Fréchet 微分入門

木原浩貴

2008 年 11 月 20 日

概要

初歩の微積分学では関数の微分を論ずる場合、関数の定義や値域は実数空間 $\mathbb{R}^n$ であるが、それをノルム空間に一般化したものが Fréchet 微分である。

Fréchet 微分の利点はいろいろあるが、その一つとして高階微分を含んだ公式が見やすくなることが挙げられる。本書ではその点に重きを置いて解説を試みた。

Fréchet 微分は無限次元ノルム空間上の微積分学への応用ももっているが、それを論ずるには関数空間の位相に関する繊細な考察を必要とするため、本書ではそれに触れる余裕がなかった。

読者は線形代数学と微分積分学の初歩的な知識をすでにもっておられることを仮定している。ただし偏微分の順序に関する Schwartz の定理と Young の定理は微積分学の入門書で必ずしも触れられているとは限らないから、本書において詳しく説明しておいた。

1　ベクトル空間のノルム

■**ノルム**　本書で $E, E', E'', \ldots$ は体 $\mathbb{R}$ 上のベクトル空間を表す。

関数 $|\cdot| : E \to \mathbb{R}, x \mapsto |x|$ は以下の条件を満たすとき E の**ノルム (norm)** と呼ばれる：

1. $|x| \geq 0$,
2. $|x| = 0 \Rightarrow x = 0$
3. $|kx| = |k||x|$,
4. $|x + y| \leq |x| + |y|$

注意. 本書ではノルムは $\|\cdot\|$ と書かず、$|\cdot|$ と書くことにする。

最後の不等式から次の不等式が導かれる：

$$||x| - |y|| \leq |x - y| \tag{1}$$

ノルム $|\cdot|$ は E 上の距離 (関数) $d(x,y) = |x-y|$ を定め、距離 d は E の位相を定める。これをノルム $|\cdot|$ が定める位相と呼ぶ。

命題 1.1. ノルム $|\cdot|$ が定める位相に関して、ノルム $|\cdot| : E \to \mathbb{R},\ x \mapsto |x|$ は連続関数である。

これは不等式 (1) から明らかであろう。

例 1.1. $E = \mathbb{R}^n$ の場合、E のノルムには次のようなものがある ($x = (x^1 \ldots x^n)$ とする)：

$$
\begin{aligned}
|x|_1 &= \sum_{i=1}^{n} \left| x^i \right|, \\
|x|_2 &= \sqrt{\sum_{i=1}^{n} \left| x^i \right|^2}, \\
|x|_\infty &= \max_{1 \le i \le n} \left| x^i \right|
\end{aligned}
$$

これらが定める $\mathbb{R}^n$ の位相はいずれも $\mathbb{R}^n$ の通常の位相 ($\mathbb{R}$ の積位相) である (確認せよ)。

■同値なノルム　E の二つのノルム $|\cdot|_\alpha, |\cdot|_\beta$ は

$$A|x|_\alpha \le |x|_\beta \le B|x|_\alpha \quad (x \in E)$$

を満たすとき、互いに**同値**であるという。ここで A, B は正定数である。
次の命題は明らかであろう。

命題 1.2. $|\cdot|_\alpha, |\cdot|_\beta$ は同値のとき E 上に同じ位相を定める。逆に $|\cdot|_\alpha, |\cdot|_\beta$ が E 上に同じ位相を定めるならば、これらは互いに同値である。

既に述べたように E のノルムは E 上の距離関数を定義するから、コーシー列、完備ということにも意味がある。これらに関しても次の命題は明らかであろう。

命題 1.3. $|\cdot|_\alpha, |\cdot|_\beta$ を E の二つのノルムとし、これらは互いに同値であると仮定する。このとき E がノルム $|\cdot|_\alpha$ に関して完備ならば $|\cdot|_\beta$ に関しても完備で、その逆も成り立つ。

例 1.2. $E = \mathbb{R}^n$ のノルムで先に挙げた $|\cdot|_1, |\cdot|_2, |\cdot|_\infty$ は、どの二つも互いに同値である。$\mathbb{R}^n$ はこれらのノルムに関して完備である (確認せよ)。

■**ノルム空間と有界線形写像**　ノルムの指定されたベクトル空間を**ノルム空間**と呼ぶ。E, E' をノルム空間とすると線形写像 $f : E \to E'$ は

$$|f(x)| \leq K|x| \quad (x \in E)$$

を満たすとき、**有界**であると言われる。ここで K は正定数である。

命題 1.4. f が有界であるための必要十分条件は f が (E, E' のノルムがそれぞれ定める位相に関して) 連続であることである。

証明はやさしいので読者に委ねる。

補題 1.5. 明確のため、$\mathbb{R}^n$ はノルム $|\cdot|_\infty$ の指定されたノルム空間と考えることにする。E を任意のノルム空間 (無限次元でもいい) とし、$f : \mathbb{R}^n \to E$ を線形写像とする。このとき f は有界 (連続) である。

証明 $(\mathsf{e}_1 \ldots \mathsf{e}_n)$ を $\mathbb{R}^n$ の標準基底とする。すなわち $x = (x^1 \ldots x^n)$ ならば $x = \sum_i x^i \mathsf{e}_i$, $f(x) = \sum_i x^i f(\mathsf{e}_i)$。よって

$$|f(x)| \leq \sum_{i=1}^{n} \left|x^i\right| |f(\mathsf{e}_i)| \leq K|x|_\infty$$

ここで $K = \sum_i |f(\mathsf{e}_i)|$ とおいた。したがって f は有界。 □

定理 1.6. E を n 次元ノルム空間 ($n < +\infty$) とし、$(e_1 \ldots e_n)$ を E の基底、$f : E \to \mathbb{R}^n$ をその逆写像が

$$f^{-1}(x^1 \ldots x^n) = \sum_{i=1}^{n} x^i e_i$$

で定義される線形同型とする。$\mathbb{R}^n$ はノルム $|\cdot|_\infty$ の指定されたノルム空間と考える。このとき f は有界である。したがって前の結果と合わせると、f は同相 (双連続) である。

証明 $Q^n = \left\{ x' \in \mathbb{R}^n \,\middle|\, |x'| < 1 \right\}$, $\partial Q^n = \left\{ x' \in \mathbb{R}^n \,\middle|\, |x'| = 1 \right\}$ とおく。∂Q^n はコンパクトだから、f^{-1} の連続性により $f^{-1}(\partial Q^n)$ はコンパクトである。ゆえに $r = \min\left\{ |y| \,\middle|\, y \in f^{-1}(\partial Q^n) \right\}$ が存在する。$0 \notin \partial Q^n$ だから $0 = f^{-1}(0) \notin f^{-1}(\partial Q^n)$ で、ここから $r > 0$ であることが分かる。
$U = \left\{ y \in E \,\middle|\, |y| \leq r/2 \right\}$ とおく。$U \cap f^{-1}(\partial Q^n) = \emptyset$ だから $f(U) \cap \partial Q^n = \emptyset$ である。すなわち $x \in U$ ならば $|f(x)| > 1$ か $|f(x)| < 1$ のどちらかでなければならない。

今、$|f(x)| > 1$ と仮定する。$a = |f(x)|, y = a^{-1}x$ とおくと $0 < a^{-1} < 1$ だから $y \in U$ である。しかるに $|f(y)| = a^{-1}|f(x)| = 1, f(y) \in \partial Q^n$ だから矛盾が生じる。ゆえに $|f(x)| < 1$ が常に成り立たなければならない。すなわち $f(U) \subset Q^n$ で、f は有界である。 □

系 1.7. E を有限次元ノルム空間、E' をノルム空間とする (無限次元でもいい)。このとき任意の線形写像 $f : E \to E'$ は有界 (連続) である。

証明 $\varphi : E \to \mathbb{R}^n$ を線形同型とすると補題 1.5 で示したように、線形写像 $f \circ \varphi^{-1} : \mathbb{R}^n \to E'$ は連続である。一方、定理 1.6 により φ は連続。したがって $f = f \circ \varphi^{-1} \circ \varphi$ は連続である。
□

系 1.8. E, E' を n 次元ノルム空間 $(n < +\infty)$ とし、$f : E \to E'$ を線形同型とする。このとき f, f^{-1} は有界 (連続) である。つまり f は同相である。

系 1.9. E を有限次元空間とすると、任意の二つのノルム $|\cdot|_\alpha, |\cdot|_\beta$ は互いに同値である。

証明 E にノルム $|\cdot|_\alpha, |\cdot|_\beta$ を指定した空間をそれぞれ E_α, E_β と記す。$f : E_\alpha \to E_\beta$ を恒等写像とすると f, f^{-1} は有界である。ゆえに $|\cdot|_\alpha, |\cdot|_\beta$ は互いに同値である。 □

ここから分かるように、E が有限次元ならば、E のノルムの定める位相はノルムに無関係な一定の位相、線形同型 $f : E \to \mathbb{R}^n$ を同相とするような位相である (ここで $\mathbb{R}^n$ には $\mathbb{R}$ の積位相を入れている)。これを有限次元空間 E の "自然な" 位相と呼ぼう。E が有限次元の場合、E の自然な位相を扱う上でノルムは便利な道具にすぎず、自由に取り換えのきくものであると言うことができる。

E が無限次元ならば事情は一変し、E 上に互いに同値でないノルムを定義することができる。すなわち、それらは E 上に異なる位相を定めることになる。したがって位相を論ずる上でノルムの指定が本質的になってくるのは根底にあるベクトル空間 E が無限次元のときである。

例 1.3 (数列空間). 有限個の i を除いて後は $x^i = 0$ なる実数列 $x = (x^1, x^2, \ldots)$ のつくるベクトル空間を $\mathbb{R}^\omega$ と記す。

$\mathbb{R}^\omega$ は無限次元である。$\mathbb{R}^n$ との類比で $\mathbb{R}^\omega$ に次の三つのノルムを考えることができる：

$$
\begin{aligned}
|x|_1 &= \sum_{i=1}^{\infty} \left| x^i \right|, \\
|x|_2 &= \sqrt{\sum_{i=1}^{\infty} \left| x^i \right|^2}, \\
|x|_\infty &= \max_{1 \le i \le \infty} \left| x^i \right|
\end{aligned}
$$

このとき $|x|_\infty \le |x|_2 \le |x|_1$ であるが、$|x|_1 \le K|x|_2$ を満たす K も、$|x|_2 \le K|x|_\infty$ を満たす K も存在しないことを示すことができる。つまり $\mathbb{R}^\omega$ においては $|\cdot|_1, |\cdot|_2, |\cdot|_\infty$ は互いに同値でなく、$\mathbb{R}^\omega$ 上に異なる位相を定義する。

以下では、とくに断らない限り有限次元ベクトル空間だけを考えることにし、$E, E', \ldots$ は有限次元ベクトル空間を表すとする。

■作用素ノルム　(E, E') で E から E' への線形写像 (作用素ともいう) 全体のつくる空間を表す。

E, E' をノルム空間とすると任意の $A \in (E, E')$ は有界だから $A \in (E, E')$ の “ノルム” $|A|$ が

$$
\begin{aligned}
|A| &= \inf\left\{ K \ge 0 \,\middle|\, |A(x)| \le K|x| \right\} \\
&= \sup_{|x|=1} |A(x)| = \sup_{|x| \ne 0} \frac{|A(x)|}{|x|}
\end{aligned} \tag{2}
$$

により定義される。

練習問題 1.1. (2) が互いに同値な定義になっていることを確かめよ。また、関数 $|\cdot| : A \in (E, E') \mapsto |A| \in \mathbb{R}$ がノルムの公理を満たすことを確かめよ。

(2) で定義される関数 $|\cdot| : A \in (E, E') \mapsto |A| \in \mathbb{R}$ を**作用素ノルム**という。通常、(E, E') は作用素ノルムの指定されたノルム空間であると考える。このとき

$$
|A(x)| \le |A||x| \quad (A \in (E, E'),\ x \in E) \tag{3}
$$

が成り立つ。

$(E, E'; E'')$ で $E \times E'$ から E'' への**双線形写像**全体のつくるベクトル空間を表す。$A \in (E, E'; E'')$ に対し、$\hat{A} \in (E, (E', E''))$ を

$$
A(x, y) = \hat{A}(x)(y)
$$

で定義すると、写像 $A \mapsto \hat{A}$ は $(E, E' ; E'')$ と $(E, (E', E''))$ の同型を与える。したがって、E, E', E'' がノルム空間ならば $(E, E' ; E'')$ も作用素ノルムの与えられたノルム空間である。このとき

$$|A(x,y)| \le |A||x||y| \quad (A \in (E, E' ; E''), x \in E, y \in E') \tag{4}$$

