, x^{n-1}) dx^n = \int_{(a^n, b^n)} \int_W |\det J_{\xi_{x^n}}(x^1, \dots, x^{n-1})| d\lambda^{n-1}(x^1, \dots, x^{n-1}) dx^n$$

である. $\xi_{x^n}(W) = \varphi_1(W \times \{x^n\})$ および帰納法の仮定と併せて

$$\begin{aligned} \int_{(a^n, b^n)} \int_W |\det J_{\xi_{x^n}}(x^1, \dots, x^{n-1})| d\lambda^{n-1}(x^1, \dots, x^{n-1}) dx^n &= \int_{(a^n, b^n)} \int_{\varphi_1(W \times \{x^n\})} d\lambda^{n-1}(x^1, \dots, x^{n-1}) dx^n \\ &= \int_V d\lambda^n(x) \end{aligned}$$

を得る. ■

問 4.17. (X, \mathcal{A}, μ) を測度空間, $f: X \rightarrow \mathbb{R}$ を可測写像とすると, $\nu(A) := \int_A f d\mu$ は X 上の測度を定める. 更に可測写像 $g: X \rightarrow \mathbb{R}$ に対して, $\int gf d\mu$ または $\int g d\nu$ の片方が定義されるならば, もう片方も定義され, 両者は同じ値を取る.

問 4.18 ().** このノートでは Spivak にならって 1 の分割を使い, 局所的に変数変換公式を証明してそれを大域的につなぎ合わせるということをした. それ以外に, 被積分関数をよりかんたんな関数で近似して結論を得る, という証明の戦略もある. この方針に基づく証明は Jost “Postmodern Analysis” にある.

ところで, 変数変換公式の「 φ が C^1 級微分同相」という仮定は弱めることができる. その正当化に役買うのが Sard の定理である.

定理 4.19 (Sard). $A \subset \mathbb{R}^n$ を開集合, $g: A \rightarrow \mathbb{R}^n$ を単射な C^1 級写像とする. このとき, $B := \{x \in A \mid \det J_g(x) = 0\}$ と置くと, $\lambda(g(B)) = 0$ が成り立つ.

Sard の定理の証明は後回しにして, 先に変数変換公式の仮定を弱めてしまうことにしよう.

定理 4.20 (\mathbb{R}^n における変数変換公式の強い形). $A \subset \mathbb{R}^n$ を開集合, $f \in L^1(\mathbb{R}^n, \mathcal{B}^n, \lambda^n)$ とし, $\varphi: A \rightarrow \mathbb{R}^n$ を単射な C^1 級写像とする. このとき,

$$\int_{\varphi(A)} f(y) d\lambda^n(y) = \int_A f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x).$$

証明. $B := \{x \in A \mid \det J_\varphi(x) = 0\}$ と置き, 連続写像 $x \mapsto \det J_\varphi(x)$ を h と書くことにすると, $h^{-1}(\{0\}) = B$ なので B は閉集合である. したがって $A \setminus B$ は開集合であるから, \mathbb{R}^n における変数変換公式より

$$\int_{\varphi(A \setminus B)} f(y) d\lambda^n(y) = \int_{A \setminus B} f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x)$$

を得る. Sard の定理より $\varphi(B)$ は零集合であるから, $\int_{\varphi(B)} f(y) d\lambda^n(y) = 0$ である. また仮定よ

り $\int_B f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x) = 0$ である。以上より

$$\int_{\varphi(A \setminus B)} f(y) d\lambda^n(y) + \int_{\varphi(B)} f(y) d\lambda^n(y) = \int_{A \setminus B} f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x) + \int_B f \circ \varphi(x) \cdot |\det J_\varphi(x)| d\lambda^n(x)$$

となって定理が得られる。 ■

問 4.21 ().** ここで扱った Sard の定理は、実際には微分可能多様体に対して拡張できる。また、Sard の定理の応用として挙がるのが写像度の概念である。写像度の応用先として有名なものは私の知る限り 2 つある；代数学の基本定理と、つむじ頭定理である。

問 4.22 (*). 実際には変数変換公式の仮定は更に弱めることができるらしい。例えば単射性を落とすといった拡張の仕方があり得るらしい。ところで極座標変換くらいだったら微分同相という仮定のついたバージョンからでも出せる（ことを後で見る）。仮定を弱めたいくなる実用的な例はいくらでもあると思うのだが、とりあえず思いついていない。

最後に、極座標変換を用いた Gauss 積分を例として扱おう。仮定を弱めておいて何なのだが、極座標変換の公式を出す上では弱い形の定理 4.16 で充分である。

例 4.23 (極座標変換). $P := (0, \infty) \times (0, 2\pi)$ と置いて、関数 $\varphi: P \rightarrow \mathbb{R}^2$ を $\varphi(r, \theta) = (r \cos \theta, r \sin \theta)$ によって定める。 φ は単射であり、Jacobi 行列は $J_\varphi(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$ なので、特に $\det J_\varphi(r, \theta) = r$ がわかるから、逆関数定理より φ を $\varphi(P)$ への写像とみなすと C^1 級微分同相である。ところで $A := \{(x, 0) \in \mathbb{R}^2 \mid x \geq 0\}$ と置くとこれは閉集合で、 $\varphi(P) = \mathbb{R}^2 \setminus A$ なので、定理 4.16 より、 $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ が非負可測であるか $f \in L^1(\mathbb{R}^2)$ であるならば

$$\int_{\varphi(P)} f(x, y) d\lambda^2(x, y) = \int_P (f \circ \varphi)(r, \theta) \cdot |\det J_\varphi(r, \theta)| d\lambda^2(r, \theta)$$

すなわち

$$\int_{\mathbb{R}^2 \setminus A} f(x, y) d\lambda^2(x, y) = \int_P (f \circ \varphi)(r, \theta) r d\lambda^2(r, \theta)$$

を得る。 A が測度 0 なので、左辺は \mathbb{R}^2 全域での積分に置き換えることができる。また右辺に対して Fubini の定理を適用することで、

$$\int f(x, y) dx dy = \int_0^\infty \int_0^{2\pi} (f \circ \varphi)(r, \theta) r d\theta dr$$

を得る。

例 4.24 (Gauss 積分). $\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi}$ を証明してみよう。そのためには、 $\int_{\mathbb{R}^2} e^{-x^2-y^2} dx dy = \pi$ を示せばよい。これが示せれば Fubini の定理より $(\int e^{-x^2} dx)(\int e^{-y^2} dy) = \pi$ となって結論が得られる

からである．極座標変換をすると，求める積分は

$$\int_0^\infty \int_0^{2\pi} e^{-r^2} r \, d\theta \, dr = 2\pi \int_0^\infty r e^{-r^2} \, dr$$

となる． $r e^{-r^2}$ は $(0, \infty)$ 上非負なので，広義 Riemann 積分として計算してよいから，値は π である．