が成り立つ。

同様に $(E_1, \ldots, E_p ; E'')$ で $E_1 \times \cdots \times E_p$ から E'' への p-**線形写像**全体のつくるノルム空間を表す。$(E_1, \ldots, E_p ; E'')$ と $(E_1, (E_2, \ldots, E_p ; E''))$ のあいだには自然な同型がある。すなわち $A \in (E_1, \ldots, E_p ; E'')$ に対して

$$A(x_1 \ldots x_p) = \hat{A}(x_1)(x_2 \ldots x_p) \tag{5}$$

で定義される $\hat{A}$ を対応させる写像がそれである。

■変換の引き起こす変換　$\psi : E' \to E''$ を線形変換とすると、ψ は

$$\psi_*(A) = \psi \circ A \tag{6}$$

で定義される線形変換

$$\psi_* : (E, E') \to (E, E'')$$

を引き起こす。ここで E は別の任意のベクトル空間であるが、E をひとつ決めたらあとはずっと固定して考える。よって ψ_* がまた線形変換

$$\psi_{**} : (E, E ; E') \to (E, E ; E'')$$

を引き起こす。ψ_{**} は ψ_{*^2} とも書かれる。以下同様で、すべての正整数 r について線形変換

$$\psi_{\underbrace{* \cdots *}_{r}} (= \psi_{*^r}) : (\underbrace{E \ldots E}_{r} ; E') \to (\underbrace{E \ldots E}_{r} ; E'')$$

が帰納的に定義される。正確に言うと、$\psi_{*^{r-1}}$ が定義されていると仮定して ψ_{*^r} が

$$\psi_{*^r}(A) = B \quad \Longleftrightarrow \quad \psi_{*^{r-1}} \circ \hat{A} = \hat{B} \tag{7}$$

として定義されるのである。ここで $\hat{A} \in (E, (\underbrace{E \ldots E}_{r-1} ; E'))$, $\hat{B} \in (E, (\underbrace{E \ldots E}_{r-1} ; E''))$ はそれぞれ $A \in (\underbrace{E \ldots E}_{r} ; E')$, $B \in (\underbrace{E \ldots E}_{r} ; E'')$ に対応する元である。

命題 1.10. 次の関係が成り立つ。

$$\psi_{*^r}(A) = \psi \circ A \tag{8}$$

証明 証明は r についての帰納法による。
$r = 1$ ： これは定義 (6) から自明。
$r > 1$ ： $r - 1$ まで成立と仮定。$\psi_{*^r}(A) = B$ とおくと、定義により

$$\psi_{*^{r-1}}(\hat{A}(\xi_1)) = (\psi_{*^{r-1}} \circ \hat{A})(\xi_1) = \hat{B}(\xi_1).$$

帰納法の仮定から

$$\psi \circ (\hat{A}(\xi_1)) = \hat{B}(\xi_1).$$

すなわち

$$\psi((\hat{A}(\xi_1))(\xi_2 \dots \xi_r)) = (\hat{B}(\xi_1))(\xi_2 \dots \xi_r),$$
$$\psi(A(\xi_1, \xi_2 \dots \xi_r)) = B(\xi_1, \xi_2 \dots \xi_r).$$

ここから $\psi \circ A = B$ を得る。 □

2　多変数微分学のおさらい

U を $\mathbb{R}^n$ の開集合とし、$a = (a^1 \dots a^n)$ を U 上の点、$x = (x^1 \dots x^n)$ を U 上を動く "多変数"、$f : U \to \mathbb{R}$ とする。$(\mathsf{e}_1 \dots \mathsf{e}_n)$ は $\mathbb{R}^n$ の標準基底とする。

■偏微分　$i = 1 \dots n$ とする。

$$\lim_{t \to 0} \frac{f(a + t\mathsf{e}_i) - f(a)}{t} = A_i \tag{9}$$

が存在するとき、f は a で変数 x^i に関して**偏微分可能**であるといい、(9) の右辺 の数 A_i を x^i に関する**偏微分係数**と呼ぶ。U の各点で偏微分可能ならばたんに偏微分可能という。そしてこのとき U の各点 x にその点における偏微分係数 A_i を割り当てる関数が定義される。この関数を

$$\frac{\partial f}{\partial x^i}, \quad \frac{\partial}{\partial x^i} f, \quad f_{x^i}, \quad \dots$$

等と書いて、変数 x^i に関する**偏導関数**と呼ぶ。

■**全微分**　f は以下の性質を満たすとき点 a で**微分可能**であるという：決まった数ベクトル $(A_1 \dots A_n)$ があり、任意の正数 ϵ に対し正数 δ があって

$$\left| f(x) - f(a) - (A_1(x^1 - a^1) + \cdots + A_n(x^n - a^n)) \right| \le \epsilon |x - a| \quad (x \in U, |x - a| \le \delta) \tag{10}$$

が成り立つ。ここで $|x - a| = \max_i \left| x^i - a^i \right|$ とする。このとき $\mathbb{R}^n$ 上の一次形式

$$(\xi^1 \dots \xi^n) \mapsto A_1\xi^1 + \cdots + A_n\xi^n$$

を f の点 a における**全微分**といい、$(df)_a$ と書く。

f が点 a で微分可能ならば、その点で各変数 x^i に関して偏微分が可能で、A_i が偏微分係数となっていることは見やすい。すなわちある点において微分可能ならば、その点ですべての変数に関して偏微分が可能である。

逆にある点ですべての変数に関して偏微分可能であっても微分可能であるとは限らない。

しかし次の定理が成り立つ。

定理 2.1. 関数 f は各変数 x^i に関して偏微分が可能で、n 個の偏導関数 $\partial f/\partial x^i$ $(i = 1 \dots n)$ がひとつを除いてすべて点 a で連続ならば、f は点 a で微分可能である。

証明はやさしいので読者に任せる (ヒント：微分法の平均値の定理を用いる)。

■**2 階の微分と微分の順序**　$1 \le i, j \le n$ とする。

$$\frac{\partial}{\partial x^j}\left(\frac{\partial f}{\partial x^i}\right) = \frac{\partial^2 f}{\partial x^i\,\partial x^j} = f_{x^i x^j}$$

等と書いて (添字の順序に注意)、f の 2 **階の偏導関数**と呼ぶ。2 階の偏導関数は n^2 通りあるが、

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i} \quad (i < j)$$

がほとんどの場合に成り立つ。つまり偏微分の順序は交換できる。実際、以下で証明するように f の 2 階偏導関数が連続ならば、つねに偏微分の順序は交換できる。

微分の順序が交換できるための十分条件として、次に述べる Schwartz の定理と Young の定理が有名である。定理とその証明では簡単のため関数 f は $\mathbb{R}^2$ の開集合 U で定義されているものとし、しかも U は矩形領域

$$U = \left\{ (x, y) \,\middle|\, |x - a|, |y - b| < r \right\}$$

であると仮定する。f の偏導関数 (あるいは偏微分係数) は f_x, f_y, f_{xy}, f_{yx} 等と書くことにする。ここでもちろん

$$f_{xy} = \frac{\partial^2 f}{\partial x\,\partial y} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

定理の証明では

$$\begin{aligned}\Delta(x,y) &= f(x,y) - f(x,b) - (\,f(a,y) - f(a,b)\,)\\ &= f(x,y) - f(a,y) - (\,f(x,b) - f(a,b)\,),\end{aligned}$$

$$\varphi(x,y) = f(x,y) - f(x,b),$$

$$\psi(x,y) = f(x,y) - f(a,y)$$

で定義される関数 $\Delta, \varphi, \psi : U \to \mathbb{R}$ を利用する。定義から明らかに

$$\Delta(x,y) = \varphi(x,y) - \varphi(a,y) \tag{11}$$
$$\Delta(x,y) = \psi(x,y) - \psi(x,b) \tag{12}$$

が恒等的に成り立つ。

f_x, f_y が存在すれば、$\varphi_x, \varphi_y, \psi_x, \psi_y$ も存在する。(11) を x の関数と見立てて平均値の定理を適用すると

$$\Delta(x,y) = \varphi_x(x',y)(x-a) = (\,f_x(x',y) - f_x(x',b)\,)\,(x-a) \tag{13}$$

ここで x' は a と x のあいだの数、すなわち $0 < \theta < 1$ なるある数 θ に対して $x' = a+\theta(x-a)$ となる数である。

定理 2.2 (Schwartz). $f : U \to \mathbb{R}$ とする。U 全体で f_x, f_y, f_{xy} が存在して、f_{xy} が点 (a,b) で連続ならば f_{yx} が点 (a,b) において存在し、$f_{xy}(a,b) = f_{yx}(a,b)$ となる。

証明 仮定から $\varphi_x, \varphi_y, \psi_x, \psi_y$ が存在し、(13) がなりたつ。f_{xy} も存在するから、(13) の $f_x(x',y) - f_x(x',b)$ にさらに平均値の定理を適用し

$$\Delta(x,y) = f_{xy}(x',y')(x-a)(y-b) \tag{14}$$

を得る。ここで y' は b と y のあいだの数である。
さて f_{xy} は (a,b) で連続だから任意の正数 ϵ に対して正数 $\delta(\leq r)$ があって

$$\left| f_{xy}(x,y) - f_{xy}(a,b) \right| < \epsilon, \quad (|\,x-a\,| < \delta,\ |\,y-b\,| < \delta) \tag{15}$$

が成り立っている。(14) と (15) から

$$\left|\frac{\Delta(x,y)}{(x-a)(y-b)} - f_{xy}(a,b)\right| < \epsilon \quad (0 < |x-a| < \delta,\ 0 < |y-b| < \delta) \tag{16}$$

を得る。
ここで $0 < |x-a| < \delta$ なる x の各々の値に対し、正数 $\delta'(x)$ $(< \delta)$ を

$$\left|\underbrace{\psi(x,y) - \psi(x,b)}_{=\Delta(x,y)} - \underbrace{\psi_y(x,b)}_{f_y(x,b)-f_y(a,b)}(y-b)\right| \leq \epsilon|x-a||y-b| \quad (|y-b| \leq \delta'(x)) \tag{17}$$

を満たすように選ぶ。(17) の両辺を $|x-a||y-b|$ で割ると

$$\left|\frac{\Delta(x,y)}{(x-a)(y-b)} - \frac{f_y(x,b) - f_y(a,b)}{x-a}\right| \leq \epsilon \quad (0 < |y-b| \leq \delta'(x)). \tag{18}$$

(16) と (18) から

$$\left|\frac{f_y(x,b) - f_y(a,b)}{x-a} - f_{xy}(a,b)\right| < 2\epsilon \quad (0 < |x-a| < \delta) \tag{19}$$

を得る。すなわち (19) は $f_{yx}(a,b)$ が存在してそれが $f_{xy}(a,b)$ に等しいことを意味し、定理が証明される。 □

系 2.3. U は $\mathbb{R}^2$ の開集合、$f : U \to \mathbb{R}$ とする。U 全体で f_x, f_y, f_{xy} が存在して、このうち f_{xy} が連続ならば、f_{yx} も存在し $f_{xy} = f_{yx}$ が成り立つ。

系 2.4. U は $\mathbb{R}^n$ の開集合、$f : U \to \mathbb{R}$ とする。任意の $i = 1 \ldots n$ に対して $\partial f/\partial x^i$ が存在し、かつ任意の $i, j = 1 \ldots n, i < j$ に対して $\partial^2 f/\partial x^i \partial x^j$ が存在して連続ならば $\partial^2 f/\partial x^j \partial x^i$ も存在し、

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$$

が成り立つ。

Schwartz の定理によれば、与えられた関数 f に対し、$\partial^2 f/\partial x^i \partial x^j$ $(i < j)$ の存在とそれらの連続性を確認すれば、$\partial^2 f/\partial x^j \partial x^i$ の存在は確認するに及ばない。

定理 2.5 (Young). $f : U \to \mathbb{R}$ とする。U 全体で f_x, f_y が存在して、f_x, f_y がともに点 (a,b) において微分可能ならば $f_{xy}(a,b) = f_{yx}(a,b)$ となる。

証明 $|t|<r$ とし、(13) の x, y にそれぞれ $a+t, b+t$ を代入すると

$$\Delta(a+t, b+t) = (\ f_x(a+\theta t, b+t) - f_x(a+\theta t, b)\ )\,t$$

を得る。ここで θ は $0<\theta<1$ なる数。この式の両辺を t^2 で割ると

$$\frac{\Delta(a+t, b+t)}{t^2} = \frac{(\ f_x(a+\theta t, b+t) - f_x(a+\theta t, b)\ )}{t} \quad (0<|t|<r) \tag{20}$$

となる。

さて f_x が (a,b) で微分可能だから、任意の正数 ϵ に対して正数 δ $(\delta \le r)$ があって、

$$\left| f_x(a+\theta t, b+t) - f_x(a,b) - \left(\ f_{xx}(a,b)\theta t + f_{xy}(a,b)t\right)\right| \le \epsilon |t| \quad (|t| \le \delta) \tag{21}$$

$$|f_x(a+\theta t, b) - f_x(a,b) - f_{xx}(a,b)\theta t| \le \epsilon |t| \quad (|t| \le \delta) \tag{22}$$

が成り立っている。(21) と (22) から

$$\left| f_x(a+\theta t, b+t) - f_x(a+\theta t, b) - f_{xy}(a,b)t\right| \le 2\epsilon |t| \quad (|t| \le \delta)$$

を得る。両辺を $|t|$ で割って

$$\left| \frac{f_x(a+\theta t, b+t) - f_x(a+\theta t, b)}{t} - f_{xy}(a,b)\right| \le 2\epsilon \quad (0<|t| \le \delta). \tag{23}$$

ここで (20) を使うと (23) は

$$\left| \frac{\Delta(a+t, b+t)}{t^2} - f_{xy}(a,b)\right| \le 2\epsilon \quad (0<|t|<\delta)$$

となる。つまりこれは

$$\lim_{t\to 0} \frac{\Delta(a+t, b+t)}{t^2} = f_{xy}(a,b). \tag{24}$$

を意味している。

今までの議論を繰り返せば同様に

$$\lim_{t\to 0} \frac{\Delta(a+t, b+t)}{t^2} = f_{yx}(a,b). \tag{25}$$

も成立するはずである。(24) と (25) から

$$f_{xy}(a,b) = f_{yx}(a,b)$$

を得て、定理が証明される。 □

本書において Young の定理は、次に述べる Fréchet 微分の高階微分が対称的であることを証明するのに用いられる。

3　Fréchet 微分

3.1　概念

■Fréchet 微分　初歩の多変数微分学では関数 f の微分を論ずる場合、定義域は n-次元数空間 $\mathbb{R}^n$ (あるいは $\mathbb{R}^n$ の開集合) であり、値域は $\mathbb{R}$ であった。$\mathbb{R}^n$ と $\mathbb{R}$ をそれぞれノルム空間 E, E' に一般化することで Fréchet 微分の概念が得られる。

Fréchet 微分を考える利点として次の三つを挙げることができると思う：

1. 関数 f の全微分 Df の全微分 $D^2 f = D(Df)$、それのまた全微分 $D^3 f = D(D^2 f)$, 等々を首尾一貫した形で定義することができる。
2. 抽象化することで高階微分を含む公式が単純になり、見やすくなる。
3. 議論のいくつかは無限次元ノルム空間にも通用する。

ただし無限次元ノルム空間の扱いは本書のレベルを超えているから、以下ではノルム空間は有限次元であると仮定している。

本当はノルム空間ではなく、ノルム空間が加法群として自由かつ推移的に作用する "アフィン空間" に f の定義域と値域を拡張することができるのだが、本書では簡単のためノルム空間とする。ノルム空間を作用群とするアフィン空間に理論を拡張する作業は形式的なものである。難しくはないからこの点に拘る読者は自ら試みられたい。

■前提事項　以下では $E, E', \ldots$ は実ノルム空間を表し、これらは (特に断りのないかぎり) 有限次元であると仮定する。

また $U \subset E$, $U(\subset E)$ 等という表記があれば、これは (特に断りのない限り)「U は E の "開集合" である」あるいは「E の "開集合" U」を意味するものとする。

$f : U(\subset E) \to E'$ は通常通り f が U で定義され E' に値をとる関数であることを意味するものとする。f が連続であるとは仮定しない。

■微分　$f : U(\subset E) \to E'$ とする。線形写像 $(Df)_x \in (E, E')$ は次の性質を満たすとき、点 $x \in U$ での f の**微分** **(Fréchet 微分)** と呼ばれる：任意の正数 ϵ に対して正数 δ が存在して

$$|f(x+\xi) - f(x) - (Df)_x \xi| \le \epsilon |\xi| \quad (x+\xi \in U,\ |\xi| \le \delta)$$

が成り立つ。

定義から容易に分かるように、微分 $(Df)_x$ は存在すれば一意に定まる。微分 $(Df)_x \in \mathcal{L}(E, E')$ が存在することを、f は x で**微分可能**であるともいう。

U の各点 x で $(Df)_x$ が存在するとき関数 $Df : U \to \mathcal{L}(E, E')$ を

$$(Df) : x \mapsto (Df)_x$$

で定義し、これをたんに f の微分 (Fréchet 微分) と呼ぶ。Df が存在することを f は微分可能であるともいう。

注意. $E = \mathbb{R}^n, E' = \mathbb{R}$ の場合、f の微分とは微分学の復習で述べた関数の「全微分」に他ならない。

定理 3.1. $f : U(\subset E) \to E'$ とする。f は微分可能ならば連続である。

証明 微分可能の仮定より、$\epsilon = 1$ に対して正数 δ があって、

$$|f(x+\xi) - f(x) - (Df)_x\xi| \le |\xi| \quad (x+\xi \in U, |\xi| \le \delta)$$

すなわち

$$|f(x+\xi) - f(x)| \le (1 + |(Df)_x|)|\xi| \quad (x+\xi \in U, |\xi| \le \delta)$$

ここから f の連続性は明らかであろう。 □

■方向微分 $f : U(\subset E) \to E', x \in U, \xi \in E$ とする。

$(D_\xi f)(x) \in E'$ は以下の性質を満たすときに f の x での **ξ-方向の微分**と呼ばれる：

$$\lim_{t \to 0} \frac{1}{t} (\, f(x+t\xi) - f(x)\,) = (D_\xi f)(x),$$

すなわち任意の正数 ϵ に対し正数 δ があって

$$\Big| f(x+t\xi) - f(x) - t(D_\xi f)(x) \Big| \le \epsilon |t| \quad (x+t\xi \in U,\ |t| \le \delta)$$

が成り立つ。ξ-方向の微分は定義から明らかに一意的である。

f が U の各点で ξ-方向の微分をもつ場合に関数 $D_\xi f : U \to E'$ を

$$(D_\xi f) : x \mapsto (D_\xi f)(x)$$

で定義し、これをたんに f の **ξ-方向の微分**と呼ぶ。$D_\xi f$ が存在する場合に、f は **ξ-方向に微分可能**ともいう。

定理 3.2. f は微分可能ならば任意の ξ に対して ξ-方向に微分可能である。そしてこのとき

$$(Df)_x\xi = (D_\xi f)(x) \quad (x \in U)$$

が成り立つ。

証明 仮定から任意の正数 ϵ に対して次の式を満たすような正数 δ がある：

$$\left| f(x+\xi') - f(x) - (\,Df\,)_x\,\xi' \right| \le \epsilon \left|\xi'\right| \quad \left(\left|\xi'\right| \le \delta \right)$$

正数 δ' を $\delta' = \delta/|\xi|$ なるように取ると、$|t| \le \delta'$ なる t に対して $t\xi$ は上式の ξ' に代入することができて

$$|f(x+t\xi) - f(x) - t(Df)_x\xi| \le \epsilon'|t| \quad (|t| \le \delta')$$

を得る。ここで $\epsilon' = \epsilon|\xi|$。すなわち $(D_\xi f)(x)$ が存在し $(Df)_x\xi$ に等しい。 □

■**高階微分**　$f : U(\subset E) \to E'$ とし、f は微分可能とする。$Df : U \to (\,E\,,\,E'\,)$ がさらに微分可能のとき、すなわち

$$D^2f(= DDf) : U \to (\,E\,,(\,E\,,\,E'\,)\,) \cong (\,E,\,E\,;\,E'\,)$$

が存在するとき、f は 2 **回微分可能**であるといい、$D^2f : U \to (\,E,\,E\,;\,E'\,)$ を f の 2 **階の微分**と呼ぶ。

詳しく言うと、作用素 $(D^2f)_x \in (\,E,\,E\,;\,E'\,) \cong (\,E\,,(\,E\,,\,E'\,)\,)$ は任意の正数 ϵ に対して正数 δ が存在して

$$\left|(Df)_{x+\xi} - (Df)_x - (D^2f)_x\xi\right| \le \epsilon|\xi| \quad (x+\xi \in U,\ |\xi| \le \delta)$$

が成り立つとき、F の $x \in U$ での 2 階の微分と呼ばれる。U 上のすべての点 x で f の 2 階微分 $(D^2f)_x$ が存在するとき、関数 $D^2f : U \to (\,E,\,E\,;\,E'\,)$ が $D^2f : x \mapsto (D^2f)_x$ により定義される。

任意の正整数 r に対して f が r 回微分可能であることと f の r 階の微分 D^rf : $U \to (\underbrace{E\ldots E}_{r};\,E'\,)$ も帰納的に定義される。すなわち f が $r-1$ 回まで微分可能で関数 $D^{r-1}f : U \to (\underbrace{E\ldots E}_{r-1};\,E'\,)$ が定義されているとき、$D^{r-1}f$ がさらに微分可能ならば f は r **回微分可能**であるといい、関数

$$D^rf\,(= DD^{r-1}f) : U \to (\,E\,,(\underbrace{E\ldots E}_{r-1};\,E'\,)\,) \cong (\underbrace{E\ldots E}_{r};\,E'\,)$$

を f の r **階の微分**と呼ぶ。

2 階以上の微分は**高階の微分**と呼ばれる。

次に連続微分可能性を定義しよう。f が r 回可微分でかつ $D^r f : U \to (\underbrace{E \dots E}_{r}; E')$ が連続関数のとき、f は r **回連続的に微分可能**であるとか C^r **級**であると呼ばれる。

次の命題は直観的に明らかで証明は不要と思われる。

命題 3.3. r を正整数とする。f が r 回 [連続] 微分可能ならば、Df は $r-1$ 回 [連続] 微分可能で、

$$D^{r-1}(Df) = D^r f$$

が成り立つ。逆に Df が存在して、それが $r-1$ 回 [連続] 微分可能ならば f は r 回 [連続] 微分可能である。

■値の線形変換 $\psi : E' \to E''$ は線形変換と仮定する。関数 $f : U(\subset E) \to E'$ に対し、関数 $\psi \circ f : U \to E''$ を本書では f の**値の** ψ **による変換**と呼ぶことにする。

次の “定理” は自明のことと思われるが、一応証明を付けておく。

定理 3.4. $\psi : E' \to E''$ は線形変換と仮定する。f が [連続] 微分可能ならば、値の変換 $\psi \circ f$ は [連続] 微分可能で

$$(\, D(\psi \circ f)\,)_x = \psi \circ (\, Df\,)_x \quad (x \in U) \tag{26}$$

が成り立つ。言い換えると

$$D(\psi \circ f) = \psi_* \circ (Df) \tag{27}$$

が成り立つ。ここで ψ_* は (6) で定義される線形変換 $(\, E\, ,\, E'\,) \to (\, E\, ,\, E''\,)$ である。

証明 仮定から

$$|\, f(x+\xi) - f(x) - (Df)_x \xi\, | \le \epsilon\, |\xi| \quad (|\xi| \le \delta),$$

ゆえに

$$|\, (\psi \circ f)(x+\xi) - (\psi \circ f)(x) - (\psi \circ (Df)_x)\xi\, | \le \epsilon\, |\psi|\, |\xi| \quad (|\xi| \le \delta).$$

つまり $\psi \circ f$ は微分可能で (26) が成り立つ。(26) は

$$(D(\psi \circ f))_x = \psi_*((Df)_x) = (\psi_* \circ (Df))_x \quad (x \in U).$$

のように書き換えることができるから、(27) が成り立つ。
連続微分可能性は 式 (27) から明らかである。 □

定理 3.5. $\psi : E' \to E''$ は線形変換と仮定する。f が r 回 [連続] 微分可能ならば、値の変換 $\psi \circ f$ も r 回 [連続] 微分可能で

$$(D^r(\psi \circ f))_x = \psi \circ (D^r f)_x \quad (x \in U), \tag{28}$$

$$D^r(\psi \circ f) = \psi_{*^r} \circ (D^r f) \tag{29}$$

が成り立つ。ここで ψ_{*^r} は (7) で定義される線形変換である。

証明 r についての帰納法による。
$r = 1$ ： これは定理 3.4 である。
$r > 1$ ： 帰納法の仮定により

$$D^{r-1}(\psi \circ f) = \psi_{*^{r-1}} \circ (D^{r-1} f)$$

$D^{r-1} f$ は可微分だから、定理 3.4 で既に示したようにその値の線形変換である $D^{r-1}(\psi \circ f)$ は可微分である。これに D を作用させると

$$\begin{aligned} DD^{r-1}(\psi \circ f) &= D(\psi_{*^{r-1}} \circ (D^{r-1} f)) \\ &= \psi_{*^r} \circ (DD^{r-1} f) \end{aligned}$$

すなわち (29) を得る。(29) を書き換えると

$$(D^r(\psi \circ f))_x = \psi_{*^r} \left((D^r f)_x \right) = \psi \circ (D^r f)_x \quad (x \in U),$$

すなわち (28) を得る。ここで命題 1.10 を用いた。 □

■高階微分と方向微分　ここではまず次の定理を証明する。

定理 3.6. $f : U(\subset E) \to E'$ とし、r を正整数、$\xi \in E$ とする。f が r 回 [連続] 微分可能ならば $D_\xi f$ は $r-1$ 回 [連続] 微分可能で、

$$\left(D^{r-1} \left(D_\xi f \right) \right)_x (\eta_1 \ldots \eta_{r-1}) = (D^r f)_x (\eta_1 \ldots \eta_{r-1}, \xi) \tag{30}$$

が成り立つ。ここで $x \in U, \eta_i \in E$.

証明 線形変換 $\psi_\xi : (E , E') \to E'$ を

$$\psi_\xi(A) = A(\xi)$$

で定義すると、

$$\Big(\psi_\xi \circ (Df) \Big)(x) = \psi_\xi \left((Df)_x \right) = (Df)_x(\xi) \\ = \Big(D_\xi f \Big)(x),$$

$$\psi_\xi \circ (Df) = D_\xi f$$

が成り立つ。すなわち $D_\xi f$ は Df の値の線形変換である。Df は $r-1$ 回 [連続] 微分可能ゆえ、定理 3.4 により $D_\xi f$ もまた $r-1$ 回 [連続] 微分可能で、

$$\Big(D^{r-1} \Big(D_\xi f \Big) \Big)_x = \psi_\xi \circ \Big(D^{r-1} Df \Big)_x$$

すなわち

$$\Big(D^{r-1} \Big(D_\xi f \Big) \Big)_x (\eta_1 \ldots \eta_{r-1}) = \psi_\xi \Big(\Big(D^{r-1} Df \Big)_x (\eta_1 \ldots \eta_{r-1}) \Big) \\ = \Big(D^{r-1} Df \Big)_x (\eta_1 \ldots \eta_{r-1})(\xi) = \Big(D^{r-1} Df \Big)_x (\eta_1 \ldots \eta_{r-1}, \xi).$$

□

系 3.7. $f : U(\subset E) \to E'$ とし、r を正整数、$\xi_1 \ldots \xi_r \in E$ とする。f が r 回 [連続] 微分可能ならば、$D_{\xi_2} \ldots D_{\xi_r} f$ が存在し [連続] 微分可能で、

$$\Big(D_{\xi_1} D_{\xi_2} \ldots D_{\xi_r} f \Big)(x) = (\ D^r f\ )_x (\xi_1 \ldots \xi_r) \quad (x \in U) \tag{31}$$

証明 $r = 1$ ： 既に 3.2 で示した。
$r > 1$ ： $D_{\xi_r} f$ は $r-1$ 回 [連続] 微分可能で

$$\Big(D^{r-1} \Big(D_{\xi_r} f \Big) \Big)_x (\xi_1 \ldots \xi_{r-1}) = (\ D^r f\ )_x (\xi_1 \ldots \xi_{r-1}, \xi_r).$$

一方、帰納法の仮定から

$$\Big(D_{\xi_1} \ldots D_{\xi_{r-1}} (D_{\xi_r} f) \Big)(x) = \Big(D^{r-1} (D_{\xi_r} f) \Big)_x (\xi_1 \ldots \xi_{r-1}).$$

これらから (31) を得る。 □

一般に Fréchet 高階微分を含んだ公式はそのままでは抽象的で難解であるが、上のように方向微分で表してみるとすっきりすることが多い。

■変数の線形変換 $f : U(\subset E) \to E'$ とし、$\varphi : E_0 \to E$ は線形変換とする。このとき関数 $f \circ \varphi : V \to E'$ を本書では関数 f の φ **による変数変換**と呼ぶことにする。ここで $V = \varphi^{-1}(U) \subset E_0$ である。値の変換と同様、変数変換についても次の定理が成り立つ。

命題 3.8. f が可微分ならば $f \circ \varphi$ は可微分で、

$$(D(f \circ \varphi))_y = (Df)_x \circ \varphi \quad (y \in V, x = \varphi(y)) \tag{32}$$

が成り立つ。

証明 仮定から

$$\Big| f(x+\xi) - f(x) - (Df)_x (\xi) \Big| \leq \epsilon |\xi| \quad (x+\xi \in U, |\xi| \leq \delta).$$

φ は連続だから正数 δ' が存在し $\eta \in E_0, |\eta| \leq \delta'$ ならば $|\varphi(\eta)| \leq \delta$ となる ($\delta' = \delta/|\varphi|$ とおけば十分である) 。ゆえに $\eta \in E_0, y+\eta \in V, |\eta| \leq \delta'$ なる η に対しては

$$| f(x+\varphi(\eta)) - f(x) - (Df)_x(\varphi(\eta)) | \leq \epsilon |\varphi(\eta)| \leq \epsilon |\varphi||\eta|$$

つまり

$$|(f \circ \varphi)(y+\eta) - (f \circ \varphi)(y) - ((Df)_x \circ \varphi)(\eta)| \leq \epsilon' |\eta| \quad (|\eta| \leq \delta')$$

である。ここで $\epsilon' = \epsilon|\varphi|$ とおいた。よって $f \circ \varphi$ は y で可微分で (32) が成り立つ。 □

線形変換 $\varphi : E_0 \to E$ に対し $\varphi^* : (E , E') \to (E_0 , E')$ を

$$\varphi^*(A) = A \circ \varphi$$

で定義すると、φ^* はまた線形変換である。これを使うと式 (32) は

$$\begin{aligned} (D(f \circ \varphi))_y &= (Df)_x \circ \varphi = \varphi^* ((Df)_x) \\ &= (\varphi^* \circ (Df))_x \end{aligned}$$

と書き換えられる。すなわち f が可微分ならば $f \circ \varphi$ は可微分で

$$D(f \circ \varphi) = \varphi^* \circ (Df) \circ \varphi \tag{33}$$

が成り立つ。

(32) を方向微分で表すと

$$\Big(D_\eta(f \circ \varphi) \Big)(y) = \Big(D_\xi f \Big)(x)$$

が得られる。ここで $\eta \in E_0, \xi = \varphi(\eta)$. これがすべての $y \in V$ で成立だから結局

$$D_\eta(f \circ \varphi) = \Big(D_\xi f \Big) \circ \varphi \tag{34}$$

命題 3.9. f が r 回 [連続] 微分可能ならば $f\circ\varphi$ は r 回 [連続] 微分可能で

$$(\, D^r(f\circ\varphi)\,)_y\,(\,\eta_1\ldots\eta_r\,) = (D^r f)_x\,(\,\xi_1\ldots\xi_r\,) \quad (y\in V,\ x=\varphi(x)), \tag{35}$$

つまり

$$D_{\eta_1}\ldots D_{\eta_r}(f\circ\varphi) = \left(\, D_{\xi_1}\ldots D_{\xi_r} f\,\right)\circ\varphi \quad (\eta_i\in E_0,\ \xi_i=\varphi(\eta_i)) \tag{36}$$

が成り立つ。ここで $\eta_i\in E_0, \xi_i=\varphi(\eta_i)\in E\ (i=1\ldots r)$ である。

証明 r についての帰納法による。
$r=1$ ： これはすでに 3.8 で証明した。
$r\geq 2$ ： $r-1$ までは成立と仮定。f が r 回可微分ならば Df は $r-1$ 回可微分で、帰納法の仮定より $(Df)\circ\varphi$ は $r-1$ 回可微分である。(33) を使い $D(f\circ\varphi)=\varphi^*\circ(Df)\circ\varphi$ は $r-1$ 回可微分と分かる (値の線形変換だから)。すなわち $f\circ\varphi$ は r 回可微分である。
さらに帰納法の仮定から

$$D_{\eta_2}\ldots D_{\eta_r}(f\circ\varphi) = \left(\, D_{\xi_2}\ldots D_{\xi_r} f\,\right)\circ\varphi$$

が成立している。 3.7 によりこれは可微分で

$$D_{\eta_1}D_{\eta_2}\ldots D_{\eta_r}(f\circ\varphi) = D_{\eta_1}\left(\left(\, D_{\xi_2}\ldots D_{\xi_r} f\,\right)\circ\varphi\right) = \left(\, D_{\xi_1}D_{\xi_2}\ldots D_{\xi_r} f\,\right)\circ\varphi$$

を得る。ここで (34) を用いた。よって (36) も証明された。 □

■関数の値の表示と成分関数 $f: U(\subset E)\to E'$ とする。線形同型 $\psi: E'\to\mathbb{R}^m$ による f の値の変換 $\psi\circ f: U\to\mathbb{R}^m$ を本書では ψ **による** f **の値の表示**と呼ぶ。
定理 3.3 から次のことは明らかである：

命題 3.10. f は f の値の表示 $\psi\circ f$ が r 回 [連続] 微分可能のとき、かつそのときに限って r 回 [連続] 微分可能で

$$(D^r(\psi\circ f))_x = \psi\circ(D^r f)_x \quad (x\in U)$$

が成り立つ。

各 $i=1\ldots m$ に対し $\pi_i:\mathbb{R}^m\to\mathbb{R}$ を

$$\pi_i: (x^1\ldots x^m)\mapsto x^i$$

で定義すると、π_i は明らかに線形である。

$F : U(\subset E) \to \mathbb{R}^m$ に対し

$$F_i = \pi_i \circ F \quad (i = 1 \dots m)$$

で定義される関数 $F_i : U \to \mathbb{R}$ $(i = 1 \dots m)$ を F の**成分関数**と呼ぶと、次の定理が成り立つ。

定理 3.11. r を正整数とする。F が r 回 [連続] 微分可能ならば F のすべての成分関数 $F_1 \dots F_m$ は r 回 [連続] 微分可能で

$$(\, D^r(\pi_i \circ F)\,)_x = \pi_i \circ (\, D^r F\,)_x \quad (x \in U) \tag{37}$$

が成り立つ。逆にすべての成分関数が r 回 [連続] 微分可能ならば F は r 回 [連続] 微分可能である。

証明 前半は値の変換についての一般論から明らか。以下、その逆を示す。証明は r についての帰納法を使う。
$r = 1$：$x \in U$ を固定する。仮定から、任意の正数 ϵ に対して、ある正数 δ があって、すべての $i = 1 \dots m$ について

$$\left| \pi_i \circ F(x+\xi) - \pi_i \circ F(x) - (\, D(\pi_i \circ F)\,)_x\, \xi \right| \le \epsilon |\xi| \quad (|\xi| \le \delta)$$

が成り立つ。このとき $A_x \in (E, \mathbb{R}^m)$ で

$$\pi_i \circ A_x = (\, D(\pi_i \circ F)\,)_x \quad (i = 1 \dots m)$$

を満たすものが唯一つある。すなわち

$$|\pi_i \circ F(x+\xi) - \pi_i \circ F(x) - \pi_i \circ A_x \xi| \le \epsilon |\xi| \quad (|\xi| \le \delta)$$

で、$\mathbb{R}^m$ のノルムの定義からただちに

$$|F(x+\xi) - F(x) - A_x \xi| \le \epsilon |\xi| \quad (|\xi| \le \delta)$$

が導かれる。よって F は可微分である。
$r > 1$： $r-1$ までは定理の主張が正しいと仮定する。すると $F_1 \dots F_m$ が r 回微分可能ならば、

$$\left| \Big(\, D^{r-1}(\pi_i \circ F)\,\Big)_{x+\xi} - \Big(\, D^{r-1}(\pi_i \circ F)\,\Big)_x - (\, D^r(\pi_i \circ F)\,)_x\, \xi \right| \le \epsilon |\xi| \quad (|\xi| \le \delta, i = 1 \dots m)$$

あるいは

$$\left| \pi_i \circ (DF)_{x+\xi} - \pi_i \circ (DF)_x - (\, D^r(\pi_i \circ F)\,)_x\, \xi \right| \le \epsilon |\xi| \quad (|\xi| \le \delta, i = 1 \dots m)$$

ところが $A_x \in (E_1, \ldots, E_r\,;\,E')$ $(E_1 = \cdots = E_r = E)$ で

$$\pi_i \circ A_x = (\, D^r(\pi_i \circ F)\,)_x \quad (i = 1 \ldots m) \tag{38}$$

を満たすものが唯一つある。ゆえに前と同様、

$$\left| (D^{r-1}F)_{x+\xi} - (D^{r-1}F)_x - A_x\xi \right| \leq \epsilon|\xi| \quad (|\xi| \leq \delta)$$

が導かれ、$D^{r-1}F$ の可微分性が証明される。すなわち D は r 回微分可能である。
式は

$$(\, D^r(\pi_i \circ F)\,)_x = \pi_{i*}(\,(\, D^rF\,)_x\,) \quad (x \in U)$$

と書き換えられる。すなわち

$$D^r(\pi_i \circ F) = \pi_{i*} \circ (D^rF) \quad (i = 1 \ldots m)$$

ここから D^rF が連続ならば $D^r(\pi_i \circ F)$ がすべて連続なること、ならびにその逆は明らかであろう。 □

系 3.12. 各 $i = 1 \ldots m$ について $\psi_i : E' \to \mathbb{R}$ は線形で $\psi(\eta) = (\psi_1(\eta) \ldots \psi_m(\eta))$ で定義される関数 $\psi : E' \to \mathbb{R}^m$ は線形同型であると仮定する。
$f : U \to E'$ とする。このとき f が r 回微分可能ならば $\psi_i \circ f$ は r 回微分可能で

$$(\, D^r(\psi_i \circ f)\,)_x = \psi_i \circ (\, D^rf\,)_x \quad (x \in U) \tag{39}$$

が成り立つ。逆にすべての $\psi_i \circ f$ が r 回微分可能ならば f は r 回微分可能である。

■関数の変数の表示 $f : U(\subset E) \to E'$ とする。線形同型 $\varphi : \mathbb{R}^n \to E$ による f の変数変換 $f \circ \varphi : V \to E'$ を本書では φ **による** f **の変数表示**と呼ぶ。ここで $V = \varphi^{-1}(U)$.
定理 3.8 から、次は明らかであろう：

命題 3.13. f は $f \circ \varphi$ が r 回 [連続] 微分可能のとき、かつそのときに限って r 回 [連続] 微分可能である。

$f : U(\subset \mathbb{R}^n) \to \mathbb{R}$ とし、$\{\, \mathsf{e}_1 \ldots \mathsf{e}_n \,\}$ を $\mathbb{R}^n$ の標準基底とすると、

$$D_{\mathsf{e}_j}f = \frac{\partial f}{\partial x^j} \quad (j = 1 \ldots n)$$

が成り立つ。これは定義から明らかであろう。以下

$$D_j = D_{\mathsf{e}_j} \quad (j = 1 \ldots n)$$

の略号を用いる。

■高階微分の対称性

補題 3.14. $f : U(\subset \mathbb{R}^n) \to \mathbb{R}$ とする。f が r 回微分可能ならば

$$D_{j_1} D_{j_2} D_{j_3} \ldots D_{j_r} f = D_{j_2} D_{j_1} D_{j_3} \ldots D_{j_r} f \tag{40}$$

証明 $j_3 \ldots j_r$ を固定し、$g = D_{j_3} \ldots D_{j_r} f$ とおくと、仮定からすべての $j_2 = 1 \ldots n$ に対して $D_{j_2} g$ が存在し可微分である。よって定理 2.5 [Young] により $D_{j_1} D_{j_2} g = D_{j_2} D_{j_1} g$ □

補題 3.15. $f : U(\subset \mathbb{R}^n) \to \mathbb{R}$ とする。f が r 回微分可能ならば任意の $j_1 \ldots j_r \in \{ 1 \ldots n \}$ に対して

$$D_{j_1} \ldots D_{j_r} f = D_{j_{\sigma(1)}} \ldots D_{j_{\sigma(r)}} f \tag{41}$$

が成り立つ。すなわち

$$(D^r f)_x \Big(\mathsf{e}_{j_1} \ldots \mathsf{e}_{j_r} \Big) = (D^r f)_x \Big(\mathsf{e}_{j_{\sigma(1)}} \ldots \mathsf{e}_{j_{\sigma(r)}} \Big) \quad (x \in U) \tag{42}$$

が成り立つ。ここで σ は 集合 $\{ 1 \ldots r \}$ の任意の置換である。

証明 r に関する帰納法による。$r = 1$ ： 証明不要。
$r > 1$ ： $r - 1$ までは命題の主張が成り立っていると仮定する。すると j_1 を固定する $j_1 \ldots j_r$ の任意の置換の下で式

$$D_{j_1} \ldots D_{j_r} f \tag{43}$$

は不変である。一方、$2 \leq s \leq r$ とすると前の補題と帰納法の仮定から

$$\begin{aligned} & D_{j_1} D_{j_2} \ldots D_{j_s} \ldots D_{j_r} f = D_{j_1} D_{j_s} \ldots D_{j_2} \ldots D_{j_r} f \\ & = D_{j_s} D_{j_1} \ldots D_{j_2} \ldots D_{j_r} f = D_{j_s} D_{j_2} \ldots D_{j_1} \ldots D_{j_r} f. \end{aligned}$$

つまり j_1 と j_s の互換の下でも (43) は不変である。結局、$j_1 \ldots j_r$ のすべての互換の下で (43) は不変で、$j_1 \ldots j_r$ の任意の置換は互換の積だから (41) が成り立つ。 □

定理 3.16. $f : U \to E'$ とし、f は r 回微分可能とする。
　このとき $(D^r f)_x \in (\underbrace{E \ldots E}_{r}; E')$ は対称的である。つまり σ を 集合 $\{ 1 \ldots r \}$ の任意の置換として

$$(D^r f)_x (\xi_1 \ldots \xi_r) = (D^r f)_x (\xi_{\sigma(1)} \ldots \xi_{\sigma(r)}) \tag{44}$$

が成り立つ。

証明 はじめに $E' = \mathbb{R}$ と仮定する。線形変換 $\varphi : \mathbb{R}^r \to E$ を

$$\varphi(\mathsf{e}_j) = \xi_j \quad (j = 1 \dots r)$$

で定義すると、仮定から $f \circ \varphi : V \to E'$ は r 回微分可能。ここで $V = \varphi^{-1}(U)$ とおいた。$y \in V, x = \varphi(y)$ とすると

$$(D^r(f \circ \varphi))_y(\mathsf{e}_{j_1} \dots \mathsf{e}_{j_r}) = (D^r f)_x(\xi_{j_1} \dots \xi_{j_r}).$$

ここで左辺が $j_1 \dots j_r$ の任意の置換の下で不変だから右辺も不変。したがって (44) が成り立つ。

E' が一般の有限次元ベクトル空間のときは同型 $E' \cong \mathbb{R}^m$ を考えて、f の値の表示 $F = (F_1 \dots F_m)$ を考えてみれば分かる。各成分関数 F_i について $(D^r F_i)_x$ が対称的だから $(D^r F)_x$ は対称的である。したがって $(D^r f)_x$ は対称的である。 □

■高階微分と方向微分 (つづき)

補題 3.17. $(e_1, \dots e_n)$ を E の基底とする。すると $A \in (\, E \,, \mathbb{R}\,)$ に対し

$$\psi(A) = (\psi_1(A) \dots \psi_n(A))$$
$$\psi_j(A) = A(e_j) \quad (j = 1 \dots n)$$

で定義される写像 $\psi : (\, E \,, \mathbb{R}\,) \to \mathbb{R}^n$ は線形同型である。

　もっと一般に $A \in (\,\underbrace{E \dots E}_{r}\,; \mathbb{R}\,)$ に対し

$$\psi(A) = \Big(\, \psi_{j_1 \dots j_r}(A) \,\Big)_{j_1 \dots j_r}$$
$$\psi_{j_1 \dots j_r}(A) = A(e_{j_1} \dots e_{j_r}) \quad (j_1 \dots j_r \in \{\, 1 \dots n \,\})$$

で定義される写像 $\psi : (\,\underbrace{E \dots E}_{r}\,; \mathbb{R}\,) \to \mathbb{R}^{n^r}$ は線形同型である。

証明 証明はやさしいので読者に委ねる。 □

定理 3.18. $f : U(\subset \mathbb{R}^n) \to \mathbb{R}$、$r \geq 2$ とする。すべての $j_2 \dots j_r \in \{\, 1 \dots n \,\}$ について $D_{j_2} \dots D_{j_r} f$ が存在し、可微分ならば $D^r f$ が存在する (つまり D^{r-1} が存在し、可微分である)。

証明 r についての帰納法を使う。

$r=2$：$D_j f$ が存在して可微分ならば、とくに $D_j f$ は連続である $(j=1\dots n)$。このとき定理 2.1 により Df が存在する。ψ_j を補題 3.17 で定義された関数とすると、

$$\psi_j \circ (Df) = D_j f \quad (j=1\dots n)$$

となる。すると仮定から $\psi_j \circ (Df)$ はすべて可微分である。よって Df は可微分。
$r>2$：$j_3 \dots j_r$ を一時固定しておくと、仮定から $D_{j_2}(D_{j_3\dots j_r})f$ はすべての $j_2=1\dots n$ について連続であるから、$D_{j_3\dots j_r}f$ は可微分だとわかる。
帰納法の仮定より D^{r-1} が存在。$\psi_{j_2\dots j_r}$ を補題 3.17 で定義された関数とすると

$$\psi_{j_2\dots j_r} \circ (D^{r-1}f) = D_{j_2}\dots D_{j_r}f$$

すると仮定から $\psi_{j_2\dots j_r} \circ (D^{r-1}f)$ は可微分。よって $D^{r-1}f$ は可微分である。 □

系 **3.19.** $f : U(\subset \mathbb{R}^n) \to \mathbb{R}$、$r \geq 1$ とする。すべての $j_1 \dots j_r \in \{1\dots n\}$ について $D_{j_1}\dots D_{j_r}f$ が存在し連続ならば、$D^r f$ が存在し連続である (f は r 回連続微分可能である)。

証明 $D_{j_1}(D_{j_2}\dots D_{j_r}f)$ が存在し連続だから $D_{j_2}\dots D_{j_r}f$ は可微分である。すると定理 3.18 により $D^r f$ が存在。仮定から

$$\psi_{j_1\dots j_r} \circ (D^r f) = D_{j_1}\dots D_{j_r}f$$

は連続だから、$D^r f$ は連続である。 □

3.2　線形性と積の微分法

■**微分の線形性**　$f, g : U(\subset E) \to E'$ とする。f, g がともに r 回 [連続] 微分可能ならば、$f+g : U \to E'$ も r 回 [連続] 微分可能で、

$$D(f+g) = Df + Dg \tag{45}$$

が成り立つ。また f が r 回 [連続] 微分可能ならば、任意の定数 c に対して cf も r 回 [連続] 微分可能で

$$D(cf) = c(Df) \tag{46}$$

が成り立つ。これらを微分の**線形性**という。線形性の証明は容易であるので、読者の演習問題とする。

■**作用素の積の微分法**　$U \subset E$, $f : U \to (\,E_2\,,\,E_3\,)$, $g : U \to (\,E_1\,,\,E_2\,)$ とする。このとき f と g の積$f \cdot g : U \to (\,E_1\,,\,E_3\,)$ が

$$(f \cdot g)(x) = f(x) \cdot g(x) = f(x) \circ g(x) \quad (x \in U)$$

で定義される。$f \cdot g$ は以下では fg と略記される。

定理 3.20. $U \subset E$, $f : U \to (\,E_2\,,\,E_3\,)$, $g : U \to (\,E_1\,,\,E_2\,)$ とする。f, g がともに微分可能ならば、積 fg も微分可能で、

$$(D(fg))_x\xi = (\,(Df)_x\xi\,)\,g(x) + f(x)\,(\,(Dg)_x\xi\,) \quad (x \in U,\ \xi \in E) \tag{47}$$

が成り立つ。

証明 $x \in U$ とする。仮定より f, g は微分可能だから任意の正数 ϵ に対して正数 δ があって

$$|\,f(x+\xi) - f(x) - A_x\xi\,| \le \epsilon\,|\,\xi\,| \quad (|\,\xi\,| \le \delta) \tag{48}$$

$$|\,g(x+\xi) - g(x) - B_x\xi\,| \le \epsilon\,|\,\xi\,| \quad (|\,\xi\,| \le \delta) \tag{49}$$

$$|\,g(x+\xi) - g(x)\,| \le \epsilon \quad (|\,\xi\,| \le \delta) \tag{50}$$

ここで $A_x = (Df)_x \in (\,E\,,\,(\,E_2\,,\,E_3\,)\,)$, $B_x = (Dg)_x \in (\,E\,,\,(\,E_1\,,\,E_2\,)\,)$ とおいた。(50) からとくに

$$|\,g(x+\xi)\,| \le |\,g(x)\,| + 1 \quad (|\,\xi\,| \le \delta). \tag{51}$$

を得る。(48) 〜 (51) から

$$\begin{aligned}
&\ |\,f(x+\xi)g(x+\xi) - f(x)g(x) - ((A_x\xi)g(x) + f(x)(B_x\xi))\,| \\
\le&\ |\,f(x+\xi)g(x+\xi) - f(x)g(x+\xi) - (A_x\xi)g(x+\xi)\,| \\
&\quad + |\,(A_x\xi)g(x+\xi) - (A_x\xi)g(x)\,| + |\,f(x)g(x+\xi) - f(x)g(x) - f(x)(B_x\xi)\,| \\
\le&\ \epsilon\,|\,g(x+\xi)\,|\,|\,\xi\,| + \epsilon\,|\,A_x\,|\,|\,\xi\,| + \epsilon\,|\,f(x)\,|\,|\,\xi\,| \\
\le&\ \epsilon M\,|\,\xi\,| \quad (|\,\xi\,| \le \delta)
\end{aligned} \tag{52}$$

ここで $M = |\,A_x\,| + |\,f(x)\,| + |\,g(x)\,| + 1$ とおいた。
左辺の $\xi \in E$ の関数

$$P_x(\xi) = (A_x\xi)g(x) + f(x)(B_x\xi) \in (\,E_1\,,\,E_3\,)$$

は ξ に関して線形である。つまり $P_x \in (\,E\,,\,(\,E_1\,,\,E_3\,)\,)$ で、要するに (52) は

$$|\,f(x+\xi)g(x+\xi) - f(x)g(x) - P_x\xi\,| \le \epsilon M\,|\,\xi\,| \quad (|\,\xi\,| \le \delta).$$

と書ける。これは fg が x で可微分で、

$$(D(fg))_x = P_x$$

であることを意味し、(47) が証明される。 □

(47) を方向微分を用いて表すと

$$D_\xi(fg) = (D_\xi f)g + f(D_\xi g) \quad (\xi \in E). \tag{53}$$

(47) はまた次のように書き直すことができる：一般に $B \in (E_1, E_2)$ に対し $\psi_B : (E_2, E_3) \to (E_1, E_3)$ を

$$\psi_B(A) = A \cdot B$$

で定義すると、ψ_B は明らかに線形である。すなわち $\psi_B \in ((E_2, E_3), (E_1, E_3))$ で、ここから写像

$$\psi : B \in (E_1, E_2) \mapsto \psi_B \in ((E_2, E_3), (E_1, E_3))$$

が定義されるが、容易に分かるようにこの ψ も線形である。
同様に $A \in (E_2, E_3)$ に対し線形変換 $\chi_A : (E_1, E_2) \to (E_1, E_3)$ を

$$\chi_A(B) = A \cdot B$$

で定義し、写像

$$\chi : A \in (E_1, E_2) \mapsto \chi_A \in ((E_1, E_2), (E_1, E_3))$$

を定義すると、χ は線形である。
ψ と χ を使うと (47) の右辺は

$$\psi_{g(x)}(\,(Df)_x\xi\,) + \chi_{f(x)}(\,(Dg)_x\xi\,)$$

と書き換えることができる。すなわち (47) から

$$(D(fg))_x = \psi_{g(x)} \cdot (Df)_x + \chi_{f(x)} \cdot (Dg_x) \quad (x \in U),$$

あるいは

$$D(fg) = (\psi \circ g) \cdot (Df) + (\chi \circ f) \cdot (Dg) \tag{54}$$

が導かれる。

定理 3.21. $U \subset E$, $f : U \to (E_2, E_3)$, $g : U \to (E_1, E_2)$ とする。f, g がともに r 回 [連続] 微分可能ならば、積 fg も r 回 [連続] 微分可能である。

証明 r に関する帰納法による。
$r = 1$ ：これは既に証明した。
$r > 1$ ：f, g が r 回 [連続] 微分可能ならば、ψ, χ は線形変換だから $\psi \circ g, \chi \circ f$ は r 回 [連続] 微分可能である。また Df, Dg は $(r-1)$ 回 [連続] 微分可能である。すると帰納法の仮定により $(\psi \circ g) \cdot (Df), (\chi \circ f) \cdot (Dg)$ は $(r-1)$ 回 [連続] 微分可能。ゆえに (54) から $D(fg)$ は $r-1$ 回 [連続] 微分可能である。すなわち fg は r 回 [連続] 微分可能である。 □

$U \subset E$, $f : U \to (E_2, E_3)$, $g : U \to (E_1, E_2)$ とする。f, g は r 回微分可能だとして、積 fg の r 回の方向微分の式を求めてみよう：
$\xi_1 \dots \xi_r \in E$ とし、以下しばらくこれを固定する。
r 回微分可能な任意の関数 $h : U \to E'$ と、$\{ 1 \dots r \}$ の任意の部分集合 $I = \{ i_1 \dots i_s \}$ (ただし $i_1 < \dots < i_s$) に対し

$$h_I = D_{\xi_{i_1}} \dots D_{\xi_{i_s}} h$$

と略記する。I が空集合のときはもちろん $h_I = h$ である。

命題 3.22. このとき

$$(f \cdot g)_{\{ 1 \dots r \}} = \sum_I f_I \cdot g_{I'} \tag{55}$$

が成り立つ。ここで I は $\{ 1 \dots r \}$ の部分集合全体に渡るものとし、I' は I の補集合とする。

証明 r に関する帰納法。
$r = 1$ ：これは (53) だから既に証明されている。
$r > 1$ ：$r-1$ までは正しいと仮定。帰納法の仮定から

$$(f \cdot g)_{\{ 1 \dots r-1 \}} = \sum_{I_0} f_{I_0} \cdot g_{I'_0},$$

ここで $I_0 \subset \{ 1 \dots r-1 \}$, I'_0 は $\{ 1 \dots r-1 \}$ に対する I_0 の補集合である。

すると (53) を使って

$$
\begin{aligned}
(f \cdot g)_{\{1 \dots r\}} &= \sum_{I_0} \Big(f_{I_0 \cup \{r\}} \cdot g_{I_0'} + f_{I_0} \cdot g_{I_0' \cup \{r\}} \Big) \\
&= \sum_{(r \in) I} f_I \cdot g_{I'} + \sum_{(r \notin) I} f_I \cdot g_{I'} = \sum_I f_I \cdot g_{I'}
\end{aligned}
$$

を得る。 □

すなわち積 fg に関して我々は次の r 回方向微分の式を得る。

$$
D_{\xi_1} \dots D_{\xi_r}(f \cdot g) = \sum_{s=0}^{r} \sum_{1 \le i_1 < \cdots < i_s \le r} \Big(D_{\xi_{i_1}} \dots D_{\xi_{i_s}} f \Big) \cdot \Big(D_{\xi_{i'_1}} \dots D_{\xi_{i'_{r-s}}} g \Big), \tag{56}
$$

ここで $\left\{ i'_1 \dots i'_{r-s} \right\}$ は $\{ i_1 \dots i_s \}$ の補集合を表す。

系 3.23. 式 (56) でとくに $\xi_1 = \dots \xi_r = \xi$ の場合、さらに s に $r-s$ を代入して次の式 (Leibniz の公式) を得る：

$$
D_\xi{}^r (f \cdot g) = \sum_{s=0}^{r} \binom{r}{s} \Big(D_\xi{}^{r-s} f \Big) \cdot \Big(D_\xi{}^s g \Big) \tag{57}
$$

■作用素とベクトルの積の微分法　次に $U \subset E$, $f : U \to (E_1, E')$、$h : U \to E_1$ とする。これらに対しても積 $f(h) : U \to E'$ が $(f(h))(x) = f(x)(h(x))$ で定義される。

より一般に p を任意の正整数、$f : U \to (E_1, \dots, E_p ; E')$, $h_i : U \to E_i$ $(i = 1 \dots p)$ として f と $h_1 \dots h_p$ の積 $f(h_1 \dots h_p) : U \to E'$ が

$$
(f(h_1 \dots h_p))(x) = f(x)(h_1(x) \dots h_p(x)) \quad (x \in U)
$$

で定義される。

命題 3.24. f, h_i $(i = 1 \dots p)$ が微分可能ならば積 $f(h_1 \dots h_p)$ は微分可能で、

$$
D_\xi(f(h_1 \dots h_p)) = (D_\xi f)(h_1 \dots h_p) + \sum_{i=1}^{p} f(h_1 \dots D_\xi h_i \dots h_p) \tag{58}
$$

が成り立つ。

f, h_i $(i = 1 \dots p)$ が r 回 [連続] 微分可能ならば積 $f(h_1 \dots h_p)$ は r 回 [連続] 微分可能である。

証明 $f(h_1 \dots h_p) = (f(h_1))(h_2 \dots h_p)$ だから、$p = 1$ の場合に還元することができる。
さらに E_1 と $(\mathbb{R}, E_1)$ の線形同型を考えれば容易に分かるように、これも作用素の積の微分法に還元される。
証明の詳細は読者の演習問題とする。 □

3.3 合成関数の微分法

■合成関数の微分可能性 ここでは $U \subset E, V \subset E', f : V \to E'', g : U \to E'$ ただし $g(U) \subset V$ とする。このとき合成関数 $f \circ g : U \to E''$ が

$$(f \circ g)(x) = f(g(x)) \quad (x \in U)$$

で定義される。

定理 3.25. f, g が微分可能ならば $f \circ g$ は微分可能で

$$(D(f \circ g))_x = (Df)_{g(x)} \cdot (Dg)_x \quad (x \in U), \tag{59}$$

あるいは

$$D(f \circ g) = (\,(Df) \circ g\,) \cdot (Dg) \tag{60}$$

が成り立つ。
f, g が r 回 [連続] 微分可能ならば $f \circ g$ は r 回 [連続] 微分可能である。

証明 $x \in U$ とし, $y = g(x)$ とおく。仮定から 任意の正数 ϵ に対して正数 δ_1, δ_2 で

$$\left| f(y + \eta) - f(y) - A_y \eta \right| \le \epsilon |\eta| \quad (|\eta| \le \delta_1, \eta \in E'), \tag{61}$$

$$|g(x + \xi) - g(x) - B_x \xi| \le \epsilon |\xi| \quad (|\xi| \le \delta_2, \xi \in E) \tag{62}$$

を満たすものがある。ここで $A_y = (Df)_y, B_x = (Dg)_x$ とおいた。また以下では $\epsilon \le 1$ と仮定する。
(62) から

$$|g(x + \xi) - g(x)| \le (\epsilon + |B_x|)|\xi| \quad (|\xi| \le \delta_2, \xi \in E) \tag{63}$$

を得る。よって δ_2 を十分小さくとればさらに

$$|g(x + \xi) - g(x)| \le \delta_1 \quad (|\xi| \le \delta_2)$$

を満足せしめることができる。
そこで (61) の η に $g(x+\xi)-g(x)$ を代入すると、(63) により

$$\left| f(g(x+\xi)) - f(g(x)) - A_y(g(x+\xi) - g(x)) \right| \le \epsilon(1 + |B_x|) \tag{64}$$

を得る。よって (62) と (64) により

$$\left| f(g(x+\xi)) - f(g(x)) - A_y B_x \xi \right| \le \epsilon \left(1 + \left| A_y \right| + |B_x| \right) |\xi| \quad (|\xi| \le \delta_2)$$

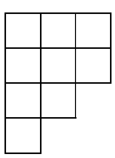
が得られ、 (59) が示される。
定理の後半は r に関する帰納法による。すなわち $r > 1$ に対しては $r-1$ まで定理が成立していると仮定する。f, g が r 回 [連続] 微分可能ならば Df, Dg は $r-1$ 回 [連続] 微分可能。よって帰納法の仮定より $(Df) \circ g$ は $r-1$ 回 [連続] 微分可能で、(60) より $D(f \circ g)$ は作用素の積だから $r-1$ 回 [連続] 微分可能である。すなわち $f \circ g$ は r 回 [連続] 微分可能で、定理が r に対しても成立していることが証明される。 □

(59) を方向微分を使って表すと

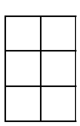
$$D_\xi(f \circ g) = [\, (Df) \circ g \,] \left(D_\xi g \right) \quad (\xi \in E) \tag{65}$$

となる。

■Young 図形　合成関数の高階微分の式を導出する前に、そのための準備として数学や物理学のいろいろな場面で登場する Young 図形について述べておこう。下の図

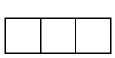


のように同じ大きさの箱を縦横に、左寄せに並べた図形を **Young 図形**という。ただし、横に並ぶ箱の数は下へ行くほど少くなっていく (同じでもいい) ものとする。横に並ぶ箱の数はずっと同じでもいいのだから



等も Young 図形である。

Young 図形における箱の横の並びを**行**、縦の並びを**列**と呼ぶことにする。下図のように一行だけから成る Young 図形も考えられる。



Young 図形の箱の総数 は Young 図形の**次数**と呼ばれる。本書では Young 図形の各行の箱の数をその行の**幅**と呼び、最左列の箱の数を Young 図形の**高さ**と呼ぶ。

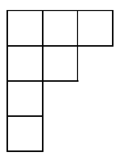
n を正整数とすると、次数が n の Young 図形 (これを略して n 次の Young 図形という) は n の分割を表していると考えることができる。ここで n の**分割**とは

$$n = n_1 + n_2 + \cdots + n_k \quad (n_1 \geq n_2 \geq \cdots \geq n_k) \tag{66}$$

のように n を段々に小さくなる (同じでもいい) 正整数の和に分けることである。例えば

$$7 = 3 + 2 + 1 + 1$$

は 7 の分割である。この分割は Young 図形



で完全に表現することができる。

n の分割 (= n 次 Young 図形) (66) を簡単に記号

$$(n_1, \ldots, n_k) \tag{67}$$

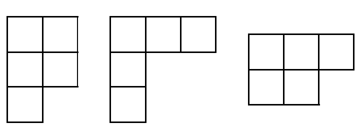
で表す。このとき k は Young 図形の高さである。

n 次 Young 図形の総数は n の分割の総数で、これは n の**分割数**と呼ばれる数である。

さて、Young 図形には (67) とは別の表記がある。すなわち、1 から n までの各整数 s に対して、幅 s の行の数を k_s とすると、n 個の非負整数の列

$$[\, k_1, k_2, \ldots, k_n \,] \tag{68}$$

が n 次 Young 図形を完全に記述している。例えば下図の次数 5 の Young 図形



はそれぞれ

$$[\, 1, 2, 0, 0, 0 \,] \quad [\, 2, 0, 1, 0, 0 \,] \quad [\, 0, 1, 1, 0, 0 \,]$$

で表される。

Young 図形 を (68) で表す場合、$\sum_s sk_s$ が Young 図形の次数で、$\sum_s k_s$ が高さである。

■集合の分割　一般に集合 S の**分割**とは S の部分集合から成る族 (集合を要素とする集合) $\mathcal{P} = \{ S_i \}$ で次の条件を満たすものをいう。

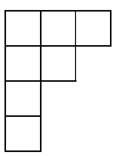
1. $S_i \neq \emptyset$,
2. $S_i \neq S_j$ ならば $S_i \cap S_j = \emptyset$,
3. $\bigcup_i S_i = S$

例えば

$$\mathcal{P} = \{ \{ 1, 2, 4 \}, \{ 3, 5 \}, \{ 6 \}, \{ 7 \} \}$$

は集合 $S = \{ 1, 2, 3, 4, 5, 6, 7 \}$ の分割である。

n 要素からなる集合の分割には n 次の Young 図形が対応している。例えば、上に挙げた分割 $\mathcal{P}$ には Young 図形



が対応している。すなわち、3 要素から成る部分集合 $\{ 1, 2, 4 \}$ を幅 3 の行で表す。2 要素から成る集合 $\{ 3, 5 \}$ を幅 2 の行で表す。最後に 1 要素から成る集合 $\{ 6 \}, \{ 7 \}$ をそれぞれ幅 1 の行で表す。

つまり、n 要素から成る集合 S の分割 $\mathcal{P} = \{ S_i \}$ に対して、s 要素から成る部分集合 S_i の数を k_s とすると $(s = 1, 2, \ldots, n)$、Young 図形 $[k_1, k_2, \ldots, k_n]$ が作られる。これを分割 $\mathcal{P}$ に対応する Young 図形と呼ぶのである。そして同時に、分割 $\mathcal{P}$ は **Young 図形 $[k_1, k_2, \ldots, k_n]$ に従う分割**と呼ぶことにする。

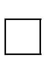
容易に分かるように Young 図形 $[k_1, k_2, \ldots, k_n]$ に従う分割は全部で

$$\frac{n!}{k_1! \cdot k_2! \cdot \cdots \cdot k_n! \cdot (1!)^{k_1} (2!)^{k_2} \ldots (n!)^{k_n}} = n! \prod_{s=1}^{n} \frac{1}{k_s! (s!)^{k_s}} \tag{69}$$

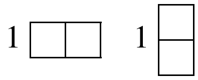
通りある。

例 3.1. 以下で各 Young 図形の左横に、それに従う分割が何通りあるかを記す。

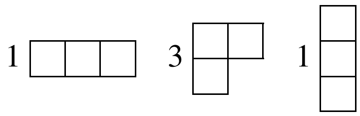
1 次の Young 図形は

1 

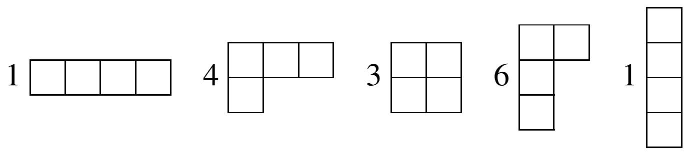
2 次の Young 図形は



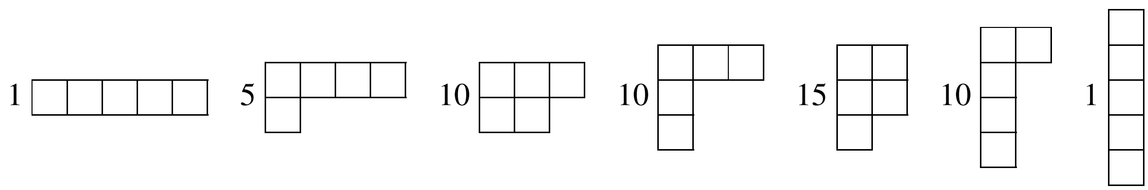
3 次の Young 図形は



4 次の Young 図形は



5 次の Young 図形は



練習問題 3.1. 6 次、7 次の Young 図形を完成させよ。各 Young 図形に従う分割の数も記すこと。

■合成関数の高階微分　合成関数の高階微分を方向微分で表す式を導出してみよう。$\xi_1, \xi_2 \dots \in E$ とする。はじめに (65) から

$$D_{\xi_1}(f \circ g) = [\,(Df) \circ g\,]\Big(\,D_{\xi_1} g\,\Big).$$

ここからまた

$$D_{\xi_1} D_{\xi_2}(f \circ g) = \Big[\,(D^2 f) \circ g\,\Big]\Big(\,D_{\xi_1} g, D_{\xi_2} g\,\Big) + [\,(Df) \circ g\,]\Big(\,D_{\xi_1} D_{\xi_2} g\,\Big).$$

式が繁雑になっていくので、以下では $f^{(s)} = D^s f$ と略記することにし、また U 上で定義された関数 h と集合 $\{\,1 \dots r\,\}$ の部分集合 $I = \{\,i_1 \dots i_s\,\}$ に対し $D_{\xi_{i_1}} \dots D_{\xi_{i_s}} h$ を、思い切って簡潔に $h_{i_1 \dots i_s}$ あるいは h_I と書くことにすると

$$(f \circ g)_1 = \Big[\,f^{(1)} \circ g\,\Big](g_1),$$

$$(f \circ g)_{12} = \Big[\,f^{(2)} \circ g\,\Big](g_1, g_2) + \Big[\,f^{(1)} \circ g\,\Big](g_{12}),$$

$$
\begin{aligned}
(f \circ g)_{123} = &\left[\, f^{(3)} \circ g \,\right](g_1, g_2, g_3) + \left[\, f^{(2)} \circ g \,\right](g_{13}, g_2) + \left[\, f^{(2)} \circ g \,\right](g_{23}, g_1) \\
&+ \left[\, f^{(2)} \circ g \,\right](g_{12}, g_3) + \left[\, f^{(1)} \circ g \,\right](g_{123}), \\
\cdots\cdots\cdots
\end{aligned}
$$

を得る。

ここで $r = 1, 2, 3$ に対し $(f \circ g)_{1\dots r}$ を表す式において、集合 $\{\,1 \dots r\,\}$ のすべての分割が各々一回ずつ現れていることに注意する。実際、このことは任意の正整数 r について成り立つのであって

命題 3.26. f, g は r 回微分可能だとすると

$$
(f \circ g)_{1\dots r} = \sum_{\mathcal{P}} \left[\, f^{(s)} \circ g \,\right](g_{I_1}, \dots, g_{I_s}) \tag{70}
$$

が成り立つ。ここで $\mathcal{P}$ は集合 $\{\,1 \dots r\,\}$ の分割全体に渡るものとし、$\mathcal{P} = \{\, I_1, \dots, I_s \,\}$ とする。

証明 r に関する帰納法による。

$r = 1$ ： これは自明である。

$r > 1$ ：帰納法の仮定から

$$
(f \circ g)_{1\dots r-1} = \sum_{\mathcal{Q}} \left[\, f^{(s)} \circ g \,\right](g_{K_1}, \dots, g_{K_s})
$$

が成り立つ。ここで $\mathcal{Q} = \{\, K_1 \dots K_s \,\}$ は集合 $\{\,1 \dots r-1\,\}$ の分割全体を渡る。

これを ξ_r で方向微分して

$$
(f \circ g)_{1\dots r} = \sum_{\mathcal{Q}} \left(\left[\, f^{(s+1)} \circ g \,\right](g_{K_1} \dots g_{K_s}, g_{\{\,r\,\}}) + \sum_{i=1}^{s} \left[\, f^{(s)} \circ g \,\right](g_{K_1} \dots g_{K_i \cup \{\,r\,\}} \dots g_{K_s}) \right). \tag{71}
$$

(71) の右辺に現れる $s+1$ 個の族 $\{\, K_1 \dots K_s, \{\,r\,\}\,\}, \{\, K_1 \dots K_i \cup \{\,r\,\} \dots K_s \,\}\,(i = 1 \dots s)$ は明らかに集合 $\{\,1 \dots r\,\}$ の分割である。これらを $\mathcal{Q}$ に対応する分割と呼ぶことにすると、次の主張は容易に確かめることができる：

1. 同じ $\mathcal{Q}$ に対応する $s+1$ 個の分割は互いに異なる分割である。
2. 異なる $\mathcal{Q}$ に対応する分割は異なる分割である。
3. 集合 $\{\,1 \dots r\,\}$ の任意の分割 $\mathcal{P}$ はある $\mathcal{Q}$ に対応する分割である。

したがって (71) 式は (70) にまとめることができる。 □

分割を Young 図形で分類すると (70) は

$$(f \circ g)_{1\dots r} = \sum_{\lambda \in Y_r} \sum_{\mathcal{P} \in \lambda} \left[\, f^{(s)} \circ g \,\right] (g_{I_1}, \dots, g_{I_s}). \tag{72}$$

ここで λ は次数 r の Young 図形全体 Y_r を、$\mathcal{P}$ は λ に従う分割全体を渡るものとする。

■Bell の多項式　式 (72) で $\xi_1 = \cdots = \xi_r = \xi$ の場合を考えてみよう。この場合、$\{\, 1 \dots r \,\}$ の部分集合 $I = \{\, i_1 \dots i_s \,\}$ に対して

$$\begin{aligned} g_I &= D_{\xi_{i_1}} \dots D_{\xi_{i_s}} g \\ &= D_\xi{}^s g \end{aligned}$$

が成り立つから、(72) において

$$(g_{I_1} \dots g_{I_s}) = D_\xi{}^{r_1} \dots D_\xi{}^{r_s} g$$

である。ここで r_i は I_i の要素数を表す。$r_1 \dots r_s$ を大きさの順に並べると明らかに $\lambda = (r_1 \dots r_s)$。つまり $(g_{I_1} \dots g_{I_s})$ は λ だけで決まり、$\mathcal{P}$ に依らない。

したがって (72) から

$$D_\xi{}^r (f \circ g) = \sum_{\lambda \in Y_r} \pi(\lambda) \left[\, (D^s f) \circ g \,\right] \left(\, D_\xi{}^{r_1} g \dots D_\xi{}^{r_s} g \,\right) \tag{73}$$

が得られる。ここで $\pi(\lambda)$ は λ に従う分割の総数で (69) で与えてある。$r_1 \dots r_s$ は λ そのものである：$\lambda = (r_1 \dots r_s)$.

$E = \mathbb{R}, \xi = 1$ の場合、(73) は

$$(f \circ g)^{(r)} = \sum_{\lambda \in Y_r} \pi(\lambda) \left[\, f^{(s)} \circ g \,\right] \left(\, g^{(r_1)} \dots g^{(r_s)} \,\right) \tag{74}$$

の形に書くことができる (再び $f^{(s)} = D^s f$ の略号を用いた)。

とくに $E = \mathbb{R}, E' = \mathbb{R}^n, \xi = 1$ の場合、(74) は

$$(f \circ g)^{(r)} = \sum_{\lambda \in Y_r} \pi(\lambda) \sum_{j_1 \dots j_s \in \{\, 1 \dots n \,\}} \left(\, f_{j_1 \dots j_s} \circ g \,\right) \left(\, g^{j_1} \,\right)^{(r_1)} \dots \left(\, g^{j_s} \,\right)^{(r_s)} \tag{75}$$

の形に書くこともできる。ここで $f_{j_1 \dots j_s}$ は $D_{j_1} \dots D_{j_s} f$ を意味する。$(g^1 \dots g^n)$ は g の成分関数である。

さらに $n = 1$ の場合、(75) は

$$(f \circ g)^{(r)} = \sum_{\lambda \in Y_r} \pi(\lambda) \left(\, f^{(s)} \circ g \,\right) g^{(r_1)} \dots g^{(r_s)} \tag{76}$$

の形をとる。ここで一風変わった書き方だが、

$$f_s = f^{(s)} \circ g, \quad g_k = g^{(k)}$$

という記法を用いると、(76) から

$$(f \circ g)_r = \sum_{\lambda \in Y_r} \pi(\lambda) f_s \,(\, g_1 \,)^{k_1} \ldots (\, g_r \,)^{k_r}\,, \tag{77}$$

あるいは

$$= \sum_{s=1}^{r} f_s \sum_{\lambda \in Y_{r,s}} \Big(\, \pi(\lambda) \;(\, g_1 \,)^{k_1} \ldots (\, g_r \,)^{k_r} \,\Big) \tag{78}$$

が得られる。ただし $Y_{r,s}$ は次数 r、高さ s の Young 図形全体の集合であり、$\lambda = [\, k_1 \ldots k_r \,]$ である。

例 3.2 (例 3.1 を参照). 以下の公式において f_r は $f^{(r)} \circ g$ を意味し、g_r は $g^{(r)}$ を意味することに注意する。右肩の添字は累乗を意味する。

$$(f \circ g)_1 = f_1 g_1,$$

$$(f \circ g)_2 = f_1 g_2 + f_2 {g_1}^2,$$

$$(f \circ g)_3 = f_1 g_3 + f_2 \,(\, 3 g_1 g_2 \,) + f_3 {g_1}^3,$$

$$(f \circ g)_4 = f_1 g_4 + f_2 \Big(\, 4 g_1 g_3 + 3 {g_2}^2 \,\Big) + f_3 \Big(\, 6 {g_1}^2 g_2 \,\Big) + f_4 {g_1}^4,$$

$$\begin{aligned}(f \circ g)_5 = f_1 g_5 + f_2 \,(\, 5 g_1 g_4 + 10 g_2 g_3 \,) + f_3 \Big(\, 10 {g_1}^2 g_3 + 15 g_1 {g_2}^2 \,\Big) \\ + f_4 \Big(\, 10 {g_1}^3 g_2 \,\Big) + f_5 {g_1}^5,\end{aligned}$$

$$\cdots\cdots\cdots$$

ここで Bell の多項式というものを紹介しておこう。r を正整数とする。$X = (X_1 \ldots X_r)$ を不定元とする多項式

$$\begin{aligned} Y_r(X) &= \sum_{\lambda \in Y_r} \pi(\lambda) X^{\lambda} \\ &= \sum_{\lambda \in Y_r} \pi(\lambda) {X_1}^{k_1} \ldots {X_r}^{k_r} \quad (\lambda = [\, k_1 \ldots k_r \,]) \end{aligned} \tag{79}$$

を r 次の **Bell の多項式**という。

$(f \circ g)_r$ は $Y_r(X_1 \dots X_r)$ の各 X_i に tg_i を代入して t を括り出したあと、t^s を f_s で置き換えたものに等しい。実際、

$$\begin{aligned}\sum_{\lambda\in Y_r}\pi(\lambda)\,(\,tg_1\,)^{k_1}\dots(\,tg_r\,)^{k_r} &= \sum_{\lambda\in Y_r}\pi(\lambda)t^s\,(\,g_1\,)^{k_1}\dots(\,g_r\,)^{k_r}\\ &\longrightarrow \sum_{\lambda\in Y_r}\pi(\lambda)f_s\,(\,g_1\,)^{k_1}\dots(\,g_r\,)^{k_r}\end{aligned}$$

となって、(77) に一致する。

例 3.3 (Bell の多項式).

$$\begin{aligned}
&Y_1(X) = X,\\
&Y_2(X) = X_2 + {X_1}^2,\\
&Y_3(X) = X_3 + 3X_1X_2 + {X_1}^3,\\
&Y_4(X) = X_4 + 4X_1X_3 + 3{X_2}^2 + 6{X_1}^2X_2 + {X_1}^4,\\
&Y_5(X) = X_5 + 5X_1X_4 + 10X_2X_3 + 10{X_1}^2X_3 + 15X_1{X_2}^2 + 10{X_1}^3X_2 + {X_1}^5,\\
&\dots\dots\dots
\end{aligned}$$

話が横道に逸れるが、r 次の Bell の多項式 $Y_r(X)$ に対し代入 $X_1 = \cdots = X_r = X$ を行ったものを $Z_r(X)$ と書くと、

$$\begin{aligned}Z_r(X) &= \sum_{\lambda\in Y_r}\pi(\lambda)X^{k_1+\cdots+k_r}\\ &= \sum_{s=1}^{r}\left(\sum_{\lambda\in Y_{r,s}}\pi(\lambda)\right)X^s.\end{aligned}$$

$Z_r(X)$ の s 次の係数は、r 要素の集合を s 個の部分集合に分割する仕方の総数で、これは**第 II 種の Stirling 数**と呼ばれる数である。記号では

$$\left\{ \begin{matrix} r \\ s \end{matrix} \right\}$$

と書く。すると

$$Z_r(X) = \sum_{s=1}^{r}\left\{ \begin{matrix} r \\ s \end{matrix} \right\}X^s.$$

例 **3.4 (第 II 種の Stirling 数 を係数にもつ多項式).**

$$Z_1(X) = X,$$

$$Z_2(X) = X + X^2,$$

$$Z_3(X) = X + 3X^2 + X^3,$$

$$Z_4(X) = X + 7X^2 + 6X^3 + X^4,$$

$$Z_5(X) = X + 15X^2 + 25X^3 + 10X^4 + X^5,$$

$$\cdots\cdots\cdots$$

さて話を本筋に戻して (75) で $n = 2$ の場合を考えてみる。添字が多くなって煩わしいため、$\varphi = g^1, \psi = g^2$ とおいて、$n = 1$ のときと同様、

$$\varphi_l = \varphi^{(l)}, \psi_m = \psi^{(m)}, \ldots$$

の記法を使うと、

$$(f \circ g)_r = \sum_{\lambda \in Y_r} \pi(\lambda) \sum_{K \subset \{\, 1 \ldots s \,\}} f_{t,s-t} \, (\, \varphi_1 \,)^{l_1} \ldots (\, \varphi_r \,)^{l_r} \, (\, \psi_1 \,)^{k_1 - l_1} \ldots (\, \psi_r \,)^{k_r - l_r}, \tag{80}$$

ここで K は $\{\, 1 \ldots s \,\}$ の部分集合全体を動き、t は K の要素数,

$$f_{t,s-t} = f_{\underbrace{1 \ldots 1}_{t} \underbrace{2 \ldots 2}_{s-t}} \circ g$$

である。$l_1 \ldots l_r$ は次のように定義される：$K = \{\, i_1 \ldots i_t \,\}$ (ただし $i_1 < \cdots < i_t$) とし、Young 図形 μ を $\mu = (r_{i_1} \ldots r_{i_t})$ で定義する。つまり μ は λ の上から i_1 番目、i_2 番目、... i_t 番目の行を選んで作られる "部分"Young 図形である。これを $\mu = [\, l_1 \ldots l_r \,]$ と書き直して、$l_1 \ldots l_r$ が定義される。

(80) で $f_{t,s-t}, l_1 \ldots l_r$ が部分 Young 図形 μ だけから決まることに注意すると、容易にわかるように (80) から

$$\begin{aligned}(f \circ g)_r = \sum_{\lambda \in Y_r} \pi(\lambda) \sum_{\mu \leq \lambda} f_{t,s-t} \binom{k_1}{l_1} \cdots \binom{k_r}{l_r} \cdot \\ \cdot (\, \varphi_1 \,)^{l_1} \ldots (\, \varphi_r \,)^{l_r} \, (\, \psi_1 \,)^{k_1 - l_1} \ldots (\, \psi_r \,)^{k_r - l_r}\end{aligned} \tag{81}$$

が得られる。

(81) はまた次のようにして得ることもできる：r 次の Bell 多項式 $Y_r(X)$ の各 X_i に $u\varphi_i + v\psi_i$ を代入して展開し、u, v を前に括り出したあと $u^t v^{s-t}$ を $f_{t,s-t}$ で置き換えれば (81) が得られるのである。実際、

$$\begin{aligned}
&\sum_{\lambda \in Y_r} \pi(\lambda) \left(u\varphi_1 + v\psi_1 \right)^{k_1} \ldots \left(u\varphi_r + v\psi_r \right)^{k_r} \\
&= \sum_{\lambda \in Y_r} \pi(\lambda) \sum_{l_1 \ldots l_r} \binom{k_1}{l_1} \ldots \binom{k_r}{l_r} u^{l_1+\cdots+l_r} v^{k_1+\cdots+k_r-(l_1+\cdots+l_r)} \\
&\qquad \left(\varphi_1 \right)^{l_1} \ldots \left(\varphi_r \right)^{l_r} \left(\psi_1 \right)^{k_1-l_1} \ldots \left(\psi_r \right)^{k_r-l_r} .
\end{aligned}$$

ここで各 l_i は $0 \le l_i \le k_i$ の範囲を自由にかつ独立にうごくとものとする。すなわち $l_1 \ldots l_r$ は $\lambda = [\, k_1 \ldots k_r \,]$ の部分 Young 図形 $\mu = [\, l_1 \ldots l_r \,]$ を表す。ゆえに

$$\begin{aligned}
&= \sum_{\lambda \in Y_r} \pi(\lambda) \sum_{\mu \le \lambda} u^t v^{s-t} \binom{k_1}{l_1} \ldots \binom{k_r}{l_r} \left(\varphi_1 \right)^{l_1} \ldots \left(\varphi_r \right)^{l_r} \left(\psi_1 \right)^{k_1-l_1} \ldots \left(\psi_r \right)^{k_r-l_r} \\
&\longrightarrow \sum_{\lambda \in Y_r} \pi(\lambda) \sum_{\mu \le \lambda} f_{t,s-t} \binom{k_1}{l_1} \ldots \binom{k_r}{l_r} \left(\varphi_1 \right)^{l_1} \ldots \left(\varphi_r \right)^{l_r} \left(\psi_1 \right)^{k_1-l_1} \ldots \left(\psi_r \right)^{k_r-l_r}
\end{aligned}$$

となって (81) が得られる。

3.4　Taylor の公式

未完

参考文献

[1] 高木貞治「解析概論 改訂第 3 版」岩波書店

[2] 森口・宇田川・一松「岩波数学公式 I」岩波書店